Simulating the Grid

Robert Samborski

Robert Walter

Seminar "Grid Computing 2"
(SE 2.0, 703822)
WS 2006/07
Presentation Overview

- Introduction
  - Overview & motivation
- Simulation basics
  - Definition & techniques
    - Emulation, discrete events, …
- Simulation of the Grid
  - Principles, problems & trends
- Simulation tools
  - Some of a large number of existing tools divided into simulators and emulators
Motivation for Simulation

- Grids are complex
- Testbeds not always available
- Speed up research
- High costs
- Specific scenarios
Simulation

- Definition:

“A simulation is an imitation of some real thing, state of affairs, or process. The act of simulating something generally entails representing certain key characteristics or behaviors of a selected physical or abstract system.”

(from Wikipedia)
Simulation techniques basics

- Mathematical simulation
  - Mathematical formulas
  - Abstract
  - Fast
Simulation techniques basics

- Emulation
  - code execution on virtualized resources
  - Real/simulated resources
  - Real/accelerated system time
  - Real life behavior (?) = closer to reality
  - Real life problems
Simulation techniques basics

- “Classical” simulation
  - Discrete event simulation
    - Time steps (granularity)
    - Evaluation on event
    - No idle time
Simulation techniques basics

- Agent driven simulation
  - Widely used computer simulation technique
    - Biotechnology, AI, …
  - Special case of discrete event simulation
    - Resources modeled as objects
    - Each object “knows” it’s behavior
Grid Simulators

- Reuse of well-known principles
- Implementation on top of existing tools
  - SimJava, Ptolemy, NS
- Representation of the Grid
  - Implicit
  - Tables, Objects
  - Structures
    - Tree, Graph
Problems of Grid Simulations

- Emulations
  - Typical test bed problems
  - Really reproducible results?

- Abstract simulator problems
  - Matching the reality?

- Using the results
  - Verification
  - Validation
Observations

- Last 5 years
  - Scheduling, replication

- Trends
  - More powerful Grids & test beds
  - Security issues gain importance

- Latest simulator development
  - Lightweight & as abstract as possible
    - CATNETS, animat agents, …
Grid Simulation Tools

- Emulators
  - MicroGrid, PlanetLab, EmuLab
  - Specialized Grid test beds

- Simulators
  - Bricks, ChicagoSim, GangSim, GridNet, OptorSim, SimGrid and GridSim
Grid Emulators

- MicroGrid
  - Project to provide virtual grid infrastructure
    - Controlled, repeatable experiments
  - Emulates Globus Grids only
    - Globus components on virtual hosts
  - Network simulated
    - Discrete events on packet level
  - Continuous time, virtual time
Grid Emulators

- PlanetLab

- 722 nodes @ 349 sites on Dec. 11 2006
- Open platform
- Real "real life" problems
Grid Emulators

- EmuLab
  - Same script language as NS simulator
    - Maybe not so hard to configure
  - Over 200 nodes, most are PCs
  - Front-end to PlanetLab
Beyond Grid Emulators

- Real life research platforms
  - Explicitly built for Grid research
    - No production cluster
  - DAS (Distributed ASCI Supercomputer)
    - DAS1: pioneer research project
    - DAS2: 5 homogeneous clusters with 200 nodes
    - DAS3: on the way
  - Grid’5000
    - Can be configured as an emulator
    - Large scale infrastructure to study under real conditions
Grid Simulators

- Bricks
  - Discrete event simulation
  - Written in JAVA
  - Purpose
    - Resource allocation strategies
    - Scheduling
    - Integration of real components
  - Associated with NWS
Grid Simulators

- **ChicagoSim**
  - Discrete event simulation
    - CPU, network, apps
  - Built on top of Parsec
    - C-based language
      - Parallel architectures, VLSI
  - Behavior of an “entire” Grid as events
    - Plug in your algorithms
Grid Simulators

- GangSim
  - Discrete event simulation
    - Periodical evaluation of all components
  - Associated with Ganglia
  - High scalability
    - 100’000’s PCs & storage systems
  - Models sites & VOs with different policies
Grid Simulators

- **GridNet**
  - Built on top of NS
    - Event driven
    - Data exchange as stream of packets
  - Written in C++
  - Purpose
    - Evaluation of scalable replication topologies
Grid Simulators

- OptorSim
  - Discrete event simulator
  - Associated with European Data Grid (EDG)
    - ITC-first, Univ.Glasgow, CERN
  - Architecture based on EDG model
    - CEs, resource brokers, routers
  - Allows different network topologies & file access patterns
Grid Simulators

- **SimGrid**
  - Toolkit with core functionalities
    - Distributed applications
    - Heterogeneous distributed environment
  - Discrete event simulator
    - Uses a lot of mathematical simulation
  - Configurable topology
  - Resources & tasks in high detail
  - SimBOINC built on top of SimGrid
Grid Simulators

■ GridSim (briefly)
  ■ Discrete event simulator
  ■ Built on SimJava
    ■ Core classes for discrete events
    ■ Distributed systems, networks, protocols, architectures
  ■ Large scale resources simulation
  ■ Economy based scheduling
    ■ Multiple concurrent submission, resource booking
  ■ Until recently missing explicit Grid topology
GridSim by example

- Part II of the presentation
  - SimJava background
  - GridSim in detail
  - Source code explanation
  - Running examples
- Presentation of part II
  - Robert Walter