

Virtual SQL Server Architecture

David Klee



About David Klee





 @kleegeek
 davidklee.net
 heraflux.com
 linkedin.com/in/davidaklee

Specialties / Focus Areas / Passions:

- Performance Tuning & Troubleshooting
- Virtualization
- Cloud Enablement
- Infrastructure Architecture
- Health & Efficiency
- Capacity Management

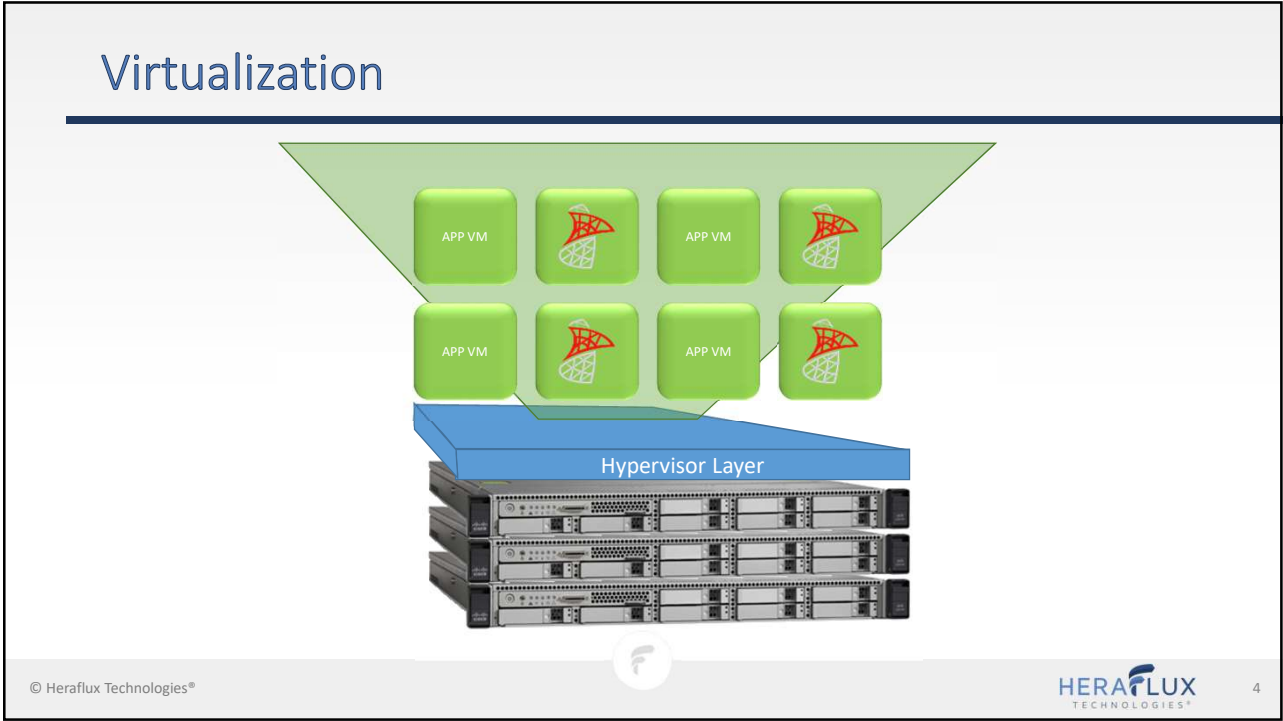
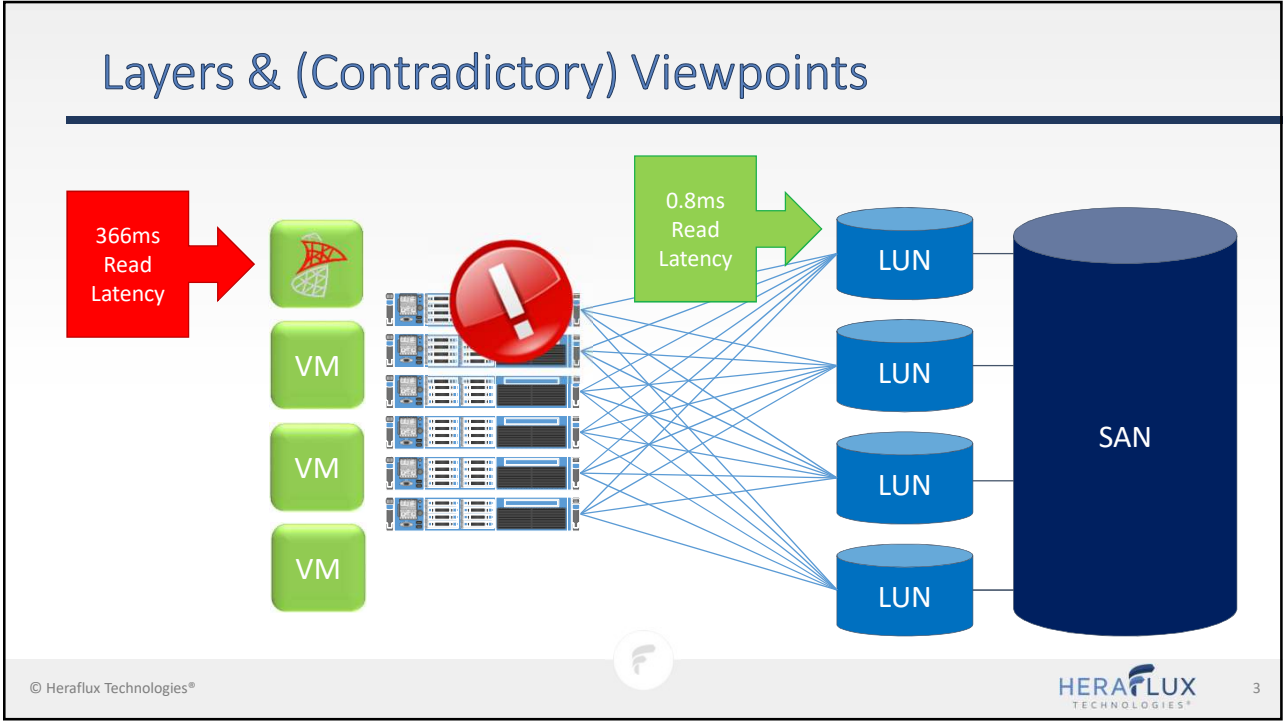


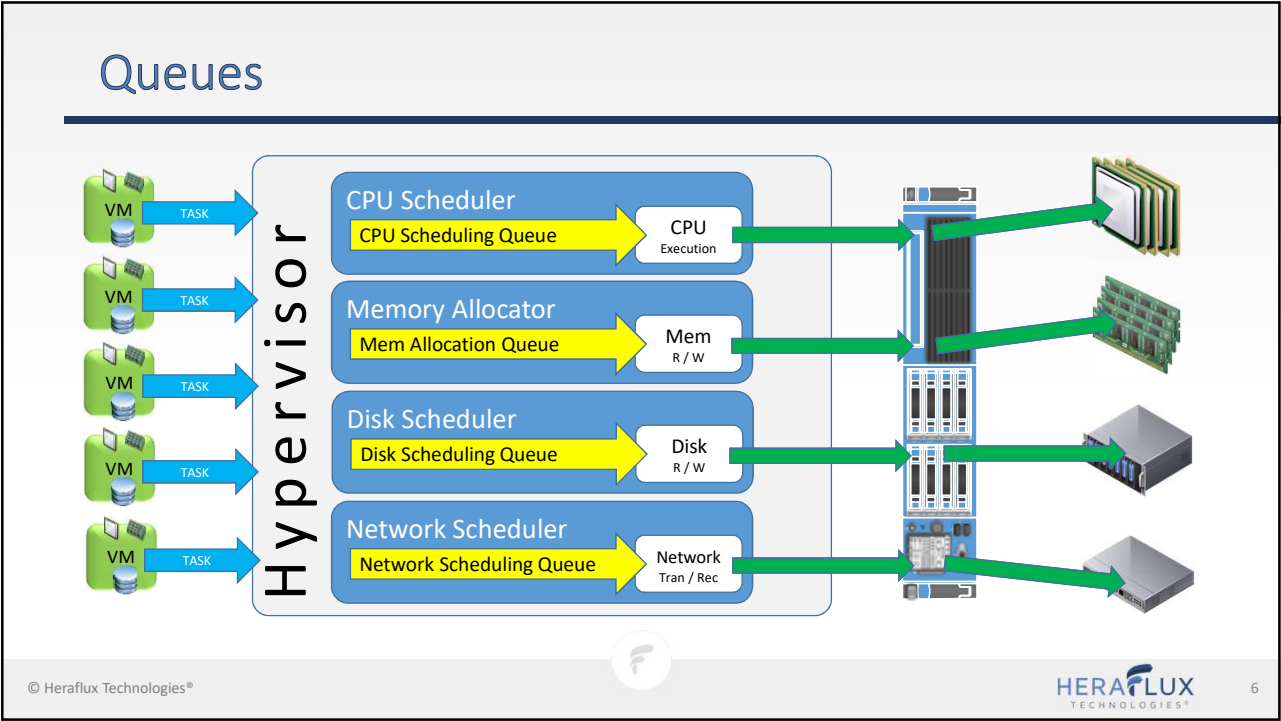
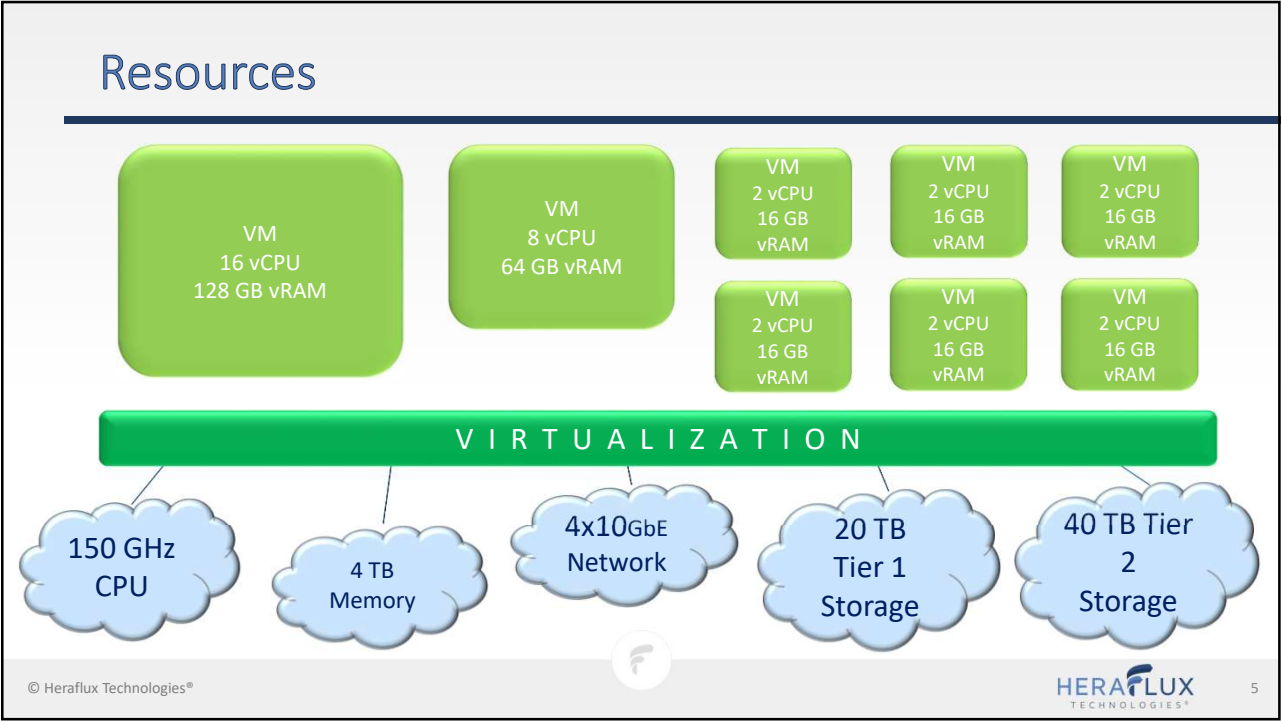
Founder & Chief Exorcist

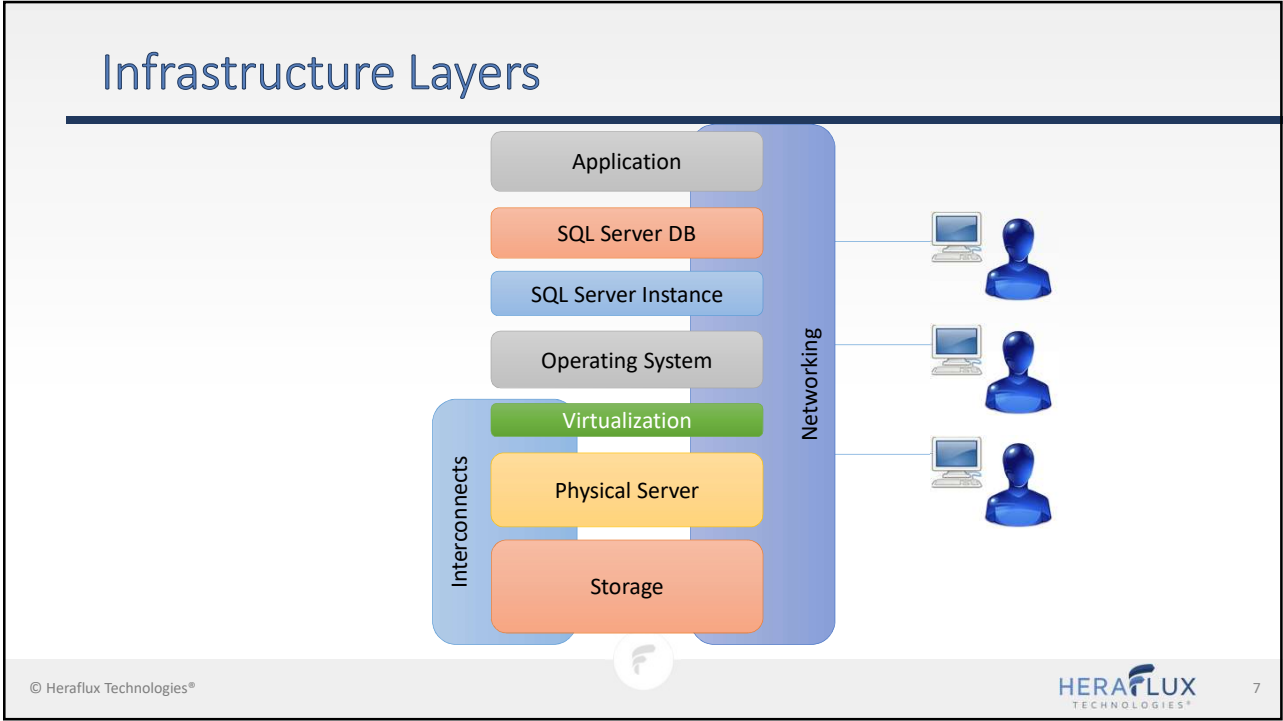
© Heraflux Technologies®



2







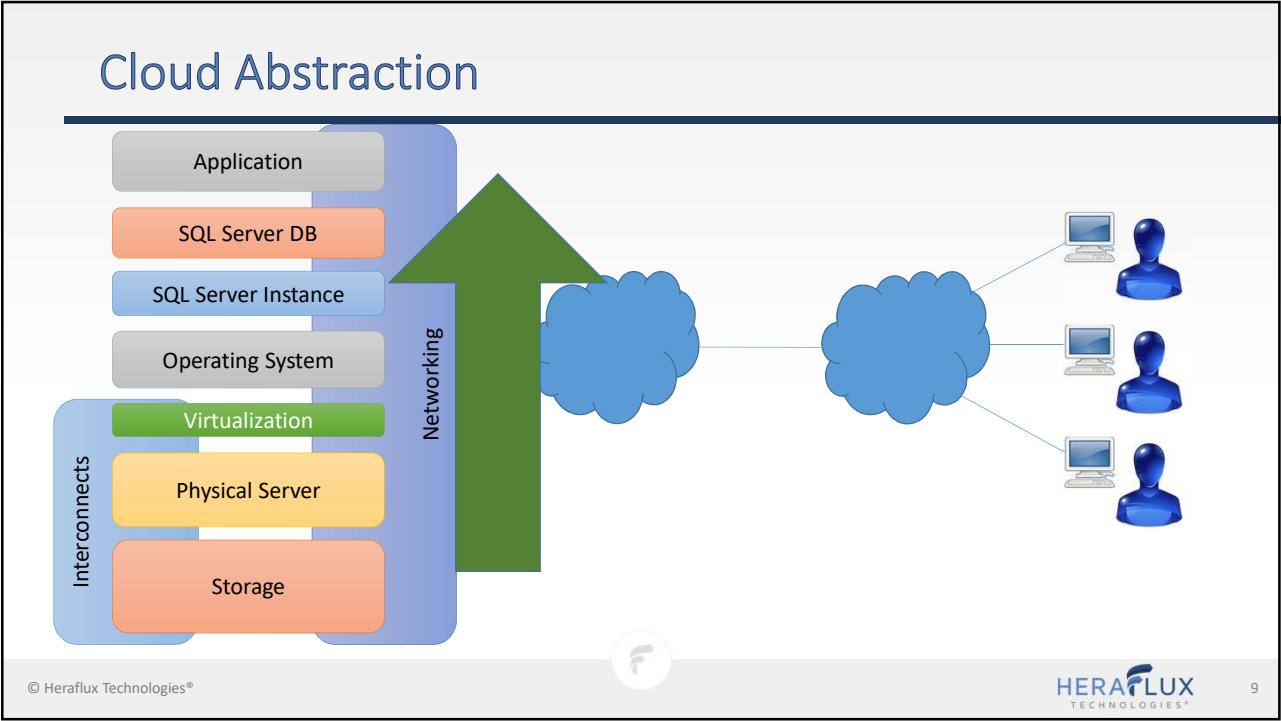
Your On-Prem VM vs. Cloud

The screenshot shows the Heraflux console interface. On the left, the 'Edit Settings' window for a VM named 'dc1' is open, showing configuration for CPU (2), Memory (8 GB), Hard disk 1 (60 GB), SCSI controller 0 (LSI Logic SAS), Network adapter 1 (VLAN 001 - Non-Lab Network), and CD/DVD drive 1 (Client Device). On the right, the 'Create virtual machine' window is open, showing a 'Choose a size' dialog. The dialog lists three recommended sizes: D4S V3 Standard (4 vCPUs, 16 GB, 8 data disks, 8000 Max IOPS, 32 GB Local SSD, Premium disk support, Load balancing, 313.97 USD/MONTH), E2S V3 Standard (2 vCPUs, 16 GB, 4 data disks, 4000 Max IOPS, 32 GB Local SSD, Premium disk support, Load balancing, 184.51 USD/MONTH), and D51 V2 Standard (1 vCPU, 3.5 GB, 2 data disks, 3200 Max IOPS, 7 GB Local SSD, Premium disk support, Load balancing, 91.51 USD/MONTH). The 'Select' button is highlighted.

© Heraflux Technologies®

HERAFLUX TECHNOLOGIES®

8



Storage

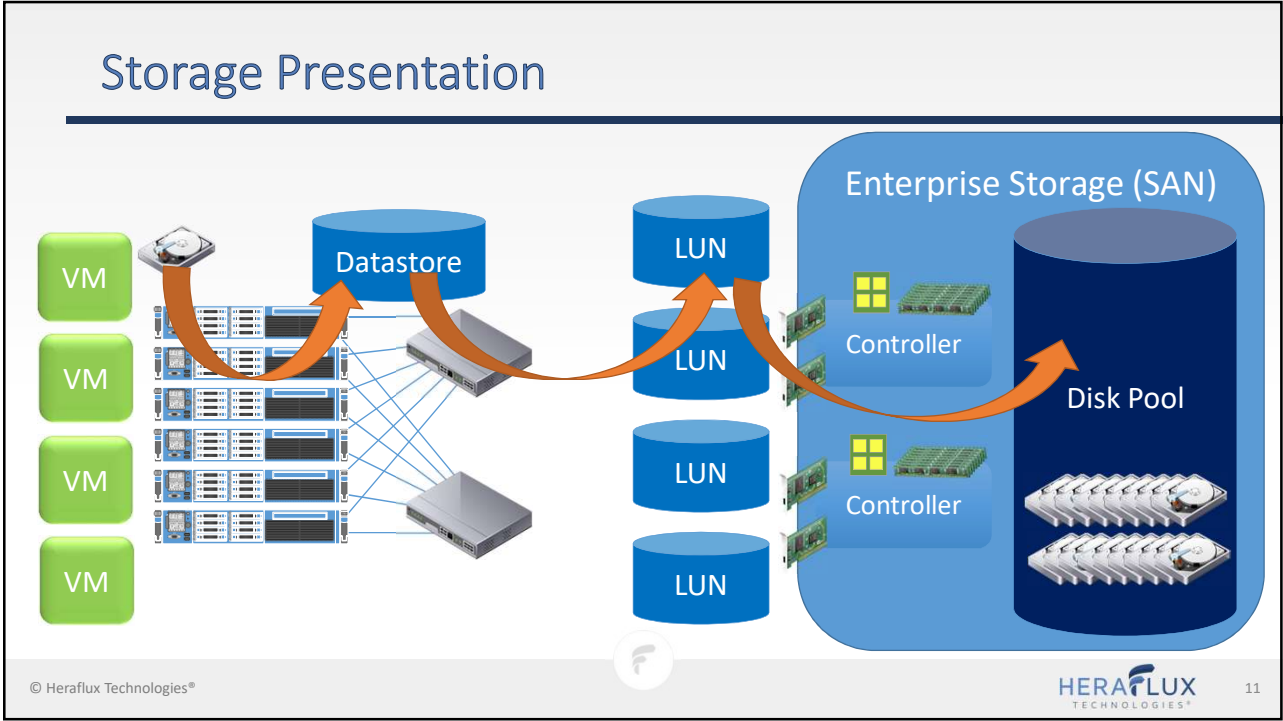
Your data has to be stored somewhere...

The diagram shows the physical server and storage layers of the architecture. It includes a yellow box for the Physical Server and an orange box for Storage, both connected to a larger blue cloud icon. This cloud is then connected to three server icons, each with a person icon next to it, representing users or clients.

© Heraflux Technologies®

HERAFLUX TECHNOLOGIES®

10



vDisk Layout

- C: - Operating System
- D: - SQL Server Instance Home
- E: - System Databases (master, model, msdb) *
- F: - User Database Data (1 of X)...
- G: - User Database Log (1 of Y)...
- H: - TempDB...
- Y: - Windows Page file **
- Z: - Backups

Distribute Load - 4 SCSI Controllers

Adjust as necessary
(but stay standardized)

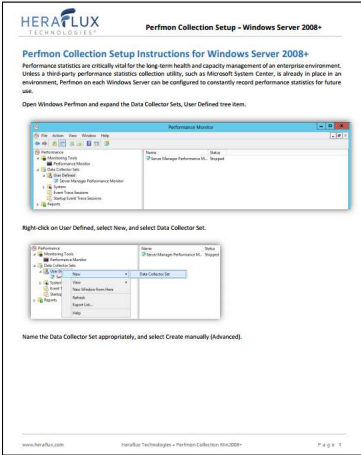
© Heraflux Technologies®

HERAFLUX TECHNOLOGIES®

12

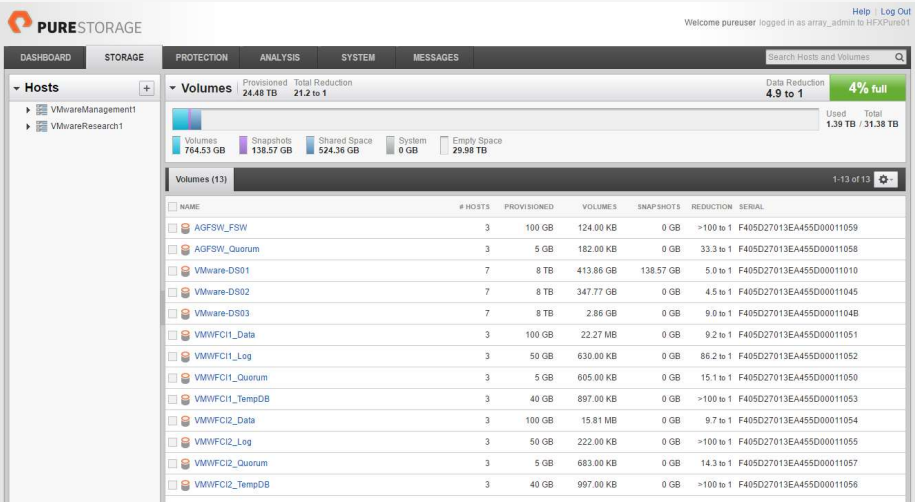
Windows Perfmon

- Collect raw data first
- Perfmon how-to guide
 - Download at hfxte.ch/perfmon
 - No strings attached
- 5-minute interval if routine
- 30-second interval if troubleshooting performance
- PowerShell for BLGtoSQL import:
 - hfxte.ch/perfmonposh



Storage Data Savings

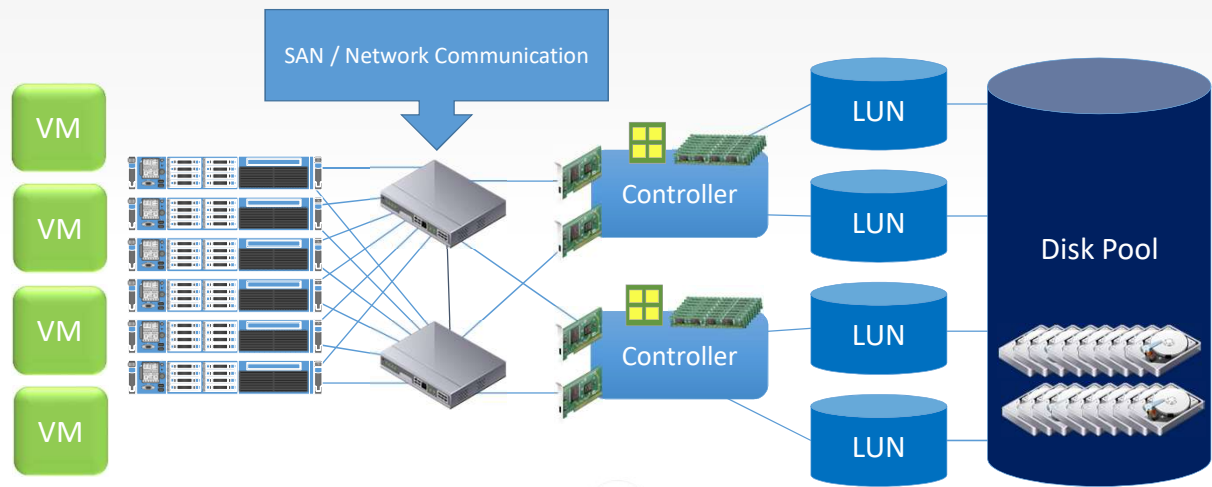
- Compression
- De-duplication
- Perf impact?
- Conflicts with SQL Server
 - Table-level compression
 - Encryption

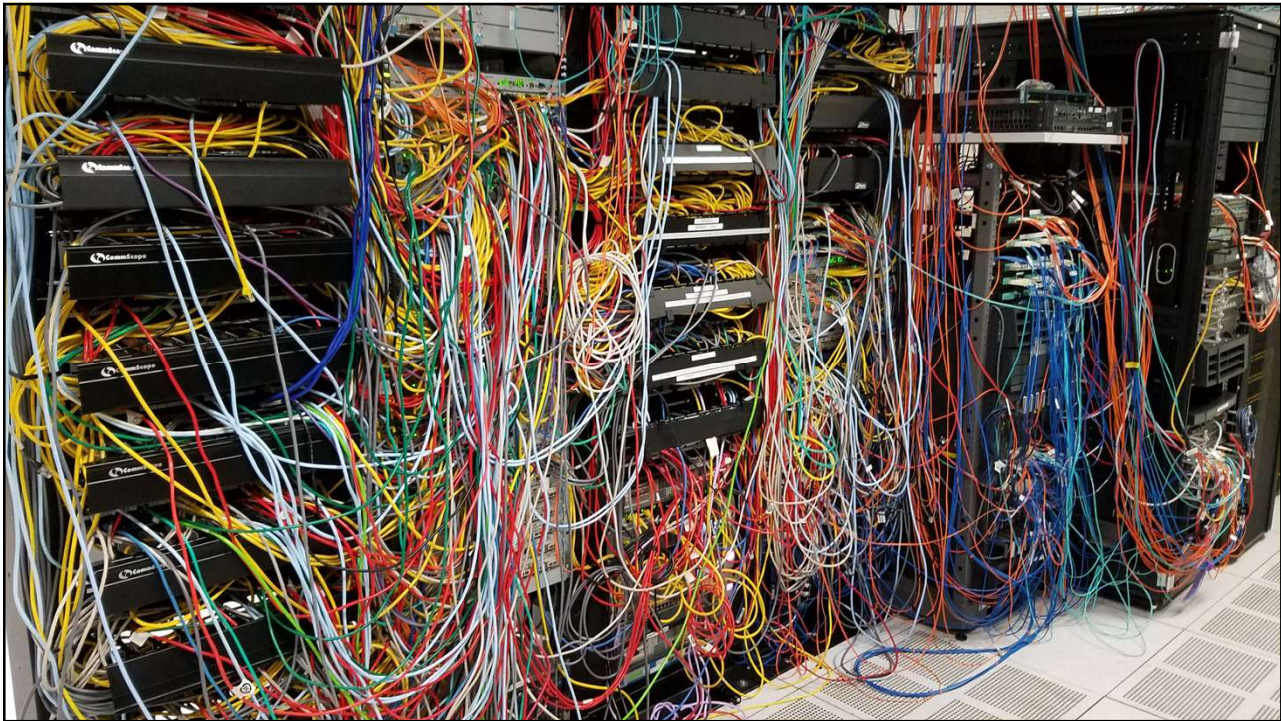


Networking



Communication Conduits





Network Throughput Testing

- Verify and test end-to-end performance
- Use iperf to test network throughput
 - Iperf.fr
- How-To Guide: hfxte.ch/iperf

Administrator: Command Prompt - iperf3 -s

```
C:\iperf>iperf3 -s
Server listening on 5201
```

Administrator: Command Prompt

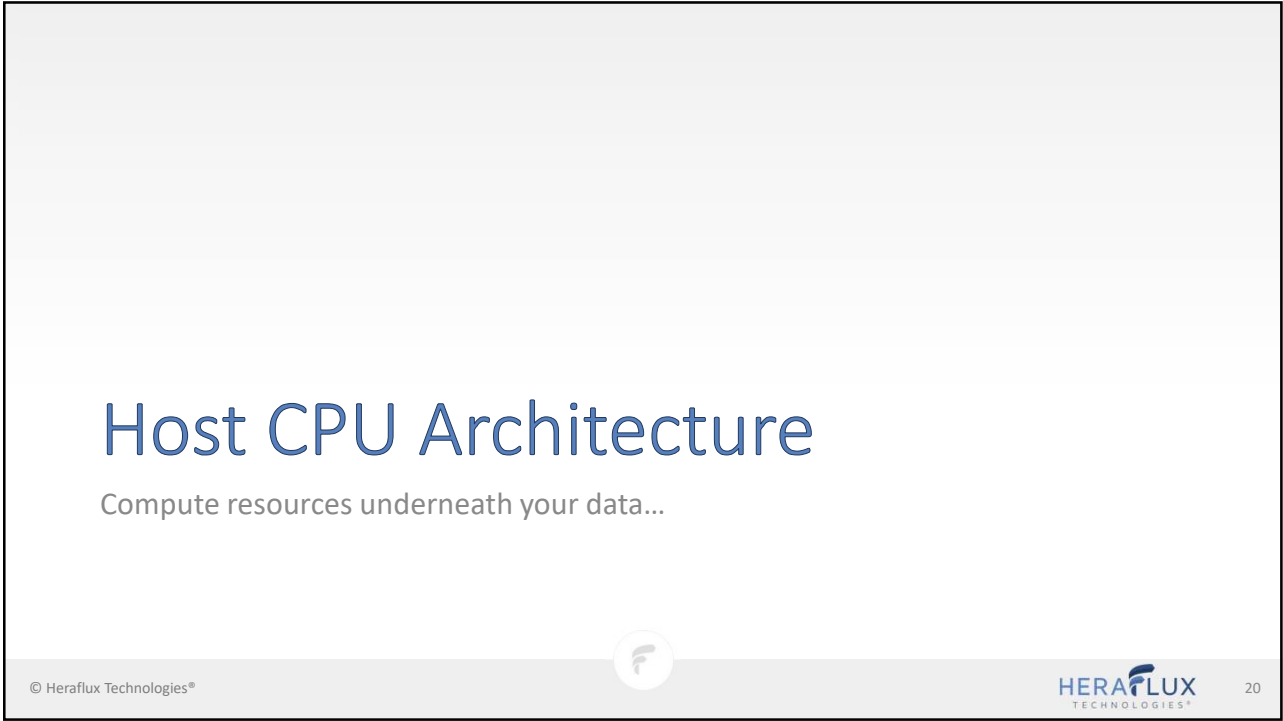
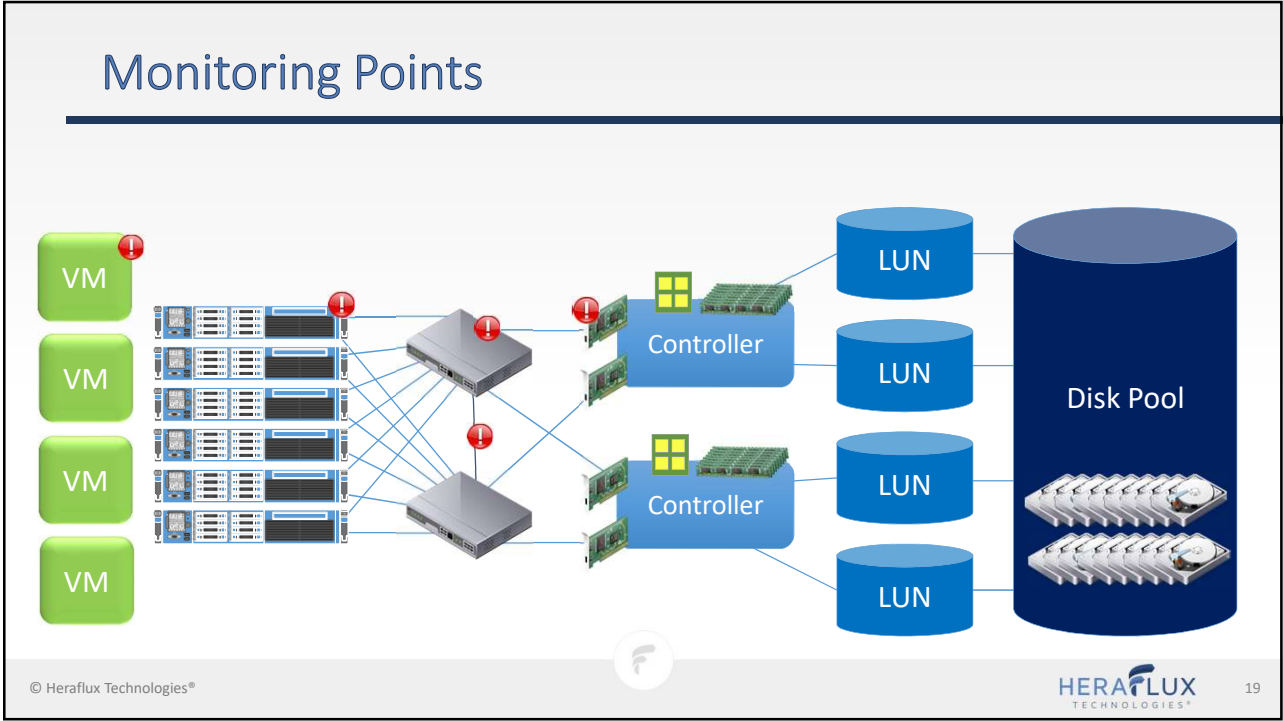
```
C:\iperf>iperf3 -c 10.100.50.62 -t 10 -P 10_
```

Administrator: Command Prompt

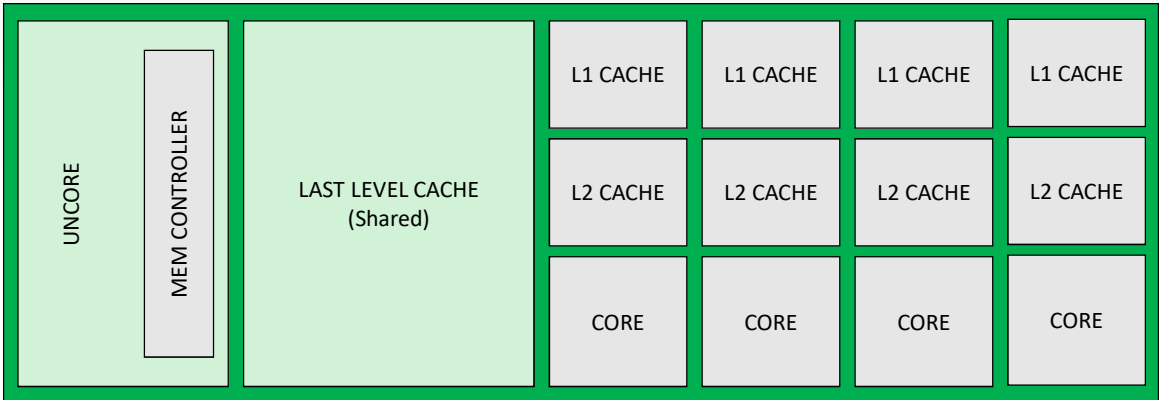
[4]	0.00-10.00	sec	1.03	GBytes	886	Mbits/sec	receiver
[6]	0.00-10.00	sec	1.06	GBytes	906	Mbits/sec	sender
[6]	0.00-10.00	sec	1.06	GBytes	906	Mbits/sec	receiver
[8]	0.00-10.00	sec	1.06	GBytes	915	Mbits/sec	sender
[8]	0.00-10.00	sec	1.06	GBytes	915	Mbits/sec	receiver
[10]	0.00-10.00	sec	1.14	GBytes	977	Mbits/sec	sender
[10]	0.00-10.00	sec	1.14	GBytes	977	Mbits/sec	receiver
[12]	0.00-10.00	sec	1.04	GBytes	896	Mbits/sec	sender
[12]	0.00-10.00	sec	1.04	GBytes	896	Mbits/sec	receiver
[14]	0.00-10.00	sec	1.03	GBytes	881	Mbits/sec	sender
[14]	0.00-10.00	sec	1.03	GBytes	881	Mbits/sec	receiver
[16]	0.00-10.00	sec	1.04	GBytes	893	Mbits/sec	sender
[16]	0.00-10.00	sec	1.04	GBytes	893	Mbits/sec	receiver
[18]	0.00-10.00	sec	1.11	GBytes	952	Mbits/sec	sender
[18]	0.00-10.00	sec	1.11	GBytes	952	Mbits/sec	receiver
[20]	0.00-10.00	sec	940	MBytes	789	Mbits/sec	sender
[20]	0.00-10.00	sec	940	MBytes	788	Mbits/sec	receiver
[22]	0.00-10.00	sec	910	MBytes	764	Mbits/sec	sender
[22]	0.00-10.00	sec	910	MBytes	764	Mbits/sec	receiver
[SUM]	0.00-10.00	sec	10.3	GBytes	8.86	Gbits/sec	sender
[SUM]	0.00-10.00	sec	10.3	GBytes	8.86	Gbits/sec	receiver

iperf Done.

```
C:\iperf>
```



CPU “Package”

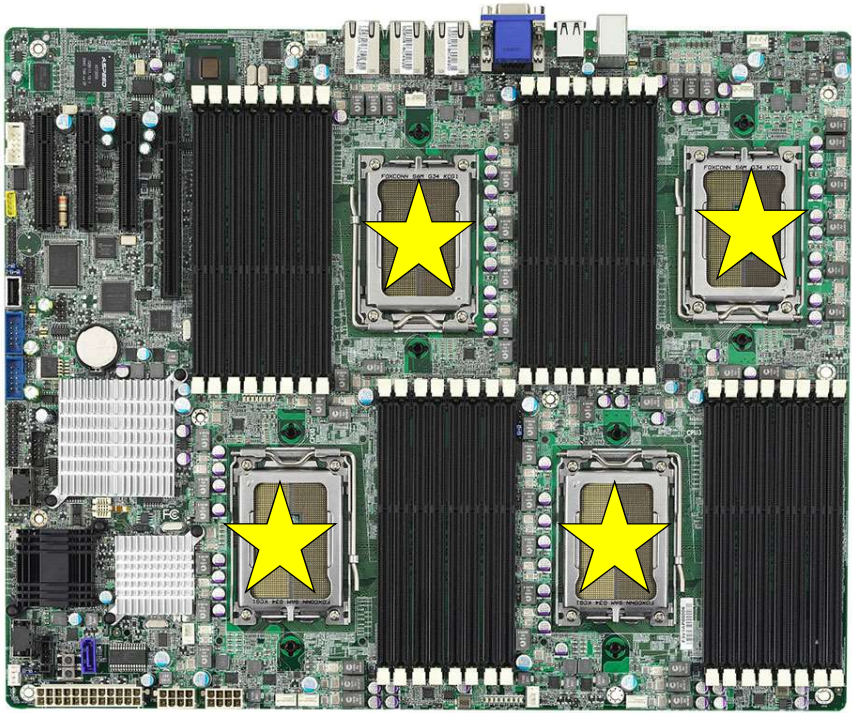


CPU Package Connectors



(Img src: <https://en.wikipedia.org/wiki/Xeon>)

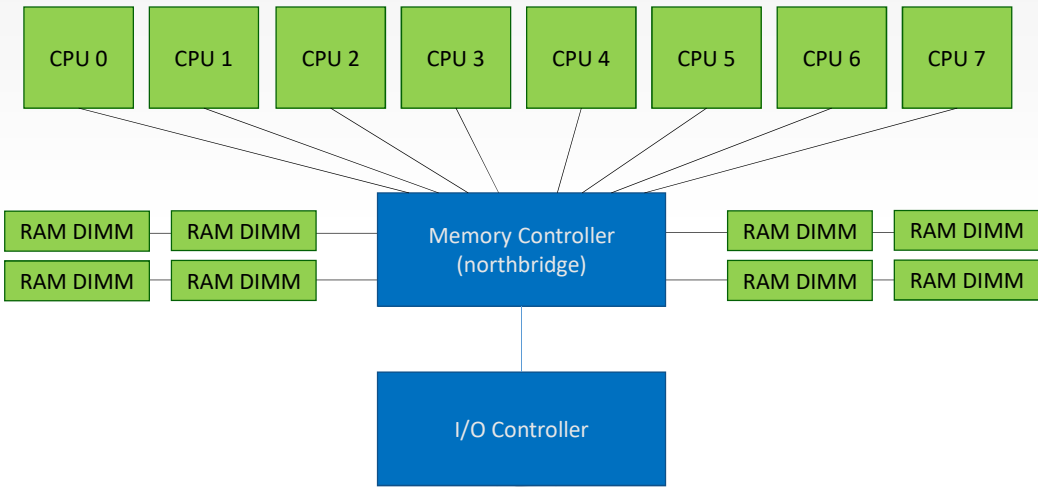
CPU Sockets



(Img src: <http://bit.ly/2tJU98k>)

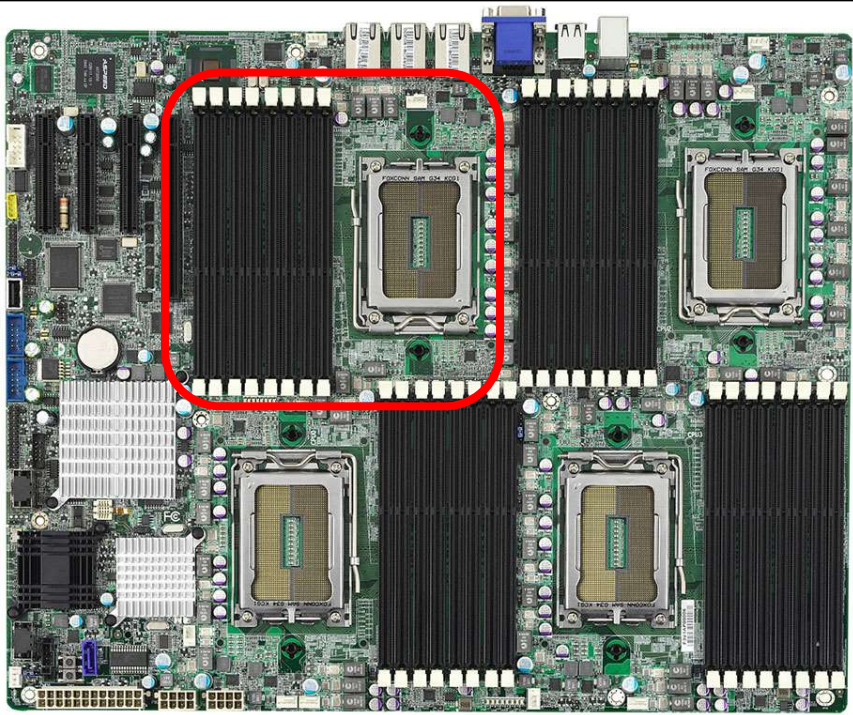
© Heraflux Technologies®

CPU UMA Architecture



© Heraflux Technologies®

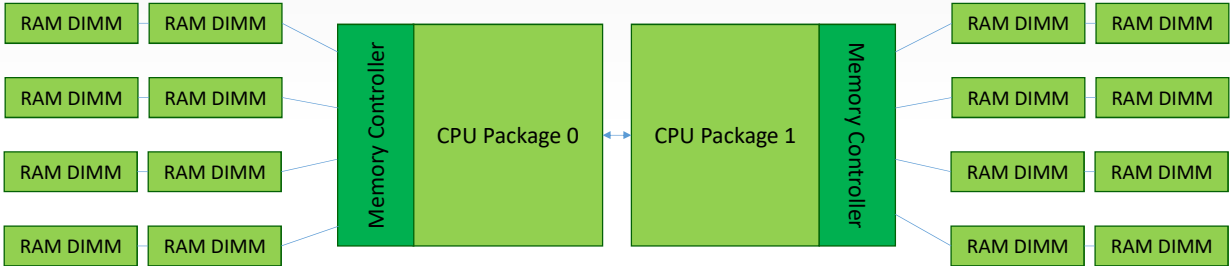
NUMA Nodes



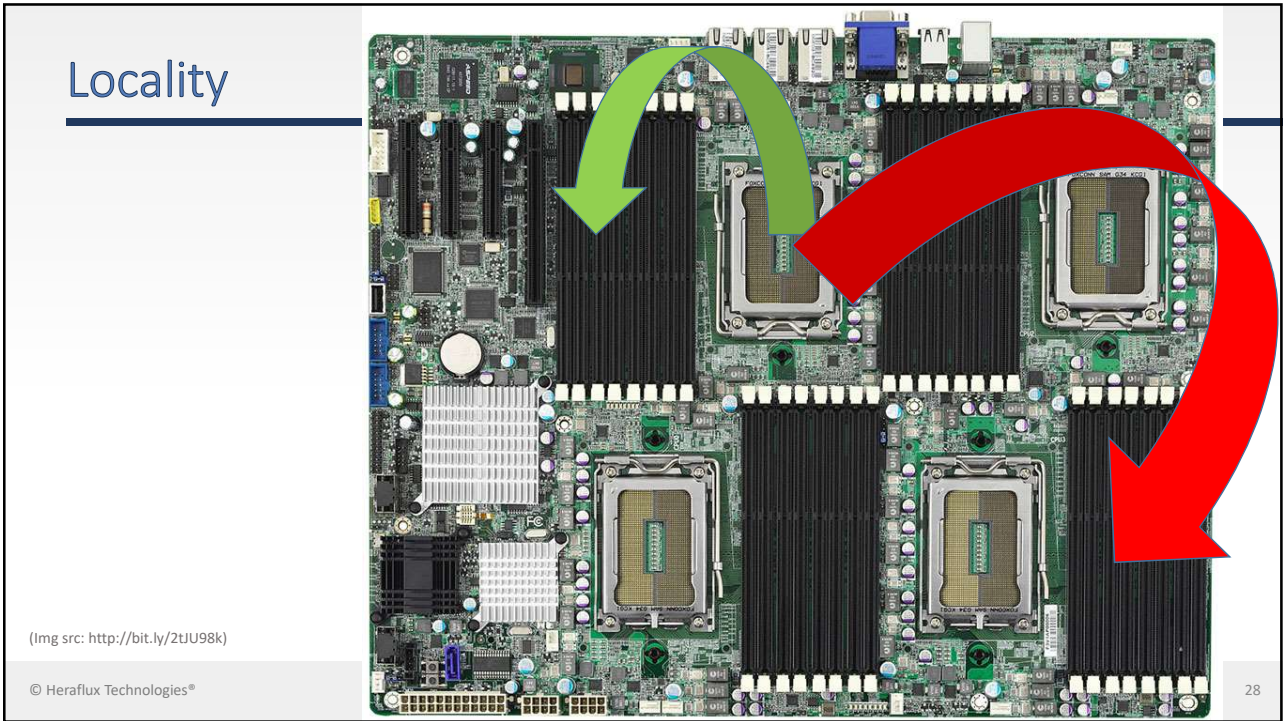
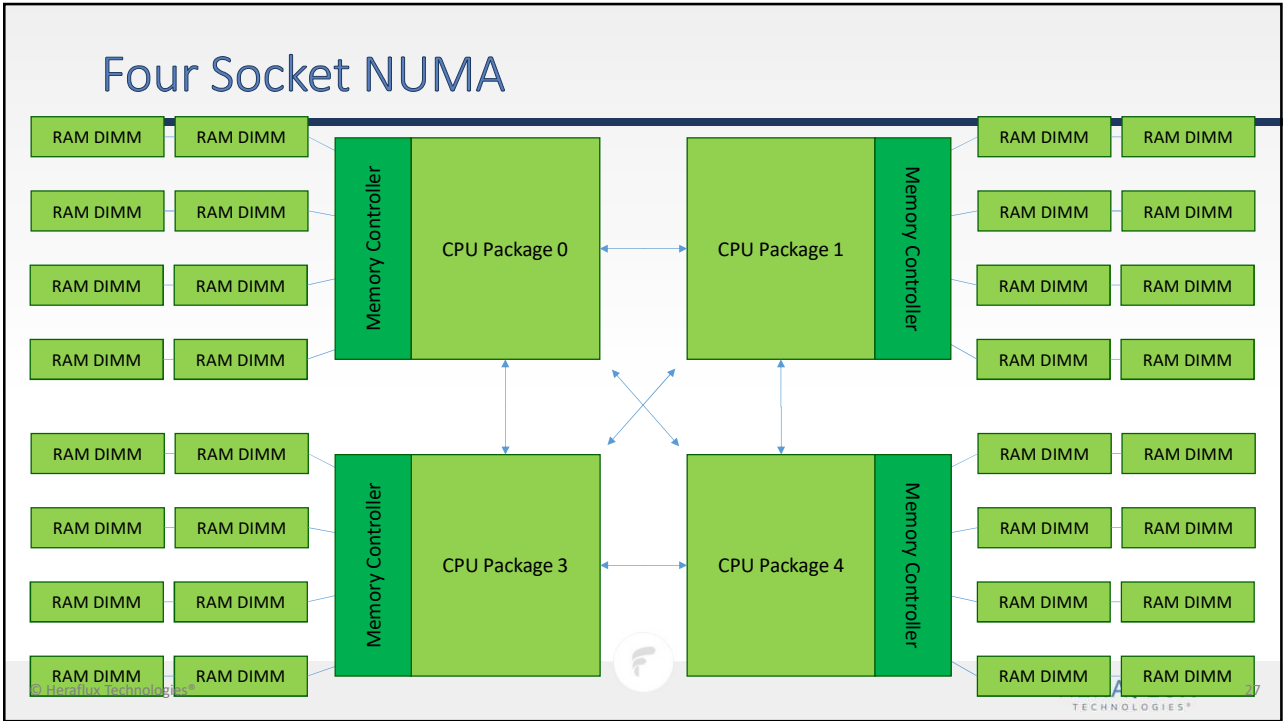
(Img src: <http://bit.ly/2tJU98k>)

© Heraflux Technologies®

CPU NUMA Architecture



© Heraflux Technologies®



Memory Speed

- Sample Cisco UCS memory config
- Slot & chip placement

Table 24 Recommended Memory Configurations for Intel Xeon E5-2600 v3 CPUs (with 2133-MHz DIMMs)¹

Total System Memory Size	CPU 1 DIMMs			CPU 2 DIMMs			DIMM Max Speed (MHz)	Total DIMMs
	Blue Slots Slot 1 (A1,B1,C1,D1)	Black Slots Slot 2 (A2,B2,C2,D2)	White Slots Slot 3 (A3,B3,C3,D3)	Blue Slots Slot 1 (E1,F1,G1,H1)	Black Slots Slot 2 (E2,F2,G2,H2)	White Slots Slot 3 (E3,F3,G3,H3)		
64 GB	4x8 GB	—	—	4x8 GB	—	—	2133	8
128 GB	4x8 GB	4x8 GB	—	4x8 GB	4x8 GB	—	2133	16
	4x16 GB	—	—	4x16 GB	—	—	2133	8
192 GB	4x8 GB	4x8 GB	4x8 GB	4x8 GB	4x8 GB	4x8 GB	1600	24
	4x16 GB ²	2x16 GB ²	— ²	4x16 GB ²	2x16 GB ²	— ²	2133	12
	4x16 GB	4x8 GB	—	4x16 GB	4x8 GB	—	2133	16
256 GB	4x16 GB	4x16 GB	—	4x16 GB	4x16 GB	—	2133	16
	4x32 GB	—	—	4x32 GB	—	—	2133	8
384 GB	4x16 GB	4x16 GB	4x16 GB	4x16 GB	4x16 GB	4x16 GB	1866	24
	4x32GB	4x16GB	—	4x32GB	4x16GB	—	2133	16
512 GB	4x32 GB	4x32 GB	—	4x32 GB	4x32 GB	—	2133	16
	4x64 GB	—	—	4x64 GB	—	—	2133	8
768 GB	4x32 GB	4x32 GB	4x32 GB	4x32 GB	4x32 GB	4x32 GB	1866	24
1024 GB	4x64 GB	4x64 GB	—	4x64 GB	4x64 GB	—	2133	16
1536 GB	4x64 GB	4x64 GB	4x64 GB	4x64 GB	4x64 GB	4x64 GB	1600	24

© Heraflux Technologies®
(Src: <http://www.cisco.com/c/dam/en/us/products/collateral/servers-unified-computing/ucs-b-series-blade-servers/b200m4-specsheet.pdf> page 44)

29

Determine vCPU Count

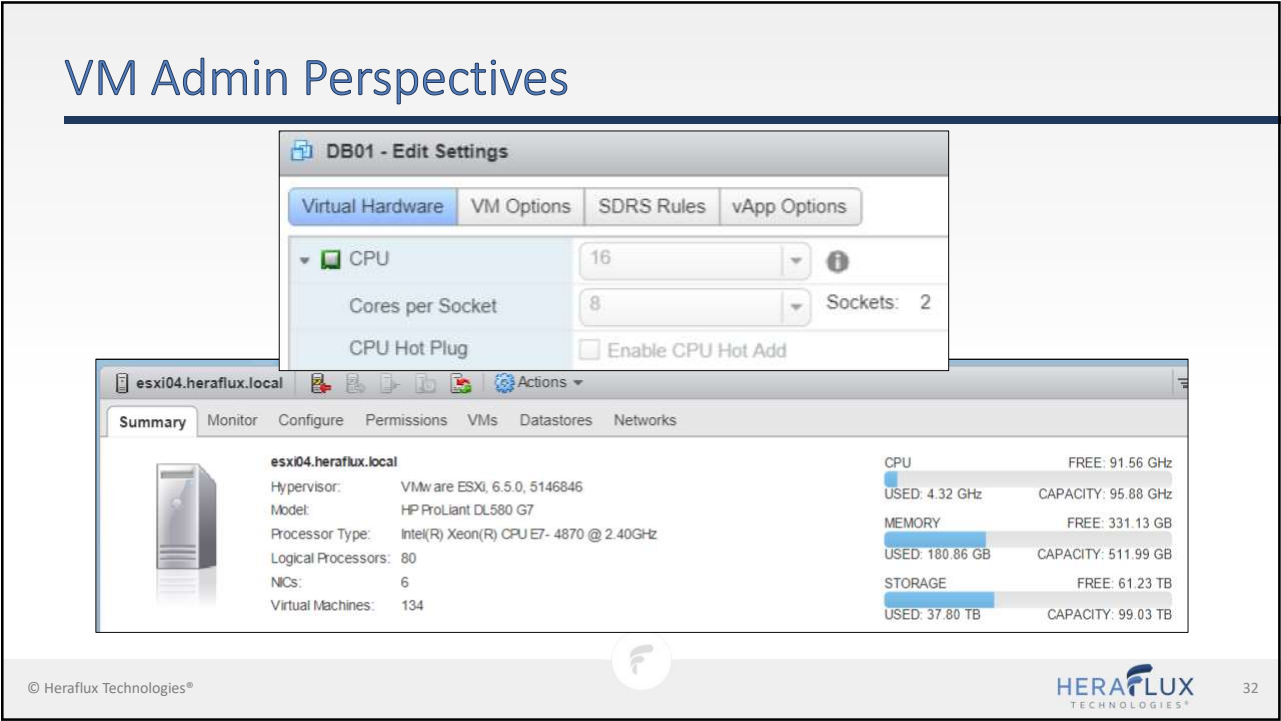
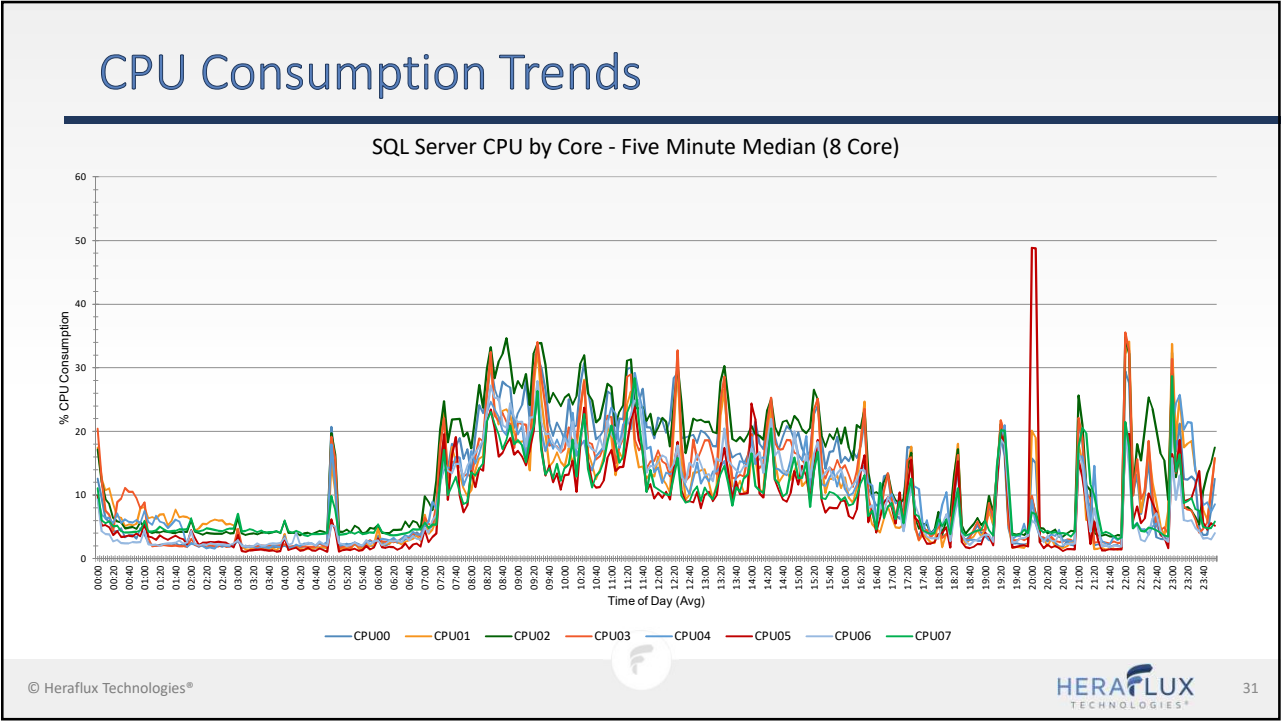
How Many Do You Need?

- “Right-sizing” analysis
- Ongoing performance baseline
- Size for now, not future
- Want target CPU utilization 40-60% during routine business operations
- Leave headroom for short-term growth
- Resize VM as necessary

Task Manager Performance tab showing CPU usage at 100% (2.58 GHz) and a list of Ethernet adapters. The CPU section shows 100% utilization across 144 logical processors. The Ethernet section shows various network adapters with their speeds and utilization.

© Heraflux Technologies®

30



What You Can See

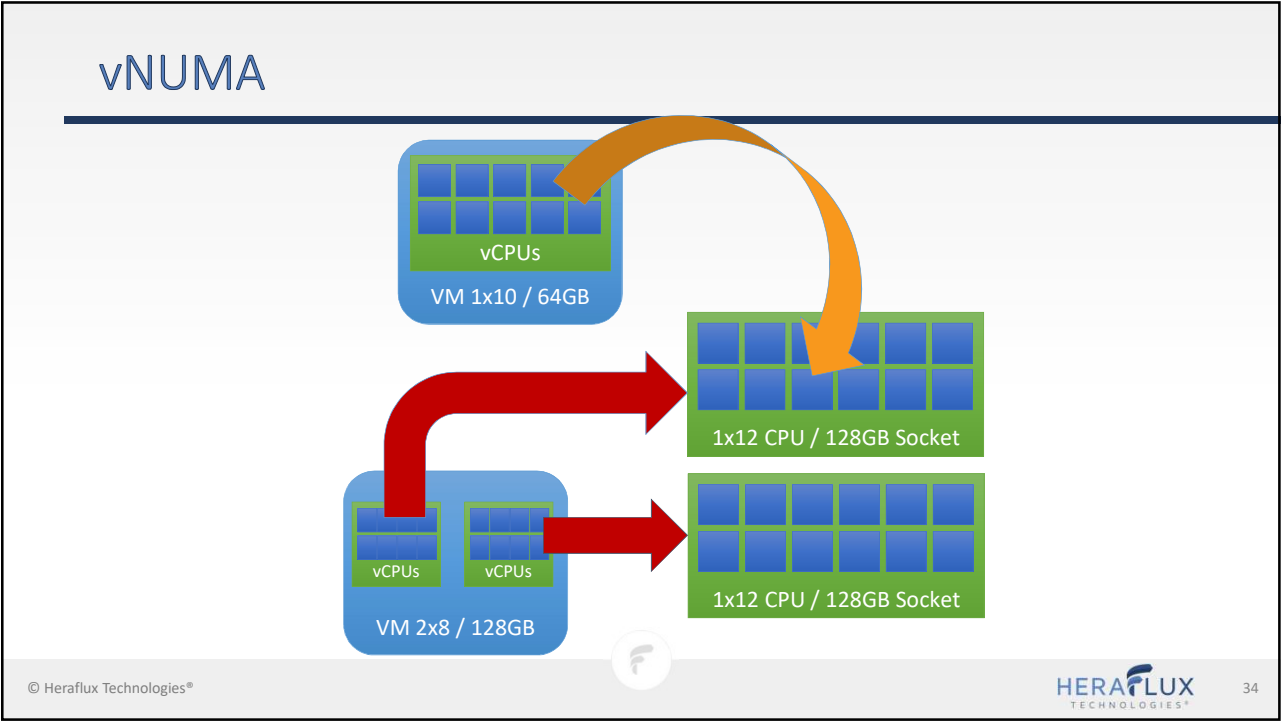
Processor	Processor Affinity	I/O Affinity
ALL	☒	☒
NumaNode0	☐	☐
NumaNode1	☐	☐
CPU8	☐	☐
CPU9	☐	☐
CPU10	☐	☐
CPU11	☐	☐
CPU12	☐	☐

Utilization: 4% Speed: 2.40 GHz Maximum speed: 2.40 GHz
Processes: 70 Threads: 1916 Handles: 31968 Sockets: 2
Virtual processors: 16 Virtual machine: Yes L1 cache: N/A

© Heraflux Technologies®

HERAFLUX TECHNOLOGIES®

33



Verify vCPU Presentation

MS CoreInfo

<http://bit.ly/1SKNcWL>

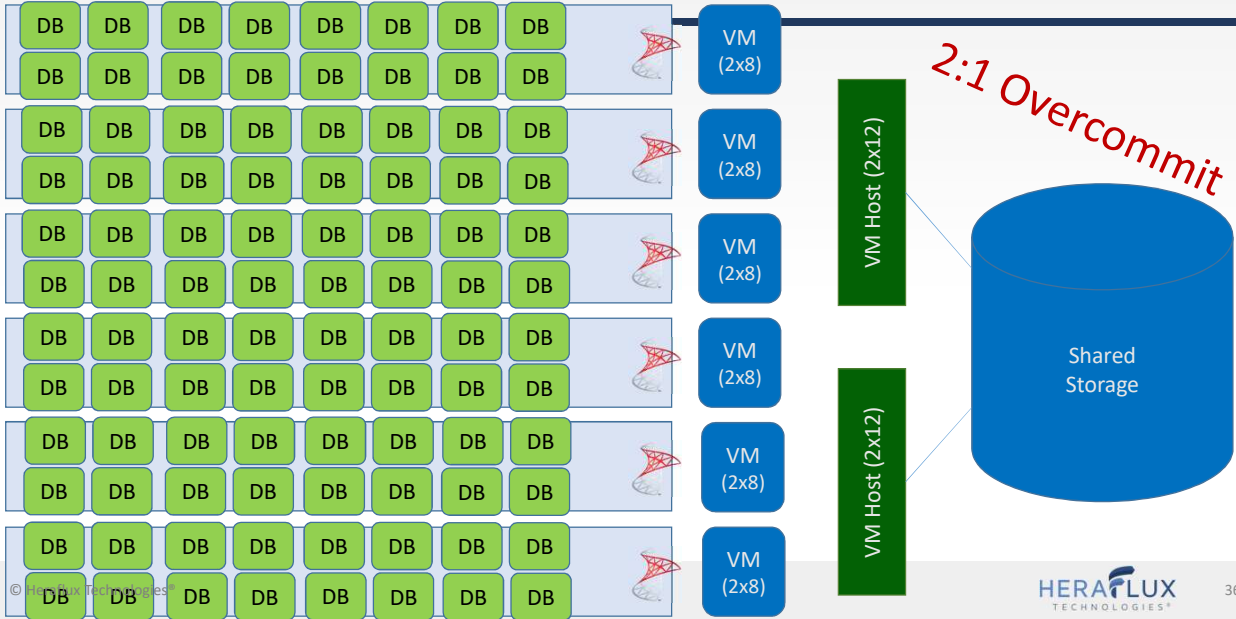
```
Logical Processor to Socket Map:
*****----- Socket 0
-----***** Socket 1

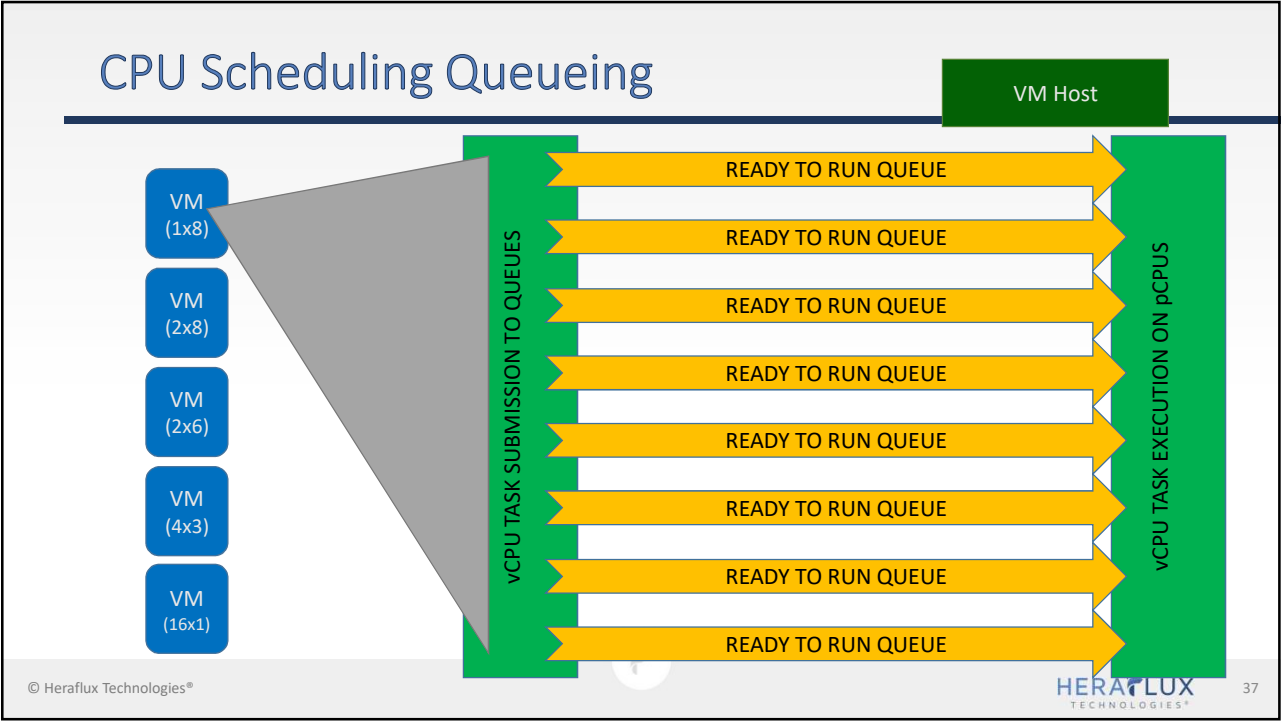
Logical Processor to NUMA Node Map:
*****----- NUMA Node 0
-----***** NUMA Node 1
```

```
Logical to Physical Processor Map:
*----- Physical Processor 0
*----- Physical Processor 1
*----- Physical Processor 2
*----- Physical Processor 3
*----- Physical Processor 4
*----- Physical Processor 5
*----- Physical Processor 6
*----- Physical Processor 7
*----- Physical Processor 8
*----- Physical Processor 9
*----- Physical Processor 10
*----- Physical Processor 11
*----- Physical Processor 12
```

```
Approximate Cross-NUMA Node Access Cost (relative to fastest):
  00 01
00: 1.2 1.4
01: 1.4 1.0
```

vCPU Overcommit





Scheduling Trouble Measurement

Hyper-V – Wait Time Per Dispatch

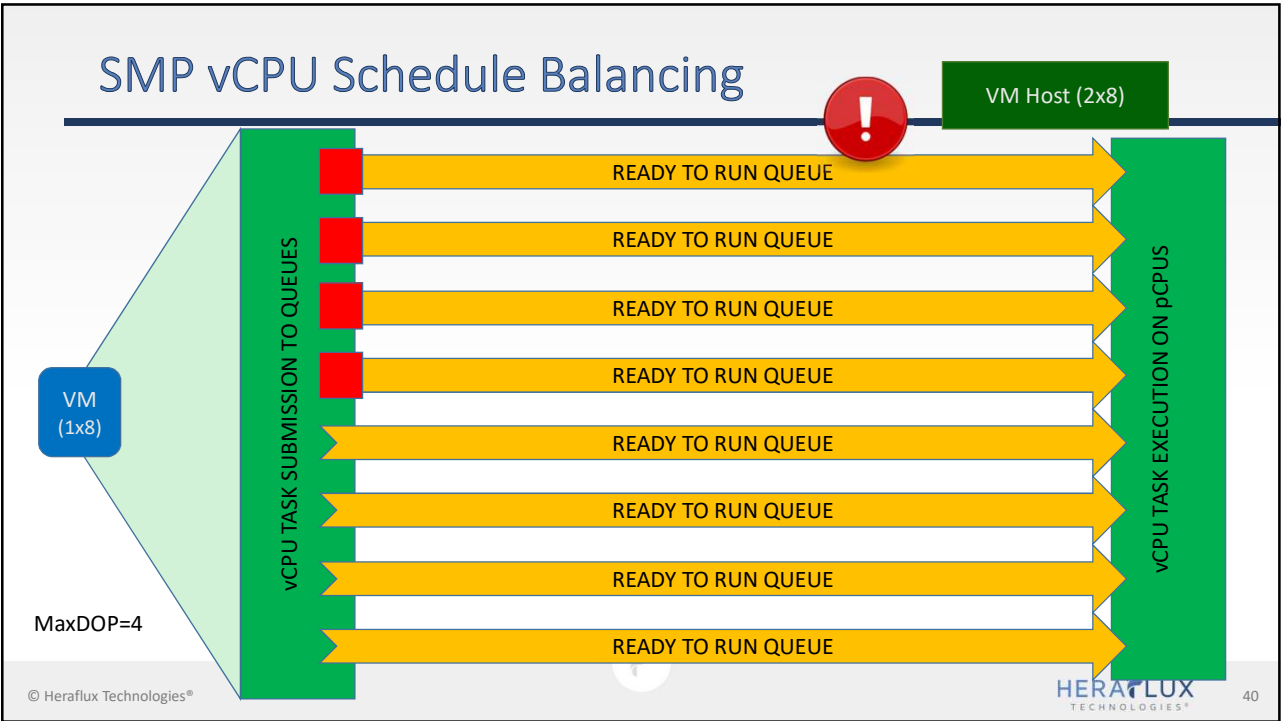
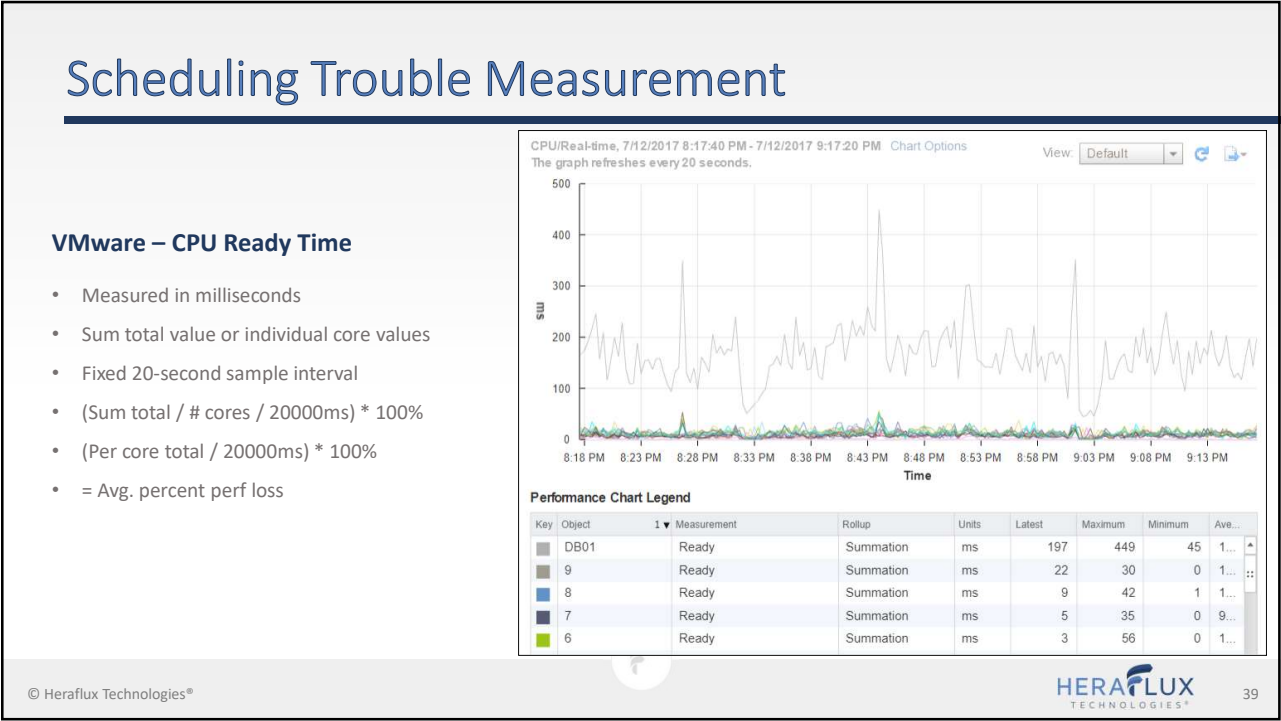
- Measured in nanoseconds
- Average value or individual core values
- Sample interval (one second)
- Avg. over collection interval (X seconds)
- (Metric value / sample interval total nanoseconds) * 100%
- = Avg. percent perf loss

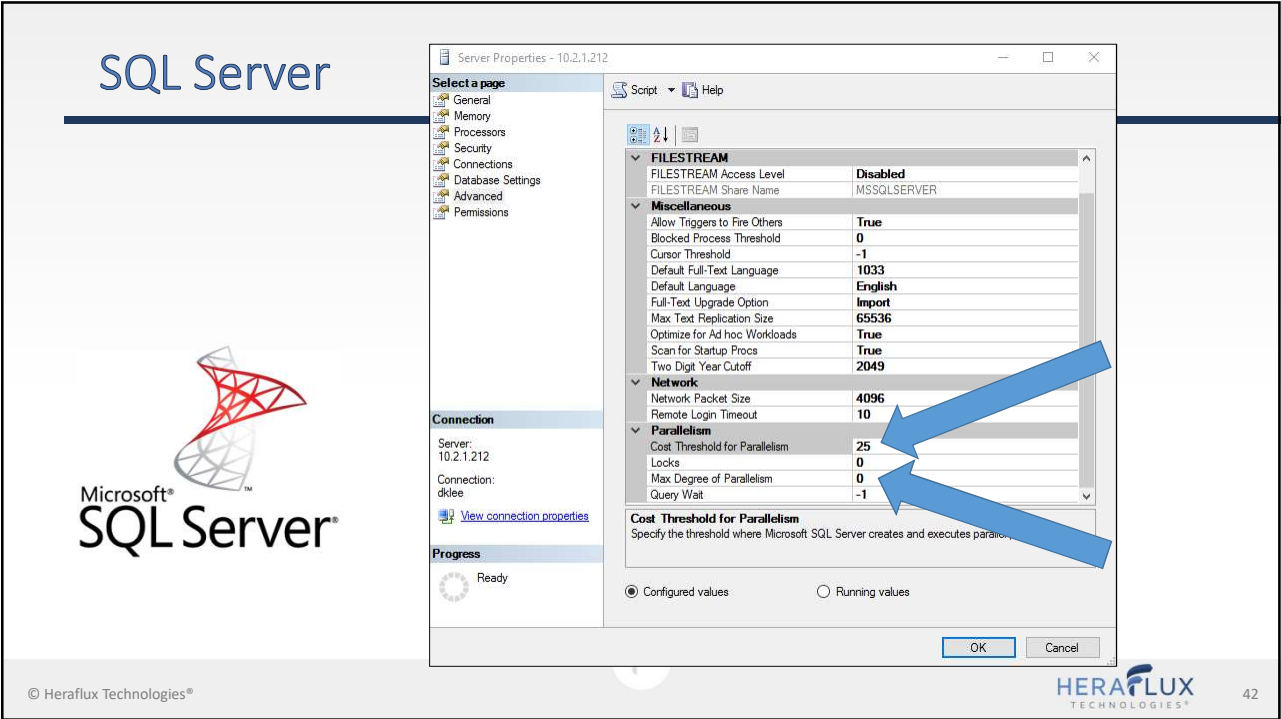
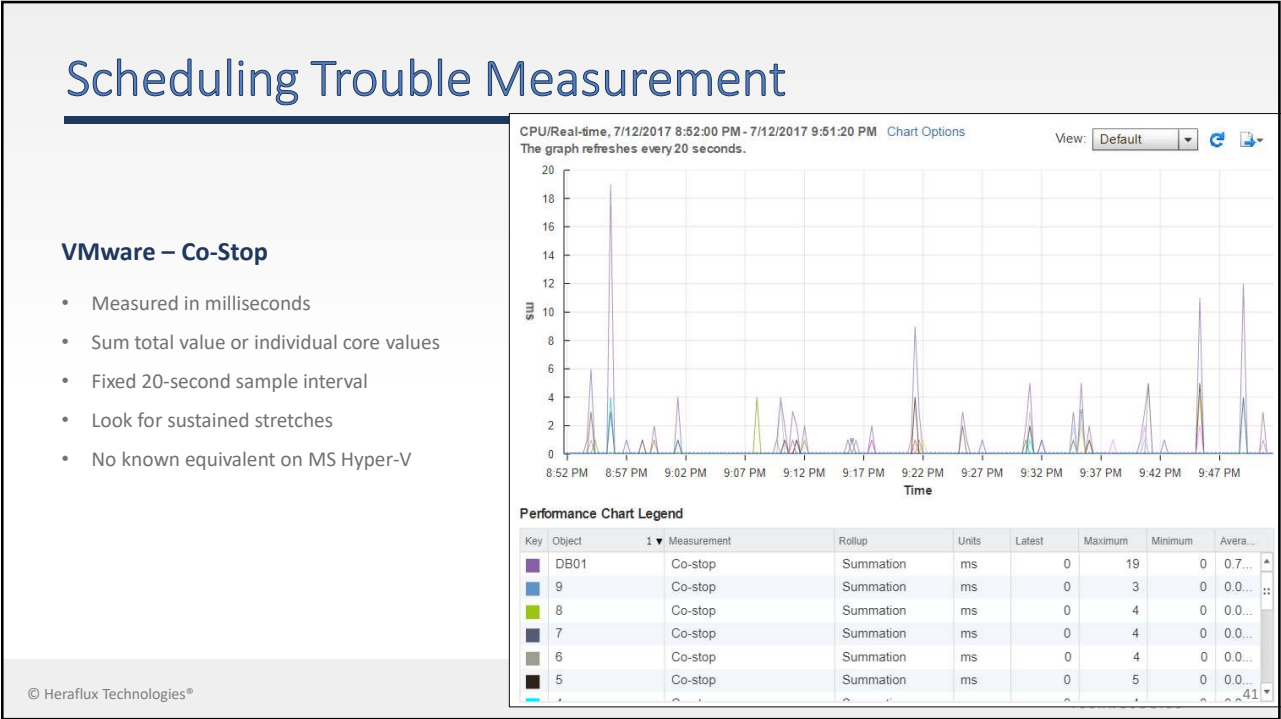
SetName	CounterName	InstanceName	CollectionDT	CounterValue
hyper-v hypervisor virtual processor	cpu wait time per dispatch	_total	2016-11-28 00:00:00.000	25749.88
hyper-v hypervisor virtual processor	cpu wait time per dispatch	_total	2016-11-28 00:15:00.000	25599.74
hyper-v hypervisor virtual processor	cpu wait time per dispatch	_total	2016-11-28 00:30:00.000	25549.12
hyper-v hypervisor virtual processor	cpu wait time per dispatch	_total	2016-11-28 00:45:00.000	25276.59
hyper-v hypervisor virtual processor	cpu wait time per dispatch	_total	2016-11-28 01:00:00.000	22561.70
hyper-v hypervisor virtual processor	cpu wait time per dispatch	_total	2016-11-28 01:15:00.000	22501.42
hyper-v hypervisor virtual processor	cpu wait time per dispatch	_total	2016-11-28 01:30:00.000	22388.80
hyper-v hypervisor virtual processor	cpu wait time per dispatch	_total	2016-11-28 01:45:00.000	22157.82
hyper-v hypervisor virtual processor	cpu wait time per dispatch	_total	2016-11-28 02:00:00.000	20973.48
hyper-v hypervisor virtual processor	cpu wait time per dispatch	_total	2016-11-28 02:15:00.000	20568.92
hyper-v hypervisor virtual processor	cpu wait time per dispatch	_total	2016-11-28 02:30:00.000	22334.37
hyper-v hypervisor virtual processor	cpu wait time per dispatch	_total	2016-11-28 02:45:00.000	22390.42
hyper-v hypervisor virtual processor	cpu wait time per dispatch	_total	2016-11-28 03:00:00.000	21617.33
hyper-v hypervisor virtual processor	cpu wait time per dispatch	_total	2016-11-28 03:15:00.000	21208.31
hyper-v hypervisor virtual processor	cpu wait time per dispatch	_total	2016-11-28 03:30:00.000	16428.44
hyper-v hypervisor virtual processor	cpu wait time per dispatch	_total	2016-11-28 03:45:00.000	14421.77

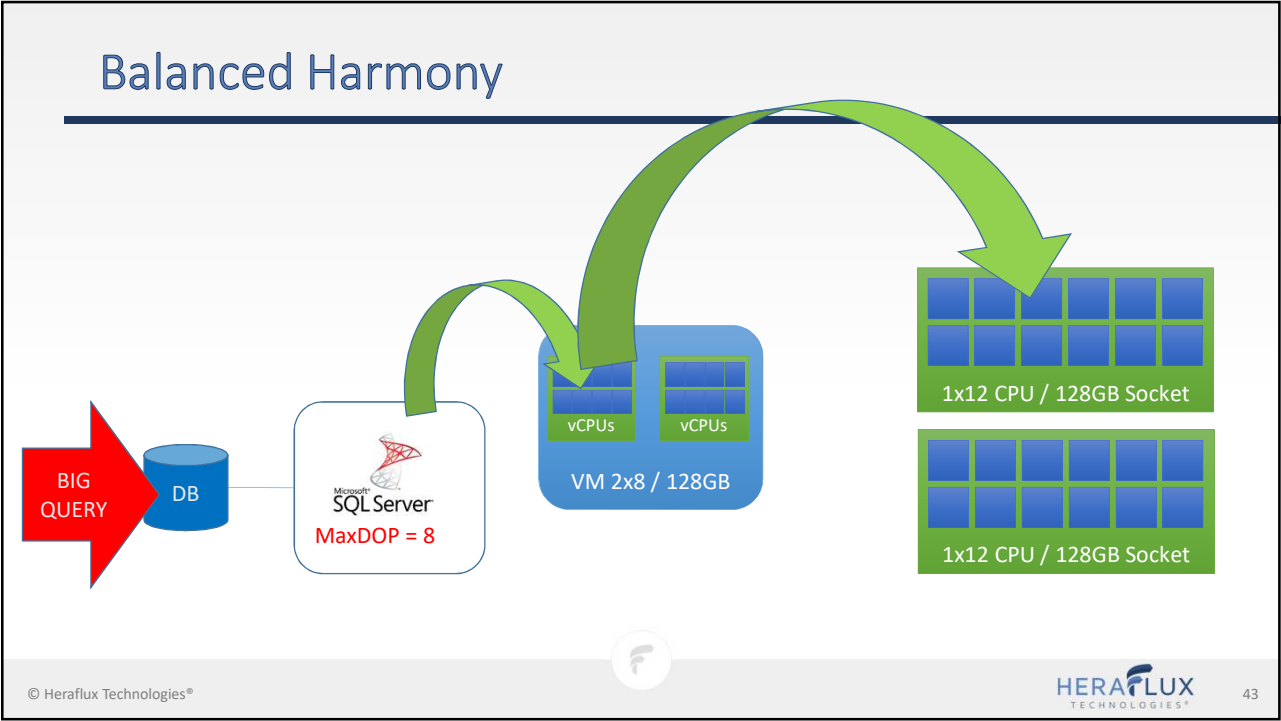
© Heraflux Technologies®

HERAFLUX TECHNOLOGIES®

38







Conclusions

- Virtualization is new beige
- Virtualization is first step to cloud
- Equivalent performance if done right
- vCPU / vNUMA configuration
- vDisk configuration
- SQL Server parallelization settings



© Heraflux Technologies®

HERAFLUX TECHNOLOGIES®

44

Questions?

 @kleegeek
 davidklee.net
 heraflux.com
 linkedin.com/in/davidaklee

© Heraflux Technologies®



HERAFLUX
TECHNOLOGIES®

45

Thanks for attending!

HERAFLUX
TECHNOLOGIES®