

# Another Year, Another Petabyte

A Look Into the Laureate Institute for Brain  
Research's CephFS Deployment

## Research At LIBR

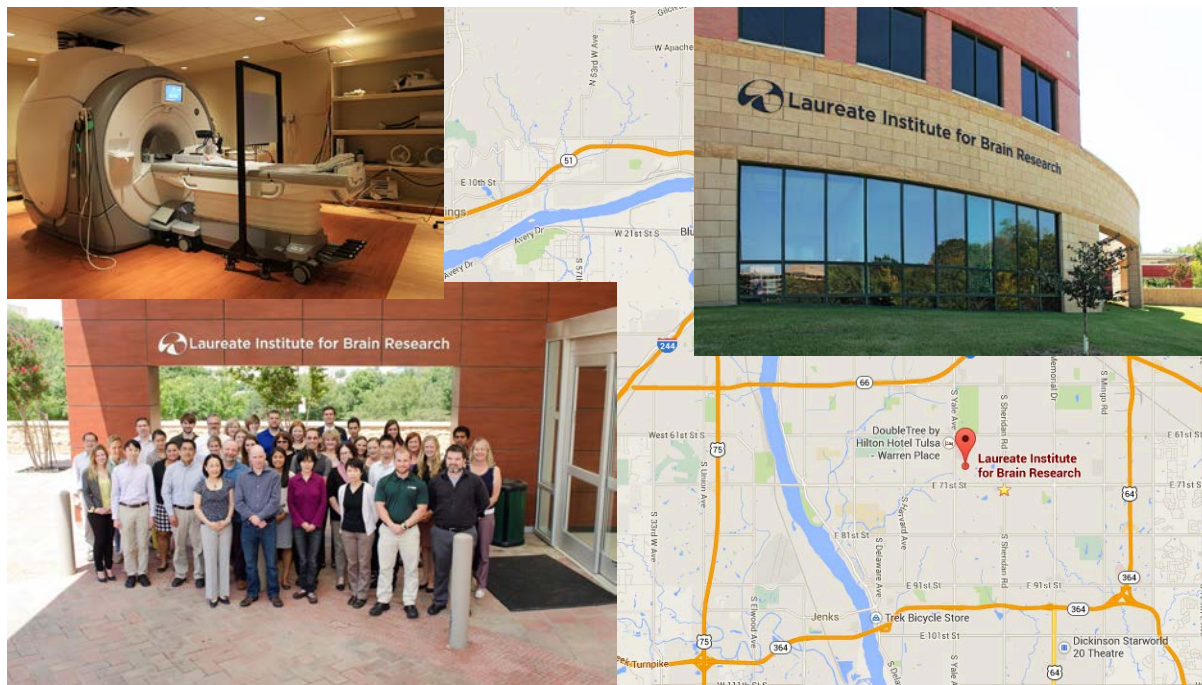
- **Neuroscience-based, clinical and developmental research:**
  1. To develop neuroscience-based individually predictive assessments.
  2. To develop novel brain-body based interventions
    - Focus: mood, anxiety, addiction, or eating disorders.
  3. To use experimental systems to quickly test assessments and interventions.



# LIBR

Laureate Institute for Brain Research

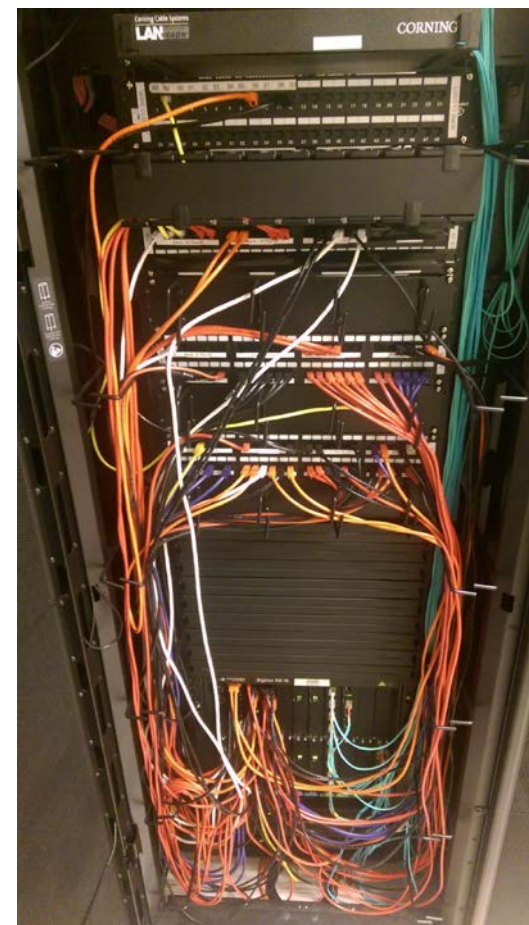
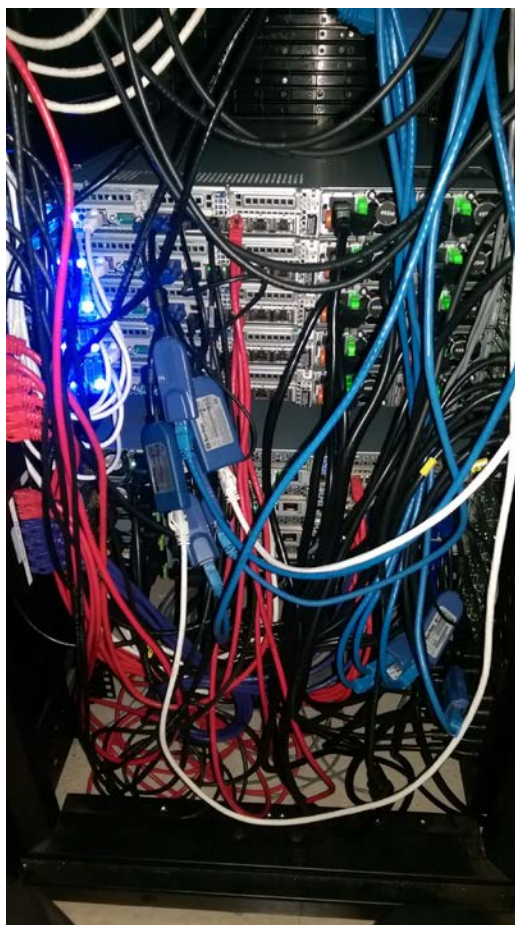
## The LIBR Facilities





## My Inheritance November 2015

- Foundry Big Iron RX-16
- EMC Isilon
- Oracle ZFS Appliance
- Spectra Logic Black Pearl
- Zmanda
- Spectra Logic T950e





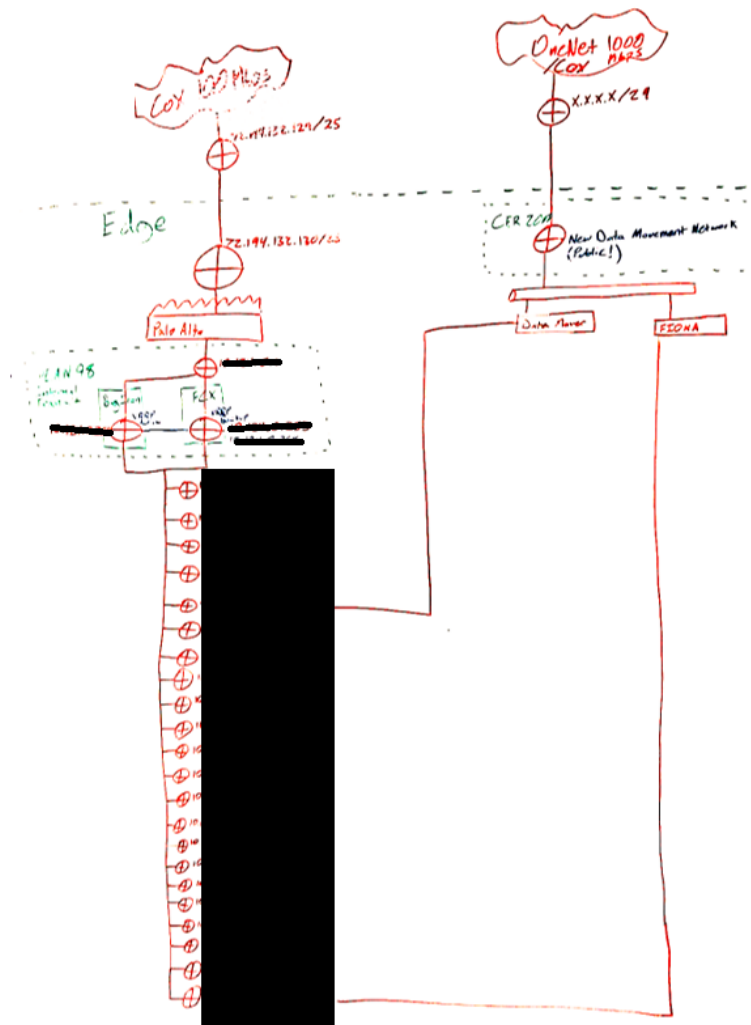
## The Problem

- Depletion of compute/storage resources
- Neglected infrastructure
- \$\$\$
- Inadequate performance
- Unreliable





## Network Diagram





## Options Presented to Stakeholders January 2016

- Scale existing Isilon with used hardware (\$1.3M)
- Panasas (\$630k)
- Oracle ZFS (\$450k)
- DiY scale-out + Network Refresh (\$450k)
  - Lustre
  - Gluster
  - Ceph
  - OrangeFS



## Slow Progress

- February – Approval to develop pilot
- June – pilot approval
- July – Pilot Equipment arrival
- July – CEO approval to proceed



# Board of Directors Approval August 2016

- Placed Order
- Brings configuration to:
  - 8x OSD (24 disk + 2 journal NVMe per node)
    - 1.1PB raw
  - 3x MON
  - 2x MDS



## Ceph Design

- 8x OSD nodes
  - 256 GB RAM
  - 2x Intel S3610 for OS
  - 24x 6TB Enterprise SATA (24 slot chassis)
  - 2x Intel P3700
  - 2x Xeon 2660 V4
    - 28x2.0GHz
  - Dual 40GB Ethernet Adapter (Mellanox ConnectX-4)



## Ceph Design (Cont...)

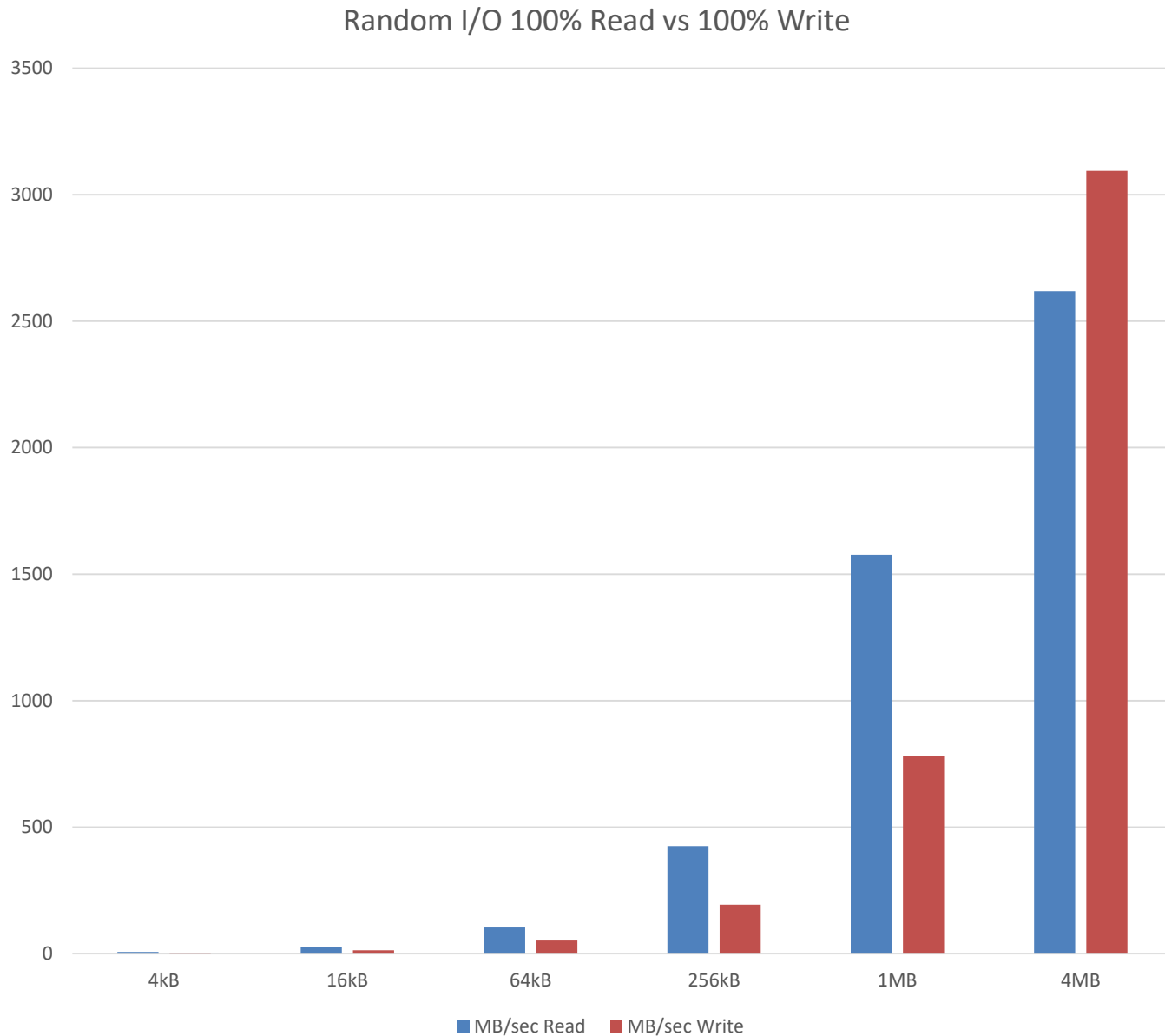
- 2x MDS + 3x MON node
  - 128 GB RAM
  - 2x Intel S3610 for OS
  - 2x Xeon 2643 V4
    - 12x3.4GHz
  - Dual 40GB Ethernet Adapter (Mellanox ConnectX-4)



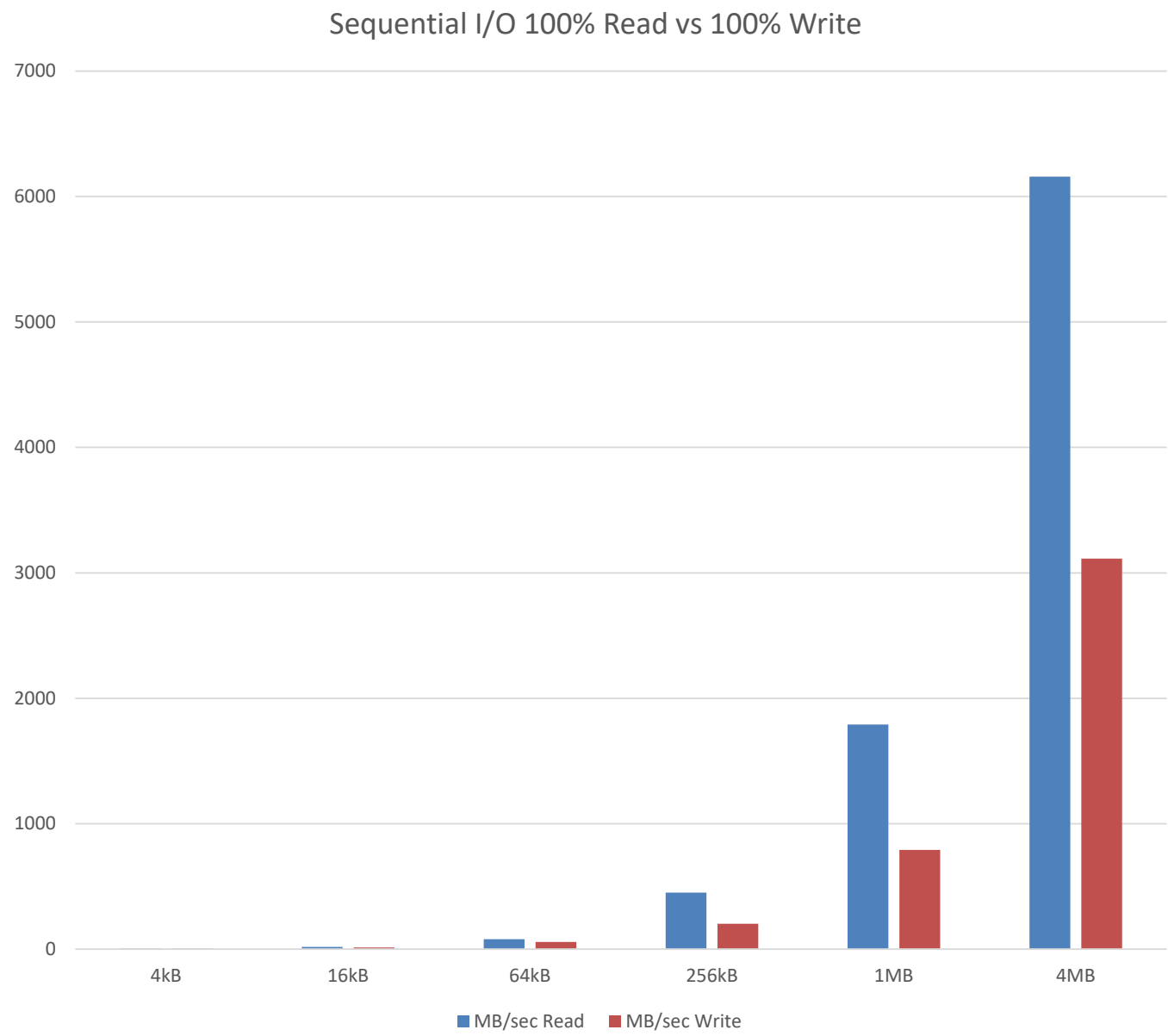
## Other Hardware and Software

- Brocade VDX
- Spectra Logic
- Zabbix
- ELK
- Nfs-ganesha
- Samba

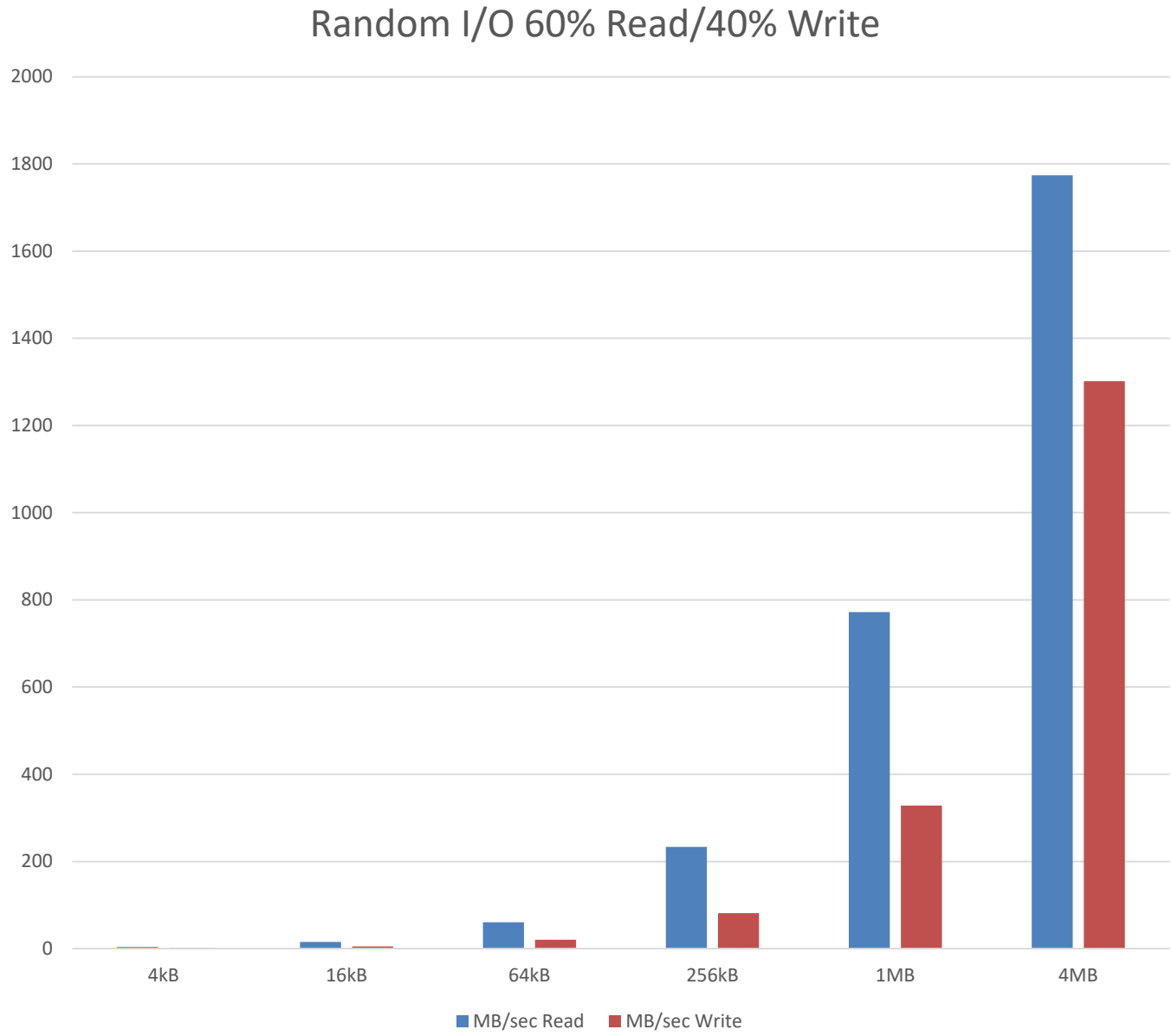
Blocksize	MB/sec Read	MB/sec Write
4kB	6.845013	3.221325
16kB	27.43865	13.15497
64kB	103.5336	51.65482
256kB	424.7271	193.5268
1MB	1576.252	782.4383
4MB	2619.374	3095.016



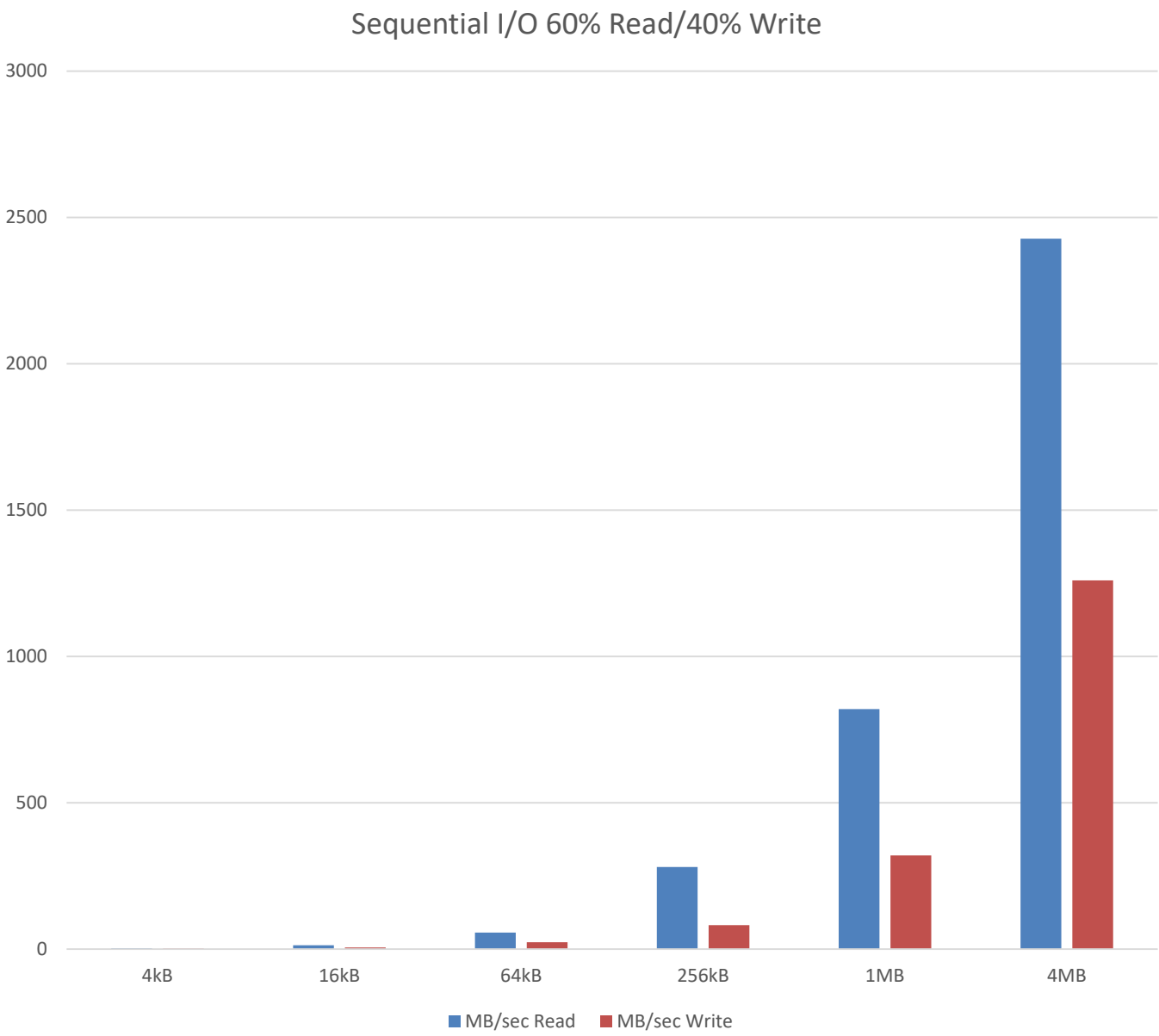
Blocksize	MB/sec Read	MB/sec Write
4kB	3.806028	3.780054
16kB	16.70582	13.7856
64kB	78.28159	57.00872
256kB	449.4757	203.128
1MB	1790.493	791.1742
4MB	6158.911	3112.754



Blocksize	MB/sec Read	MB/sec Write
4kB	3.873204	1.331676
16kB	14.99433	4.84942
64kB	60.45827	20.30537
256kB	233.5852	81.27407
1MB	771.8829	327.8226
4MB	1773.824	1301.71



Blocksize	MB/sec Read	MB/sec Write
4kB	3.082823	1.739016
16kB	12.86438	5.733841
64kB	56.4047	23.77207
256kB	280.6739	81.68536
1MB	820.0511	320.4626
4MB	2427.747	1260.187





## Current Status

- The Good
  - Power users and their projects migrated
    - No more “file not found” errors
    - Computation > 100% faster
  - Backups running and performing nicely
  - Everybody is fairly happy
  - Very resilient to many types of failure
- The Bad
  - Snapshots