

RAID controllers under the microscope

DOING IT THE HARD WAY

Linux supports a wide range of SCSI RAID controllers.

Linux Magazine benchtested ten examples, together with two top-class servers.

Bernhard Kuhn delivers the verdict.

BERNHARD KUHN

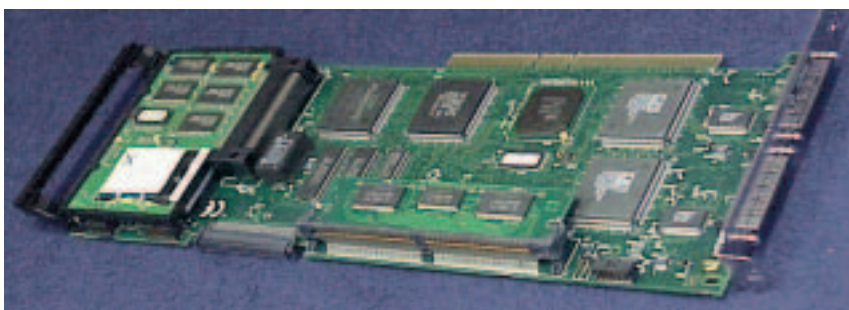
Fig. 1:
IBM ServeRAID

There are a great many RAID-compatible host adapters for Linux. For this test we concentrated solely on current SCSI-3 RAID controllers for PC servers. Besides pure performance, handling was of

particular interest as was what monitoring tools were supplied. However, it often turned out that what was true about one controller could be applied to others in the same family.

For testing of the ServeRAID controller from IBM and the competing products from Compaq, our hardware lab was also supplied with high-powered servers from these same manufacturers. Thanks to this, we took the opportunity to have a look at these as well.

Any moderately recent Linux distribution should recognise the RAID controllers on installation, provided that the RAID array has previously been prepared using the controller BIOS or support CD as required and the controller BIOS is a sufficiently recent version.



IBM Netfinity 7600

The Netfinity 7600 from IBM comes in a classy black housing that can accommodate up to 8 units, which means that at least two grown men are needed to move it about. The device can, however, be completely dismantled within minutes without the aid of any tools (see below). If necessary it can also be dismantled and transported as individual components. The machine we tested had a XEON 550 processor, 512Mb RAM and three 9.1 Gb hard disks, which leaves ample room for additional expansion. Up to four redundant power supplies provide plenty of current and can obviously be changed while the server is running. As with other manufacturers, and as a general rule, all components marked in bright red can be changed when the server is running. This list of red components includes all the fans in the system, whose accumulated noise level remains, perhaps surprisingly, within limits. However, IBM should have a think about the lack of guide rails for the swappable components.

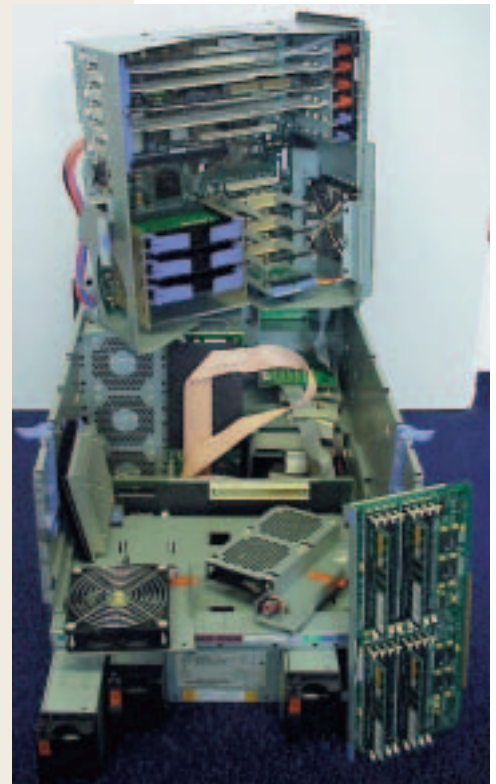
Four out of the six 64-bit PCI slots are hot-swap-capable. This means that slot-in cards can be exchanged while the computer is running. Unfortunately, Linux isn't yet able to exploit this useful characteristic – at least, we couldn't find any information about this at IBM itself or anywhere else in the Linux universe.

As is usual in a four-Xeon mainboard machine, voltage regulation modules for the processors are present. They bump up the cost and are perhaps ultimately unnecessary. There are a few manufacturers who include more than the absolute minimum – after all, they want to be able to give customers tailor-made processor upgrade kits, in return for the usual small change, of course.

In combination with the ServeRAID-controller, hot-swapping the SCA hard disks worked as expected, although they were marked light blue and so strictly should only be swapped in the "cold" condition. They were this colour because of the standard and otherwise normal Adaptec 7896 controller which is present onboard. Also on the motherboard is a S3 graphics chip, for which, however, we were unable to find a usable X server configuration.

You don't get the usual BIOS menu with the Netfinity. Instead it has to be initialised using one of the Windows bootable support CDs. During the ServeRAID set-up you're asked about the operating system you're going to install – of the selection on offer, we were forced to choose "other".

In all, the Netfinity 7600 left a good impression and proved fully compatible as a Linux server. With IBM's repeated announcements that it is going to align its products better with Linux it shouldn't be too long before we get PCI hot plugging features too.



The Netfinity 7600 can be completely dismantled within minutes.

Fig. 2: Successful hot-swap: the most colourful monitoring tool at present comes from IBM for its ServeRAID controller – but it demands Java.

IBM ServeRAID

Commendably, IBM has some web pages dedicated to Linux ServeRAID support and this makes it considerably easier to get started. We tested the RAID controller in its native IBM environment in the IBM server also reviewed here. The BIOS screen of the host adapter only allows, via the set RAID configuration option, one choice: loading the factory default settings. For the initial set-up the IBM ServeRAID Support CD must be used (see Netfinity 7600.)

With the monitoring tools for Linux, IBM leaves it up to the administrator to choose between command line operation (*IPSSEND*) and the ServeRAID Manager which boasts a GUI (see Figure 2). Using the latter you can remotely administer Linux Netfinity boxes that have the ServeRAID adapter installed. The necessary daemon is automatically installed at the same time as part of the RPM package (which runs to 11 MByte).

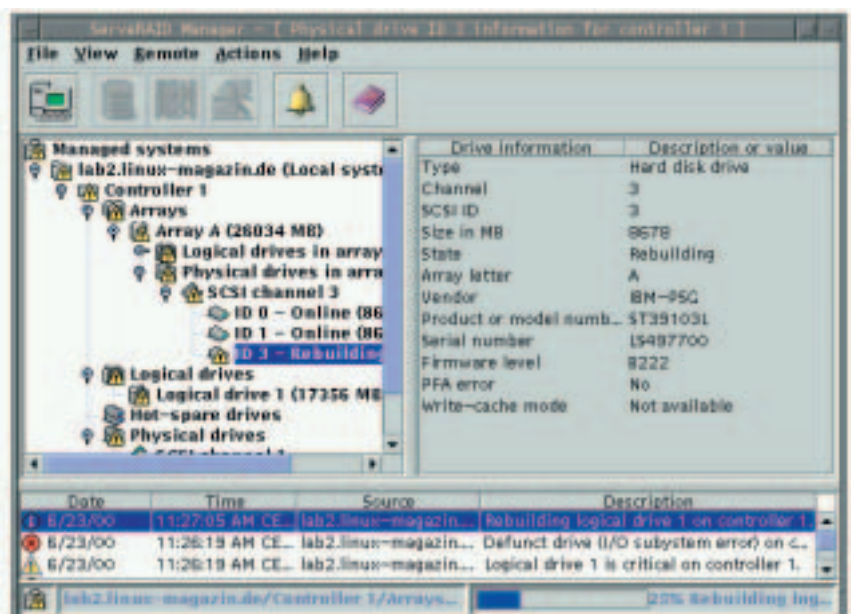
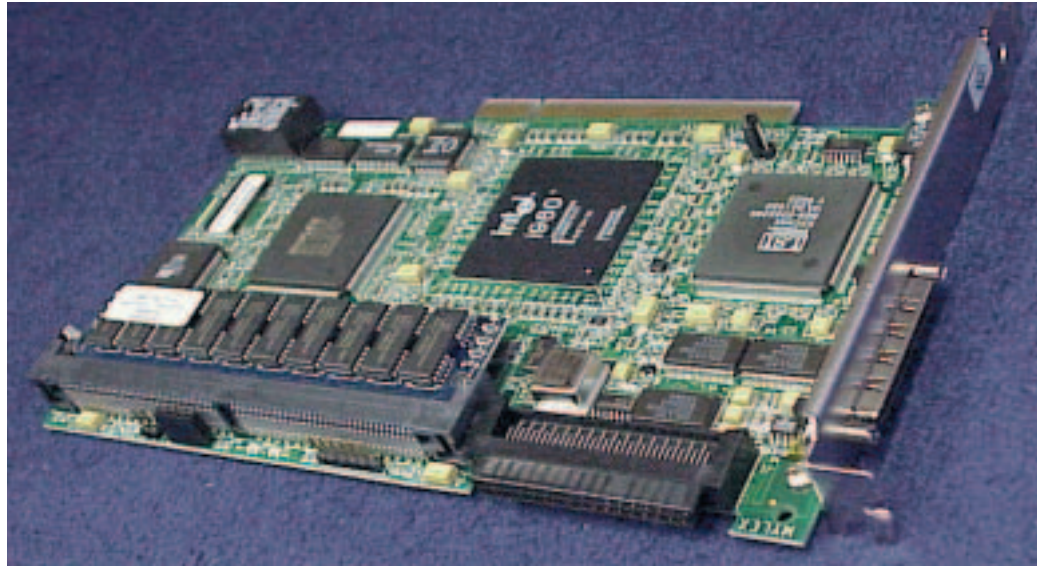
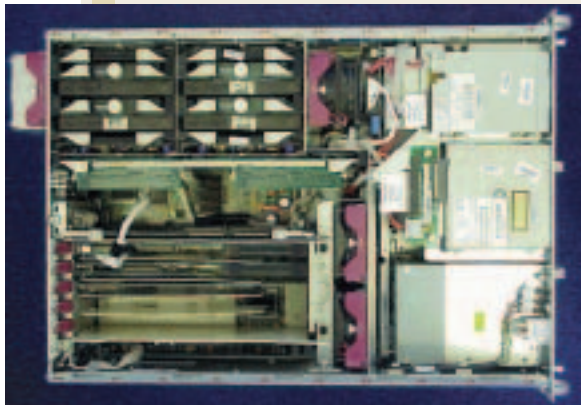


Fig. 3: Mylex
AcceleRAID 250



Compaq Proliant 6400R

The Proliant server comes from the Compaq 6400 series. It seems a case of Compaq in name and compact in nature as the inner workings of the computer are extremely densely packed but don't fundamentally restrict the functionality. The only limitation we found is that the drive bay can only be fitted with a maximum of four hard disks. The low height of the device may



You could hardly get more compact than this – but in the Proliant 6400R it's still not a tight fit.

Like the Netfinity, the Compaq server has no BIOS menu and is initialised using the tools on the supplied bootable Windows CD. Here, too, Linux does not appear in the list of supported operating systems. Compaq's Linux support Web pages advised us to enter "UnixWare", which did the trick.

The onboard SCSI controller from Symbios Logic serves the SCA hard disks as usual. However, our test configuration was also equipped with three Compaq RAID controllers which were able during the test to make use of the three 9.1 Gb media one after another.

The compact server makes a splendid platform for web services based on Linux. However, the deafening noise of the fans might make a fleet of servers too loud for many. ■

well force you to buy a secondary hard disk unit.

All mechanical components give the impression of being solidly built and are fitted with guide rails if they can be swapped during operation (the hot-swap colour is brown-red). The Proliant we tested had only two spaces for redundant power supplies, though most buyers should find this adequate.

All six PCI slots (64 bit) are hot plug capable. A little door in the cover of the housing grants easy access to the slot-in cards. But it's a shame that hot swapping can only be done on systems running Motorola's CompactPCI-Bus and the corresponding Linux kernel patches.

Mylex AcceleRAID 250

At present, Mylex is keeping Linux at arm's length. Linux drivers are available, but this is not down to Mylex but to Leonard Zubkoff, who created the drivers by means of reverse engineering when the hardware was released. As customer demand increased, Zubkoff was commissioned by Mylex to develop the official software. But on the manufacturer's web pages Linux is still not mentioned anywhere. Nevertheless, Linux is aware of almost all the AcceleRAID- and extremeRAID family and can therefore cope with them in the usual fashion.

AcceleRAID 250 has, like its big brother, an easy-to-understand BIOS menu with which you can perform the initial RAID set-up. The configuration which is necessary after that for higher RAID levels occurs in the background. The priority to be used for the synchronisation procedure can be selected as required – 50% is the default setting.

After installing and booting the operating system only the Proc-Filesystem is available for monitoring and configuration purposes which is very unsatisfactory. In an emergency few administrators would have the nerve to delve into the README in order to find the magic solution. But, thanks to SAF-TE (SCSI Accessed Fault-Tolerant Enclosure) this should rarely be necessary: all Mylex controllers obviously recognise defective disks and take them out of the group. After changing a disk during operation (assuming there is an appropriate swap cradle) the controller immediately begins the reconstruction.

Thus we have to conclude that Mylex controllers can be used under Linux within certain limits. In the near future an easy monitoring and configuration tool for Linux should be available with Mylex's own Global Array Manager.

In our tests, the AcceleRAID 250 refused to perform in the Netfinity test machine (the SCSI drives could not be found). With an old BIOS from 1998 on the extremeRAID 1100, Linux did not know where to start with either test system.

Compaq Smart Array 221/3200/4200

Initialisation of the RAID stack on the Smart Array was child's play, thanks to the "SmartStart" support CD supplied. Five minutes later the Linux installation could begin. However, monitoring and reconfiguration on the fly is a bit thin: unfortunately `/proc/array/ida0` gives no indication of the status of the RAID system. The only Linux monitoring tools – `status` and `gtkstatus` – only make one `ioctl()` query. They could, it must be said, be compiled very easily, but the hoped-for results (see Figure 7) did not appear. Also, after removing a hard disk during operation the report from `status` was still "Logical Drive OK". Nevertheless, the hot-swap took place with no problems. During reconstruction the middle

Hot swap with SCA

RAID 5 makes a disk stack safe in the event of a crash but without special swap cradles the server still has to be turned off to replace a defective medium. However, the high signal frequencies present on a SCSI-3 bus cause trouble with the cradle to hard disk connector. For this reason, larger server systems provided by big name manufacturers have tacitly started using fast SCSI hard disks with the Single Connection Attachment (SCA) interface. This is an 80-pole Low-Voltage Differential (LVD) connector that includes the voltage supply. The SCSI ID of the disk is also no longer defined using a jumper but is determined automatically from the socket on the SCA backplane in the drive bay.



Normally the SCA connector is found only in OEM systems and module recognition is done optically.

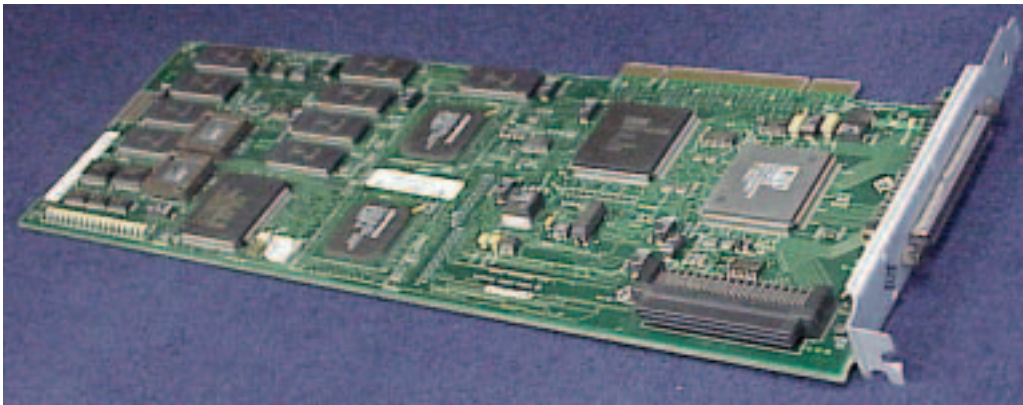


Fig. 4: Compaq Smart Array 221

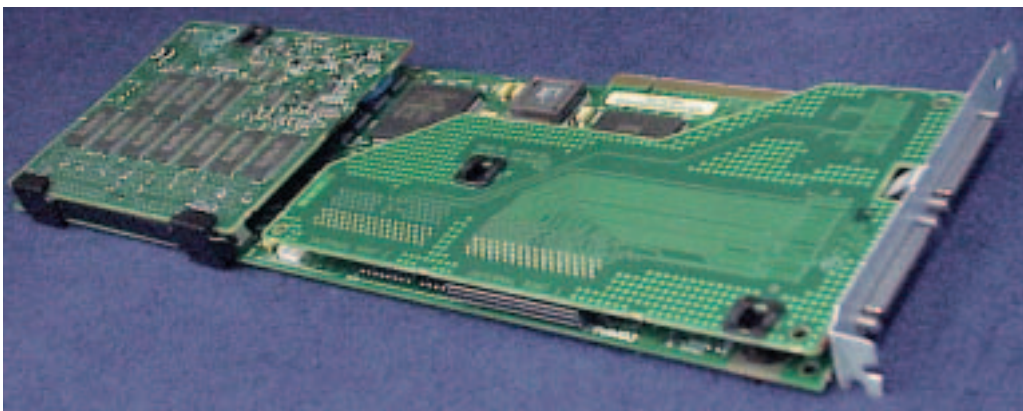
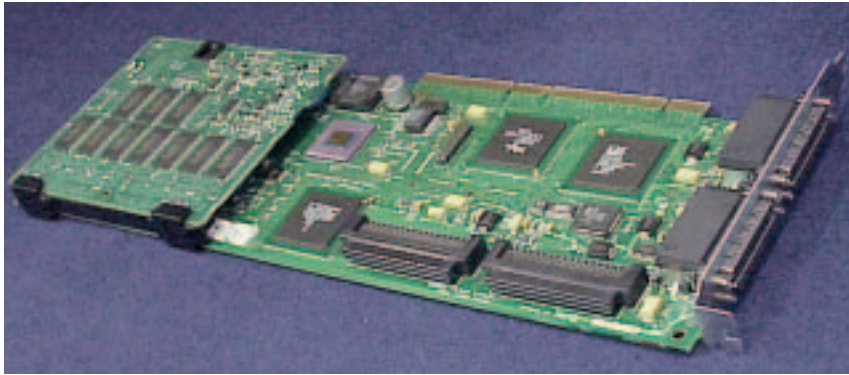


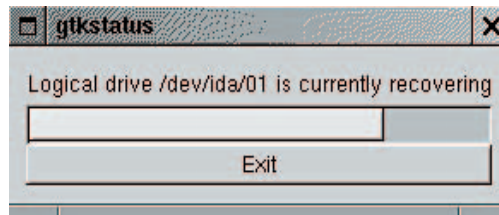
Fig. 5: Compaq Smart Array 3200

Table 1: Features											
Manufacturer	Product	Pre-OS configuration	Monitoring/Configuration-Tool	User Interface for M/C-Tools	RAID Levels	Block-Device	Hardware Support	PCI Bit	LVD/SE intern	LVD/SE extern	Cache [MB]
DPT/Adaptec	SmartRAID VI Decade	BIOS/Floppy-Disk	Storage Manager Utility	Lesstif-GUI and command line	0,1,5,10,50	SCSI	+	32	2	2	128
DPT/Adaptec	SmartRAID V Century	BIOS/Floppy-Disk	Storage Manager Utility	Lesstif-GUI and command line	0,1,5,10,50	SCSI	+	32	1	1	128
AMI	MegaRAID 1400	BIOS	MegaRAID Manager	Text-UI (Slang)	0,1,3,5,10,30,50	SCSI	+	32	3	3	16-64
IBM	ServeRAID	Bootable Support-CD	ServeRAID Manager	Java-GUI or command line	0,1,5	SCSI	+	64	1	2	4
Mylex	AcceleRAID 250	BIOS	/proc (Global Array Manager i.V.)	Proc-filesystem	0,1,3,5,10,30,50	/dev/rdl	0	32	1	1	64
IPC Vortex	GDT 7538RN	BIOS	lpcmon und ICPD	Text-UI (Ncurses)	0,1,5,10	SCSI	++	32	3	2	8-128
IPC Vortex	GDT 6538RS	BIOS	lpcmon und ICPD	Text-UI (Ncurses)	0,1,5,10	SCSI	++	64	4	3	8-128
Compaq	Smart Array 221	Bootable Support-CD	- (Gtkstatus)	ioctl()-System invocation	0,1,4,5,10	/dev/ida/	-	32	1	1	6
Compaq	Smart Array 3200	Bootable Support-CD	- (Gtkstatus)	ioctl()-System invocation	0,1,4,5	/dev/ida/	-	32	1	2	16-64
Compaq	Smart Array 4200	Bootable Support-CD	- (Gtkstatus)	ioctl()-System invocation	0,1,4,5,10	/dev/ida/	-	64	2	4	16-64



[above]
Fig. 6: Compaq Smart Array 4200:

[right]
Fig. 7: At present there are no more monitoring options from Compaq



LED flashes and so the end of the approximately 30 minute synchronisation process (in the 4200 using RAID 5 with 3 times 9.1 Gb drives) could at least be seen.

AMI MegaRAID 1400

AMI's web pages provide the Linux driver and monitoring software for its MegaRAID family. They're

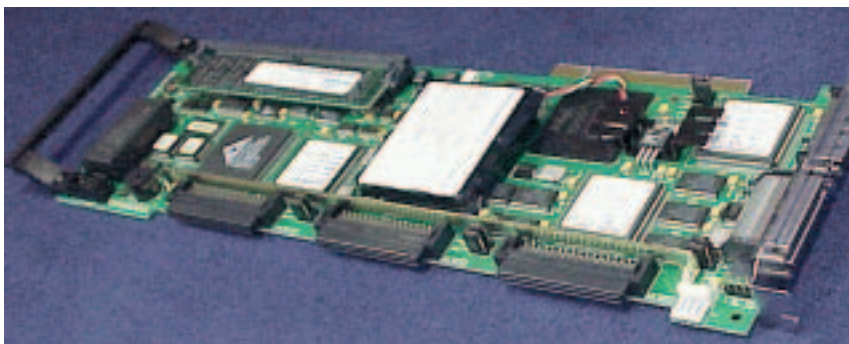


Fig. 8: AMI MegaRAID 1400

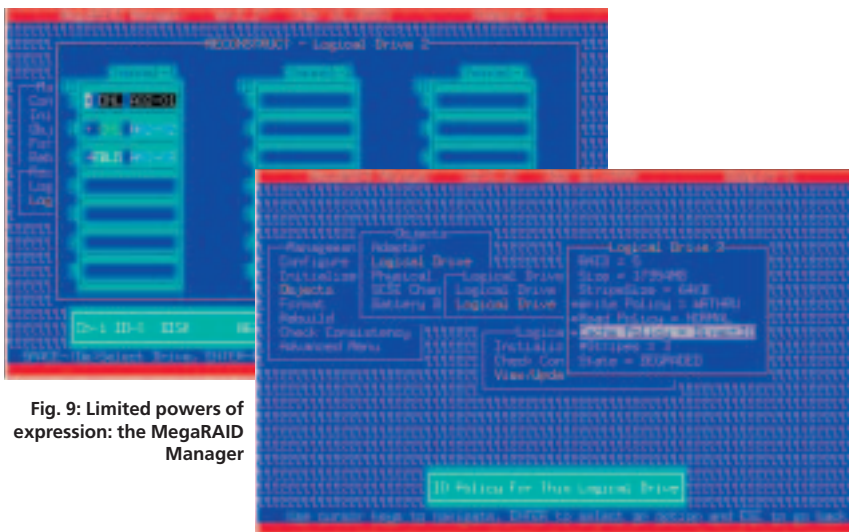


Fig. 9: Limited powers of expression: the MegaRAID Manager

shown as having equal status with SCO, Solaris and Windows NT – hooray!.

The initial set-up of the RAID system occurs in much the same way as usual with an overview menu in the controller BIOS. Synchronisation for a RAID 5-configuration consisting of three 9.1 GByte-hard disks took about five minutes in the IBM test system. During the installation of Red Hat 6.2 the controller hesitated a little ("SCSI-input/output error"), but after repeatedly ignoring the error in the popup window, the controller finally gave in. Maybe the synchronisation was not quite perfect, but no more anomalies appeared later. As is the case with the Compaq Smart Array, the MegaRAID-controller also requires a separate boot partition (because of the troublesome 1024 cylinder LILO limits).

After the Linux installation the hard disks can be monitored and/or reconfigured using "MegaRAID Manager", a text-based user interface. Hot swapping, tested with the Compaq machine, was a success. However, the monitoring tool merely reported the status of the system (*degraded*, see Figure 9). It didn't say how much longer the rebuild would take. The end can only be determined by observing the activity of the hard disk.

DPT SmartRAID Decade and Century

Distributed Processing Technologies, who make the SmartRAID Decade and Century, treat Linux like any other operating system. And that's been the case for a while. So together with the obligatory device drivers there are also Lesstif-based configuration and monitoring programs as well as boot diskettes for Red Hat Linux ready for downloading. Anyone who won't tolerate a graphical user interface on his server can opt for the opulently equipped command line instruction *dptutil*.

The BIOS is used for setting up the RAID stack. However, it is only possible to install one maximum size logical device per RAID group. The initialisation of a level 5 configuration is run from the BIOS menu and takes more than two hours for the Decade using three 9.1 Gb disks. Although this is performed in background, the BIOS made three engineers believe that leaving the Menu will abort the synchronisation.

Let's hope that after the take-over of DPT by Adaptec the promising development of Linux drivers and tools will continue.

IPC Vortex 6538 and 7538

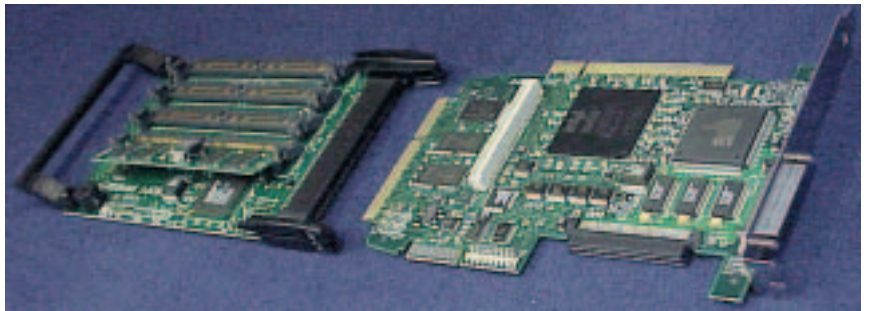
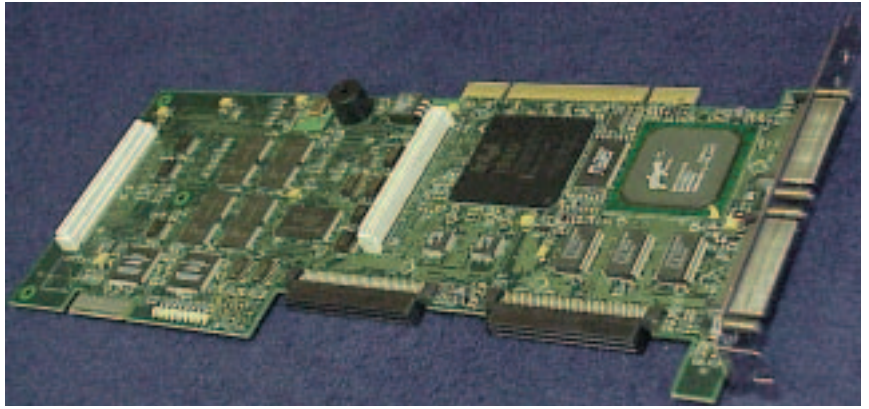
This company started development of the Linux driver on day one. This is why Vortex controllers have far and away the best support for the free operating system. Apart from the comprehensive monitoring/configuration tool *ipcmn* (formerly *gdtmon*), you also get *ipcd*. RAID server systems can be remotely monitored via a network (using TCP/IP) with this daemon. This means that monitoring for an entire fleet of servers

can be performed centrally with just one *ipcmom* process running on the administrator's workstation. As with some products from other manufacturers, the controllers from Intelligent Computer Peripherals also have a comprehensive BIOS for the initial configuration of the RAID stack.

The flag ship GDT7538RN was reluctant to do its duty at first in either of the two test systems. A tip from the support hotline – they told us to open Jumper S4 – at least allowed us to make it work in the Compaq Proliant. Its little brother GDT6538RS was also unable to make friends with the special BIOS variants from IBM and Compaq. After some tweaking, the initial set-up was extraordinarily fast and easy, but then Linux refused to boot up.

Benchmarks

During the performance tests of the RAID controllers, *bonnie* was used for booting (Option `-s 1000`). The virtual file system layer with its caching algorithm has a considerable effect on the measured values. This can largely be prevented using the boot option `mem=32M`. The RAID 5 systems, consisting of three fast SCA-SCSI hard disks with a maximum throughput of 80 Mb/sec, were prepared with the ext2 file system for the 1 Gb test file (approx). Since not all the controllers in both computers wanted to run, the measurement results in Table 2 should be taken with a pinch of salt. The GDT6358, for example, did not want to do its duty in either of the two high-end servers so we had to switch to a third system from a lower performance class ("IBUS").



The measured values for Adaptec and Symbios Logic relate respectively to the onboard controllers of the two test machines. The test with single hard disk drives ("Single HDD") shows that the hard disk drives of both systems are approximately equal in speed. This means that the test results of the controllers can be compared with each other, despite the fact they were calculated with different computers. Interestingly, SoftRAID 5 comes out astonishing

[top]
Fig. 10: DPT SmartRAID VI Decade

[above]
Fig. 11: DPT SmartRAID V Century

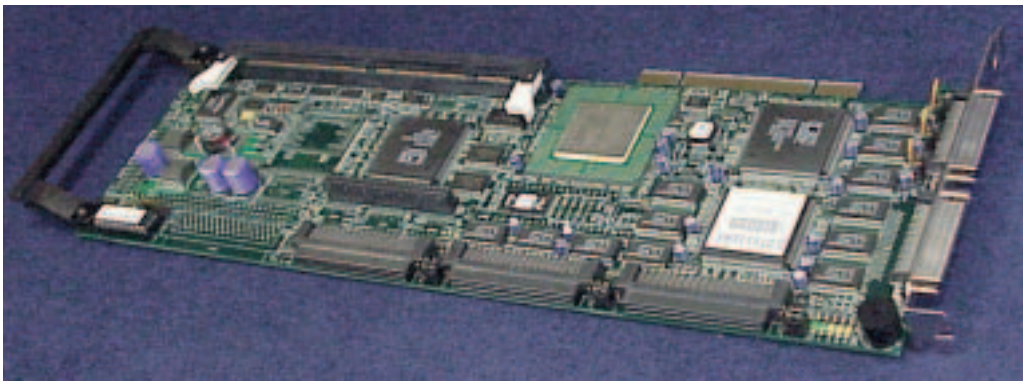


Fig. 12:
IPC Vortex GDT6538RS

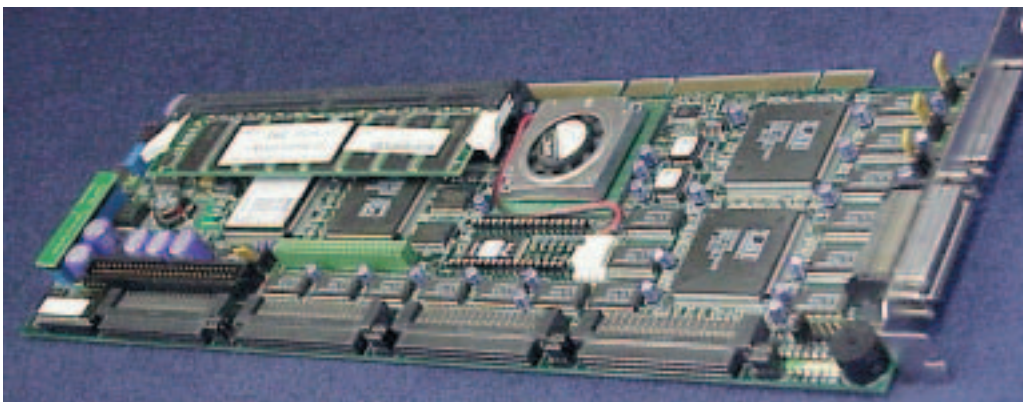


Fig. 13:
IPC Vortex GDT7538RS

Table 2: Test results											
Manufacturer	Controller	Mode	hdparm	Char-out	Block-out	Rewrite	Char-in	Block-In	Seeks	normal geom.	
			MB/sec	MB/sec	MB/sec	MB/sec	MB/sec	MB/sec	/sec	average/1000	
DPT	SmartRAID IV Decade	RAID5	11.59	3378	3431	2504	18892	26804	231.0		5.8
DPT	SmartRAID V Century	RAID5	11.59	10289	10202	5349	19924	27913	243.6		9.4
AMI	MegaRAID 1400	RAID5	11.99	3713	3692	3604	18708	25399	246.0		6.2
IBM	ServeRAID	RAID5	14.61	15258	15501	8441	19981	24479	206.8		10.7
ServeRAID	AcceleRAID 250	RAID5		8614	8623	5045	26777	36786	249.9		10.0
IPC Vortex	GDT 7538RN	RAID5	37.21	19656	20868	9459	25765	36244	265.4		14.3
IPC Vortex	GDT 7538RN	RAID5	26.67	22608	30414	8939	33310	35918	272.8		16.2
IPC Vortex	GDT 6538RS	RAID5	34.41	18344	18423	8845	33143	35877	270.3		14.1
Compaq	Smart Array 221	RAID5		6901	6881	4730	23752	31606	230.0		8.6
Compaq	Smart Array 3200	RAID5		6871	7001	6076	23050	36011	246.8		9.3
Compaq	Smart Array 4200	RAID5		7032	6923	5714	23278	36499	252.0		9.3
IBM	Adaptec 7896	Single-HDD	21.77	22916	24228	10967	24612	26104	166.4		13.0
IBM	Adaptec 7896	SoftRAID5	30.19	16021	18734	11512	27511	33343	250.3		13.6
Compaq	Symbios 53c876	Single-HDD	22.38	20243	28207	11825	27004	28822	175.4		14.0
Compaq	Symbios 53c876	SoftRAID5	31.68	15753	15894	4249	6048	6509	263.3		5.7
		Weight		1	2	1	3	1	1		

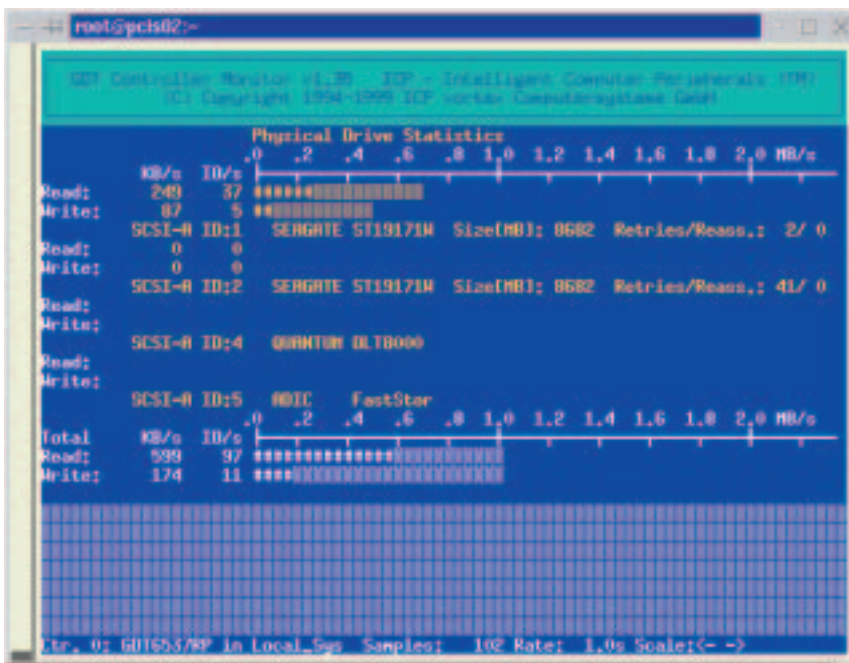


Fig. 14: the ipcmom for the GDT controller series is teeming - especially with statistical functions

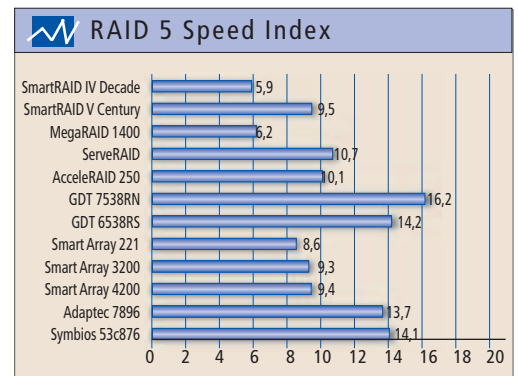


Fig. 15: Speed Index: the controller in overview

ly well on the Netfinity but constantly loads the processor to full capacity. This could be counteracted by having more CPUs, although it would be cheaper to acquire a hardware RAID controller than a higher-performance XEON processor.

Missing values in the "hdparm" column, by the way, indicate that the corresponding controllers did not register as adapters for storage media in the system. The benchmark findings were created with the co-operation of Dipl.-Ing. Axel Dittrich and Dr.-Ing. Hans Pfefferl, systems administrators at AGFA-Gevaert Munich, Germany.

Conclusion

Apart from one Mylex controller with its old BIOS, all the test products ran on the whole satisfactorily. Only the extremely poor performance of the DPT Decade (which, by the way, had an Ultra 160 interface) left a lot to be desired.

The fact that quite a few of the controllers had problems with the test systems may have been due to incompatibilities, in particular with the BIOS of the servers. The GDT controllers in particular were better able to make friends with normal PC servers and demonstrated their full power. With the ServeRAID from IBM and the Compaq Smart Arrays one must rely on the bootable support CD to perform configuration.

Info

- Linux IBM ServeRAID device driver**
<http://www.developer.ibm.com/welcome/netfinity/serveraid.html>
- Dandelion Digital's Linux DAC960 Page**
<http://www.dandelion.com/Linux/DAC960.html>
- Mylex Disk-Array-controller under Linux**
http://sdb.suse.de/sdb/de/html/mylex_dac960.html
- Linux Driver for Compaq Smart-2**
<http://www.insync.net/~frantz/cpqarray.html>
- Compaq Linux pages** <http://www.compaq.com/linux>
- MegaRAID Drivers and Utilities**
http://www.ami.com/megaraid/mr_drivers.html
- Linux DPT Hardware RAID HOWTO:**
http://www.ram.org/computing/linux/dpt_raid.html
- Driver and Monitoring/Configuration-Tools:**
<http://www.dpt.com/techsup/sr5drv.html#LINUX>
- IPC-Vortex Linux drivers and Monitoring/Configuration Tools**
http://www.vortex.de/supp/down/linux/linux_d.html