

HEC FSIO Session 2: Metadata Gaps Roadmap

James Nunez and Rob Ross
August 2008



Metadata Gaps

- Scaling
- Extensibility and Name Spaces
- File system/archive metadata integration
- Hybrid devices exploration
- Data transparency and access methods

Presentations on Metadata

- High Throughput I/O for Large Scale Data Repositories
 - Ali Saman Tosun, University of Texas SA
- Petascale I/O for High End Computing
 - Matt Wolf, GA Tech for Karsten Schwan, GA Tech
- An Update-Aware Disk I/O Interface for High-Throughput Database Indexes
 - Tzi-cker Chiueh, SUNY
- Techniques for Streaming File Systems And Databases
 - Michael Bender, SUNY
- SAM² Toolkit: Scalable and Adaptive Metadata Management for High-End Computing
 - Yifeng Zhu, U Maine

2007 Metadata Gap Area

Area	Researcher	FY 07	FY 08	FY 09	FY 10	FY 11	FY 12	Rankings
Scaling	Bender/Farach-Colton	■	■	■				<p>All existing work is evolutionary. What is lacking is revolutionary research; no fundamental solutions proposed.</p>
	Leiserson	■	■	■				
	Maccabe/Schwann	■	■	■				
	SciDAC - PDSI	■	■	■	■	■	■	
	HECEWG HPC Extensions	■	■	■	■	■	■	
	UCSC's Ceph	■	■	■	■	■	■	
	Lustre	■	■	■	■	■	■	
	ANL/CMU – Large Directory PVFS	■	■	■	■	■	■	
Extensibility and Name Spaces	Bender/Farach-Colton	■	■	■				<p>All existing work is evolutionary.</p>
	Leiserson	■	■	■				
	Tosun	■	■	■	■			
	Wyckoff	■	■	■				
	UCSC – LIFS/facets	■	■	■				
	ANL/CMU - MDFS	■	■	■				
	SciDAC PDSI	■	■	■	■	■	■	
File System/ Archive Metadata Integration	Lustre HSM	■	■	■	■			<p>Extended Attributes, although not standardized, could solve problem.</p>
	UMN Lustre Archive	■	■					
Hybrid Devices Exploitation	<i>None</i>							<p>Research is being done, but no research focused on metadata</p>
Data Transparency and Access Methods	<i>None</i>							<p>No research focused on metadata</p>

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