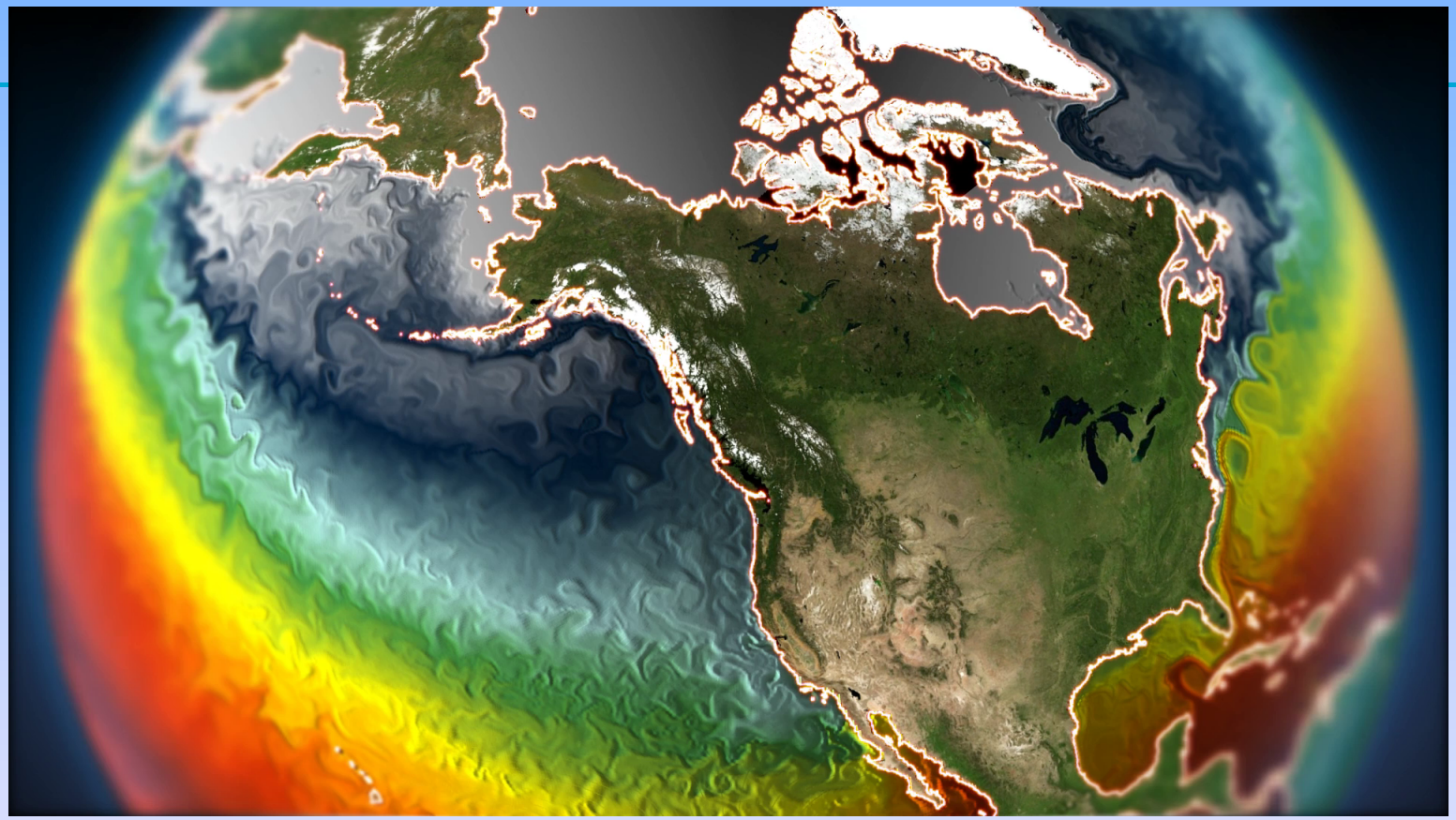
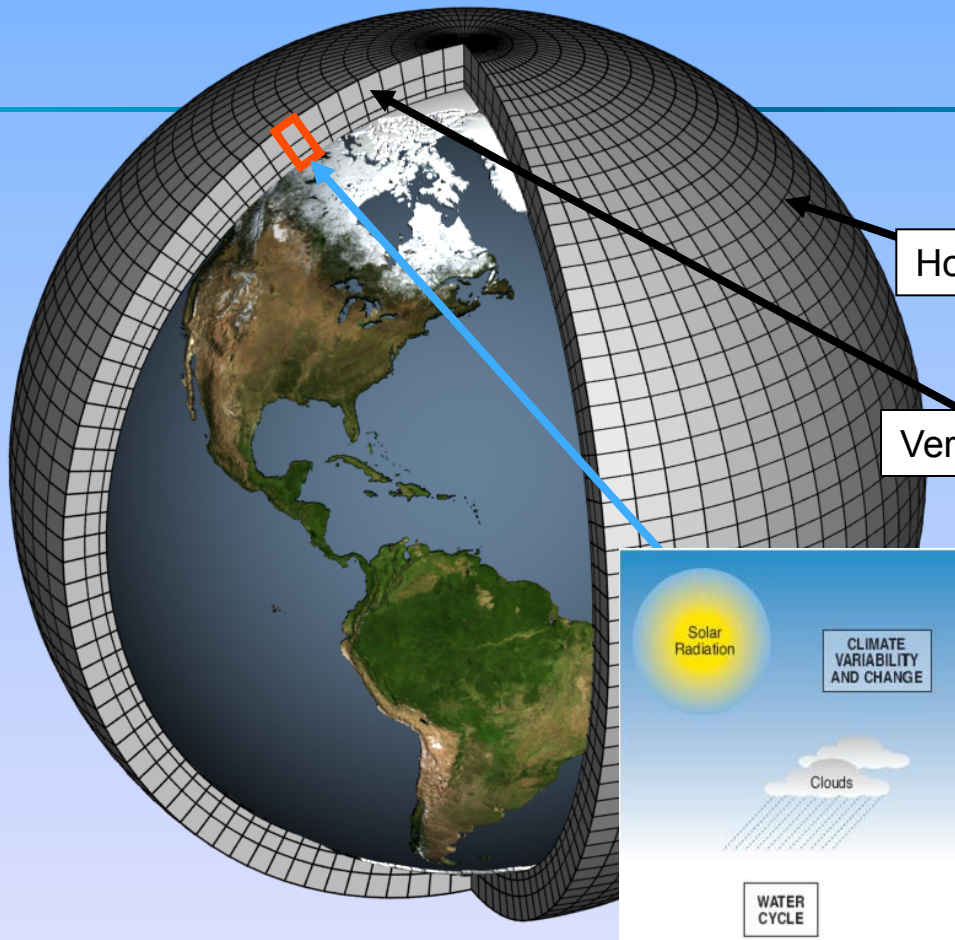


# Scientific User's Perspective on Lustre

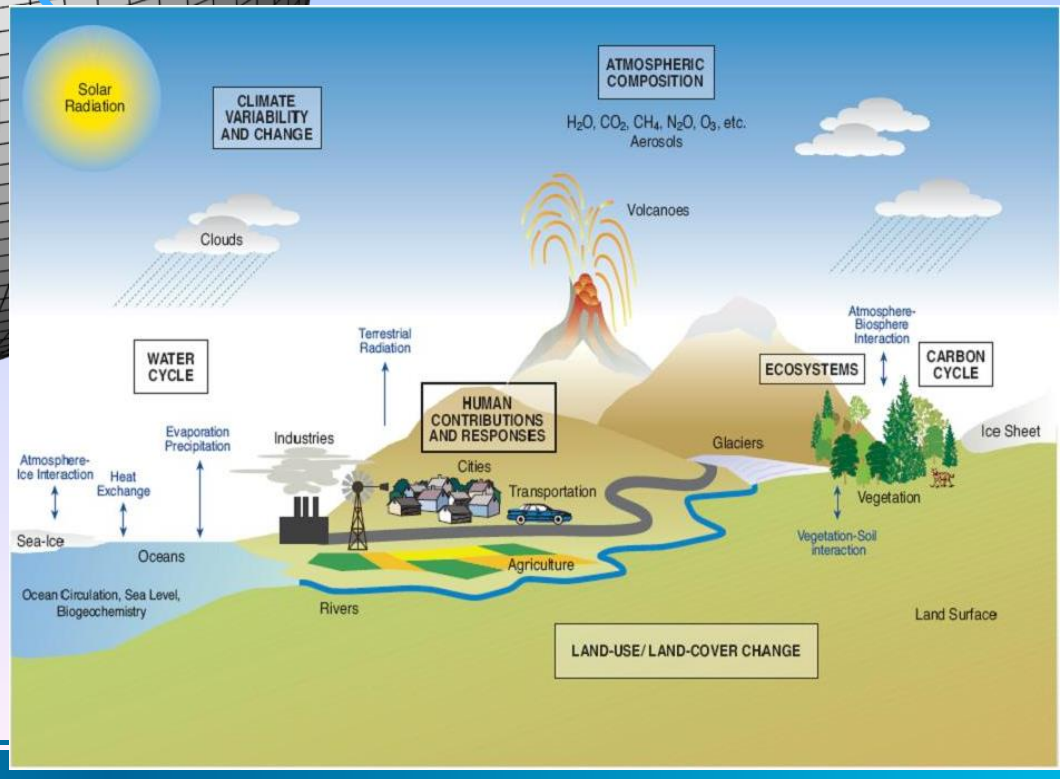
Frank Indiviglio – HPTi @ GFDL

# Schematic Global Climate Model



Horizontal Grid (Latitude-Longitude)

Vertical Grid (Height or Pressure)




A GCM is a mathematical representation of the major climate system components and their interactions. The GCM equations operate on a global grid and are solved on a computer.

# The User Community

At GFDL we have a variety of user experience levels with lustre.

Some have years of experience with large installations, others have just recently been introduced to lustre on Gaea.

I obtained information on the user perspective by interviewing members of the GFDL user community.

# First, the Positive


# What do the users think?

- Users have a generally good perception of Lustre.
- Lustre provides the scientists with a platform on which to do large scale experimentation.
- Lustre also provides the scalability that supports the users needs to scale the science.

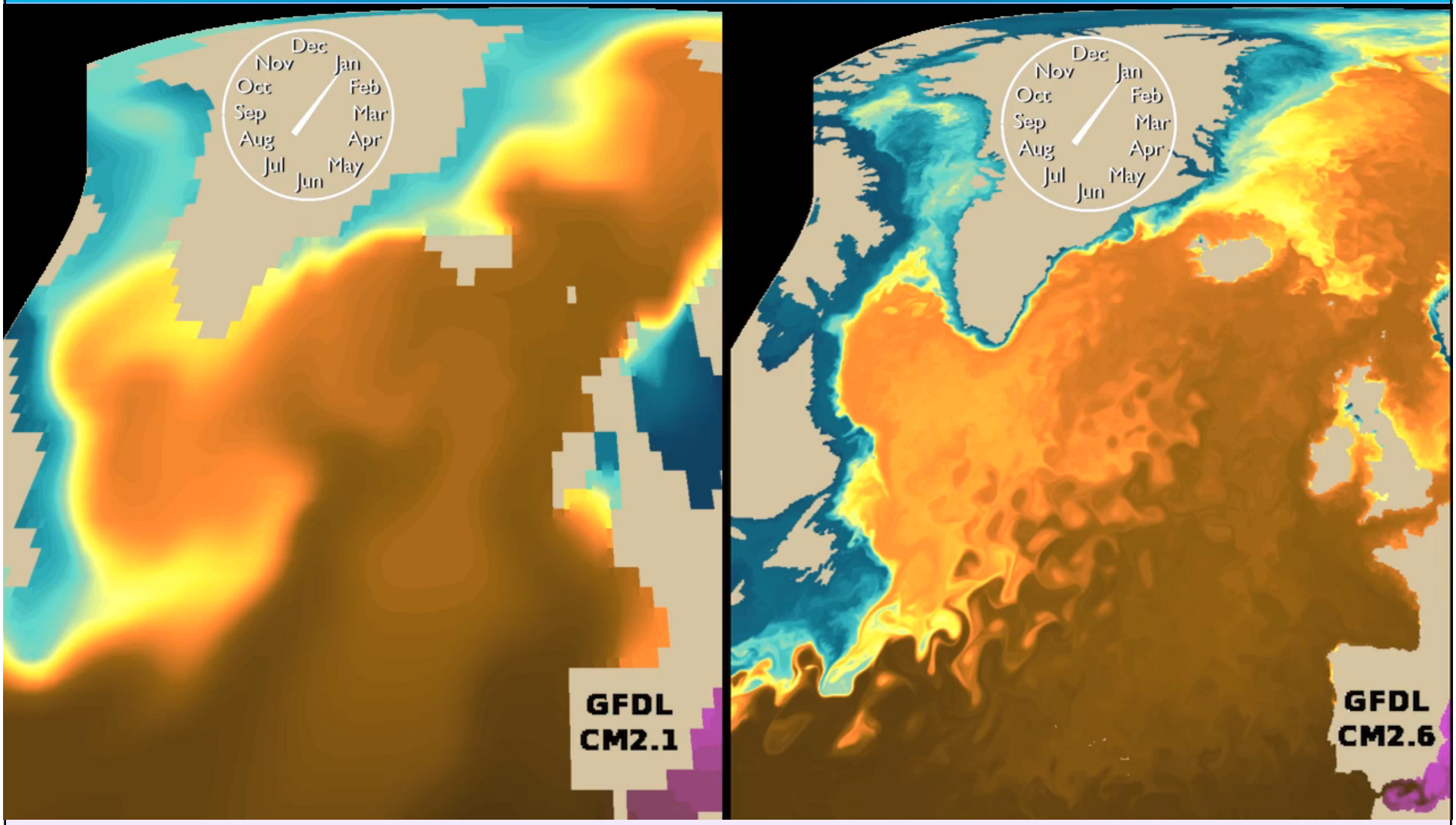
# Users Perceptions

- Performance
  - Model initialization took 15 mins before now it takes 8 mins.
- Reliability
  - Generally it's a reliable and stable filesystem.
- Size
  - Scalability allow for large filesystems, less data movement, and larger experiments.

# Transition to Lustre

- Education is important in the process:
  - Most users have an established methodology in there workstreams that depend on standard Unix tools.
    - Find replacement tools for the workstream.
  - The filesystem isn't monolithic.

# Research HPC




CM2.1: 200 km atmosphere, 100km ocean  CM2.6: 50km atmosphere, 10km ocean

# Now, what needs improvement

# Other Perspectives

# Users Perspective

- They are not insulated from the bad practices of misbehaving users.
- Do not have the necessary tools to manage the filesystems and user behavior.
  - Quotas
  - Slowness and potential issues with using standard unix commands
    - du, ls, find, etc.


# Users Perspective

- Buffering Sensitivity
  - When moving from 1.8.4 to 1.8.5 users started to see holes in their data.
  - I/O patterns changed in applications tend to stay around a long time.
    - Users are reluctant to change I/O for what is perceived as a transient problem.
  - This issue usually surfaces intermittently, so it becomes difficult to plan for all end cases.


# Users Perspective

- Confused with problems in their jobs resulting from OST or OSS failures.
- Users don't know if the I/O error they receive in their output is permanent or transient.
- If parts of the filesystem are offline, users and management want the ability to quickly see this and adjust the running workload to it.
  - Ideally, this would be automated.

# Users Perspective

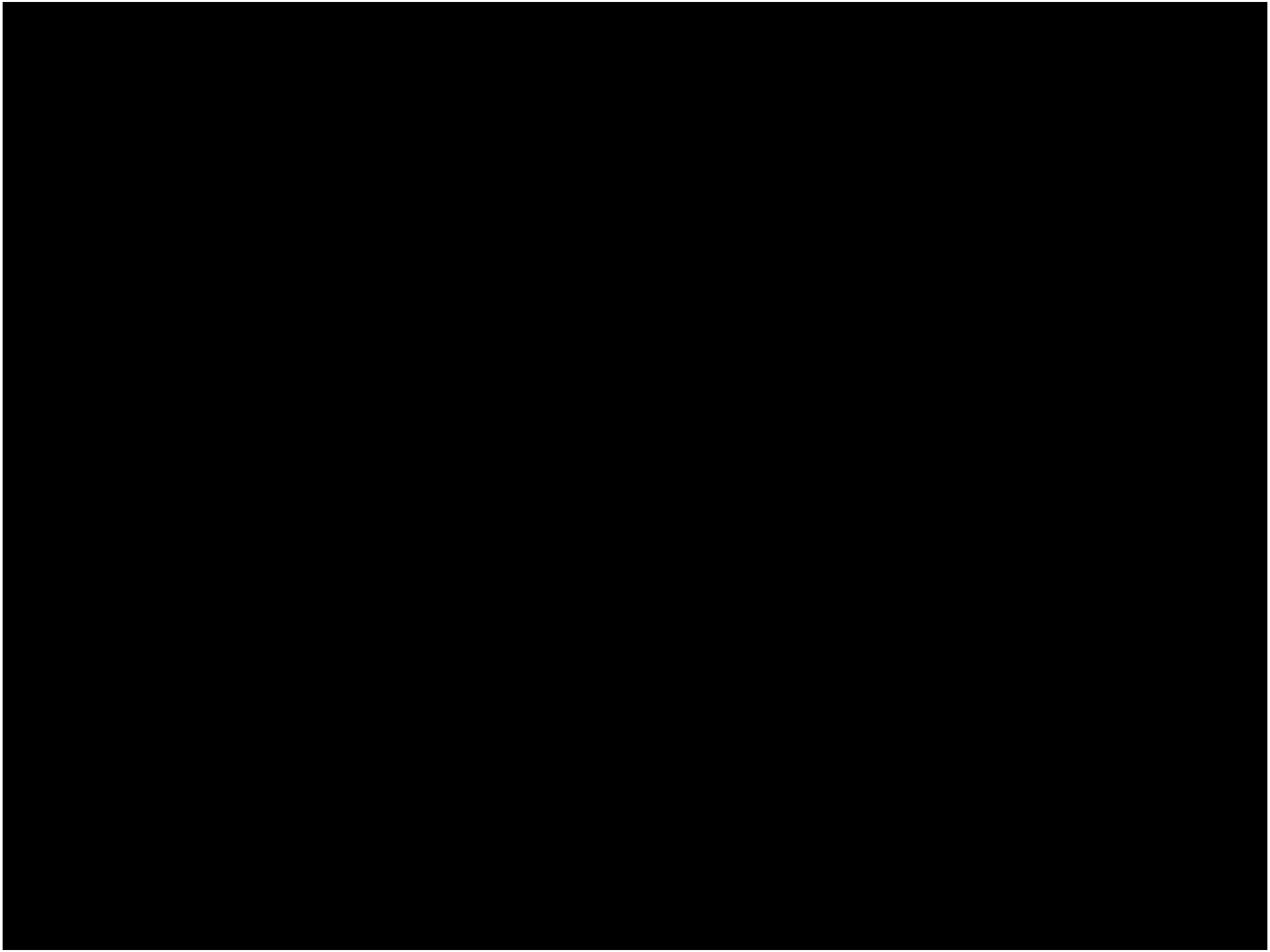
- Want API enhancements to signal jobs and/or the scheduler when there are filesystem problems.
- Want better error messaging and reporting.
- All liked the fact the system was resistant to hardware issues.


# Users Perspective

- Users and Management both view growing a filesystem is a painful process.
  - Growth requires taking a long outage to rebuild the filesystem.
  - Both view this is something that would ease scientific growth and allow more flexibility in planning.

# Summary

- GFDL users have a generally good perception of Lustre.
- Scalability and performance were viewed favorably.
- Educating the traditional user is important.
- The users perspective can be improved: better reporting, better error messages, allow for the use of standard Unix tools.
- Options for growth are important for planning and flexibility of an organization.




# Thank You

Frank Indiviglio

HPTi @ GFDL

frank.indiviglio@noaa.gov

findiviglio@hpti.com

Special thanks to

Rich Brueckner

from

InsideHPC