I/O Challenges: Todays I/O Challenges for Big Data Analysis

Henry Newman
CEO/CTO
Instrumental, Inc.
April 30, 2013
The Challenge is Archives

- Big data in HPC means archive and archive translates to a tape archive at least today
  - Cost per PB per year

<table>
<thead>
<tr>
<th>Drives Per PB RAID-6 or 58+x</th>
<th>ES4 Drive SAS interface 4TB</th>
<th>Enterprise Value CS 1TB</th>
<th>Savio 15K 300 GB</th>
<th>Savio 10K 900GB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drive Count</td>
<td>Typical</td>
<td>Idle</td>
<td>Typical</td>
<td>Idle</td>
</tr>
<tr>
<td>Kw Cost</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.08</td>
<td>$2,771.38</td>
<td>$2,391.21</td>
<td>$5,457.49</td>
<td>$4,192.61</td>
</tr>
<tr>
<td>0.10</td>
<td>$3,464.23</td>
<td>$2,989.01</td>
<td>$6,821.87</td>
<td>$5,240.76</td>
</tr>
<tr>
<td>0.12</td>
<td>$4,157.08</td>
<td>$3,586.81</td>
<td>$8,186.24</td>
<td>$9,060.29</td>
</tr>
</tbody>
</table>

- What about reliability?

<table>
<thead>
<tr>
<th>Device</th>
<th>Hard error rate in bits</th>
<th>Equivalent in bytes</th>
<th>PB equivalent</th>
</tr>
</thead>
<tbody>
<tr>
<td>SATA consumer</td>
<td>10E14</td>
<td>1.25E+13</td>
<td>0.01</td>
</tr>
<tr>
<td>SATA Enterprise</td>
<td>10E15</td>
<td>1.25E+14</td>
<td>0.11</td>
</tr>
<tr>
<td>Enterprise SAS/FC</td>
<td>10E16</td>
<td>1.25E+15</td>
<td>1.11</td>
</tr>
<tr>
<td>LTO and some Enterprise SAS SSDs</td>
<td>10E17</td>
<td>1.25E+16</td>
<td>11.10</td>
</tr>
<tr>
<td>Enterprise Tape</td>
<td>10E19</td>
<td>1.25E+18</td>
<td>1110.22</td>
</tr>
</tbody>
</table>

- For the foreseeable future, tape is the archive medium of choice
Archive challenges

- Networking challenges
- Cost challenges
- RAS challenges
- Scalability challenges
- I/O performance challenges
- Metadata challenges
- Final thoughts
Networking Challenges

Those who own the archive own the big data solutions as you cannot move data around
Cannot move archive data

• Networking data rates are not increasing as fast as data generation

• In a perfect world:

<table>
<thead>
<tr>
<th>Network</th>
<th>1 PB</th>
<th>10 PB</th>
<th>40 PB</th>
<th>100 PB</th>
</tr>
</thead>
<tbody>
<tr>
<td>OC-3</td>
<td>802</td>
<td>8018</td>
<td>32071</td>
<td>80178</td>
</tr>
<tr>
<td>OC-12</td>
<td>200</td>
<td>1998</td>
<td>7992</td>
<td>19980</td>
</tr>
<tr>
<td>OC-48</td>
<td>51</td>
<td>506</td>
<td>2023</td>
<td>5057</td>
</tr>
<tr>
<td>OC-192</td>
<td>13</td>
<td>126</td>
<td>506</td>
<td>1264</td>
</tr>
<tr>
<td>OC-384</td>
<td>3</td>
<td>32</td>
<td>126</td>
<td>316</td>
</tr>
<tr>
<td>OC-768</td>
<td>1</td>
<td>8</td>
<td>32</td>
<td>79</td>
</tr>
</tbody>
</table>

• So if you wanted to read the whole archive to search for new information even with an OC-192 channel, it may take years
Cost Challenges or Business Opportunity

Those who control the archive own the big data solutions and the resulting benefits
• Justifying keeping data you might or might not need today is difficult
  – In the 80s for example the DNA between chromosomes was called junk DNA
    • We know today it is not junk
• Do we know what we need to keep?
  – We do not know what information we will find in the data we collect today
• Some industries have a policy of keeping everything forever
  – Geosciences is a good example
  – Satellite imagery is another
• These industries build in the cost of archives into the operational costs
  – These costs are well defined and well known based on industry trends
RAS Challenges

Those who own the archive own the big data solutions, as it is difficult to maintain data integrity and availability over long periods of time, will benefit
RAS is not easy

• Though tapes might last 30 years:
  – Interfaces, OSs and hardware do not
    • Try to find a supported fibre channel 1 Gbit interface
    • Try to find an operating system where this hardware is supported
    • How many motherboards still support PCI-X?

• What about things like silent corruption?
  – Shameless plug for tutorial I am giving on this topic at IEEE Mass Storage Conference next week
    – [http://www.storageconference.org/2013/Presentations.html](http://www.storageconference.org/2013/Presentations.html)
    » Slides will be up in a few weeks

• What about the time to migration to new technology?
  – Architecting for availability and meeting SLAs is difficult when migrating
    • Becomes a bigger issue when namespaces are broken up into smaller chunks
Scalability Challenges

Those who control the archive own the big data solutions, and as more data is collected and the size of the archive grows, and the organization will continue to win
Archives are hard

• Today’s largest single namespace archives have over 55 PB
  – Scaling for archives technologies choices
    • How many namespaces do you want to manage?
      – 1, 5, 10, 100, 1000?

• Large namespaces are important for more than just searching for your file
  – You can always put in a DB interface to search for a file
  – Large namespace size is important for management efficiency
    • Namespace load leveling is labor intensive
I/O Performance Challenges

Those who understand technology direction will be able to use the archive effectively and will be successful
• Performance is not increasing with density growth for any technology (SSD, disk, tape)
  – This trend is not expected to change
• This impacts archives in many different ways
  – Migration to new technology
  – Access time for data
    • Architecture and design must take this into account
MB/sec per GB of storage

Bandwidth per GB of Capacity

- Single Disk 1977: 37.5000 MB/sec
- Single SATA 2 TB Drive: 12.5000 MB/sec
- LTO-6 Compressed: 0.0800 MB/sec
- T10000C Uncompressed: 0.0420 MB/sec

Copyright 2013, Instrumental, Inc.
**Metadata Challenges**

Those who own the archive own the big data solutions as re-indexing metadata remotely at future scale will be impossible
• Bringing back a large archive over the network and indexing for new information is impossible
  – If someone has a new idea for how to use data and the archive is owned by someone else, you would not want to run the re-indexing on their machines
    • You would need to bring the database back to your location
• There seems an industry goal of outsourcing archives
  – This could have a long term negative financial impact on those companies
Final Thoughts

• Archives are and will be the central point for big data analysis
  – To be successful with archives, planning for costs, reliability, scalability and how the archive will be used is critical to the success
    • Those that have successful archives, like the geosciences industry, do all of the above

• Archive performance must be matched with data usage requirements
  – Do you need to use 1 TB per day or 100 TB per day?
  – Do you need to ingest 10 TB a day or 500 TB per day?

• Archive management and design requires careful planning
  – Spend the time and money to do it right the 1st time as doing the second time is far more costly
• Without owning the archive, you are dependent on a 3rd party for your future
  – That 3rd party might or might not be a competitor

• Ownership of the raw data is critical to your business and potential to the future of a nation
Thank you

Thanks for listening