



Suse Linux Enterprise Server 9

Green Light

Dirk Wetter

Starttime 04082004: After the Redhattians of System Americus had been attacking the interplanetary confederation of United Linux with preliminary version 2.6 starships for 0.75 sidereal years, there has now been a counterattack with an improved 2.6 fleet under a new constellation: The old confederation has been broken since those doubtful Santa Cruzian actions before the intergalactic court of justice. With the Novellians as their new replicator wage giver, the Germanic core crew now hopes to besiege the Redhattians, particularly outside the Europa system.

Suse's Linux Enterprise Server 9 (SLES 9) was introduced to the public at LinuxWorld Expo in San Francisco at the beginning of August, giving a two-fold impression: The last product from Nuremberg – Suse 9.1 – had starting problems of some technical sort [1]. In addition, it was offered exclusively with kernel 2.6, which was quite a sig step up from the previous version.

Now, shelf distributions are a bit on the side for their creators. The core business are the products for the enterprise and services around it. Errors there would be much more critical both for the reputation and the finances. Neither Red Hat nor Suse develop their enterprise distributions from scratch: United Linux 1, to some better known as SLES 8 [2], was based on Suse 8.1, and Suse 9.1 acted as

godfather for the current SLES 9. According to Chris Schläger, Suse's Research and Development Manager, the list of improvements for the enterprise version alone was as long as your arm. The features sound rather promising: first-time use of kernel 2.6 in the enterprise field, support of enterprise file systems, which existed already in the predecessor, as well as Linux Kernel Crash Dumps (LKCD), Samba 3, and a few new and revised YaST modules that should make work easier. In short: It will be interesting to see what SLES 9 offers and how things will be about its stability and internal life.

64-Bit on the run

An enterprise distribution should be tested under adequate conditions. A Celestica Quad Opteron system (AMD-64-848 CPUs with 2.2 GHz each) was available as test computer number one. Among other things, the system had 8 GByte memory and four U320 SCSI disks on a Fusion MPT controller [3]. A brand-new HP Proliant DL380 G4 Dual-Xeon (3.4 GHz HT-EM64T CPUs) with 1 GB of RAM and an integrated Compaq Smart Array RAID controller served as second test system. Two SCSI disks sat at this RAID controller waiting for SLES. More details about the HP computer will be included in a test in the next issue. Both machines had dual-gigabit Ethernet cards and received their network setup via DHCP over a gigabit link connected by a switch.

The exciting thing about Celestica is that its memory is not "flat"; it is a NUMA (Non-Uniform Memory Architecture). While Intel specifies up to 4 CPUs per cell for its IA64 processors, each of the Opterons does no longer consider CPU-own memory to be local. From the Linux kernel perspective this leads to three problems: The kernel has to know the hardware memory location. Memory should ideally be allocated only locally. The scheduler should try not to migrate a process that has memory allocated to it to another CPU, since a potential memory access of that process would then again not be local and take more time.

The x86-64 architecture of the two test computers is compatible with the older i386 architecture, which would allow you to install the SLES i386 version, if you okay a funny warning message (Figure 1). The test did without this, because the "small" 64-bit archi-

tures have become an affordable alternative for computer centers in view of the 4 Gbyte limitation of 32-bit operating systems.

All SLES versions (i386, x86-64, IA64, IBM I/P/Z series) come on six CDs. The first (“Suse SLES”) offers a handful of RPMs and booting documentation. This “auxiliary CD” can be arbitrarily exchanged, similarly to the first United Linux CD. Of the remaining five “Suse Core” CDs the last two include source packages. In addition, ISO images with an SDK and Eclipse for i386 can be downloaded for all platforms from developer.novell.com/ndk/susesdk.htm. Apart from several “Devel RPMs” the SDK includes software development programs, such as KDevelop, Subversion, as many as 130 Perl modules, XEmacs – Gnu Emacs is on CD number two – C# compiler, including tools from the DotGNU project, and additional goodies such as Gimp 2, MrProject, Quanta, and TeX. Ximian's or Novell's Mono, respectively, is included in the SDK for i386, ppc32, and s390 only.

Installed in 15 minutes

The first candidate in the test was Proliant. It was expected to undergo a fresh manual installation from the CD. The default consists of a swap plus Reiserfs root partition, and a software standard selection. The YaST installer can be used to change the default values in menus. The YaST installer's look and feel hasn't changed since Suse 9.1. However, the top menu in the software selection offers the “Full Installation” option and, a menu level lower, there is a package selection, clearly adapted to the potential purpose of use, which gives the whole thing a

X-TRACT

- SLES 9 is the first Linux enterprise server with kernel 2.6 and replaces the almost two-year old version 8.
- Thanks to Novell's better market presence in the US, there's hope that the SLES 9 may be marketed better than formerly under the Suse directorship.
- A number of future enterprise products by Novell/Suse and Sun will be based on the new server Linux.

tidier appearance. 15 minutes later, 576 packages – almost all of them from the first Core CD – had been installed in RAID-1. The IDE CD-ROM hadn't been running in DMA “UnmaskIrq” or 32-bit-I/O-mode during that time; `hdparm -d1 -u1 -c1 /dev/hda` in a shell speeded things up a little, and later on the tester saved this in `/etc/sysconfig/ide` to ensure faster operation in the future

The network configuration followed after rebooting and entering the super-user password, as usual. By default SLES still uses the relatively weak but compatible DES encryption. Since `nullok` is missing in `/etc/security/pam_unix2.conf`, at least it doesn't allow empty passwords any more. The first of the two internal gigabit cards was happily preconfigured with DHCP, so that no changes were required. As usual, there was an optional network connection test by downloading the release notes and installing online updates. The next two steps are new to SLES 9 and remind you a little of the standard server from the same company: Configuring the OpenLDAP server and deploying a CA (Certificate Authority). YaST uses an imported or – if required – even a self-created CA (“YaST CA”) to create a server certificate, but it configures it only for the OpenLDAP server. For additional services requiring certificates, such as Courier-IMAP, Apache, or VPN, you have to do some extra work manually or let YaST do that job. This may tempt some system administrators to cast a glance at the standard server [4], which does part of this job automatically. The defaults of the YaST installer for the CA and LDAP servers form a good starting point. The following fine-tuning included the assignment of other passwords for the LDAP server and the CA, because the previously entered root password would apply otherwise.

Automatic login is now activated by default for the obligatorily created local user, in contrast to Suse 9.x. When changing the `uid` and adding a second group the group/user dialog got a little mixed up.

Printer automatically integrated

The subsequent hardware detection of the graphics card and the monitor, with 1600 x 1200 pixels and 75 Hz, re-

sulted for the first time with any distribution in a reasonable and functioning resolution for the Sony monitor. The CUPS server present in the network announces its available printers via UDP broadcast, which YaST detected correctly and added to the CUPS client configuration.

Further installation attempts followed. The Proliant received an XFS root file system, which worked nicely despite the use of a CVS snapshot of XFS. A hiccup occurred with an encrypted `/home` file system when the Proliant booted: The init script hang when the pass phrase was entered and the `ext3fs` mount was confirmed, but this didn't occur when we used Reiserfs. As a “workaround” it was recommended to simply wait for the timeout to expire and then to mount the file system manually by `/etc/init.d/boot.cryptostart` after booting.

Enhanced administration tool

One of the new helpers in YaST is the installation server module, which handles all work, including the configuration and startup of required services (NFS, HTTP, FTP). As a prerequisite, the module requires CDs in the drive, either as files or ISO images. Additional software, such as SDK or Eclipse, can be deposited on the install server, just as well as installation CDs for other architectures. The non-CD start didn't ask much; all we had to do is prepare the non-SLES TFTP and DHCP servers, cross-references on the parameters required for a VNC, serial, or SSH installation in the manual would have been helpful. A step further is undertaken by the `Autoyast` module; the administrator can use it to create an XML file with all options for the system to be installed. The new `YOUserver` module (YaST Online Update) installs a patch proxy for the Suse portal. If necessary, it starts the service (FTP, HTTP) required by the client and announces the YOU server via SLP, similarly to the install server's announcement.

The next step involved the AMD Celestica. A little problem occurred here: Due to the additional installation of the SDK – which integrated seamlessly into the YaST software selection dialog – the product name appeared to have changed to SLES-SDK, which caused YOU to search the Suse

portal in the wrong directory for its online updates. After a brief hunt for bugs, we were able to solve this problem by removing the sinister `/var/adm/YaST/ProdDB/prod_0000003` file. Meanwhile, there are appropriate directories in the portal.

Painting and innards

Apart from technical fusions, such as the NLD (Novell Linux Desktop), which will be coming up at the end of this year, according to Novell, Novell's influence on the corporate identity has already shown in the SLES, – though timidly: As shown in the lead picture, this reflects in discreet indications, the Suse logo (“Suse” – a Novell Company) and in the KDE desktop.

Now for a look at the inner values: In contrast to what says the manual and the “release notes” Apache is available not only in Version 2, which is better performing for kernel 2.6; the disks also include its Version 1.3. Moreover, Jserv, Tomcat, and JBoss are included, together with Sun's JVM 1.4.2. The native JDK 5.0 for x86-64 (formerly 1.5.0) was introduced as Candidate 1 release at copy deadline. For i386 and IA64 systems, the user can additionally choose IBM's Java (1.4.2) and BEAWebLogic. SLES doesn't take care well of the desktop user and, according to Suse, doesn't plan to do so: Open Office and KOffice are

missing entirely, and Mozilla 1.6 – patched against a large number of bugs by means of an online update – is not included in the software default, just as it isn't in Suse 9.1. The mailer, calendar, and address book parts are missing even in the full installation, as is the Red Carpet daemon, including the *rug* front-end, which offered one channel only: the program itself.

With regard to the software versions (see “Technical Data and Prices” below), SLES follows largely Suse 9.1 [1]: XFree86 is present in Version 4.3.99; unfortunately, *gcc* didn't get an update to the faster 3.4 version; no trace about Gnome 2.6.x. According to SLES Project Manager Ralf Flaxa, however, the SLES-*gcc* does include all important performance improvements from 3.4.

Security officer, be on guard

Not much has changed since Suse 9.x from the security standpoint: The Suse Firewall2 is intelligent and logs a lot, compared to its competitors. On the other hand, the system still replies to ICMP packets at broadcast addresses. Similarly, the port mapper listens to port 111, which is LSB-compliant according to Suse, although it represents an unnecessary potential risk in non-NFS environments.

In contrast, `/usr/local/bin` or `/usr/local/sbin` at the beginning of the system path can turn into a trap for root and average Unix users in NFS environments, which are typical for some computer centers. An SSH trojan, for example, that captures all keystrokes could be placed on the NFS server and compromise all user accounts on the NFS client. It is hard to understand what the superuser shell has to do there at all: Root would hardly be able to output commands, let alone logging itself on when the NFS server hangs. Suse should better not anchor the “local” paths, – they are empty anyway – , in the shell environment.

For IPv4 connection tunneling across external networks, SLES offers both FreeS/WAN and the KAME tools, based on the IPsec stack that comes with kernel 2.6. And there is a new YaST helper to configure the KAME tools. As regards SE Linux, there is a kernel-2.6 type of support, in addition to *libselinux* and its utilities. The policy compiler *checkpolicy* and the (*policy*-)*coreutils* are missing.

At first startup, the new YaST module for HA configuration appears to be a bit misleading: It offers the “Heartbeat on” or “Heartbeat off” option. The configuration hides a bit modestly behind “Continue”. A bug fix is in the works. Using a null modem cable and a few mouse clicks we managed to have the two test computers mutually sense their heartbeats via UDP unicast serially and over their second gigabit interfaces. An interruption was detected by SLES.

The heart of it all

Though Suse designated the version to be 2.6.5, the SLES kernel borrows some features from later “mainstream kernels” and patches that are not yet found there. One actually shouldn't take Suse's version numbers as a measure of things in other (important) parts of the SLES either. Suse doesn't emphasize its improvements clearly, at its own and the technically interested people's disadvantage. What really hides behind the number can be drawn only from the sources included in the product.

SLES 9 includes the NUMA API patch of Andi Kleen, as has been the case since SLES 8 with SP3, or kernel 2.6.7 rcl. This patch enables user-level tools, for example, to bind processes to CPUs to let them allocate memory by a specific cell only. Initial tests were promising: long running processes remained on their CPUs, and the statistics in `/sys/devices/system/node/node?/numastat` were pretty impressive. The NUMA scheduler is currently at the 2.6.7 level, according to Suse.

Here's another difficulty on SMP architectures, particularly in I/O-intensive use: The first CPU (CPU0) serves almost all interrupts, which might put the S in SMP in question when it comes to busy machines. The Nuremberg server installed and started *irqbalance* automatically, which meant that timers, or interrupts, originating, for example, from network, SCSI, or IDE devices, were distributed relatively evenly.

“AA” strikes back

Together with kernel 2.6, the virtual memory concept of Suse developer Andrea Arcangeli (AA-VM) was replaced by Riek van Riel's reverse mapping VM (RMAP-VM), with all of its

TECHNICAL DATA AND PRICES

Suse Linux Enterprise Server 9

Software: Kernel 2.6.5 with extensions, Glibc 2.3.3, GCC 3.3.3, XFree86 4.3.99, KDE 3.2.1, Gnome 2.4.1, Mozilla 1.6, Samba 3.0.4, Apache 2.0.49/1.3.29, 4 Binary and 2 Source-CD-ROMs.

Service

Installation support via e-mail and Web; security fixes, patches, updates, and upgrades for the purchased runtime via YaST online update.

For further options and CPU architectures see www.suse.de/de/business/products/server/sles/pricing.html

Free 30-day evaluation from www.novell.com/de-de/products/linuxenterprise-server/eval.html

Prices

x86, x86-64: \$418 per year, up to 2 CPUs,
\$958 per year, up to 16 CPUs

SUPERMICRO®

Machen Sie Veränderungen in der
Geschäftswelt nicht einfach nur mit.Profitieren Sie
von ihnen!

Auch in schwarz lieferbar!

**Supermicro 6113M-i/6113L-8**

- 19" / 1HE Rackmount
- Dual Intel® Titanium™ 2 Prozessor bis 1.5 GHz mit 6MB L3 Cache
- Intel® EB870 Chipset mit 400 MHz FSB
- ATA/100 oder SCSI U320 on board
- Intel® 82546EB Dual Port Gigabit Ethernet on board
- ATI Rage XL BMB PCI Grafik on board
- max. 4x 250GB IDE oder 4x 147GB U320 Festplatten
- PCI-X 64-Bit 133/100 MHz Erweiterungsslot
- 1x Slim CD-ROM
- 500W Cold Swap Netzteil

Auch in schwarz lieferbar!

**MidNAS 3415/9033 NAS-System**

- bis zu 15x 250GB S-ATA Festplatten (3,7TB)
- RAID 0, 1, 10 oder 5 möglich
- 760W Triple redundantes Netzteil
- Linux basierendes NAS-Betriebssystem
- Unterstützte Netzwerk-Protokolle: TCP/IP, NetBEUI, SMB, CIFS, NFS, FTP
- Unterstützte Clients: Microsoft Windows, Linux, Unix, Apple OS X
- Windows Active Directory/NT Primary Domain Controller Support
- Snapshot-Funktion
- Unterstützung von USV (unterbrechungsfreie Stromversorgung)
- keine Einzelizenzen für Clients oder Server notwendig

Weitere Informationen erhalten Sie auf unserer homepage unter www.cpigmbh.de oder kontaktieren Sie unseren Vertrieb per e-mail unter sales@cpigmbh.de oder unter der kostenfreien Telefonnummer 0800-100 82 69.

Alle Marken- und Produktbezeichnungen sind Warenzeichen oder eingetragene Warenzeichen des entsprechenden Unternehmens. Druckfehler, Irrtümer und Änderungen vorbehalten.

**CPI**

Computer Partner Handels GmbH

storage unlimited

Kapellenstraße 11 • D-85622 Feldkirchen b. München
 Fon: (+49)-0 89/96 24 41-0 • Fax: (+49)-0 89/96 24 41-33
 Büro Marburg • Fon: (+49)-0 64 21/94 814-0

sales@cpigmbh.de • www.cpigmbh.de

benefits and drawbacks. The so-called RMAP chains offer a chain of reverse pointers, enabling the kernel to quickly find out which virtual memory page belongs to what physical address. This is critical when it comes swapping. Otherwise, the kernel has to do the rounds of all PTEs ("Page Table Entries") to check their allocations, which can be pretty time-consuming, depending on the number of pages in memory. However, the original RMAP-VM suffered from the fact that its linked list needed a lot of space in the valuable "low-memory" area below 896 Mbytes (32 bits), which meant additional performance losses. For this reason, Andrea Arcangeli had fallen back on the old "objrmap" concept [5], and improved it by adding various patches ("anon_yama", "prio_tree"). From 2.6.6 onwards, you will find parts in the official kernel again; the VM has been identical with the current SLES since version 2.6.8.

A relatively young project was introduced into the SLES kernel, though it has been of academic value only, since only the "kernel hooks" are present, while userland tools are missing: Class-based Kernel Resources Management (CKRM). CKRM can best be described as a QoS on system level, or for the kernel, respectively. The CKRM project (ckrm.sf.net) seeks to provide kernel resource schedulers, such as task schedulers, I/O schedulers, VM page manager, and inbound network connection managers, which can be defined to schedule and prioritize address ranges, users, or a group of tasks.

Bottom Line

Suse didn't miss out on anything. SLES' benefits include the comfortable YaST installer, which doesn't leave system administrators doubtful as to where exactly they are within the installation process. Subsequently, they can use YaST, which was expanded again by additional functions, to complete many tasks. Despite the kernel being named 2.6.5, it includes a number of important features for enterprises, and it can score points with LKCD, InfiniBand, as well as good NUMA and SMP support. SLES 9 is even up to 50 percent cheaper than its predecessor, depending on the product.

But there are also a few drawbacks: *ALT_SYSRQ* – a useful feature – is not activated. The electronic installation

IX-EVALUATION

Suse Linux Enterprise Server 9

- ⊕ Kernel 2.6.5 plus NUMA API and Scheduler, InfiniBand support, LKCD, CKRM
- ⊕ YaST modules facilitate work
- ⊖ Security shortcomings, weaknesses in manual

and administration manual included in the product package – there is no print version – should deal with some of the topics more thoroughly, though it comprises 865 pages. In the breadth, some topics are missing, e.g., material on LKCD, UML, how to boot with what I/O scheduler, and information about the extolled Red Carpet novelties, the supported cluster file system Lustre-Lite, or the ZENworks Linux Management (ZLM).

As a technical basis, the SLES 9 rings the bell for the start of a fresh round in Suse's respective Novell Enterprise products: Novell Linux Desktop is forthcoming, as is Sun's Java Desktop Release 3. What should be worth further observation is the development of Open Exchange Server; its kernel has been put under the GPL recently. Under these conditions, it might be interesting to see what siblings the spaceship dockyard will churn out in the future. (avr)

DR. DIRK WETTER

is a consultant for Linux business solutions and IT security.

References

- [1] Dirk Wetter; Desktop-Linux; Grüner Bock im roten Kleid; Suse Linux 9.1 Professional; iX 6/2004, p. 68.
- [2] Ralph Hülsenbusch; 64-Bit-Server; Massentransporter; Rack-mounted Server mit HPs Alpha EV7 und AMDs Opteron; iX 12/2003, p. 64.
- [3] Dirk Wetter; Server-Linux; Familienzuwachs; Suse Linux Standard Server 8; iX 10/2003, p. 88.
- [4] Dirk Wetter; United Linux 1.0/Suse Linux Enterprise Server 8; United Colors of Linux; iX 1/2003, p. 64.
- [5] Interview: Andrea Arcangeli; kernel-trap.org/node/view/3148.

(Translation: Angelika Shafir)

