

BIG DATA : PAST, PRESENT AND FUTURE - *AN ANALYST'S PERSPECTIVE*

Carl Olofson : Research Vice President, IDC

Mark Simmonds, IBM Enterprise Architect and Senior Product
Marketing Manager, IBM Software Group

January 21 2014



Webcast Agenda

- Big data in the enterprise – Carl Olofson, IDC
- Making big data a reality - IBM System z Information Management solutions
- Summary and Call to Action
- Q&A





Big Data in the Enterprise

Carl Olofson

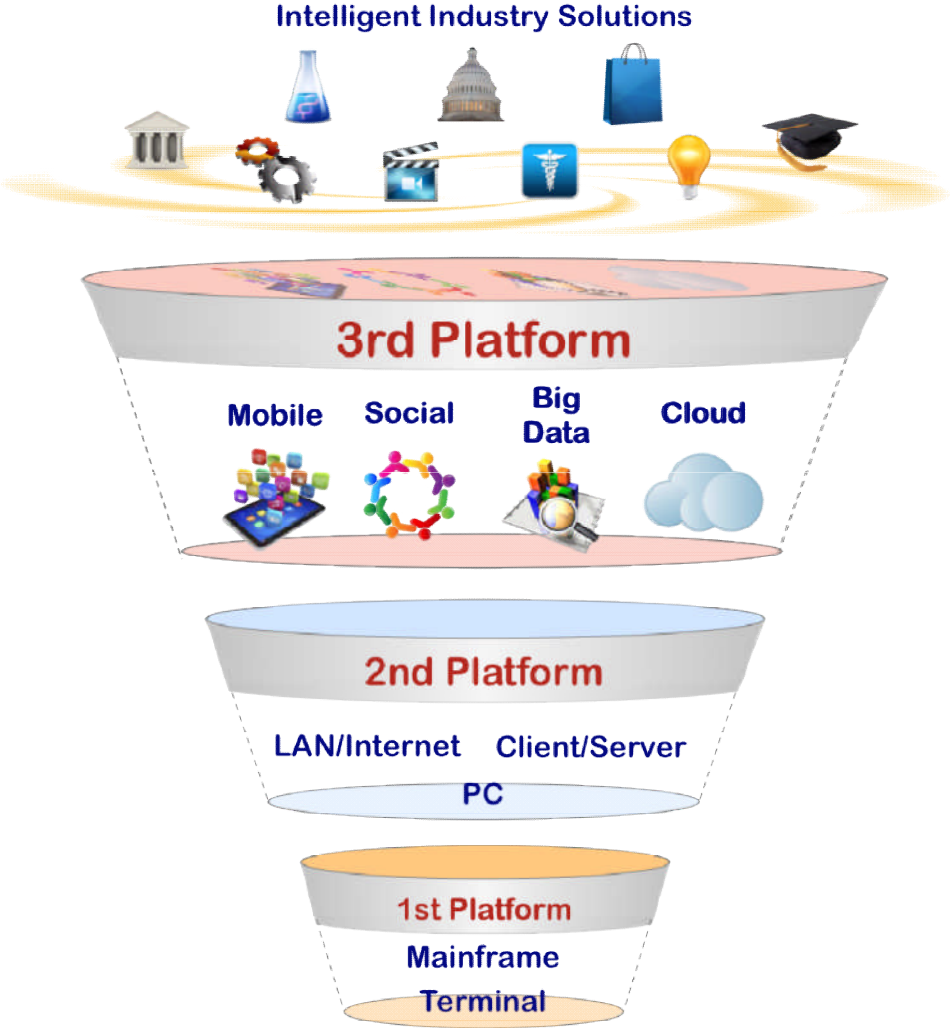
Research Vice President

IDC

Agenda

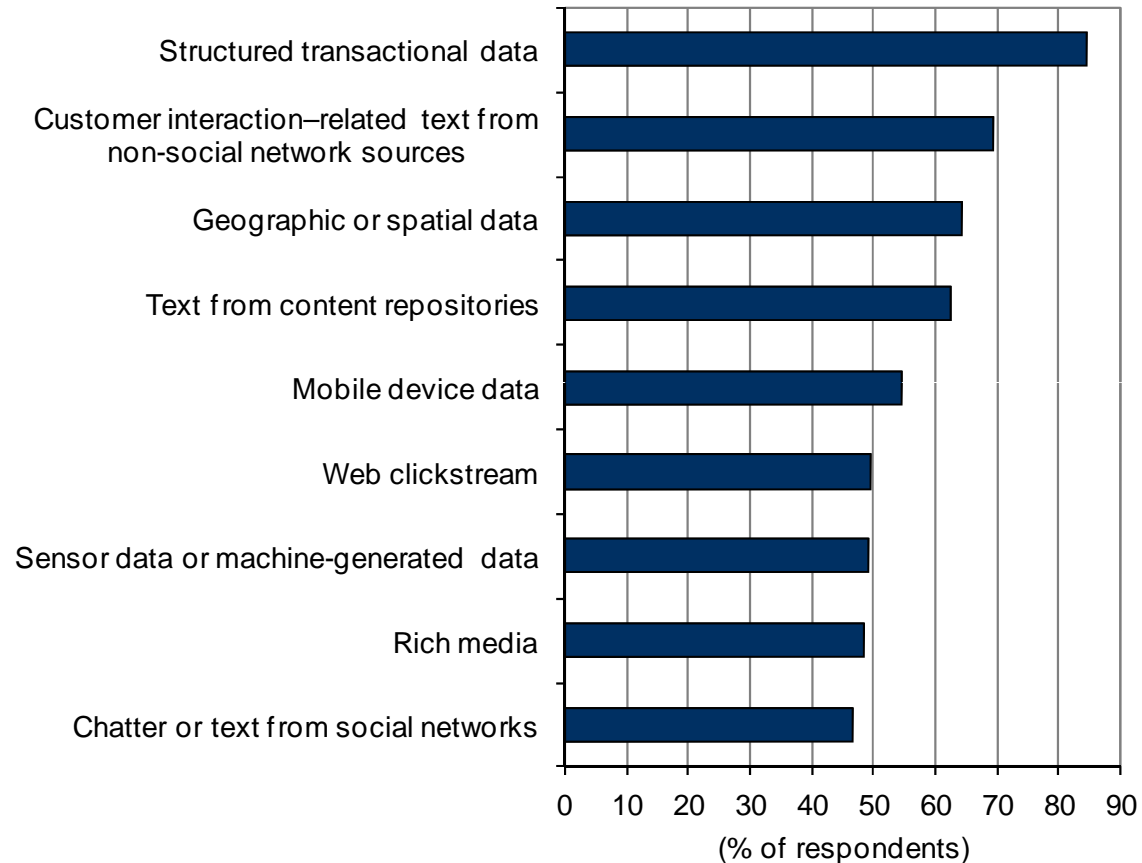
- Market Demand for Big Data
- The Maturing of Big Data in the Enterprise
- Big Data Technologies and Their Uses
- Future Directions
- Conclusions and Recommendations

Market Forces: The 3rd Platform



The Demand for Broad-Based Analytics

Q. *Are the following types of data analyzed in your organization?*



What Does Big Data Mean To You?

CPG company is conducting experiments in digital marketing with online retail and social media partners.

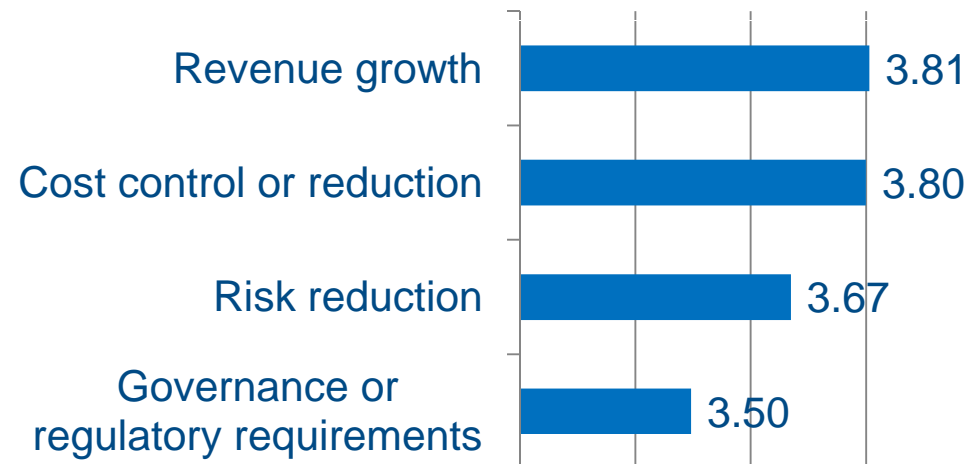
Telecommunications company is improving predictability of customer churn

Manufacturer of wind turbines uses weather data and big data technology to identify optimal locations

Retailer is scoring customers based on their on- and off-line behavior

Credit card company is improving fraud detection capabilities

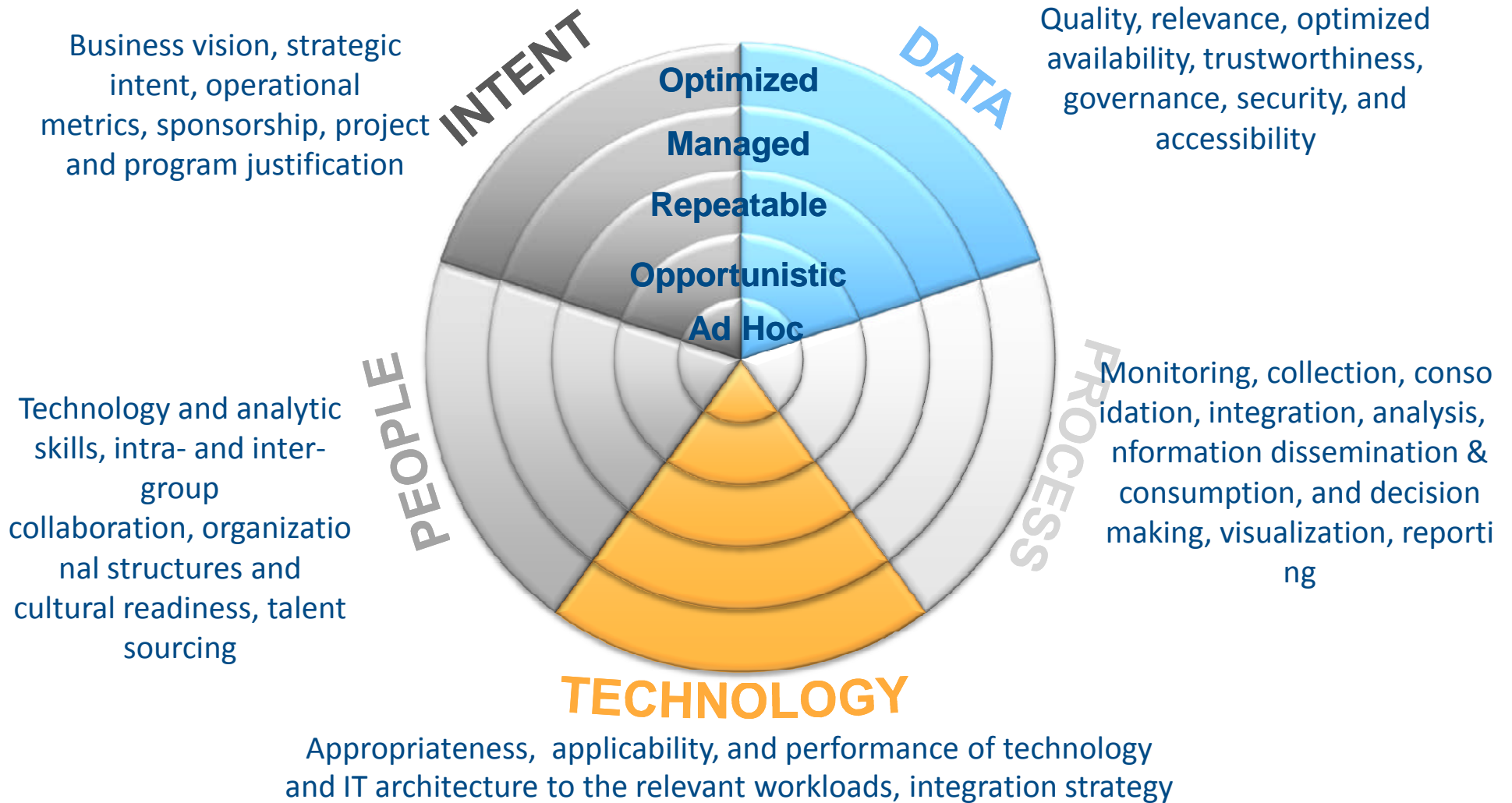
Mean response



How important are each of the following outcomes as drivers of business analytics initiatives at your organization?

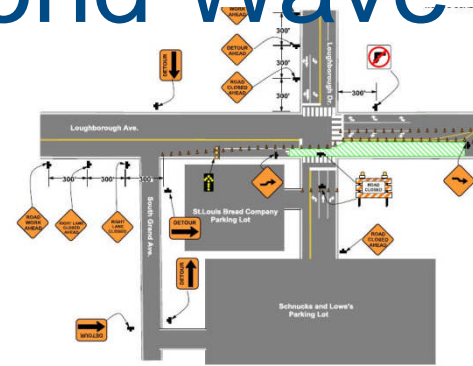
IDC and Computerworld 2013 Business Analytics Survey

IDC Big Data & Analytics Maturity Model



Big Data: The Second Wave

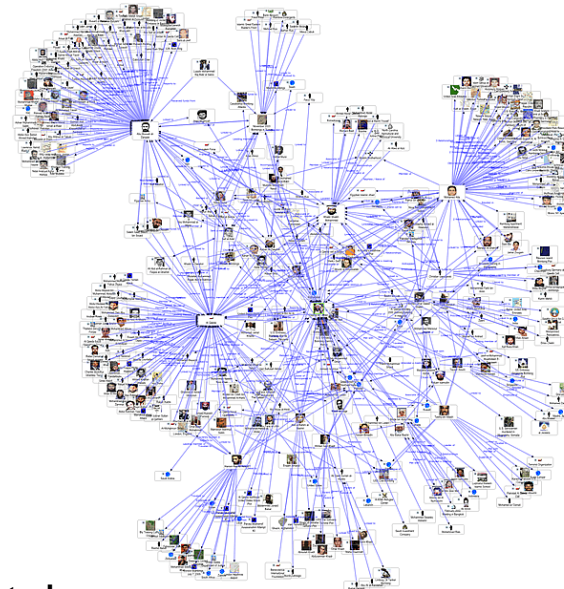
- Hadoop will share the Big Data space with other technologies.
- Some conventional scalable data warehouse (RDBMS) platforms may also have to yield to these newer technologies.
- A whole new workload, based on pattern and relationship analysis will be served
- Hadoop is especially useful for large scale ingestion of unorganized data, and for bulk aggregation and other deep analytics.
- Other technologies will address very demanding analytic workloads not well served by Hadoop or RDBMS.



“NewSQL” for Complex Transactions
(traffic management example)

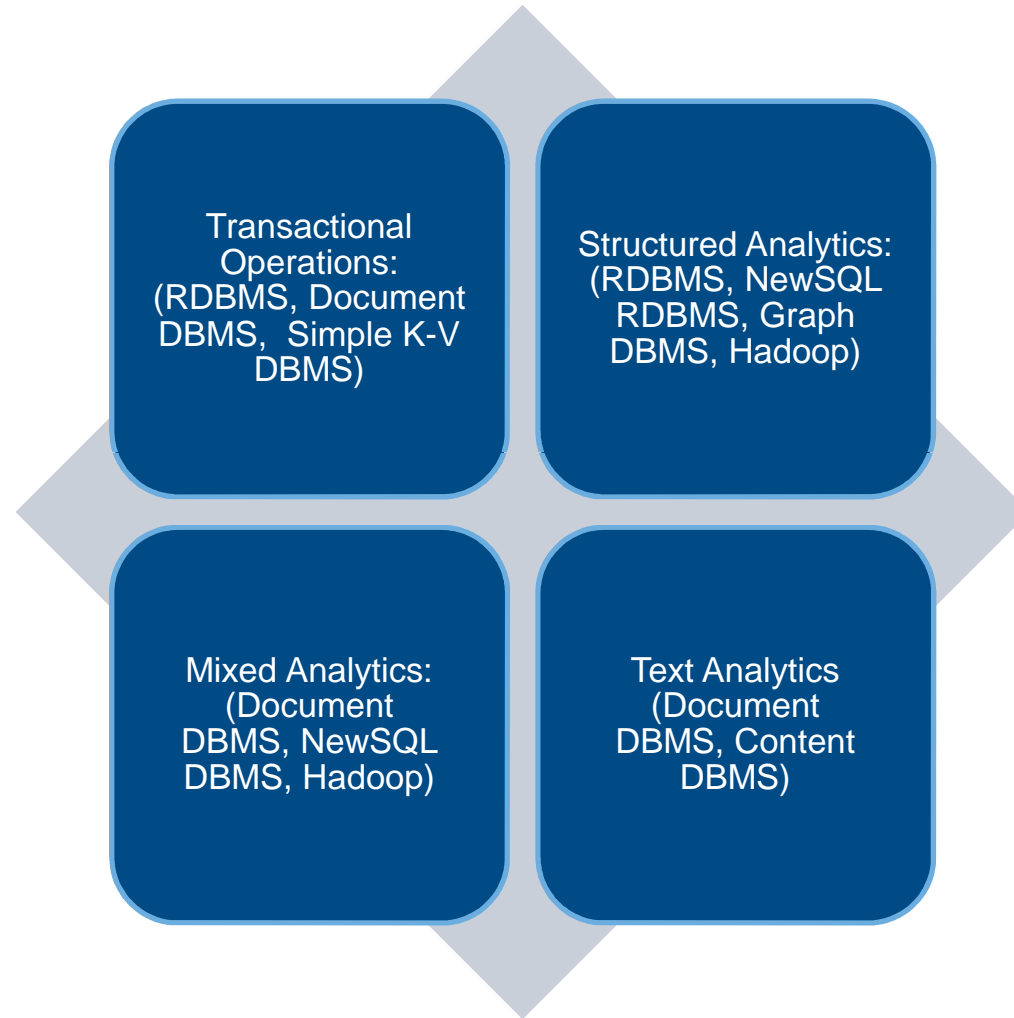


Hadoop for Data Ingestion and Bulk Processing

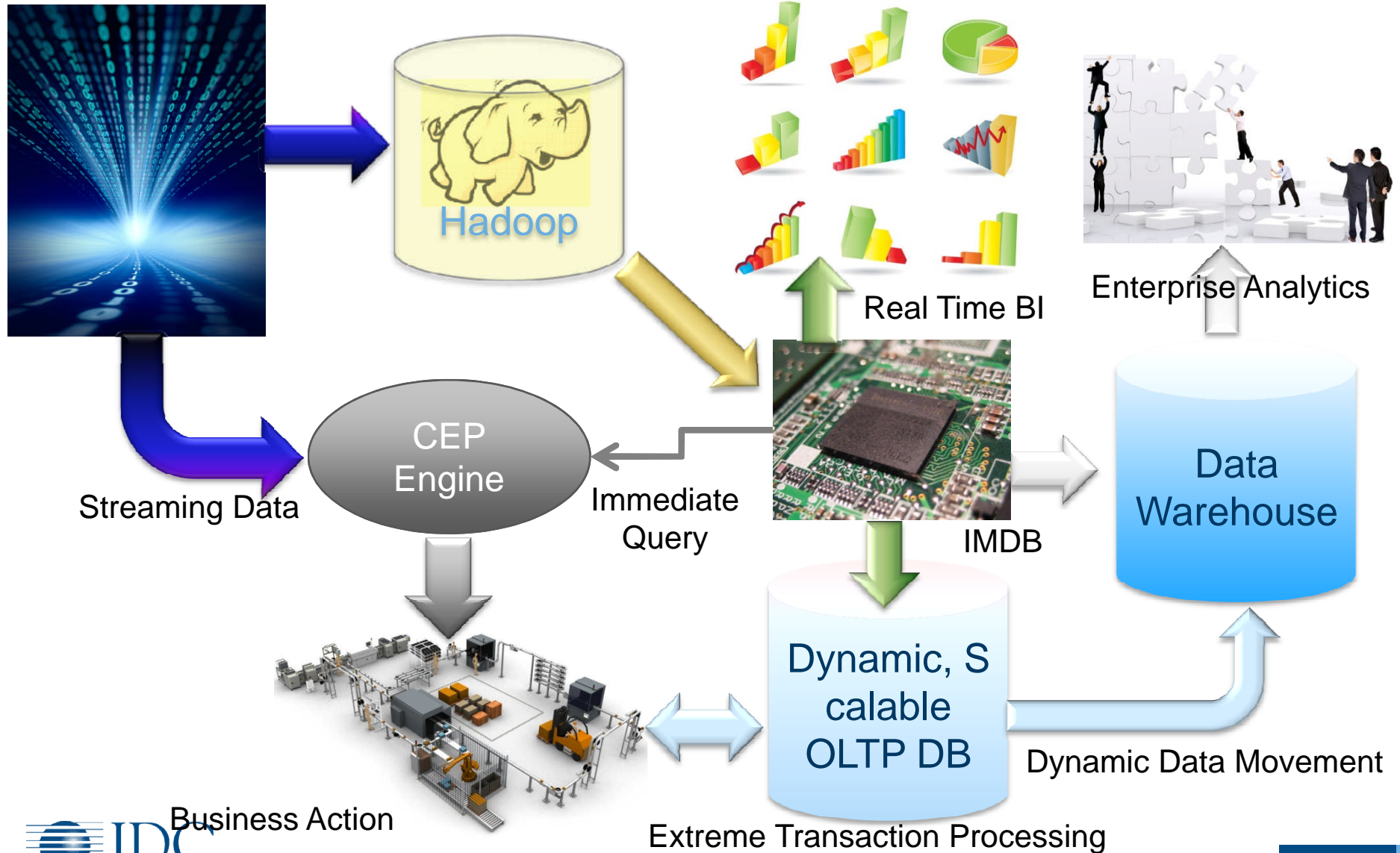


Graph Database

Use Case Classes



Extreme Transaction Processing in the Real Time Enterprise



Important Technology Advances

- RDBMS, Including NewSQL
 - Memory-optimized technology
 - Columnar compression that actually improves performance
 - Vector processing
- Document-Based DBMS
 - Improved compression and IMDB support
 - Improved consistency and recoverability
 - Schema elements that blur the lines with NewSQL
- All DBMS technology (hardware)
 - Advances in memory addressability
 - Non-volatile memory (NVM)

Key Challenges

- **System Level Management**
 - Existing technologies are disjoint.
 - No overall control is possible.
- **Granular Data Security**
 - Sensitive data can be accessed easily.
 - Protection of data at the field level is difficult.
- **Integrated Systems and Functions**
 - Various Hadoop, NoSQL, Graph DB, etc., technologies are not designed to work together or with conventional IT resources.
- **Holistic, Cloud-Based System Architecture**
 - An economical, manageable approach to Big Data requires virtualization, flexibility, and elastic scalability.
 - The ideal is a core platform that exudes these qualities.

Conclusions / Recommendations

■ Conclusions

- We are still in the early days of Big Data deployment.
- Technologies continue to evolve.
- Applications and tools will make these technologies more nearly address real-world workloads with manageability, in time.

■ Recommendations

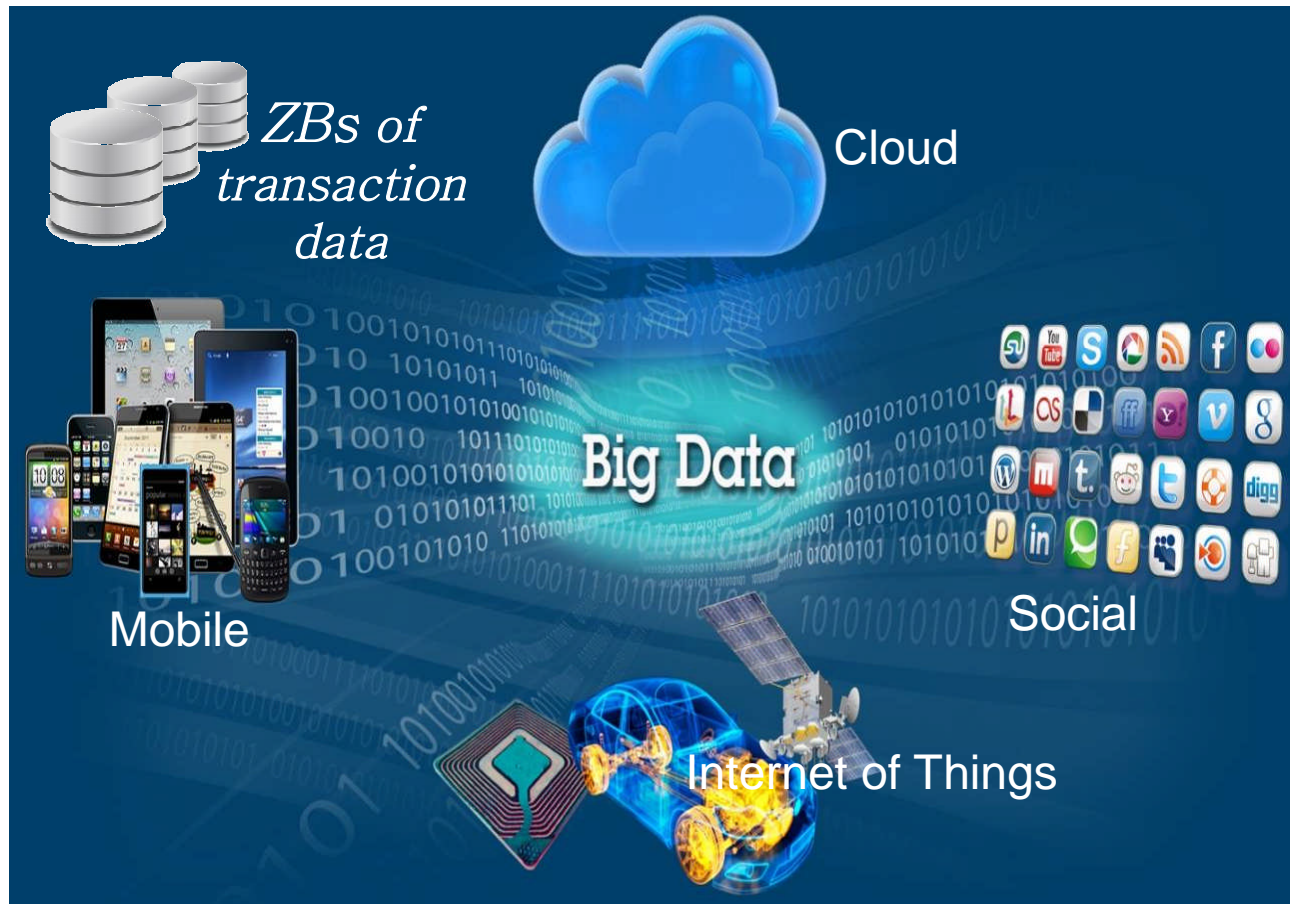
- Learn about all the emerging technologies and determine their applicability to various problem spaces.
- Assess your organization's needs for both deep and real-time analytics, and for large scale data management for extreme transaction processing.
- Develop an architectural vision, and a strategy for getting from where you are to where you want to be.
- Identify a key vendor as your partner in this journey who has the right vision, and can support your objectives.

Webcast Agenda

- Big data in the enterprise – Carl Olofson, IDC
- [Making big data a reality - IBM System z Information Management solutions](#)
- Summary and Call to Action
- Q&A



Big Data – Helping to Create a Smarter Planet



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



Majority of today's analytics based on relational / structured data

- Analytics and decision engines reside where the DWH / transaction data is
- “Noise” (veracity) surrounds the core business data
 - Social Media, emails, docs, telemetry, voice, video, content
- What data are you prepared to **TRUST?**
- Where do you put your trusted Data?



“Circle of trust”



zEnterprise and Big Data

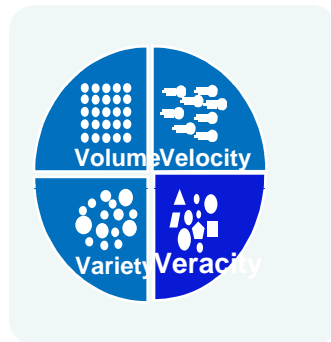
A significant data source for today's business critical analytics

- **Data that originates and/or resides on 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
- **The downtime of an application running on zEnterprise = approx 5 minutes per yr**
- **1,300+ ISVs run zEnterprise today**
 - More than 275 of these selling over 800 applications on Linux

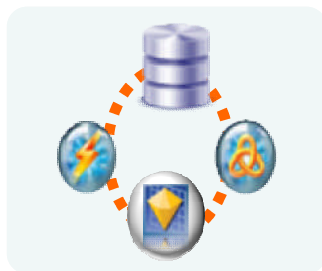


Forward Thinking Organizations are creating value from Big Data

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



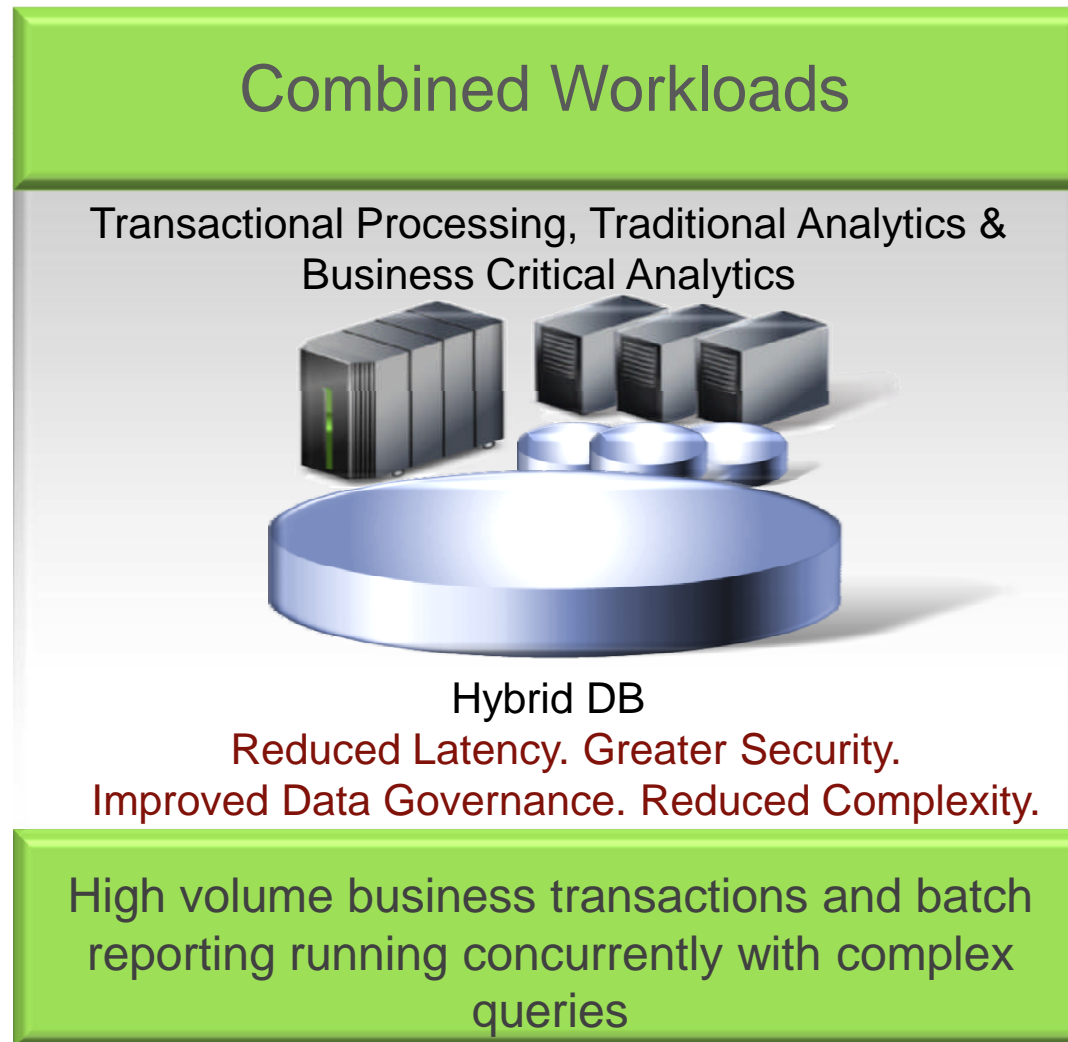
Big Data & Analytics on IBM zEnterprise

Bringing it all together > Driving greater insight



The Hybrid Vision

Delivering **business critical analytics**



Core data management solutions for the 21st Century



DB2 11 for z/OS

Unmatched availability, reliability, and security for business critical information



IMS 13

Delivering the highest levels of performance, availability, security, and scalability in the industry

- **Up to 40% CPU reductions and performance improvements** for (OLTP), batch, and business analytics
- **Improved data sharing performance** and efficiency
- **Integration with InfoSphere BigInsights™ / Hadoop and noSQL support**
- **Improved utility performance** and additional zIIP eligible workload



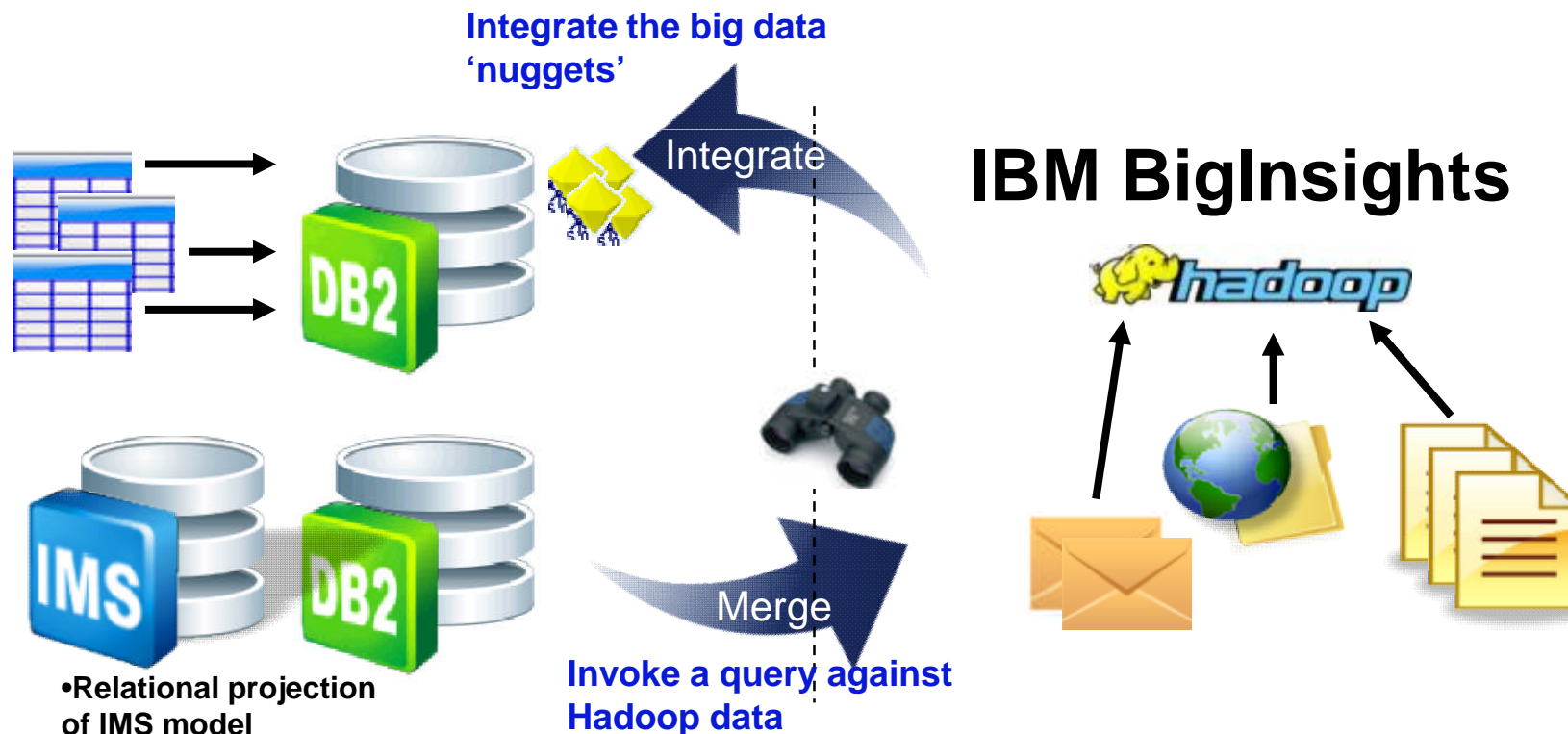
- **Breaking through 100k TPS** 800% greater than IMS 12
- **CPU reductions up to 62% for Java Apps**
- **SQL access to IMS data from both .NET and COBOL applications**
- **Greater flexibility and faster deployment** for new applications with database versioning



Enhancing Big Data Analytics with IMS and DB2 for z/OS

| | |
|---|--|
| <ul style="list-style-type: none"> ▪ Much of the world's operational data resides on z/OS | <ul style="list-style-type: none"> ▪ Unstructured data sources are growing fast |
|---|--|

1. Merge this data with trusted OLTP data from zEnterprise data sources
2. Integrate this data so that insights from Big Data sources can drive business actions
 - IMS & DB2 - connectors & DB capability to allow BigInsights to easily & efficiently access each data source
 - DB2 - connectors & DB capability to allow DB2 apps to easily and efficiently access Hadoop data sources



DB2 11 Support for Big Data

- Goal: integrate DB2 for z/OS with IBM Hadoop based BigInsights Bigdata platform
 - Enabling traditional applications on DB2 z/OS to access Big Data analytics.
- Analytic jobs can be specified using JSON Query Language (Jaql)
 - Submitted to BigInsights
 - Results stored in Hadoop Distributed File System (HDFS).
- A table UDF (HDFS_READ) reads the Bigdata analytic result from HDFS, for subsequent use in an SQL query.
- Must have a variable shape of HDFS_READ output table
 - DB2 11 supports generic table UDF, enabling this function



IBM PureData System for Hadoop

Accelerate Hadoop analytics with appliance simplicity



Accelerate Big Data projects with built-in expertise

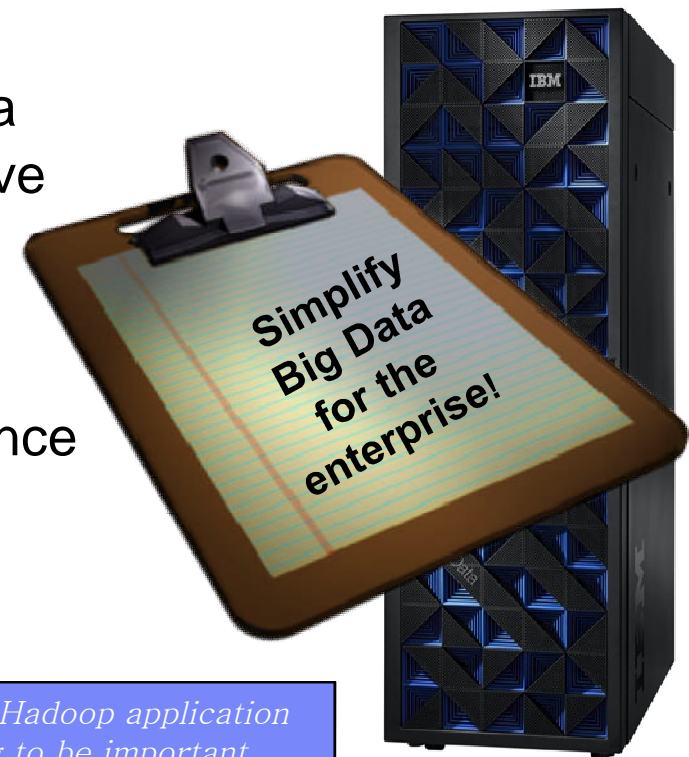
- Explore new ways to use all data
- Unlock new insights from unstructured data
- Establish a cost efficient on-line data archive

Simplify with integrated system management

- InfoSphere BigInsights software
- Compute and Storage hardware

Ensure production grade security and governance

Easily integrate with other systems
in the IBM big data platform



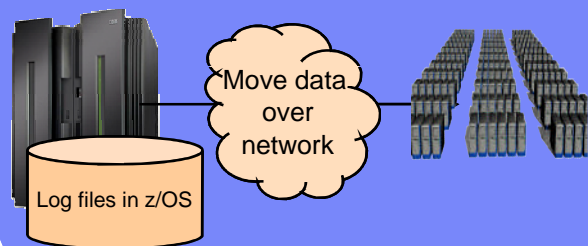
CIO:Insight Apr 29 2013...*Issues surrounding how long it takes to get a Hadoop application into production coupled with a lack of real-time capabilities are proving to be important barriers to deployment. As a result, the respondents are reporting that both the number of Hadoop applications and the size of the overall Hadoop environment remain relatively small.*



Approaches

1

Processing done outside z
(Extract and move data)

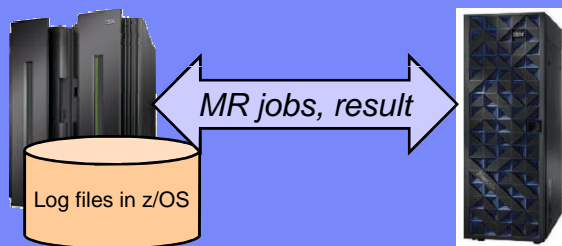


\$\$\$\$

Additional infrastructure.
Challenges with
scale, governance, ingestion.

2

Processing in Appliance
(z remains the master)

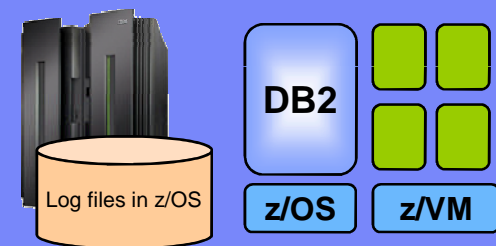


\$\$

Appliance approach with
PureData System for
Hadoop.
High speed load.
z is the control point.

3

Processing done on
System z
(MR cluster on zLinux)



\$\$\$

Provision new node quickly
Near linear scale.
High speed load.
z is the control point.



TECHNOLOGY

Deliver dramatically faster analysis

IBM DB2 Analytics Accelerator 4.1

A high performance appliance that integrates Netezza technology with zEnterprise technology, to deliver dramatically faster business analysis

- Speeds up complex queries up to 2000x
- Lowers the cost of long term storage
- Minimizes data latency
- Improves security and reduces risk
- Complements existing investments



Big Data Innovation

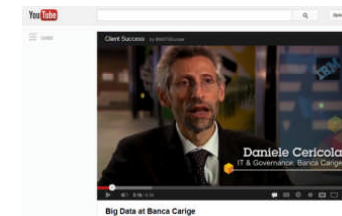
with IBM zEnterprise



Banco do Brasil purchases the **largest ever** DB2 Analytics Accelerator solution to drive customer insight from operational data. The 120-way system can hold 1.28 Petabytes of data. Queries that previously took **11 hours to run now complete in 26 seconds**, over 1500 times faster!



Banca Carige chooses System z to provide real time analytics as part of their Big Data client solution



GPS & sensor information volumes exceeded the capabilities of the existing system. It was redesigned as an enterprise mission critical application using DB2 for z/OS and System z data sharing to now provide the availability and scalability to meet the current and future requirements for this solution.



Webcast Agenda

- Big data in the enterprise – Carl Olofson, IDC
- Making big data a reality - IBM System z Information Management solutions
- *Summary and Call to Action*
- Q&A



IBM understands your big data requirements



IBM understands all kinds of data

- **Game-Changing Innovation** – such as Watson, IBM DB2 Accelerator for z/OS, streaming analytics, Hadoop implementations, expert integrated systems; 20 years of patent leadership, best of breed databases / data warehouses – *DB2 for z/OS, IMS...*
- **Business-Ready Capabilities** – big data and analytics capabilities, integrated and hardened for serious use, with mission and business critical qualities of service unmatched in the industry



IBM knows how to turn data into value

- **Client Expertise** – deep industry know-how and solutions with global reach
- **Strong Ecosystem** – growing investment with 360+ business partners & 100+ universities
- **Build on Current Investments** – enhance existing analytics and information infrastructure with unparalleled breadth and depth of new capabilities



IBM continued investments in big data and analytics processing

- **\$16B+ in Acquisitions** – coupled with game-changing innovations
- **IBM zEnterprise** - The platform with the highest levels of integrity, qualities of service and governance for your business critical applications and processes.
- **Analytics Solution Centers** – visited by 4000+ organizations accessing global expertise



Take Action Now!

Next steps:

- For additional information including whitepapers and demos, please visit:
 - [IBM big data and business analytics](#)
 - www.ibm.com/bigdata/z
- Education
 - Free online education at bigdatauniversity.com
 - 70,000+ registered students
- Community
 - Join Big data and World of DB2 for z/OS on LinkedIn
- Further developments:
 - Watch for future webcast and announcements
- Wanting to experiment on a big data integration project but not sure where to start?
Partner with IBM Silicon Valley and other laboratories nearest you..

[Who We Are](#) [Our Mission](#) [Our Courses](#)



Thank You

