**Definition/Taxonomy Meeting Minutes August 12, 2013**

Big Data Definition

Bob – amount of data exceeded limits of single platform

Scale out horizontally, need an architecture to distribute data and processors

New considerations, reliability, fault tolerance, … dealing with large numbers of commodity processors

What’s new is how to handle the large scale-out, since can’t scale vertically any more

Cost savings at the expense of complexity

Key is horizontal scalability

Question is what broke to cause people to move into a new paradigm?

Analogy to supercomputer days, integrating more processors

Amazon’s approach was to move to a collection of commodity processors

Most people think Hadoop or NoSQL (who are also trying to scale)

Orit – beyond traditional capacities

Many implications – security, processing, etc

Wo – networking is the computing

Parallel processing had to work out messaging (MPI)

Can’t send data across the network because it’s too big

How do you efficiently manage data when it can’t move

Distributed processing – grid computing with multiple nodes

Have to have more efficient data processing

New architecture to handle scalability

Mechanisms for fault tolerance and scalability

Parallel processing of data

Bob – CAP Theorem needs to be referenced in the taxonomy

With the network involved, can’t have all three

You have to determine your trade-off, tune between all three

BASE/ACID

Have to define big data in contrast to these

No joins because of horizontal scale (expensive across the network)

William Miller

Time synchronization – 1588, being extended

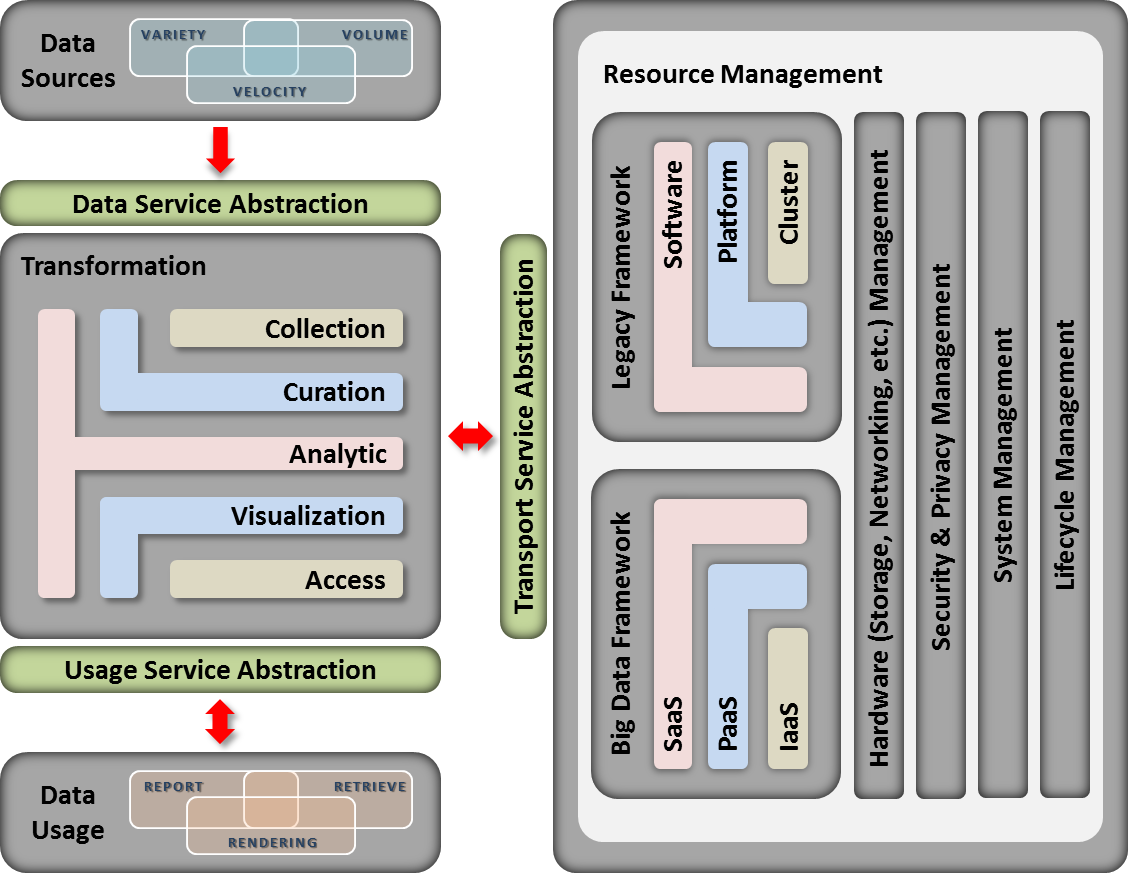
Have to assume you don’t have clocks in sync

Common sense of time, of looking at data, and of expressing data

The machines have to express as convenient for them

Bob – vector clocks should be in the taxonomy

Look at reference architecture draft diagram



Good overall architecture, but insufficient detail to differentiate vendor solutions.**Captured Chat (note second half lost during several site restarts)**

(12:06 PM) Pw Carey (Compliance Partners, LLC): On behalf of my Entire Team....we think you all are doing a good job.....Respectfully yours, Pw

(12:07 PM) Pw Carey (Compliance Partners, LLC): Is your diagram on the NIST BD Web site?

(12:12 PM) Pw Carey (Compliance Partners, LLC)86: It's available at the NIST BD Web Site-M0100 pptx

(12:13 PM) Pw Carey (Compliance Partners, LLC)86: Oops different Ref. Arch....our apologies...

(12:18 PM) Pw Carey (Compliance Partners, LLC)86: And the 'Service Abstraction' is referring to the 3 services available within the Cloud: IaaS, PaaS and SaaS....no?

(12:19 PM) Pw Carey (Compliance Partners, LLC)86: Infrastracture as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS) yes...?

(12:20 PM) Pw Carey (Compliance Partners, LLC)86: And these would be through a Public Cloud, Private Cloud, Community Cloud or a Hybrid Cloud....no?

(12:24 PM) PavithraKenjige (PK Technology)25: there is echo .. can not undestand

(12:25 PM) Pw Carey (Compliance Partners, LLC)86: You can change your microphine settings via the sprocket icon at the top right of your session window....

(12:26 PM) Orit: May be "Legacy" -> "Traditional"

(12:27 PM) Orit: I think that Hadoop "stack" would be unde or use the CC Framework.

(12:31 PM) Pw Carey (Compliance Partners, LLC)86: There is a Cloud Conceptual Reference Model within NIST SP 500-292, not sure if it's still in DRAFT MODE....

(12:40 PM) Pw Carey (Compliance Partners, LLC)86: Good point....such as the following reference...Getting Started with Apache Hadoop on Rackspace Cloud ...

(12:40 PM) Pw Carey (Compliance Partners, LLC)86: Or any other cloud vendor working with Hadoop.....will do....

(12:42 PM) Pw Carey (Compliance Partners, LLC)86: What if the term 'Internet' is becoming extinct and will be replaced by the term..."Cloud"...what if....?

(12:44 PM) Pw Carey (Compliance Partners, LLC)86: And the term 'Network' will be replaced by the term 'Cloud Network', 'Cloud Database', 'Cloud Data Marts'...et cetera.....and how long will this translation require....?

(12:44 PM) Pw Carey (Compliance Partners, LLC)86: Oops...Cloud Data Marts..or Cloud Data Warehouses....

(12:49 PM) Pw Carey (Compliance Partners, LLC)86: We believe you're on the right tack.....