INTELLIGENT PARALLEL JOB SCHEDULERS FOR COMPUTATIONAL MULTI-CLUSTERS

Multi-cluster Paradigm
- Computational resource
- Geographic co-location
- Dedicated network
- Multiple participants
- Shared resources
- Distributed processing
- Job scheduling

Co-allocation Model
- Required link bandwidth
  \[ BW_i = \left( \frac{1}{n_i} + \frac{1}{n_j} \right) (P_i \cdot BW_i) + w_j \]
- Bandwidth saturation
  \[ BW_{sat} = BW_{max} \cdot \sum_{i,j} BW_i \]
- Communication slowdown
  \[ BW_{com} = BW_{max} \cdot \sum_i BW_i \]

Parallel Job Scheduling
- Resource allocation
  - Where to run job?
  - When to run job?
- Local execution
- Job migration
- Move entire job
- Job co-allocation
- Map across clusters
- Share resources
- Network contention

Execution-time Dynamics
- State changing events
  - New co-allocated job
  - Co-allocated job terminates
  - Variation w.r.t. time
  \[ T = Q + B \]

Intelligent Scheduling
- Global optimizations that improve job response time
- Efficiently manages both node and network resources
- Controls network contention during job co-allocation

Scheduler Model
- Scheduler decision
  - Bandwidth-aware co-allocation

Scheduling Effectiveness

Improved Performance

Future Work
- Verify existing models and update if necessary
- Add additional parallel job types and topologies
- Enhance workload generation with trace-files
- Integrate configuration GUI into visualizer

Acknowledgements: This work was supported by the ERC program of the National Science Foundation under Award Number EEC-9731680.