

Key Attributes and Considerations For Business Intelligence Applications in an IBM i Environment

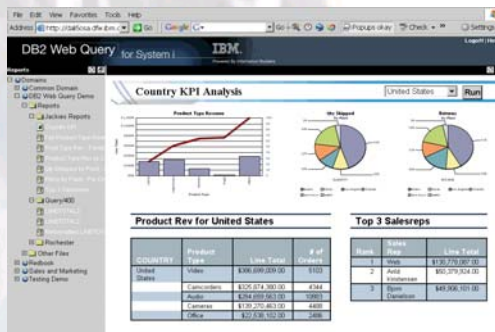
Business Intelligence is a broad term relating to applications designed to gather, consolidate, cleanse and analyze data for purposes of understanding and acting on the key metrics that drive profitability in an enterprise. Examples include:

- Business Performance Measurement reporting, including Balanced Scorecard reporting and standard financial reports.
- Budget and planning applications that simplify the process and insure integrity of data during each planning cycle.
- Analytical CRM (Customer Relationship Management) applies BI practices to the marketing function to optimize budgets, execute and track campaign strategies, and generate more profit from marketing activities.
- Sales Analysis to track sales vs. forecast, to build plans based on profit or packaging strategies, and nurture premier customers with better service through a more informative sales force.
- Even Human Resource departments leverage BI applications to monitor and analyze payroll classifications, training programs or labor utilization.

Armed with timely, intelligent information that is easily understood (because it is delivered in business terms) the Business Analyst is enabled to affect change and develop strategies to drive higher profits.

Key Benefits of IBM i for BI

- **DB2 Web Query for i makes analyzing DB2 data EASY**



IBM's DB2 Web Query for i provides capabilities to query or build reports against data stored in DB2 for i databases through browser based user interface technologies. Build new reports with ease through wizard like Report and Graph Assistant components. Simplify the management of reports by leveraging parameterized reporting to reduce the number of reports that require maintenance. Deliver data to end users in many different formats, including spreadsheets, PDF, HTML or powerful web visualizations. Automate the report distribution function through DB2 Web Query Report Broker. Enable BI as part of the day to

day process by integrating reports into line of business applications with the DB2 Web Query Software Development Kit.

DB2 Web Query allows you to import Query/400 definitions and enhance their look and functionality. Add OLAP-like slicing and dicing to the reports, or view reports in disconnected mode for the end user on the go, including mobile device users. Access DB2 through one of several native DB2 adapters.

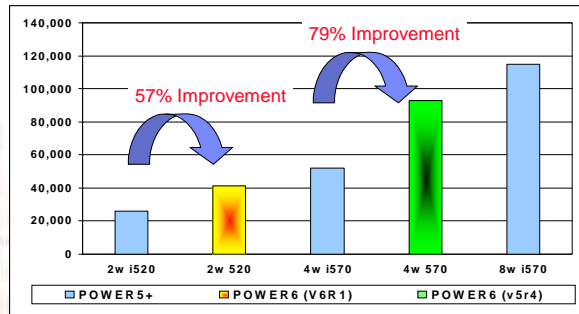
Build reports pulling data from multiple DB2 for i databases running remotely or in separate servers or LPARs (Logical Partitions). Leverage the meta data layer design to simplify report authoring, improve productivity, and provide consistent data definitions across the enterprise, eliminating ambiguity and distrust of the data in reports.

- **DB2 for i Query Processing OUTPERFORMS!**

IBM has been enhancing DB2 for i over the last decade, and continues to add new capabilities to enable the most efficient processing of the BI workload while still maintaining the low total cost of ownership value proposition.

The sophisticated cost based **query optimizer** built into DB2 for i is the basis of achieving optimal performance in a BI application. The optimizer provides the “brains” behind the SQL processing which most BI applications depend on. The optimizer’s goal is to build the best plan for accessing data. DB2 for i’s cost-based optimizer uses statistics stored automatically in the database as well as other information, such as system configuration, parallelism settings, and available indexes in its development of a data access plan.

A great proof point of the performance capabilities of DB2 for i in a Business Intelligence workload is the recent results of SAP certified benchmarks. The chart here shows the consistent performance of DB2 for i



*See detailed certified benchmark results at http://www.sap.com/solutions/benchmark/bid_results.htm

over time, running SAP’s BI Data Mart benchmark on IBM i. The most recent results include IBM’s new Power Systems Server based on POWER6 processor technology and DB2 for i 6.1. The POWER6 benchmark showed a 57% performance boost from a similarly configured Power5+ model! The metrics of the benchmark are “dialog steps per hour” which are similar to queries per hour. Another unique advantage of DB2 for i is IBM’s patented **Encoded Vector Indexing (EVI)** technology, built into DB2. EVIs use a

symbol table and vector array indexing technology that can make complex data warehouse queries scream! Many customers have reported that through EVI usage they have been able to reduce query jobs from hours to minutes, or in some cases, even to seconds. EVIs can be built right on top of your DB2 databases, and maintained real time – no need to move data to a specialized query accelerator database.

Additional enablers for processing complex query workloads include:

- Sophisticated processing of star schema or snow flake data models with Lookahead Predicate Generation, Query Re-Write, Dynamic Bit Mapped Indexing and EVIs

Star schema and snowflake data models are very common in data warehouse implementations. DB2 for i combines a number of optimizer techniques to add special processing of queries in this environment. Lookahead Predicate Generation (LPG) will analyze the FACT and DIMENSION tables specified in the query and minimize I/O by re-writing the query using only predicates that are going to be referenced in the FACT table. Once the query is rewritten, EVIs and Dynamic Bit Mapped Indexing take over to process the query.

- Auto recognition of summary data with Materialized Query Tables (MQTs)

MQTs are “optimizer aware” summary tables. For large relational data warehouses, building of summary tables often speeds up commonly run queries. DB2’s ability to **re-write queries** allows the database engine to automatically redirect an end user query to use the MQT for faster processing.

- Near Real-time Data Warehousing with Remote Journaling

“Near” real-time data warehousing can be accomplished through the use of a built in DB2 feature called Remote Journaling. Remote Journaling allows capturing of changed database records real time into a log (Journal Receiver). Through IBM i directives, the log can be shipped to a 2nd system or partition running IBM i, and the change can be applied to the data warehouse based on timings that meet the requirements of the application.

- Autonomic Functions

Autonomic is a term associated with a self learning, self adapting organism. With DB2 6.1, the query optimizer can learn about performance characteristics while executing a query, and on the fly adapt the query execution plan based on what it learns. Autonomic Indexes were first delivered

in IBM i 5.4, and allows DB2 to create indexes where necessary to improve query performance while preserving those indexes for use by other jobs and users.

• DB2 for i Delivers Low Total Cost of Ownership

IBM i's unique architecture, including Single Level Storage and a tightly integrated object based operating system, provide the foundation for its leadership in autonomic computing. DB2 for i leverages those architectural attributes to reduce the tasks required by a database administrator (DBA) in comparison to other relational database management systems. For example, with DB2 for i:

- Data Partitioning for performance/scalability: **NOT REQUIRED**
- Moving data or indexes to avoid disk hot-spots: **NOT REQUIRED**
- Re-balancing Indexes: **NOT REQUIRED**
- Monitoring table spaces, log buffers, lock contention buffers: **NOT REQUIRED**
- Running integrity checkers or statistics collection routines: **NOT REQUIRED**

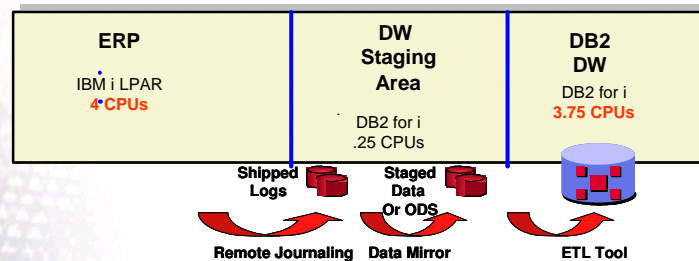
“The tools alone made the move from SQLServer to DB2 worth it. While SQLServer provides easy to use tools, the depth to really analyze what is happening with SQL based programs was very limited compared to DB2.” - Steve Hester, Senior Development Manager, Merillat Corp.

The **On Demand Performance Center**, built into DB2 for i, provides capabilities to understand how DB2 is processing the BI workload. Real time analysis of query jobs through the SQL Plan Cache, or post execution analysis of SQL statements against collected database monitor data provide IT with graphical tools to understand how DB2 is processing SQL. Visual Explain, Index Advice, and additional analysis features make it very easy to tune the workload to improve performance and ensure your system is running optimally.

• Virtualization attributes of PowerVM™ and IBM i provide FLEXIBILITY

PowerVM allows you to dynamically allocate resources (e.g., CPU or memory) across multiple copies of IBM i (or Linux or AIX operating systems) within a single system footprint. The “on demand” features of PowerVM also provide automatic resource allocation based on policies and priorities that you define.

PowerVM technology can be applied to a Business Intelligence environment to reduce the Extract/Transformation/Load (ETL) and summary table/cube calculation times often used in a data warehouse architecture. These workloads, typically running at night, can steal processor and memory from operational application partitions that are going into batch or reduced workload modes. Once the ETL work is completed, those resources can be automatically moved back into the production partition for normal 8 to 5 workloads.



The above picture depicts how you could use PowerVM in a BI environment.

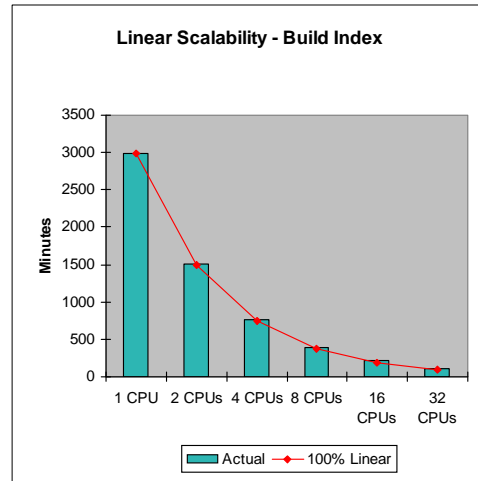
Capacity Upgrade on Demand (CoD) is an IBM feature that allows you to quickly add additional processing power to your system, on a permanent or temporary basis. One of the challenges with BI applications is that usage of the application can grow like a snowball. Being able to react quickly to additional demands placed on the database server can be the difference between the success and failure of the application. With CoD, you have almost instantaneous ability to turn on additional processors or acceleration features to meet these demands.

- **Scalability of DB2**

A key consideration in any database platform in support of BI applications is its ability to scale while maintaining acceptable performance. DB2 for i offers a variety of techniques to address scalability/performance.

Parallel database operations through the DB2 Symmetric Multiprocessing (SMP) feature of IBM i provide the ability to split single database tasks/requests across multiple processors. The SMP feature provides significant performance enhancements to database queries, index builds and other database tasks invoked from the BI application.

The Parallel I/O design of IBM i uses intelligent I/O Processors with their own levels of caching and parallelism. Combined with Single Level Store, the I/O subsystem of the IBM i removes a common bottleneck with data warehouse workloads: the I/O subsystem.



The graph here represents an actual test of linear scalability. This particular database task is an index build, but it shows that each time the job was run with the number of CPUs doubled, the performance indeed nearly doubled, reflecting a “nearly 100%” linear scalability.

“We built identical models over identical sets of data. There was no comparison in performance. In fact, the bigger the tables, the bigger the advantage of DB2 over SQL Server. That proved to us that DB2 for i was a better choice.” – Sam Gottlieb, Project Lead, Elie Tahari

- **Combining High Availability or Disaster Recovery with Business Intelligence**

Customers are often looking to enhance the availability of the IBM i based applications by leveraging a 2nd server and appropriate mirroring software to create a 24 x 7 or other disaster recovery scenario (in addition to being able to maintain availability during PLANNED downtime).

In addition, one of the challenges of the data warehouse architecture is how you TRANSPORT the data from source system to target while not impacting the production systems.

As with Remote Journaling (mentioned above), some availability software solutions for can also be leveraged to create an architecture that can provide near real time TRANSPORTATION of data from source to target systems, with minimal impact to the source side database server. An Operational Data Store (ODS), commonly used for real time operational reporting, is a BONUS of this architecture. The ODS can then also become the source for feeds into the data warehouse.

The benefits include:

- Data Transport becomes a non issue
- Leverage an otherwise under utilized 2nd server for the data warehouse workload
- Achieve a Disaster Recovery (DR) or High Availability (HA) solution for the data warehouse (as well as the operational applications)
- SAVE MONEY by combining the acquisition of a server for both HA and/or DR and Business Intelligence

- **Leverage tools or application tiers running in Linux, AIX or Windows environments**

Power Systems Servers also support Linux and AIX operating systems. For example, multiple IBM i, Linux, or AIX partitions can be defined to a single Power Systems server, reducing complexity, leveraging dynamic allocation of resources and other benefits such as virtual I/O, high speed interconnects between the partitions, and disk virtualization.

How does this relate to support of Business Intelligence applications? Suppose you are using an ETL tool whose TRANSFORMATION component does not support IBM i but can run in Linux or AIX. Through support of multiple operating systems within a single server, the ETL code logic remains within the framework of the server along with the data warehouse and operational database. This implies that no data has to be moved across a network, and high speed bus communications can be leveraged to optimize the entire ETL process. All data – source databases, target databases, and any intermediate data storage areas, all remain within the framework of the Power Systems server.

Similarly, the use of IBM BladeCenter® solutions can provide similar benefits to software designed for Intel processors by its ability to leverage the virtual I/O and high speed connections to the server.

- **Reduce Risk**

Business Intelligence Applications can be a very powerful enabler for any company looking to gain insights into their data. Risk in any BI project is significantly reduced if you can leverage attributes common with current operational systems. These attributes include sharing of common hardware and software components, leveraging existing operational and administrative skills and policies, and being able to optimize investments across both operational and business intelligence applications through sharing resources and skills.

IBM i + Power Systems servers + DB2 for i provide an excellent platform for minimizing risk and costs associated with implementation of this powerful technology. For additional information, please refer to:

DB2 for i Home Page: <http://www.ibm.com/systems/i/db2>