Microsoft Hypervisor Top-Level Functional Specification (Hv#1) implementation in vSphere

hypervisor.cpuid.v0

- HWV 7 on all guests.
- CPUID.1.ECX - bit 31, “hypervisor bit”
  - Indicates CPUID leaves 0x4000000 - 0x40000ff maybe available by the hypervisor.
- Necessary for SVVP certification for Windows Server products.

Hypervisor bit side effects

- QueryPerformanceCounter (QPC) on Windows 7 and later will not use RDTSC by default if the “hypervisor bit” is set.
  - RDTSC is supported on Win8 with Hv#1 Reference Time.
- Hot-add memory is enabled on all Windows SKUs.
- IE9 initially disabled hardware acceleration when not on a MSFT compliant hypervisor.
- Prevents Hyper-V enablement.

hypervisor.cpuid.v1

- HWV9 with Win8 / Server 2012
- CPUID.0x40000001.EAX - Hv#1
- CPUID.0x40000003.EDX:3 - VCPU hot-add
- CPUID.0x40000004.EAX:5 - Disable watchdog BSOD timers

hypervisor.cpuid.v2

- HWV10 on all Windows guests
- Adds support for the minimal Microsoft compliant hypervisor requirements.
- Timer frequencies available
- Reference Time support
- Guest OS ID
- Hypercall support
- VP Index support

Timer frequencies available

- Windows 8 / Server 2012 support
- CPUID.0x40000003.EDX:8 - Timer TSC/APIC frequencies available.
- MSR_HYPERV_TIME_TSC_FREQUENCY (0x40000022)
- MSR_HYPERV_TIME_APIC_FREQUENCY (0x40000023)
- Eliminates need for guest timer calibration.

Reference Time

- Alternative time source for QPC.
- Returns time in 100-ns units.
- Uses the Reference TSC if possible, then fallsback to the Reference Counter.
- Reference TSC

```
typedef struct HvRefTscPage {
  volatile uint32    tscSequence;
           uint32    reserved1;
  volatile uint64    tscScale;
  volatile int64     tscOffset;
} HvRefTscPage;
```

Reference Time (cont)

- CPUID.0x40000003.EAX:9
- MSR 0x40000021
- Calls RDTSC and reads the TSC page for inputs.
- ReferenceTime = ((RDTSC * TscScale) >> 64) + TscOffset
- TSC sequence valid values are 0 – 0xFFFFFFFE. A value of 0 causes the reference counter to be used.

Reference Time (cont)

- Reference Counter
  - CPUID.0x40000003.EAX:1
  - MSR 0x40000020
  - Returns the Reference Time calculation using VCPU-0’s TSC value.

QueryPerformanceCounter Sources

	Windows Vista / Windows Server 2008	Windows 7 / Windows Server 2008 R2	Windows 8 / Windows Server 2012
HWV8 / HWV9	HPET	HPET	HPET
HWV10	HPET	Reference Time	TSC

Other features

- Guest OS ID MSR
  - MSR_HYPERV_GUEST_OS_ID (0x40000000)
  - Guest reported value

63:48	47:40	39:32	31:24	23:16	15:0
Vendor ID	OS ID	Major Ver	Minor Ver	Service Ver	Build Num

- Hypercall MSR
  - CPUID.0x40000003.EAX:5
  - MSR_HYPERV_HYPERCALL (0x40000001)
  - Maps a no-op code page into guest memory that returns HV_STATUS_INVALID_HYPERCALL_CODE.

Other features (cont)

- VP Index MSR
  - CPUID.0x40000003.EAX:6
  - MSR_HYPERV_VP_INDEX (0x40000002)
  - Returns vcpu id

Future work

- > 64 vcpu support.
  - Windows Server 2008 R2 and earlier are limited to 64 vcpus on an MSFT compliant hypervisor.
  - Windows Server 2012 and later have an option to remove this limit.
  - CPUID.40000005.EAX must be set to 0xFFFFFFFF.
- PMC feature enablement
  - CPUID.0x40000003.EDX:2
  - Performance Monitor support is available.

Future work (cont)

- Intelligent Timer Tick Distribution (ITTD)
  - Restricts timer interrupts to the BSP.
  - Requires support for Virtual Processor Idle Sleep State
  - CPUID.0x40000003.EAX:10
  - MSR_HYPERV_GUEST_IDLE (0x400000F0)

References

- Hypervisor Top-Level Functional Specification 2.0A: http://www.microsoft.com/en-us/download/details.aspx?id=18673
- Requirements for Implementing the Microsoft Hypervisor Interface: http://msdn.microsoft.com/en-us/library/windows/hardware/hh975392.aspx
- https://wiki.eng.vmware.com/MicrosoftHypervisorTLFS