



IBM SOA^{*} Summit

^{*} Informations valorisées et SOA,
le couple gagnant.



IBM SOA*
Summit



DB2 9, XML, and SOA Session Number

Brian Haan
Senior Architect IBM SWG



The Best Foundation for Data: DB2 9 “Viper”

We want to get the most out of our database and reduce the cost.



**On Demand Merchant
CIO**

The foundation for your data management solutions should be DB2 9 “Viper.”



IBM



Why Is DB2 9 the Best Foundation?



New!

- **Best Native XML Support**—Beats Oracle and Microsoft in all key areas of XML storage and handling, including query performance.



New!

- **Best Compression**—Lowers storage requirements, saves money.
- **Best Performance**—Leader in benchmark tests.
- **Best Autonomics**—Automation features ease configuration, tuning, and administration.
- **Best Security**—Best architecture and privilege scheme, easy to configure, easy to administer.
- **Highest Availability**—Best availability and disaster recovery capability.
- **Highest Scalability**—Leader in scale up and scale-out clustering.
- **Best Database for SAP Solutions**—Optimized for SAP, plus inherent advantages.

- **Flexible**
 - No fixed format or syntax—for structured data, semi-structured data, schema-less data.
 - Structures can be easily changed.
 - Easy to extend: define new tags as needed.
- **Platform Independent**
 - Not tied to any platform, operating system, language, or software vendor.
 - XML can be easily exchanged.
- **Easy to "validate" XML—i.e., to check compliance with a schema; any XML parser can do it!**
- **Easy to transform XML documents into other formats (HTML, etc.).**
- **Fully Unicode compliant.**



- **More XML data generated than in databases.**
- **Growing at twice the rate of the total database market.** [IDC]
- **XML is pervasive in all kinds of organizations.**
- **Almost every sector has XML-based standards.**

Figure VI.2: Market Size by XML Data Store Solution Type

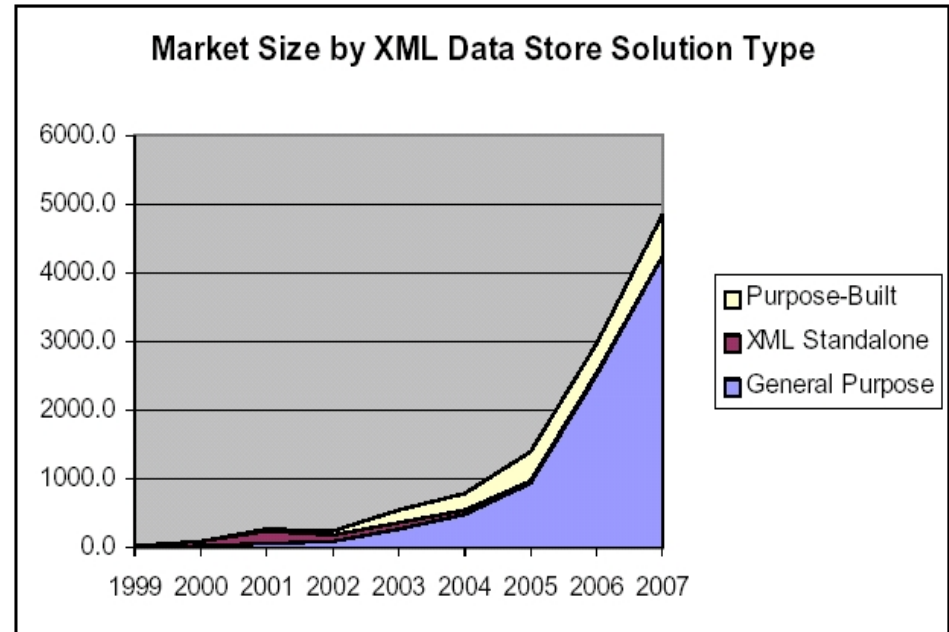


Chart Sources: XML Market Opportunities, Forecasts and Strategies, 2004-2009
WinterGreen Research Inc. ZapThink

“XML is emerging as the standard way to represent data for the purposes of interchange between applications and enterprises across structured and unstructured sources.”—Gartner Group, “Gartner Study on EIM Highlights Early Adopter Trends and Issues,” David Newman, February 7, 2006.

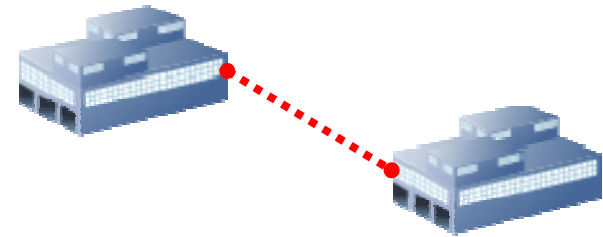
- **SOA**

- Web Services messages are XML

- **Business-to-Business Integration**

- Platform-independent transport mechanism.

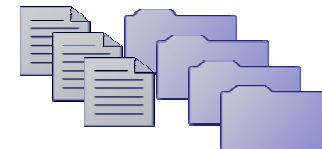
Transaction orders may be defined in XML



- **Forms and Document Processing**

- Government and legal industry require digital signature

Tax forms require signature & change year to year



- Documents often contain sub-documents

Literary materials contain books, chapters, and sub-chapters

XML – The Change Is Fundamental



■ Relational is a data model

- Relations (tables)
- Attributes (columns)
- Set based w/ some sequences
- Strict schema

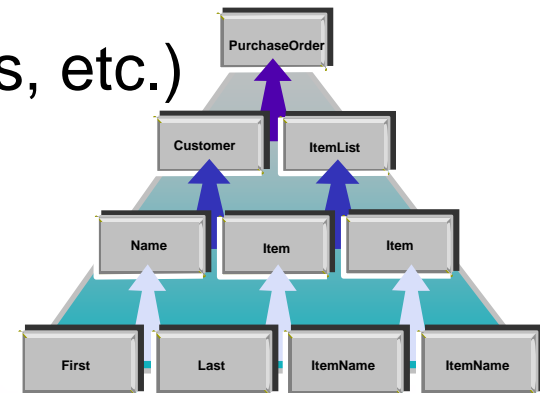
POID	CustomerID	ItemID
12	1	2
162	3	4
162	3	5

Id	LastName	FirstName	Street	City	State	Zip
1	Pirahesh	Hamid	1 Harry Rd	San Jose	CA	95141
3	Selinger	Pat	555 Bailey Ave	San Jose	CA	95141

ItemID	Name
2	#6 wire nut
5	Small Walrus
4	Apollo moon rocket

■ XML is a data model

- Hierarchical tree structure
- Nodes (elements, attributes, comments, etc.)
- Relationships between nodes
- Sequence based w/ some sets
- Flexible schema



Ways to Store XML



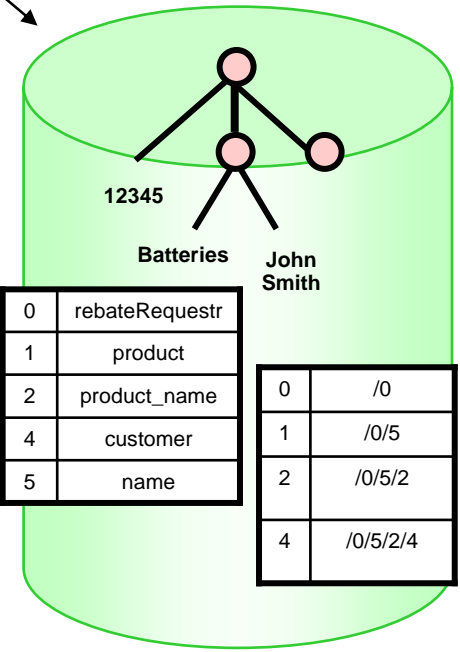
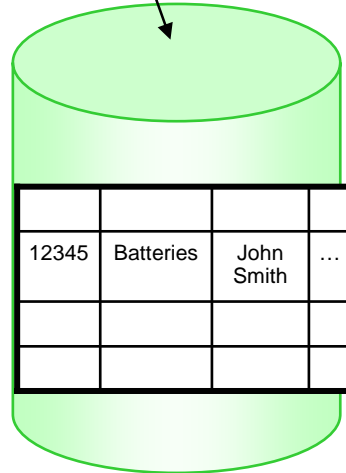
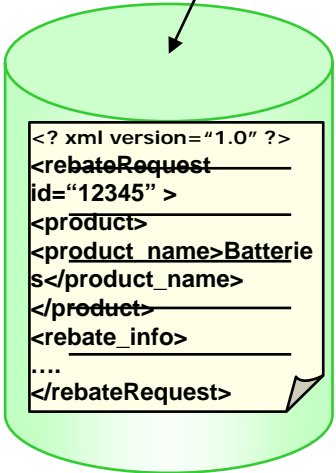
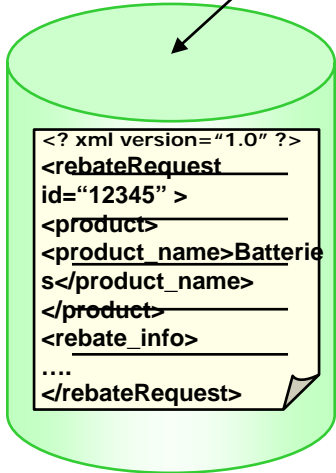
```
<? xml version="1.0" ?>
<rebateRequest id="12345" >
  <product>
    <product_name>Batteries</product_name>
  </product>
  <rebate_info>
    <rebate_amt>$1.00</rebate_amt>
  </rebate_info>
  <customer>
    <name>John Smith Co</name>
  </customer>
  ....
</rebateRequest>
```

XQuery
pureXML

File Name

SQL XML

SQL XML



File System

BLOB

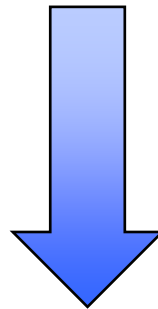
Shredded

Native Integration

- Problems with BLOB's
 - Storing XML as a binary large object (BLOB) or character large object (CLOB) means the database remains unaware of the internal document structure.
 - Searching or retrieving a part of an XML document requires parsing at query time.
 - Parsing at query time is expensive and slows performance.
- Problems with “Shredding”
 - Document integrity is not guaranteed.
 - Because the XML structure is “hard-wired” to specific tables and columns, changing schemas is costly.
 - Retrieving all or part of an XML document can require complex, costly, multi-table joins to reconstruct the various pieces.



Native XML Database With an XML
Query Language



DB2 version 9 Has It! pureXML®



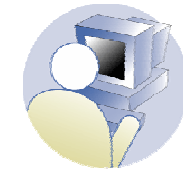
■ XML Data Gets Needed Protection

- Backup and recovery features to ensure continuity
- Data is protected using database security
- Inherits RDBMS transactional support



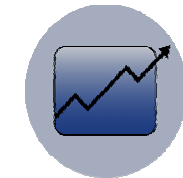
■ Simplified XML Data Access

- Centrally store and access difficult-to-retrieve data
- SQL or XQuery can be used to retrieve data
- Join XML data with its related relational data



■ Search Speed

- Search documents quickly and efficiently using proven search optimization engine of mature database



■ Optimize Existing Investments

- Use existing technology infrastructure and skills to store and manage both relational and XML



How Does a Programmer Access XML Data?

DB2 9 Supports Two Ways to Access XML Data

1) XQuery

- A pure access language designed for accessing XML data
- Is built on XPath expressions
- An emerging standard defined by the W3C
- Is supported by all the major database engines (IBM, Oracle, Microsoft, etc.)

2) SQL/XML

- Allows SQL to access XML
- XML extensions to traditional SQL language
- Defined as part of the SQL 2003 standard



- **Toolkit for DB2 Development**
- **Eclipse-based tool (replaces previous DB2 Development Center)**
- **XML Support**
 - Tool for creating XQueries
- **Stored procedure debugger**
- **Supports Java and .NET development**
- **Can create a routine for Unix, Linux, Windows and deploy to DB2 for z/OS**



ODM Needs to Support New Business Processes

We want to let customers apply for rebates from the manufacturers online

We can connect to the manufacturer's systems with web services

The rebate information is in XML format



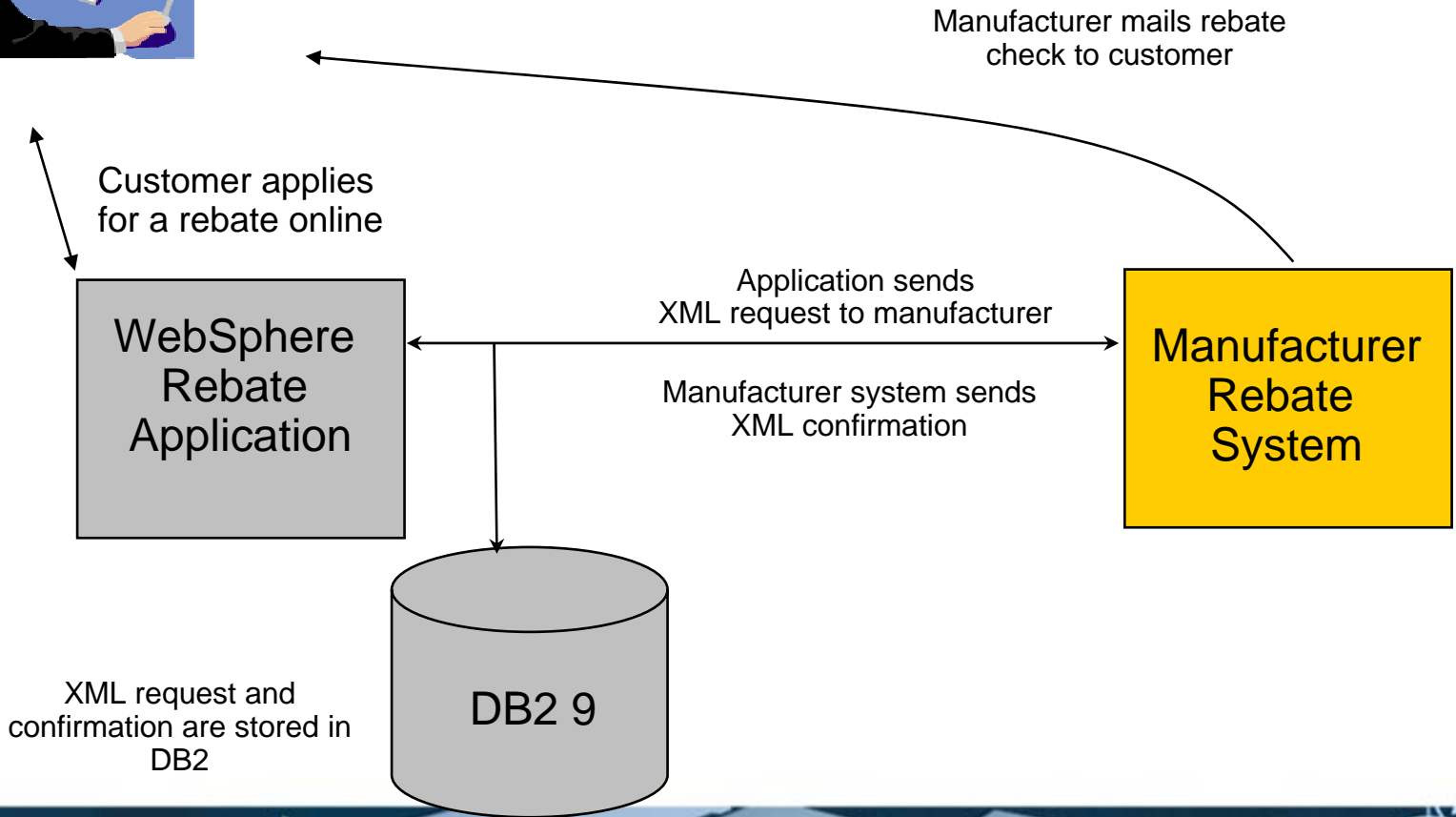
CIO

With DB2 9's pureXML you can use one database to manage both your relational and XML data.



IBM

ODM's New Manufacturer's Rebate Process



- Using XQuery to retrieve XML data

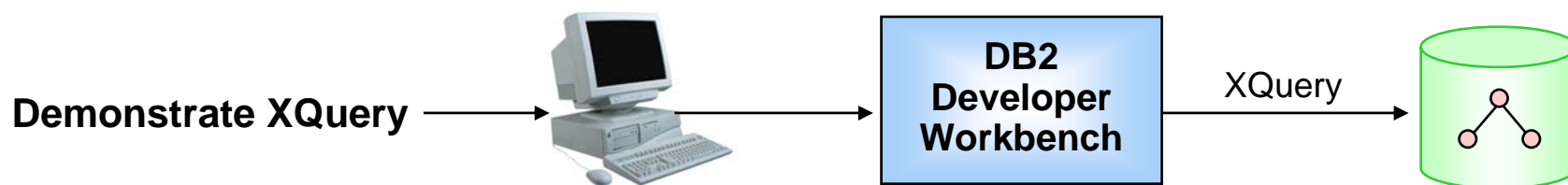
The screenshot shows the XQuery editor interface. On the left, a tree view displays the structure of an XML document named 'person.xml'. The root is '#document', which contains a 'ppl:people' element. This element has two attributes: 'xmlns:pp1' and 'xmlns:xsi'. It contains a 'person' element with three children: 'firstName [Joe]', 'lastName [Smith]', and 'nickName [Joey]'. On the right, the 'For Logic (FLWOR)' editor is active. It shows a 'For Each' loop over the 'person' element. The 'Let' section is empty. The 'Where' section contains a table with the following content:

Operand 1	Operator	Operand 2	Conjunction
<input checked="" type="checkbox"/> firstName	=	<input checked="" type="checkbox"/> firstName	and
<input checked="" type="checkbox"/> lastName	=	<input checked="" type="checkbox"/> lastName	

The 'Return' section shows a table with one row: 'zipAndNickName'. Below the table, a note says 'Build an XML fragment to be returned with each qualified iteration'.

The screenshot shows the XQuery code in the editor. The code is as follows:

```
values(XMLQUERY('
declare boundary-space strip;
declare namespace tns0="http://www.example.org/address";
declare namespace pp10="http://www.example.org/people";
for $address0 in db2-fn:xmlcolurn("XQJOINADDR.ADDRESS")/tns0:addr
for $person0 in db2-fn:xmlcolurn("XQJOINPERS.PERSON")/pp10:people
where $person0/firstName/fn:string(.) = $address0/firstName/fn:string(.)
return
<zipAndNickName >
  (
    ($address0/zipCode,
     $person0/nickName)
  )
</zipAndNickName>
' RETURNING SEQUENCE))
```



Native XML

- XML as a native data type
- pureXML storage and indexing
- XQuery
- XML Schema Repository
- Schema validation
- Application Support (Java, C/C++, .NET, PHP, etc.)
- Visual Tooling, Control Center Enhancements

Traditional

- SQL/XML support
- Still supports shredding and CLOBs

DB2 9

Secure and Resilient Infrastructure for a New Breed of Agile Applications

What Do People Say About DB2 9's XML Technology?



- ***"...this leaves Oracle and Sybase ...well behind the curve, with Microsoft and the others more or less out of sight."***

– Philip Howard, Bloor Research, The Register



- ***"You want to be able to take those data-centric things in XML and put them into a database without a loss of fidelity, and this is one area where IBM is going further than Oracle and Microsoft."***

– Peter O'Kelley Burton Group, Internet News





- **SQL Server uses BLOB-based storage**
 - DB2 9 uses a new storage mechanism tailored expressly for XML.
- **SQL Server requires tables for at least one relational column in order to index the XML data**
 - DB2 enables users to create tables that contain only XML columns rather than forcing users to create tables with both relational and XML columns.
- **SQL Server translates XQuery statements into SQL**
 - DB2 9 has full XQuery support.
- **DB2 9's performance is better than SQL Server's at storing and retrieving XML data**
 - DB2 can retrieve full XML documents 5X faster than SQL Server. For some document types, 100X faster.
 - DB2 can insert full XML documents 2-3X faster than SQL Server when indexes are defined.



Oracle Stores XML in Two Ways: Structured and Unstructured

■ Structured Storage

- Documents for only one schema can be stored in an XML column, and the schema cannot be changed.
 - DB2 9's doesn't require mapping XML data into SQL columns; furthermore, DB2 can accommodate multiple XML schemas so that changing business needs won't "break" the existing database design.
- XPath/XQuery is translated into SQL for execution.

■ Unstructured Storage

- Queries that reference specific XML elements or attributes cause Oracle to parse the XML documents at runtime – an expensive proposition.
 - DB2 stores XML data in a parsed hierarchical representation, and never parses the XML data at runtime.
- Indexes possible, but only on *non-repeating* elements/attributes, XPath/XQueries require XML parsing at run-time.



DB2 vs. Oracle for Industry Standard Schema Support



Schema	DB2 9	ORACLE 10G R2 Error message:
FIXML v4.4 (Financial, Trading)	✓	ORA-30951; Exceeding length
FPML v4.1 (Financial, Derivatives)	✓	ORA-01792 ; Columns required > 1000
ACCORD (Insurance)	✓	ORA-01792; Columns required > 1000
HL7 (Healthcare)	✓	ORA-00600 ; Internal Error / Cycles
PureEdge (XML Forms)	✓	OCI-31167 Text Nodes > 64K
GJXDM (Government)	✓	ORA-31079; Reference Error / Instances

01/2007 IBM Information Management Evaluation Tests



DB2 9 Is the Best Choice for XML Handling



XML Storage Model	IBM DB2 9	MicrosoftSQL Server	Oracle 10g
True Native XML Storage	YES	Limited Parse XML at Runtime	Limited Parse XML at Runtime
Schema Flexibility	YES	NO	NO
Full Standard XQuery Support	YES	Limited Non-std	Limited Non-std

All “Native” XML Support Is Not Created Equal!



Challenge

- Reduce development time and improve response time when using XML data

Solution

- DB2 9 a combined XML / Relational Information Server

Benefits

- Reduce development time and improve response time associated with storing XML
- Schema changes in **minutes** not days
- New database search and retrieval methods using native XML in **minutes...**



Implications

- Deeply integrated XML support enables unique client application opportunity
- XML data services critical to SOA implementations

The Best Foundation for Data: DB2 9 “Viper”

The pureXML™ sounds great, but what about helping in reducing costs.



**On Demand Merchant
CIO**

DB2 9's compression, autonomics, and availability enhancements can help reduce your TCO.



IBM



Why Is DB2 9 the Best Foundation?



New!

- **Best Native XML Support**—Beats Oracle and Microsoft in all key areas of XML storage and handling, including query performance.



New!

- **Best Compression**—Lowers storage requirements, saves money.
- **Best Performance**—Leader in benchmark tests.
- **Best Autonomics**—Automation features ease configuration, tuning, and administration.
- **Best Security**—Best architecture and privilege scheme, easy to configure, easy to administer.
- **Highest Availability**—Best availability and disaster recovery capability.
- **Highest Scalability**—Leader in scale up and scale-out clustering.
- **Best Database for SAP Solutions**—Optimized for SAP, plus inherent advantages.



- Tables will compress in the range of 55% - 75%.
- Overall database storage savings average around 40% - 45%.
- **That's 45% less disk space** needed to support a DB2 9 database!
- Saves storage and money
- Usually performs better overall



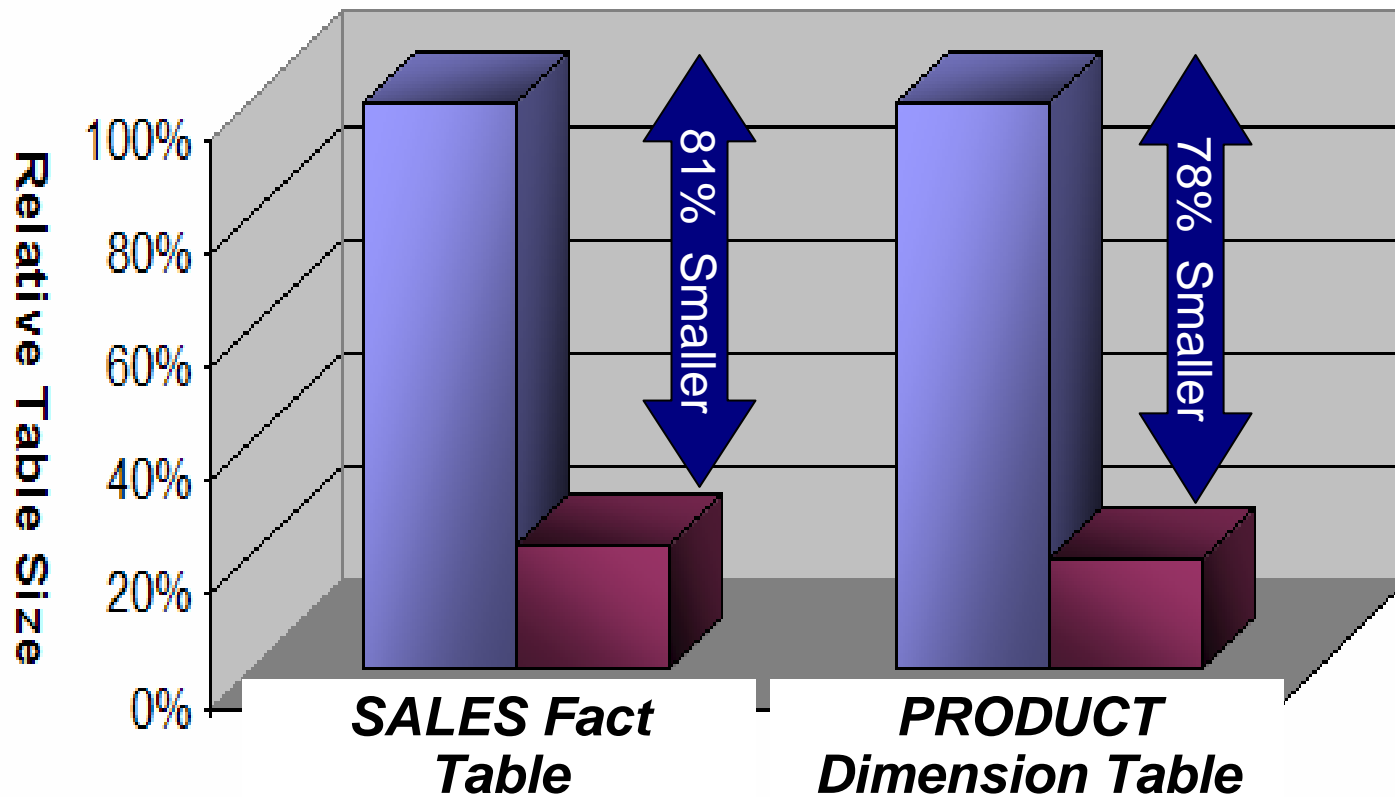
- **Implemented by software**
 - Hardware compression with DB2 for z/OS

- **Compression looks for repeating patterns across the entire table**
 - When it finds a pattern, it automatically replaces that string in the rows with a 12-bit symbol
 - Symbols are stored in a dictionary for fast lookup

- **Data resides compressed on pages (both on-disk and in memory)**
 - Significant I/O bandwidth savings
 - Significant memory savings
 - Some CPU costs
 - Rows must be decompressed before being processed for evaluation

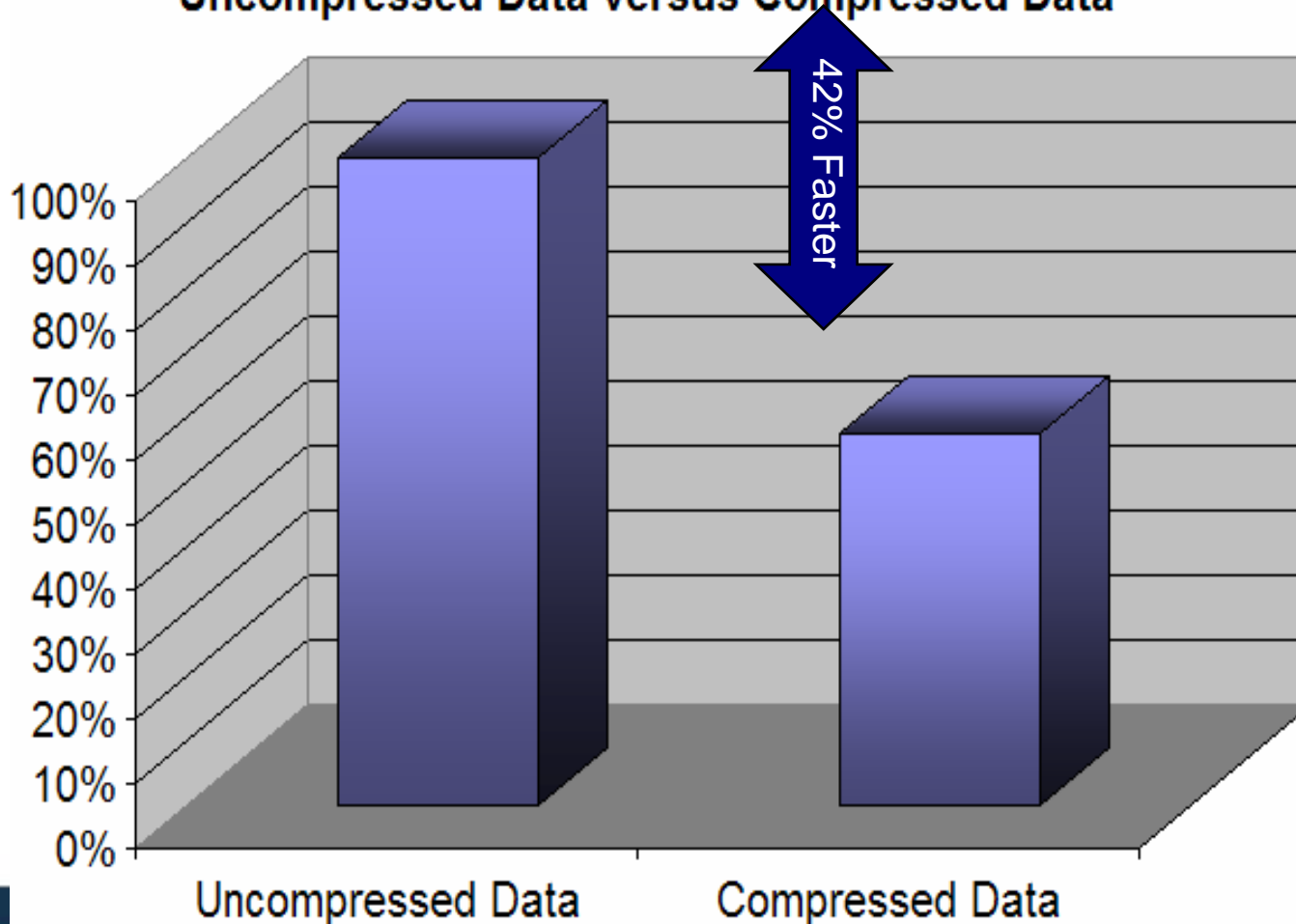


Overall Storage Savings: Uncompressed Data versus Compressed Data





Overall Workload Performance: Uncompressed Data versus Compressed Data



DB2 – Better Compression Than Oracle

- **TPC-H is a well known data warehouse benchmark**
 - Each vendor uses the same tables and same data
 - Oracle published their compression rates for TPC-H tables at the VLDB conference in 2003
 - IBM ran the same tests on the same tables

- **Test results – DB2 reduces cost by requiring less storage**

Table	Compression Ratio	
	Oracle	DB2
LINEITEM	38%	58% (1.5x better)
ORDERS	18%	60% (3x better)
Entire Database	29%	59% (2x better)



The Three Key Performance Benchmarks

- **There are two large-scale standardized performance benchmark tests that measure the performance of a transactional-type workload:**
 - **TPC-C** – industry standard OLTP benchmark.
 - **SAP SD Standard Application Benchmark** – represents an SAP R/3 Sales and Distribution application.
- **There is one large-scale standardized performance benchmark test that measures the performance of a reporting or decision support type workload:**
 - **TPC-H** – industry standard BI/DSS/DWH benchmark.

DB2 Is the Only Data Base to Win the Performance Triple Crown



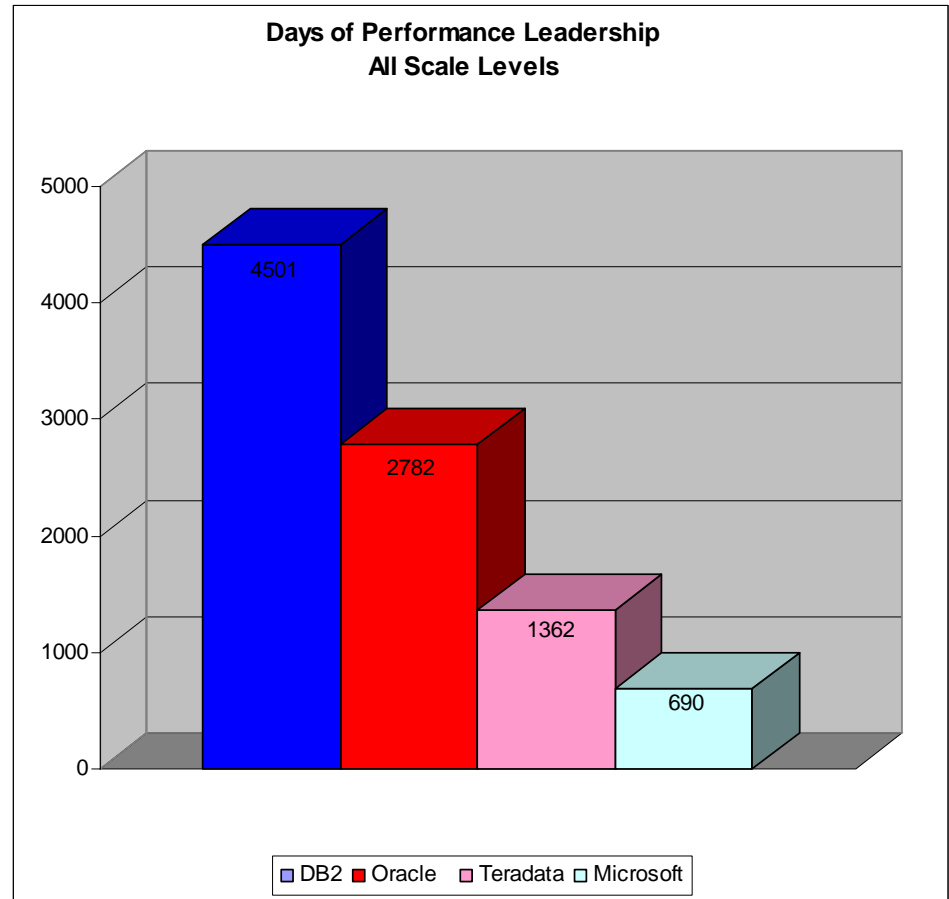
- #1 in TPC-C Performance
 - Leading industry standard OLTP benchmark by 2x over Oracle
- #1 in SAP SD 3-tier Standard Application Performance
 - Leading SAP SD R/3 performance benchmark by 68% over Oracle
- #1 in 10TB TPC-H Performance
 - Largest data warehousing benchmark leading Oracle by 66% using 11% fewer CPUs



Longevity in Performance Leadership



- **First to publish 10TB TPC-H.**
- **Only vendor to hold 10TB TPC-H and top TPC-C and top SAP 3-tier simultaneously.**
- **In the leapfrog game that is the world of Benchmarks, DB2 leads year after year.**



Results current as of July 14, 2006, Check <http://www.tpc.org> for latest results



Reduces Cost and Makes Administrators' Tasks Easier

- **Simplifies**

- Installation
- Maintenance
- Backup and Restore
- Administration
- Utility Scheduling
- Tuning and Operation
- Storage Administration



Everyday tasks are simply automatic!

- **Statistics Collection**
- **Backup**
- **Table Reorganization**

Status as of:	4/14/04 7:34 AM	Refresh
DBM State:	Started	Stop
Last Backup:	4/13/04 9:00 AM	Backup Database
Size:	19 MB	Manage Storage
Capacity:	5316 MB	
	<input type="text" value="1%"/>	
Health:	Normal	Monitor DB Health
Maintenance:	Fully automated	Maintenance

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

It's Automatic!

Online maintenance window

Online automatic maintenance can occur during the following window

Time	00:00 - 05:00 (5 hours)
Days of the week	All
Days of the month	All
Activities using this window	Backup database (BACKUP), Optimize data access (RUNSTATS)

[Change...](#)

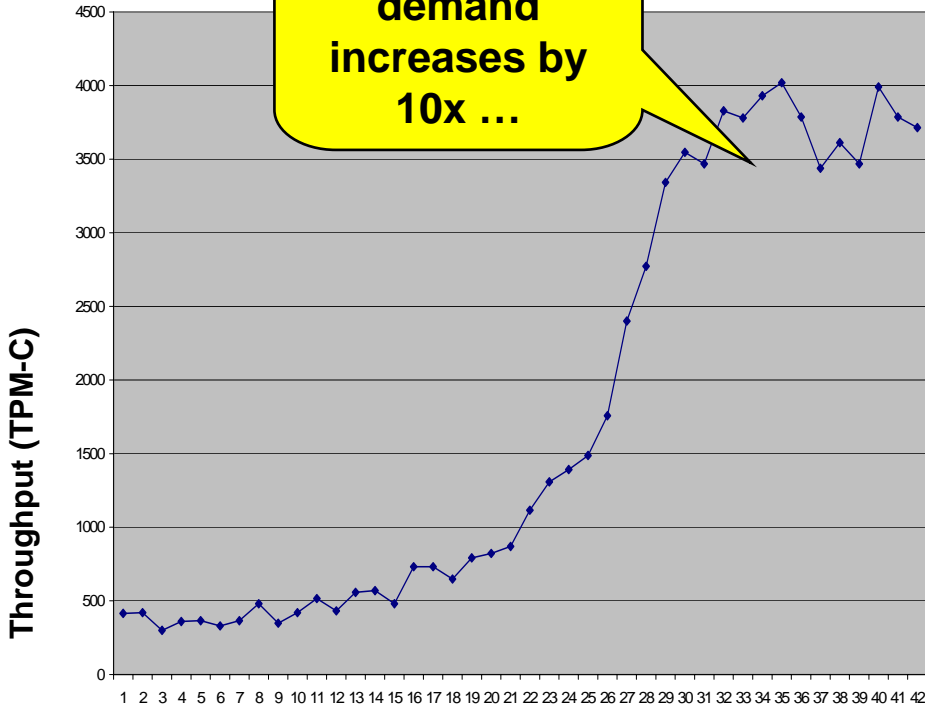
10:00
11:00
12:00
1:00
2:00

Adaptive Self Tuning Memory

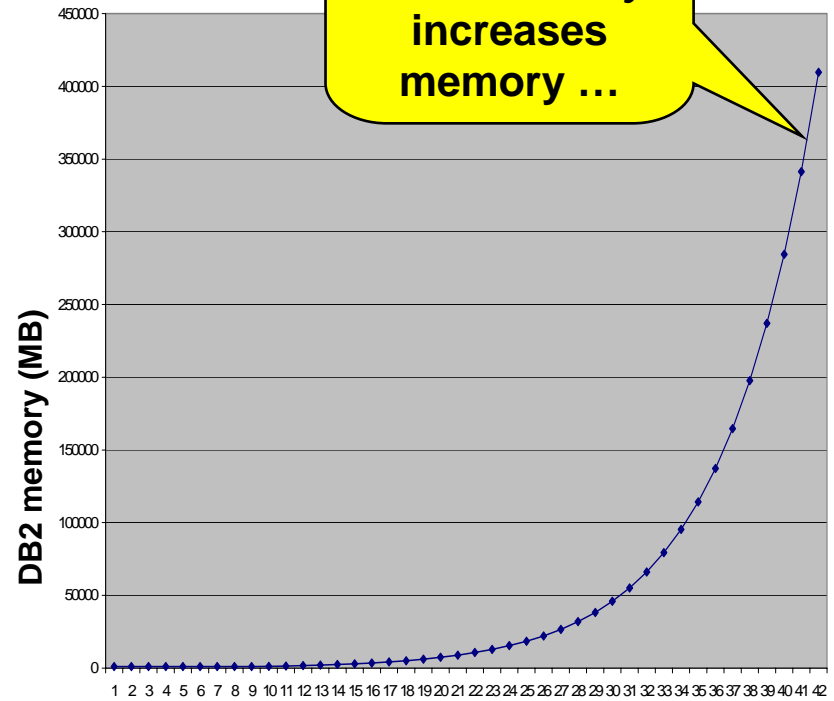


Increases Business Value, Decreases DBA Tuning Tasks

As workload demand increases by 10x ...



...DB2 automatically increases memory ...

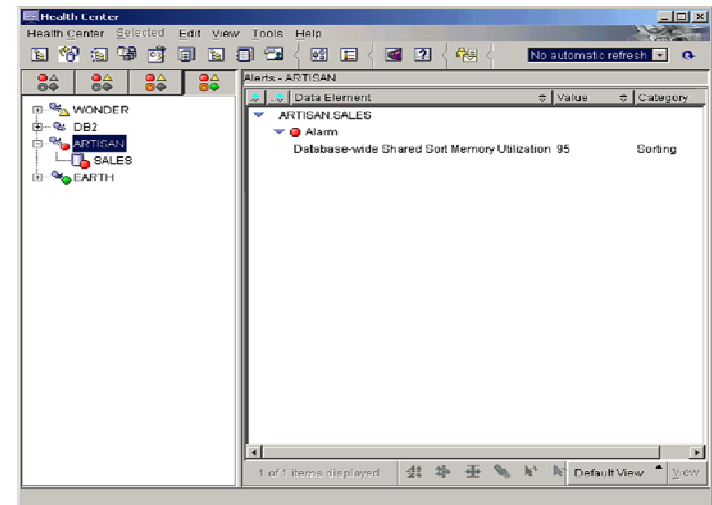


Time (10s intervals)

Time (10s intervals)

1. Show Health Center

- Show using the health center to determine the status of one of the database systems.
- Show alerts and Recommendation Advisor.
- Show how user can customize settings for alerts.
- Show how alerts can be set to go to e-mail.



- **In Control Center, show other options for Automatic Memory and Space Management.**



Provides Fine-Grain Access

- **Label Based Access Control (LBAC)**

- A “label” is associated with both user sessions and data rows
- Rules for comparing users and data labels provide access control at the row level

- **Labels may consist of multiple components**

- Hierarchical, group or tree types
- Row labels appear as a single additional column in a protected table.
- User labels are granted by a security administrator



- **Limit access based on job function**

- **Avoid maintaining separate databases**



Availability: Oracle is Plagued by Patches

Frequent security patches *increase* planned downtime



COMPUTERWORLD Security

JUMP TO SEARCH **GO**

- Home
- News
- E-mail Newsletters
- + Shark Bait
- Knowledge Centers
 - + Business Intelligence
 - + Careers
 - + Development
 - + Government
 - + Hardware
 - + IT Management
 - + Mobile/Wireless
 - + Networking

Oracle has 65 fixes in latest security update

COMPUTERWORLD Software

JUMP TO SEARCH

COMPUTERWORLD Business Intelligence

JUMP TO SEARCH **GO**

- Home
- News
- E-mail Newsletters
- + Shark Bait
- Knowledge Centers
 - Business Intelligence
 - Databases
 - Data Mining
 - Data Warehousing
 - + Careers
 - + Development
 - + Government
 - + Hardware
 - + IT Management
 - + Mobile/Wireless
 - + Networking

Oracle releases 101 patches in quarterly update

Oracle releases 51 patches in quarterly security update

26 are for database products

Jaikumar Vijayan [Today's Top Stories](#) > or [Other Business Intelligence Stories](#) >

Comments (0) Recommendations: 5 — [Recommend this article](#)

January 16, 2007 (Computerworld) -- As expected, Oracle Corp. today released 51 new security fixes for flaws across its database and application server products as well as its collaboration software and e-business suites.

Of these, 26 fixes addressed flaws in the company's database products, including 10 that the company said could be remotely exploited without the need for a username or a password. Oracle typically assigns its highest severity ratings to such flaws.

Today's Critical Patch Update (CPU) from Oracle also contained 12 fixes for vulnerabilities in Oracle's Application Server software, eight of which were rated "critical" because they can be remotely exploited without any user authentication. Also included in the update were three patches for holes -- including one that could be remotely

MORE RELATED CONTENT

- Oracle will give early notice of security updates
- Bug hunter cancels plan to publicize Oracle Database flaws
- Oracle Tops Century Mark With Latest Patch Release

TOP STORIES
y vendors:



- **Better XML support**
- **Better Compression**
- **Better Performance**
- **Built-in Automation for Easier Configuration and Management**
- **Better Security**
- **More Scalable Architecture**
- **Higher Availability**
- **Lower Total Cost of Ownership**
 - Not only is software licensing and maintenance less expensive, but hardware requirements are less and performance is better.



Challenge

- Draw additional value from existing database and storage investments
- Increase productivity of database administrators
- Implement a platform that manages risk and compliance

Solution

- Optimized data platform by upgrading to IBM DB2® 9 data server
- Took advantage of the deep compression capabilities of DB2 9, as well as features that improve performance
- Installed Label Based Access Control feature of DB2 9 to manage risk and compliance



Business Benefits

- Expects an **initial cost savings of US\$2 million with additional savings of US\$500,000 each year**
- Reduces storage requirements by increasing compression rate to 83 percent on the data warehouse tables
- **Improves system performance, yielding faster query responses**
- Protects data and nets additional savings with improved security

धन्यवाद

Hindi

多謝

Traditional Chinese

ขอบคุน

Thai

Спасибо

Russian

Gracias

Spanish

شكراً

Arabic

Thank You

Obrigado

Brazilian Portuguese

Grazie

Italian

Danke

German

Merci

French

நன்றி

Tamil

多谢

Simplified Chinese

감사합니다

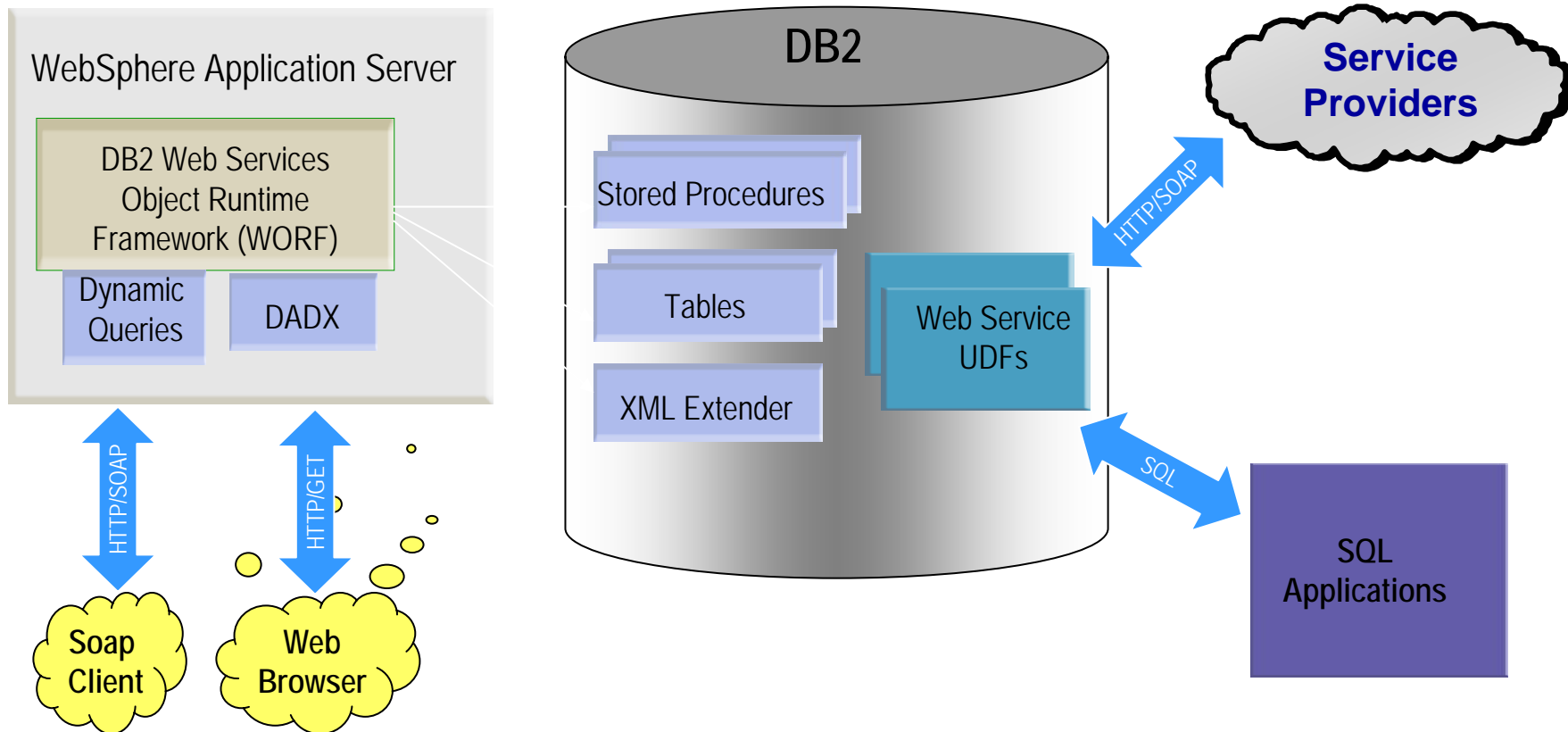
Korean

ありがとうございました

Japanese



DB2 Web Services Object Runtime Framework (WORF)



DB2 provides Web Services data

DB2 consumes Web Services data