

Supervised job preemption methodology for controlled memory consumption of jobs running in the ALICE Grid

Marta Bertran Ferrer¹, Kalana Wijethunga²

¹marta.bertran.ferrer@cern.ch, CERN

²kalana.16@cse.mrt.ac.lk, University of Moratuwa



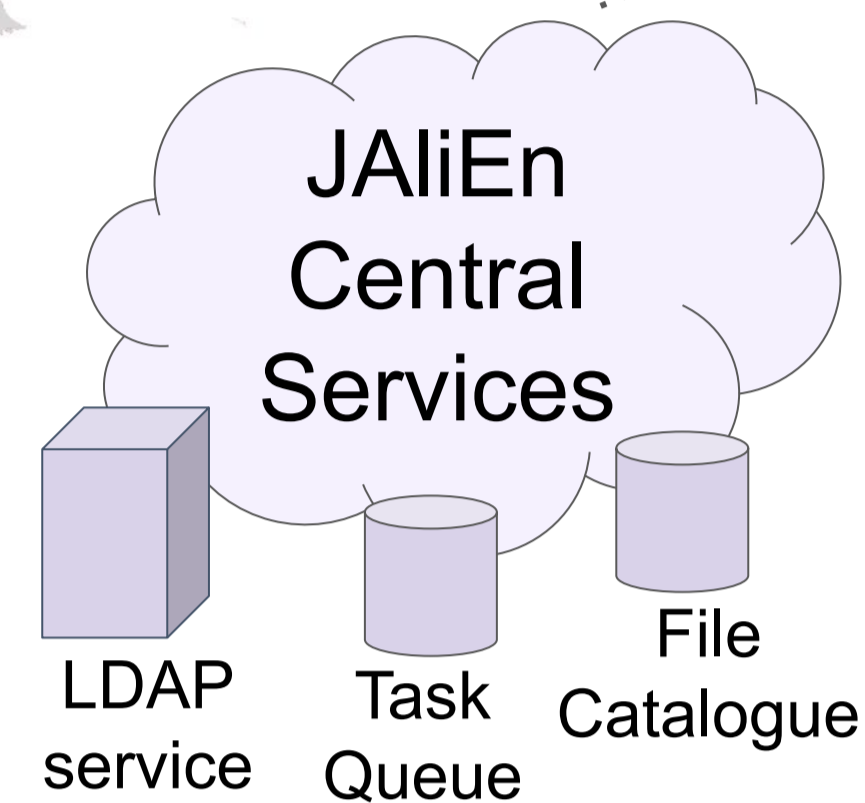
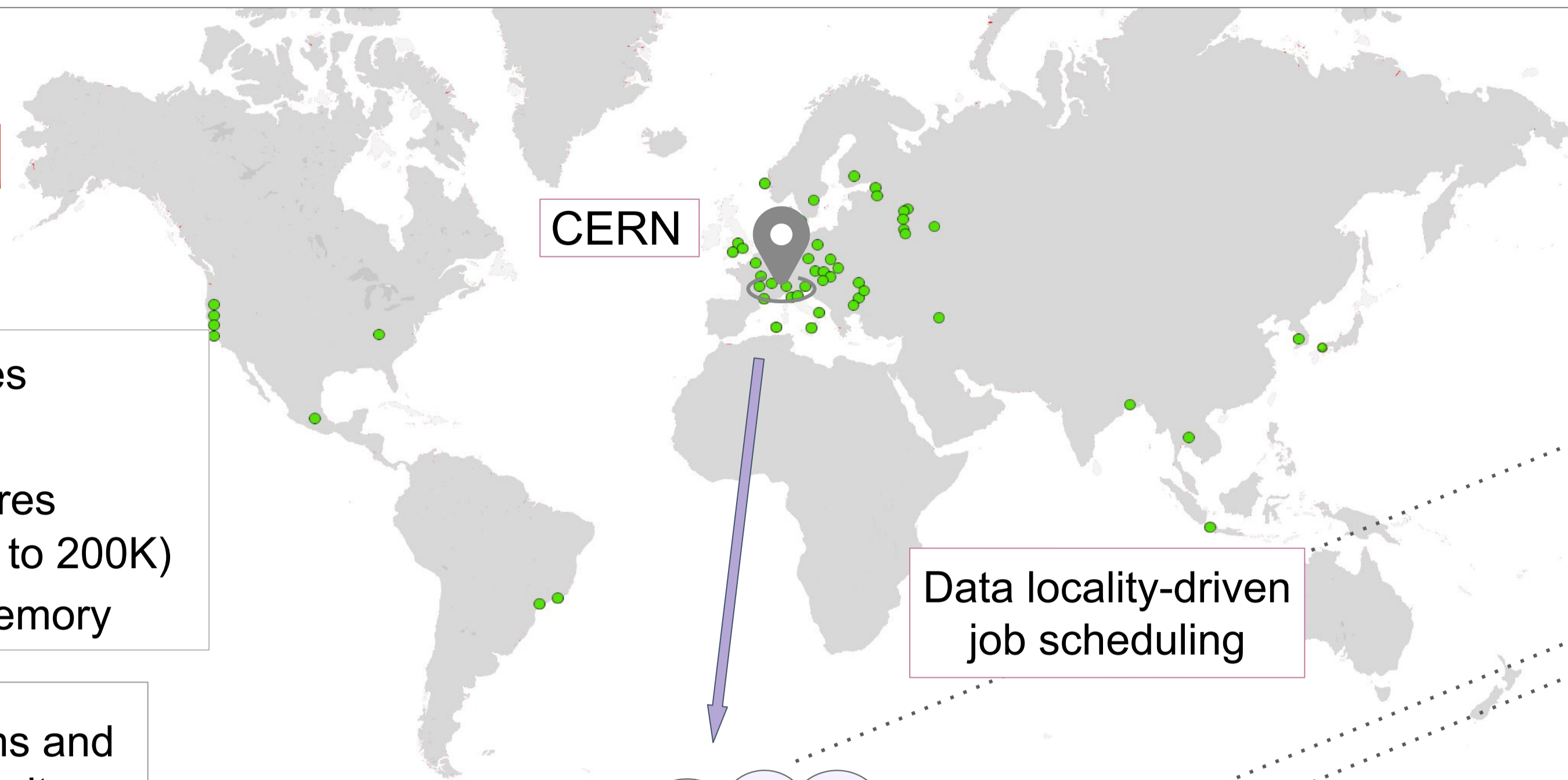
The ALICE Grid

54 WLCG Grid sites
9441 worker nodes
608K shared CPU cores (ALICE gets up to 200K)
2.05 PB shared RAM memory

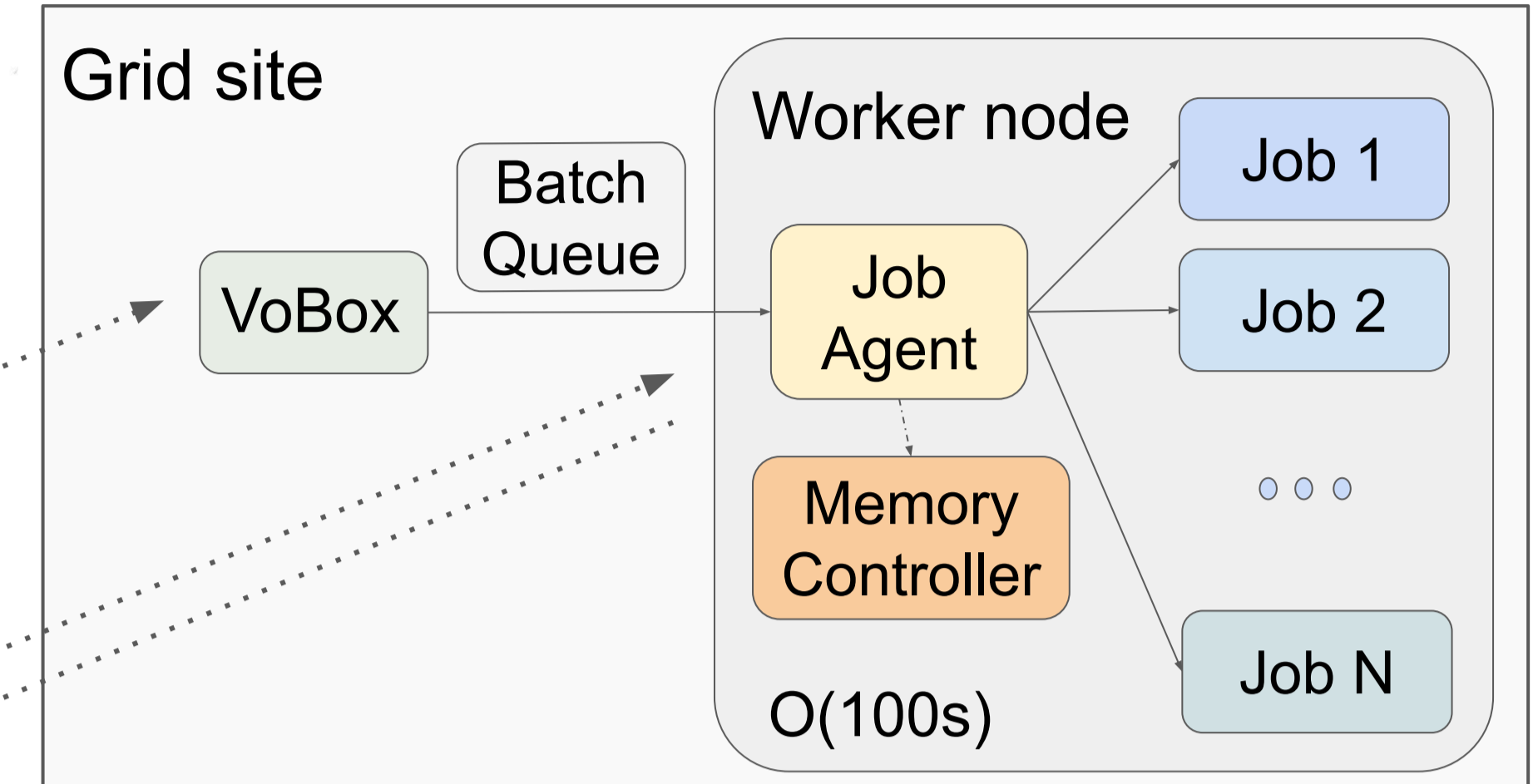
Heterogeneous systems and configurations across sites

Running workloads with different resource needs (CPU, memory, disk...)

Grid worker nodes are typically shared with other VOs



Data locality-driven job scheduling

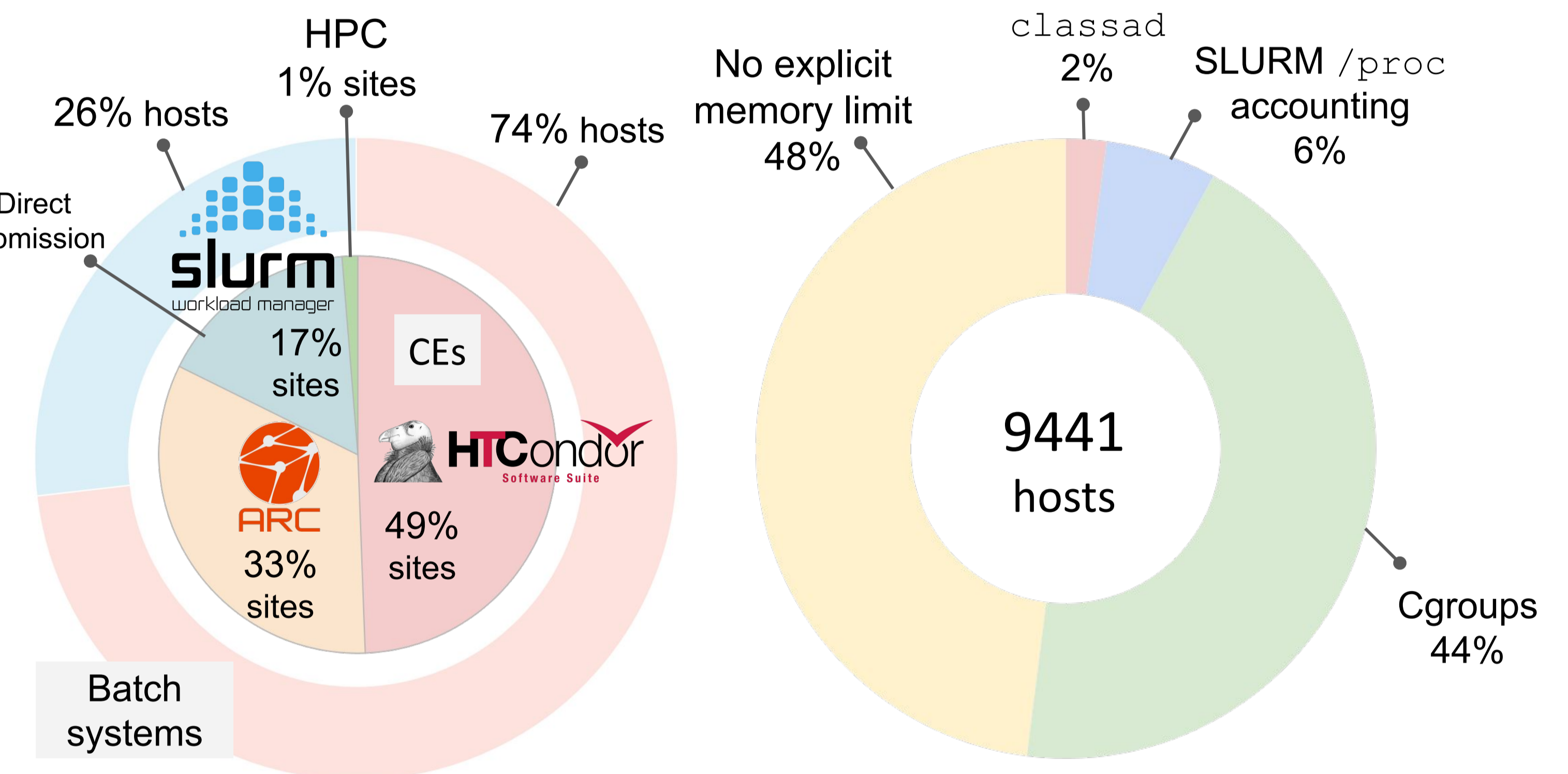
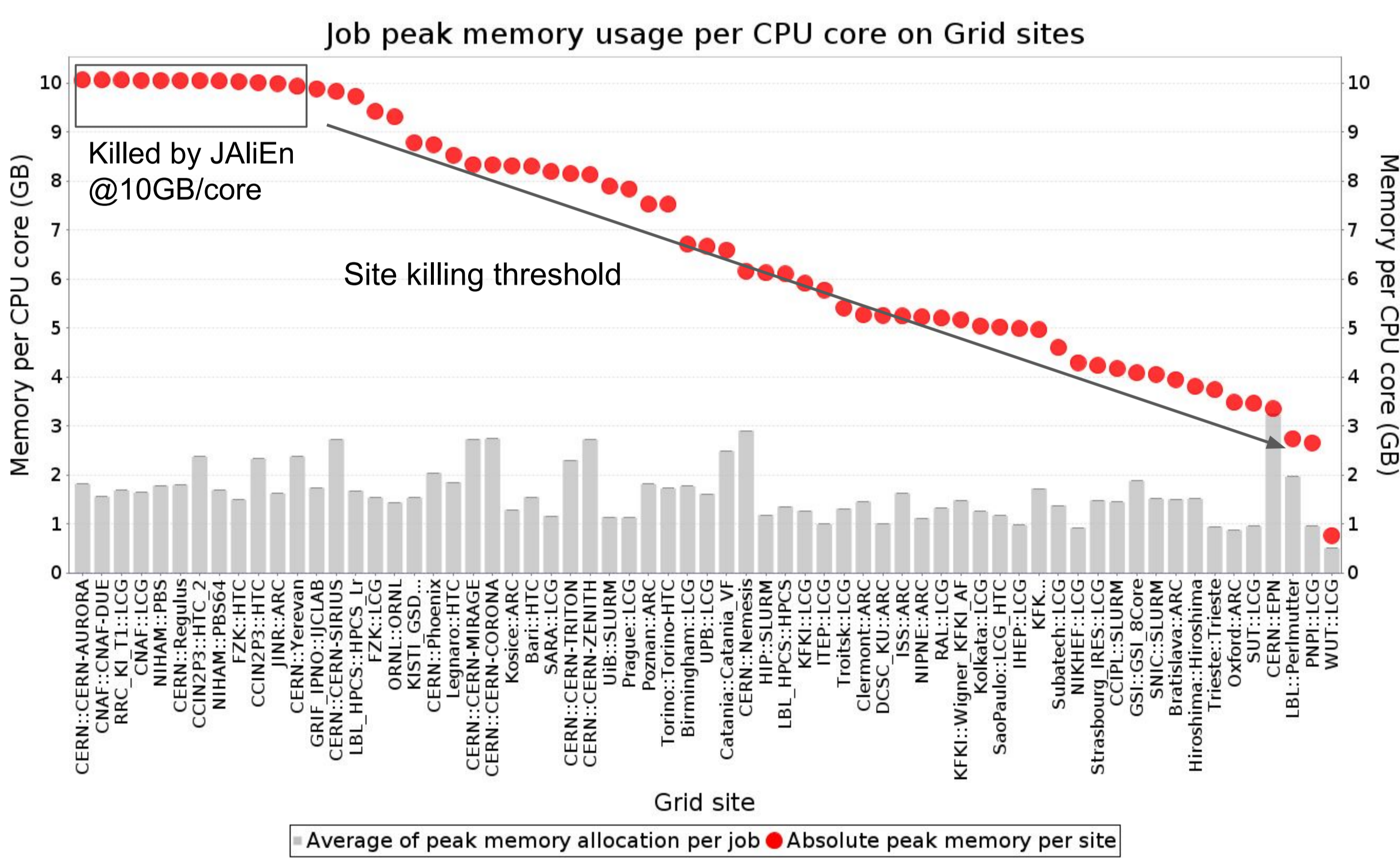


Slot resources can be sub-partitioned among the running jobs
- Jobs executed in the same slot **share resources** (CPU, RAM)
Resource **over-consumption** by any of the jobs can impact the whole slot
- The whole slot is killed → all jobs go into EXPIRED JAIEn state, or
- System kills a running process → affected job goes to ERROR_E state

ALICE payloads have a **high memory footprint**
- Jobs can be killed due to memory over-consumption

Need to **anticipate** killing decisions to protect well-behaved jobs

Memory management on Grid sites



Number of Computing Elements (inner ring) and Batch Systems (outer ring) used in the Grid

Memory limiting on Grid hosts

ALICE Sites need to provide **at least 2GB RAM/core**
Allocation impacted by fluctuating machine memory usage
Allocation limits on memory resources limited by:
- Physical machine specs
- Constraining policies on the operating system or batch system level

Memory management depending on configuration of batch systems
- Sites with static thresholds - **Predictable** behaviour, **perfect candidates** for controlled job preemption
- Sites with no memory limit or soft limit - **Fluctuating** allocation, depending on machine utilization levels

Memory constraints need to be identified on individual basis to apply the appropriate preemption method based on slot memory consumption

The memory controller module

