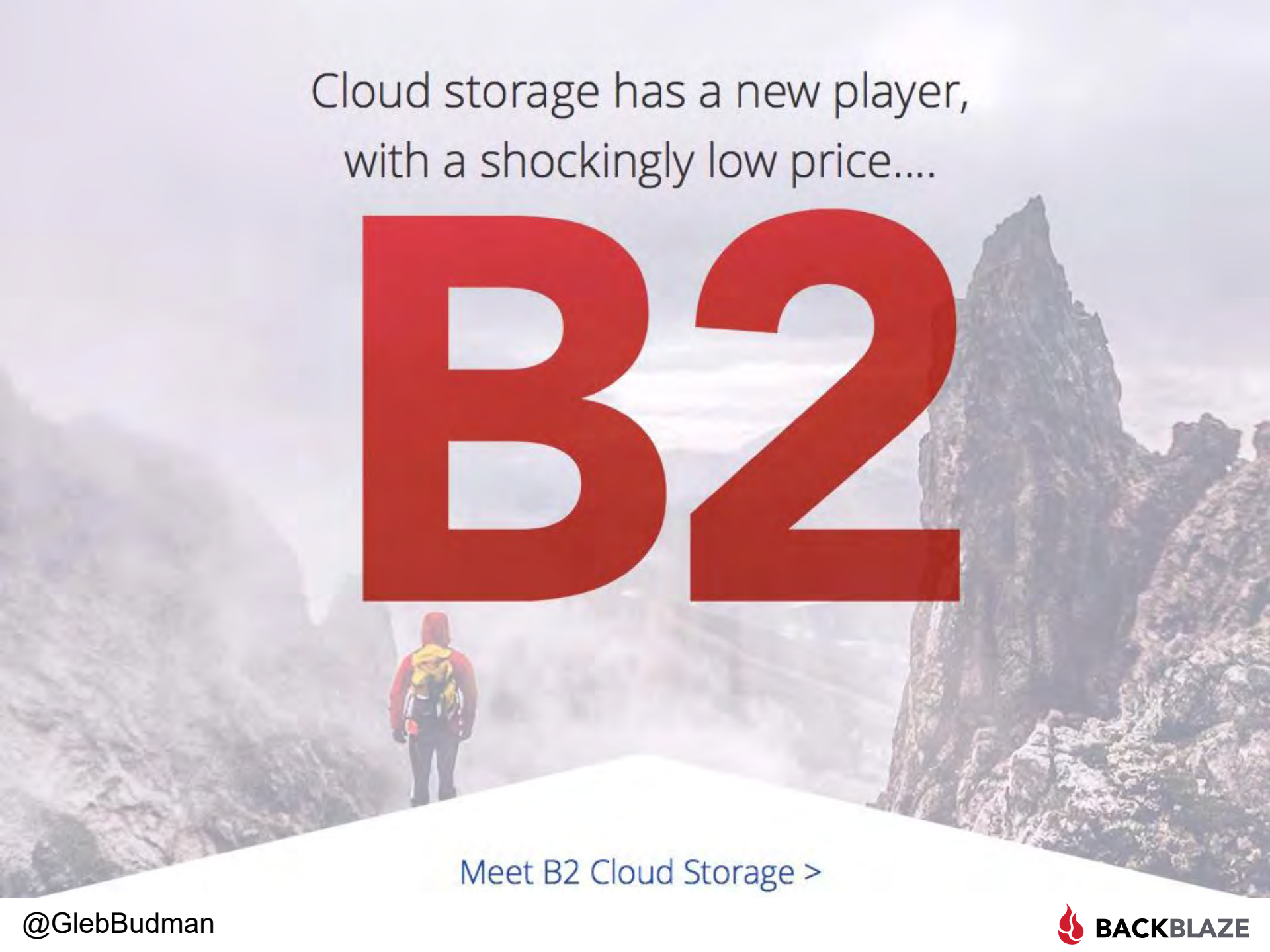
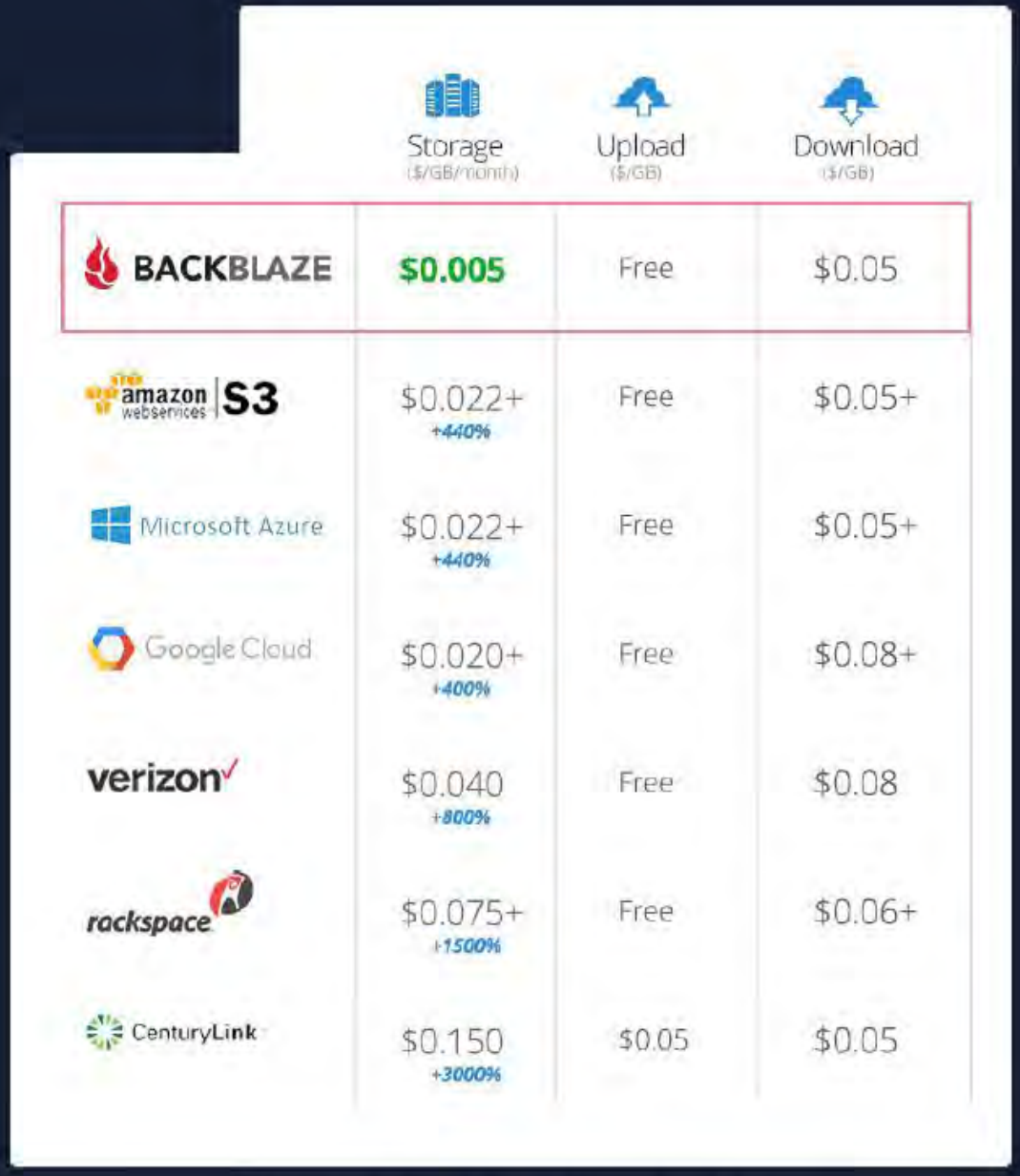
BACKBLAZE

Learnings from Operating 200 PB of Disk-Based Storage

What	MSST
When	May 2016
Who	Gleb Budman, CEO


Cloud storage has a new player,
with a shockingly low price....
B2
Meet B2 Cloud Storage >

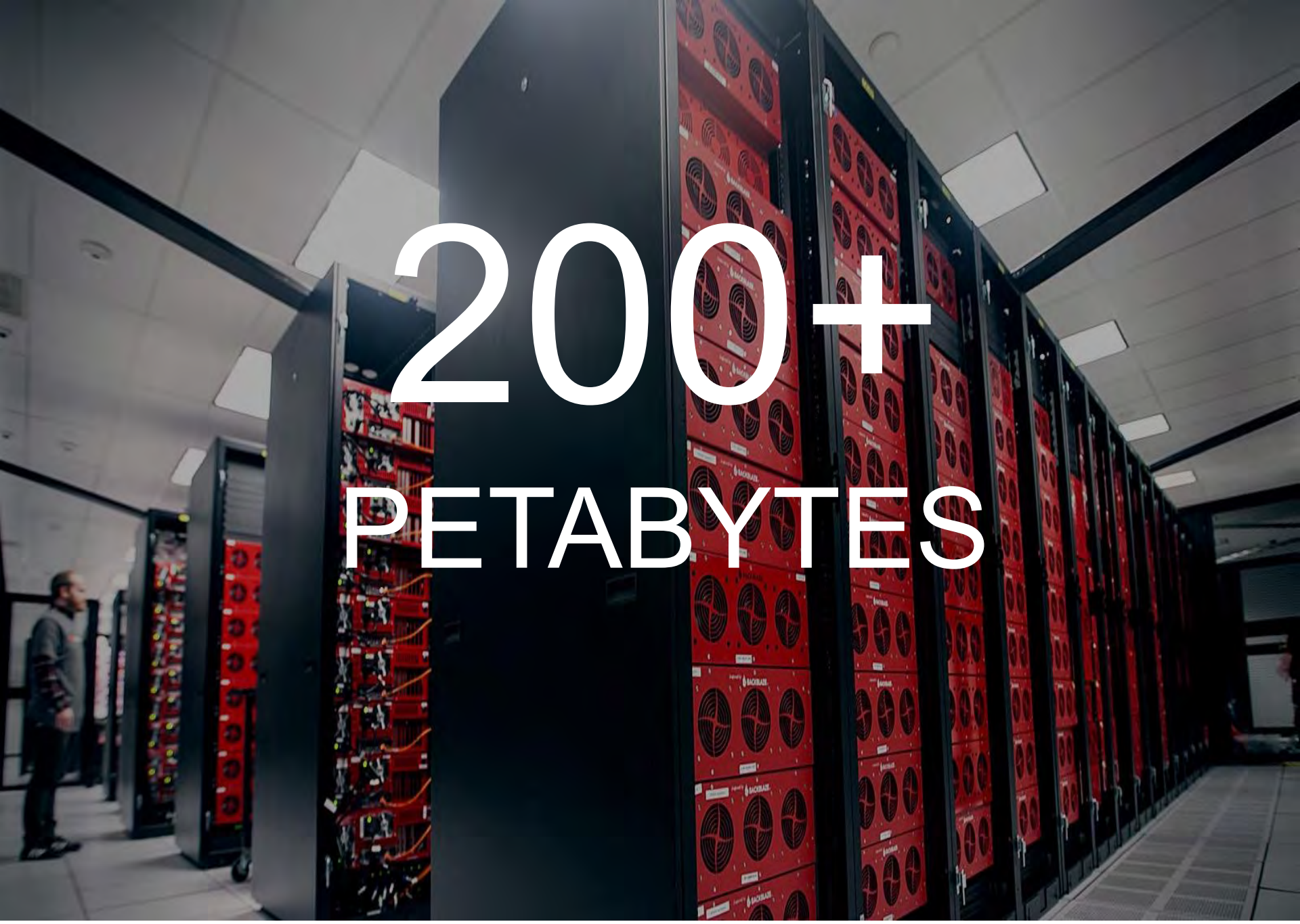


Lowest cost shown for real-time cloud storage.

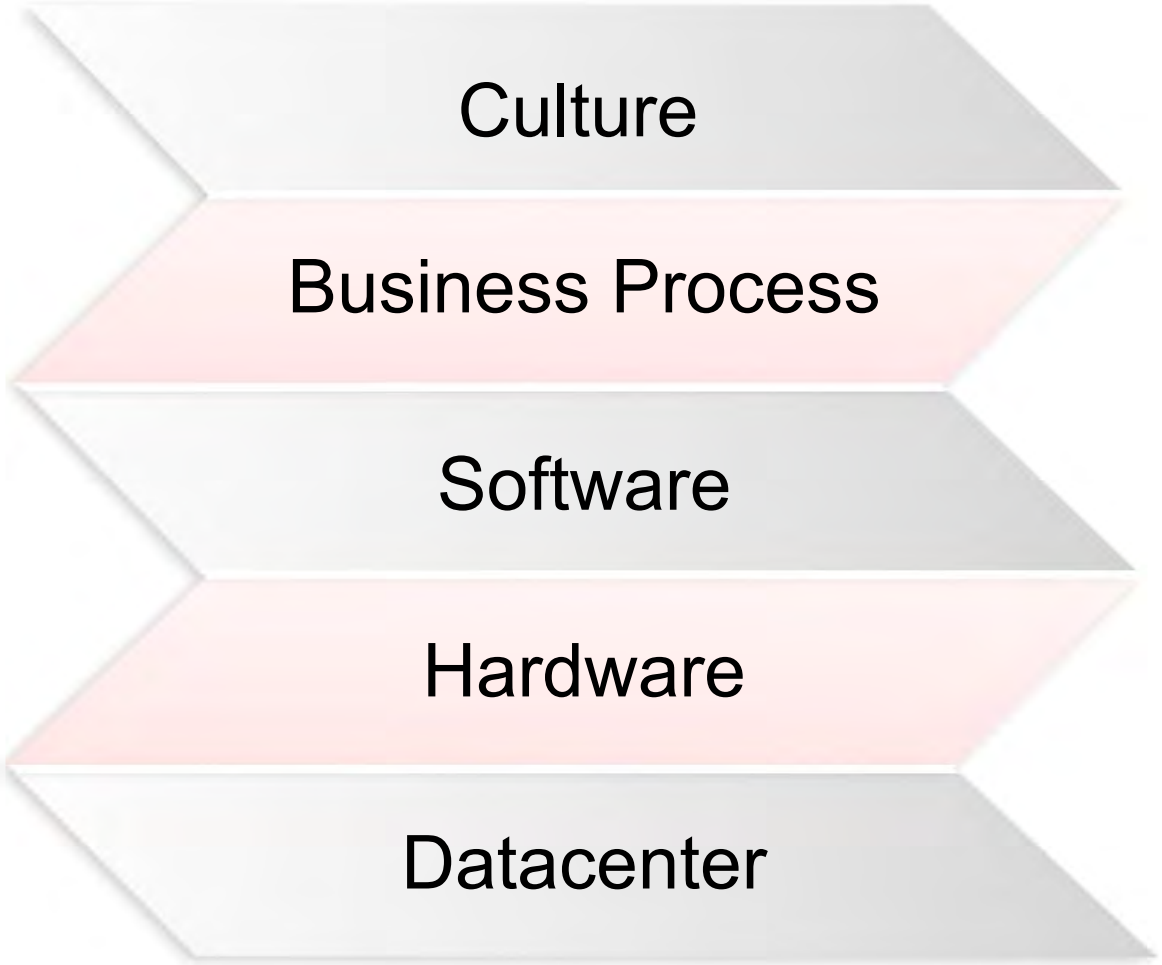
BACKBLAZE

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
- Upload transactions are free


200+
PETABYTES

The Cloud Storage Stack




Datacenter
Convert Kilowatts-to-Kilobits

BACKBLAZE

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

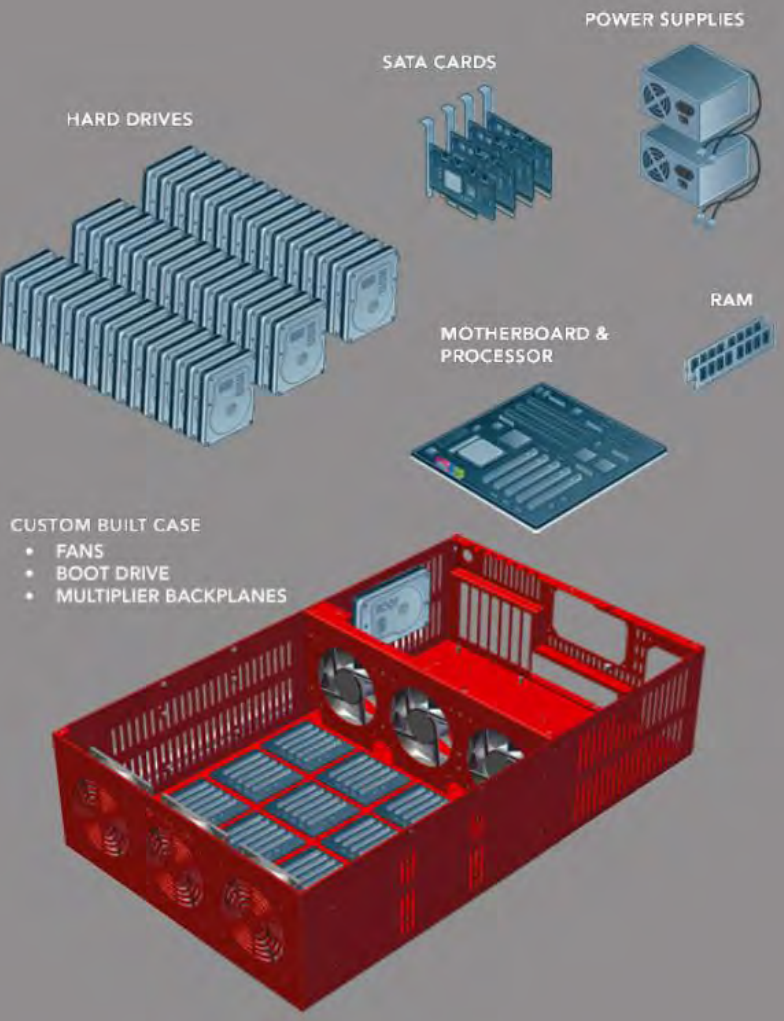
https://www.backblaze.com/blog/backblaze-datacenter-grows-1000x-datacenter-2-0-needed-apply-within/

Hardware
Connect Hard Drives to the Internet

BACKBLAZE

Backblaze Storage Pod

Don’t Make Hardware Redundant



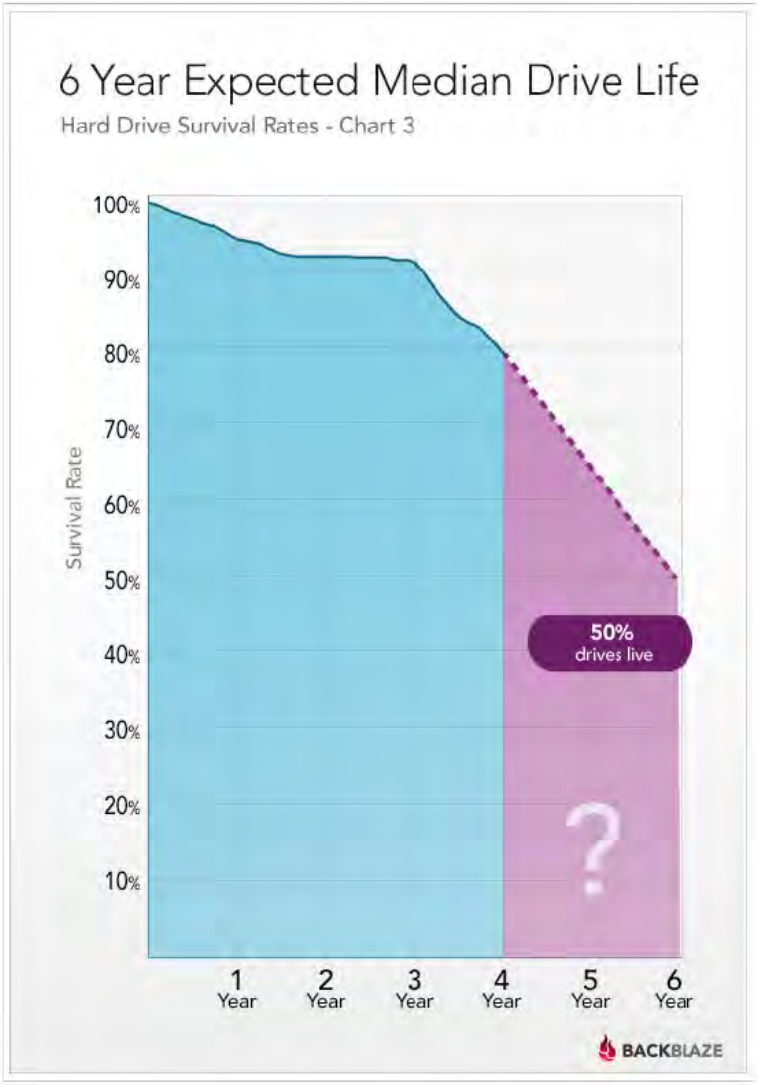
Use Commodity Parts

Server Power Supply

VS

Desktop Power Supply

Use Consumer Hard Drives



Backblaze Hard Drive Failure Rates

Cumulative by Quarter (Q1 2014 - Q2 2015)

Name/Model	Size	2013 Q4	2014 Q1	2014 Q2	2014 Q3	2014 Q4	2015 Q1	2015 Q2
HGST Deskstar 7K2000 (HDS722020ALA330)	2TB	1.10%	1.08%	1.09%	1.03%	1.06%	1.15%	1.90%
HGST Deskstar 5K3000 (HDS5C3030ALA630)	3TB	0.90%	0.85%	0.70%	0.73%	0.74%	0.74%	1.10%
HGST Deskstar 7K3000 (HDS723030ALA640)	3TB	0.90%	1.54%	1.46%	1.55%	1.81%	1.83%	0.50%
HGST Deskstar 5K4000 (HDS5C4040ALE630)	4TB	1.50%	1.33%	1.25%	1.06%	1.17%	1.16%	1.10%
HGST Megascale 4000 (HGST HMS5C4040ALE640)	4TB		2.67%	1.90%	1.86%	1.43%	1.18%	1.60%
HGST Megascale 4000.B (HGST HMS5C4040BLE640)	4TB		20.29%	1.23%	0.59%	0.52%	0.48%	0.80%
Seagate Barracuda 7200.11 (ST31500341AS)	1.5TB	25.40%	22.27%	22.98%	23.02%	23.41%	24.12%	23.90%
Seagate Barracuda LP (ST31500541AS)	1.5TB	9.90%	9.87%	9.67%	9.56%	9.93%	10.18%	10.50%
Seagate Barracuda LP (ST32000542AS)	2TB	7.20%	8.03%	8.18%	9.96%	9.63%	9.93%	10.10%
Seagate Barracuda 7200.14 (ST3000DM001)	3TB	9.80%	13.92%	17.65%	27.15%	28.31%	28.26%	28.20%
Seagate Barracuda XT (ST33000651AS)	3TB	7.30%	6.53%	6.33%	6.08%	5.59%	5.27%	5.30%
Seagate Barracuda XT (ST4000DX000)	4TB		0.75%	0.56%	0.45%	1.12%	1.61%	1.70%
Seagate Desktop HDD.15 (ST4000DM000)	4TB		3.83%	3.03%	2.73%	2.75%	2.83%	3.00%
Seagate 6 TB SATA 3.5 (ST6000DX000)	6TB						1.70%	3.80%
Toshiba DT01ACA Series (TOSHIBA DT01ACA300)	3TB		4.63%	3.48%	4.20%	4.81%	4.23%	4.60%
Toshiba MD04ABA-V Series (TOSHIBA MD04ABA400V)	4TB						0.00%	3.50%
Toshiba MD04ABA-V Series (TOSHIBA MD04ABA500V)	5TB						0.00%	6.50%
Western Digital Red 3 TB (WDC WD30EFRX)	3TB	3.20%	8.78%	9.07%	6.96%	6.49%	7.90%	8.40%
Western Digital 4 TB (WDC WD40EFRX)	4TB						9.01%	1.90%
Western Digital Red 6 TB (WDC WD60EFRX)	6TB				13.75%	3.07%	6.64%	6.20%

BACKBLAZE

Hardware

Goal:

- Connect hard drives to the Internet

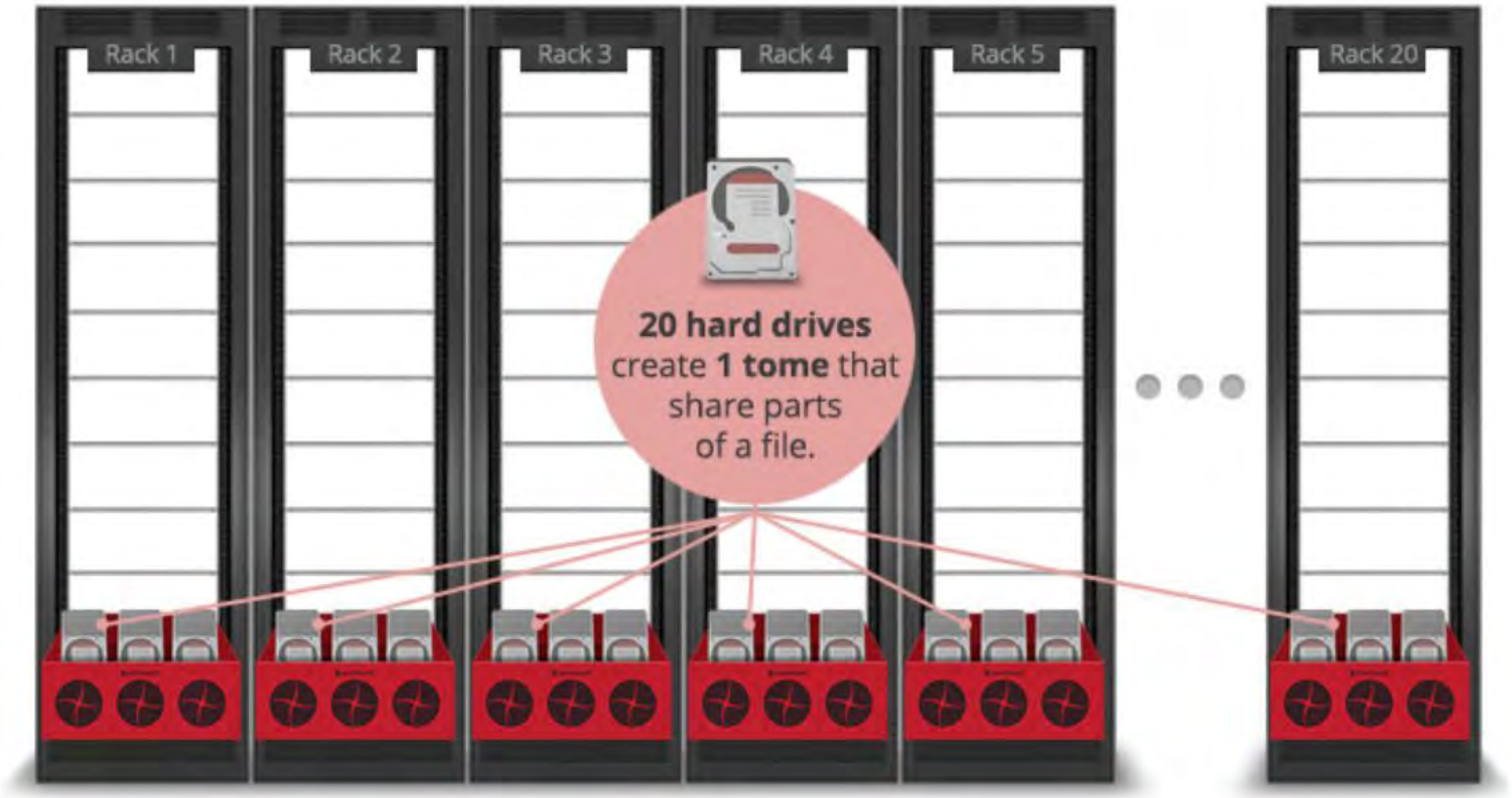
Considerations:

- Don’t make hardware redundant
- Use commodity parts
- Use consumer hard drives

Resource: Backblaze Storage Pod
https://www.backblaze.com/storage-pod.html/


Software
Put All Intelligence Here

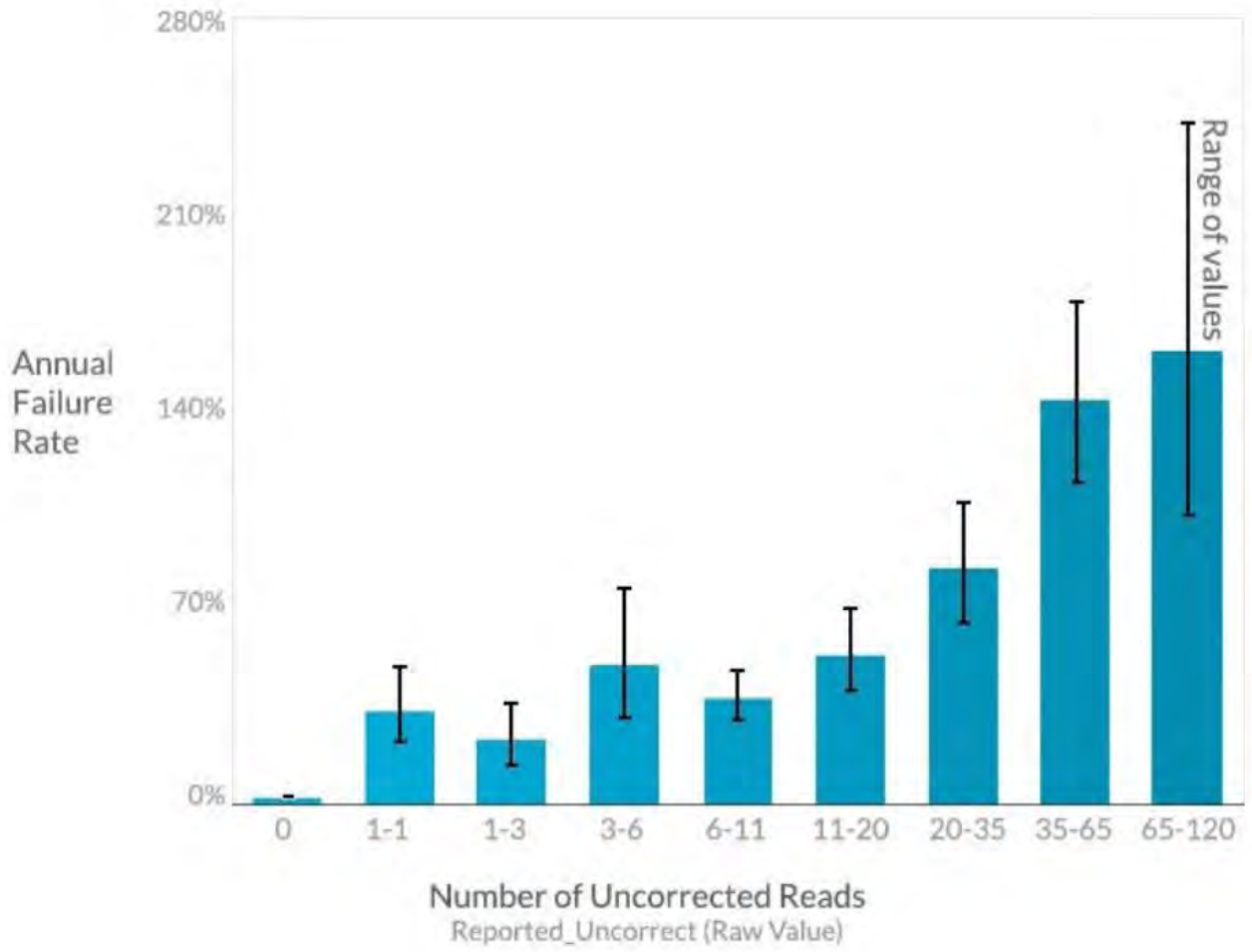
Backblaze Vault




Avoid Choke Points

SMART 187

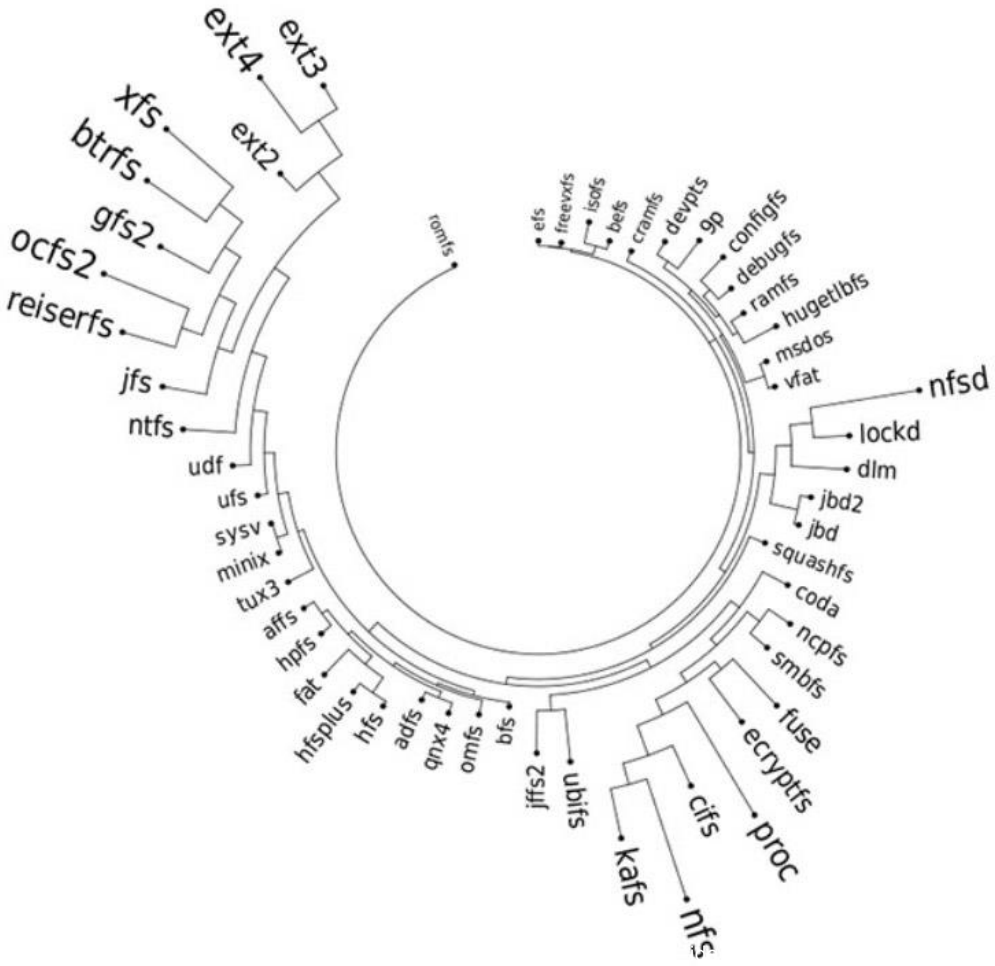
Correlated to Annual Failure Rate. As the number of read errors increase, it is more likely the drive will experience a failure.



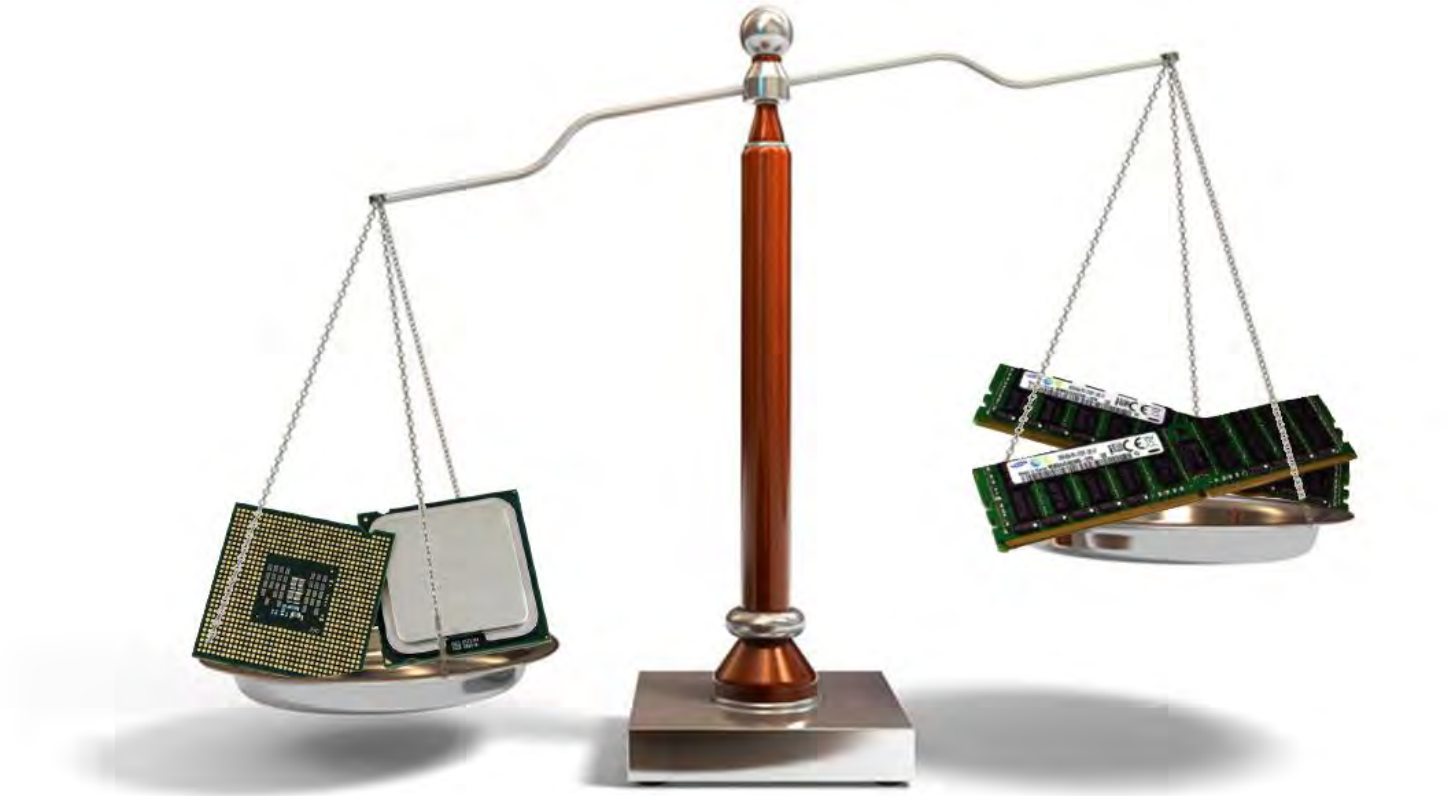
Plan for Silent Corruption

VS

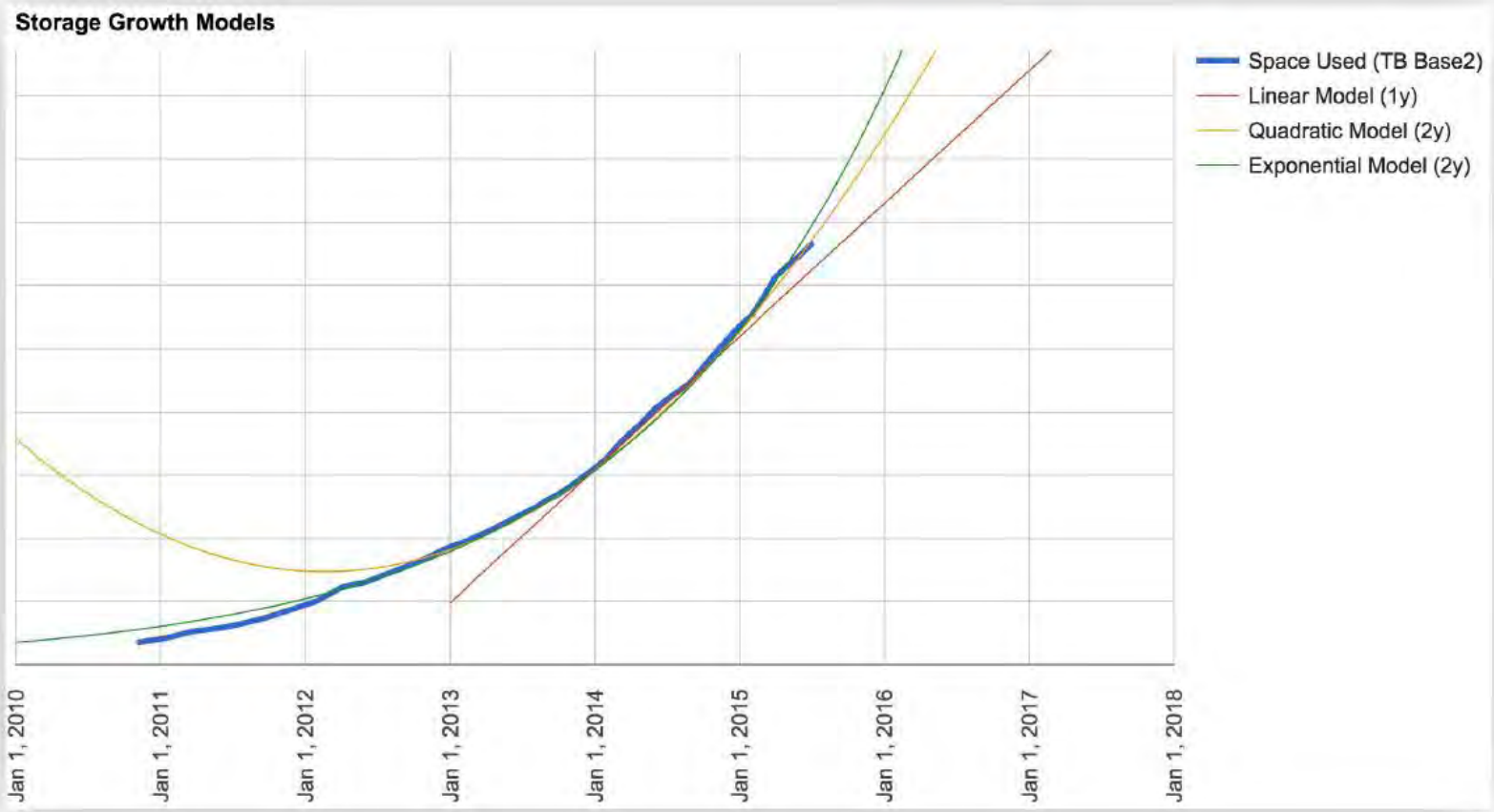
Put Replication Above the File System



Run Out of Resources Simultaneously



Model & Monitor Storage Burn



BACKBLAZE

Software

Goal:

- Put all intelligence here

Considerations:

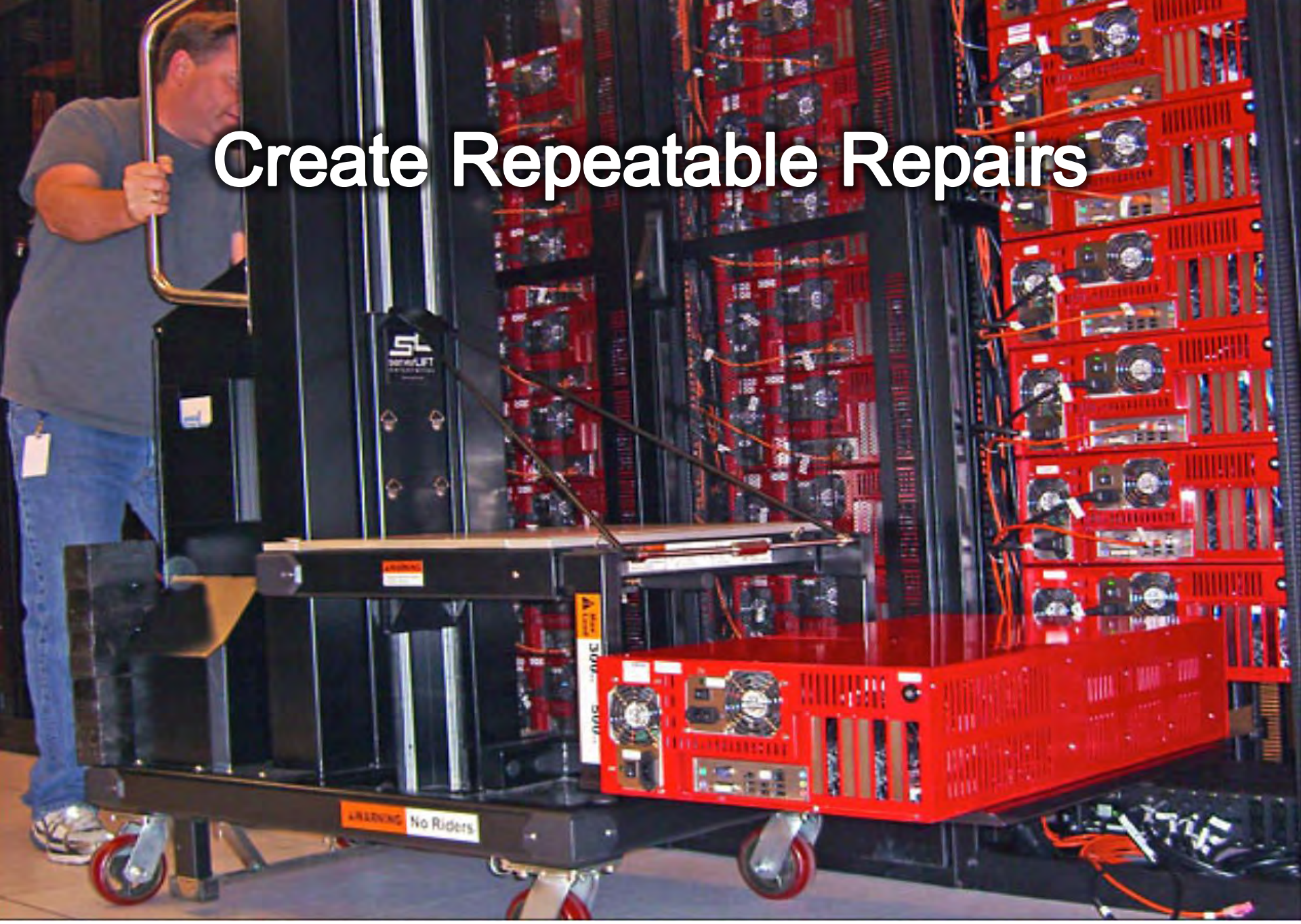
- Avoid choke points
- Algorithmically monitor SMART stats
- Plan for Silent Corruption
- Put replication above the file system
- Run out of resources simultaneously
- Model & monitor storage burn

Resource: Backblaze Vault
https://www.backblaze.com/blog/vault-cloud-storage-architecture/

Business Processes Optimize for TCO

Design for Failure…

But Fix

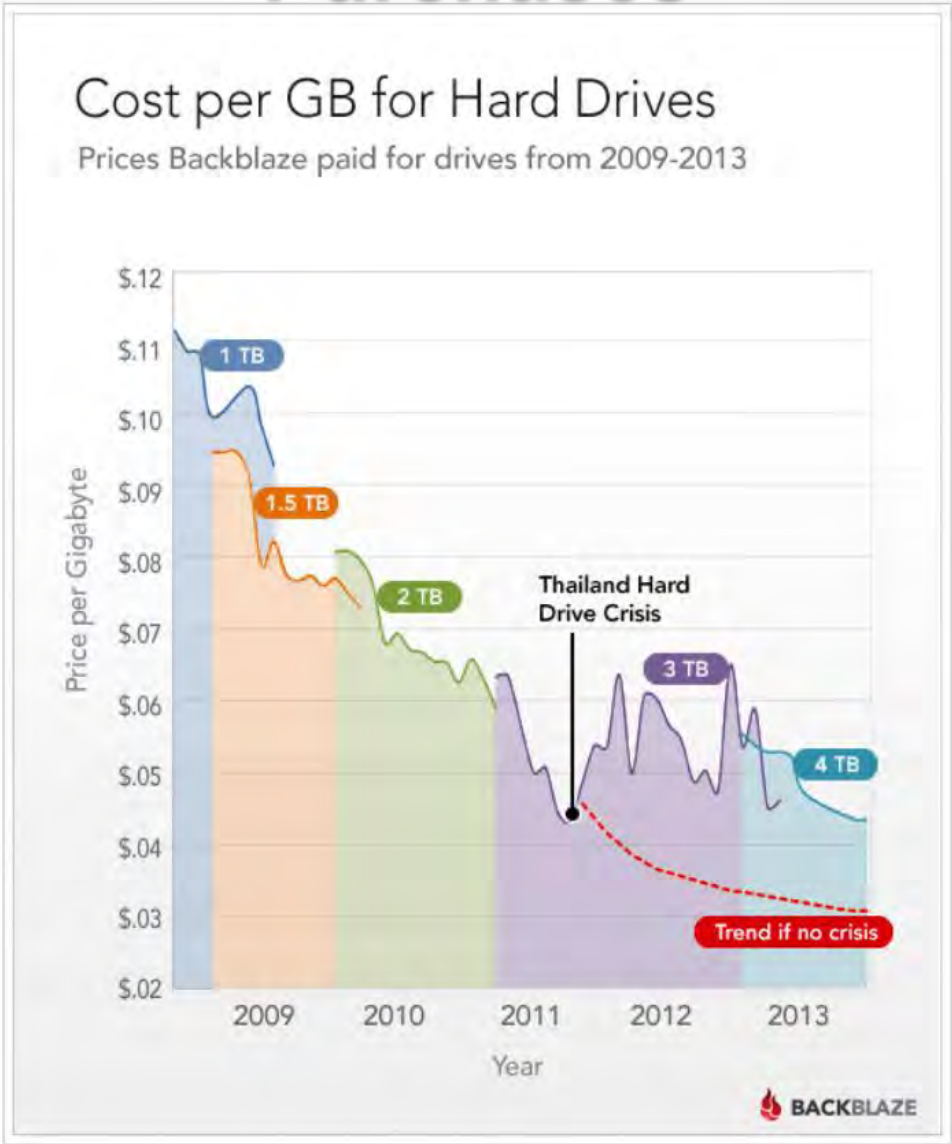

Create Repeatable Repairs
No Riders

Standardize Pod Chassis


ROI Drives Automation

Workflow for Storage Buffer

Purchases


Cost per GB for Hard Drives
Prices Backblaze paid for drives from 2009-2013
Price per Gigabyte
$.12
$.11
$.10
$.09
$.08
$.07
$.06
$.05
$.04
$.03
$.02
1 TB
1.5 TB
2 TB
Thailand Hard Drive Crisis
3 TB
4 TB
Trend if no crisis
2009
2010
2011
2012
2013
Year
BACKBLAZE

BACKBLAZE

Business Processes

Goal:

- Optimize for TCO

Considerations:

- Design for failure, but fix failures quickly
- Create repeatable repairs
- Standardize pod chassis
- ROI drives automation
- Workflow for storage buffer
- Insource & use math for drive purchases

Resource: Backblaze Drive Process
https://www.backblaze.com/blog/alas-poor-stephen-is-dead/

Culture
Question “Conventional Wisdom”

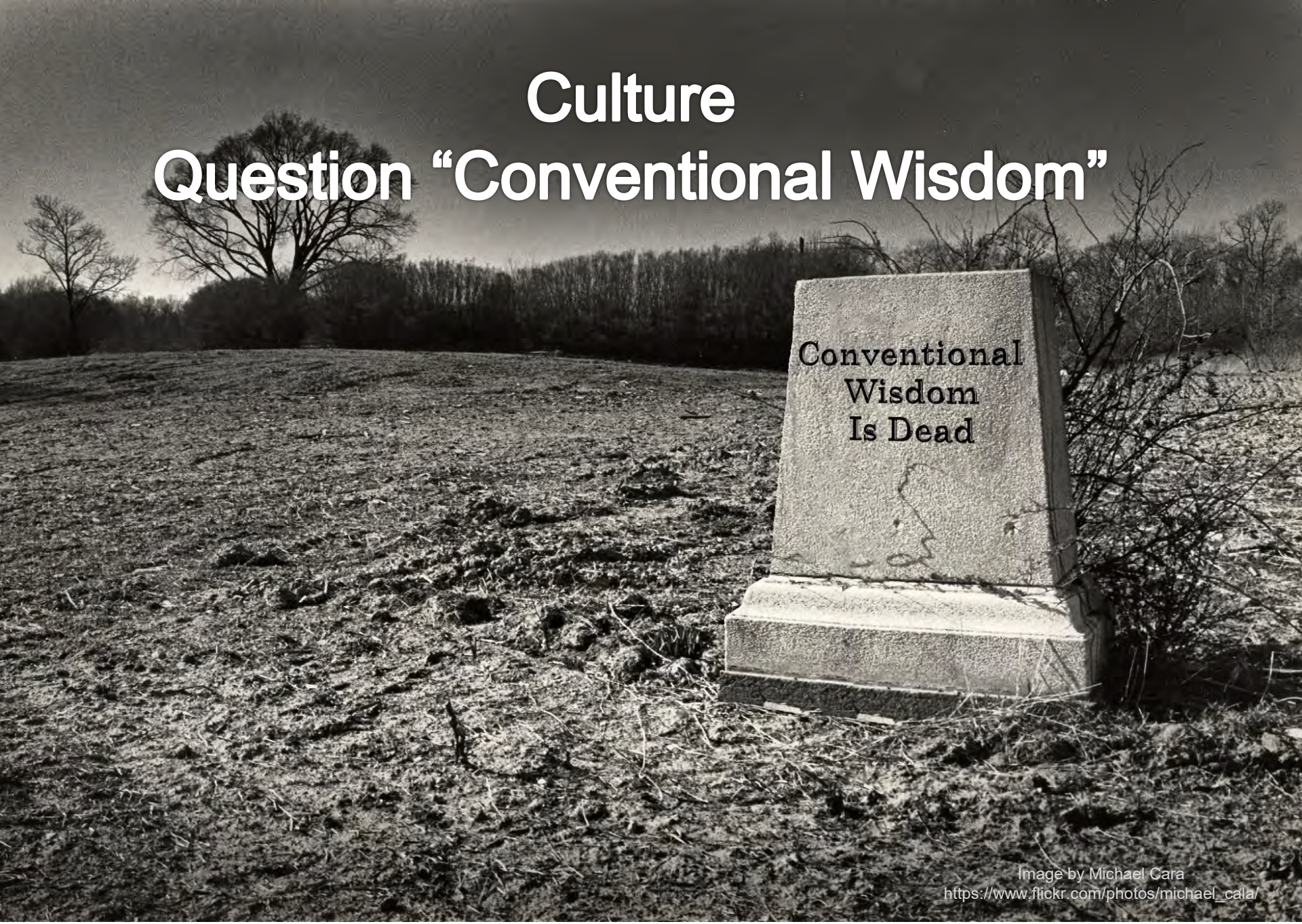
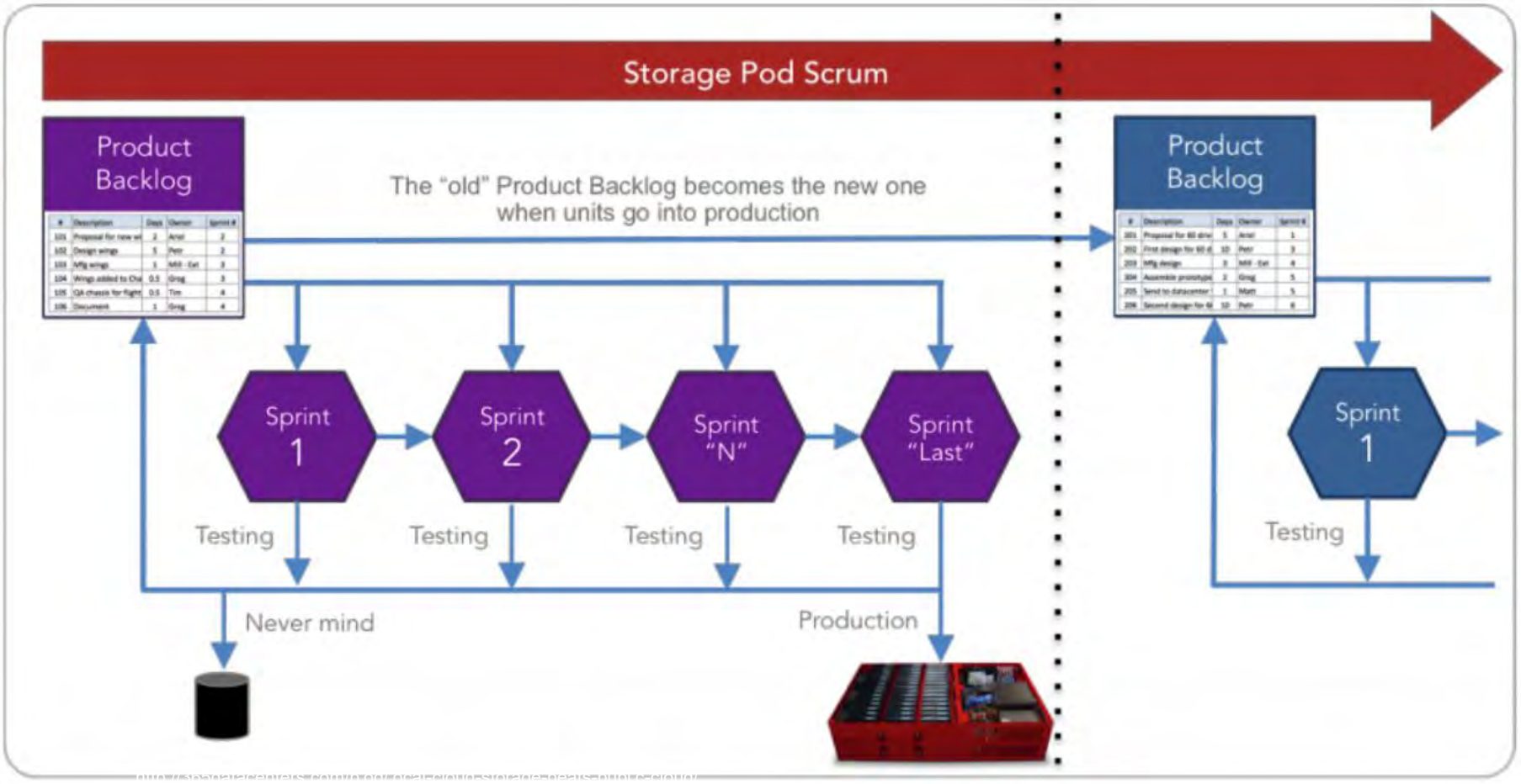


Image by Michael Cara
https://www.flickr.com/photos/michael_cala/

No Hardware Worshippers

Agile Extends to Hardware



BACKBLAZE

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?

BACKBLAZE

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/

BACKBLAZE

Backblaze B2 Use Cases

When cost matters…

- Archive
- Backup
- Tape Replacement
- Compliance
- Disaster Recovery

But also…

- Any storage that is ok with Internet-latency

BACKBLAZE

Learnings from Operating 200 PB of Disk-Based Storage

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

Gleb Budman
@glebbudman
Gleb.Budman@Backblaze.com