Mero: Co-Designing an Object Store for Extreme Scale

Presented at PDSW'2016(SC'16)

Presented by
Sai Narasimhamurthy

Co-authors: Nikita Danilov, Nathan Rutman [All Seagate]
John Bent[Seagate Government Solutions]
Storage Software for BDEC

Exascale Computing

Big (Massive!) Data Analysis

Exascale Data Centric Computing (Big Data Extreme Computing, or BDEC)

Mero: Object Storage Software

✓ Build from the ground up to cater to BDEC
✓ Involved co-design with the community
Mero Simplicity

State of the Art Example

Mero Stack
Mero
Co-Design Timeline

Munich QAW
- API Hints
- Tiered Storage Management
- Data Layouts
- Data Compression
- Plugins
- Telemetry

2009
Lustre Problems going into Exascale
Paris Lustre Architects Meeting

2010
EIOW/E10 Setup through EOFS

2011

2012
Tokyo QAW
- Plug-ins
- Lazy Commit and Rollback
- Is-I problem
- High Availability

Portland QAW
- Data “Containerisation”
- Distributed Transaction Management
- Experimental Data Handling

2013

2014

2015

2016
SAGE Co-Design
- Nuclear Fusion data
- Human Brain Science
- Bio-informatics
- Synchrotron data
- Space Weather
- Climate science
- Data Analytics
Everything you’d expect in an Exascale Object Store.. PLUS

Complex Layouts
PGAS and MPI One sided Comms

Peer-Peer Caching

In-Storage Compute
Analytics

Event Log Subscription & Diagnostics

Expect More!
Ongoing Work & Next Steps


The SAGE storage system, will be capable of querying, retrieving and sharing immense volumes of data at Extreme scales, with the added function and capability to accept and perform user defined computations integral to the storage media. The SAGE Cluster will be built around the HDFS and other distributed file system technologies, that will work together to provide the required functionalities and analytics required by Extreme scale workflows. The SAGE system will seamlessly integrate a new generation of storage devices, including high bandwidth flash memories as they become available. The SAGE system will also offer a powerful API and a powerful framework that is suitable for easy extensibility by third parties.

This white paper provides a technical overview of the SAGE system and describes its key component pieces and the extended capabilities and tools required to support it

http://sagestorage.eu/

http://www.sagestorage.eu/content/public-deliverables