Application Execution using Hybrid Resources

Mentor: Matthew Woitaszek
Zhifeng Yun
Louisiana State University
July 30, 2010
Outline

• Introduction
• Objective
• DA-TC Model
• Cloud resources
• Summary and Future Work
Daymet

- Science: Terrestrial ecosystem model
  - Daymet component: surface observation interpolation
  - Originally built on a private NCAR grid (2005-2007)
  - Status: Maintenance (2007-present)
  - UI: Web portal (Java)
  - Back-end: GT 4.0 Grid Services, GRAM; short single tasks

- Current work
  - Evaluating new execution strategies for this type of “many-task” workflow
The Daymet application

- composed of many tasks
- coordinated by single-threaded Perl Script
- each task has variable length
- can be individually scheduled on different computing resources across multiple administrative boundaries to achieve short execution time
Challenges

- Require excessive resource-specific knowledge and software development
  - Independent administration
  - Varying performance and characteristics
- Main issues
  - Heterogeneity of computational performance
  - Fluctuating resource availability
  - Bottleneck of slow resources
  - Long queuing time
Objective

Use the LONI DA-TC Execution Model for multi-site runs:

• Reduced turnaround time
• Reliability
• Execution steering

This work emphasizes using a number of computing resources over short periods of time to execute computationally demanding applications that are composed of a large number of tasks
DA-TC Execution Model

- A model for efficiently executing an ensemble of jobs without the necessity to queue each individual job
  - Use centralized queue to accommodate users’ jobs
  - Submit Task containers (TCs) instead of real jobs to the local resource management system (batch queue)
  - Provide a light-weight hosting environment once the TC is assigned resources
  - Send centrally queued workload to first available resources
- DA-TC decouples resource allocation from resource binding
- Support multiple ways of Task Container scheduling
Runtime Scenario

Task Queue

- Task N
- Task 21
- Task 20
- Task 19

Queue On Cluster 1
- TC0: R
- TC1: R
- TC2: R
- Other: R
- Other: Q

Queue On Cluster 2
- Other: R
- Other: R
- TC3: Q
- TC4: Q
- TC5: Q
- Other: Q
- Other: Q

Queue On Cluster 3
- Other: R
- TC6: R
- TC7: R
- TC8: Q
- TC9: Q
- Other: Q
- Other: Q

AEA
DA-TC Advantages

• A user-space execution model
• Dynamic load balancing
  ‣ Fast clusters will be assigned more tasks
  ‣ Slow resources: beneficial factors, not bottleneck any more
• Significant reduction of turnaround time
• Enhanced reliability of application execution
• Easily applied to Daymet model workflow
• Result:
  ‣ Run the Daymet tasks on two 8-core systems and one external cluster
  ‣ Finish all 800 tasks in 3 hours with 11 TCs
Cloud Computing

- Give users access to compute/data resources that they do not own

- Dynamic provision of services/resource pools in a co-ordinated fashion
  - On demand computing – No waiting period
  - Location of resource is irrelevant
    - May be relevant from performance (network latency) perspective, data locality

- Web interfaces
Clouds vs Grids

Grid

- Cluster of loosely coupled, networked computers to perform very large tasks
- Resources are shared by large number of users
- Jobs are submitted to the scheduler
- Lacks automation, agility, simplicity and SLA guarantees

Cloud

- Each VM is dedicated to one user
- Jobs are assigned to a VM by the user
- Physical resources are provisioned to VM instead of workload
Clouds

• Commercial Clouds

• Science Cloud Software
  ‣ Nimbus (Keahey and Freeman, Argonne & Univ. of Chicago)
  ‣ Eucalyptus (Wolski, UCSB)
  ‣ Both are
    - Cloud Computing on Clusters
    - Amazon Web Services compatible
    - Supports KVM and Xen
Nimbus

• Cloud service
  ‣ Initiate and terminate IaaS
  ‣ Cumulus storage service

• Cloud client
  ‣ Launch, query, terminate workspaces

• Workspace service
  ‣ Hardware implementation and virtualization

• Dynamically provision resources and environment
  ‣ Publish information of each workspace

• Web service

• Rely on
  ‣ Xen or KVM
  ‣ Libvirt
  ‣ GT4
  ‣ DHCP
Nimbus Implementation

- Deploy VMs
- Deploy Virtual Clusters
Issues Encountered

• Surprisingly more complex than anticipated

• Lots of debug effort
  ‣ Client VM and Host VMM incompatibility
  ‣ Image unportability
  ‣ Insufficient documentation

• Needs extensive knowledge about
  ‣ System administration
  ‣ Networking
  ‣ Virtual machine
Integrating with DA-TC

• Each VM can be viewed as a TC

• Difference
  ‣ Each VM needs a copy of binary executable and grids file
  ‣ Each VM needs specify running time
  ‣ VM scheduling

• Using virtual cluster as one of the multiple resources
  ‣ Fit DA-TC paradigm
  ‣ One-click cluster -- easy to boot
    ◌ Need configuration
  ‣ Once booted, dedicated resources -- no queuing time
Integration

Cluster 1

Cluster 2

Virtual Cluster

Waiting Queues
Summary and Future Work

- Implemented DA-TC model for Daymet to achieve reduced turnaround time
- Deployed Nimbus cloud service on NCAR evaluation clusters
  - Zero to Cloud
  - One-Click Clusters
- Future work
  - Learn the scalability of configured image on different OS
  - Learn the scheduling and load balancing of VM deployment