

# GridPP

UK Computing for Particle Physics

## Tier1 - Disk Failure stats and Networking

Martin Bly

Tier1 Fabric Manager



Science & Technology Facilities Council

e-Science

- Tier1 Disk Storage Hardware
- Disk Failure Statistics
- Network configuration

- Tier1 provides disk storage for Castor
  - Particle Physics data (mostly)
- Storage is:
  - Commodity components
  - Storage-in-a-box
  - Lots of units

- Typically 16- or 24-disk chassis
  - Mostly SuperMicro
- One or more hardware RAID cards
  - 3ware/AMCC, Areca, Adaptec, LSI
  - PCI-X or PCI-e
  - Some active backplanes
- Configured as RAID5 or more commonly RAID6
  - RAID1 for system disks, on second controller where it fitted
- Dual multi-cores CPUs
- 4GB, 8GB, 12GB RAM
- 1 GbE NIC
- IPMI

Year	Chassis	CPU	CPU x Cores	RAM /GB	NIC	IMPI	Units
2005	24-bay	AMD 275	2 x 2	4	1 GbE	✗	21
2006	16-bay	AMD 275	2 x 2	4	1 GbE	✗	86
2007 A	16-bay	AMD 2220	2 x 2	8	1 GbE	✓	91
2007 B	16-bay	Intel 5310	2 x 4	8	1 GbE	✓	91
2008 A	16-bay	L5410	2 x 4	8	1 GbE	✓	50
2008 B	16-bay	E5405	2 x 4	8	1 GbE	✓	60
2009 A	16-bay	E5504	2 x 4	12	1 GbE	✓	60
2009 B	16-bay	E5504	2 x 4	12	1 GbE	✓	38

Year	Controller	Data Drive	Data Drives	Raid	Capacity	Units	Drives
2005	Areca 1170	WD5000	22	6	10TB	21	462
2006	3ware 9550SX-16ML	WD5000	14	5	6TB	86	1204
2007 A	3ware 9650SE-16	WD7500	14	6	9TB	91	1274
2007 B	3ware 9650SE-16	WD7500	14	6	9TB	91	1274
2008 A	Areca 1280	WD RE	22	6	20TB	50	1100
2008 B	3ware 9650SE-24	WD RE	22	6	20TB	60	1320
2009 A	LSI 8888-ELP	WD RE4-GP	22	6	38TB	60	1320
2009 B	Adaptec 52445	Hitachi	22	6	40TB	38	836
						<b>497</b>	<b>8790</b>

- Servers are assigned to a Service Classes
  - Each Virtual Organisation (VO) has several Service Classes
  - Usually between 2 and many servers per service class
- Service Classes are either:
  - D1T1: data is on both disk and tape
  - D1T0: data is on disk but NOT tape
    - If lost, VO gets upset :-)
  - D0T1: data is on tape - disk is a buffer
  - D0T0: ephemeral data on disk
- Service Class type assigned VO depending on their data model and what they want to do with a chunk of storage

- Want to make sure data is as secure as possible
- RAID5 is not secure enough:
  - Only 1 disk failure can put data at risk
  - Longer period of risk as rebuild time increases due to array/disk sizes
  - Even with host spare, risk of double failure is significant
- Keep D1T0 data on RAID6 systems
  - Double parity information
  - Can lose two disks before data is at risk
- Tape buffer systems need throughput (network bandwidth) not space
  - Use smaller capacity servers in these Service Classes





# Disk Statistics

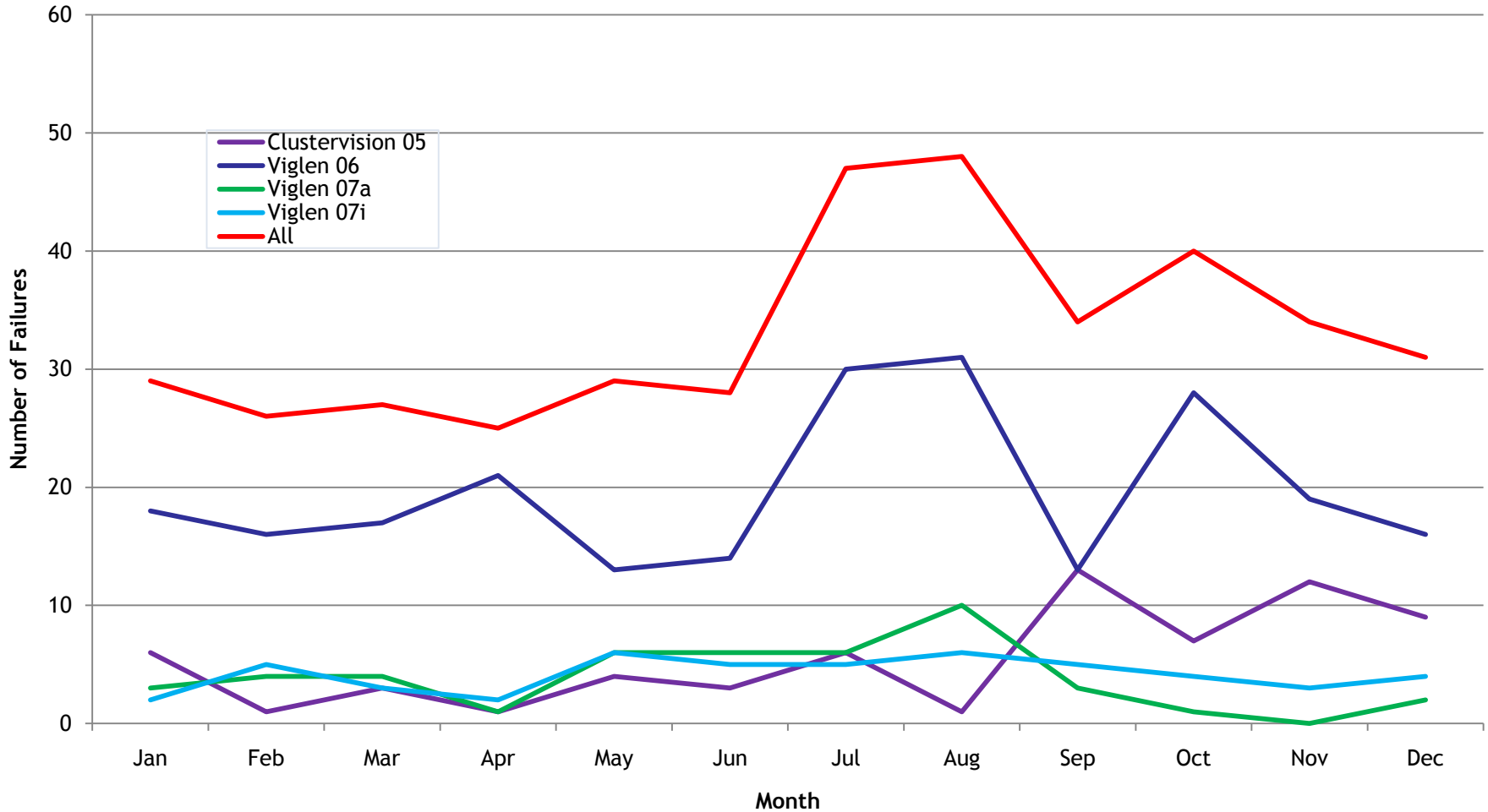


- Drives failed / changed: 398 (9.4% !)
- Multiple failure incidents: 21
- Recoveries from multiple failures: 16
- Data copied to another file system: 1
- Lost file systems: 4

Month	Clustervision 05	Viglen 06	Viglen 07a	Viglen 07i	WD 500GB	WD 750GB	Seagate 750GB	All
Jan	6	18	3	2	24	3	2	29
Feb	1	16	4	5	17	4	5	26
Mar	3	17	4	3	20	4	3	27
Apr	1	21	1	2	22	1	2	25
May	4	13	6	6	17	6	6	29
Jun	3	14	6	5	17	6	5	28
Jul	6	30	6	5	36	6	5	47
Aug	1	31	10	6	32	10	6	48
Sep	13	13	3	5	26	3	5	34
Oct	7	28	1	4	35	1	4	40
Nov	12	19	0	3	31	0	3	34
Dec	9	16	2	4	25	2	4	31
<b>Total</b>	<b>66</b>	<b>236</b>	<b>46</b>	<b>50</b>	<b>302</b>	<b>46</b>	<b>50</b>	<b>398</b>
Average/month	5.50	19.67	3.83	4.17	25.17	3.83	4.17	33.1667
Base	462	1204	1274	1274	1688	1274	1274	4214
% fail	14.3%	19.6%	3.6%	3.9%	17.9%	3.6%	3.9%	9.4%

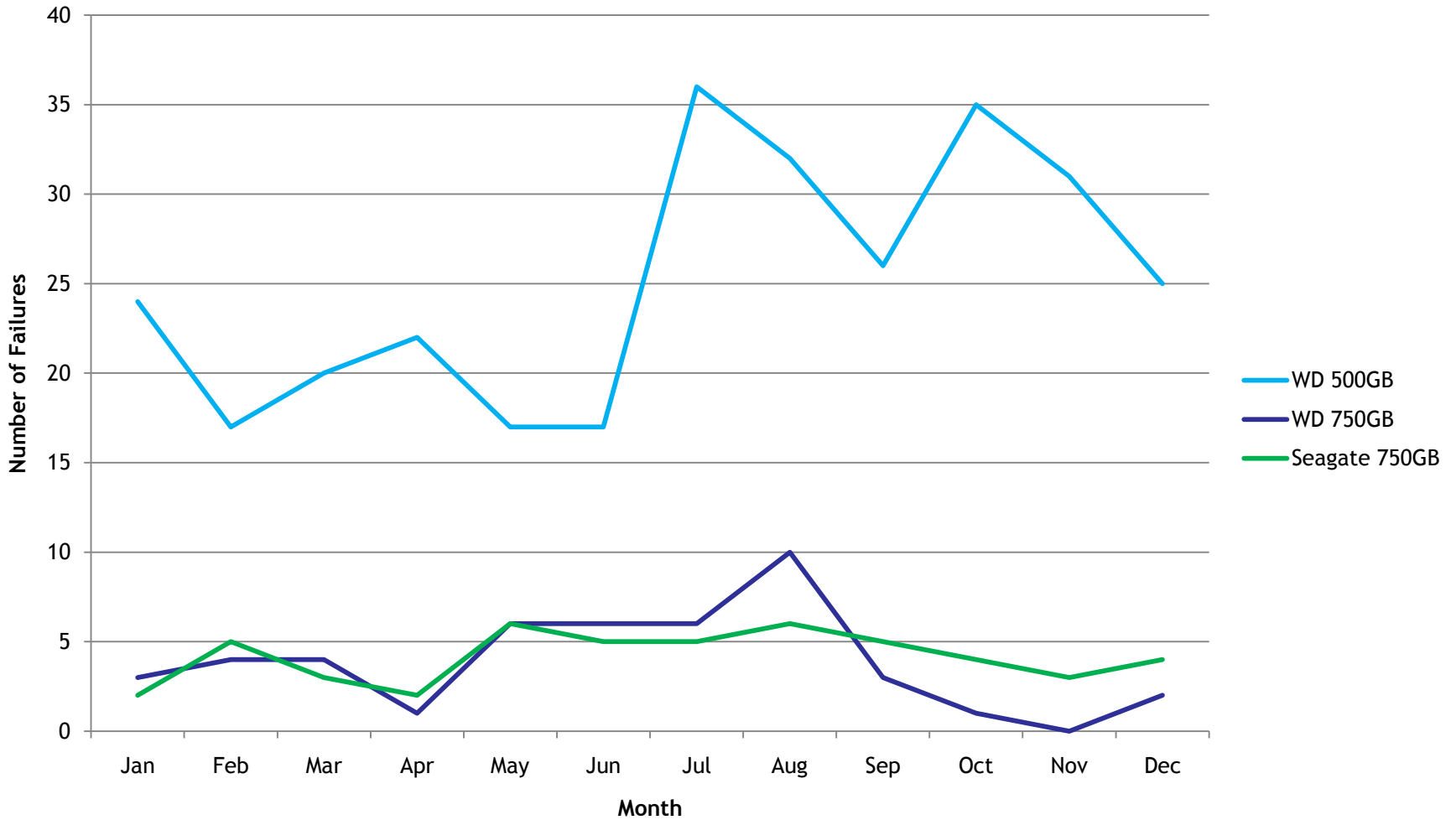


## Drive Failures by Generation - 2009



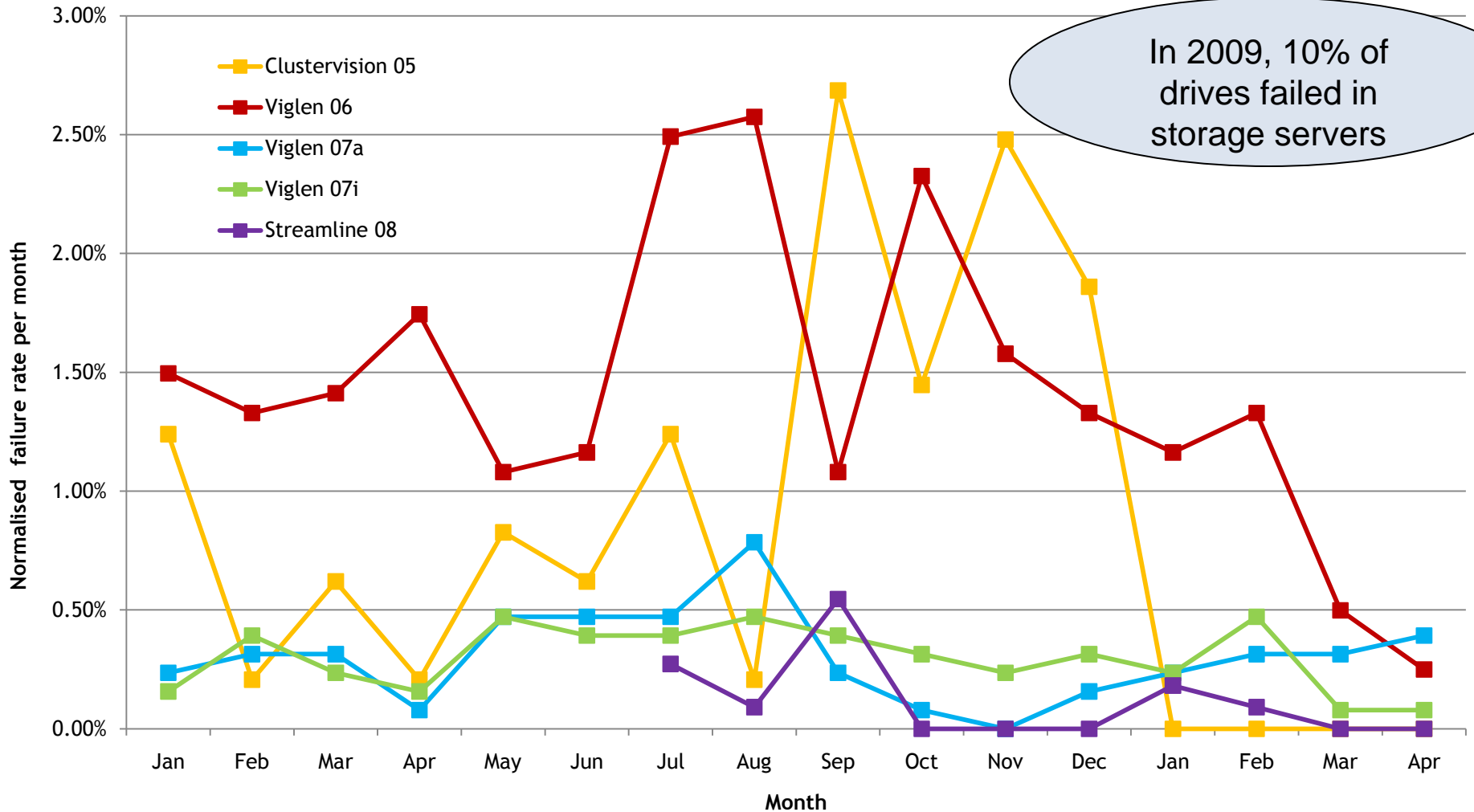


## Dive Failures by Drive Type - 2009





## Failure Rate Normalised by Installed Base 2009-10





Days	Month	Clustervision 05	Viglen 06				Viglen 07a				Viglen 07i				All				
			Saved	Copied	Lost	Total Doubles	Saved	Copied	Lost	Total Doubles	Saved	Copied	Lost	Total Doubles	Saved	Copied	Lost	Total Doubles	
31	Jan	0	1																
28	Feb	0	1			1													
31	Mar	0	1					1											
30	Apr	0	1							1									
31	May	0																	
30	Jun	0																	
31	Jul	0	1			1													
31	Aug	0	4	1		1				1									
30	Sep	0																	
31	Oct	0	2																
30	Nov	0																	
31	Dec	0																	
365	Total	0	11	1	4	16	3	0	0	3	2	0	0	2	16	1	4	21	

Average/  
month 0.00

1.33

0.25

0.17

1.33

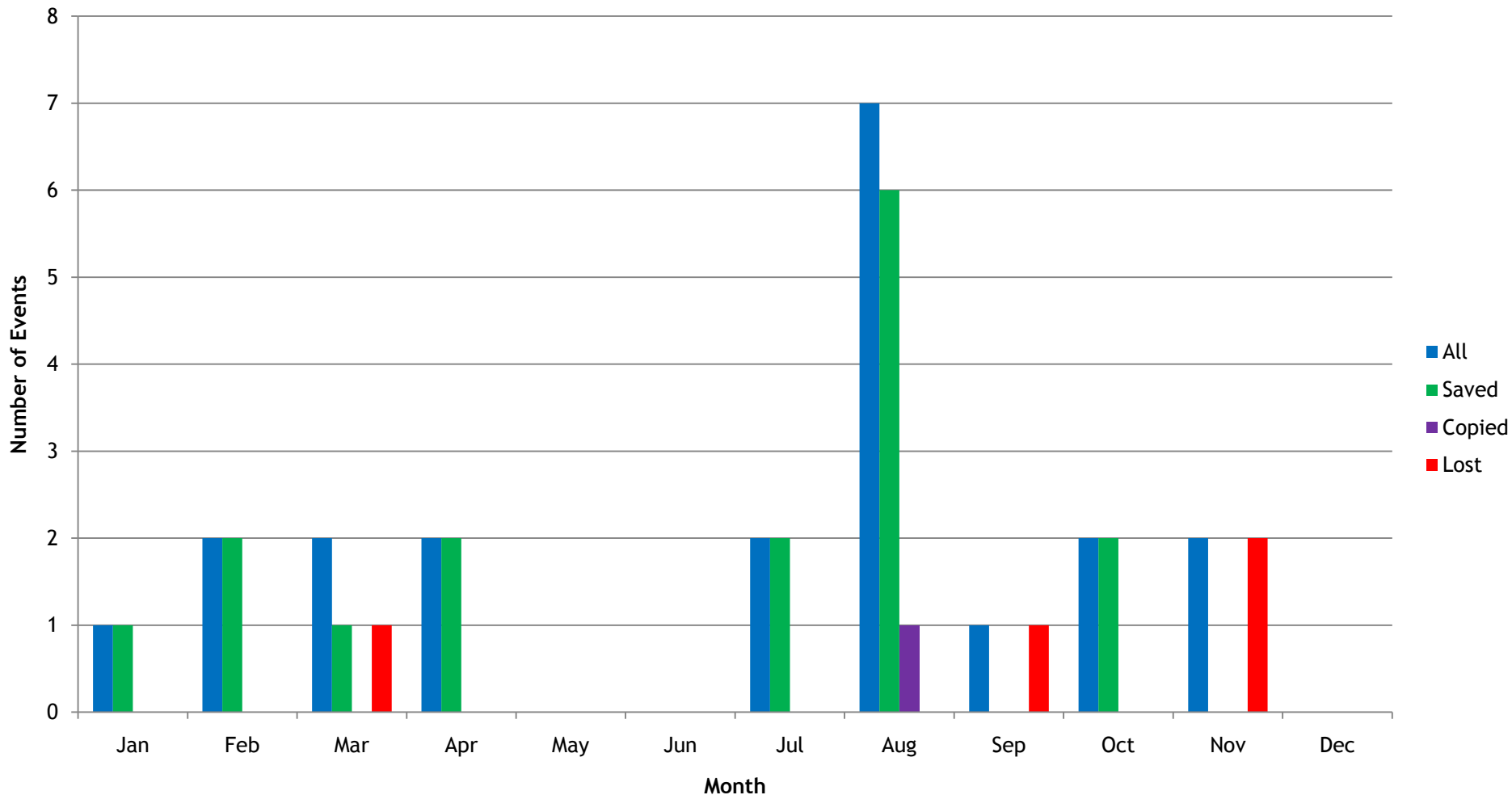
0.08

0.33

1.75



## Multiple-disk failure events 2009

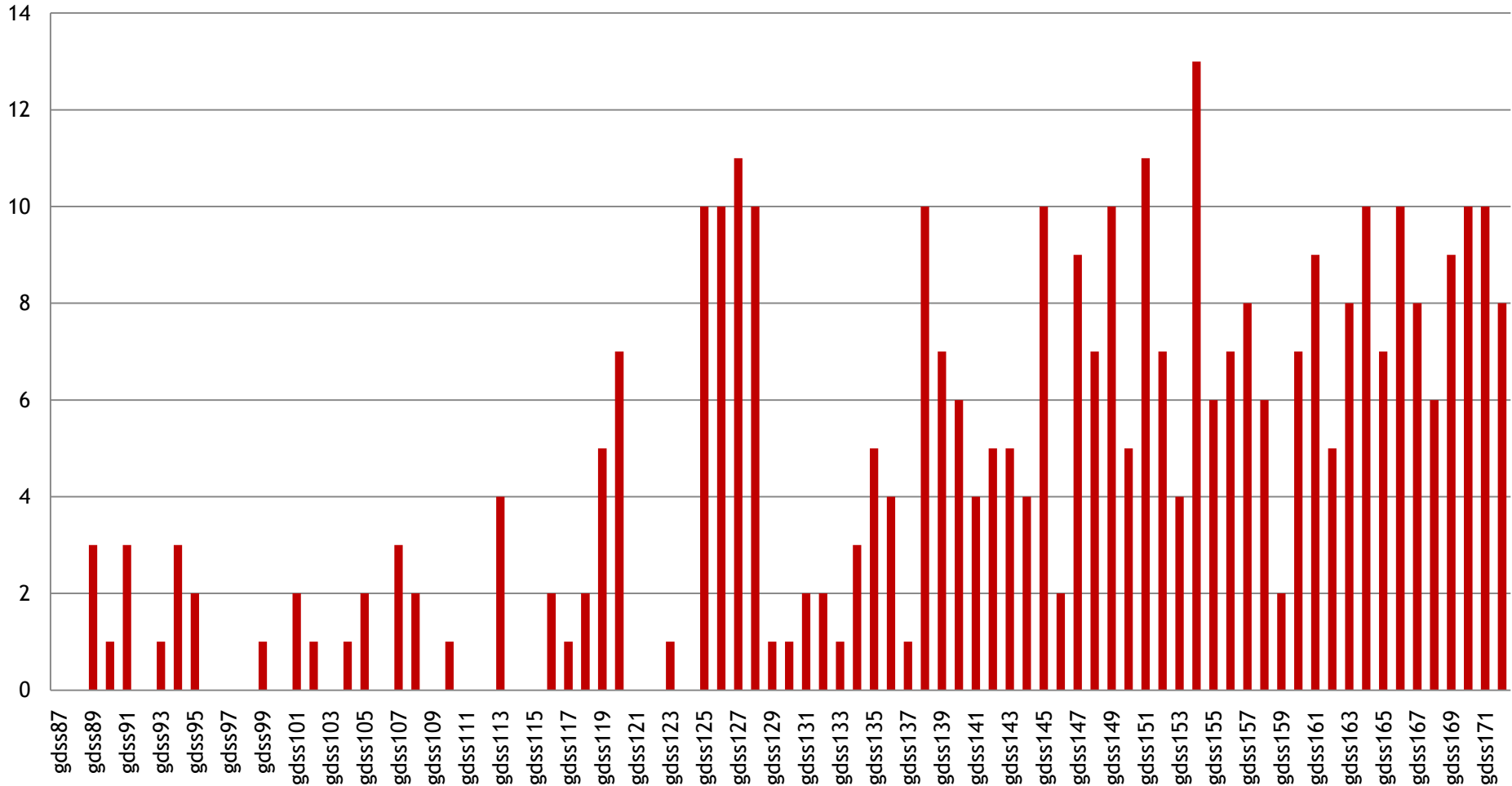




- Why are Viglen 06 servers less reliable?
- RAID5?
  - More vulnerable to double disk failure causing a problem
- Controller issues?
  - Yes: Failure to start rebuilds after a drive fails
  - Yes: Some cache RAM issues
- System RAM?
  - Yes: ECC not set up correctly in BIOS
- Age?
  - Possibly: difficult to assert with confidence
- Disks?
  - Hmmmm....

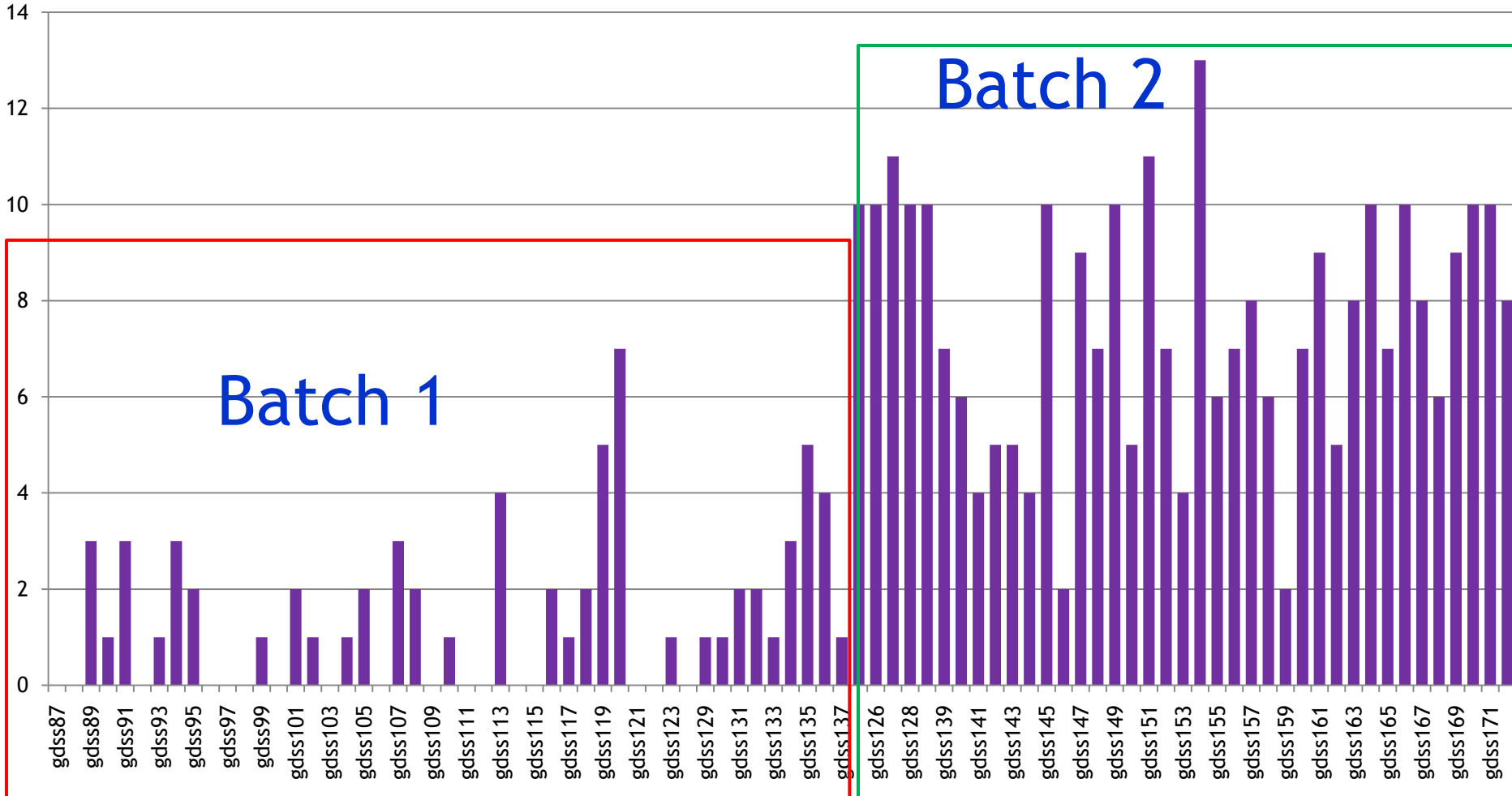


## Disk failures - Viglen 06 - June '08 to April '10





## Viglen Disk Data - Sorted



- Overall failure rate of drives is high (9.4%)
- Dominated by WD 500GB drives (19.6%)
- Rate for 750GB drives much lower (3.75%)
- Spike in failures:
  - at the time of the Tier1 migration to R89
  - when air conditioning in R89 failed
- Batch effect for Viglen 06 generation

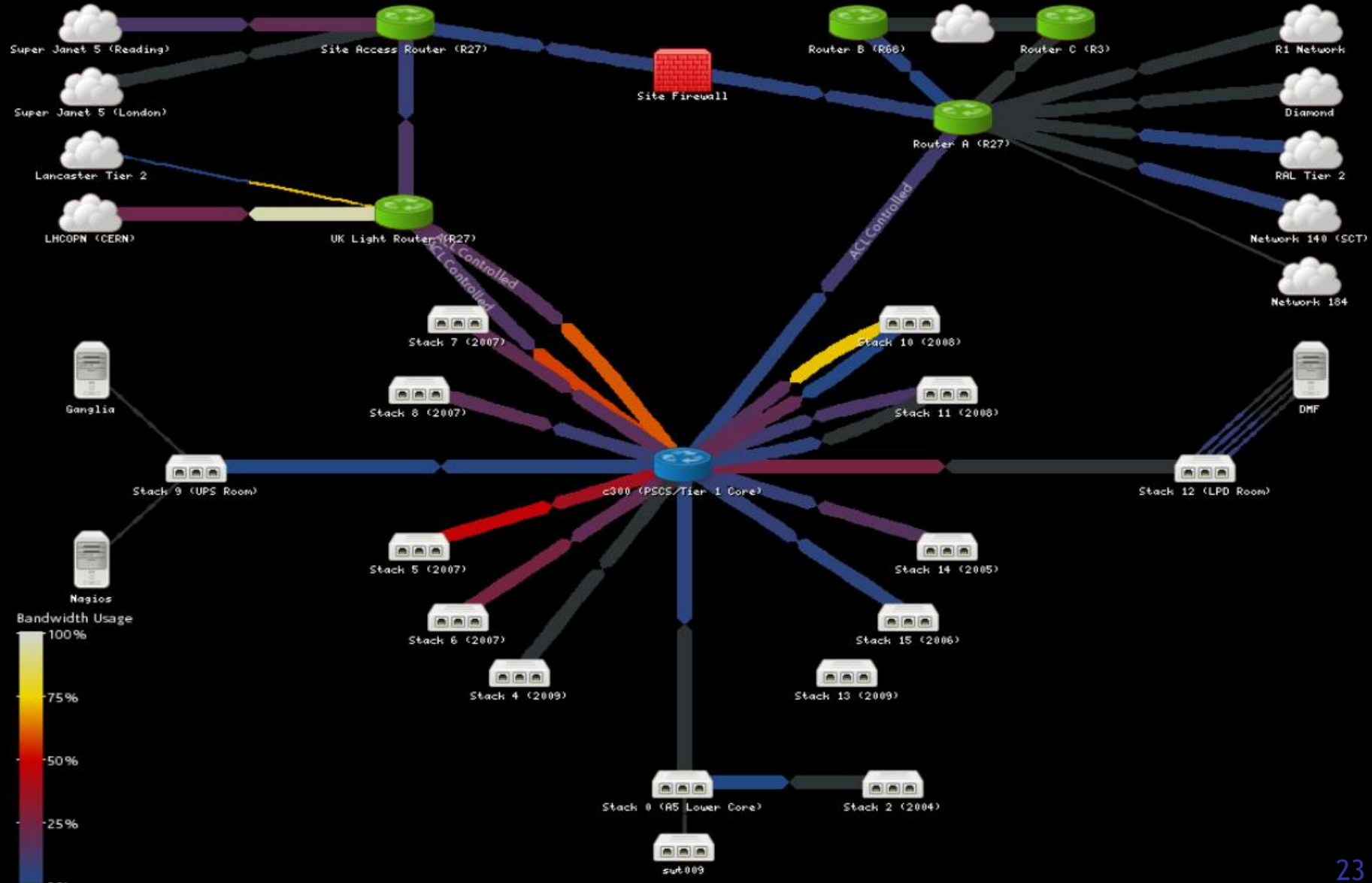
# Networking

- Data network
  - Central core switch: Force10 C300 (10GbE)
  - Edge switches: Nortel 55xx/56xx series stacks
  - Core-to-Edge: 10GbE or 2 x 10GbE
  - All nodes connected at 1 x GbE for data
- Management network
  - Various 10/100MbE switches: NetGear, 3Com
    - New and salvaged from old data network
- Uplinks
  - 10Gb/s to CERN (+ failover)
  - 10Gb/s to Site LAN
  - 10Gb/s + 10Gb/s failover to SJ5 (to be 20Gb/s each soon)
  - Site routers: Nortel 8600 series

RAL Tier 1 Network Load

Updated: Apr 07 2010 14:45:17

RAL Tier 1 Network Load

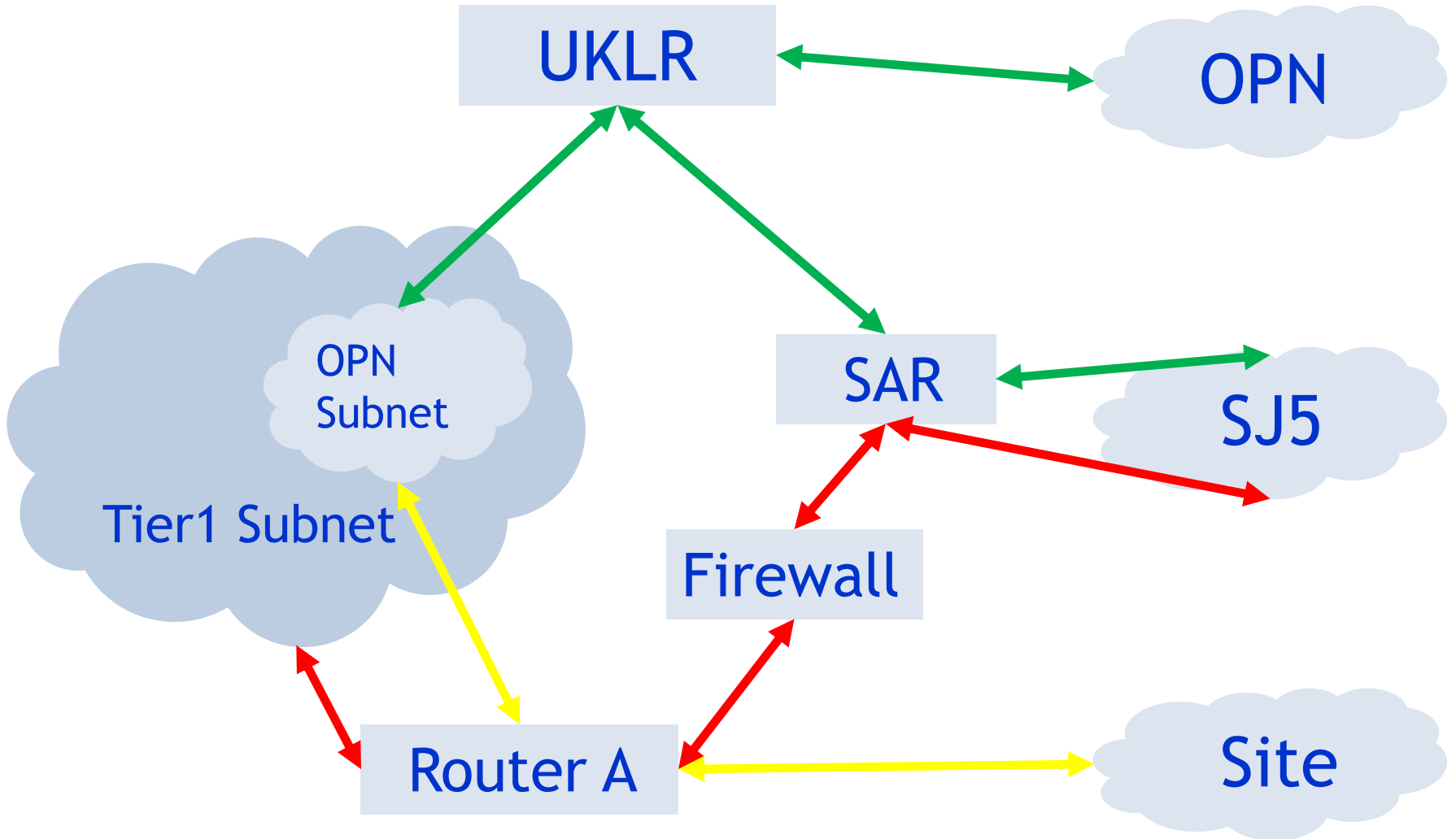


- Tier1 has a class 21 network within the site range
  - ~2000 addresses
- Need to have a dedicated IP range for the LCHOPN
  - Class 23 network within existing subnet
  - ~500 addresses, Castor disk servers only

130.246	176.0					
130.246	177.0					
130.246	178.0			T1-Extra		
130.246	179.0		OPN		Tier1	
130.246	180.0					
130.246	181.0			T1-Original		
130.246	182.0					
130.246	183.0					
130.246	184.0					
130.246	185.0			T1 Dev		
130.246	186.0					
130.246	187.0					
130.246	188.0					
130.246	189.0			SSD		
130.246	190.0					
130.246	191.0					



- General nodes
  - Gateway address on site router
  - All traffic goes to Router A
    - From there to site or off site via firewall and Site Access Router (SAR)
- OPN nodes
  - Gateway address on UKLight router
  - Special routes to Router A for site only traffic
  - All off-site traffic to UKLight router
- UKLight router
  - BGP routing information from CERN
  - T0, T1<->T1 traffic directed to CERN link
  - Lancaster traffic -> link to Lancaster
  - Other traffic from RAL Tier1 up to SAR (the bypass)
- SAR
  - Filters inbound traffic for Tier1 subnet
    - OPN machines to Core via UKLight router
    - Others to Router A via Firewall
    - Non-Tier1 traffic not affected



- Firewall is a NetScreen
  - 10Gb/s linespeed
- Site FW policy:
  - closed inbound except by request on port/host (local/remote) basis
  - Open outbound
  - Port 80 + kin redirected via web caches...
- Tier1 subnet:
  - As Site except port 80 + kin do not use web caches
- OPN subnet:
  - Traffic is filtered at SAR - only Castor ports open