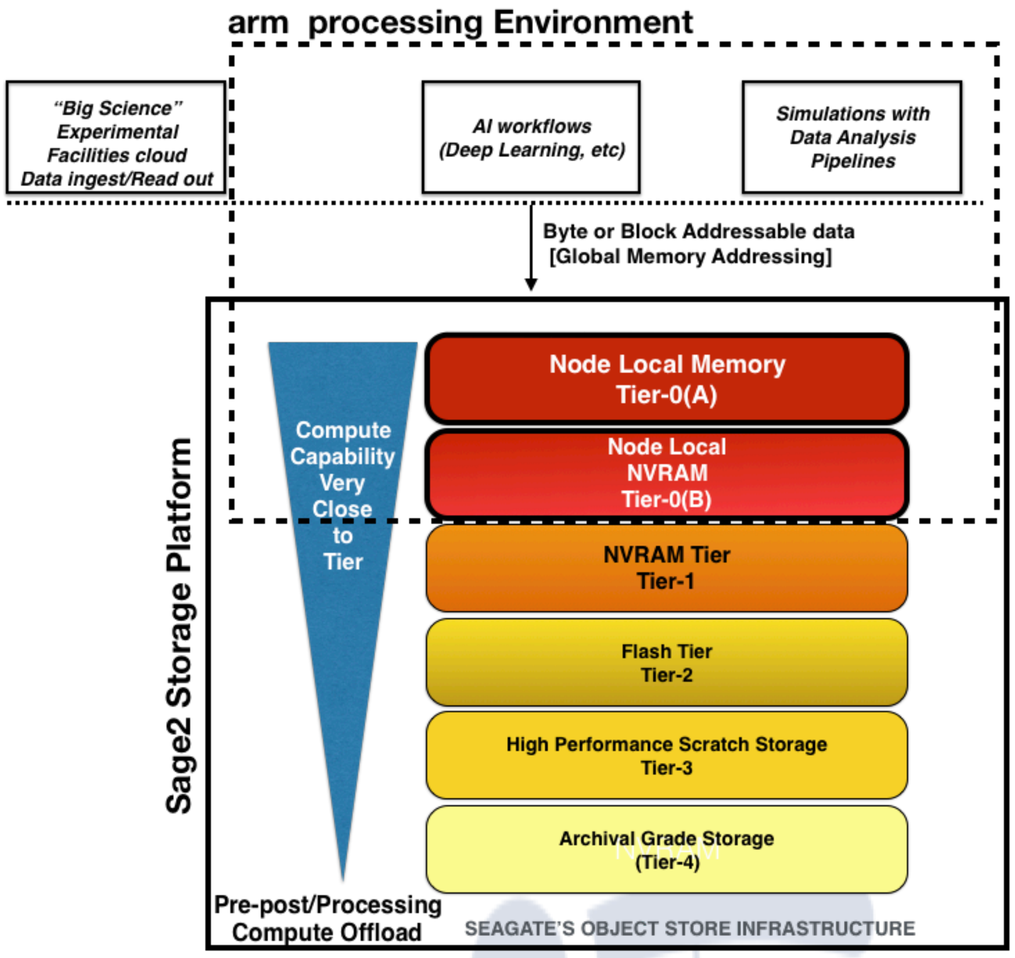
Sage2 – Percipient StorAGe for Exascale Data Centric Computing

S. de Witt[1], D. Samaddar[1], A.Davis[1], S. Narasimhamurthy[2], G. Umanesan[2], D. Pleiter[3], M. Salem El Sayed[3]

[1] *United Kingdom Atomic Energy Authority*
[2] *Seagate (UK) Ltd*
[3] *Jülich Supercomputing Centre*



Consortium: 9 partners led by Seagate (UK) Ltd. covering technology providers, computing centres and application providers and users

Objective: Design an exascale ready architecture based on hierarchical storage layers and object storage (MERO), adaptable to novel storage media .

Justification: I/O can represent a bottleneck in many applications . This bottleneck can be addressed by facilitating use of fast but expensive storage media in an hierarchical architecture. The use of storage-class memory additionally to spinning disks blurs the line between byte addressable and block data store

Unique Features:

- Containers for treating many objects the same
- Distributed Transaction Management
- Supports global memory abstraction
- Integrated AI framework
- Function shipping to move processing to the data
- Data Aware Scheduling based on SLURM
- QoS based Architecture

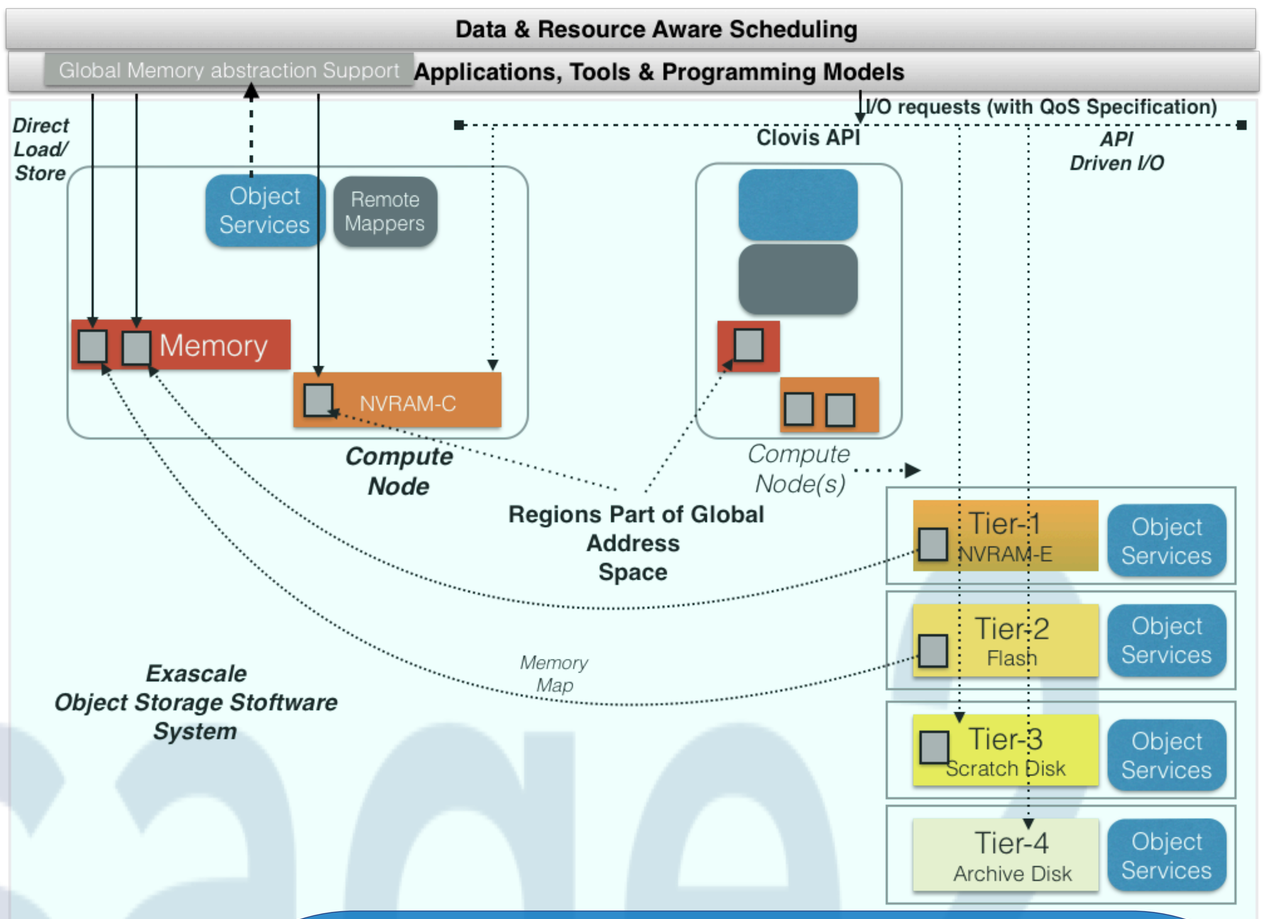
Global Memory Abstraction: applications, tools and programming models access data on the SAGE system either directly by mapping objects to memory in the tier nodes, or, by accessing them as a storage system through the Clovis API.

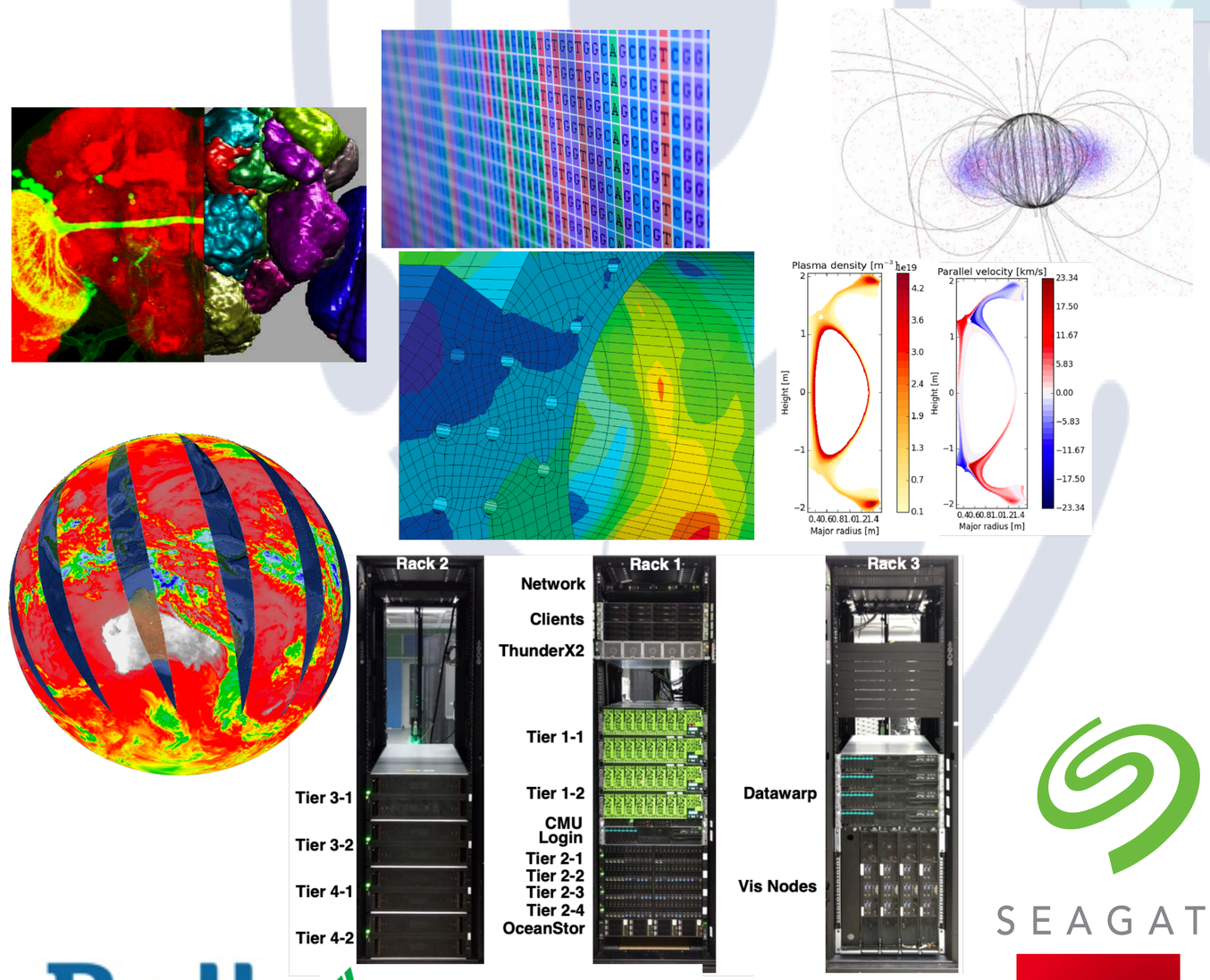
Three scenarios envisaged:

- Make NVRAM-C (NVRAM within compute nodea) appear as a single addressable space
- For larger objects, make use of external NVRAM (NVRAM-E), directly accessing using RDMA
- Using standard MERO APIs to access data in any tier

Under Development:

- Differential checkpointing of objects to back them up to lower level storage
- Data Reshaping to allow restart with different numbers of nodes





Co-Design: SAGE architecture has been co-designed with application developers and users. Applications cover

- Space weather (IPIC3D)
- Satellite Data Processing (JÜRASSIC)
- Finite Element Modelling (paraFEM)
- CFD (BOUT++)
- Neural Simulation Tool (NEST)
- CAT (Savu)

Application Support and Programming Models

ARM have developed extensions to their Allinea Forge toolkit and are supporting integration of MERO onto ARM processors. Kitware are adapting their visualization tools to optimize user experience when analysing results on SAGE.

Current and Future Status:

- Prototype system installed and operational at Jülich Supercomputing Centre
- MERO object store under consideration for Open Sourcing

The sage2 Project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement 800999