DB2 UDB for iSeries Update
The next generation iSeries...
...simplicity in an on demand world

January, May 2003
DB2 UDB for iSeries May 2003 Update - Agenda

- Overview of OS/400 V5R2 SQL support
- Update on SQL query performance - the V5R2 SQL Query Engine
- ORACLE Migration Toolkit
- OS/400 Database and Autonomic Computing
- Information References

For coverage of the extensive V5R2 DB2 UDB for iSeries enhancements available prior to May 2003, refer to the Database presentation within the V5R2 Technical Overview presentation set dated September/October 2002:

- http://www.ibm.com/eserver/iseries/support
- Select Technical Databases -> iSeries Technical Overviews
- Select V5R2 -> DB2 for iSeries

See also:
- IBM Server Sales and PartnerInfo websites for Servers -> iSeries for complete, detailed coverage of iSeries database topics.
- January V5R2 Performance Capabilities Reference manual
Notes: Agenda

This foil lists the topics covered in this presentation.

See the Additional Resources foil for a complete list of DB2 UDB for iSeries database capabilities.
OS/400 Database History Lesson

IBM Delivers Robust SQL Function

1970
✓ Codd Invents RDBMS

1973
✓ System R Project

1980
✓ S/38 Delivered with RDBMS Based on System R

1988
✓ IBM Defines COMMON SQL Language

1988
✓ AS/400 Delivered with SQL/400

1995/96
✓ SQL Optimization
✓ Referential Integrity
✓ DRDA DUW
✓ Net.Data
✓ Triggers
✓ Long Names
✓ Stored Procedures
✓ SMP Parallelism

1995/98
✓ Cluster parallelism
✓ Data mining
✓ Parallel Index Build
✓ UCS-2
✓ BI Solutions
✓ Dynamic Bitmaps
✓ Encoded Vector Indexes
✓ Parallel Data Loader

1997/98
✓ Large Objects
✓ SQLJ
✓ Data Links
✓ UDTs, UDFs
✓ Op Nav Performance Monitor
✓ DB2 OLAP
✓ Visual Explain
✓ Java (tm) Stored Procedures

1999/2000
✓ Large Objects
✓ SQLTriggers
✓ Database Navigator
✓ Generate SQL
✓ 2G LOBs
✓ 1TB tables
✓ ODBC 3.5
✓ DB2 Extenders

2001
✓ Self Optimizing Query Engine
✓ Migration Toolkits
✓ Multiple Name Spaces
✓ Switchable Disk Clustering
✓ Index Advisor
✓ New Statistics Manager
✓ Adaptive e-Business Transaction Services
✓ Enhanced DB2 Family Compatibility
✓ Add’l SQL Standards Support

1970:
RS/400

1973:
DB/400

1980:
OS/400

1988:
AS/400

1995:
RDBMS is named DB2/400

1997:
Renamed to DB2 UDB

2002/3:
Self Optimizing Query Engine
Migration Toolkits
Multiple Name Spaces
Switchable Disk Clustering
Index Advisor
New Statistics Manager
Adaptive e-Business Transaction Services
Enhanced DB2 Family Compatibility
Add’l SQL Standards Support
Notes: OS/400 Database History Lesson

This chart summarizes the evolution of database and SQL functions to that database on the iSeries. Note that the iSeries has had since DAY ONE a very robust relational database management system.

SQL functionality first started showing up at V3R1 in 1995, when database parallelism, stored procedures, and triggers were first introduced. At V3R1, IBM gave the database a name: DB2!!

Since that time, a significant amount of function has been added (and will continue to be added) to the adherence to the ANSI SQL Standards in each new OS/400 release. At V4R2, the database name was changed again slightly - the Universal Database (UDB) acronym was added - to reflect the database's ability to extend into capabilities beyond typical relational data elements and to remain consistent with the other DB2 family members.

Some things to consider:
- You or your customer may not know whether they need the DB2 Symmetric Multiprocessing feature of OS/400 to get parallel database tasks
- DB2 UDB for iSeries was the first database running on any operating system to introduce SQL Stored Procedures, based on the industry standard
- Encoded Vector Indexes were introduced in V4R2 and have delivered significant performance improvements over several releases
- See the April 2002-October 2002 Database presentation for a full description of the V5R2 functional enhancements in DB2 UDB for iSeries
**Additional background information**

While all DB2 UDB family members share research and development across various IBM laboratories, the delivery of DB2 UDB may differ. There are actually three core DB2 UDB family code bases - one for iSeries, zSeries and distributed platforms.

However, it is important to note that much of the code and technology is shared across all DB2 UDBs.

DB2 UDB for iSeries has the unique advantage of being integrated into the iSeries Operating System (OS), maintaining the iSeries' value proposition - low cost of ownership through INTEGRATION.

The commonality is in these areas:

- **Shared SQL Standards**: and this website contains a paper on details (http://www.ibm.com/servers/eserver/iseries/db2/common.html)
- **Shared DB Utilities**: and if you CLICK on the target it'll take you to an optional slide to provide an example. Also, the website links you to a document describing these utilities (http://www.ibm.com/servers/eserver/iseries/db2/util.htm)
- **Shared research**: For instance, Encoded Vector Indexes (EVIs), a V4R2 feature of DB2 UDB for iSeries, was actually developed out of our Silicon Valley Lab

Where the products differ:

- Each DB2 UDB is optimized for the platform. For instance, on iSeries, much of the database code is built into OS and microcode levels which is very efficient.
- **Packaging**: How the function is delivered may be slightly different in the DB2 UDBs. If you CLICK on the target, it'll take you to a slide that shows an example of this.
- **DBA Facilities**: Because DB2 UDB for iSeries is built into the OS, many of the DB2 management facilities are also built into the OS, and delivered through native operational menus or via Operations Navigator.
- Some functional differences due to development schedules/resources and customer requirements.
SQL Access: The DB2 Optimizer

- **Native Access**
  - Programmer Controlled
  - Read a Record, do something, write a record
  - Select next record

- **SQL Access**
  - Build/Update Access Plan
    - Determine query complexity
    - Review database statistics
    - Build/Update Plan to get data
      - physical table scan?
      - use/create an index?
      - build a temporary file?

**Only the Optimizer knows!!**
Notes: SQL Access: The DB2 Optimizer

A review of the "optimizer responsibilities" for finding the rows and columns that match the specified search criteria with operating system-specific efficiency is illustrated on this foil.

To the left of the Optimizer we show the OS/400 "native file access" (for example READ, WRITE) and on the right a representation of the SQL access (for example SELECT, INSERT).

Since the industry standard interface to database is through SQL (it has much more powerful facilities than the OS/400 native database interfaces), it is imperative to constantly improve the efficiency of the "optimizer."
Self-Optimizing SQL Query Availability

- New State of the Art SQL Query Engine (SQE)
  - In most cases provides 2 times or more performance improvement for read only complex queries
  - Phased Approach providing
    - Improved SQL Performance
    - Minimizes system resource utilization during SQL Processing
  - Delivery Method - see Informational APAR II13486
    - DB2 PTFs planned to be available May 2003
      - DB2 UDB for iSeries group PTF: SF99502, level 8
Notes: Self-Optimizing SQL Query Availability

The query optimizer of previous releases is being improved in V5R2.

Note that while some improvements were included in the original V5R2 general availability software, additional performance improvements are planned to be made available through software fixes (PTFs) planned to be available May 23, 2003. The new optimizer is planned to be fully enabled as part of the May 2003 DB2 UDB for iSeries group PTF SF99502, level 8.

In V5R2, DB2 UDB for iSeries redesigned the query engine, which may provide performance improvement for SQL read-only queries. When the new DB2 PTFs become available, there will be significant information available providing details on the performance improvements, which types of queries can take advantage of the redesign, and how to aid the optimizer in taking advantage of the new improvements.

These sources of information include:
- Info APAR II13486 describes how to order the improvements
- [www.ibm.com/eserver/iseries/db2/sqe.html](http://www.ibm.com/eserver/iseries/db2/sqe.html) on preparing for the new optimizer
- Red Draft (formerly Redpiece) SG24-6598, to be updated 1H 2003

**Important:** The query optimizer of previous releases handles queries from a number of different interfaces. The V5R2 SQE query optimizer only handles queries from SQL interfaces. For example, queries from OPNQRYF and Query/400 are not supported. Ensure you read the information referenced in this presentation when the support becomes available.
V5R2 Performance Summary with May 2003 PTFs

- Performance improvement for read-only queries involving complex joins and selection logic, such as Business Intelligence and CRM analytics sets of data
  - May provide up to 5 times performance improvement: most workloads tested averaged a 2 X improvement

- Some simple queries showed a slight degradation

- Individual results may vary

- See the following sources of performance information:
  - Info APAR II13486 describes how to order and install the improvements
  - www.iseries.ibm.com/db2/sqe.html on preparing for the new optimizer
  - Red Draft (formerly Redpiece) SG24-6598, to be updated 1H 2003
Notes: V5R2 Performance Summary with May 2003 PTFs

This is a summary of expectations when the new Query Optimizer set of PTFs becomes available which is planned for the first half of 2003. Many queries will improve, but some will not. A few may actually degrade.

You need to consult the available documentation listed on this chart when the PTFs become available to ensure you have appropriate expectations and instructions on how to use the new support.

The next foil shows a set of Rochester laboratory test results of a specific set of "long running" queries through May 2003.

The foil following the long running query test results shows a set of results for "quick response" queries.
119 queries exercising a wide range of function were selected for study. In this set of test cases:

- Most queries demonstrated significant response time improvement, but 2 queries showed degradation.
- The average response time improvement on a per query basis was 2.6 times across the query group.
- The most improved query showed a response time improvement of 18 times.
- The most degraded query showed a degradation of 2 times.

All equal or degraded queries are being examined.

Consult the available performance documentation before installing PTFs when they become available.
Notes: Long Running Query Response Time Comparison

This chart provides a view of the new query engine/optimizer performance by comparing response times between the new engine/optimizer and the current query engine/optimizer - for long running queries.

119 longer running (up to 600 seconds) Business Intelligence (BI) type queries were measured and a comparison of response time between the new engine/optimizer and the old engine/optimizer is graphed in the chart.

- When "Times Improvement" is positive, the new query engine/optimizer outperformed the current support by the amount shown
- When "Times Improvement" is negative, the old engine and optimizer perform better than the new query engine /optimizer by the amount shown

Here are important observations to set proper perspectives for the new support coming May 2003. The data shows across the range of 119 queries:

- The new query engine/optimizer in general outperforms the existing engine/optimizer with all but two of the queries performing better
- On a per query basis, the new engine/optimizer response times were on average 2.6 times faster than the current

The query comparisons were made in the following environment. Results in different environments may vary:

- These queries were run in a single user environment where there was only one job active
- Individual queries were run sequentially, one query at a time
- The queries were run in a 2 Gigabyte memory pool with Query Degree set to *Optimize
- There were two CPUs present in the environment
- All of the data needed by the queries was on disk (as is typical in most BI Environments)

As stated in earlier foils, refer to the following sources for more up to date information when the new support becomes available:

- Info APAR II13486 describes how to order the improvements
- www.ibm.com/eserver/iseries/db2/sqe.html on preparing for the new optimizer
- Red Draft (formerly Redpiece) SG24-6598, to be updated 1H 2003
118 queries exercising a wide range of function were selected for study. In this set of test cases:
- 91 queries showed significant response time improvement.
- 19 queries showed approximately equal performance
- 8 queries were degraded
- Of the 118 queries studied, the average improvement was 2.5 times
- The most improved query showed a response time improvement of 20 times
- The most degraded query showed a response time degradation of a 1/2 time (or 50%)

All equal or degraded queries are being examined

Consult the available performance documentation before installing PTFs when they become available
The chart provides a view of the new query engine/optimizer performance by comparing response times between the new engine/optimizer and the current query engine/optimizer for short running queries.

118 short running queries were measured and a comparison of response times between the new engine/optimizer and the old engine/optimizer is graphed in the chart.

- When "Times Improvement" is positive, the new query engine/optimizer outperformed the current support by the amount shown.
- When "Times Improvement" is negative, the old engine/optimizer performed better than the new query engine/optimizer by the amount shown.

These queries are similar (in function) to simple short running transactional queries. However, these queries were run in a non-transactional environment where none of the data needed by these queries existed in memory. (In memory access is often the case in transactional environments.)

Here are important observations to set proper perspectives for the new support coming May 2003. The data shows that:

- The new query engine/optimizer, in general, outperforms the existing engine/optimizer across the range of 118 queries.
- Although some queries performed worse with the new engine/optimizer, the new engine/optimizer was on average better.
- On a per query basis, the new engine/optimizer response times were on average 2.5 times faster than the current engine/optimizer.

The query comparisons were made in the following environment. Results in different environments may vary:

- These queries were run in a single user environment where there was only one job active.
- Individual queries were run sequentially, one query at a time.
- The queries were run in a 2 Gigabyte memory pool with Query Degree set to *Optimize.
- There were two CPUs present in the environment.

As stated in earlier foils, refer to the following sources for more up to date information when the new support becomes available:

- Info APAR II13486 describes how to order the improvements.
- Red Draft (formerly Redpiece) SG24-6598, to be updated 1H 2003.
Performance Considerations

- When performing SQL performance analysis or sizing:
  - Sizing is not useful until performance analysis techniques have been applied
  - Use of iSeries SQL-based performance analysis tools is often required
  - Application type examples:
    - Consider how your application uses JDBC, ODBC, SQL Call Level Interfaces, or other SQL-based interfaces
    - Consider how you have defined your database interfaces using SQL in a WebSphere Application Server instance

- iSeries query-based analysis tools include:
  - Query Optimizer information available through job debug, STRDBMON command
  - iSeries Navigator SQL Performance Monitors and Visual Explain

- SQL Optimization skills are not prevalent among iSeries application developers
  - Get iSeries-based SQL optimization skills
  - An excellent source is: http://www-1.ibm.com/servers/eserver/iseries/db2/db2educ_m.htm
    - Excellent set of presentations, labs, and web-based education resources
    - Includes:
      - DB2 UDB for iSeries: Coding for SQL Performance
      - DB2 UDB for iSeries: SQL Performance Basics
      - Piloting DB2 UDB for iSeries with iSeries Navigator on V5R2 or V5R1
      - Performance Tuning DB2 UDB for iSeries with OS/400 job debug mode messages or with iSeries Navigator and Visual Explain graphics and messages
Notes: Performance Considerations

This foil lists some SQL-based application performance analysis and sizing considerations and performance analysis tools. The considerations listed lead to the subjects of iSeries SQL and query based performance analysis tools and iSeries specific SQL performance tuning techniques and education resources.

Further discussion on these tips and tools are beyond the scope of this presentation.

However, you are given a URL to an iSeries web page that has several SQL performance oriented presentations, classes, and workshops.

Note: There may be other application implementation inefficiencies than just SQL-based interfaces. For example, a database design that holds locks on rows of data for a period of time, causing other jobs/threads to wait, is independent of any interface to the database.

See the Additional References foil for a summarized list of iSeries Database and SQL information resources.
Oracle Migration
DB2 UDB Migration Toolkit for Oracle

- No Charge Download

- Automate Migration from Oracle to DB2 UDB for iSeries
  - Replicate Database Schemas
  - Move/Convert Data
  - Convert Stored Procedures, Triggers, Functions
  - Convert SQL Calls

- What this will NOT convert
  - Application Code in High Level Languages (C, C++)
  - Oracle Forms (but utilities exist for this)
  - OLAP functionality
  - Engage e-Custom Technology Center
DB2 UDB for Oracle Migration Toolkit web page

- Downloadable toolkit for Oracle to DB2 UDB for iSeries:

**DB2 Universal Database**
for iSeries

**DB2 UDB for Oracle Migration Toolkit**

This toolkit is designed to automate many of the steps involved in the migration of a database from Oracle to DB2 UDB. This no-charge download automates many of the tasks of migration Oracle objects to DB2 UDB for iSeries, including tables, views, indexes, triggers, and stored procedures.

- Oracle to DB2 Migration Toolkit Beta (Windows version - 28 MB)
- Toolkit Experience Report
- Migration Toolkit Presentation (PDF 2.5 MB)
- DB2 UDB for iSeries porting information
- Support & service information
  - General questions about the tool, or problems with the tool can be reported to the following e-mail address: db2mtki@us.ibm.com
  - DB2 UDB for iSeries migration and porting services are available from the eServer Custom Technology Center
DB2 UDB for Oracle Migration Toolkit - Steps

- Steps:
  - Manage migration projects
  - Specify the source
  - Convert source metadata
  - Refine the metadata conversion
  - Generate data transfer scripts
  - Deploy to DB2 for iSeries
Notes: DB2 UDB for Oracle Migration Toolkit - Steps

**Manage migration projects**
The Oracle Migration Toolkit (MTK) lets you organize your work into projects. Each migration project is kept separate, so you can easily manage all the files that are generated during each migration. Once you have created a project, you can come back at any time and open a previously created project. When you start the MTK, you can choose to create a new project or open a project that you have created previously.

**Specify the source**
It is important to keep in mind that the data itself is not being processed during this stage. You are simply identifying the Oracle objects whose definitions will be translated into a form so that the same objects can be created into DB2 on your iSeries system. The movement of the data that is contained in these objects is part of a later step. This step creates an empty database on DB2 that closely matches your Oracle database.

**Convert source metadata**
In the previous steps, you created the migration project and . This step uses the defined set of objects in your Oracle database that will be migrated to DB2 (R) Universal Database (TM) for iSeries.

**Refine the metadata conversion**
The Refine step gives you the opportunity to view the results of the conversion and to make changes. For example, you can change source stored procedures and some DB2 object names. However, you must return to the Convert step to apply these changes.

**Generate data transfer scripts**
Now you are ready to generate the scripts that move the data from your Oracle tables into your DB2 tables on the iSeries. The scripts that we generated previously on the Convert step were only concerned with the creation of the database objects. This step produces the scripts that import the data from Oracle and populate your DB2 files. If you are interested only in deploying the script to create the objects, and do not want to move any data from the Oracle database into DB2, you can skip this step.

**Deploy to DB2 UDB for iSeries**
In the final step in the migration process, you have created all the scripts necessary to create the DB2 objects, extract the data from the Oracle tables, and import the data into the DB2 tables.
DB2 UDB for iSeries: Autonomic Computing summary

- DB2 UDB for iSeries is a NATIVE, Object Based file system TIGHTLY Integrated with OS/400
  - Not built on top of a Unix/PC File System like Oracle
  - Single Level Store and Multiple I/O Processor Design is a HUGE Advantage for DB2 UDB for iSeries

- Self Configuring
  - No Installation required
  - No Tablespace, data partitioning, management of buffers
    - No Data Hot Spots

- Self Healing
  - Systems Managed Access Path Protection
  - Self Managed Database Logging
  - Auto Index Rebalancing
  - Database Locking Issues (Oracle FUD)

- Self Tuning
  - Auto Tuner, Sophisticated Cost Based Query Optimizer
  - Automatic Index Advisor

- Self Protecting
  - Object Based Kernel in OS/400

- See: "DB2 UDB for iSeries: Autonomic Computing FAQs"
  - System Sales ->
While all DB2 UDB family members share research and development across various IBM laboratories, the delivery of DB2 UDB may differ. There are actually three core DB2 UDB family code bases - one for iSeries, zSeries and distributed platforms.

However, it is important to note that much of the code and technology is shared across all DB2 UDBs.

DB2 UDB for iSeries has the unique advantage of being integrated into the iSeries Operating System (OS), maintaining the iSeries' value proposition - low cost of ownership through integration.

This chart shows at a high level where the DB2 UDB family members have commonality, and key areas where there are differences.

The commonality is in these areas:

- Shared SQL Standards - and this website contains a paper on details (http://www.ibm.com/servers/eserver/iseries/db2/common.html)
- Shared DB Utilities - and if you CLICK on the target it'll take you to an optional slide to provide an example. Also, the website links you to a document describing these utilities (http://www.ibm.com/servers/eserver/iseries/db2/util.htm)
- Shared research. For instance, Encoded Vector Indexes (EVIs), a V4R2 feature of DB2 UDB for iSeries, was actually developed out of our Silicon Valley Lab

Areas where the products differ

- Each DB2 UDB is optimized for the platform. For instance, on iSeries, much of the database code is built into OS and microcode levels which is very efficient.
- Packaging. How the function is delivered may be slightly different in the DB2 UDBs. If you CLICK on the target, it'll take you to a slide that shows an example of this.
- DBA Facilities. Because DB2 UDB for iSeries is built into the OS, many of the DB2 management facilities are also built into the OS, and delivered through native operational menus or via Operations Navigator.
- Some functional differences due to development schedules/resources and customer requirements.

NOTE that DB2 UDB for iSeries was the first DB2 family member to introduce industry standard SQL Stored Procedures, EVIs, as well as other elements such as two-phase commit over DRDA. There are also features that may be in DB2 UDB for NT/Unix that are NOT currently in DB2 UDB for iSeries. See the iSeries Database web page in Server Sales example shown at the end of this presentation, which shows how to find information on this in the Frequently Asked Questions topic.
Additional References

- DB2 UDB for iSeries CD
  - Briefing on a CD
  - Strategy video clips
  - Education
  - White Papers
    - http://www-1.ibm.com/servers/eserver/iseries/db2/db2educ_m.htm

- Indexing Strategies White Paper

- SQE Red Draft
  - *SQL Query Enhancements on DB2 Universal Database for iSeries*, SG24-6598 (to be available 1H 2003)

- DB2 Home Page

- iSeries Performance Management Home Page (Performance Capabilities Reference and more)
IBM Systems Sales

Systems Sales site search

Search for [iSeries and Database] in [iSeries] [GO]

Narrow your search by picking a subject from one of the following categories:

- [By Industry:]
- [By Solution:]
- [By Product:]
- [By Document Type:]

95 hits sorted by relevance

- [DB2 UDB for iSeries]
  This 40-page presentation provides a strategic overview of DB2 UDB on iSeries. (Presentation) 9
  Sep 2002
  [--> list categories]

- [IBM DB2 Universal Database for iSeries FAQ]
  This three-page document covers 10 of the most frequently asked questions about IBM DB2
  Universal Database (UDB) for iSeries. (Q&A) 10 Apr 2002
  [--> list categories]

- [IBM DB2 Universal Database for iSeries - Business Intelligence concepts]
  This 13-page document offers an overview of BI concepts as they pertain to IBM DB2 Universal
  Database for iSeries. (White paper) 20 Mar 2002
  [--> list categories]

© 2000-2003 IBM Corporation
improvements equivalent to the ratios stated here. Photographs shown are of engineering prototypes. Changes may be incorporated in production models.

Therefore, no assurance can be given that an individual user will achieve throughput or performance equivalent to the ratios stated here. Photographs shown are of engineering prototypes. Changes may be incorporated in production models.

Performance is based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput or performance that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve throughput or performance equivalent to the ratios stated here.