IBM

Workload Optimized Systems
Tuning POWER7 for Analytics

- Strengthens IBM Cognos and SPSS for analytics workloads
- Helps optimization, security, scaling, capacity, and resilience
- Shows model for analytics on IBM Power Systems

Dino Quintero
Thomas Ancel
Graeme Cassie
Rodrigo Ceron
Amr Darwish
Guilherme G. Felix
Jian Jun He
Bharathraj Keshavamurthy
Sreenivas Makineedi
Girish Nikalje
Sudipto Pal
Zubair Salie
Ashu Tiwary

Redbooks

ibm.com/redbooks

International Technical Support Organization

Workload Optimized Systems: Tuning POWER7 for Analytics

October 2012

SG24-8057-00

Note: Before using this information and the product it supports, read the information in “Notices” on page v.

First Edition (October 2012)

This edition applies to IBM SPSS Modeler Premium 15.0, SPSS Text Analytics Client 15.0 for Windows, SPSS Text Analytics Administration Console, SPSS Modeler Premium Server 15.0 for AIX, SPSS Text Analytics Server 15.0 for AIX, and SPSS Data Access Release 6 Service Pack 3 for AIX.

Contents

Notices

This information was developed for products and services offered in the U.S.A.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not grant you any license to these patents. You can send license inquiries, in writing, to:
IBM Director of Licensing, IBM Corporation, North Castle Drive, Armonk, NY 10504-1785 U.S.A.

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law: INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM websites are provided for convenience only and do not in any manner serve as an endorsement of those websites. The materials at those websites are not part of the materials for this IBM product and use of those websites is at your own risk.

IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Any performance data contained herein was determined in a controlled environment. Therefore, the results obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurements may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs.

Trademarks

Preface

This IBM® Redbooks® publication addresses topics to help clients to take advantage of the virtualization strengths of the POWER® platform to solve system resource utilization challenges and maximize system throughput and capacity. This publication examines the tools, utilities, documentation, and other resources available to help technical teams provide business solutions and support for Cognos® Business Intelligence (BI) and Statistical Package for the Social Sciences (SPSS®) on Power Systems™ virtualized environments.

This book addresses topics to help address complex high availability requirements, help maximize the availability of systems, and provide expert-level documentation to the worldwide support teams.

This book strengthens the position of the Cognos and SPSS solutions with a well-defined and documented deployment model within a POWER system virtualized environment. This model provides clients with a planned foundation for security, scaling, capacity, resilience, and optimization.

This book is targeted toward technical professionals (BI consultants, technical support staff, IT Architects, and IT Specialists) who are responsible for providing Smart Analytics solutions and support for Cognos and SPSS on Power Systems.

The team who wrote this book

This book was produced by a team of specialists from around the world working at the International Technical Support Organization, Poughkeepsie Center.

Dino Quintero is an IBM Senior Certified IT Specialist with the ITSO in Poughkeepsie, NY. His areas of knowledge include enterprise continuous availability, enterprise systems management, system virtualization, and technical computing, and clustering solutions. He is currently an Open Group Distinguished IT Specialist. Dino holds a Master of Computing Information Systems degree and a Bachelor of Science degree in Computer Science from Marist College.

Thomas Ancel is an AIX® Senior Product Services Specialist working for the IBM EMEA AIX Development Support Team, and is located in Noisy Le Grand, France. He has 11 years of experience on AIX support and in the AIX performance-tuning field. He holds a degree in Computer and IT Engineering from IG2I, a department of Ecole Centrale de Lille. His areas of expertise include AIX tuning, AIX Kernel debugging, AIX File Systems and Logical Volume Manager (LVM) debugging. He participated in writing patent disclosures and publications for IBM AIX and wrote articles in the AIX Support quarterly publication intended for French clients.

Graeme Cassie is Technical Director for Barrachd Ltd, which is an IBM Business Partner in the UK that specializes in BI and Analytics solutions. He has almost 30 years experience in the IT industry, working 15 years for IBM. The last 14 years, he worked for IBM Business Partners, always in a technical role. He specializes in Power Systems™, Storage, System x, and is a Certified IBM Technical Architect. He holds many IBM certifications on infrastructure and was part of the team that wrote high-end storage certifications.

Rodrigo Ceron is a Master Inventor and Consultant in IBM Lab Services and Training Latin America, in Brazil. He has 12 years of experience in the UNIX/Linux area and eight years working at IBM, where he created seven patents in multiple areas. He graduated with honors in Computer Engineering from the University of Campinas (UNICAMP) and holds an IEEE CSDA credential. His areas of expertise include Power Systems high availability, performance, and analytics. He also published papers about Operations Research in IBM developerWorks®.

Amr Darwish is an Advisory Software Engineer, Application Integration & Middleware Solutions Specialist, based at the IBM Cairo, Egypt facility. He has over 13 years of experience in the IT field and participated in several extensive development and maintenance cycles as part of several IBM internal and external engagements. He holds a Bachelor of Science degree in Computer and Systems Engineering from Faculty of Engineering - Computer and Systems Department, Ain Shams University. He has expertise in the following areas and held these positions: Java Developer, Database Developer, Application Designer, Application Architect, Technical Leader, and Project Manager.

Guilherme G. Felix is a Power System, Virtualization, and AIX Specialist working for IBM Brazil. He has more than 12 years of experience with UNIX and virtualization. He holds a Bachelor degree in Information Systems. He is a Certified Advanced Technical Expert (CATE), Red Hat Certified Engineer (RHCE), and a Master Certified IT Specialist. His articles were published in IBM developerWorks, and he has co-authored other IBM Redbooks publications.

Jian Jun He is an Advisory IT Specialist in IBM CSTL Lab Service and Training in China. He has nine years of experience in the UNIX/Linux area and IBM Power Systems. He holds a Master degree in Software Technology from Beijing University of Posts and Telecommunications. He also has IBM Certifications in IBM AS/400 and IBM WebSphere® Application Server. He published several papers in IBM developerWorks.

Bharathraj Keshavamurthy is a Performance Architect for Enterprise Solutions from Bangalore, India. He works with the software group and has over five years of experience in the performance engineering of IBM cross-brand products, mainly in WebSphere Application Server integration areas. He holds a Bachelor of Engineering degree from University of RVCE, Bangalore, India. His areas of expertise include performance benchmarking IBM products, end-to-end performance engineering of enterprise solutions, performance architecting, designing solutions, and sizing capacity for solutions with IBM product components. He wrote many articles that pertain to performance engineering in developerWorks and in international science journals.

Sreenivas Makineedi is an AIX Advisory Software Engineer and Senior AIX Development Support Specialist working for the IBM AIX Development Support Team based in Austin, TX, US. He has more than 13 years of experience in AIX Development and Support. Sreenivas holds a Master of Technology degree in Electronics and Communication Engineering from Indian Institute of Technology, Kharagpur, India. He specializes in Development and Support of various AIX Operating System components. His areas of expertise include AIX Virtual Memory Management, File Systems Development, and Kernel Debugging. He has several patents and publications in AIX Operating System Technologies.

Girish Nikalje is an IT Specialist based at the IBM CIO Innovation Systems PUNE, India. He has seven years of experience in UNIX and WebSphere administration. He holds an IBM Certification for IBM AIX, Microsoft Certified Systems Engineer (MCSE), CISCO Certified Network Associate (CCNA), and ITIL. His articles about WebSphere are published on the Knowledge Community forum. For the last three years, Girish worked extensively on WebSphere Administration, which includes application deployment, problem solving, performance tuning, and the implementation of fail-safe runtime environments.

Sudipto Pal is a Senior System Analyst in the Business Analytics and Optimization (BAO) team working for IBM Global Services in India. He leads support, enhancement, migration, and performance tuning of the Cognos BI environment for several IBM clients, such as Disney Corporation and NiSource America. Sudipto also has in-depth knowledge in AIX PowerVM® virtualization. He was a Technical Consultant in the STG independent software vendor (ISV) enablement team, working for IBM Global Services in India. He works with the software group to benchmark several products. He published several white papers with collateral for WebSphere Application Server and WebSphere MQ Server with virtualization features, such as logical partition (LPAR), workload partition (WPAR), and Active Memory Sharing (AMS). He has experience coding in C/C++ under the UNIX environment for several networking protocols, such as LDAP, Domain Name System (DNS), Dynamic Host Configuration Protocol (DHCP), and Simple Network Management Protocol (SNMP). Sudipto has 12 years in the industry, and holds a Master degree in Computer Applications.

Zubair Salie is an Integration Manager who works for a major insurance company in South Africa. He has 14 years experience in the IT industry, 10 years in BI. His areas of expertise range from architecting solutions on IBM Cognos, TM1®, and RTM to BI application development and dimensional modelling for data warehouses.

Ashu Tiwary is an IBM employee with the AIX Development Support organization, working mostly to resolve Critical Situations. He provides leadership for the global AIX Development Support mission. He worked with AIX for more than 15 years as a Systems Administrator, Technical Support Engineer, and Developer.

Thanks to the following people for their contributions to this project:

David Bennin, Richard Conway
International Technical Support Organization, Poughkeepsie Center

Robbie Crawford, Brian Henderson
Barrachd UK

Vincent Salewyn
IBM France

Charles Herman, Barb Parker, John Mullin, Beth Hoffman, Wayne Huang, Mark Pilon, Mark McFadden, Alan Nelson, Jim Gosack
IBM USA

Now you can become a published author, too!

Here's an opportunity to spotlight your skills, grow your career, and become a published author - all at the same time! Join an ITSO residency project and help write a book in your area of expertise, while honing your experience using leading-edge technologies. Your efforts will help to increase product acceptance and customer satisfaction, as you expand your network of technical contacts and relationships. Residencies run from two to six weeks in length, and you can participate either in person or as a remote resident working from your home base.

Find out more about the residency program, browse the residency index, and apply online at:

ibm.com/redbooks/residencies.html

Comments welcome

Your comments are important to us!

We want our books to be as helpful as possible. Send us your comments about this book or other IBM Redbooks publications in one of the following ways:

- Use the online **Contact us** review Redbooks form found at:

  ibm.com/redbooks

- Send your comments in an email to:

  redbooks@us.ibm.com

- Mail your comments to:

  IBM Corporation, International Technical Support Organization
  Dept. HYTD Mail Station P099
  2455 South Road
  Poughkeepsie, NY 12601-5400

Stay connected to IBM Redbooks

- Find us on Facebook:

  http://www.facebook.com/IBMRedbooks

- Follow us on Twitter:

  http://twitter.com/ibmredbooks

- Look for us on LinkedIn:

  http://www.linkedin.com/groups?home=&gid=2130806

- Explore new Redbooks publications, residencies, and workshops with the IBM Redbooks weekly newsletter:

  https://www.redbooks.ibm.com/Redbooks.nsf/subscribe?OpenForm

- Stay current on recent Redbooks publications with RSS Feeds:

  http://www.redbooks.ibm.com/rss.html

1

Introducing analytics on IBM Power Systems

Analytics concepts and the usage of two IBM solutions on POWER hardware to achieve enhanced business results are introduced:

- IBM Cognos 10 Business Intelligence (BI) and Cognos TM1
- IBM Statistical Package for the Social Sciences (SPSS) Collaboration and Deployment Services

The following topics are described:

- Analytics and the advantages to businesses
- IBM Cognos: Cognos 10 BI and Cognos TM1
- IBM SPSS suite and the Collaboration & Deployment Service
- POWER7 hardware, AIX 7, and PowerVM features

1.1 Introduction to Business Analytics

The job of business leaders is to ensure that their business is as successful as possible. Business leaders decide on a daily basis where to invest to obtain the best returns, how to make a marketing campaign reach customers with more buying potential, how to optimize operations, how to improve logistics, and other decisions. These decisions help their businesses increase profitability, reduce costs, and more efficiently manage and serve their customer base, which is critical to maintain competitiveness and efficiency in this global economy.

Business Analytics helps organizations to understand, anticipate, and shape business outcomes. Business Analytics efficiently organizes data to let organizations understand how they are doing at a particular moment, why they are doing so, and what they must do to achieve better outcomes. These data-driven insights help business leaders make more accurate and confident decisions.

Business Analytics helps you turn *relevant information* into *business insights* to make *smarter decisions* and achieve *better outcomes*.

For example, imagine you are a large retail chain. You buy products from providers in multiple locations and sell those products at various stores across the country. What if you have a highly accurate view of product cost by provider, transportation costs by carrier, and consumption trends by area? You can use this information to decide what combination of provider, shipment, and storage costs best suits each of your retail stores. Even better, what if this information you consume is provided in real time? You can then quickly adapt to price changes in the chain and learn which one is now the optimal combination to feed your stores. This decision process can even be automated to process thousands or even millions of daily transactions. If your competition is unable to also understand these logistics and act, you have a competitive advantage over them. This scenario is a simple example of what you can do with analytics.

In this world where information is overwhelming, comes from multiple sources, and changes rapidly, it is becoming imperative for businesses to use available information to apply intelligence to the business. According to research[1], organizations that achieve competitive advantage with analytics are 2.2 times more likely to substantially outperform their industry peers.

When you are able to align your organization around information, it helps you see and predict business outcomes by the use of analytics, which provides you the confidence you seek to act at the points of impact to optimize outcomes. By doing so, you transform your business, and learn from the solutions that get smarter with every outcome, thus feeding a *virtuous cycle*. Figure 1-1 on page 3 shows the three-step, virtuous cycle paradigm of Business Analytics.

[1] Source: The New Intelligent Enterprise, a joint MIT SLOAN Management Review and IBM Institute of Business Value analytics research partnership. Copyright Massachusetts Institute of Technology 2011.

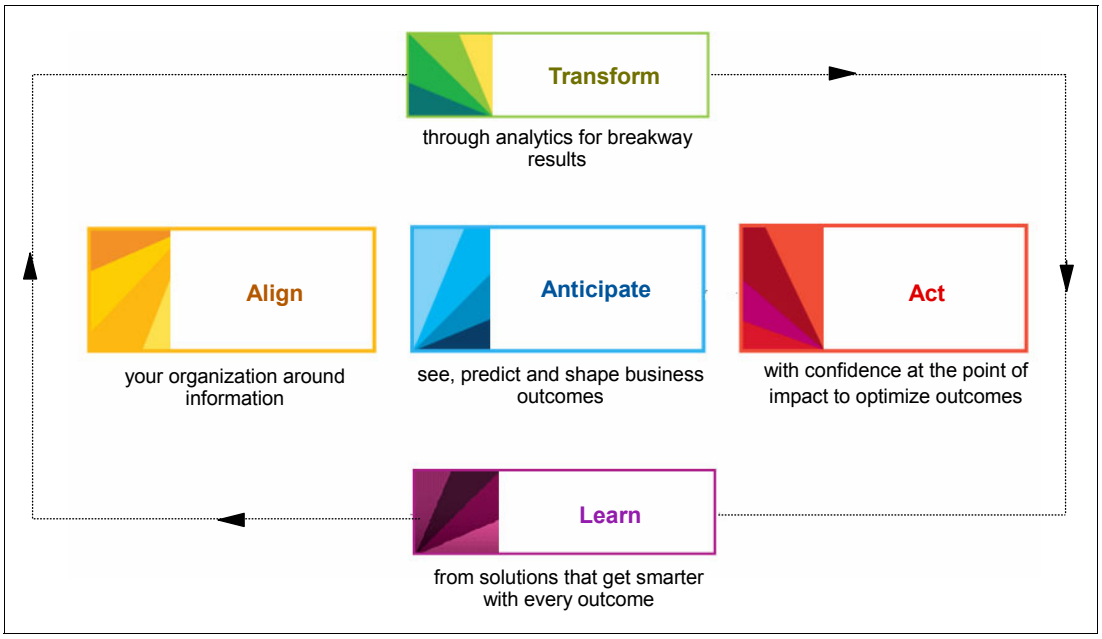

Figure 1-1 The paradigm of business analytics

A critical factor in the virtuous cycles in Figure 1-1 is being able to *anticipate*, which is the area of predictive analytics. With it, you can, for example, predict the next purchase of a customer based on past data that is available from other customer purchase patterns. Predictive analytics serves the purpose of scoring how likely an action is to happen based on known past patterns and current input data; it gives you more confidence to take the best action.

Probability: Predictive analytics helps you to score the *probability* of a future event to happen. It is based on past historical data patterns that are applied to current input data to help you to make the best decision.

Business Analytics is in the ramp-up of exploitation by businesses worldwide. Until a few years ago, most business people believed that it was not a critical factor to business success. As more leaders experiment with it, this scenario is changing and the importance of analytics is a primary area in Information Technology (IT).

There are different levels of maturity when it comes to the adoption of analytics. The more experienced you are, the more you know what data is important to your business and the quicker you can adapt to changes. Also, analytics-experienced players tend to use it in every aspect of the business in a widespread fashion, benefiting from cross-organizational insights. More mature players perform well in three areas:

- Information management
- Analytics skills and tools
- Data-oriented culture

The first area, information management, refers to how well you can gather and use the information. It also refers to how well spread this information gets within the organization so that everyone can use it toward the organizational goals. Having information available to all of the business areas allows for better data integration and provides better views about how processes are interrelated. You can more easily determine how a change impacts the business as a whole.

The second area, analytics skills and tools, refers to the set of technology that you use to help you with insights on more complex questions. These tools predict what might happen in the current scenario soon, assist you in “what-if” scenarios, and are able to automate operational tasks based on business models.

The third area, data-oriented culture, means that the company culture is to expect that a data-driven analysis exists to back up ideas and actions. That is, base the day-to-day decisions and analysis of the organization less around intuition and more around facts from data.

To help your organization step into or mature in this new way of handling business, IBM has five categories within its integrated portfolio of Business Analytics solutions that can assist you:

- Business Intelligence: IBM Cognos Business Intelligence (BI):
  - Enables users to discover, understand, and share the right information to drive better decisions across the organization.
  - Allows organizations to respond to market opportunity and challenges with agile and collaborative decision making processes.
  - Empowers users to make the best decisions and act on them from anywhere.
  - Permits a quick and effective response to changing business demands.
- Predictive and advanced analytics: IBM SPSS:
  - Creates forecasts.
  - Simulates what-if scenarios.
  - Predicts future outcome and behavior.
  - Allows the analysis of structured data (tables and databases) and unstructured (text).
- Financial performance and strategy management: IBM Cognos BI, IBM Cognos Controller and IBM Cognos TM1, and IBM Cognos Financial Statement Reporting (FSR):
  - Supports performance reporting and *scorecarding*.
  - Creates planning, analysis, and forecasting.
  - Performs profitability modeling and optimization.
  - Consolidates financial and statutory reporting.
  - Embeds control and compliance policies, practices, and metrics into operational processes and automates workflow and other key processes in the production of financial and regulatory reports.
- Governance, risk, and compliance: OpenPages® software products:
  - Enables organizations to acquire deep insight into cross-domain business operations and to embed risk management practices into them to manage unexpected outcomes and reduce the impact of risks that actually occur.
- Analytics applications:
  - Ready templates for reporting and analyzing results of specific business domains, such as customer, workforce, supply chain, or financial management. They integrate with enterprise environments with little effort. These applications are good entry points for analytics starters.

The use of analytics is not constrained to a few segments in industry. IBM Business Analytics solutions already help organizations from multiple industries benefit from the use of analytics:

- Transportation
- Healthcare
- Energy
- Public sector
- Marketing
- Communications
- Banking
- Insurance

The following examples show different areas within an organization that use analytics:

- Finance: Create and manage reports, scorecards, plans, and forecasts with data from many sources. Align resource plans for intelligent growth and profit, and also comply with confidence.
- Human resources: Optimize staffing mix and benchmark benefits.
- Product development: Reduce portfolio gaps and reduce development risk.
- Operations: Improve production capacity (earnings) and reduce inventory buffer (savings).
- Sales: Maximize pipeline, effectiveness, and customer profitability. Understand who are the key customers for the business.
- Marketing: Improve competitive positioning, prioritize delivery of more profitable products, and drive greater demand by achieving a more responsive niche of consumers.
- Customer service: Reduce customer churn, increase satisfaction and loyalty, and understand customer sentiment and behavior. Predict customer behavior.

Analytics can be applied to any business. It is a critical investment area to achieve success in the current economy.

The following sections provide an overview of two of the IBM Business Analytics software products, IBM Cognos and IBM SPSS, that can bring you the business advantage that you are seeking. We also discuss a few business scenarios that can benefit from using them.

In Chapter 2, “Implementing Cognos on IBM Power Systems” on page 49 and Chapter 3, “Implementing IBM SPSS on IBM Power Systems” on page 79, we provide implementation scenarios. We outline the preferred practices and tuning to achieve the best performance for IBM Cognos and IBM SPSS on the Power hardware.

1.2 Introduction to IBM Cognos

Organizations are constantly under pressure to understand and react quickly to new information. In addition, the complexity and volumes of data for all aspects of the environments in which organizations operate are increasing. Markets, regulatory environments, customer and supplier data, competitive information, and internal operational information all affect how data is viewed and interpreted. It is becoming imperative for organizations to react correctly, dynamically, and in a timely fashion to answer key business questions and to outperform the competition.

From BI to financial performance and strategy management to analytics applications, Cognos software can provide what your organization needs to become top-performing and analytics-driven. With products for the individual, workgroup, department, mid-sized business, and large enterprise, Cognos software is designed to help everyone in your organization make decisions that achieve better business outcomes for now and in the future.

1.2.1 How can IBM Cognos help

Cognos solutions are designed to help you make smarter decisions and manage performance for optimized business results. Combine your financial and operational data into a single, seamless source of information in the environment of your choice for transparent and timely reporting, analysis, and action.

1.2.2 What can IBM Cognos offer

IBM Cognos introduces a list of featured products to meet the increasing demands of organizations.

IBM Cognos Enterprise

IBM Cognos Enterprise meets the needs of business, finance, and IT with powerful and scalable BI and performance management solutions that support all business decisions, across the entire organization. Cognos Enterprise helps users freely explore information, analyze key facts, quickly collaborate to gain alignment with key stakeholders, and plan and act with confidence to drive better business outcomes.

With Cognos Enterprise, everyone is equipped with what they need to help their organization outperform, including reports, analysis, statistics, dashboards, what-if analysis, plans, and budgets.

Cognos Enterprise is a portfolio of integrated BI and performance management solutions that supports the way users want to work and can help everyone in your organization meet strategic objectives. It is part of the Cognos 10 family of products that enables your organization to start anywhere to quickly satisfy urgent business requirements and grow your solution over time to address future needs.

IBM Cognos Express

IBM Cognos Express delivers the essential reporting, analysis, dashboard, scorecard, planning, budgeting, and forecasting capabilities that workgroups and mid-sized companies need at a price they can afford. Everything is included in a pre-configured solution that is easy to install, easy to use, and easy to buy.

With Cognos Express, you can equip everyone with the tools to help your organization outperform, including reports, analysis, statistics, dashboards, what-if analysis, plans, and budgets.

Cognos Express: IBM Cognos Express is not available on the AIX operating system. However, it is listed here for completeness.

IBM Cognos Insight

IBM Cognos Insight puts the power to solve individual and workgroup challenges directly in the hands of business users. More than merely a discovery tool, IBM Cognos Insight helps you turn your insights into action and solve problems or seize new opportunities quickly, all from the convenience of your desktop.

Cognos Insight: IBM Cognos Insight is available as part of IBM Cognos Express Planner and IBM Cognos Express Advisor, and therefore is not available on AIX. It is listed here for completeness.

Cognos TM1

IBM Cognos TM1 radically transforms your entire planning cycle, from forecasting and analysis processes to budgeting and profitability modeling.

Cognos TM1 facilitates preferred practices, such as driver-based planning, rolling forecasts, and profitability analysis that are integrated with planning. It elevates the ability of Finance to manage performance proactively. And with the full range of planning processes tied to the operational frontline of an organization, Cognos TM1 enables planners and analysts to uncover hidden business options, increase profits, and manage capital effectively.

Cognos FSR

IBM Cognos Financial Statement Reporting (FSR) transforms the way that organizations manage and deliver critically important management reports and external disclosures. Cognos FSR capabilities help CFOs reduce risk and improve efficiency in their most critical reporting processes.

Cognos FSR: IBM Cognos FSR is not available on the AIX operating system. It is listed here for completeness.

Cognos Controller

IBM Cognos Controller sets a new standard for financial consolidation software with a comprehensive suite of capabilities for financial consolidation and management and financial reporting. Operated and managed by the office of finance, Cognos Controller ensures the accurate collection and reconciliation of financial results. It delivers the timely, accurate information that finance organizations need for regulatory compliance, planning, analysis, and performance management.

Cognos platform

The common foundation for your Cognos software solution. The Cognos platform ensures that your BI and performance management systems have access to all relevant data sources throughout your organization.

The Cognos platform can access the complete mix of data sources present in most organizations today, whether they are data warehouses, technology platform applications, online analytical processing (OLAP) cubes, historical data sources, or modern data sources.

It provides a broad reach and open access to all data sources and an easy modeling environment for building access to all needed information.

Two products of the Cognos family are described:

- Cognos BI (as a member of Cognos Enterprise group)
- Cognos TM1

1.2.3 Cognos 10 BI

For the best business outcomes, your company needs agile BI capabilities that are designed to help users assess facts, reinvent strategy, and adjust plans. With the wide range of IBM BI capabilities right-sized to satisfy all user types and working styles, everyone in your organization can freely explore data, analyze key facts, collaborate to gain alignment with key stakeholders, and act with confidence to drive your business forward.

Cognos BI features

We describe the Cognos BI features:

- Reports

  IBM BI software helps ensure that users are equipped with the reports they need to make fact-based decisions in a system that is simpler, faster, and easier to manage. From professional report authors who design one-to-many reports for the enterprise, to business users who need to create their own ad hoc queries or customize existing reports, IBM reporting capabilities fit the needs of users throughout your organization.

- Analysis

  With the analysis capabilities of IBM BI software, users can explore information and different perspectives easily and intuitively to make sure that they are making the right decisions. General business users can easily view, assemble, and analyze the information that is required to make better decisions. And, business and financial analysts can take advantage of more advanced, predictive, or what-if analysis capabilities.

- Scorecards

  Scorecards enable your organization to capture corporate strategy and communicate that strategy at the operational level. Executives and managers can define quantifiable goals and targets and track performance for business units, operating subsidiaries and geographic regions to quickly identify the areas that need attention.

- Dashboards

  With dashboards, users can access, interact, and personalize content in a way that supports the unique way that they make decisions. Security-rich access to historic, current, and projected data means that users can quickly move from insight to action.

- Statistics

  Statistics capabilities help you incorporate statistical results with core business reporting, reducing the time that it takes to analyze data and prepare business presentations that are based on that analysis.

- Mobile BI

  Mobile BI capabilities make it possible for your mobile workforce to interact with information in new ways by delivering relevant BI wherever the workers are. Users interact with trusted BI through a rich and visual experience, whether offline or online. The flexible platform ensures mobile decision-making is simple, reliable, and safe.

- Real-time monitoring

  Real-time monitoring capabilities provide your employees on the frontline with a rich view of operational KPIs[2] and measures as they happen to support up-to-the-moment decision making.

[2] Key performance indicators (KPIs) are used by companies to better evaluate their current level of business success and to help plan for the future.

- Collaboration

  Collaboration capabilities help individuals, key stakeholders, workgroups, and teams align their strategic objectives, build stronger relationships, learn from history, and use resources for important decision-making most effectively.

- Planning and budgets

  Get the right information to the right people in the form they need it to plan, budget, and forecast. Planning and budgeting capabilities in the IBM BI software support a wide range of requirements, from high-performance, on-demand customer and profitability analysis and flexible modeling to enterprise contribution for a broad range of users.

Cognos BI components overview

IBM Cognos BI is an integrated BI suite that provides a wide range of functionality to help you understand the data of your organization. Everyone in your organization can use IBM Cognos BI to view or create business reports, analyze data, and monitor events and metrics so that they can make effective business decisions.

IBM Cognos BI integrates the following BI activities in one web-based solution as stated in Table 1-1.

Table 1-1 IBM Cognos BI list of components

Activity	Component
Publishing, managing, and viewing content	IBM Cognos Connection
Interactive workspaces	IBM Cognos Business Insight
Simple reporting and data exploration	IBM Cognos Business Insight Advanced
Ad hoc querying	IBM Cognos Query Studio
Managed reporting	IBM Cognos Report Studio
Event management and alerting	IBM Cognos Event Studio
Scorecarding and metrics	IBM Cognos Metric Studio
Analyzing your business	IBM Cognos Analysis Studio
Working with IBM Cognos BI content in Microsoft Office	IBM Cognos for Microsoft Office

Before you use IBM Cognos BI, you must understand how each of the components that make up the IBM Cognos BI user interfaces can help you do your job.

More about Cognos BI components

The following section is a brief description about each of Cognos BI components.

Cognos Connection

IBM Cognos Connection is the web portal for IBM Cognos BI. It is the starting point to access your BI information and the functionality of IBM Cognos BI.

Use the portal to publish, find, manage, organize, and view the BI content of your organization, such as reports, scorecards, and agents. To view the reports, you use IBM Cognos Viewer.

If you have the necessary permissions, you can access the various studios from the portal and use the portal for content administration, including scheduling and distributing reports, and creating jobs.

System administrators also use the portal to manage servers, optimize performance, and set access permissions.

Business Insight

With IBM Cognos Business Insight, you can build and use sophisticated, interactive workspaces by using IBM Cognos content, as well as external data sources.

A *workspace* is a visual display of the most important information that you need to quickly gain insight about your business objective. You can monitor a workspace at a glance. In Business Insight, objects that you add to a workspace are displayed in widgets. These widgets contain content and interact with each other.

Within a Business Insight workspace, you work with existing content and perform basic analysis, data exploration, and collaborative decision making. You can also perform enhanced searches of IBM Cognos BI content to quickly find relevant information that is presented in reports, analyses, workspaces, metric information, and events.

Business Insight Advanced

With IBM Cognos Business Insight Advanced, you can create simple reports and explore your data.

When you are in an IBM Cognos Business Insight workspace and want to perform deeper analysis and report authoring, you can seamlessly upgrade to Business Insight Advanced. In Business Insight Advanced, you can perform more advanced data exploration, such as adding more measures, conditional formatting, and advanced calculations. You can also launch Business Insight Advanced directly from the IBM Cognos Connection portal.

With Business Insight Advanced, you can create reports with relational or dimensional data sources, containing data in lists, crosstabs, and charts. You can also use your own external data source.

When you are using Business Insight Advanced, if you open a report that was authored in IBM Cognos Report Studio, you can see the objects that can be inserted only in Report Studio, such as maps. However, you cannot modify these objects.

Report Studio

IBM Cognos Report Studio is a report authoring tool that professional report authors and developers use to create sophisticated and manageable reports.

With Report Studio, users can create any BI report that an organization requires, such as weekly sales and inventory reports, by using any data source (relational or multidimensional).

Use Report Studio for reports that fit these characteristics:

- Are intended for a wide audience
- Exist long enough to require maintenance for changing requirements and data
- Require detailed control over the appearance

Report Studio provides powerful functionalities, such as bursting, prompts, maps, and advanced charting, and provides many ways to customize reports.

Query Studio

IBM Cognos Query Studio is the reporting tool for creating simple queries and reports in IBM Cognos BI.

A casual or novice user can use Query Studio to create self-serve reports that answer simple business questions. With minimal steps, you can view data, author basic reports, change the report layout, filter and sort data, add formatting, and create charts.

When you use Query Studio, you interact directly with your data. With Query Studio, you can work only with relational data sources.

Event Studio

IBM Cognos Event Studio is the notification tool that is used to alert decision-makers in your organization of events as they happen so that they can make timely and effective decisions.

You can use Event Studio to create agents that monitor status changes, priority customers, the data of your organization to detect occurrences of business events, or any other factor that is important to your business.

Specify the event condition, or a change in data, that is important to you. When an agent detects an event, it acts, such as sending an email, by adding information to the portal, and by running reports.

Metric Studio

IBM Cognos Metric Studio is the scorecarding and metrics tool in IBM Cognos BI.

You can use Metric Studio to create a customized scorecarding environment to monitor and analyze business metrics throughout your organization. Metric Studio helps you translate your organizational strategy into relevant, measurable goals that align the actions and accountability of each employee with a strategic plan.

The rich scorecarding environment of Metric Studio shows you where your organization is successful and where it needs improvement. It tracks performance against predetermined targets and indicates the status of the business so that decision-makers at every level of the organization can react and plan.

With Metric Studio, you can readily see how the organization is progressing against its strategy. You can set priorities for your own actions and understand how your decisions affect the performance of the company.

You can use the flexibility of Metric Studio to model metrics and their relationships based on any standard or proprietary scorecarding and management methodology that you already use.

Analysis Studio

Managers and analysts use IBM Cognos Analysis Studio to better understand their business and to get answers to questions that they have about their business. Users can quickly and easily perform analysis to emphasize what is behind an event or action so that they can improve business performance.

With analysis, it is possible to see trends and understand anomalies or variances that might not be evident with other types of reporting. Analysis Studio users can easily focus on what is important even with large volumes of dimensional data.

When you use Analysis Studio, you interact directly with visible data, knowing that Analysis Studio can work only with dimensional data sources.

IBM Cognos for Microsoft Office

When you use IBM Cognos for Microsoft Office, you can work with secure IBM Cognos BI content in your familiar Microsoft Office environment.

You can retrieve report content from various IBM Cognos applications, including IBM Cognos BI and IBM Cognos PowerPlay®. IBM Cognos for Microsoft Office provides access to all IBM Cognos report content, including data, metadata, headers, footers, and charts. You can use predefined reports, or create new content by using IBM Cognos Query Studio, IBM Cognos Analysis Studio, or IBM Cognos Report Studio.

By importing content into Microsoft Excel spreadsheet software, you can work with the data and use Microsoft Excel formatting, calculation, and presentation capabilities. You can also use the formatting and charting features of Microsoft Excel.

By importing content into Microsoft PowerPoint and Microsoft Word, you can include reports and charts to enhance your presentations and documents.

1.2.4 Cognos TM1

Many Finance departments short change their strategy support, enterprise performance management, and risk management activities because they are too busy reining in and validating data. They rely on systems that range from spreadsheets to enterprise resource planning (ERP) solutions. These systems are disconnected, expensive, managed by IT, and lack the performance and data reliability essential for on-demand analytics.

IBM Cognos TM1 capabilities span personal, functional, and enterprise requirements for planning, analysis, and reporting to help finance drive timely decision-making.

What can Cognos TM1 do

Cognos TM1 raises planning and what-if analysis to new levels of responsiveness. It ties Finance to operations to help uncover new business options, increase profits, and manage capital effectively. And it facilitates dynamic planning, forecasting and profitability analysis. Cognos TM1 offers these benefits:

- A personal analytics interface, including on-demand scenario modeling that is infused with planning capabilities
- Rapid development and deployment of flexible models, including profitability models, without batch processing
- Enterprise-wide reporting on important financial metrics
- Integration with IBM Cognos BI, including dashboards, scorecards, and standard reporting for general ledger, accounts receivable, and accounts payable
- Support for in-depth predictive and risk analytics
- Support for rolling forecasting with participation from a broad range of contributors
- Control by the finance organization, while business users manage their own work without the need for constant IT support

IBM Cognos TM1 radically transforms your entire planning cycle, from forecasting and analysis processes to budgeting and profitability modeling.

Cognos TM1 facilitates preferred practices, such as driver-based planning, rolling forecasts, and profitability analysis that are integrated with planning. It elevates the ability of Finance to manage performance proactively. And with the full range of planning processes tied to the operational frontline, planners and analysts can uncover hidden business options, increase profits, and manage capital more effectively.

IBM Cognos TM1 features

Cognos TM1 enhances your planning, analytics, and related reporting processes:

- Personalized analysis and planning desktop: This component enables analytics, planning, and reporting.
- Innovation modeling: Guided modeling, application design, and deployment for planning, analysis, and forecasting.
- Choice of interfaces: Spreadsheet, web, or Cognos Insight for non-spreadsheet users.
- High-performance analytics: Patented, 64-bit, in-memory OLAP for enterprise scalability and real-time scenario modeling to facilitate enterprise-wide planning and analysis.
- Security: Role-based security that supports multiple users and user types. Advanced predefined selection options ensure that users see only the necessary portions of the plan.
- Microsoft Excel formatting on the web: All planners and stakeholders can read, write, perform “what-if” analysis, and report with Microsoft Excel sheets right on the web.
- Administrative tools: Utilities for managing server performance and role-based access ease administrative tasks.
- Global support: Full Unicode support accommodates string lengths that are expanded beyond 255 characters. Language support for Japanese and simplified Chinese is included.
- Link to external applications: Because Cognos TM1 is Cognos 10 ready, you can link to enterprise reporting, scorecarding, and other desktop applications so that users can work with familiar interfaces.
- Connectors: Extend your ERP investments with the IBM Cognos data orchestration tool, which easily connects to data warehouses, Microsoft Excel worksheets, or any Open Database Connectivity (ODBC)-compliant databases.
- Cognos Performance Blueprints: Templates for improving your planning processes, optimizing the use of Cognos TM1 and reducing your time to value at less cost and with less risk. Available Cognos Performance Blueprints include these templates:
  - Capital Project Planning Performance Blueprint Powered by TM1
  - Expense Planning and Control Performance Blueprint Powered by TM1
  - Integrated Financial Planning and Forecasting Performance Blueprint Powered by TM1
  - Strategic Finance Performance Blueprint Powered by TM1
  - Workforce Planning Performance Blueprint Powered by TM1

These advanced capabilities help companies become highly collaborative and adaptive. Multiple users in functions, such as sales, marketing, and HR can analyze planning results – either online or offline – and explore alternative scenarios with an easy-to-use, graphical interface. Users can share plans in dashboards or reports, and easily prompt team members for comments and validation.

When a plan model is complete, a financial or business analyst can publish the analysis to the enterprise, then develop the model further, link it to other models to form a cohesive system, or deploy it to workgroups or others for their contributions. Cognos TM1 also enables planning team members to perform complex analysis, such as profitability, predictive, or risk analytics for even greater insight.

IBM Cognos TM1 supporting components

The following section is an overview of the components that IBM Cognos TM1 supports.

IBM Cognos Insight: The personal Business Analytics desktop

Cognos Insight, a new interface through which users can access Cognos TM1, transforms planning and analytics for the individual contributor, accommodating the non-spreadsheet user. Not locked in by spreadsheets, a planner can push the boundaries of participation as deeply and as widely as required and halt the development of shadow planning systems that hinder cooperation. Users who are intimate with the operations of the business can examine causes, trends, and options in greater depth, and develop plans that better reflect the complex dynamics of the business, by spotlighting drivers that need to be incorporated into the planning process.

With Cognos Insight, users are empowered to freely explore new plan models. They can pull the data of their choice onto their desktop, from managed or personal sources and without the assistance of IT. Users can build plan prototypes with personalized dimensions and hierarchies, and run what-if scenarios on the plan data. They can dive deeply into data to examine the profitability of the business and analyze hierarchies, such as product families down to the SKU level or customer family sets. They can view on demand the effect of a costing change. The quick, easy development and experimentation with the plans is aided by intuitive gestures for exploration, instant recalculations, and rich visualization.

To share planning ideas, users can collaborate with multiple groups throughout the enterprise:

- Promote their findings in a tailored dashboard in Cognos Business Insight for broad distribution.
- Publish the material to be read as a report in Cognos BI.
- Move the plan to the web as an application for either limited or widely distributed access.

Cognos TM1 Performance Modeler: Solution design and deployment

Cognos TM1 Performance Modeler, with its advanced modeling environment, provides the seasoned planner with a full array of analytics and model building tools as well as planning application deployment all with unprecedented ease of use. Because Cognos TM1 Performance Modeler eases application design with its guided development process, planners can employ the following preferred practices:

- Driver-based planning and rolling forecasting to enhance the contribution of the Finance department to enterprise performance management
- Broad involvement in the modeling process
- Frequent, rapid prototyping
- Deep linking of models to meet the broad range of enterprise planning and profitability analysis

Within Cognos TM1 Performance Modeler, the finance or business analyst can design a planning and analysis solution, such as an expense plan, or bring in an existing model from Cognos Insight. Designing the model includes these tasks:

- Define new dimensions, such as units or pricing for a profitability model.
- Build cubes, such as sales planning or customer profitability analysis, to support the planning requirements of business users.
- Create or modify a calculation by using preconfigured calculation expressions.
- Efficiently import data from multiple sources, mapping it to the model with drag-and-drop gestures.

With a user-friendly graphical interface, the analyst can link applications (or cubes) with simple drag-and-drop functionality, coupling financial resource management with operational performance commitments.

Cognos TM1 Performance Modeler includes these features:

- Three dedicated panes for content, working area, and properties.
- Wide use of drag-and-drop and what you see is what you get (WYSIWYG) functionality.
- Dockable editors.
- Simple spreadsheet-like dimension calculations.
- Graphical mapping for moving data between models and across dimensions within the models.
- An inter-editor refresh to highlight changes.

Cognos TM1 Performance Modeler also provides a process for deploying managed planning, analysis, and reporting that spans personal, workgroup, and enterprise requirements.

Cognos TM1 Performance Modeler helps increase user participation across departments and increase user commitment to achievable goals. Application design guides the planner through steps to set workflow, preferred client, and application views and rights.

For planning contribution, Cognos TM1 Performance Modeler provides a broad array of application styles and interfaces through Cognos Insight and Cognos TM1 Web Sheets.

Cognos Analytic Server: Distributed architectural advances for scale and interactivity

Cognos Analytic Server lets an organization choose centralized or distributed deployment methods for its planning activities, depending on the level of interactivity that you want. The “thin client”, Cognos TM1 Web, supports centralized planning. The distributed process supports higher levels of planning participation.

The high-performance OLAP engine at the core of Cognos Analytic Server enables on-demand, in-memory analytics for large or complex data sets with real-time query and calculation performance. Users can analyze frequently changing data for millions of items, such as SKUs or transactions. For example, users can perform these functions:

- Review sales daily or even hourly, by store, brand, or sales representative.
- Analyze day-by-day profitability and customer churn for thousands or even millions of customer accounts.
- Track trends in spending over weeks or months.

The 64-bit OLAP technology of Cognos Analytic Server meets even the most complex, multidimensional analytics needs of large-scale operations. A user can query data whenever needed, no matter how large the data set. In addition, the user can view instant updates from streamed data and drill through to transaction systems for added context and thus greater accuracy in decision making.

To rapidly access valuable information that is locked away in back-office applications, Cognos Analytic Server includes data connectors that are shared with IBM Cognos BI to streamline the loading of disparate data into models. Users can also create their own processes to automate data importation, metadata management, and other tasks.

The Cognos TM1 Operations Console, a web-based tool, makes it easy to collectively manage Cognos TM1 servers on a daily basis. Administrators can dynamically monitor, sort, and filter thread activity on Cognos TM1 servers, including monitoring threads that run on multiple Cognos TM1 server instances. The Cognos TM1 Operations Console also provides a health check feature, which determines the current state of each monitored server.

1.3 Predictive Analytics in business: SPSS

The ability to predict outcomes with a reasonable degree of confidence is of great strategic importance in current businesses. To address this need, the Statistical Package for the Social Sciences (SPSS) is available as part of the IBM Business Analytics solution portfolio.

SPSS consists of a comprehensive set of tooling that uses data for decision-making purposes. It is a solid product with over 40 years of presence in the market. It can be used to collect data, provide statistics, create predictive models, and deploy all of these analysis into your business.

Figure 1-2 on page 17 summarizes what you can accomplish with SPSS. The product encompasses various techniques from statistics, data mining, and text mining that analyze current and historical facts to predict future events. Notice that the software suite helps you answer the important question of “Which action is likely to drive the best outcome?”

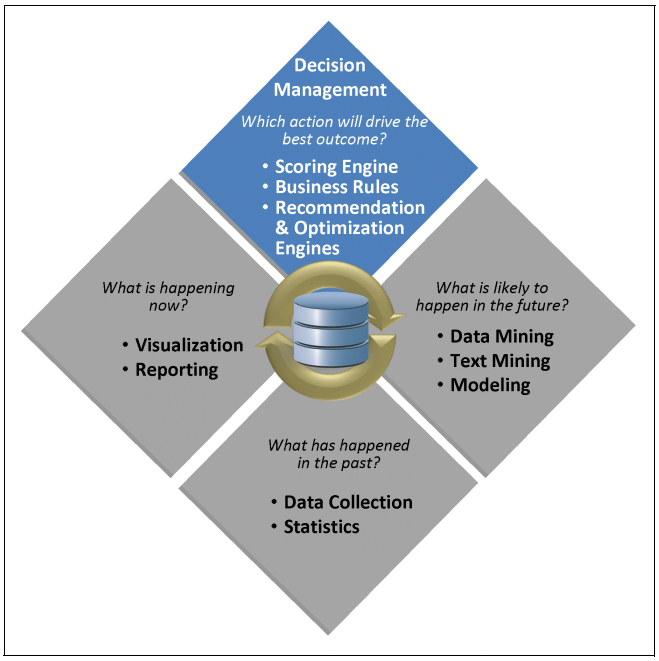

Figure 1-2 Predictive analysis with SPSS

The capability to predict is an advantage in the return on investment (ROI). You might have a good sense of your business, or eventually some pre-built rules to help with decision making. However, it is not until you use analytics in decision management that you can, sustainably, choose the best path to follow, for every point of impact of your business.

Predictive analytics can capture insights from historic data patterns and provide you with evidence from this data. Predictive analytics is flexible in its modeling to understand changes in trends and can analyze massive amounts of structured and unstructured data. Also important, it can promptly avail the information learned from the insights to everyone who needs access to it. It increases ROI because business areas are not analyzed independently. Cross-departmental analyses are performed. These analyses do not leave out any business relationships that were uncovered or considered too small. As a result, every aspect of the business is optimized with these analyses.

The ROI in organizations that do not use predictive analytics is much lower than the ROI in organizations that do use it. An opportunity is better taken if you are prepared for it. Alternatively, money can be saved if you simply prevent the situations in which you must spend more to fix their outcomes, such as customer churn.

It is with these advantages in mind that we now present the IBM SPSS software suite.

1.3.1 The IBM SPSS Software suite

Figure 1-3 on page 18 describes the classification of the IBM SPSS portfolio of products that you can use for predictive analytics.

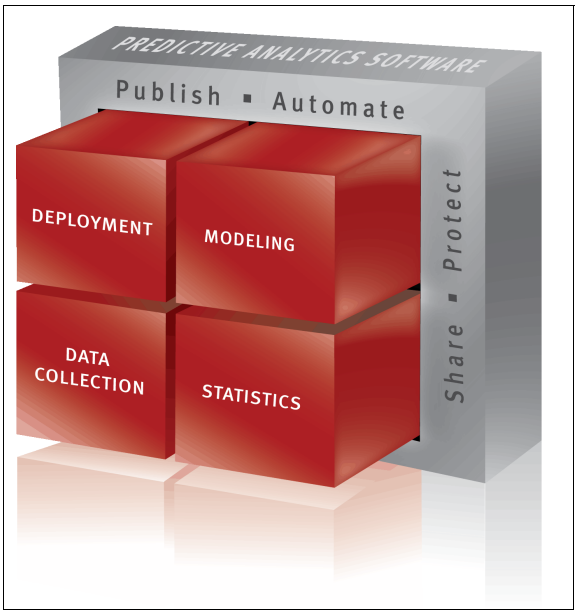

Figure 1-3 IBM SPSS portfolio classification

Predictive analytics with the SPSS software stack is designed to work according to the three steps of the analytical paradigm: align, anticipate, and act. It helps with data collection, uses statistics and modeling to predict outcomes, and helps you deploy your insights. With IBM SPSS, you can publish, automate, share, and protect your analytic assets.

Figure 1-4 gives you a summarized view of the SPSS predictive analytics methodology.

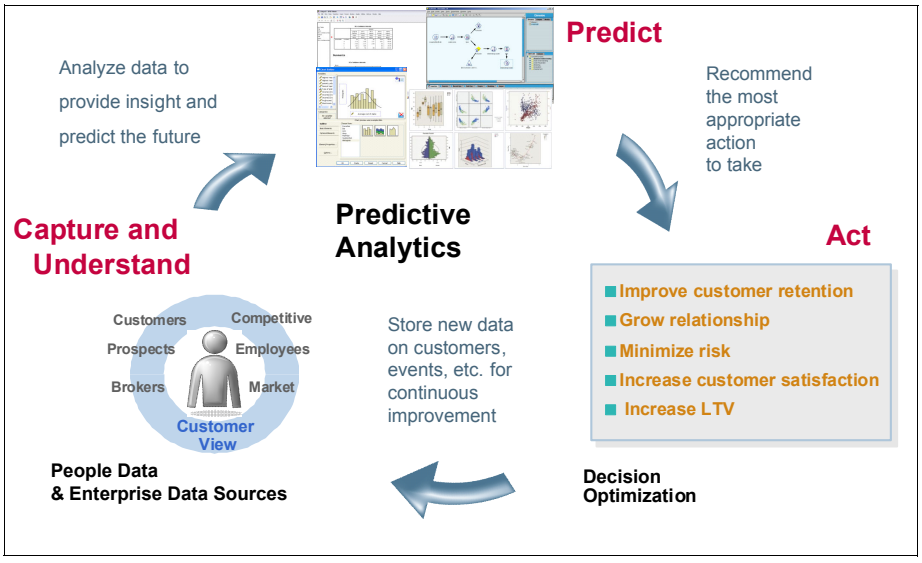

Figure 1-4 SPSS predictive analytics methodology

IBM SPSS software is organized in four product families:

- Data collection

  Allows the collection of data from customers and enables an accurate view of attitudes and opinions

- Statistics

Provides evidence that is based on data and verifies hypotheses

- Modeling

  Works out accurate predictions to aid decision making

- Deployment

  Helps you act upon the impact points in your operations

The following sections overview and highlight the products from each of these families.

Data collection

This family is represented by the IBM SPSS Data Collection software. It is intended to be used by survey, market, and business researchers intent on acquiring clean data from many data sources and by using different methods. The result is a set of data about attitudes and preferences (potentially of your customers). With this information, you can gain valuable insights about the behavior of your customers.

With the IBM SPSS Data Collection software, you can create surveys easily. You can apply them by deploying interviews through many technologies (web, phone, or in-person), report the results of these surveys, and effectively manage each survey.

Statistics

This family is represented by the IBM SPSS Statistics software suite. It is based on sophisticated mathematics to validate hypotheses and assumptions and is widely used by government, commercial, and academic institutions to solve business and research problems. You can use it to test an opinion on a new product, predict the acceptance of ideas, experiment with allocation within a supply chain, or test the efficiency of a medical treatment. This software uses data to back up (or not) your theories. With this backup, you are more confident when making a decision.

Modeling

This family is represented by the IBM SPSS Modeler software. The previous family, statistics, is used to test hypotheses. With the Modeler software, you can create business models to predict future outcome. It uncovers hard-to-identify relationships within structured and unstructured data that seems unrelated at first. You can predict the future and understand what happens based on what happened before. This capability is useful to prevent customer churn, for example, and to help people consistently make decisions. Another benefit of this software is that it can also explain which factors drive future outcomes. You can use it to mitigate risks and take advantage of opportunities.

With IBM SPSS Modeler, you can create models in an intuitive and quick fashion, without programming. Modeler includes the advanced, interactive visualization of models. Multiple techniques can be included within a model, and the results are easy to understand and communicate to your staff.

Deployment

This family is represented by the following software:

- IBM SPSS Decision Management
- IBM SPSS Collaboration and Deployment Services

The first software, Decision Management, is intended to automate and optimize small decisions that are made in day-to-day business operations in real time. It combines predictive analytics with business rules. The models are created in an easy-to-use interface with which the business user interacts without the specialized help of an analyst, statistician, or data

miner. This more independent process allows people at any level of the organization to create automated models for making small decisions, thus helping to optimize every aspect of the overall business operation.

The second software, Collaboration and Deployment Services, is analyzed in detail in an upcoming section.

IBM SPSS products and features

IBM SPSS offers a wide range of products to perform various tasks. Table 1-2 summarizes the available software and features.

Table 1-2 IBM SPSS products and features

Products	Features
IBM SPSS Statistics Base	The core capabilities that you need to take the analytical process from start to finish.
IBM SPSS Advanced Statistics	More accurately analyze complex relationships by using powerful univariate and multivariate analysis.
IBM SPSS Bootstrapping	Create more reliable models and generate more accurate results.
IBM SPSS Categories	Predict outcomes and reveal relationships in categorical data.
IBM SPSS Complex Samples	Correctly and easily compute statistics for complex sampling.
IBM SPSS Conjoint	Easily discover what people value.
IBM SPSS Custom Tables	Easily analyze and communicate your analytical results.
IBM SPSS Data Preparation	Improve data preparation for more accurate results.
IBM SPSS Decision Trees	Easily identify groups and predict outcomes.
IBM SPSS Direct Marketing	Improve your marketing campaigns and maximize the ROI of your marketing budget.
IBM SPSS Exact Tests	More accurately analyze small data sets or data sets with rare occurrences.
IBM SPSS Forecasting	Build expert time-series forecasts in a flash.
IBM SPSS Missing Values	Build better models when you estimate missing data.
IBM SPSS Neural Networks	Find more complex relationships in your data.
IBM SPSS Regression	Improve the accuracy of predictions with powerful regression procedures.
IBM SPSS Advantage for Excel 2007	Specialized tools extend the capabilities of Excel to manage and analyze business data sets.
IBM SPSS Amos	Take your analysis to the next level and get your research noticed.
IBM SPSS Statistics Programmability Extension	Dramatically increase the power and capabilities of IBM SPSS Statistics.
IBM SPSS SamplePower	Get the right sample size the first time.
IBM SPSS Statistics Developer	A complete workbench for Python programmers to integrate analytical functionality and procedures.

Products	Features
IBM SPSS Statistics Server	Faster performance-customized analytic access for business users and greater security in enterprise deployments.
IBM SPSS Text Analytics for Surveys	Analyze survey text and discover valuable, hidden insights.
IBM SPSS Visualization Designer	Charting your course just got easier.
IBM SPSS Modeler	A data mining workbench that helps you build predictive models quickly and intuitively, without programming.
IBM SPSS Modeler Text Analytics	Sophisticated text analytics helps you to unlock concepts that are trapped in unstructured data.

Note: The following software from Table 1-2 is not available for AIX: IBM SPSS Amos, IBM SPSS Text Analytics for Survey, and IBM SPSS Visualization Designer. An updated list of operating system requirements for SPSS software is at this website:

http://publib.boulder.ibm.com/infocenter/prodguid/v1r0/clarity/osForProduct.html

1.3.2 IBM SPSS Collaboration and Deployment Services (C&DS)

The IBM SPSS C&DS falls into the Deployment family of the SPSS suite, as explained in Figure 1-3 on page 18.

The role of IBM SPSS C&DS is to enable widespread use and deployment of predictive analytics. It provides a centralized, secure, and auditable placeholder of analytical assets and advanced capabilities for management and control of predictive analytic processes, as well as sophisticated mechanisms to deliver the results of these assets to users.

IBM SPSS CS&D has three key functionalities:

- Collaboration
- Automation
- Deployment

Collaboration

Collaboration enables efficient sharing and reusing of analytics assets. It protects them in ways that meet internal and external requirements. Collaboration ensures that results from these assets are published to a greater number of users, who can view and interact with them. An illustration of this functionality is depicted in Figure 1-5 on page 22.

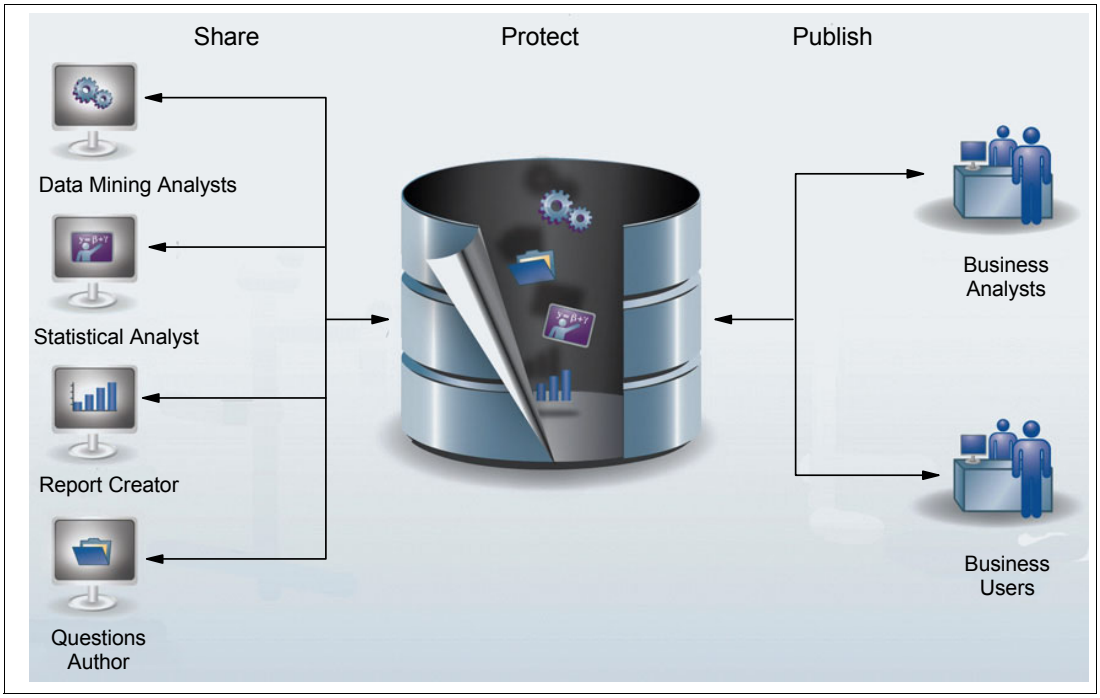

Figure 1-5 IBM SPSS C&DS: Collaboration

Automation

Automation allows the use of analytical assets as a core component of your organization decision-making process. This capability fosters the achievement of consistent results across the entire organization. An illustration of this feature is depicted in Figure 1-6 on page 23.

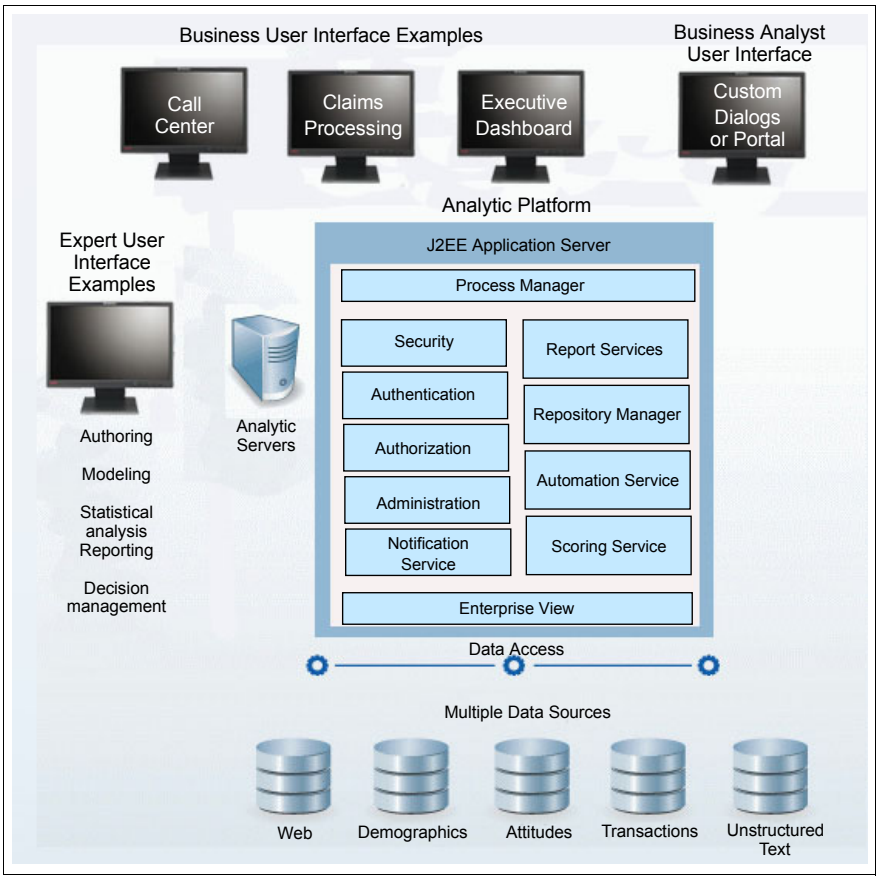

Figure 1-6 IBM SPSS C&DS: Automation

Deployment

Deployment makes it possible for organizations to use analytics in front-line business processes, thus closing the align, anticipate, then *act* paradigm of using analytics in the business world. IBM SPSS Collaboration and Deployment Services facilitates the integration of all of the SPSS predictive features into your existing infrastructure by using standard programming tools and interfaces. An illustration of this feature is depicted in Figure 1-7 on page 24.

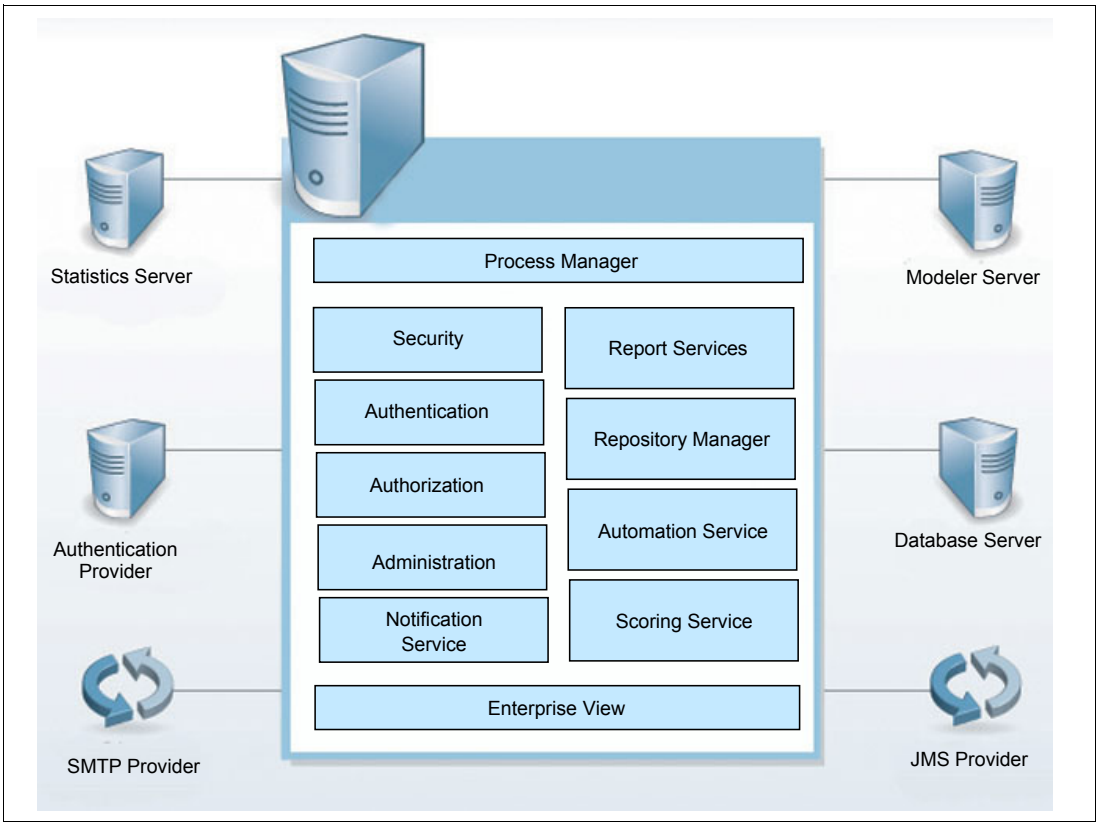

Figure 1-7 IBM SPSS C&DS: Deployment

Benefits of using IBM SPSS Collaboration and Deployment Services

This list summarizes the advantages of using IBM SPSS C&DS as part of the predictive analytics software stack offered by IBM:

- Control analytical processes by centralizing and automating the evaluation and deployment of models
- Enhance model accuracy by performing champion and challenger tests
- Generate real-time scorecards that are based on models to support decisions
- Integrate key business processes and analytics
- Safeguard the value of analytical assets
- Foster the reuse of analytical assets, resulting in widespread use and minimizing the costs of asset management

IBM SPSS C&DS architecture

The diagram in Figure 1-8 on page 25 illustrates the typical architecture of an SPSS C&DS deployment, which consists of these components:

- The central SPSS C&DS repository
- The database server
- Execution servers
- Client servers that access the SPSS C&DS repository

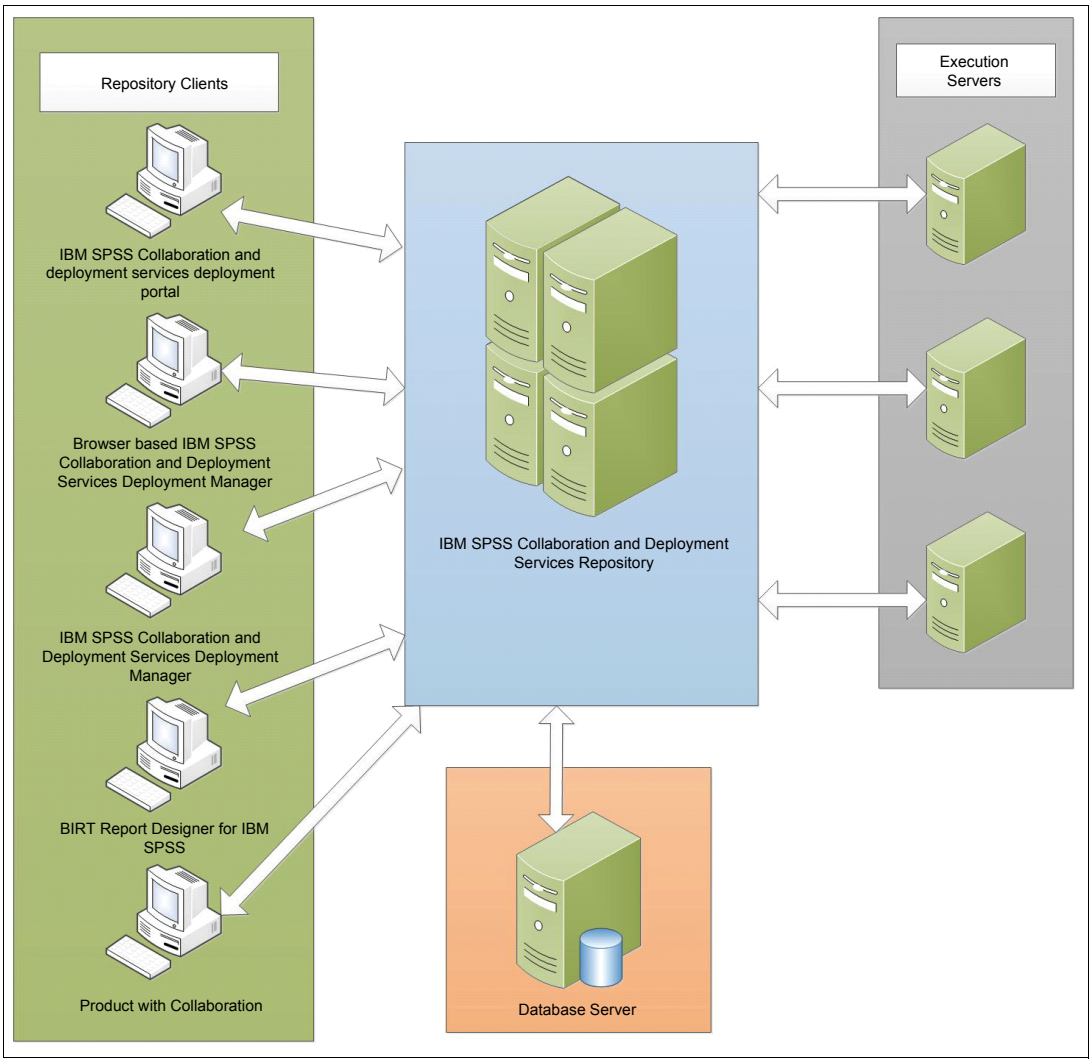

Figure 1-8 SPSS C&DS architecture

In this architecture, all of the analytics assets are stored on the SPSS C&DS repository. Clients can access these assets through web services or by using specialized client tooling for communications.

The requests that are performed by these clients are sent to execution servers, which perform all of the work on top of the analytics data. The results are then stored onto C&DS and can be accessed by the requesting clients.

IBM SPSS C&DS architecture components

The following list describes the components that are presented in the architecture of Figure 1-8:

- IBM SPSS Collaboration and Deployment Services Repository

  Used for collecting and storing analytical assets. It includes models and data at a centralized location.

- IBM SPSS Collaboration and Deployment Services Deployment Manager

  Responsible for creating, executing, and automating the analytical task. It includes updating the model stored in repository by users.

- IBM SPSS Collaboration and Deployment Services Deployment Portal

  This web browser-based thin-client interface accesses the IBM SPSS Collaboration and Deployment Services Repository, running analyses, and viewing output.

- BIRT Report Designer for IBM SPSS

  Ad hoc reports against relational and file-based data sources can be created by using BIRT Report Designer for IBM SPSS.

- IBM SPSS Collaboration and Deployment Services Enterprise View Driver

  To Access IBM SPSS Collaboration and Deployment Services Enterprise View objects stored in the repository, including IBM SPSS Statistics and third-party applications.

- Browser-based Deployment Manager

  Used by the SPSS administrator for performing, tuning, and updating system management tasks.

IBM SPSS C&DS prerequisites

This publication refers to IBM SPSS C&DS Version 4.2.1.0. The prerequisite software for this version is summarized in Table 1-3. Only one application server, database, and browser are required. However, Table 1-3 lists all of the supported alternatives.

Table 1-3 IBM SPSS prerequisite software

Application servers	**Version**
WebSphere Application Server	6.1, 7.0
JBoss	4.2, 5.1
JBoss EAP	4.3, 5.0
Oracle WebLogic Server 11g	(10.3.1)
Oracle WebLogic Server 11gR1	PS1 (10.3.2)[a]
Oracle/BEA WebLogic Server	10.0
Databases	**Version**
DB2® Enterprise Server Edition	9.1, 9.5, 9.7
DB2 for i5/OS®	V5R4, 6.1, 7.1
Microsoft SQL Server	2005, 2008
Microsoft SQL Server Enterprise Edition	2008 R2
Oracle Database 10g Standard/Enterprise Editions	Release 2
Oracle Database 11g Standard/Enterprise Editions	Release 1, Release 2
Web browsers	**Version**
Apple Safari	4.0
Microsoft Internet Explorer	7.0, 8.0, 9.0
Mozilla Firefox	2.0, 3.0, 4.0

a. Requires IBM SPSS C&DS 4.2.1 Fix Pack 2 or higher.

IBM SPSS C&DS related software

The software in the following table can be used with IBM SPSS C&DS as a data source or as an optionally integrated collaboration product. These products are not prerequisites for running IBM SPSS C&DS. This table relates to IBM SPSS C&DS Version 4.2.1.0. Table 1-4 lists all of the related software.

Table 1-4 IBM SPSS related software

Data sources	**Version**
DB2 Enterprise Server Edition	9.1, 9.5, 9.7
DB2 Express Edition	9.1, 9.5
DB2 Workgroup Server Edition	9.1, 9.5
DB2 for i5/OS	V5R4, 6.1, 7.1
DB2 for z/OS®	9.1
InfoSphere Classic Federation Server for z/OS	V9.5, V10.1
InfoSphere Master Data Management Server	10.0
WebSphere Classic Federation Server for z/OS	9.1
HP Neoview	2.4
Microsoft SQL Server	2005, 2008
Microsoft SQL Server Enterprise Edition	2008 R2
MySQL	5.1
Netezza® Data Warehouse	4.5, 5.0, 6.0
Oracle Database 10g Standard/Enterprise Editions	Release 2
Oracle Database 11g Standard/Enterprise Editions	Release 1, Release 2
Sybase IQ	12.7, 15.1
Teradata	12, 13
Collaboration	**Version**
SPSS Data Collection Desktop	6.0.1.0
SPSS Modeler	15.0
SPSS Statistics Desktop	20.0.0.0
SPSS Data Collection	5.6
SPSS Decision Management	6.1, 6.2, 6.0
SPSS Statistics	18.0

SPSS Statistics Desktop: SPSS Statistics Desktop is not available for AIX. An updated list of operating system requirements for SPSS software is at this website:

http://publib.boulder.ibm.com/infocenter/prodguid/v1r0/clarity/osForProduct.html

1.3.3 SPSS business scenarios

We describe two business scenarios in which you can use IBM SPSS to gain a business advantage with predictive analytics:

- Insurance companies
- Retail

The business of insurance companies

Two of the greatest problems that insurance companies face are customer churn and claim handling. The goal of a company is to retain the most profitable customers and minimize the costs of handling the less profitable customers.

Companies that choose to investigate every claim in depth to avoid paying for the fraudulent claims pay for a costly and time consuming process. So, a balance must exist. Customer churn can arise when a company investigates an honest customer. The customer might become dissatisfied if it takes too long for the claim to complete or if the customer senses suspicion on the part of the insurance company. These situations are false positives. They happen because claims for review are usually selected based on the feeling of the handlers.

IBM SPSS Customer Retention and Growth is part of the SPSS software suite and uses advanced algorithms to identify customers at risk of churning. The tool is able to suggest retention strategies and strengthen customer relationships by analyzing the attitudes, preferences, and needs of customers. The retention offers are customized, which results in customer loyalty and increased satisfaction and profitability levels.

IBM SPSS Predictive Analytics and Reporting for Claims can be used to minimize fraud and unnecessary payments. The algorithms use multiple data as input, such as customer data, interaction data, campaign data, and internal expertise, to score the claims according to risk. The claims are classified as low, medium, and high risk. Low-risk claims can then be given preference in processing (fast-tracked). This preferential treatment results in a high level of quality of service to these customers, driving satisfaction and loyalty to higher levels. High-risk claims can be sent to the correct investigation channels. The process of scoring each claim becomes smarter with each iteration, because the predictive model can learn from the claims that it processes. The expected result is to see fewer false positives over time.

The business of retail companies

Retailers face challenges to create successful promotions on their products. If they fail by targeting the wrong customers or the wrong products, the invested time and money do not provide the expected return on marketing investments.

When creating a promotion, executives and marketers need to balance inventory costs against lost sales opportunities, and also understand the buying preferences of their customers. Customer loyalty is built when a consistent service is delivered, by offering the right products to the right people.

IBM SPSS Market Basket Analysis is part of the IBM SPSS portfolio of solutions. It uses predictive algorithms to find product combinations that increase sales. It ensures that offers are targeted to the right customers by crossing the data against the needs and preferences of the customer.

IBM SPSS Market Basket Analysis analyzes point-of-sale data to identify a pattern of which products tend to be bought together. The result is the creation of sets of products that can be joined in a promotion to foster a combination sale. Offers might include discounts on the second product, for example. By using information about your customers, such as demographics, behavior, interaction, and attitude, you can target these offers to the right set of consumers, thus maximizing the return on marketing investments.

Also, with *IBM SPSS Assortment Planning*, you can optimize inventory based on predicted sales of a certain item in a certain store location based on customer buying behavior.

1.4 IBM Power Systems

We see how important analytics is in business. Analytics allows us to act at the business points of impact by applying a powerful and efficient data analysis. It helps us to use historic data to uncover trends. Analytics helps us consume real-time data to understand what is happening at that moment so we can make a decision that yields a better outcome. And to achieve outstanding results, the better hardware we run our analytics solutions on, the quicker we can make these decisions. Because timing is a critical variable to success, we suggest that you run analytics on the IBM Power hardware.

The hardware, virtualization, and operating system features that make the Power hardware the infrastructure on which to run analytics are highlighted.

1.4.1 Overview of IBM POWER7 family features

The IBM POWER7® processor is the latest generation of the IBM POWER family of processors. It has a distinctive architecture in terms of cache setup and the number of cores on the same chip. It also supports a higher level of parallelism in thread execution. In the following sections, we explore each characteristic that makes the POWER7 architecture the best choice for running your workloads.

Number of cores per chip

Before the POWER7 line of processors, we used dual-core chips in commercial servers. Dual-core chips allowed these cores to share some of the resources on the chip on which they both reside, such as L2 cache, and increased the speed of communication between them. Today, the POWER7 architecture supports a configuration of either 4, 6, or 8 cores on the same chip. There are more cores on the same chip, thus making it faster to transfer data among these various cores. There is also the benefit of on-chip cache sharing among more cores.

Another advantage of having more cores on the same chip is a reduced energy consumption by core. This reduced energy consumption by core allows the Power 795 with 256 cores to consume the same energy that a POWER6® 595 with 64 cores consumes.

These improvements to the POWER7 chip architecture present potential gains in the processing speed of computational intense workloads, such as analytics.

L2 and L3 cache

The L2 and L3 cache are high-speed memory areas that processors use to store information that they are currently processing. In the previous POWER chip family, the POWER6, the L3 cache was stored outside of the chip. So, accessing this cache layer was slower, because memory transfers happened from the outside of the chip to the inside of the chip through an L3 cache bus controller, as depicted in Figure 1-9 on page 30.

In the POWER7 architecture, the L3 cache was brought onto the chip, so access to its contents happen much faster because it is an on-chip operation. Also, the bus that controls the data flow of the L3 cache was enhanced: the Power GX Bus. Multiple memory controllers are also now available, as opposed to only two, to exchange data with the main memory of the system (RAM memory). The new architecture is depicted in Figure 1-9.

Each core still has access to its own, private L2 cache, which you can check on Figure 1-9. The total amount of L3 cache is 32 MB for 8-core chips, 24 MB for 6-core chips, and 16 MB for 4-core chips. Each core can access 4 MB of L3 cache.

These enhancements to the L2/L3 cache chip architecture also present a potential computational gain to intensive workloads. It is thus another advantage to analytics workloads.

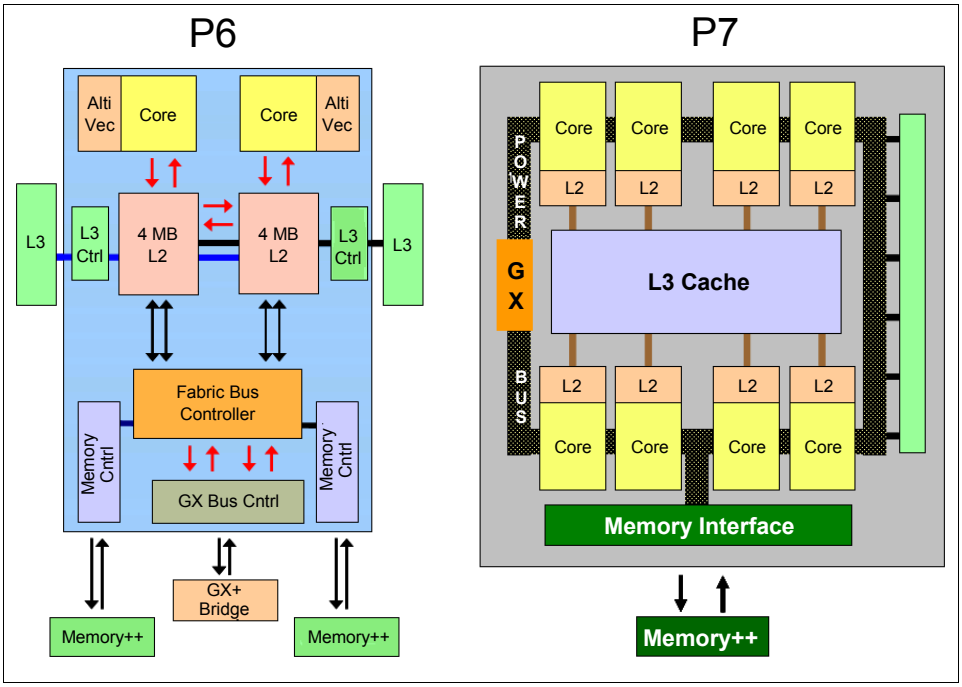

Figure 1-9 Comparison between the POWER6 and the POWER7 chip architectures

Multithreading evolution: SMT4

Before the advent of the POWER7 chip, commercial systems were composed of cores that supported only two simultaneous threads (SMT2), such as the POWER5 and POWER6 cores. A *thread* is an instance of a running process or program that can concurrently run with its counterpart on the same core. The concurrency is possible when each of the threads uses a different set of processor registers at a particular point in time. They can be executed in parallel, thus allowing the program to run faster.

The POWER7 chip introduced the concept of four simultaneous threads. That is, four of your program threads request the use of four different sets of processor registers. These four threads can execute in parallel, allowing your program to run even faster than it would using two simultaneous threads. This behavior is depicted in Figure 1-10 on page 31.

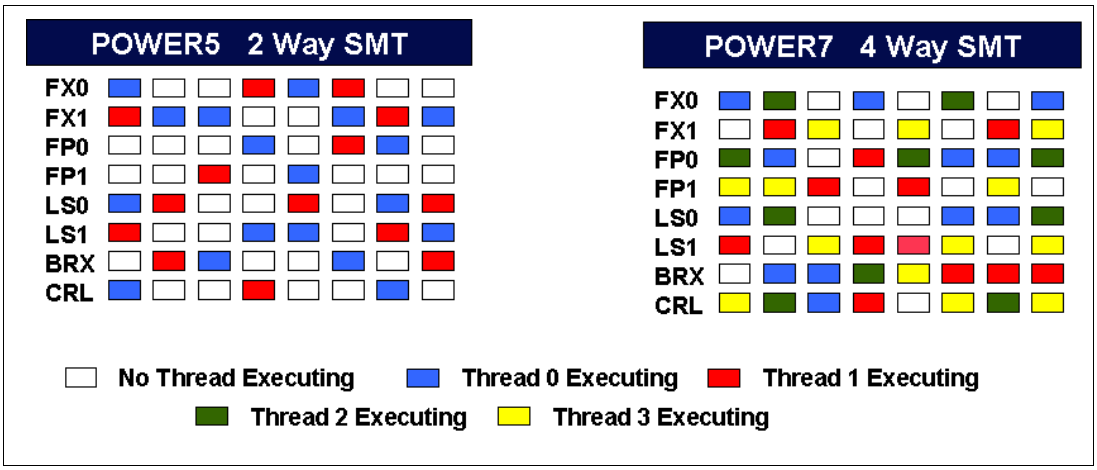

Figure 1-10 Threading model enhancement in the POWER7 architecture: SMT4

An increased number of thread executions means that your program executes faster if it is able to take advantage of this parallelism. Again, only the Power hardware provides this advantage to execute a workload, such as analytics.

TurboCore mode on IBM Power Systems

An innovation of the POWER7 processor architecture is the TurboCore mode on IBM Power Systems. This mode applies only to the 8-core chips, and it is available for the Power 780 and Power 795 machine models only.

Recall the POWER7 processor architecture from Figure 1-9 on page 30. In normal mode, all of the cores on the chip are active. When you turn on TurboCore mode, half of your chip cores are turned off. TurboCore mode offers these advantages:

- The L3 cache is now shared by half of the cores only; therefore, each core now has access to 8 MB of L3 cache as opposed to 4 MB.
- The deactivated half of the cores stop consuming power; thus, they also stop generating heat.

Because the chip stops generating a certain amount of heat when half of its cores are shut down, the remaining cores can be “overclocked”. This extra-generated heat is compensated by the savings in heat generation of the shutdown cores. The overclocking makes the core frequency go from 3.8 GHz to 4.1 GHz.

Figure 1-11 on page 32 depicts an 8-core chip in TurboCore mode.

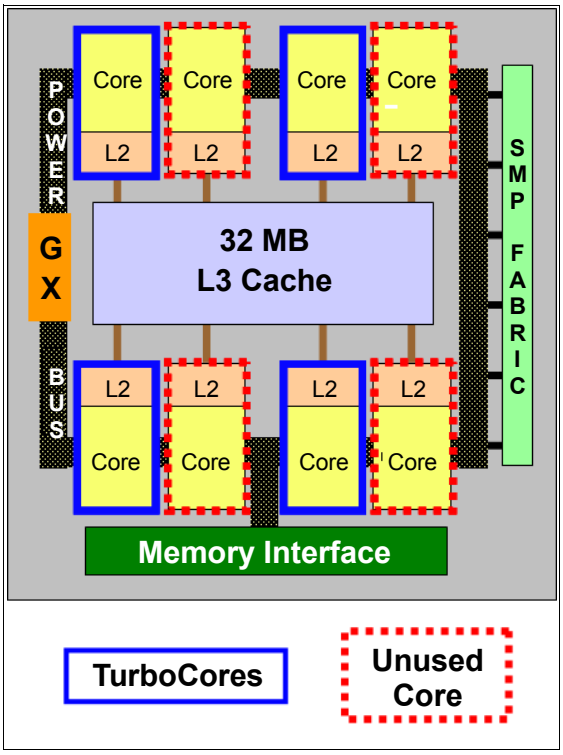

Figure 1-11 POWER7 chip in TurboCore mode

TurboCore mode on IBM Power Systems can boost your processing due to more L3 cache available to your cores (higher cache hit ratio), and faster processing due to overclocking. These benefits are a significant advantage to your analytics environment.

1.4.2 PowerVM

PowerVM is the set of virtualization products and features available on the Power hardware. Most of these products were introduced in the launch of the Power 5 hardware in the last decade. These technologies are improved and widely accepted by clients as mature, safe, and stable.

Virtualization serves the purpose of hosting multiple systems on the same piece of hardware. Each system is contained within a logical partition (LPAR) and has access to processor, memory, and I/O resources, which are the three required elements for a system to exist. The virtualization of processor and memory is controlled by the Power hardware firmware, called the POWER Hypervisor™. I/O virtualization also occurs through the help of an auxiliary system, the Virtual I/O Server (VIOS).

With virtualization, you can achieve higher hardware utilization levels, because it minimizes the amount of time that resources are idle through the *sharing* of resources. It also enables a phased plan for resource deployment, because it makes it possible to *aggregate* resources dynamically or to increase the amount of resources of a logical partition dynamically. Virtualization can also provide compatibility with existing equipment, because a piece of physical hardware can be delivered to a logical partition with different characteristics. This function is called hardware emulation or *extension*. Finally, virtualization allows a *transparent change* of physical resources during a maintenance window, for example. The four advantages that are provided by virtualization on the Power platform are depicted in Figure 1-12 on page 33.

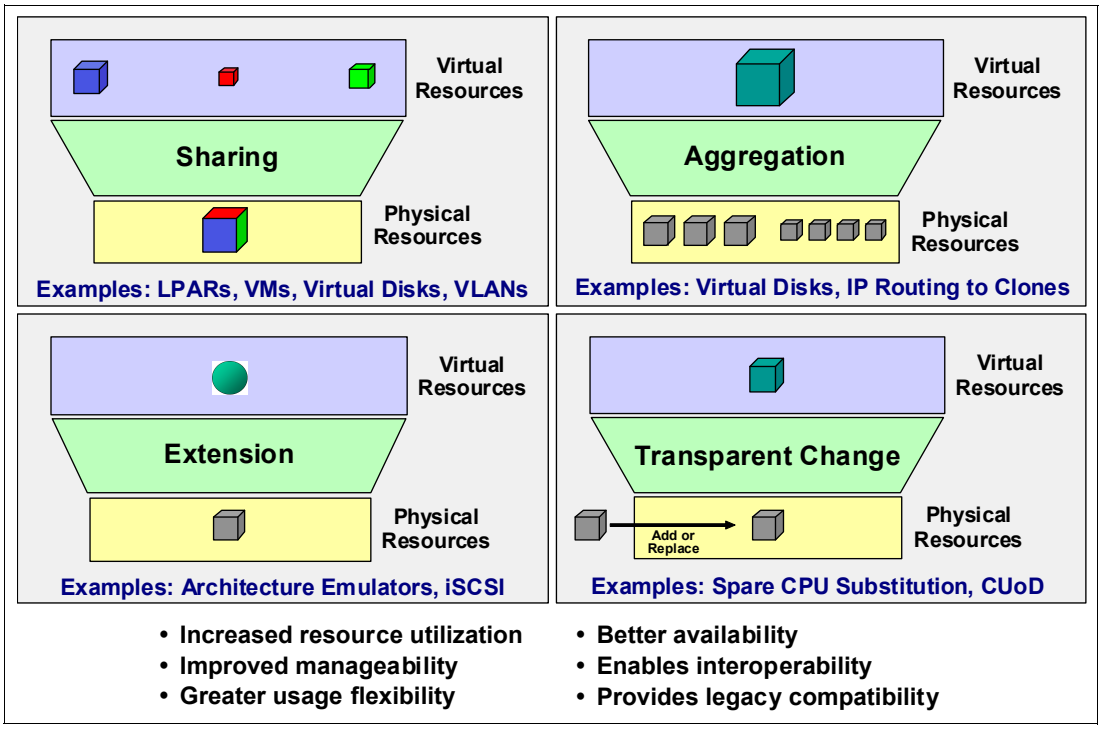

Figure 1-12 Advantages of virtualization

A PowerVM feature allows one LPAR to be moved from one physical Power machine to another, in a dynamic manner that incurs no downtime, called Live Partition Mobility (LPM). This technology permits that you move your systems out of a frame, so that you can, for example, provide maintenance to it. This technology is perfect for analytic systems that need to be up and running all the time.

The next few sections provide an overview of the particular PowerVM features that are used in our Cognos and SPSS environment setup. For a complete and thorough understanding of the PowerVM technology, see *IBM PowerVM Virtualization Introduction and Configuration*, SG24-7940.

POWER Hypervisor

The POWER Hypervisor is the core of IBM PowerVM technologies. It is a firmware layer of code between the machine bare metal and the logical partition operating systems. It controls multiple features that you can use when virtualizing your Power hardware:

- Processor virtualization and Micro-Partitioning
- Processor affinity
- Memory virtualization and sharing
- Communication between virtual adapters for network and disk I/O, and system consoles

The Hypervisor is an abstraction layer on top of the physical hardware, as shown in Figure 1-13 on page 34. It enforces partition integrity by providing a security layer between the logical partitions so that one does not interfere with the other. It controls the hardware I/O interrupts and management facilities for the partitions. The Hypervisor organizes how much and which physical resources are assigned to each one of your logical partitions, based on the configuration of your partitions.

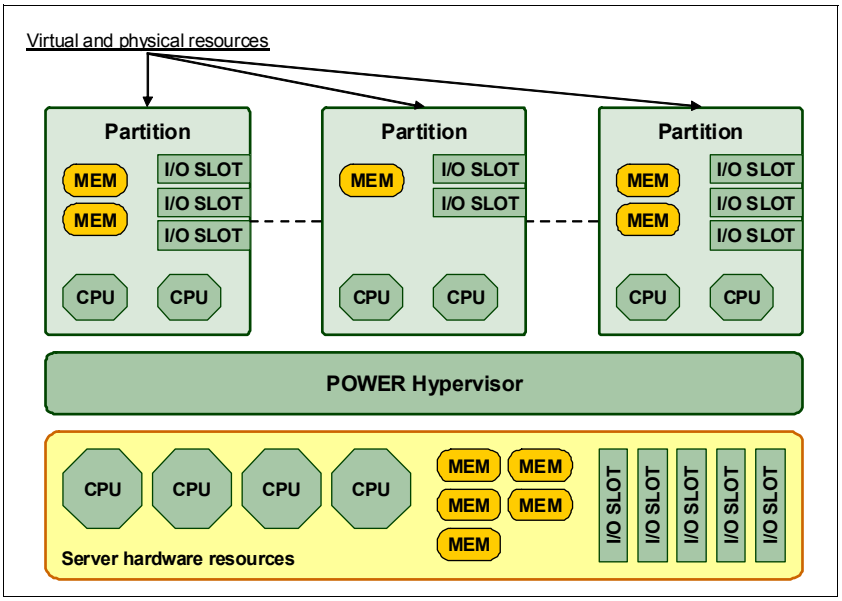

Figure 1-13 The Hypervisor abstraction layer

Processor virtualization and Micro-Partitioning

Power Systems can be equipped with as many as 256 processor cores (Power 795 full processor capacity). It is impractical, though, that all of these processors are assigned to the same workload. Usually, these cores are distributed to various workloads, each one contained within a logical partition. So, the overall processing capacity of a Power system is divided up to address the requirements of each logical partition. This division of the overall processing capacity is controlled by the Hypervisor, as explained in “POWER Hypervisor” on page 33.

Processors can be assigned to logical partitions as either an amount of physical processor cores (dedicated processors) or as a fractionate amount of processing power (shared processors).

In the first scenario, an entire processor core is provided to the logical partition to process its workload. The operating system sees a specific number of cores that correspond exactly to the number of physical cores at the disposal of that logical partition. Eventual idle processing cycles can be donated to other LPARs, but they are usually not configured in that manner. Dedicated processors can provide slightly more performance than shared ones, but also usually result in lower machine utilization rates due to unused idle cycles.

In the second scenario, a fraction of as low as 1/10th of a processor core can be assigned to a logical partition. You can have a more fine-grained tuning of the processor requirements of your workloads by assigning a more accurate amount of processing power to each one of the logical partitions within the Power frame. This approach is called *Micro-Partitioning,* which is how we can virtualize the use of processors on the Power Systems. It allows for better processor utilization levels and processor sharing among logical partitions (LPARs).

For a more thorough understanding of how processor virtualization works on the Power hardware and the architecture behind dedicated and shared processors, see *IBM PowerVM Virtualization Introduction and Configuration*, SG24-7940.

Processor affinity

Due to the possibility of Micro-Partitioning and processor virtualization on the Power hardware, the Hypervisor needs to control which amount of processing power is delivered to each of the logical partitions on the system. This control is according to a time-base processor dispatching cycle, as explained in Figure 1-14. For example, your LPAR uses Micro-Partitioning and is assigned 1.4 processing units. In an interval of 10 ms of a Hypervisor processor dispatch cycle, your LPAR is able to use 14 ms of processing time, which is the equivalent processing capacity of 1.4 processors in that cycle. See Figure 1-14.

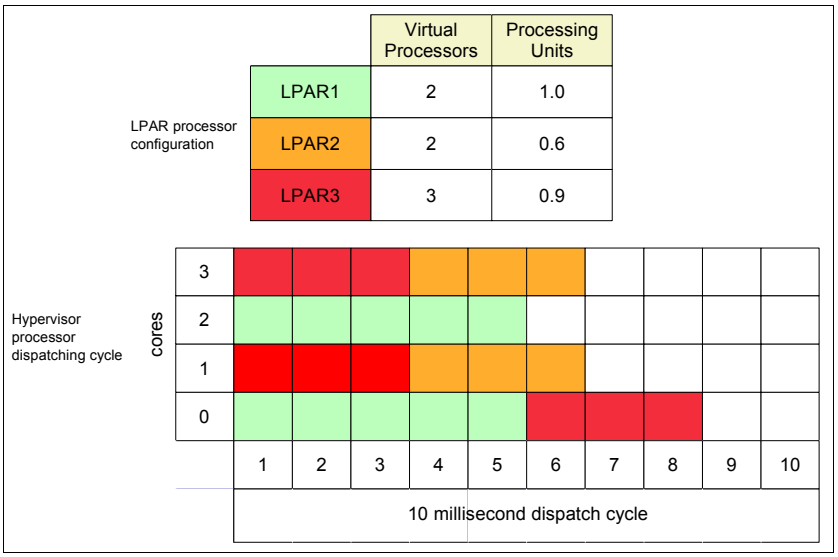

Figure 1-14 Hypervisor dispatch cycle: Processor affinity applied

When using Micro-Partitioning, the processing power of an LPAR can come from any physical processor that has available cycles during the Hypervisor processor dispatch window. However, to minimize the number of processor context switches and to better use the L2/L3 processor cache, an LPAR must receive its processing cycles from the same underlying physical processor as often as possible. The Hypervisor has a mechanism to optimize this behavior, which is calle*d processor affinity*.

Micro-Partitioned processors are seen by the LPAR operating system by using an abstraction layer: virtual processors (VP). VPs are a representation in the operating systems for their physical core counterparts. A VP can represent up to 10 fractions of 1/10th of a processor. The Hypervisor always attempts to dispatch the cycles of a VP to the same underlying physical processor.

The Hypervisor uses this algorithm to provide processor affinity in this order:

1. The same processor core
2. Another core on the same chip
3. Another core on the same chip module
4. Another core on the same processor card/book

In addition to the Hypervisor processor affinity, the AIX operating system can also fine-tune the level of affinity of a running process with a certain VP. For further explanation, see “Dynamic optimization by Active System Optimizer” on page 43 and “Memory affinity API enhancements” on page 43.

Memory virtualization and sharing

The concept of memory virtualization is simple. The main memory of the system needs to be divided among the logical partitions that are on the Power frame. Remember, memory is one of the three components that are required for a system to exist: processor, memory, and I/O.

The Hypervisor controls which portions of the main memory are assigned to each of the logical partitions. A logical memory map is maintained within the Hypervisor, which is also known as the Hypervisor Page Table (HPT), for each of the running LPARs on the frame. Figure 1-15 shows memory virtualization in Power Systems.

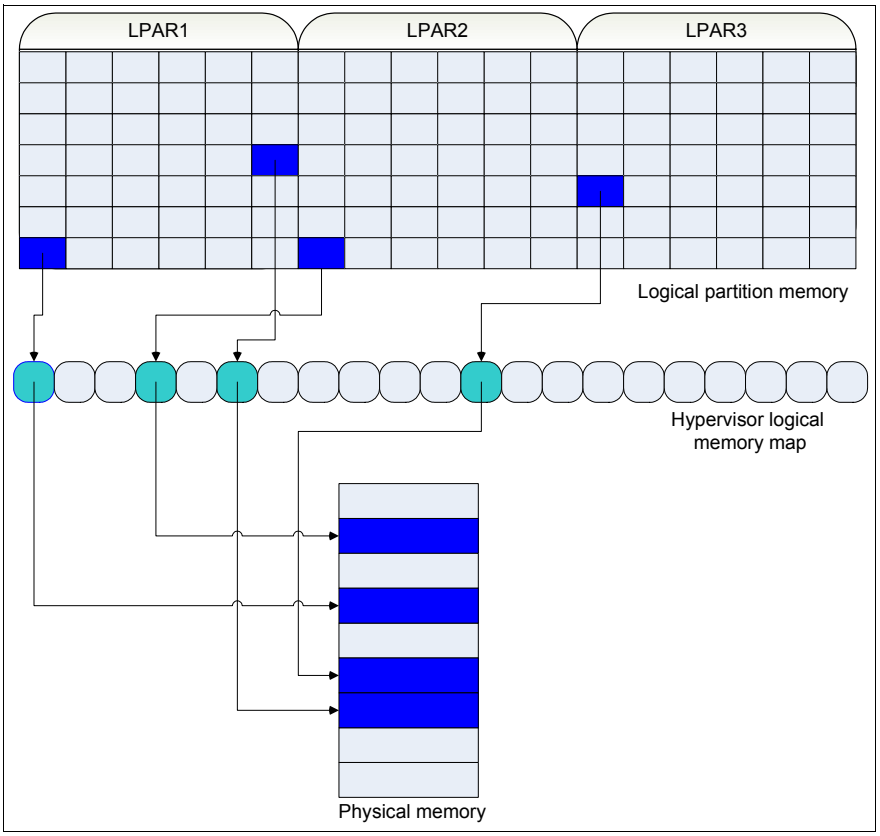

Figure 1-15 Memory virtualization: Hypervisor logical memory map

Active Memory Sharing and Active Memory Deduplication

Memory can also be shared among multiple LPARs. This feature is called *Active Memory™ Sharing* (AMS). There is a pool of physical memory from which the LPARs get their memory. This virtualization technique saves physical memory if you have LPARs with memory consumption behaviors that complement one another during peak memory consumption windows. Also, this virtualization technique saves physical memory if you know that only a few of the LPARs are expected to have memory consumption peaks at a certain time.

Figure 1-16 on page 37 depicts the use of AMS. LPARs 3 - 6 use shared memory and get their memory from the shared memory pool. Together, they are configured to request 10+4+6+8 = 28 GB of memory from the Hypervisor. However, the shared pool has only 24 GB available.

There are three situations that can happen in this scenario:

- No memory overcommitment: The state of the requested memory for all the active shared memory LPARs does not exceed 24 GB (when some of them are not activated, for example).
- Logical memory overcommitment: The state in which the true memory consumption of the shared memory LPARs is under 24 GB.
- Physical memory overcommitment: The state in which the true memory consumption of the shared memory LPARs is over 24 GB. If all the LPARs consume all of its memory, they are consuming 28 GB. In this case, 4 GB of memory is paged out to disk. Well-planned AMS environments do not get to this state.

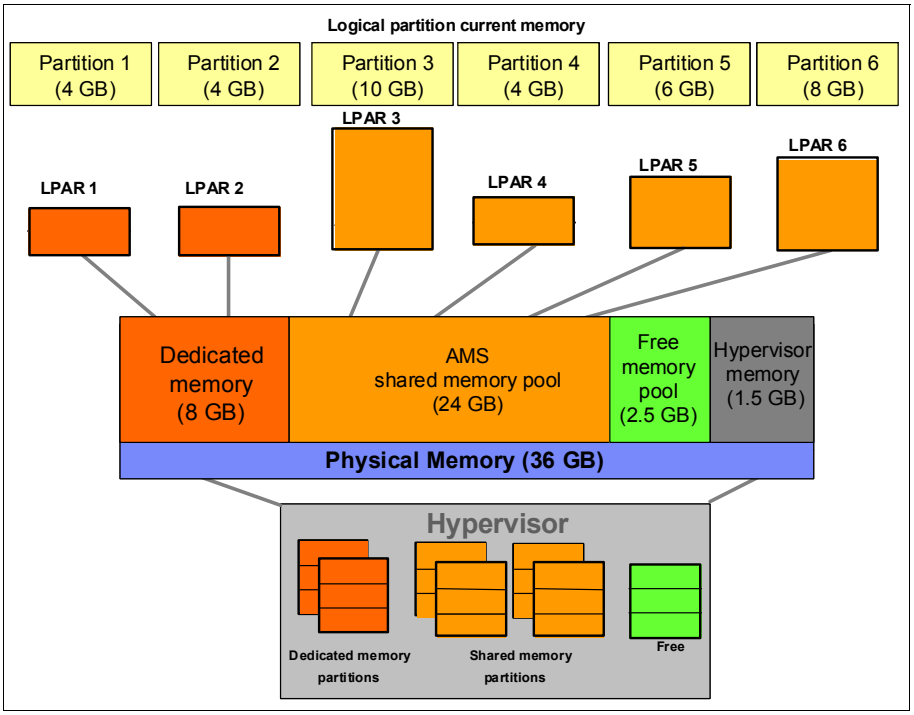

Figure 1-16 Active Memory Sharing

For more information about AMS, see *IBM PowerVM Virtualization Active Memory Sharing*, REDP-4470.

Active Memory Deduplication

Another technology that is associated with shared memory is Active Memory Deduplication (AMD). It eliminates duplicated memory pages by creating another level of memory indirection in memory map tables. It enhances the performance of memory sharing.

Figure 1-17 on page 38 depicts how the memory map is handled with deduplication active. The logical memory pages with equal content all point to the same physical memory page, thus reducing overall memory footprint usage. Compare this scenario with the scenario where Active Memory Deduplication is not in use, as shown in Figure 1-15 on page 36.

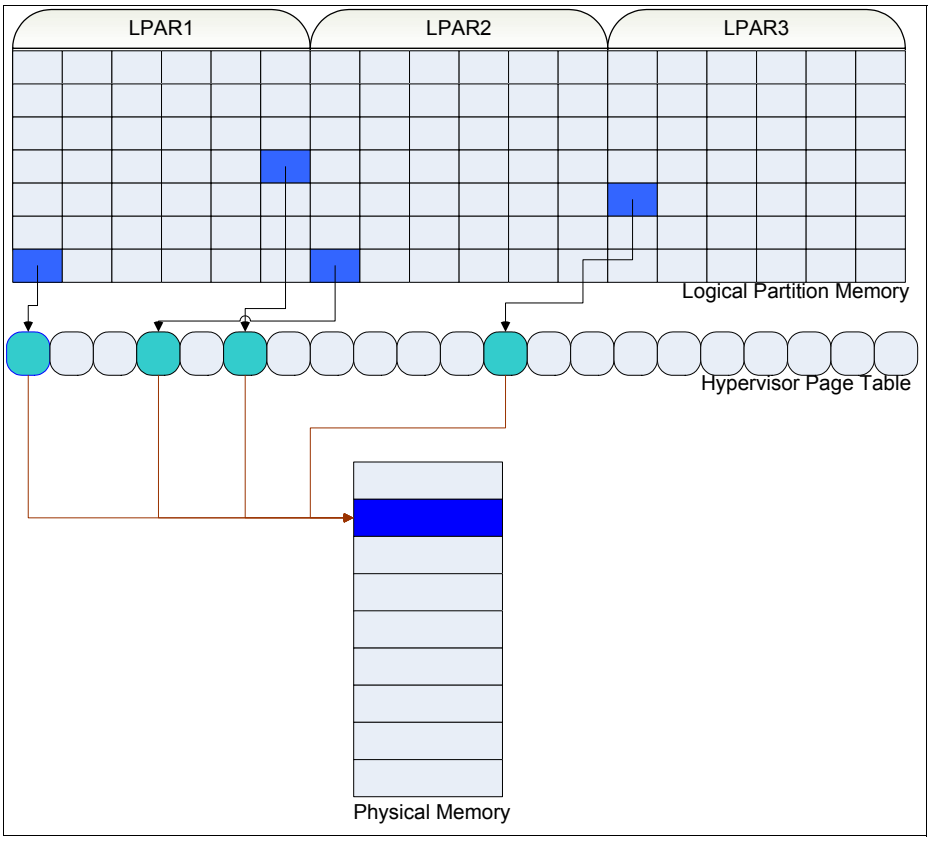

Figure 1-17 Physical memory savings with the use of memory deduplication

Because of the reduction in memory footprint overall usage, AMD helps shared memory systems avoid getting to the state of physical memory overcommitment. So, it is a good practice to turn on deduplication in a shared memory environment.

For more information about AMD, see *Power Systems Memory Deduplication*, REDP-4827.

Active Memory Expansion

Another feature of PowerVM related to memory is Active Memory Expansion (AME). This technology allows your LPAR to dynamically compress and decompress its memory. AME saves memory space at the expense of using processing power to perform these operations.

AME compresses the memory areas that are not recently used to free up memory space. This operation starts to happen after a threshold in free memory is reached. In this way, you can either use these savings to reduce the logical partition memory footprint, or to handle more memory than your LPAR conventionally is able to handle with the same amount of physical memory.

AME works is by selecting a compression ratio. The ratio to use is driven by a prior analysis of data compressibility and the amount of processor power that you want to exchange to compress and decompress data. This prior analysis is performed by the **`amepat`** tool in AIX.

Figure 1-18 on page 39 demonstrates the use of AME in an LPAR.

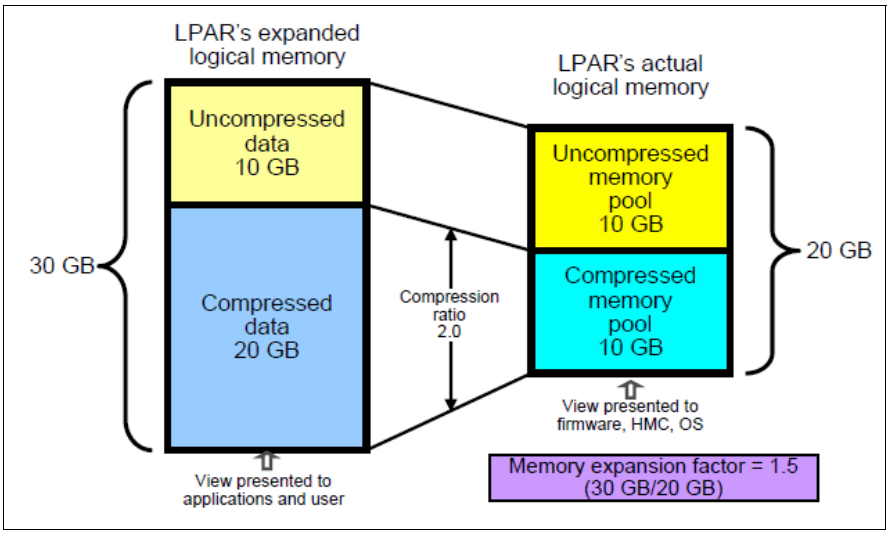

Figure 1-18 Active Memory Expansion

Virtual I/O Server

After processor and memory, I/O is the third essential operation of a system to perform useful work. In the virtual world of Power Systems, you use the Virtual I/O Server (VIOS).

The VIOS is an auxiliary logical partition on your Power frame. It is an operating system, but it specializes in sharing its I/O devices with other logical partitions. The VIOS owns the physical I/O resources: network cards, disks, or Fibre Channel (FC) adapters. Whenever an LPAR with virtual I/O adapters needs to perform an I/O operation, it sends the data to its virtual adapters. The Hypervisor then copies this data from the requesting LPAR virtual adapters to the corresponding virtual adapters on the VIOS. The VIOS maps which adapter it uses to communicate with each LPAR in terms of I/O. After that, the VIOS performs the I/O operation by using its own physical I/O adapters. This part is the data flow of an "output" I/O operation. An "input" I/O operation happens in the same way, but with the opposite direction of data flow.

The VIOS is able to virtualize and share the following types of I/O cards:

- Networking cards
- Internal and external storage area network (SAN) disks
- FC cards

Networking virtualization

One advantage of consolidating a stack of workloads on the same Power Systems frame by the use of virtualization is that these workloads can take advantage of a high-speed virtual network. Figure 1-19 on page 40 depicts how TCP/IP communication happens in this environment, which is the most basic way of configuring it. Production systems use a more sophisticated setup to ensure high availability with two VIOS and network traffic throughput with the use of link aggregations (IEEE 802.3ad).

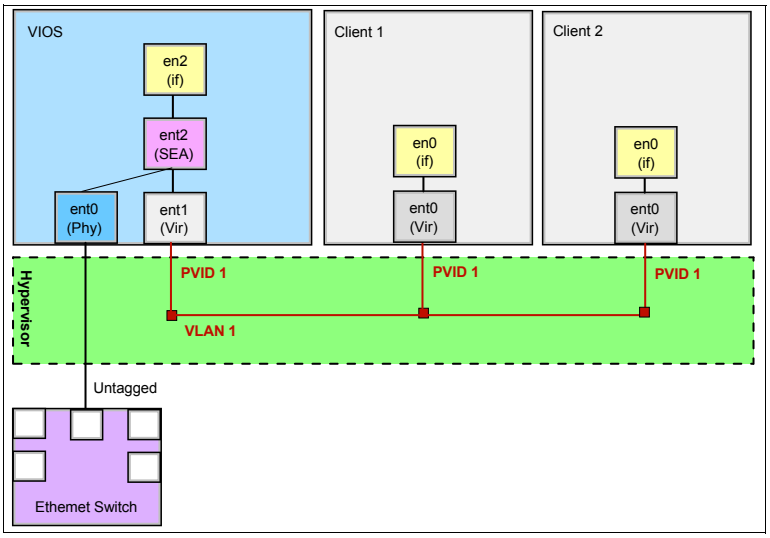

Figure 1-19 Virtual networking on a Power Systems machine

Whenever the origin and destination LPARs are on the same physical hardware, there is no need to send a data packet out to the physical network. The physical network is limited by the speed of the available network card (1 Gbps or 10 Gbps). During a network transfer, the packet is put into the send queue of the sender on the virtual network card, which is really a buffer in memory. Because the destination LPAR is on the same frame, you need to copy this packet to the recipient receive queue on the other side, which is another buffer area in memory because it is also a virtual adapter. Because both LPARs reside on the same frame, this memory-to-memory copy is handled by the Hypervisor. It happens at memory transfer speeds, which are faster than a transfer through actual network cards, cables, and switches.

If the recipient is not on the same frame, the packet is routed through the VIOS to the outside physical network, through a physical adapter on the VIOS, as shown in Figure 1-19. This routing occurs through the setup of a Shared Ethernet Adapter (SEA) on the VIOS.

By consolidating your analytics workloads on a Power frame, you can take advantage of the fast in-memory transfers when one logical partition needs to send data to another one over TCP/IP.

Disks virtualization (vSCSI)

Disks can be virtualized by the use of a VIOS, as depicted in Figure 1-20 on page 41. This schema is also simplified. Production systems rely on higher availability levels by using dual VIOS to serve SAN disks. Or, they create internal mirrors from two disks that are served by different VIOS when serving internal disks.

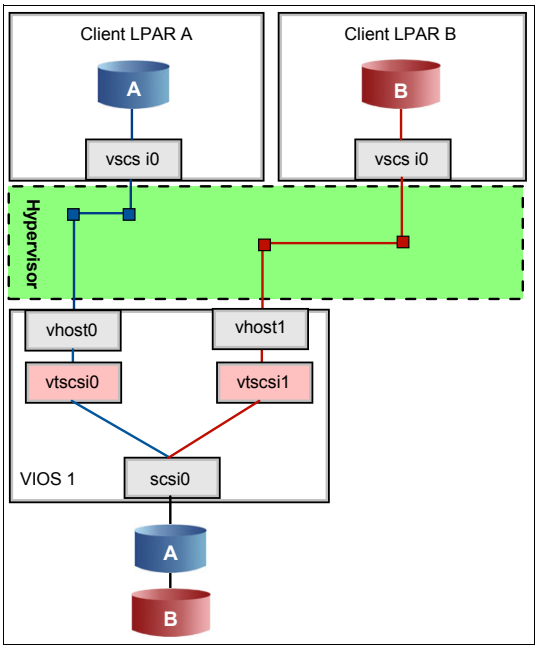

Figure 1-20 Disk virtualization with vSCSI

The target physical disk, either internal or external SAN, is owned by the VIOS. By using virtual Small Computer System Interface (vSCSI) adapter pairs, one on the VIOS and the other on the client LPAR, this disk can be virtualized and presented to the client LPAR.

This process applies by using logical volumes (LV) on the VIOS as the “source disk” as opposed to using an entire physical disk. The client sees the disk in the same way in either case as a vSCSI disk. A third option, called *file-backed storage*, consists of creating a storage pool of disks in the VIOS. In file-backed storage, a file system is configured and the files that are created in it are used as the backing devices for the virtual disks. File-backed storage is similar to LV-backed devices, but it uses file system files as opposed to a logical volume.

VIOS shared storage pools

Shared storage pools is a technology that is based on vSCSI (available since the end of 2010) and that consists of creating a storage pool on SAN disks. The client LPAR sees the disk as an ordinary vSCSI disk. The pool can be available to multiple VIOS servers, even to VIOSs on different machines, and the space is allocated as it is required. Due to this latter characteristic, shared storage pools are considered a means of thin provisioning of virtual storage at the VIOS level.

Fibre Channel virtualization (NPIV)

You can virtualize FC cards by using a VIOS, which is known as N-Port ID Virtualization (NPIV). Figure 1-21 on page 42 depicts NPIV. It is a simplification of an actual implementation that relies on a dual VIOS implementation to ensure high availability.

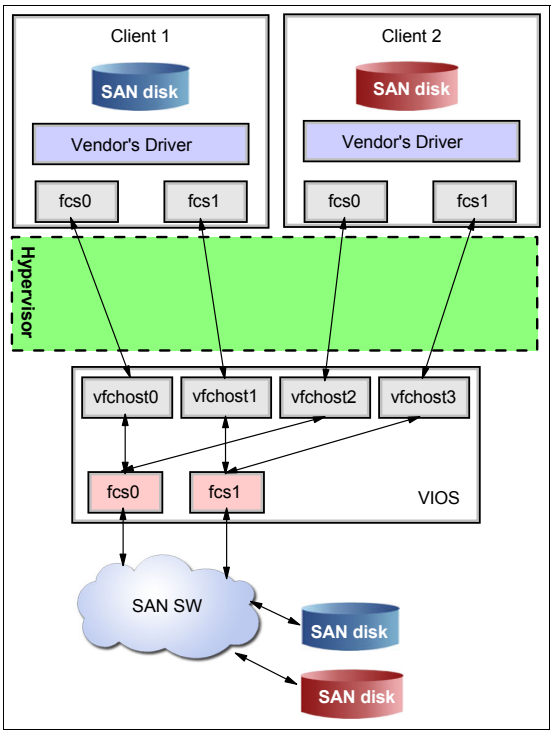

Figure 1-21 FC virtualization with NPIV

The VIOS is the owner of the FC adapter. It virtualizes and shares it with the LPARs by using a pair of virtual adapters, one on the VIOS and the other on the client LPAR.

The concept of virtualizing and sharing FC adapters is called N_port ID Virtualization (NPIV) and is a more recent technology than vSCSI. There are advantages in the use of NPIV as opposed to the use of vSCSI:

- Performance during an I/O because data transfers bypass the VIOS I/O buffers.
- SAN disks are zoned directly to the client LPAR as opposed to being zoned to the VIOS for further vSCSI virtualization. This design results in a more organized zoning scheme and maintenance.
- Client LPARs see a real FC adapter and can apply the vendor host bus adapter (HBA) drivers to obtain load balancing and other features.

1.4.3 Overview of AIX 7.1 features

The POWER7 platform provides a new set of capabilities and facilities that are aimed at improving the performance of various applications and workloads. The base enablement of the processor, as well as the new features, drives changes in AIX to recognize, start, and enable the functionalities provided by the processor.

AIX Version 7.1 introduces many new features to use the capabilities and facilities provided by POWER7 and its virtualization technologies. The following section briefly describes some of the important features provided by the AIX 7.1 operating system that enhance application performance.

Dynamic optimization by Active System Optimizer

Active System Optimizer (ASO) is an AIX 7 and POWER7 performance tuning tool that automatically tunes your logical partition to optimally use POWER7 cache affinity and memory affinity. ASO is similar to a performance tuning agent that tunes your machine 24 hours a day. It is simple to set up and use.

POWER7 Systems feature increasing numbers of available processor cores, with memory closely associated with each group of cores. The various potential groupings of cores across different chips and nodes (“books”) leads to a non-uniform memory access (NUMA) topology in which access to memory has a non-uniform cost for a specific core based on its location in the system.

AIX 7.1 permits different policies for associating memory or processes with specific affinity domains to reduce the need for a process to access memory that is not local to the core on which it is running. However, depending on the type of workload, the default policies used might not result in optimal placement all the time. The optimal placement is difficult for a user or system administrator to identify without an extensive analysis.

The ASO addresses these issues by automatically making dynamic workload placement decisions based on heuristics and detailed analysis data provided by AIX and the hardware Performance Monitoring Unit (PMU). At the time of writing this book, ASO optimization focuses on improving cache and memory affinity primarily on the placement of workloads, but also through the redistribution of memory when required.

The ASO is expected to improve the performance of all workloads, although it is able to recognize scenarios where it offers limited benefit. When this situation occurs, the ASO temporarily hibernates until workload characteristics change to be more prone to optimization.

The ASO is supported on POWER7 Systems only where Enhanced Affinity support is provided. Support for older hardware, or POWER7 running in P6 compatibility mode, is not provided, and in these configurations, the ASO automatically shuts itself down.

1024 hardware thread enablement

AIX 7.1 provides support to up to 1,024 logical processors in a running logical partition. Processors can be configured in dedicated or shared processor modes. The earlier limit on the number of supported logical processors was 256 on AIX 6.1 Technology Level (TL) 4 on POWER7 Systems.

Memory affinity API enhancements

AIX 7.1 allows an application to request a *strict* attachment from a thread to a Scheduler Resource Allocation Domain (SRAD) for memory affinity. This new form of attachment is similar to the current SRAD attachment APIs, except that the thread is not moved to a different SRAD for load balancing by the dispatcher.

The following comparison shows the differences between the new strict attachment API and the existing advisory attachment API:

- When a thread has an *advisory* SRAD attachment, the AIX thread dispatcher is free to ignore the attachment if the distribution of load across various SRADs justifies the migration of the thread to another SRAD. The new *strict* attachment overrides any load balancing efforts of the dispatcher.

- The current *advisory* SRAD attachment APIs allow SRAD attachments to the R_PROCESS, R_THREAD, R_SHM, R_FILDES, and R_PROCMEM resource types. The new *strict* SRAD attachment allows SRAD attachment to the R_THREAD resource type only. Any other use of strict SRAD attachment results in an EINVAL error code.
- The pthread_attr_setsrad_np API is modified to accept a new flag parameter that indicates whether the SRAD attachment is in strict or advisory mode.

Improved performance by using 1 TB segments

In AIX V7.1, 1 TB segments are an autonomic operating system feature designed to improve performance of 64-bit large memory applications. This enhancement optimizes performance when using shared memory regions (shmat/mmap). New restricted **vmo** options are available to change this operating system policy. A new VMM_CNTRL environment variable is available to alter per process behavior.

One TB segment aliasing improves performance by using 1 TB segment translations on Shared Memory Regions with 256 MB segment size. This support is provided on all 64-bit applications that use Shared Memory Regions. Both directed and undirected shared memory attachments are eligible for 1 TB segment aliasing.

If an application qualifies to have its Shared Memory Regions use 1 TB aliases, the AIX operating system uses 1 TB segment translations without changing the application. This capability requires the use of the shm_1tb_shared **vmo** tunable, the shm_1tb_unshared **vmo** tunable, and the esid_allocator **vmo** tunable.

LVM enhanced support for solid-state disks

Solid-state disks (SSDs) are an option for enterprise storage requirements. SSDs are unique in that they do not have any moving parts. They perform at electronic speeds without the mechanical delays (moving heads or spinning platters) associated with traditional spinning hard disk drives (HDDs). Compared to traditional HDDs, the characteristics of SSDs enable a higher level of I/O performance in terms of greater throughput and lower response times for random I/O. These devices are ideal for applications that require high IOPS/GB, low response times, or both.

AIX V7.1 includes enhanced support in the AIX Logical Volume Manager (LVM) for SSD. This support includes the capability for LVM to restrict a volume group (VG) to contain SSDs only, and the ability to also report that a VG contains only SSDs. This feature is also available in AIX V6.1 with the 6100-06 TL.

Traditionally, a VG can consist of physical volumes (PVs) from a mix of storage devices, such as HDDs. There was no method to restrict the disks of a VG to a specific type of storage device. The LVM is enhanced to allow the creation of a VG to a specific storage type, in this case, SSDs. The ability to restrict a VG to a particular type of disk can help to enforce performance goals for the VG.

For example, a DB2 database can be on a set of SSDs for best performance. Without restricting the storage device type, slower devices can be included on the VG. Then, reads and writes on that VG perform only as fast as the slowest disk. Therefore, it is best to restrict this VG to SSDs only. To maximize performance, restrict mixing SSDs and HDDs in the same VG.

Hot files detection in JFS2

SSDs offer a number of advantages over traditional HDDs. With no seek time or rotational delays, SSDs can deliver substantially better I/O performance than HDDs. To maximize the benefit of SSDs, it is important to place data on them that requires high throughput and low response times only. This data is referred to as *hot* data or *hot* files. Typically, a *hot* file can be described as a file that is read from or written to frequently. It can also be a file that is read from or written to in large chunks of data.

Before deciding to move suspected hot files to faster storage (for example, to SSDs), users of a file system need to determine which files are hot. The files must be monitored for a while to identify the best candidates.

AIX 7.1 includes enhanced support in the Enhanced Journaled File System (JFS2) for SSDs. JFS2 is enhanced with the capability to capture and report per-file statistics related to the detection of hot files. These statistics can be used to determine whether a file needs to be placed on an SSD. These capabilities enable applications to monitor and determine optimal file placement. This feature is also available in AIX V6.1 with the 6100-06 TL.

Processor interrupt disablement

AIX 6.1 TL6 and AIX 7.1 provide a facility to quiesce external I/O interrupts on a specific set of logical processors. This facility helps to reduce interrupt jitter that affects application performance.

When co-scheduling Parallel Operation Environment (POE) jobs, or even in a non-POE commercial environment, administrators can control the process scheduling and interrupt handling across all the processors. It is important to quiesce interrupts on the simultaneous multithreading (SMT) threads that are running POE jobs to avoid interrupting the jobs. By quiescing interrupts, your applications can run on a specific set of processors without being affected by any external interrupts.

The processor interrupt disablement function can be configured by using the following kernel service, system call, or user command:

- Kernel service: **`k_cpuextintr_ctl()`**
- System call: **`cpuextintr_ctl()`**
- Command line: **`cpuextintr_ctl()`**

This functionality is supported on POWER5, POWER6, and POWER7 and any future POWER processor-based hardware. It is supported on both dedicated or shared processor LPARs.

Cluster Aware AIX

The Cluster Aware AIX (CAA) services help in creating and managing a cluster of AIX nodes to build a highly available and ideal architectural solution for a data center. IBM cluster products, such as Reliable Scalable Cluster Technology (RSCT) and PowerHA®, use these services. CAA services can help in the management and monitoring of an arbitrary set of nodes or in running third-party cluster software.

Details about each of these services, together with examples that use commands to configure and manage the cluster, are described. CAA services are a set of commands and services that the cluster software can use to provide high availability and disaster recovery support to external applications.

The CAA services are broadly classified into the following categories:

- Cluster-wide event management

  The AIX Event Infrastructure allows event propagation across the cluster so that applications can monitor events from any node in the cluster.

- Cluster-wide storage naming service

  When a cluster is defined or modified, the AIX interfaces automatically create a consistent shared device view across the cluster. A global device name, such as *cldisk1*, refers to the same physical disk from any node in the cluster.

- Cluster-wide command distribution

  The **clcmd** command provides a facility to distribute a command to a set of nodes that are members of a cluster. For example, the command **clcmd date** returns the output of the **date** command from each of the nodes in the cluster.

- Cluster-wide communication

  Communication between nodes within the cluster is achieved by using multicasting over the IP-based network and also by using storage interface communication through FC and serial-attached SCSI (SAS) adapters. A new socket family (AF_CLUST) is provided for reliable, in-order communication between nodes. When all network interfaces are lost, applications that use these interfaces can still run.

The nodes that are part of the cluster must have common storage devices, either through the SAN or SAS subsystems.

Lightweight Directory Access Protocol enhancements

Multiple enhancements were made to AIX 7.1 for Lightweight Directory Access Protocol (LDAP) use. The next sections outline them in detail.

AIX LDAP authentication enhancements

AIX LDAP authentication is enhanced with the following new features:

- chpasswd support for LDAP:

  The **chpasswd** command manages user passwords. The root user can supply or change user passwords specified through standard input. The **chpasswd** command is enhanced to set LDAP user passwords in an ldap_auth environment by specifying **-R LDAP** and by not specifying the **-e** flag for encrypted format. If you specify the **-e** option for the encrypted format, the **chpasswd** encrypted format and LDAP server-crypted format must match.

- Case-sensitive LDAP user names:

  The LDAP UID and CN attributes are used to store user account name and group account name. Both the UID and the CN attributes are defined as directory strings and were not case sensitive. Starting with AIX 6.1 TL06 and AIX 7.1, both the UID and CN can be case-sensitive by enabling the **caseExactAccountName** configuration parameter in the `/etc/security/ldap/ldap.cfg` file.

LDAP alias support

This feature allows AIX users to log in with an alias name defined in the LDAP directory entry.

LDAP caching enhancement

The AIX LDAP *secldapclntd* client daemon caches user and group entries retrieved from the LDAP server. AIX 6.1 TL06 and AIX 7.1 offer the ability to control the caching mechanism through a new attribute called TO_BE_CACHED. This change translates into having an additional column in the existing mapping files in the `/etc/security/ldap` directory. All attributes in the LDAP mapping files have a value of `yes` in the TO_BE_CACHED new field, by default. Administrators can selectively set an attribute to `no` to disable the caching of that attribute.

The following additional LDAP enhancements are available:

- AIX LDAP supports Windows 2008 Active Directory (AD) and Active Directory application mode (ADAM).
- The **`ls1dap`** command lists the following information:
  - Users
  - Groups
  - Network Information Service (NIS) entities (hosts, networks, protocols, services, rpc, and netgroup)
  - Automount maps
  - Role-based access control (RBAC) entries (authorizations, roles, privileged commands, and devices)

  This command is extended to cover advance accounting.

The AIX LDAP module is a full-function module that covers both authentication and identification. It cannot be used as an authentication-only module, which some clients want. This functionality is enhanced to have the same module support as a full-function module or an authentication-only module.

AIX V7.1 minimum system requirements

Information about the minimum system requirements to install and run AIX 7.1 is provided.

Required hardware

Only 64-bit Common Hardware Reference Platform (CHRP) machines are supported by AIX 7.1. The following processors are supported:

- PowerPC® 970
- POWER4
- POWER5
- POWER6
- POWER7

Minimum firmware levels

Update your systems to the latest firmware level before migrating to AIX 7.1. For the AIX V7.1 release notes for information about the minimum system firmware levels required for AIX V7.1, see this website:

`http://publib.boulder.ibm.com/infocenter/aix/v7r1/index.jsp?topic=/com.ibm.aix.ntl/releasenotes_kickoff.htm`

For the latest Power system firmware updates, see the following website:

`http://www14.software.ibm.com/webapp/set2/firmware/gjsn`

Memory requirements

The minimum memory requirement for AIX 7.1 is 512 MB.

The current minimum memory requirements for AIX 7.1 vary based on the configuration of a system. It might be possible to configure a smaller amount of memory for a system with a few devices or small maximum memory configuration.

The minimum memory requirement for AIX 7.1 can increase as the maximum memory configuration or the number of devices scales upward.

Paging space requirements

For all *new* and *complete overwrite* installations, AIX 7.1 creates a 512 MB paging space device named /dev/hd6. As a rule, your system paging space must be at least half of your maximum memory configuration.

Disk requirements

A minimum of 5 GB of physical disk space is required for a default installation of AIX 7.1. This space includes all devices, the graphics bundle, and the system management client bundle. Table 1-5 provides information about disk space usage with a default installation of AIX 7.1.

Table 1-5 File system space requirements for AIX 7.1

File system	Allocated (used)
/	196 MB (181 MB)
/usr	1936 MB (1751 MB)
/var	380 MB (264 MB)
/tmp	128 MB (2 MB)
/admin	128 MB (1 MB)
/opt	384 MB (176 MB)
/var/adm/ras/live/dump	256 MB (1 MB)

2

Implementing Cognos on IBM Power Systems

The details of the IBM Cognos Business Intelligence (BI) installation and configuration with other software stacks are described. A database is the integral component for any BI software to fetch required data. Security is another important aspect to maintain data security, integrity, and unauthorized access. Thus, BI infrastructure needs a secured environment.

We describe several data sources, such as relational database (DB2), PowerCube, and TM1 Cubes. We describe the security that is implemented through Tivoli® Lightweight Directory Access Protocol (LDAP) server. We also cover various client tools and web interfaces to access Cognos reports.

The following topics are described:

- Cognos vertical and horizontal scalability
- Cognos single dispatcher deployment
- Cognos multiple dispatcher deployment

2.1 Cognos vertical and horizontal scalability

Cognos requires several components to perform all the necessary tasks for an analytical reporting system. These components are spread across three tiers.

BI uses the following components:

- Cognos Gateway (web tier)
- Dispatcher, including report server, batch report server, and so on (application tier)
- Application Server (application tier)
- Cognos Content Manager (application tier)
- Cognos Content Store (data tier)
- Cognos Query DB (data tier)
- Cognos Audit DB (data tier)
- Data Warehouse (source)
- LDAP Server (data tier)

Several of these components can be scaled both vertically or horizontally to increase performance. *Vertical scaling* adds more resource to the server or logical partition (LPAR) on which the component runs. Horizontal scaling adds another server or LPAR with the same installed component that performs the same function with the workload balanced across all servers that run the same component.

Vertical scaling is performed first and is more suited to lower-end implementations. At a certain point, horizontal scaling is a more effective way to increase performance. It also depends on the reasons for scaling, such as scaling for more users or scaling to run the reports more quickly. Licensing and cost also influence when you decide whether to scale vertically or horizontally. Some of the components scale better vertically and some components scale better horizontally.

The Cognos server is the component that scales best horizontally. When you create multiple Cognos servers, there are different ways that you can balance the workload across them. One method is to split the work tasks that each Cognos server processes and another method is to round-robin the requests to each server. Which method works best depends on your workload pattern. We created two scenarios to test and optimize. One scenario has a single dispatcher where we can test vertical scalability, and the second scenario has multiple dispatchers where we can test horizontal scalability. The following sections describe the two scenarios in more detail.

2.1.1 Vertical scalability

Cognos installation starts with the basic configuration to accommodate a limited workload. It starts with one dispatcher with all services configured. As the workload increases, the Cognos server and associated software stack are tuned to achieve the following objectives:

- Higher workload
- Better throughput
- Faster turnaround time

Vertical scalability identifies the best system profiling information, which consists of the following factors:

- Processing power: As the workload increases, the processor allocation to that LPAR must be increased. IBM Cognos throughput and response time must be mapped with the increased processing ability.
- Memory allocation: Memory usage changes as workload changes on the IBM Cognos server. Memory allocation must be changed as the workload increases.
- Storage and query optimization: Database activity increases with the increase of the workload. The Content Store, audit database, and query database are some of the most active databases for a busy Cognos environment. Performance tuning for faster storage I/O retrieves more data in a shorter time. Also, query optimization ensures faster data retrieval.
- IBM Cognos concurrent processes and threads: IBM Cognos servers must be tuned to create multiple concurrent processes with concurrent threads. Multiple users and reports execute through this setting.

Vertical scalability provides the optimal performance from the IBM Cognos environment to the users. Later, to address a higher workload and better performance, you can replicate a similar stack and configuration over several LPARs.

2.1.2 Horizontal scalability

Through vertical scalability, the Cognos server achieves high watermark performance. However, to accommodate better performance, horizontal scalability is needed. Multiple software stacks are the primary feature in horizontal scalability. Horizontal scalability achieves the following objectives:

- Performance scalability beyond a single software stack: Multiple instances of the server are required to improve the Cognos environment performance beyond single server throughput. Multiple server instances provide more BI Buses, processing power, and memory, and redundancy, failover, and workload distribution.
- Redundancy: The configuration of redundant services offers uninterrupted service to clients, even during a partial environment outage. For example, if the Cognos environment contains only one content manager and that server stops responding, the entire Cognos service stops serving. However, if multiple content managers are configured, if the primary content manager is unavailable, the secondary content manager takes over as the primary and the client experiences uninterrupted service.
- Redundant report server: Multiple report servers execute multiple concurrent reports. Performance improves.
- Load balancing: Depending on the workload, a report request is routed through a server with a lower workload. Response time is faster.

2.1.3 Cognos transaction flow

We explore the transaction flow through the Cognos environment, which is primarily driven by HTTP requests. The IBM Cognos environment consists of three components: the gateway module, the authentication and authorization module, and the report execution model, as shown in Figure 2-1 on page 52.

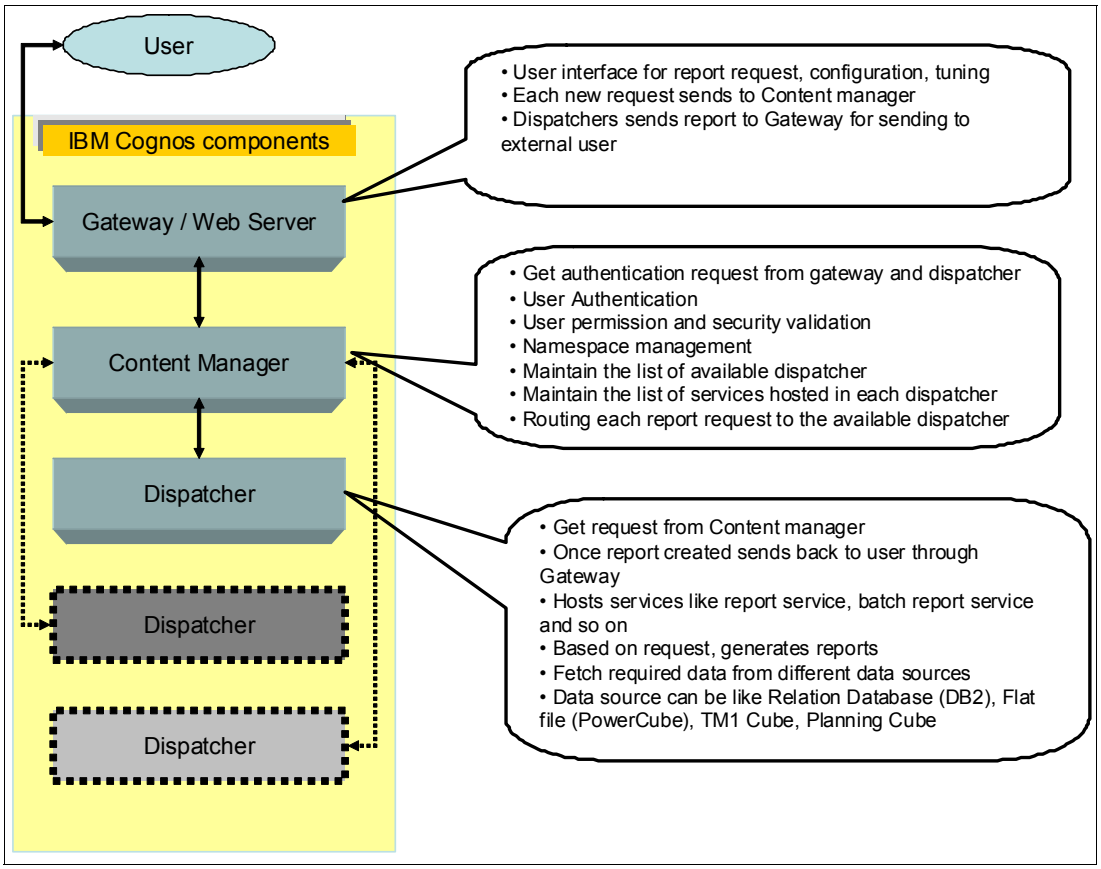

Figure 2-1 Cognos transaction and workflow diagram

Cognos gateway

The Cognos gateway is the primary user interface and is accessed through the Cognos connection and portal URL. User authentication and namespace validation happen at this layer. However, internally, the gateway communicates with the content manager for authentication and validation. After successful authentication, a request is sent to the content manager for further processing. After the report execution completes, it reaches the gateway and is sent to the client.

Content manager

In an IBM Cognos environment, the main controlling unit is the content manager. It performs several important functions. Every Cognos environment consists of one primary content manager only. However, a secondary content manager can be configured. The secondary content manager takes over as primary only when the primary content manager is unavailable. Even in a busy environment, all requests and transactions are handled by the single primary content manager. Some of the key functions of the content manager are explained next.

User validation

The content manager maintains the namespace information and users that are listed. For any new report request, the user details are validated through the content manager. The Cognos administrator assigns the required access and permission for users to access several reports and functionality.

Maintain dispatcher list and services

Dispatcher is the component that runs each report and sends the output to the user. Each dispatcher registers itself to the content manager during startup. Each content manager is configured to host several services. Some of the most common services are the report service, batch report service, query service, and annotation service. The content manager stores each dispatcher and associated service details in the Content Store. Figure 2-2 shows the registered dispatcher and services.

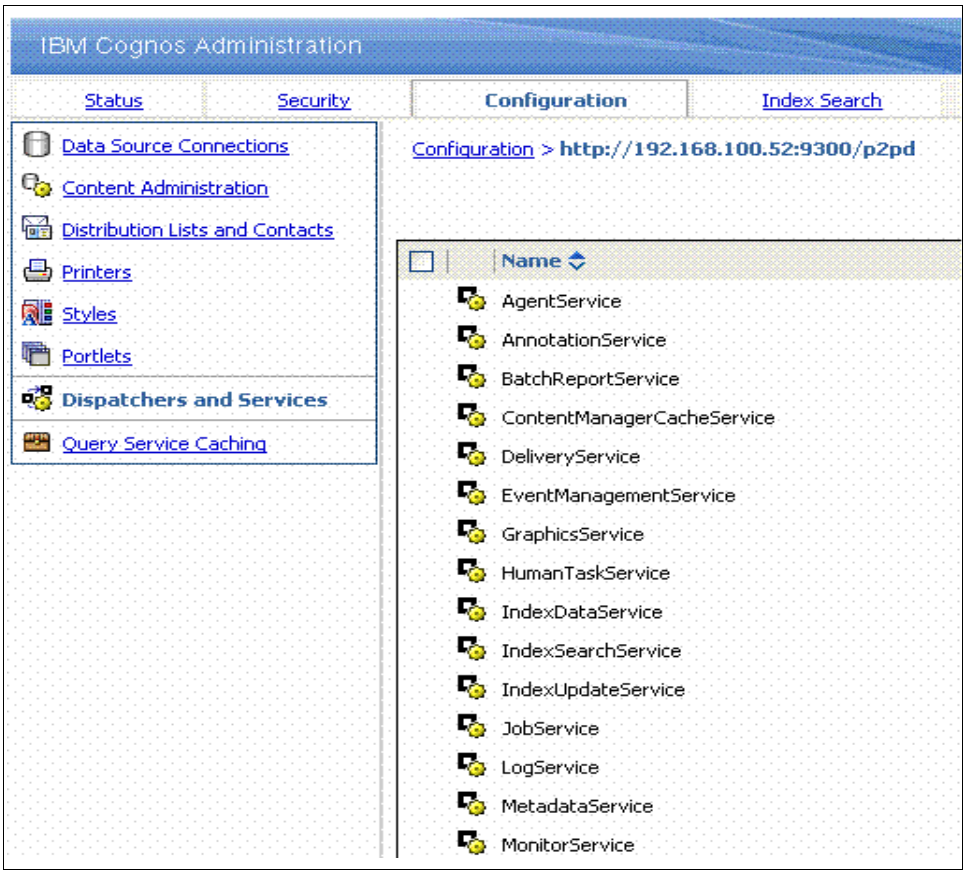

Figure 2-2 Registered dispatcher and service view from content manager

Request routing

The content manager routes a report request to the available dispatcher in a multiple dispatcher IBM Cognos deployment. For drill-through reports that are based on the affinity setting, the content manager routes multiple report requests to the same dispatcher for better performance. Figure 2-3 on page 54 shows the content manager view with multiple registered dispatchers.

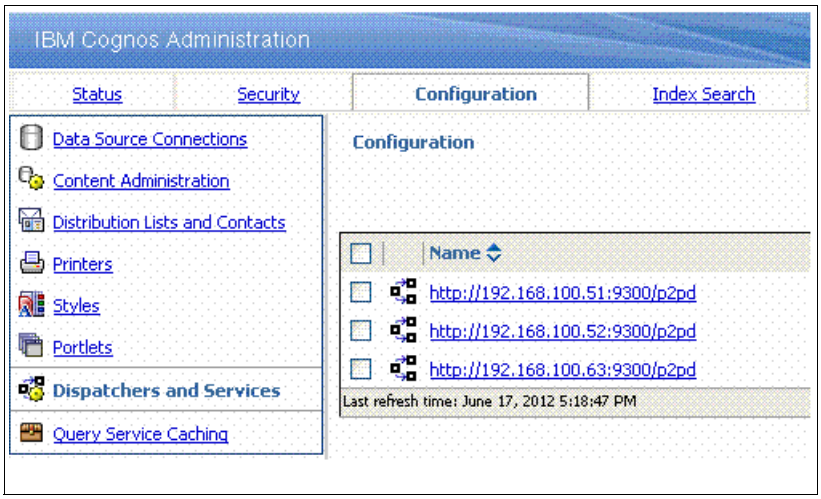

Figure 2-3 Content manager view of multiple registered dispatchers

Data source management

A *data source* is the logical interface between the Cognos environment and the data. Data can be pulled in from a relational database (for example, IBM DB2), the planning cube data in a flat PowerCube, or from the TM1 cube data. Data sources are managed from a web-based Administration console of the Cognos connection portal. Figure 2-4 shows the data source view from the Cognos Administration console.

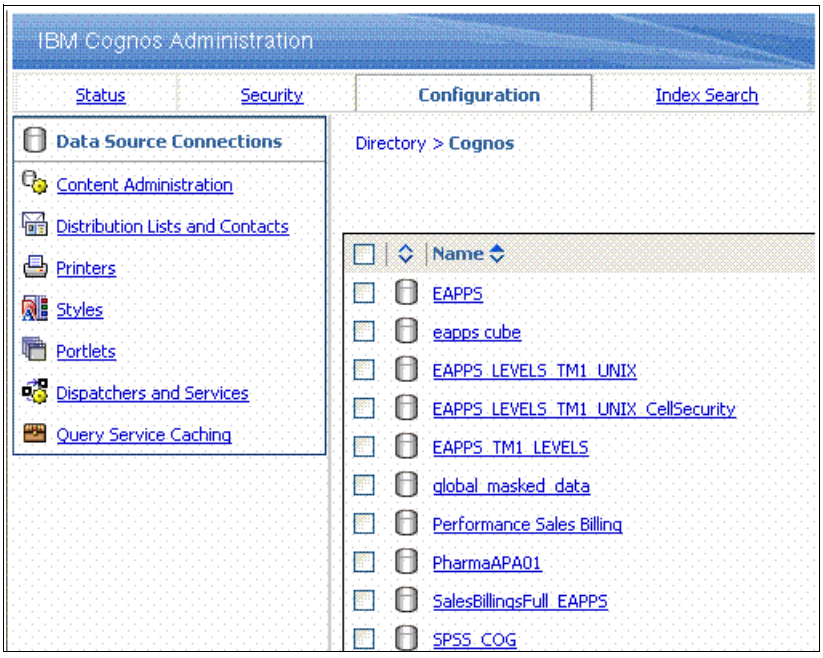

Figure 2-4 Data source view from the Cognos connection

Application tier and dispatcher

The dispatcher is the main working thread in the Cognos environment that is used to generate reports that are based on user requests. Each dispatcher hosts several services, as shown Figure 2-2 on page 53. The dispatcher fetches the required data from the different data sources, and it renders this data in the report based on the specification. Each dispatcher creates multiple BI Bus processes, and each process handles one request at a time. Each idle BI Bus process stops based on an idle timeout setting.

2.2 Cognos single dispatcher deployment

The gateway, application tier, and content manager are the three primary components in an IBM Cognos deployment. Based on your business requirements, these components can be installed and configured on the same or multiple servers. We describe different deployment landscapes.

The IBM Cognos installation needs more software stacks, such as the web server, application server, database, and LDAP server. The performance optimization of these software products is out of the scope of this IBM Redbooks publication. However, we describe the software stack installation and configuration in the following sections. Also, the same set of these servers can be used for single and multiple dispatcher IBM Cognos deployments.

Figure 2-5 on page 56 shows the single dispatcher layout for a typical Cognos system.

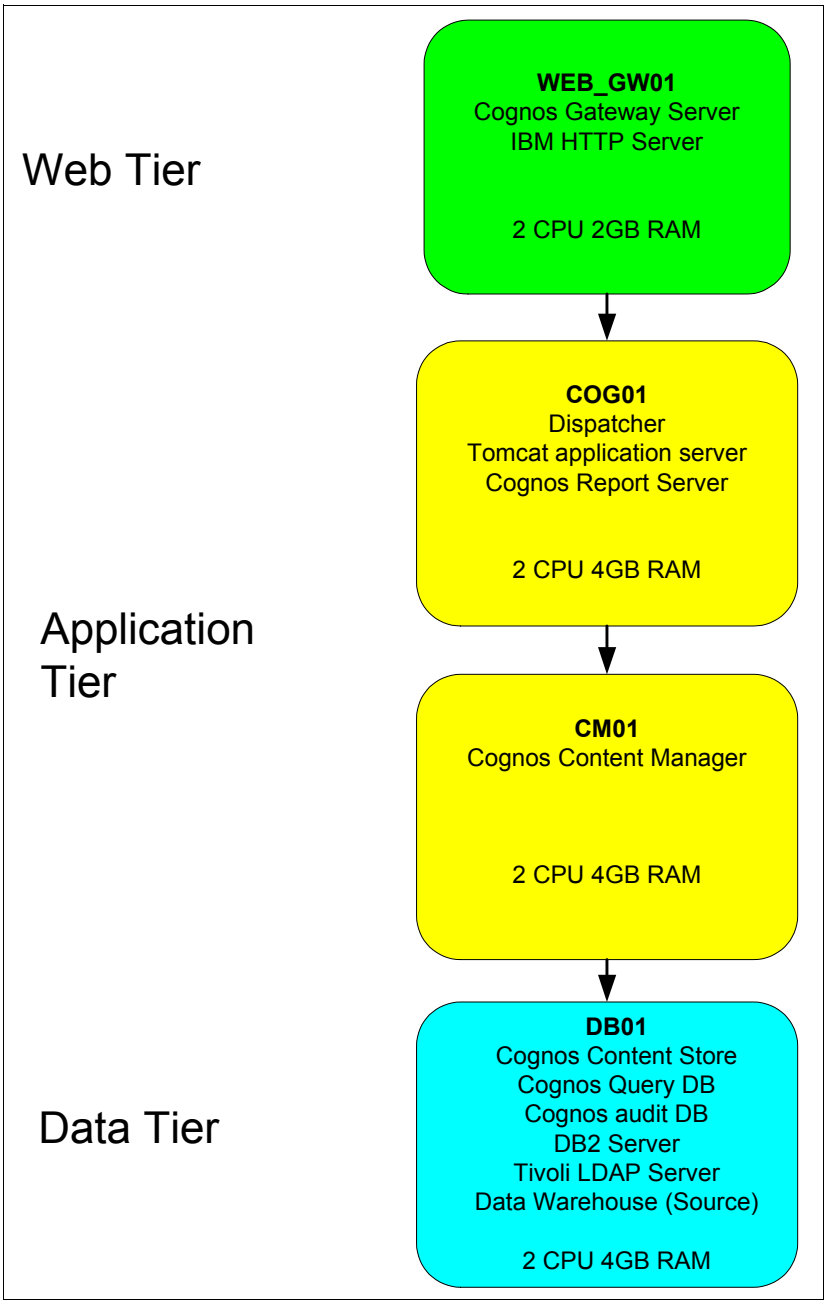

Figure 2-5 Single dispatcher Cognos layout

2.2.1 Primary landscape and associated software configuration

Information about the primary landscape and associated software configuration is described. See Figure 2-7 on page 58.

Content Store

In the Content Store, IBM Cognos puts all of the vital details, such as the configuration details, security information, authentication namespace, cryptographic keys, and metadata for published packages. Any IBM Cognos configuration starts with the Content Store, which can be created on any relational database. In this book, we used DB2 Enterprise Server as the database.

To achieve optimal performance, the Content Store database "CSO2" is created with the required user authentication, table spaces, and buffer pool settings. The configuration details are shown in Figure 2-6.

```
CREATE DATABASE CSO2 AUTOMATIC STORAGE YES ALIAS CSO2 USING CODESET
UTF-8 TERRITORY US COLLATE USING SYSTEM PAGESIZE 32768 WITH 'Cognos
CSO2'

CONNECT TO CSO2
CREATE SCHEMA "DB2COGNOS" AUTHORIZATION DB2INST1
CREATE BUFFERPOOL "COGNOS_04KBP"  SIZE 250 PAGESIZE 4096
CREATE BUFFERPOOL "COGNOS_08KBP"  SIZE 250 PAGESIZE 8192
CREATE BUFFERPOOL "COGNOS_16KBP"  SIZE 250 PAGESIZE 16384
CREATE BUFFERPOOL "COGNOS_32KPB"  SIZE 250 PAGESIZE 32768
CONNECT RESET
CONNECT TO CSO2
CREATE TEMPORARY TABLESPACE TSN_SYS_COGNOS IN DATABASE PARTITION
GROUP IBMTEMPGROUP PAGESIZE 32768 MANAGED BY AUTOMATIC STORAGE
EXTENTSIZE 16 PREFETCHSIZE 16 BUFFERPOOL COGNOS_32KPB OVERHEAD
12.670000 TRANSFERRATE 0.180000 FILE SYSTEM CACHING DROPPED TABLE
RECOVERY OFF
CREATE USER TEMPORARY TABLESPACE TSN_USR_COGNOS IN DATABASE
PARTITION GROUP IBMDEFAULTGROUP PAGESIZE 16384 MANAGED BY AUTOMATIC
STORAGE EXTENTSIZE 16 PREFETCHSIZE 16 BUFFERPOOL COGNOS_16KBP
OVERHEAD 10.500000 TRANSFERRATE 0.140000 FILE SYSTEM CACHING DROPPED
TABLE RECOVERY OFF

CREATE REGULAR TABLESPACE TSN_REG_COGNOS IN DATABASE PARTITION GROUP
IBMDEFAULTGROUP PAGESIZE 16384 MANAGED BY AUTOMATIC STORAGE
AUTORESIZE YES INITIALSIZE 32 M MAXSIZE NONE EXTENTSIZE 16
PREFETCHSIZE 16 BUFFERPOOL COGNOS_16KBP OVERHEAD 10.500000
TRANSFERRATE 0.140000 DROPPED TABLE RECOVERY ON

GRANT CREATETAB ON DATABASE  TO USER "DB2USER"
GRANT BINDADD ON DATABASE  TO USER "DB2USER"
GRANT CONNECT ON DATABASE  TO USER "DB2USER"
GRANT IMPLICIT_SCHEMA ON DATABASE  TO USER "DB2USER"
GRANT USE OF TABLESPACE "TSN_USR_COGNOS" TO USER "DB2USER"

COMMIT WORK
CONNECT RESET
```

Figure 2-6 Content Store setup and tuning

Web server

IBM Cognos provides a web-based interface for administrative activities and to access reports and tools. Thus IBM Cognos gateway component and a web server are integrated. IBM Cognos is compliant with any web server. However, the preferred practice web servers are Internet Information Services (IIS) on Microsoft Windows, IBM HTTP Server for AIX, and Apache for other UNIX variants.

Application server

The IBM Cognos distribution contains the Apache Tomcat server to host the content manager and dispatcher in the application tier. However, IBM Cognos can be integrated with other application servers, such as WebSphere Application Server. We configured IBM Cognos with the default Apache Tomcat distribution.

LDAP server

User-based access control and profiling are key features of IBM Cognos. By default, Cognos uses anonymous access. Based on the business requirements, the user profiles are integrated by using LDAP, SAP, and Active Directory. We used the IBM Tivoli LDAP server for namespaces. The preferred practice is to use the IBM Tivoli LDAP server installation and user profile import process.

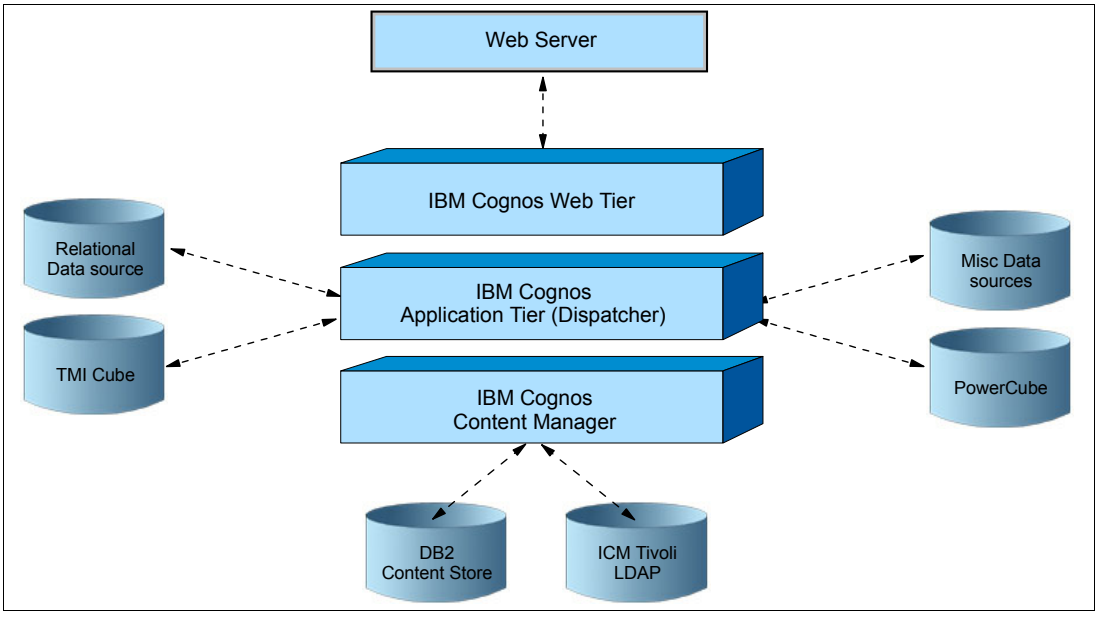

Figure 2-7 Cognos deployment with the associated software stack

2.2.2 Default configuration with gateway, application tier, content manager, and TM1

We describe an IBM Cognos deployment with a single dispatcher that uses a 64-bit Cognos 10.1.1 BI distribution. Under a single dispatcher configuration, three major components of IBM Cognos, web, application tier, and content manager, are installed in three LPARs. The details of the installation and configurations are provided.

Content manager deployment

From the installer, select only the content manager component, as shown in Figure 2-8 on page 59. The file system installation path must have at least 4 GB of disk space.

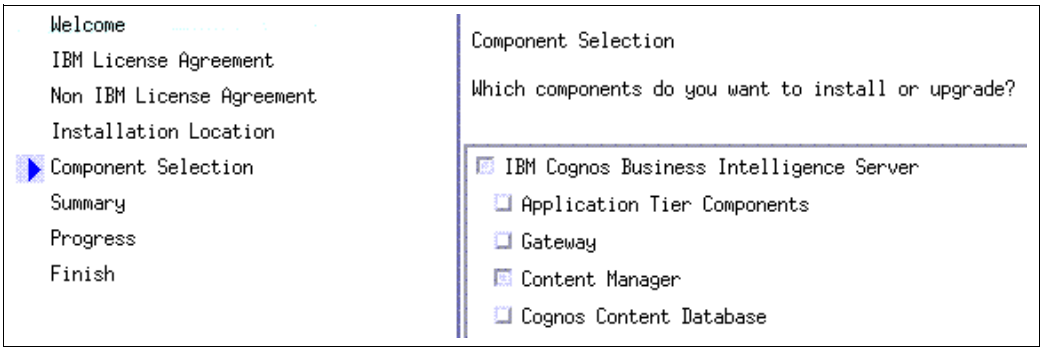

Figure 2-8 Content manager option selection from Cognos installer

After the installation completes successfully, open `cogconfig.sh` from <*installation location*>/`bin64`. We change the configuration to the settings that are described next.

Environment section

The dispatcher and content manager URLs are configured in this section.

The content manager publishes two web URLs. The first web URL is the dispatcher URL. It is registered in the content manager and used to set the runtime setting from the Cognos connection. The second web URL is the content manager URL, which is accessed from the dispatchers to communicate with the content manager.

In Cognos, the configuration replaces the localhost parameter in the URL sections with the IP address or the fully qualified host name of the LPAR that hosts content manager. See Figure 2-9.

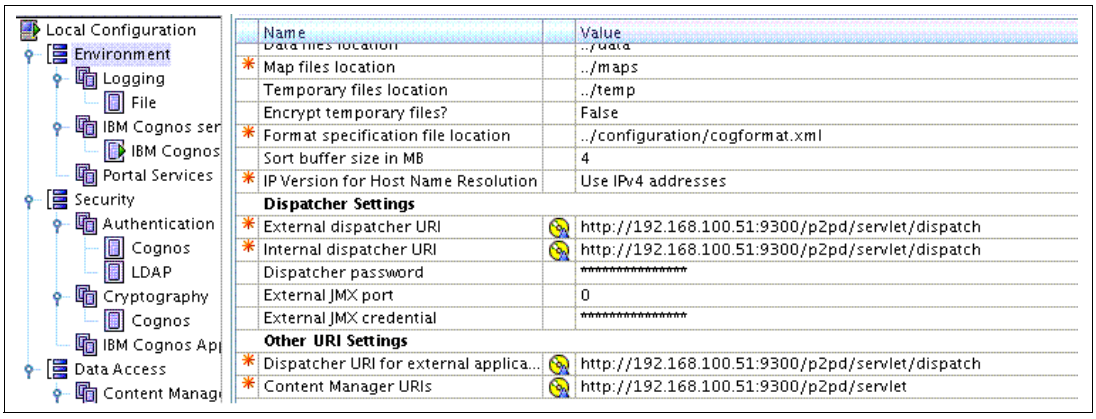

Figure 2-9 Content manager configuration and environment setting

Service setting

Each IBM Cognos dispatcher hosts several services and each service performs several predefined functions. For the IBM Cognos content manager, keep only the relevant services (as shown in Figure 2-10 on page 60) that run from the content manager server.

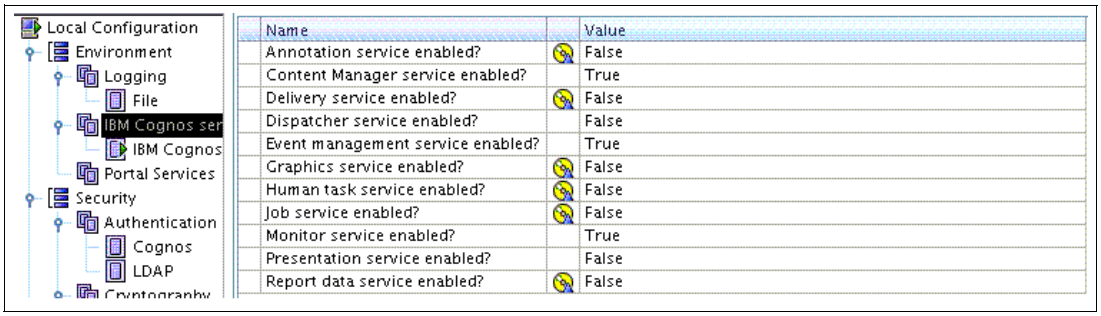

Figure 2-10 Content manager service configuration

Content Store configuration

The content manager is the only component that directly interacts with the Content Store to save and retrieve important Cognos information. We provide the Content Store type, database server, port number, user credential, and database name under the Content Store configuration. Validate the Content Store configuration after the configuration is done. To perform the test, select the Content Store from the left pane and select **Test**, as shown in Figure 2-11.

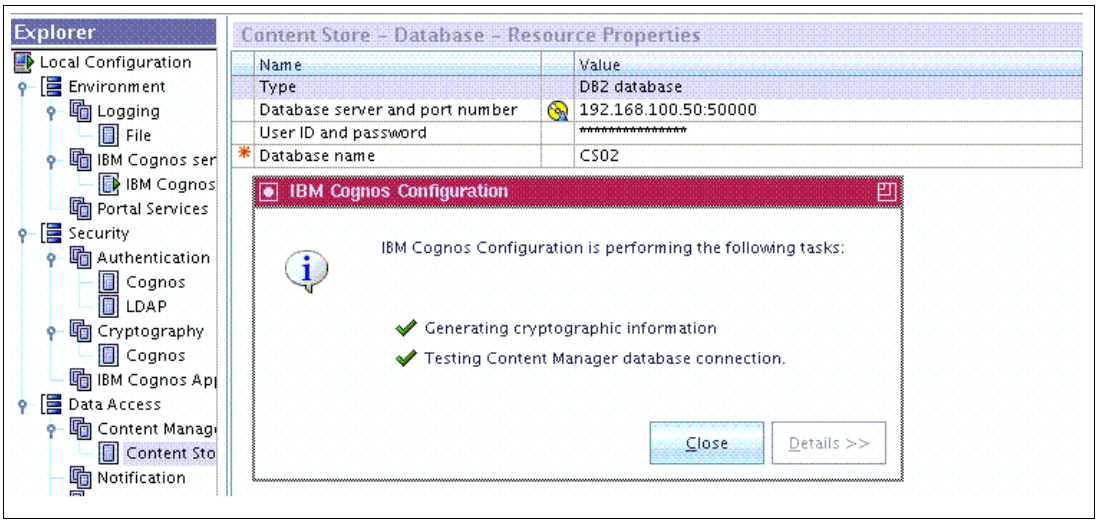

Figure 2-11 Content Store configuration and validation

Authentication

Unless a namespace is created anonymously, access must be limited. After a namespace is created, the access parameter is set to `false` so that only valid users can access the Cognos environment. See Figure 2-12 on page 61.

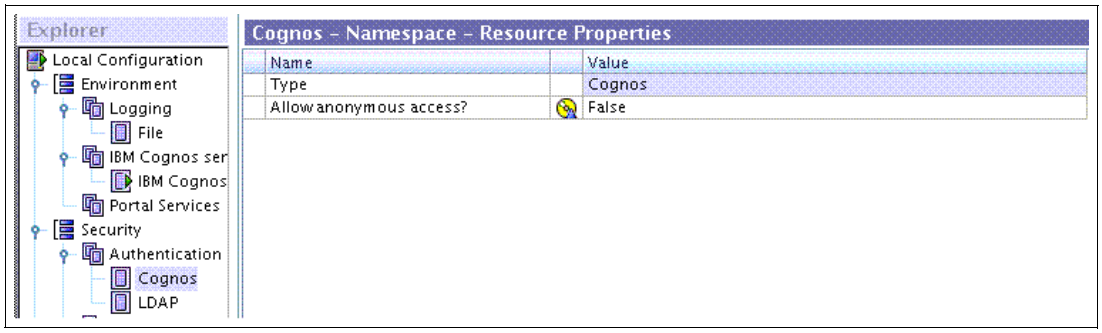

Figure 2-12 Anonymous login disabled

In this residency, we configured an LDAP namespace, which is added under the Authentication LDAP, as shown in Figure 2-13.

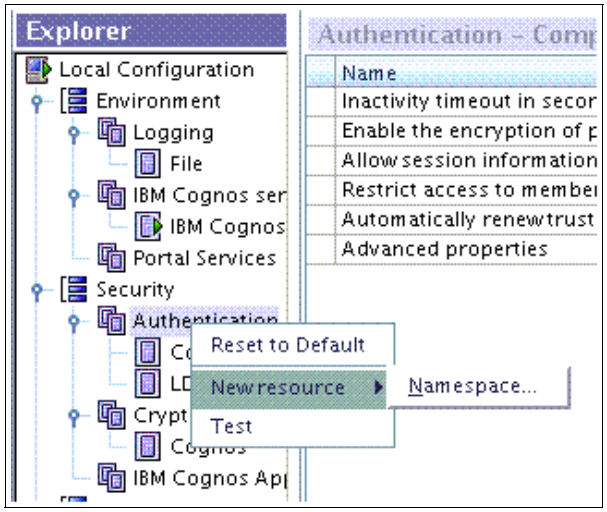

Figure 2-13 LDAP namespace add

The LDAP namespace is configured with the LDAP server IP address and port. You must provide values for the Base Distinguished Name and User lookup. Also, validate the configuration before you proceed to the next configuration. To validate the configuration, select the LDAP namespace from the left pane, right-click, and select **Test**. The test must be successful. See Figure 2-14 on page 62.

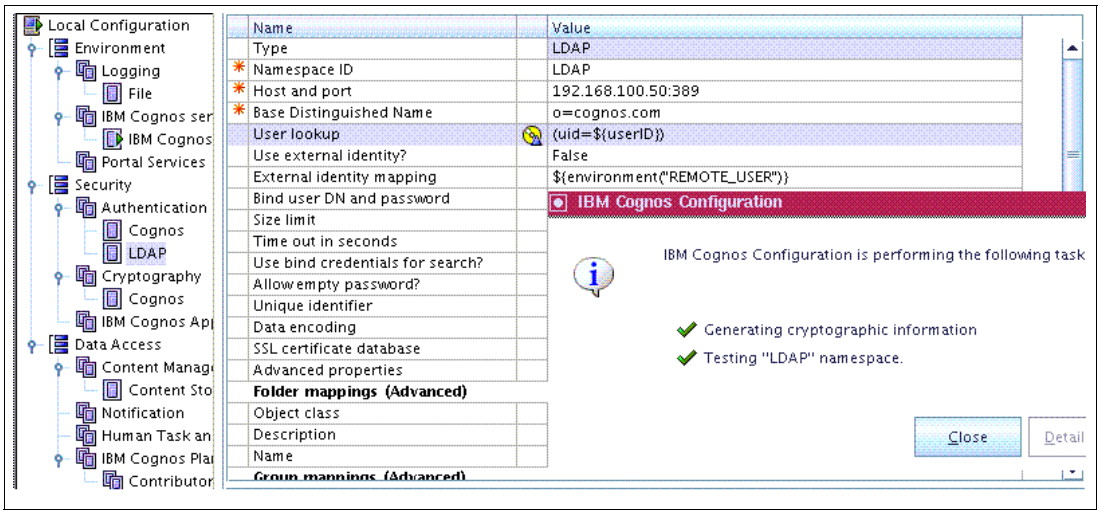

Figure 2-14 LDAP namespace configuration and validation

After the configuration is complete, save the configuration and start the content manager service. On a successful start, validate the content manager URL (as it is configured under the environment section) from a web browser, as shown in Figure 2-15.

Figure 2-15 Validation of the content manager

Application tier deployment

The IBM Cognos application tier contains the dispatcher (Figure 2-16 on page 63), which has several services, where reports are generated and delivered based on the delivery mode and schedules. To install the dispatcher, select **Application Tier Component** from the installer and maintain a minimum disk space of 4 GB.

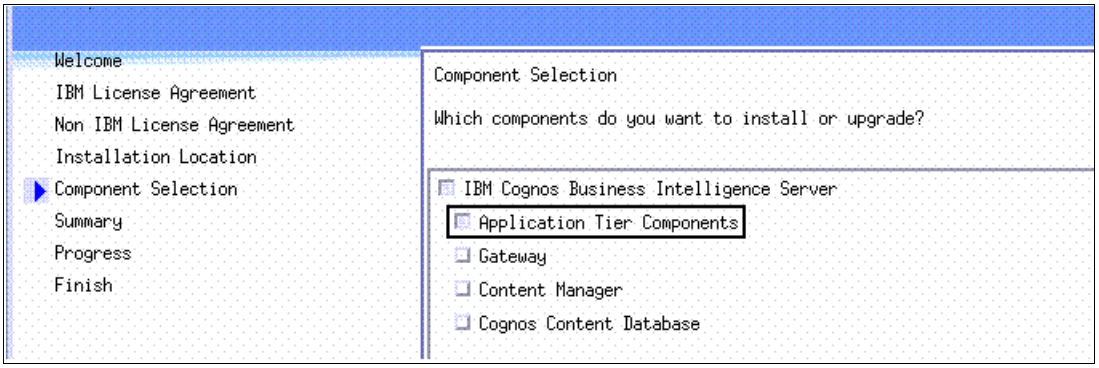

Figure 2-16 Application tier component selection

After the software is installed, the configuration starts with environment parameters, where localhost is replaced by the LPAR IP address or host name. In the dispatcher setting, provide the content manager URL and the gateway URL, as configured in earlier steps. See Figure 2-17.

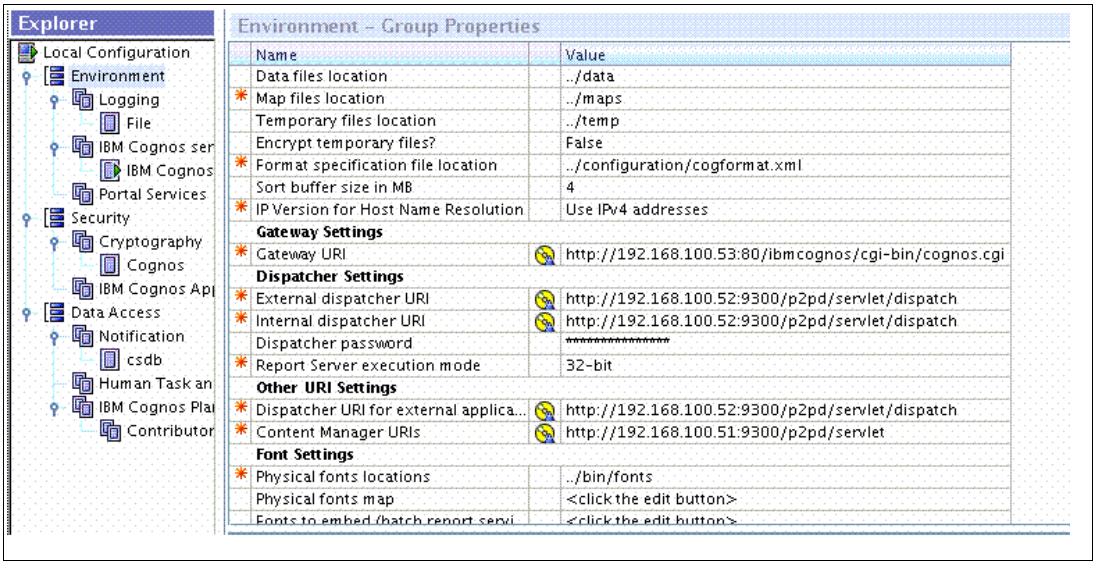

Figure 2-17 Application tier environment configuration

As a single dispatcher deployment, all the services are enabled from the IBM Cognos service configuration, as shown in Figure 2-18 on page 64.

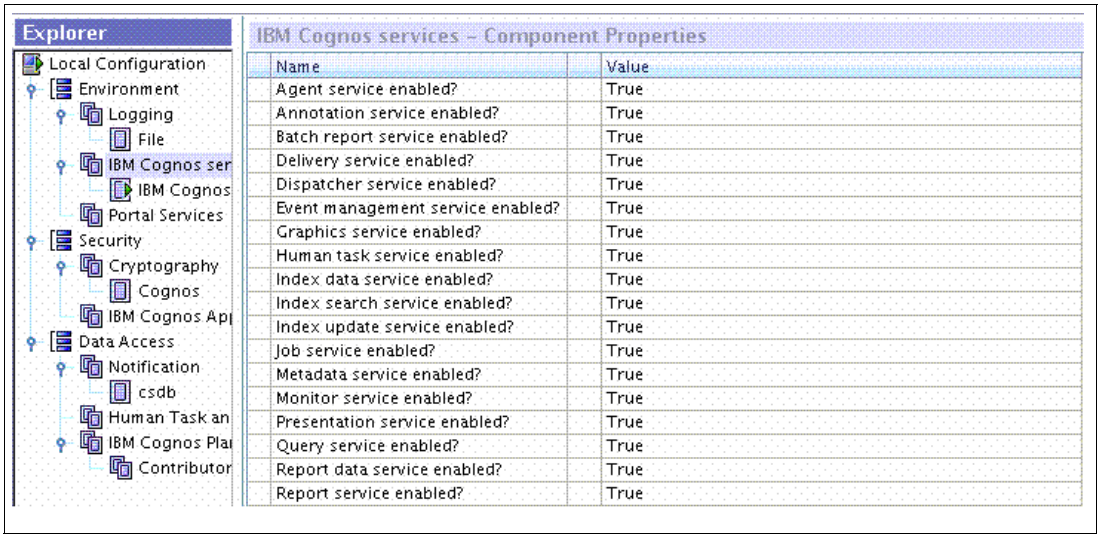

Figure 2-18 All services are activated for a single dispatcher environment

Save the configuration and start the service. After the service starts, verify the dispatcher URL from a web browser. We configured the LDAP name in the content manager, so the dispatcher URL prompts for credentials, as shown in Figure 2-19.

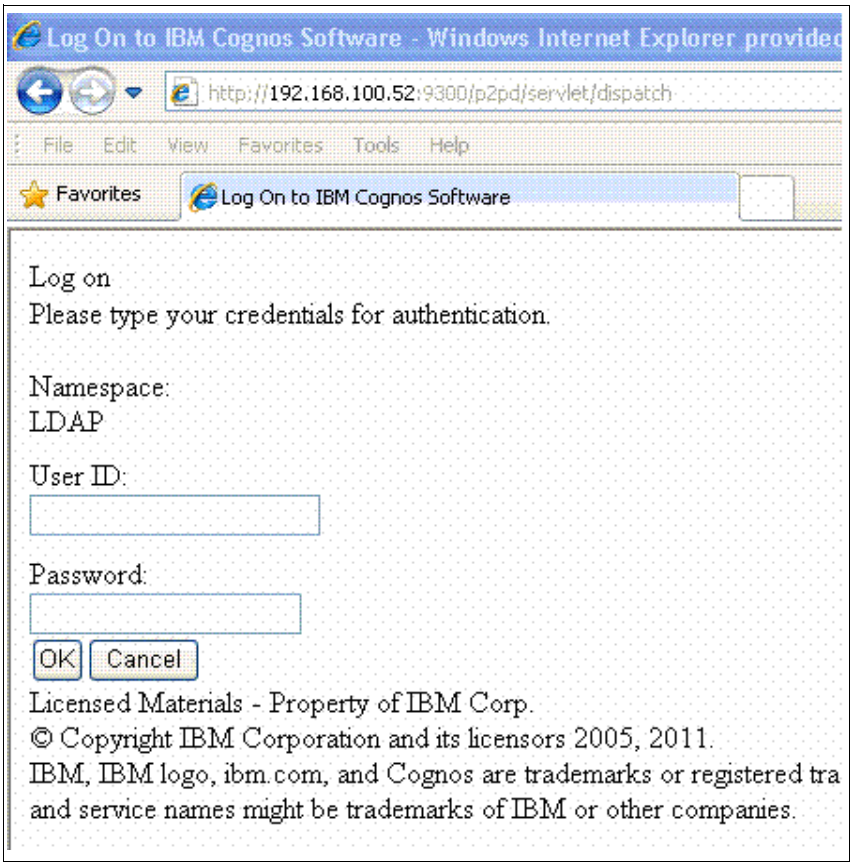

Figure 2-19 Validation of dispatcher URL

Web tier deployment

The web tier configuration starts with the installation gateway component in the `opt/IBM/Cognos/cgi-bin/` path. See Figure 2-20.

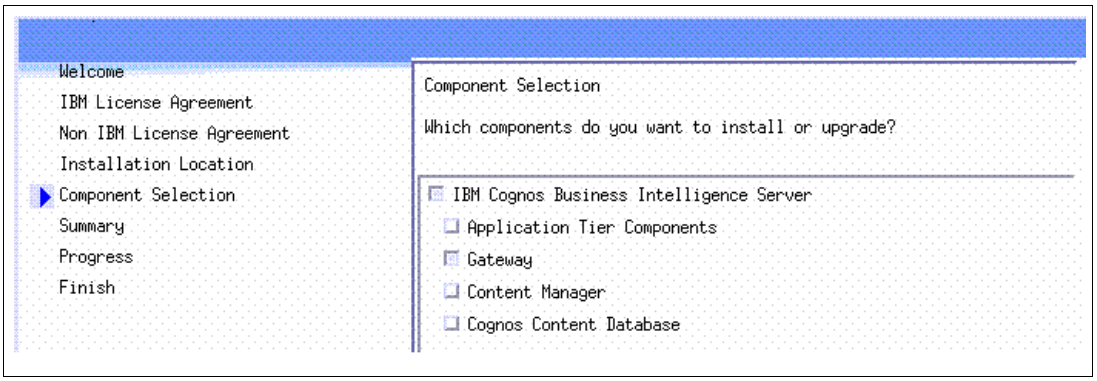

Figure 2-20 Cognos gateway installation

After the installation starts, provide the dispatcher URLs. If there are multiple dispatchers, all of them can be added in the sequence of preference, as shown in Figure 2-21. In the controller URL configuration, replace the localhost with the IP address or host name of the LPAR.

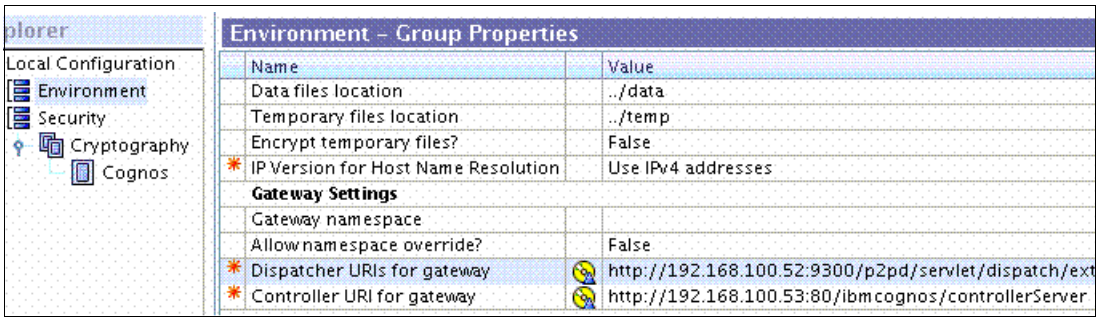

Figure 2-21 Cognos gateway configuration

To integrate IBM Cognos with the HTTP server, modify the `httpd.conf` file to add virtual directories, as shown in Figure 2-22 on page 66.

```
   ScriptAlias /ibmcognos/cgi-bin "/opt/IBM/Cognos/cgi-bin/"

   <Directory "/opt/IBM/Cognos/cgi-bin/">
       AllowOverride None
       Options +ExecCGI
       Order allow,deny
       Allow from all
   </Directory>

   Alias /ibmcognos "/opt/IBM/Cognos/webcontent"

   <Directory "/opt/IBM/Cognos/webcontent/">
       AllowOverride None
       Options Indexes FollowSymLinks MultiViews
       Order allow,deny
       Allow from all
</Directory>
```

Figure 2-22 Virtual directory configuration in the httpd.conf file

After you save the `httpd.conf` file, the web server restarts to activate all changes. Validate the gateway URL from a web browser, which prompts for credentials from the LDAP server.

TM1 server deployment

The TM1 server configuration on an AIX environment starts with the installation. Use the standard installer. To install the TM1 server on AIX, the display must be exported to your local personal computer by using either the X Windows server program or a Virtual Network Computing (VNC) server. Ensure that the *JAVA_HOME* environment variable is set to a valid Java Runtime Environment (JRE) location.

You must ensure that the memory resource limits are set to the maximum available memory:

```
ulimit -d -H
```

This installation modifies the `/etc/services` file. It also adds a startup script named `S89IBM_TM1` to the `/etc/rc.d/rc2.d` directory.

From the installation media, run the installation script:

```
./install_aix.bin
```

The installation script starts a graphical welcome window. Press Enter to progress to the window where you choose the products to install.

The TM1 server has two primary components, such as the admin server and the data servers, as shown in Figure 2-23 on page 67. One TM1 server deployment must have one admin server and one or more data servers. Admin servers and data servers can be configured on different hardware and operating system environments. The server where the TM1 admin server is configured is called `admin host`. As the TM1 data servers start, they register themselves to the TM1 admin server. The admin server checks the availability of the data servers through a heartbeat acknowledgement that is sent every 60 seconds. The default port number for the admin server is `5495`; however, this port can be reconfigured. Several TM1 clients connect to the admin server to reach registered data servers. The TM1 data server spawns an admin server local system if an admin server local system is not available when the data server starts. If the admin server stopped, it stops all the registered data servers.

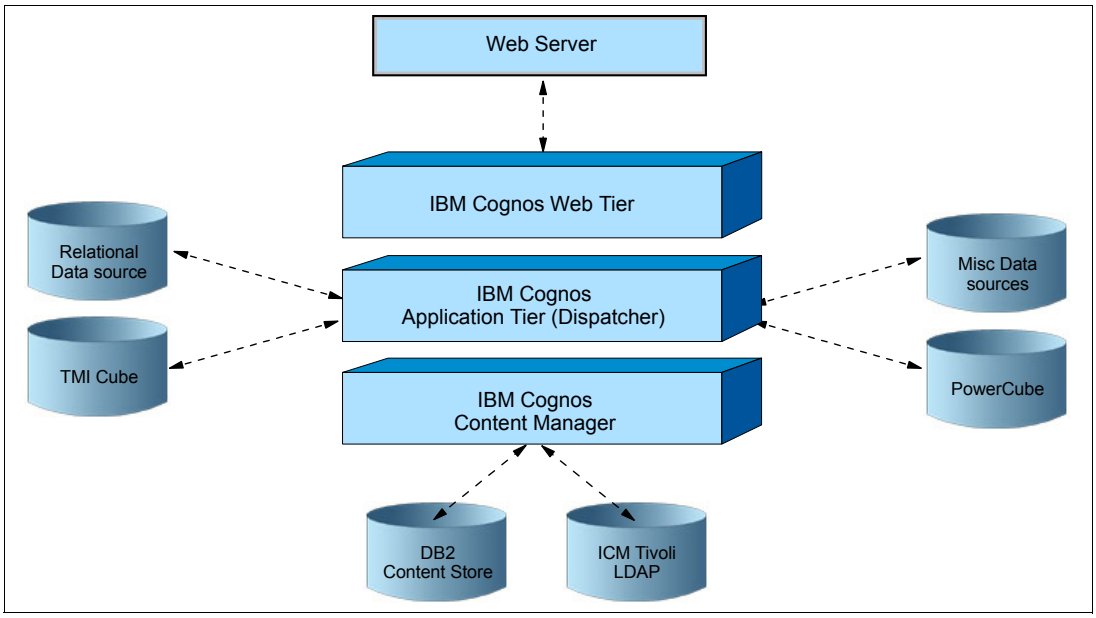

Figure 2-23 TM1 admin server and data servers

TM1 is deployed with a two-tier architecture. The web and gateway tier provides the front-end and web server for the TM1 server. The TM1 web and gateway tier runs on Microsoft Windows only because it requires IIS to provide web and application services. The other component is the TM1 data tier. Figure 2-24 on page 68 shows the logical diagram of the TM1 architecture.

Important: When you unpack the TM1 packages, ensure that you unpack them into separate directories before you untar the images. The packages have common named files and cannot be combined into a single installation directory.

When you install the TM1 server 10.1 on AIX, ensure that the *PATH* and `JAVA_HOME` variables are set to use the 64-bit version of Java. Otherwise, the configuration tool fails with an unclassified link error.

Ensure that all the prerequisite software and features are enabled on the Windows portion of the installation.

You must pre-configure Windows before you install TM1.

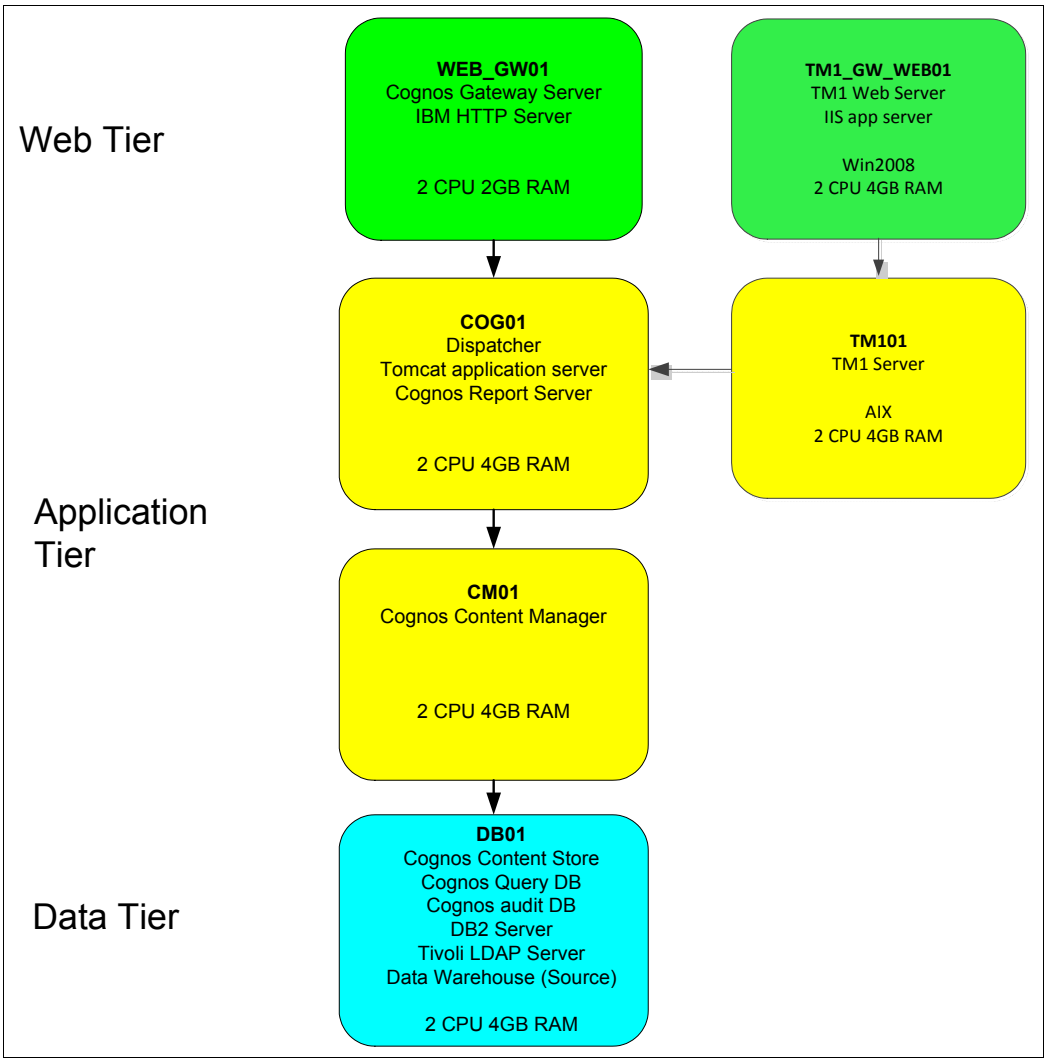

Figure 2-24 Logical diagram for TM1

TM1 server configuration

Accept the license agreement and select the installation directory. Then, click **Next** to accept the default administration ports.

Enter the following information to complete the TM1 server installation:

- Server name: The name that you want to use to identify the TM1 server.
- Server port number: The port on which the TM1 server runs. This parameter is used to distinguish multiple TM1 servers that run on the same computer. The valid port number values are `5000 - 49151`. The default is `12345`.
- TM1 admin host: The name of a computer on which a TM1 admin server is installed. When a TM1 server is running, it registers itself on this admin server. TM1 clients connect to this admin server to obtain information about which TM1 servers are available on a network. The admin server can be on the same computer on which the TM1 server is installed or another computer on your network.
- Do you want to install the sample data: Select this option if you want to install the TM1 sample data, which is referenced extensively in the TM1 documentation. Clear this option if you do not want to install the sample data.

- Path to the TM1 database: Specify the path to the TM1 server data directory. Verify that the summary window details are correct and there is enough installation space. Then, click **Next** to finish the TM1 server installation.

TM1 web

The TM1 web component must run on Windows because it uses Windows IIS as a web and application server. The IIS component within Windows must be activated and configured.

Integration of TM1 web with Cognos BI

We explored the integration of Cognos BI and the TM1 environment. In this environment, the TM1 servers use Cognos BI authentication through the IBM Cognos Application Firewall (CAF) authentication method, as shown in Figure 2-25. The Cognos BI dispatcher CAF authentication enables the TM1 web server to be added as a valid domain or host.

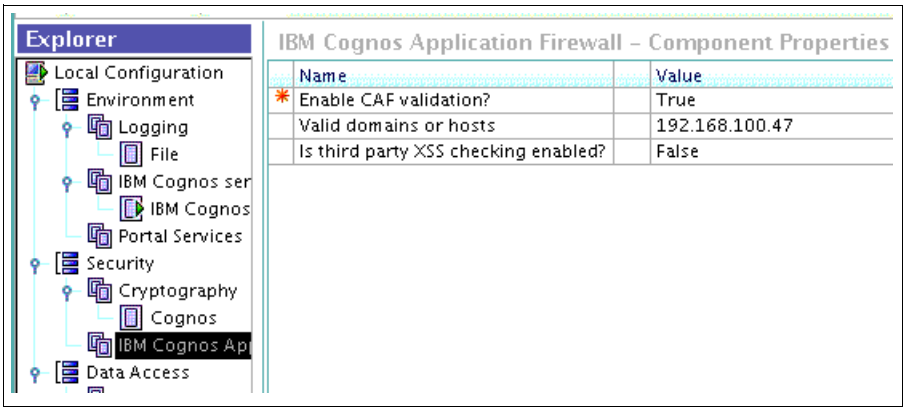

Figure 2-25 Cognos BI dispatcher configuration for CAF validation

You must edit the `variables_TM1.xml` file to provide TM1 web URL and server details. After all modification is done, this file is saved in the *<Cognos 10 installation>*`/templates/ps/portal` path. In the same location, a sample file, which can be modified and renamed as `variables_TM1.xml`, is available.

TM1 tools

TM1 has Windows based desktop utilities, some of which are web-based tools. The desktop tools are used for TM1 server monitoring and server administration activities, such as setting security, user management, and backup. The web clients are mostly used to browse the TM1 server, views, and cubes.

TM1 architect

The TM1 architect is the primary TM1 client tool. It is a desktop utility that is mostly used for TM1 server administration, backup, security, and access control. The TM1 architect connects to the TM1 admin server to obtain the list of the available TM1 data servers.

TM1 perspective

The TM1 perspective is an Excel extension for TM1. It uses Microsoft Excel to open reports.

TM1 web client

The TM1 web client is a browser-based solution for server monitoring. The user can access all cubes and views through the TM1 web.

2.2.3 Baseline performance with DB2, TM1, and PowerCube data sources

IBM Cognos reports represents data in a custom format. A data source contains the information to connect to a database, such as the IP address or host name of the database, user credentials, and timeout parameter. We create data sources by using DB2, PowerCube, and TM1 cubes. All data sources are created through the Cognos connection portal. Log in to Cognos connection portal and select the **Configuration** tab from the IBM Cognos Administration window to show all available data sources and options to create a data source.

DB2 data source

In the data source creation wizard, select **IBM DB2** as the type. The *isolation level* specifies the transaction to modify the database. By default, it is **Use the default object gateway**, as shown in Figure 2-26.

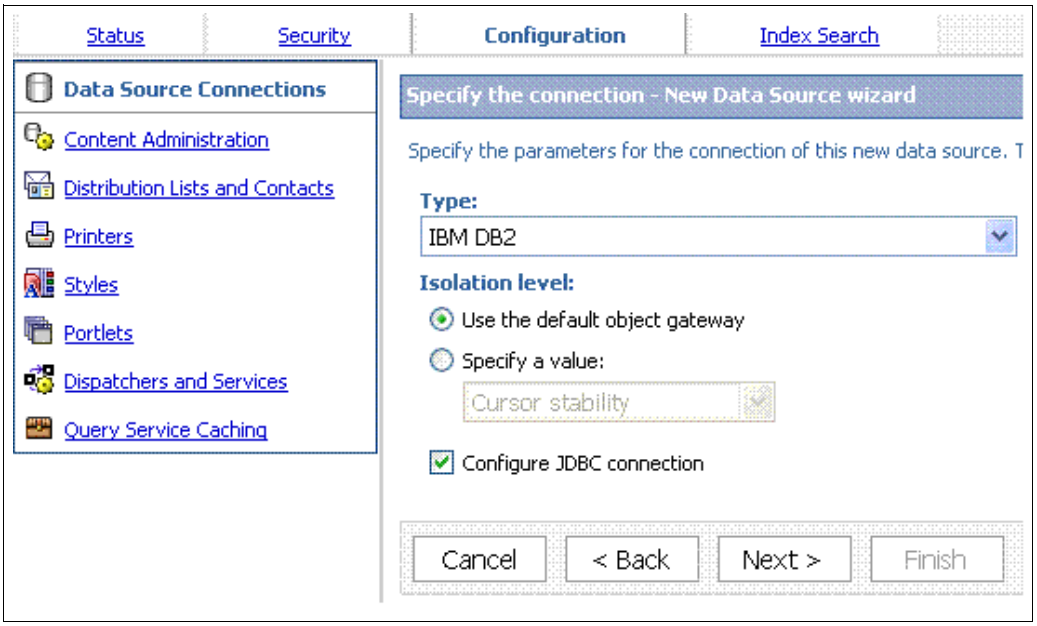

Figure 2-26 DB2 data source creation

There are two types of DB2 data sources: Java Database Connectivity (JDBC) and the DB2 command-line interface (CLI). The classic Cognos Query engine uses the DB2 CLI. Configure the CLI data source by using the database name and a sign-on that is created by using the credential of the user. Validate the connection before you save it. See Figure 2-27 on page 71.

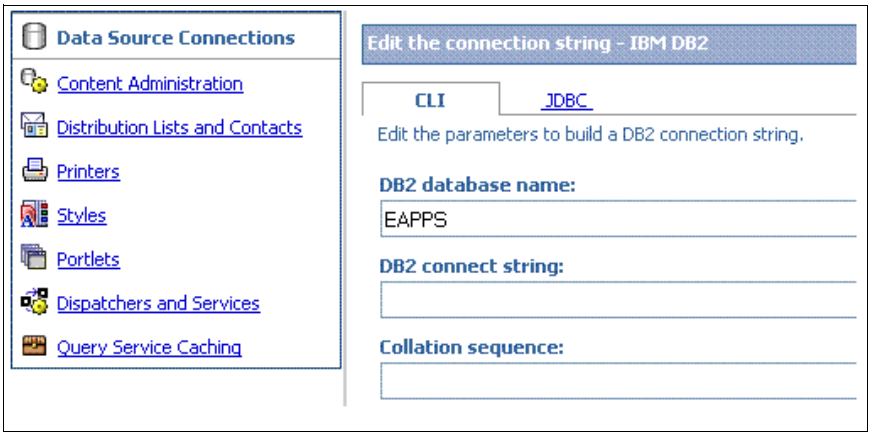

Figure 2-27 DB2 CLI data source

The DB2 JDBC data source connectivity is configured with the database server name, port number, and database name. Validate the connectivity and save the data source, as shown in Figure 2-28.

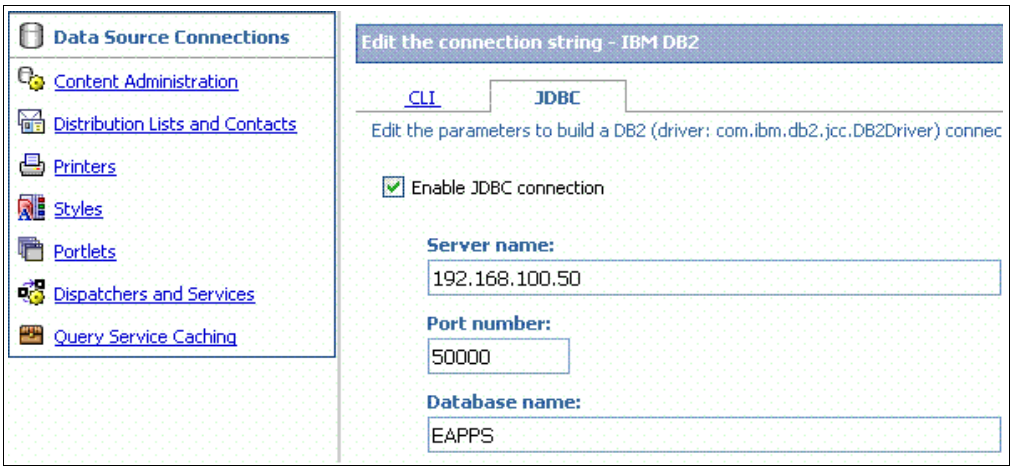

Figure 2-28 DB2 JDBC data source connectivity

PowerCube data source

The IBM Cognos PowerCube is created through the transformer. The PowerCube data sources refer to the path of the cube from the file system. Reports are installed on the specific path in the AIX environment. But the data source wizard does not validate an empty Windows location. So, you must enter a value. The PowerCube location must have *read* permission. The default read cache size is 80 MB, which can be extended up to 1 GB for optimal query performance, as shown in Figure 2-29 on page 72.

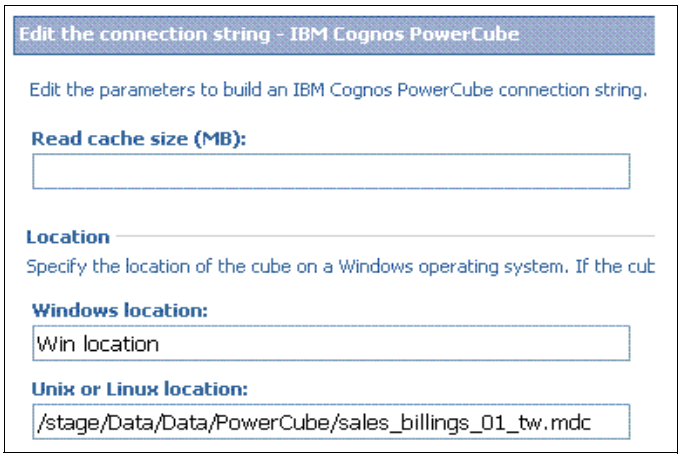

Figure 2-29 PowerCube data source

2.2.4 Cognos dispatcher tuning and BI Bus tuning

Cognos dispatcher tuning and the BI Bus tuning are described.

Java virtual machine memory resource for IBM Cognos service

By default, 768 MB of memory is assigned to the IBM Cognos process initially. Dispatchers that host the report service or batch report service are tuned with 1536 MB as Java virtual machine (JVM) memory for higher throughput. The tuning is done from the Cognos configuration console, and it restarts the dispatcher to implement the change. The steps are shown next.

Start the Cognos configuration from the `bin64` directory in the IBM Cognos installation path. In the reference environment, start the Cognos configuration through the `cogconfig.sh` from `/opt/IBM/Cognos/bin64`. Select **IBM Cognos service** from left pane and modify the maximum memory in MB to `1536`, as shown in Figure 2-30. Save the configuration and restart the dispatcher.

Note: For the maximum memory field, and as a starting point for IBM Cognos V10.1.1 and later, we recommend at least 4 GB (4096) of memory.

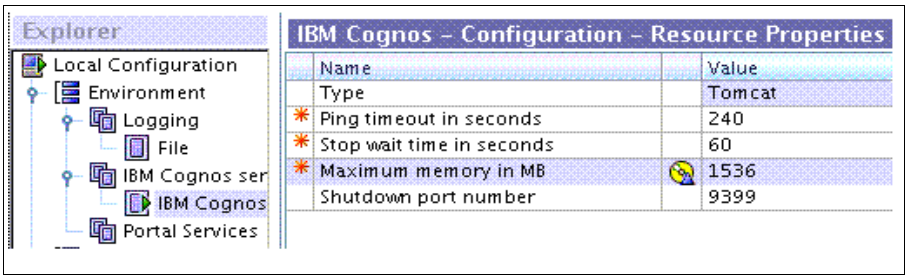

Figure 2-30 JVM memory setup

Report service and batch report server BI Bus workload thread tuning

Report service and batch report service are among the primary services that are hosted by a dispatcher. Interactive reports are served by the report server and scheduled/batch reports are supported by the batch report service. To handle multiple concurrent reports, each dispatcher spawns several concurrent threads that are known as *BI Buses*. The higher the

number of BI Buses, the more concurrent reports are executed. However, as the number of concurrent users increases, the IBM Cognos server needs more resources, as well. The configuration steps are shown.

Log on to the Cognos portal, start IBM Cognos administration, and select the **Configuration** tab. Select **Dispatchers and Services** from the left pane to show the number of registered dispatchers in the right pane. See Figure 2-31.

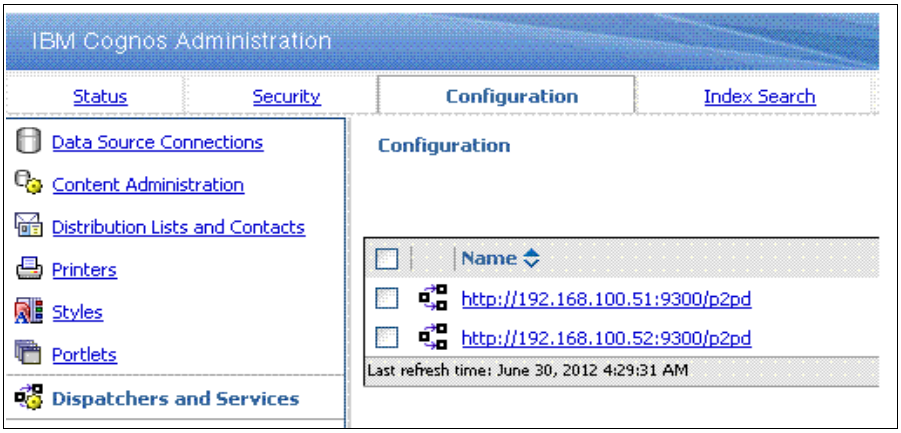

Figure 2-31 Dispatcher list from the Cognos Administration portal

Select any of the available dispatchers to see a list of the services that are hosted in that dispatcher, such as the report service, which shows in the list of registered services. See Figure 2-32.

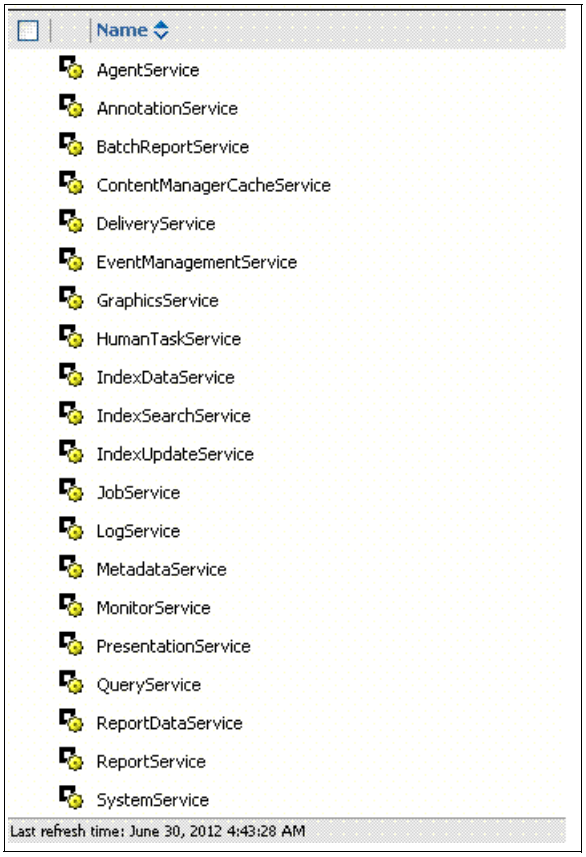

Figure 2-32 Service list from the dispatcher

Select the edit property for the service. Then, select the “setting” section, which contains several tunables. Increase the maximum number of processes for the report service during a peak period (Figure 2-33). Save the configuration and the modified value is activated without a Cognos server restart.

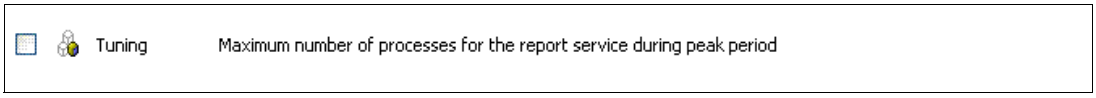

Figure 2-33 BI Bus at peak period

High and low affinity connections in report and batch report server

A *high affinity request* is a transaction that can benefit from previously processed requests. High affinity configuration helps reports that use cached data. This tuning is mostly used by report servers. *Low affinity connections* are primarily used for batch reports for scheduled activities (Figure 2-34). These tunables are managed through the Cognos web portal.

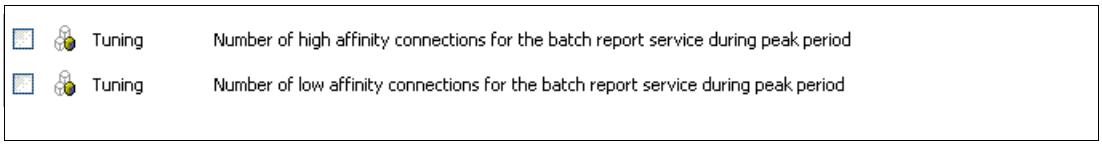

Figure 2-34 High and low affinity tuning for report and batch report services

Report bursting

The *report bursting* method produces reports that are based on a common report definition with personalized content. Bursting performs a single execution of a report and sections the content as required, typically based on security access. Then, it distributes the sections to the appropriate users based on the report content. Bursting is an important part of any large-scale enterprise reporting solution.

Report bursting improves the performance scalability of the Cognos server through the effective management of resources. It reduces network traffic, minimizes database queries, and enables IBM Cognos BI to process multiple personalized reports in parallel.

Dynamic query

Dynamic query is a Java based query mode with the following advanced features:

- Query optimization to address complex queries, data volume, and faster response time
- Capability to generate advanced and efficient Multidimensional Expression Language (MDX) to use the local and remote processing ability to improve complex online analytical processing (OLAP) queries
- JDBC connectivity for relational databases
- Dimensionally modeled relational (DMR) package for OLAP functionality
- Taking advantage of 64-bit processing
- Security-aware caching

2.2.5 PowerVM, AIX tuning, and performance statistics

The AIX and PowerVM features that are necessary to improve the performance and scalability are described. The associated tuning options are explained.

The baseline measures are taken with the AIX default settings. To improve the performance and scalability, we set up the following AIX environment variables for Cognos 10.1.1:

```
AIXTHREAD_MINKTHREADS = 32
AIXTHREAD_MNRATIO = 1:1
AIXTHREAD_MUTEX_FAST = ON
AIXTHREAD_SCOPE = S
SPINLOOPTIME = 4000
YIELDLOOPTIME = 20
```

These environment variables control the way that the multi-threading and locking are handled on the system.

With `AIXTHREAD_SCOPE = S` and `AIXTHREAD_MNRATIO = 1:1`, the scheduling of the threads is done directly by the system instead of the POSIX thread library (libpthreads). This scheduling ensures that one POSIX thread maps to one kernel thread to avoid serializing several pthreads to one kernel thread.

The variable `AIXTHREAD_MINKTHREADS = 32` ensures that a minimum of 32 kernel threads are available for a process. This number increases the scheduling possibilities over the default value of 8.

By using `SPINLOOPTIME = 4000` and `YIELDLOOPTIME = 20`, you can control the number of times that a pthread spins when it attempts to obtain a lock before blocking on a mutex. You can control the number of times that a pthread yields itself before blocking on a mutex. With these settings, you can avoid pthreads that retry too infrequently to obtain the locks. You can avoid pthreads that go inactive (asleep) too quickly and for an extended period.

The last `AIXTHREAD_MUTEX_FAST = ON` variable enables the pthreads to use an optimized mutex-locking mechanism, which provides a performance boost. This mechanism is available for private mutexes only. Only mutexes that are initialized by pthread_mutex_init are optimized. The effect in transactions per second (TPS) of applying these environment variable settings is shown in Table 2-1.

Table 2-1 Test results with the number of users and environment variable settings

	Environment variables not set	**Environment variables set**
300 users	40 TPS	58 TPS
500 users	Not tested	50 TPS
500 users reduced think time	Not tested	48 TPS

Sample results only: The number of transactions per second displayed is the result of tests in our environment. It is only an example, and these figures might not compare to what you can obtain in your environment.

These variables show a significant improvement during our tests. We managed to increase the number of concurrent users from 300 to 500 and increased throughput from 38 reports per second to 50 reports per second.

We tried to increase the number of users further, but then we noticed that we needed more dispatchers on the Cognos side. These environment variables are important for the scalability of the solution. Use them to take advantage of the Power Systems architecture.

2.3 Cognos multiple dispatcher deployment

A typical multiple dispatcher Cognos deployment consists of several application tiers or dispatchers where each dispatcher can be configured to host all or a selected number of services. Also, the services can be stopped or started from the Cognos Administration console. As shown in Figure 2-35, the web-tier and content manager remain the same, similar to a single dispatcher configuration, but new dispatchers are added to the infrastructure.

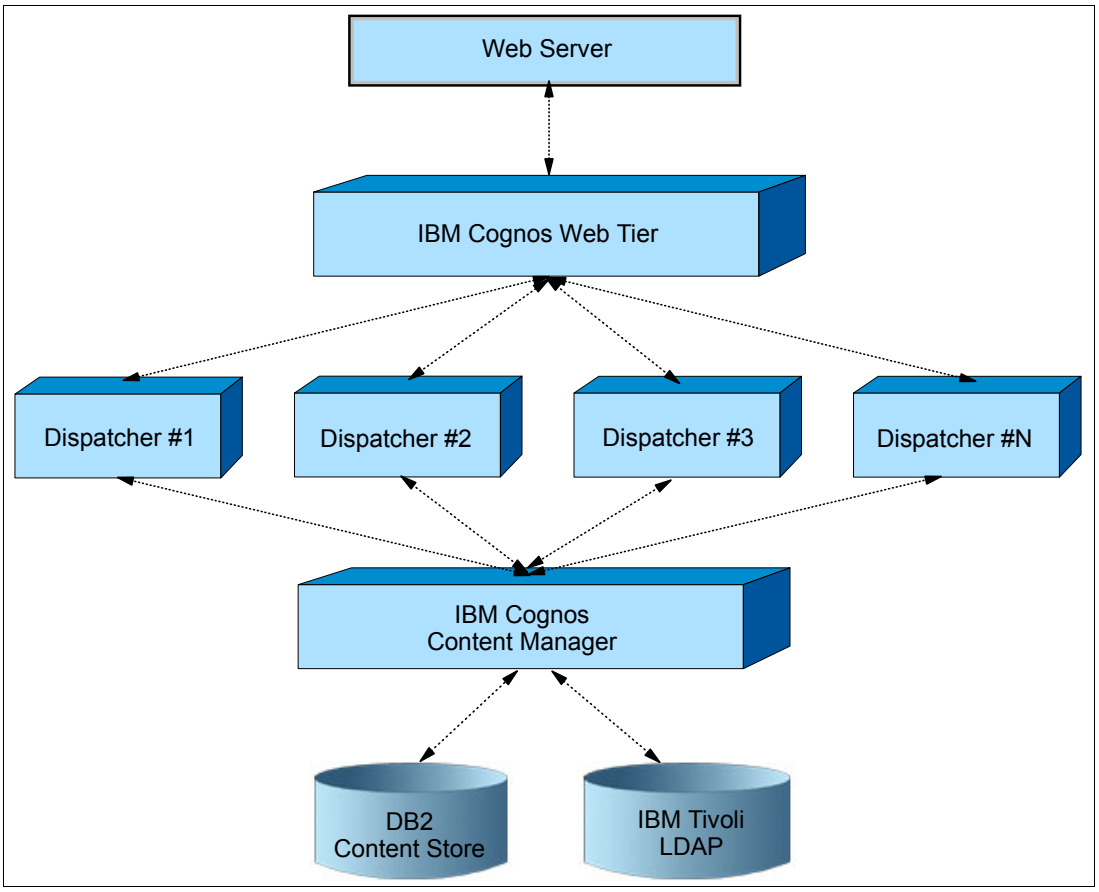

Figure 2-35 Multiple dispatcher environment

The multiple dispatcher setup offers the following primary advantages:

- Higher processing capacity: Each dispatcher registers itself to the content manager with all hosted services. A new request is routed to the next available dispatcher in a round-robin algorithm. With this workload distribution, the users get better performance and faster throughput.
- Redundancy and failover: In a multiple dispatcher environment, similar services are hosted through several other dispatchers. Even if one dispatcher stops responding, reports are served by other available dispatchers. Sometimes, the servers need outages for several reasons. In this scenario, the Cognos environment keeps serving users through other available dispatchers.
- Dispatcher with designated services: Each dispatcher can host one or many services. The services of each dispatcher can be started selectively based on the requirement.

2.3.1 Description of default setup

The multiple dispatcher setup is similar to a single dispatcher environment. The only difference is the additional number of dispatchers, for example, the application tier. We can install the application tier by using the IBM Cognos installer, followed by the required configuration. However, another practice is the installation from the master copy. In this approach, one dispatcher is installed and configured with all required information. Then, the entire installation path is archived to a shared location. During the new dispatcher installation, the archive is extracted in the same path where the earlier dispatcher is installed. Later with minimum changes, the dispatcher starts. Figure 2-36 shows the multiple dispatcher layout for a typical Cognos system.

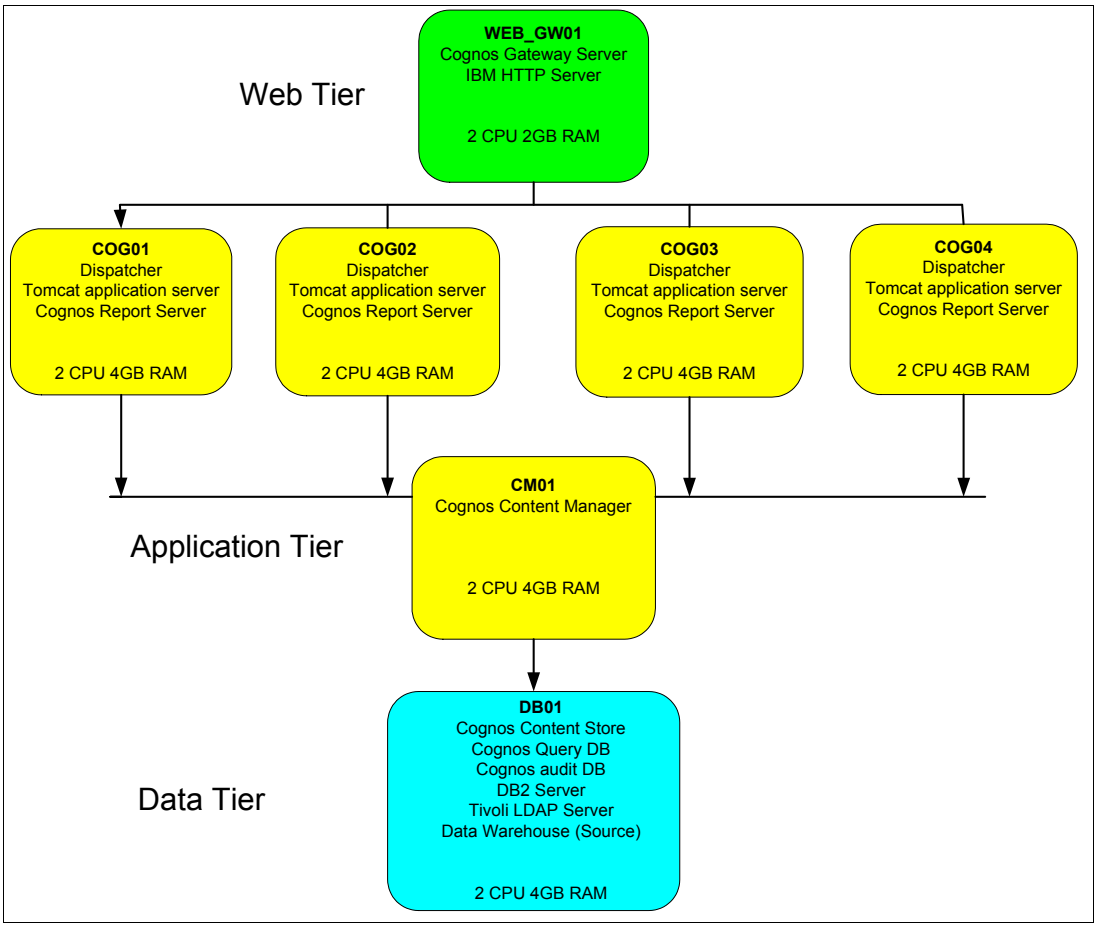

Figure 2-36 Multiple dispatcher Cognos layout

2.3.2 PowerVM, AIX tuning, and performance statistics

This section covers the additional PowerVM and AIX tuning that is applied to provide the best performance with multiple dispatchers. It also shows the statistics that are gathered.

For the multiple dispatcher tests, we used the environment variables. We observed that they improve the scalability significantly. Table 2-2 on page 78 illustrates the various configurations for the tests that we performed.

Table 2-2 Test results with the number of users and dispatchers

TPS results	Two dispatchers	Four dispatchers
1,000 users 8 logical processors 8 GB memory	60 TPS	80 TPS
2,000 users 8 logical processors 8 GB memory	88 TPS	80 TPS
2,000 users 16 logical processors 8 GB memory	118 TPS	120 TPS

Disclaimer: The number of transactions per second displayed here is the result of tests in our environment. The numbers are only examples. These figures might not compare to the results that you can obtain in your environment.

For these tests, we started runs with 1,000 users because we noticed that we needed multiple dispatchers to increase from 500 users. With two dispatchers and 1,000 users, we increased the average number of transactions per second to 60 TPS. We compared that result to 50 TPS with 500 users and a single dispatcher. So, we can see that the number of dispatchers also has a significant impact on the overall scalability.

Trying to further push the scalability tests, we decided to increase the number of concurrent users to 2,000 with two dispatchers. We managed to increase the throughput further with an average of 88 TPS with some peaks at 112 TPS.

After the test, the processor resources became constrained, and we wanted to see the effect of a processor capacity increase on scalability. We noticed improved results on the throughput on the application side.

We changed the next testing steps to use four dispatchers. We returned to the initial processor capacity and 1,000 users as shown in the first test with two dispatchers. We managed to increase the number of transactions per second to 80 TPS. So, we even improved scalability over the two dispatcher solution with a result of 60 TPS. This result further confirms that the number of dispatchers significantly affects the overall performance.

Running 2,000 users with four dispatchers is the next iteration. The results are lower than the results with 2,000 users and only two dispatchers with a throughput of 80 TPS.

Therefore, with 2,000 concurrent users, and the default configuration, scaling up the number of dispatchers might not help because we are processor-bound. We need to increase the processor resources available.

The last test used four dispatchers, 2,000 users, and doubled the processor capacity available. We noticed again an increase in average throughput to 120 TPS for that test, which shows that processor power and the number of dispatchers are important for the scalability.

These tests showed that the solution is scalable and that both the number of dispatchers and the processor power play a role in the scalability.

3

Implementing IBM SPSS on IBM Power Systems

Details about how to install and configure IBM Statistical Package for the Social Sciences (SPSS) on IBM Power Systems that use the AIX operating system are described. This chapter serves as an implementation quick reference guide to deploy IBM SPSS on Power Systems. It also suggests application and system tuning configurations that are based on the testing scenarios that are described in this chapter.

This chapter contains the following topics:

- SPSS Collaboration & Deployment Services
- IBM SPSS Modeler
- SPSS integration with Cognos

3.1 SPSS Collaboration & Deployment Services

We describe the details of setting up SPSS Collaboration & Deployment Services (C&DS). We provide detailed information about installing individual components of the solution, the context diagram, the deployment topology, and the mechanism of interaction between the solution components.

The installation consists of multiple components. Therefore, you must follow the correct order during the installation process. We document this order step-by-step for your ease of understanding and implementation.

We also document all the test scenarios that were simulated, and we document the suggested tuning parameters.

3.1.1 Description of the setup and considerations

We provide details about the setup for the IBM SPSS solution.

Context Diagram

Figure 3-1 shows the context diagram of the SPSS solution.

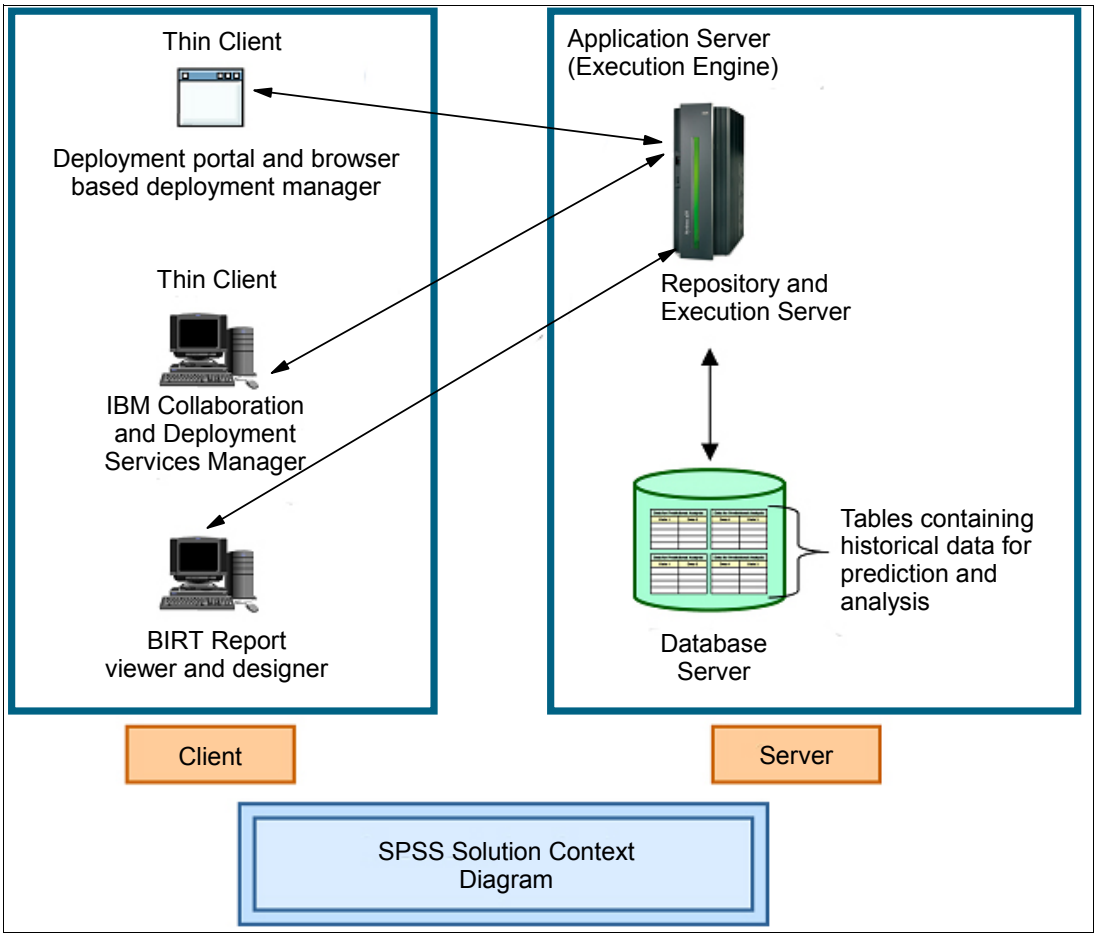

Figure 3-1 Context diagram of the SPSS solution

The SPSS solution typically consists of a centralized IBM SPSS Collaboration and Deployment Services (C&DS) Repository, which runs on a single instance logical partition (LPAR). This repository receives requests from clients, and processes these requests by using execution servers. The requests are generally meant to process analytical data and carry out historical analysis in an attempt to predict future behavior.

The SPSS solution consists of following components:

- IBM SPSS Collaboration and Deployment Services Repository

  This centralized repository is used to host the historical data for predictive analysis. The repository processes data and provides new insights, which can be used to take calculative decisions by organizations. It also integrates the results of analysis and provides customized decisions, which can influence key strategies that are adopted by organizations.

  The repository provides a centralized location for storing analytical assets, such as models and data. The repository includes facilities for these functions:

  - Security
  - Version control
  - Searching
  - Auditing

  The repository requires an installation of a relational database, such as Oracle, IBM DB2 Universal Database (UDB), or Microsoft SQL Server.

  Configuration options for the repository are defined by using the deployment manager or the browser-based deployment manager. The contents of the repository are managed with the deployment manager and accessed with the deployment portal.

  The repository needs to be installed on the servers that host the SPSS solution, in our case, the IBM Power Systems servers.

  Resource: For more information about SPSS Collaboration and Deployment Services features and resources, see the product documentation:

  http://www-01.ibm.com/software/analytics/spss/products/deployment/cds/features.html

- IBM SPSS Collaboration and Deployment Services Deployment Manager (thick client setup)

  This client is used to connect to the repository server and send requests that bear current statistics with an intent to obtain a score and the probability of the outcome about an extrapolated variable.

  It allows users to schedule, automate, and execute analytical tasks, such as updating models or scores, by using the repository. Deployment manager allows a user to perform these tasks:

  - View any existing files within the system, including reports, serial-attached Small Computer System Interface (SCSI) (SAS) syntax files, and data files.
  - Import files into the repository.
  - Schedule jobs to be executed repeatedly by using a specified recurrence pattern, such as quarterly or hourly.
  - Modify existing job properties in a user-friendly interface.
  - Determine the status of a job.
  - Specify email notification of job status.

In addition, deployment manager allows a user to perform administrative tasks for collaboration and deployment services:

- User management.
- Security provider configuration.
- Role and action assignment.
- The client installs on any Microsoft Windows based machine.
- IBM SPSS Collaboration and Deployment Services Deployment Portal.

The deployment portal is the thin client interface for accessing the repository. Unlike the browser-based IBM SPSS Collaboration and Deployment Services Deployment Manager, which is intended for administrators, the deployment portal is a web portal that serves various users. The web portal includes the following functionality:

- Browsing the repository content by folder
- Opening published content
- Running jobs and reports
- Generating scores by using models that are stored in the repository
- Searching repository content
- Viewing content properties
- Accessing individual user preferences, such as email address and password, general options, subscriptions, and options for output file formats

- Browser-based IBM SPSS Collaboration and Deployment Services Deployment Manager

  This Deployment Manager is the easiest way of accessing the IBM SPSS Collaboration and Deployment Services Repository. The user is provided with a web user interface to enter the data and request predictive analysis. The user must log on to the solution and use it like the thick client version.

  The following tasks can be carried out through the browser-based IBM SPSS Collaboration and Deployment Services Deployment Manager:

  - Configuring the system
  - Configuring security providers
  - Managing Multipurpose Internet Mail Extensions (MIME) types

  Non-administrative users can perform any of these tasks. They must have the appropriate actions that are associated with their login credentials. The actions are assigned by an administrator.

- IBM SPSS Collaboration and Deployment Services Enterprise View

  Enterprise View provides a single, consistent view of enterprise data. Enterprise View allows users to define and maintain a common view of warehoused and transaction data that is needed to perform analytics, optimization, deployment, and reporting. Underlying data can come from various sources, including a data warehouse, an operational data store, and an online transaction database. Enterprise View ensures a consistent use of enterprise data and hides the complexities of stored data structures from the user. Enterprise View is the data backbone for the predictive enterprise.

  Data discovery requires a major investment of resources from the organizations that deploy predictive analytics. The process is labor-intensive. It can involve representatives from departments across the organization and often entails resolving differences in data structure and semantics across organizational boundaries. Enterprise View provides a mechanism for recording the outcomes of the data discovery process, versioning and securing the resulting schema, and tracking changes over time.

  Enterprise View includes the Enterprise View Driver component that is designed to provide other applications access to Enterprise View objects that are stored in the repository. The driver operates similarly to Open Database Connectivity (ODBC) drivers. However, it does not directly query a physical data source but rather references Enterprise

View data provider definitions and application views. Although Enterprise View is installed as part of Deployment Manager, Enterprise View Driver must be installed separately.

- Execution servers

  Execution servers can execute resources that are stored in the repository. When a resource is included in a job for execution, the job step definition includes the specification of the execution server that is used for processing the step. The execution server type depends on the resource.

  The following execution servers are currently supported by Predictive Analytics Software (PASW®) Collaboration and Deployment Services:

  - The SAS execution server, which is the SAS executable file `sas.exe` that is included with the Base SAS software. Use this execution server to process SAS syntax files.
  - A remote process execution server, which allows processes to be initiated and monitored on remote servers. When the process completes, it returns a success or failure message. Any machine that acts as a remote process server must have the necessary infrastructure installed for communicating with the repository. Execution servers that process other specific types of resources can be added to the system by installing the appropriate adapters. During job creation, the remote process server assigns an execution server to each step that is included in the job. When the job executes, the repository uses the specified execution servers to perform the corresponding analysis.

- BIRT Report Designer for IBM SPSS

  This component enables the reporting functionality of IBM SPSS Collaboration and Deployment Services, where the user can customize the data, input requests, and view graphical results with varying trends. The user can also create scripts by using this designer and create customized reports.

  The reporting functionality of SPSS Collaboration and Deployment Services is enabled by Business Intelligence and Reporting Tools (BIRT), which is an open source package that is distributed by the Eclipse Foundation under the Eclipse Public License. BIRT provides core reporting features, such as report layout, data access, and scripting.

  For more information about BIRT, see the BIRT project page:

  `http://eclipse.org/birt`

The SPSS Collaboration and Deployment Services installation includes the BIRT reporting engine server components, which enable the execution of BIRT report syntax files as part of the SPSS Collaboration and Deployment Services reporting job steps. SPSS BIRT Report Designer is a stand-alone application that can be used with SPSS Collaboration and Deployment Services. It provides a rich user interface with many advanced features for creating reports and must be installed separately.

If an SPSS BIRT Report Designer report requires a Java Database Connectivity (JDBC)-based database connection, a corresponding JDBC driver must be installed with the repository. For application server-specific information about the location of the JDBC drivers, see the corresponding section of the repository installation instructions.

To start SPSS BIRT Report Designer, execute the file `BIRT.exe` in the installation directory. For information about using PASW BIRT Report Designer, see the documentation that is installed with the application.

Installation and configuration of the SPSS repository

In this section, we describe the installation procedure for the SPSS Collaboration and Deployment Services Repository.

Important: We suggest that you look at the SPSS product documentation before you proceed with the installation. Use this link:

http://www-01.ibm.com/software/analytics/spss/

Details of the hardware that was used in the SPSS solution installation

Table 3-1 consists of the hardware details that we used to install the SPSS solution.

Table 3-1 Hardware details that we used in the installation of the SPSS solution

Component	Requirement
Processor	Power System 750 with 2 POWER7 8-core 3.3-GHz processors
Hard drive	Minimum of 10 GB space
Memory	Minimum of 4 GB RAM

Details of the operating system that was used in the SPSS solution installation

Table 3-2 shows the details of the operating system that was used to install the SPSS solution.

Table 3-2 Details of the operating system that was used in the installation of the SPSS solution

Component name	Operating system	Version	Processor type	Word size
IBM SPSS Collaboration and Deployment Services Repository (Server)	AIX	7.1 Technology Level 1 Service Pack 4	POWER7	64 bit
IBM SPSS Collaboration and Deployment Services Deployment Manager (Client)	Microsoft Windows	2003 Enterprise Edition R2	Intel Xeon X7560	32 bit

Details of the application server that was used in the SPSS solution installation

Table 3-3 shows the details of the application server that was used in the installation of the IBM SPSS solution.

Table 3-3 Details of the application server that was used in the SPSS solution installation

Application server name	Version	Java Development Kit (JDK) version
WebSphere Application Server	7.0.0.23	1.6 SR6

Details of the database server that was used in the SPSS solution installation

Table 3-4 shows the details of the database server that was used in the installation of the IBM SPSS solution.

Table 3-4 Details of the database server that was used in the installation of the SPSS solution

Database server name	Version
IBM DB2 Enterprise Edition	9.7

Other requirements

The following products are required for the environment:

- Any X Windows server system utility, such as the Virtual Network Computing (VNC) server or the XManager to export the display from the AIX operating system into the terminal or remote desktop of the user.
- Mozilla Firefox must be installed on AIX. It is required by the WebSphere Application Server installer.

> **Requirements:** For more hardware and software requirements for installing the IBM SPSS Collaboration and Deployment Services solution, see this document:
>
> ftp://ftp.software.ibm.com/software/analytics/spss/documentation/cds/4.2.1/en/RepositoryConfigUNIX.pdf

Deployment topology

Figure 3-2 shows the deployment topology for a single LPAR implementation for the SPSS solution.

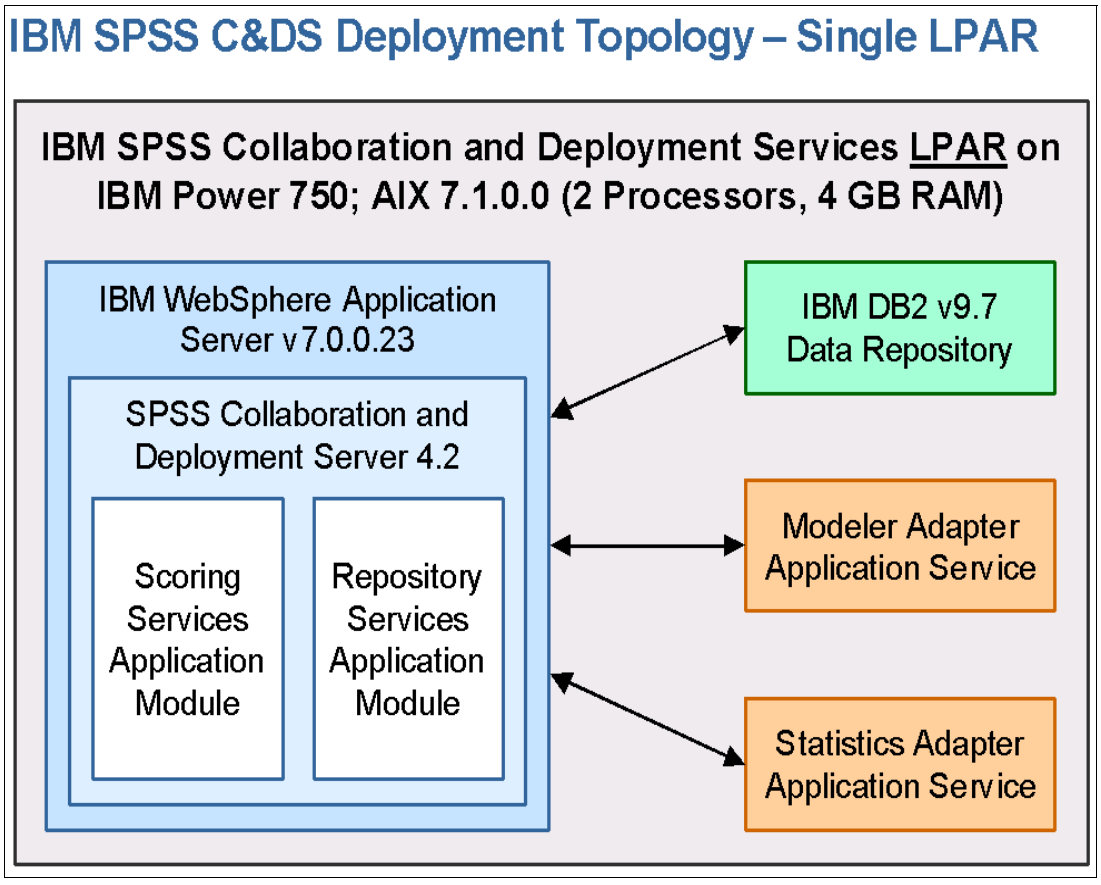

Figure 3-2 Deployment topology for a single LPAR installation of the SPSS solution

Installation procedure

The installation of the SPSS C&DS solution is divided into sections:

- WebSphere Application Server Network Deployment Version 7.0.0.23
- DB2 Database Version 9.7
- SPSS C&DS Version 4.2.1

- SPSS adapter suite for C&DS 4.2.1:
  - Statistics adapter
  - Modeler adapter
  - Decision management adapter

WebSphere Application Server Network Deployment 7 installation

We do not cover WebSphere Application Server Network Deployment Version 7 in detail. For a detailed installation guide, see the link that is mentioned in the footnote[1]:

http://publib.boulder.ibm.com/infocenter/ieduasst/v1r1m0/topic/com.ibm.iea.was_v7/was/7.0/InstallationAndMigration/WASv7_InstallationLab.pdf

Consider the following important information when you install WebSphere Application Server Network Deployment V7 for the IBM SPSS C&DS solution.

Do not enable administrative security on the WebSphere Application Server. It is not required for the C&DS installation. Figure 3-3 shows where to disable this option.

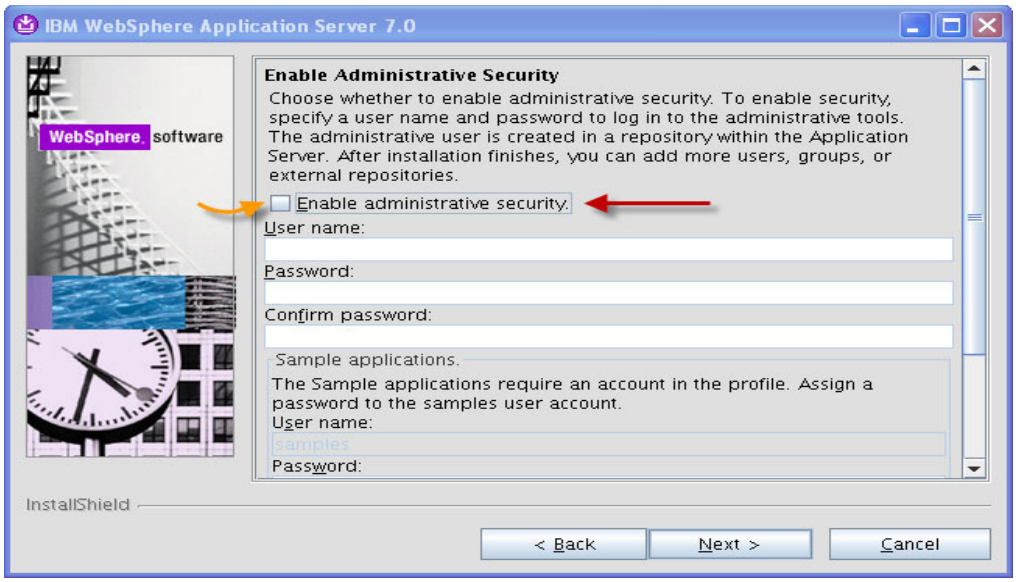

Figure 3-3 Disable WebSphere Application Server administrative security

Do not create a centralized repository for the centralized installation managers. Clear the option to create a centralized repository for the centralized installation manager, as shown in Figure 3-4 on page 87.

[1] *WebSphere Application Server Network Deployment Version 7 Installation Guide*: http://publib.boulder.ibm.com/infocenter/ieduasst/v1r1m0/topic/com.ibm.iea.was_v7/was/7.0/InstallationAndMigration/WASv7_InstallationLab.pdf

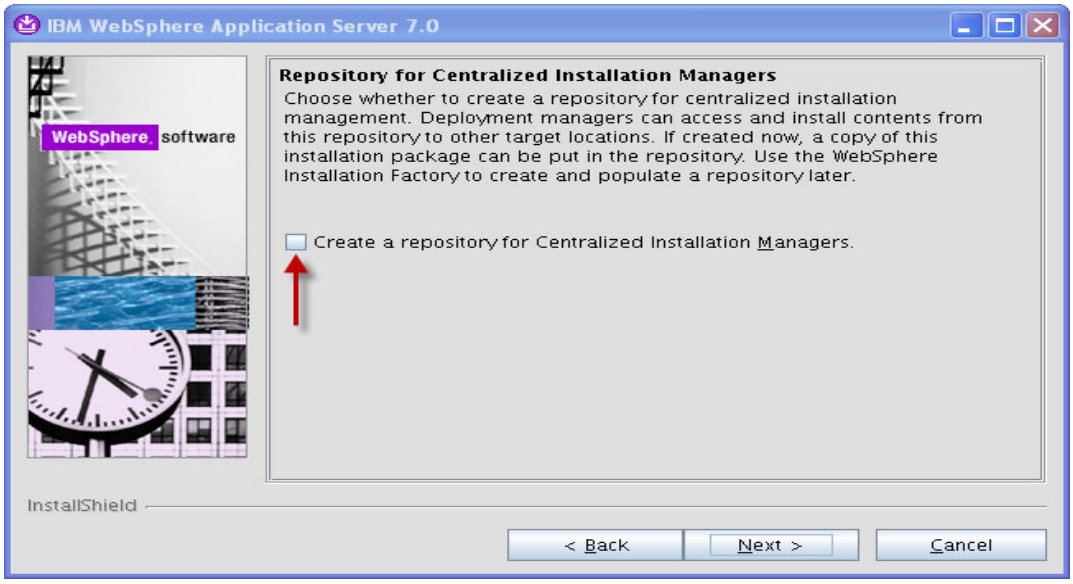

Figure 3-4 Clear repository for centralized installation managers option

Do not create the application server in a cluster configuration. Install the application server only in a stand-alone configuration.

Verify the installation summary, and click **Next** to complete the WebSphere Application Server Network Deployment Version 7 installation.

We suggest that you update WebSphere Application Server Version 7.0 to Version 7.0.0.23. And, we suggest that you update the Java Software Development Kit (SDK) for WebSphere Application Server to Fix Pack 7 before you install the SPSS C&DS solution.

To update WebSphere Application Server Version 7.0, download WebSphere Application Server 7 Fix Pack 23 at this website:

`http://www-304.ibm.com/support/docview.wss?uid=swg24032493`

To download the Java SDK Fix Pack 7, see this website:

`http://www-01.ibm.com/support/docview.wss?rs=180&uid=swg27014463#7007`

Important: The update installer software is required to update the WebSphere Application Server. You can download this software at this website:

`http://www-01.ibm.com/support/docview.wss?rs=180&uid=swg24020212`

Installing DB2 Version 9.7

We do not describe the detailed installation steps for DB2 Version 9.7. For the DB2 detailed installation guide, see this website:

`http://publib.boulder.ibm.com/infocenter/db2luw/v9r7/index.jsp?topic=%2Fcom.ibm.db2.luw.qb.server.doc%2Fdoc%2Ft0008099.html`

We describe the key considerations and information to enter as input during the IBM DB2 installation for the SPSS C&DS solution:

- Select **Default** on the SAMP window.
- Provide the default options for the db2 instance user with `db2inst1` and *`password`*; `db2iadm1` is the group user ID.

- Provide the default options for db2 fenc user with `db2fenc1` as the user ID and *`password`*.
- Use a single partition instance.
- Do not set up any tools catalog.

Verification: When you select the default option, you see a warning or an error message. This message can be ignored. The user can verify whether DB2 is installed correctly by running the **`db2val`** command after the installation. The following sample shows the output of the command:

```
$ db2val
DBI1379I  The db2val command is running. This can take several minutes.
DBI1335I  Installation file validation for the DB2 copy installed at
      /usr/IBM/db2/V9.7 was successful.
DBI1339I  The instance validation for the instance db2inst1 was
      successful.
DBI1343I  The db2val command completed successfully. For details, see
      the log file /tmp/db2val-06_07_16:02:36.log.
```

Setting up the database and tables

The SPSS C&DS solution requires that the database tables are created before the installation. These tables are used during the setup by the SPSS installer to populate configuration data, and also to set up the database as the data source with the WebSphere Application Server.

Save the SPSS database SQL statements as `CreateSPSSDB_AIX.sql`, as shown in Example 3-1.

Database: The name of the database in this script is `SSKTEST`.

Example 3-1 DB2 setup of SQL scripts for the SPSS solution

```
CREATE DATABASE SSKTEST ON '/home/db2inst1' USING CODESET UTF-8 TERRITORY US
COLLATE USING SYSTEM;
CONNECT TO SSKTEST;
CREATE Bufferpool SPSS8K IMMEDIATE SIZE 250 AUTOMATIC PAGESIZE 8 K ;
CREATE REGULAR TABLESPACE SPSS8K PAGESIZE 8 K MANAGED BY AUTOMATIC STORAGE
EXTENTSIZE 8 OVERHEAD 10.5 PREFETCHSIZE 8 TRANSFERRATE 0.14 BUFFERPOOL SPSS8K
DROPPED TABLE RECOVERY ON;
COMMENT ON TABLESPACE SPSS8K IS '';
CREATE Bufferpool SPSSTEMP IMMEDIATE SIZE 250 PAGESIZE 32 K ;
CREATE SYSTEM TEMPORARY TABLESPACE SPSSTEMP PAGESIZE 32 K MANAGED BY AUTOMATIC
STORAGE EXTENTSIZE 16 OVERHEAD 10.5 PREFETCHSIZE 16 TRANSFERRATE 0.14 BUFFERPOOL
"SPSSTEMP";
COMMENT ON TABLESPACE SPSSTEMP IS '';
CONNECT RESET;
```

Log in as `db2inst1` and create the `SSKTEST` database by using the command that is shown in Example 3-2.

Example 3-2 SPSS DB2 SQL execution output

```
$ db2 -tvf ./CreateSPSSDB_AIX.sql
Below is the output of the command after execution:
$ db2 -tvf ./CreateSPSSDB_AIX.sql
```

```
CREATE DATABASE SSKTEST ON '/home/db2inst1' USING CODESET UTF-8 TERRITORY US
COLLATE USING SYSTEM
DB20000I  The CREATE DATABASE command completed successfully.

CONNECT TO SSKTEST

   Database Connection Information

 Database server        = DB2/AIX64 9.7.0
 SQL authorization ID   = DB2INST1
 Local database alias   = SSKTEST


CREATE Bufferpool SPSS8K IMMEDIATE SIZE 250 AUTOMATIC PAGESIZE 8 K
DB20000I  The SQL command completed successfully.

CREATE REGULAR TABLESPACE SPSS8K PAGESIZE 8 K MANAGED BY AUTOMATIC STORAGE
EXTENTSIZE 8 OVERHEAD 10.5 PREFETCHSIZE 8 TRANSFERRATE 0.14 BUFFERPOOL SPSS8K
DROPPED TABLE RECOVERY ON
DB20000I  The SQL command completed successfully.

COMMENT ON TABLESPACE SPSS8K IS ''
DB20000I  The SQL command completed successfully.

CREATE Bufferpool SPSSTEMP IMMEDIATE SIZE 250 PAGESIZE 32 K
DB20000I  The SQL command completed successfully.

CREATE SYSTEM TEMPORARY TABLESPACE SPSSTEMP PAGESIZE 32 K MANAGED BY AUTOMATIC
STORAGE EXTENTSIZE 16 OVERHEAD 10.5 PREFETCHSIZE 16 TRANSFERRATE 0.14 BUFFERPOOL
"SPSSTEMP"
DB20000I  The SQL command completed successfully.

COMMENT ON TABLESPACE SPSSTEMP IS ''
DB20000I  The SQL command completed successfully.

CONNECT RESET
DB20000I  The SQL command completed successfully.

$
```

Obtain the values of the WebSphere variables, as shown in Table 3-5 on page 90.

Table 3-5 WebSphere variables for SPSS C&DS installation

Variable	Value that we used	Remarks
Java Runtime Environment (JRE) location	*/<WebSphere_install_root>/AppServer/java/jre*	JRE location must be same as the location that the WebSphere Application Server V7.0.0.23 uses.
WebSphere HTTP port	9081	Check this port in the Port listing in the Administration console under **Servers** → **Application servers** → *<server_name>* → **Ports** → *WC_defaulthost*
SOAP Connector port	8881	Check this port in the Port listing in the Administration console under **Servers** → **Application servers** → **Server1 Ports** → *SOAP_CONNECTOR_ADDRESS*
WebSphere Server name	Server1	Check this name in the Administration console under **Servers** → **Application Servers**
WebSphere Application Server node name	spss02Node01	Check this name in the Administration console under **Servers** → **Nodes**
WebSphere Application Server cell name	spss02CellManager01	Check this name in the Administration console under **Servers** → **Nodes**
Database server name	spss02	Run the **hostname** command in the DB2 server to discover this value
Database port	50000	Assigned by default during the DB2 installation
Database user ID	db2inst1	DB2 instance user ID; the default is db2inst1
Database user ID password	*******	Password of DB2 instance user

Installing SPSS C&DS

We describe the detailed installation steps for SPSS C&DS Version 4.2.1:

1. Verify that the WebSphere Application Server instance and the DB2 instance are active and running.
2. Mount the ISO image on the AIX file system and run the `install.bin` file in the path:

   <mount_point_loc>/Server/Disk1/InstData/AIX/NoVM

   See Example 3-3 on page 91.

Example 3-3 Output of the installation binary (install.bin)

```
root@spss02:/stage/spss/Disk1/InstData/AIX/NoVM # ./install.bin
Preparing to install...
Extracting the installation resources from the installer archive...
Configuring the installer for this system's environment...

Launching installer...
```

3. The initial window opens, as shown in Figure 3-5.

Figure 3-5 Initial window of SPSS C&DS installer

4. Click **OK** and proceed to see the introduction window, as shown in Figure 3-6.

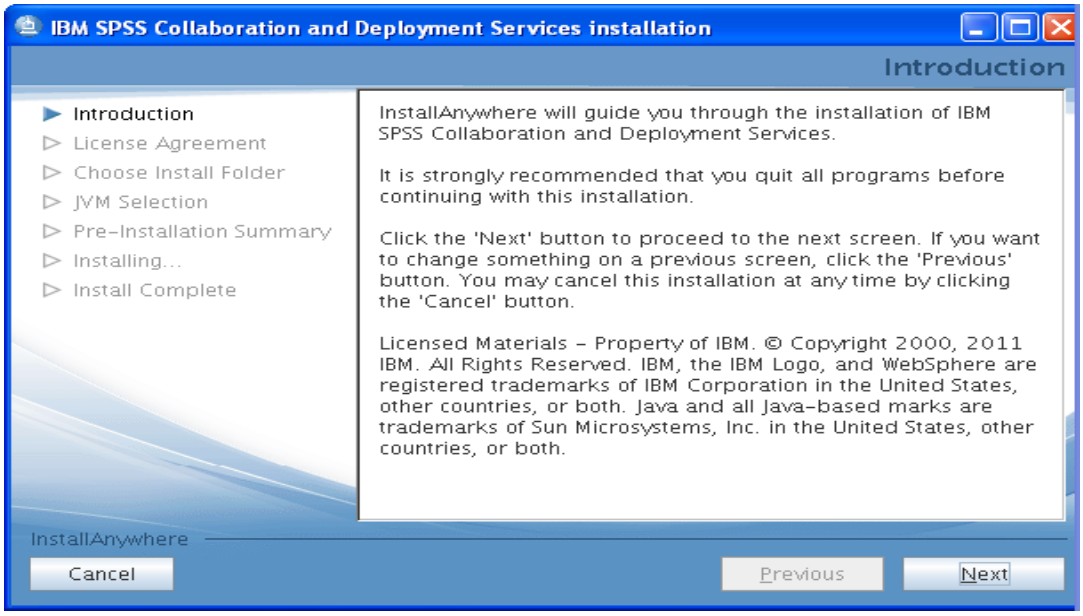

Figure 3-6 Introduction window of SPSS C&DS installer

5. Click **Next** and accept the licensing agreement, as shown in Figure 3-7 on page 92.

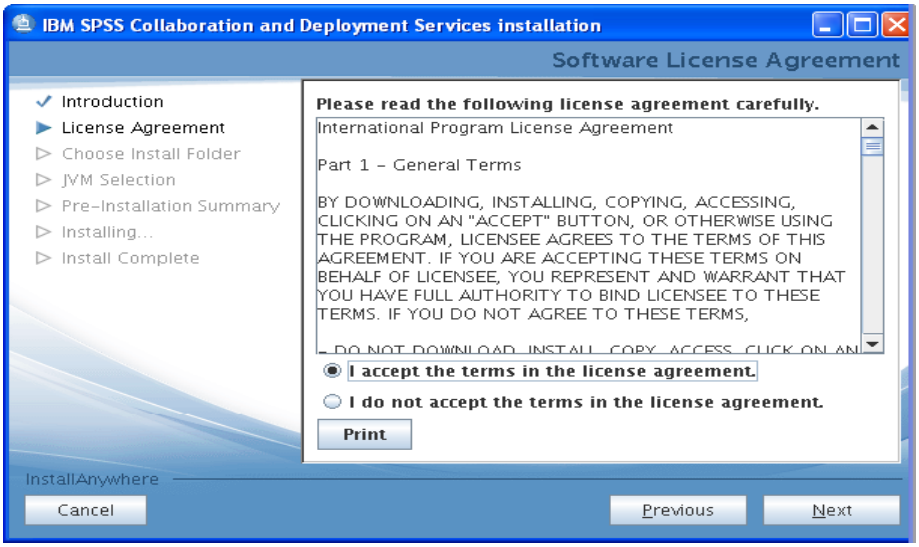

Figure 3-7 Accept license agreement

6. Choose the installation location. Choose the Java virtual machine (JVM) location, as shown in Figure 3-8. Use the WebSphere variables that are shown in Table 3-5 on page 90.

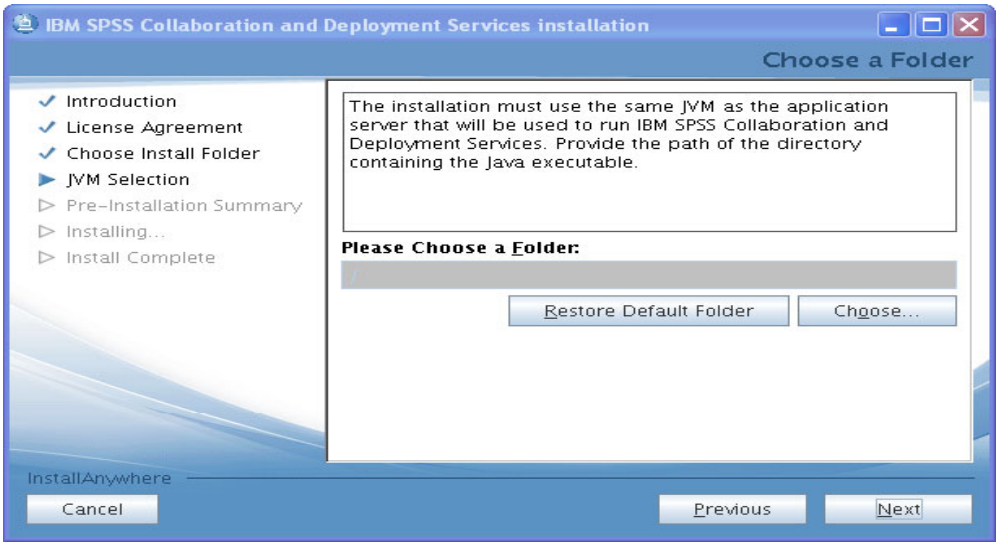

Figure 3-8 WebSphere Application Server JVM installed location

7. Review the pre-installation summary information and click **Install**, as shown in Figure 3-9 on page 93.

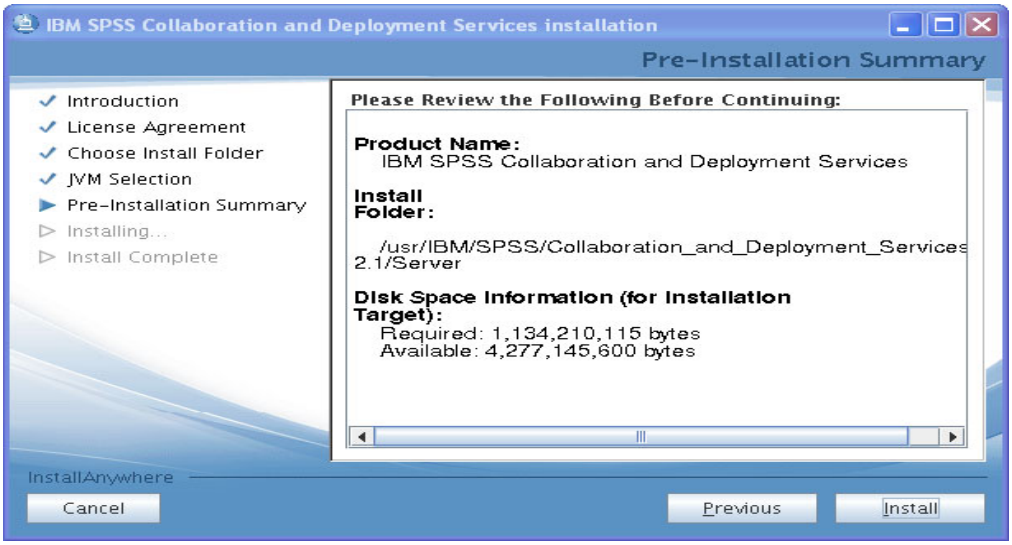

Figure 3-9 Pre-installation summary

8. Figure 3-10 shows the installation and extraction of the packages.

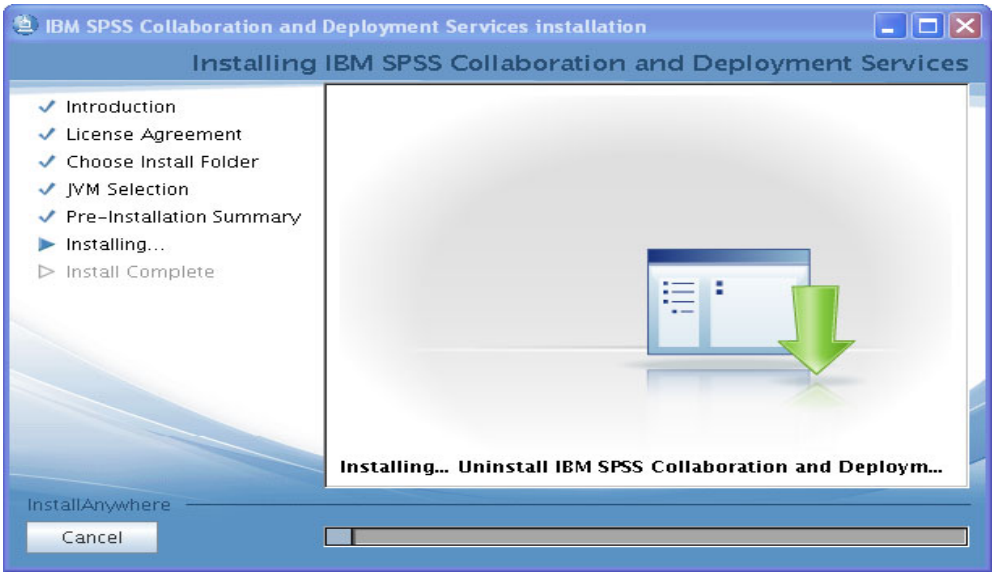

Figure 3-10 Installation and extraction of packages in progress

9. After the installation and extraction completes, go to the directory `/<spss_install_root>/setup/`. Run the `setup.sh` file from this location. The SPSS C&DS solution installer window opens, as shown Figure 3-11 on page 94. Click **Next**.

Figure 3-11 SPSS C&DS Setup Tool

10. Enter the keystore location and password, and click **Next**. See Figure 3-12.

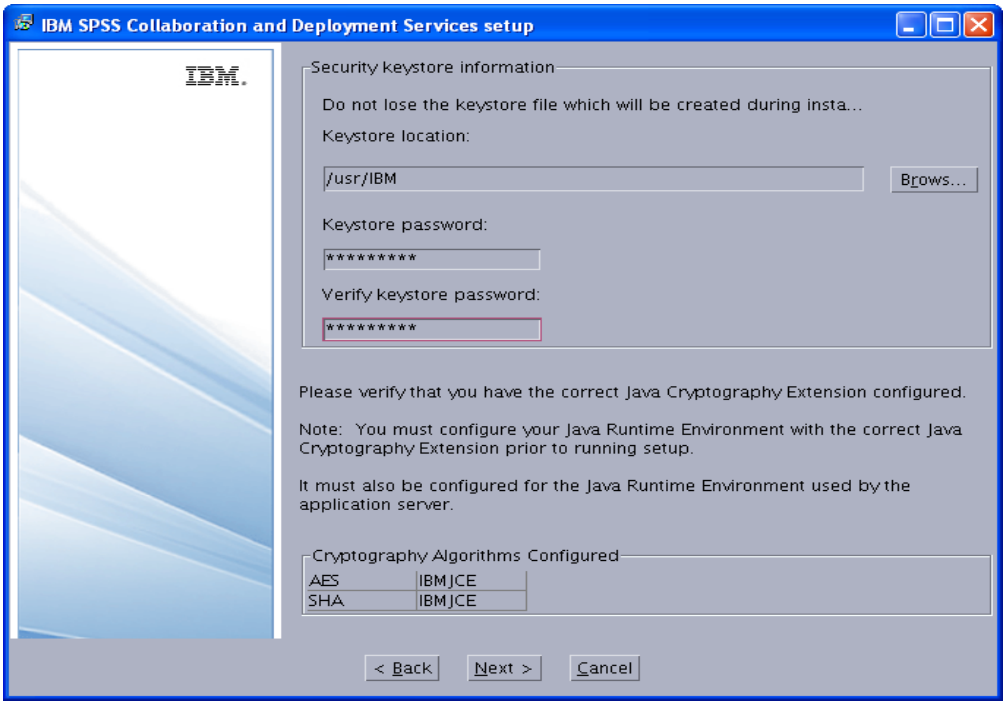

Figure 3-12 Keystore location and password

11. Choose the application server environment type, as shown in Figure 3-13 on page 95. We selected **Standard Standalone Application Server (Recommended)**.

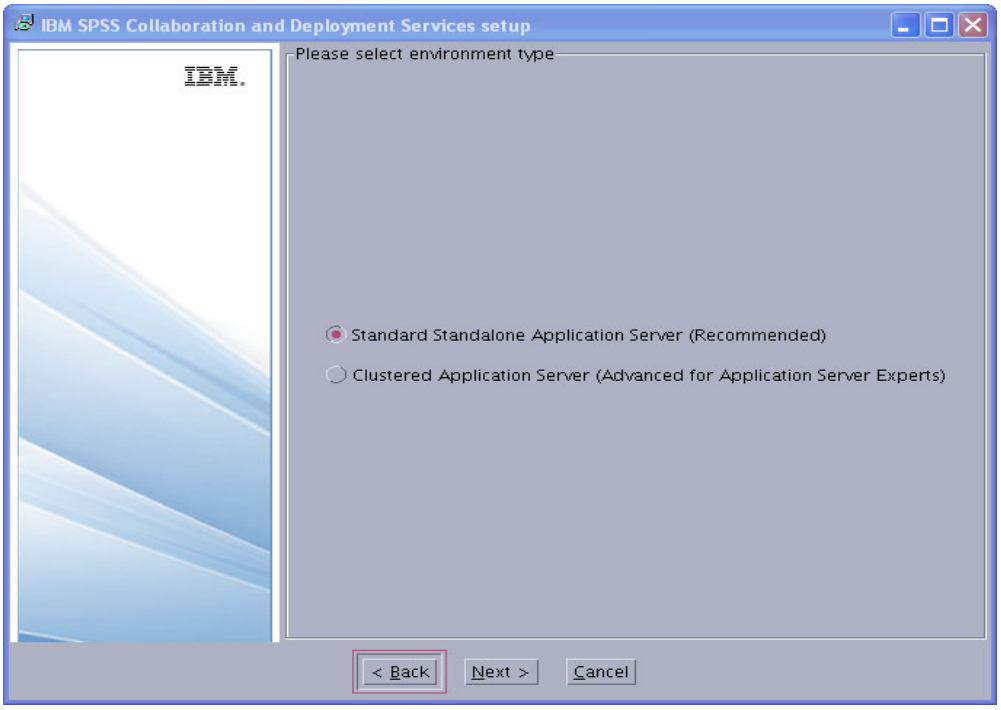

Figure 3-13 Choose the standard stand-alone server type

12. Enter the WebSphere Application Server parameters, as shown in Figure 3-14. Use Table 3-5 on page 90.

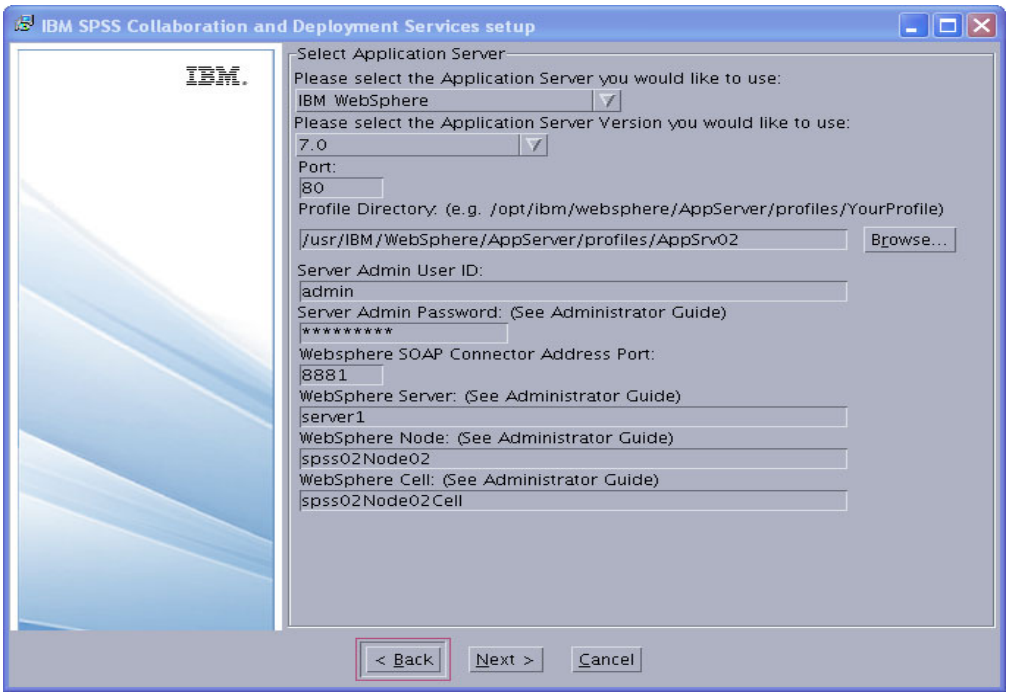

Figure 3-14 WebSphere Application Server parameters

13. Provide the database name, as shown in Figure 3-15 on page 96.

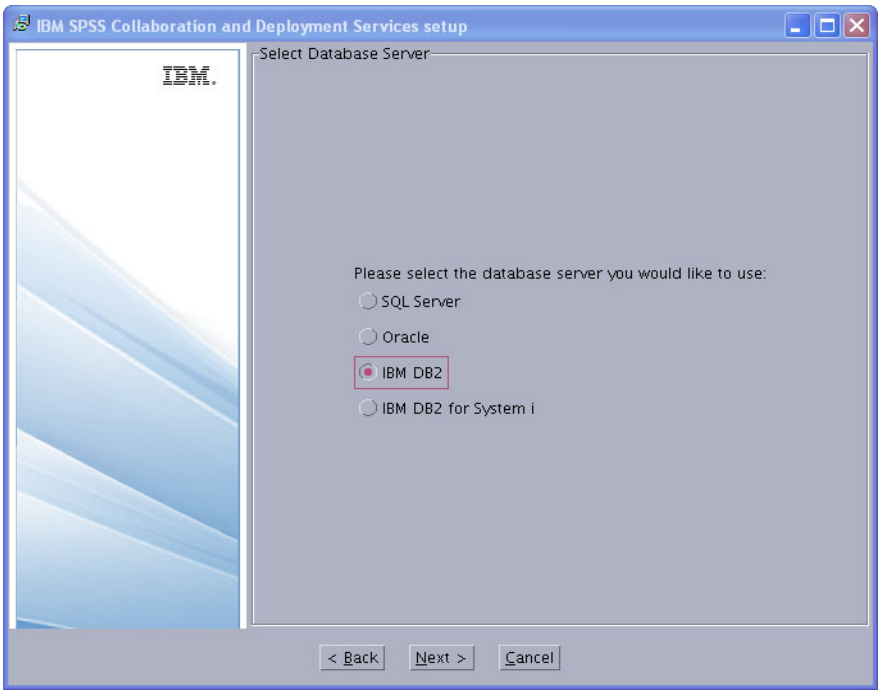

Figure 3-15 IBM DB2 database

14. Provide the DB2 parameters (Figure 3-16). Use Table 3-5 on page 90.

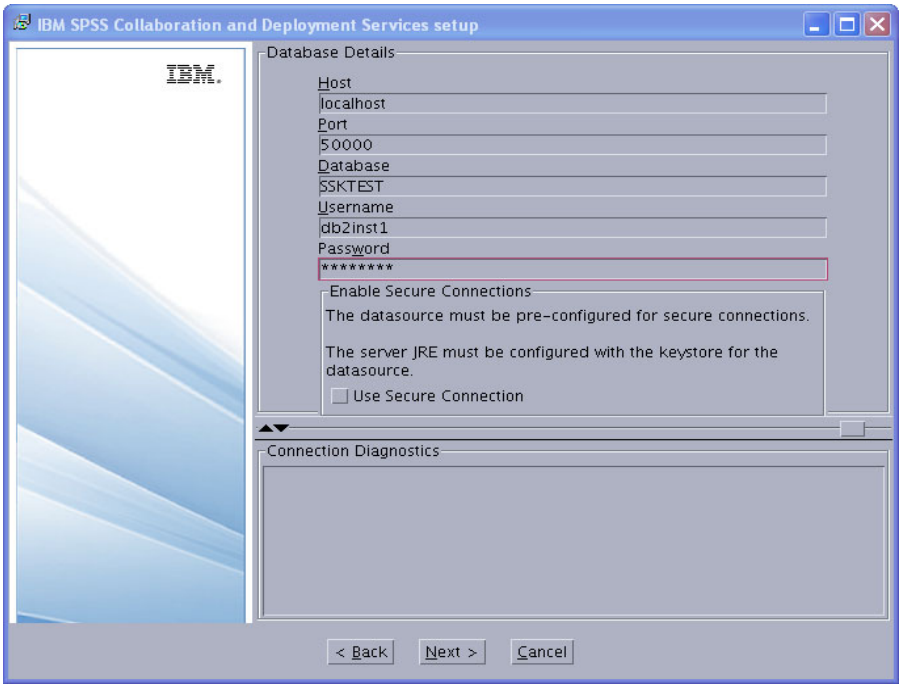

Figure 3-16 IBM DB2 parameters

15. Provide the security password for the C&DS administration, as shown in Figure 3-17 on page 97.

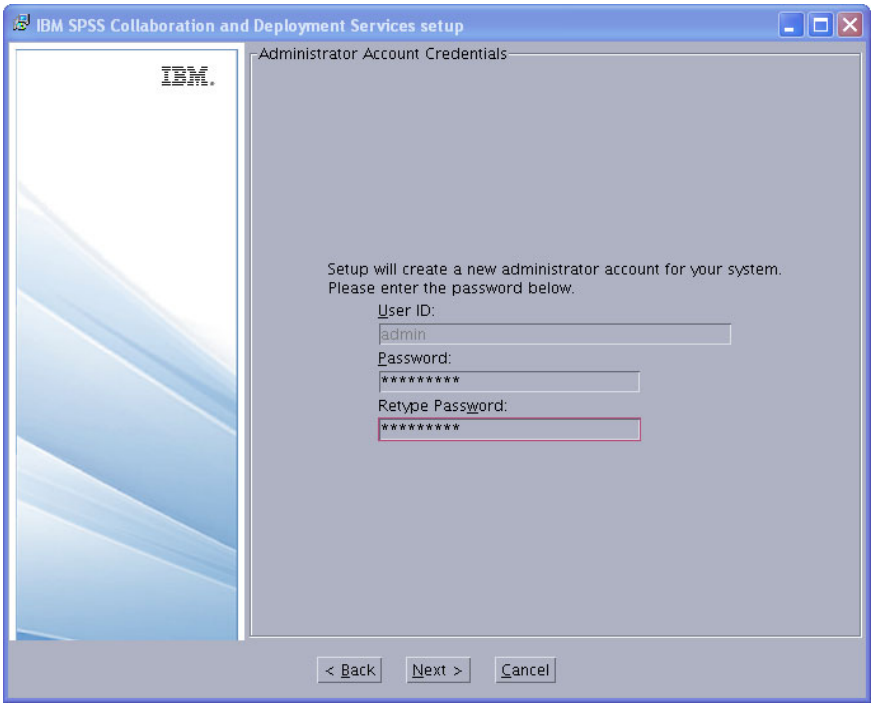

Figure 3-17 Security password for C&DS administration

16. Enter the contact details, as shown in Figure 3-18.

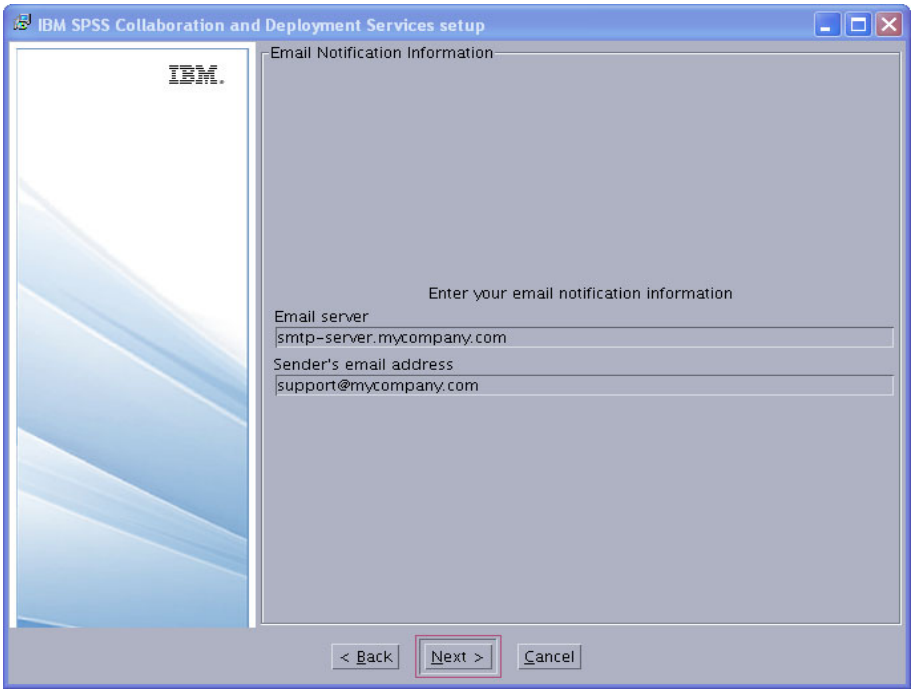

Figure 3-18 Contact details

17. Click **Next** to begin the installation, as shown in Figure 3-19 on page 98.

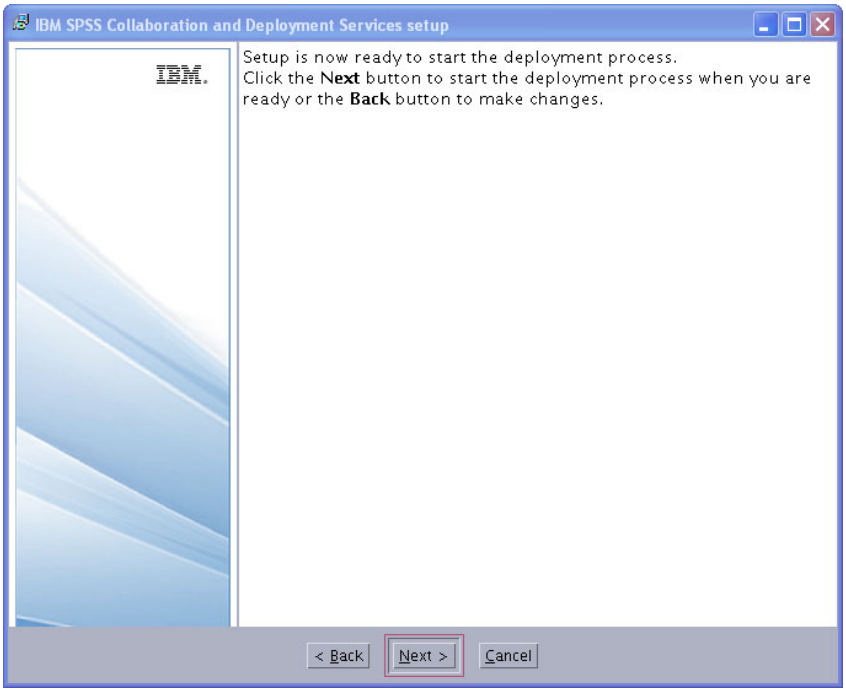

Figure 3-19 Ready for the installation

18. When the setup completes without errors, click **Finish**, as shown in Figure 3-20.

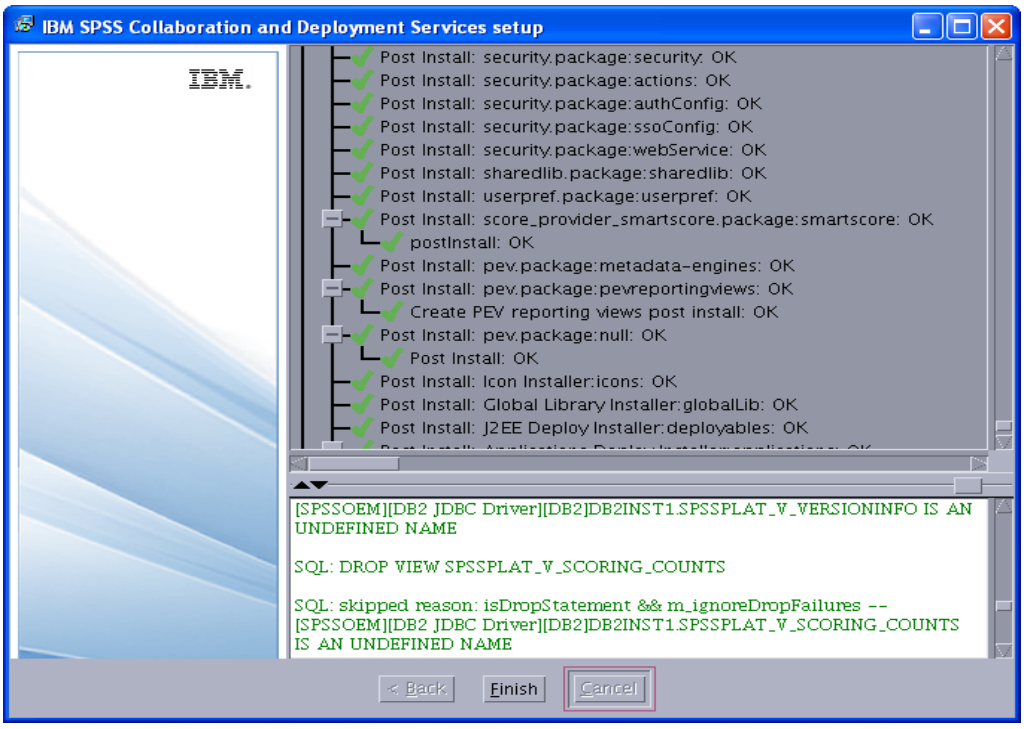

Figure 3-20 Setup complete

Installing the SPSS Statistics adapter

This section describes in detail the installation of the SPSS Statistics adapter for the C&DS solution.

Run the `SPSS_CnDS_42_StatisticsServerAdapter_19_Aix64.bin`. Use the same process as described in this section to install the Statistics adapter on SPSS.

Installing the modeler adapter and decision management adapter

To install the modeler adapter and the decision management adapter, we use the package manager software to install the modeler and the decision management packages.

Extract the modeler and decision management packages:

- `ModelerAdapters_14_20_014_21.zip`
- `DM_7_0_0_0_442_GA.zip`

The following packages are available in the modeler zip file:

- `ScenarioPlugin_14_2_2.package`
- `ModelerAdapters_14_2_2.package`
- `ModelerBase_14_2_2.package`

The following packages are available in the decision management zip file:

- `DecisionManagement_7_0_0.package`
- `CampaignOptimization_7_0_0.package`
- `ClaimsManagement_7_0_0.package`
- `CustomerInteractionManagement_7_0_0.package`
- `ModelerAdvantage_7_0_0.package`
- `RulesManagement_7_0_0.package`
- `DM_ModelerExt_aix64_7_0_0.packageDM_Redirect_7_0_0.package`

Installing by using the package manager

This section shows the installation by using the package manager:

1. Go to the path:

   `<`*`spss_install_root`*`>//opt/IBM/SPSS/Collaboration_and_Deployment_Services/4.2.1/Server/setup/`

2. Start the file `packagemanager.sh`. The window in Figure 3-21 on page 100 opens.

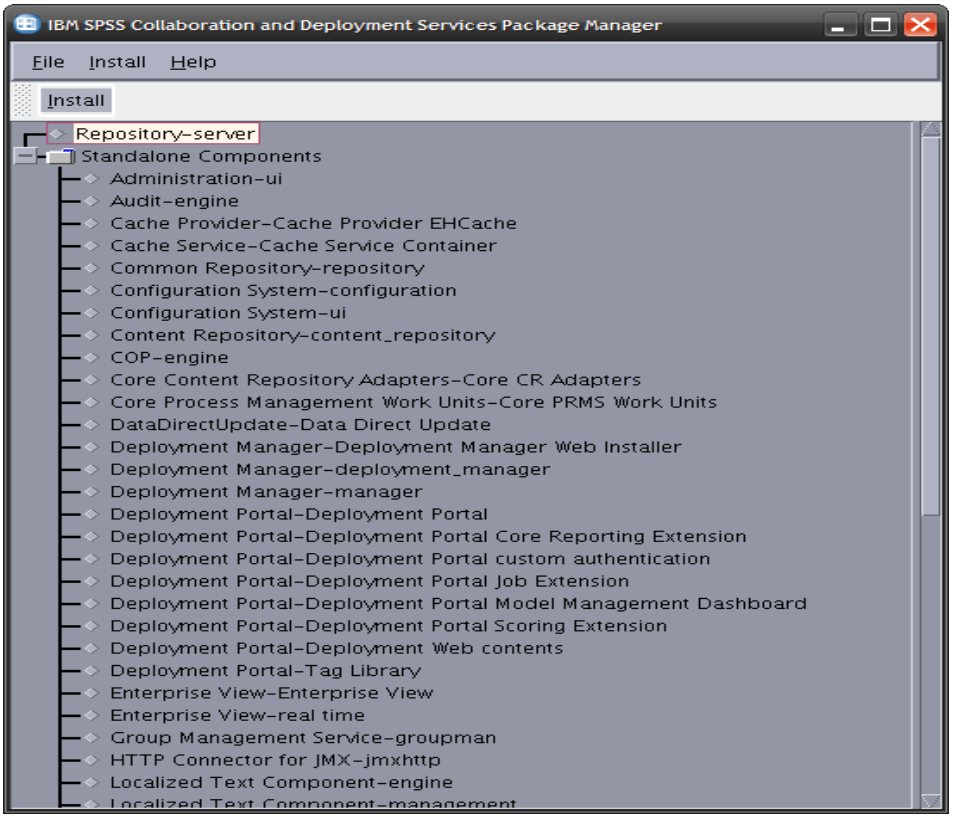

Figure 3-21 Package manager window

3. Click **Install** and browse to the location of the packages, as shown in Figure 3-22 on page 101. Select multiple packages to install by holding the Ctrl key.

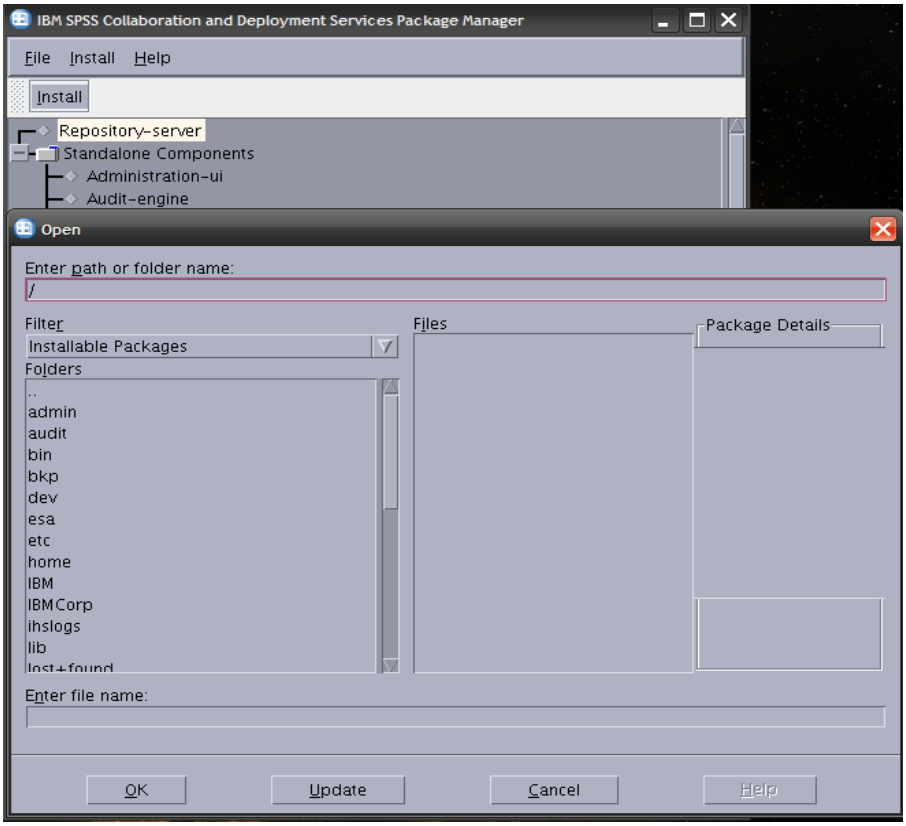

Figure 3-22 Browse to the location and select the packages to install

4. The selected packages to install are shown in Figure 3-23. Click **OK**.

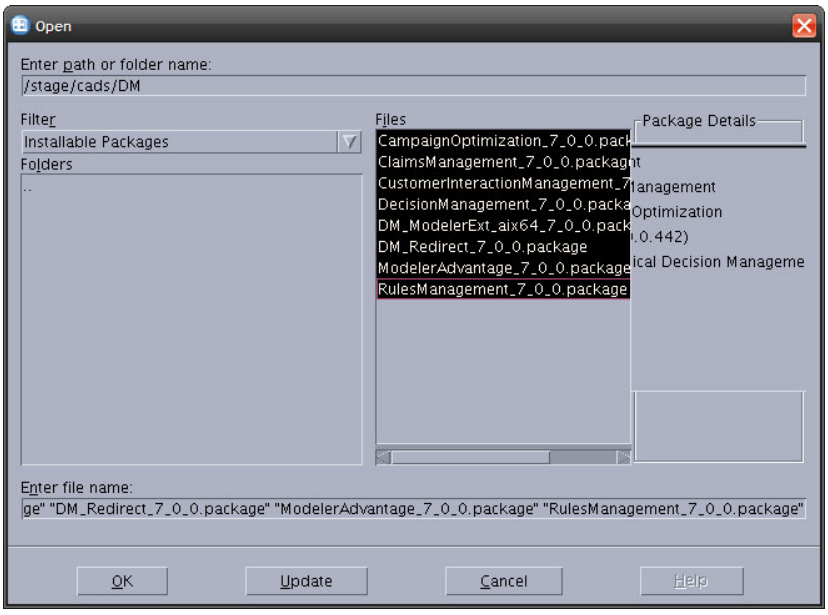

Figure 3-23 Selecting multiple packages to install

5. Click **Finish**, as shown in Figure 3-24 on page 102.

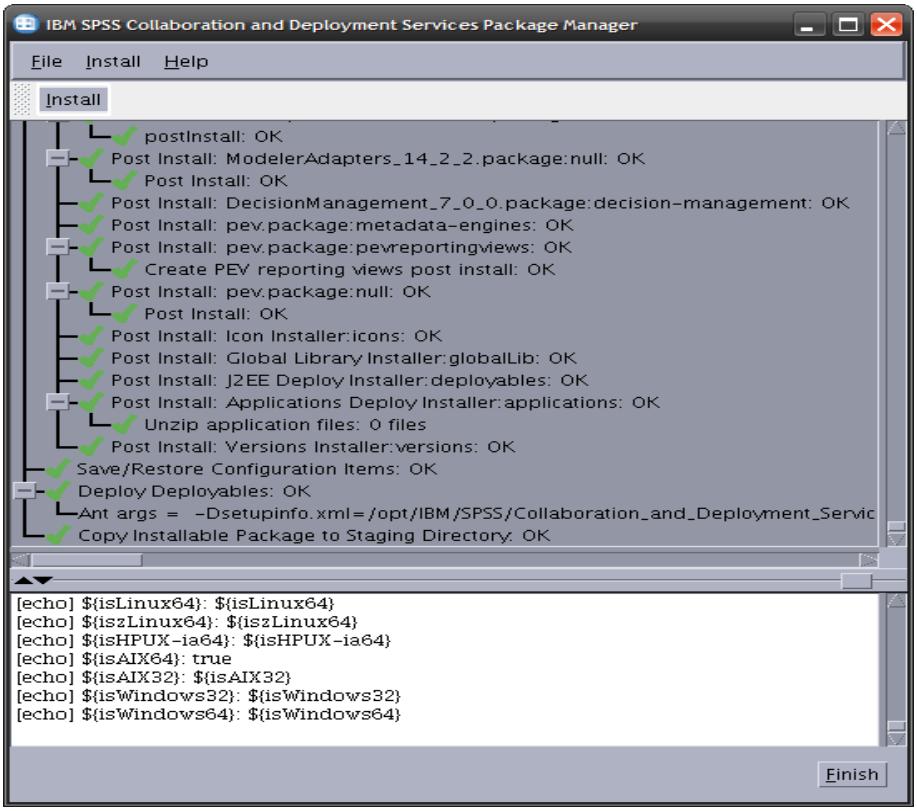

Figure 3-24 Click Finish upon completion

Installing the SPSS C&DS Deployment Manager thick client

We document the installation procedure of the IBM SPSS C&DS Deployment Manager that is used to configure and access the solution from remote desktops:

1. Download the file `SPSS_CnDS_421_DeploymentMgr_win64(32).exe` and run it. Figure 3-25 shows the SPSS C&DS Deployment Manager setup.

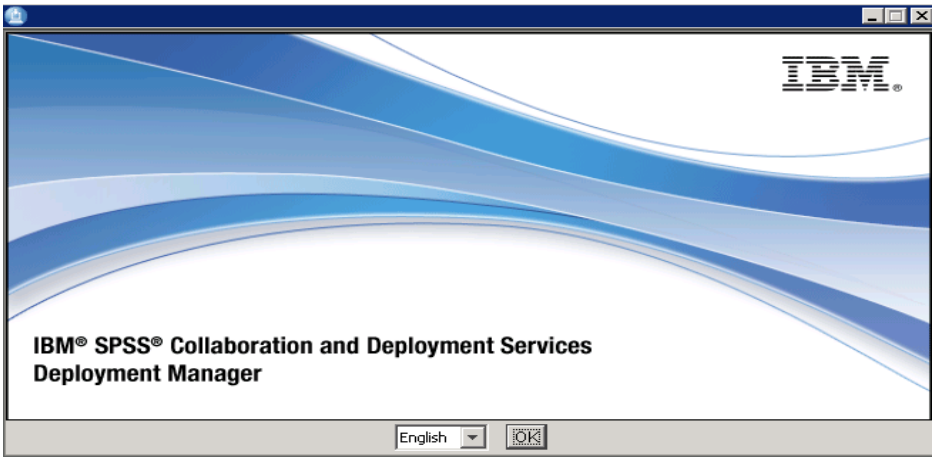

Figure 3-25 SPSS C&DS Deployment Manager setup

2. Figure 3-26 on page 103 shows the license agreement window.

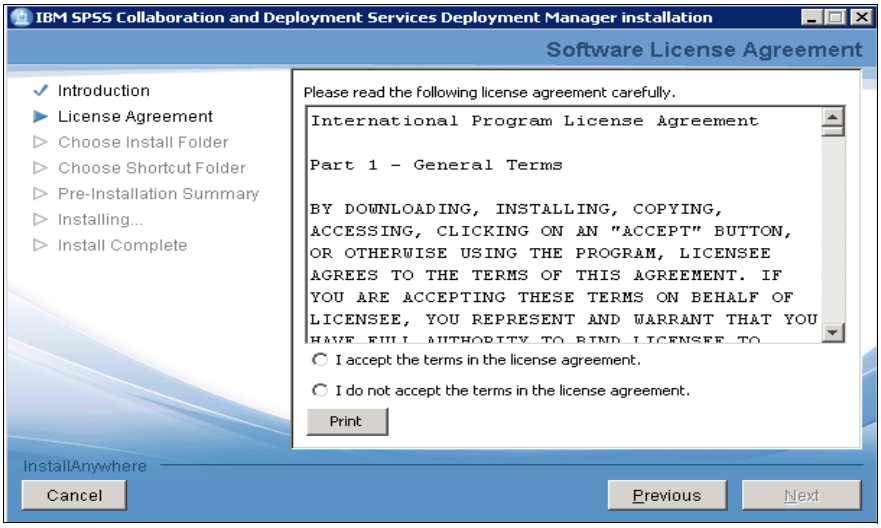

Figure 3-26 License agreement window

3. Figure 3-27 shows the introduction window.

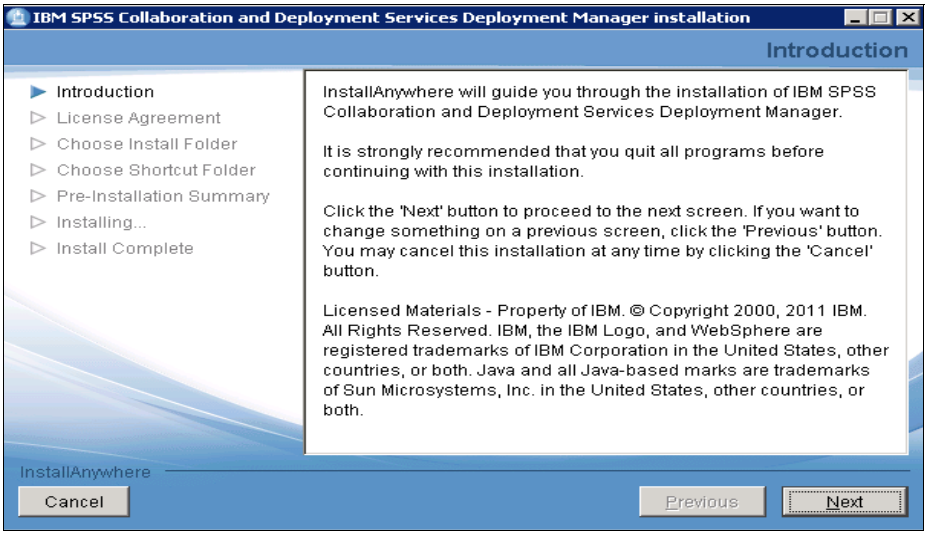

Figure 3-27 Introduction window

4. Figure 3-28 on page 104 shows the window to select the software installation location.

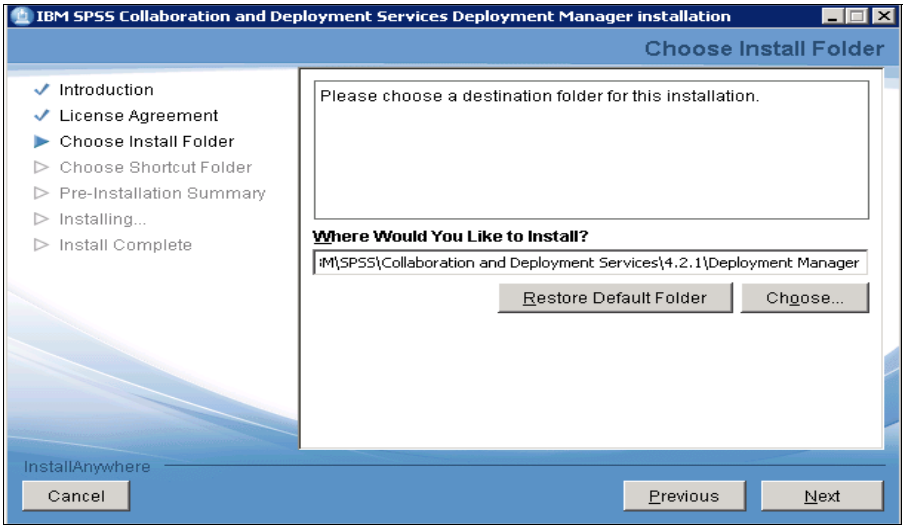

Figure 3-28 Choose your installation location

5. Figure 3-29 shows the installation completion window. Click **Done**.

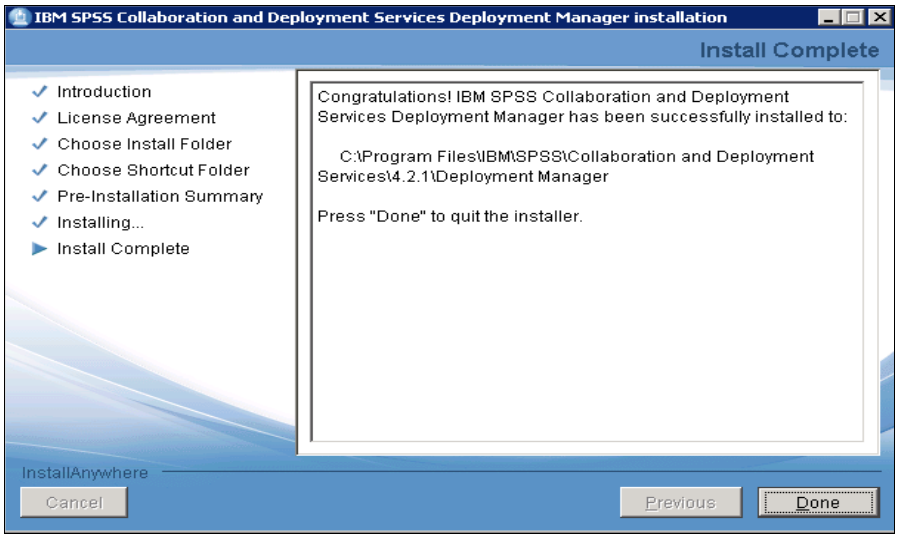

Figure 3-29 Completing the installation

The installation of the SPSS suite of products is complete.

3.1.2 SPSS performance in a single WebSphere Application Server instance

We describe the performance of the SPSS C&DS solution without any performance tuning of the solution.

We then tuned the solution for optimal performance until we achieved the best settings for our workloads under test. After we achieved the best performance of the solution, we compared our settings and describe this comparison in a step-by-step tuning approach.

Our goal in this section is to optimize the SPSS C&DS solution to obtain the maximum transaction throughput rate (scores). The real-time scoring service enables the delivery of on-demand scores to applications, making it easy to deploy analytical scores within business processes.

Scoring models, such as a model that generates a credit risk score that is based on customer data, can be created within IBM SPSS Modeler, and then stored in the platform repository. An outside application can then pass input variables to the web services interface and receive a score in real time.

During the performance analysis, we experimented with various test scenarios to achieve maximum scores, identify tunable parameters, identify scoring bottlenecks, and work to remove the bottlenecks.

Monitoring tools that are used during performance test execution

We used the following tools to monitor the SPSS C&DS solution during the execution of the performance test. We identified performance bottlenecks and conducted in-depth analysis of the performance issues in the solution:

- NMON is a monitoring tool that is used for monitoring the operating system resource usage. This tool is used for monitoring AIX operating system behavior and resource usage. This online and offline monitoring tool provides real-time operating system utilization statistics. It also logs them in an output file to provide historical data to plot charts for easier analysis.

  The NMON tool is part of AIX Version 6.1, and later. For more information about NMON, see this website:

  `http://www.ibm.com/developerworks/aix/library/au-analyze_aix/`

- The **`tprof`** command reports processor usage for individual programs and the entire system. This command is a useful tool for Java, C, C++, or Fortran programs that might be processor-bound and for situations where the user wants to know which sections of the program use the processor the most.
- The IBM thread and monitor dump analyzer for Java tool is used to help analyze Java `java.core.txt` thread and monitor dump files. This tool is used to identify deadlocks and lock contentions within the application threads at run time.
- Perfmon is an application for monitoring Windows operating system performance. It measures the operating system resource utilization, and it provides real-time graphical data that relates to the utilization.
- Tivoli performance viewer is an online monitoring utility that is built into the WebSphere Application Server Administration console. It is used to monitor many fine-grained parameters, such as the Java heap memory utilization and the JDBC connection pool size. It can also capture the monitored data in a historical file and later be used to analyze the behavior of the application server JVM.
- The db2top monitoring utility is used for monitoring the DB2 database behavior and SQL performance during their executions at run time. This online monitoring utility monitors and analyzes the behavior of the database. This tool is also used to quickly identify and diagnose performance issues and poorly performing SQL queries.

Performance test tool

The performance test tool that is used to test the SPSS C&DS solution is the soapUI tool. For more information about the soapUI tool, see this website:

`http://www.soapui.org`

LPAR configuration details

For a single instance LPAR deployment, we used the LPAR configurations that are shown in Table 3-6.

Table 3-6 LPAR configuration

Number of cores	Memory
2	4 GB
4	8 GB
6	12 GB

Test approach

The testing methodology focused on achieving the highest scoring transaction rate on a stand-alone server and on a cluster environment. Test results are measured by using the average number of scores per second during the test run.

The main goal is to create an environment that replicated the scoring service from a user perspective and that produced an optimal level of scores per second. We address the test scenario that was used to performance-test the SPSS solution. We also compare the baseline performance with the solution that is optimized for performance.

Single-instance configuration

The performance testing initially focused on a single instance of WebSphere with a Collaboration and Deployment Services server running. After we captured the baseline data, we experimented with different Java options and memory allocation settings on a single instance of WebSphere.

During this test execution, we analyzed performance profiling reports and discovered contention in sending or receiving web service messages in Collaboration and Deployment Services. We capture the various tuning parameters and the performance improvement with each parameter change.

We initially tested with Java 6 SR6, which ships with WebSphere Network Deployment 7.0.0.23. With this version, the test team set 64K page usage manually for the Java binary. The WebSphere Network Deployment 7.0.0.17 ships with Java 6 SR9 FP1. The Java 6 SR9 FP1, by default, has a page size that is set to 64K.

We executed several trial runs before we determined the optimized set of performance scores per second for different Collaboration and Deployment Services test variations. We used different model cache settings with different numbers of virtual users initially. Based on the test result analysis, we achieved the highest scores per second by applying optimized AIX environment settings.

The test results documented in this publication include baseline performance results. We also capture the performance results in an optimized environment. The test results of Collaboration and Deployment Services from the baseline are compared to the results obtained by AIX-tuned settings to access performance improvements by using AIX environment settings. The workload is driven to achieve nearly 99 - 100% processor utilization in an optimized environment.

All the tests ran with a warm-up period of 300 seconds and a measurement period of 180 seconds, with which the test results with system metrics are captured. Before we execute each new test run, we stop DB2 and the WebSphere Application Servers and start them. If clustering with DB2 and WebSphere Application Servers, Node Manager and Deployment Manager are also stopped and started. Therefore, we cleaned up the environment before each test run.

Test workload

This section explains the workload characterization of SPSS Collaboration and Deployment Services performance testing. The various workload tuning parameters that are set from the client to drive more load to the system also are explained in this section.

ProjectProfile1 is a medium-complex model that is used in all performance variation testing in this book. The analytic algorithms that are used in ProjectProfile1 are Bayesian Network, Logistic Regression, and Discriminate. This type of stream was used to drive the test. The type of the stream that was used in ProjectProfile1 is Auto Model with no rules. The Auto Model with no rules stream is created by using SPSS Modeler Advantage. An auto-model is built with a target and eight input fields. No expressions or rules were used to limit the model.

With the SPSS Modeler Advantage, the interface is simple so that business users can build a model without requiring the skills of a typical IBM SPSS Modeler user. SPSS Modeler Advantage allows business users to select data, build models, and apply models in the decision management framework without any high-level analyst skill.

Rather than being a stand-alone product, SPSS Modeler Advantage generates the model creation stream in the background. This shared underlying model allows modelers and business people to work together. The modeler adds expertise to those streams that are generated by business users, for instance, or integrates standard elements that are already in use within the company.

Scoring the configuration model cache setting

The model cache is the key performance tuning setting for SPSS Collaboration and Deployment Services scoring tests, which helps to drive the workload. During the performance testing, different model cache settings (2, 100, 200, and 300) were applied to determine the optimal model cache. Based on the scores that were obtained for different model cache settings, the 300 model cache value performed the best among the other settings. All the test results that were documented were used with the 300 model cache setting.

Figure 3-30 on page 108 presents the improvement in performance of the solution relative to the baseline performance as a result of applying different model cache settings.

Model cache setting: To set the scoring configuration model cache in the browser, go to the SPSS C&DS browser page. Go to **Configuration** → **Scoring Service** → **WorkerPoolMaximumSize**. Set this value to 300 or the value that you want and update the settings.

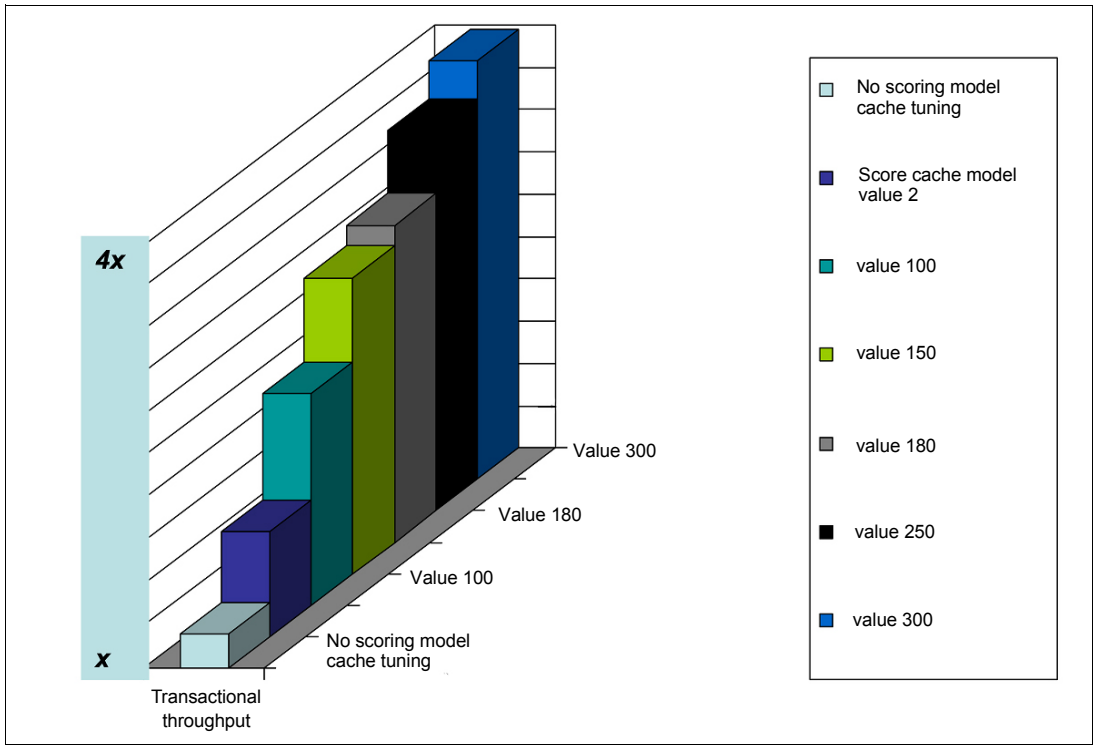

Figure 3-30 Baseline performance versus Cache model tunable performance

WebSphere WebContainer thread pool setting

The WebSphere Application Server web container manages all HTTP requests to web services. These requests are processed by a pool of server threads. The minimum and maximum thread pool size for the web container can be configured for optimal performance.

Settings: It is suggested to set the same value for web container settings as that of scoring model cache.

Scoring configuration response data setting

The performance characteristics of real-time scoring for a specific stream are influenced by the number of fields that are provided as predictor inputs and responses. The number of predictor fields that is used as input is determined by the stream function during model building. A preferred practice is to limit the response fields to the minimum. Specifically, if possible, do not echo the input fields in the response fields. This approach might not be necessary if all the input field values are provided in the web service invocation. Figure 3-31 on page 109 represents the performance improvement over retrieving one response output variable compared to two output variables.

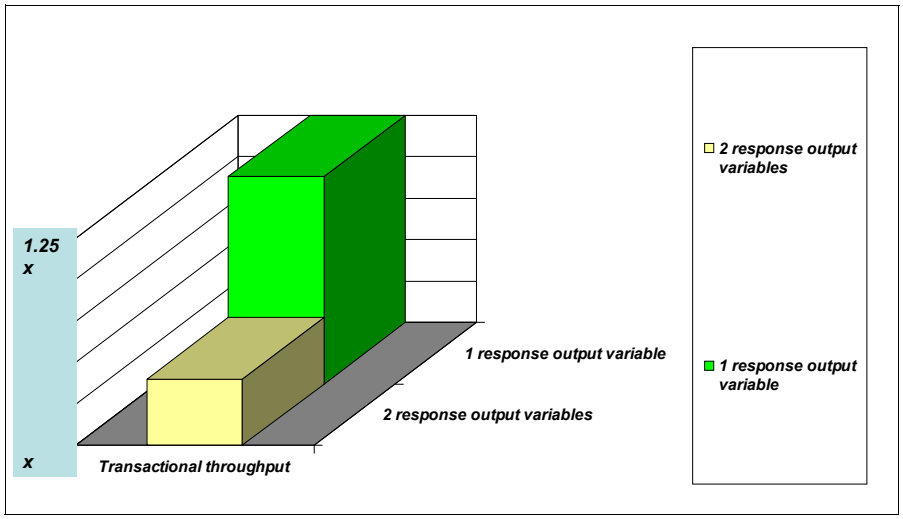

Figure 3-31 Performance of two response output variables versus one response output variable

WebSphere Application Server additional memory-related tuning

We show the tuning configuration for WebSphere Application Server that helped achieve better results for the workload tested:

- Java heap and garbage collection policy settings:
  - Java heap size was set to 4096 MB.
  - The JDK options that were used were **`-Xgcpolicy:gencon -Xcodecache32m -Xlp64k`**:
    - **`gcpolicy:gencon`**: This option sets the garbage collection policy to gencon. This option handles short-lived objects differently than long-lived objects. Applications that have many short-lived objects can see shorter pause times with this policy and still achieve good throughput.
    - **`codecache32m`**: This option adds 32-MB code cache segments to the JVM memory. It sets the unit size of the memory that is allocated to store the native code of compiled Java systems.
    - **`lp64k`**: This option is used to configure the JVM with the specific page size as parameter. In our tests, we used the value of 64 KB. The use of larger page sizes can reduce the processor overhead to track heap memory, therefore, increasing the overall throughput of the application.

Important: If you decide to use page sizes that are equal to or larger than 16 MB, ensure that all AIX tuning is performed to use large page sizes. For more information about the use of the **`vmo`** command to set large page sizes, see this website:

`http://pic.dhe.ibm.com/infocenter/wasinfo/v6r1/index.jsp?topic=%2Fcom.ibm.websphere.base.doc%2Finfo%2Faes%2Fae%2Ftprf_tuneaix.html`

- MALLOCOPTIONS: Initially, tests are run with different MALLOCTYPE (default, watson, and 3.1_64BIT) and MALLOCOPTIONS (multiheap, buckets, and pool) AIX environment setting combinations. Each setting was analyzed and we determined the final setting of a 512-MB pool size and 32 heaps (`pool:0x20000000,multiheap:32`) combination. These settings gave optimal performance for both stand-alone and clustered Collaboration and Deployment Services.

  The Collaboration and Deployment Services server is a multithreaded application where multiple heaps are required so that the server process can issue **`malloc()`** and **`free()`** system calls from more than one thread. With a single heap, this process is not achieved as seen from the baseline result, which is a serious performance bottleneck. With MALLOCOPTIONS set to `multiheap`, it improved the performance of Collaboration and Deployment Services by nearly 19 times over the baseline results:

  ```
  export MALLOCOPTIONS=multiheap
  ```

  `tprof` reports are analyzed with these settings. We observed that the server process issues many small allocation requests to AIX. So, we tried to set MALLOCOPTIONS to `pool`. Coupled with the multiheap option, we observed that the performance improved significantly:

  ```
  export MALLOCOPTIONS=pool:0x20000000,multiheap:32
  ```

  The result of each MALLOCOPTIONS setting for single and multiple WebSphere testing is presented in 3.1.3, “Optimal performance with clustered WebSphere Application Server” on page 112.

- Memory affinity

  We enabled memory affinity so that the memory that is allocated is close to the processor. To enable memory affinity, the MEMORY_AFFINITY environment variable was set to `MCM`:

  ```
  export MEMORY_AFFINITY=MCM
  ```

- AIX user limits

  It is important to remove any limits that exist on the user profile that runs the WebSphere Application Server. Example 3-4 shows how to check the limits.

  Example 3-4 Checking user limits on AIX

  ```
  root@spss01:/ # ulimit -a
  time(seconds)        unlimited
  file(blocks)         unlimited
  data(kbytes)         unlimited
  stack(kbytes)        unlimited
  memory(kbytes)       unlimited
  coredump(blocks)     unlimited
  nofiles(descriptors) unlimited
  threads(per process) unlimited
  processes(per user)  unlimited
  ```

  If any of the user limit parameters that are shown are not set to `unlimited`, you must change the parameter to `unlimited`.

- IOO and VMO settings

  We tried various settings for the AIX tunables to better achieve SPSS C&DS performance. Example 3-5 on page 111 shows the suggested values for IOO and VMO settings.

Example 3-5 AIX IOO and VMO tunables

```
vmo:
        maxfree = "1216"

ioo:
        maxpgahead = "64"
        j2_maxPageReadAhead = "256"
        j2_minPageReadAhead = "16"
```

For **vmo**, we changed the **maxfree** parameter. This parameter specifies the number of frames on the free list at which page-stealing is to stop.

The Virtual Memory Manager (VMM) maintains a list of free real-memory page frames. These page frames are available to hold the virtual-memory pages that are needed to satisfy a page fault. When the number of pages on the free list falls under the number that is specified by the **minfree** parameter, the VMM begins to steal pages to add to the free list. The VMM continues to steal pages until the free list has at least the number of pages that is specified by the **maxfree** parameter.

This **vmo** setting is changed because the **ioo** changes that we made required it to be bigger than the default value. For **ioo**, we changed the **j2_maxPageReadAhead** and **j2_minPageReadAhead** values.

If a process appears to be reading sequentially from a file, the values that are specified by the **j2_maxPageReadAhead** and **j2_minPageReadAhead** parameters determine the number of pages to be read ahead when the condition is first detected. The value that is specified by the **j2_maxPageReadAhead** and **j2_minPageReadAhead** parameters sets the maximum and minimum number of pages that are read ahead, regardless of the number of preceding sequential reads.

The difference between **minfree** and **maxfree** must always be equal to or greater than **j2_maxPageReadAhead**. If run time decreases when the value of **j2_maxPageReadAhead** increases, ensure that the performance of the other applications does not deteriorate.

Figure 3-32 shows is the improvement in performance that is achieved through the IOO and VMO settings.

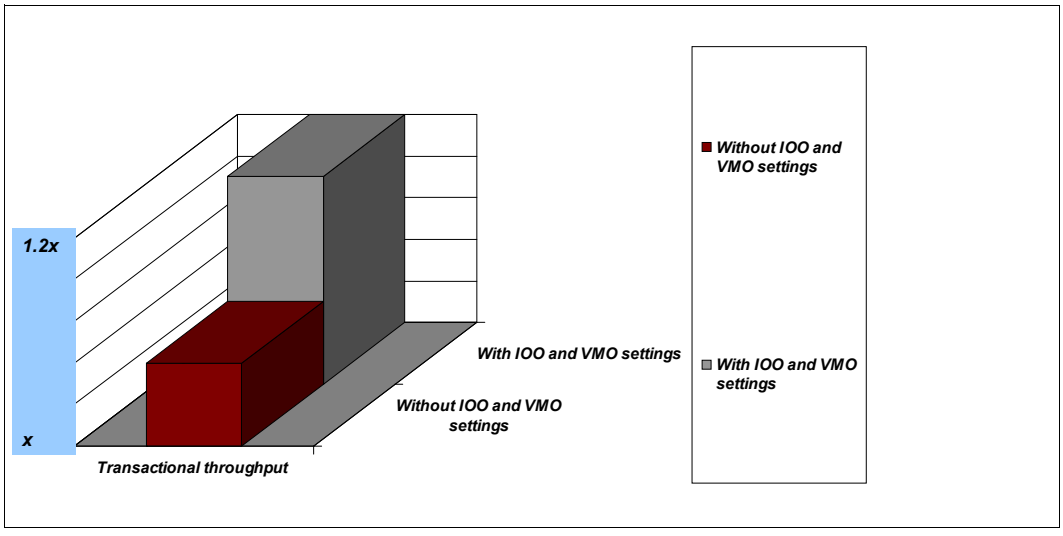

Figure 3-32 Performance gain by using IOO and VMO tuning parameters in AIX

Performance comparison between two, four, and six POWER7 cores

This section provides sizing guidance by comparing the performance of the SPSS C&DS solution among two, four, and six POWER7 cores.

By using a single WebSphere Application Server instance, we observed a linear scoring in the service workload, which is graphically represented in Figure 3-33.

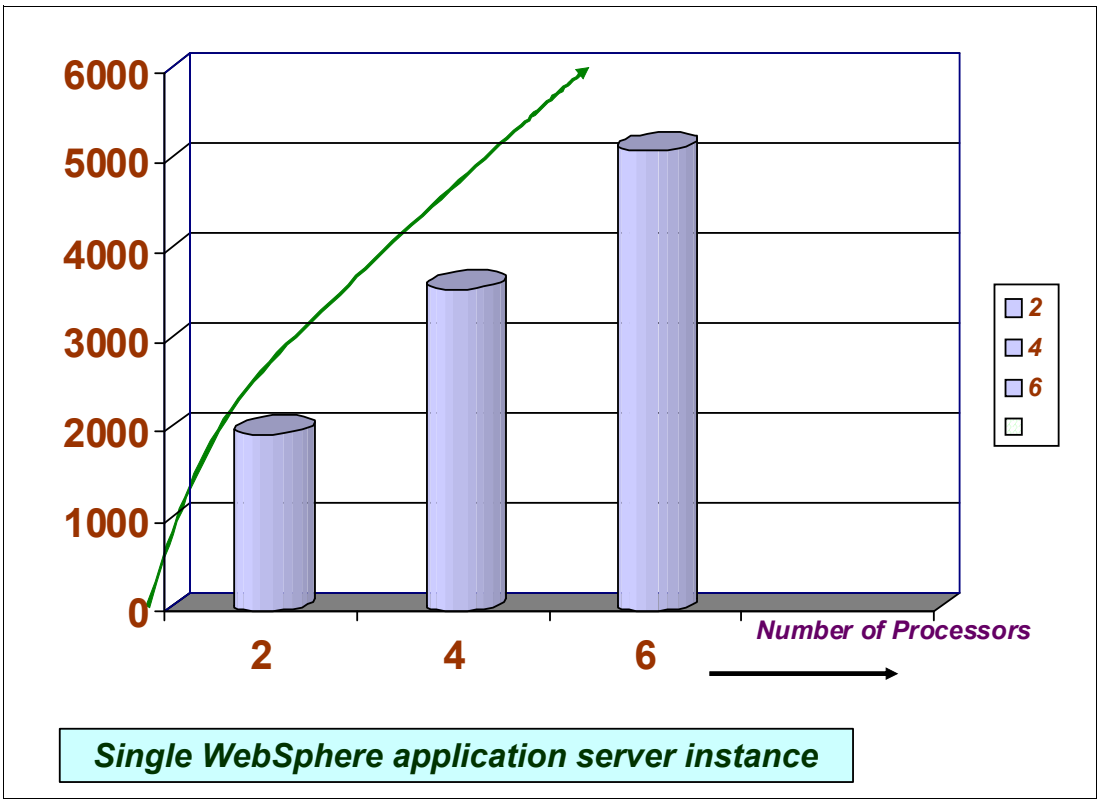

Figure 3-33 Performance comparison of two, four, and six processors

3.1.3 Optimal performance with clustered WebSphere Application Server

We describe the SPSS performance on a WebSphere Application Server cluster environment. For this scenario, we configured the SPSS C&DS solution in a single LPAR with multiple WebSphere Application Server instances that work in a clustered topology.

Setting up the SPSS C&DS cluster

We describe the performance of the SPSS solution in a clustered environment with multiple WebSphere Application Server instances that serve the incoming requests.

The repository can be deployed into an environment of clustered Java 2 Platform, Enterprise Edition (J2EE) application servers. Each application server in the cluster must have the identical configuration for the hosted application components. The repository is accessed through a hardware-based or software-based load balancer.

This architecture allows processing to be distributed among multiple application servers. It also provides redundancy in the case of a single server failure.

Figure 3-34 on page 113 shows the topology of a clustered SPSS C&DS solution.

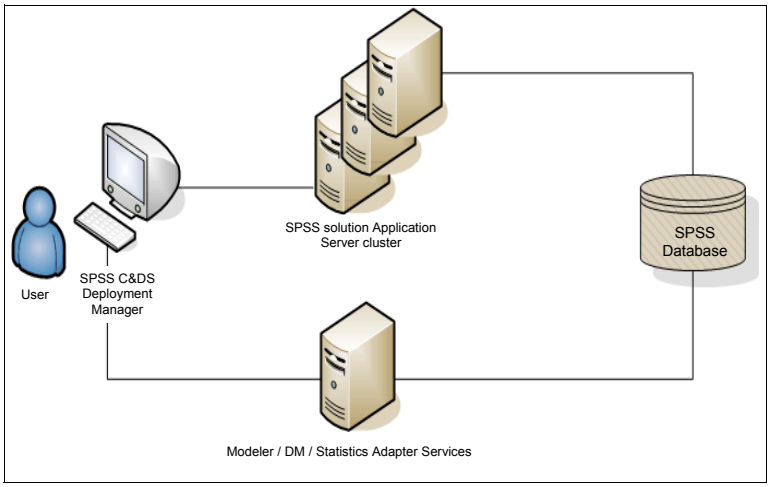

Figure 3-34 SPSS Cluster Deployment topology

We describe how to configure a clustered environment after you install a stand-alone application environment.

Scripted cluster deployment

Several scripts can be used to automate the cluster setup and the cluster deployment into WebSphere. These scripts are in the `<cluster install location>/scripts` directory.

The `config.ini` file contains the parameters that are used by the Jython scripts (described next) to automatically create a cluster. The file contains the following sections and properties.

The Cluster section contains these properties:

- *name* of the cluster (for example, *spss_cluster*)
- *cell* name of the WebSphere cell for the cluster (for example, `spss01Cell01`)
- *singletonServer* name of the server to deploy applications that can run on a single server only
- *singletonNode* name of the node for the singleton server

The Servers section contains these properties:

- *name* of the server (for example, `spss01`)
- *node name* of the node for the server (for example, `spss01Node01`)
- *javaInitHeapSize* initial Java heap size (for example, `256`)
- *javaMaxHeapSize* maximum Java heap size (for example, `1024`)
- *platformOS* operating system: Valid values include `aix`, `aix64`, `hpux64`, `linux`, `linux64`, `solaris64`, `windows`, and `windows64`
- *platformSharedDir* shared platform installation directory (for example, `\\machine\shared\platform_install`)
- *platformLocalDir* directory of the local platform installation (for example, `C:\platformLocal`)
- *platformKeystoreLocation* location of the platform keystore that is created at installation time
- *platformKeystorePassword* password of the platform keystore that is created at installation time

The Java Message Service (JMS) section contains this property:

- *dataStoreSchema* name of the schema to use for the JMS data store

The Platform section contains these properties:

- *database.name* name of the database that was selected in the setup
- *database.driver* platform database driver class name
- *database.host* host for the database server
- *database.library* database library (IBM System i™ only)
- *database.user* platform database user
- *database.password* platform database password (might be encrypted)
- *database.url* platform database URL
- *deploy.directory* platform to deploy directory

These configuration parameters are used:

- *platformDeploy.py* creates a cluster that contains all the servers that are defined in the Servers section. The configuration in the Platform section is used to create the platform data source and deploy the platform applications to the cluster.
- *platformClean.py* is used to remove the components that are installed by the `platformDeploy.py` script.

platformDeploy.py

This script is used to deploy platform components to a clustered WebSphere domain as configured in `config.ini`. The script uses the following arguments:

- *all* Deploy everything (default)
- *sharedLibrary* Deploy the shared library only
- *cluster* Deploy the cluster, servers, and shared library
- *servers* Deploy the servers only
- *virtualHosts* Update the virtual host aliases for the cluster
- *components* Deploy the data source and JMS components
- *datasource* Deploy the data source components only
- *jms* Deploy the JMS components only
- *applications* Deploy the platform applications
- *patch* Deploy the updated platform applications

platformClean.py

This script is used to undeploy platform components from a clustered WebSphere domain as configured in `config.ini`. The script uses the following arguments:

- *all* Undeploy everything (default)
- *sharedLibrary* Undeploy the shared library only
- *cluster* Undeploy the cluster, servers, and shared library
- *servers* Undeploy the servers only
- *virtualHosts* Update the virtual host aliases for the cluster
- *components* Undeploy the data source and JMS components
- *datasource* Undeploy the data source components only
- *jms* Undeploy the JMS components only
- *applications* Undeploy the platform applications
- *patch* Undeploy the updated platform applications

Important: If you clear your JMS components, you also need to remove the database tables before you can re-create them. Delete the tables that start with `SIB` in your platform database. They are re-created on server startup after you re-create your JMS components.

The executable scripts for running Jython scripts are in the <*cluster install location*>/scripts directory:

- *setEnv* Sets up the environment
- *wsadmin* Executes a specified Jython script
- *installNode* Installs the necessary platform components on the local file system

After the previous scripts are edited, follow these steps to deploy a WebSphere Application Server cluster:

1. Install the same version of WebSphere Application Server Network Deployment on each node in the cluster:
   - Set up a single WebSphere Deployment Manager. If you patched WebSphere after you created the Deployment Manager profile, you might need to re-create the Deployment Manager profile for the wsadmin scripts to run correctly.
   - Federate all nodes in the cluster by using the Deployment Manager.
2. Set up the platform installation directory as a shared directory for each node in the cluster.
3. Update /bin/setEnv to set the values for the following environment variables:
   - *DM_PROFILE_HOME* Location of your WebSphere Deployment Manager profile
   - *WSADMIN_LANG* Language of your scripts (leave the default of Jython)
   - *WSADMIN_SECURITY* User name and password if administrative security is enabled
4. Update config.ini to set up your cluster configuration.
5. Run the script to deploy the platform components into a WebSphere cluster. The configuration is read from config.ini.
   - Open a command prompt to the <*cluster install location*>/bin directory.
   - Execute **wsadmin -f ../scripts/platformDeploy.py**
   - The following components are deployed:
     - Shared library is targeted to the cluster.
     - JDBC data source is targeted to the cluster.
     - JDBC persistent stores are targeted to a single server.
     - JMS server is targeted to a single server.
     - A JMS connection factory is targeted to the cluster.
     - Several JMS queues; all platform applications (EARs, WARs, and RARs).
6. Start all nodes in the cluster by using the Administration console.

Important: If the ports are manually changed for any server in the cluster, the corresponding changes must be made to the default_host virtual host aliases to ensure that cluster communication functions correctly.

SPSS Collaboration & Deployment on a clustered WebSphere

After the cluster is created and deployed, we assign the processor affinity to the individual WebSphere Application Servers that run on the system.

Example 3-6 on page 116 shows the command that is used to configure the processor affinity from 0 to the seventh logical processor to WebSphere Application Server 1, which is one of the many application server instances. It creates a resource set with logical processors 0 - 7 and starts the WebSphere Application Server instance1. It binds WebSphere Application Server to two cores of the POWER7 processor (eight logical processors in SMT4 mode).

Example 3-6 execrset command syntax

```
root@spss01:/ # execrset -c 0-7 -m 0 -e
/usr/IBM/WebSphere/AppServer/profiles/AppSrv01/bin/startServer.sh server1
```

Parameters: The same JVM and system tuning parameters that are used on the tests with a single WebSphere Application Server instance are kept for the clustered environment.

The use of the resource sets command allows WebSphere Application Server to run on an assigned number of logical processors and to use the same physical processors always. Therefore, it creates affinity between the WebSphere Application Server process and the physical processor cores.

Based on the number of multiple WebSphere Application Server instances that are running, you need to run the command that is shown in Example 3-6 for each instance. Change the specified set of logical processors so that each instance uses a different set of cores. Ensure that each instance uses only cores from one POWER7 processor chip.

Now, the WebSphere Application Server instance is limited to run on a defined number of logical processors. It is always suggested to enable memory affinity so that memory can be allocated closest to the processor. To enable memory affinity, the *MEMORY_AFFINITY* environment variable is set to `MCM`, as shown on Example 3-7.

Example 3-7 Exporting environment variable for memory affinity

```
root@spss01:/ # export MEMORY_AFFINITY=MCM
```

In our test scenarios, we created the affinity of two processor cores per WebSphere instance, as shown in Table 3-7.

Table 3-7 Summary of the number of WebSphere instances per partition size

Partition size	Number of WebSphere Application Server Instances
4 cores	2 (2 per core)
8 cores	4 (2 per core)
16 cores	8 (2 per core)
32 cores	16 (2 per core)
64 cores	32 (2 per core)

Test results

The following section shows the test results.

Clustered WebSphere Application Server

Our tests focused on determining the scalability of SPSS C&DS under a clustered WebSphere environment, and the results are shown on Figure 3-35 on page 117.

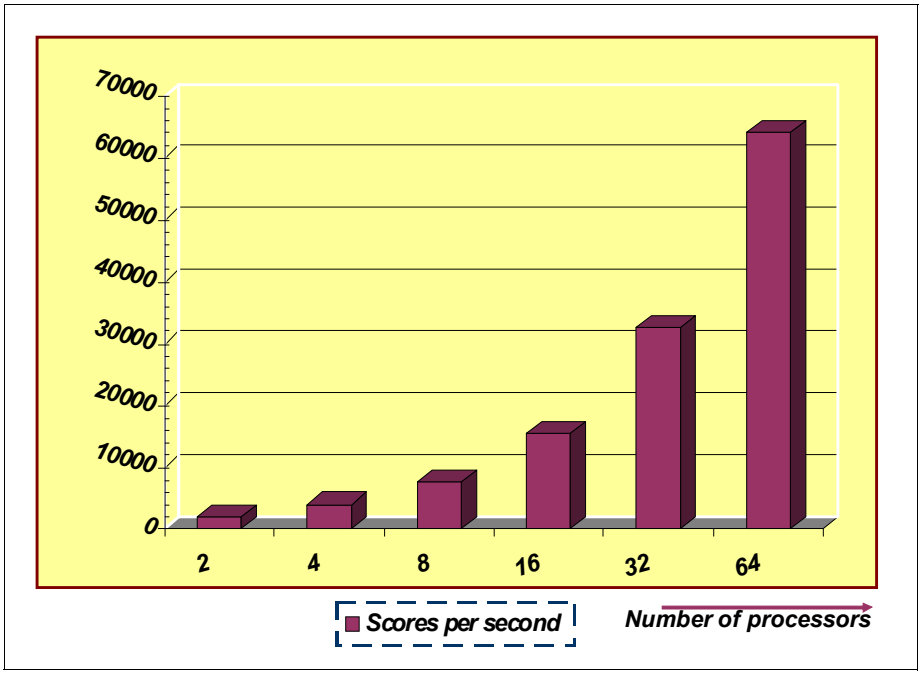

Figure 3-35 SPSS scalability from 2 - 64 processors on the well-tuned environment

We infer from Figure 3-35 that for the specific workload that was tested (“Test workload” on page 107), every processor addition resulted in nearly 1,000 scores/second. The processor in Figure 3-35 is the Power 780 processor with simultaneous multithreading (SMT) enabled. Figure 3-35 shows a linear growth between scores/second and processors.

Clustered WebSphere Application Server with the POWER7 TurboCore feature

By using the same test cases that we used in “Clustered WebSphere Application Server”, we ran the tests on a Power System 780. We enabled the POWER7 TurboCore feature (“TurboCore mode on IBM Power Systems” on page 31) on an 8-core partition.

TurboCore mode on IBM Power Systems is a special processing mode of these systems in which four cores only per chip are activated. With only four active cores, the ease of cooling allows the active cores to provide a frequency that is faster (~7.25%) than the nominal rate.

The POWER7 cache design has 4 Mbytes of L3 cache per core. Although it might appear as though there is a private L3 cache per core, this cache can be shared among cores. The cache state from an active core L3 cache can be saved into the L3 cache of less active cores.

With TurboCore mode, the cache state from the four active cores can be saved into the L3 cache of the TurboCore mode inactive cores. The result is more accessible cache per core.

Both the higher frequency and the greater amount of cache per core are techniques for providing better performance. It is common for a longer-running, even multithreaded, workload that accesses largely private data to experience a performance benefit in excess of the benefit that might be expected from the better frequency alone. Even more complex workloads that reside in a partition with affinity to the cores and memory of a certain processor can experience similar benefits.

With the multiple threaded architecture and design of SPSS, the TurboCore feature shows a 12% improvement over SPSS C&DS throughput. Figure 3-36 on page 118 shows a comparison graph.

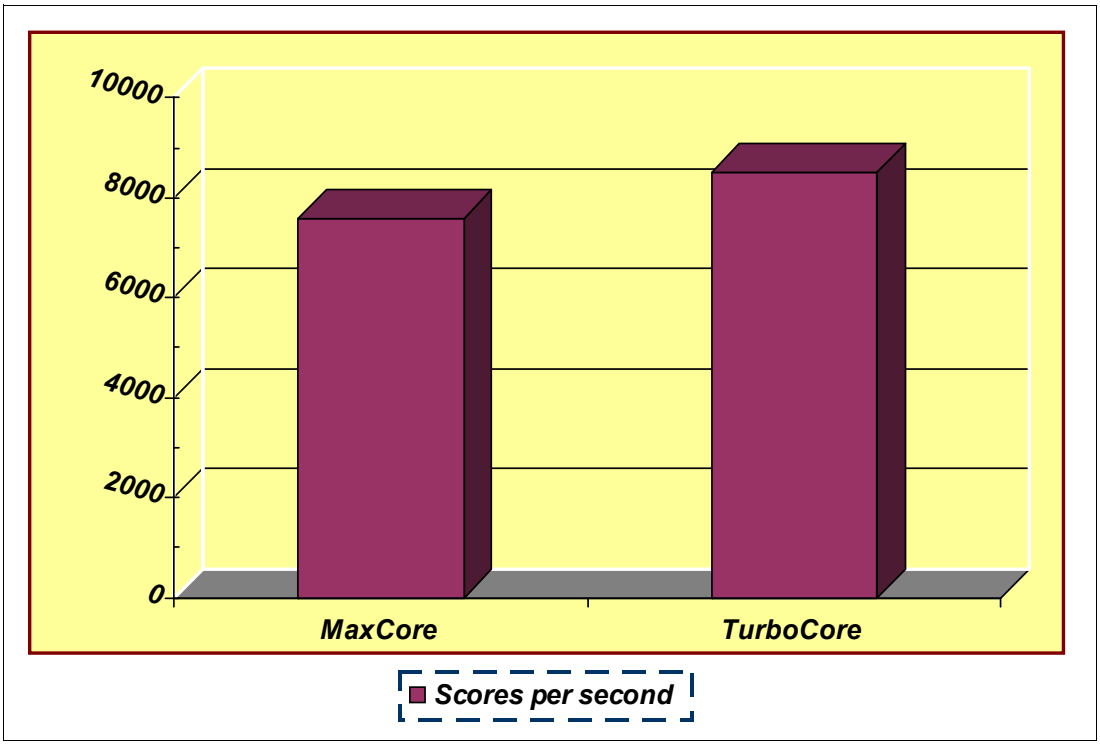

Figure 3-36 MaxCore versus TurboCore performance comparison for SPSS C&DS

Clustered WebSphere Application Server with AIX Active System Optimizer

By using the same test cases that we used in “Clustered WebSphere Application Server”, we enabled AIX 7.1 Active System Optimizer (ASO) on an 8-core partition.

To run ASO, the system must be at AIX 7.1 Technology Level (TL) 01 or greater, on POWER7 hardware that runs in POWER7 mode. ASO is installed, by default, on AIX 7.1.

ASO is designed to improve cache and memory affinity by dynamically evaluating and changing tuning options dynamically. It pre-monitors and post-monitors to ensure that the changes improve performance. ASO provides cache affinity, aggressive cache affinity, and memory affinity. It monitors performance and detects situations where threads can be moved from one chip to another to use the closer L3 cache. ASO operates best on multithreaded workloads. The jobs that it monitors must be stable and long-running so that it can most effectively change the workload.

To start ASO, follow the procedure that is shown in Example 3-8.

Example 3-8 Starting ASO

```
root@spss01:/ # startsrc -s aso
0513-059 The aso Subsystem has been started. Subsystem PID is 13369556.
root@spss01:/ # aso -o aso_active=1
Setting aso_active to 1
```

Two log files are in the `/var/log/aso/*` directory:

- `aso.log`, which has on/off/hibernating information
- `aso_process.log`, which provides the details of actions and modified processes

Although we used AIX ASO in our tests, we started WebSphere Application Server instances without **execrset** because we wanted ASO to perform all of the affinity work.

AIX ASO was able to set memory and cache affinity within our workloads automatically, and it increased SPSS C&DS throughput by 8%.

3.1.4 Using PowerVM shared processor pools for license management

PowerVM provides industrial-strength virtualization environments on IBM POWER processor-based systems. PowerVM technology is designed for clients to build a dynamic infrastructure to reduce their costs and improve their service levels, including license management.

License management

An important part of any solution is the cost. It is important to optimize the solution for licensing and performance.

SPSS can be purchased through Passport Advantage® from either IBM or an authorized IBM Business Partner. The way that SPSS is licensed depends on the SPSS products that you use.

The following SPSS products are licensed by either authorized user, authorized user value unit, or concurrent user:

- IBM SPSS Modeler Professional
- IBM SPSS Modeler Premium
- IBM SPSS Collaboration and Deployment Services - Scoring
- IBM SPSS Collaboration and Deployment Services Deployment Manager
- IBM SPSS Collaboration and Deployment Services Deployment Portal

The following SPSS products are licensed by Processor Value Unit (PVU):

- IBM SPSS Modeler Server Professional
- IBM SPSS Modeler Server Premium
- IBM SPSS Collaboration and Deployment Services - Automation

PVU units are calculated from the processor type and multiplied by the number of processors to which the software has access, or the number of processors in the server. Licensing fewer than the total number of processors in the server requires *sub-capacity licensing*, which has special rules and conditions. The following link describes those rules and how to calculate the required PVU licenses:

http://www-01.ibm.com/software/lotus/passportadvantage/Counting_Software_licenses_using_specific_virtualization_technologies.html

The following basic core calculations are for sub-capacity licensing with features:

- Capped LPAR cores are calculated by Entitled Value (EV).
- Uncapped LPAR cores are calculated by Virtual Processor (VP).
- Cores are capped by the number of cores that is available in the shared system pool.
- IBM Tivoli License Manager must be running, although there are some exceptions.

With some of the PowerVM features, you can license your software efficiently. The main PowerVM feature that enables this licensing is the *Multiple Shared Processor Pool* feature. With this feature, you can limit the number of cores that is required to be licensed. You still can take advantage of the flexibility of uncapped LPARs.

3.1.5 SPSS tuning and performance statistics

We describe the various tuning configuration parameters on the SPSS solution that helped with the performance testing that is shown in this book. Table 3-8 shows the tuning parameters and their significance on performance.

Table 3-8 Tuning parameters and their significance on performance

Sl. No	Layer or component	Parameter	Value (from - to)	Performance improvement potential	Remarks
1	**SPSS application settings**	Scoring configuration model cache setting	1 - 300	Very high	The default value is 100 in the configuration manager. This value needs to be increased to 300.
2		Scoring configuration	2 - 1	High	By reducing the response variables, you can increase the scoring throughput proportionately.
3	**WebSphere Application Server**	Web container thread pool size	10 - 300	Very high	We advise you to keep this setting the same as the scoring model cache size setting.
4		Default thread pool size	10 - 50	Medium	The default thread pool contains the thread pool for TCP socket connections. We advise you to keep this setting at 50 for a value of 300.
5		JVM option: -Xlp64K	4 - 64	Medium	By changing the allocation page size of JVM from 4 to 64 KB, you help improve the overall object allocation time for the JVM.
6		JVM option: -Xcodecache32m	32	Low	Adds code cache segments.
7		JVM option: Max heap size	4096	High	Sets the maximum size of the JVM heap.
8		JVM option: -Xgcpolicy:gencon	optthruput -gencon	High	Reduced garbage collection overhead by 20%.
9		Malloc options	pool multiheap 32	Very high	By changing the memory allocation option in the operating system, you see better performance results.
10		Memory affinity	MCM	Medium	There was marginal improvement in performance with this setting.

Sl. No	Layer or component	Parameter	Value (from - to)	Performance improvement potential	Remarks
11	**AIX operating system settings**	user limits	unlimited	Medium	Enables extra resource allocation for the application server instances.
12		j2_maxPageReadAhead	128 - 256	Medium	Changes the maximum number of pages to be read ahead sequentially when it accesses a file when you use Journaled File System 2 (JFS2).
13		j2_minPageReadAhead	2 - 16	Medium	Changes the minimum number of pages to be read ahead sequentially when it accesses a file when you use JFS2.
14		maxfree	1088 - 1216	Medium	Enables the increase of the j2_maxPageReadAhead value.
15		ASO	Set to active	High	See “Clustered WebSphere Application Server with AIX Active System Optimizer” on page 118.
16		**execrset** command	Processor affinity	High	Creates a resource set (rset) for the processes. It limits the processes that use **execrset** to run on a defined set of cores.
17	**Power System settings**	Turbo core	Off - on	High	Enables TurboCore mode on a Power System.
18	**DB2 settings**	DB2 buffer pool	1024 - 16384	Low	This setting helped reduce physical disk reads on the database and increased the cached/logical reads. It improved the overall performance of the solution.

3.2 IBM SPSS Modeler

The last few sections focused on deploying SPSS by using SPSS Collaboration and Deployment. This method provides the flexibility of an available repository, but it is not essential if you need to run Modeler only.

There are two ways other than Collaboration and Deployment to run IBM SPSS Modeler. The first method is to run it on a single Microsoft Windows system. The second method is to run the server part of Modeler on a server, either Windows, Linux, or UNIX or on AIX on a POWER7 system. All the data can remain in the data center where it is more secure and can typically be transferred across a faster network with more bandwidth. The streams and models are still run on the Windows portion of Modeler.

In this scenario, we set up SPSS with the environment that is shown in Figure 3-37 on page 122.

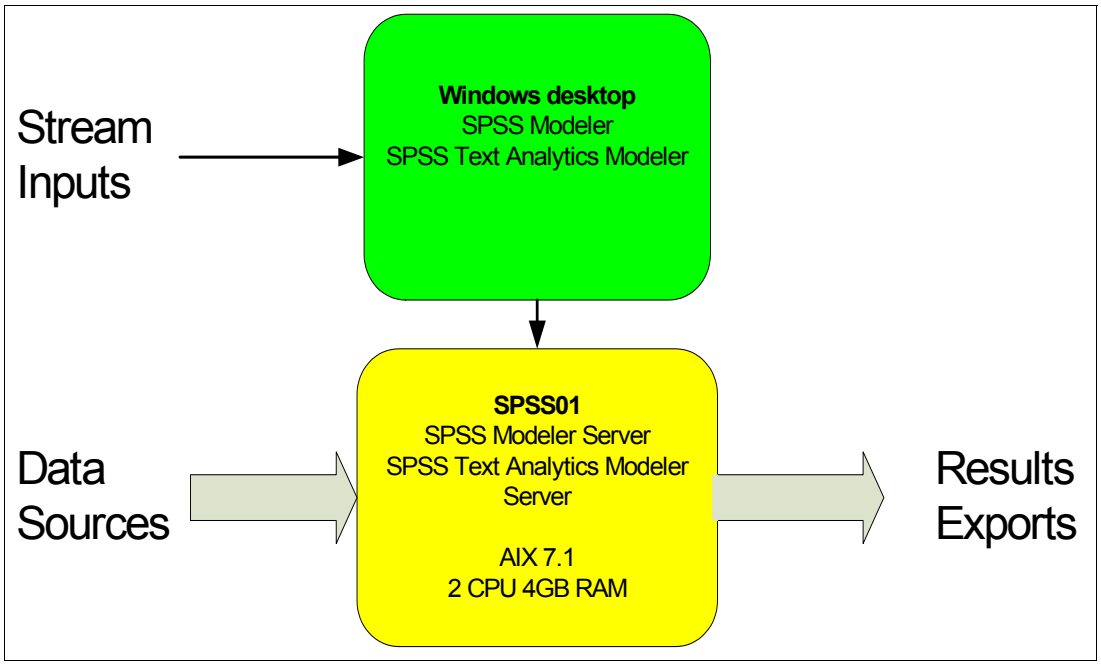

Figure 3-37 SPSS Modeler Server logical diagram

We set up SPSS to analyze the sentiment of 10,000 - 1,000,000 short text messages. The SPSS components that we used are SPSS Modeler Premium and SPSS Modeler Server Premium. The premium version includes text analytics, which we used for these tests.

The IBM SPSS Modeler runs on Windows only so we required SPSS Modeler Server to run the analytics on our POWER7 system.

We are required to install the SPSS Modeler Premium on a Windows desktop, and SPSS Modeler Server Premium is installed on an AIX LPAR on a Power 750. The LPAR is configured with 4 GB of RAM, 2.0 processor entitlement, and two virtual processors.

We installed the modeler server first, which requires that you export your display by using either X Windows server program or VNC viewer. After we installed the SPSS Modeler Server, we installed the Text Analytics plug-in.

Important: Both SPSS Modeler Server and SPSS Text Analytics Modeler Server must be installed on the AIX LPAR to perform the text analytics workload.

After the installation, the Modeler Server must be started as shown in Example 3-9, from the Modeler Server installation directory.

Example 3-9 Starting SPSS Modeler Server

```
root@spss03:/ # modelersrv.sh start
```

The Modeler Server starts to run and listens for client connections on port 28052. The next step is to install the IBM SPSS Modeler and SPSS Text Analytics on a Windows system.

Important: The evaluation license works on desktop versions of Windows only. To install on a Windows server, a full server license key is required.

When the IBM SPSS Modeler and the SPSS Text Analytics products are installed on your Windows system, you need to connect the Windows system to the Modeler Server. When the modeler is opened, select **tools** from the top menu and then select **Server Login**. The window that is shown in Figure 3-38 opens.

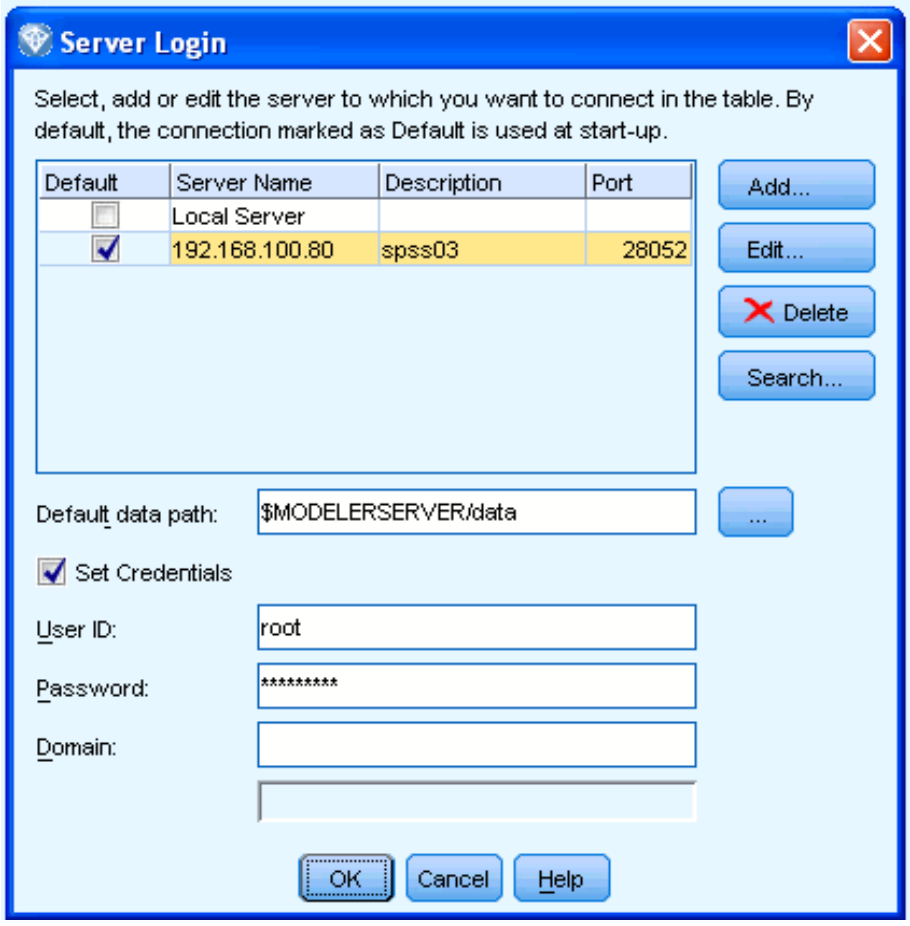

Figure 3-38 Server Login window

Click **Add** to add the server to which you want to connect. Enter the server IP address or host name, leave the port at the default (28052), and type a description. Click the newly added server and add the login set credentials and default data path.

Important: When you use a distributed Modeler server, all the data sources must be on the server and not on the Windows machine.

When you click **OK**, the Modeler client logs in to the server and the window closes. If the login fails, you see a message and the window remains open.

After the IBM SPSS Modeler logs in to the server, you are ready to start processing streams. For our scenario, we used a simple stream. This stream takes short text messages from an Excel file and performs sentiment processing on each message and outputs the sentiment results to a table. The Excel file is transferred to the AIX LPAR that runs on a Power 750. See Figure 3-39 on page 124.

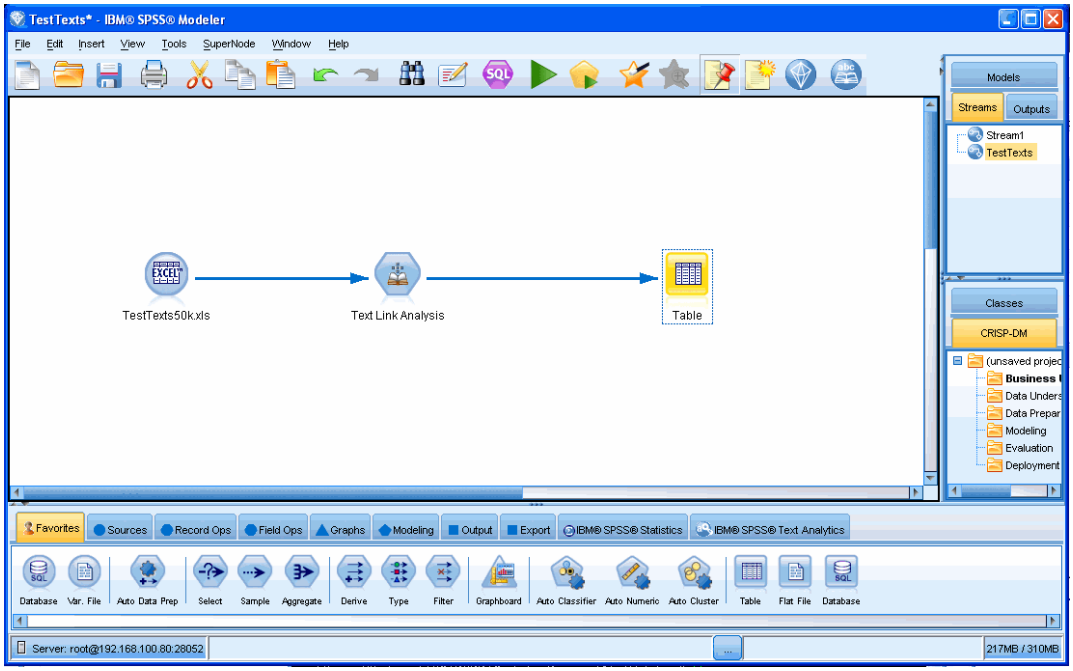

Figure 3-39 SPSS Text Analytics stream

We tested different amounts of text messages per file and also tested different processor numbers and memory combinations.

3.2.1 Results

We discover that the more text messages are in the file, the quicker they are processed. We tested with 10,000, 50,000, and 500,000 messages per file and found that a 10 times increase in workload entailed only around a 6.72 times increase in time.

Our tests also compare the amount of time that is spent to process the same input file with different amounts of cores (processing units) that are assigned to the LPAR. This test checks how much parallelism can be obtained in the processing of a stream.

Figure 3-40 on page 125 compares the results for 1, 2, 4, and 8-core scenarios.

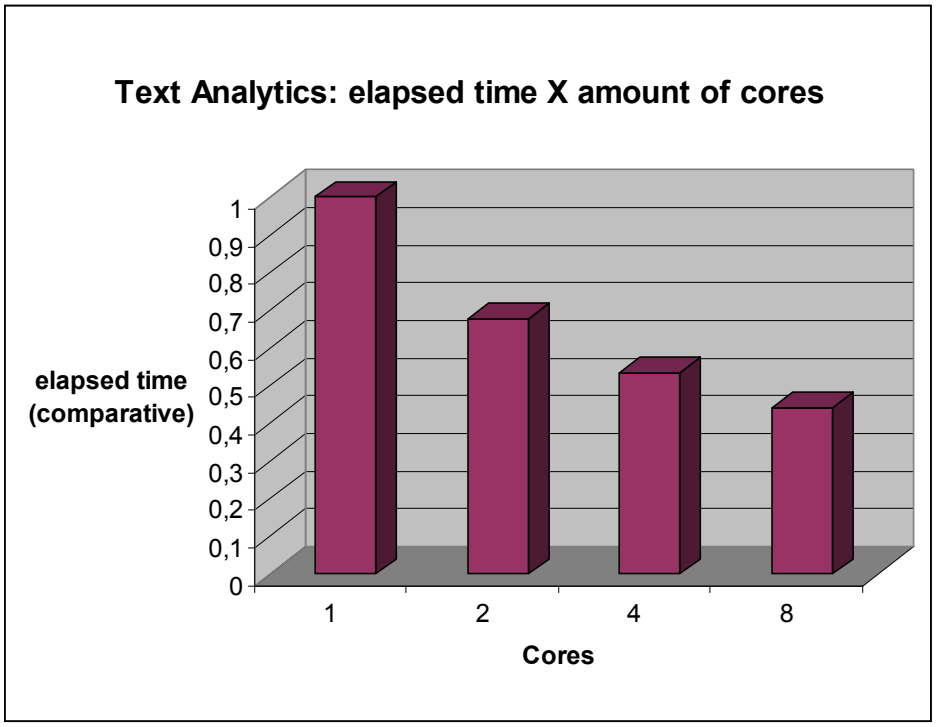

Figure 3-40 Elapsed time to complete stream analysis with different core configurations

As shown in the results in Figure 3-40, the gains by adding more cores to the processing of a text stream are not linear. Not all parts of the processing of our text stream can be parallelized.

In the previous results, we obtained a decrease of elapsed time of 33% by adding one more core compared to the initial scenario. Elapsed time decreased 47% by adding three more cores, and elapsed time decreased 56% by adding seven more cores.

Version: The tests in this section are run with SPSS Modeler Server 15 Premium with Fix Pack 1.

3.3 SPSS integration with Cognos

This section describes the integration of SPSS and Cognos solutions. The integration with Cognos is a direct connection to Cognos as a data source for the IBM SPSS Modeler. You can also export the results to Cognos directly to allow Cognos to report on the results. The benefit of using Cognos as a source is that your data is all in one place and is formatted and neat, suitable for SPSS analytics. By exporting the results to Cognos, you can report on the results by using the familiar Cognos reporting formats and various types of Cognos reports, such as dashboards and active reports. See Figure 3-41 on page 126.

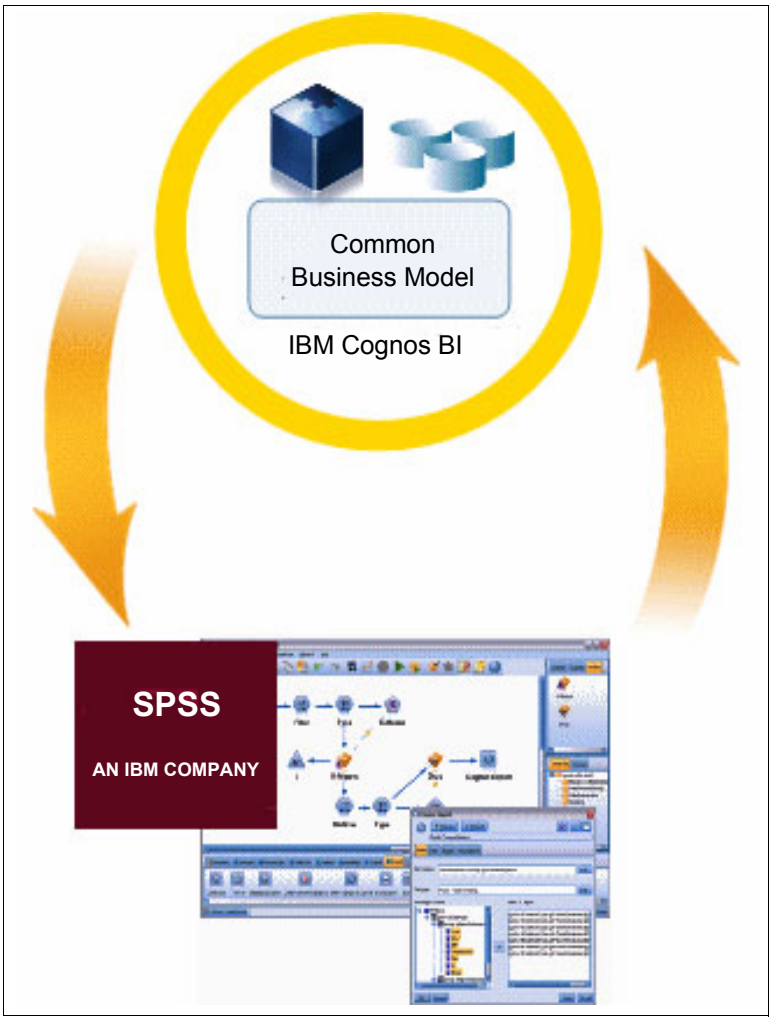

Figure 3-41 Cognos and SPSS integration

To implement this integration, you select the Cognos source node or the Cognos export node and drag it onto your stream.

Figure 3-42 on page 127 shows a basic model with a Cognos data source node and a Cognos export node.

Important: When you integrate SPSS with Cognos, remember that you also need to have an ODBC connection directly from the server to the database. This connection must have the same name as the data source on the Cognos server.

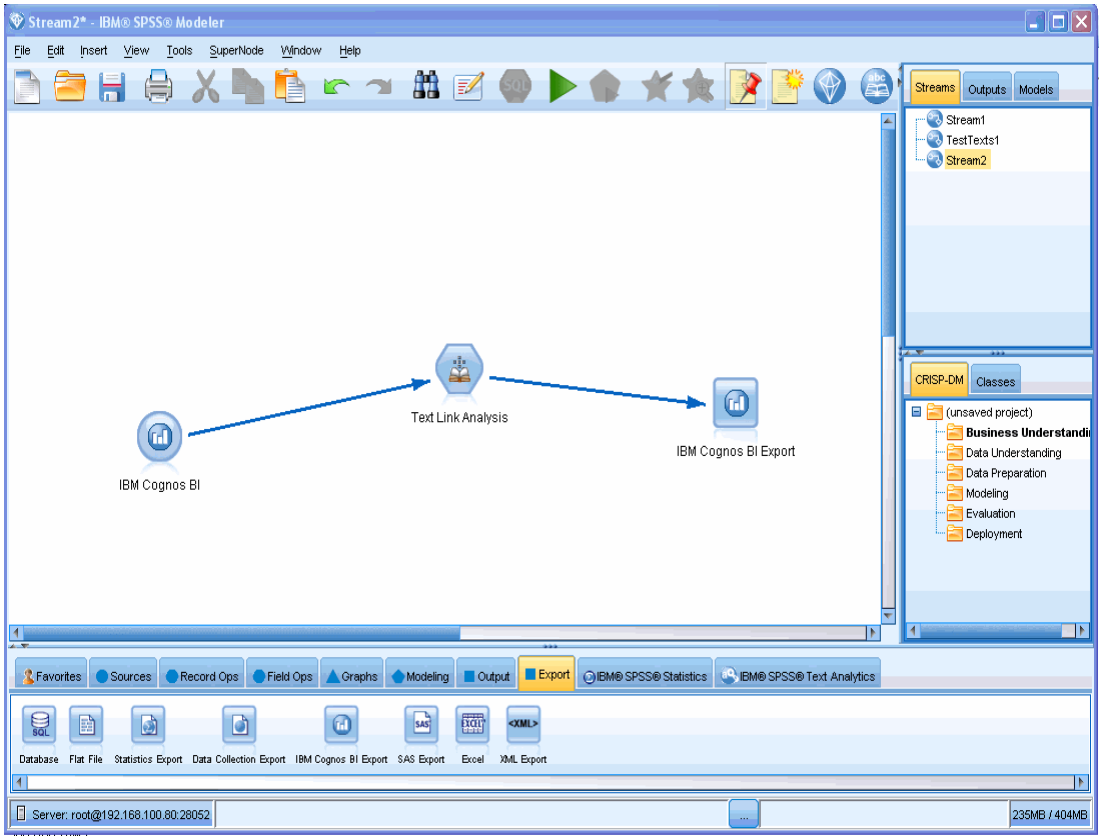

Figure 3-42 Cognos integrated stream

You then must edit the Cognos node and add the Cognos server IP address details and logon credentials, as shown on Figure 3-43 on page 128.

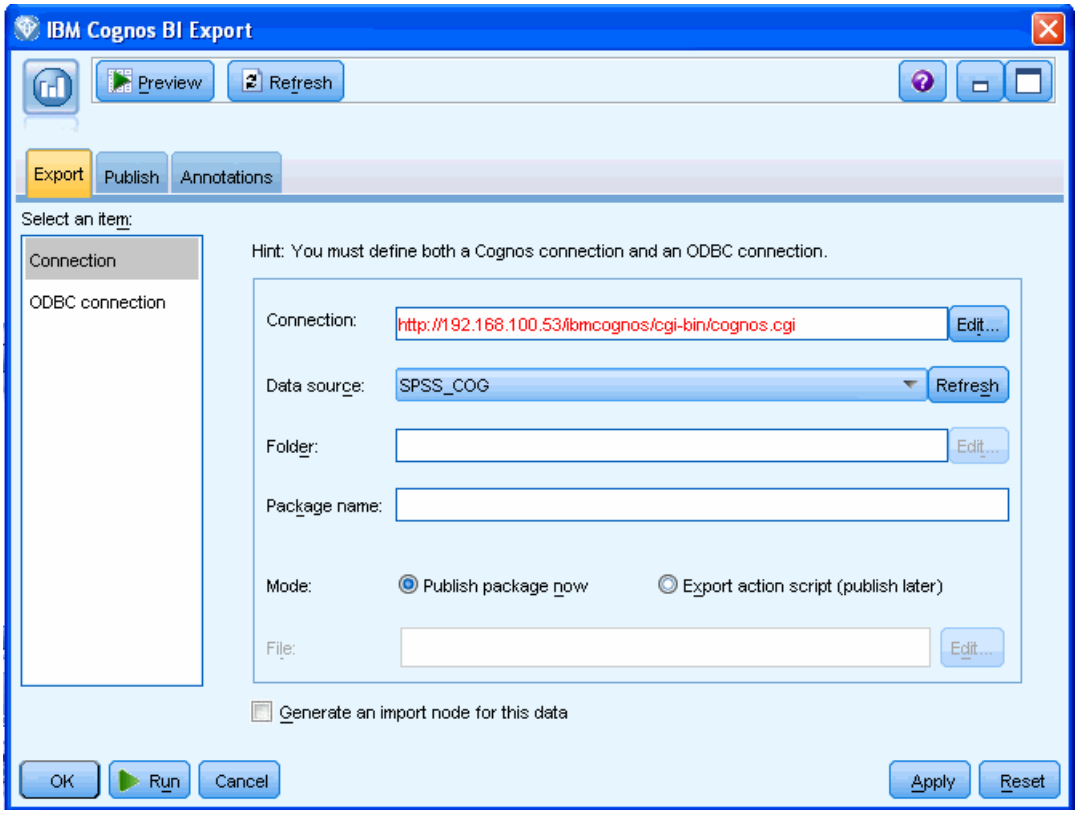

Figure 3-43 Adding a Cognos server to the Cognos node

You also must make an ODBC connection to the Cognos data warehouse from the SPSS server. This connection must have the same name and details as the Cognos data source. If these details do not match on the Cognos and the SPSS server, the integration does not work. See Figure 3-44 on page 129.

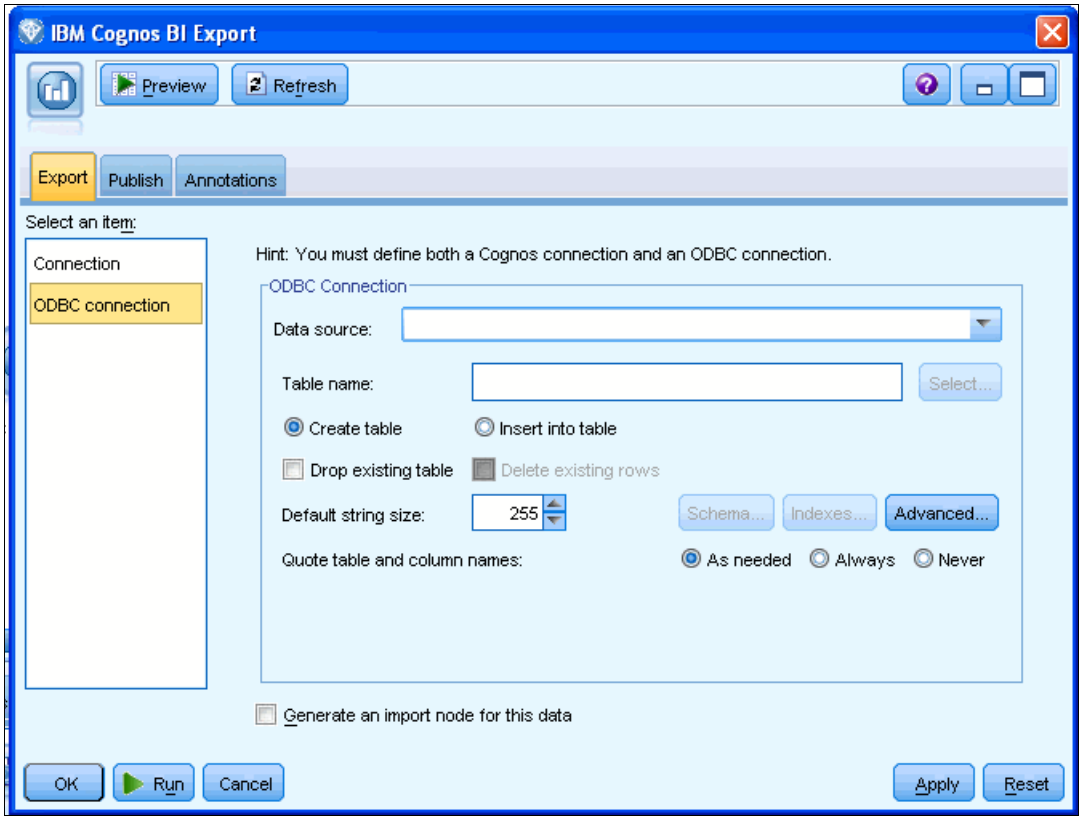

Figure 3-44 Adding an ODBC connection to the Cognos node

The benefit of using Cognos as a source is that your data is all in one place, formatted, and neat, suitable for SPSS analytics. Exporting the results to Cognos has the benefit of being able to report on the results by using the familiar reporting format and various types of Cognos reports that your company already uses.

Important: To either export to Cognos or use Cognos as a data source, an ODBC connection to the Cognos data warehouse must be established in addition to the Cognos Server connection.

3.3.1 Setting up an ODBC connection to DB2

To have the SPSS Modeler Server connect to a Cognos data source, you must install the SPSS Data Access Pack (DAP) on the SPSS Modeler Server system. We used DAP Version 6.1 SP3 for out tests.

The Data Access Pack is targeted to multiple platforms, and all of the binaries are in a single compressed file. The process that is shown in Example 3-10 is for an AIX installation of DAP. The last line of the example outlines the binary executable file for installation on an AIX system.

Example 3-10 Extracting the SPSS Data Access Pack and identifying the AIX installation binary

```
root@spss03:/tmp/dap # ls
spss_dap_6.1_sp3_mp_en.zip
root@spss03:/tmp/dap # unzip spss_dap_6.1_sp3_mp_en.zip
Archive:  spss_dap_6.1_sp3_mp_en.zip
```

```
   creating: Aix/
  inflating: Aix/SDAP_6.1_Aix64.bin
   creating: HP-UX/
  inflating: HP-UX/SDAP_6.1_HPUX.bin
   creating: InstallationInstructions/
   creating: InstallationInstructions/Linux/
  inflating: InstallationInstructions/Linux/IBM SPSS Data Access Pack Installation
Instructions.pdf
   creating: InstallationInstructions/UNIX/
  inflating: InstallationInstructions/UNIX/IBM SPSS Data Access Pack Installation
Instructions.pdf
   creating: InstallationInstructions/Windows/
  inflating: InstallationInstructions/Windows/IBM SPSS Data Access Pack
Installation Instructions.pdf
   creating: Linux/
  inflating: Linux/SDAP_6.1_Linux32.bin
  inflating: Linux/SDAP_6.1_Linux64.bin
   creating: Solaris/
  inflating: Solaris/SDAP_6.1_Solaris.bin
   creating: Win32/
  inflating: Win32/SDAP_6.1_win32.exe
   creating: Win64/
  inflating: Win64/SDAP_6.1_win64.exe
   creating: zLinux/
  inflating: zLinux/SDAP_6.1_zLinux.bin
root@spss03:/tmp/dap # cd Aix/
root@spss03:/tmp/dap/Aix # ls
SDAP_6.1_Aix64.bin
```

The `SDAP_6.1_AIX64.bin` file is a compressed executable file. After it is run, it prompts you for an installation directory and extracts its contents. Example 3-11 shows this step of the DAP installation on AIX. We used a different target installation directory in our tests other than the default one. We suggest that you install the DAP on to the file system that hosts your application binary files. The last command that is run in Example 3-11 is a listing of the contents of the installation directory.

Example 3-11 Executing the DAP installation binary on AIX

```
root@spss03:/tmp/dap/Aix # ls
SDAP_6.1_Aix64.bin
root@spss03:/tmp/dap/Aix # chmod 755 SDAP_6.1_Aix64.bin
root@spss03:/tmp/dap/Aix # ./SDAP_6.1_Aix64.bin
Preparing to install...
Extracting the JRE from the installer archive...
Unpacking the JRE...
Extracting the installation resources from the installer archive...
Configuring the installer for this system's environment...

Launching installer...


Graphical installers are not supported by the VM. The console mode will be used
instead...

===============================================================================
```

```
IBM SPSS Data Access Pack 6.1                    (created with InstallAnywhere)
-------------------------------------------------------------------------------

Preparing CONSOLE Mode Installation...

===============================================================================

    International Program License Agreement

    Part 1 - General Terms

[... snip ...]

Press Enter to continue viewing the license agreement, or enter "1" to
   accept the agreement, "2" to decline it, "3" to print it, or "99" to go back
   to the previous screen.: 1

===============================================================================
Choose Install Folder
---------------------

Where would you like to install?

  Default Install Folder: /SDAP61

ENTER AN ABSOLUTE PATH, OR PRESS <ENTER> TO ACCEPT THE DEFAULT
      : /SDAP61_redbook

INSTALL FOLDER IS: /SDAP61_redbook
   IS THIS CORRECT? (Y/N): Y

===============================================================================
Installing...
-------------

 [==================|==================|==================|==================]
 [------------------|------------------|------------------|------------------]
root@spss03:/tmp/dap/Aix # ls /SDAP61_redbook
IBM_SPSS_Data_Access_Pack_6.1_InstallLog.log  license
ODBC64README.TXT                              sdap_6.1_aix64.tar.gz
documentation
```

After you extract it to the target installation directory, you need to extract the ODBC drivers that are in the `sdap_6.1_aix64.tar.gz` file. This file also contains the configuration scripts that you need to make the SPSS Modeler Server recognize the ODBC drivers (Example 3-12).

Example 3-12 Extracting the DAP ODBC drivers

```
root@spss03:/SDAP61_redbook # ls
IBM_SPSS_Data_Access_Pack_6.1_InstallLog.log  license
ODBC64README.TXT                              sdap_6.1_aix64.tar.gz
documentation
root@spss03:/SDAP61_redbook # gunzip sdap_6.1_aix64.tar.gz
root@spss03:/SDAP61_redbook # tar xvf sdap_6.1_aix64.tar
x .
x ./help
```

```
x ./help/salesforce_connect.txt, 622 bytes, 2 tape blocks

[... snip ...]
x ./ODBC64README.TXT, 30914 bytes, 61 tape blocks
x ./tools
x ./odbc.csh, 177 bytes, 1 tape blocks
x ./NOTICES.TXT, 54149 bytes, 106 tape blocks
root@spss03:/SDAP61_redbook # ls
DDXE.LIC                                    lib
IBM_SPSS_Data_Access_Pack_6.1_InstallLog.log  license
NOTICES.TXT                                 locale
ODBC64FIXES.TXT                             odbc.csh
ODBC64README.TXT                            odbc.ini
bin                                         odbc.sh
bind                                        odbcinst.ini
documentation                               samples
fixes.txt                                   sdap_6.1_aix64.tar
help                                        setodbcpath.sh
include                                     tools
```

The ODBC drivers are in the `lib` folder. The driver for DB2 is named `XEdb225.so` in this version of the SPSS Data Access Pack (SDAP).

The next step is to run the setup scripts to set up the environment variables that the SPSS Modeler Server needs to know about to use the ODBC drivers. To accomplish this step, run the `setodbcpath.sh` script that was extracted in the previous step. This script identifies and saves the correct environment variables to the `odbc.sh`. After this step completes, go to the installation directory of the SPSS Modeler Server and edit the `modelersrv.sh` file. Add the four lines that are emphasized as shown in Example 3-13. Remember to use the path into which you installed DAP with the line that sources the `odbc.sh` script.

Example 3-13 Enabling the IBM SPSS Modeler to find and use the ODBC drivers

```
#!/bin/sh
##############################################################################
# Licensed Materials - Property of IBM
# IBM SPSS Products: Modeler Common
# (C) Copyright IBM Corp. 1994, 2011
# US Government Users Restricted Rights - Use, duplication or disclosure
restricted by GSA ADP
# Schedule Contract with IBM Corp.
##############################################################################
INSTALLEDPATH=/usr/IBM/SPSS/ModelerServer/15.0
DLLIBPATH=/usr/IBM/SPSS/ModelerServer/15.0/bin
SCLEMDNAME=modelersrv_15_0

. /opt/SDAP61_redbook/odbc.sh
ulimit -d unlimited
ulimit -m unlimited
ulimit -f unlimited

MALLOCOPTIONS=pool:0x20000000,multiheap:32; export MALLOCOPTIONS

[...snip...]
```

The last step before you start the SPSS Modeler Server to use the ODBC drivers to change the SPSS Modeler Server default ODBC handling library to use the ODBC data direct library. These libraries are inside the `bin` directory of your SPSS Modeler Server installation. Perform the steps that are described in Example 3-14 to change the library.

Example 3-14 Setting the SPSS Modeler Server to use the ODBC data direct library

```
root@spss03:/usr/IBM/SPSS/ModelerServer/15.0/bin # mv libspssodbc.so
libspssodbc.so.bak
root@spss03:/usr/IBM/SPSS/ModelerServer/15.0/bin # ln -s libspssodbc_datadirect.so
libspssodbc.so
root@spss03:/usr/IBM/SPSS/ModelerServer/15.0/bin #
root@spss03:/usr/IBM/SPSS/ModelerServer/15.0/bin # ls -la libspssodbc.so
lrwxrwxrwx    1 root     system           25 Jun 28 16:44 libspssodbc.so ->
libspssodbc_datadirect.so
```

Stop and start your SPSS Modeler Server for the changes to take effect. Upon the successful completion of these steps, you can see options when you open a connection to a database by using an ODBC connection within your IBM SPSS Modeler client. See Figure 3-45.

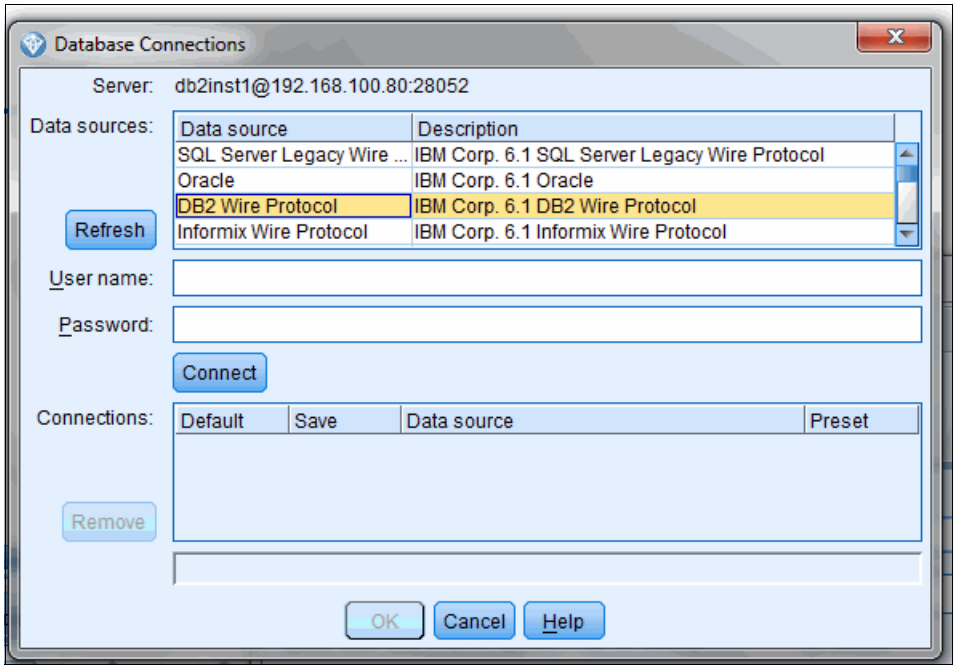

Figure 3-45 SPSS Modeler Server that uses DAP ODBC drivers

Credentials: When you use the IBM SPSS Modeler client to log on to the AIX SPSS Modeler Server, use root credentials to gain access to the system. However, when you connect to a data source as shown in Figure 3-45, use the database credentials.

Your last step to successfully connect to a DB2 database from within SPSS by using the ODBC drivers is to create an entry in the `odbc.ini` file that points to your database.

Two ways are available to establish a connection to a DB2 database by using the ODBC driver layer:

- Use the wired protocol.

  In this scenario, you configure the SPSS Modeler Server where you installed SDAP to directly query the DB2 server.

- Use the non-wired protocol.

  In this scenario, you configure the SPSS Modeler Server where you installed SDAP to use a local DB2 client to connect to a DB2 server.

We describe the wired protocol method only because it performs better than the non-wired protocol method. To use non-wired protocols means that your ODBC driver talks to the DB2 client, which in turn talks to the server through a catalog mapping. This additional communication is the reason that the non-wired protocol does not perform as well as a wired protocol.

To add an entry to the `odbc.ini` file, you first define it in the file header section and then create the entry, as specified in Example 3-15.

Example 3-15 Creating an ODBC entry in the odbc.ini file

```
[ODBC Data Sources]
SQL Server Legacy Wire Protocol=IBM Corp. 6.1 SQL Server Legacy Wire Protocol
Oracle=IBM Corp. 6.1 Oracle
DB2 Wire Protocol=IBM Corp. 6.1 DB2 Wire Protocol
Informix Wire Protocol=IBM Corp. 6.1 Informix Wire Protocol
Oracle Wire Protocol=IBM Corp. 6.1 Oracle Wire Protocol
Sybase Wire Protocol=IBM Corp. 6.1 Sybase Wire Protocol
Teradata=IBM Corp. 6.1 Teradata
SQL Server Wire Protocol=IBM Corp. 6.1 SQL Server Wire Protocol
MySQL Wire Protocol=IBM Corp. 6.1 MySQL Wire Protocol
PostgreSQL Wire Protocol=IBM Corp. 6.1 PostgreSQL Wire Protocol
Greenplum Wire Protocol=IBM Corp. 6.1 Greenplum Wire Protocol
Salesforce=IBM Corp. 6.1 Salesforce
SPSSCOG=IBM DB2 ODBC Driver 1

[... snip ...]

[SPSSCOG] 2
Driver=/opt/SDAP61/lib/XEdb225.so
Description=IBM DB2 ODBC Driver
Database=SPSS_COG
IpAddress=192.168.100.50
LogonID=db2inst1
Password=db2inst1
TcpPort=50000
```

In Example 3-15, the last line of the header, inside the *[ODBC Data Sources]* entry, describes a new data source named *SPSSCOG*. This name is a name that you assign to it. However, when you later create the entry in the file, the entry name must be the same name that you used to describe it earlier, as shown by the *[SPSSCOG]* entry later in **2**. You can give this new data source any description you want after the "=" character as shown by **1**.

There are several important pieces in the entry creation:

- Driver: Use the DB2 driver from SDAP. The version might vary according to the SDAP version that you use.
- Description: This description is a simple description. We used the same description for both the data source definition and the entry creation.
- Database: This name is the database name in the DB2 server to which you want to point. This database must be an existing database or the communication does not work.
- IpAddress: The IP address of the DB2 server.
- LogonID: The user ID that has access to the database as described by the previously mentioned database field.
- Password: The password of this user ID with access to the database.
- TcpPort: The port on which the DB2 server listens for incoming connections.

After you modify these fields, restart the SPSS Modeler Server for the changes to take effect.

Now, when you open the IBM SPSS Modeler client, you can see the newly created data source as a database connection option, as shown in Figure 3-46.

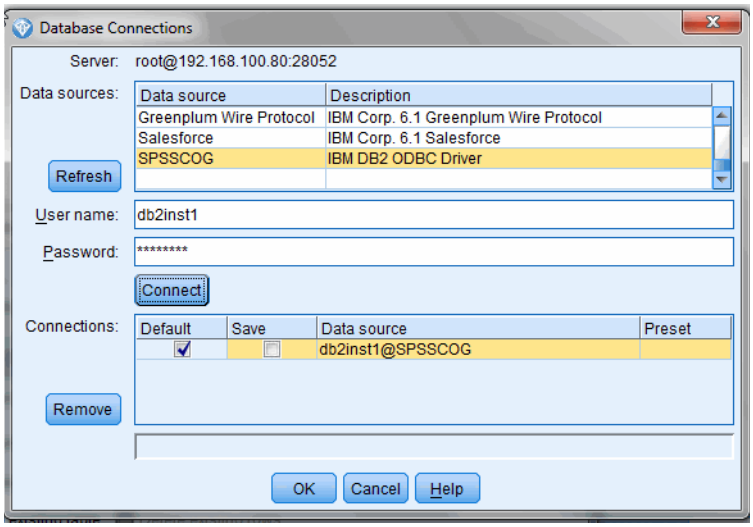

Figure 3-46 Successful connection to DB2 through the SDAP ODBC drivers

3.3.2 Troubleshooting the ODBC connection with SDAP

There are certain points in the configuration of an ODBC connection that might require debugging. We explain several of the known points based on our experience.

Pluggable Authentication Module

Your AIX system on which the SPSS Modeler Server is running might be configured to use Pluggable Authentication Module (PAM) as the authentication module. In this case, PAM must be configured to grant access to an incoming IBM SPSS Modeler client logon request.

To verify whether your PAM configuration is preventing the Modeler client from accessing the system, configure your syslog daemon to log debug information on the system. If your logon attempt is blocked, you see a message in syslog that is similar to the message that is shown in Example 3-16 on page 136.

Example 3-16 PAM log messages

```
Jul  3 10:51:44 spss03 auth|security:debug
/usr/IBM/SPSS/ModelerServer/15.0/modelersrv_15_0 PAM: pam_authenticate: error
Authentication failed
```

Tracing the activity of the ODBC drivers

It is possible to configure SDAP to trace the activity of the ODBC drivers. Edit the `odbc.ini` file in the *[ODBC]* stanza. Configure the tracing parameter to activate tracing (set it to `1` as opposed to `0`), as shown in Example 3-17. The trace file is stored within the “data” folder of your SPSS Modeler Server installation on AIX.

Example 3-17 Turning on tracing of the activity of the ODBC drivers

```
[ODBC]
IANAAppCodePage=4
InstallDir=/opt/SDAP61
Trace=1
TraceFile=odbctrace.out
TraceDll=/opt/SDAP61/lib/XEtrc25.so
```

Modeler Client connection refused error

Your Modeler Client might inform you of a connection refused error when you attempt to establish an ODBC connection to a database. Check that your DB2 database entry, as explained in Example 3-15 on page 134, points to the DB2 server IP address, not to the Modeler Server IP address, if these two IP addresses differ.

4

Preferred practices

In this chapter, we describe the preferred practices that we followed. This chapter contains the following topics:

- PowerVM configuration and tuning preferred practices
- AIX 7.1 configuration and tuning preferred practices
- Using POWER7 features in an analytics environment effectively
- SPSS configuration and tuning preferred practices
- Tools and their usage for performance analysis
- Tools and their usage for performance analysis

4.1 PowerVM configuration and tuning preferred practices

IBM Power Systems servers are designed to run mission-critical applications in highly, virtualized, consolidated operating environments that offer the flexibility that analytics workloads need. IBM Power Systems are a leader in performance, energy efficiency, flexible virtualization features, and reliability, availability, and serviceability (RAS) capabilities. Companies that deploy IBM Cognos Business Intelligence (BI) and other business analytics workloads can benefit from these superior features. The use of advanced virtualization features can help lower the total cost of ownership (TCO), can provide flexibility in systems management, and can ensure that business analysts and other BI users receive continued service without interruptions.

IBM Power Systems offers a feature called *IBM PowerVM* to achieve higher resource utilization rates through the dynamic allocation of processors and memory.

With the flexibility and features offered by PowerVM, there are a number of choices you can make. There is more than one correct way to design and implement your virtual environment. Reaching the preferred practices for your environment is a matter of choosing the option that best fits the workloads that you plan to run.

This section describes preferred practices and tools to use in fine-tuning the environment to maximize performance for a specific application or workload.

Chapter 2, “Implementing Cognos on IBM Power Systems” on page 49 and Chapter 3, “Implementing IBM SPSS on IBM Power Systems” on page 79 explain the various components of Cognos BI and Statistical Package for the Social Sciences (SPSS)™ software suits and their partition configurations. But the actual resources that are associated with these partitions might vary depending on the volume of the data and type of output requirements of the individual organizations. In addition to the basic configuration details that are provided, you can also use a tool called System Planning Tool (SPT) to configure and plan your environment.

4.1.1 System Planning Tool

The *System Planning Tool* (SPT) is a browser-based application that helps you plan, design, and validate system configurations. It is a good practice to use this tool to create your initial setup and to distribute the hardware resources among the dedicated and virtual partitions. This tool has many functions and built-in checks to assist you with your system planning.

It is also a good practice to use this tool to plan your partitions. You can deploy the system plans that are generated with the SPT to create the partition profiles on your Hardware Management Console (HMC) or the Integrated Virtualization Manager (IVM).

Another benefit of using the SPT is that it can help you calculate how much of the system memory the Power Hypervisor uses, depending on your configuration. This way, you can plan how much system memory you have left that is available for use in your partitions. Finally, it is also a good practice to print the SPT validation report to ensure that you meet all the system requirements. The SPT is at this website:

`http://ibm.com/systems/support/tools/systemplanningtool`

After you allocate the partitions and resources for the various components in your software suit by following the guidelines that are provided earlier chapters, it is a good practice to fine-tune the resource allocation and tuning parameters of various logical partitions (LPARs). This fine-tuning helps you to maximize the performance of each component and partition in your environment. There are a few useful tools that can be used to fine-tune your PowerVM environment:

- Virtualization Performance Advisor to optimize an LPAR
- Virtual I/O Server (VIOS) advisor to optimize the VIOS

The goal of the advisor tools is not to provide another monitoring tool. These expert systems view performance metrics that are already available to you and assess and suggest based on the expertise and experience available in the IBM systems performance group.

Download and setup instructions are at the following locations:

- Virtualization Performance Advisor (VPA):

  `https://www.ibm.com/developerworks/wikis/display/WikiPtype/PowerVM+Virtualization+performance+advisor`

- VIOS advisor:

  `http://www.ibm.com/developerworks/wikis/display/WikiPtype/VIOS+Advisor`

To get better recommendations and tuning from these tools, you must run them during a peak workload on the partition.

The VPA tool generates a `.xml` output file at the end of its run. The output `.xml` file has many sections that indicate the current values, and it also indicates the recommended values to maximize the performance.

For example, if an LPAR has imbalanced processor and memory allocation, the output `.xml` file that is generated by the virtualization performance advisor indicates this imbalance, as shown in Figure 4-1 on page 140.

LPAR PROCESSOR OPTIMIZATION

	Name	Current Value	Recommended Value	First Observed	Last Observed	Risk 1=lowest 5=highest	Impact 1=lowest 5=highest
	Lpar Placement Optimization	Placement Global Domain / Chip Domain / Memory / CPU 0 0 63228.560 0-11 0 6 60756.000 1 1 63246.000 12-15 1 2 63246.000 16-19 2 3 63223.000 20-23 2 4 63478.690 24-27 3 5 61503.000 28-31 3 7 52539.000	-		-	NA	NA
	Lpar Placement Optimization	Only memory assigned from Global Domain 0, Chip Domain 6 . Only memory assigned from Global Domain 3, Chip Domain 7 .	Imbalanced resource allocation. However changing this allocation requires help from IBM. Also refer to P7 Virtualization Performance best practice document.		-	NA	NA
	Remote/Distant Memory Access	Only memory assigned from Global Domain 0, Chip Domain 6 . Only memory assigned from Global Domain 3, Chip Domain 7 .- Local Memory access - 0.00 and Distant memory access - 9275.13 Distant memory access is high.	Imbalanced resource allocation. However changing this allocation requires help from IBM. Also refer to P7 Virtualization Performance best practice document.-		-	NA	NA
	SMT Effectiveness	SMT-4 is set	-	Tue Mar 6 20:14:28 2012	-	NA	NA
	Virtual Processor Folding Optimization	Virtual Processor Folding Threshold - 49%	Rerun when lpar is busy	Tue Mar 6 20:14:33 2012	-	NA	NA

Figure 4-1 LPAR processor optimization

It is important to balance the allocation of memory and processor to attain higher memory affinity. When the VPA reports that the LPAR has imbalanced resource allocation, consider the following actions to get the balanced resources to the LPAR.

Important: It is a preferred practice to activate the critical LPAR, such as the Cognos server, first.

LPARs that require high performance, such as the Cognos server, can be forced to get the best resources by activating the critical LPAR first before activating any other LPARs, including the VIOS. The VIOS hosts I/O for all the LPARs so it must be up to service other LPARs. However, the LPAR resource assignment at the hypervisor occurs before any access to storage. Therefore, the critical LPAR can be activated first, which goes through the hypervisor resource assignment and waits for the VIOS to be activated before proceeding further.

PowerVM, when all the resources are free (an initial machine state or reboot of the CEC), allocates memory and cores as optimally as possible. At the partition start time, PowerVM is aware of all the LPAR configurations, so the placement of processors and memory are irrespective of the order of activation of the LPARs.

However, after the initial configuration, the setup might not stay static. There are numerous operations that take place:

- Reconfiguring the existing LPARs with new profiles
- Reactivating the existing LPARs and replacing them with new LPARs
- Adding and removing resources to LPARs dynamically (dynamic LPAR operations)

Any of these changes can result in memory fragmentation that causes the LPARs to spread across multiple domains.

From an AIX LPAR, the **lssrad** command can be used to display the number of domains to which an LPAR is spread.

The following example shows the **lssrad** syntax:

```
lssrad -av
```

If all the cores and memory are in a single book:

```
REF1 SRAD    MEM     CPU
0    0     31806.31  0-31
     1     31553.75  32-63
```

REF1 is the second-level domain, in this case (Power 795), REF1 refers to book. Scheduler resource allocation domain (SRAD) refers to socket numbers. However, the **lssrad** command does not report the actual physical domains (book numbers and chip numbers). The output of the **lssrad** command indicates that the LPAR allocated 16 cores from two chips in the same book.

There are ways to minimize or even eliminate the spread. The spread can be minimized by releasing the resources that are currently assigned to the deactivated LPARs.

Resources of an LPAR can be released by using the following commands:

```
chhwres -r mem -m <system_name> -o r -q <num_of_Mbytes> --id <lp_id>
chhwres -r proc -m <system_name> -o r --procunits <number> --id <lp_id>
```

The first command frees the memory and the second command frees cores.

If all the partitions are inactive, there is another way to clear the resources of all the existing configurations before you create a configuration. In this situation, all the resources of all the partitions can be cleared from the HMC with the following procedure:

1. Shut down all partitions.
2. Create the all-resources partition.
3. Activate the all-resources partition.
4. Shut down the all-resources partition.
5. Delete the all-resources partition.

Fragmentation due to frequent movement of memory or processors between partitions can be avoided through planning ahead of the time. Dynamic LPAR actions can be done in a controlled way so that the performance impact of resource addition or deletion is minimal. Planning for growth helps to alleviate the fragmentation that is caused by dynamic LPAR operations. By knowing the LPARs that need to grow or shrink dynamically and by placing them with LPARs that can tolerate nodal crossing latency (less critical LPARs) are approaches to handle the changes of critical LPARs dynamically. In such a configuration, when growth is needed for the critical LPAR, the resources that are assigned to the non-critical LPAR can be reduced so that the critical LPAR can grow.

For further details about the preferred practices to be followed in a shared processor LPAR (SPLPAR) environment, see the *POWER7 Virtualization Best Practice Guide*:

https://www.ibm.com/developerworks/wikis/download/attachments/53871915/P7_virtualization_bestpractice.doc?version=1

If the VPA report indicates that an LPAR requires more memory, add the memory to the LPAR to maximize the throughput. Our basic Cognos configuration in Chapter 2, “Implementing Cognos on IBM Power Systems” on page 49 indicates 4 GB of RAM for the Cognos server component LPAR, VPA, and other monitoring tools, such as nmon. Cognos performs better with more memory. The Cognos server that is configured with 8 GB of RAM maximized throughput in the number of scores/second and it also scaled to a high number of users with the same performance.

To get a similar XML report on the VIOS server, the VIOS Advisor needs to be used. The XML report that is generated by the VIOS Advisor includes many parameters that are relevant to the VIOS environment.

As a preferred practice in configuring your LPARs, start with the basic configurations that are suggested in Chapter 3, “Implementing IBM SPSS on IBM Power Systems” on page 79. Then, fine-tune their resource allocations by using the tools that are suggested in this chapter.

Active Memory Sharing

Based on our observations with Cognos and SPSS testing, not all LPARs are equally loaded or utilized at a specific time. These environments can benefit from implementing *Active Memory Sharing* (AMS) because it helps to use system memory better.

AMS is a PowerVM advanced memory virtualization technology that provides system memory virtualization capabilities to IBM Power Systems. It allows multiple LPARs to share a common pool of physical memory.

A system administrator configures an LPAR with enough memory to satisfy the workload, and to avoid paging activity. However, such memory demands might be needed for a short time only, for example, during workload peak times. At times, the changes in these workloads can be difficult to predict. When a system is configured with AMS, it automatically adjusts the amount of memory that is needed by the virtual client, depending on the workload.

With dedicated memory, the allocation of memory to each partition is static. The amount of memory, regardless of demand, does not change. With AMS, the amount of memory that is available to a partition changes over the run time of the workload, based on the memory demands of the workload.

As a preferred practice, we suggest that you use AMS in an environment that supports workloads for different regions. AMS is suggested for geographically separated workloads where LPAR memory demands differ for day and night or for workloads that consist of many LPARs that are used sporadically.

For further details about AMS and how to implement it in an analytics environment, see *Exploiting IBM PowerVM Virtualization Features with IBM Cognos 8 Business Intelligence*, SG24-7842-00.

Active Memory Deduplication

Active Memory Deduplication is a PowerVM technology that minimizes the existence of identical memory pages in the main memory space. When you run workloads on traditional partitions, multiple identical data is saved across different positions in main memory. Active Memory Deduplication can improve the use of physical memory.

As a preferred practice, we suggest that you enable Active Memory Deduplication if you use AMS and a VIOS that is not processor-constrained.

The Power Hypervisor reserves extra memory for page tracking and the calculation of the deduplication table ratio. Therefore, it is suggested to keep the maximum shared memory pool size parameter to a minimum.

For more information about the setup and configuration of Active Memory Deduplication, see *Power Systems Memory Deduplication*, REDP-4827.

Live Partition Mobility

Live Partition Mobility (LPM) is a feature of PowerVM Enterprise Edition that can move live AIX, Linux, and IBM i LPARs from one frame to another without any downtime. It can also be used to move inactive LPARs. The mobility process transfers the Virtual I/O client partition environment, including the processor state, memory, attached virtual devices, and connected users.

The following list shows preferred practice scenarios, which can benefit by the use of Live Partition Mobility:

- Live Partition Mobility can balance the frame workload across multiple frames. For example, if you need to increase the processor or memory allocation for a partition, and the frame does not have the required capacity, you can move partitions to another frame to free up resources.
- Live Partition Mobility offers flexibility in repair actions and updates. For example, you can migrate all the running partitions to another frame, while you perform VIOS or frame firmware updates.
- Live Partition Mobility can be used to migrate partitions to a new frame when a hardware refresh is done, or to consolidate Power System environments. For information about how to implement Live Partition Mobility, see *IBM PowerVM Live Partition Mobility*, SG24-7460.

4.2 AIX 7.1 configuration and tuning preferred practices

AIX 7.1 is latest release of the AIX operating system that runs on IBM Power Systems. It is a preferred practice to keep the system firmware and the AIX operating system release levels up to their latest released versions.

Tip: It is advisable to keep all AIX restricted tunable variables to their defaults.

The two main areas to consider when tuning AIX for an IBM analytics application are memory management and thread handling.

The following settings are suggested settings for AIX parameters to support IBM Analytics applications:

- `AIXTHREAD_MINKTHREADS = 32`
- `AIXTHREAD_MNRATIO = 1:1`
- `AIXTHREAD_MUTEX_FAST = ON`
- `AIXTHREAD_SCOPE = S`
- `SPINLOOPTIME = 4000`
- `YIELDLOOPTIME = 20`

Processor bindings in a shared LPAR

With AIX 6.1 Technology Level (TL) 5 and AIX 7.1, *binding processors* is available to applications that run in a shared LPAR. An application process can be bound to a virtual processor in a shared LPAR. In a shared LPAR, a virtual processor is dispatched by the PowerVM hypervisor. The PowerVM hypervisor maintains three levels of affinity for dispatching, such as core-level, chip-level, and node-level affinity in eFW7.3 and later firmware versions. By maintaining affinity at the hypervisor level and in AIX, applications can achieve higher-level affinity through processor bindings.

The following example demonstrates how to use **`execrset`** to create an rset with CPUs 4 - 7 and start the WebSphere Application Server `instance1`:

```
execrset –c 4-7 –m 0 –e
/usr/IBM/WebSphere/AppServer/profiles/AppSrv01/bin/startServer.sh instance1
```

In addition to running the application that is attached to an rset, the *MEMORY_AFFINITY* environment variable must be set to `MCM` to assure that the application private and shared memory are allocated from memory that is local to the logical processors of the rset:

```
MEMORY_AFFINITY=MCM
```

In general, rsets must be created on core boundaries. The **`smtctl`** command can be used to determine which logical processors belong to which core, as shown in Example 4-1.

Example 4-1 Using the smtctl command

```
smtctl
This system is SMT capable.
This system supports up to 4 SMT threads per processor.
SMT is currently enabled.
SMT boot mode is not set.
SMT threads are bound to the same physical processor.
proc0 has 4 SMT threads.
Bind processor 0 is bound with proc0
Bind processor 1 is bound with proc0
Bind processor 2 is bound with proc0
```

```
Bind processor 3 is bound with proc0
proc4 has 4 SMT threads.
Bind processor 4 is bound with proc4
Bind processor 5 is bound with proc4
Bind processor 6 is bound with proc4
Bind processor 7 is bound with proc4
```

The **smtctl** output in Example 4-1 on page 144 shows that the system runs in simultaneous multithreading (SMT) 4 mode with bind processors (logical processor) 0 - 3 that belong to `proc0` and bind processors 4 - 7 that belong to `proc1`. An rset with four logical processors must be created either for processors 0 - 3 or for processors 4 - 7. To achieve the best performance with rsets that are created across multiple cores, all cores of the rset must be in the same scheduler resource allocation domain (SRAD).

The **lssrad** command can be used to determine which CPUs belong to which SRAD, as shown in Example 4-2.

Example 4-2 The lssrad command

```
lssrad -av
REF1 SRAD MEM CPU
0
   0 22397.25 0-31
1
   1 29801.75 32-63
```

Example 4-2 shows a system that has two SRADs. CPUs 0 - 31 belong to the first SRAD. CPUs 32 - 63 belong to the second SRAD. In this example, an rset with multiple cores must be created by either using the CPUs of the first or second SRAD.

Important: A user must have root authority or have CAP_NUMA_ATTACH capability to use rsets.

For more information about how to use rsets in an analytics environment, see *Best practices of IBM SPSS Collaboration and Deployment Services 4.2 on IBM POWER7/AIX 7.1*:

http://www.ibm.com/systems/be/power/advantages/

4.3 Using POWER7 features in an analytics environment effectively

The Power 780 and Power 795 can operate in two modes:

- Standard mode (maxcore mode)
- TurboCore mode

What makes this IBM high-end server truly unique is the ability to switch between its standard throughput mode and its unique, optimized TurboCore mode. In TurboCore mode, performance per core is boosted with access to both additional L3 cache and more clock speed.

In standard mode, which is also known as maxcore mode, each 8-core Single Chip Module (SCM) operates at 4.0 GHz with 32 MB of L3 cache and can scale up to 256 cores. When it operates in TurboCore mode, each 8-core SCM operates with up to four active cores per SCM at 4.25 GHz and 32 MB of L3 cache, twice the L3 cache per core that is available in maxcore mode.

In TurboCore mode, the Power 795 server can scale up to 128 cores. Each 6-core SCM operates at 3.72 GHz only in standard mode with 4 MB of L3 cache per core. Both the 8-core and 6-core SCMs are supported by 256 KB of L2 cache per core.

TurboCore mode delivers the highest performance per core, and it can help save on software costs for applications that are licensed per core. Switching in and out of TurboCore mode requires a system reboot. The entire system must operate in either standard or TurboCore mode.

If the primary goal is to get the maximum performance out of your POWER7 server, it is a preferred practice to operate it in TurboCore mode.

Simultaneous multithreading (SMT)

The POWER7 processor starts with increased simultaneous multithreading capability. Each POWER7 core can have four SMT threads.

AIX represents each SMT thread as a logical processor. Therefore, the number of logical processors in an LPAR depends on the SMT mode. For example, an LPAR with four virtual processors running in SMT4 mode has 16 logical processors. An LPAR with the same number of virtual processors running in SMT2 mode has only eight logical processors.

Table 4-1 shows the number of SMT threads and logical processors that are available in single thread (ST), SMT2, and SMT4 modes.

Table 4-1 SMT threads and logical processors

SMT mode	Number of SMT threads	Number of logical processors
ST	1	1
SMT2	2	2
SMT4	4	4

The default SMT mode on POWER7 depends on the AIX version and the compatibility mode with which the processors run. Table 4-2 shows the default SMT modes.

Table 4-2 Default SMT modes

AIX version	Compatibility mode	Default SMT mode
AIX 7.1/6.1	POWER7	SMT4
AIX 7.1/6.1	POWER6/POWER6+	SMT2
AIX 5.3	POWER6/POWER6+	SMT2

Most applications benefit from SMT. However, some applications do not scale with an increased number of logical processors on a SMT-enabled system. One way to address an application scalability issue is to change to a lower SMT mode with fewer logical processors. For example, change from SMT4 to SMT2.

In general, as a preferred practice, two cores per WebSphere Application Server instance are suggested for use in a 6-core per POWER7 processor chip hardware. It is also suggested to use two cores per WebSphere Application Server instance on systems that use 4-core or 8-core POWER7 processor chips.

4.4 SPSS configuration and tuning preferred practices

This section provides information to help with SPSS configuration and tuning.

4.4.1 IBM SPSS repository configuration and its components

IBM SPSS Collaboration and Deployment Services has many components to customize and configure the repository to the requirements of your organization. Log on to IBM SPSS Collaboration and Deployment Services Deployment Manager (Figure 4-2).

To access the browser-based IBM SPSS Collaboration and Deployment Services Deployment Manager, open any supported browser and type `http://<repository host>:<port number>/security/login`.

Specify the administrator login credentials. The credentials are established during the repository configuration.

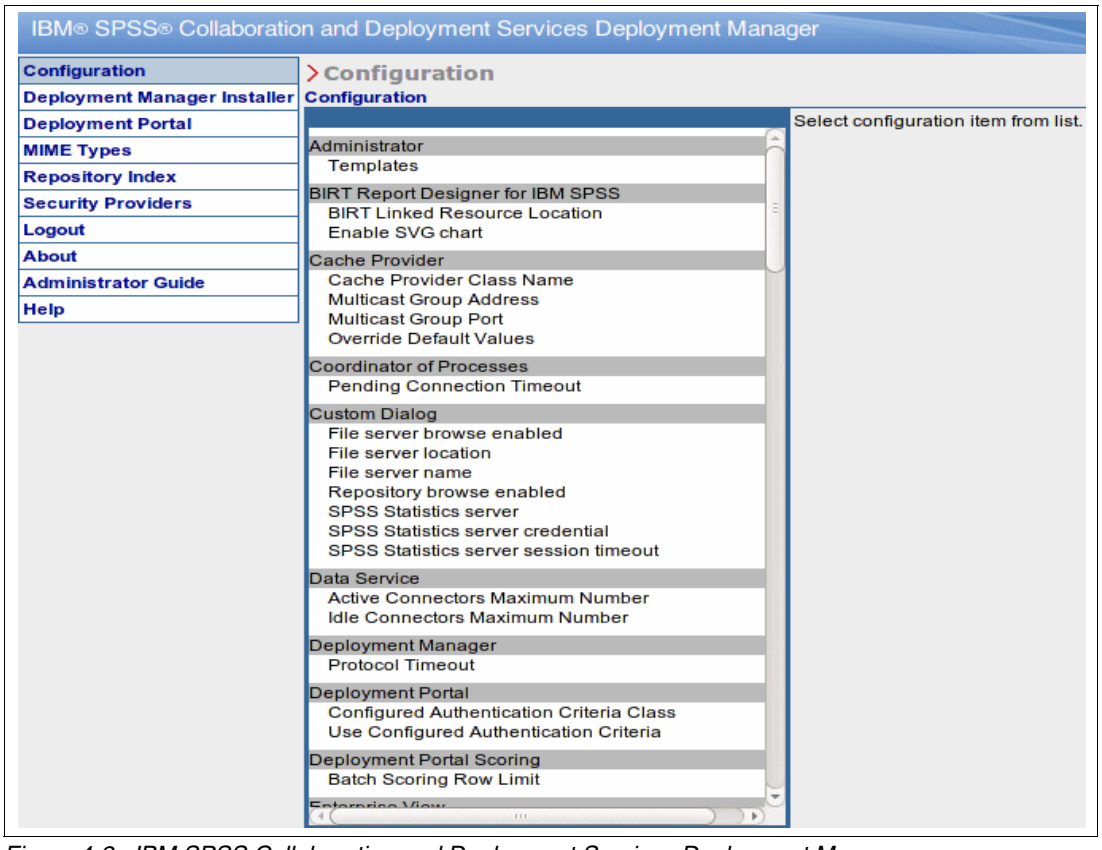

Figure 4-2 IBM SPSS Collaboration and Deployment Services Deployment Manager

Administrator

You can configure and specify the location of the templates that are used for the administrative user interface. To list the default template directory, go to **Administration** → **Templates**. You can see on the right pane the following default path of the template:

```
/opt/IBM/SPSS/Collaboration_and_Deployment_Services/4.2.1/Server/components/adminTemplates
```

BIRT Report Designer for IBM SPSS

This location on the server file system is where the external resource images, such as cascading style sheets (CSSs), are stored. The default location is:

```
/opt/IBM/SPSS/Collaboration_and_Deployment_Services/4.2.1/Server/components/reporting-engines/BIRT/Resources
```

You can modify the setting by clicking the options under BIRT Report Designer for IBM SPSS.

Configuration options for BIRT Report Designer for IBM SPSS

The following configuration options are explained.

BIRT linked resource location

This location on the server file system is where the external resource images, such as CSS, are stored.

Enable SVG chart

If you select the check box, the provider uses Multicast Group Address and Multicast Group Port to override the default values.

Cache provider

With the cache provider options, you can specify the cache provider class name. By default, this setting is "`com.spss.cache.service.ehcache.EhcacheProvider`".

Coordinator of processes

This option defines a delay (in seconds) for the coordinator of processes to time out connections, for which a get request is received but the server did not respond after specified delay. The default timeout value is set to `5` seconds.

Custom dialog

In the IBM SPSS Statistics custom dialog, you can use the configuration options to specify settings to run custom dialogs. To modify the settings, click the corresponding option under Custom Dialog in the Configuration list.

Data service

You can specify required parameters for optimizing the data services connections with the help of data service configuration options.

Deployment Manager

In Deployment Manager, you can specify the protocol timeout for communication between IBM SPSS Collaboration and Deployment Services Deployment Manager and the repository. The default protocol value is set to 180 seconds. To change the protocol timeout value, click **Deployment Manager** → **Configuration**, click **Protocol timeout**, enter the required value in seconds, and click **Set**.

To set the system-defined value, click **Use Default** to restore the default value.

Deployment portal

To specify an authentication setting and a report timeout limit, use the deployment portal. You can modify the setting with the help of the following subcomponents.

Configured authentication criteria class

The class can be used to provide authentication information for the deployment portal. The default class is "`com.spss.er.internal.configuration.ConfiguredAuthenticationImpl`".

Use configured authentication criteria

This option allows the user to pass authentication information to the deployment portal and to bypass the Login window.

Deployment portal scoring

With the Batch Scoring Row Limit configuration option, you can specify the maximum number of rows that can be batch-scored from a selected data set.

Enterprise view

You can use the enterprise view configuration option to configure the setting for the IBM SPSS Statistics data file server.

Help

With the Help configuration option, you specify the location of the Help system for IBM SPSS Collaboration and Deployment Services Deployment Manager. By default, the system uses the path that is established by the installation program.

Notification

Under the notification configuration option, you can configure various Simple Mail Transfer Protocol (SMTP)-related options, such as the email setting. You can enable notification service performance tuning.

Pager

To specify the amount of time in minutes for paged data, the pager option is used. The default pager timeout value is set to 30. Any changes in this value require a repository restart.

Process management

With the process management configuration options, you can specify job execution settings and define the web service endpoints.

Reporting

With the Reporting configuration option, you specify the path for writing out debugging information (as XML output) for visualization processing. Use this setting for capturing Visualization specifications, including the data, for debugging purposes. The default of blank means do not capture the data.

Repository

With the repository configuration options, you define the web service endpoints and toggle connection validation.

SPSS Statistics data file driver

Use this option to define the data file driver details for SPSS Statistics, which include following subcomponents.

Host or IP address

Use this option to specify the host name or IP address of the computer on which the service driver is running. If the service runs on the same machine as the SPSS Statistics Adapter, you can use the default setting of local host.

Port number

Specify the port number on which the SPSS Statistics data file driver service listens for a connection. Keep this default setting if there is no explicit port change.

Scoring service

With the scoring service configuration options, you can specify settings for scoring. To modify the settings, click the corresponding option under Scoring Service in the Configuration list.

Search

With the Search configuration option, you specify the number of hits to display per page in IBM SPSS Collaboration and Deployment Services Deployment Manager search results, and the result set size and whether searches get logged in audit views.

Security

With the Security configuration options, you can specify repository access settings. You can specify the cache login, cache session timeout, and cached login revalidation interval. You can disable clients, encrypt a password, specify lowercase user IDs, specify messages, and specify normalize principles.

Setup

With the Setup configuration option, you specify the miscellaneous setup settings for the repository, such as the URL prefix that is used in references to IBM SPSS Collaboration and Deployment Services, Java Message Service (JMS) queue setting, and JMS message bus settings.

MIME

Multipurpose Internet Mail Extensions (MIME) refers to an Internet standard that specifies how messages must be formatted so that they can be exchanged between different email systems. MIME originated as an extension of email, but it is also used by HTTP to define the content that is delivered by a server.

To configure these mappings, use the MIME Types and File Type Icons page of IBM SPSS Collaboration and Deployment Services Deployment Manager, which you can access by clicking **MIME Types** in the navigation list.

There are many predefined MIME types, such as GIF graphics files and PostScript files. It is also possible to define your own MIME types.

On the MIME Types and File Type Icons page, you can perform the following tasks:

- Add MIME-type mappings to the server.
- Edit existing MIME-type settings, including the assignment of images to files.
- Delete MIME-type mappings from the server.

Repository index

Use the repository index to optimize the IBM SPSS Collaboration and Deployment Services Repository search, By default, the old index is cleared and rebuilt when the repository is upgraded. See Figure 4-3.

To access the repository index, click **Repository Index** in the navigation list. The Content Repository Indexing page opens. As shown in Figure 4-3, there are two options: Disable Clients while indexing is running and Clear the entire index before reindexing.

Select the first option if users are still logged in to the repository. Select the second option if no users are logged in to the repository.

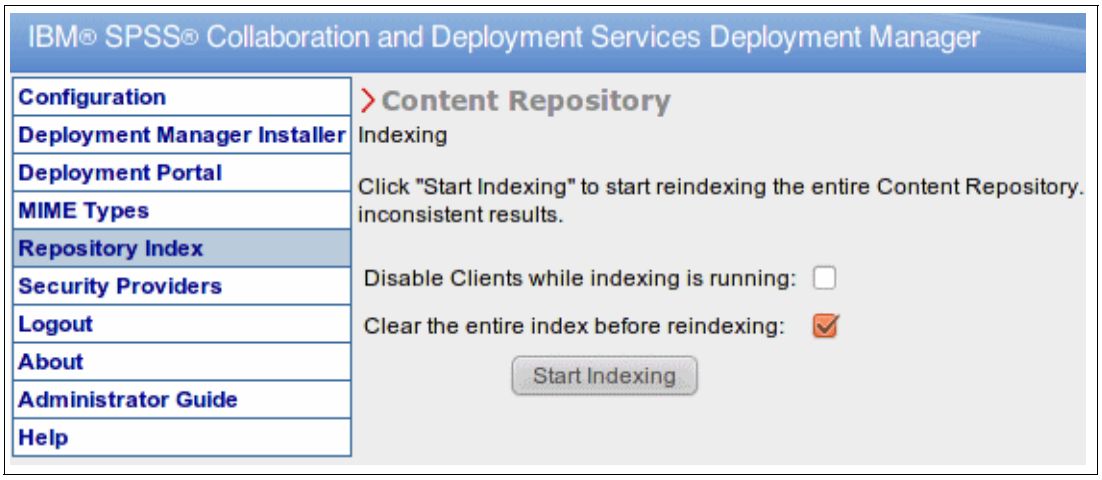

Figure 4-3 Repository Index

Click **Start Indexing**. The Content Repository Indexing Status page displays the statistics of the processed objects while the index rebuild is in progress, as shown in Figure 4-4.

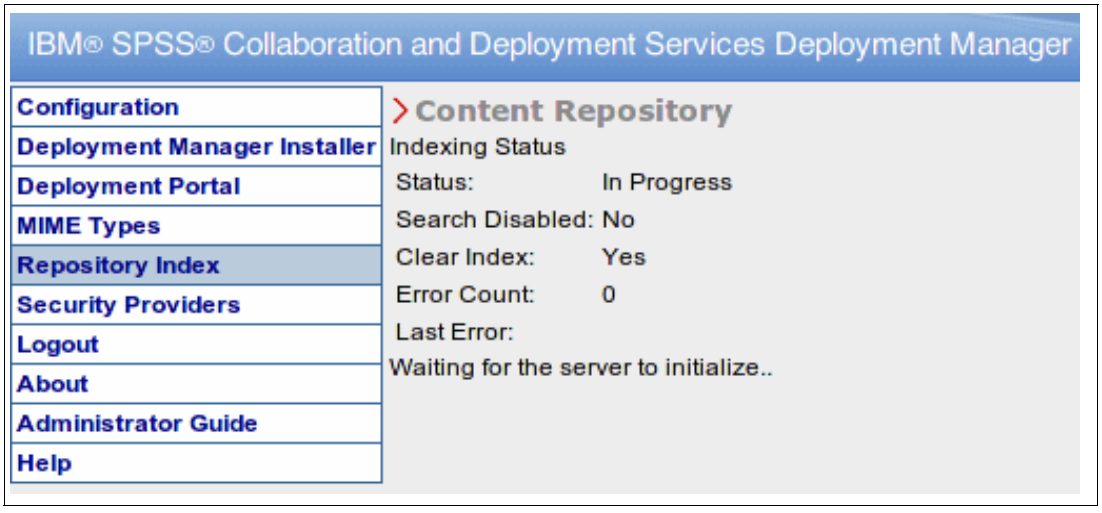

Figure 4-4 Repository Index rebuild is in progress

After the repository index rebuild completes, the status page opens, as shown in Figure 4-5 on page 152.

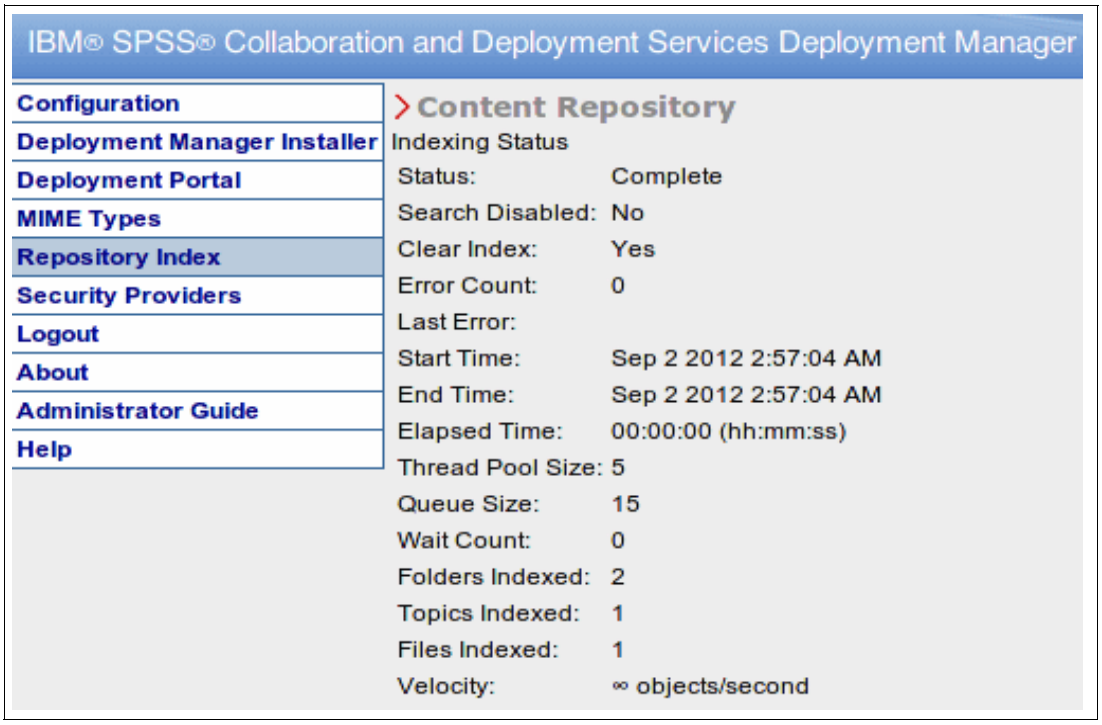

Figure 4-5 Indexing status is complete

Security providers

A *security provider* is responsible for authenticating user credentials. You can define users and groups either locally, or by using Windows Active Directory or Open Lightweight Directory Access Protocol (LDAP). There are different types of security providers through which you can configure users and groups, as shown in Figure 4-6.

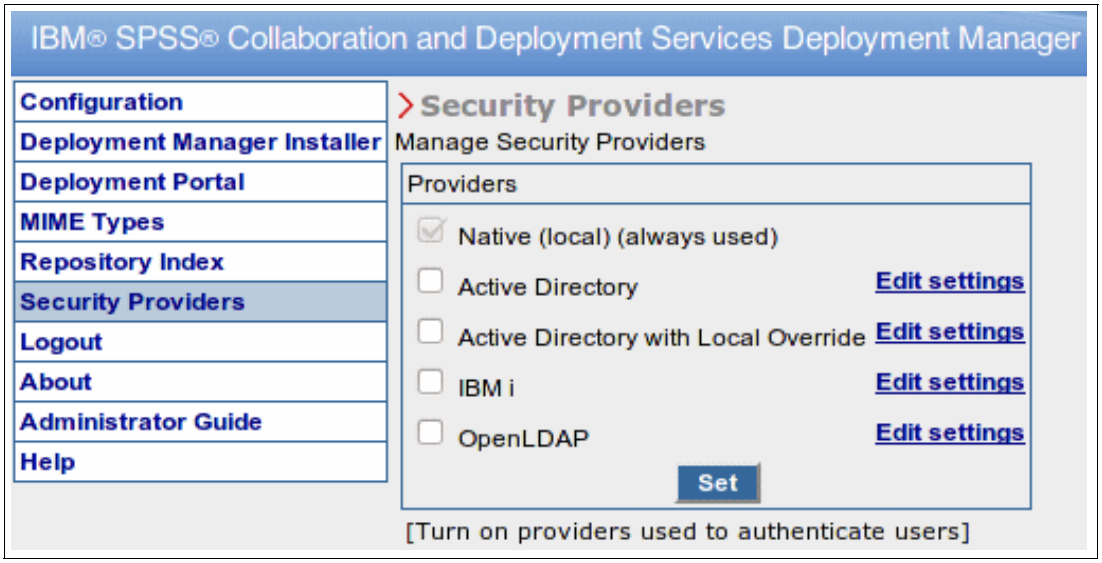

Figure 4-6 Security Providers window

4.5 Tools and their usage for performance analysis

A business analytics solution, such as Cognos and SPSS on Power, consists of multiple technology layers. We can classify them in three major areas:

- Hardware
- Operating system
- Software stack

To fine-tune the pieces of this solution, you must be able to monitor variables, such as the application response time, operating system resource utilization, and hardware resources. The hardware resources can be measured from within the operating system. In the next sections, we describe which tools you can use to monitor these aspects, and how you can invoke them.

4.5.1 AIX operating system tools for performance analysis

There are multiple tools to monitor the consumption of various AIX resources, such as processors, memory, I/O throughput in the adapters, and disk utilization. There are overlaps in the information that is displayed by each of the tools. You can obtain the same information by using different tools.

We present the tools organized by section:

- Data collection for non-real-time analysis
- Data collection for real-time analysis

The first type of analysis is to monitor the system behavior for a longer time to determine the peak resource consumption windows that can lead to a bottleneck or cause performance issues. This type of analysis is used for capacity planning purposes.

After a particular problem is identified, use the second type of analysis. Use the monitoring tools in real time to see what is happening in the system immediately, change settings, and quickly observe the system behavior again.

The tools that we present in the next sections cover the monitoring of the processor, memory, and I/O.

NMON monitoring tool

NMON (Nigel's MONitor) is one of the default monitoring tools for the AIX operating system. Since AIX 6.1, it is integrated into the operating system. You do not need to install it in the system any longer. NMON can be used to collect data for further review or real-time analysis.

To collect data for further review, NMON must be called with arguments. Each argument instructs the program to collect a particular type of information from the system. We explain the arguments that we used to monitor the parameters of interest to tune Cognos and SPSS on Power. Example 4-3 shows how we started NMON in our tests for further data analysis. You might include the call to NMON in your system crontab to get continuous data.

Example 4-3 Starting NMON for data collection

```
/usr/bin/nmon -f -s <interval> -c <samples> -Y -d -^ -A -m <output_directory>
```

We describe the meaning of each argument:

-f	Specifies that the output must be written in spreadsheet format.
-s	Specifies the sampling rate in seconds.
-c	Specifies the number of samples to take.
-Y	Collects the processes data utilization of a processor. Commands with the same name are grouped in the recording.
-d	Collects the data for disk servicing time.
-^	Collects data about Fibre Channel (FC) device utilization.
-A	Collects data for Asynchronous I/O (AIO).
-m	Specifies the output directory.

The data is collected at every sample interval and saved to an output file in the specified directory. The format of the output file is *hostname_YYMMDD_HHMM*.nmon, where *YY* is the year, *MM* is the month, *DD* is the day, *HH* is the hour, and *MM* is the minute in which the NMON data collection started.

The output file is a comma-separated value file (CVS) that can be easily imported into a spreadsheet. In the spreadsheet, you can use this data to create charts to better visualize the trends in resource utilization and to identify bottlenecks in an easier fashion.

Fortunately, there is already a macro that you can use to create the charts in a spreadsheet file automatically from an NMON output file, as long as you used the **-f** flag to collect your data. The NMON analyzer[1] is a spreadsheet file that you can use to accomplish this task.

After you open the NMON analyzer spreadsheet, click **Analyze nmon data** and select the NMON output file that you collected from your systems. This option generates charts about processor utilization and memory utilization. Each type of information appears in a tab on the resulting spreadsheet report. You can see the following types of data:

- Processor utilization
- Processor folding activity (shared processor LPARs only)
- Threading activity (utilization of logical processors)
- Run queue activity
- Memory utilization
- Paging space
- Disk I/O statistics (service time, disk utilization, throughput, and block size)
- Network adapter statistics (throughput and number of packets)
- I/O adapter statistics (throughput)

NMON can also be used in real time. All of the data that it can capture during a data collection can also be displayed on an AIX shell. Call the **nmon** command from the command line. The interface initially shows a text-mode splash window. Then, you select the information that you want to display by pressing the corresponding letters on your keyboard. Example 4-4 on page 155 and Example 4-5 on page 155 show the splash window and a summary window with the hot keys for data visualization.

[1] NMON analyzer is available for download from this website:
http://www.ibm.com/developerworks/aix/library/au-nmon_analyzer

Example 4-4 NMON splash window

```
+-------------------------------------------------------------------------------+
|  ------------------------------                                               |
|  N    N  M    M   OOOO   N    N   For online help type: h                     |
|  NN   N  MM  MM  O    O  NN   N   For command line option help:              |
|  N N  N  M MM M  O    O  N N  N      quick-hint  nmon -?                     |
|  N  N N  M    M  O    O  N  N N    full-details  nmon -h                     |
|  N   NN  M    M  O    O  N   NN   To start nmon the same way every time?     |
|  N    N  M    M   OOOO   N    N    set NMON ksh variable, for example:      |
|  ------------------------------    export NMON=cmt                          |
|    TOPAS_NMON                                                                 |
|                             16 - CPUs currently                               |
|                             16 - CPUs configured                              |
|                           3300 - MHz CPU clock rate                           |
|                  PowerPC_POWER7 - Processor                                   |
|                           64 bit - Hardware                                   |
|                           64 bit - Kernel                                     |
|                       8,SPSS03 - Logical Partition                            |
|                   7.1.1.15 TL01 - AIX Kernel Version                          |
|                           spss03 - Hostname                                   |
|                           spss03 - Node/WPAR Name                             |
|                         100C99R - Serial Number                               |
|                    IBM,8233-E8B - Machine Type                                |
|                                                                               |
+-------------------------------------------------------------------------------+
```

Example 4-5 shows the hot keys for data visualization. You can get to this help window by pressing h anytime after you called the **nmon** command.

Example 4-5 NMON hot keys for data visualization

```
+-HELP---------most-keys-toggle-on/off--------------------------------------------+
|h = Help information     q = Quit nmon             0 = reset peak counts       |
|+ = double refresh time  - = half refresh          r = ResourcesCPU/HW/MHz/AIX  |
|c = CPU by processor     C=upto 128 CPUs           p = LPAR Stats (if LPAR)    |
|l = CPU avg longer term  k = Kernel Internal       # = PhysicalCPU if SPLPAR   |
|m = Memory & Paging      M = Multiple Page Sizes  P = Paging Space             |
|d = DiskI/O Graphs       D = DiskIO +Service times o = Disks %Busy Map         |
|a = Disk Adapter         e = ESS vpath stats       V = Volume Group stats      |
|^ = FC Adapter (fcstat)  O = VIOS SEA (entstat)    v = Verbose=OK/Warn/Danger  |
|n = Network stats        N=NFS stats (NN for v4)   j = JFS Usage stats          |
|A = Async I/O Servers    w = see AIX wait procs   "="= Net/Disk KB<-->MB       |
|b = black&white mode     g = User-Defined-Disk-Groups (see cmdline -g)         |
|t = Top-Process --->     1=basic 2=CPU-Use 3=CPU(default) 4=Size 5=Disk-I/O    |
|u = Top+cmd arguments    U = Top+WLM Classes       . = only busy disks & procs|
|W = WLM Section          S = WLM SubClasses        @=Workload Partition(WPAR)  |
|[ = Start ODR            ] = Stop ODR                                          |
|~ = Switch to topas screen                                                     |
|Need more details?  Then stop nmon and use: nmon -?                            |
|                                                                               |
+-------------------------------------------------------------------------------+
```

For a more detailed description of NMON, see this website:

http://www.ibm.com/developerworks/aix/library/au-analyze_aix

topas

The *topas* tool also serves multiple parameter monitoring. Since AIX 6.1, topas and NMON are integrated. You can alternate from one tool to the other tool by pressing ~.

Topas can be used to monitor various system resource utilization in the same way as NMON. Topas is called from the command line. As opposed to NMON, it does not show a splash window when it is started and it displays data from your system. Example 4-6 shows the topas interface after it is called.

Example 4-6 Topas interface

```
Topas Monitor for host:spss03                    EVENTS/QUEUES    FILE/TTY
Tue Jun 19 16:42:26 2012   Interval:2            Cswitch     410  Readch      206
                                                 Syscall     368  Writech     428
CPU     User% Kern% Wait% Idle%   Physc  Entc%  Reads        16  Rawin         0
Total     0.0   0.2   0.0  99.8    0.01   0.41  Writes        1  Ttyout      202
                                                 Forks         0  Igets         0
Network    BPS  I-Pkts  O-Pkts    B-In   B-Out  Execs         0  Namei        20
Total    400.0    1.50    1.50   81.50   318.5  Runqueue      0  Dirblk        0
                                                 Waitqueue   0.0
Disk    Busy%      BPS     TPS  B-Read  B-Writ                   MEMORY
Total     2.5        0       0       0       0  PAGING           Real,MB    8192
                                                 Faults        0  % Comp      43
FileSystem          BPS    TPS  B-Read  B-Writ  Steals        0  % Noncomp   45
Total             12.2K   8.00   12.2K       0  PgspIn        0  % Client    45
                                                 PgspOut       0
Name            PID  CPU%  PgSp Owner           PageIn        0  PAGING SPACE
sshd        7864340  0.1 1.23M root              PageOut       0  Size,MB      512
reaffin      589846  0.0  640K root              Sios          0  % Used       3
topas       6291770  0.0 2.99M root                              % Free      97
java        3932398  0.0 20.8M root              NFS (calls/sec)
java        6881474  0.0 54.4M root              SerV2         0  WPAR Activ     0
swapper         264  0.0  448K root              CliV2         0  WPAR Total     0
getty       1442232  0.0  640K root              SerV3         0  Press: "h"-help
modelers    7471260  0.0 5.40M root              CliV3         0         "q"-quit
wrapper     4980846  0.0 1.20M root
modelers    5177498  0.0 5.46M root
```

The displayed data for CPU, Network, and Disk in Example 4-6 can be expanded to display information for all of the existing processors, network cards, and disks. Press the corresponding hot keys on the keyboard to cycle through the display modes for these sections.

A summary of the available hot keys in topas is documented:

a — The a key shows all of the variable subsections that are monitored (CPU, network, disk, Workload Manager (WLM), and process). Pressing the a key always returns the **topas** command to the initial main display.

c — The c key toggles the CPU subsection between the cumulative report, off, and a list of the busiest processors. The number of busiest processors that is displayed depends on the space that is available on the window.

d — The d key toggles the disk subsection between a list of the busiest disks, off, and the report on the total disk activity of the system. The

	number of the busiest disks that is displayed depends on the space that is available on the window.
h	Shows the help window.
n	The n key toggles the network interfaces subsection between a list of the busiest interfaces, off, and the report on the total network activity of the system. The number of the busiest interfaces that is displayed depends on the space that is available on the window.
w	The w key toggles the Workload Manager (WLM) classes subsection on and off. The number of the busiest WLM classes that is displayed depends on the space that is available on the window.
p	The p key toggles the hot processes subsection on and off. The number of busiest processes that is displayed depends on the space that is available on the window.
P	The uppercase P key replaces the default display with the full-window process display. This display provides more detailed information about the processes that are running on the system than the process section of the main display. When the P key is pressed again, it toggles back to the default main display.
W	The uppercase W key replaces the default display with the full-window WLM class display. This display provides more detailed information about the WLM classes and processes that are assigned to the classes. When the W key is pressed again, it toggles back to the default main display.
L	The uppercase L key replaces the current display with the LPAR display.
f	Moving the cursor over a WLM class and pressing the f key displays the list of top processes in the class at the bottom of the WLM window. This key is valid only when topas is in the full-window WLM display (by using the W key or the **-wlms** flag).
q	Quit the program.
r	Refresh the display.
Arrow and Tab keys	Subsections from the main display, such as the CPU, Network, Disk, WLM Classes, and the full-window WLM and Process displays, can be sorted by different criteria. Positioning the cursor over a column activates sorting on that column. The entries are always sorted from the highest value to the lowest value. The cursor can be moved by using the Tab key or the arrow keys. Sorting is valid for 128 disks and 16 network adapters only.

mpstat

You can use the **mpstat** command to monitor statistics about the logical processors of your system. Although NMON and topas can provide this information, the **mpstat** command is an alternative to this real-time resource monitoring.

The tool is called from within the command-line shell. The easiest and typical way to call the tool is illustrated in Example 4-7 on page 158.

Example 4-7 mpstat command invocation and output

```
root@spss03:/ # mpstat 1 2

System configuration: lcpu=8 ent=2.0 mode=Capped

cpu  min  maj  mpc  int   cs  ics   rq  mig lpa sysc us sy wa id   pc  %ec  lcs
  0   15    0    0  251  258    0    1    0 100  758 33 59  0  8 0.01  0.4  163
  1    0    0    0    9    0    0    0    0   -    0  0  1  0 99 0.00  0.1    9
  2    0    0    0    9    0    0    0    0   -    0  0  1  0 99 0.00  0.1   10
  3    0    0    0    9    0    0    0    0   -    0  0  1  0 99 0.00  0.1    9
  U    -    -    -    -    -    -    -    -   -    -  -  -  0 100 1.99 99.4    -
ALL   15    0    0  278  258    0    1    0   0  758  0  0  0 100 0.01  0.6  191
--------------------------------------------------------------------------------

cpu  min  maj  mpc  int   cs  ics   rq  mig lpa sysc us sy wa id   pc  %ec  lcs
  0    7    0    0  254  258    0    1    1 100  451 20 69  0 11 0.01  0.3  163
  1    0    0    0   10    0    0    0    1 100    0  0  2  0 98 0.00  0.1   10
  2    0    0    0    9    0    0    0    0   -    0  0  1  0 99 0.00  0.1    9
  3    0    0    0    9    0    0    0    0   -    0  0  1  0 99 0.00  0.1    9
  U    -    -    -    -    -    -    -    -   -    -  -  -  0 100 1.99 99.5    -
ALL    7    0    0  282  258    0    1    2   0  451  0  0  0 100 0.01  0.5  191
```

In Example 4-7, the tool is started as **mpstat *<interval> <samples>***, where you can enter the interval for the data collection and the number of samples to collect.

One useful flag for monitoring thread processor activity with **mpstat** is the **-s** flag. It groups the logical processors under a virtual processor and shows you the activity of each logical processor. You can also detect whether your workload uses your configured threading mode efficiently by using this output. Example 4-8 illustrates an **mpstat** command that is called with the **-s** flag on an idle system.

Example 4-8 mpstat -s invocation and output

```
root@spss03:/ # mpstat -s 1 2

System configuration: lcpu=8 ent=2.0 mode=Capped

              Proc0                          Proc4
               1.37%                          0.00%
 cpu0    cpu1    cpu2    cpu3    cpu4    cpu5    cpu6    cpu7
  0.77%   0.21%   0.18%   0.20%   0.00%   0.00%   0.00%   0.00%
--------------------------------------------------------------------------------
              Proc0                          Proc4
               2.92%                          0.01%
 cpu0    cpu1    cpu2    cpu3    cpu4    cpu5    cpu6    cpu7
  1.80%   0.37%   0.38%   0.37%   0.00%   0.00%   0.00%   0.00%
```

The output shows information about the statistics of each logical processor. For more information about the meaning of each column, see the man page or *IBM PowerVM Virtualization Managing and Monitoring*, SG24-7590. This publication also provides insight about how to use this monitoring tool to identify potential bottlenecks that relate to processor use.

vmstat

The **vmstat** command is used to report memory statistics. It is run from the command line and provides real-time information. The output of **vmstat** provides an overview of the following information:

- Memory utilization levels
- File system activity (pages in and out)
- Paging activities (pages in and out)
- Memory page faults
- Basic processor utilization metrics

In addition to this information, if you are running a system with any of the advanced memory features, such as AMS, AME, or Active Memory Deduplication, information about parameters that are related to these operations is also displayed.

The vmstat tool is usually called in the same way as mpstat, which is described in "mpstat" on page 157. You start it in the form **vmstat <*interval*> <*samples*>**. Example 4-9 illustrates two samples that are taken by vmstat that use intervals of one second.

Example 4-9 vmstat command invocation and output

```
root@spss03:/ # vmstat 1 2

System configuration: lcpu=8 mem=8192MB ent=2.00

kthr    memory              page              faults              cpu
----- ----------- ------------------------ ------------ -----------------------
 r  b   avm   fre  re  pi  po  fr   sr  cy  in   sy  cs us sy id wa    pc    ec
 0  0 914857 213535   0   0   0   0    0   0   3  702 398  0  0 99  0  0.01   0.6
 0  0 914857 213535   0   0   0   0    0   0   1  295 384  0  0 99  0  0.01   0.3
```

For a thorough understanding of the meaning of each parameter that is reported by vmstat, see *IBM PowerVM Virtualization Managing and Monitoring*, SG24-7590. This publication also provides insight about how to use this monitoring tool to identify potential bottlenecks that relate to memory use.

iostat

The **iostat** command is the tool that is used to measure disk I/O statistics. The way to start it follows the same pattern as for mpstat and vmstat: **iostat <*interval*> <*sample*>**. This tool reports the following real-time statistics:

- Percentage of the time that a disk is busy
- Disk transfer rates for read and write operations
- Transactions per second
- Basic processor utilization metrics

Example 4-10 illustrates the use of this monitoring command. The example collects three samples by using a collection interval of one second between them.

Example 4-10 iostat command invocation and output

```
root@spss03:/ # iostat 1 3

System configuration: lcpu=8 drives=2 ent=2.00 paths=8 vdisks=0
\
tty:      tin         tout    avg-cpu: % user % sys % idle % iowait physc % entc
          1.0         68.0                0.1   0.2   99.7      0.0   0.0    0.6
```

```
Disks:        % tm_act     Kbps      tps    Kb_read   Kb_wrtn
hdisk1           0.0       0.0       0.0          0         0
hdisk0           0.0       0.0       0.0          0         0

tty:      tin         tout    avg-cpu: % user % sys % idle % iowait physc % entc
          0.0        357.0                 0.0   0.1   99.8      0.0   0.0    0.4

Disks:        % tm_act     Kbps      tps    Kb_read   Kb_wrtn
hdisk1           0.0       0.0       0.0          0         0
hdisk0           0.0       0.0       0.0          0         0

tty:      tin         tout    avg-cpu: % user % sys % idle % iowait physc % entc
          0.0        357.0                 0.0   0.2   99.8      0.0   0.0    0.4

Disks:        % tm_act     Kbps      tps    Kb_read   Kb_wrtn
hdisk1           0.0       0.0       0.0          0         0
hdisk0           0.0       0.0       0.0          0         0
```

For a comprehensive description of each output field of the **iostat** command, see *IBM PowerVM Virtualization Managing and Monitoring*, SG24-7590, or go to the command man page.

Network monitoring tools

In this section, we describe the most commonly used monitoring tools for network traffic and configuration.

netstat

The **netstat** command is used to display information about network traffic and configuration. Although you can use it to measure throughput, measuring throughput is more common with the NMON and topas tools that are presented in “NMON monitoring tool” on page 153 and “topas” on page 156.

The netstat tool can be started through a command line to display information that is collected every interval. Example 4-11 displays a collection of data every second.

Example 4-11 netstat output for data collection

```
root@spss03:/ # netstat 1
    input    (en0)     output           input   (Total)    output
 packets  errs  packets  errs colls  packets  errs  packets  errs colls
 1629153     0  2508426     0     0  1806351     0  2685623     0     0
       3     0        4     0     0        3     0        4     0     0
       1     0        1     0     0        1     0        1     0     0
       1     0        1     0     0        1     0        1     0     0
```

The **netstat** command is used to debug networking information by starting it with other parameters, such as **netstat -rn**, to display the routing table information. Or, you can call it without any arguments to display information about the active connections in the system. Example 4-12 on page 161 displays the state of some of the connections in one of our test systems. Some parts of the output are deleted from the example.

Example 4-12 Using netstat to check the state of connections

```
root@spss03:/ # netstat
Active Internet connections
Proto Recv-Q Send-Q  Local Address          Foreign Address        (state)
tcp        0      0  *.*                    *.*                    CLOSED
tcp4       0      0  *.*                    *.*                    CLOSED
tcp4       0      0  *.daytime              *.*                    LISTEN
tcp        0      0  *.ftp                  *.*                    LISTEN
tcp4       0      0  *.ssh                  *.*                    LISTEN
[...]
tcp4       0      0  spss03.38441           172.16.253.30.1141     ESTABLISHED
tcp4       0      0  spss03.38463           172.16.253.30.dbreport ESTABLISHED
ttcp4       0      0  spss03.38329           172.16.253.30.cas-mapi ESTABLISHED
tcp4       0      0  spss03.ssh             172.16.253.18.42591    ESTABLISHED
[...]
tcp        0      0  loopback.32784         loopback.acsisrv       CLOSE_WAIT
[...]
Active UNIX domain sockets
SADR/PCB        Type   Recv-Q Send-Q     Inode          Conn           Refs
Nextref      Addr
f1000e0000015408 dgram      0      0 f1000a05802d8420              0
0                0 /dev/.SRC-unix/SRCI-ajUc
f1000e0002512900
f1000e0002719c08 dgram      0      0 f1000a0580318020              0
0                0 /dev/.SRC-unix/SRCIPajUa
[...]
```

entstat

The **entstat** command is used to gather the network transmission statistics of a specific network card. By using this command, you can verify whether your packets are dropping or whether there are network transmission errors.

The command is started in the format **enstat -d <*interface*>** as demonstrated in Example 4-13. Some of its output is deleted from the example.

Example 4-13 enstat invocation and output

```
root@spss03:/ # entstat -d ent0
-------------------------------------------------------------
ETHERNET STATISTICS (ent0) :
Device Type: Virtual I/O Ethernet Adapter (l-lan)
Hardware Address: 7a:40:c1:92:55:02
Elapsed Time: 1 days 8 hours 49 minutes 27 seconds

Transmit Statistics:                          Receive Statistics:
--------------------                          -------------------
Packets: 2508780                              Packets: 1630714
Bytes: 2614606691                             Bytes: 1252877904
Interrupts: 0                                 Interrupts: 1143140
Transmit Errors: 0                            Receive Errors: 0
Packets Dropped: 0                            Packets Dropped: 0
                                              Bad Packets: 0

Max Packets on S/W Transmit Queue: 0
S/W Transmit Queue Overflow: 0
Current S/W+H/W Transmit Queue Length: 0
```

```
Broadcast Packets: 553                      Broadcast Packets: 185738
Multicast Packets: 537                      Multicast Packets: 6583
No Carrier Sense: 0                         CRC Errors: 0
DMA Underrun: 0                             DMA Overrun: 0
Lost CTS Errors: 0                          Alignment Errors: 0
Max Collision Errors: 0                     No Resource Errors: 0
Late Collision Errors: 0                    Receive Collision Errors: 0
Deferred: 0                                 Packet Too Short Errors: 0
SQE Test: 0                                 Packet Too Long Errors: 0
Timeout Errors: 0                           Packets Discarded by Adapter: 0
Single Collision Count: 0                   Receiver Start Count: 0
Multiple Collision Count: 0
Current HW Transmit Queue Length: 0

General Statistics:
-------------------
No mbuf Errors: 0
Adapter Reset Count: 0
Adapter Data Rate: 20000
Driver Flags: Up Broadcast Running
        Simplex 64BitSupport ChecksumOffload
        DataRateSet
[...]
```

For a more detailed description of this command, see its man page.

Virtual I/O Server monitoring tools

When you have a VIOS to virtualize disks and networking to your client LPARs, it is important that you understand how to monitor the VIOS resources so that it does not become a bottleneck to the other LPARs. The next sections describe tools that you can use for disk and network monitoring.

viostat

It is important to measure the health of the VIOSs of your Power Systems because LPARs with virtual I/O resources relate closely to the VIOS health.

Although you can use all of the previously described commands to monitor the VIOS performance, the VIOS also provides a command called **viostat**.

The output from viostat is similar to the output that is shown by iostat. The same basic information is shown: percentage of disk utilization, transactions per second, throughput, and some basic processor metrics. The way to start viostat is **viostat *<interval> <sample>***, which is the same way as with iostat.

The goal with monitoring the VIOS disks is to ensure that the physical layer, if you use vSCSI to virtualize them, is not a bottleneck to the virtual disks that it serves to your LPARs. Sometimes, you might think that your LPAR virtual disk has a bottleneck, but the real problem might be in the configuration of the parameters of the underlying physical disk in the VIOS. Performance analysis of virtual disk resources must also include an analysis in the VIOS level.

Important: The **viostat** command must be started by the padmin user.

seastat

The **seastat** command is used in the VIOS to gather information about the Shared Ethernet Adapter (SEA) statistics. It provides statistics on a virtual client basis. In this way, you know how much traffic each LPAR client consumes.

The invocation of the command is **seastat -d <*sea adapter*>**. Example 4-14 displays its output. The machine where this example is run has only one virtual client on top of the VIOS, so we can see only one client section.

Example 4-14 seastat command output

```
$ seastat -d ent19

================================================================================

Advanced Statistics for SEA
Device Name: ent19

================================================================================
MAC: EE:AF:0E:10:D0:02
----------------------

VLAN: None
VLAN Priority: None
IP: 192.168.100.56

Transmit Statistics:            Receive Statistics:
--------------------            -------------------
Packets: 1672399                Packets: 1677389
Bytes: 431334555                Bytes: 1222742739

================================================================================
MAC: EE:AF:0E:10:D0:02
----------------------

VLAN: None
VLAN Priority: None

Transmit Statistics:                    Receive Statistics:
--------------------                    -------------------
Packets: 15                             Packets: 192
Bytes: 900                              Bytes: 12571

================================================================================
```

You can clear the statistics before you call the command if you want to start to monitor fresh data. After you clear the statistics, you must turn them on again as shown in Example 4-15.

Example 4-15 Clearing the statistics

```
$ seastat -d ent19 -c
Statistics for ent19 have been cleared
$ chdev -dev ent19 -attr accounting=enabled
ent19 changed
```

Important: The **seastat** command must be started by the padmin user.

4.5.2 Performance analysis tools

In this section, we describe various tools to use to measure resources performance:

- IBM whole system analysis of ideal time (WAIT)
- IBM heap analyzer
- pGraph

IBM whole system analysis of ideal time (WAIT)

The IBM WAIT tool is designed to diagnose performance and scalability issues mostly in deployed enterprise environments. It can be run against any running Java virtual machine (JVM).

The primary input that is used by WAIT is Javacore, which is available from any running JVM without using any command line or environment changes. By using the WAIT tool, you can identify performance and resource utilization in workloads with the central Java component.

WAIT quickly identifies the problem component and provides accurate results so that you can delve in-depth into the source to identify the issue. See Figure 4-7 on page 165.

```
Found javacore in directory: [/opt/IBM/WebSphere/AppServer/profiles/AppSrv01/].
Collected 1 IBM javacores

Triggering snapshot 2: ( 20120630 11:50:06:%N GMT)
Triggering kill -3 for 11534494
Collected 1 IBM javacores

Triggering snapshot 3: ( 20120630 11:50:46:%N GMT)
Triggering kill -3 for 11534494
Collected 1 IBM javacores

Triggering snapshot 4: ( 20120630 11:51:26:%N GMT)
Triggering kill -3 for 11534494
Collected 1 IBM javacores

Triggering snapshot 5: ( 20120630 11:52:07:%N GMT)
Triggering kill -3 for 11534494
Collected 1 IBM javacores

Triggering snapshot 6: ( 20120630 11:52:47:%N GMT)
Triggering kill -3 for 11534494
Collected 1 IBM javacores

Triggering snapshot 7: ( 20120630 11:53:27:%N GMT)
Triggering kill -3 for 11534494
Collected 1 IBM javacores

Triggering snapshot 8: ( 20120630 11:54:07:%N GMT)
Triggering kill -3 for 11534494
Collected 1 IBM javacores

Triggering snapshot 9: ( 20120630 11:54:47:%N GMT)
Triggering kill -3 for 11534494
Collected 1 IBM javacores
```

Figure 4-7 WAIT generating Java core information

WAIT analyzes the following components

This section provides details about the components that WAIT analyzes:

- App Server
- Database
- Network
- Deadlock
- Livelock
- Garbage collection
- Memory leak
- File system
- Insufficient load

Using WAIT

WAIT requires three steps:

1. Collect the data.
2. Upload the data to the WAIT server by using your browser (Firefox, Chrome, Internet Explorer, or Safari), as shown in Figure 4-8 on page 166.

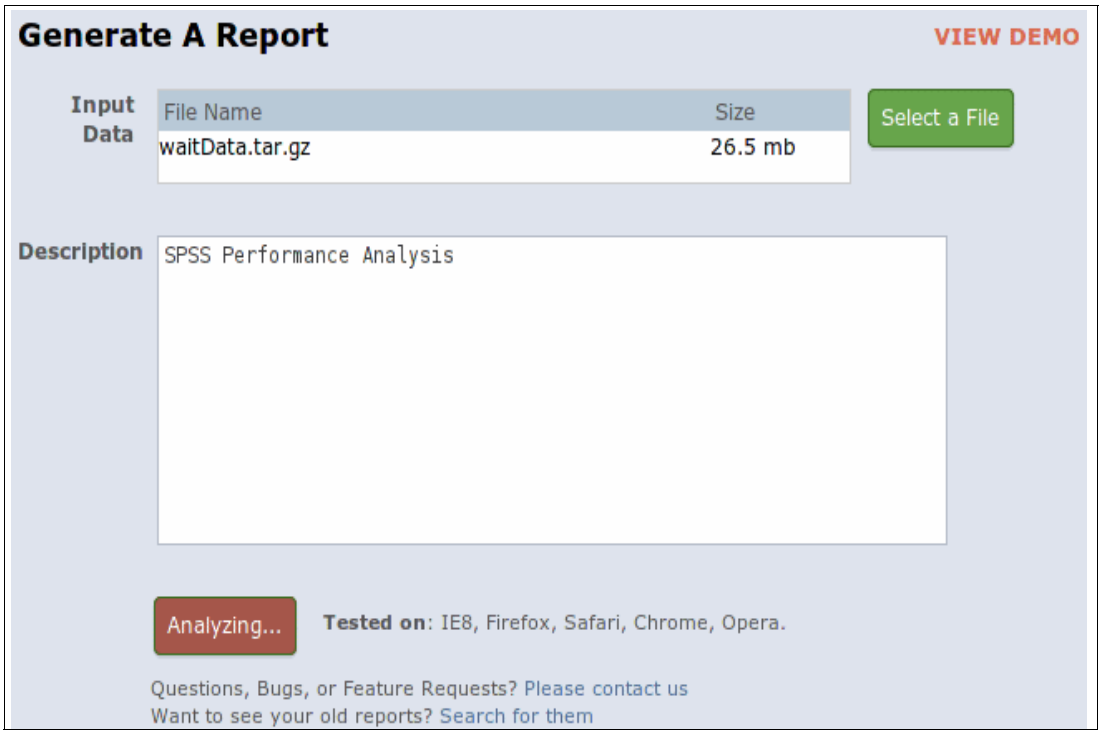

Figure 4-8 Generating reports with the collected WAIT data

3. View the interactive WAIT report with your browser, as shown in Figure 4-9.

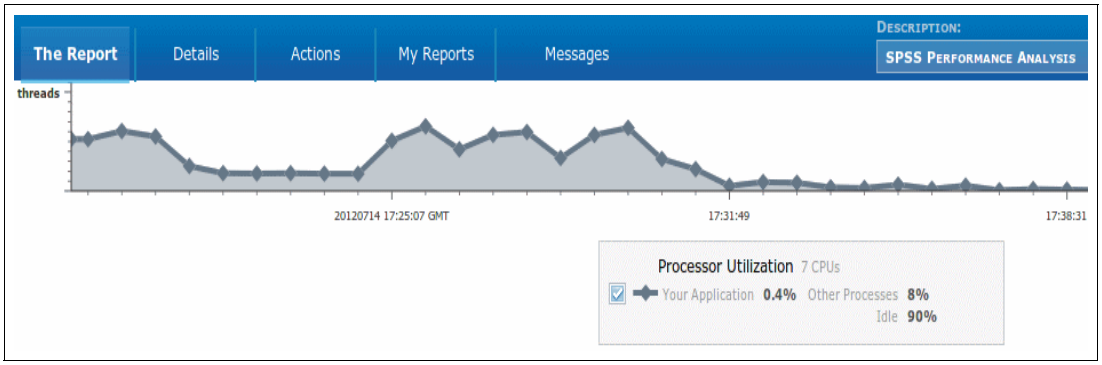

Figure 4-9 Processor utilization report that is generated by the WAIT tool

Framework supported by the WAIT tool

The following framework is supported by the WAIT tool:

- DB2
- Java Database Connectivity (JDBC)
- WebSphere Application Server
- FileNet®
- WebSphere Commerce
- Portal
- WebSphere MQ
- Oracle
- Remote Method Invocations (RMIs)
- Apache commons
- Tomcat

IBM heap analyzer

The IBM *heap analyzer* helps you to find possible Java heap memory leakage, and it analyzes the Java heap dump in the Java application.

The IBM heap analyzer analyzes Java heap dumps by breaking down the Java heap dump and creates directional graphs and directional trees. With the help of the graphs and trees, you can identify the suspected area and location of possible memory leakage. See Figure 4-10.

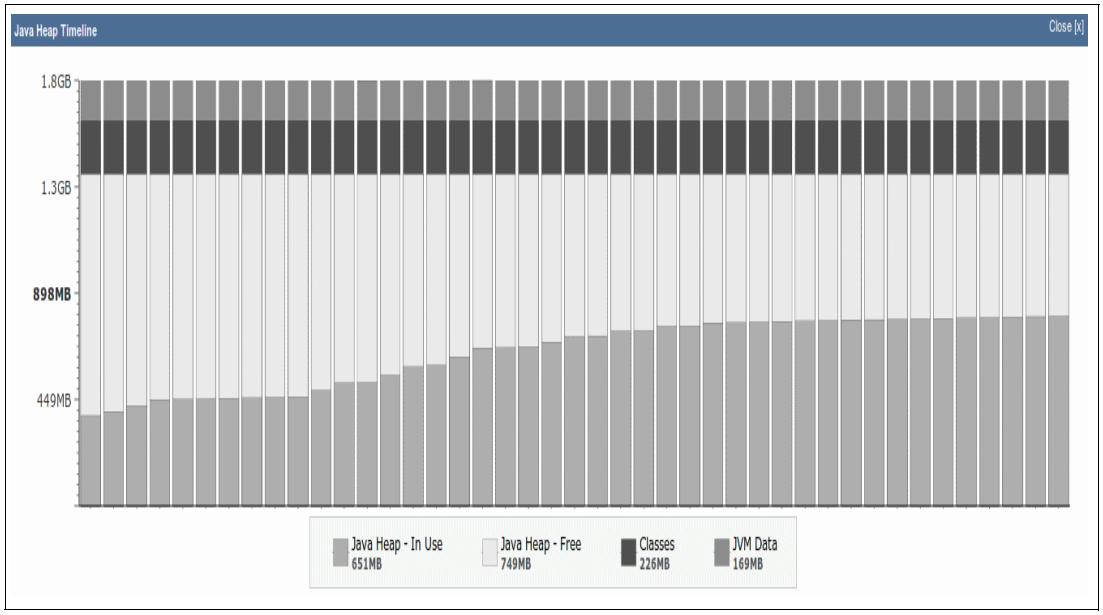

Figure 4-10 Java heap dump generation by using WAIT

Prerequisite for IBM heap analyzer

The following list provides the prerequisites for the IBM heap analyzer:

- Java Runtime Environment 6 or higher
- Minimum of 3 GB of RAM or higher if 64 bit of Java Runtime Environment is for large Java heap dump
- Java heap dump in `.txt` or `.phd` format

Portable Heap Dump format: The IBM heap analyzer uses a `.phd` or `.txt` file to analyze heap dump. The `.phd` contains the *Portable Heap Dump* format, which contains records of all Java heap objects that are used for debugging application errors. Portable heap dump can be generated by setting environment variable parameters in the IBM WebSphere Console, as shown in Figure 4-11 on page 168.

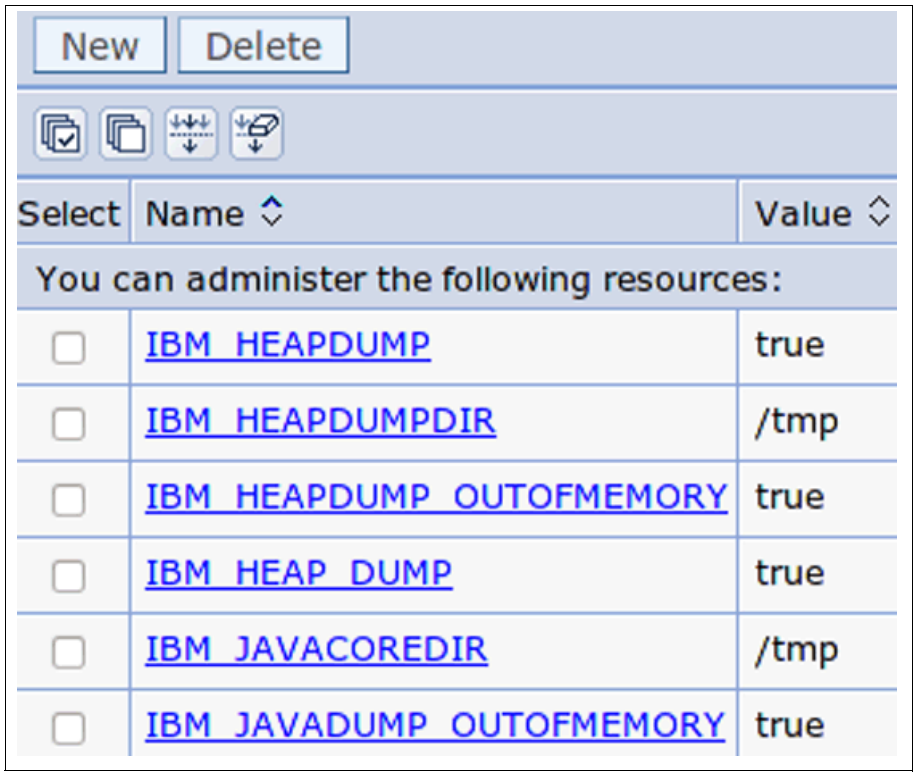

Figure 4-11 Environment variable settings in IBM WebSphere

The following variable settings can be set up, as shown in Figure 4-11:

- `IBM_HEAP_DUMP=true`
- `IBM_HEAPDUMP=true`
- `IBM_HEAPDUMPDIR=<Directory to store heap dump>`
- `IBM_HEAPDUMPOUTOFMEMORY=true`
- `IBM_JAVACOREDIR=<Location of java core dump>`
- `IBM_JAVADUMP_OUTOFMEMORY=true`

Running the IBM heap analyzer

To run the IBM heap analyzer, use the Java 2 Platform Standard Edition Version 6 or higher Java Runtime Environment (JRE). See Figure 4-12 on page 169.

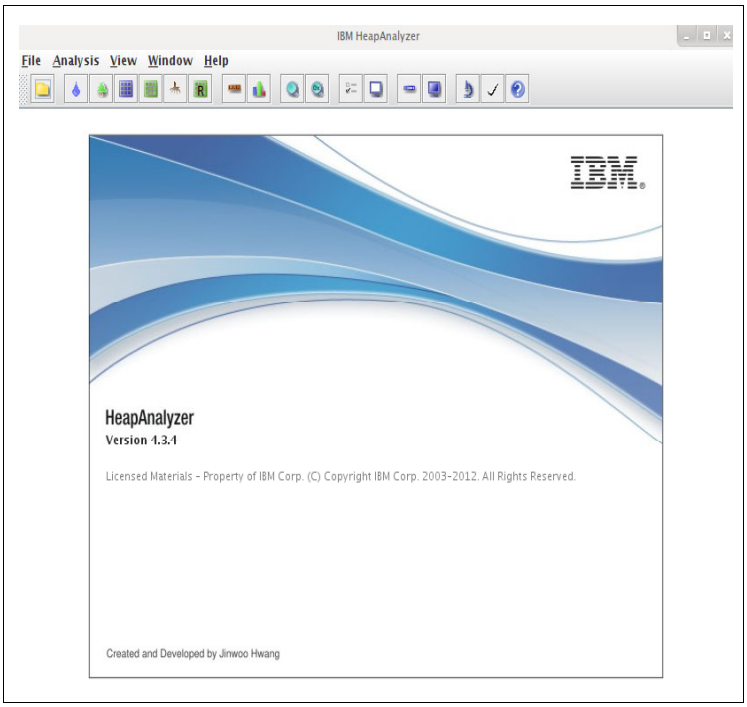

Figure 4-12 IBM HeapAnalyzer

Use the following syntax (Figure 4-13):

`<Java path>java –Xmx[heapsize] –jar ha<HeapAnalyzer version>.jar <heapdump file>`

The following example shows the syntax:

```
java –Xmx1000m –jar ha434.jar heapdump1234.txt
```

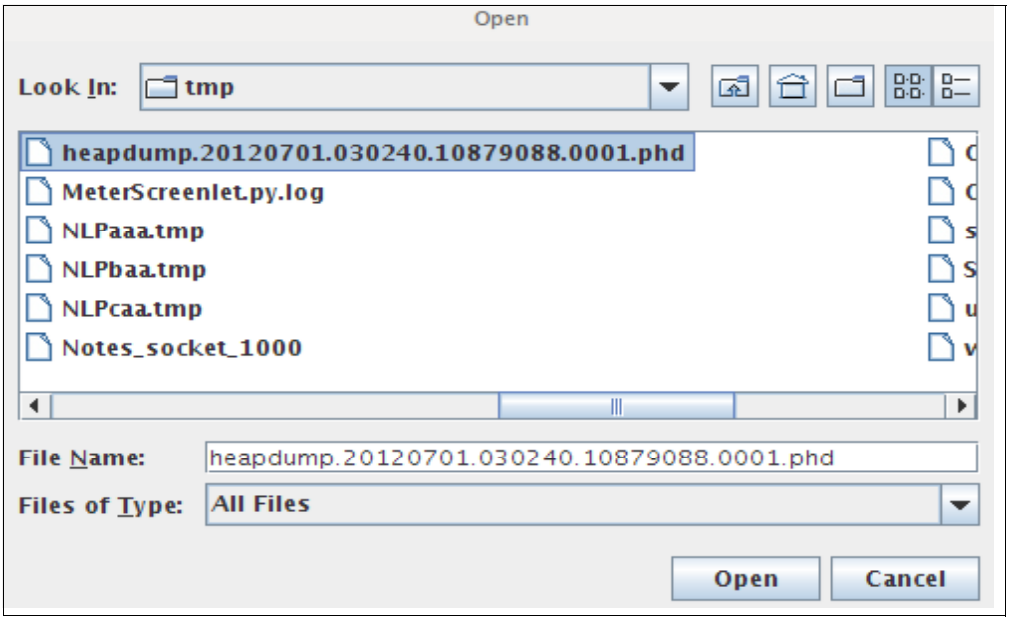

Figure 4-13 Provide the path for the heap dump to generate the report

Figure 4-14 on page 170 shows the main view of the IBM heap analyzer.

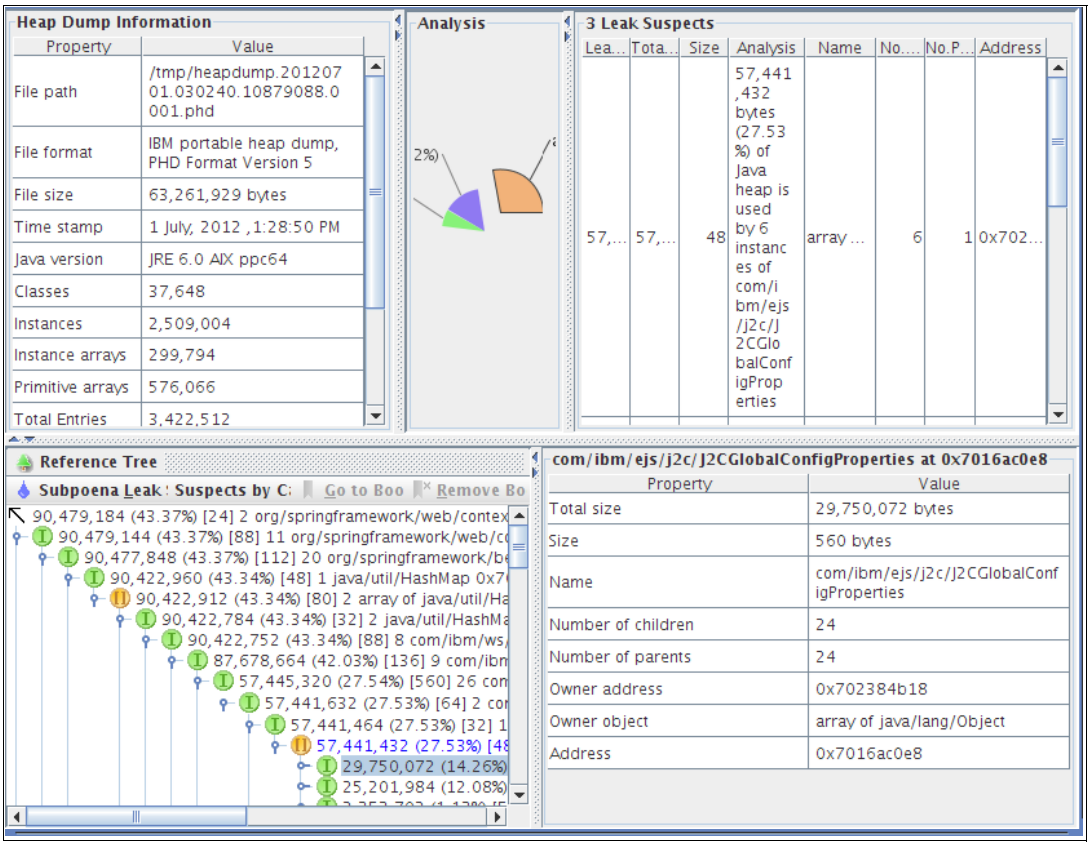

Figure 4-14 Main view of IBM heap analyzer

pGraph

This section describes the uses of the pGraph for performance monitoring of resources, such as the memory, processor, hard disk, and network. pGraph produces the resources utilization graph from collected performance data that includes nmon, vmstat, or iostat.

Features of pGraph

pGraph has these features:

- Multiple data format support, which includes nmon, mvstat, iostat, and sar.
- Can work on multiple operating systems, such as AIX, Linux, Solaris, and HP-UX.
- pGraph can read multiple files to produce the performance reports.
- Can export a graph to CSV and PNG file formats.

How to run pGraph

pGraph requires a Java virtual machine (JRE or Java Development Kit (JDK)) Version 1.5 or later. You can run pGraph in interactive mode, batch mode, or applet mode. You might see the "`out-of-memory`" error. If you see this error, you need to increase the memory that is available to Java by using the **`-Xmx<size>`** parameter, for example:

```
java -Xmx1000
```

In this example, 1000 MB of memory is allocated to this Java process.

Create the pGraph directory, download the pGraph tool, and run the command:

```
#java -cp pGraph.jar pGraph.Viewer xyz123.nmon (where xyz123.nmon is the file
created by nmon)
```

You can download the latest pGraph tool from the following website:

https://www.ibm.com/developerworks/wikis/display/WikiPtype/pGraph#pGraph-download

Figure 4-15 shows the sample output for various resource utilization that is captured by pGraph.

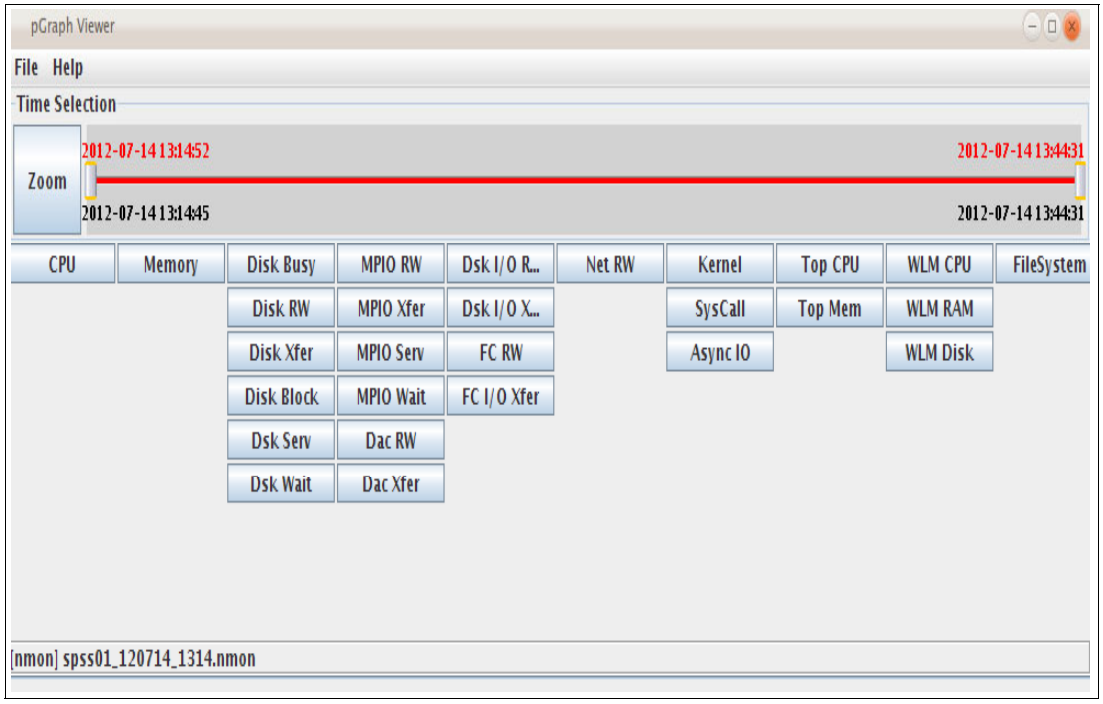

Figure 4-15 pGraph viewer that contains various options to view resource utilization

By using the pGraph viewer, you can see various resources and the associated utilization.

Figure 4-16 on page 172, Figure 4-17 on page 173, and Figure 4-18 on page 174 show the various resource utilization outputs taken from pGraph, which include CPU utilization, average disk busy during load, and memory utilization.

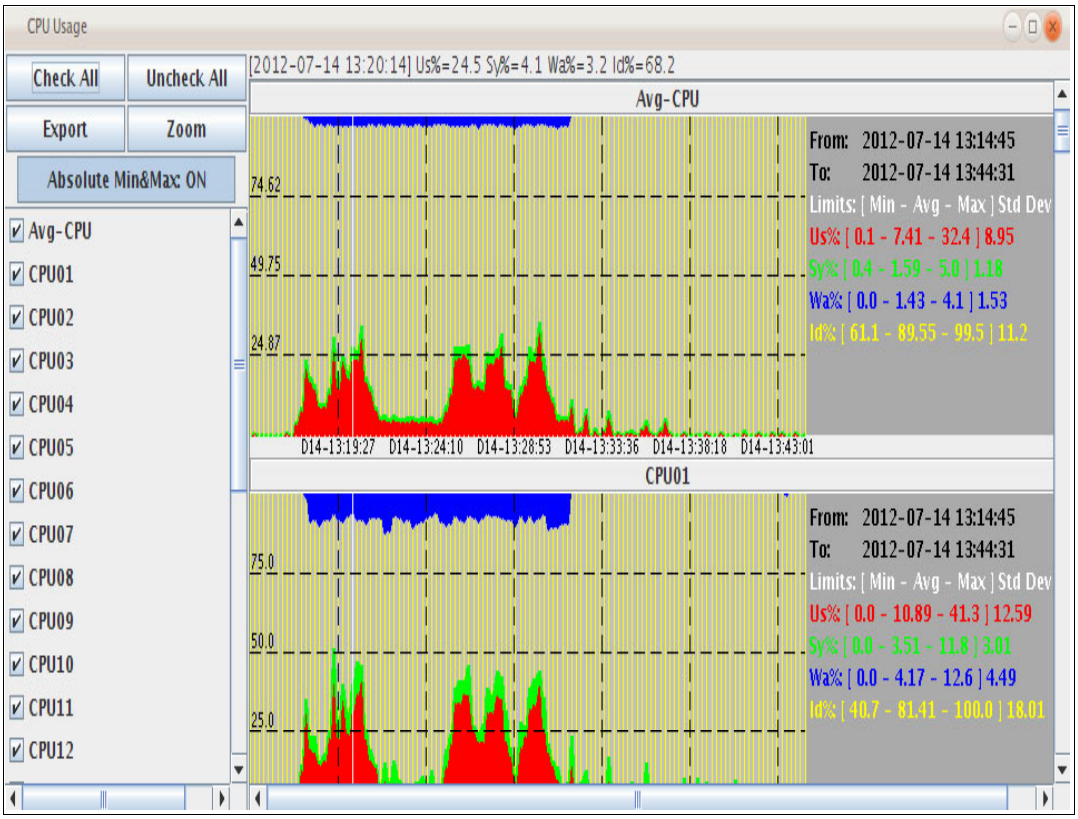

Figure 4-16 pGraph shows CPU utilization

Figure 4-17 pGraph shows average disk busy output along with hdisk2 and hdisk3

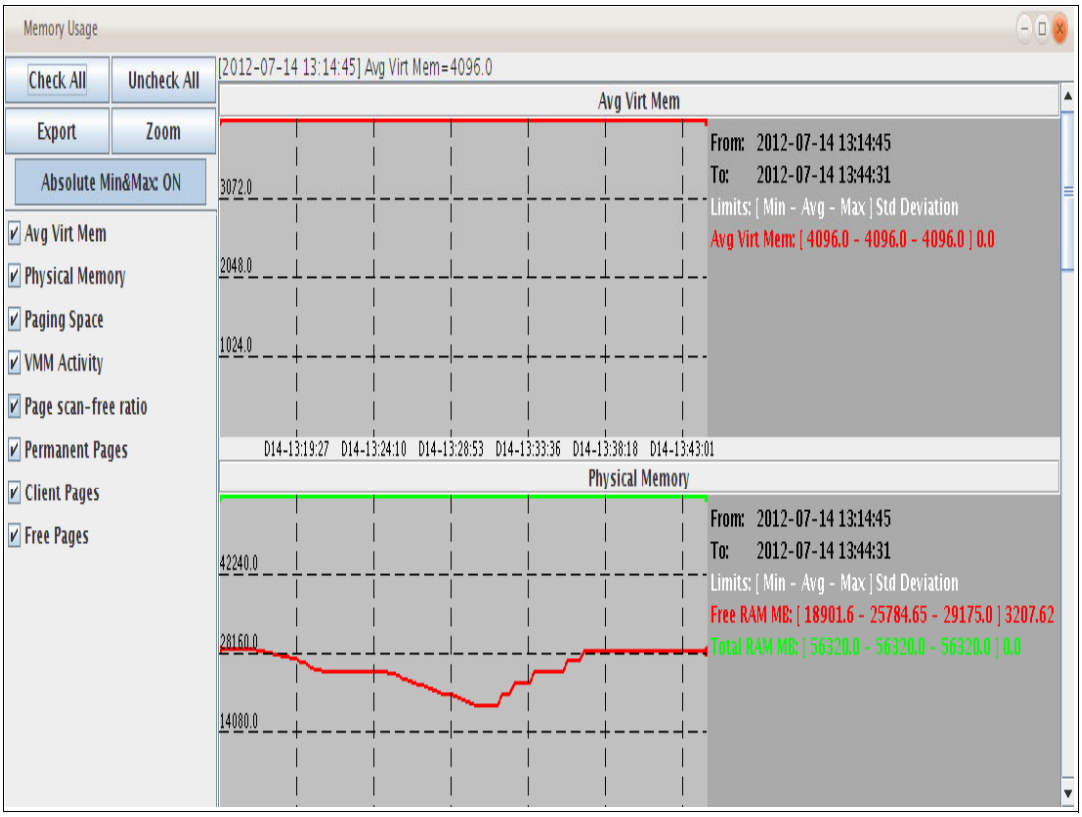

Figure 4-18 Memory consumption view that uses pGraph

4.5.3 IBM SPSS monitoring tools

Knowing how to find bottlenecks in the operating system and hardware configuration is important. Measuring whether a change affects your environment positively or negatively is important, which is why we describe many tools in 4.5.1, “AIX operating system tools for performance analysis” on page 153.

However, tuning the operating system or the hardware is only in response to an application need. If the expectation for the execution of an application is met (amount of processed data, estimated time of completion, or number of concurrent users), there is no need to further tune it. If not, you must tune it. We need to know:

- How to measure if a change affected the application positively or negatively
- How to tune the application

This section provides guidance about which tools you can use to monitor the performance of your SPSS environment. The particular tunables that you can change to obtain gains are explained in 3.1.5, “SPSS tuning and performance statistics” on page 120.

In the following sections, we describe two of the tools that you can use to gain an insight about the performance of your server:

- The IBM SPSS Collaboration & Deployment Services interface
- The soapUI tool

The soapUI is not a tool to monitor a production environment. It is intended to test a pre-production environment with test requests so that you can measure the system performance before beginning production. The values that it reports are received from the server as a result of tests, and not reported by the server. It can provide a clear overview of performance as it is seen by SPSS clients.

IBM SPSS Collaboration and Deployment Services Deployment Manager

In addition to being the tool to interact with to set up models and their execution on the SPSS server, the IBM SPSS C&DS Deployment Manager can also provide performance statistics that you can use for analysis.

Figure 4-19 depicts the interface of the Deployment Manager that displays information for server `spss01`. There are many models that are stored in that SPSS C&DS server. There is one scoring that is already configured from model *ProjectProfile1.str*, which is outlined in red.

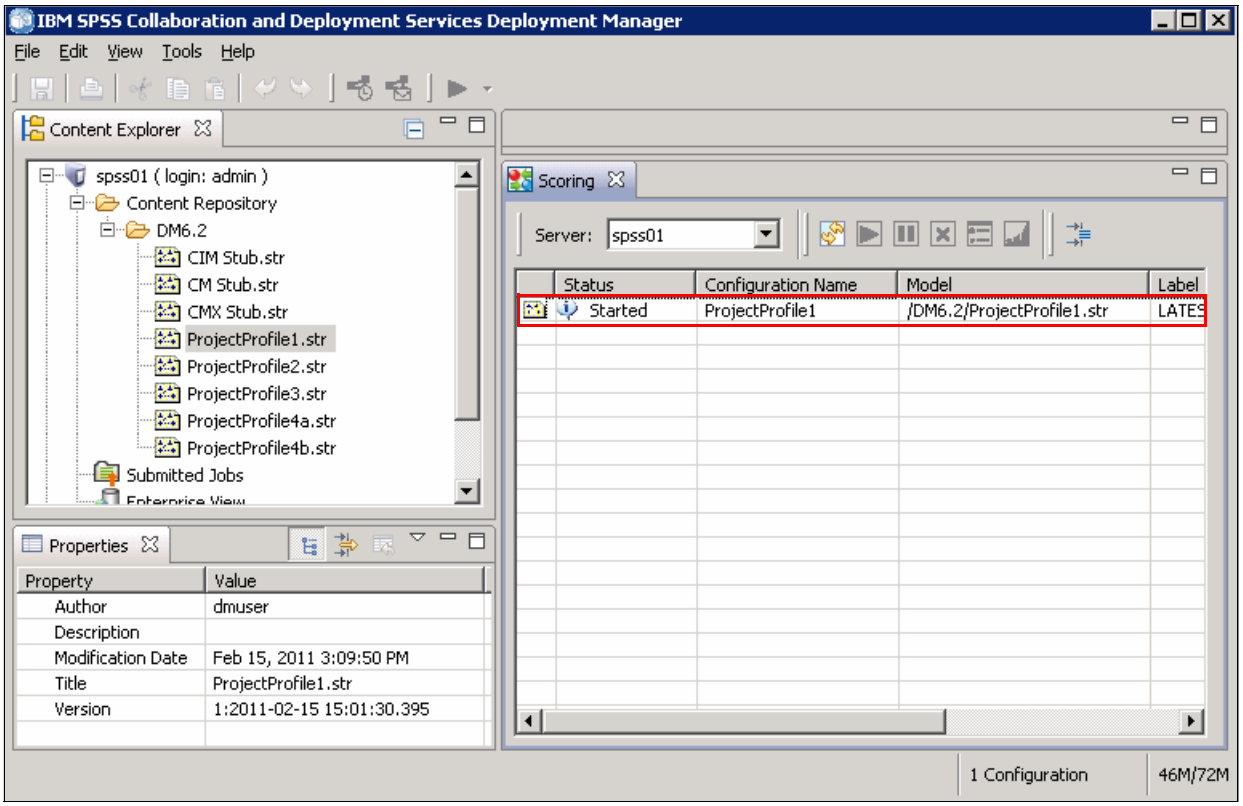

Figure 4-19 IBM SPSS C&DS Deployment Manager interface

When your server has any load to process, you can click the **Refresh** icon, as shown in Figure 4-20 on page 176.

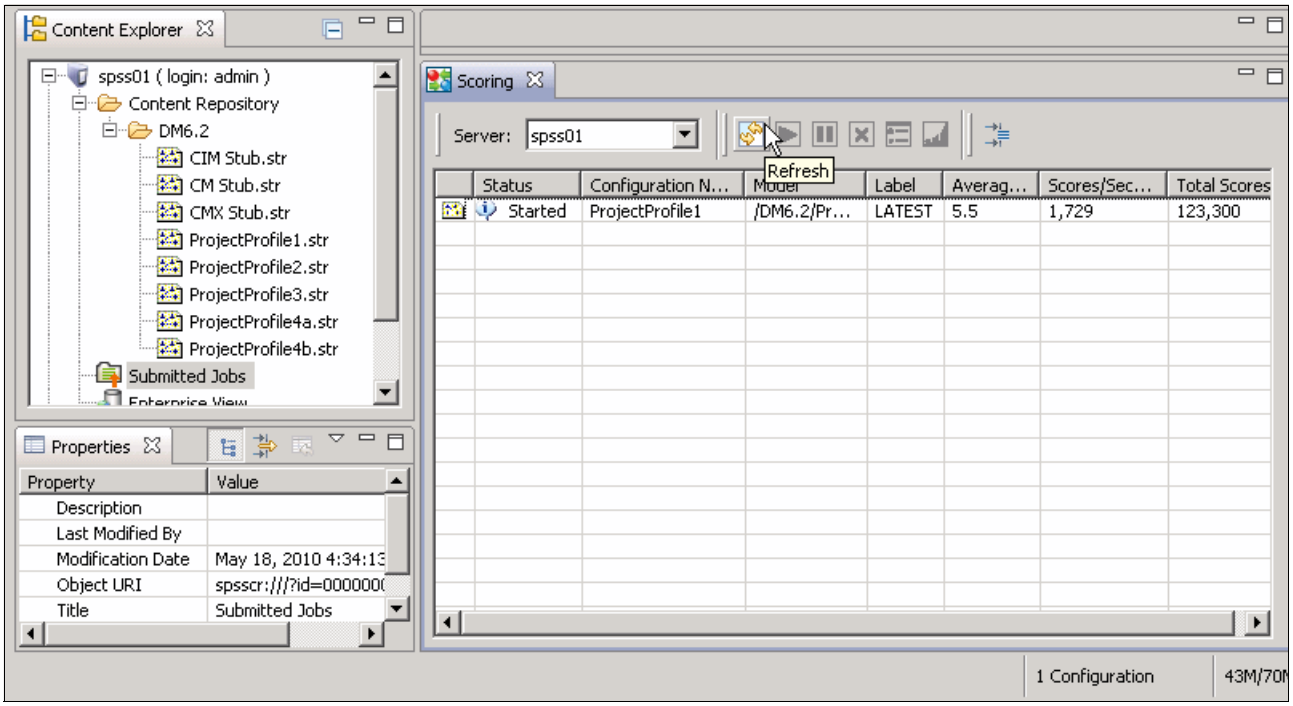

Figure 4-20 Collecting basic statistics from the Deployment Manager

The tool reports statistics as a whole for the average response time of a request and the number of scores per second. If you have more than one JVM listening to answer to requests, you can see the aggregate results in this view. Also, it provides punctual information about these statistics when you click the Refresh icon.

If you want to view the value for some of the statistics variables in real time, click the **Graphic View** icon in the Scoring tab, as shown in Figure 4-21 on page 177. This icon shows the area that is outlined in red. You can choose which statistic to follow by selecting it in the Content Shown selection box.

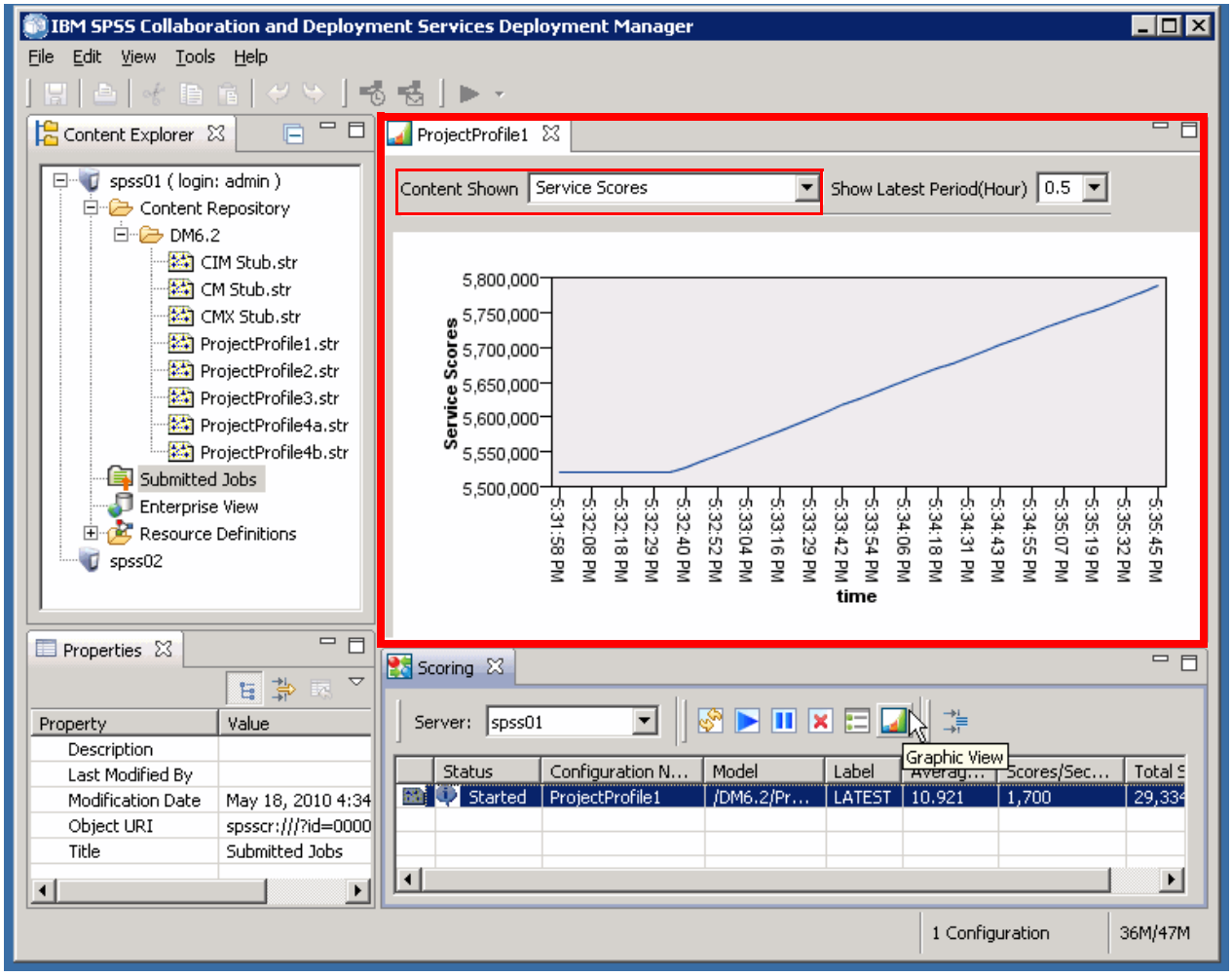

Figure 4-21 Deployment Manager graphic view

The soapUI tool

The soapUI is an open source tool that can be run on multiple platforms. It is used to automate multiple types of tests, as described at http://www.soapui.org. We are interested in its capability to send a burst of requests to an IBM SPSS C&DS server to obtain the number of scores per second that the server currently handles under its current configuration.

This tool is not a tool to monitor a production environment. It is intended to test a pre-production environment with test requests so that you can measure the system performance before going into production.

Our test environment is composed of an SPSS server with historical data. We query it with a request that uses a random entry for which we want to predict the value of two of its fields according to a prediction model that was created. This request is based on XML, and so is the response that we obtain from the server. If you plan to use soapUI to measure metrics from your SPSS server, you need to use an XML request that is compatible with your historical data. You also must define a model in the SPSS C&DS Configuration Manager to use with the tests, as explained in“Scoring the configuration model cache setting” on page 107.
Figure 4-22 on page 178 displays a request that is sent to the server and its response.

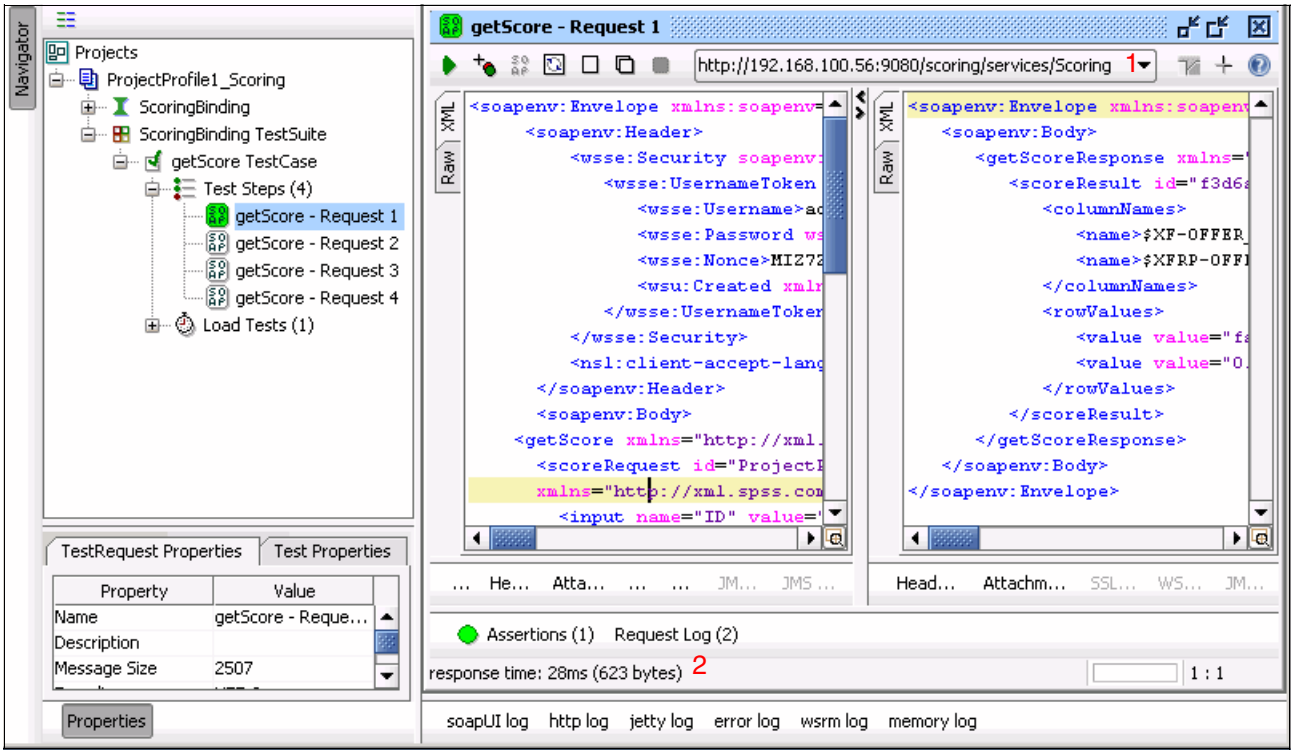

Figure 4-22 soapUI interface: Sending a request to the server

The information in which we are interested for this single test case is outlined in Figure 4-22:

1. The URL of the server that is used to score the query, and the port number[2]
2. The response time of that particular query

Issuing a single request to the server is useful to test the connectivity of the test tool with the server. A better approach in using soapUI is to create a test case in which multiple requests are sent to the server (simulating a certain amount of simultaneous users contacting the system), and configure a certain amount of time to keep querying the server (simulating a sustained load). This example simulates an actual scenario.

Figure 4-23 on page 179 displays a running test where four JVMs are configured to handle the load that is sent to the SPSS server. Each request test (1 - 4) is targeted at a different JVM to process the batch that comes from it.

[2] You can configure your SPSS server to run multiple JVM instances to handle requests. Each one listens to a port of your choice.

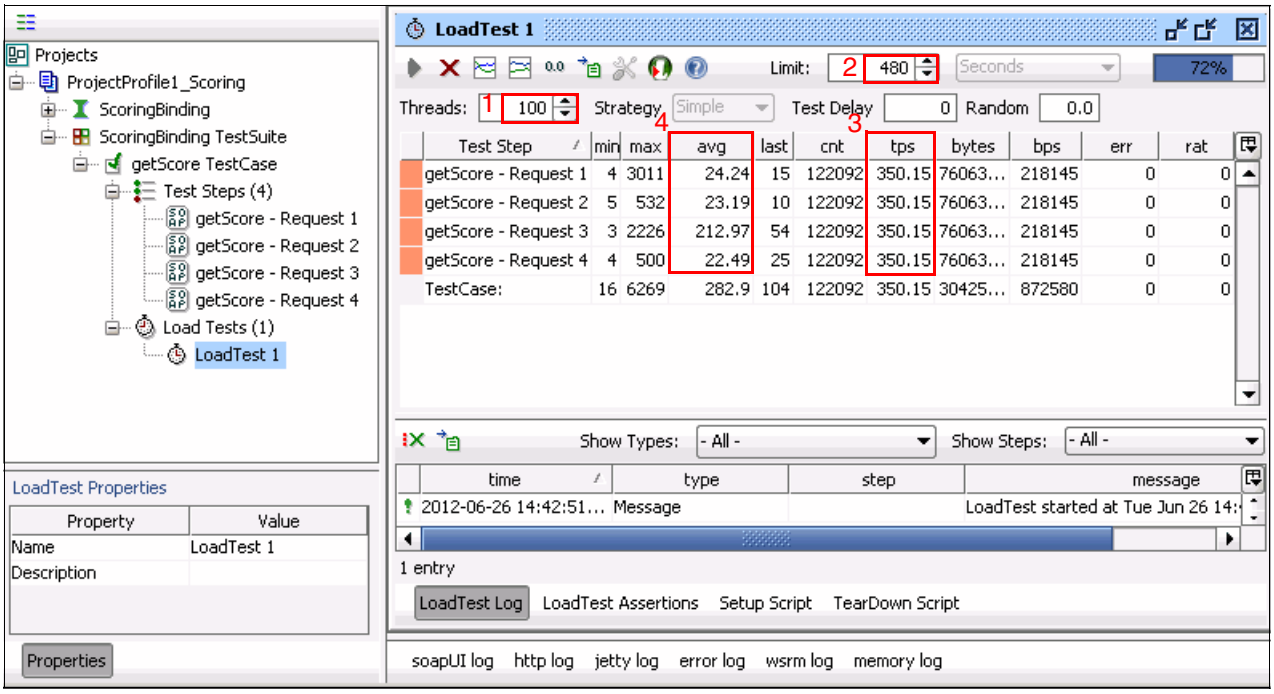

Figure 4-23 soapUI: Measuring the number of scores per second

Figure 4-23 shows relevant information (the numbers relate to the numbers in Figure 4-23):

1. The number of threads (simultaneous requests) that is sent to each JVM
2. The amount of time that the test runs
3. The number of transactions per second (scores) that is performed by *each* JVM
4. The average response time on the requests that are sent to that JVM

The overall scores per second of the system are the sum of the scores of each JVM. In this case, where each JVM reports to handle 350 scores per second, the overall server handles a total of 1,400 scores per second. You can use this number to determine how the performance of your system is affected by a specific change in the configuration of your environment.

Depending on where you run your tests (for example, a partition on the same Power frame, or another system outside of it), some external aspects might affect your results. The server on which the test tool is running must not be a bottlenecked server. If it is a bottlenecked server, your SPSS server never gets enough requests per second to reach its full potential.

You can see whether the load generation test system is the bottleneck in tests by checking whether your SPSS system operates under 100% of its processing capacity (assuming that the server does not have I/O or memory bottlenecks). If so, it usually means that the load it is receiving is not enough to drive it to higher processing levels. In this case, consider tuning your load generation environment or add another client machine to generate more load. In our scalability tests, we needed to run the soapUI tool from multiple smaller systems to achieve 100% of processor utilization of an SPSS server with 64 cores, as well as to tune their networking parameters (TCP window sizes).

To determine whether your SPSS server has a bottleneck, use the tools that are described in 4.5.1, “AIX operating system tools for performance analysis” on page 153.

Related publications

The publications listed in this section are considered particularly suitable for a more detailed discussion of the topics covered in this book.

IBM Redbooks

The following IBM Redbooks publications provide additional information about the topic in this document. Note that some publications referenced in this list might be available in softcopy only.

- *IBM PowerVM Getting Started Guide*, REDP-4815-00
- *IBM PowerVM Virtualization Introduction and Configuration*, SG24-7940-04
- *IBM PowerVM Virtualization Managing and Monitoring*, SG24-7590-03
- *IBM AIX Version 7.1 Differences Guide*, SG24-7910
- *Exploiting IBM PowerVM Virtualization Features with IBM Cognos 8 Business Intelligence*, SG24-7842-00
- *IBM PowerVM Live Partition Mobility*, SG24-7460
- *IBM PowerVM Virtualization Active Memory Sharing*, REDP-4470
- *Power Systems Memory Deduplication*, REDP-4827

You can search for, view, download or order these documents and other Redbooks, Redpapers, Web Docs, draft and additional materials, at the following website:

ibm.com/redbooks

Other publications

These publications are also relevant as further information sources:

- *SPSS Modeler UNIX Installation Guide*

  ftp://public.dhe.ibm.com/software/analytics/spss/documentation/modeler/15.0/en/ServerInstallationUNIX.pdf

- *SPSS Modeler Client Installation Guide*

  ftp://public.dhe.ibm.com/software/analytics/spss/documentation/modeler/15.0/en/ClientInstallation_ConcurrentNetworkLicense.pdf

- *SPSS Administration Console User Guide*

  ftp://ftp.software.ibm.com/software/analytics/spss/documentation/modeler/15.0/en/ModelerAdministrationConsole.pdf

- *SPSS Modeler Administration and Performance Guide*

  ftp://ftp.software.ibm.com/software/analytics/spss/documentation/modeler/15.0/en/ServerandPerformanceGuide.pdf

Online resources

These websites are also relevant as further information sources:

- Cognos Optimized for Power Systems

  http://www-03.ibm.com/systems/power/solutions/analytics/

- Best practices of IBM SPSS Collaboration and Deployment Services 4.2 on IBM POWER7/AIX 7.1

  http://www.ibm.com/common/ssi/cgi-bin/ssialias?infotype=SA&subtype=WH&htmlfid=POW03077USEN

- POWER7 Virtualization Best Practice Guide

  https://www.ibm.com/developerworks/wikis/download/attachments/53871915/P7_virtualization_bestpractice.doc?version=1

- Cognos: BI: Best practices and advantages of IBM Power Systems for running IBM Cognos Business Intelligence

  http://www-01.ibm.com/common/ssi/cgi-bin/ssialias?infotype=SA&subtype=WH&htmlfid=YTW032O9CAEN

Help from IBM

IBM Support and downloads

ibm.com/support

IBM Global Services

ibm.com/services

Index

S

T

V

W

IBM Redbooks

Workload Optimized Systems: Tuning POWER7 for Analytics

(0.2”spine)
0.17”<->0.473”
90<->249 pages

Workload Optimized Systems
Tuning POWER7 for Analytics

Redbooks®

Strengthens IBM Cognos and SPSS for analytics workloads

Helps optimization, security, scaling, capacity, and resilience

Shows model for analytics on IBM Power Systems

This IBM Redbooks publication addresses topics to help clients to take advantage of the virtualization strengths of the POWER platform to solve system resource utilization challenges and maximize system throughput and capacity. This publication examines the tools, utilities, documentation, and other resources available to help technical teams provide business solutions and support for Cognos Business Intelligence (BI) and Statistical Package for the Social Sciences (SPSS) on Power Systems virtualized environments.

This book addresses topics to help address complex high availability requirements, help maximize the availability of systems, and provide expert-level documentation to the worldwide support teams.

This book strengthens the position of the Cognos and SPSS solutions with a well-defined and documented deployment model within a POWER system virtualized environment. This model provides clients with a planned foundation for security, scaling, capacity, resilience, and optimization.

This book is targeted toward technical professionals (BI consultants, technical support staff, IT Architects, and IT Specialists) who are responsible for providing Smart Analytics solutions and support for Cognos and SPSS on Power Systems.

INTERNATIONAL TECHNICAL SUPPORT ORGANIZATION

BUILDING TECHNICAL INFORMATION BASED ON PRACTICAL EXPERIENCE

IBM Redbooks are developed by the IBM International Technical Support Organization. Experts from IBM, Customers and Partners from around the world create timely technical information based on realistic scenarios. Specific recommendations are provided to help you implement IT solutions more effectively in your environment.

For more information:
ibm.com/redbooks

SG24-8057-00 ISBN 0738437328