

SUSE Linux Enterprise High Availability Extension

Piotr Szewczuk

konsultant

pszewczuk@novell.com

Novell®

SUSE® Linux Enterprise Product Lines and Products

Server Product Line

- **SUSE Linux Enterprise Server**
- SUSE Linux Enterprise Server for System z
- SUSE Linux Enterprise Real Time Extension
- **SUSE Linux Enterprise High Availability Extension**
- SUSE Linux Enterprise Mono Extension
- SUSE Linux Enterprise Point of Service

Desktop Product Line

- SUSE Linux Enterprise Desktop

Virtualization Product Line

- SUSE Linux Enterprise Server

Appliance Product Line

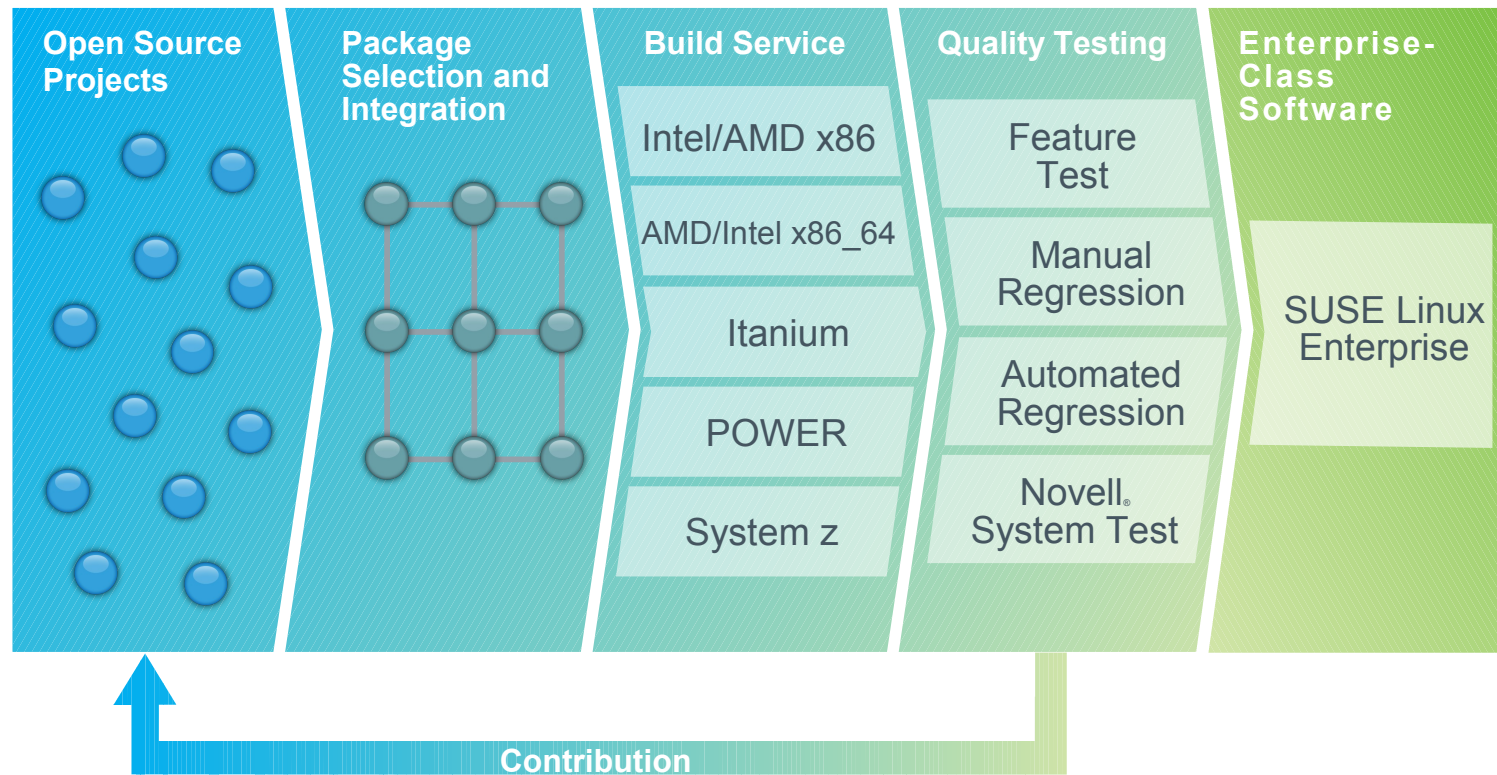
- SUSE Linux Enterprise JeOS (just enough operating system)
- SUSE Studio Online
- SUSE Appliance Toolkit

Specialized Support Offers

- SUSE Linux Enterprise Server Subscription with expanded support (for Red Hat migrations)
- SUSE Linux Enterprise Server Priority Support for SAP Applications



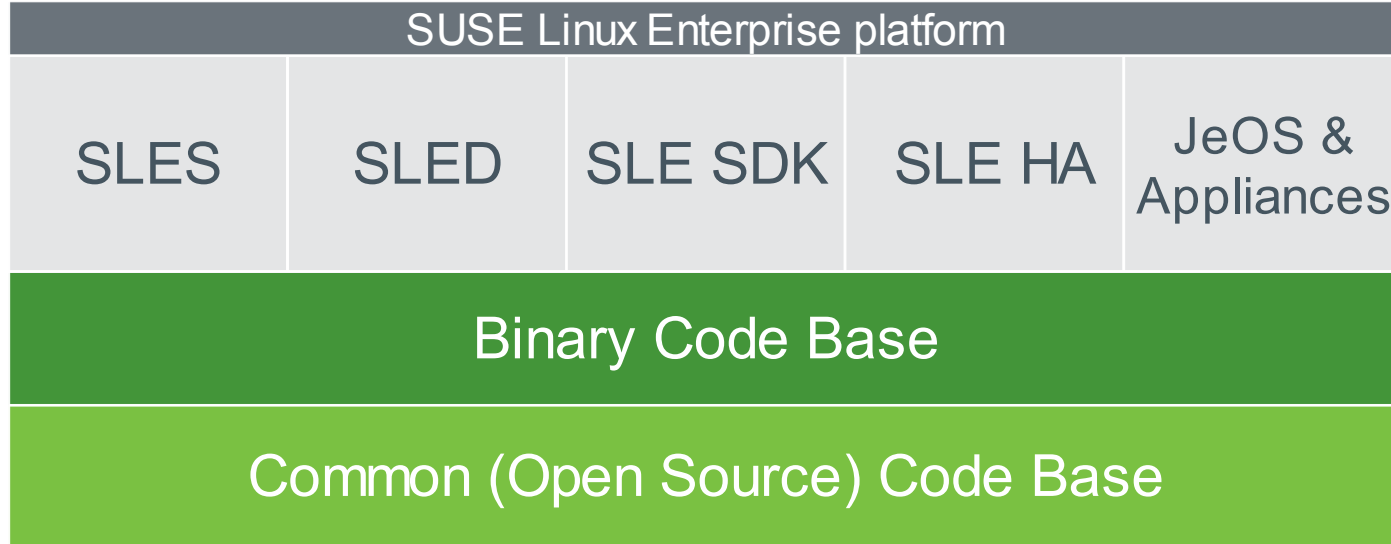
The SUSE® Build Service* Advantage



- Reduces production problems
- Consolidates IT skills across disparate systems
- Delivers critical updates in hours – not days or weeks

* SUSE Build Service is the internal entity of the openSUSE® BuildService infrastructure

SUSE® Linux Enterprise Common Code Base



- Foundation for SUSE Linux Enterprise products
- Fully supported core system: L3 support commitment
- Architectures:
Intel/AMD 32-bit, AMD64/Intel 64, Itanium,
POWER, System z

Introduction to the High Availability Extension

The bottom of the slide features a series of five horizontal, glowing white lines that fade out towards the right, set against the blue background.

Data Center Challenges



Minimize unplanned downtime



Ensure quality of service



Contain costs



Utilize resources



Effectively manage multiple vendors



Minimize risk

SUSE® Linux Enterprise High Availability Extension Value Proposition and Benefits

- An affordable, integrated suite of robust open source clustering technologies that you can use to implement highly available physical and virtual Linux services.
- Used with SUSE Linux Enterprise Server, it helps you maintain business continuity, protect your data, and reduce unplanned downtime for your mission critical Linux workloads.
- Benefits
 - Cost effectively meet your service-level agreements
 - Ensure continuous access to your mission-critical systems and data
 - Maintain data integrity
 - Increase resource utilization



What Do You Get?

- High Availability clustering solution
 - Service availability 24h a day
- Oracle Cluster File System 2 (OCFS2)
 - Sharing and Scaling data-access by multiple nodes
- Distributed Remote Block Device 8 (DRBD)
 - Provide low-cost “SAN” through disk replication via TCP
- IP Load-balancing (Linux Virtual Server)
 - Scale network services using multiple nodes
- User-friendly tools
- Documentation and whitepapers



Notable Features

- Up to 16 nodes per cluster generally supported
 - Specific support cases beyond that are possible
- Local Data Center clusters only
 - But see road map for SP1
- Shared storage supported, but not required
 - MPIO fully supported
- Available on all architectures for SUSE® Linux Enterprise Server

Virtualization-aware

- Xen-aware cluster resource manager extends high availability to virtual workloads
- Recognize, monitor and manage services running within virtual servers, as well as services running in physical servers
- Guest systems can be managed as services by the cluster
- Cluster virtual servers with physical servers for higher availability and higher utilization



SUSE® Linux Enterprise High Availability Extension

Competitive Landscape



	HP HP-SG	IBM HACMP	Veritas VCS	MSFT Cluster	Steeleye Lifekeeper	RHAT Ad. Plat.	Novell SLES10	Novell SLE HA 11	Novell SLE HA 11 SP1
Upper Node Limit	↔	↓	↔	↔	↔	↑	↔	↔	↑
Network Load-Balancing	↓	↓	↓	↓	↓	↔	↔	↔	↑
System Recovery	↓	↓	↓	↓	↓	↓	↓	↓	↑
Disk Mirroring	↔	↓	↓	↓	↓	↑	↔	↑	↑
Platform Support	↔	↓	↑	↓	↑	↔	↔	↔	↔
HW Support	↔	↓	↑	↑	↑	↑	↑	↑	↑
Storage Support	↔	↓	↔	↔	↔	↑	↑	↑	↑
ISV Support	↔	↓	↑	↑	↔	↔	↔	↔	↔
Setup, Installation and Configuration	↔	↔	↑	↑	↔	↔	↓	↔	↑
GUI	↔	↔	↑	↑	↑	↔	↓	↔	↑
Command line	↑	↔	↑	↓	↔	↔	↔	↑	↑

Installation

Media

Available as an ISO image than can be used as is, burned to a single CD, or integrated into a network installation source.

The SUSE® Linux Enterprise Server can access multiple install sources in addition to the standard media.

This means that the High Availability Extension can be integrated into the initial install of a new server whether it is done manually or automated with AutoYaST.



The High Availability Pattern

The screenshot shows the YaST2 installation tool interface. On the left, a list of patterns is displayed under the filter 'Patterns'. The 'High Availability' pattern is selected. On the right, a table lists packages associated with this pattern. The 'drbd' package is highlighted, and its details are shown in a pop-up window at the bottom.

Package	Summary	Installed (V)
<input type="checkbox"/> drbd	Distributed Replicated Block Device	(8.2.7-4.2)
<input type="checkbox"/> ipvsadm	A Utility for Administering the Linux ...	(1.24-297.1)
<input type="checkbox"/> lvm2-clvm	Clustered LVM2	(2.02.39-1)
<input type="checkbox"/> ocfs2console	Oracle Cluster Filesystem 2 GUI to...	(1.4.1-8.20)
<input checked="" type="checkbox"/> ocfs2-tools	Oracle Cluster File System 2 Core ...	(1.4.1-8.20)
<input type="checkbox"/> openais	The OpenAIS Standards-Based Cl...	(0.80.3-16)
<input type="checkbox"/> pacemaker	The Pacemaker scalable High-Ava...	(1.0.1-8.1)
<input type="checkbox"/> pacemaker-pygui	Management Tools for Pacemaker...	(1.4-6.6)
<input type="checkbox"/> sle-hae-release	SUSE Linux Enterprise High Availa...	(11-1.10)
<input type="checkbox"/> yast2-cluster	Configuration tool for HA clusters	(2.15.0-6.3)
<input type="checkbox"/> yast2-drbd	YaST2 - DRBD Configuration	(2.13.1-21)

drbd - Distributed Replicated Block Device

Drbd is a distributed replicated block device. It mirrors a block device over the network to another machine. Think of it as networked raid 1. It is a building block for setting up clusters.

The level of support is unspecified

Buttons:

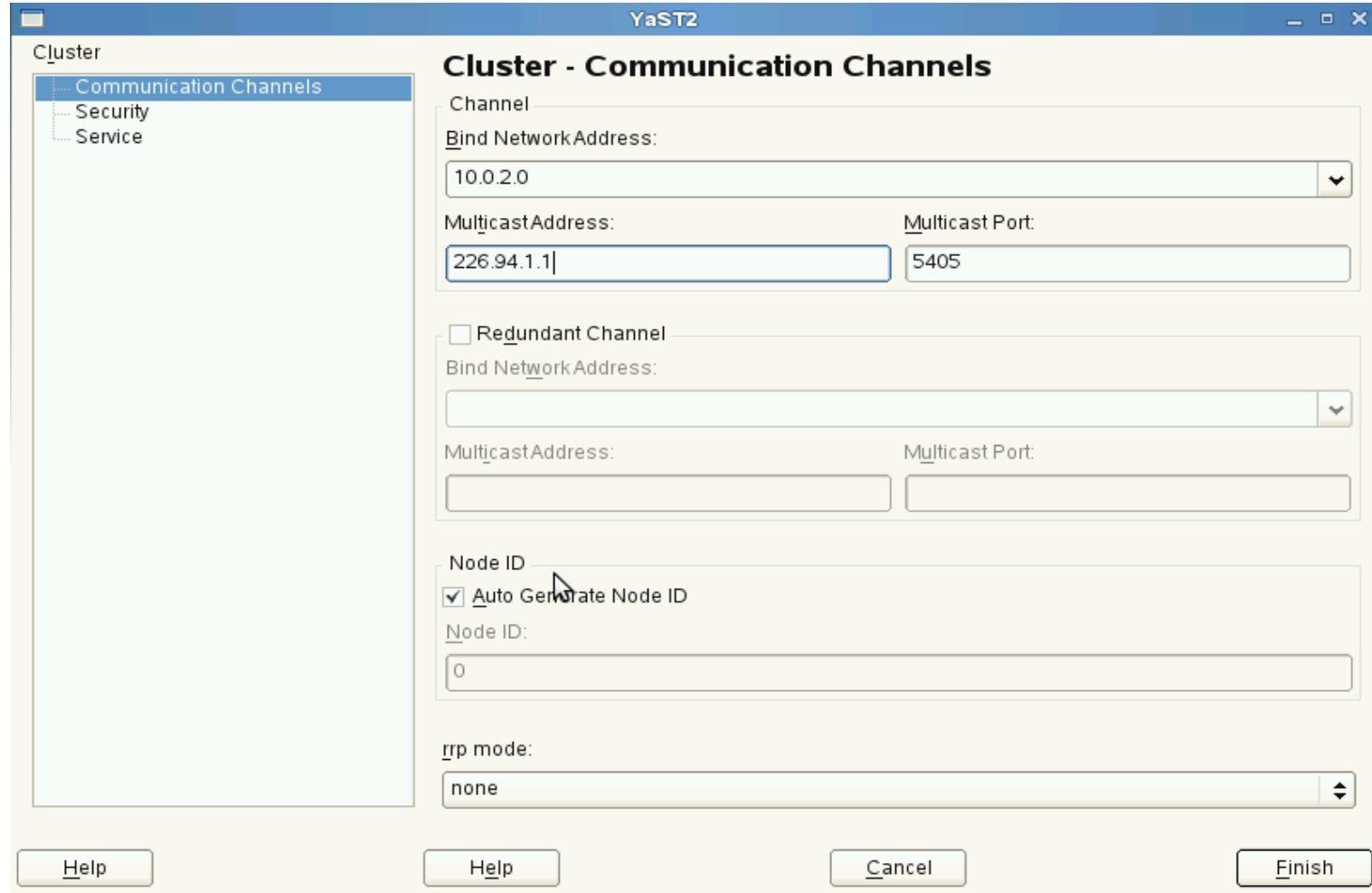
OpenAIS

- Open Source Initiative's certified implementation of the Service Availability Forum Application Interface Specification
- Leading standards-based communication protocol for server and storage clustering
- Messaging and membership layer

Details

- The underlying protocol is called Totem
- Multicast only
- Uses one mcast group only in SUSE® Linux Enterprise Server 11
- Multiple multicast groups on the way for SP1
- Push redundancy down the stack – NIC bonding
- Config file is `/etc/ais/openais.conf`

Configuring OpenAIS



The image shows the YaST2 configuration window for OpenAIS communication channels. The window title is "YaST2" and the main title is "Cluster - Communication Channels". On the left, a tree view shows "Cluster" with sub-items "Communication Channels", "Security", and "Service". The "Communication Channels" section is selected. The main area contains the following fields:

- Channel**: A dropdown menu.
- Bind NetworkAddress:** A text field containing "10.0.2.0".
- MulticastAddress:** A text field containing "226.94.1.1".
- Multicast Port:** A text field containing "5405".
- Redundant Channel**: A checkbox.
- Bind NetworkAddress:** An empty text field.
- MulticastAddress:** An empty text field.
- Multicast Port:** An empty text field.
- Node ID**: A section containing:
 - Auto Generate Node ID**: A checked checkbox.
 - Node ID:** A text field containing "0".
- rrp mode:** A dropdown menu containing "none".

At the bottom, there are four buttons: "Help", "Help", "Cancel", and "Finish".

SUSE® Linux Enterprise High Availability Extension Features



Cluster aware file system and volume manager

- OCFS2

- > Shared-disk POSIX-compliant generic cluster file system
- > Cluster-aware POSIX locking
- > Parallel I/O

- cLVM2 (Clustered Logical Volume Manager)

- > Convenient, single, cluster-wide view of storage
- > Clustering extensions to the standard LVM2 toolset
- > Eliminates need to learn a new set of tools



SUSE® Linux Enterprise High Availability Extension Features (continued)



User-friendly tools

- Unified command line interface
 - > Powerful tool for installing, configuring and managing Linux clusters
 - > For more experienced IT professionals
- Graphical user interface
 - > Simple tool for monitoring and administering clustered environment
 - > Does not require in-depth knowledge
- YaST modules
 - > DRBD
 - > OpenAIS
 - > Multipath



SUSE® Linux Enterprise High Availability Extension Features (continued)



Resource Agents

- For popular third-party applications included at no extra charge
 - > SAP Instance and Database, IBM WebSphere Application Server, DB2, and Informix, Oracle and VMware
- For popular open source applications included at no extra charge
 - > Apache, Ipv4 and IPv6, LVM, RAID, Pure-FTPd, Route, ServeRAID, Squid, VIPArip, Xen, Xinted, DRBD, Novell® eDirectory™, iscsi, mysql, nfsserver, and postgres, sfex, tomcat, filesystems
- For the most up to date list of resource agents, visit www.novell.com/products/highavailability



Fencing

- Isolating resources – usually storage - from nodes
- Two ways of doing this
 - Isolating the resource – SAN switch remapping
 - Isolating the node – power off or panic

STONITH

- Shoot The Other Node In The Head
- We are doing node fencing here
- A daemon run as part of the cluster stack
- Has support for a number of fencing methods

```
stonith -L
```

- Displays the list

DRBD

Continuous data replication

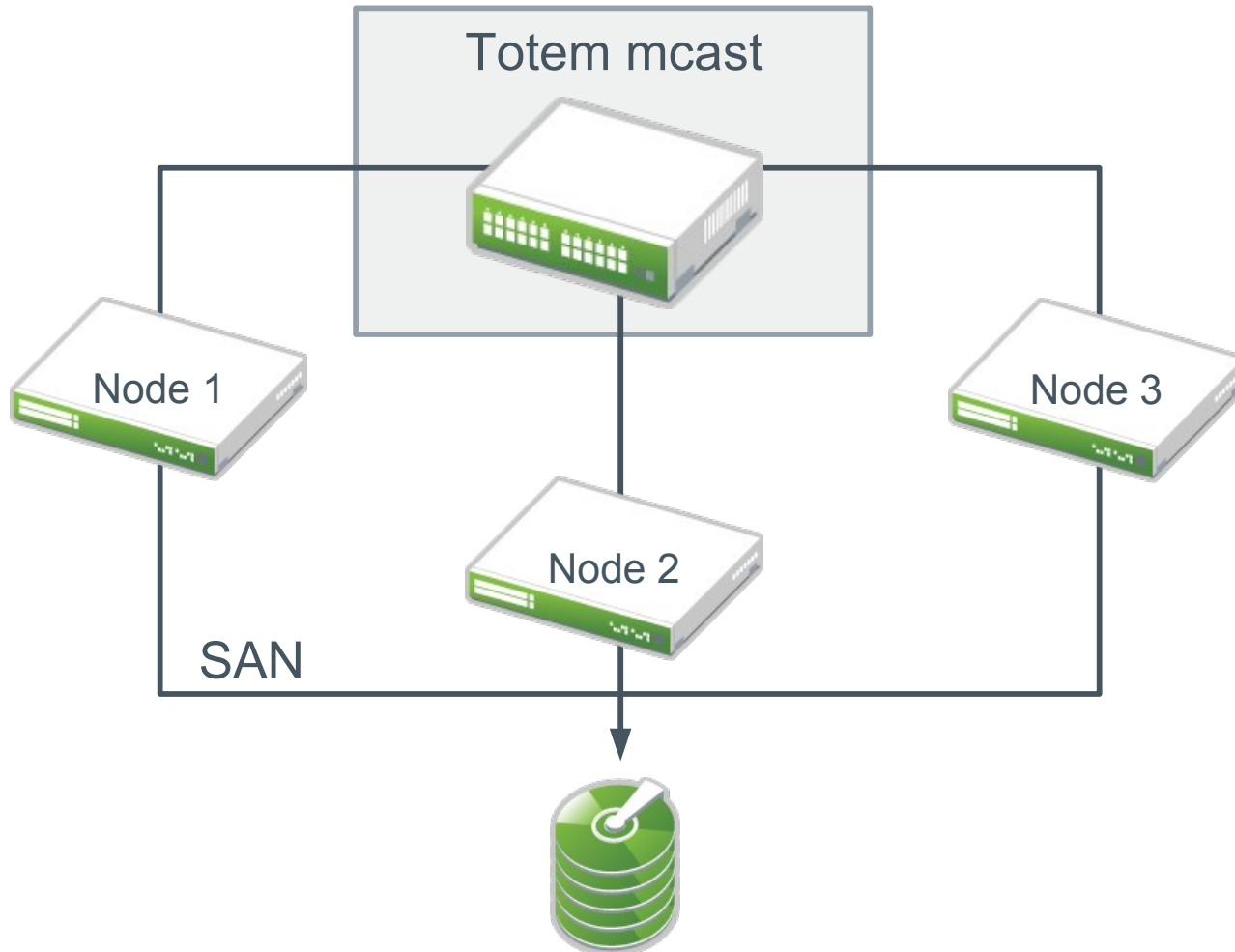
- DRBD8 (Distributed Replicated Block Device)
 - > Leading open source networked disk management tool
 - > Build single partitions from multiple disks that mirror each other
 - > Fast data resynchronization capabilities
 - > Supports both synchronous and asynchronous mirroring
 - > Provides replicated storage area network (SAN) semantics, allowing cluster-aware file systems to be used without additional SANs



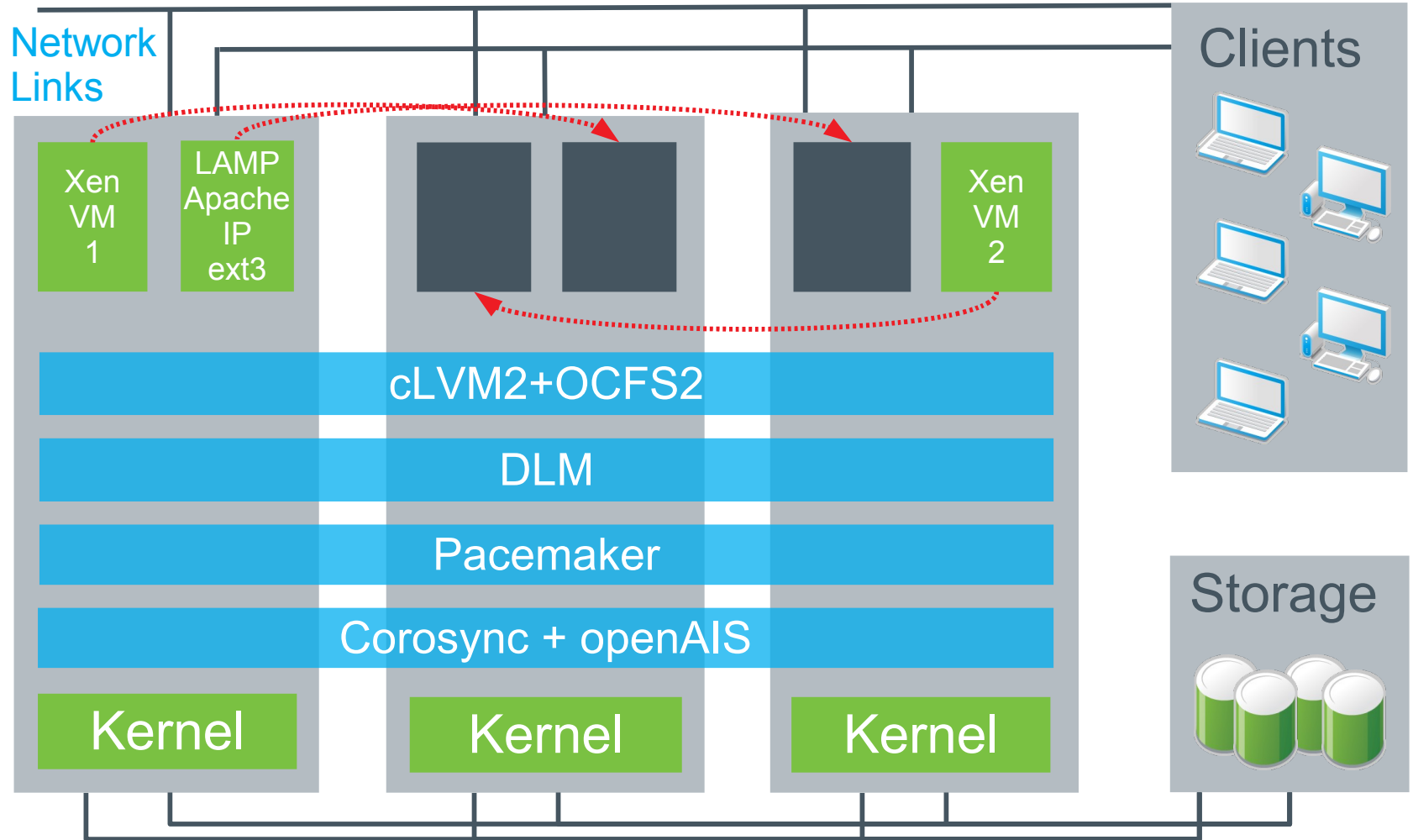
Cluster Architecture

The bottom of the slide features a series of five horizontal, glowing white lines that fade out towards the right, set against the blue background.

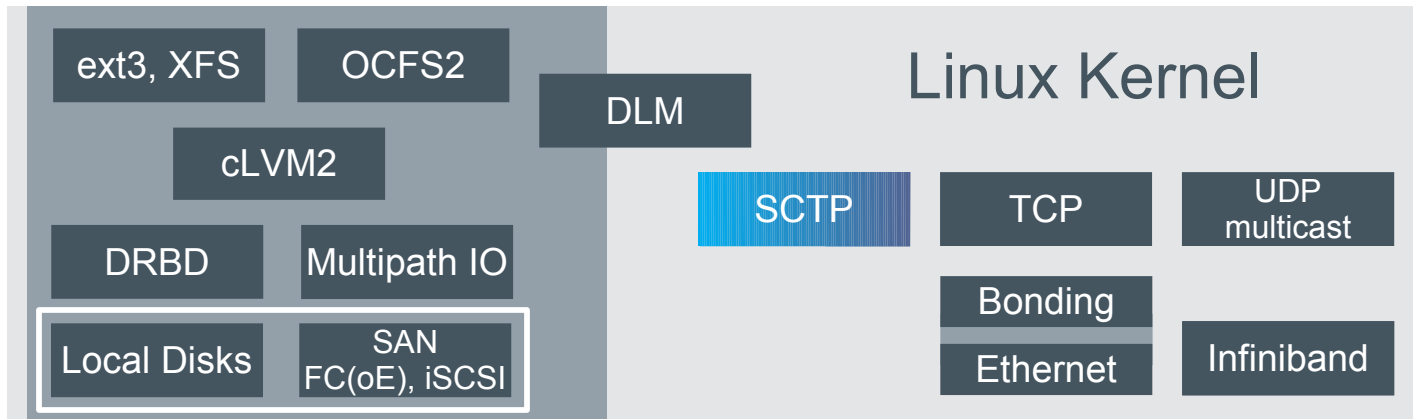
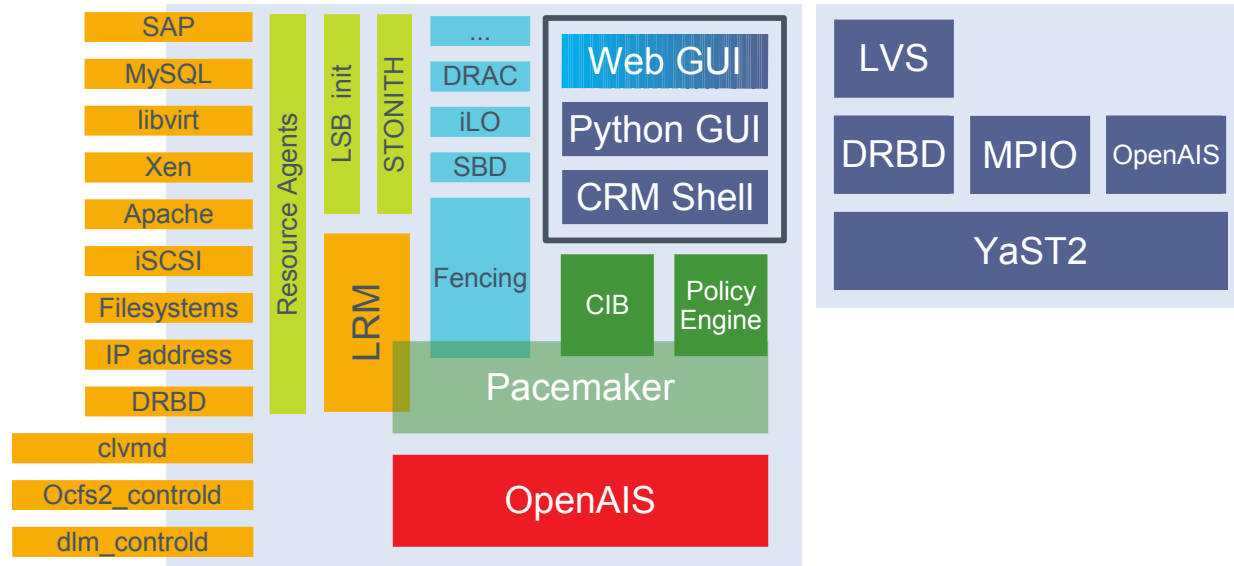
A Simple Cluster



3 Node Cluster Overview



Detailed View of Components Per Node:



High Availability Extension SLES11 / SLES11 SP1 / Future

SUSE® Linux Enterprise High Availability Extension 11

Key Capabilities Today



- **Service Availability 24/7**
 - Policy driven clustering
 - > OpenAIS messaging and membership layer
 - > Pacemaker cluster resource manager
- **Sharing and Scaling Data-access by Multiple Nodes**
 - OCFS2 Clusterfile system
 - Clustered logical volume manager
- **Disaster Tolerance**
 - Continuous data replication via IP
 - Distributed replicated block device
- **Scale Network Services**
 - IP load-balancing
- **User-friendly Tools**
 - Graphical user interface
 - Unified command line interface



Key capabilities Tomorrow

- **Storage Based Quorum**
 - Enabling the use of a storage device as a quorum instance to match traditional Unix setups and to prevent split brain scenarios
- **Integrated Samba Clustering**
 - Integration of Samba with OCFS2 for higher throughput and scale out of SMB access
- **Metro-Area Clusters**
 - Supporting clustering between different data center locations
- **Multilevel Administration Rights**
 - More fine-grained control of cluster operation and administration
- **Enhance Data Replication**
 - Improved DRBD capabilities
- **DR Framework**
 - A tool set (ReaR) for node recovery

SUSE® Linux Enterprise High Availability Extension

Key capabilities Future



- **Usability and Management**
 - Full web GUI
 - Improved access control
- **Ease of Use**
 - Guided and automated configuration
 - Prepackages applications
 - Preloaded clusters
- **Clusters Functionality**
 - Wide area clusters
 - Improved data replication
 - Unix cluster stack leadership
- **Backup and Disaster Recovery**
 - Backup integration
 - DR automation

Key Features in Service Pack 1

Web GUI – Enabling cross platform management

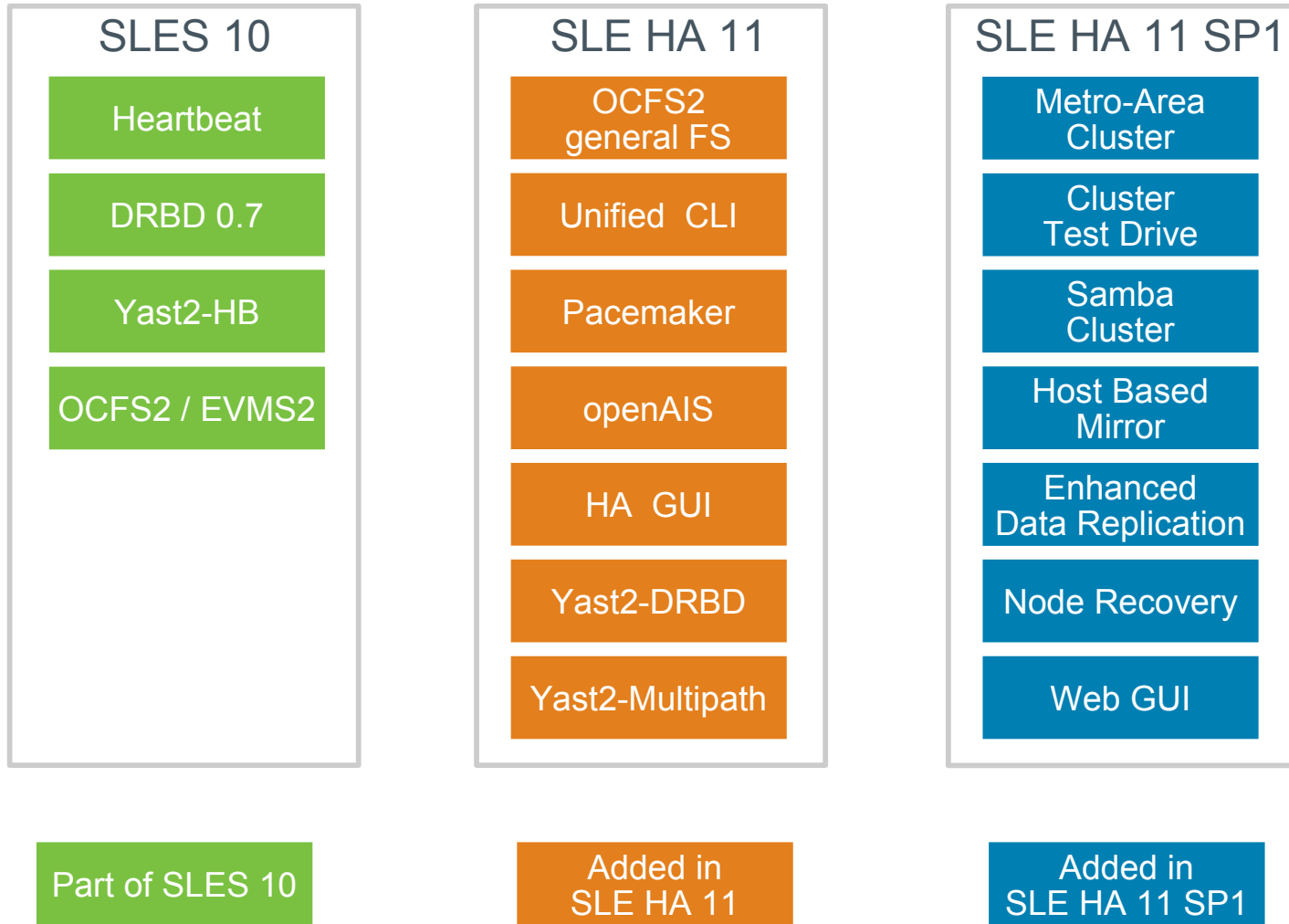
Integrated Samba Clustering – Integration of Samba with OCFS2 for higher throughput and scale out

Metro-Area Clusters – Supporting clustering between different data center locations

Enhance Data Replication – DRBD with Linbit

Node Recovery – ReaR for node recovery

SUSE® Linux Enterprise High Availability Extension HA Stack from 10 to 11



Use Cases

SUSE Linux Enterprise High Availability Extension

- Metro Area Cluster
 - Protection against site impact
 - Leverage single name space for Storage & Network
 - Distance up to 20 miles / 30 km
- Node Recovery
 - Quickly achieve full operational status after node failure
 - Based on ReaR—Relax and Recover
 - Integration into backup solution
 - Automated and unattached operation

Use Cases

SUSE Linux Enterprise High Availability Extension

- Third off-site storage replication
 - Disaster protection
 - Asynchronous data replication
- Host-based Mirroring
 - Affordable storage protection
 - Storage vendor independent
 - active-active / active passive

Use Cases

SUSE Linux Enterprise High Availability Extension

- Clustered Samba network file system (CIFS)
 - High available & scalable using multiple nodes
 - Transparent failover using cluster wide locking
 - Integration with OCFS2
- Storage Quorum “Coverage”
 - Protection against *Split Brain*
 - Storage access controls cluster behavior
 - “SBD” poison-pill mechanism for node isolation via shared storage eases fencing configuration

Operation Improvements

SUSE Linux Enterprise High Availability Extension

- Web GUI
 - Cross platform access to cluster status
- IP load-balancer YaST module
 - Tool chain integrated configuration
- Cluster Test Drive
 - Dry run for cluster changes and reaction
- Cluster Configuration Management
- Cluster Configuration Sync
 - Keep configuration consistent across nodes using *csync2*

Feature Improvements

SUSE Linux Enterprise High Availability Extension

- Cluster File System: OCFS2
 - File operations up to **5x** faster, Reduced File Fragmentation by **95%**
 - On-line resizing
 - Snapshots and linked clones (reflinks)
- Data Replication: DRBD
 - Partnership with Linbit for development, support and maintenance
 - Faster operation by block bitmap compression
- New and Updated Resource Agents
 - DRBD, Clustered Samba, postfix, Squid, iSCSI, SAP, libvirt (Virtualization/KVM)

SAP certification status

SUSE® Linux Enterprise Server 11 is fully certified for SAP

<http://www.sap.com/linux/> → supported platforms

This includes also all extensions to SUSE Linux Enterprise Server 11

Status of High Availability Stack @SAP Linux Lab

- SAP Linux Lab (which includes Novell® staff) is working on defining the high availability stack for SAP on SUSE® Linux Enterprise Server 11
- Goal: reliable and well-tested stack
- Base: SUSE Linux Enterprise High Availability Extension 11 replaces the High Availability Storage Infrastructure from SUSE Linux Enterprise Server 10
- What remains in version 1: MD-RAID, SFEX
- Scheduled for evaluation: cLVM, OCFS2, SDB

Online Demo

<http://www.novell.com/media/content/suse-linux-enterprise-server-high-availability-extension-demo.html>

Novell.[®]

Unpublished Work of Novell, Inc. All Rights Reserved.

This work is an unpublished work and contains confidential, proprietary, and trade secret information of Novell, Inc. Access to this work is restricted to Novell employees who have a need to know to perform tasks within the scope of their assignments. No part of this work may be practiced, performed, copied, distributed, revised, modified, translated, abridged, condensed, expanded, collected, or adapted without the prior written consent of Novell, Inc. Any use or exploitation of this work without authorization could subject the perpetrator to criminal and civil liability.

General Disclaimer

This document is not to be construed as a promise by any participating company to develop, deliver, or market a product. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions. Novell, Inc. makes no representations or warranties with respect to the contents of this document, and specifically disclaims any express or implied warranties of merchantability or fitness for any particular purpose. The development, release, and timing of features or functionality described for Novell products remains at the sole discretion of Novell. Further, Novell, Inc. reserves the right to revise this document and to make changes to its content, at any time, without obligation to notify any person or entity of such revisions or changes. All Novell marks referenced in this presentation are trademarks or registered trademarks of Novell, Inc. in the United States and other countries. All third-party trademarks are the property of their respective owners.