Per User Lustre File Systems Leveraging ZFS Pools

Lustre User Group 2015

April 2015 – Denver, CO

Marc Stearman
Parallel File Systems Operations Lead

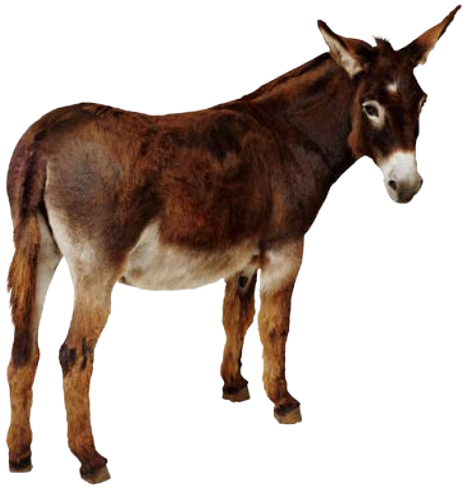
Lawrence Livermore National Laboratory

LLNL-PRES-669094

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344. Lawrence Livermore National Security, LLC

Users Want It ALL!

Users don’t want to change



“You want me to change my I/O?”

“Someone else wrote that code”

“I downloaded it off the Internet”

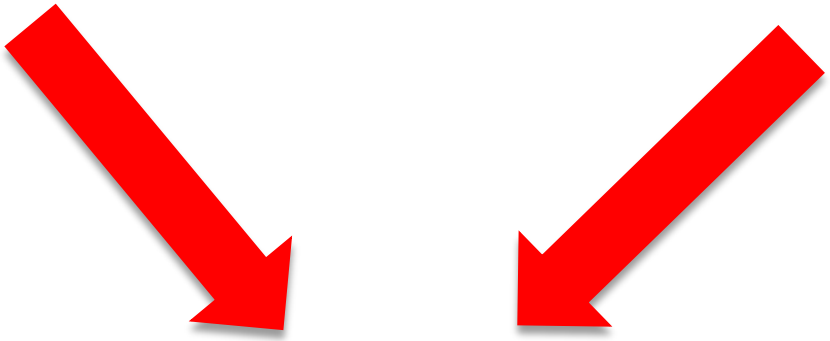
“I’d rather optimize my compute algorithms”

Legacy apps written for decades old file systems

Where does that leave us?

Small I/O

Billions of Files



Overwhelmed MetaData Servers!!

We Need More MetaData Ops!!!

- DNE v1 is an option. Not too different from multiple file systems.
- Users could write to multiple file systems in their apps!
- Oh yeah…they don’t want to change their I/O
- DNE v2 is better, but we have problems today.

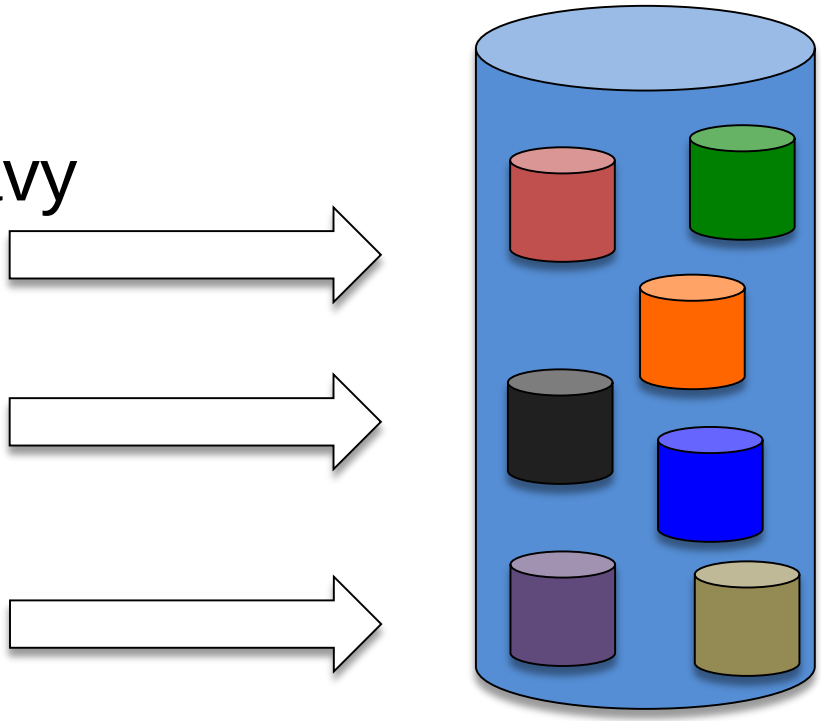
Every app is different!

Some users are write heavy

Some are read heavy

Some small I/O

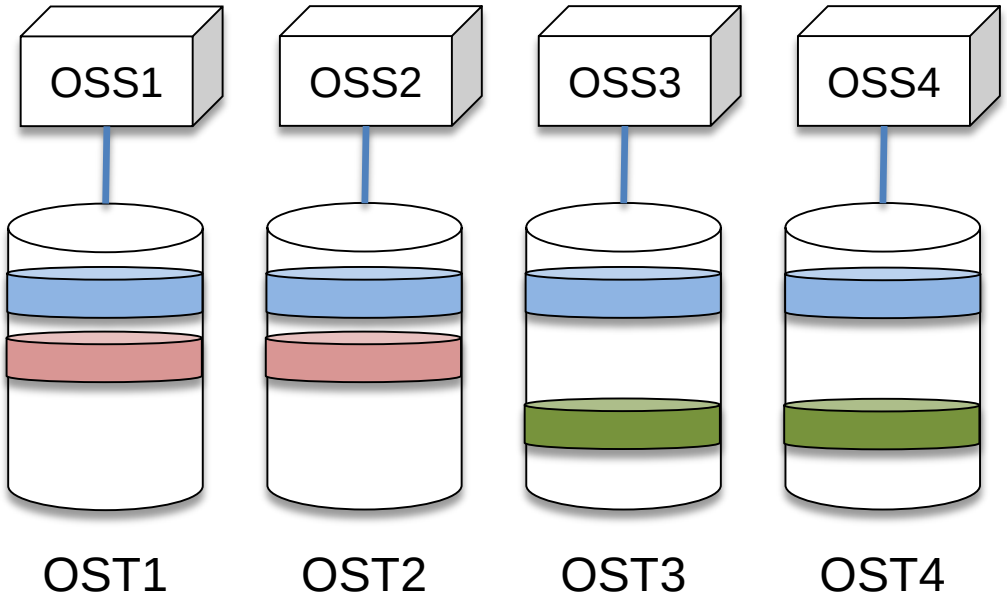
Some large I/O

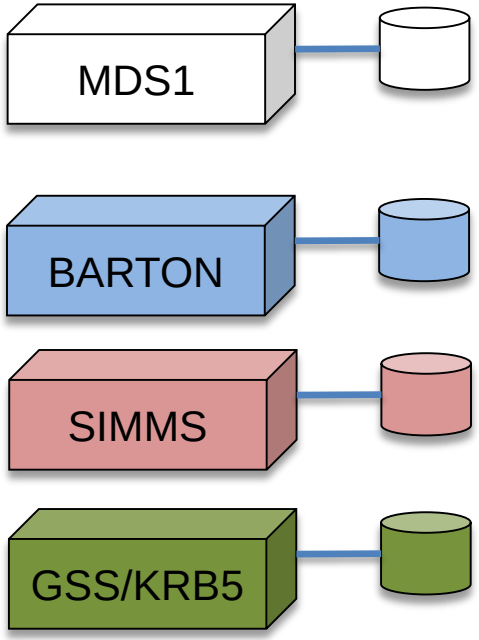


Give a user their own file system

- ZFS Pools to the rescue!
- Create multiple Lustre file systems in the pool
- LLNL has done this before
  - Reformat a file system
  - Migrate the data

Multiple File Systems In A Pool





Forget Lustre Level Quotas

- ZFS allows you to limit the size of data sets in the pool
- User quotas can be managed by the size of the space you give them.
- Lustre Quota may still be needed if you have a general purpose file system

Don’t mount globally!

- Scaled all the way, thousands of file systems can present a problem on a node
- Compute nodes are easy
  - mount the user’s file system at job start
  - unmount when job completes
- How to handle login or data transfer nodes?
  - VMs?
  - Automounter?

Wait! There's more!

- Why just Lustre?
- Create NFS shares in the pool
- Let Users run on the OSS, local to the pool (Hadoop?)

Questions?

Ideas?

Marc Stearman

stearman2@llnl.gov

925-423-9670