Connection

Novell Connection Magazine
SECOND QUARTER 2006 VOLUME 17 NUMBER 2

REPRINT

ARTICLES FIRST PUBLISHED IN NOVELL CONNECTION MAGAZINE.
NOVELL.COM/CONNECTIONMAGAZINE
REPRINTED WITH PERMISSION.

Novell®

By Ken Baker

ARTICLE REPRINT 01

A Hardened Back End

SUSE Linux Enterprise Server 10

A hundred years from now historians will look back at us as the "Super Size Me" generation. But it's not just at the fast food joints; we want more time, more fun, more money, more of our fair share, more of the good things of life. We're a society obsessed with getting more. Of course, in the business world it goes beyond obsession. The life and death of an enterprise hinges on its ability to get more from its people, more for its money and more from its IT investments.

When it comes to more performance, more scalability, more manageability, more reliability, more security, more support and more for your money, Novell has a strong reputation of delivering all that and more. But the new release of SUSE Linux Enterprise 10 takes supersized leaps in giving businesses so much more of what they are looking for in a world-class enterprise server operating system. As with any new release, the server component of the platform, SUSE Linux Enterprise Server, has of course more features than its predecessors, but there are three significant additions to the server that go a long way toward pleasing the ever-growing demand for more.

- AppArmor Application Security
- Storage Foundation
- Server Virtualization

> More Application Security

Security has long been one of the strengths of the Linux operating system, but sometimes, the vulnerabilities of individual applications open the door to security breaches. The full integration of the AppArmor application security framework in SUSE Linux Enterprise closes the door on those breaches, in essence, wrapping specialized layers of security around each individual application.

AppArmor is not new to SUSE Linux Enterprise Server. Version 9 shipped with an open source kernel module and a component that handled security mediation, but the capability to create and interpret policies came as a separate proprietary piece. Now AppArmor is completely open sourced under the GNU General Public License, and is tightly integrated within the SUSE Linux Enterprise framework.

So the big question is, how does AppArmor secure enterprise applications? It essentially employs a white listing methodology that defines what actions applications are allowed to take. This is opposed to black listing techniques that define all the things an application shouldn't do. Black lists work well until new vulnerabilities pop up that you don't know about, leaving you unprotected until vendors create new patches for them.

The white lists in AppArmor are essentially policies that mediate the file and directory access of an application, as well as an application's POSIX capabilities. POSIX is an IEEE standard that partitions root privileges into distinct sets. In other words, even though a server application might need root privileges to operate properly, instead of allowing the application to run with unbarred root privileges, AppArmor manages the application's file accesses and POSIX capabilities at the kernel level to limit the application to the set of resources and root privileges it actually needs.

This white list concept of mandatory access control per application has been around for a while. In fact, the National Security Agency created a similar solution called Security-Enhanced Linux, but its negative impact on performance and difficultly in implementing have been barriers to its deployment. On the other hand, AppArmor has minimal impact on performance, and the way it is integrated into SUSE Linux Enterprise Server makes it easy to deploy.

> Out of the Box Protection

Included inside of SUSE Linux Enterprise Server is a set of predefined AppArmor policies for common operating system services and applications, including Apache Web server, Postfix mail server, Sendmail mail server, OpenSSH, squid, ntpd, nscd and more. These policies will work out of the box without modification, except for Apache, which will require you to tell it things like where your home directory is located for your Web pages.

You create an AppArmor policy for an application in a four-phase process that you kick off by clicking the Novell AppArmor icon in the YaST management console.

> Server Analyzer

AppArmor provides a tool to help you determine which applications on your system should have an AppArmor policy associated with them. From the AppArmor menu in YaST, click on the AppArmor Reports icon and run the Application Audit Report. This automated process scans your server for applications that listen on open network ports and that don't already have an AppArmor policy defined. When the Server Analyzer finishes, it provides you a list of applications that can be accessed from outside the network that need policies.

> Policy Template Generator

To start the profiling process, click on the Add Profile Wizard icon from the main AppArmor menu. Choose which application you want to profile, then AppArmor will perform a fast static analysis that, in turn, creates a policy template specific for that application.

> Learning Mode

Once the policy template is created, the profile wizard automatically goes into learning mode, which is where the majority of the policy development takes place. It's during this phase that the AppArmor framework monitors how the application works, what directories and files it needs to access, and what type of accesses it needs. Of course,

AppArmor manages the application's file accesses and POSIX capabilities at the kernel level to limit the application to the set of resources and root privileges that it actually needs.

for the learning mode to work, you need to use your application as it would be during normal operation. Depending on the complexity of the application, this phase can take anywhere from a couple of hours to a couple of days to complete.

During the learning period, no policy rules are being enforced for the application, so make sure you're running the application in an isolated environment where you know your system won't become subject to attack while it's trying to learn good behavior.

> Interactive Optimizer

During the learning period, AppArmor creates a large log of events associated with the normal and acceptable behavior of the application. The Interactive Optimizer in AppArmor parses all of these events in a way that allows the profile wizard to ask you a series of questions that will enable it to quickly define the appropriate policy for that application.

For example, during learning mode the application might have accessed a certain file in a certain directory, so the profile wizard will ask you to choose from the following options to determine how the policy should govern that access:

Allow: This gives the application explicit access to that file in that directory, allowing the application to run as expected. Access can include or exclude read, write or execute modes.

Deny: This denies access to the file in that directory and could prevent the application from working properly.

Glob: This takes the last entry in the path and puts a wild card in its place, making everything in that directory accessible to the application using the defined access mode. This is useful when you know there are no security concerns with the application accessing other files in that directory and that it will indeed need to access those other files. Creating a wild card policy definition like this can significantly shorten the interview process because the Interactive Optimizer will automatically recognize that it

Figure 1 *The AppArmor profile wizard asks a series of questions that enable it to quickly define an appropriate policy for an application.*

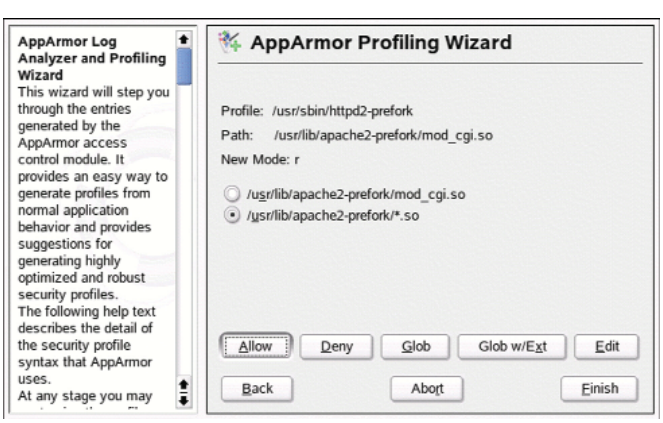


Figure 2 *AppArmor profiles can be viewed in a colorized format in vim, which highlights rules to which administrators should pay particular attention.*

```
/usr/sbin/ntpd {
  #include <abstractions/base>
  #include <abstractions/nameservice>
  capability ipc_lock,
  capability net_bind_service,
  capability sys_time,
  capability sys_chroot,
  capability setuid,
  /etc/ntp.conf                    r,
  /etc/ntp/drift*                  rwl,
  /etc/ntp/keys                    r,
  /etc/ntp/step-tickers            r,
  /tmp/ntp*                        rwl,
  /usr/sbin/ntpd                   rix,
  /var/log/ntp                     w,
  /var/log/ntp.log                 w,
  /var/run/ntpd.pid                w,
```

doesn't need to ask you about the application accessing other files in the directory. You can do a double glob by hitting the Glob button twice. This makes every file in that directory and all of its subdirectories accessible to the application.

Glob w/Ext: This is similar to the Glob, but instead of putting a complete wildcard in the last entry of the path, it puts a wild card with an extension in the last entry. For example, you could specify that all files with the extension .so in that directory can be accessed by the application. (SEE FIGURE 1.)

Edit: This pops up a dialog window that lets you edit the directory path and filename as desired.

As alluded to above, every time you answer a question about an operation in the application's event log, the profile wizard creates a policy rule for that event. If that rule satisfies other events in the log, the wizard skips those events during the interview process. The profile wizard also recognizes when existing policies for other applications might apply to the application on which you're working. So when it encounters an event similar to an existing policy, it will simply ask you if you want to apply that policy to your application. It's this type of intelligence inherent to the AppArmor profile wizard that gives it the ability to ask a few dozen policy questions to address a thousand or more application events.

When you finish answering the questions, the profile wizard will let you view and edit the policy in a colorized format, highlighting in yellow, policy rules that might be a cause for concern, enabling you to quickly analyze the policy. (SEE FIGURE 2.)

Once you finish creating policies, which are basic text files, you can easily distribute them to other servers in your environment that use those same applications and require those same policies. Also, AppArmor includes an update profile wizard that makes it easy to update existing application policies so you can implement changes to your system or simply add new rules. Once an application policy is in force, AppArmor keeps a log of application events that the policy rejects. Like the profile wizard, the update profile wizard parses the log and asks you questions about the event in a manner that lets you supplement your existing policy. (See *AppArmor Community.)*

> More Robust, More Scalable and More Available Storage

The new storage foundation in SUSE Linux Enterprise Server is all about delivering more in an enterprise server as well. It's about being a more robust and manageable foundation that can support small file systems or millions of files in terabytes of storage. It's about having a more flexible foundation that can support a wide range of applications from Web applications to databases. It's about having more high availability (HA) with improved clustering capabilities. It's about being a more scalable foundation with the ability to scale out with its parallel cluster file system.

All of this ability to do more is based on three tightly integrated storage foundation pillars:

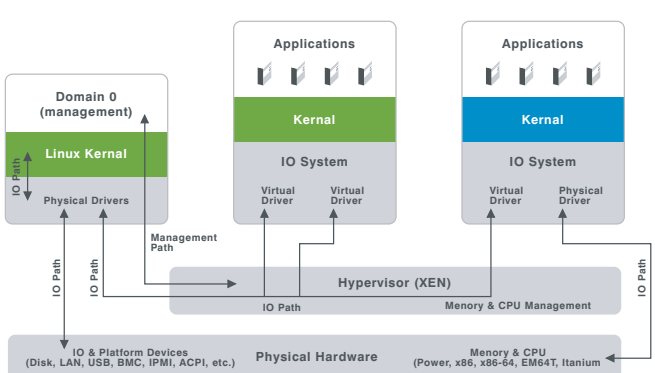
- HA Cluster Resource Manager
- Cluster Volume Manager
- Cluster Parallel File System

> HA Cluster Resource Manager

The HA Cluster Resource Manager aspect of the storage foundation is what you have historically known as high availability fail-over clustering. In other words, the ability for the cluster software to recognize that a service or server has failed, and then move that service to a surviving node in the cluster so that a server or service can remain highly available. The key component of the HA Cluster Resource Manager is the integration of the newly introduced Heartbeat 2.

An open source creation of the High Availability Linux Project, Heartbeat 2 increases clustering scalability dramatically from its

Figure 3 *The Xen hypervisor allocates resources for the different domains, presenting them with a virtual machine that acts like the domains' native architecture.*



Application Support

All of the existing applications that work in the SUSE Linux Enterprise Server 9 environment and run on what is often referred to as the high level LAMP stack (Linux, Apache, MySQL and PHP/Perl/Python), have no problems running on top of SUSE Linux Enterprise Server 10. The same is true for Java applications that don't make direct OS calls and support the Java Runtime Environment, which is included in the distribution.

But applications that call Linux threads might need some minor modifications. In SUSE Linux Enterprise Server 10, Novell has moved away from the older Linux Threads model in favor of the newer Native POSIX Threads Library (NPTL). Most developers will likely not notice this change because SUSE Linux Enterprise Server 9 used NPTL by default. Whether an application uses NPTL or Linux threads is typically determined at runtime when the LD_ASSUME_KERNEL environment variable is set. It really only becomes an issue for applications that were deliberately programmed to be aware of Linux Threads behavior.

Also relative to the area of application support, SUSE Linux Enterprise Server now shares a common code base with its desktop counterpart. As a result, Novell can provide its partners a single SDK for both the server and the desktop, simplifying development efforts by enabling them to code and test for both platforms in a single environment.

The bottom line is that with the combined support of 900+ software and hardware vendors committed to the platform, the support is there for SUSE Linux Enterprise Server and continues to grow.

Heartbeat 1 predecessor. Instead of being limited to two-node clusters, clusters can be increased to sixteen nodes. Actually, there is no set limit on the number of supported nodes, but it has been tested with up to sixteen cluster nodes.

A new advanced resource monitor capability has also been added to Heartbeat 2. This essentially allows an application vendor to incorporate a small monitoring agent into its application that can tell the HA Cluster Resource Manager if the application has stopped working properly. This is important for those times when an application doesn't necessarily crash, but has stalled or isn't responding as it should. The resource monitoring agent can detect this bad behavior and tell the HA Cluster Resource Manager that the application either needs to be restarted or moved to another server in the cluster.

Heartbeat 2 has also been tightly integrated with the other storage foundation pillars in SUSE Linux Enterprise Server so they can interact with each other. For example, if the HA Cluster Resource Manager fails over a resource it will also be able to fail over a file system with it. If changes occur on a volume, the Cluster Volume Manager will recognize the change and make sure it's reflected across the cluster. Another key aspect of the integration is the manageability that has been wrapped around the storage foundation. SUSE Linux Enterprise Server uses the open standards-based Common Information Model (CIM), greatly facilitating the management of the storage foundation.

> Cluster Volume Manager

The volume managers included in SUSE Linux Enterprise Server are basically the same ones that were provided in SUSE Linux Enterprise Server 9, which are Multi-Disk, Device Mapper, Logical Volume Manager and Logical Volume Manager 2. It still includes the Enterprise Volume Management System 2, which is an extensible enterprise level volume manager with plug-in capabilities and is cluster aware. In this release, Enterprise Volume Management System 2 is tightly integrated with Heartbeat 2 in the HA Cluster Resource Manager and the Oracle Cluster File System 2 (OCFS 2) in the Cluster Parallel File System.

> Cluster Parallel File System

SUSE Linux Enterprise Server still includes the non-parallel, but cluster-safe robust file systems of ReiserFS v3, XFS, and EXT3. But it's the Cluster Parallel File System addition that really gives businesses the increased scalability and reliability that they're looking for in an enterprise file system. A cluster parallel file system is not only cluster safe, but it is cluster aware. It allows multiple nodes to access the same volume and the same data simultaneously.

While SUSE Linux Enterprise Server supports partner solutions from Polyserve and IBM, it's the inclusion of Oracle's OCFS 2 that is the big news. While OCFS 2 was available in version 9, it only supported Oracle's Real Application Clusters (RAC). Beginning in January of this year, OCFS 2 was included in the mainline kernel, allowing it to be tightly integrated with the entire storage foundation and making it the basis for data center manageability. OCFS 2 support in SUSE Linux Enterprise Server has been extended to support other applications and databases as well. Linux, Apache, MySQL, and Perl and PHP stacks can all run on top of it. The OCFS 2 integration also acts as a critical piece to the Xen virtualization available in this release, allowing all nodes in a cluster to access the same virtualization image.

> Virtually More Servers

Easily the most exciting addition to the SUSE Linux Enterprise Server distribution is the new server virtualization capability. Based on the Xen open source project hosted by the University of Cambridge, server virtualization, in essence, does away with the need to tie server applications, services and files systems to a particular machine.

The power of virtualization is seen when you leverage its ability to run several of these self-contained virtual machines on a single compute server. (A compute server is a type of parallel processor that has no I/O except via a bus or other connection to a front-end processor that handles all I/O to disks, terminals, networks, etc.) This enables workload isolation, where instead of having multiple applications running on top of the same fat OS, you isolate each application to run on its own virtual machine. If an application happens to crash, since it's isolated, it won't affect any of the other services and applications running in their own virtual machines on the compute server. Additionally, since a virtual machine might only be running a single application or service, you will only need to load those operating system services and components that the application specifically needs. It also creates the opportunity for ISVs and integrators to develop highly customized virtual machines for the solutions they offer.

New Features

Xen server virtualization, high availability enhancements to the storage foundation, and the AppArmor application security framework are by far the most prominent features new to SUSE Linux Enterprise Server 10. But other new improvements also enrich the latest enterprise server offering from Novell.

In the area of performance and scalability, it ships with the latest Linux Kernel 2.6.16 that can scale up to 1024 CPUs. Other CPU performance and scheduler enhancements include the addition of multicore/hyperthreading support and the ability to partition multiprocessor machines by execution areas. Machines can be split in terms of CPUs and processes can be bound to certain CPUs. Pluggable I/O schedulers are now optimized for specific workloads and the included kernel patches enable support for the upcoming Networking Acceleration Technology from Intel. It also takes advantage of the latest open source hardware drivers for things like IPMI, power management, USB, Firewire, RAID, SAS, SATA, Multipathing and Fibre channel. It includes both iSCSI initiator and target so you can create a fully functional SAN using commodity server hardware.

Manageability enhancements include a new Novell Customer Center that lets you easily manage your subscriptions. Novell ZENworks Linux Management helps you manage all your Linux servers and desktops. You can now manage large user bases and access control to networks and applications using OpenLDAP. You can fully integrate SUSE Linux Enterprise Server into Novell eDirectory environments, as well as into Active Directory infrastructures.

In addition to AppArmor security, it includes other security enhancements such as MIT Kerberos 1.4.3 for authentication, Snort version 2.4.3 for network intrusion detection and the AIDE file manipulation monitoring system. In the area of new services, SUSE Linux Enterprise Server ships with SAMBA 3.0.21b, NFS v4, Cyrus IMAP Deamon Version 2.2, MySQL 5.0, PostgreSQL 8.1, Apache Web Server 2.2.0 and PHP 5.1.

Paravirtualization is a virtualization technique that uses a software interface to virtual machines that is similar but not identical to that of the underlying hardware.

Server virtualization adds another level to High Availability through Virtual Machine Live Migration. HA clustering allows for automatic restart of a failed service. This means the application or service is down for a short period of time, although usually not long enough to cause production problems. Virtual Machine Migration allows an application or service running in a virtual machine to be moved from one physical machine to another in a cluster without any restart. This means there is no downtime and the complete application running state is preserved when it is moved. This is a great advantage by allowing normal maintenance of the system during production hours.

When it comes to server consolidation, the ability to run multiple virtual machines on a single compute server combined with the ability to have virtual machines running different guest operating systems can also greatly simplify those efforts. Even though SUSE Linux Enterprise Server is the host operating system for server virtualization, virtual machines can run different paravirtualized guest operating systems. (Paravirtualization is a virtualization technique that uses a software interface to virtual machines that is similar but not identical to that of the underlying hardware.) So, for those legacy applications that need to run on top of legacy operating systems, they can be contained in their own individual virtual machine and consolidated onto a single compute server. SUSE Linux Enterprise Server can support full virtualization when coupled with forthcoming hardware technologies from both AMD and Intel.

The following are some of the key architectural components of Xen virtualization. (SEE FIGURE 3.)

Domain: Domains are the containers for self-contained virtual machines.

Hypervisor: At the heart of Xen, the hypervisor lies on top of the physical layer running at the most privileged hardware protection ring and has the responsibility to allocate resources for the domains, presenting the domains with a virtualized view of the domains' native architecture.

Domain 0: As a privileged domain, Domain 0 hosts the management framework for Xen virtualization. It is the first domain started by the hypervisor at boot time. It manages all other domains. Domain 0 can host all the physical drivers on which other virtual machines rely. SUSE Linux Enterprise Server serves as the host OS running in Domain 0 and may also run as a guest on the same physical hardware.

Unprivileged Domain: This is any domain other than Domain 0, sometimes referred to as DomU.

Driver Domain: Domains other than Domain 0 can be granted specific access to a particular hardware I/O device so access does not need to be mediated by Domain 0. These driver domains, while optional, can reduce traffic at Domain 0, improving performance.

Paravirtualization: A guest operating system that has been paravirtualized is one that has been modified to recognize that it is running on top of a hypervisor to improve performance.

Full virtualization: Fully virtualized operating systems do not realize they have been virtualized and as a result the hypervisor traps and emulates every I/O and hardware instruction.

Out of the box, SUSE Linux Enterprise Server will provide fully virtualized and paravirtualized guest operating system support for itself. Later in the year, Novell plans to add paravirtualized support for SUSE Linux Enterprise Server 9 SP3 and support for NetWare running in an Open Enterprise Server environment.

Notwithstanding official statements of support, kernels that have been ported by the open source community to run as paravirtualized guests on Xen, include Linux 2.4, Linux 2.6, NetWare 6.5, NetBSD, FreeBSD, Plan9 and OpenSolaris. Even though its stated support for guest hosts is limited, Novell has made it clear that it is firmly committed to making Xen on SUSE Linux Enterprise Server the best virtualization platform available. So in the future, you can definitely expect to see more.

> More, more, more

If all you really want is more fries and a drink to go with your burger, just tell the person at the drive-up window you want a Combo Meal. But if that new raise, car or house depends on getting much more than you dreamed possible out of your existing enterprise server, then SUSE Linux Enterprise Server will deliver. **N**

AppArmor Community

To make it even easier for organizations to take advantage of the application security benefits AppArmor provides, Novell initiated an open source project and invites you and other community members to contribute to the future development of AppArmor as well as to submit AppArmor profiles you have created for your own applications. The goal is to build a large repository of predefined application security policies that can help you and others in the open source community to quickly and easily put AppArmor to work to protect your IT environments. For more information on this project, visit *opensuse.org/apparmor.*

Novell

Novell Connection Magazine
1800 South Novell Place
Provo, UT 84606

Tel: (801) 861.7000
E-mail: editor@novell.com
Online: www.novell.com/
connectionmagazine