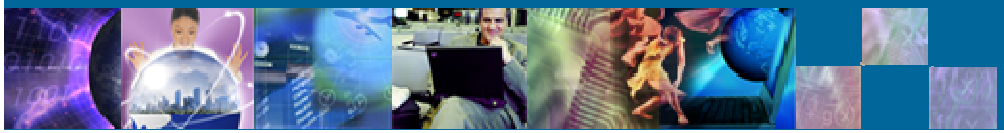


pureXML en DB2 9

ON DEMAND BUSINESS™

What is XML?

- ***XML***
  - eXtensible Markup Language
  - Self-describing data structures
  - XML tags describe each element and their attributes
- ***Benefits***
  - Flexible
    - No fixed format or syntax
    - Structures can be easily changed
  - Platform Independent
    - Not tied to any platform, operating system, language or software vendor
    - XML can be easily exchanged
  - Fully Unicode compliant

```
<? xml version="1.0" ?>
<purchaseOrder id="12345" >
    <customer id="A6789">
        <name>John Smith Co</name>
        <address>
            <street>1234 W. Main</street>
            <city>Toledo</city>
            <state>OH</state>
            <zip>95141</zip>
        </address>
    </customer>
    ...
```

XML Solves Business Problems Today

- ***Business to Business Integration***
  - Platform independent transport mechanism

    Purchase order triggers transactions flowing over a service oriented architecture

- ***Forms and Document Processing***
  - Government and legal industry require digital signature

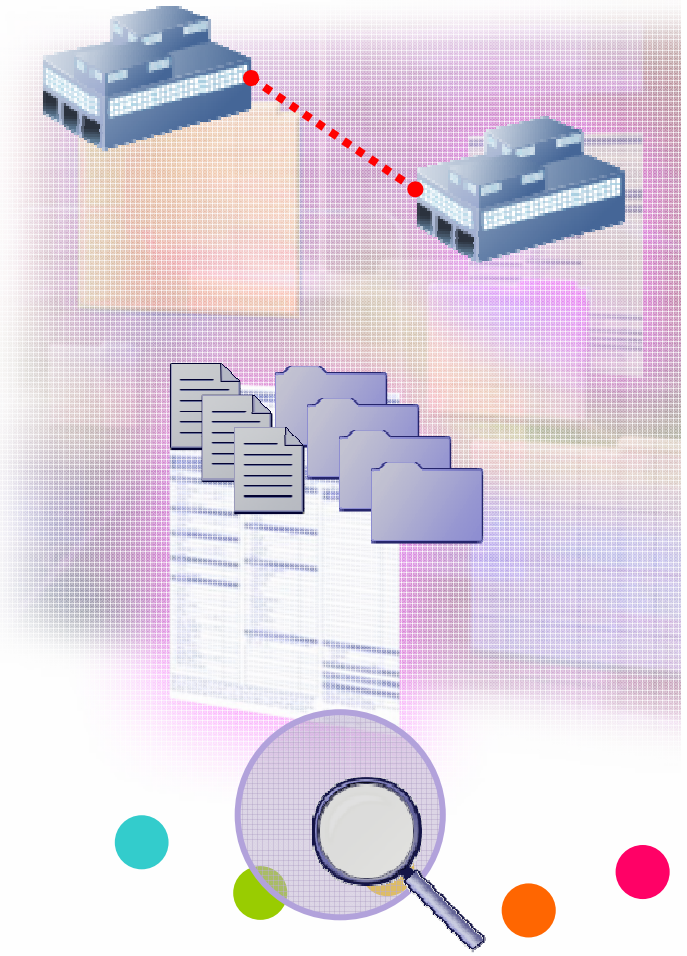
    Tax forms require signature & change year to year

  - Documents often contain sub-documents

    Literary materials contain books, chapters, and sub-chapters

- ***Business Insight***
  - Universal representation from multiple sources

    Claims adjustor reviews damage estimates from multiple garages with consideration of original format



Importance of XML?

- ***More XML data generated than in databases***
- ***Growing at 2X rate of total database market***
  [IDC]
- ***XML is pervasive in all kinds of organizations***
- ***Almost every sector has XML based standards***

XML is the Language of Business

- **Banking and Financial Markets**
  - **IFX** - Interactive Financial Exchange – Trades, banking, consumer transactions, etc
    http://www.ifxforum.org
    .
  - **XBRL** – Extensible Business Reporting Languaje - Communication of business and financial data
    http://www.xbrl.org
- **Insurance**
  - **ACORD** – Policy management, underwriting, indemnity, claims, etc.
    http://www.acord.org
- **Health Care**
  - **HL7** – Patient Management – Diagnosis, treatments, prescriptions, etc.
    http://www.hl7.org
- **Retail**
  - **IXRetail** – Inventory, customer transaction and employee management
    http://www.nrf-arts.org/
- **Cross-industry, Business-to-Business**
  - **SWIFT** – B2B Transaction, invoices, purchase orders, inventory status, etc.
  - http://www.swift.com

XML Example: Financial Data (FIXML)

- Buying 1000 Shares of IBM Stock..

```
8=FIX.4.2^9=251^35=D^49=AFUNDMGR^56=ABROKER^34=2
^52=20030615-01:14:49^11=12345^1=111111^63=0^64=2003
0621^21=3^110=1000^111=50000^55=IBM^48=459200101^22=
1^54=1^60=2003061501:14:4938=5000^40=1^44=15.75^15=USD
^59=0^10=127
```

Old FIX Protocol

New FIXML Protocol

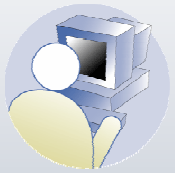
- extensible
- lower appl development & maintenance cost

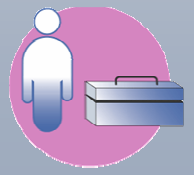
```
<FIXML>
   <NewOrdSingle  ClOrdID="123456"
     Side="2"
     TransactTm="2003-06-15T01:14:49-05:00"
     OrderType="2"
     Price="93.25"
     Acct="26522154">
      <Header  Sent="2001-06-21T01:31:28-05:00"
       PosDup="N"
       PosRsnd="N"
       SeqNum="521">
       <Sender  ID="AFUNDMGR"/>
       <Target  ID="ABROKER"/>
      </Header>
      <Instrument  Symbol="IBM"
            ID="459200101"
            IDSrc="1"/>
      <OrderQuantity  Qty="1000"  Cur="USD"/>
   </NewOrdSingle>
</FIXML>
```

XML Data Needs Relational Maturity

Complementing XML Processing

- ***XML Data Needs Protection***
  - Backup and recovery features to ensure continuity
  - Data is protected using database security
- ***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 it's 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





DB2 9 – A Pure XML, Relational Hybrid

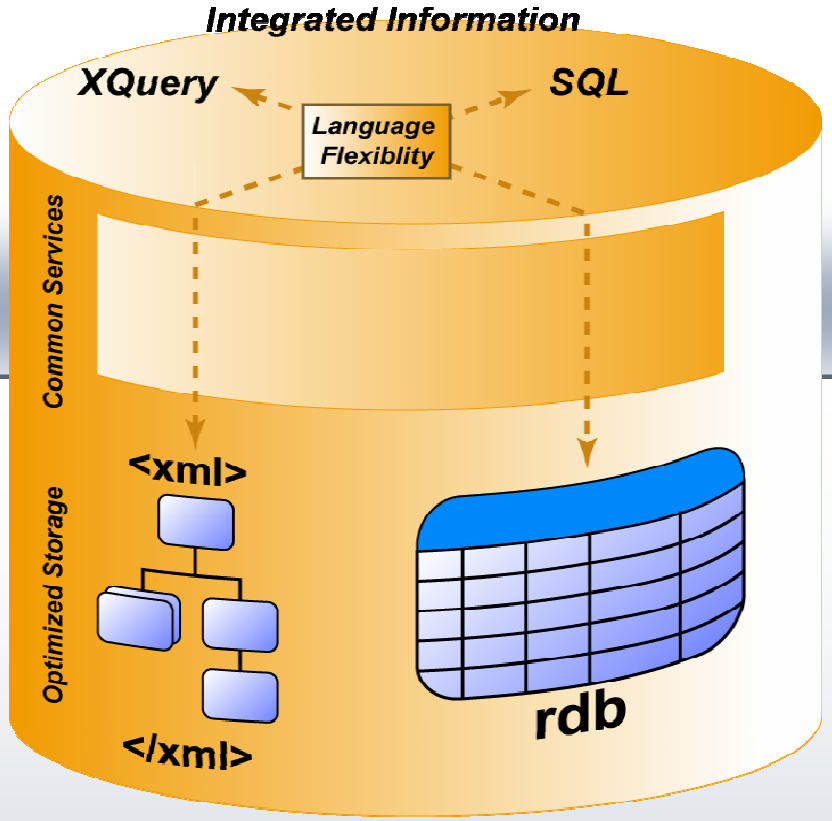
XML Developer

"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

Familiar Tooling

Optimized Storage Models

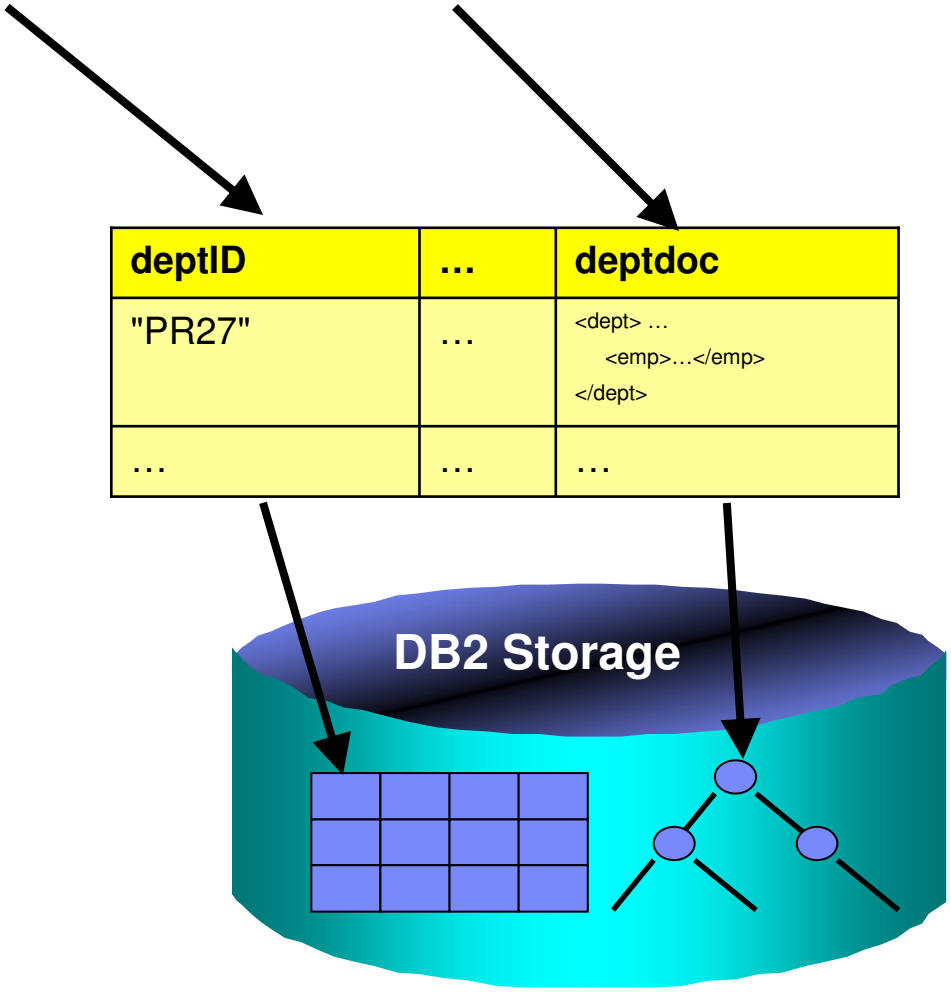
Optimized Performance & Scale

Native XML Storage

- DB2 stores XML in **parsed hierarchical** format (~DOM)

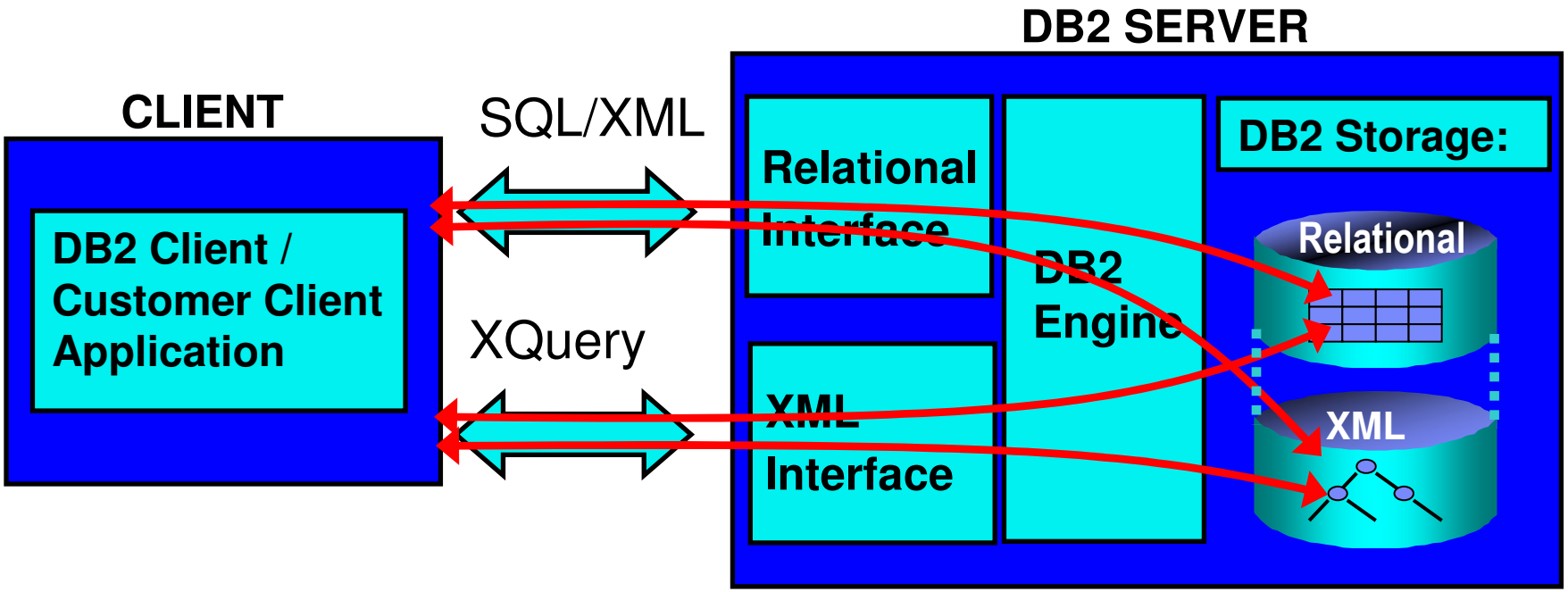
create table dept (deptID char(8),…, deptdoc xml);

- Relational columns are stored in relational format (tables)
- XML columns are stored **natively**
- **No XML parsing for query evaluation!**

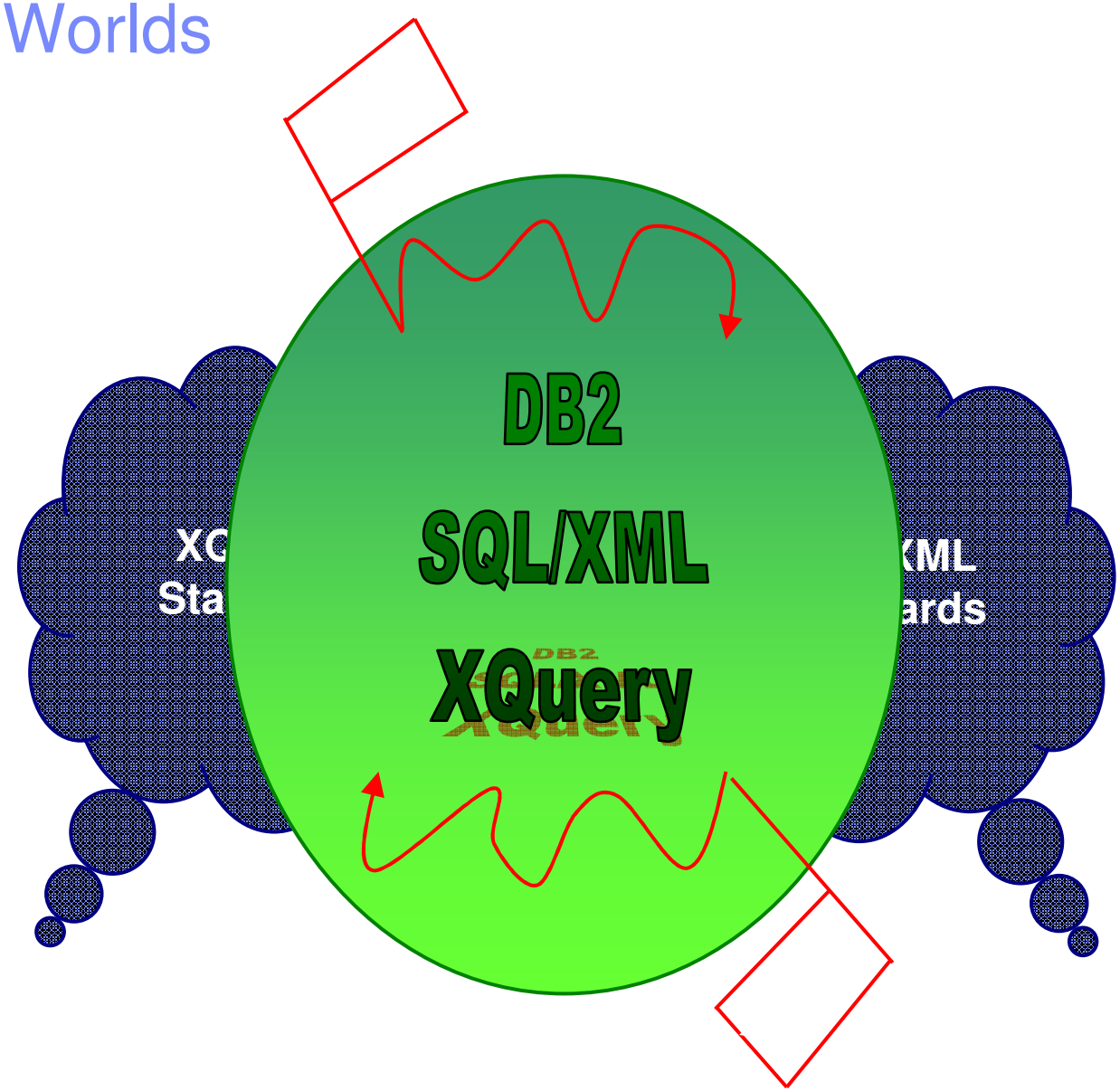


Integration of XML & Relational Capabilities

- Applications combine XML & relational data
- Native XML data type (server & client side)
- XML Capabilities in all DB2 components



Two Worlds



XQuery XPath Samples

Table "PEOPLE"
Column "INFO"

```
<dept bldg="101">
   <employee id="901">
      <name>John Doe</name>
         <phone>408 555 1212</phone>
         <office>344</office>
    </employee>
    <employee id="902">
         <name>Peter Pan</name>
         <phone>408 555 9918</phone>
         <office>216</office>
    </employee>
</dept>
```

```
XQUERY db2-fn:xmlcolumn ("PEOPLE.INFO")
        Returns all Documents in column INFO

XQUERY db2-fn:sqlquery ("select info from people")
        Returns all Documents in column INFO
```

XQUERY	Result
XQUERY db2-fn:xmlcolumn ("PEOPLE.INFO") /dept/employee[@id="902"]/name	<name>Peter Pan</name>
XQUERY db2-fn:xmlcolumn ("PEOPLE.INFO") /dept[@bldg="101"]/employee[office >"300"]/name	<name>John Doe</name>
XQUERY db2-fn:xmlcolumn ("PEOPLE.INFO") //employee[office="344" OR office="216"]/@id	901 902
XQUERY db2-fn:xmlcolumn ("PEOPLE.INFO") /dept/employee[2]/@id	902

New SQL/XML Functions in SQL 2006

- XMLPARSE
  - parses character/BLOB data, produces XML value
- XMLSERIALIZE
  - converts an XML value into character/BLOB data
- XMLVALIDATE
  - validates XML value against an XML schema and type-annotates the XML value
- XMLEXISTS
  - determines if an XQuery returns a results (i.e. a sequence of one or more items)
- XMLQUERY
  - executes an XQuery and returns the result sequence
- XMLTABLE
  - executes an XQuery, returns the result sequence as a relational table (if possible)
- XMLCAST
  - cast to or from an XML type

XML Indexes for High Query Performance

- Define 0, 1 or multiple XML Value Indexes per XML column
- XML index maps: (pathID, value) → (nodeID, rowID)
- Index any elements or attributes, incl. mixed content
- Index definition uses an XML pattern to specify which elements/attributes to index (and which not to)
- Can index all elements/attributes, but not forced to do so
- Can index repeating elements
  ⇒ 0 , 1 or multiple index entries per document
- New XML-specific join and query evaluation methods, evaluate multiple predicates concurrently with minimal index I/O

xmlpattern = XPath without predicates, only child axis (/) and descendent-or-self axis (//)

XML Full Text Search

- DB2 Net Search Extender enhanced for XML
- Full XML-aware index of entire doc
  - Create index myIndex for text
    on myTable (xmlcol);
  - Can also index partial documents
- Simple usage
  - select xmlcol from myTable where
    contains(xmlcol,‘sections("/book/section") "Brazil" ')= 1
  - Complex search criteria also supported
- Management is integrated in the DB2 Control Center

XML Schema Repository (XSR)

- Database needs a Schema repository:
  - Stable & high performance access to Schemas for validation at XML insert/update time
  - Support for XML Schema management
- XSR
  - XML Schemas are registered:
    - Consistent set of .xsd document
  - Also DTDs and External entities
    - Used for entity reference resolution and defaults
    - NOT used for validation
  - Registered Schema identification:
    - A SQL 2-part name
    - The URL the Schema is externally known as (e.g. used in schemaLocation attributes)
    - The "primary namespace"
  - Also used by Shred
    - Stores annotated Schema
    - Internal formats to make Shredding efficient.
  - Implementation:
    - All data is stored in the DB2 Catalog
    - Binary representation of the Schema for fast validation

Shredding into relational tables

- There are still reasons to shred XML.
  - Co-existence with legacy applications
  - Relational processing is faster than XML
  - Analytics/cubes work over non-XML data
- Mapping from XML to relational:
  - Annotate the XML schema
  - Register XML schemas in the schema repository
  - Shred via CLP commands or stored procedure calls

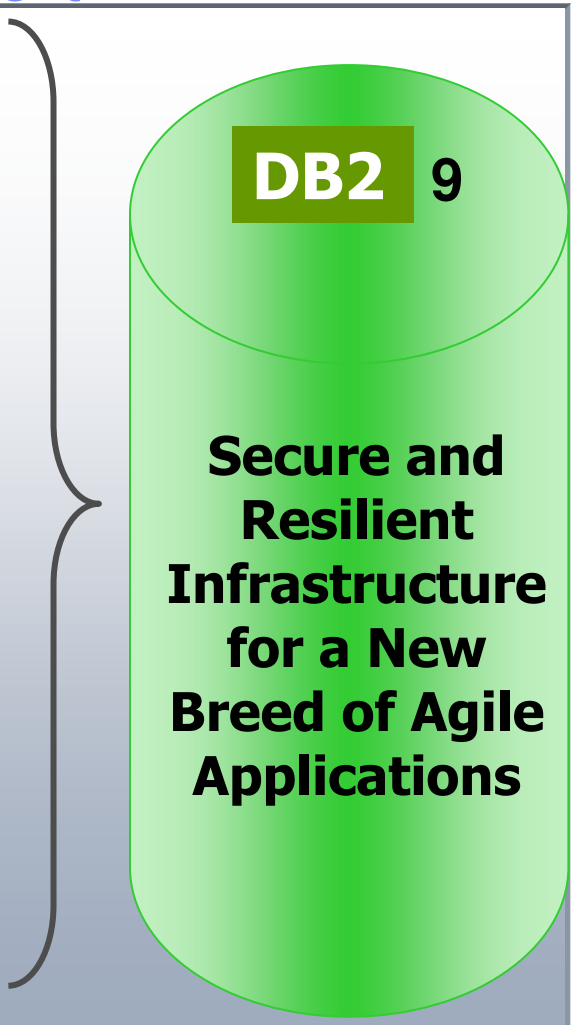
Annotation Example:

```
<xsd:element name="phone" type="xsd:string"
             db2-xdb:rowSet="employee_tab"
             db2-xdb:column="phone_col"/>
```

- Replaces XML Extender shred (XML collection)
  - Faster; using XML Schema

DB2 9 – Summary of pureXML 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



Business Benefits of DB2 9 with pureXML technology

- ***Lower Development Costs***
  - Reduced code and development complexity
  - Improved developer productivity

  → *Quicken solution development and gain cost savings*

- ***Greater Business Agility***
  - Easily accommodate changes to data and schemas
  - Update applications rapidly and reduce maintenance costs

  → *Respond quickly to dynamic conditions and get faster time to value*

- ***Improved Business Insight***
  - Access to “hidden gems” (data) in unexploited documents
  - Unprecedented application performance

  → *Gain competitive advantage through better and quicker information*

Reduce Code Complexity with DB2 pureXML

Make Changes Easily with DB2 pureXML

```
<DEPARTMENT  deptid="15" deptname="Sales">
      <EMPLOYEE>
         <EMPNO>10</EMPNO>
         <FIRSTNAME>CHRISTINE</FIRSTNAME>
         <LASTNAME>SMITH</LASTNAME>
         <PHONE>408-463-4963</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
[illegible]	415-010-1234
[illegible]	408-463-4963

Department

DEPTID	DEPTNAME
15	Sales

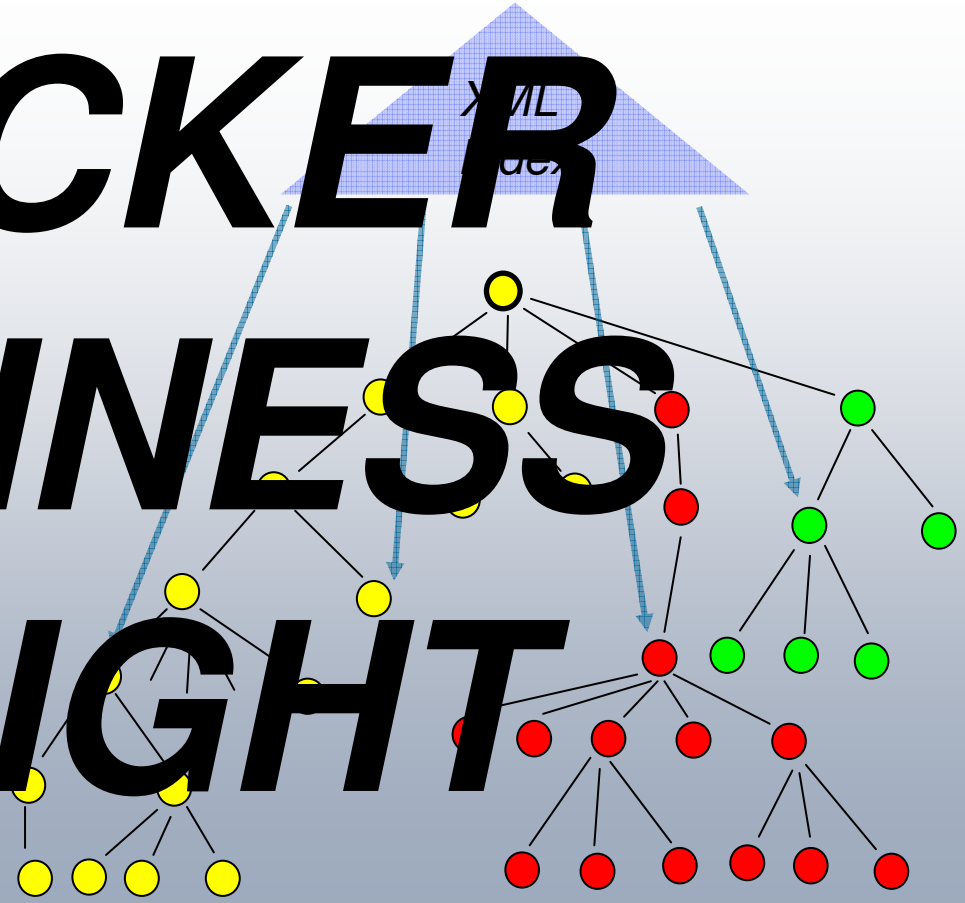
Costly!

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 pureXML

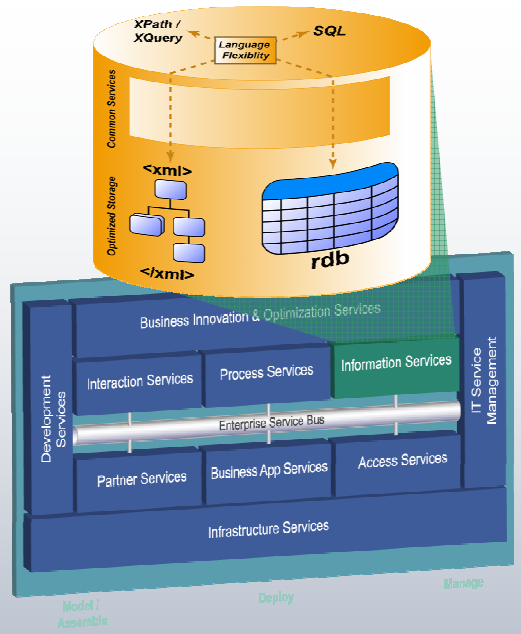
ID	
123	<? xml version="1.0" ?><purchaseOrder id="123"><customer id="A6789"><name>John Smith Co</name><address> <street>1234 Main</street> <city>Toledo</city> <state>OH</state> <zip>95141</zip> </address> </customer>
456	…



QUICKER BUSINESS INSIGHT

DB2 9: Early adopter feedback

storebrand



Proto-type results using DB2 9 based SOA solution

Task	with relational server	with DB2 9 pureXML
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

Industry Experts on DB2 pureXML 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
- *"… enable users to work with both types of data via SQL or Xquery requests. It could also open up the the world of database applications to developers and ISVs with experience in other data types and applications."*
  - Barbara Darrow, CRN

Words from DB2 9 Beta Customers and Partners

"Our development time using Viper's native XML store is a radical improvement over existing XML 'shred' technology. We are now able to make schema changes in minutes rather than days and will dramatically improve our customer response time."

- Thore Thomassen, Senior Enterprise Architect for Storebrand

"We wanted to be able to support queries that just were based on information in the e-records that had not been indexed. The way we have to do that at the moment is not terribly efficient … [Native XML support in DB2 Viper] is going to enable us to store things more compactly and access them easier ... and make it easy for us to be able to ingest and then export data in XML when we're able to migrate to that version of DB2."

- Dave Richards, Chief Technology Officer for The Research Libraries Group Inc.

"We are delighted to be partnering with IBM to integrate our platform with DB2 Viper. The combination of industrial strength database management for native XML by DB2 Viper and Skytide's ability to provide direct multidimensional analysis of XML data, removes two key barriers to widespread adoption of XML and the transformation of this data into actionable business information."

- Joseph Rozenfeld, Vice President of Products for Skytide

"The upcoming release of DB2, offers leading-edge technology for storing, managing and searching XML data in a secure, highly scalable environment. The new multi-structure hybrid architecture offered by DB2 combines the best of relational database management technology with the best of XML data management."

-Tim Harvey, CEO of XAware.

Selected Partners enabling pureXML based solutions with DB2 9

NEXTANCE®

SKYTIDE

BeaconIT
Empower Your Data

JustSystem

exegenix®

mdxsys
xml productivity tools

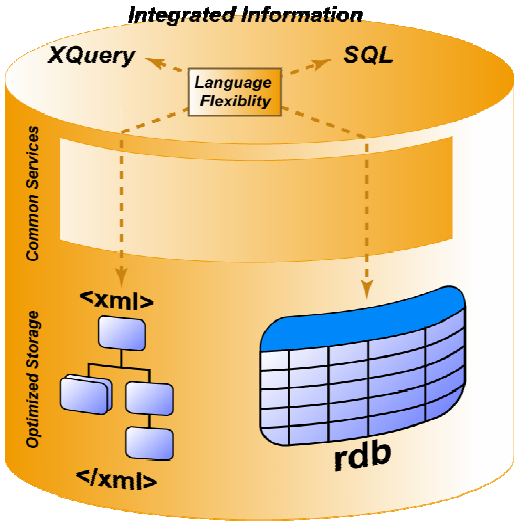
xfy technology

kbrix
xml reporting

strada

TEMENOS™

Napersoft™



EMBARCADERO TECHNOLOGIES

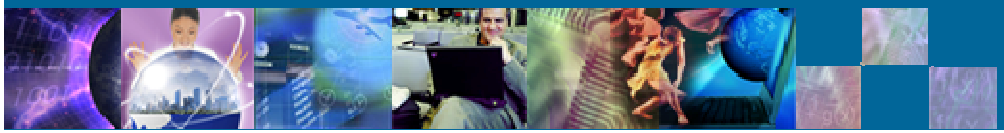
AdminServer
Policy Administration without limitations.

XAware™

DIS

BlueWare
Strategic HealthWare Solutions

ibm.com/db2/xml



DB2
Label-Based Access Control

ON DEMAND BUSINESS™

Introducción

- Label Based Access Control (LBAC), en general, se refiere a la posibilidad de controlar el acceso a nivel de registro basándose en etiquetas de seguridad
- DB2 LBAC => Seguridad a nivel de registro (fila) y a nivel de campo (columna)
- Incluye la posibilidad de controlar el acceso a nivel de tabla
- Compatible con DB2 Discretionary Access Control (DAC)
- Similar al implementado por DB2 zOS

Etiquetas de Seguridad

- Tipos de etiquetas
  - Usuario: Etiqueta asociada a un usuario
  - Fila: Etiqueta asociada a un registro determinado
  - Columna: Etiqueta asociada a una columna de una tabla
- Componentes
  - Set (conjunto): ej. Proyectos = (‘Desarrollo’, ‘Implantación’, ‘Consultoría’)
  - Array (lista ordenada): ej: Nivel = (Confidencial, Interno, Público)
  - Tree (árbol): ej: Departamentos = {D1, D2, D3}, donde D1 es elemento raíz y D2 y D3 son “hijos” de D1

```
CREATE SECURITY LABEL COMPONENT Nivel
ARRAY [‘Confidencial', ‘Interno', ‘Público']
```

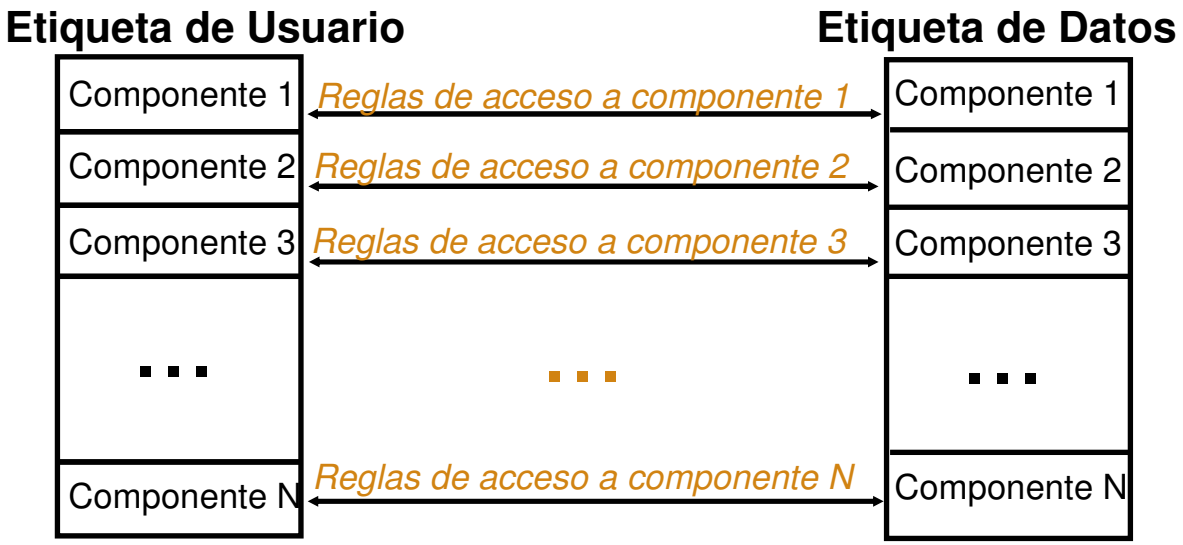
Reglas de Acceso

- Conjunto de reglas predefinidas:
  - Reglas de acceso de lectura: aplicables en sentencias SQL SELECT, UPDATE y DELETE
  - Reglas de acceso de escritura: aplicables en sentencias SQL INSERT, UPDATE y DELETE

Políticas de Seguridad

- Conjunto de Etiquetas de seguridad con uno o varios componentes y reglas de acceso asociadas

- Cuando se accede a un objeto protegido por una política de seguridad, DB2 aplica las reglas de acceso adecuadas, basadas en el tipo de acceso y de componente



Exenciones

- Privilegio especial que permite a un usuario obviar una regla determinada de una política de seguridad

- El SECADM puede conceder una exención de cualquier combinación de reglas de acceso

GRANT EXEMPTION ON RULE DB2LBACWRITEARRAY WRITEDOWN

FOR *PolíticaSeg1* ***TO USER*** *Fernandez*

Administrador de Seguridad (SECADM)

- Autoridad necesaria para:
  - Crear y eliminar etiquetas de seguridad
  - Crear y eliminar componentes de etiquetas de seguridad
  - Crear y eliminar políticas de seguridad
  - Asignar (Grant) y revocar (Revoke) etiquetas de seguridad
  - Asignar (Grant) y revocar (Revoke) exenciones
- Sólo SYSADM puede asignar SECADM a un usuario
- SECADM no puede acceder a los datos de una tabla protegida por defecto
- SECADM no puede concederse permisos a sí mismo

Ventajas principales de DB2 LUW LBAC

- Etiquetas de Columna
  - Independientes de las etiquetas de fila
  - Permiten control de acceso a columnas de una tabla
- Etiquetas de Tabla
  - Mejora rendimiento y ahorro espacio
  - Protección de los datos incluso del SYSADM
- Definición de Políticas de acceso flexible
  - Posibilidad de definir políticas ‘a medida’
  - Posibilidad de utilizar tres tipos de componentes predefinidos (set, array, tree)
  - Una etiqueta puede tener cualquier número de componentes de cualquier tipo. (No fuerza a introducir jerarquías).
  - DB2 LBAC selecciona y aplica únicamente las reglas relevantes, basándose en la política de seguridad (No aplicación sistemática de reglas de lectura y escritura)

Ventajas de DB2 LUW LBAC (cont.)

- Flexibilidad de Asignación de Etiquetas a Usuarios
  - Etiquetas sólo para lectura
  - Etiquetas sólo para escritura
  - Etiquetas para lectura y escritura
- Granularidad en el acceso
  - Los Administradores de Seguridad pueden proteger distintas tablas con distintas políticas de seguridad dentro de la misma base de datos
- Codificación de las Etiquetas de Seguridad
  - Los usuarios gestionan las etiquetas en formato legible, pero DB2 LBAC almacena las etiquetas en formato interno eficiente

Más información…

- Manuales de DB2 9
- http://www.ibm.com/**developerworks**/