THE FILLMORE GROUP
Relational Database Solutions

IBM Data Interoperability Tools

For Oracle Users

Premier Business Partner IBM

THE FILLMORE GROUP
Relational Database Solutions

The Fillmore Group, Inc.

- Frank Fillmore, DB2 Gold Consultant, IBM Champion
- DB2 Technical Support and Consulting
- IBM Training Partner
- IBM Information Management Software Reseller

IBM CHAMPION

THE FILLMORE GROUP
Relational Database Solutions

Services:

- IBM Information Management consulting
- Information Integration – replication and federation
- Big Data and Netezza
- Optim Data Governance

Premier Business Partner IBM

THE FILLMORE GROUP
Relational Database Solutions

Agenda:

- DB2 vs. Oracle – Market Strategy
- IBM's Acquisitions
- What's in the IBM Data Interoperability Toolbox?
- What do the replication and federation tools do??
  - IBM InfoSphere Data Replication – Q Rep and ICDC
  - IBM InfoSphere Federation Server
  - IBM InfoSphere Data Explorer
- Next steps

Premier Business Partner IBM

THE FILLMORE GROUP
Relational Database Solutions

3 Takeaways:

1. An understanding of how data interoperability tools can help maximize the value of your data.

2. What replication and federation data interoperability solutions are available from IBM

3. What to do next if you are interested in learning more about IBM's solutions

Premier Business Partner IBM

THE FILLMORE GROUP
Relational Database Solutions

DB2 vs. Oracle - the history

IBM vs. Oracle, Mobility, Cloud: What IT Pros Need to Know for 2011

Big Blue vs. Big Red

That would be IBM versus Oracle. The successful strategy of IBM built on services versus Oracle's strategy of buying up the enterprise software universe is a major act for business-to-business in 2011. CIOs know the value of buying a fully integrated software offering in terms of compatibility but are wary of getting tied to one vendor. The rest of the B2B tech vendor community has to find a way to outmaneuver IBM and Oracle. This will be fun to watch on the outside, less fun on the inside if you make the wrong choice.

IBM

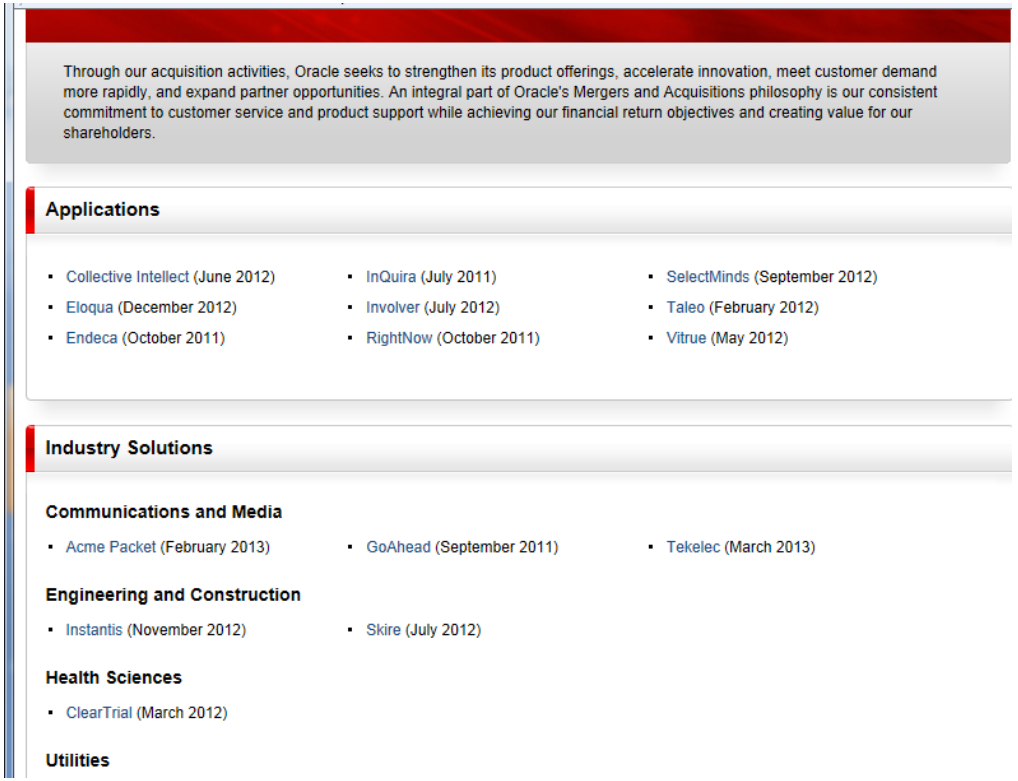
VS

ORACLE

DB2 is the highest value database in the market.

1. **DB2 has one of the lowest people cost requirements.**
   - ✓ *DB2 requires 43% less DBA time vs. Oracle, Solitaire Interglobal analysis; average cost per DBA saved: $100 /hr x 40 hrs/week x 52 weeks/yr = $208,000/yr burdened. If you save 5 DBA's, That's $1M / yr*
2. **DB2 is the smartest database in the market.**
   - ✓ *DB2 Autonomics; self-optimizing, self-healing, self-configuring and self-protecting, DB2 has the ability to monitor its own health status - independent of the database administrator (DBA).*
3. **DB2 reduces storage costs better than any other vendor.**
   - ✓ *"With DB2 9, we're seeing compression rates up to 83 percent on the data warehouse tables. The projected cost savings are more than US$2 million initially with ongoing savings of US$500,000 a year." - Michael Henson, DB2 Unix Team Lead, Data-base Delivery Services, SunTrust Bank, Inc.*
4. **DB2 is the fastest database.**
   - ✓ *DB2 at largest SAP/Oracle customer in Germany*
     - *DB2 up to 9 times faster*
     - *DB2 on average over 40% faster*
5. **DB2 has the highest availability.**
   - ✓ *"DB2 was heavily favored in this regard... with fewer and shorter outages for normal operational activities, the overall availability and reliability of the DBMS shows some clear differentiation." — Solitaire Interglobal*
6. **DB2 is the most scalable database.**
   - ✓ *DB2 pureScale reduces the risk and cost of business growth by providing unlimited capacity, continuous availability, and application transparency. Scaling your system is simply a matter of connecting a new node and issuing two simple commands. Oracle RAC scalability tails off dramatically after 4 nodes, DB2 purescale performance stays at 95% scalability at 32 nodes, dropping only slightly as you add more nodes (81% at 112 nodes)*

Oracle has historically focused on acquiring specialty application developers and industry solutions.



Through our acquisition activities, Oracle seeks to strengthen its product offerings, accelerate innovation, meet customer demand more rapidly, and expand partner opportunities. An integral part of Oracle's Mergers and Acquisitions philosophy is our consistent commitment to customer service and product support while achieving our financial return objectives and creating value for our shareholders.

Applications

- Collective Intellect (June 2012)
- Eloqua (December 2012)
- Endeca (October 2011)
- InQuira (July 2011)
- Involver (July 2012)
- RightNow (October 2011)
- SelectMinds (September 2012)
- Taleo (February 2012)
- Vitrue (May 2012)

Industry Solutions

Communications and Media

- Acme Packet (February 2013)
- GoAhead (September 2011)
- Tekelec (March 2013)

Engineering and Construction

- Instantis (November 2012)
- Skire (July 2012)

Health Sciences

- ClearTrial (March 2012)

Utilities

While IBM Has Invested in the "Toolbox"

$14B invested in acquisitions and development over the past 3 years
25,000 developers WW
$20B additional investment in acquisitions and development over the next 3 years

Trusted Information Platform

Business Analytics & Optimization Platform

Business Analytics & Optimization Solutions

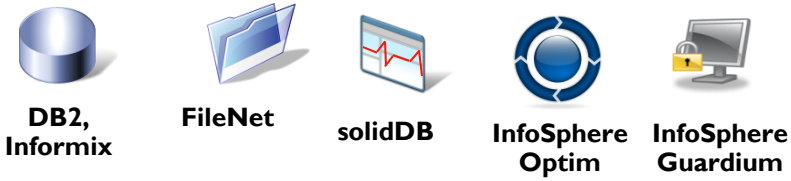
NETEZZA an IBM Company
Unica
SPSS
EXEROS
princeton softech
COGNOS
solid.
FileNet
DataMirror
dwl
iphrase
Ascential
SRD
unicorn

What Does This Mean to YOU?

- Oracle offers a broader range of off-the-shelf application options.

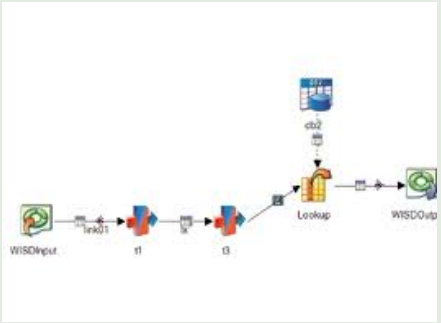
- IBM offers a broader range of tools to support in-house application development.

IBM's Information Integration



Organizations Use Integration for...

DataStage, QualityStage

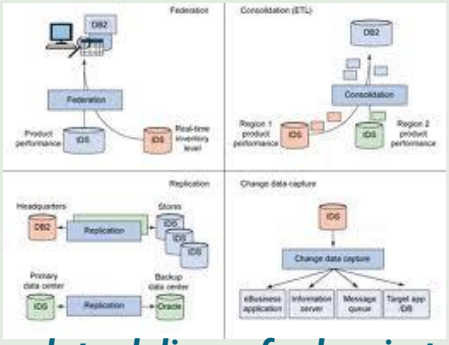


...data loading, cleansing, migrating, movement, high availability and high performance

Foundation Tools

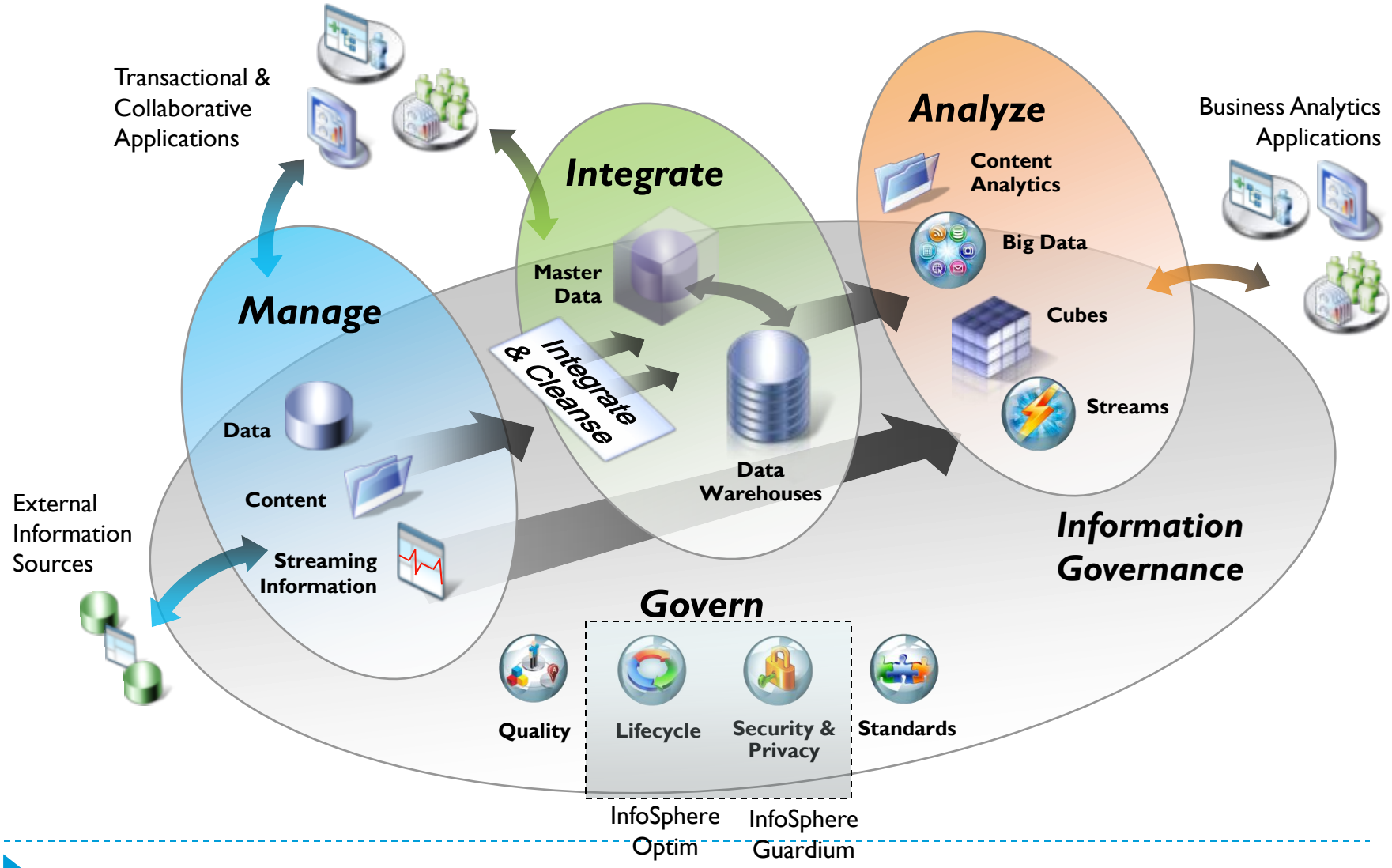
...deep analysis, design, metadata, collaboration and data governance

CDC, Federation, Replication



...data delivery for low impact, timely access to critical business data and operational data requirements

The "Information Supply Chain"



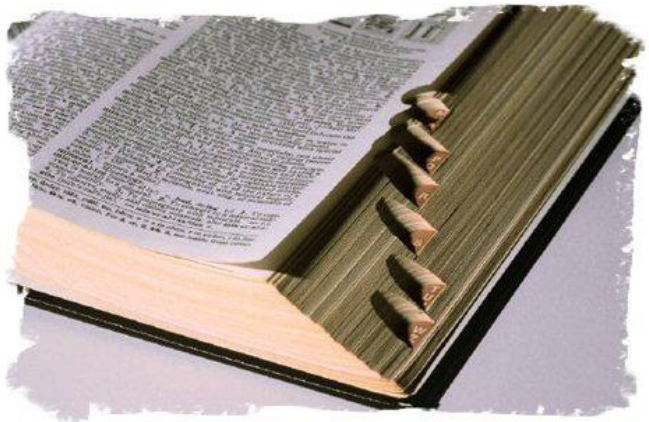
Poll Question #1

IBM Information Management

THE FILLMORE GROUP
Relational Database Solutions

What is Data Interoperability?

The ability of IT systems to be able to provide data and services to each other. This ability to provide data and services needs to be considered at the levels of business processes, information, applications and technology.



How do we enable data interoperability?

- Understand how data is defined and stored in one database versus another.
- Understand the potential need for transformation and translation, so data will appear native when it moves from the source database to the target database.
- Understand the need for data governance standards to control the quality of data.
- Ensure data security policies are consistent and enforced.

Deliver data to where it's most valuable, when it's needed

IBM Information Management Product Families

Information Management = DB2, Informix, solidDB

InfoSphere = Integration and Data Interoperability
Replication, Federation and ETL

Optim = Data Governance

IBM InfoSphere Replication

Product name: IBM InfoSphere Data Replication (IIDR)
Combines two heritage products:

1. Queue Replication (Q Rep)
2. InfoSphere Change Data Capture (ICDC/DataMirror)

THE FILLMORE GROUP
Relational Database Solutions

Reasons to Implement Replication

- Maintain production system availability 24X7
- Disaster Recovery Hot Site
- Database migrations – vendor, version-to-version, platform
- Loading a data warehouse
- Load balancing and facilitating extended use
- Integrating data from disparate systems

Premier Business Partner IBM

THE FILLMORE GROUP
Relational Database Solutions

What exactly IS replication?

From the Information Management Glossary:

The process of copying a portion of a database from one environment to another and keeping the subsequent copies of the data in sync with the original source. Changes made to the original source are propagated to the copies of the data in other environments.

IBM's Replication Options – How We Got Here

- **SQL Replication** – released in the mid-1990's as DataPropagator Relational (still around in DB2 for no charge)
- **Queue Replication** - released along with WebSphere Information Integrator V8.2 in 2005, incorporating MQ and SQL Replication to deliver high-speed replication
- **DataMirror/ICDC** – IBM acquired DataMirror in 2007 and rebranded DataMirror Transformation Server as InfoSphere Change Data Capture (ICDC) in 2008.

THE FILLMORE GROUP
Relational Database Solutions

How to chose which is best for you

- Understand the basic replication terminology
- Evaluate and prioritize your replication goals
- Understand each solution's strengths
- Apply your product understanding to your priorities
- Consult IBM and Business Partner experts

Premier Business Partner IBM

Terminology

- **Latency**
  - The time it takes for data to get from one point to another
  - Synonymous with delay; measured in microseconds, seconds, minutes
- **Source and Target databases**
  - The "source" database on which data is initially stored and the "target" is the database to which data is replicated
- **Capture and Apply**
  - Terms used to describe the process of collecting and delivering changed data
- **Transport**
  - The methodology used to move the data from source to target
- **Uni- directional, bi-directional and peer-to-peer**
  - Description of data movement from either source to target (uni) or from both source to target and target to source (bi), or to n-tier

Evaluate your replication goals – consider:

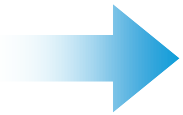
- Source and target databases
  - Are the source and target homogeneous or heterogeneous?
- Uni-directional, bi-direction or peer-to-peer replication?
- Speed and latency
  - Is there an SLA for replication speed?
  - What is the data volume on the source and target databases?
    - GB/TB
  - What is your anticipated transaction volume?
    - INSERTS/UPDATES/DELETES

Evaluate your replication goals – continued:

- Resiliency
  - Are outages acceptable? How often? For how long?
  - Do you need automated failover in the event of an unplanned outage?
- Ease of use
  - Monitoring, installation, administration
  - How frequently do you change your data model?
  - How complex is conflict resolution? Is it based on:
    - Value/Source/Timestamp/Application Logic
- Cost
  - All IBM replication solutions have the same cost.

Q Replication

- Sources: DB2 for z/OS and LUW, and Oracle on all platforms
- Targets: DB2 for z and LUW, Oracle, Sybase, others with Federation



Can invoke Stored Procedures or publish XML

- Uni-direction, bi-directional and peer-to-peer replication supported
- Latency – 3X faster than SQL Replication; recommended for very large transaction volumes; uses MQ Series for speedy delivery
- Resiliency – works with HADR, Q Replication Dashboard monitoring
- Ease of use – ASNCLP scripting language

THE FILLMORE GROUP
Relational Database Solutions

InfoSphere Change Data Capture

•Sources: DB2 (all platforms), Oracle, SQL Server, Sybase, Informix
•Targets: DB2, Oracle, SQL Server, Sybase, Teradata, Netezza

And non-relational data targets such as DataStage and MQ

•Uni-direction and bi-directional replication supported

•Latency – faster than SQL Rep; vs. Q Replication scalability issues may impact performance at very high volumes

•Resiliency – works with HADR for DB2

•Ease of use – great GUI interface; CHCCLP nascent equivalent scripting to ASNCLP in v10

THE FILLMORE GROUP
Relational Database Solutions

IBM Replication Pricing

- IBM Processor Value Units for the source and target databases
- Current pricing (full retail) = $173/PVU

Contact Kim for more information

Poll Question #2

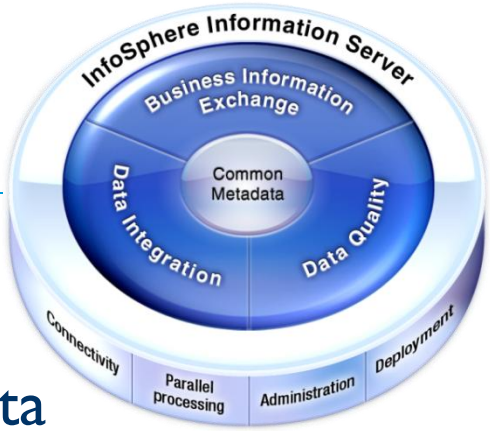
IBM Information Management

IBM Federation Solutions

Product name: IBM InfoSphere Federation Server

Product name: IBM InfoSphere Data Explorer

THE FILLMORE GROUP
Relational Database Solutions



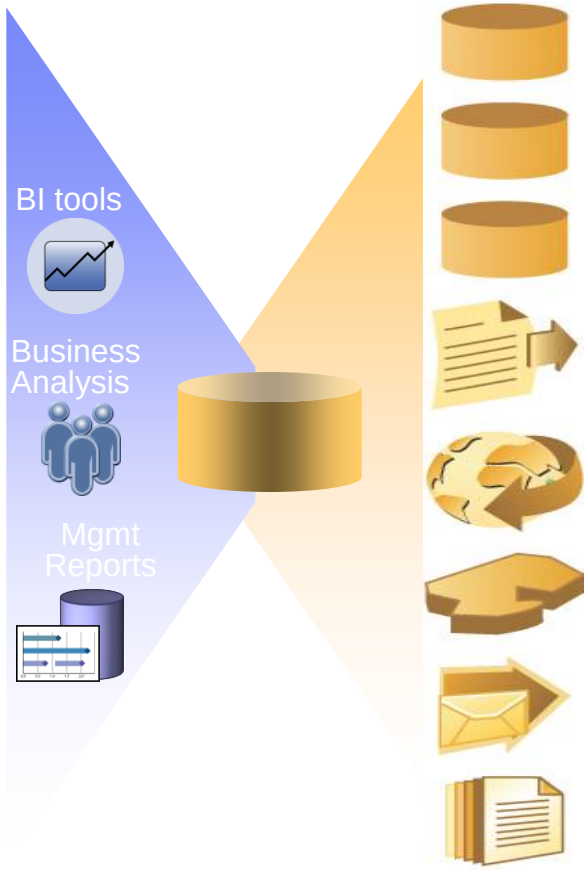
InfoSphere Federation Server

- InfoSphere Federation Server V10.1 offers data virtualization through a robust SQL interface and a full relational database engine designed for global optimization of data access.
- IFS helps you quickly prototype and implement solutions involving disparate databases.
- Integrates and transforms data and content to deliver accurate, consistent, timely and complete information on a single platform unified by a common metadata layer
- Unifies a broad range of structured and unstructured data sources.

What If You Could…

- Access data anywhere in your enterprise
  - No matter where it resides
  - Regardless of what format it is in
  - Regardless of vendor
  - Without creating new databases and without disruptive changes to existing ones
  - Using standard SQL and any tool that supports JDBC/ODBC…
  - All while looking to the user like a single database!

…***this is what InfoSphere Federation Server is designed to do!***



InfoSphere Federation Server Supported Data Sources

Relational data sources
- IBM DB2 (for z/OS®, i, LUW)
- IBM Informix®
- Oracle
- Sybase Adaptive Server Enterprise
- Microsoft® SQL Server™
- Teradata
- Netezza
- DB2 Server for VSE and VM• Sources accessible by open database connectivity (ODBC) or Java database connectivity (JDBC)

Extensibility
- C++ and Java Software Development Kits

Other sources and formats
- Web services
- WebSphere MQ message queues
- Microsoft Excel® spreadsheets
- Table-structured flat files
- XML documents
- OLE DB–accessible data sources
- Script output data (Perl, Python and others)
- BioRS

InfoSphere Federation Server Product Features

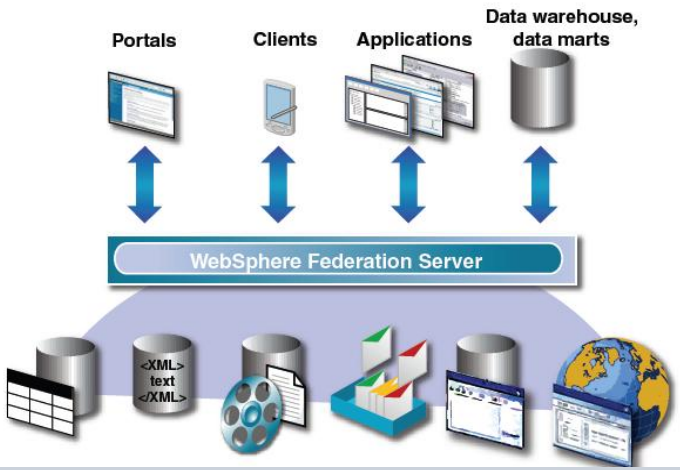
- Transparent
  - Appears to be one source
  - Independent of how and where data is stored
  - Applications continue to work despite of any change in how data is stored
- Heterogeneous
  - Accesses data from diverse sources
  - Relational, Structured, XML, messages, Web content …
- Extensible
  - Bring together almost any data source.
  - Wrapper Development Toolkit
- High Function
  - Full query support against all data
  - Capabilities of sources as well
- Autonomous
  - Non-disruptive to data sources, existing applications, systems.
- High Performance
  - Optimization of distributed queries

Deliver

IBM InfoSphere Federation Server

Access and integrate heterogeneous information across multiple sources as if they were a single source

Extend value of existing analytical applications by providing real-time access to integrated information



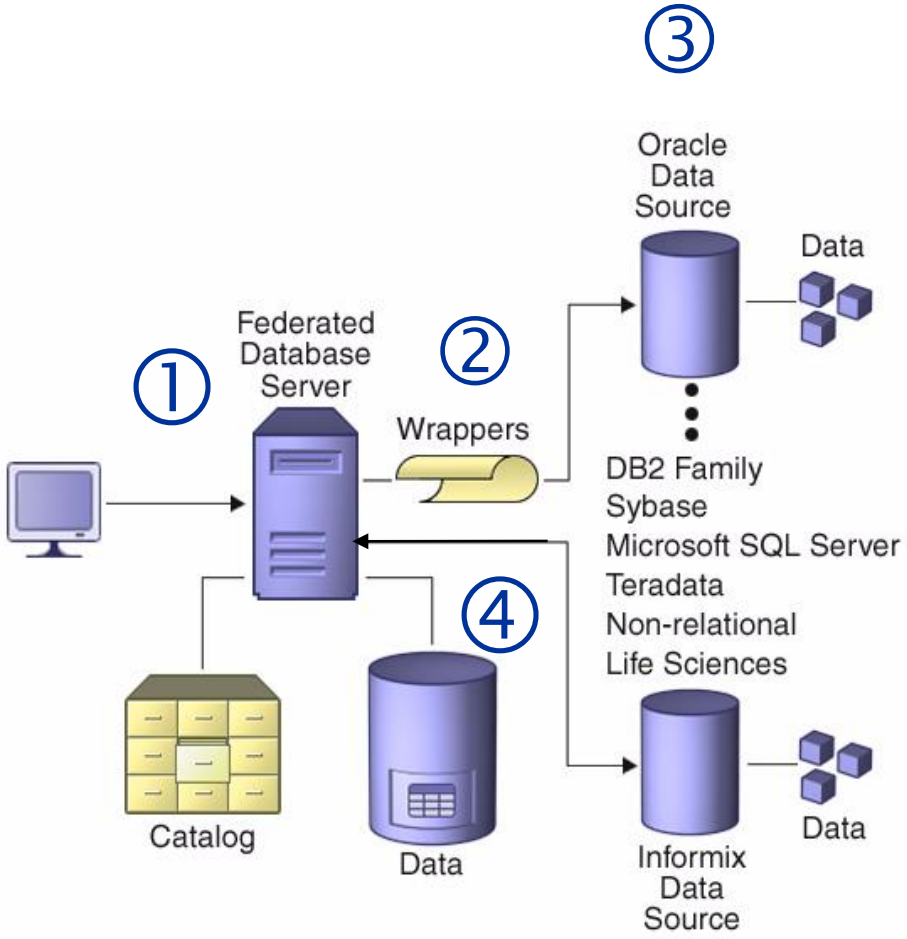
Data Federation Approach

1. User's client sends a request which goes to the Federation server interface

2. The Federation server parses the request and creates query fragments to send to the sources

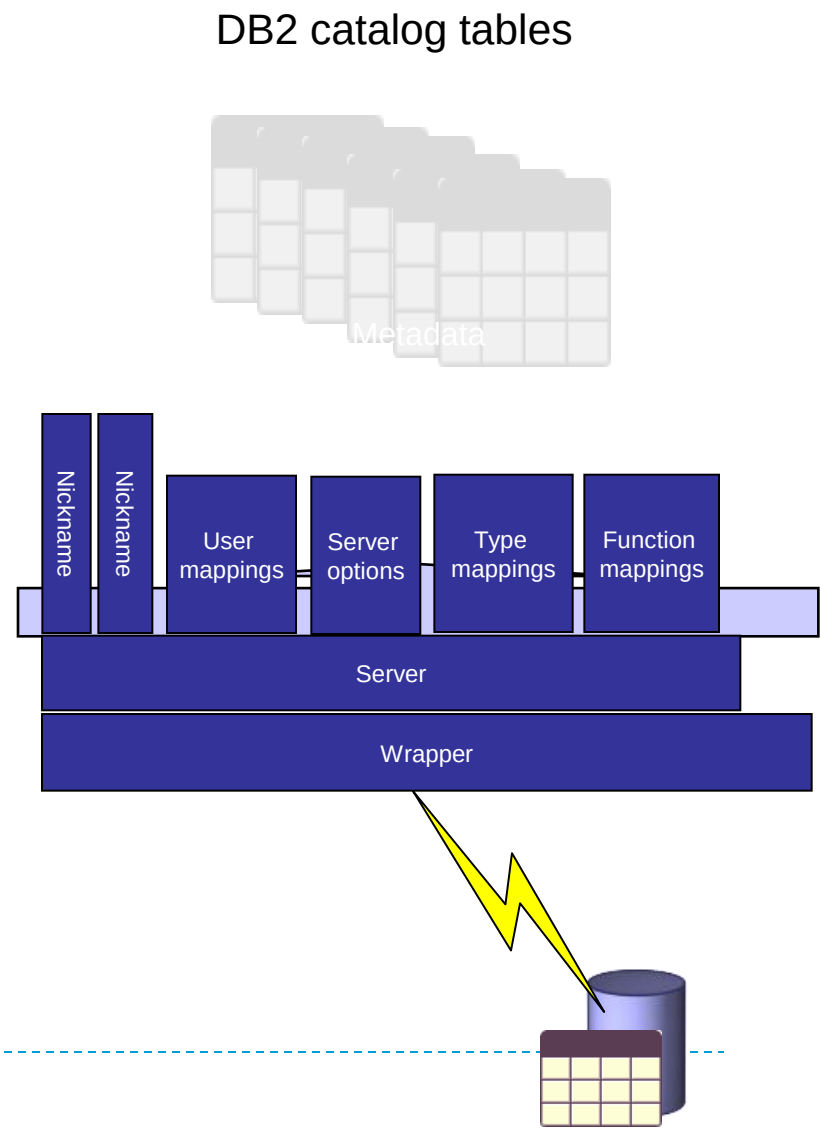
3. Each source system receives its query fragment, processes it and returns the result

4. Federation server assembles the final result (which may mean additional processing) and the result is sent back to the client



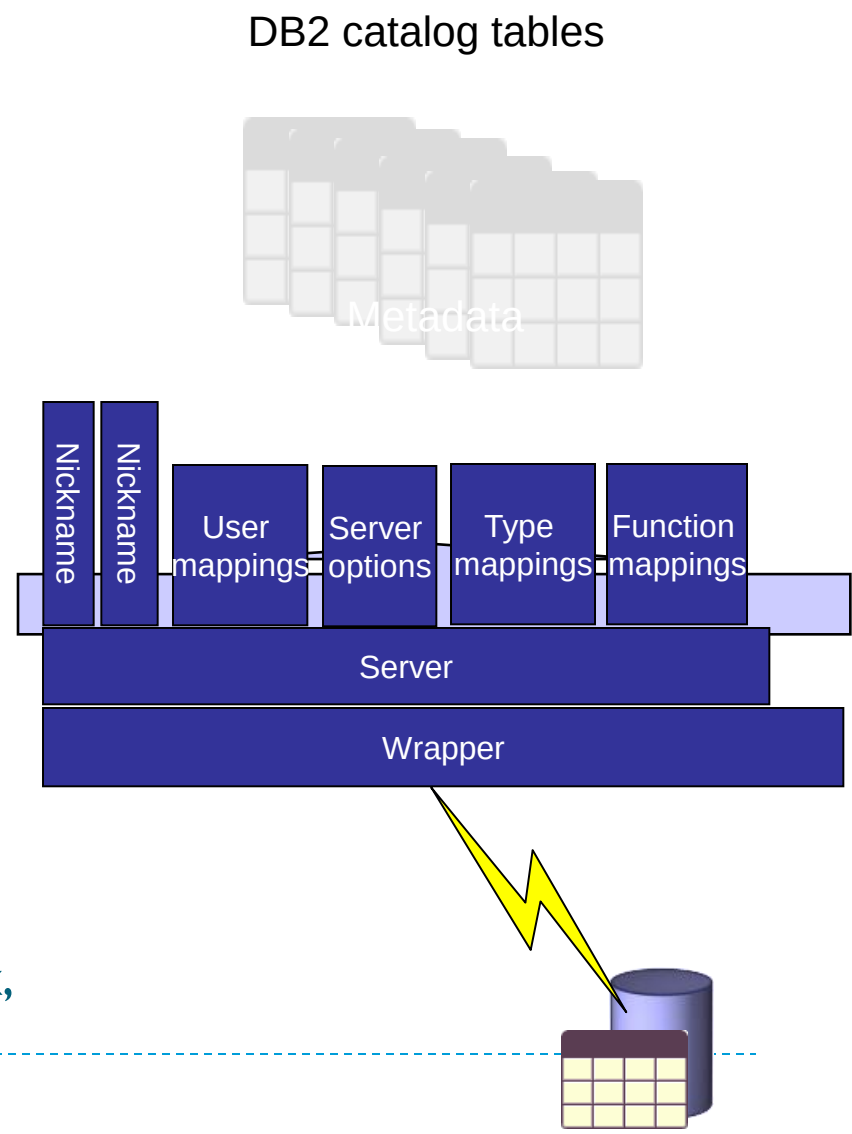
Federation Metadata

- Servers:
  - Type/version of data source ? (Oracle, Sybase, Excel, ODBC… Nonrelational…)
  - Capabilities of the data source
- User mappings:
  - How to map the IFS userid/password to the data source userid/password
  - Password is encrypted
- Server options defining the characteristics of the data source:
  - Does it sort strings the same way as IFS? Does it care about trailing blanks?
  - How does the remote machine speed compare to this one? How fast is the network connection?
  - Query processing capabilities
    - Can it evaluate a predicate? Can it compute aggregates? Can it do a left outer join?



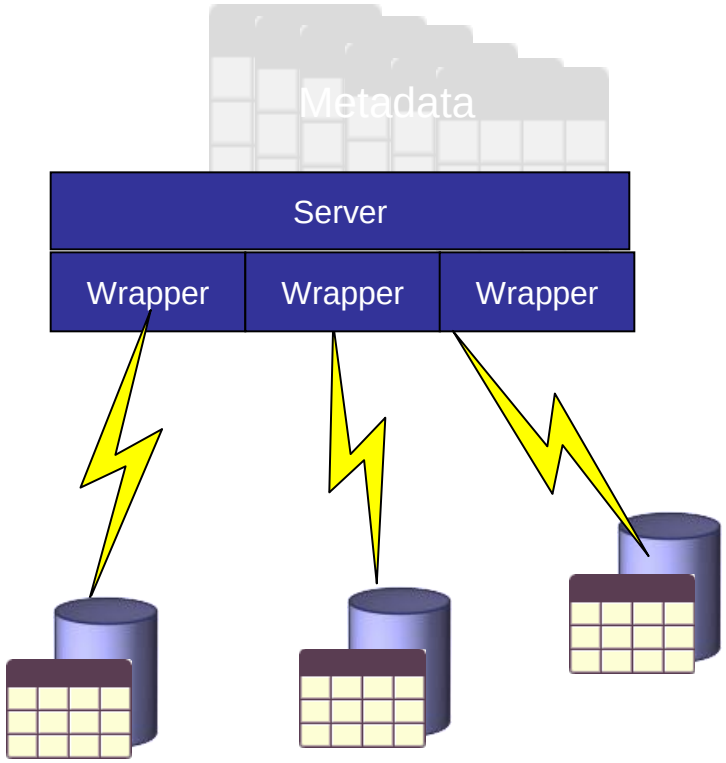
Federation Metadata – Cont.

- When a nickname is created, the data source catalog is 'sniffed' and appropriate metadata collected:
  - What are the names and data types of each column on the remote source?
  - What's the mapping between remote data types and IFS data types?
  - What indexes are available? Are they unique? (tables)
- Statistical information is also retrieved at create nickname time:
  - Tables: How many rows are there?
  - Columns: How many distinct values, what is the range of values?
  - Indexes: Number of distinct values in index, 1st key, 2nd key etc…



Big Data Access

- Script Wrapper
  - Widest access
  - Most maintenance
- UDFs
  - Narrowest Access
  - Highest Performance
- ODBC Wrapper
  - Requires Third Party
  - Least maintenance



Security

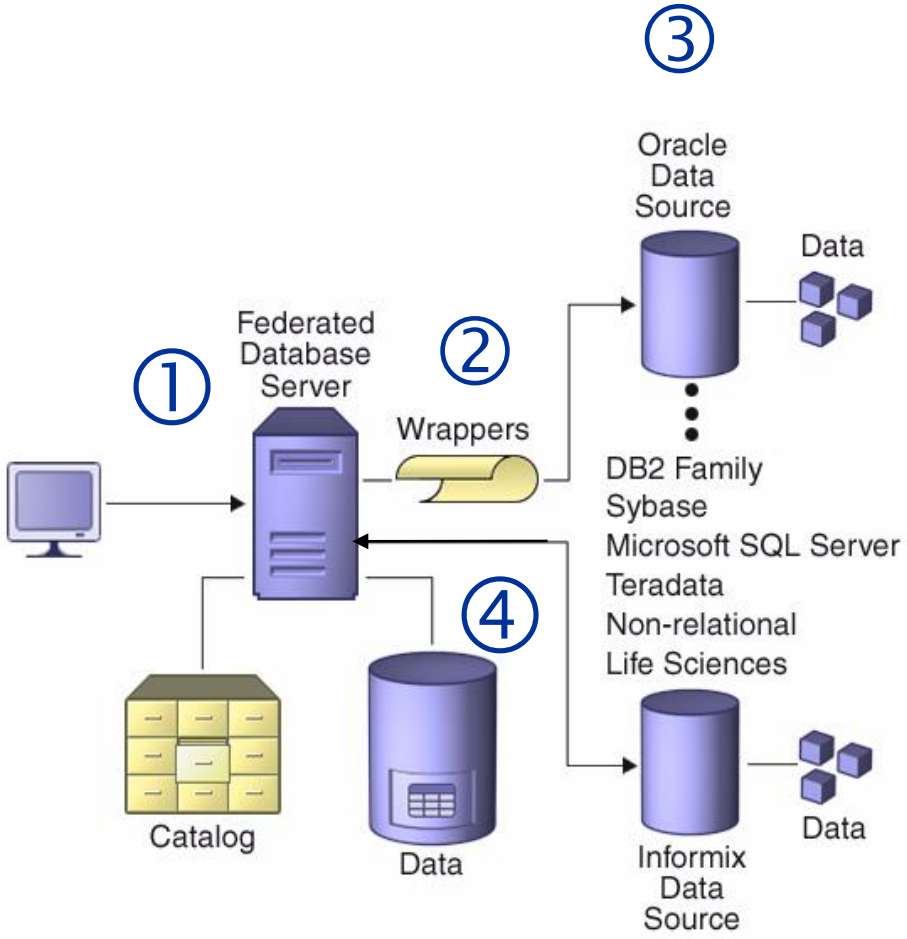
1. Client Authentication

2. Federation Server Privileges

3. Data Source Security

4. LDAP Authentication

+ SSO Support and Trusted Context



THE FILLMORE GROUP
Relational Database Solutions

IBM InfoSphere Federation Server Pricing

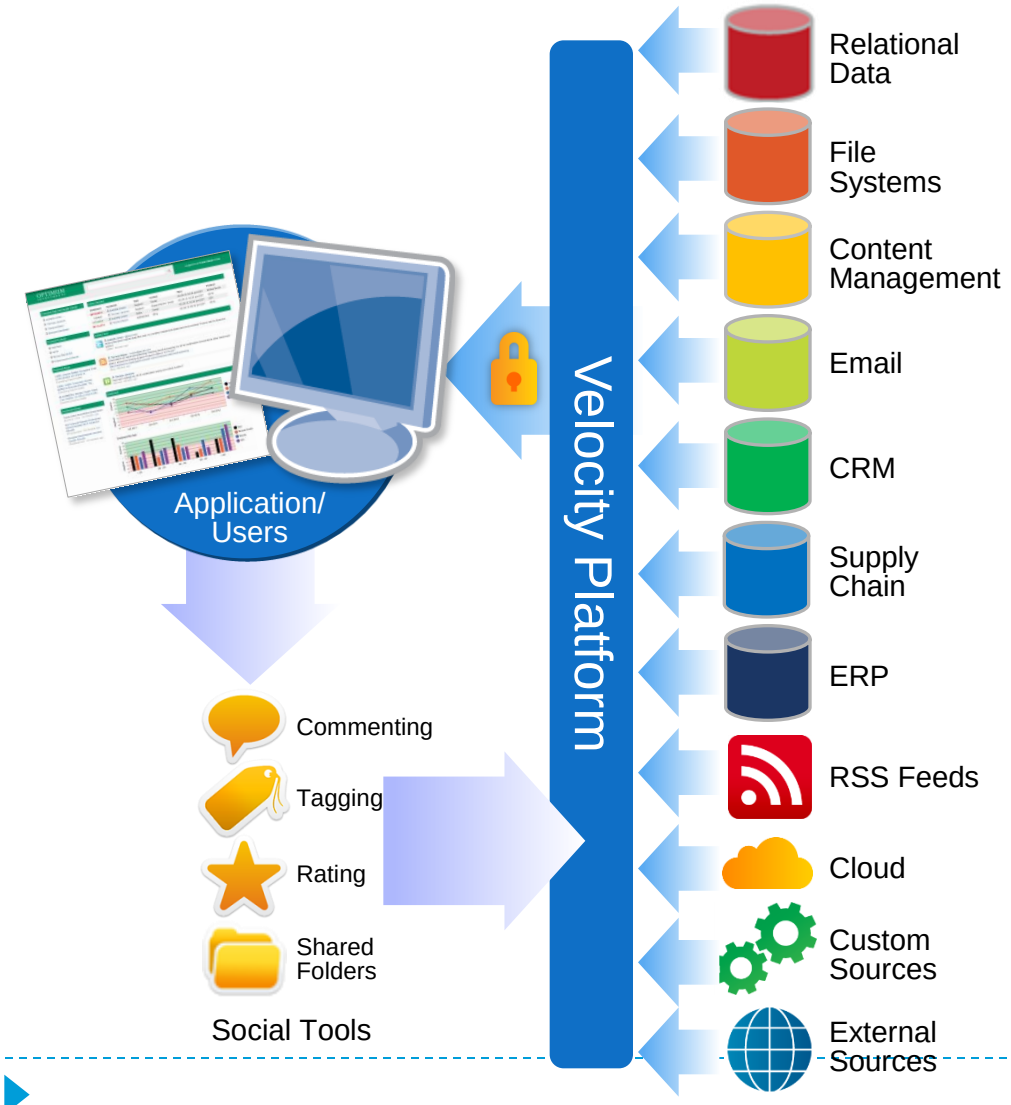
- IBM Processor Value Units for the serve on which InfoSphere Federation Server is installed
- Current pricing (full retail) = $737/PVU

Contact Kim for more information

InfoSphere Data Explorer

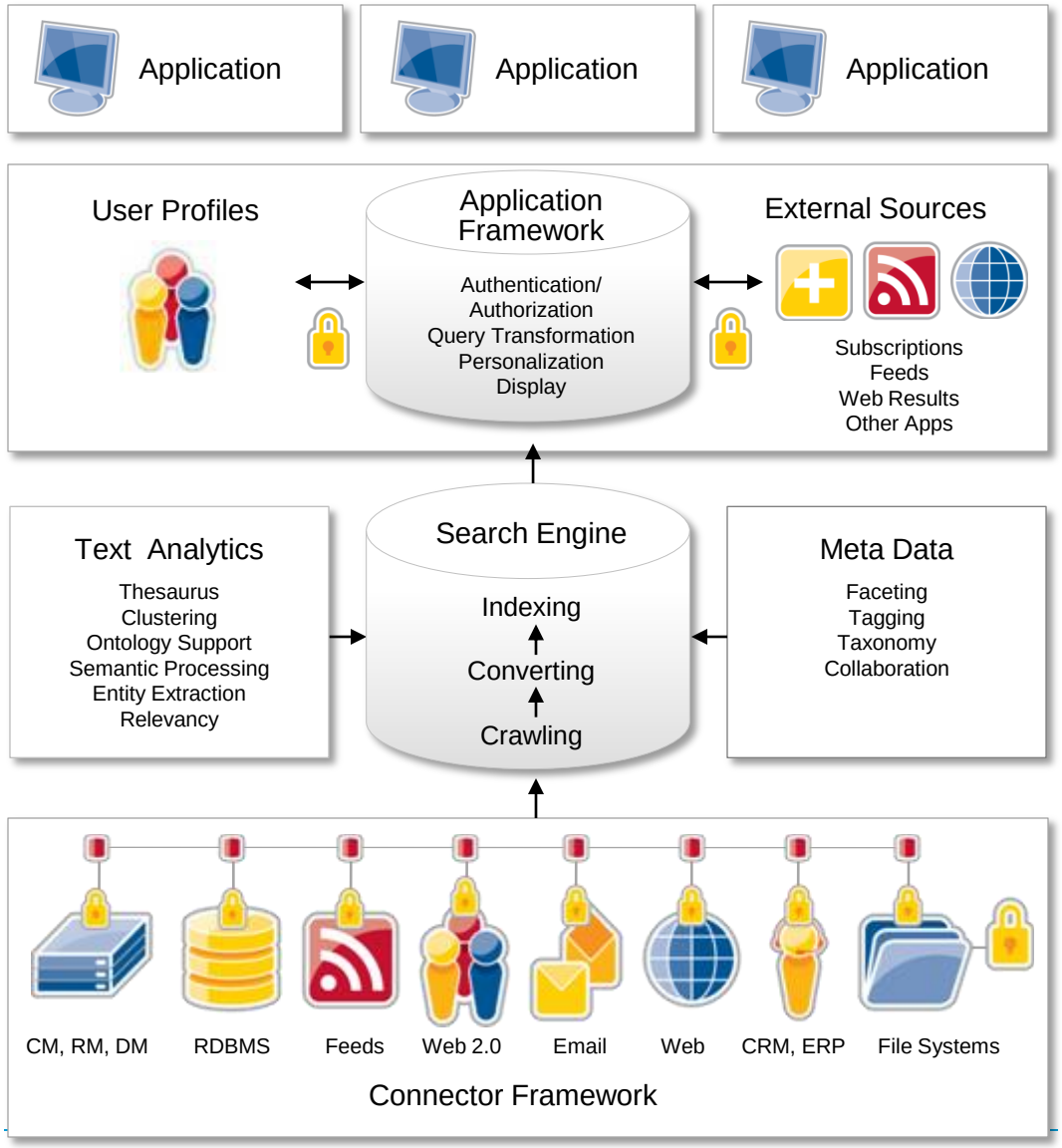
- Acquired by IBM in May, 2012
- Founded in 2000 by three computer science researchers at Carnegie Mellon University
- IBM is focused on integrating Data Explorer with products in their Big Data portfolio

InfoSphere Data Explorer in the Enterprise

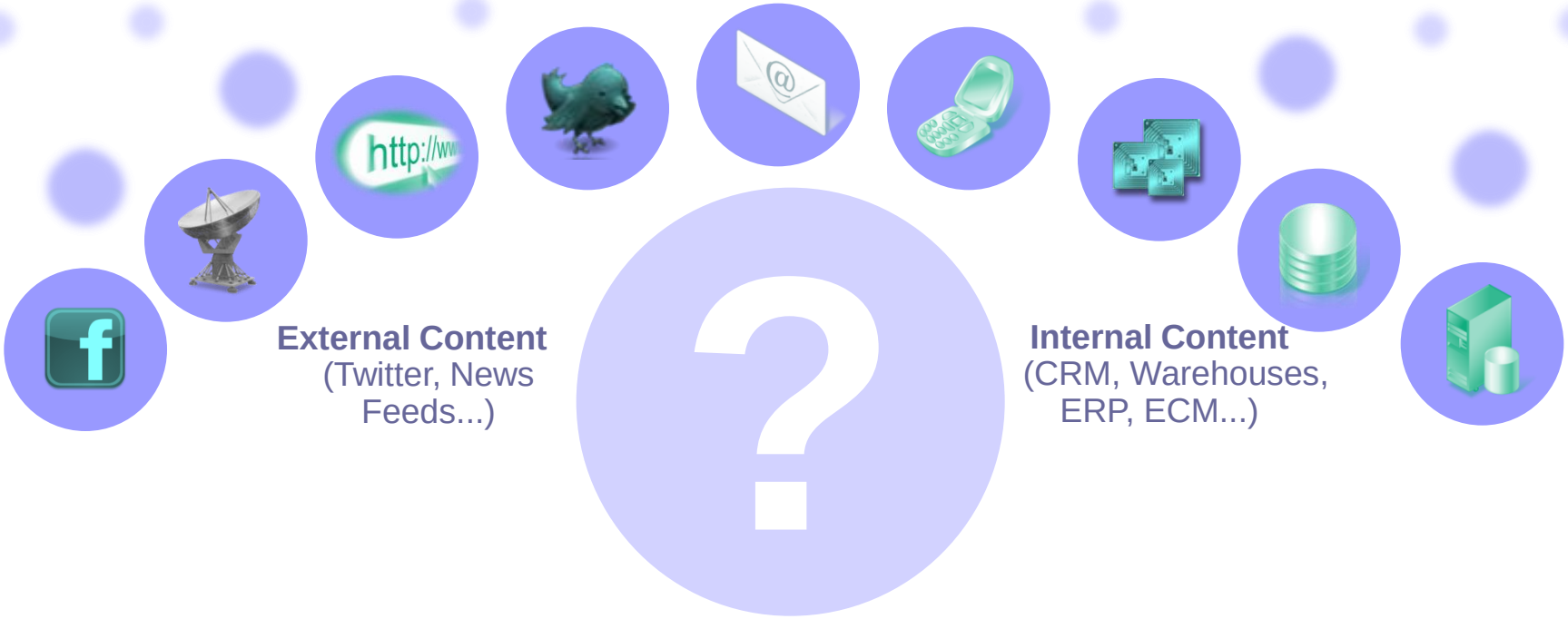


- Secure for access to large number of applications and data stores
- Discovery and navigation across the entire enterprise
- Fusion of structured and unstructured information to drive:
  - Better decisions
  - More efficient operations
  - Better understanding of customers
  - Innovation
- Social tools for collaboration and re-use

Data Explorer Platform Architecture



Data is everywhere, but may be unusable at the point of impact



External Content (Twitter, News Feeds...)

Internal Content (CRM, Warehouses, ERP, ECM...)

"I can't find the right answers fast enough to support my customers."

"Innovation is falling short as I am unable to see the full research picture."

"I can't unlock the value in my data to drive economic value to my business."

"I don't know what I don't know – where is my business exposed?"

"I am monitoring all angles – yet I can't connect the dots."

With Data Explorer - access information *when* you need it most

InfoSphere Data Explorer

Providing unified, real-time access and fusion of big data unlocks greater insight and ROI

Data access & integration
- Index structured & unstructured data—in place
- Support existing security
- Federate to external sources
- Leverage MDM, governance, and taxonomies

Discovery & navigation
- Clustering & categorization
- Contextual intelligence
- Easy-to-deploy applications
- All at the scale required for today's big data challenges

Improve customer service & reduce call times

Increase productivity & leverage past work increasing speed to market

Analyze customer data to unlock true customer value

Identify areas of information risk & ensure data compliance

Create unified view of ALL information for real-time monitoring

Global Aerospace Manufacturer

Increases knowledge worker efficiency resulting in saving USD36 million per year in one department alone.

80,000 internal users, 40,000+ external users

Accessing dozens of separate data sources throughout the enterprise

Single-point of data fusion for all employees

Improved operat rmance for multiple departm

© 2013 IBM Corporation

THE FILLMORE GROUP
Relational Database Solutions

IBM InfoSphere Data Explorer Pricing

- IBM Resource Value Units for all data potentially being federated
- Current pricing (full retail) = $428 per RVU

Contact Kim for more information

Poll Question #3

IBM Information Management

THE FILLMORE GROUP
Relational Database Solutions

Next steps:

- Determine where your organization may benefit from an IBM Data Interoperability solutions.
- Request a product review session for your team
- Identify competing solutions
- Identify and document the factors to evaluate to determine which solution is best
- Let the competition begin!

Premier Business Partner IBM

THE FILLMORE GROUP
Relational Database Solutions

Thank you

Frank Fillmore, President

Kim May, Vice President Business Development

www.thefillmoregroup.com/blog

Premier Business Partner IBM