

Table of Contents

ScFourServiceTechnicalFactors.....	1
SC4 High Level Architecture and Flows.....	1
Product Components.....	1
Service Capacity Data.....	2
Product Availability Data.....	2
Product Network Data.....	3
Detailed Configuration.....	4
Server process accounts registered names and uid/gid values.....	4
Reserved usernames for specific (local) services, as reserved in CRA:.....	5
Critical and High Services.....	5
Service Class Criteria.....	8
Product Evaluation.....	8
Issues.....	9
Assumptions.....	9

ScFourServiceTechnicalFactors

This chapter provides additional data on what components, capacities and constraints are involved in the delivery of the Service Challenge 4 services.

While this is independent of the service level requirements, the technical implementation of a product may influence if or how a particular service level is achieved for that product.

SC4 High Level Architecture and Flows

SC4 consists of a set of services based around 5 major groups. The architectures of LCG-2. and EGEE can be used as references.

Grouping	Products Covered	Description	Flows
WMS	RB CE GRPK	Workload Management System	WmsFlows
DMS	SE FTS LFC	Data Management System	
IS	BDII	Information System	
AAS	PX VOMS	Authentication and Authorisation Services	
MS	MONB GRVW SFT	Monitoring System	

Product Components

Each product consists of a set of software implementing the function and dependent middleware. The dependent middleware must be running in order for the service to function.

This table defines the components for each product as to be installed at CERN. Networking and Linux are assumed. Other sites implementations may differ depending on local skills and policies.

Grouping	Product	Notes	Implementation	Database	Web Server	LDAP	!GridFTP
WMS	RB	RbNotes	RbWlcg	MySQL		GRIS:2135	Yes:2811
WMS	CE	CeNotes	CeWlcg	MySQL (empty)		GRIS	Yes:2811
WMS	GRPK						
DMS	SE						Yes:2811
DMS	FTS	FtsNotes	FtsWlcg	Oracle	Tomcat5	GRIS:2135	
DMS	LFC	LfcNotes	Memorandum LfcWlcg	Oracle			
IS	BDII	BdiiNotes	BdiiWlcg			GIIS EGEE.BDII	
AAS	PX	PxNotes	PxWlcg			Yes:2135	
AAS	VOMS	VomsNotes	VomsWlcg	Oracle	Apache Tomcat		
MS	RGMA	RgmaNotes		MySQL			

MS	SFT			MySQL Oracle			
MS	GRVW				Apache		

Service Capacity Data


The product capacity indicates the current requirements for each product.

For the number of machines in each category, see Detailed Configuration.

Service	Memory (GB)	File Storage (GB)	Oracle (GB)	Criticality
RBP	2	9000		C
PXP	2	40		C
BDIIP	2	80		C
BDIIL	2	80		H
BDIIE	2	80		C
CEP	2	80		C
RGMAP	4	160		M
MONBP	4			M
GRVWP				M
SFTP	4	160		M
GRPKP				M
VOMSP	4	80		C
LFCP-ALICE	4		1000	H
LFCP-ATLAS	4		1000	H
LFCP-CMS	4		1000	H
LFCP-LHCB	4		1000	C
FTSP	4		1000	C
CTRGP				C
GRVWP				M

where

- File storage includes local files and MySQL data (stored in local databases).

 No test services have been discussed. These are not within the scope of SC4 but will be required for the complete production service.

Product Availability Data

Many grid services record state information as part of their operations. A failure of a component or a failover within a high availability configuration may lead to a loss of state data. This section covers what state data is required to be stored in shared storage in the event of a high availability configuration being selected to reach the required service levels.

Product	HA Approach	Impact of Downtime	File State Data	Database State Data
RB	Filesystem Takeover			
PX	Application Replication	Long running jobs cannot renew proxy and fail. Users cannot create new proxies. FTS transfer suspend		

BDII	Multiple independent instances with DNS round robin	no automatic failover to external BDII's if CERN site down. Some sites have their own BDII's. State kept (4MB) in memory and on disk		
CE	Filesystem Takeover	New jobs cannot be submitted to run at the site. Status of jobs running at the site will not be reported.		
RGMA				
MONB		Permanently lose monitoring data		
GRVW				
SFT		Site status cannot be monitored. New or fixed sites cannot join. Broken sites will not be detected.		
GRPK		Job output cannot be viewed by users		
VOMS	Master/Slave with IP address takeover	VOMS permissions are allocated with a lifetime of 24 hours. 90 minutes before expiration, a renew operation is tried. Therefore, after 90 minutes of downtime, 5% of jobs will fail every hour.		
LFC	DNS Round Robin			File Catalog
FTS	DNS Round Robin	Single machine does not affect service. No file transfers initiated by site performed if entire service is down.		
CTRG	DNS Round Robin	Single machine does not affect service. No access to Castor if entire service is down.	None	None

where

- Impact of downtime defines what is the result of the machine not being available (such as a reboot or repairable hardware problem)
- Stateful defines if the server running the product requires state information which may be lost in the event of a failure of the data storage devices.

The High Availability solution portfolio is described in ScFourHighAvailabilityPortfolio.

Product Network Data

Each product has its own network requirements regarding

- Network capacity (High>100Mbit/s, Medium>10Mbit/s, Low<10Mbit/s)
- External network accessibility (outgoing means low ports protected by firewall, incoming low ports accessible by all services)
- Aliases supported (can the service be identified using a network alias which is different from the hostname of the machine). OK means aliases are supported. LB means that aliases and load balanced (i.e. a list of machines can be given) or NO means not supported.
- Ports In provide the list of ports > 1024 for which connectivity is required

Product	Capacity	Accessibility	Aliases	Ports In
RB	M	Incoming		
PX	L	Incoming		
BDII	L	Outgoing	LB	2170
CE	L	Incoming		
RGMA	M	Incoming		
MONB	M	Incoming		
ARCH	M	Incoming		
GRVW	M	Incoming		
SFT	M	Incoming		

GRPK	M	Outgoing		
VOMS	M	Outgoing		
LFC	H	Incoming		
FTS	H	Incoming		
CTRG	H	Outgoing		

Detailed Configuration

Server process accounts registered names and uid/gid values

The following server account names and uid and gid values have been reserved in the CERN central account registration data base CRA in order to prevent them being used by other users or groups. The values installed on a server are not taken from CRA but from a SINDES database managed by FIO group. The service group name is not stored in CRA and the associations between the uids and gids are not stored in CRA.

Service account	uid	gid	Service group name	CRA group name	primary or secondary group
edguser	17680	2747	edguser	g01	Primary
edguser	17680	2761	infosys	g15	Secondary
edginfo	17695	2748	edginfo	g02	Primary
edginfo	17695	2761	infosys	g15	Secondary
rgma	17696	2749	rgma	g03	Primary
rgma	17696	2761	infosys	g15	Secondary
dpmmgr	17697	2750	dpmmgr	g04	Primary
lfcmgr	17700	2751	lfcmgr	g05	Primary
ceuser	17719	2752	ceuser	g06	Primary
condor	17728	2753	condor	g07	Primary
wmsuser	17856	2754	wmsgroup	g08	Primary
hacluser	11774	2755	hacluser	g09	Primary
gridview	15257	2756	gridview	g10	Primary
glite	21086	2757	glite	g11	Primary

Here are the /etc/passwd lines:

```
edguser:x:17680:2747::/home/edguser:/bin/bash
edginfo:x:17695:2748::/home/edginfo:/bin/bash
rgma:x:17696:2749:RGMA user:/opt/edg/etc/rgma:/bin/bash
dpmmgr:x:17697:2750:DPM manager:/home/dpmmgr:/bin/bash
lfcmgr:x:17700:2751:LFC manager:/home/lfcmgr:/bin/bash
ceuser:x:17719:2752::/home/ceuser:/bin/bash
condor:x:17728:2753::/home/condor:/bin/bash
wmsuser:x:17856:2754:/home/wmsuser:/bin/bash
hacluser:x:11774:2755:/home/hacluser:/bin/bash
gridview:x:15257:2756:/home/gridview:/bin/bash
glite:x:21086:2757:/home/glite:/bin/bash
```

And the lines in /etc/group:

```
edguser:x:2747:
edginfo:x:2748:
rgma:x:2749:
dpmmgr:x:2750:
lfcmgr:x:2751:
ceuser:x:2752:
```

```
condor:x:2753:
wmsgroup:x:2754:
haclient:x:2755:
gridview:x:2756:
glite:x:2757:
infosys:x:2761:rgma,edginfo,edguser
```

Reserved usernames for specific (local) services, as reserved in CRA:

Service account	uid	gid	Owner
samops	23550	1028	Judit Novak
samdteam	23551	1028	Judit Novak
samatlas	23552	1028	Piotr Nyczyk
samcms	23554	1028	Andrea Sciaba
samalice	23763	1028	Patricia Mendez
dirac	25133	1470	Joel Closier
jabber	25134	1470	Joel Closier
tomcat	none	1028	Production Grid-Service
mysql	none	1028	Production Grid-Service
atlsrv	28475	1028 (local 1307)	Production Grid-Service

Critical and High Services

Service	Masters	Passive	Clones	Spares	FCports	Comment
RB-ALICE	1	0	0	0	1	Spare shared with RBP-PROD
RB-ATLAS	1	0	0	0	1	Spare shared with RBP-PROD
RB-CMS	1	0	0	0	1	Spare shared with RBP-PROD
RB-LHCB	1	0	0	0	1	Spare shared with RBP-PROD
RB-PROD	1	0	0	1	2	
PX	2	0	2		2	Replicated
BDIIL	1	1	0	0		LCG BDII
BDIIP	1	1	0	0		PROD BDII (CERN Site)
BDIIE	1	1	0	0		Experiment BDII
CE	1	1	0	0	2	
VOMS	2	1	0	0	0	
FTS	7	0	0	2		Spare shared between VOs
LFC-LHCB	2	0	0	0		Spare shared between VOs
LFC-ALICE	1	0	0	0		Spare shared between VOs
LFC-ATLAS	1	0	0	0		Spare shared between VOs
LFC-CMS	1	0	0	0		Spare shared between VOs
LFC-SHARED	1	0	0	0		Shared server for other VOs
LFC-PROD	1	0	0	0		Backup lfc server for all
GRVW	1	1	0	0		Grid View

Thus

- 28 spaces for master mid range servers in the LCG area
- 11 spaces for backup/slaves in the LCG area
- 10 fibre channel ports required

The full list of machines is therefore

ScFourServiceTechnicalFactors < LCG < TWiki

Machine	Service	CDB Cluster	Purpose	Area	Config	Comment
bdii001	BDIIL	gridbdii	LCG BDII Master	UPS	Basic Midrange Server	In prod.To be logically moved to LCG
bdii002	BDIIL	gridbdii	LCG BDII Backup	UPS	Basic Midrange Server	In prod.To be logically moved to LCG
bdii101	BDIIL	gridbdii	LCG BDII Master		Basic Midrange Server	Switch1.Add to load balancing then stop bdii001. Priority 1
bdii102	BDIIL	gridbdii	LCG BDII Backup		Basic Midrange Server	Switch2. Add to load balancing then stop bdii002. Priority 1
bdii103	BDIIP	gridbdii	Site BDII Master		Basic Midrange Server	Switch2. Priority 1
bdii104	BDIIP	gridbdii	Site BDII Backup		Basic Midrange Server	Switch1. Priority 1
bdii105	BDIIE	gridbdii	Experiment BDII Master		Basic Midrange Server	Switch1. Priority 1
bdii106	BDIIE	gridbdii	Experiment BDII Master		Basic Midrange Server	Switch2. Priority 1
ce101	CEP	gridce	Production CE Master	NFC	Basic Midrange Server	Switch2.Leave unused ce001 in UPS area for now. Priority 2
ce102	CEP	gridce	Production CE Backup	NFC	Basic Midrange Server	Switch1. Priority 2
fts101	FTSP	gridfts	production FTS Transfer Agent Master		Large Memory Midrange Server	Switch1. Priority 4
fts102	FTSP	gridfts	production FTS Transfer Agent Hot Spare		Large Memory Midrange Server	Switch2. Priority 4
fts103	FTSP	gridfts	production FTS Web Server Master		Large memory Midrange Server	Switch1. lb name prod-ftsws. Priority 4
fts104	FTSP	gridfts	production FTS Web Server Master		Large Memory Midrange Server	Switch2. lb name prod-ftsws. Priority 4
fts105	FTSP	gridfts	production FTS Alice agent		Basic Midrange Server	Switch1. alias prod-ftsvo-alice. Priority 4
fts106	FTSP	gridfts	production FTS Atlas agent		Basic Midrange Server	Switch2. alias prod-ftsvo-atlas. Priority 4
fts107	FTSP	gridfts	production FTS CMS agent		Basic Midrange Server	Switch1. alias prod-ftsvo-cms. Priority 4
fts108	FTSP	gridfts	production FTS LHCb agent		Basic Midrange Server	Switch2. alias prod-ftsvo-lhcb. Priority 4
fts109	FTSP	gridfts	production experiment agent Hot Spare		Basic Midrange Server	Switch1. Priority 4
grvw001	GRVWP	gridgrvw	production GRIDVIEW Web server		Basic Midrange Server	Switch1. Priority 5
grvw002	GRVWP	gridgrvw	production GRIDVIEW data		Basic Midrange Server	Switch2. Priority 5

ScFourServiceTechnicalFactors < LCG < TWiki

			mining server			
lfc101	LFC-LHCB	gridlfc	production LHCb LFC		Basic Midrange Server	Switch1. alias prod-lfc-lhcb. Priority 7
lfc102	LFC-LHCB	gridlfc	production LHCb LFC Backup		Basic Midrange Server	Switch2. Priority 7
lfc103	LFC-ALICE	gridlfc	production Alice LFC		Basic Midrange Server	Switch2. alias prod-lfc-alice. Priority 7
lfc104	LFC-ATLAS	gridlfc	production Atlas LFC		Basic Midrange Server	Switch2. alias prod-lfc-atlas. Priority 7
lfc105	LFC-CMS	gridlfc	production CMS LFS		Basic Midrange Server	Switch2. alias prod-lfc-cms. Priority 7
lfc106	LFCP	gridlfc	production shared LFC		Basic Midrange Server	Switch1. alias prod-lfc-shared. Priority 7
lfc107	LFCP	gridlfc	production LFC backup		Basic Midrange Server	Switch1. Priority 7
rb101	RB-ALICE	gridrb	RB for Alice	NFC	Extra disk Midrange Server	Switch1. Priority 8
rb102	RB-ATLAS	gridrb	RB for Atlas	NFC	Extra disk Midrange Server	Switch1. Priority 8
rb103	RB-CMS	gridrb	RB for CMS	NFC	Extra disk Midrange Server	Switch1. Priority 8
rb104	RB-LHCB	gridrb	RB for LHCB	NFC	Extra disk Midrange Server	Switch1. Priority 8
rb105	RB-PROD	gridrb	RB for other VOs	NFC	Extra disk Midrange Server	Switch1. Priority 8
rb106	RB-PROD	gridrb	RB spare	NFC	Extra disk Midrange Server	Switch2. Priority 8
px101	PXP	gridpx	Production MyProxy Master	NFC	Basic Midrange Server	Switch2. Priority 3
px102	PXP	gridpx	Production MyProxy Slave	NFC	Basic Midrange Server	Switch1. Priority 3
px103	PXP	gridpx	Production MyProxy Master for FTS		Basic Midrange Server	Switch2. Priority 3
px104	PXP	gridpx	Production MyProxy Slave for FTS		Basic Midrange Server	Switch1. Priority 3
voms101	VOMSP	gridvoms	Production VOMS Master	NFC	Large Memory Midrange Server	Switch1. Priority 6
voms102	VOMSP	gridvoms	Production VOMS Slave	NFC	Large Memory Midrange Server	Switch2. Priority 6
voms103	VOMSP	gridvoms	Production VOMS ldap publisher	NFC	Basic Midrange Server	Switch2. Priority 6

- A Basic Midrange Server has 2GB memory and 160GB internal mirrored disk. A configuration larger than this would also be ok.

- A large memory midrange server has the same configuration as a basic midrange server but with 4GB memory.
- An extra disk midrange server has the same configuration as a basic midrange server but with two extra 250GB disks run mirrored.
- The Resource Brokers should have extra disks. The plan is to replace the first servers with recuperated tape servers, which have extra memory and an HBA in and will not need the extra disks, when the SAN infrastructure is in place.
- UPS means in the diesel backed critical area
- NFC means that the machine needs to be near a fibre channel switch in the LCG network area
- Priority 1 is highest. Items to priority 5 should be completed in 2005.

Service Class Criteria

Attribute	Class U	Class L	Class M	Class H	Class C
Facilities					
Controlled physical access			Badge	Badge	Badge
Power into Data Centre				Redundant	Redundant
Physical					
Power connection on UPS				Yes	Yes
If HA, only 1 machine required on UPS					
Machine in rack			Yes	Yes	Yes
Hardware					
Redundant power supply in PC				Yes	Yes
Internal system disks mirrored			Yes	Yes	Yes
Console remotely accessible		Yes	Yes	Yes	Yes
Storage					
Minimum RAID Levels for data		5	5	5	5
Redundant Controllers / Paths				Yes	Yes
Backup					
Off-site copies of backup data					
Yearly backup/restore test					
Networking					
Redundant network cards					
Monitoring					
Status command for each component			Yes	Yes	Yes
Automatic Event reported to console if component down			Yes	Yes	Yes
Configuration					
Automatic configuration from database/xml				Yes	Yes
High Availability					
Standby Levels			Cold	Warm	Hot
Procedures for failover			Administrator	Operator	Automatic

Product Evaluation

In order to assess what technical factors may cause problems to deliver the quality of service requested, the ScFourTechnicalQuestionnaire has been written. With these questions, an assessment of the readiness of the application and infrastructure to provide the requested service level can be made.

The current servers involved in delivering the service are defined at PreSc4ServersInfo.

Issues

The following items have been raised as part of the evaluation of the technical solutions.

Nr	Description	Status	Open Date	Who	Log
1	Service definition for MySQL	inprogress	2005/09/15	Bernd	IssueMySQLService
2	RB disk space estimates are very large	inprogress	2005/09/15	Maarten	IssueRbDiskSpace

Assumptions

In order to accelerate the definition of the services, some assumptions have been made by the fabric team. This section documents these.

Nr	Description	Status	Open Date	Who	Log
1	BDII is outgoing connectivity only	closed	2005/09/22	Tim	Port 2170 required which is covered by outgoing connectivity
2	CE has MySQL installed and an empty database.	closed	2005/09/22	Tim	Database created by install
3	CEmon is not included in SC4	closed	2005/09/26	Maarten	To be reviewed in Dec 2005
4	myproxy does not require external connectivity for low ports	open	2005/09/26	Tim	Need to identify contact
5	myproxy data is all stored in /var/myproxy	closed	2005/09/26	Tim	Review replication procedure with NCSA developers.

-- TimBell - 05 Sep 2005 s

This topic: LCG > ScFourServiceTechnicalFactors

Topic revision: r75 - 2007-10-08 - HarryRenshall



Copyright &© 2008-2019 by the contributing authors. All material on this collaboration platform is the property of the contributing authors.

Ideas, requests, problems regarding TWiki? Send feedback