A blurred photograph of a crowd of people crossing a street at a crosswalk. The people are in motion, creating a sense of a busy, high-volume environment. The crosswalk has white stripes on a dark asphalt surface.

# 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]



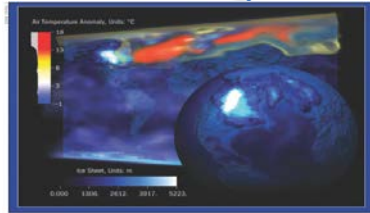
SEAGATE

# Storage Software for BDEC



Exascale  
Computing

Big (Massive!)  
Data Analysis



Exascale Data Centric Computing  
(**B**ig **D**ata **E**xtrême **C**omputing, or *BDEC*)

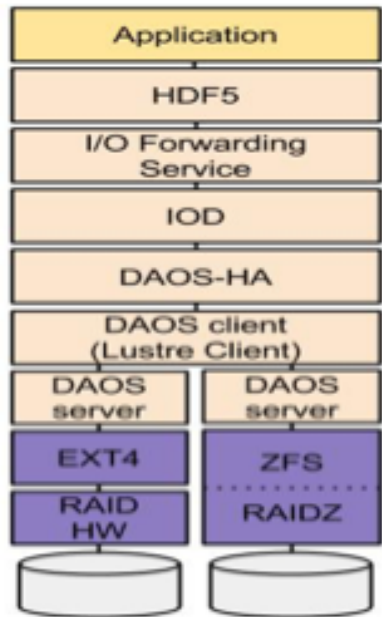


Mero: Object Storage Software

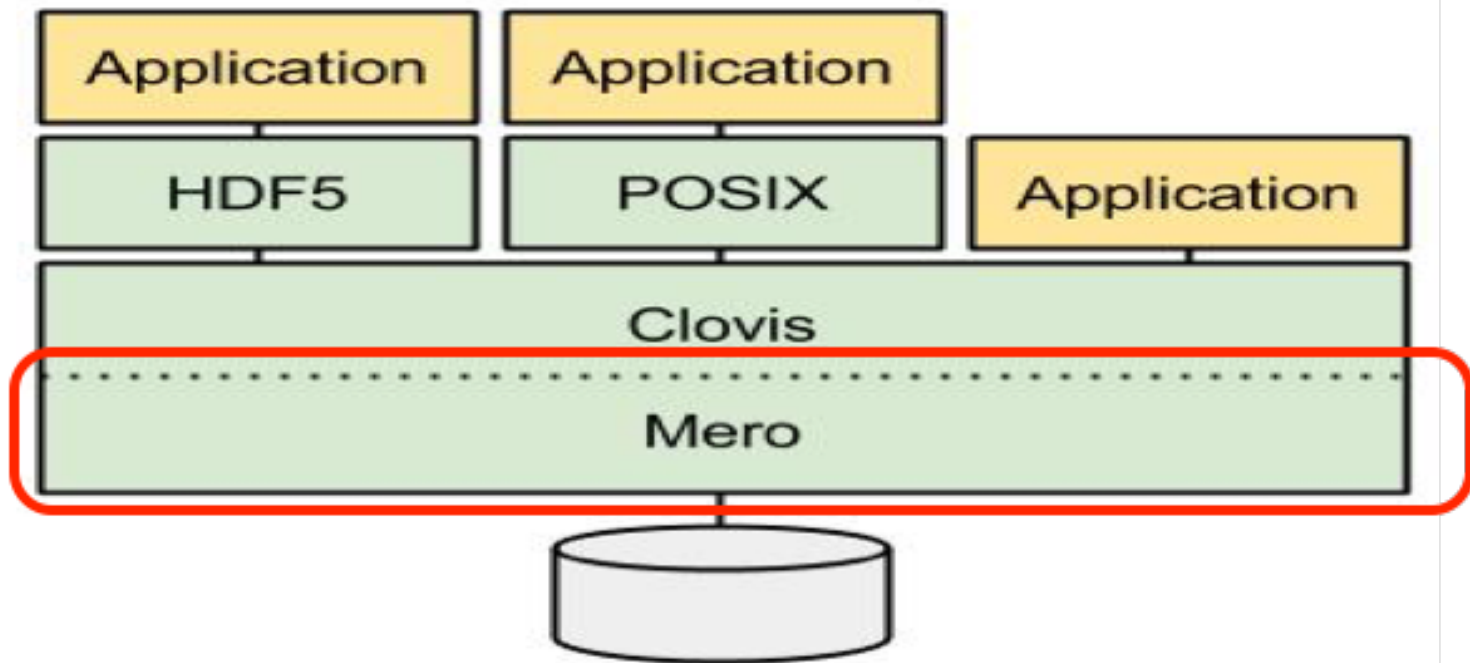
- ✓ Build from the ground up to cater to BDEC
- ✓ Involved co-design with the community



# Mero Simplicity



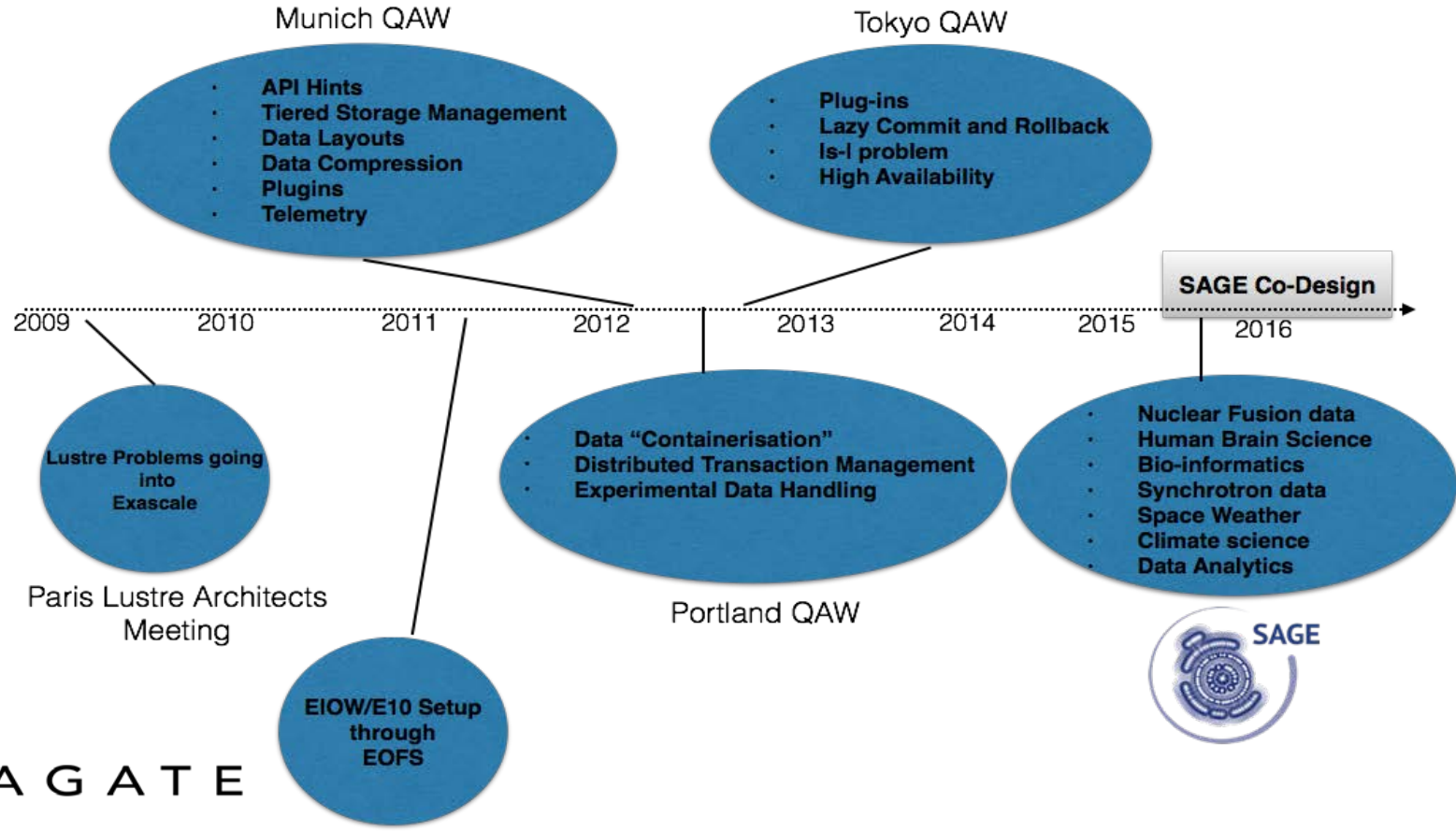
**State of the Art  
Example**



**Mero Stack**

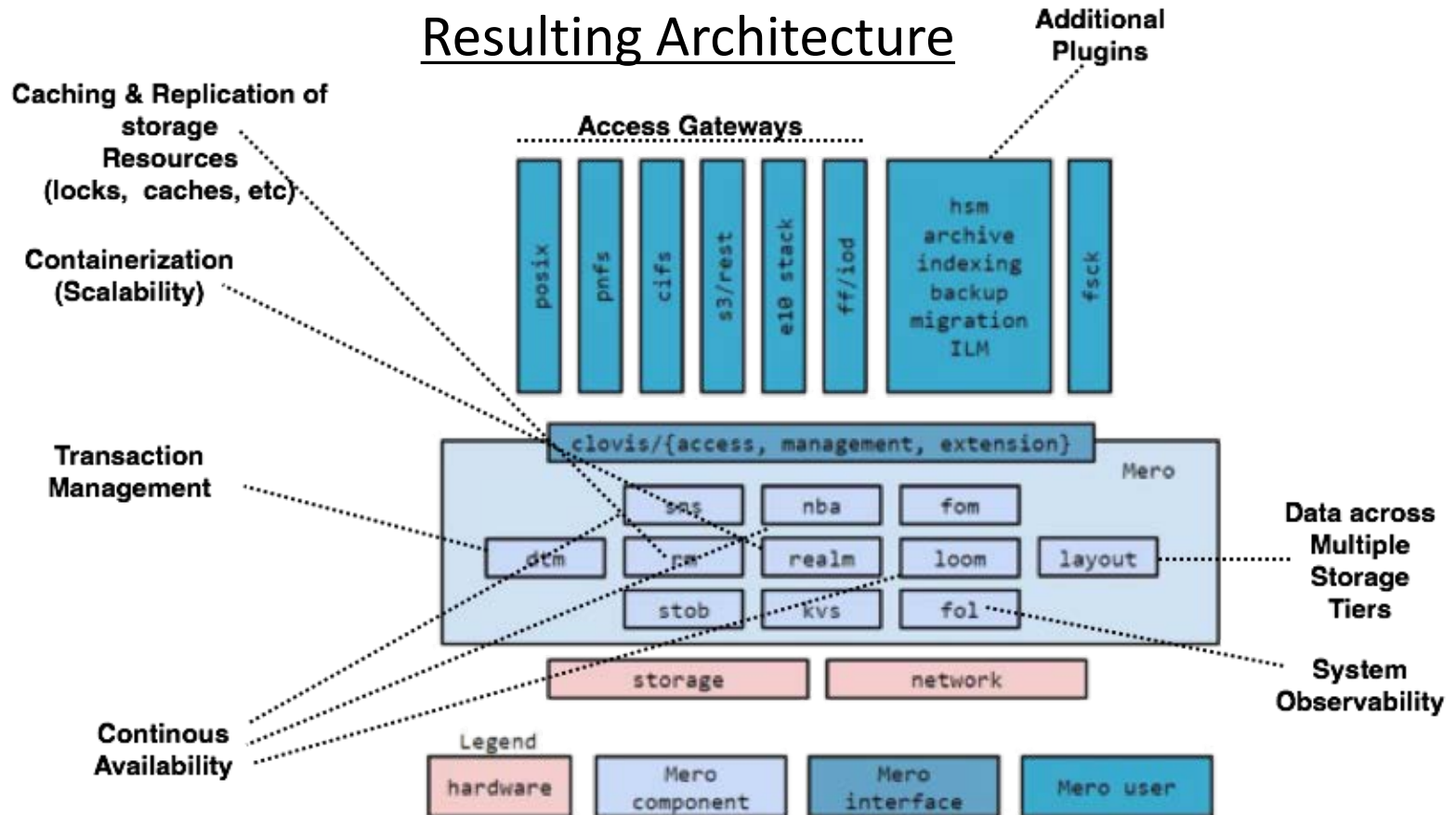
# Mero

## Co-Design Timeline



# Mero

## Resulting Architecture



# Everything you'd expect in an Exascale Object Store ..PLUS

Complex Layouts

PGAS and MPI One sided Comms

Peer-Peer Caching

Expect More!

In-Storage Compute

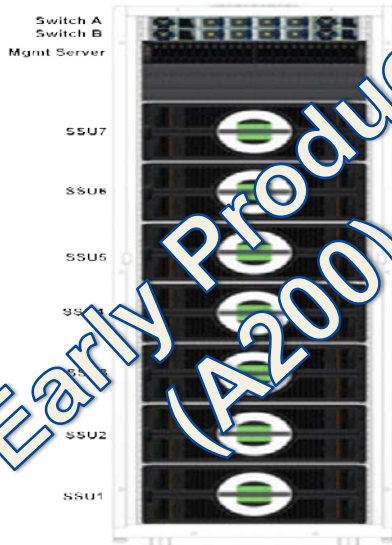
Analytics

Event Log Subscription & Diagnostics



# Ongoing Work & Next Steps

[http://www.seagate.com/files/www-content/product-content/xyratex-branded/clustered-file-systems/\\_shared/datasheets/seagate-clusterstor-a200-datasheet.pdf](http://www.seagate.com/files/www-content/product-content/xyratex-branded/clustered-file-systems/_shared/datasheets/seagate-clusterstor-a200-datasheet.pdf)



A200 Base Rack



SEAGATE

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

The StorAGE for Exascale Data Centric Computing (SAGE) system, research and built as part of the SAGE project, aims to implement a Big Data/Extreme Computing (BDEC) or High Performance Data Analytics (HPDA) capable infrastructure suitable for Extreme scales - including Exascale and beyond. Increasingly, overlaps occur between Big Data Analysis and High Performance Computing, caused by the proliferation of massive data sources, such as large, dispersed scientific instruments, sensors, and social media data, whose data needs to be processed, analysed and integrated into computational simulations to deliver scientific and innovative insights.

The SAGE storage system, will be capable of efficiently storing and retrieving immense volumes of data at Extreme scales, with the added functionality of the capability to accept and perform user defined computations integral to the storage system. The SAGE system will be built around the Mero object storage software platform and its supporting ecosystem of tools and techniques, that will work together to provide the required functionalities and capabilities required by Exascale workflows. The SAGE system will seamlessly integrate a new generation of storage device technologies, including non-volatile memories as they become available. The SAGE system will also offer a flexible API and a powerful software framework 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 created to support it

<http://sagestorage.eu/>

Future Results & Demos

Watch this space!