Sonexion GridRAID Characteristics

CUG 2014

Mark Swan, Cray Inc.

Safe Harbor Statement

This presentation may contain forward-looking statements that are based on our current expectations. Forward looking statements may include statements about our financial guidance and expected operating results, our opportunities and future potential, our product development and new product introduction plans, our ability to expand and penetrate our addressable markets and other statements that are not historical facts. These statements are only predictions and actual results may materially vary from those projected. Please refer to Cray's documents filed with the SEC from time to time concerning factors that could affect the Company and these forward-looking statements.

Sonexion GridRAID Characteristics

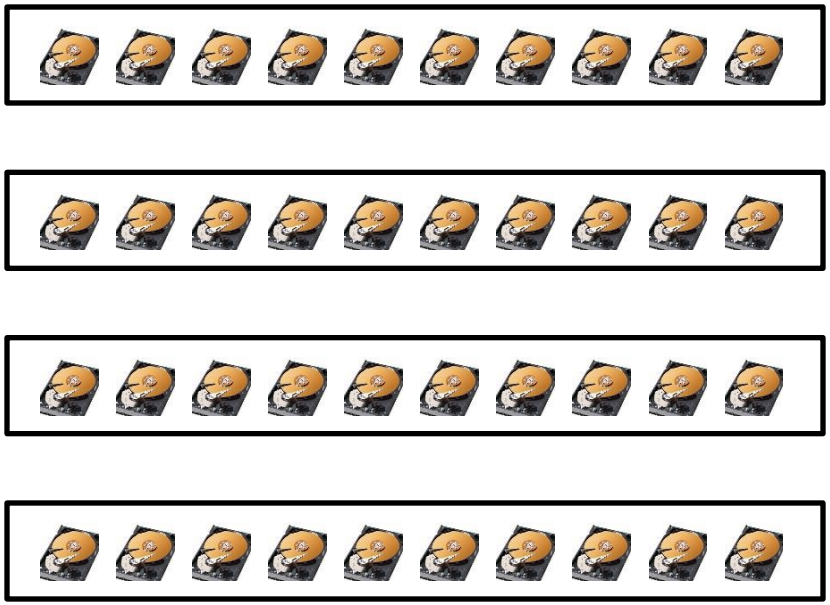
Architecture of the OST

Performance

Degraded Modes

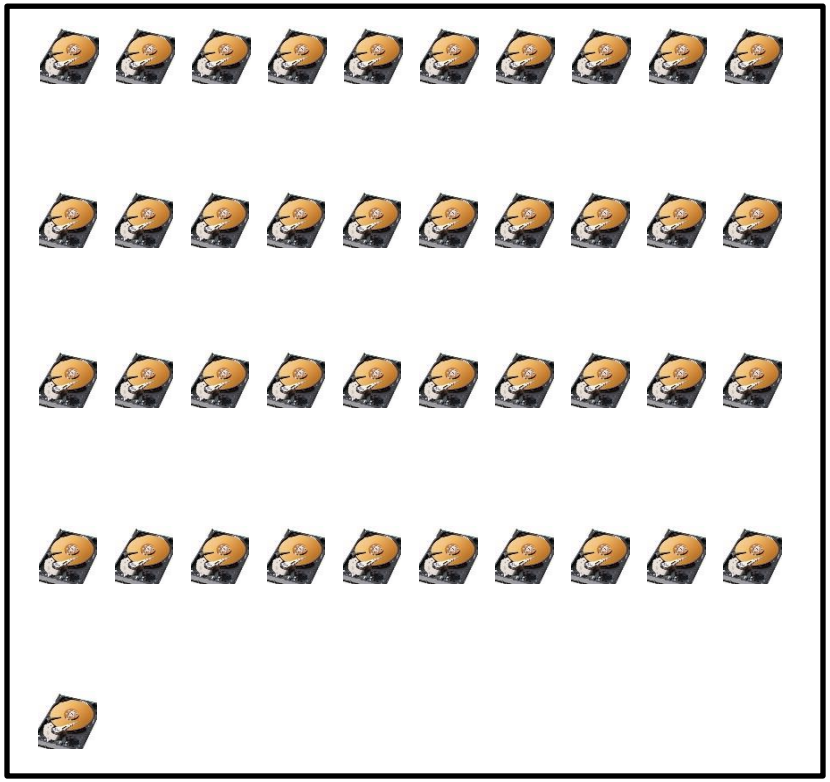
Architecture of the OST

MDRAID
4 OSTs per OSS
RAID 6 – 10 drive (8+2)



Global Hot Spares

GridRAID
1 OST per OSS
RAID 6 – 41 drive (8+2+2)



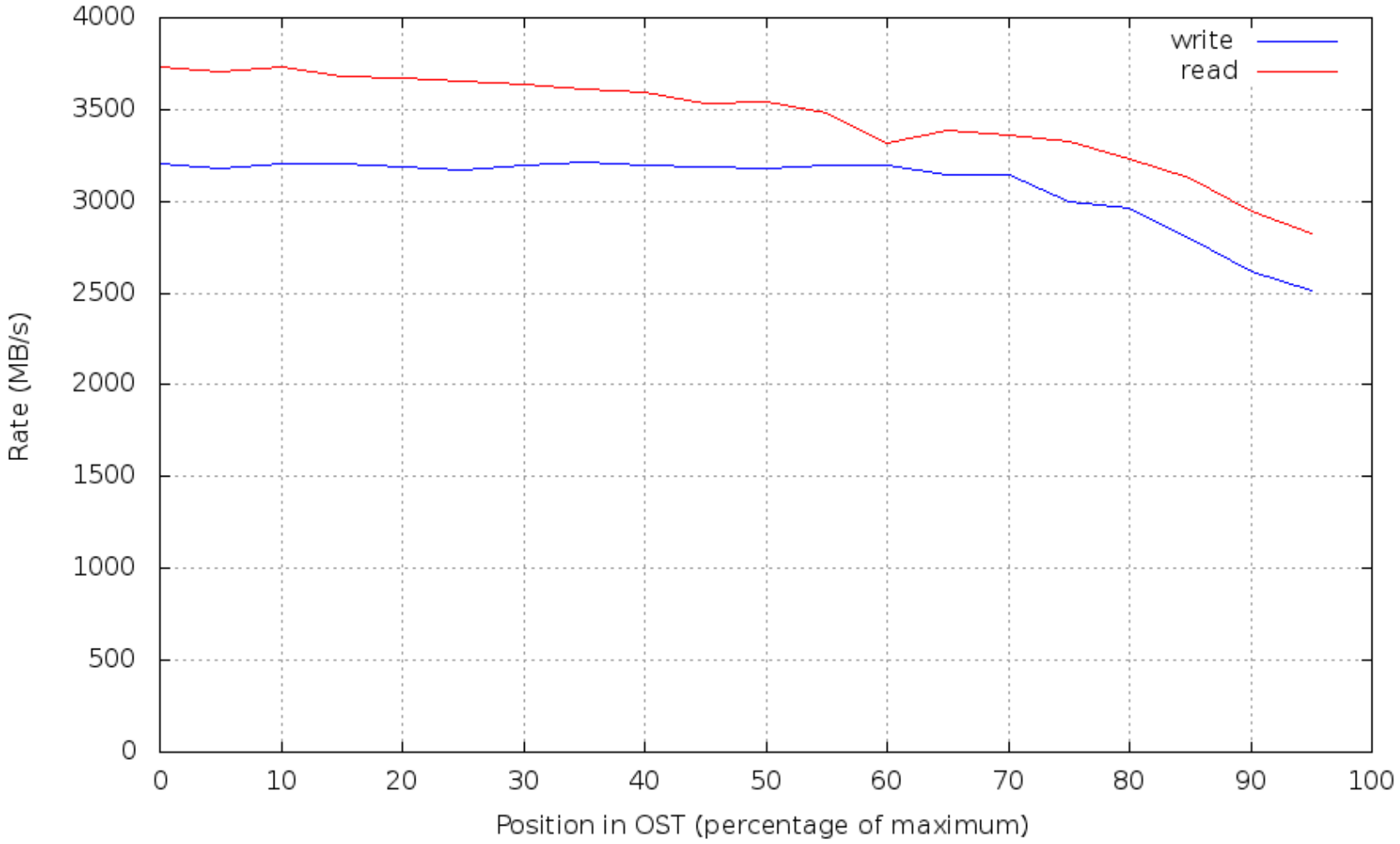
Performance

Where data is on the OST

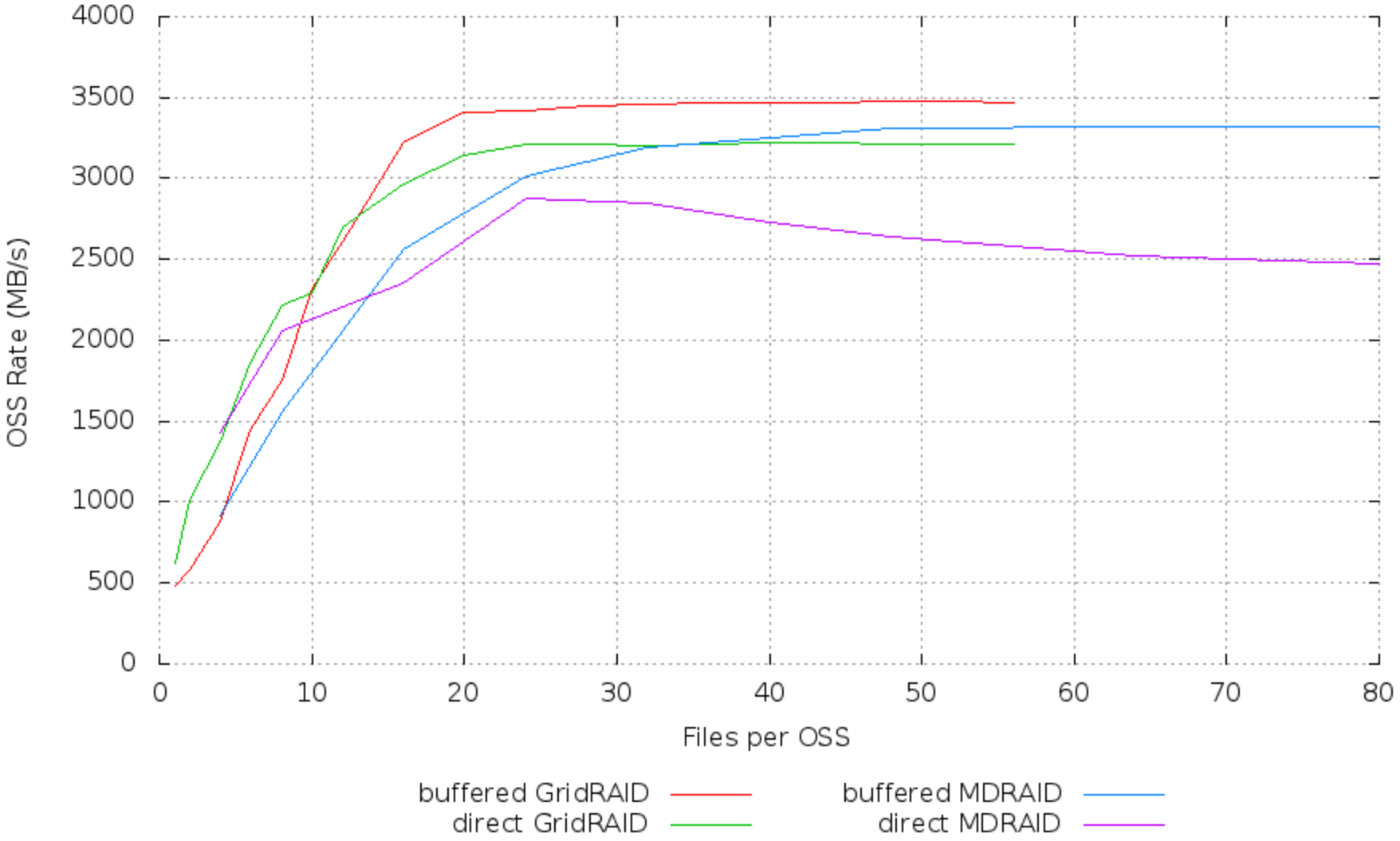
How data gets to the OST

How data is arranged on the OST

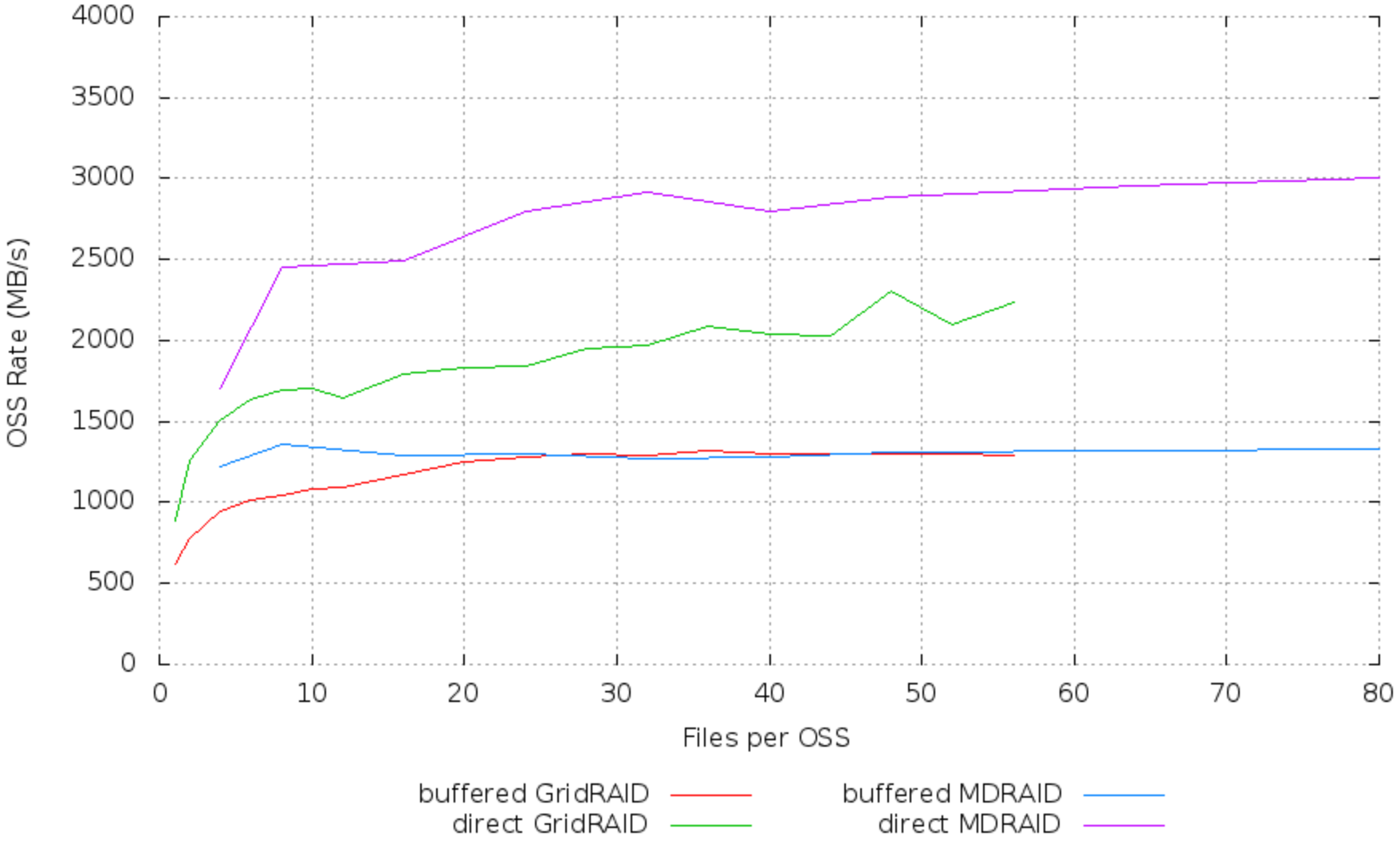
Edge to edge performance curve obdfilter-survey results single GridRAID OST, 3 TB Hitachi drives



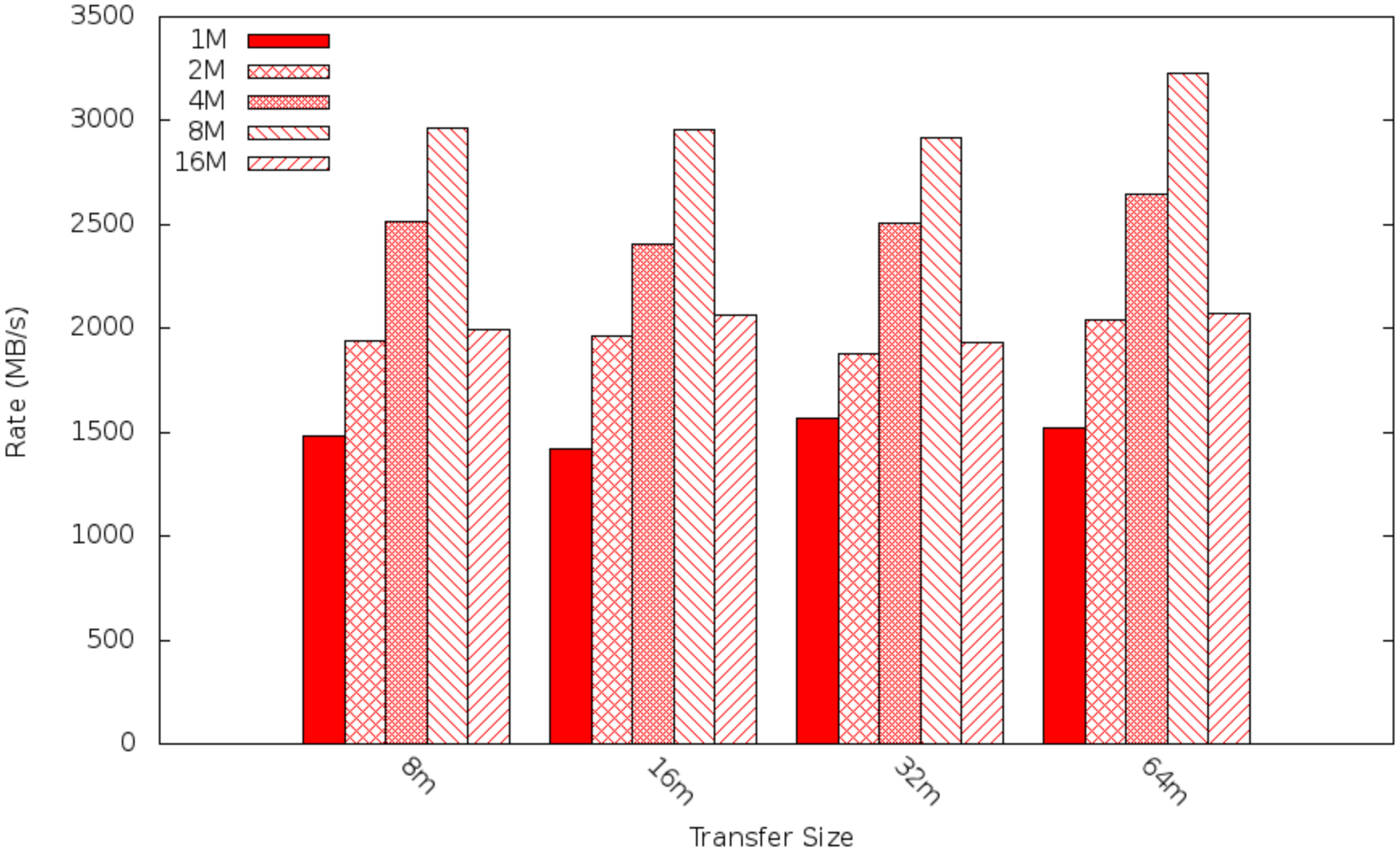
Comparing MDRAID/GridRAID write rates 32 MB transfers



Comparing MDRAID/GridRAID read rates 32 MB transfers



Effects of OST preallocation sizes single GridRAID OST, IOR buffered read 16 files per OST

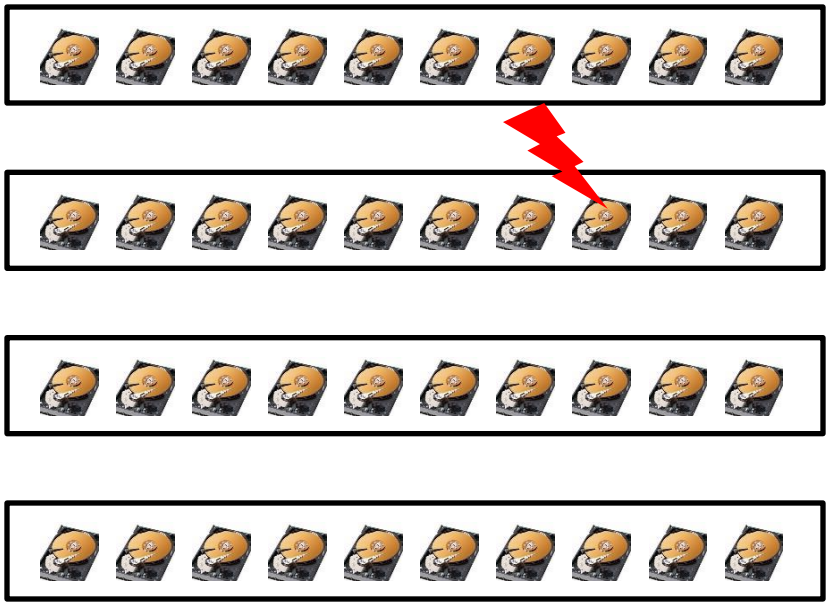


Degraded modes

MDRAID	GridRAID
Repair	Reconstruct
	Rebalance

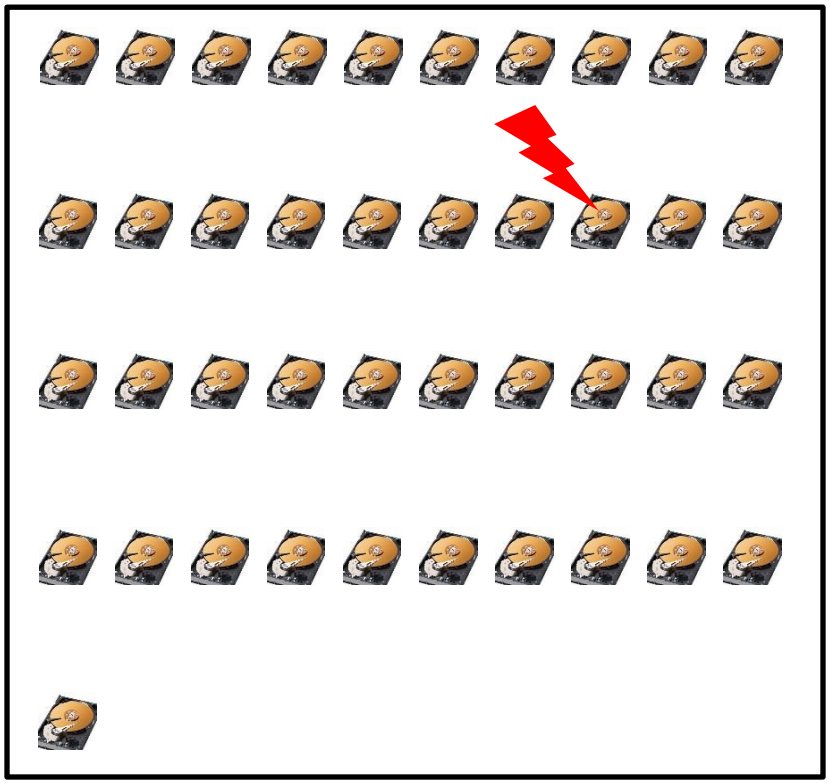
Degraded Mode – losing a drive

MDRAID – repair
4 OSTs per OSS
RAID 6 – 10 drive (8+2)



Global Hot Spares

GridRAID – reconstruct
1 OST per OSS
RAID 6 – 41 drive (8+2+2)



Degraded Mode - reconstruct

MDRAID Repair Time

TIME = (size of drive) / (minimum drive bandwidth)

Example: 4 TB / 50 MB/s = ~22 hours

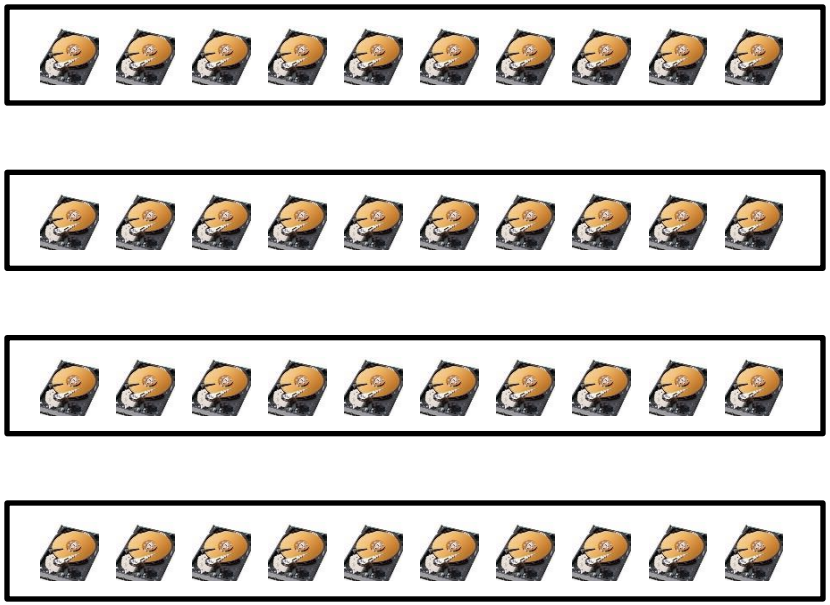
GridRAID Reconstruct Time

(size of drive) / (40 * (minimum drive bandwidth) / 9)

Example: 4 TB / (40 * 50 MB/s / 9) = ~5 hours

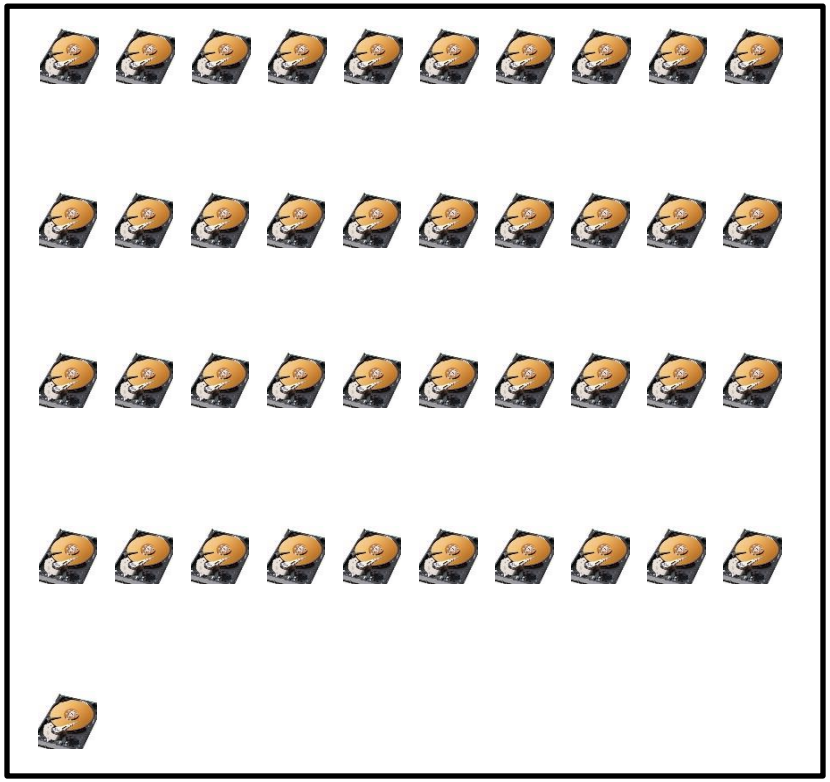
Degraded Mode – replacing a drive

MDRAID – new hot spare
4 OSTs per OSS
RAID 6 – 10 drive (8+2)



Global Hot Spares

GridRAID – rebalance
1 OST per OSS
RAID 6 – 41 drive (8+2+2)



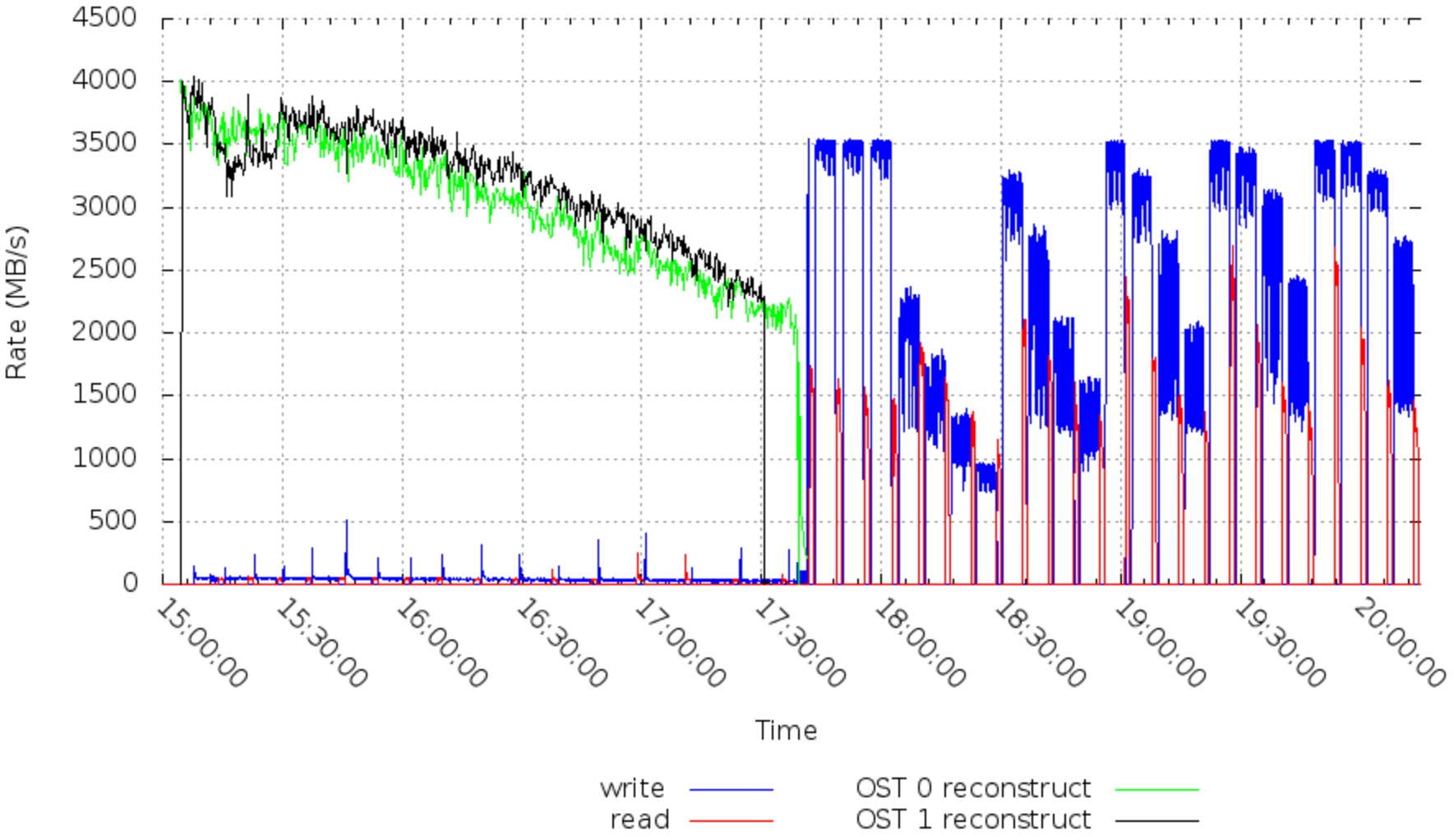
Degraded Mode - rebalance

GridRAID Rebalance Time

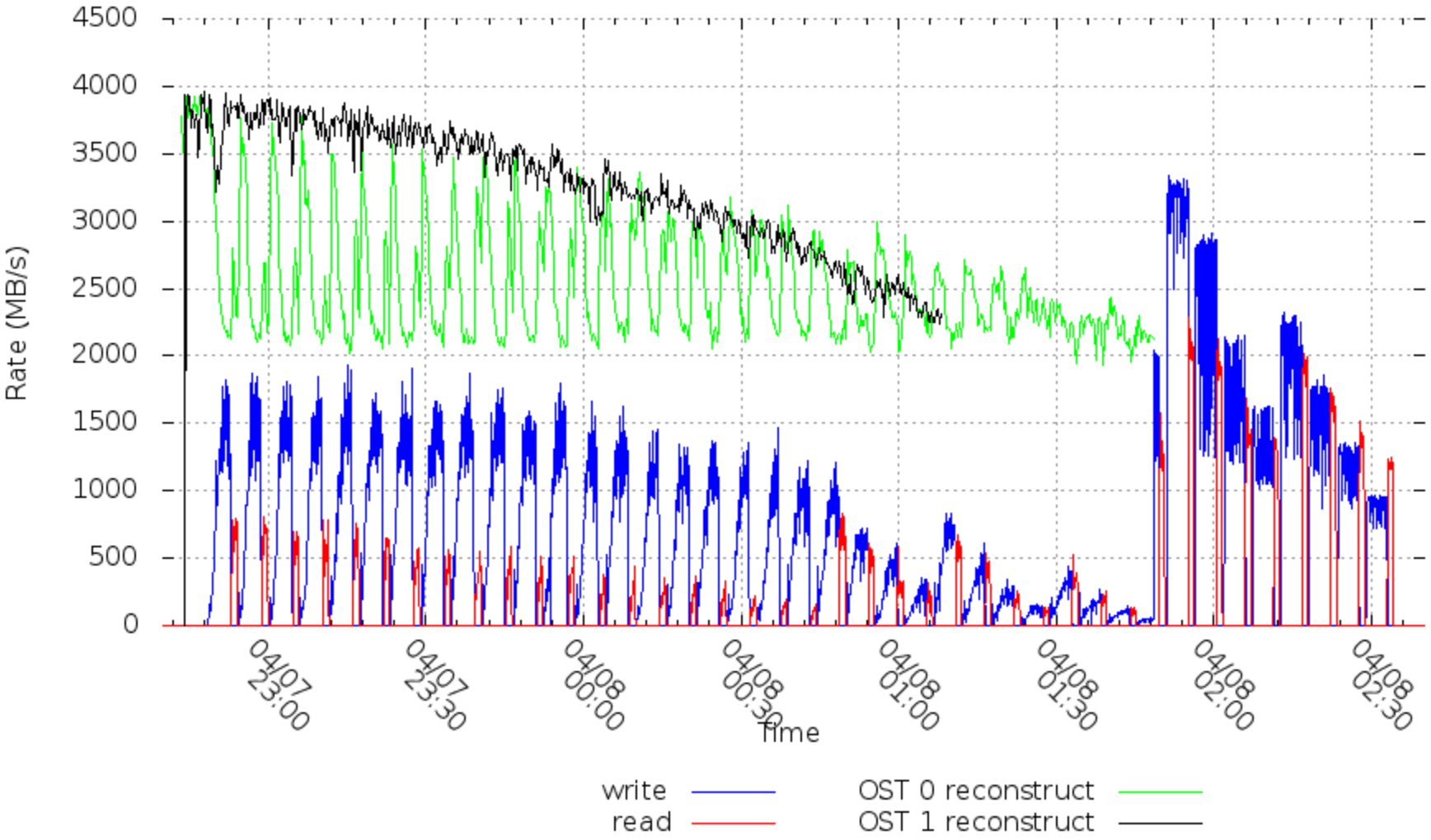
(size of drive) / (minimum drive bandwidth)

Example: 4 TB / 50 MB/s = ~22 hours

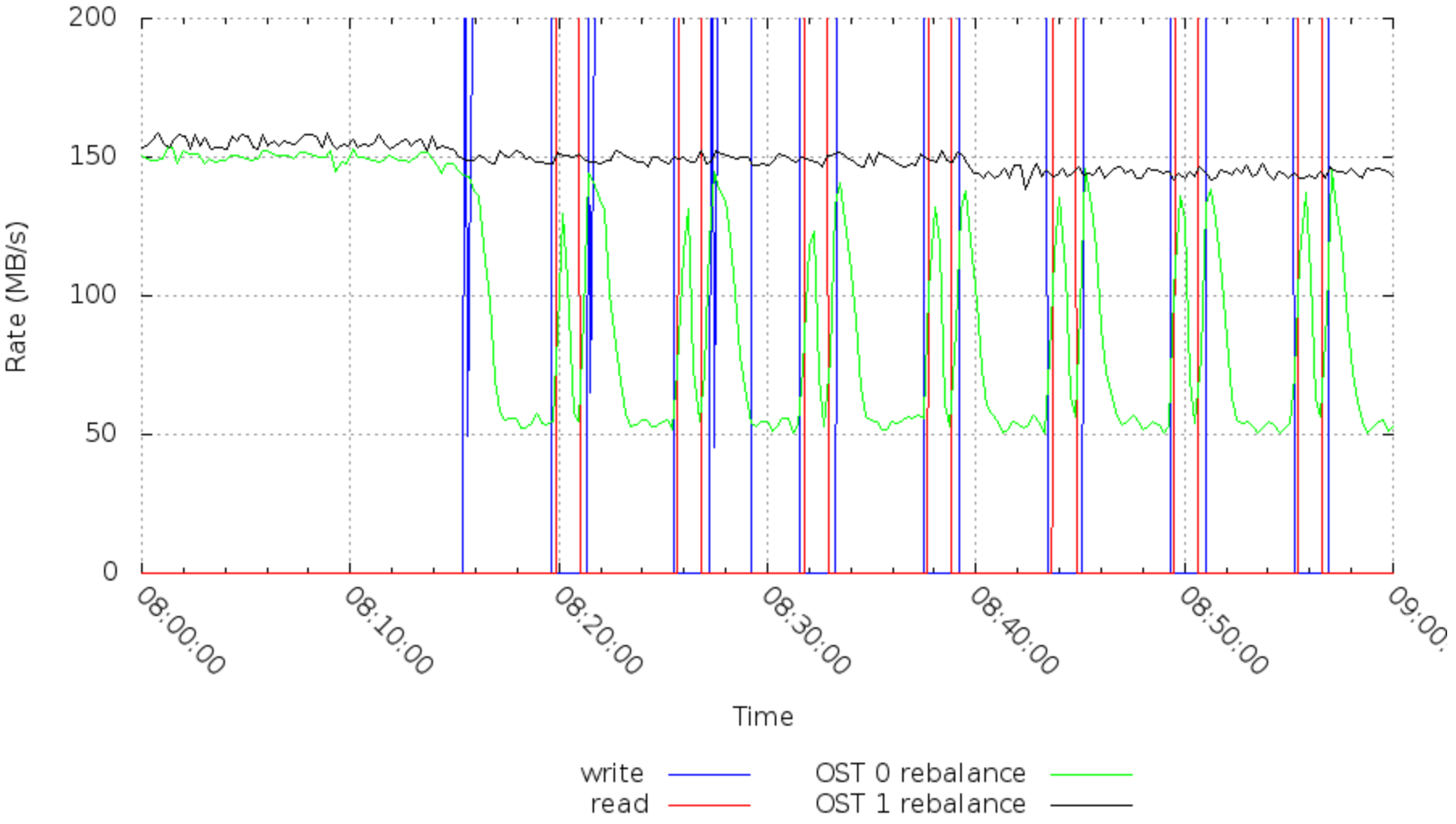
Reconstruct of OST 0 and OST 1
Minimum speed=100000 KB/sec/disk
IOR mini survey to OST 0 only



Reconstruct of OST 0 and OST 1
Minimum speed=50000 KB/sec/disk
IOR mini survey to OST 0 only



Rebalance of OST 0 and OST 1
Minimum speed=50000 KB/sec/disk
IOR mini survey to OST 0 only



Summary

High end performance on par with MDRAID

Performance ramps up faster than MDRAID

One-fourth as many OSTs to stripe data across

One-fourth less time recovering from single disk failure

Legal Disclaimer

Information in this document is provided in connection with Cray Inc. products. No license, express or implied, to any intellectual property rights is granted by this document.

Cray Inc. may make changes to specifications and product descriptions at any time, without notice.

All products, dates and figures specified are preliminary based on current expectations, and are subject to change without notice.

Cray hardware and software products may contain design defects or errors known as errata, which may cause the product to deviate from published specifications. Current characterized errata are available on request.

Cray uses codenames internally to identify products that are in development and not yet publically announced for release. Customers and other third parties are not authorized by Cray Inc. to use codenames in advertising, promotion or marketing and any use of Cray Inc. internal codenames is at the sole risk of the user.

Performance tests and ratings are measured using specific systems and/or components and reflect the approximate performance of Cray Inc. products as measured by those tests. Any difference in system hardware or software design or configuration may affect actual performance.

The following are trademarks of Cray Inc. and are registered in the United States and other countries: CRAY and design, SONEXION, URIKA, and YARCDATA. The following are trademarks of Cray Inc.: ACE, APPRENTICE2, CHAPEL, CLUSTER CONNECT, CRAYPAT, CRAYPORT, ECOPHLEX, LIBSCI, NODEKARE, THREADSTORM. The following system family marks, and associated model number marks, are trademarks of Cray Inc.: CS, CX, XC, XE, XK, XMT, and XT. The registered trademark LINUX is used pursuant to a sublicense from LMI, the exclusive licensee of Linus Torvalds, owner of the mark on a worldwide basis. Other trademarks used in this document are the property of their respective owners.