



Mission Critical Open Source Software Implementation in the Enterprise

*Mr. David Chow
ShaoLin Microsystems Ltd.*

Personal Background

- Over 10 years experience in the field
- System integration (system engineer,pre/post sales)
- Education & Academic (supervisor for undergraduates and postgraduates thesis)
- Academic background in system software (OS & distributed systems), real-time systems (Microcontrollers)
- 6 years full time in system architect in system software research and development
- Expertise: File systems, SAN storage drivers, thin client platforms, replication, distributed clusters

The Leading Enterprise Linux Utility Computing
From Desktop to Data Center



ShaoLin Microsystems Overview

The Leading Enterprise Linux Utility Computing
From Desktop to Data Center



Company Background

- Hong Kong software company founded in year 2000
- Developed packaged enterprise utility system software for Linux
- Focus on storage, infrastructure and system management software
- Products ranges from desktop to data centre
- Technology area: distributed computing, file system, data storage, networking, clustering

The Leading Enterprise Linux Utility Computing
From Desktop to Data Center



Market Positioning

- Targets public sector and enterprise who move to Linux (from desktop to data centre)
- Provides enterprise infrastructure management utilities to leverage OSS advantages (desktop to datacenter)
- Leading Linux support and compatibility (support 30+ Linux distributions, 1000+ kernel releases, 10+ CPU types and multiple hardware platforms)
- Packaged software with one click installation, minimal set-up works, and no kernel patches or manual intervention needed

The Leading Enterprise Linux Utility Computing
From Desktop to Data Center



Product Lines

ShaoLin InfiniCluster



Cluster
Management
Software

ShaoLin Aptus



Desktop
Management
Software

ShaoLin Volume Replicator



Real time
Data Replication
Software

ShaoLin CogoFS



Storage
Management
Software

The Leading Enterprise Linux Utility Computing
From Desktop to Data Center



Agenda

- Looking into Open Source Software (OSS)
- Challenges
- Understanding HA & Clustering Technologies
- Virtualization
- Where are we?
- Conclusion

The Leading Enterprise Linux Utility Computing
From Desktop to Data Center



Looking Into OSS

The Leading Enterprise Linux Utility Computing
From Desktop to Data Center



Is OSS free?

- In terms of licensing (yes)
- Cost = 0 when skill and time cost nothing
- Freedom (Yes)

The Leading Enterprise Linux Utility Computing
From Desktop to Data Center



What is OSS good for Enterprise?

- Direct cost saving in licensing
- Lower maintenance cost
- Infrastructure (OS, storage, web platform, DB, J2EE)
- High security and reliability (e.g. Apache, Linux)
- User control of product life cycle
- Strategic benefits on removing vendor lock-in

The Leading Enterprise Linux Utility Computing
From Desktop to Data Center



What's good about Linux in the Enterprise?

- Platform consolidation - Cross platform compatibility, supports all types of commercial platforms (PC, Itanium, PPC, AMD, Alpha, Sparc, Arm... and many)
- Higher performance, cheaper hardware - Workload consolidation (fewer servers)
- Strong in next generation(parallel) applications

The Leading Enterprise Linux Utility Computing
From Desktop to Data Center



Challenges

The Leading Enterprise Linux Utility Computing
From Desktop to Data Center



Challenges

- ✦ **Cost**
 - Additional software and hardware
 - Additional tools and training
- ✦ **Complexity**
 - Management and provisioning
 - Application-specific business continuity needs
- ✦ **Reliability**
 - Complex solutions are hard to test
 - Requires specialized training for personnel

The Leading Enterprise Linux Utility Computing
From Desktop to Data Center



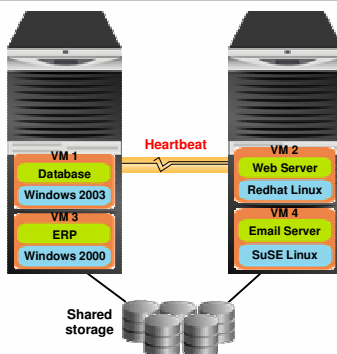
The Real & Compelling Challenges

- ✦ “On-demand” scalability of IT landscapes
- ✦ Consolidation (Scale-up) vs Expansion (Scale-out): Which is right?
- ✦ Mixed computing environment (multiple OS, multiple hardware platform)
- ✦ Evolution of new technologies – Grid computing, parallel applications...

The Leading Enterprise Linux Utility Computing
From Desktop to Data Center



Consolidation (Scale-Up) Problem



- ✦ Consolidate multi-OS and multi-apps with virtualization
- ✦ Max. 50% utilization for reserving capacity to failover all applications

The Leading Enterprise Linux Utility Computing
From Desktop to Data Center



Leverage Scale-Up

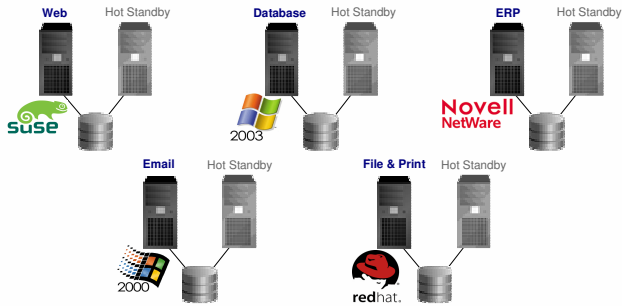
- ✦ Consolidate to use big box instead of many small boxes
- ✦ Difficult to achieve both good utilization with redundancy
- ✦ Difficult to cope with on-demand concepts
- ✦ Get more savings with load management technologies to consolidate, and balance out utilization and availability

The Leading Enterprise Linux Utility Computing
From Desktop to Data Center



Expansion (Scale-Out) Problem

Typical scale-out “HA” infrastructure – max. 50% utilization



The Leading Enterprise Linux Utility Computing
From Desktop to Data Center



Leverage Scale-Out

- ✦ Architect with multi-tier applications and systems
- ✦ Maintain scalability, cope with on-demand concepts
- ✦ Complexity and cross platform issues as number of server grows
- ✦ Savings with cluster management technologies to achieve high availability and simplified single view cluster administration

The Leading Enterprise Linux Utility Computing
From Desktop to Data Center



Parallel Application Management

- ✦ Unable to manage using traditional failover or high availability software
- ✦ “HA” is not a mandatory feature of parallel app.
- ✦ Applications has built-in HA will still need cluster management software to failover system resources (i.e. Disk volumes, IP addresses)
- ✦ Often isolate from traditional infrastructures

The Leading Enterprise Linux Utility Computing
From Desktop to Data Center

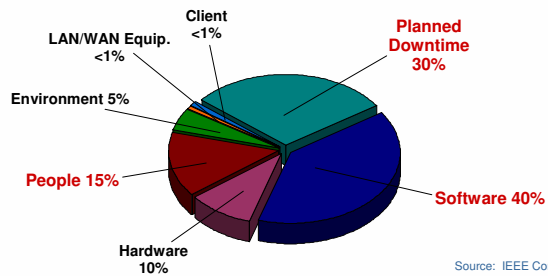


Understanding Challenges of High Availability Implementation

The Leading Enterprise Linux Utility Computing
From Desktop to Data Center



Causes of Downtime



Application management is the kill!

The Leading Enterprise Linux Utility Computing
From Desktop to Data Center



Motivation of Clustering

→ Things breaks

- Applications have un-reproducible bugs
- Servers crash, network connections fail
- Storage devices become inaccessible
- Operators and administrators make errors

→ Things changes

- Application needs outgrow server capacity
- Ad hoc applications arise

→ Things are expensive to manage

- Configuration control (versions and patches)
- Many applications = many servers = many administrators

The Leading Enterprise Linux Utility Computing
From Desktop to Data Center



High Availability Technologies

- Fault Tolerance Machines
 - Hardware centric
 - Redundant components (Dual CPUs, Storage, Memory)
 - Transparent to software
- Clustering
 - Software centric
 - Performs failover by cluster management software
 - Provides application monitoring and control
 - Improves online manageability

The Leading Enterprise Linux Utility Computing
From Desktop to Data Center



Goal of Clustering

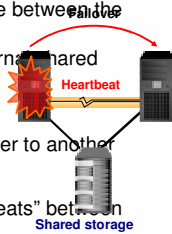
- Elimination of Single-Point-of-Failure (SPOF) by redundancy
- Provides software and application level monitoring and recovery
- Use standard hardware components and peripherals
- Maintain transparency, applications often do not have to be modified being implemented
- Improves online manageability for reduction of planned downtime
- Consolidation of services on a high availability infrastructure

The Leading Enterprise Linux Utility Computing
From Desktop to Data Center



How Failover Works?

- The simplest setup: 2 servers and external storage
 - Each server runs its own applications
 - Management utility is used to share the storage between the servers
 - All high availability data are maintained in external shared storage
- Failover of a service
 - Migrate failed service (resource) from one server to another
- Peer fault detection
 - Daemons are typically used to monitor "heartbeats" between servers

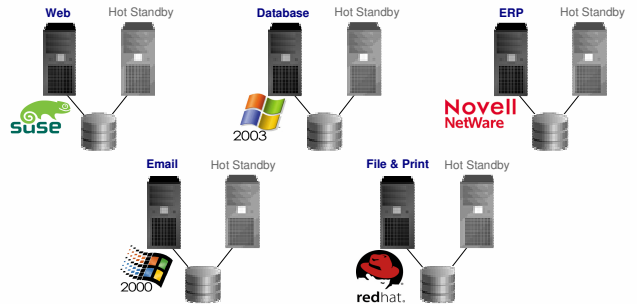


The Leading Enterprise Linux Utility Computing
From Desktop to Data Center



2-node Style HA Infrastructure

"HA" infrastructure – max. 50% utilization

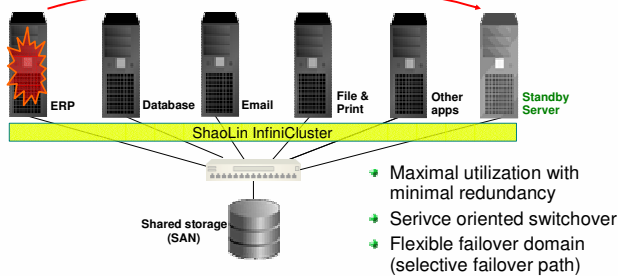


The Leading Enterprise Linux Utility Computing
From Desktop to Data Center



Multi-node N+1 Clusters

Selective service oriented switch/fail over



- Maximal utilization with minimal redundancy
- Service oriented switchover
- Flexible failover domain (selective failover path)

The Leading Enterprise Linux Utility Computing
From Desktop to Data Center



Advantages of Software Clustering

- System & hardware failure detection, avoid SPOF
- Application monitoring and recovery
- Support multiple types of applications
- Automate administration and online manageability
- Single view of management for multiple servers and services
- Low cost with no hardware investment

The Leading Enterprise Linux Utility Computing
From Desktop to Data Center



Cluster Manage Software Design

- Reliable heartbeats and accurate fault detection
- False failure can result in split-brain (serious corruption - disaster)
- Synchronisation logics - Simple in 2-node clusters, complex in multi-node clusters, and very complex with parallel applications
- Rely on network heartbeat messages for node keep-alive and sync (e.g Ethernet, IP, RS-232)
- Complex logic to keep service states in-sync ASAP
- Cluster management software is a software. Software is always vulnerable (how to get around?)
- On-demand scalability, cross platform operability

The Leading Enterprise Linux Utility Computing
From Desktop to Data Center



Automation and Manageability

- Automatic failover without human intervention
- Provides policy-based, cascading, multi-directional failover paths
- Only recovers failed applications without full server power-down
- Notification of exceptions and events through standard protocols (SMTP & SNMP)

The Leading Enterprise Linux Utility Computing
From Desktop to Data Center



Split-Brain

- Caused by race conditions due to failure of the cluster management software
- Reliable network protocols can have delays (e.g. TCP/IP)
- Heartbeat daemon breaks – False failure detection from peer
- Heartbeat process can be affected by run-away applications or exceptions (Out of memory)

The Leading Enterprise Linux Utility Computing
From Desktop to Data Center



Userland Heartbeats

- Redundant heartbeat paths – Can get away SPOF
- Userland heartbeat daemons are NOT reliable
 - Affected by run-away apps
 - Vulnerables to exceptions
 - Rely on OS built-in network stacks (i.e. TCP/IP)

The Leading Enterprise Linux Utility Computing
From Desktop to Data Center



Other methods to avoid Split-Brain

- Build shared quorum disks – Shared disk with continuous timestamp updates of all nodes
 - Generate excessive I/O traffic, degraded I/O performance
 - Quorum disks can corrupt or fail (SPOF)
- External power switch to physically power-off other nodes
 - The power supply unit can fail (SPOF)
 - Only good for 2 node clusters, difficult to implement on multi-node clusters
- Quorum election – Not allow less than 50% nodes to run
 - Impossible to cause split brain on multi-node cluster unless more than half of the nodes dead

The Leading Enterprise Linux Utility Computing
From Desktop to Data Center



Internode Messages (INM)

- The core cluster management facility relies on internode messages
- Synchronizes service and resource states, correct failover decisions
- Performance can affect failover time = downtime
- Must use reliable transmission protocol
- Must be 2-way acknowledged to ensure peer synchronization
- If failed, can cause split-brain

The Leading Enterprise Linux Utility Computing
From Desktop to Data Center



INM Mechanisms

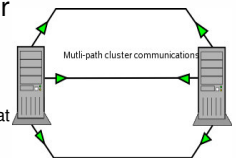
- Redundant communication paths
- Performance vs Reliability vs Redundancy
- Broadcast + Collective Acknowledgments
 - Limited scalability due to dynamic creation of peer lists barrier
 - All states must sync to all nodes in the cluster
- Unicast with reliable with end points:
 - Unlimited scalability
 - Selective synchronization, can optimize performance for multiple cluster sub-domains

The Leading Enterprise Linux Utility Computing
From Desktop to Data Center



Cluster Infrastructure

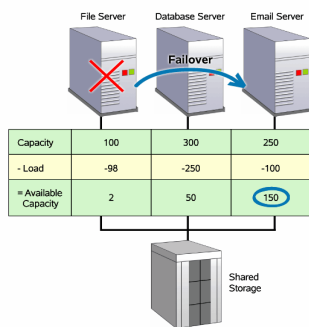
- Split heartbeat and communication out of the cluster management engine
- Kernel heartbeat mechanism
 - heartbeat process lives with the kernel (heartbeat stops = kernel stops = system hangs)
- Kernel level cluster membership service
- Implements kernel-to-kernel multi-path internode communications in layer 2 to 4



The Leading Enterprise Linux Utility Computing
From Desktop to Data Center



Workload Management



- Load balanced across cluster nodes during failover
- Proactive failover on predetermined policy
- Threshold-based system load and health monitoring
- Suitable for service consolidation and QoS monitoring

The Leading Enterprise Linux Utility Computing
From Desktop to Data Center



Application Monitoring & Recovery

- Implementation of agents or small piece of software for application monitoring and recovery
- Script-based agents
 - Easy to implement (quick & dirty)
 - Draw excessive CPU power, no good for system software and hurt performance
 - Less accuracy, unable to detect system level or low level faults

The Leading Enterprise Linux Utility Computing
From Desktop to Data Center



Online Manageability & Scalability

- Online manageability can as much as 90% of planned downtime
- Online cluster expansion – Add/Remove nodes to/from the cluster online
- Most of the broadcast-based cluster does not allow beyond 32-nodes
- Migrate to new hardware platform without destruct the system

The Leading Enterprise Linux Utility Computing
From Desktop to Data Center



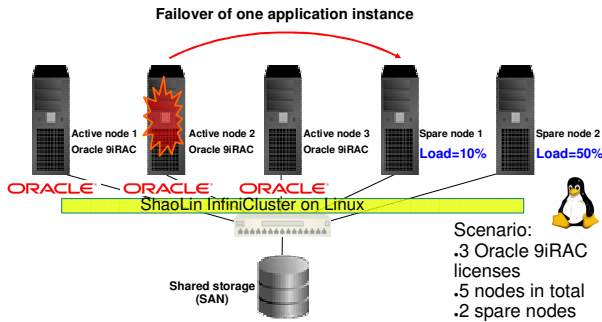
Parallel Applications Support

- Support multi-instance parallel applications, N-to-M failover configuration
- Switchover both application resources and system resources (i.e IP address, application restarts and recovery)
- Optimize performance for parallel applications
- Maintain and control min/max number of application instant over N cluster nodes

The Leading Enterprise Linux Utility Computing
From Desktop to Data Center



HA + Parallel Application Sample



The Leading Enterprise Linux Utility Computing
 From Desktop to Data Center



Virtualization

The Leading Enterprise Linux Utility Computing
 From Desktop to Data Center



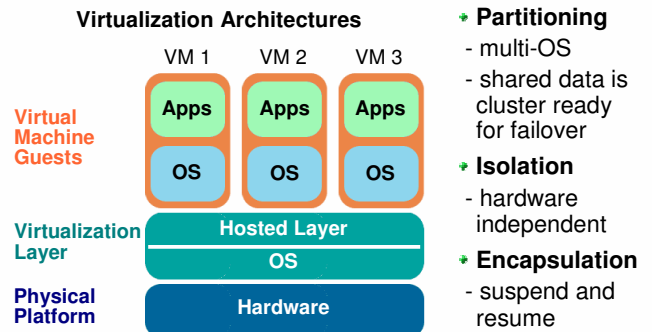
Virtualization Basics

- ✦ Breaks hardware dependencies, separating BIOS, OS and applications from physical hardware (Storage & Host CPU)
- ✦ Allows multiple OSes and applications (many virtual systems) on one physical machine
- ✦ Wraps OS and its applications into a transportable virtual machine
- ✦ Provides security outside the OS context
- ✦ Standardizes multiple generations of multi-vendor hardware

The Leading Enterprise Linux Utility Computing
 From Desktop to Data Center



Virtualization Features



The Leading Enterprise Linux Utility Computing
 From Desktop to Data Center



Virtualization Benefits

- ✦ **Transparency**
No modification and no changes in end-user environment
- ✦ **Software/Hardware Independence**
Dynamic, cross-system re-mapping of software to hardware resources
- ✦ **Online manageability**
On the fly virtual hardware infrastructure modifications without taking down services
- ✦ **Instant Provisioning**
Provision servers from templates instantly

The Leading Enterprise Linux Utility Computing
From Desktop to Data Center



Virtualization Opportunities

- ✦ **Server consolidation** - Consolidate servers into virtual machines hosted on fewer but more powerful physical machines
- ✦ **Achieve high availability with lower cost** - Deploy high availability infrastructure on virtual infrastructure to failover virtual machines
- ✦ **Match IT supply with resource demand** - Deploy applications and balance workloads between existing computer resources
- ✦ **Extend life of legacy environments** - Run legacy OSes and apps on modern, high performance servers

The Leading Enterprise Linux Utility Computing
From Desktop to Data Center

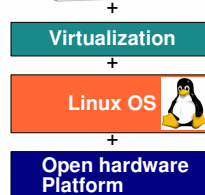


Where are we?

The Leading Enterprise Linux Utility Computing
From Desktop to Data Center



Virtualization + Cluster Infrastructure



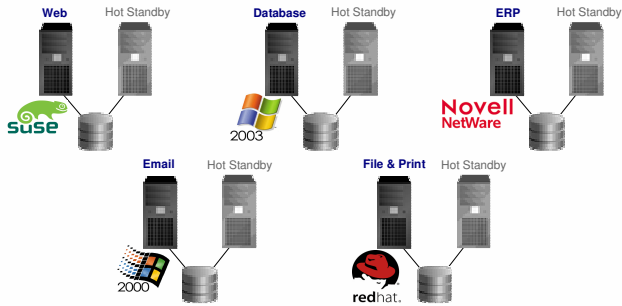
- ✦ High availability cluster management solution over virtual infrastructure
- ✦ Can cope with complex IT landscape – Multi-platform, legacy platforms and parallel applications
- ✦ On-demand N-node clusters, suitable for small to large distributed environment
- ✦ Hardened cluster infrastructure
- ✦ Consolidate all work loads on open platform - Linux

The Leading Enterprise Linux Utility Computing
From Desktop to Data Center



Consolidation of Diversed Infrastructure

10 servers , 50% utilization, deversed management

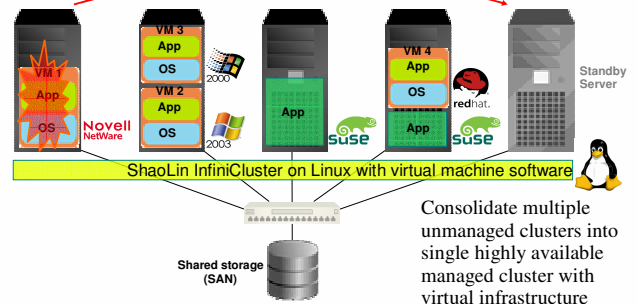


The Leading Enterprise Linux Utility Computing
From Desktop to Data Center



Multi-Node Clustering on Virtual Infrastructure

Failover



The Leading Enterprise Linux Utility Computing
From Desktop to Data Center



Multi-node Clustering on Virtual Infrastructure

- Share redundant elements across workloads
- Virtual machines are hardware-independent, don't need redundant systems or identical components
- Maximize utilization for mixed environment
- Add machines (both physical and virtual) on demand
- Single view cluster administration and storage management
- Lowest cost protection from hardware and software failures
- Migrate to new hardware without major downtime

The Leading Enterprise Linux Utility Computing
From Desktop to Data Center



Open Platform Cluster

- Runs on Open Source OS - Linux
- Multi-platform support (IA-32, IA-64, EM64T, x86 compatible 32/64, PPC)
- Supports SAN, NAS, shared RAID, replicated data storage
- Transparent to standard storage software technologies (software RAID, EVMS, LVM)
- Supports most commercial Linux distributions and kernel versions (2.4.x and 2.6.x)

The Leading Enterprise Linux Utility Computing
From Desktop to Data Center



Conclusion

The Leading Enterprise Linux Utility Computing
From Desktop to Data Center



Benefits Summary

- Maximize utilization, cope with both consolidation(scale-up) and on-demand(scale-out) strategies
- Load management, “on the fly” reallocation of server loads with automatic overload detection
- Lower cost
 - Physical resource reduction, sharing of redundant components
 - Encapsulate complex application resources as one single VM
 - Simplified cluster administration

The Leading Enterprise Linux Utility Computing
From Desktop to Data Center



Benefits Summary

- Reduce planned downtime with online manageability and scalability
- Reduce unplanned downtime elimination of SPOF and automatic fault recovery and notification
- Simpler to implement and execute
- Convert distributed unmanaged computing environment into managed, highly available open computing infrastructure

The Leading Enterprise Linux Utility Computing
From Desktop to Data Center



The End

Thank You

Scale Your Business To Infinite