Introduction to SLA & Scheduling Features for Version 3.5
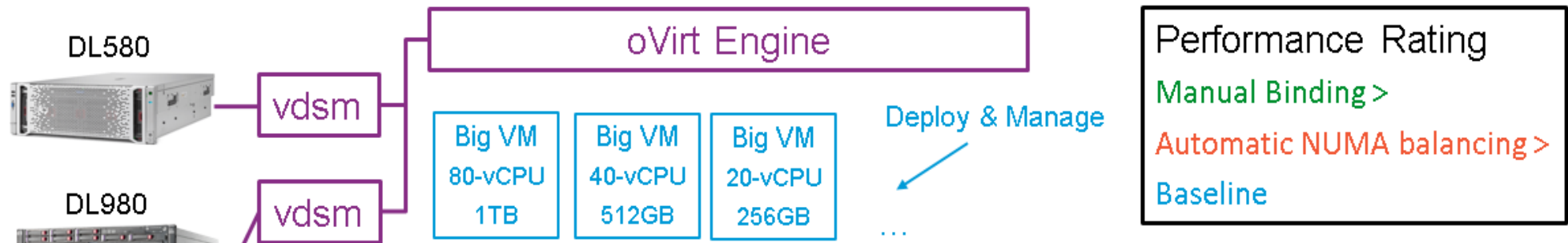
April 2014

Agenda

- NUMA integration, David and Vinod (HP), 10min
- Opta planner integration, Martin, 10min
- Hosted Engine on SAN support, Jirka, 10min
- CPU limits, Kobi, 10min
- blkio limits, Gilad, 10min
- Q&A, rest of time

NUMA control for oVirt

DL580

DL980

vdsm ── oVirt Engine

| Big VM 80-vCPU 1TB | Big VM 40-vCPU 512GB | Big VM 20-vCPU 256GB |

Deploy & Manage

Performance Rating
Manual Binding >
Automatic NUMA balancing >
Baseline

Feature 1: Query target host's NUMA topology

Ability from the UI, RESTful API and other APIs to gather a given host's NUMA topology(# of NUMA nodes, CPUs & total memory per node, NUMA node distances), NUMA statistics(Free memory per node, CPUs & memory usage per node). Besides consuming this information for planning/provisioning guests and oVirt scheduler etc. There may be other likely consumers for now and in the future.

Feature 2: NUMA bindings of guest resources (vCPUs & memory)

Ability from the UI, RESTful API and other APIs to optionally specify the bindings for backing memory of a guest (i.e. via numatune with mode set to: strict, preferred or interleave) along with the vCPU pinning across a desired set of host NUMA nodes. Automatic NUMA balancing feature will be introduced in kernel 3.13. As this technology matures it should reduce the need for having to specify explicit NUMA bindings.

Feature 3: Virtual NUMA topology

Ability from the UI, RESTful API and other APIs to specify virtual NUMA nodes for a mid/large sized guest. This helps the OS running in the guest to do NUMA aware allocation of data structures and scale better. Automatic NUMA balancing in the guest kernel can take advantage of this too.

http://www.ovirt.org/Features/NUMA_and_Virtual_NUMA
http://www.ovirt.org/Features/Detailed_NUMA_and_Virtual_NUMA

…

3

Feature's Patches and Status

No.	Story	Task	Owner	Status	Design Latest Updated Date	Submit Date	Merge Date	Comments
patch 1	VDSM patches	Get capability of NUMA information	HP	Code in Review	3/28/3014	1/26/2014		http://gerrit.ovirt.org/#/c/23703/
patch 2		Get statistics of NUMA information	HP	Develop				
		NUMA API for MOM	HP					included in patch 1 and patch 2
patch 3		Set numatune and virtual numa topology	HP	Merged	3/21/3014	3/3/2014	3/24/2014	http://gerrit.ovirt.org/#/c/25254/
patch 4	Engine core patches	Business Entities for NUMA feature	HP	Design in Review	4/4/2014	2/24/2014		http://gerrit.ovirt.org/#/c/23702/
patch 5		Query and Action for NUMA Entities	HP	Design in Review				depend on patch 4
patch 6		VDSM/engine integration	HP	Design in Review				depend on patch 4
patch 7	Database pathes	Tables and Views change for NUMA feature	HP	Design in Review	4/4/2014			
patch 8		Query SQLs and Action SQLs for NUMA	HP	Design in Review				depend on patch 7
patch 9	Scheduler patches	VM scheduler with NUMA aware	HP	Design in Review	4/4/2014			
patch 10	Restful API patches	Get NUMA information from restful API	HP?	Design in Review	4/8/2014			
	UX patches	POC for d&d, and d&d ui infrastructure	Redhat					
		gui: support basic functionally	Redhat					
	Test patches	functional tests for vdcClient	HP					benchmark compare report

Optaplanner Integration

- External service that gets data from engine (REST, db)
- Based on Optaplanner optimization engine
- Computes the optimal Vm to hosts assignments
- Runs constantly to refine the solution while adapting to changes
- Reports the solution to administrator using an UI plugin

http://www.ovirt.org/Features/Optaplanner

Self Hosted Engine on SAN

- The iSCSI support is planned for 3.5
  - The he-setup will provide the means to setup the iSCSI storage during the initial setup
  - Patches are already in gerrit going through the review
- We're also considering to add support for the fiber channel (but can't commit to it for 3.5)
- Patches are always welcome! :)

http://www.ovirt.org/Feature/Self_Hosted_Engine_iSCSI_Support

CPU Limits

oVirt

- libvirt quota and period (cpu tuning) will be supported by oVirt.
- Create a CPU QoS object to hold a number that represent limited CPU workload (calculation of both fields).
- CPU Profile entity to be coupled with every VM to hold QoS object, and other (future) CPU related fields.
- Supporting upgrade.

http://www.ovirt.org/Feature/CPU_SLA

blkio Limits

oVirt

- libvirt throughput and IOps (ioTune support) will be supported by oVirt.
- Create a Storage QoS object to hold max, read & write throughput and IOps (6 fields).
- Disk Profile entity to be coupled with every disk to hold QoS object, and other storage related fields.
- Supporting upgrade.

http://www.ovirt.org/Feature/blkio-support

Questions?

oVirt

THANK YOU !

http://www.ovirt.org
http://www.ovirt.org/Category:SLA

http://lists.ovirt.org/mailman/listinfo
devel@ovirt.org

#ovirt irc.oftc.net
shangchun.liang@hp.com (David)
chegu_vinod@hp.com (Vinod)
msivak@redhat.com (Martin)
jmoskovc@redhat.com (Jirka)
kianku@redhat.com (Kobi)
gchaplik@redhat.com (Gilad)