

Scalable Storage and I/O

Rob Ross

Senior Computer Scientist

Mathematics and Computer Science Division

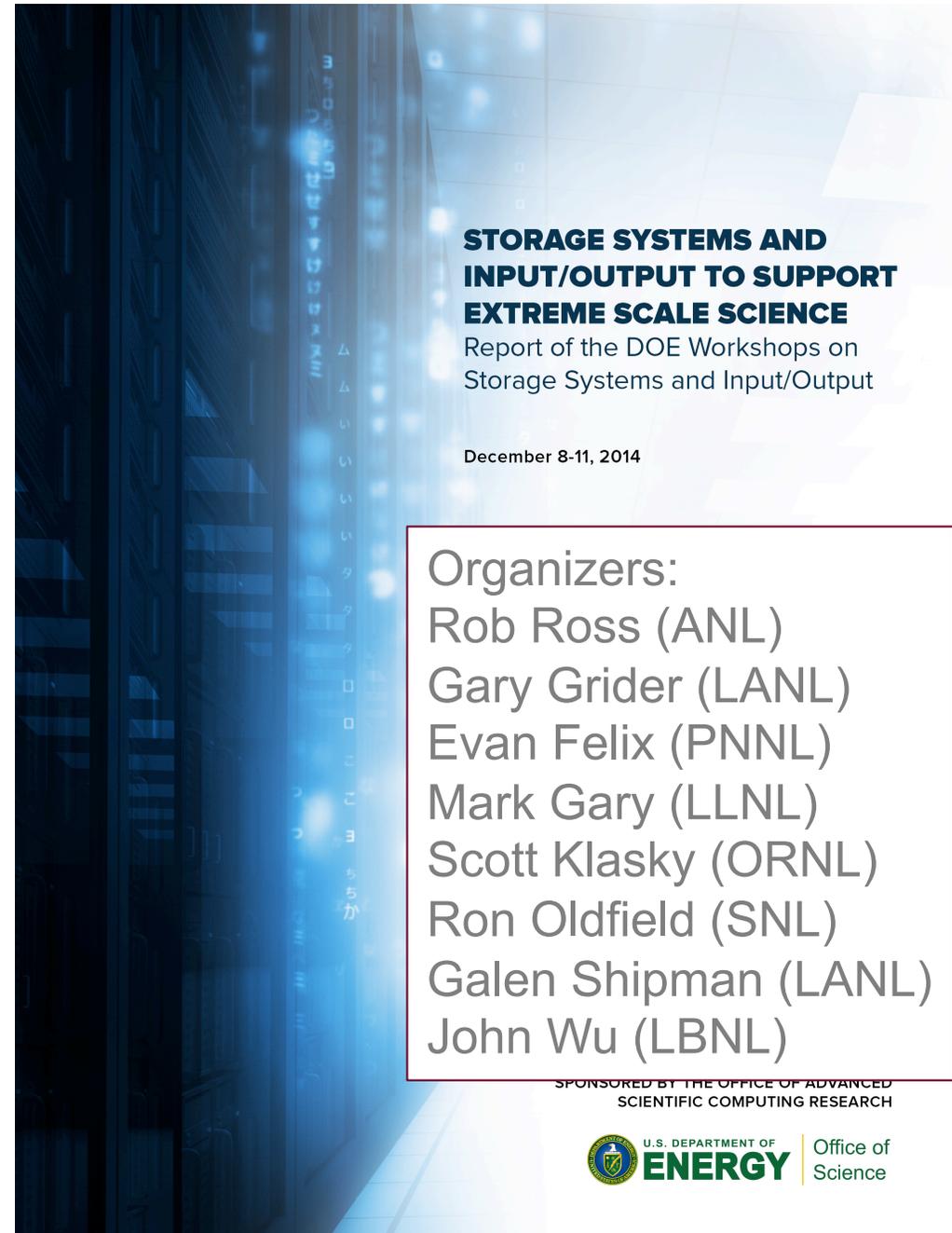
Argonne National Laboratory

ross@mcs.anl.gov

Thanks to Gary Grider for creating some of this material.

Recent SSIO Workshop(s)

- Workshop(s) back in December 2014
 - Day 1: apps people talking with SSIO and software stack people
 - Day 2: software stack and SSIO people using apps requirements to talk about cross cuts
 - Days 3-4: SSIO researchers talk about needed R&D given previous 2 days input



Organizers:
Rob Ross (ANL)
Gary Grider (LANL)
Evan Felix (PNNL)
Mark Gary (LLNL)
Scott Klasky (ORNL)
Ron Oldfield (SNL)
Galen Shipman (LANL)
John Wu (LBNL)

SPONSORED BY THE OFFICE OF ADVANCED
SCIENTIFIC COMPUTING RESEARCH



Topics in HW/SW Architectures for Storage and I/O

- **Networks**
 - Topology
 - Understanding Impact
- **Deep Storage Hierarchies**
 - Technology Integration
- **Nonvolatile**
 - Use: Burst Buffers, On-Node
 - Memory/Storage Convergence (incl. Prog. Models)
- **Active Storage**
 - Programmable FS/Storage
- **Resilience/Integrity/Availability**
 - End-to-end Data Integrity
 - Resilience to Component Failures
 - Availability
 - Coding Techniques
 - Checkpointing Strategies
- **Autonomics**
 - Garbage Collection
 - Adapting to User Demands/Behavior
- **Understandability (also another session)**
 - Architecting for Understandability
 - Acting on Prediction
- **Security**
 - Security and HPC Services
 - Avoiding Information Leakage
- **New (to us) Paradigms/Capabilities**
 - Key-Value
 - Scheduling of I/O
 - Coupling with Experimental/Observational Facilities



SSIO Workshop Findings

- **In situ data analysis** is already an important component of many applications.
- New **solid state and disk storage layers** are complicating the storage hierarchy.
- Scientists need a **coherent view and management methods of the storage resources**.
- Current SSIO designs are hindered by their **isolation from system-level resource management**, monitoring, and workflow systems.
- Many important aspects of **app/system SSIO behavior aren't well understood**.
- New requirements for results validation may change the role of SSIO systems.
- New programming models/systems drive new persistence mechanisms.
- Scientists desire increasingly complex data abstractions that improve productivity.
- Community access to data on apps and systems needed, as well as test environments, for new technology evaluation and bringing new talent into the community.

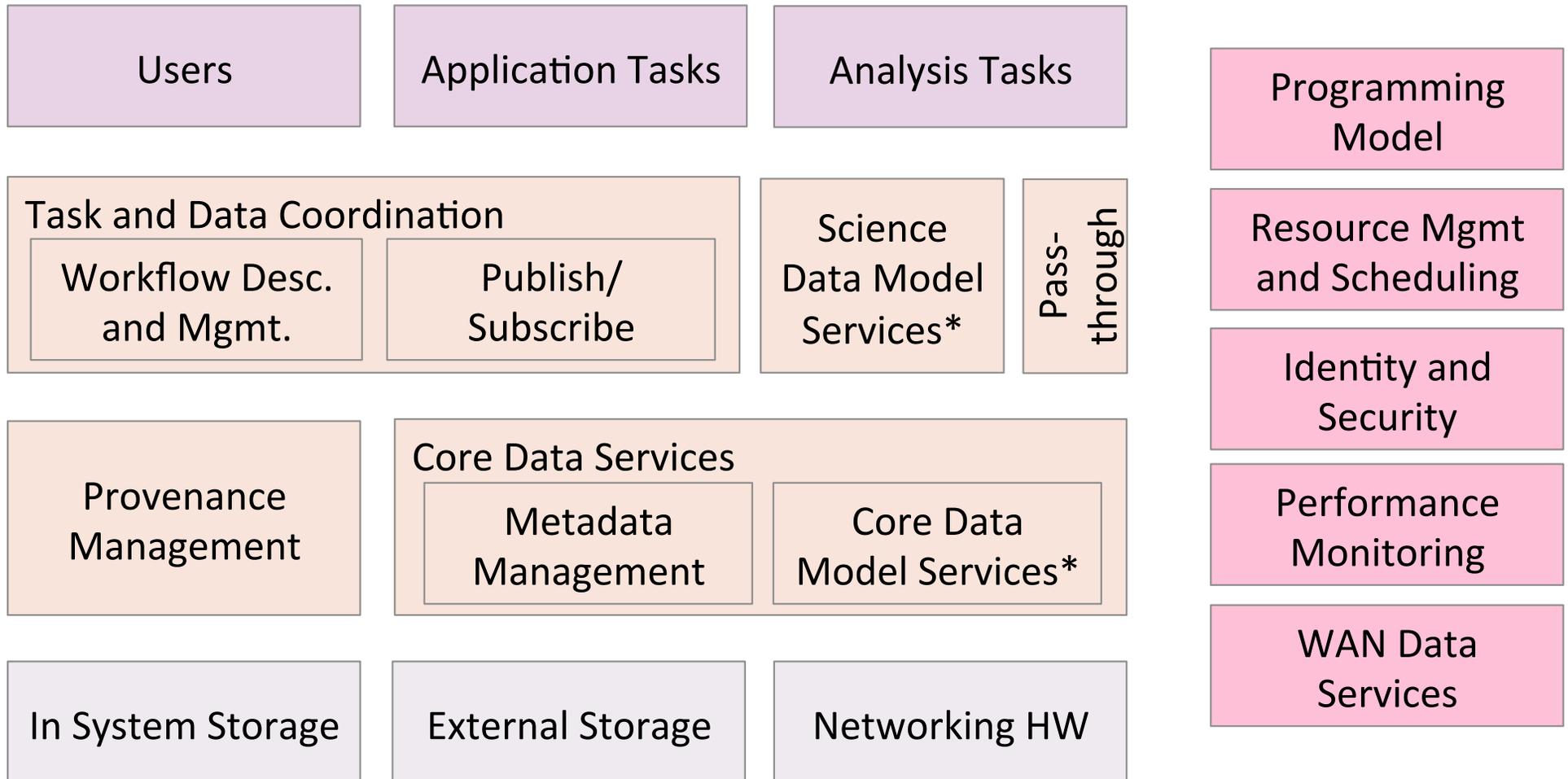


Near Term SSIO Research Priorities from Workshop

- SSIO architectures research for:
 - managing deep and heterogeneous storage hierarchies
 - alternative management paradigms to the file system model
- In metadata, name spaces, and provenance:
 - new methods of management of rich metadata
 - breaking away from the current file model
- In the area of supporting science data :
 - develop the next generation of I/O middleware and services to support new programming abstractions and workflows
- In the area of understanding SSIO:
 - improve our ability to characterize storage activities to model and predict the behavior of SSIO activities on future systems.



Somewhat Bigger Picture: Data Management Services



Functionality in an SSIO solution

- Data/metadata storage, persistence, resilience, naming
- Bridging between networks
- Performance monitoring
- Executing user code (presumably in some sandbox)
- Probably built using a custom RPC system
- Possibly built using a custom programming language

- Block management on storage resources
- Group membership, STOMITH (kill misbehaving resource)
- Client interface (usually embedded in node OS)

- Lots of potential for sharing of components with runtime/OS

Some Questions (in Conclusion)

- In what ways are runtime systems consumers of storage and I/O resources/services?
 - What data models and capabilities are needed to support those uses?
- How do runtimes and SSIO cooperate to connect the pieces of the deep memory hierarchy?
- How can SSIO, runtimes, and other components capture adequate provenance for validation of results?
- What's the relationship between runtimes and active storage?
- What common building blocks are shared by runtimes and SSIO? Can technology be shared?
- How do we coordinate cross-cutting activities?