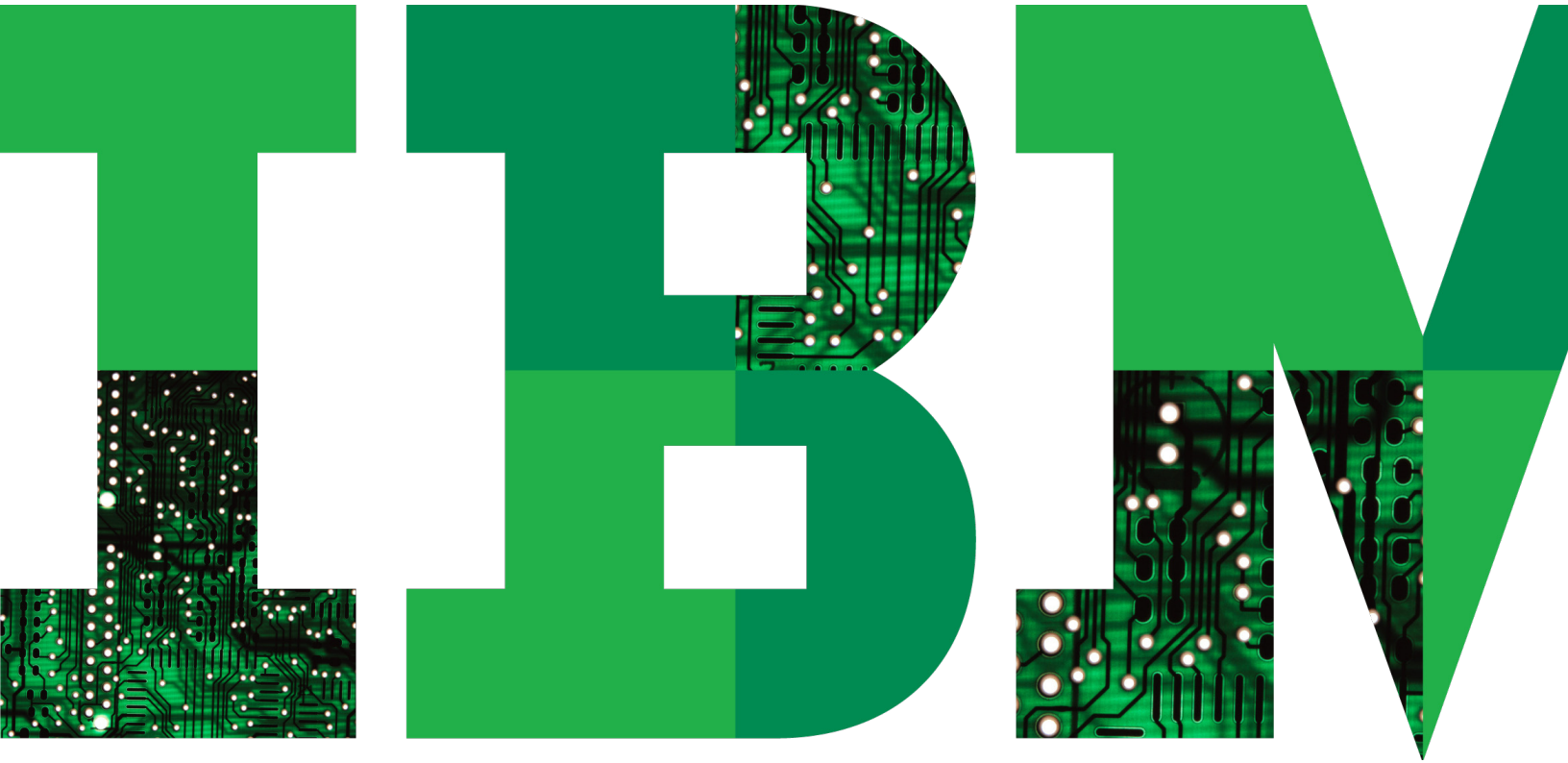


Data warehousing and business intelligence for IBM System z

*Effective use of System z in an Information Agenda
for a Smarter Planet*



Executive summary

Companies are faced with massive amounts of data that, when managed strategically, can be transformed into a wealth of information for forward-thinking business decisions. Organizations have increasingly depended on data warehouses for collecting, storing and staging their corporate data in order to provide a solid, trusted source for decision making. Importantly, the underlying infrastructure of the data warehouse delivers the data needed for today's critical business intelligence (BI) strategies, helping to enable business insight for decision making through reporting, analysis, visualization, statistical data mining and other key functions. However, the traditional data warehouse is only one element in an enterprise's information resource. Many organizations are exploring unstructured data—the information hidden in text files, external Web sites, memos and more. It's also becoming clear that there is BI value in data that few organizations would consider storing in a data warehouse. The Smarter Planet™ initiative from IBM shows how to broaden the scope of the enterprise to include information with a global reach as well as being hosted on global sites.

Today, the mainframe plays a central role in operational BI scenarios and enterprise data warehouses since most operational data is captured and hosted there. Operational BI is geared toward access to operational data and delivery of BI processes to individuals, such as customer service representatives. This trend extends the value of data as a corporate asset, while extending the reach and value of BI tools to a more holistic population.

The growing focus on operational BI, coupled with business demands, such as power and cooling constraints, consolidation objectives, budget limitations, risk management initiatives and virtualization goals, is generating a new interest in the mainframe as the optimal data warehousing and BI solution.

IBM has invested billions of dollars and assigned many new individuals to the support of the mainframe as its popularity has increased. Customer-driven initiatives and recommendations by industry analysts have spurred on this investment. IBM offers not only a flexible, scalable, cost-effective mainframe platform—IBM® System z®—but also delivers a portfolio of innovative BI and data warehousing solutions for the mainframe. Many new products have been unveiled during the past 18 months with more to follow. New solutions such as cloud computing are being introduced on System z due to its ability to provision and host large numbers of users within a secure and continuously available framework. IBM has announced the Smart Analytics Cloud for System z as a means to help customers rapidly and seamlessly implement and deploy large numbers of end users on a global basis.

The emerging trends for BI tools to access data on demand and the surge in software and hardware enhancements for IBM System z and IBM z/OS® are encouraging many enterprises to re-examine their current data warehouse and BI strategies, and investigate how they can maximize the value of the IBM System z for strategic advantage.

Building an Information Agenda

Building an Information Agenda is about the effective interoperation between IT and the business units, having mutual goals and a support infrastructure in place to address information needs globally. It transcends internal data and information sources and looks at information in all sources and locations as possibly being relevant to the business. However, one needs to begin by putting the internal data infrastructure in place. BI is touted by many CIOs as “mission critical” and they perceive it to be imperative to building and maintaining a competitive advantage. BI tools provide historical, current and predictive views of business operations—such as sales, production and finance through reporting, analysis, visualization, statistical data mining and other BI functions. IBM recently closed its acquisition of SPSS as a move to enhance its portfolio in the advanced analytics space. IBM has emerged as one of the most comprehensive BI provider in the industry with a full portfolio of products and solutions to meet practically any business and technology requirement.

When selecting the best platform for data warehousing and BI, companies must consider the platform’s ability to deliver the best value to the enterprise and its end users. They need useful, enterprise-wide data that is stored securely and readily available to help increase productivity, competitiveness and customer satisfaction. Time and again, System z is the platform of choice.

Business considerations

Various industry trends and business considerations drive an enterprise to examine and sometimes restructure their data warehouse and BI strategies and implementation. Some important considerations include:

- Information Agenda infrastructure creation for agility and growth
- Security and regulatory compliance
- “Server farm” maintenance costs and environmental issues, such as power consumption
- Server consolidation for better control and lower costs
- Data warehouse building and updating cycles
- BI tools standardization and operational BI: real-time BI
- New hardware and software for System z (e.g., z/VM[®] virtualization and Linux[®])
- Federated (heterogeneous) data access and new data formats (e.g., XML)

Total cost of ownership considerations

When choosing a BI and data warehousing platform and assessing associated costs, many organizations focus mainly on hardware and software. However, hidden costs need to be weighed accordingly. For example, every server in an enterprise must be managed, which involves labor costs. Each server will also require software, maintenance, upgrades, systems administration, monitoring, backup and various associated IT processes. Some companies also face power issues, where increasing their current energy consumption is not an option.

The total cost of ownership for any application platform is based on many factors. Hidden costs can greatly impact the total cost of platform ownership and are often overlooked. Consider, for example, a system’s value to an enterprise when costly downtime is a regular occurrence. Security is also a growing economical concern today, with security breaches at times resulting in significant monetary fines. The following industry quotes substantiate some of the concerns enterprises face—including management, security and availability—when considering platform options:

- **Management and administration**
 - Any system requires some amount of software, maintenance, time and attention. This typically

demands human intervention and resource. The costs of supporting distributed environments have soared in comparison with the mainframe (see Figure 1).

- **Security breaches**
 - There has been “a 43-percent rise in costs compared to 2005 as affected companies scrambled to notify customers.”¹
 - “The total average cost of a data breach grew to \$197 per compromised record, an increase of 8 percent since 2006 and 43 percent compared to 2005.”¹
 - “The average total cost per reporting company was more than \$6.3 million per breach and ranged from \$225,000 to almost \$35 million.”¹

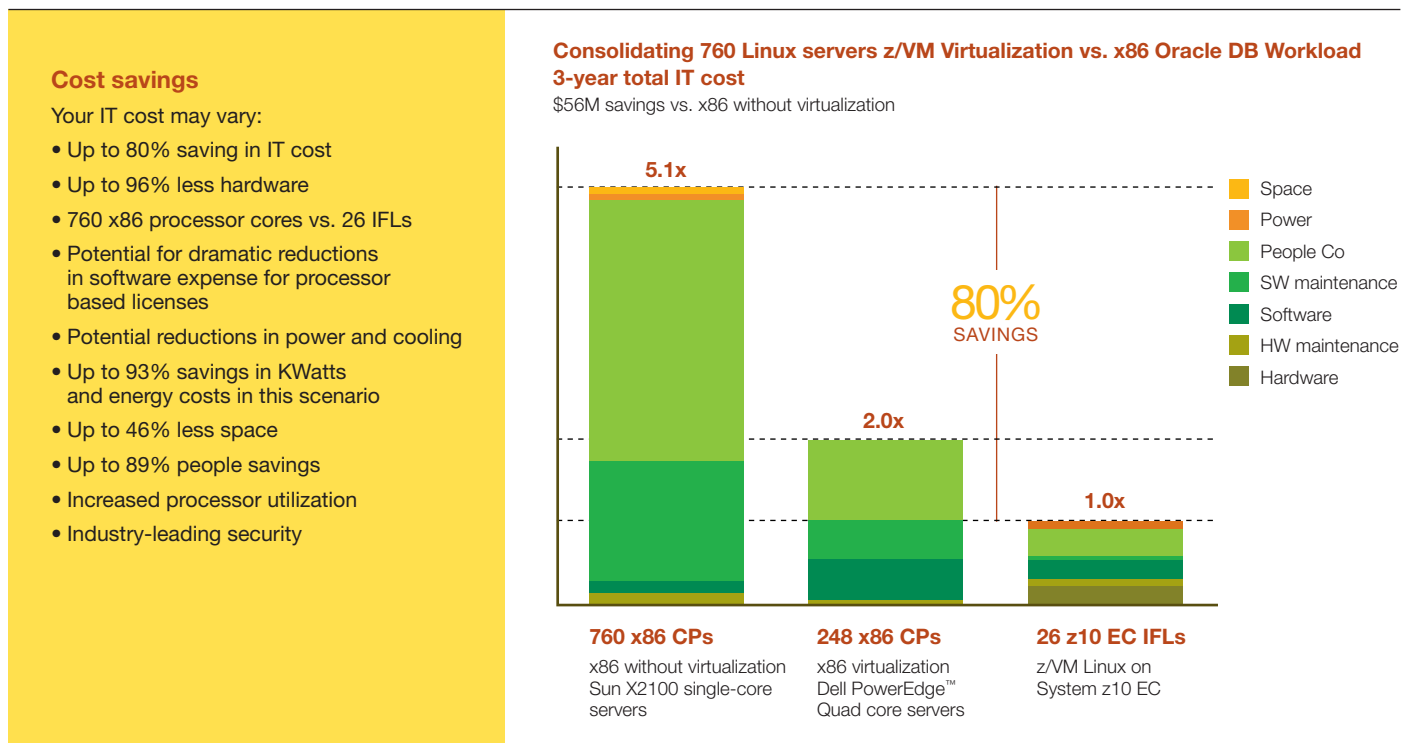


Figure 1: Three-year total IT cost savings comparison of System z to a distributed environment

- **Downtime**

- “Cost of downtime can vary by industry and can range from hundreds of thousands to millions of dollars per hour.”²

In assessing the potential costs of employing BI and data warehousing on the various platforms available, companies should align their platform choices with their plans for business growth and expansion. They should ultimately select and build data warehousing architectures based on how they use information—or more importantly, how they can optimize their corporate data day to day to empower them in their highly competitive marketplace today, and in the future. They need to heavily factor in changes in business requirements, such as the build cycle time for the current warehouse or changes in data currency (real-time) requests.

In a distributed server environment, adding capacity requires adding servers, thus costs go up linearly with expanding workloads. In a mainframe environment, an enterprise will run a “mixed” workload of production, data warehousing and more. The mainframe is optimized for heavy use and companies can add incremental capacity without adding resources to manage the environment. In the end, companies need to select the platform that will deliver the highest return on investment possible, given costs specific to their environment.

The versatile mainframe

The mainframe is important to businesses that require reliability, security, performance, low total cost of ownership, resiliency, scalability and more, and is arguably the most successful business-computing platform in history.

Given the prevalence of the distributed platform beginning in the 1990s, along with the solutions to support it, some companies wonder if the data warehouse can be built and maintained on the mainframe today with the same support and infrastructure as on a distributed platform. Until recently, only some components of the data warehousing software infrastructure were available on the mainframe. Therefore, companies previously opting for a mainframe data warehouse had to customize their own solutions. Today, customer demand and a surging interest in the mainframe for data warehousing and BI have resulted in new, significant end-to-end offerings for running data warehousing and BI on the mainframe platform. Some of the new offerings are:

IBM InfoSphere Information Server for Linux on System z

A fully integrated software platform that profiles, cleanses and transforms information from mainframe and distributed data sources to drive insight without added z/OS operational costs. It offers a full array of data warehousing functionality from data definition to transformation and loading (Figure 2).

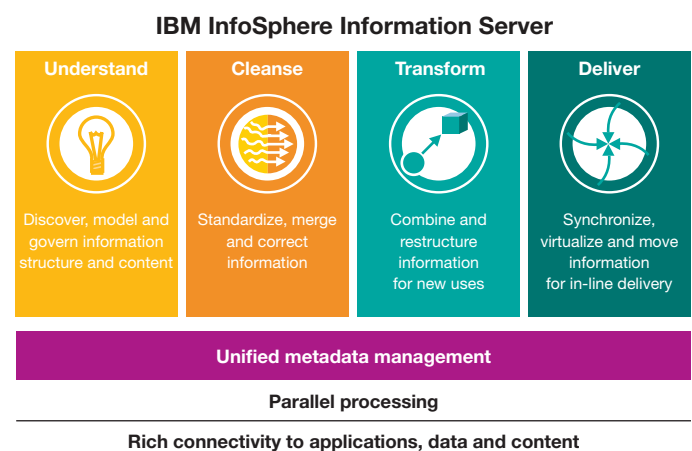


Figure 2: The IBM InfoSphere Information Server offers full-range functionality.

IBM InfoSphere Warehouse for Linux on System z

This data warehouse accelerator complements Information Server by offering a DB2®-based set of building and query execution enhancements that make it easier to create a new data warehouse or data mart as well as improve multi-dimensional (MDX) query execution. It simplifies operational complexity by deploying both operational and warehouse data on a single platform (see Figure 3).

IBM InfoSphere Master Data Management Server

Provides the strategic architecture companies need to solve critical enterprise master data management issues by managing customer, account and product data centrally. Gives your organization an operational environment to store, maintain and update your organization’s critical information about customer, product and account data.

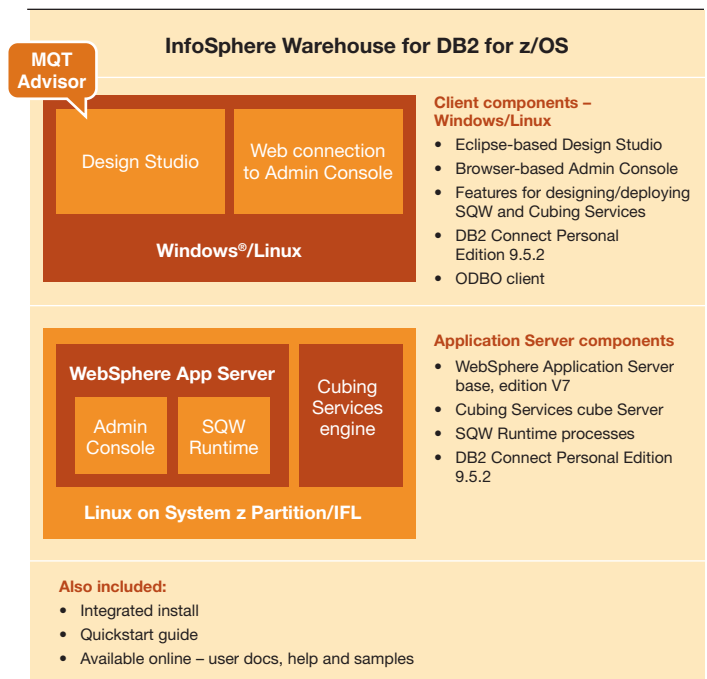


Figure 3: The IBM InfoSphere Warehouse

IBM Cognos 8 BI for Linux on System z

Cognos® offers a full array of BI and performance management tools and solutions that have been well integrated within the IBM portfolio. Cognos 8 provides a BI platform running under Linux on System z with reporting, analysis and dashboard capabilities in easy-to-deploy, manage and use software that runs on your mainframe platform (see Figure 4).

IBM Smart Analytics Cloud for System z

A suite of customizable products and services targeted at large-scale BI deployments via a private cloud. Enterprises seeking a means to standardize their BI platform, reduce costs, easily deploy and administer such environments will find this offering strongly appealing.

Cognos 8 BI for Linux on System z

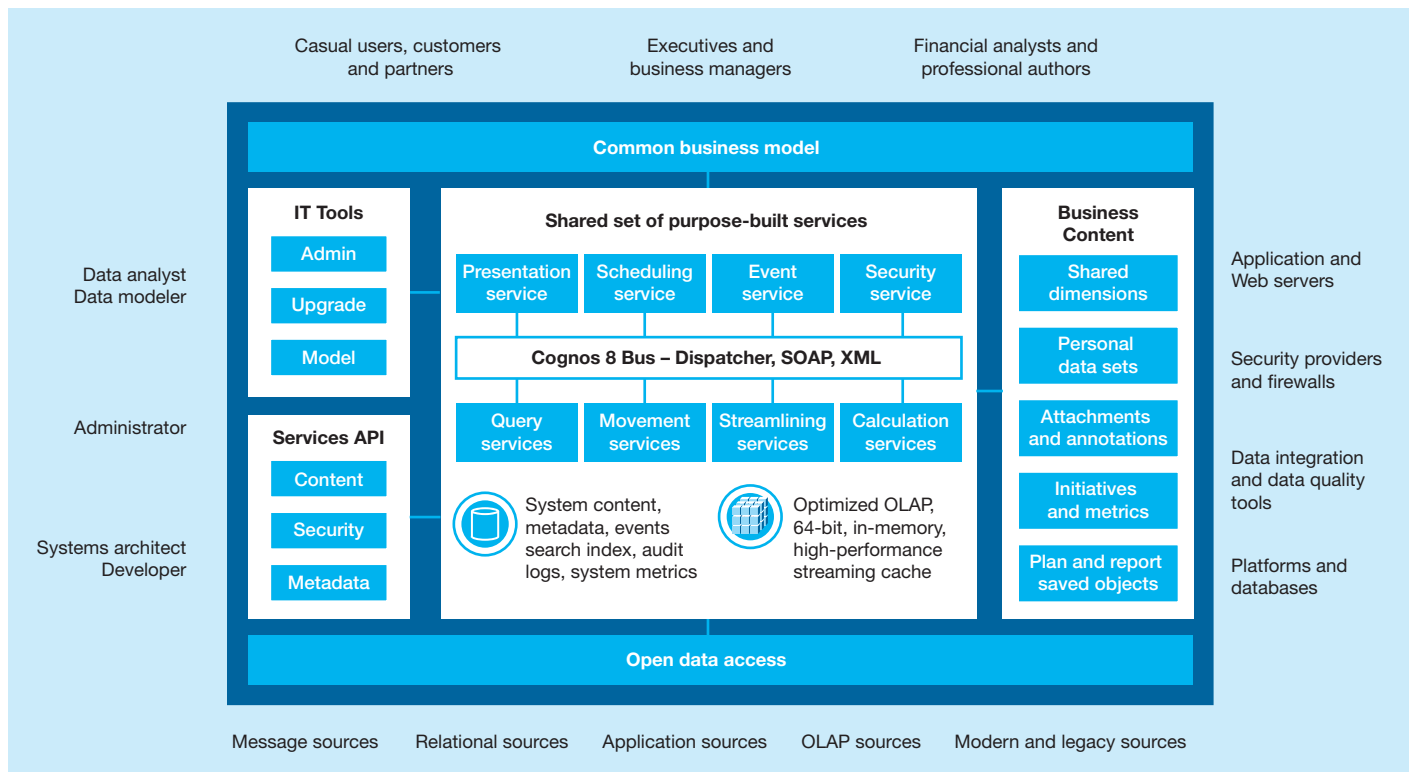


Figure 4: Cognos 8 BI for Linux on System z has a large selection of BI and performance management tools and solutions.

IBM Smart Analytics Optimizer for System z

A hardware offering currently under development and being tested within IBM labs, the Optimizer is a new blade server which will be used to intercept certain large, complex queries from DB2 for z/OS and re-route them to provide extremely improved performance.

IMS Version 11

IMS™ has added significant new features including a native JDBC™ driver to provide direct access to IMS data from tools that support it. In scenarios where operational BI may be of keen interest, this removes the requirement to either extract data from IMS or having intermediary software required for data access.

IBM SPSS

Advanced analytics (predictive analytics, etc.) has long been a requirement for anyone's well-rounded BI infrastructure. IBM opted to acquire SPSS to provide one of the industry leading solutions as part of its portfolio and is in the process of finalizing all the necessary requirements.

Delivering value to the enterprise

The benefits of implementing a data warehousing and BI solution on the mainframe will vary by company and the characteristics of a particular enterprise. According to BI and data warehousing experts such as Claudia Imhoff, Ph.D., the mainframe may be the optimal platform for data warehousing and BI solutions for many organizations, while other companies can benefit from leveraging their existing mainframes as a more integral component to the enterprise data warehousing architecture. Imhoff suggests key mainframe benefits to consider, include: location of captured data; security requirements; availability; scalability; mixed workload requirements; near real-time BI operation; environmental initiatives, server consolidations, total cost of ownership.

Mainframes are designed to run at maximum capacity to derive benefits and value for every available cycle. Mainframes are optimized for mixed workloads and for automated systems management with facilities such as Workload Manager (WLM), as well as new hardware features—Integrated Facilities for Linux (IFLs), System z Integrated Information Processor, or zIIP (DB2 DRDA® execution) and zSeries® Application Assist Processor, or zAAP (Java support processors)—for offloading significant processing costs.

Facilitating useful, reliable information through sound data management

In any data warehousing and BI environment, companies must consider many important data-related issues. Some key considerations include:

- Extract, transform and load (ETL) requirements
- Intended BI use
- Periodicity of data capture, replication and synchronization
- Monitoring and managing the data warehouse efficiency
- Security, access, backup and recovery demands
- Scalability over time and ease of deployment

To meet those demands, IBM has created an integrated portfolio of ETL and BI products that incorporate the many data warehousing processes as a set of integrated processes.

The role of the mainframe in distributed data warehousing

Many companies have a mainframe infrastructure where the bulk of operational data is collected within applications, used in building the data warehouses and leveraged to run the business. A percentage of installations provide data warehouse and BI solutions on distributed platforms. The mainframe data is either accessed remotely or is replicated and transformed into a form that can be

optimized for BI analysis. There are numerous replication strategies supported by IBM. The customer's business process requirements will determine the optimal configuration for synchronization of data between two platforms.

Implementing a hybrid solution which features a mixed workload carried by both mainframe and distributed platforms, involves considerations and costs that specifically apply to the distributed platform. In some cases a hybrid makes perfect sense, such as where the mainframe does not support a particular BI process or a dedicated resource may prove necessary due to intense processing requirements and a need for isolation exists. Factors to consider:

- Additional hardware requirements: processors, disks and telecommunications gear
- ETL and other supporting software
- BI tools
- Administrative and personnel costs
- Environmental costs
- Coordination and synchronization requirements
- Performance monitoring

Increased data warehouse and BI usage in a hybrid environment will require the addition of servers and the associated costs. An investment in System z may be a good alternative with the growing availability today of sophisticated BI and data warehousing/BI solutions.

Operational BI: Leveraging the mainframe for expanded business insight

An emerging trend taking advantage of the capability of the mainframe platform is operational BI, also known as operational intelligence. Operational BI complements traditional BI systems by allowing more people to share more information. With operational intelligence strategies, BI results are delivered to non-technical staff as part of a packaged solution. In those scenarios, the data used by the operations staff is more detailed and granular than the data stored in the data warehouse. The data may not need to go through the ETL cycle, but just needs to be available. This is a good example of real-time or near real-time information access. As mentioned earlier, IBM is extending its portfolio to provide better access to IMS, for example, to support this emerging initiative.

Research suggests that an estimated 80 to 85 percent of enterprise end users are not serviced by their existing data warehousing and BI solutions. Operational BI solutions can alleviate this disparity and may deliver immediate value to an often overlooked user population (see Figure 5).

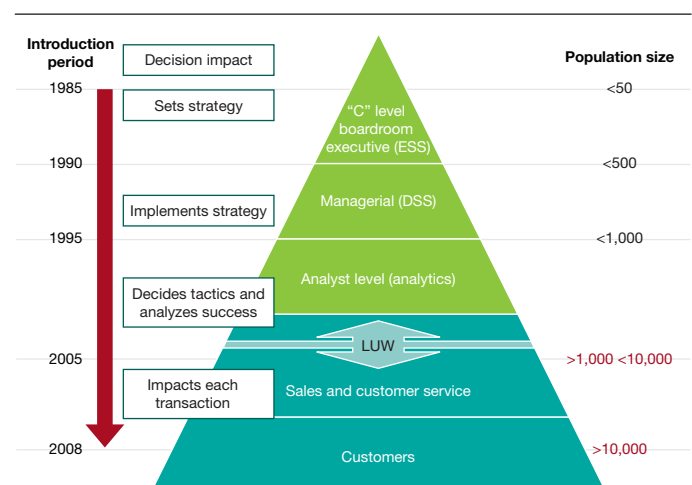


Figure 5: Operational Intelligence complements traditional BI systems and may deliver much-needed value to overlooked user populations.

Strategic BI solutions tend to be more internally focused and provide insights into the past, present and future of a business and its operations. Data at this level is operated on by more technically savvy individuals, such as analysts, or by top levels in the enterprise looking at the business through dashboards and other executive views.

Tactical or operational BI is targeted toward those individuals who interface with the customer and require immediate data access. Their value is more in their ability to directly interact with a paying customer. Thus direct access to operational data on the mainframe is extremely important. By doing so, employees are empowered to proactively make strategic, analytical and operational decisions to support corporate and departmental initiatives.

The mainframe plays an enormous role in operational BI scenarios. The IBM mainframe, including the powerful new capabilities offered by IBM System z10™, continues to be a reliable, scalable and secure enterprise platform.

IBM offers an end-to-end data warehousing and BI solution suite for System z

IBM now offers a series of innovative products and extensions to existing products that provide the rich ETL, systems management and BI tools necessary to create, maintain and grow a modern infrastructure for data warehousing and BI on IBM System z. These new and enhanced IBM solutions include:

- **IBM DB2 for z/OS VUE (Value Unit Edition)**—A new value point that facilitates the deployment of eligible application workloads on the System z platform that will provide the same robust DB2 for z/OS data server at a one-time charge price.
ibm.com/software/data/db2/zos/edition-vue.html
- **IBM InfoSphere™ Information Server for Linux on System z**—A data integration software platform that helps organizations profile, cleanse, extract, transform and load trusted information into a data warehouse to derive more value from the complex, heterogeneous information spread across their systems.
ibm.com/software/data/integration/info_server_system_z
- **IBM InfoSphere Warehouse for Linux on System z**—A data warehouse accelerator providing rapid building processes and enhanced query execution capabilities, plus new administrative functions optimized for performance and ease of use.
ibm.com/software/data/infosphere/warehouse-z
- **IBM Cognos 8 BI for Linux on System z**—A comprehensive System z offering for enterprise BI, it helps enable all user communities to receive relevant information to drive more informed, faster and more aligned business decisions.
ibm.com/software/data/cognos/products/cognos-8-business-intelligence

- **IBM DataQuant for z/OS**—A powerful business analytics tool delivering a host of BI capabilities, such as dashboards and data visualization.
ibm.com/software/data/db2imstools/db2tools/dataquant
- **IBM DB2 QMF™ (Query Management Facility)**—Transforms business data into a visual information platform for the entire enterprise with visual data on demand.
ibm.com/software/data/qmf
- **IBM InfoSphere Master Data Management Server for Linux on System z**—A strategic architecture to centrally manage customer, account and product master data for use across the enterprise.
ibm.com/software/data/infosphere/mdm_server
- **IBM System z10 Enterprise Class Processor**—A world-class server helping businesses create a new enterprise data center for data warehousing and BI.
ibm.com/systems/z/news/announcement/20080226_annnc.html

Conclusion

Today, data warehousing and BI are considered mission-critical functions for many organizations of all types and sizes. Companies are rethinking their BI strategies. Many companies are deciding to employ a more holistic platform approach so they can extend BI capabilities to the entire organization. Executives and operational teams alike can excel at making more informed, faster, more aligned business decisions.

When choosing where and how to implement a data warehousing architecture and BI applications, companies need to carefully weigh their options. Most importantly, companies must consider how they can meet the requirements of handling their current, diverse workloads, along with data warehouse and BI processes. They must focus on aligning the relative cost, complexity and capabilities of an advanced computing environment with their future application needs and vision for the company down the road.

Recent advances in mainframe architectures, along with development of the data warehousing and BI software that support them, are empowering organizations with a reliable, secure, scalable, cost-effective architecture for their new data warehousing and BI functionality. IBM System z supports the entire data warehousing cycle, along with BI software to deliver a total solution infrastructure. The disparity between the mainframe and distributed platform solutions is disappearing, opening the door to a wider range of data warehousing and BI options for today's forward-thinking companies. IBM continues to invest in System z with new offerings, solutions, pricing actions and acquisitions to continue its drive toward excellence as a top provider of data warehousing and BI solutions.



© Copyright IBM Corporation 2009

IBM Global Services
Route 100
Somers, NY 10589
U.S.A.

Produced in the United States of America
December 2009
All Rights Reserved

IBM, the IBM logo, ibm.com, Cognos, DB2, DRDA, IMS, InfoSphere, QMF, Smarter Planet, System z, System z10, zSeries, z/OS and z/VM are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries or both. If these and other IBM trademarked terms are marked on their first occurrence in this information with a trademark symbol (® or ™), these symbols indicate U.S. registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the Web at “Copyright and trademark information” at ibm.com/legal/copytrade.shtml

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

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

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

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

References in this publication to IBM products or services do not imply that IBM intends to make them available in all countries in which IBM operates. Offerings are subject to change, extension or withdrawal without notice.

All statements regarding IBM future direction or intent are subject to change or withdrawal without notice and represent goals and objectives only.

The information contained in this document is provided for informational purposes only. While efforts were made to verify the completeness and accuracy of the information contained in this document, it is provided “as is” without warranty of any kind, express or implied. In addition, this information is based on IBM’s current product plans and strategy, which are subject to change by IBM without notice. IBM shall not be responsible for any damages arising out of the use of, or otherwise related to, this document or any other documents. Nothing contained in this document is intended to, nor shall have the effect of, creating any warranties or representations from IBM Software.

¹ Data breach costs soar, Bill Brenner, Senior News Writer, SearchSecurity.com, 29 November 2007

² Robert Francis Group, 2005



Please Recycle