

Open Problems and Gaps in High End Computing File Systems and I/O (HEC FSIO):

James Nunez
Los Alamos National Lab

ASCR PI Meeting
April 1, 2008

LA-UR-07-5890
Unclassified

Introduction

- The HEC-IWG FSIO
 - What, Who, Why?
- Open problems and gaps in existing research

The Formation of the HEC FSIO

- Created in response to U.S. Presidential commission report that noted under funding in high tech
- The High End Computing Revitalization Task Force (HECRTF)
 - engage in planning activities to guide future investments
- Interagency Working Group on HEC(HECIWG)
- The HECIWG decided to do some pilot coordination efforts and thus begat the HEC FSIO Technical Advisory Group (HEC FSIO)

HEC FSIO

- Who
 - DoE Office of Science and NNSA
 - NASA
 - Others
- To coordinate government spending on File Systems and I/O R&D by all the government agencies that are involved in HEC
 - NSF
 - DOE Office of Science & NNSA
 - DoD
 - NASA
 - Others
- Listing, categorizing, and prioritizing HEC I/O and File Systems R&D needs.

FSIO Gaps and Barrier Areas

- Metadata
- Quality of Service
- Measurement and Understanding
- Security
- Next Generation I/O Architectures
- Manageability
- RAS
- Communications and Protocols
- Archive
- Assisting R&D

Metadata

- **Definition:** Metadata is structured, encoded data that describe characteristics of information-bearing entities to aid in the identification, discovery, assessment, and management of the described entities
- **Sample Problem:** Groups samples into single files whose names are a concatenations of sample attributes:

04142007_LANL_V3E_proton_52_122

Date

Facility

Device

Radiation source

Angle

Tilt

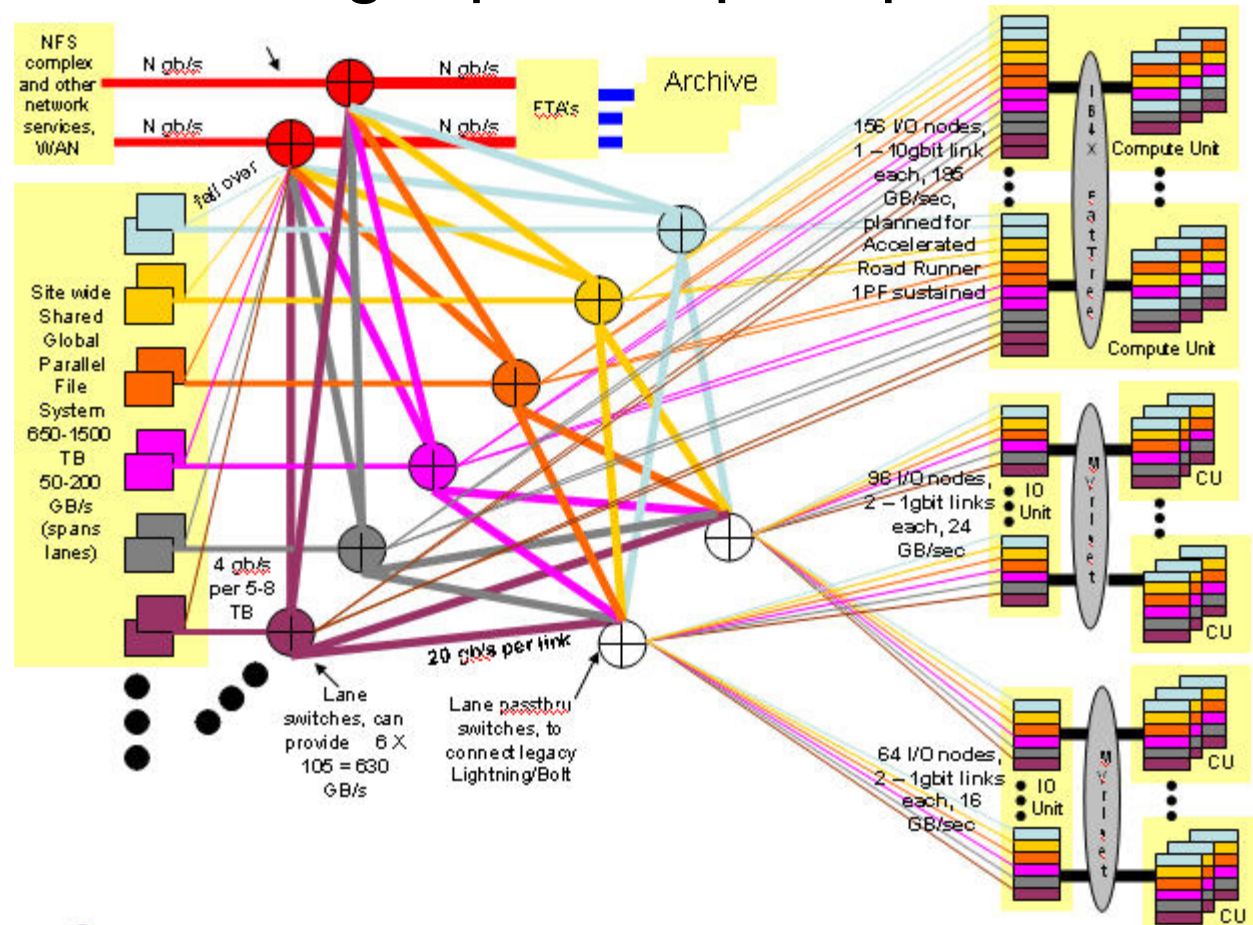
Metadata

- Open problems
 - Extensibility
 - Scalability
 - Exploitation of hybrid storage devices
 - File system and Archive integration and coherency
 - Transparency and access methods

Quality of Service

- **Definition:** Quality of Service can be defined as features of a storage architecture which allow a user or administrator to recommend policies for data movement during Input/Output operations.

- **Problem:**



Quality of Service

- Open problems
 - End-to-end QoS
 - Guaranteed I/O performance
 - Specific Redundancy Requirements
 - A standard API
 - Hardware Support

Measurement and Understanding

- **Definition:** Methods to gain insight into the design, use and optimization of FSIO for users and designers.
- **Problem:** My I/O is slower than it was yesterday. Why?

Measurement and Understanding

- Open problems
 - Understanding Layering Contribution
 - End-to-end Standardized Benchmarks
 - End-to-End (Efficient) Tracing Methods
 - Visualization and analysis tools to large scale traces
 - System workload in the enterprise
 - Testbeds

Security

- Definition: Authentication and authorization
- Open problems
 - Usability
 - Long-term key management
 - End-to-end encryption
 - Overhead and scaling
 - Tracking of information flow, provenance
 - Ease of management, quick recovery, APIs

Next Generation I/O

- **Definition:** Applying/Generating new technologies to apply to existing/future problems
- **Problem:**



Unclassified

Next Generation I/O

- **Open Problems:**
 - Small record access
 - Understanding abstractions, naming, organization
 - Architectures
 - Self-* components; Assembly, reconfiguration, healing
 - Managing millions of components
 - Hybrid devices

Communication and Protocols

- Impact from networks and protocols
- Open problems
 - Active networks
 - Alternative transport schemes
 - Coherent schemes



Archive

- “Definition”: How to never have to delete anything
- Open problems
 - APIs and standards for interface, searches, attributes, staging, ...
 - Long term, attribute-driven, security
 - Data reliability and management
 - Metadata scaling
 - Policy-driven management

Management and RAS

- Config, deploy, debug, tune, and reliability, availability, servicability
- Open problems
 - Existing Reliability Techniques
 - Automated analysis and modeling
 - Formal failure analysis
 - Scalability
 - Power consumption and efficiency

Road Map Example

2007 QoS Gap Area								
Area	Researchers	CY06	CY07	CY08	CY09	CY10	CY11	Rankings
End to End QoS in HEC	Brandt	Partial Calendar Year Funding	Full Calendar Year Funding	Partial Calendar Year Funding				 Good research, but much work needed to get a standards based solution.
	Chiueh	Partial Calendar Year Funding	Full Calendar Year Funding	Full Calendar Year Funding	Partial Calendar Year Funding			
	Ganger	Partial Calendar Year Funding	Full Calendar Year Funding	Partial Calendar Year Funding				
Standard API for QoS	SciDAC - PDSI	Partial Calendar Year Funding	Full Calendar Year Funding	Full Calendar Year Funding	Full Calendar Year Funding	Full Calendar Year Funding	Partial Calendar Year Funding	 Very partially addressed by HECEWG. Will be driven by above "End to End QoS in HEC".
	POSIX HPC Extensions	On-Going Work	On-Going Work	On-Going Work	On-Going Work	On-Going Work	On-Going Work	
	PVFS	On-Going Work	On-Going Work	On-Going Work	On-Going Work	On-Going Work	On-Going Work	

-  Very Important
-  Greatly Needs Research
-  Greatly Needs Commercialization
-  Medium Importance
-  Needs Research
-  Needs Commercialization
-  Low Importance
-  Does Not Need Research
-  Does Not Need Commercialization
-  Full Calendar Year Funding
-  Partial Calendar Year Funding
-  On-Going Work

Resources

- HEC FSIO planning site
 - <http://institute.lanl.gov/hec-fsio/>
 - Documents
 - Workshop output
 - Gaps/analysis
 - R&D descriptions
 - Etc.

Conclusion

- The field is ripe for a paradigm shift
 - In architectures, protocols, and physical devices