IBM System p for Oracle Data Warehousing

---

**Highlights**

- Reduces the complexity and risk of Oracle Data Warehousing deployments
- Incorporates IBM System p™ and System Storage™ solution options into IBM and Oracle Data Warehousing reference configurations developed collaboratively as part of the Oracle Information Appliance Initiative
- Offers a family of validated and tuned pre-sized configuration building blocks to support a range of raw data, user and query workload requirements
- Provides repeatable, balanced building blocks to scale-out the data warehouse
- Seamlessly incorporates Extract, Transform and Load (ETL) and application loads into the data warehouse

---

**IBM System p**

A reliable, scalable, resilient server infrastructure is critical for application deployments that rely on an Oracle database. IBM System p technology is a smart choice for UNIX® and Linux® operating system-based Oracle database deployments and for businesses that demand powerful, flexible, reliable and secure computing solutions. An extensive system family from 1- to 64-core scalability is competitively priced with 64-bit IBM POWER5™, POWER5+™, and POWER6™ technology designed to lower software, energy and space costs through leadership performance and unique IBM capabilities that can enable a dramatic increase in individual system utilization. A choice of more than 10,000 IBM AIX® and Linux applications supports a broad array of business requirements. And models can be selected for specific workloads, from front-end business intelligence (BI) applications to back-end data warehouse (DW) infrastructure.

**System p technology leadership**

IBM System p models are based on advanced IBM POWER™ Dual-Core chip technology and deliver outstanding price/performance, mainframe-inspired reliability features and innovative virtualization capabilities. Add to that IBM’s unique simultaneous-multithreading
technology allowing two application threads to be run at the same time, and it’s easy to see why the IBM System p server delivers outstanding performance.

**System p virtualization**

System p virtualization technology can dramatically increase server utilization allowing workloads to be easily consolidated and enabling you to do more than ever before with a single server. Available as an option on all System p models, Advanced POWER Virtualization allows you to run multiple applications on AIX and Linux operating systems at the same time in separate, secure partitions using as little as one-tenth of a processor — allowing a reduction in the number of servers and peripheral devices needed. And the System p Capacity on Demand function available on selected models is designed to scale-up and scale-down processing power and memory as business needs fluctuate with no disruption to your business.

**IBM System Storage**

Disk storage is a critical element in an Oracle Data Warehousing environment. Multiple solution options are available including:

- **IBM System Storage DS8000™ family** - designed to deliver robust, flexible, highly available, and cost-effective disk storage to support continuous operations for mission-critical workloads
- **IBM System Storage DS4000™ family** - scalable, modular Fibre Channel disk storage designed with growth, reliability, and availability in mind, from entry-level to enterprise environments and performance-oriented to capacity-oriented applications

For more information regarding IBM System Storage product compatibility with the Oracle database please visit: [ibm.com/servers/storage/oraclecompatibility.html](http://ibm.com/servers/storage/oraclecompatibility.html)

**IBM and the Oracle Information Appliance Initiative**

IBM has developed data warehouse reference configurations, called Information Appliance Foundations, as part of the Oracle Information Appliance Initiative. These foundations combine Oracle Data Warehousing components with IBM System p and System Storage products.

These Oracle Information Appliance Foundations provide validated, balanced configurations for data warehouses that combine system resources such as hardware, storage, I/O and networking into data warehouse building blocks that can be combined to address different scalability needs in a linear fashion. They’re designed to support varying loads based on varying raw data size, concurrent user load and varying query complexity.
The development and use of repeatable building blocks allows high-performance data warehouses to scale through the use of a modular design approach to the data warehouse, ETL functionality, business intelligence tools and applications. The modular building blocks are designed to deploy into existing client business intelligence and OLTP (On Line Transaction Processing) infrastructures. For more information about the Oracle Information Appliance Initiative please visit: oracle.com/solutions/business_intelligence/oiai.html

Business intelligence reference architecture for Oracle on IBM Systems.

Figure 1 outlines the IBM reference architecture for the deployment of Oracle business intelligence components on IBM Systems and IBM System Storage products. The architecture is a high-level system design. It is free of implementation details and provides a high level description of the solution components.

The key elements of the reference architecture are the:

- **Software Architecture Components** - these define the overall structure and relationships among the key functional elements of the data warehouse, the infrastructure software and BI server repository
- **Systems Architecture** - this defines a proven approach for insuring that a balanced set of system resources are in place to deliver expected performance based on the primary drivers of data warehouse performance - including computer power, network bandwidth and storage capacity and bandwidth

The software architecture is made up of three primary components or groups. The first group is made up of those components that comprise the data warehouse. These include the relational database (Oracle's latest Enterprise Edition release with Oracle Partitioning), storage management (Oracle Advanced Storage Management product for logical volume management of the database objects) and ETL functionality (Oracle Warehouse Builder). For existing deployments the architecture is flexible enough to allow third party ETL products such as IBM Ascential, Informatica, SAS or Business Objects to be used as the ETL driving mechanism.

The second group within the software architecture comprises the infrastructure including administration and management (Oracle Enterprise Manager), cluster control (Oracle Real Application Clusters (RAC)) as well as hardware (IBM System p) and storage (IBM System Storage).

---

**Figure 1. Business intelligence reference architecture for Oracle on IBM Systems**
The third software architecture group consists of the BI server repository (Oracle BI Enterprise Edition) and the individual applications that the client uses to access the data warehouse. These applications can be specific Oracle applications or third party tools from other vendors such as SAS, Business Objects, Cognos or Microstrategy.

The systems architecture that IBM has developed which supports the Oracle Information Appliance Foundation and the IBM business intelligence reference architecture for Oracle on IBM Systems will be described in the next section.

**Functional components of the business intelligence reference architecture for Oracle on IBM Systems**

The logical flow of data from the data source to delivery can be defined as a series of processes that accommodate data integration, data warehousing and data analytics. Each of these areas exhibit different workload and resource characteristics. These characteristics can be defined and resource requirements articulated through best practices and workload sizing and capacity planning methodologies.

Figure 2 outlines these different areas and defines the approach of repeatable building blocks (called Oracle nodes) for ETL, data warehousing and analytics applications. These building blocks or nodes relate back to the reference architecture and comprise the systems architecture described earlier. The use of nodes in this fashion allows the client to integrate existing ETL and application deployments into the data warehouse using appropriate sizing techniques.

Each of the nodes are designed using IBM System p and System Storage hardware with associated interconnect technology to scale in an Oracle RAC-managed environment. The nodes provide flexibility allowing a deployment to start in an SMP environment and then to scale-out in an Oracle RAC deployment as the data warehouse grows.

Each functional node (Figure 3) is sized based on workload characteristics in terms of I/O bandwidth, memory requirement and CPU utilization. The holistic design is that data integration nodes are balanced not only across the internal resources available but also balanced in terms of performance with the other nodes in the entire business intelligence solution. The Oracle data warehouse node is designed to provide balanced, scalable

---

**Figure 2. Business intelligence reference architecture for Oracle on IBM Systems**

- **Data Sources**: Enterprise, Unstructured, Informational, External
- **Data Integration**: Extraction, Transforms, Load / Apply, Synchronize, Transport / Messaging, Information Integrity
- **Data Repositories**: Operational Data Stores, Data Warehouses, Data Marts, Staging Areas, Metadata
- **Analytics**: Collaboration, Query & Reporting, Data Mining, Modeling, Scorecard, Visualization, Embedded Analytics
- **Access**: Web Browser, Portals, Devices, Web Services

- Oracle Informatica, SAS, Business Objects, IBM
- Oracle Siebel Analytics, Hyperion, SAS, Business Objects, Cognos, Microstrategy
performance of the data warehouse as multiple nodes are connected together into a RAC-enabled data warehouse deployment.

The individual nodes are the core components of the Oracle Information Appliance Foundation (Figure 4). The business intelligence solution is deployed based on a detailed sizing exercise which defines the number of Oracle data integration nodes, Oracle data warehouse nodes and Oracle analytics nodes that are required. The deployment uses the infrastructure components outlined in the reference architecture based on IBM System p, System Storage, Oracle RAC, Oracle Partitioning and Enterprise Manager.

**Sizing functional building blocks**

**How to size Oracle Business Intelligence and Oracle Data Warehousing solutions**

IBM offers a process for sizing future hardware requirements when a client is looking to run Oracle Business Intelligence and data warehouse solutions on IBM hardware. This process is based on performance data and other information gathered from the client’s existing environment. This input is used to estimate the resources required to support one or more of the following scenarios:

- **New Oracle BI or DW installations**
- **Additional applications for an existing Oracle BI or DW production environment**

In order to start the sizing estimate process, follow the instructions on the cover page of the Oracle database sizing questionnaire. The questionnaire provides information on what needs to be completed and where to send the completed document for processing. Please work with your IBM representative or IBM Business Partner in order to obtain a sizing estimate. To access the Oracle database sizing questionnaire please visit: [ibm.com/erp/sizing](http://ibm.com/erp/sizing)

---

*Figure 3. Business intelligence functional building blocks*

*Figure 4. Business intelligence deployment comprising functional nodes*
Once there, select the Oracle Database sizing questionnaire from the list of supported application systems.

On the same web page just mentioned you can access the Oracle Business Intelligence quick sizer from the same list. This is a simplified tool that provides general sizing guidelines on possible IBM hardware configurations when running Oracle BI. It does not replace the standard IBM sizing process but can be an excellent complement or starting point for the sizing estimate.

The Oracle BI quick sizer is a tool developed by IBM and Oracle. It is used as an initial sizing reference with clients who wish to build a new, or extend an existing Oracle Database data warehouse. The BI quick sizer is designed to provide a reference point for discussion around the deployment of Oracle Database 10g and, where appropriate Oracle RAC on IBM System p or IBM System x server technology. The Oracle BI quick sizer provides a discussion structure whereby the client can consider a number of factors:

- Comparisons of growth scenarios around scale-out or scale-up strategies
- Comparisons between System x with Linux and System p with AIX solutions.
- Comparisons of servers for different query workloads
- Storage configurations based on amount of usable disk and Host Bus Adapters (HBAs aka Fiber Channel adapters)

In addition to sizing methodologies for Oracle Data Warehousing, ETL and business intelligence applications, IBM has sizing methods and practices for many of the third party tools that can be seen in a complex heterogeneous business intelligence deployment including SAS, Business Objects, Cognos and Microstrategy. These sizing practices and methods are used to design the analytics nodes as part of the complete deployment. The availability of all these sizing paths provides a comprehensive sizing solution for the client sizing needs in this market segment.

**Oracle data warehouse node for IBM System p and System Storage**

As noted, the Oracle data warehouse node is the core component in

| Table 1. Infrastructure elements of a balanced building block for an Oracle data warehouse node. |
| --- | --- |
| Architecture / OS | POWER5 2.2 GHz / AIX 5.3 |
| Server | IBM System p5 570 |
| Processors per Block | 8 |
| Memory per Block | 32 GB |
| Internal Storage | 4 x 73 GB 15K rpm HDD |
| Fiber Channel HBAs per Block | 8 x Dual Port 4 Gbps |
| Interconnects per Block | 4 x 10/100/1000 Ethernet |
| SAN Switch | 16 port SAN Fibre Channel Switch |
| External Storage Configuration | 1 x DS4800 with 10 x EXP810 Expansion Units – 160 x 145GB 15K rpm HDD |
| Raw / Usable Storage Space | 5 TB / 18 TB |
constructing a scalable data warehouse which is the underlying driver of the business intelligence solution. Table 1 identifies the infrastructure elements of an Oracle data warehouse node that can be found in the Oracle BI quick sizer.

The node provides 5TB of raw data warehouse space providing the flexibility for a moderate deployment to grow in a scale up fashion to 5TB before then scaling outwards in an Oracle RAC deployment for data warehouses greater than 5TB. The node is designed to provide balanced memory, I/O and processing power to efficiently execute application queries. The node is designed to service between 15 and 90 concurrent users depending on the complexity of the user queries.

Summary
One of the major areas of concern for business is the effort involved in architecting, developing and deploying complex data warehouses and the associated components of the business intelligence solution around ETL and analytic applications. IBM and Oracle have developed a powerful architectural model and approach to simplifying the deployment and reducing the risk of data warehouses based on the IBM System p for Oracle Data Warehousing. This solution fully supports the Oracle Information Appliance Initiative and provides a family of optimized and validated pre-sized configuration building blocks for data warehouses to support a range of raw data, user and query workload requirements.

For more information
To explore other System p and Oracle Data Warehousing solutions or to find out more about other joint solutions from IBM and Oracle, please contact an IBM sales representative at 1 866 426-9989, or visit us at: ibm.com/solutions/oracle

For more information about the IBM System p please visit: ibm.com/systems/p

For more information about IBM System Storage please visit: ibm.com/storage