IBM, PeopleSoft, and Société Générale Business Intelligence Solution

A Performance Analysis of PeopleSoft8 Enterprise Performance Management with DB2 UDB EEE and IBM @server pSeries 690

April 2002

Acknowledgements

Authors

Denis Grimaud, PeopleSoft Consulting
Jacques Milman, IBM Global Business Intelligence Solutions

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

Société Générale

Jean-Louis Tribut: project director
Lotfi Klouche: project manager
Jean-Pierre Marcoff: project sponsor
Yves Lambert: architect
Oussama Charabeh: functional expert
Christelle Josserand: functional expert
Isabelle Albinet: integration manager
Christian Sonolet: AIX/DB2 expert

PeopleSoft

Daniel Moukouri: PeopleSoft Consulting
Christian Watbled: PeopleSoft Consulting
Michael Dipersio: PeopleSoft Consulting
Kottresh Kogali: PeopleSoft Consulting
Olivier Tolon: PeopleSoft Consulting
Prasada Alokam: PeopleSoft Consulting
Jim Errant: PeopleSoft Product Strategy

IBM

Dominique Auclair: GBIS Sales
Joel Bogart: pSeries Benchmark Center
Guy Delaporte: GBIS
Priti Desai: DB2/PeopleSoft Benchmark Group
Rosa Fernandez: IBM pSeries
Eric Fleury: Data Management/DB2
Alain Lot: IBM, GBIS Technical Operations
Bob Mc Namara: Business Intelligence Teraplex Center
Merrilee Osterhoudt : Business Intelligence Teraplex Center
Ron Sherman: DB2 WW Benchmark Group
Nick Venidis: pSeries Benchmark Center

IBM, PeopleSoft, and Société Générale formed a team with a diversity of skills including project architecture, systems integration, Business Intelligence solution architecture, EPM design, implementation and tuning, DB2 database design, optimization, and administration.

This document is provided with permission by IBM, PeopleSoft, and Société Générale Corporations. The information cannot be used or copied without inclusion of the Notices section from the back of this document.

Executive summary

Société Générale asked IBM and PeopleSoft for a solution to track their business results. It required large volumes of data to be processed quickly to meet their reporting schedule. The joint recommendation included an IBM @server pSeries™ 690 hosting DB2® UDB and PeopleSoft8 Enterprise Performance Management (EPM). This system was tested and found to exceed the customer's expectations.

Business and technical drivers

Founded in 1864 and headquartered in Paris, France, Société Générale is one of the leading banks in the euro zone. The Group employs 80,000 people worldwide.

The bank launched a complete re engineering of its organization, the 4D program (**D**ispositif de **D**istribution de la Banque de **D**étail de **D**emain), with the objective of reorganizing the information system in the context of a multi-channel organization (e.g. branch offices, call centers, Internet). The Pilotage project, one of the six projects of the 4D program, focused on the adoption of their day-to-day reporting and analysis tools to a multi-channel distribution framework.

IBM/PeopleSoft solution

Société Générale selected PeopleSoft Enterprise Warehouse as their platform for enterprise-wide business analysis. Of key concern was their ability to track business results, evaluate the efficiency of deployed resources, and monitor the performance of critical EPM batch processing.

Validating the solution

Prior to the solution implementation, Société Générale requested a performance test in order to assess the capabilities of the EPM solution to handle the batch processing with the high volumes of data associated with the Pilotage 4D project, within the processing window available during the bank's reporting schedule. The performance test also had to demonstrate the solution performance and throughput in a technical environment similar to that being planned for the Pilotage 4D project.

The focus of the performance test was confined to the financial reporting processes. Test scenarios spanned the monthly processing cycle in its entirety, starting from the ETL processing for the initial load of the enterprise warehouse to the building of data marts. These steps included the following:

- Extraction of data into a staging area by DB2 Autoloader
- ETL data transformation and cleansing
- EPM table load by DB2 Autoloader
- Analytical processing and Data Mart build by EPM

The configuration consisted of an IBM pSeries 690 server with 12 processors, 12 terabytes (TB) of IBM ESS storage system, PeopleSoft EPM V8.3, DB2 UDB Enterprise Extended Edition V7.2 and PowerMart ETL for PeopleSoft. Details on the configuration can be found in the appendices.

Société Générale provided representative sample data and the expansion rules for generating the target volumes. A database representing eight months of history data was built utilizing a total of 2.4 TB of disk for the enterprise warehouse and final set of data marts.

Summary of results

Performance results exceeded Société Générale's expectations. Table 1 lists processes, target performance, and performance results. Times are in hours and minutes.

Table 1 Summary of results

Process	Target	Result	Margin
Customers and Services Operations monthly loads	12:00	6:39	42.92%
Analytic computation processing	3:00	2:53	3.89%
Data Mart builds and updates	3:00	2:07	29.44%
Total	**18:00**	**11:31**	**36.02%**

Detailed results are documented in the main body of this paper.

Recommendations and lessons learned

One of EPM's strengths is the ability to allow parallelized analytical computations in the Data Manager process, using job-streaming techniques. Tests run at the conclusion of the performance test demonstrated that using job streams could further optimize many Data Manager processes, but not all. It is recommended that you run tests during design and implementation in order to determine when job streaming can be beneficial for individual processes, and evaluate when indexing can be utilized.

Conclusions

Société Générale validated that PeopleSoft EPM, running on an infrastructure of the IBM DB2® Universal Database Enterprise Edition Extended and pSeries™ 690 server, utilizing the Enterprise Storage Solution disk, was a robust and strategic solution that would support the needs and future goals of the Pilotage 4D project.

Contents

Introduction

This paper documents a validation test conducted by IBM, PeopleSoft and Société Générale at IBM's pSeries Business Intelligence Teraplex Center, utilizing PeopleSoft Enterprise Performance Management solution with IBM DB2 Universal Database Enterprise Edition Extended (DB2 UDB EEE) and @server pSeries 690 server.

The purpose of this paper is to provide readers with information about the performance and operational characteristics of an IBM/PeopleSoft Business Intelligence solution environment.

The intended audience for this document includes database and data warehouse architects, database performance specialists and solution integrators.

This paper begins with a summary of Société Générale's business and technical drivers and a description of the IBM and PeopleSoft solution, followed by an overview of the performance test plan and the test environment. The remainder of the paper details the test methodologies and test results. The paper concludes with a discussion of the lessons learned and the final conclusions based on the results of the tests.

Background

Société Générale is one of the leading banks in the euro zone. The Group employs 80,000 people worldwide in three core businesses:

- Retail banking: 13.4 million customers in France and abroad
- Asset management and private banking: with EUR 298 billion in assets under management at the end of 2001, the Group is the third largest euro-zone bank in asset management
- Corporate and investment banking: SG is the fourth largest player in the euro zone by net banking income

Société Générale is included in the world's four major sustainable development indices.

Société Générale business drivers

The business strategy of Société Générale is underpinned by four core values:

- **Selectivity**

  The organization is developing its activities selectively in terms of its businesses, markets and customers.

- **Development**

  Société Générale will continue to develop through a combination of organic growth, acquisitions, and the creation of new activities.

- **Innovation**

  Innovation lies at the heart of the development of the Group's product mix and the adjustment of its distribution channels, in particular the Internet.

- **Sustainable profitability**

  This is achieved by enhanced productivity and effective risk management.

The bank launched a complete reengineering of its organization, the 4D program (**D**ispositif de **D**istribution de la Banque de **D**étail de **D**emain, which translates to Tomorrow's Consumer Banking Dispatching Architecture) with an objective of reorganizing the information system in the context of a multi-channel organization (e.g. branch offices, call centers, Internet).

The Pilotage project, one of the six projects of the 4D program, focused on the adoption of their day-to-day reporting and analysis tools to a multichannel distribution framework, including four key elements:

- Bank staff activity: trading and non-trading activity
- Global profitability: customers, offerings, structures, actions
- Risks
- Quality

Enterprise Performance Management (EPM) from PeopleSoft was selected by Société Générale as the tool for enterprise-wide business analysis, to provide management with the capability to track business results, evaluate the efficiency of deployed resources, and measure their performance.

Société Générale technical drivers

With the large volumes of data that would be loaded and processed, the performance and management of critical EPM batch processing was a key concern. The underlying database and data processing infrastructure had to provide the performance, reliability and manageability required to handle the high volume of data within the strict confines of the bank's reporting schedule.

With these requirements in mind, the IBM DB2 UDB EEE and pSeries 690 server were obvious choices for the processing infrastructure, with Enterprise Storage Solutions providing the disk.

IBM and PeopleSoft solution

The PeopleSoft and IBM solution proposed for the Pilotage 4D project included the following key hardware and software components:

- PeopleSoft EPM/Enterprise Warehouse V8.3
- DB2 UDB/EEE V7.2 Database System
- pSeries 690 server running AIX 5.1
- ESS 2105-F20 Storage System

PeopleSoft EPM

The PeopleSoft8 EPM suite of performance management applications was developed to maximize the competitive advantage, profitability, and value of the enterprise. EPM facilitates optimal decision making in terms of cost and risk management, product and channels mix optimization, and client relationship value. The EPM suite is integrated on PeopleSoft Enterprise Warehouse (EW), a high-performance warehouse platform for enterprise business intelligence.

Enterprise Warehouse architecture

The Enterprise Warehouse information architecture includes an Operational Data Store (ODS) for near-real-time transaction-level consolidation, Data Warehouse for time-series analysis, and data marts for subject and role-based analysis. This architecture provides the foundation for reporting, planning, and strategic applications.

PeopleSoft8 EPM and Enterprise Warehouse product details can be found at the PeopleSoft Web site at:

`http://www.peoplesoft.com/`

IBM infrastructure

IBM recommended DB2 UDB Enterprise Extended Edition (DB2 UDB EEE), running on IBM pSeries 690, with the AIX operating system, backed by IBM Enterprise Storage System.

DB2 UDB EEE Version 7.2

DB2 UDB is the IBM VLDB (Very Large Database) offering. DB2 has been optimized for business intelligence architectures such as data warehouses and data marts. These have many integrated functions that specifically address decision-support solutions such as:

- High volume databases
- Large numbers of concurrent users
- Mixed workloads
- Complex query processing
- Multidimensional applications

DB2 provides significant systems management efficiencies and functionality to minimize the administrative costs of high volume databases, such as:

- Automatic processing parallelization
- Advanced statistic optimizer
- Administration tools ease-of-use

Details on the DB2 UDB EEE architecture can be found on IBM's Web site at:

`http://www.ibm.com/software/db2/`

IBM pSeries 690 server

The pSeries 690 server is the first member of the IBM UNIX 64-bit Symmetric MultiProcessing (SMP) server family equipped with POWER4 technology processors. This is a multipurpose server, well suited for decision support as well as transactional commercial applications. The pSeries 690 has been designed with many of the industry-leading features of the IBM zSeries™ servers that facilitate consolidation of critical applications onto a single server, such as Logical Partitioning (LPAR), Auto-configuration, Auto-healing and Auto-optimization.

Details on the pSeries 690 architecture and features can be found in Appendix E, "The IBM pSeries 690" on page 37.

AIX 5L operating system

AIX V5.1 was used for this project. Future versions, with new Web-based systems management functions and workload management enhancements, will support 64-bit UNIX for POWER and IA-64 processors.

Enterprise Storage Systems

The IBM 2105 Enterprise Storage Server (ESS) facilitates the consolidation of data from multiple server platforms (S/390, UNIX, Windows, AS/400) simultaneously on a single storage system. ESS provides high performance and addresses the requirement for high bandwidth that characterizes Business Intelligence applications. With a usable storage capacity of 420 GB to 22.4 TB with RAID5 security, ESS facilitates the consolidation of diverse workloads with varied activity patterns without mutual disruption.

Details on the Enterprise Storage Systems architecture can be found on IBM's Web site at:

`http://www.storage.ibm.com`

Validating the solution

Prior to the solution implementation, Société Générale requested a validation test to prove the capabilities of the solution. The validation test, to be conducted on the configuration proposed for Pilotage 4D, would include the following activities:

- Building an application environment similar to the one that will be implemented at the Customer site
- Building data volumes similar to those expected for the project, starting from test cases provided by Société Générale, and using a predefined data model
- Implementing and optimizing the processes to be measured, following scenarios provided by Société Générale
- Communicating test measurements in the context of Société Générale's expectations

> **IBM Business Intelligence Teraplex Centers**
>
> The IBM Business Intelligence Teraplex Centers are facilities established by IBM to prove very large data warehouse implementations on IBM server platforms.
>
> Teraplex Centers are used to integrate and test IBM hardware and software solutions, along with products developed by IBM Business Partners, for very large, end-to-end customer solutions. Large-scale customized proofs of concept are executed using real customer data on very large-scale configurations.
>
> To learn more about the IBM Business Intelligence Teraplex Centers, ask your IBM representative or visit the Web site:
> `http://www.ibm.com/software/bi/teraplex`

IBM and PeopleSoft teamed with Société Générale to conduct the test at the pSeries Business Intelligence Teraplex Center in Poughkeepsie, New York.

Test objectives

The objective of the test was to validate the capabilities of the EPM solution to handle the batch processing of high volumes of data associated with the Pilotage 4D project, within the processing window available to meet the bank's reporting schedule. The test had to demonstrate the solution's performance and throughput in a technical environment similar to the one planned for the Pilotage 4D project.

Test scope

Taking into account the banking priorities, and the critical factor attached to building the processes of financial reporting, the test scope was limited to financial piloting.

The test scenario spanned the monthly processing cycle in its entirety, starting from the initial load to the building of data marts. A test database would be created with an equivalent of eight months of data history.

Success criteria

The processing time for each performance test phase should not exceed the following durations:

- Customers and Services/Operations monthly loads: 12 hours
- Analytic computation processing: 3 hours
- Data mart builds and updates: 3 hours

Test team

A team with members from IBM, PeopleSoft and Société Générale brought together a diversity of skills including Pilotage 4D project architecture, systems integration, Business Intelligence solution architecture, EPM design, implementation and tuning, DB2 database design, optimization and administration.

Test configuration

Figure 1 shows an overview of the validation test hardware and software configuration.

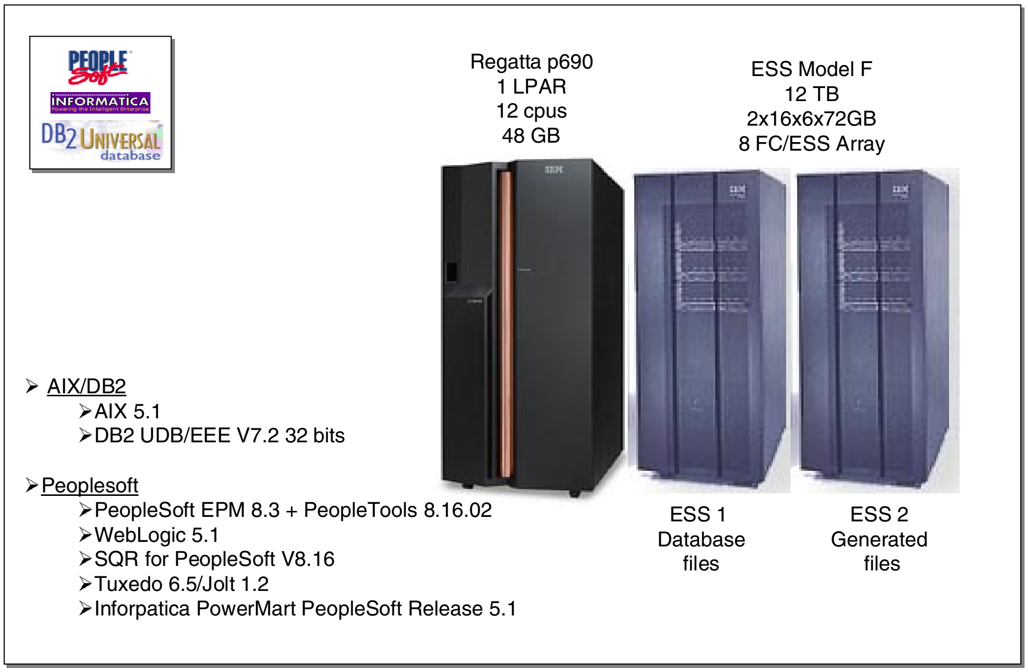

Figure 1 Performance test configuration

Processors

The tests were run on a pSeries 690-681 with 32 1.3 GHz rs and 128 gigabytes of real memory. The tests were run inside a unique logical partition, using 12 processors spread across 2 MCM and 48 GB of memory. Refer to "Performance test configuration details" on page 28 for details on the server configuration.

Disk

The server was configured with forty 36 GB internal disks, a total of 720 GB useful storage secured in RAID/1.

Two ESS 2105-F20 units were used. The first one hosted the DB2 database (data + log) and the ETL files, and the second was used for data generation. Each unit was directly connected to the pSeries server with 8 Fiber Channel links, in a point-to-point topology. Refer to "Performance test configuration details" on page 28 for details on the disk configuration.

Software

IBM

- AIX Version 5.1
- DB2 Universal Database EEE V7.2

PeopleSoft

- EPM/Enterprise Warehouse 8.3
- PeopleTools 8.16.02

Other

- WebLogic 5.1
- Tuxedo 6.5/Jolt 1.2
- ETL

Data

The amount of data generated for the performance test was based on estimations of the actual volume of data that was expected in the framework of the monthly customer profitability processing. A history base covering eight months was loaded into the Enterprise Warehouse before the monthly processing was launched. Société Générale provided Excel source files for 1/10000 of expected volumes (less than 1 MB) and expansion rules for generating the target volumes.

A total of 2.4 terabytes of disk was utilized to build the enterprise warehouse and final set of data marts.

Five files provided by Société Générale were loaded into the Enterprise Warehouse:

- **Customers:** to be loaded in the CUSTOMER_D00 table.
- **Services:** to be loaded in the FI_INSTR, FI_ISTATUS, FI_IBALANCE, FI_IINC, and FI_IVOL tables.
- **Transactions:** to be loaded in the FI_IVOL, FI_IINC tables.
- **Refinancing:** to be loaded in the FI_IINC table.
- **Mailing Customer-Teller/Customer:** to be loaded in a table created specifically to handle customer transcoding.

The loading process also included the following actions:

- Transcoding (customer code, product code)
- Controls (such as date coherency with considered timeframe, dimensions instantiated in the repository, services instantiation)
- Aggregations (aggregation of transactions by date) as well as data enrichments. Error tables are updated in case of failure.

Analytic computations

The 40 million instrument (account) files and corresponding transaction files are loaded in the database. They are used as input for the customer profitability computation process and for the management cost calculation of services and operational costs.

For the costs, four functional rules have to be implemented by the Data Manager tool:

- Management cost of each instrument
- Opening cost of each instrument
- Transaction cost
- Capital needs refinancing cost computation

The first three functional rules are processed by a single Data Manager rule. A specific rule has been created for the last computation. A fifth rule is used, allowing the load of the

PF_LEDFER_F00 table (main source for the profitability data mart) and the merge if necessary of various entities used in case of parallel processing (job streams).

Data mart creation

The data loaded by ETL, plus new data attributes calculated by the Data Manager processes, is used as source to a data mart analyzing the profitability by market segment:

- Dimensions: customers, products, network, time, section, department
- Facts: volume of services, volume of transactions, average balance, profitability elements

Descriptions of the tables, records and EPM enterprise warehouse history log can be found in Appendix C, “Data Structures” on page 31.

Test strategy

During the planning phase in France, tests were conducted on a small sample of the data in order to validate the installation and functionality before going to the Teraplex. Prior to the actual start of the validation test, a second set of tests were conducted at the Teraplex using a subset of the total volumes in order to establish baseline performance numbers for the three target processes of ETL for monthly loads, analytical computation processing, and data mart build and update. The final validation test was executed using the full eight months of data.

Prior to the start of the test, the test team identified the key elements on the critical path of the application architecture, determined the appropriate DB2 configuration, and developed the disk placement strategy.

Application architecture

The application contained four components: DB2 Autoloader, ETL, a Data Manager, and a data mart builder. Figure 2 on page 12 is a representation of the processing chain that was tested during the performance test.

DB2 Auto Loader

Before being processed by ETL, DB2 Autoloader would load the data files into a set of DB2 tables with identical structures. This set of tables is referred to as the “Staging Area”. The load is performed using the fast-loading DB2 autoloader tool.

PowerMart ETL for PeopleSoft

ETL handles the entire functional logic for:

- Data mapping between data source structures and table structures to be loaded
- Transcoding of customer and product codes
- Checking for availability of dimensions related to the repositories
- Checking the date coherency with processed timeframe
- Enriching transactions and refinancing files with the involved service dimension
- Placing invalid records in error tables
- Aggregating transactions by date
- Spreading data on several BU if needed (see the following section)

The output of the ETL processing consists of files in a format similar to the one used by target tables. These files, containing transformed data, are then loaded in the EPM tables by the DB2 autoloader without further transformation.

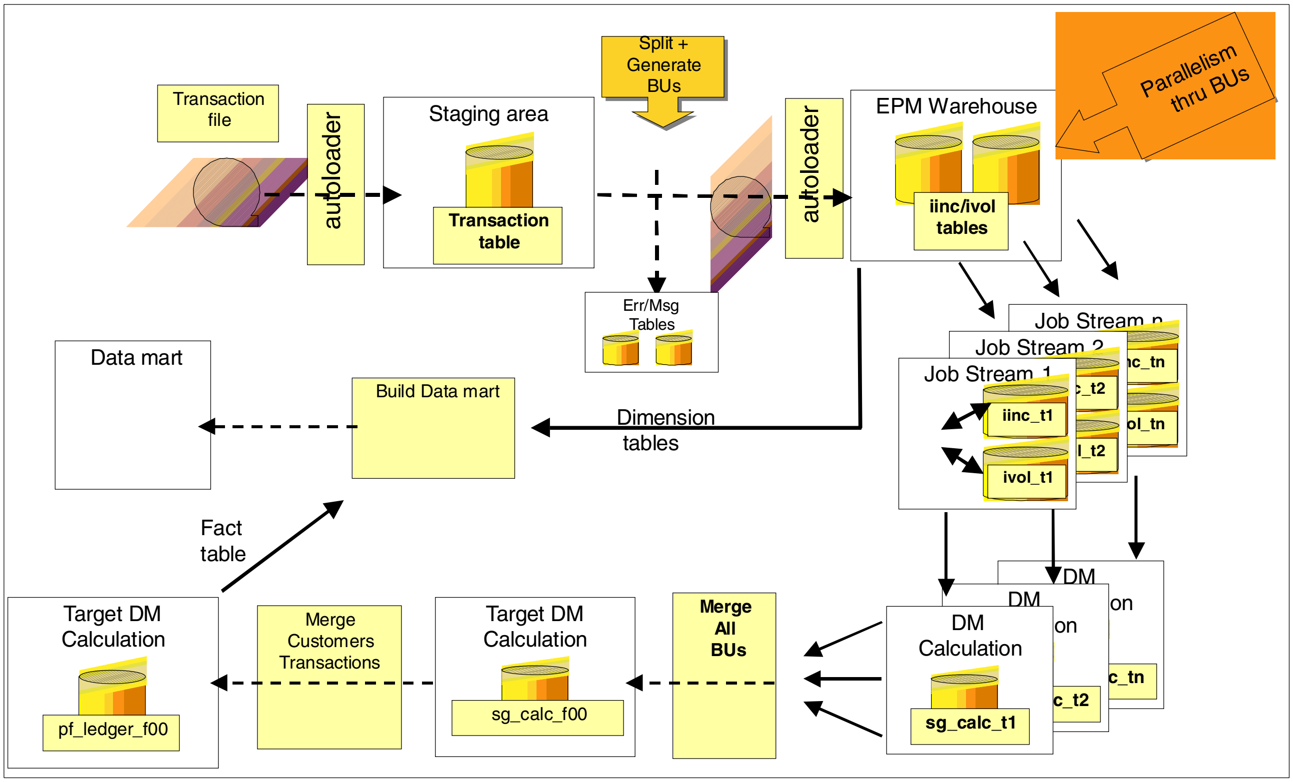

Figure 2 Application architecture

Data Manager

- The Data Manager performs the analytic computation processing. The Business Unit (BU) can do this in parallel job streams, if necessary. A group of rules is applied, aiming at computing the costs and aggregating the net banking revenue (PNB) elements.
- Each job stream copies the data to a temporary table (nom_de_la_table_Tn), then computes, then writes the results in the target table SG_CALC_F00. This table contains results spread across one or more BUs, depending on the level of parallelism.
- A second Data Manager process allows you to merge the data from a single BU into the target table PF_LEDGER.

Data mart builder

Data mart builder is the process for loading the data mart (fact and dimension tables), in an optimized way, for the query processing. It reads the repository tables to load the dimension tables. The fact tables are loaded by:

- PF_LEDGER for the Profitability Elements table
- FI_INSTR, FI_IVOL and SG_IBAL for the volume and average balance fact tables.

Note: If a single BU is used, this second processing is useless. In this case, the target table from the first group of rules is PF_LEDGER.

DB2 configuration

DB2 was configured to optimize utilization of processing resources such as CPU, memory, and available disk (12 CPUS, 48 GB memory and 128 disks). The following were defined:

- Optimized number of DB2 partitions
- Data placement method on ESS disks
- Partitioning keys for DB2 tables
- Indexing method

A discussion on DB2 instance partitioning architecture is provided in Appendix D, "DB2 partition disk mapping" on page 34.

Disk placement

The objective for the disk configuration was to make sure the workload would be evenly spread across the ESS disks (control units, disk adaptors, RAID racks) in order to:

- Spread I/Os across all the disk drives to maximize response times and avoid wait I/Os during large reads or updates.
- Facilitate DB2 prefetch. Even though AIX sees the physical disks as clusters (LUNs on 6 disk arrays) and ESS manages parallel disk access, it is preferable to create several containers per tablespace and per partition to allow DB2 prefetch mechanisms, which are highly recommended when performing large scans and mass inserts.

These objectives were achieved by spreading the data (data, index and temporary space) evenly across the disk units.

Details on the mapping of the DB2 partitions to the ESS disk are provided in Appendix D, "DB2 partition disk mapping" on page 34.

Physical model

When loading the data, decisions were made regarding partitioning, indexing, and log management.

Table partitioning

Given the large number of tables generated by EPM at installation time (more than 21000), only the major tables we used for this performance test were partitioned. Those were the staging stables, the EWH tables and the data mart tables. All other tables were defined with the default options used by the EPM installation process.

With DB2 UDB/EEE, table partitioning is based on a hash-coding mechanism. This technique uses one or more columns to calculate the partition where a row should be stored. To decide on the partitioning key, two criteria should be taken into account:

- High cardinality

  To insure even distribution of data and maximize parallelism, it is recommended that you choose a column with high cardinality (large number of distinct values).

- Local process

  In general, it is recommended that tables be partitioned based on popular join columns. Thus the joins can execute locally in the partition before merging the results.

Two key categories of tables have been identified, based on different partitioning logic:

- The EWH tables, which are instrument-driven, are partitioned on the column «fi_instrument_id ».
- The data mart tables, which are customer-driven, are partitioned on the column «cust_id ».

Small tables that do not benefit from parallelism have been created on non-partitioned tablespaces.

Indexing

In the Pilotage 4D project, most EPM processes scan entire tables. In this context, defining indexes does not make the system perform better because the DB2 optimizer favors prefetch mechanisms that do not use indexes.

Therefore, we dropped all unused indexes and kept only unique ones, which insure the uniqueness of primary keys and thus enable the consistency of the database.

Log management

To manage aging data in the large historic tables of the EW, the oldest month data is removed from the database each month and the current month data is added. Because these tables can be very large (4.6 billion rows and more than 450 GB of index data for the largest table, PS_FI_IINC_R00), the process of cleaning aged data can consume a lot of resources.

In order to simplify the management of the tables, we used UNION ALL views, using a separate table for each processing period. We then created a view that does the UNION of all the tables. The application has access to the view, and not to the physical tables. This technique offers a lot of flexibility because it allows segmentation into smaller jobs of heavy maintenance tasks, such as data population, indexing, statistics collection, and deletes.

In the performance test environment, 3-month periods were used as follows:

- Each month new data is loaded into the table, corresponding to the most recent period.
- Every 3 months, the oldest table is dropped.
- After each update process, a view is recreated to UNION all the period tables of the database.

With period tables, we do not need to use deletes to purge data. Deletes can be very expensive when a table has several indexes. Instead, we DROP the entire period table.

To optimize table access, constraints are defined on date columns (column asof_dt or pf_trans_dt). Constraints allow the DB2 optimizer to eliminate certain tables in the access path selection if the statement violates the restriction on the time period. These optimizations are transparent to the applications, such as EPM or query tools.

Test results

Prior to taking performance measurements on the target volumes, the application processes were run against small data volumes and the results extrapolated in order to estimate the elapse times that would be expected when the processes were run against the actual volumes. If the extrapolations were not within an acceptable range, reoptimization was done and the processes were rerun against the small volumes; those results were then extrapolated. This approach allowed the test team to achieve optimal performance with a single final run against the full volume of data.

The following table summarizes the overall results achieved in each processing category. All times are expressed in hours and minutes.

Process	Target	Result	Margin
ETL	12:00	6:39	42.92%
Data Manager	3:00	2:53	3.89%
Data mart	3:00	2:07	29.44%
Total	18:00	11:31	36.02%

All results met expectations, and performance measured for each of the individual processes exceeded expectations, as shown in Figure 3.

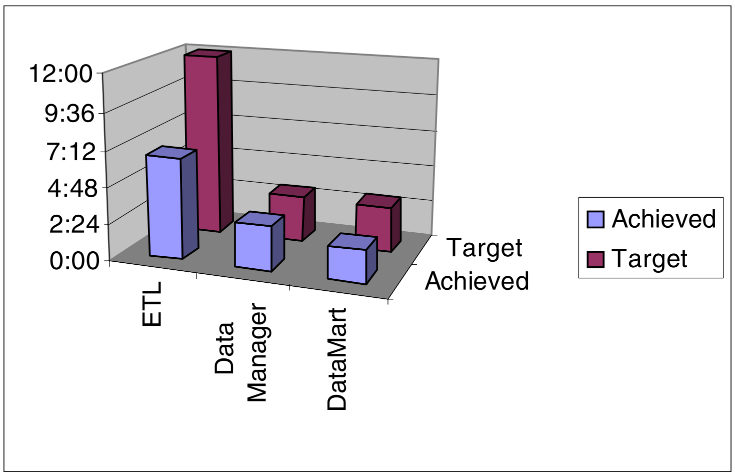

Figure 3 Achieved numbers exceeded expectations

Refer to "Detailed results by test scenario" on page 20 for more specifics. We discuss the optimization tests below.

Summary of optimizations

The test team experimented with a number of different ways to optimize the various solution components. The results are depicted in subsequent figures. ETL, Data Manager and Data Mart are addressed.

ETL optimizations

Figure 4 on page 16 shows the savings achieved by optimizing the ETL processes for the Services, Transaction and Refinancing mappings:

Test 1. Extrapolated duration of direct input process into DB2 without parallelism.

Test 2. Duration of direct input process into DB2 with 12 parallel sessions for Service mapping, and 10 and 5 parallel sessions, respectively, for the Transaction and Refinancing mappings.

Test 3. Duration of the file write process and DB2 load with 12 parallel sessions for Service mapping, and 10 and 5 parallel sessions respectively for the Transaction and Refinancing mappings.

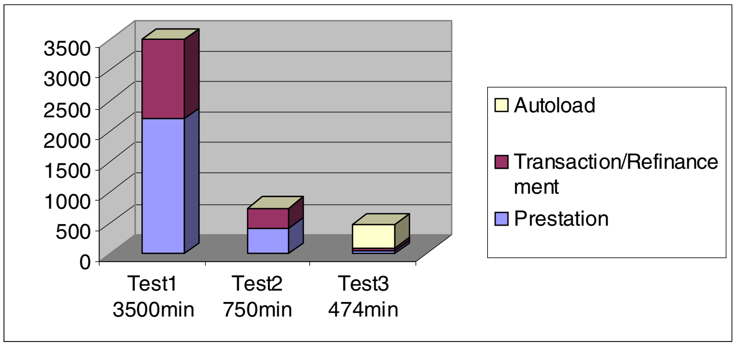

Figure 4 ETL Process optimization

Data Manager optimizations

Figure 5 shows the savings achieved by optimizing the Data Manager:

Test 1. Drop indexes and disable logging on temporary tables, collect statistics.

Test 2. Use only one job stream.

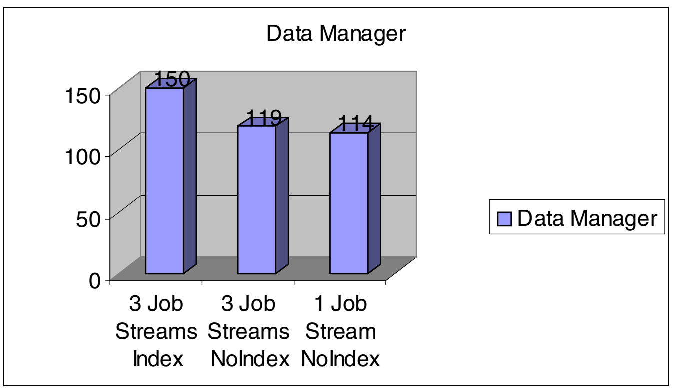

Figure 5 Data Manager optimization

Data mart optimizations

Figure 6 shows the savings achieved by dropping unused indexes on the data mart tables:

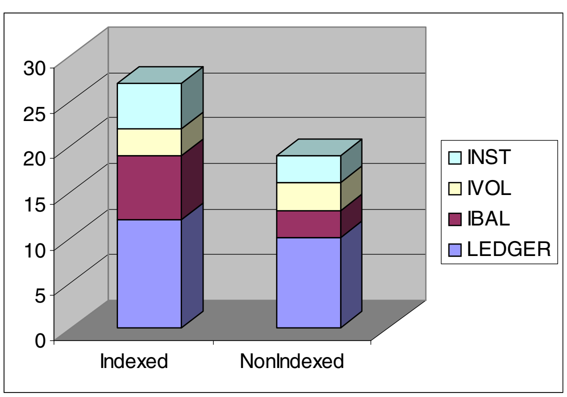

Figure 6 Data mart optimization

Recommendations and lessons learned

One of EPM's strengths is the ability to parallelize the analytical computations in the Data Manager process by using job streaming techniques. At the conclusion of the performance test, another set of tests was performed to determine how job streaming could improve process execution times for the Pilotage 4D project.

Figure 7 depicts the flow of executions when applying job streaming techniques (the picture shows 1 rule out of 4).

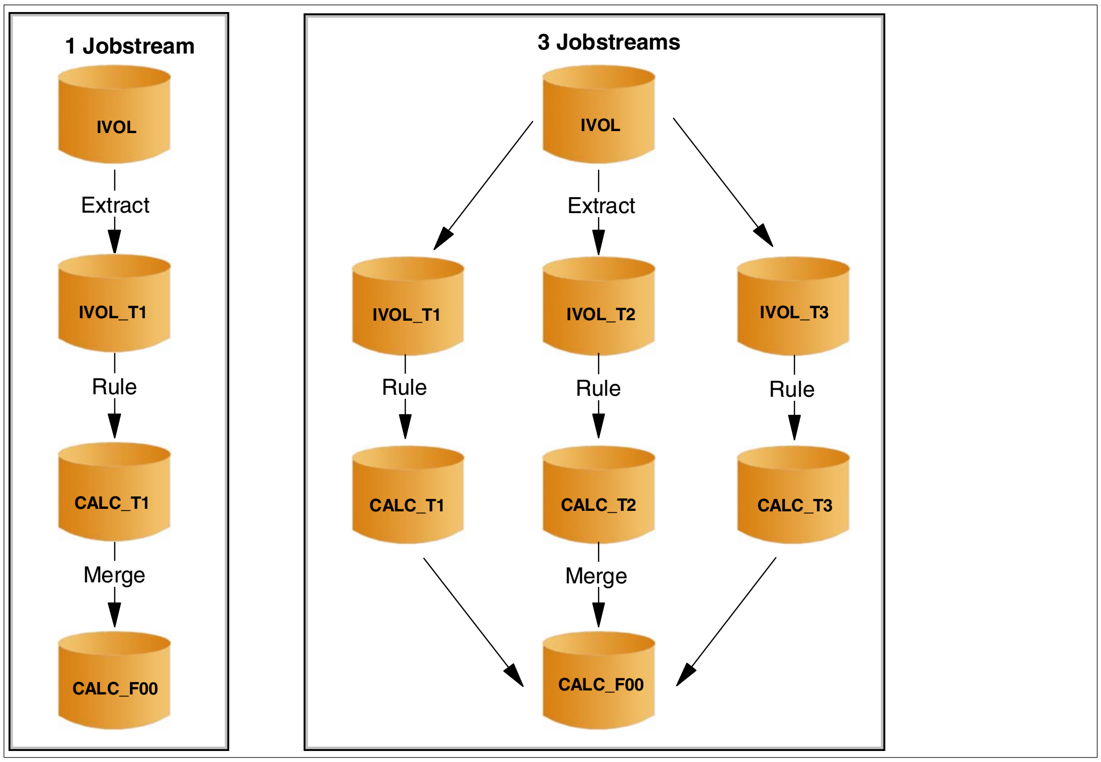

Figure 7 Execution flow with job streaming

This process is executed for the four rules to be tested during the performance test on these tables: IVOL (one rule), IBAL (one rule), and IINC (two rules). Table 2 details the execution times using one job stream on actual volumes.

Table 2 Execution times with one job stream

Query	**Jobstream 1**
insert IVOL_T from IVOL	0:09:00
insert CALC_T (Rule 1 on IVOL_T)	0:13:00
insert IBAL_T from IBAL	0:07:00
insert CALC_T (Rule 2 on IBAL_T)	0:02:00
insert IINC_T from IINC	0:21:00
insert CALC_T (Rule 3 on IINC_T)	0:08:19
insert CALC_T (Rule 4 on IINC_T)	0:19:00
Total	**1:19:19**

Table 3 details the execution times using three job streams on actual volumes.

Table 3 Execution times with three job streams

Query	Jobstream 1	Jobstream 2	Jobstream 3
insert IVOL_T from IVOL	0:09:00	0:09:00	0:10:00
insert CALC_T (Rule 1 on IVOL_T)	0:08:00	0:09:00	0:08:00
insert IBAL_T from IBAL	0:05:00	0:05:00	0:05:00
insert CALC_T (Rule 2 on IBAL_T)	0:03:00	0:02:00	0:02:00
insert IINC_T from IINC	0:22:00	0:22:00	0:22:00
insert CALC_T (Rule 3 on IINC_T)	0:10:00	0:10:00	0:10:00
insert CALC_T (Rule 4 on IINC_T)	0:16:00	0:16:00	0:16:00
Total	**1:13:00**	**1:13:00**	**1:13:00**

A slight gain of 6 minutes (73 vs. 79) was observed. However, two data extraction statements over three were penalized (IVOL and IINC). The data extraction statements scan entire tables and saturate the ESS subsystem. These statements can benefit from indexes, because they apply restrictions on the date and business unit columns.

Execution times after creating indexes are shown in Table 4.

Table 4 Execution times after creating indexes

Query	Jobstream 1	Jobstream 2	Jobstream 3
insert IVOL_T from IVOL	0:07:00	0:07:00	0:06:00
insert CALC_T (Rule 1 on IVOL_T)	0:08:00	0:08:00	0:08:00
insert IBAL_T from IBAL	0:04:00	0:04:00	0:04:00
insert CALC_T (Rule 2 on IBAL_T)	0:02:00	0:02:00	0:02:00
insert IINC_T from IINC	0:20:00	0:20:00	0:20:00
insert CALC_T (Rule 3 on IINC_T)	0:10:00	0:10:00	0:10:00
insert CALC_T (Rule 4 on IINC_T)	0:16:00	0:16:00	0:16:00
Total	**1:07:00**	**1:07:00**	**1:06:00**

Another 6 minutes of savings were observed as a result of using the index.

The total savings realized as a result of using both indexes and job streams was 12 minutes. The test took over 79 minutes without these techniques.

We ran a complete test by varying the number of jobstreams from 1 to 6 (including a MERGE phase that was not part of the previous test):

Jobstream 1	Jobstream 3	Jobstream 4	Jobstream 6
1:19:09	1:08:35	1:15:04	1:13:30

In summary, these tests demonstrated that, while performance was very good using DB2 parallelism exclusively, using job streams could further optimize the Data Manager processes. Job streaming can provide performance benefits for several Data Manager steps; however, data extraction processes can be I/O-bound and may be penalized. In those cases, the use of indexes can benefit data extraction processes.

It is good practice to run tests on large data sets during design and implementation in order to determine when job streaming can be beneficial for individual processes, and to evaluate when indexing can be utilized.

Conclusions

The performance test conducted at the IBM pSeries Teraplex Center met or exceeded all the objectives and success criteria defined by Société Générale prior to the start of the project.

Société Générale validated that PeopleSoft EPM, implemented on an infrastructure of the IBM DB2 Universal Database Enterprise Extended edition on the pSeries 690 server, utilizing the Enterprise Storage Solution disk, was a robust and strategic solution that would support the needs and future goals of the Pilotage 4D project.

Appendix A

Detailed results by test scenario

ETL

Scenario 1: Volume 1/100

The initial machine configuration is 16 CPUs and 96 GB of RAM.

Test 1: Running each ETL mapping without optimization.

Scenario 1 – Test 1	
Client	1 min
Service	22 min
Transactions and Refinancing (run in parallel)	13 min

Test 2: Running the Service mappings in 12 parallel sessions.

This allowed important elapsed time gains.

Scenario 1 - Test 2	
Service	10 min

Scenario 2: Volume 1/10

Test 1: Processing with parallelized mappings:

- Services - 12 sessions
- Transactions and Refinancing - 10 and 5 sessions, respectively

Test 2: After modifying indexes and reorganizing input tables

Scenario 2 – Test 1	
Client	2 min 30 sec
Lookup Client	1 min 30 sec
Service	120 min
Transaction and Refinancing (run in parallel)	57 min (36 min for refinancing)

Scenario 2 – Test 2	
Client	2 min 30 sec
Lookup Client	1 min 30 sec
Service	38 min
Transaction and Refinancing (run in parallel)	28 min

Figure A-1 shows the results of these tests with and without indexes.

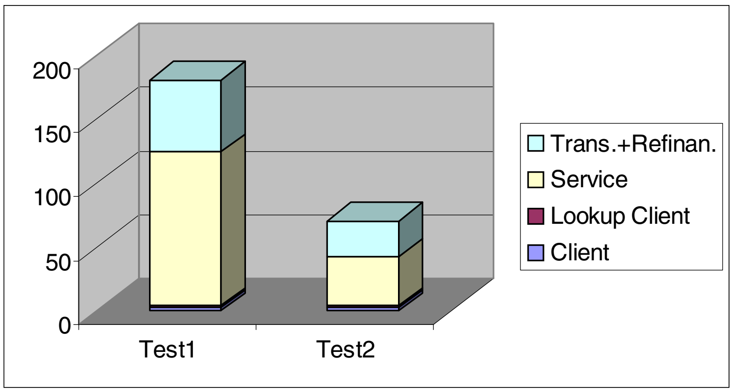

Figure A-1 ETL test with and without indexes

Test 3: Configuration with 12 CPUs and 48 GB of RAM

- The machine configuration now matches the Société Générale's machine
- Two months of history data has been added to the database.

Scenario 2 – Test 3	
Client	2 min 15 sec
Lookup Client	1 min 40 sec
Service	47 min
Transaction and Refinancing (run in parallel)	32 min

Scenario 3: Real volumes

Test 1: Actual volumes

- ETL writes directly in the DW tables
- Two months of history data

Scenario 3 – Test 1	
Clients	20 min
Lookup Clients	20 min
Service	7 hrs
Transaction and Refinancing (run in parallel)	5 hrs 30 min

Test 2: Modifications for updating the tables

Updating the tables was the longest process and we needed to make some modifications. ETL updates some files for Service, Transactions and Refinancing mappings. These files are then loaded into the Data Warehouse tables using the DB2 autoloader function. The initial load time of the files in the staging tables (using the DB2 autoloader) is also measured.

Scenario 3 – Test 2	
Loading source data in staging tables with Autoload	14 min 30 sec
Client	23 min
Lookup Client	21 min
Service	41 min
Autoload Service File + collect statistics	43 min 30 sec
Transaction and Refinancing (run in parallel)	54 min
Autoload Tables IINC, IVOL, ISTATUS, IBALANCE + collect statistics	3 hrs 36 min
Total	6 hrs 39 min

While executing this process, the following measurements were done:

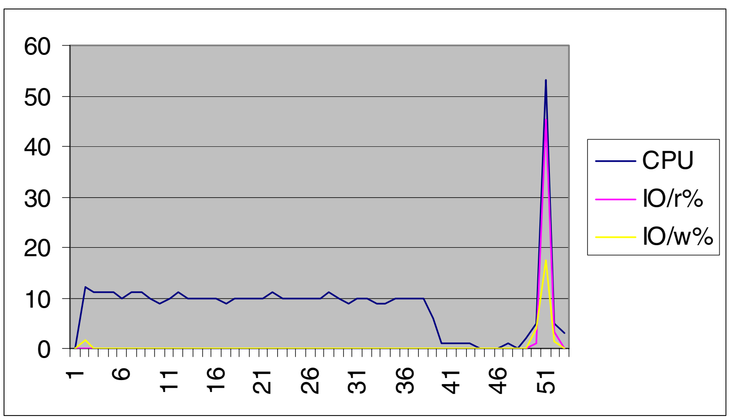

Figure A-2 CPU and I/O resource utilization during customer lookup

Figure A-2 shows that there is little use of system resources. The ETL process on the CLIENT table is not parallelized and CPU utilization is around 10%. If we had parallelized the process on the CLIENT table, we would have significantly decreased the elapsed time on this table. We didn't do this optimization because it was a relatively small segment of the overall ETL process time (45 minutes over a total of 6:30 hours).

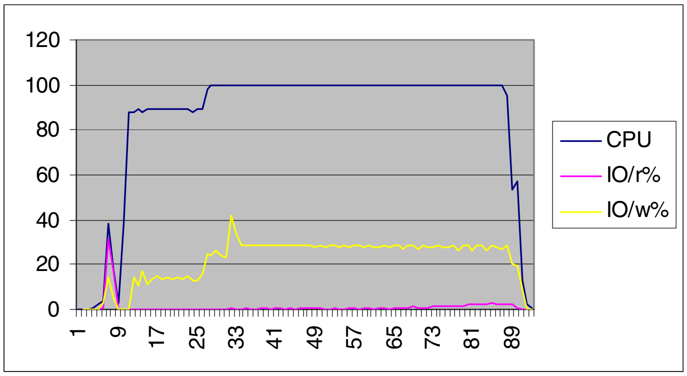

Figure A-3 CPU and I/O resource utilization during Service mapping

In Figure A-3, we see that the CPU is saturated at 100% through almost the entire process. This observation verified the efficiency of parallelism on the ETL process.

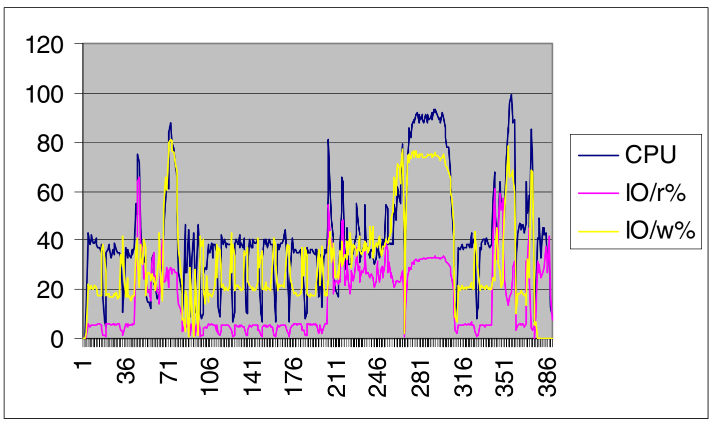

Figure A-4 CPU and I/O resource utilization during the autoloader process

In Figure A-4, we see more heavy I/O activity, including IVOL/IINC/IBALANCE/ISTATUS, with read peaks at 75%. Nevertheless, the CPU is not saturated. It would have been possible to gain more time during this phase if we parallelized the loads of the tables (they were done one after the other).

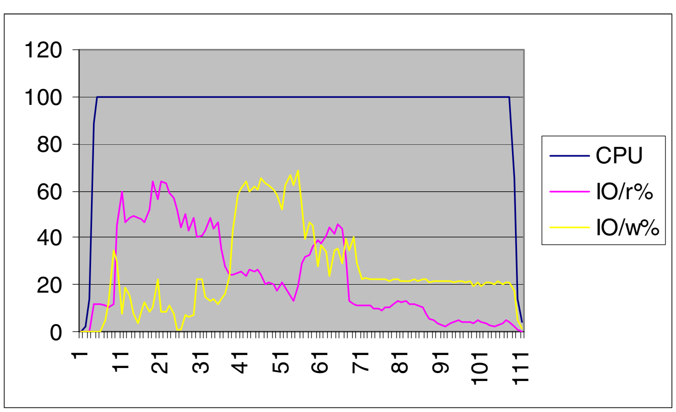

Figure A-5 CPU and I/O resource utilization during transaction and funding mapping

In Figure A-5 on page 23, as in the Service mapping case, the CPU is saturated at 100%. The system is at maximum utilization. At this point, only the addition of another processor can decrease the process elapsed time.

Data Manager

Scenario 2: Volume 1/10

The Data Manager process has two phases: The calculations phase, called Data Manager in the tables below; and the merging of the results phase, called Posting in the tables.

The Posting phase is a process achieved with the Data Manager function of EPM.

Test 1: Without optimization

- Three Business Units, hence three job streams, executed in parallel
- Machine configuration with 16 CPU and 96 GB of RAM

Scenario 2 – Test 1	
Data Manager	21 min 19 sec
Posting	18 min

Test 2: Machine configuration with 16 CPUs and 96 GB of RAM

- Three Business Units, hence three job streams, executed in parallel
- Drop indexes on temporary tables

Scenario 2 – Test 2	
Data Manager	12 min 48 sec
Posting	7 min 21 sec

Test 3: Hardware configuration reduced to 2 CPUs and 48 GB of RAM

- Three Business Units, hence three job streams, executed in parallel
- Two months of history data
- Drop indexes on temporary tables

Scenario 2 – Test 3	
Data Manager	16 min
Posting	7 min 20 sec

Scenario 3: Real volumes

Test 1

- Actual volumes
- Three Business Units

– Complete data history

Scenario 3 – Test 1	
Data Manager	2 hrs 30 min
Posting	Not recorded

Test 2: Drop indexes on PF_LEDGER table

– No indexes on temporary tables
– Tables are "not logged initially"
– Three Business Units and a date index on the table

Scenario 3 – Test 2	
Data Manager	1 hr 59 min
Posting	59 min

Test 3: No indexes on PF_LEDGER

– No indexes on temporary tables
– Tables are "not logged initially"
– One Business Unit

Scenario 3 – Test 3	
Data Manager	1 hr 54 min
Posting	59 min
Total Analytic Computations	2 hrs 53min

Measurements were taken when executing this process, as shown in Figure A-6.

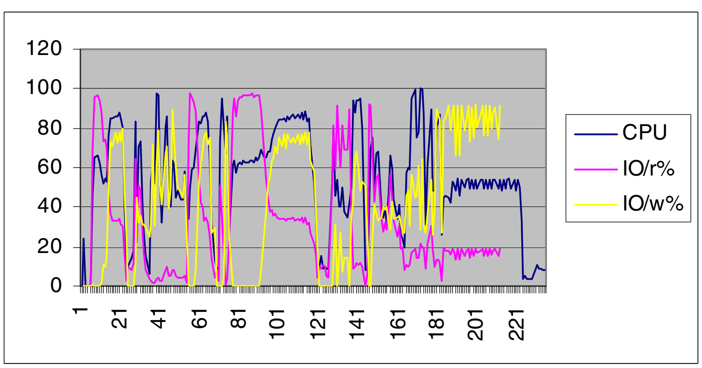

Figure A-6 CPU and I/O utilization during the Data Manager process, one job stream

The system is heavily loaded but not saturated. There is heavy I/O activity, both in read and write. CPU consumption is reasonable. By optimizing the I/Os (either by using indexes or by adding more disk/memory resources), it should be possible to save elapsed time by running several job streams.

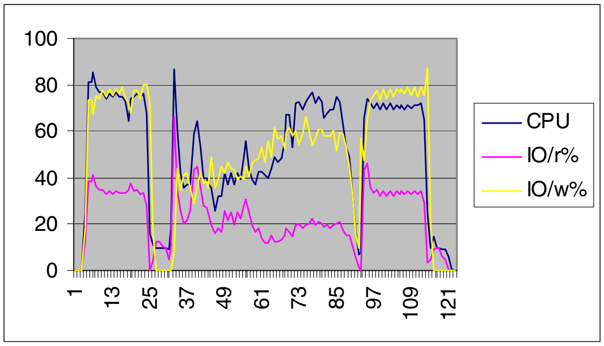

Figure A-7 CPU and I/O resource utilization during posting process

Figure A-7 shows that I/O activity is very high. Adding disk or memory resources could shorten the elapsed time.

Data Mart

Scenario 2: Volume 1/10

Test 1: Four Facts tables processed in parallel

- Hardware configuration: 12 CPUs and 48 GB of RAM
- Two months of history data

Scenario 2 – Test 1	
LEDGER	12 min
IBAL	7 min
IVOL	3 min
INST	5 min

Test 2: Drop unnecessary indexes

Scenario 2 – Test 2	
DIMENSION	5 min
LEDGER	10 min
IBAL	3 min
IVOL	3 min
INST	3 min

Scenario 3: Real volumes

Test 1: Processes are executed serially, with a complete data history

Scenario 3 – Test 1	
Dimensions	4 min
LEDGER	83 min
IBAL	18 min
IVOL	14 min
INST	14 min

Test 2: No LEDGER phase

We dropped the LEDGER phase, which replaced the Deptid value with a “#”, because it was redundant with the ETL controls performed during the load phases.

Scenario 3 – Test 2	
Dimensions	5 min
LEDGER	71 min
IBAL	15 min
IVOL	14 min
INST	22 min
Total	**2hrs 7 min**

We obtained these results by executing the processes serially. Elapsed time could be improved by parallelizing the processes IBAL, IVOL, and INST.

We measured the following resource consumptions, shown in Figure A-8.

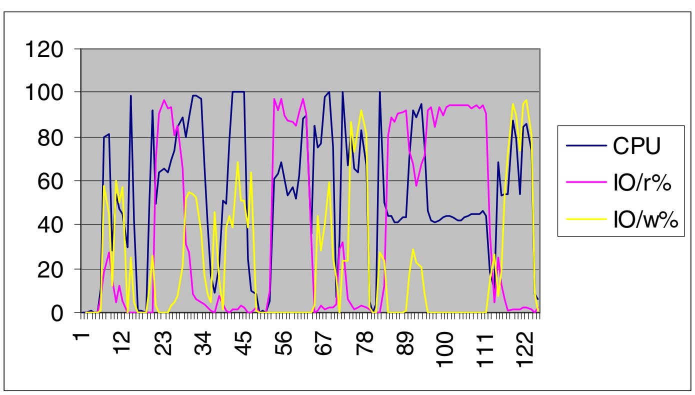

Figure A-8 Resource consumption

I/O activity is very high but the system is not saturated. Here again, adding disk/memory resources can shorten the elapsed time.

Appendix B

Performance test configuration details

pSeries 690 internal configuration

The tests were run on a pSeries model p690-681 with 32 1.3 GHz rs and 128 GB of real memory. The server also used 40 x 36 GB internal disks, giving a total of 720 GB of useful storage secured in RAID/1. The tests were run inside a unique logical partition, using 12 processors spread across 2 MCM and 48 GB of memory; see Figure B-1.

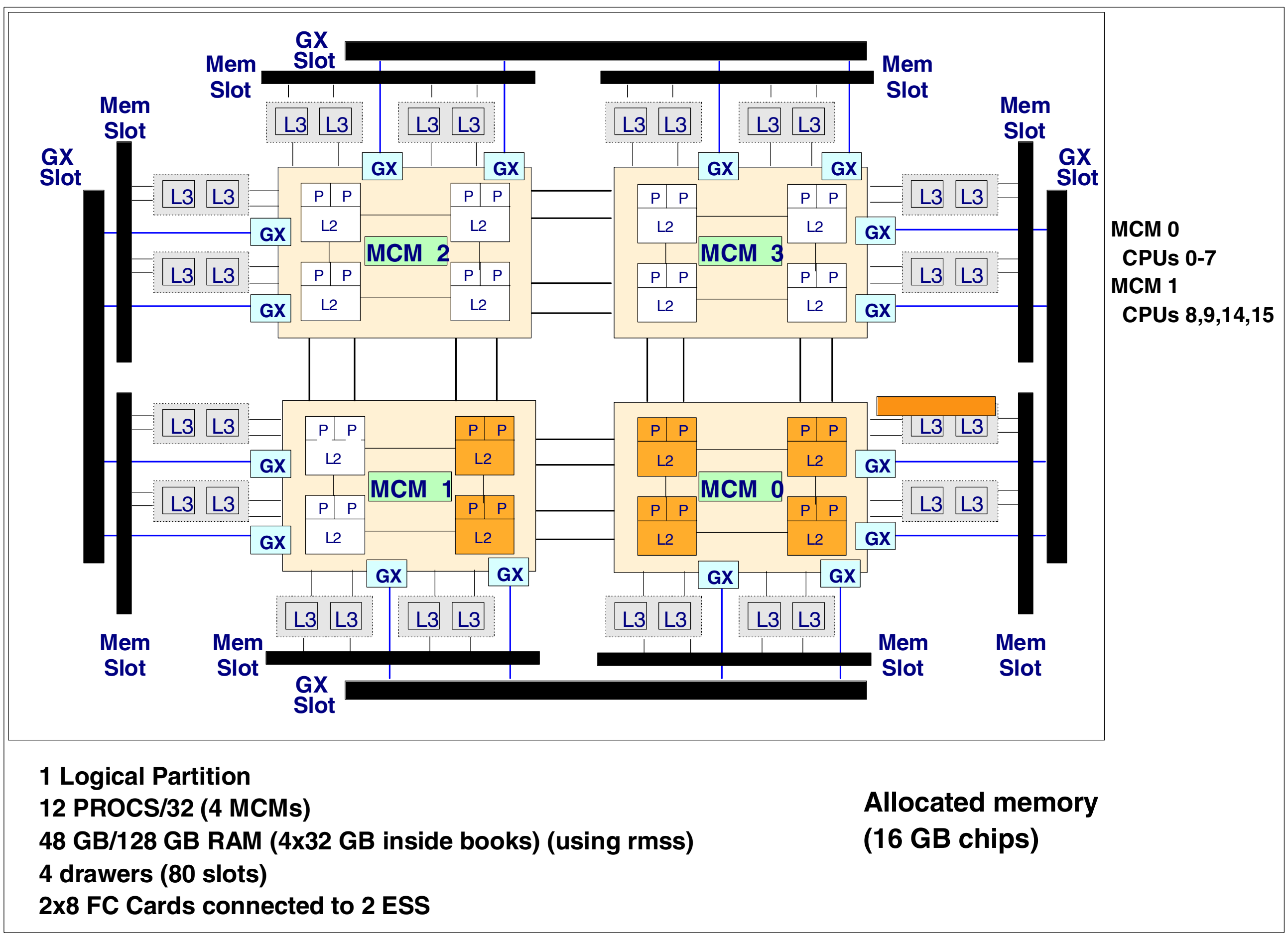

Figure B-1 p690 layout

ESS configuration

Two ESS 2105-F20 units were used for the performance test. Each unit was directly connected to the Regatta server with 8 Fiber Channel links in a point-to-point topology. Each ESS 2105-F20 unit included:

- 8 Fiber Channel cards
- 32 GB of cache memory
- 384 MB of NVS (non-volatile storage)
- 16 drawers of 8 x 72 GB disks each

Also:

- Each drawer had a useful capacity of 432 GB.
- The usable capacity was 6.9 TB.

Figure B-2 illustrates the schematic.

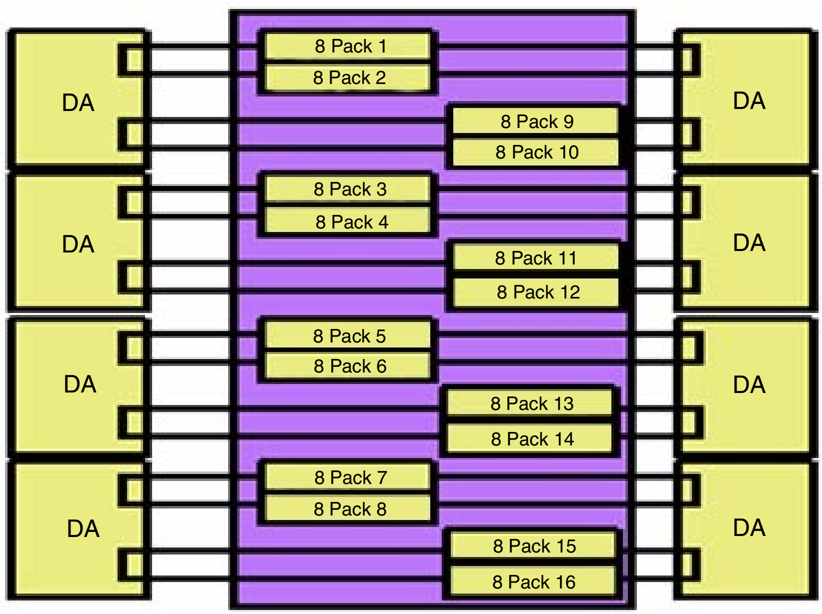

Figure B-2 ESS schematic

In an ESS unit, disks are grouped in *arrays*, on which are defined logical disks called LUNs. The LUNs are then seen as physical disks from the AIX operating system. Thus an AIX I/O can translate into several simultaneous physical disk accesses in the ESS unit. The fully populated ESS 2105-F20 unit contained 16 arrays, each made of 6 disks for the data, 1 disk for parity within RAID/5 protection, and 1 spare disk. On each array, two LUNs of 210 GB each were created, a total of 32 LUNs, having an addressable total storage of 6.7 tera-octets.

Configured in the AIX Logical Volume Manager were:

- 32 logical volumes, containing a 150 GB file system on each LUN, for the DB2 tablespaces - overall 32 file systems for a total of 4.8 terabytes
- A unique stripped file system of 1 TB in size, using 35 GB on each LUN

Figure B-3 on page 30 illustrates the storage layout of the ESS unit:

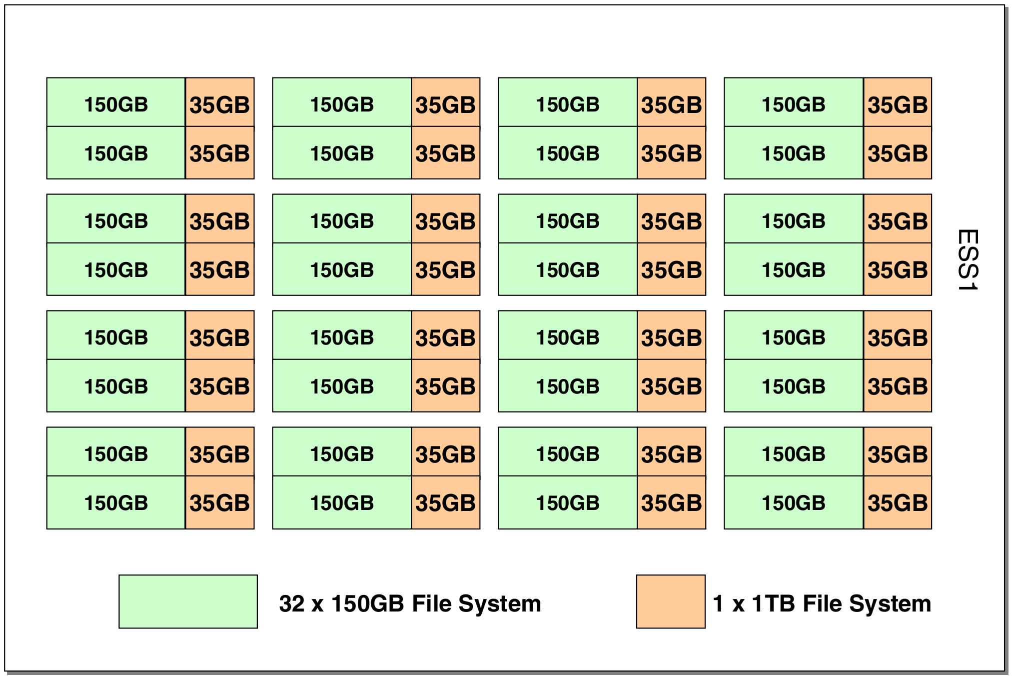

Figure B-3 ESS storage layout

Appendix C

Data Structures

Tables

Table C-1 Tables used in the tests

Data	Rows
Customer/Branch	15,000,000
Customer	5,000,000
Products	1146
Products Analysis	200
Branch	692
Media	5
Accounts	13
Standard Costs	4,422
DEC	47
DR	8
Market	9
Segment Groups	11
Segment	26
Sub Segment	86
Product Groups	4
Product Family	38
Segment Family	16
Customer Type	3
Transaction Type	4

Monthly flow

Table C-2 Volumes of source files to be loaded for the monthly flow

Data	Rows
Clients-Guichets	15 million
Clients	5 million
Prestations	40 million
Transactions	100 million
Refinancement*	40 million

* However, every record for the refinancing file has two amounts (management rate and economy rate), leading to a total volume of 80 million.

Table C-3 Source files loaded in the staging area

Table	Rows	Data (MB)	Index (MB)	Data+Index (MB)
CLIENT	5,030,000	390	103	493
PRESTATION	40,010,000	4,305	2,342	6,647
REFINANCEMENT	40,010,000	2,457	2,472	4,929
TRANSACTION	100,070,000	6,533	3,293	9,826
TRANSCO_CLIENT	15,230,000	937	0	937
TOTAL		**14,622**	**8,210**	**22,832**

Table C-4 EW tables loaded by the monthly processing

Table	Rows	Data (MB)	Index (MB)	Data+Index (MB)
PS_FI_IINC_R00_CURR	419,466,114	68,883	41,144	110,027
PS_FI_IVOL_R00_CURR	140,355,293	24,604	12,974	37,578
PS_FI_IBALANCE_R00_CURR	119,970,258	20,795	604	21,399
PS_FI_INSTR_F00_CURR	39,990,086	23,995	1,575	25,570
PS_FI_ISTATUS_R00_CURR	39,990,086	5,776	1,575	7,352
PS_CUSTOMER_D00_CURR	5,030,000	1,649	230	1,879
TOTAL		**145,703**	**58,102**	**203,805**

EPM EW history log

The performance test had to simulate data warehouse stored volumes, corresponding to one year of activity. The following tables were thus loaded with eight months worth of history data:

- Customers (CUSTOMER_D00)
- Services (FI_INSTR_F00, FI_ISTATUS, FI_IBALANCE)
- Transactions (FI_IVOL_F00)
- PNB detailed elements (FI_IINC_F00)

Table C-5 Tables loaded with eight months of history data

Table	Rows	Data (MB)	Index (MB)	Data+Index (MB)
PS_CUSTOMER_D00	40,513,728	13,190	3,197	16,387
PS_FI_IINC_R00	3,350,013,792	551,059	329,208	880,267
PS_FI_INSTR_F00	319,367,040	191,957	31,362	223,319
PS_FI_ISTATUS_R00	319,367,040	46,211	12,602	58,813
PS_FI_IVOL_R00	1,121,017,248	196,830	98,404	295,234
PS_FI_IBALANCE_R00	958,101,120	166,360	49,084	295,936
TOTAL		**1,165,607**	**523,855**	**1,689,461**

Data Manager tables

Table C-6 Tables loaded by Data Manager monthly processing

Table	Rows	Data (MB)	Index (MB)	Data+Index (MB)
SG_CALC_F00	366,266,432	61,232	0	61,232
PF_LEDGER_F00	366,266,432	61,214	0	61,214
TOTAL				**122,446**

Data mart tables

Table C-7 Tables loaded by data mart monthly processing

Table	Rows	Data (MB)	Index (MB)	Data+Index (MB)
PS_SG_CUST_DIM	5,030,000	2,139	357	2,496
PS_SG_IBAL_FACT	39,967,116	5,321	5,276	10,598
PS_SG_INSTR_FACT	39,967,116	5,321	5,276	10,598
PS_SG_IVOL_FACT	22,846,849	3,040	3,014	6,055
PS_SG_LEDGER_FACT	219,269,641	29,196	52,104	81,299
TOTAL		**45,018**	**66,027**	**111,045**

Appendix D

DB2 partition disk mapping

DB2 Instance partitioning

The strength of DB2 UDB EEE resides in its “shared-nothing” architecture, which allows for the division of large volumes of data into multiple partitions, thus making parallel executions possible. Partitioning is especially beneficial for EPM, which extensively uses the INSERT/SELECT technique.

The INSERT/SELECT technique permits inserting directly into a table, the result set of a SELECT statement executed on other tables. In the case of EPM, it is used for the update of the EW, data mart, and temporary tables (job streams, merge). The advantages of the INSERT/SELECT technique are twofold:

- The database engine directly handles operations such as joins, aggregations, and sorts. This avoids any data transfer between the DBMS and the application.
- Parallelism is managed entirely by the database engine, which optimizes its performance and scalability when additional resources, such as processors and nodes, are provided.

To process an INSERT/SELECT statement, DB2 UDB/EEE uses a partitioning process to achieve parallel inserts. To optimize the benefits of parallelism, we designed a DB2 partition per processor, giving us 12 partitions in total. Thus, all processors are used to process a statement, optimizing resource utilization. Figure D-1 on page 35 illustrates the DB2 UDB/EEE instance architecture.

Disk placement

Four arrays (eight LUNs) were allocated to each group of three DB2 partitions. Tablespaces were defined with eight containers per partition spread over the 8 LUNs allocated to each partition, thus giving a total of 96 containers for each tablespace. All partitioned tablespaces were created using the DMS file mode.

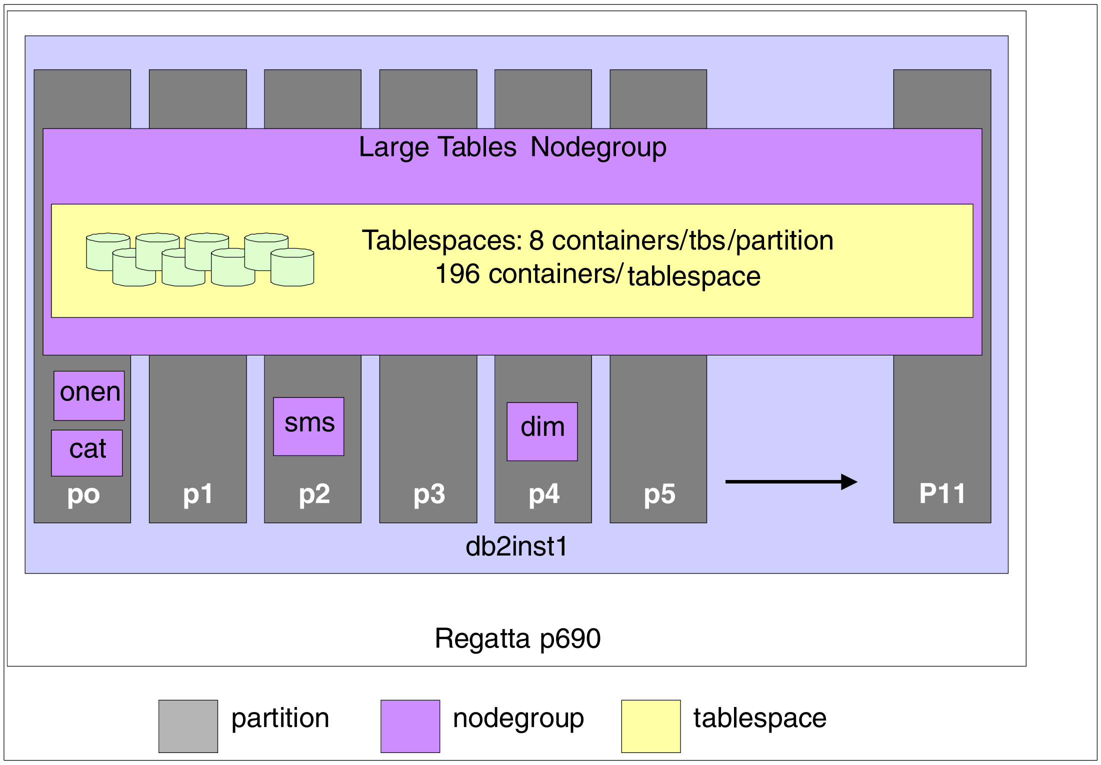

Figure D-1 DB2 UDB/EEE instance architecture

Figure D-2 shows the mapping of DB2 partitions to ESS LUNs.

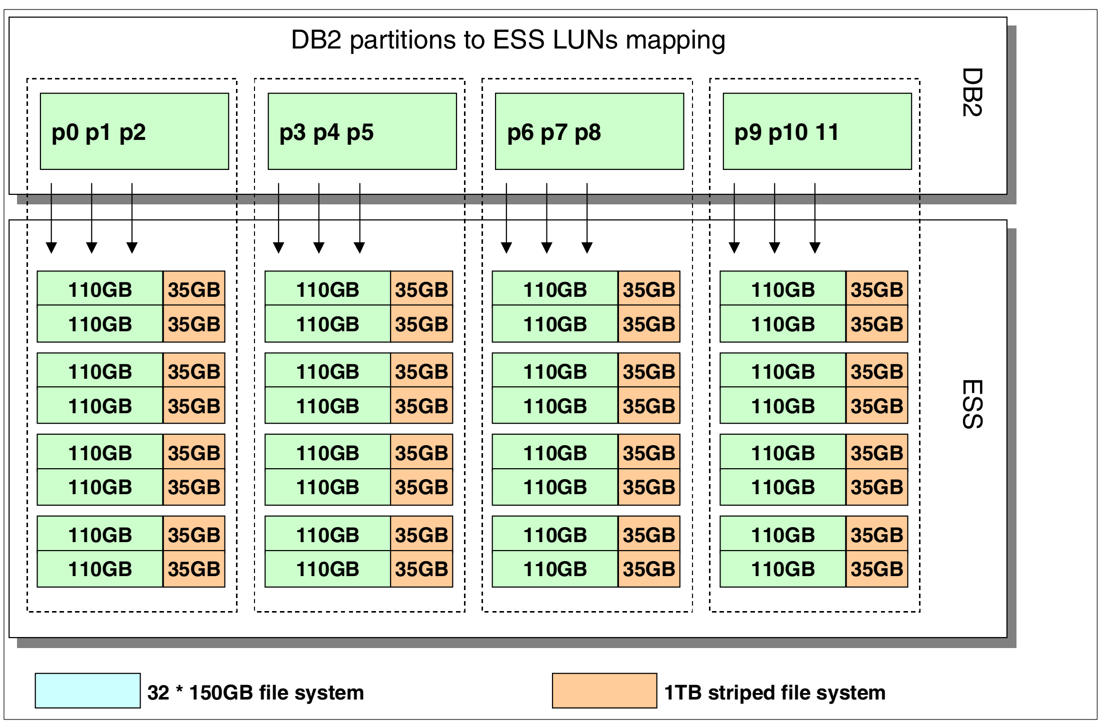

Figure D-2 Mapping of DB2 partitions to ESS LUNs

Figure D-3 on page 36 shows the containers created on the first array allocated to partitions 0, 1, and 2 for tablespaces tbs1, tbs2, and temptbs.

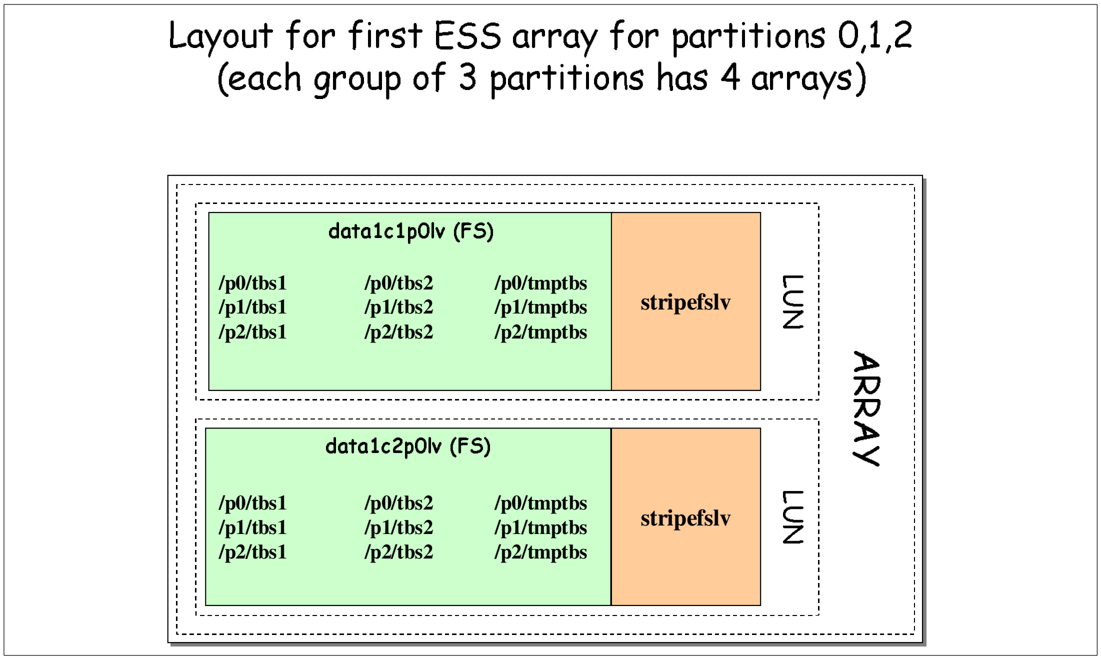

Figure D-3 ESS array layout

Figure D-4 shows how the containers are arranged.

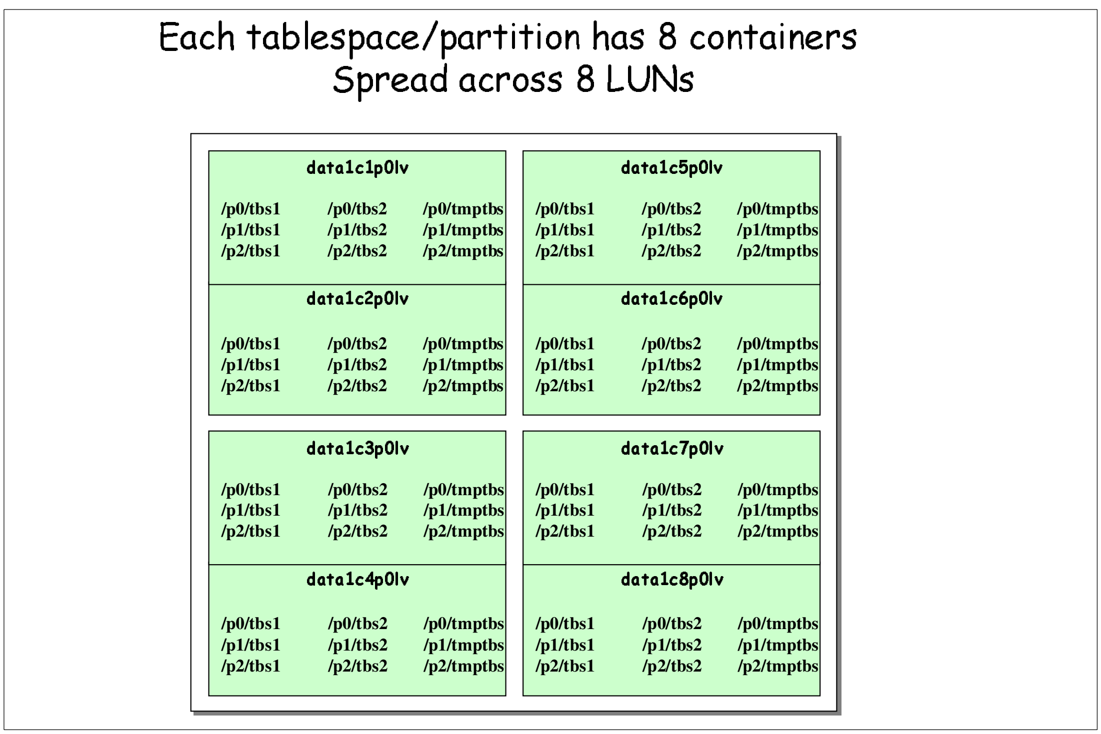

Figure D-4 Containers spread across LUNs

The I/Os were balanced and all the disk resources were used evenly during heavy reads and inserts.

We verified the efficiency of this configuration by measuring read rates close to 300 MB/sec, which is the average rate of the ESS storage system.

Appendix E

The IBM pSeries 690

The IBM pSeries™ 690 is the first member of the IBM UNIX 64-bit Symmetric Multi-Processing (SMP) server family equipped with POWER4 technology processors. This is a multi-purpose server that is especially qualified for commercial applications, either decisional or transactional.

The pSeries was designed with features to enable the consolidation of critical applications on a single server:

- Logical partitioning (LPAR): The system can be configured to host independent environments, each having its own copy of the operating system.
- Auto-configuration: Starting with AIX 5.2, the system will allow hot reconfiguration or extension causing no operation disruption. In the 5.1 version, this feature is already applicable for all PCI slots and disks.
- Auto-healing: The system can detect problems before they occur and correct them.
- Auto-optimization (WLM): The system can dynamically balance resources and workloads to optimize response time and throughput.

A key feature of the pSeries 690 is logical partitioning, as shown in Figure E-1 on page 38. Logical partitioning permits the definition of up to 16 different system partitions. Each runs a virtual machine that can use operating system and software versions that are different from the other LPARs. Each virtual machine has dedicated resources (CPU, memory and peripheral devices), and is completely isolated from others. Each has one Power4 processor, one GB of memory, and one PCI slot.

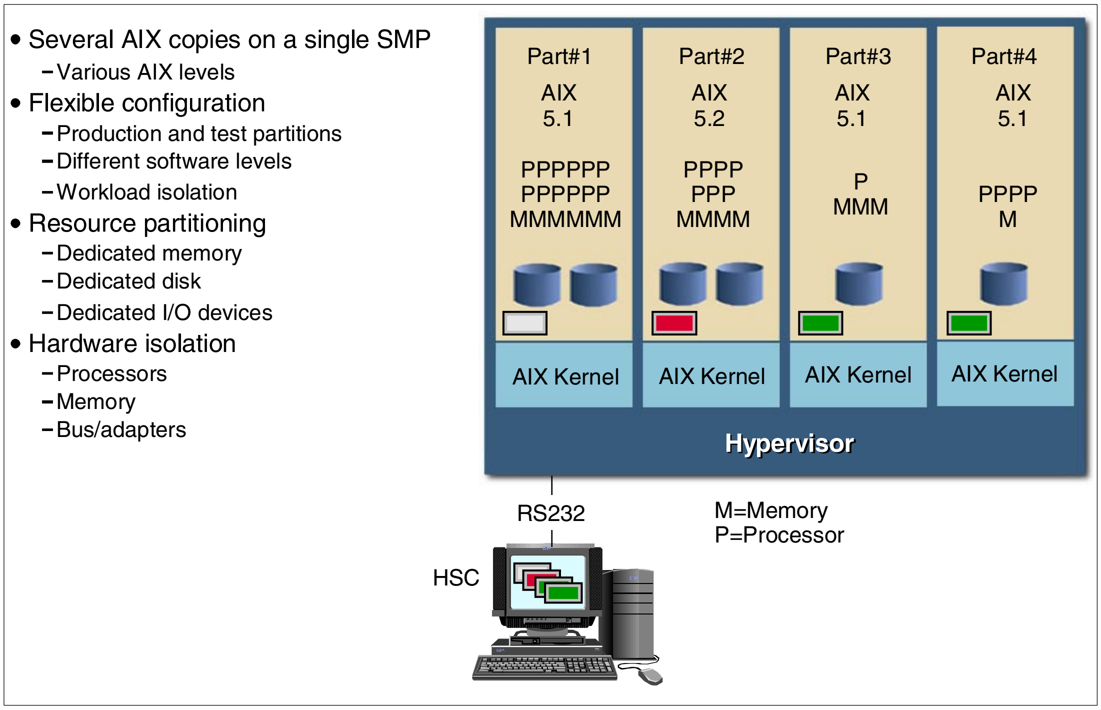

Figure E-1 pSeries 690 logical partitions

The pSeries 690 exploits the AIX 5L operating system, which includes:

- 64-bit UNIX for POWER and IA-64 processor support
- Up to 64 petabyte file systems
- Up to 1-terabyte files
- Workload Management, including I/O management
- New Web-based System Manager functions
- Default Java V1.3.0 JDK
- System resources controller

Information about the complete IBM @server line can be found at:

`http://www.ibm.com`

Notices

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

Information in this presentation concerning non-IBM products was obtained from the suppliers of these products, published announcement material or other publicly available sources. Sources for non-IBM list prices and performance numbers are taken from publicly available information including D.H. Brown, vendor announcements, vendor WWW Home Pages, SPEC Home Page, GPC (Graphics Processing Council) Home Page and TPC (Transaction Processing Performance Council) Home Page. IBM has not tested these products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

IBM may have patents or pending patent applications covering subject matter described 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:
IBM Director of Licensing, IBM Corporation, North Castle Drive Armonk, NY 10504-1785 U.S.A.

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

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. The use of this information or the implementation of any techniques described herein is a customer responsibility and depends on the customer's ability to evaluate and integrate them into the customer's operational environment. Customers attempting to adapt these techniques to their own environments do so at their own risk.

All statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only. Contact your local IBM office or IBM authorized reseller for the full text of a specific Statement of General Direction.

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

IBM is not responsible for printing errors in this presentation that result in pricing or information inaccuracies.

The information contained in this presentation represents the current views of IBM on the issues discussed as of the date of publication. IBM cannot guarantee the accuracy of any information presented after the date of publication.

IBM products are manufactured from new parts, or new and serviceable used parts. Regardless, our warranty terms apply.

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

Customer examples cited or described in this document are presented as illustrations of the manner in which some customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics will vary depending on individual customer configurations and conditions, similar results cannot be guaranteed elsewhere.

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

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

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

For general inquiries about IBM:
800-426-4868 (within the United States)
404-238-1234 (outside the continental United States)
Printed in the United States 2002

Send us your comments in one of the following ways:

- Use the online **Contact us** review redbook form found at:
  http://www.ibm.com/redbooks
- **Send** your comments in an Internet note to:
  redbook@us.ibm.com
- Mail your comments to:
  IBM Corporation, International Technical Support Organization
  Dept. HYJ Mail Station P099
  2455 South Road
  Poughkeepsie, NY 12601-5400 U.S.A.

Trademarks

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

AIX®
AIX 5L™
Application Region Manager™
AS/400®
CICS®
ClusterProven®
DataHub®
DataJoiner®
DB2®
DB2 OLAP Server™
DB2 Universal Database™
DEEP BLUE®
DYNIX®
Enterprise Storage Server™
ESCON®
GigaProcessor™
IBM®
Intelligent Miner™
iSeries™
LANStreamer®
LoadLeveler®
Magstar®
MediaStreamer®
Micro Channel®
MQSeries®
Netfinity®
Network Station®
NUMACenter™
Parallel Sysplex®
PartnerLink®
Perform™
POWERparallel®
PowerPC®
PowerPC 604™
pSeries™
S/390®
Scalable POWERparallel Systems®
Sequent®
Sequent (logo)™
SequentLINK™
Service Director™
SmoothStart™
SP™
SP2®
ThinkPad®
TotalStorage™
TURBOWAYS®
VisualAge®
Visualization Data Explorer™
WebSphere®
xSeries™
zSeries™

The following terms are trademarks of other companies:

ActionMedia, LANDesk, MMX, Pentium and ProShare are trademarks of Intel Corporation in the United States, other countries, or both.

Java and all Java-based trademarks and logos are trademarks or registered trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.

The eServer brand consists of the established IBM e- business logo followed by the descriptive term “server”.

PeopleSoft is a registered trademark and PeopleSoft EPM is a trademark of PeopleSoft Corporation.

C-bus is a trademark of Corollary, Inc. in the United States, other countries, or both.

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

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.