

# Escala & Oracle 10g

## Data consolidation

This document focuses on data consolidation, introduces a number of database consolidation scenarios and takes stock of the very latest technologies.



Architect of an Open World™

# WHITE PAPER



## Table of content

▪ Introduction .....	3
▪ Escala's benefits: Partitioning & virtualisation .....	4
▪ Oracle10g: Real Application Clusters, ASM, Resource Manager .....	5
▪ Oracle database consolidation models.....	7
▪ Oracle database deployments .....	8
▪ Oracle consolidated architecture on Escala.....	11
▪ Simplified Administration and support .....	13
▪ Conclusion .....	14

# Introduction

The fully-distributed deployment model in which each application accesses its own resources (Servers, Databases, etc.) is now in widespread use. Nevertheless, it does have some disadvantages that can be addressed by today's technologies without having to modify applications.

The major disadvantages of the distributed model are:

- Under-utilization of available computing power to the extent where servers are often oversized purely to cope with peaks in workload
- Operational complexity, in terms both of backups and data flows between databases to ensure data consistency
- High administration costs as a result of the need to monitor and maintain a large number of environments (database releases, operating system releases, etc.)
- Integration costs generated by the multiplicity of (OS and database) releases, and the impossibility of developing all environments simultaneously.

Compared with this model, which we might describe as "hyper-distributed", today's technologies give us the opportunity to design and implement models based on consolidation.

**Bull is among the very first manufacturers to have chosen open standards and systems to liberate information systems and reduce their complexity.** In 1992, Bull opted for Escala and the AIX™ operating system, and has been substantially involved in AIX's technological development.

AIX 5L™, recognized by analysts as the UNIX® operating system with the greatest growth dynamics, was the first operating system to obtain UNIX 2003 certification. The latest generation of Escala servers with Power5™ processors and running AIX 5L v5.3 is characterized by the flexibility of their architecture, innovative virtualization functions, reliability and excellent availability, as well as their easy maintainability.

**For many years, Bull has developed a close relationship with key partners, such as Oracle**, whose database is known as the reference on the market in terms of performance and reliability, while providing highly enhanced functions. Oracle Database 10g is the first database designed for Grid Computing for corporate use. It complies with industry standards and enables connectivity with applications and tools on the market. It has advanced self-administration functions and optimizes the use of resources, which significantly reduces its operating costs. Finally, since it is associated with Escala servers and leading EMC2, Brocade, Overland and StorageTek secure storage solutions which support them, Oracle Database 10g is the preferred choice of many customers who need to manage critical applications with voluminous databases.

**Whether geographical consolidation, data consolidation, storage consolidation or application consolidation, these models combine flexibility with resource control, cost control, performance and robustness. This document focuses on data consolidation, introduces a number of database consolidation scenarios and takes stock of the very latest Oracle databases and Bull Escala server technologies.**

# Escala's benefits: Partitioning & virtualisation

The innovative technology of Escala systems with POWER5 and AIX 5L v5.3 processors is the preferred platform for Oracle database environments.

## Partitioning and Load Management

Two methods of resource distribution are possible on the Bull Escala range models:

- Distributing system resources (CPU, Memory, I/O) between several Operating System instances via the concept of partition
- Distributing system resources (CPU, Memory, I/O) between different applications using the same Operating System instance via the use of a Workload Manager (WLM).

## Partitioning and Virtualisation on Escala / AIX™ 5L

With the Logical Partition notion (LPAR), the system's resources – CPU, Memory Zones and I/O Controllers – are distributed over the different partitions of the system. The Dynamic Partitioning function (DLPAR) increases flexibility by allowing dynamic change of the resources without disturbing operation.

The partitioning notion was extended with the Micro-Partitioning (SPLAR) and I/O virtualisation (VIO) functions which allow an even more accurate usage of the CPU resources and considerably optimizes the I/O resources.

Two types of DLPAR Partitions exist and can cohabit on the Escala range:

- Partitions with dedicated processors or DEDICATED PARTITIONS
- Micro Partitions which share a set of physical processors or SHARED PROCESSORS PARTITIONS.

In SHARED PROCESSORS PARTITIONS Mode, several partitions share a set of physical processors. One assigns a certain amount of Virtual Processors (VP) and an Execution Capacity (EC) to a partition, which is expressed in tenths of processor.

The number of virtual processors assigned to a partition can exceed the number of shareable physical processors (within a limit of 64 virtual processors). The number of virtual processors allows to increase or decrease the parallelism of the applications' executions.

The execution capacity of the set of SHARED PROCESSORS PARTITIONS is necessarily lower or equal to the number of physical shareable processors.

With the concept of Virtualisation, it is also possible to share an I/O Controller between several partitions.

Different partitions can run under different releases of AIX thus allowing a high level of logical isolation between applications running in different partitions. The crash of the Operating System running in a partition has no incidence on other partitions.

## Load Management under Escala / AIX 5L

When several applications or modules of an application are running in the same instance of the Operating System, WorkLoad Manager helps guarantee that all the resources will not be allocated to an application or a module to the prejudice of other applications.

WLM helps control the processor resources, the physical memory and the I/O bandwidth.

# Oracle 10g: Real Application Clusters, ASM, Resource Manager

**Version 10g of the Oracle database includes new self-management and virtualization functions, providing optimum management of resources, and the implementation of virtualized infrastructures that are essential in corporate environments today.**

With Oracle 10g RAC, the same database can be accessed simultaneously by multiple servers, each of which hosts an Oracle instance. The Oracle cluster components (Cluster Ready Services or CRS) manage the cluster configuration and certain events, such as server booting or recovery...

An ASM pseudo-instance runs on each cluster node: when booting an ASM-based Oracle database, the Oracle instance calls the local ASM pseudo-instance to obtain the file allocation table.

The Oracle instances exchange data between each other via the Cache Fusion protocol.

In the RAC environment, services are deployed across the entire cluster in accordance with their individual properties. Services may have the following properties:

- One or more preferred servers
- One or more backup servers.

If a preferred server fails, the service will be redeployed to a backup server.

At any time, the administrator may:

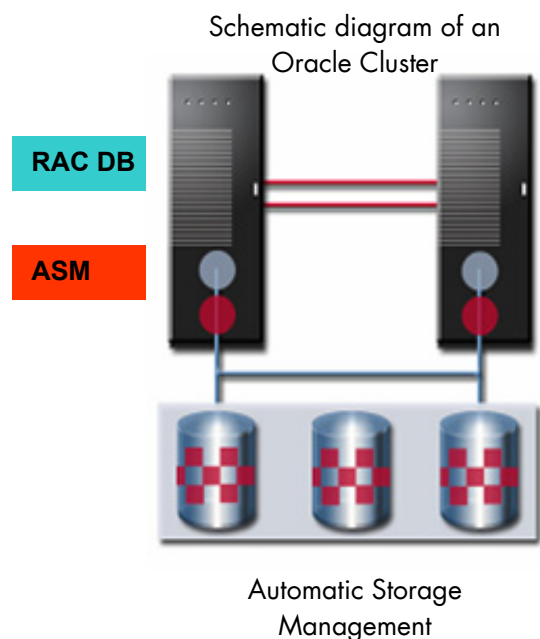
- Stop or start a service on one or all servers
- Move a service from one server to another.

It is possible to activate the same resource plan for all Oracle instances or to activate a specific resource plan for each instance.

In a cluster environment, the workload is distributed to the various instances on the basis of the service configuration. Each instance controls the way resources are used by implementing the active resource plan.

The previous example could therefore be modified as follows:

- A TP service on both cluster nodes
- A batch service on node 1, with node 2 acting as backup
- Resource plans giving priority to the TP service during the day and/or distributing resources equally between TP and batch at night
- TP service on node 2 halted during the closing period.



### **Automatic Storage Management or ASM**

Automatic Storage Management or ASM functionality introduced in Oracle 10g simplifies Storage Management for Oracle Databases while integrating VOLUME MANAGER functionalities in a single product: Disk and Redundancy management, I/O repartition and CLUSTER FILE SYSTEM Management.

ASM distributes storage blocks automatically to the different disks affected to a volume or DISKGROUP in the ASM terminology.

Volume management operations are dynamical: addition or withdrawal of a disk in an active DISKGROUP – rebalancing of the I/O are automatic.

The same volume can lodge data from several bases. Any file generated by Oracle can be managed by ASM: control files, spfile, data files, redo logs, archive logs, backup sets, and recovery area.

An ASM pseudo-authority manages metadata which are secured on disks.

When booting, an Oracle instance using ASM interrogates the ASM instance in order to obtain an extent map of its files. Then, it will carry out on itself its own I/O.

ASM can only manage extent maps. The ASM instance is thus used during the following operations:

- Opening Database files
- Extension of a Database file
- Data rebalancing within a DISKGROUP after its modifications: addition or withdrawal of disks.

### **Oracle 10g Resource Manager**

Oracle 10g Enterprise Edition integrates a resource manager, Oracle 10g Resource Manager. Oracle 10g Resource Manager is comparable to AIX WLM for Databases so that you control the use of resources within an Oracle instance. The resources are affected to resource groups based on the following criterias:

- Explicit i.e. via SQL command
- Name of the service used to connect to the Database
- User Name ( Oracle or Operating System)
- Name of the program, the module or the client machine.

The Service notion which is totally integrated to the Database lies within the responsibility of the administrator. A service can correspond to an application or a type of use of the Database.

Oracle 10 g Resource Manager controls the processor resources, the number of simultaneously active transactions, the authorized degree of parallelism, the maximum execution time of an SQL request, the consumption of canceling spaces or even the inactivity time.

The activation of a resource plan is fully dynamic.

Oracle 10g integrates a performance follow-up reference frame (Automatic Workload Repository or AWR). With AWR, service functionalities and Oracle 10 g Resource Manager, the administrator can follow the consumption and adapt the resource management to the needs. For example:

- A TP service and a batch service
- Resource plans
- Diurnal giving the priority to the TP
- Nocturnal, distributing the resources equitably between TP and batch
- Monthly, giving an absolute priority to the closure module.

# Oracle database consolidation models

## **Physical Consolidation**

The physical consolidation model is the simplest to implement, and consists of grouping several databases together on a single Bull Escala server, each database being hosted on a separate partition. This consolidation model aims to exploit the benefits to be gained from those functions delivering dynamic resource allocation between partitions, and therefore makes better use of the available computing power. On the other hand, it delivers no improvement in respect of any other issues, such as operational complexity, administration costs and integration costs.

## **Logical Consolidation**

In the Logical Consolidation model, several databases coexist within the same Operating System instance, and resources are shared between basic instances by the WLM. As with the previous model, this consolidation mode aims to make better use of available computing power, whilst reducing administration costs slightly (only one operating system instance to administer). However, there remains the necessity of managing multiple databases or even multiple Oracle releases.

## **Basic Consolidation in Oracle 10 g Enterprise Edition**

In this consolidation model, a database becomes a schema in an Oracle 10g database. Instead of accessing a database, the application accesses a service. Oracle 10g Resource Manager allocates resources between services.

This consolidation model delivers more efficient use of available computing power, at the same time as reducing administration and integration

costs, since there is only one database to be administered.

In order to maintain application autonomy in terms of backup and restore, the schema associated with a particular application will be deployed in a set of Oracle TABLESPACES specific to that application.

Using the TRANSPORTABLE TABLESPACE concept, the administrator can extract an application and integrate it into an autonomous database at any time. Such operations may be required when upgrading to a new release, if one of the applications cannot implement the update required by the other applications consolidated in the same database.

Data flows between databases then become internal flows. Some, such as repository flows, can be deleted and replaced using SYNONYM.

## **Basic Consolidation in Oracle 10 g RAC**

In this Consolidation mode, a database becomes a schema in an Oracle 10g RAC database. Instead of accessing a database, the application accesses a service that may be hosted on one or more servers in the cluster. This consolidation mode, which is similar to the previous one, also exploits the availability and load balancing functionalities of the Oracle 10g RAC cluster.

# Oracle database deployments

## Generic deployment

This chapter examines how candidate applications, initially deployed using the distributed model, run in multi-level mode. This consolidation model groups the application server and database levels together on a set of Bull Escala servers.

Micro-partitioning and the virtualisation of I/O controllers and the network are particularly well-suited to the application server level, but the same is not necessarily true for the database server level. The application server level will be deployed on micro-partitions within a dedicated server or on a server that also includes the database server level.

For the remainder of this article, we will focus exclusively on the database server level.

## Deployment for database server levels

The purpose of deploying Oracle databases on a Bull Escala server as part of a consolidation (multiple databases running on the same server) is to mutualise or share the maximum amount of resources.

Each Oracle database requires the following resources:

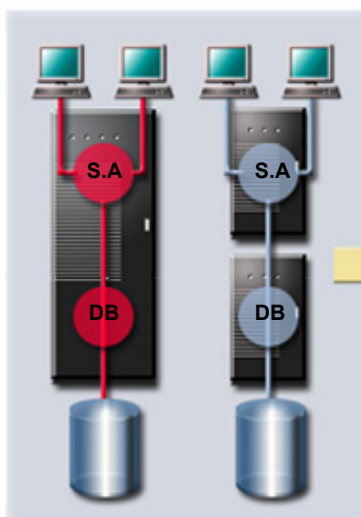
- Disks to store data and transactions
- Memory to manage connections, data caches, instructions and transactions
- CPUs to run processes and transactions.

Regardless of database size, there is a minimum disk resource requirement of approximately 300 Mb for the Oracle catalogue.

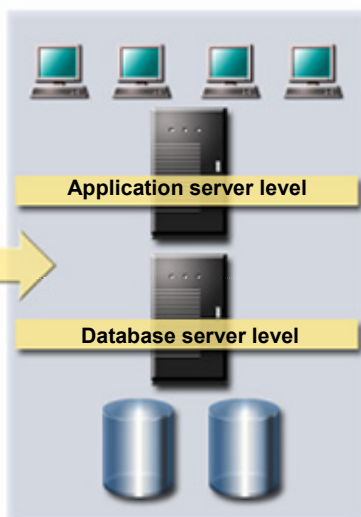
The number of logical partitions available on Bull Escala servers enables processor and memory resources, and even I/O controller resources, to be distributed between partitions. ASM enables storage space to be shared between:

- Multiple databases hosted on the same Bull Escala server or partition
- Multiple databases hosted on different nodes of an Oracle 10g RAC cluster.

Distributed model



Consolidated model



### Deployment of smaller databases

Deploying a smaller database (between ten to a few hundred Gb) on a dedicated partition or micro-partition is not an optimum solution, a management by partition or micropartition is required with:

- A storage space for the binaries and the Oracle database
- An Oracle instance
- CPU, memory and controller resource allocation policy.

Moreover Logical consolidation (in which multiple databases share the same partition of a Bull Escala server under WLM control) optimizes only storage management:

- A single file system for Oracle Binaries
- A single set of ASM DISKGROUPS for all databases.

As a matter of fact physical and logical consolidation of smaller databases shares the same disadvantages:

- Multiple databases to be administered
- Multiple copies of the Oracle catalogue
- Multiple instances to be booted, etc.).

Resulting in an excessive consumption of processor and memory resources.

Oracle 10g EE or Oracle 10g RAC database consolidation is the model best suited to this type of database.

The following table summarizes the components and the contribution they make to this type of consolidation.

Deployment of smaller databases			
Advantages	Database consolidation under Oracle 10g EE	Consolidation under Oracle 10g RAC between 2 partitions of the same server	Consolidation under Oracle 10g RAC between partitions of different servers
ASM	<ul style="list-style-type: none"> <li>- I/O Optimization</li> <li>- Development within RAC is possible at any time</li> </ul>	<ul style="list-style-type: none"> <li>- I/O Optimization</li> <li>- Disaster Recovery ("failure Disk group" on different disks)</li> </ul>	
Oracle Resource Manager	<ul style="list-style-type: none"> <li>- Controls resources between services or schemas</li> </ul>		
Service	<ul style="list-style-type: none"> <li>- Linking to ORM</li> <li>- Vacation Management</li> </ul>	<ul style="list-style-type: none"> <li>- Linking to ORM</li> <li>- Vacation Management</li> <li>- Distribution of services between instances</li> </ul>	
RAC	N/A	<ul style="list-style-type: none"> <li>- Enhanced transactional performance</li> <li>- Availability: service accessible as long as an instance is running</li> <li>- Option to dedicate an instance to a service temporarily to cope with peak loads</li> </ul>	
		<ul style="list-style-type: none"> <li>- Evolution towards inter-server RAC.</li> </ul>	<ul style="list-style-type: none"> <li>- Enhanced scalability</li> </ul>

### Deployment of large databases

With large databases of several Terabytes, the resource sharing problem is primarily one of memory and processor resources.

This type of database may be deployed in its own partition on a Bull Escala server, or in Oracle cluster mode between two Bull Escala servers.

The choice is then between « physical consolidation models in Oracle 10g EE mode or Oracle 10g RAC inter-server mode » and « database consolidation under Oracle 10g RAC ».

The decision of whether to choose a dedicated model « physical consolidation under Oracle 10g EE, with or without RAC » or a mutualized model « database consolidation under Oracle 10g with RAC » will essentially depend on database size and anticipated I/O flow.

The processor power requirement can be covered by adding nodes to an Oracle 10g RAC cluster in either the dedicated or consolidated models. In database consolidation mode, it is also possible to dedicate some cluster nodes temporarily to one of the services.

Deployment of large databases			
Advantages	Dedicated DLPAR	RAC between dedicated inter-server DLPARs	Oracle 10g RAC consolidation
ASM	<ul style="list-style-type: none"> <li>- I/O Optimization</li> <li>- Evolution towards RAC is possible at any time</li> </ul>	<ul style="list-style-type: none"> <li>- I/O Optimization</li> <li>- Disaster Recovery ("failure Disk group" on different disks)</li> </ul>	
Oracle Resource Manager	<ul style="list-style-type: none"> <li>- Controls resources between services or schemas</li> </ul>		
Service	<ul style="list-style-type: none"> <li>- Associated vacation management.</li> <li>- Linking to ORM</li> </ul>		<ul style="list-style-type: none"> <li>- Linking to ORM</li> <li>- Vacation Management</li> <li>- Distribution of services between instances</li> </ul>
RAC	N/A	<ul style="list-style-type: none"> <li>- Evolution towards inter-server RAC.</li> </ul>	<ul style="list-style-type: none"> <li>- Performance</li> <li>- Availability</li> <li>- Scalability</li> </ul>

# Oracle consolidated architecture on Escala

The preceding chapter adopted the viewpoint of the database administrator who must answer the question « which deployment model should I adopt for which type of database? »

This chapter attempts to answer the question « how should a set of Bull Escala servers be structured to consolidate Oracle databases? »

Bull Escala servers deliver a transactional power level of over 3,000,000 TPM-C. A single server can therefore host partitions with large non-consolidated databases (NC) alongside partitions with large and small consolidated databases (CD).

ASM is systematically used for database storage in order to retain the option of evolving to Oracle 10g RAC.

## ASM Virtualisation

ASM also allows storage space virtualisation for a set of databases hosted on the same Operating System instance or on the same cluster: this kind of cluster is called an ASM Virtualisation Cluster.

In nominal mode, a DLPAR (DEDICATED PARTITION or SHARED PROCESSOR PARTITION) hosts one ASM instance and one Oracle database, whether consolidated or not. A very large database might have a dedicated:

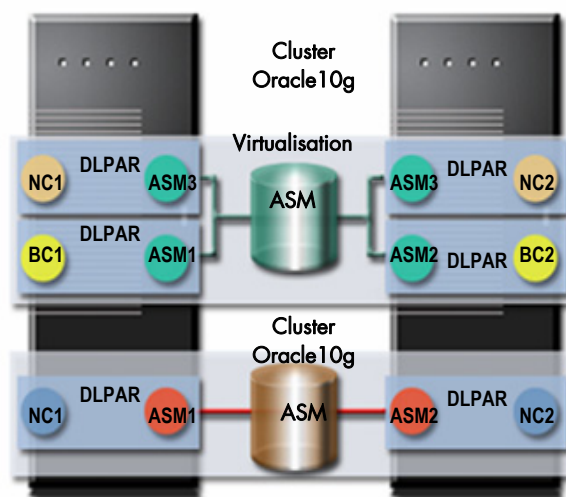
- ASM storage space
- Oracle 10g cluster between inter-server DLPARs.

## ASM Virtualisation Cluster

This architecture offers a high level of flexibility. BC (Consolidated) databases and NC (non-consolidated databases) run on the same Oracle 10g cluster.

There are therefore two possible options for coping with temporary load peaks in the NC database:

- Using the Oracle 10g RAC grid concept by halting the NC1 instance on node 3 to start the BC3 instance (for example)
- Using the ability of the DLPARs to migrate dynamically from the node 3 processors to node 1 and from node 4 to node 2.



### Logical Consolidation

Where there are frequent, significant and/or unpredictable variations in peak load, it would be possible to deploy multiple databases in logical consolidation mode within an Oracle 10g RAC cluster.

In nominal mode, partitions containing multiple databases host:

- One ASM instance
- Two or more database instances.

Where databases (typically two) with complementary load profiles occupy the same DLPAR, the peak load periods of one should coincide with the minimal load periods of the other. Resources will be distributed by the WLM on the basis of the Unix process name, which enables easy database identification. It is always possible to halt the instance of a database in order to temporarily dedicate a cluster node to a single database. This architecture offers high availability and is extremely responsive to variations in workload.

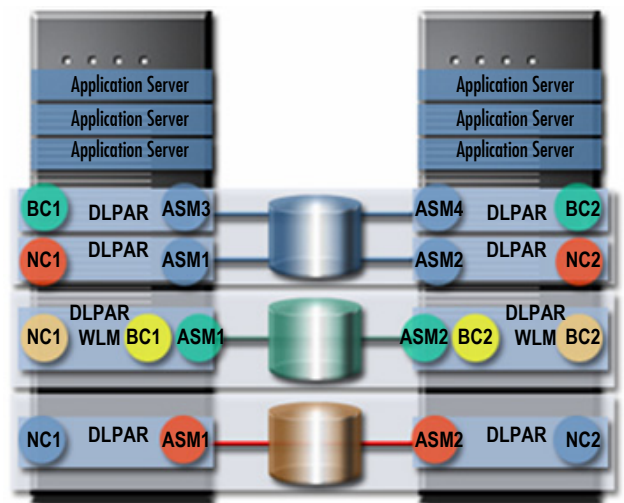
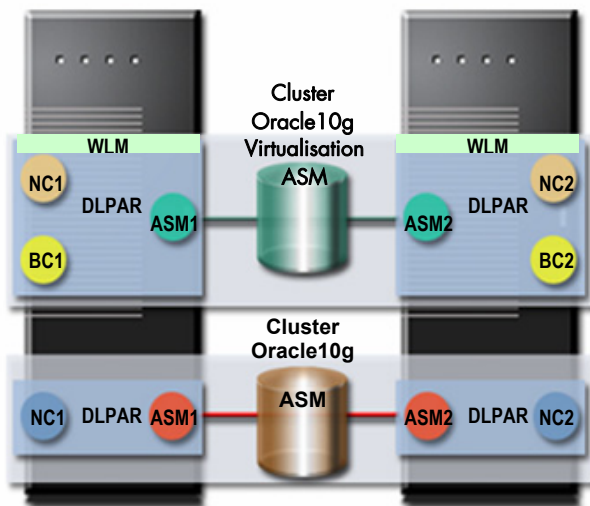
### Global Infrastructure

In global terms, a Bull Escala server contains one set of partitions for the application server level and another set of partitions for the database server level.

Within a Bull Escala server, databases allocated to the same ASM virtualisation space will be deployed using one of the following two modes:

- Logical consolidation within a partition, where processor resources are controlled by WLM
- ASM Virtualisation cluster comprising several DLPARs.

In principle, micro-partitions are used for the application server level, whilst the database server level may use either type of partition: DLPAR (DEDICATED PARTITION) or SHARED PROCESSOR PARTITION.



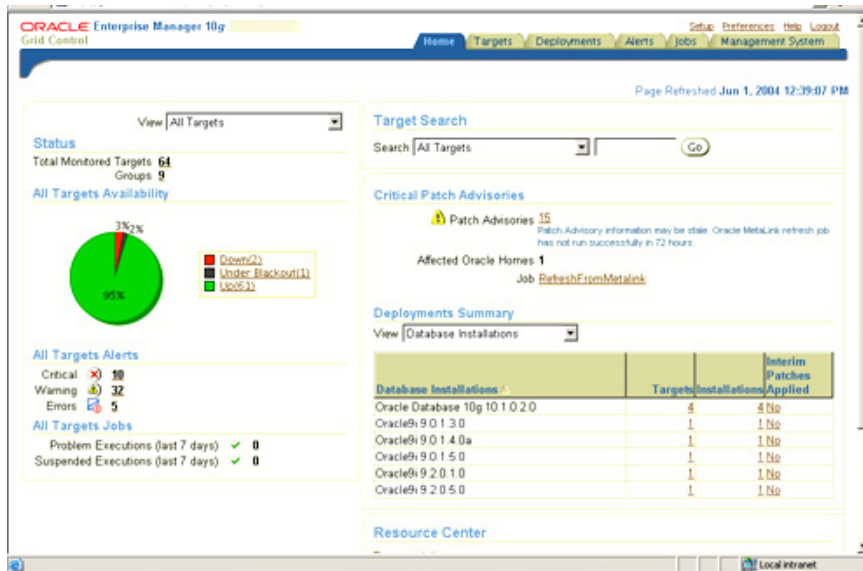
# Simplified Administration and support

## Simplified Administration

The supervision and administration of this type of environment is considerably simplified by the use of Oracle Enterprise Manager Grid Control, which administers, supervises and provides configuration management for the entire environment, including servers, databases (whether clustered or not) and listening processes (LISTENERS).

## Bull's expertise on Oracle

Bull has developed a long term expertise on Oracle, allowing it to design, develop and deploy complete Oracle solutions. Bull uses its high-level technological expertise and relies on a global approach to optimize the operation of Oracle solutions, particularly in an Escala environment, by making cost control one of its priorities.



## Consulting services in the decision-making field

Bull is a major player in information systems. Always listening to its customers, it was one of the first to invest in infocentres, the precursors to today's decision-making systems. The Business Intelligence division of Bull Services responds to the needs of companies and administrations in the field of integrating decision-making solutions. This mastery of decision-making projects consists of forming multidisciplinary teams to gather a full range of skills – technical, functional and methodological.

## Bull's Oracle Competence Centre

Bull's European Oracle Competence Centre masters even the most complex Oracle infrastructure solutions. Bull's experts commit to providing perennial Oracle infrastructures to customers, while optimizing the profitability of their investments.

# Conclusion

The deployment of large databases and the consolidation of databases in Oracle 10g RAC clusters enable the dynamic partition technologies (DLPARs) of Bull Escala servers to be combined with Oracle 10g RAC cluster technologies. This method obtains maximum performance at minimum cost through:

1. simplified storage management via ASM Virtualisation,
2. fewer databases to administer as a result of database consolidation in the form of a multi-schema database,
3. reduction in the over-consumption of resources as a result of reducing the number of databases,
4. improved use of processor and memory resources as a result of DLPAR dynamic migration functionalities,
5. ability to cope with temporary load peaks with no need for over-sizing as a result of the grid functionality offered by Oracle 10g RAC or DLPAR dynamic migration functionalities,
6. ability to distribute processor resources automatically via WLM, where databases are deployed in the same partition,
7. protection of investment as a result of Oracle 10g RAC cluster scalability to cope with increased load,
8. simple configuration monitoring and administration delivered by the Oracle Enterprise Manager Grid Control administration tool.

*For more information :*

[www.bull.com](http://www.bull.com)



The power of choice