Modeling the Allocation of Heterogeneous Storage Resources on HPC Systems

**PhD Student:**
Julien Monniot - julien.monniot@inria.fr
INRIA Rennes / Université Rennes 1

**Advisors:**
François Tessier - INRIA Rennes
Gabriel Antoniu - INRIA Rennes

---

**Context and motivations**
Examples:
- Perlmutter 35PB all flash storage at NERSC (2021)
- Aurora DAOS Storage with Intel Optane persistent memory at ALCF (2022)
- Summit compute nodes with 1.6TB of NVRAM at OLCF (2018)

---

**Main axes of work**
- Dealing with resource heterogeneity
- Fair and efficient use of resources

---

**Solution design**
- Data deluge from new large-scale scientific workflows
- PFlops ∗ TBps
- Underutilization of resources
- Complexity & Underutilization of resources

---

**Simulation results**
Showcase problem: What is the correct sizing for a burst buffer capacity, when accounting for the effect of a storage scheduling algorithm and strategy (splitting large requests)?

---

**In situ visualisation**
- Total allocated volume at time t
- Number of concurrent allocations at time t
- Allocation details per disk / code

---

**Implementation**
- Python3
- Docker
- OpenMPI (MPI simulation)
- ZMQ (Messaging)
- bokeh (Real-time plotting)
- seaborn (Plotting)
- Darshan (IO traces)

---

**Storage-aware job scheduler simulator**
- Extensible with new algorithms
- Independent components design
- Common messaging interface
- Abstraction of heterogeneous storage hardware