Genias Global Resource Director

Customer Scenarios for Large Multiprocessor Environments

F. Ferstl, GENIAS Software GmbH - http://www.genias.de
Customer Scenarios

Content Overview

- GRD - What is it?
- How GRD helps BMW
- How GRD helps ARL
- Conclusions
Global Resource Director
GRD
Unique Capabilities

- Share-based control (vs priority)
- Global resource model
- Dynamic control
- Usage-based management
- Deadline capabilities
- Modular approach
- Operational controls
No reprioritization after dispatching
- Resource utilization by individual processes deviates from intent
- Doesn’t adapt to changing needs
- Limits execution control policies
GRD with Global Dynamic Scheduler

Maintains active low level control of workload during execution
- Supports multiple policies (functional priority, share-based, urgency)
- Correlates all workload elements
- Keeps resource utilization aligned with policies
- Responds to ad hoc needs

Dynamic process control
Usage/Load data collection
Dynamic scheduler
Enhanced Codine job management

Budgets & Policies
Share redistribution
Dispatching
Policy enforcement
Usage
Resource specific load sensing
Process specific utilization
GRD Policy Capabilities

**Functional Priority**
- User
- Department
- Job Class
- Project

**Share-based Usage**
- User share tree
- Project share tree

**Urgency-based Priority**
- Initiate Time
- Deadline

**Override System**
- User
- Department
- Job Class
- Job
- Project

**Global Dynamic Scheduler**
- Correlates & tracks workload elements
- Dynamically enforces policies
- Manages global resource utilization
Technical Approach

GRD

Develop/integrate advanced scheduler (GDS)

- Multiple policies (Functional priority, share-based, urgency-based)
- Highly customizable
- Automated policy enforcement
- Global dynamic resource management

Start with proven foundation (Codine)

- Defacto standard in Europe
- Independently rated as top tier JMS
- Multiple job types (scalar, parallel, interactive, etc.)
- User access controls
- Load balancing
- Checkpointing & automigration
- X-window/motif-based GUI

Integrate low overhead dynamic data collector

- www client
- Job dependencies
- Multiple clusters
- Extensible load sensors
- POSIX 1003.2d
- NQS command support
- Interoperation with other JMSs
BMW – CODINE/GRD installation

Crash simulations with PAM-CRASH
Department EK-20, Dr. Holzner
SGI multiprocessor environment with 350+ CPUs
BMW – Planning

Situation:
- No information on total utilization
- No capacity planning

CODINE:
- Full utilization of workstations
- Facilitates hardware upgrades

Result:
- CODINE made growth possible:
  1994: 32 CPUs
  1998: 350+ CPUs
- Million DM savings
- 83% utilization through CODINE jobs

CUG/10
BMW – End–User

**Situation:**
- Crash simulations with PAM-CRASH
- Number of computations has doubled since 1994
- Size of 3D models has increased by factor 3

**CODINE:**
- Short turnaround times
- Better computational performance by optimized scheduling
- Standardization, complexity is reduced/avoided

**Result:**
- Productivity boost
- End–user can focus on engineering tasks
BMW – System Administration

Situation:
- Scheduling by phone calls
- Number of users increased from 3 to > 45
- Growth of computing environment

CODINE:
- Automatization
- Global system view and control

Result:
- Reduced time needed for administrative tasks
- Increase in availability
BMW – Application Services

Situation/Requirements:
- grant resource shares for users
- jobs with different priorities
- several applications with different usage profiles

GRD:
- granted resource shares
- dynamic control of resource usage
- automated response to priority work

Result:
(Planned)
- high quality supercomputing services
- shorter product cycles through integration
- increase in utilization to over 90%
GRD at Army Research Laboratory (ARL)
Aberdeen, MD, USA
ARL Configuration
Unclassified System

- 1 Cray T90 8 IEEE processors
- 4 x 64 processor Origin 2000
- 1 x 32 processor Origin 2000

User Profile
- Approx. 500 engineers, developers & researchers from ARL and other DoD agencies
- Parallel jobs (mainly MPI) on Origins
- Defined resource share entitlements for a range of departments (Army, Navy, Air Force, ...)

CUG/15
ARL Resource Management Goals

- Fair distribution of resources over a sliding time window
- Short term over-commitment of resources for the price of later compensation
- Express & dead-line jobs
- Automatic enforcement of policies
- Manual override capabilities
Denice Brown - Mgr. of Operations & Customer Services

The user can be assured of getting his or her job run in the most fair and efficient manner according to set sharing policies;

... the system administrator has the benefit of having the resource allocation done automatically as well as being able to monitor resources at both the site and individual job level.”

Previously, resources were degraded for everyone when a few users “overused” the system.
Conclusions

- GRD targeted for
  - HPC centers
  - Enterprise or large department computing facilities
  - Application/Computing service providers
- Unique capabilities:
  - Support for heterogeneous environments
  - Utilization policies across an enterprise computing environment
  - Flexible administration
- GRD helps to
  - Gain overview on resource utilization profile
  - Distribute resources fairly
  - Implement level-of-service agreement