ProLiant Clusters: SteelEye LifeKeeper for Linux, a Technical Overview
Session 2256

Dan Cox

Manager, Linux Clusters
Hewlett-Packard Company

Eddie Williams

Senior Software Engineer
SteelEye Technology, Inc

HP WORLD 2003
Solutions and Technology Conference & Expo

Agenda

- LifeKeeper Introduction
  - Who, What, Why
- A Closer Look at LifeKeeper for Linux
  - Concepts
  - Architecture
- Building a LifeKeeper for Linux Cluster

LifeKeeper for Linux

- **Features**

  - Full Certification across ProLiant
  - Support to 16/32 Nodes
  - Multiple configuration options

  - Full Storage Support
    - SCSI
    - Fibre
    - Enterprise

  - Graphical GUI

  - Full Function SDK

  - Range of Application Recovery Kits
    - ERP
    - DBMS
    - Web
    - Mail
    - File Print

- **Benefits**

  - Range of Server Choices
  - Cluster scalability
  - Cluster versatility

  - Storage Option Choices
    - Entry to Low End
    - Mid Range
    - Data Center Class

  - Ease of admin / control

  - Ease of cluster configuration

  - Support for major user needs
    - SAP
    - ORACLE, DB2, MySQL, Postgress
    - Apache
    - Sendmail, Lotus
   - Samba

LifeKeeper for Linux

•**Features**

•online reconfiguration for cluster nodes

•mix and match cluster nodes

•multiple failover conditions
- • IP@
- •node failure
- • network failure
- • application/service failure

•multiple failover configurations
- • mutual takeover
- • cascading
- • n+1

•heartbeat over ethernet (multi link support)

•rolling upgrade infrastructure

•**Benefits**

•ease of administration and scale out

•ease of scalable growth

•comprehensive failover detection
- • client access
- • server or storage
- • network failure
- • application failure

•sophisticated configurations
- • servers back each other up
- • application failover thru the cluster
- • one server backing up many

•multiple heartbeats for continuous availability

•ease of administration

LifeKeeper for Linux on ProLiant: Supported Configurations

Replicated Storage Cluster
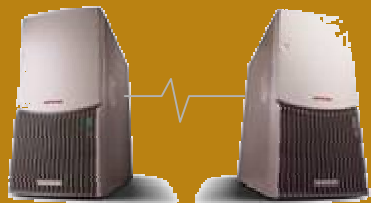
Servers:
- Entry level ProLiant DL or ML

Storage:
- Internal server drives

Max nodes: 2

Typical Applications:
- Distributed mission critical

Departmental Cluster including DL380 Packaged Cluster

Servers:
- Selected ProLiant DL or ML

Storage:
- Smart Array Cluster Storage

Max nodes: 2

Typical Applications:
- Distributed business critical

Infrastructure Cluster

Servers:
- Selected ProLiant BL, DL, or ML

Storage:
- MSA1000 & RA4100

Max nodes: 16

Typical Applications:
- Mid-range business critical

Multipath: Secure Path

Enterprise Cluster

Servers:
- Selected ProLiant BL DL, or ML

Storage:
- EVA3000, EVA5000 & MA8000

Max nodes: 16

Typical Applications:
- Consolidated mission critical

Multipath: Secure Path

Why HP & SteelEye

- SteelEye and HP product teams work together to develop cluster solutions with LifeKeeper, ProLiant servers, and HP storage arrays.

- SteelEye performs all product certification testing of the ProLiant servers and StorageWorks Storage.

- LifeKeeper's use of SCSI reservations for I/O fencing is an ideal complement to HP DL380/DL580 servers.

- SteelEye's LifeKeeper and HP's Secure Path deliver "No single point of failure" solution for MSA1000 and EVA clusters.

- SteelEye works with HP solution teams to develop solutions with Oracle, SAP, Sendmail, and Lotus Notes.

- SteelEye has worldwide marketing and sales that are working with resellers, customers, and hp account teams.

New Features in LifeKeeper v4.3

- Enhancements targeted at large resource configurations

- Easy install CCISS recovery module

- Flexibility in storage configurations
  - Loosen up Secure Path restrictions
    - Single to dual upgrades
    - Persistent device node support

- Certification with latest RH and UL kernels

- SAP support with DB2 UDB/SLES, Oracle 8/RHEL AS 2.1

- Certification with EVA, Hitachi 9xxx storage units

HP LifeKeeper Offerings

- **ProLiant Cluster HA/L100 – LifeKeeper for Linux Kit**
  - Includes HP generated installation documentation and a production level copy of LifeKeeper for Linux for use with 30-day free evaluation licenses or permanent licenses which can be purchased separately
  - HA/L100 kit pn is 303523-B22
  - Check product page for ProLiant and storage certification

- **ProLiant LifeKeeper for Linux Cluster Solutions in NA**
  - Solutions for Oracle (pn 331514-001), SAP (pn 331517-001), and Sendmail (pn 331518-001)
  - Solutions include DL380 G3 Packaged Cluster, Redundant Smart Array Cluster Storage Controller, HA/L100 LifeKeeper Kit, and the appropriate Application Recovery Kit.

ProLiant LifeKeeper for Linux Cluster Solution for Oracle

- Provides protection for Oracle 8i and 9i database services

- Operates with either Standard Edition or Enterprise Edition

- Operates in a RAC environment
  - RAC protects DBMS, LifeKeeper protects application and other system services

- Can operate in either Active/Stand-by or Active/Active configurations

- Can be used as component of integrated solution to provide protection for all tiers of application architecture
  - mySAP as an example

ProLiant LifeKeeper for Linux Cluster Solution for SAP

- Performs monitoring and switchover for the mySAP Central Instance services including Enqueue and Messaging

- Has the ability to shutdown a mySAP Application Server that is running on the backup server when the Central Instance fails over, and restart the Application Server on the backup server once the Central Instance is in-service anywhere else in the cluster.

- Works seamlessly with other LifeKeeper recovery modules to provide a total HA solution for mySAP environments

  - Database Recovery Kit

  - NFS Recovery Kit

  - IP Recovery Kit

# ProLiant LifeKeeper for Linux Cluster Solution for Lotus

- HP ProLiant High Availability Solution for Lotus Domino using SteelEye LifeKeeper for Linux
  - SteelEye LifeKeeper clustering solution, industry-standard HP ProLiant ML350 or ML370 servers, HP StorageWorks Modular SAN Array 1000 storage solutions and rapid provisioning for Linux with RDP
  - Whitepaper describing how to enhance the availability of Lotus Domino enterprise servers using LifeKeeper with ProLiant servers and the MSA1000 - http://h71019.www7.hp.com/aa_downloads/6/100/225/1/72305.pdf
  - HP's pre-configured and tested solution
- LifeKeeper products can be purchased from HP via HP Direct in the U.S.
- Reselling LifeKeeper products in EMEA is being investigated.

LifeKeeper Summary

- **Coverage across entire ProLiant Brand:** Support for SMB to enterprise ML, DL, and BL series where appropriate. Coverage for two node SCSI with SACS, and multi node FC with MSA1000, and EVA.

- **Continuing to grow with ProLiant servers and storage and OS support:** LifeKeeper is evolving with the current HP Linux strategy and positioning and is growing as a key partner with HP

- **Continuing to develop Application Recovery Kits to met Enterprise Application needs:** What solution offerings are currently available for business-critical applications and where LifeKeeper is going next

A Closer Look at LifeKeeper for Linux
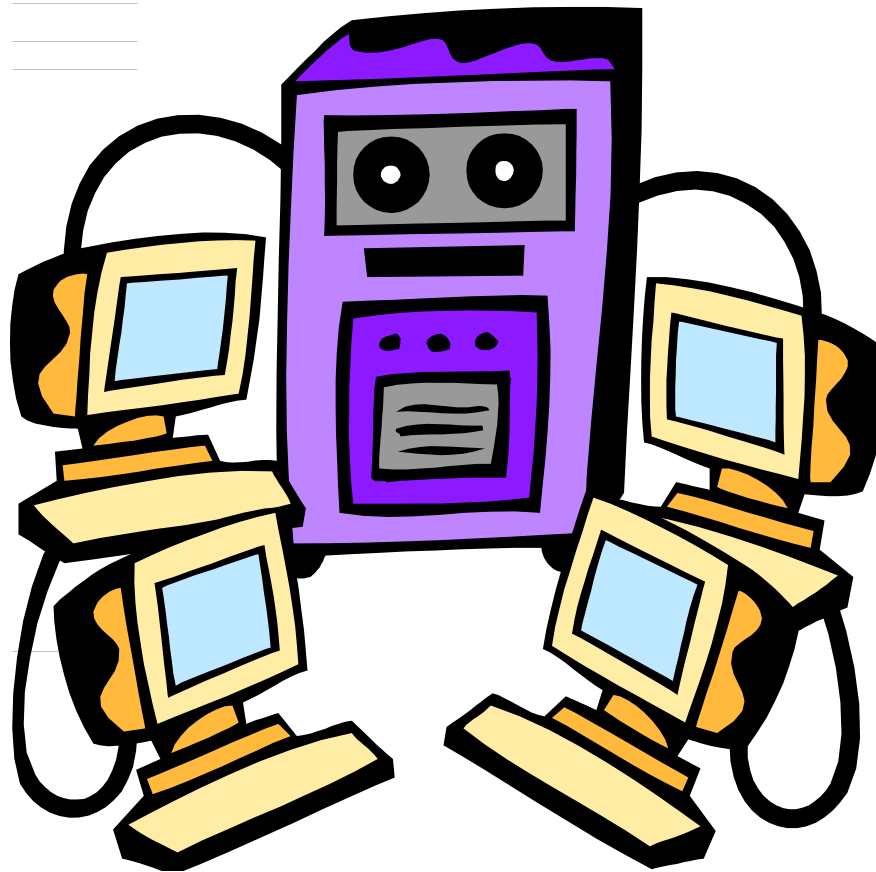
LifeKeeper Concepts: Resources

- Each cluster item that LifeKeeper protects is called a resource:
  - System asset to be protected by LifeKeeper

- Resources are grouped into a resource hierarchy:
  - Set of dependent resources with parent/child relationships
  - Example – an application may need both a file system and an IP interface available for it to be usable

- LifeKeeper will monitor the health of each individual resource and will take appropriate actions to keep them available.
  - This action may be recovering them locally on the same server or moving them to a backup server.
  - LifeKeeper ensures that all resources in a hierarchy are moved together.

LifeKeeper Concepts: Types of Recovery

- ## Failover
  - Process of restoring a resource when the original server or resource has failed. Normally, a failover is unplanned and occurs when failure is detected by a backup system.

- ## Switchover
  - Process of shutting a resource or server down in an orderly manner and then recovering the resource or server on a backup system. This is normally performed in maintenance or test mode and is planned.

LifeKeeper Concepts: Heartbeats and Comm Paths

- Heartbeat
  - Periodic message between servers used to monitor health of cluster members
  - Interval between heartbeats and the number of heartbeats that must be missed before a system is considered "dead" is configurable
  - Heartbeats are sent across defined Communication Paths
- Communication Paths
  - Two are required to avoid SPOF

  (Single Point of Failure)
  - Communication path types:

    TCP - LAN connection (multiple allowed)

       up to six supported

    TTY - serial port connection (one allowed)
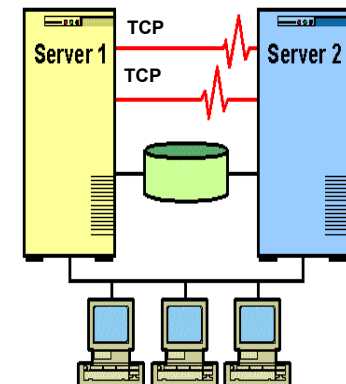
LifeKeeper Clusters: Failover Configurations

- Because LifeKeeper supports more than 2 nodes in a cluster, many failover configurations are possible.

- Supported failover configurations are:
    - Multi-Directional Failover
    - Cascading Failover
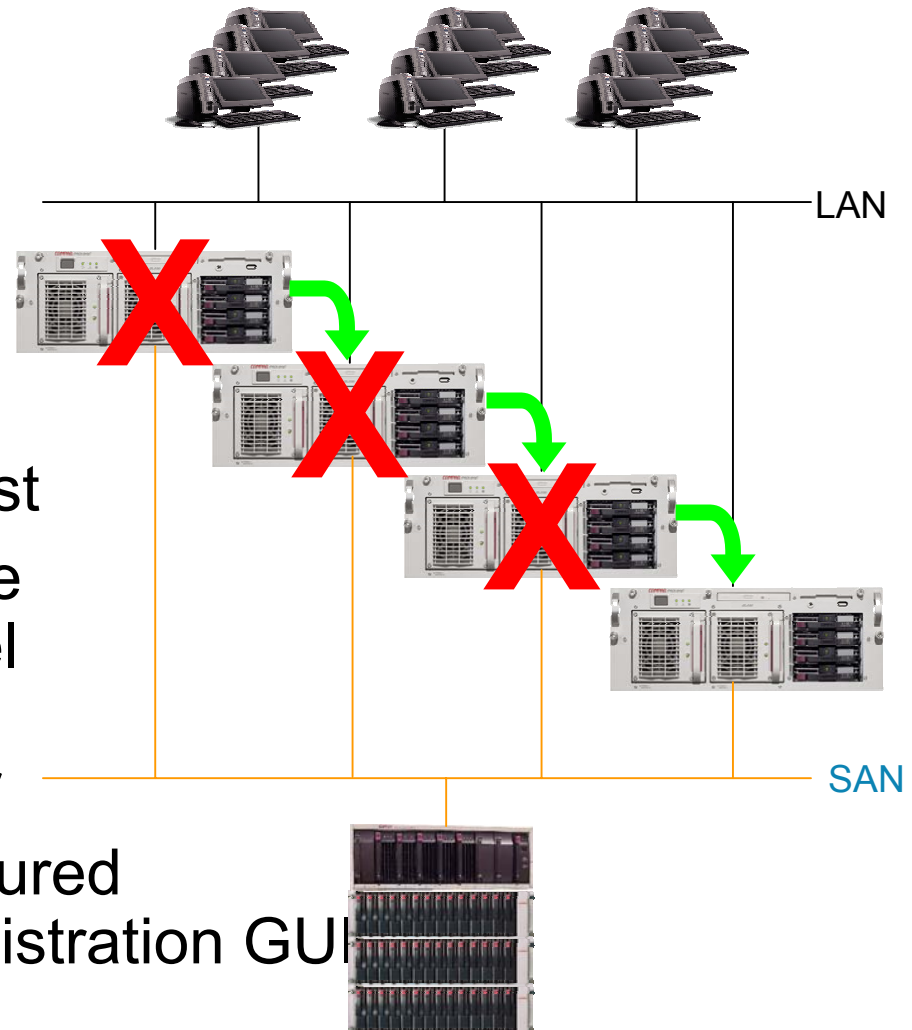   - N-to-1 Failover

LifeKeeper Concepts: Types of Failover

- In a multi-directional failover configuration:

  - Resources running on a certain server will move to separate servers on failover

  - The server to which each resource moves is configured thru the LifeKeeper Administration GUI.

  - Also called 1-to-Many

LAN

SAN
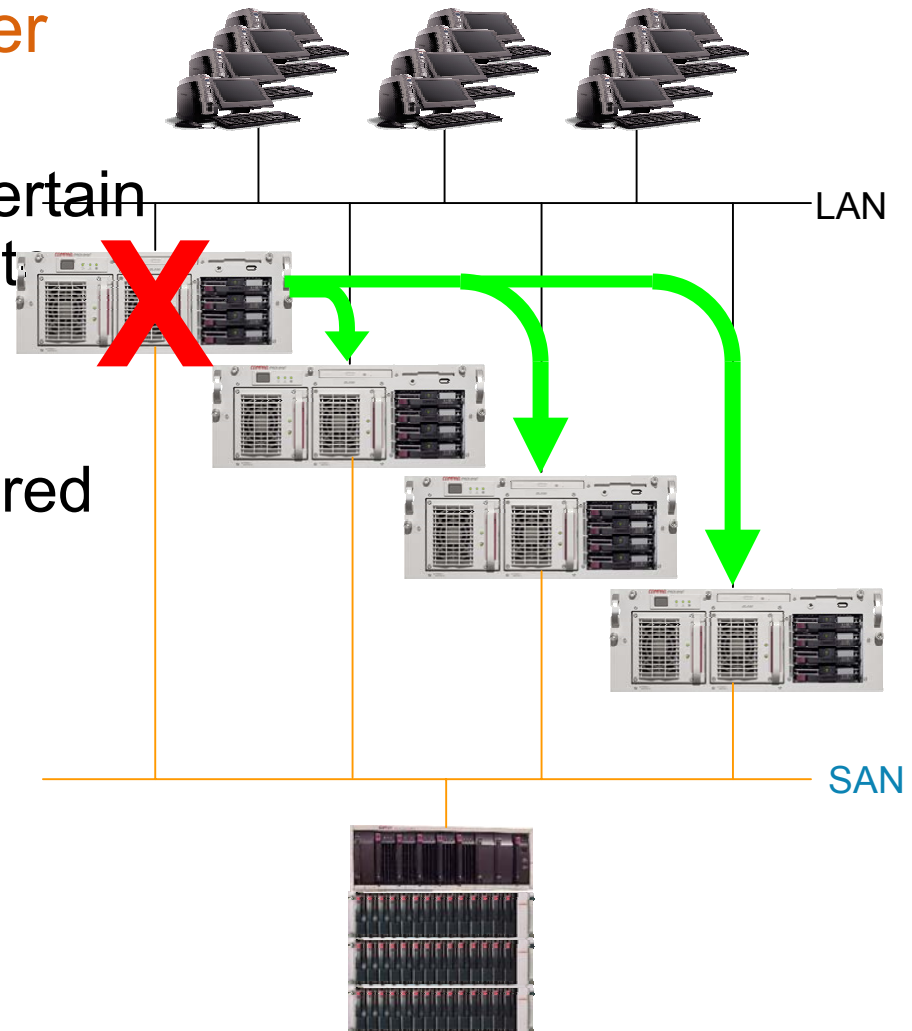
LifeKeeper Concepts: Types of Failover

- In a cascading failover configuration:
  - Multiple backup servers can be defined
  - Each backup is given a priority which indicates where it is in the failover list
  - Priorities are defined at the resource, not system, level so resources can cascade independent of each other
 - These priorities are configured thru the LifeKeeper Administration GUI

LAN

SAN

LifeKeeper Concepts: Types of Failover

- In an N+1 failover configuration:
  - A single server can be acting as backup for multiple (N) others
  - The backup server must be able to access the data being used by all of the servers that it protects.
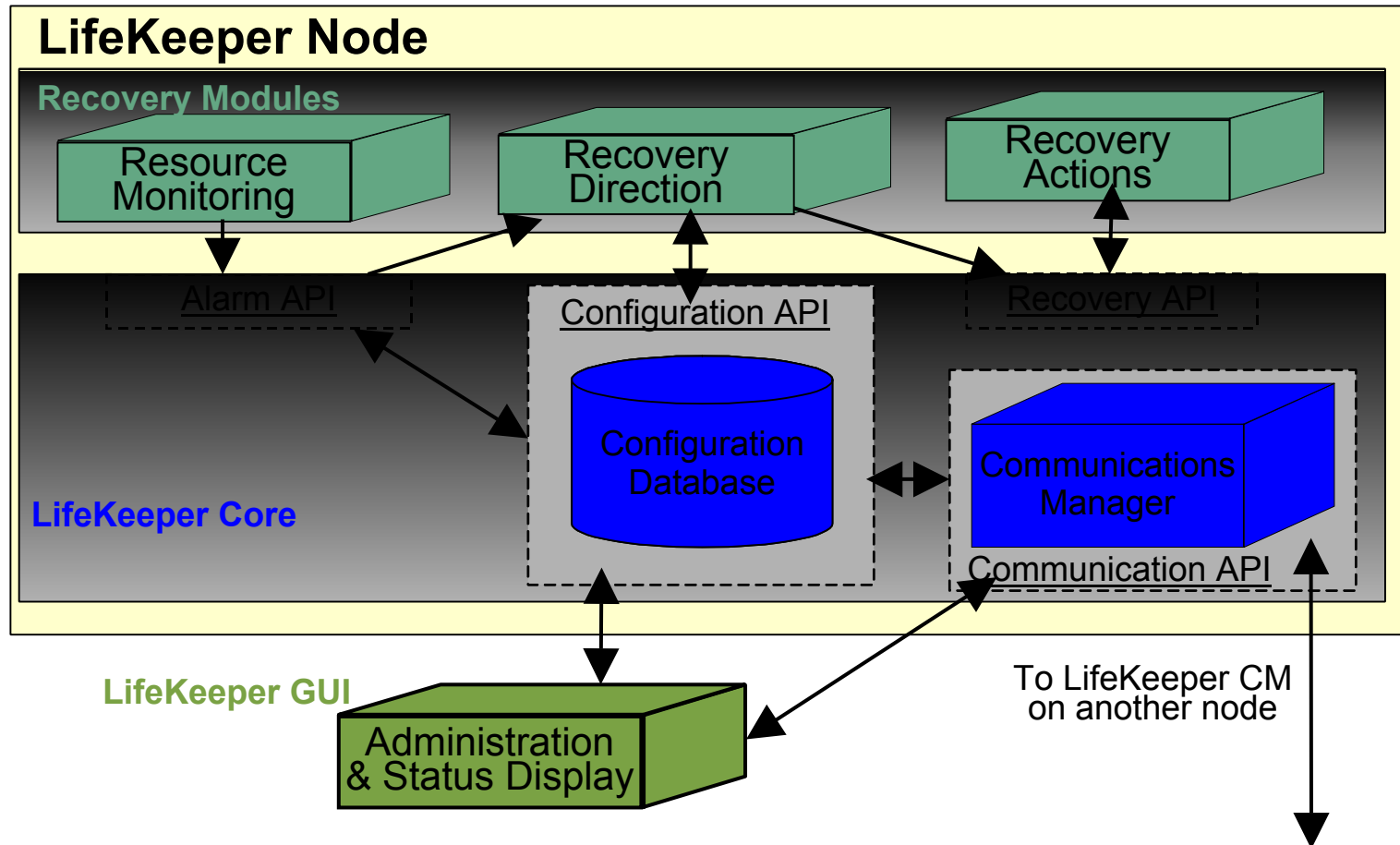 - Also called Many-to-1
- LK offers alternatives:
- ✓ Multi-directional
- ✓ Cascading
- ✓ N+1

LAN

SAN
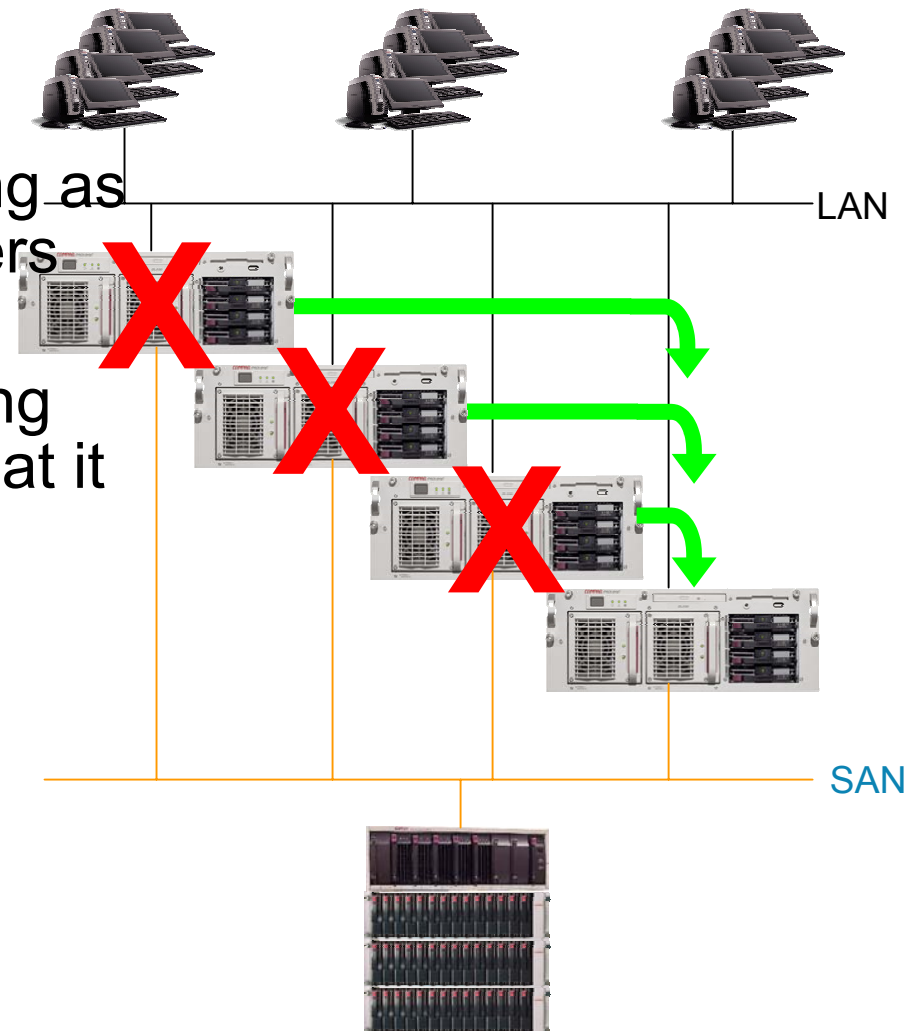
LifeKeeper Architecture: Major Components

- LifeKeeper is composed of 3 major components:
  - LifeKeeper Core
    - LK Configuration Database (LCD)
    - LK Communication Manager (LCM)
    - lkcheck Daemon
    - Alarm and Recovery APIs
    - SNMP Trap Generator

  - LifeKeeper GUI

  - Optional Recovery Module(s)

LifeKeeper Architecture:
Major Components

LifeKeeper Node

Recovery Modules

Resource Monitoring

Recovery Direction

Recovery Actions

Alarm API

Configuration API

Recovery API

Configuration Database

Communications Manager

LifeKeeper Core

Communication API

LifeKeeper GUI

Administration & Status Display

To LifeKeeper CM on another node

LifeKeeper Architecture: Graphical User Interface

- LifeKeeper GUI is used to
  - Define cluster membership
  - Create/delete communication paths
  - Define resource hierarchy relationships
  - "Extend" hierarchies to other cluster members
  - Put resources "in-service" and "out-of-service"
  - Monitor status of cluster communication paths
  - View status of cluster members
  - View status of individual resources
  - Read log files

- GUI is written in JAVA as client/server program

LifeKeeper Architecture: Graphical User Interface

- **LifeKeeper GUI provides a visual indication of:**
  - Server Status:

Alive   /   Dead

Alive, but:

- >= 1 comm paths from this node are marked as DEAD, or
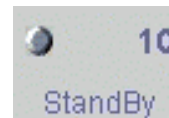- no redundant comm path from to another node

  - Resource Status:

In-Service   In-Service   Out-of-Service   Out-of-Service   Unknown
Protected   Unprotected Unimpaired   Failed   State
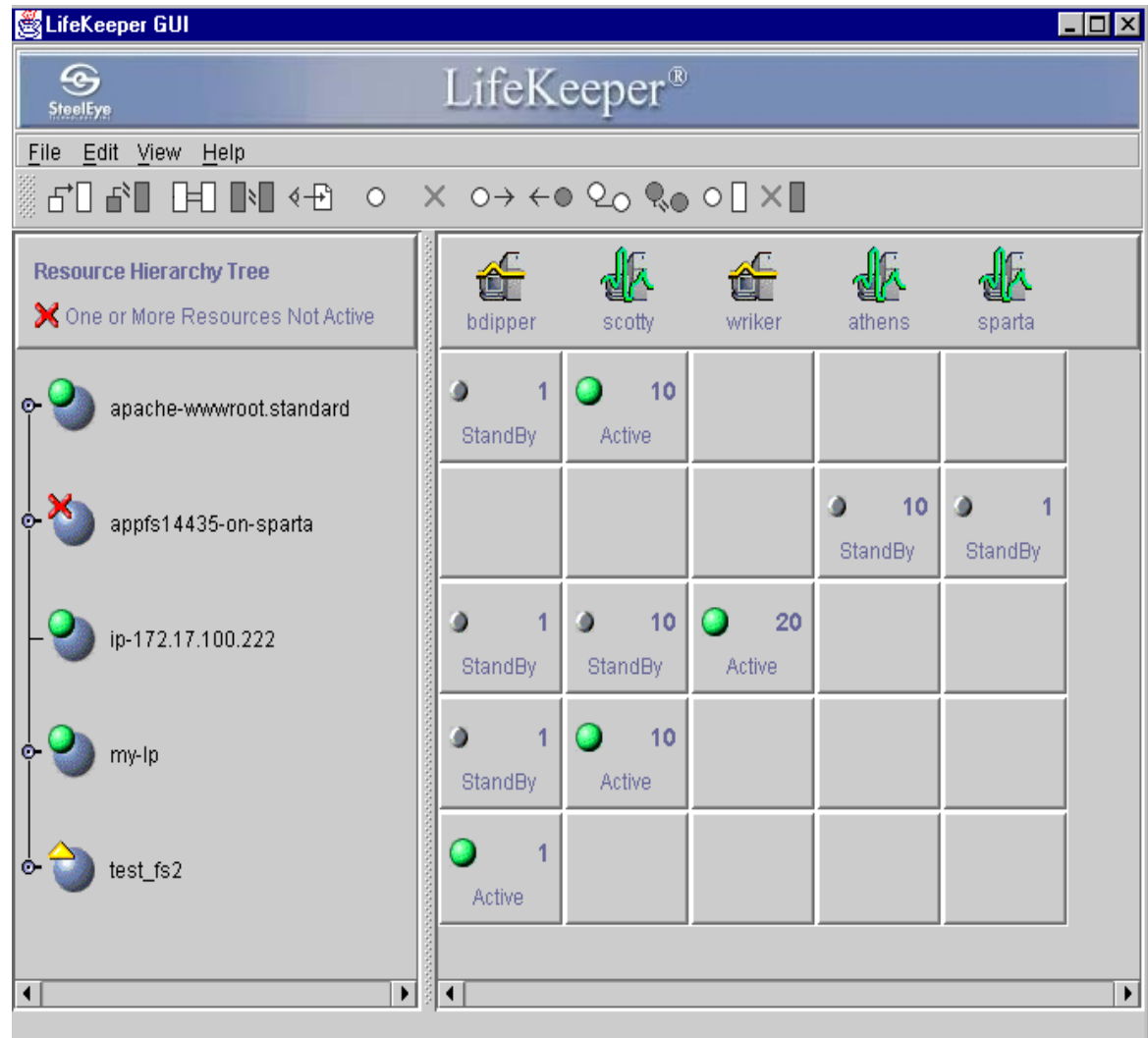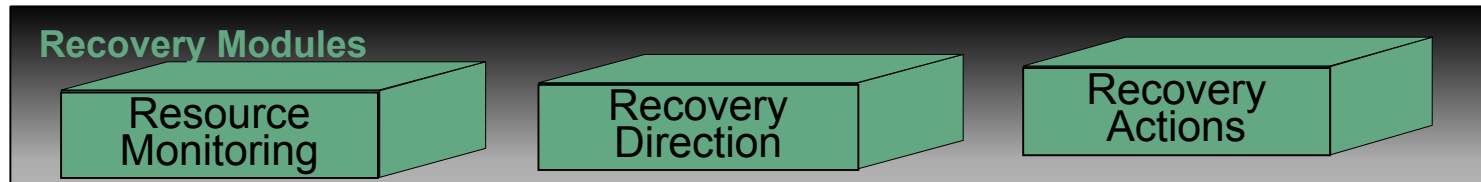
LifeKeeper Architecture: Graphical User Interface

What does this display show us?

LifeKeeper Architecture: Graphical User Interface

LifeKeeper Architecture: Recovery Modules

Recovery Modules

Resource Monitoring | Recovery Direction | Recovery Actions

- Each module must have 4 components:
  - Instructions to monitor health of the resource being protected
  - Policies on what to do if a problem is detected
  - Scripts to perform desired actions
  - JAVA property files for integration into GUI
- Modules are written in a script language (perl, bash, …) and use command line utilities to communicate with the LifeKeeper core.
- A Software Developer's Kit is provided so that RMs for custom applications can be easily built and plugged into LifeKeeper.

LifeKeeper Packages

LifeKeeper Core Package Cluster

- LifeKeeper Core Software
- LifeKeeper GUI Software
- LifeKeeper Man Pages
- LifeKeeper IP Recovery Module
- Raw I/O Recovery Module
- LifeKeeper Online Help

Recovery Modules

- Apache Web Server Recovery Module
- Informix Recovery Module
- MySQL Recovery Module
- Oracle RDBMS Recovery Module
- DB2 Recovery Module
- NFS Server Recovery Module
- Samba Server Recovery Module
- SAPDB Recovery Module

- mySAP Recovery Module
- Print Services Recovery Module
- Sendmail Recovery Module
- SAMS Recovery Module
- NAS Recovery Module
- Software Developer's Kit
- LifeKeeper for Linux Data Replication

Building a ProLiant Linux HA Cluster
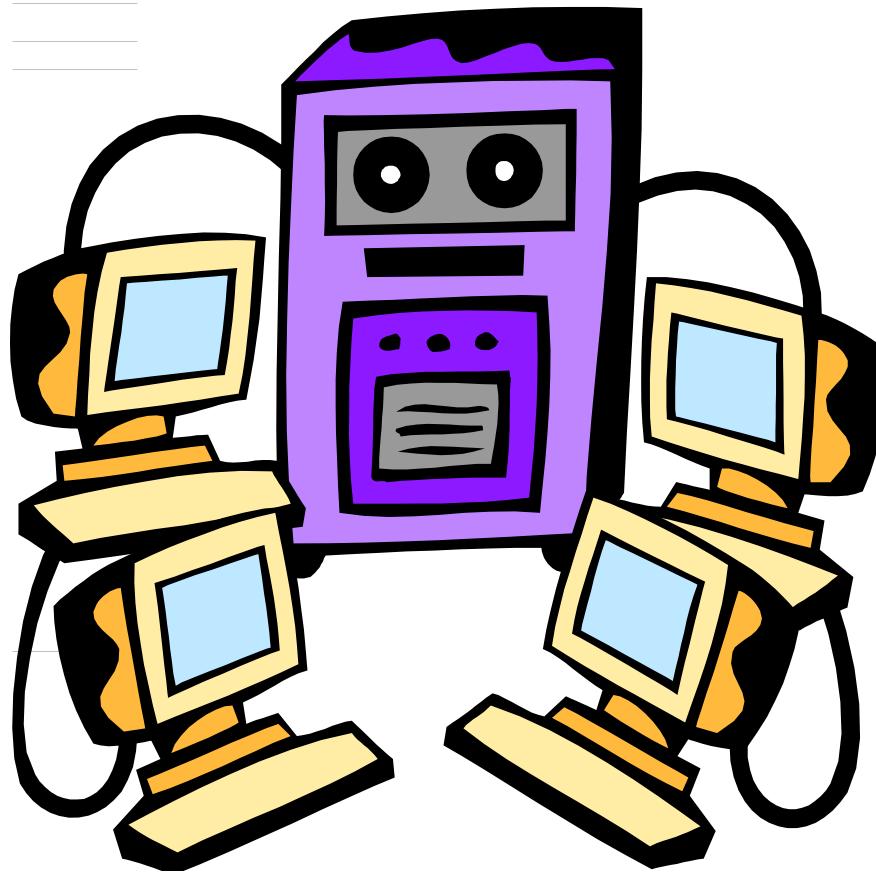
Configuration Planning

- Cluster servers
  - All servers must be running same operating system
  - Servers may have different hardware configurations
- Communication paths
  - Minimum of two communication paths to avoid single point of failure
- Data resources
  - Shared or replicated ?
- Shared communication resources
  - IP addresses
- Application resources
  - Off-the-shelf kits or need to develop kit for custom application?

Operational Considerations

- System grouping arrangement?
  - Active/Active
  - Active/Standby
- Failover configuration?
  - Multi-directional
  - Cascading
  - N+1
- Switchback configuration?
  - Automatic switchback
  - Intelligent (manual) switchback
- Shutdown strategy?
  - Switchover resources on manual shutdown?

Networking Considerations

- Ethernet host bus adapters
- IP addresses for each system and one for Virtual IP address
- Two network interfaces are recommended
  - Private LAN for LifeKeeper communication
  - Public LAN
  - Public LAN can also be a secondary communication path
- Network addresses must be resolvable
  - Recommend /etc/hosts entries for all LifeKeeper servers
  - DNS is fine for systems outside the cluster
- Verify the network is functional before starting the communication path configuration

Gather the Components – Hardware

- DL or ML Servers
  - DL320, DL360, DL380, DL560, DL580
  - DL380 G3/G2 Packaged Cluster
  - ML330, ML370, ML530
  - BL20P and 40P – this fall
- SCSI Storage
  - Smart Array Cluster Storage w/ HP 64-bit PCI Dual Channel Adapter Smart 532 or internal Smart 5i
- Fibre Storage
  - StorageWorks MSA1000, EVA3000, EVA5000, MA8000, w/ HP 64-bit/66Mhz Host Adapter FCA2214, Part Number 281541-B21
  - StorageWorks RA4100 w/HP 64-bit Mhz Host Adapter Kit 120186-B21
  - Switches

Gather the Components - Software

Operating System

redhat   SuSE

Application Software

SENDMAIL®   ORACLE®   PostgreSQL

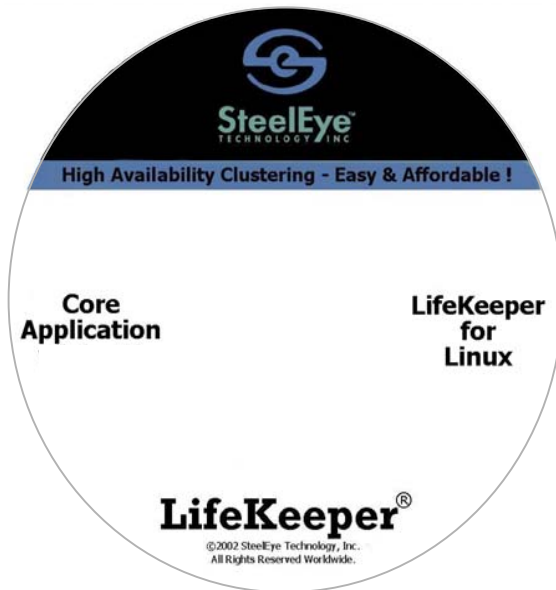Powered by APACHE   SAP   samba

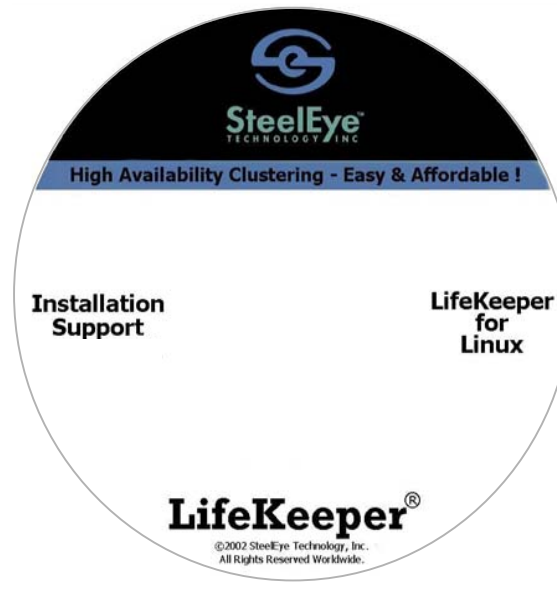Custom App

Gather the Components - Software

- Multipath Software for MSA1000, EVA3000, and EVA5000 configurations – Secure Path for Linux
  - Currently supported in 2 node configurations only
  - Limited OS support for Secure Path. Broader support coming in 4Q03.

Gather the Components - Software

LifeKeeper

Core

Installation Support

Recovery Modules

Cluster Installation

- The following steps are required before installing LifeKeeper:

1) Install Linux and associated packages

2) Connect servers and shared storage

3) Configure shared storage

4) Verify network configuration

5) Install and set up database application (if applicable)

6) Set up LifeKeeper environment using LifeKeeper Installation Support CD

Note: Consult the LifeKeeper Planning and Installation Guide and LifeKeeper Release Notes for additional installation details.

Cluster Installation

- **After the initial cluster setup:**

1) Install LifeKeeper and required Recovery Modules on all cluster systems

2) Start LifeKeeper on all cluster systems

3) Start LifeKeeper GUI

4) Create redundant communication paths between each unique pair of cluster systems

5) Create application hierarchy on primary system

6) Extend application hierarchy to one or more backup systems

7) Test manual switchover of application

8) Test failover of application

Custom Recovery Module Development

- **Two options for supporting a new or custom application:**
  - Generic Recovery Module
    - Application treated as a "generic" resource type
    - You provide scripts for starting, stopping, monitoring, and local recovery of the application
    - Template scripts provided
  - Software Developer's Kit
    - Used to build a complete recovery kit for the application
    - Start, stop, monitoring, and local recovery scripts
    - Scripts and properties files for complete GUI integration
    - Template scripts and script generation tool provided

Thanks for attending the session