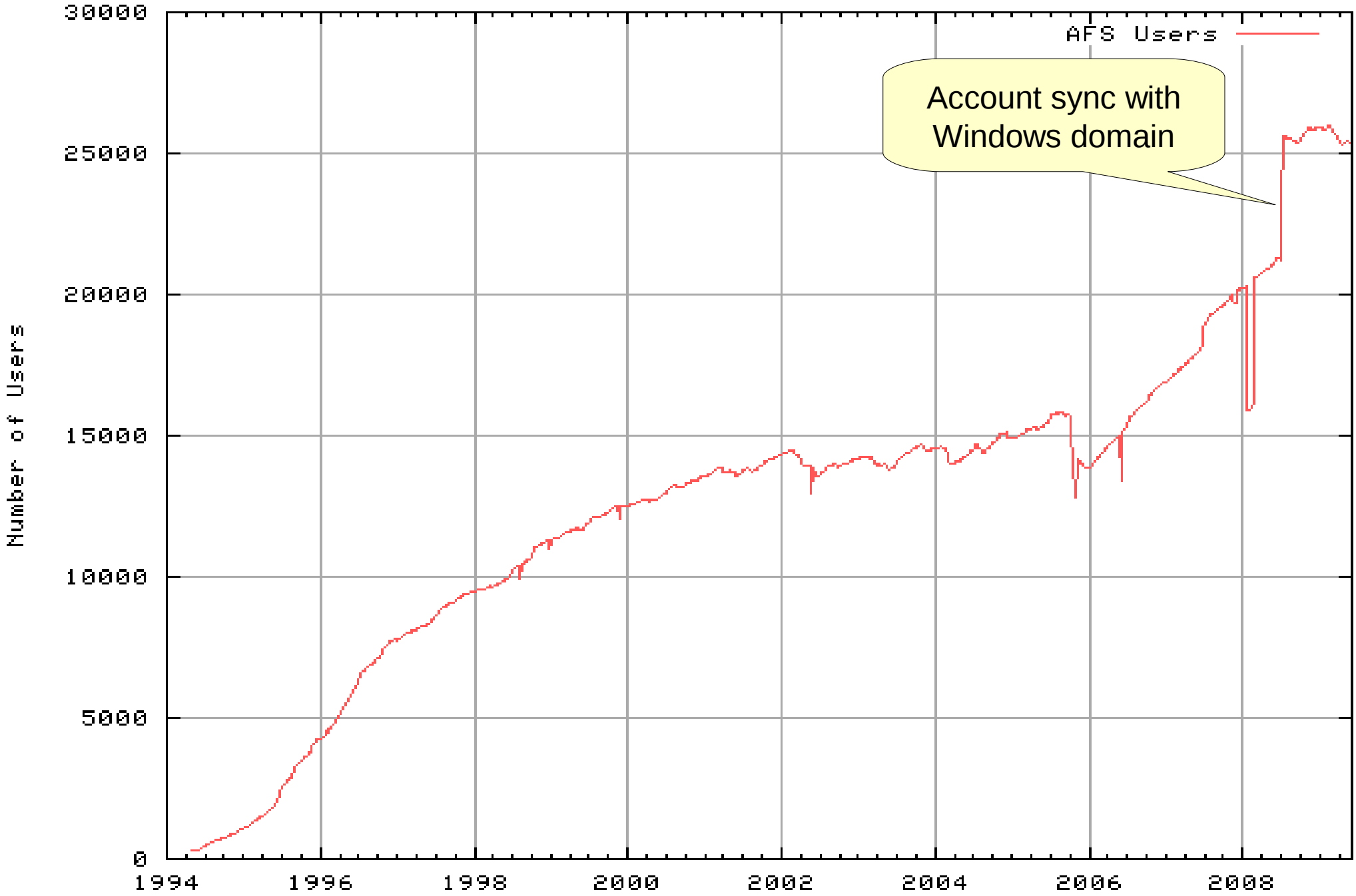
2009 AFS & Kerberos Workshop
CERN Site Report

KELEMEN Péter
CERN IT FIO

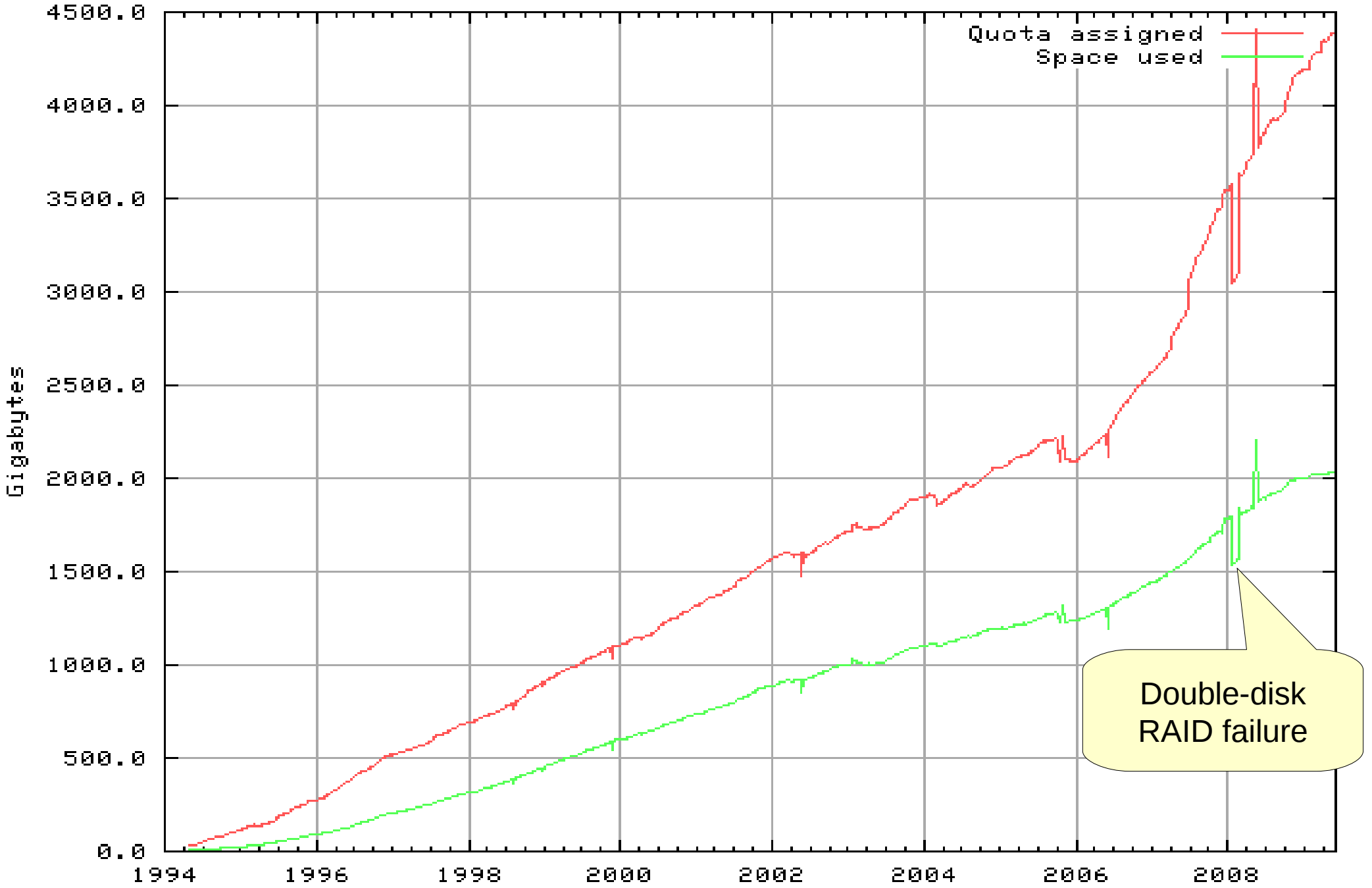
CERN AFS in numbers

- **25'500** AFS users
- **48'779** AFS volumes
- **407.8 million** files (avg 65 KiB)
- **25/112 TiB** space (used/total)
- **1500 million** accesses/day (avg)
- **40** AFS servers in production
- **13'000** active AFS clients
- **325:1** client/server ratio

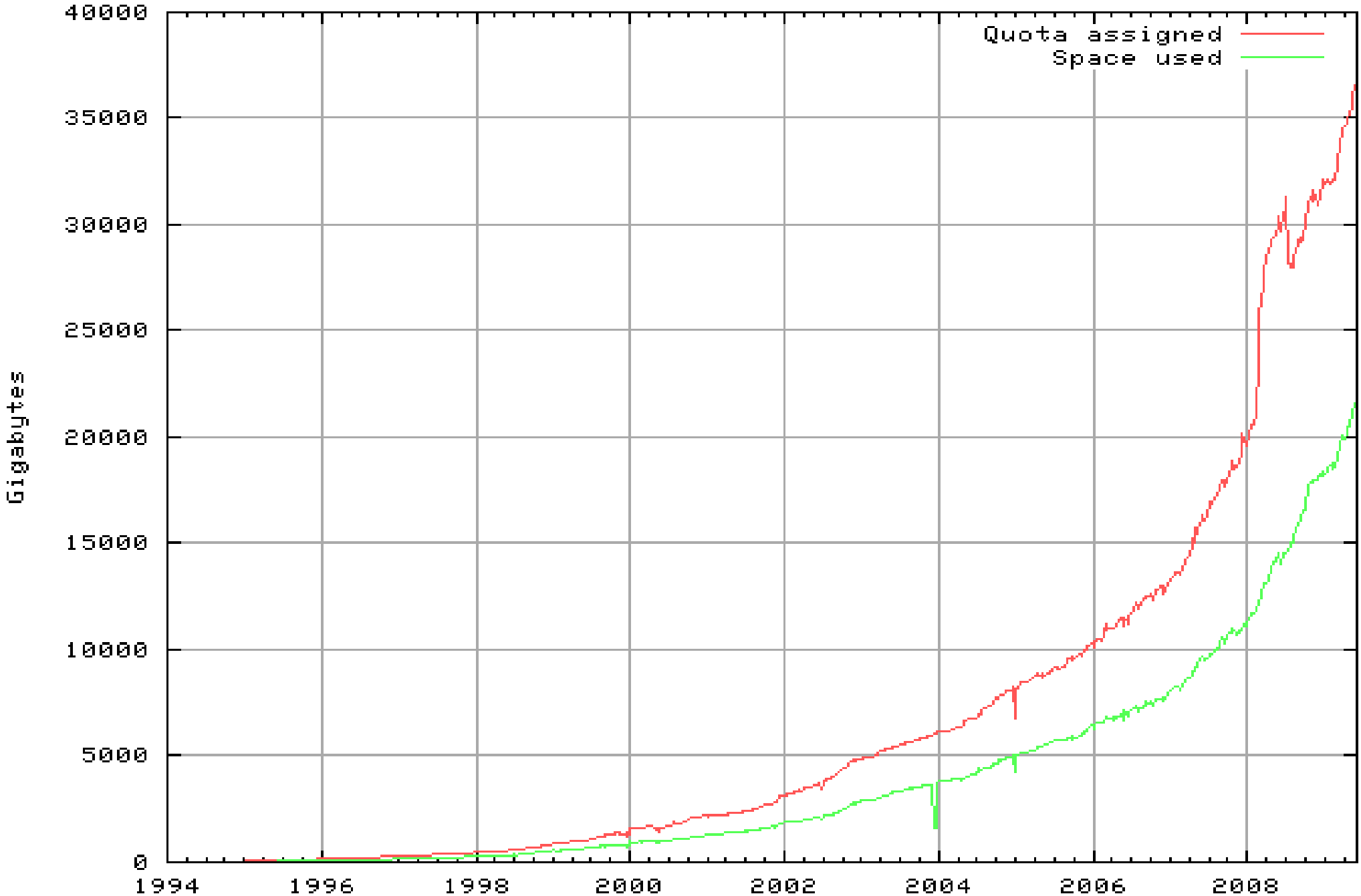
of users


AFS Users
Account sync with Windows domain
Number of Users
30000
25000
20000
15000
10000
5000
0
1994
1996
1998
2000
2002
2004
2006
2008

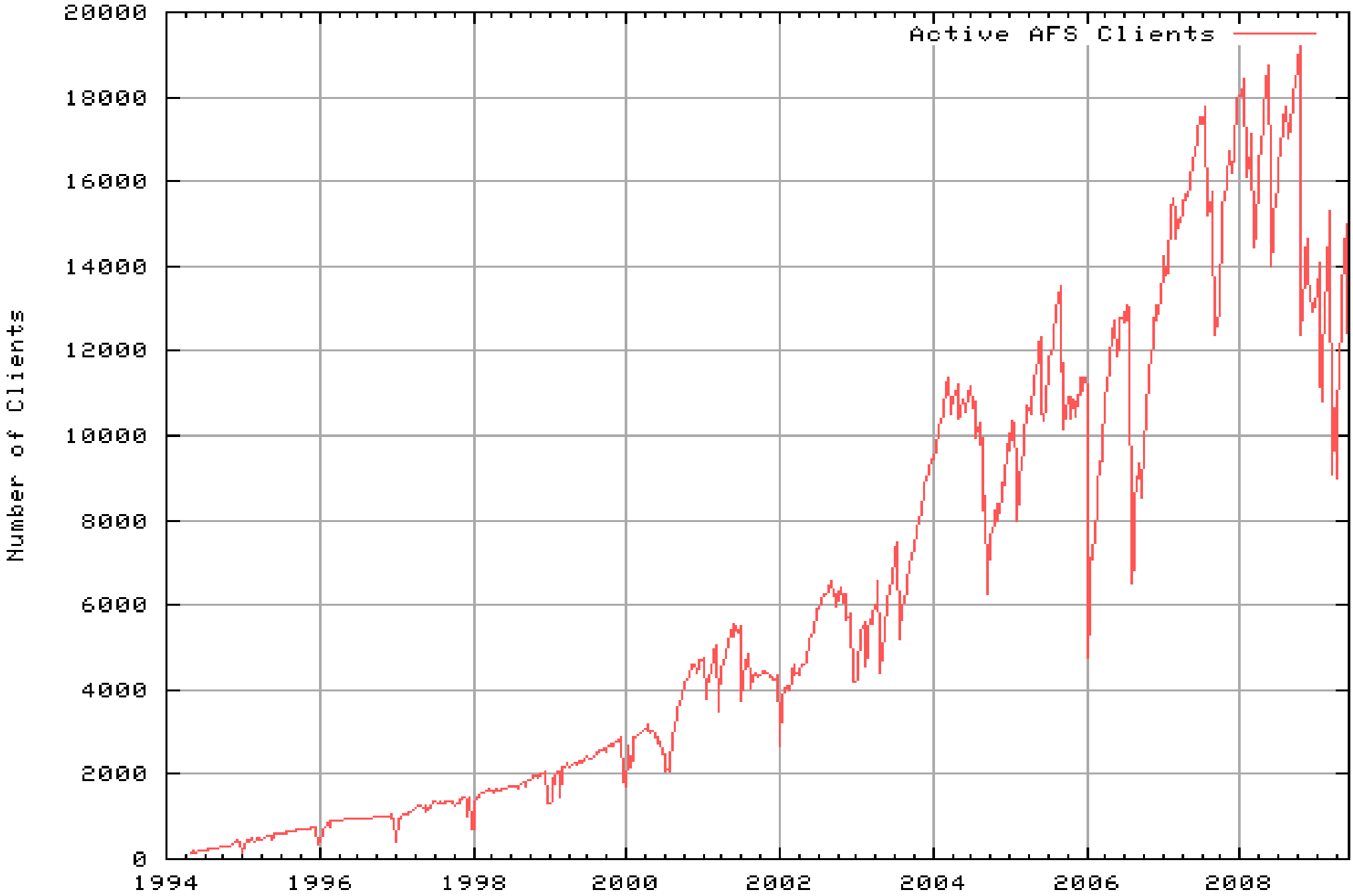
Home directories



Project space



of clients

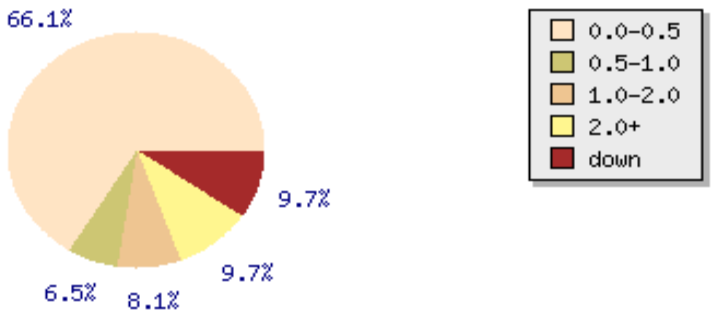

Active AFS Clients
Number of Clients
20000
18000
16000
14000
12000
10000
8000
6000
4000
2000
0
1994
1996
1998
2000
2002
2004
2006
2008

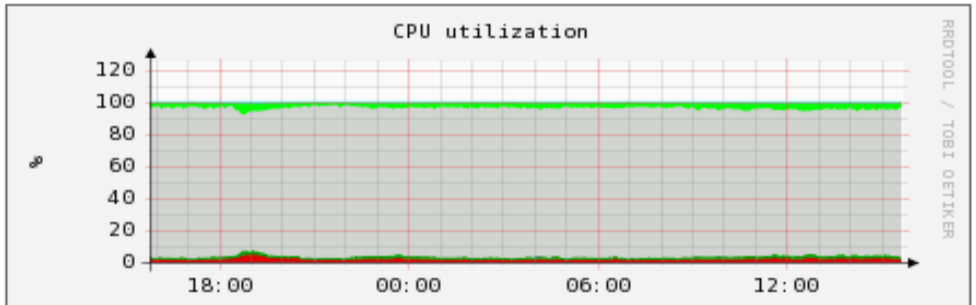
Information for group of hosts: afs

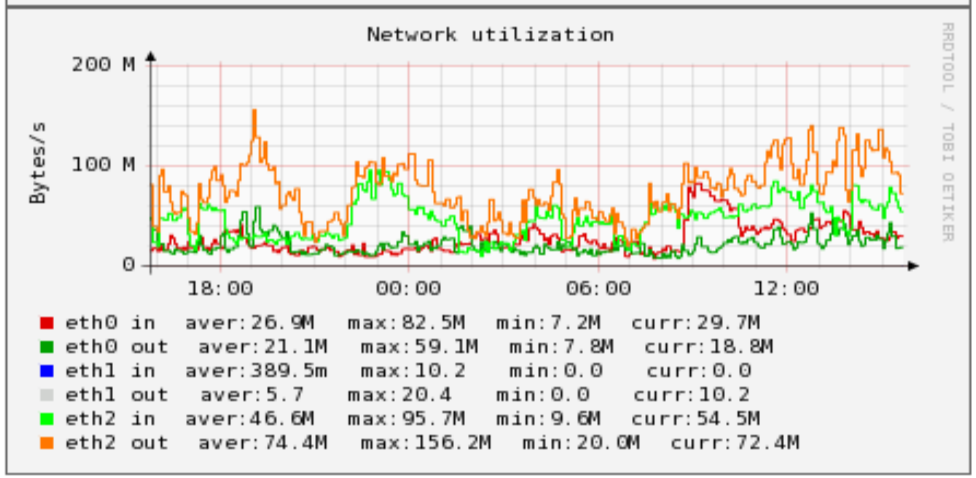
Cluster information

number of hosts (down)	81 (0)
operating system(s)	Scientific Linux SL release 5.1 (Boron), Scientific Linux SL release 5.2 (Boron), Scientific Linux CERN SLC release 5.2 (Boron), Red Hat Enterprise Linux Server release 5.2 (Tikanga), Red Hat Enterprise Linux ES release 4 (Nahant Update 6), Red Hat Enterprise Linux Server release 5.1 (Tikanga), Red Hat Enterprise Linux ES release 4 (Nahant Update 4), Red Hat Enterprise Linux Server release 5.3 (Tikanga), Scientific Linux CERN SLC release 4.4 (Beryllium), Scientific Linux CERN SLC release 4.5 (Beryllium), Scientific Linux CERN SLC release 5.3 (Boron)
average of up times	**269 days, 21h:55m**
hosts down	**afs87, afs90, afsdb6, afsdb7, afsdb8, lxfsrd4008**
select from hosts	None

Load average distribution



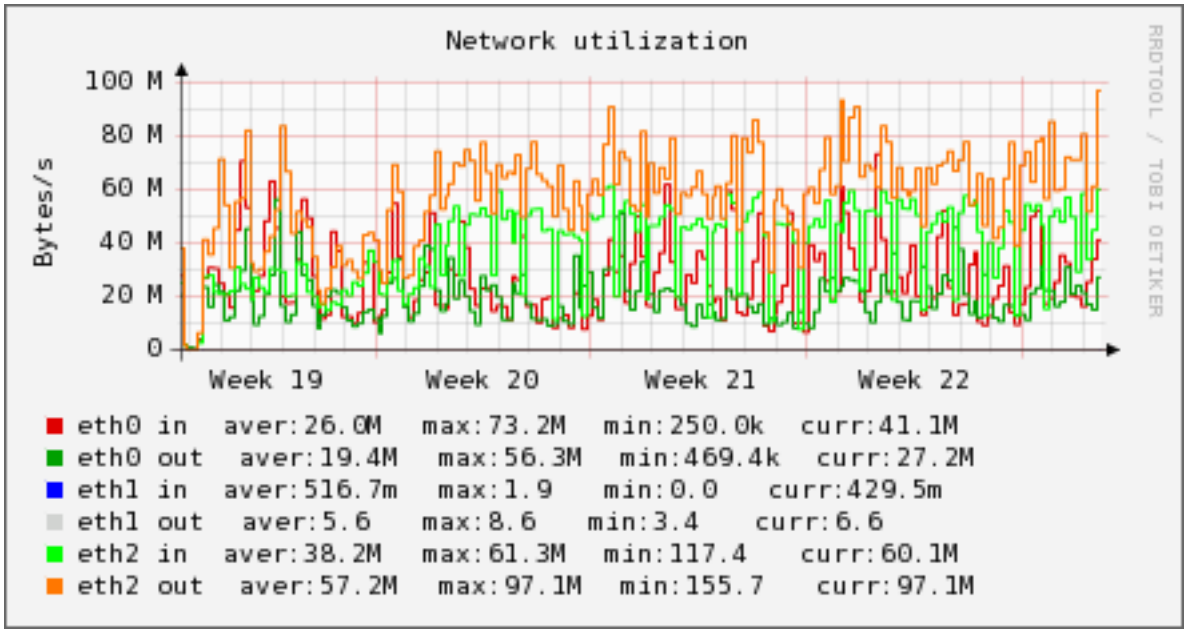


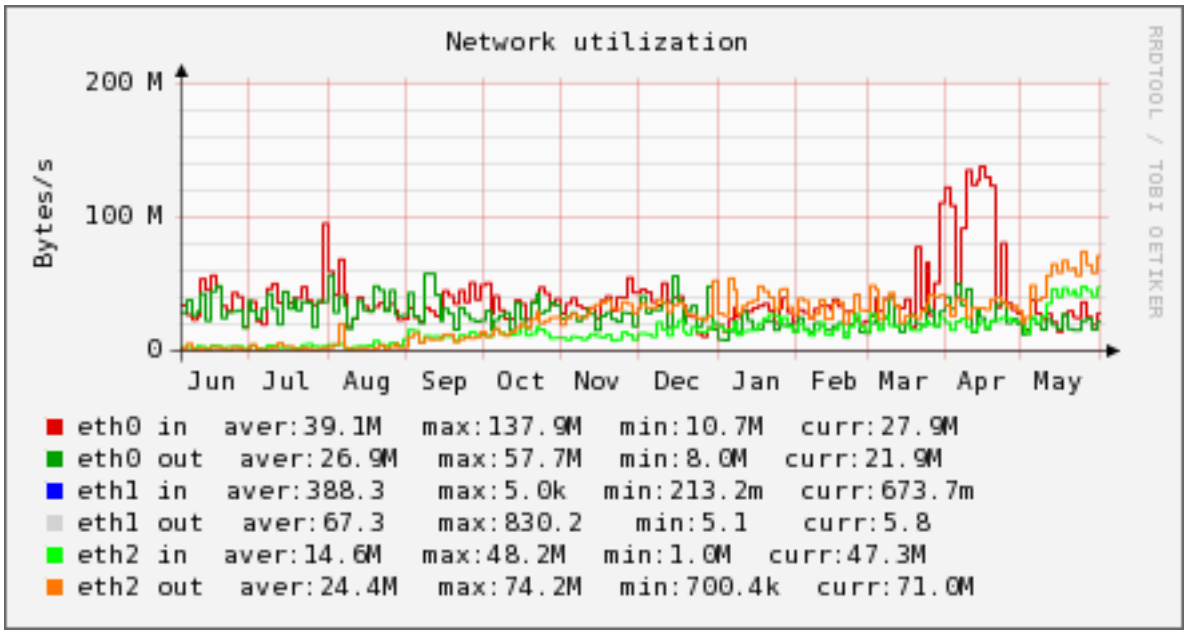


Set span: day **offset:** null

Search entities: Virtual Clusters | Clusters | Racks | Hardware types | Virtual organizations | Services

Traffic... last month, last year



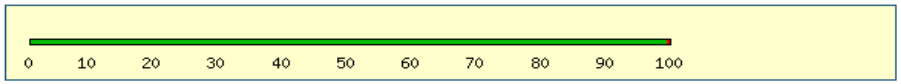


Fabric Infrastructure and Operations

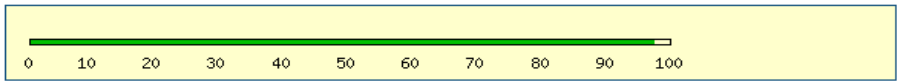
CERN AFS "Console"

CERN IT Department

AFS Availability



AFS Available Performance



AFS Console's Promptness

Partition related information is from 2009-06-03/15:46:29
Volume related information is from 2009-06-03/15:39:57

AFS Alarms / Warnings

afs43 has a local response time for 64Kb of 55.81 ms (avg last hour)
List problematic volumes
List hottest volumes

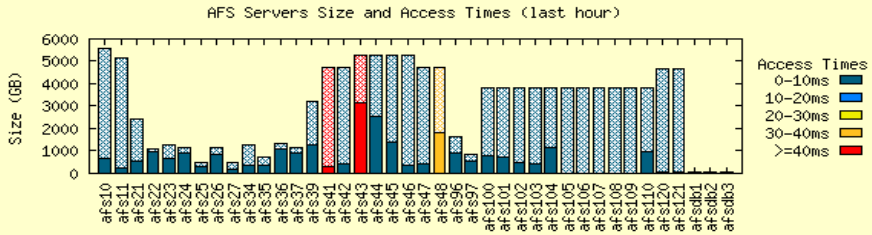
AFS Statistics

Lemon
SLS

Report Generator

Volumes Report by server by project global
Servers Report
Partitions Report
Pool Report

Servers Overview



Projects Overview

atlas alice cms lhcb compass gd swlcg user sw cvs afs

AFS Historic Monitoring Data

Volume Statistics / Volume History
Partition Statistics / Partition History

Service Incidents

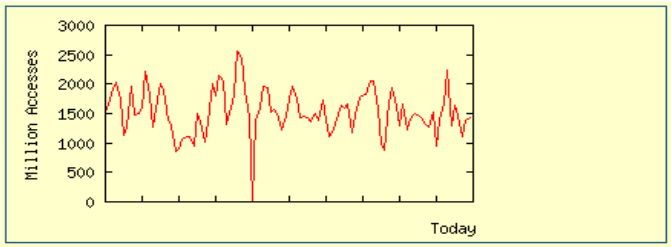
Get Incidents Log

AFS CERN Cell Logs

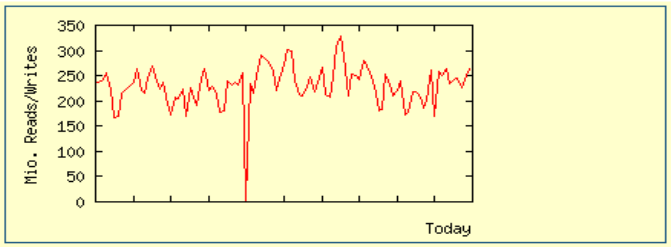
Simple entry Log...
Complete Log...

Get Log from date: 2009-06-03 Submit

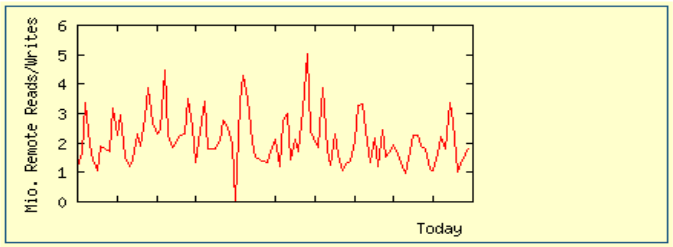
AFS Total Accesses (Last 100 days)



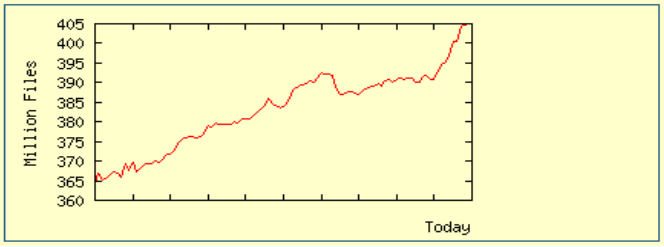
AFS Total Reads+Writes (Last 100 days)



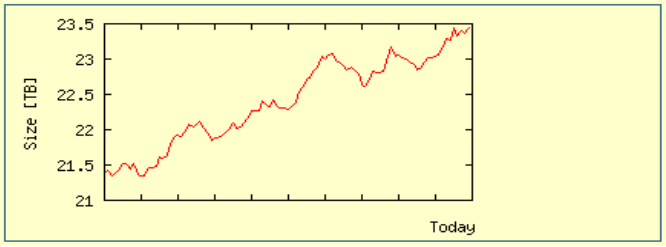
AFS Remote Reads+Writes (Last 100 days)



AFS Files (Last 100 days)



AFS Used Space (Last 100 days)



Example: hot volumes

Timestamp	Volume Id	Volume Name	Server	Partition	Type	Size [kByte]	Files	Status	Quota [kByte]	Current Accesses [1/s]	Current Reads [1/s]	Current Writes [1/s]	Current Reads+Writes [1/s]	Total Accesses	Total Reads	Total Writes	Project
2009-06-03/15:39:57	537699722	s.atlas.lcg3.Re-bu3	afs10	d	RW	1624597	64880	On	4800000	326.753	10.259	208.725	**218.984**	796161	39539	323593	atlas
2009-06-03/10:06:29	537252478	p.cvs.atlas	afs41	am	RW	8940614	209078	On	16679745	1060.587	202.868	2.56	**205.428**	6046209	769488	10045	cvs
2009-06-03/15:39:57	1933854940	q.atlas.b.mi10.Ev1	afs42	eo	RW	850416	63151	On	1500000	440.65	11.637	148.814	**160.451**	1194241	49592	475257	atlas
2009-06-03/15:39:57	1933843193	q.atlas.b.mi7.Ev1	afs48	bl	RW	2578475	64342	On	3000000	390.812	17.829	110.604	**128.433**	946433	42889	320207	atlas
2009-06-03/15:39:57	537419062	user.rlyzwa	afs22	fa	RW	90160	3766	On	100000	164.691	7.514	100.69	**108.204**	964737	51191	57535	user
2009-06-03/15:39:57	537615019	p.atlas.gd.lcg.readonly	afs37	aa	RO	95683017	2312930	On	100000000	646.454	97.633	0	**97.633**	1325825	148348	0	atlas
2009-06-03/15:39:57	1933873955	s.atlas.b.151XY-rVl3	afs100	d	RW	3432376	36699	On	4000000	251.1	10.298	83.945	**94.243**	1189889	97061	283557	atlas
2009-06-03/15:39:57	537615019	p.atlas.gd.lcg.readonly	afs27	cc	RO	95683017	2312930	On	100000000	594.824	89.451	0	**89.451**	1298305	145873	0	atlas
2009-06-03/15:39:57	537402497	user.fernando	afs22	fa	RW	265263	6833	On	300000	180.467	.001	58.83	**58.831**	5766529	3541	1865928	user
2009-06-03/15:39:57	1933843670	p.atlas.releas.14225	afs48	cl	RW	21155308	348811	On	25000000	576.299	50.398	0	**50.398**	121691521	5974615	0	atlas
2009-06-03/13:35:33	537547478	q.cms.slc4_ia_32_c	afs45	bj	RW	10019893	414341	On	15000000	440.478	42.851	0	**42.851**	18105217	964635	2	cms
2009-06-03/15:39:57	537721998	u.atlas.atlasdqm.0	afs44	ia	RW	98483882	11938157	On	100000000	2713.337	33.256	8.108	**41.364**	12012929	150330	470095	atlas
2009-06-03/15:39:57	537202112	p.alice.offline_v02	afs45	bj	RW	3108648	144625	On	4000000	262.334	9.93	30.971	**40.901**	1855489	49496	296789	alice
2009-06-03/15:39:57	537529366	user.vmiccio	afs35	cg	RW	142824	4768	On	200000	113.539	0	37.556	**37.556**	6233857	2	2076006	user
2009-06-03/15:39:57	1933879341	q.atlas.152XYPr-V3	afs48	bl	RW	3468086	25500	On	7500000	155.13	1.859	30.036	**31.895**	2575617	65921	43373	atlas
2009-06-03/15:39:57	537693933	p.lsf.logdir	afs46	dk	RW	28363750	2164	On	40000000	87.365	28.694	.444	**29.138**	5075329	1663791	22496	lsf
2009-06-03/15:39:57	537286803	p.cvs.lhcb	afs46	dk	RW	1784219	103751	On	2297479	121.905	18.753	9.669	**28.422**	5091969	775074	329116	cvs
2009-06-03/15:39:57	537547506	q.cms.slc4_ia32_l	afs45	bj	RW	6906604	179413	On	12000000	437.303	18.711	0	**18.711**	22479105	831940	12	cms
2009-06-03/15:39:57	537688416	q.atlas.d.An-dvl3	afs26	fb	RW	6662392	98587	On	13000000	243.092	18.61	0	**18.61**	7892481	275675	1355749	atlas
2009-06-03/15:39:57	1933841276	q.lhcb.LHCB_v26r0	afs43	hc	RW	10182245	72315	On	10190000	100.273	18.436	0	**18.436**	159489	28025	0	lhcb

Last year changes

- All servers run OpenAFS **1.4.8**
  - Almost all servers run various flavors of Linux (RHEL, SL, SLC)
  - Single Solaris box remain for restores (replacement planned)
- Backup user/project vol's to **TSM**
- KDCs run **Heimdal 1.2**
  - CERN clients however use MIT (RedHat Enterprise, 1.3.4/1.6.1)
- Kerberos 4 phase-out
  - Not yet complete, dealing with the “long tail”
- XFS used for scratch volumes
- Few SSDs in production
  - STEC MACH8iops, 3000 wIOPS sustained

Future outlook

- AFS service is very stable
  - Few incidents that are handled efficiently: largely automated procedures
  - Service is “critical” but runs without a piquet
- Currently 2.75 FTE, go to 1.75?
- Massive space expansion?
  - At least 1.5 PiB, possibly more
- Active Directory integration?
  - Currently we control the KDCs → batch authentication with `arc`
  - Windows realm is called CERN.CH as well...
  - Possible unification of realms under AD – hot discussion

Current issues and worries

- Max. number of servers in cell
  - Currently avg. 2.8 TiB/server → 1.6 PiB/250 servers = 6.4 TiB/server
- Rx latency issues
  - Scalability tests show that massive load causes too many retransmits
  - (in 2008) 2500 clients R/W reqs to a dual-GigE fileserver: 1.5 GiB/s
- Users want large volumes
  - “I go to (insert your local electronics store here) and buy a terabyte!”
  - Current max. allowed is 250 GiB (a few volumes exceptionally larger)
  - Painful for backup, moving around for load balancing (partition size!)
- Volume corruptions
  - Investigation in progress, long-term issue

- Kerberos replay cache too slow*