Challenges with Data Integrity

Henry Newman
Instrumental, Inc.
hsn@instrumental.com
What is covered

• Problem Statement
• Flash concerns
• The numbers
Problem Statement

 Significant changes are needed in error encoding
The Problem

• Fibre channel originally developed at 25 MB/sec and it is now 64 times faster at 1600 MB/sec

• IDE channel originally was .625 MB/sec and it is now 480 times faster at 300 MB/sec

• The channel error rate for both is $10^{12}$ bits
  • Corrected to higher value
    – IB copper is the same
    • Optical is claimed to be better
    – 10GbE is the same
10GbE is a concern

• Research has shown rates of errors undetected by link CRC’s and TCP checksums ranging from one in 16 million to 10 billion packets
  – 16 Million 9K packets can be sent in less than 2 minutes on a 10GE link
  – 10 Billion 9K packets can be sent in less than 24 hours on a 10GE link
  – “When compared to un-detected error rates for local I/O (e.g., disk drives), these rates are disturbing”

• This is with TCP/IP checksums!
• Both disks and tape have far more error encoding than the channels
  – The encoding for tape is far more robust; for example, LTO is around 8 orders of magnitude greater than FC
  – Enterprise tape is at least 10 orders of magnitude better

• The channel error encode was not a consideration 20 years ago as things were too slow and too expensive to have lots of channels
  – No longer true
Error Encoding

• Robustness of error encoding has not changed for either storage channel type over the life of the channel
  – Between 20 (FC) and 25+ (SATA) years
  – It needs to be changed in ethernet
  – Seagate recently published SAS/FC undetectable rates

• This has resulted in a situation where organizations are starting to see actual data loss as we have hit the wall with error encoding
  – This does not even consider PCIe, memory issues or other part of the path
Flash SSDs and reliability

• Everyone thinks these are the ultimate solution for metadata and logs
• SSD have wide performance range for read and especially write
• NAND flash does not support writes over 100K times to a specific location and flash will fail
  – What happens to reliability at 70K, 90K and just before failure of the write?
  – What historical data do we have?
SSD and SMART Monitoring

• SMART is a standard that was developed for disk drives
  – Some of the error conditions found in flash do not fit within the framework for SMART
  – It took RAID vendors 3-5 years to accomplish predictive failure in controllers for disk drives

• No standard for SMART statistics for flash
  – New proposal to ANSI for flash but in early stages
### Annual Failure Rates at Different Sustained Transfer Rates Per Second.

<table>
<thead>
<tr>
<th>UDBER</th>
<th>0.5 GB/sec</th>
<th>1 GB/sec</th>
<th>10 GB/sec</th>
<th>100 GB/sec</th>
<th>1 TB/sec</th>
<th>10 TB/sec</th>
<th>100 TB/sec</th>
</tr>
</thead>
<tbody>
<tr>
<td>Est.T10 PI Detection</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.E-28</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>1.E-27</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>1.E-26</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>1.E-25</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>1.E-24</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>1.E-23</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>1.E-22</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>FC/SAS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.E-21</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.3</td>
<td>2.7</td>
<td>27.1</td>
</tr>
<tr>
<td>1.E-20</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.3</td>
<td>2.7</td>
<td>27.1</td>
<td>270.9</td>
</tr>
<tr>
<td>1.E-19</td>
<td>0.0</td>
<td>0.0</td>
<td>0.3</td>
<td>2.7</td>
<td>27.1</td>
<td>270.9</td>
<td>2708.9</td>
</tr>
<tr>
<td>1.E-18</td>
<td>0.1</td>
<td>0.3</td>
<td>2.7</td>
<td>27.1</td>
<td>270.9</td>
<td>2708.9</td>
<td>27089.2</td>
</tr>
<tr>
<td>SATA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.E-17</td>
<td>1.4</td>
<td>2.7</td>
<td>27.1</td>
<td>270.9</td>
<td>2708.9</td>
<td>27089.2</td>
<td>270892.2</td>
</tr>
<tr>
<td>1.E-16</td>
<td>13.5</td>
<td>27.1</td>
<td>270.9</td>
<td>2708.9</td>
<td>27089.2</td>
<td>270892.2</td>
<td>2708921.8</td>
</tr>
<tr>
<td>1.E-15</td>
<td>135.4</td>
<td>270.9</td>
<td>2708.9</td>
<td>27089.2</td>
<td>270892.2</td>
<td>2708921.8</td>
<td>27089217.7</td>
</tr>
</tbody>
</table>

- These annual failure rates are for a perfect world where the channels are operating at the specified rate of 10E^-12 and corrected to 10E^-17/19.
- What happens when the world is not perfect?
- What about GbE?
### Hard Error Rates and I/O

<table>
<thead>
<tr>
<th>Technology</th>
<th>Unrecoverable read error per bits read</th>
<th>1 PB</th>
<th>10 PB</th>
<th>40 PB</th>
<th>100 PB</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 TB Consumer SATA</td>
<td>10E14</td>
<td>9.007</td>
<td>90.07</td>
<td>360.288</td>
<td>900.720</td>
</tr>
<tr>
<td>1 TB</td>
<td>10E15</td>
<td>0.901</td>
<td>9.007</td>
<td>36.029</td>
<td>90.072</td>
</tr>
<tr>
<td>450 GB</td>
<td>10E16</td>
<td>0.090</td>
<td>0.901</td>
<td>3.603</td>
<td>9.007</td>
</tr>
<tr>
<td>LTO-4/TS1130</td>
<td>10E17</td>
<td>0.009</td>
<td>0.090</td>
<td>0.360</td>
<td>0.901</td>
</tr>
<tr>
<td>T10000B</td>
<td>10E19</td>
<td>0.000</td>
<td>0.001</td>
<td>0.004</td>
<td>0.009</td>
</tr>
</tbody>
</table>

- Clearly this is a problem that needs to be addressed
  - Vendors do not seem to be improving these values as it is on required in the commodity world
- To ensure data reliability other methods need to be investigated
### Will Clouds and Replication Work

<table>
<thead>
<tr>
<th>Network</th>
<th>Data Rate (Gb/sec)</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>0.15</td>
<td>802</td>
<td>8018</td>
<td>32071</td>
<td>80178</td>
</tr>
<tr>
<td>OC-12</td>
<td>0.61</td>
<td>200</td>
<td>1998</td>
<td>7992</td>
<td>19980</td>
</tr>
<tr>
<td>OC-48</td>
<td>2.40</td>
<td>51</td>
<td>506</td>
<td>2023</td>
<td>5057</td>
</tr>
<tr>
<td>OC-192</td>
<td>9.60</td>
<td>13</td>
<td>126</td>
<td>506</td>
<td>1264</td>
</tr>
<tr>
<td>OC-384</td>
<td>19.20</td>
<td>6</td>
<td>63</td>
<td>253</td>
<td>632</td>
</tr>
<tr>
<td>OC-768</td>
<td>38.40</td>
<td>3</td>
<td>32</td>
<td>126</td>
<td>316</td>
</tr>
</tbody>
</table>

- Given hard error rates and time to replicate in the event of a disaster at a site data will be lost
- I do not believe that “Hadoop method” will work given these considerations especially with the cost of power for CPUs and memory
  - 5 year costs with power is huge compared with other methods and risk of data loss in case of disaster is an issue
Final thoughts

• If there is corruption most people blame the file system first and the hardware last
  – That might have been a good plan in the 1970s-1990s but it is no longer true in most cases
• Some questions we could discuss as I have some thoughts and opinions:
  – Does error correction belong in the file system?
  – What should be done about hard error rate?
  – What will happen to tape given Dedup impact?