

# FermiGrid-HA Phase 2 (FermiGrid-HA2)

Keith Chadwick  
18-Mar-2010

Abstract:

This document discusses the requirements and design for FermiGrid-HA (High Availability) Phase 2 (FermiGrid-HA2).

**Document Revision History:**

<b>Version</b>	<b>Date</b>	<b>Author</b>	<b>Comments</b>
0.1	16-Mar-2010	Keith Chadwick	Initial highly draft version.
1.0	18-Mar-2010	Keith Chadwick	Draft placed in CD-DocDB

## **Table of Contents**

<b>Introduction .....</b>	<b>1</b>
<b>Core Services Rack Configuration .....</b>	<b>3</b>
<b>Gatekeeper Rack Configuration .....</b>	<b>4</b>
<b>OSG Services Rack Configuration .....</b>	<b>5</b>
<b>FermiGrid-HA2 Requirements .....</b>	<b>6</b>
<b>FermiGrid-HA2 Network Design.....</b>	<b>6</b>
<b>FermiGrid-HA2 Rack Configuration.....</b>	<b>8</b>
<b>FermiGrid-HA2 Transition Plan .....</b>	<b>9</b>

## Introduction

FermiGrid has previously architected the high availability FermiGrid-HA ([reference CD-DocDB Document # 3579](#)). As part of FermiGrid-HA planning, the potential requirements of a future enhancement to extend the FermiGrid-HA design and deployment to support Redundant and/or Resilient Services (FermiGrid-RS) were also considered. Due to the previously demonstrated reliability of the computer rooms in the Feynman Computing Center following the installation of the building wide UPS and generator, the FermiGrid-RS extensions were not viewed as being needed until the CY2011-2012 timeframe. The principle issues with the computer room had been the availability of cooling capacity, and due to the relatively low power required to operate the systems that provided the FermiGrid-HA services, FermiGrid personnel had been allowed to operate these systems during FCC cooling outages.

Unfortunately, there were two significant power interruptions in February 2010 for the computer rooms in the Feynman Computing Center. The exact root cause is not pertinent to this document, but these two incidents raised the need for continued operation of the set of FermiGrid-HA services, and thus the planning process was initiated for FermiGrid-RS. In addition, the name of FermiGrid-RS was changed to FermiGrid-HA Phase 2 (HA2).

The figure below shows the FermiGrid-HA “Service Communication Diagram”:

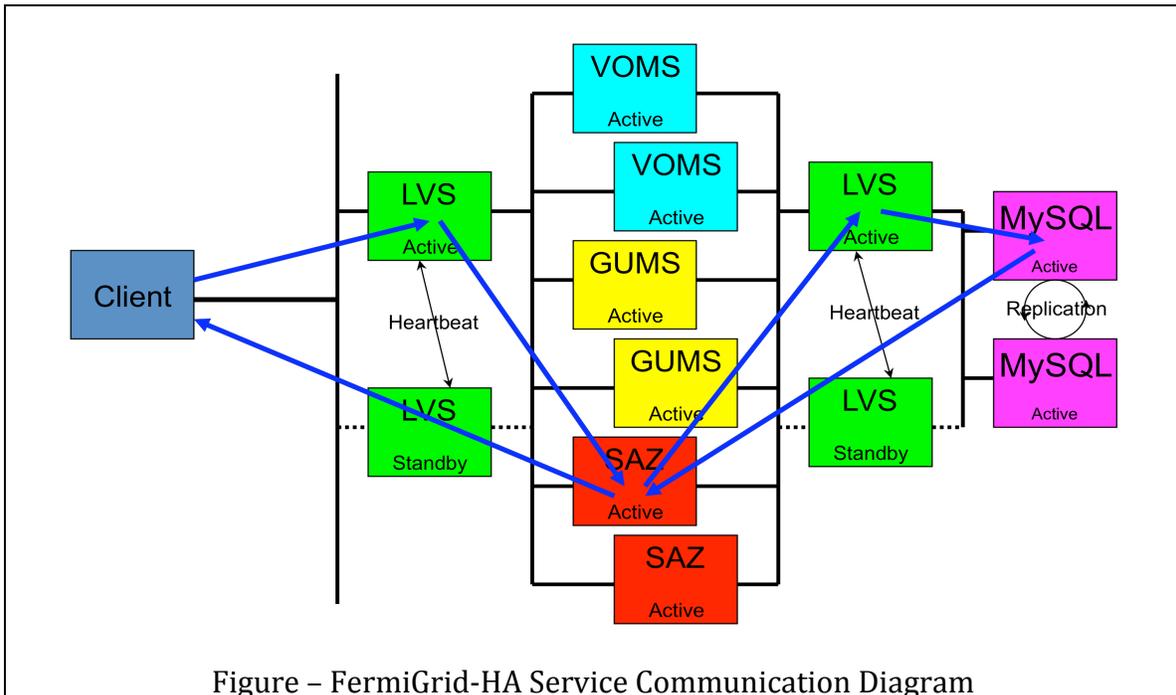


Figure – FermiGrid-HA Service Communication Diagram

Each of the services identified above were deployed in Xen virtual machines under Scientific Linux as show in the figure below:

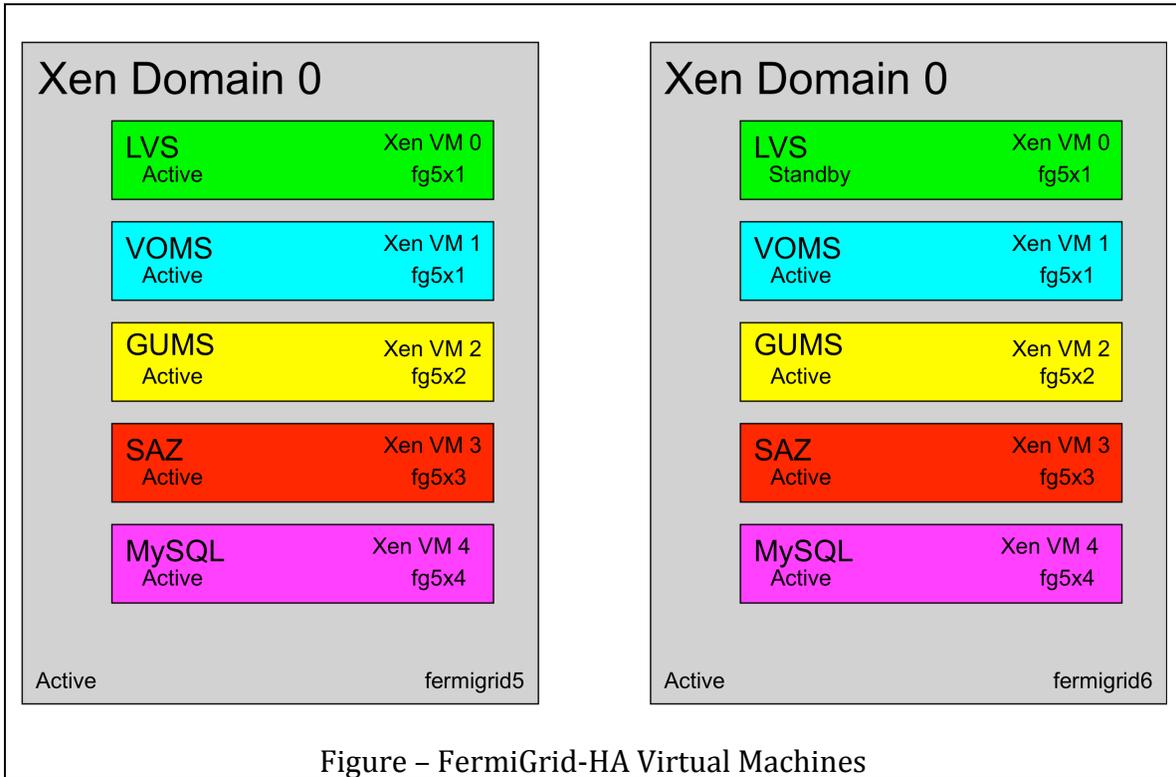


Figure – FermiGrid-HA Virtual Machines

Beyond the FermiGrid-HA / FermiGrid Core Services systems above, FermiGrid operates three classes of systems:

**Core Services:**

fermigrd[0-6] and associated virtual machines fg0x[0-7] and fg[1-6]x[0-5]

**Gatekeeper Services:**

fcdsrv[0-5] and associated virtual machines fcdf[0-5]x[0-5]  
d0osgrv[1,2] and associated virtual machines d0osg[1,2]x[0-4]  
fnpcsrv[3-9] and associated virtual machines

**OSG Services:**

Gratia[10-13] and associated virtual machines gr[10-13]x[0-5]  
Ress[01,02] and associated virtual machines ress[1,2]x[0-3]

The following three sections show the current physical machine layout of these three classes of systems across the three FermiGrid racks:

## Core Services Rack Configuration

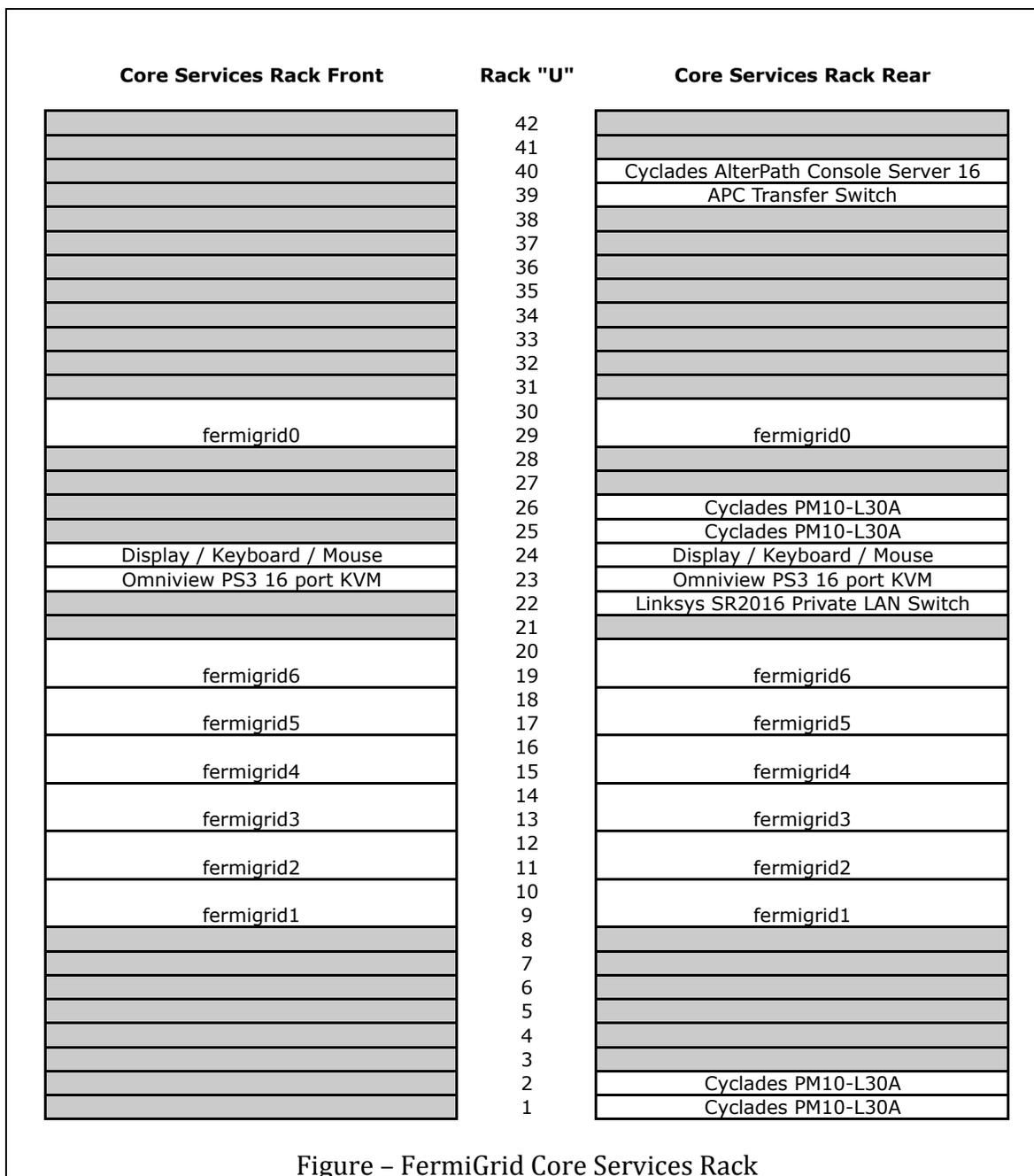


Figure – FermiGrid Core Services Rack

## Gatekeeper Rack Configuration

Gatekeeper Rack Front	Rack "U"	Gatekeeper Rack Rear
	42	
	41	Public LAN Switch Catalyst 2960G
	40	Cyclades AlterPath Console Server 32
	39	
d0osgsrv2	38	d0osgsrv2
	37	
	36	
	35	
d0osgsrv1	34	d0osgsrv1
	33	
fcdfsrv4	32	fcdfsrv4
	31	
fcdfsrv5	30	fcdfsrv5
	29	
fcdfsrv3	28	fcdfsrv3
	27	
	26	Cyclades PM10-L30A
	25	Cyclades PM10-L30A
Display / Keyboard / Mouse	24	Display / Keyboard / Mouse
Omniview PS3 16 port KVM	23	Omniview PS3 16 port KVM
	22	
	21	
	20	
	19	
	18	
	17	
fcdfsrv2	16	fcdfsrv2
	15	
fcdfsrv1	14	fcdfsrv1
	13	
fcdfsrv0	12	fcdfsrv0
fnpcsrv3	11	fnpcsrv3
fnpcsrv4	10	fnpcsrv4
fnpcsrv5	9	fnpcsrv5
	8	
	7	
fnpcsrv8	6	fnpcsrv8
fnpcsrv9	5	fnpcsrv9
	4	
	3	Cyclades PM10-L30A
	2	
	1	

Figure – FermiGrid Gatekeeper Rack Layout

# OSG Services Rack Configuration

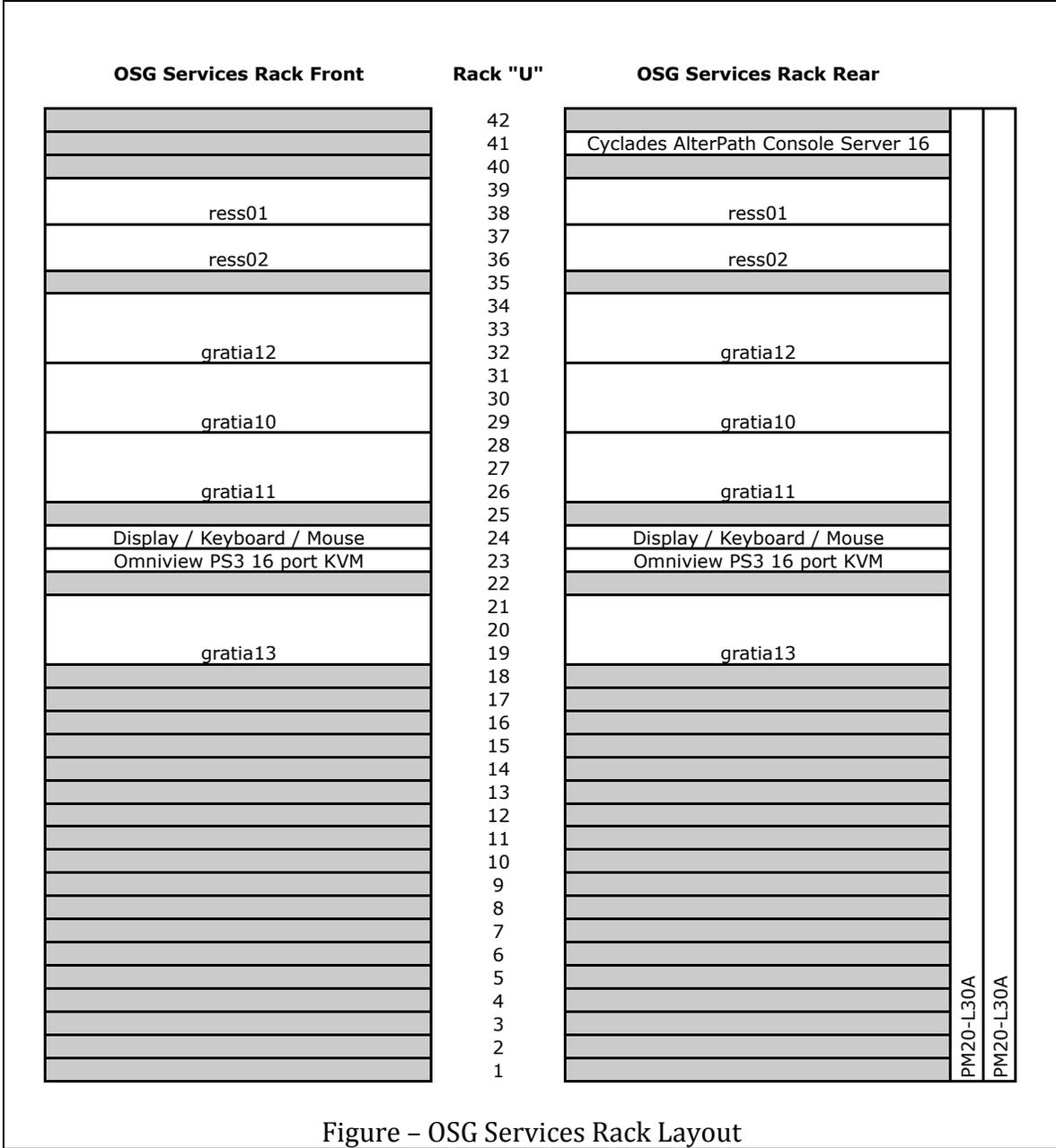


Figure – OSG Services Rack Layout

## FermiGrid-HA2 Requirements

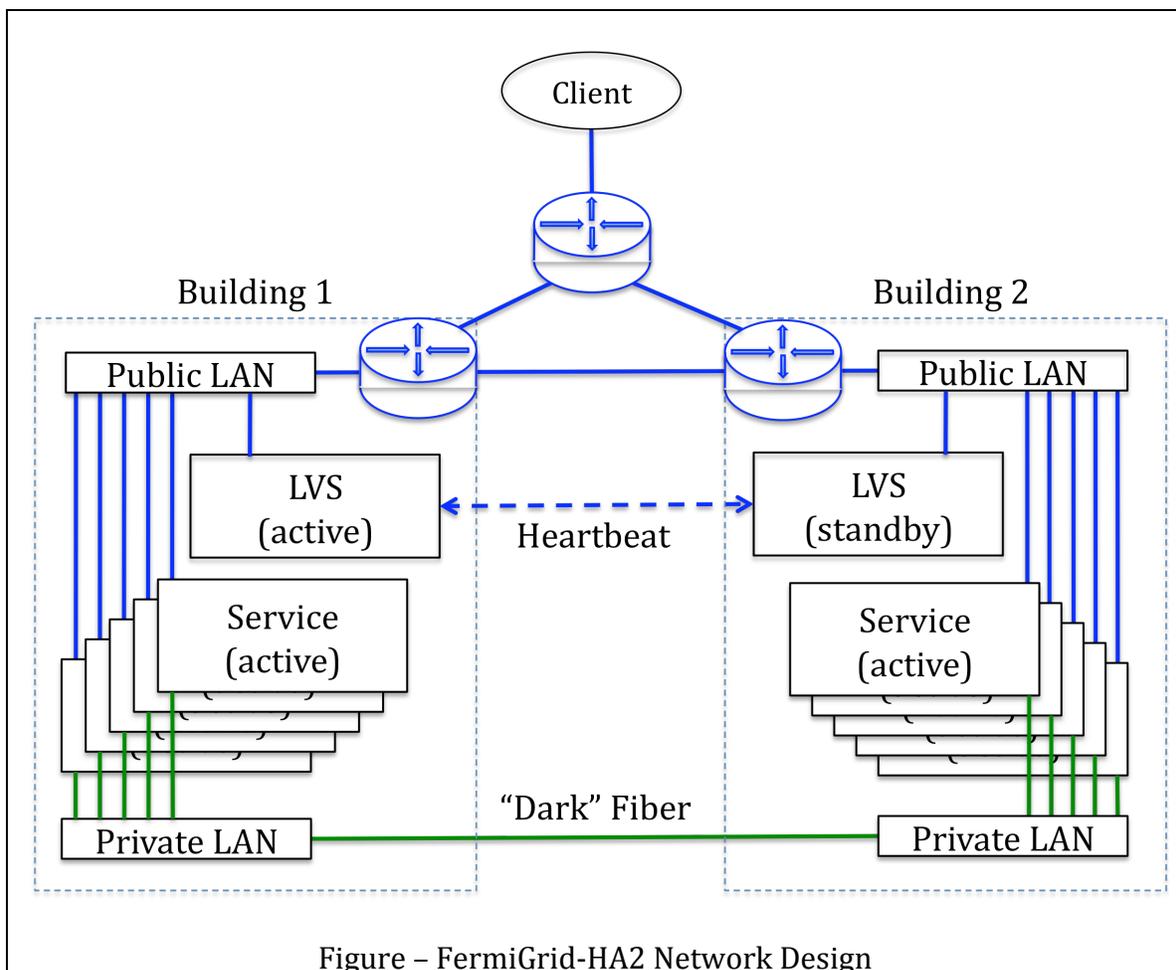
The requirements of FermiGrid-HA2 are very simple:

- Continued operation of the set of FermiGrid-HA services if either building hosting the FermiGrid-HA services is down within the limits of the service operation.

This simple requirement places significant constraints on the network design and service deployment.

## FermiGrid-HA2 Network Design

The figure below shows the proposed FermiGrid-HA2 network design:



This design uses both the “public” Fermilab network (routers and switches attached to various workgroups) as well as an extended private LAN (to allow for secure replication of various “secrets” between the services). An example of such a “secret” is the replication of the drbd blocks that hold the private keys of the credentials that users have chosen to place in the FermiGrid MyProxy repository.

The “Public LAN” connections shall be the standard data communications switch and workgroup ports, and the services that require heartbeat (such as LVS) shall exchange “heartbeat” information across the public LAN interfaces.

The “Private LAN” connections shall be two Linksys SR2024 switches (unmanaged 24 port gigabit Ethernet switches with optional dual miniGBIC optical interfaces), connected via miniGBIC interfaces over “dark” fiber between building 1 (the Feynman Computing Center) and building 2 (Grid Computing Center).

## FermiGrid-HA2 Rack Configuration

Based on the above FermiGrid-HA2 network design, the current FermiGrid systems can be allocated into two identically configured racks as shown in the Figure below:

FermiGrid-HA2 Rack Front	Rack "U"	FermiGrid-HA2 Rack Rear
blank - 2U	42	Cisco Nexxus / 2960G Public LAN Switch
fnpcsrv1 / fnpcsrv2	41	Cyclades AlterPath Console Server 16
fnpcsrv3 / fnpcsrv4	40	fnpcsrv1 / fnpcsrv2
fnpcsrv5 / fnpcsrv6	39	fnpcsrv3 / fnpcsrv4
fgitbsrv1 / fgitbsrv2	38	fnpcsrv5 / fnpcsrv6
	37	fgitbsrv1 / fgitbsrv2
	36	
d0osgsrv1 / d0osgsrv2	35	d0osgsrv1 / d0osgsrv2
blank - 2U	34	
	33	blank - 2U
	32	
fcdfsrv5 / fcdfsrv6	31	fcdfsrv5 / fcdfsrv6
	30	
fcdfsrv3 / fcdfsrv4	29	fcdfsrv3 / fcdfsrv4
	28	
fcdfsrv1 / fcdfsrv2	27	fcdfsrv1 / fcdfsrv2
	26	Cyclades PM10-L30A
blank - 2U	25	Cyclades PM10-L30A
Display / Keyboard / Mouse	24	Display / Keyboard / Mouse
Omniview PS3 16 port KVM	23	Omniview PS3 16 port KVM
	22	Linksys SR2024 Private LAN Switch
blank - 2U	21	APC Transfer Switch
	20	
ress01 / ress02	19	ress01 / ress02
blank - 1U	18	blank - 1U
	17	
	16	
gratia12 / gratia13	15	gratia12 / gratia13
	14	
	13	
gratia10 / gratia11	12	gratia10 / gratia11
blank - 1U	11	blank - 1U
	10	
fermigrid5 / fermigrid6	9	fermigrid5 / fermigrid6
	8	
fermigrid2 / fermigrid3	7	fermigrid2 / fermigrid3
	6	
fermigrid1 / fermigrid4	5	fermigrid1 / fermigrid4
	4	
fermigrid0 / fermigrid8	3	fermigrid0 / fermigrid8
	2	Cyclades PM10-L30A
blank - 2U	1	Cyclades PM10-L30A

Figure – FermiGrid-HA2 Rack Layout

## FermiGrid-HA2 Transition Plan

The transition between the current FermiGrid Services deployment across the three current racks to the FermiGrid-HA2 deployment across the two racks shall consist of the following steps:

1. Identify and acquire the necessary equipment to implement the transition.
2. Identify the (proposed) rack locations:

Rack	Proposed Location
FermiGrid-HA2 Rack 1	FCC 2 Computer Room
FermiGrid-HA2 Rack 2	GCC Network Room A

3. Identify power sources and circuits in the (proposed) rack locations:

FermiGrid-HA2 Rack 1	Power Source	Circuit
Cyclades PM10-L30A #1	FCC UPS-1	tbd
Cyclades PM10-L30A #2	FCC UPS-2	tbd
Cyclades PM10-L30A #3	FCC UPS-1	tbd
Cyclades PM10-L30A #4	FCC UPS-2	tbd

FermiGrid-HA2 Rack 2	Power Source	Circuit
Cyclades PM10-L30A #1	GCC UPS-1	tbd
Cyclades PM10-L30A #2	GCC UPS-2	tbd
Cyclades PM10-L30A #3	GCC UPS-1	tbd
Cyclades PM10-L30A #4	GCC UPS-2	tbd

4. Identify the racks that will be used to perform the migration – The FermiGrid Core Services rack and the OSG Services rack appear to be the best candidates. For the purposes of the remainder of this document, the following rack identification shall be used:

Rack	Current Identification	Future Identification
FCC-1-3465	FermiGrid Core Services	FermiGrid-HA2 Rack 2
FCC-1-3467	FermiGrid Gatekeepers	none
FCC-1-3470	OSG Services	FermiGrid-HA2 Rack 1

5. Install the “base” infrastructure equipment in the FermiGrid Core Services and OSG Services racks (Cyclades ACS-32, Cyclades PM10-L30A, Keyboard-Mouse-Monitor, KVM switches, APC power transfer switches, Linksys SR2024 “private LAN” network switches, public LAN network switches).

6. Connect the Linksys SR2024 switches between the two racks.

7. Begin the system-by-system migration of the selected FermiGrid Core Services systems from the current FermiGrid Core Services rack to the current OSG Services Rack as shown in the FermiGrid-HA2 rack layout above. Verify full service functionality after the selected systems have been relocated.

<b>System</b>	<b>Initial Location</b>	<b>Target Location</b>
fermigrid0	Core Services – 29/30	OSG Services – 3/4
fermigrid1	Core Services – 9/10	OSG Services – 5/6
fermigrid2	Core Services – 11/12	OSG Services – 7/8
fermigrid5	Core Services – 17/18	OSG Services – 9/10

8. Reorganize the remaining FermiGrid Core Service systems in the current FermiGrid Core Services rack as shown in the FermiGrid-HA2 rack layout above.

<b>System</b>	<b>Initial Location</b>	<b>Target Location</b>
fermigrid8	WH8SE Fgtest Rack	Core Services – 3/4
fermigrid4	Core Services – 15/16	Core Services – 5/6
fermigrid3	Core Services – 13/14	Core Services – 7/8
fermigrid6	Core Services – 19/20	Core Services – 9/10

9. Reorganize the Fermilab and OSG Gratia collector systems in the current OSG Services rack as shown in the FermiGrid-HA2 rack layout above.

<b>System</b>	<b>Initial Location</b>	<b>Target Location</b>
gratia10	OSG Services – 29/30/31	OSG Services – 12/13/14
gratia12	OSG Services – 32/33/34	OSG Services – 15/16/17

10. Reorganize the Fermilab and OSG Gratia reporter systems in the current FermiGrid Core Services rack as shown in the FermiGrid-HA2 rack layout above.

<b>System</b>	<b>Initial Location</b>	<b>Target Location</b>
gratia11	OSG Services – 26/27/28	Core Services – 12/13/14
gratia13	OSG Services – 19/20/21	Core Services – 15/16/17

11. Reorganize the OSG Resource Selection Services (ReSS) systems across the OSG Services and FermiGrid Core Services racks as shown above.

<b>System</b>	<b>Initial Location</b>	<b>Target Location</b>
ress01	OSG Services – 38/39	OSG Services – 19/20
ress02	OSG Services – 36/37	Core Services – 19/20

12. At this point, the two racks presently known as FermiGrid Core Services and OSG Services should be fully functioning copies of one another (within the service deployment limitations). A set of tests shall be performed to demonstrate the service failover between the two racks – power down individual systems, power

down the one of the two racks, etc.. If any services fail to failover as expected, the cause shall be identified and rectified prior to proceeding to the next step in the FermiGrid-HA2 transition.

13. Once the service failover tests have been successfully completed, the racks formerly known as the FermiGrid Core Services and OSG Services racks shall be renamed to FermiGrid-HA2 rack 2 and FermiGrid-HA2 rack 1 respectively.
14. At this point, the remainder of the FermiGrid-HA2 repackaging shall be performed – the individual gatekeeper systems in the FermiGrid Gatekeepers rack shall be moved one-by-one to the allocated location in the FermiGrid-HA2 racks. Systems shall be moved one system at a time, and full system functionality shall be verified prior to the movement of any additional systems.

<b>System</b>	<b>Initial Location</b>	<b>Target Location</b>
fnpcsrv3	Gatekeeper – 11	FermiGrid-HA2 Rack 1 –
fnpcsrv4	Gatekeeper – 10	FermiGrid-HA2 Rack 2 –
fnpcsrv5	Gatekeeper – 9	FermiGrid-HA2 Rack 1 –
???	???	FermiGrid-HA2 Rack 2 –
fnpcsrv8	Gatekeeper – 6	FermiGrid-HA2 Rack 1 –
fnpcsrv9	Gatekeeper – 5	FermiGrid-HA2 Rack 2 –
fcdfsrv0	Gatekeeper – 12/13	FermiGrid-HA2 Rack 1 – 27/28
fcdfsrv1	Gatekeeper – 14/15	FermiGrid-HA2 Rack 2 – 27/28
fcdfsrv2	Gatekeeper – 16/17	FermiGrid-HA2 Rack 1 – 29/30
fcdfsrv3	Gatekeeper – 28/29	FermiGrid-HA2 Rack 2 – 29/30
fcdfsrv4	Gatekeeper – 32/33	FermiGrid-HA2 Rack 1 – 31/32
fcdfsrv5	Gatekeeper – 30/31	FermiGrid-HA2 Rack 2 – 31/32
d0osgsrv1	Gatekeeper – 34/35	FermiGrid-HA2 Rack 1 – 35/36
d0osgsrv2	Gatekeeper – 38/39	FermiGrid-HA2 Rack 2 – 35/36

15. Once all systems have been removed from the FermiGrid Gatekeepers rack, FermiGrid personnel shall remove the FermiGrid Gatekeepers rack.
16. The re-racked systems shall be operated for at least 24 hours in the “final” configuration prior to proceeding to the next step in the transition plan.
17. Relocate “FermiGrid-HA2 Rack 1” from the current location to the agreed location identified in Transition Step 2 above. Due to weight restrictions, the systems in the upper half of the rack will need to be physically removed from the rack prior to movement and reinstalled in the rack when the rack has been moved to it’s new location.

18. Re-establish private LAN network connectivity between FermiGrid-HA2 Rack 1 and FermiGrid-HA2 Rack 2.
19. Re-establish public LAN network connectivity between FermiGrid-HA2 Rack 1 and the site network.
20. The racks shall be operated for at least 24 hours in these locations prior to proceeding to the next step in the transition plan.
21. Relocate “FermiGrid-HA2 Rack 2” from the current location to the agreed location identified in Transition Step 2 above. Due to weight restrictions, the systems in the upper half of the rack will need to be physically removed from the rack prior to movement and reinstalled in the rack when the rack has been moved to its new location.
22. Re-establish private LAN network connectivity between FermiGrid-HA2 Rack 2 and FermiGrid-HA2 Rack 1.
23. Re-establish public LAN network connectivity between FermiGrid-HA2 Rack 2 and the site network.
24. The racks shall be operated for at least 24 hours in these locations prior to proceeding to the next step in the transition plan.
25. Transition Complete.