

Usage Record Fields -- Survey Results and Proposed Minimum Set

Status of This Memo

This memo provides information to the Grid community in the area of usage records and accounting. It does not define any standards or technical recommendations. Distribution is unlimited.

Copyright Notice

Copyright © Global Grid Forum (2002). All Rights Reserved.

Abstract

This document summarizes the usage record fields used at a sampling of different sites. The purpose of this document is to get an overview of the types of data that are being tracked and used for accounting in existing systems, and to determine a minimum set of usage record fields anticipated to represent the accounting needs for contemporary systems requesting or supplying resources to a GRID. A survey performed by the author collected usage record terms from NASA, NPACI, PNNL, NCSA and ANL. These usage records fields were gathered for resources tracked on vector {Cray C90, SV1, T3E}, parallel/shared memory processors (SMP) {SGI Origin 2000, SGI Origin 3000}, and distributed memory processors (DMP) {IBM SP} systems.

Table of Contents

Abstract	1
Table of Contents.....	1
1. Background.....	2
2. Survey Results.....	2
3. Common Usage Record Fields	10
4. Security Considerations.....	11
5. Author Information	11

6. Intellectual Property Statement.....	11
7. Full Copyright Notice.....	11

1. Background

In order to allow users to execute application on remote computer systems using some mechanism established between sites, sites should be able to provide the resource usage information in commonly defined record format. Currently, each site has it's own accounting record and resource attributes to describe the resources that are being tracked at individual sites. Thus, the common terms to describe a resource attributes must be identified for all resources that are being tracked at sites to exchange resource usage information between sites. To accomplish this goal, survey was conducted to collect all the resources that are being tracked and terms that are used to describe resources at different sites. Based on responses received from sites (NASA, NPACI, PNNL, NCSA, and ANL), resources are grouped together to show the terms used to describe the same resources in following tabularized "Survey Results" section.

Based on survey results, a proposed common set of Usage Record fields are identified with the common terms to describe the resources. This proposed set doesn't limit the extension of the set as the new and/or additional resource attributes are encountered. An individual sites is not required to report all the resource attributes listed in the proposed set but should report resources that are currently being tracked at the individual site. These survey results are intended for use by other groups in the Grid community to reference resources that are currently being tracked and indicate common terms used to describe such resources.

2. Survey Results

Terms used at different sites	Globus Resource Specification Language(RSL)	Datatype	Reference/Description
LOGIN_ NAME, user, userName,		Text/ char	User's login name corresponding to user Id in /etc/ passwd file

username, UserName, USERNAME			
uid, userId		Int	User identification number from the /etc/ passwd file
Type		String	This field indicates the type of transaction being recorded, such as doWithdrawal, doDeposit, doTransfer, modifyAllocation, deleteAccount, etc.
AuthName		String	Authorized userid performing the transaction
ACCOUNT, project, AccountName, projectName, GROUPNAME	project	Text/ char	Users account name where usage will be charged
projid, projectName		int	The account ID
JOB_ ID, jid, jid_ num, jobId, jobid		Number/ int	Job id where job was submitted to the batch queue.
Session_ id		Number	session id from the originating system
Id		char	Identifier indicating the job_ id, session_ id, reservation_ id, quote_ id, allocation_ id, etc according to context

Terms used at different sites	Globus Resource Specificatio n Language(RSL)	Datatype	Reference/ Description
Pid		Number/ int	The process identifier assigned by the operating system during the life of the process.
client, hostname		Text/ char	Name of the system job was executed.
Machine		String	Machine name (This could be a list of machines (systems) for a job which spans clusters and each machine could be a composite name composed of the host, partition, cluster, site, and/ or enterprise)

QUEUE_ID		Text	Queueing system identification code. NQS id, LoadLeveler cluster id or LSF id
QUEUE, qname, queue, Queue, queueName	queue	Text/ char	Queue name where job was executed. (LSF - job was submitted)
QWAIT		Number	Queue wait time for batch jobs
QUEUE_DATE, submitTime, QueueTime		Number/ long	The date the job was queued to the batch system. Number of seconds since the Epoch in GMT.
JOB_QUEUE_DATE		Date	This is the date the job was submitted in the date format
START_DATE, start_time, beginTime, StartTime		Number/ long	The date the job was started by the system. Number of seconds since the Epoch in GMT
JOB_START_DATE		Date	This is the date the job started running in date format.
END_DATE, end_time, Event Time, EndTime, finishTime		Number/ long	The date the job was completed by the system. Number of seconds since the Epoch in GMT
JOB_END_DATE		Date	This is the date the job ended in date format.
REQUESTED_PROCS, ncpus, Processors, limitNpe, numProcessors	count	Number/ int	Number of processors requested at job submission time.

Terms used at different sites	Globus Resource Specification Language(RSL)	Datatype	Reference/ Description
nprocs, peakNpe, maxNumProcesors, MAXPROCS		Number/ int	The number of CPU's used.
MINPROCS		Number/ int	
NODES, Nodes		Number	Cumulative sum of all nodes allocated to the job. (number_of_nodes times cpu_per_node)

num_nodes		int	Number of nodes used. $\max((\text{cputime} + \text{process_per_node} - 1) / \text{process_per_node}, (\text{memory} + \text{mb_per_node} - 1) / \text{mb_per_node})$
Nodemask		char	A hexadecimal string representing the bit mask specifying the nodes (a pair of processors) to be associated with this job.
NodeType		String	Type of node might factor into performance and charge rate
MAXPAR		Number	Maximum node partition. Largest number of processors allocated to parallel applications within the job. On all systems except Cray T3E systems this number will be the same as NODES. On Cray T3E systems multiple parallel applications per job can be run, therefore, MAXPAR will describe the largest number of NODES allocated for the entire job.
Cpupercent		percent	The maximum percentage of a cpu which the job used. A value of 100 means 1 cpu. This cannot be set, it is only reported.
ProconsumptionRate		Number	Percentage of Total CPU used for prorating charge – a decimal number between 0 and 1
CPU_TIME, cput, connect_time, cputime, CPUTime, cpuTime		Number/ long	CPU time used by all processes of job

Terms used at different sites	Globus Resource Specification Language(RSL)	Datatype	Reference/Description
CONNECT_TIME		Number	Connect time for interactive session
Pcput	max_cpu_time	long	Maximum amount of CPU time used by any single process in the job.
user_cpu, ru_utime		long/ double	The user CPU time in seconds
sys_cpu, ru_stime		long/ double	The system CPU time in seconds

interactive_cpu		double	Interactive cpus used (user_cpu + sys_cpu)
Batch_cpu		double	Batch cpus used (user_cpu + sys_cpu)
mt_user_cpu		double	The total user CPU time in seconds in multitasking (MT) queue (user_cpu)
mt_sys_cpu		double	total system CPU time in seconds in MT (sys_cpu)
mt_connect		double	total connect time in MT queue
mt_nconnect		double	The sum of (connect_time * nprocs) in each of the CPU's in MT queue.
mt_non- mt		double	number of seconds which are not multitasking in MT queue. (user_cpu - mt_nconnect)
WALLCLOCK, walltime, Wallclock, runTime	max_wall_time	Number/ long	Wall clock time which elapsed while the job was in running state. For clusters where a node is exclusively allocated the wallclock is multiplied by the number of processors yielding wallclock processor hours. Therefore, on an IBM SP system this is actually the wallclock node hours or "wallclock * number of cpus"

Terms used at different sites	Globus Resource Specification Language(RSL)	Datatype	Reference/Description
REQUESTED_TIME, limitRuntime		Number/ long	Amount of time requested at queue submission time for resource time, either wallclock time for parallel jobs or cpu time for vector/ DMP systems
MAXMEMORY, high_mem		Number /long	Memory high water mark for entire job
MEMORY		Number	Memory usage in Kcore- hours

REQUESTED_ MEM, limitMem		Number	Amount of memory requested at job submission time
Pmem		size	Maximum amount of virtual memory (workingset) used by any single process of the job.
vmem, memory, maxRSwap, mem, Memory, peakMem	max_ memory	size	Maximum amount of virtual memory used by all concurrent processes in the job.
workingset		size	Maximum amount of physical memory used by any single process of the job.
maxRMem		size	Maximum amount of resident memory used by all processes in the job.
NUMMPPJOBS		Number	Number of parallel applications run in this job. On all systems except Cray T3E systems this number will be one. On Cray T3E systems multiple parallel applications per job can be run, therefore, NUMMPPJOBS will describe the number of parallel applications run in this job.
kword_ sec, kword- minutes		double	memory integral in seconds and in minutes(kword_ sec/ 60)
I_ O, Mbytes I/ O, io_ kbytes, IO		Number/ double	IO usage in megabytes or kbytes transferred
IOread		number/ double	total number of bytes read by the job
IOwrite		Number/ double	total number of bytes written by the job

Terms used at different sites	Globus Resource Specification Language(RSL)	Datatype	Reference/ Description
Iobread		Number/ double	total number of bytes read by the job to block devices
Iobwrite		Number/ double	total number of bytes written by the job to block devices.
io_ physreq(Physical I/ O)		double	The number of physical I/ O requests
DISK, Disk		Number	Disk storage used or Disk Charge in units defined by CPU: disk blocks or

			other.
Network		int	Network used (withdrawals) or requested (reservations) by job [could be AVG, TOT, or MAX]
EXPF		Number	Expansion factor. (QWAIT+ WALLCLOCK)/WALLCLOCK This gives whether queue times are proportional to job size
File		Size	The largest size of any single file that may be created by the job.
Fsblkused		long	The number of file system blocks consumed during the job.
Nice		int	The nice value under which the job is to be run.
PRIORITY		Number	Priority weight value.
JOB_ COMP_ STATUS, Status, jStatus		Number	Number representing completion status of the job.
ExitStatus		int	UNIX exit status of the job.
KillReason		Text/ char	if killed, reason the job was killed(npe, mem, cputime, runtime)
command, Executable		char	The name of the executable or system command
APP_ NAME, jobName, JobName	executable	Text/ char	Job or Application name.

Terms used at different sites	Globus Resource Specification Language(RSL)	Datatype	Reference/ Description
Class		String	Class of job (batch, interactive, etc.)
JobType		String	Here you could distinguish between RMS job types, NQS, PBS, LSF, LL, etc.
QOS		String	Quality of Service

Total_charge, Charge		double, float	The total charge of the job in system's billing unit.. Amount debited or credited to account or allocation/ reservation/ quotation
SU		Number	Total charge for this job in System Billing Units. (seconds)

Datatypes

The resource datatype corresponds to the following units.

- Date: Date in human readable format
- Number: specifies the maximum amount in terms or time period, integer, long integer, or double.
- Text: Specifies the character representation of string
- time: specifies a maximum time period the resource can be used. Time is expressed in seconds as an integer, or in the form [[hours:] minutes:] seconds[. milliseconds]
- size: specifies the maximum amount in terms of bytes or words. It is expressed in the form integer[suffix]. The suffix is a multiplier defined in the following table, The size of a word is the word size on the execution host.
 b or w bytes or words.
 kb or kw Kilo (1024) bytes or words.
 mb or mw Mega (1,048,576) bytes or words.
 gb or gw Giga (1,073,741,824) bytes or words.
- unitary: The maximum amount of a resource which is expressed as a simple integer.
- int: specifies the numeric representation in integer
- long: specifies the numeric representation in long integer.
- double: specifies the numeric representation in float.
- char: specifies the character representation.
- String: specifies the character representation.
- percent: specifies the numeric representation in percentage (i. e 0- 100)

3. Common Usage Record Fields

Table 1: Usage Record Fields

Resource Name	Descriptions
Username	User's login name corresponding to user Id in /etc/ passwd file
ProjectName	Users account name where usage will be charged
JobId	Job id where job was submitted to the batch queue for batch jobs. Process Id for the interactive jobs
Queue	Queue name where job was executed or submitted depending on the batch system
GridId	User's global unique Id that identifies the user. Distinguish Name in the user's X509 certificate
FromHost	System name where job was submitted from
execHost	System name where job ran on.
StartTime	The date job started running in date time format (UTC timezone)
EndTime	The date job completed in date time format (UTC timezone)
Processors	Number of processors either used or requested that each center uses for billing purpose..
NumNodes	Number of nodes used.
Cputime	CPU time used, summed over all processes in the job.
Walltime	Wall clock time which elapsed while the job was in the running state.
Memory	Maximum amount of virtual memory used by all concurrent processes in the job.
Disk	Disk storage used or Disk Charge in units defined by CPU: disk blocks or other.
Network	Network used (withdrawals) or requested (reservations) by job [could be AVG, TOT, or MAX]
jobName	Job or Application name
Status	Number representing completion status of the job.
Charge	The total charge of the job in system's allocation unit

4. Security Considerations

NONE

5. Author Information

Mi young Koo
NASA Ames Research Center
mkoo@nas.nasa.gov
650-604-4528

6. 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.

7. Full Copyright Notice

Copyright (C) Global Grid Forum (date). All Rights Reserved.

This document and translations of it may be copied and furnished to others, and derivative works that comment on or otherwise explain it or assist in its implementation may be prepared, copied, published and distributed, in whole or in part, without restriction of any kind, provided that the above copyright notice and this paragraph are included on all such copies and derivative works. However, this document itself may not be modified in any way, such as by removing the copyright notice or references to the GGF or other organizations, except as needed for the purpose of developing Grid Recommendations in which case the procedures for copyrights defined in the GGF Document process must be followed, or as required to translate it into languages other than English.

The limited permissions granted above are perpetual and will not be revoked by the GGF or its successors or assigns.

This document and the information contained herein is provided on an "AS IS" basis and THE GLOBAL GRID FORUM DISCLAIMS ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE."