IBM Software Group

Putting IBM databases on Rails

IBM Toronto Laboratory

Leon Katsnelson leon@ca.ibm.com

Agenda

- Why use DB2 in your Ruby on Rails projects
- How to get DB2 on Rails

What are “IBM databases”?

A Portfolio of Data Servers

Superior Capabilities Across the Spectrum of Needs

DB2 Everyplace	Cloudscape	U2	IDS	IMS	DB2
OLTP	OLTP	OLTP	OLTP	OLTP	OLTP & Analysis
Relational	Relational	Multi-Value	Relational	Hierarchical	Relational & XML
Mobile Embedded	Intra-App / Single-App	Intra-App / Single-App	Intra-App / Single-App	Single / Multi-App	Single / Multi-App
Linux PalmOS Symbian…	Java	AIX, etc. Linux Windows	AIX, etc. Linux Windows	z/OS	z/OS, I5/OS AIX, Linux, Windows…

DB2 Everyplace

- DB2 database for mobile and embedded platforms
  - Very small: ~200KB on the device
  - Rich relational capabilities - SQL92/99 compliant subset
  - Per table data encryption
  - Extensive device support
  - Easy Application: .Net, C/C++, VB, Java
- Synchronization Server:
  - Multi-platform: Windows, Linux, AIX, Solaris
  - Multi-data source: DB2 Family, MS SQL Server, Oracle, Informix, Sybase, Domino DB

Broad Platform Support

- Pocket PC 2000/2002, Windows Mobile 2003, 2005
- Windows CE 2.11/3.0, WinCE .Net 4.2
- Windows 95/98/ME/NT/2K/XP/Tablet/Media/Embedded
- Palm OS 3.5+, Palm OS 5.0+
- Symbian V6+
- Linux
- QNX Neutrino
- Sync client for J2ME/RIM
- DB2 Everyplace Sync Server Platforms
  - Windows NT/2000/XP/2003
  - AIX
  - Linux and Solaris

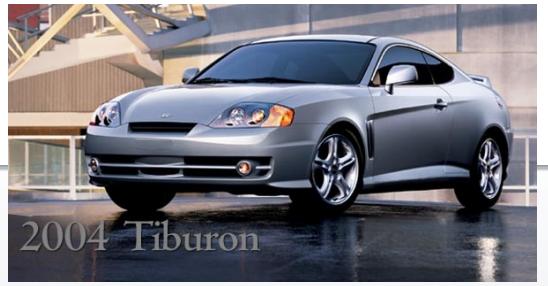
- Replication data sources
  - DB2 UDB for Windows, Linux, UNIX, OS/390 & zSeries, and AS/400 & iSeries
  - Oracle, Microsoft SQL Server, Informix, Sybase, Other JDBC-based sources, and Domino Databases

Broad Platform Support

- Pocket PC 2000/2002, Windows Mobile 2003, 2005
- Windows CE 2.11/3.0, WinCE .Net 4.2
- Windows 95/98/ME/NT/2K/XP/Tablet/Media/Embedded
- Palm OS 3.5+, Palm OS 5.0+
- Symbian V6+
- Linux
- QNX Neutrino
- Sync client for J2ME/RIM
- DB2 Everyplace Sync Server Platforms
  - Windows NT/2000/XP/2003
  - AIX
  - Linux and Solaris

- Replication data sources
  - DB2 UDB for Windows, Linux, UNIX, OS/390 & zSeries, and AS/400 & iSeries
  - Oracle, Microsoft SQL Server, Informix, Sybase, Other JDBC-based sources, and Domino Databases

Hyundai Motor Company



- **Challenge:**
  - HMC wanted to implement built-in wireless devices in its high end automobiles. To achieve this goal, it needed to build telematic devices along with a mobile database and synchronization solution.
- **Solution:**
  - DB2 Everyplace combined with WebSphere Everyplace Connection Manager, enables seamless roaming and secure connections.
  - DB2 Everyplace synchronizes three tables of legacy Oracle source data and stores it in the built-in mobile devices. It also manages address books, destination data and appointment data and was an ideal choice for the customer because it supports the QNX device client.
- **Value:**
  - The customer selected IBM because DB2 Everyplace was easy to use and develop, had strong performance capabilities and was highly scalable.

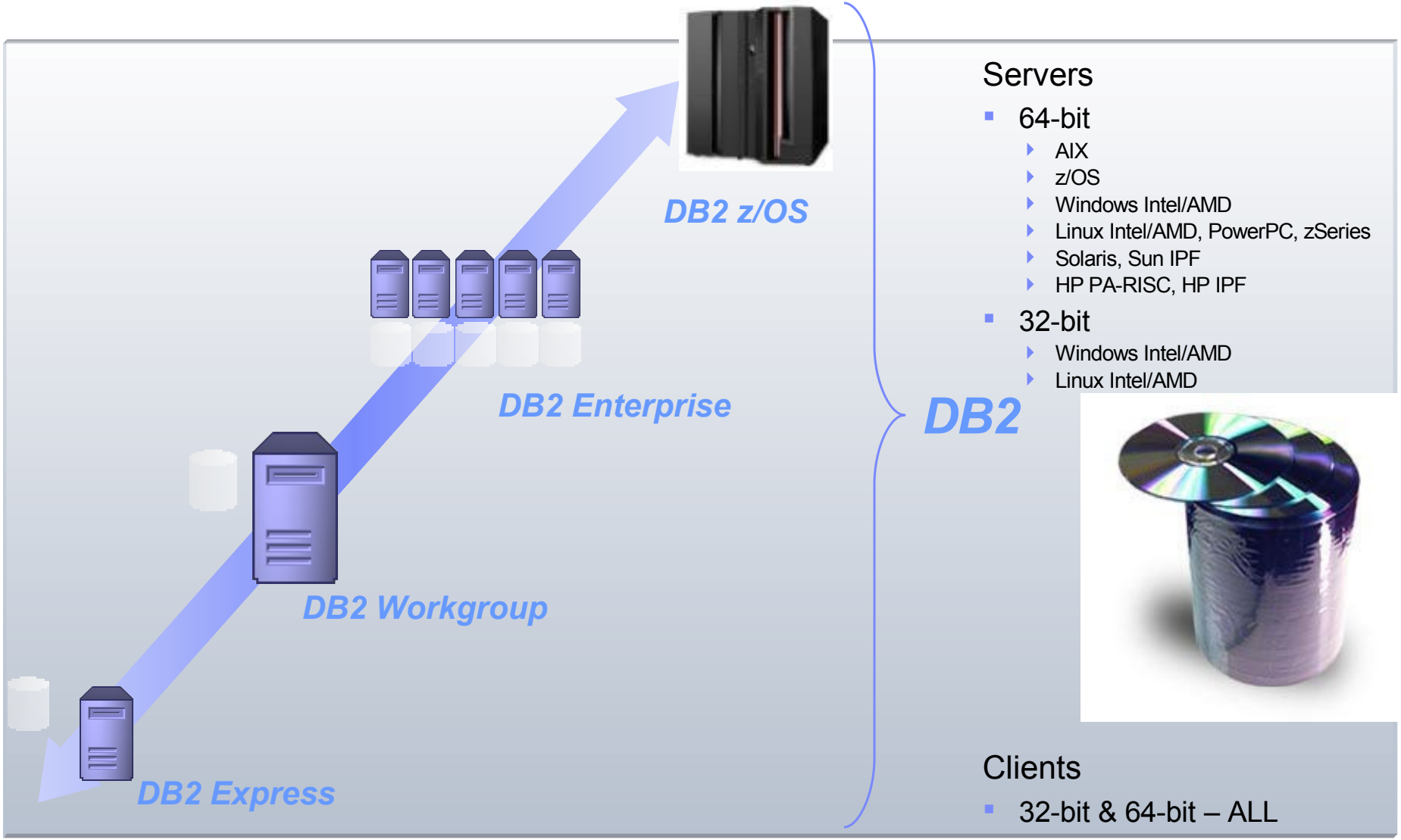
Apache Derby? IBM Cloudscape?

- Apache Derby
  - Complete relational database
  - Pure Java
  - Standards-based
  - Small footprint
    - Embeddable
  - Zero admin
    - Easy to use, easy to deploy
  - Secure
  - Open Source
  - Apache 2.0 License
- IBM Cloudscape
  - Snapshot image of Derby code +
  - Support available from IBM
  - Redistributable, free license
  - Adds:
    - User-friendly installers
      - Java
      - Platform specific (Windows & Linux)
    - JRE (Windows & Linux)
    - Translated docs
    - Sample databases
    - Support for PHP and C (Windows) applications via optional client install

Cloudscape – Small footprint, but powerful solution!!

- Functionality
  - Engine
    - Multi-user, multi-threaded, transactions, row locking, isolation levels
    - crash recovery, backup & restore
  - SQL
    - Tables, temp tables, indexes, views, triggers, procedures, functions
    - Foreign key and check constraints
    - joins, cost-based optimizer
  - Security
    - Encryption, authentication, authorization
  - Multiple databases per system
- Standards
  - SQL
    - SQL92, SQL:1999, SQL:2003
  - Java
    - J2SE 1.3, 1.4, 1.5
    - JDBC 2.0 & 3.0 / J2EE 1.3/1.4
    - J2ME / OSGI
- Requirements
  - J2SE 1.3 or higher
    - JRE fine for running Derby
    - JDK for developing applications
  - **Any hardware that runs J2SE 1.3 or higher**

DB2 Data Server Editions



Why DB2?

Did You Know …

- IBM DB2 is at the core of business applications in …
  - 25 of the Top 25 World Wide Banks
  - 23 of the Top 25 US Retailers
  - 9 of the Top 10 Global Life/Health Insurance Providers

What is Gartner's analysis on the DBMS market?

Gartner Study on DBMS Market
Identifies Spending and Deployment Trends

"The biggest surprise is DB2 from IBM, with 19 percent planning to install DB2. This is nearly 50 percent higher than the number planning to install Oracle."

"Even more interesting is that IBM DB2 is at 27 percent planned [deployments] in the A/P region, and Oracle is only at 10 percent."

"The overall numbers so strongly in favor of DB2 indicate a pending increase in IBM's market share for DBMS during the next few years."

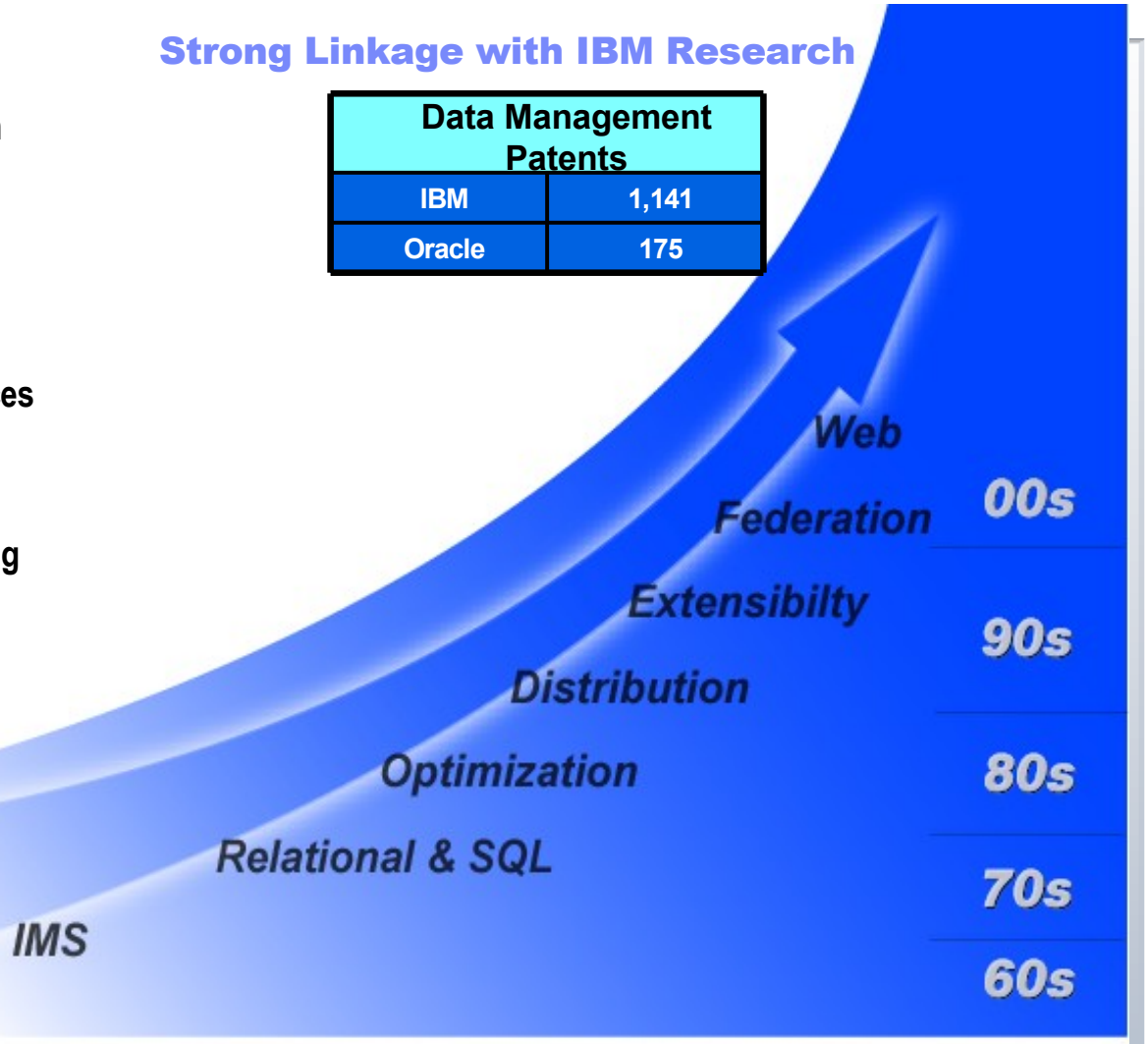
Gartner Group - Colleen Graham, Donald Feinberg, March 29, 2006

DB2 Invention & Innovation Leadership

- Invented the Relational Model & SQL
- First RDBMS with Cost Based Optimization
- First RDBMS with Object Extensions
- First Federated RDBMS
- First RDBMS with Java Support
- First RDBMS with In-Memory Text Search
- First RDBMS with Industry Std. Web Services
- First RDBMS with SMP Support
- First RDBMS with Query Rewrite
- First RDBMS with Integrated OLAP & Mining
- First RDBMS to Publish BI Benchmarks
- First RDBMS to Publish Linux Benchmarks
- First RDBMS Certified for Windows 2000
- First RDBMS cluster on Linux
- First RDBMS on Linux for the Mainframe
- First RDBMS on Linux for IA-64
- First RDBMS on Linux for AMD-64
- First RDBMS to validate for United Linux
- *…and more*

Strong Linkage with IBM Research

Data Management Patents	
IBM	1,141
Oracle	175



Why DB2?
Things you may have expected …

- **Fast**: consistently beats competitors on industry standard and application- specific benchmarks
- **Scalable**: from the smallest smart phone to the largest mainframe. From a few megabytes of data to terabyte data warehouses.
- **Secure**: the best security record of all commercial RDBMS
- **Reliable**: several high availability options for continuous application availability

DB2 Is The Performance Leader
Better performance means more for your money

- #1 TPC-C
  - Outperforms Oracle by 2x
  - SQL Server by 2.7x
- #1 SAP SD 3-tier
  - Outperforms Oracle by 1.7x
  - SQL Server by 1.8x
- #1 Cluster TPC-H
  - IBM leads at 100GB, 1TB, 3TB, 10TB

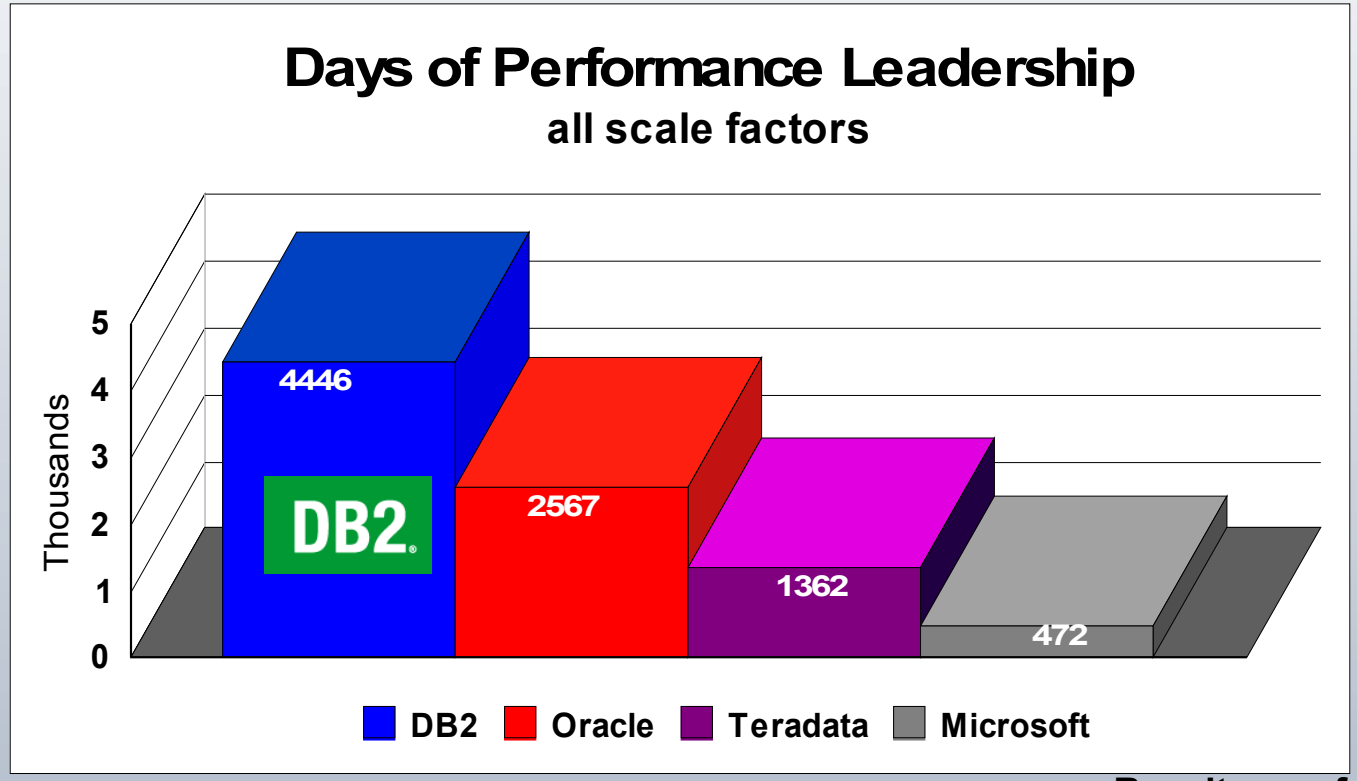
- http://www.tpc.org

- http://www.sap.com/benchma

- http://www.tpc.org

Longevity in Data Warehouse Performance

- First to publish 10TB TPC-H
- Only vendor to hold 10TB TPC-H, Top TPC-C, **and** Top SAP 3-tier simultaneously
- In the leapfrog game that is the world of Benchmarks, DB2 has longevity



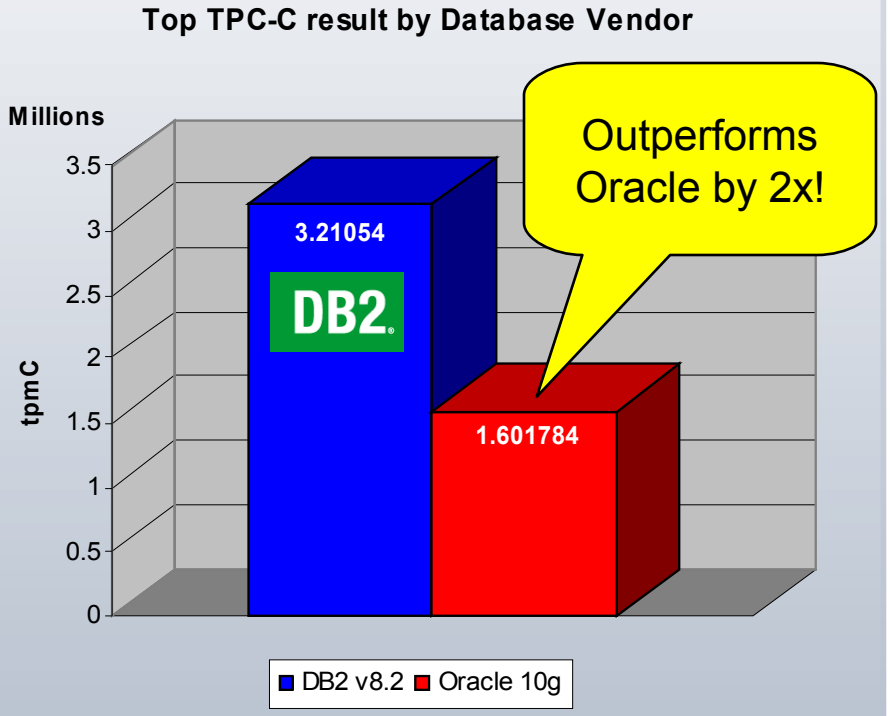
Results as of Nov 30, 2005

DB2 Outperforms Competitors in Overall TPC-C

Scale to support the most demanding workloads

Superior OLTP and Mixed Workload Performance

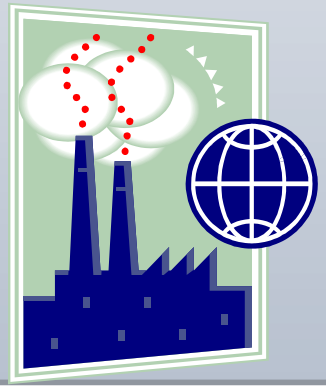
- OLTP Performance Is No Longer Just About Traditional Transaction Workloads
- Important when feeding real-time data into the Data Warehouse
- Mixing ODS Applications into the Warehouse includes OLTP-like Queries



Results current as of April 16, 2006 See Appendix for details.
Check http://www.tpc.org for latest results

Interesting facts on the DB2 TPC–C result

- This fictitious company (TPC-C simulated company) is so large that it would have 270 distribution warehouses in every country in the world
- The performance delivered by DB2 would be sufficient for this company to process a new purchase order from every person on the planet in only 14 days
- If this fictitious company were a popular fast food hamburger restaurant, the sign out front would go from 0 to 3 Billion Sold in **one week**



Why DB2?
… and things you may not have expected

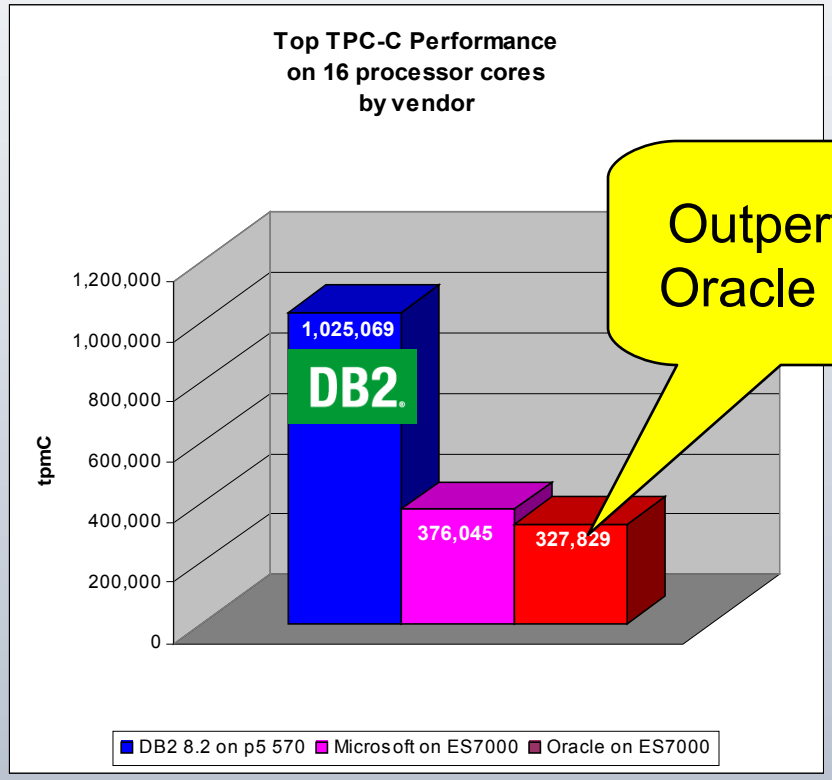
- Programmable:
  - integrated with Java, .NET, PHP, XQuery, SQL/XML etc.
  - complete set of programming tools at no charge
  - Stored procedures, triggers, user defined functions - choice of languages: SQL PL, Java, C#, VB.NET, C, C++, COBOL etc.
- Very inexpensive:
  - low license fees,
  - Free version available
  - Uses less resources (eg. Storage compression)
  - Unbundled features = buy what you need
- XML and Web Services enabled:
  - Hybrid data server that supports both relational and unstructured data natively
  - SQL and XML programming model
  - Delivers data via industry standard web services (SOAP, XML)
  - Consumes data generated by Web Services
- Simple to run:
  - Significant Ease-of-Use Enhancements
  - Faster development & Open standards
  - Autonomic Object Maintenance
  - Automatic Statistics Collection
  - Self-Tuning Resource Management

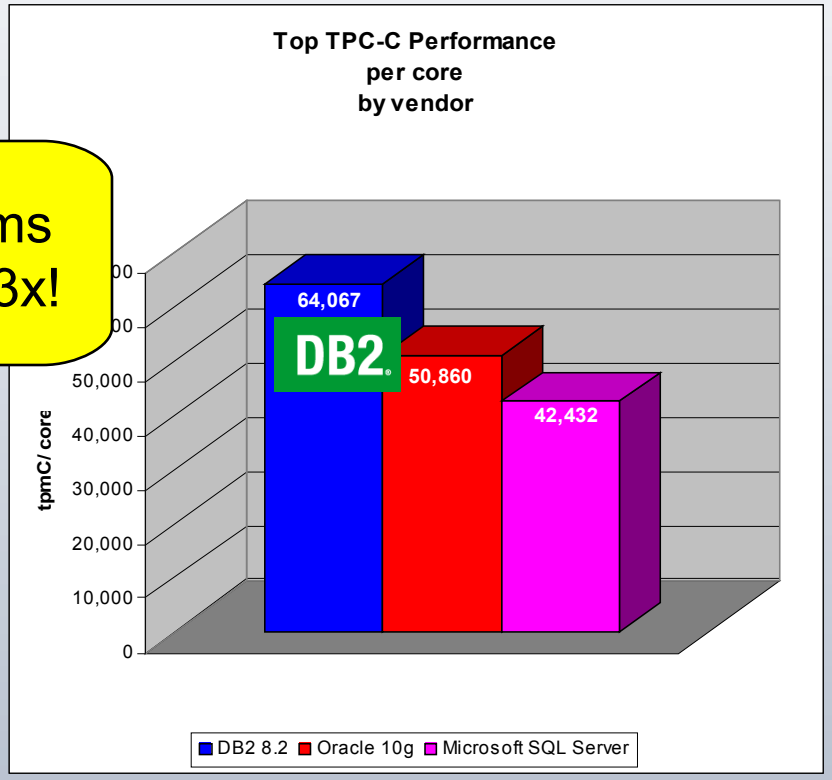
DB2 Uses Less CPU to deliver better performance

Fewer cores means lower software and hardware cost

Core-for-core DB2 delivers more!

- DB2 on 16 core p5 outperforms Oracle and Microsoft by a wide margin
- DB2 delivers best performance per core = fewer processors = cost savings





Results as of 02/14/2006. See Appendix for details

Row Compression

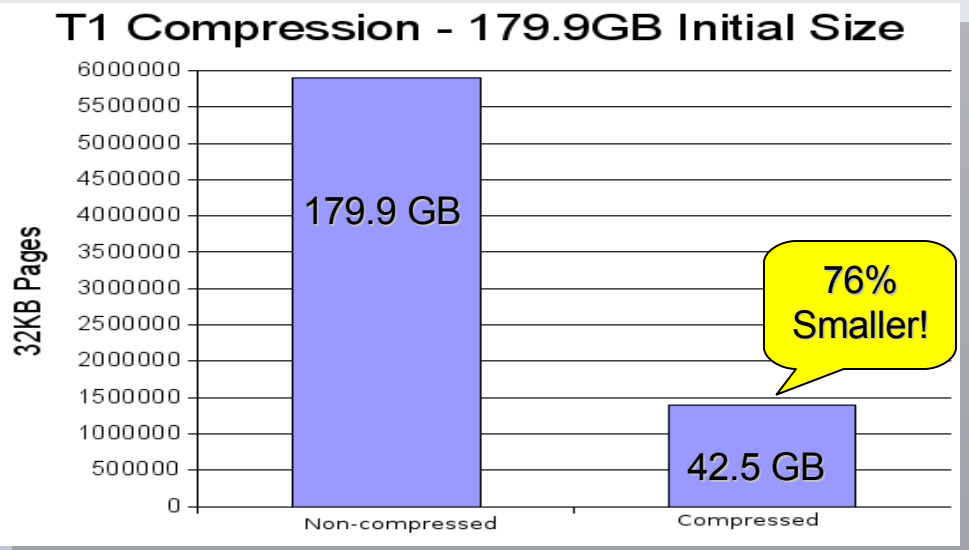
Reduces data storage costs without impact on performance

Fred, Dept 500, 10000, Plano, TX, 24355…

John, Dept 500, 20000, Plano, TX, 24355, Site 3

Fred, (01), 10000, (02),
John, (01), 20000, (02)

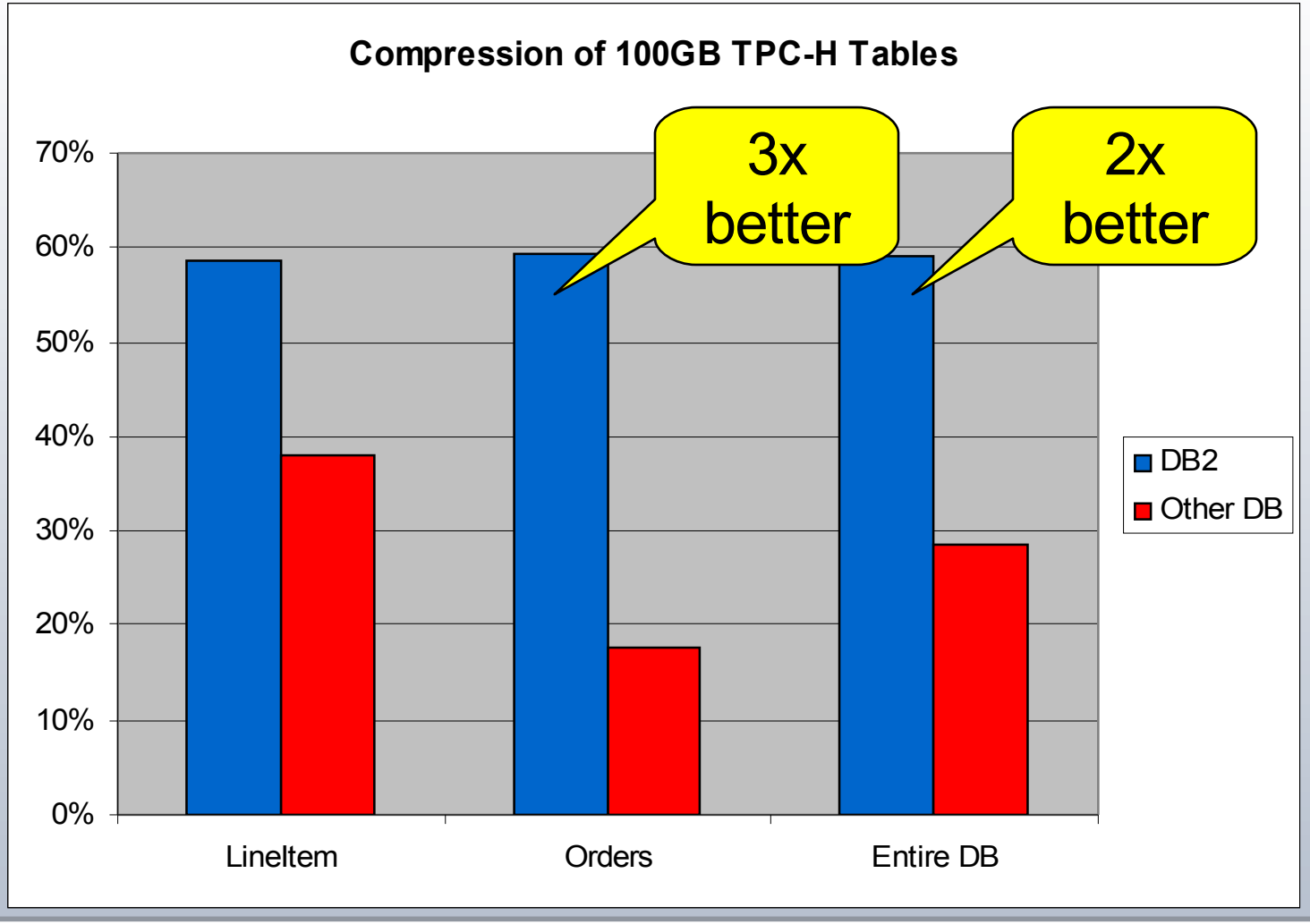
0	Dept 500
1	
0	Plano, TX,
2	24355
…	…



Dictionary contains repeated information from the rows.

How Does DB2 Compression Compare?

Better compressions means lower cost

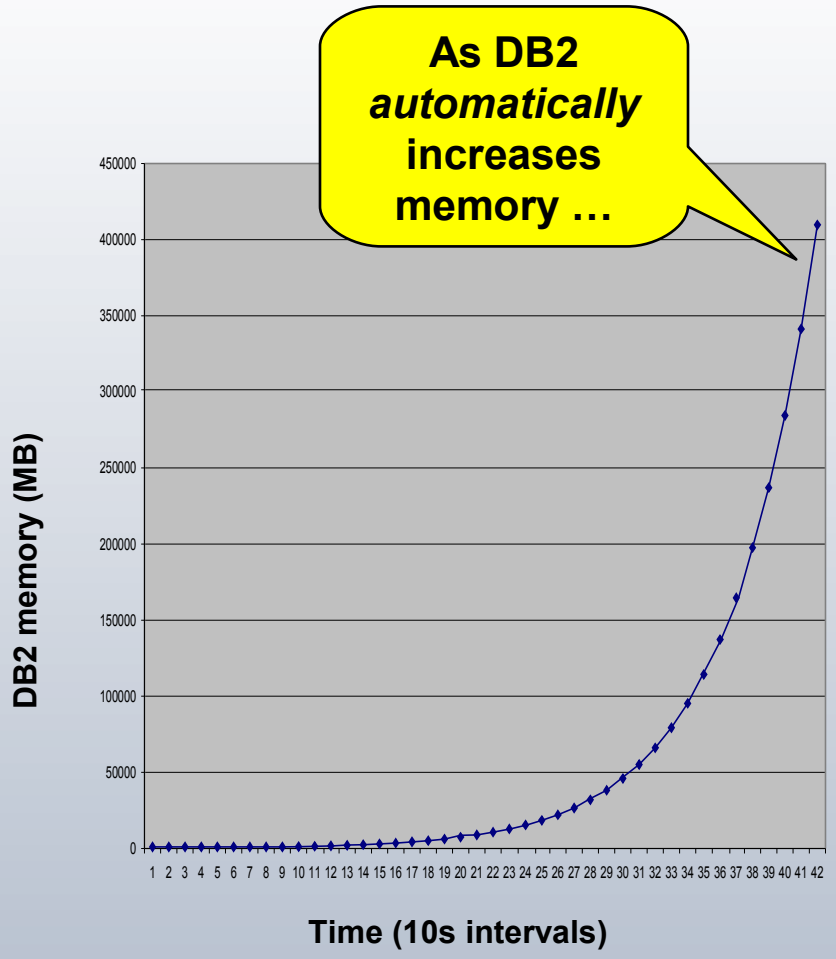


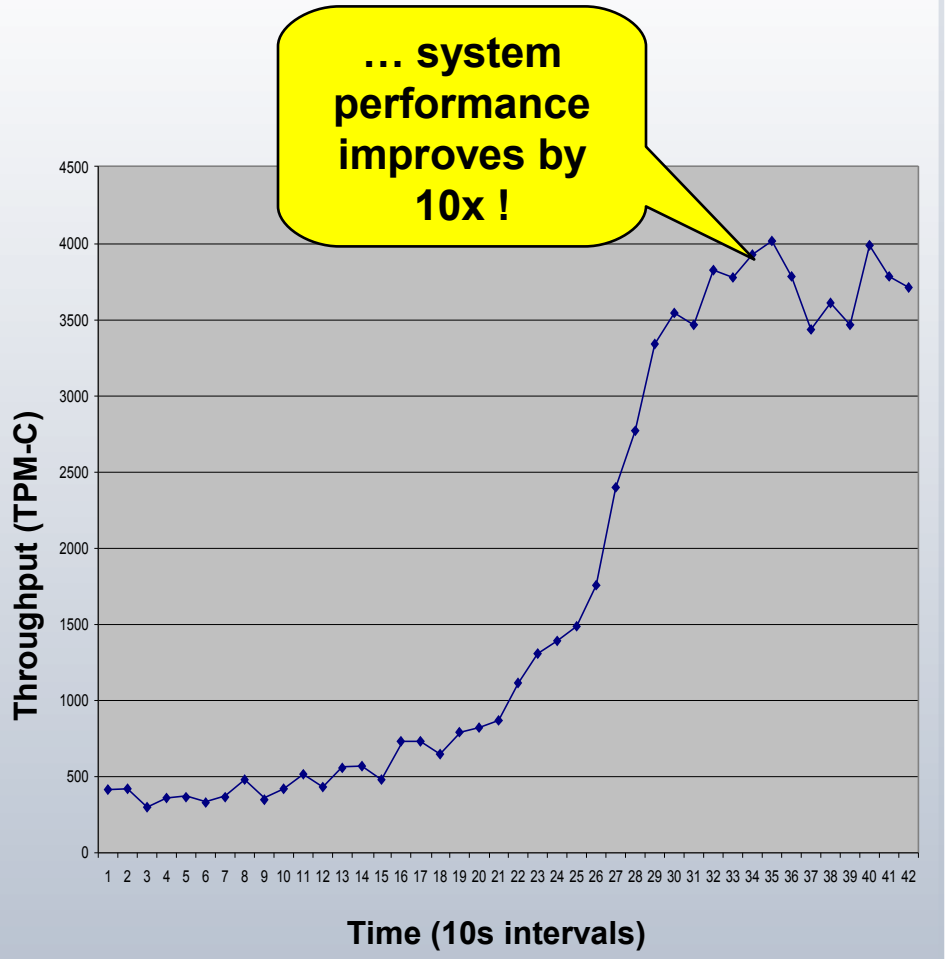
DB2 Simple to Run

Increases Ease of Use and Reduces Cost

- **Configuration adviser**: configure your server based on your workload
- **Design adviser**: optimizes database design (indexes, MQTs, MDCs, partitioning) based on workload
- **Automatic storage management**: don't need to worry about running out of disk
- **Adaptive Self Tuning Memory**: no need to learn how to best allocate memory

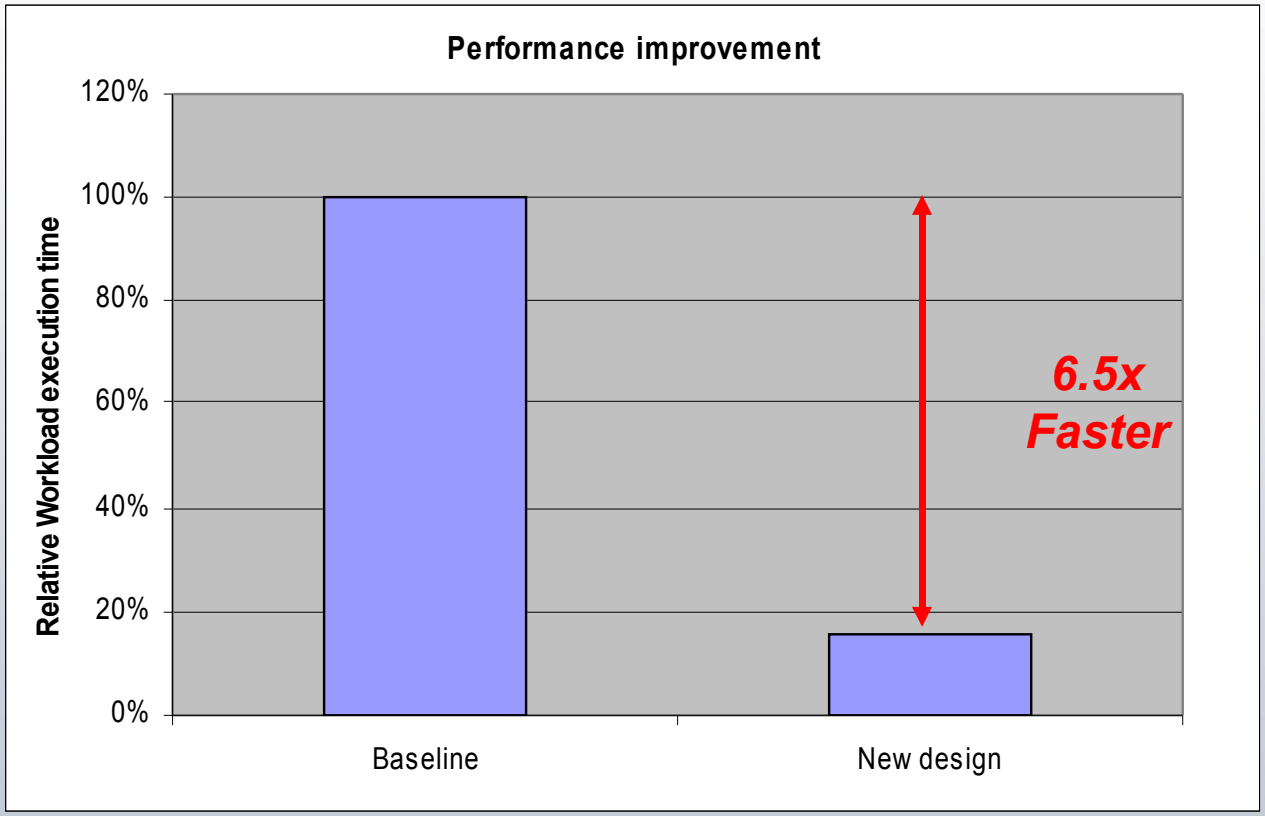
Adaptive Self Tuning Memory





DB2 Design Advisor

Optimizes database design based on workload



DB2 Design Advisor

Recommendation summary:

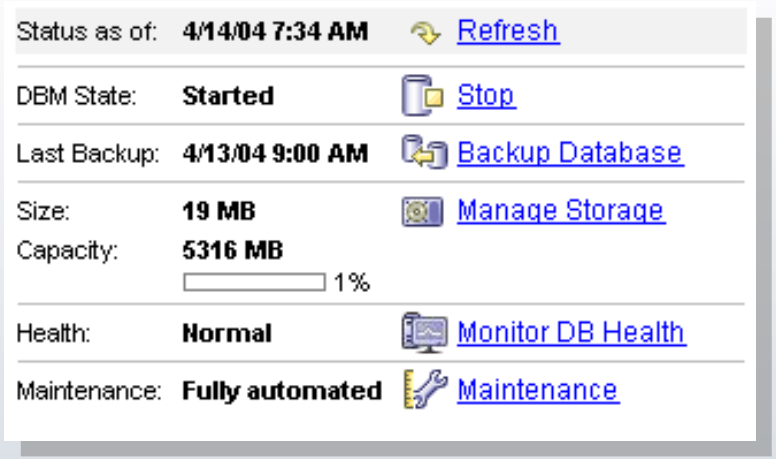
- **20 new indexes**
- **6 new MDC dimensions**
- **4 new partitioning keys**
- **2 new MQTs**

Environment: 1 TB complex query workload in 4 Logical Partitions running on AIX in a 8 CPUs SMP

DB2 is Simple to Run

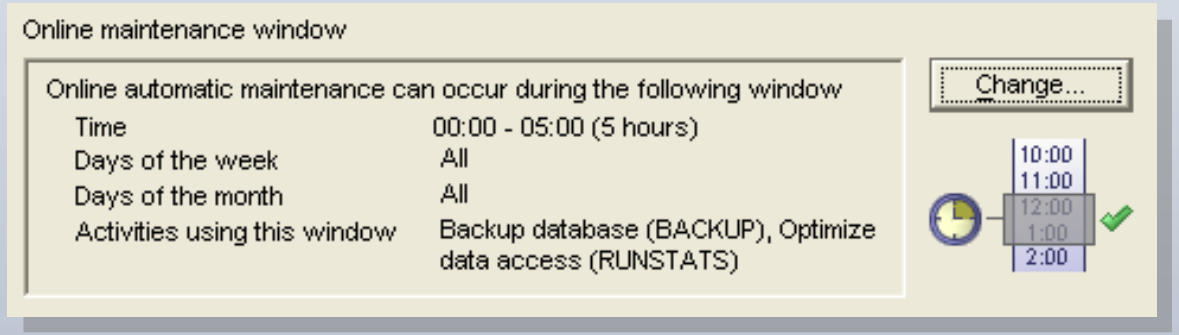
Everyday tasks are simply automatic!

- Backup
- Table Reorganization
- Statistics Collection

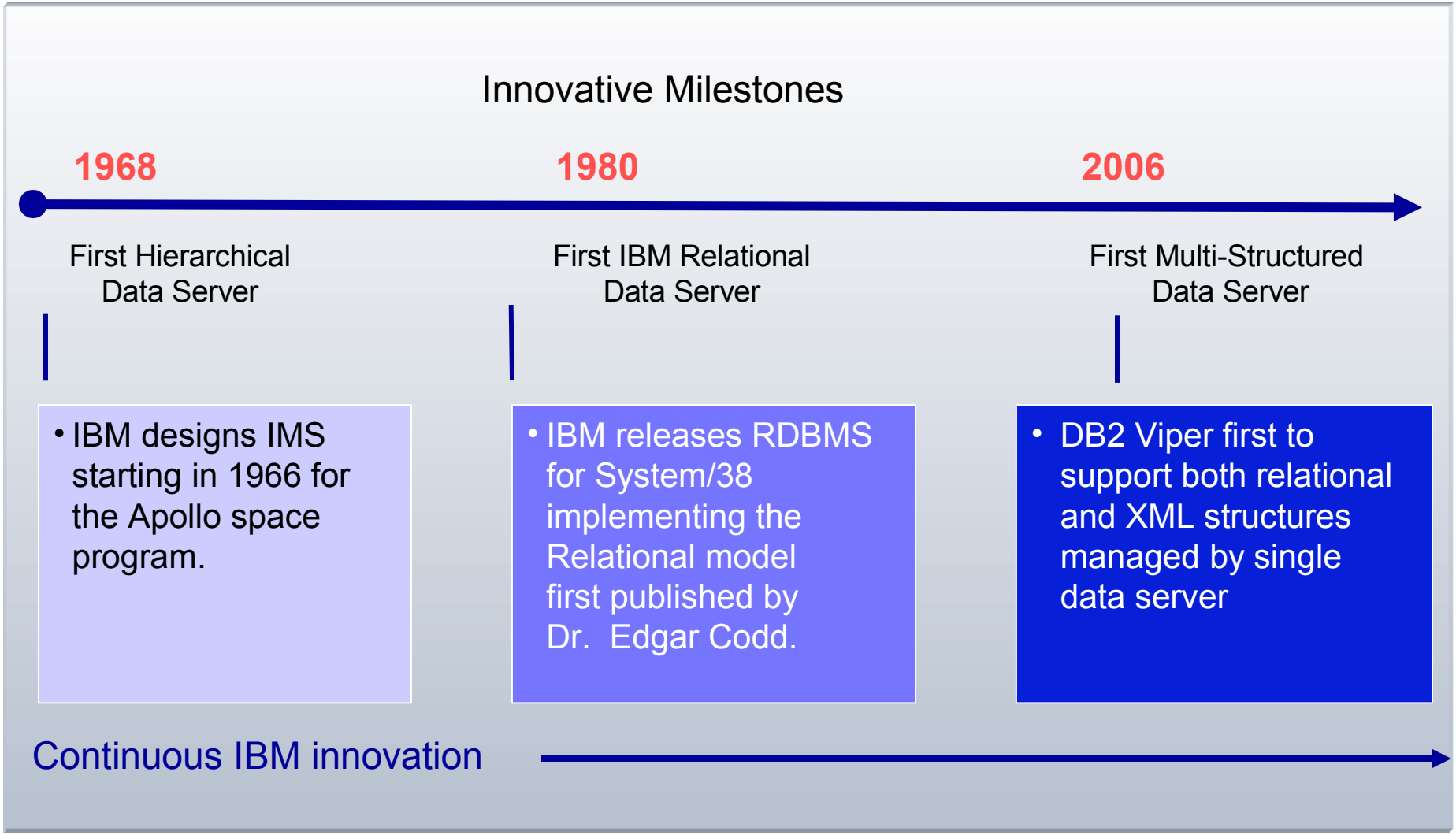


No need to wonder when it's needed to run these utilities

It's Automatic!



IBM leads data server creation and evolution



DB2 Viper – A Pure XML, Relational Hybrid

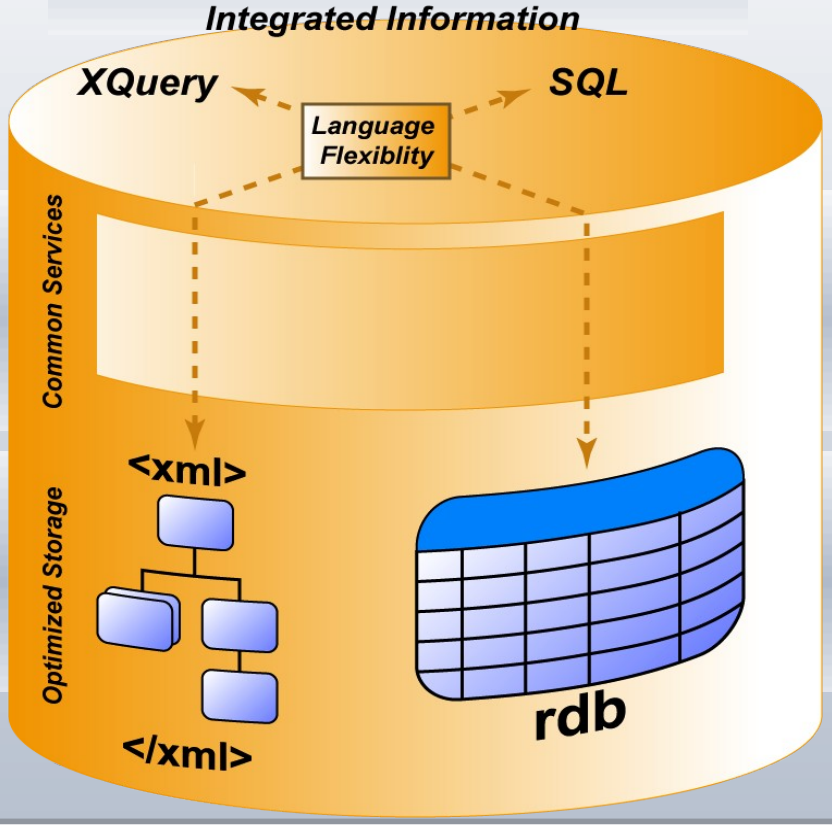
XML Developer

"I see a sophisticated XML repository that also supports SQL."

Familiar Programming Models

SQL Developer

"I see a sophisticated RDBMS that also supports XML."



Mature Services

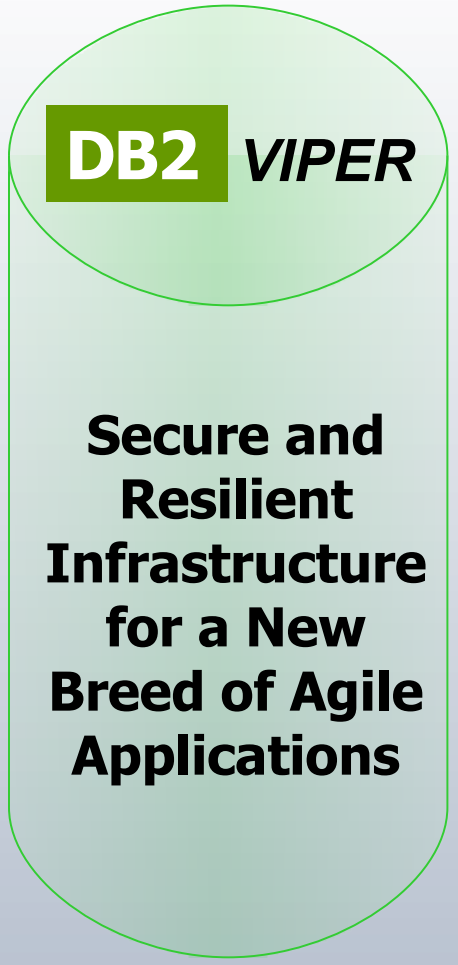
Optimized Storage Models

Familiar Tooling

Optimized Performance & Scale

DB2 Viper – Summary of XML Support

- XML as a native data type
- Pure XML storage and indexing
- XQuery and SQL/XML support
- XML Schema Repository
- Schema validation
- Application Support (Java, C/C++, .NET, PHP, etc.)
- Visual Tooling, Control Center Enhancements
- Annotated schema shredding
- DB2 Utilities: Import/Export, HADR, etc.
- …and more



Does the way you store XML really matter?

"In an interview Wednesday, an Oracle exec took exception to a question about Oracle's own handling of XML in the database. "How we store XML on the database is, excuse me, none of your business. The point is you can write an app using XML standards," said Mark Drake, manager of product management for XML technology for the Redwood Shores, Calif. vendor. ."

Barbara Darrow, *CRN*

http://www.crn.com/showArticle.jhtml?articleID=184429174

If you care about:

- Performance
- Schema evolution
- Programmer productivity

you will care about how XML is stored

Reduce Code Complexity with DB2 Viper

Make Changes Easily with DB2 Viper

```
<DEPARTMENT  deptid="15" deptname="Sales">
      <EMPLOYEE>
         <EMPNO>10</EMPNO>
         <FIRSTNAME>CHRISTINE</FIRSTNAME>
         <LASTNAME>SMITH</LASTNAME>
         <PHONE>408-463-4963</PHONE>
         <PHONE>415-010-1234</PHONE>
         <SALARY>52750.00</SALARY>
      </EMPLOYEE>
      <EMPLOYEE>
         <EMPNO>27</EMPNO>
         <FIRSTNAME>MICHAEL</FIRSTNAME>
         <LASTNAME>THOMPSON</LASTNAME>
         <PHONE>406-463-1234</PHONE>
         <SALARY>41250.00</SALARY>
      </EMPLOYEE>
</DEPARTMENT>
```

Requires:
- **Normalization of existing data !**
- **Modification of the mapping**
- **Change of applications**

IMPROVE BUSINESS AGILITY

Phone

EMPNO	PHONE
27	406-463-1234
10	415-010-1234
10	408-463-4963

Department

DEPTID	DEPTNAME
15	Sales

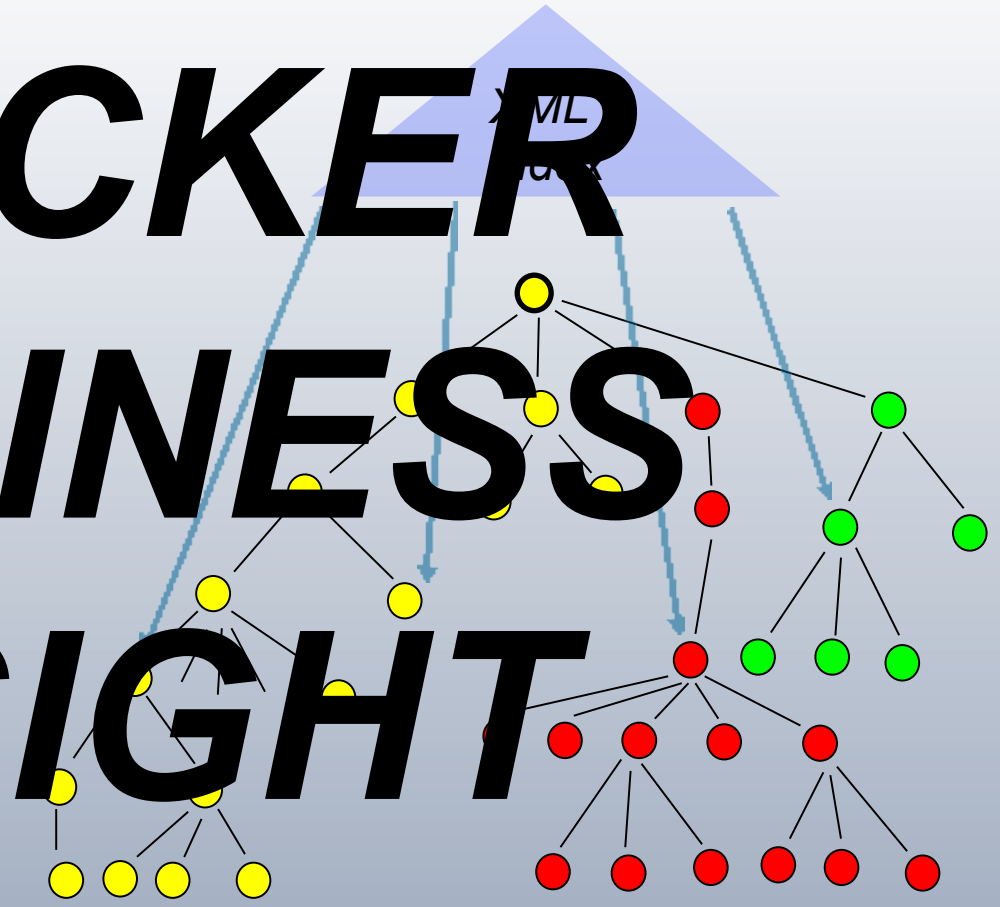
Costly!

Employee

DEPTID	EMPNO	FIRSTNAME	LASTNAME	PHONE	SALARY
15	27	MICHAEL	THOMPSON	406-463-1234	41250
15	10	CHRISTINE	SMITH	408-463-4963	52750

Access more information faster with DB2 Viper

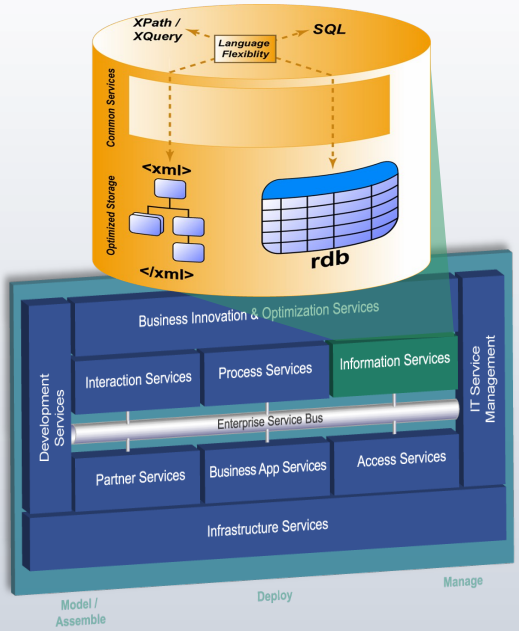
ID	
123	<? xml versi n="1. ? ><purchaseO id="123"><customer id="A6789"><name>John Smith Co</name <street>1 ain /street> <cit Tol o city <state>OH</state> <zip>95141</zip> </address> </customer>
456	…



QUICKER BUSINESS INSIGHT

DB2 Viper: Early adopter feedback

storebrand



Proto-type results using DB2 Viper based SOA solution

Task	With relational DB	With DB2 Viper
Development of search & retrieval business processes	CLOB: 8 hrs Shred: 2 hrs	30 min.
Relative lines of I/O code (65% reduction)	100	35
Add field to schema	1 week	5 min.
Queries	24 - 36 hrs	20 sec - 10 min
Query non-shredded XML element	1 week	½ day

Business Benefits

- Quickly create customized products that customers want
- Expected to process five times more business
- Fast, easy access to richer product & client information

All of the power of DB2 for FREE

- Use for development, deploy in production or distribute as part of our product for free.
- Install on Linux or Windows servers with 2 CPUs and 4B of memory
- No limits on size of database

How to put DB2 on Rails?

Announcing …

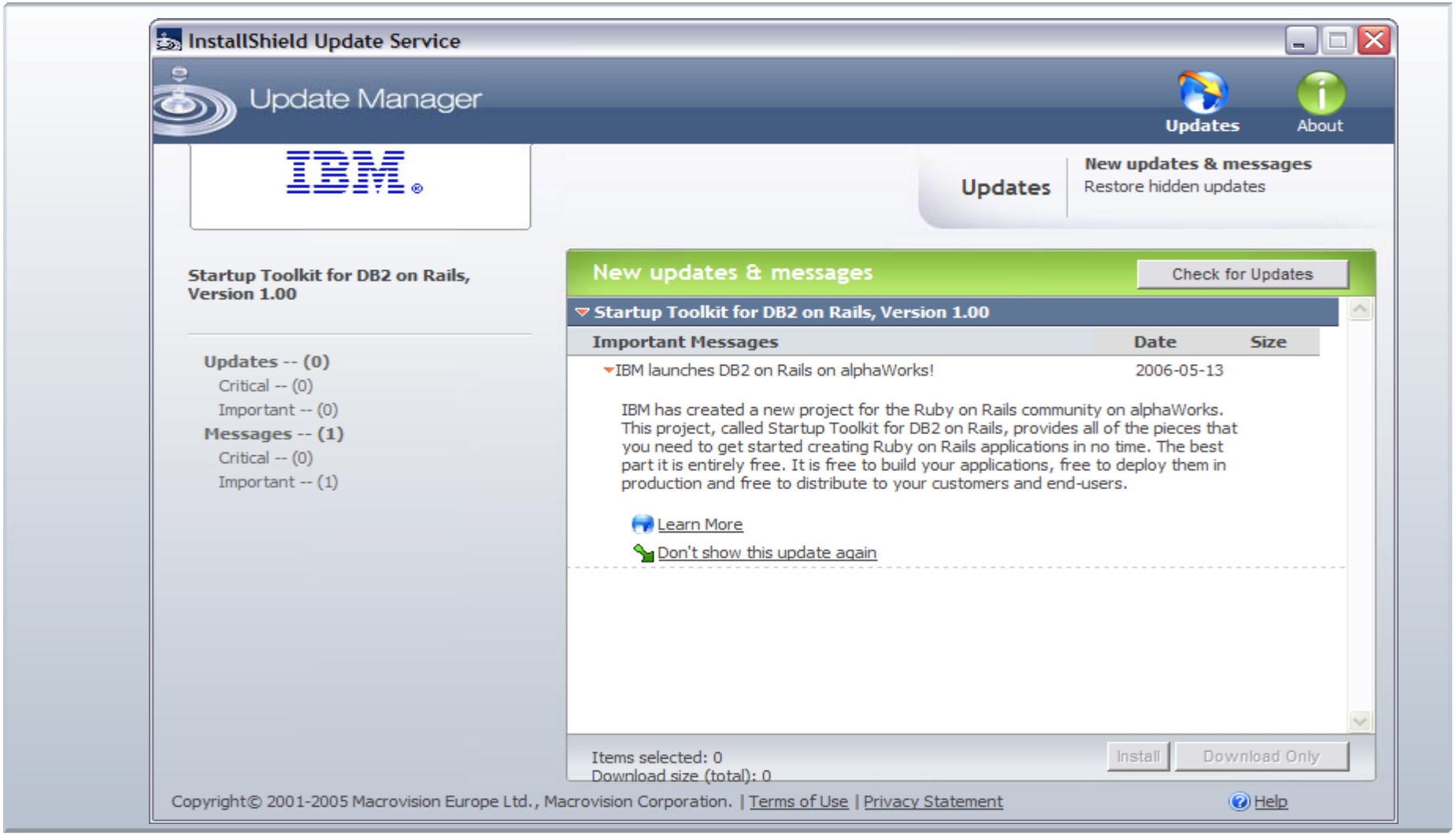
Startup Toolkit for DB2 on Rails

- Available on IBM alphaWorks
  http://www.alphaworks.ibm.com/tech/db2onra
- Integrated installer that creates a complete DB2 Ruby on Rails development environment on a Windows PC
- Source code available for you to do your own builds on other platforms

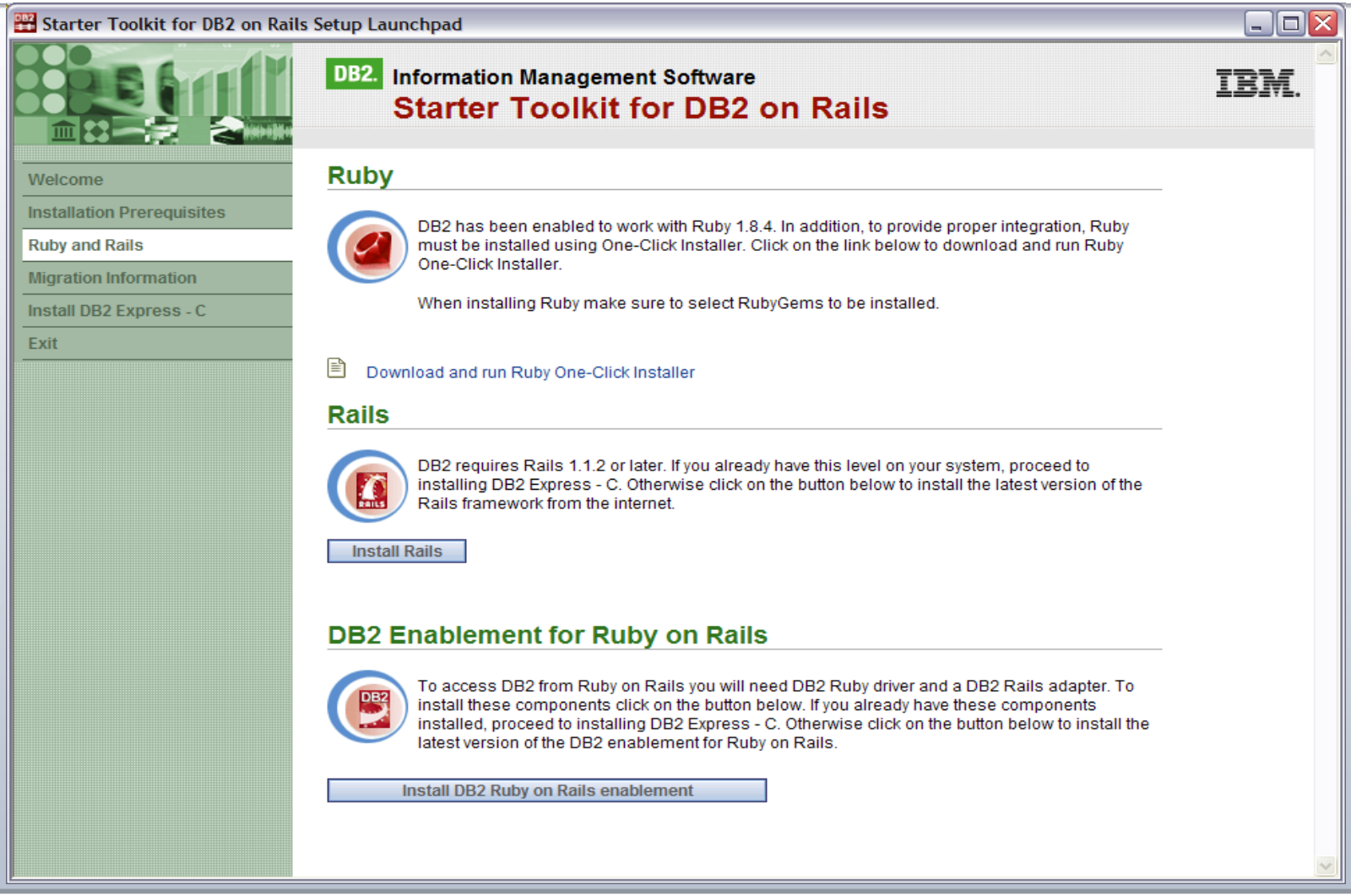
Startup Toolkit for DB2 on Rails
What do you get

- Integrated installer
- Helps you install Ruby 1.8.4 and Rails 1.1.2
- Installs DB2 Express – C Viper and tools for managing database
- IBM developed DB2 Ruby driver and DB2 Rails Adapter
- Demos and tutorials – more to come soon!

Keeps you up to date



What does it look like



What do you need

- Windows XP or Windows Server 2003
- 500 MB of free disk space
- 512 MB of memory but I recommend 1GB if you plan to use DB2 Control Center
- Some free time to learn your way around DB2 and Ruby on Rails

What to expect in the future
No guarantees but this is what we are thinking …

- Expanded platform support for the installer – think Linux
- Expanded platform support for DB2 severs – think DB2 for z/OS and DB2 for i5/OS
- Additional IBM Data Servers – think IDS
- DB2 Driver and adapter on RubyForge
- DB2 driver/adapter as GEMs
- Better support for XML programming model in DB2
- … something wonderful ☺

What do you want to see us deliver?

- Visit DB2 on Rails forum to tell us
  http://www.alphaworks.ibm.com/tech/db2onra
- Send me an email leon@ca.ibm.com
- Some thoughts/questions:
  - What about Mac OS X as development platform?
  - Would you write server-side code (procedures, functions, triggers) if you could do it in Ruby?
  - Do you/would you use Eclipse as your IDE for RoR?
  - How do you see yourself using XML in DB2?
  - …

Summary

- Why use DB2 in your Ruby on Rails projects:
  - Enterprise quality data server: secure, reliable, fast and easy to run.
  - A different kind of data server: handles SQL and XML data in its native form
  - … and you get it for FREE
- How to get DB2 on Rails:
  - Come talk to us in the vendor exhibit area
  - Get Startup Toolkit from IBM alphaWorks
    http://www.alphaworks.ibm.com/tech/db2onrails
  - See the demos

Appendix of required TPC information

TPC-C is a trademark of the Transaction Performance Processing Council (TPC).

Results current as of February 14, 2006

Top 16way results

[1] 1,025,069 tpmC, $4.42/tpmC, available as of 05/31/06, run on a 16 processor 2.2GHz IBM eServer p5 570 running DB2 8.2

[2] 376,045 tpmC, $3.97/tpmC, available as of 01/03/06, run on a 16 processor Unisys ES7000 running MS SQL Server 2005.

[3] 327,829 tpmC, $4.40/tpmC, available as of 07/29/05, run on a 16 processor Unisys ES7000 running Oracle Database 10g.

Top performance per core

[1] 1,025,069 tpmC, $4.42/tpmC, available as of 05/31/06, run on a 16 processor 2.2GHz IBM eServer p5 570 running DB2 8.2

[3] 203,439 tpmC, $3.93/tpmC, available as of 10/17/05, run on a 4 processor IBM eServer p5 570 running Oracle Database 10g.

[2] 42,432 tpmC, $1.96/tpmC, available as of 03/29/05, run on a 1 processor HP Proliant ML350 running MS SQL Server 2000.

Appendix

Required TPC-C information

2. Oracle 10g on 32way p5 595; 1,601,784 tpmC, $5.05/tpmC, Availability 04/20/2005
3. DB2 v8.2 on 64way p5 595; 3,210,540 tpmC, $5.09/tpmC, Availability 05/14/2005
4. Oracle 10g RAC on HP Integrity rx5670 Cluster; 1,184,893 tpmC, $5.52/tpmC, Availability 04/30/2004
5. Oracle 10g on 8way p5 570; 371,044 tpmC, $5.26/tpmC, Availability 09/30/2004
6. DB2 v8.1 on 8way p5 570; 429,899 tpmC, $4.99/tpmC, Availability 09/30/2004

Required TPC-H information

9. DB2 V8.2 on 8 8way p5 575; 104,100 QphH@10000GB, $61/QphH@10000GB; Availability 08/15/2005
10. Oracle 10g on 64way HP Integrity Superdome; 49,104 QphH@10000GB, $118/QphH@10000GB; Availability 03/25/2004
11. Oracle 10g with RAC on 2 64way HP Integrity Superdome; 86,282 QphH@10000GB, $161/QphH@10000GB; Availability 04/06/2005

TPC-C,TPC-H, QphH and tpmC are trademarks of the Transaction Processing Performance Council.

Required SAP Information

- For more information regarding these results and SAP benchmarks, visit www.sap.com/benchmark.
- These benchmark fully complies with the SAP Benchmark Council regulations and has been audited and certified by SAP AG
- The SAP SD standard R/3 Enterprise 4.70 application benchmark has been certified with the following data:100,000 SD users;Avg response time:1.75 sec;Fully processed order line items/hour:10,210,330;OS HPUX 11i; RDBMS Oracle 10g ; HP Integrity Model SD64A, 64-way SMP, Intel Itanium 2 1.6 GHz, 32 KB L1 cache, 256 KB L2 cache, 9 MB L3 cache , 256 GB main memory. Certification 2005021
- The SAP SD standard R/3 Enterprise 4.70 application benchmark has been certified with the following data:168,300 SD users;Avg response time:1.95 sec;Fully processed order line items/hour:16,896,670;OS AIX 5.3; RDBMS DB2 8.2 ; IBM eServer p5 Model 595, 32-way SMP, POWER5, 1.9 GHz, 32 KB(D) + 64 KB(I) L1 cache per processor, 1.92 MB L2 cache and 36 MB L3 cache per 2 processors, 256 GB main memory Certification 2004068