

Survey about the setup of clusters in OSG

Marco Mambelli, OSG Site Coordination, December 15, 2010

Summary and analysis

This survey is aimed to understand the technologies used by OSG sites at the cluster level, specifically to manage the hosts, the jobs, to replicate system files and to share files. The results will be useful information to share within OSG and help in providing recommendations for new sites, especially Tier 3s. Furthermore it will allow recognizing and contacting experts.

The survey is in 3 parts with similar structure: Cluster management, Local Resource Manager (a.k.a. Queue manager) and Shared file system.

The First question of each part asks how well known is each tool and for personal judgments about performance, ease of use and documentation. System administrators are asked to compare the tools in question with other similar tools that they tested or with their expectations. To help using an uniform scale, some voting guidelines were provided:

- Performance depends on how well the system satisfies your need and on how it compares to others: poor or average if you'd need better, average or good if it satisfies your requirements.
- A First installation/configuration is simple if it works out of the box or requires little documented configuration steps; average if you had to follow long procedures or do little trial and error; it is difficult if it took long time or a lot of guesswork to find a working setup or if you had to modify the code or took a long time to learn the system.
- The Following installation/upgrades are simple if there is an automatic procedure; average if you need some time/effort to re-tune the system; difficult if it is still time consuming or requires new guesswork.
- The Operation is easy if the system simply works; average if it requires little documented periodic tasks; difficult if it requires an expert or requires constant attention or is otherwise time consuming.
- The Documentation is poor if there is little or no documentation or it is obsolete or bad; average if it is mostly there but you'd like more or sometime is confusing; good if it has all that you need and is clear.

A second question in each part asks about a hypothetical involvement in community support within OSG.

There were 32 valid surveys after excluding one empty survey. About one third of the respondents were from major sites, Tier 2s, long-term OSG participants; one third from smaller sites, Tier 3s, resources of smaller or newer VOs; one third from outreach areas, like Brazil, Colombia and South Africa (Johannesburg).

Almost all respondents took between 5 and 10 minutes to answer the survey, between the opening of the first page and the final submission, with few exception under one minute and few other over 30 minutes or one hour.

The most used cluster management tool is Rocks, sometime customized, followed by Puppets, Cobbler and Cfengine. Users are happy with the performance of the tools they use, especially Rocks and Puppets. The difficulty of the first installation is average (sometime long or with some guesswork) to easy (works out of the box); same for the updates. The operation is automatic or requires simple documented tasks. Rocks seem the easiest to operate while Puppet is the easiest to install/update. Available documentation is good for Rocks, average (there could be more or sometime is confusing) for the others.

There are some long time users of Rocks and Cfengine while Puppet gained popularity in recent times.

Other tools suggested, beside homegrown tools, include Platform Cluster Manager (<http://www.platform.com/> and Dell), Chef (<http://www.opscode.com/chef/>) and a NFSRoot/stateless setup.

Condor is the most common LRM (Local Resource Manager), mostly requiring some customization, followed by PBS variations (specially TORQUE/MAUI), mostly used as they are. Users are generally happy with the performance. LRM are generally average to install, average to easy to upgrade and operate. Several users consider Condor first installation difficult (took long time or a lot of guesswork to find a working setup or if users had to modify the code or took a long time to learn the system). The available documentation is average to good: Condor and PBSPro are better documented than the others.

Condor has mostly long time users while PBS adoption is more uniform, from more than 5 years to few months.

One respondent mentioned SLURM (<https://computing.llnl.gov/linux/slurm/>), another LRM supported by Globus. Other notes from respondents: Torque provides better support for parallel jobs (MPI/SMP) and is well supported also by gLite allowing to share worker nodes.

While users tend to focus on one or few tools in the previous categories, there is instead a wider adoption of different file systems. NFS is the most used, followed by dCache, Lustre, xrootd, AFS and Hadoop all close to each other. AFS is the worst as far as performance. NFS has average to poor performance. On the other end of the scale Lustre, Hadoop and GPS are considered performing well, followed by xrootd (good to average). NFS is generally considered easy to install, update and operate. Hadoop, xrootd and GPFS are considered between easy and average to install, update and operate. Lustre ranges from easy to difficult. dCache is the most difficult file system to install and use. NFS, GPFS and Hadoop are generally considered well documented, dCache could be improved and xrootd documentation is sometime obsolete or incomplete.

The respondents are long time users of NFS, AFS and dCache, while Lustre, Hadoop and xrootd have more recent users.

One of the respondents is worried about future support/development for Lustre since Oracle bought SUN. Many other file systems have been suggested: Isilon Systems (<http://www.isilon.com/>), Ceph (<http://ceph.newdream.net/>), ZFS

(<http://www.sun.com/software/solaris/zfs.jsp>), Gluster (<http://www.gluster.org/>) and MogileFS (<http://www.danga.com/mogilefs/>).

People using a tool are generally available in supporting fellow users, mostly via mailing lists. Mailing list support is followed by personal emails, followed by chat participation, workshop participation and document writing. People opting for providing documentation are mostly newer members of the OSG community, not involved in regular OSG meetings and the documentation process (from outreach programs).

In cluster management the most support is available for Rocks, followed by Puppets, followed by Cobbler and Cfengine.

The LRM with most support offers is Condor, especially if looking for something more than email support, followed by TORQUE/MAUI.

In the file systems part most of the respondents offered to help with NFS, followed by Hadoop, Lustre and xrootd close to each other. Anyway at least email support is available for all file systems, probably showing that most site administrators spent effort on more than one.

There are also two questions about the use of native packages. Almost half of the respondents is interested in using only in native packages (13), 12 people will keep using both Pacman and native packages, 1 person will use only Pacman, 4 are indifferent. Out of the native package users, the vast majority (23) is interested in RPMs, 8 in Debian packages, 2 people would like the native packages integrated with a cluster/configuration management system (Rocks).

Survey data

Cluster setup survey

Define your level of knowledge and what you like/dislike in the following cluster management systems. A blank answer is the same as N/A, does not apply (e.g. You can leave blank the answers if you don't know that cluster management system)							
Your level of knowledge							
Answer Options	N/A - Unknown	Read about it	Evaluated it	Use it	Did some customization	Are official developer	Response Count
Bcfg2	12	2	0	0	0	0	14
Cfengine	8	4	2	4	1	0	19
Cobbler	9	2	0	4	0	0	15
Modules	10	1	0	1	1	0	13
Perceus/Warewulf	11	1	0	0	0	0	12
Puppet	8	5	0	5	2	0	20
Quattor	11	2	0	0	1	0	14

Performance

Answer Options	N/A	Good	Average	Poor	Response Count
Bcfg2	3	0	0	0	3
Cfengine	1	2	3	0	6
Cobbler	2	2	2	0	6
Modules	2	2	0	0	4
Perceus/Warewulf	2	0	0	0	2
Puppet	2	4	3	0	9
Quattor	2	0	1	0	3
Rocks	0	11	5	0	16

First installation/configuration

Answer Options	N/A	Easy	Average	Difficult	Response Count
Bcfg2	3	0	0	0	3
Cfengine	1	0	3	2	6
Cobbler	2	3	1	0	6
Modules	2	2	0	0	4
Perceus/Warewulf	2	0	0	0	2
Puppet	3	2	2	1	8
Quattor	2	0	0	1	3
Rocks	1	4	10	1	16

Following installation/update

Answer Options	N/A	Easy	Average	Difficult	Response Count
Bcfg2	3	0	0	0	3
Cfengine	1	1	4	0	6
Cobbler	2	3	0	0	5
Modules	2	2	0	0	4
Perceus/Warewulf	2	0	0	0	2
Puppet	3	3	2	0	8
Quattor	2	0	0	1	3
Rocks	0	6	8	1	15

Operation

Answer Options	N/A	Easy	Average	Difficult	Response Count
Bcfg2	3	0	0	0	3
Cfengine	1	2	3	0	6
Cobbler	2	3	0	0	5
Modules	2	1	1	0	4
Perceus/Warewulf	2	0	0	0	2
Puppet	2	2	4	0	8
Quattor	2	0	0	1	3
Rocks	0	9	6	0	15

Available documentation

Answer Options	N/A	Good	Average	Poor	Response Count
----------------	-----	------	---------	------	----------------

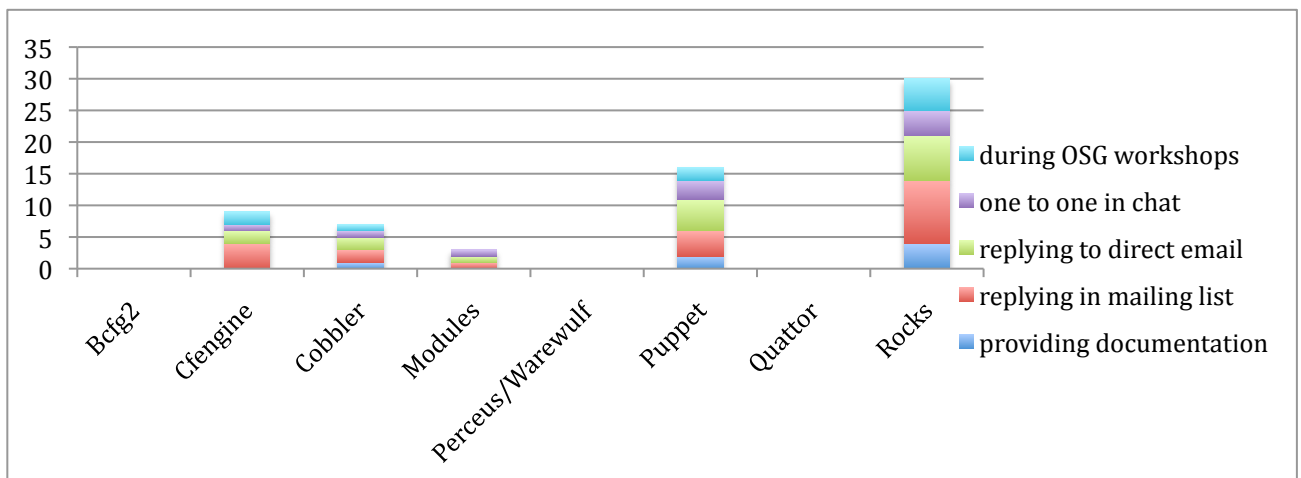
Bcfg2	3	0	0	0	3
Cfengine	1	2	2	1	6
Cobbler	2	2	1	0	5
Modules	2	1	1	0	4
Perceus/Warewulf	2	0	0	0	2
Puppet	1	3	4	0	8
Quattor	2	0	1	0	3
Rocks	0	9	5	1	15

How long have you been using it?

Answer Options	N/A - Never	Less than 1 month	Months	1 year	2-4 years	5 years or more	Response Count
Bcfg2	3	0	0	0	0	0	3
Cfengine	1	0	0	1	2	2	6
Cobbler	2	0	1	1	1	0	5
Modules	2	0	0	0	2	0	4
Perceus/Warewulf	2	0	0	0	0	0	2
Puppet	2	1	1	3	1	0	8
Quattor	2	0	0	1	0	0	3
Rocks	2	0	1	3	4	5	15

Would you help fellow OSG site administrators using any of the following management system? If yes you may check also your preferred ways to help.

Answer Options	no	yes	providing documentation	replying in mailing list	replying to direct email	one to one in chat	during OSG workshops	Resp. Count
Bcfg2	6	0	0	0	0	0	0	6
Cfengine	6	2	0	4	2	1	2	10
Cobbler	6	3	1	2	2	1	1	10
Modules	6	2	0	1	1	1	0	8
Perceus/Warewulf	6	0	0	0	0	0	0	6
Puppet	5	5	2	4	5	3	2	11
Quattor	7	0	0	0	0	0	0	7
Rocks	7	9	4	10	7	4	5	18



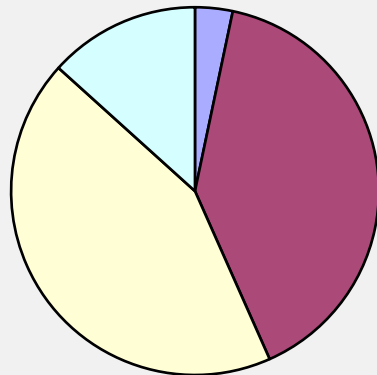
VDT is working hard on providing native packages (DEB and RPM). Are you interested in native packages?

Answer Options	Response Percent	Response Count
No, I will use only Pacman packages (or a tar based distribution)	3.3%	1
Yes, but I will keep using also Pacman	40.0%	12
Yes, I will switch to only native packages	43.3%	13
Don't care	13.3%	4
answered question		30
skipped question		3

If you are interested in native packages, are you... (Please check all that applies)

Answer Options	Response Percent	Response Count
Interested in RPM	88.5%	23
Interested in DEB (Debian) packages	30.8%	8
Would like native packages integrated with your cluster/configuration management system (add below which one)	7.7%	2
Other (please specify)		1
answered question		26
skipped question		7

VDT is working hard on providing native packages (DEB and RPM). Are you interested in native packages?



- No, I will use only Pacman packages (or a tar based distribution)
- Yes, but I will keep using also Pacman
- Yes, I will switch to only native packages
- Don't care

Define your level of knowledge and what you like/dislike in the following local resource managers. A blank answer is the same as N/A, does not apply (e.g. You can leave blank the answers if you don't know that local resource manager)

Your level of knowledge

Answer Options	N/A - Unknown	Read about it	Evaluated it	Use it	Did some customization	Are official developer	Response Count
Condor	0	1	4	11	9	0	25
LSF	5	3	0	1	2	0	11
OpenPBS	1	4	1	4	0	0	10
PBS Pro	3	5	0	3	0	0	11
TORQUE/MAUI	2	4	1	12	0	0	19
SGE	2	4	3	3	2	0	14

Performance

Answer Options	N/A	Good	Average	Poor	Response Count
Condor	1	18	4	2	25
LSF	2	2	1	0	5
OpenPBS	0	4	1	1	6
PBS Pro	4	3	0	0	7
TORQUE/MAUI	1	10	3	1	15
SGE	1	2	2	0	5

First installation/configuration

Answer Options	N/A	Easy	Average	Difficult	Response Count
Condor	2	3	14	6	25
LSF	1	0	1	1	3
OpenPBS	1	0	4	1	6
PBS Pro	4	1	2	0	7
TORQUE/MAUI	2	1	9	2	14
SGE	1	0	3	0	4

Following installation/update

Answer Options	N/A	Easy	Average	Difficult	Response Count
Condor	3	9	10	1	23
LSF	1	0	2	0	3
OpenPBS	1	0	5	0	6
PBS Pro	4	1	2	0	7
TORQUE/MAUI	2	4	7	0	13
SGE	1	2	1	0	4

Operation

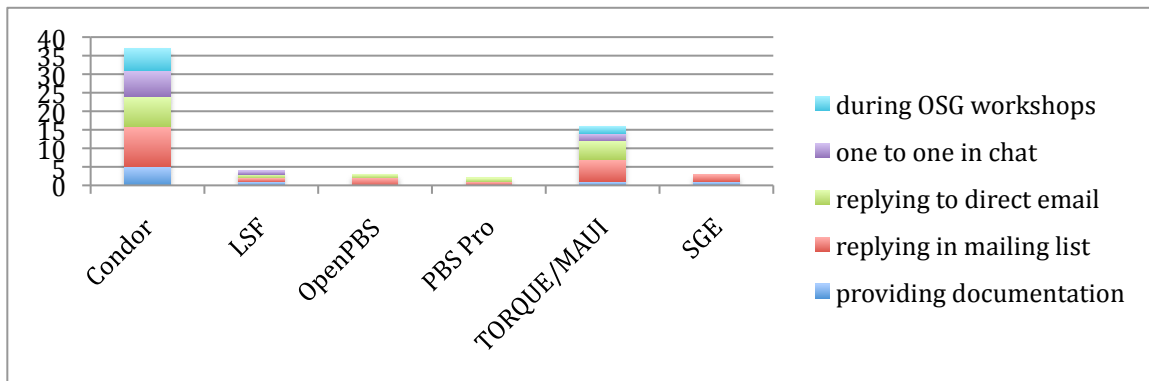
Answer Options	N/A	Easy	Average	Difficult	Response Count
Condor	1	7	13	3	24
LSF	1	1	2	0	4
OpenPBS	0	2	3	0	5

PBS Pro	4	3	0	0	7
TORQUE/MAUI	3	2	8	0	13
SGE	1	0	3	0	4

Available documentation					
Answer Options	N/A	Good	Average	Poor	Response Count
Condor	1	16	6	1	24
LSF	2	2	0	0	4
OpenPBS	0	3	2	1	6
PBS Pro	3	4	0	0	7
TORQUE/MAUI	2	1	9	0	12
SGE	2	1	1	0	4

How long have you been using it?							
Answer Options	N/A - Never	Less than 1 month	Months	1 year	2-4 years	5 years or more	Response Count
Condor	2	2	1	1	11	7	24
LSF	1	0	0	0	2	1	4
OpenPBS	0	1	3	1	0	1	6
PBS Pro	2	0	0	1	1	1	5
TORQUE/MAUI	2	0	2	5	3	2	14
SGE	1	0	0	1	1	1	4

Would you help fellow OSG site administrators using any of the following local resource managers? If yes you may check also your preferred ways to help.								
Answer Options	no	yes	providing documentation	replying in mailing list	replying to direct email	one to one in chat	during OSG workshops	Response Count
Condor	10	13	5	11	8	7	6	24
LSF	8	1	1	1	1	1	0	9
OpenPBS	4	5	0	2	1	0	0	9
PBS Pro	7	1	0	1	1	0	0	8
TORQUE/MAUI	7	8	1	6	5	2	2	15
SGE	8	2	1	2	0	0	0	10



Define your level of knowledge and what you like/dislike in the following shared file systems. A blank answer is the same as N/A, does not apply (e.g. You can leave blank the answers if you don't know that specific file system)

Your level of knowledge

Answer Options	N/A - Unknown	Read about it	Evaluated it	Use it	Did some customization	Are official developer	Response Count
AFS	1	6	1	4	2	0	14
dCache	1	4	0	8	2	0	15
GPFS	2	4	0	2	0	0	8
Hadoop	2	7	3	3	2	0	17
Lustre	1	5	1	8	0	0	15
NFS	0	0	0	24	0	0	24
xrootd	1	3	1	6	0	1	12

Performance

Answer Options	N/A	Good	Average	Poor	Response Count
AFS	2	1	3	3	9
dCache	2	3	5	1	11
GPFS	3	2	0	0	5
Hadoop	3	7	0	0	10
Lustre	3	8	0	0	11
NFS	0	6	13	2	21
xrootd	2	4	3	0	9

First installation/configuration

Answer Options	N/A	Easy	Average	Difficult	Response Count
AFS	2	3	2	2	9
dCache	3	0	2	7	12
GPFS	2	0	0	1	3
Hadoop	2	3	3	1	9
Lustre	4	1	2	2	9
NFS	0	18	2	0	20
xrootd	2	2	5	0	9

Following installation/update

Answer Options	N/A	Easy	Average	Difficult	Response Count
AFS	2	3	2	1	8
dCache	3	1	0	8	12
GPFS	2	0	1	0	3
Hadoop	2	3	2	1	8
Lustre	4	1	3	1	9
NFS	0	18	1	0	19
xrootd	2	3	2	1	8

Operation

Answer Options	N/A	Easy	Average	Difficult	Response Count
AFS	2	2	2	2	8
dCache	3	0	3	6	12
GPFS	2	0	1	0	3
Hadoop	2	4	2	0	8
Lustre	3	2	2	2	9
NFS	0	16	2	1	19
xrootd	2	3	3	0	8

Available documentation					
Answer Options	N/A	Good	Average	Poor	Response Count
AFS	3	2	1	1	7
dCache	3	2	5	2	12
GPFS	2	1	0	0	3
Hadoop	2	4	2	0	8
Lustre	4	2	2	0	8
NFS	1	12	4	2	19
xrootd	2	0	4	2	8

How long have you been using it?							
Answer Options	N/A - Never	Less than 1 month	Months	1 year	2-4 years	5 years or more	Response Count
AFS	2	0	0	1	0	4	7
dCache	3	0	0	0	6	2	11
GPFS	2	0	0	0	1	1	4
Hadoop	2	0	2	2	1	1	8
Lustre	3	0	1	3	2	0	9
NFS	0	0	0	0	9	10	19
xrootd	2	0	1	2	2	0	7

Would you help fellow OSG site administrators using any of the following shared file systems? If yes you may check also your preferred ways to help.								
Answer Options	no	yes	providing documentation	replying in mailing list	replying to direct email	one to one in chat	during OSG workshops	Response Count
AFS	9	3	0	3	2	2	1	12
dCache	11	3	0	3	2	1	1	14
GPFS	7	1	0	1	1	0	1	8
Hadoop	7	5	3	5	5	4	5	12
Lustre	8	5	1	5	3	2	1	13
NFS	6	14	4	13	9	5	6	20
xrootd	10	4	0	4	4	3	3	14

