

Samba HA Cluster on SLES 9

High Availability with Samba 3 on Linux

Rolf Schmidt

Support Engineer

Novell, Inc.

Rolf.Schmidt@novell.com

May 29, 2006



Novell.[®]

Agenda

Introduction

Samba Setup

Definitions

High Availability Cluster

Opensource packages used

detailed Server Setup

Test

Discussion of Results

Planning of a High Availability Cluster Samba 3 on SLES9

Introduction

What we would like to have

- Samba 3 Fileserver
- High Availability
- No third party software
- works 'out of the box' with SLES9
- independent of architecture

(I use 'architecture' to describe different hardware like S390, i386 or PowerPC)

- password synchronisation
- update strategies

Introduction

What we will use

- Software
 - SLES9
 - Windows© XP Professional Client
- Hardware
 - i386 Machine
 - 2 LPARs on an OpenPower 720, Power 5 Architecture

Resources

- Configuration files and Howto's (in progress) at http://beta.suse.com/private/rschmid/samba_ha_project/
- Contact rschmid@suse.de

Introduction

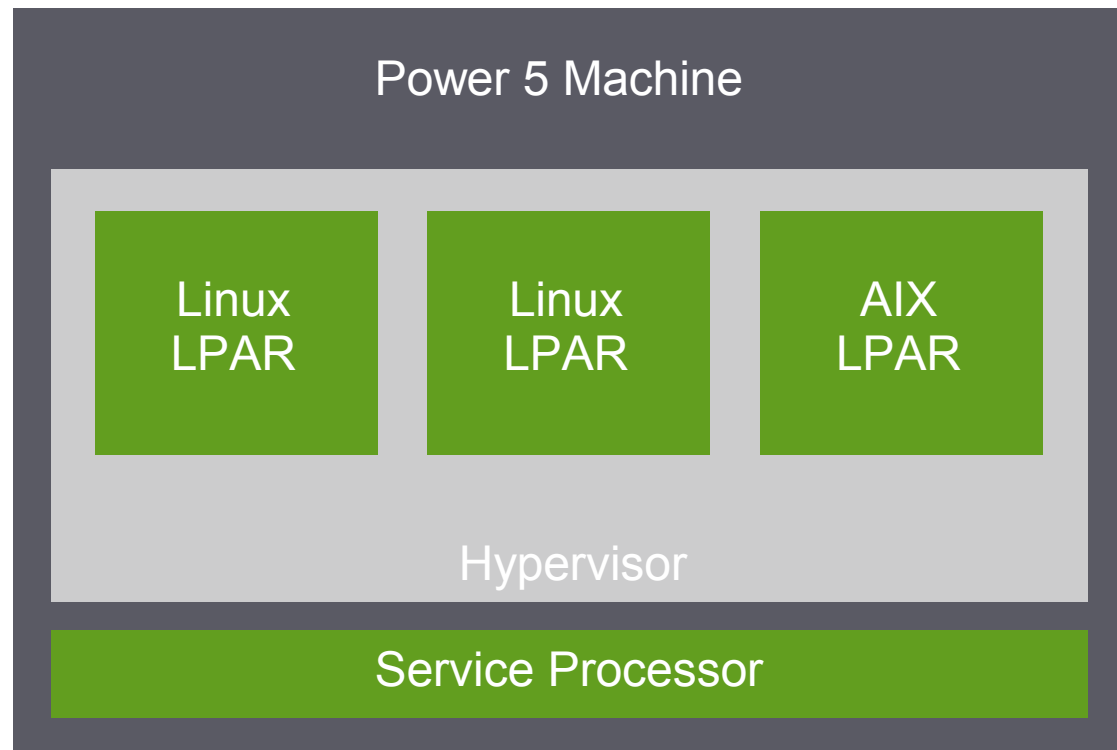
Power5 Architecture

- Power5 Processor provides SMT, Simultaneous Multi-threading
- Power5 provides micropartitioning of Resources
- The System in a Power5 Machine is in one or more LPARs and can be Linux, AIX or i5
- Power5 Machines provide fully transparent Virtualisation
- All IO not dedicated to a LPAR and all Systemresources are controlled by the Hypervisor
- The Administrator controls the Hypervisor via the Service Processor

Introduction

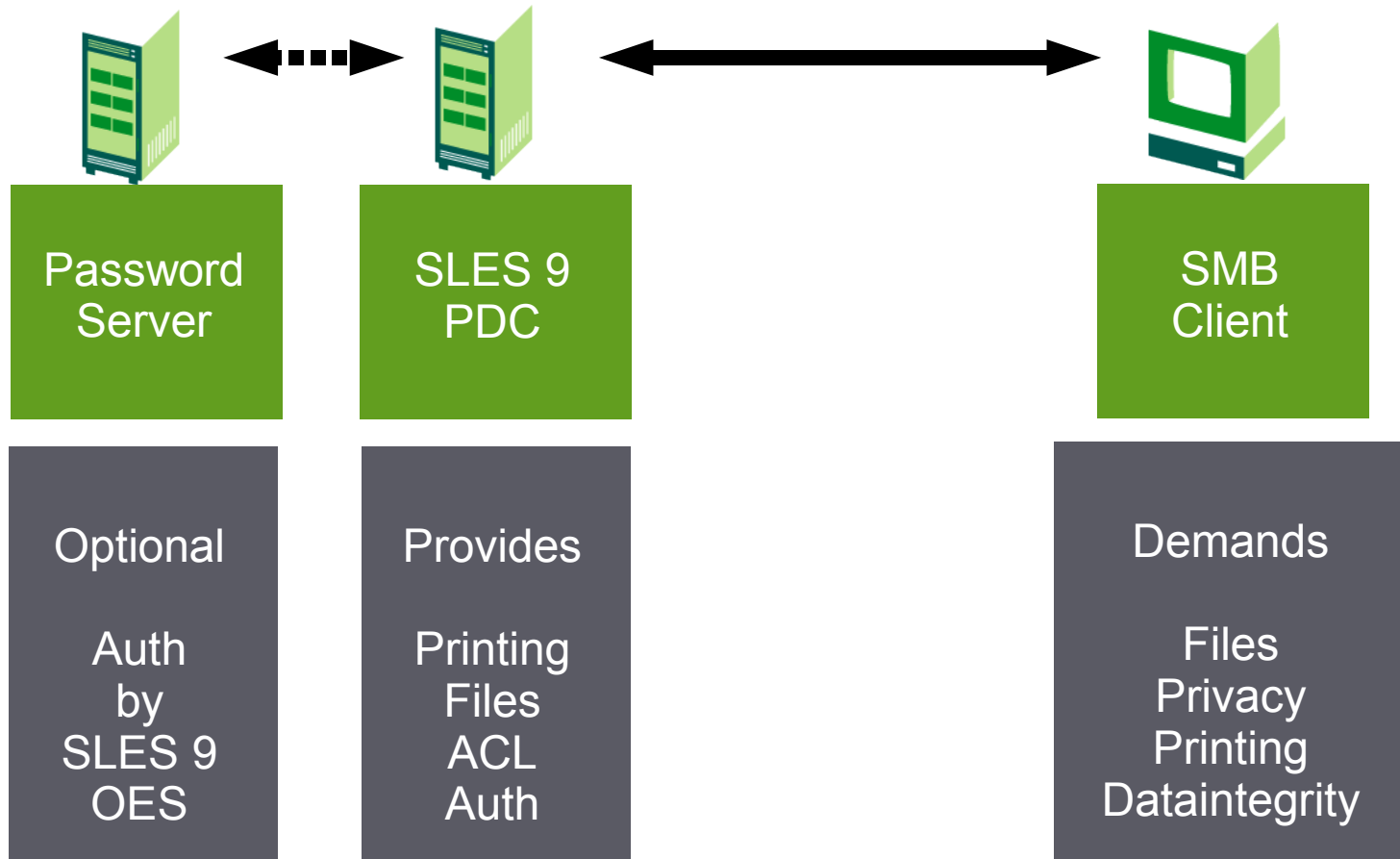
Power 5 Architecture

- The LPARs 'float' in the Hypervisor Layer



Samba Setup

Domain Example with 'Linux only Servers'



Samba Setup

Our Server will provide the following

- Fileserving for smb Clients
- User Authentication
- Access Control
- Data Integrity

Our failure settings are

- Server crashes / dies – Heartbeat
- Service dies – Nagios

Definitions

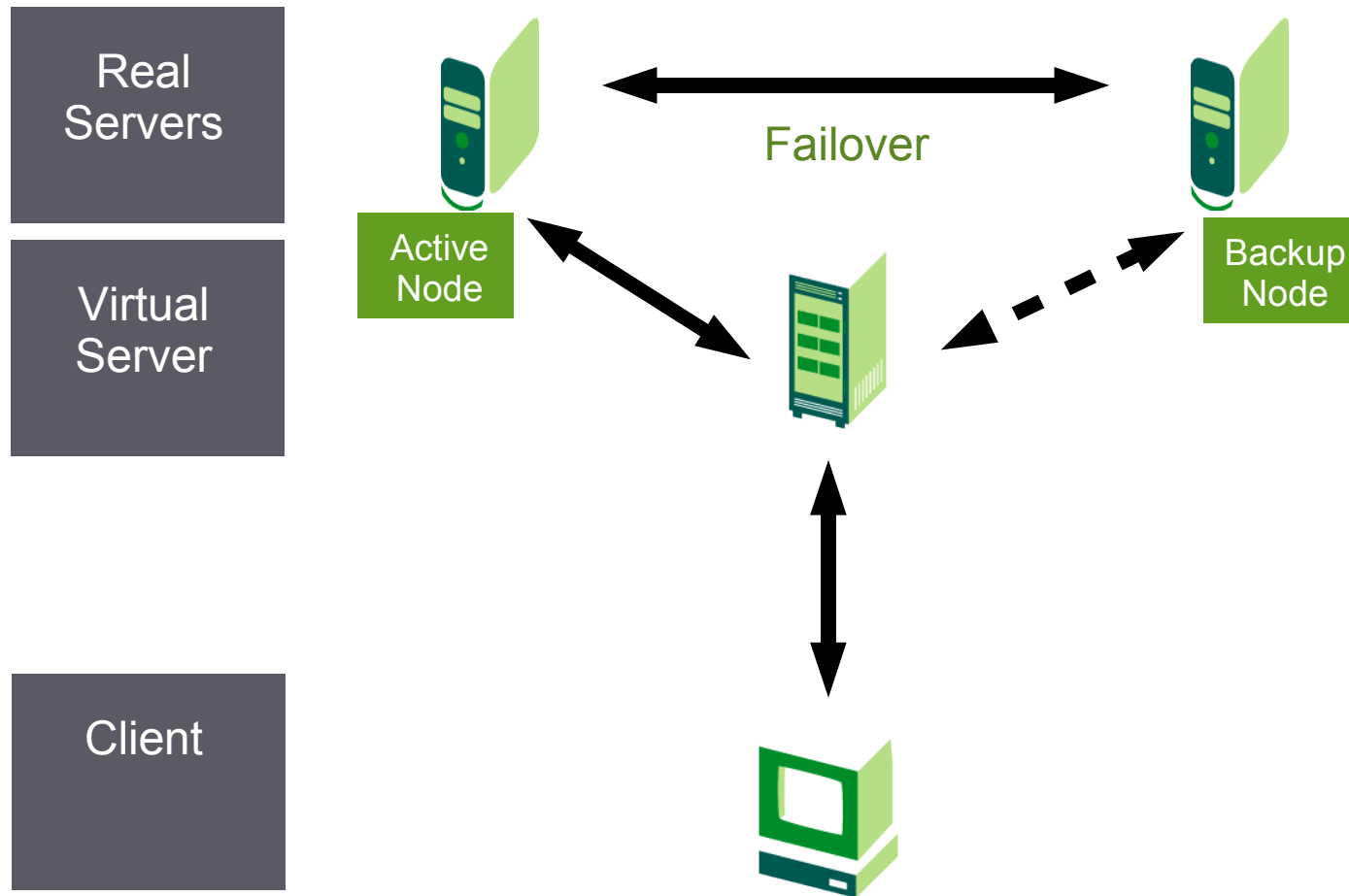
High Availability vs High Performance Clustering

- What is HA?
- What is HPC?
- HA as 'cheap' or 'small' HPC can be dangerous!



High Availability Cluster

Visualisation of a High Availability Samba Server



Definitions

Stateless vs Stateful

- **What is a Stateless Protocol?**

Stateless is an application or service that does not maintain state. Meaning it can process requests individually, without any dependence on previous requests. Each interaction request has to be handled based entirely on information that comes with it. An example would be NFS.

- **What is a Stateful Protocol?**

Stateful means that an application or service keeps track of the state of interaction, usually by setting values in a storage field in the computers memory designated for that purpose. An example would be smbfs/cifs. An explanation of the different information stored on a Samba Server can be found at <http://ictd.undp.org/rc/forums/tech/sdnptech/msg02093.html> written by Gerald Carter.

- **Consequences**

Samba is by no means well suited for High Availability as the information needed for the stateful protocol stored on the dying node is lost during failover.

Opensource packages used

Samba – samba3.0.20b-3.4

- Server

DRBD – drbd-0.7.14-0.3

- Raid 1 over Network
- Alternative: Storage

Heartbeat – heartbeat-1.2.3-2.9

- Node monitoring

Nagios – nagios-1.2-73.1

- Service monitoring

LDAP – openldap2-2.2.24-4.12

- User authentication

detailed Server Setup - DRBD

DRBD

- Distributed Replicated Block Device
- Provides RAID1 device over Network
- Better would be a storage connected via FC

Setup and Configuration

- `/etc/drbd.conf`
- copy to second node
- `uname -n`
- `insserv drbd`

detailed Server Setup - DRBD

Setup and Configuration - continued

- `drbdsetup /dev/drbd0 primary`
- `watch cat /proc/drbd`
- `mkfs.reiserfs /dev/drbd0`

tuning and caveats

- watch the port you are using
- use high enough syncer rate
- always use protocol c
- dont forget to set acl's

detailed Server Setup - Heartbeat

Heartbeat

- Node monitoring – not services
- STONITH

Setup and Configuration

- `uname -n`
- `/etc/ha.d/authkeys`
- `/etc/ha.d/ha.cf`
- `/etc/ha.d/haresources`
- `/etc/ha.d/rpc.cfg` and `/etc/ha.d/ibmhmc.cfg` (stonith)

detailed Server Setup - Heartbeat

Setup and Configuration - continued

- copy files to second node
- modify `/etc/ha.d/ha.cf` on second node
- `rheartbeat start`

Caveats

- `/etc/ha.d/authkeys` must be 0600 and owned by root
- `/etc/ha.d/haresources` must be the same on both nodes
- don't make heartbeat start by default now

detailed Server Setup - Nagios

Nagios

- Nagios checks for services
- Nagios can start a failover if a service fails
- configuration of Nagios can be painful

Setup and Configuration

- `/etc/nagios/contacts.cfg`
- `/etc/nagios/hostgroups.cfg`
- `/etc/nagios/hosts.cfg`
- `/etc/nagios/services.cfg`

detailed Server Setup - Nagios

Setup and Configuration

- `/etc/nagios/nagios.cfg`
- `/etc/nagios/checkcommands.cfg`
- `/usr/lib/nagios/eventhandlers/failover`

Settings

- We check only TCP 139 for connection
- should be LDAP and all the ports used by the Samba Server
- failover will trigger heartbeat to hand over to other node
- test your settings, as said, nagios is tricky

detailed Server Setup - LDAP

LDAP

- provides User Authentication
- Can be replicated with DRBD (NOT GOOD as it creates a single point of failure) or slurpd

Setup and Configuration

- Just setup normally during node installation
- don't add your users and groups yet
- if migrating use an Idif to feed ldap with data

detailed Server Setup - Samba

Samba

- Fileserver

Setup and Configuration

- Setup as PDC and WINS Server
- use LDAP as first passdb backend
- set netbios name to 'host IP_FROM_HARESOURCES'
- set LDAP password with `smbpasswd -w PASSWORD`

detailed Server Setup - Migration

Files and Directories to move onto DRBD Device

- /home
- /var/lib/ldap
- /var/lib/samba
- /etc/samba/smb.conf
- /etc/samba/smbpasswd
- create links from files and directories on DRBD device to system
- failover and create symlinks on second node

Test

Failover

- trigger with `/usr/lib/heartbeat/hb_*` - manual failover
- disable service on node – test of Nagios

Samba

- connect to Server. Edit file. Failover during editing.
- Edit again, further and failover again

Failover

- disable network device on node – test of heartbeat

Discussion of Results

provided Services

- Authentication of User and Access Control
- Password synchronization
- File integrity

Update and configuration

- central configuration files on both servers in sync
- Update possibilities
- manual failover as tool for updates

Data integrity is essential

- We moved to the Server because our Data is the most important.
- Data corruption defeats the purpose of the Server

Planning of a High Availability Cluster Samba 3 on SLES9

What We Need

- The Serenity to accept the things we cannot change
- The Courage to change the things we can
- The wisdom to know the difference

What might help

- Avoid single points of failure
- Don't forget backups

Questions and Answers

Thank you for your attention and feedback!

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. 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. 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.

