It’s dangerous to go alone!

How about *we* do this!?

Questions!

Our needs

Sustainability

Recognition of the benefits of preservation by decision makers

Process for ***Selecting*** digital materials with long-term value

Incentive for decision makers to preserve in the public interest

appropriate ***Organization*** and ***Governance*** of digital preservation activities

mechanisms to secure an ongoing, efficient ***Allocation*** of ***Resources*** for digital preservation activities

timely actions to ensure access

Locally or Consortially?

Source: Adam Jacob, OpsCode: Choose Your Own Adventure: http://www.youtube.com/watch?v=UD7qOfa4QDk

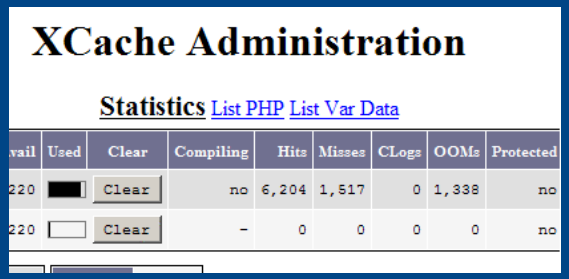
Technology ??

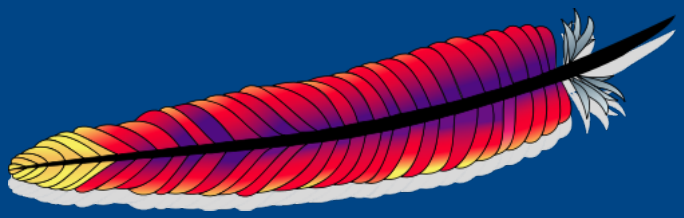
KEEPALIVED

KVM

HA
PROXY

VARNISH
CACHE

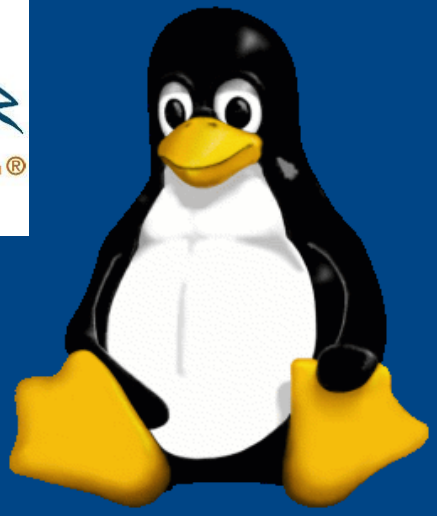

XCache Administration
Statistics List PHP List Var Data
vail Used Clear Compiling Hits Misses CLogs OOMs Protected
220 Clear no 6,204 1,517 0 1,338 no
220 Clear - 0 0 0 0 no



php

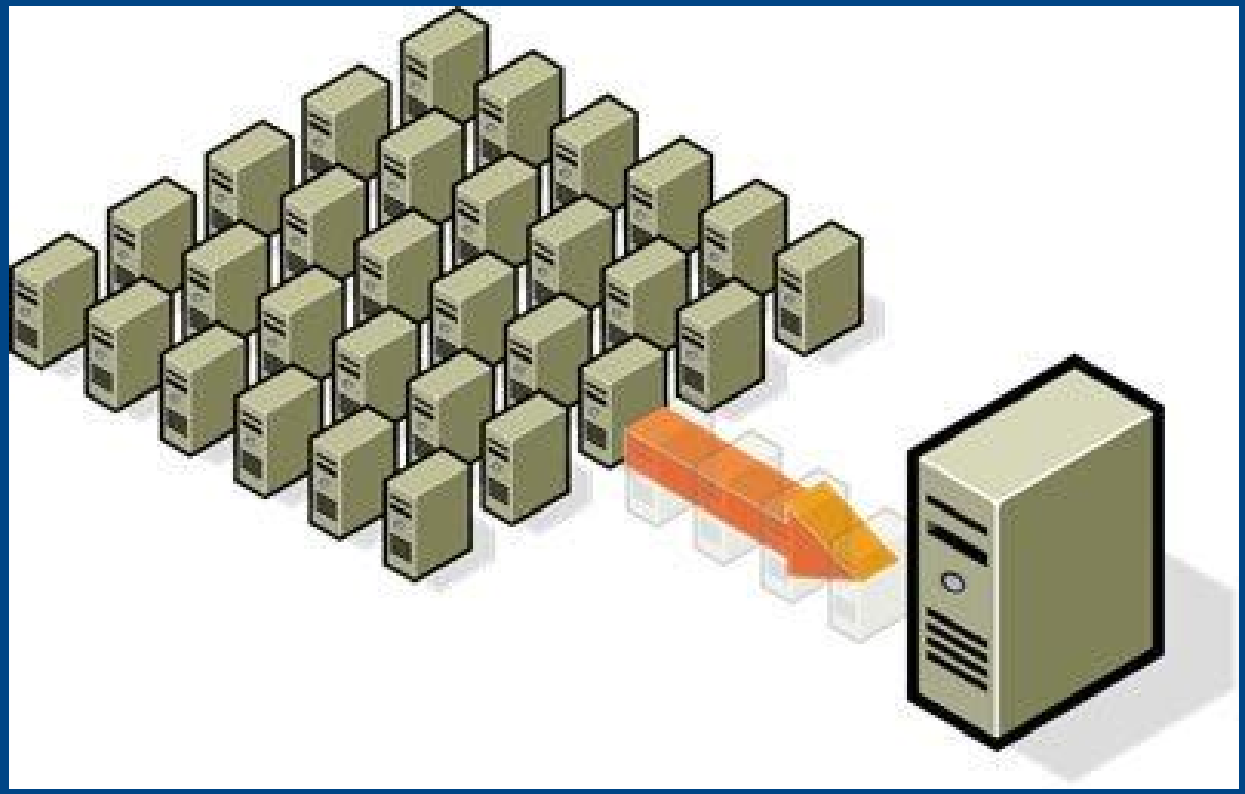
ubuntu

MySQL®

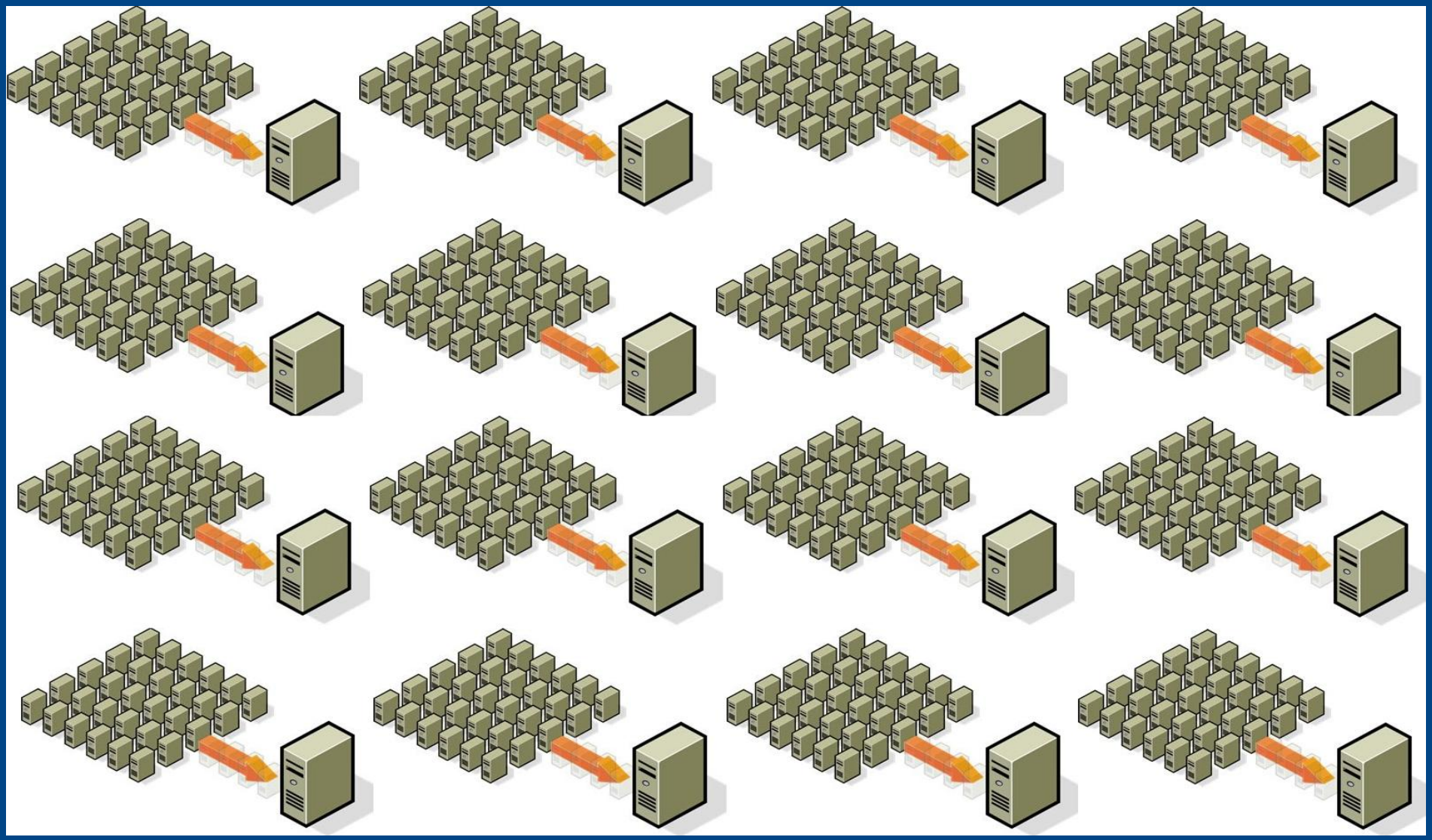


MEMCACHED

Server Virtualization

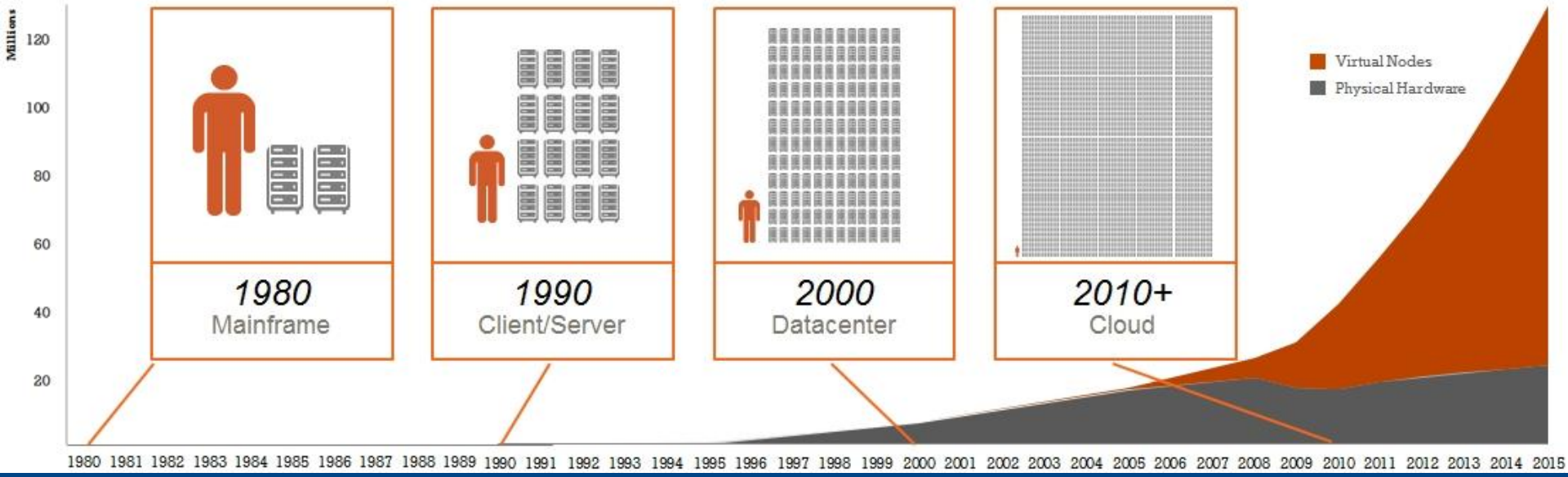


Source: http://www.techcrates.com/use-of-virtual-servers/



Source: http://www.techcrates.com/use-of-virtual-servers/

Scale x Complexity > Skills

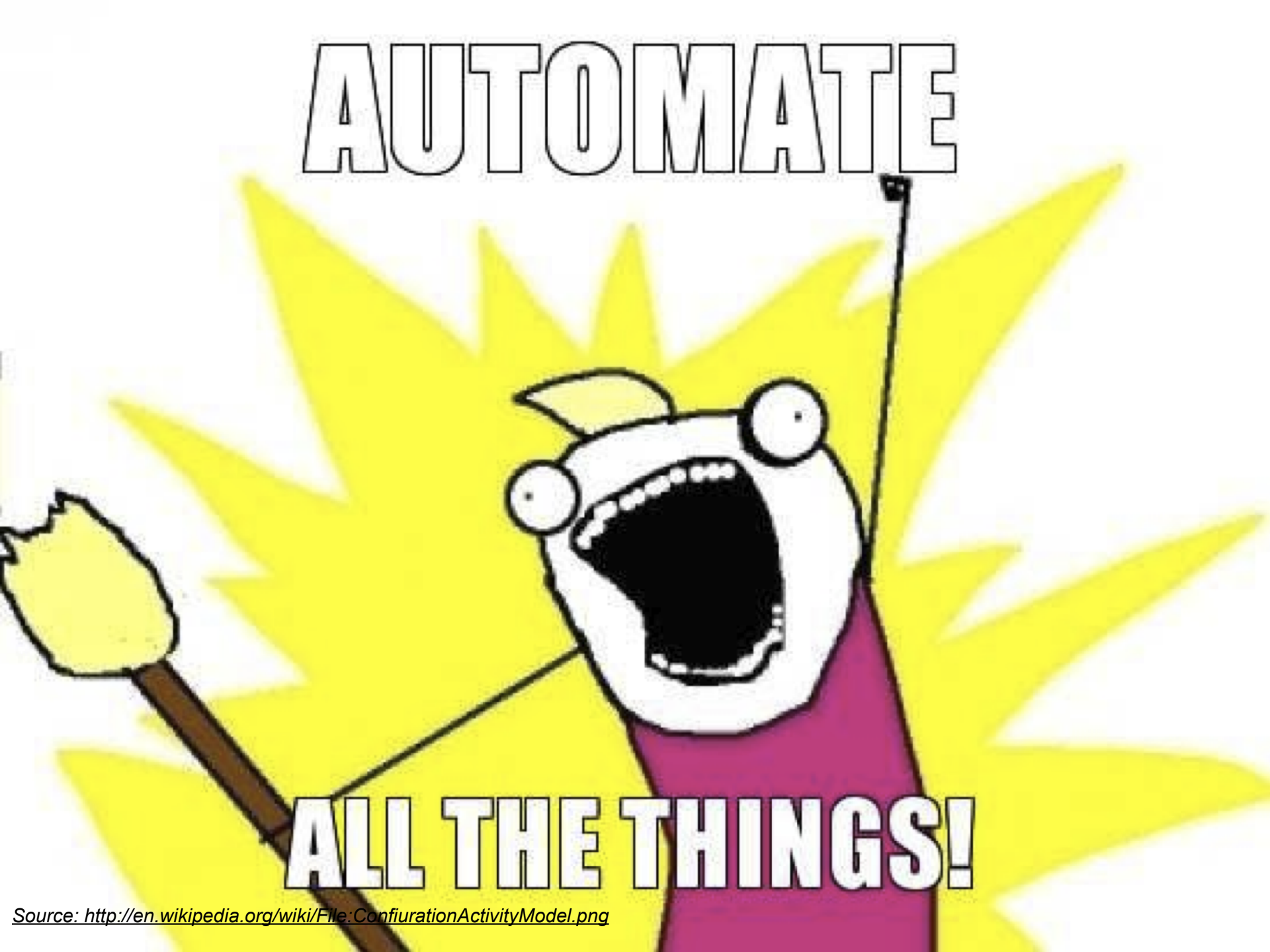


Source: http://www.theregister.co.uk/2012/03/26/opscode_funding_chef_momentum/

Configuration Management / DevOps

puppet labs

OPSCODE Chef



Source: http://en.wikipedia.org/wiki/File:ConfiurationActivityModel.png

Hardware commoditization / open source hardware

Intel-based x86 servers dominant in data centres for 10 years

Nothing particular to distinguish the products of different hardware vendors

Software processes,
virtualization,
configuration management
all well defined

Hardware: Does it really matter anymore?

Moderated by Lawrence Dignan | January 9, 2012 -- 07:00 GMT (23:00 PST)

Summary: *At the 2012 Consumer Electronics Show, hardware will take a backseat to software -- perhaps for good. Is hardware innovation alive and well?*

David Chernicoff

It's alive and well

or

Hardware? Ho-hum

Christopher Dawson

Best Argument: It's alive and well

List of Events: | *Opening Statements* | *Rebuttal* | *Closing Statements*

The moderater has delivered his final verdict.

Opening Statements

Hardware is what makes it possible

David Chernicoff: No matter how interesting software is, or how exciting the user experience becomes, without continued innovation in hardware technologies, software will begin to

Merely a matter of preference

Chris Dawson: As the world moves online, hardware has become nothing more than a religion. I tend to use Macs because I like the OS and the hardware takes the abuse I give it. On

ZDNet | Follow @zdnet | Join | Log In | Privacy | Cookies

Source: http://www.zdnet.com/debate/hardware-does-it-really-matter-anymore/10086653/

As Companies Focus on Services, Will Hardware Become Irrelevant?

Published: November 07, 2012 in Knowledge@Wharton

In recent weeks, Amazon.com, Microsoft and Apple have all held splashy events to tout their newest tablets -- the Kindle Fire HD, the Surface and the iPad Mini. But even as Amazon CEO Jeff Bezos talked up the specs of the new Kindle Fire, he noted that the market is becoming less about the gadgets themselves, and more about how companies can make money off additional services and purchases that appeal to device users.

Technology companies now see the buying of a tablet as merely the beginning of a long relationship with a consumer -- a relationship that will extend across multiple platforms and content types, with users periodically paying subscription fees or purchasing individual applications, music, movies and television shows. And if the firm is assured of a future payoff, it can offset the losses from charging less for the device.

But making such a transition is easier said than done, according to Wharton legal studies and business ethics professor Andrea Matwyshyn. "At present, this model is limited to only a few players," she says. "The services model definitely works for Amazon, given its

Print
Get as a PDF
Send a comment
Share this Article

Source: http://knowledge.wharton.upenn.edu/article.cfm?articleid=3106

Rethinking the hardware entirely: open source hardware

Source: http://www.opencompute.org/

5 System View

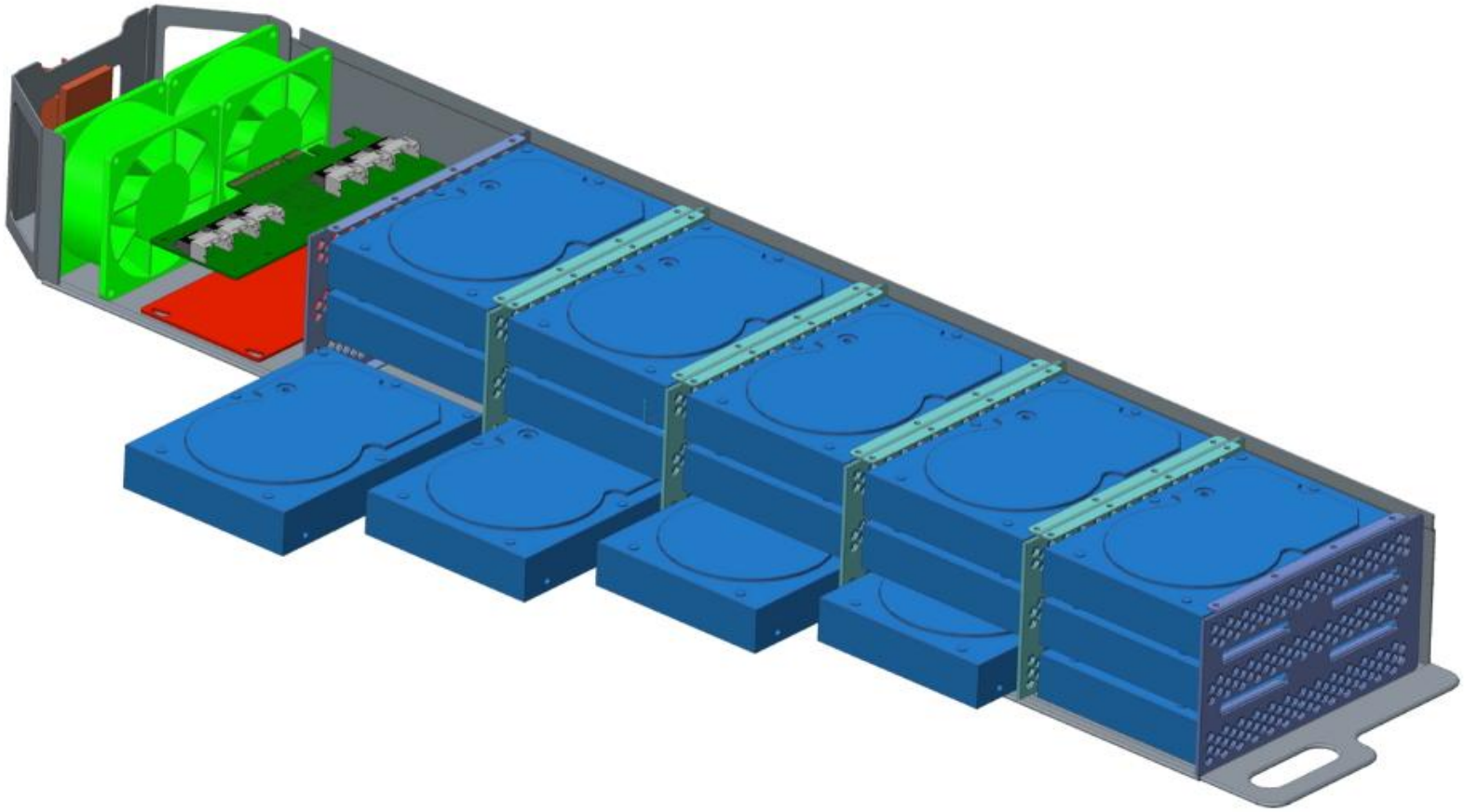
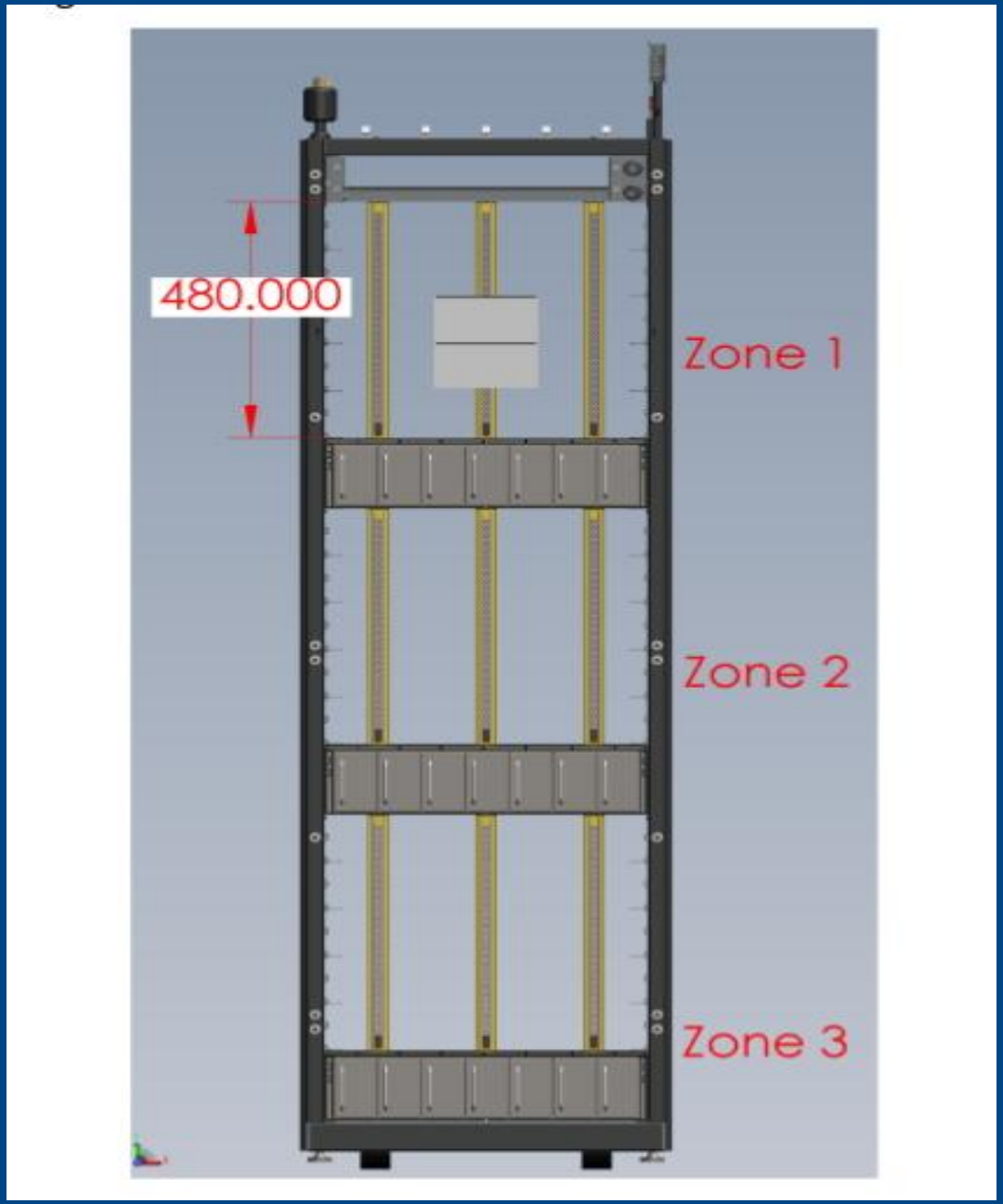


Figure 1 System Overview

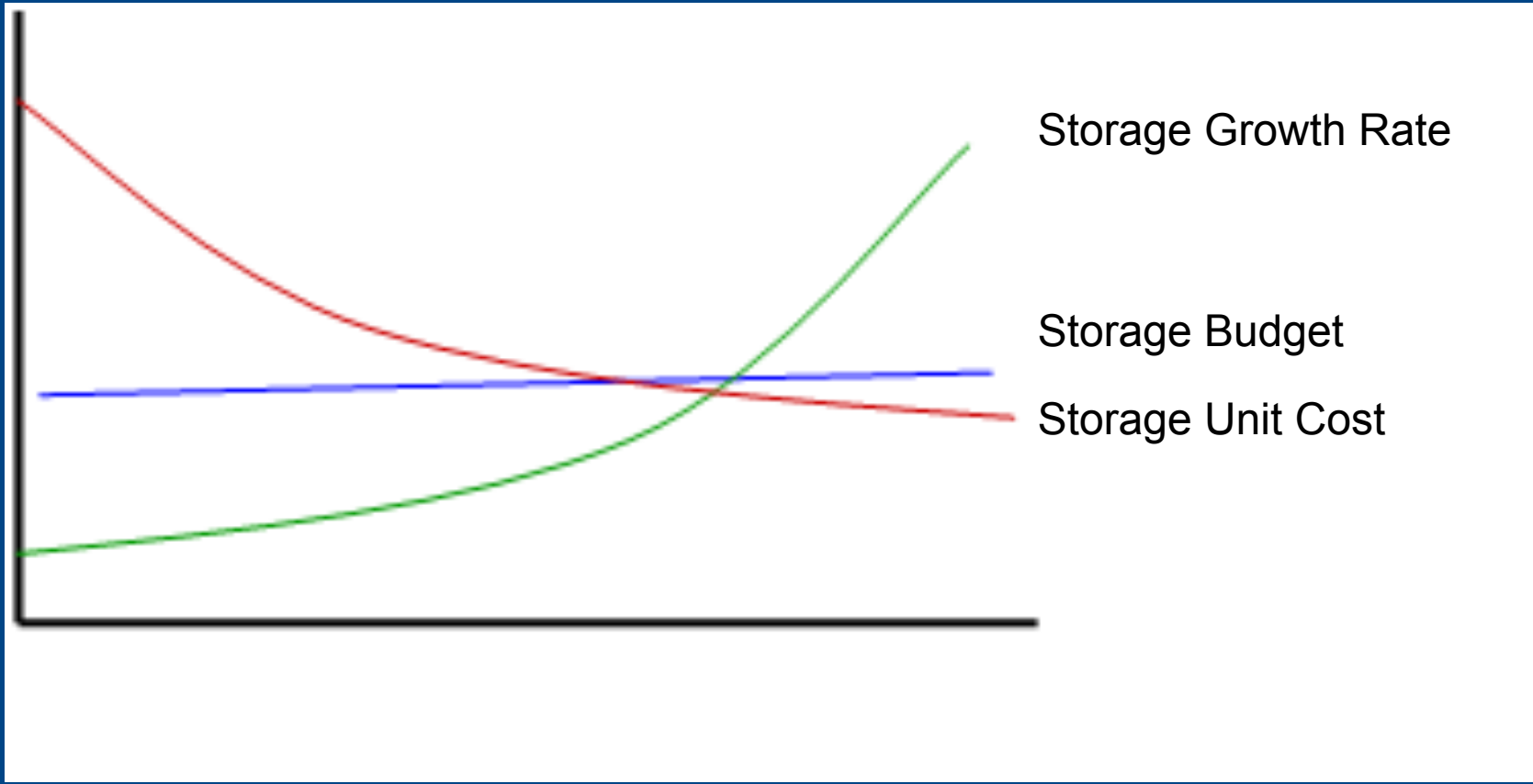
Source: http://www.opencompute.org/wp/wp-content/uploads/2013/01/FIO-3-2TB-I-O-Acceleration-HW-v06-OCP-24Jan2013.pdf



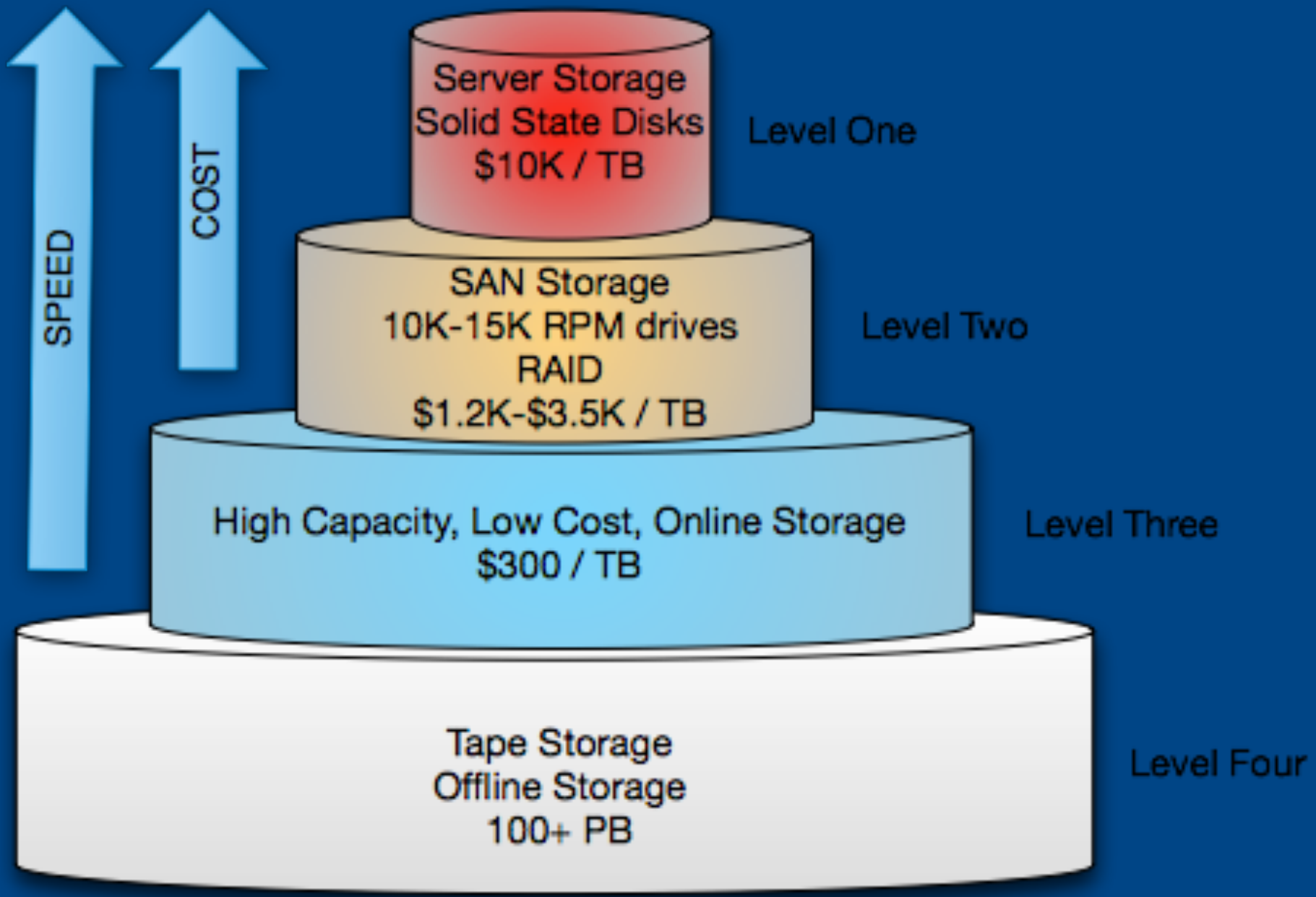
Source: http://www.opencompute.org/wp/wp-content/uploads/2013/02/Open_Compute_Project_Intel_Server_Open_Rack_v0.3.pdf

Software on commodity hardware can replicate costly vendor systems: Software defined storage

Vendor storage: cost vs. budget vs.growth



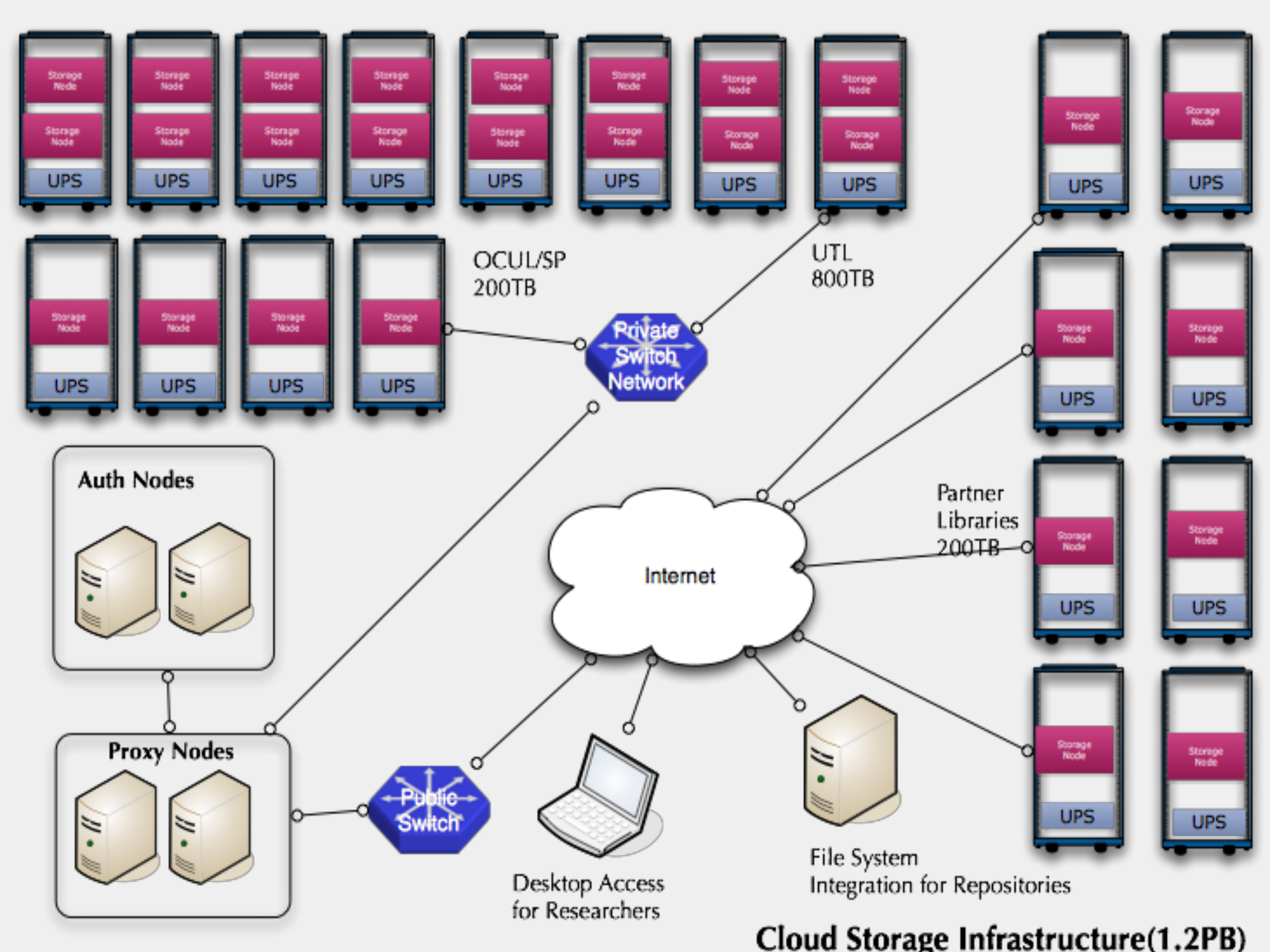
Software defined storage


SPEED
COST
Server Storage
Solid State Disks
$10K / TB
Level One
SAN Storage
10K-15K RPM drives
RAID
$1.2K-$3.5K / TB
Level Two
High Capacity, Low Cost, Online Storage
$300 / TB
Level Three
Tape Storage
Offline Storage
100+ PB
Level Four

Storage Options & Cost

Storage Type	Price/TB/Yr	Price/1.2PB/Yr
Amazon S3 (1)(5)	$829.59	$995,512
Rackspace (2)(5)	$960.10	$1,152,123
Duracloud Preservation (3)(5)	$1,000.63	$1,200,750
Microsoft Azure(4)(5)	$829.59	$995,512

(1) http://aws.amazon.com/s3/pricing/
(2) http://www.rackspace.com/cloud/files/pricing/
(3) http://www.duracloud.org/pricing
(4) http://www.windowsazure.com/en-us/pricing/details/storage/
(5) USD on a Tiered ” Pay As You Go” Plan, data is geographically redundant



Cloud Storage Infrastructure(1.2PB)

Storage Node

Commodity Hardware

- Off the shelf consumer grade disks
- Software Replication (3X)
- Fixity & Self healing
- Geographically distributed
- Scalable… HIGHLY Scalable!
- Object Storage

Block vs. Object Storage

Block Storage

- Traditional storage
- SCSI or ATA or iSCSI or FC
- Access data via OS
- “C:\” or “/dev/sda” or “Macintosh HD”
- Attached to one server at a time

Block vs. Object Storage

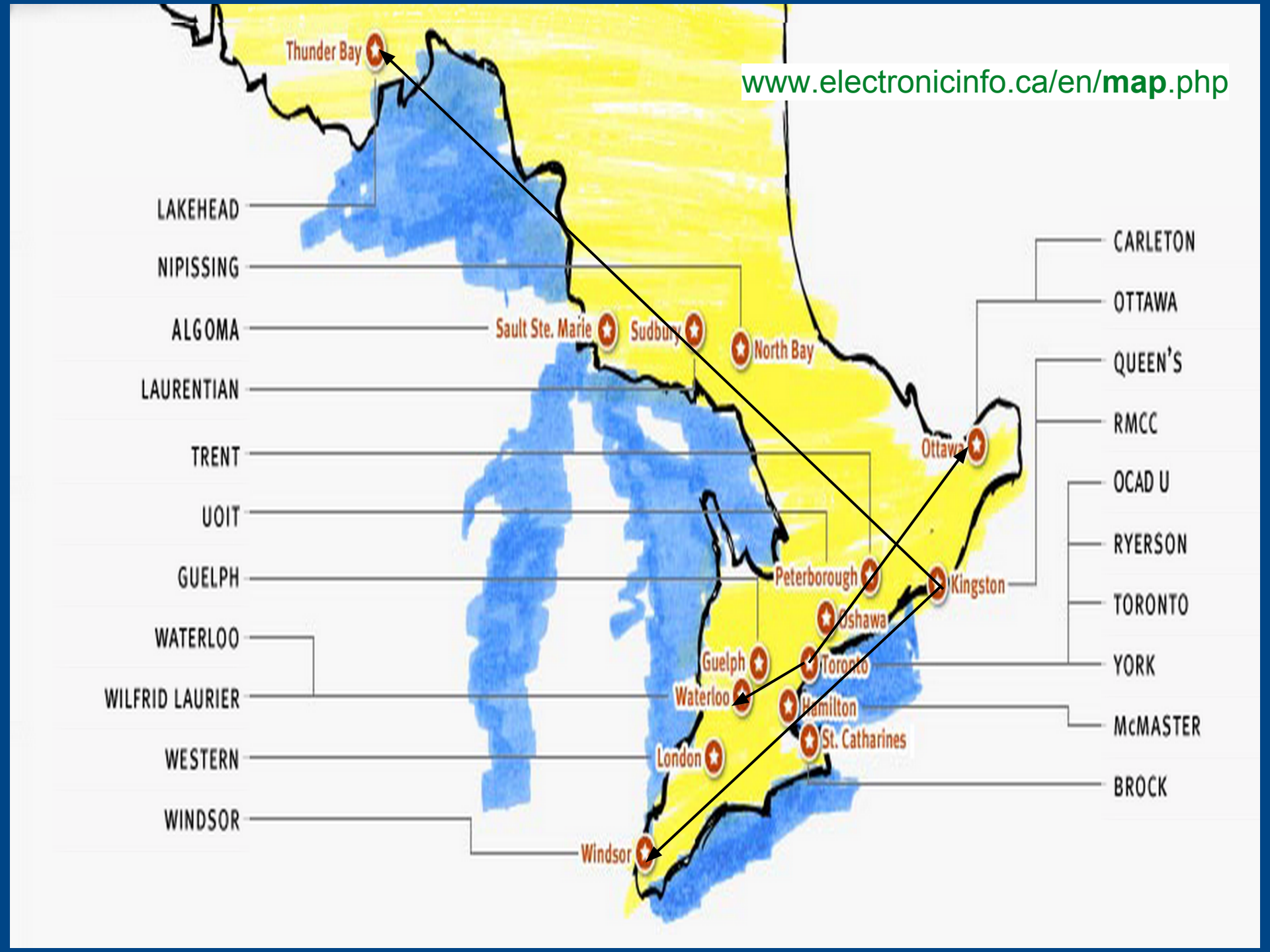
Object Storage

- Access data via HTTP
- GET, POST, DELETE
- Build applications using REST API

```
swift -A https://oculcloud/auth/v1.0 -U system:root -K testpass upload thumbnails dataverse/logo.png
swift -A https://oculcloud/auth/v1.0 -U system:root -K testpass upload myfiles dataverse/head_right.png
```

- Static content (PDFs, jpgs)
- All files have a URL

Cyberduck Demo


www.electronicinfo.ca/en/map.php
Thunder Bay
LAKEHEAD
NIPISSING
ALGOMA
Sault Ste. Marie
Sudbury
North Bay
LAURENTIAN
TRENT
UOIT
GUELPH
WATERLOO
WILFRID LAURIER
WESTERN
WINDSOR
Windsor
London
Waterloo
Guelph
Peterborough
Oshawa
Toronto
Hamilton
St. Catharines
Kingston
Ottawa
CARLETON
OTTAWA
QUEEN'S
RMCC
OCAD U
RYERSON
TORONTO
YORK
McMASTER
BROCK

Storage Options & Cost

Storage Type	Price/TB/Yr	Price/1.2PB/Yr
Amazon S3 (1)(5)	$829.59	$995,512
Rackspace (2)(5)	$960.10	$1,152,123
Duracloud Preservation (3)(5)	$1,000.63	$1,200,750
Microsoft Azure(4)(5)	$829.59	$995,512
OCUL(6)	$350	$420,000

(1) http://aws.amazon.com/s3/pricing/
(2) http://www.rackspace.com/cloud/files/pricing/
(3) http://www.duracloud.org/pricing
(4) http://www.windowsazure.com/en-us/pricing/details/storage/
(5) USD on a Tiered ” Pay As You Go” Plan, data is geographically redundant
(6) Includes cost of additional staff

So what?

Interesting because:

We're envisioning a solution that can help us out in a lot of different areas:

- Scholars Portal platforms
- Dataverse/other RDM initiatives
- Institutional Repositories
- New applications
- Just space in general!

with lots of different functional requirements:

- Autonomic digital preservation activities
- Lots of different interfaces
- Integration with HPC centers

Important because:

- Cost
- Jurisdiction
- Self-sufficiency

Which brings us to the point….

IT'S DANGEROUS TO GO
ALONE! TAKE THIS.

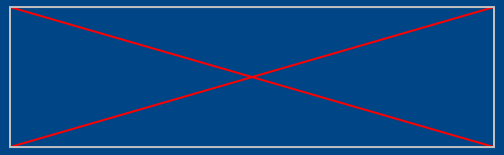
IT'S DANGEROUS TO GO
ALONE! TAKE THIS.

IT'S DANGEROUS TO GO
ALONE! TAKE THIS.

IT'S DANGEROUS TO GO
ALONE! TAKE THIS.

IT'S DANGEROUS TO GO
ALONE! TAKE THIS.

If it's worth doing here….



University of Waterloo

McMaster University

Queen's University

YORK UNIVERSITÉ UNIVERSITY

UNIVERSITY OF TORONTO

University of Windsor

LAURIER Inspiring Lives.

UNIVERSITY of GUELPH

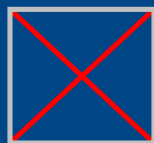
uOttawa

….it’s worth doing here!

Contact

Nick Ruest
ruestn@yorku.ca/@ruebot

Steve Marks
steve@scholarsportal.info/@sbmarks

Amaz Taufique
amaz@scholarsportal.info/@amazofficial

Graham Stewart graham.stewart@utoronto.ca