HPC & Big Data: Trends to Watch

Dave Ellis
HPC Architect
U. S. Public Sector, NetApp, Inc.

September 3, 2012
What is “Big Data”?

“Big Data” refers to datasets whose volume, speed and complexity is beyond the ability of typical tools to capture, store, manage and analyze.

Coined by Francis Diebold, professor of economics at the University of PA in 2000, when “Big” meant Gigabytes / day\(^1\)

Big Data Solution Portfolio

Analytics
Insight from extremely large datasets

Big Data

Content
Secure boundless data storage

Bandwidth
Performance for data intensive workloads
Quantifying The Big Data Challenge

60 Zettabytes
Estimated size of the digital universe in 2020

5 Billion
Smart phones

30 Billion
Pieces of new content to Facebook per month

1000 PB – Some perspective…

- > 20 “Normal” Data Centers
- > 750 Million Customers
- > 48,000 Emails / Second
- > 1 Million Disks
- 24 Full Time Admins (> 40 PB/FTE)
Experience Managing Data at Scale

NetApp’s Largest Customer 1000 Petabytes

- 100 PB, 4 Customers
- 50 PB, 10 Customers
- 20 PB, 50 Customers
- 10 PB, 100 Customers

© 2012 NetApp, Inc. All rights reserved.
State-of-the-Science in HPC Today

**LLNL Sequoia** – Currently the fastest supercomputer and storage combination on the planet

- 20 Peak PetaFLOPS \((20 \times 10^{15})\)
- 1.6 Million Cores
- 1.6 PB main memory
- **55 PB (usable) NetApp E-Series storage**
  - Lustre Cluster Parallel File System
  - QDR (40 Gigabit) Infiniband
  - > 1 Terabyte/sec aggregate bandwidth
- ~8 MW of Power
- Simulations for
  - Nuclear Weapons Viability
  - Counter Terrorism
  - Energy Security
  - Climate Change

Now #1 of the Top 500 HPC Systems in the world.

http://top500.org

Dispelling the Misconceptions About Big Data
Big Data Is NOT New

National Oceanographic and Atmospheric Administration

3.5 billion observations per day from NOAA sensors

6-hour Update Cycle

Global Atmospheric Model (HIRAM)

Other High Res Specialized Models:
- Hurricane (3 km)
- Thunderstorms
- Tornadoes
- Fire Weather
- Ocean Models
- Volcanic Ash
- Etc.

HPC Models

Forecast & Warning Guidance to Public and Private Sector Forecasters

15 million information products per day

30 PB of New Data Annually

National Weather Service Data Stream

© 2012 NetApp, Inc. All rights reserved.
Big Data = Big Analytics = Hadoop?

- That’s What The Media Hype Implies, but it is NOT true!
- Traditional analytics (BI/DSS/DW) dominates the analytics market
- Like other technologies vying to gain broad adoption in Enterprise IT (e.g., Traditional Analytics, HPC & Cloud), it shows promise
## Worldwide Market Numbers

<table>
<thead>
<tr>
<th></th>
<th>2011</th>
<th>2012</th>
<th>2013</th>
<th>2014</th>
<th>2015</th>
<th>2016</th>
<th>CAGR</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hadoop</td>
<td>0.077</td>
<td>0.123</td>
<td>0.198</td>
<td>0.317</td>
<td>0.507</td>
<td>0.812</td>
<td>60%</td>
<td>Source: IDC - <a href="http://bit.ly/S9FcP">http://bit.ly/S9FcP</a></td>
</tr>
<tr>
<td>All Analytics</td>
<td>31.8</td>
<td>34.9</td>
<td>38.3</td>
<td>42.1</td>
<td>46.2</td>
<td>50.7</td>
<td>9.8%</td>
<td>Source: IDC - <a href="http://bit.ly/S9FcP">http://bit.ly/S9FcP</a></td>
</tr>
<tr>
<td>HPC</td>
<td>27.5</td>
<td>29.4</td>
<td>31.4</td>
<td>33.6</td>
<td>36.0</td>
<td>38.5</td>
<td>7.0%</td>
<td>Source: Intersect360 - <a href="http://bit.ly/S9bQWf">http://bit.ly/S9bQWf</a></td>
</tr>
<tr>
<td>Public Cloud Services</td>
<td>17.0</td>
<td>22.5</td>
<td>29.4</td>
<td>36.6</td>
<td>45.0</td>
<td>52.9</td>
<td>25.5% w/o BPaaS.</td>
<td>Source: Gartner - <a href="http://bit.ly/S9HpPK">http://bit.ly/S9HpPK</a></td>
</tr>
</tbody>
</table>

In Billions US$

These are Worldwide Numbers
Analytics
Hadoop

- Runs on a collection of cheap, commodity servers, in a distributed, shared nothing architecture

- Two key components
  - HDFS
    - Hadoop Distributed File System
  - MapReduce
    - Programming model for processing and generating large datasets
More than just Map/Reduce

- **Enterprise Support** that gives you the expertise and responsiveness you need to run Apache Hadoop in Production
- **Management Suite** that improves quality of service, increases compliance and reduces administrative costs
- **Customer Portal** access to submit and manage cases with Subject Matter Experts
- **Knowledge Base** that your team can use to expand their Hadoop knowledge with hundreds of articles and tech notes
- **Optimized Connectors** that integrate your Hadoop clusters with existing data systems
Faster Performance on Shared Storage

Increasing Hadoop Cluster Efficiency with NetApp
(Hadoop TeraGen utility, 8 through 24 node cluster)

Cluster Throughput (MB/sec)

Hadoop Data Nodes

- Replication 2 with NetApp
- Replication 3
Faster Job Completion During Disk Failure

- Compared to a Hadoop cluster with the same number of nodes, but using only internal disk (no RAID) and triple copies, the NetApp Open Solution for Hadoop shows only a slight performance decrease during a disk failure with less impact to job completion.

<table>
<thead>
<tr>
<th>Table 4. Drive Failure Recovery Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test Scenario</td>
</tr>
<tr>
<td>Healthy cluster</td>
</tr>
<tr>
<td>NetApp E2660 drive failure</td>
</tr>
<tr>
<td>Internal data node drive failure</td>
</tr>
</tbody>
</table>

**AutoSupport: Hadoop Use Case at NetApp**

- “Call-home” service for all NetApp® systems
- Foundation of NetApp proactive support strategies
- Machine-generated data doubles every 16 months

<table>
<thead>
<tr>
<th>CHALLENGE</th>
<th>NETAPP SOLUTION</th>
<th>BENEFITS</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 weeks to run a query on 24 billion unstructured records</td>
<td>10-node Hadoop Cluster w/ shared Storage</td>
<td>Time reduced from 4 weeks to 10.5 hours</td>
</tr>
<tr>
<td>Impossible to run a query: 240 billion unstructured records</td>
<td></td>
<td>Previously impossible, now achievable in just 18 hours</td>
</tr>
</tbody>
</table>

“NetApp ASUP is a mission-critical application”
Analytics Roadmap

- Traditional & Big Analytics side-by-side for years to come
- Hadoop moves to shared, virtualized infrastructure, for better efficiency and ease of management:
  - Hadoop remains logically distributed, shared nothing, but runs on a virtualized shared everything architecture (e.g., FlexPod for VMware + eSeries)
  - Same as above, except Hadoop becomes logically shared everything, as HDFS is replaced by a parallel file system (e.g., Lustre Cluster, StorNext or GPFS)
- Enterprise class resiliency (no SPoF) and reliability with HPC-like performance (no need for triplicas)
- Use of a single copy of data for the map phase (higher storage utilization)
- Natural intersection with Cloud (Analytics as a Service)
Bandwidth
Big Bandwidth Use Cases

**Full Motion Video/ISR**
Scalable density and performance to ingest and simultaneously analyze UAV, satellite and other data

**Video Surveillance**
High bandwidth & density supporting hundreds or thousands of HD cameras

**Media Content Management**
High ingest & play-out rates with support for media and entertainment workflows

**HPC & Seismic**
Massively parallel distributed file system for large scale cluster computing and O&G Seismic Processing
Bandwidth Driven by HPC

- THE driver for innovations in bandwidth are clearly high performance computing (HPC) applications.
- HPC falls into two camps:
  - High Performance Technical Computing (HPTC),
    - Supports scientific and engineering modeling and simulation.
    - HPTC makes up about 70% of the HPC market
    - Checkpoint/Rerstart
  - High Performance Business Computing (HPBC)
    - Financial Services, Media & Entertainment, Complex Event Processing, business process optimization and ultrascale business computing.
HPC Roadmap

- National Labs are already hard at work architecting the next system, pushing technologies to their limits
- Next stop: Exascale ($10^{18}$ FLOPS) in 2016-2018
- This will require 50 – 100X Sequoia
  - 80-160 million CPU cores
  - 80-160 PB’s of main memory
  - 266 – 533 TB/sec burst of sequential I/O from main memory during 5 minute checkpoint
- > 50 MW of power
- Existing storage technologies won’t scale within bandwidth, power and budget constraints
Trends Shaping the Big Content Market

- **Data at scale**
  - Objects replace files to overcome filesystem limitations
  - Containers 10’s PB in size, 10’s billions of objects, Millions of users
  - “Life time retention” of enterprise and consumer data

- **Policy driven infrastructure - “Set it and forget it”**
  - Object/VM granular automated data mobility/management

- **Simpler, efficient, self-managing systems; low price**
  - Systems with lower performance & resiliency for less $/GB
  - Space, power, density @ lower infrastructure cost (~2 PB/rack)

- **Centralized → distributed, cloud-friendly “system”**
  - Multi-location repositories, data mobility, local view/access
Distributed, Object-Based Repository

MULTIPLE: APPLICATIONS + SITES + LOCAL PROTOCOLS

- Site 1: APPLICATIONS
  - CIFS
  - NFS
  - HTTP

- Site 2: APPLICATIONS
  - CIFS
  - NFS
  - HTTP

- Site 3: APPLICATIONS
  - CIFS
  - NFS
  - HTTP

... Site N: APPLICATION
  - CIFS
  - NFS
  - HTTP

MULTIPLE: TENANTS + POLICIES + ADMINISTRATORS

SINGLE OBJECT ID + SINGLE OPEN GLOBAL PROTOCOL: CDMI

MULTIPLE: TARGETS + TIERS

- High Density Storage
- Tape

© 2012 NetApp, Inc. All rights reserved.
Distributed Content Repository Roadmap

- Federated CDMI (Cloud to Cloud)
- Strong ecosystem of ISVs supporting CDMI
- Exabyte size containers
- Trillions of objects
- Rack densities > 5 PB/rack
- Power densities < 1 kW / PB
- Costs < $150,000 / PB ($0.15 / GB)
Pre-configured, Certified Solutions
NetApp E-Series Storage Solutions

Dense and Modular Storage to Meet Any Scale

- Pre-Configured, Tested, Certified Solutions
- NetApp E-Series Hardware
- NetApp Professional Services
- NetApp Support Services

60 drives  120 drives  180 drives  240 drives  300 drives  360 drives
180TB      360TB      540TB      720TB      900TB      1080TB
E-Series Modular Design

DE6600

- Modular packaging architecture for maximum configuration flexibility
- Two optimized controller families

- 4U/60 x 3.5” drives
- Highest throughput
- Largest capacity/density
- 3TB NL-SAS drives

Disk Shelves

- 8Gb/s FC, Infiniband
- 12/24GB Cache per system
- 24.8Gb/s writes actual performance
- High Bandwidth / High IOPs

Controllers

- FC or iSCSI
- 4GB Cache per system
- 11.2Gb/s writes actual performance
- Cost effective performance

Systems

E-5460

5400 2600

E-2660
Two Sample Configurations: **E-Series 2660 and 5460**

**Single E2660 Array with Expansion Shelves**
- 3 4U-60 Enclosures
- 2 2600 Controllers
- **180 drives = 540TB raw**
- 8 ports iSCSI HIC, up to 1Gbps each port
- or 4 ports iSCSI HIC, up to 10Gbps each port
- Max write throughput is 9.6Gbps (theoretical)
- 220V Power

**Single E5460 Array with Expansion Shelves**
- 6 4U-60 Enclosures
- 2 5400 Controllers
- **360 drives = 1,080TB raw**
- 8 on board FC ports or 16 FC ports total with FC HIC, up to 8Gbps each port
- Max write throughput is 24Gbps (theoretical)
- 220V Power
Summary
So, What Does All This Mean to You?

You are also at an Inflection Point: You also have a decision to make, as “business as usual” may not cut it!
Summary

- Despite the hype, **Big Data is not new and is more than just analytics!** (Many agencies and private companies have struggled with Big Data for decades)

- **Analytics:** Traditional BI/DSS analytics still dominate. Importance of newer NoSQL & Columnar DB applications, enabled by MapReduce will grow with the growth of multi-structured data

- **Data bandwidth** will be driven by HPC Exascale initiatives

- Continued data growth that outstrips network bandwidth growth and file system scalability will drive low cost, geographically distributed, object-based, **Content Repositories** with federated access through CDMI

- Big Data applications, such as Hadoop, will need to adopt shared, virtualized infrastructure (and its management benefits) if they are to be widely adopted by Enterprise IT
Thank you

Dave Ellis
HPC Architect, PSC
NetApp, Inc.

Dave.Ellis@netapp.com
770-330-9486