

I/O Challenges



- Large data volume and high velocity
- Data management practice does not scale and is not portable
 - ▶ Cannot easily manage file placement and knowledge of what file contains
 - ▶ Hierarchical namespaces does not reflect use cases
 - ▶ Individual strategies at every site
- Data conversion/merging is often needed
 - ▶ To combine data from multiple experiments, time steps, ...
- The storage stack becomes more inhomogeneous
 - ▶ Non-volatile memory, SSDs, HDDs, tape
 - ▶ Node-local, vs. global shared, partial access (e.g., racks)
- Suboptimal performance & performance portability
 - ▶ Users cannot properly exploit the hardware / storage landscape
 - ▶ Tuning for file formats and file systems necessary at the *application* level