Distributed Caching in the WLCG

Max Fischer, Eileen Kühn, René Caspart, Tabea Fessenbecker, Manuel Giffels, Christoph Heidecker, Günter Quast, Andreas Petzold

Applicability of Caching in the WLCG

- Increased usage of network-bound resources
  - Remote data access across Grid sites
  - University clusters with limited bandwidth
  - Opportunistic compute resources
- WLCG is a unique use-case for data access
- Distributed caching reduces network usage
  - Replicate data close to compute sites
  - Trade network versus storage capacity
  - Exploit volatile, expendable resources
- WLCG requires diverse caching approaches

Distributed Cache Management

- Predict placement for both jobs and data
  - Ideal placement given enough metadata
  - Limited scalability and fault tolerance
- Iteratively converge job and data placement
  - Accuracy of placement limited by design
  - Ideal scalability and fault recovery

Opportunistic Data Provisioning

- Caches face complex data access patterns
  - Network is powerful and often uncongested
  - Many workflows do not need high throughput
- Need for locally adaptive cache management
  - Significant impact from time and location
  - Different requirements for various workflows

Distributed Cache Management in the WLCG

- Various approaches to cache management
  - Extremes of predictive versus reactive
  - Trade precision against scalability
- Diverse data access scenarios
  - Perfect cache not defined by hit rate
  - Optimise capacity, suitability, agility, …
- Robust foundation for distributed caching
  - Underlying technology well established
  - Research into orchestration/coordination
- Design architectures for WLCG and beyond
  - Support existing infrastructure and workflows
  - Expect opportunistic and regional resources

Max Fischer
Karlsruhe Institute of Technology
max.fischer@kit.edu

COBaLD