Learnings From Nearly A Decade Of Building Low-Cost Cloud Storage

What: SNIA Presentation
When: Sep 2015
Who: Gleb Budman, CEO
150+ Petabytes
Cloud storage has a new player, with a shockingly low price....

Meet B2 Cloud Storage >
<table>
<thead>
<tr>
<th>Storage ($/GB/month)</th>
<th>Upload (Free)</th>
<th>Download ($/GB)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>BACKBLAZE</strong></td>
<td>$0.005</td>
<td>$0.05</td>
</tr>
<tr>
<td>Amazon S3</td>
<td>$0.022+</td>
<td>$0.05+</td>
</tr>
<tr>
<td>Microsoft Azure</td>
<td>$0.022+</td>
<td>$0.05+</td>
</tr>
<tr>
<td>Google Cloud</td>
<td>$0.020+</td>
<td>$0.08+</td>
</tr>
<tr>
<td>Verizon</td>
<td>$0.040+</td>
<td>$0.08</td>
</tr>
<tr>
<td>Rackspace</td>
<td>$0.075+</td>
<td>$0.06+</td>
</tr>
<tr>
<td>CenturyLink</td>
<td>$0.150+</td>
<td>$0.05</td>
</tr>
</tbody>
</table>

Lowest cost shown for real-time cloud storage.
Savings That Really Add Up.

1 Petabyte of data stored for 12 months

- Backblaze: $60,000 (-75%)
- Amazon S3: $264,000
- Microsoft Azure: $288,000

@GlebBudman
Glacier? Nearline? S3-Infrequent Access?

- Backblaze B2 is still 40% - 250% lower cost
- Doesn’t make you wait for your data
- Doesn’t penalize for deletes
- Doesn’t penalize for access
The Cloud Storage Stack

Culture
Business Process
Software
Hardware
Datacenter
Datacenter
Convert Kilowatts-to-Kilobits
Datacenter

Goal:
• Convert kilowatts-to-bits

Considerations:
• Local cost of power and real estate
• Don’t ignore taxes, economic zones
• Climate
• Building and system efficiency
• Proximity to ops team & good people
• Connectivity

Resource: Backblaze Datacenter RFP
https://www.backblaze.com/blog/backblaze-datacenter-grows-1000x-datacenter-2-0-needed-apply-within/
Hardware
Connect Hard Drives to the Internet
Backblaze Storage Pod
Don’t Make Hardware Redundant
Use Commodity Parts

Server Power Supply VS Desktop Power Supply

@GlebBudman
Use Consumer Hard Drives

### Backblaze Hard Drive Failure Rates
Cumulative by Quarter (Q1 2014 - Q2 2015)

<table>
<thead>
<tr>
<th>Name/Model</th>
<th>Size (TB)</th>
<th>2013 Q4</th>
<th>2013 Q1</th>
<th>2014 Q4</th>
<th>2014 Q1</th>
<th>2014 Q2</th>
<th>2014 Q3</th>
<th>2014 Q4</th>
<th>2015 Q1</th>
<th>2015 Q2</th>
</tr>
</thead>
<tbody>
<tr>
<td>HGST Deskstar 7K2000 (HDS722020ALA330)</td>
<td>2TB</td>
<td>1.10%</td>
<td>1.00%</td>
<td>1.09%</td>
<td>1.03%</td>
<td>1.06%</td>
<td>1.15%</td>
<td>1.40%</td>
<td>2015 Q1</td>
<td>1.90%</td>
</tr>
<tr>
<td>HGST Deskstar 5K3000 (HDS5C300ALZA630)</td>
<td>3TB</td>
<td>0.90%</td>
<td>0.85%</td>
<td>0.73%</td>
<td>0.73%</td>
<td>0.74%</td>
<td>0.74%</td>
<td>0.74%</td>
<td>0.74%</td>
<td>1.10%</td>
</tr>
<tr>
<td>HGST Deskstar 7K3000 (HDS723030ALA640)</td>
<td>3TB</td>
<td>0.90%</td>
<td>1.54%</td>
<td>1.54%</td>
<td>1.55%</td>
<td>1.81%</td>
<td>1.83%</td>
<td>0.50%</td>
<td>0.50%</td>
<td>0.50%</td>
</tr>
<tr>
<td>HGST Deskstar 5K4000 (HDS5C400ALAE640)</td>
<td>4TB</td>
<td>1.50%</td>
<td>1.33%</td>
<td>1.25%</td>
<td>1.06%</td>
<td>1.17%</td>
<td>1.16%</td>
<td>1.11%</td>
<td>1.11%</td>
<td>1.11%</td>
</tr>
<tr>
<td>HGST Megascale 4000 (HGST HM5C4044LE640)</td>
<td>4TB</td>
<td>2.67%</td>
<td>1.90%</td>
<td>1.86%</td>
<td>1.43%</td>
<td>1.18%</td>
<td>1.16%</td>
<td>1.66%</td>
<td>1.66%</td>
<td>1.66%</td>
</tr>
<tr>
<td>HGST Megascale 4000 B (HGST HM5C4044LE640)</td>
<td>4TB</td>
<td>20.29%</td>
<td>1.23%</td>
<td>0.59%</td>
<td>0.52%</td>
<td>0.48%</td>
<td>0.80%</td>
<td>0.80%</td>
<td>0.80%</td>
<td>0.80%</td>
</tr>
<tr>
<td>Seagate Barracuda 7200.11 (ST315000415AS)</td>
<td>1.5TB</td>
<td>25.40%</td>
<td>22.27%</td>
<td>22.98%</td>
<td>23.02%</td>
<td>23.41%</td>
<td>24.12%</td>
<td>23.90%</td>
<td>23.90%</td>
<td></td>
</tr>
<tr>
<td>Seagate Barracuda LP (ST31500415AS)</td>
<td>1.5TB</td>
<td>9.90%</td>
<td>9.97%</td>
<td>9.67%</td>
<td>9.56%</td>
<td>9.93%</td>
<td>10.18%</td>
<td>10.50%</td>
<td>10.50%</td>
<td></td>
</tr>
<tr>
<td>Seagate Barracuda LP (ST32000421AS)</td>
<td>2TB</td>
<td>7.20%</td>
<td>8.09%</td>
<td>8.18%</td>
<td>9.46%</td>
<td>9.65%</td>
<td>9.93%</td>
<td>10.10%</td>
<td>10.10%</td>
<td></td>
</tr>
<tr>
<td>Seagate Barracuda 7200.14 (ST3000DM001)</td>
<td>3TB</td>
<td>9.80%</td>
<td>13.92%</td>
<td>17.65%</td>
<td>27.15%</td>
<td>28.31%</td>
<td>28.26%</td>
<td>28.20%</td>
<td>28.20%</td>
<td></td>
</tr>
<tr>
<td>Seagate Barracuda XT (ST3000DMAS1AS)</td>
<td>3TB</td>
<td>7.30%</td>
<td>6.33%</td>
<td>6.33%</td>
<td>6.08%</td>
<td>5.59%</td>
<td>5.27%</td>
<td>5.30%</td>
<td>5.30%</td>
<td></td>
</tr>
<tr>
<td>Seagate Barracuda XT (ST4000DMAS1AS)</td>
<td>4TB</td>
<td>0.75%</td>
<td>0.56%</td>
<td>0.45%</td>
<td>1.12%</td>
<td>1.61%</td>
<td>1.70%</td>
<td>1.70%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seagate Desktop HD.15 (ST4000DM000)</td>
<td>4TB</td>
<td>3.33%</td>
<td>3.03%</td>
<td>2.73%</td>
<td>2.75%</td>
<td>2.83%</td>
<td>3.00%</td>
<td>3.00%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seagate 6TB SATA 3.5 (ST6000VM000)</td>
<td>6TB</td>
<td>1.70%</td>
<td>3.80%</td>
<td>3.80%</td>
<td>3.80%</td>
<td>3.80%</td>
<td>3.80%</td>
<td>3.80%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Toshiba D01A1CA Series (TOSHIBA D01A1CA800)</td>
<td>3TB</td>
<td>4.63%</td>
<td>4.63%</td>
<td>4.63%</td>
<td>4.63%</td>
<td>4.63%</td>
<td>4.63%</td>
<td>4.63%</td>
<td>4.63%</td>
<td></td>
</tr>
<tr>
<td>Toshiba MD040ABA-V Series (TOSHIBA MD040ABA800V)</td>
<td>4TB</td>
<td>0.00%</td>
<td>3.50%</td>
<td>3.50%</td>
<td>3.50%</td>
<td>3.50%</td>
<td>3.50%</td>
<td>3.50%</td>
<td>3.50%</td>
<td></td>
</tr>
<tr>
<td>Toshiba MD04ABA-V Series (TOSHIBA MD040ABA800V)</td>
<td>5TB</td>
<td>0.00%</td>
<td>6.50%</td>
<td>6.50%</td>
<td>6.50%</td>
<td>6.50%</td>
<td>6.50%</td>
<td>6.50%</td>
<td>6.50%</td>
<td></td>
</tr>
<tr>
<td>Western Digital Red 3TB MDC WD30EFZ</td>
<td>3TB</td>
<td>3.20%</td>
<td>8.78%</td>
<td>9.07%</td>
<td>6.96%</td>
<td>6.45%</td>
<td>7.00%</td>
<td>8.40%</td>
<td>8.40%</td>
<td></td>
</tr>
<tr>
<td>Western Digital Red 4TB MDC WD40EFRX</td>
<td>4TB</td>
<td>9.07%</td>
<td>1.90%</td>
<td>1.90%</td>
<td>1.90%</td>
<td>1.90%</td>
<td>1.90%</td>
<td>1.90%</td>
<td>1.90%</td>
<td>1.90%</td>
</tr>
<tr>
<td>Western Digital Red 6TB MDC WD60EFRX</td>
<td>6TB</td>
<td>13.75%</td>
<td>3.07%</td>
<td>6.64%</td>
<td>6.20%</td>
<td>6.20%</td>
<td>6.20%</td>
<td>6.20%</td>
<td>6.20%</td>
<td>6.20%</td>
</tr>
</tbody>
</table>
Insorce & Do the Math for Drive Purchases

Cost per GB for Hard Drives
Prices Backblaze paid for drives from 2009-2013

- 1 TB
- 1.5 TB
- 2 TB
- 3 TB
- 4 TB

Year
2009 2010 2011 2012 2013

@GlebBudman
Hardware

Goal:
• Connect hard drives to the Internet

Considerations:
• Don’t make hardware redundant
• Use commodity parts
• Use consumer hard drives
• Insourse & use math for drive purchases

Resource: Backblaze Storage Pod
https://www.backblaze.com/storage-pod.html/
Software
Put All Intelligence Here

@GlebBudman
Backblaze Vault

20 hard drives create 1 tome that share parts of a file.
Avoid Choke Points
Algorithmically Monitor SMART Stats

SMART 187

Correlated to Annual Failure Rate. As the number of read errors increase, it is more likely the drive will experience a failure.

@GlebBudman
Plan for Silent Corruption

VS

@GlebBudman
Put Replication Above the File System
Run Out of Resources Simultaneously
Model & Monitor Storage Burn

Storage Growth Models

- Space Used (TB Base2)
- Linear Model (1y)
- Quadratic Model (2y)
- Exponential Model (2y)

@GlebBudman
Goal:
• Put all intelligence here

Considerations:
• Avoid choke points
• Algorithmically monitor SMART stats
• Plan for Silent Corruption
• Put replication above the file system
• Run out of resources simultaneously
• Model & monitor storage burn

Resource: Backblaze Vault
https://www.backblaze.com/blog/vault-cloud-storage-architecture/
Business Processes
Optimize for TCO
Design for Failure...
...But Fix Failures Quickly
Create Repeatable Repairs
Standardize Pod Chassis
ROI Drives Automation
Goal:
- Optimize for TCO

Considerations:
- Design for failure, but fix failures quickly
- Create repeatable repairs (drive swap, chassis swap)
- Standardize pod chassis
- ROI drives automation
- Workflow for storage buffer
  (different buffer types; e.g. 11 PB in a day)

Resource: Backblaze Drive Process
https://www.backblaze.com/blog/alas-poor-stephen-is-dead/
Culture
Question “Conventional Wisdom”

Image by Michael Cara
https://www.flickr.com/photos/michael_cala/

@GlebBudman
No Hardware Worshippers

@GlebBudman
Agile Extends to Hardware

Storage Pod Scrum

The “old” Product Backlog becomes the new one when units go into production.

Sprint 1 → Sprint 2 → Sprint “N” → Sprint “Last” → Production

Never mind → Testing
Relentless Focus on Cost

- Is this required?
- Is there a comparable lower cost option?
- Can business processes work around this?
- Can software work around this?
Culture

Goal:
• Question “conventional wisdom”

Considerations:
• No hardware worshippers
• Agile extends to hardware
• Relentless focus on cost

Resource: Backblaze Agile Hardware Process
https://www.backblaze.com/blog/designing-the-next-storage-pod/
INTRODUCING BACKBLAZE

B2 Cloud Storage

The lowest cost cloud storage on the planet: $.005/GB a month.
Try it and get the first 10 GB free on us.

Create Account & Get B2 Invite

B2 is in private beta. Sign up free to be among the first to try it.
No credit card required
(Already have a Backblaze account?)
Over the past year we've been releasing hard drive reliability statistics based on the drives we use to store customer data for our online backup business. As of the end of Q1 2015 we had 44,252 hard drives spinning in our datacenter. If we subtract boot drives and drive models with less than 45 drives from that total, we get 42,749 hard drives remaining spread across 21
Learnings from nearly a decade of building low-cost cloud storage.

Backblaze.com/B2 => Lowest cost cloud storage
Backblaze.com/blog => Cloud storage resources

Gleb Budman
@glebbudman
Gleb.Budman@Backblaze.com