

BIG DATA NETWORK ENVIRONMENTS IDC HPC USER FORUM

Scott Pearson – Director Big Data Solutions

April 29, 2013



Exponential Data Growth – Why Big Data?

Big Data: “A Revolution That Will Transform How We Live, Work, and Think”

90% world’s data produced past two years

15 billion mobile devices by the 2016

2.5 billion Internet Users in 2012

Internet traffic will increase 400% by 2016

Estimated Global Data Volume:

Ⓢ 2011: 1.8 Zettabytes

Ⓢ 2015: 7.9 Zettabytes



iPhone



iPad

What is Big Data?

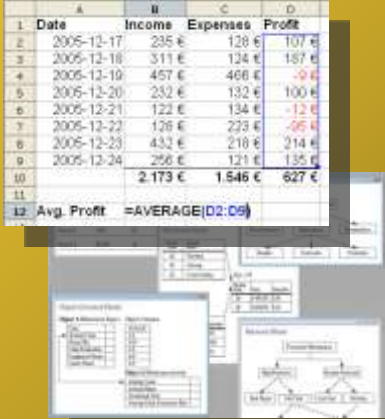
Pervasive & Mainstream

- **Datasets whose size is beyond the ability of typical database software**
- **Horizontal Solution - Crosses all verticals**
- **Uses traditional and non-traditional data sets**
- **Growing 7x faster than the overall IT business**



Sample Data Types and Sources

80% of Data
Unstructured



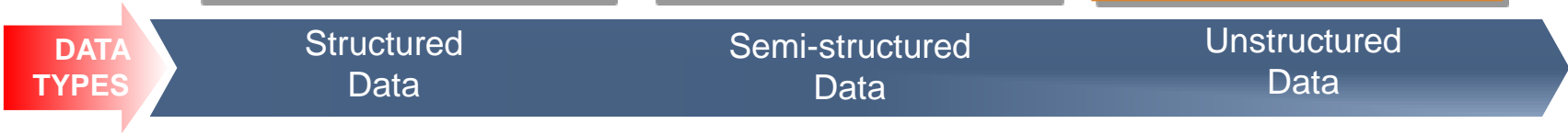
*Spreadsheets
Databases*



*XML
eMail*

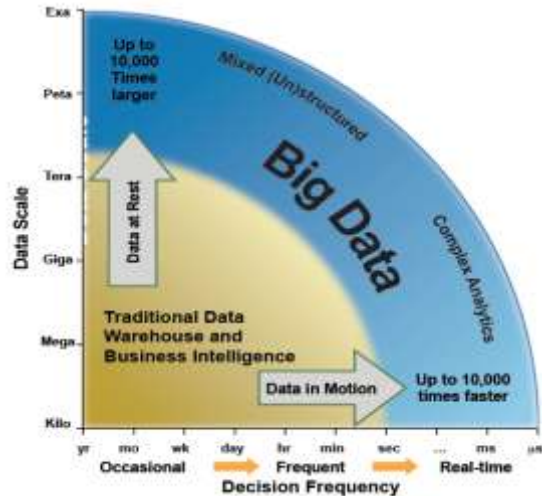


*Voice
Video
Print*



Big Data Analytics Is Comprised Of Two Components

Big Data

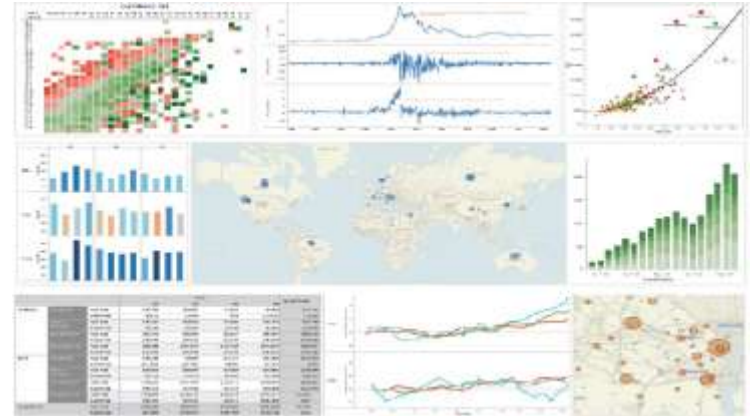


Defined by: Volume,
Velocity, Variety

HIPPO



Data Analytics



Communication of meaningful
patterns in data



“What is Big Data?”

“How do you make money in Big Data?”

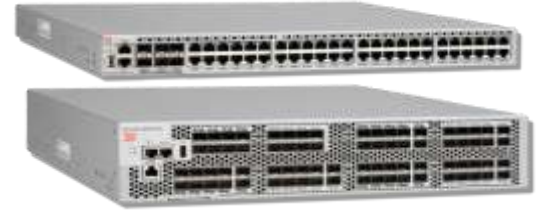
- “.. If they want to drink Merlot, we're drinking Merlot..”



Role of Networks in Big Data

Big Data built on Ethernet Fabrics

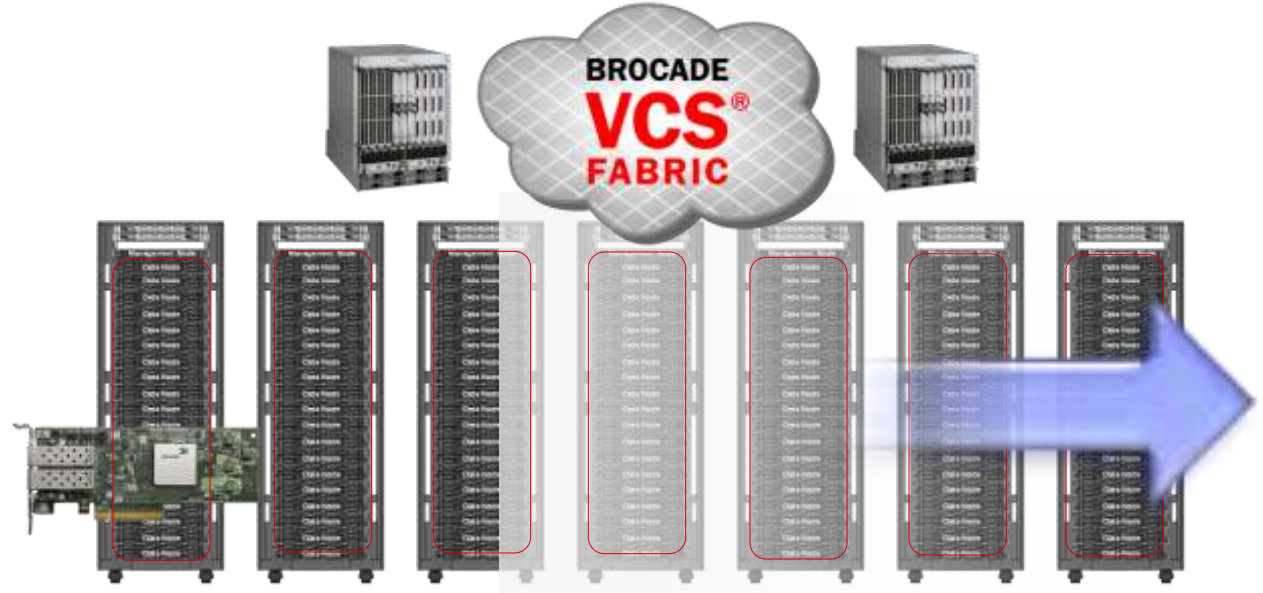
- Networks are crucial to process, transport, analyze and manage Big Data
- Network is key in multi-rack Big Data deployments
- Big Data Deployments Scale Incrementally and are Modular
- Ethernet Fabrics:
 - Flexible, Elastic, and Resilient
 - Self-forming & Smart
 - Self-healing – Feature Rich
 - Easy Management
 - Zero Down-time



Enterprise Grade Platform for Exploring Big Data

Reference Architecture for Big Insights Hadoop with Fabrics

- ⑩ IBM System X Servers
- ⑩ Fabric Adapters
- ⑩ Ethernet Fabric TOR Switches
- ⑩ Fabric Aggregation Chassis Switches



IBM
InfoSphere
BigInsights



Intelligent Cluster Solution

Networks in HPC & Big Data Compute Systems

• Networks Fits:

• Inside the Cluster

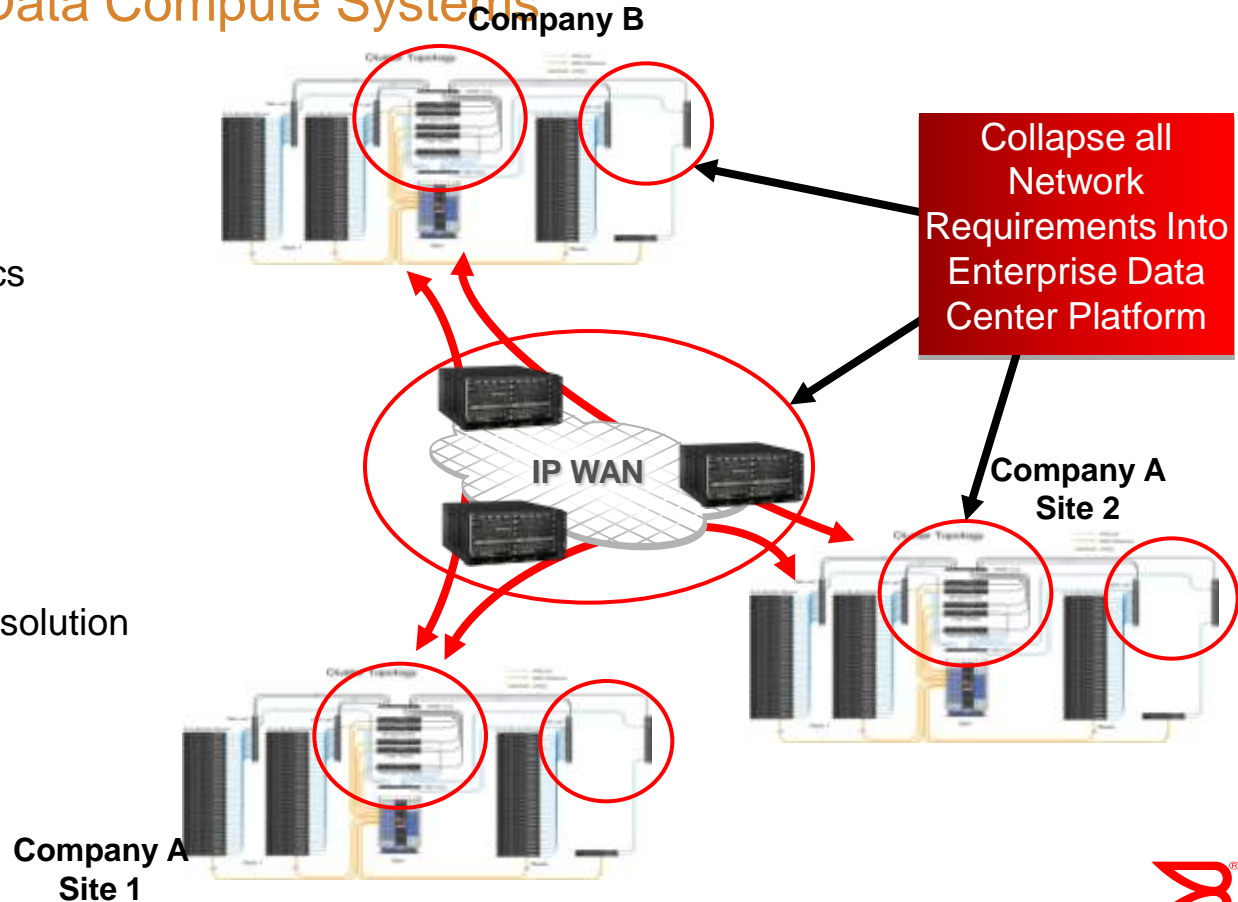
- Management / Service
- Compute Network with Fabrics

• Outside the Cluster

- Data Distribution
- Research / Collaboration

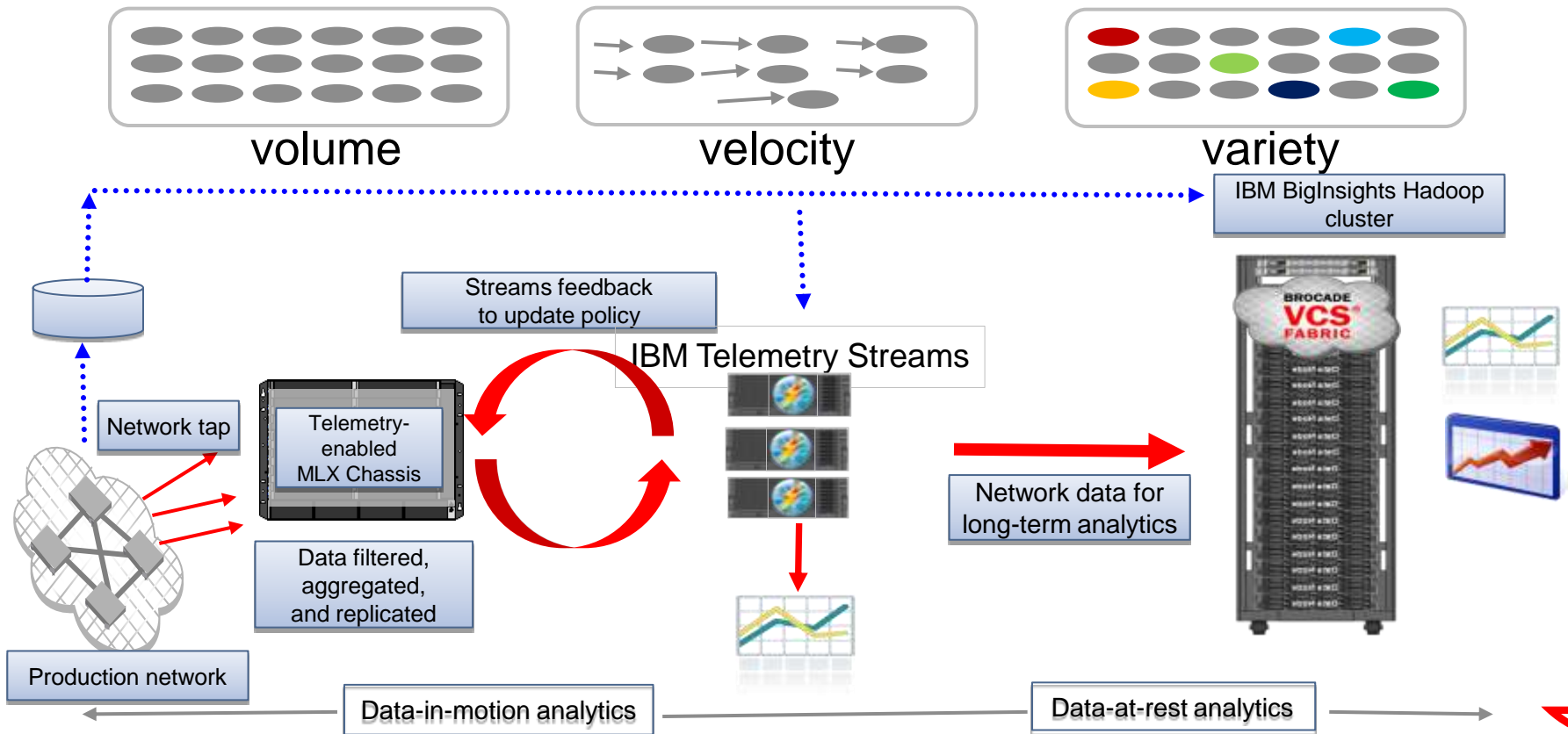
• Network Requirements:

- Modular platforms for enterprise solution
- Highly Scalable
- Full Routing & IPv6
- Jumbo Frames
- Inter-site connectivity



End-to-End Big Data Analytics

Analytics for data-in-motion and for data-at-rest



Networks = Big Data Solutions

Unstructured Data Sources



Data Center



IBM Big Data Platform



InfoSphere
BigInsights



InfoSphere
Streams

Adding Value to Industry



Brocade Big Data Solutions

Launched Big Data Solutions

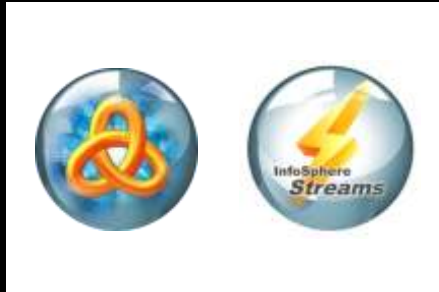
IBM Intelligent Cluster

- ⑩ High Performance Computing
- ⑩ Big Data Compute Systems



IBM BigData Solutions

- ⑩ InfoSphere Streams
 - ⑩ Data in Motion Analytics
- ⑩ BigInsights Hadoop
 - ⑩ Data at Rest Analytics



Hyve Solutions

- ⑩ High Performance Scale Out File System with Intel Lustre



Big Data Solutions Eco-system

1000s of Permutations – Flexibility & Adaptability



Brocade Big Data Story

Present Customers with Guidance and Direction for Analytics Decisions and Deployments

Brocade Big Data Key Tenets

- Networks are an essential part of Big Data Solutions
- Big Data Deployments Scale Incrementally and are Modular
- Ethernet is a primary networking protocol for Big Data Networks
- Entry into the Datacenter via Integrated Solutions
- Eco-system of partners with dozens of Application & Systems providers
- Education component trains people on how to look for value in data
- Provide Partners and Users with Seamless Simplified Flexible Program
- Academic backed (UCSD) Big Data Education Program (**Guidance & Direction**)



SDSC Center for Large Scale Data Systems Research (CLDS)

National Science Foundation (NSF) – Industry research center focused on Data

Sponsors: Brocade, NetApp, Seagate, Mellanox, EMC/Greenplum – **TPC Affiliation**

Objective: Work closely with industry to help solve current and future technology challenges



Workshop for Big Data Benchmarking (WBDB) – Brocade hosted first WBDB in May 2012

Industry Forum on Big Data and Data Value – Brocade hosted November 2012



BROCADE

Current Events and Projects:

- Workshop on Big Data Benchmarking (July 2013 – Xian, China)
- CLDS Big Data Industry Symposium (April 2013 – Brocade, San Jose)
- Workshop on Big Data Benchmarking (October 2013 – Brocade, San Jose)

SDSC
SAN DIEGO SUPERCOMPUTER CENTER

Center Lead – Dr. Chaitan Baru (cbaru@sdsc.edu) & Dr. James Short (short@ucsd.edu)



Brocade Big Data Summary

Mission – Message -- Model

Big Data Mission:

- **Communicate to verticals and markets that Big Data Solutions are built on Ethernet Fabrics Networks via Datacenter Solutions with Partners (Integrators & Systems Companies)**

High Level Message:

- **Ethernet Fabrics (VCS) are essential for Big Data solutions large and small.**
- **Brocade partnered with leading Big Data companies to provide best-of-breed, fully integrated, hardware and software, and supported solutions**
- **Brocade has rich and versatile product portfolio to build Big Data Networks**

Integrator & Systems Partner Centric Sales Model:

- **Identify and recruit Datacenter Integrators for design win and solution delivery**
- **Integrators & Systems Companies deliver Brocade Products via Solutions into the Datacenter**



THANK YOU

