

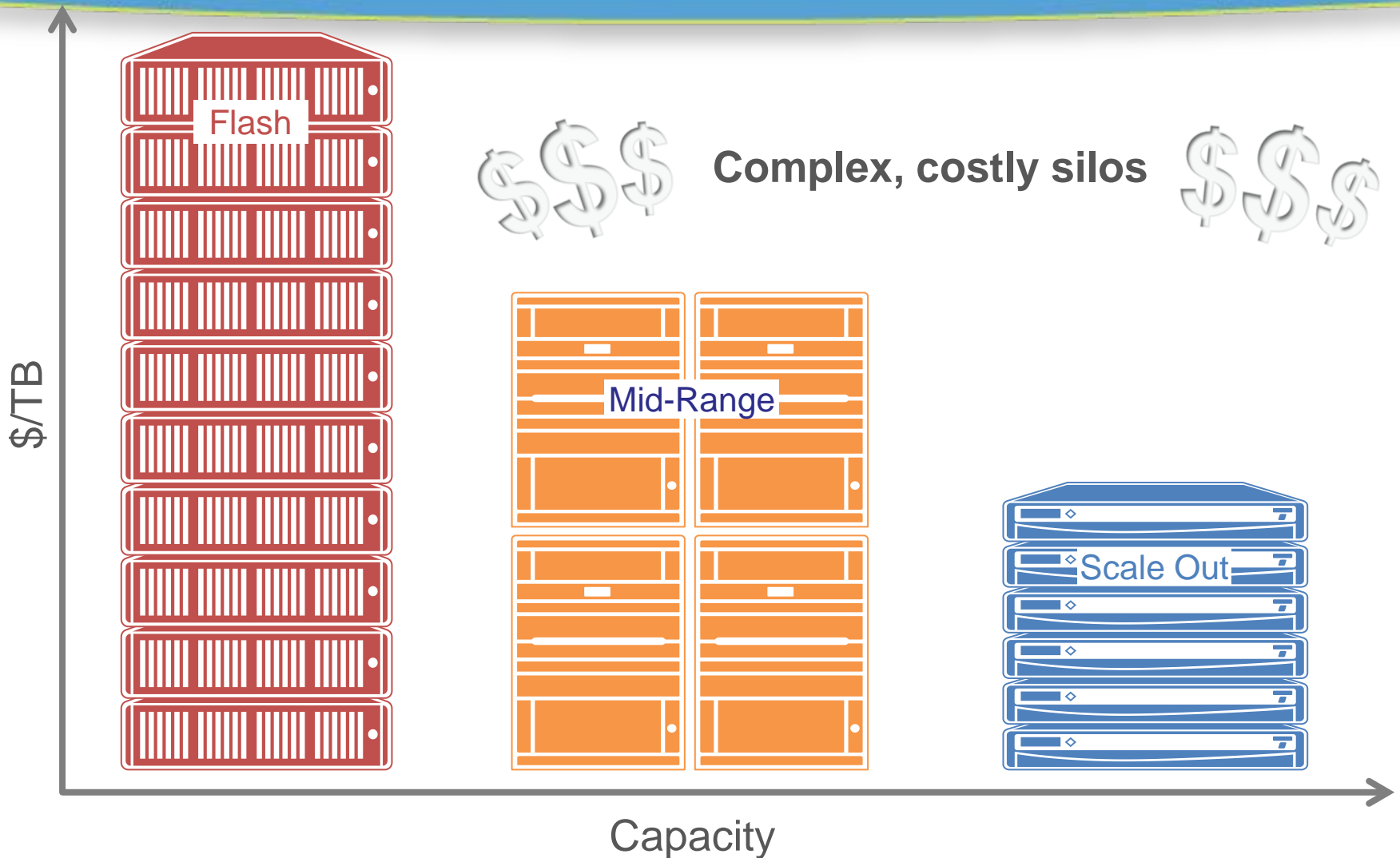


Rethink Cloud Strategies for Cost Effective Enterprise Storage Management

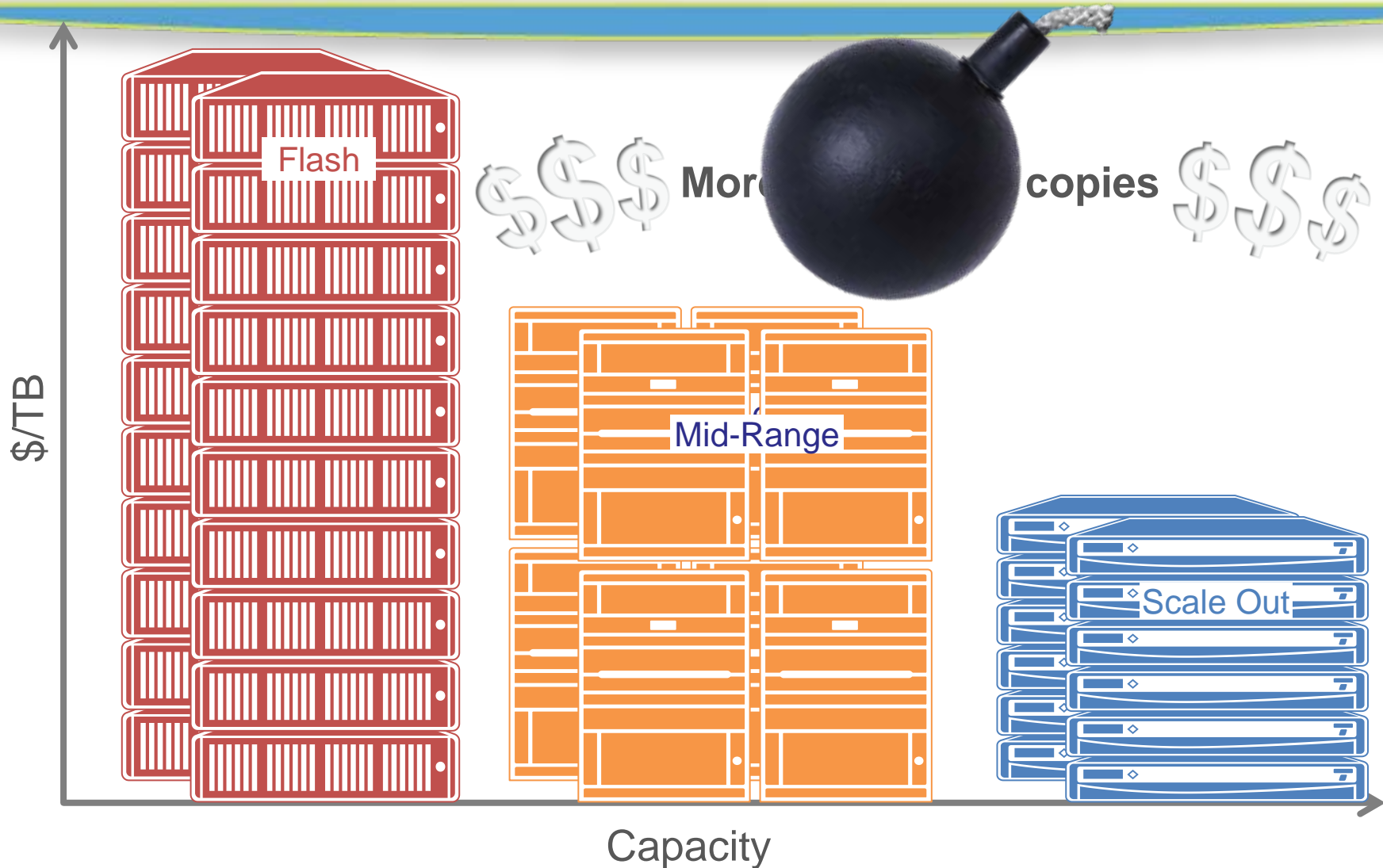
Laz Vekiarides
ClearSky Data



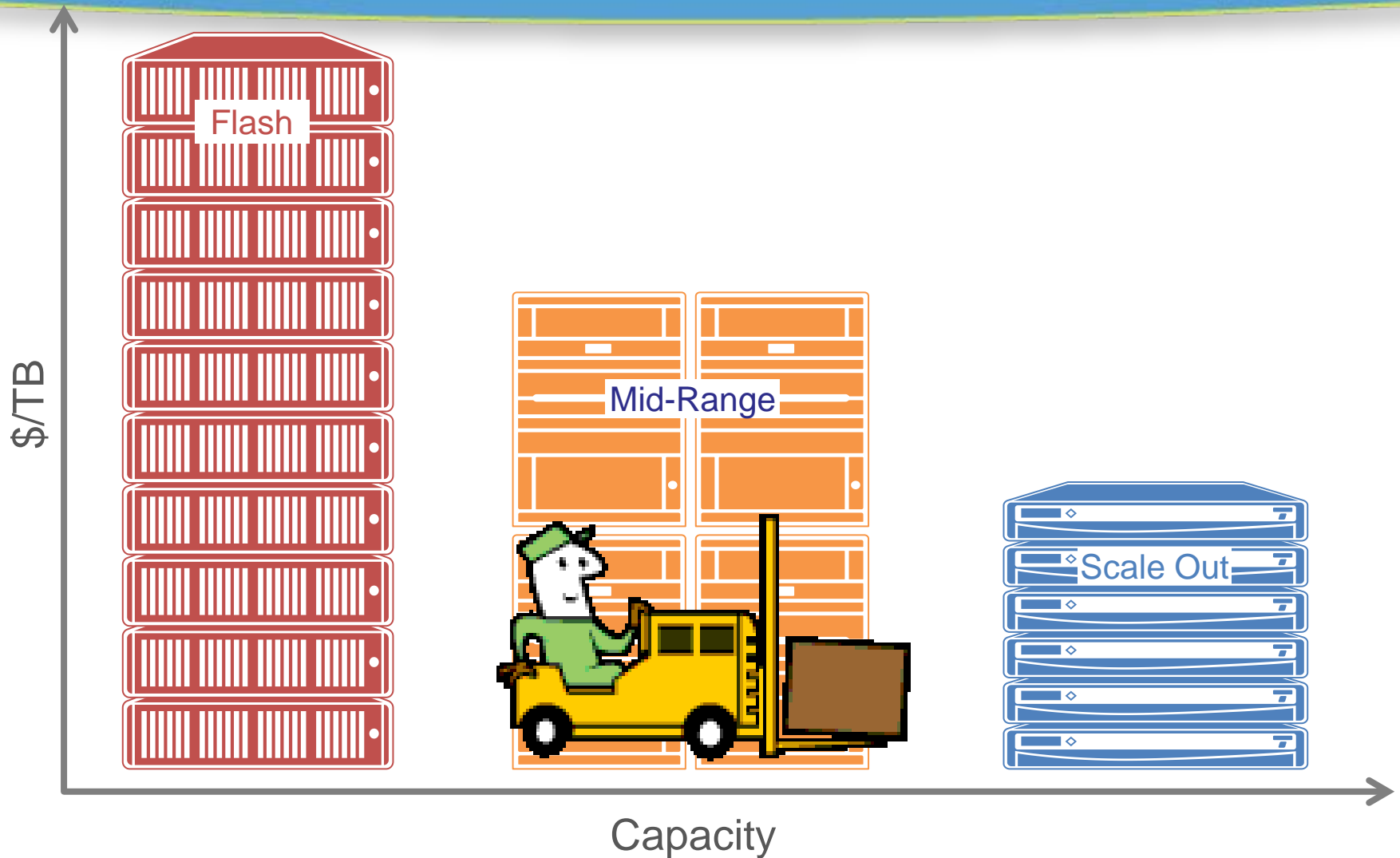
Enterprise Storage Today



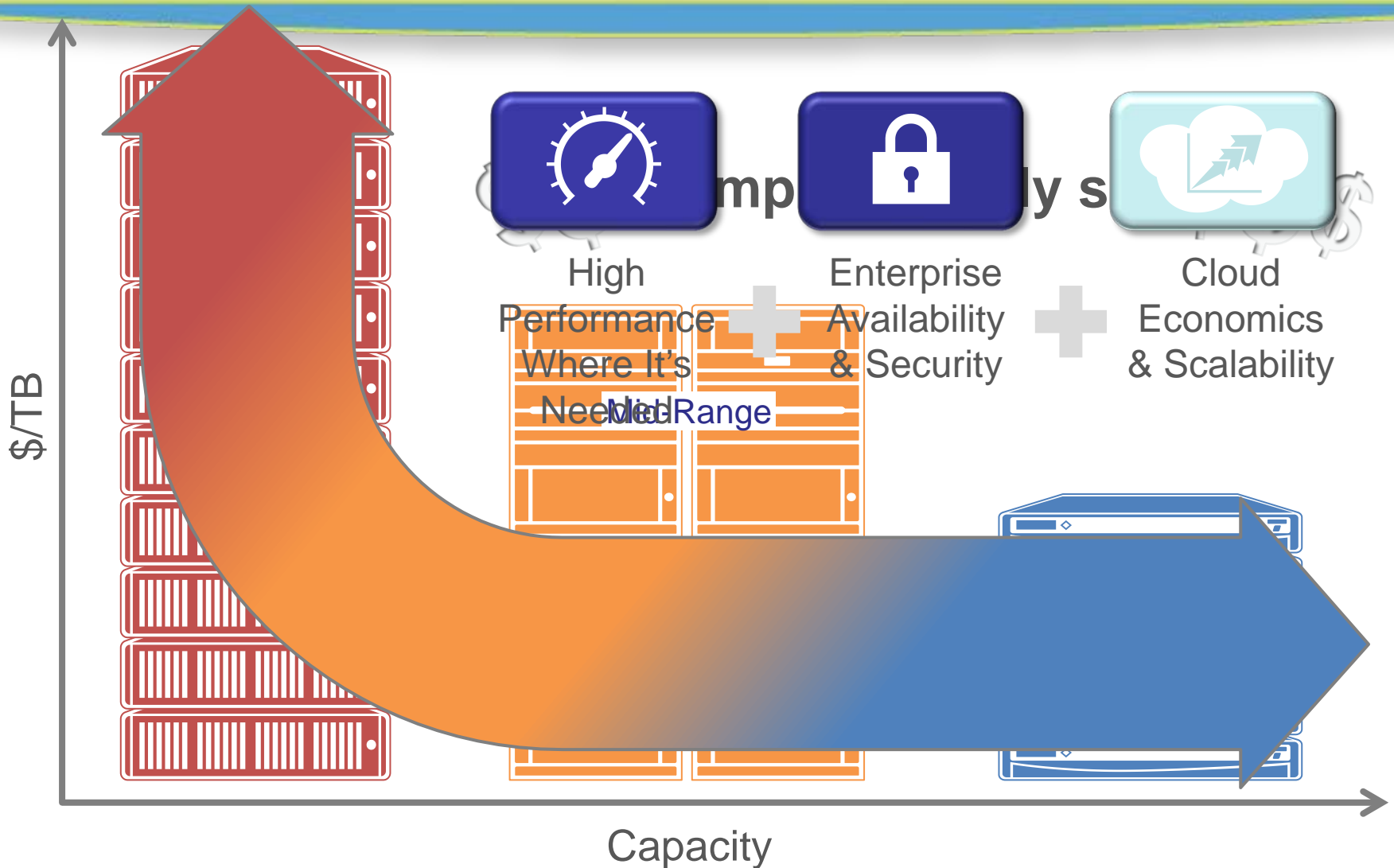
Then There is Disaster Recovery



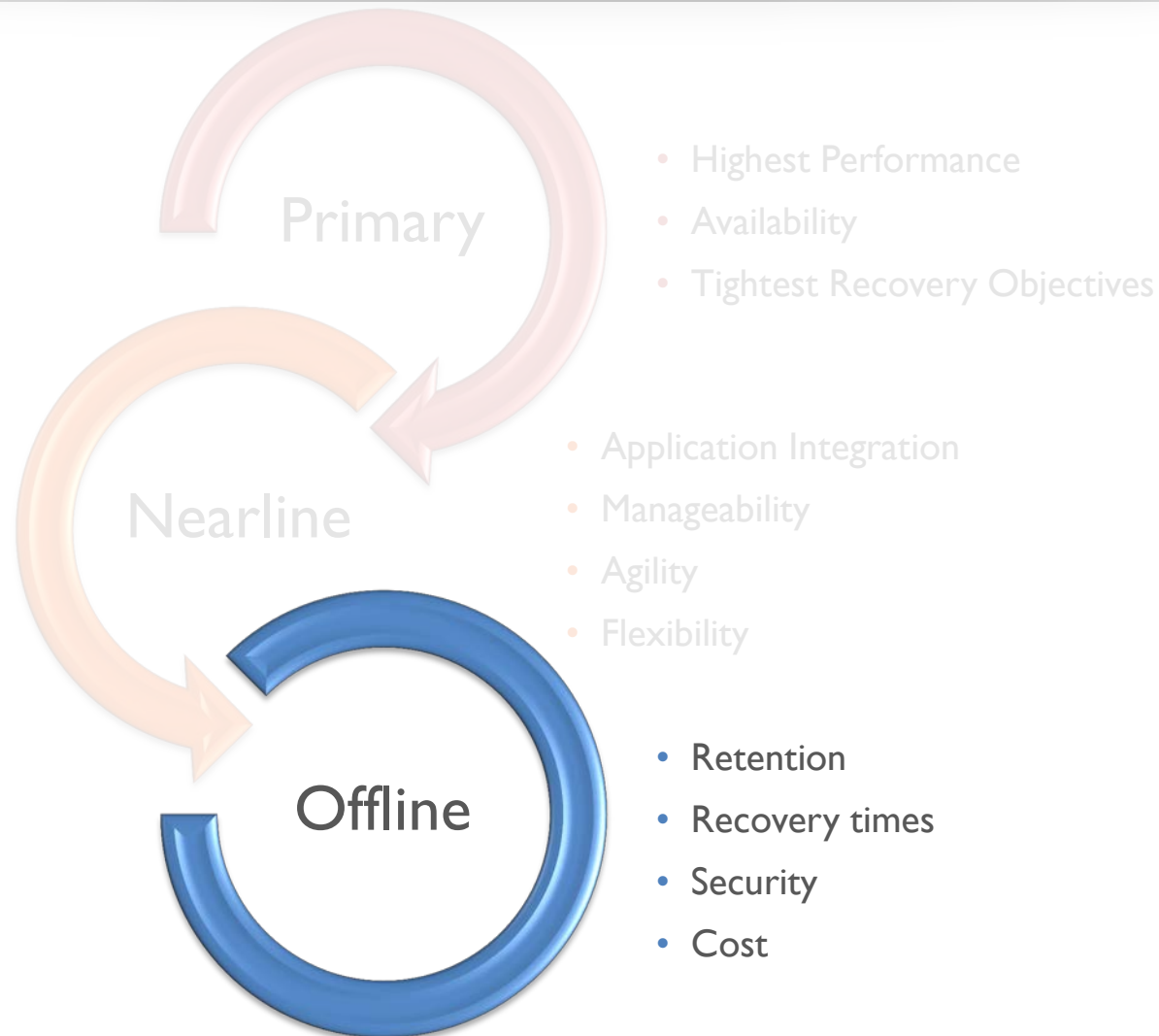
... And Then the Upgrades



What Enterprises Really Want



Today, Cloud Storage is “Where Data Goes to Die”

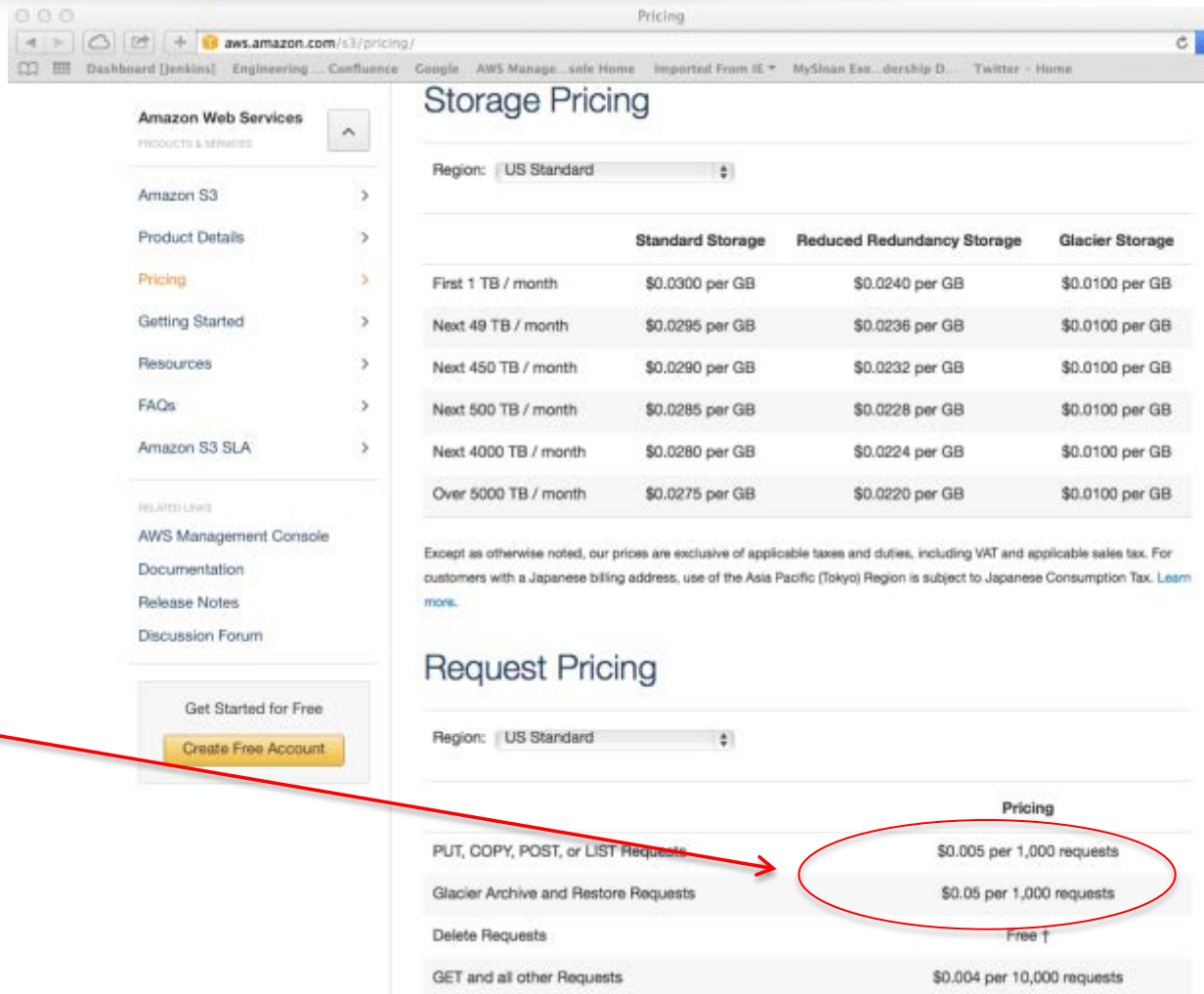


Closer Look At Storage Clouds

- Elastic, reliable, durable
- Pay as-you-go
- No hardware – mostly
- Really “feels” like software
- Seems well priced

BUT...

Look at the transaction costs!



The screenshot shows the AWS S3 Pricing page for the US Standard region. It includes a navigation menu on the left and two main pricing sections: Storage Pricing and Request Pricing.

Storage Pricing Table:

	Standard Storage	Reduced Redundancy Storage	Glacier Storage
First 1 TB / month	\$0.0300 per GB	\$0.0240 per GB	\$0.0100 per GB
Next 49 TB / month	\$0.0295 per GB	\$0.0236 per GB	\$0.0100 per GB
Next 450 TB / month	\$0.0290 per GB	\$0.0232 per GB	\$0.0100 per GB
Next 500 TB / month	\$0.0285 per GB	\$0.0228 per GB	\$0.0100 per GB
Next 4000 TB / month	\$0.0280 per GB	\$0.0224 per GB	\$0.0100 per GB
Over 5000 TB / month	\$0.0275 per GB	\$0.0220 per GB	\$0.0100 per GB

Request Pricing Table:

	Pricing
PUT, COPY, POST, or LIST Requests	\$0.005 per 1,000 requests
Glacier Archive and Restore Requests	\$0.05 per 1,000 requests
Delete Requests	Free †
GET and all other Requests	\$0.004 per 10,000 requests

A red arrow points from the text "Look at the transaction costs!" to the "Request Pricing" section, specifically highlighting the "Glacier Archive and Restore Requests" row which is circled in red.

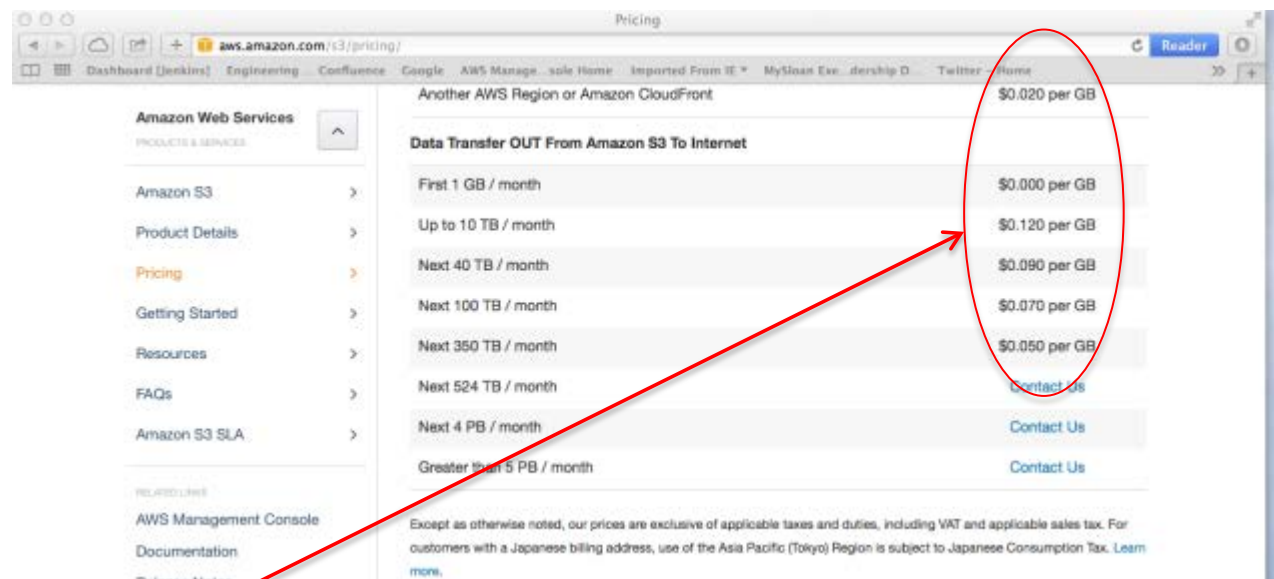
Closer Look At Storage Clouds

- Elastic, reliable, durable
- Pay as-you-go
- No hardware – mostly
- Really “feels” like software
- Seems well priced

BUT...

Look at the transaction costs!

Don't forget transfer costs!



Usage	Price
Another AWS Region or Amazon CloudFront	\$0.020 per GB
Data Transfer OUT From Amazon S3 To Internet	
First 1 GB / month	\$0.000 per GB
Up to 10 TB / month	\$0.120 per GB
Next 40 TB / month	\$0.090 per GB
Next 100 TB / month	\$0.070 per GB
Next 350 TB / month	\$0.050 per GB
Next 524 TB / month	Contact Us
Next 4 PB / month	Contact Us
Greater than 5 PB / month	Contact Us

Except as otherwise noted, our prices are exclusive of applicable taxes and duties, including VAT and applicable sales tax. For customers with a Japanese billing address, use of the Asia Pacific (Tokyo) Region is subject to Japanese Consumption Tax. [Learn more.](#)

A Simple “Enterprise” Cost Example

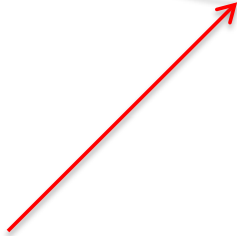
◆ Assumptions:

- 100 TB of storage
- 10K 8K write IOPS average
- 30K 8K read IOPS
- Trivial mapping between objects and I/Os
- No caching
- Use S3 rates

◆ Results get VERY expensive with heavy use

◆ Caching and packing provide only minimal cost benefit

Storage (100TB @ \$29/mo)	\$ 2900
PUT requests (10K/s-30 days)	129,600
GET requests (30K/s-30 days)	31,104
Bytes read	31,104
Bytes written	0
<hr/>	
TOTAL	\$ 194,708

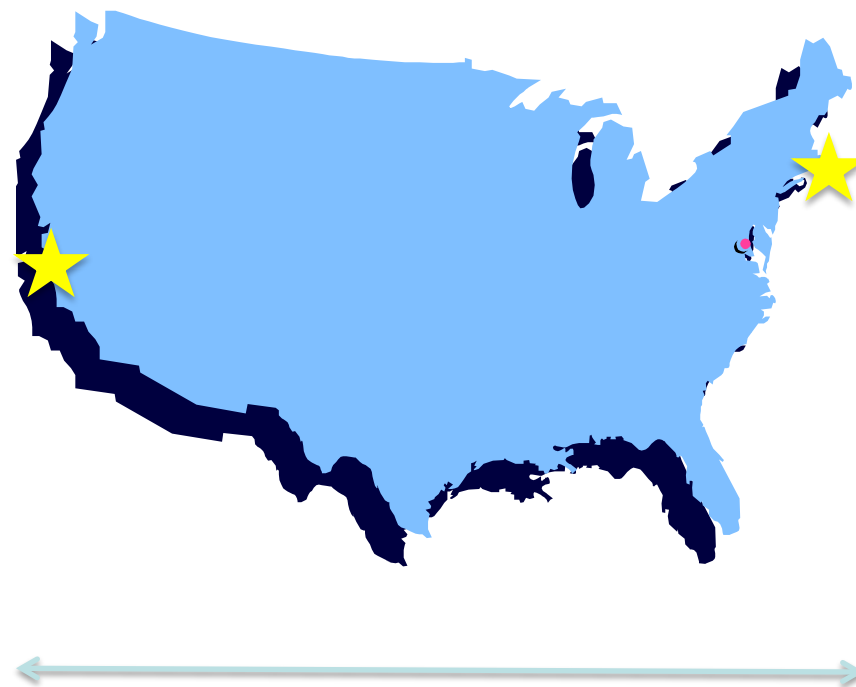


That's per month!

It's the Latency, Stupid

(Apologies to Stuart Cheshire)

- Data travels at the speed of light
- Fast - but finite
 - ◆ 3×10^8 meters per second
 - ◆ 186000 miles per second
- Example: Boston to San Francisco
 - ◆ 2740 miles
 - ◆ 29.4 milliseconds RT
- There are more delays
 - ◆ Light travels more slowly in fiber
 - ◆ Fiber-optic repeaters every few hundred miles
 - ◆ Switches, routers
 - ◆ Protocols, virtualization, etc.
- End result: ~60ms



So, Where Exactly Is “The Cloud”?

- Amazon East is near Ashburn, VA
- West is in San Francisco
- Boston is closest to East
- Best case numbers:
 - ◆ ~12ms round trip
 - ◆ From Markley via Direct Connect Ethernet
 - ◆ Does not include time to actually access the storage



What Latency Do Applications Expect?

- ◆ Applications see end-to-end latency
 - ◆ OS/Hypervisor
 - ◆ LAN/Network
 - ◆ Storage
 - ◆ Compute
- ◆ Storage is largest part
- ◆ Best practice <10ms

vmware®

Exchange

Microsoft®
SQL Server® 2012

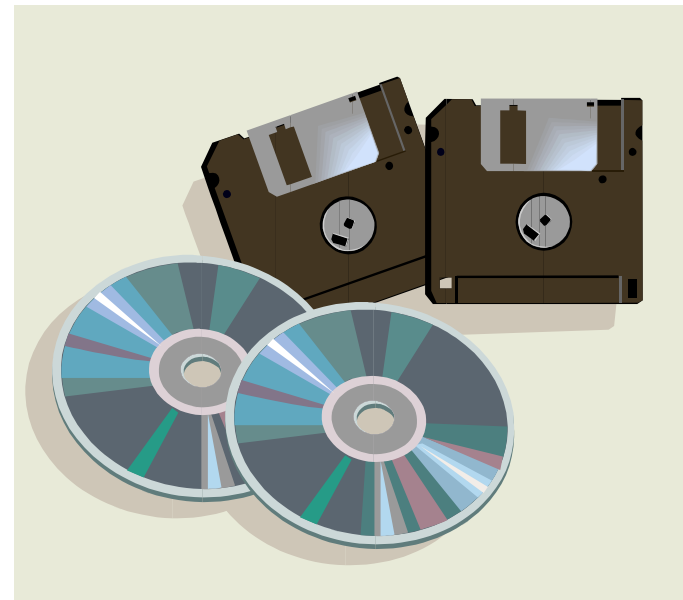
ORACLE®
DATABASE

Obvious conclusion: Compute must live near data

Microsoft

Some Applications Can Work Today

- ◆ Low performance, low intensity
 - ◆ Archive storage
 - ◆ Backup
 - ◆ Cold data
- ◆ Predominantly files
- ◆ Could be objects
 - ◆ Images
 - ◆ Movies
- ◆ Tape/removable disk use cases



So the conclusion?

- Cloud data is reasonably cheap as long as you don't access it much
- Data, apps, and users are happiest when they are nearest to each other
 - ◆ Network latency matters. Pick a good carrier.
 - ◆ Location matters. Pick a good datacenter 😊
- With a gateway, using cloud storage remotely is only suitable for cold data

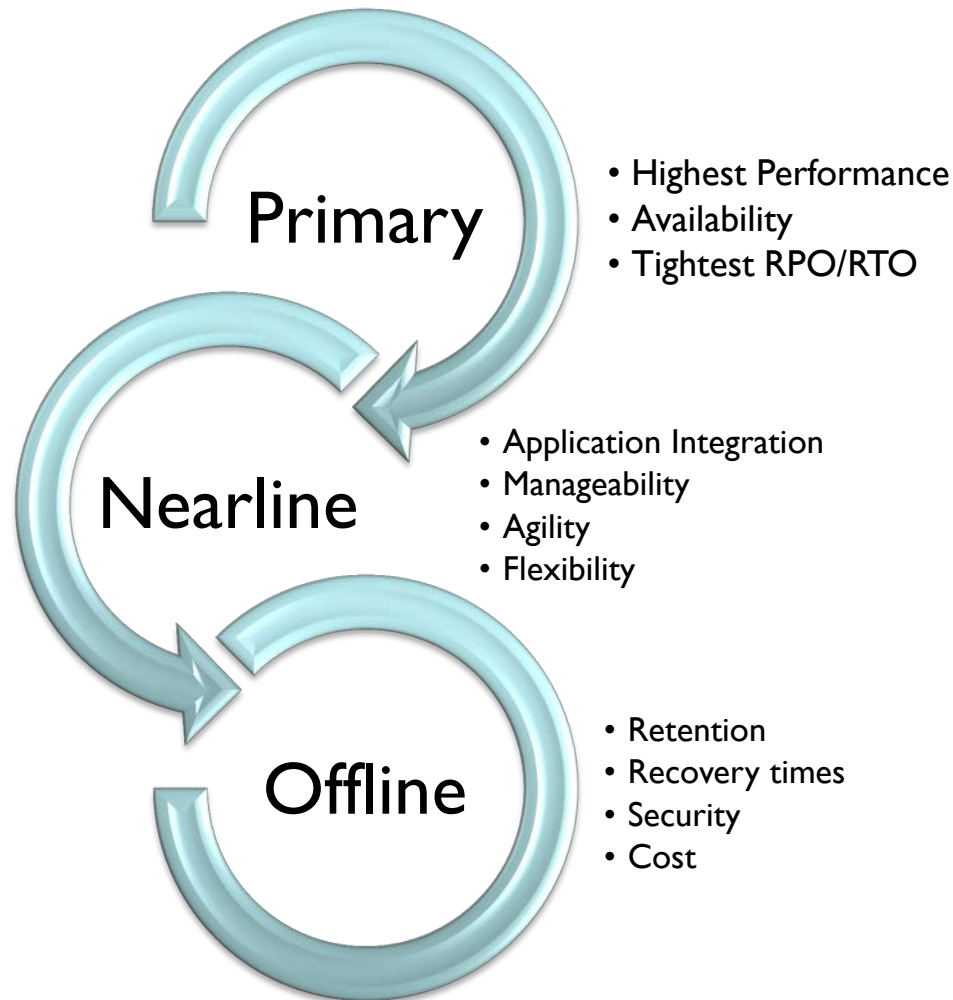
ClearSky Data: Enterprise Storage Is About to Change Forever.

Backups

What is “Cloud Storage”?

- **Scale-out object storage:**
 - ◆ Blobs of data and tags (Name, size, MD5 hash, etc)
 - ◆ Simple PUT, GET, DELETE operations
 - ◆ Accessed remotely or within a public cloud
- **Shamelessly lifted from S3 literature:**
 - ◆ Reliable
 - ◆ Scalable
 - ◆ Fast
 - ◆ Inexpensive
 - ◆ Simple
- **Caveats:**
 - ◆ NOT a file system
 - ◆ NOT a block (i.e. disk) device

Requirements For Data Storage Evolve Over The Course Of Its Life



Today, We Use Cloud For Data Only At The End of Life

