GWD-I
Category: INFORMATIONAL
Grid Scheduling Dictionary WG (SD-WG)
http://www.fz-juelich.de/zam/RD/coop/ggf/sd-wg.html
http://www.fz-juelich.de/zam/RD/coop/ggf/draft.pdf

M. Roehrig, Sandia National Laboratories
Wolfgang Ziegler, Fraunhofer-Institute for Algorithms and Scientific Computing
Philipp Wieder, Research Centre Jülich

June-2002
Revised 02-Oct-2002
Revised 30-Oct-2002

Grid Scheduling Dictionary of Terms and Keywords

Status of This Memo

This memo provides information to the Grid community regarding a dictionary for Grid Scheduling. Distribution is unlimited.

Copyright Notice

Abstract

The GGF Scheduling and Resource Management Area is concerned with various issues relating to resource scheduling and resource management in Grid environments. In examining current scheduling systems, it became apparent that a dictionary of terms and keywords would assist in discussions. This information could then be used to develop a Grid language that would provide seamless communication and advance reservation capabilities to the various schedulers in use within a Grid.

Contents

mroehr@sandia.gov
wolfgang.ziegler@scai.fraunhofer.de
ph.wieder@fz-juelich.de

1. General Remarks

The dictionary is focusing on scheduling in a Grid environment, it does not address scheduling terms in general. The terms are used in the manner typically found in Grid schedulers. The authors are aware that there may be multiple definitions of some terms depending e.g. on the environment where a term is used.
A context is given for each term indicating where a term usually is employed. If a term appears in several contexts, the most relevant ones are listed together with the term.
The following contexts are used:

- Access
- Application
- Network
- Process
- Resource
- Resource Management
- Scheduler

2. List of Terms and Definitions

Term	Context	Definition
Account	Access	The permissions for a particular user to use various resources.
Authentication	Access	The process by which one entity verifies the identity of another.
User	Access	A person authorized to submit jobs to *High Performance Computing* resources.
Application Software	Application	The actual programs used by an application.
Checkpoint	Application	Saving the current state of a program and its data (including intermediate results) to disk or other non-volatile storage, so that if interrupted the program could be restarted at the point at which the last Checkpoint occurred. This is a feature provided by some schedulers, but not all.
Fork	Application, Resource Management, Scheduler	Making a copy of a process for execution.
Preemption	Application	Preemption of individual threads or the entire job during the execution of the job (and potential relocation).
Re-runable	Application	If a *Batch Job* can be terminated and its execution restarted from the beginning without harmful side effects, then the job is said to be Re-runable.
Bandwidth	Network	Bandwidth is a property of the network hardware (the transmission medium), higher Bandwidth allows theoretically more data to be transmitted per time unit (unit is bit per second). The achievable *Data Rate* depends on and is limited by the Bandwidth. *QoS* parameter.
Data Rate	Network	The Data Rate defines the amount of data which can be transmitted per time unit via a transmission medium from one application to another (unit is bit per second). *QoS* parameter.
Delay	Network	A Delay or latency is the time interval between sending and receiving a signal (unit is second; the ITU recommends a delay < 100ms in WANs). *QoS* parameter.
IP Address	Network,	Internet Protocol Address. Every system within a network using

mroehr@sandia.gov
wolfgang.ziegler@scai.fraunhofer.de
ph.wieder@fz-juelich.de

Term	Context	Definition
	Resource	TCP/IP (Transmission Control Protocol/Internet Protocol, also called Internet protocol family) has an unambiguous IP address assigned.
Jitter	Network	Phase distortion, changing the signal frequency. Jitter occurs especially at high frequencies and may result in loss of data. *QoS* parameter.
LDAP	Network	Lightweight Directory Access Protocol.
Packet Size	Network	Size of a data packet (unit is bit). A data packet is a well defined order of characters which are treated as a unit. *QoS* parameter.
Port Number	Network, Resource	Port Numbers are used for data transfer between the transport protocol layer and the application process. Also ranges of port numbers can be scheduled.
Protocol	Network	A protocol is a complete and unambiguous set of rules (formats, their semantics & syntax, parameters, timing, error handling, ...) defining the communication between two or more entities
QoS	Network, Resource	Quality of Service (QoS) refers to the performance attributes of an end-to-end connection (like a TCP/IP network connection), Covered QoS parameters are *Bandwidth*, *Data Rate*, *Delay*, *Jitter* and *Packet Size*.
Application	Process	A combination of program tasks and data manipulations that make use of computing resources.
Batch Job	Process	A Batch Job is submitted to a *Batch System* and is processed without any further interaction from the *User* (except for status changes like *Cancel*, *Hold*, ...)
Cancel	Process	Or *Delete*. Executing this action on a job results in the job being removed from the system. On a *Batch System* this is analogous to removing the job from the *Queue*.
Collaborative Computing	Process	To work jointly with other computing devices.
Delete	Process	Or *Cancel*. Executing this action on a job results in the job being removed from the system. On a *Batch System* this is equivalent to removing the job from the *Queue*.
Hold	Process	If a job is waiting in a queue for execution the action Hold prevents the job from being scheduled. Performing a *Release* allows the held job to be scheduled again. *PBS* also enables holding running jobs which is equivalent to *Suspend*.
Interactive Job	Process	A *Job* whose input and output are interleaved, allowing the user's input to depend on earlier output from the same run.
Job	Process	An application or task performed on *High Performance Computer* resources. A Job may be composed of steps/sections as individual schedulable entities.
Release	Process	Allows a held job to be scheduled again (see *Hold*).
Resume	Process	A Resume action can be performed on a *Suspended* job. The job execution is restarted.
Suspend	Process	The current job status is stored, resources are released and the execution of the job is interrupted. A Suspended job can be *Resumed*.
Task	Process	A specific piece of work required to be done as part of a job or application.
Cluster	Resource	A set of execution "servers" or hosts on which a single *Batch Server* manages *Batch Jobs*. A cluster may be made up of a set of workstations, multiple *CPU* systems, or a set of nodes in a

mroehr@sandia.gov
wolfgang.ziegler@scai.fraunhofer.de
ph.wieder@fz-juelich.de

Term	Context	Definition
		parallel system.
Computational Grid	Resource	Large-scale high performance distributed computing environments that provide access to high-end computational resources.
CPU	Resource	Central Processing Unit - the part of a computer system that operates on information or data. Some schedulers require a specification of how many CPUs are available on a node.
Disk	Resource	A non-volatile device that is used to store information and that provides read/write access.
HPC	Resource	High Performance Computing - The technology that is used to provide solutions to problems that: • require significant computational power • either need to access, or process, very large amounts of data quickly • need to operate interactively across a geographically distributed network
HPSS	Resource	High Performance Storage System.
Load	Resource	The quantity of resources required for a job to run.
Memory	Resource	The capacity for storing information during the lifetime of a program.
Name Space	Resource, Resource Management	Name Spaces provide (locally) unique names for (a hierarchy of) objects (e.g. resources, services, applications). The information is often organized in a tree structure and accessible through directory services. A global Name Space aims to unify multiple or conflicting naming schemes allowing to use the same name across an entire *Grid*.
Node	Resource	A point of connection on a network. Some schedulers require a job to specify how many Nodes it requires. A Node consists of one or more Central Processing Units *(CPUs*). Each processor may have multiple threads running on it that share code and data but have different stacks.
Processor	Resource	The part of a computer which performs the manipulation of data from one state to another.
Queue	Resource	A collection of schedulable entities, e.g. jobs (or job-related tasks) within the (batch) queuing system. Each Queue has a set of associated attributes that determine which actions are to be performed upon each job within the Queue. Typical attributes include queue name, queue priority, resource limits, destination(s), and job count limits. Selection and scheduling of jobs are implementation-defined. The use of the term "queue" does not imply the ordering is "first in, first out."
Visualization	Resource	A device that displays data in picture form. The most basic Visualization is that of turning transaction data and summary information into charts and graphs.
Advance Reservation	Resource Management	Advance reservation is the process of negotiating the (possibly limited or restricted) delegation of particular resource capabilities over a defined time interval from the resource owner to the requester.
Allocation	Resource Management	The quantity of resources required by a job or reserved for a job.
Broker	Resource Management	A process which performs resource quoting (producer) or resource discovery (consumer) and selection based on various strategies, assigns application task(s) to those resources, and distributes data or co-locates data and computations. *Cost*

Term	Context	Definition
		Models may be used for negotiations before selecting/requesting resources.
Co-Allocation	Resource Management	Ensures that a given set of resources is available for use simultaneously.
Condor	Resource Management	A computing environment that makes use of and manages very large collections of distributed, private workstations.
Condor-G	Resource Management	A version of *Condor* that works with *Globus*.
Co-Scheduling	Resource Management	The ability to schedule various resources for a given time or for a series of sequential times (supported by *Co-Allocation)*.
Cost Model	Resource Management	Defines the cost-benefit ratio of a resource, determines supply and demand. Cost Models exist for both producers and consumers.
CPU Time	Resource Management	The amount of time a process takes to run assuming that it has exclusive and uninterrupted use of the *CPU*. In praxis this means adding up all the small amounts of time the CPU actually works on the process.
Data Co-Location	Resource Management	For task-parallel programs, computations or data may reside in distinct locations, and the scheduler must determine which need to be moved.
Data Distribution	Resource Management	For data-parallel programs, all computational resources execute the same program, and the complexity of the scheduling process lies in the determination of a performance-efficient distribution or decomposition of data.
Duration	Resource Management	A specification of how long a job is expected to run.
End Time	Resource Management	A specification of the time when a job is expected to quit running or actually quit running. May be specified as an actual time or as an offset from the current system time.
Fair Share	Resource Management	A *Scheduling Policy* that ensures each user or group of users receive a configurable share of (computing) resources over a configurable period of time called "history hours". The priority of each job is based on a dynamic calculation based on weighted usage of one or more resources during the history hours for the user/group. A decay factor is applied so recently used time weights more heavily than the time used in the distant past.
Globus Project	Resource Management	Software tools that make it easier to build computational *Grids* and grid-based applications.
GLUnix	Resource Management, Scheduler	GLUnix is a scheduling and resource management system. It supports *Batch* and *Interactive Jobs* and performs execution time *Load Balancing*.
GRD	Resource Management, Scheduler	The Global Resource Director is a scheduling and resource management system.
Grid	Resource Management	Grids are persistent environments that enable software applications to integrate instruments, displays, computational and information resources that are managed by diverse organizations in widespread locations.
Legion	Resource Management	Builds system components on a distributed object-oriented model and provides a uniform programming model.
Migration	Resource Management	Migration describes the rearrangement of allocated resources within a resource pool.
Pexec	Resource	Pexec is an executable and a support library for simplifying the

mroehr@sandia.gov
wolfgang.ziegler@scai.fraunhofer.de
ph.wieder@fz-juelich.de

Term	Context	Definition
	Management	implementation and execution of SPMD programs in a Beowulf environment. It is build on the top of the Parallel Virtual Machine (PVM).
POE	Resource Management	Parallel Operation Environment – POE manages the execution of parallel applications across multiple nodes including the communication between the nodes.
Prun	Resource Management	Prun is used to launch parallel processes on a Cluster, e.g. MPI applications.
Reservation	Resource Management	The act of specifying a resource, a *Start Time* and an *End Time*.
Resource Discovery	Resource Management	The process of locating a set of resources on which to schedule the task(s) of an application.
Resource Quotes	Resource Management	Bidding of resources together with costs. May be mediated through a *Broker*.
Resource Selection	Resource Management	The process of selecting candidate resources from a pool.
Start Time	Resource Management	A specification of the time when a job is expected to run or actually began running. May be specified as an actual time or as an offset from the current system time.
Time	Resource Management	Usually refers to execution time of a job.
UNICORE	Resource Management	UNICORE (UNiform Interface to COmputing REsources) provides a science and engineering *Grid* combining resources of supercomputer centers and making them available through the Internet.
Usage	Resource Management	A measurement of the compute resources accessed by a user. This is typically a combination of *Nodes* and execution time.
Wallclock Time	Resource Management	The elapsed time between when a process starts to run and when it is finished. The Wallclock Time is normally longer than the *CPU Time* due to time-sharing *CPUs* and I/O.
Application Schedule	Scheduler	An assignment of tasks, data, and communication to resources, ordered in time – based on the rules of the *Scheduling Policy*, and evaluated as “performance efficient” under the criteria established by the performance model.
Batch	Scheduler	A group of jobs (as programs) which are submitted for processing on a computer and the results of which are obtained at a later time.
Batch Processing	Scheduler	The capability of running jobs outside the interactive login session and providing for additional control over job scheduling and resource contention.

mroehr@sandia.gov
wolfgang.ziegler@scai.fraunhofer.de
ph.wieder@fz-juelich.de

Term	Context	Definition
Batch Queue	Scheduler, Resource	An execution queue where the request actually is started from/runs.
Batch Server	Scheduler, Resource	A persistent subsystem (daemon) upon a single host that provides *Batch Processing* capability.
Batch System	Scheduler, Resource	A set of *Batch Servers* that are configured for processing. The system may consist of multiple hosts, each with multiple servers.
EASY	Scheduler	EASY (Extensible Argonne Scheduling System) was originally written to enable parallel jobs to be scheduled efficiently on IBM SP2 systems, however, various versions supporting other systems are now available.
EASY-LL	Scheduler	EASY-LL is an API that can combine *LoadLeveler* and *EASY*.
Gang Scheduling	Scheduler	The gang scheduler permits a set of processes, or multiple threads from a single process, to be scheduled concurrently as a group.
Grid Engine	Scheduler	A scheduling system that accepts job submission by users and schedules them for execution on appropriate systems in the *Grid* based on resource management policies.
High-Performance Scheduler	Scheduler, Resource Management	A software system that uses scheduling models to predict performance, determine application schedules based on these models, and take action to implement the resulting schedule. The goal is to optimize the performance experienced by the application on computational grid. The result is an *Application Schedule*.
Job Scheduling	Scheduler	Establishing a job queue to run a sequence of jobs (programs) over any period of time
Load Balancing	Scheduler	The process of maintaining balanced workloads across multiple *CPUs* or systems.
LoadLeveler	Scheduler	A workload management and scheduling system for serial and parallel jobs over a cluster of servers.
LSF	Scheduler	Load Sharing Facility – load sharing and sophisticated batch scheduling across distributed UNIX and Windows NT computing environments.
Machine Scheduler	Scheduler	A method used to schedule jobs for execution on a machine. Priority, length of time in the job queue and available resources are examples of criteria used.
Mapping	Scheduler, Resource Management	The allocation of computation and data “in space”.
Meta-Scheduler	Scheduler	A scheduler that allows to request resources of more than one machine for a single job. May perform *Load Balancing* of workloads across multiple systems. Each system would then have its own local scheduler to determine how its job queue is processed. Requires *Advance Reservation* capability of local schedulers (see *Super Scheduler*).
NQS	Scheduler	Network Queuing System – allows a user to submit *Batch Jobs* to various *Queues* on local or remote machines, to monitor the job progress, and to have the log file returned to the originating machine or another machine. The *Queues* can be set up according to a variety of parameters such as job size and required resources.
PBS	Scheduler	Portable Batch System – a batch queuing and workload management system.

mroehr@sandia.gov
wolfgang.ziegler@scai.fraunhofer.de
ph.wieder@fz-juelich.de

Term	Context	Definition
Scheduling	Scheduler	The process of ordering tasks on compute resources and ordering communication between tasks. Also, known as the allocation of computation and communication “over time.”
Scheduling Algorithm	Scheduler, Resource Management	A procedure used by a scheduler to determine when a job can run.
Scheduling Model	Scheduler	Consists of a *Scheduling Policy*, a program model, a performance model, and a performance measure.
Scheduling Policy	Scheduler	A set of rules for producing schedules.
Staging	Scheduler, Resource Management	The process of moving a file or files to the host before the *Batch Job* begins execution.
Submit Jobs	Scheduler	The process of placing a job into a *Queue* for execution.
Super Scheduler	Scheduler	The process that will (1) discover available resources for a job, (2) select the appropriate system(s), and (3) submit the job. Each system would then have its own local scheduler to determine how its job queue is processed (see *Meta-Scheduler*).

mroehr@sandia.gov
wolfgang.ziegler@scai.fraunhofer.de
ph.wieder@fz-juelich.de

3. Summary of Terms by Context

Context	Term
Access	Account
Access	Authentication
Access	User
Application	Application Software
Application	Checkpoint
Application	Fork
Application	Preemption
Application	Re-runable
Network	Bandwidth
Network	Data Rate
Network	Delay
Network	IP Address
Network	Jitter
Network	LDAP
Network	Packet Size
Network	Port Number
Network	Protocol
Network	QoS
Process	Application
Process	Batch Job
Process	Cancel
Process	Collaborative Computing
Process	Delete
Process	Hold
Process	Interactive Job
Process	Job
Process	Release
Process	Resume
Process	Suspend
Process	Task
Resource	Batch Queue
Resource	Batch Server
Resource	Batch System
Resource	Cluster
Resource	Computational Grid
Resource	CPU
Resource	Disk
Resource	HPC
Resource	HPSS
Resource	IP Address
Resource	Load
Resource	Memory
Resource	Name Space
Resource	Node

mroehr@sandia.gov
wolfgang.ziegler@scai.fraunhofer.de
ph.wieder@fz-juelich.de

Context	Term
Resource	Port Number
Resource	Protocol
Resource	Processor
Resource	QoS
Resource	Queue
Resource	Visualization
Resource Management	Advance Reservation
Resource Management	Allocation
Resource Management	Broker
Resource Management	Co-Allocation
Resource Management	Condor
Resource Management	Condor-G
Resource Management	Co-Scheduling
Resource Management	Cost Model
Resource Management	CPU Time
Resource Management	Data Co-Location
Resource Management	Data Distribution
Resource Management	Duration
Resource Management	End Time
Resource Management	Fair Share
Resource Management	Fork
Resource Management	Globus Project
Resource Management	GLUnix
Resource Management	GRD
Resource Management	Grid
Resource Management	High-Performance Scheduler
Resource Management	Legion
Resource Management	Mapping
Resource Management	Migration
Resource Management	Name Space
Resource Management	Pexec
Resource Management	POE
Resource Management	Prun
Resource Management	Reservation
Resource Management	Resource Discovery
Resource Management	Resource Quotes
Resource Management	Resource Selection
Resource Management	Scheduling Algorithm
Resource Management	Staging
Resource Management	Start Time
Resource Management	Time
Resource Management	UNICORE
Resource Management	Usage
Resource Management	Wallclock Time
Scheduler	Application Schedule
Scheduler	Batch
Scheduler	Batch Processing
Scheduler	Batch Queue

mroehr@sandia.gov
wolfgang.ziegler@scai.fraunhofer.de
ph.wieder@fz-juelich.de

Context	Term
Scheduler	Batch Server
Scheduler	Batch System
Scheduler	EASY
Scheduler	EASY-LL
Scheduler	Fork
Scheduler	Gang Scheduling
Scheduler	GLUnix
Scheduler	GRD
Scheduler	Grid Engine
Scheduler	High-Performance Scheduler
Scheduler	Job Scheduling
Scheduler	Load Balancing
Scheduler	LoadLeveler
Scheduler	LSF
Scheduler	Machine Scheduler
Scheduler	Mapping
Scheduler	Meta-Scheduler
Scheduler	NQS
Scheduler	PBS
Scheduler	Scheduling
Scheduler	Scheduling Algorithm
Scheduler	Scheduling Model
Scheduler	Scheduling Policy
Scheduler	Staging
Scheduler	Submit Jobs
Scheduler	Super Scheduler

4. Security Considerations

Security issues are not discussed in this document.

5. Author Information

Mary P. Roehrig
Sandia National Laboratories
mroehr@sandia.gov

Philipp Wieder
Research Centre Jülich GmbH
D – 52425 Jülich
ph.wieder@fz-juelich.de

Wolfgang Ziegler
Fraunhofer Institute for Scientific Computing and Algorithms
D – 53754 Sankt Augustin
Wolfgang.Ziegler@scai.fhg.de

6. Glossary

A glossary is not required for this document.

7. Intellectual Property Statement

The GGF takes no position regarding the validity or scope of any intellectual property or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; neither does it represent that it has made any effort to identify any such rights. Copies of claims of rights made available for publication and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the GGF Secretariat.

The GGF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights which may cover technology that may be required to practice this recommendation. Please address the information to the GGF Executive Director.

8. Full Copyright Notice

mroehr@sandia.gov
wolfgang.ziegler@scai.fraunhofer.de
ph.wieder@fz-juelich.de