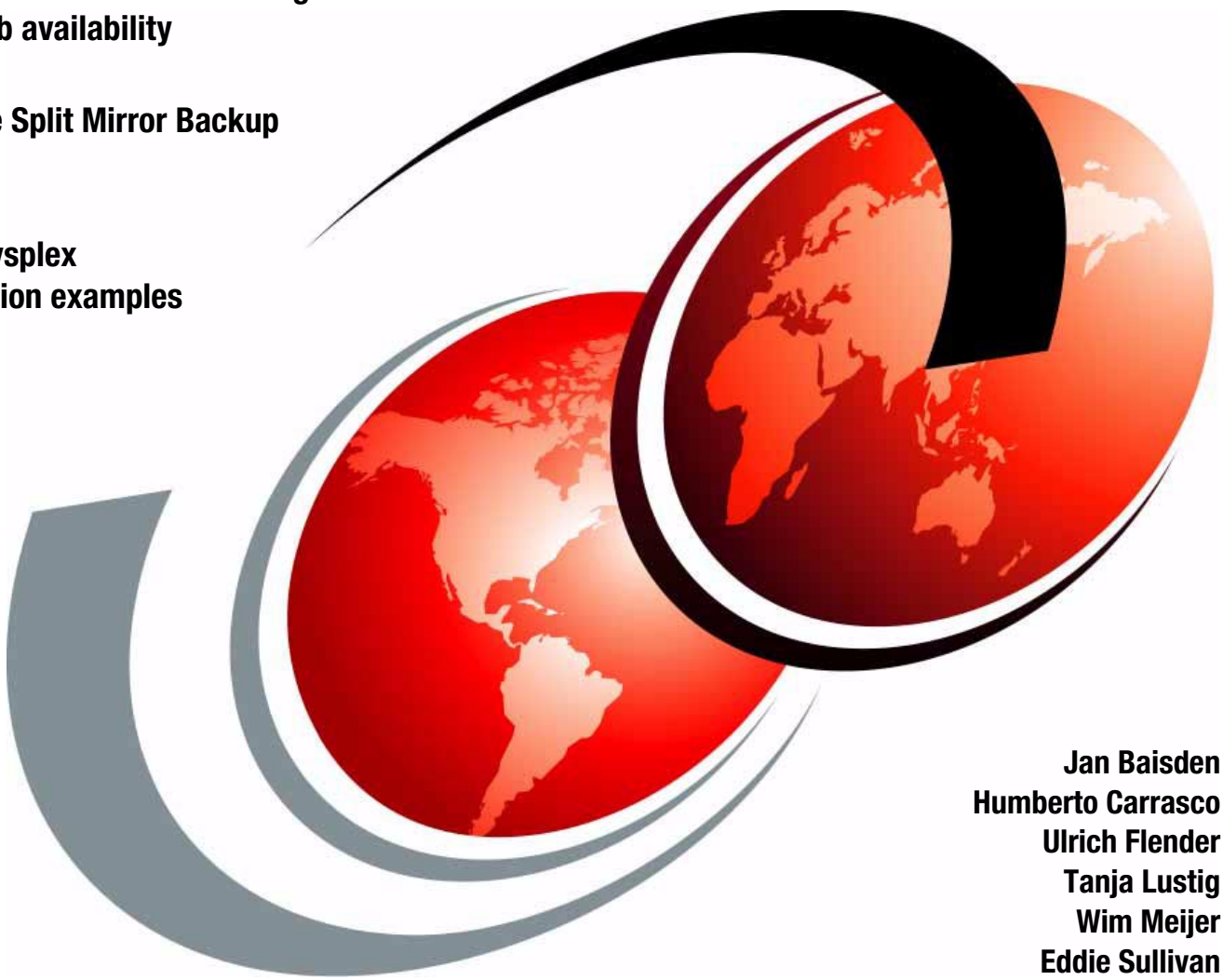
IBM

SAP R/3 on DB2 UDB for OS/390: Database Availability Considerations

- Define and use DB2 data sharing to improve db availability
- Exploit the Split Mirror Backup Facility
- Parallel Sysplex configuration examples



Jan Baisden
Humberto Carrasco
Ulrich Flender
Tanja Lustig
Wim Meijer
Eddie Sullivan

Redbooks

ibm.com/redbooks

International Technical Support Organization

SG24-5690-00

SAP R/3 on DB2 UDB for OS/390: Database Availability Considerations

February 2001

Take Note!

Before using this information and the product it supports, be sure to read the general information in Appendix D, “Special notices” on page 145.

First Edition (February 2001)

This edition applies to SAP R/3 4.6B for use with DB2 UDB for OS/390 Version 6 (5645-DB2) as used with OS/390 Release 2.8 (5645-001); external servers use AIX 4.3.3 (5765-603) as an operating system.

Comments may be addressed to:
IBM Corporation, International Technical Support Organization
Dept. HYJ Mail Station P099
2455 South Road
Poughkeepsie, NY 12601-5400

When you send information to IBM, you grant IBM a non-exclusive right to use or distribute the information in any way it believes appropriate without incurring any obligation to you.

Contents

Preface

This IBM Redbook describes availability procedures for the database of an SAP R/3 4.6B installation that uses DB2 UDB for OS/390 as the SAP R/3 database server. The redbook is one of a series that focuses on SAP R/3 when DB2 UDB for OS/390 is used as a database server.

This redbook reflects the considerable development effort, both by SAP AG and by IBM, that has been focused on the area of database availability due to the reliance on SAP R/3 by cutomers who expect a high degree of availability, in some cases 24 hours per day, every day of the year.

Since the subject of this book is database availability, we concentrate on the database that applies to this solution, DB2 UDB for OS/390. Much effort was spent on looking at data sharing with DB2 to see how that improves availability of the database, and to investigate how it is defined and used. We also review database administrative activity that must be periodically run, to see how that can be done with as little disruption as possible.

We also investigate the Split Mirror Backup Facility as defined by SAP R/3 and implemented with DB2 UDB for OS/390. A particular area of interest is the minimizing of time to perform a backup, since this regularly scheduled activity can be disruptive. Obviously, no discussion of backup is complete without a discussion of using the backups for recovery, so we address the subjects of full and partial recovery as well.

The team that wrote this redbook

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

Jan Baisden is a Senior Market Support Representative at the International Technical Support Organization, Poughkeepsie Center. He concentrates on ERP solutions and particularly on AIX support of S/390 usages in those solutions. He currently maintains five redbooks on SAP R/3 on DB2 for OS/390. Before joining the ITSO six years ago, Jan worked at the IBM International Systems Center in Gaithersburg, Maryland, supporting projects in Latin America, Asia, and Australia.

Humberto Carrasco Ancochea is an S/390 System Specialist and certified SAP Basis consultant in Spain. He has 4 years of experience in DB2 implementation and administration projects. He holds a degree in Physics from the University of Madrid. His areas of expertise include Data Sharing environments and SAP R/3 with DB2 on S/390.

Ulrich Flender is a systems programmer at DVG Hannover in Germany. He has 13 years of experience in customizing several system software components in OS/390. He holds a certificate degree in computer science. His areas of expertise include UNIX System Services and TCP/IP network access on OS/390.

Tanja Lustig is responsible for Technical Marketing of SAP R/3 on S/390 in the IBM SAP International Competence Center in Walldorf, Germany. Tanja is a

certified SAP R/3 Basis Consultant and has 4 years of experience with SAP on the AIX and S/390 platforms. She holds an economics degree with specialization in computer science and has worked at IBM for 7 years.

Wim Meijer is a Consulting IT Specialist in ERP Solutions of The Netherlands. Wim is a certified SAP Basis Consultant and has 5 years of experience with SAP. He has almost 20 years of experience in IT. His areas of experience includes, besides consulting on SAP infrastructure, installation, system administration, training, problem analysis and performance tuning of SAP systems. Wim worked for IBM partners and is now working for IBM for 1 year.

Eddie Sullivan is an SAP BASIS Administrator in the USA for IBM Global Services. Eddie has over 20 years IT experience in systems programming and database administration. His areas of expertise include SAP and DB2 on OS/390.

Nick Vallance is an SAP Basis Technician in Portsmouth, England. He has six years of experience in the S/390 systems programming field and is a certified SAP Basis Consultant. He holds a degree in Information Systems Management from Bournemouth University. His areas of expertise include Parallel Sysplex, CMOS and R/3. He has previously written on Complex Systems Availability and Recovery.

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

Vasilis Karras
International Technical Support Organization, Poughkeepsie Center

Bob Haimowitz
International Technical Support Organization, Poughkeepsie Center

Alfred Schwab
International Technical Support Organization, Poughkeepsie Center

Terry Barthel
International Technical Support Organization, Poughkeepsie Center

Andreas Maier
Joint SAP & IBM Porting Team, Walldorf, Germany

Namik Hrle
Joint SAP & IBM Porting Team, Walldorf, Germany

Don Geissler
IBM SAP Competency Center

Patrick Horkan
IBM SAP Competency Center

Michael Sheets
IBM SAP Competency Center

Michael Gordon
Gordon Consulting and IBM SAP Competency Center

Comments welcome

Your comments are important to us!

We want our Redbooks to be as helpful as possible. Please send us your comments about this or other Redbooks in one of the following ways:

- Fax the evaluation form found in “IBM Redbooks review” on page 159 to the fax number shown on the form.
- Use the online evaluation form found at **ibm.com**/redbooks
- Send your comments in an Internet note to redbook@us.ibm.com

Chapter 1. Introduction

This chapter provides background for the chapters that are later in the book. We describe the SAP R/3 and DB2 UDB for OS/390 definitions and concepts that apply to high availability. We rely on other redbooks in the SAP R/3 on DB2 UDB for OS/390 series for the architecture description of this environment. Books in that series are:

- *SAP R/3 on DB2 for OS/390: Implementing with AIX or Windows NT Applications Servers*, SG24-4945
- *SAP R/3 on DB2 UDB for OS/390: OS/390 Application Server*, SG24-5840
- *Database Administration Experiences: SAP R/3 on DB2 for OS/390*, SG24-2078
- *SAP R/3 on DB2 for OS/390: Disaster Recovery*, SG24-5343
- *High Availability Considerations: SAP R/3 on DB2 for OS/390*, SG24-2003

At the time of writing this redbook, the most recent redbook in the series is *SAP R/3 on DB2 UDB for OS/390: OS/390 Application Server*, SG24-5840. It contains the most current description of SAP R/3 architecture and the environment when DB2 UDB for OS/390 is used as the database server.

1.1 Availability background and terms

This section discusses considerations in achieving high availability. The ability to have a system with availability approaching 24x7 (24 hours per day, 7 days per week) depends on planning, hard work, money, and proper use of the resources the money buys. Ultimately, it is up to systems planners to determine the value of availability to a business enterprise; this determination will lead to investments in people resources and computer resources (hardware and software) that achieve the required availability levels.

1.1.1 Availability definitions

To insure that the terms we use are consistent with standard industry terminology, we provide the definitions and explanations in this section.

1.1.1.1 High availability (HA)

Availability is that characteristic of a system that states that an agreed-upon level of service will be provided during scheduled periods. The availability level is usually stated in a service level agreement between a representative of end users and a service provider; this is the *agreed level*. Since service is at the agreed level only during scheduled periods, there is an implication that it is possible to perform maintenance and apply modifications outside these hours.

One obvious characteristic of a system with high availability is reliable hardware and software, especially the system software on which applications are built. This category includes operating system software and a database management system (DBMS).

The implication of high availability is that points of failure have alternatives that allow the agreed level to be maintained. This *elimination of single points of failure* is the most important activity in planning for a highly available system.

1.1.1.2 Continuous operation (CO)

Continuous operation describes a system that is in operation “around the clock”, that is, 24 hours per day each day of the year; with no scheduled outages. A system in continuous operation is not necessarily operating with high availability - there could be a high number of unscheduled outages. A system with continuous operation (since it allows for no planned outages) will be one in which hardware maintenance, software maintenance, and administrative activities occur without the necessity of stopping applications.

1.1.1.3 Continuous Availability (CA)

This is the extension of high availability to 24 hours per day every day of the year. CA combines the characteristics of continuous operations and high availability to eliminate outages or mask the effect of those outages from the end user. Availability is measured at the level of the end user, so that a system is considered unavailable if any of the hardware, software, network resource, or user data required cannot be supplied.

Continuous availability systems approach “24x7” operation. For this to occur, there must be ways to perform administrative work, especially the work performed by a database administrator (DBA) without stopping applications. There must also be ways to perform hardware and software maintenance while the applications remain available.

1.1.2 Outage characteristics

Decreasing the number and duration of outages increases availability. Consequently, this is an important objective of system management. Outages are categorized to make it easier to measure and attack their causes or effects. One such categorization is planned and unplanned outages. Our definitions for these are:

- Planned outage

  An interruption in availability known early enough that decisions can be made on how to notify users, how to sequence the stopping of components, and how to organize recovery actions.

- Unplanned outage

  An interruption in availability that occurs without prior knowledge or that is expected to occur before one or more of the items in a planned outage can be performed.

1.1.2.1 Planned outages

When analysts and administrators know of an interruption to availability, it becomes possible to minimize the impact. It can be determined whether steps in the recovery plan (the “Unplanned Outage Plan”) are necessary; since normal shutdown or backups have occurred, it is possible that some steps will not have to be performed.

Here are some examples of planned outages:

- Hardware or software changes

  Upgrading hardware to new devices, adding resources such as additional processor memory or more disk devices, or applying changes to hardware

microcode are activities that have potential to be an outage; proper planning should allow operations to continue when this work is performed.

SAP R/3, DB2 UDB for OS/390, and OS/390 itself should all allow operations to continue while maintenance is performed. This is usually an item for the availability test plan (see Appendix A, “Availability test scenarios” on page 131). With OS/390 operating in a Parallel Sysplex, it is possible to stop an LPAR, apply OS/390 changes, restart the LPAR, stop another LPAR, apply changes to this second member, and restart the second LPAR; all without stopping the operation of SAP R/3.

- Data reorganization

  Database reorganization during SAP R/3 operation should be a normal activity, see “Data reorganization” on page 104. This is another item for an availability test plan.

- Administrative work

  Normal administrative functions should not require an outage. We address the considerations for this kind of activity in Chapter 4, “Database administration - reducing planned outages” on page 83.

- Batch functions

  Sometimes applications will preclude updating, thus producing the effect of a planned outage. For example, when a series of items becomes obsolete and no longer orderable (such as at the end of a year), the mass removal of those items from the database might preclude order entry.

- Time changes

  When system clocks are required to be reset, stopping and restarting processors may be necessary. Changing the time and verifying that the change has occurred correctly sometimes forces a planned outage.

1.1.2.2 Unplanned outages

The unplanned outage is caused by a failure in hardware, software, or procedures. The disruption of an unplanned outage can be minimized or totally eliminated by duplexing of hardware or software. Duplexing one or more components adds to the expense. Consequently, increased availability relative to the unplanned outage becomes a business decision (the benefit of uninterrupted service to the user versus the cost to provide that service).

Following are considerations for minimizing or avoiding altogether, the impact on the user of an unplanned outage. For every software, hardware, or network component, an outage plan should exist. The totality of these individual plans is referred to as the *Unplanned Outage Plan*. This plan should be the basis for operations scripts and test activity. It should be available to operations staff at all times; it should also be available to analysts for planned outage planning. The plan should be an audit item.

When new hardware is to be added to the configuration, when software levels are to be changed, and when audits are performed, the plan should be examined for possible changes. Also, at the conclusion of both test activities and true outage situations, a review of the plan should occur, looking for possible improvements to it.

Typical items addressed in this plan are:

- Hardware failure
- Software failure
- Network failure
- Performance bottleneck

1.1.3 Attacking planned outages

There are several strategies that allow an enterprise to either eliminate planned outages or to reduce the effect of a planned outage. Sometimes the reduction of the effect is so dramatic that users see no difference in operational characteristics when a planned outage occurs. Among the strategies that may be employed are:

- Utilization of alternative hardware components

  The alternate hardware (and the software that makes it possible to use alternates) that is so valuable for unplanned outages also is important for planned outages. Consider secondary disks and paths to those disks; the use of these facilities means that maintenance can be done to primary disks and paths without interfering with data base activity. Similarly, the presence of multiple processors in a Parallel Sysplex means that hardware maintenance or repair to one processor does not interrupt the work of the complex.

- Use of software that tolerates lower-level components

  In the case of a Parallel Sysplex, it is important that OS/390 can run with fixes applied to software on some of the processors but not on others. This means that the installation of fixes can occur to one member at a time; that is, the complex does not need to be stopped to perform operating system maintenance. It is also important for DB2 data sharing to have the same characteristic, you need not stop all the DB2 members; you can apply fixes to each member of the sharing group in turn, thus no outage occurs.

- Performing administrative tasks without application disruption

  Conditions will arise during the running of your SAP R/3 system that will require administrative tasks to be performed. For example, you may discover that performance will be improved if a table is reorganized. DB2 UDB for OS/390 allows you to do this without a planned outage. Several improvements have been made in this area and we point out recent developments that allow administrative functions to be done without a planned outage. This information is mainly found in Chapter 4, “Database administration - reducing planned outages” on page 83.

- Creating backups without service interruption

  The creation of backups, especially of database backups, has been consideration requiring attention in all installations of SAP R/3. The systems are highly active, so creation of backups usually must be done while applications execute. The backup must be done very quickly, otherwise tables backed up early in the process may be “out of sync” with tables backed up later. Also the backup function must allow access to both the system processors and to the database; it cannot “lock out” those resources for long periods.

Fortunately, the use of ESS or RVA hardware along with some recent changes in DB2 UDB for OS/390 provides all these characteristics. We discuss this in Chapter 5, “Creating database backups” on page 115.

1.1.4 Recovering from unplanned outages

The usual necessity for recovery is to get back to an operational state as quickly as possible. In the case of an unplanned outage, information to be used in determining the cause of the outage should be quickly gathered, then the actions in the Unplanned Outage Plan should be done. Note that the operational state reached may not be the normal state of SAP R/3 operations, and performance may not be at normal levels since some normal resources may not be available.

Actions that may be taken within the plan are:

- Failover to alternate

  This is automatic or manual use of an alternate component for one that has become unavailable. The alternate may have other functions or it may be totally inactive except for its replacement role. Also, the alternate may require some initialization or it may be ready for its takeover function. This latter case is referred to as *hot standby*.

- Component Restart

  Some components (an example is the SAP R/3 enqueue server) may not have multiple instances running. In this case, a restart of that function, perhaps on a different SAP R/3 application server, is necessary.

- Database Recovery

  This is an involved subject more thoroughly covered in Chapter 5, “Creating database backups” on page 115.

1.2 Scope of this redbook

This redbook is limited to a subset of availability planning, *database* availability. While database availability is by no means the only topic for an availability plan that applies to an SAP R/3 environment, it will make up the largest section of the plan; more than half the possible outages covered by the plan may be oriented to addressing the problem of how to access the database when critical items fail.

We limited the scope in this way because of our own limitations in time and resources. We also wanted to concentrate on database availability because we believe that DB2 UDB for OS/390 offers significant capabilities in this area; we wanted to test these capabilities to be certain that SAP R/3 installations can obtain these benefits and to determine if there are steps to be performed that have not been well documented elsewhere.

The limitations in time and resources forced us to make decisions to limit our approach to a subset of options that an SAP R/3 installation might examine. For example, in Chapter 2, “SAP R/3 availability benefits from S/390” on page 15, we limit our discussion to SAP R/3 availability; a full discussion of the benefits of S/390 could be (and in fact is) the subject of a redbook.

Another example is our implementation in Chapter 3, “Database high availability for SAP R/3” on page 61 of only one of the data sharing options described in this

chapter; we felt the other options would have most of the same considerations and our time would be best spent in going into option 2 in depth.

You will also recognize how we were limited when you read Chapter 4, “Database administration - reducing planned outages” on page 83; where the most important administrative tasks absorbed our attention. In Chapter 5, “Creating database backups” on page 115, we were limited to testing on the hardware available, although we do include some discussion of other hardware.

1.3 SAP R/3 architecture and database availability

This section describes how database availability is affected by the way SAP R/3 works. Since our scope is limited to database availability, in this section we concentrate on two areas:

1. The functions within SAP R/3 that assure the integrity of the database. This is done because the functionality impacts database availability.
2. The particular advantages in database availability an installation can realize by using features of DB2 UDB for OS/390.

1.3.1 Database Integrity

Two principles that are basic to database integrity are *locking* and the concept of a *unit of work*.

Locking prevents concurrent users from accessing inconsistent data. For example, by getting a lock for a database record before modifying it, you prevent any other user transaction from modifying the same record at the same time.

Locking in SAP R/3 is done by the enqueue server that runs as a process in the central instance SAP R/3 application server. It is the application's responsibility to request an enqueue (ENQ) for a data object before accessing it.

Changes to a database made by an application are not visible to other applications until the changes are committed. The application can commit its changes explicitly, or the commit can occur implicitly when the application terminates.

If the application terminates abnormally, then all changes to the database made since the last commit point are “rolled back”. The processing between the start of the application and the commit point is called a *logical unit of work* (LUW).

Database integrity is maintained by ensuring that all changes to a database made during a unit of work are either committed or rolled back; there can be no partial updates.

1.3.1.1 Transaction concept in SAP R/3

Each dialog step of a business transaction in SAP R/3 can (in theory) be processed within a different work process in an application server and would therefore use a different ICLI thread to communicate with DB2 UDB for OS/390.

The database changes made in a dialog step are all committed at the end of the dialog step. This is acceptable to some applications, but most business transactions consist of multiple dialog steps. In other words, the SAP R/3 LUW is

expected to consist of the complete business transaction, while DB2 treats each dialog step as a discrete unit of work.

The SAP R/3 system has special update services to manage this difference in scope; see Figure 1.

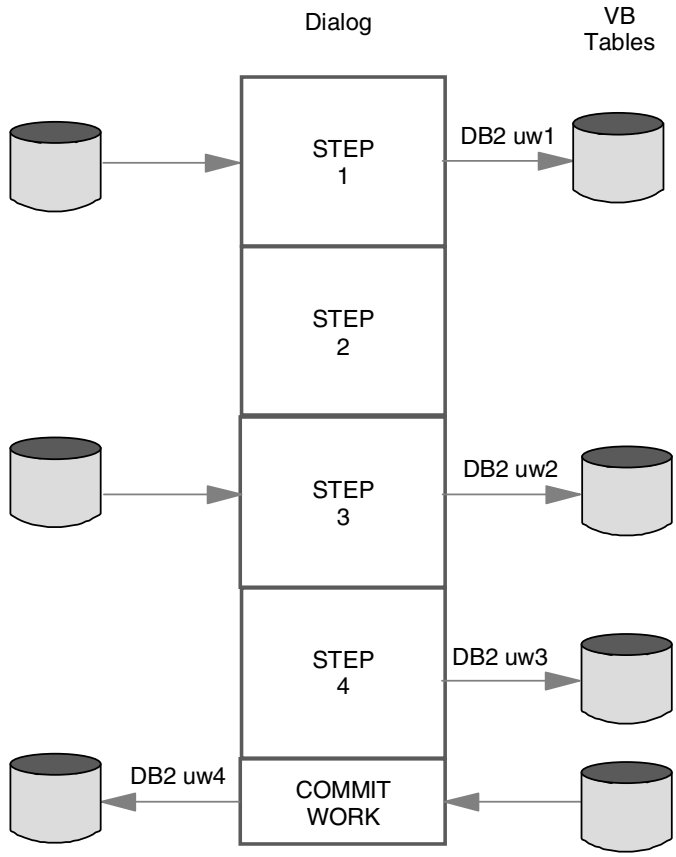


Figure 1. SAP R/3 transaction

The following notes refer to Figure 1:

- Steps 1, 3, and 4 intend to update the application database; the actual writing is to VBLOG (or the three tables that replace VBLOG: VBHDR, VBDATA, and VBMOD).
- Step 2 is user interaction; no database processing occurs.
- Step 4 performs an insert with no (additional) database input.
- The SAP R/3 enqueue server has the responsibility to lock out other users' access to affected application data from the time update intent is recognized until the SAP R/3 COMMIT WORK is complete.
- Note the difference in the DB2 units of work (DB2 uw1-uw4) and the SAP R/3 LUW (the entire business transaction).
- The work performed in steps 1 through 4 is done by an SAP R/3 dialog process (it could also be done by a batch process if batch input rather than online is used), while the COMMIT WORK function and the associated database updates occur in an SAP R/3 update process.

- This is a common technique used in SAP R/3 programming; however, it is *not*: the only way dialogs function.

If an outage occurs, SAP R/3 relies on DB2 UDB for OS/390 to roll back the database to a consistent point (from the DB2 LUW point of view). For example, in Figure 1 this could be at the completion of db2uw2. Note that the consistent database includes the VB tables. When SAP R/3 begins to operate, the VB tables reflect an SAP R/3 incomplete LUW, and units of work (db2uw1 and db2uw2) are rolled back by SAP R/3.

1.3.2 DB2 UDB for OS/390 availability features

DB2 UDB for OS/390 has many features that increase the availability of database access to an SAP R/3 installation. Some of these features are significant development items that were engineered into the product. Others are small enhancements that make DB2 UDB for OS/390 particularly useful to users of SAP R/3.

A feature of DB2 UDB for OS/390 that is most important for availability is data sharing. This feature allows multiple instances of DB2 to provide access to databases, and (through the facilities of OS/390 Parallel Sysplex) to recognize if one of the DB2 instances is no longer available. As a result of this recognition, action can be taken so that the database service is provided by other instances, and user interaction with the database is not interrupted. We describe possible scenarios for using data sharing in “Database failover planning” on page 8.

DB2 UDB for OS/390 contains a series of utilities to perform administrative tasks against a database. Of particular interest to an enterprise that uses SAP R/3 is the unobtrusive nature of these utilities, that is, the ability to use the utilities to maintain the database or modify the organization of it without stopping SAP R/3. We examine some common functions and show how impact of administrative activities is minimized in Chapter 4, “Database administration - reducing planned outages” on page 83.

1.4 Database failover planning

In this section, we define four options for the use of DB2 UDB for OS/390 data sharing. These options correspond to implementations found in SAP R/3 installations or to configurations that were considered before an enterprise implemented one that is more complex. Although our major interest in data sharing is in how it provides database availability, we also discuss other aspects of the use of data sharing; how it can provide better performance and how it can provide some relief from a virtual storage constraint.

1.4.1 Data sharing configuration options - in general

DB2 for OS/390 has a function called *data sharing*; users take advantage of this function for many advantages: performance, capacity, flexibility, and availability. We are particularly interested in the use of data sharing to provide increased database availability.

In the case of application servers outside OS/390, the Integrated Call Level Interface (ICLI) component of SAP R/3, working with facilities of TCP/IP may recognize that a DB2 request is not being handled in expected time, and thus

cause the use of a different ICLI, that routes requests to a different data sharing member. It may also happen that ICLI client-to-server communication has become impossible, so an alternate client and server combination is used, communicating to a different data sharing member to perform requests.

If the application server is on OS/390, the program within the application server communicates to the data sharing group, with DB2 UDB for OS/390 then performing the function of routing the request to a surviving member of the group if the normal group member has become unavailable.

1.4.2 Single DB2 with inactive standby - option 0

Option 0 is basic data sharing for the objective of availability (see Figure 2 on page 9). That is, a standby data sharing member has no function assigned other than to wait for failover to occur from a primary member. Note that the architected data sharing function provides the ability to attack several possible single points of failure or planned outage situations: hardware failure or necessary maintenance in the MVS1 system, OS/390 failure or maintenance in the MVS1 system, an error in DB2 UDB for OS/390 in the DSN1 subsystem, maintenance necessity in DB2 UDB for OS/390, ICLI failure in the primary ICLI server, or hardware connectivity failures in the ICLI client-to-server path from the application server to OS/390 system MVS1.

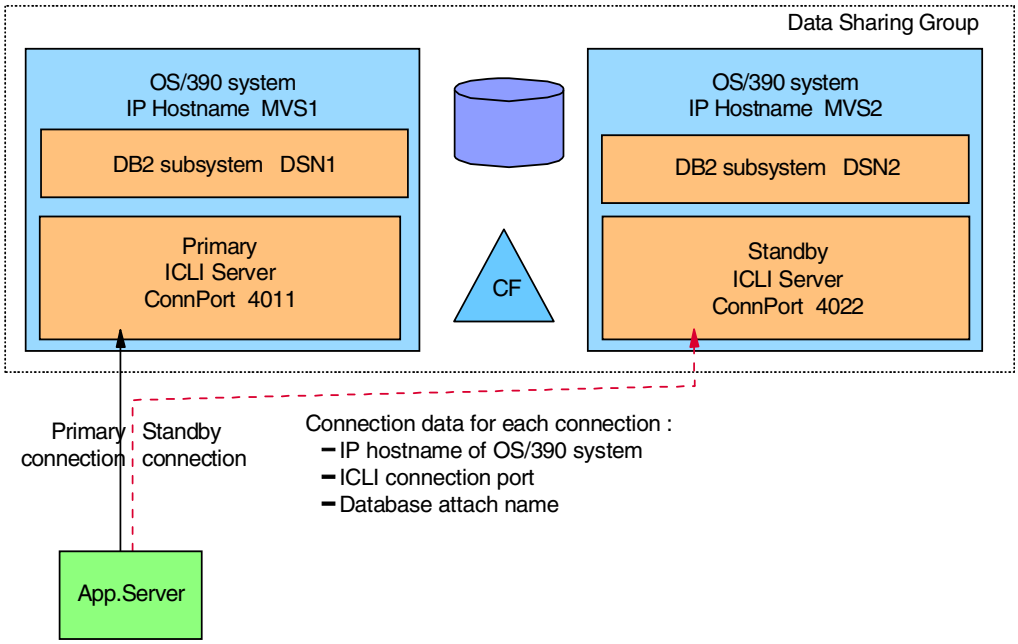


Figure 2. Data sharing option 0

Note that the S/390 hardware and OS/390 found in MVS2 may have other work assigned, but the load there should be light enough that the occurrence of failover will not cause SAP R/3 to suffer degraded performance.

1.4.3 Two active DB2s, virtual storage relief or performance - option 1

Option 1 is basic data sharing with a different objective: the DB2 member DSN2 performs work on behalf of SAP R/3 in normal operation; it provides access to the additional processing power found in MVS2, and it provides an additional address

space of 2 Gigabytes for storage of DB2 buffer pools, locks, and prepared SQL. Refer to Figure 3 on page 10.

This should *not* be viewed as an availability possibility because the occurrence of failover would cause degraded performance in the best case (since the resources for operation would be significantly lessened). In the worst case, failover would cause system failure, since additional resource (virtual storage) requirements in the DBM address space could exceed 2 Gigabytes.

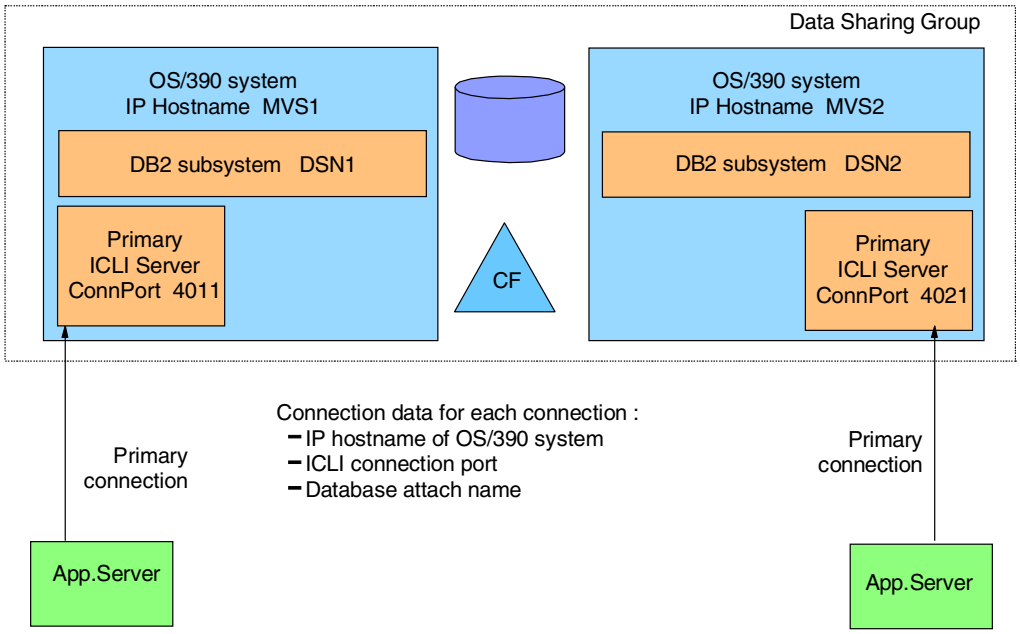


Figure 3. Data sharing option 1

We include option 1 only because it is a starting point for many large SAP R/3 installations that are ready to begin availability planning; option 1 is not an availability solution.

1.4.4 Two active DB2s, inactive standby in opposite LPAR - option 2

Data sharing option 2 is a realistic implementation for a large SAP R/3 installation to allow database availability. With option 2, standby data sharing members are provided for each of two primary members, see Figure 4 on page 11.

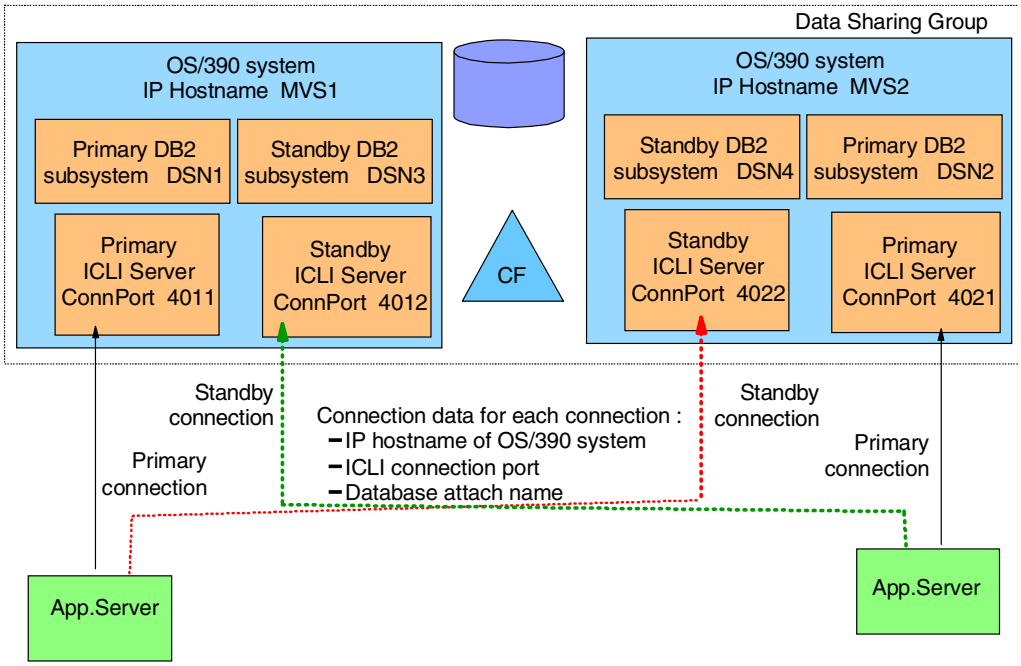


Figure 4. Data sharing option 2

Note that the necessity for multiple data sharing members usually exists in SAP R/3 installations as is discussed in “Two active DB2s, virtual storage relief or performance - option 1” on page 9. You could view option 2 as a combination of options 0 and 1, with the multiple primary members providing performance and virtual storage resources, and the standby members providing availability. As in the other options, it is possible that the MVS1 and MVS2 machines (or LPARs) are performing other work, as long as the load is light enough to allow adequate performance if activation of one of the standby members occurs.

We used option 2 in our testing. Figure 4 is a generalized view of an implementation with option 2. To see the exact port numbers, addresses, system names, and DB2 subsystems in our configuration, refer to Figure 15 on page 62 with the explanation found in Chapter 3, “Database high availability for SAP R/3” on page 61.

We chose this option for several reasons. One of our objectives was to encounter the same difficulties an SAP R/3 installation would meet in an HA implementation. Another objective was to have an implementation that reflected the usage of most customers. We also wanted a configuration that would illustrate the new advances in HA for SAP R/3 users who choose DB2 UDB on OS/390 as their database management system. Finally, we wanted to use as simple a configuration as possible so that implementation of the option could be completed in the time we had.

1.4.5 Two active DB2s, inactive standby in independent LPAR - option 3

Option 3 represents the option 2 solution carried to the next level. Not only are there standby data sharing members, there are also standby machines or LPARs in which the members reside. Refer to Figure 5 on page 12.

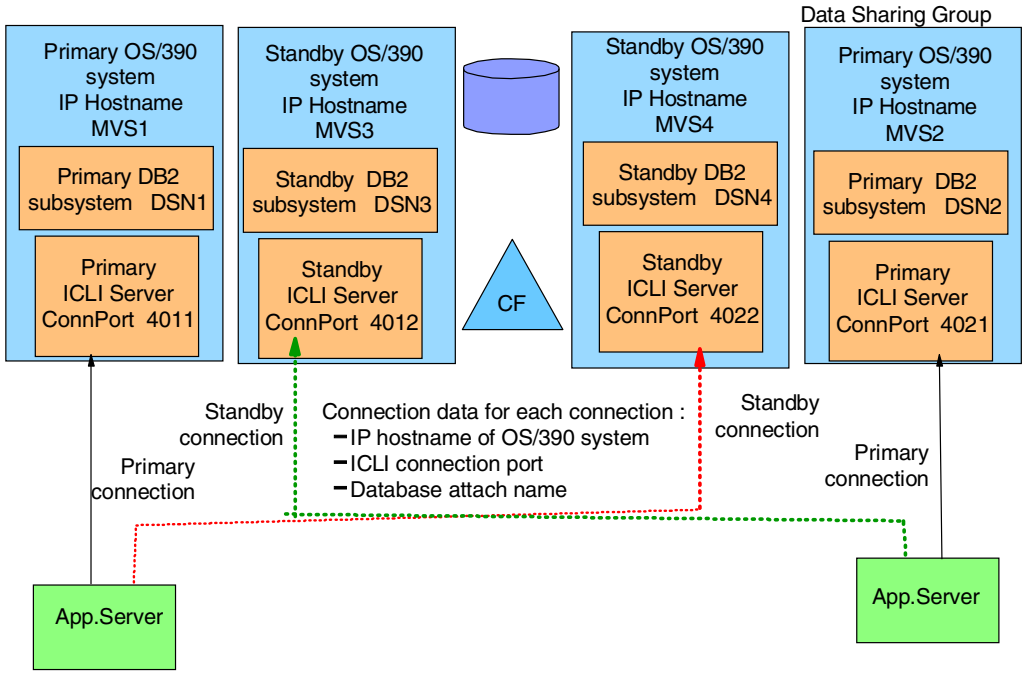


Figure 5. Data sharing option 3

At first glance it may appear that additional LPARs for the standby function is a heavy price to pay for availability, but most installations that incorporate option 3 assign the standby function to one of the LPARs in their “test-to-production landscape”. That is, the normal function of MVS3 or MVS4 could be as a test or quality assurance machine, functions that can be interrupted if an availability situation arises.

The use of machines in the test-to-production landscape is not without risk; they could be using a level of OS/390 or DB2 UDB for OS/390 that has not yet been put into production. Failover scenarios could then encounter situations with this new code that have software failures, operational differences, or performance impacts that are unexpected.

1.4.6 Database failover impact on enterprise objectives

As we discussed in explaining data sharing configuration option 1 (see “Two active DB2s, virtual storage relief or performance - option 1” on page 9), an SAP R/3 installation will usually have requirements for data sharing other than database availability. In considering failover scenarios, you must keep the objectives of your installation in mind. “Can some degradation in performance be tolerated in this situation?” is a question that will recur continuously. Also remember the failover scenarios must be tested; we’ve mentioned several times how a requirement for virtual storage beyond the 2 Gigabyte limit could occur if a standby DB2 had production resource requirements in addition to its standby role.

Note also that your plans for minimizing multi-system interest could be impacted in failover situations. For example, look again at Figure 5 on page 12. Note that there could be an application on MVS2 that interacts with the SAP R/3 database or (with application server on OS/390) with an SAP R/3 application. In a failover scenario, the database processing occurs on MVS4. This means that the two

machines must interact in order for the applications to be able to communicate. This interaction creates additional overhead, so performance could be degraded.

1.4.7 Choosing a data sharing configuration option

Table 1 summarizes the data sharing configuration pros (+) and cons (-).

Table 1. Pros and cons of data sharing configuration options

	Option 0	Option 1	Option 2	Option 3
Advantages				
Maintain tight control of dispatching priorities	+	+	+	
Failover dialogs steal cycles from batch ICLIs		+	+	
No impact to surviving DBM1 virtual storage or buffer pool hit rates.			+	+
Non-disruptive resumption of normal configuration			+	+
No impact to available virtual storage or LPAR cycles	+			+
Greater than 2 Gigabytes of central storage may be used				+
Disadvantages				
Reserve thread storage required for reconnect		-		
Increased demand on survivor buffer pools and Dynamic Statement Cache (DSC)		-		
Resumption requires thread/storage cleanup	-	-		
Increased demand on central storage, expanded storage and processor cycles		-	-	
Higher memory and LPAR management costs				-
Diffuses tight control of dialog-versus-batch dispatching priority				-

1.5 Recent developments in database availability

This section refers to some recent items for DB2 UDB for OS/390 and for SAP R/3 that address database availability. While the items are in fix documentation, they are not fixes for program “bugs”, but a method of introducing software changes that address availability issues.

1.5.1 Log write suspend

This function, implemented as a fix in DB2 UDB for OS/390 V6.1 (APAR PQ31492, PTF UQ36695), allows a command, SET LOG SUSPEND. The command forces out log buffers, updates the high written RBA in the BSDS, and sets a latch that prevents writes to the log and to application data. This creates a point of consistency which allows a completely consistent backup copy.

1.5.2 Split mirror backup

This is a facility that uses the log write suspend function along with the mirroring capability of IBM ESS to allow a backup copy to be made using mirrors of the logs, the BSDS, and the ICF catalog. The process is completely described in a white paper available at: *www.storage.ibm.com/hardsoft/diskdrls/technology.htm*

1.5.3 Storage cushion for DB2 master address space

This function implements a ZPARM called CONTSTOR through PTF PQ14391 to DB2 UDB for OS/390 V5. The ZPARM causes the thread control blocks to be periodically examined and inactive ones discarded. This aids availability by reducing the possibility of DB2M1 address space abends due to exceeding the 2 GB address space limit. However, there is potential increased CPU impact, since there may be more frequent calls to OS/390 FREEMAIN.

1.5.4 Parameter DEFINE NO in CREATETABLESPACE

This parameter allows a tablespace to be created, but does *not* cause the initial allocation of space until the tablespace is opened. this is described in “New function: Deferring the definition of data sets” on page 101.

Chapter 2. SAP R/3 availability benefits from S/390

This chapter addresses the usage of SAP R/3 on a single S/390 system and in a System/390 Parallel Sysplex, and describes how you can achieve high availability objectives in such an environment. We discuss physical hardware configuration considerations and also show how processor failure can occur without losing availability of SAP R/3. Finally we show how certain Coupling Facility failures will have no impact on SAP R/3 database server availability.

2.1 Availability features of single system configurations

The SAP R/3 database is in every SAP R/3 system the component with the highest demands on availability. Actually, the most frequently articulated requirement of SAP R/3 users is the need for continuous data availability with integrity. Therefore, providing a hardware and software infrastructure with the highest possible availability for the SAP R/3 database is of essential importance to every enterprise.

The goal of such configuration considerations is to eliminate any possible single point of failure (SPOF), which can be based in the hardware and software. The elimination of SPOFs can mostly be achieved through redundancy, both on the hardware and the software side. Of course, the easiest way to eliminate failures is by using components that are less likely to fail. For hardware, this can be easily achieved by using reliable components. Furthermore when a failure occurs, the system should record sufficient information about it so that the problem can be fixed before it recurs. For software, it should not only be written to avoid failures, but also to identify and recover those that occur. Automation also eliminates failures, by ensuring that procedures are followed accurately and quickly every time.

The availability features of the S/390 platform are derived from these concepts. S/390 was designed with a reliability, availability and serviceability (RAS) philosophy. Its availability features result from more than 30 years of evolution and are incorporated in the S/390 hardware, the OS/390 operating system and in DB2 UDB for OS/390. This chapter will elaborate on those features which are incorporated in a standalone S/390 system. It will be shown how SAP R/3 on DB2 UDB for OS/390 can benefit from those features. Certainly a Parallel Sysplex configuration, which typically consists of several S/390 systems, will benefit from those availability features of a single S/390 system as well. The next section, 2.2, “Additional availability features and benefits with S/390 Parallel Sysplex” on page 32 will show which features are available on top of those of a single system.

2.1.1 Features of the S/390 hardware architecture

Several RAS features are implemented in the S/390 hardware to provide a high availability of the application. Every S/390 generation over the years contained more and further enhanced features as the platform matured, and there will surely be more of those features coming in future generations. Many of S/390 RAS features were developed at a time when the failure of hardware elements was more frequent, and they are very rare nowadays. However, these hardware availability features are just as valuable today. The design concept is to avoid SPOFs. Therefore most S/390 hardware elements have built-in redundancy or can be circumvented if they fail.

This chapter is going to explain in detail which availability features are imbedded in the S/390 hardware architecture and how SAP R/3 benefits. Each of the availability features is explained in more detail in the IBM publication *System Overview, S/390 9672 Generation 6,* GA22-1030. After having explained those hardware availability features, we will do a practical test at the end of this chapter to verify that they work as designed also for SAP R/3. In this test, we will simulate a the loss of a CPU and document the results on SAP R/3.

2.1.1.1 List of S/390 availability features

Table 2 on page 16 summarizes the features which are implemented in the S/390 architectural design. It shows which availability features apply to the frequency, duration, and scope of an outage. It further explains whether this feature helps eliminate planned or unplanned outages, or both.

Table 2. S/390 availability features matrix

Availability feature	Reduces outage frequency	Reduces outage duration	Reduces outage scope	Type of addressed outage
Enhanced Processor Design	X			unplanned
Fault Tolerant Design	X			unplanned
Alternate Support Element		X	X	unplanned
Transparent CPU Sparing	X			unplanned
Dynamic SAP Sparing / Reassignment	X			unplanned
Enhanced Application Preservation	X			unplanned
Cryptographic Coprocessor Enhancements	X	X		unplanned
Error Correction Code	X			unplanned
Scrubbing	X			unplanned
Dynamic Memory Sparing	X			unplanned
LPAR Dynamic Storage Reconfiguration	X			planned
Partial Memory Restart	X	X		unplanned
Dynamic I/O Configuration	X			planned
Concurrent Channel Upgrade	X			planned
Partial I/O Restart	X		X	planned
Dual Power Feeds	X			unplanned
Redundant Power Supply Technology	X			unplanned
Storage Recovery	X			unplanned
External Time Reference	X			unplanned
Concurrent Hardware Maintenance	X			planned
Capacity Upgrade on Demand	X			planned
Concurrent Licensed Internal Code Patch	X			planned

Availability feature	Reduces outage frequency	Reduces outage duration	Reduces outage scope	Type of addressed outage
Internal Battery Feature	X			unplanned
Local Uninterruptible Power Supply	X			unplanned

Legend: X=Applies

In addition to those availability features, S/390 also incorporates features that address the reliability and serviceability of the platform for a complete RAS approach. All of the standard features of S/390 in terms of RAS are in detail:

1. **Reliability** features:
   - High-reliability technology components
   - Parts integration to reduce the number of parts in the machine
   - MCM System Run-In Test in manufacturing to simulate 9000 hours of extreme customer operation
2. **Availability** features:
   - **Enhanced Processor Design**
     S/390 processors are provided with an enhanced processor design. Each CPU contains dual Instruction / Execution Units, which operate simultaneously. Results are compared, and in the event of an unequal compare, Instruction Retry is invoked. If the Instruction Retry doesn't succeed, the CPU is considered failed and put offline. This design simplifies checking, and virtually eliminates CPU failures due to soft errors.
   - **Fault Tolerant Design**
     Fault tolerant design allows hardware recovery to be performed in most cases totally transparent to customer operation and eliminates the need for a repair action, or defers a repair action to a time which is more convenient for you.
   - **Alternate Support Element**
     From S/390 G6 on a second Support Element is installed in the server frame, as a backup to the primary Support Element. In the event of a hardware malfunction, the mirrored alternate will take over for the failing Support Element when a switch located in the server's back panel is manually set.
   - **Processor Unit (PU) Sparing**
     All S/390 G5 and G6 servers provide at least one spare PU except the maximum configured models. Using the 10-way ZX7 model as an example, it is equipped with fourteen PU chips. Ten of the PUs are S/390 Central Processor Units (CPUs), two PUs are system assist processors (SAPs), and the other two PUs are spare. Availability scenarios for CPU and SAP are:
     - **Transparent CPU Sparing**
       In the event of a CPU failure, and when a spare PU is available, in most cases the system will initialize a spare PU to take over as a CPU. This process also covers a failure of an Internal Coupling Facility. CPU Sparing is transparent in all modes of operation and requires no

operator intervention to invoke a spare PU. In most cases, the application that was running on the failed CPU will be preserved and will continue processing on a new CPU with no customer intervention required. See Enhanced Application Preservation below.

- **Dynamic SAP Sparing / Reassignment**
  Dynamic recovery is provided for failure of the System Assist Processor (SAP). In the event of a SAP failure, if a spare Processor Unit (PU) is available, in most cases the spare PU will be dynamically activated as a new SAP. If there is no spare PU available, and the CPC has more than one CPU, an active CPU will be reassigned as a SAP. In either case, there is no customer intervention required. This capability eliminates an unplanned outage and permits a service action, if necessary, to be deferred to a more convenient time.

– **Enhanced Application Preservation**
Application Preservation was introduced on the S/390 G4 servers and is enhanced on the S/390 G5 servers to provide more comprehensive application recovery should a CPU fail. Application Preservation captures the machine state in the event of a CPU failure and will, in most cases, switch processing to another active CPU without customer intervention. On the S/390 G4 servers, the uniprocessors were not covered by Application Preservation; G5 and G6 uniprocessor models will now recover work due to a CPU failure, in most cases, on a spare PU using Application Preservation. Also the maximum configured G5 and G6 models, although without CPU Sparing, will attempt to recover an application that was running on a failed CPU on another active CPU in the configuration. This capability helps eliminate unplanned outages, eliminates customer intervention in the recovery process, and preserves the customer's application processing environment.

– **Cryptographic Coprocessor Enhancements**
Recovery of a cryptographic coprocessor element is done by the operating system; i.e., the operating system reschedules and dispatches the failed instruction on the other cryptographic coprocessor element.
Another availability feature on S/390 servers is a second path from each cryptographic coprocessor element to a spare PU. Normally, each crypto element is configured to a primary CPU. Should a primary CPU fail, the alternate PU would transparently replace (spare) the failed primary CPU, maintaining the crypto element's operation. However, an alternate PU is available only if that PU is not configured into the system as another CPU, ICF or SAP. The two PUs associated with the alternate path from each crypto element are the last to be assigned as CPUs, SAPS or ICFs.

– **Error Correction Code**
Memory error checking and correction code detects and corrects single bit errors. Also, because of the memory structure design, errors due to a single memory chip failure are corrected. This feature ensures that data corruption cannot go undetected and provides superior reliability of S/390 main memory. ECC also includes checking of the system memory buses.

– **Scrubbing**
During normal operations, the system monitors and records accumulation of failing bits in memory chips that are corrected by ECC. Storage background scrubbing provides continuous monitoring of storage for the correction of detected faults before the storage is used.

– **Dynamic Memory Sparing**
In addition to the use of Error Correction Code (ECC) and storage scrubbing, S/390 memory cards are equipped with spare memory chips. Before the amount of failing bits in memory chips reaches a threshold which could result in an uncorrectable error, the system invokes a spare memory chip in place of the one with the accumulated failing bits. This action may prevent an unscheduled outage for replacement of the memory card. It further reduces the possibility of a memory failure that would cause loss of data and a system outage.

– **Storage Recovery**
In the event of cache and directory errors, the recovery design of S/390 processors provides line deletes and relocates to remove the affected areas from the configuration. Processing will continue and the service action, if required, will be scheduled for a convenient time. Storage recovery complements ECC, background scrubbing and dynamic memory sparing.

– **LPAR Dynamic Storage Reconfiguration**
PR/SM LPAR storage reconfigurations of both central and expanded storage can occur allowing nondisruptive add or removal to any partition with a cooperating guest. Furthermore, the restriction of storage reconfigurations only being possible from an adjacent and above logical partition is removed.

– **Partial Memory Restart**
In the event of a memory card failure, the system can be restarted with half of the original memory. Processing can be resumed until a replacement memory card is installed. This feature minimizes the length of an outage caused by a memory card failure.

– **Dynamic I/O Configuration**
Dynamic I/O configuration enhances system availability by supporting the dynamic addition, removal, or modification of channel paths, control units, I/O devices, and I/O configuration definitions to both hardware and software without requiring a planned outage. This capability, which is also referred to as “hot plugging”, allows configuration changes like the replacement of a channel card non-disruptively during the system operation. Simply the CHPIDs on the affected card would have to be configured offline.

– **Concurrent Channel Upgrade**
It is possible to concurrently add ESCON, Parallel, Coupling Facility (CF), and OSA channels provided a Self-Timed Interface is installed and there are unused channel positions and a channel driver card or CF channel adapter card that controls the unused channel positions. In addition, Integrated Cluster Bus, FICON channel cards, and OSA-Express channel cards and their associated STI cables can be added provided there are unused channel and cable jack positions. This feature helps eliminate an outage to upgrade the channel configuration.

– **Partial I/O Restart**
In the event of a failure of a Memory Bus Adapter, the system can be restarted to run with only the I/O connections associated with the failed adapter deconfigured. In a system configured for maximum availability, alternate paths will maintain access to critical I/O. This capability enables

the system to run, partially degraded, until the part is replaced, restoring full capacity.

- **Dual Power Feeds**
  The power system offers dual primary (AC) power feeds. Each feed is electrically isolated and enables redundant power paths to each server. Customers may elect to provide a dual electrical service to the server, further minimizing any outage due to a single path power interruption.
- **Redundant Power Supply Technology**
  The AC and DC power subsystems are designed with N+1 redundancy. Failure of a power thermal component does not cause a system outage. Concurrent replacement of the failed component results in an avoidance of a planned outage.
- **External Time Reference**
  ETR connection provided across PU sets to enable the set degrade capability. The design includes both the full- and half-populated module cases.
- **Concurrent Hardware Maintenance**
  Concurrent maintenance enables the replacement of failed units concurrently with system operation. This enhances the processor availability by eliminating the need for system outage to effect the repair. Concurrent maintenance capability exists for example for ESCON and FICON channels, coupling links, OSA2 and OSA-Express channels, power supplies and further elements.
- **Capacity Upgrade on Demand**
  Capacity Upgrade on Demand enables you to add non-disruptively processing or coupling capacity, an engine at a time, to most of the recent Parallel Enterprise Server models. The additional capacity can be invoked non-disruptively, eliminating the need for a scheduled outage.
- **Concurrent Licensed Internal Code (LIC) Patch**
  Concurrent Code Patch allows the activation of a patch concurrent with system operation, thereby increasing the availability of the processor by reducing scheduled outage for LIC maintenance. This capability exists for code for elements like CPU, SAP, LPAR, ESCON and FICON channels, coupling links, OSA2 and so on.

Furthermore there are some optional features that provide a high level of availability. They include the Internal Feature, which provides backup input power, and the Local Uninterruptible Power Supply, which can be installed as a supplement or an alternative to a central UPS.

3. **Serviceability** features:
   - Automatic error detection and fault isolation concurrent with system operation
   - Automatic remote support capability
   - High degree of concurrent maintenance capability in hardware and code
   - Multiple Channel Swap - an enhancement for channel problem determination allowing up to four channels to be swapped concurrently with system operation

– Status Panel showing status of N+1 power system

– Enhanced diagnostics for Coupling Links

2.1.1.2 Scenario: CPU Engine Failure

All of the preceding features provide a higher reliability of S/390 hardware and therefore a better availability of SAP R/3. As shown in Appendix A, “Availability test scenarios” on page 131, certain tests have been performed by IBM to show that failures of those features will have no impact on SAP R/3. Concurrent maintenance of a channel card, the loss of the redundant power supply or cooling unit or a failure of the support element all didn’t impact SAP R/3 which went on running.

In this chapter, we perform a sample test of one of the availability features of the S/390 hardware. We will simulate the loss of a CPU and see what kind of effect it has on SAP R/3. However, keep in mind that this is just meant to be an example instead of a typical scenario which causes an outage. S/390 CPUs tend to be much more reliable than on any other platform. The mean time between critical failures (MTBF) of a single S/390 CPU is 40 years. For a fully-equipped S/390 system with 12 CPUs, however, the average time until the first of these CPUs fails is less than 40 years. So we will test the loss of a CPU even though it is very unlikely to happen.

Purpose:

This test demonstrates S/390’s ability to run with less capacity in the event of a processor failure, while SAP R/3 is up and running. It will verify the features Transparent CPU Sparing and Enhanced Application Preservation.

Method:

To simulate the loss of a CPU, one engine will be configured offline. In our installation, we have 2 LPARs SC54 and SC42. Both of them have more than one CPU, which is a requirement for this test. SC54 has 2 CPUs, SC42 has 3 CPUs. We are going to use SC42 for this test.

Tasks:

1. Produce workload on the system. We do that by running an especially developed benchmark tool.
2. Configure 1 CPU offline on SC42.

```
15:22:14.30 HAIMO    00000290  CF CPU(2),OFFLINE
15:22:15.62 HAIMO    00000090  IEE505I CPU(2),OFFLINE
15:22:15.62 HAIMO    00000090  IEE712I CONFIG   PROCESSING COMPLETE
```

3. Display that both LPARs are running, SC42 with 2 CPUs and SC54 with 2 CPUs.

```
D M=CPU
IEE174I 15.22.24 DISPLAY M 213
PROCESSOR STATUS
ID  CPU                 SERIAL
0    +                  0593059672
1    +                  1593059672
2    -

CPC ND = 009672.R76.IBM.02.000000049305
CPC SI = 9672.R76.IBM.02.0000000000049305
CPC ID = 00

+ ONLINE    - OFFLINE    . DOES NOT EXIST

CPC ND  CENTRAL PROCESSING COMPLEX NODE DESCRIPTOR
CPC SI  SYSTEM INFORMATION FROM STSI INSTRUCTION
CPC ID  CENTRAL PROCESSING COMPLEX IDENTIFIER


/RO SC54,D M=CPU

RESPONSE=SC54
 IEE174I 15.22.30 DISPLAY M 247
 PROCESSOR STATUS
 ID  CPU                 SERIAL
 0    +                  0108229672
 1    +                  1108229672

 CPC ND = 009672.X77.IBM.02.000000050822
 CPC SI = 9672.X77.IBM.02.0000000000050822
 CPC ID = 00

 + ONLINE    - OFFLINE    . DOES NOT EXIST

 CPC ND  CENTRAL PROCESSING COMPLEX NODE DESCRIPTOR
 CPC SI  SYSTEM INFORMATION FROM STSI INSTRUCTION
 CPC ID  CENTRAL PROCESSING COMPLEX IDENTIFIER
```

4. Verify that the SAP R/3 workload is still running. We will take a screenshot to display that.

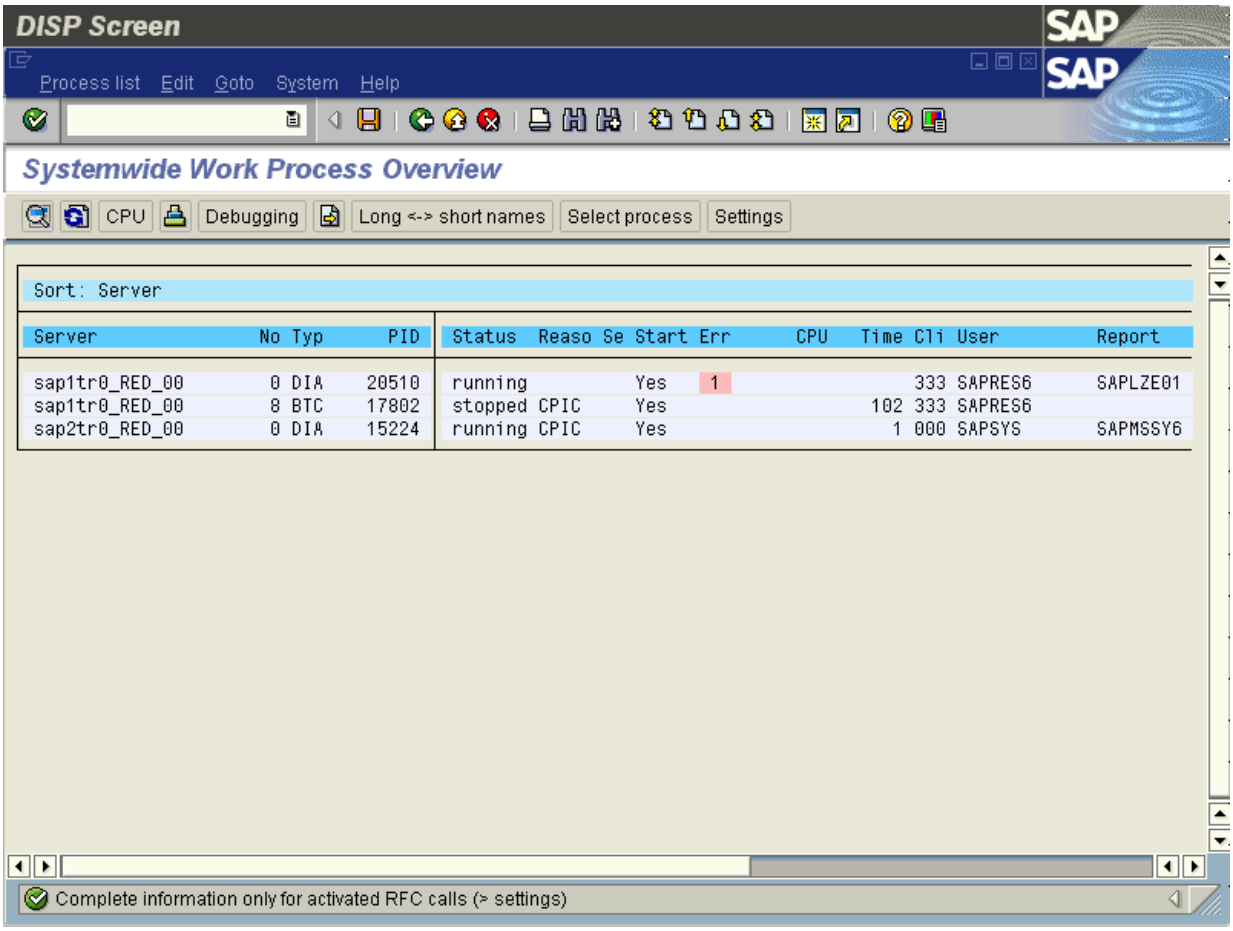


Figure 6. SM66: Online SAP R/3 workload during CPU outage

5. Configure the offline CPU back online.

```
15:22:34.01 HAIMO    00000290  CF CPU(2),ONLINE
15:22:34.45 HAIMO    00000090  IEE504I CPU(2),ONLINE
15:22:34.45 HAIMO    00000090  IEE712I CONFIG   PROCESSING COMPLETE
```

6. Display that all three CPUs are online again.

```
D M=CPU
IEE174I 15.22.38 DISPLAY M 218
PROCESSOR STATUS
ID  CPU                 SERIAL
0    +                  0593059672
1    +                  1593059672
2    +                  2593059672

CPC ND = 009672.R76.IBM.02.000000049305
CPC SI = 9672.R76.IBM.02.0000000000049305
CPC ID = 00

+ ONLINE    - OFFLINE    . DOES NOT EXIST

CPC ND  CENTRAL PROCESSING COMPLEX NODE DESCRIPTOR
CPC SI  SYSTEM INFORMATION FROM STSI INSTRUCTION
CPC ID  CENTRAL PROCESSING COMPLEX IDENTIFIER
```

7. Verify again that the SAP R/3 workload is still running and make another screenshot.

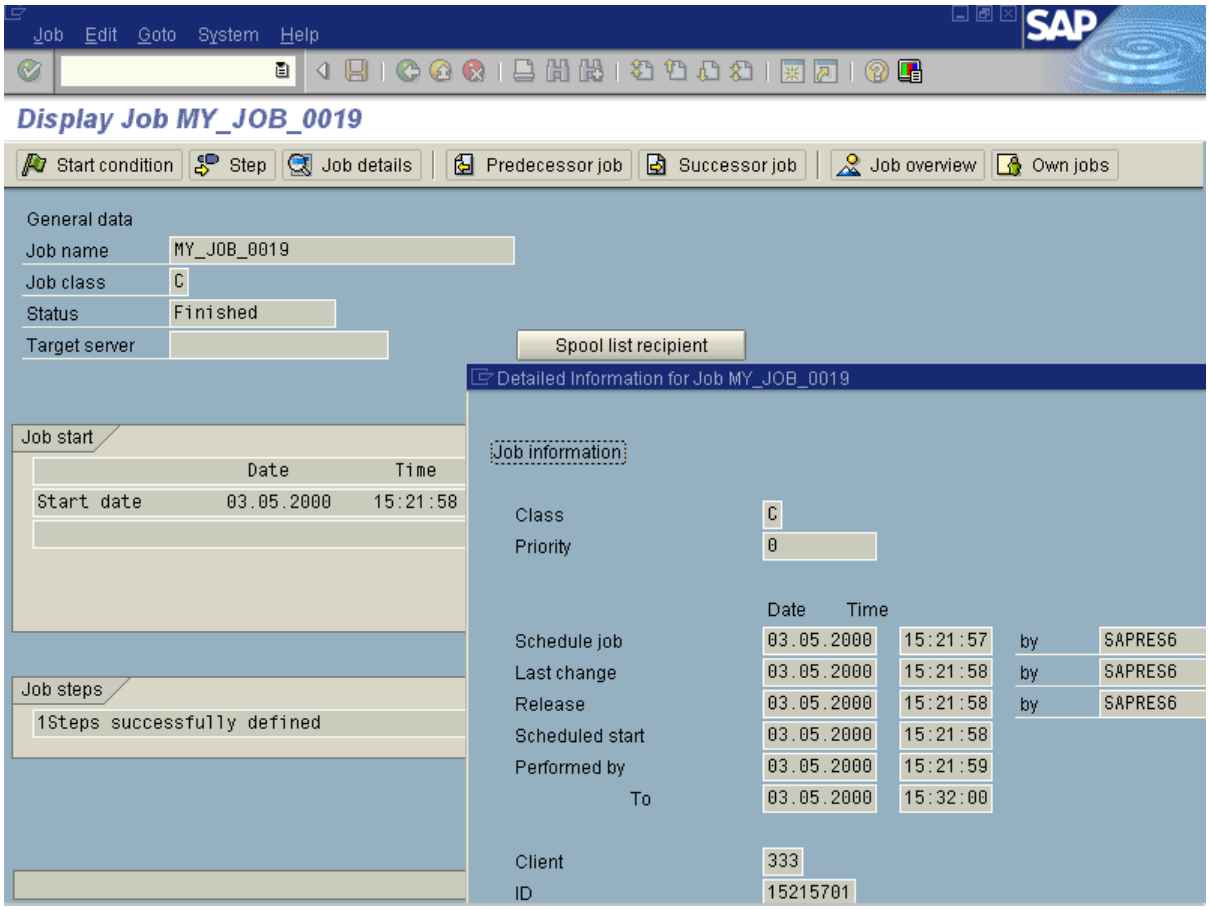


Figure 7. SM37: Successful completion of SAP R/3 workload

Result:

The successful completion of the test proved that SAP R/3 will go on running without impact if you lose a CPU in your system.

2.1.2 Features of OS/390

The OS/390 operating system has a philosophy of reliability which recognizes the inevitability of errors. This philosophy dictates a comprehensive approach of error isolation, identification and recovery rather than a simplistic automatic restart approach. In support of this approach, OS/390 provides a vast array of software reliability and availability features, far beyond that currently provided by any other operating system. Applications running on OS/390 such as SAP R/3 with its database server on DB2 UDB for OS/390 will certainly benefit from the reliability of the underlying architecture.

A large portion of the OS/390 kernel operating system exists solely to provide advanced reliability, availability and serviceability capabilities. OS/390 platform code is written to a set of RAS guidelines. These define how to provide superior RAS by fully exploiting the facilities provided by OS/390. Some of the RAS guidelines are:

- All code must be covered by a recovery routine, including the code of recovery routines themselves. Multiple layers of recovery are therefore supported.
- All control areas and queues must be verified before continuing.
- Recovery and retry must be attempted if there is hope of success.
- All failures that cannot be transparently recovered must be isolated to the smallest possible unit, for example the current request, a single task or a single address space.
- Diagnostic data must be provided which, as an objective, will allow the problem to be identified and fixed after a single occurrence. The diagnostic data is provided even when retry is attempted and successful.

In practice, about 30% of the OS/390 kernel operating system code is devoted to following the RAS guidelines. This is in addition to the code within each component that supports the RAS guidelines or deals with processor, storage or I/O error recovery.

In total, about 50% of the OS/390 kernel code deals with the inevitability of errors: isolating them, recovering from them and documenting them so they can be fixed and avoided in the future. This was truly a monumental effort, with millions of lines of code developed over 30 years, and it is a key reason for the high availability provided even by a single stand-alone OS/390.

As the RAS principles are so tightly integrated in OS/390, we will not show a matrix of OS/390 availability features here. Instead, we will just explain the most outstanding availability elements of OS/390.

2.1.2.1 List of OS/390 availability features

OS/390 delivers an integrated set of MVS, UNIX, LAN, distributed computing, and application enablement services through its base elements. These base services provide a rich set of operating system functions, all based on the RAS concept. The services can be divided into certain functional groups. All of the following services groups contribute to the high availability of OS/390 and are utilized by

SAP R/3 on S/390:

1. **System services** provide the classic strengths of MVS reliability and availability, and support for high transaction workloads and many users. Some of the key elements to provide those RAS characteristics are:

   – **System Integrity**
   OS/390 system software has a total commitment to system integrity. Data and system functions are protected from unauthorized access, whether accidentally or deliberately with sinister intent. IBM provides an integrity warranty for OS/390 which is unique in the industry. Without system integrity, there can be no assurance of security, and therefore reliability.

   – **Dynamic Operating System Customization**
   To provide maximum continuous operations, many functions and major subsystems in OS/390 are designed to be dynamically operated, configured, and tuned. In addition, customer experiences and requirements for continuous operations are seriously considered for each release of OS/390 to remove reasons for planned outages for functions such as customization and tuning changes. This is a continuous improvement process.

   – **Storage Key Protection**
   This mechanism allocates a key to each program and each piece of storage. It then ensures a match between the key of a program and the storage it is accessing. This protects the operating system from the applications it is running. It also protects the kernel of the operating system from the higher subsystems, and those subsystems from each other. It protects middleware components, such as transaction managers, which operate on behalf of many users, from errors in the application programs that are run in the same address space. Storage key protection also protects against storage overlays due to erroneous I/O operation. S/390 I/O channel processors ensure that the key of the storage being read from or to matches the key specified for the I/O operation. Without storage key protection it would be impossible for IBM to provide the OS/390 integrity warranty.

   – **Low Address Protection**
   OS/390 stores its most critical status and control information at the low end of each address space. This feature protects these vital areas from erroneous modification, even by the operating system kernel itself. A wide variety of potential system hangs are thus prevented.

   – **Functional Recovery Routines**
   A major philosophical design guideline of OS/390 from its beginning is that every part of the operating system should include recovery facilities. These facilities handle both hardware and software errors by utilizing a specific Functional Recovery Routine (FRR) of the component of OS/390 in control at the time of the error. The signal takes the form of an interrupt to a pre-specified FRR. The job of each FRR is to assess any damage, generate diagnostic information, and either repair the problem or remove the offending work unit from the system. Particularly serious problems beyond the scope of a specific FRR may be percolated up to a higher level FRR, but a major attempt is to isolate the error for minimal impact. So if an error in OS/390 itself occurs, FRRs will take over control, diagnose, and repair the failed parts of OS/390 or isolate them from other parts of OS/390 and other work in the OS/390 system.

– **Isolation from Outboard Errors**
OS/390 has defenses against outboard errors that can impact the system. One example is called hot I/O detection. I/O devices can malfunction in such a way as to harass the server with extremely frequent unsolicited interruptions. When that happens, it can stop the system from processing any useful work. OS/390 protects against hot I/O by fencing the device that is causing the problem.

2. **Security server** includes RACF and the Unix System Services DCE security server to provide full function security services. With SAP R/3 on S/390, the OS/390 security server is used to permit authorizations to users in OS/390 and UNIX System Services. The ICLI server, by default, requires the ICLI client to use Pass Ticket signon, which requires RACF Version 2.2 (or higher) or a compatible security product to be installed on OS/390.
3. **Systems management services** provide a window to enterprise-wide systems management of SAP R/3 on S/390 to keep running during most of the maintenance tasks. System Resource Manager (SRM) or Workload Manager (WLM) for example are being utilized by SAP R/3 on S/390 to set the dispatching and I/O priorities. With WLM you could give different priorities to different ICLI servers. You could distinguish batch from online work, or if you had one ICLI server for each application server, WLM could indirectly perform workload management for all of your application servers.
4. **UNIX system services** provide a reliable environment with S/390 “industrial strength” for the ICLI server.
5. **Communications server** includes TCP/IP, which is utilized as the strategic communication protocol between SAP R/3 application and database server on DB2 UDB for OS/390.

2.1.3 Features of DB2 UDB for OS/390

DB2 UDB for OS/390 was designed so that you should not have to take DB2 down in order to perform traditional database activities. Every new version of DB2 delivers new functions that are designed to ensure high availability. The development of the latest two versions V5 and V6 has been especially focused on key vendors like SAP, and it contains several high impact features which make SAP R/3 more reliable, scalable and easier to manage. Another key area of improvement was performance, based on benchmarks and customer experiences. In this chapter we discuss the main features that are built into DB2 UDB for OS/390 to improve high availability and continuous operation of the database.

2.1.3.1 List of DB2 UDB for OS/390 availability features

In this section we only discuss DB2 UDB for OS/390 availability features that apply to standalone DB2s. For DB2 data sharing and related features that apply to Parallel Sysplex configurations, and which are probably the most important DB2 availability features, please see chapter Chapter 2.2, “Additional availability features and benefits with S/390 Parallel Sysplex” on page 32.

Table 3 on page 28 summarizes the features which are implemented in the design of DB2 UDB for OS/390. It shows which availability features apply to the frequency, duration, and scope of an outage. It further explains whether this feature helps eliminate planned or unplanned outages, or both.

Table 3. DB2 UDB for OS/390 availability features matrix

Availability feature	Reduces outage frequency	Reduces outage duration	Reduces outage scope	Type of addressed outage
Online Backup with SHRLEVEL CHANGE option		X		planned
CONCURRENT Option in COPY		X		planned
CHANGELIMIT Option in COPY	X	X		planned
COPYDDN and RECOVERYDDN	X			planned
Backing up indexes		X		both
Fast Log Apply		X		unplanned
Alter Checkpoint Frequency		X		unplanned
Parallel COPY and Parallel RECOVER		X		both
ARM Support		X		unplanned
Automatic Recovery at Restart			X	unplanned
Online Reorg		X		planned
COPYDDN Option in LOAD/REORG	X	X		both
Inline Statistics		X		planned
ALTER column lengths in Indexes	X			planned
Partition Independence		X	X	planned
Partition Rebalancing	X			planned
RRS Attachment Facility		X		planned
Type 2 Indexes and Row-Level Locking	X		X	both

Legend: X=Applies

The main DB2 UDB for OS/390 availability features are in detail:

1. **Backup and Recovery**
   One of the most significant continuous operation features of DB2 is the COPY utility. It can be used to produce an online backup of a tablespace, while SAP R/3 continues to read and write the data.

   – **Online Backup with SHRLEVEL CHANGE option**
   The use of the SHRLEVEL CHANGE option produces a fuzzy image copy during concurrent SAP R/3 workload. To recover to a point of consistency, you may apply necessary log records. An important aspect of the online backup is the “incremental” online backup. This is a copy of only those tablespace data pages that have been changed since the last backup. Except for a small processor and DASD overhead, the online backup has no impact on the concurrent SAP R/3 activities.

– **CONCURRENT Option in COPY**
If you perform an offline backup of a tablespace, concurrent write activity on this particular tablespace is not allowed. The usage of the option CONCURRENT however can significantly reduce the time during which the tablespace is unavailable for write activity. The database activity will be quiesced and made available again automatically. This method does not need the separate quiesce and restart steps.

– **CHANGELIMIT Option in COPY**
The CHANGELIMIT Option in COPY allows DB2 to determine whether to take a full or incremental image copy, depending on the number of pages changed since the last image copy. With this option you can avoid running image copy jobs when very few or no pages of a table space have changed since the last image copy was taken. The savings in time can be used to maximize the use of batch windows.

– **COPYDDN and RECOVERYDDN**
The options COPYDDN and RECOVERYDDN allow you to create up to four identical copies of the tablespace.

– **Backing up indexes**
In previous DB2 versions you could not make image copies of indexes. Therefore you could recover indexes only by rebuilding the indexes from existing data. This process could be lengthy, especially if index recovery had to wait until the data was recovered, making those indexes unavailable until the rebuild was complete. In DB2 V6 you can take a full image copy or a concurrent copy of an index, just as you have always done for table spaces. To recover those indexes, you use the RECOVER utility, which restores the image copy and applies log records.

– **Fast Log Apply**
A faster log apply process in DB2 UDB for OS/390 V6 improves restart and recovery times up to 5 times in order to reduce unplanned outages. The new process sorts out log records so that changes that are to be applied to the same page or same set of pages are together. Then, using several log-apply tasks, DB2 can apply those changes in parallel. This feature requires fewer I/O operations for the log apply and can reduce CPU time.

– **Alter Checkpoint Frequency**
The new SET LOG LOGLOAD(n) command allows you to dynamically change the LOGLOAD system parameter. This parameter controls the frequency of checkpoints. The more frequent checkpoints, the faster the DB2 restart after abnormal termination. On the other hand, too frequent checkpoints negatively affect performance. The new command allows you to adjust the frequency according to your site objectives and do it dynamically, without restarting the system.
Another interesting aspect of the new command is initiating checkpoint on demand by specifying SET LOG LOGLOAD(0). E.g. it is recommended to issue this command before suspending log writes (see Chapter 5, “Creating database backups” on page 115 about using this technique for taking a system backup) in order to reduce the number of data base writes during the log write suspension and consequently reducing the exposure of generating inconsistent 32K size pages.

– **Parallel COPY and Parallel RECOVER**
Previously objects were restored serially, and you could never copy multiple objects in a single COPY job. Now with DB2 UDB for OS/390 V6,

when you specify a list of table spaces and index spaces, you can also copy and recover some or all of those objects in parallel using the new PARALLEL option of COPY and RECOVER. This feature enables a faster backup and restore of DB2 as it reduces the elapsed time of those jobs.

- **Automatic Restart Manager (ARM) Support**
  ARM is a component that helps availability by providing a fast, efficient restart of critical applications. The purpose of using automatic restart is to reduce the time a particular system is down. When DB2 stops abnormally, the surviving OS/390 systems analyze the situation to determine whether the corresponding OS/390 has failed too and where DB2 should be restarted. It then restarts DB2 appropriately, either on the same or on a different system.

- **Automatic Recovery at Restart**
  When a subsystem failure occurs, a restart of DB2 automatically restores data to a consistent state by backing out uncommitted changes and completing the processing of the committed changes.

2. **Utility Operations**

   - **Online Reorg**
     Since DB2 V5, the REORG utility allows the reorganization of an SAP R/3 tablespace or index during online operation. The keyword SHRLEVEL allows you to choose standard, read-only online, and read-write online reorganization. With the SHRLEVEL CHANGE option, you have both read and write through almost the entire reorganization process.

     The process involves these activities:

   a. The utility unloads the data from the original tablespace, sorts the data by clustering key, and reloads the data into a shadow copy. Concurrently, SAP R/3 has read and write access to the original tablespace, and changes are recorded in the log.
   b. The utility reads the log records and applies them to the shadow copy iteratively. During the last iteration, SAP R/3 has only read access.
   c. The utility switches the application access to the shadow copy by renaming the data sets for the table and the indexes. During the actual switch, SAP R/3 has no access to the data.
   d. SAP R/3 reads and writes to the renamed shadow copy.

   DB2 UDB for OS/390 V6 contains several features which further improve online reorg. You can now select rows to be discarded during a REORG, and optionally write the discarded records to a file. An external UNLOAD can be performed faster in comparison to the formerly used DSNTIAUL sample program. Furthermore, processor and elapsed time is greatly reduced.

   - **COPYDDN Option in LOAD/REORG**
     If you run REORG or LOAD REPLACE and use LOG(NO), then an image copy is required for data integrity. By default, DB2 will place the tablespace in copy-pending status, and you have to perform an image copy before you can further change the tablespace. If you run REORG or LOAD REPLACE with the COPYDDN option, a full image copy is produced during the execution of the utility and DB2 does not place the tablespace in copy-pending status. This eliminates the period of time when the table

space is in copy-pending status and a separate COPY step is not required. Therefore the data is available sooner.

– **Inline Statistics**
Prior releases of DB2 require the user to update statistics by executing RUNSTATS after common utility operations on table spaces such as LOAD, REORG and REBUILD INDEX. Version 6 lets you include RUNSTATS within the execution of those utility operations. This avoids the need for separate RUNSTATS jobs and uses less processing resource by making fewer passes over the data. Furthermore, table spaces will be made available sooner.

– **ALTER column lengths in Indexes**
Changing the length of a column of a partitioning index can affect partition boundaries and the limit key of a partition. With DB2 V6, the length of the limit key can be changed. If the length of a column that is part of the limit key is altered, and the changed column length affects the partition boundary, the limit key can change. Changing the length of the limit key ensures that the partition boundary does not change.

– **Partition Independence**
A key availability feature of DB2 UDB for OS/390 is the ability to partition a DB2 tablespace into as many as 256 partitions. In DB2 V6, each partition can be up to 64 GB. This means that one tablespace can hold up to 16 terabytes of data in a compressed or an uncompressed format.

– **Partition Rebalancing**
When data in a partitioned tablespace becomes heavily skewed, performance can be negatively affected because of contention for I/O and other resources. In that case you might want to shift data among partitions. DB2 V6 enables you to rebalance those partitions more efficiently. Partitioning key ranges can be altered while applications that access data not affected by the rebalance continue to run. The actual redistribution is accomplished by running the REORG utility for affected partitions after adjusting the key ranges. The REORG both reorganizes and rebalances the range of partitions. In prior releases of DB2, this task would require that the tablespace be unloaded, dropped, recreated, and reloaded.

– **RRS Attachment Facility**
The ICLI server exploits the Recoverable Resource Manager Services Attachment Facility (RRSAF). The RRSAF is a DB2 attachment facility that relies on an OS/390 component called OS/390 Transaction Management and Recoverable Resource Manager Services (OS/390 RRS). OS/390 RRS provides system-wide services for coordinating two-phase commit operations across OS/390 products. It is an OS/390 system logger application that records events related to protected resources. The system logger allows an application to log data from one system or from multiple systems across a Parallel Sysplex.

– **Type 2 Indexes and Row Level Locking**
Until Version 4, DB2 supported a single index structure, now known as a type 1 index. Version 4 introduced a new type of index known as a type 2 index. Type 2 indexes eliminate locks on index pages. Because there are usually fewer rows to a data page than there are index entries on a index page, locking only the data when you lock pages is likely to cause less contention than locking the index. This way you can avoid most of the deadlock and timeout problems on the index that often caused application

abends because data was not available.
Another advantage of type 2 indexes is that they let you use other function such as processing multiple parallel tasks, improved partition independence, the read-through locks, and row locking. SAP R/3 makes heavy use of row-level locking, which allows you to lock every row in a table separately. These functions improve concurrency in accessing data, and therefore avoid outages due to data not being available to SAP R/3.

2.2 Additional availability features and benefits with S/390 Parallel Sysplex

In a single system configuration with just one copy of any system component, hardware or software, you are inevitably exposed to system outages because of either failure of the component or because of planned changes to the component that require it to be taken offline. Therefore, the availability of SAP R/3 on S/390 gets even more enhanced through Parallel Sysplex features, the S/390 clustering technology. S/390 Parallel Sysplex has a design point of 99.999%, which equates to less than five minutes of unplanned downtime per year and delivers near-continuous availability. S/390 Parallel Sysplex combines S/390 hardware, OS/390 software and the DB2 database into a uniquely scalable and available infrastructure as the base for SAP R/3. Certainly the S/390 Parallel Sysplex clustering solution will leverage all the availability features of a single system, as mentioned in the preceding chapter. However it will further complement those availability features and make S/390 Parallel Sysplex the ultimate choice for SAP R/3 customer who depend on 24x7 availability for their system. This chapter will explain the availability features of S/390, OS/390 and DB2 especially targeting S/390 Parallel Sysplex. It will especially show how your SAP R/3 installation can benefit from all those features and which SAP R/3 scenarios are possible. Later in this chapter we will test the S/390 Parallel Sysplex availability features and describe the impact on SAP R/3.

S/390 Parallel Sysplex technology uses a form of clustering where all servers appear to the business application as a single server. This form of clustering, known as “single system image,” enables Parallel Sysplex to provide industry leading availability by allowing workloads to be balanced across multiple servers. Up to 32 servers can be linked together with near linear scalability, building on and extending the strengths of the S/390 hardware architecture. Every server in a S/390 Parallel Sysplex cluster has access to all data resources and every “cloned” application can run on every server. Using S/390’s coupling technology, Parallel Sysplex technology provides a “shared data” clustering technique that permits multi-system data sharing with high performance read/write integrity. This “shared data” (as opposed to “shared nothing”) approach enables workloads to be dynamically balanced across all servers in the Parallel Sysplex cluster.

SAP R/3 with its database on DB2 UDB for OS/390 can fully utilize S/390 Parallel Sysplex. It allows SAP R/3 to take advantage of the aggregate capacity of multiple servers for the probably most critical part of an SAP R/3 installation, the database server. S/390 Parallel Sysplex helps to ensure maximum system throughput and performance during peak processing periods. In the event of a hardware or software outage, either planned or unplanned, workloads can be dynamically redirected to available servers providing continuous application availability.

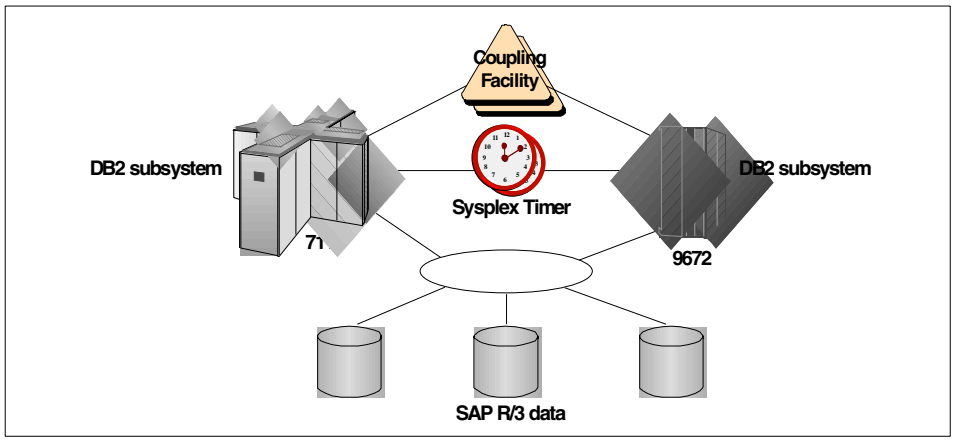


Figure 8. SAP R/3 database on DB2 UDB for OS/390 in a Parallel Sysplex

As S/390 Parallel Sysplex is designed for continuous operation, you should make sure that your configuration follows the principle of avoiding SPOFs to really achieve your availability goal. The key design point is that the configuration should tolerate a failure in any single major component. The same design concept of redundancy that applied to single S/390 systems therefore also applies to S/390 Parallel Sysplex. If at least two instances of a resource exist, then a failure of one allows the application to continue. Sysplex Timer, Coupling Facility, and Coupling Facility Links are elements that should be duplicated.

Sysplex Timer

A Sysplex Timer is used to synchronize all clocks in all S/390 processors connected to a S/390 Parallel Sysplex. It is a single point of control outside of all processors which makes sure that all Sysplex members use the same time source. This is necessary in order to guarantee the correct order of database updates. The 9037 Sysplex Timer is the unit designed to perform this task. Sometimes it is also called the External Timer Reference (ETR).The 9037 is a tabletop unit that connects via fiber optic cable to a Sysplex Timer hardware feature on the processors it is to provide timing signals for. This means that if the 9037 fails, the sysplex also fails. The 9037 is now a single point of failure for the Parallel Sysplex. To avoid that, an Expanded Availability Feature can be ordered for the 9037 which eliminates the 9037 as a single point of failure. This Expanded Availability Feature provides a second 9037 and special communication cards to allow both timers to check the other timer's status.

Coupling Facility

The function of a Coupling Facility can be performed by a standalone specialized S/390 processor classified as the 9674, or it can be operated within an LPAR of a standard S/390 processor. The 9674 CF is based on S/390 9672 hardware design and it therefore inherits all of the continuous availability function like Enhanced Processor Design, Dynamic SAP Sparing, Dynamic Memory Sparing and so on. The Internal Coupling Facility allows the same functions to be performed within an LPAR, therefore reducing cost. However, at least two CFs are required for full redundancy. Otherwise a loss of the CF in a Parallel Sysplex without a backup would mean loss of the Parallel Sysplex. To avoid that, you need a second CF. This does not mean that a second coupling facility can only be used when a failure occurs, and normally sits doing nothing. This second coupling facility can be configured to perform useful work also, but care must be taken to

ensure that sufficient spare capacity is maintained in each to allow the takeover of the other's work. This could be following a failure, or following a planned close down of the other coupling facility. Splitting the work between two facilities also allows a faster takeover, because only half the contents need to be rebuilt, not the complete contents. Remember that each member of the sysplex is responsible for rebuilding its associated information.

Coupling Facility Link

A Coupling Facility Link is the hardware which forms the communication path between the coupling facility and the S/390 processor. Fiber optic cable is used to connect the two devices. This is a high speed, high bandwidth design and although it uses the same standards of fiber optic cable used by ESCON channels, it does not use normal ESCON channel protocols. In a configuration where a coupling facility is connected to an active sysplex member by only a single coupling link, any failure in this link has exactly the same result as losing the coupling facility or the complete hardware/software of the member. That member can no longer communicate with the coupling facility, so it can no longer be part of the Parallel Sysplex. Therefore, it is essential that sufficient links are included for redundancy in this area.

2.2.1 List of S/390 Parallel Sysplex availability features

Table 4 summarizes the features which are implemented in the design of DB2 UDB for OS/390. It shows which availability features apply to the frequency, duration, and scope of an outage. It further explains whether this feature helps eliminate planned or unplanned outages, or both.

Table 4. Parallel Sysplex availability features matrix

Availability feature	Reduces outage frequency	Reduces outage duration	Reduces outage scope	Type of addressed outage
Data Sharing	X	X	X	both
Non-Disruptive Hardware Changes	X	X	X	both
Non-Disruptive Software Changes	X	X	X	both
Non-Disruptive Policy Changes	X	X	X	both

Legend: X=Applies

As these headers already express, the Parallel Sysplex solution targets all kinds of possible outages and is clearly designed to provide continuous availability. Here we will list the availability features of S/390 Parallel Sysplex in detail and explain how the SAP R/3 database server on DB2 UDB for OS/390 will benefit from these features:

1. DB2 Data Sharing

   Data sharing is a key element in the Parallel Sysplex continuous availability design. It allows the redundancy needed to overcome both the failure of a member which is processing database updates, and allows the scheduled removal of a member for service or a similar action. In each case, the service provided by the Parallel Sysplex is unaffected.

DB2 data sharing is based on the “shared everything” approach: There can be multiple DB2 subsystems belonging to one DB2 data sharing group. Each DB2 member of the data sharing group is assigned both read and write access to all data in the database. Therefore, all data must reside on shared DASD. The members of a DB2 data sharing group use coupling facility services to communicate and move data between each other. The Coupling Facility (CF) is further used by the individual members of a data sharing group to exchange locking information, system and data object status information, and to cache shared data pages with the other members of the group.

DB2 uses special data sharing locking and caching mechanisms to ensure data consistency. When one or more members of a data sharing group have opened the same table space, index space, or partition, and at least one of them has been opened for writing, then the data is said to be of "inter-DB2 R/W interest" to the members. To control access to data that is of inter-DB2 interest, DB2 uses the locking capability provided by the CF. As already mentioned, DB2 also caches the data in a storage area in the CF called a group buffer pool structure. Group buffer pools are used for caching data of interest to more than one DB2. There is one group buffer pool for all local buffer pools of the same name. When a particular page of data is changed by one DB2 subsystem, DB2 caches that page in the group buffer pool. The CF invalidates any image of the page in the local buffer pools of all the members. Then, when a request for that same data is subsequently made by another DB2, it looks for the data in the group buffer pool. The access from each DB2 to the CF is handled in a matter of microseconds, so that overall linear scalability is reached.

2. Non-disruptive Hardware Changes

   When extra capacity is needed, S/390 Parallel Sysplex gives you the ability to dynamically add new processing power to the Parallel Sysplex. The capacity can be added in incremental steps: each processor, LPAR and CEC can be added without the need to close the Parallel Sysplex first. Even the combination of bi-polar and CMOS machines is supported. The current maximum limit is 32 members, where each member can be up to 12 processors. Certainly the possibility of non-disruptive hardware changes also covers the removal of a system member from the Parallel Sysplex.

3. Non-disruptive Software Changes

   Both OS/390 and DB2 UDB for OS/390 have the ability to support non-disruptive software changes. This means that individual instances of an element can be upgraded by removing that element from the sysplex and adding the upgraded element back when ready. This demands that both the old and the new versions must happily co-exist and work together within the Parallel Sysplex. The convention for describing this release tolerance is n +1, n +2 or n+3, where n is the base.

   OS/390 supports n+3 releases, so it tolerates a range of four consecutive OS/390 releases for which coexistence is supported. For example, if you installed OS/390 Release 9 after it became generally available in March this year, this release can co-exist in a Parallel Sysplex with the next three releases of OS/390. This range rolls on with each new release of OS/390. In early releases of OS/390, coexistence was limited to three consecutive levels of OS/390. With OS/390 2.6, IBM had provided some relief of this limitation by extending the supported number of consecutive releases from 3 to 4. New

OS/390 releases are available every six months, they are currently scheduled for March and September of each year.

DB2 UDB for OS/390 supports n+1 releases. New DB2 UDB for OS/390 releases are introduced, in general, about every 18 to 24 months.
For SAP R/3 on S/390 however, there are dependencies between the SAP R/3 release you intend to run and the necessary OS/390 and DB2 UDB for OS/390 releases. For example DB2 V5 has been used by SAP R/3 on S/390 from the time on when the solution was launched, which was more than 2 years ago. DB2 UDB for OS/390 V6 is now a requirement from SAP R/3 4.6A on, and it can also be used for releases <4.6A. To find out which OS/390 and DB2 UDB for OS/390 version is required, check out OSS Note 138906 or the corresponding Info APAR II11352 on the Web page:

`http://techsupport.services.ibm.com/support/s390`

In addition, the recommended service level is documented in SAP's OSS Note 81737. This note should be reviewed on a regular basis to ensure that all PTFs fixing problems that were experienced with SAP R/3 on DB2 on OS/390 are applied.

4. Non-disruptive Policy Changes

   The Sysplex Failure Manager is used to describe a set of actions that the Parallel Sysplex should follow in the event of certain failures. These can range from the loss of a LPAR, where the remaining active LPARs can be allowed to automatically take the storage from the failing LPAR, to failures within database subsystems. The active set of instructions is known as an Sysplex Failure Manager Policy, and this policy can be changed dynamically without a service interruption.

2.2.2 Continuous Parallel Sysplex improvements

Providing one logical image to SAP R/3 across multiple servers and DB2 subsystems is only possible because S/390 hardware, OS/390 and DB2 UDB for OS/390 all play together to make DB2 data sharing on S/390 Parallel Sysplex work. Continuous development on all those impacted components makes sure that S/390 Parallel Sysplex stays a unique solution in the industry and undergoes permanent further improvement. New S/390 hardware developments and new releases of OS/390 and DB2 UDB for OS/390 all contain special features especially targeting performance, scalability and availability of S/390 Parallel Sysplex. Some of the highlights of the recent developments in this area are going to be explained in this chapter.

2.2.2.1 S/390 Hardware improvements

The S/390 hardware architecture is the underlying basis for S/390 Parallel Sysplex. Evolutionary enhancements to the system hardware provide core cluster technology functions, such as the coupling links, channel subsystem, and processor microcode supporting execution of the CF Control Code. The following features and functions exemplify the continuous improvements of Parallel Sysplex hardware technology.

- Internal Coupling Facility (ICF)

  Unique to IBM S/390 servers G3 and above with a spare, idle processor, the ICF provides the capability of trading off up to two of the spare PUs as a coupling facility. It can be used as the backup CF when an external CF is used as primary, reducing the cost for a backup coupling facility. The ICF can also

be used as a production CF even in a single-server Parallel Sysplex configuration, running two or more OS/390 logical partitions which share resources via the ICF. This configuration provides protection from software outages due to maintenance or release upgrade.

- Dynamic Internal Coupling Facility (ICF) Expansion

  Dynamic ICF Expansion permits an ICF to acquire additional processing power from the LPAR pool of shared processors being used to execute production and test work on the system. This function is very useful when the ICF partition backs up another Coupling Facility. In this event, the ICF, using Dynamic ICF Expansion, can acquire additional processing capacity to handle the full Coupling Facility workload. Also, Dynamic ICF Expansion can be used to handle a peak workload situation when the ICF is being used as the regular coupling facility. This feature provides greater flexibility to the configurability of the ICF and the Parallel Sysplex cluster, and it optimizes the use of the processing power in the system. Dynamic ICF Expansion is available to every ICF partition in S/390 G3 servers and above.

- Dynamic Coupling Facility Dispatching

  The Dynamic Coupling Facility (CF) Dispatching function helps enable continuous computing in the event of a coupling facility failure without requiring a standalone backup coupling facility. Enhanced dispatching algorithms enable you to define a backup coupling facility in a logical partition (LPAR) on your system. While this logical partition is in backup mode, although it is sharing resources with other LPARs running other active workload, it uses very little processor resource. When the backup CF becomes active, only the resource necessary to provide coupling is allocated. This feature is available from all S/390 G3 models on.

- Coupling Facility Enhancements

  Over time the CF command architecture has been extended and the CF architecture has been enhanced. When sufficient new function is defined, a new version of the CF microcode is released, supporting the new functions as defined by the latest architecture CFLEVEL. CFLEVELs 0, 1, and 2 provided base support for IBM’s first Parallel Sysplex data sharing environments, for improved performance and availability. CFLEVEL 3 targeted IMS, and CFLEVEL 4 contained enhancements also being exploited by DB2. With CFLEVEL 5, the CF supports the customer data sharing requirement for continuous availability with duplexing of DB2 cache structures. A level 6 coupling facility provides TPF record level locking in the coupling facility, and CFLEVEL 7 improves the processing efficiency of the cache structure command, exploited by DB2. A level 8 coupling facility provides a system-managed rebuild capability for list and cache structures whose users do not support user-managed rebuild, with the object of moving a structure's contents from one structure instance to another as quickly as possible. Also, two new commands -- "compute list-structure parameters" and "compute cache-structure parameters" -- have been added to replace the coupling-facility storage-allocation formulas. Further substantial enhancements to the CF command architecture and IBM's CF component are underway.

- Enhanced Links

  In 1997, IBM introduced its second-generation coupling channels, HiPerLinks (High Performance Coupling Links), to provide improved channel efficiency

and response times in processing Coupling Facility requests. Further, Coupling Facility link capacity can be improved by up to 20 percent. HiPerLinks are available on S/390 G3 servers and above. With the launch of the S/390 G5 servers, IBM introduced coupling enhancements to ensure system response matched the speed of the new CMOS processors, which are up to twice as fast as G4 Servers. Also, two new coupling links got introduced: the Internal Coupling Channel (IC) and the Integrated Cluster Bus (ICB). The IC Channel provides a high-speed internal communication link between a CF LPAR and OS/390 LPAR on the same server, eliminating the need for external hardware links. The ICB provides an external, bi-directional, high-speed bus connection very close in speed to the IC Channel. ICB is able to use nearly three-times the bandwidth of HiPerLinks, reducing the latency of synchronous send messages and increasing efficiency by one-to-three percent.

- Redundant Coupling Links

  Redundant Coupling Links are an optional feature. They can be configured between a processor and the CF. This potentially removes a SPOF for the processor's data sharing capability in the Parallel Sysplex environment.

2.2.2.2 OS/390 improvements

With the introduction of OS/390 Version 2, the focus of S/390 Parallel Sysplex clustering technology has shifted to support new application environments such as SAP R/3. These new environments are building on the broad Parallel Sysplex support already being provided by the S/390 database managers, the base operating system OS/390 and underlying S/390 hardware. Continuous improvement works toward improving the Parallel Sysplex environment for these applications. The most recent OS/390 Release 9 is an example of this process. Simplification of management, structures, and recovery and enhancements in functionality make it possible to run a more efficient Parallel Sysplex configuration. Some of the highlights are:

- UNIX System Services file system support of HFS

  In previous releases, OS/390 UNIX System Services hierarchical file systems could only be shared across a Parallel Sysplex cluster in read-only mode. To have write access to data, system users had to reside on the same system as their data. With Release 9, this constraint will be removed. Simultaneous read/write access of the same HFS from instances of UNIX System Services running on different OS/390 images in the Parallel Sysplex cluster is now possible.

- XES enhancements

  New features are a better management of CF structure resources, CF recovery enhancements, XES Event Suppression, XES Structure Full Monitoring, and more. These OS/390 enhancements will aid the applications programmers and systems programmers in troubleshooting when they face a problem. In addition, these new capabilities will help to avoid problems that can affect productivity.

- OS/390 Wizards

  OS/390 wizards can help simplify and guide you through key OS/390 tasks. Two web-based tools have been delivered, in addition to an enhancement for Release 1.3 of System Automation for OS/390, to simplify the Parallel Sysplex implementation.

– S/390 Parallel Sysplex Configuration Assistant

  Intended for those who are setting up a Parallel Sysplex for the first time, this wizard builds tailored instructions, jobs, and parmlib members for configuring a Parallel Sysplex.

– S/390 Coupling Facility Structure Sizer

  This is a web based application that will return structure sizes for the IBM products that exploit the Coupling Facility.

2.2.2.3 DB2 UDB for OS/390 Improvements

The initial delivery of DB2 data sharing came with Version 4 of DB2 for OS/390. With each new DB2 release, high availability and performance characteristics of data sharing have been improved. DB2 UDB for OS/390 Version 6 delivers the following features:

- Group Buffer Pool (GBP) Duplexing

  The failure of a CF or a CF link prior to DB2 V6 made a manual or automatic GBP recovery or a GBP rebuilt necessary. The possibility to duplex group buffer pools enables you to have a hot standby copy of a GBP ready and waiting. Each GBP is allocated in a different CF. Changed pages to both a primary and a secondary group buffer pool are written at the same time, where overlapped writes to both GBPs provide for good performance. If the primary GBP fails, DB2 UDB for OS/390 can recover quickly by switching over to the secondary GBP. Also if the secondary GBP fails, DB2 just drops back to simplex mode. If both the primary and the secondary GBP are damaged, DB2 can still use automatic GBP recovery for a duplexed GBP. GBP duplexing allows for faster recovery in the unlikely event of a CF failure and also for partial or 100% loss of connectivity. It makes the data sharing subsystems more reliable, and it reduces the time to recover the GBP.

- P-lock Overhead Reduction

  In previous DB2 versions, page P-lock latch contention could sometimes cause performance and concurrency problems on a busy system. Thanks to an IRLM performance enhancement in DB2 UDB for OS/390 V6, page P-lock processing now runs faster, improving concurrency for the entire data sharing system.

- Faster DISPLAY DATABASE

  The DISPLAY DATABASE command with the LOCKS or USE keyword sometimes could cause concurrency problems in previous versions of DB2. It now allows greater concurrency due to avoiding the main latch in IRLM query requests.

- More caching options

  DB2 gives you more options for using the coupling facility, which can reduce its overhead and improve performance. One new option lets you use the coupling facility strictly for cross-invalidation purposes, not for caching data. Data is written synchronously to DASD at commit time, so there is some performance cost for using the new option. However for applications that do sequential batch updates to a large table space, use of this option can reduce coupling facility overhead costs.

2.2.3 Parallel Sysplex failure scenarios

In this section we concentrate on tests that include the CF, and measure the impact of failures of this component on SAP R/3. Chapter 3, “Database high availability for SAP R/3” on page 61, describes tests in the DB2 data sharing environment that cover failures of application servers, DB2 subsystems, and the connections between them.

Altogether we have 3 CFs in use in our Parallel Sysplex: CF05, CF06 and CF07. Normally you would only need two CFs to perform these tests. However keep in mind that three CFs are useful for planned maintenance of a CF.

If you have a third Coupling Facility specified in the CFRM policy, it is possible to continue duplexing during the outage of the target Coupling Facility. We make use of this feature in our tests. When an operator command causes the structure to drop from duplex to simplex mode, OS/390 avoids automatically reduplexing the structure back into the same Coupling Facility from which one of the duplexed instances of the structure was just deallocated.

Instead, it duplexes into that third coupling facility. So after dropping back into simplex mode, OS/390 will automatically start duplexing the GBPs again in the third CF. If you just have 2 CFs in your installation, you will fall back in simplex mode during those tests. However, you could also still make use of automatic reduplexing. For example, if a CF fails causing a duplexed GBP to revert back to a simplexed GBP, when those CF resources become available, duplexing can be automatically restarted.

For this scenario, where we demonstrate how GBP duplexing protects from unplanned outages in the case of a crash, we use the CFs according to our preference list. Based on our preferences, they will be used in the following order: CF07, CF06, CF 05. All of our primary GBPs are allocated in CF07, and DUPLEX is set to ALLOWED in the CFRM policy. We enable duplexing by issuing the following command:

```
SETXCF START, REBUILD, DUPLEX, STRNAME=DB2V610K_GBPn
```

This command is issued for all of our GBPs: DB2V610K_GBP0, DB2V610K_GBP2, DB2V610K_GBP3, DB2V610K_GBP8K0 and DB2V610K_GBP32K. Structures for the secondary GBPs get allocated in CF06, to which we will transition during the simplex mode. CF05 will be used for the automatic reduplexing.

For the screenshots in this scenario, we will show the messages for GBP0 as an example for all GBPs, and we will concentrate on DBK4 as an example for all four DB2 subsystems.

In order to produce workload in SAP R/3, we use a tool specially developed for performance and benchmark measurements.

2.2.3.1 Scenario: Coupling Facility failure

In our SAP R/3 on S/390 installation we utilize the new feature of DB2 UDB for OS/390 V6, group buffer pool (GBP) duplexing. This feature is going to be tested in this chapter by simulating a failure, in this case even the loss of a CF.

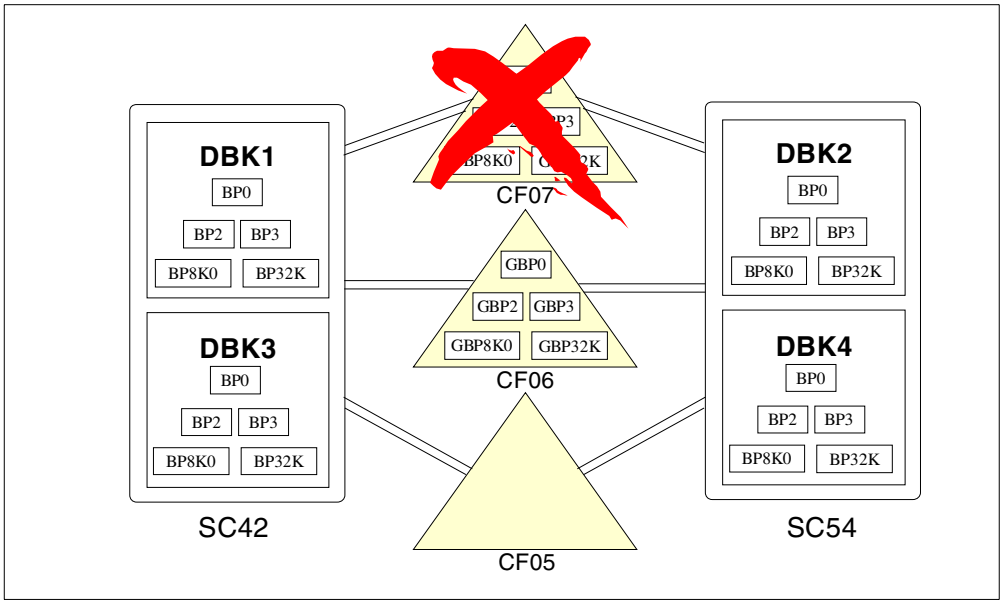


Figure 9. Scenario: Loss of Coupling Facility

Purpose:

This test demonstrates that Parallel Sysplex with GBP duplexing can tolerate the loss of a CF without affecting the SAP R/3 workload.

Method:

Two scenarios for a CF outage are possible: Either the CF with the primary or the CF with the secondary GBPs can fail. Here we will test the worst-case scenario: The failure of the CF containing the primary GBPs, which will initiate the transition to the secondary GBP. Prior to DB2 UDB for OS/390 V6, the outage of a CF would have made a manual or automatic recovery of the GBPs necessary. In that case, DB2 would have to recover any lost data from its logs, which is very time-consuming. With GBP duplexing, the time to recover the GBP is significantly reduced. The actual times can be seen in the system log.

Tasks:

1. Produce workload in the SAP R/3 system.

2. Display and print the structure information of the CF07.

```
CFNAME: CF07
  COUPLING FACILITY     :   009672.IBM.02.000000050822
                            PARTITION: E   CPCID: 00
  POLICY DUMP SPACE SIZE:   2048 K
  ACTUAL DUMP SPACE SIZE:   2048 K
  STORAGE INCREMENT SIZE:   256 K

  CONNECTED SYSTEMS:
  SC04     SC42     SC43     SC47     SC48     SC49     SC50
  SC52     SC53     SC54     SC55     SC61     SC62     SC66
  SC67     SC69

  STRUCTURES:
  DB2V610K_GBP0    DB2V610K_GBP2    DB2V610K_GBP3    DB2V610K_GBP32K
  DB2V610K_GBP8K0
```

3. Display and print the structure information of the GBPs in CF07, here GBP0 as an example.

```
D XCF,STR,STRNAME=DB2V610K_GBP0
IXC360I  15.08.57  DISPLAY XCF 740
STRNAME: DB2V610K_GBP0
 STATUS: REASON SPECIFIED WITH REBUILD START:
           OPERATOR INITIATED
         DUPLEXING REBUILD
           METHOD     : USER-MANAGED
         REBUILD PHASE: DUPLEX ESTABLISHED
  POLICY SIZE    : 32000 K
  POLICY INITSIZE: 16000 K
REBUILD PERCENT: 5
 DUPLEX         : ENABLED
 PREFERENCE LIST: CF07     CF06     CF05
 ENFORCEORDER   : NO
 EXCLUSION LIST IS EMPTY

DUPLEXING REBUILD NEW STRUCTURE
-------------------------------
 ALLOCATION TIME: 05/01/2000 11:25:31
 CFNAME         : CF06
 COUPLING FACILITY: 009672.IBM.02.000000049305
                    PARTITION: F   CPCID: 00
 ACTUAL SIZE    : 16128 K
 STORAGE INCREMENT SIZE: 256 K
 PHYSICAL VERSION: B3FA08CB 0DF71202
 LOGICAL  VERSION: B3FA08CB 0DF71202
 SYSTEM-MANAGED PROCESS LEVEL: 8
 DISPOSITION    : DELETE
 ACCESS TIME    : 0
MAX CONNECTIONS: 32
 # CONNECTIONS  : 4

DUPLEXING REBUILD OLD STRUCTURE
-------------------------------
 ALLOCATION TIME: 04/28/2000 14:47:39
 CFNAME         : CF07
 COUPLING FACILITY: 009672.IBM.02.000000050822
                    PARTITION: E   CPCID: 00
 ACTUAL SIZE    : 16128 K
 STORAGE INCREMENT SIZE: 256 K
 PHYSICAL VERSION: B3F67060 844A5043
 LOGICAL  VERSION: B3F67060 844A5043
 SYSTEM-MANAGED PROCESS LEVEL: 8
 ACCESS TIME    : 0
 MAX CONNECTIONS: 32
 # CONNECTIONS  : 4

CONNECTION NAME  ID VERSION  SYSNAME  JOBNAME  ASID STATE
---------------- -- -------- -------- -------- ---- ----------------
 DB2_DBK1         01 0001000B SC42     DBK1DBM1 00AA ACTIVE NEW,OLD
 DB2_DBK2         04 00040004 SC54     DBK2DBM1 006E ACTIVE NEW,OLD
 DB2_DBK3         02 00020006 SC42     DBK3DBM1 009A ACTIVE NEW,OLD
 DB2_DBK4         03 00030002 SC54     DBK4DBM1 0073 ACTIVE NEW,OLD
```

4. We powered off CF07. The system log shows that the connection from both SC42 and SC54 to CF07 gets lost. Note the timestamp 15:10:07.

```
 15:10:07.12 IXC518I SYSTEM SC54 NOT USING 322
         COUPLING FACILITY  009672.IBM.02.000000050822
                            PARTITION: E    CPCID: 00
         NAMED CF07
        REASON: CONNECTIVITY LOST.
        REASON FLAG: 13300002.
*IXL158I PATH 24 IS NOW NOT-OPERATIONAL TO CUID: FFFD 396
                    COUPLING FACILITY 009672.IBM.02.000000050822
                                      PARTITION: E  CPCID: 00
*IXL158I PATH 84 IS NOW NOT-OPERATIONAL TO CUID: FFFD 397
                    COUPLING FACILITY 009672.IBM.02.000000050822
                                      PARTITION: E  CPCID: 00
*IXL158I PATH 24 IS NOW NOT-OPERATIONAL TO CUID: FFFD 796
                    COUPLING FACILITY 009672.IBM.02.000000050822
                                      PARTITION: E  CPCID: 00
*IXL158I PATH 84 IS NOW NOT-OPERATIONAL TO CUID: FFFD 797
                    COUPLING FACILITY 009672.IBM.02.000000050822
                                      PARTITION: E  CPCID: 00


D XCF,CF
COUPLING FACILITY 009672.IBM.02.000000050822
                  PARTITION: E  CPCID: 00
                  CONTROL UNIT ID: FFFD
 NAMED CF07
 NO COUPLING FACILITY SPACE DATA AVAILABLE

 SENDER PATH       PHYSICAL            LOGICAL       CHANNEL TYPE
         24         PATH NOT AVAIL      ONLINE        CFS
         84         PATH NOT AVAIL      ONLINE        CFS
```

5. DBK1 to DBK4 report a loss of connectivity to the GBPs. DB2 stops duplexing first and runs in simplex mode. For the simplex mode, DB2 automatically reverts to the secondary GBPs in CF06.
   Capture information from the OS/390 system log which documents the transition to CF06 due to the lost connection to CF07.

```
DSNB228I =DBK4 DSNB5STS GROUP BUFFER POOL GBP0 323
           CANNOT BE ACCESSED FOR READ_STGSTATS
           MVS IXLCACHE REASON CODE=X'0C1C0C06'

DSNB303E =DBK4 DSNB1REE A LOSS OF CONNECTIVITY WAS 332
DETECTED TO GROUP BUFFER POOL GBP0

IXC522I REBUILD FOR STRUCTURE DB2V610K_GBP0 327
IS BEING STOPPED TO SWITCH TO THE NEW STRUCTURE DUE TO
LOSS OF CONNECTIVITY TO THE STRUCTURE

DSNB744I =DBK4 DSNB1GBR DUPLEXING IS BEING STOPPED 334
FOR GROUP BUFFER POOL GBP0
           SWITCHING TO SECONDARY
           REASON = LOSSCONNPRI
           DB2 REASON CODE = 00000000

DSNB745I =DBK4 DSNB1GBR THE TRANSITION BACK TO 353
SIMPLEX MODE HAS COMPLETED FOR
           GROUP BUFFER POOL GBP0
```

6. If the CFRM policy specifies DUPLEX(ALLOWED) or DUPLEX(DISABLED), the structure simply goes into simplex mode as shown in the step before. Our CFRM policy specifies DUPLEX(ENABLED), so OS/390 tries to automatically restart duplexing. Note that duplexing was successfully reestablished at 15:11:52.

```
DSNB740I =DBK4 DSNB1RBQ ATTEMPTING TO ESTABLISH 355
DUPLEXING FOR
           GROUP BUFFER POOL GBP0
           REASON = POLICY


IXC529I DUPLEX REBUILD NEW STRUCTURE DB2V610K_GBP0 775
IS BEING ALLOCATED IN COUPLING FACILITY CF05.
OLD STRUCTURE IS ALLOCATED IN COUPLING FACILITY CF06.
REBUILD START REASON: MVS INITIATED BASED ON POLICY
INFO108: 00000007 00000000.
DSNB302I =DBK3 DSNB1RBC GROUP BUFFER POOL GBP0-SEC IS 7


IXL014I IXLCONN REBUILD REQUEST FOR STRUCTURE DB2V610K_GBP0 377
WAS SUCCESSFUL.  JOBNAME: DBK4DBM1 ASID: 0073
CONNECTOR NAME: DB2_DBK4 CFNAME: CF05
DSNB332I =DBK4 DSNB1PCD THIS MEMBER HAS COMPLETED 378
CASTOUT OWNER WORK FOR GROUP BUFFER POOL GBP0


15:12:51.86   DSNB742I =DBK4 DSNB1GBR DUPLEXING HAS BEEN 403
SUCCESSFULLY ESTABLISHED FOR
           GROUP BUFFER POOL GBP0


IXL157I PATH 24 IS NOW OPERATIONAL TO CUID: FFFD 838
                  COUPLING FACILITY 009672.IBM.02.000000050822
                                    PARTITION: E  CPCID: 00
 IXL157I PATH 84 IS NOW OPERATIONAL TO CUID: FFFD 839
                  COUPLING FACILITY 009672.IBM.02.000000050822
                                    PARTITION: E  CPCID: 00


D XCF,STR
DB2V610K_GBP0    05/01/2000 15:11:52 DUPLEXING REBUILD NEW STRUCTURE
                                     DUPLEXING REBUILD
                                       METHOD     : USER-MANAGED
                                     REBUILD PHASE: DUPLEX ESTABLISHED
```

7. We verified that the SAP R/3 workload is still running. We took a screenshot to display that; see Figure 10.

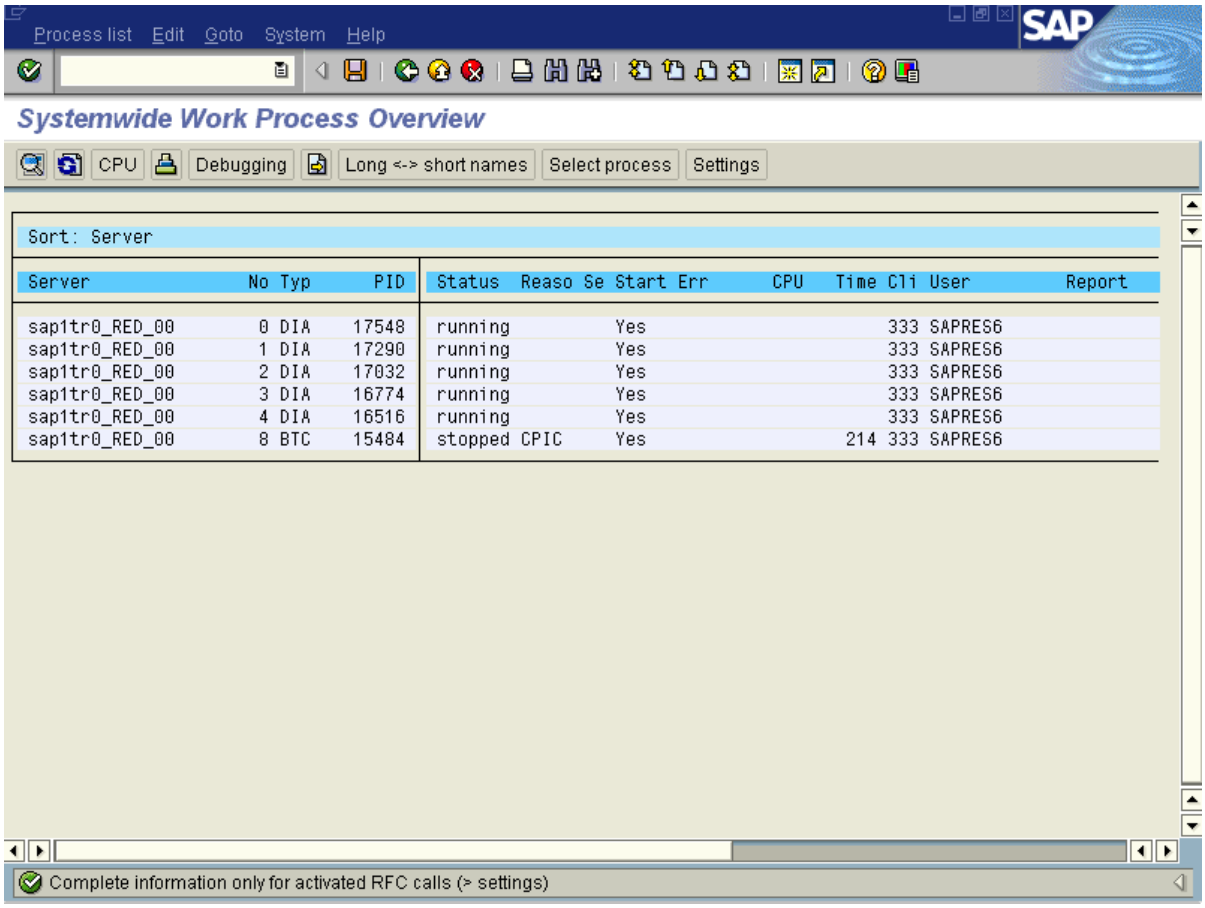


Figure 10. SM66: Online SAP R/3´workload during CF outage

8. We powered on CF07 and printed the message that indicates that it is available for use.

```
D XCF,CF,CFNAME=CF07
IXC362I  15.16.59  DISPLAY XCF 932
CFNAME: CF07
  COUPLING FACILITY     :   009672.IBM.02.000000050822
                            PARTITION: E   CPCID: 00
  POLICY DUMP SPACE SIZE:   2048 K
  ACTUAL DUMP SPACE SIZE:   2048 K
  STORAGE INCREMENT SIZE:   256 K
CONNECTED SYSTEMS:
  SC04     SC42     SC43     SC47     SC48     SC49     SC50
  SC52     SC53     SC54     SC55     SC61     SC62     SC66
  SC67     SC69

  NO STRUCTURES ARE IN USE BY THIS SYSPLEX IN THIS COUPLING FACILITY
```

This command is used to populate a Coupling Facility that is newly brought into service or is being returned to service after a maintenance operation:

```
SETXCF START,REBUILD,POPULATECF,CFNAME=cfname
```

9. We displayed the new allocation of the GBPs in CF06 and CF05.

```
D XCF,STR,STRNAME=DB2V610K_GBP0
IXC360I  15.25.15  DISPLAY XCF 896
STRNAME: DB2V610K_GBP0
 STATUS: REASON SPECIFIED WITH REBUILD START:
           POLICY-INITIATED
         DUPLEXING REBUILD
           METHOD     : USER-MANAGED
         REBUILD PHASE: DUPLEX ESTABLISHED
  POLICY SIZE    : 32000 K
  POLICY INITSIZE: 16000 K
  REBUILD PERCENT: 5
  DUPLEX         : ENABLED
  PREFERENCE LIST: CF07     CF06     CF05
  ENFORCEORDER   : NO
  EXCLUSION LIST IS EMPTY

 DUPLEXING REBUILD NEW STRUCTURE
 -------------------------------
 ALLOCATION TIME: 05/01/2000 15:11:52
 CFNAME         : CF05
 COUPLING FACILITY: 009672.IBM.02.000000050822
                    PARTITION: F   CPCID: 00
 ACTUAL SIZE    : 16128 K
 STORAGE INCREMENT SIZE: 256 K
 PHYSICAL VERSION: B3FA3B62 B4015286
 LOGICAL  VERSION: B3FA3B62 B4015286
 SYSTEM-MANAGED PROCESS LEVEL: 8
 DISPOSITION    : DELETE
 ACCESS TIME    : 0
 MAX CONNECTIONS: 32
 # CONNECTIONS  : 4

DUPLEXING REBUILD OLD STRUCTURE
-------------------------------
 ALLOCATION TIME: 05/01/2000 11:25:31
 CFNAME         : CF06
 COUPLING FACILITY: 009672.IBM.02.000000049305
 PARTITION: F   CPCID: 00
 ACTUAL SIZE    : 16128 K
 STORAGE INCREMENT SIZE: 256 K
 PHYSICAL VERSION: B3FA08CB 0DF71202
 LOGICAL  VERSION: B3FA08CB 0DF71202
 SYSTEM-MANAGED PROCESS LEVEL: 8
 ACCESS TIME    : 0
 MAX CONNECTIONS: 32
 # CONNECTIONS  : 4

CONNECTION NAME  ID VERSION  SYSNAME  JOBNAME  ASID STATE
---------------- -- -------- -------- -------- ---- ----------------
DB2_DBK1         01 0001000B SC42     DBK1DBM1 00AA ACTIVE NEW,OLD
DB2_DBK2         04 00040004 SC54     DBK2DBM1 006E ACTIVE NEW,OLD
DB2_DBK3         02 00020006 SC42     DBK3DBM1 009A ACTIVE NEW,OLD
DB2_DBK4         03 00030002 SC54     DBK4DBM1 0073 ACTIVE NEW,OLD
```

10. Normally with three CFs there would be no need to activate the primary GBP again because of the automatic reduplexing that occurred. We could continue running with CF06 and CF05.

    However since we plan for another Coupling Facility outage scenario, we will move the GBPs back to CF07 after completion of this test, in order to have the same situation like at the beginning of this test.

To move the GBPs back to CF07, we simply stop duplexing again. As DB2 automatically tries reduplexing, it will duplex all GBPs based on our preference list for the structures. CF07 has the highest priority in this list, so DB2 will create the duplexed GBPs in CF07. Therefore we issue the following command for all GBPs:

```
SETXCF,STOP,REBUILD,DUPLEX,STRNAME=DB2V610K_GBPn,KEEP=OLD
```

Note: Keep in mind that if you ran this scenario with just two CFs without automatic reduplexing, you would still be in simplex mode right now and should restart duplexing with the SETXCF START command.

The following messages appeared for DBK1 to DBK4:

```
SETXCF STOP,REBUILD,DUPLEX,STRNAME=DB2V610K_GBP0,KEEP=OLD

DSNB743I =DBK4 DSNB1GBR DUPLEXING IS BEING STOPPED 947
 FOR GROUP BUFFER POOL GBP0
             FALLING BACK TO PRIMARY
             REASON = OPERATOR
             DB2 REASON CODE = 00000000

DSNB745I =DBK4 DSNB1GBR THE TRANSITION BACK TO 953
SIMPLEX MODE HAS COMPLETED FOR
            GROUP BUFFER POOL GBP0

DSNB740I =DBK4 DSNB1RBQ ATTEMPTING TO ESTABLISH 955
DUPLEXING FOR
            GROUP BUFFER POOL GBP0
            REASON = POLICY

IXC529I DUPLEX REBUILD NEW STRUCTURE DB2V610K_GBP0 790
IS BEING ALLOCATED IN COUPLING FACILITY CF07.

IXL014I IXLCONN REBUILD REQUEST FOR STRUCTURE DB2V610K_GBP0 957
WAS SUCCESSFUL.  JOBNAME: DBK4DBM1 ASID: 0073
CONNECTOR NAME: DB2_DBK4 CFNAME: CF07

10:44:35.86    DSNB742I =DBK4 DSNB1GBR DUPLEXING HAS BEEN 966
SUCCESSFULLY ESTABLISHED FOR
            GROUP BUFFER POOL GBP0
```

11. We displayed and printed CF information to verify that all structures are correctly spread across both CFs.

12.

```
D XCF,CF,CFNAME=CF07
IXC362I  10.20.47  DISPLAY XCF 215
CFNAME: CF07
  COUPLING FACILITY     :   009672.IBM.02.000000050822
                            PARTITION: E   CPCID: 00
  POLICY DUMP SPACE SIZE:   2048 K
  ACTUAL DUMP SPACE SIZE:   2048 K
  STORAGE INCREMENT SIZE:   256 K

  CONNECTED SYSTEMS:
  SC04      SC42      SC43      SC47      SC48      SC49      SC50
  SC52      SC53      SC54      SC55      SC61      SC62      SC66
  SC67      SC69

  STRUCTURES:
  DB2V610K_GBP0     DB2V610K_GBP2     DB2V610K_GBP3     DB2V610K_GBP32K
  DB2V610K_GBP8K0
```

13. We again verify that the SAP R/3 workload is still running. The screenshot shows that the job ran from 15:09:50 to 15:18:01 and completed successfully.

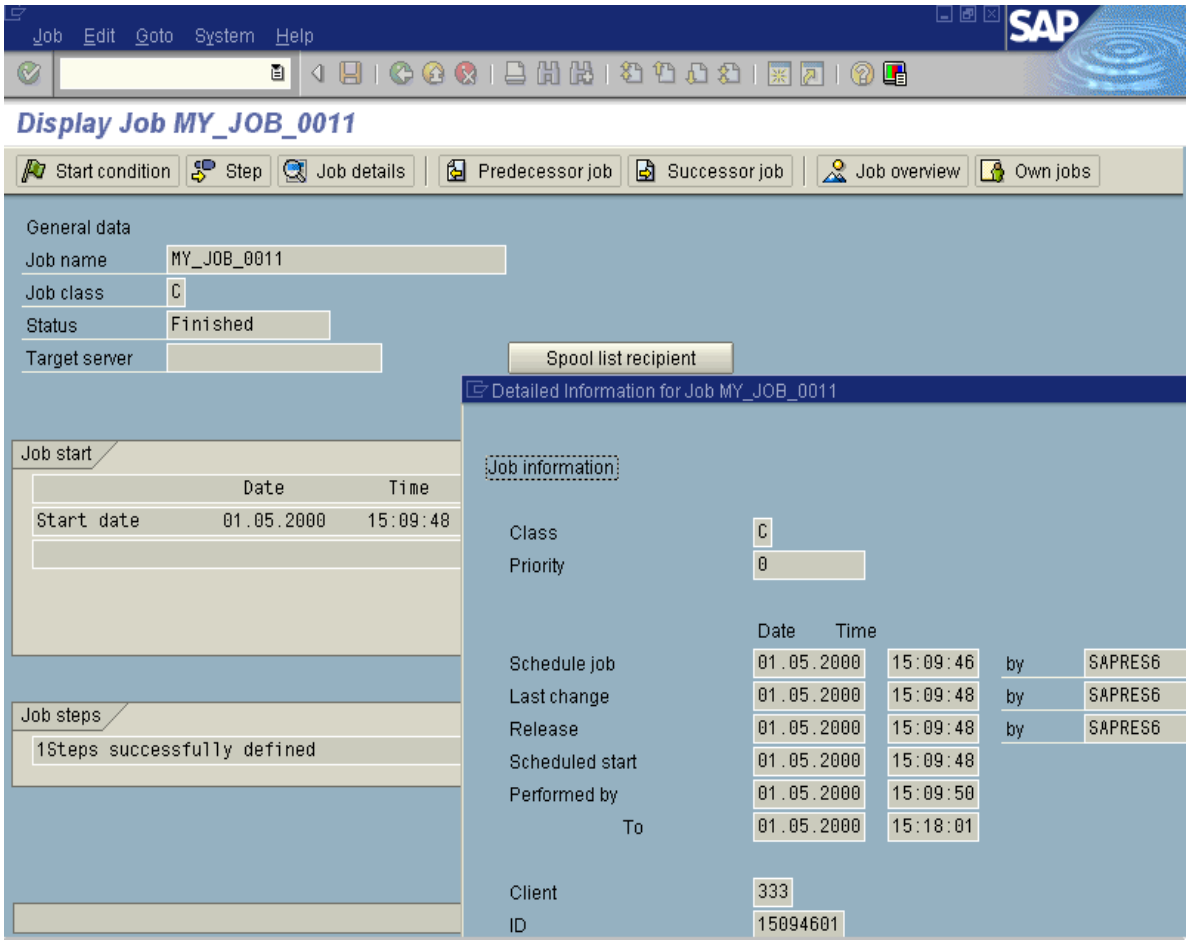


Figure 11. SM37: Successful completion of the SAP R/3 workload

Result
The successful completion of the test proved that an outage of a CF does not have any impact on running SAP R/3 workload.

2.2.3.2 Scenario: Coupling Facility link failure

This test scenario is again based on DB2 group buffer pool (GBP) duplexing. This time we will test GBP duplexing by simulating the partial loss of connectivity to the CF.

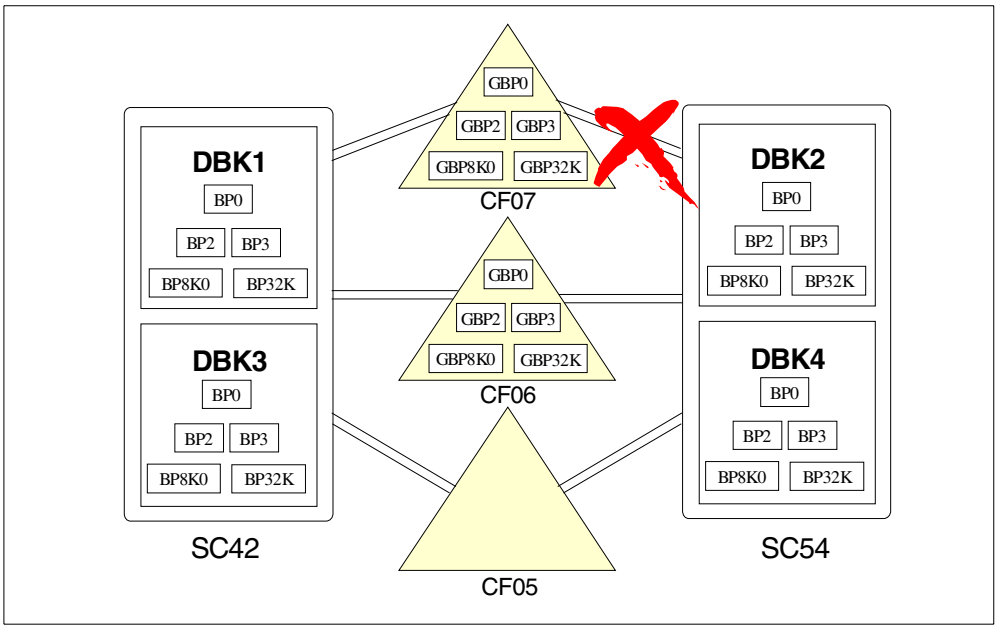


Figure 12. Scenario: Partial loss of connectivity to Coupling Facility

Purpose
This test demonstrates that Parallel Sysplex with GBP duplexing can tolerate a CF link failure without affecting the SAP R/3 workload.

Method
Again, certain scenarios for a CF link failure are possible: either the links to the CF with the primary, or the links to the CF with the secondary GBPs, can fail. Here we will test the partial loss of connectivity to the CF with the primary GBPs.

The transition to the second CF will again be initiated, and automatic reduplexing in the third CF will occur. Prior to DB2 UDB for OS/390 V6, this kind of disruption would have made a GBP rebuild necessary. A GBP rebuild requires connectivity to that GBP by at least one member. During GBP rebuild, the member that still has the connectivity copies only the changed data in the GBP to a newly allocated GBP.

With GBP duplexing, DB2 automatically switches to the secondary GBP. The actual processing time can be seen in the screenshots from the system log.

Just as at the beginning of the last test, our primary GBPs are allocated in CF07 and the secondary GBPs in CF06, to which we will transition during the simplex mode. CF05 will be used for the automatic reduplexing during the outage of CF07.

As CF07 is an Internal Coupling Facility (ICF), we cannot physically unplug the links that connect the S/390 to the CF. Instead, we will simulate the partial loss of connectivity by configuring the CHPIDs offline.

Tasks

1. Produce workload in the SAP R/3 system.
2. Display and print the structure information of the CF07. We see that all our GBPs are allocated in this CF, and that CHPIDs 01 and 32 are used.

```
D XCF,CF,CFNAME=CF07

RESPONSE=SC54
 IXC362I  10.34.28  DISPLAY XCF 065
 CFNAME: CF07
   COUPLING FACILITY     :   009672.IBM.02.000000050822
                             PARTITION: E   CPCID: 00
   POLICY DUMP SPACE SIZE:   2048 K
   ACTUAL DUMP SPACE SIZE:   2048 K
   STORAGE INCREMENT SIZE:   256 K

   CONNECTED SYSTEMS:
   SC04     SC42     SC43     SC47     SC48     SC49     SC50
   SC52     SC53     SC54     SC55     SC61     SC62     SC66
   SC67     SC69

   STRUCTURES:
   DB2V610K_GBP0    DB2V610K_GBP2    DB2V610K_GBP3    DB2V610K_GBP32K
   DB2V610K_GBP8K0


D CF,CFNAME=CF07
 IXL150I  10.39.37  DISPLAY CF 067
 COUPLING FACILITY 009672.IBM.02.000000050822
                   PARTITION: E  CPCID: 00
                   CONTROL UNIT ID: FFFC
 NAMED CF07
 COUPLING FACILITY SPACE UTILIZATION
  ALLOCATED SPACE                 DUMP SPACE UTILIZATION
   STRUCTURES:     144640 K        STRUCTURE DUMP TABLES:        0 K
   DUMP SPACE:       2048 K                  TABLE COUNT:        0
  FREE SPACE:      100352 K       FREE DUMP SPACE:            2048 K
 TOTAL SPACE:      247040 K      TOTAL DUMP SPACE:            2048 K
                                MAX REQUESTED DUMP SPACE:        0 K
    VOLATILE:         YES         STORAGE INCREMENT SIZE:      256 K
     CFLEVEL:           9
COUPLING FACILITY SPACE CONFIGURATION
                          IN USE          FREE          TOTAL
CONTROL SPACE:          146688 K      100352 K      247040 K
NON-CONTROL SPACE:           0 K           0 K           0 K

SENDER PATH       PHYSICAL           LOGICAL       CHANNEL TYPE
        01         ONLINE             ONLINE        ICS
        31         ONLINE             ONLINE        ICS

COUPLING FACILITY DEVICE     SUBCHANNEL     STATUS
                  FFD6         11CF         OPERATIONAL/IN USE
                  FFD7         11D0         OPERATIONAL/IN USE
                  FFE2         11D1         OPERATIONAL/IN USE
                  FFE3         11D2         OPERATIONAL/IN USE
```

3. Display and print the structure information of the GBPs in CF07, here GBP0 as an example. We see that the primary GBP is allocated in CF07, the secondary in CF06.

```
D XCF,STR,STRNAME=DB2V610K_GBP0
IXC360I  10.30.26  DISPLAY XCF 061
STRNAME: DB2V610K_GBP0
 STATUS: REASON SPECIFIED WITH REBUILD START:
           POLICY-INITIATED
         DUPLEXING REBUILD
           METHOD     : USER-MANAGED
         REBUILD PHASE: DUPLEX ESTABLISHED
  POLICY SIZE    : 32000 K
  POLICY INITSIZE: 16000 K
  REBUILD PERCENT: 5
  DUPLEX         : ENABLED
  PREFERENCE LIST: CF07     CF06     CF05
  ENFORCEORDER   : NO
  EXCLUSION LIST IS EMPTY

DUPLEXING REBUILD NEW STRUCTURE
-------------------------------
 ALLOCATION TIME: 05/02/2000 10:03:50
 CFNAME         : CF07
 COUPLING FACILITY: 009672.IBM.02.000000050822
                    PARTITION: E   CPCID: 00
 ACTUAL SIZE    : 16128 K
 STORAGE INCREMENT SIZE: 256 K
 PHYSICAL VERSION: B3FB3866 19049286
 LOGICAL  VERSION: B3FB3866 19049286
 SYSTEM-MANAGED PROCESS LEVEL: 8
 DISPOSITION    : DELETE
 ACCESS TIME    : 0
 MAX CONNECTIONS: 32
 # CONNECTIONS  : 4

DUPLEXING REBUILD OLD STRUCTURE
-------------------------------
 ALLOCATION TIME: 05/01/2000 11:25:31
CFNAME         : CF06
  COUPLING FACILITY: 009672.IBM.02.000000049305
                     PARTITION: F   CPCID: 00
  ACTUAL SIZE    : 16128 K
  STORAGE INCREMENT SIZE: 256 K
  PHYSICAL VERSION: B3FA08CB 0DF71202
  LOGICAL  VERSION: B3FA08CB 0DF71202
  SYSTEM-MANAGED PROCESS LEVEL: 8
  ACCESS TIME    : 0
  MAX CONNECTIONS: 32
  # CONNECTIONS  : 4

 CONNECTION NAME  ID VERSION  SYSNAME  JOBNAME  ASID STATE
 ---------------- -- -------- -------- -------- ---- ----------------
 DB2_DBK1         01 0001000B SC42     DBK1DBM1 00AA ACTIVE NEW,OLD
 DB2_DBK2         04 00040004 SC54     DBK2DBM1 006E ACTIVE NEW,OLD
 DB2_DBK3         02 00020006 SC42     DBK3DBM1 009A ACTIVE NEW,OLD
 DB2_DBK4         03 00030002 SC54     DBK4DBM1 0073 ACTIVE NEW,OLD
```

4. Configure the CHPIDs of SC54 and CF07 offline. The system log shows that the connection from SC54 to the first CHPID of CF07 gets lost at 10:42:48. The connection is completely interrupted somewhat after 10:43:03, after the manual confirmation “CONTINUE” is received from the operator.

```
10:42:48.49 HAIMO    00000290  CF CHP(01),OFFLINE,FORCE

IEE503I CHP(1),OFFLINE
IEE712I CONFIG   PROCESSING COMPLETE

10:43:03.28 HAIMO    00000290  CF CHP(31),OFFLINE,FORCE

*IXL126I CONFIG WILL FORCE OFFLINE LAST CHP(31) TO COUPLING FACILITY
 CF07
*020 IXL127A REPLY CANCEL OR CONTINUE
 R 20,CONTINUE
 R 20,CONTINUE
 IEE600I REPLY TO 020 IS;CONTINUE
 IXC518I SYSTEM SC54 NOT USING 076
         COUPLING FACILITY  009672.IBM.02.000000050822
                            PARTITION: E    CPCID: 00
          NAMED CF07
        REASON: CONNECTIVITY LOST.
        REASON FLAG: 13300002.

IEE503I CHP(31),OFFLINE
IEE712I CONFIG   PROCESSING COMPLETE
```

5. DBK1 to DBK4 report a loss of connectivity to the GBPs. DB2 stops duplexing first and runs in simplex mode. For the simplex mode, DB2 automatically reverts to the secondary GBPs in CF06.
   The information from the OS/390 system log documents the transition to CF06 due to the lost connection to CF07.

```
IXC522I REBUILD FOR STRUCTURE DB2V610K_GBP0 077
IS BEING STOPPED TO FALL BACK TO THE OLD STRUCTURE DUE TO
LOSS OF CONNECTIVITY TO THE STRUCTURE

DSNB303E =DBK4 DSNB1REE A LOSS OF CONNECTIVITY WAS 084
DETECTED
           TO GROUP BUFFER POOL GBP0-SEC

DSNB743I =DBK4 DSNB1GBR DUPLEXING IS BEING STOPPED 086
FOR GROUP BUFFER POOL GBP0
           FALLING BACK TO PRIMARY
           REASON = LOSSCONNSEC
           DB2 REASON CODE = 00000000

IXC579I PENDING DEALLOCATION FOR STRUCTURE DB2V610K_GBP0 IN 882
        COUPLING FACILITY  009672.IBM.02.000000050822
                           PARTITION: E    CPCID: 00
HAS BEEN COMPLETED.

DSNB745I =DBK2 DSNB1GBR THE TRANSITION BACK TO 107
SIMPLEX MODE HAS COMPLETED FOR
           GROUP BUFFER POOL GBP8K0
```

6. If the CFRM policy specifies DUPLEX(ALLOWED) or DUPLEX(DISABLED), the structure simply goes into simplex mode as shown in the step before. Our CFRM policy specifies DUPLEX(ENABLED), so OS/390 tries to automatically restart duplexing. Note that it successfully completes reduplexing at 10:44:19.

```
IXC536I DUPLEXING REBUILD OF STRUCTURE DB2V610K_GBP0 122
INITIATED.
REASON: PREVIOUS REBUILD PROCESS COMPLETED

DSNB740I =DBK4 DSNB1RBQ ATTEMPTING TO ESTABLISH 120
DUPLEXING FOR
           GROUP BUFFER POOL GBP0
           REASON = POLICY

IXC529I DUPLEX REBUILD NEW STRUCTURE DB2V610K_GBP0 126
 IS BEING ALLOCATED IN COUPLING FACILITY CF05.
 OLD STRUCTURE IS ALLOCATED IN COUPLING FACILITY CF06.
 REBUILD START REASON: MVS INITIATED BASED ON POLICY
 INFO108: 00000007 00000000.
 IXL014I IXLCONN REBUILD REQUEST FOR STRUCTURE DB2V610K_GBP0 127
 WAS SUCCESSFUL.  JOBNAME: DBK4DBM1 ASID: 0073
 CONNECTOR NAME: DB2_DBK4 CFNAME: CF05
 IXL015I REBUILD NEW STRUCTURE ALLOCATION INFORMATION FOR 128
 STRUCTURE DB2V610K_GBP0, CONNECTOR NAME DB2_DBK4
  CFNAME     ALLOCATION STATUS/FAILURE REASON
  --------   --------------------------------
  CF07       INSUFFICIENT CONNECTIVITY 00000002
  CF06       RESTRICTED BY REBUILD OTHER
  CF05       STRUCTURE ALLOCATED
 DSNB302I =DBK4 DSNB1RBC GROUP BUFFER POOL GBP0-SEC IS 129
            ALLOCATED IN A VOLATILE STRUCTURE
DSNB332I =DBK4 DSNB1PCD THIS MEMBER HAS COMPLETED 130
CASTOUT OWNER WORK FOR GROUP BUFFER POOL GBP0

DSNB333I =DBK4 DSNB1GBR FINAL SWEEP COMPLETED FOR 140
GROUP BUFFER POOL GBP0

DSNB742I =DBK4 DSNB1GBR DUPLEXING HAS BEEN 165
SUCCESSFULLY ESTABLISHED FOR
GROUP BUFFER POOL GBP0

D XCF,STR
DB2V610K_GBP0    05/02/2000 10:44:19 DUPLEXING REBUILD NEW STRUCTURE
                                     DUPLEXING REBUILD
                                       METHOD     : USER-MANAGED
                                     REBUILD PHASE: DUPLEX ESTABLISHED
```

7. We verified that the SAP R/3 workload is still running, and took a screenshot to display that; see Figure 13.

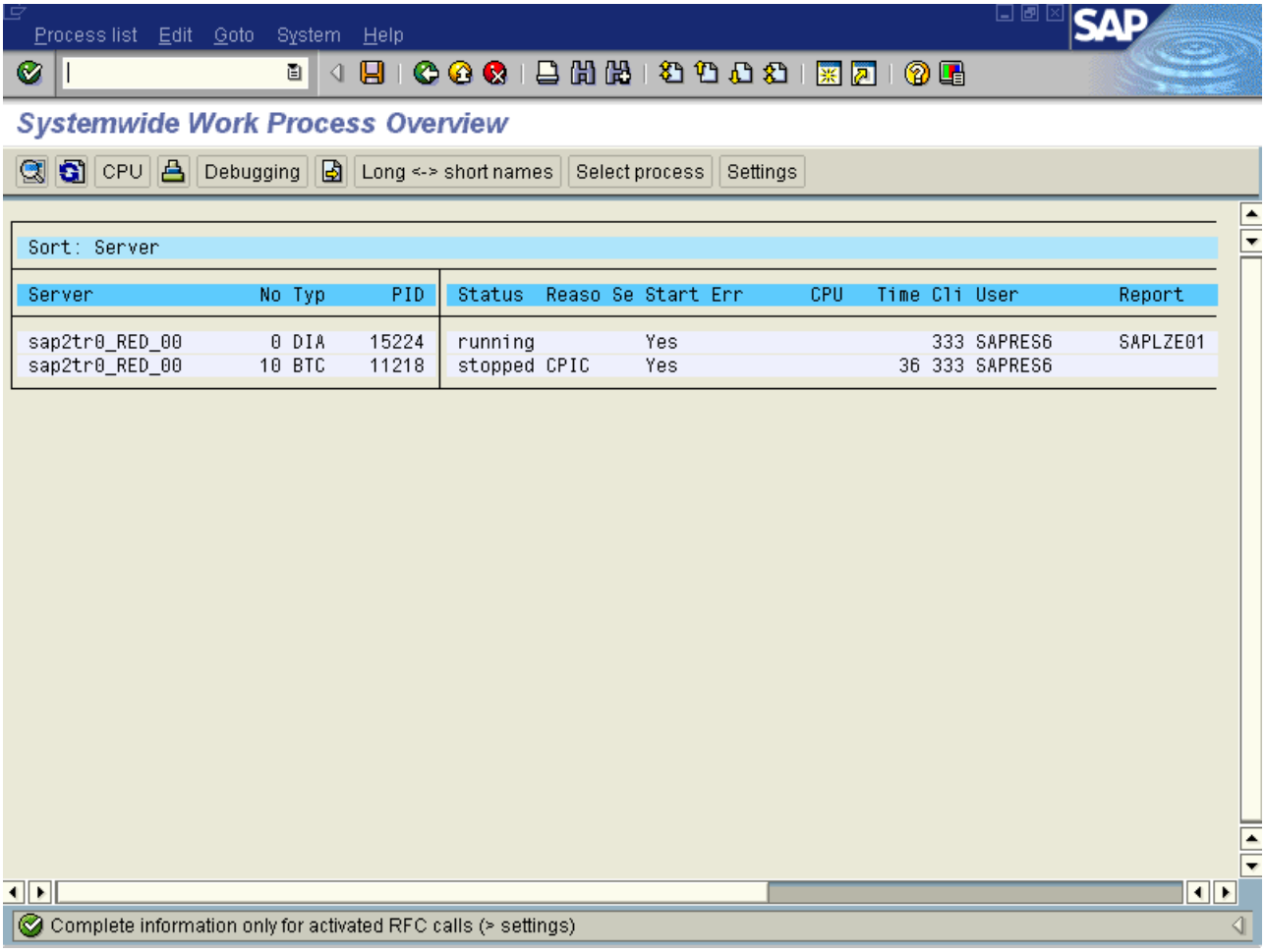


Figure 13. SM66: Online SAP R/3´workload during CF outage

8. We configured the CHPIDs online and printed the message that indicates that the connections are available for use. After the first of the two CHPIDs is online again, the system notified us that the connection to CF07 from SC54 is again established.

```
CF CHP(31),ONLINE
IXL157I PATH 31 IS NOW OPERATIONAL TO CUID: FFFC 174
        COUPLING FACILITY 009672.IBM.02.000000050822
                                    PARTITION: E  CPCID: 00
IEE502I CHP(31),ONLINE
IEE712I CONFIG   PROCESSING COMPLETE
IXC517I SYSTEM SC54 ABLE TO USE 177
        COUPLING FACILITY  009672.IBM.02.000000050822
                           PARTITION: E    CPCID: 00
        NAMED CF07
CF CHP(01),ONLINE
IEE502I CHP(1),ONLINE
IEE712I CONFIG   PROCESSING COMPLETE
```

9. In our case there is no need to activate the primary GBP again because of the automatic reduplexing that occurred. We will continue running with CF06 and CF05.

   However, if you ran this scenario with just two CFs, you would still be in simplex mode right now and should restart duplexing with the SETXCF START command. After that, you should display and print CF information to verify that all structures are correctly spread across both CFs.

   For our case, we will demonstrate that the CHPIDs 01 and 31 are online and therefore the connections to CF07 are available again. CF07 could now be used for GBP duplexing again.

```
D CF,CFNAME=CF07
IXL150I  10.46.08  DISPLAY CF 182
COUPLING FACILITY 009672.IBM.02.000000050822
                  PARTITION: E  CPCID: 00
                  CONTROL UNIT ID: FFFC
NAMED CF07
COUPLING FACILITY SPACE UTILIZATION
 ALLOCATED SPACE                 DUMP SPACE UTILIZATION
  STRUCTURES:          0 K        STRUCTURE DUMP TABLES:        0 K
  DUMP SPACE:       2048 K                  TABLE COUNT:        0
 FREE SPACE:       244992 K       FREE DUMP SPACE:            2048 K
TOTAL SPACE:       247040 K      TOTAL DUMP SPACE:            2048 K
                                MAX REQUESTED DUMP SPACE:        0 K
   VOLATILE:          YES         STORAGE INCREMENT SIZE:      256 K
    CFLEVEL:            9

COUPLING FACILITY SPACE CONFIGURATION
                          IN USE          FREE          TOTAL
CONTROL SPACE:            2048 K      244992 K      247040 K
NON-CONTROL SPACE:           0 K           0 K           0 K

 SENDER PATH       PHYSICAL            LOGICAL       CHANNEL TYPE
        01         ONLINE              ONLINE        ICS
        31         ONLINE              ONLINE        ICS



D XCF,CF,CFNAME=CF07
IXC362I  16.53.23  DISPLAY XCF 149
CFNAME: CF07
  COUPLING FACILITY     :   009672.IBM.02.000000050822
                            PARTITION: E   CPCID: 00
  POLICY DUMP SPACE SIZE:   2048 K
  ACTUAL DUMP SPACE SIZE:   2048 K
  STORAGE INCREMENT SIZE:   256 K

  CONNECTED SYSTEMS:
  SC04     SC42     SC43     SC47     SC48     SC49     SC50
  SC52     SC53     SC54     SC55     SC61     SC62     SC66
  SC67     SC69

  NO STRUCTURES ARE IN USE BY THIS SYSPLEX IN THIS COUPLING FACILITY
```

10. We again verified that the SAP R/3 workload is still running. The screenshot shows that the SAP R/3 workload ran uninterrupted from 10:42:50 to 10:47:52 and completed successfully; see Figure 14.

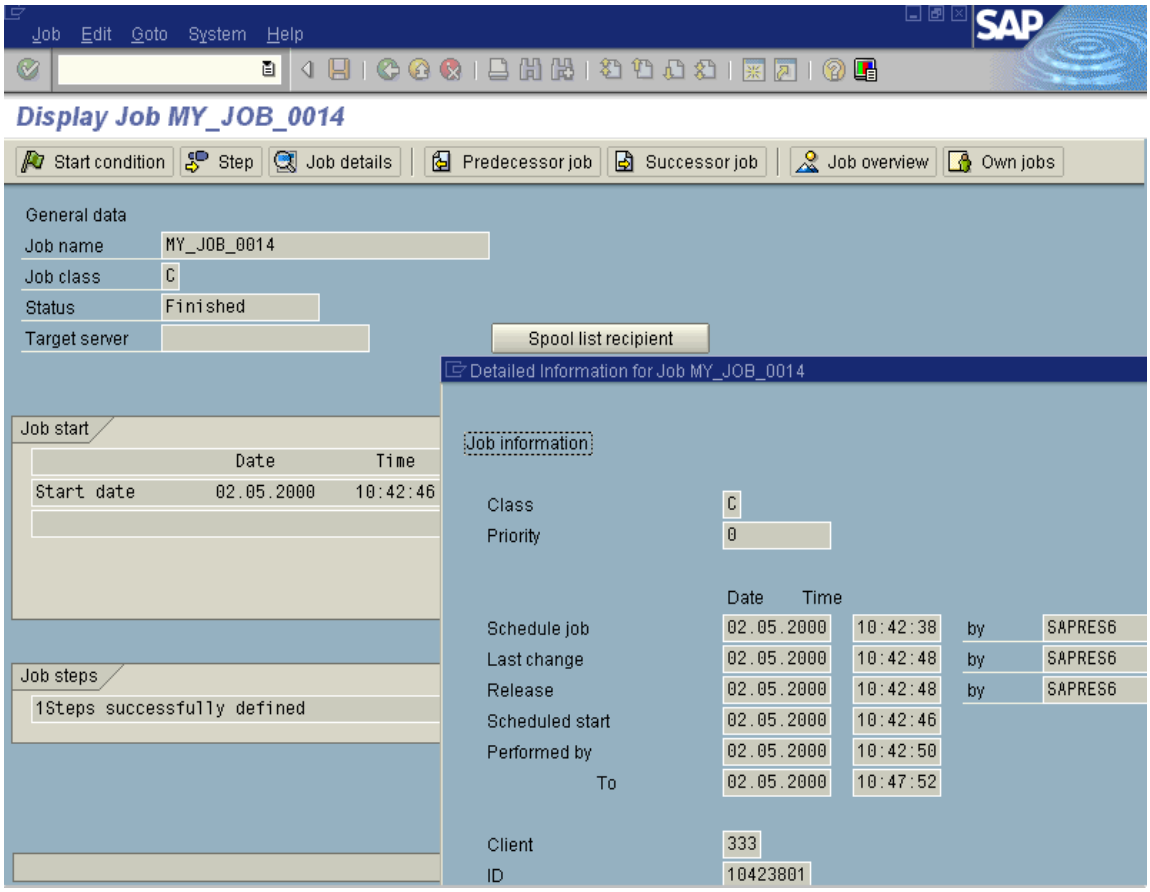


Figure 14. SM37: Successful completion of the SAP R/3 workload

Result

The successful completion of the test proved that an outage of a CF link does not have any impact on running SAP R/3 workload. Compared to the preceding CF scenario, the failure response is the same for both structure failures and for lost connectivity.

2.2.4 SAP R/3 benefits and availability scenarios

1. S/390 Parallel Sysplex and all of the previously mentioned features allow more or less *uninterrupted operation* of SAP R/3. If a DB2 subsystem that is a member of the DB2 data sharing group goes down, access is possible to the SAP R/3 data through those members that continue to work normally. This addresses planned and unplanned downtime.

   The failover mechanism of the ICLI server ensures that SAP R/3 users connected to SAP R/3 application servers will automatically be connected to one of the other DB2 subsystems through the secondary ICLI server. They can access their data from this other DB2 subsystem.

With this concept, neither the S/390 hardware, the OS/390 software nor the DB2 subsystem on one of the Sysplex members are a SPOF anymore. Thanks to S/390 Parallel Sysplex and DB2 data sharing, the following hardware and software failure scenarios won't cause downtime to SAP R/3:

- Outage of a CEC in the Parallel Sysplex
- Coupling Facility Outage
- OS/390 outage on a Sysplex member
- DB2 subsystem outage on a Sysplex member
- S/390 hardware upgrade in the Parallel Sysplex
- Installation of additional Coupling Facility
- OS/390 upgrade on a Sysplex member
- DB2 upgrade on a Sysplex member
- ICLI Failure
- Coupling Facility Link Failure

2. DB2 data sharing enables nearly unlimited growth of the SAP R/3 database server, which otherwise could be the bottleneck for an SAP R/3 installation. The *linear scalability* has been proven by running the SD Parallel benchmark that SAP developed especially for parallel, clustered or distributed databases.

   Furthermore, customer sites where SAP R/3 on S/390 is in productive use in a 2-way Parallel Sysplex have reported coupling efficiencies of more than 90%. In other words, less than 10% data sharing overhead has been seen for global locking and intersystem cache coherency. Thus, nearly the full power of the additional system may be applied to SAP R/3 processing. This vertical and horizontal scalability of S/390 provides for support of multi-TB SAP R/3 databases.

2.3 Summary

The hardware, microcode, and software products associated with a Parallel Sysplex have all been designed with availability and ease of management in mind. The aims of automated detection, analysis, and either resolution or bypass are key focus areas of the Parallel Sysplex solution.

Our tests in our DB2 data sharing environment have shown that certain failures of S/390 Parallel Sysplex components don't impact ongoing SAP R/3 workload, no disruption occurs. Therefore, S/390 Parallel Sysplex clustering is the business solution for helping to ensure that SAP R/3 is available through any downtime event and that revenues from sales and other business opportunities are not lost to the competition.

Chapter 3. Database high availability for SAP R/3

This chapter details our experiences in implementing and using the data sharing feature of DB2 UDB for OS/390. We used an existing installation of SAP R/3 that was implemented with data sharing on a Parallel Sysplex, but where usage was defined for only one DB2 subsystem. What we did was to implement a full data sharing environment.

The essentials regarding the data sharing feature DB2 UDB for OS/390 are described in 3.1, “Data sharing concepts” on page 61.

The changes we made to the DB2 environment are described in 3.2, “Implementing our test environment using data sharing” on page 62.

The changes we made to the SAP R/3 environment are described in 3.3, “Activating database high availability for SAP R/3” on page 63.

After implementing, we performed extensive testing that showed we could stop one member to add a PTF while the other member performed the database server functions. We also include in this chapter our procedure for returning to the original configuration, that is, having both DB2 instances perform database server functions.

3.1 Data sharing concepts

The data sharing function of DB2 UDB for OS/390 enables applications that run on more than one DB2 subsystem to read from and write to the same set of data concurrently.

DB2 subsystems that share data must belong to a DB2 data sharing group that runs on a Parallel Sysplex. A *data sharing group* is a collection of one or more DB2 subsystems that access shared DB2 data. A Parallel Sysplex is a collection of S/390 systems that communicate and cooperate with one each other.

Each DB2 subsystem that belongs to a particular data sharing group is a member of that group. All members of a data sharing group use the same shared DB2 catalog and directory.

Data sharing concepts for DB2 UDB for OS/390 are explained in more detail in the following publications:

DB2 UDB for OS/390 Data Sharing: Planning and Administration Version 6, SC26-9007

High Availability Considerations: SAP R/3 on DB2 for OS/390, SG24-2003

Database Administration Experiences: SAP R/3 on DB2 for OS/390, SG24-2078

If you run SAP R/3 with multiple DB2 subsystems in different LPARs in a Parallel Sysplex with data sharing, SAP R/3 will profit from improved availability, incremental growth and especially workload balancing.

3.2 Implementing our test environment using data sharing

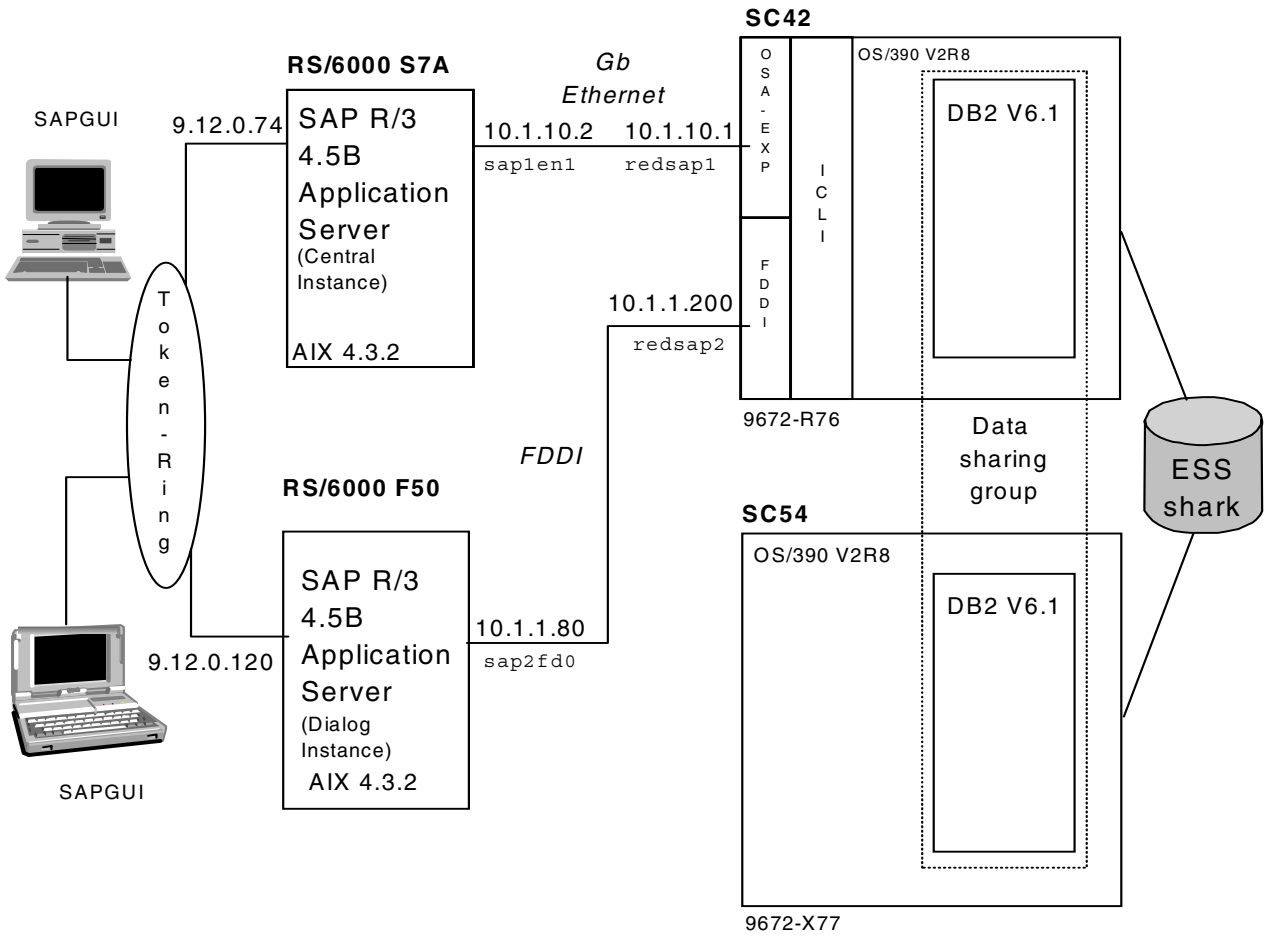


Figure 15. View of our data sharing configuration for SAP R/3

Figure 15 shows our implemented configuration.On a Parallel Sysplex with data sharing we implement four DB2 subsystems called DBK1, DBK2, DBK3 and DBK4. DBK1 and DBK3 are on LPAR SC42. DBK2 and DBK4 are on LPAR SC54.

The ICLI servers on S/390 are OS/390 started tasks.

The central instance (CI) communicates via primary connection ICLIRED to DBK1 using port 6531, and via a standby connection ICLIRED3 to DBK4 using port 6533.

The application server on F50 communicates via primary connection ICLIRED2 to DBK2 using port 6531, and via a standby connection ICLIRED1 to DBK3 using port 6533.

In our environment, as described in Chapter 1, “Introduction” on page 1 we refer to option 2: failover into a standby DB2 member. We chose this options because of its possibilities as described in 1.4.4, “Two active DB2s, inactive standby in opposite LPAR - option 2” on page 10.

However, the connection between the application server sap2tr0 and the second OS/390 SC54 system in our environment is made by a token ring network. To provide the speed you need in a production environment, the connection should be a faster one, like Gb Ethernet.

3.3 Activating database high availability for SAP R/3

To change from a data sharing non-high availability database system for SAP R/3 to a high availability system, we applied changes as described in *SAP R/3 on DB2 for OS/390: Planning Guide SAP R/3 Release 4.6B,* SC33-7966.

Some of the most important changes we applied are specified in the following figures. Figure 16 on page 64 shows the relevant portion of OSS Note 98051 that we consulted. This note describes the process SAP R/3 uses to implement failover (the note describes the situation as of the time of writing; always remember to refer to the latest version for current information).

The process is called *database reconnect*, and is described in more detail in Figure 16. It is important to understand that such a failover impacts both online and batch processing

Figure 17 on page 65 shows important settings for an application server. Figure 18 on page 65 shows the most important of our settings for the central instance.

It is essential to verify that the ports you specify are not used by any other application.

The port number for the primary connection must be specified in the /etc/services file using the name sapdb2<sid>.

In a sysplex environment, it is important to set the following SAP R/3 instance parameters:

- To enable failover and failback, set `rsdb/reco_symmetric` to `ON`.
- To prevent failing over all other application servers to their standby database, set `rsdb/reco_sync_all_server` to `OFF`.

For more details related to enabling Sysplex failover support refer to the corresponding chapter in *SAP R/3 on DB2 for OS/390: Planning Guide SAP R/3 Release 4.6B,* SC33-7966.

```
 a) The reconnect scenario is triggered by the breakdown of the
    database when the R/3 System is in operation. Those work
    processes, which already started an R/3 transaction before the
    database breakdown, receive a database (SQL) error.

b) To keep the R/3 System in a consistent status, a rollback is
    carried out within R/3 and on in the database in which the
    cancelled transaction is reset. This is also mostly documented in
    a short dump. If the work process finds the return code in the
    reconnect class, a message of the following structure is first
    stored in the developer trace dev_w*:
    B ***LOG BYM=> severe db error 12571    work process is in the
    reconnect state [dbsh     0649]
    A similar message (with BYM) is then entered in the SYSLOG.

c) After the R/3 rollback, the work process attempts several times to
    set up the database connection again. The profile parameters
    rsdb/reco_trials and rsdb/reco_sleep_time determine here the
    number of attempts and the pauses in between. In the case of
    short-term problems, for example when the network or database
    capacity is overloaded, the work process will be able to set up
    the connection again when the appropriate settings are made.

d) If the connection setup fails, the user involved receives the
    following dialog box: "Connection to database lost, session
    terminated". The user session is terminated with this message.
    Note: the work process itself is not exited, but is ready to
    receive the next user request.

e) In addition, all work processes of the application server are
    informed about the database problem. The WPs informed write the
    following messages in their developer traces:
    B  ***LOG BV4=> reconnect state is set for the work process

f) From there on, all work processes are in reconnect status.
    Whenever they are assigned a new user request in this status, they
    attempt to set up the database connection again before they start
    processing the request. After the attempts determined in
    rsdb/reco_trials, the user receives the dialog box "Connection to
    database lost, session terminated" and the user's session is
    terminated. The WP however remains in the reconnect status and
    waits for the next request...
    In the developer trace, the futile reconnect attempts have the
                                                              Page 3

    following effect:
    M  *** ERROR => ThHdlReconnect: db_reconnect failed (-1) [thxxhead
    1304]

g) If the database(connection) is available again, the reconnect
    attempt is also successful and the above-mentioned loop is broken.
    This is documented in the developer trace:
    B  ***LOG BYY=> work process left reconnect state [dblnk    1467]
    An equivalent message also appears in the SYSLOG.

h) A work process can only restore a database connection and write a
    SYSLOG or developer trace message when it executes a request. If
    the system load is very low and a work process edits a user
    request hours after the database loss, the reconnect of this work
      process and the respective message also only occur at this later
```

Figure 16. Part of OSS note 98051

```
SAPSYSTEMNAME                           RED
INSTANCE_NAME                           D00
SAPSYSTEM                               00
SAPDBHOST                               WTSC54OE
dbs/db2/hosttcp                         WTSC54OE
dbs/db2/ssid_standby                    DBK3
dbs/db2/ssid                            DBK2
dbs/db2/hosttcp_standby                 redsap2
dbms/type                               db2
rsdb/reco_tcp_service
rsdb/reco_ping_cmd
rsdb/reco_sync_all_server               OFF
rsdb/reco_symmetric                     ON
rsdb/db2_host_standby                   REDSAP2
rsdb/reco_sleep_time                    1
rsdb/reco_trials                        1
rsdb/db2_port_standby                   6533
```

Figure 17. Important configuration parameters for an application server on F50

```
SAPSYSTEMNAME          RED
INSTANCE_NAME         DVEBMGS00
SAPSYSTEM              00
SAPDBHOST                                redsap1
dbs/db2/hosttcp                          redsap1
dbs/db2/ssid_standby                     DBK4
dbs/db2/ssid                             DBK1
dbs/db2_port                             6531
dbs/db2/hosttcp_standby                  wtsc54oe
dbms/type                                db2
rsdb/reco_tcp_service
rsdb/reco_ping_cmd
rsdb/reco_sync_all_server                OFF
rsdb/reco_symmetric                      ON
rsdb/db2_host_standby                    wtsc54oe
rsdb/reco_sleep_time                     1
rsdb/reco_trials                         1
rsdb/db2_port_standby                    6533
```

Figure 18. Important configuration parameters for the Central Instance on S7A

3.4 Testing database high availability

In the following section we describe what we tested regarding DB2 UDB for OS/390 and SAP R/3. We specify the method, actions, and results for each test.

During the test we ran a batch job on one or both servers, and used SAP R/3 transactions to simulate user activity.

In some cases we applied additional load using RFC-based batch.

3.4.1 ICLI failure impact on the application server

Purpose

Demonstrate the ability to fail over an SAP R/3 application server and continuing transactions when the active ICLI started task is going down.

Method

Cancel the primary ICLI started task**.**

Criteria

The successful completion of the tasks described below will demonstrate that the SAP server will automatically reconnect to the database via the failover ICLI server. Active or logged-on users will *not* be required to log off the server and will be able to continue running transactions after the reconnect has occurred. However, any tasks running during the failover will abend with an sqlcode 0 return code.

```
Subsystem  DBK2   at   14:39:28   24.04.2000     Data since   DB start
Last reset                                       DB system    DB2
Last refresh                                     DB release   6.1.0
```

Figure 19. DB2 Subsystem activity before failing over

Tasks

1. Start all ICLI started tasks.
2. Create a batch workload at the application server.
3. Start the online transactions.
4. Cancel the ICLI started task ICLIRED2 (see Figure 20).
5. Restart the ICLI started task ICLIRED2.
6. Cancel the ICLI started task ICLIRED1.

```
Display  Filter  View  Print  Options  Help
-------------------------------------------------------------------------------
SDSF STATUS DISPLAY ALL CLASSES                        LINE 1-4 (4)
COMMAND INPUT ===> /ro sc54,p iclired2                        SCROLL ===> CSR
PREFIX=ICLIRED*  DEST=(ALL)  OWNER=*
NP   JOBNAME  JOBID    OWNER    PRTY QUEUE      C  POS  SAFF  ASYS STATUS
     ICLIRED1 STC14798 ICLIRUN1   15 EXECUTION          SC42  SC42
     ICLIRED  STC14799 ICLIRUN    15 EXECUTION          SC42  SC42
     ICLIRED3 STC14800 ICLIRUN3   15 EXECUTION          SC54  SC54
     ICLIRED2 STC14801 ICLIRUN2   15 EXECUTION          SC54  SC54
```

Figure 20. Cancel ICLIRED2 started task

Results

1. The application server fails over from DB2 subsystem DBK2 to DBK3 (see Figure 21 on page 67).
2. The online SAP R/3 transactions did not notice any impact.
3. The batch job was canceled. Manual action will be needed to restart it.

4. A short dump is generated by SAP R/3 because of the abnormal end of batch job (see Figure 22 on page 67).
5. There are no lost update records.
6. There are no remaining lock entries.

After the fallback, we see the following results:

7. The batch job was canceled. Manual action will be needed to restart it.
8. No short dump is generated by SAP R/3.
9. Database changes are not rolled back, because the updates are committed during the run. In this case our run has restart capability. However, for each abandoned batch job, application logic should be taken in account if restart is possible or alternative action is required.
10. There are no lost update records.
11. There are no remaining lock entries.

```
Subsystem  DBK3   at   15:17:48   24.04.2000     Data since   DB start
Last reset                                       DB system    DB2
Last refresh                                     DB release   6.1.0
```

Figure 21. DB2 Subsystem activity after failing over

```
24.04.2000 15:17:20 sap2tr0 SAP* 123 C DBIF_RSQL_SQL_ERROR


Database error text.........: "ICLC2511E Could not receive data from server -
 Connection reset by peer (errno=73); ICLC2507E Receiving data failed in
 transmission phase; ; Details: normal transmission mode ;"
Internal call code..........: "[RSQL/FTCH/TADIR ]"
Please check the entries in the system log (Transaction SM21).
```

Figure 22. Short dump from batch job

```
15:06:05 BTC 10 123 SAP*              BYM SQL error 0 (possibly a network error);
                     WP in reconnect status
15:06:05 BTC 10 123 SAP*              R68 Perform rollback
15:06:05 BTC 10 123 SAP*              BYY Work process has left reconnect status
15:06:09 BTC 10 123 SAP*              AB0 Run-time error "DBIF_RSQL_SQL_ERROR" occurred
15:06:12 DIA  0                       BV4 Work process is in reconnect status
15:06:12 DIA  0                       BYY Work process has left reconnect status
15:06:14 BTC 10 123 SAP*              AB1 > Short dump "000424 151720 sap2tr0 SAP* " generated
15:06:14 BTC 10 123 SAP*              D01 Transaction termination 00
                     (DBIF_RSQL_SQL_ERROR 20000424151720sap2tr0 SAP* 1231 )
15:06:14 BTC 10 123 SAP*              R68 Perform rollback
15:06:33 DIA  1                       BV4 Work process is in reconnect status
15:06:33 DIA  1                       BYY Work process has left reconnect status
15:06:35 DIA  2                       BV4 Work process is in reconnect status
15:06:36 DIA  2                       BYY Work process has left reconnect status
```

Figure 23. Systemlog after failover

3.4.2 ICLI failure impact on the central instance

The behavior of the central instance during an ICLI failure is the same as the behavior of an application server. However, because some services are provided only by the central instance that is failing over, jobs running on application servers other than the central instance are also impacted. Batch jobs are also canceled on those application servers and active dialog processes are interrupted as shown in Figure 24.

```
Clnt  User       Date        Time      TCode             Information      Status

333   SAPRES1    25.04.2000  13:30:29  SESSION_MANAGER                    Err
333   SAPRES1    25.04.2000  13:30:29  SESSION_MANAGER                    Err
```

Figure 24. SAP cancelled batch jobs after failover

During the restart of the interrupted batch jobs there will be an SAP express message, which informs the user that his job has failed to complete.

3.4.3 Adapter failure impact on the central instance on AIX

Purpose
Show what happens when the network connection on the central instance fails.

Method
Shut down the network connection with the `ifconfig <interface> down` command at the central instance server.

Tasks
1. Create a batch workload at the central instance.
2. Use dialog processes on central instance and application server.
3. Shut down the network interface using the `ifconfig <interface> down` command.

Results
You can find the full tracefile of the work process, which is generated during the above tasks, in Appendix C, “Tracefile of the work processes after adapter failure” on page 135. In Appendix C, referencing Figure 48 on page 143, the following notes apply:

As shown in [1], the ICLI client is not able to receive data from the server. This error message appears several times in the trace file.

As shown at [2], the work process is set to the reconnect state.

As shown at [3], the transaction is rolled back.

As shown at [4], the work process is performing a reconnect.

As shown at [5] to [6], the actual failover to the standby database takes place.

As shown at [7] to [8], all other work processes are reconnected to the standby database, and normal SAP R/3 operation is possible again.

Note: After we stopped the Gigabit Ethernet adapter, the elapsed time between adapter failure and failover to the standby database was approximately 8 minutes. We are not sure what caused this delay. It seems to be related to the normal TCP/IP time-out when an adapter or network failure takes place.We will open an OSS request to SAP to investigate this issue.

Other information

1. Batch jobs failed and were abandoned (see Figure 25).
2. SAP R/3 detected a failure of the network connection. This is shown in the system log (see Figure 26 on page 70).
3. Work processes are left in reconnect status (see Figure 24 on page 68) and restarted after the failover.

```
Date       Time     Host         User        Clt X Error ID

04.05.2000 13:51:36 sap1tr0      SAP*       456 C DBIF_RSQL_SQL_ERROR
04.05.2000 13:51:38 sap1tr0      SAPRES1     333 C DBIF_REPO_SQL_ERROR
04.05.2000 13:51:41 sap2tr0       WIMM        666 C TIME_OUT
04.05.2000 13:51:51 sap1tr0      SAPSYS      000 C DBIF_RTAB_SQL_ERROR
04.05.2000 13:52:14 sap1tr0      SAP*       456 C ZDATE_ILLEGAL_LOCTIME
```

Figure 25. Dumps after the network failure

```
Time     Ty. Nr Cl. User         Tcod MNo Text
Date : 04.05.

13:42:53 S-A    000 redadm            E07 Error 00006 : 9Network is down in Module
rslgsend(042)
13:48:12 DIA  1 456 SAP*         SCC3 BYM SQL error 0 (possibly a network error); WP in
reconnect status
13:48:12 DIA  1 456 SAP*         SCC3 R68 Perform rollback
13:48:13 DIA  1 456 SAP*         SCC3 BYY Work process has left reconnect status
13:48:13 DIA  1 456 SAP*         SCC3 AB0 Run-time error "DBIF_RSQL_SQL_ERROR" occurred
13:48:14 DIA  0 333 SAPRES1      ST04 R68 Perform rollback
13:48:14 DIA  0 333 SAPRES1      ST04 BYM SQL error 0 (possibly a network error); WP in
reconnect status
13:48:14 DIA  0 333 SAPRES1      ST04 D01 Transaction termination 00 (0 )
13:48:14 DIA  0 333 SAPRES1      ST04 AB2 > Include ??? line 0000.
13:48:14 DIA  0 333 SAPRES1      ST04 AB3 Error 00 951 in error recovery
13:48:14 DIA  0 333 SAPRES1      ST04 AB0 Run-time error "TIME_OUT" occurred
13:48:14 DIA  0 333 SAPRES1      ST04 AB2 > Include ??? line 0000.
13:48:14 DIA  0 333 SAPRES1      ST04 AB0 Run-time error "DBIF_DSQL2_SQL_ERROR" occurred
13:48:14 DIA  0 333 SAPRES1      ST04 AB1 > Short dump "000504 135137 sap1tr0 SAPRES1 "
generated
13:48:14 DIA  0 333 SAPRES1      ST04 R68 Perform rollback
13:48:14 DIA  0 333 SAPRES1      ST04 BYY Work process has left reconnect status
13:48:15 DIA  0 333 SAPRES1      ST04 AB0 Run-time error "DBIF_REPO_SQL_ERROR" occurred
13:48:15 DIA  0 333 SAPRES1      ST04 AB1 > Short dump "000504 135138 sap1tr0 SAPRES1 "
generated
13:48:15 DIA  0 333 SAPRES1      ST04 D01 Transaction termination SY (SQL error 0 occurred
when accessing program SAPMS380 )
13:48:15 DIA  0 333 SAPRES1      ST04 R68 Perform rollback
13:48:17 DIA  1 456 SAP*         SCC3 AB1 > Short dump "000504 135136 sap1tr0 SAP* "
generated
13:48:20 DIA  1                       BV4 Work process is in reconnect status
13:48:20 DIA  1                       BYY Work process has left reconnect status
13:48:21 ENQ  7                       BV4 Work process is in reconnect status
13:48:28 DIA  2 000 SAPSYS            BYM SQL error 0 (possibly a network error); WP in
reconnect status
13:48:28 DIA  3 000 SAPSYS            BYM SQL error 0 (possibly a network error); WP in
reconnect status
13:48:28 DIA  2 000 SAPSYS            R38 Error at DB commit, return code 001024
13:48:28 DIA  3 000 SAPSYS            R68 Perform rollback
13:48:28 DIA  2 000 SAPSYS            R68 Perform rollback
13:48:28 DIA  3 000 SAPSYS            BYY Work process has left reconnect status
13:48:28 DIA  3 000 SAPSYS            AB0 Run-time error "DBIF_RTAB_SQL_ERROR" occurred
13:48:29 DIA  3 000 SAPSYS            AB1 > Short dump "000504 135151 sap1tr0 SAPSYS "
generated
13:48:29 DIA  3 000 SAPSYS            D01 Transaction termination 00 (DBIF_RTAB_SQL_ERROR
20000504135151sap1tr0 SAPSYS 0001 )
13:48:29 DIA  3 000 SAPSYS            R68 Perform rollback
13:48:29 DIA  2 000 SAPSYS            BYY Work process has left reconnect status
13:48:51 DIA  4 456 SAP*         ST04 BYM SQL error 0 (possibly a network error); WP in
reconnect status
13:48:51 DIA  4 456 SAP*         ST04 R68 Perform rollback
13:48:51 DIA  4 456 SAP*         ST04 BYY Work process has left reconnect status
13:48:51 DIA  4 456 SAP*         ST04 AB0 Run-time error "ZDATE_ILLEGAL_LOCTIME" occurred
13:48:52 DIA  0                       BV4 Work process is in reconnect status
13:48:52 DIA  0                       BYY Work process has left reconnect status
13:48:52 DIA  4 456 SAP*         ST04 AB1 > Short dump "000504 135214 sap1tr0 SAP* "
generated
```

Figure 26. Syslog taken during the interface downtime

3.4.4 Line failure impact on the central instance on AIX

Purpose

Show what happens when the network connection between the central instance and the database server is disconnected, for whatever reason.

Method

We unplug the interface cable from the network adapter to shut down the network connection.

Tasks

1. Create batch workload at the central instance.
2. Use dialog processes on the central instance and the application server.
3. Unplug the interface cable.
4. Plug in the interface cable again after diagnostics.

Results

We disconnected the cable at 9:10. As you can see in Figure 24 on page 68, the trace shows the work process tracefile contains errors that are written starting at 9:17. User interaction was possible until that process needed access to the database. From that moment on, the work process is in reconnect status.

At 9:18, the ICLI client decided it could not connect to the primary ICLI server (see Figure 27 on page 72) and instead connects to the standby host (see Figure 28 on page 73). After that, all processes are reconnected to the database and the SAP R/3 system is available again.

Work on the application server can continue until commit to the database is required, then the process hangs.

Batch jobs failed and were abandoned (see Figure 30 on page 75).

```
C  Thu Apr 27 09:17:20 2000
C  *** ERROR =>
[00004594] ICLC2511E Could not receive data from server - Connection timed out (errno=78)
           Time: 956841440.547981 = (Thu) 2000 Apr 27 09:17:20.547981
 [dbsldb2. 2141]
C  *** ERROR =>
[00004594] ICLC2507E Receiving data failed in transmission phase

           Details: normal transmission mode
           Time: 956841440.548159 = (Thu) 2000 Apr 27 09:17:20.548159
 [dbsldb2. 2141]
C  *** ERROR => ERROR in sldb2.c line 5798 : ICLI ended with returncode -5
 [dbsldb2. 2141]
C  *** ERROR => ICLC2511E Could not receive data from server - Connection timed out
(errno=78)
 [dbsldb2. 2141]
C  *** ERROR => ICLC2507E Receiving data failed in transmission phase
 [dbsldb2. 2141]
C  *** ERROR =>
 [dbsldb2. 2141]
C  *** ERROR =>            Details: normal transmission mode
 [dbsldb2. 2141]
B  db_con_set_reco_flag: 0: name = R/3, handle = 0
B  Reconnect state is entered by connection:
B  0: name = R/3, concnt= 000000 state = ACTIVE      , perm = YES, reco = YES
B  ***LOG BYM=> severe db error 0          work process is in the reconnect state [dbsh
0580]
B  ***LOG BY4=> sql error 0      performing SEL on table DDLOG      [dbsyntsp 0430]
B  ***LOG BY0=> ICLC2511E Could not receive data from server - Connection timed out
(errno=78); ICLC2507E Receiving data failed in transmissio
B  ***LOG BZY=> unexpected return code 2          on DbSlExeRea-call [dbsync   2503]
B  collect_new : unexpected returncode 512
B  db_con_set_reco_flag: 0: name = R/3, handle = 0
B  Reconnect state is entered by connection:
B  0: name = R/3, concnt= 000000 state = ACTIVE      , perm = YES, reco = YES
B  ***LOG BYM=> severe db error 0          work process is in the reconnect state [dbsh
0580]
```

Figure 27. Extract from the tracefile of the work process - part 1

```
B  hdl_error_on_commit_rollback: Connection details
B  0: name = R/3, concnt= 000000 state = ACTIVE      , perm = YES, reco = YE
B  ***LOG BY2=> sql error 0      performing COM [dbcon    2283]
B  ***LOG BY0=>  [dbcon    2283]
B  ***LOG BYJ=> Database COMMIT     for connection R/3        failed [dbcon    2295
M  ***LOG R38=> ThICommit3, db_commit98 ( 001024) [thxxhead 9237]
M  *** ERROR => ThEndBg: ThICommit3 failed [thxxhead 1072]
M  ***LOG R68=> ThIRollBack, roll back () [thxxhead 9404]
S  *** ERROR => task rolled out -> no rollback(b) for spool possible [rspoooh. 0944
B  db_con_set_reco_flag: 0: name = R/3, handle = (
B  Reconnect state is entered by connection:
B  0: name = R/3, concnt= 000000 state = ACTIVE      , perm = YES, reco = YE
B  ***LOG BYM=> severe db error 0          work process is in the reconnect state [dbsh     0580
B  hdl_error_on_commit_rollback: DB-ROLLBACK ok. RECONNECT stat
B  DB_RECONNECT: rsdb/reco_ping_cmd is not set or too lon
B  DB_RECONNECT: rsdb/reco_tcp_service not set or too lon
B  DB_RECONNECT: Currently active DB-Instance is on primary hos
B  Disconnected from connection 0
B  find_con found the following connection for reuse
B  0: name = R/3, concnt= 000000 state = DISCONNECTED, perm = YES, reco = YE
C  ********* Profile Parameter for connection 0 *********
C  DBIF Versions:
C    R/3 Version       : 12 (4.6B)
C    API Version       : 9
C  ICLI-Client Versions:
C    R/3 Version       : 12 (4.6B)
C    API Version       : 9
C    PROTOCOL Version  : 1
C    Internal Version  : 0
C  Workprocess         : 0
C  USE_NO_HOLD         : 0
C  USE_OPT_1_ROW       : 0
C  DB_UID              : SAPR3
C  DB_ID               : REL
C  SAPSYSTEM           : 00
C  DB_SSID             : DBK1
C  DBHOST              : redsap1
C  PORT                : see /etc/services entry sapdb2REL
C  CCSID               : 819
C  KEEP_X_LOCK         : 1
C  RO_COMMIT_SUPPRESS  : 0
C
C  Used Bugfixes:
C  ************************************
C  Network-Protocol    : TCP
C
C  Thu Apr 27 09:18:06 2000
C  *** ERROR =>
[00004594] ICLC2514E Could not connect to ICLI server - Interrupted system call (errno=4
           Used ICLI connection port 6531 and communications protocol TCP/I
           Time: 956841486.238439 = (Thu) 2000 Apr 27 09:18:06.23843
 [dbsldb2. 2141]
```

Figure 28. Extract from the tracefile of the work process - part 2

```
B  DB_RECONNECT: connect test to current instance not successful
B  Disconnected from connection 0
B  db_con_set_reco_flag: 0: name = R/3, handle = 0
B  Reconnect state is entered by connection:
B  0: name = R/3, concnt= 000000 state = DISCONNECTED, perm = YES, reco = YES
B  DB_RECONNECT: current instance is declared as dead
B  db_con_reconnect performing the reconnect for con:
B  0: name = R/3, concnt= 000000 state = DISCONNECTED, perm = YES, reco = YES
C  ********* Profile Parameter for connection 0 **********
C  Connecting to standby-Host
C  DBIF Versions:
C    R/3 Version       : 12 (4.6B)
C    API Version       : 9
C  ICLI-Client Versions:
C    R/3 Version       : 12 (4.6B)
C    API Version       : 9
C    PROTOCOL Version  : 1
C    Internal Version  : 0
C  Workprocess         : 0
C  USE_NO_HOLD         : 0
C  USE_OPT_1_ROW       : 0
C  DB_UID              : SAPR3
C  DB_ID               : RED
C  SAPSYSTEM           : 00
C  Definitions for standby host
C  DB_SSID             : DBK4
C  DBHOST              : wtsc54oe
C  PORT                : 6533
CCSID              : 819
C  KEEP_X_LOCK         : 1
C  RO_COMMIT_SUPPRESS  : 0
C
C  Used Bugfixes:
C  ************************************
C  Network-Protocol    : TCP
C
C  Thu Apr 27 09:18:07 2000
C  *** ERROR =>
[00004594] ICLC2520W Security is switched off - it should be switched off on the server too!
           Time: 956841487.198046 = (Thu) 2000 Apr 27 09:18:07.198046
 [dbsldb2. 2141]
C
C  Thu Apr 27 09:18:09 2000
C  Your DB2 system works with ASCII code page 819 and EBCDIC code page 37
B  Connection 0 opened
B  Wp  Hdl ConName              ConCnt ConState     TX  PRM RCT Date     Time
B  000 000 R/3                  000000 ACTIVE       NO  YES NO  20000427 091806
B  001 000 R/3                  000000 ACTIVE       YES YES NO  20000427 091806
B  002 000 R/3                  000000 CONNECTING   NO  YES YES 20000427 091806
B  003 000 R/3                  000000 CONNECTING   NO  YES YES 20000427 091747
B  004 000 R/3                  000000 ACTIVE       NO  YES NO  20000426 200544
B  005 000 R/3                  000000 INACTIVE     NO  YES NO  20000426 200544
B  006 000 R/3                  000000 INACTIVE     NO  YES NO  20000426 200544
B  007 000 R/3                  000000 INACTIVE     NO  YES NO  20000426 200544
B  008 000 R/3                  000000 CONNECTING   NO  YES YES 20000427 091711
B  009 000 R/3                  000000 INACTIVE     NO  YES NO  20000426 200544
B  010 000 R/3                  000000 ACTIVE       NO  YES NO  20000426 200544
B  011 000 R/3                  000000 INACTIVE     NO  YES NO  20000426 200544
B  012 000 R/3                  000000 INACTIVE     NO  YES NO  20000426 200544
B  Successful reconnect for connection:
B  0: name = R/3, concnt= 000000 state = ACTIVE      , perm = YES, reco = NO
B  ***LOG BYY=> work process left reconnect state [dblink   1671]
S  *** ERROR => task rolled out -> no rollback(a) for spool possible [rspoooh. 0952]
M  *** ERROR => PfStatWrite: missing STAT_END for opcode STAT_COMMIT [pfxxstat 2912]
M
M  Thu Apr 27 09:18:11 2000
M  ***LOG R49=> ThCPICSyncWait, CPIC-Error (002679) [thxxcpic 3082]
M  ***LOG R64=> ThCPICSyncWait, CPIC-Error ( CMINIT(SAP)) [thxxcpic 3095]
```

Figure 29. Extract from the tracefile of the work process - part 3

```
27.04.2000 09:21:09 sap1tr0       WIMM         333 C TIME_OUT
27.04.2000 09:21:15 sap1tr0       SAPSYS       000 C DBIF_RSQL_SQL_ERROR
27.04.2000 09:21:17 sap1tr0       WIMM         333 C TIME_OUT
27.04.2000 09:21:31 sap1tr0       SAP*         123 C DBIF_RSQL_SQL_ERROR
27.04.2000 09:21:36 sap1tr0       SAP*         123 C TIME_OUT
27.04.2000 09:21:38 sap1tr0       SAP*         123 C TIME_OUT
27.04.2000 09:21:57 sap1tr0                    000 C DBIF_RTAB_SQL_ERROR
27.04.2000 09:21:57 sap1tr0       SAPSYS       000 C TIME_OUT
27.04.2000 09:24:23 sap1tr0       SAPSYS       000 C DBIF_RSQL_SQL_ERROR
```

Figure 30. Dumps after the line failure

```
09:17:11 BTC  8 123 SAP*              BYM SQL error 0 (possibly a network error); WP in reconnect
status
09:17:11 BTC  8 123 SAP*              R68 Perform rollback
09:17:11 BTC  8 123 SAP*              F6H Database error: TemSe->XRTAB(3)->1 for table TST01 key
09:17:20 DIA  0 000 SAPSYS            BYM SQL error 0 (possibly a network error); WP in reconnect
status
09:17:20 DIA  1 333 WIMM         SM50 BYM SQL error 0 (possibly a network error); WP in reconnect
status
09:17:20 DIA  2 000 SAPSYS            BYM SQL error 0 (possibly a network error); WP in reconnect
status
09:17:20 DIA  0 000 SAPSYS            R38 Error at DB commit, return code 001024
09:17:20 DIA  0 000 SAPSYS            R68 Perform rollback
09:17:20 DIA  2 000 SAPSYS            R68 Perform rollback
09:17:20 DIA  1 333 WIMM         SM50 R68 Perform rollback
09:17:47 DIA  3 123 SAP*         SCC3 BYM SQL error 0 (possibly a network error); WP in reconnect
status
09:17:47 DIA  3 123 SAP*         SCC3 R68 Perform rollback
09:17:48 ENQ  7                       BV4 Work process is in reconnect status
09:18:07 DIA  1 333 WIMM         SM50 BYY Work process has left reconnect status
09:18:07 DIA  1 333 WIMM         SM50 AB0 Run-time error "TIME_OUT" occurred
09:18:09 DIA  0 000 SAPSYS            BYY Work process has left reconnect status
09:18:09 DIA  2 000 SAPSYS            BYY Work process has left reconnect status
09:18:11 DIA  0 000 SAPSYS       SESS R49 Communication error, CPIC return code 002, SAP return code
679
09:18:11 DIA  0 000 SAPSYS       SESS R64 > CPI-C function: CMINIT(SAP)
09:18:12 DIA  2 000 SAPSYS            AB0 Run-time error "DBIF_RSQL_SQL_ERROR" occurred
09:18:13 DIA  2 000 SAPSYS            AB1 > Short dump "000427 092115 sap1tr0 SAPSYS " generated
09:18:13 DIA  2 000 SAPSYS            D01 Transaction termination 00 (DBIF_RSQL_SQL_ERROR
20000427092115sap1tr0 SAPSYS 0001 )
09:18:13 DIA  2 000 SAPSYS            R68 Perform rollback
09:18:13 DIA  1 333 WIMM         SM50 AB1 > Short dump "000427 092109 sap1tr0 WIMM " generated
09:18:14 DIA  1 333 WIMM         SM50 R68 Perform rollback
09:18:14 DIA  1 333 WIMM         SM50 AB0 Run-time error "TIME_OUT" occurred
09:18:14 DIA  1 333 WIMM         SM50 AB1 > Short dump "000427 092117 sap1tr0 WIMM " generated
09:18:15 DIA  1 333 WIMM         SM50 D01 Transaction termination SY (Time limit exceeded. )
09:18:15 DIA  1 333 WIMM         SM50 R68 Perform rollback
09:18:28 BTC  8 123 SAP*              BYY Work process has left reconnect status
09:18:29 BTC  8 123 SAP*              AB0 Run-time error "DBIF_RSQL_SQL_ERROR" occurred
09:18:31 BTC  8 123 SAP*              AB1 > Short dump "000427 092131 sap1tr0 SAP* " generated
09:18:31 BTC  8 123 SAP*              D01 Transaction termination 00 (DBIF_RSQL_SQL_ERROR
20000427092131sap1tr0 SAP* 1231 )
09:18:31 BTC  8 123 SAP*              R68 Perform rollback
09:18:34 DIA  3 123 SAP*         SCC3 BYY Work process has left reconnect status
09:18:34 DIA  3 123 SAP*         SCC3 AB0 Run-time error "TIME_OUT" occurred
09:18:35 DIA  3 123 SAP*         SCC3 AB1 > Short dump "000427 092136 sap1tr0 SAP* " generated
09:18:35 DIA  3 123 SAP*         SCC3 R68 Perform rollback
09:18:35 DIA  3 123 SAP*         SCC3 AB0 Run-time error "TIME_OUT" occurred
09:18:35 DIA  3 123 SAP*         SCC3 AB1 > Short dump "000427 092138 sap1tr0 SAP* " generated
09:18:35 DIA  3 123 SAP*         SCC3 D01 Transaction termination SY (Time limit exceeded. )
09:18:35 DIA  3 123 SAP*         SCC3 R68 Perform rollback
09:18:54 DIA  4 000 SAPSYS            BYM SQL error 0 (possibly a network error); WP in reconnect
status
09:18:54 DIA  4 000 SAPSYS            R68 Perform rollback
09:18:55 DIA  4 000 SAPSYS            BYY Work process has left reconnect status
09:18:55 DIA  4 000 SAPSYS            AB0 Run-time error "DBIF_RTAB_SQL_ERROR" occurred
09:18:55 DIA  4 000 SAPSYS            AB1 > Short dump "000427 092157 sap1tr0             " generated
09:18:55 DIA  4 000 SAPSYS            D01 Transaction termination SY (#INCLUDE DBIF_RSQL_SQL_ERROR )
09:18:55 DIA  4 000 SAPSYS            R68 Perform rollback
09:18:55 DIA  4 000 SAPSYS            AB0 Run-time error "TIME_OUT" occurred
09:18:55 DIA  4 000 SAPSYS            AB1 > Short dump "000427 092157 sap1tr0 SAPSYS " generated
09:18:55 DIA  4 000 SAPSYS            D01 Transaction termination SY (Time limit exceeded. )
09:18:55 DIA  4 000 SAPSYS            R68 Perform rollback
09:18:55 DIA  4 000 SAPSYS            R5C Terminal type 032: Incorrect output destination/message type
TERM
09:18:55 DIA  4 000 SAPSYS            R68 Perform rollback
09:18:55 DIA  4 000 SAPSYS            R47 Delete session 001 after error 005
09:18:59 DIA  0                       BV4 Work process is in reconnect status
09:18:59 DIA  0                       BYY Work process has left reconnect status
09:19:06 DIA  1                       BV4 Work process is in reconnect status
09:19:06 DIA  1                       BYY Work process has left reconnect status
09:19:32 DIA  3                       BV4 Work process is in reconnect status
09:19:32 DIA  3                       BYY Work process has left reconnect status
```

Figure 31. Part of the syslog taken during the line failure

3.4.5 DB2 stop with force mode impact on the OS/390 database server

Purpose

Show what happens when the DB2 subsystem DBK1 is stopped with force.

Tasks

1. Create batch workload at the central instance.
2. Use dialog processes on central instance and application server.
3. Issue the command `/=DBK1 stop DB2 MODE(FORCE)`.

Results

After stopping DB2 subsystem DBK1 with force mode, a failover took place almost immediately. The work process trace file shows an SQL error -924, as shown in Figure 32 on page 78. The system log file is shown in Figure 33 on page 79.

Work on the application server can continue until commit to the database is required, then the process hangs.

Batch jobs failed and were abandoned (see Figure 30 on page 75).

```
B  Fri Apr 28 16:26:13 2000
B  ***LOG BV4=> reconnect state is set for the work process [dblink   1228]
B  ***LOG BYY=> work process left reconnect state [dblink   1671]
B
B  Fri Apr 28 16:26:14 2000
B  ***LOG BV4=> reconnect state is set for the work process [dblink   1228]
B  ***LOG BYY=> work process left reconnect state [dblink   1671]
B
B  Fri Apr 28 16:26:15 2000
B  db_con_set_reco_flag: 0: name = R/3, handle = 0
B  Reconnect state is entered by connection:
B  0: name = R/3, concnt= 000000 state = ACTIVE      , perm = YES, reco = YES
B  ***LOG BYM=> severe db error -924       work process is in the reconnect state [dbsh     0580]
B  ***LOG BY4=> sql error -924   performing SEL on table TFDIR      [dbtrtab  6242]
B  ***LOG BY0=>  DSNT408I SQLCODE = -924, ERROR:  DB2 CONNECTION INTERNAL ERROR, 0001, 0100,
00F30018
M  ***LOG R68=> ThIRollBack, roll back () [thxxhead 9404]
B  db_con_set_reco_flag: 0: name = R/3, handle = 0
B  Reconnect state is entered by connection:
B  0: name = R/3, concnt= 000000 state = ACTIVE      , perm = YES, reco = YES
B  ***LOG BYM=> severe db error -924       work process is in the reconnect state [dbsh     0580]
B  hdl_error_on_commit_rollback: DB-ROLLBACK ok. RECONNECT state
B  db_con_reconnect performing the reconnect for con:
B  0: name = R/3, concnt= 000000 state = INACTIVE    , perm = YES, reco = YES
C  DBSL TRACE: ICLI client functions successfully loaded.
(lib=/usr/sap/RED/SYS/exe/run/ibmiclic.o,sldb2.c,l=8994)
C  ********* Profile Parameter for connection 0 **********
C  Connecting to standby-Host
C  DBIF Versions:
C    R/3 Version       : 12 (4.6B)
C    API Version       : 9
C  ICLI-Client Versions:
C    R/3 Version       : 12 (4.6B)
C    API Version       : 9
C    PROTOCOL Version  : 1
C    Internal Version  : 0
C  Workprocess         : 0
C  USE_NO_HOLD         : 0
C  USE_OPT_1_ROW       : 0
C  DB_UID              : SAPR3
C  DB_ID               : RED
C  SAPSYSTEM           : 00
C  Definitions for standby host
C  DB_SSID             : DBK4
C  DBHOST              : wtsc54oe
C  PORT                : 6533
C  CCSID               : 819
C  KEEP_X_LOCK         : 1
C  RO_COMMIT_SUPPRESS  : 0
C
C  Used Bugfixes:
C  ************************************
C  Network-Protocol    : TCP
C  *** ERROR =>
[00003BE6] ICLC2520W Security is switched off - it should be switched off on the server too!
           Time: 956953575.232768 = (Fri) 2000 Apr 28 16:26:15.232768
 [dbsldb2. 2141]
C  Your DB2 system works with ASCII code page 819 and EBCDIC code page 37
B  Connection 0 opened
B  Wp  Hdl ConName              ConCnt ConState     TX  PRM RCT Date     Time
B  000 000 R/3                  000000 ACTIVE      NO  YES NO  20000428 162615
B  001 000 R/3                  000000 INACTIVE    NO  YES NO  20000428 155843
B  002 000 R/3                  000000 INACTIVE    NO  YES NO  20000428 155843
B  003 000 R/3                  000000 INACTIVE    NO  YES NO  20000428 155843
B  004 000 R/3                  000000 INACTIVE    NO  YES NO  20000428 155843
B  005 000 R/3                  000000 INACTIVE    NO  YES NO  20000428 155843
B  006 000 R/3                  000000 INACTIVE    NO  YES NO  20000428 155843
B  007 000 R/3                  000000 INACTIVE    NO  YES NO  20000428 155843
B  008 000 R/3                  000000 ACTIVE      NO  YES NO  20000428 162614
B  009 000 R/3                  000000 INACTIVE    NO  YES NO  20000428 155843
B  010 000 R/3                  000000 INACTIVE    NO  YES NO  20000428 155843
B  011 000 R/3                  000000 INACTIVE    NO  YES NO  20000428 155843
B  012 000 R/3                  000000 INACTIVE    NO  YES NO  20000428 155843
B  Successful reconnect for connection:
B  0: name = R/3, concnt= 000000 state = ACTIVE      , perm = YES, reco = NO
B  ***LOG BYY=> work process left reconnect state [dblink   1671]
A
A  Fri Apr 28 16:26:20 2000
A  Fri Apr 28 16:26:20 2000
A  ABAP/4 Program SAPLSTU3                          .
A  Source LSTU3F01                            Line 158.
A  Error Code DBIF_RTAB_SQL_ERROR.
A  Module  $Id: //bas/46B/src/krn/runt/ablink.c#1 $ SAP.
A  Function ab_tfdir Line 139.
A  &INCLUDE DBIF_RSQL_SQL_ERROR.
A
B
B  Fri Apr 28 16:26:24 2000
B  ***LOG BV4=> reconnect state is set for the work process [dblink   1228]
B  ***LOG BYY=> work process left reconnect state [dblink   1671]
B
B  Fri Apr 28 16:28:03 2000
B  ***LOG BV4=> reconnect state is set for the work process [dblink   1228]
B  ***LOG BYY=> work process left reconnect state [dblink   1671]
```

Figure 32. Extract from tracefile of work process after forced stop of DB2

```
Time     Ty. Nr Cl. User         Tcod MNo Text
Date : 28.04.00

16:26:08 BTC  8 123 SAP*              BYM SQL error -923 (possibly a network error); WP in
reconnect status
16:26:08 BTC  8 123 SAP*              R68 Perform rollback
16:26:08 BTC  8 123 SAP*              F6H Database error: TemSe->XRTAB(3)->1 for table TST01 key
16:26:13 DIA  0                       BV4 Work process is in reconnect status
16:26:13 DIA  0                       BYY Work process has left reconnect status
16:26:14 DIA  0                       BV4 Work process is in reconnect status
16:26:14 DIA  0                       BYY Work process has left reconnect status
16:26:15 BTC  8 123 SAP*              BYY Work process has left reconnect status
16:26:15 DIA  0 123 SAP*         ST04 BYM SQL error -924 (possibly a network error); WP in
reconnect status
16:26:15 DIA  0 123 SAP*         ST04 R68 Perform rollback
16:26:15 DIA  0 123 SAP*         ST04 BYY Work process has left reconnect status
16:26:17 ENQ  7                       BV4 Work process is in reconnect status
16:26:17 ENQ  7                       BYY Work process has left reconnect status
16:26:17 DIA  0 123 SAP*         ST04 AB0 Run-time error "DBIF_RTAB_SQL_ERROR" occurred
16:26:17 BTC  8 123 SAP*              AB0 Run-time error "DBIF_RSQL_SQL_ERROR" occurred
16:26:20 DIA  0 123 SAP*         ST04 AB1 > Short dump "000428 162924 sap1tr0 SAP* " generated
16:26:20 BTC  8 123 SAP*              AB2 > Include LSCC1U03 line 0028.
16:26:20 BTC  8 123 SAP*              AB0 Run-time error "DBIF_DSQL2_KEY_ALREADY_EXISTS" occurred
16:26:20 BTC  8 123 SAP*              AB1 > Short dump "000428 162924 sap1tr0 SAP* " generated
16:26:20 BTC  8 123 SAP*              D01 Transaction termination 00 (DBIF_RSQL_SQL_ERROR
20000428162924sap1tr0 SAP* 1231 )
16:26:20 BTC  8 123 SAP*              R68 Perform rollback
16:26:22 DIA  1                       BV4 Work process is in reconnect status
16:26:22 DIA  1                       BYY Work process has left reconnect status
16:26:22 DIA  1 000 SAPSYS            BYM SQL error -924 (possibly a network error); WP in
reconnect status
16:26:22 DIA  1 000 SAPSYS            R68 Perform rollback
16:26:22 DIA  1 000 SAPSYS            BYY Work process has left reconnect status
16:26:23 DIA  1 000 SAPSYS            AB0 Run-time error "DBIF_RSQL_SQL_ERROR" occurred
16:26:23 DIA  1 000 SAPSYS            AB1 > Short dump "000428 162930 sap1tr0 SAPSYS " generated
16:26:23 DIA  1 000 SAPSYS            D01 Transaction termination 00 (DBIF_RSQL_SQL_ERROR
20000428162930sap1tr0 SAPSYS 0001 )
16:26:23 DIA  1 000 SAPSYS            R68 Perform rollback
16:26:24 DIA  0                       BV4 Work process is in reconnect status
16:26:24 DIA  0                       BYY Work process has left reconnect status
16:27:03 DIA  2                       BV4 Work process is in reconnect status
16:27:03 SPO 10                       BV4 Work process is in reconnect status
16:27:03 SPO 10                       BYY Work process has left reconnect status
16:27:03 DIA  2                       BYY Work process has left reconnect status
16:27:03 DIA  2 000 SAPSYS            BYM SQL error -924 (possibly a network error); WP in
reconnect status
16:27:03 DIA  2 000 SAPSYS            R68 Perform rollback
16:27:04 DIA  2 000 SAPSYS            BYY Work process has left reconnect status
16:27:04 DIA  2 000 SAPSYS            AB0 Run-time error "DBIF_RSQL_SQL_ERROR" occurred
16:27:04 DIA  1                       BV4 Work process is in reconnect status
16:27:04 DIA  1                       BYY Work process has left reconnect status
16:27:05 DIA  2 000 SAPSYS            AB1 > Short dump "000428 163011 sap1tr0 SAPSYS " generated
16:27:05 DIA  2 000 SAPSYS            D01 Transaction termination 00 (DBIF_RSQL_SQL_ERROR
20000428163011sap1tr0 SAPSYS 0001 )
16:27:05 DIA  2 000 SAPSYS            R68 Perform rollback
16:27:16 ENQ  7                       BV4 Work process is in reconnect status
16:27:16 ENQ  7                       BYY Work process has left reconnect status
16:28:03 DIA  0                       BV4 Work process is in reconnect status
16:28:03 SPO 10                       BV4 Work process is in reconnect status
16:28:03 DIA  0                       BYY Work process has left reconnect status
16:28:03 SPO 10                       BYY Work process has left reconnect status
16:29:03 DIA  3                       BV4 Work process is in reconnect status
16:29:03 DIA  3                       BYY Work process has left reconnect status
16:29:03 SPO 10 000 SAPSYS            BYM SQL error -924 (possibly a network error); WP in
reconnect status
16:29:03 DIA  3 000 SAPSYS            BYM SQL error -924 (possibly a network error); WP in
reconnect status
16:29:03 DIA  3 000 SAPSYS            R68 Perform rollback
16:29:03 DIA  3 000 SAPSYS            BYY Work process has left reconnect status
16:29:03 DIA  3 000 SAPSYS            AB0 Run-time error "DBIF_RSQL_SQL_ERROR" occurred
16:29:04 SPO 10 000 SAPSYS            R68 Perform rollback
16:29:04 DIA  3 000 SAPSYS            AB1 > Short dump "000428 163210 sap1tr0 SAPSYS " generated
16:29:04 DIA  3 000 SAPSYS            D01 Transaction termination 00 (DBIF_RSQL_SQL_ERROR
20000428163210sap1tr0 SAPSYS 0001 )
16:29:04 DIA  3 000 SAPSYS            R68 Perform rollback
16:29:04 SPO 10 000 SAPSYS            BYY Work process has left reconnect status
16:29:09 DIA  1                       BV4 Work process is in reconnect status
16:29:09 DIA  1                       BYY Work process has left reconnect status
16:29:16 ENQ  7                       BV4 Work process is in reconnect status
16:29:16 ENQ  7                       BYY Work process has left reconnect status
16:30:03 DIA  0                       BV4 Work process is in reconnect status
16:30:03 DIA  0                       BYY Work process has left reconnect status
16:30:03 BTC  8                       BV4 Work process is in reconnect status
16:30:03 BTC  8                       BYY Work process has left reconnect status
16:34:03 DIA  2                       BV4 Work process is in reconnect status
16:34:04 DIA  2                       BYY Work process has left reconnect status
```

Figure 33. Extract from system log after forced stop of DB2

```
Date       Time     Host          User         Clt X Error ID

2000/04/28 16:30:11 sap1tr0       SAPSYS       000 C DBIF_RSQL_SQL_ERROR
2000/04/28 16:32:10 sap1tr0       SAPSYS       000 C DBIF_RSQL_SQL_ERROR

Contents of dump 1 and 2
Database error text........: " DSNT408I SQLCODE = -924, ERROR: DB2 CONNECTION
 INTERNAL ERROR, 0001, 0100, 00F30018 DSNT418I SQLSTATE = 58006 SQLSTATE RETURN
 CODE DSNT415I SQLER
```

Figure 34. Generated dump and contents of dump

3.4.6 . Apply PTF in a running SAP R/3 system

Purpose

Show the approach we used to apply a PTF to the ICLI client and server in a running system.

Criteria

Our goal is to apply maintenance to LPAR SC54 with no impact to the availability of our SAP systems, and then repeat the test for LPAR SC42.

Tasks

1. Create batch workload at the central instance.
2. Use dialog processes on central instance and application server.
3. Issue the command `/P ICLIRED2` to stop the second ICLI server, which failed over the workload to LPAR SC42 and DBK3.
4. Issue shutdown to ICLIRED3 via command `/P ICLIRED3`.
5. Issue shutdown to DBK2 via command `/=DBK2 SHUTDOWN DB2`.
6. Issue shutdown to DBK4 via command `/=DBK4 SHUTDOWN DB2`.
7. Apply PTF UW68831 on LPAR SC54 and IPL.

   To accomplish this task, we applied the fix to our OS/390 build system via SMP/E. Then we copied the updated build SYSRES and HFS volumes to our alternate IPL volumes and IPLed.

   For more information regarding this technique, refer to *Parallel Sysplex - Software Management for Availability,* SG24-5451.
8. Issue command `/=DBK2 START DB2` on SC54.
9. Issue command `/=DBK4 START DB2` on SC54.
10. Run the DB2 BIND and GRANT jobs for DBK2 and DBK4.
11. Logon to `sap2tr0` and create FTP connection to LPAR SC54.
12. Download ICLI Client code ‘SYS1.SFOMDATA(FOME46BA)’ and rename it to ibmiclic.o into a /tmp file.
13. Issue command `STOPSAP` on application server `sap2tr0`.
14. Copy ibmiclic.o to SAP executable library /usr/sap/RED/SYS/exe and issue AIX command `chmod 755 ibmiclic.o`.
15. Issue startup of ICLIRED2 via command `/S ICLIRED2`.

16. Issue startup of ICLIRED3 via command `/S ICLIRED3`.
17. Issue command `STARTSAP` on application server sap2tr0.
18. Issue the command `/P ICLIRED` to stop the primary ICLI server which failed over the workload to LPAR *SC54* and DBK4.
19. We repeated steps 4 through 17 on LPAR SC42 and central instance server sap1tr0.

Results

This approach allowed us to apply maintenance to one LPAR while the other LPAR was carrying the full SAP database workload. We experienced no impact to the central instance during failover of application server sap2tr0 to DBK3 on LPAR SC42.

However, the PTF we chose caused us to apply maintenance on OS/390 and SAP on AIX. During steps 13 through 17, we did have to shut down SAP on the application server to apply a new ICLI client executable, but the outage was minimal and the central instance experienced no interruptions.

We then repeated these steps for LPAR SC42 and experienced no interruptions until steps 13 through 17, when we applied maintenance to our central instance on sap1tr0. Both the SAP central instance and application server had to be shut down for a short period of time to copy the executable in to sap1tr0s SAP executable library and run the DB2 BIND and GRANT jobs for DBK1 and DBK3.

3.5 Database high availability summary

Our conclusion after the testing is as follows:

Using data sharing on a Parallel Sysplex with SAP R/3 configured as used in our configuration, it is possible to improve availability and diminish otherwise necessary outage. This otherwise necessary outage can be prevented by failing over the database to the other LPAR in a time frame without batch jobs on the active application server and LPAR where maintenance is needed. This is necessary because failing over terminates active batch jobs and cancels and rolls back active online transactions on the target application server / DB2 subsystem. The same applies if maintenance is needed on OS/390.

In our configuration maintenance on AIX or SAP R/3 on the central instance is not possible without an outage, because the central instance itself is a single point of failure. Using HACMP or Microsoft clustering will offer some possibilities to apply this kind of maintenance. These issues are not in our scope.

Chapter 4. Database administration - reducing planned outages

This chapter describes those functions of a database administrator that have potential impact on availability. In particular, we concentrated on functions that must be accomplished, on either a scheduled or unscheduled basis, without an interruption to SAP R/3 usage. Examples of such functions are definition changes, database reorganization and collection of statistical data. We show how these activities were done in our installation during times that a simulated production load existed; we document the impact of performing the function on that load.

4.1 Altering the database layout

In this section we describe the database layout of a SAP R/3 on DB2 UDB for OS/390 system. We enumerate the main tasks that administrators and developers do in which DDL is involved, like creation of tables and indexes and modification of definition parameters and storage options. We also discuss DDL/DML concurrency, introduced with DB2 Version 6, and give the options that allow greater availability of data.

4.1.1 Requirements for the current database layout

An important aspect to consider when planning for the continuous availability requirements of the SAP R/3 database system is the need to introduce changes to the different database components. These changes include both the creation of new objects like tables or indexes and changes to the definition parameters of already existing objects.

Starting with the 4.5a release of SAP R/3 on DB2 UDB for OS/390 and continuing in release 4.6b on which this redbook is based, there have been substantial changes to the database layout with respect to preceding SAP R/3 releases. The changes were introduced in order to optimize the accomplishment of the following requirements:

- General requirements
  - High end performance
  - Efficient monitoring and administration
  - Unlimited use of DB2 code pages
- SAP R/3 requirements
  - Identical definitions for structure and names of transparent tables and views in the ABAP Dictionary (DDIC) and in the DB2 Database (DB).
  - Index key definitions also must be the same in both.
- DB2 for OS/390 requirements
  - Not more than 100 tables per database
    (considered optional in terms of performance and operation)
  - Efficient applications of DB2 utilities, such as RUNSTATS, REORG, and COPY.
  - Flexible and simple naming convention that is non-destructive to customer modifications.

- OS/390 requirements when SMS is used for managing the extensions of DB2 data sets
  - Relevant storage information included in the data set names of tablespaces and indexes.
  - If SAP R/3 needs to move a table to a newly created tablespace, the new database and tablespace names should be such that there is no need to change data sets qualifiers in the ACS routines.

As a result of all these requirements the developed database layout has significantly increased the number of databases and tablespaces. This has the disadvantage of having to control a larger number of objects in order to achieve the advantage of improved concurrency and availability properties of the database.

Table 5 lists the types and total numbers of the objects in our recently installed SAP R/3 system.

Table 5. Total objects created by in the SAP R/3 database.

Type of objects	Total number
Stogroup	26
Databases	9166
Tablespaces	9167
Tables	20001
Indexes	23242
Total size SAP R/3 Database	25 GB

In order to understand all the topics referred in this chapter we include a description of the database layout based in the SAP publication, *SAP R/3 Database Administration Guide: DB2 for OS/390 Release 4.6 B* Material Number 51008497.

4.1.2 The database layout

A table, as the elementary structure in the database, is going to be the basis for the naming of the rest of associated components. Each time a new table is created in the R/3 database using transaction SE11, by a transport or during an upgrade, the system will ask for a set of technical information that anticipates how it is going to be used by the system. With these technical settings, and the table's name, R/3 is able to obtain the database name, the tablespace name, the stogroup used to store the tablespace the stogroup used to store all the indexes and the index names.

4.1.2.1 R/3 technical settings

The R/3 technical settings of a table allow you to categorize the table according to its space requirements (usage/growth), I/O rates and so on. You can specify the following technical attributes using transaction SE11:

Data class (TABART)

The data class is a table's attribute that describes the use of a table. For example, it allows you to differentiate between a table with master data and a table with

transactional data. Each data class has an associated 2-character ID (Storage ID) that is used internally by R/3 to construct the names of DB2 databases. The data class is also directly mapped to DB2 stogroups.

Table 1 in *SAP R/3 Database Administration Guide: DB2 for OS/390 Release 4.6 B* (Material Number 51008497) lists available data classes and the corresponding storage IDs and DB2 stogroups.

Size category (TABSIZE)

The size category (between 0 and 4) describes the estimated space requirements of a table. A higher number identifies a larger table.

SAP R/3 buffering

SAP R/3's table buffering is one of the most important performance tuning mechanisms. Under certain circumstances, SAP R/3 enables storing of tables locally on each application server. The time consuming process of repeatedly accessing the database is thus avoided. The following types of buffering are possible:

- Full buffering

  In case of read access, the entire table is loaded into the R/3 buffer.

- Generic buffering

  The generic key is defined as part of the primary key. During the access of a record from a table that is generically buffered, all records whose generic field key fields correspond to this record are loaded into the R/3 buffer.

- Single record buffering
  Only those records actually being accessed are loaded into the R/3 buffer.

4.1.2.2 Mapping between SAP R/3 and DB2

The following rules describe the way SAP R/3 implement the naming convention within the database.

- DB2 Storage Groups
  There is a fixed number of DB2 storage groups defined for an SAP R/3 system (in our newly installed R/3 system there were 24). All of them start with the letters SAP followed by two characters associated with a Data Class and, in the last position, the character D for Data and I for indexes (i.e. SAPU1D is the stogroup name for tablespaces of user defined tables -Data Class U1- and SAPU1I is the stogroup name for the same table's indexes).

  All DB2 stogroups used by SAP R/3 are associated with the same Integrated Catalog Facility Catalog (parameter VCAT within CREATE STOGROUP statement). This VCAT is a parameter of SAP R/3 R3setup installation tool.

- Databases
  In the SAP R/3 environment, the DB2 database is an object that consists of a DB2 tablespace and all the indexes on the tables contained in the tablespace. Specifically, for a single-table tablespace, the database includes one table and all its indexes.

  The real value of this DB2 object in the SAP R/3 environment is in its name; it carries all the technical settings of the contained tables. For a given table the associated database name is constructed as follows:

[STORAGEID] [TABSIZE] [BUFFER] X [ABC]

where the parameters have the following meaning:

[STORAGEID]	SAP R/3 Data Class identified by 2 characters A0, A1, A2, and so on. The unique relation between data class and storage ID is defined in the R/3 table TADB2.
[TABSIZE]	SAP R/3 size category (1 digit: 0,1,2,3,4)
[BUFFER}	Actual SAP R/3 buffering on the application server at creation time (1 digit): - 0 = no buffering - 1 = full buffering - 2 = generic buffering - 3 = single-record buffering
X	Placeholder for future developments (earlier releases use '#' as a placeholder)
[ABC]	3 random characters ([A-Z] [0-9])

With the naming convention in mind you can determine table's technical settings by looking at the appropriate DB2 catalog table (SYSIBM.SYSTABLES), or, more importantly, at the underlying dataset name (remember that a database name is a constituent part of the underlying data set name). Tat allows you to control the physical placement of tables and indexes, for example, by passing the technical setting details to an ACS routine.

- Tablespaces
  DB2 tablespace is a common name for one or more data sets used to store single or multiple tables. In the SAP R/3 environment the SAP R/3 buffering technical settings is the only criterion for deciding whether a tablespace stores a single table or multiple tables. If a given table is not SAP R/3 buffered at creation time, it is stored in a dedicated tablespace that does not contain any other table. Otherwise, the table is put in multi-table tablespace.

  A multi-table tablespace holds up to 100 tables. R/3 assigns the name XSAP to this type of tablespace (up to SAP R/3 Release 4.5B the default name #SAP was used). For a single-table tablespace, the system determines the tablespace name by taking the first seven characters of the associated table name. Naming space qualifiers are ignored and the character _ is replaced by X.

  The database and tablespace name of a table in a single-table tablespace are only slightly modified if the table needs to be converted due to structural changes, for example, during an upgrade. In that case, the character "X" is appended to the tablespace name. Therefore, if you use the qualifier

  [VCAT].DSNDBD.[DATABASE].[TABLESPACE]*.[SUFFIX}

  in your ACS routines, there is no need to adjust these routines after a table conversion.

- Indexspaces
  Each DB2 index occupies its own DB2 indexspace placed in the DB2 database that contains the base table of index. Default values for stogroups or primary and secondary quantities are defined in table IGDB2.

4.1.3 Availability enhancements for administration

SAP R/3 on DB2 UDB for OS/390 takes advantage of recent enhancements introduced in versions 5 and 6 of DB2.

Data Definition Language (DDL) is a set of SQL instructions which allow the definition of new objects like databases, tablespaces, tables and indexes.

A Database Descriptor (DBD) contains definitions of tablespaces, tables and indexes that belong to the same database. Each time that a DDL statement is executed, the DBD of the associated database is deleted and recreated again to reflect the change. The Database Manager prevents access to the DBD with an exclusive lock between the time the sentence is executed and the next commit.

As long as this lock is held, any other DDL statement executed against the same database must wait for the lock to be released. In order to reduce the duration of the lock, there have been improvements to the functionality and performance of these sentences:

- Locking isolation for the internal plan to execute DDL changed from RR to RS improving concurrency across different databases.
- Parallel tasks enabled to do Data Set management activities:
  - 20 tasks enabled for define data set (CREATE).
  - 20 tasks enabled for delete/define data set (RESET used by REORG/LOAD/RECOVER y REBUILD).
  - 20 tasks enabled for delete data set (DROP).
- Defer Delete and re-insert DBD until COMMIT.
- There is no need to STOP TABLESPACE when altering PRIQTY, SECQTY, FREEPAGE, PCTFREE, CLOSE and ERASE.
- A new index on SYSIBM.SYSSTOGROUP has improved the performance of the statement DROP STOGROUP.
- It is possible to avoid acquiring S-DBD locks for Dynamic SQL when statement caching is allowed.
- ALTER INDEX has been enhanced to allow altering of limit key values on partitioning indexes.
- It is possible to extend a VARCHAR column length without dropping and re-creating the table.
- It is possible to support up to 16 different versions before REORG or REBUILD.

SAP R/3 on DB2 UDB for OS/390 has improved DDL concurrency by reducing the number of objects within the DB2 databases. This has improved and released the work of administration.

4.1.4 Partitioning a large tablespace

There are situations in which it becomes necessary to perform actions over the structure or definition of different objects in the database that might affect availability of the corresponding set of data. Clearly, one of these situations is having a tablespace whose size is reaching the 64 GB limit for a non-partitioned tablespace.

The first option in this case would be to compress the tablespace with the statement ALTER TABLESPACE COMPRESS YES. If the tablespace space is still growing and reaching the 64 GB limit the only solution is partitioning the tablespace.

Following are the current space addressability limitations for tablespaces in DB2 UDB for OS/390:

1. For non-partitioned tablespaces the maximum size is 64 GB. distributed in as many as 32 data sets of 2 GB of data each.
2. For partitioned tablespaces which are not defined with DSSIZE greater than 2 GB, the maximum number of partitions that can be defined is 64 with the following maximum sizes for each partition:
   - 4 GB for 1 to 16 partitions
   - 2 GB for 17 to 32 partitions
   - 1 GB for 33 to 64 partitions

   In any case the maximum allowed space is always 64 GB.
3. For partitioned tablespaces defined with DSSIZE greater than 2 GB, the maximum number of partitions is 254 with a maximum size for each partition of 64 GB. The requirements to create this kind of tablespaces are:
   - DB2 is running with DFSMS Version 1 Release 5.
   - The data sets for the tablespace are associated with a DFSMS data class that has been specified with extended format and extended addressability.

SAP R/3 on DB2 UDB for OS/390 supports partitioning of tablespaces. This means that it is possible to let the R/3 Dictionary know the partitioning index, the partitioned key and all the storage attributes associated with the new definition of the tablespace.

Following we explain the steps we followed to partition the tablespace for the table “TODIR” of 1286132 rows and 128 MB. It is not a very large table but it can describe the process and give some orientation about unavailability times. Most of the steps were done with SAP R/3 screens. The exceptions were the unload and reload of rows of data that we controlled from a TSO session with DB2 utilities.

1. We checked the planned partitioning key for access by any updated statement within an SAP R/3 program:
   a. In transaction SE11 we chose Utilities and Where-used list.
   b. We selected Programs and continued.
   c. We searched for the string “UPDATE” in the expanded where-used list. If a partitioning key is used in a UPDATE statement, each update locks the complete table. Therefore, you should consider using a different index for partitioning.
2. We collected limitkey data.
   The *SAP R/3 Database Administration Guide: DB2 for OS/390 Release 4.6 B* says that this step is optional, but we used it in order to optimize the distribution of rows.
   a. We chose the table in transaction SE14.

b. We chose storage parameters. In the next figure we show the storage attributes for the tablespace as it is currently defined.

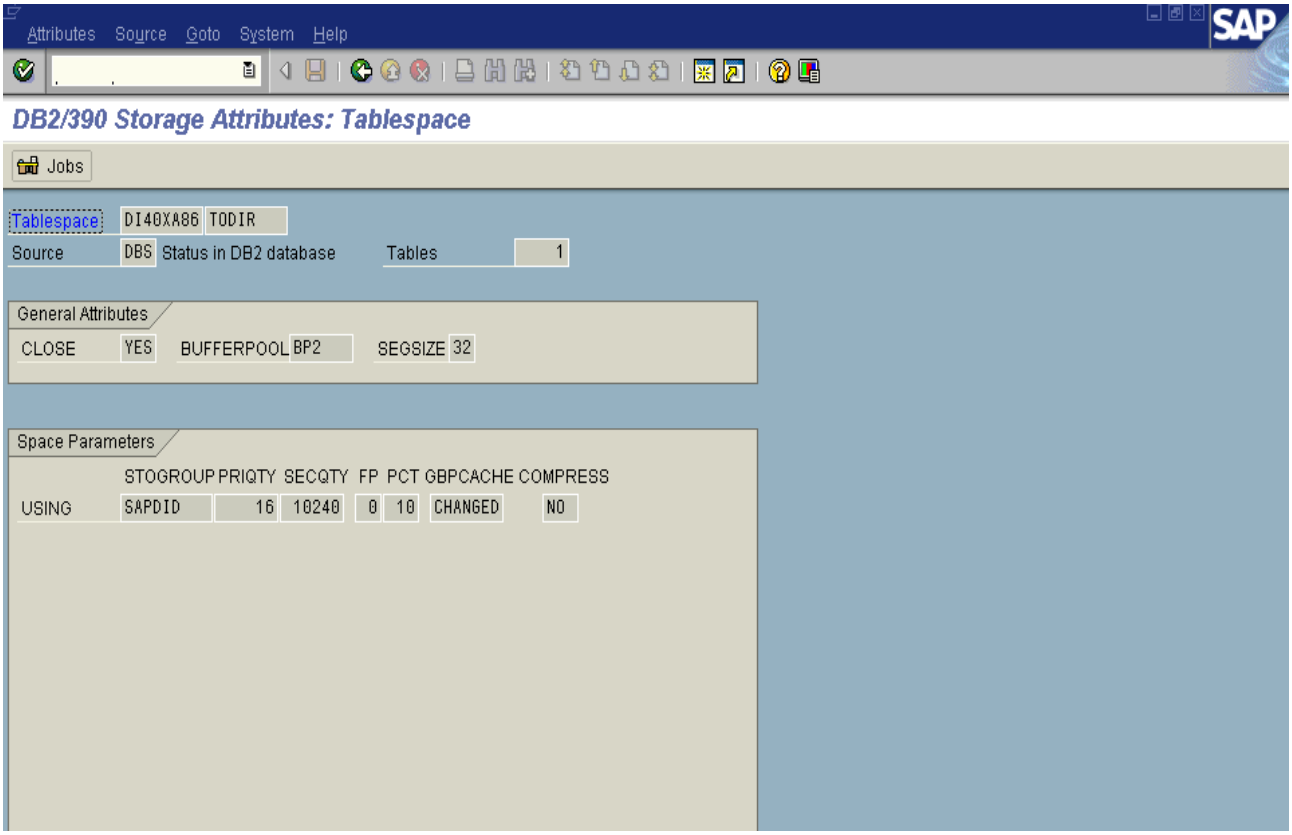


Figure 35. SE14 Transaction. DB2/390 Storage Attributes: Tablespace

c. In the menu bar: Goto and Limitkey Data.

d. We entered Index ID and a number of partitions. SAP R/3 supports here more than 64 partitions, up to a limit of 254 partitions. Entering a higher value is corrected by SAP R/3 to 254. For our example we chose 25 partitions.

e. We started a job to collect the limitkey data.

Collecting limitkey data can take several hours depending on the size of the table, but it doesn't affect availability of the related table because the table is scanned with Isolation UR.

3. From this point on, the tablespace will be unavailable for users. We issued the following DB2 command using transaction "DB2" :

```
-START DB(DI40XA86) SPACE(*) ACCESS(RO)
```

This allowed us to be sure that the database was unavailabe for updates. SAP R/3 transactions would fail while trying access the tablespace in write mode.

We recommend that you establish a system recovery point with a backup of the entire system as is explained in "External backups with RVA SnapShot" on page 119 of this redbook.

We also obtained a backup copy of the tablespace using DSN1COPY. We queried the identifiers DBID, PSID and OBID to be able to recover if something went wrong.

4. In order to get access to run utilities we started the database with access UT.
5. We unloaded the data using REORG UNLOAD EXTERNAL. We found that the space required for the SYSREC file was bigger than the tablespace size itself. Here we show the JCL we used for this special REORG.

```
//JOBLIB   DD  DISP=SHR,
//             DSN=DSN610.SDSNLOAD
//DSNTICX EXEC PGM=DSNUTILB,PARM='DB2K,RUNLD',REGION=1024K
//UNLD     DD  DSN=DB2V610K.REORG.STEP1.TODIR,DISP=(NEW,CATLG,CATLG),
//             UNIT=SYSDA,SPACE=(CYL,(800,200)),VOL=SER=(SAP016,SAP010)
//SYSPRINT DD  SYSOUT=*
//SYSPUNCH DD  SYSOUT=*
//SYSIN    DD  *
  REORG TABLESPACE DI40XA86.TODIR UNLOAD EXTERNAL
        UNLDDN (UNLD)
        WORKDDN (WORK)
//*
```

6. The next step was establishing the partitioning attributes for the tablespace.
    - In the SE14 main transaction panel we chose *storage parameter* and got into edit mode to set the tablespace as partitioned.
    - At the question "*use collected limitkey data?*" we answered "*yes*" and the system used the set of key ranges previously collected. If you miss at this point definition parameters for tablespaces or indexes, OSS Notes 184399 and 162818 refer to necessary transports to DDIC in order to get support for the definition options PIECESIZE and DSSIZE.

   The following screen shows the attributes just saved:

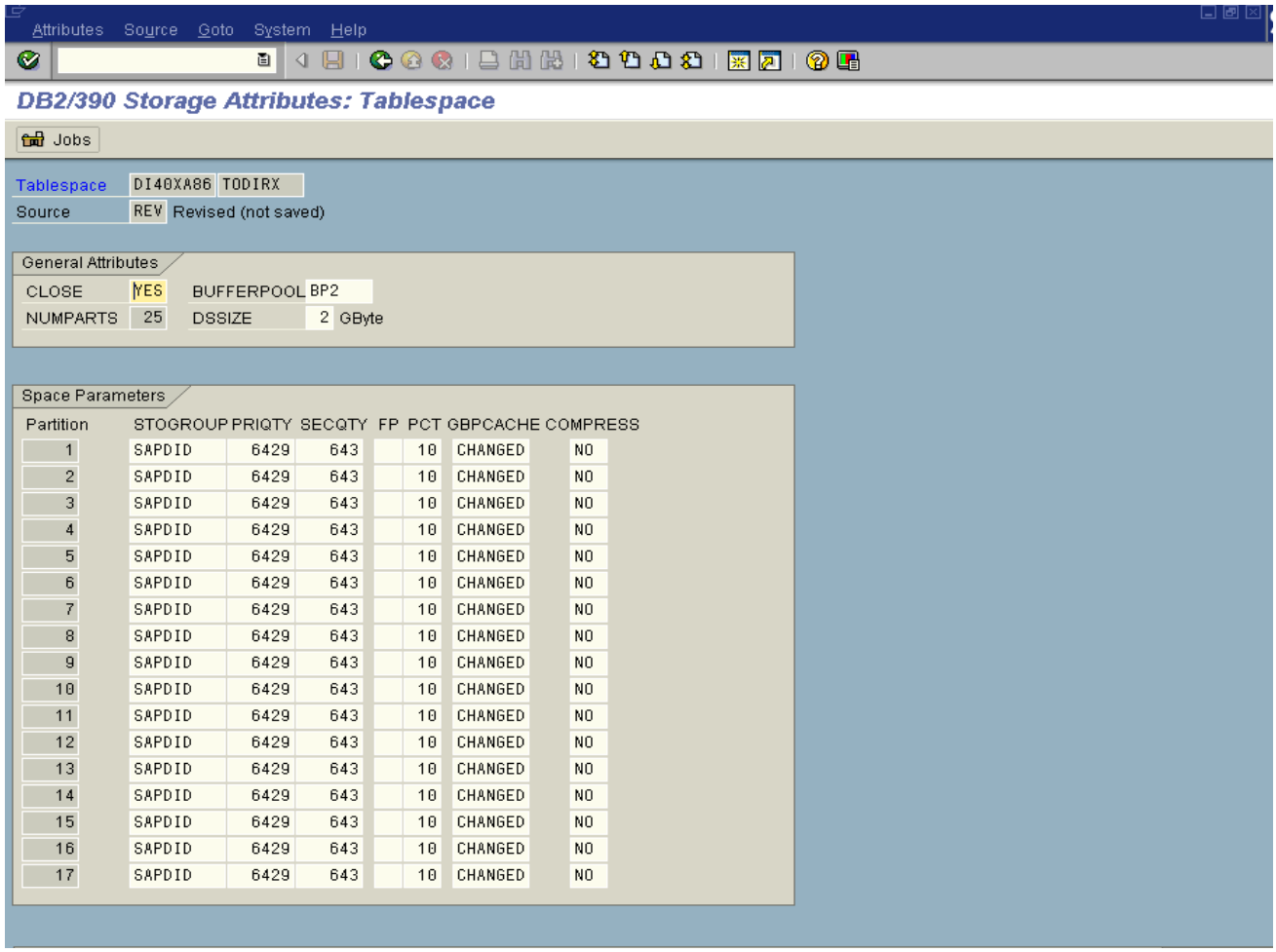


Figure 36. SE14 Transaction. Storage attributes for the partitioned tablespace.

7. We saved the new data attributes.
8. Afterwards we went back to the main screen of the transaction and chose ‘Activate and adjust database’ with the option ‘Delete data’. At this time, DB2 will need to create as many datasets as the number of partitions plus another set for index partitions. If you take an SQL trace (transaction ‘ST05’) of the process you will see a CREATE DATABASE, CREATE TABLESPACE, CREATE TABLES and as many CREATE INDEX entries as necessary.

   It is important to consider at this time that the name of the previous database has changed in its last three characters, and also the name of the tablespace now carries an ‘X’ at the end.
9. We loaded the saved table contents into the newly created table. The utility control statement can be found in the “SYSPUNCH” output from REORG. For this, we used a LOAD REPLACE. During the execution of the utility, we took an inline full image copy of the tablespace.

   Following we include the output of the LOAD execution:

```
LOAD DATA INDDN UNLD LOG NO REPLACE COPYDDN MYCOPY1 ASCII CCSID(819,0,0)
 -   INTO TABLE "SAPR3   "."TODIR             " WHEN(4:5=X'0003')
 -    ("FRAGID            " POSITION(7:12) VARCHAR,
 -     "FRAGMENT          " POSITION(13:18) VARCHAR,
 -     "FRAGNAME          " POSITION(19:140) VARCHAR,
 -     "PGMID             " POSITION(141:146) VARCHAR,
 -     "OBJECT            " POSITION(147:152) VARCHAR,
 -     "OBJ_NAME          " POSITION(153:194) VARCHAR,
 -     "EXOBJECT          " POSITION(195:197) VARCHAR,
 -     "STATUS            " POSITION(198:200) VARCHAR,
 -     "MODSTATUS         " POSITION(201:203) VARCHAR,
 -     "PREMOD            " POSITION(204:206) VARCHAR,
 -     "POSTMOD           " POSITION(207:209) VARCHAR,
 -     "CHANGEREL         " POSITION(210:217) VARCHAR,
 -     "GENFLAG           " POSITION(218:220) VARCHAR,
 -     "RELMAP            " POSITION(222:227) VARCHAR NULLIF(221)=X'FF')
T - EXISTING RECORDS DELETED FROM TABLESPACE
COPY PROCESSED FOR TABLESPACE DI40XA86.TODIRX
NUMBER OF PAGES=32185
AVERAGE PERCENT FREE SPACE PER PAGE = 10.46
PERCENT OF CHANGED PAGES =100.00
ELAPSED TIME=00:01:29 -
(RE)LOAD PHASE STATISTICS - NUMBER OF RECORDS=1286132 FOR TABLE SAPR3.TODI
DB2 IMAGE COPY SUCCESSFUL FOR TABLESPACE DI40XA86.TODIRX
(RE)LOAD PHASE STATISTICS - NUMBER OF INPUT RECORDS PROCESSED=1286132
(RE)LOAD PHASE COMPLETE, ELAPSED TIME=00:01:31
SORT PHASE STATISTICS -
NUMBER OF RECORDS=2572264
ELAPSED TIME=00:00:57
NUMBER OF KEYS=51445 FOR INDEX SAPR3.TODIR~0  PART 1
NUMBER OF KEYS=51445 FOR INDEX SAPR3.TODIR~0  PART 2
NUMBER OF KEYS=51445 FOR INDEX SAPR3.TODIR~0  PART 3
NUMBER OF KEYS=51445 FOR INDEX SAPR3.TODIR~0  PART 4
NUMBER OF KEYS=51445 FOR INDEX SAPR3.TODIR~0  PART 5
NUMBER OF KEYS=51445 FOR INDEX SAPR3.TODIR~0  PART 6
NUMBER OF KEYS=51445 FOR INDEX SAPR3.TODIR~0  PART 7
NUMBER OF KEYS=51445 FOR INDEX SAPR3.TODIR~0  PART 8
NUMBER OF KEYS=51445 FOR INDEX SAPR3.TODIR~0  PART 9
NUMBER OF KEYS=51445 FOR INDEX SAPR3.TODIR~0  PART 10
NUMBER OF KEYS=51445 FOR INDEX SAPR3.TODIR~0  PART 11
NUMBER OF KEYS=51445 FOR INDEX SAPR3.TODIR~0  PART 12
NUMBER OF KEYS=51445 FOR INDEX SAPR3.TODIR~0  PART 13
NUMBER OF KEYS=51445 FOR INDEX SAPR3.TODIR~0  PART 14
NUMBER OF KEYS=51445 FOR INDEX SAPR3.TODIR~0  PART 15
NUMBER OF KEYS=51445 FOR INDEX SAPR3.TODIR~0  PART 16
NUMBER OF KEYS=51445 FOR INDEX SAPR3.TODIR~0  PART 17
NUMBER OF KEYS=51445 FOR INDEX SAPR3.TODIR~0  PART 18
NUMBER OF KEYS=51445 FOR INDEX SAPR3.TODIR~0  PART 19
NUMBER OF KEYS=51445 FOR INDEX SAPR3.TODIR~0  PART 20
NUMBER OF KEYS=51448 FOR INDEX SAPR3.TODIR~0  PART 21
NUMBER OF KEYS=51445 FOR INDEX SAPR3.TODIR~0  PART 22
NUMBER OF KEYS=51442 FOR INDEX SAPR3.TODIR~0  PART 23
NUMBER OF KEYS=51445 FOR INDEX SAPR3.TODIR~0  PART 24
NUMBER OF KEYS=51452 FOR INDEX SAPR3.TODIR~0  PART 25
NUMBER OF KEYS=1286132 FOR INDEX SAPR3.TODIR~1 DSNURBXD
BUILD PHASE STATISTICS - NUMBER OF INDEXES=2
DSNURBXD - BUILD PHASE COMPLETE, ELAPSED TIME=00:03:27
DSNUGBAC - UTILITY EXECUTION COMPLETE, HIGHEST RETURN CODE=0
```

10. Finally, we verified that SAP R/3 could access the table (“SE16” transaction) and that the number of rows didn’t change during the process.

In Table 6 we present the results of our testing in terms of unavailability time.

Table 6. Unavailability time measurements for partition a 1 million rows table.

	Data Availability	**Elapsed Time**
Collect limitkey	YES	0 sec.
Database Backup	NO	30 sec.
Reorg unload external	READ ONLY	26 sec.
Delete and Re-create the table	NO	120 sec.
Reload table contents	NO	309 sec.
Total Time	**NO**	**485 sec.**

4.2 Space management

This section describes the potential availability problems related to lack of space, and ways of controlling and monitoring the status of the database objects. It will also refer to Storage Management System (SMS) as a solution for the management of the high number of data sets defined by the SAP R/3 database. There are new options in DB2 UDB for OS/390 that can be specified in order to save all the space allocated for empty tables.

4.2.1 Out of DASD space or extent limit reached

A broad area that can have a severe impact on the availability of the different objects is controlling the space left in the volumes available and the number of extensions in data sets.

4.2.1.1 Description of the problem

When an insert or update requires additional space, but that space is not available in the current table or index space, DB2 issues the following message:

```
DSNP007I - DSNPmmmm - EXTEND FAILED FOR
          data-set-name. RC=rrrrrrrr
          CONNECTION-ID=xxxxxxxx,
          CORRELATION-ID=yyyyyyyyyy
          LUWID-ID=logical-unit-of work-id=token
```

A look-ahead warning occurs when there is enough space for a few inserts and updates, but the index space or tablespace is almost full. On an insert or update at the end of a page set, DB2 tries to extend the data set if the amount of space available is less than all of the following values:

- The primary space quantity from the integrated catalog facility catalog
- The secondary space quantity from that catalog
- The allocation unit size times 10

If the extend request fails, then DB2 issues the following message:

```
DSNP001I - DSNPmmmm - data-set-name IS WITHIN
       nK BYTES OF AVAILABLE SPACE.
       RC=rrrrrrrr
       CONNECTION-ID=xxxxxxxx,
       CORRELATION-ID=yyyyyyyyyyy
       LUW-ID=logical-unit-of-work-id=token
```

If one of these conditions occurs, the system takes the following actions:

- For a demand request failure during restart, the object supported by the data set (an index space or a table space) is stopped with deferred restart pending.
- Otherwise, the state of the object remains unchanged. Programs receive a -904 SQL return code (SQLSTATE '57011')

The appropriate choice of actions depends on particular circumstances. If the data set has not reached its maximum size, you can enlarge it. (For the maximum size of data sets see 4.1.4, "Partitioning a large tablespace" on page 87. Otherwise you have to add a new data set to the tablespace or, if the object is partitioned, redefine the partitions.

A full description of the actions can be found in the publication: *DB2 UDB for OS/390 Administration Guide*, SC26-9003.

4.2.1.2 SAP R/3 Monitor for Tables and Indexes

In order to prevent these situations it is very important to constantly monitor the status of the objects in the database. SAP R/3 provides the database administrator with tools to perform this control.

The Tables and Indexes Monitor collects status information about the SAP R/3 system. It gets some of this information directly from the DB2 catalog, for example the number of tables. However, some information has to be collected at the operating system level, for example all space information.

Therefore, some of the information presented by the monitor is obtained using OS/390 utilities (LISTCAT function of the IDCAMS program). To do this, a JCL job that performs a LISTCAT over the whole subsystem is generated, uploaded to the host and submitted from within the SAP R/3 System in synchronous mode; that is, the process waits until the output of the JCL job has been transferred to the R/3 System. FTP is used to transfer this job to the host.

The program that triggers this JCL is RSDB2TM which is run in batch processing mode. This program is released each time you choose *refresh* in the Table and Index Monitor main screen, or what is more convenient, it can be scheduled to run on a periodic basis using the DBA Planning Calendar.

The transaction that calls *Database Performance: Tables and Indexes* is DB02. In the following figure we present the initial screen:

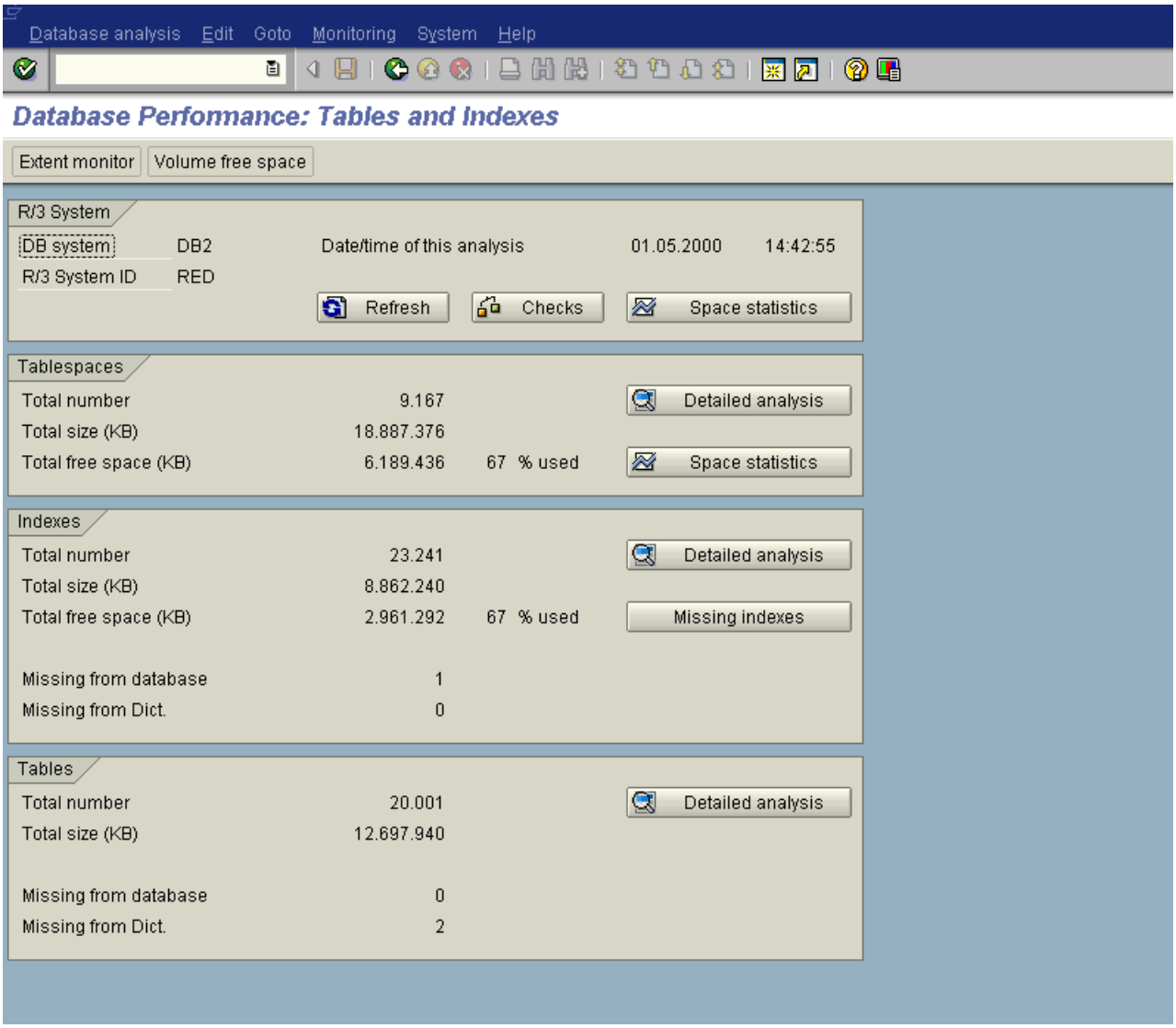


Figure 37. DB02. Database Performance: Tables and Indexes

This screen provides us with data about the name of the DB system and SAP R/3 system, the last date the statistics were updated and information about total number of objects and total amount of space allocated and used for tablespaces, indexes and tables. It is possible to get detailed information at an object level from this screen.

The Extent monitor

From the initial screen of the Table and Index Monitor, we can reach the Extent Monitor, which shows a table of all tablespaces and indexes that at the time of the last update of the data that required more extents than a specified threshold.

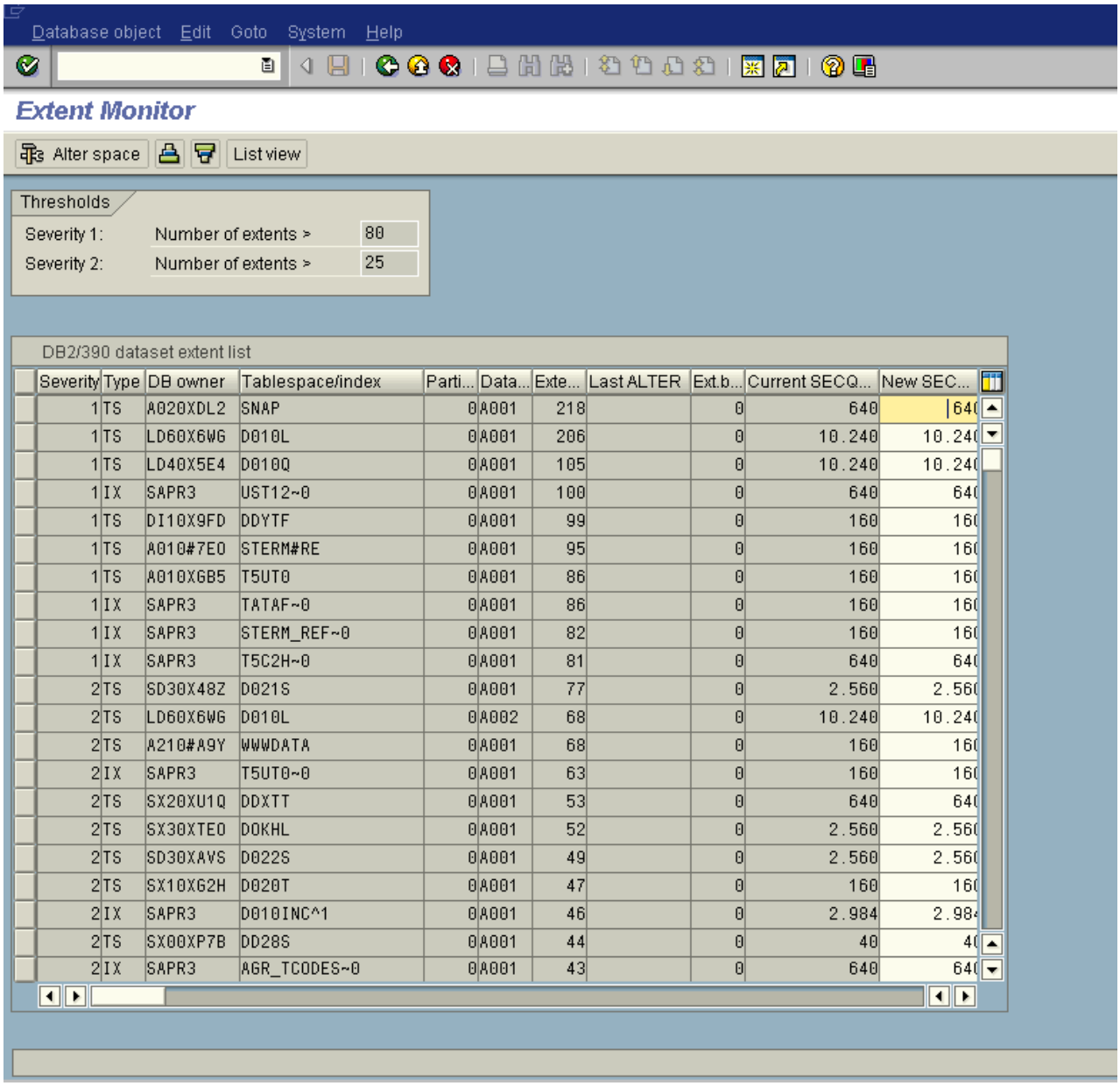
The objects are split into two levels of urgency, according to the number of extents:

- Severity 1:
  Number of extents > high threshold
  If the number of extents exceeds the high threshold, a reorganization will be required.
- Severity 2:
  high threshold > number of extents > low threshold
  The object is marked for observation. You may have to increase the secondary quantity.

The objects are specified using their name and the database (for tablespaces, type TS) or the creator (for indexes, type IX). For partitioned tablespaces or indexes the affected partition is identified, otherwise the *Partition* column contains a 0. For tablespaces or indexes that have several data sets, the data set in question is identified in the Dataset column by its last qualifier.

In Figure 38 on page 97, we show the output of the Extent Monitor after an update of the Table and Indexes monitor data

.



Severity	Type	DB owner	Tablespace/index	Parti...	Data...	Exte...	Last ALTER	Ext.b...	Current SECQ...	New SEC...
1	TS	A020XDL2	SNAP	0	A001	218		0	640	640
1	TS	LD60X6WG	D010L	0	A001	206		0	10.240	10.240
1	TS	LD40X5E4	D010Q	0	A001	105		0	10.240	10.240
1	IX	SAPR3	UST12~0	0	A001	100		0	640	640
1	TS	DI10X9FD	DDYTF	0	A001	99		0	160	160
1	TS	A010#7EO	STERM#RE	0	A001	95		0	160	160
1	TS	A010XGB5	T5UT0	0	A001	86		0	160	160
1	IX	SAPR3	TATAF~0	0	A001	86		0	160	160
1	IX	SAPR3	STERM_REF~0	0	A001	82		0	160	160
1	IX	SAPR3	T5C2H~0	0	A001	81		0	640	640
2	TS	SD30X48Z	D021S	0	A001	77		0	2.560	2.560
2	TS	LD60X6WG	D010L	0	A002	68		0	10.240	10.240
2	TS	A210#A9Y	WWWDATA	0	A001	68		0	160	160
2	IX	SAPR3	T5UT0~0	0	A001	63		0	160	160
2	TS	SX20XU1Q	DDXTT	0	A001	53		0	640	640
2	TS	SX30XTEO	DOKHL	0	A001	52		0	2.560	2.560
2	TS	SD30XAVS	D022S	0	A001	49		0	2.560	2.560
2	TS	SX10XG2H	D020T	0	A001	47		0	160	160
2	IX	SAPR3	D010INC^1	0	A001	46		0	2.984	2.984
2	TS	SX00XP7B	DD28S	0	A001	44		0	40	40
2	IX	SAPR3	AGR_TCODES~0	0	A001	43		0	640	640

Figure 38. DB02 Transaction; Extent Monitor

Note the tablespace A020XDL2.SNAP with severity 1 at the head of the list.

A quick verification using ISPF screens shows that the Data Set has 218 extensions:

```
Menu  Options  View  Utilities  Compilers  Help
sssssssssssssssssssssssssssssssssssssssssssssssssssssssssssssssssssssssssssssssss
DSLIST - Data Sets Matching DB2V610K.DSNDBD.A020XDL2.SNAP.*         Row 1 of 1
Command ===>                                                   Scroll ===> PAGE

Command - Enter "/" to select action                Tracks %Used XT  Device
------------------------------------------------------------------------------
         DB2V610K.DSNDBD.A020XDL2.SNAP.I0001.A001       3039    ? 218  3390
***************************** End of Data Set list ****************************
```

From the Extent Monitor, you can change the secondary quantity. To change this value, select the corresponding lines in the table, enter the required size in the field New SEC..., and choose **Alter space**. If successful the column Current SECQ... shows the new (changed) value, the column last ALTER contains the date of the last change (directly after the last change, that is, the current date). The column Ext.b... shows the number of extents before the last change (initial: 0), so that you can see how quickly the tablespace or index occupies new extents after ALTER.

Volume Free Space

You can use this function to see how much free space remaining on a volume.

1. Call transaction DB02.
2. Choose *Volume freespace.*
3. Enter the required volume on the following screen.

An OS/390 job is generated and executed under your user ID. This obtains the information directly from the OS/390 host.

After executing the job, whose runtime depends on the size of the volume, you will see the free space on the volume in cylinders, as well as the number of free data set control blocks (DSCBS) on the volume.

To see a display of all volumes relevant to the SAP R/3 system, that is, all volumes on which the SAP R/3 system keeps data sets, choose Help (or press F4). See Figure 39 on page 99. If these are assigned to a storage group created by the user SAPR3 the name of the storage group and the creator are displayed; otherwise these fields contain an asterisk. If one volume belongs to several storage groups, it will also appear several times in the list.

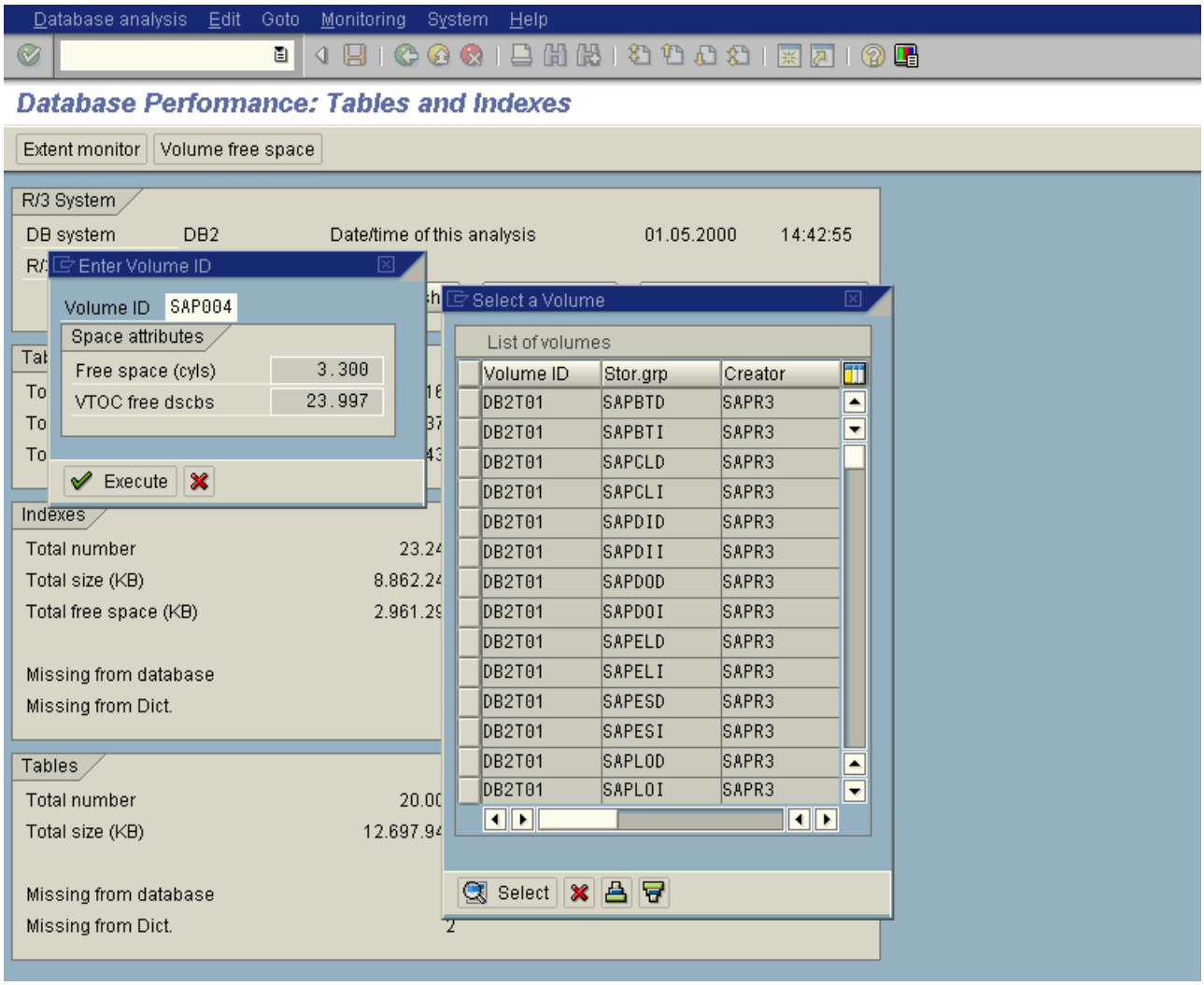


Figure 39. Pop-up menus for DB02 transaction

4.2.2 DB2 UDB for OS/390 and storage management

A fundamental question requires an answer when preparing to use DB2 UDB for OS/390 as the database server for SAP R/3: *Who is going to be responsible for the auxiliary storage management?*

On one hand, database administrators worry about the separation of performance-critical data sets, like data from indexes, data from log, copies of log and BSDS, spreading workfiles. They use multiple storage groups and are careful with the association of volumes to storage groups.

On the other hand, the need for automated methods ways to administer DASD space is growing, as new requirements are imposed by the new layout of the database (with an explosion in the number of the objects), and the extensions to data warehousing (which needs very large tables).

Furthermore, as processors and disk control units provide more capacity and more memory, DB2 UDB for OS/390 exploits its larger buffer pools as a second level of cache for I/O execution, reducing the I/O frequency and making it mostly

asynchronous. This implies that the criticality of dataset placement is greatly reduced.

DFSMSdfp's storage management subsystem (SMS) can be used to manage DB2 DASD data sets. The purpose of SMS is to automate as much as possible the management of physical storage by centralizing control, automating tasks, and providing interactive controls for system administrators. SMS can reduce users' needs to be concerned about physical details of performance, space, and device management.

As explained in 4.1, “Altering the database layout” on page 83, one of the requirements for the final layout of the SAP R/3 V4.6 on DB2 UDB for OS/390 database is to include storage-relevant SAP R/3 information in the OS/390 data set name of tablespaces and indexspaces. This enables the possibility of developing SMS management policies coherently with the expected SAP R/3 utilization of the tables.

4.2.2.1 Benefits of DFSMS for availability of auxiliary space

The following benefits have been extracted from the redbook: *Storage Management with DB2 for OS/390*, SG54-5462. In this book the reader can find documentation about how to develop a project for migrating DB2 data sets to be SMS managed.

Using DFSMS, the DB2 administrator gains the following benefits:

- Improved allocation control
  Free space requirements can be set using SMS across a set of disk volumes. Sufficient levels of free space can be guaranteed to avoid space abends. The system automatically places data on a volume containing adequate freespace.
- Automated disk space management
  SMS has a facility that automatically reclaims space which formerly was allocated to old and unused data sets. Policies can be defined that determine how long an unused data set is allowed to reside on level 0 volumes (active data). It is important to use this option very carefully, since allowing DB2 access to migrated tablespaces may impose unacceptable performance levels.
- Improved data availability management
  With SMS, different backup requirements can be provided for data residing on the same primary volume. Therefore, all data on a single volume can be treated independently. The HSM component can be used to automatically back up data. The ABARS facility can be used to group data sets into logical components, so that the group is backed up at the same time, allowing for recovery of an application.
- Simplified data movement
  SMS permits the movement of data to new volumes without the necessity for users to amend their JCL. Because users in DFSMS environment do not need to specify the unit name and volume which contain their data, it does not matter to them if their data resides on a specific volume or device type. This allows the replacement of old devices with minimum intervention from the user.

4.2.2.2 Summary

A detailed analysis of the different types of DB2 UDB for OS/390 data sets shows that DFSMS can automatically manage all of the data sets DB2 UDB for OS/390 uses and requires. The new layout and naming convention of the SAP R/3 database gives to the storage administrator the full range of possibilities to implement policies which meet the space requirements of the SAP R/3 on DB2 UDB for OS/390. All in all, this can be seen as an important step to provide the system with increased levels of availability and eliminate points of failure related to space management.

4.2.3 New function: Deferring the definition of data sets

Starting with release 4.6C of SAP R/3, the installation and upgrade processes create tablespaces with the DEFINE=NO option. Using the option means that the underlying data sets are not created until the first row is inserted into the corresponding table. For most SAP R/3 installations. a large number of tables are empty (for example. those tables used in applications that do not apply to this enterprise) which means that a significant number of data sets are not created. This is beneficial in reducing DASD space utilization; it also simplifies many database administration tasks.

The objects created with DEFINE=NO are fully supported by all the functions within SAP R/3. Potential problems may arise if you use third party tools that do not have support for this new kind of object. Check with your tool provider to determine if the tools implement the necessary support.

The remaining portion of this section deals with implementing DEFINE=NO on SAP R/3 releases prior to 4.6C.

During the installation of the SAP R/3 system, it is necessary to create the whole SAP R/3 database with the complete set of objects; tablespaces, tables, indexes and data sets. Many of the tables in the SAP R/3 database stand empty depending on which SAP R/3 modules are active. Even if modules are not used, there is always an underlying definition of the associated data sets, since it is always possible to activate the modules.

As the number of objects has increased gradually in recent versions of SAP R/3, it could be a waste of space and effort to deal with a number of data sets associated to not-used objects. Also massive administrative process, that may have impact on the availability of the database, can be delayed by the presence of this not-used data sets.

DB2 UDB for OS/390 (through APAR PQ30999) provides a new function to allow users to create tablespaces or index spaces with an option to defer the physical creation of underlying VSAM data sets until the very first write. (such as through an SQL insert or the use of a load utility).

4.2.3.1 How the DEFINE NO option works.

The new option is called DEFINE NO and it is included in the CREATE TABLESPACE and CREATE INDEX SQL statement. Using this option, a newly defined tablespace or index will have a catalog entry, but is considered empty when accessed by SELECT or FETCH operations (SQLCODE +100). A value of -1 in the SPACE column of catalog tables marks the tablespace or index as ‘undefined’.

Following we list usability features related to this option.

- At the first write, DB2 resets the ‘undefined’ status in the catalog and creates the underlying VSAM data set to allow the write operation.
- The default attribute is DEFINE YES which means to define VSAM datasets at creation time.
- The DEFINE NO option is only applicable for stogroup defined datasets (DB2 will define and manage the data set). The option is ignored for non-stogroup defined data or LOB table spaces.
- Workfile page sets cannot use this option.
- For the creation of indexes, if the associated table the index to be created for is already defined and has data inside, the DEFINE NO option will be ignored. The option will also be ignored for AUX INDEX.
- In general:
  - For utilities which read objects with row processing, these objects will appear the same as if the data set exists with no rows or keys.
  - For utilities that write to objects, the definition will be done implicitly on behalf of the utility.
- Currently, it is possible to run LOAD REPLACE and LOAD RESUME YES|NO on undefined objects. LOAD utilities invoke the reset function to create the undefined objects.
- For utilities that require a check of the OBDs during initialization: COPY, RECOVER, QUIESCE, RUNSTATS, REPORT, MODIFY, REPAIR (except REPAIR DBD), CHECK, REORG, REBUILD and MERGECOPY, a new message of DSNU185I will be issued and the utilities will end with a return code of 8 if any target object is undefined.
- A value of -1 in the SPACE column in SYSTABLEPART and SYSINDEXPART indicates that the data set is deferred for the creation of underlying VSAM data sets. Apart from that there are no changes to the catalog, and the meaning of columns is the same than before this option was supported.
- For compatibility of this option with OEM software and earlier releases of DB2 UDB for OS/390 please consult the APAR PQ30999 documentation. Also refer to APARs PQ32213, PQ34030, PQ34386, PQ34592 and PQ34029.

4.2.3.2 Advantages of DEFINE NO

We see the following advantages for using these option:

- Faster installation of the SAP R/3 system.
  Once this option is included in the DDL generated by the SAP R/3 installation tool that generates all the objects (this occurred in SAP R/3 4.6c), installation time will be reduced in the phase of object definition. Also such functions as REORG, RUNSTATS etc.... that are done in installation will require less time.
- Simplified DBA operations.
- Better DASD utilization, not only for the data, but also for backup copies.
- Management and virtual storage relief for the subsystem on DD limits.
- Improved performance of DDL statements.

All of these advantages can been seen as beneficial for increasing the availability of the SAP R/3 database. They decrease the administration time and simplify operations, which ultimately implies better disposal of data.

4.2.3.3 Considerations for using DEFINE NO

In order to get an idea of the amount of space that can be saved using the option we generated the following tablespace and index queries to the catalog:

For tablespaces:

```
SELECT SUM(SPACE) FROM SYSIBM.SYSTABLESPACE
WHERE NAME IN (
SELECT TSNAME FROM SYSIBM.SYSTABLES
WHERE CREATOR = 'SAPR3' AND NPAGES = 0);
```

And for indexes:

```
SELECT SUM(SPACE) FROM SYSIBM.SYSINDEXES
WHERE TBNAME IN (
SELECT NAME
FROM SYSIBM.SYSTABLES
WHERE  CREATOR ='SAPR3' AND NPAGES = 0);
```

For the first query we obtained 7.2 GB of allocated space for 10199 out of 24105 tables without rows. For the second query we obtained 0.7 GB of allocated space. As our database had 25 GB of allocated space, 31% of the total space could be saved using DEFINE NO.

This option will be fully usable when it becomes part of the DDL generated by the installation tool R3setup, which will happen in SAP R/3 V4.6C. As it is impossible to use an ALTER statement to activate the option, the only possible way of using it during run time is deleting and recreating the associated table.

If there are no rows in a table there is no problem deleting and recreating it, using the database utility transaction ‘SE14’, which keeps all the definitions in the SAP R/3 dictionary.

In our SAP R/3 system, there is no support for this option in any of the storage attributes menus. An way of implementing the option would be to capture the generation SQL with an SQL trace (transaction ‘ST05’), and later, with the database manager interactive panels, DROP the tablespace and recreate all the related objects. You would incorporate the parameter when you perform this recreate step.

4.3 Data reorganization

This section describes considerations and impact on availability of the SAP R/3 database server for doing a tablespace reorganization, the impact of reorganization on users, and the benefits of the DB2 UDB for OS/390 online reorganization.

4.3.1 Considerations for doing tablespace reorganizations

The following is a list of reasons that could necessitate a tablespace reorganization:

- SQL queries that use sequential processing of the tablespace begin to experience degraded performance.

  This situation typically shows up during batch sequential processing of the tablespace.

- Free space is exhausted, causing the sequence of the data to degrade.

  This tends to be the root cause of the degraded performance of the sequential processes previously discussed.

- The underlying VSAM file that contains the DB2 tablespace reaches the 119 extent limit.
- The DASD volume where the VSAM file resides that contains the DB2 tablespace runs out of space.

Although there are many considerations for doing a tablespace reorganization, one major consideration is how the tablespace is accessed and processed. If you are doing direct processing, such as fully qualified SQL select of a single row using a unique key (normal case with SAP R/3), then there should be little need to do a tablespace reorganization. The physical order of the data is unimportant when doing a fully-qualified single row, direct processing. However, if you are doing sequential processing, then the physical order of the data may influence performance. In that case, a tablespace reorganization will most likely improve performance.

The traditional concern about reorganization is the unavailability of the whole set of data while the utility REORG is running. This fact has forced many installations to find time windows of no application activity during which all these data maintenance tasks can be done. The availability requirements of such installations are increasing since 24 x 7 operation is becoming more standard. It is becoming more difficult to get time windows broad enough for the amount of administrative activity necessary to keep the system performing adequately.

DB2 V5 introduced online REORG, which provided a great improvement in data availability for applications during a REORG. The keyword SHRLEVEL was added to the control statement to allow the setting of different degrees of concurrency between the utility and other work running at the same time.

- SHRLEVEL NONE
  This is the default value. It specifies that reorganization operates by unloading from the area being reorganized (while applications can read but cannot write to the area), reloading into that area (while applications have no access), and then allowing read/write access again.

- SHRLEVEL REFERENCE
  This specifies that reorganization operates by unloading from the area being reorganized, reloading into a shadow copy of the area (while applications can read but not write to the original copy), switching applications' future access from the original copy to the shadow copy by exchanging the names of the data sets, and then allowing read/write access again.
- SHRLEVEL CHANGE
  This specifies that reorganization operates by unloading from the area being reorganized, reloading into a shadow copy of that area (while applications have read/write access to the original copy of the area), applying the log of the original copy to the shadow copy (while applications can read and usually write to the original copy), switching applications' future access from the original copy to the shadow copy by exchanging the names of the data sets, and then allowing read/write access again. This last option is what is called online REORG.

As it is explained in the white paper: *Implementing Online Reorg in a Production Environment, by Craig Friske (*available in the Web site `http://www-4.ibm.com/software/data/db2/os390/pdf/oreorg.pd`f) using this new function requires planning and a careful evaluation almost at an object level, for it introduces complexity that didn't previously exist.

The trade-off between the elapsed time of no-allowed access to the data imposed by the utility and the tolerance of applications to deal with this elapsed time is the key factor to consider when planning for online REORG.

In order to evaluate the data unavailability period, it is important to understand the phases of the online REORG.

4.3.2 Phases of the Online Reorg Utility

The detailed description of the phases of Online Reorg can be found in *DB2 UDB for OS/390 Utility Guide and Reference*, SC26-9015.

- UNLOAD
  Unloads data to be reorganized into a sequential data set (SYSREC). Data is sorted in clustering sequence if a clustering index exists and either SORTDATA or SHRLEVEL CHANGE is specified.
- RELOAD
  Reloads the data from the unload phase and extracts keys to build indexes. Specifying SORTKEYS allows the extracted keys to be passed to sort in memory and sorted in parallel with the reload phase. This reduces elapsed time by avoiding I/O, in addition to reduction gained by sorting in parallel. For SHRLEVEL CHANGE, SORTKEYS is the default.

  You can use the COPYDDN (and optionally the RECOVERYDDN) option to take an inline image copy (and optionally a remote copy) during the reload phase. This is a choice with SHRLEVEL NONE and mandatory in SHRLEVEL REFERENCE and SHRLEVEL CHANGE.
- SORT
  Sorts the extracted index keys.
- BUILD
  Builds indexes.

If REORG TABLESPACE PART SHRLEVEL NONE is specified, this phase also corrects the logical partition of any nonpartitioning index.

- LOG
  This phase is executed only if SHRLEVEL CHANGE is specified. During this phase, log records are applied to the shadow copy. This step executes iteratively, each iteration processing a sequence of log records. During the last iteration, applications have only read access to the original copy.

  Changed pages are appended to the inline full-image copy taken during the reload phase by two incremental copies taken during the last iteration - one just before writers are drained, and one just after.

- SWITCH
  Executes only if SHRLEVEL REFERENCE or SHRLEVEL CHANGE is specified.

  During this phase, DB2 switches access to shadow copies of the table and index spaces. This switching is done by IDCAMS rename, and there is no access to the data during this phase. The more physical data sets IDCAMS has to process, the longer the data remains unavailable to applications.

  If the object being reorganized is a partition of a partitioned tablespace, the actual partition of the table space and of the partitioning index are renamed. Logical partitions of a nonpartitioning index are not processed.

- BUILD2
  Executes only if SHRLEVEL REFERENCE or SHRLEVEL CHANGE is specified and the object that is being reorganized is a partition of a partitioned tablespace with one or more nonpartitioning indexes.

  During this phase, DB2 corrects logical partitions in nonpartitioning indexes rather than replacing them. This method is used to allow concurrent reorganizations of several partitions. It uses the shadow created for the logical part of the nonpartitioning index to correct the RIDs in the original copy of the nonpartitioning index. When completed, the shadow of the logical partition is deleted.

  During this operation, no access, other than read only is allowed to the logical partition being corrected in the nonpartitioning index. Read and update are allowed in other logical partitions (those not being reorganized), however, insert and delete are permitted only for Type 2 indexes.

  This is the most time-consuming operation with limited application access to the data. As a result, trying to optimize the working conditions for this process seems worthwhile. The option PIECESIZE, introduced in DB2 V5 in the CREATE INDEX statement, allows you to divide nonpartitioning indexes into smaller pieces and spread these pieces evenly on the DASD available. This avoids I/O contention and thereby reduces the elapsed time for the corrections, which is the time when the reorganized object is unavailable to applications.

Following is the output of the utility which executed with SHARELEVEL CHANGE without any concurrent workload. You can see the time spent in each phase.

```
OUTPUT START FOR UTILITY, UTILID = RUNLD
 REORG TABLESPACE LD40X5E4.D010Q LOG NO COPYDDN(SYSCOPY) SORTDEVT SYSDA SORTNUM 4 SHRLEVEL
 2000-5-4-11.30.00 MAPPINGTABLE SAPRES4.MAP_TBL MAXRO 240 LONGLOG DRAIN DELAY 900
 UNLOAD PHASE STATISTICS - NUMBER OF RECORDS UNLOADED=92521 FOR TABLESPACE LD40X5E4.D010Q
 UNLOAD PHASE COMPLETE, ELAPSED TIME=00:02:56
 INDEXES WILL BE BUILT IN PARALLEL, NUMBER OF TASKS = 4
 (RE)LOAD PHASE STATISTICS - NUMBER OF RECORDS=92521 FOR TABLE SAPR3.D010Q
 (RE)LOAD PHASE STATISTICS - NUMBER OF INPUT RECORDS PROCESSED=92521
 (RE)LOAD PHASE COMPLETE, ELAPSED TIME=00:04:10
 SORTBLD PHASE STATISTICS - NUMBER OF KEYS=92521 FOR INDEX SAPR3.D010Q^0
 SORTBLD PHASE STATISTICS - NUMBER OF KEYS=92522 FOR INDEX SAPRES4.XMAP_TBL
 SORTBLD PHASE STATISTICS. NUMBER OF INDEXES = 2
 SORTBLD PHASE COMPLETE, ELAPSED TIME = 00:00:12
 LOG PHASE STATISTICS. NUMBER OF ITERATIONS = 1, NUMBER OF LOG RECORDS = 0
 LOG PHASE COMPLETE, ELAPSED TIME = 00:00:01
 COPY PROCESSED FOR TABLESPACE LD40X5E4.D010Q
 NUMBER OF PAGES=37621
 AVERAGE PERCENT FREE SPACE PER PAGE = 19.08
 PERCENT OF CHANGED PAGES =100.00
 ELAPSED TIME=00:04:23
 SWITCH PHASE COMPLETE, ELAPSED TIME = 00:00:05
 DB2 IMAGE COPY SUCCESSFUL FOR TABLESPACE LD40X5E4.D010Q
 UTILITY EXECUTION COMPLETE, HIGHEST RETURN CODE=0
```

Our next testing was to perform the REORG with concurrent workload activity. This activity was created by a simulator tool which generates SAP R/3 transactions through RFC. The sequence of the activity generated by the tool against the table was: select - update - commit - select - insert - commit. This sequence was repeated for one hour, so the update rate was very high.

While the tool was running, we submitted the utility job. Following we present the JCL:

```
//DSNTICX EXEC PGM=DSNUTILB,PARM='DB2K,RUNLD2',REGION=1024K
//SYSCOPY  DD SPACE=(CYL,(50,20),,,ROUND),
//            VOL=SER=(SAP004),DISP=(NEW,CATLG,DELETE),
//            DSN=DB2V610K.COPY.ZPK2TB0
//SYSREC   DD  DSN=DB2V610K.REORG.ZPK2TB0,DISP=(NEW,CATLG,DELETE),
//             UNIT=SYSDA,SPACE=(CYL,(5,20)),VOL=SER=SAP012
//SYSPRINT DD  SYSOUT=*
//UTPRINT  DD  SYSOUT=*
//SYSPUNCH DD  SYSOUT=*
//SYSIN    DD  *
  REORG TABLESPACE A130X1F2.ZPK2TB0 LOG NO COPYDDN(SYSCOPY)
     SORTDEVT SYSDA SORTNUM 4
     SHRLEVEL CHANGE
     DEADLINE 2000-5-7-11.30.00
     TIMEOUT TERM
     MAPPINGTABLE
   SAPRES4.MAP_TBL MAXRO 10 LONGLOG DRAIN DELAY 0
//*
```

We took several displays to check the concurrency of the processes. As can be seen in this display both processes were running in different data sharing members:

```
DSNT360I =DBK2 ***********************************
 DSNT361I =DBK2 *  DISPLAY DATABASE SUMMARY
                *    GLOBAL LOCKS
 DSNT360I =DBK2 ***********************************
 DSNT362I =DBK2     DATABASE = A130X1F2  STATUS = RW
                    DBD LENGTH = 4028
 DSNT397I =DBK2
 NAME     TYPE PART STATUS             CONNID   CORRID       LOCKINFO
 -------- ---- ---- ------------------ -------- ------------ ---------
 ZPK2TB0  TS        RW,UTUT            RRSAF    010001010002 H-IX,S,C
     -                  MEMBER NAME DBK1
 ZPK2TB0  TS        RW,UTUT            RRSAF    010001010002 W-S,W,01
     -                  MEMBER NAME DBK1
 ZPK2TB0  TS        RW,UTUT                                  H-IS,PP,I
     -                  MEMBER NAME DBK1
 ZPK2TB0  TS        RW,UTUT                                  H-SIX,PP,I
     -                  MEMBER NAME DBK2     (CO)
 ZPK2TB0  TS        RW,UTUT            UTILITY  SAPRES4X     H-IX,W,A
     -                  MEMBER NAME DBK2
 3        TB                           RRSAF    010001010002 H-IX,T,C
     -                  MEMBER NAME DBK1
 3        TB                           UTILITY  SAPRES4X     H-IX,W,A
     -                  MEMBER NAME DBK2
 ***
 ZPK2TB00 IX        RW,UTUT            UTILITY  SAPRES4X     H-X,R,A
     -                  MEMBER NAME DBK2
 ZPK2TB00 IX        RW,UTUT            UTILITY  SAPRES4X     H-X,C,A
     -                  MEMBER NAME DBK2
 ZPK2TB00 IX        RW,UTUT            UTILITY  SAPRES4X     H-X,W,A
     -                  MEMBER NAME DBK2
 ******* DISPLAY OF DATABASE A130X1F2 ENDED      **********************
 DSN9022I =DBK2 DSNTDDIS 'DISPLAY DATABASE' NORMAL COMPLETION
```

In the switch phase, the utility tried to drain the update activity. This took quite a long time. Finally, the utility finished successfully.

```
OUTPUT START FOR UTILITY, UTILID = RUNLD2
 REORG TABLESPACE A130X1F2.ZPK2TB0 LOG NO COPYDDN(SYSCOPY) SORTDEVT SYSDA SORTNUM 4 SHRLEVEL
 CHANGE DEADLINE 2000-5-7-11.30.00 TIMEOUT TERM MAPPINGTABLE SAPRES4.MAP_TBL
 MAXRO 10 LONGLOG DRAIN DELAY 0
 DSNU301I =DBK2 DSNURFIT - KEYWORD 'SORTDATA' SPECIFIED AND/OR KEYWORD
 'NOSYSREC' WAS SPECIFIED BUT NO CLUSTERING INDEX EXISTS, KEYWORD IS IGNORED
 UNLOAD PHASE STATISTICS - NUMBER OF RECORDS UNLOADED=10000 FOR TABLESPACE A130X1F2.ZPK2TB0
 UNLOAD PHASE COMPLETE, ELAPSED TIME=00:00:01
 INDEXES WILL BE BUILT IN PARALLEL, NUMBER OF TASKS = 4
 (RE)LOAD PHASE STATISTICS - NUMBER OF RECORDS=10000 FOR TABLE SAPR3.ZPK2TB002
 (RE)LOAD PHASE STATISTICS - NUMBER OF INPUT RECORDS PROCESSED=10000
 (RE)LOAD PHASE COMPLETE, ELAPSED TIME=00:00:09
  SORTBLD PHASE STATISTICS - NUMBER OF KEYS=10000 FOR INDEX SAPR3.ZPK2TB002~0
  SORTBLD PHASE STATISTICS - NUMBER OF KEYS=10000 FOR INDEX SAPRES4.XMAP_TBL
  SORTBLD PHASE STATISTICS. NUMBER OF INDEXES = 2
  SORTBLD PHASE COMPLETE, ELAPSED TIME = 00:00:01
 LOG PHASE STATISTICS. NUMBER OF ITERATIONS = 1, NUMBER OF LOG RECORDS = 50
 LOG PHASE COMPLETE, ELAPSED TIME = 00:00:00
 COPY PROCESSED FOR TABLESPACE A130X1F2.ZPK2TB0
 NUMBER OF PAGES=2421
 AVERAGE PERCENT FREE SPACE PER PAGE = 22.83
 PERCENT OF CHANGED PAGES =100.00
 ELAPSED TIME=00:00:10
 SWITCH PHASE COMPLETE, ELAPSED TIME = 00:10:16
 DB2 IMAGE COPY SUCCESSFUL FOR TABLESPACE A130X1F2.ZPK2TB0
 UTILITY EXECUTION COMPLETE, HIGHEST RETURN CODE=0
```

In this test, due to the high update rate, the utility was not able to drain the updaters and the readers and the unavailability time (616 seconds) was higher than the timeout value (600 seconds) so the transaction failed with SQLCODE -911. It is possible to set the online reorg parameters so that the transaction gets higher priority than the utility.

4.3.3 Recommendations for using online REORG

Following are some considerations to keep in mind when moving reorganizations outside the batch maintenance window that aid in minimizing impact on normal activity.

- Run the Online Reorg Utility when dialog processing is as lowest. Even the busiest SAP R/3 system has a period of lowest dialog activity.
- Before running the Online Reorg Utility verify that there is no outstanding update activity on the objects.
- Run the Online Reorg Utility when there is *no* SAP R/3 Batch running, unless the batch programs issue COMMITs every 100 or so records. Read-only Batch work holds pages of the database until you issue a COMMIT. Without the COMMIT, the SAP R/3 transaction could die with a SQLCODE -911. (Remember the programming consideration about reestablishing the cursor after the COMMIT.)
- Use the following settings for the new parameters of the DB2 UDB for OS/390 Online Reorg utility. The DRAIN parameter was introduced with APAR/PTF PQ22944/UQ27129 in DB2 UDB for OS/390: See *DB2 UDB for OS/390 Utility Guide and Reference,* SC26-9015.
  - MAXRO 10
    This parameter specifies in number of seconds, the maximum amount of

time for the last iteration of log processing. During that iteration, the SAP R/3 system has only read access to the tablespace or index.

- TIMEOUT TERM
  This parameter specifies the action to be taken if the REORG utility gets a time out condition while trying to drain objects in either the LOG or SWITCH phases. The option TERM means that when the time-out condition occurs DB2 issue an implicit TERM UTILITY command, causing the utility to end with a return code 8, issuing the DSNU590I and DSNU170I messages, and leaving the object in a RW state.
- DRAIN ALL
  This parameter specifies the drain behavior at the end of the log phase after the MAXRO threshold is reached and when the last iteration of the log is to be applied. And the option ALL specifies that DB2 drain all readers and writers during the log phase, after the MAXRO threshold is reached.
- LONGLOG DRAIN
  This option specifies the action DB2 UDB for OS/390 performs after sending a message to the console, if the number of records that the next iteration of the log process is not sufficiently lower than the number of that the previous iterations processed. DRAIN specifies that DB2 drains the write claim class after the delay specified by the DELAY parameter.
- DELAY 0

  Specifies the minimum interval between the time that DB2 sends the LONGLOG message to the console and the time REORG performs the action specified by the LONGLOG parameter.

4.3.4 Summary

The use of online REORG is a necessary and unavoidable step if the availability requirements of a SAP R/3 installation tends to 24 x 7. Implementing this solution requires careful planning and evaluation to ensure that it is done in a controlled manner to avoid unpleasant surprises. There must be an agreement between the time that the utility needs to complete, recognizing the short bursts when data is unavailable, and the affordable time that the SAP R/3 system can admit. The approach must be directed almost at a table level, and for this the SAP R/3 transaction “ST10” may become an important help.

4.4 Catalog statistics

In this section we explain what RUNSTATS does and when its use is necessary. We discuss the different situations that imply maintenance of catalog statistics, and how RUNSTATS can affect availability. We introduce In-line RUNSTATS and we describe the options that optimize availability of data. You should also reference the OSS Notes and IBM APARs that are relevant to this topic.

4.4.1 Collecting statistical data for the optimizer

The RUNSTATS utility updates the DB2 catalog tables with statistical information about tables, tablespaces, and indexes. There are two classes of statistics, those used by DB2 to determine the access path of SQL statement, and those used by a database administrator to assess the status of a particular tablespace or index.

By updating the DB2 catalog, RUNSTATS provides DB2 with current information from which to perform path selection. A massive inserts process can change significantly table's size and the cardinalities of its columns. If RUNSTATS has not been run after such an activity, the input into DB2 Optimizer access path selection is outdated and that can result in an access path that is less than optimal.

The impact on the availability of the data referred when this utility needs to be run depends on the choice for the SHARELEVEL option. Using SHARELEVEL REFERENCE all the data used during the collection of statistics is committed, but only read activity is allowed. Using SHARELEVEL CHANGE not all the data used is committed but there is concurrency with read and write activity. This is the normal and recommended option.

The cost of running this utility is mainly that of CPU consumption. For stable tables there is no need to update their statistics often. It is important, therefore, not to use this utility indiscriminately on all tables of the SAP R/3 database.

SAP R/3 provides a very useful tool for monitoring wokload on a per-table basis: transaction ST10. The output of the transaction is a list of tables and the operations statistics for the tables, such as the number of inserts, updates and deletes for various time intervals: since startup, today, previous day, this week, previous week, this month, previous month. You can use these statistics to assess which tables need RUNSTATS; in general the tables will be those with a considerable number of changes (relative to the table's size) since the last time RUNSTATS was done.

To ensure that information in the catalog is current, invoke RUNSTATS in the following situations:

- As soon as convenient, for the tables identified by ST10 as considerably changed.
- After the initial load, migration and upgrade. There is a separate step in the R3setup procedure where RUNSTATS is performed for all the tables in the system. This is the time when RUNSTATS should be run for catalog tables as well, because of a large number of new database objects.
- If Batch Input is used to import data into the system, it is important to do RUNSTATS not only after, but also at least once during the process (in the first quarter, not too soon after the start). Since batch Input includes queries as well as inserts, the queries need current statistics in order to use optimal access paths.
- For tablespaces and indexes that have just been reorganized.
- For tablespaces and indexes that have just been recovered.
- After loading a tablespace, and before binding application plans and packages that access the tablespace.
- After creating an index with a CREATE INDEX statement, in order to update catalog statistics related to the new index.

For a full description of the RUNSTATS options and job specifications, see DB2 V6 Utility Guide and Reference. For a description of RUNSTATS in the SAP R/3 environment, see OSS Note 0113008.

4.4.2 Improving the performance of RUNSTATS

The following hints will improve the performance of RUNSTATS:

- You can specify the STATISTICS keyword in LOAD, REBUILD INDEX, and REORG utility statements, which results in updated table space or index space catalog statistics for the objects the utility was run on.
- Another method of improving RUNSTATS performance is to specify the SAMPLE option on tablespaces that were defined with the LARGE option, which reduces the number of rows sampled for statistics.
- When you run RUNSTATS concurrently against partitions of a partitioned table space or index, the sum of the processor time for the concurrent jobs will be roughly equivalent to the processor time it takes to run a single RUNSTATS job against the entire table space or index. However, the total elapsed time for the concurrent jobs can be significantly less than when you run RUNSTATS against an entire table space or index.
- When requesting non indexed column statistics, provide a list of columns that might be used in queries as search conditions in a WHERE clause.

Collecting statistics on all columns of a table is costly and might not be necessary.

4.4.3 The volatile tables access path problem

If RUNSTATS is executed on a table that happened to be empty or very small (occupying just few pages), the statistics collected at that time can be very misleading if the table subsequently gets heavily inserted. This can often happen on the tables such update protocol tables (VBDATA, VBMOD, VBHDR) or any kind of queue tables in general. This can lead to serious performance problems like choosing wrong access path or heavy lock contention including deadlocks.

A current solution for this problem is *manual catalog updates* which is used in other situations (cluster tables, archived tables, batch input, etc....). See OSS Note 0113008 for details about catalog updates. But in the volatile tables case this solution presents the following problems:

- Volatile table are not always known.
- Updates are complex from an operational point of view.
- They produce lock contention on the catalog.

As this last point may produce concern about availability of DB2 catalog data, there is a new option introduced in DB2 UDB for OS/390 that can help greatly in these kind of situations.

APAR PQ33429 of DB2 UDB for OS/390 introduced a new system parameter, NPGTHRSH, which value is taken into account during access path selection. For a given table, if NPAGES is less than the NPGTHRSH value, an index access for the table will be preferred over a tablespace scan.

It is strongly recommended to use the new parameter. It will not only ensure that very serious performance problems are avoided, but also remove the need for manual catalog updates that are supposed to prevent tablespaces scan on large tables.

The parameter can take the following values:

- A value of ‘0’ means to use the standard cost optimization (default).
- A value of ‘n’ means that if NPAGES is less than n, then prefer index access. The recommended value is 10.
- A value of ‘-1’ means that always prefer index access (not recommended).

Although there will be situations in which applying this option could lead to a non-optimal access path when there are so few rows in a table that a tablespace scan is better than index access, the advantages in the opposite situation overcome the disadvantages.

Chapter 5. Creating database backups

One of the challenges for an installation of DB2 UDB for OS/390 as a database server for SAP R/3 is to be able to produce database backups on a scheduled basis with no interruption in SAP R/3 services. In this chapter we examine alternatives for creating these backup copies and additionally discuss the procedures for restoring the database, or portions of it, from backup copies.

5.1 High Availability (HA) in an SAP R/3 environment

Increasingly, the demand for SAP R/3 systems to be constantly available can be in direct conflict with the requirement to take adequate backups for recovery purposes. As more and more businesses rely on SAP R/3 systems, their Service level agreements now request that the systems are available almost continuously. However, while High Availability can be measured in service hours, it must also relate to the need for good backup and recovery strategies.

These dual pressures seem to conflict as time spent taking backups eats into online availability; reducing the number of backups taken reduces the ability to recover the application in the event of failure.

Good backup and recovery strategies require processes and procedures that ensure an SAP R/3 database can be reinstated with minimal disruption to the service levels after any kind of hardware, software, operational or environmental error or outage. For a general overview of backup and recovery procedures, constantly updated with the latest information, readers should refer to *OSS note 83000*.

5.2 Backup and recovery procedures

Developing appropriate backup and recovery procedures, since they are a key factor in system availability and reliability, require clear processes that are well planned and practiced. In general, the recovery process consists of selecting a backup of the tablespace and applying all the changes that are recorded in the log from the time the backup was created to the restore time.

The backup and recovery procedures at a particular site have to be set up by the database administrators for each individual SAP R/3 database. These processes will depend on:

- System availability requirements
- SAP R/3 database size
- The rate of change to the SAP R/3 database
- The available hardware and software resources
- The tools used to produce the backups

The optimal backup process is a trade-off between its usage of resources, such as CPU, DASD and tapes, and any increased contention with other concurrent activities in the system on one hand, and the required speed of recovery on the other. The shorter the time needed to apply any changes from the logs, the faster the recovery.

Frequent backups of all data pages containing committed rows only, provide the fastest recoveries. However, this has a higher cost in resources and will impact the concurrent activity in the system, to an extent that may not be acceptable.

We will now consider the two main types of backup available for DB2 UDB for OS/390 systems, online and offline backups. For more detailed information refer to *DB2 UDB for OS/390 Administration Guide,* SC26-9003.

5.2.1 Online backups

An online backup copies tablespace data while the tablespace remains online, allowing concurrent reads or writes to it. Therefore, except for a small processing and disk access overhead, the online backup has no impact on the SAP R/3 System. As it can contain uncommitted data, such a backup alone is never enough for tablespace recovery. Therefore, DB2 UDB for OS/390 automatically applies the recovery log on top of the online backup, as necessary.

From a DB2 UDB for OS/390 perspective, an online backup is a sharelevel change full image copy of the tablespace. An important aspect of the online backup is an “incremental” online backup, which only backs up data pages that have changed since the last backup.

Incremental copies have FULL NO specified, and only record pages in the tablespace that have been changed since the last image copy. These incremental image copies can be combined with the full image copy using the MERGECOPY utility, or they can be processed by the RECOVERY utility when required.

The DB2 UDB for OS/390 COPY utility with the SHRLEVEL (CHANGE) option is an efficient tool for creating online backups. Other COPY options are briefly described here; for a full explanation of all parameters and usage requirements, refer to *DB2 UDB for OS/390 V6 Utility Guide and Reference,* SC26-9015.

FULL

Specifies whether a full (YES) or incremental (NO) image copy is to be created.

CHANGELIMIT

Allows you to let DB2 UDB for OS/390 decide whether to take a full or incremental image copy, depending on the number of pages that have been changed since the last image copy.

COPYDDN or RECOVERYDDN

Allows you to create up to four identical copies of the tablespace for DR purposes. COPYDDN is used for taking two local copies, and RECOVERYDDN creates remote copies.

To assist in minimizing the time necessary for recoveries, you should consider taking full image copies of the index spaces as well. This will mean that in a recovery situation you will not have to rebuild the index in the traditional manner, which on very large tablespaces could take considerable time.

However, to take advantage of this feature the indexes have to be defined with COPY YES. If they are not you will have to run an ALTER INDEX for each one specifying COPY YES.

Note: This is a feature of DB2 UDB for OS/390 version 6, and in older versions the RECOVER INDEX command had the same function as the new REBUILD INDEX command, so any JCL in which you wish to continue using the tablespace scan method of recovery must be replaced with REBUILD INDEX control statements.

It is good practice that for any SAP R/3, catalog and directory tablespaces online backups are regularly created. Producing these regular online backups should really be considered mandatory. How often the backup should be taken and whether it should be full or incremental is dependent on the rate of change for that tablespace. To make this decision you need to categorize the tables into heavily, moderately or lightly updated and the ST10 transaction can be used to help in this categorization by access pattern.

5.2.1.1 Creating online backups

For our relatively simple test we took an online backup of a tablespace and its index, so we could then recover them later illustrating the new RECOVER INDEX facilities of DB2 UDB for OS/390 version 6.

The following illustration shows the example JCL required to create the online backup of the small tablespaces:

```
//SAPRES1C JOB (999,POK),'NICK VALLANCE',CLASS=A,MSGLEVEL=(1,1),
// MSGCLASS=T,NOTIFY=&SYSUID,TIME=1440,REGION=0M
/*JOBPARM L=9999,SYSAFF=SC42
//STEP1    EXEC DSNUPROC,
//         UTPROC='',SYSTEM=DB2K
//COPYTB1  DD DSN=SAPRES1.A140X8EU.DMSXPHR.TBSP,UNIT=SYSDA,
//         SPACE=(CYL,(15,1)),DISP=(NEW,CATLG,CATLG)
//COPYIX1  DD DSN=SAPRES1.A140X8EU.DMSRPHRE.INDEX,UNIT=SYSDA,
//         SPACE=(CYL,(15,1)),DISP=(NEW,CATLG,CATLG)
//SYSIN DD *
COPY
     TABLESPACE A140X8EU.DMSXPHR COPYDDN(COPYTB1)
     INDEXSPACE A140X8EU.DMSRPHRE COPYDDN(COPYIX1)
SHRLEVEL(CHANGE)
/*
```

The following illustration shows the output from the online backup of both tablespaces

```
OUTPUT START FOR UTILITY, UTILID = SAPRES1.SAPRES1C
COPY TABLESPACE A140X8EU.DMSXPHR COPYDDN(COPYTB1) INDEXSPACE A140X8EU.DMSRPHRE
GE)
COPY PROCESSED FOR TABLESPACE A140X8EU.DMSXPHR
NUMBER OF PAGES=2
AVERAGE PERCENT FREE SPACE PER PAGE =114.00
PERCENT OF CHANGED PAGES =  0.00
ELAPSED TIME=00:00:00
DB2 IMAGE COPY SUCCESSFUL FOR TABLESPACE A140X8EU.DMSXPHR
COPY PROCESSED FOR INDEXSPACE A140X8EU.DMSRPHRE
NUMBER OF PAGES=5
AVERAGE PERCENT FREE SPACE PER PAGE = 45.00
PERCENT OF CHANGED PAGES =  0.00
ELAPSED TIME=00:00:01
DB2 IMAGE COPY SUCCESSFUL FOR INDEXSPACE A140X8EU.DMSRPHRE
UTILITY EXECUTION COMPLETE, HIGHEST RETURN CODE=0
```

5.2.2 Offline backups

A tablespace offline backup is a copy of the tablespace taken while no write activity is allowed to that tablespace. As a result all of the data is in a committed status, which means that this backup alone could form the basis for a tablespace recovery but only to the point in time that the backup was taken.

Offline does not mean that either the DB2 UDB for OS/390 subsystem or the SAP R/3 system must be offline. On all the other tablespaces concurrent read and write activity can continue, while the tablespace being backed up allows read only access. However, for tablespaces that are heavily updated by many users concurrently, removing them from update activity could be extremely disruptive to the SAP R/3 system. This would especially be the case if the backup method was not fast enough.

From a DB2 UDB for OS/390 perspective, an offline backup is a sharelevel reference full image copy of the tablespace. As in the case of an online backup, the offline backup can also be incremental. The FULL, CHANGELIMIT, COPYDDN and RECOVERYDDN are also equally applicable to offline backups.

Another example of creating an offline backup of the SAP R/3 database uses the CONCURRENT parameter on the COPY statement. This has the advantage that it does not require a separate quiesce and restart step, as the database activity will be quiesced and made available again automatically. However, on the downside, the concurrent copy may fail in the phase of hardening the data (making the physical copy), once the logical copy has completed successfully and the SAP R/3 database is made available again for read and write. In this instance, the copy as a whole has not completed and the offline backup has not been created. To successfully create this backup, the cause of the problem has to be removed and the process must be repeated.

A special type of offline backup is an SAP R/3 database offline backup. This is a set of copies of all the tablespaces, the DB2 catalog and directory taken with no write activity allowed in the system. This has traditionally been almost too restrictive to online processing, especially in a high availability environment, to prevent it being carried out regularly. However, with the new log suspension techniques discussed in Chapter 5.3, “New DB2 UDB for OS/390 database backup option” on page 120 this method is now a very powerful way of backing up a database in its entirety with a tool such as RVA SnapShot, providing an excellent prior point in time recovery target.

Since some installations do not have RVA installed, we’ll now discuss the two methods for taking offline backups. However, our recommendations fall heavily on the backup procedures discussed in Chapter 5.2.2.2, “External backups with RVA SnapShot” on page 119, relying on RVA to take a full system backup for an accurate point of consistency, rather than working at the tablespace level.

5.2.2.1 Creating offline backups

The first step requires that no updates are allowed in the system. Unfortunately, when using the DB2 UDB for OS/390 COPY utility, the log suspend feature discussed in the next section, Chapter 5.3, “New DB2 UDB for OS/390 database backup option” on page 120, is not available due to the need to update entries in the SYSCOPY tablespace. Therefore for offline backups at the tablespace level you must first stop all of the updates in the system with the following commands:

```
START DATABASE (*) ACCESS (RO)
START DATABASE (DSNDB01) ACCESS (UT)
START DATABASE (DSNDB06) ACCESS (UT)
```

Once these commands have been entered, a job similar to the following sample JCL could be submitted:

```
//SAPRES1C JOB (999,POK),'NICK VALLANCE',CLASS=A,MSGLEVEL=(1,1),
// MSGCLASS=T,NOTIFY=&SYSUID,TIME=1440,REGION=0M
/*JOBPARM L=9999,SYSAFF=SC42
//STEP1    EXEC DSNUPROC,
//         UTPROC='',SYSTEM=DB2K
//SYSCOPY  DD DSN=SAPRES1.A140X8EU.DMSXPHRA.BACKUP.OFF,UNIT=SYSDA,
//         SPACE=(CYL,(15,1)),DISP=(NEW,CATLG,CATLG)
//SYSIN DD *
COPY TABLESPACE A140X8EU.DMSXPHR SHRLEVEL(REFERENCE)
/*
```

To take a completely recoverable backup it is necessary to copy all the SAP R/3 tablespaces together, to at the very least the key tablespaces in the system. To speed this process up a number of COPY jobs can be created and run in parallel. This method will take a long time, and it would still be best to try and use RVA where available, or DFSMSdss DUMP of the volumes.

Once the job has completed successfully, and all the requisite tablespaces have been dumped at a point of consistency, the DB2 databases need to be restarted allowing normal access.

5.2.2.2 External backups with RVA SnapShot

When the SAP R/3 database and the DB2 UDB for OS/390 system volumes are all stored on IBM RAMAC Virtual Array (RVA), the RVA's virtual disk architecture enables you to produce almost instantaneous copies of data sets, and volumes. This is the SnapShot facility, which produces copies without actual data movement.

Conventional methods of copying data on DASD consist of making a physical copy of the data on either DASD or tape. Host processors, channels, tape, and DASD controllers are involved in these conventional copy processes. Copying may take a lot of time, depending on the availability of system resources.

With SnapShot, copying is achieved with no data movement so it can take just seconds rather than minutes or hours, because data is not moved and the host processor and channels are not involved in data transfer. Furthermore, additional physical disk space is not required to accommodate the snap until either the source or the target is updated. As far as the operating system is concerned, the snap is a real copy of the data; as far as the RVA hardware is concerned, it is a virtual copy of the data.

See the JCL in the following illustration for performing the SnapShot of volume DB2T01:

```
//SNAPDUMP JOB (999,POK),'RVA SNAP DUMP',CLASS=A,MSGCLASS=T,
// NOTIFY=&SYSUID,TIME=1440,REGION=0M
/*JOBPARM L=999,SYSAFF=SC42
//SNAP1    EXEC PGM=SIBBATCH
//STEPLIB  DD   DSN=IXFP210.PROD.SIBLOAD,DISP=SHR
//         DD   DSN=IXFP210.PROD.STKLOAD,DISP=SHR
//         DD   DSN=IXFP210.PROD.SACLINK,DISP=SHR
//SYSTERM  DD   SYSOUT=*
//SYSPRINT DD   SYSOUT=*
//SYSIN    DD   *
 SNAP VOLUME( -
             SOURCE(VOLUME(DB2T01)) -
             TARGET(VOLUME(DB2X01)) -
             COPYVOLID(NO) -
             REPLACE(YES) -
            )
```

The following screen shows the output from a successful snap taken of the volume:

```
SIB0715I   IXFP 2.1.1 (SIBBATCH) invoked on Sun Apr 30, 2000 16:22:35.
SIB0700I   16:22:35  SNAP VOLUME( -
SIB0700I   16:22:35              SOURCE(VOLUME(DB2T01)) -
SIB0700I   16:22:35              TARGET(VOLUME(DB2X01)) -
SIB0700I   16:22:35              COPYVOLID(NO) -
SIB0700I   16:22:35              REPLACE(YES) -
SIB0700I   16:22:35              TOLENQF(YES) -
SIB0700I   16:22:35             )
SIB4630I   A VARY ONLINE request has been issued for device 25B1.
SIB4617I   16:22:37 SnapShot completed, rc=0.
```

5.3 New DB2 UDB for OS/390 database backup option

In order to achieve a reliable point in time for recovery, in the past a quiesce point has been required. This would necessitate the use of either the quiesce utility, the ARCHIVE LOG command or a DB2 shutdown, none of which are easy to achieve with the size or availability requirements of an SAP R/3 4.6 database.

This has been addressed in DB2 UDB for OS/390 v6.1 by the introduction of a new function with APAR/PTF PQ31492/UQ36695, allowing DB2 update activity to be periodically suspended. This provides a method for controlled ‘freezing’ or suspending the updates to the DB2 subsystem while the logs and database can be copied. This copying can be done without having to stop the DB2 subsystem and once the copy has been completed, the database logging can be resumed and the copied data can be archived to tape.

5.3.1 Activating log suspension

To activate this facility, new options have been added to the -SET LOG command to be able to SUSPEND and RESUME logging and therefore updates for a DB2 subsystem. When a SUSPEND request is issued the following events occur:

- A system checkpoint will be taken (in a non-data sharing environment)
- Any unwritten log buffers will be written to DASD

- the BSDS will be updated with the high- written RBA
- the log-write latch is obtained to prevent any further log records from being created. This will prevent any further updates to the data base until update activity is resumed. The latch will be held until a -SET LOG command is issued to RESUME logging, or until a -STOP DB2 command is issued.

See the results of the suspension commands in the following screen:

```
 DSN
-set log suspend
 DSNJ372I =DBK4 DSNJC009 UPDATE ACTIVITY HAS BEEN SUSPENDED FOR DBK4
 AT RBA 000000A0D662
 DSN9022I =DBK4 DSNJC001 '-SET LOG' NORMAL COMPLETION
 DSN
```

The scope for these commands is single-subsystem only, so in a data sharing environment such as ours the commands have to be entered for each member. The output from a -DISPLAY LOG command, as shown in the following figure Fig xx, also indicates that logging is suspended after a -SET LOG SUSPEND command has been issued. While updates are suspended in the system it should still be possible to carry out normal read-only activity providing there is no other resource contention.

See the results of the -DISPLAY LOG command in the following illustration:

```
 DSN
-dis log
 DSNJ370I =DBK4 DSNJC00A LOG DISPLAY
 CURRENT COPY1 LOG = DB2V610K.DBK4.LOGCOPY1.DS01 IS 9% FULL
 CURRENT COPY2 LOG = DB2V610K.DBK4.LOGCOPY2.DS01 IS 9% FULL
 H/W RBA = 000000A0D662, LOGLOAD = 100000
 FULL LOGS TO OFFLOAD = 0 OF 6, OFFLOAD TASK IS (AVAILABLE)
 DSNJ371I =DBK4 DB2 RESTARTED 09:37:59 APR 21, 2000
 RESTART RBA 00000001D000
 DSNJ372I =DBK4 DSNJC00A UPDATE ACTIVITY HAS BEEN SUSPENDED FOR DBK4
 AT RBA 000000A0D662
 DSN9022I =DBK4 DSNJC001 '-DIS LOG' NORMAL COMPLETION
```

See the results of resuming updates in the system in the following figure:

```
 DSN
-set log resume
 DSNJ373I =DBK4 DSNJC009 UPDATE ACTIVITY HAS BEEN RESUMED FOR DBK4
 DSN9022I =DBK4 DSNJC001 '-SET LOG' NORMAL COMPLETION
 DSN
```

For a consolidated view of the suspensions and resumptions, the DB2 started task output can be checked for the following messages:

```
16.35.18 STC17632 *DSNJ372I =DBK1 DSNJC09A UPDATE ACTIVITY HAS BEEN
                   SUSPENDED FOR DBK1 AT RBA 0015D29DC8E8
16.35.18 STC17632  DSN9022I =DBK1 DSNJC001 '-SET LOG' NORMAL COMPLETION
16.37.17 STC17632  DSNJ373I =DBK1 DSNJC09A UPDATE ACTIVITY HAS BEEN RESUMED FOR
16.37.17 STC17632  DSN9022I =DBK1 DSNJC001 '-SET LOG' NORMAL COMPLETION
```

5.3.2 Our tests with the log suspension feature

Now we have discussed the different backup options in Chapter 5.2, “Backup and recovery procedures” on page 115, and introduced the log suspension feature, it would be useful to consider how the two can be used together for backup and recovery in the future.

As part of our testing we used both the log suspension and RVA SnapShot to backup all the volumes. The order of events for our test were:

1. Enter the log suspend command for all four DB2 subsystems in our data sharing group.
2. Once this is complete, submit the job to SNAP all of the SAP volumes, which runs on average around 15 seconds to dump 8 volumes.
3. Once the job has completed successfully for all volumes, resume the DB2 updates with the set log resume command. DB2 is now running normally again.
4. A logical copy of the whole system has now been made, and you can recover to this point in time. You can start a DFSMSdss volume dump to copy these DASD backup volumes to tape. This may take some time, but the creation of the tapes can continue while the SAP R/3 system is running.

As you can see from the timings for this test, it is a very quick and simple way of backing up the entire DB2 database, without the need to wait for occasionally long quiesce jobs to complete. Using this method results in the least time were DB2 UDB for OS/390 is unavailable for updates. During our testing of this procedure, as the update function was suspended, we also attempted updates to the SAP R/3 system using our workload generation tool, to see how this was handled. This resulted in no roll backs or shortdumps, the only noticeable result was a ‘hang’ on the SAP R/3 GUI. We could also observe in DB2 PM that all update activity had ceased.

Once the set log resume command was issued the updates continued as normal, also visible in DB2PM.

5.4 High Availability Split Mirror Backup Recovery

While the log suspension facilities will have a major impact on the methods of backing up DB2 UDB for OS/390 systems, these new features can be adapted further and used in a new method for achieving high availability and Disaster Recovery (DR) capabilities. This High Availability Split Mirror Backup and Recovery solution is a new approach developed jointly by IBM and SAP which takes into consideration the fact that many customers are faced with the need for very large databases and an increased demand for high availability.

As a result, in the case of a disaster the system has to be available again within minutes. From the systems management point of view this is a contradiction, since the fastest way to recover to a certain point in time requires a consistent database backup, which can only be achieved during a quiet system period. However, continuous system availability can prevent any quiet periods.

Since high availability of the live system, in general, has the highest priority for most customers, they prepare themselves for a DR by taking backups that in the case of recovery will require updates being reapplied from the logs to establish database consistency and to rebuild the indexes. This can result in very long recovery times.

One way to avoid these drawn out recoveries would be the use of a remote copy, which is constantly synchronized with the live database. The split mirror backup takes this backup on a remote RVA that is connected to a primary RVA by the peer-to-peer remote copy (PPRC) function. The PPRC function is a hardware-based remote copy service that provides a synchronous volume copy across storage subsystems.

Normally the PPRC connection is suspended (split mirror) and will only be resumed for the resynchronisation of the remote database copy. This remote copy can also be used in some way by the customer, for example as a read only information system.

For more detailed information on how the split mirror solution on either RVA or the more recent ESS DASD can be implemented, view the Web site:

```
http://www.storage.ibm.com/hardsoft/diskdrls/technology.htm
```

5.5 Recovery

In the event of any failure it is vital that the database can be restored quickly. Some recovery operations are done automatically by DB2 UDB for OS/390 without any outside intervention, such as recovering the database to a consistent state before an operating system or database failure had occurred. Consistent state means that all database changes made by incomplete units of work will be rolled back, and the changes made to complete units of work will be committed. In this case, automatic recovery happens at the next DB2 for OS/390 start.

It is important to bear in mind that an SAP R/3 database includes all the tablespaces, indexes and DB2 catalog and directory entries (practically all the catalog and directory tablespaces and indexes) that are related to the SAP R/3 system. From the operational and database integrity's viewpoint an SAP R/3 database as a whole needs to be considered a single unit of recovery. In other words, if a single SAP R/3 tablespace needs to be recovered to a point in time, all the other SAP R/3 tablespaces and indexes need to be either also recovered to the same point in time, or already be at the state that they had at that time.

To allow recovery of DB2 data, the RECOVER utility is provided by DB2, allowing recovery of DB2 objects such as tablespaces, indexes, partitions, individual datasets, and individual pages. The RECOVER utility can recover the data to a particular backup, a point in time or the current state depending on the parameters passed to it.

Recovery of the database can be classified in two ways, depending on the time to which the data needs to be restored:

- Current state recovery
- Point in time recovery

We'll now review the two types of recovery.

5.5.1 Recovery to current state

A recovery to the current state is generally less demanding and is usually needed more often than a point-in-time recovery. Often the recovery to current state is automatic, a typical example is volume failure in a direct access storage device (DASD), resulting in data loss. You need to find out which tablespace and indexes were located on the volume and recover only these tablespace and indexes, or even only partitions or individual datasets that are affected. The rest of the system is already at the current state and need not be recovered.

5.5.1.1 Example of recovering with an online backup

To try a simple test of recovering from an online backup we recovered the tablespace and indexspace backups we created in Chapter 5.2.1.1, "Creating online backups" on page 117 to a current state. This is also a chance to try the new RECOVER INDEXSPACE command, which removes the need in the past for a sometimes lengthy table scan and new index being created.

The following JCL was used for this test:

```
//SAPRES1C JOB (999,POK),'NICK VALLANCE',CLASS=A,MSGLEVEL=(1,1),
// MSGCLASS=T,NOTIFY=&SYSUID,TIME=1440,REGION=0M
/*JOBPARM L=9999,SYSAFF=SC42
//STEP1    EXEC DSNUPROC,
//         UTPROC='',SYSTEM=DB2K
//SYSIN DD *
RECOVER TABLESPACE A140X8EU.DMSXPHR
        INDEXSPACE A140X8EU.DMSRPHRE
/*
```

This job successfully completed with the following messages:

```
OUTPUT START FOR UTILITY, UTILID = SAPRES1.SAPRES1C
RECOVER TABLESPACE A140X8EU.DMSXPHR INDEXSPACE A140X8EU.DMSRPHRE
RECOVER TABLESPACE A140X8EU.DMSXPHR   START
THE IMAGE COPY DATA SET SAPRES1.A140X8EU.DMSXPHR.TBSP WITH DATE=20000503 AND TIM
IS PARTICIPATING IN RECOVERY OF TABLESPACE A140X8EU.DMSXPHR
MERGE STATISTICS FOR TABLESPACE A140X8EU.DMSXPHR   -
NUMBER OF COPIES=1
NUMBER OF PAGES MERGED=2
ELAPSED TIME=00:00:02
RECOVER INDEXSPACE A140X8EU.DMSRPHRE   START
THE IMAGE COPY DATA SET SAPRES1.A140X8EU.DMSRPHRE.INDEX WITH DATE=20000503 AND T
IS PARTICIPATING IN RECOVERY OF INDEXSPACE A140X8EU.DMSRPHRE
MERGE STATISTICS FOR INDEXSPACE A140X8EU.DMSRPHRE   -
NUMBER OF COPIES=1
NUMBER OF PAGES MERGED=5
ELAPSED TIME=00:00:01
RECOVERY COMPLETE, ELAPSED TIME=00:00:04
UTILITY EXECUTION COMPLETE, HIGHEST RETURN CODE=0
```

5.5.2 Recovery to a previous point in time

This type of recovery is used to reinstate the SAP R/3 database to the condition it was in at an earlier point in time. All changes after that time are lost. You must carefully consider the decision to set the system back in time. Typically, a recovery to a previous point in time is needed when an application program logic error introduced unwanted and irreversible changes into the system.

There are now three different methods to accomplish a previous point in time recovery of an SAP R/3 database:

- Recovery with conditional restart
- Recovery to the time of an offline backup
- Recover to the time of a log suspend or a quiesce point

We will now discuss these three in turn.

5.5.2.1 Recovery with Conditional Restart

Recovery with conditional restart has the least impact on everyday operations in terms of preparation for a previous point in time recovery of the SAP R/3 database. The main characteristics of the method are that neither offline backups nor quiesce points need to be provided which makes it the prime choice in High Availability SAP R/3 environments. It can also bring the system closest to the time when the SAP R/3 database is known to be consistent.

We'll now describe the steps necessary to perform a conditional restart.

1. Find out the Log Record Sequence Number (LRSN) or a Log Relative Byte Address (RBA) that is close to the time in which the system must be restored. This will depend on whether the system is running in data sharing mode or not.
   Data sharing mode
   Translate the time you require, represented as a timestamp, into its STCK format. This equates to the LRSN to which you want to restart the data sharing group.
   Non data sharing mode

First identify which log data set covers the interval containing the required time. To do this, run the print log map utility (DSNJU004). Once this has been identified, use DSN1LOGP SUMMARY on the log data set and calculate which RBA is closest to the required time. This must be a multiple of 4096.

2. Create a list of tablespaces that have been changed since the RBA or LRSN identified in (1). This can be produced either with the REPORT RECOVERY utility for all the tablespaces in the system, or by running the DSN1LOGP SUMMARY report.
   Finding the tablespaces that have been changed since the point in time to which you want to recover can significantly reduce the elapsed time for the recovery.

3. Stop DB2.

4. Copy the Boot Strap Data Set (BSDS) and all the logs that contain RBAs or LRSNs that are later than those found in step (1). This will allow the process to be repeated if it is later decided to recover the data again, but to a more recent point in time.

5. Use the DSNJU003 utility to create a conditional restart record in the BSDS, specifying the ENDRBA or ENDLRSN from step (1) and leaving other CRESTART options to default.

6. Update the system parameters to specify DEFER=ALL.
   This option means that all the objects that were in the started state at the point in time specified in step (5), will not be started at the next DB2 UDB for OS/390 start and so will not go through the normal restart process. DEFER does not affect processing of the log during the system restart, so DB2 still processes the log range for both the forward and backward log recovery phases of restart, but the logged operations are not applied to the deferred start data sets.

7. Start DB2. When restarting DB2 with an active CRESTART record it is initialized as if it is now running at the time indicated by the ENDRBA or ENDLRSN.

8. RECOVER the catalog, directory and all the tablespaces identified in step (2) to their current state, without a TORBA/TOLOGPOINT or TOCOPY parameter.

9. REBUILD any indexes on the tablespaces recovered in step (8), or if DB2 UDB for OS/390 v6 is being used the indexes can also be restored with RECOVER.

10. Finally reinstate RESTART ALL, allowing DB2 UDB for OS/390 to make a normal restart again in any future restarts of the subsystem. The databases should now be back at the time indicated by the RBA or LRSN chosen, and it would now be a good idea to take a new backup of the system.

5.5.2.2 Recovery with an offline backup

Recovery to the state at the time an offline backup was taken is the simplest and fastest point in time recovery method, but it is also the most restrictive. In the first instance it requires taking the original backup, which in the past may have resulted in an extended outage that would be unacceptable. Additionally, depending on the frequency of backups, it could restore the system much further than is necessary, for example, if the offline backups of the SAP R/3 database were taken weekly on a Sunday, but the data was damaged on Friday, restoring to the last backup taken would unnecessarily result in almost a week of lost work.

Recovering to an offline backup can be carried out by using the RECOVER utility with the TOCOPY option, a DFSMSdss restore or RVA snap, depending on the method used for taking the original backup. We recommend the use of RVA with Snapshot for performing backups and recoveries, especially now the log suspension option is available.

If RVA SnapShot is not available, then the recovery will require the complete set of offline backups taken with COPY and shrlevel reference. To return all the tablespaces to exactly this point in time will require control statements specifying TOCOPY with the name of the backup dataset for each tablespace.

5.5.2.3 Recovery with SnapShot

This is the method we recommend for backing up and recovering an SAP R/3 system. Using the method described in Chapter 5.3, “New DB2 UDB for OS/390 database backup option” on page 120, very fast backups were taken with only a very small reduction in update capabilities. Since taking backups with this method is now so fast and simple, it should be possible to take these backups more regularly, and therefore recovering to a point in time is a case of restoring the snapped volumes and restarting DB2 UDB for OS/390.

If the recovery is to the time of the latest backup, the data on the backup volumes is still valid, and there is no requirement to restore from tape. If the recovery is necessary from a previous snap copy, firstly those tapes must be restored to the backup DASD volumes after which the volumes must be snapped back.

5.5.2.4 Example of recovering with a SnapShot backup

With DB2 UDB for OS/390 down, the volumes need to be snapped back onto the production volumes with the commands shown in the following JCL extract:

```
//SNAPREST JOB (999,POK),'RVA SNAP REST',CLASS=A,MSGCLASS=T,
// NOTIFY=&SYSUID,TIME=1440,REGION=0M
/*JOBPARM L=999,SYSAFF=SC42
//SNAP1    EXEC PGM=SIBBATCH
//STEPLIB  DD   DSN=IXFP210.PROD.SIBLOAD,DISP=SHR
//         DD   DSN=IXFP210.PROD.STKLOAD,DISP=SHR
//         DD   DSN=IXFP210.PROD.SACLINK,DISP=SHR
//SYSTERM  DD   SYSOUT=*
//SYSPRINT DD   SYSOUT=*
//SYSIN    DD   *
 SNAP VOLUME( -
             SOURCE(VOLUME(DB2X01)) -
             TARGET(VOLUME(DB2T01)) -
             COPYVOLID(NO) -
             REPLACE(YES) -
            )
```

In our test of this recovery procedure, this snap took only 14 seconds, resulting in the whole DB2 UDB for OS/390 environment now being returned to the time the backup was taken.

The output from the snap restore is shown in the following illustration

```
SIB0700I   16:06:13  SNAP VOLUME( -
SIB0700I   16:06:13              SOURCE(VOLUME(DB2X07)) -
SIB0700I   16:06:13              TARGET(VOLUME(DB2T07)) -
SIB0700I   16:06:13              COPYVOLID(NO) -
SIB0700I   16:06:13              REPLACE(YES) -
SIB0700I   16:06:13             )
SIB4630I   A VARY ONLINE request has been issued for device 251B.
SIB4617I   16:06:14 SnapShot completed, rc=0.

SIB0700I   16:06:14  SNAP VOLUME( -
SIB0700I   16:06:14              SOURCE(VOLUME(DB2X08)) -
SIB0700I   16:06:14              TARGET(VOLUME(DB2T08)) -
SIB0700I   16:06:14              COPYVOLID(NO) -
SIB0700I   16:06:14              REPLACE(YES) -
SIB0700I   16:06:14             )
SIB4630I   A VARY ONLINE request has been issued for device 2504.
SIB4617I   16:06:15 SnapShot completed, rc=0.
```

Recovery with this method has the advantages of recovering the complete DB2 UDB for OS/390 environment, including the DB2 UDB for OS/390 catalog, to a single point of consistency. This is a very fast recovery technique, and as long as backups are now being taken with more regularity, the system shouldn't be restored to a point too far in the past.

5.5.2.5 Recover to the time of a log suspend or a quiesce point

As we've already discussed in this chapter, the new log suspend feature provides a reliable method for ensuring a consistent point for recovery. With this feature, it is no longer necessary to run the quiesce utility. This can now be considered in conjunction with the previous section, Chapter 5.5.2.4, "Example of recovering with a SnapShot backup" on page 127. They should be considered together because the log suspend feature now represents the best way of achieving a point of consistency, and will usually be used to stop updates in the system prior to a backup being taken.

5.6 Backup and recovery summary and recommendations

As we described earlier in this chapter, each SAP R/3 installation should choose a backup and recovery strategy that is best suited to its particular needs and conditions.

We have discussed the different types of backups that are available, both online and offline. Additionally we have considered some of the new developments in this area, namely the log suspend feature, and recovery of indexes.

We feel the greatest benefits to availability in terms of backups and recovery will occur when customers are able to fully exploit the log suspend feature. When this is combined with the capabilities of RVA SnapShot it is now possible to take very fast backups of the complete SAP R/3 database for the easiest point in time recovery.

Our recommendations for an installation's backup strategy are, where possible, to take regular online (shrlevel change) image copies, combined with regular log suspends, during which RVA SnapShot should be used to take full volume copies. Taking regular online image copies of the data is a mandatory requirement for any well managed system. Also for index recovery available in DB2 UDB for OS/390 version 6 you should continue using the RECOVER control statement, but now for INDEXSPACEs. For this to be available you must ensure that the indexes have been defined with COPY YES, if not an ALTER INDEX is required. This can make major improvements in recovery time and hence availability, as there is no longer the need to perform expensive table scans to rebuild indexes of recovered tablespaces.

Now the log suspend function is available, making it so easy to obtain a point of consistency, it makes the traditional requirement for obtaining quiesce points and therefore extended outages obsolete. The implication of this to a backup strategy in the future is that offline backups taken at the tablespace level with sharelevel reference would no longer be necessary.

Appendix A. Availability test scenarios

This appendix is a listing of defined test scenarios for availability. It should serve as a reference list to assist you in your availability planning. The test results shown here are the results for a single specific environment. We believe they show what is achievable, but you should select the items most important in your installation and test those rather than rely on these results. We are particularly indebted to Mike Gordon, Don Geissler, and Patrick Horkan of the IBM SAP Competency Center for providing this list.

Table 7 shows an extensive list of failure scenarios. Tests were performed in a specific environment to determine the impact of such a failure. The impact shown in the “Effect” column assumes the availability plan for that environment is followed.

Table 7. High availability test scenarios

High availability scenario	Effect
OS/390 failure	No impact / SQL 0000
OS/390 system upgrade	No impact / SQL 0000
DB2 failure and automatic restart	No impact / SQL 0000
DB2 upgrade	No impact / SQL 0000
Coupling Facility failure	No impact
Coupling Facility link failure	No impact
Coupling Facility takeover	No impact
Channel path (CHPID) failure	No impact
ICLI failure	No impact / SQL 0000
ESCON failure	No impact / SQL 0000
Central processor complex failure	No impact / SQL 0000
CPU engine failure	No impact
DB2 online full image copy utility	No impact
Incremental copies and merge copy full	No impact
DB2 online REORG utility	Slow responses
Recover SAP tablespace to current state	Short term outage
DB2 point in time recovery	Outage during recovery
Dynamically adjust the hardware CPU capacity	No impact
Dynamically add XCF signaling paths	No impact
Dynamically adjust dispatching priority	No impact
Loss of redundant power supply or cooling unit	No impact
Loss of redundant utility power	No impact

High availability scenario	Effect
Dynamically add DASD volumes	No impact
Dynamically add DASD work space (2 TESTS)	No impact
Hardware management console concurrent patch	No impact
Support Element (SE) failure	No impact
Support Element concurrent patch	No impact
Daylight savings time spring ahead	No impact
Daylight savings time fall back	No impact
Sysplex timer link failure	No impact
Sysplex timer failure	No impact
* Control workstation failure	No impact
* High Speed Switch adapter failure	Lost session
* Mixed AIX levels	No impact
* HACMP central instance node failure	Lost session
* HACMP application server failure	Lost session
* HACMP central instance failure	Lost session
* HACMP central instance failure and DB2 failure	Lost session
* Install maintenance to PSSP	No impact
* Add or remove an SP node or frame	No impact

(* Indicates this item is outside the scope of this redbook. It also indicates items that are dependent on the application server platform; therefore Windows NT, SUN Solaris, or OS/390 application server users would substitute similar tests appropriate to their application server.)

Appendix B. SAPNet R/3 - Notes (OSS) & Informational APARs

This appendix lists selected frontend notes and PTFs that are useful in constructing and maintaining a high availability SAP R/3 system on the OS/390 platform. It should serve as a reference list to assist you in your availability planning.

Table 8. OSS Notes

OSS Note	Description
81737	APAR List
183311	Automated PTF Check
138906	Planning Info (Release, Connectivity Matrix)
153073	Failover to Standby DB2 in same LPAR
186882	First Connect to Standby ICLI Server
108469	DB2 Datasharing clarification
46272	Implement new data class in technical settings
163449	SAP Rules for Self Defined Objects
83000	SAP DB2 Database Recovery Options
162818	DDIC Corrections (4.5)
184339	DDIC Corrections (4.6)
113008	Maintaining Catalog Statistics
103135	Manual Installation of saposcol

APARs	Description
II11352	OS/390 Release Matrix
II12111	OS/390 Misc. Information
II11173	DB2 DBD: Hints and tips
PQ14391	DB2 Storage Contraction Enhancement - CONTSTOR
PQ30999	Defer DEFINE Support
PQ31492	Suspend DB2 update activity
PQ32213	Extend failed during insert
PQ34592	Retrofit defect for DEFINE NO
PQ34030	REPAIR DBD REBUILD fix
PQ34029	DEFER DEFEINE Enhancement's and restrictions
PQ33429	Access path adjustment
PQ34386	Defer DEFINE fix

Appendix C. Tracefile of the work processes after adapter failure

```
------------------------------------------------
trc file: "dev_w0", trc level: 1, release: "46B"
------------------------------------------------
*
*  ACTIVE TRACE LEVEL           1
*  ACTIVE TRACE COMPONENTS      all, M
*
M
M  Thu May  4 13:21:54 2000
M  relno      4620
M  patchlevel 0
M  patchno    133
M  intno      0
M  pid        17864
M
M  ***LOG Q01=> tskh_init, WPStart (Workproc 0 1 17864) [thxxhead 0911]
M  ThSysUsrCheck: clear T0
M  calling db_connect ...
B  DB_RECONNECT: standby_is_def is set
B  DB_RECONNECT: Reconnect for parallel DB is active
M  *** ERROR => ThGetDBConn: no connect info saved [thxxtool 6801]
B  find_con found the following connection for reuse:
B  0: name = R/3, concnt= 000000 state = DISCONNECTED, perm = YES, reco = NO
C  DBSL TRACE: ICLI client functions successfully loaded.
(lib=/usr/sap/RED/SYS/exe/run/ibmiclic.o,sldb2.c,l=8994)
C  ********* Profile Parameter for connection 0 **********
C  DBIF Versions:
C    R/3 Version       : 12 (4.6B)
C    API Version       : 9
C  ICLI-Client Versions:
C    R/3 Version       : 12 (4.6B)
C    API Version       : 9
C    PROTOCOL Version  : 1
C    Internal Version  : 1
C  Workprocess         : 0
C  USE_NO_HOLD         : 0
C  USE_OPT_1_ROW       : 0
C  DB_UID              : SAPR3
C  DB_ID               : RED
C  SAPSYSTEM           : 00
C  DB_SSID             : DBK1
C  DBHOST              : redsap1
C  PORT                : see /etc/services entry sapdb2RED
C  CCSID               : 819
C  KEEP_X_LOCK         : 1
C  RO_COMMIT_SUPPRESS  : 0
C  --------------------------------------
C  Used Bugfixes:
C  **************************************
C  Network-Protocol    : TCP
C  *** ERROR =>
[000045C8] ICLC2520W Security is switched off - it should be switched off on the server too!
           Time: 957460914.104366 = (Thu) 2000 May 04 13:21:54.104366
 [dbsldb2. 2141]
C
C  Thu May  4 13:21:56 2000
C  Your DB2 system works with ASCII code page 819 and EBCDIC code page 37
B  Connection 0 opened
```

Figure 40. Extract from the tracefile of the work processes - part 1 of 9

```
M  db_connect o.k.
X  EmInit: MmSetImplementation( 2 ).
B
B  Thu May  4 13:21:57 2000
B  db_con_shm_ini:  WP_ID = 0, WP_CNT = 13
B
B  Thu May  4 13:22:08 2000
B  ***LOG BB0=> buffer TABL       started with length  50000000  bytes [dbtbxbuf 1521]
B  ***LOG BB0=> buffer TABLP      started with length 030720000  bytes [dbbfe    0859]
B  Layout of object buffer shared memory:
B  0: 1 * 4 = 4
B  1: 1 * 196 = 196
B  2: 13 * 20 = 260
B  3: 5003 * 20 = 100060
B  4: 2500 * 128 = 320000
B  5: 5003 * 4 = 20012
B  6: 1 * 100 = 100
B  7: 193 * 4 = 772
B  8: 18275 * 256 = 4678400
B  ***LOG BB0=> buffer CUA        started with length 5000k      bytes [dbexpbuf 1869]
B  Layout of object buffer shared memory:
B  0: 1 * 4 = 4
B  1: 1 * 196 = 196
B  2: 13 * 20 = 260
B  3: 10007 * 20 = 200140
B  4: 5000 * 128 = 640000
B  5: 10007 * 4 = 40028
B  6: 1 * 100 = 100
B  7: 65 * 4 = 260
B  8: 25885 * 128 = 3313280
B  ***LOG BB0=> buffer EIBUF      started with length 4096k      bytes [dbexpbuf 1869]
B  ***LOG BB0=> buffer CALE       started with length 500000     bytes [dbcalbuf 1583]
B  dbtran INFO (init_connection '<DEFAULT>' [DB2:46B.00]):
B   max_blocking_factor =  10,  max_in_blocking_factor      =  35,
B   min_blocking_factor =   3,  min_in_blocking_factor      =   6,
B   prefer_union_all    =   1,  prefer_union_for_select_all =   0,
B   prefer_fix_blocking =   1,  prefer_in_itab_opt          =   1,
B   convert AVG         =   1,  alias table FUPD           =   0,
B   escape_as_literal   =   0,  opt GE LE to BETWEEN       =   0,
B   select *            = 0x3,
B   use_hints           = abap->0, dbif->1, upto->4294967295, rule_in->0,
B                         rule_fae->0, concat_fae->0, concat_fae_or->0
M
M  Thu May  4 13:22:11 2000
M  SsfSapSecin: automatic application server initialization for SAPSECULIB
M  SsfSapSecin: Looking for PSE in database
M  SsfPseLoad: started...(path=/usr/sap/RED/DVEBMGS00/sec, AS=sap1tr0, instanceid=00)
M  SsfPseLoad: Downloading file /usr/sap/RED/DVEBMGS00/sec/SAPSYS.pse (client:    , key: SYSPSE, len: 1579)
M  SsfPseLoad: ended (1 of 1 sucessfully loaded, 1 checked...
A  PXA allocated (address 43b7e000, size 419332344)
A  System name
A  DB2.............................RED.....................
```

Figure 41. Extract from the tracefile of the work processes - part 2 of 9

```
A  is used for RFC security.
A  RFC Destination> destination sap1tr0_RED_00 host sap1tr0 system RED systnr 0 (sap1tr0_RED_00)
A  RFC Options> H=sap1tr0,S=00
A  RFC Method> initialize RemObjDriver for ABAP Objects
S
S  Thu May  4 13:22:13 2000
S  initialize debug system
T  Stack direction is downwards.
T  debug control: prepare exclude for printer trace
T  new memory block 21277580
S  spool kernel/ddic check: Ok
S  1 spool work process(es) found
S  frontend print via spool service enabled
S  printer list size is 150
S  printer type list size is 50
S  queue size (profile)   = 300
S  hostspool list size = 3000
S  option list size is 30
S      found processing queue enabled
S  found spool memory service RSPO-RCLOCKS at 5cbfba70
S  doing lock recovery
S  setting server cache root
S  found spool memory service RSPO-SERVERCACHE at 5cbfbc80
S    using messages for server info
M  SsfSapSecin: putenv(SECUDIR=/usr/sap/RED/DVEBMGS00/sec): ok
M  SsfSapSecin: PSE /usr/sap/RED/DVEBMGS00/sec/SAPSYS.pse found!
M
M  ================================================
M  === SSF INITIALIZATION:
M  ===...SSF Security Toolkit name SAPSECULIB .
M  ===...SSF trace level is 0 .
M  ===...SSF library is /usr/sap/RED/SYS/exe/run/libssfso.o .
M  ===...SSF hash algorithm is SHA1 .
M  ===...SSF symmetric encryption algorithm is DES-CBC .
M  ===...sucessfully completed.
M  ================================================
E
E  Thu May  4 13:22:15 2000
E  Replication is disabled
S
S  Thu May  4 13:23:46 2000
S  server @>SSRV:sap2tr0_RED_00@< appears or changes
C
C  Thu May  4 13:48:14 2000 ----------------------------------------------------> [1]
C  *** ERROR =>
[000045C8] ICLC2511E Could not receive data from server - Network is down (errno=69)
           Time: 957462494.827625 = (Thu) 2000 May 04 13:48:14.827625
 [dbsldb2. 2141]
C  *** ERROR =>
[000045C8] ICLC2507E Receiving data failed in transmission phase
           Details: stream transmission mode
           Time: 957462494.827802 = (Thu) 2000 May 04 13:48:14.827802
 [dbsldb2. 2141]
C  *** ERROR =>
[000045C8] ICLC2184E Internal error in line 4048 and file ibmiclio.c; Processing continues.
           Details: Transmit failed for performance monitor functions; retcode = -5
           Time: 957462494.827921 = (Thu) 2000 May 04 13:48:14.827921
```

Figure 42. Extract from the tracefile of the work processes - part 3 of 9

```
A  is used for RFC security.
A  RFC Destination> destination sap1tr0_RED_00 host sap1tr0 system RED systnr 0 (sap1tr0_RED_00)
A  RFC Options> H=sap1tr0,S=00
A  RFC Method> initialize RemObjDriver for ABAP Objects
S
S  Thu May  4 13:22:13 2000
S  initialize debug system
T  Stack direction is downwards.
T  debug control: prepare exclude for printer trace
T  new memory block 21277580
S  spool kernel/ddic check: Ok
S  1 spool work process(es) found
S  frontend print via spool service enabled
S  printer list size is 150
S  printer type list size is 50
A  is used for RFC security.
A  RFC Destination> destination sap1tr0_RED_00 host sap1tr0 system RED systnr 0 (sap1tr0_RED_00)
A  RFC Options> H=sap1tr0,S=00
A  RFC Method> initialize RemObjDriver for ABAP Objects
S
S  Thu May  4 13:22:13 2000
S  initialize debug system
T  Stack direction is downwards.
T  debug control: prepare exclude for printer trace
T  new memory block 21277580
S  spool kernel/ddic check: Ok
S  1 spool work process(es) found
S  frontend print via spool service enabled
S  printer list size is 150
S  printer type list size is 50
S  queue size (profile)   = 300
S  hostspool list size = 3000
S  option list size is 30
S      found processing queue enabled
S  found spool memory service RSPO-RCLOCKS at 5cbfba70
S  doing lock recovery
S  setting server cache root
S  found spool memory service RSPO-SERVERCACHE at 5cbfbc80
S    using messages for server info
M  SsfSapSecin: putenv(SECUDIR=/usr/sap/RED/DVEBMGS00/sec): ok
M  SsfSapSecin: PSE /usr/sap/RED/DVEBMGS00/sec/SAPSYS.pse found!
M
M  ================================================
M  === SSF INITIALIZATION:
M  ===...SSF Security Toolkit name SAPSECULIB .
M  ===...SSF trace level is 0 .
M  ===...SSF library is /usr/sap/RED/SYS/exe/run/libssfso.o .
M  ===...SSF hash algorithm is SHA1 .
M  ===...SSF symmetric encryption algorithm is DES-CBC .
M  ===...sucessfully completed.
M  ================================================
E
E  Thu May  4 13:22:15 2000
E  Replication is disabled
S
S  Thu May  4 13:23:46 2000
S  server @>SSRV:sap2tr0_RED_00@< appears or changes
C
C  Thu May  4 13:48:14 2000
C  *** ERROR =>
[000045C8] ICLC2511E Could not receive data from server - Network is down (errno=69)
           Time: 957462494.827625 = (Thu) 2000 May 04 13:48:14.827625
 [dbsldb2. 2141]
C  *** ERROR =>
[000045C8] ICLC2507E Receiving data failed in transmission phase
           Details: stream transmission mode
           Time: 957462494.827802 = (Thu) 2000 May 04 13:48:14.827802
 [dbsldb2. 2141]
C  *** ERROR =>
[000045C8] ICLC2184E Internal error in line 4048 and file ibmiclio.c; Processing continues.
           Details: Transmit failed for performance monitor functions; retcode = -5
           Time: 957462494.827921 = (Thu) 2000 May 04 13:48:14.827921
```

Figure 43. Extract from the tracefile of the work processes - part 4 of 9

```
A  is used for RFC security.
A  RFC Destination> destination sap1tr0_RED_00 host sap1tr0 system RED systnr 0 (sap1tr0_RED_00)
A  RFC Options> H=sap1tr0,S=00
A  RFC Method> initialize RemObjDriver for ABAP Objects
S
S  Thu May  4 13:22:13 2000
S  initialize debug system
T  Stack direction is downwards.
T  debug control: prepare exclude for printer trace
T  new memory block 21277580
S  spool kernel/ddic check: Ok
S  1 spool work process(es) found
S  frontend print via spool service enabled
S  printer list size is 150
S  printer type list size is 50
A  is used for RFC security.
A  RFC Destination> destination sap1tr0_RED_00 host sap1tr0 system RED systnr 0 (sap1tr0_RED_00)
A  RFC Options> H=sap1tr0,S=00
A  RFC Method> initialize RemObjDriver for ABAP Objects
S
S  Thu May  4 13:22:13 2000
S  initialize debug system
T  Stack direction is downwards.
T  debug control: prepare exclude for printer trace
T  new memory block 21277580
S  spool kernel/ddic check: Ok
S  1 spool work process(es) found
S  frontend print via spool service enabled
S  printer list size is 150
S  printer type list size is 50
A  is used for RFC security.
A  RFC Destination> destination sap1tr0_RED_00 host sap1tr0 system RED systnr 0 (sap1tr0_RED_00)
A  RFC Options> H=sap1tr0,S=00
A  RFC Method> initialize RemObjDriver for ABAP Objects
S
S  Thu May  4 13:22:13 2000
S  initialize debug system
T  Stack direction is downwards.
T  debug control: prepare exclude for printer trace
T  new memory block 21277580
S  spool kernel/ddic check: Ok
S  1 spool work process(es) found
S  frontend print via spool service enabled
S  printer list size is 150
S  printer type list size is 50
S  queue size (profile)   = 300
S  hostspool list size = 3000
S  option list size is 30
S      found processing queue enabled
S  found spool memory service RSPO-RCLOCKS at 5cbfba70
S  doing lock recovery
S  setting server cache root
S  found spool memory service RSPO-SERVERCACHE at 5cbfbc80
S    using messages for server info
M  SsfSapSecin: putenv(SECUDIR=/usr/sap/RED/DVEBMGS00/sec): ok
M  SsfSapSecin: PSE /usr/sap/RED/DVEBMGS00/sec/SAPSYS.pse found!
M
M  ===============================================
M  === SSF INITIALIZATION:
M  ===...SSF Security Toolkit name SAPSECULIB .
M  ===...SSF trace level is 0 .
M  ===...SSF library is /usr/sap/RED/SYS/exe/run/libssfso.o .
M  ===...SSF hash algorithm is SHA1 .
M  ===...SSF symmetric encryption algorithm is DES-CBC .
M  ===...sucessfully completed.
M  ===============================================
E
E  Thu May  4 13:22:15 2000
E  Replication is disabled
S
S  Thu May  4 13:23:46 2000
S  server @>SSRV:sap2tr0_RED_00@< appears or changes
C
C  Thu May  4 13:48:14 2000
C  *** ERROR =>
[000045C8] ICLC2511E Could not receive data from server - Network is down (errno=69)
          Time: 957462494.827625 = (Thu) 2000 May 04 13:48:14.827625
 [dbsldb2. 2141]
C  *** ERROR =>
[000045C8] ICLC2507E Receiving data failed in transmission phase
          Details: stream transmission mode
          Time: 957462494.827802 = (Thu) 2000 May 04 13:48:14.827802
 [dbsldb2. 2141]
C  *** ERROR =>
[000045C8] ICLC2184E Internal error in line 4048 and file ibmiclio.c; Processing continues.
          Details: Transmit failed for performance monitor functions; retcode = -5
          Time: 957462494.827921 = (Thu) 2000 May 04 13:48:14.827921
```

Figure 44. Extract from the tracefile of the work processes - part 5 of 9

```
A  is used for RFC security.
A  RFC Destination> destination sap1tr0_RED_00 host sap1tr0 system RED systnr 0 (sap1tr0_RED_00)
A  RFC Options> H=sap1tr0,S=00
A  RFC Method> initialize RemObjDriver for ABAP Objects
S
S  Thu May  4 13:22:13 2000
S  initialize debug system
T  Stack direction is downwards.
T  debug control: prepare exclude for printer trace
T  new memory block 21277580
S  spool kernel/ddic check: Ok
S  1 spool work process(es) found
S  frontend print via spool service enabled
S  printer list size is 150
S  printer type list size is 50
A  is used for RFC security.
A  RFC Destination> destination sap1tr0_RED_00 host sap1tr0 system RED systnr 0 (sap1tr0_RED_00)
A  RFC Options> H=sap1tr0,S=00
A  RFC Method> initialize RemObjDriver for ABAP Objects
S
S  Thu May  4 13:22:13 2000
S  initialize debug system
T  Stack direction is downwards.
T  debug control: prepare exclude for printer trace
T  new memory block 21277580
S  spool kernel/ddic check: Ok
S  1 spool work process(es) found
S  frontend print via spool service enabled
S  printer list size is 150
S  printer type list size is 50
A  is used for RFC security.
A  RFC Destination> destination sap1tr0_RED_00 host sap1tr0 system RED systnr 0 (sap1tr0_RED_00)
A  RFC Options> H=sap1tr0,S=00
A  RFC Method> initialize RemObjDriver for ABAP Objects
S
S  Thu May  4 13:22:13 2000
S  initialize debug system
T  Stack direction is downwards.
T  debug control: prepare exclude for printer trace
T  new memory block 21277580
S  spool kernel/ddic check: Ok
S  1 spool work process(es) found
S  frontend print via spool service enabled
S  printer list size is 150
S  printer type list size is 50
C  *** ERROR =>
 [dbsldb2. 2141]
C  *** ERROR =>            Details: normal transmission modele
 [dbsldb2. 2141]
B  db_con_set_reco_flag: 0: name = R/3, handle = 0 ----------------------> [2]
B  Reconnect state is entered by connection:
B  0: name = R/3, concnt= 000000 state = ACTIVE      , perm = YES, reco = YES
B  ***LOG BYM=> severe db error 0          work process is in the reconnect state [dbsh     0580]
B  ***LOG BY4=> sql error 0      performing SEL on table D010LINF   [dbrepo   4114]
B  ***LOG BY0=> ICLC2508E Could not send data to ICLI server - Broken pipe (errno=32)=6; ICLC2506E Sending data failed in
transmission phasese;    ;            Details: normal transmission modele; [dbrepo   4114]
B  *** ERROR => missing return code handler
 caller does not handle code 1 from dbrepo.c[2437]
 ==> calling sap_dext to abort transaction
 [dbeh.c   0082]
```

Figure 45. Extract from the tracefile of the work processes - part 6 of 9

```
A
A ** RABAX: could not save part LEV_RX_STANDARD of dump.
B db_con_set_reco_flag: 0: name = R/3, handle = 0
B Reconnect state is entered by connection:
B 0: name = R/3, concnt= 000000 state = ACTIVE      , perm = YES, reco = YES
B ***LOG BYM=> severe db error 0          work process is in the reconnect state [dbsh     0580]
B ***LOG BY2=> sql error 0      performing TIM [dblink   1014]
B ***LOG BY0=>  [dblink   1014]
B db_con_set_reco_flag: 0: name = R/3, handle = 0
B Reconnect state is entered by connection:
B 0: name = R/3, concnt= 000000 state = ACTIVE      , perm = YES, reco = YES
B ***LOG BYM=> severe db error 0          work process is in the reconnect state [dbsh     0580]
B ***LOG BY2=> sql error 0      performing TIM [dblink   1014]
B ***LOG BY0=>  [dblink   1014]
B db_con_set_reco_flag: 0: name = R/3, handle = 0
B Reconnect state is entered by connection:
B 0: name = R/3, concnt= 000000 state = ACTIVE      , perm = YES, reco = YES
B ***LOG BYM=> severe db error 0          work process is in the reconnect state [dbsh     0580]
B ***LOG BY2=> sql error 0      performing TIM [dblink   1014]
B ***LOG BY0=>  [dblink   1014]
C dbdsdb2: DBSL99 SQL0
C
B ***LOG BY2=> sql error 0          performing OPC       [dbds     0462]
B ***LOG BY0=>  [dbds     0462]
C dbdsdb2: DBSL99 SQL0
C
B ***LOG BY2=> sql error 0          performing OPC       [dbds     0462]
B ***LOG BY0=>  [dbds     0462]
C dbdsdb2: DBSL99 SQL0
C
B ***LOG BY2=> sql error 0          performing OPC       [dbds     0462]
B ***LOG BY0=>  [dbds     0462]
C dbdsdb2: DBSL99 SQL0
C
B ***LOG BY2=> sql error 0          performing EXE       [dbds     0809]
B ***LOG BY0=>  [dbds     0809]
A
A ** RABAX: could not save part LEV_SN_PROGRAMS of dump.
C dbdsdb2: DBSL99 SQL0
C

A is used for RFC security.
A RFC Destination> destination sap1tr0_RED_00 host sap1tr0 system RED systnr 0 (sap1tr0_RED_00)
A RFC Options> H=sap1tr0,S=00
A RFC Method> initialize RemObjDriver for ABAP Objects
S
S Thu May  4 13:22:13 2000
S initialize debug system
T Stack direction is downwards.
T debug control: prepare exclude for printer trace
T new memory block 21277580
S spool kernel/ddic check: Ok
S 1 spool work process(es) found
S frontend print via spool service enabled
S printer list size is 150
S printer type list size is 50
A is used for RFC security.
A RFC Destination> destination sap1tr0_RED_00 host sap1tr0 system RED systnr 0 (sap1tr0_RED_00)
A RFC Options> H=sap1tr0,S=00
A RFC Method> initialize RemObjDriver for ABAP Objects
S
S Thu May  4 13:22:13 2000
S initialize debug system
T Stack direction is downwards.
T debug control: prepare exclude for printer trace
T new memory block 21277580
S spool kernel/ddic check: Ok
S 1 spool work process(es) found
S frontend print via spool service enabled
S printer list size is 150
S printer type list size is 50
A is used for RFC security.
A RFC Destination> destination sap1tr0_RED_00 host sap1tr0 system RED systnr 0 (sap1tr0_RED_00)
A RFC Options> H=sap1tr0,S=00
A RFC Method> initialize RemObjDriver for ABAP Objects
S
```

Figure 46. Extract from the tracefile of the work processes - part 7 of 9

```
S  Thu May  4 13:22:13 2000
S  initialize debug system
T  Stack direction is downwards.
T  debug control: prepare exclude for printer trace
T  new memory block 21277580
S  spool kernel/ddic check: Ok
S  1 spool work process(es) found
S  frontend print via spool service enabled
S  printer list size is 150
S  printer type list size is 50
C  *** ERROR =>
 [dbsldb2. 2141]
C  *** ERROR =>           Details: normal transmission modele
 [dbsldb2. 2141]
B  db_con_set_reco_flag: 0: name = R/3, handle = 0
B  Reconnect state is entered by connection:
B  0: name = R/3, concnt= 000000 state = ACTIVE      , perm = YES, reco = YES
B  ***LOG BYM=> severe db error 0          work process is in the reconnect state [dbsh     0580]
B  ***LOG BY4=> sql error 0      performing SEL on table D010LINF   [dbrepo   4114]
B  ***LOG BY0=> ICLC2508E Could not send data to ICLI server - Broken pipe (errno=32)=6; ICLC2506E Sending data failed in
transmission phasese;    ;            Details: normal transmission modele; [dbrepo   4114]
B  *** ERROR => missing return code handler
 caller does not handle code 1 from dbrepo.c[2437]
 ==> calling sap_dext to abort transaction
B  ***LOG BY2=> sql error 0          performing OPC        [dbds     0462]
B  ***LOG BY0=>  [dbds     0462]
C  dbdsdb2: DBSL99 SQL0
C
B  ***LOG BY2=> sql error 0          performing OPC        [dbds     0462]
B  ***LOG BY0=>  [dbds     0462]
C  dbdsdb2: DBSL99 SQL0
C
B  ***LOG BY2=> sql error 0          performing OPC        [dbds     0462]
B  ***LOG BY0=>  [dbds     0462]
B  db_con_set_reco_flag: 0: name = R/3, handle = 0
B  Reconnect state is entered by connection:
B  0: name = R/3, concnt= 000000 state = ACTIVE      , perm = YES, reco = YES
B  ***LOG BYM=> severe db error 0          work process is in the reconnect state [dbsh     0580]
B  hdl_error_on_commit_rollback: Connection details:
B  0: name = R/3, concnt= 000000 state = ACTIVE      , perm = YES, reco = YES
B  ***LOG BY2=> sql error 0      performing COM [dbcon    2283]
B  ***LOG BY0=>  [dbcon    2283]
B  ***LOG BYJ=> Database COMMIT     for connection R/3        failed [dbcon    2295]
A  Thu May  4 13:48:14 2000
A  ABAP/4 Program SAPLSTU3                                .
A  Source  Line 0.
A  Error Code TIME_OUT.
A  Module  $Id: //bas/46B/src/krn/runt/abinit.c#2 $ SAP.
A  Function ab_chstat Line 1426.
A  Error TIME_OUT occured. P1=rdisp/max_wprun_time P2=300.
A
B  db_con_set_reco_flag: 0: name = R/3, handle = 0
B  Reconnect state is entered by connection:
B  0: name = R/3, concnt= 000000 state = ACTIVE      , perm = YES, reco = YES
B  ***LOG BYM=> severe db error 0          work process is in the reconnect state [dbsh     0580]
B  ***LOG BY4=> sql error 0      performing SEL on table D010TINF   [dbrepo   4185]
B  ***LOG BY0=>  [dbrepo   4185]
M  ***LOG R68=> ThIRollBack, roll back () [thxxhead 9404]
B  db_con_set_reco_flag: 0: name = R/3, handle = 0
B  Reconnect state is entered by connection:
B  0: name = R/3, concnt= 000000 state = ACTIVE      , perm = YES, reco = YES
B  ***LOG BYM=> severe db error 0          work process is in the reconnect state [dbsh     0580]
B  hdl_error_on_commit_rollback: DB-ROLLBACK ok. RECONNECT state ----------------------> [3]
B  db_con_reconnect performing the reconnect for con: ------------------------------> [4]
B  0: name = R/3, concnt= 000000 state = INACTIVE    , perm = YES, reco = YES
```

Figure 47. Extract from the tracefile of the work processes - part 8 of 9

```
-------------------------------------------------------------------------------> [5]
C  ********* Profile Parameter for connection 0 **********
C  Connecting to standby-Host
C  DBIF Versions:
C    R/3 Version       : 12 (4.6B)
C    API Version       : 9
C  ICLI-Client Versions:
C    R/3 Version       : 12 (4.6B)
C    API Version       : 9
C    PROTOCOL Version  : 1
C    Internal Version  : 1
C  Workprocess         : 0
C  USE_NO_HOLD         : 0
C  USE_OPT_1_ROW       : 0
C  DB_UID              : SAPR3
C  DB_ID               : RED
C  SAPSYSTEM           : 00
C  Definitions for standby host
C  DB_SSID             : DBK4
C  DBHOST              : wtsc54oe
C  PORT                : 6533
C  CCSID               : 819
C  KEEP_X_LOCK         : 1
C  RO_COMMIT_SUPPRESS  : 0 ----------------------------------------------------------> [6]

C  --------------------------------------
C  Used Bugfixes:
C  **************************************
C  Network-Protocol    : TCP
C  *** ERROR =>
[000045C8] ICLC2520W Security is switched off - it should be switched off on the server too!
          Time: 957462494.914146 = (Thu) 2000 May 04 13:48:14.914146
 [dbsldb2. 2141]
C  Your DB2 system works with ASCII code page 819 and EBCDIC code page 37
B  Connection 0 opened
B  Wp  Hdl ConName             ConCnt ConState     TX  PRM RCT Date     Time ----------> [7]
B  000 000 R/3                 000000 ACTIVE      NO  YES NO  20000504 134814
B  001 000 R/3                 000000 ACTIVE      YES YES NO  20000504 134812
B  002 000 R/3                 000000 ACTIVE      NO  YES NO  20000504 132154
B  003 000 R/3                 000000 ACTIVE      NO  YES NO  20000504 132154
B  004 000 R/3                 000000 ACTIVE      NO  YES NO  20000504 132154
B  005 000 R/3                 000000 INACTIVE    NO  YES NO  20000504 132154
B  006 000 R/3                 000000 INACTIVE    NO  YES NO  20000504 132154
B  007 000 R/3                 000000 INACTIVE    NO  YES NO  20000504 132154
B  008 000 R/3                 000000 ACTIVE      YES YES NO  20000504 132154
B  009 000 R/3                 000000 INACTIVE    NO  YES NO  20000504 132154
B  010 000 R/3                 000000 ACTIVE      NO  YES NO  20000504 132154
B  011 000 R/3                 000000 INACTIVE    NO  YES NO  20000504 132154
B  012 000 R/3                 000000 INACTIVE    NO  YES NO  20000504 132154
B  Successful reconnect for connection: ----------------------------------------------> [8]
B  0: name = R/3, concnt= 000000 state = ACTIVE      , perm = YES, reco = NO
B  ***LOG BYY=> work process left reconnect state [dblink   1671]
A
A  Thu May  4 13:48:15 2000
A  Thu May  4 13:48:15 2000
A  ABAP/4 Program ??????????????????????????????????????????.
A  Source                                      Line 0.
A  Error Code DBIF_REPO_SQL_ERROR.
A  Module  $Id: //bas/46B/src/krn/runt/abdbif3.c#1 $ SAP.
A  Function ab_loadtext Line 241.
A  SQL error 0 occurred when accessing program SAPMS380                                part TEXT..
A
M  ***LOG R68=> ThIRollBack, roll back () [thxxhead 9404]
Y  *** ERROR => invalid parameter [dynpmain 2360]
Y  *** ERROR => invalid parameter [dynpmain 2395]
M
M  Thu May  4 13:48:16 2000
M  *** ERROR => PfStatWrite: missing STAT_END for opcode STAT_REPLOAD [pfxxstat 2912]
S
S  Thu May  4 13:48:23 2000
S  found spool memory service RSPO-ACTIONS at 5cbffeb8
B
B  Thu May  4 13:48:52 2000
B  ***LOG BV4=> reconnect state is set for the work process [dblink   1228]
B  ***LOG BYY=> work process left reconnect state [dblink   1671]
B
B  Thu May  4 13:49:25 2000
B  ***LOG BV4=> reconnect state is set for the work process [dblink   1228]
B  ***LOG BYY=> work process left reconnect state [dblink   1671]
B
B  Thu May  4 13:50:26 2000
B  ***LOG BV4=> reconnect state is set for the work process [dblink   1228]
B  ***LOG BYY=> work process left reconnect state [dblink   1671]
```

Figure 48. Extract from the tracefile of the work processes - part 9 of 9

Appendix D. Special notices

This publication is intended to help database administrators, analysts, consultants, and managers to create and implement availability plans for the database of an SAP R/3 on DB2 UDB for OS/390 installation. The information in this publication is not intended as the specification of any programming interfaces that are provided by SAP R/3, DB2 UDB for OS/390, or OS/390. See the PUBLICATIONS section of the IBM Programming Announcement for the relevant software for more information about what publications are considered to be product documentation. See SAP announcement material for more information about what publications are considered to be product documentation for SAP R/3.

References in this publication to IBM products, programs or services do not imply that IBM intends to make these available in all countries in which IBM operates. Any reference to an IBM product, program, or service is not intended to state or imply that only IBM's product, program, or service may be used. Any functionally equivalent program that does not infringe any of IBM's intellectual property rights may be used instead of the IBM product, program or service.

Information in this book was developed in conjunction with use of the equipment specified, and is limited in application to those specific hardware and software products and levels.

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

Licensees of this program who wish to have information about it for the purpose of enabling: (i) the exchange of information between independently created programs and other programs (including this one) and (ii) the mutual use of the information which has been exchanged, should contact IBM Corporation, Dept. 600A, Mail Drop 1329, Somers, NY 10589 USA.

Such information may be available, subject to appropriate terms and conditions, including in some cases, payment of a fee.

The information contained in this document has not been submitted to any formal IBM test and is distributed AS IS. The information about non-IBM ("vendor") products in this manual has been supplied by the vendor and IBM assumes no responsibility for its accuracy or completeness. The use of this information or the implementation of any of these techniques is a customer responsibility and depends on the customer's ability to evaluate and integrate them into the customer's operational environment. While each item may have been reviewed by IBM for accuracy in a specific situation, there is no guarantee that the same or similar results will be obtained elsewhere. Customers attempting to adapt these techniques to their own environments do so at their own risk.

Any pointers in this publication to external Web sites are provided for convenience only and do not in any manner serve as an endorsement of these Web sites.

Any performance data contained in this document was determined in a controlled environment, and therefore, the results that may be obtained in other operating

environments may vary significantly. Users of this document should verify the applicable data for their specific environment.

This document contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples contain 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.

Reference to PTF numbers that have not been released through the normal distribution process does not imply general availability. The purpose of including these reference numbers is to alert IBM customers to specific information relative to the implementation of the PTF when it becomes available to each customer according to the normal IBM PTF distribution process.

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

IBM ®	Redbooks
	Redbooks Logo 
S/390	Wizard
SP	XT
SP1	3890
SP2	400
Sysplex Timer	Lotus
System/390	Notes

The following terms are trademarks of other companies:

Tivoli, Manage. Anything. Anywhere.,The Power To Manage., Anything. Anywhere.,TME, NetView, Cross-Site, Tivoli Ready, Tivoli Certified, Planet Tivoli, and Tivoli Enterprise are trademarks or registered trademarks of Tivoli Systems Inc., an IBM company, in the United States, other countries, or both. In Denmark, Tivoli is a trademark licensed from Kjøbenhavns Sommer - Tivoli A/S.

C-bus is a trademark of Corollary, Inc. in the United States and/or other countries.

Java and all Java-based trademarks and logos are trademarks or registered trademarks of Sun Microsystems, Inc. in the United States and/or other countries.

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

PC Direct is a trademark of Ziff Communications Company in the United States and/or other countries and is used by IBM Corporation under license.

ActionMedia, LANDesk, MMX, Pentium and ProShare are trademarks of Intel Corporation in the United States and/or other countries.

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

SET, SET Secure Electronic Transaction, and the SET Logo are trademarks owned by SET Secure Electronic Transaction LLC.

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

Appendix E. Related publications

The publications listed in this section are considered particularly suitable for a more detailed discussion of the topics covered in this redbook.

E.1 IBM Redbooks

For information on ordering these publications see “How to get IBM Redbooks” on page 151.

- *SAP R/3 on DB2 for OS/390: Implementing with AIX or Windows NT Applications Servers*, SG24-4945
- *SAP R/3 on DB2 for OS/390: Disaster Recovery*, SG24-5343
- *Database Administration Experiences: SAP R/3 on DB2 for OS/390*, SG24-2078
- *High Availability Considerations: SAP R/3 on DB2 for OS/390*, SG24-2003
- *SAP R/3 on DB2 UDB for OS/390: OS/390 Application Server,* SG24-5840
- *Storage Management with DB2 for OS/390*, SG54-5462

E.2 IBM Redbooks collections

Redbooks are also available on the following CD-ROMs. Click the CD-ROMs button at http://www.redbooks.ibm.com/ for information about all the CD-ROMs offered, updates and formats.

CD-ROM Title	Collection Kit Number
System/390 Redbooks Collection	SK2T-2177
Networking and Systems Management Redbooks Collection	SK2T-6022
Transaction Processing and Data Management Redbooks Collection	SK2T-8038
Lotus Redbooks Collection	SK2T-8039
Tivoli Redbooks Collection	SK2T-8044
AS/400 Redbooks Collection	SK2T-2849
Netfinity Hardware and Software Redbooks Collection	SK2T-8046
RS/6000 Redbooks Collection (BkMgr Format)	SK2T-8040
RS/6000 Redbooks Collection (PDF Format)	SK2T-8043
Application Development Redbooks Collection	SK2T-8037
IBM Enterprise Storage and Systems Management Solutions	SK3T-3694

E.3 Other resources

These publications are also relevant as further information sources:

E.3.1 SAP AG Resources

- *Upgrade to Release 4.6B: OS/390 UNIX System Services*, Material Number 51009018
- *R/3 Installation on UNIX - OS-Dependencies*, Material Number 51008168
- *R/3 Installation on OS/390 UNIX System Services: DB2 UDB for OS/390*, Material Number 51008496

- *SAP R/3 Database Administration Guide: DB2 for OS/390 Release 4.6 B* Material Number 51008497

E.3.2 IBM Resources

- *SAP R/3 on DB2 for OS/390: Planning Guide SAP R/3 4.6B*, SC33-7966
- *System Overview, S/390 9672 Generation 6*, GA22-1030
- *DB2 UDB for OS/390 Data Sharing: Planning and Administration Version 6* SC26-9007
- *Parallel Sysplex - Software Management for Availability,* SG24-5451
- *DB2 UDB for OS/390 Administration Guide*, SC26-9003
- *DB2 UDB for OS/390 Utility Guide and Reference*, SC26-9015

E.4 Referenced Web sites

These Web sites are also relevant as further information sources:

- `http://techsupport.services.ibm.com/support/s390`

  S/390 TechSupport Online

- `http://www.storage.ibm.com/hardsoft/diskdrls/technology.htm`

  Technology for disk storage systems

- `http://www-4.ibm.com/software/data/db2/os390/pdf/oreorg.pdf`

  White paper: *Implementing Online Reorg in a Production Environment*

How to get IBM Redbooks

This section explains how both customers and IBM employees can find out about IBM Redbooks, redpieces, and CD-ROMs. A form for ordering books and CD-ROMs by fax or e-mail is also provided.

- **Redbooks Web Site** ibm.com/redbooks

  Search for, view, download, or order hardcopy/CD-ROM Redbooks from the Redbooks Web site. Also read redpieces and download additional materials (code samples or diskette/CD-ROM images) from this Redbooks site.

  Redpieces are Redbooks in progress; not all Redbooks become redpieces and sometimes just a few chapters will be published this way. The intent is to get the information out much quicker than the formal publishing process allows.

- **E-mail Orders**

  Send orders by e-mail including information from the IBM Redbooks fax order form to:

	e-mail address
In United States or Canada	pubscan@us.ibm.com
Outside North America	Contact information is in the “How to Order” section at this site: http://www.elink.ibmlink.ibm.com/pbl/pbl

- **Telephone Orders**

United States (toll free)	1-800-879-2755
Canada (toll free)	1-800-IBM-4YOU
Outside North America	Country coordinator phone number is in the “How to Order” section at this site: http://www.elink.ibmlink.ibm.com/pbl/pbl

- **Fax Orders**

United States (toll free)	1-800-445-9269
Canada	1-403-267-4455
Outside North America	Fax phone number is in the “How to Order” section at this site: http://www.elink.ibmlink.ibm.com/pbl/pbl

This information was current at the time of publication, but is continually subject to change. The latest information may be found at the Redbooks Web site.

IBM Intranet for Employees

IBM employees may register for information on workshops, residencies, and Redbooks by accessing the IBM Intranet Web site at http://w3.itso.ibm.com/ and clicking the ITSO Mailing List button. Look in the Materials repository for workshops, presentations, papers, and Web pages developed and written by the ITSO technical professionals; click the Additional Materials button. Employees may access MyNews at http://w3.ibm.com/ for redbook, residency, and workshop announcements.

IBM Redbooks fax order form

Please send me the following:

Title	Order Number	Quantity

First name	Last name	
Company		
Address		
City	Postal code	Country
Telephone number	Telefax number	VAT number

☐ Invoice to customer number

☐ Credit card number

Credit card expiration date	Card issued to	Signature

We accept American Express, Diners, Eurocard, Master Card, and Visa. Payment by credit card not available in all countries. Signature mandatory for credit card payment.

Abbreviations and acronyms

ABAP	advanced business application programming (programming language used by SAP)
AC	alternating current
ACS	access control system
AIX	Advanced Interactive Executive (an IBM version of the UNIX operating system)
APAR	authorized program analysis report
ARM	Automatic Restart Manager
BSDS	bootstrap dataset (DB2)
CA	continuous availability
CEC	central electronics complex
CF	Coupling Facility
CFRM	Coupling Facility Resource Manager
CHPID	channel path id
CMOS	complementary metal oxide semiconductor
CO	continuous operation
CPU	central processing unit
DASD	direct access storage device
DB2	Database 2 (an IBM relational database management system)
DB2 UDB for OS/390	Database 2 universal database for OS/390 (OS/390 implementation of the DB2 product)
DBA	database administrator
DBD	database descriptor
DBMS	database management system
DC	direct current
DCE	distributed cell environment
DDIC	data dictionary
DDL	data definition language
DML	data manipulation language
DR	disaster recovery
DVG Hannover	(Datenverarbeitungsgesel--schaftt Hannover) German Banking IT solutions company
ETR	external time reference
ECC	error correction code
ENQ	enqueue
ERP	enterprise resource planning
ESCON	enterprise systems connection (architecture, IBM S/390)
ESS	Enterprise Storage System (IBM System/390)
FDDI	Fiber Distributed Data Interface (100 Mbits/s fiber optic LAN)
FICON	fiber connection
FRR	functional recovery routine
Gb	gigabit (one Gb =10^9 bits)
GB	gigabyte (one GB =10^9 bytes)
GBP	group buffer pool
HA	high availability
HACMP	high availability cluster multi-processing (AIX)
HFS	hierarchical file systems
HiPerLink	high performance coupling link
I/O	input output
IBM	International Business Machines Corporation
IC	internal coupling
ICB	integrated cluster bus
ICF	1.integrated catalog facility (OS/390) 2. internal coupling facility (S/390 architecture)
ICLI	integrated call level interface
ID	identifier
IDCAMS	the program name for access method services (OP SYS)
IPL	initial program load
IRLM	integrated resource lock manager
IT	information technology
ITSO	International Technical Support Organization
JCL	job control language
LAN	local area network
LIC	licensed internal code
LPAR	logical partition

LRSN log record sequence number

LUW logical unit of work

MCM multiple chip module

MTBF mean time between failures

MVS multiple virtual storage (IBM System 370 & 390)

OS/390 Operating System (for the) IBM System/390

OSA open systems adapter (IBM System/390)

OSS online support service (SAP)

PPRC peer-to-peer remote copy

PR/SM processor resource/system manager

PSSP AIX Parallel System Support Programs (IBM program product for SP1 and SP2)

PTF program temporary fix

PU processor unit

R/W read-write

RACF Resource Access Control Facility

RAMAC Raid Architecture with Multi-Level Adaptive Cache

RAS reliability, availability, and serviceability

RBA relative byte address

RFC remote function call

RID Record IDentifier (DB2)

RRS recoverable resource services

RRSAF Recoverable Resource Manager Services Attachment Facility

RVA RAMAC Virtual Array

SAP 1.Systems, Applications, Products in Data Processing (software vendor) 2.system assist processor (S/390 architecture)

SAP AG SAP Aktiengesell (German name of software vendor)

SAP R/3 SAP Release 3 (software product of SAP)

SD Sales and Distribution (SAP application)

SMP/E system modification program/extended (MVS)

SMS System Managed Storage

SP IBM RS/6000 Scalable POWERparallel Systems (RS/6000 SP)

SPOF single point of failure

SQL structured query language

SRM system resources manager

STCK store clock

STI self-timed interface

TB terabyte (one TB = 10^{12} bytes)

TCP/IP Transmission Control Protocol/Internet Protocol

TPF Transaction Processing Facility

UDB universal database

UNIX An operating system developed at Bell Laboratories

UPS uninterruptible power supply

VSAM Virtual Storage Access Method (IBM)

WLM Workload Manager

XES cross-system extended services (MVS)

Index

T

U

V

W

Z

IBM Redbooks review

Your feedback is valued by the Redbook authors. In particular we are interested in situations where a Redbook "made the difference" in a task or problem you encountered. Using one of the following methods, **please review the Redbook, addressing value, subject matter, structure, depth and quality as appropriate.**

- Use the online **Contact us** review redbook form found at **ibm.com**/redbooks
- Fax this form to: USA International Access Code + 1 845 432 8264
- Send your comments in an Internet note to redbook@us.ibm.com

Document Number **Redbook Title**	SG24-5690-00 SAP R/3 on DB2 UDB for OS/390: Database Availability Considerations
Review	
What other subjects would you like to see IBM Redbooks address?	
Please rate your overall satisfaction:	O Very Good O Good O Average O Poor
Please identify yourself as belonging to one of the following groups:	O Customer O Business Partner O Solution Developer O IBM, Lotus or Tivoli Employee O None of the above
Your email address: The data you provide here may be used to provide you with information from IBM or our business partners about our products, services or activities.	O Please do not use the information collected here for future marketing or promotional contacts or other communications beyond the scope of this transaction.
Questions about IBM's privacy policy?	The following link explains how we protect your personal information. **ibm.com**/privacy/yourprivacy/

IBM Redbooks

SAP R/3 on DB2 UDB for OS/390: Database Availability Considerations

(0.2”spine)
0.17”<->0.473”
90<->249 pages

SAP R/3 on DB2 UDB for OS/390: Database Availability Considerations

Define and use DB2 data sharing to improve db availability

Exploit the Split Mirror Backup Facility

Parallel Sysplex configuration examples

This IBM Redbook describes availability procedures for the database of an SAP R/3 4.6B installation that uses DB2 UDB for OS/390 as the SAP R/3 database server. The redbook is one of a series that focuses on SAP R/3 when DB2 UDB for OS/390 is used as a database server.

This redbook reflects the considerable development effort, both by SAP AG and by IBM, that has been focused on the area of database availability due to the reliance on SAP R/3 by cutomers who expect a high degree of availability, in some cases 24 hours per day, every day of the year.

Since the subject of this book is database availability, we concentrate on the database that applies to this solution, DB2 UDB for OS/390. Much effort was spent on looking at data sharing with DB2 to see how that improves availability of the database, and to investigate how it is defined and used. We also review database administrative activity that must be periodically run, to see how that can be done with as little disruption as possible.

We also investigate the Split Mirror Backup Facility as defined by SAP R/3 and implemented with DB2 UDB for OS/390. A particular area of interest is the minimizing of time to perform a backup, since this regularly scheduled activity can be disruptive. Obviously, no discussion of backup is complete without a discussion of using the backups for recovery, so we address the subjects of full and partial recovery as well.

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

SG24-5690-00 ISBN 0738417181