State of Virtualization at SUSE

Michal Svec
Product Manager
msvec@suse.com

Mike Latimer
Engineering Manager
mlatimer@suse.com
Agenda

- Virtualization @SUSE
- Enhancements in XEN/KVM
- Enhancements in Linux Containers and Docker
- Virtualization with VMware
- Virtualization with Microsoft
Virtualization @SUSE
SUSE Virtualization Ecosystem

- Containers
- Live Patching
- SUSE Studio
- SUSE Storage
- SUSE Manager
- SUSE OpenStack Cloud
- SUSE Linux Enterprise Server
SUSE Virtualization

Virtualization is a key component in SUSE strategy!

Xen
• SUSE first to deliver Xen to the Enterprise in SLES 10 GA
• SUSE continues to support Xen in SLES 12 SP2

KVM
• SUSE first to deliver KVM to the Enterprise in SLES 11 GA
• SUSE first to deliver KVM on IBM System z in SLES 11 SP3

SUSE first to deliver Xen & KVM in OpenStack
Virtualization use cases

<table>
<thead>
<tr>
<th>In Your Data Center</th>
<th>In the Clouds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perfect Guest</td>
<td>Available in Public Clouds</td>
</tr>
<tr>
<td>Dual Hypervisor Support</td>
<td>Cloud Hosts</td>
</tr>
</tbody>
</table>
Virtualization Host

**Full Support** for leading open source hypervisors KVM and Xen

Available for major architectures:
- Intel 64/AMD64 (x86_64), IBM System z (s390x), IBM POWER8 (ppc64le) and ARMv8-A (AArch64)!

Complete virtualization solution:
- VM save/restore, snapshots, hotplugging, live migration, etc...

OS-level or application virtualization with **Linux Containers (LXC)** and **Docker**
Perfect Guest

Perfect Guest strategy, operating system tuned to run great as a guest on all major hypervisors
Virtualization in the Cloud

Public Cloud:
• Deploy workloads in the public cloud or cloud service provider of your choice

Private Cloud:
• Deliver cloud infrastructure solution powered by SUSE OpenStack Cloud

Hybrid Cloud:
• SUSE Manager to manage servers on premise and in the public cloud
Virtualization in the Cloud Ecosystem

SUSE Storage
Software defined storage for physical and cloud

SUSE OpenStack Cloud
Highly flexible and adaptable cloud infrastructure

SUSE Manager
- Provisioning
- Management
- Monitoring

SUSE Linux Enterprise
The foundation for your datacenter workloads and virtualization
Virtualization Enhancements in Xen and KVM
SLES 12 SP2 Virtualization

Major changes were introduced in SLES12

- xend → libxl
- lxc → libvirt-lxc
- qemu 1.4.1 → 2.0.2

Ongoing feature development in SP2

- Updated Virtualization Packages
  - xen 4.7.0
  - qemu 2.6.1
  - libvirt 2.0.0
Bug fixes, new features and enhancements...

• Updated libguestfs
  – Rewrite of Xen to KVM migrations (virt-v2v)
    • Increased speed, reliability and security
  – Physical to virtual conversions (virt-p2v (technology preview))

• Improved SPICE support in tools
  – USB redirection

• Guest Agent improvements

• Updated VMDP, Hyper-V drivers, VMware tools, etc...
SLES 12 SP2 Xen Highlights

General Status of Xen:
• Xen 4.7.0

Libvirt preferred management layer
• xl supported for thin management (not reflected in libvirt)

New Features
• PV-OPS kernel
• qemu-xen replaced by qemu-system-x86_64
• UEFI support through qemu-ovmf
• PVUSB support
Paravirt_ops (PV-OPS) Kernel?

Previously...
• kernel-xen was the only kernel providing paravirtual functionality

PV-OPS kernel
• Operates as both a normal Linux kernel and a paravirtual kernel
  – Supports bare-metal (non-virtual); Xen dom0/domU (PV and HVM); and KVM (fully virtual) guest
• Much easier to maintain
  – No more kernel-xen specific QA tests, security backports, kernel trees, etc...
• Offers additional capabilities
  – Indirect descriptors (now backported to SP1)
  – PVHVM mode
PV-OPS specific notes

Checking for Xen environments:

```bash
xen151:~ # uname -r
4.4.21-64-default
xen151:~ # xl list
<table>
<thead>
<tr>
<th>Name</th>
<th>ID</th>
<th>Mem</th>
<th>VCPUs</th>
<th>State</th>
<th>Time(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Domain-0</td>
<td>0</td>
<td>30153</td>
<td>88</td>
<td>r-----</td>
<td>3811.3</td>
</tr>
<tr>
<td>sles12sp2</td>
<td>2</td>
<td>1024</td>
<td>1</td>
<td>-b----</td>
<td>440.4</td>
</tr>
<tr>
<td>sles12sp2-fv</td>
<td>4</td>
<td>1024</td>
<td>1</td>
<td>-b----</td>
<td>443.9</td>
</tr>
</tbody>
</table>
```

- Dom0 has /proc/xen/privcmd
- PV or HVM? Check dmesg - e.g. “Booting paravirtual kernel on Xen HVM”

Block device naming:

- PV block devices **MUST** be named using the xvd* naming convention
  - Names such as sd*, hd* or vd* are no longer valid
## Xen Guest Modes

Virtualization modes are changing!

<table>
<thead>
<tr>
<th>Mode</th>
<th>Hypervisor</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fully Virtual</td>
<td>HVM</td>
<td>Requires virtualization hardware extensions</td>
</tr>
<tr>
<td>PV on HVM</td>
<td>HVM</td>
<td>Paravirtual disk, net drivers in HVM domain (e.g. SLES &lt;= 12 SP1)</td>
</tr>
<tr>
<td>PVHVM</td>
<td>HVM</td>
<td>PV on HVM, plus paravirtual interrupts and timers (e.g. SLES &gt;= 12 SP2+)</td>
</tr>
<tr>
<td>PVH, PVHv2, HVMlite</td>
<td>PV</td>
<td>Paravirtual kernel and drivers, plus virtualization hardware extensions</td>
</tr>
<tr>
<td>Fully Paravirtual</td>
<td>PV</td>
<td>Paravirtual kernel and drivers (cannot run in non-PV environments)</td>
</tr>
</tbody>
</table>
SLES 12 SP2 QEMU / KVM Highlights

General Status of QEMU/KVM:
- QEMU v2.6.1 and KVM in v4.4.21 kernel
- Available on Intel 64 / AMD 64, IBM System z, POWER 8 and AArch64!

Libvirt preferred management layer ...
- QEMU command line supported for access to more features

New features and changes:
- libiscsi support for improved iscsi access
- 3D acceleration using virtio-gpu
- UEFI support through qemu-ovmf
- Post-copy live migration
- PCI-passthrough is deprecated in favor of VFIO
QEMU Post-copy Live Migration

Traditional migration
• Machine stays on source until memory on host and target converges.
  – With very busy machines, memory may never converge!
    • Guest CPU throttling may be used to force convergence.

Post-copy live migration
• Machine is immediately transferred to target machine.
• Memory is synchronized, with page faults given priority.
  – Network outages during this phase can cause unpredictable failures!
Mixed Virtualization Environment?

SLE 11 and SLE 12?
• Migration is supported from product to product+1
  – SLES 12 SP1 → SLES 12 SP2
  – SLES 11 SP4 → SLES 12
    • No live migration between SLE 11 and SLE 12

Xen and KVM?
• Domains can be converted from Xen to KVM...
  – virt-v2v
    • Supported for SLE 11 or SLE 12 Xen host
    • Conversions are non-destructive as source disk is copied
### SLES 12 SP2 Supported Hosts

<table>
<thead>
<tr>
<th>Host</th>
</tr>
</thead>
<tbody>
<tr>
<td>SLES 12 SP2 (Xen + KVM)</td>
</tr>
<tr>
<td>SLES 12 SP1 (Xen + KVM)</td>
</tr>
<tr>
<td>SLES 11 SP4 (Xen + KVM)</td>
</tr>
<tr>
<td>VMware vSphere (ESXi) 6.5/6.0/5.5</td>
</tr>
<tr>
<td>Microsoft Windows 2016</td>
</tr>
<tr>
<td>Microsoft Windows 2012 / 2012 R2</td>
</tr>
<tr>
<td>Microsoft Windows 2008 SP2 / R2 SP1</td>
</tr>
<tr>
<td>Citrix XenServer 7 / 6.5</td>
</tr>
<tr>
<td>Oracle VM 3.4 / 3.3</td>
</tr>
</tbody>
</table>
Virtualization with VMware

VMware tools and drivers **integrated with SLES 12 SP2** for best out-of-the-box experience

- open-vm-tools (10.0.7)
- In-tree kernel modules:
  - `vm[wx]*.ko`
    - Ballooning, network, scsi, graphics, etc...
- New (workstation) features include:
  - Shared folders
  - drag-and-drop/copy-paste

Fully supported by VMware via L3 support agreement
Virtualization with Microsoft

Hyper-V
- Latest drivers and features supported in SLES
  - SUSE works directly with Microsoft to enhance and improve drivers
- Drivers are included with SUSE kernels
  - External Linux Integration Services (LIS) package is not required
  - Version numbers (such as LIS 3.5/4.1) are not applicable!
- Support for Windows Server 2016

Azure
- SLES 12 Linux RDMA image
  - Very low-latency network connection
  - Tuned for HPC workloads
## SLES 12 SP2 Supported Guests (1/2)

<table>
<thead>
<tr>
<th>SLES 12 SP2</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>SLES 12 SP1</td>
<td></td>
</tr>
<tr>
<td>SLES 12</td>
<td></td>
</tr>
<tr>
<td>SLES 11 SP4</td>
<td></td>
</tr>
<tr>
<td>SLES 10 SP4</td>
<td></td>
</tr>
<tr>
<td>SLED 12 SP1 (Tech Preview)</td>
<td></td>
</tr>
<tr>
<td>OES 2015 SP1</td>
<td></td>
</tr>
<tr>
<td>OES 2015</td>
<td></td>
</tr>
<tr>
<td>OES 11 SP3</td>
<td></td>
</tr>
<tr>
<td>NetWare PV 6.5 SP8 (32-bit)</td>
<td></td>
</tr>
<tr>
<td>RHEL 7.2+ / 6.8+ / 5.11+ (*)</td>
<td></td>
</tr>
</tbody>
</table>
## SLES 12 SP2 Supported Guests (2/2)

<table>
<thead>
<tr>
<th>MS Windows 2016</th>
</tr>
</thead>
<tbody>
<tr>
<td>MS Windows 2012 R2+</td>
</tr>
<tr>
<td>MS Windows 2012+</td>
</tr>
<tr>
<td>MS Windows 2008 R2 SP1+</td>
</tr>
<tr>
<td>MS Windows 2008 SP2+</td>
</tr>
<tr>
<td>MS Windows 10+ (best effort)</td>
</tr>
<tr>
<td>MS Windows 8.1+ (best effort)</td>
</tr>
<tr>
<td>MS Windows 8+ (best effort)</td>
</tr>
<tr>
<td>MS Windows 7 SP1+ (best effort)</td>
</tr>
</tbody>
</table>
VMDP 2.4 for best guest support

• Support for SUSE Linux Enterprise Server 12 SP2
• Support for Microsoft Windows Server 2016, 2012 R2 and Windows 10
• Unified driver for easier migration from one hypervisor to another (Xen → KVM)
• Simple host to guest access (Windows Guest Agent)
• Numerous bug fixes and enhancements
Outlook – SLES 12 SP3

- Continuing support for both Xen and KVM
- Continuing hardware enablement
- Further enhancements in Xen guest modes (HVMLite)
- Nested virtualization with KVM?
Best Practices

Virtualization can be complicated!
• Normal physical machine issues + virtualization layer

Virtualization Best Practices Guide
• Regularly updated tips directly from SUSE Virtualization developers
  – Recommended CPU and memory features and usage
    • Pinning, NUMA, etc...
  – Block and filesystem recommendations
• Covers SLES11SP4 and SLES12 (GA → SP2)

Also see: SLES Virtualization Guide
Virtualization Enhancements
Container Technologies
Linux Containers

• Lightweight virtualization
  - Faster provisioning, less downtime
  - Higher virtualization density

• Flexibility and agility
  - Containerized apps can be deployed anywhere
  - Normal I/O, no congestion

• Near native performance
  - IBM research: http://ibm.com/Search/?q=rc25482
Linux Containers

- **System containers**
  - Full system in the container (no kernel)
  - libvirt-lxc

- **Application containers**
  - One process per container
  - Docker, rkt, ...
  - Rich ecosystem
Container Lifecycle

Create

Run

Maintain

Zypper

Zypper Docker

KIWI

docker

Portus

docker

kubernetes

openstack

YaST
Collaborate Securely with Portus

Open Source
Created by SUSE

Authenticate
Authorize
Control access to your images

Easy to Use
Navigate Image Catalog

Audit
Keep everything under control
Agile Platform Ideal for Containers

- Mini SUSE Linux Enterprise Server
- Ideal for Bimodal
- Designed for the Future

- small fast agile
- Mode 1
- Mode 2
Surgically Patch Container Application

Check Containers & Images (both!)

Identify Vulnerable Apps

Update Image

Update app

```
sles12sp1:~ # zypper docker list-patches --cve=CVE-2014-6271 stable_image:1

<table>
<thead>
<tr>
<th>Issue</th>
<th>No.</th>
<th>Patch</th>
<th>Category</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>cve</td>
<td>CVE-2014-6271</td>
<td>slessp3-bash-9740</td>
<td>security</td>
<td>needed</td>
</tr>
</tbody>
</table>
```

Shellshock vulnerability
Orchestration Enhances Business Value of Containers

- **Containerized Micro-server**
  - Zero downtime,
  - Highly available,
  - Easy to migrate across hosts

- **On-demand Self-services**
  - Agile delivery,
  - Continuous development

- **Continuous Integration**
  - App Lifecycle Pipeline
  - Dev, Test, Prod

---

- **Deployment**
- **Networking**
- **High Availability**
- **Scaling**
- **Monitoring**
Bimodal Datacenter

Mode 1
- Compute
- Storage
- Network

Mode 2
- Magnum
- Kubernetes
- Containers

Deployment  Networking  High Availability  Scaling  Monitoring
Docker Resources

- Docker mini-course videos
  - https://www.suse.com/promo/sle/docker.html
- Try SUSE Linux Enterprise Server 12
- SUSE Docker QuickStart
  - https://www.suse.com/documentation/sles-12/singlehtml/dockerquick/docke
    rquick.html
- More information in SUSE Linux Enterprise 12