Utilizing KVM in production
KVM Gets Down to Real Business

Qumranet Desktop Virtualization Solutions by Dor Laor

June 12, 2008
Agenda

- SolidIce Quick view
  - Motivation – challenge - solution
- Product components & features
  - KVM hosts
  - Management
  - Spice – Remote desktop protocol
  - Storage
- KVM usage in detailed
  - Problems solved
  - Problems to be solved
  - Stable branch & testing
- Performance & Conclusions
At a glance

- SolidIce is desktop virtualization solution
  - Uses KVM hypervisor
  - Uses Linux

- Strives to preserve the user experience
  - Uses home grown remote desktop protocol - Spice

- Various deployment options
  - Server Based Computing
  - Client Based Computing
  - Branch offices
Motivation For Desktop Virtualization

**PROVISIONING**
- Provisioning new machines
- Supporting new hardware

**SUPPORT AND MAINTENANCE**
- Desk-side visits
- Patches and updates

**FAILURES**
- Desktop backups
- Failures (HD, PS, fans etc)
- Disaster Recovery

**SECURITY AND COMPLIANCE**
- Enforcing security policies
- Compliance and audit trails

**POWER**
- Increasing power costs

- **Total cost of maintaining and managing a PC environment is $3,000 - $5,000 +**

- **Complexity increases with multiple sites and lack of trained desktop personnel.**

* Gartner 2007
The Challenge – End User Experience

Approaches to solve desktop delivery challenges

- **Terminal Services**
  - Not enough isolation
  - Apps need to be multi-user enabled
  - Complex to roll out new applications

- **VDI**
  - End-user experience is not sufficient
  - System cost is prohibitive
  - Not built/ready for desktops

End-user experience vs. Ease of Management: Solid ICE and Terminal Services compared to PC and “VDI.”
Solid Ice: Conceptual Overview

- User's desktop runs inside a KVM Virtual Machine in the data center
- User connects to virtual desktop using SPICE

1. Virtual Desktop Server (VDS) Cluster
2. Virtual Desktop Controller (VDC)
3. Storage

Thin Client or repurposed PC

SPICE Remote rendering technology
How Stateless Solid Ice works

USER VIEW

SYSTEM FUNCTIONING

User powers on the power-client, and sees the log-in screen
How Stateless Solid Ice works

**USER VIEW**

**SYSTEM FUNCTIONING**

- User logs in with AD username and password

![Diagram showing system functioning with VDC, AD, Power-client, and Storage]
How Stateless Solid Ice works

**USER VIEW**

**SYSTEM FUNCTIONING**

VDC authenticates the user against AD

- VDC
- AD
- Power-client
- Storage
How Stateless Solid Ice works

VDC presents the user with a list of virtual desktops he/she is allowed to log into (in case of multiple)
How Stateless Solid Ice works

USER VIEW

SYSTEM FUNCTIONING

VDC
AD

Power-client

Storage

User selects the desktop that he/she wants to log into (if they have multiple desktops)
How Stateless Solid Ice works

USER VIEW

SYSTEM FUNCTIONING

VDC provides the power-client with relevant information and credentials
How Stateless Solid Ice works

USER VIEW

SYSTEM FUNCTIONING

The power-client streams/runs the image from central storage
How Stateless Solid Ice works

The power-client launches the virtual desktop with a local spice session
Solid Ice: Virtual Desktop Server

- **Leverages KVM & Linux**
  - Networking, bridging, taps
  - Priorities
  - Memory - Page-sharing, ballooning, swapping
  - Encryption for controller connection & remote clients
  - Uses NFS
  - Unix sockets for monitor & vmchannel

- **High Density of Virtual Machines/Desktops**
  - Shared pages, ballooning and swapping
  - Spice with hw acceleration
  - 2-5x other solutions (workload dependent)

- **High Availability: Operates in clustered model**
  - Controller can kill (fense) a server and re-run VMs on another

- **Easy to Use**
  - Automatic installation of packages
  - Single Sign On
Solid Ice: Virtual Desktop Controller

- **Built from the ground-up for desktops.**
  - 90% reduction in storage
  - Load balance highly utilized server
  - AD/LDAP integration, multiple domains
  - Integrates with desktop build process
  - HA Monitoring of VDSs.
  - Audit trails, reports and more

- **Manage every aspect of the virtual desktop infrastructure from one central location**

- **Search driven User Interface**
  - Cross-correlate information instantly (virtual desktops, servers, users etc)

- **Power-shell based scripting API**
Spice – Remote Desktop Protocol

- **User-experience indistinguishable from physical PC**
  - 30+ frames per second video
  - Native color, full resolution
  - Bi-directional audio and video for video-conferencing/VoIP
  - Multi-monitor support (4+ monitors)
  - USB 1.1 and 2.0 support

- **SPICE is an adaptive remote rendering solution**
  Utilizes graphics processing capabilities wherever they exist
  - Offloads graphics processing/rendering to client (if powerful) and/or server

- **Improves desktop density on the server**
  - Minimizes rendering on the server side (adaptive)
Cool Features

- **Integrated**
  - Client-VM-Server-Active directory
- **Super dense**
- **Enhanced remote desktop experience**
  - Video / Multiple monitors / Audio / USB redirection
- **Search based management rocks**
  - Need to manage $10^5$ objects
- **Image management**
  - Templates, shrinkage, collapse-backup, sync with remote site
- **External hibernation, live migration, live snapshots**
- **Many use cases**
  - Server based / Client based / Branch office
  - Runs Windows*, Linux* guests
## Benefits Of Solid Ice For All Stakeholders

<table>
<thead>
<tr>
<th><strong>End-users</strong></th>
<th><strong>IT administrators</strong></th>
<th><strong>Enterprise</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>User experience indistinguishable from physical PC</td>
<td>Centralized management</td>
<td>Full control over costs of desktops</td>
</tr>
<tr>
<td>Non-intrusive to user</td>
<td>Fully integrated system built for virtual desktops</td>
<td>TCO savings of &gt;50%</td>
</tr>
<tr>
<td>Access work PC from anywhere</td>
<td>Open, architecture (no high-end servers, proprietary file systems etc)</td>
<td>IRR &gt;100%</td>
</tr>
<tr>
<td>Computing power on demand</td>
<td>Non intrusive to IT</td>
<td>Improved auditing ability/ compliance</td>
</tr>
<tr>
<td></td>
<td>Significantly reduced storage requirements</td>
<td>Power saving approach, green enterprise initiative</td>
</tr>
</tbody>
</table>
# Solid Ice Potential Use Cases

<table>
<thead>
<tr>
<th>Use-case</th>
<th>Customer Benefits</th>
</tr>
</thead>
<tbody>
<tr>
<td>IT Training Rooms</td>
<td>Eliminate provisioning and re-provisioning time</td>
</tr>
<tr>
<td>Consultants &amp; Contractors</td>
<td>Security, provisioning time, capex</td>
</tr>
<tr>
<td>Developer Desktops</td>
<td>Linux PC to WinXP VM, or vice-versa, on-demand</td>
</tr>
<tr>
<td>Outsourcing</td>
<td>Intellectual property and data security</td>
</tr>
<tr>
<td>Primary Desktop</td>
<td>Significant manageability benefits, ROI</td>
</tr>
<tr>
<td>Work-from-home</td>
<td>Access work PC (virtual desktop) from home, capex.</td>
</tr>
<tr>
<td>Disaster Recovery</td>
<td>Pool of DR desktops, or individual desktops</td>
</tr>
</tbody>
</table>
Use KVM :)  
Use RHEL/SLES Linux for Virtual Desktop Servers  
Windows standard HAL performs better  
Time drift fix is a must  
- -no-kvm-pit --> -no-irq-chip :(  
Migration keep breaking every qemu merge  
- Regression tests  
Tpr optimization/FlexPriority is a must for windows acpi  
Use templates based snapshots  
Monitor/vmchannel unix socket  
External hibernation – migrate into file  
Every device has a fallback  
- Virtionet-rtlnet, spice-vnc-rdp, networking- vmchannel,  
Upgrades using cdrom (change cd)  
Massive tests are paramount
Developed Features

- Good hypervisor to form a great community!
- VMDK image format support
  - Actually we now use it only as a migration path
- Time drift fix
  - Rtc, pit
- Pvnet, virtionet, virtio network drivers for windows
- E1000 network drivers
- Migration (along with Anthony Liguori)
- In-kernel apic (beginning)
- Tpr optimization
- Balloon driver (+Marcello, Rusty)
- Vmchannel driver
- KSM
- Usb isochronous; 2.0

- Task switching (Ghost)
- MMU notifiers
  - Kprobes
- AMD support (with amd of course)
- > 3.75G support
- Zero block de-dup
- Image sync – dirty bits
- Monitor shared library for Stable machine readable protocol
- Various time drift fixes
  - RTC, Get rid of in-kernel pit
- Migration issues
  - AMD – Intel migration
    - Choose performance vs flexibility
  - Least common denominator
  - Masquerade the cpuid bits
- Use libvirt
- Virtio based vmchannel, balloon (windows)
- Static pci slots
- Stabilize scsi for windows
- Stabilize the various Linux guests
Versions, Stability & Tests

- Our mainline features are massively tested by internal QA
  - However we do not test all types of Linux guests.
  - If things break in mainline tree we discover it way too late
  - Until May 2008 we used to branch mainline in a random point

- Now we base it over community published stable branch
  - Official kernel release – starting at 2.6.26
  - Use latest 'stable' userspace – kvm-68
  - Only apply bug fixes
  - Until it reaches standard distributions it can serve as community stable branch

- Build open source test framework
  - Stabilize KVM
  - Discover regression fast
Results - Office workload

- Started batches of 5 Virtual Desktops (VD) at a time, every 5 minutes
- Each of the VDs ran a script of office workload continuously
- One VD was constantly manned by a real user doing office workload
- The test stopped when the user could not do natural office work
Office workload – Solid ICE Memory Usage

> Memory % used (server = 16GB)

Adjusted Time

Solid ICE virtual desktops limit (52)
Office workload – Solid ICE CPU Usage

Adjusted Time

CPU % used (server = 2x4 cores)

Solid ICE virtual desktops limit (52)
How we tested – Mixed Multimedia workload

- Started batches of 5 Virtual Desktops (VD) at a time, every 5 minutes
- 4 of the VDs ran a script of office workload continuously
- The 5th VD ran full screen video
- One VD was constantly manned by a real user doing office workload
- The test stopped when the user could not do natural office work or the video was completely unwatchable
Video workload – Solid ICE Memory Usage

Memory % used (server = 16GB)

Adjusted Time

Solid ICE virtual desktops limit (41)
Video workload – Solid ICE CPU Usage

Adjusted Time

CPU % used (server = 2x4 cores)

Solid ICE virtual desktops limit (41)
Conclusions

- Qumranet enjoys from KVM
- KVM enjoys from Qumranet
- We are the community
- KVM is ready for mainline
Thank You

June 12, 2008