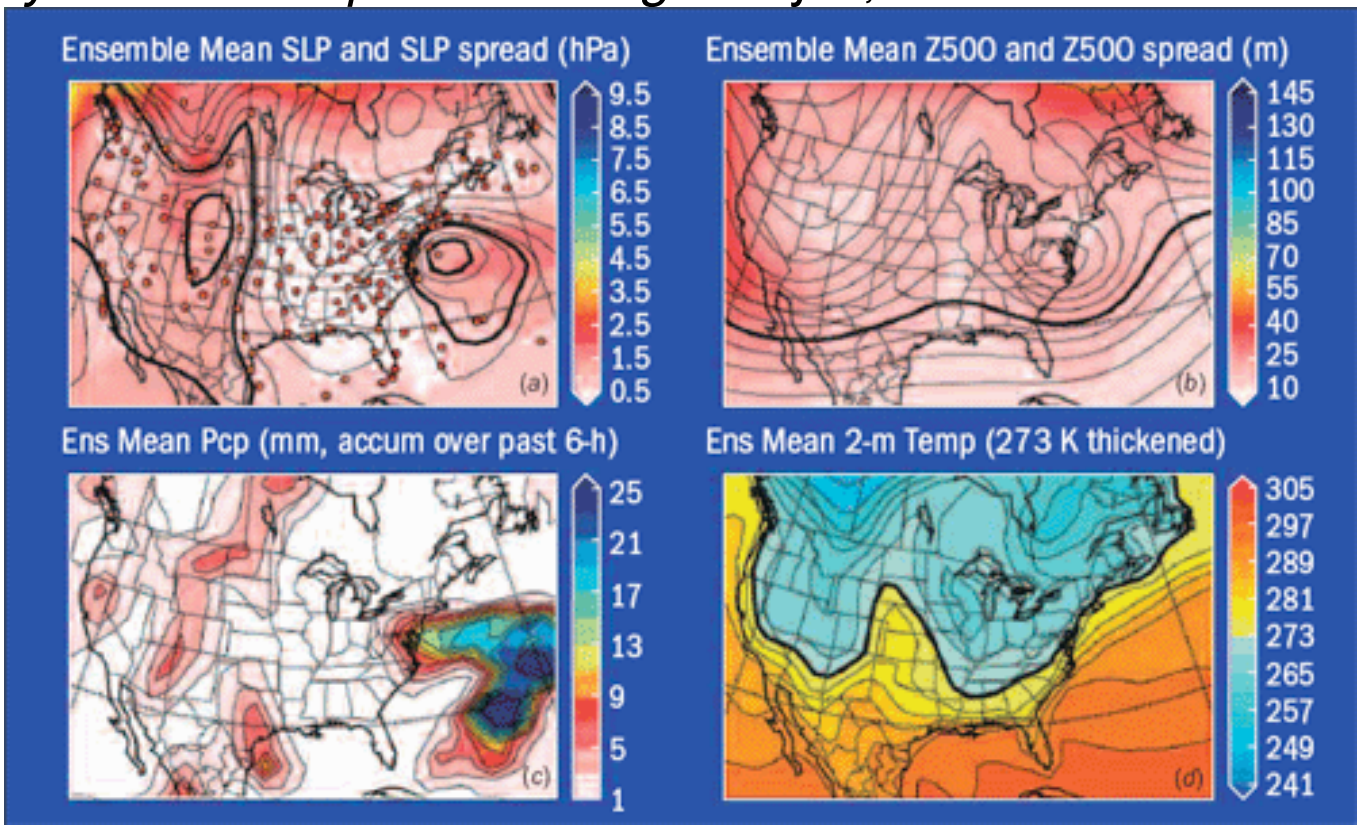
Deploying Server-side File System Monitoring at NERSC

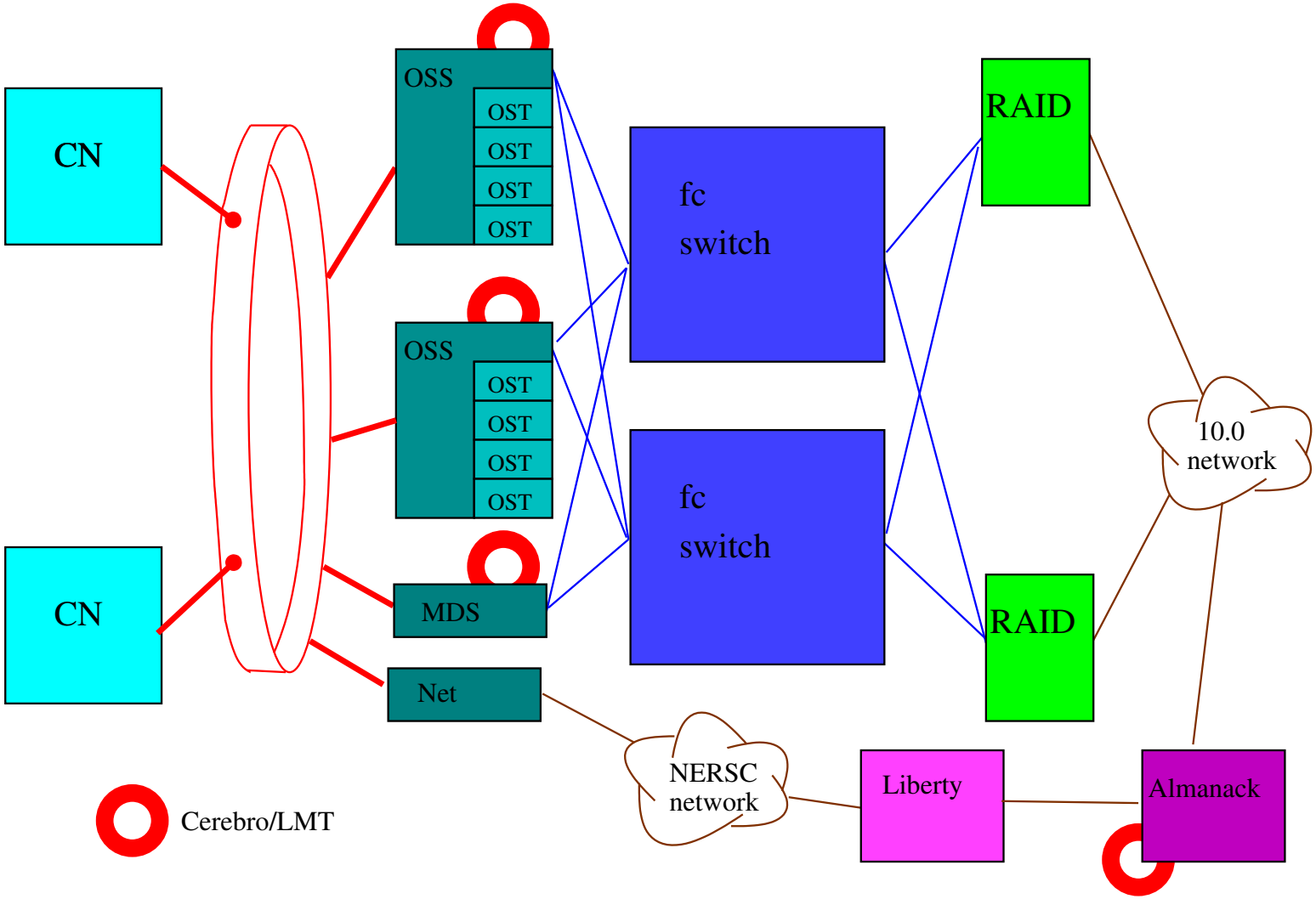
Cray Users Group Proceedings May 7, 2009



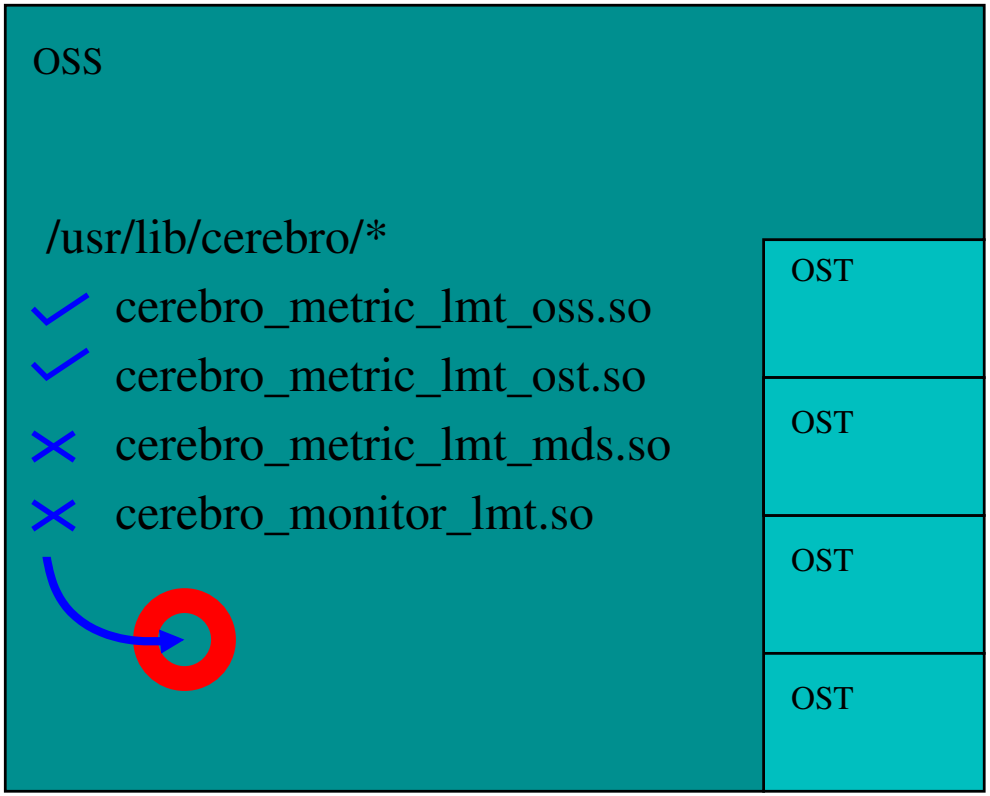
Andrew Uselton
National Energy Research Scientific Computing Center
Lawrence Berkeley National Lab

Contents

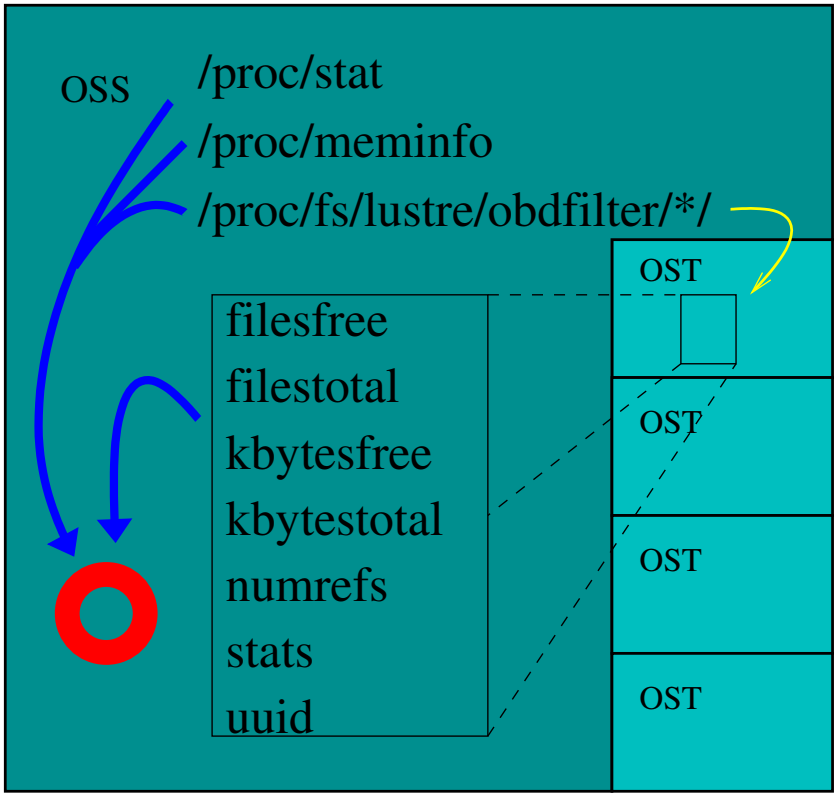
Monitoring the I/O Subsystem



Cerebro



LMT



An OSS Tuple

Cerebro Protocol Version
Host Name
CPU Utilization
Memory Utilization

1.0;nid04187;4.990020;39.303989

OST Data Values

Cerebro Protocol Version
Host Name
UUID
Bytes Read
Bytes Written
Kbytes Free
Kbytes Used
Inodes Free
Inodes Used

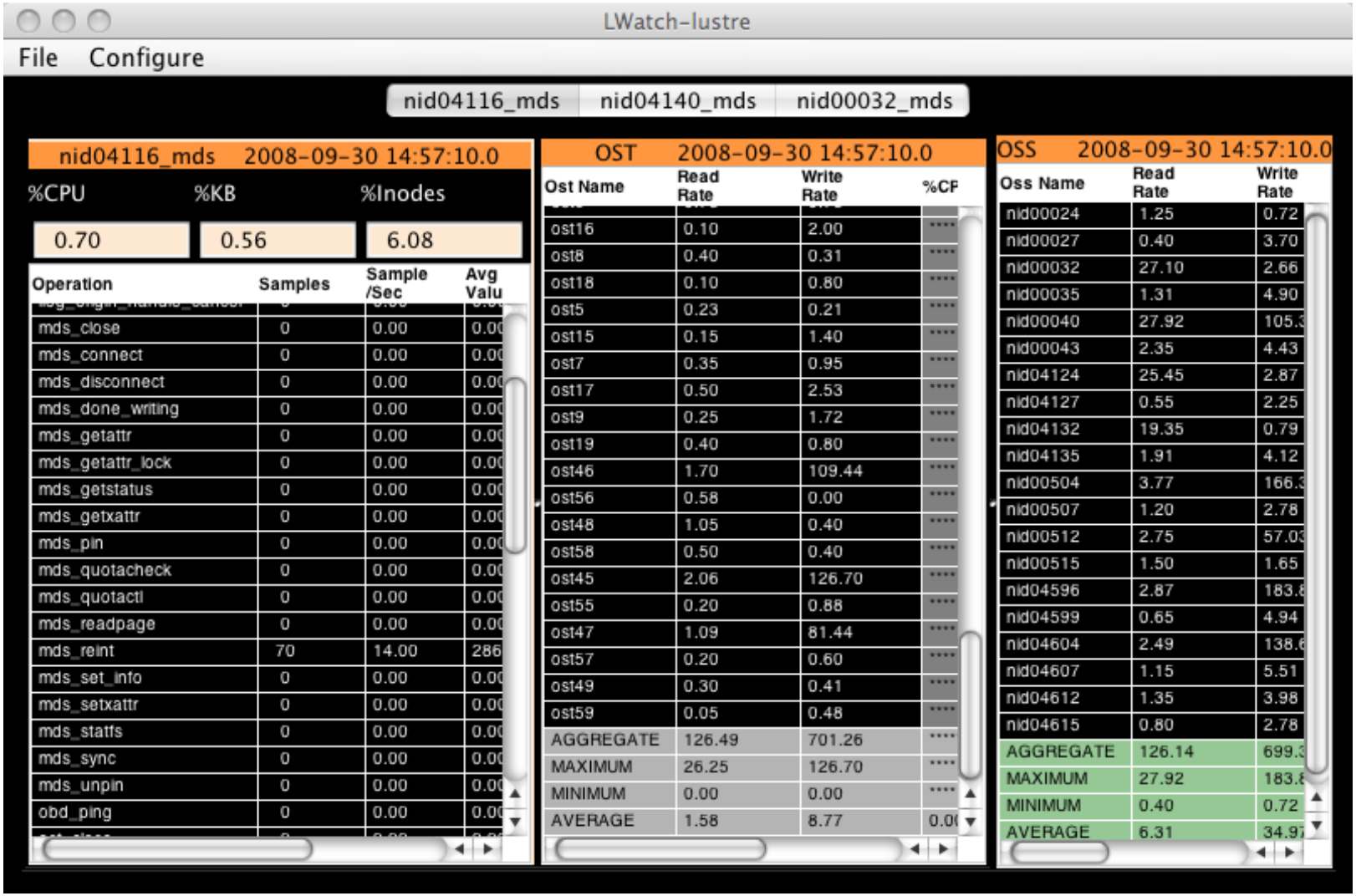
MDS Operations

mysql> select * from OPERATION_INFO;

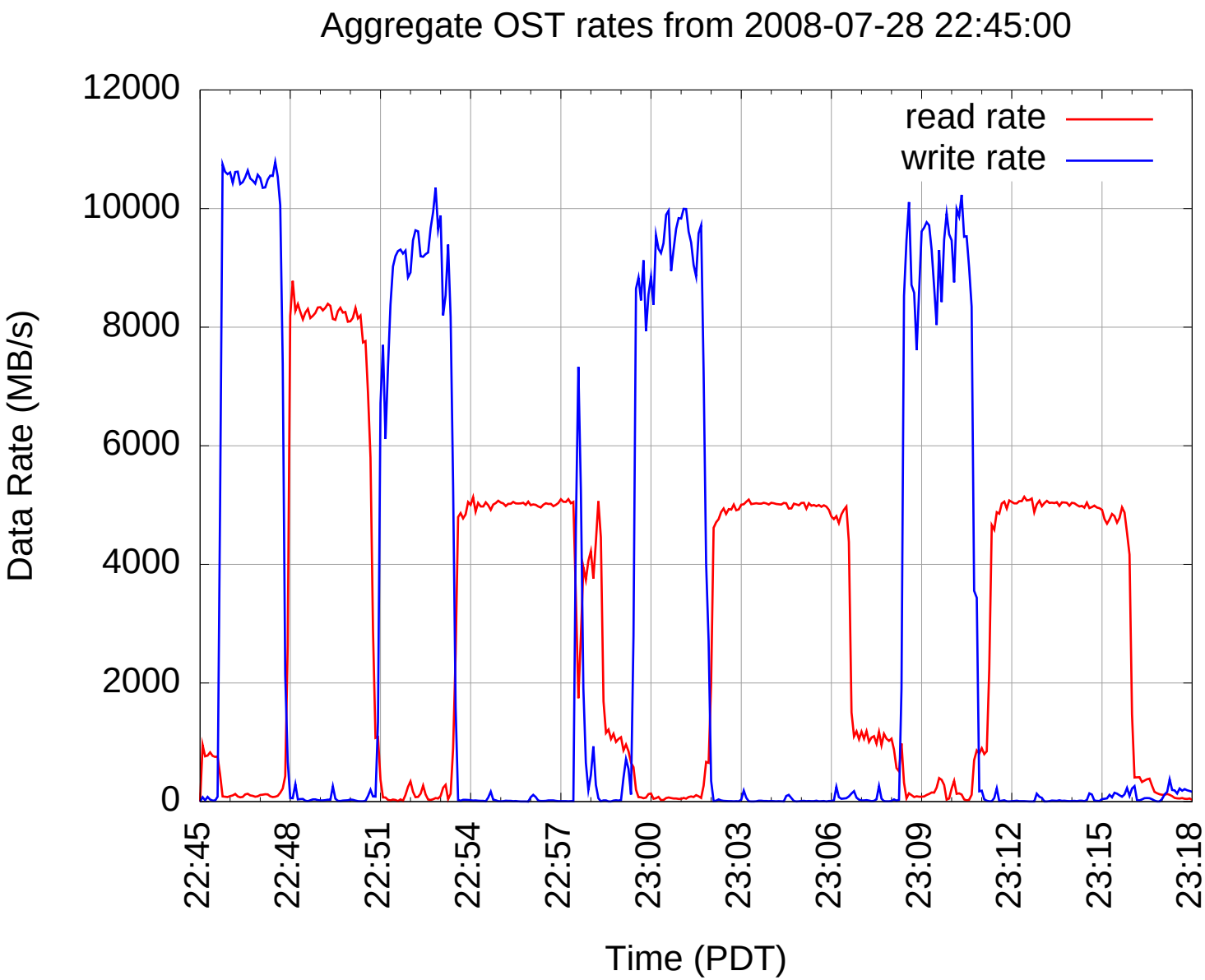
OPERATION_NAME	UNITS	OPERATION_NAME	UNITS
req_waittime	usec	mds_getattr_lock	usec
req_qdepth	reqs	mds_close	usec
req_active	reqs	mds_reint	usec
reqbuf_avail	bufs	mds_readpage	usec
ost_reply	usec	mds_connect	usec
ost_getattr	usec	mds_disconnect	usec
ost_setattr	usec	mds_getstatus	usec
ost_read	bytes	mds_statfs	usec
ost_write	bytes	mds_pin	usec
ost_create	usec	mds_unpin	usec
ost_destroy	usec	mds_sync	usec
ost_get_info	usec	mds_done_writing	usec
ost_connect	usec	mds_set_info	usec
ost_disconnect	usec	mds_quotacheck	usec
ost_punch	usec	mds_quotactl	usec
ost_open	usec	mds_getxattr	usec
ost_close	usec	mds_setxattr	usec
ost_statfs	usec	ldlm_enqueue	usec

...

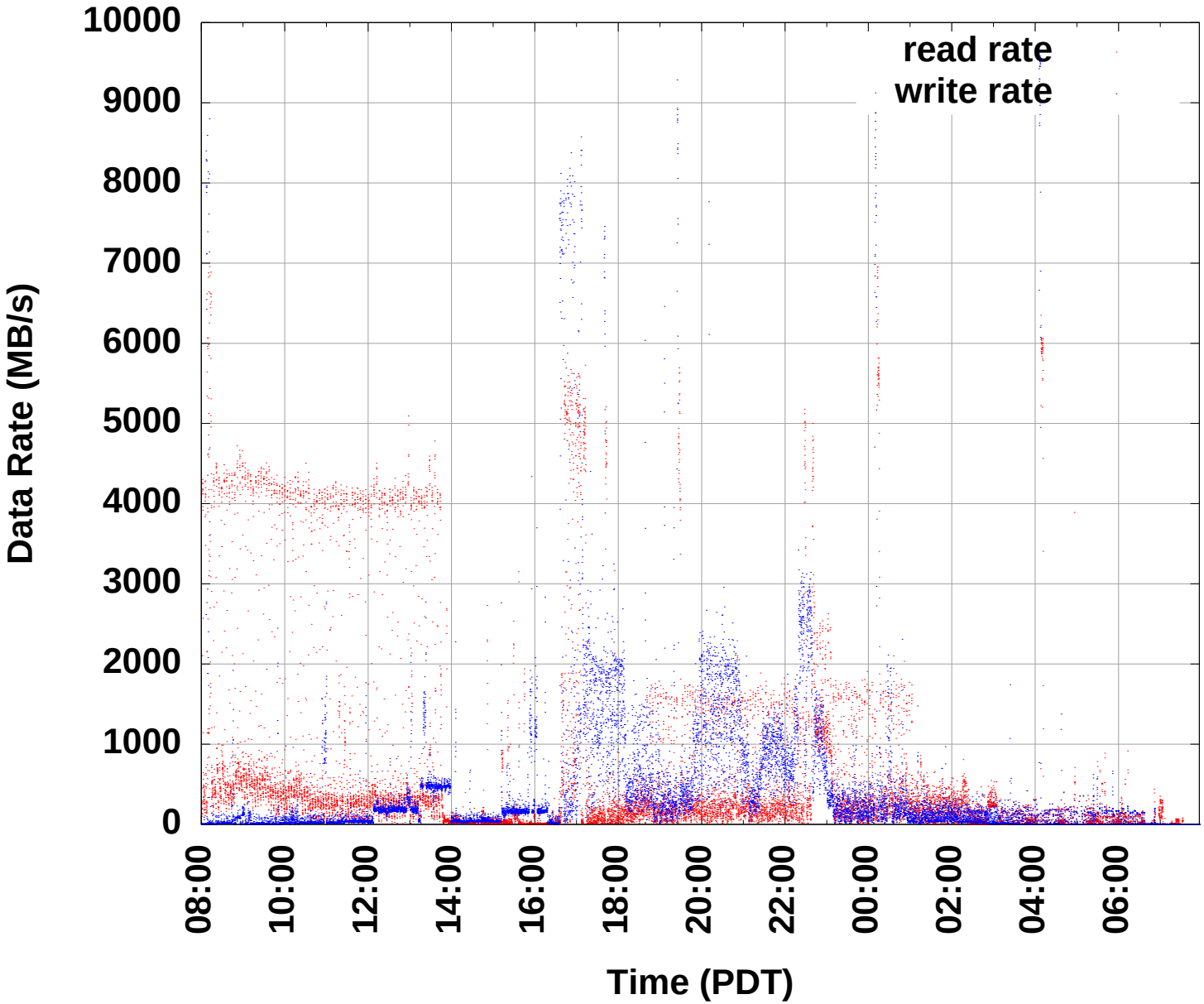
The Lustre Dashboard



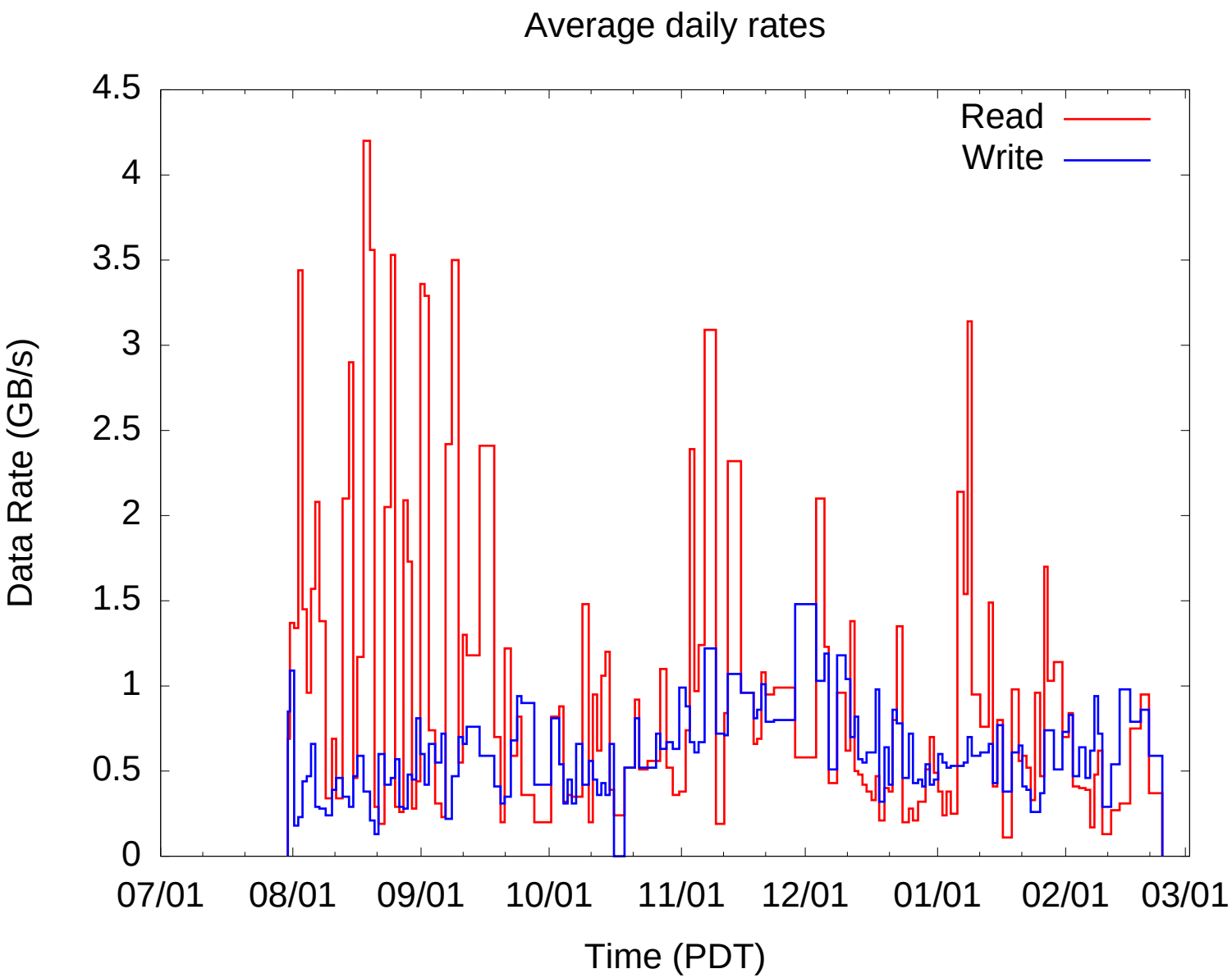
Four IOR Tests



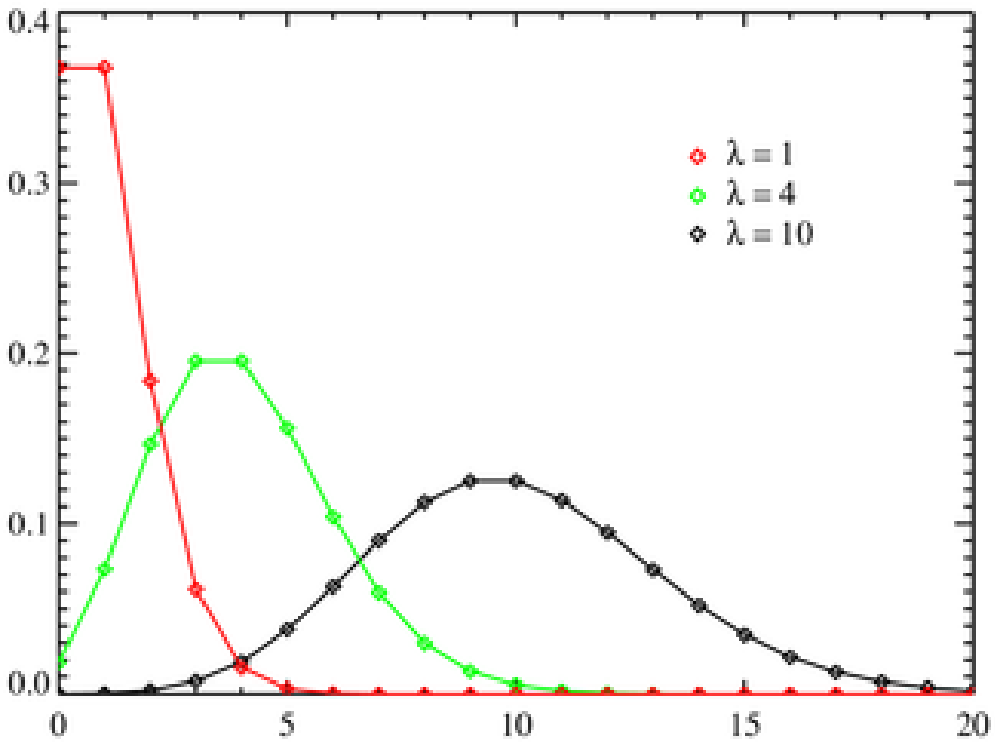
24 Hours of LMT Data



Daily Averages

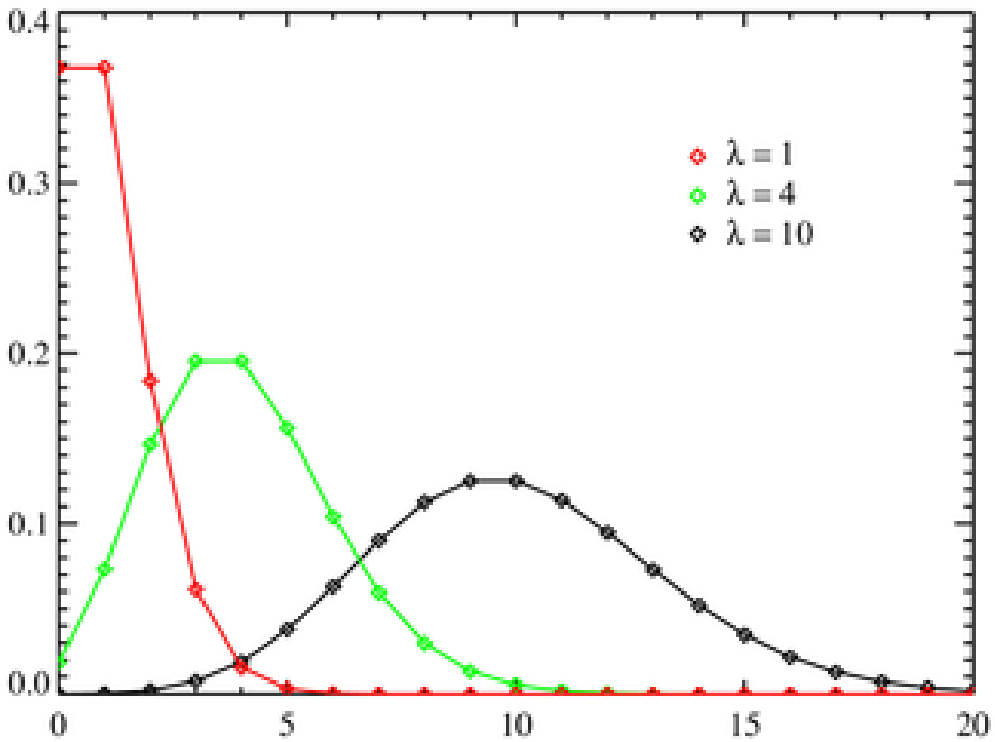


http://en.wikipedia.org/wiki/Poisson_distribution:



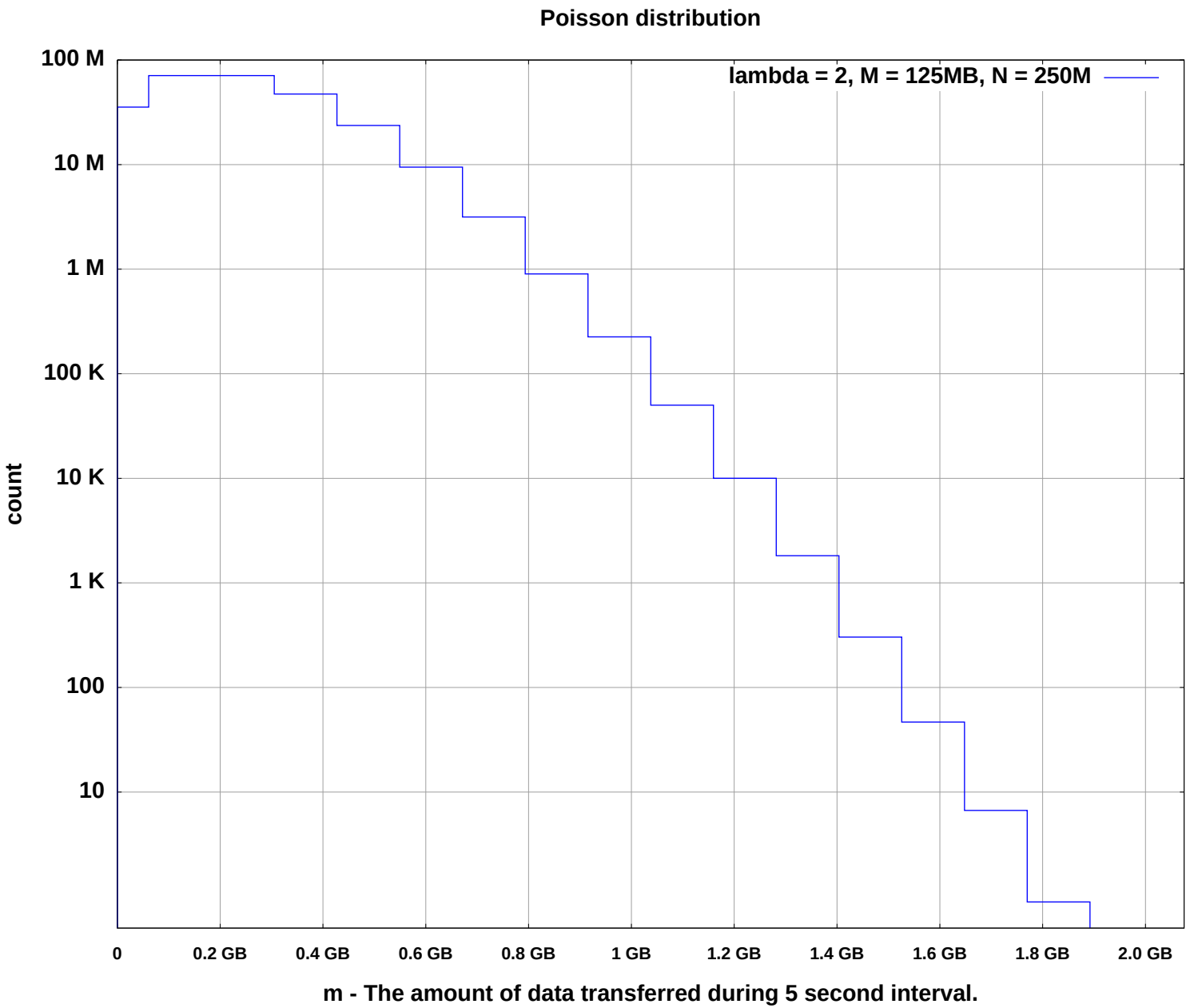
- $$f_\lambda(k) = \frac{\lambda^k e^{-\lambda}}{k!}$$

http://en.wikipedia.org/wiki/Poisson_distribution:

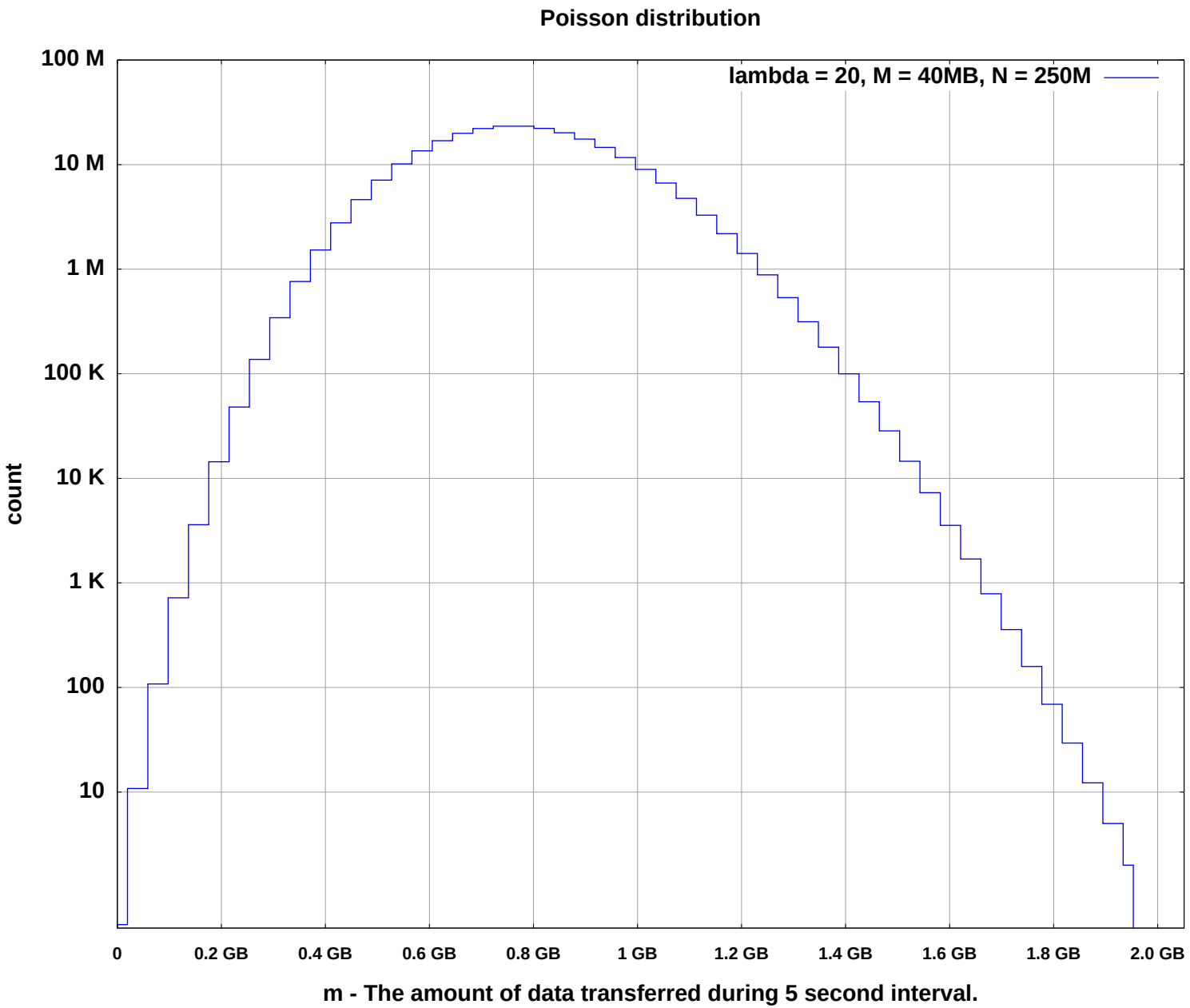


- $C(m) = N \times f_\lambda(\text{int}(m/M))$

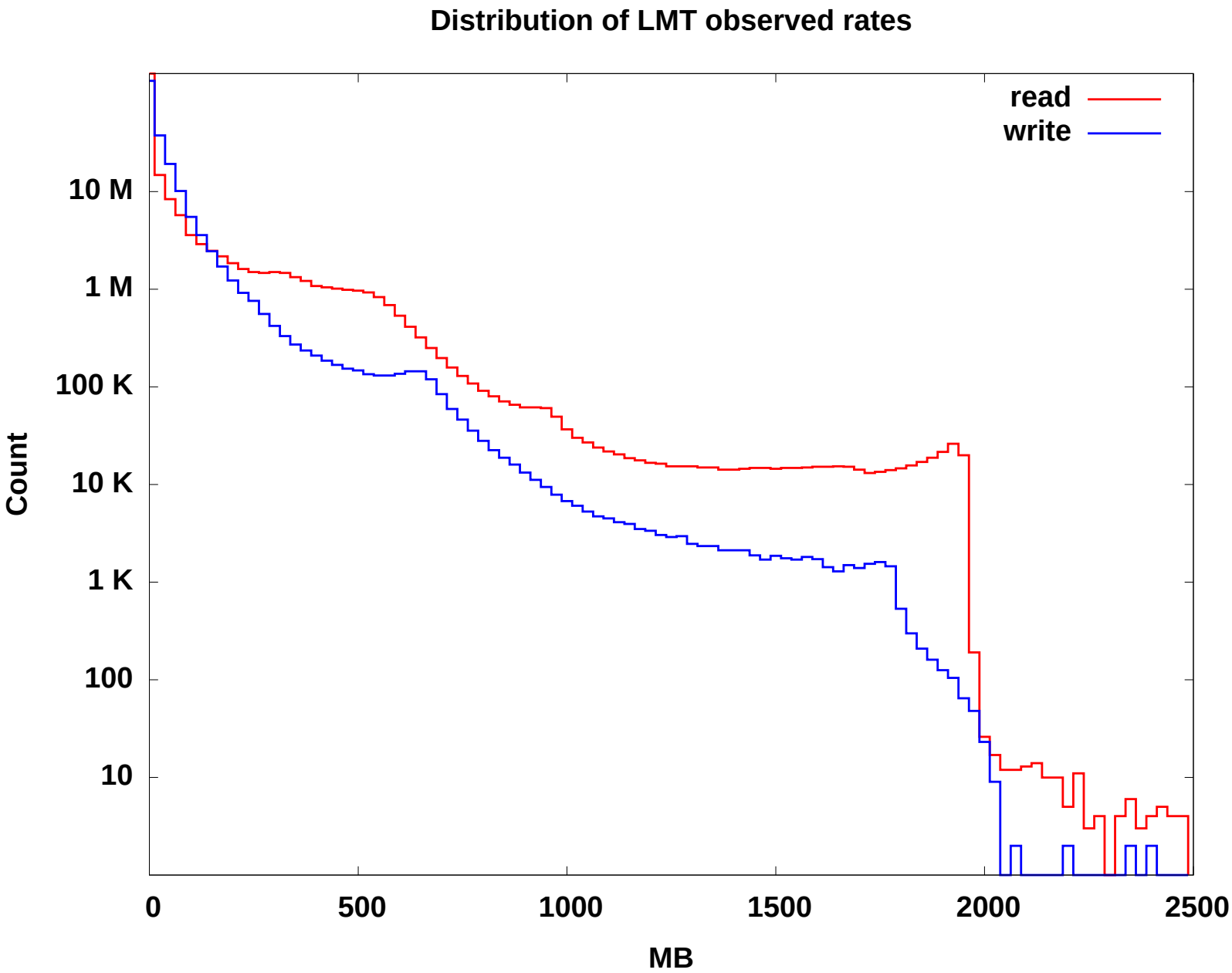
Poisson Distribution: $\lambda = 2$



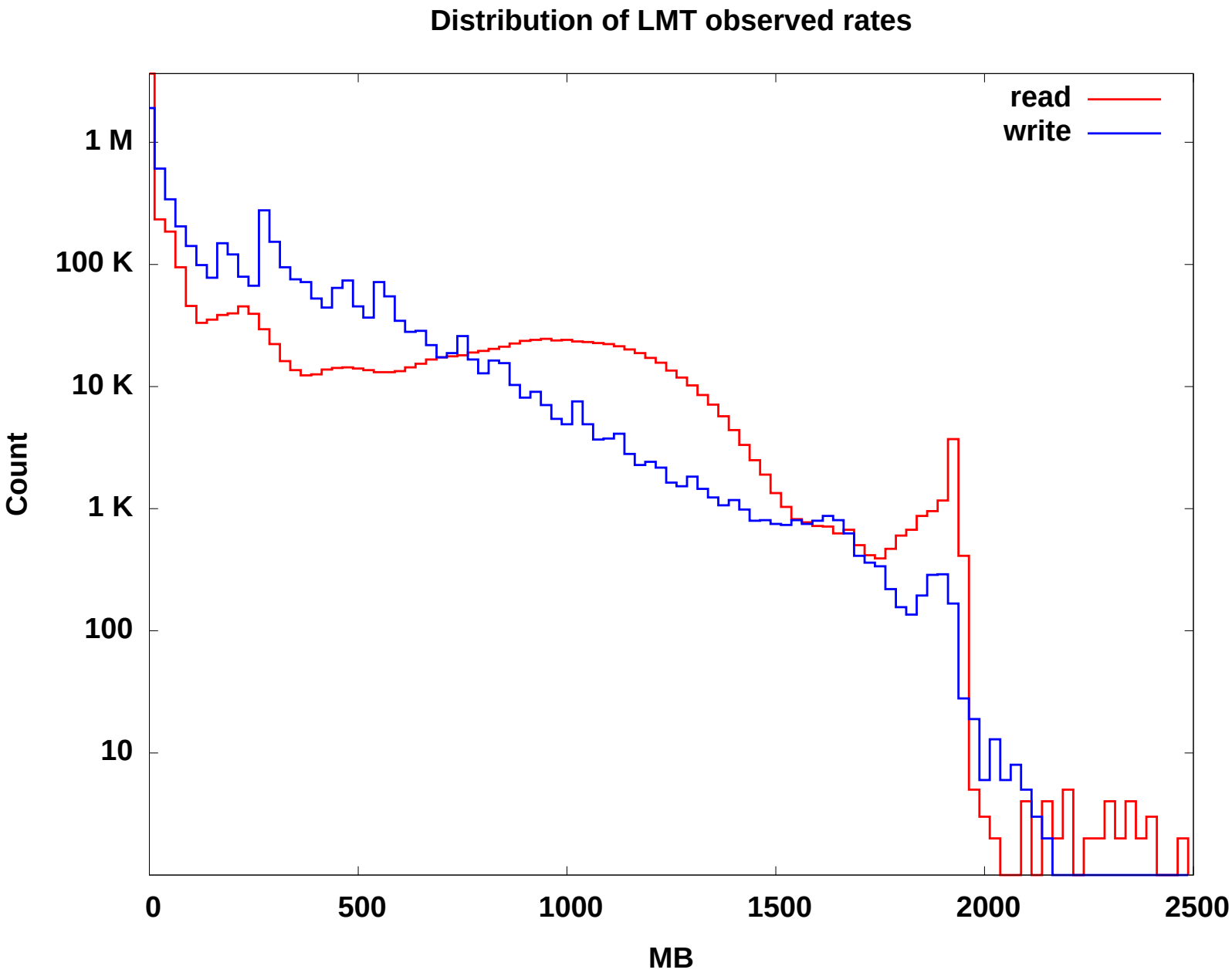
Poisson Distribution: $\lambda = 20$



250 M LMT Observations



Two weeks of recent observations



I would like to acknowledge and thank:

Al Chu The author of Cerebro.

Herb Wartens The author of the Lustre Monitorining Tool plug-ins.

Both work at Lawrence Livermore National Lab, which supported the development of these tools. Both were very generous with their time as I deployed the software on Franklin.

The software is available from:

Both applications are open source and available from Sourceforge.

Cerebro `http://sourceforge.net/projects/cerebro`

LMT `http://sourceforge.net/projects/lmt/`

If you would like hints and encouragement with getting this software deployed, contact me: Andrew Uselton (acuselton@lbl.gov)

If you get results from your deployment that you would like to share, please do so.