



Big Data and IMS



Please note

IBM's statements regarding its plans, directions, and intent are subject to change or withdrawal without notice at IBM's sole discretion.

Information regarding potential future products is intended to outline our general product direction and it should not be relied on in making a purchasing decision.

The information mentioned regarding potential future products is not a commitment, promise, or legal obligation to deliver any material, code or functionality. Information about potential future products may not be incorporated into any contract. The development, release, and timing of any future features or functionality described for our products remains at our sole discretion.

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

Acknowledgements and Disclaimers

Availability. References in this presentation to IBM products, programs, or services do not imply that they will be available in all countries in which IBM operates.

The workshops, sessions and materials have been prepared by IBM or the session speakers and reflect their own views. They are provided for informational purposes only, and are neither intended to, nor shall have the effect of being, legal or other guidance or advice to any participant. While efforts were made to verify the completeness and accuracy of the information contained in this presentation, it is provided AS-IS without warranty of any kind, express or implied. IBM shall not be responsible for any damages arising out of the use of, or otherwise related to, this presentation or any other materials. Nothing contained in this presentation is intended to, nor shall have the effect of, creating any warranties or representations from IBM or its suppliers or licensors, or altering the terms and conditions of the applicable license agreement governing the use of IBM software.

All customer examples described are presented as illustrations of how those customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics may vary by customer. Nothing contained in these materials is intended to, nor shall have the effect of, stating or implying that any activities undertaken by you will result in any specific sales, revenue growth or other results.

© *Copyright IBM Corporation 2014. All rights reserved.*

• ***U.S. Government Users Restricted Rights - Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM Corp.***

IBM, the IBM logo, ibm.com, IMS are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. If these and other IBM trademarked terms are marked on their first occurrence in this information with a trademark symbol (® or ™), these symbols indicate U.S. registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the Web at “Copyright and trademark information” at www.ibm.com/legal/copytrade.shtml

Other company, product, or service names may be trademarks or service marks of others.

On a Smarter Planet, Unprecedented Changes are Occurring



- Business models under constant pressure
- Customers are more demanding and connected
- Great relationships trump great products

And leaders are responding by...



**Providing a
Great Experience**



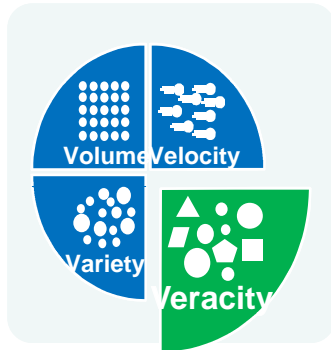
**Offering Value
In Every
Interaction**



**Innovating
Across
the Ecosystem**

Forward thinking organizations are creating value from Big Data

The power of Data coming together...



...with the power of Technology...



...to deliver improved business outcomes



1. Enrich your information base
with Big Data Exploration



2. Improve customer interaction
with Enhanced 360° View of the Customer



3. Optimize operations
with Operations Analysis



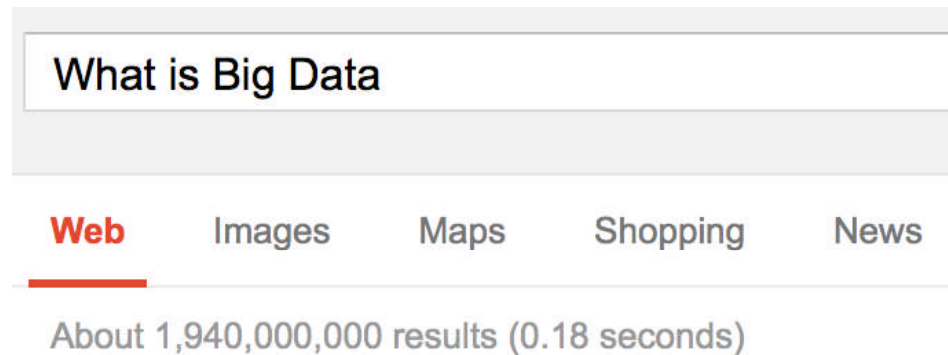
4. Gain IT efficiency and scale
with Data Warehouse Augmentation



5. Prevent crime
with Security and Intelligence Extension

But what is Big Data?

- *Google can give you nearly 2 Billion options*
- *Vendors have even more definitions*



Here is how Gartner defines Big Data

- *Big data is high-volume, high-velocity and high-variety information assets that demand **cost-effective, innovative information processing for enhanced insight and decision making.***

Gartner research note "Survey Analysis - Big Data Adoption in 2013 Shows Substance Behind the Hype" Sept 12 2013
Analyst(s): Lisa Kart, Nick Heudecker, Frank Buytendijk

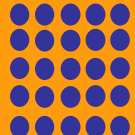
We've moved into a new era of computing

- V^4

Radical Flexibility

12 terabytes

of Tweets create daily

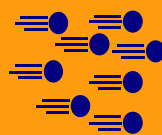


Volume

Extreme Scalability

5 million

trade events per second



Velocity

100's

Of video feeds from surveillance cameras



Variety



Veracity

Only **1 in 3**

Decision makers trust their information

Information from everywhere

“We have for the first time an economy based on a key resource [Information] that is not only renewable, but self-generating.

Running out of it is not a problem, *but drowning in it is.*”

– John Naisbitt

The 3 V's of Big Data: Volume, Velocity and Variety

- **Volume:** Enterprises are awash with ever-growing data of all types, easily amassing terabytes—even petabytes—of information.
 - Turn **12 terabytes of Tweets created each day into improved product sentiment analysis**
 - **Convert 350 billion annual meter readings to better predict power consumption**
- **Velocity:** Sometimes 2 minutes is too late. For time-sensitive processes such as catching fraud, big data must be used as it streams into your enterprise in order to maximize its value.
 - **Scrutinize 5 million trade events created each day to identify potential fraud**
 - **Analyze 500 million daily call detail records in real-time to predict customer churn faster**
 - **The latest I have heard is 10 nano seconds delay is too much.**
- **Variety:** Big data is any type of data - structured and unstructured data such as text, sensor data, audio, video, click streams, log files and more. New insights are found when analyzing these data types together.
 - **Monitor 100's of live video feeds from surveillance cameras to target points of interest**
 - **Exploit the 80% data growth in images, video and documents to improve customer satisfaction**
- And a 4TH V:
- **Veracity:** Unstructured Big Data can be of uncertain accuracy

Demand for differently structured data to be seamlessly integrated, to augment analytics / decisions

- Analytics and decision engines reside where the DWH / transaction data is
- “Noise” surrounds the core business data
 - Social Media, emails, docs, telemetry, voice, video, content
- Expanding our insights – getting closer to the “truth”
 - Lower risk and cost
 - Increased profitability

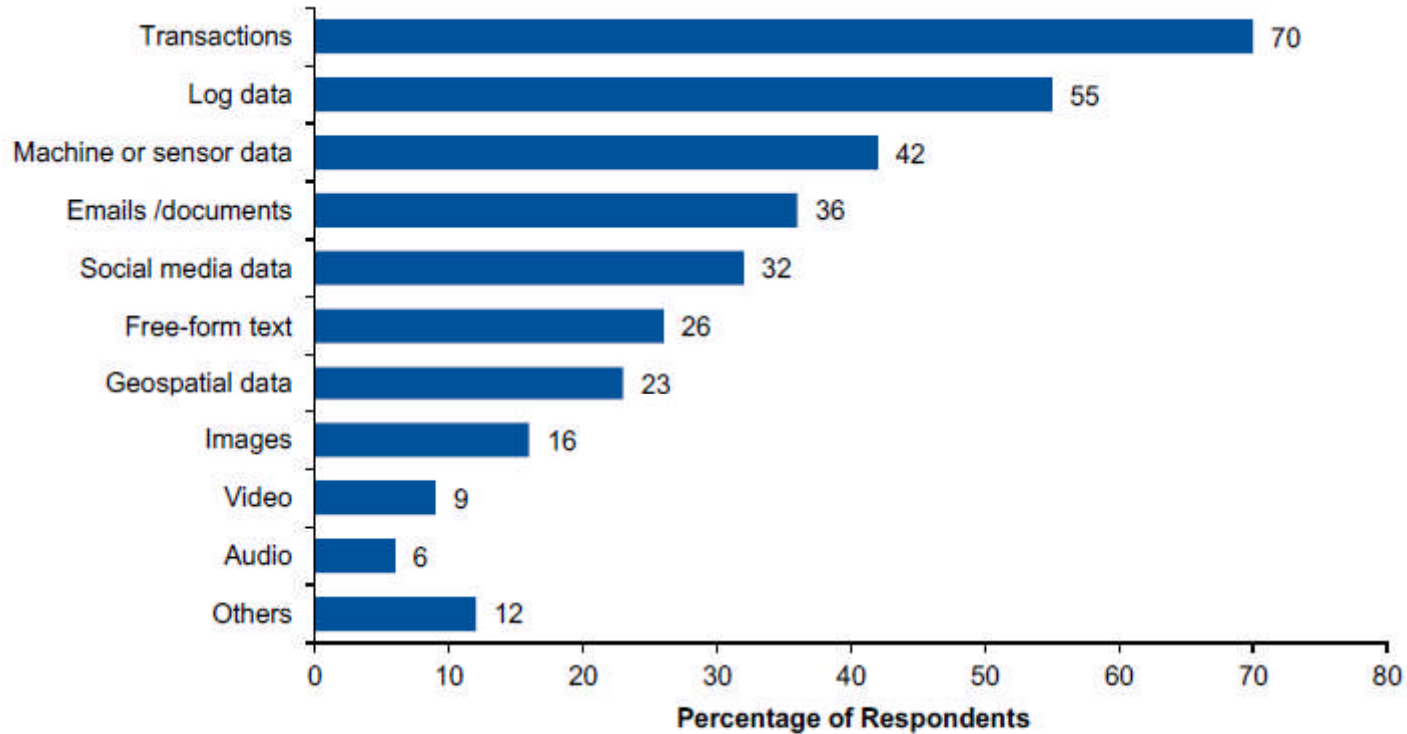


“Circle of trust” widens

The Big Data starting point

Types of Data Analysed

Transactional sources are the dominant data types analyzed in big data initiatives



N =465 (multiple responses allowed)

Source: Gartner (September 2013)

Gartner research note "Survey Analysis - Big Data Adoption in 2013 Shows Substance Behind the Hype" Sept 12 2013
Analyst(s): Lisa Kart, Nick Heudecker, Frank Buytendijk

The Big Data starting point

Types of Big Data Analysed by Industry

Transactional sources are the dominant data types analyzed in big data initiatives

	Manufacturing and Natural Resources	Media/ Communications	Services	Government	Education	Retail	Banking	Insurance	Healthcare	Transportation	Utilities
Transactions	73%	62%	67%	67%	54%	93%	83%	81%	75%	79%	80%
Log data	44%	57%	58%	59%	54%	40%	66%	61%	33%	71%	60%
Machine or sensor data	53%	38%	35%	33%	31%	27%	27%	48%	42%	50%	40%
Emails /documents	27%	43%	43%	41%	46%	27%	34%	39%	17%	29%	20%
Social media data	32%	52%	39%	26%	54%	73%	27%	13%	-	50%	-
Free-form text	17%	24%	28%	30%	31%	20%	34%	35%	67%	21%	40%
Geospatial data	27%	14%	19%	19%	38%	27%	27%	26%	8%	29%	40%
Images	19%	24%	17%	11%	38%	13%	5%	16%	25%	7%	-
Video	8%	29%	12%	7%	31%	13%	-	6%	8%	7%	-
Audio	10%	19%	8%	4%	8%	-	-	6%	-	-	-
Other	8%	14%	13%	15%	8%	7%	10%	16%	42%	14%	-
n =	59	21*	127	27*	13*	15*	41	31	12*	14*	5*

Note: Highlighted cells indicate the top three data types by industry. Multiple responses allowed

Source: Gartner (September 2013)

Gartner research note "Survey Analysis - Big Data Adoption in 2013 Shows Substance Behind the Hype" Sept 12 2013
Analyst(s): Lisa Kart, Nick Heudecker, Frank Buytendijk

Big Data use cases



Big Data Exploration

Find, visualize, understand all big data to improve decision making



Enhanced 360° View of the Customer

Extend existing customer views (MDM, CRM, etc) by incorporating additional internal and external information sources



Security/Intelligence Extension

Lower risk, detect fraud and monitor cyber security in real-time



Operations Analysis

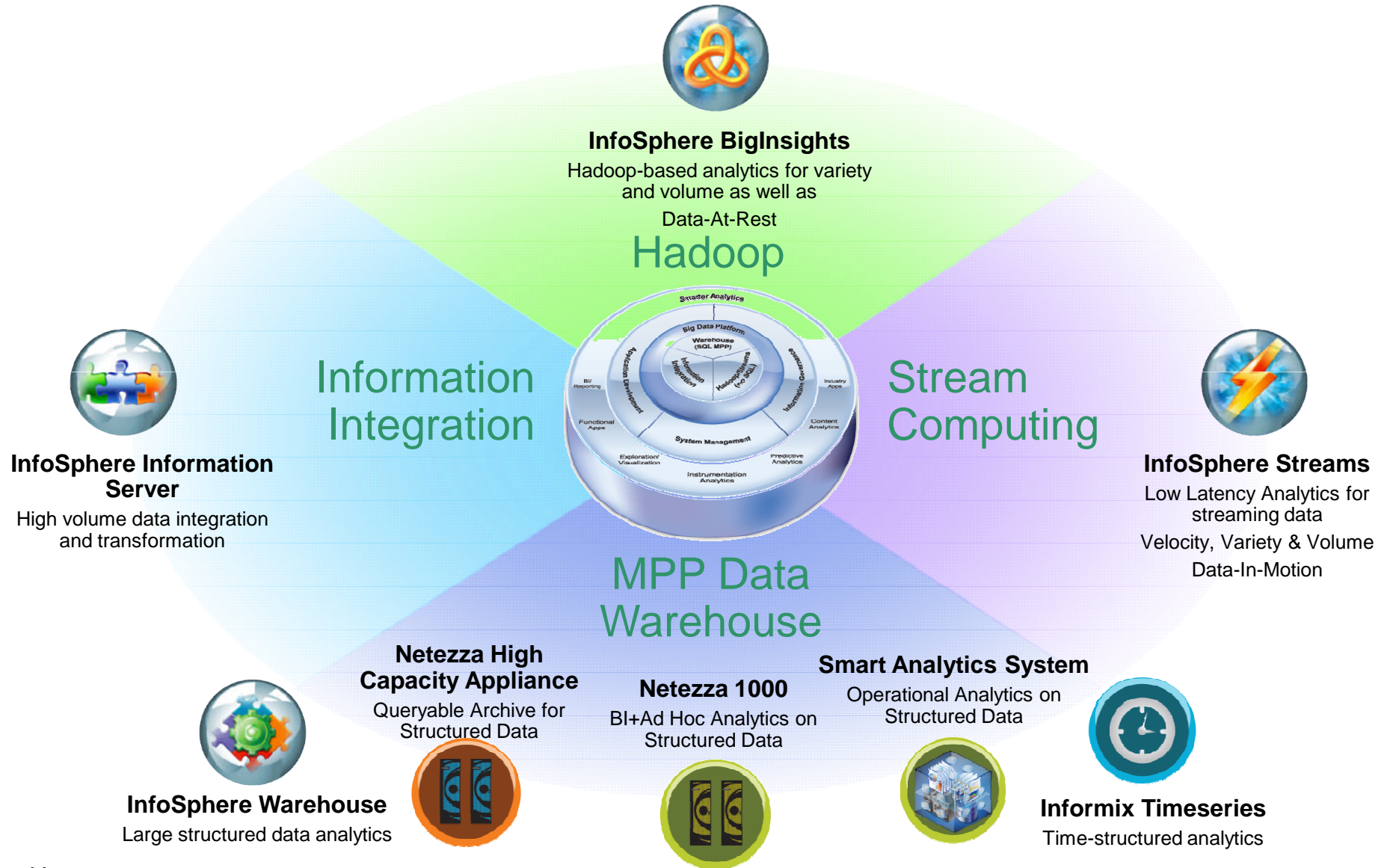
Analyze a variety of machine data for improved business results



Data Warehouse Augmentation

Integrate big data and data warehouse capabilities to increase operational efficiency

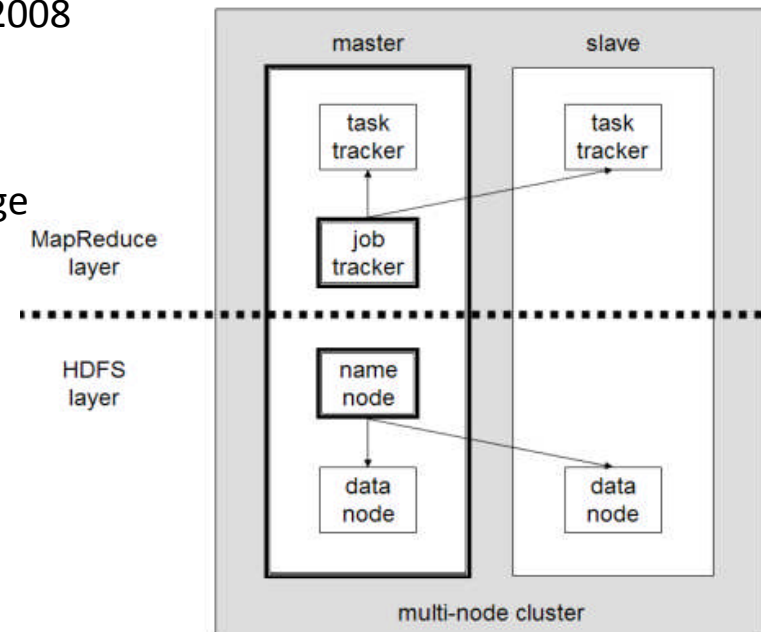
The IBM Big Data platform



What is



- An open source software framework that supports data-intensive distributed applications
 - High throughput, batch processing
 - runs on large clusters of commodity hardware
 - Yahoo runs a 4000 nodes Hadoop cluster in 2008
- Two main components
 - Hadoop distributed file system
 - self-healing, high-bandwidth clustered storage
 - MapReduce engine



Hadoop: The underlying principle

- Lots of redundant disks – really inexpensive disks
- Lots of cores – inexpensive cores working all the time
- Disks crash – that's ok – just replace them
- Processors fail – that's ok – just replace them
- Network errors happen – that's ok - just retry
- Disks, processors networked

Hadoop Distributed File System (HDFS)

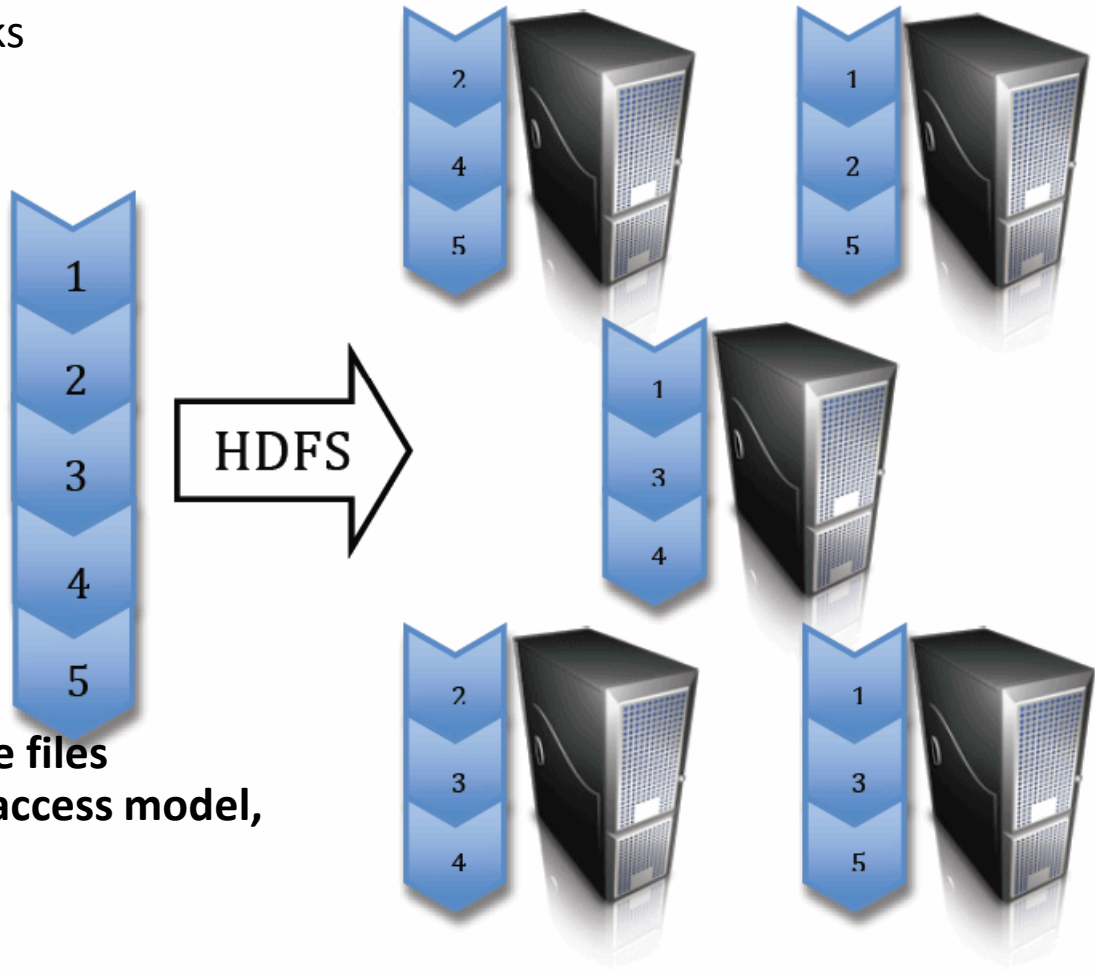
Files are broken in to large blocks
(default=64MB).

Blocks are replicated
(default=3 times) and
distributed across the
cluster.




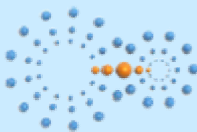

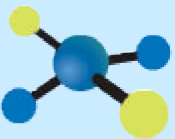
- **Durability**
- **Availability**
- **Throughput**

Optimized for

- **Streaming reads of large files**
- **write-once-read-many access model,
append-only**

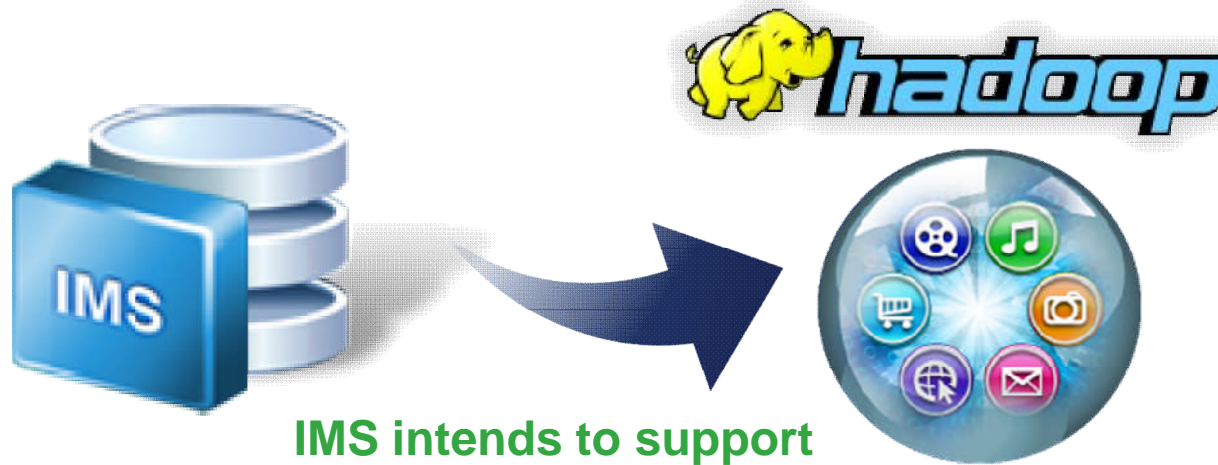


BIG DATA is not just HADOOP

Understand and navigate federated big data sources		Federated Discovery and Navigation
Manage & store huge volume of any data		Hadoop File System MapReduce
Structure and control data		Data Warehousing
Manage streaming data		Stream Computing
Analyze unstructured data		Text Analytics Engine
Integrate and govern all data sources		Integration, Data Quality, Security, Lifecycle Management, MDM

Enhancing IMS analytics on System z with Big Data

- Much of the world's operational data resides on z/OS
- Unstructured data sources are growing fast
- There is a need to merge this data with trusted OLTP data from System z data sources
- IMS provides the connectors and the DB capability to allow BigInsights v2.1.2.0 to easily and efficiently access the IMS data source

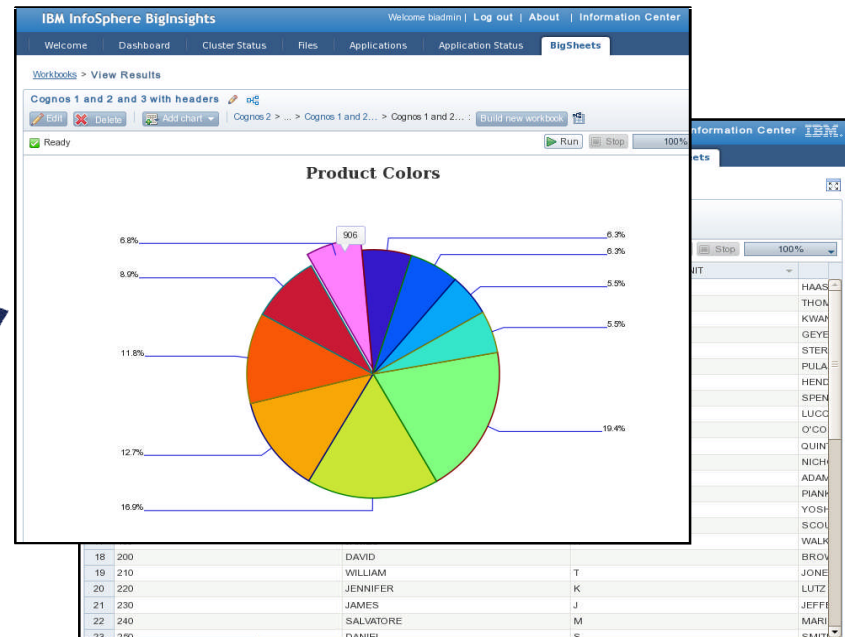
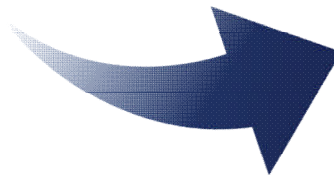


Enhancing IMS analytics on System z with Big Data

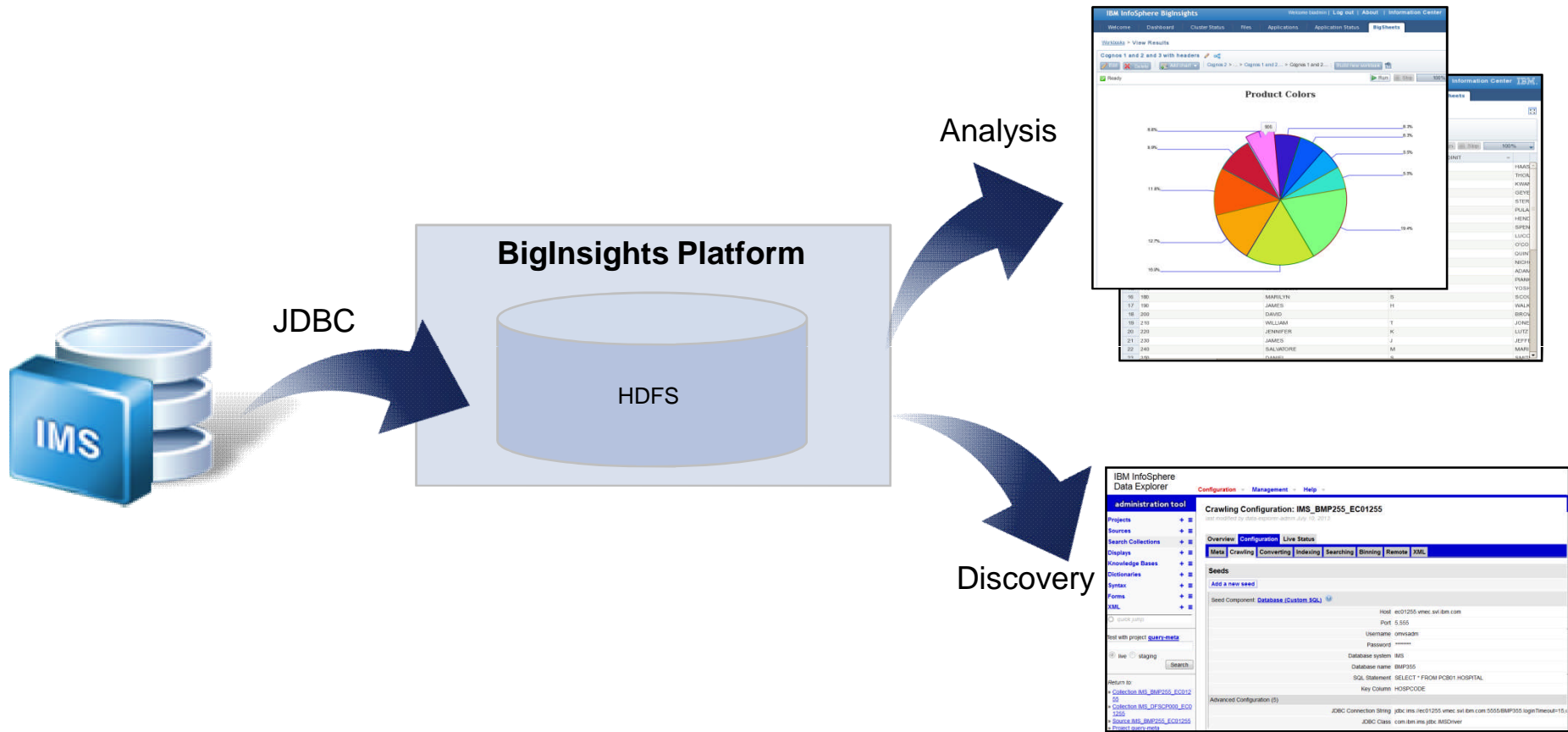
- Observation points lead to new business opportunities
- Observation points gleaned from both archived data and live data
- Score business events, track claims evolution, and more
- *Make the data available to people who can do something meaningful with it*



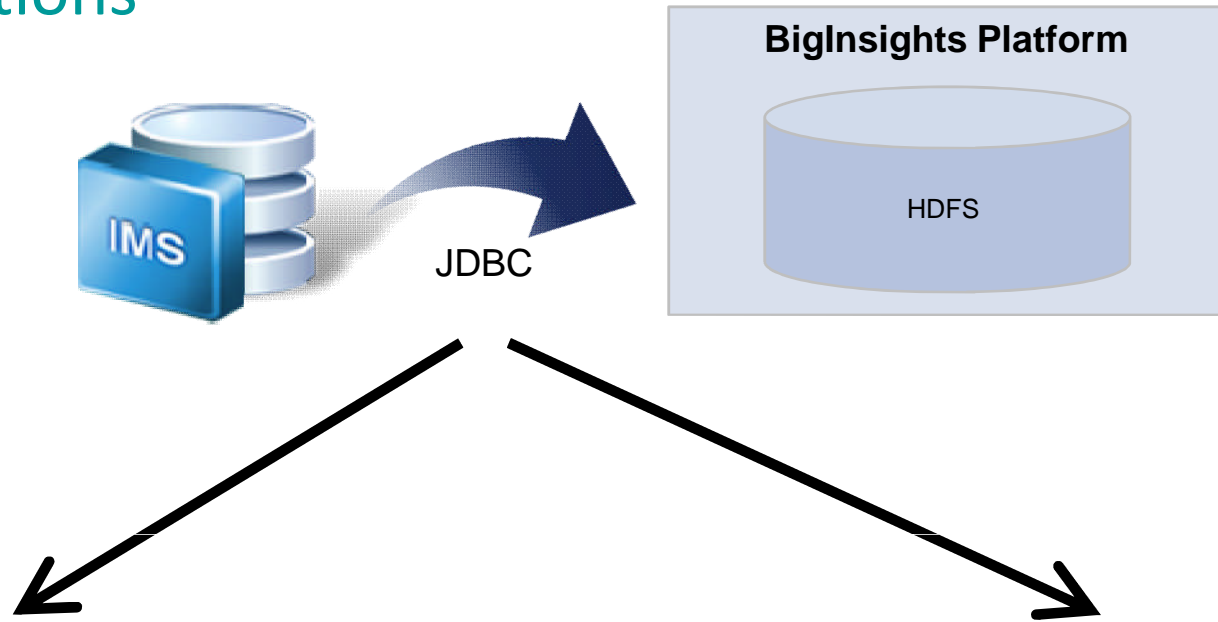
IMS intends to support



High level overview



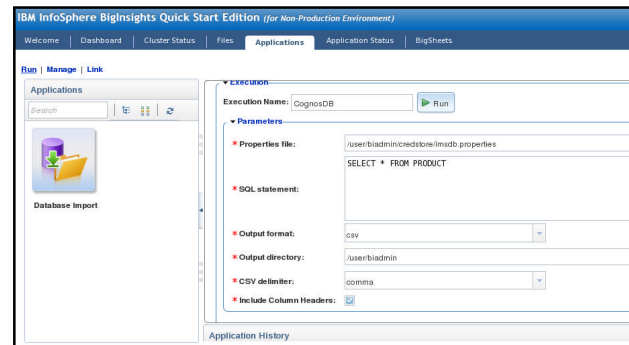
Import options



Sqoop Import

	Header
1	000070,EVA , DPULASKI ,D21,7831,2003-05-26,MANAGER , -3594,F,1985-07-09,-2022459.00,-1687590.56,-23
2	000090,EILEEN , WHENDERSON ,E11,5498,1971-05-15,MANAGER , -3594,F,1985-07-09,-2022459.00,-168759
3	000110,VINCENZO , GLUCCHESSI ,A00,3490,1988-05-16,SALEREP , -3591,M,1985-07-09,-2022459.00,-16875
4	000120,SEAN , O'CONNELL ,A00,2167,1993-12-05,CLERK , -3596,M,1985-07-09,-2022459.00,-1687590.56,-2
5	000140,HEATHER , ANICHOLLS ,C01,1793,2006-12-15,ANALYST , -3593,F,1985-07-09,-2022459.00,-1687590.
6	000170,MASATOSHI , JYOSHIMURA ,D11,2890,1999-09-15,DESIGNER , -3594,M,1985-07-09,-2022459.00,-168
7	000220,JENNIFER , KLUTZ ,D11,0672,1998-08-29,DESIGNER , -3592,F,1985-07-09,-2022459.00,-1687590.56
8	000270,MARIA , LPEREZ ,D21,9001,2006-09-30,CLERK , -3595,F,1985-07-09,-2022459.00,-1687590.56,-2352

POJO Import



Sqoop Import

- Command line interface application for transferring data between RDBMS and HDFS.
- Import into Hive and Hbase
- Export from HDFS back into RDBMS
- Import:
 - Divides table into ranges using primary key max/min (can use split-by param)
 - Creates mappers for each range
 - Mappers write to multiple HDFS nodes
 - Creates text or sequence files
- Export:
 - Reads files in HDFS directory via MapReduce
 - Bulk parallel insert into database table.

Sqoop Import

- Import into HDFS using the below command:

```
./sqoop import --connect
jdbc:ims://ecwas09.vmec.svl.ibm.com:5555/BIGDATP1 --driver
com.ibm.ims.jdbc.IMSDriver --table EMPLOYEE -m 3 --split-by
EDLEVEL --username 'OMVSADM' -P
13/06/19 17:50:27 INFO db.DataDrivenDBInputFormat: BoundingValsQuery:
SELECT MIN(EDLEVEL), MAX(EDLEVEL) FROM EMPLOYEE
13/06/19 17:50:46 INFO mapreduce.ImportJobBase: Transferred 5.123 KB in
20.3762 seconds (257.4572 bytes/sec)
13/06/19 17:50:46 INFO mapreduce.ImportJobBase: Retrieved 43 records.
```

- HDFS Output (below)

	Header
1	000070,EVA , ,DPULASKI ,D21,7831,2003-05-26,MANAGER , -3594,F,1985-07-09,-2022459.00,-1687590.56,-23
2	000090,EILEEN , ,WHENDERSON ,E11,5498,1971-05-15,MANAGER , -3594,F,1985-07-09,-2022459.00,-1687590
3	000110,VINCENZO , ,GLUCCHETTI ,A00,3490,1988-05-16,SALEREP , -3591,M,1985-07-09,-2022459.00,-16875
4	000120,SEAN , , O'CONNELL ,A00,2167,1993-12-05,CLERK , -3596,M,1985-07-09,-2022459.00,-1687590.56,-23
5	000140,HEATHER , ,ANICHOLLS ,C01,1793,2006-12-15,ANALYST , -3593,F,1985-07-09,-2022459.00,-1687590.
6	000170,MASATOSHI , ,JYOSHIMURA ,D11,2890,1999-09-15,DESIGNER , -3594,M,1985-07-09,-2022459.00,-168
7	000220,JENNIFER , , KLUTZ ,D11,0672,1998-08-29,DESIGNER , -3592,F,1985-07-09,-2022459.00,-1687590.56
8	000270,MARIA , , LPEREZ ,D21,9001,2006-09-30,CLERK , -3595,F,1985-07-09,-2022459.00,-1687590.56,-2352

BigInsights Database Import Application

- Utilize the built in Database Import Application by providing the database connection parameters:

The screenshot displays the IBM InfoSphere BigInsights Quick Start Edition interface. At the top, a blue navigation bar contains the following items: Welcome, Dashboard, Cluster Status, Files, Applications (highlighted), Application Status, and BigSheets. Below this, a secondary navigation bar includes Run, Manage, and Link. The main content area is divided into two panels. The left panel, titled 'Applications', features a search bar and a 'Database Import' icon. The right panel, titled 'Execution', shows the configuration for an application named 'CognosDB'. It includes a 'Run' button and several parameter fields: 'Properties file' (set to /user/biadmin/credstore/imsdb.properties), 'SQL statement' (set to SELECT * FROM PRODUCT), 'Output format' (set to csv), 'Output directory' (set to /user/biadmin), 'CSV delimiter' (set to comma), and 'Include Column Headers' (checked).

BigInsights Database Import Application

- Once the run is completed, view the data in HDFS:

Application History

Status	Execution Name	Progress	Start Time	Elapsed Time (sec)	Output	Details
No filter applied						
✓	CognosDB	100%	2013 10 3 20:53:20	14		
✓	CognosDB	100%	2013 10 3 20:41:22	18		

IBM InfoSphere BigInsights Quick Start Edition (for Non-Production Environment) About | Info

Welcome | Dashboard | Cluster Status | **Files** | Applications | Application Status | BigSheets

HDFS

- hdfs://bivm:9000/
 - biginsights
 - hadoop
 - hbase
 - tmp
 - user
 - applications
 - biadmin
 - .staging
 - credstore
 - oozie-biad
 - output.txt**

Path: /user/biadmin/output.txt

Name	Size	Block Size	Time	Permission	Owner
output.txt	7.8 KB	128.0 MB	Oct 3, 2013 8:53:31 PM	rw-r--r--	biadmin

Edit Viewing Size: 10KB Text Sheet

hdfs://bivm:9000/user/biadmin/output.txt
Comma Separated Value (CSV) Data

✓ Ready

	GOSALES_ROOTKEY	COUNTRY_CODE	COUNTRY_EN	FLAG_IMAGE
1	1	1003	United States	F03.jpg
2	1	1004	Canada	F04.jpg
3	1	1020	Mexico	F19.jpg
4	1	1021	Brazil	F20.jpg
5	1	4011	Japan	F11.jpg
6	1	4012	Singapore	F12.jpg
7	1	4013	Korea	F13.jpg
8	1	4014	China	F14.jpg

BigInsights BigSheets



- This data can be saved as BigSheets workbook for further analytics

The screenshot displays the IBM InfoSphere BigInsights interface. At the top, there is a navigation bar with 'Welcome biadmin | Log out | About | Information Center' and the IBM logo. Below this is a secondary navigation bar with tabs for 'Welcome', 'Dashboard', 'Cluster Status', 'Files', 'Applications', 'Application Status', and 'BigSheets'. The main content area shows 'Workbooks > View Results' for a workbook named 'Employee'. The interface includes a toolbar with 'Delete', 'Add chart', 'Employee', and 'Build new workbook' buttons. Below the toolbar, there are status indicators ('Ready'), 'Refresh', 'Fit column(s)', 'Export as', 'Run', 'Stop', and a progress indicator at '100%'. The data table has the following structure:

	EMPNO	FIRSTNME	MIDINIT	
1	10	CHRISTINE	I	HAAS
2	20	MICHAEL	L	THON
3	30	SALLY	A	KWAN
4	50	JOHN	B	GEYE
5	60	IRVING	F	STER
6	70	EVA	D	PULA
7	90	EILEEN	W	HEND
8	100	THEODORE	Q	SPEN
9	110	VINCENZO	G	LUCC
10	120	SEAN		O'CO
11	130	DELORES	M	QUIN
12	140	HEATHER	A	NICH
13	150	BRUCE		ADAM
14	160	ELIZABETH	R	PIAN
15	170	MASATOSHI	J	YOSH
16	180	MARILYN	S	SCOU
17	190	JAMES	H	WALK
18	200	DAVID		BROV
19	210	WILLIAM	T	JONE
20	220	JENNIFER	K	LUTZ
21	230	JAMES	J	JEFF
22	240	SALVATORE	M	MARI
23	250	DANIEL	S	SMIT

Elevated demand for business analytics drives new requirements and focus

More aggressive requirements

- Enterprise-level scale & performance
- Mission critical availability
- Faster access to operational data
- Rapid, cost effective deployment & expansion
- More integrated view of data across the environment

Driving new focuses

- Modernization
- Standardization & Consolidation
- Operational BI
- Data Governance
- Cloud Computing



The role of zEnterprise in Big Data analytics

- A large percent of the data that is accessed for analytics originates/resides on IBM zEnterprise
 - 2/3 of business transactions for U.S. retail banks
 - 80% of world's corporate data
- Businesses that run on zEnterprise
 - 66 of the top 66 worldwide banks
 - 24 of the top 25 U.S. retailers
 - 10 of the top 10 global life/health insurance providers
- 1,300+ ISVs run zEnterprise today, more than 275 of these selling over 800 applications on Linux
- The downtime of an application running on System z equates to approximately 5 minutes per year
- The System z mainframe can run over a thousand virtual Linux images on a single frame the size of a refrigerator

DB2 Analytics Accelerator

Accelerating decisions to the speed of business

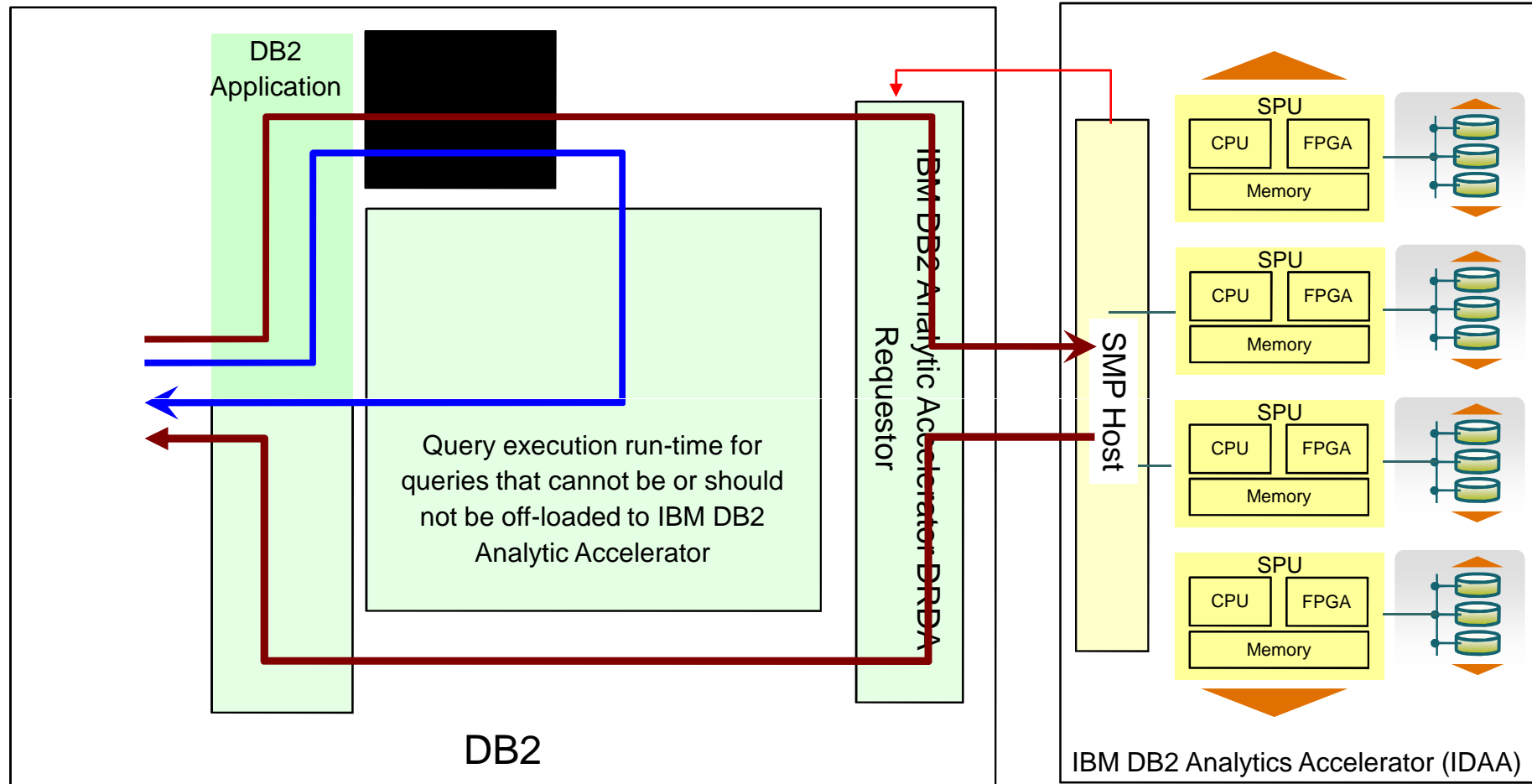
Blending System z and Netezza technologies to deliver unparalleled, mixed workload performance for complex analytic business needs.






Get more insight from your data timely

- Fast, predictable response times for “right-time” analysis
- Accelerate analytic query response times
- Improve price/performance for analytic workloads
- Minimize the need to create data marts for performance
- Highly secure environment for sensitive data analysis
- Transparent to the application

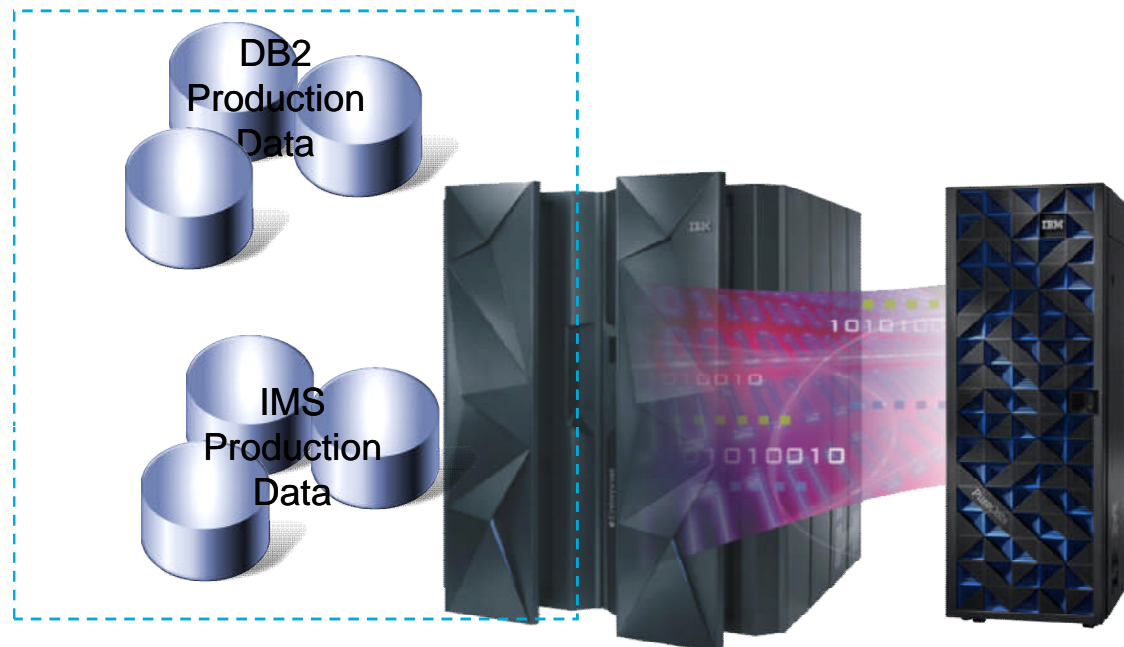
Query execution process flow with DB2 & IDAA



-  Heartbeat (IBM DB2 Analytics Accelerator availability and performance indicators)
-  Queries executed without IBM DB2 Analytics Accelerator
-  Queries executed with IBM DB2 Analytics Accelerator

System z data consolidation

Leveraging your DB2 & IMS information infrastructure



Potential support for IMS

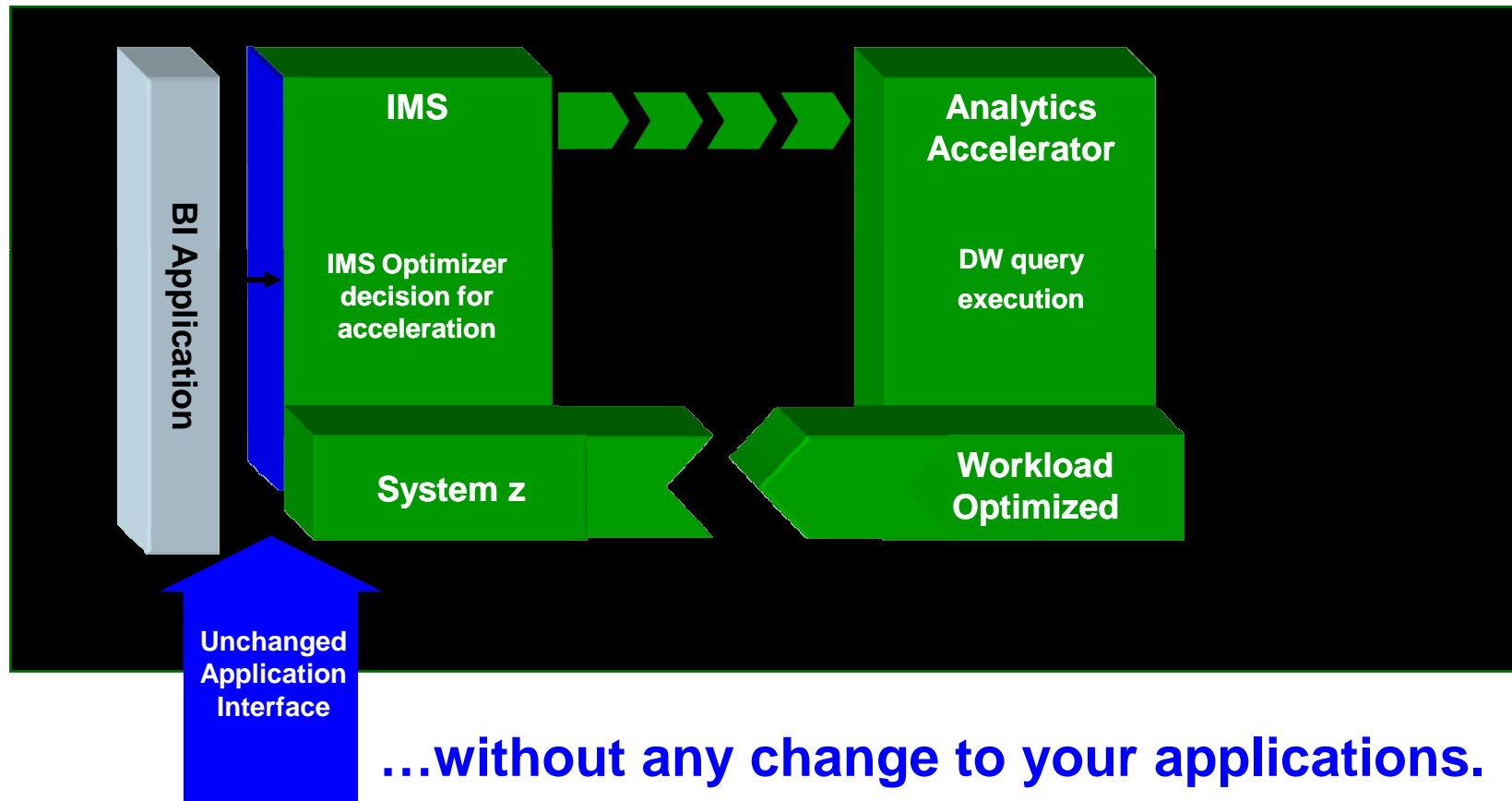
- Organized to support high speed transaction environment with historical and analytic information
- Increase agility by rapidly responding with immediate, accurate information, eliminating the need to search for answers with analysis that is timely
- Enables the organization to become more nimble by placing fact-based information into the hands of decision makers
- Removes the never-ending search for the right information

Another option

Marrying the best of each environment

Potential support

Total solution remains centrally managed by System z...



...without any change to your applications.

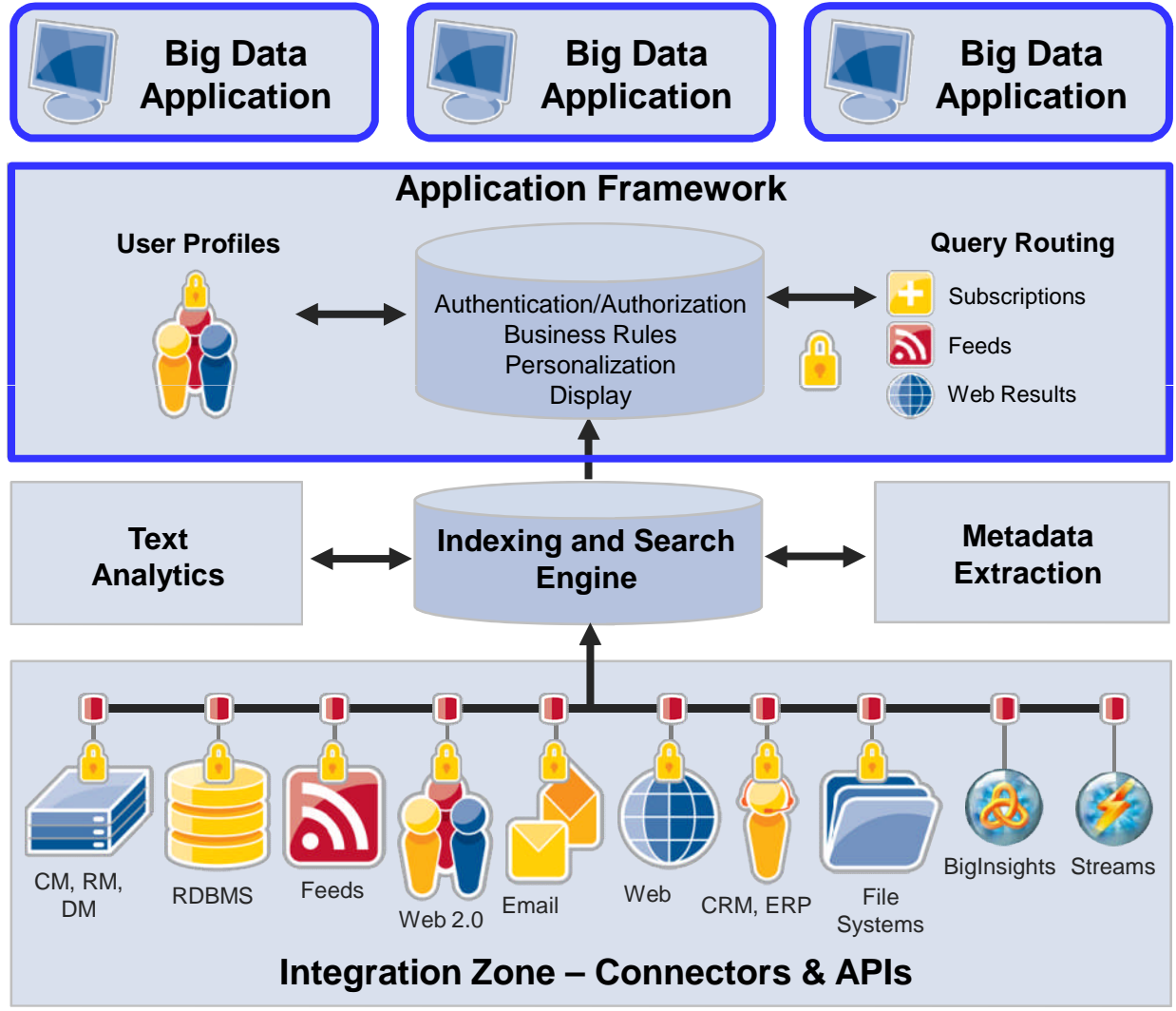
InfoSphere Data Explorer

- ✓ Data Explorer is the **visualization & discovery** capability for IBM's comprehensive **big data platform**
- ✓ Data Explorer is a **key component of all the big data use cases** with greatest impact in **Big Data Exploration & Enhanced 360 View of the Customer**

Data Explorer : visualization & discovery across all your data sources : "Integration at the glass"



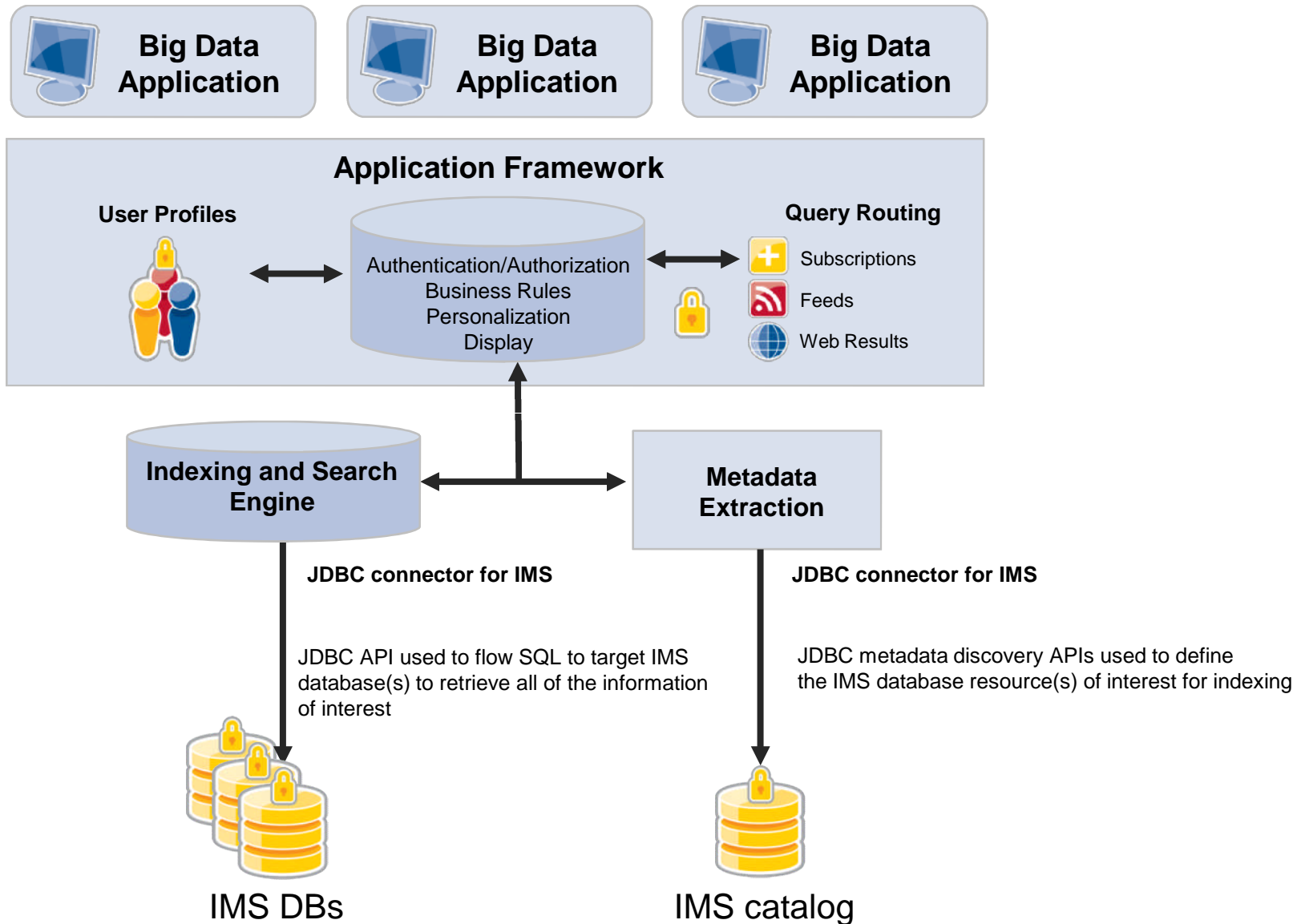
InfoSphere Data Explorer product architecture and differentiators



Differentiators

- ✓ Federated discovery and navigation
- ✓ Scalable architecture
- ✓ Accurate results
- ✓ Secure connectivity
- ✓ Powerful development tools
- ✓ Unique application framework
- ✓ Fast time to value

Seamless IMS integration



IMS + Data Explorer

-Configuring the IMS source

- After deploying the IMS JDBC driver, create a new Database seed

Seed Component: Database (Custom SQL)	
Host	ec01255.vmec.svl.ibm.com
Port	5,555
Username	omvsadm
Password	*****
Database system	IMS
Database name	BMP355
SQL Statement	SELECT * FROM PCB01.HOSPITAL, WARD, PATIENT
Key Column	PATNAME
Advanced Configuration (5)	
JDBC Connection String	jdbc:ims://ec01255.vmec.svl.ibm.com:5555/BMP355:dpsbOnCommit=true;
JDBC Class	com.ibm.ims.jdbc.IMSDriver

IMS + Data Explorer

-Setting up the data transformation

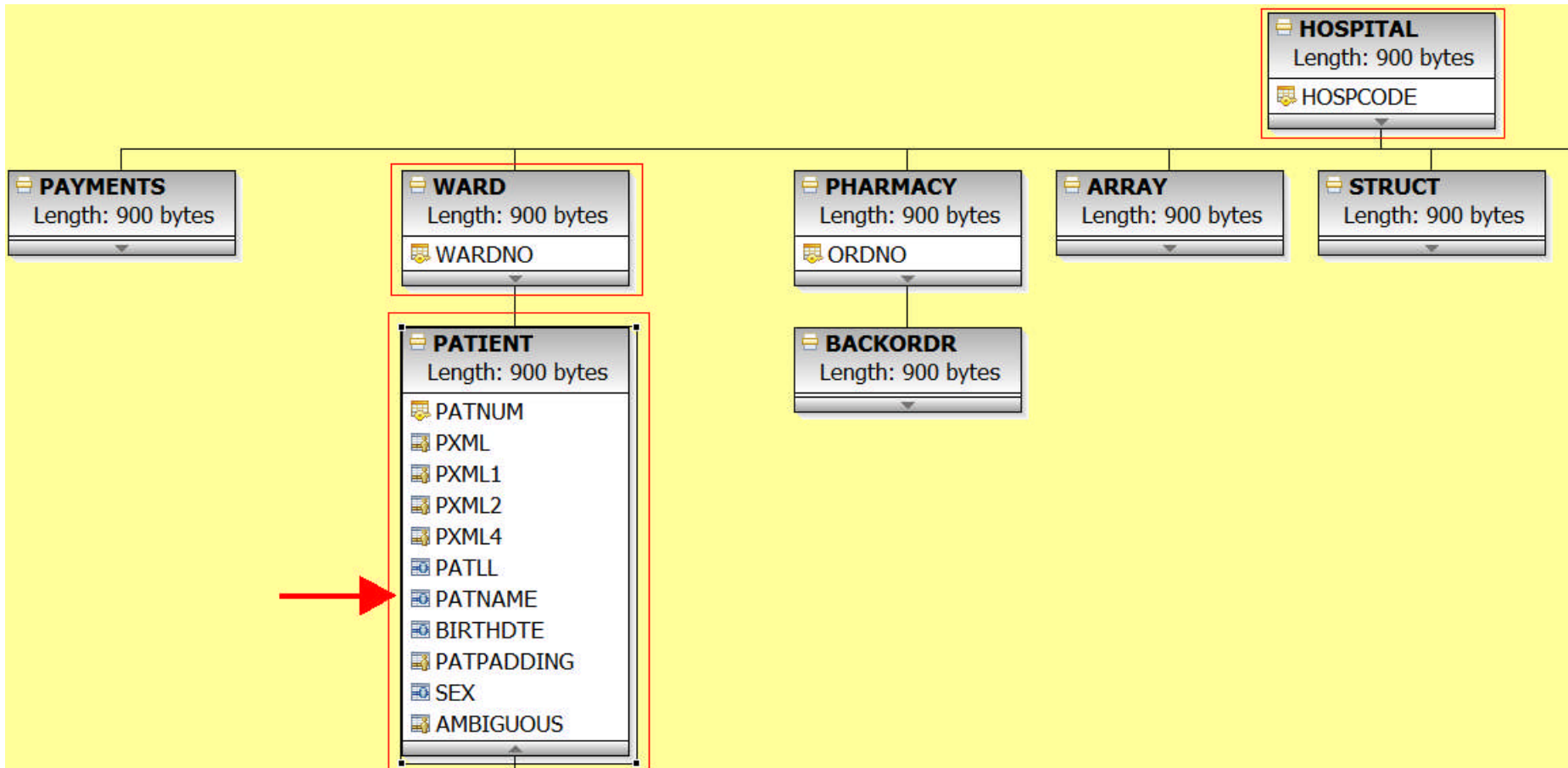
- After creating a new seed, a converter needs to be configured using standard XPATH

Converter Component: [Database seeds support](#)

Type-In	application/vxml-db	▼
Type-Out	application/vxml-db	▼
Fallback	(unset)	▼
Output forking	(unset)	▼
Name	<input type="text"/>	

Original IMS hierarchy for hospital database

- Hierarchy: HOSPITAL->WARD->PATNAME
- Goal: Get a patient centric view



Why use Data Explorer?

- Previously to change the schema so that the PATIENT information is at the top, a logical database needs to be created
- This requires a DBA to be involved and a time window when the new database resources can be brought online
- Data Explorer allows indexes to be created dynamically and for better searching that is not restricted to IMS Segment Search Arguments (SSAs)

Searching the IMS database with Data Explorer

- Query: Who are the patients in the Alexandria hospital

IBM InfoSphere
Data Explorer

Alexandria

Topic Clusters
Not enough text to cluster

- [BOB DAVIS](#) [new window](#) [preview](#)
Hospital_hospcode: R1210010000A, R1210010000A, R1210010000A
Ward_wardno: 0004, 0004, 0004
HospII: 900
Hospcode: R1210010000A
Hospname: ALEXANDRIA
WardII: 900
Wardno: 0004
Wardname: SURGICAL
Patcount: rrr
PatII: 900
Patnum: 0001
[ims://ec01255.vmec.svl.ibm.com:5555/...ey-val=BOB DAVIS - 2K - cache - IMS_BMP355_EC01255](#)
- [KEVIN HITE](#) [new window](#) [preview](#)
Hospital_hospcode: R1210010000A, R1210010000A, R1210010000A
Ward_wardno: 0004, 0004, 0004
HospII: 900
Hospcode: R1210010000A
Hospname: ALEXANDRIA
WardII: 900
Wardno: 0004
Wardname: SURGICAL
Patcount: rrr
PatII: 900
Patnum: 0002
[ims://ec01255.vmec.svl.ibm.com:5555/...ey-val=KEVIN HITE - 2K - cache - IMS_BMP355_EC01255](#)

Searching the IMS database with Data Explorer

- Query: Who are the patients currently in dermatology

IBM InfoSphere
Data Explorer

Dermatology

Topic Clusters

Not enough text to cluster

1. [WILLIAM LI](#) [new window](#) [preview](#)

Hospital_hospcode: R1210020000A, R1210020000A, R1210020000A
Ward_wardno: 0002, 0002, 0002
Hospil: 900
Hospcode: R1210020000A
Hospname: SANTA TERESA
Wardil: 900
Wardno: 0002
Wardname: DERMATOLOGY
Patil: 900
Patnum: 0001

<ims://ec01255.vmec.svl.ibm.com:5555/...ey-val=WILLIAM LI - 2K - cache>

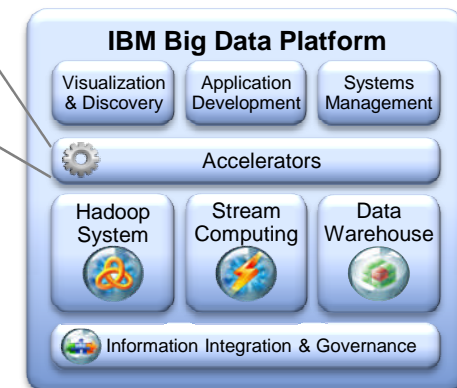
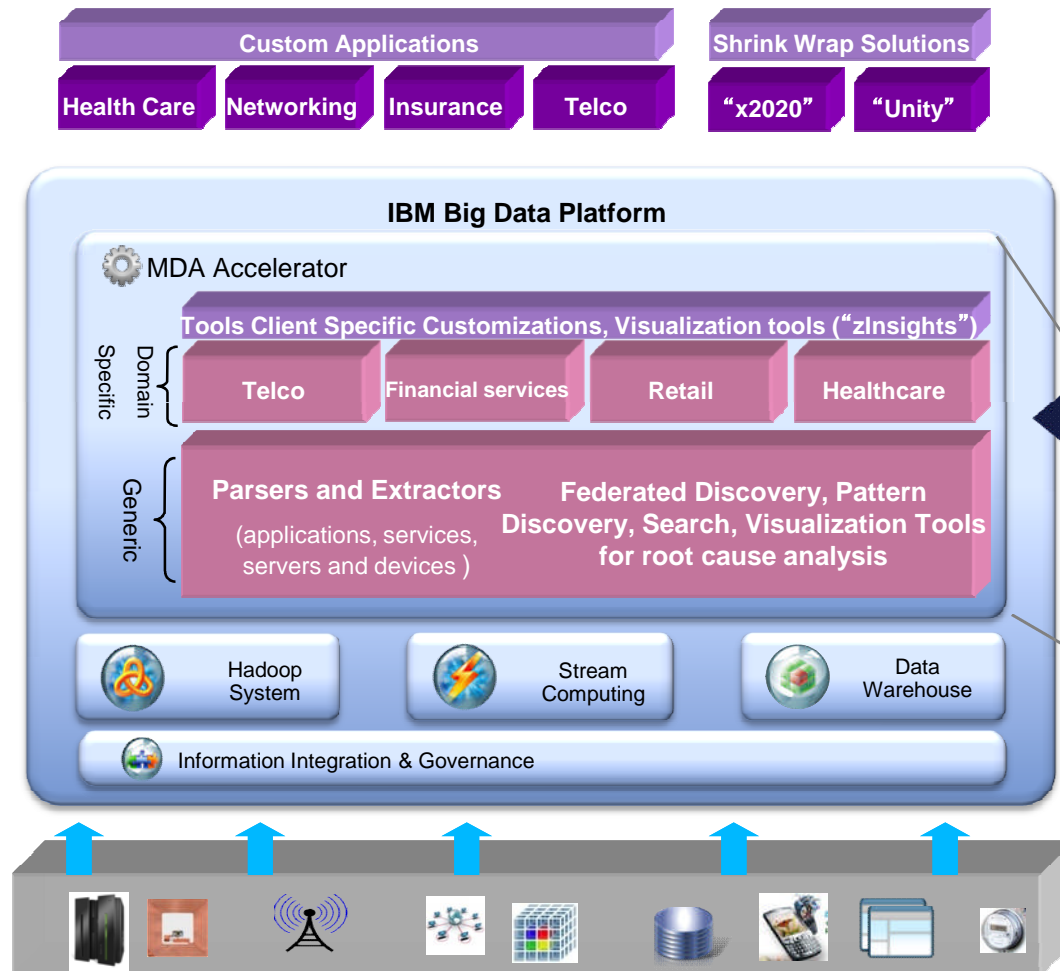
Machine Data Analytics Accelerator

IT use cases:

- Server, performance, troubleshooting

Business use cases:

- Click stream and transaction analysis
- Optimize production, advance planning

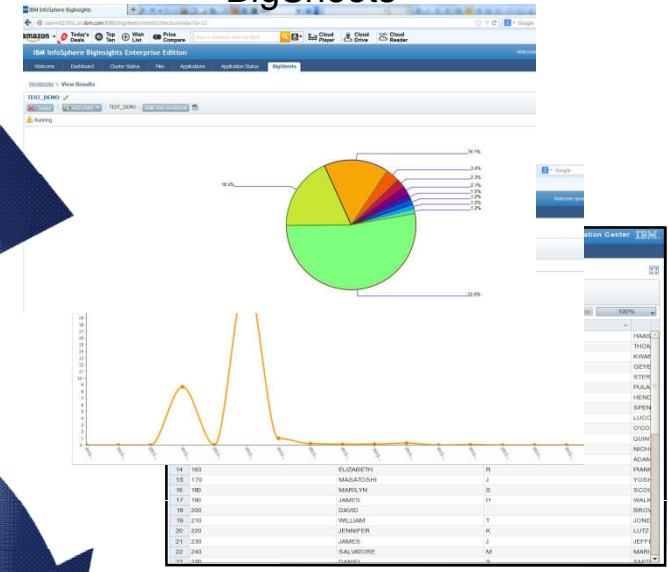


IMS and IBM Accelerator for Machine Data Analytics

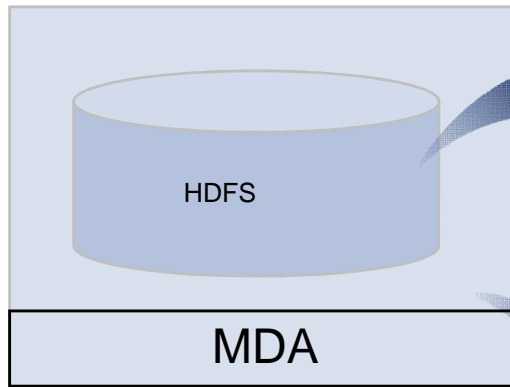
- Consume log data produced by Transaction Analysis Workbench
- Index and link transactions together across products (IMS, DB2, MQ, CICS, WebSphere)
- Make large amounts of IMS transactional log data available to the suite of BigInsights tools.

IMS Log and MDA Overview

BigSheets



BigInsights Platform



Index, Extraction

SFTP

Data Explorer

```

IBM InfoSphere
Data Explorer

Home | Users

Log Type: Results: 2204 results, 0.476 seconds

-- Row (2204)
2013-08-11 22:01:38.244507;CCUTIL ; DFSMTCNT;C9D4E2C6404040000100BA00000000;IMSF ;0.000223;5.770438;5.770661;0;;;;;;DB2A
DFSMTCNT;C9D4E2C6404040000100BB00000000;IMSF ;0.009176;0.041509;0.050685;0;;;;;;DB2A
DFSMTCNT;C9D4E2C6404040000100BC00000000;IMSF ;8.716070;0.536743;9.252959;0;;;;;;
DFSMTCNT;C9D4E2C6404040000100BD00000000;IMSF ;0.044868;0.638912;0.683318;0;;;;;;DB2A
DFSMTCNT;C9D4E2C6404040000100C000000000;IMSF ;24.965797;0.467127;25.432924;0;;;;;;
DFSMTCNT;C9D4E2C6404040000100C100000000;IMSF ;0.130944;0.033681;0.035123;0;;;;;;
DFSMTCNT;C9D4E2C6404040000100C200000000;IMSF ;0.142469;0.952377;0.967167;0;;;;;;DB2A
DFSMTCNT;C9D4E2C6404040000100C300000000;IMSF ;0.302760;1.535334;1.545423;0;;;;;;QY2A
DFSMTCNT;C9D4E2C6404040000100BE00000000;IMSF ;1.011752;0.269375;1.281127;0;;;;;;
  
```

Transaction Analysis Workbench

Conversion to ASCII in CSV format

```

backlog | messagecontext.txt | WSSecurityLibertyPluginInterceptor.java | passwd.txt | fileNames.txt | linuxMachineInfo.txt | imsbig.csv | imscs1.csv | imsd2.csv | imsm1.csv | MSCSV3

1 TIME;TranCode;UserId;RecToken;IMSID;InputQ;Process;TotalTm;CPUtime;FGGets;FFUpdates;FFCalls;FGGets;FFUpdates;TPESAF;ESAFName
2 2013-08-11 22:01:38.244507;CCUTIL ; DFSMTCNT;C9D4E2C6404040000100BA00000000;IMSF ;0.000223;5.770438;5.770661;0;;;;;;DB2A
3 2013-08-11 22:01:44.465656;CCUTIL ; DFSMTCNT;C9D4E2C6404040000100BB00000000;IMSF ;0.009176;0.041509;0.050685;0;;;;;;DB2A
4 2013-08-11 22:04:59.984936;TSSIM ; ;C9D4E2C6404040000100BC00000000;IMSF ;8.716070;0.536743;9.252959;0;;;;;;
5 2013-08-11 22:05:09.223175;AUTOQRY ; ;C9D4E2C6404040000100BD00000000;IMSF ;0.044868;0.638912;0.683318;0;;;;;;DB2A
6 2013-08-11 22:06:01.494244;PDMG2 ;PDMEMPLE ;C9D4E2C6404040000100C000000000;IMSF ;24.965797;0.467127;25.432924;0;;;;;;
7 2013-08-11 22:06:26.796034;P61 ;PDMEMPLE ;C9D4E2C6404040000100C100000000;IMSF ;0.130944;0.033681;0.035123;0;;;;;;
8 2013-08-11 22:06:27.796171;PD3 ;PDMEMPLE ;C9D4E2C6404040000100C200000000;IMSF ;0.142469;0.952377;0.967167;0;;;;;;DB2A
9 2013-08-11 22:06:27.595731;EMAILPR;PDMEMPLE ;C9D4E2C6404040000100C300000000;IMSF ;0.302760;1.535334;1.545423;0;;;;;;QY2A
10 2013-08-11 22:06:23.509104;TSSIM ; ;C9D4E2C6404040000100BE00000000;IMSF ;1.011752;0.269375;1.281127;0;;;;;;
  
```

Transaction Analysis Workbench

- Log conversion

trace.log	messagecontext.txt	WSSecurityLibertyPluginInterceptor.java	passwd.txt	fileNames.txt	linuxMachineInfo.txt	imsbigd.csv	imscics1.csv	imsdb2.csv	imsims1.csv	IMSCSV3
1	TIME;TranCode;Userid;RecToken;IMSID;InputQ;Process;TotalTm;CPUtime;FFGets;FFUpdats;FPCalls;FPGets;FPUpdats;TPESAF;ESAFName									
2	2013-08-11 22:01:38.244507;CCUTIL ;DFSMT CNT;C9D4E2C640404040000100BA00000000;IMSF ;0.000223;5.770438;5.770661;0;;;;;;;;;DB2A									
3	2013-08-11 22:01:44.465656;CCUTIL ;DFSMT CNT;C9D4E2C640404040000100BB00000000;IMSF ;0.009176;0.041509;0.050685;0;;;;;;;;;DB2A									
4	2013-08-11 22:04:59.984936;TSSIM ; ;C9D4E2C640404040000100BC00000000;IMSF ;8.716070;0.536743;9.252959;0;;;;;;;;;									
5	2013-08-11 22:05:09.223175;AUTOQRY ; ;C9D4E2C640404040000100BD00000000;IMSF ;0.044868;0.638912;0.683318;0;;;;;;;;;DB2A									
6	2013-08-11 22:06:01.494244;PDMSG2 ;PDBMPLE ;C9D4E2C640404040000100C00000000;IMSF ;24.965797;0.467127;25.432924;0;;;;;;;;;									
7	2013-08-11 22:06:26.796034;PS61 ;PDBMPLE ;C9D4E2C640404040000100C100000000;IMSF ;0.130944;0.033681;0.035123;0;;;;;;;;;									
8	2013-08-11 22:06:26.796171;PD32 ;PDBMPLE ;C9D4E2C640404040000100C200000000;IMSF ;0.142469;0.952377;0.967167;0;;;;;;;;;DB2A									
9	2013-08-11 22:06:27.595731;EMAILBMP;PDBMPLE ;C9D4E2C64040404000000000C000001C4;IMSF ;0.302760;1.535334;1.545423;0;;;;;;;;;QY2A									
10	2013-08-11 22:06:23.509104;TSSIM ; ;C9D4E2C640404040000100BE00000000;IMSF ;1.011752;0.269375;1.281127;0;;;;;;;;;									

Machine Data Analytics Accelerator

IBM InfoSphere Data Explorer
bootstrap-user | Help

Machine Data
🔍

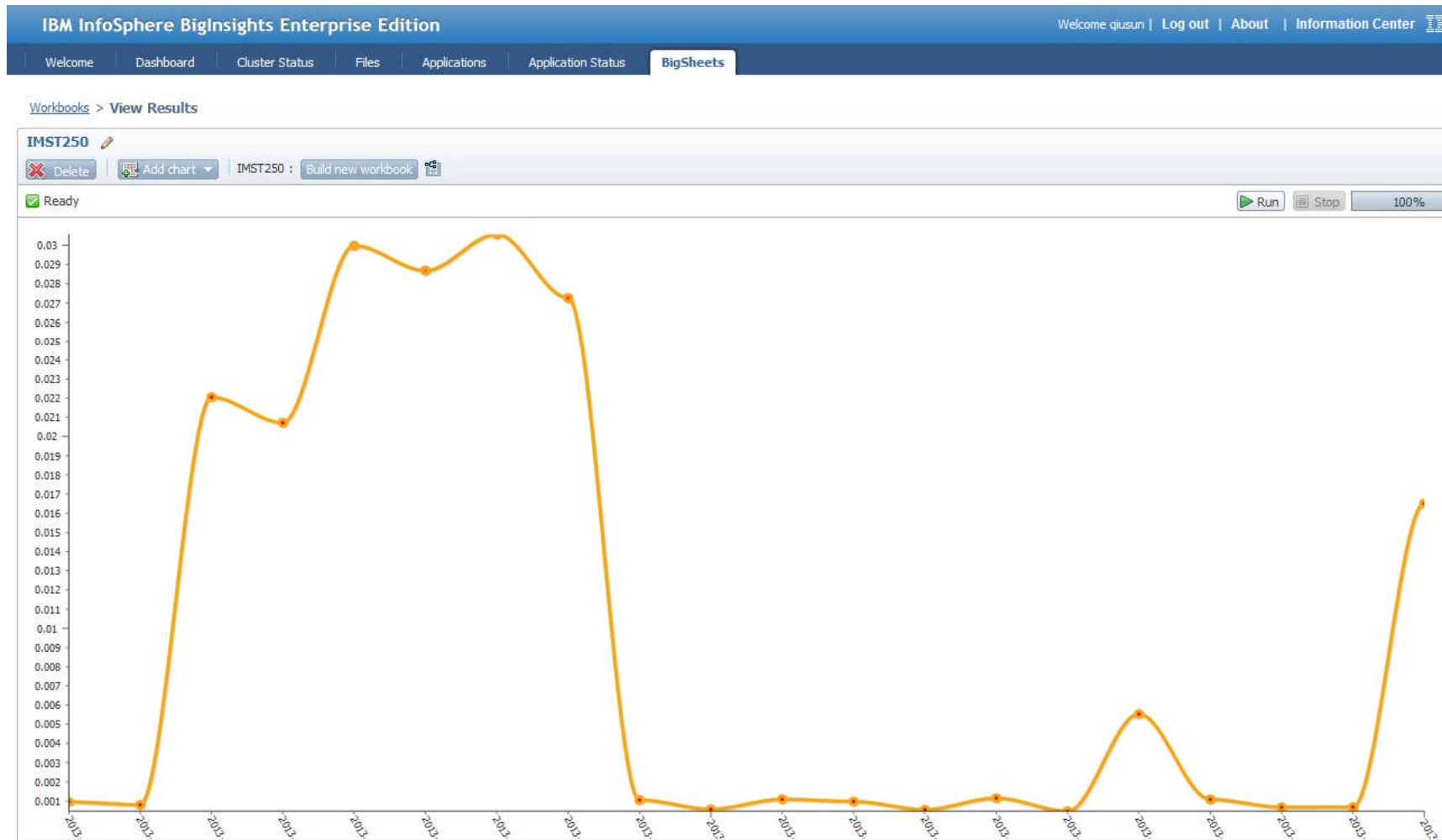
Home
Users

Home >

Log Type	Results for: - (2924 results, 0.067 seconds)
<ul style="list-style-type: none"> ▶ Csv (2,924) 	<p>2013-08-11T22:01:38-07:00 0,DB2A,IMSF,0.000223,5.770438,C9D4E2C640404040000100BA00000000,2013-08-11 22:01:38.244507,5.770661,CCUTIL,DFSMTCNT 284182159d4115f39d277118180e77b6</p>
<ul style="list-style-type: none"> ▶ Imsmq (2,924) 	<p>Add Tag 2013-08-12T02:06:19-07:00 0,IMSF,0.002689,0.094098,C9D4E2C6404040400001029B00000000,2013-08-12 02:06:19.646477,0.097343,TSSIM,437c10379fafc518dccb0741020207a</p> <p>Add Tag 2013-08-12T02:06:19-07:00 0,DB2A,IMSF,0.021295,0.263024,C9D4E2C6404040400001029C00000000,2013-08-12 02:06:19.722494,0.263668,AUTOQRY,a31e9e117248507ea15de6520d21037b</p> <p>Add Tag 2013-08-12T02:06:27-07:00 0,IMSF,0.000758,0.948535,C9D4E2C6404040400001029D00000000,2013-08-12 02:06:27.794712,0.949293,PDMSG2,PDBMPLE,d79ca926b662a68744b01f58d2a9cacc</p> <p>Add Tag 2013-08-12T02:06:28-07:00 0,IMSF,0.171319,0.275632,C9D4E2C6404040400001029E00000000,2013-08-12 02:06:28.561491,0.278735,PS61,PDBMPLE,5de1c08f0dad39453cba2a54aff3d065</p> <p>Add Tag 2013-08-12T02:06:37-07:00 0,IMSF,0.024757,0.342069,C9D4E2C6404040400001029F00000000,2013-08-12 02:06:37.164637,0.366992,PCCBRYS2,X0OPSMVS,5c5fb8bbbbe0a9c90f05c56d0f7092fb</p> <p>Add Tag 2013-08-12T02:07:01-07:00 0,IMSF,0.461442,1.557907,C9D4E2C640404040000102A000000000,2013-08-12 02:07:01.871772,2.019529,PCCBRYS2,X0OPSMVS,2ca0e9535f83699e6ecd28669b5be3a3</p> <p>Add Tag 2013-08-12T02:07:21-07:00 0,IMSF,0.003015,0.003173,C9D4E2C640404040000102A100000000,2013-08-12 02:07:21.855832,0.006188,TSSON,*****,f4d3ffc4fb0f27f5ffd32d5e7a2642d0</p>

Machine Data Analytics Accelerator

– Data Analytics using BigSheets



Thank You for Joining Us today!

Go to www.ibm.com/software/systemz/events/calendar to:

- ▶ **Replay this teleconference**
- ▶ **Replay previously broadcast teleconferences**
- ▶ **Register for upcoming events**