Policies for Automatically Dedicating CPUs to Workloads

Tomáš Golembiovský
Software Developer

09/2021
CPU and Topology

- Socket – physical connector on motherboard for CPU package
- Die – piece of semiconducting material on which cores are fabricated \((\textit{not configurable in oVirt})\)
- Core – a processor
- Thread – logical unit sharing resources with other threads on core

- Often you don’t really care where your virtual CPUs run
- Virtual topology does not have to match physical topology
CPU and Topology

- **Socket** – physical connector on motherboard for CPU package
- **Die** – piece of semiconducting material on which cores are fabricated *(not configurable in oVirt)*
- **Core** – a processor
- **Thread** – logical unit sharing resources with other threads on core

- Often you don’t really care where your virtual CPUs run
  … *but sometimes you do!*
- Virtual topology does not have to match physical topology
  … *but it can help performance if it does.*
NUMA – Non-Uniform Memory Access

- Each node has separate:
  - CPUs
  - memory controller and memory
  - IO controllers and devices
- Locality matters
- Typically NUMA node = Socket, but this is not a rule

Source: HPC Wiki (CC BY-SA)
CPU Assignment in oVirt
CPU Pinning

- Specified by pinning string
- Difficult to understand
- Difficult to write
- Requires host pinning

0#3_2#1-2,12_5#3,4,10,^10_6#6-9,^8_9#13-15
NUMA Auto Pinning

- Assigns CPUs based on host topology
- Only one policy “Resize and Pin” that resizes the CPU topology of VM based on our advices for SAP HANA users
- Effective on VM edit
- Does not change on VM start
Limitations

- Static or evaluated on VM edit
- Require host pinning
- CPUs are shared (!)
Dedicating CPUs
Dedicating CPUs

- Pins virtual CPUs (vCPUs) to physical CPUs (pCPUs) — 1 to 1
- Adapted on each VM start
- CPUs are assigned to VM exclusively
- VM runs on any host that satisfies topology requirements

- Policies:
  - Dedicated
  - Isolate Threads
  - Siblings
Intended Use

For performance tuning.

Not a security feature!
Dedicating CPUs

1. Engine analyses available resources
2. VM scheduled to start on a host
3. Other VMs are evicted from dedicated CPUs
Shared Pool

- VMs without policy use CPUs from shared pool
- Initially contains all CPUs on host
- Never empty – contains at least one CPU where VDSM process runs
Policy: Dedicated

- Simplest form of CPU assignment
- No special requirements on allocation
- Only CPU topology is taken into consideration

Photo: Rainer Knäpper (Free Art License)
Policy: Dedicated

Situation 1

vCPUs: 3
Topology: 1:3:1

Shared pool
Assigned to VM

Situation 2
Policy: Dedicated

- Only one option
- vCPUs: 2
  Topology: 1:1:2

- Shared pool
- Assigned to VM
Policy: Isolate Threads

- Each vCPU is placed on a separate core
- Other VMs cannot share allocated cores
- Emulates host architecture without SMT
- For hosts without SMT it is same as dedicated policy
- Allocates whole CPU cores
- vCPUs are pinned to pCPUs 1:1
Policy: Isolate Threads

vCPUs: 3
Topology: 1:3:1

vCPUs: 2
Topology: 1:1:2

- Shared pool
- Assigned to VM
- Blocked
Policy: Isolate Threads vs. Dedicated

Isolate Threads

Dedicated

vCPUs: 3
Topology: 1:3:1

Shared pool
Assigned to VM
Blocked
Policy: Siblings

- Requires host with SMT
- Places all vCPUs on same CPU core(s)
- Other VMs cannot share allocated cores

- Allocates whole CPU cores
- vCPUs are pinned to pCPUs 1:1
Policy: Siblings

vCPUs: 3
Topology: 1:3:1

vCPUs: 2
Topology: 1:1:2

- Shared pool
- Assigned to VM
- Blocked
Policy: Siblings vs. Dedicated

**Siblings**

- vCPUs: 3
- Topology: 1:3:1

**Dedicated**

- Shared pool
- Assigned to VM
- Blocked
Final words...

- Expected to work for all normal flows
- CPU set can change during:
  - Migration
  - Hibernation and resume
  - Snapshot and restore
- CPU set will not change for:
  - Pause and resume
- UI aspects and visualisation of CPU assignment on host or VM level not yet decided
- Planned for oVirt 4.5
  https://ovirt.org/develop/release-management/features/virt/dedicated-cpu.html
Thank you!

https://ovirt.org/

users@ovirt.org

@ovirt