Novell.
SUSE. Linux Enterprise 10
Your Linux is ready™

Quick Reference Card 16 UNIX/Linux HASF and Virtualization

Your?
host logo

Definitions #Xen Introduction (x86 32 & 64 bits only)
HASI #**H**igh **A**vailability **S**torage **I**nfrastructure
Hypervisor #Domain 0, host, microkernel, VM Monitor
#Priviliged domain
Virtual Machine (vm) #DomU, guest, unpriviliged domain
Migration #Relocation of vm; about 100ms downtime
STONITH #**S**hoot **T**he **O**ther **N**ode **I**n **T**he **H**ead
Full virtualization drivers #Licensed Driver Pack Performance boost
Xen best practices #novell.com/documentation **sles10 guide:**
Disable power management #Disable on host and guests
Do not use Network Manager #Use traditional network setup
dom0__mem=2048M #Set max. dom0 mem (grub kernel line)
Close VMM, when unused (leak) #SLES10 SPx features:
http://www.novell.com/documentation/vmserver/index.html
NOTE: Do not edit the initial startup files stored in /etc/xen/vm, because they are used only during the creation of a new virtual machine (vm).
xm list -l vmname > filename #Modify vm settings in filename
xm new -F filename #Import new vm settings into xend.
xm start vmname #Start the vm with its new settings
No Full Virtualization guarantee #Intel VT or AMD V cpu's work:
http://wiki.xensource.com/xenwiki/HVM_Compatible_Processors
Hot-add-memory not yet #Will be available via update (release note)
Max 16GB memory is stable #PAE kernel for memory > 4GB
Enable Multiple networks (bridges, ...):
http://pronetworkconsulting.com/linux/scripts/network-multinet.html
Live migration checks #Xen will try to migrate a VM without ever attempting to determine if the VM is capable to run on the destination:
vi /etc/xen/xend-config.sxp #Enable xend-relocation options
#Are the Dom0 kernels matched? #(32->32, 32PAE->32PAE, 64->64)
#Are the CPUs matched? #(P4->P4, Core2->core2, etc.)
#Can both the source and destination server see all of the VM's disks
#Is there enough free memory on the destination server?
#HA xen migration: http://www.novell.com/coolsolutions/feature/19676.html
#Same bridge names/numbers? #**Supported?**:
http://www.novell.com/surveyutil/survey.jsp?id=204
Xen VM Migration setup #Based on SLES10GA documents.
Install sles10 on 3 machines (2 nodes)
#Reserve >=4 GB unpart. space #Use bare part./LVM/EVMS/or sparse file
#Include High Availability & XEN pattern, and yast2-vm, yast2-heartbeat
#Enable ip_forward #By using yast or sysctl.conf
Packages iSCSI #iSCSI for demo purposes
open-iscsi & iscsitarget #Do not plan initiator & target on 1 node
Packages XEN #Physical Address Extension: 32bit & 4G
kernel-xen[pae], xen, xen-libs, xen-tools, xen-tools-ioemu
Survive a reboot
/etc/init.d/powersaved stop #Prevent power management conflicts
chkconfig powersaved off #Do not survive a reboot
/etc/init.d/xend start #Start XEN daemon & survive reboot:
chckconfig --level 35 xend on #Possibly troubleshoot network aft reboot
#If required, change eth0 settings in three steps:
/etc/xen/scripts/network-bridge stop netdev=eth0
yast2 lan
/etc/xen/scripts/network-bridge start netdev=eth0
Disable SuSEfirewall #Or enable xenbr0 forwarding:
vi /etc/sysconfig/SuSEfirewall2
FW_FORWARD_ALWAYS_INOUT_DEV="xenbr0"
/etc/init.d/SuSEfirewall2_setup restart

Default bootmenu #default title nr starts with 0
vi /boot/grub/menu.lst
title SUSE Linux Enterprise XEN Server 10
root (hd0,4)
kernel /boot/xen-pae.gz dom0_mem=2048M
module /boot/vmlinuz-xenpae root=/dev/sda5 vga=0x31a splash=silent showopts
module /boot/initrd-xenpae
http://en.opensuse.org/Installing_Xen3
brctl show #Output after a reboot, e.g.:

```
bridge name  bridge id          STP enabled     interfaces
xenbr0       8000.feffffffffff  no              vif0.0   peth0
```

Setup a Virtual Machine #Prepare XEN installation source, e.g.
#Copy SLES10 ISO file to host
fdisk -l #Show unpartitioned space
#No cluster support for SLE10 and LVM vg. Use EVMS container(s) instead.
vgdisplay #Show free space for LVM
lvdisplay #Show Logical Volumes
mount #Show mounted volumes
evms_activate #If EVMS is not used during installation
evmsgui #Show free space for EVMS
dmsetup ls #Map Volume Name to mapper device
system-vm2(253,3) #Maps to: /dev/dm-3
Yast, System, Virtual Machine Management (Xen)
xm create /etc/xen/vm/vm1 #Start VM. Create domain in memory.
xm list #Get <DomId>
xm console 2 #Open screen <DomId>
#Disconnect: <**Ctrl-]**>
<Ctrl><Ctrl><Ctrl><Alt><Fx> #Pass through keys (**3x**<**Ctrl**>) in VMM
xm top #Monitor domains. Base info: xm info
xm shutdown 2 #Shutdown VM <DomId>
xm reboot 2 #Restart VM <DomId>
xm destroy 2 #Kill VM <DomId>
xm migrate 2 host #Migrate <DomId> hostname
xm pause 2 #Standby VM <DomId>
xm unpause 2 #Start VM <DomId> from standby
xm save 2 file #Suspend VM <DomId> to filename
xm restore file #Resume VM from filename
xm mem-set 0 512 #Set memory of domain 0 to 512MB

Heartbeat 2 Setup #At least 2 ethernet cards, 2 nodes
#Recommended redundant NICs #Software RAID 1 is not supported
#Optional STONITH Agent for power supply
http://www.novell.com/documentation/sles10/hb2/data/hb2_config.html
Heartbeat 2 definitions #Four Layers
Messaging/Infrastructure Layer #"I am alive" "Heartbeat" Layer
Membership Layer #Cluster Consensus Membership service
Resource Allocation Layer #Resource Administration Services
Resource Layer #(RA) **R**esource **A**gents with scripts
Standby node #Node with ability to run a resource
Cluster **R**esource **M**anager #(CRM) Master of Ceremony. One is DC:
Designated **C**oordinator #(DC) Own/react on master CIB changes
Cluster **I**nformation **B**ase #(CIB) XML cluster setup/view
Policy **E**ngine #(PE) Step ordering
Transition **E**ngine #(TE) Step execution
Local **R**esource **M**anger #(LRM) Call Resource Agents
Setup name resolution #Make all hostnames resolvable on nodes
vi /etc/hosts #Or setup DNS. Check via: ping nodename

Setup time sync #Initial time set: ntpdate ntp.srv.com, or:
ssh node1 date $(date +%m%d%H%M)
ssh node2 date $(date +%m%d%H%M)
vi /etc/ntp.conf #Time syncs only with max 1000 sec delta
  server ntp.srv.com #Comment out server and fudge lines
/etc/init.d/ntp start #Start time sync and wait 5 minutes.
ntpq -p #Check time sync. (or ntptrace)
Initial heartbeat setup #Add nodes & propagate configuration
yast2 heartbeat #On one of the nodes
Add node(s), Next #Unlimited. Tested to 16
Select authentication key, Next #Same on all nodes (none, md5, or sha1)
On and Survive reboot #chkconfig --level 35 heartbeat on
/usr/lib/heartbeat/ha_propagate #Replicate configuration to nodes
cat /etc/ha.d/authkeys #View configuration file
cat /etc/ha.d/ha.cf #View configuration file
cat /var/lib/heartbeat/crm/cib.xml #Is replicated aft heartbeat start
/etc/init.d/heartbeat start #On the other node(s)
chkconfig --level 35 heartbeat on #On the other node(s)
passwd hacluster #Give user(s) hacluster a password for:
hb_gui #Add a resource (from any node), e.g.:
+ (add new item), native type, OK,
Resource ID: test-ip, IPaddr (OCF RA) as Type, Param. Value: ip:172.17.0.170,
optional: Add Parameter, Name: nic, Value: eth0, OK,
Start Resource (MB2) #Check via ping, ifconfig and unplug cable
Heartbeat administrative tools #From node wherever in the cluster
hb_gui #GUI for config. and monitoring cluster
crmadmin #Provide node related details
cibadmin #Query/modify current configuration
crm_verify #Check configuration validity
crm_mon #Show cluster status in text or HTML
crm_resource #Query/modify cluster resources/services
crm_standby #Control a node's standby status
cl_status #Show node-centric connectivity info
http://linux-ha.org/v2 #For detailed information

iSCSI Setup #Used as cheap SAN for Image Store
Initiator #User of the block level iSCSI device
Target #Host sharing the block device
Add a new partition on target #e.g. /dev/hda6, /dev/vg/lv,
yast disk #/dev/evms/lvm2/cont/lv
Do not mix LVM & EVMS #Best practices
No cluster support for SLE10 and LVM vg. Use EVMS container(s) instead,
but iSCSI provides a block device not an LVM volume group
dmsetup ls #Map Volume Name to mapper device
  system-vm2(253,3) #Maps to: /dev/dm-3
Configure iSCSI target #File, block device, RAID or LVM device
yast iscsi-server #Provide /dev/sd device for client
When booting, (Open Firewall) #In Service tab
Delete demo target #In Targets tab
Add target #e.g. /dev/vg-usb/lv-xen1
No authentication in demo #Next, OK, Finish, Restart, Yes
Configure iSCSI initiator #From first node
yast iscsi-client #Connect IET-VIRTUAL-DISK /dev/sda
When booting #In Service tab
Discovery, Fill in IP of target #In Discovered Targets tab
Login, no authentication #
Toggle startup, Finish #Survive reboot

yast disk #Add 3 partitions (IET-VIRTUAL-DISK):
Create part. for VM image #Leave e.g. 300MB free. No mount point
Create part. for VM config. files #Use e.g. 200MB. No mount point
Create part. for VM data storage #Use e.g. 100MB. No mount point
Configure iSCSI initiator #From other node(s) (discovery only)

OCFS2 Setup #Oracle Cluster File System
Can run in pure OCFS2 Cluster #Integration with heartbeat2 (user space)
ocfs2console #GUI for setup and propagation
Initialize the native OCFS2 stack #Only on ONE node
Cluster, Configure nodes, Close
Add nodes (incl. first), Close #Name, IP and port (7777)
Close, Cluster, Propagate Config. #Copy via ssh and close ocfs2console
/etc/init.d/o2cb configure #Add heartbeat and bootconfig (all nodes)
/etc/init.d/o2cb force-reload #On first node only (because of failure)
cat /sys/o2cb/heartbeat_mode #Check for 'user'
find /sys/kernel/config/cluster #Interface between kernel & user space
mkfs.ocfs2 /dev/sda1 #Create OCFS2 file systems (sda1 & sda2)
mounted.ocfs2 -d /dev/sda1 #Ask for UUID
OCFS2 in heartbeat cluster #Integrate
#Clonesets can run concurrent and on all nodes
#OCFS2 via clone File System RA on each node
#Notify resource stop/start, node join/leave
#Demo stonith device is ssh reboot:
#Simulate node crash by killing heartbeat and not unplug cable
vi cibbootstrap.xml #Create XML blobs for the CIB
cibadmin -C -o crm_config -x ./cibbootstrap.xml
cibadmin -C -o crm_config -x ./stonithcloneset.xml
cibadmin -C -o crm_config -x ./imagestorecloneset.xml
cibadmin -C -o crm_config -x ./configstorecloneset.xml
#Check for /var/lib/xen/images & /etc/xen/vm via:
mount
crm_mon -1 #List cluster resources
Test by manually umount ocfs2 #Remount should occur
Setup ssh keys for root #For unattended ssh stonith
Enable atd #For ssh stonith
pkill heartbeat #Test by node crash emulation
VM as Cluster Resource
Change sync mode of loop device #Undo after installing VM
vi /etc/xen/scipts/block #Remove -y at do_or_die
Create VM on node1 #Use default Sparse File
Restore -y #Sync mode
Stop VM on node1 #Inside VM is a not cluster safe fs
yast xen #Check availability on node2
cibadmin -C -o crm_config -x ./vmlocation.xml
cibadmin -C -o crm_config -x ./vm1.xml
crm_mon -1 #xm list
cibadmin -C -o crm_config -x ./vm1orderconstraints.xml
pkill heartbeat #Test by node1 crash emulation
crm_mon -1 #xm list on node2
cibadmin -Ql > cib-xen.txt #Dump all HA settings

Exend maximum loop mounts #Default max loops is 8
rmmod loop #Extend without reboot. **Rem**ove **mod**ule.
modprobe loop max_loop=64 #Extend without reboot
vi /etc/modprobe.conf #Was max_loop=64 as SLES9 boot par.
  options loop max_loop=64