Can Storage Fix Hadoop

John Webster, Senior Partner
Agenda

- What is the Internet Data Center and how is it different from Enterprise Data Center?
- How is the Apache Software Foundation (ASF) addressing the issues?
- What needs fixing from the perspective of Enterprise Storage vendors and the Enterprise Storage world?
- What are the proposed fixes?
- Can Hadoop fix Enterprise storage?
- Can the Internet Data Center/Enterprise Data Center Chasm be Crossed?

FYI: I will use vendor names and products as examples only—no explicit or implied endorsements
The Data Center Chasm

Internet Data Center

Enterprise Data Center
Defining the Data Center Chasm

Internet Data Center
- Embraces open source
- Automates IT
- Comfortable with systems that run in “failure mode”
- “Cheep and deep” – hardware inefficiency not an obvious issue
- More willing to build their own systems and self-support
- Manages storage from a systems perspective

Enterprise Data Center
- Prefers proprietary but learning open source
- Approaches IT automation conservatively
- Doesn’t get “failure mode”
- Hardware efficiency-conscious
- More willing to buy from proprietary vendors and deal with them for support
- Sees value in storage environment as a place for data and storage management
The Datacenter as a Computer

An Introduction to the Design of Warehouse-Scale Machines
Second Edition

Luiz André Barroso
Jimmy Clidaras
Urs Hölzle

Synthesis Lectures on Computer Architecture
Mark D. Hill, Series Editor
What has the ASF Fixed in HDFS?

- **NameNode SPOF**
  - NameNode active/standby failover support

- **Snapshot**
  - Read-only Copy on Write (COW) included in latest v2 Beta (2.1.0)

- **NFS support**
  - Support for NFSv3 in latest v2 Beta (2.1.0)

- **DR Support**
  - Distributed Copy (distcp)
What Needs to be Fixed – the Enterprise Storage Vendor Perspective

- Hadoop NameNode is a single point of failure in V1. Manual failover in v2 (Beta).
- JobTracker is also a single point of failure.
- For data integrity and protection, HDFS creates three full clone copies of data:
  - 3x the storage for each file – slow and inefficient
  - If all three copies are corrupted, you’re still hosed (reload and start over)
- 60% of Enterprise Hadoop projects fail or are put on hold.
- Steep learning curve—six months is not uncommon for those that actually go from pilot to production.
- No storage tiering.
- Limited (if any) ways to respond to corporate security and data governance policies.
- Difficult to move between cloud and data center.
- Fundamentally a batch process.
- Data in/out processes can take longer than the actual query process.
- Inability to dis-aggregate storage from compute so that the two can be scaled independently.
- Dearth of applications built on top.
- Dearth of people available in the job market to run this beast and the ones that can go for big bucks.
- ....and more leading some analysts to believe that Big Data has entered the “trough of disillusion”
What Needs to be Fixed— the Enterprise Storage Vendor Perspective

- **Hadoop NameNode is a single point of failure**
  - JobTracker is also a single point of failure
- **For data integrity and protection, HDFS creates three full clone copies of data**
  - 3x the storage for each file – slow and inefficient
  - If all three copies are corrupted, you’re still hosed (reload and start over)
- **60% of Enterprise Hadoop projects fail or are put on hold**
- Steep learning curve—six months is not uncommon for those that actually go from pilot to production
- No storage tiering
- **Limited (if any) ways to respond to corporate security and data governance policies**
- Difficult to move between cloud and data center
- Fundamentally a batch process
- **Data in/out processes can take longer than the actual query process**
- **Inability to dis-aggregate storage from compute so that the two can be scaled independently**
- Dearth of applications built on top
- Dearth of people available in the job market to run this beast and the ones that can go for big bucks
- ....and more leading some analysts to believe that Big Data has entered the “trough of disillusion”
Hadoop External Storage - EMC Isilon Example

- Shared storage replaces node-level DAS
- HDFS implemented as “over the wire” protocol on OneFS
- Isilon cluster nodes emulate NameNodes and DataNodes
- NameNode SPOF eliminated
- Decoupled storage and compute layers
- Data protection and DR by OneFS
Hadoop External Storage - NetApp Example

- Preserves shared nothing architecture and HDFS
- Decouple compute and storage
- Hardware RAID: reduction in copies from 3 to 2
- NameNode metadata in separate array for faster NameNode recovery
- Datanode drive failures do not “blacklist” the Datanode
- Apply built-in enterprise data and storage management functionality to Hadoop data

Source: NetApp
Shared Storage as Secondary Storage

Network Layer

Compute Layer

Primary Storage Layer

Secondary Storage Layer

SAN/NAS Secondary Storage Layer
Data mirrored or migrated from primary to secondary storage
Storage services also live here
Progression of Hadoop @ Yahoo!

Source: Yahoo!
Is Hadoop a new Storage Platform?

No
- It’s a distributed computing platform for analytics

Yes
- HDFS - Embedded, distributed file system (like scale-out NAS)
- Data protection and management built-in (like Enterprise Storage)
- Storage performance at scale and low cost and with native intelligence
- Growing use case as data repository for existing enterprise BI and Data Warehousing apps – the “Data Lake”
What Does the Enterprise Want from Big Data?

“If we could harness all of our data, we would be a much stronger business.”*

* From CompTIA survey where two thirds of respondents either agreed or strongly agreed with the statement
Can the Chasm Be Crossed?

Internet Data Center

Enterprise Data Center