

Center for Information Services and High Performance Computing (ZIH)

Monitoring and System Management @ ZIH

TIMaCS Workshop Berlin 10-2009

Zellescher Weg 16 Trefftz-Bau (HRSK-Anbau) Room HRSK/151 Tel. +49 351 - 463 - 39871

> Center for Information Services & High Performance Computing

Guido Juckeland (guido.juckeland@tu-dresden.de)

Monitoring concepts

Bootup checks

Runtime checks

Do not start services, if not all requirements are met

> ECHNISCHE NIVERSITAT



Monitor Service states (Nagios)



Guido Juckeland – Slide 2

Bootup checks

 Servers do not start the service(s) they provide automatically → instead run sanity check that starts the service(s) if all checks out ok

deimos101.hrsk.tu-dresden.de - PuTTY	
p2s223:~ # /etc/init.d/node startup start	A
Node startup check	
O. waiting 30 seconds for IB to settle	done
1. Checking IB Port 1 cable	done
2. Checking IB Port 1 speed	done
3. Checking IB Port 2 cable	done
4. Checking IB Port 2 speed	done
5. mounting /opt/cluster	done
6. fixing default route	done
7. loading missing kernel modules	done
8. fixing resolv.conf	done
9. checking ypbind	done
10. restarting ntp	done
11. restarting postfix	failed
12. Checking if I am a special node:	
Am I a login node? No	done
Am I a Lustre MDS server? No sandbox server node No	done
Am I a Lustre OSS server? No	done
Am I node 1sf1? No	done
Am I node 1sf? No	done
13. mounting Lustre file system	done
14. mounting SGI NFS	done
15. checking number of cores (found 2 cores)	done
16. checking that /tmp is writable	done
17. setting dimm labels	done
18. checking disk size (found 2 cores, 74 GB disk)	done
19. changing permissions on /dev/perfctr for PAPI	done
20. checking ram size (found 2 cores, 3.93 GB ram)	done
21. updating cron (removing updatedb)	done
22. checking hw clock (O seconds offset)	done
23. Checking IB Firmware	done
24. Setting IB-Hostname	done
25. Enabling SysRq	done
26. Checking for running Clusterworx vx processes -> not found	done
27. checking disk speed 54.3 MB/s	done
Looking good. Bringing up LSF.	done
p2s223:~ #	· · · · · · · · · · · · · · · · · · ·



Runtime Monitoring (Ganglia)



Guido Juckeland - Slide 4

High Performance Computing

SAN Monitoring (Live Monitor)



Center for Information Services & High Performance Computing

LSF Monitoring (LLview – W.Frings, FZ Jülich)



LSF Monitoring (LLview – right half of the window)



Temperature monitoring



tina

System state monitoring (Nagios)

🕲 Currei	nt Network Statu	ıs - Mozill	a Firefox						
<u>D</u> atei <u>B</u> e	arbeiten <u>A</u> nsicht	<u>C</u> hronik	<u>L</u> esezeichen	E <u>x</u> tras <u>H</u> ilfe				$\langle \rangle$	
A C A C A A A A A A A A A A A A A A									
🕒 Nagios	S	×	Current	Network Status				•	
Service Status Details For Host Group 'hrsk-pc-farm'									
Host 个	Service	$\uparrow \downarrow$	Status ↑	↓ Last Check ↑↓	Duration ↑↓	Attempt 🛧	Status Information		
deimos00) <u>1</u> /file system	<u>m</u>	ОК	18-04-2007 10:47:59	15d 15h 59m 42s	1/3	DISK OK - free space: / 4361 MB (30%):		
	/boot file sy	<u>ystem</u>	ок	18-04-2007 10:46:08	15d 16h 2m 36s	1/3	DISK OK - free space: /boot 65 MB (66%):		
	/opt file sys	stem	ок	18-04-2007 11:02:38	5d 22h 58m 16s	1/3	DISK OK - free space: /opt 42290 MB (50%):		
	/scratch file	<u>e system</u>	ОК	18-04-2007 11:02:59	5d 21h 59m 37s	1/3	DISK OK - free space: /scratch 20748 MB (72%):		
	<u>/tmp file sy</u>	<u>/stem</u>	OK	18-04-2007 11:02:59	15d 16h 5m 34s	1/3	DISK OK - free space: /tmp 6762 MB (70%):		
	/var file sys	<u>stem</u>	ОК	18-04-2007 11:04:47	0d 20h 2m 1s	1/3	DISK OK - free space: /var 3599 MB (75%):		
	PING		ок	18-04-2007 11:04:45	1d 3h 44m 2s	1/3	PING OK - Packet loss = 0%, RTA = 1.60 ms		
	<u>SSH</u>		ОК	18-04-2007 11:02:59	15d 16h 5m 54s	1/3	SSH OK - OpenSSH_4.2 (protocol 1.99)		
	Temperatu	<u>[</u>	OK	18-04-2007 11:05:19	0d 5h 21m 28s	1/3	OK - 22 Grad (max. Luftzufuhr)		
deimos00	2 💭 🗾 / file system	<u>n</u>	CRITICAL	18-04-2007 11:04:50	12d 16h 52m 29s	1/3	CHECK_NRPE: Socket timeout after 30 seconds.		
	PING		CRITICAL	18-04-2007 11:04:55	12d 16h 50m 36s	1/3	CRITICAL - Plugin timed out after 10 seconds		
	<u>SSH</u>		CRITICAL	18-04-2007 11:04:46	12d 16h 52m 29s	1/3	CRITICAL - Socket timeout after 10 seconds		
<u>deimos10</u>	11 💭 Check der	Nodes		18-04-2007 10:59:40	0d 1h 52m 7s	3/3	LSF says 14 nodes down - p1d010 p1d042 p1d046 p1d061 p1d088 p1q005 p1s012 p1s017 p2d001 p2d081 p2q013 p2q069 p2q088 p2s204		
	Lustre Fas	<u>tfs</u>	💥 ок	18-04-2007 10:47:11	0d 3h 19m 36s	1/3	OK		
	Lustre Wor	<u>'k</u>	💥 ок	18-04-2007 10:49:51	0d 0h 56m 56s	1/3	OK		
	NFS moun HPC-SAN	<u>ts des</u> /fastfs	💥 ок	18-04-2007 10:53:14	0d 3h 13m 33s	1/3	ок		
	NFS moun HPC-SAN	<u>ts des</u> /work	🐹 ок	18-04-2007 11:03:21	0d 0h 23m 26s	1/3	ок		
	PING		OK	18-04-2007 11:04:45	1d 3h 44m 2s	1/3	PING OK - Packet loss = 0%, RTA = 6.02 ms		
	<u>SSH</u>		OK	18-04-2007 11:02:37	7d 19h 12m 12s	1/3	SSH OK - OpenSSH_4.2 (protocol 1.99)		
	UNICORE	TSI	OK	18-04-2007 11:02:59	7d 20h 59m 55s	1/3	TCP OK - 0,001 second response time on port 5578		
	tmp		OK	18-04-2007 10:49:51	1d 4h 57m 4s	1/3	ОК		
deimos10	2 PING		OK	18-04-2007 11:04:52	1d 3h 44m 3s	1/3	PING OK - Packet loss = 0% RTA = 6 16 ms		
30110010	SSH		OK	18-04-2007 11:02:37	7d 19h 43m 8s	1/3	SSH OK - OpenSSH 4.2 (protocol 1.99)		
								-	

SAN monitoring(Nagios)

TECHNISCHE

DRESDEN







Things left unsaid

- Bootup checks
 - Do not (yet) cover all possible checks (e.g. MPI missing)
 - Should be repeated regularly to ensure QoS
- Ganglia
 - Pretty long load times for 700 nodes
- Nagios
 - Still using version XXX
 - Checks are not executed in parallel
 - Grown prototype ightarrow needs redesign
 - Currently open as Master Thesis
- Admin needs a lot of screens to see everything
- Open Question of scalability





System Management



System management (GUI): Clusterworx



Guido Juckeland - Slide 13

System Management (Console): SGI Tempo

Cascaded Management

- Sub-Managers per rack
- Allows cluster boot really in 10 minutes
- Information spread over the whole system

Yet another set of tools

• C-Tools (cpower, console, cexec) instead of P-Tools (powerman, conman, pdsh)

Only supports SLES

No version management for images





Infiniband Monitoring



ONE "THING" TO RULE THEM ALL.





Guido Juckeland - Slide 16