**Inter-Operating Grids through Delegated MatchMaking**

Alexandru Iosup, Dick Epema, Hashim Mohamed, Mathieu Jan, Ozan Sonmez
Parallel and Distributed Systems Group, TU Delft
Todd Tannenbaum, Matt Farrellee, Miron Livny
CS Department, U. Wisconsin-Madison

Agenda
- Context: our grid infrastructure
- Inter-operating grids
  - Topology
  - Mechanisms
- Conclusion

**Grid Infrastructure**
- DAS-2
  - Homogeneous
  - 5 sites, 5 clusters
    - 200 nodes, Dual 1GHz Pentium III
    - Myrinet and Surfnet
    - PBS, then SGE
- DAS-3
  - Heterogeneous
  - 4 sites, 5 clusters
    - 260 nodes, 2 Single/Dual core AMD Opteron
    - Myrinet/Gigabit Ethernet and Surfnet
    - SGE, then ?
- Grid’5000
  - Heterogeneous
  - 9 sites, 15 clusters
    - >2000 nodes
    - RENATER
    - OAR

**Inter-Operating Grids**
- More computing power
- Handling demand surges
- Good research environment
Inter-Connecting Grids: Topology (1/4)

- Grids not sharing load
- User selects where to submit jobs
- Q: Can the user efficiently select resources? What is the administrative overhead for multiple accounts?

Inter-Connecting Grids: Topology (2/4)

- Inter-connect any grid sites
- Towards full-mesh topology
- Q: Too many redundant paths? Routing loops?

Inter-Connecting Grids: Topology (3/4)

- Create grid root, which routes between grids
- Fully hierarchical topology

Inter-Connecting Grids: Topology (4/4)

- Inter-connect existing grid roots
- Not fully hierarchical topology
Inter-Connecting Grids: Mechanisms (1/3)

- Independent Koala schedulers, no load sharing
- Two choices:
  - Let the user select the grid
  - Assist the user’s grid selection (observational scheduling)

Inter-Connecting Grids: Mechanisms (2/3)

- Policy: use remote grid only if local grid saturated
- Operation choices:
  - Koala operates on top of local resource managers (OAR)
  - Koala deploys its own environment, by-passing OAR

Inter-Connecting Grids: Mechanisms (3/3)

- Policy: use remote grid only if local grid saturated
- Dynamic environment build-up
- Routing choices:
  - To parent, to children, to siblings

The Delegated MatchMaking Mechanism

- Delegate resources, not jobs
- Dynamic environment build-up
- Peer-to-Peer exchange (negotiation?)

With M. Livny, T. Tannenbaum, M. Farrelllee (U. Wisc-Madison)
Inter-Connecting Grids: Mechanisms (3/3)

Why should DMM work?

- Overall workload imbalance: normalized daily load (5:1)
- Temporary workload imbalance: hourly load (1000:1)

Inter-Connecting Grids: Mechanisms (3/3)

Experimental Results: Performance

- High goodput
- Low wait time
- Finishes all jobs
- Even better for load imbalance between grids

Inter-Connecting Grids: Mechanisms (3/3)

Experimental Results: Overhead

- DMM
  - Overhead ~16%
  - 93% more control messages
  - Constant number of delegations per job until 80% load
  - DMM Threshold to control overhead.

Conclusion

Inter-operating DAS-2, DAS-3, and Grid’5000

- First steps accomplished
  - Koala (inter-)operates DAS-2 and DAS-3
  - GRAM interface to OAR
  - Also… analysis of Grid’5000 usage patterns
- On-going work
  - Koala operating over OAR
  - “DAS-2” image on Grid’5000: Globus, KOALA, OAR
  - DMM implementation and experimental validation
- Future work
  - Experiments
  - DMM research, SLAs
  - Virtualization, more application types
Information

• Publications
  • see PDS publication database at www.pds.ewi.tudelft.nl

• Web site
  • KOALA: www.st.ewi.tudelft.nl/koala