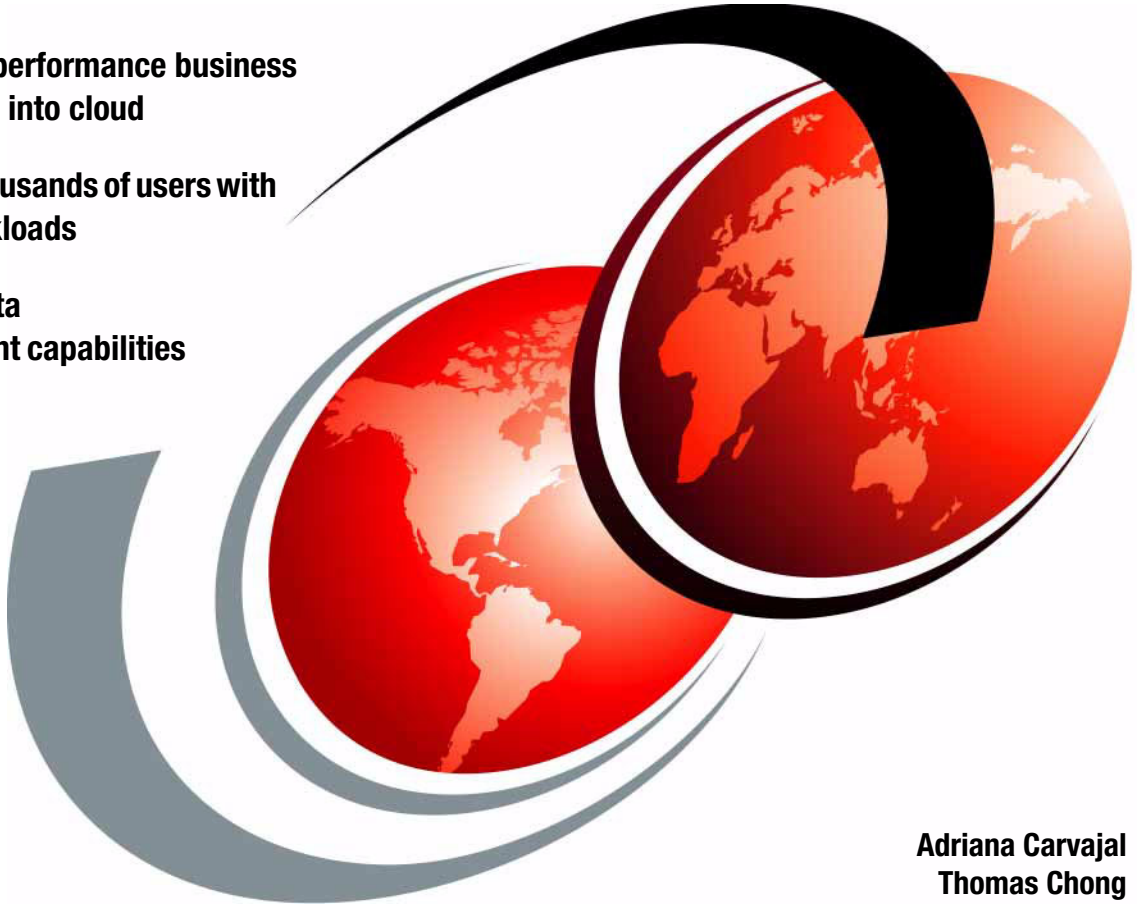
IBM® Information Management Software

Virtualized Business Intelligence with InfoSphere Warehouse

- Bring high performance business intelligence into cloud
- Manage thousands of users with mixed workloads
- Use rich data management capabilities



Adriana Carvajal
Thomas Chong

Redbooks

ibm.com/redbooks

International Technical Support Organization

Virtualized Business Intelligence with InfoSphere Warehouse

October 2012

SG24-8019-00

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

First Edition (October 2012)

This edition applies to Version 10.1 of IBM InfoSphere Warehouse (product number: 5724-E34).

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. 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 IBM's application programming interfaces.

Trademarks

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. These and other IBM trademarked terms are marked on their first occurrence in this information with the appropriate symbol (® or ™), indicating US registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the Web at http://www.ibm.com/legal/copytrade.shtml

The following terms are trademarks of the International Business Machines Corporation in the United States, other countries, or both:

BladeCenter®
ClearCase®
CloudBurst®
Cognos®
DataStage®
DB2®
Guardium®
IBM SmartCloud™
IBM®
Informix®
InfoSphere®
Intelligent Miner®
Optim™
Passport Advantage®
PowerVM®
POWER®
PureSystems™
QualityStage®
Rational Rose®
Rational®
Redbooks®
Redbooks (logo)  ®
System p®
System x®
System z®
Systems Director VMControl™
Tivoli®
WebSphere®
z/OS®

The following terms are trademarks of other companies:

Intel Xeon, Intel, Intel logo, Intel Inside logo, and Intel Centrino logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

Linux is a trademark of Linus Torvalds in the United States, other countries, or both.

Microsoft, Windows, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

Java, and all Java-based trademarks and logos are trademarks or registered trademarks of Oracle and/or its affiliates.

UNIX is a registered trademark of The Open Group in the United States and other countries.

Other company, product, or service names may be trademarks or service marks of others.

Preface

With the benefit of advanced analytics such as online analytical processing (OLAP), data mining, and text analytics, the IBM® InfoSphere® Warehouse Enterprise Edition brings sophisticated business intelligence (BI) to warehouse users. InfoSphere Warehouse allows you to run extreme concurrent query volumes that can help answer questions for all types of business users, while consistently meeting service level requirements.

Combined with a virtualization platform and a solid BI solution, such as IBM Cognos®, you can deliver BI cloud services with improved flexibility and speed to your clients, thereby presenting a new avenue for which your services can be offered.

This IBM Redbooks® publication discusses the deployment of a BI cloud solution. It includes details such as understanding the architecture of a cloud, planning implementation, integrating various software components, and understanding the preferred practices of running a cloud deployment. Essentially, this book can be used as a guide by anyone who is interested in deploying a virtualized environment for a BI cloud solution.

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, San Jose Center.

Adriana Carvajal is a software developer in the Information Management Technology Ecosystem team, where she provides support for IBM Information Management products on Linux and virtualization platforms. She has worked with IBM DB2® implementations since 2007, and recently became involved in activities related to data warehousing, business intelligence and cloud computing. Adriana's expertise in InfoSphere Warehouse and Cognos knowledge has allowed her to assist, teach, and set up various business intelligence implementations for customers and Business Partners.

Thomas Chong is an Automation and Integration Lead in the Information Management Technology Ecosystem team, specializing in cloud computing enablement for various IBM middleware products such as IBM DB2. His tasks include providing education, best practices, and enablement and integration of IBM middleware to IBM Business Partners looking to deliver their solutions on virtualized and cloud platforms.

Acknowledgements

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

Boris Bialek
Melody Ng
Kit Man Cheung
Jason Chan
Piotr Proski
Chris Moon
Hua Lu
Daniel Piao
Manasa Seethepalli
IBM Toronto Lab

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

Introduction to virtualized business intelligence

Moving traditional business intelligence (BI) implementations into the cloud can accelerate and simplify deployment, give more users access to services on demand, and address users' overall needs faster and more cheaply than hosting the infrastructures themselves. Transitioning into the cloud brings challenges, but these challenges can be minimized with clear scope and business objectives.

This chapter introduces the concept of virtualized business intelligence and how it can support a cloud infrastructure to provide decision-making services anytime, anywhere.

The chapter includes the following sections:

- The evolution to cloud computing
- Advantages and disadvantages of using a public cloud versus a private cloud
- Market growth and challenges
- Using IBM cloud enablement technologies for business intelligence
- Scope and business objectives

1.1 The evolution to cloud computing

From software developers to marketing and sales reps to business decision makers, everyone in the industry seems to have a different perspective on what, exactly, cloud computing is. This book adopts the perspective that cloud computing is not a new technology invented or built from scratch but rather an evolution of many combined technologies that have emerged over the past several decades. The evolving model was built upon virtualization, grid computing, utility computing, and software as a service (SaaS), as illustrated in Figure 1-1.

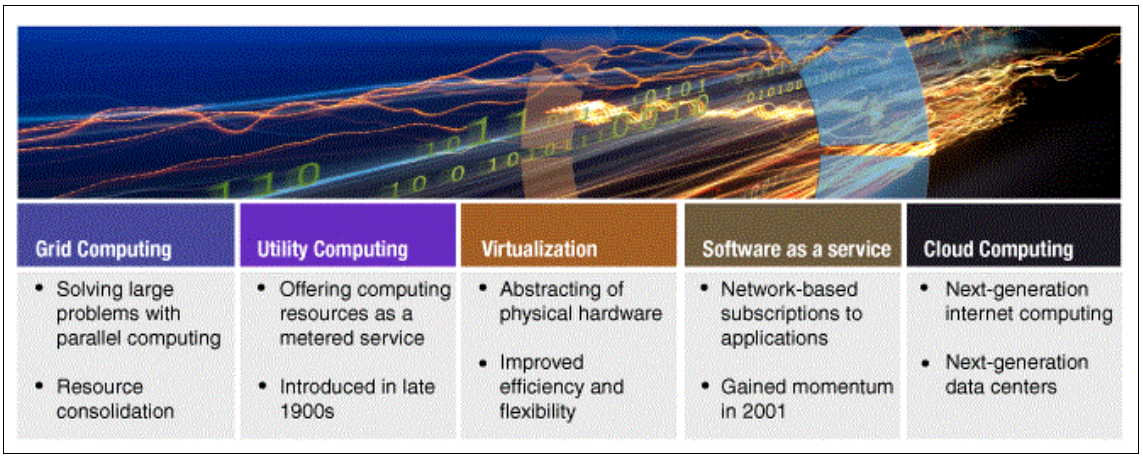


Figure 1-1 The evolution to cloud computing

The origins of cloud computing can be traced back to as early as the 1980s, when the idea of *grid computing* emerged. Grid computing was put forth as an idea to solve particularly complex, individual problems. Grid computing technology pooled together a large number of systems using parallel computation.

The following decade saw the rise of utility computing, as popularized by IBM. By implementing control facilities and metering capabilities, servers and mainframes were offered as time-shares to large corporations, thereby enabling a renting model for computing resources.

In the 1990s, *virtualization* emerged as a technology able to abstract physical hardware into subcomponents such as storage and network resources. Later on, *utility computing* combined with virtualization by adding metering capabilities, thereby enabling a virtual hardware rental model.

Eventually, virtualization paved the way for the development of *software as a service* (SaaS). In the early 2000s, SaaS became more prevalent in a consumer's everyday computing as services such as email, picture hosting, web hosting, and so on became more accessible to the general public.

The concept of cloud computing has evolved from the concepts of grid, virtualization, utility computing, and SaaS. The emerging model lets users access cloud services, whether SaaS or other *as-a-service* services, and also lets users deliver their own cloud services as well. The delivery platform has grown into its own marketplace and has evolved into a technology ecosystem where new software and services are being built and delivered every day.

1.1.1 Understanding cloud computing

Consider the following technologies that support cloud computing to gain a better understanding of its key concepts:

- Virtualization

  From a computer science and engineering perspective, *virtualization* is a technique for simulating a physical computing environment using generated virtual resources. Although there are many types of virtualization, all virtualization accomplishes the following benefits:

  - Create a smaller working environment
  - Create a larger working environment

  Creating multiple working environments from a single physical computing environment results in a smaller but similar working environment. In addition, building a working environment upon many physical computing environments creates a larger working environment. So virtualization, in a general sense, either creates a smaller or larger working environment that is similar to the underlying hardware.

- Grid computing

  Grid computing handles computationally large problems or applications using the combined resources of a clustered set of servers, in parallel. Grid applications generally must be customized to the requirements of the environment to be processed.

  Grid computing requires the use of software that can divide and farm out pieces of a program as one large system image to several thousand computers. With grid computing, if one piece of the software on a node fails, other pieces of the software on other nodes might fail. This concern can be alleviated if a component has a failover component on another node. However, problems can still arise if components rely on other pieces of software to accomplish one or more grid computing tasks. Large system

images and the associated hardware to operate and maintain them can contribute to large capital and operating expenses.

- Cloud computing

  Whereas grid computing focuses on processing one large problem or application, *cloud computing* focuses on using pooled computing resources to deliver a service to a user. A cloud environment virtualizes local infrastructure to dynamically provide servers, storage, network, and other hardware, otherwise known as *infrastructure as a service* (IaaS). From there, subservices can be delivered, such as platform as a service (PaaS) and software as a service (SaaS). These services are usually delivered and accessed over a network, either the public Internet or a private intranet.

Cloud computing attributes

From the perspective of a client or a user, an effective cloud computing platform has the following attributes:

- Self-service on demand: A client must be able to gain access to a cloud's resources or services. Access must be available 24x7, without requiring the help of a middleman, as though the resources and services belong to the client. This type of access is usually accomplished through a web interface or client-side software.
- Task automation: Clients should be able to start resources and services with the push of a button. The processes that launch these services should be modularized in a way that allows arbitrary technical steps to be grouped together, automated, and abstracted away from the client, creating a seamless and hassle-free user experience.
- Elastic scaling: As workloads increase, so do the amounts of resources that are required to process them. Clients with applications running on the cloud must be able to scale resources to the needs of the workload. This scaling involves features for adding memory, storage, and computing capacity. More advanced cloud solutions can offer a dynamically scaling infrastructure and can allow clients to add infrastructure resources without any downtime.
- Metering: Clients need to know what they are using and the costs that are incurred. Metering is especially important if a cloud solution incorporates billing-by-the-hour or more fine-grained payment models. Metering is also important because it allows clients to understand their business operations better.
- Focus on customer service: Any successful cloud computing service needs to have market demand and be easily usable from the client viewpoint.

The main aspect of cloud computing is the hosting of the physical infrastructure or the computing resources that provide the hosted services. Instead of being managed locally, the infrastructure is managed remotely, typically offsite, by a

third-party provider. This remote management allows the user to eliminate the initial cost of the physical computer resources, along with the cost in both money and time of managing and maintaining those resources. In a sense, cloud computing can be treated as an extension of virtualization.

Cloud computing models

Cloud computing has often been separated into the following types:

- The *public cloud* is the model that comes to mind when most people hear the term *cloud computing*. When a company chooses to implement a public cloud for hosting a solution, the solution will use cloud services that are provided by a third-party vendor through the public Internet, and the solution's data will be stored remotely on the vendor's site.
- A *private cloud* is similar to a public cloud. However, the term is often misinterpreted because clients have trouble differentiating between a local data center and a private cloud. Although a private cloud is indeed local, meaning that data is stored on-site at the company (the solution provider), it still exhibits the cloud computing attributes of an effective cloud solution described previously. In addition, services and resources of a private cloud, although delivered over a network, are generally accessible only through a private intranet. Private clouds can be company owned and operated or they can be managed by a third-party vendor. The commonly accepted definition is that a private cloud solution's data resides in-house.
- A *hybrid cloud* deployment uses both private cloud (in-house) resources and public cloud (external third-party) resources.

Figure 1-2 illustrates these cloud computing models. Each model has advantages and disadvantages, which are described in greater depth in 1.2, “Advantages and disadvantages of using a public cloud versus a private cloud” on page 10.

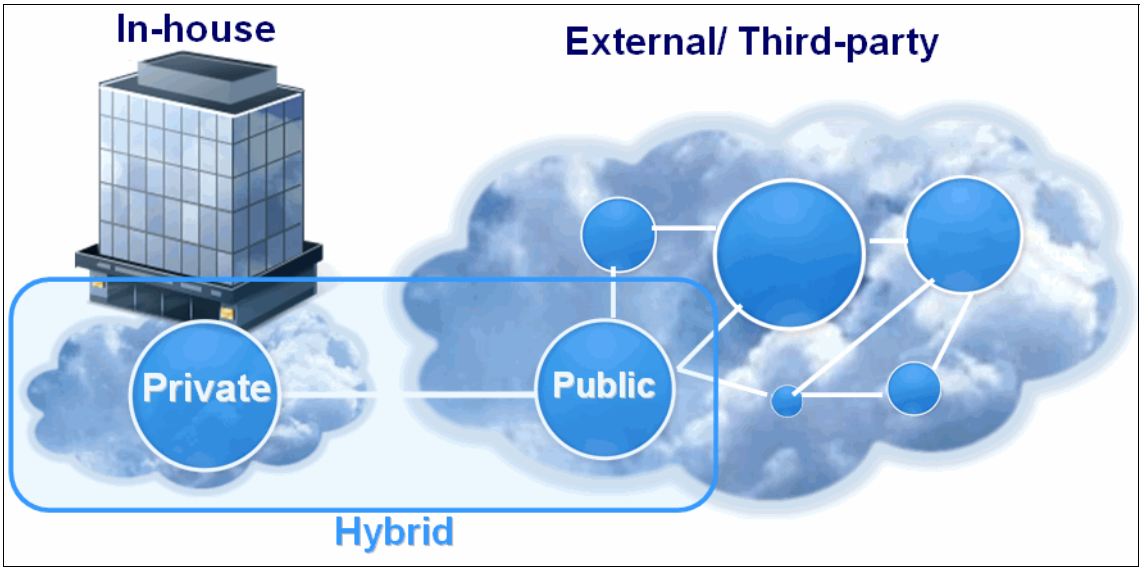


Figure 1-2 Cloud computing models

Cloud service models

The following cloud service models are available today:

- Software as a service (SaaS)

  In the SaaS model, cloud consumers use web-delivered applications without any control of the middleware, operating system, hardware, or networking behind them. Services range from web-based email to inventory control and even business processing. Because the applications and the data are hosted in the cloud, the user can use the service from anywhere over the Internet. Consumers of SaaS tend to be users. Business intelligence SaaS focuses on people working in business development and on line-of-business managers.

- Platform as a service (PaaS)

  PaaS is a solution stack that consists of application development tools, database management systems, development and test environments, and other application services that aid in design, development, collaboration, testing, and so on. PaaS is usually offered as an Internet service to software and application developers to use to create applications.

- Infrastructure as a service (IaaS)

  IaaS provides computing resources, such as virtual server instances with unique IP addresses, blocks of storage, and processing power on demand. Application program interfaces (APIs) are made available to start, stop, access, and configure the virtual servers and storage. Consumers of IaaS tend to be system administrators and software developers.

Figure 1-3 illustrates the typical elements that support each cloud service model.

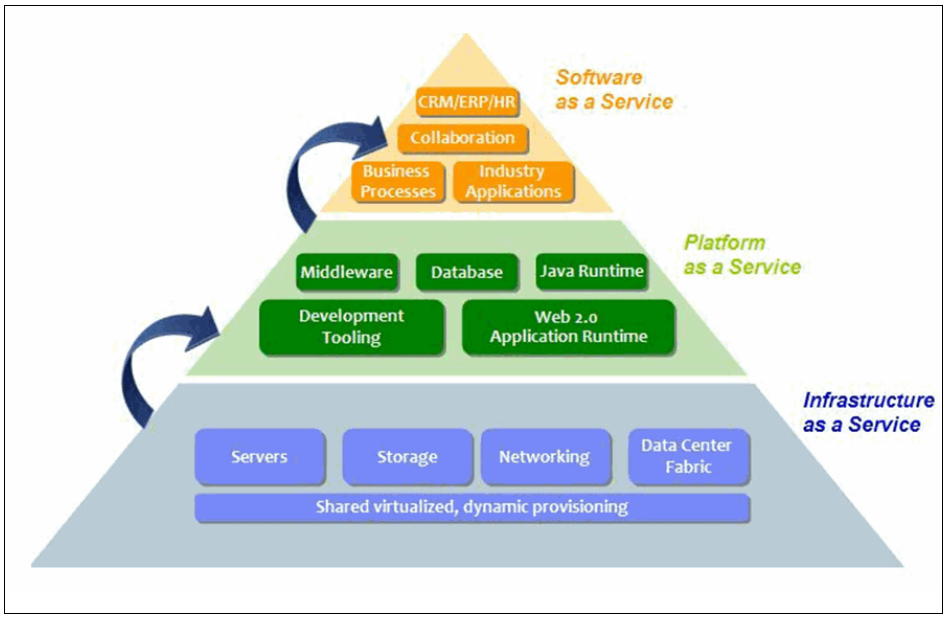


Figure 1-3 Cloud service models: IaaS, PaaS, and SaaS

1.1.2 Reasons for adopting cloud computing

Simply put, cloud is hot right now, with a future that foresees mass adoption by organizations of all sizes, geographies, and industries. These organizations view cloud as a way to reduce costs and complexity of operating traditional IT solutions. Applications for email, music, games, social networking, productivity, customer relationship management (CRM), and other services that were traditionally designed for desktops and servers are now being converted into cloud-based services.

This type of conversion is ultimately the result of the commodization of hardware over the past decade through virtualization. Cloud computing is essentially the vessel that has delivered virtualization to the masses. Thus, as cloud computing matures and as advances are made in the power, features, and usability of the

cloud, the infrastructure will shift from client to solution provider, ultimately allowing clients to focus more on the services that they want to use than on the hassle of managing those services.

Keep in mind the following reasons for adopting cloud computing:

- You use services only when you need to and only pay for a resource or service when you use it.
- Clients do not need to maintain or own the hardware to power the services they need, so they face lower up-front costs. Clients only deal with operational expenditure, not capital expenditure. Furthermore, clients can begin to see positive return on investment much more quickly than with a traditional datacenter deployment.
- Resource efficiency and utilization are improved. Cloud computing, which uses virtualization, allows you to easily allocate resources dynamically.
- Being self-serviceable means cloud services are accessible at all times without the hassle of installing, setting up, and maintaining the service.

Figure 1-4 illustrates how cloud adoption can accelerate its << whose? >>business value.

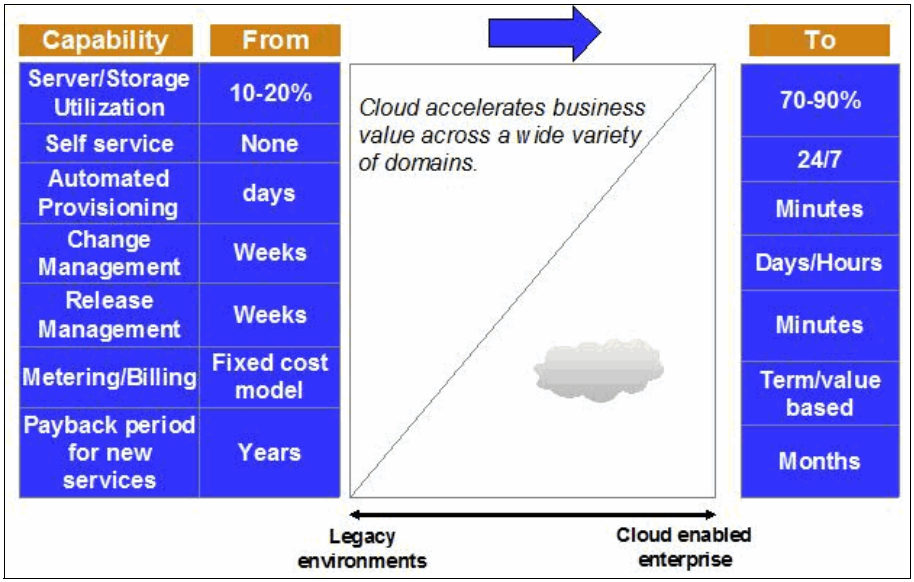


Figure 1-4 Adopting cloud computing

The main hindrance to cloud adoption remains the issue of security for sensitive data. Even industries that handle a lot of sensitive data, such as governments

and financial and healthcare companies, have use for cloud services for their non-sensitive data.

Today, companies are shifting from consuming in-house software to SaaS. By 2015, it is expected that companies will see cloud as a much greater force to impact their business models with more software being offered as services, as shown in Figure 1-5. This shift ultimately will reinvent how companies do business in IT by increasing the number of projects implemented with some form of cloud computing.

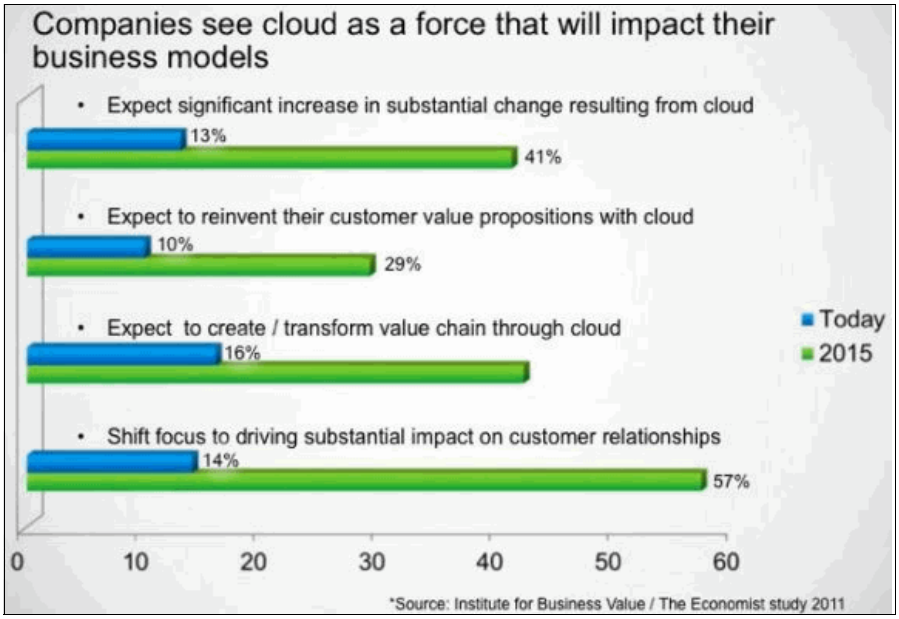


Figure 1-5 Impact of cloud computing on corporate business models

To view a discussion about cloud computing trends, visit:

http://thoughtsoncloud.com/index.php/2011/12/executive-corner-dont-settle-for-the-check-mark/

Small to medium sized businesses that find it too expensive to procure and maintain enterprise-level hardware might look to public cloud services as a cost-effective option for obtaining infrastructure to power enterprise-level services. Larger companies can implement private or public cloud services to enhance pre-existing solutions by adding scalability or redundancy or by offloading other workloads. Cloud services are so broad and useful that there is always a way to incorporate them to enhance an IT solution.

1.2 Advantages and disadvantages of using a public cloud versus a private cloud

When considering whether to use a public or private cloud, you need to consider both the advantages and disadvantages of both types of environments.

1.2.1 Public cloud

A public cloud uses externally-owned resources from third-party service providers and is accessed through a subscription over the internet.

Advantages

A public cloud offers the following advantages:

- Elasticity
  - Use what you need.
  - Turn off (and stop paying for) resources that are unused.
  - Add or change functionality without impacting existing functionality.
- High scalability
  - Access the public cloud's limitless resources.
  - No workload is too big or too small.
- Flexible pricing
  - Lowest up-front cost and capital expenditure.
  - Pay as you go.
  - Lowest risk.
- Faster deployment
  - Resources and services already available and ready to use.
  - Custom services can be created quickly using pre-existing platform tools that are available on most public cloud providers.
- Automatic and rapid provisioning

  Properly configured, automated systems are geographically agnostic, offering a seamless user experience.
- Global availability

  Most public cloud providers have many data centers around the world, enabling network-efficient, global access to your services.

Disadvantages

A public cloud also has the following disadvantages:

- Lower degree of security, control, and technology
  - Data is hosted off-site. Customers with sensitive data requirements might not be able to use the service.
  - Systems are hosted off-site. Although many safeguards and features help ensure that services keep running, control is limited with respect to what you can and cannot do with your systems.
  - Public clouds offer generic or cookie-cutter services that are usually not as customizable as an in-house deployment. For example, there is no ability to choose specific hardware components.
- Highest long-term cost

  A public cloud offering is the most expensive choice for handling a base (predictable) workload. You are essentially renting the infrastructure fully.

When the term *cloud computing* is used, the public cloud deployment model is usually the first model that comes to mind. This model has single-handedly changed the face of modern computing. The public cloud has commoditized enterprise computing, which was once out of reach for small businesses. In addition, the public cloud model helps to overcome challenges, such as low server-utilization rates, operational inefficiencies, complex procurement processes, building and maintaining server environments, long application deployment times, and rising costs.

Public cloud target users

The following companies are suitable for using public cloud services:

- Small to medium sized companies that are looking to reduce initial project spending by reducing capital expenditure and maintenance costs and consuming resources only when needed

  Public cloud services offer flexible pricing schemes that range from pay-as-you-go to fixed-term commitments. Resources are largely maintained by the hosting provider, meaning reduced spending on in-house IT work. You can spin up the necessary cloud servers whenever customers request the service, and then shut down the resources when they are not needed.
- Companies looking for an agile development environment

  Infrastructure, platforms, and software are already created and ready to run at the push of a button. For example, IBM DB2 images are offered on public cloud vendors such as Amazon Web Services, IBM SmartCloud™ Enterprise, and GoGrid cloud. You can get up and running with these images within minutes.

- Companies looking for demonstration environments

  If you have technology that needs to be showcased to a customer, the public cloud can fit perfectly with your need to build proofs of concept or demonstrations. In addition, because services are accessed through the Internet, the services can be deployed anywhere, at any time.

- Companies looking to deploy globally

  Public cloud vendors often have data centers spread around the globe. Thus, you can design a solution to provide services to clients everywhere.

- Large companies that are looking for ways to offload non-critical workloads

  For example, if certain data requires auxiliary batch processing, a public cloud can be employed to handle the work while the in-house data center concentrates on more important activities.

Figure 1-6 illustrates a high-level architecture of a public cloud for business intelligence.

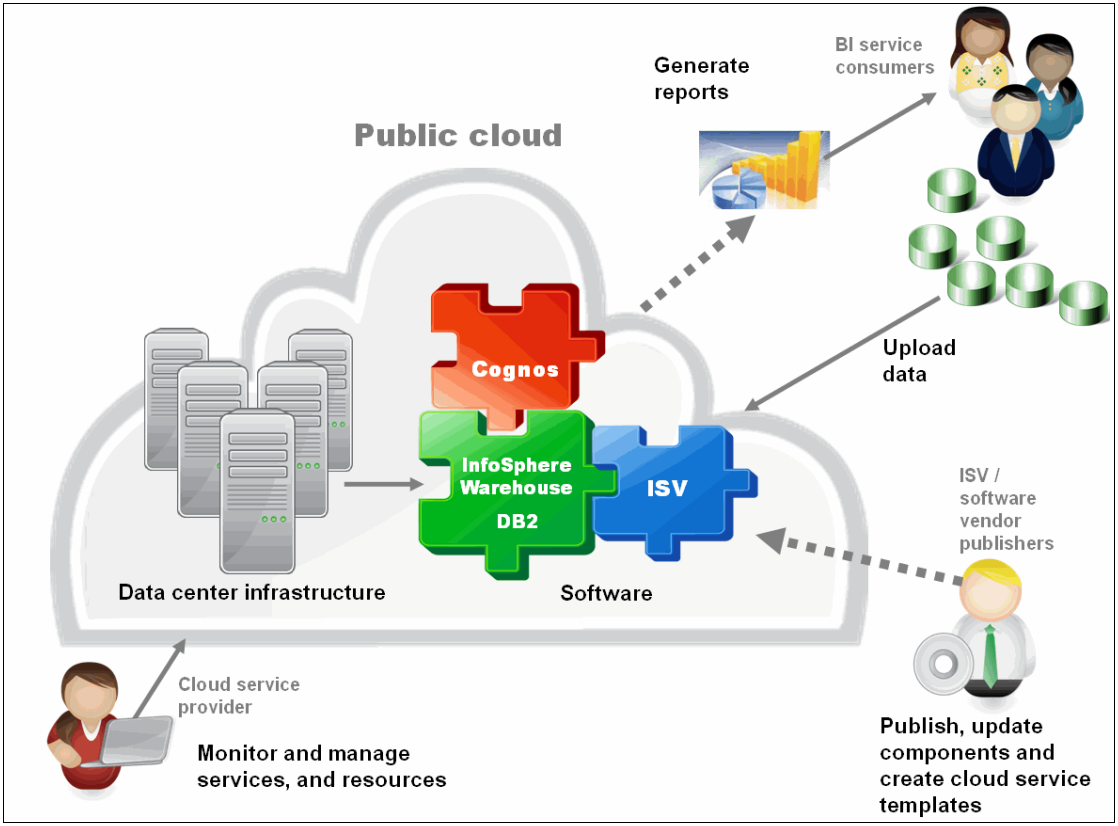


Figure 1-6 Business intelligence public cloud overview

1.2.2 Private cloud

With a private cloud, data is stored on site at the company. In addition, although services and resources are delivered through a network, these services and resources are generally accessible only through a private company intranet.

Advantages

A private cloud offers the following advantages:

- Capitalize on unused hardware
  - Virtualization allows pool resources to be adapted for changing environments, for example from development to test and then to production
  - Savings can be gained by reducing the number of inactive hardware instances
- Highest degree of security, control, and technology customization
  - Data stored locally, systems run locally
  - Restrict access easily
  - Start and stop resources
  - Choose hardware, such as high-bandwidth networks, switches, consoles, CPU, memory, storage, and other types of hardware
- Excellent resource efficiency

  Provisioning is elastic, similar to the public cloud deployment
- Lowest long-term cost
  - Most cost-effective choice for handling base (predictable) workloads, if system is fully used
  - Reduced operational and capital costs thanks to improved cloud lifecycle and reduced complexity
- Automatic and rapid provisioning
  - Properly configured automated systems are geographically agnostic, offering a seamless end-user experience
  - Drives operational efficiency and standardizes leading practices
  - Resources are provisioned on demand to optimize resource utilization

Disadvantages

A private cloud also has the following disadvantages:

- Limited scalability

  Can scale only within the capacity of internal hosted resources

- Inflexible pricing
  - Highest up front cost
  - Hardware must be purchased before meeting any demand
  - Resource use is less efficient if supply is greater than demand

- Restricted to local use

  Can be expensive and difficult to deploy at multiple locations globally

- Slower to deploy than public cloud

  Many tools and services must be built from the ground up because it is a private deployment

- Requires additional skills

  Virtualization and cloud computing expertise is required to create and maintain a private cloud deployment

Private cloud target users

The following companies might benefit from a private cloud:

- Companies that handle sensitive data

  Data in a private cloud is stored in house for the highest degree of security.

- Companies with large amounts of capital to spend on an IT infrastructure

  With a private cloud, hardware must be purchased up front.

- Companies where workloads are predictable

  A private cloud is not as flexible as a public cloud because purchased infrastructure has limitations.

- Companies that are looking to save the most money in the long run

  Private clouds do not have a rental model. Thus, no price markup is involved. You own the resources that you deploy.

- Companies whose solutions require the highest degree of performance and customizability

  All components of a private cloud can be chosen by the company.

1.2.3 Hybrid cloud

A hybrid cloud deployment is a cloud solution that is built on a base that uses a private cloud infrastructure backbone. It incorporates additional resources as needed from a public cloud infrastructure.

Advantages

A hybrid cloud offers the following advantages:

- Base workload requirements are met with the highest degree of reliability
- Ability to respond to increasing workloads by using additional public cloud resources
- Ability to fall back on private cloud resources if public cloud resources are not available
- Ability to comply with specific local laws, such as when sensitive financial and healthcare data must be kept in-house (less sensitive operations can be performed using a public cloud infrastructure)
- Additional layer of disaster recovery when applications and data are in both the private and public cloud, allowing for failover

Disadvantages

The hybrid cloud also has the following disadvantages:

- An increased complexity in application programming due to the need to create interfaces and automation scripts for both private and public cloud infrastructure, and then integrating them in a seamless manner
- Data communication between the private and public cloud infrastructure will not be optimal due to geographical discrepancies
- Best suited for stateless applications

Although public cloud deployments are elastic, scalable, and reduce capital expenditure, these benefits come at the cost of reduced performance and increased security risk. Private cloud solutions perform better and are more secure, but the infrastructure must be paid for up front. The hybrid cloud model attempts to bridge the gap between the two cloud models by extracting their advantages while reducing the impact of their disadvantages.

1.3 Market growth and challenges

Cloud computing is not just a technology. It is also a software delivery platform. As more software is made available as services, companies will shift from

consuming in-house resources to using externally-hosted resources. The cloud landscape will continue to evolve and grow.

New technology, such as cloud computing, is often accompanied by changes in its market and how the technology is consumed and delivered. These changes can create challenges when working with cloud environments. With respect to storage on public cloud, market predictions indicate that security will be a hurdle due to the nature of data ownership. Also, an increasing number of data centers are incorporating virtualization and cloud to increase agility and scalability while decreasing costs through resource efficiency.

Because private cloud deployments are already quite popular today, companies have high expectations for hybrid cloud deployments, either from building a hybrid solution from the ground up or from converting previous private cloud deployments to hybrid.

Finally, the mobile platform market is currently booming, with the popularization of smartphones. The cloud has basically instantiated itself as the de-facto method for delivering multimedia to these mobile devices. As the mobile market continues to grow, so too will the need for cloud computing services.

When you are designing cloud applications: With all this talk about the cloud, it might be hard to believe that cloud computing, as it stands today, is still in its infancy. In fact, despite all the vendors and hosting companies that dabble in cloud, no standardized API for managing cloud resources exists. This situation will change as cloud technology matures. For now, when developing cloud applications, design them with the cloud framework in mind but not necessarily with a specific cloud technology in mind, because the technology is constantly changing.

Cloud computing has also drawn attention to certain areas of security and how data is handled and kept. More specifically, new laws and regulations might be implemented regarding where clients' sensitive data can be kept and who can keep it. For example, is it lawful to store financial data records on a third-party vendor infrastructure? Details regarding these issues can vary from region to region. So, be sure to understand the laws for the location where you are deploying your solution. Other issues can arise with respect to data privacy, especially when incorporating public cloud services.

SaaS is a result of a major trend in the evolution of software, specifically, to make software easier to access and use. As such, cloud services, although easily accessible and usable by today's standards, might not be judged the same way years from now. Software must constantly be updated, methodologies must evolve, and the way customers consume this software will change. So offering a

BI cloud solution, like any IT solution, involves constant change and evolution as the IT landscape changes.

Keep in mind the following challenges that you might encounter when implementing a cloud solution:

- Scaling to future demands

  Demand for cloud services will inevitably grow as more software and services are made available on the cloud. Business that offer cloud services will look for BI services that are easy to access.

- Competing against other BI cloud solutions

  The ISV solution for added value combined with a BI cloud deployment is essential to providing easier access for customers to consume services. Because the landscape for the cloud and BI services is constantly evolving, companies must continue to develop value-added services to stay competitive.

- Changes in hosting infrastructure

  To ensure competitive performance of a BI cloud solution, the underlying infrastructure eventually needs to change or be upgraded. Consider how to handle these changes with the least impact on the solution and business.

- Changes in software and middleware

  New software and middleware offer new features to help keep a BI cloud solution competitive. Consider how to incorporate these changes in a seamless and automated manner to ensure that business operations remain uninterrupted.

1.4 Using IBM cloud enablement technologies for business intelligence

IBM cloud enablement technologies offer integrated solutions of hardware and software in areas of business intelligence, data warehousing, virtualization, and cloud computing. The design and architecture of these solutions minimize the risk and costs, from deployment to maintenance, by allowing you to invest in more efficient ways to deploy software and utilize IT resources.

Decision makers are looking for ways to access real-time and reliable BI services across the entire organization without the complex and expensive challenge of acquiring, deploying, and maintaining a traditional BI infrastructure. With IBM cloud enablement technologies, you can accelerate the deployment of your BI applications when you need to deliver business value quickly. These

technologies can add a greater value to the traditional enterprise software and allow you to deploy your BI applications using different methods and techniques, such as virtualization and cloud computing.

Figure 1-7 illustrates IBM deployment options for BI and data warehousing solutions.

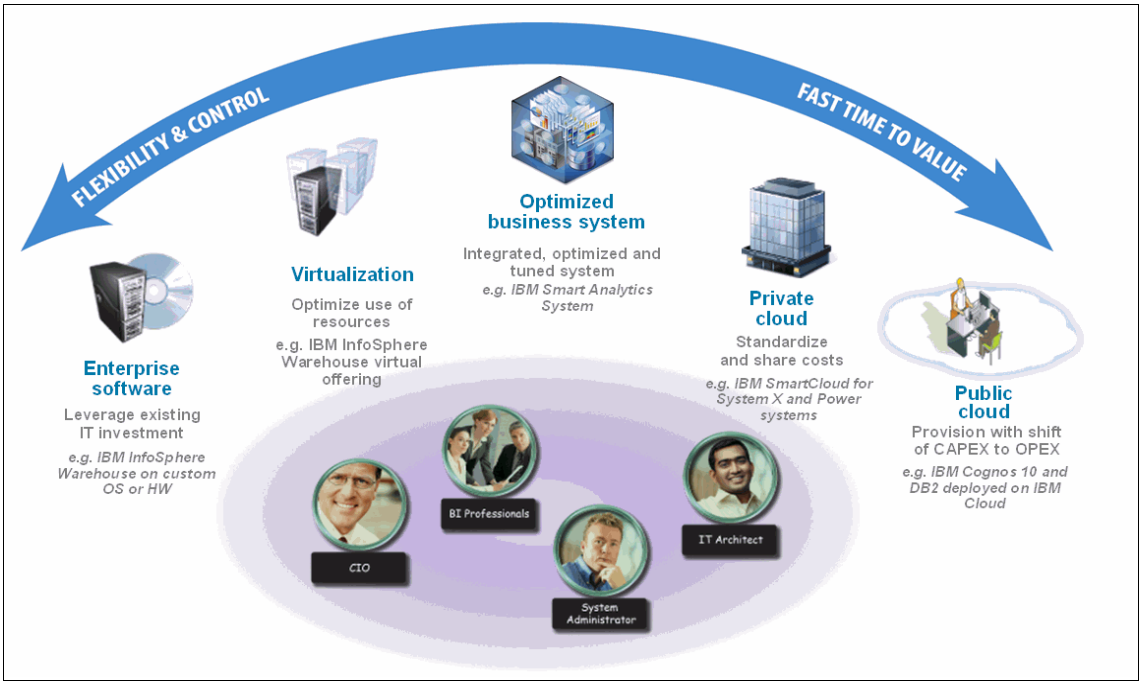


Figure 1-7 IBM deployment options for business intelligence and data warehousing

With virtualization technology, you can ensure that your systems are fully utilized to optimize consumption and that they can accelerate the delivery of desired BI features. Using a virtualization strategy is one of the most common methods of delivering more efficient and affordable BI computing services. For more information, see:

`http://www.ibm.com/systems/virtualization/index.html`

1.4.1 IBM cloud enablement technologies (hybrid cloud)

The underlying technology for the cloud is not specified to the cloud service consumers; however, the cloud service provider needs to ensure that the cloud components can bear and adapt to the vision and strategy of a cloud infrastructure. Thus, these components should be able to support multitenancy

and scale on demand using a set of protocols and standards defined by the cloud service provider.

The cloud physical infrastructure consists of a data center with a pool of servers and storage systems. Those systems are connected to a high-speed network that can sustain high demand communication between all components inside the cloud. The cloud physical infrastructure can be used with virtualization software to bring together all or a subset of computing resources into a unified environment capable of running several operating systems in parallel using a single centralized administration unit.

This administration unit can be integrated with cloud management software to provide control, security, and monitoring capabilities to the overall infrastructure. Additionally middleware software and applications can be installed within the operating system instances to help cloud consumers perform specific tasks. The combination of all these components represents the foundation that lies beneath any cloud offering for BI.

For more information about the cloud infrastructure, see:

http://www.ibm.com/cloud-computing/us/en/cloud-infrastructure.html

Figure 1-8 on page 20 illustrates the typical technologies needed to enable BI for IaaS, PaaS, and SaaS cloud service models. Each of the cloud service models can be designed with integrated IBM hardware and software solutions, and leading practices, to help you build a private or hybrid cloud for BI.

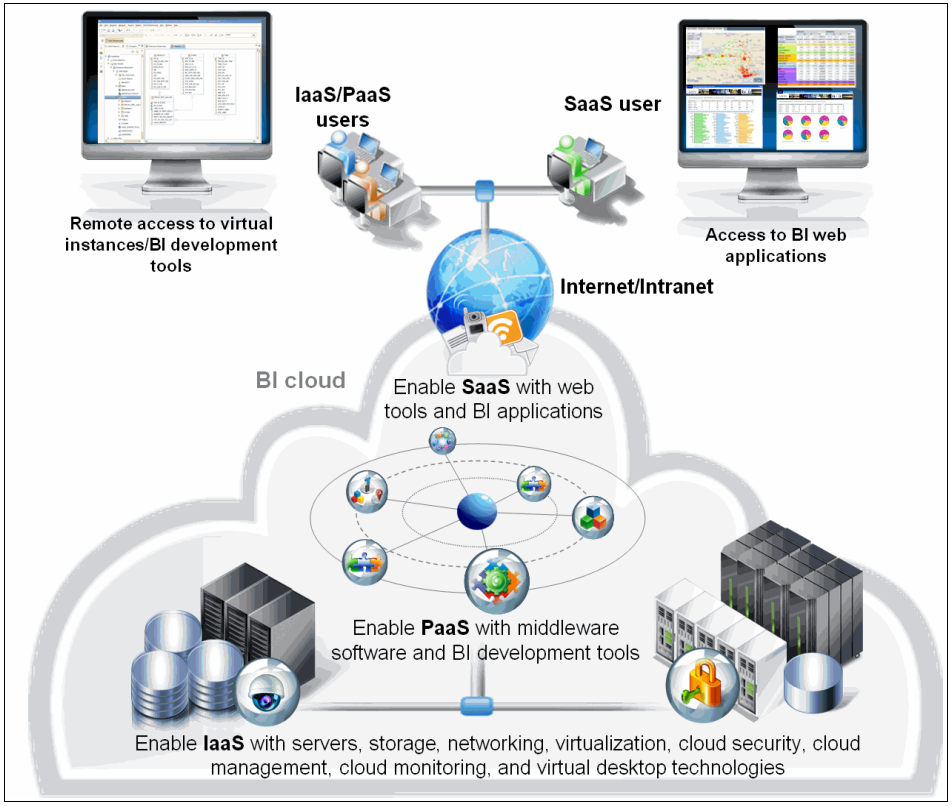


Figure 1-8 Typical technologies to enable BI for IaaS, PaaS, and SaaS delivery models

Technologies to enable your IaaS model

A set of servers, storage, and networking hardware provides a shared pool of resources (memory, disk storage capacity, and CPU) to offer computing power for an IaaS model. IBM hardware can be clustered and virtualized to maximize hardware efficiency, power, cooling, and space requirements of the computing resources. Some technologies that can help you enable your IaaS model are:

- Servers

  A key benefit of the IBM servers is their broad support for a wide variety of hardware, operating systems, and virtualization software from IBM and third-party vendors. The physical server component can include, but is not limited to, IBM System z®, IBM System p®, IBM System x®, and IBM BladeCenter®. For more information, see:

  http://www.ibm.com/systems/

- Storage

  IBM Disk Storage systems support large volumes of data capacity at fast performance rates. For more information, see:

  http://www.ibm.com/systems/storage/index.html

- Networking

  IBM 10 Gigabit Ethernet switches provide high-speed communications between all hardware components for bandwidth-intensive applications. For more information, see:

  http://www.ibm.com/systems/networking/

- Virtualization software

  Control and improve the productivity of cross-hypervisor platforms with IBM Systems Director VMControl™. For more information, see:

  http://www.ibm.com/systems/software/virtualization.html

- Cloud security

  Traditional security technologies are not optimized to work with a virtual environment. IBM Security Virtual Server Protection for VMWare, IBM Injection Logic Engine, IBM Shellcode Heuristics and IBM Security Identity and Access Assurance allow you to secure and protect your virtual environment from security threats. For more information, see:

  http://www.ibm.com/software/tivoli/products/virtual-server-protection/

- Cloud management

  Manage automation, provisioning, and delivery of resources with IBM SmartCloud Provisioning, Tivoli® Service Automation Manager, IBM Service Delivery Manager, or Tivoli Usage and Accounting Manager. For more information, see the IBM SmartCloud provisioning website at:

  http://www.ibm.com/software/tivoli/products/smartcloud-provisioning/

- Cloud monitoring

  Gain visibility and control of your cloud infrastructure with IBM Tivoli Application Performance Management. For more information, see:

  http://www.ibm.com/software/products/us/en/ibmsmarmoni

- Service portal and service catalog

  IBM SmartCloud Control Desk provides management applications for service desk, self-service catalog, and business support services (customer account, order management, subscription management, billing, entitlement management, and IT asset management). For more information, see:

  http://www.ibm.com/software/tivoli/products/smartcloud-controldesk/

- Pre-integrated solutions

  Combine servers, storage, networking and virtualization, and management into a single infrastructure system with IBM SmartCloud Entry on Power, IBM SmartCloud Entry on x86, or IBM BladeCenter Foundation for cloud, and IBM PureSystems™. For more information, see the IBM SmartCloud Entry website at:

  http://www.ibm.com/systems/x/solutions/cloud/starterkit/index.html

  Additional information is available at the IBM PureSystems website:

  http://www.ibm.com/ibm/puresystems/us/en/productfamily.html#!/subtab:ptab-product-family-rpanel-pureflex-system-link

- Virtual desktop solutions

  Simplify the adoption of a virtual desktop infrastructure with IBM Smart Cloud Desktop Infrastructure. For more information, see:

  http://www.ibm.com/systems/x/solutions/desktopvirtualization/index.html

- VM deployer appliance

  IBM Workload Deployer is a hardware appliance with access to IBM middleware virtual images and database patterns. For more information, see:

  http://www.ibm.com/software/webservers/workload-deployer/index.html

- Public cloud

  IBM SmartCloud Enterprise is the enterprise public cloud to access on-demand virtual servers and storage resources for development and test workloads. For more information, see:

  http://www.ibm.com/services/us/en/cloud-enterprise/uses-development-and-test.html

Technologies to enable your PaaS model

Technologies to enable your PaaS model consist of database, application server, and development software to help business users to build applications. The following IBM middleware software offerings provide enterprise technologies to enable BI capabilities for a PaaS delivery model:

- Data warehouse software

  IBM InfoSphere Warehouse is a data management solution that supports high-performance BI applications. For more information, see:

  http://www.ibm.com/software/data/infosphere/warehouse/

- Application server software

  IBM WebSphere® Application Server is a Java application server that provides a runtime environment for web applications. For more information, see:

  http://www.ibm.com/software/webservers/appserv/wasfamily/

- Business intelligence software

  IBM Cognos Business Intelligence is a reporting and analysis solution that provides a collaborative BI workspace for all user reporting types. For more information, see:

  http://www.ibm.com/software/analytics/cognos/business-intelligence/

- Information integration software

  IBM InfoSphere DataStage® and QualityStage® are data integration and cleansing solutions that collect, transform, and clean large volumes of data, ranging from simple to complex data structures. For more information, see:

  http://www.ibm.com/software/data/infosphere/datastage/
  http://www.ibm.com/software/data/infosphere/qualitystage/

- Database security software

  IBM InfoSphere Guardium® is a database security and monitoring solution that provides fine-grained database activity monitoring, database vulnerability assessments, and data encryption. For more information, see:

  http://www.ibm.com/software/data/guardium/

- Performance monitoring software

  IBM Optim™ Performance Manager is a real-time monitoring and performance warehouse solution. For more information, see:

  http://www.ibm.com/software/data/optim/performance-manager/

- Identity management software

  Tivoli Directory Server is an enterprise identify management solution that uses the Lightweight Directory Access Protocol to provide trusted-identity authentication. For more information, see:

  http://www.ibm.com/software/tivoli/products/directory-server/

- Pre-built data structures and applications

  Industry Models and InfoSphere Warehouse Packs are data model blueprints that can accelerate the development of BI applications. For more information, see:

  http://www.ibm.com/software/data/industry-models/
  http://www.ibm.com/software/data/infosphere/warehouse/packs/

- Concurrent versions system

  IBM Rational® ClearCase® is a version control solution that provides workspace management and parallel development support. For more information, see:

  http://www.ibm.com/software/awdtools/clearcase/

- Front-end development tools

  InfoSphere Warehouse Design Studio, Data Studio, and InfoSphere Data Architect are design and development tools that can be used to create BI applications and content on the cloud. For more information, visit the following websites:

  http://www.ibm.com/software/data/optim/data-studio/
  http://www.ibm.com/software/data/optim/data-architect/

Technologies to enable your SaaS model

In the SaaS delivery model, users can access applications using a thin client through a web browser, most commonly through a subscription fee. The SaaS applications are typically cloned to support scalability and manage a large number of users. Although the applications can share the underlying IaaS and PaaS infrastructure, each cloud user's application is isolated to a sufficient degree that it can provide a unique BI application sandbox to the cloud service consumer. IBM Software provides some web tools to deliver BI and data warehousing capabilities:

- Front-end web tools

  Data Studio Web Console, InfoSphere Warehouse Administration Console, Optim Perfomance Web Console, and IBM Cognos Connection (Report Studio, Query Studio, Analysis Studio, and Business Insight) are web tools that can be used as a cloud front-end to run, administer, and monitor BI

applications. Many of these web tools are deployed as part of the IBM software to enable a SaaS model.

- Front-end mobile tool

  IBM Cognos Mobile extends IBM Cognos Business Intelligence by allowing users to consume its content on mobile devices without compromising security. For more information, see:

  http://www.ibm.com/software/analytics/cognos/mobile/

1.4.2 IBM InfoSphere Warehouse 10.1

IBM InfoSphere Warehouse provides a foundation of trusted information that is built upon rich data management capabilities. This foundation can support BI in a cloud environment with high performance satisfaction. InfoSphere Warehouse enables massively parallel processing (MPP), which ensures that the initial investment can continue to perform as the volume of data grows and can satisfy multiple requests for large amounts of information. InfoSphere Warehouse is a suitable fit for BI on a cloud because of its ability to scale, share multiple tenants, and manage thousands of users with a variety of workloads and requirements.

Figure 1-9 illustrates the InfoSphere Warehouse key components for a BI foundation.

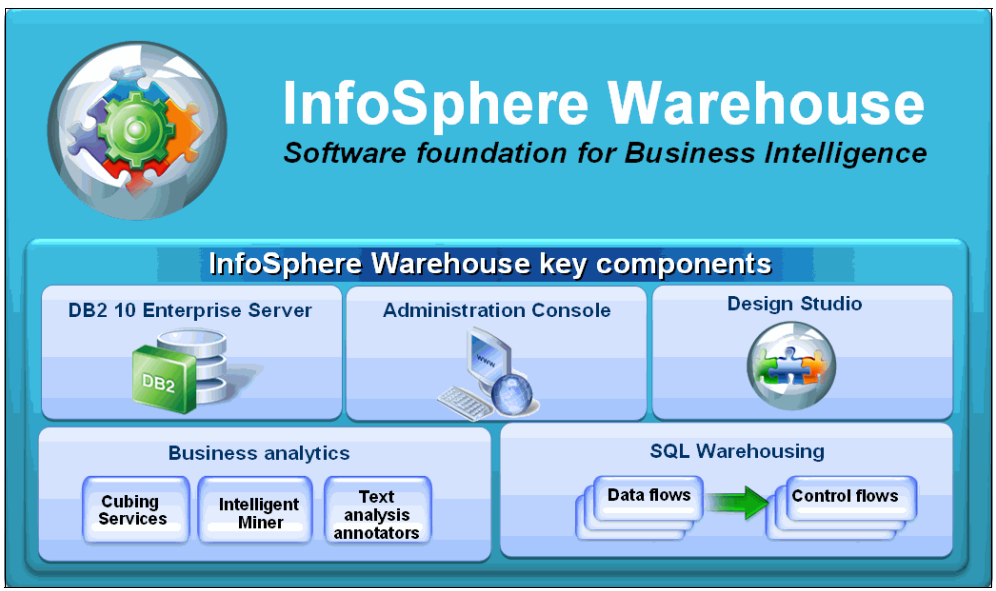


Figure 1-9 InfoSphere Warehouse as a software foundation for BI on cloud

You can use InfoSphere Warehouse to build a complete data warehousing and business analytics solution. The sections that follow highlight key components of the InfoSphere Warehouse solution.

IBM DB2 10.1 Enterprise Server

IBM DB2 is an industry-leading relational database management system (RDBMS) that ships with InfoSphere Warehouse to provide a common repository. This repository can be used to combine data from different systems, formats, structures, and vendors; it eliminates the need for heterogeneous queries against many operational data stores. DB2 has many characteristics that can help you to achieve the following benefits:

- Lower the cost of managing data by automating administration.
- Reduce the cost of storage with data compression technologies.
- Optimize workload execution and deliver scalability with data partitioning.

The following DB2 10.1 features support and enhance BI cloud requirements:

- The database partitioning feature

  This feature addresses performance requirements for large data warehouses by distributing data over multiple servers, also referred to as *nodes*. More nodes means more processor, memory, and disk capacity is available to process more complex database workloads. These workloads can be distributed over many processors with their own memory and disk resources to enable hardware parallelism (see Figure 1-10). This type of distribution is called a *share nothing architecture*. Each system processes only its part of the query and returns the results to the administration node, which then performs the final processing before returning to the user application.

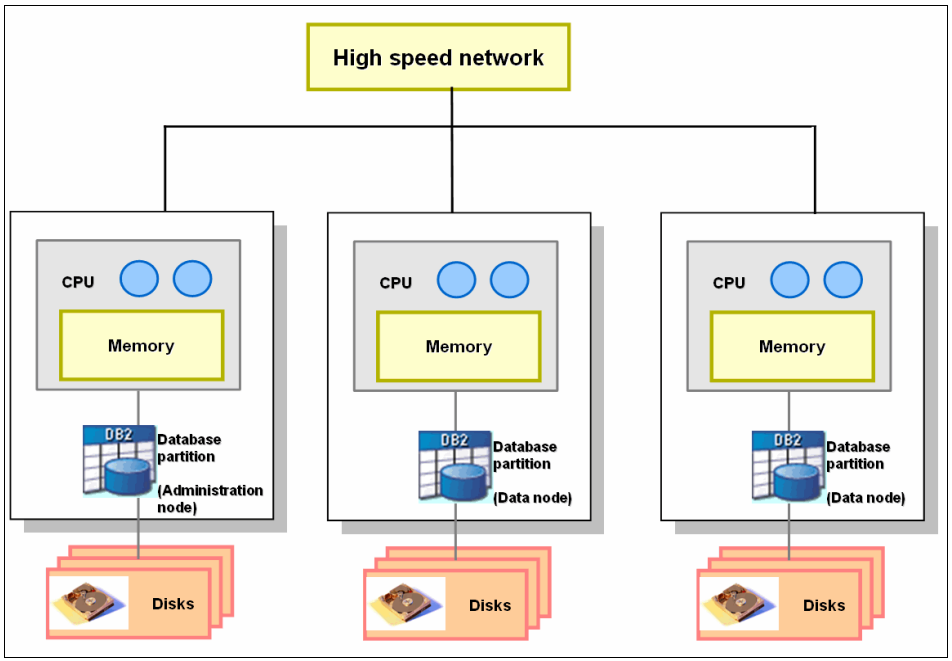


Figure 1-10 The DB2 database partitioning feature

The database partitioning feature is suitable for large tables, such as fact tables, which often contain hundreds of millions or billions of rows, especially if there is significant scanning, aggregation, or complex calculations of the data before obtaining the final result. Taking advantage of the database partitioning feature allows you to achieve near linear scalability that lets you manage petabyte volumes of data while still meeting your SLAs.

Important: The DB2 database partitioning feature is available only for use within InfoSphere Warehouse. It cannot be acquired separately for inclusion with DB2.

- Workload management

  InfoSphere Warehouse was designed for high levels of user concurrency and mixed workloads, or *random I/O*, which is common in an enterprise data warehouse. DB2 Workload Manager allows administrators to prioritize queries coming from different users and applications and to control the number of underlying resources dedicated to those processes.

  This feature can help reduce the risk of SLA failure by prioritizing execution of business-critical workloads. Load balancing can result in better utilization of existing hardware, which can help delay a server upgrade. The warehouse's

design for high levels of concurrency together with workload management ensures business continuity, especially on peak hours.

- Autonomics for performance tuning

  The autonomics features in DB2 can help lower the cost of managing data in a warehouse by automating administration, increasing storage efficiency, and improving performance. By automating tasks, such as memory allocation, storage management, and business policy maintenance, DB2 can perform many management tasks itself, freeing up database administrators (DBAs) to focus on new projects.

  Autonomics includes the following features:

  – A self-tuning memory feature responds to significant changes in workload by dynamically distributing available memory resources among several memory consumers for the database. The DB2 self-tuning memory manager (STMM) can help you avoid critical impact to database response times that threaten your SLAs.

  – A self-configuration feature automatically configures memory, storage, and maintenance operations in DB2 databases. This allows users to be confident that the DB2 databases and DB2 instances are configured correctly and enabled for a better automated operation.

- Deep compression

  Disk storage systems can often be the most expensive components of a database solution. Even a small reduction in the storage subsystem can result in substantial cost savings.

  The DB2 Storage Optimization feature transparently compresses data on disk to decrease disk space and storage infrastructure requirements and improve query performance. The DB2 row compression feature allows storage savings using a dictionary-based symbol table that records repeated values and replaces that data with a small pointer to the values in the dictionary. Because data warehouses often contain many repeating patterns, high compression levels can be achieved. The feature is complemented by value, XML, index, and backup compression.

SQL Warehousing

InfoSphere Warehouse uses a data warehouse industry process called extract, transform, and load (ETL), which allows you to extract data from disparate data sources, such as flat files and database management systems from different vendors, and to consolidate relevant information into a centralized data warehouse, as shown in Figure 1-11. When a more complete set of information is used, business users can make better decisions in the everyday business operations.

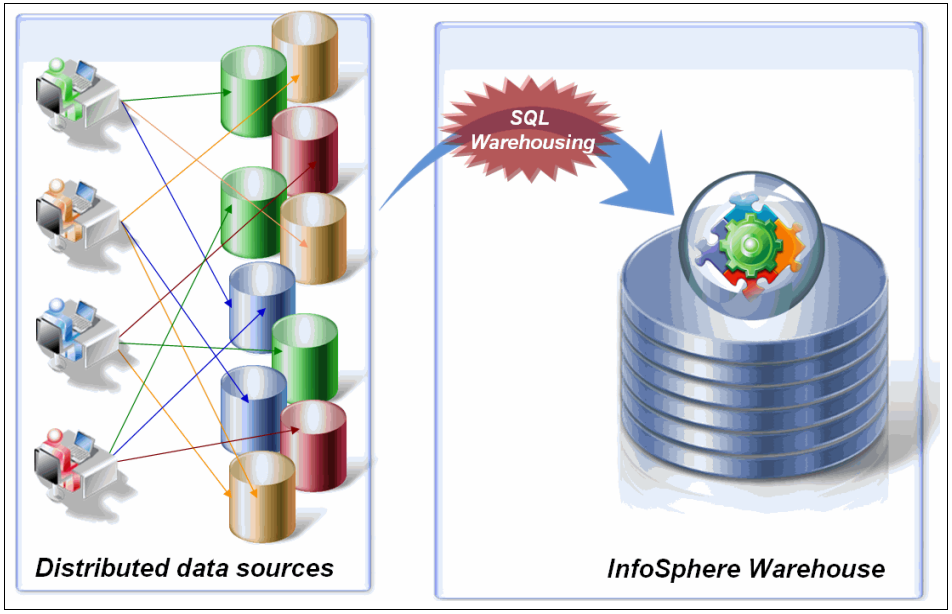


Figure 1-11 SQL Warehousing to consolidate distributed data sources into InfoSphere Warehouse

SQL Warehousing allows you to accelerate the deployment of reusable and powerful warehouse applications that can execute multiple functions to consolidate individual silos of information into a common structure that is optimized for BI. SQL Warehousing provides you with the ability to visually design SQL-based data movement and transformation flows using a prebuilt library of functions.

The SQL Warehousing component consists of data flows, control flows, and data warehousing applications that can be optimized and executed in DB2. For more information, see “ETL development services with Design Studio” on page 200.

Business analytics

InfoSphere Warehouse is used as a data warehouse repository and it also integrates enhanced business analytics features, such as OLAP cube servers, data mining, and text analysis, as illustrated in Figure 1-12.

Figure 1-12 InfoSphere Warehouse business analytics components

InfoSphere Warehouse gives the user the ability to execute these complex analytics inside the DB2 database server, without requiring data extraction. The data-consuming process becomes less complex, accessible to the latest data, and scalable to large volumes of analysis.

InfoSphere Warehouse includes the following business analytics components:

- IBM Intelligent Miner®

  This component is an extension to the SQL Warehousing component. It adds the ability to develop data mining models graphically, using mining flows to do data preparation, model creation, visualization, extraction, and scoring.

- Cubing Services

  Cubing Services is a multidimensional analytics component built into InfoSphere Warehouse to deliver standard OLAP access using multidimensional expressions (MDX) language, which enables support for drill-down, and slice-and-dice analysis. This component also enables the design of OLAP structures based on facts and dimensions physical tables, and provides wizards to optimize the OLAP processing.

- Text analysis annotators

  Unstructured text integration can enable a greater amount of knowledge because it is the fastest growing set of information for enterprises. Text analysis annotators are modules that analyze unstructured information to discover hidden knowledge from text-based data, such as call center logs, reports, claims, records, social media, and so on. They are based on a industry standard called *Unstructured Information Management Architecture* (UIMA), which defines a common interface that enables the development of new and reuse of existing annotators.

Design Studio

Design Studio is an integrated development environment (IDE) used to design, organize, and maintain your data warehouse projects for InfoSphere Warehouse. Design Studio has a unified graphical environment with built-in functions, which supports all users tasks for a typical data warehouse and BI project. This type of project often goes through multiple iterations of design when user requirements change over time. Consequently, Design Studio can create warehouse projects that can be replicated, recycled, and modified multiple times to accelerate the typical data warehouse development lifecycle tasks, such as dimensional modeling, physical database design, ETL design, and BI design, as illustrated in Figure 1-13 on page 32.

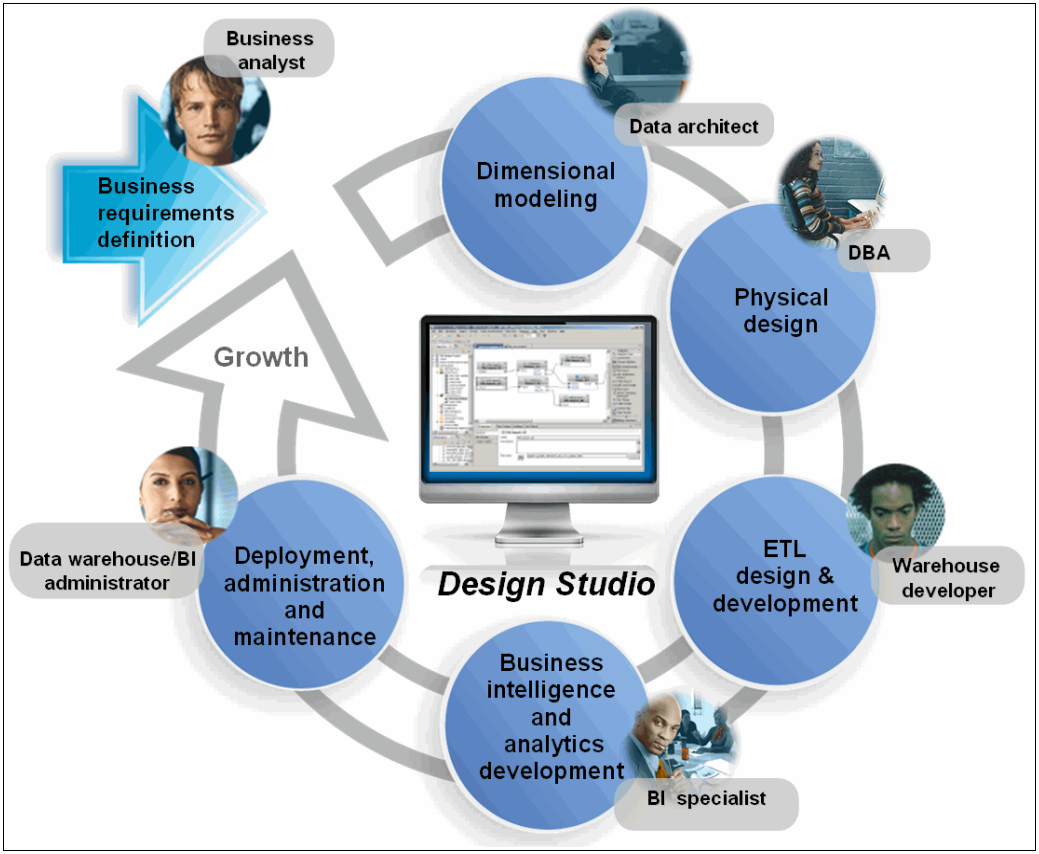


Figure 1-13 Design Studio accelerates the data warehouse development lifecycle tasks

Administration Console

Subsequent to the warehouse and BI applications design, the production environment of InfoSphere Warehouse is deployed, administered, and maintained from a web-based administration tool, called *Administration Console*. This web tool allows the deployment and execution of warehouse applications and OLAP structures. It also enables administration of system resources, data source connections, and mining models, as shown in Figure 1-14.

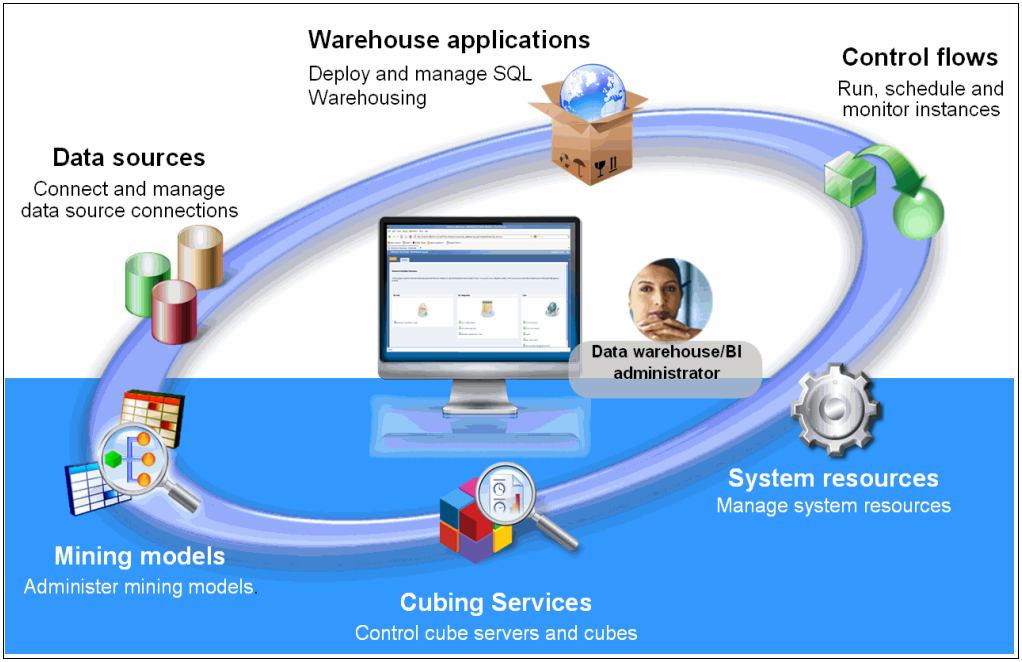


Figure 1-14 Administration Console capabilities

1.4.3 Feature enhancements in DB2 10.1

Today's warehouse solutions demand features to handle different business insights and are increasingly challenged to contain costs and to demonstrate business value faster than before. DB2 10.1, as a part of InfoSphere Warehouse 10.1 includes the following enhancements that can help achieve these objectives:

- Continuous data ingest

  The ingest utility allows the high-speed ingest of data from files and pipes into DB2 tables, using SQL-like commands to process large amounts of real-time data without affecting availability. It is a specialized utility for continually populating data warehouse tables with minimal impact on concurrent user workload and data server resources. Thus, it reduces the overall, end-to-end time from when the source data arrives to when it is in the warehouse and available for queries.

- Multi-temperature storage

  Data warehouses retain large volumes of data for real-time business analytics. However, there is a strong tendency for a relatively small proportion of data to be hot data, and the majority of the data to be cold data. Current

data is often considered to be hot data, but it typically becomes cold as it ages. The challenge faced by organizations is prioritizing the storage of data.

DB2 10.1 introduces a multi-temperature storage feature to achieve a better tradeoff between performance and cost for hot and cold data. Assign high-speed storage (for example, SSD) to critical business analytics and decision support systems (DSS).

- Adaptive compression

  Enhancements to the industry-leading compression technologies of InfoSphere Warehouse come with the adaptive compression feature, which can help further reduce storage needs. The enhancements are designed to deliver efficient compression of high amounts of new and changing data. The improved compression ratio can further help to reduce storage needs and allows for more data in memory, thus helping to increase performance. The adaptive compression approach also reduces the need for table reorganization. As a consequence, the overall maintenance of compressed data is reduced, which can help achieve additional cost savings.

- Row and column access control (RCAC)

  DB2 10.1 provides critical enhancements to security and auditing. RCAC can help you further secure your data. RCAC is sometimes referred to as *fine-grained access control*.

  RCAC security allows you to create varying security rules at the data level. These security rules ensure that users in approved roles or groups see only the data that they are allowed to see. These rules also remove the security constraints and performance headaches that result from complex views and predicates.

- Multi-core parallelism

  Data warehouse workloads can benefit from parallelization. Multi-core parallelism is an improvement from query parallelism for a single-partition database or for partitions of a partitioned database. Multi-core parallelism is also called *intra-partition parallelism* or *SMP parallelism*. Components can be run in parallel and factors are interrelated to enhance performance.

- Star schema optimization

  DB2 10.1 has better detection and performance of star schemas, hash join improvements and a new zig-zag join mechanism, which result in minimal tuning on complex large-volume ad hoc queries. The zig-zag join calculates the cartesian product of the dimension tables, and then probes the fact table using a multi-column index, which efficiently queries large fact tables.

- Time travel

  Temporal tables allow you to view and report on data at a specific time in the past, so you can better understand how the business was doing before and

what has changed in the present. Temporal tables are managed by DB2 but also support custom business time data and manipulation.

1.4.4 InfoSphere Warehouse companion products

Additional IBM products are included or available in special InfoSphere Warehouse 10.1 editions, such as data management, data integration, and DB2 advanced recovery solutions, Cognos Business Intelligence, and InfoSphere Warehouse Packs, as shown in Figure 1-15.

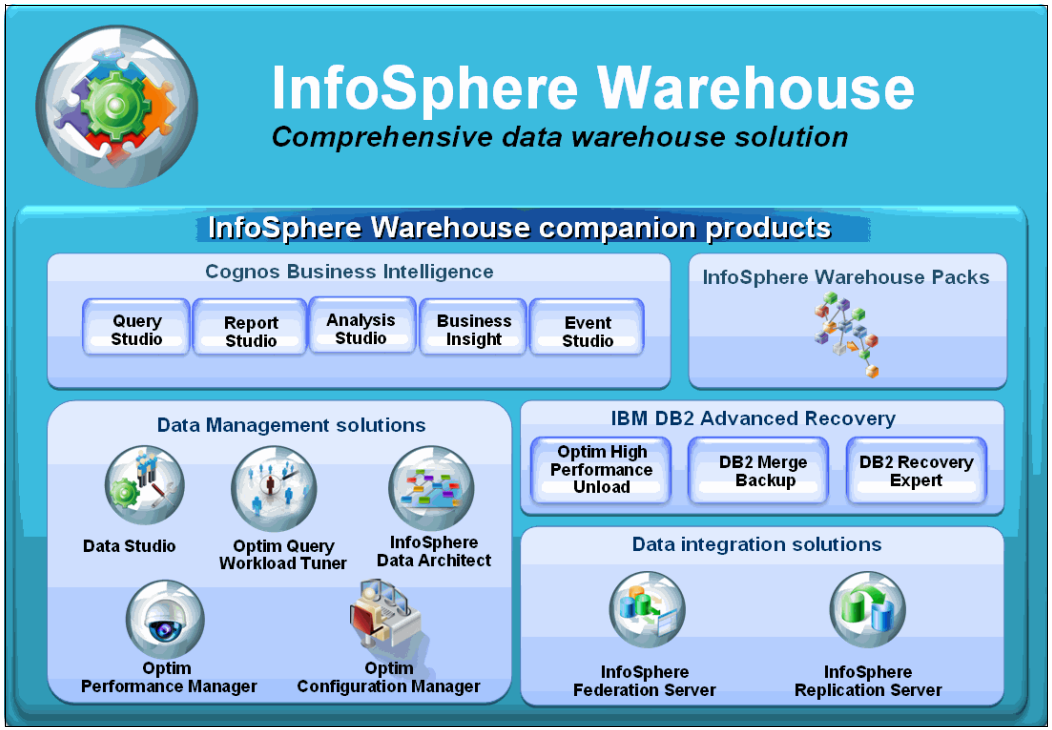


Figure 1-15 InfoSphere Warehouse components and features

You need to download, install, and configure some of these components. For additional details, refer to the licensing information for the InfoSphere Warehouse editions, as described in 1.4.5, “InfoSphere Warehouse editions and differences” on page 39.

IBM Data Management solutions

The IBM data management solutions described in this section provide the foundation and tools for enhancing the administration and maintenance of data.

These tools can help you meet SLAs and solve performance problems before they impact the business.

- IBM Data Studio

  IBM Data Studio is not just another integrated development environment (IDE) for application development, it is a full data lifecycle management tool. Data Studio provides a basic set of tools to administer and develop your database; Data Studio Health Monitor allows you to view health information, alerts, application connections, and utilities related to the system.

- InfoSphere Data Architect

  Data modeling is the first step in creating a data warehouse. You might realize that the current physical definitions of your data model need to be modified. You might need additional columns or referential integrity that is not currently implemented.

  A subset of the data modeling component was added to InfoSphere Warehouse using an Eclipse plug-in (physical data modeling). With this plug-in, existing data structures can be reverse-engineered and new physical data structures can be defined. One tool that can help is InfoSphere Data Architect, which can be used to extend the data modeling capabilities into logical modeling, logical dimensional modeling, and business glossary features.

- InfoSphere Optim Performance Manager

  Optim Performance Manager can be used to isolate performance problems by monitoring bufferpools, I/O, locking, logging, memory, active SQL, system utilities, and workload. It offers highly accessible monitoring with a new browser-based user interface. InfoSphere Optim Performance Manager Extended Edition provides extensive dashboards and alerts to identify a problem at any time from any location. This alert includes identifying locking conflicts, deadlocks, problematic applications or queries, constrained buffer pool cache and heap sizes, CPU, memory, file system shortages, and data skew conditions. DBAs can also set thresholds so alerts can be automatically triggered by a particular set of circumstances.

- InfoSphere Optim Query Workload Tuner

  IBM InfoSphere Optim Query Workload Tuner empowers administrators to more efficiently manage performance by proactively optimizing the performance of SQL queries and query workloads. Built on an open-source, Eclipse-based environment, InfoSphere Optim Query Workload Tuner makes it easy to access candidate queries and define workloads from a number of common sources. DBAs can efficiently capture, format, and analyze SQL statements and workloads, creating graphic visualizations of query plans and costs. For administrators, the solution removes much of the tedium and time

from query analysis by formatting queries for readability, annotating them with statistics, visualizing the access plan, and more.

- InfoSphere Optim Configuration Manager

  IBM InfoSphere Optim Configuration Manager for DB2 for Linux, UNIX, and Windows offers centralized management of database and client configuration based on IBM DB2 technology and is designed to manage data more effectively and efficiently. Centralized client connection definitions can improve manageability by tracking client access to data servers and providing information, such as host name, IP address, user ID, client version information, and more. Properties of deployed database clients can be modified to improve database transaction performance, enforce security and privacy controls, allocate data clients among data servers for desired workload balance, and redirect data clients to connect to a different data server.

Data integration solutions

As organizations try to view their customers and business in a unified way, they are quickly discovering that the individual silos of information that have been created or acquired over the years are now a major obstacle in optimizing their business processes and relationships with customers. The following IBM Data Integration solutions can help you to synchronize information across heterogeneous environments and provide a consistent view of information without impacting the source system performance:

- InfoSphere Replication Server

  IBM InfoSphere Replication Server provides a high-volume, low-latency data replication solution that uses WebSphere MQ message queues for high availability and disaster recovery, data synchronization, and data distribution.

- InfoSphere Federation Server

  IBM InfoSphere Federation Server allows organizations to virtualize their data and provide information in a form that applications and users need while hiding the complexity of the underlying sources. Data virtualization allows information to be accessed through a common interface that centralizes the control of data access.

IBM DB2 Advanced Recovery solutions

The following IBM DB2 Advanced Recovery solutions offer an effective backup and recovery strategy to help meet SLAs, reduce errors, and reduce recovery time:

- Optim High Performance Unload

  The primary function of the DB2 High Performance Unload tool is unloading very large tables from DB2 rapidly; often, four to six times faster than when

using the DB2 export command. It achieves this speed by reading the table space files directly and processing them itself rather then using standard DB2 services. As a result, it cannot do complex SQL query-type processing. It trades away that capability for high speed unloading of the data.

High Performance Unload can filter columns and rows with simple predicates and even do some basic transformations of the data while it is being unloaded. High Performance Unload is designed to be an unload utility rather than a query tool. It is highly threaded to exploit the hardware to perform the unloads in the shortest possible time through parallel processing.

- DB2 Recovery Expert

  DB2 Recovery Expert is separate from the database, which allows it to run without taking operations offline. It only needs the artifacts of the database, such as the database logs and the backup files, to figure out the best path to recovery, and reduces the amount of disruption during the recovery process.

- DB2 Merge Backup

  DB2 Merge Backup can help reduce application downtime, manage backups more efficiently and easily, and recover more quickly to meet SLAs. For example, if backups are large and are now taking longer, it is often perceived that you can solve the problem by buying a newer, faster tape drive or higher performance storage. So, when grappling with these kinds of issues, consider a well thought out backup and recovery strategy.

IBM Cognos Business Intelligence

IBM Cognos Business Intelligence is the preferred reporting tool to integrate with InfoSphere Warehouse as the front-end of a virtualized business intelligence environment.

IBM Cognos Business Intelligence delivers reporting capabilities in a single product, with one platform that can address a large range of users, from executives to line managers to sales and support staff. Business users can access any data or combination of data and visualize it through a number of access modes, including zero-footprint web browser, enterprise search, and even on a mobile device. The platform is founded on a web-services-based service-oriented architecture to maximize integration and delivery across all products.

InfoSphere Warehouse Packs

Few organizations have the time, budget, and specialized IT staff to build all capabilities from scratch. Many departments or branch offices within larger enterprises also lack the resources to implement business analytics from scratch. These organizations can use the following IBM solutions that are ready

for immediate use to simplify and speed deployment of critical business analytics while minimizing risk and cost.

IBM InfoSphere Warehouse Packs are designed to speed time-to-value by allowing organizations to focus on business issues and innovation rather than database plumbing. The packs are accelerator add-ons for InfoSphere Warehouse installations. They include complete physical data models and sample reports based on specific business issues that organizations face every day.

The following packs are designed to reduce data warehousing project time, business risk, and deployment cost:

- Customer Insight Warehouse Pack

  This pack addresses a range of business issues common to organizations that are trying to better understand the profitability and profiles of their typical customers. It includes sample Cognos reports in the areas of communications analysis, transaction analysis, periodic analysis, and profitability analysis.

- Marketing and Campaign Insight Pack

  This pack is designed to help organizations understand the overall details of their market and determine how to acquire new customers in a competitive environment. It includes sample Cognos reports in the areas of market analysis and campaign analysis.

- Supply Chain Insight Pack

  This pack is designed to help organizations ensure that they have the correct products or components in the correct place at the correct time, without hurting cash flow by overinvesting in inventory. For many organizations, efficient supply chain management is fundamental to success.

1.4.5 InfoSphere Warehouse editions and differences

The following editions of InfoSphere Warehouse 10.1 are available:

- Departmental Edition

  This edition is ideal for small and medium organizations or departments that have moved beyond basic querying and reporting and now need tools to discover new business opportunities. Capabilities include mixed workload management and advanced analytics, such as cubing services, data mining, text analytics, and Cognos 10 BI with no physical server licensing restrictions. The Departmental Edition offers the ideal platform for consolidating your data marts and reducing your overall software, hardware, and maintenance costs.

- Advanced Departmental Edition

  This edition targets the same organizations as the Departmental Edition, except that the advanced edition also includes the customer insight, campaign insight, and supply chain warehouse packs.

- Enterprise Edition

  This edition is ideal for enterprise-scale data warehouses in medium to large organizations that require scalability and manageability along with a complete range of analytic capabilities. This edition provides a complete enterprise offering at a cost that is substantially less than the sum of its components, providing real value in terms of price and solution completeness.

- Advanced Enterprise Edition

  This edition targets the same organizations as Enterprise Edition, except the Advanced Enterprise Edition delivers the following set of additional management tools that let you deploy, manage, grow, and architect increasingly business-critical and complex warehouse environments:

  – InfoSphere Data Architect
  – Optim Query Tuner
  – Optim High Performance Unload
  – DB2 Merge Backup
  – DB2 Recovery Expert

  This edition also includes InfoSphere Warehouse Packs to accelerate the insight into markets and campaigns, customers, and supply chains.

- Developer Edition

  This edition provides developers a non-production environment in which to design and build innovative capabilities for InfoSphere Warehouse.

Table 1-1 illustrates the components and features included in each InfoSphere Warehouse 10.1 edition.

Table 1-1 Edition comparison table

	Departmental	**Advanced Departmental**	**Enterprise**	**Advanced Enterprise**	**Developer**
DB2 10 Enterprise Server Edition with Database Partitioning	Yes	Yes	Yes	Yes	Yes
Continuous Data Ingest	N/A	Yes	Yes	Yes	Yes

	Departmental	Advanced Departmental	Enterprise	Advanced Enterprise	Developer
Multi-temperature Data Management	N/A	Yes	Yes	Yes	Yes
Storage Optimization Feature (with Adaptive Compression)	Yes	Yes	Yes	Yes	Yes
Label and Row Based Access Control	Yes	Yes	Yes	Yes	Yes
DB2 Workload Management	Yes	Yes	Yes	Yes	Yes
SQL Warehousing Tool	Yes	Yes	Yes	Yes	Yes
Intelligent Miner	Yes	Yes	Yes	Yes	Yes
Cubing Services	Yes	Yes	Yes	Yes	Yes
Text Analysis	Yes	Yes	Yes	Yes	Yes
Design Studio	Yes	Yes	Yes	Yes	Yes
Administration Console	Yes	Yes	Yes	Yes	Yes
Data Studio	Yes	Yes	Yes	Yes	Yes
InfoSphere Federation Server	Yes	Yes	Yes	Yes	Yes
InfoSphere Replication Server[a]	Yes	Yes	Yes	Yes	Yes
Optim Performance Manager	Yes	Yes	Extended Edition included	Extended Edition included	Extended Edition included
Cognos 10 BI	5 users	5 users	5 users	5 users	1 user

	Departmental	Advanced Departmental	Enterprise	Advanced Enterprise	Developer
InfoSphere Data Architect	N/A	N/A	N/A	10 users	N/A
Optim Query Workload Tuner	N/A	Yes	N/A	Yes	N/A
Optim High Performance Unload	N/A	N/A	N/A	Yes	N/A
DB2 Merge Backup	N/A	N/A	N/A	Yes	N/A
DB2 Recovery Expert	N/A	N/A	N/A	Yes	N/A
DB2 Configuration Manager	Yes	Yes	N/A	Yes	N/A
InfoSphere Warehouse Pack Customer Insight	available	Yes	Available	Yes	Available
InfoSphere Warehouse Pack Market and Campaign Insight	Available	Yes	Available	Yes	Available
InfoSphere Warehouse Pack Supply Chain Insight	Available	Yes	Available	Yes	Available

a. Certain usage restrictions apply to these components. For details, refer to the program's license information document.

Table 1-2 lists the platform support, licensing, and pricing options in each InfoSphere Warehouse 10.1 edition.

Table 1-2 InfoSphere Warehouse platform support, licensing, and pricing

	Departmental	Advanced Departmental	Enterprise	Advanced Enterprise	Developer
Maximum DB2 server instance memory	64 GB	64 GB	Unrestricted	Unrestricted	Unrestricted
Maximum processor cores used by DB2 server	16	16	Unrestricted	Unrestricted	Unrestricted
Operating System	Windows, UNIX, Linux	Windows, UNIX, Linux	Windows, UNIX, Linux	Windows, UNIX, Linux	Windows, UNIX, Linux
Terabyte Pricing	Maximum 15 TB	Maximum 15 TB	Yes	Yes	N/A
Processor Value Unit (PVU)	N/A	N/A	Yes	Yes	N/A
Limited-Use Socket	Maximum 4 sockets	Maximum 4 sockets	N/A	N/A	N/A
Authorized Users	Minimum 30 users per socket	Minimum 30 users per socket	N/A	N/A	Limited for use only in a non-productio n environment

To evaluate which InfoSphere Warehouse edition you need, refer to the latest updates at the following website:

http://www.ibm.com/software/data/infosphere/warehouse/edition-comparison.html#compatibility-pricing

1.4.6 Subcapacity licensing options for virtualization

Because virtualization allows you to have smaller logical servers, IBM clients can acquire subcapacity licensing for software that only uses a subset of the resources from one physical server. In this case, virtualization impacts only the PVU-based offering where clients acquire licenses for all the physical processor cores in the server.

Subcapacity licensing permits the software license to be lowered to match the virtualization capacity of the VMs that restrict the number of processors you can use, as shown in Figure 1-16. This flexible licensing option allows you to more easily embrace the virtualization technologies and take advantage of the benefit of lowering the overall cost of ownership (TCO).

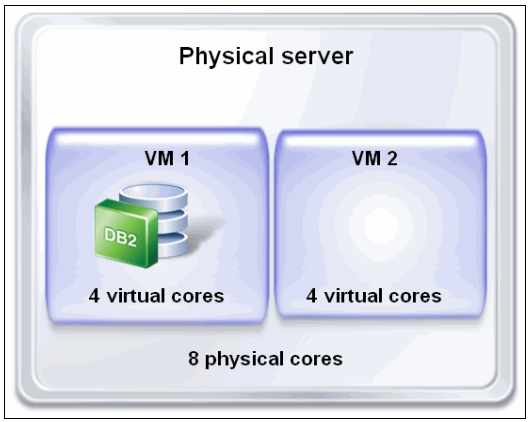


Figure 1-16 DB2 licensed with four virtual cores

The subcapacity license also applies to a cluster of physical servers configured to act as one virtual environment, as shown in Figure 1-17.

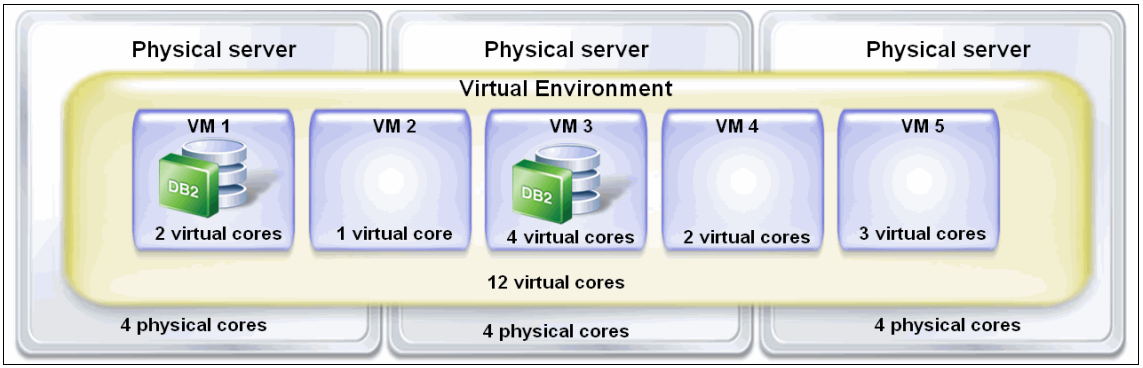


Figure 1-17 DB2 licensed with six virtual cores

Subcapacity licensing applies only when using selected virtualization technologies, eligible subcapacity products, and eligible processor technologies. Since the subcapacity announcement in April 2009, nearly all IBM PVU-based products are eligible for subcapacity licensing, although there are some exceptions for products that do not run in virtual machines.

There have been some enhancements in the license ordering and tracking process, in which the full capacity part numbers can be used for full capacity or subcapacity modes. For additional information, review the latest updates at these subcapacity licensing websites:

- http://www.ibm.com/software/lotus/passportadvantage/subcaplicensing.html
- http://www.ibm.com/software/support/virtualization_policy.html

To learn more about the virtualized environments that support a specific IBM product, refer to the software product compatibility report at:

http://publib.boulder.ibm.com/infocenter/prodguid/v1r0/clarity/index.jsp

1.4.7 IBM SmartCloud for System x and POWER platforms (private cloud)

Finally, you want to have an affordable hardware system that allows you to deploy virtual machines easily. IBM SmartCloud Entry for System x and IBM POWER® Systems are entry-level offerings for deploying private clouds designed for mid-sized businesses. These systems provide self-service workload provisioning and virtualized environment management functions such as provisioning, metering, and reporting on IBM servers and storage systems.

The cloud software stack easily integrates with the System x and POWER platforms and optimizes the data center with increased automation, structured security, comprehensive management, and resource sharing, as illustrated in Figure 1-18 on page 46. The use of these solutions greatly accelerates the deployment of any virtualized asset for business intelligence private clouds.

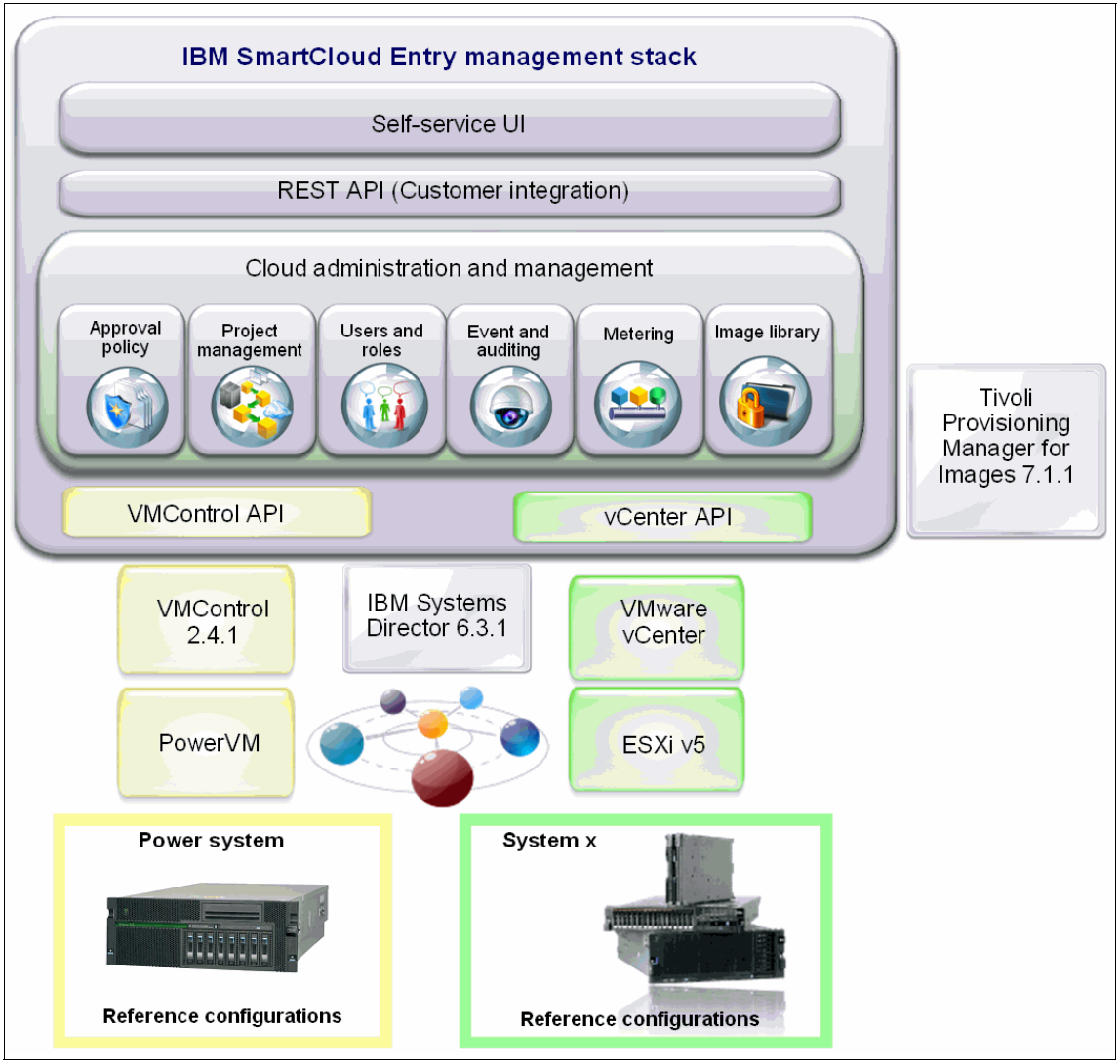


Figure 1-18 IBM SmartCloud for System x and POWER systems

1.4.8 IBM SmartCloud Enterprise (public cloud)

Another option for accelerating your development and testing of BI applications is to take advantage of an enterprise-class public cloud infrastructure such as IBM SmartCloud Enterprise. It is well suited for on-demand access to virtual servers and competitive PaaS offerings. You can start running your software stacks right away and stop paying when you stop using them. IBM Smart Cloud features

guide you through the process of designating servers and triggering the automated process that builds them for you, as shown in Figure 1-19.

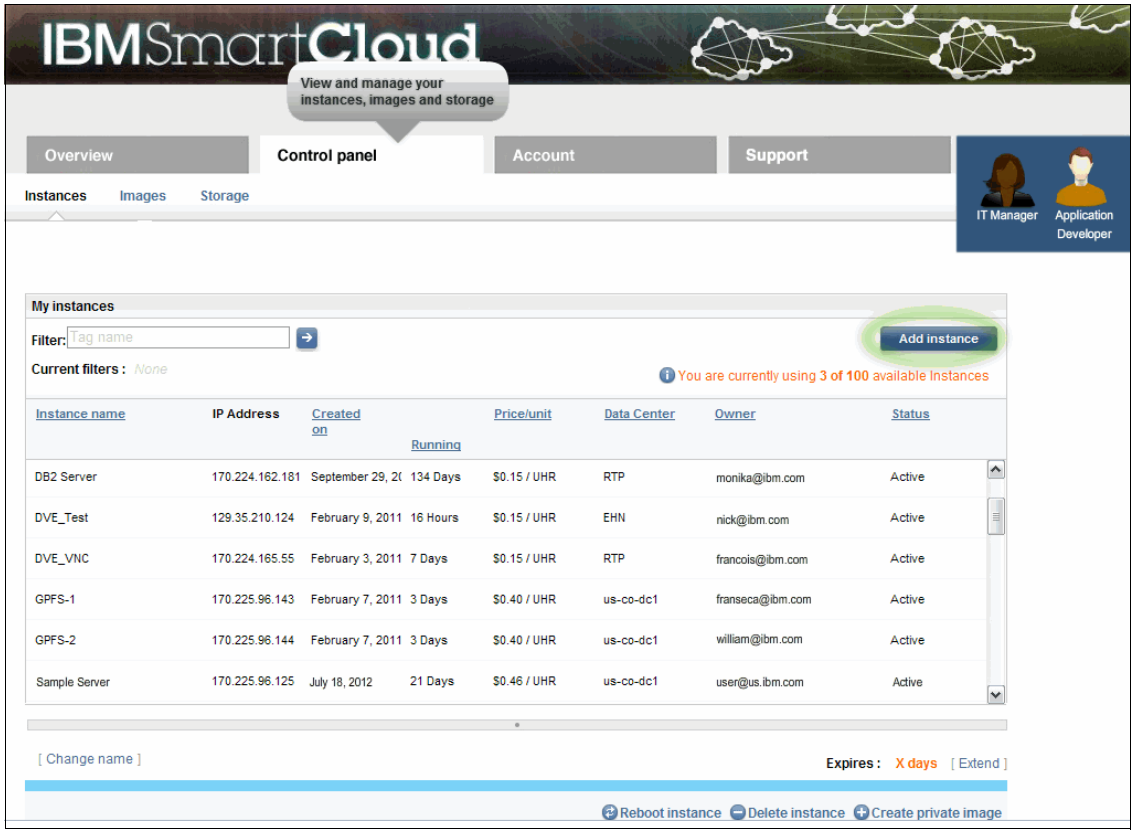


Figure 1-19 IBM SmartCloud Enterprise self service portal

1.4.9 IBM Cloud Computing Reference Architecture

IBM Cloud Computing Reference Architecture (CC RA) is an IBM blueprint that provides a set of preferred practices and methodologies for delivering cloud services and enabling management across IaaS, SaaS, and PaaS solutions. The examples in this book use the architecture principles as defined by CC RA to architect and design a virtualized business intelligence infrastructure for cloud using InfoSphere Warehouse 10.1 and other IBM cloud enablement technologies.

The following elements define a high-level architecture of a cloud environment using the CCRA architecture:

- Cloud service consumer

  A cloud service consumer is an individual or organization that subscribes to a cloud membership, which allows it to consume particular services from a service offering catalog. Technical and administration roles can be enabled to integrate existing in-house IT systems with cloud services for better interoperability and usage, as shown in Figure 1-20.

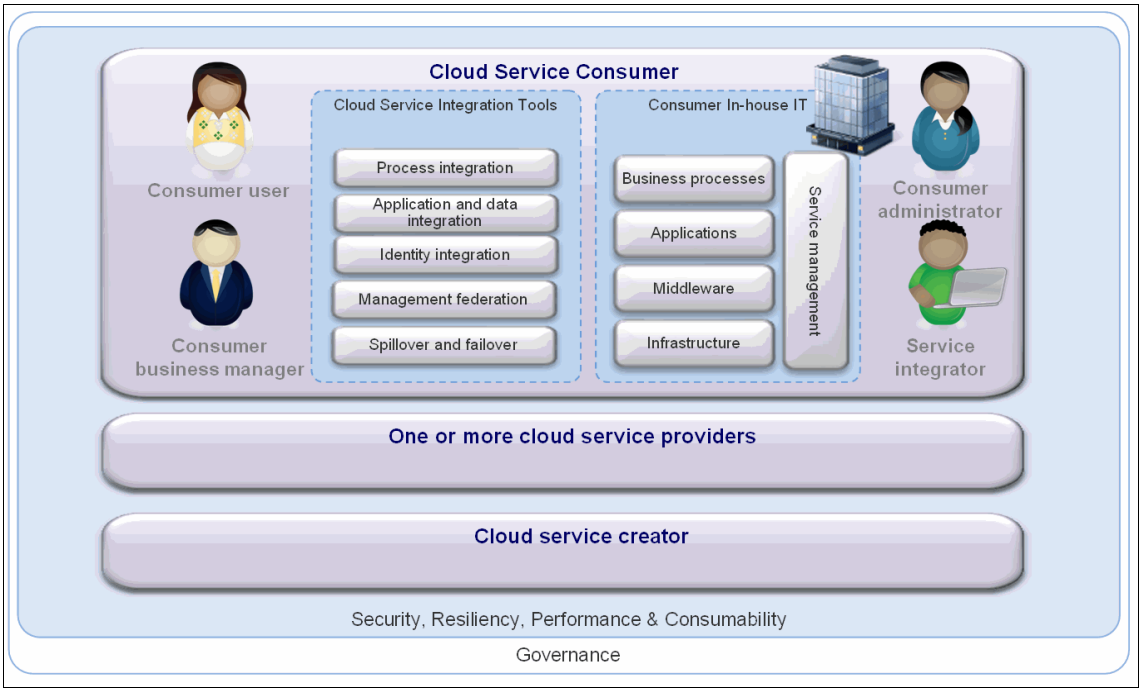


Figure 1-20 IBM Cloud Computing Reference Architecture for the cloud service consumer

- Cloud service provider

  The cloud service provider is an organization or many organizations that partner together to enable the physical infrastructure, but also manage it through a common cloud central management platform (CCMP) for operational and business support. This CCMP delivers cloud management services to sustain the IaaS, PaaS, and SaaS self-service instances and handles all business-related services, such as metering, billing, and pricing, as shown in Figure 1-21 on page 49.

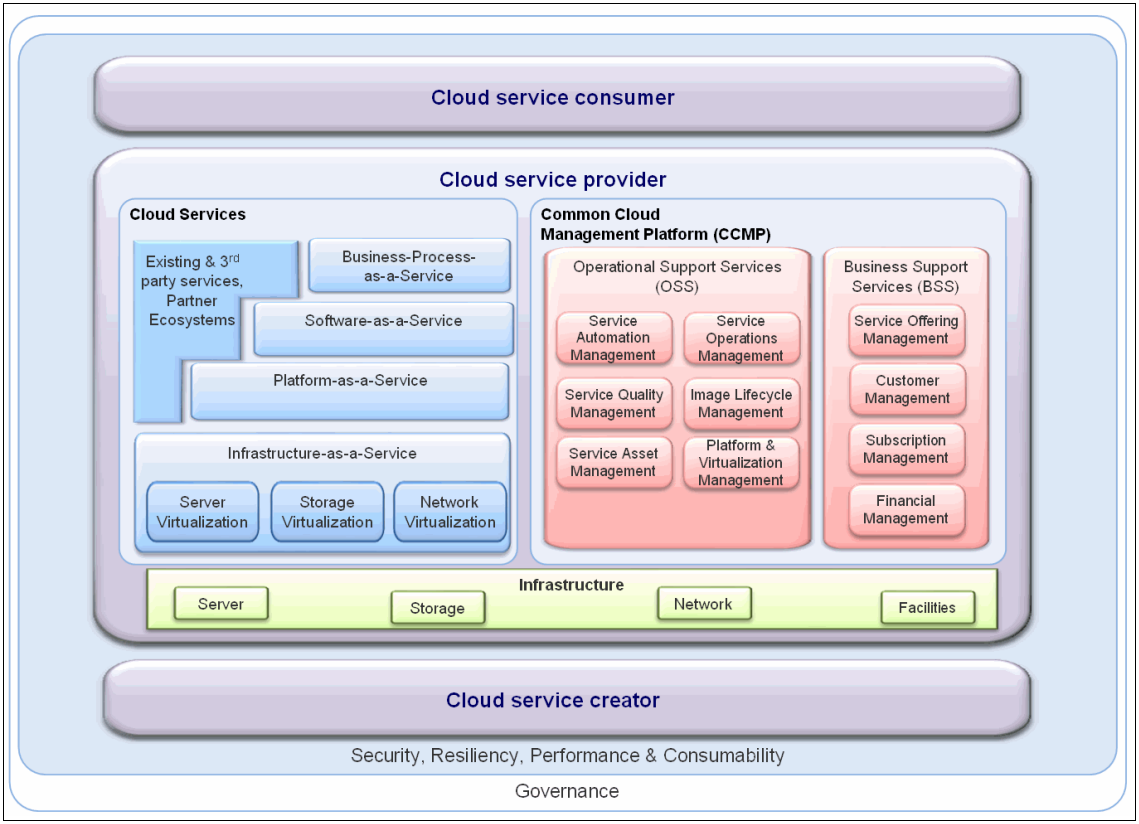


Figure 1-21 IBM Cloud Computing Reference Architecture for the cloud service provider

- Cloud service creator

  The *cloud service creator* is an individual or organization that is responsible for delivering services content. The cloud service creator continuously improves the quality and efficiency of the services. Services can be enhanced by staff from the cloud service provider or can be acquired by an ISV company. Cloud service creators use development tools to create, design, and develop different runtime and management artifacts that are particular to a cloud service, as illustrated in Figure 1-22 on page 50.

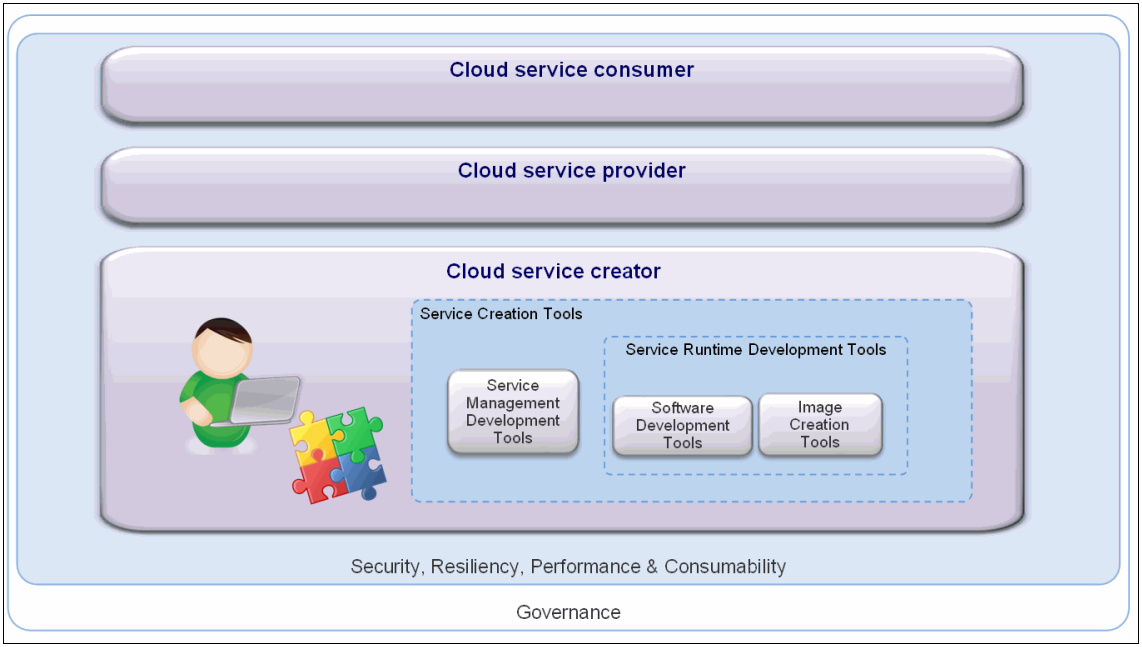


Figure 1-22 IBM Cloud Computing Reference Architecture for the cloud service creator

CC RA also highlights important elements that need to be understood and addressed to support the overall cloud environment, such as security, resiliency, performance, consumability, and governance. For a detailed description of the CC RA, see:

http://public.dhe.ibm.com/common/ssi/ecm/en/ciw03078usen/CIW03078USEN.PDF

1.5 Scope and business objectives

The remaining chapters of this book provide a high- to mid-level explanation of the deployment of a BI cloud solution. They include details such as the architecture of a cloud, planning implementation, integration of various software components (such as Infosphere Warehouse and Cognos), and understanding the leading practices involved in running a cloud deployment. The book is intended as a guide to deploying an IBM middleware-based BI cloud stack.

Both BI and cloud computing as individual technologies are massive and ever-evolving. By combining the two, many possibilities emerge with respect to

the types of services that can be realized and delivered. As such, deciding your target market can sometimes be a daunting task.

Consider the following checklist before designing or planning the implementation:

- How large are the target customers? If they are large companies, expect to deliver enterprise-ready software and services. If they are small companies, entry-evel, production-ready offerings might be sufficient. What about development offerings?
- How tech-savvy are the target customers? The more features your solution offers, the higher the complexity. More features are great for customers that are well versed with technology but might disappoint less tech-focused companies.
- Will the customers try to build on top of your offering or use it for in-house purposes? This type of use can change the way that you design the solution.

2

Planning the implementation and architecture of a virtualized BI cloud

This chapter provides a high-level, step-by-step guide for planning the deployment of a virtualized business intelligence (BI) cloud. It includes the following topics:

- Planning implementation
- Cloud architecture
- The role of virtualization and preferred practices
- Virtualized business intelligence architecture

2.1 Planning implementation

Planning the implementation of a virtualized cloud is a multi-step process that must not be overlooked. Careful planning will ensure that your project meets the established requirements and business objectives in a predictable and timely manner. High-quality implementations are almost always the result of accurately defining the scope and a timeline of the project before it begins.

This section describes how to define the scope of the project and introduces the four stages of implementation:

- Design
- Development
- Testing
- Deployment

2.1.1 Defining the project scope

The scope of an IT solution defines the technical goals, target audience, and business objectives of the solution, and lays the groundwork for building it.

Consider these specific IT issues and how addressing them can help define the scope of a cloud solution:

- How sensitive is your customer data?

  Data sensitivity is often a top priority for customers. Customers implementing cloud services are forced to have their data outsourced, therefore giving up control. So the data must be monitored to ensure it is protected.

  Securing sensitive data can be thought of as a two part operation. First, customer data needs to be protected from entities *outside* the cloud. One customer using a cloud solution must not have access to, or even be aware of, the data of other customers using the solution. Second, data must be protected from *inside* the cloud: The database and cloud administrators running the cloud service must not have access to the customer data either. This second part is often overlooked.

  Implementing a proper multi-tenant architecture is the preferred practice for isolating sensitive data. By designing a proper multi-tenant environment, the residing customer data can be protected against unauthorized internal and external access.

  Virtualized and cloud environments are highly dynamic, allowing for rapid provisioning and de-provisioning of server instances. This, in turn, can lead to escalating hacking threats if the environment is not properly secured. Therefore, a cloud solution must ensure that every server instance that is

provisioned is a self-contained environment capable of locking down access to rightful users.

Most private, in-house solutions implement a network-sniffing security mechanism to filter out possible security threats. This model works especially well within an on-premise, intranet-based network. However, a public or even a hybrid cloud can change the nature of your solution. What additional security measures must be implemented to provide secure access to server instances for rightful users? Options such as secure browsing using private and public key-based credentials might need to be implemented.

To ensure the highest degree of security within the cloud, users that have been granted indirect access to the data must not be able to gain direct access to it. For example, database administrators must be able to perform database maintenance activities but not access sensitive data on those databases. Software products such as IBM Infosphere Warehouse allow advanced authorities and privileges to be granted to specific users to separate these types of duties from actual data access.

Finally, consider the geography where your solution will be deployed. Certain countries have more stringent requirements for sensitive customer data such as financial and health records. Is it acceptable for a third-party vendor to have indirect access to this information?

With the information gathered through this type of analysis, the common practice is to utilize private cloud services for solutions that store and process sensitive data, because private cloud implementations offer the highest degree of control.

- How much are you willing to invest in the solution?

  Total cost of ownership is a consideration in any IT solution, and even more so for cloud services due to their ever-changing workloads. How much money must be spent up front versus how much money must be spent for ongoing operations?

  The cost of a BI cloud solution can be divided into two categories:

  – Software and services
  – Infrastructure

  The first question that needs to be asked is: What service are you offering? The second question is: What supporting software and features are required to provide this service?

  A BI cloud solution consists of many different software products, not just BI software but also data warehousing software, virtualization software, operating system software, infrastructure management software, and other support software and services. Each component product has editions, versions, and features to select, which also affect the cost of the solution. What features are needed or beneficial, versus what are the consequences of

not having those features? For example, not opting for high availability features and implementing a manual disaster recovery solution can save money, but it can also reduce the robustness for your service level agreement. Another issue: How much money must be allocated for IT personnel to maintain the solution?

Each cloud computing delivery model must be evaluated with respect to how much money will need to be spent. One of the driving decision-makers for choosing a public cloud infrastructure is the initial cost savings. Consider the advantages and disadvantages of the main cloud computing delivery models.

- Private cloud: Pay more now and less later. This model allows for complete ownership of a solution, allowing services to be added at will with minimal additional cost.
- Public cloud: Pay less now and more later. This model allows for less initial capital expenditure because the hardware is not purchased outright. Workloads are processed using virtualized hardware resources that are, effectively, rented.
- Hybrid cloud: With this mixture of public and private clouds, costs remain relatively steady throughout the solution lifecycle. Base-level workloads are processed as usual, with additional resources rented as needed for larger workloads.

- What type of workloads will be processed?

  Consider the workloads the BI cloud solution will process. If they are online transactional processing (OLTP) workloads, a database solution can be the best fit. Online analytical processing (OLAP) workloads, in contrast, can benefit from a data warehousing solution. Also investigate the tools that will be needed when analyzing the data. Does the solution require analysis using data mining or cubing services? Will large amounts of data be stored for long periods of time and require database partitioning for scale-out?

  In-house, private cloud hardware is best utilized when the workload demand is equal to its capacity. Public cloud hardware is best suited for changing and inconsistent workloads. Consider utilizing both private and public cloud resources to minimize costs and maximize efficiency.

  Following are some general guidelines for which workloads are best-suited to each type of cloud architecture:

  - Workloads are less than your resources can handle: Private cloud
  - Workloads are more than your resources can handle: Public cloud
  - Consistent workloads: Private cloud
  - Variable workloads: Public cloud

- What level of security is required?

  The issue of security often is tied to protecting sensitive customer data, although there are subtle differences. Although data sensitivity has to do with isolation and preventing non-malignant, unintentional access threats, security deals with malignant and unauthorized access to data.

  How well is your solution able to secure its data? More specifically, what hardware, software, and features will be implemented to make your cloud solution secure? Will the cloud solution incorporate a hardware or software firewall? Is a virtual private network (VPN) required at a per-tenant level? And are these services that your BI cloud can provide, on-premise, or will they need to be outsourced to a third party?

  For example, although a private cloud offering can initially seem to be adequately secure because its data is inaccessible by third parties, private cloud deployments often lack the tools and resources necessary for high-level security. Public clouds or companies that specialize in offering third-party hosting resources usually have additional services such as firewalls or VPNs that can supplement an existing cloud solution. Furthermore, these services are managed for you, which can reduce cost and complexity, but allows for less customization and carries higher risk because the services and resources are not under your control.

- How reliable must be the cloud services you offer?

  Customers using BI reporting services will demand that the services they pay for be highly reliable, even as reliable as their own, in-house-provided services.

  Examine every component of the cloud solution, from hardware to software. In order for services to be deemed reliable, each component must be redundant and available. Although it is generally accepted that the greatest reliability comes from a properly implemented, in-house solution that is highly available, this cannot be achieved without a high cost, so consider using services provided by a third-party public cloud infrastructure. Most public cloud vendors already have pre-existing service level agreements (SLA) in place that your cloud solution can inherit.

- What type of cloud service is the solution providing?

  BI customers are looking only to consume the software or service that is being provided. They are not concerned with the underlying hardware, tools, or supporting software components within the solution. They only want access to reporting services, so services must be software-as-a-service (SaaS)-focused.

- What, if any, cloud services are required to deliver the solution?

  To effectively deliver a SaaS solution, use the following items:

  - Hardware (either virtualized or physical, externally or internally hosted), or infrastructure-as-a-service (IaaS)
  - Development tools such as an operating system, database, business analytics, and automation software, or platform-as-a-service (PaaS)

- How do you expect demand to grow in the future?

  Demand for services drives the need for developing new solutions. How do you expect your solution to grow? Will the infrastructure you have in place today be able to accommodate processing demands years down the road?

- What kind of pricing model is being offered?

  Fixed term or flexible pricing? This decision is related to the choice of monitoring and metering tools that your solution will utilize and the level of flexibility it will offer.

- How will the services be maintained?

  Maintenance, like other aspects of cloud computing, can be divided into two parts: software and hardware. In-house software services must be maintained and monitored to ensure they are running properly. IT specialists such as database administrators must be on site to do this. The hardware infrastructure must also be maintained and checked periodically. On-premise, in-house hardware will need to be checked more frequently than public cloud infrastructure.

  Periodic maintenance is required to keep customer-provided services running and problem-free. Consider these questions when planning for maintenance:

  - When is the best time to schedule maintenance?
  - Which maintenance tasks must be performed during down time?
  - How will down time affect the service level agreement?

Given the issues and suggestions listed here, you can now formulate an implementation plan. Before moving forward to the development stage, ensure that you have clearly defined what your solution will deliver and what will be required to create it.

2.1.2 Implementing the solution

Although cloud and virtualization technologies are a solid part of today's IT industry, they are constantly evolving. Do not underestimate the time required to architect or design a cloud solution.

The application development lifecycle in a cloud environment is not much different than in a traditional IT environment. The amount of development will largely depend on whether services are being built from scratch or provided by a third-party vendor. Generally, leveraging pre-built solutions is more efficient than building from scratch and frees up time for you to focus on gathering requirements, but the trade-off is reduced customizability.

An *application* is a collection of programs that satisfies certain specific requirements (resolves certain problems). The *solution* can reside on any platform or combination of platforms from a hardware or operating system point of view.

The project timeline for a BI cloud solution includes the following major steps (as illustrated in Figure 2-1):

- Gather requirements
- Perform analysis
- Design the solution (high-level and detailed views)
- Develop and test the solution
- Perform user and system tests
- Go to production
- Maintain the solution

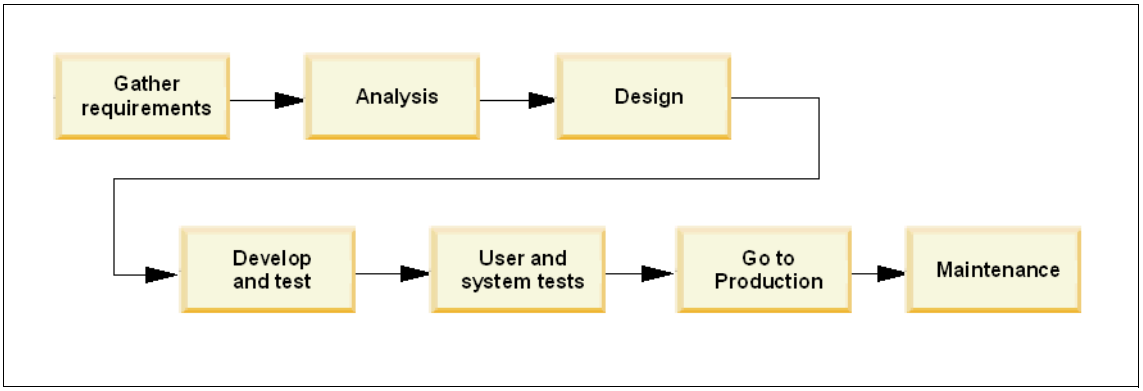


Figure 2-1 Application development lifecycle

Four of the steps in the development timeline deserve additional explanation.

Design

Solution design is a process of conceptualization, verification, and revision, as shown in Figure 2-2.

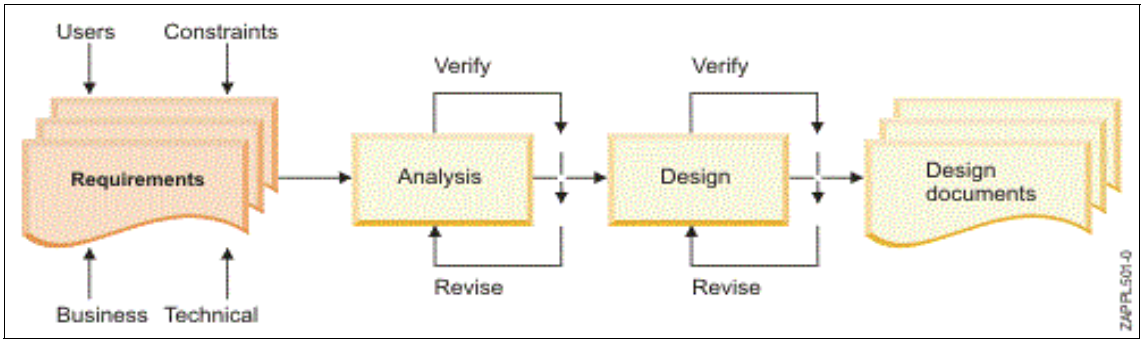


Figure 2-2 Application design flow

This phase consists of determining what the BI cloud solution must accomplish and then deciding what software, services and hardware will be required. A set amount of time must be allocated for design because research will be needed to decide which components are compatible and best suited to power the solution.

After a final design has been adopted, the programming requirements are passed on to the development team.

Development

The development phase consists of integrating the ideas and components selected in the design phase. The developers take the programming requirements, usually in the form of design documents, and proceed with the iterative process of coding, testing, revising, and testing again, as shown in Figure 2-3.

Development is usually the longest and most time-variable implementation phase because getting particular components to work together for the first time can be quite difficult. Consider the various areas of expertise that will be required to build the BI cloud solution:

- Warehousing
- Business intelligence
- Availability, security, reliability, and scalability
- Virtualization
- Graphical interfaces and APIs
- Optimization and tuning

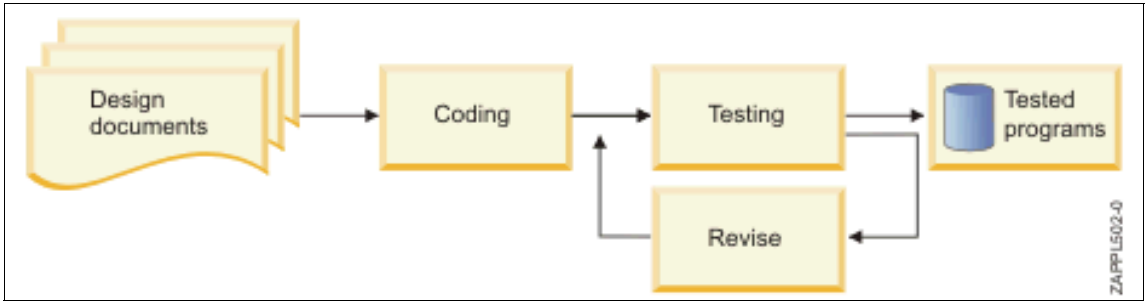


Figure 2-3 Solution development

Testing

During the testing phase, test cases are created and run to determine if the BI cloud solution performs at a satisfactory level. After the programs have been tested by the developers, they undergo a series of formal user and system tests, as shown in Figure 2-4. These tests verify usability and functionality from the user's point of view and confirm the functions of the separate applications within a larger framework.

The testing phase is usually the easiest to plan and estimate a timeline, provided that sufficient preliminary testing was performed in the development phase. As with many software projects, design, development, and testing are cyclical, so time must be allocated for this as necessary.

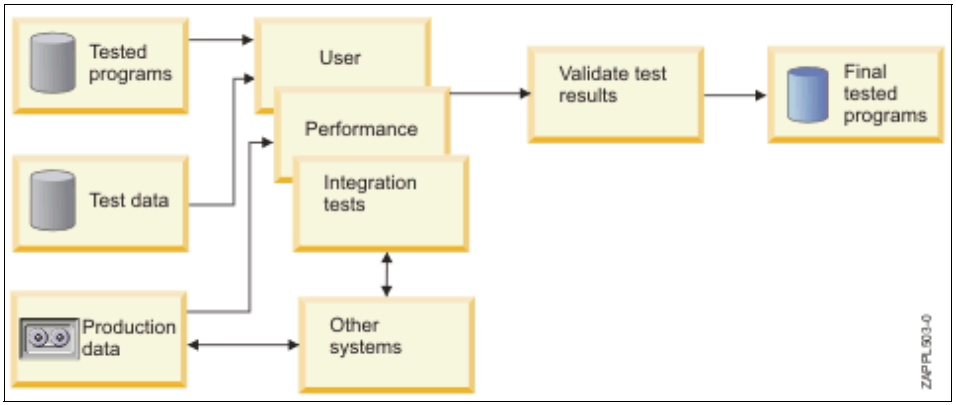


Figure 2-4 Solution testing

Production

Figure 2-5 shows the final phase in the development lifecycle, which is to promote the solution to production in a steady state. As a prerequisite to going to production, the development team needs to provide documentation, typically

user training and operational procedures. The user training familiarizes users with the new application. The operational procedures enable operations personnel to run the solution on an ongoing basis.

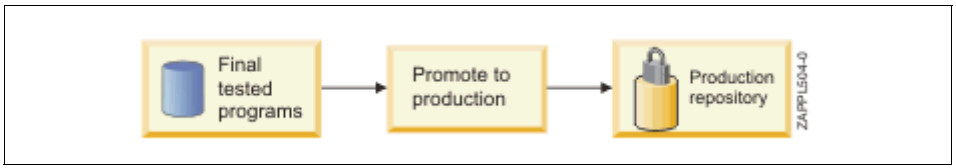


Figure 2-5 Moving a solution to production

In production, changes and enhancements to the solution are handled by a group (possibly the original developers) that performs ongoing maintenance. At this point in the lifecycle of the solution, changes are tightly controlled and must be rigorously tested before being put into production.

Do not set a finite amount of time for the deployment or move-to-production phase, as it is an on-going process. For deployment, make sure to schedule operations and tasks such as routine backup, maintenance, cleanup, and so on.

2.2 Cloud architecture

By now, the business benefits gained by adopting cloud technologies should be clear, as is the process of implementing a basic cloud solution. But to fully understand the potential technical benefits of the cloud approach, the architecture of a cloud must be examined in detail. Knowing the different types of cloud deployments will help you select the correct one for your enterprise.

This section describes the architecture and components of each type of cloud:

- Public cloud
- Private cloud
- Hybrid cloud

2.2.1 Public cloud

Public cloud infrastructure uses many different IT resources. These resources are often delivered by third-party vendors over virtualized platforms, meaning specific network infrastructure, hard disk, memory, and other requirements with respect to manufacturers and peripherals cannot be customized.

Figure 2-6 shows an example of a public cloud architecture.

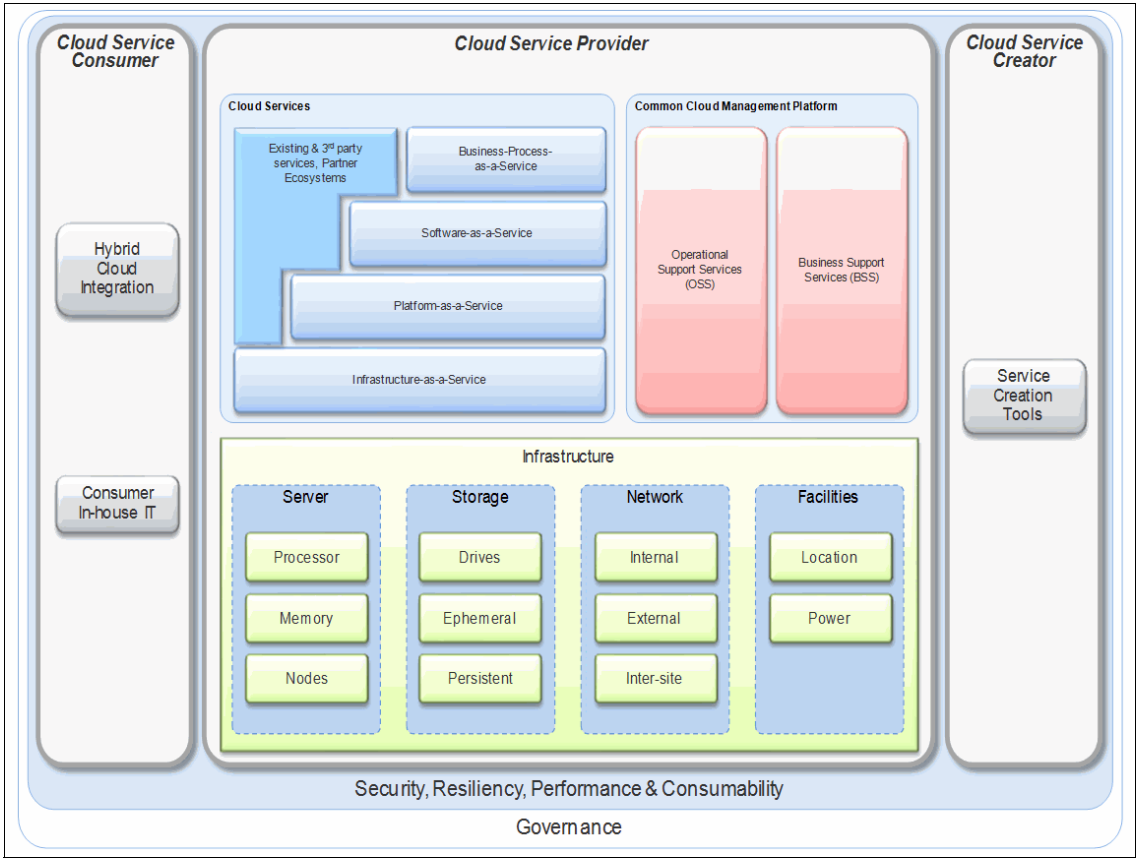


Figure 2-6 Sample cloud infrastructure

A public cloud includes the following typical components:

- Server instances

  Stand-alone virtualized server instances that provide a bare-bones computing solution. Usually, server instances come with different computing capacities to suit differing technical needs, allowing for customization of CPU, memory, and storage.

  Server instances come with a pre-installed operating system, which is usually also customizable. Server instances can be started, stopped, and discarded as necessary to maximize resource efficiency and save on costs. Multiple server instances can be launched within the same zone to reduce network latency between co-located servers.

- Storage

  Although base storage is usually provided in a cloud's server instance, there can be times when additional storage is needed. This additional storage, unlike the base server instance storage, is usually customizable in terms of size and performance. It is advantageous in other ways, too, because server instances are often seen as volatile (non-resilient), which makes data on auxiliary storage more persistent. Auxilary storage is often replicated across several data centers in different locations, which provides data redundancy and makes auxiliary storage ideal for storing backups and for keeping costs low when server instances are shut down or discarded.

- Networking

  Public cloud vendors often have multiple data centers in the same region that are interconnected by a high-speed LAN. However, to achieve redundancy, there can be multiple regions in the same zone, and therefore latency between regions can be slightly more significant than it is within a single region. Vendors often provide premium network hardware options for premium cloud server offerings.

- Other infrastructure resources

  Public cloud vendors usually offer a number of other resources to supplement the infrastructure backbone, such as services that improve security, resilience, monitoring, load-balancing, and so on. Examples of such resources and services include cloud VPNs, firewalls, and dedicated network options such as static IP addresses.

- Platform-as-a-service

  Many public cloud vendors offer a variety of different platform-as-a-service (PaaS) options, ranging from middleware software and databases to complete development environments, application lifecycle management, and web servers.

- Cloud API and Web UI

  A web console is commonly employed to facilitate fundamental operations within a cloud computing environment. The web console serves as a central hub for managing and monitoring the cloud resources. To automate and templatize the launching of resources and services, a cloud API can be used. Note that because a public cloud is still a maturing technology, standardized APIs do not yet exist among the different public cloud vendors. Therefore, care must be taken in deciding which public cloud vendor can best fulfill your cloud solution requirements.

At a minimum, an effective UI console and API performs these tasks:

- Provisioning and management of server instances and other resources
- Client billing and metering
- Image cataloging
- Cloud infrastructure monitoring
- Account management

Figure 2-7 shows these and other important features of an effective cloud computing solution.

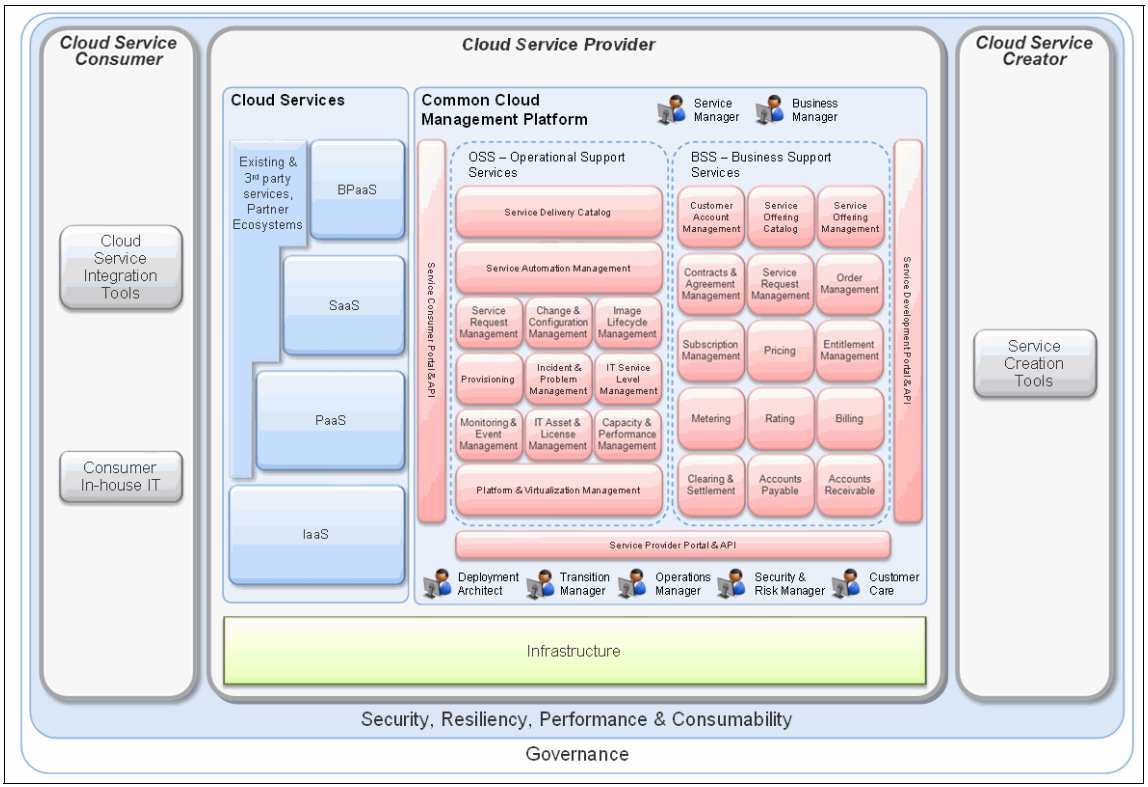


Figure 2-7 Features of an effective cloud solution

IBM middleware is cloud-ready on multiple public cloud platforms that are available today. Pre-installed images for some of today's most popular cloud platforms can be integrated into an existing solution, thereby reducing time-to-value. Here is a list of cloud vendors that host IBM middleware:

- IBM SmartCloud Enterprise
  - Enterprise-focused
  - Highest degree of support and integration with IBM software

– Largest pre-existing library of IBM software cloud images
– Data centers in North America, Europe, and Asia
– Instance sizing to suit any hardware requirement

For clients interested in private cloud deployment, IBM also offers SmartCloud Enterprise+, an IBM-managed IaaS cloud.

IBM SmartCloud services can be divided into application services, enterprise services, and enterprise+ services, as shown in Figure 2-8.

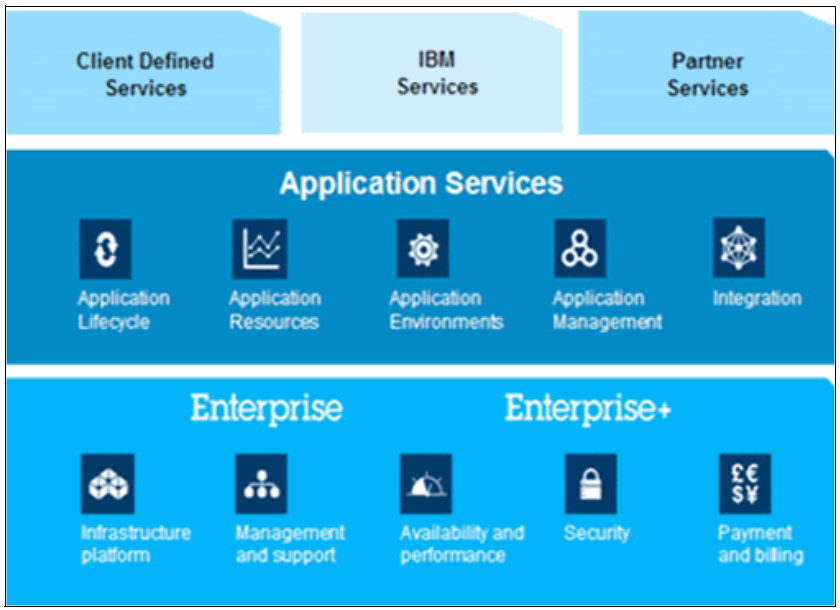


Figure 2-8 IBM SmartCloud infrastructure and platform services

- Amazon Web Services Elastic Compute Cloud
  – Pre-bundled IBM software cloud images available since February 2009
  – Flexible payment options including pay-as-you-go by-the-hour billing, or you can use your own IBM middleware license
  – Most widely used and mature public cloud platform available today
  – Offers the most advanced technical features including VPNs, segregated data zones and centers for redundancy, static IP addressing, load-balancing and storage
  – Highest degree of geographical accessibility with data centers spread across the globe

- GoGrid cloud
  - ISV-focused, with easy to build on pre-built images to speed new offerings to market; GoGrid handles the billing
  - Dynamically resizable instance capacity
  - Hardware VLAN for network segregation, dedicated firewall options, available VPN, and user-friendly API
  - Offers both by-the-hour and subscription-based servers, plus hybrid solutions
  - Large library of available software images
  - Data centers in North America and Europe
  - GoGrid Exchange area to showcase your offerings

All of the public cloud vendors listed here offer many options and configurations for hardware to meet a wide range of requirements for CPU, storage, networking, security, and accessibility.

2.2.2 Private cloud

A particular advantage of implementing a private cloud solution is the ability to choose and customize any component within the solution. Components suitable for customization include the following:

- Hardware

  Choosing the correct middleware components for a BI cloud solution is imperative for a successful deployment. Is the identified middleware cloud-ready? Do the middleware components support the requirements for scalability, virtualization, varying workload sizes, security, and performance?

- Support software

  These components glue together the various pieces of your BI cloud solution. For example, consider what operating system to use and its support for various software packages. What software and services are required to maintain things such as time, availability, web server, LDAP, and DNS services.

- Client software

  This software or service is the client's specialization. It can be anything from optimizing performance to adding new capabilities to existing middleware components. This is the added value that customers ultimately seek in a truly innovative BI cloud solution. Consider how this software is accessed today. Is it already software-as-a-service ready, or does it require additional development to make it cloud service-compliant?

- Virtualization technology

  What are the requirements for virtualization? Virtualization enables multiple working environments to be created from a single physical computing environment. Good virtualization software provides a cloud solution with abilities such as multi-tenancy, scalability, isolation, high availability, monitoring, improved efficiency, and more. Consider which middleware components and services require virtualization. Some of them are obvious candidates to be virtualized but others require too high a level of performance to perform virtually. For example, if specific software requires high network connectivity that virtualized hardware cannot provide, then it might need to be hosted by physical infrastructure instead. There are many virtualization options available today, including VMware, KVM, and Xen virtualization, each with their own unique features and advantages.

2.2.3 Hybrid cloud

A *hybrid cloud* is a combination of at least one private cloud and at least one public cloud. The result is an environment that provides transparent user access that is capable of dynamic scalability to manage varied workloads.

2.3 The role of virtualization and preferred practices

Although it is possible to deliver a cloud solution without using virtualization, doing so can compromise important cloud features such as rapid provisioning, elasticity, and self-service.

Expertise in virtualization is particularly important for deploying private cloud solutions because it is key to fulfilling the promise of provisioning, elastic scaling, and automation. In contrast, solutions that use a public cloud already have virtualization built in, so they do not require as much expertise in that field.

All of the deployment examples and suggestions in this book assume that virtualization is being used, with no emphasis on any specific virtualization technology. It is also assumed that Type 1 virtualization (which is explained next) will be used with respect to private cloud deployments.

A hypervisor, also known as a *virtual memory monitor* (VMM), is a virtualization technology that controls the resources between the physical hardware and virtual machines. The hypervisor manages the machines, each of which has its own operating system environment, so the machines can run concurrently but be isolated from each other.

The following types of hypervisors are available:

- Type 1 runs directly on top of the hardware, as shown in Figure 2-9. This hypervisor model allows for the best performance on virtualized hardware.

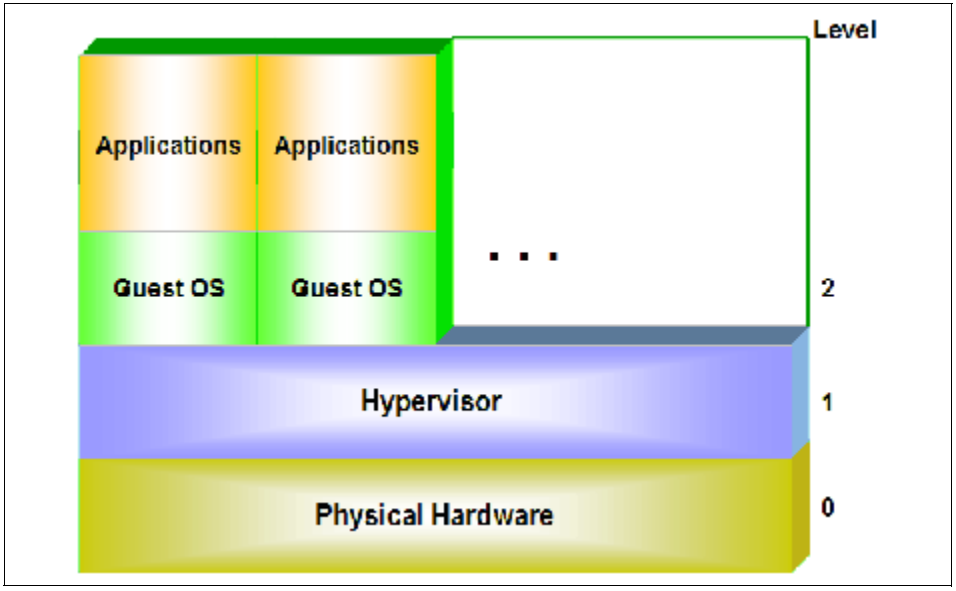


Figure 2-9 Type 1 hypervisor

- Type 2 runs on a host operating system that sits on top of the hardware, as shown in Figure 2-10 on page 70. This hypervisor model allows for flexibility because it sits on top of an OS, but virtualization performance is not as good as with Type 1.

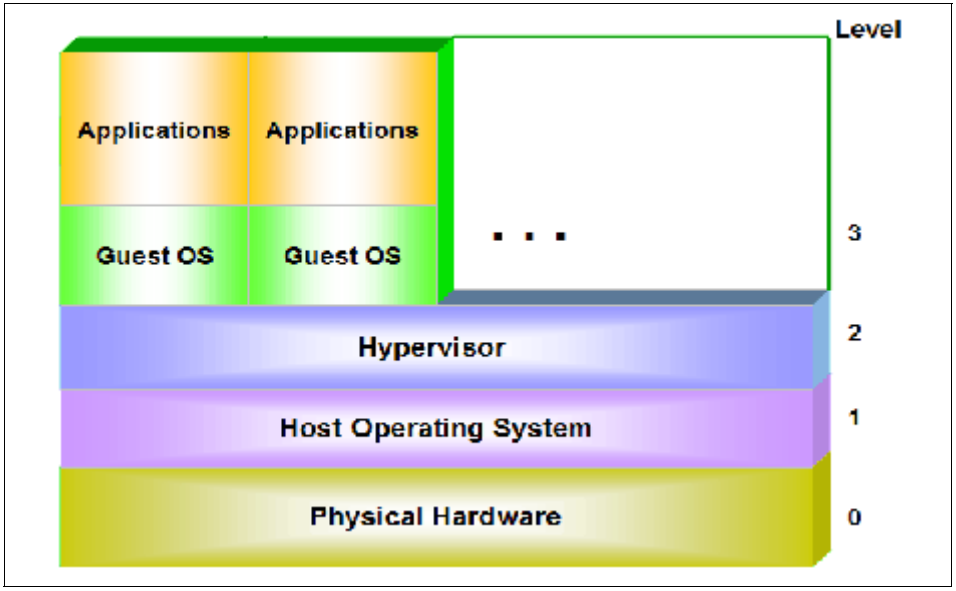


Figure 2-10 Type 2 hypervisor

Using virtualization allows you to create an ecosystem of servers to perform all the tasks necessary to deliver a true cloud solution. Everything from BI to warehousing to web services to management can be virtualized. But how do you know what software to virtualize? How do you decide what software and services can be virtualized together? How do you choose the optimal configurations for virtualized environments while preserving availability, isolation, and maintainability?

2.3.1 Virtual Machine Templating

Virtual machine (VM) templating allows for more efficient deployments. By pre-performing repetitive configuration tasks, you can create a VM template for those tedious tasks, such as installation of software and middleware. VM templates usually allow for some form of automation, such as response-file input. For example, when launching a VM template, you can submit an XML file with the required database attributes such as instance name and database name, and end up with an environment that is tuned to your liking. Figure 2-11 on page 71 shows a sample workflow for creating a VM template.

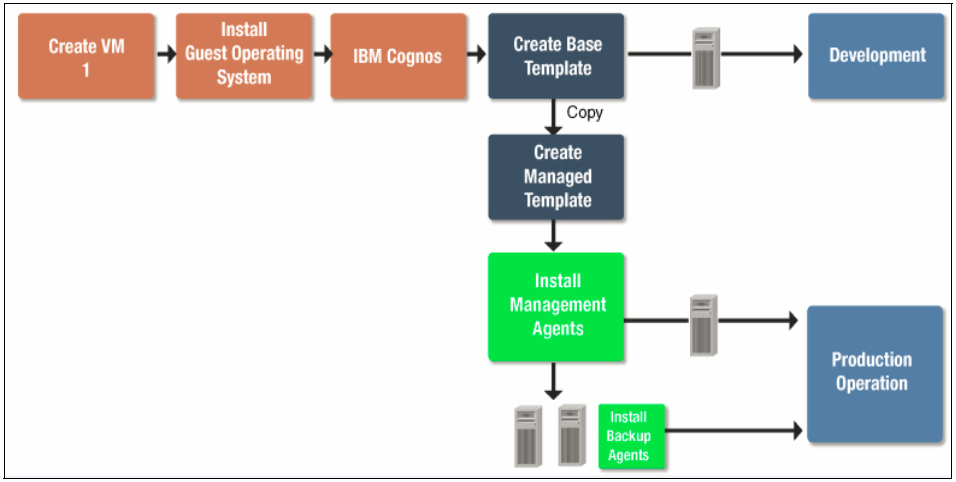


Figure 2-11 Creating a virtual machine template to automate tasks

2.3.2 Deciding what to virtualize

Before deciding to virtualize a specific software component, be sure that the software works well with virtualization. For example, some software requires specialized hardware that is not yet able to be virtualized. This can produce lower than expected performance or cause errors, so always do your research before choosing to virtualize specific software.

You must also evaluate the software you are running in a virtualized environment to confirm that it is performing to the standard required for your cloud solution. Many virtualized solutions today are only semi-virtualized. More specifically, only certain components or tiers are virtualized. For example, a solution's middleware application might be virtualized while the back-end database sits on non-virtualized infrastructure. In such situations, you might find that bare metal hardware is a better option.

2.3.3 How to virtualize

The motivation behind virtualization is to improve resource efficiency in an IT solution. However, virtualizing every piece of software so that it runs in its own virtual machine is not resource-efficient. To achieve the most efficient design, you must carefully plan the specific software and middleware components that will run together in the same VM.

Consider the following issues when deciding how to segregate and pair software together in a virtualized environment:

- Tier location

  A cloud solution is similar to a traditional IT solution in that they share similar design principles, most notably the *N*-tier or multi-tier design.

  When grouping software within the same VM, each program must reside within the same tier for optimal redundancy. Also, do not consolidate tiers through virtualization; some tiers are so complex that consolidating them can compromise maintainability and even security. For example, InfoSphere Warehouse is considered data tier software, whereas Cognos is application tier software. Therefore, these two software components must never be combined in the same virtualized environment.

- Isolation

  Does a particular software component require a high degree of isolation? If so, you must consider this when planning what to virtualize.

  To achieve proper VM isolation with respect to software, design your virtualized environment from the perspective of a recovery scenario. For example, if you have a series of components that perform integral tasks and fall under the same category of services, such as a domain name system server and a time server, it can be a good idea to group these services together. In contrast, database software that is vital to keeping your data intact must be kept within its own operating environment. In this way, if a database failure occurs, you only need to worry about recovering the database. If the database is coupled with other important software in the same VM, the database will be subject to the recovery process even though there is nothing wrong with it.

- Performance

  Virtualization allows for performance throttling on single, isolated VMs. You need to determine if the specific virtualized environment you are planning has enough performance capacity for your software, or, from the opposite perspective, if you can group additional software components in the same VM for greater resource efficiency.

2.3.4 Multi-tier architecture

The multi-tier architecture in many of today's applications can be applied to virtualized application architecture as well. Consider the typical three-tiered BI solution shown in Figure 2-12.

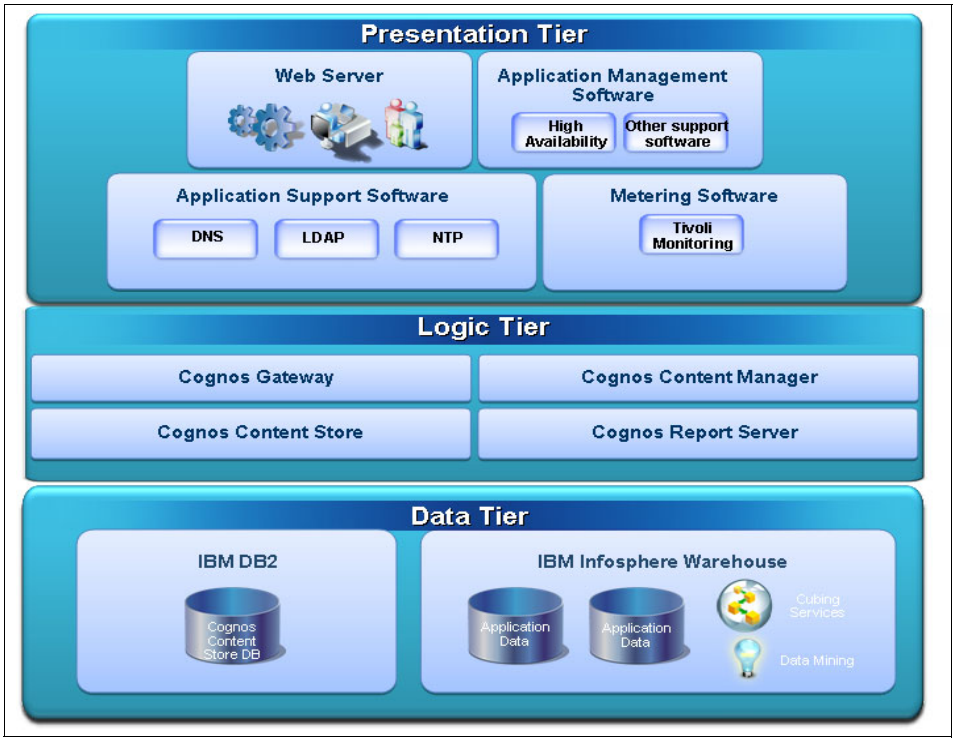


Figure 2-12 Virtualized multitier architecture for a BI solution

This approach splits the BI solution into tiers that focus on different aspects of functionality:

- Presentation

  The presentation tier gives clients access to the solution's primary services (in this case, business intelligence functions such as reporting). It contains VMs with the software for the services, which exist at the highest level of the solution stack. The services at this level can be thought of as the glue that holds together the overall solution.

- Logic tier

  Otherwise known as the *application tier*, the logic tier is the engine of the overall solution. It is where incoming data is processed into more useful forms of outgoing data, such as BI reporting, insight, and analysis.

- Data tier

  The data tier is where processed data from the logic tier is stored. Additional lower-level, complex services (for example, data mining or cubing services) can be provided in this layer by premium database or warehousing software such as Infosphere Warehouse.

This sample architecture is only a rough guideline for how to componentize software and services within a virtualized environment. You might need additional, supporting services and software to enable your solution. If so, follow these steps each time:

1. Determine the tier in which each software component belongs.
2. Decide if virtualization is advantageous for each specific component.
3. Plan and design how to virtualize it: stand-alone or shared with another VM.

2.3.5 Physical architecture

After designing your virtualized architecture around the necessary software and services, the next step is to decide how to host the virtualized environment on the physical infrastructure.

To achieve the highest degree of isolation, performance, and maintainability, each tier must exist on its own physical hardware with its own storage and servers. However, if hardware spending must be limited, it is possible for one physical server to host all three tiers. Virtualization, by its nature, already improves the modularity of any solution.

Keep these considerations in mind as you weigh the requirements and limitations of your solution to find the best balance between using one server or multiple ones.

2.3.6 Architecting for high availability

So far, the discussion of architecture has been solely from a functional standpoint, with no regard to issues of availability. Yet providing cloud services to clients requires at least some degree of high availability. A key attribute of cloud computing is that clients can access their cloud services on demand, so any cloud solution must be architected with high availability in mind.

By incorporating the recommendations in this chapter, your solution can achieve a high degree of isolation, maintainability, and redundancy. This, in turn, will make it easier to perform disaster recovery on any portion of the solution, whether it be at the software level, the VM level, or in the physical hardware.

Achieving high availability is relatively simple and requires only a mirroring of the infrastructure, virtualized environment, and underlying software components. Virtualization management software, such as VMware VSphere, can monitor physical infrastructure and virtual machine health. If the server hosting the presentation tier fails, VSphere can detect that the server is offline and switch to the tier's standby server, as shown in Figure 2-13.

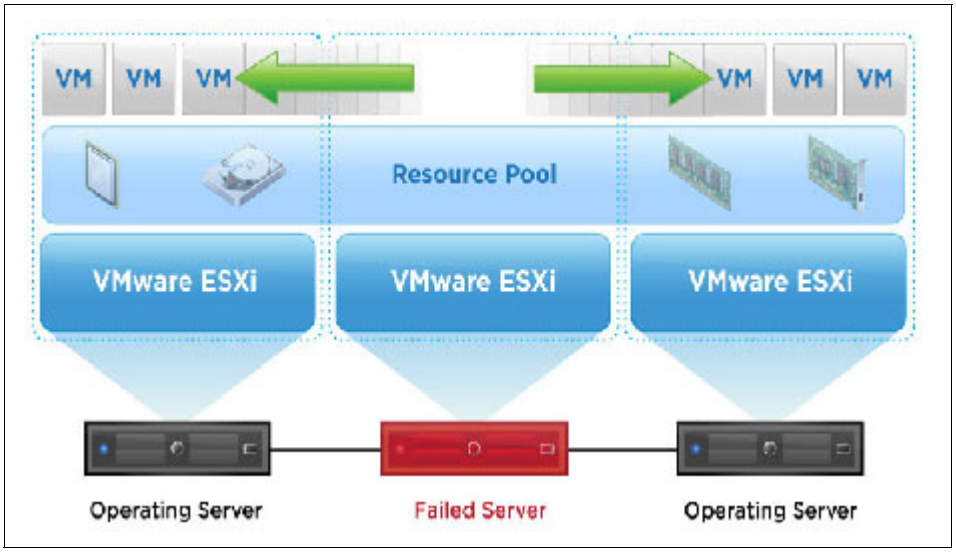


Figure 2-13 Virtual machine with VMware vSphere for high availability

In a clustered solution with VSphere enabled, both the standby and primary servers have access to the shared storage where the VM resides.

2.3.7 Architecting for future demand

In the past, vertical scaling of infrastructure was the only way to handle increased workloads. As hardware advancements were made, horizontal scalability became a more feasible option. The commoditization of hardware away from highly scalable symmetric multiprocessors (vertical scaling) to more cost effective but less scalable x86 architecture has also invoked a change in the way enterprise applications are designed today, in order to meet scaling demands.

To scale a virtualized solution reliably and effectively, all components of the cloud must be designed with horizontal scalability in mind. For example, database management systems, such as IIBM Infosphere Warehouse, are capable of being scaled horizontally across many different x86 servers using database partitioning. By using virtualization, an environment of commoditized x86 servers is created to fulfill horizontal scalability requirements. Each and every component of the cloud solution must be able to scale within the means of this hardware.

2.4 Virtualized business intelligence architecture

The operational infrastructure of a BI cloud consists of a number of underlying hardware and software components.

2.4.1 Hardware components

With a wide array of choices available in today's server and mainframe market, it can be difficult to choose the appropriate hardware for your solution. Consider the hardware options when hosting a BI cloud solution:

- IBM BladeCenter is an integrated and available system of blade servers and storage management. It can help to reduce energy costs with its efficient cooling technologies. It is designed to handle flexible scalability with multiple building blocks.
- IBM System x is both affordable and scalable, making it suitable as an initial investment in a cloud. The x86 Intel-based System x servers are designed to maximize memory, minimize cost, and improve application performance while also managing growth, complexity, and risk.
- IBM POWER Systems offer the benefits of consolidating workloads with more performance and capacity. POWER systems use virtualization technology to consolidate multiple workloads onto fewer systems and can increase server utilization and reduce cost.
- IBM System z supports large scale consolidation of multiple environments. Availability, scalability, and security are key cloud-related features of the System z platform. The System z platform can run multiple instances simultaneously, provision new instances quickly, and track its capacity efficiently.

2.4.2 Software (middleware) components

Choosing the correct middleware for a BI cloud solution is imperative for a successful deployment. Is the middleware cloud-ready? Are the components

capable of handling requirements such as scalability, varying workload sizes, security, and performance?

InfoSphere Warehouse architecture

InfoSphere Warehouse 10.1 is implemented using a three-tier topology: the database server, the application server, and the client. These components can be arranged in a variety of ways:

- Single-tier architecture puts the database server, application server, and a client together on one system (a physical server or a virtualized server). This configuration is typically reserved for proof-of-concept purposes or for development and testing. It is not recommended for production environments.
- Two-tier architecture puts the database server and application server on one system and the client or clients on another or multiple other systems. This configuration is used for small or medium sized workloads with multiple developers. The server components are installed on a powerful system with processors and memory assigned to them, whereas the clients are installed on each warehouse developer's personal computer or virtual machine. The client typically requires fewer computing resources than the servers.
- Three-tier architecture puts the server components and the clients all on different systems, as shown in Figure 2-14. Processing power and data capacity can be increased with additional server installations. This configuration is often used when there is a heavy query load and many users or applications connected to InfoSphere Warehouse. High-speed networking and communications are required across all tiers.

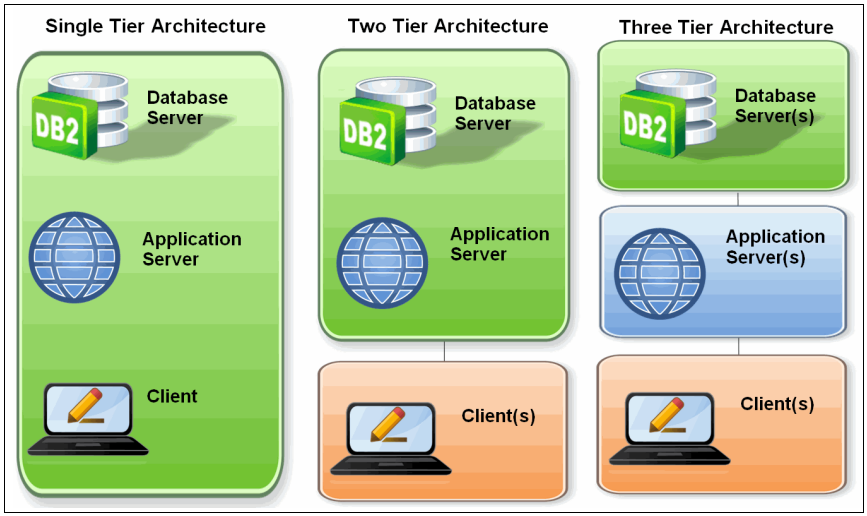


Figure 2-14 InfoSphere Warehouse topology

IBM Cognos Business Intelligence architecture

The IBM Cognos Business Intelligence environment includes the following components:

- Content Store

  Content Store is a relational database that contains data that IBM Cognos Business Intelligence products need to operate, such as report specifications, published models (and the packages that contain them), connection information for data sources, information about the external namespace (and the Cognos namespace itself) and information about scheduling and bursting reports. The relational database can be Cognos Content Database or another supported database.

- Content Manager

  Content Manager is the service that manages the storage of customer application data, including security and configuration data, models, metrics, report specifications, and report output. Content Manager is needed to publish packages, retrieve or store report specifications, manage scheduling information, and manage the Cognos namespace.

- Application server

  The application or logic tier contains one or more servers to run requests, such as reports, analyses, and queries, that are forwarded by a gateway. An IBM Cognos Business Intelligence server also renders the IBM Cognos Connection and studio interfaces.

- Gateway server

  Communication in the IBM Cognos architecture is typically done through gateways, which reside on one or more web servers. A gateway is an extension of a web server program that transfers information from the web server to another server. Gateways are often CGI programs, but can follow other standards such as Internet Server Application Program Interface (ISAPI) and Apache Modules (apache_mod).

2.4.3 Example mainframe architecture

Mainframes allow end-to-end workloads and business processing in both transactional and BI environments, with great tolerance for scalable workloads. They reduce complexibility and risk by enabling unified management tools for multi-platform environments. This allows all processes to be integrated into one machine rather than many distributed servers.

Figure 2-15 illustrates how a System z mainframe can be deployed for cloud computing solutions.

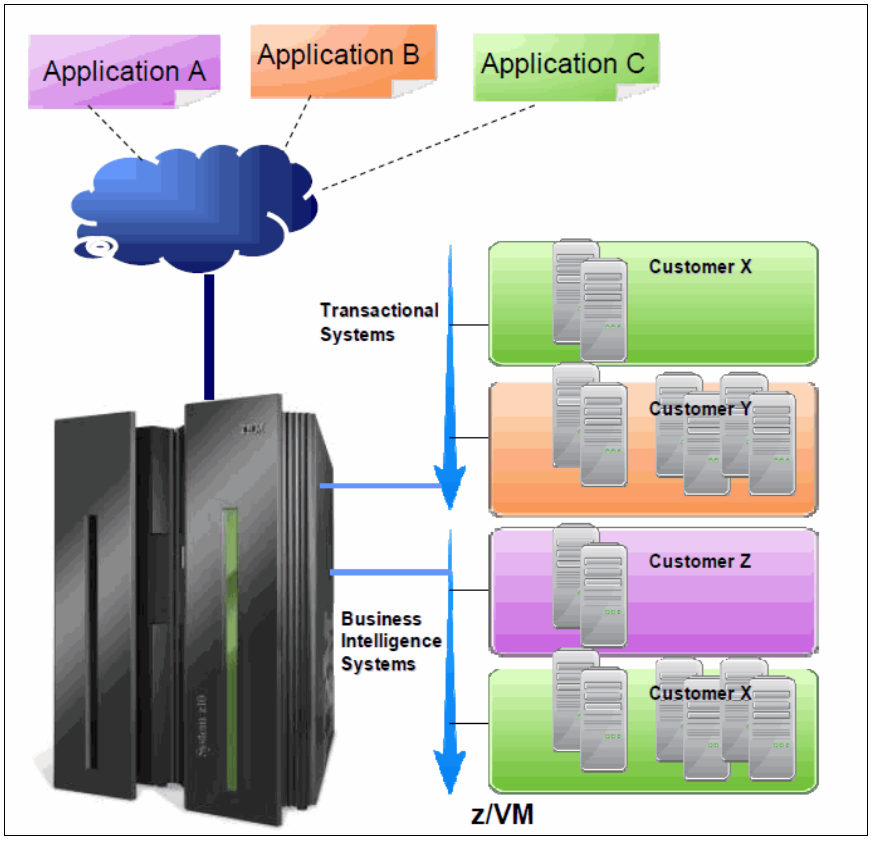


Figure 2-15 System z cloud architecture

2.4.4 Example server architecture

Consider the following cell-based architecture, also known as clustering. It is a proven practice that increases the resilience, performance, and availability of cloud solutions.

A cell is composed of a pair of servers that function in conjunction with each other to take advantage of the shared or pooled resources available within the cell to load-balance the work and achieve high availability. Depending on the functionality of the cell, the servers' specifications can change to meet performance needs. Figure 2-16 on page 80 shows how cell architecture can be used to meet scalability requirements in the Cognos, gateway, and database server tiers.

The server clusters must be able to meet increasing performance and capacity demands. IBM System x and IBM BladeCenter servers can scale vertically and horizontally, which can make it easier to adopt cloud computing at a measured pace. Some IBM server models can even scale up by adding more CPU, memory, or disk space to avoid performance bottlenecks.

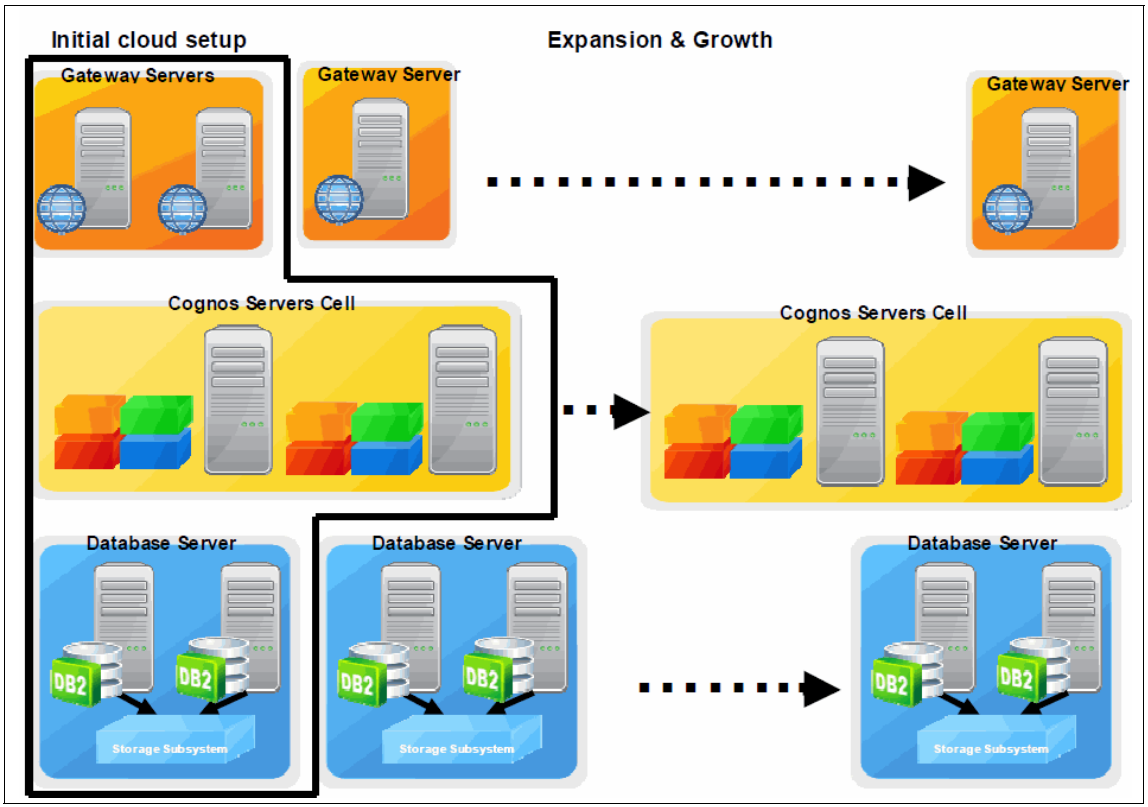


Figure 2-16 Cell architecture for gateway, Cognos, and database server

For example, if a cell is required to support a database server, two servers must be implemented with adequate memory and CPU capacity to avoid latency while still maintaining performance for database queries and updates. Inside the database server cell, a storage subsystem is typically shared by the two servers to configure a sufficient number of disk drives to match the processing power of the CPU and to separate the logs, temporary tablespaces, and user data into different disk arrays. The external storage subsystem is added in the database server cell due to the anticipated high number of I/O requests to bring data records into memory, and maximizes overall reporting and query performance.

Database server cell

Scalability in the cell architecture is achieved by integrating multiple cells within the cloud. For example, if you need to increase the capacity of your database, additional cells can be added consisting of two servers sharing a single external storage subsystem.

Figure 2-17 shows the physical breakdown of a database server cell.

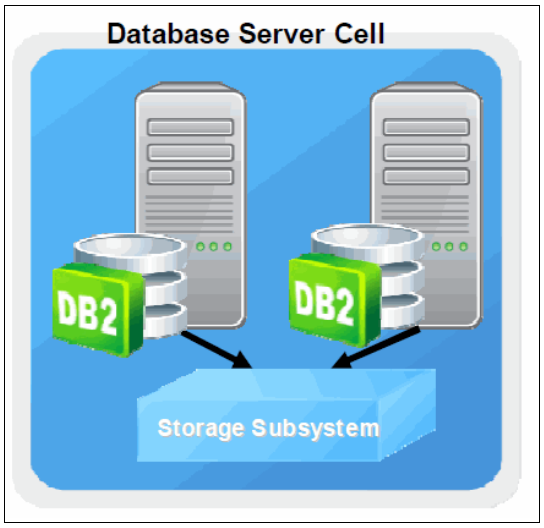


Figure 2-17 Database server cell

The database server cell includes the following components:

- Hardware:
  - Two servers (System x or Blade)
  - At least one storage subsystem
  - Optional additional (*N#*) storage subsystems
- Software (high availability):
  - One DB2 database server per physical server
  - Optional additional (*N#*) DB2 database servers
- Database instances (customer or application isolation):
  - Multiple instances per database server
  - Two DB2 instances per customer (user data and Cognos metadata)
  - Mirrored instances on both servers, one active and one standby
- Databases:
  - One user database per customer
  - One Cognos Content Store database per customer

- One InfoSphere Warehouse control database per customer
- Mirrored databases on both servers, one active and one standby
- Half of the databases active on Server 1 and the rest active on Server 2
- Two load-balancing resource pools

- Disk layout (reduce latency on I/O):
  - Data, logs, and backup and temporary spaces in external disks
  - Instance home and catalog tables in internal disks

Cognos server cell

If a cell is required to support Cognos servers for BI, then the initial cluster is composed of two servers. The Cognos software components are the Content Manager, the Cognos dispatchers, and the Cognos gateways. To provide scalability, the Cognos components are installed in different instances.

Each of the Cognos components must have a clustered or standby component to make the BI environment highly available and to load balance the incoming requests, as shown in Figure 2-18. Each BI environment is dedicated to a particular customer to achieve isolation for secured access and reports.

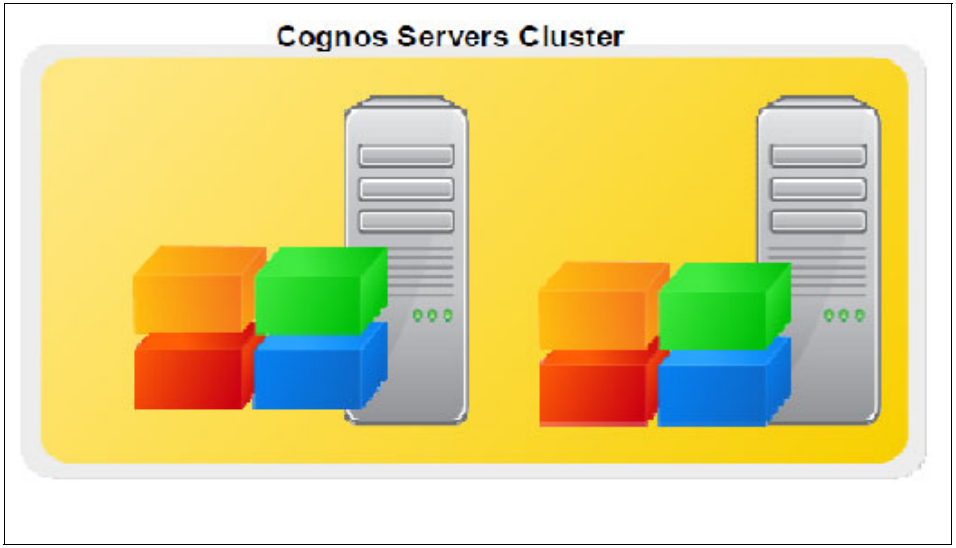


Figure 2-18 Cognos server cluster

The Cognos server cluster has the following configuration:

- Hardware

  Two Servers (System x or Blade)

- Software
  - Distributed Cognos installations and configurations
  - Multiple Cognos Content Manager installations per server
  - Two Cognos Content Manager installations per customer, one in each server (active and standby)
  - Multiple Cognos Dispatcher installations (at least two) per customer for load balancing
  - One DB2 client installation on each server (if the Cognos customer installations are isolated from each other, then one DB2 client installation is needed for each Cognos Content Manager installation)

Figure 2-19 shows how high availability and load balancing are performed in a Cognos environment that implements a cell architecture.

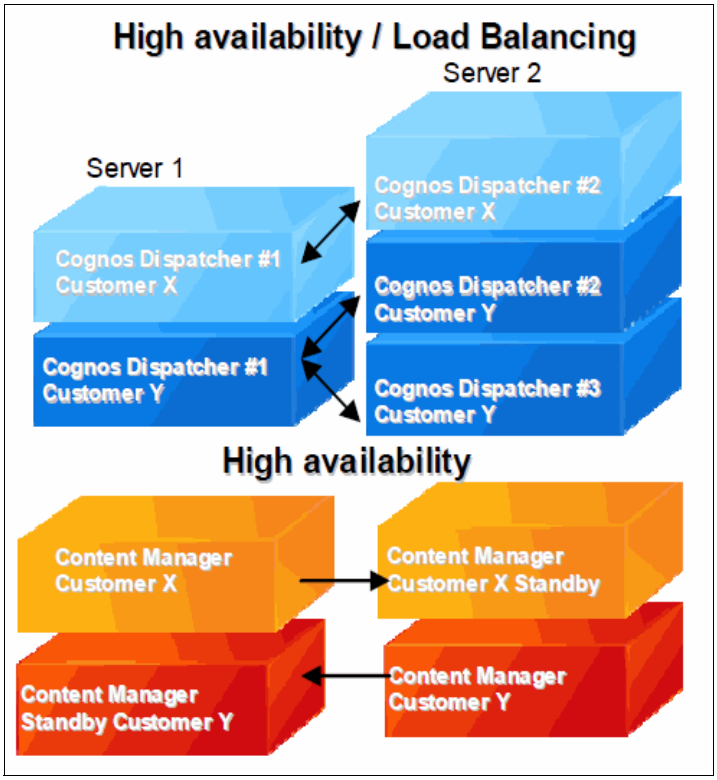


Figure 2-19 Cell-based Cognos environment with high availability and load balancing

Because the Cognos gateway components receive all user requests, they can create a bottleneck if resources in the layer are not sufficient. Therefore, in some cases, you can separate the gateways from the rest of the components using two additional servers, creating a *gateway server cluster* as shown in Figure 2-20 on page 85. If more than one gateway server is added to the BI environment, additional load balancing software will be needed on top of the gateways.

Figure 2-20 Gateway servers cluster

Components of a gateway server cluster include the following:

- Hardware

  Two servers (System x or Blade)

- Software (high availability)
  - Multiple Cognos gateway installations (at least two) per customer for load balancing user requests
  - Web server (WebSphere or Apache)

Virtualization of cell architecture

Enabling technologies such as virtualization increase the flexibility of hardware and allow for increased automation and faster provisioning of software. Virtualization enables multiple virtual machines to host and power rich client-server software. In addition, virtualization allows for resources to be more efficiently used during times when the software and services are not needed.

A virtualized cell needs at least two servers and a storage subsystem, as shown in Figure 2-21. The virtual machines and VM templates are stored in standby mode in inside storage. When a request is initialized, a VM starts consuming resources such as memory and CPU. The rich client software is dedicated to designing and developing advanced BI analytics. Another possibility is to set up other VMs with monitoring software or server capabilities.

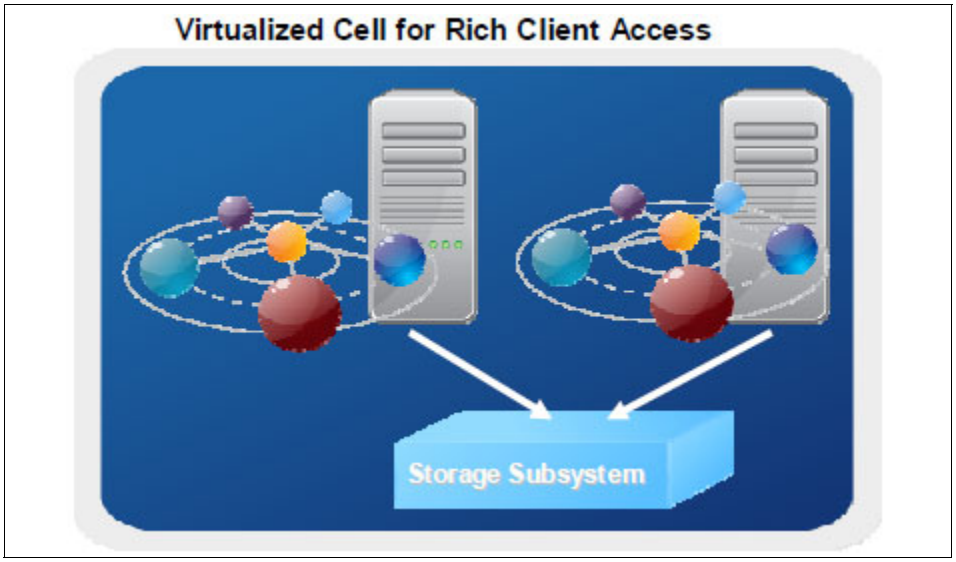


Figure 2-21 Virtualized cell for client access

A virtualized cell includes the following components:

- Hardware:
  - Two servers (System x or Blade)
  - At least one storage subsystem
  - Optional additional storage subsystems for scaling
- Software:

  Virtualization software, for example, VMWare
- VM software:
  - Multiple VMs with InfoSphere Warehouse Design Studio available
  - Optional multiple VMs per customer with InfoSphere Warehouse Application Server (at least one for cubing services and SQL Warehousing scheduling)
  - Optional multiple VMs with Optim Performance Manager
  - Multiple VMs with Cognos Transformer

  - Multiple VMs with Cognos Framework Manager
  - Optional multiple VMs with InfoSphere DataStage Designer
- Multiple VMs with InfoSphere DataStage Designer (Optional)

This architecture is just one example of how to set up a cloud infrastructure. IBM System x and IBM BladeCenter provide flexibility to customize any environment to fit budget restrictions and growth requirements.

Example server hardware: BladeCenter configuration

The BladeCenter shown in Figure 2-22 has the capacity for additional computing power as needed through the addition of extra Blade servers into the chassis. Features of the BladeCenter H Chassis include the following:

- 9U rack-optimized chassis
- Support for up to 14 hot-swappable HS22 Blades and up to 10 switches or 8 bridges for internal communication
- Support for Virtual Fabric, FcOE, 10 Gb Ethernet, 8 Gb Fibre Channel, InfiniBand, and more through up to four existing module bays, up to four high-speed module bays and up to four bridge module bays
- Up to four 2980 W hot-swappable and redundant load-balancing power supply modules

Figure 2-22 BladeCenter H chassis

HS22 Blade servers

Each Blade server provides additional dedicated or virtual computing capacity to process more software or larger workloads. The HS22 Blade servers include the following configuration:

- Two Intel Xeon 5670 processors (up to 3.06 GHz)
- Up to 192 GB of memory with 12 VLP DDR-3 memory DIMMs
- One CIOv slot (standard PCIe daughter card) and 1 CFFh slot (high-speed PCIe daughter card)
- Virtual Fabric Adapter integrated in some models
- Broadcom 5709S onboard NIC with dual Gigabit Ethernet ports with TOE
- RAID-0, -1 and -1E (optional RAID-5 with battery-backed cache)
- Integrated LSI 1064 SAS/SATA Controller that supports up to two internal SAS or SATA HDDs of 148 GB 10k RPM or two internal SSDs with RAID 0/1 support
- Blade server arrangement

Table 2-1 shows an example non-virtualized Blade server configuration for a Cognos or Infosphere Warehouse environment.

Table 2-1 BladeCenter H configuration with Blade servers

Blade server # (Bay #)	Server hostname	Main function
Blade Server 1 (Bay 1)	SERVER01	ISW Primary dataserver
Blade Server 2 (Bay 2)	SERVER02	ISW Secondary dataserver
Blade Server 3 (Bay 10)	SERVER10	Cognos Content Manager and dispatchers
Blade Server 4 (Bay 11)	SERVER11	Cognos Content Manager and dispatchers
Blade Server 5 (Bay 12)	SERVER12	Cognos gateway and services
Blade Server 6 (Bay 13)	SERVER13	Cognos gateway and services
Blade Server 7 (Bay 14)	SERVER14	Client virtual machines

DS3524 Express Dual controller storage system

The DS3524 Express Dual controller storage system has the following configuration:

- Eight 8-Gbps FC Ports and twenty-four 300-GB hard disks
- 2 GB cache upgrade
- 600 GB 3.5-inch 15K 6-Gb SAS hard disk
- 8 Gb FC 4-Port daughter card
- Turbo Performance
- 2.8 m, 10A/230V, C13 to CEE7-VII (Europe)
- 5 m fiber optic cable LC-LC
- Three-year, 24x7 on-site repair with 4-hour response

This hardware configuration is used in an actual BI private cloud configuration of an IBM business partner. The middleware components (Infosphere Warehouse and Cognos) are non-virtualized. However, the data warehouse and BI services layers are duplexed and mirrored to achieve high availability, and the non-critical support software and services are virtualized.

The configuration achieves maximum performance in the BI and data warehousing layers because there is no virtualization underneath. Drawbacks are that there is less flexibility and resource allocation is not as efficient. Multi-tenancy in database environments can still be achieved thanks to how InfoSphere Warehouse and Cognos are architected. Scalability is still achievable because of the BladeCenter hardware architecture, which allows new Blade servers and storage to be added to meet additional workload demand.

Integrating public cloud services

Deploying an application using public cloud services is similar to using in-house infrastructure. In fact, the more software-as-a-service components that are being used, the less complex it is to implement the final solution. Cloud infrastructure and services abstract many of the fine details away, so decision making is easier. Businesses that use public cloud infrastructure and services are generally less concerned about performance and hardware details, and more concerned about delivering reliable, functioning, value-driven solutions.

An example os software-as-a-service that is commonly used by today's solutions is something called *database on the cloud*. DB2 is available on today's most popular cloud hosting platforms, including IBM SmartCloud Enterprise, Amazon Web Services, and GoGrid cloud. The DB2 images on these infrastructures are easy to set up because the software is pre-loaded into the images and requires

little configuration. When deciding to use public cloud computing services, one must consider the impact of the service on the overall solution.

For example, because database functionality is an integral part of any solution, network access to the database must not only be secure, but also fast. As such, other underlying components and services that revolve around the cloud database service must also be co-located at the same site as the infrastructure hosting the database. In most cases, when using a database on a public cloud, an entire cloud computing solution must be built around that same public cloud infrastructure as a way to ensure optimal performance, security, and ease of access from one layer to another.

Introducing a public cloud service into a private cloud solution is essentially like opening a door to the outside world with your solution. When using third-party cloud services, one of the biggest concerns is how to keep the data and your solution highly secured.

3

Implementation of a virtualized business intelligence cloud

This chapter explains how to set up IBM InfoSphere Warehouse on virtualized environments as the foundation of a business intelligence (BI) cloud platform. It includes the following topics:

- Implementing IBM InfoSphere Warehouse on virtualized environments
- Implementing InfoSphere Warehouse Departmental Edition virtual image
- Implementing IBM Cognos Business Intelligence
- Integrating additional software

3.1 Implementing IBM InfoSphere Warehouse on virtualized environments

You can use IBM InfoSphere Warehouse on virtualized environments to provide scalability and to enhance performance.

VMware environments: In VMware environments, you can use an InfoSphere Warehouse Departmental Edition virtual image on a SUSE Linux Enterprise Server operating system to speed deployment. For more information, see 3.2, “Implementing InfoSphere Warehouse Departmental Edition virtual image” on page 111.

This section explains how to install and configure InfoSphere Warehouse for the following environments:

- Virtualization platforms other than VMware, such as KVM or IBM PowerVM®
- Operating systems other than a SUSE Linux Enterprise Server operating system, such as Red Hat or Windows operating systems
- Distributed installations of InfoSphere Warehouse server components that use a three-tier architecture to support a more robust environment

To comply with a scalable cloud architecture, you can install each of the InfoSphere Warehouse components (database, application server, and client) on separate virtual machines (VMs) to scale up or scale out when the performance of a specific component must be improved. InfoSphere Warehouse can scale in the following ways:

- Scaling up (vertically) can be achieved by adding virtual resources (CPU, memory, and disk space) to the VM that is hosting the software to increase performance.
- When scaling up is not possible, scaling out (horizontally) is another way to continue to improve performance. This method is achieved by adding more VMs with similar hardware specifications.

  InfoSphere Warehouse tiers can be scaled out in the following ways:

  - By provisioning more database partitions to store large amounts of user data or allocate more complex workloads
  - By provisioning more application servers to host more cube servers, thus providing better OLAP processing
  - By provisioning more VMs with IBM Design Studio to support a larger number of warehouse developers

Figure 3-1 illustrates how InfoSphere Warehouse tiers can scale out with virtualization.

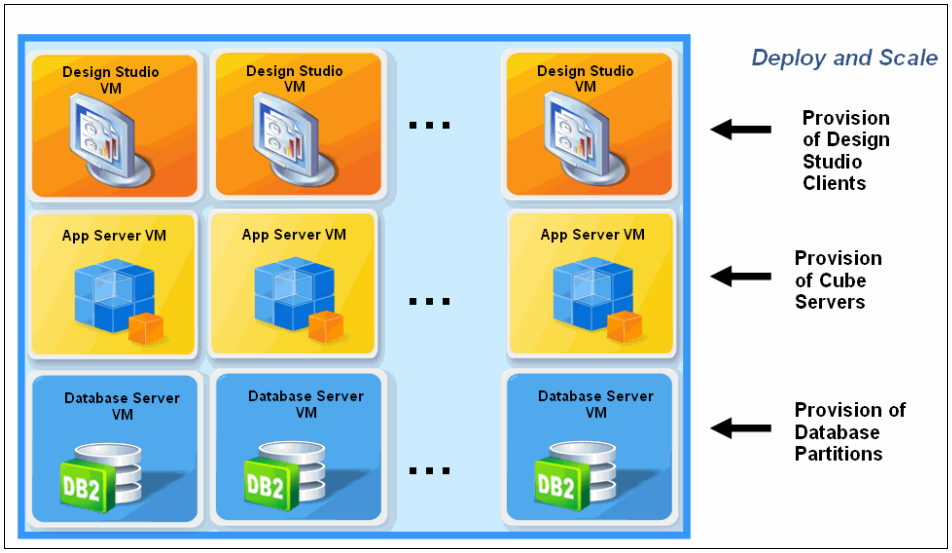


Figure 3-1 Scaling out InfoSphere Warehouse tiers using virtualization

Additional components: InfoSphere Warehouse includes additional key components, depending on the edition purchased, that you can install and configure to work in conjunction with InfoSphere Warehouse to expand its functionality. For more information, refer to 1.4.4, “InfoSphere Warehouse companion products” on page 35.

Before you install InfoSphere Warehouse, ensure that each component is installed on a supported hypervisor, hardware architecture, and guest operating system, as detailed in the following requirements:

http://www.ibm.com/developerworks/wikis/display/im/InfoSphere+Warehouse+Virtualization+Support

Ensure that your system satisfies at least the minimum requirements. A production system needs significant planning to determine the appropriate configuration, based on factors such as expected size of the warehouse, user population, and type of queries or reports requested, especially during peak hours. For a detailed list of the system requirements and installation instructions, see:

http://www.ibm.com/support/docview.wss?uid=swg21255099

You can obtain the InfoSphere Warehouse server product packages by accessing the product physical disk media or downloading the files from the IBM Passport Advantage® website. For more details about how to download IBM InfoSphere Warehouse 10.1, follow the instructions at this website:

http://www.ibm.com/support/docview.wss?uid=swg24032059

You can install InfoSphere Warehouse tiers using either of the following methods:

- Attended installation

  This installation method allows you to specify installation preferences using a graphical interface, also referred as an *installation wizard*. You can set the following preferences:

  - Component to install

    Choose the InfoSphere Warehouse component that you want to install, either the database server, application server, or client.

  - Activation code

    Specify where the installation program can find the activation or licensing files to proceed with the installation.

  - Software license agreement

    Read the license agreement carefully, and accept the terms to continue with the installation wizard.

  - Installation method

    Determine which of the following tasks you want to execute:

    - Proceed with installation now.
    - Save your selected preferences in a response file, without installing the product now.
    - Proceed with installation now, and save your selected preferences in a response file.

  - Installation directory

    Accept the default installation directory or specify a different installation directory.

  - Features to install

    Select the features that you want to install on a particular system.

  - Custom information

    Provide the custom information that is needed for each InfoSphere Warehouse component.

- Pre-installation summary

  Before the components are installed, review a summary of the preferences that you chose in the installation wizard.

- Silent installation using a response file

  The silent installation method automates the installation and requires no manual intervention or interaction. Silent installation of the InfoSphere Warehouse components requires a *response file*. A response file contains the values that the installation program uses to set up the component that is required. You can use either the sample response file or create your own response file by running the installation wizard.

In a cloud environment, the installation of software components can occur on multiple VMs at a time. For this purpose, you can simplify and automate the installation process. You can embed the execution of an InfoSphere Warehouse silent installation into an operating system shell script, and the script can be executed without any user input. You can call this script from a centralized system that can access many VMs remotely, typically through a local area network or an enterprise intranet, anytime that a new InfoSphere Warehouse component is required.

The following example shows how to prepare for a remote installation of InfoSphere Warehouse on a Linux guest VM:

1. To begin an installation, have the hypervisor and the guest operating system up and running. Ensure that you have root or administrator authority in the guest operating system.
2. Remote access is required as part of the installation. Connect to the guest operating system using secure shell (SSH), which allows you to execute command line based functions. For example, run the following command to open SSH to the Linux guest operating system:

   ```
   ssh root@iswserver.ibm.com
   ```

3. If you are installing the data server on a Linux operating system, set the default file creation permission by running the following command in the shell:

   ```
   umask 022
   ```

4. Download the installation image of InfoSphere Warehouse server and the activation image of the license profile to a temporary directory on the VM, for example `/temp/SW_Download`.

   If your VMs have no access to the Internet, use any secured file transfer protocol, for example SCP or SFTP. Example 3-1 on page 96 shows how to use SCP to copy the product and activation code packages from a library server across an SSH connection.

Example 3-1 Copy a file using SCP

```
deployserver:~ #scp root@library.ibm.com:/SW/iswserver.tar.gz
root@iswdbserver.ibm.com:/temp/ISW/
deployserver:~ # scp root@library.ibm.com:/SW/iswlicense.tar.gz
root@iswdbserver.ibm.com:/temp/ISW/
```

5. Create a temporary directory on which you can extract the installation images for the InfoSphere Warehouse components. For example, you can call the directory `wh_server` for the server or `wh_client` for the client installation files. The directory must not contain blank spaces and should have enough disk space. For example, run the following command to create a directory on Linux:

   ```
   mkdir /temp/wh_server
   ```

6. Extract the contents of the installation image into the folder using a compression tool.

   Example 3-2 shows how to extract the contents into the `warehouse_server` directory on Linux.

Example 3-2 Extract the contents of a file using GNU tar compression tool

```
iswdbserver:~ #cd /temp/ISW
iswdbserver:~ #tar -xzf iswserver.tar.gz /temp/wh_server
iswdbserver:~ #tar -xzf iswlicense.tar.gz
/temp/wh_server/isw/profile
```

3.1.1 The database server tier

The Infosphere Warehouse database server tier installs DB2 Enterprise Server Edition, which is used to store and consolidate data on a single enterprise data warehouse repository to support reporting and analysis. The database server is the base layer of an InfoSphere Warehouse implementation and operates independently of any tier. Each database server implementation can coordinate one administration node and multiple data nodes to handle a massively parallel processing (MPP) architecture. This type of implementation is ideal for large databases and when significant aggregation and complex analytics are needed.

The database server tier also includes the InfoSphere Federation Server Relational Wrappers and the Intelligent Miner components:

- InfoSphere Federation Server

  This component helps you with the creation of a federated data warehouse. This type of warehouse can send distributed requests to multiple data sources (local and remote) within a single SQL statement, as though all data

is stored locally in the federated data warehouse. This Federation function takes precedence even when the data sources use different SQL from DB2 or do not support SQL at all. A consolidated or federated data warehouse creates a common business model and a shared staging area to keep information consistent across all business analysis.

- Intelligent Miner

  This component is embedded into DB2. With it, you can perform in-database mining based on DB2 tables, stored procedures, and user-defined functions. This type of data mining style is useful for handling large data volumes and for taking advantage of the DB2 parallelism. It also hides most of the complexities of data mining and allows users who are not statisticians or programmers to productively mine data using pre-programmed DB2 functions.

Figure 3-2 depicts the InfoSphere Warehouse database server tier components.

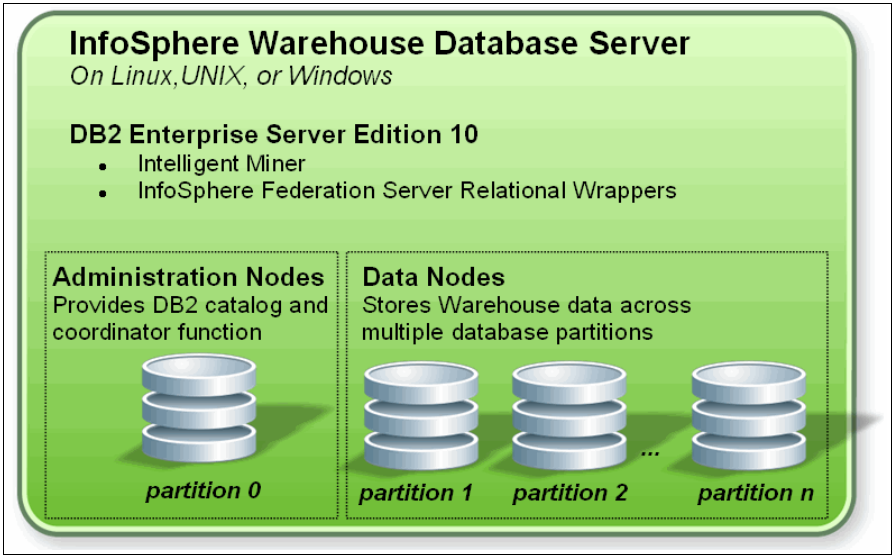


Figure 3-2 InfoSphere Warehouse Database Server components

Physical servers have more resources available and, in most case, are not fully utilized. Using virtualization, the administration node and each data node can be installed in separate VMs with dedicated resources (CPU, I/O, and memory) to emulate an MPP architecture. Each virtualized partition works in parallel, which can result in query performance improvement on data servers that contain large amounts of data.

Additionally, virtualization allows you to share or consolidate pools of system resources that can provide flexibility to create mixed environments. For example, you can reduce the number of physical servers that are needed by virtualizing the physical server into two or more database server installations. As a result, this virtualized environment can support a data warehousing testing environment in one VM and a data warehousing production environment on another VM.

To perform a silent installation of the database server tier on a Linux guest:

1. Begin by editing the response file to meet your installation requirements. The sample `PROC_ISW.rsp` response file is located in the `ResponseFiles` directory of the InfoSphere Warehouse installation image.

   Example 3-3 illustrates a sample response file generated by the installation wizard to install InfoSphere Warehouse database server.

   Example 3-3 Sample response file to install InfoSphere Warehouse Database Server

```
#Profile using variables
#-----------------------
PRODUCT_NAME=IBM InfoSphere Warehouse
EXCLUDED_FEATURES=

#Disable OS Pre-requisite checking
#---------------------------------
DS_IGNORE_OSPREREQ=FALSE

#Has the license been accepted
#-----------------------------
#LICENSE_ACCEPTED=TRUE

#Installation Directory
#----------------------
USER_INSTALL_DIR=/opt/IBM/ISWarehouse
INSTALL_NEW_COPY=true

#Selected Features in the Choosen Install Sets panel
#---------------------------------------------------
CUSTOM_CHOSEN_INSTALL_FEATURE_LIST=DATA,ESE,IM,RCON

#DB2 Selection Panel
#-------------------
DB2_COPY_LOCATION=/opt/IBM/ISWarehouse/db2/V10.1

#DB2 DAS User Panel
#------------------
```

```
DB2DAS_PASSWORD.encrypted=true
DB2DAS_NAME=dasusr1
DB2DAS_GROUPNAME=dasadm1
DB2DAS_HOMEDIR=/home/dasusr1
DB2DAS_PASSWORD=axfeO2ki3kdi383
DB2DAS_PANEL_TYPE=DAS
DB2DAS_CONPASSWORD=axfeO2ki3kdi383

#DB2 Instance User Panel
#----------------------
DB2INSTANCE_PASSWORD.encrypted=true
DB2INSTANCE_NAME=db2inst1
DB2INSTANCE_GROUPNAME=db2iadm1
DB2INSTANCE_HOMEDIR=/home/db2inst1
DB2INSTANCE_PASSWORD=axfeO2ki3kdi383
DB2INSTANCE_PANEL_TYPE=INST
DB2INSTANCE_CONPASSWORD=axfeO2ki3kdi383

#DB2 Fenced User Panel
#--------------------
DB2FENCE_PASSWORD.encrypted=true
DB2FENCE_NAME=db2fenc1
DB2FENCE_GROUPNAME=db2fadm1
DB2FENCE_HOMEDIR=/home/db2fenc1
DB2FENCE_PASSWORD=axfeO2ki3kdi383
DB2FENCE_PANEL_TYPE=FENCE
DB2FENCE_CONPASSWORD=axfeO2ki3kdi383
```

You can edit the following information in this file:

– Product license

  The product license must be accepted to continue with the installation. The default value for the LICENSE_ACCEPTED is TRUE.

– Installation path

  The installation path can be changed in the USER_INSTALL_DIR variable. By default, the installation path for InfoSphere Warehouse on Linux is the `/opt/IBM/ISWarehouse` directory.

– Individual components

  You can customize the components that are to be installed by adding or removing values from the CUSTOM_CHOSEN_INSTALL _FEATURE_LIST variable.

  Example 3-4 on page 100 shows the list of features that you can add into a database server installation.

Example 3-4 Edit the features list variable to add or remove features

```
#To add Features, add the following values to the variable Custom
#Choosen Install Feature List:
#  DB2 Enterprise Server Edition: ESE
#  InfoSphere Federation Server Relational Wrappers: RCON
#  Intelligent Miner: IM
#-----------------------------------------------
CUSTOM_CHOSEN_INSTALL_FEATURE_LIST=DATA,ESE,IM,RCON
```

– DB2 installation path

You can install DB2 to an installation path of your choice. To change the installation path, change the value of the DB2_COPY_LOCATION variable from the response file. By default, the installation path for InfoSphere Warehouse on Linux is the `/opt/IBM/ISWarehouse/db2/v10.1` directory.

If you need to have multiple copies of DB2 on the same system, you can change the default path to the `/opt/IBM/ISWarehouse/db2/v10.1_##` directory, where `##` represents a sequential number in the range 01 to 99.

– DB2 users details

The DB2 administration server, DB2 instance owner, and DB2 fenced users are created automatically on the operating system while installing the database server. Before installation, you need to add details to the response file for each DB2 user, including the user name, group name, home directory, password, DB2 user type, and password encryption type.

These database users have the following characteristics:

- The DB2 instance user performs administration on DB2 and owns the home directory where the DB2 instance is created. This user controls all DB2 processes and owns all file systems and devices that are used by the databases that are contained within the instance.

  Instance names must be unique for all DB2 copies. The default user is db2inst1, and the default group is db2iadm1. If the default user already exits, you can change the default name to db2inst#, where # is a sequential number.

- The DB2 fenced user has privileges to run user-defined functions and stored procedures outside of the address space by the DB2 database. The default user is db2fenc1, and the default group is db2fadm1.

- The DB2 administration is used to run the DB2 administration server on the system. This role is deprecated in V9.7 and might be removed in a future release. In the meantime, only one DB2 Administration Server can be available on the same system.

The user names have restrictions and requirements. For more information consult the DB2 Information Center:

`http://publib.boulder.ibm.com/infocenter/db2luw/v10r1/topic/com.ibm.db2.luw.qb.server.doc/doc/c0011931.html`

– Passwords

When you create a response file using the installation wizard, the passwords are encrypted by default. However, if you need to modify the passwords before the silent installation, you can set the encryption property value to false and set the value to the actual passwords that you want to use.

Example 3-5 shows how to change the password on the response file.

Example 3-5 How to change the password on the response file

```
DB2INSTANCE_PASSWORD.encrypted=false
DB2INSTANCE_PASSWORD=apassw0rd
DB2INSTANCE_PANEL_TYPE=INST
DB2INSTANCE_CONPASSWORD=apassw0rd
```

The passwords length can be up to the maximum number of characters supported by your operating system.

2. After you have customized or generated the response file according to your requirements, create a shell script. Then, embed the data server installation command to install the InfoSphere Warehouse database server silently on a new VM.

   Example 3-6 shows the command to pass the response file that you created to install the Database Server on Linux.

   Example 3-6 Command to install the database server with a response file

   ```
   iswdbserver:/temp/wh_server # isw/install.bin -i silent -f
   /temp/PROC_ISW.rsp -Dprofile=/temp/wh_server/isw/profile/
   ```

3. When the installation is completed, a message indicates the outcome of the installation. If the installation fails, investigate the log files to troubleshoot the problem. The logs are located in the default directory for installation, the `/opt/IBM/ISWarehouse/logs` directory on Linux and UNIX.

3.1.2 The application server tier

The application server tier of the InfoSphere Warehouse topology is used to support the InfoSphere Warehouse *Administration Console*. The Administration Console is a web browser-based client for administering functions, such as SQL

Warehousing processing, Cubing Services, Intelligent Miner, and warehouse system resources. In the Administration Console, the warehouse application packages are deployed into executable modules that can be run or scheduled for execution on a production environment. All configurations added into the Administration Console rely on a metadata database hosted in the InfoSphere Warehouse database server and on a WebSphere Application Server environment:

- Metadata Database

  This component is a required repository database for storing InfoSphere Warehouse configurations that are added in the Administration Console. You can either create the database before starting the installation or the installation program can create it for you. The metadata database must be created in a non-partitioned database.

- WebSphere Application Server

  This component is an IBM runtime environment to support web accessible, distributed, and Java-based applications. This application server layer serves as a platform to deliver business application services to clients based on an SOA architecture. Every service that is created within WebSphere Application Server is secured, scalable, and reusable. The Integrated Solutions Console is the administration interface to manage the InfoSphere Warehouse Administration Console application and its role-based security.

The application server also includes the Administration Console command-line client, DB2 10 client, and Cubing Services components. The DB2 client is a lightweight software that is installed to provide remote connectivity to the database server sources. Using the DB2 client, you can catalog database servers and databases and add them within the Administration Console.

Figure 3-3 depicts the InfoSphere Warehouse application server components.

Figure 3-3 InfoSphere Warehouse Application Server components

To install the application server component of InfoSphere Warehouse on a Linux guest:

1. First install the Installation Manager and WebSphere Application Server. The following examples provide sample commands for installing IBM Installation Manager and WebSphere Application Server V8.0 silently.

   a. Install the Installation Manager as shown in Example 3-7.

   Example 3-7 Command to install the installation manager

   ```
   /temp/installc -acceptLicense
   ```

   b. Record a response file to install the WebSphere Application Server as shown in Example 3-8.

   Example 3-8 Command to record the install of WebSphere Application Server

   ```
   /opt/IBM/InstallationManager/eclipse/IBMIM -skipInstall
   /var/temp/imRegistry -record /var/temp/install_response_file.xml
   ```

   c. Use the response file to install the product as shown in Example 3-9.

   Example 3-9 Command to install using the Installation Manager

   ```
   /opt/IBM/InstallationManager/eclipse/tools/imcl -acceptLicense
   input /var/temp/install_reponse_file.xml
   ```

2. Edit the response file to meet your installation requirements. The sample `PROC_ISW.rsp` response file is located in the `ResponseFiles` directory of the installation image.

   Example 3-10 illustrates a sample response file that is generated by the installation wizard to install the InfoSphere Warehouse application server.

Example 3-10 Sample response file to install application server

```
#Profile using variables
#-----------------------
PRODUCT_NAME=IBM InfoSphere Warehouse
EXCLUDED_FEATURES=

#Disable OS Pre-requisite checking
#---------------------------------
DS_IGNORE_OSPREREQ=FALSE

#Has the license been accepted
#-----------------------------
#LICENSE_ACCEPTED=TRUE

#Installation Directory
#----------------------
USER_INSTALL_DIR=/opt/IBM/ISWarehouse
INSTALL_NEW_COPY=true

#Selected Features in the Choosen Install Sets panel
#---------------------------------------------------
CUSTOM_CHOSEN_INSTALL_FEATURE_LIST=WHADM,CLIENT,WAS,DSWeb,WEB_IM,WEB
_CS,WEB_SQW,WEB_UTA,ADMINCLIS,Cubing

#WebSphere Applicaiton Server
#----------------------------
was.installLocation=/opt/IBM/WebSphere/AppServer

#WebSphere Applicaiton Server Profile Name
#-----------------------------------------
PROFILE_NAME=AppSrv01

#WebSphere Security
#------------------
#set below variable if create new WAS profile
WAS_SECURITY_USER=root

#WebSphere Application Server Global Security
```

```
#-------------------------------------------
#set below variable if work with existing WAS profile
SecurityUser=dweadmin
SecurityPassword=axfeO2ki3kdi383
SecurityPasswordEncrypted=true

#Meta Data Panel
#---------------
MetaData.database.name=DSSDB
MetaData.database.port=db2inst1
MetaData.database.host=iswdbserver
#Option of metadata database
#   -CREATE_NEW_FOR_NEW_DB : create new database
#   -NEWESE_REUSEDB : re-use existing database
#   -CREATE_NEW_FOR_LOCAL_DB : re-use ESE and create new database
#   -REUSEESE_REUSEDB : re-use ESE and re-use existing database
MetaData.database.dbStatus=CREATE_NEW_FOR_NEW_DB
MetaData.database.user=db2inst1
MetaData.database.password=axfeO2ki3kdi383
MetaData.database.encrypted=true
#Database type
#   - DB2LUW
#   - DB2Z
#   - INFORMIX_JDBC
MetaData.database.type=DB2LUW
MetaData.database.idsServer=
MetaData.database.createNewDB=true
#set below variable re-use ESE and create database
REPDB_DB2INST_LOCATION=/opt/IBM/ISWarehouse/db2/V10.1
REPDB_DB2INST_NAME=DB2
```

You can edit the following information:

– Product license

  The product license must be accepted to continue with the installation. The default value for the LICENSE_ACCEPTED parameter is TRUE.

– Installation path

  You can change the installation path in the USER_INSTALL_DIR variable. By default the installation path for InfoSphere Warehouse on Linux is the `/opt/IBM/ISWarehouse` directory.

– Individual components

  You can customize the components that are to be installed by adding or removing values from the CUSTOM_CHOSEN_INSTALL_FEATURE _LIST variable.

  Example 3-11 shows the list of features that you can add for an application server installation.

  Example 3-11 Edit the features list variable to add or remove features

```
#To add or remove Features, edit the following values from the
variable Custom #Choosen Install Feature List:
#   -WHADM: installs the administration console
#   -CLIENT: installs the db2 client
#   -WAS: creates a WAS profile
#   -WEB_IM: Intelligent Mining administration
#   -WEB_CS: Cubing Services administration
#   -WEB_SQW: SQL Warehouse administration
#   -WEB_UTA: Unstructured Text Analysis
#   -ADMINCLIS: Administration Console Command Line Client
#   -Cubing: Cubing Services
#-------------------------------------------------
CUSTOM_CHOSEN_INSTALL_FEATURE_LIST=WHADM,CLIENT,WAS,DSWeb,WEB_IM,
WEB_CS,WEB_SQW,WEB_UTA,ADMINCLIS,Cubing
```

– WebSphere Application Server instance, profile, and operating system user

  Specify the location of the WebSphere Application Server installation, a name for a WebSphere Application Server profile, and a local operating system user to enable WebSphere Application Server global security.

– Metadata connection

  Provide the connection details to log in to the database server where the repository database is storing warehousing metadata. You can either create the database before starting the installation, or the installation program can create it for you. The default name of this database is DSSDB.

– Passwords

  When you create a response file using the installation wizard, the passwords are encrypted by default. However, if you need to modify the passwords before the silent installation, you can set the encryption property value to false and set the value to the actual passwords that you want to use.

Example 3-12 shows how to change the password in the response file.

Example 3-12 How to change the password on the response file

```
MetaData.database.user=db2inst1
MetaData.database.password=apassw0rd
MetaData.database.encrypted=false
```

The password length can be up to the maximum number of characters that are supported by your operating system.

3. After you customize or generate the response file according to your requirements, create a shell script. Embed the command to install the InfoSphere Warehouse application server silently on a new VM when it is required to be replicated.

   Example 3-13 shows the command to pass the response file that you created to install the application server on Linux.

   Example 3-13 Command to install the application server with a response file

   ```
   iswappserver:/temp/wh_server # isw/install.bin -i silent -f
   /temp/PROC_ISW.rsp -Dprofile=/temp/wh_server/isw/profile/
   ```

4. When the installation is completed, a message indicates the outcome of the installation. If the installation fails, investigate the log files to troubleshoot the problem. The logs are located in the default directory for installation, the `/opt/IBM/ISWarehouse/logs` directory for Linux and UNIX.

3.1.3 The client tier

The client tier is responsible for installing Infosphere Warehouse Design Studio, which provides a design environment for creating physical data models, OLAP modeling, SQL data flows, and control flows. Infosphere Warehouse Design Studio is a customizable design environment that is built upon the Eclipse platform. It integrates all the tasks of users of a data warehouse and business intelligence project within a unified environment.

Figure 3-4 depicts the InfoSphere Warehouse client components.

Figure 3-4 InfoSphere Warehouse client components

To install the client component of InfoSphere Warehouse on a Linux guest:

1. Edit the response file to meet your installation requirements. The sample `CLIENT_ISW.rsp` response file is located in the `ResponseFiles` directory of the installation image.

   Example 3-14 illustrates a sample response file generated by the installation wizard to install the client.

   Example 3-14 Sample response file to install InfoSphere Warehouse client

```
#Profile using variables
#-----------------------
PRODUCT_NAME=IBM InfoSphere Warehouse
EXCLUDED_FEATURES=

#Disable OS Pre-requisite checking
#---------------------------------
DS_IGNORE_OSPREREQ=FALSE

#Has the license been accepted
#-----------------------------
LICENSE_ACCEPTED=TRUE

#Installation Directory
#----------------------
```

```
USER_INSTALL_DIR=/opt/IBM/ISWarehouse

#Selected Features in the Choosen Install Sets panel
#-------------------------------------------------
CUSTOM_CHOSEN_INSTALL_FEATURE_LIST=WHDEV,CLIENTD,DSW_Main,OLAP_LUW,S
QW_LUW,MINING,TA,MININGB,IMVis,ADMINCLI,Doc,Samples

#DB2 Selection Panel
#-------------------
DB2_COPY_LOCATION=/opt/IBM/ISWarehouse/db2/V10.1

#DB2 Instance User Panel
#-----------------------
DB2INSTANCE_PASSWORD.encrypted=true
DB2INSTANCE_NAME=db2inst2
DB2INSTANCE_GROUPNAME=db2iadm1
DB2INSTANCE_HOMEDIR=/home/db2inst2
DB2INSTANCE_PASSWORD=axfeO2ki3kdi383
DB2INSTANCE_PANEL_TYPE=INST
DB2INSTANCE_CONPASSWORD=axfeO2ki3kdi383

#Select Install Package
#----------------------
New.install.package=true
PROFILE_ID=IBM Design Studio
SS_INSTALL_LOC=/opt/IBM/ISWarehouse
SHELLSHARE=false

#Extend Existing Eclipse
#-----------------------
Eclipse.JVM=
EXTEND_ECLIPSE_LOCATION=
EXTEND_ECLIPSE_OPTION=
```

You can edit the following information:

– Product License

  The product license must be accepted to continue with the installation. The default value for the LICENSE_ACCEPTED is TRUE.

– Installation path

  You can change the installation path in the USER_INSTALL_DIR variable. By default, the installation path for InfoSphere Warehouse on Linux is the `USER_INSTALL_DIR=/opt/IBM/ISWarehouse` directory.

– Individual components

You can customize the components that are to be installed by adding or removing values from the CUSTOM_CHOSEN_INSTALL_FEATURE _LIST variable.

The Example 3-15 shows the list of features that can be added for a Database Server installation.

Example 3-15 Edit the features list variable to add or remove features

```
#To add or remove Features, edit the following values from the
variable Custom #Choosen Install Feature List:
#   -WHDEV: installs the Warehouse Client
#   -CLIENTD: installs the db2 client
#   -DSW_Main: installs Design Studio
#   -OLAP_LUW: installs Cubing Services tools
#   -SQW_LUW: installs SQL Warehouse tools
#   -MINING: Intelligent Mining tools
#   -TA: installs the text analysis tools
#   -IMVis: installs Intelligent Miner visualization
#   -ADMINCLI: installs Administration Console CLI Client
#   -Doc: installs documentation
#   -Samples: install samples
#-------------------------------------------------
CUSTOM_CHOSEN_INSTALL_FEATURE_LIST=WHDEV,CLIENTD,DSW_Main,OLAP_LU
W,SQW_LUW,MINING,TA,MININGB,IMVis,ADMINCLI,Doc,Samples
```

– IBM Installation Manager

For the client, the IBM Installation Manager is required to perform this installation. If it is not already on your system, the installation program downloads it from IBM and installs it.

– Shared resources directory

The shared resources directory is used only by the Eclipse-based components of the Design Studio as part of the client component group. If the computer on which you install the client components does not already have a shared resources directory from a previous installation of a product that is compatible with InfoSphere Warehouse, the installation program creates the shared resources directory in the default location, which is the `/opt/IBM/ISWhared` directory on Linux and UNIX.

– Eclipse environment

For the client, unless you already have an Eclipse environment, simply leave Extend Existing Eclipse unspecified. Otherwise, you need to specify information about the Eclipse IDE that already is present.

– Passwords

When you create a response file using the installation wizard, the passwords are encrypted by default. However, if you need to modify the passwords before the silent installation, you can set the encryption property value to false and set the value to the actual passwords that you want to use.

Example 3-16 shows how to change the password in the response file.

Example 3-16 How to change the password in the response file

```
DB2INSTANCE_PASSWORD.encrypted=true
DB2INSTANCE_PASSWORD=axfeO2ki3kdi383
DB2INSTANCE_CONPASSWORD=axfeO2ki3kdi383
```

The password length can be up to the maximum number of characters that are supported by your operating system.

2. After you customize or generate the response file according to your requirements, create a shell script. Embed the command to install the InfoSphere Warehouse client silently on a new VM, when it is required.

   Example 3-17 shows the command to pass the response file that you create to install the application server on Linux.

   Example 3-17 Command to install the client with a response file

   ```
   iswclient:/temp/wh_client # isw/install.bin -i silent -f
   /temp/CLIENT_ISW.rsp -Dprofile=/temp/wh_client/isw/profile/
   ```

3. When the installation is completed, a message indicates the outcome of the installation. If the installation fails, investigate the log files to troubleshoot the problem. The logs are located in the default directory for installation, the `/opt/IBM/ISWarehouse/logs` directory for Linux and UNIX.

3.2 Implementing InfoSphere Warehouse Departmental Edition virtual image

InfoSphere Warehouse Departmental Edition is available on a virtual image. You can use this image to speed up and simplify deployment of a business intelligence solution. This virtual image offering includes the Infosphere Warehouse middleware. It allows you to bypass installation and configuration activities, which can reduce time-to-value and free database administrator resources to focus on other activities.

3.2.1 The virtual image OVF template

The virtual image comes as an OVF template. You can deploy this OVF template on any VMWare vSphere server with ESXI 4 or later for Intel or AMD-based platforms. The InfoSphere Warehouse Departmental Edition virtual image, purchased on a terabyte license basis, provides the flexibility to manage fluctuations in database activity. With InfoSphere Warehouse Departmental Edition, you have the freedom to adjust CPUs, data storage capacity, and hardware to adapt to changing business needs.

The OVF template provides an integrated and pre configured environment of InfoSphere Warehouse servers with a guest operating system of SUSE Linux Enterprise Server 11 SP2. This virtual offering enables the implementation of the database partitioning feature, in various single or multiple node configurations with multiple database partitions, on each single VM and in multiple VMs. This offering provides an integrated operating system and InfoSphere Warehouse software deployment in a virtualized environment. Using this virtual image can speed up the process of planning, installing, and configuring a system by yourself.

Before you deploy the OVF template through the vSphere client, you need to decide how many OVF templates are required for your setup. This decision depends on whether you want to use the database partitioning feature in a multi-node configuration.

After you download the OVF template for InfoSphere Warehouse, you can deploy the `.ovf` template and accompanying disks through your vSphere client environment. Keep in mind that for a multipartition environment, the number of virtual resources such as disks, CPU, and memory might increase as shown in Table 3-1.

Table 3-1 Minimum requirements for each VM

Topology	vHard Disk (OS)	vHard Disks (user data in GB)[a]	vCPUs[b]	vMemory[c]
Single VM (Single Partition)	1 x 60 GB	1 x # GB	1	1 GB + 4 GB
Single VM (Multiple Logical Partitions)[d]	1 x 60 GB	# of partitions x # GB[e]	# of partitions	1 GB + #partitions x 4 GB RAM
Multiple VMs (Multiple Cross-VM Partitions[d]	1 x 60 GB	# of partitions x # GB[e]	# of partitions	1 GB + # partitions x 4 GB RAM

a. Have each virtual hard disk on a separate data store.
b. Add 1 vCPU for every disk added to a VM.
c. Increase memory by 4 GB for each logical partition.
d. 8 partitions maximum per VM.
e. Do not add more than 8 disks to a single VM.

In production environments, you can minimize I/O disk contention when each vSphere data store that hosts a database partition maps to a separate underlying physical storage array. As per DB2 recommendations, configure the underlying storage disks in a RAID 5 configuration with a minimum configuration of three data disks and one parity disk.

3.2.2 Installing the vSphere administration client

To perform many basic tasks, such as reverting snapshots or increasing memory on VMs, you need to have the vSphere administration client installed on your development system or notebook. To download and install the administration client from your ESXI Server:

1. Open a browser and point it to the host on which your EXSI server is located (for example, `http://hostname`).
2. When the download page for VMWare ESXi 4 displays, click **download vSphere client**.
3. Save the installer to your desktop.
4. When the download is complete, launch the installer, choose the client installation, and follow all the default installation options.

3.2.3 Deploying the InfoSphere Warehouse OVF template

To deploy the InfoSphere Warehouse OVF template:

1. Launch the vSphere client and enter the host name, user name, and password parameters.
2. Click **File** → **Deploy OVF Template**.
3. In the Deploy OVF wizard, browse the file system, and locate the `.ovf` file. Click **Next** to continue.
4. Ignore the warning, and click **Yes** to continue.
5. Verify the OVF template details. Click **Next** to continue.
6. Provide a name up to 80 characters. The name must be unique within the inventory. Click **Next** to continue.
7. Select the resource pool within which you want to deploy this template. A resource pool allows hierarchical management over the VMs and child pools that it owns. Click **Next** to continue.

8. Select the data store in which you want to store this VM. The OVF template registers the following disks when deployed:
   - A default disk of 60 GB for the operating system files
   - A default disk of 10 GB for the first database partition
9. Select the data store that you want to host your operating system disk. The database partition disk is deployed automatically to the same data store. Afterwards, you can choose to move it to its own data store location. Click **Next** to continue.
10. Decide the format in which you want to store the virtual disks. The “Thin provisioned format” option allocates on demand as data is written to the virtual disks, allowing you to save space. For a production environment, use the “Thick provisioned format” option, which allocates storage immediately, to reduce disk errors and for improved performance. Click **Next** to continue.
11. Map the networks used in this OVF template to networks in your inventory. Click **Next** to continue.
12. Review the deployment settings, and click **Finish** to continue. Wait until the deployment is finished.
13. When the deployment is complete, right-click the deployed VM through your vSphere client, and select the item **Edit Settings** link from the Commands section. You need to add or remove computing resources, such as CPUs, memory, and virtual disks, depending on your ideal configuration, as shown in Figure 3-5 on page 115.

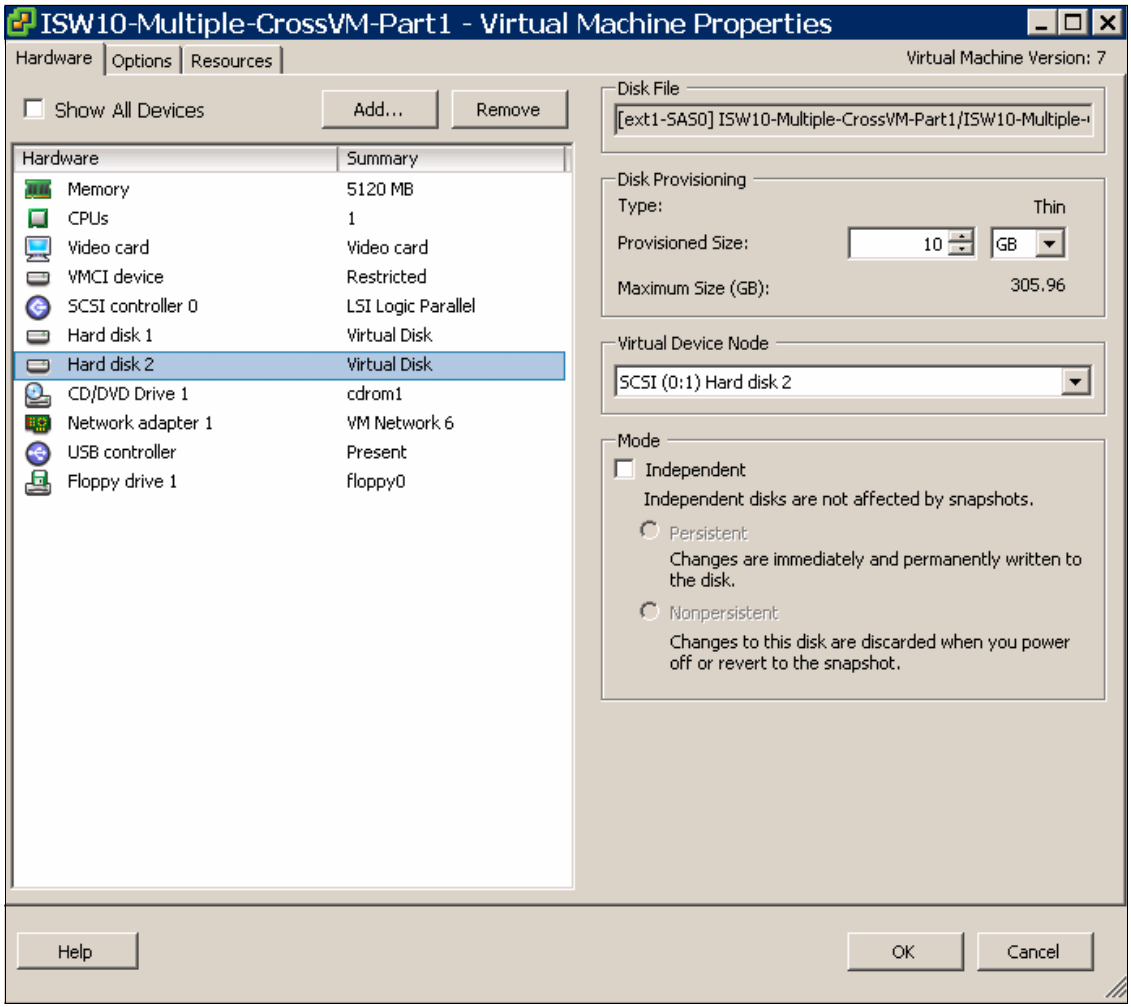


Figure 3-5 Resources configuration for the InfoSphere Warehouse virtual offering

Remove the default virtual hard disk

For each VM, the default disk that is assigned to the first data partition is 10 GB. If you plan to have a different data size, you need to remove the default “Hard disk 2” option as follows:

1. Select the virtual disk called **Hard Disk 2** and click **Remove**.
2. From the removal options, select **Remove from VM and delete files from disk** as shown in Figure 3-6, and click **OK** to accept the disk removal.

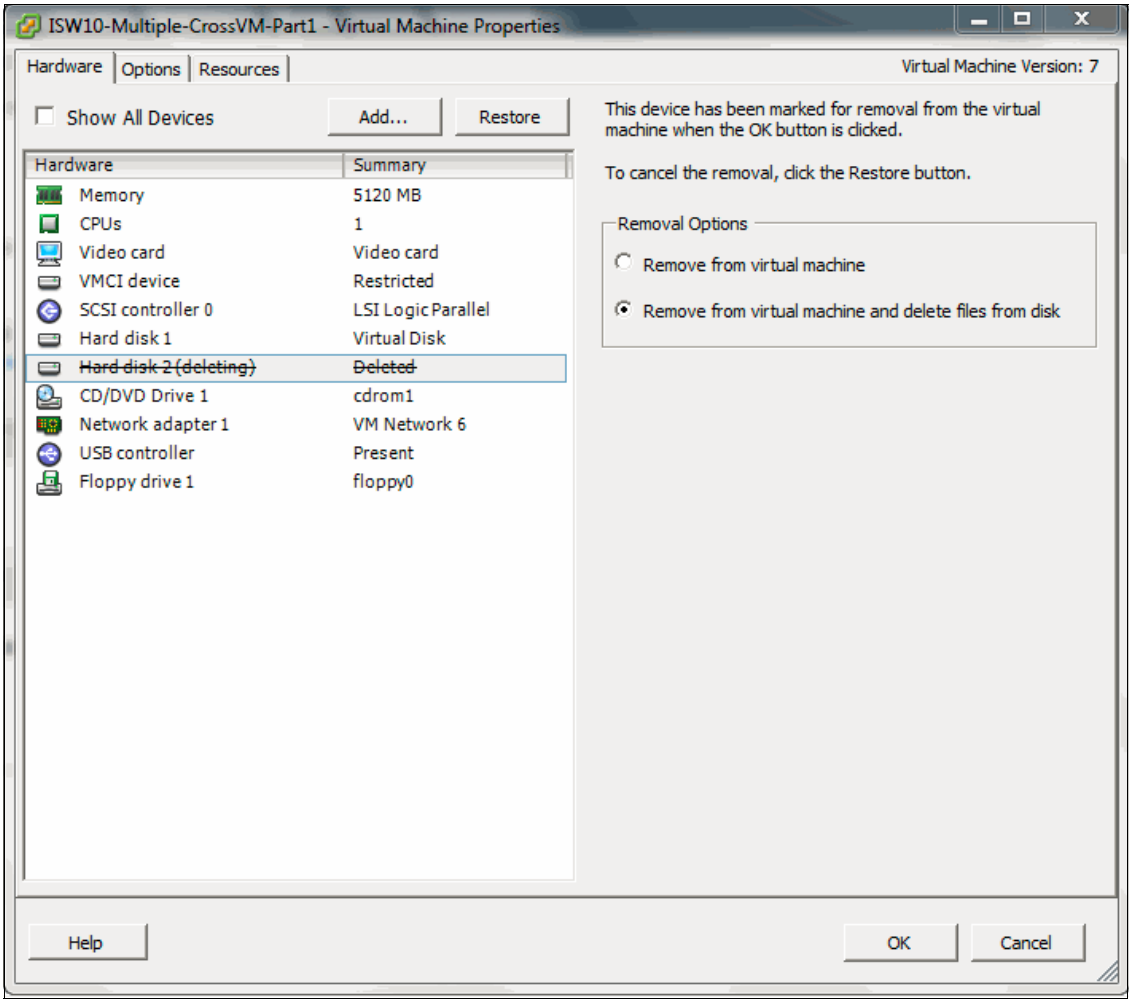


Figure 3-6 Remove the Hard disk 2 from VM and delete files from disk

Add more virtual hard disks

Before adding virtual hard disks, estimate how much usable data is required in GB. Usable data is the amount of space that is available for storing application data, including indexes, aggregates, derivations, and summary data. It also includes raw data space. If you are using logical partitioning within the VM, the virtual disk capacity for each disk comes from the total usable data size divided among all the partitions; see Table 3-1 on page 112.

To add a new hard disk with the user data size that you require:

1. Right-click the deployed VM through your vSphere client and select the item **Edit Settings** link from the Commands section.
2. Click **Add** to open the Add Hardware wizard (Figure 3-7 on page 118).
3. Select **Hard Disk** from the Device Type section as the desired hardware component to add to the VM. Click **Next** to continue.
4. A virtual disk is composed of one or more files on the host file system. Select **Create a new virtual disk** option as the type of disk to use. Click **Next** to continue.
5. Specify the virtual disk capacity as the space of usable data you need. In a production environment, ensure the **Allocate and commit space on demand (Thin Provisioning)** option is cleared.

   Tip: In the Location section, specify an external data store to locate the operating system hard disk.

6. Specify any advanced options for the virtual disk. Normally these options do not need to be changed. Keep the default option and click **Next** to continue.
7. Review the summary of selected options, and click **Finish** to complete the addition of the new hard disk.
8. Repeat these steps to add more disks for logical partitioning.

   Note: Do not add more than eight disks to a single VM because the current version of vSphere Server (at the time of this writing) does not allow more than eight vCPU allocation per VM.

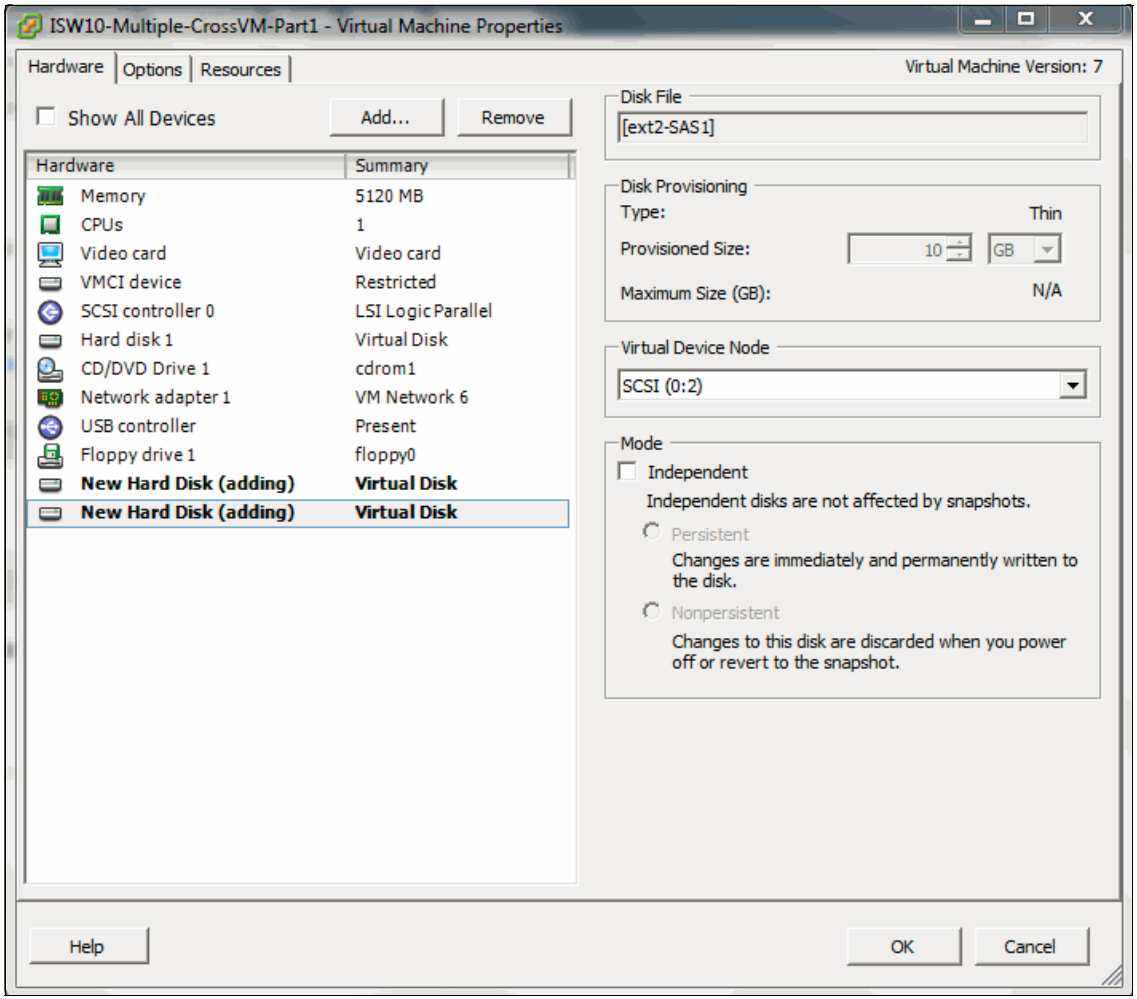


Figure 3-7 Add a new hard disk in the VM

Increase the virtual memory

To add more virtual memory:

1. Right-click the deployed VM through your vSphere client, and select the item **Edit Settings** link from the Commands section.
2. Select the **Memory** component available and increase the memory size as shown in Figure 3-8 according to the configuration chart in Table 3-1 on page 112. Click **OK** to complete the increase of memory.

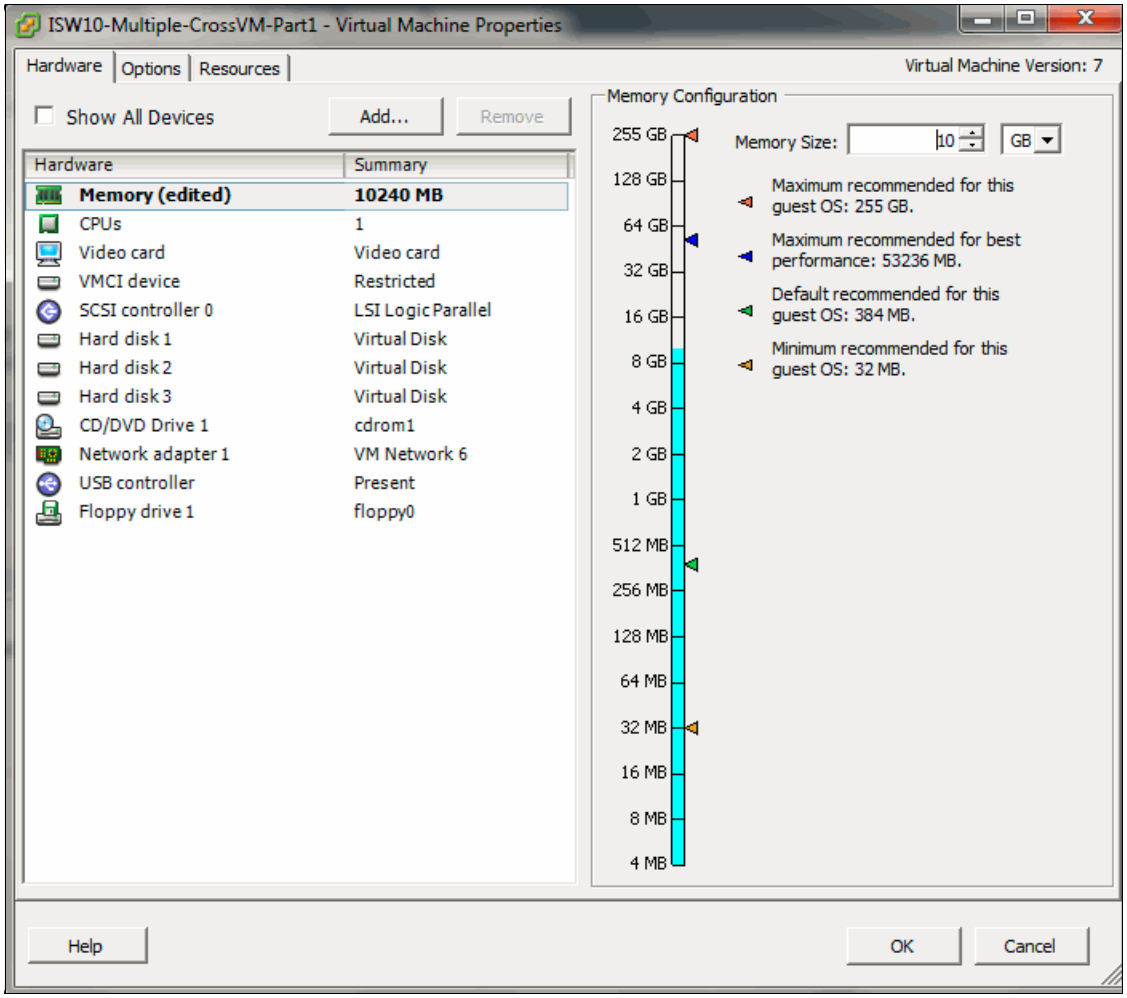


Figure 3-8 Increase the memory if required

Increase the number of vCPUs

To add more vCPUs:

1. Right-click the deployed VM through your vSphere client, and select the item **Edit Settings** link from the Commands section.
2. Select the CPUs component available and increase the number of CPUs as shown in Figure 3-9 according to the configuration chart in Table 3-1 on page 112. Click **OK** to accept the increase of CPUs.

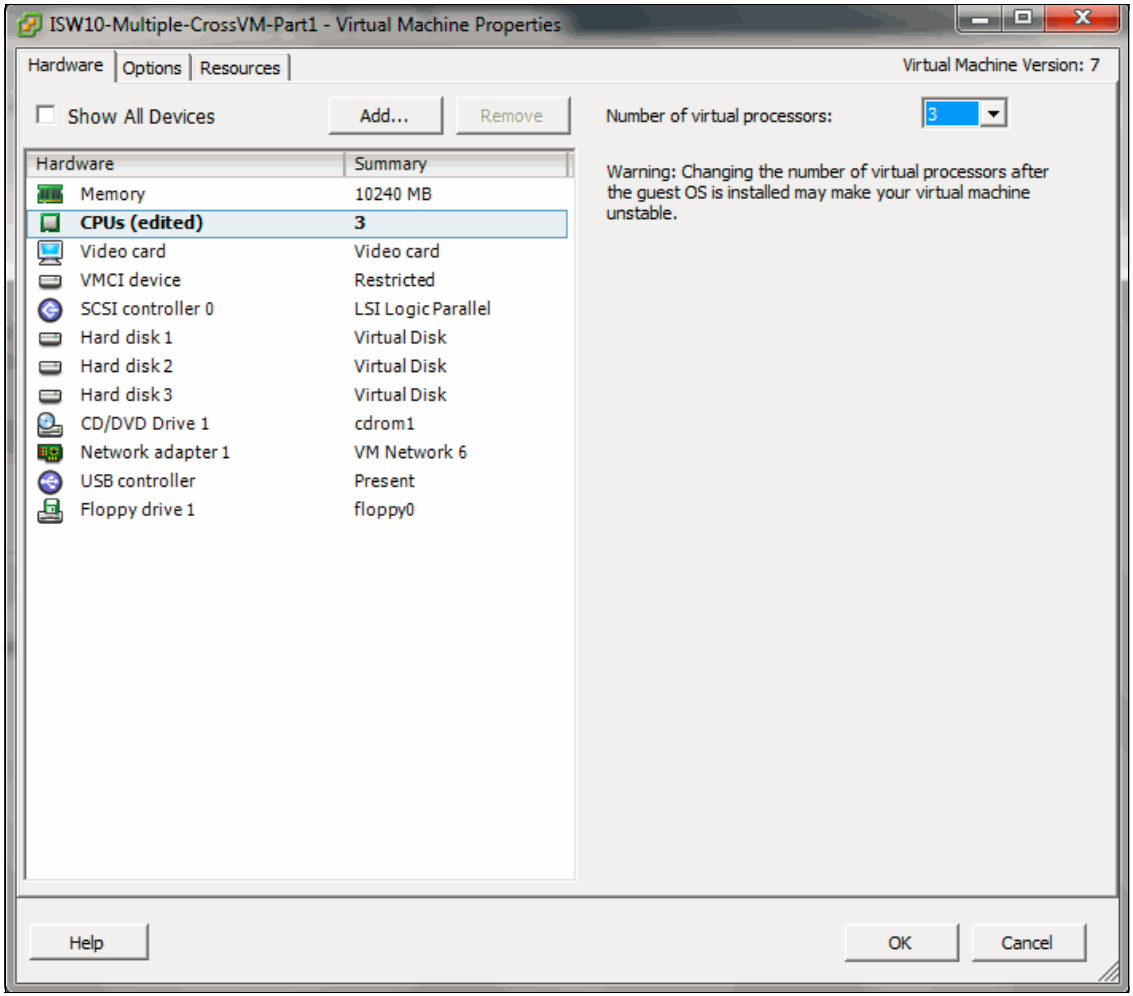


Figure 3-9 Increase the vCPU of the VM if required

Set up and configure the first VM

For deployment of data partitioning on a multi-VM topology, always power on and configure the first VM before the subsequent VMs:

1. Power on the first VM in the multi-node setup, and select the **IBM InfoSphere Warehouse Departmental Edition [VMX]** option from the boot menu.
2. Wait until your machine finishes booting and the Welcome window displays. Click **Next** to continue.
3. Accept the **SUSE Linux Enterprise license agreement** option to continue. If you do not accept the license agreement, the VM halts. Click **Next** to continue.
4. Accept the **IBM InfoSphere Warehouse Departmental Edition virtual image** license agreement. If you do not accept the agreement, the VM will halt. Click **Next** to continue.
5. Change the host name of your VM in the following configuration window. Configure a static IP address for this VM, because the host name and IP address is recorded in the `/etc/hosts` file. Clear the **Change hostname via DHCP** option. The next configuration window allows you to set the static IP. Click **Next** to continue.
6. The next configuration window allows you to configure the network card in your VM. It is preset to use DHCP; use the CHANGE button to choose a static IP address and subnet mask. Click **Next** to continue.
7. Change the time zone of your machine; use the CHANGE button to choose the preferred configuration.
8. The next window prompts you to choose a password for the root user. Choose a secure password, and click **Next** to continue.
9. The next window allows you to choose a multi-node setup, as shown in Figure 3-10 on page 122. Choose one of the following options:
   - If you are setting up a single VM, with or without logical database partitions, choose the **Single VM deployment setup** option.
   - If you are setting up a multi VM configuration, where this is the first VM in this setup, choose the **First VM in a multiple VM deployment setup** option.
   - If you are setting up a multi VM configuration and this happens to be a subsequent (second, third, and so on) VM, select the **Subsequent VM in a multiple VM deployment setup** option.

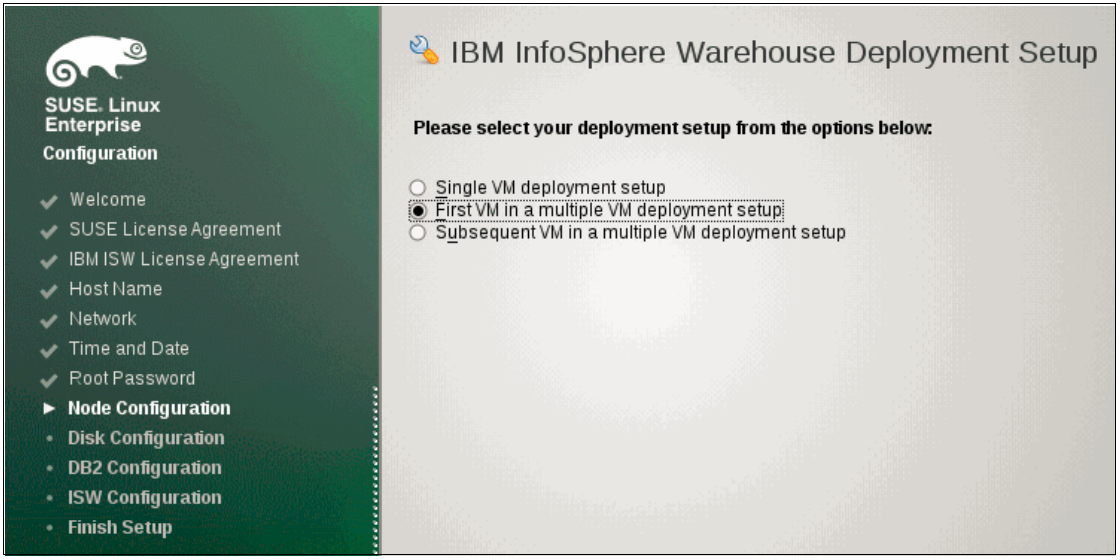


Figure 3-10 Deployment setup option for the InfoSphere Warehouse virtual offering

10. The next window allows you to choose the number of database partitions. Your choice here depends on whether you have previously added empty virtual hard disks to your VM.

    For example, if you have four data disks added, all four of these data disks are proposed to be used as database partitions on this machine. You can reduce the number of partitions by clearing the option in the “Available disks – Dedicated mount points” section, but you must always use at least one disk per partition.

    The upper portion of the window displays warnings and information messages if your configuration is running below minimal requirements, as listed in Table 3-1 on page 112. In such cases, abort the configuration and satisfy the warning messages. You can, however, choose to ignore the warnings and continue. Click **Next** to continue.

11. The next window configures the raw disks and creates a single partition on each. It creates an ext3 file system on each disk. After the file systems are created, the DB2-specific configuration uses them as mount points for the number of partitions that you decide to use. You can also create a 2 GB swap file for any emergency situations, such as when the machine runs out of memory. Wait for the configuration to complete.

12. When the tasks are complete, a summary of the status of the execution displays. Assuming all items report a status of OK, click **Next** to continue.

13. The next window prompts you to choose passwords for the DB2 instance user, DB2 fenced user, and DB2 administration server users. You can change the default user and group values by clearing the **Use default values** option. Click **Next** to continue.

14. The next window shows the progress of the DB2 configuration. The DB2 system users and instance users are created and the previously created file systems are used as database storage for the multi-partitioned DB2 instance. The database storage paths are in the form of `/db2fs/%INSTANCE%/NODE0000` to `/db2fs/%INSTANCE%/NODE00##`, where `%INSTANCE%` is the DB2 instance name you choose in the previous window and the number of `NODE000#` directories depends on the number of hard disks that you add to be used as database partitions.

    The InfoSphere Warehouse metadata database DSSDB is created during the DB2 configuration process and the DB2 database manager is started. The DSSDB database only is stored on the first database partition.

15. The next window shows a summary of the DB2 configuration. All tasks report a status of OK, as shown in Figure 3-11. Click **Next** to continue.

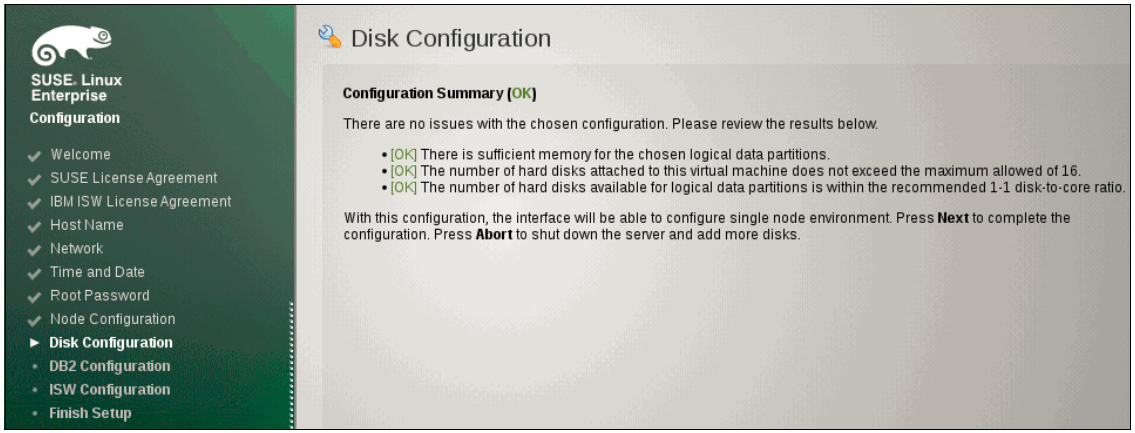


Figure 3-11 Disk configuration summary

16. The final input required is to choose the InfoSphere Warehouse and WebSphere Administration user. The user name defaults to *wasadmin* and the group defaults to *dweagrp* (Administrators group for InfoSphere Warehouse). You can change the default user and group values by clearing the **Use defaults values** option. Remember your choices because you need this user name to log in to the InfoSphere Warehouse and WebSphere Administration Consoles. Click **Next** to continue.

17. The next window creates the chosen administration users and performs a final configuration of the InfoSphere Warehouse components, including WebSphere and DB2.
18. When the configuration is complete, a summary of the configurations performed is displayed. Click **Next** to finalize all your configurations.
19. The login window of the system displays. Log in to your first VM as root by using the password that you set up during the configuration. The login is in runlevel 3.
20. Execute the **`ifconfig`** command to retrieve the network cards settings. Look into the eth0 configuration and write down the IP address.
21. Also, you can always type the **`startx`** command to start the Gnome Desktop Manager.

Set up and configure a subsequent VM

If you need to set up a subsequent VM in this configuration, you must retrieve the IP address of the first VM to have the subsequent VM make an SSH connection. To set up and configure a subsequent VM:

1. Power on the subsequent VM in the multi-node setup and select the **IBM InfoSphere Warehouse Departmental Edition [VMX]** option from the boot menu.
2. Continue with the basic setup of the VM. When you reach the Node Configuration window, select the **Subsequent VM in a multi VM deployment setup** option as shown in Figure 3-12 on page 125. After you select the option, two text fields are displayed where you can enter the IP address and root password of the first VM in this multi-node setup. Click **Next**, and the software sets up the SSH connection between the VMs.

Figure 3-12 Selecting the "Subsequent VM in a multiple VM deployment setup" option

3. When the SSH connection is verified, you are taken automatically to the next step, which shows you the found disks to be used as database partitions. The disks continue the NODE00XX count from the first VM.

   For example, if the first VM has two database partitions (NODE0000 – NODE0001), the second VM (with two data disks created) will have two database partitions mounted on NODE0002 through NODE0003. Click **Next** to continue the configuration.

4. Wait until the configuration completes and review the summary. All items should report a status of OK. Click **Next** to continue the configuration.

5. The DB2 configuration, which is the next task, starts automatically. Because the second VM mounts the first VM's instance user home directory (using the network file system protocol, allowing for shared storage), you do not need to specify user names or passwords and wait for the configuration to complete. The DB2 instance and DB2 fenced user names and passwords are inherited automatically in the second VM. When the process finishes, a summary is displayed. All items report a status of OK. Click **Next** to continue.

6. After you click **Next**, the InfoSphere Warehouse configuration takes place automatically. You are not prompted for a WebSphere Application Server or InfoSphere Warehouse Administrator user because these fields were set up in the first VM. Therefore, when accessing the warehousing and application server admin consoles, you access them using the first VM's host name. Wait for the configuration to complete and review the summary. Click **Next** to complete the full configuration of the second VM.

3.3 Implementing IBM Cognos Business Intelligence

The basic components in the IBM Cognos Business Intelligence installation are:

- Content store
- Content Manager
- Application server
- Gateway server

To provide a scalable and robust environment for cloud computing, the server components can be installed on separate VMs using a distributed configuration. Use a transfer specification file (`.ats`) to install IBM Cognos Business Intelligence without being prompted for user input. If you do not use the regular installation wizard to generate an `.ats` file, you can use the `.ats` file that is available on the installation disk.

In a distributed environment, the sequence in which you configure computers is important. Use the following sequence:

1. Create the content store database.
2. Install, configure, and start the active Content Manager.
3. Install and configure the standby Content Manager.
4. Install and configure the application tier component.
5. Configure the gateway computer so that cryptographic keys are shared and secure communication can take place among the components. The server specified for the External Dispatcher URI property on the gateway computer must be the last server that you start.

3.3.1 Installing IBM Cognos Business Intelligence 10 silently

To perform a silent installation of IBM Cognos Business Intelligence on a Linux guest:

1. Create a Cognos Content Store. If you are using DB2 server, Example 3-18 illustrates how to create a Cognos Content Store with the DB2 commands.

Example 3-18 Commands to create a Cognos Content Store in DB2

```
DB2START;
CREATE DATABASE "csdb001" automatic storage yes on /csdb001data
dbpath on /home/csuser001;

CHANGE DATABASE "csdb001" COMMENT WITH 'Cognos 10 Content Store for
Client 1';
```

```
CONNECT TO "csdb001";
UPDATE DATABASE CONFIGURATION USING APPLHEAPSZ 1024 DEFERRED;
UPDATE DATABASE CONFIGURATION USING LOCKTIMEOUT 240 DEFERRED;
CONNECT RESET;
CONNECT TO "csdb001";
CREATE BUFFERPOOL "csdb001_08KBP" IMMEDIATE SIZE 1000 PAGESIZE 8K;
CREATE BUFFERPOOL "csdb001_32KBP" IMMEDIATE SIZE 1000 PAGESIZE 32K;
CONNECT RESET;
CONNECT TO "csdb001";
CREATE SYSTEM TEMPORARY TABLESPACE "TEMP_TSP_csdb001" IN DATABASE
PARTITION GROUP IBMTEMPGROUP PAGESIZE 32K MANAGED BY SYSTEM USING
('CNT_SYS_csdb001') BUFFERPOOL "csdb001_32KBP";

CREATE USER TEMPORARY TABLESPACE "USR_TSP_csdb001" IN DATABASE
PARTITION GROUP IBMDEFAULTGROUP PAGESIZE 8K MANAGED BY SYSTEM USING
('/tempspace/db2cs001') BUFFERPOOL "csdb001_08KBP";

CREATE REGULAR TABLESPACE "REG_TSP_csdb001" IN DATABASE PARTITION
GROUP IBMDEFAULTGROUP PAGESIZE 8K MANAGED BY SYSTEM USING
('CNT_REG_csdb001') BUFFERPOOL "csdb001_08KBP";

ALTER TABLESPACE "REG_TSP_csdb001" DROPPED TABLE RECOVERY ON;
CONNECT RESET;
CONNECT TO "csdb001";
CREATE SCHEMA db2cs001 AUTHORIZATION db2cs001;
COMMENT ON SCHEMA db2cs001 IS 'Cognos 10 Content Store';
GRANT CREATETAB,BINDADD,CONNECT,IMPLICIT_SCHEMA ON DATABASE  TO USER
csuser001;

GRANT CREATEIN,DROPIN,ALTERIN ON SCHEMA db2cs001 TO USER csuser001
WITH GRANT OPTION;
GRANT USE OF TABLESPACE "USR_TSP_csdb001" TO csuser001;
CONNECT RESET;
DB2STOP;
```

2. On the VM, insert the appropriate disk or copy the contents of the installation files.

3. The .ats file is ts-BISRVR-yyyymmdd_hhmm.ats. Copy it where you plan to install the IBM Cognos Business Intelligence server component. Example 3-19 is a sample .ats file that installs all IBM Cognos Business Intelligence components.

Example 3-19 IBM Cognos Business Intelligence sample .ats file

```
; Licensed Materials - Property of IBM
; BI and PM: is
; (C) Copyright IBM Corp. 2004, 2011
; US Government Users Restricted Rights - Use, duplication or
disclosure restricted by GSA ADP Schedule Contract with IBM Corp
[Dialog1]
Title=Welcome to the Installation Wizard
DE=0
EN=1
FR=0
JA=0
ES=0
NL=0
SV=0
FI=0
IT=0
PT_BR=0
KO=0
ZH_CN=0
ZH_TW=0
[Dialog2]
Title=IBM License Agreement
;Failure to accept the license agreement and non-IBM terms, if any,
will abort the installation.
;Please return the software to the point of acquisition and obtain a
refund, if applicable.
I Agree=y
[Dialog3]
Title=Non IBM License Agreement
;Failure to accept the license agreement and non-IBM terms, if any,
will abort the installation.
;Please return the software to the point of acquisition and obtain a
refund, if applicable.
I Agree=y
[Dialog4]
Title=Installation Location
APPDIR=/opt/ibm/cognos/c10_64
BACKUP=0
```

```
[Component List]
C8BISRVR_APP=1
C8BISRVR_APPLICATION_TIER=1
C8BISRVR_GATEWAY=1
C8BISRVR_CONTENT_MANAGER=1
C8BISRVR_CONTENT_DATABASE=0
```

4. To perform a silent installation, edit the response file to meet your installation requirements. You can edit the following information:
   - License Agreement

     The product license must be accepted to continue with the installation. Change the value for the “I Agree” variable to `y` to accept the IBM License Agreement.
   - Installation Location

     You can change the installation path in the APPDIR variable. By default, the installation path for IBM Cognos Business Intelligence 64 bits on Linux is the `/opt/ibm/cognos/c10_64` directory.
   - Component List

     You can customize the components that are to be installed by specifying a boolean value, 1 or 0. Example 3-20 identifies the features that you can add or remove from the installation program.

     Example 3-20 Edit the features list variable to add or remove features

     ```
     [Component List]
     C8BISRVR_APP=1
     C8BISRVR_APPLICATION_TIER=1
     C8BISRVR_GATEWAY=1
     C8BISRVR_CONTENT_MANAGER=1
     C8BISRVR_CONTENT_DATABASE=0
     ```

5. Install IBM Cognos Business Intelligence.

   On UNIX or Linux systems, change to the directory where you copied the contents of the disk. In the directory for your guest operating system, type the following command, where *`location`* is the directory where you copied the transfer specification file and *`filename`*`.ats` is the name of the file:

   `./issetup -s` *`location/filename`*`.ats`

6. If a return status other than zero (0) is returned, check the log files for error messages. Errors are recorded in the `c10_location/instlog` directory in a summary error log file.

3.3.2 Installing Cognos Framework Manager

You can install the IBM Cognos Business Intelligence modeling tool, Framework Manager, either on a virtual desktop or in the BI developer personal computer. This component requires a Windows operating system. To install Framework Manager:

1. Insert the Cognos Framework Manager disk or download the installation files from the Passport Advantage website. For more information about how to download the product, see:

   http://www.ibm.com/support/docview.wss?uid=swg24029297

2. If you download the installation files, extract them into a directory. Go to the directory where you extract the files, and double-click the `issetup.exe` file.
3. Follow the directions in the installation wizard to copy the required files to your computer.

   Directory names: Be sure to name the installation using only ASCII characters. Some servers do not support non-ASCII characters in directory names. For example, installing Cognos Framework Manager in a directory that has an apostrophe in the path name might result in the help not opening properly.

4. When notified with an information message about installing the Supplementary Languages Documentation, click **OK**.

   The Supplementary Languages Documentation contains translated versions of Guidelines for Modeling Metadata. If you want this guide available in languages in addition to English, install the contents of the disk. Otherwise your installation includes only the English version of the guide.

5. On the last page of the installation wizard, select **Start IBM Cognos Configuration** to configure Cognos Framework Manager immediately. You can choose to configure Cognos Framework Manager later by starting the configuration tool from the Windows Start menu shortcut in the IBM Cognos program folder.
6. Secure the installation directory from unauthorized access.

You must configure the modeling tool to communicate with the application tier components using one of these routes:

- Connect to the application tier components dispatcher using the Internal dispatcher URI. This route is the preferred route.
- Connect to an additional, dedicated gateway that is configured to connect to the dispatcher using the Internal dispatcher URI. You must configure

appropriate security for this gateway. This method is useful when the modeling tool is outside a network firewall.

Important: Do not change your main gateway to use the internal dispatcher URI. Doing so reduces the security of the IBM Cognos Business Intelligence portal and studios.

To ensure that the modeling tool can communicate with IBM Cognos Business Intelligence components on the computer where the tool is installed, configure cryptographic properties and the following environment properties:

- Gateway URI
- Dispatcher URI for external applications

To model DB2 data sources with Cognos Framework Manager, you must install and configure the DB2 Client on each Cognos Framework Manager computer. This installation allows for direct access from each system to the DB2 data source.

3.4 Integrating additional software

For the main middleware components to operate smoothly together, supporting software must be implemented. For example, if a Cognos server needs to retrieve data from a database, a name server service must be implemented for server address lookup. This name server is one of many software services that are essential to the operation of the overall BI cloud solution.

This section describes technical features and issues for software services that can often be overlooked but that should be considered when building a cloud environment.

3.4.1 Web server

Whether a cloud deployment is public or private, virtualized or bare metal, delivery of cloud services must be done through an interface that a client can access, such as a graphical user interface (GUI) or a command-line interface (CLI). This interface provides the solution's clients with capabilities, such as provisioning services and tracking usage, billing, and accounting.

Usually, this interface can be delivered through a web browser or a web server. Although you can choose from a variety of web servers to host a cloud solution, this section discusses generic web server software (such as Apache Web

Server) and provides details about the important features and functions, and how they should be tuned and configured.

Expected load on a web server

When configuring the web server, consider the amount of load or traffic a web server is expected or capable of handling. Will a single web server be enough to handle all the incoming requests? If not, consider multiple web servers. Based on these considerations, configure the web server as necessary, and deploy as many web servers as needed to handle the expected load.

When dealing with a large variance of web traffic, virtualization can be an ideal solution to meet scaling demands. During peak hours, more resources can be pooled to power a greater number of web servers. After hours, these resources can be freed to perform maintenance operations.

Load balancing

A proxy server that redirects incoming traffic can be implemented to optimize the web server configuration, as depicted in Figure 3-13. This configuration ensures that no single web server is a bottleneck and that resource use is optimized.

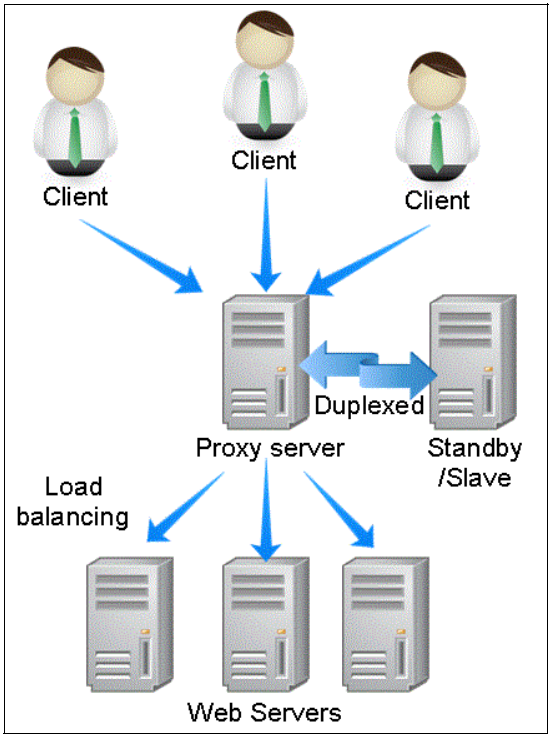


Figure 3-13 Load balancing web servers

Availability

Ensure that there is always a minimum of two web servers to handle requests. These web servers do not need to synchronize or interact with each other, but any reliable cloud service will require available web servers running at all times. In addition to having at least two virtualized web servers up and running, consider having the servers hosted on separate physical server nodes to mitigate a hardware failure scenario.

Secure Sockets Layer

Secure Sockets Layer (SSL) is a security protocol that provides communication security over the Internet. This service is essential for clients that are accessing BI cloud services to prevent unauthorized access. Clients are authenticated by providing a private X.509 certificate that matches the public certificate that the server references.

3.4.2 Domain Name System server

A Domain Name System (DNS) server provides a means of communication between servers within the cloud infrastructure. DNS servers are usually implemented behind a solution's firewall to prevent unauthorized outside access. A centralized DNS server provides a directory of addresses for each server, thereby allowing other servers to focus on their specific tasks rather than fumbling with their own localized directory of addresses.

High availability (HA) is a requirement for a DNS server in any cloud deployment. If the only DNS server goes down, servers within the infrastructure will have no way of communicating with each other. Therefore, a DNS server should be duplexed in a primary and standby, active and passive, or master and slave configuration. HA for DNS servers can usually be provided by the DNS server software itself, or by using third-party HA software, such as LinuxHA.

Because host names and addresses are added and removed throughout operation, periodic syncing and mirroring between the active and passive servers is suggested. Furthermore, consider hosting the DNS primary and standby servers on separate physical nodes to mitigate a physical server failure.

3.4.3 Time server

Time servers are a requirement of any IT solution that is composed of multiple servers. A network of multiple servers can have their time settings off, where each server has slightly different time settings. Queries or operations performed on one server to another can yield errors or failures as a result. Therefore, a centralized Network Time Protocol (NTP) server must be implemented.

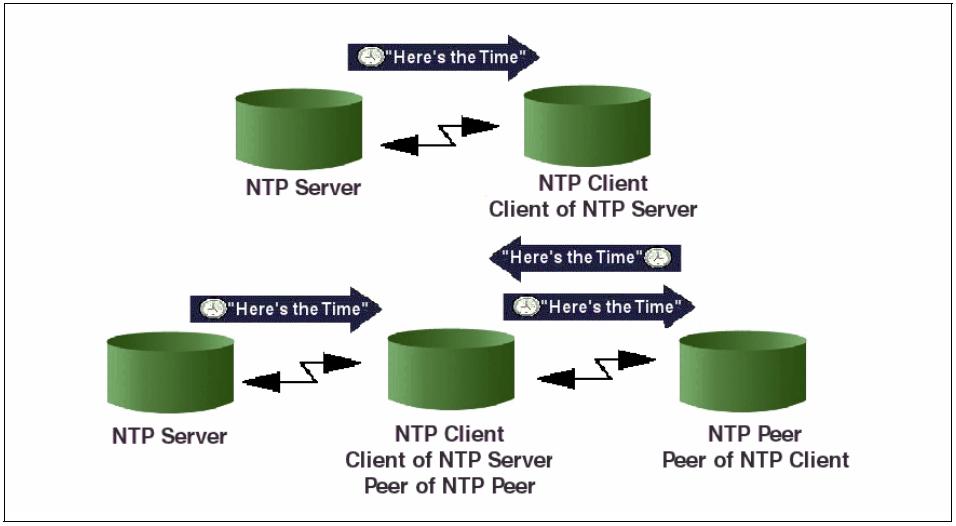


Figure 3-14 NTP server, client, and peer interaction

One option is to have the NTP server be accessible by all other servers within the network to achieve consistency with time across all infrastructure. Another option is to set up the NTP server to have client/peer communication so that the NTP server does not need to communicate directly with each server in the network. Instead, the NTP server communicates the time to its clients, and the clients communicate the time to their peers, as depicted in Figure 3-14.

High availability (HA) of an NTP server can be achieved by having an additional redundant NTP server. This can act as a client that receives time information from the main server. In the event of a failure on the main server, the client can take over as the new NTP server, and start communicating with other clients.

3.4.4 Lightweight Directory Access Protocol

Lightweight Directory Access Protocol (LDAP) provides an accessible directory for managing and verifying security credentials of different network-oriented middleware components. This provides a layer of security for crucial middleware components such as the InfoSphere Warehouse and Cognos layers. It prevents non-authorized clients from accessing sensitive data or services. Figure 3-15 depicts a simplified LDAP topology with multiple middleware components accessing the LDAP directory.

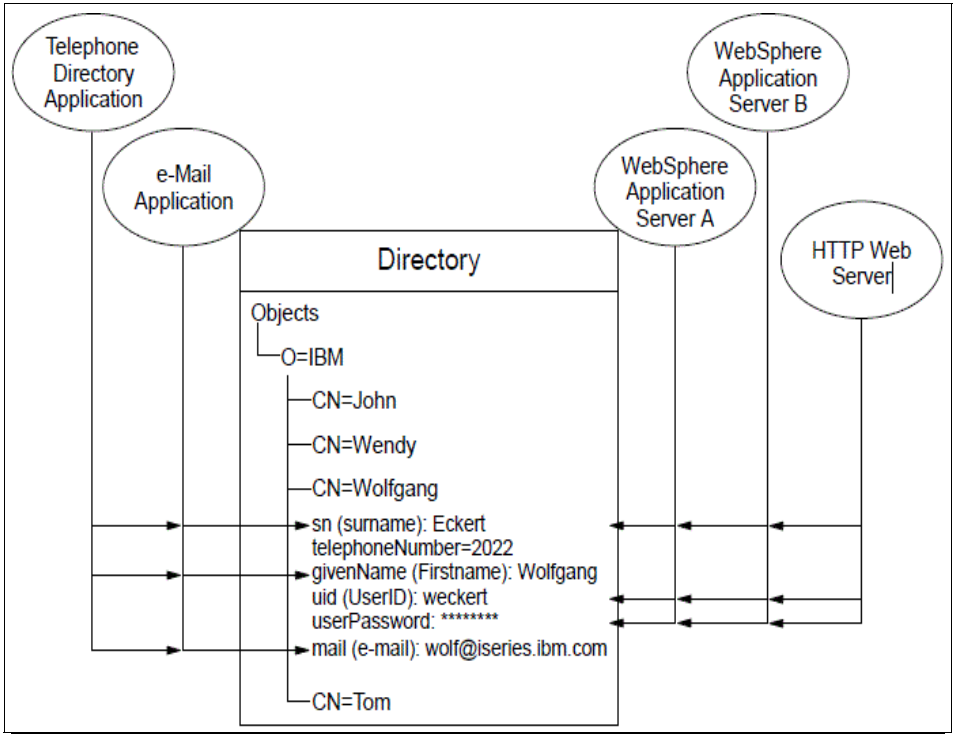


Figure 3-15 LDAP implementation example

Similar to the time server and DNS, duplex or mirror the LDAP server to mitigate the risk of a server failure, whether it be virtual or physical. Further mitigation from hardware failure is possible by keeping each of the primary or standby servers on different physical nodes.

3.4.5 Support software for providing high availability and resiliency

Although some IBM middleware such as DB2 (and InfoSphere Warehouse) come with HA capability, many of the software services listed in this section do not have native HA features. Therefore, these features might have to be sourced from third-party software. Although virtualization and cloud computing resources provide HA from a server-level standpoint, HA can also be provided at the *service* level.

The advantages of additionally providing HA at the service level are quicker recovery time, more fault tolerant services, and reducing strain on the virtualized environment. Server resources at the virtualization level do not need to be

redeployed because the underlying services were “healed.” Consider software that is capable of providing both HA and automated failover capabilities.

High availability and automated failover for services

Software, such as Tivoli System Automation for Multiplatforms and LinuxHA, can provide HA on just about any service or daemon running within the network.

Within any operating system are services and daemons that are running. Automation software can query these services to check whether the status of each service is running, stopped, or if there are problems. LDAP, NTP, DNS, web server, database, and Cognos are all examples of services. You can configure this automation software to perform a variety of operations if a failure occurs on the services that its designed to maintain.

Consider the following possible configurations for HA:

- Active/cold standby
  - One server handles active connections. The other server is “sleeping” as standby and is ready to take over as the active server in a failover scenario.
  - This configuration provides the lowest manageability requirements and the lowest resource efficiency.
- Active/hot standby (master/slave)
  - One server handles active connections (master). The other server has a slave process of its paired active server.
  - This configuration provides moderate manageability requirements and low resource efficiency.
- Active/active (master/master)
  - An active/active configuration load balances servers within the configuration. When a server fails, the remaining healthy server or servers continue processing connections. This HA configuration usually requires a homogenous software configuration and higher manageability requirements, but also yields the highest resource efficiency because all servers are being used. This configuration provides the highest manageability requirements and the best resource efficiency.

From these HA configuration choices, determine the most appropriate method for deploying your HA clustered software components. Automation software has many configurable parameters for fine-tuning HA and failover. By implementing a sound design for your HA software, you can ensure that your solution’s underlying services are running healthy, thereby allowing your solution to serve its customers unhindered.

Distributed Replicated Block Device (DRBD) is a distributed storage system for the Linux platform, as shown in Figure 3-16. It provides HA for storage (block devices) in either master/slave or master/master configurations. DRBD involves replication between two or more nodes.

DRBD is ideal for integrating shared storage file system in software services that use master/slave or active/passive HA configuration such as DNS and LDAP. This is due to the fact that both master and slave can access the same file system in the event of a failure. DRBD has mechanisms to ensure that the data stored on the DRBD device is consistent and accurate.

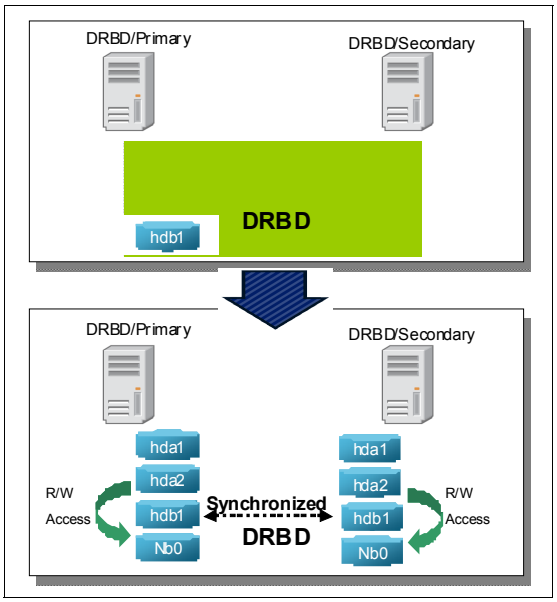


Figure 3-16 DRBD implementation

Example high availability components implementation

Consider the HA solution depicted in Figure 3-17 on page 138. This solution has the following configuration:

- Services
  - LDAP, DNS, HA Proxy (web server load balancer), and Clustered (virtual) IP services are handled by a generic automation software and DRBD configuration. Whenever a service is detected to be not running or unhealthy, the service on the standby server is brought up to resume work.
  - Services and the Cognos gateway reside in the same physical machine as an active/standby pair. The active server contains all active services and

active Cognos gateway. The standby server contains the same services as the active server, but on standby, though Cognos Gateway on this server is also active (so that the web server load balancer can feed incoming clients to either the active or standby machines).

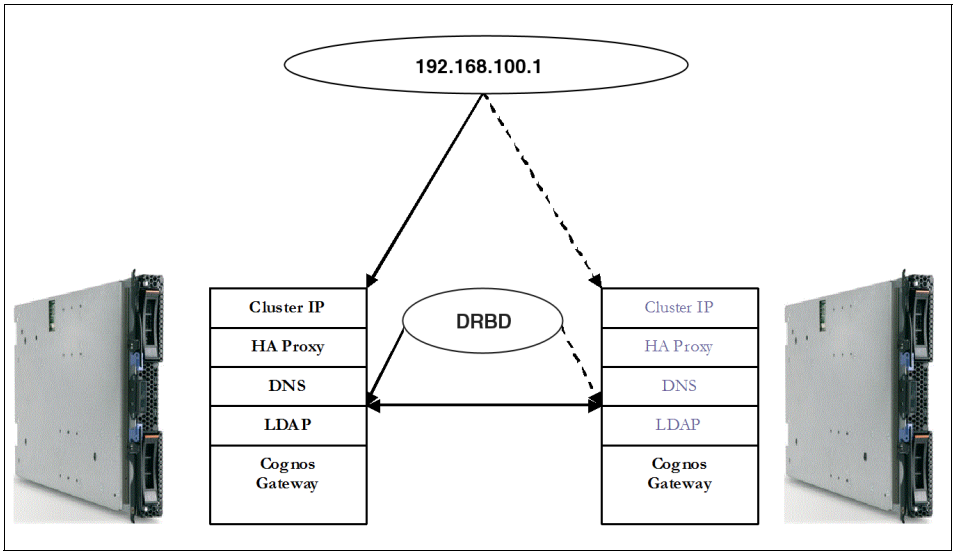


Figure 3-17 High availability implementation

– Services directories and configuration files can change due to ongoing configuration and addition of clients and servers. These files and directories are stored on a DRBD device that can be read by only the current active server. Setup of the DRBD device is as follows:

  - Create partitions on both primary and standby.
  - Create logical volumes (size determined by service space requirements) on both primary and standby.
  - Create an DRBD device that encompass the logical volumes on both primary and standby.
  - Determine which service files and directories must be placed on the DRBD device and place them there.
  - Set up DRBD as primary/standby.

Failure scenarios

These HA configurations can help mitigate downtime of an overall solution by providing automated failover capabilities. The automated failover capabilities kick in when a service or hardware component malfunctions. The following sample scenarios show how the HA software reacts when certain failures occur:

- Failure Scenario 1: Service failure

  If a particular service is stopped unexpectedly and cannot start again, the automation software fails over to the standby server. The DRBD device is made readable on the new active server.

  All services are started on the standby server, and the cluster IP now points to the new active server. All requests are redirected to the standby server. HA Proxy can still direct Cognos traffic to the old active server because the private network is still available. The Cognos gateway is still running.

- Failure Scenario 2: Cognos Gateway failure

  The automation software does not do anything. The only change is that HA Proxy stops directing Cognos traffic to the standby server's Cognos gateway.

- Failure Scenario 3: Network failure on active server

  A split-brain scenario is possible if there is no network quorum. *Split-brain* occurs when communication between the primary and standby is terminated, making both the primary and the standby servers think each server should take over operations.

  To avoid split-brain scenarios, a network quorum must be achievable on both servers. A network quorum should be an address in the network that both servers can ping and that is reliable. If the automation software on a particular server cannot achieve quorum, all services are shut down. The standby server still has quorum, can detect that the active server is unreachable, and can take over accordingly.

- Failure Scenario 4: Power failure on active server

  The automation software detects a time out on the active server and makes the standby the new active server. The DRBD device and all services are made active on the new active server.

4

Management fundamentals and preferred practices

There are many areas where management is a vital activity to maintain the overall performance and health of data warehouse and business intelligence services. This chapter includes the following topics:

- Provisioning and automation
- Multi-tenancy
- Maintenance and monitoring

4.1 Provisioning and automation

To meet your need for rapid deployment of business intelligence (BI) services with minimal effort, use virtualized BI software to provide provisioning and automation. Can you start new virtual resources and services and expect them to be utilized and ready for production in a few minutes?

4.1.1 Architecture patterns

One of the main reasons why virtualization is interesting to decision makers is the benefit of lowering costs. Virtualization can provide multiple BI environments without the overhead of operating systems, software, and services deployment costs. The main force behind provisioning depends on the ability to standardize and automate installation and configuration tasks. These tasks can be cloned and deployed into multiple virtual machines. Therefore, owners of public clouds can control the infrastructure costs by provisioning a large number of ready-to-go business intelligence applications in a matter of minutes. Another cost benefit is the reduction in maintenance tasks, because cloud services can be provided from a self-service catalog.

In InfoSphere Warehouse and Cognos implementations, many of the early tasks can easily be automated by creating deployment scripts inside virtual templates. Because business intelligence is considered a robust environment with many layers (such as an application server, a web server, and a database server), it is important to create a diverse set of deployment scripts that can satisfy different architectures, from a single virtual machine to multiple virtual machines. In this way, you can provide a mechanism to support some level of customization and mixed-workload environments.

After you develop and test the first virtual machines with deployment scripts, these can be stored passively into a image library server. When they are located on the image library server, they consume little or no resources until they are awakened by a cloud consumer's requests. You can create architecture patterns or topologies that can be easily deployed as virtual machines. You can deploy them on a single server topology or a multiple server topology.

InfoSphere Warehouse and Cognos architectures

InfoSphere Warehouse and Cognos architectures can be categorized in architecture patterns as illustrated in Figure 4-1.

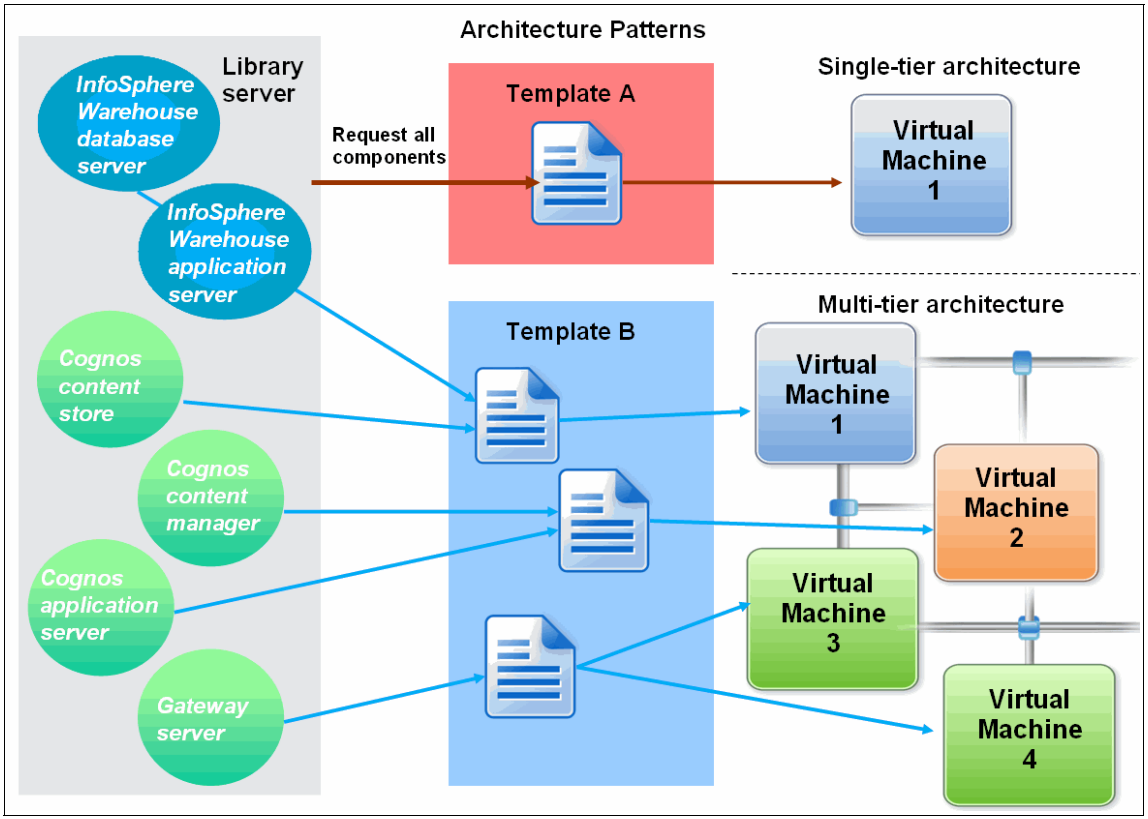


Figure 4-1 InfoSphere Warehouse and Cognos architecture patterns

In a multi-tier architecture, pay close attention to the dependency of components and network communications between them. For example, Cognos implementations rely on the discovery of components to match cryptography keys and to start services. In addition, InfoSphere Warehouse Administration Console relies on the database server tier to function.

In a multi-tier architecture, the deployment of multiple machines of the same type is possible as well. For example, it might be possible to deploy multiple database server nodes of an InfoSphere Warehouse database server with the database partitioning feature by setting an architecture pattern (ISW_DB_DPF = 8) to deploy eight data nodes. The use of an advanced network configuration ensures that the architectures for each cloud consumer are isolated from all others.

IBM Workload Deployer

IBM Workload Deployer (previously know as IBM CloudBurst®) is a hardware appliance that provides access to IBM middleware virtual images and patterns. These patterns allow repeated secure deployment of virtualized application environments that can be managed in a private cloud.

Figure 4-2 shows the lifecycle of deploying images on IBM Workload Deployer.

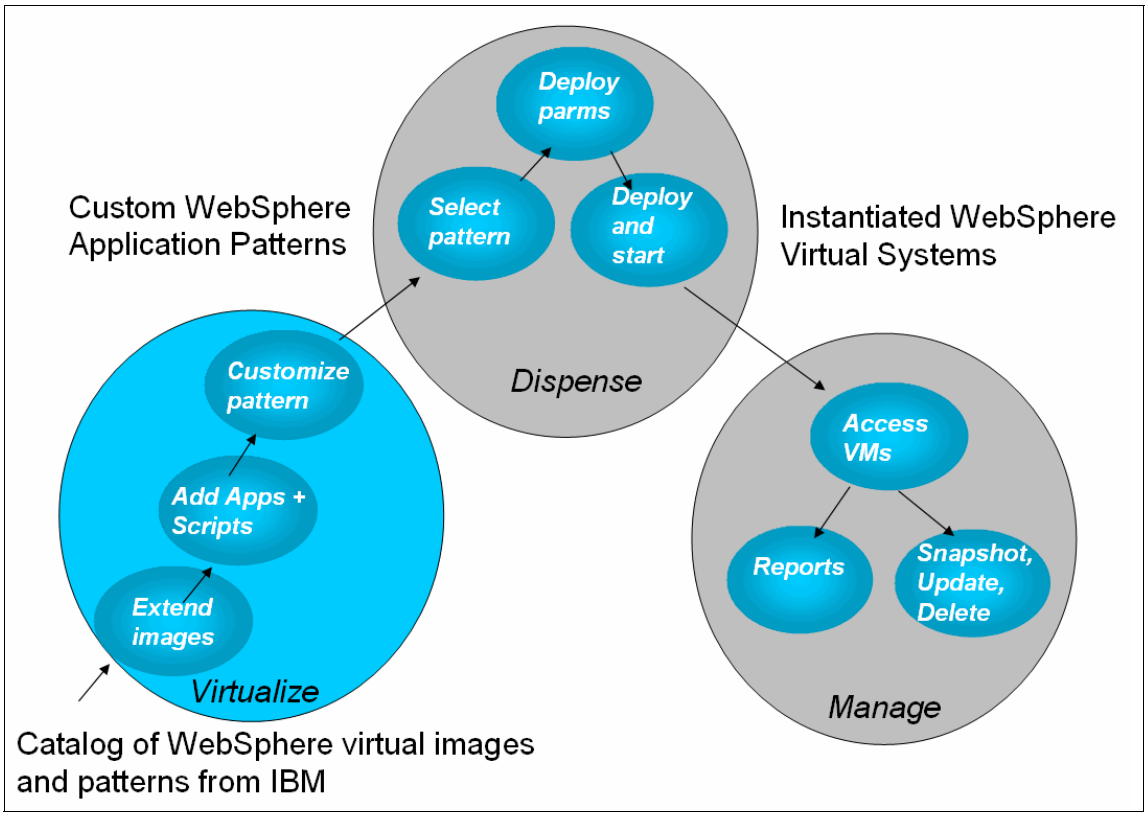


Figure 4-2 IBM Workload Deployer lifecycle of deploying images

For more information, refer to *Virtualization with IBM Workload Deployer: Designing and Deploying Virtual Systems*, SG24-7967.

4.1.2 Order of the startup operations

A repeatable procedure for running and starting a solution in the cloud environment can ensure that specific services and software function properly. To incorporate a repeatable procedure, complete the following operations and tasks in specific software layers:

1. Assuming the overall solution uses virtualization, start and enable the virtualization software first, before running any other software components. A properly implemented type 1 hypervisor (bare metal or native) runs directly on the host hardware. Thus, when the host hardware is turned on, the virtualization layer is activated.
2. If the supporting software components, such as the DNS server, LDAP, and web servers are virtualized, start these components next. Middleware components, such as Cognos and InfoSphere Warehouse, rely on addressing directory services to communicate with each other. Also, enable any high availability (HA) or resiliency software or storage devices now. Often, when systems automation software is enabled for managing these types of support services, the start of these services is automated according to the HA implementation.
3. Ensure that all the necessary database instances are started before connecting to a database. Make sure that all databases are in a consistent state. Also, enable any HA mechanisms because it is difficult to enable these features during online operation.

   If you need to start a particular instance in InfoSphere Warehouse database server, log in as the DB2 instance owner and run the **db2start** command. This command starts the current database manager instance background processes on a single database partition, or on all the database partitions defined in a partitioned database environment.
4. Start WebSphere Application Server with the InfoSphere Warehouse metadata database already running. On UNIX and Linux systems, start WebSphere Application Server with the **./startServer.sh server1** command.
5. Start the Cognos server components last. Ensure that the Cognos content manager has access to the Cognos store database. On UNIX and Linux systems, start the IBM Cognos Business Intelligence server with the **./cogconfig.sh -s** command.

4.1.3 Tasks requiring automation

Task automation is an integral part of any cloud computing solution. An effective cloud solution is designed to be self-serving, meaning that customers do not

need to interact with intermediaries to gain direct access to services and software. Benefits of implementing a highly automated cloud solution include reduced labor costs, higher reliability, quicker and more agile deployment, and increased ease of implementing new automated features down the road.

A wide variety of configuration options exist for every software component within a BI cloud solution. You need to template the most common methods to automate deployment for specific services, such as LDAP, firewall, DNS, server, and so on. As always, interaction between the virtualized software servers and the hypervisors is necessary to automate deployment of standby or passive servers. A possible solution is to deploy a centralized management or monitoring server, which ensures that specific services are running. Refer to “Monitoring” on page 150.

The InfoSphere Warehouse componentized environment provides database automation within a cloud environment. Its DB2 component has several features that can sense and respond to changes in the environment. These features can adjust the system to optimize its operation, such as self-configuring, self-optimizing, and self-managing. For more information about implementing autonomics, consult *DB2 Virtualization*, SG24-7805.

4.2 Multi-tenancy

Multi-tenancy enables a pool of sharing resources to allocate many tenants and absorb workloads in peak times. The term *tenant* can refer to a user, a department, or an organization. Multi-tenancy allows better use of computing resources and reduces the costs of the overall infrastructure. It becomes especially important when the cloud environment needs to support multiple users (tenants), as the previously mentioned attributes become more complex to maintain.

Each client can have an isolated environment, even when the underlying infrastructure is shared. This isolation environment ensures acceptable levels of security, performance, and accurate metering. Virtualization is a good way to share and isolate the underlying infrastructure (hardware, storage, and networking). However, it is also possible to use multi-tenancy at the platform level, where the tenant can share the same database or application server software. The configuration of a multi-tenant environment relies on an agreement between the cloud provider and the cloud consumer. Most of the time, the cloud consumer is not bothered with specific details about where its database is stored, as long as the database is available and SLAs are understood by both parties.

The deeper the level of multi-tenancy, the more efficient the implementation and the lower the cost, but at the price of flexibility and performance. You can

customize the InfoSphere Warehouse database environment to be multi-tenant at different levels, such as table, schema, database, instance, and system. For example, two database server installations can be hosted within the same virtual or physical server.

4.2.1 Database instances

Any particular InfoSphere Warehouse data server can host multiple *database instances*, as shown in Figure 4-3. Within each database instance, a database manager configuration file dictates the attributes and behavior of the instance. You can automate this database configuration file to ensure consistency in reproducible data server environments and to add a degree of flexibility and customization to the virtualized data server image. You can automate database instances during the startup sequence of the virtualized server.

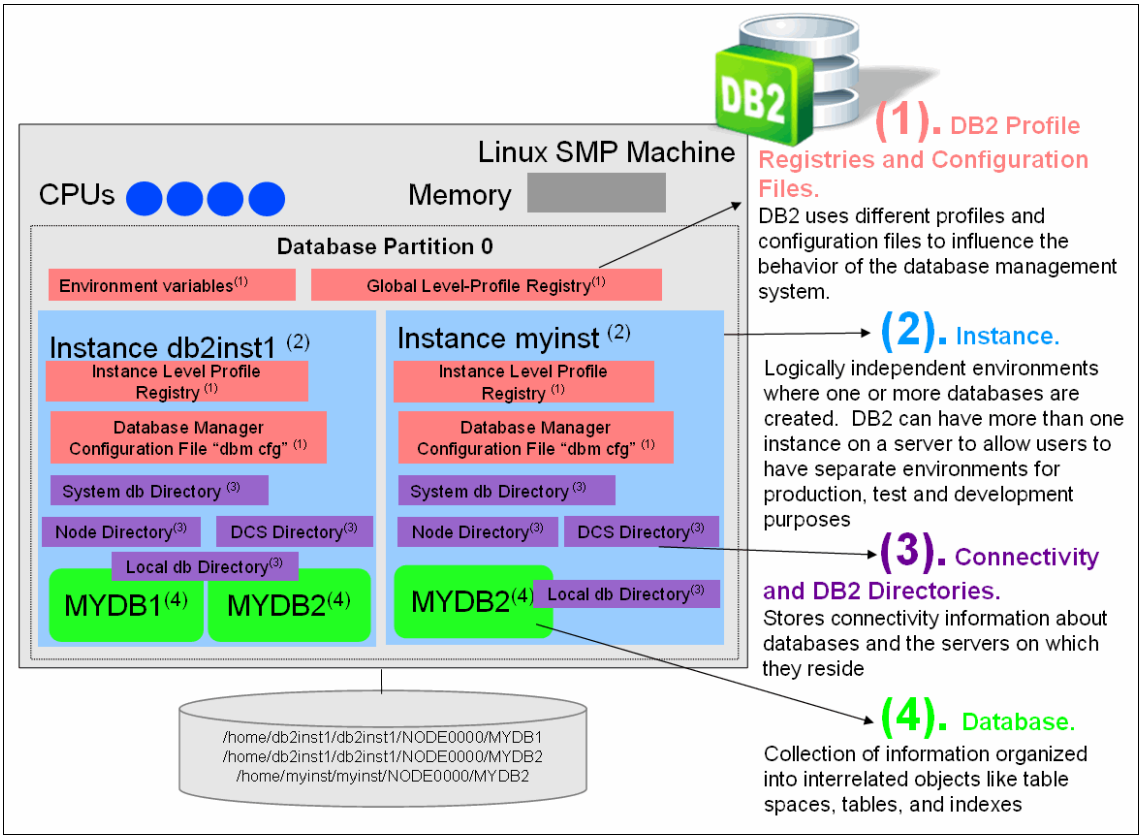


Figure 4-3 DB2 database server environment

4.2.2 Databases

A database instance can include multiple *databases*. Each of these databases contains its own database *configuration file* (not to be confused with a database manager configuration file). A database contains a set of objects that are used to store, manage, and access data according to the relational model of the data. DB2 completes the following tasks for you:

- Set up the system catalog tables for the database.
- Allocate the database recovery log.
- Create the database configuration file and set the default values.
- Bind the database utilities to the database.

4.3 Maintenance and monitoring

Maintenance and monitoring are crucial to the healthy, ongoing operation of a cloud environment. This section describes preferred maintenance and monitoring practices for various middleware and software components that are part of a BI cloud solution.

4.3.1 Maintenance

Tasks that fall under maintenance include performing backups, health checks, reporting, and any other tasks that are required to sustain the ongoing health of the overall solution. Maintenance tasks often require downtime on a particular piece of software, therefore taking the overall BI cloud solution offline. As a result, choosing an appropriate time for performing routine maintenance is crucial to decrease the disruptiveness that maintenance can cause. When is the cloud environment expecting the most traffic? When does it receive the least traffic? Ideally, maintenance should be scheduled during the period of lowest activity for the solution.

Periodically complete the maintenance operations that are described in the following sections on crucial middleware in any BI cloud solution.

Database

You need to complete maintenance tasks such as backing up data. Determine the level of database archiving that is required and at what times the data must be backed up. What level of backup is needed? Is a full backup required, or are incremental or delta backups sufficient?

Incremental and delta backups: An *incremental* backup image is a copy of all database data that has changed since the most recent, successful, full backup operation. A *delta* or *incremental delta* backup image is a copy of all the database data that has changed since the last successful backup (full, incremental, or delta) of the table space in question.

Incremental backups provide reduction in the storage of backup files but require pre-existing backups. Full backups offer whole database backups, but they require the maximum amount of storage. Mix and match the different types of backups according to the requirements of your solution.

Virtualized environment

Consider periodic analysis of the virtualized environment, including the essential software services that power the BI cloud solution. Virtualized environments often use database software to keep track of environment changes. Back up this database periodically to provide recovery ability in the event of a disaster. Determine the level of database archiving that is required and at what times the data should be backed up.

You can also back up entire VMs. Because a virtualized environment database keeps track of only the administration and configuration of an environment, it does not store the underlying VMs that run within the environment. These virtualized environments power essential services, such as LDAP, NTP, web services, and so forth, that can sometimes be intricately set up or configured. The ability to back up VMs is essential for saving (and templating) these configurations for future use or recovery purposes.

Are these software services, such as the LDAP, NTP, DNS server, and web servers, working properly? Do they need to be load tested or have their logs been pruned? Establish a set of acceptable performance benchmarks for each software service, and take periodic measurements to ensure that the services are running optimally.

Recovery

In any production environment, failures inevitably occur, whether at a virtual or physical level. How these failures are handled and how swiftly they are dealt with should be the main focus.

Consider the level of severity of the disaster. For example, if one database server goes down, another database server can be re-instantiated using the database server template. The database can be recovered using a database backup and then rolled forward to a peer-state with the primary server.

For more serious disaster scenarios, more specific and detailed procedures must be in place to ensure prompt recovery. For example, a database cluster where both database servers go down requires a complete recovery. Bring up a stand-alone primary database as soon as possible and the standby database immediately after that. Do not process any transactions with just the primary database alone, because if another failure occurs, an inconsistency in the data might occur. For every software service within the BI cloud solution, you need a procedure for dealing with complete and partial recovery.

4.3.2 Monitoring

A good monitoring methodology is based on having an understanding of the overall system performance and analyzing monitored data. Pay attention to the number of transactions executed, the behavior of those transactions in terms of buffer pool hit ratios, physical reads and writes, and the time spent sorting and aggregating data. In addition, by monitoring the system, you can detect disk, CPU, network, and memory bottlenecks, which allows you to determine whether an activity is suspended or is just taking a long time.

DB2 monitoring

You can associate the following InfoSphere Warehouse monitoring tasks with examining operational status and overall database performance:

- Table functions to collect DB2 metrics with SQL
- Event monitors to capture related point-in-time information
- Snapshot monitor to determine the status of a database system
- The DB2 problem determination tool, db2pd

The DB2 monitoring framework includes the following perspectives and collection points:

- System

  System aggregations can provide data regarding the total work being done. These aggregations are rolled up and defined along with the components that make up the infrastructure of the Workload Manager (WLM), such as workloads and service classes. System perspective metrics can include the following access points:

MON_GET_UNIT_OF_WORK	Return metrics for one or more units of work.
MON_GET_WORKLOAD	Return metrics for one or more workloads.
MON_GET_CONNECTION	Return metrics for one or more connections.

MON_GET_SERVICE SUBCLASS	Return metrics for one or more service subclasses.

The following query displays connections that return the highest volume of data clients, ordered by rows returned:

```
db2 "SELECT application_handle, rows_returned, tcpip_send_volume,
evmon_wait_time, total_peas, total_connect_request_time FROM
TABLE(MON_GET_CONNECTION(cast(NULL as bigint), -2)) AS t ORDER BY
rows_returned DESC";
```

- Data objects

  Statistics from data objects can provide information about actions taken indirectly in the realm of data objects, such as tablespaces, bufferpools, containers, and so on. Data objects also provide a perspective on the impact of application work on data objects. Data object monitoring includes the following table functions:

MON_GET_BUFFERPOOL	Monitor buffer pool efficiency, hit ratio, activity.
MON_GET_CONTAINER	Monitor container activity, rank, enumerate.
MON_GET_INDEX	Monitor index use, for example the number of index scans and how many scans are index-only scans.
MON_GET_TABLE	Monitor activity on table reads, updates, inserts, and overflow activity. (This data is always collected regardless of the database configuration.)
MON_GET_TABLESPACE	Monitor table space activity (read and writes), and buffer pool activity.

  The following query lists the utilization of container file systems, ordered by highest utilization:

```
SELECT varchar(container_name, 65) as container_name,
SUBSTR(fs_id,1,10) fs_id, fs_used_size, fs_total_size, CASE WHEN
fs_total_size > 0 THEN
DEC(100*(FLOAT(fs_used_size)/FLOAT(fs_total_size)),5,2) ELSE
DEC(-1,5,2) END as utilization FROM TABLE(MON_GET_CONTAINER('',-1))
AS t ORDER BY utilization DESC
```

- Activity

  Activity monitoring consists of statistics associated with SQL statements. The activity monitoring includes the following table functions:

  – MON_GET_PKG_CACHE_STMT

    Return a point-in-time view of both static and dynamic SQL statements in the database package cache.

  – MON_GET_ACTIVITY_DETAILS (XML)

    Return details about an activity, including general activity information (such as statement text) and a set of metrics for the activity. It can also create a query that captures information about all the activities currently running on a system.

  The following SQL statement lists all the dynamic SQL statements from the database package cache, ordered by the average CPU time:

```
db2 "SELECT MEMBER, SECTION_TYPE,
TOTAL_CPU_TIME/NUM_EXEC_WITH_METRICS as AVG_CPU_TIME,
LOCK_WAIT_TIME, SUBSTR(STMT_TEXT,1,40) STMT_TEXT FROM
TABLE(SYSPROC.MON_GET_PKG_CACHE_STMT('D', NULL, NULL, -2)) as T
WHERE T.NUM_EXEC_WITH_METRICS <> 0 ORDER BY AVG_CPU_TIME".
```

Data Studio web console

IBM Data Studio web console, included with DB2, provides health availability monitoring and the means to make database changes. You can use the Data Studio web console to create alerts for health and availability monitoring. You can also use the job manager page to create and manage script-based jobs for databases.

Data Studio web console implements the following monitor features:

- Health summary
- Alert list
- Current application connections
- Data sharing members
- Current utilities
- System log
- Current table spaces

InfoSphere Optim Performance Manager

IBM Optim Performance Manager is a web-based interface that you can use to view system health at any time from any location (Figure 4-4). This tool aids in the prevention of problems by monitoring performance indicators and displaying extensive dashboards and alerts to identify a problem. With this tool, you can identify locking conflicts, deadlocks, problematic applications or queries, or constrained buffer pool cache, as well as heap size, CPU, memory, file system shortages, and data skew conditions.

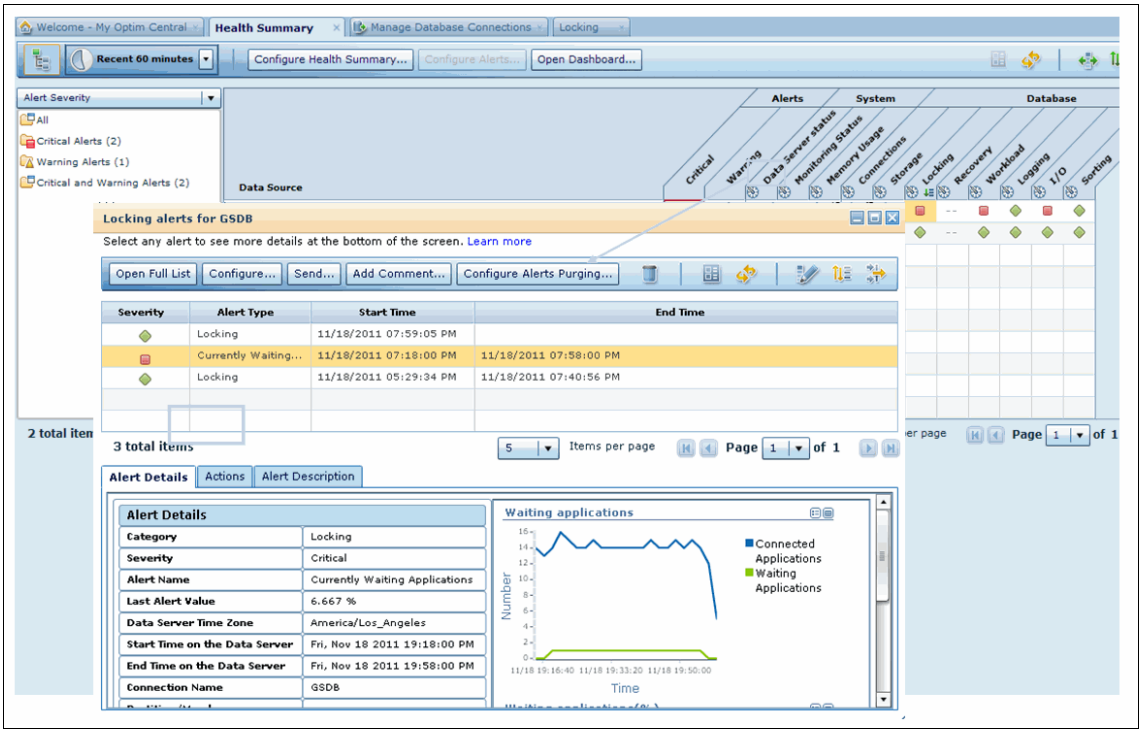


Figure 4-4 IBM Optim Performance Manager web interface

Optim Performance Manager implements the following features:

- You can monitor any DB2 environment, from single-instance SMPs to 100+ partition warehouse deployments 24x7 to ensure that nothing is missed.
- You can use built-in templates for common workloads to get up and running quickly. Templates are available for Cognos, Information Server, or for more generic BI deployments. Templates include default values for which metrics to collect, the collection frequency, and key performance indicator thresholds and alerts.

Rather than just a cross-database view, you can also get an overview for a specific database on the Overview Dashboard, as shown in Figure 4-5.

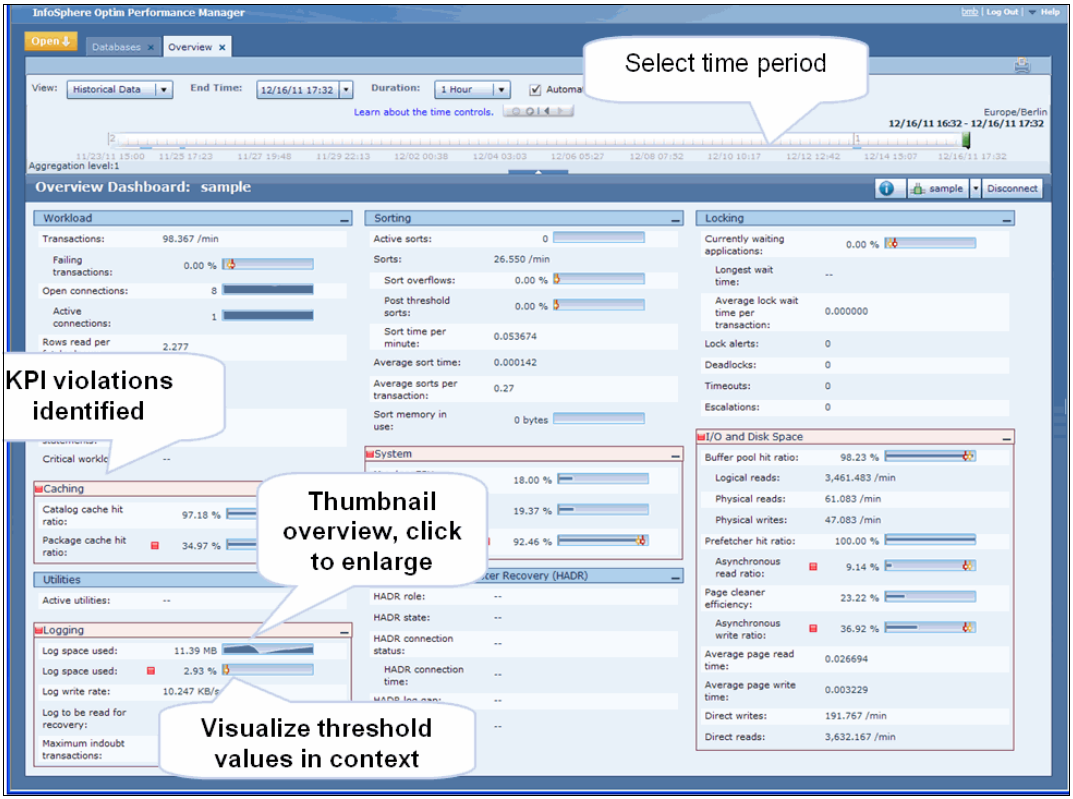


Figure 4-5 Key performance indicators using Optim Performance Manager

You can see key performance indicators for critical database activities, including workload, sorting, locking, logging, I/O, and so on. Key performance indicators (KPIs) can provide information about the activity in each category.

Optim Performance Manager performs metrics aggregations automatically at 15-minute, one-hour, and one-day intervals and supports automated retention managed at each aggregate level. These aggregations make it easier and more cost-effective to save data in the long term. You can go back to stored data for any time period in the repository by adjusting the time slider to the period of interest. Compare current values to prior values to see if things are out of the normal range. Mitigate symptoms first, and defer root-cause analysis with historical data. Widen the viewed interval to hours, days, or weeks to instantly view metric trends on dashboards to spot spikes, emergent problems, or significant changes to the steady state.

For more information, refer to *IBM Optim Performance Manager for DB2 for Linux, UNIX, and Windows*, SG24-7925.

IBM Cognos Business Intelligence performance and monitoring

Over time, an IBM Cognos Business Intelligence environment changes. User populations grow, processing requests tend to increase in number and complexity, and network capacity and other aspects of infrastructure might be modified. These changes can affect performance. As a result, it is important to monitor and tune performance regularly.

Monitoring performance means regularly checking the status of your IBM Cognos Business Intelligence installation and its resources. IBM Cognos Business Intelligence provides metrics for checking the performance of the system, servers, dispatchers, or services. You can set thresholds for metrics to identify when the performance exceeds or falls short of expected ranges. You can configure the system to notify anyone who should be made aware of the problem when a performance issue occurs.

Tuning can involve adjustments in the following areas:

- Databases

  Keep databases optimized for querying and reporting.

- Application servers

  Adjust memory and connection settings as required for better performance.

- Web servers

  Tune the servers for maximum performance.

- IBM Cognos Business Intelligence

  Monitor and tune various aspects of your IBM Cognos system

5

Service delivery model for virtualized business intelligence

The service delivery model is the final piece to consider during setup of a cloud environment. At this stage, you can describe the data management and business intelligence (BI) services that are available in your cloud. Cloud consumers can subscribe to these services and must understand and agree to the policies and responsibilities. Ultimately they should follow your preferred practices to use your BI cloud environment.

This chapter includes the following topics:

- Services delivery lifecycle
- Service delivery model
- Cloud services catalog using InfoSphere Warehouse

5.1 Services delivery lifecycle

Before an enterprise decides to use BI services from your cloud, it needs to understand its requirements and needs for BI capabilities. Involve the cloud system administrator in early planning with the enterprise that is acquiring the cloud services. Early involvement includes providing consulting and opening discussions about possible scenarios during the cloud adoption.

The following process can complete a service delivery lifecycle for BI:

- Business requirements evaluation for BI

  The potential cloud consumer needs to create a detailed description of the BI requirements (type of queries and reports), scope, and stakeholders.

- Assessment

  The potential cloud consumer can access the information about your BI cloud capabilities from your web portal, community FAQ, customer service staff, consulting staff, or service demonstrations. The enterprise needs to have clear expectations of your BI cloud capabilities.

- Planning and design

  The potential cloud consumer can estimate the number of virtual resources needed. This task can be accomplished by evaluating the size of the expected data warehouse and the number of expected cloud consumers (developers and business users) needed.

- Generation of the service level agreement (SLA) and policies contract agreement

  The cloud provider agrees to meet SLAs and service ratings to ensure the cloud consumer receives overall quality service. Quality service also provides enough resources and an underlying disaster recovery strategy for business continuity.

  The cloud provider also agrees to protect the cloud consumer's hosted sensitive data and its identity privacy by applying security boundaries, firewalls, and port monitoring technology.

  Alternatively, the cloud consumer agrees to use the BI services for legal purposes << ?? >>and to protect the user account information provided to access the cloud. Security is a shared responsibility between the cloud consumer and the cloud provider. Both work together to keep the cloud environment secure from malicious access, and respectful of client and enterprise privacy.

- Customer account and subscriptions

  The cloud administrator can open a customer account for every entity (individual or enterprise) accepting the contract. This entity can choose between Software as a Service (SaaS) or Platform as a Service (PaaS). One of these services is needed to subscribe and provision computing capabilities automatically from a self-service catalog. Every entity is responsible to request enough capability and user access to support their user requirements.

- Development and testing on a PaaS model

  When the cloud consumer requests PaaS services, the cloud provisioning software identifies which virtual resources and components are available to be allocated and started inside the cloud. All these components are requested in a batch process to clone virtual machines (VMs), install software, set up the communication links, and start services.

  To deliver the application development and testing services, the cloud must carry out the following activities:

  - Provision the environment to support development and testing services.

    The cloud consumer can start receiving services for BI application development and testing. The cloud creator can also use these types of services to create new applications, and offer a more diverse services catalog on the SaaS model.

  - Provide access credentials to the BI development tools available on your PaaS cloud model.

    When the underlying environment gets provisioned, the cloud consumer receives access credentials to connect to the cloud environment. These credentials can include one or many host names, user names, and passwords.

- Applications on a SaaS model

  When the cloud consumer requests SaaS services, the cloud provisioning software can allocate and start the resources needed to support the applications on a runtime environment. To deliver the application runtime services, the cloud must carry out the following activities:

  - Provision the environment to support the default SaaS applications or migrated applications created from the development environment.

  - Provide access to web tools available to interact with the BI applications hosted on your SaaS cloud model.

- Operation and management

  Centralized management of system resources helps to control the cloud infrastructure and the distribution of services with optimal configurations. Use

performance metrics to monitor whether the cloud performance is meeting the SLAs established in the contract agreement.

To the cloud consumer, the cloud appears to be an unlimited source of computing power. The cloud must scale out quickly to increase capacity on demand so new service requests can always be allocated. When you operate and monitor your cloud, include these standard activities:

- Provide customer support
- Provide clear information in usage reports, billing, and invoices.

  Cloud consumers typically receive a report of resource usage and pay for what they use.
- Cancel services and release resource allocation.

  When cloud consumers cancel their services, the previously allocated resources are released to fulfill other user requests and workloads.

Figure 5-1 illustrates the service delivery lifecycle.

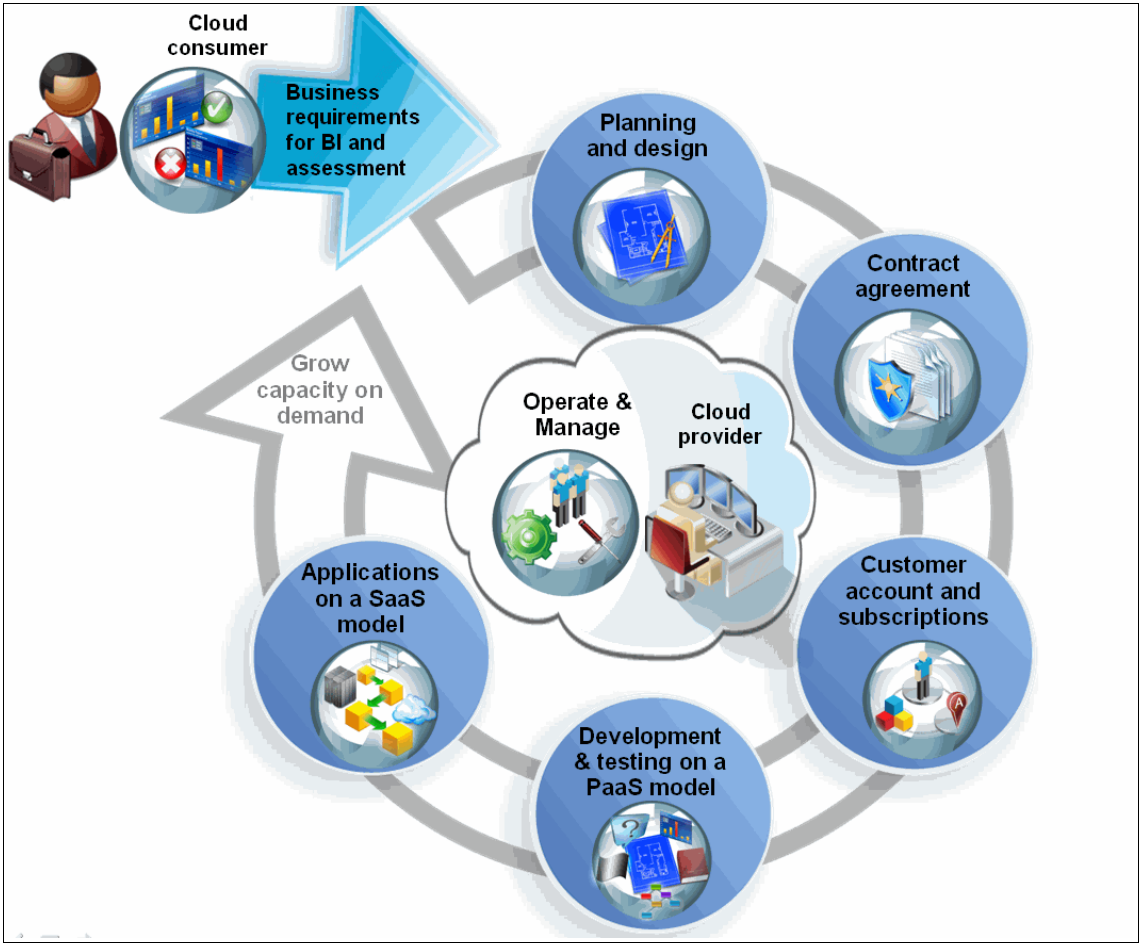


Figure 5-1 Service delivery lifecycle

5.2 Service delivery model

The creation of a service delivery model is an important step in setting up a virtualized business intelligence platform. A business user does not need to be overwhelmed by technical aspects of the supporting infrastructure as long as its capabilities empower the user to improve decision making and business outcomes.

In the cloud, the business intelligence capabilities are standardized and delivered as a consistent service experience. BI services are delivered on demand by Internet protocols, and consumers pay for what they use. The cloud provider designs a service delivery model that can describe the principles, standards, policies, and constraints that guide the operation of the BI cloud environment. This model ensures the cloud consumers can consume efficiently what they need, when they need it.

The data warehouse and BI software that you install and configure on a virtual infrastructure is called Infrastructure as a Service (IaaS). IaaS can be delivered effectively within either the SaaS or PaaS service models, as shown in Figure 5-2.

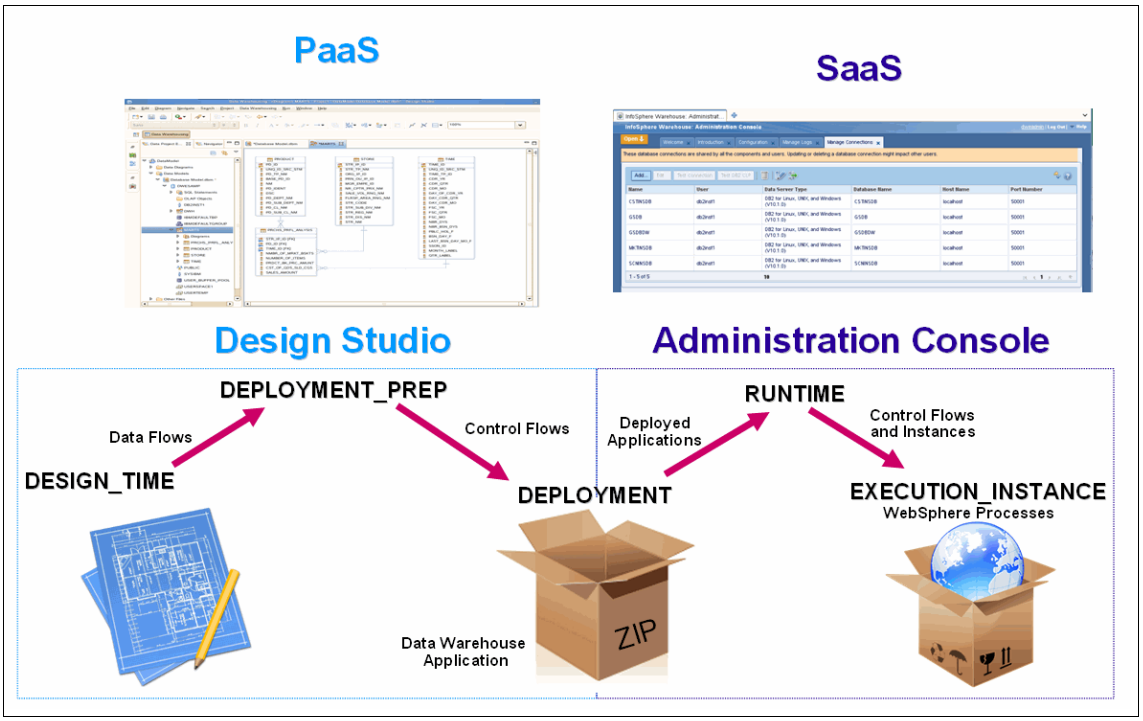


Figure 5-2 PaaS and SaaS service models using InfoSphere Warehouse tools

Depending on the type of model that you offer, you can allow the cloud consumer to have different levels of customization or interaction. The following models are available:

- In a PaaS model, the cloud consumer has access to development and configuration tools that are needed to create and build complex applications that can run on a SaaS model. The developer can do much with the development tools; however, they are not responsible for maintaining the

virtual or hardware infrastructures. Development tools are for high use clients that might need more resources. In this case, you can choose among the following options:

- Install these tools on a virtual desktop that the cloud consumer can access remotely.
- Let the users install the development tools on their computers to access the server components hosted in the cloud. If servers are located on a network that is isolated behind a firewall, cloud providers can provide a secured virtual private network (VPN) to access the isolated network.

You can use the following front-end development tools to deliver data warehousing and BI on a PaaS model:

- Design Studio
- Data Studio
- Cognos Framework Manager

- In a SaaS model, the cloud consumer interacts with the hosted applications and middleware software using a thin client, generally a web browser. SaaS provides the most complete environment for BI, where users are not responsible for maintaining applications, software, and virtual infrastructure health.

  You can use the following front-end web tools to deliver data warehousing and BI on a SaaS model:

  - Administration Console
  - Data Studio Web Console
  - Optim Performance Manager
  - Cognos Connection

You can use a combination of a SaaS and a PaaS model to fulfill different user, developer, and business requirements. When a cloud offers a PaaS model in conjunction with SaaS model, users are more likely to create applications that can be extended and enriched beyond what is available from a SaaS model.

5.2.1 Remote access

When delivering development services inside the cloud, you need to provide remote access connections based on sorting.

Desktop virtualization for cloud

Desktop virtualization has opened up new opportunities for remote access to more complex developments tools. Hosting the development tools in the same network where the middleware software and databases are has benefits.

Performance and security can improve because the data does not have to leave the data center where the cloud infrastructure is hosted.

Desktop virtualization provides the developers a web connection to work whenever they want, from wherever they are geographically located. This technology centralizes and virtualizes the environment for management and control of software upgrades and patching, and standardizing administrative policies.

Secure Shell Service

Secure Shell Service (SSH) is a tool to perform remote login, tunneling, proxy, file transfer, and remote administration on Linux systems. It not only offers an encrypted and secured way to run commands, such as **`scp`** and **`sftp`**, but also has the ability to access a graphical tool using X11 tunneling.

Windows Terminal Services

Windows systems use a different protocol called remote desktop protocol (RDP), to access a Windows graphical desktop remotely. If you are connecting from another Windows environment, you can open a remote desktop connection. If you are connecting from a Linux environment, you can use the **`rdesktop`** and **`tsclient`** tools.

X Window System

X Window System is a system for remote graphical user interfaces that is built upon a client/server model to provide controllable windows across the Internet. A solution available at no cost is Cygwin, but other commercial solutions are available, such as Virtual Network Computing (VNC). VNC transmits the keyboard and mouse movements over a network and across different operating systems through a graphical user interface.

5.2.2 Development tools that are available for a PaaS model

Both cloud service creators and consumers can use the development tools to design, develop, and customize data warehouse and BI applications that are hosted on the cloud.

Design Studio

Design Studio is an integrated development environment (IDE) that provides a customizable design environment. The design environment is built upon the Eclipse platform for InfoSphere Warehouse projects. All Design Studio functionality is supplied by Eclipse plug-ins and supports features enabling cloud consumers to perform these activities:

- Create database source connections
- Design physical database models
- Generate DDL to create and update physical database objects
- Design, test, and debug SQL Warehousing data flows and control flows
- Create warehousing applications for deployment
- Create cube models and cubes for online analytical processing (OLAP)
- Design data mining flows

Figure 5-3 illustrates the Design Studio framework.

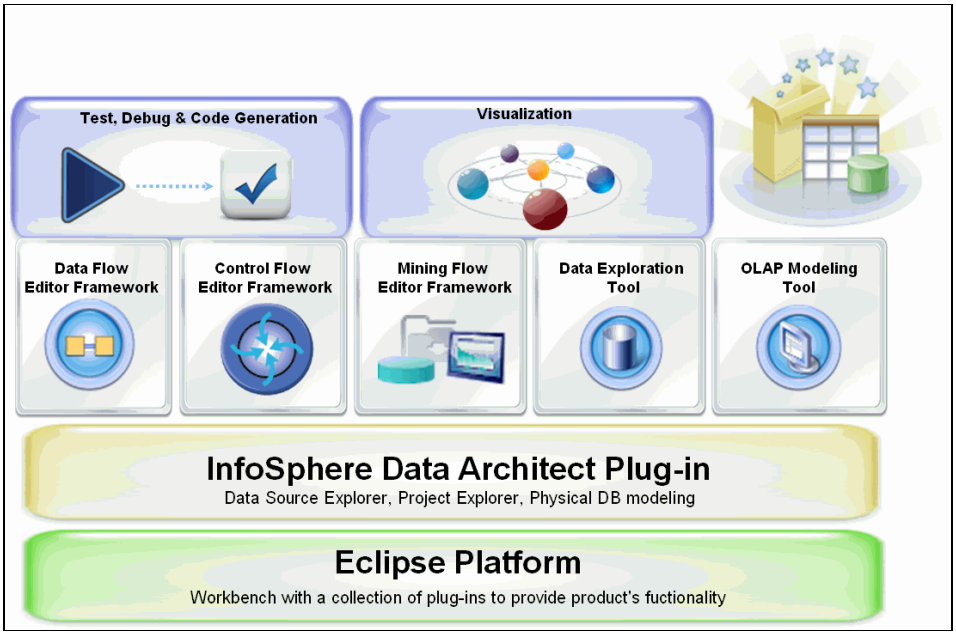


Figure 5-3 Design Studio framework

The Data Warehousing perspective is the default perspective in Design Studio. It contains views to support editing and navigation during the development of data design and data warehousing projects, as shown in Figure 5-4.

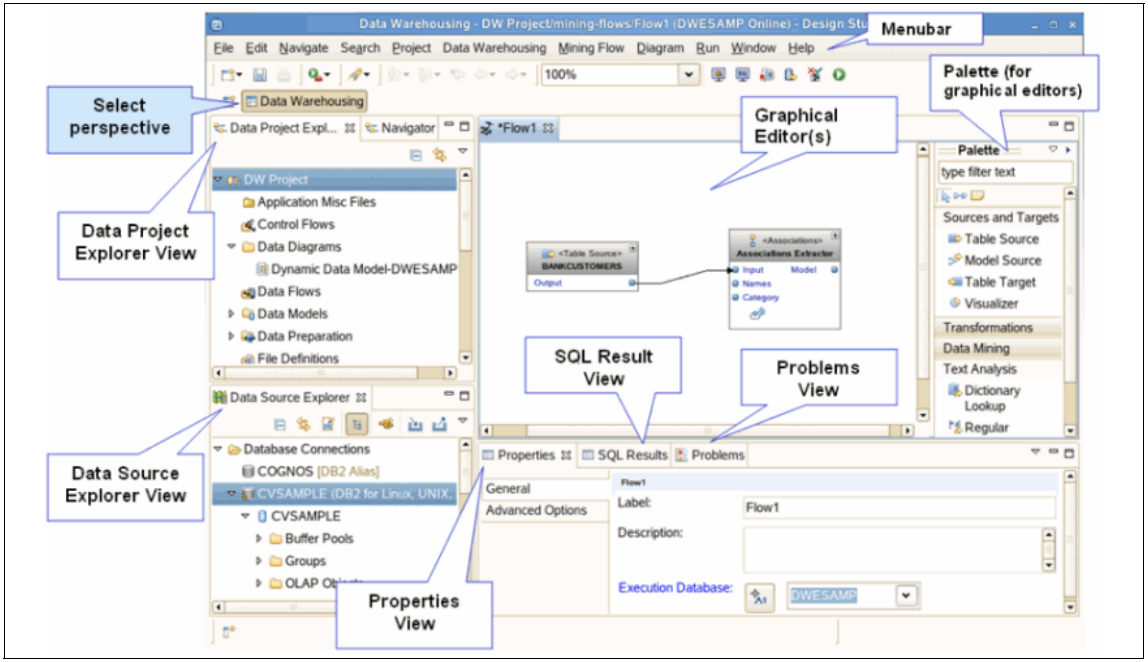


Figure 5-4 Data Warehousing perspective views

The Data Warehousing perspective includes the following views:

- Data Source Explorer
- Data Project Explorer
- Properties
- SQL results
- Problems
- Execution status
- Graphical editor

Cloud consumers can open, close, move, restore, maximize, and minimize the Data Warehousing perspective views. The following sections describe the capabilities for each view.

Data Source Explorer view

The Data Source Explorer is a tree view in Design Studio. From this view, you can manage connections from different relational databases and their database objects. Database objects are schemas, tables, views, and so on, as shown in Figure 5-5.

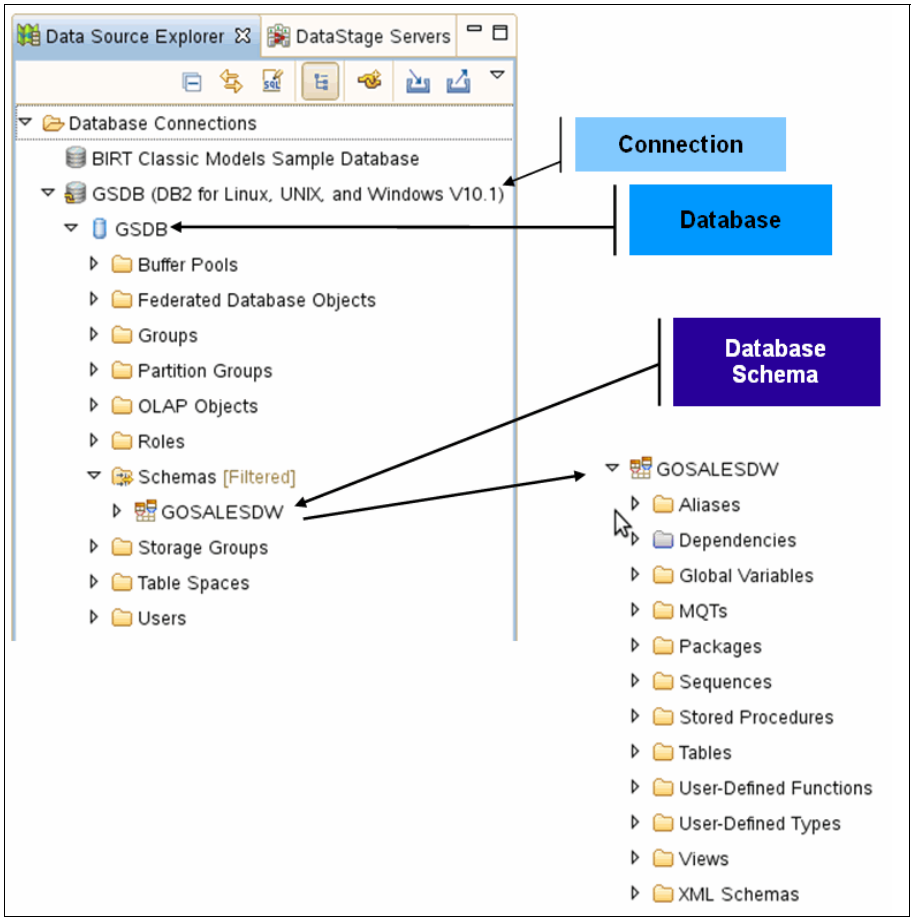


Figure 5-5 Data Source Explorer view

You can use the following options in Data Source Explorer view:

- Display all known database connections.
- Connect and disconnect to databases.
- Define new database connections.
- Filter the objects display.
- Display schema structures and properties for database objects.
- Sample data.

- Explore data value distributions.
- Compare and synchronize data objects.
- Analyze the impact of changes.
- Generate a DDL file.

Data Project Explorer view

The Data Project Explorer is a tree view that shows the files and metadata that is associated with data warehousing and data design projects, as shown in Figure 5-6.

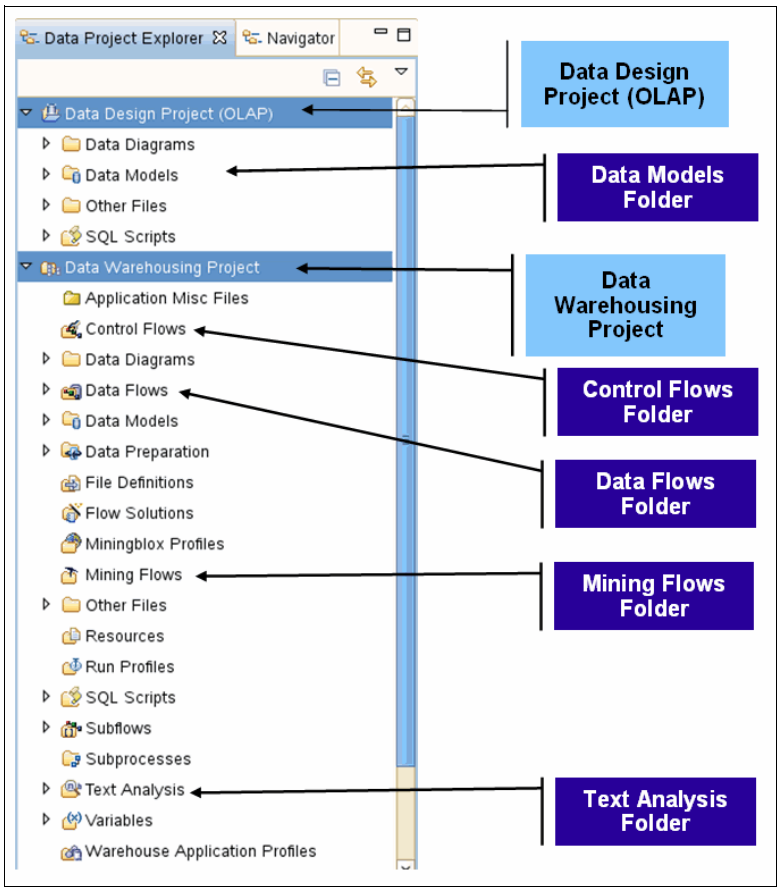


Figure 5-6 Data Project Explorer view

The following options are available in the Data Project Explorer view:

- Save SQL scripts and physical database models.
- Drag table schemas from the Data Project Explorer view and drop them in the data flow editor to create source or target operators.

- Import database metadata from a data source connection.
- Create new database objects into the data models.
- Create data flows and control flows.

Properties view

The Properties view is a tabbed view that opens by default at the bottom of the Design Studio panel. This view is actionable and updates itself when a particular object is selected, specifying its detailed behavior. In other words, the information displayed is dependant on the type of the database object you select. For example, when you select a table, you can view properties such as column definitions, privileges, table spaces, and so on, as shown in Figure 5-7.

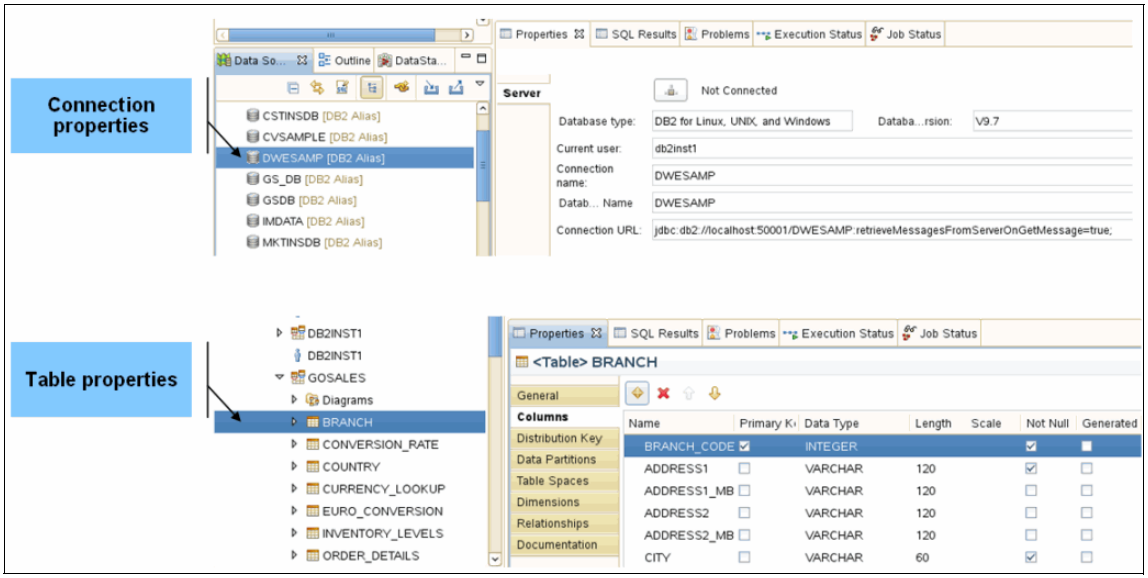


Figure 5-7 Properties view when a connection or table is selected

Most object properties from the Data Source Explorer view are read-only. You must use SQL scripts, DDL files, or the Data Studio tool to make any changes to the database objects.

Graphical Editor view and palette

The Graphical Editor view is a canvas where the cloud consumers can create data diagrams, data flows, control flows, and mining flows. The Graphical Editor view has a palette with built-in operators that can be dragged and dropped into the canvas, as shown in Figure 5-8.

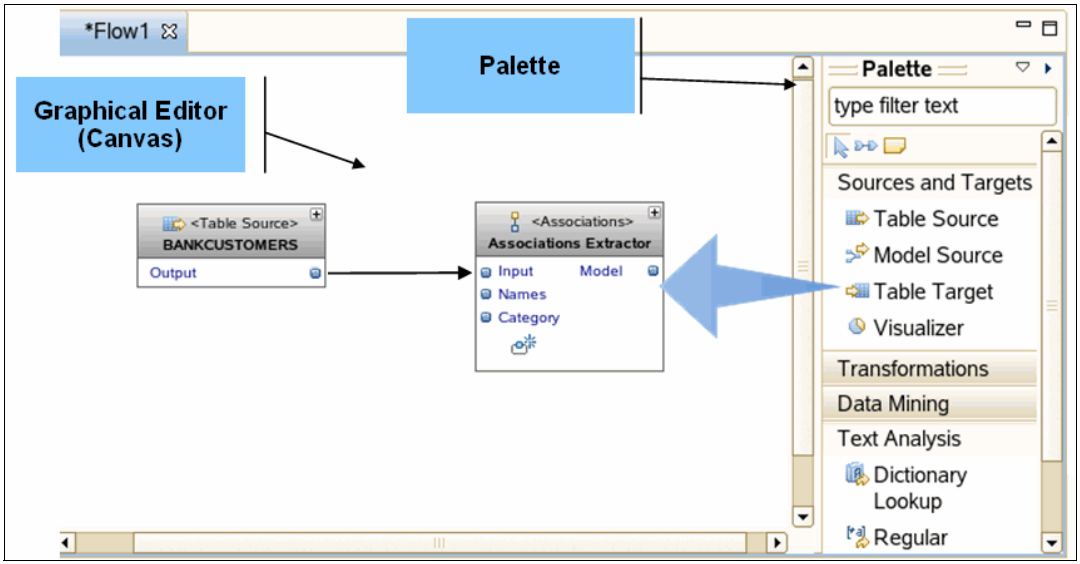


Figure 5-8 Graphical Editor view of canvas and palette

Cloud consumers can access the following palettes depending on the object they choose to create:

- Diagram palette
- Data flows palette
- Mining flows palette
- Control flows palette

Design Studio basics

Before cloud consumers create projects, prepare the Design Studio environment:

1. Open Design Studio:
   - In a Windows environment, click **Start** → **Programs** → **IBM InfoSphere Warehouse** → ***ISWCOPY01*** → **Design Studio**. *ISWCOPY01* represents the name of the client instance that was specified when the InfoSphere Warehouse was installed.
   - In a Linux environment, execute the following command: `/opt/IBM/ISWarehouse/ds/eclipse`

2. Specify the workspace folder to save the data warehouse projects, as shown in Figure 5-9.

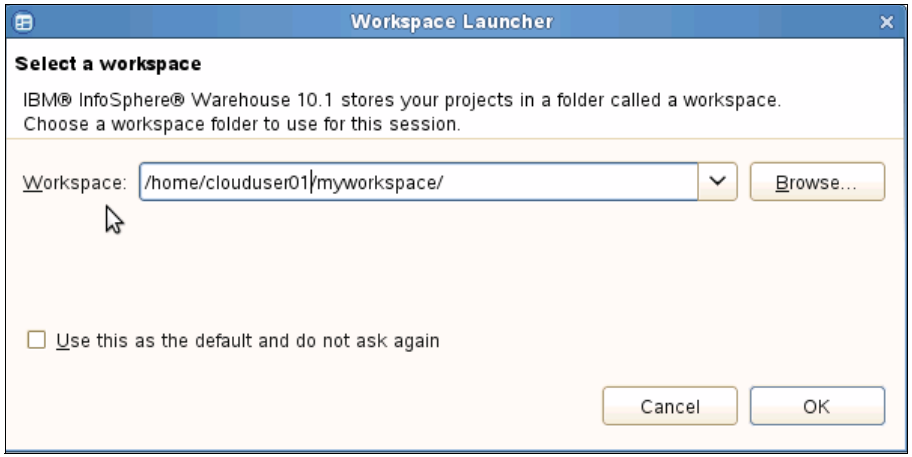


Figure 5-9 Workspace launcher

Design Studio opens and displays the welcome page. The cloud consumer can use the workbench basics and tutorials from the welcome page, as shown in Figure 5-10 on page 171.

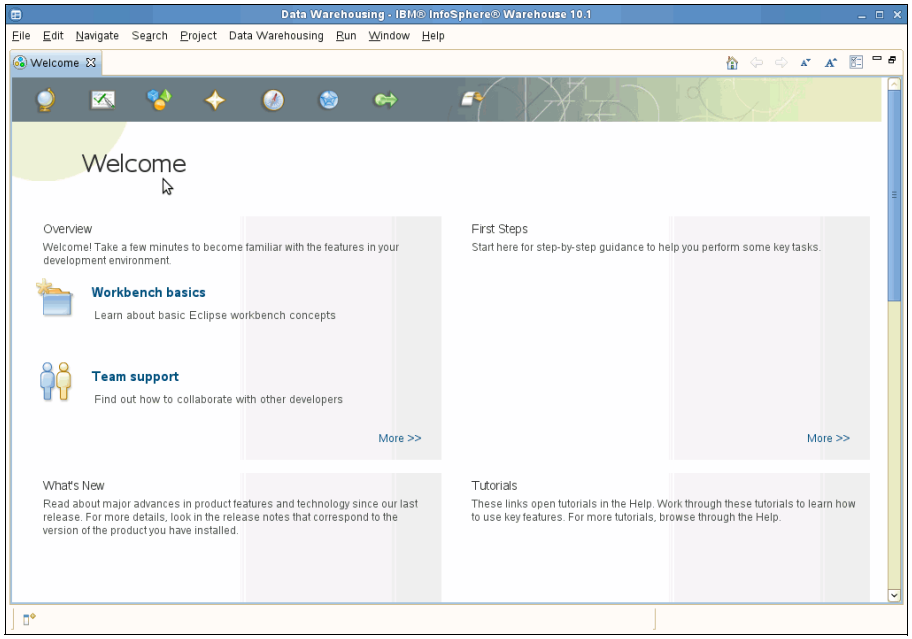


Figure 5-10 Design Studio welcome page

Upon closing the welcome page, you are by default placed in the Data Warehousing perspective. You can navigate to any other view from that point.

3. Create or edit the driver definitions.

   The drivers for the IBM databases, such as DB2 for Linux, UNIX, and Windows, DB2 for IBM z/OS®, and IBM Informix®, are preinstalled with InfoSphere Warehouse. The cloud consumer can create data source connections directly using one of the preinstalled drivers. InfoSphere Warehouse also supports the following drivers:

   – Microsoft SQL Server
   – Oracle
   – Derby
   – MySQL
   – Sybase
   – Teradata
   – Generic JDBC

   The only supported JDBC driver is the Type 4 driver.

   If the developer wants to use a non-IBM database, such as Microsoft SQL Server or Oracle driver, configure the driver first. For each new driver, create a definition on the Design Studio and on the Administration Console. Drivers are classified by their system vendors; therefore, you must add the driver class by vendor and by version.

   In Design Studio, you can create or edit the driver definition at **Window** → **Preferences** → **Data Management** → **Connectivity** → **Driver Definitions** from the menu bar.

4. Create data source connections.

   After you define the driver, cloud consumers can create database connections on the Data Source Explorer view. To catalog operational databases using a JDBC connection:

   a. In the Data Source Explorer view, right-click the **Database Connections** folder and click **New**.

   b. The New Connection wizard opens. Complete the connection parameters as described in Table 5-1.

Table 5-1 New connection parameters

Field	Field description	Paremeter example
Select the database manager	Select any of the database managers available from the list.	DB2 for Linux, UNIX, and Windows[a]
JDBC Driver	Select the JDBC driver.	IBM Data Server Driver for JDBC and SQLJ (JDBC 4.0) Default[b]

Field	Field description	Paremeter example
Database	Type the database name to which you are connecting.	DWHDB
Host	Type the host name where the database is located.	ISWSERVER
Port number	Type the port number to connect with the database manager.	50001
User name	Type the user name to connect with the database.	db2inst1
Password	Type the password to connect with the database.	db2pass

a. Choose from the database managers supported.
b. Choose from the JDBC drivers supported.

c. Click **Test the connection** and click **OK**. If the connection test is successful, click **Finish**. If the connection is unsuccessful, ensure that the database is accessible and verify the connection parameters.

5. Perform tasks on objects.

   When you right-click a data or project object inside a view in Design Studio, a context-based menu usually displays. This menu lists all the tasks that you can perform on the selected object. Perform any tasks needed specific to the context of your setup.

Data Studio

Data Studio is a development and administration client built on the Eclipse framework. All of the Data Studio functionality is supplied by the Eclipse plug-ins, It also supports features that can enable the cloud consumers to do the following activities:

- Create data overview diagrams
- Create, run, and schedule scripts
- Manage instances
- Create, alter, or drop database objects
- Manage data
- Grant or revoke security privileges
- Develop applications
- Tune SQL statements

The default perspective is the Database Administration perspective, as shown in Figure 5-11.

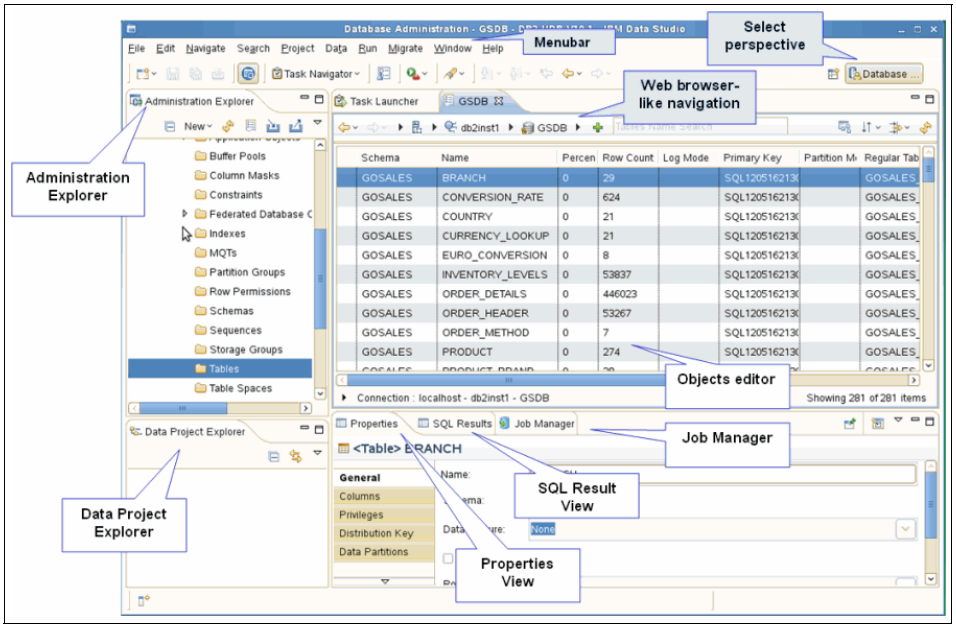


Figure 5-11 Database Administration perspective

The following views in the Data Administration perspective have different functions and purposes:

- The Administration Explorer view (upper left pane) lists all instances, database connections, and database objects.
- The Objects Editor view (upper right pane) acts as a web browser to go back and forth to different pages.
- The Properties view (bottom right pane) lists the properties of any selected database object.

When you right-click a data object inside a view in Data Studio, a context-based menu usually appears, and it lists all the tasks you can perform on the selected object. When the cloud consumer selects a task, the task assistant appears from the context-based menu. Task assistants are dialog boxes that can help to create and run database administration commands for objects in DB2 for Linux, UNIX, and Windows databases, as shown in Figure 5-12.

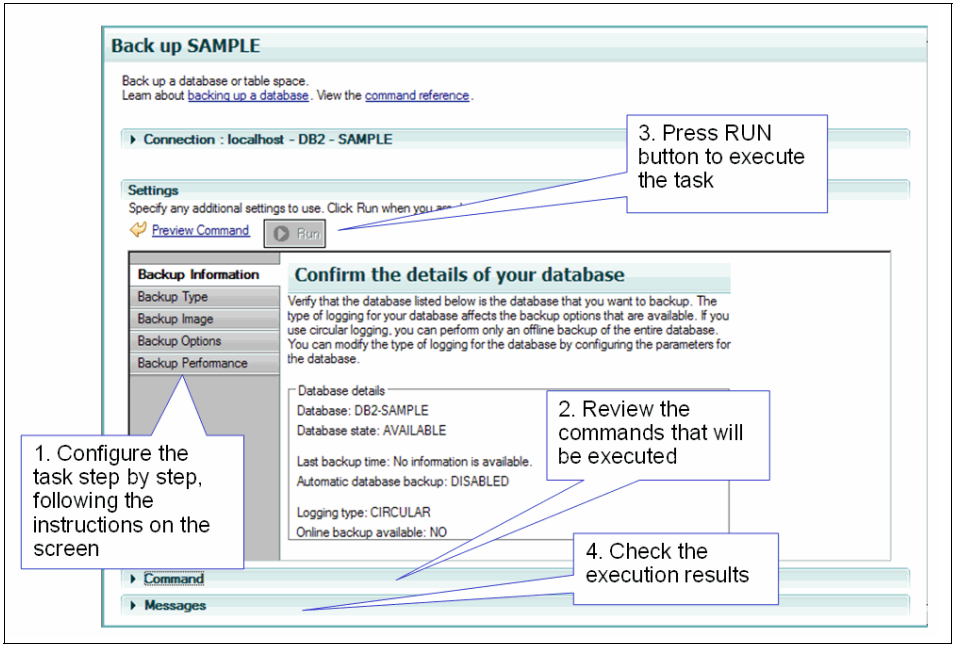


Figure 5-12 Task assistant example

Before the cloud consumer can create database objects, use the following steps to prepare the Data Studio environment:

1. Start Data Studio.
   - For Windows operating systems, click **Start** → **All Programs** → **package_group_name** → **IBM Data Studio**.
   - For a Linux environment, execute the following command:

     ```
     /opt/IBM/DS3.1.1/eclipse
     ```

2. Create data source connections.

   All saved database connections are listed in the Administration Explorer view. If DB2 aliases are configured on the system, for example all databases catalogued on the machine, database connections are created automatically when Data Studio is launched. If you want to change this behavior, use the preferences menu.

To create a database connection manually, use the New Connection wizard from the Administration Explorer as shown in the Figure 5-13. The cloud consumer can connect to different database servers, not just to DB2. To establish the connection, specify the database driver, database name, host name, port number, and user login credentials.

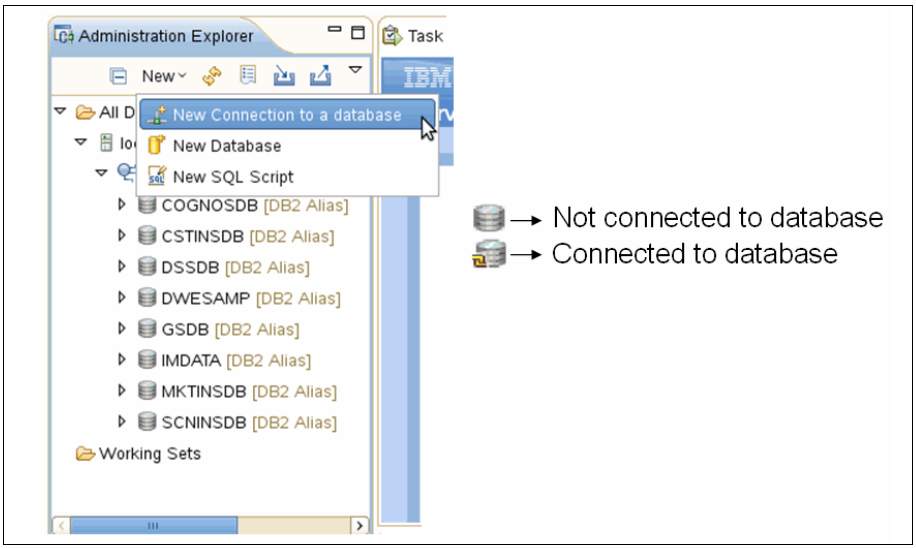


Figure 5-13 New connection to a database

Cognos Framework Manager

Cognos Framework Manager is a metadata modeling tool. It is used to create and publish query packages of relational and OLAP data sources in the IBM Cognos Connection web portal. It helps to reduce the complexity of the underlying data source by creating data views for the business users and hiding the structural complexity of your underlying data sources.

The methods to access data for reporting purposes is determined by the structure and content of the data itself. You can connect to the warehouse database directly; however, if the information is required for multidimensional analysis on a large set of aggregated data, consider connecting to an OLAP data source instead. InfoSphere Warehouse Cubing Services is one data source option, as shown in Figure 5-14.

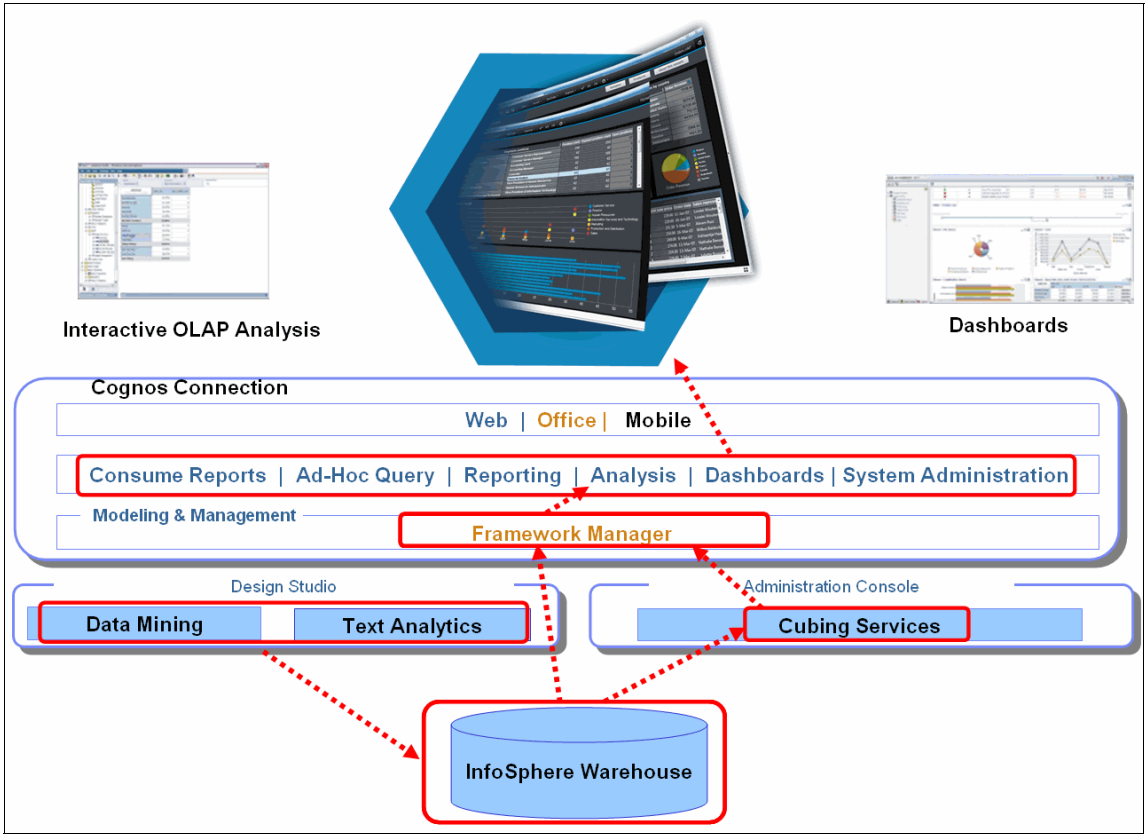


Figure 5-14 Cognos reporting and analysis using InfoSphere Warehouse database and cubing services

The Framework Manager perspective, shown in Figure 5-15 on page 178, provides the following panes with specific functions and information displays:

- The Project viewer pane (upper left corner) provides access to all the project's objects in a tree format.
- The project information pane (center) provides access to three tabs (Explorer, Diagram, and Dimension Map) that allows the cloud consumers to create, edit, configure, or delete objects.
- The Properties pane (center bottom) allows the cloud consumers to configure various properties for any of the project's objects.
- The Tools pane (upper right corner) allows the cloud consumers to switch the project language, view project statistics and perform common tasks for selected objects.

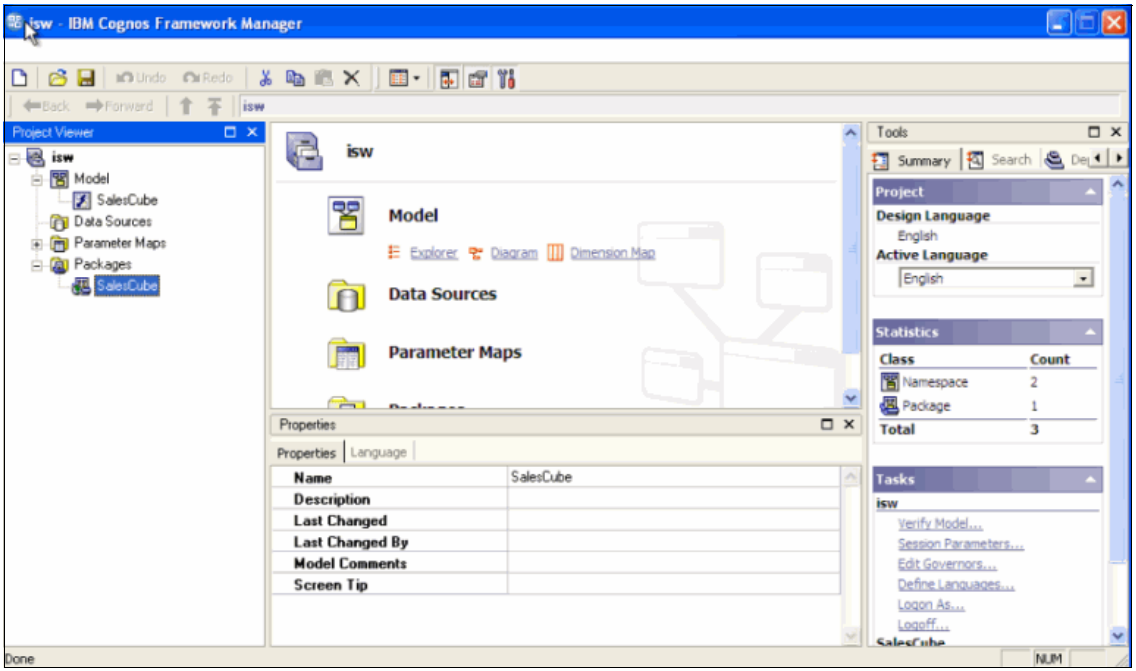


Figure 5-15 Framework Manager perspective

Project viewer

The Project viewer includes the following containers:

- The Model namespace contains data model objects.
- The *Data Sources* folder contains data source connections.
- The *Parameter Maps* folder contains parameter maps that allow data substitution at run time.
- The *Packages* folder contains packages that can be published to Cognos Connection. Publishing makes data sources available to cloud consumers using IBM Cognos Business Intelligence. See “Create and publish a package” on page 219 for more information.

Framework Manager basics

Before the cloud consumer can work with a Framework Manager model, prepare the Framework Manager environment:

1. Open Framework Manager by clicking **Start** → **All programs** → **IBM Cognos 10** → **Framework Manager**.
2. Under Projects, click **Create a new project**, browse to the location where you want to save the project, and click **OK**.

3. In the Project name field, enter the name of the project, and click **OK**.
4. In the Select Languages dialog box, select the language that is required.
5. The Metadata wizard opens. You can import the metadata in a later task. Click **Cancel**. The project opens. Refer to “Model packages publishing with Cognos Framework Manager” on page 217 for more advanced tasks.

5.2.3 Web tools that are available for SaaS

The cloud service consumer can use web tools to administer, run, and query BI applications that you host in your BI cloud.

InfoSphere Warehouse Administration Console

The InfoSphere Warehouse Administration Console is a rich text web application built using Flash/Flex. It supports web accessibility and improved process workflows for better usability. You can use this tool to administer SQL Warehousing, Cubing Services, and Intelligent Miner components.

The InfoSphere Warehouse Administration Console interface is hosted by WebSphere Application Server. Thus, WebSphere Application Server must be running before you can use the Administration Console. Use a URL that points to the InfoSphere Warehouse Administration Console, such as `http://<hostname>:<portnumber>/ibm/warehouse` (for example `http://localhost:9080/ibm/warehouse`).

From the page as shown Figure 5-16, you can access the InfoSphere Warehouse Administration Console by selecting its hot link under My Tools.

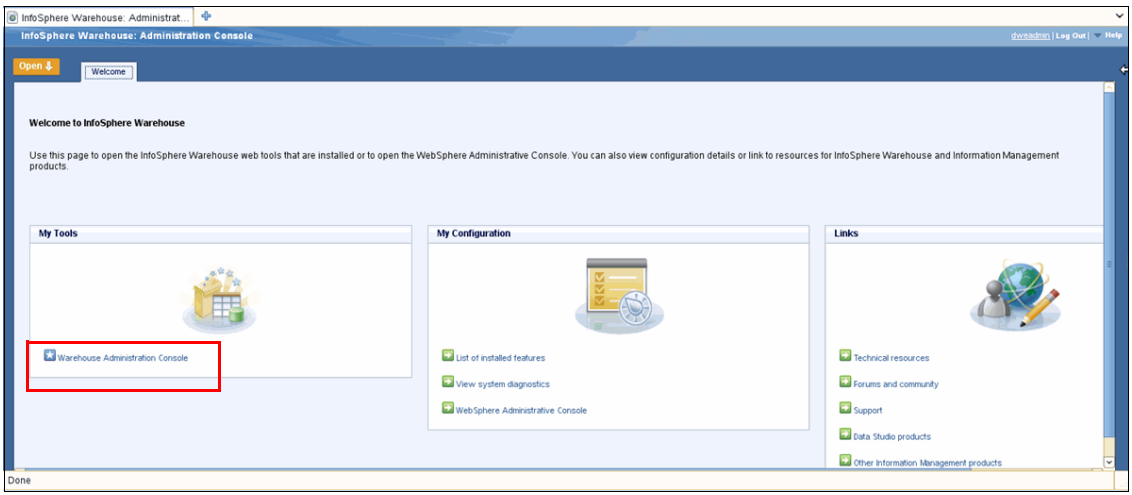


Figure 5-16 Administration Console Welcome page

The Administration Console supports the following features:

- Maintain InfoSphere Warehouse services for notification and logging.
- Manage data server drivers and data source connections.
- Manage system resources, such as FTP servers, DataStage servers, and plain servers.
- Deploy warehousing applications created in Design Studio.
- Schedule, run, and change variables of control flows.
- Monitor all scheduled, running, and completed instances.
- Deploy cube models and manage cubing services.
- Administer mining models.

Access the InfoSphere Warehouse Administration Console menu by selecting **Open**, as shown in Figure 5-17.

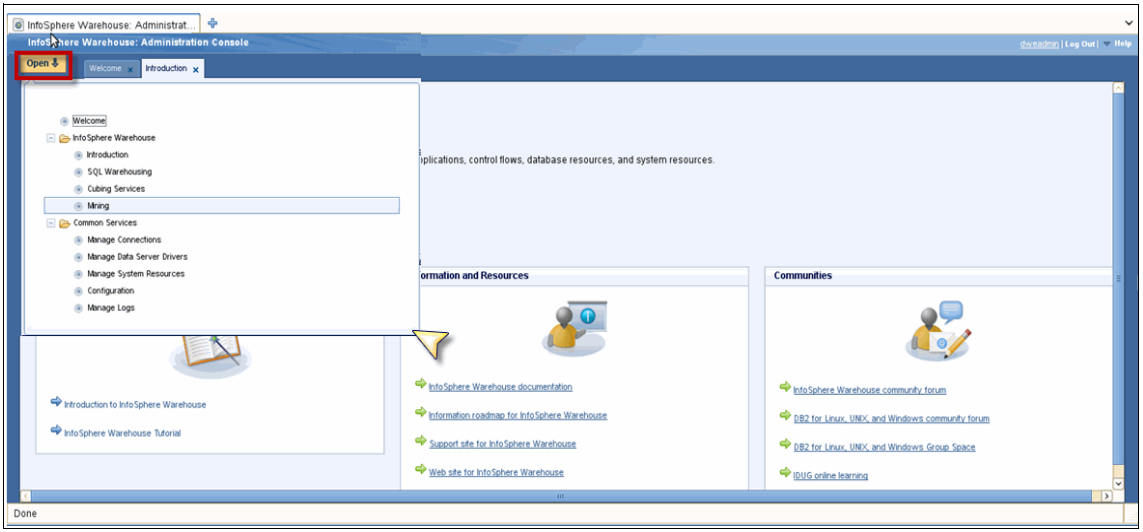


Figure 5-17 Manage connections at the Administration Console

Cognos Connection

IBM Cognos Business Intelligence 10 delivers a web tool called Cognos Connection. This tool provides a BI workspace to support a diverse reporting environment and can help cloud consumers to make comprehensive business decisions. Cognos Connection has a customizable portal interface to access Cognos content and everyday reports, live or saved output versions (Figure 5-18). From here you can launch various tools and perform administration on IBM Cognos Business Intelligence content.

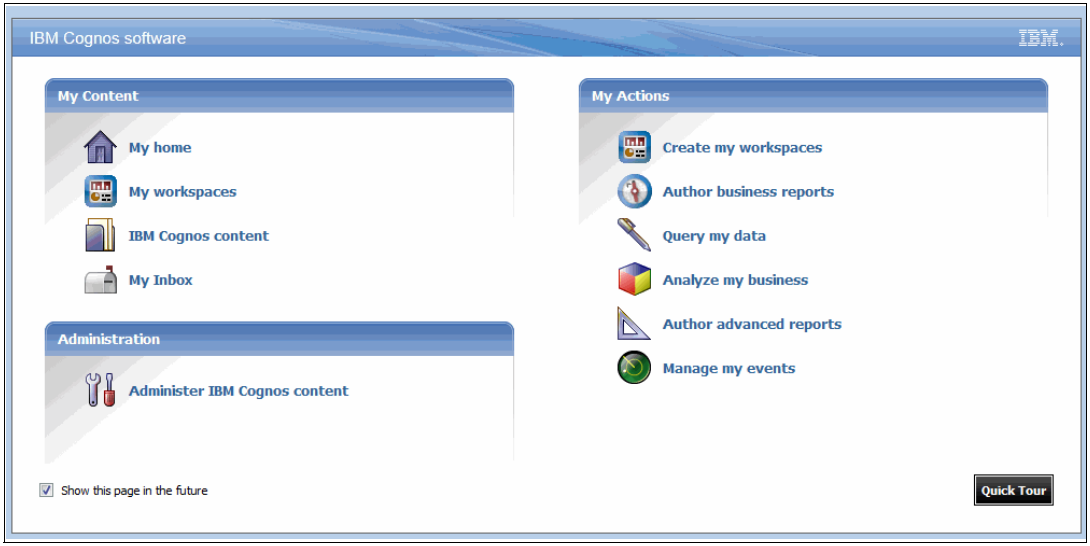


Figure 5-18 Cognos Connection web portal

When using Cognos Connection, use a URL such as `http://<hostname>/ibmcognos` (for example `http://localhost/ibmcognos`).

5.3 Cloud services catalog using InfoSphere Warehouse

To offer a comprehensive set of services to business users, you should create a *services catalog*. This services catalog itemizes the types of services that the cloud can offer and establishes delivery procedures. To make it easier to navigate and find BI services, you can organize the catalog into the following key areas:

- Development and testing on a PaaS model to consume the following services:
  - Data modeling
  - Data warehousing
  - ETL development
  - Business Intelligence development
- Production on a SaaS model to consume the following services:
  - ELT running
  - Data warehousing
  - Business intelligence
  - Performance management

5.3.1 Services catalog in an InfoSphere Warehouse virtualized environment

Figure 5-19 shows the type of services that you can deliver on PaaS and SaaS models using InfoSphere Warehouse and InfoSphere Warehouse companion solutions.

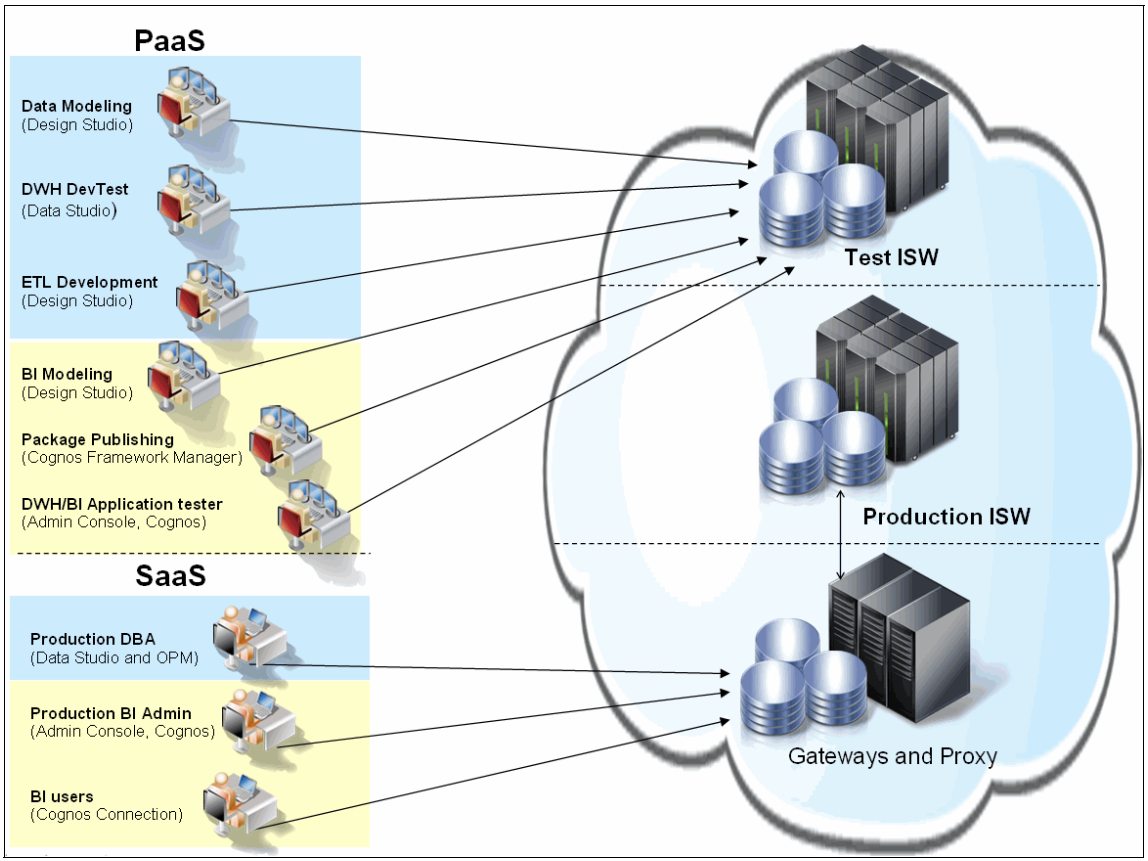


Figure 5-19 Types of services a virtualized BI cloud using InfoSphere Warehouse can deliver

The cloud consumer can access virtual desktops with the most familiar operating system, whether that is the Linux or Windows operating system. This virtual desktop can have an installation of Design Studio to provide the BI capabilities to cloud users. The cloud provider sets up the appropriate VMs to run the services and provide access details.

The following example service catalog illustrates the capabilities that InfoSphere Warehouse can provide within a virtualized environment:

- Development and testing services (PaaS)

  Your view and definition of business analytics varies from the view of other people in the enterprise depending upon your role within the enterprise. If your position requires you to work with getting the data in shape for BI, the issues you face differ from those of the individuals trying to use the data that you have created.

  BI success is measured in business impact and by the improvements in critical areas that can be attributed to its development and customization.

  Table 5-2 is an example of the development and testing services that you can provide to the cloud consumers so that they can create customized applications for BI.

Table 5-2 Development and testing services catalog using InfoSphere Warehouse

Service	Service description	Service technology
Data modeling	Design data models to meet specific business analytics requirements.	Includes Design Studio access.
Data warehouse (DevTest)	Store data into a DB2 database ready for development and testing.	Includes Data Studio and user database access to store data on an InfoSphere Warehouse database server environment.
ETL development	Develop SQL extraction, transformation, and loading (ETL) data flows for development and testing.	Includes Design Studio and SQL execution database access for basic ETL processing on an InfoSphere Warehouse database server environment.
OLAP modeling	Develop OLAP metadata objects like cube models, measures, dimensions, hierarchies, level and cubes. You can validate OLAP modeling, create role security models, and run the optimization advisor.	Includes Design Studio and predefined star schemas access stored on an InfoSphere Warehouse database server.

Service	Service description	Service technology
Data mining modeling	Develop mining flows for modeling, scoring, and visualizing data mining models, such as Clustering, Associations, Sequential Rules, Classification, Time Series Forecasting, and Regression.	Includes Design Studio and an SQL execution database enabled for mining access stored on an InfoSphere Warehouse database server.
Unstructured text modeling	Develop mining flows for text analysis such as dictionary lookup, regular expressions, and frequent terms.	Includes Design Studio and an SQL execution database enabled for mining access stored on an InfoSphere Warehouse database server.
Model packages publishing	Develop Cognos packages that can be consumed by Cognos Connection.	Includes Cognos Framework Manager access and Cognos Business Intelligence environment.
Data Warehouse and BI application testing	Test the models created in Design Studio and Cognos Framework Manager.	Includes administration tools (Administration Console and Cognos Connection portal) access on an InfoSphere Warehouse application server environment and a Cognos Business Intelligence environment.

- Production services (SaaS)

  Cloud services for SaaS are built and standardized so that cloud consumers can subscribe from a self-service catalog, pay-per-use, and consume services over the Internet using web front-end tools. SaaS users have access only to applications that are hosted over the SaaS infrastructure. They are not allowed to make changes to the applications, unless they have access to development tools from a PaaS infrastructure.

Table 5-3 is an example of the production services that you can provide to cloud consumers to create and run applications for BI.

Table 5-3 Production services catalog using InfoSphere Warehouse

Service	Service description	Service technology
Data warehouse (Production)	Store data into a DB2 database ready for production reporting and analysis. Monitor database health and availability, and run jobs against the database, such as SQL, DB2 CLP, and executable or shell scripts.	Includes Data Studio Web Console access and an InfoSphere Warehouse database server environment to store the user database and metadata repository to store control access and user privileges.
ETL processing	Connect to data sources and run predefined InfoSphere Warehouse applications with control flows to extract, transform, and load (ETL) data into the data warehouse. Execute control flows for initial data load and schedule loads when needed to update warehouse data.	Includes Administration Console access, an InfoSphere Warehouse database server tier component to store an SQL Execution Database, and an InfoSphere Warehouse metadata database for basic ETL processing.
ETL advanced processing	Connect to DataStage jobs and run them along with InfoSphere Warehouse applications.	Includes Administration Console access, predefined DataStage jobs from DataStage server, and an InfoSphere Warehouse application server to store a metadata control database for advanced ETL processing.
Production BI administration	Create cube servers that can support data cubes ready for multidimensional analysis (MDX), administer mining models, deploy predefined InfoSphere Warehouse applications, and publish Cognos packages.	Includes Administration Console and Cognos Administration access, an InfoSphere Warehouse database server to store the user database, and the InfoSphere Warehouse control database to store the predefined cube models and databases enabled for mining.
Ad-hoc reporting	Access self-service reporting to create ad-hoc reports and quickly get answers to business questions.	Includes Query Studio access, predefined Cognos Packages, and a connection to a DB2 warehouse database.

Service	Service description	Service technology
Analytical reporting	Perform multidimensional exploration with the capability to drill down, drill up, and slice and dice data.	Includes Analysis Studio access, a Cognos package, and a connection to InfoSphere Warehouse database server and InfoSphere Warehouse application server to access cubing services.
Advanced reporting	Access professional authoring environment with great flexibility, to create professional looking reports.	Includes Report Studio access, a predefined Cognos package, and a connection to an InfoSphere Warehouse database server.
Assembling reporting	Assemble content on a dashboard with reports in an intuitive authoring experience. Allows the user to drag objects into the report authoring window.	Includes Business Insight access, predefined Cognos reports, and a connection to an InfoSphere Warehouse database server.
Mobile reporting	Access Cognos content from a mobile device.	Includes the mobile applications and predefined Cognos reports and Cognos Mobile server.
Data warehouse performance management	Improve database and application performance, and view reports for data capacity planning and growth.	Includes a web tool (Optim Performance Manager), a DB2 database repository to store historical performance metrics, and tracking information for a data warehouse.
Data warehouse advanced performance management	Delivers end-to-end database and application monitoring for Java and IBM DB2 Call Level Interface applications, with configurations for IBM Cognos, IBM InfoSphere DataStage, and IBM InfoSphere SQL Warehouse applications.	Includes Optim Performance Manager extended edition access, DB2 database repository to store historical performance metrics, and tracking information.

For more information about InfoSphere Warehouse capabilities, see *InfoSphere Warehouse: A Robust Infrastructure for Business Intelligence*, SG24-7813.

5.3.2 ETL and data warehouse services catalog

The warehouse building lifecycle, shown in Figure 5-20, begins with data modeling services, where the structure of the data warehouse is defined. From there, the warehouse and associated data marts must be populated with data. To populate the warehouse and data marts with data, you need to create data flows, mining flows, and control flows in Design Studio.

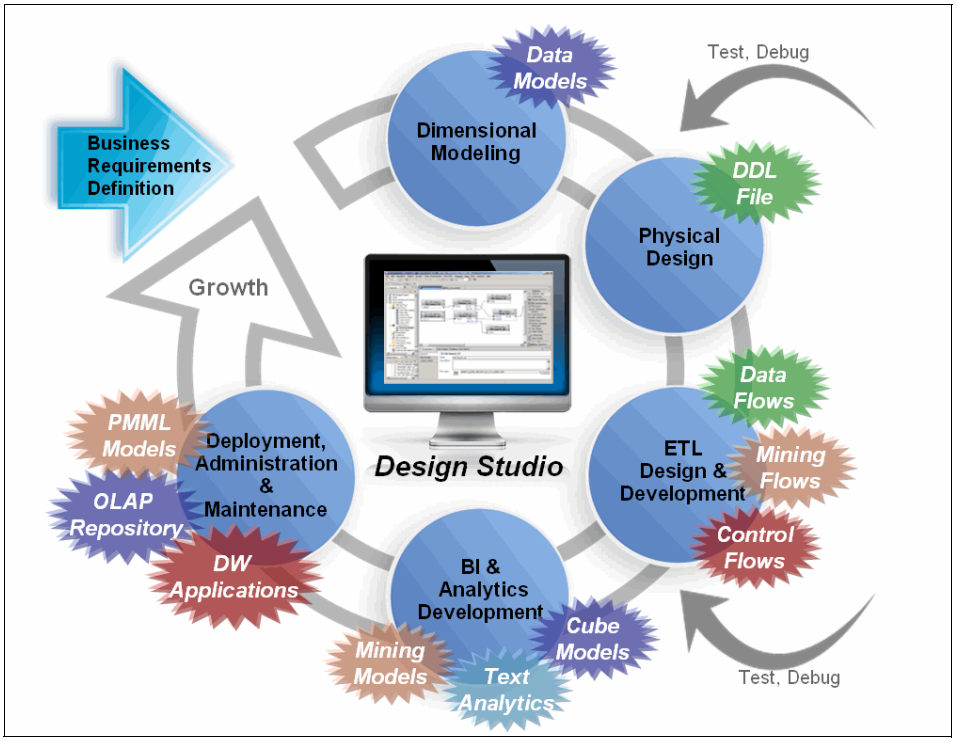


Figure 5-20 InfoSphere Warehouse development lifecycle using Design Studio

Data modeling services with Design Studio

When you create a data warehouse, you can load data in the following data layers:

- The staging area
- The enterprise data warehouse in third normal form (3NF)
- The dimensional structures

The choice depends on the warehouse developer and warehouse administrator requirements, because the choice mainly impacts the loading and modeling tasks. If you create the data marts only, those can be controlled only by particular lines of business. However, the primary concern when using data marts is the

currency of data or how to refresh them. There is a risk of inconsistency of reports due to different update cycles because usually data comes from a variety of data sources. This is where the enterprise data warehouse is useful, because the data is loaded to the consolidated enterprise data warehouse. Thus, the data can be consistent for all data marts.

The most common approach is combining the enterprise data warehouse and the data marts, where the data marts receive data from the enterprise data warehouse, rather than directly from the staging area or data sources. The enterprise data warehouse has a high degree of data access and usage for departments or lines of business because it is a common repository for decision-support data, which is available for the entire organization. You can allow the developers and analysts to query the warehouse directly or the data marts instead. The staging area can be replaced by the enterprise data warehouse, combining the extraction, transformation, and the target enterprise data warehouse model in one database.

The multi-dimensional structures (usually star schemas) are used to support OLAP. Its workload is characterized by performing complex queries and multiple aggregations. The star schemas can be created within the data warehouse to optimize the retrieval of much larger data sets that are derived from many tables through join operations. The star schema has one fact table that measures data from any business process. It is associated to multiple dimension tables that complement the fact with attributes and more descriptive information. Also, there are some variations of star schemas, such as the snowflake model or multi-star schema model.

With Design Studio, you can create or reverse engineer physical data models and view them graphically. Also you can continue to refine the model either using the graphic editor or modifying the project tree. At any step in modifying the model, you can compare the modeled data objects with other objects that are either in the model or in the physical database, and also look at the change's impact on other objects. After the data architect is satisfied with the model, it can be converted into a data definition language (DDL) file.

Physical data models are the abstraction of the database schema that capture implementation constraints of the specific database. It is implemented as an entity relationship model and its visual representation of models uses either information engineering (IE) or unified modeling language (UML) notation. The physical data model includes diagrams, database objects (tables, columns, indexes), and constraints (primary key, unique, and foreign key).

Figure 5-21 illustrates a project tree in the Design Studio workspace that contains the metadata of a newly created physical data model.

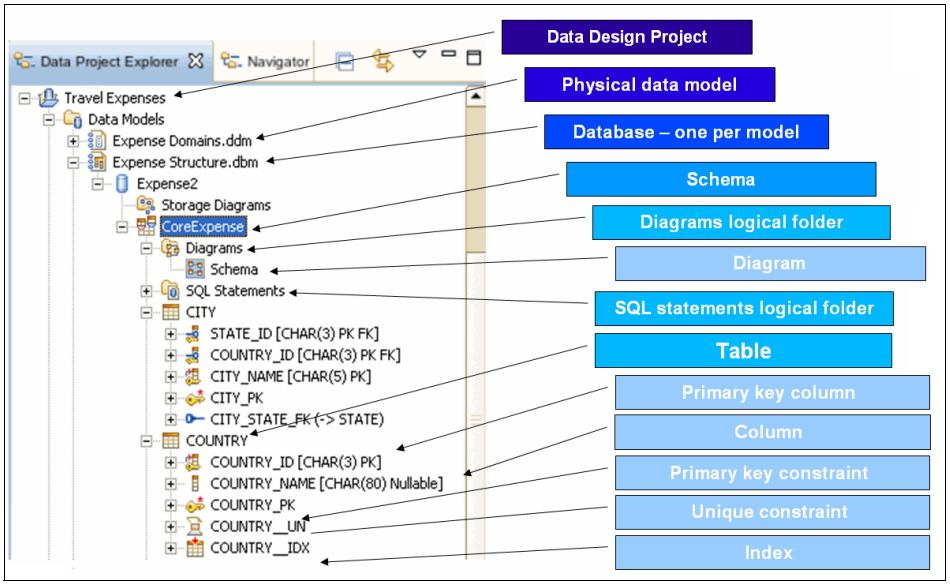


Figure 5-21 Data Model structure

When performing physical data modeling in Design Studio, you can first define a Data Project, where the new physical data model resides. All the objects that are associated to that particular physical data model, starting with the database, are organized under this Data Project folder.

Create a Data Design project

Before you create a physical data model or other data design objects, you must create a Data Design project in which to store design objects. You can store various types of design objects in a data design project, including data models, SQL scripts, and XML schemas.

To create a data design project:

1. Click **File** → **New** → **Data Design Project (OLAP)**.
2. Provide a project name in the New Data Design Project wizard.
3. Click **Finish**.
4. The new project displays in the Data Project Explorer view. Expand the project, and explore the subfolders from the tree.

Create a physical data model

A physical data model is created to contain table definitions from the database. You can work on the model without accessing the database directly, which makes creating and deleting tables easier and ensures that you do not unintentionally damage the physical database. In some cases, the person who designs the model might not have access to the database, or the database might be down.

Figure 5-22 illustrates the methods to create a physical data model using Design Studio and the sections that follow provide step-by-step instructions for each method.

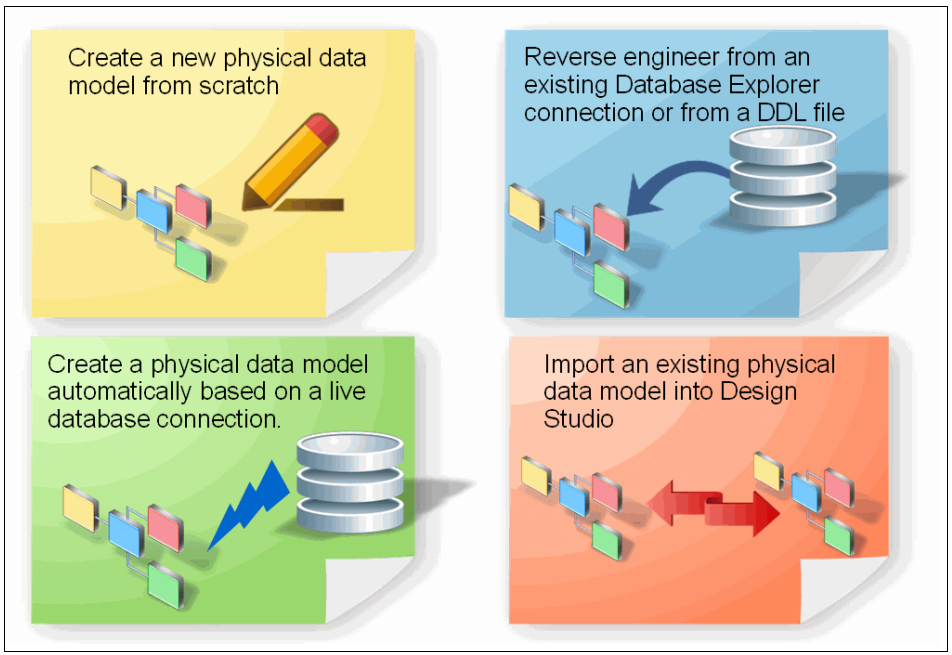


Figure 5-22 Different ways to create a physical data model in Design Studio

- To create a physical data model from scratch:
  a. Click **File** → **New** → **Physical Data Model**.
  b. Complete the New Physical Data Model wizard by entering the parameters described in Table 5-4 on page 191.

Table 5-4 New Physical Data Model wizard field parameters

Field	Field description	Parameter example
Destination folder	Location of the new model	MyProject
File Name	Name of the physical data model	DataModel
Database	Database vendor	DB2 for Linux, UNIX, and Windows
Version	Database version	V10.1

c. Select **Create from template** and click **Next**.

d. Choose **Empty Physical Model Template** and click **Finish**.

- A physical data model can be reverse engineered from an existing database connection or using a DDL script.

  A. To create the physical data model by reverse engineering using an existing database connection, follow these steps:

  a. From the Data Project Explorer view, expand your Data Modeling Project.

  b. Right-click the **Data Models** folder and select **New** → **Physical Data Model**.

  c. Complete the New Physical Data Model wizard as described in Table 5-5 on page 191.

Table 5-5 New physical data model field settings

Field	Setting description	Setting example
Destination Folder	Location of the new model	MyProject
File Name	Name of the physical data model	OLTP Model
Database	Database vendor	DB2 for Linux, UNIX, and Windows
Version	Database version	V10.1

  d. Select **Create from reverse engineering** and click **Next**.

  e. In the Select Connection section, select the data source connection and click **Next**.

  f. In the Select Objects section, select the schemas that you require and click **Next**.

g. In the Database Elements section, select the elements that you require and click **Next.**

h. In the Options section, select **Overview** and click **Finish**.

The schema objects are scanned and information in the catalog is copied as metadata into your workspace.

B. To create a physical data model using reverse engineering and a DDL script, follow these steps:

a. From the Data Project Explorer view, expand your Data Modeling Project.

b. Right-click the **Data Models** folder, and select **New** → **Physical Data Model**.

c. Complete the New Physical Data Model wizard as described in Table 5-6.

Table 5-6 New physical data model field settings

Field	Setting description	Setting example
Destination Folder	Location of the new model	MyProject
File Name	Name of the physical data model	OLTP Model
Database	Database vendor	DB2 for Linux, UNIX, and Windows
Version	Database version	V10.1

d. Select **Create from reverse engineering** and click **Next**.

e. In the Select Connection section, select the data source connection and click **Next**.

f. In the Select Objects section, select the schemas that you require and click **Next**.

g. In the Database Elements section, select the elements that you require and click **Next.**

h. In the Options section, select **Overview** and click **Finish**.

The schema objects are scanned and information in the catalog is copied as metadata into your workspace.

- Create a Physical Data Model from a live database connection.

  If you want a physical data model to be created automatically based on a live database connection, select **Online mode** when you create your data flows or mining flows.

- Create a Physical Data Model from an existing physical data model.

  If physical data models already exist but they are in a proprietary format, such as ERwin, use Design Studio to import these models and transform them to the SQL model format. Design Studio incorporates the Meta Integration Model Bridge (MIMB) feature, also known as the MITI Bridge. This bridge transforms proprietary physical data models, such as ERwin, IBM Rational Rose®, and Sybase models, to the native SQL model format that is used in the InfoSphere Warehouse Design Studio.

  To import a data model, click **File** → **Import** → **Data** → **Data Model Import Wizard**.

  To export a data model, click **File** → **Export** → **Data** → **Data Model Export Wizard**.

Create database objects in the data models

To create database objects in the data models:

1. In the Data Project Explorer view, expand the Data Models folder. Right-click the database, and select **Add Data Object Schema**. From the list of objects that displays, select the **Schema** object.
2. From the Properties view, rename the new schema in the name field. The empty schema is created and no tables exist in this schema.
3. Right-click the schema, and select **Add Data Object**. From the list of objects that displays, select the **Table** object.
4. From the Properties view, rename the new table. The empty table is created and no columns exist in this table.
5. Right-click the table, and select **Add Data Object**. From the list of objects that displays, select the **Column** object.
6. In the Properties view, rename the new column in the name field. Switch to the **Type** tab, and provide the column data type and length that you need.

Create an overview diagram and create data objects

To create an overview diagram and create data objects:

1. In the Data Project Explorer view, right-click the **Diagrams** folder, and select **New Overview Diagram**.
2. When prompted, select the elements from the schema that you want to include in the diagram and click **OK**.
3. The Diagram Editor and the palette opens as shown in Figure 5-23 on page 194. Right-click any blank space in the Diagram editor, and select **Add Data Object** to add data objects. You can also add data objects into the Diagram Editor from the palette. You can connect a table with different types

of relationships, such as Identifying, Non-identifying Optional, Non-identifying Mandatory, Non-identifying One-to-One, or View Relationship.

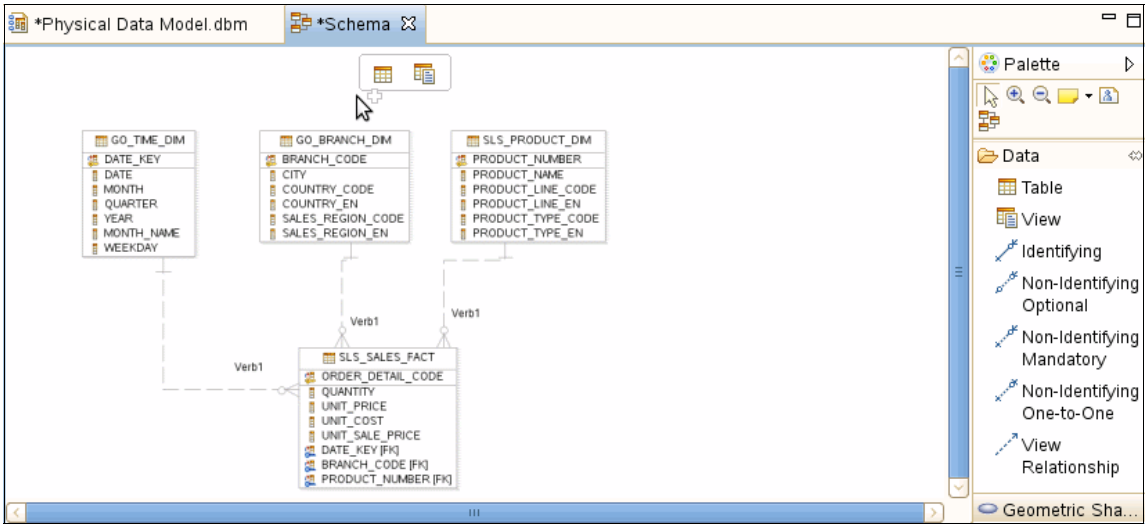


Figure 5-23 Diagram editor and palette

The Properties view, shown in Figure 5-24 on page 195, includes the following filters that you can apply to the diagram:

- Compartment display options
- Foreign key relationship display options
- Table display options
- Column display options
- Column filtering options
- Filtering columns by name
- Filtering columns by data type

For example, if you select the **Show key** and **Show non-key** options, all the columns and primary keys display in the diagram.

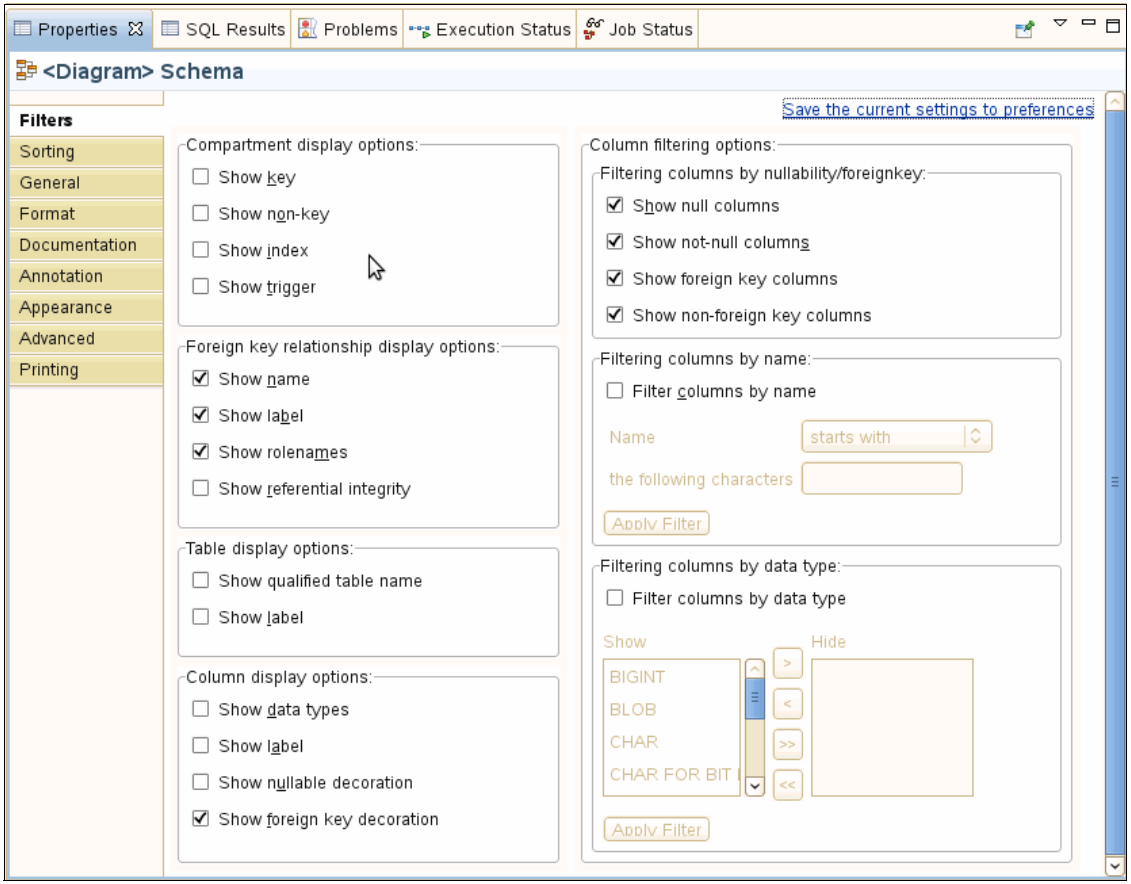


Figure 5-24 Diagram filter properties view

Analyze model

You can analyze models or the database for standards, normalization, and enterprise rules. The Model Analysis tool is built into Design Studio. To use the tool, right-click the model, and select **Analyze Model**. When performing a model analysis, you can analyze the following types of constraints and the syntax of the rules of the model:

- Rules for design and normalization
- First, second, and third normal form
- Naming standards
- Index and storage checks
- Syntax rules
- Design suggestions
- Model completeness check

When you select a rule, its description displays in the bottom pane.

Follow these steps to analyze a model:

1. In the Data Project Explorer view, right-click the **Database** object and select **Analyze Model**.
2. In the Analyze Model wizard, select a category in the Rules categories list to see the rules that are available in that category.
3. Click **Finish**.

Analyze impact

Before you make changes and synchronize those changes, you might need to use the Impact Analysis feature to view all of the dependencies of a selected element. This feature allows you to have a full view of the objects that will be affected by the changes that you are considering. The results of an impact analysis display as both a diagram and a model report list. The information includes both what the object is dependent upon and elements that are dependent upon it.

The following options to analyze impact are shown in Figure 5-25:

- Dependent Object: Selected object directly depends on other objects
- Impacted Object: Other objects depend on selected object
- Both: Dependency and impact

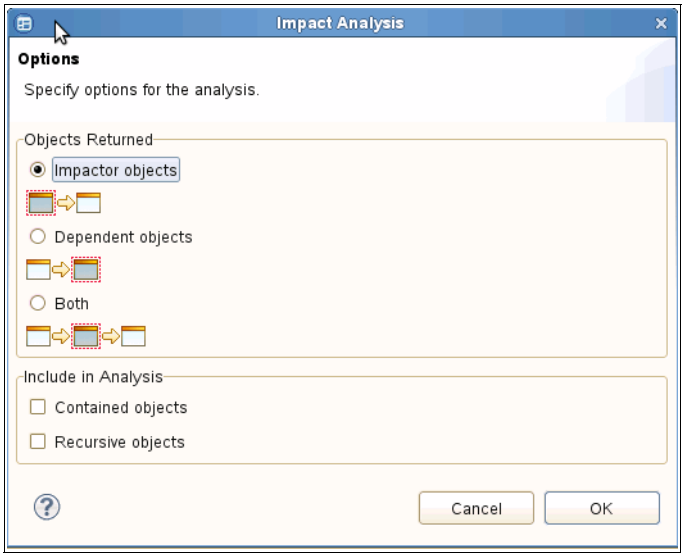


Figure 5-25 Impact Analysis options

Compare and synchronize

Comparing and synchronizing provides two capabilities. First, you can see the differences between an object in the model and another object (in the database or in the model). Second, you can transfer these differences so that the objects are identical.

The changes are in the form of DDL and can be coded to replace the target object or coded to update it. You can use the **Compare with** option to compare a model object with the original source, compare it with another data object, or compare two selected data objects, as shown in Figure 5-26.

You can update an existing database only with the compare and synchronize function using the following process:

1. Compare the model with the database.
2. Select the database objects that you need to compare and synchronize.
3. Optionally, export the differentiation script.
4. Create and update the DDL.
5. Execute the DDL on the database server.

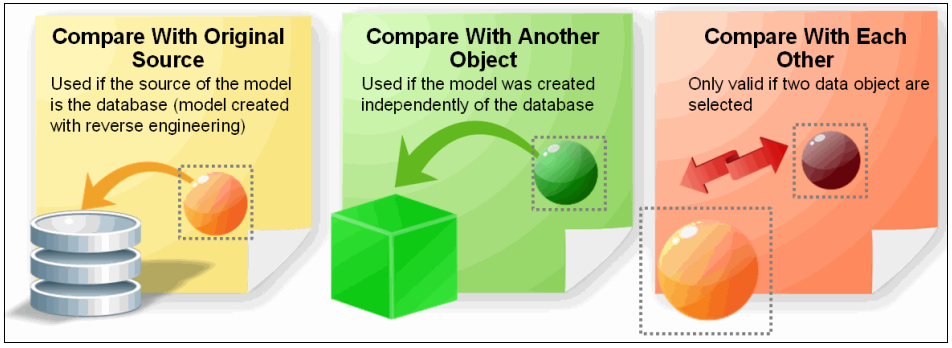


Figure 5-26 Compare and synchronize

Generate DDL file

To generate a DDL file:

1. In the Data Project Explorer view, right-click the **Database** object and select **Generate DDL**.
2. In the Generate DDL wizard, select the type of model elements that you want to include in the DDL script, and click **Next**.
3. In the Objects section, select the model objects that you want to include in the DDL script as shown in Figure 5-27 on page 198, and click **Next**.

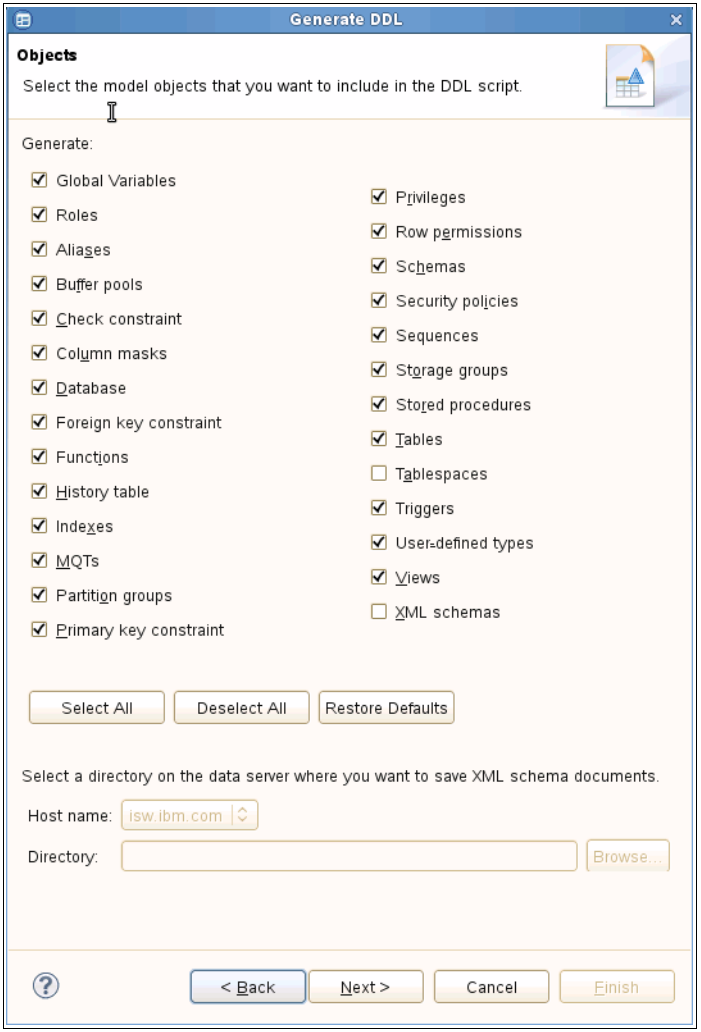


Figure 5-27 Select the model objects to include in the DDL script

4. The wizard previews the DDL that is generated. Specify a path to save the generated DDL script. If you want to run the DDL script, provide your database connection information. Click **Next**.
5. In the next section, a summary displays that shows the settings that you chose in the Generate DDL wizard. Click **Finish**.

Data Warehouse repository services with Data Studio

To deploy a warehouse DDL, the warehouse developer must create a database on an InfoSphere Warehouse database server environment and have access to a Data Studio administration client.

The InfoSphere Warehouse environment uses Data Studio, a comprehensive data management tool. The cloud consumer can administer the data warehouse repository from a virtual desktop with Data Studio installed. The developer must understand the basic concepts of DB2 from the perspective of the database administrator.

Managing databases with Data Studio

Figure 5-28 illustrates some of the options for managing a database. You can also disconnect, start or stop an instance, back up and restore a database, manage HADR, monitor the database health, and complete other tasks.

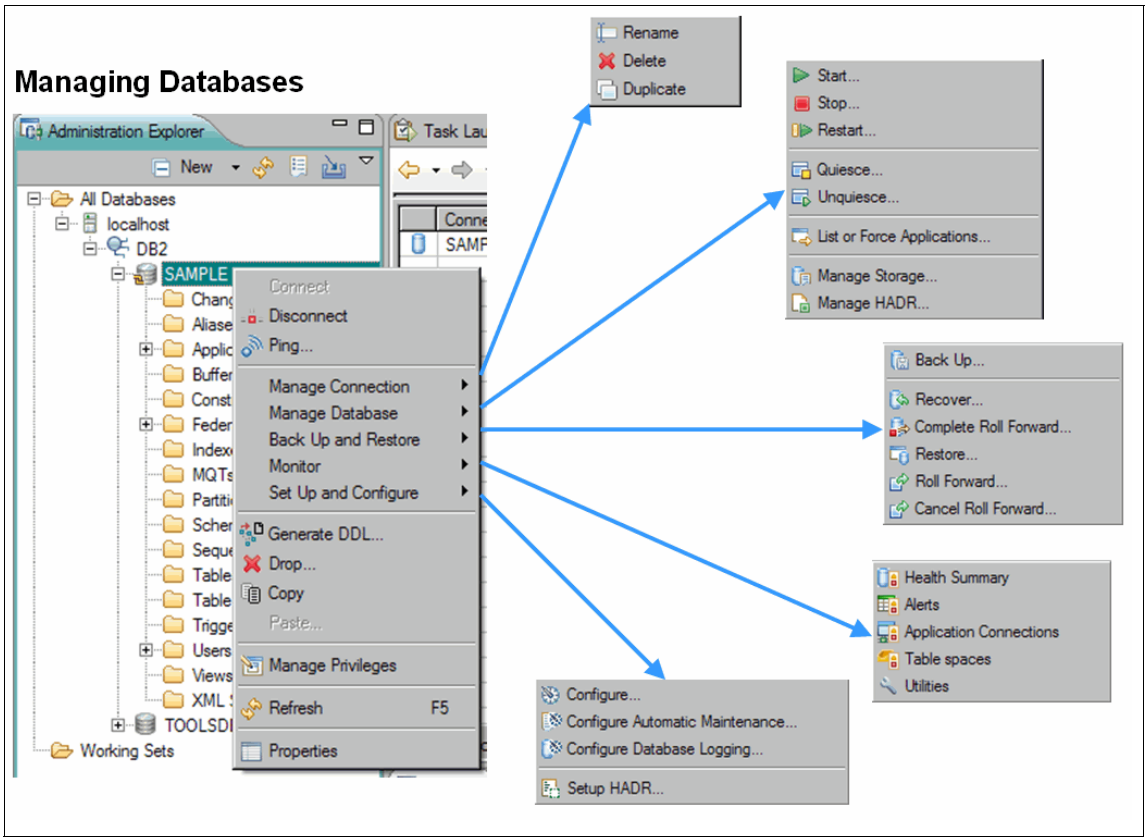


Figure 5-28 Managing databases options

Managing database objects

Database objects are grouped by type in the Administration Explorer view. If you select one type of object, the list of objects displays in the Editors area. For example, in Figure 5-29, the Tables objects are selected, and the list of tables is displayed, along with their basic information, such as name and row count. The navigation in the object list is powerful. It includes a multi-purpose navigation bar that provides flexibility on how you can explore database objects.

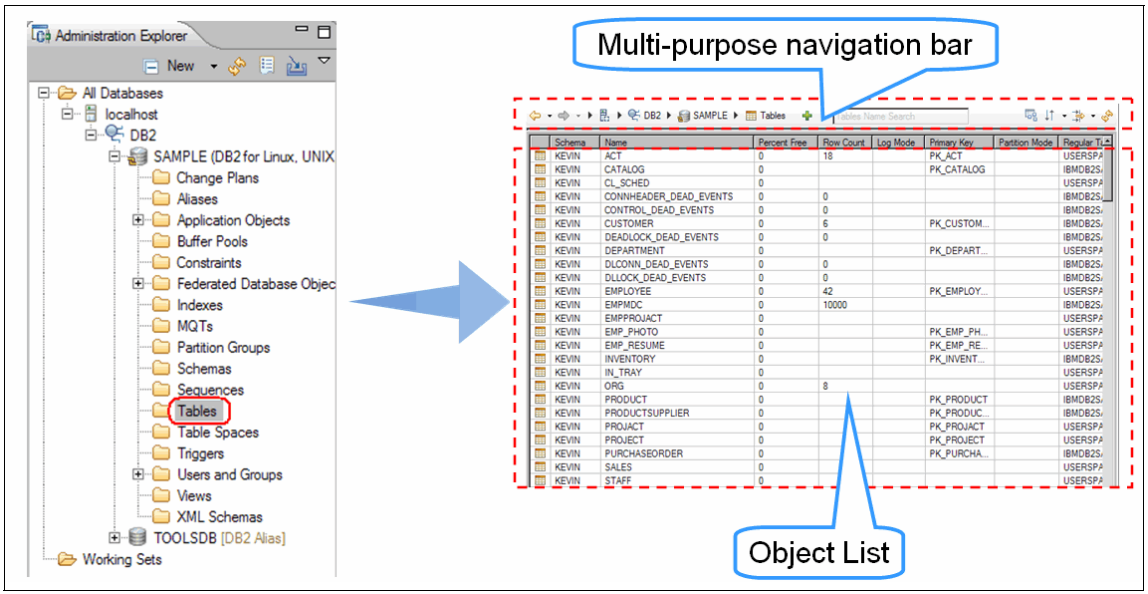


Figure 5-29 Exploration of tables category and multi-purpose navigation bar

ETL development services with Design Studio

After the data models are deployed in the data warehouse, you might need to manipulate data using data transformation processes to comply with the analytical business requirements. The developer can use Design Studio to design data flows that perform high-level login SQL extractions and transformations, such as aggregation, basic conversion, joins, normalization, merge, pivoting, sorting, and so on. The code produced by the data flow is run and optimized within the DB2 database to achieve the best query plan possible.

After the data flows are designed and tested, they can be wrapped in control flows that describe their logical execution, as shown in Figure 5-30 on page 201. These control flows are a container model that sequences one or more data flows and integrates other data processing rules and activities. Many control flows can be packaged inside a warehouse application for deployment and scheduling. The warehouse applications from InfoSphere Warehouse can be easily deployed and administered using the Administration Console.

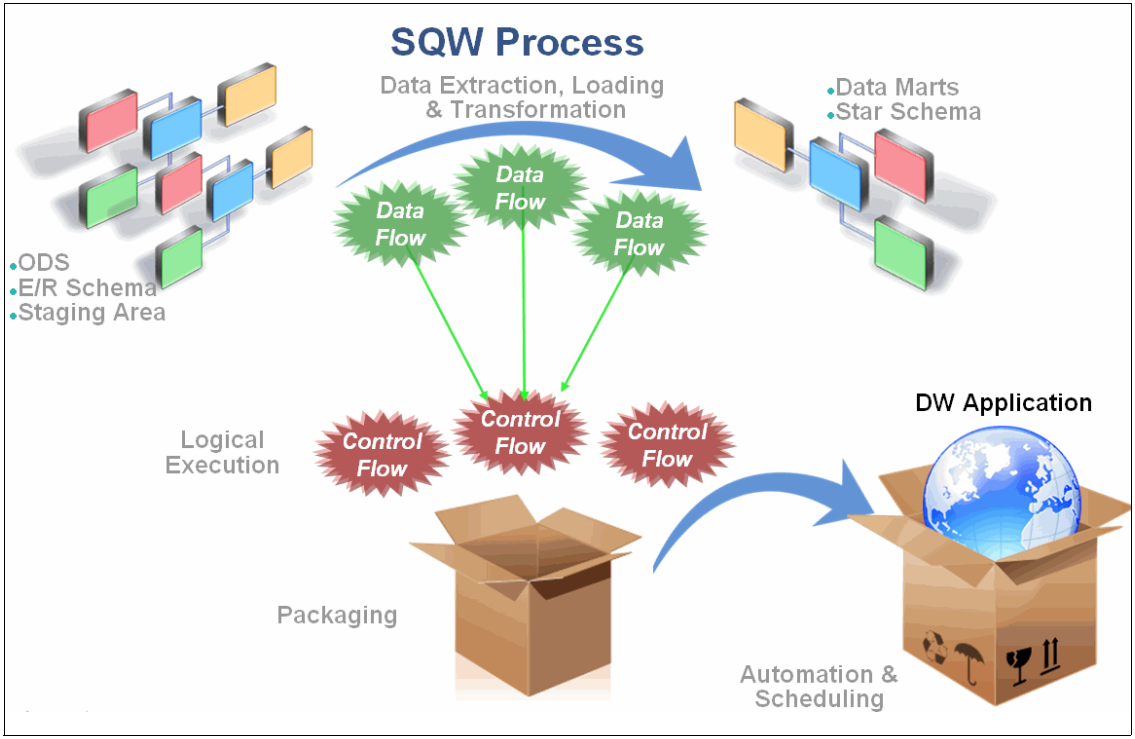


Figure 5-30 SQL Warehousing process

When you are setting up an SQL warehousing environment, make sure that you can connect to the local and remote databases where your source, staging, and target tables reside, especially the DB2 databases that you intend to use as SQL execution databases. The execution database is where all work (JOINS, UNIONS, and other SQL functions) occurs.

Create a data warehousing project

Before you design SQL Warehousing flows, you must create a data warehousing project in which to store design objects. You can store various types of design objects in a data warehousing project, including data flows, control flows, mining flows, text analysis, variables, and SQL scripts.

To create a data warehousing project:

1. Click **File** → **New** → **Data Warehousing Project**.
2. Provide a project name in the New Data Warehousing Project wizard.
3. If the Data Warehouse Project is going to access data models from a Data Design Project, select **Data Design Project as Referenced project**.

4. Click **Finish**. The new project displays in the Data Project Explorer view. Expand the project to see the subfolders from the tree.

Create a data flow

To create a data flow:

1. In the Data Project Explorer, under a data warehousing project, right-click the **Data Flows** folder. Select **New** → **Data Flow**.
2. Complete the New Data Flow wizard by specifying the parameters described in Table 5-7. Click **Finish**.

Table 5-7 New Data Flow wizard field parameters

Field	Parameter description	Parameter example
Select a data warehousing project	Select the project and folder where the data flow should be created.	MyDWProject
Data flow name	Type the name of the data flow.	MyDataFlow
Select the working mode for the flow	A data flow can work online or offline. If you choose online, the data flow needs a live connection to the databases used in the data flow. If you choose offline, the data flow works using the data models from the project.	Online
SQL execution database	The in-database approach allows you to perform fast transformations (because the data is all stored in the local database and you can take advantage of the DB2 optimization technology), and limit the amount of external data access that is performed. Therefore, every Data Flow is linked to a SQL DB2 Execution Database. The SQL Execution Database does the transformation work when the SQL code in a data warehousing application runs. This database must be a DB2 database.	The SQL Execution Database does not need to be the same database that is used to access source tables, or the target database that is loaded when applications run. You can use the approach that best fits yours needs.

3. Drag operators from the palette into the canvas.

   Operators are chosen from the palette and dragged to the work area canvas. Operators typically fall into the following groups:

   – *Sources* normally represent a table, but they can also be a flat file. The tables can be DB2 tables or tables in other supported relational database systems that can be accessed through JDBC or a federated nickname.

   – *Targets* are also normally a table, either in DB2 or another relational database system that supports JDBC, but they can also be a flat file.

   Figure 5-31 displays the source and target operators available.

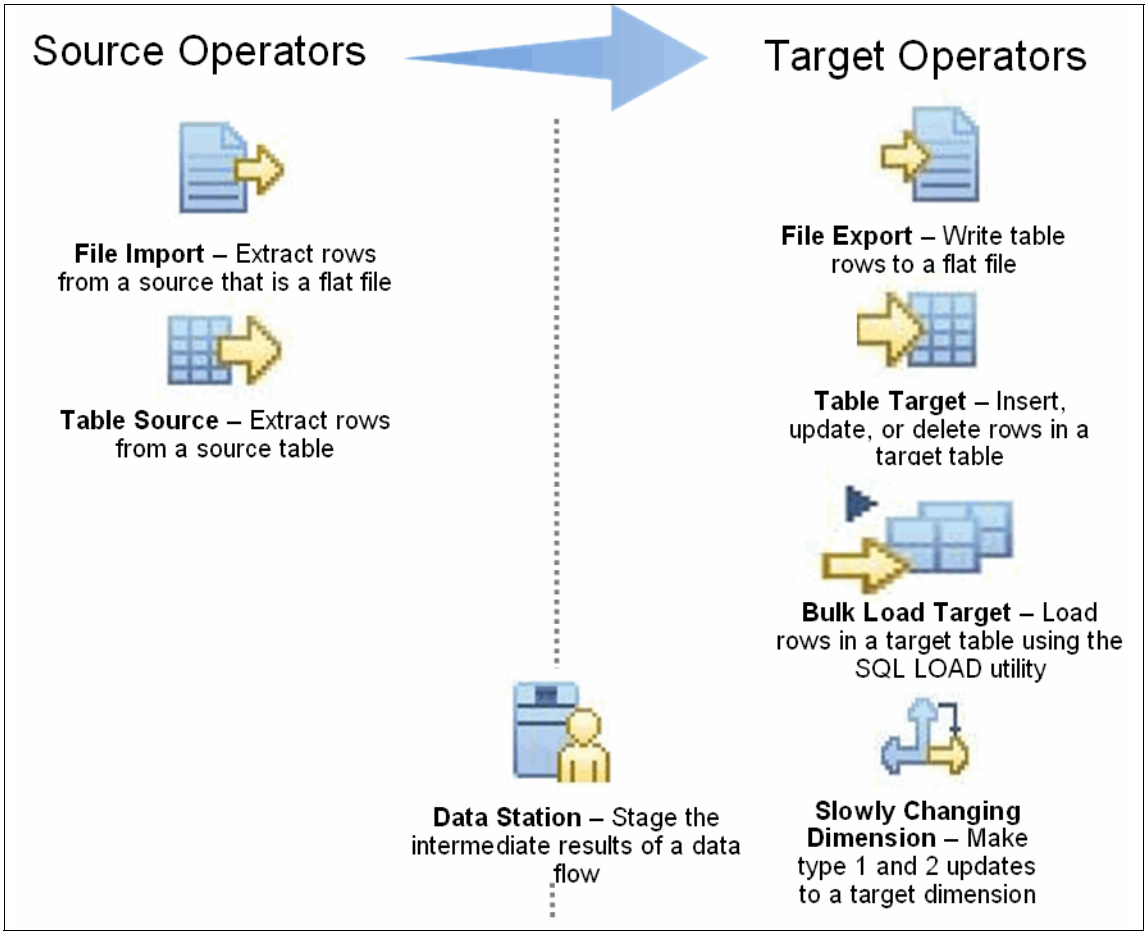


Figure 5-31 Source and target operators

– All of the supported processes that act upon the data as it is moved from a sources to a target would fall into the *transformers* category. These transformers, as shown in Figure 5-32, might be anything from basic SQL (a GROUP BY or a JOIN) to something more complex, such as using a PIVOT operator or updating a slowing changing dimension.

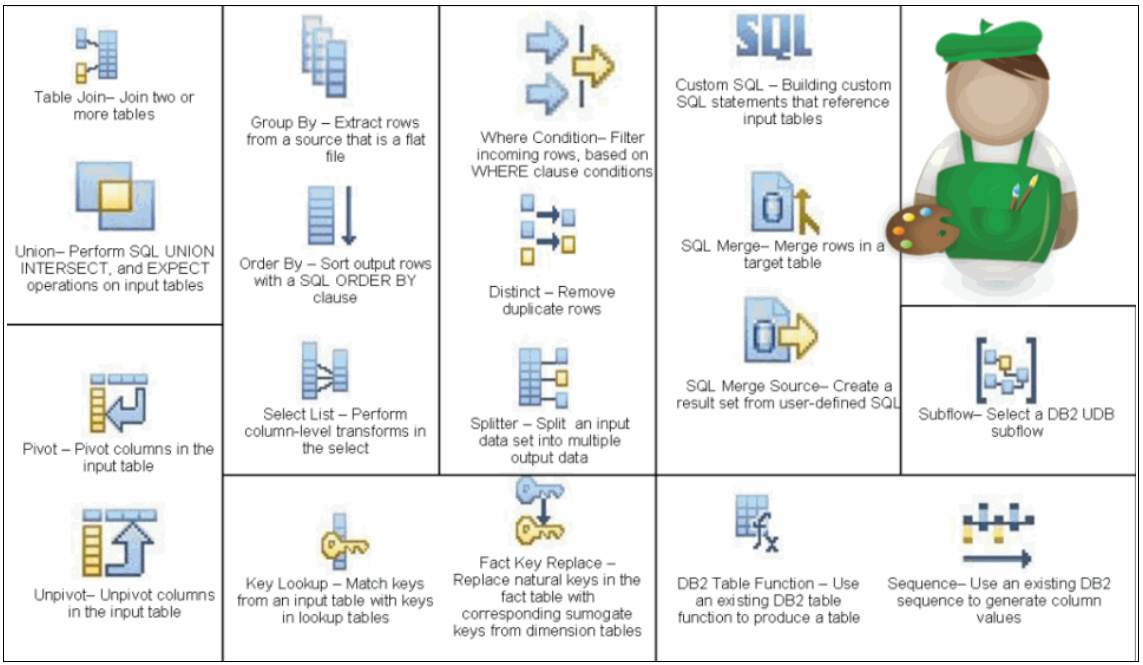


Figure 5-32 Data flow transformation operators

4. Connect the operators as the example in Figure 5-33 shows and define operator properties.
5. Save, validate, and debug the data flow.

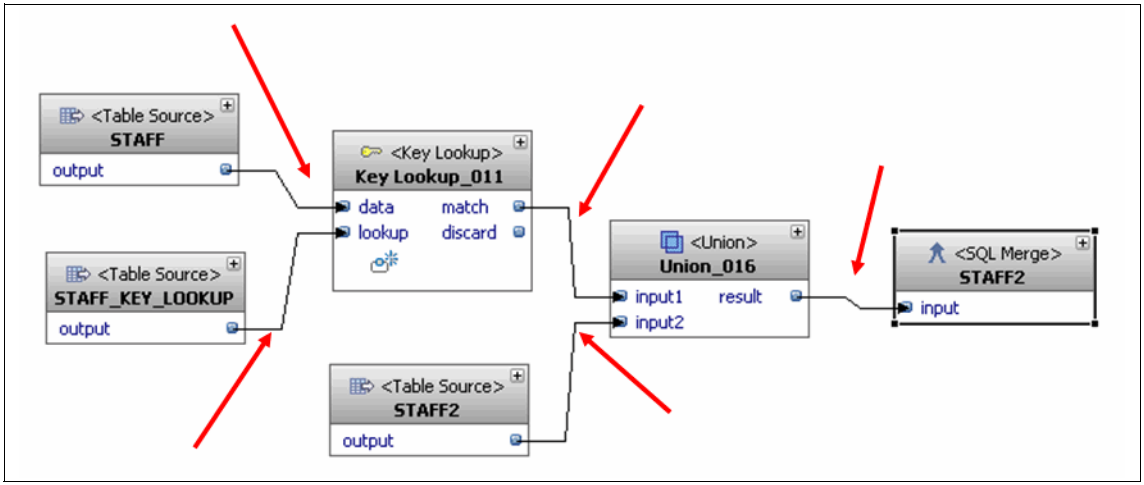


Figure 5-33 Example of a data flow

Create a control flow

To create a control flow:

1. In the Data Project Explorer, under a Data Warehousing Project, right-click the **Control Flows** folder. Select **New** → **Control Flow**.
2. Complete the New Control Flow wizard as described in Table 5-8 and click **Finish**.

Table 5-8 New Control Flow wizard field settings

Field	Setting Description	Setting Example
Select a data warehousing project	Select the project and folder where to create the data flow	MyDWProject
Control flow name	Type the name of the control flow	MyControlFlow

3. Drag operators from the palette into the canvas.

   Common operators include data flows, mining flows, FTP, email, iterators, and stored procedures. More complex operators include two for invoking an external DataStage Job, one for running a number of operators in parallel, and a number of operators that allow various forms of flow control.

Most control flow operators represent individual business activities within a process. There are two types of operators:

- Operators that perform activities in the runtime environment, such as data flow and email
- Operators that define the control flow process structures, such as start and end

Figure 5-34 displays a list of control flow operators.

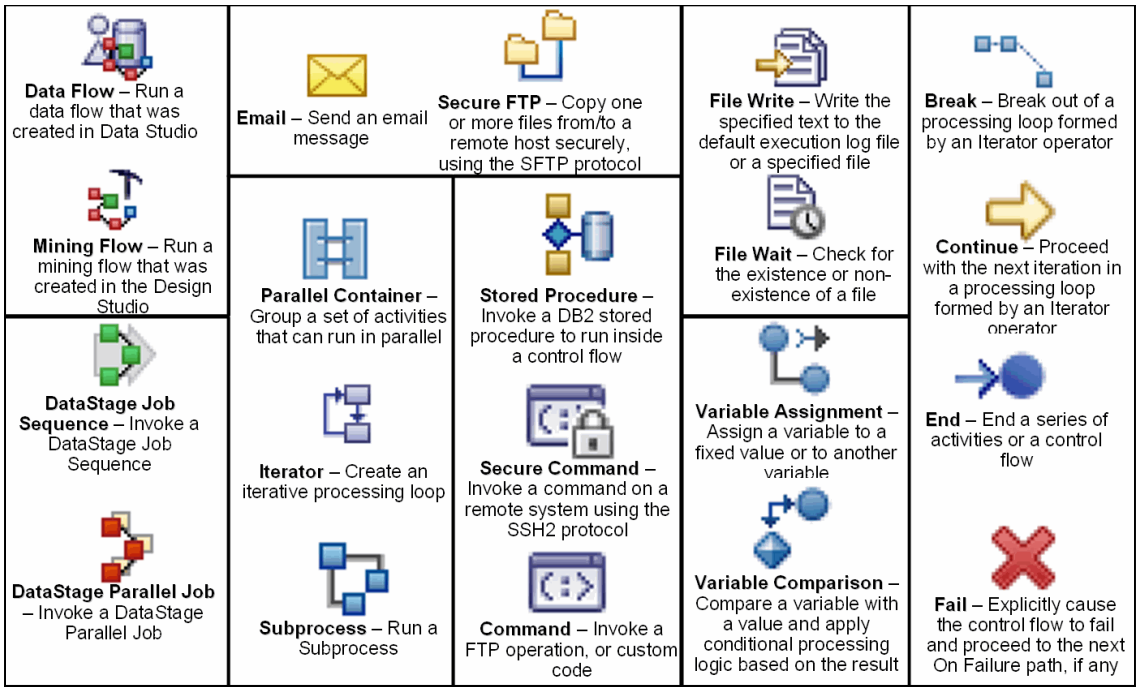


Figure 5-34 Control flow operators

Figure 5-35 shows additional control flow operators and icons that allow the inclusion of a number of DB2 operations as part of a control flow. The actual functions provided are mostly those that are used by a DBA, but they also can be of use in many application control flows.

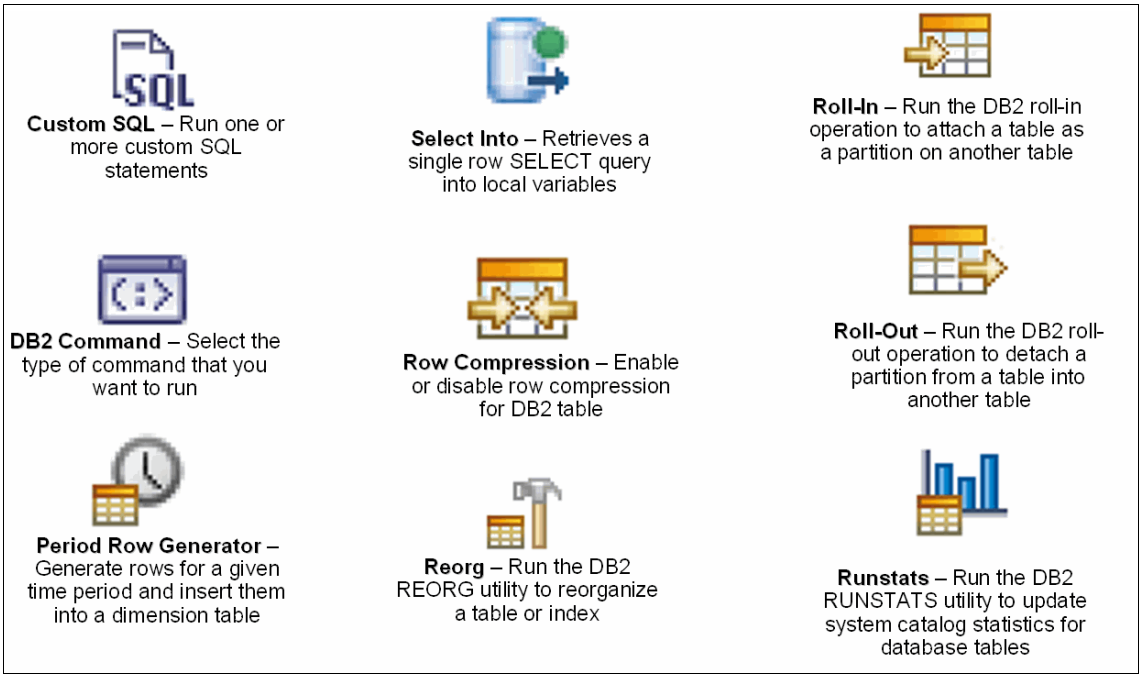


Figure 5-35 Control flow operators for DB2

4. Connect the operators as in the example shown in Figure 5-36.

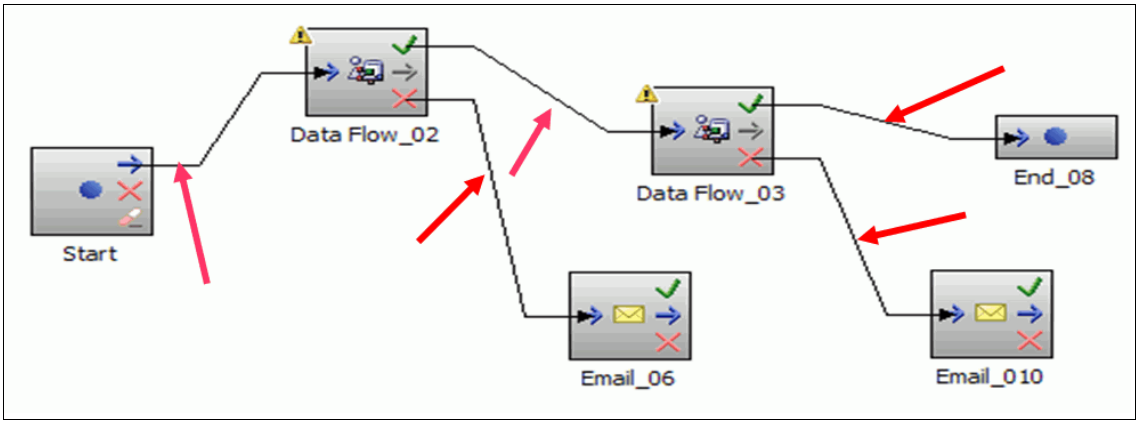


Figure 5-36 Example of a control flow

Only one start operator is allowed per control flow and it is generated automatically when you first create the control flow. End operators are optional and can be used anywhere to terminate a portion of the control flow. The exit ports for the Start Operator, as shown in Figure 5-37, are defined differently than other operators. These exits are used depending upon the success or failure of the whole control flow.

The Start Process port points to the first operation to be executed in the control flow. The Process On-Failure port gains control when a failure has occurred anywhere in the control flow. When it does take control, the rest of the control flow is not processed. The Cleanup Process port is invoked whenever a branch reaches a terminal state. In many cases, this points to a data flow operator that performs database clean-up processes, such as dropping interim tables.

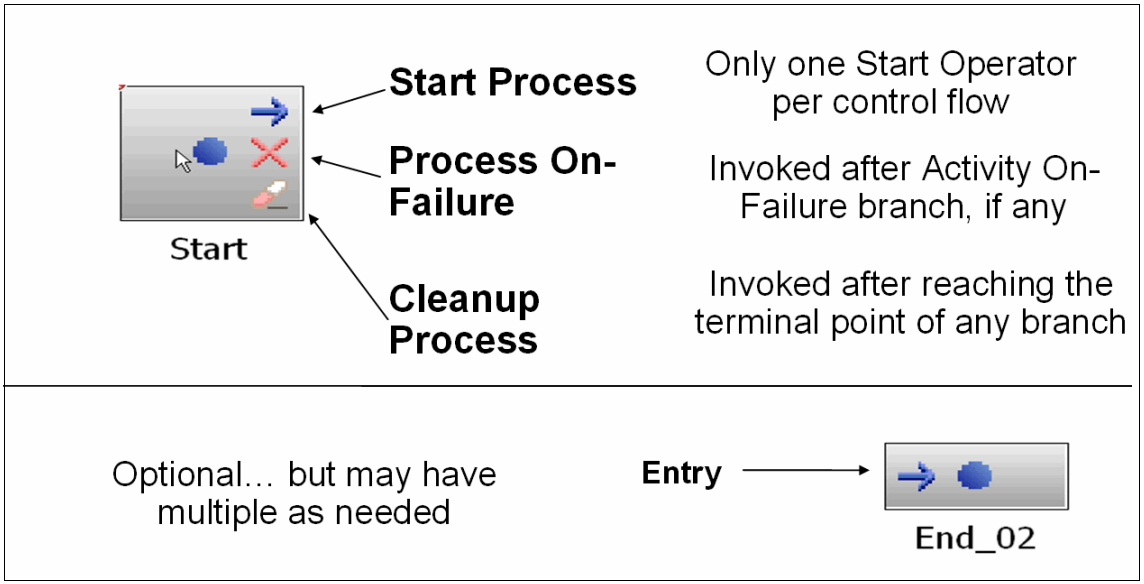


Figure 5-37 Star and end operator ports

Similar to data flows, the ports of an operator define entry and exit points. With the exception of the Start operator, each operator has one input port and three output ports (shown in Figure 5-38 on page 209) that always represent the same functions. The On-Success exit is used if the processes in the flow operator finish successfully. The On-Failure exit is used if errors occur in any of the processes. The Unconditional exit overrides both the on-success and on-failure exits and is used unconditionally to pass control regardless of the success or failure of the operator.

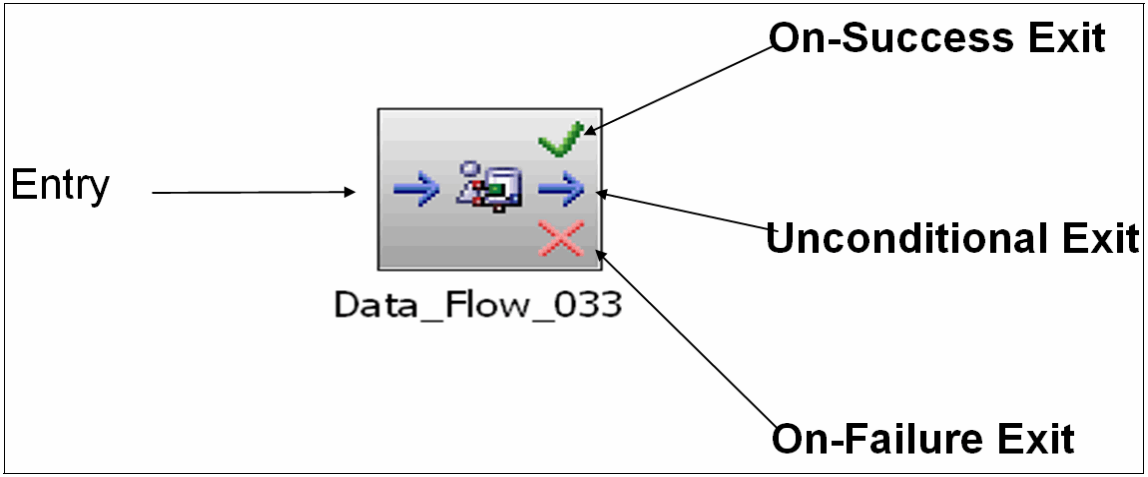


Figure 5-38 Operator ports

5. Define the operator's properties.
6. Save, validate, and debug the control flow.

Create a Data Warehousing Application package

After you have created some control flows, proceed with the following steps to create a Data Warehouse Application package:

1. In the Data Project Explorer view, right-click a project, and select **New** → **Data Warehousing Application**.
2. Provide a name and a location, and click **Finish.** A data warehouse application is packaged in a compressed file.
3. This package can be imported into the Administration Console for execution and scheduling; see "ETL and Data Warehouse services catalog" on page 220.

5.3.3 Business intelligence services catalog

Additional design and development might be necessary if you plan to use analytics against the data, such as designing mining models, text analysis annotators, cube models, and cubes in Design Studio.

OLAP modeling services with Design Studio

To create cubes for OLAP, you need to create cube models, which are the logical abstraction of the business information and contain all OLAP metadata, such as cubes, hierarchies, levels, and attributes, as shown in Figure 5-39. Most of the

information can be reverse engineered and mapped directly into the star schemas from the physical database. After creating the OLAP metadata in Design Studio, the information is saved in the InfoSphere Warehouse repository database.

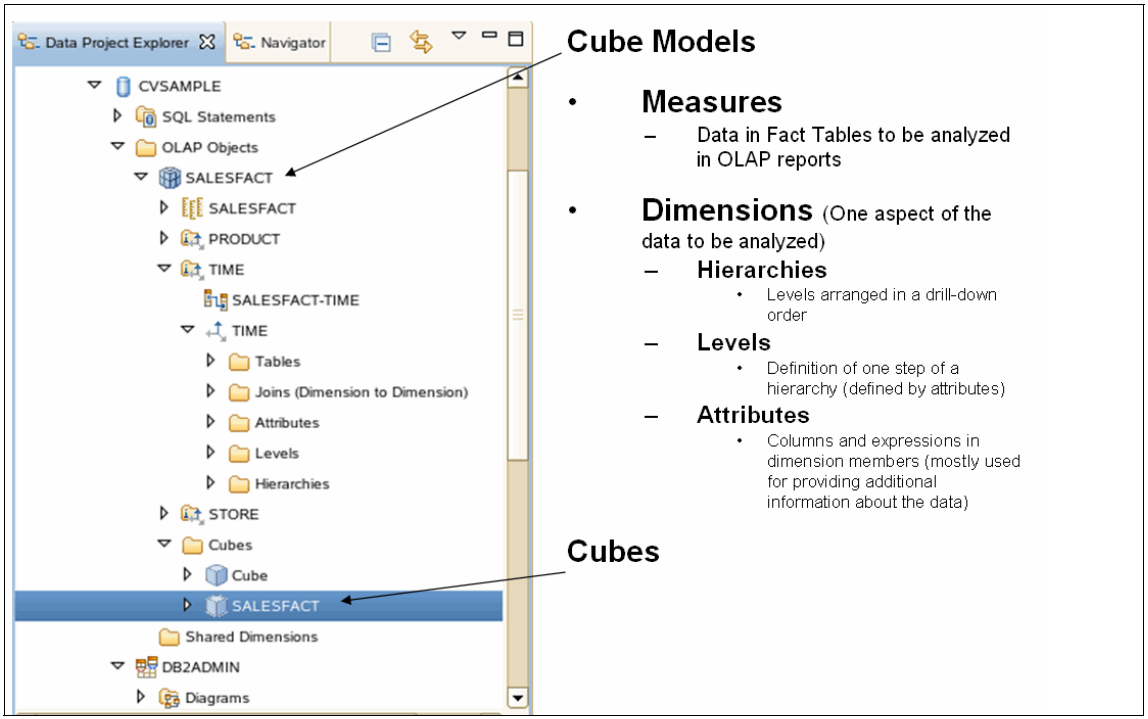


Figure 5-39 OLAP metadata for cube modeling in Design Studio

Data mining with Design Studio

InfoSphere Warehouse can also provide a more advanced capability to the warehouse developer, such as data mining. InfoSphere Intelligent Miner is the data mining component that is embedded into DB2 as extenders. Thus, after a database is enabled for mining, a number of user-defined functions, user-defined methods, stored procedures, and tables are created in the database. You can then use this database to build and test models and to score data in DB2 tables.

Modeling is the process of creating mining models based upon existing data. The following major grouping of data mining models are available in InfoSphere Warehouse:

- Discovery methods

  These methods helps the cloud consumer to find patterns in data that can improve decision making. The discovery methods are divided into:

  - Clustering

    A clustering model groups data values into various containers (or buckets) that can be further analyzed with other information within the group. You can specify a number of parameters (including the number of buckets you want to use), or let the modeler perform the clustering automatically.

    This method can use the following algorithms:

    - Demographic
    - Kohonen neural

    A Demographic algorithm can work better when most variables are categorical, and the Kohonen neural algorithm can work better when most variables are numeric. If you are using the Demographic algorithm and it yields one large cluster whose size cannot be satisfactorily reduced by increasing the similarity threshold, the Kohonen neural algorithm might yield a better model.

  - Associations

    Associations is probably the most familiar discovery method to find associations among data items, based on the number of times the association rules appear on a set of transactions.

  - Sequential rules

    Sequential rules develop sequence rules to identify affinities (or conditions, outcomes, and so on) of data items and the order in which they occur over time.

- Predictive methods

  These methods can help the cloud consumer to predict trends in data and possible outcomes. These methods can be trained << fine-tuned >>using a known data set to validate the model accuracy and quality. The predictive methods are divided into:

  - Classification

    Classification is a predictive technique to categorize data values by current and expected outcome. This method can use the following algorithms:

    - Decision tree
    - Naive Bayes

– Regression

  Regression is a method to uncover the factors that determine the class in which a record belongs. This method can use the following algorithms:

  - Transform
  - Linear
  - Polynomial

– Time series forecasting

  Time series forecasting is similar to regression but predicts values over a period of time. It can be used to predict both the overall trend and seasonal fluctuations.

Before a database can be used with Intelligent Miner, it should be enabled for mining. A database can be enabled for Data Mining in three different ways:

- From a DB2 command line
- From the Design Studio (mostly for Development Databases)
- From the Administration Console (mostly for Production Databases)

After the database is enabled for mining, follow these steps to design a data mining flow:

1. Right-click the Mining Flows folder and select **New** → **Mining Flow**.
2. The wizard is similar to the wizard used to create SQL Warehousing data flows. Provide a name and a mining database.
3. The Flow displays in the Project Explorer view with the name of the associated database connection next to it.
4. Add mining operators. Figure 5-40 is a list of the operators that are available in the data mining palette.

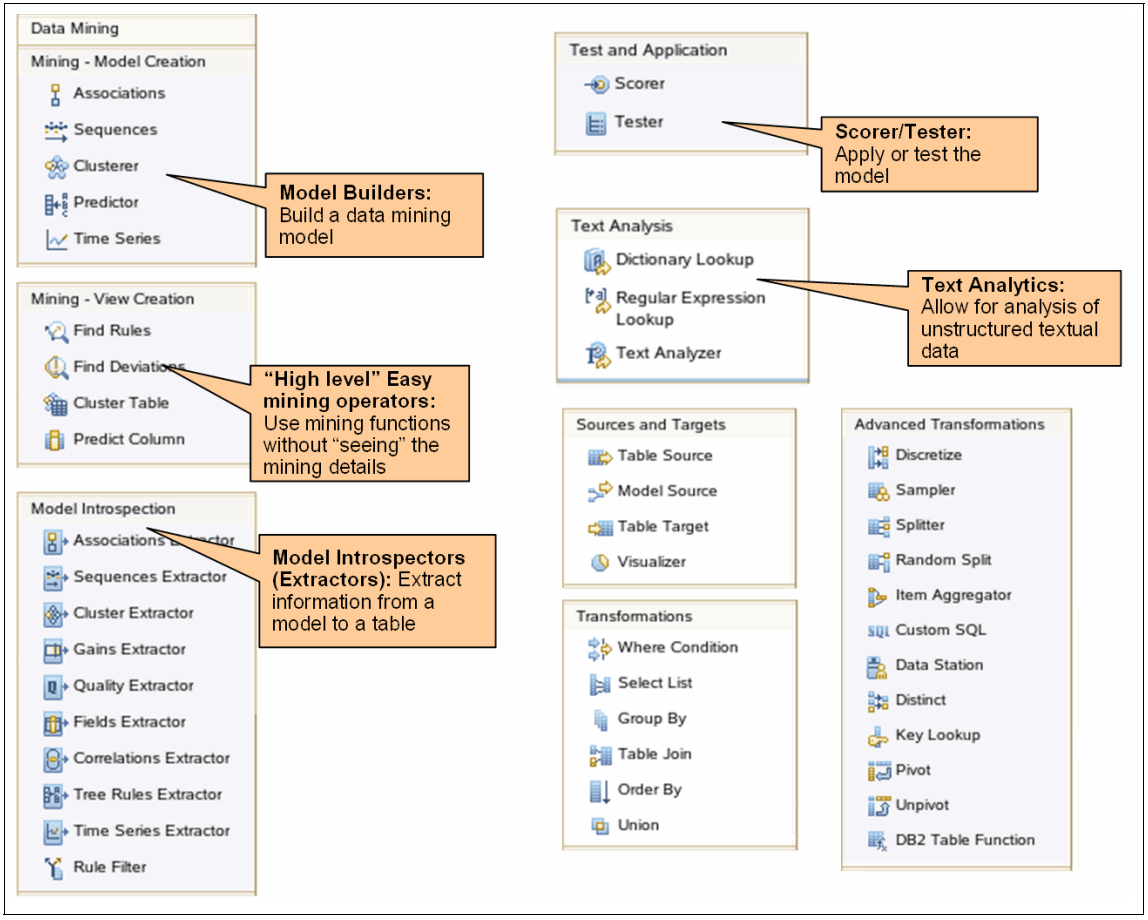


Figure 5-40 Data mining flow operators

5. Connect the operators and control their behavior in the Properties view. Figure 5-41 is an example of data mining.

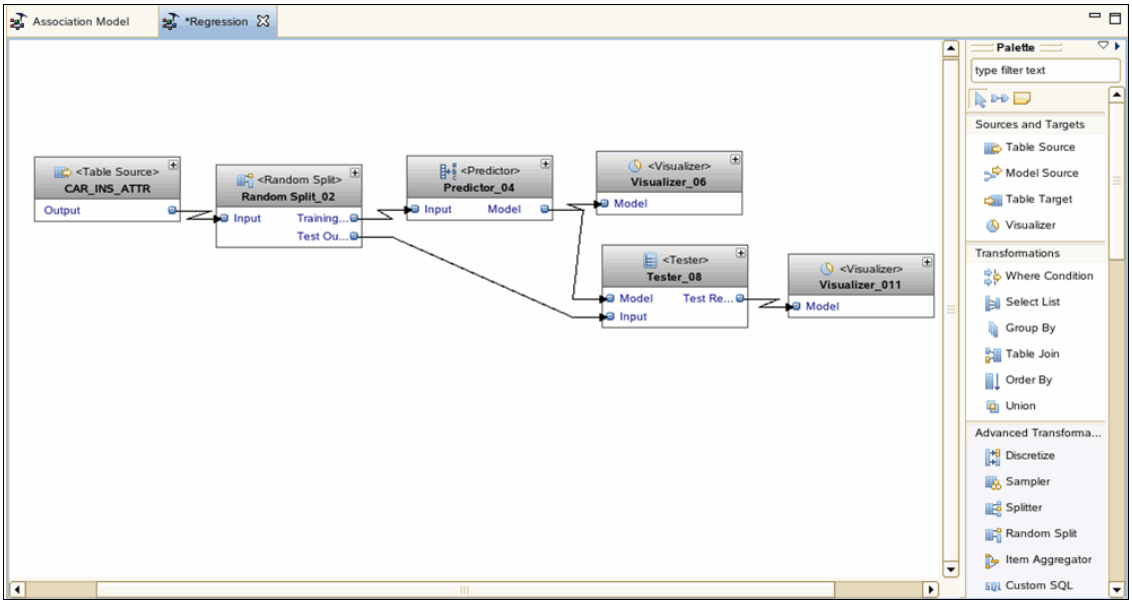


Figure 5-41 Example of a data mining flow to create a prediction data mining model

6. Save, validate, and debug the mining flows.
7. Run the mining flow. If you are using a visualizer operator, you can visualize the data mining model results.

Unstructured text modeling with Design Studio

Unstructured text analysis refers to the ability to visually identify and explore trends, patterns, and statistically relevant facts found in columns with unstructured text. Design Studio includes the following text analysis functions:

- Rule-based entity extraction based on regular expressions

  Extract words from text based on rule-based patterns. Use cases include telephone numbers, postal codes, credit card numbers, and so forth.

Figure 5-42 is an example of a rule-based pattern to extract geographic coordinates from a block of text.

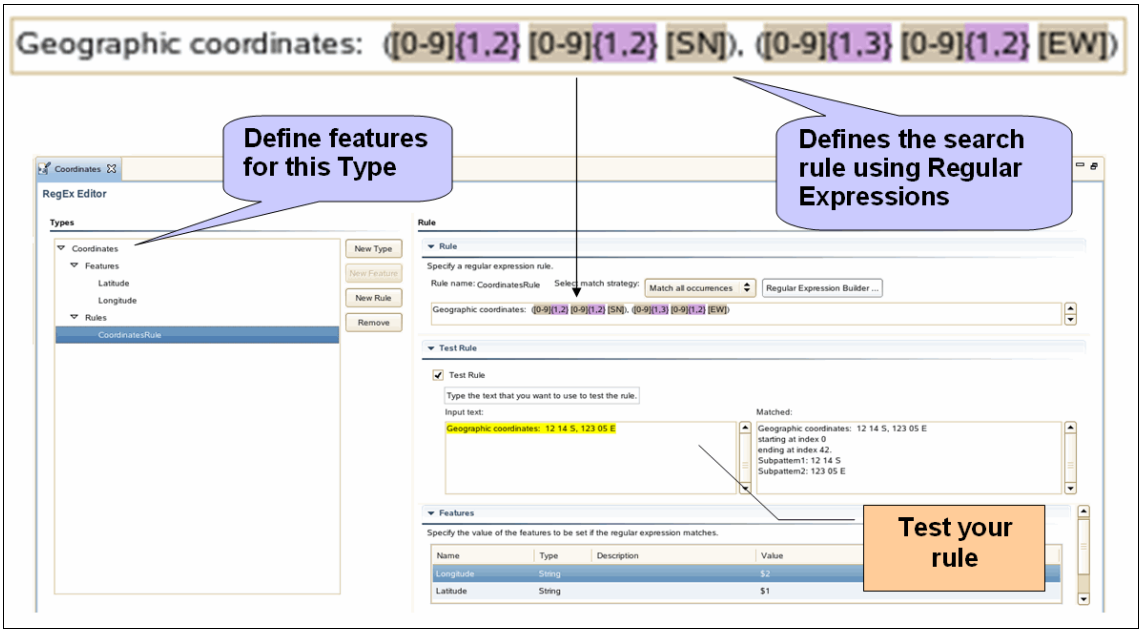


Figure 5-42 Example of a rule-based pattern to extract geographic coordinates from a block of text

- List-based entity is an extraction application combined with advanced linguistic text segmentation. Its capabilities include:
  - Search using a dictionary

    A dictionary is a list of all values that are of interest. Each term can be extracted and organized in the form of a dictionary, including the word's base form and its variants, as shown in Figure 5-43.

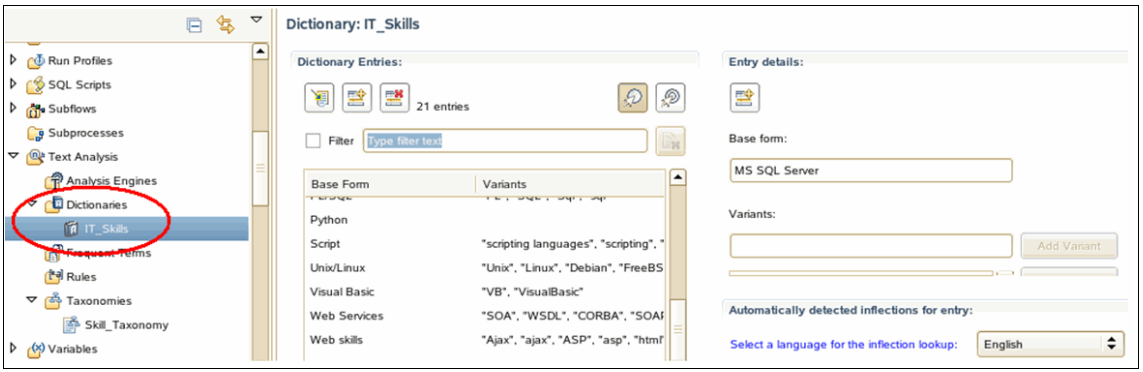


Figure 5-43 Example of a dictionary that lists computer-related words and variants

  - Frequent-terms analysis

    This analysis application counts term occurrences in a block of text that can be used to find the most repetitive words in the database and then to create a dictionary with the most frequent terms. Figure 5-44 on page 217 shows an example of frequent-terms analysis in Design Studio.

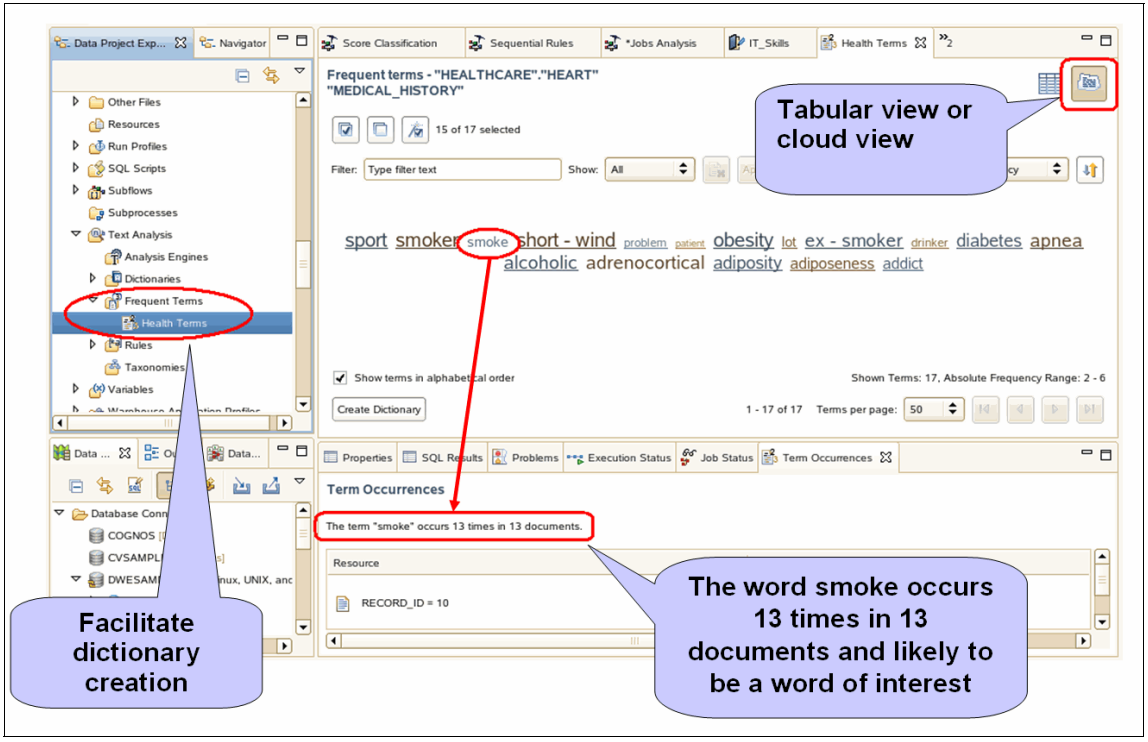


Figure 5-44 Example of the results found after the frequent term analysis

- Support for UIMA annotators

  Standard for including other annotators (text analysis functions)
- Combination of text analysis and data mining

  Using information extraction to create input for data mining
- Combination of text analysis and multidimensional analysis

  Creating additional dimensions from text data.

Model packages publishing with Cognos Framework Manager

Cognos Framework Manager is used to integrate the InfoSphere Warehouse and Cognos environments. It allows the developer to create and maintain a metadata connection to InfoSphere Warehouse databases and Cubing Services. It also facilitates the publication of Cognos packages ready to be consumed by Cognos services, as shown in Figure 5-45.

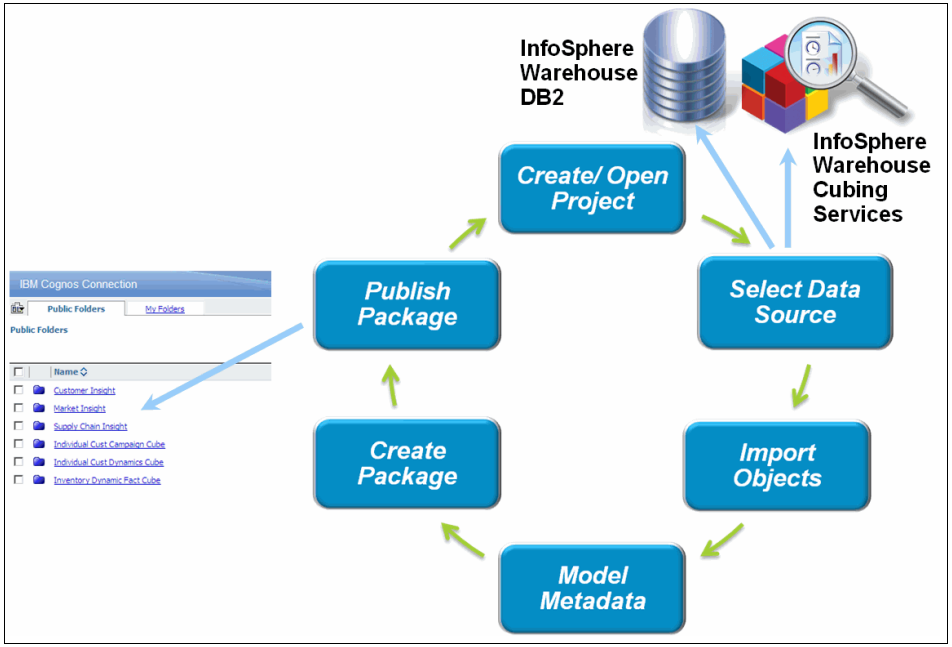


Figure 5-45 Steps to publish a package in Cognos Framework Manager

Create a data source to DB2

To report data from the DB2 tables using SQL, follow these steps:

1. Right-click the model namespace, then click **Run Metadata Wizard**.
2. In the Metadata Wizard, select the source from which to import the metadata. Ensure that Data Sources is selected, then click **Next**.
3. Click **New**.
4. In the Welcome page of the New Data Source Wizard, click **Next**.
5. Enter the name of your data source connection and click **Next**.
6. In the Type list, select IBM DB2, leave the default isolation level, and click **Next**.
7. In the Database name box, enter the name of the database to which you are connecting. Then, under Signons, select the **Password** option and enter the DB2 user name and password.
8. Click **Test the connection**, then click **Test**. When a message indicates that the test succeeded, click **Close**. Click **Close** again.
9. Click **Finish**, and then click **Close**. A new data source now displays in the list.

Create a data source to Cubing Services

If you need to analyze data from the Cubing Services using MDX, follow these steps:

1. Right-click the Model namespace, then click **Run Metadata Wizard**.
2. In the Metadata Wizard, select the source from which to import the metadata. Ensure that Data Sources is selected, and click **Next**.
3. Click **New**.
4. In the Welcome page of the New Data Source Wizard, click **Next**.
5. Enter the name of your data source connection and click **Next**.
6. In the Type list, select **IBM InfoSphere Warehouse cubing services (XMLA)**, leave the default isolation level, and click **Next**.
7. In the Server URL box, enter the URL of the cube server in the format `http://<hostname>:<cube server xmla port>/IBMXmlAnalysis` or `https://<hostname>:<cube server xmla port>/IBMXmlAnalysis/`.

   An example of `<hostname>:<cube server xmla port>` can be `iswcube1:80`.
8. To use secure sockets, select this option. To use unsecure sockets, do not select it. The setting must be the same as the setting on the server.
9. Under Signons, if a user ID and password is required in the connection string, select the User ID and Password options. Enter the cube user ID, and in the Password and Confirm Password boxes, enter the cube password.
10. Click **Test the connection**, then click **Test**. When a message displays that the test succeeded, click **Close**, and then click **Close** again.
11. Click **Finish**, then click **Close**. A new data source now displays in the list.

Create and publish a package

If you want your users to have access to the underlying data source, follow these steps to publish the framework manager package SQL:

1. In the Project Viewer pane, right-click **Packages** → **Create** → **Package**.
2. In the Create Package Wizard, enter a name for the package and click **Next**.
3. Expand and select the objects that you want to include in your package and click **Next**.
4. Here you have the option to specify which function sets are published with the package. Click **Finish**.
5. A dialog displays indicating that the package was created successfully and prompts you to open the Publish Package wizard. Publishing places a package in the IBM Cognos Content Store, where it can be access from IBM Cognos Connection and the IBM Cognos Studios. Click **Yes**.

6. IBM Cognos allows you to maintain multiple versions of the package. This is done to prevent breaking reports when re-publishing a package with changes that affect the reports. Click **Next**.
7. On the Add Security page, click **Next**, and then click **Publish**. The package is verified by default before the package in published.
8. When verified and published, a message indicates that the package was published successfully.
9. Click **Finish** to close the wizard, then save the project.

5.3.4 ETL and Data Warehouse services catalog

The data models structures, data flows, and control flows can be standardized toward different business purposes. For example, it is possible to have specific data flows for a specific industry or business purpose that can be deployed many times for different cloud consumers in the same industry.

In the InfoSphere Warehouse Administration Console, from the SQL Warehousing tab shown in Figure 5-46, you can deploy warehouse applications that you have created in the Design Studio. Select **Deploy** under the **Manage Applications** tab. A wizard prompts you for all the necessary information to deploy the application (including the location of the compressed file that was packaged in Design Studio).

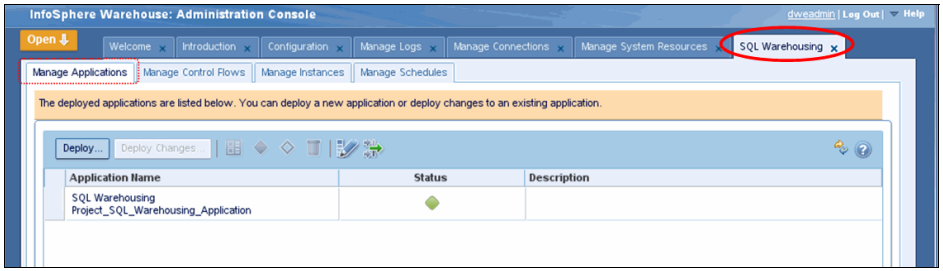


Figure 5-46 Manage Applications tab in the Administration Console

Within each control flow, specific values can be changed to update the control flow variables. The administration console also provides the means to monitor the execution status and result status of SQL Warehousing control flows. Log monitoring and searching is also supported to investigate error results or conditions.

5.3.5 Business intelligence services catalog

When business users think about business intelligence, they are not thinking about database structures or data models. They need to ask basic business-oriented questions that can make the difference between success and failure. Accessing pre-built business intelligence through a SaaS cloud model could provide a fast and flexible way of accessing an enriched decision support system from consolidated data stored in a data warehouse. Reporting tools can take advantage of cubes by sending MDX queries to cubing services as well as sending traditional SQL queries to the warehouse tables.

OLAP with InfoSphere Warehouse Cubing Services

OLAP allows you to analyze data from many different perspectives. It allows you to easily drill up, drill down, and slice and dice through the data, facilitating your ability to retrieve meaningful information about your business.

To provide OLAP processing and drill-down capability, prebuilt cube services and cubes must be started in the Administration Console. Cubing services cubes are built based on star schemas from the data warehouse and reside in memory (or a combination of memory and disk) for fast query access.

For more information, see *Multidimensional Analytics: Delivered with InfoSphere Warehouse Cubing Services*, SG24-7679.

Reporting services with IBM Cognos Business Intelligence

Cognos Business Intelligence software fits the reporting needs of all users in your organization, from business users who need to create their own ad-hoc queries, to professional report authors responsible for designing one-to-many reports, to IT administrators who must deploy and manage the application in a centralized, streamlined manner. IBM Cognos Business Intelligence 10.1 provides the following report services:

- Ad-hoc reporting

  Query Studio allows business users to execute ad-hoc queries and author basic reports with capabilities to filter, sort data, add formatting, and create charts.

- Professional reporting

  Report Studio enables professional report authors to produce professional quality content for others to consume as a service. These advanced report building and format capabilities enable enhancement, customization, and management of professional reports. This type of reporting is intended for creating reports that need to be consumed for large volumes of users and its output a directed to a number of different file types.

- Analytical reporting

  Analysis Studio allows users to perform multidimensional analysis, with full flexibility to drill down to increasingly lower levels of details and drill through to other reports to get a better view of the overall picture. Using InfoSphere Warehouse Cubing Services and Cognos Analysis Studio allows a combination of analysis with high performance.

- Active reporting

  Cognos also extends the reach to include non-BI consumers who are not connected to the infrastructure. IBM Cognos Active Report enables users to access and interact with report and dashboard content, regardless of whether they are connected to the BI network.

- Assemble reporting

  Business Insight allows business users to author, modify, and share reports with minimal training or IT input. The flexible interface gives business users the ability to arrange report objects by dragging and dropping them into the report authoring window, in an easy-to-use interface, using trusted corporate information.

- Mobile reporting

  You can use IBM Cognos Mobile in the cloud environment to allow cloud consumers to receive their Cognos Business Intelligence content on a personal mobile device. Cognos Mobile uses Cognos Content and also its security. It allows the user to respond and initiate action immediately and supports disconnected content access.

For more information about Cognos Business Intelligence capabilities, refer to *IBM Cognos Business Intelligence V10.1 Handbook*, SG24-7912.

Based on feedback, monitoring, and refinement of requirements, the business intelligence lifecycle starts over as changes and corrections are identified and used to repeat the process to create new, revised applications and business intelligence structures.

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 publications

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.

- *WebSphere Application Server V8.5 Concepts, Planning, and Design Guide*, SG24-8022
- *WebSphere Application Server V8.5: Technical Overview Guide*, REDP-4855
- *IBM CloudBurst on System x*, SG24-7981
- *Virtualization with IBM Workload Deployer: Designing and Deploying Virtual Systems*, SG24-7967
- *Implementing IBM System Networking 10Gb Ethernet Switches*, SG24-7960
- *IBM Optim Performance Manager for DB2 for Linux, UNIX, and Windows*, SG24-7925
- *IBM Cognos Business Intelligence V10.1 Handbook*, SG24-7912
- *IBM Cognos Business Intelligence V10.1: Intelligence Unleashed*, REDP-4693
- *IBM Smart Analytics Cloud*, SG24-7873
- *InfoSphere Warehouse: A Robust Infrastructure for Business Intelligence*, SG24-7813
- *Using Integrated Data Management To Meet Service Level Objectives*, SG24-7769
- *Multidimensional Analytics: Delivered with InfoSphere Warehouse Cubing Services*, SG24-7679
- *DB2 Workload Manager for Linux, UNIX, and Windows*, SG24-7524
- *Eclipse Development using the Graphical Editing Framework and the Eclipse Modeling Framework*, SG24-6302
- *IBM BladeCenter Foundation for Cloud: Integration Guide*, REDP-4773

- *IBM System x Private Cloud Offering: Solution and Component Guide*, REDP-4731
- *IBM System x Private Cloud Offering: Architecture and Implementation Guide*, REDP-4732
- *IBM System x Private Cloud Offering, Advanced Configuration: Architecture and Implementation Guide*, REDP-4771
- *Data Studio and DB2 for z/OS Stored Procedures*, REDP-4717
- *Adopting Cloud Computing using the WebSphere CloudBurst Appliance*, REDP-4708
- *Identity Management for IBM Cognos 8 with IBM Tivoli Identity Manager*, REDP-4643
- *Cloud Security Guidance IBM Recommendations for the Implementation of Cloud Security*, REDP-4614
- *Enabling Robust Business Analytics with InfoSphere Warehouse Cubing Services*, REDP-4514
- *TCP/IP Tutorial and Technical Overview*, GG24-3376
- *DB2 Virtualization,* SG24-7805

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:

- *Barrie Sosinsky, Cloud Computing Bible, John Wiley & Sons, 2011, ISBN 0470903568*

Online resources

These websites are also relevant as further information sources:

- IBM - Data Warehouse - InfoSphere Warehouse

  http://www-01.ibm.com/software/data/infosphere/warehouse/

- Executive Corner: Don't settle for the check mark | Thoughts on Cloud Blog

  http://thoughtsoncloud.com/index.php/2011/12/executive-corner-dont-settle-for-the-check-mark/

- Securing a multi-tenant SaaS application

  http://www.ibm.com/developerworks/java/library/j-saas/

- Convert your web application to a multi-tenant SaaS solution

  http://www.ibm.com/developerworks/cloud/library/cl-multitenantsaas/

- SmartCloud Provisioning

  http://www-142.ibm.com/software/products/us/en/smartcloud-provisioning

- Thought Leadership White Paper. Getting cloud computing right. April 2011

  http://public.dhe.ibm.com/common/ssi/ecm/en/ciw03078usen/CIW03078USEN.PDF

- CCRA IBM Submission.

  http://www.opengroup.org/cloudcomputing/uploads/40/23840/CCRA.IBMSubmission.02282011.doc

- IBM Cognos Business Intelligence 10.1.0 Information Center

  http://publib.boulder.ibm.com/infocenter/cbi/v10r1m0/index.jsp?topic=%2Fcom.ibm.swg.im.cognos.cbi.doc%2Fwelcome.html

- IBM Cognos Business Intelligence Version 10.1.1 Installation and Configuration Guide

  http://public.dhe.ibm.com/software/data/cognos/documentation/docs/en/10.1.1/inst_cr_winux.pdf

- IBM Data Studio v3.1.1 information center

  http://publib.boulder.ibm.com/infocenter/dstudio/v3r1/index.jsp?topic=%2Fcom.ibm.datatools.ds.release.doc%2Ftopics%2Fgetstarted.html

- IBM InfoSphere Warehouse edition comparison

  http://www-01.ibm.com/software/data/infosphere/warehouse/edition-comparison.html

- IBM InfoSphere Warehouse information center

  http://pic.dhe.ibm.com/infocenter/db2luw/v10r1/index.jsp?topic=%2Fcom.ibm.dwe.welcome.doc%2Fc_release_whatsnew.html

- Judith Myerson, Cloud computing versus grid computing. March 3, 2009.

  http://www.ibm.com/developerworks/web/library/wa-cloudgrid/

- Judith M. Myerson. Craft a cloud performance metrics policy. September 6, 2011.

  http://www.ibm.com/developerworks/cloud/library/cl-cloudperformmetrics/

- Judith M. Myerson. Craft a SaaS-oriented web application vulnerability mitigation policy. June 27, 2012

  http://www.ibm.com/developerworks/cloud/library/cl-saasmeasurepolicy/

- Cloud computing on IBM developerWorks

  http://www.ibm.com/developerworks/cloud/

- IBM Application Development Services for Cloud

  http://www.ibm.com/services/us/en/business-services/ibm-application-development-services-for-cloud.html

- Cognos Proven Practices Team, IBM Cognos Proven Practices: IBM Cognos 10.1 BI JMeter Stability Test. August 23, 2011

  http://www.ibm.com/developerworks/data/library/cognos/performance/load_testing/page584.html

- IBM Rational Performance Tester

  http://www.ibm.com/software/awdtools/tester/performance/

- developerWorks: Wikis - Systems - HOME

  https://www.ibm.com/developerworks/wikis/display/WikiPtype/HOME

- Solution Brief. Entry Cloud Reference Configuration for Mid Markets on System x

  http://public.dhe.ibm.com/common/ssi/ecm/en/xss03096usen/XSS03096USEN.PDF

- IBM Virtualization Policy for IBM Software

  http://www-01.ibm.com/software/support/virtualization_policy.html

- IBM Using IaaS: IBM SmartCloud Enterprise

  http://www.ibm.com/services/us/en/cloud-enterprise/uses-development-and-test.html

- Inside the hybrid cloud, Part 1

  http://www.ibm.com/developerworks/cloud/library/cl-hybridcloud1/index.html

- Executive Corner: Don’t settle for the check mark

  http://thoughtsoncloud.com/index.php/2011/12/executive-corner-dont-settle-for-the-check-mark/

- Getting cloud computing right

  http://public.dhe.ibm.com/common/ssi/ecm/en/ciw03078usen/CIW03078USEN.PDF

Help from IBM

IBM Support and downloads

ibm.com/support

IBM Global Services

ibm.com/services

IBM Redbooks

Virtualized Business Intelligence with InfoSphere Warehouse

(0.2”spine)
0.17”<->0.473”
90<->249 pages

Virtualized Business Intelligence with InfoSphere Warehouse

Bring high performance business intelligence into cloud

Manage thousands of users with mixed workloads

Use rich data management capabilities

With the benefit of advanced analytics such as online analytical processing (OLAP), data mining, and text analytics, the IBM InfoSphere Warehouse Enterprise Edition brings sophisticated business intelligence (BI) to warehouse users. InfoSphere Warehouse allows you to run extreme concurrent query volumes that can help answer questions for all types of business users, while consistently meeting service level requirements.

Combined with a virtualization platform and a solid BI solution, such as IBM Cognos, you can deliver BI cloud services with improved flexibility and speed to your clients, thereby presenting a new avenue for which your services can be offered.

This IBM Redbooks publication discusses the deployment of a BI cloud solution. It includes details such as understanding the architecture of a cloud, planning implementation, integrating various software components, and understanding the preferred practices of running a cloud deployment. Essentially, this book can be used as a guide by anyone who is interested in deploying a virtualized environment for a BI cloud solution.

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

SG24-8019-00 ISBN 0738437417