mission
critical
linux

Product Description

Convolo™ Cluster Software
Version 1.2

December 2000

Welcome to Linux's leading high availability cluster software product – Convolo cluster software. If you have just purchased the product, thank you. We are confident that you will be completely satisfied with the software and appreciate the benefits it brings to your Linux environment. If you are considering purchasing Convolo cluster software, we hope that this document will provide the answers to any technical questions you may have.

This Product Description provides a technical overview of Convolo cluster software and its features.

It also outlines supported hardware, prerequisite software, and other information that will be useful to you when configuring your Linux cluster. Information regarding the product Warranty and Support Services is also included.

Sincerely,

Mission Critical Linux, Inc.
100 Foot of John Street
Lowell, Massachusetts 01852
www.missioncriticallinux.com
(+1) 978.606.0200
(+1) 877.62LINUX

Description

Convolo cluster software provides a high availability computing environment for multiple Linux-based systems running static (read-only) and dynamic update (read/write) applications. It provides support for two systems accessing shared mass storage (not including boot disks) under a single application management environment. Within this highly integrated environment, systems retain their independence because they use private copies of the Linux operating system. Thus, the Linux systems can boot and shut down independently while benefiting from common storage resources. Applications running on servers in a Convolo cluster configuration access shared disk resources in a fully coordinated manner, guaranteeing the highest levels of data integrity.

Because disk storage resources are shared, Convolo cluster configurations offer higher availability than standalone Linux systems. Properly configured Convolo cluster configurations can withstand the shutdown or failure of various components. For example, if one server in a Convolo cluster is shut down for planned maintenance, applications and network clients can continue operation after restarting on or reconnecting to the remaining cluster member. Because mass storage can be shared cluster wide, restarted applications and clients are able to access their original data. Applications and network clients can be designed to survive these events automatically.

Physical configurations can be designed to have no single-point-of-failure, and will provide very high levels of application availability. Convolo cluster software provides comprehensive integrity guarantees for applications and data through multiple polling mechanisms and programmable power management.

Features

Convolo cluster software includes the following features:

- Failover support of data-centric application services such as databases and network file systems (NFS), while maintaining the highest level of data integrity.

- Management of applications and their required resources by means of service definitions and control scripts, called *Cluster Services*. Cluster Services are provided as part of the product or created by the system administrator or application provider. Cluster Service definitions include a service name, the application name, and required resources for the application, such as IP network addresses, disks, mount points, and NFS exports. When a failure condition or system administration action requires the cluster to migrate an application from one cluster member to the other, the complete service is migrated. This ensures that the application has access to all its required resources.

- Application failover support. In a Convolo cluster application, availability is maintained using a mechanism known as *failover*. Application failover occurs under system administration control, or when a server shuts down, fails, or (optionally) boots and joins the cluster. In a controlled failover, a Cluster Service and its associated application is shutdown cleanly using a service shutdown script; it is then restarted on the other cluster node using a service startup script. Failure of one cluster member running a Cluster Service is detected by the other cluster member, which automatically invokes the Cluster Service startup script to host the application. Cluster Services can be assigned a Preferred server – when that Preferred server joins the cluster, its preferred services will be automatically migrated to it. Cluster Services can be shut down and restarted on either cluster node at any time by a system administrator.

- Support for two systems connected to shared SCSI or Fibre Channel storage subsystems,

such that both systems can directly access the same disk storage with no intermediate disk or file server.

- For all standard Linux filesystems, access granularity is at the disk partition level. The Convolo infrastructure ensures that filesystems associated with Cluster Services are only mounted by one server at a time, thereby guaranteeing data coherency.

- Two system administration interfaces: a command line interface and a Web browser-based interface. The command line interface is suitable for script activities, while the Web interface is suitable for day-to-day management activities. The interfaces permit a system administrator to configure and monitor all aspects of the cluster.

- Distributed service definition. Cluster services are defined by the system administrator and maintained by Convolo software as part of the configuration's Quorum data. A service must be defined only once; the service is discovered dynamically by the other cluster member through the Quorum verification mechanism or at boot time if that other node is shut down at the time of service definition.

- A comprehensive configuration verification utility to simplify correct cluster setup. The utility checks the hardware and software setup including, for example, serial device names, power control options, Ethernet heartbeat channels, and accessability of all heartbeat ports. It also checks the completeness and integrity of information in the cluster configuration database. In addition, it runs status checks on the different pieces of the clustering software: cluster daemon operation, power control status, and quorum partition validation.

- Quorum-based cluster membership. A shared cluster disk is configured with two small partitions that are used to maintain Quorum information. Quorum information is used by the server nodes to validate cluster membership. If a server cannot access the Quorum partitions, it will not join the cluster. The two partitions are shadowed (contain identical data) to provide redundancy of Quorum information, thereby increasing cluster availability.

- Multiple polling interconnect support. Cluster nodes poll each other using Local Area Network (LAN) and/or RS-232 serial lines. Convolo cluster software automatically polls across all configured LAN and serial interconnects, thereby increasing the resilience of the cluster configuration in the face of interconnect failures. At least one polling interconnect must be configured. In addition, a Quorum-partition based polling mechanism is used by the cluster members to validate cluster membership in the event that all polling interconnects fail.

- Concurrent access to raw disk partitions. Both servers can access raw disk partitions concurrently. An application that wishes to take advantage of this feature must be designed to operate in a distributed environment.

- Support for controlling programmable controllers. To facilitate the highest level of data integrity, support is provided to control programmable power controllers to power cycle a failed node during catastrophic system hangs.

Configuration Support

The following hardware, software, and application infrastructures are supported by Convolo cluster software.

Application Service Support

Cluster Service startup and shutdown scripts are provided for the following applications:

- Oracle 8i

- MySQL

- Apache Web Server

- Zeus Web Server

- IBM DB2

- NFS Server

Template Cluster Services are provided to permit creation of scripts to control any site-specific application.

Hardware Requirements

Minimum hardware requirements are as follows:

- Two Linux servers configured with one Ethernet interface, one RS-232 serial interface, and a local boot disk.

- One shared SCSI or Fibre Channel storage subsystem for cluster services, applications, and data.

- Optional: two programmable power controllers. Use of programmable power controllers is strongly recommended in environments that require stringent data integrity guarantees.

Computer System

Convolo cluster software supports Intel® 32-bit (IA-32) hardware platforms. Convolo is supported on a wide range of Linux distributions. Refer to the system support list in the Software Requirements section of this document for more information.

Cluster nodes do not need to be identical. They can be of different speed, memory size, and SMP configuration, and can be supplied by different vendors. Selecting the appropriate configuration for each server depends on the applications and failover strategy to be deployed. The shared disk storage namespace must be identical for both cluster nodes, so it is recommended that the nodes have symmetric storage subsystems.

Additionally, it is recommended that each node in the cluster run the same Linux distribution and version.

Storage Subsystem

The Linux boot device may be configured using any commodity hard disk, including IDE and SCSI. RAID support for the boot device may be provided by means of host software, adapter or controller based methods.

The shared storage subsystem must be carefully configured to meet the availability requirements of the overall Convolo Cluster configuration. The following guidelines should be observed:

- Convolo Cluster software supports SCSI and Fibre Channel for the shared storage subsystem.

- Production environment guidelines: For robustness, ease of maintenance, and correct handling of storage bus reset conditions, it is strongly recommended that the shared storage be configured using a multi-port storage controller. A multi-port controller provides a separate physical bus for each cluster node. The multi-port controller must support concurrent access to all logical units by all cluster nodes. These controllers are easier to configure and terminate correctly, and greatly simplify removal of a node from the configuration while the cluster remains operational. Because of the bus isolation provided by multi-port controllers, it is possible to use any modern, good quality host adapter in these configurations.

- Development environments guidelines: In development environments, it is often acceptable to use a lower cost, single-port storage controller. In these configurations, all cluster nodes as well as the storage controller are connected to the same storage bus. Note that depending on the host adapter and Linux device driver used, these configurations may not reliably recover from cluster node failures and reboots.

Mission Critical Linux, Inc. is actively evaluating adapters and Linux device drivers that provide reliable recovery in these configurations.

Please refer to the www.missioncriticallinux.com/products/convolo Web site for the latest adapter and device driver information.

- Convolo cluster software does not use SCSI Reserve/Release commands or SCSI Target Mode, so support for these functions is not required.

- It is not possible to use host/server-based RAID software or adapter-based RAID firmware for the shared storage subsystem with Convolo cluster software. RAID capabilities, if required, must be provided by a storage controller.

Network Hardware

Convolo cluster software supports Local Area Network and serial line hardware supported by the specific Linux distribution on which Convolo is installed. Refer to the network hardware support list for the Linux distribution for additional information.

Power Controller

Server systems in Convolo cluster configurations can optionally be connected to programmable power controllers. The controllers are cross-coupled between the systems so that each system can control the power supplied to the other system. This feature permits Convolo cluster software to ensure data integrity under certain rare failure scenarios. It is strongly recommended that power controllers are included in production environments.

Development environments with less stringent data integrity requirements may be configured without programmable power controllers. For more information on configuring Convolo cluster solutions without power controllers, refer

to the Convolo cluster *Installation and Administration* manual.

Convolo cluster software provides support for the following power controllers:

- Western Telematic Inc:

  U.S.A. part number:

  > RPS-10 M/HD

  European part number:

  > RPS-10 M/EC

  Western Telematic Inc.
  5 Sterling
  Irvine, California 92618
  Tel: (+1) 949.586.9950
  www.wti.com

Software Requirements

Convolo Cluster software supports the following distributions:

- Red Hat Linux 6.1, 6.2, or 7.0
- Linux-Mandrake 7.1
- SuSE Linux 6.3 or 6.4
- TurboLinux Server 6.0
- Debian GNU/Linux 2.2
- Corel Linux 1.2
- Caldera OpenLinux eServer 2.3

Note: When using other distributions not listed above, Convolo cluster software requires Linux Kernel 2.2.16. The Convolo cluster software distribution CD contains a complete 2.2.16 kernel source pool and a corresponding set of patches that are required to ensure correct cluster behavior. These patches include:

- NFS Server and utilities
- Raw I/O support
- Intel EEPro network driver fixes
- I/O fairness algorithm enhancement

Documentation

Comprehensive configuration, installation, and system administration information for Convolo

cluster software is provided in the Convolo Cluster *Installation and Administration* manual. This hardcopy manual is supplied with the software product kit. The manual is also provided electronically in PDF and HTML format on the Convolo cluster software distribution CD.

Distribution

Convolo cluster software is provided as a boxed software product including the following items:

- Convolo cluster Product Description.

- License terms and conditions certificate, with one or two license numbers (depending on product order).

- Distribution CD-ROM.

- Hardcopy Convolo Cluster *Installation and Administration* manual.

There is no requirement to build Convolo cluster software prior to installation. The software is provided as fully built images in RPM and DEB formats. However, it is necessary to rebuild the Linux kernel to include mandatory features (for example – support for Raw I/O operations and NFS failover).

Licensing

Convolo cluster software is based on Kimberlite cluster technology, which has been made available by Mission Critical Linux, Inc. to the Open Source community under the GPL (GNU Public License) license. Information on Kimberlite can be obtained at `oss. missioncriticallinux.com`.

Convolo cluster software is provided with one or two server licenses, depending on the specific product order. *Both servers in the cluster must be licensed separately.* A valid license is required during software installation and to obtain Warranty services.

Kimberlite and Convolo Comparison

The following list outlines the major differences between the Convolo cluster software product and the Open Source Kimberlite cluster technology:

- Convolo cluster Software provides full support for NFS server failover, including the correct handling of file locking. This support is not provided in the Kimberlite source code.

- Although both Convolo and Kimberlite provide a Web browser-based system administration GUI, only Convolo provides the ability to configure and monitor a Convolo cluster. With Kimberlite, this interface is limited to monitoring functions only.

- The Configuration Verification utility provided with Convolo cluster software is not provided with Kimberlite.

- There is no requirement to build from source either Convolo cluster software or any of the software packages on which it is dependent. Kimberlite, however, is available only in source format.

- Convolo cluster software is provided as a boxed product with printed documentation and CD-ROM. Kimberlite is available only for download over the Internet.

- Convolo cluster software is provided with a full product warranty and provides a known, tested clustering solution.

Warranty

Convolo cluster software is provided with 90-days warranty and support services. Refer to the license agreement at the front of the Convolo Cluster *Installation and Administration* manual for warranty details.

Support Services

Support for Convolo cluster software is provided by Mission Critical Linux, Inc. Support Services are provided using our Web browser-based LifeLine customer relationship management system, email, and telephone. Service is provided during business hours (8:00 a.m. – 8:00 p.m. EST) Monday through Friday. Warranty services include assistance with Convolo cluster configuration, installation and setup, and problem resolution. Critical problems receive a response within an hour.

To receive warranty services, please have your Convolo cluster software license number available. Please call or send email to enable your LifeLine account.

Additionally, Mission Critical Linux, Inc. offers a range of support service contracts to meet any support requirement, including 7 day x 24 hour coverage as well as long-term support for Convolo cluster software and the complete Linux software environment. For additional information, please contact the Sales organization at Mission Critical Linux, Inc.

Training

Training courses are available for Convolo cluster software and cover configuration, installation, management, and theory of operation. For additional information, please contact the Sales organization at Mission Critical Linux, Inc.

Ordering

Convolo cluster software is orderable as follows:

- On the Internet at: www.missioncriticallinux.com/store/cart.cgi

- Directly from the Mission Critical Linux, Inc. Sales organization.

Part Numbers:

Convolo cluster software boxed package with:

Licenses	Part Number
1 Server License	CCS-120-001
2 Server Licenses	CCS-120-002

Contact Information

Phone: (+1) 978.606.0200
FAX: (+1) 978.446.9470
Toll Free: (+1) 877.62LINUX

Email:
sales@missioncriticallinux.com
support@missioncriticallinux.com

Internet:
www.missioncriticallinux.com

Acknowledgements

Mission Critical Linux, Inc. would like to thank the following companies for their kind provision of hardware during the development of Convolo Cluster software. The listed hardware was used extensively during the development of the product and can be thoroughly recommended:

- Intel Corporation – Engineering support platforms (IA32 and IA64)

- IBM Corporation – Netfinity servers

- VA Linux Systems, Inc – FullOn servers

- Penguin Computing, Inc. – Servers

- Winchester Systems, Inc. – RAID storage controllers

- Western Telematic, Inc. – Power controllers

Linux is a registered trademark of Linus Torvalds.

Convolo is a trademark of Mission Critical Linux, Inc.

All other trademarks are the property of their respective owners.