The IT Performance Management & Capacity Planning Company

Customer VMWare

Performance Check

Author: Walter Verhoeven

Doc. Ref.: customer_vmware_perf_check.odt

Version: 1.0

Date: 27 April 2017

I. Document History

Version	Date	Name	Description
1.0	27/04/2017	W. Verhoeven	First Copy

II. Stakeholders

Company	Responsible Stakeholder
customer	Mr. X
customer	Mr. Y
CREATIVE Associates	Walter Verhoeven

CREATIVE Associates

Dondervelstraat 52
1651 Lot/Beersel
Belgium

III. Contents

Table of Contents

1.MANAGEMENT SUMMARY....4
1.1Introduction....4
2.RECOMMENDATIONS & REMARKS....5
2.1In General....5
2.1Attention Points....5
3.OVERALL VMWARE INFRASTRUCTURE....7
3.1Cluster Level Changes....8
4.CPU RESOURCE....9
4.1Cluster Level....9
4.2Resource Pool....9
4.1ESX Host Level....10
4.1ESX VM Level....14
5.MEMORY RESOURCE....22
5.1Cluster Level....22
5.1ESX Host Level....22
5.1ESX VM Level....24
6.I/O RESOURCE....27
6.1ESX Host Level....27
6.1ESX VM Level....30
7.NETWORK RESOURCE....31
7.1ESX Host Level....31
7.1ESX VM Level....33

1. Management Summary

1.1 Introduction

At customer a VMware infrastructure has been set-up for internal company use. It is a similar solution as deployed at customer sites.

customer internal staff and customers complain about intermittent slowness of interactive applications.

As such we have been asked to provide figures for the performance behaviour of this environment and possibly a root cause too. In order to accomplish this we deployed our TeamQuest Vityl Monitor solution on March the 10th and collected data until March the 30th. The smallest data point granularity is 5 minutes.

The data at customer has been stored in a PostgreSQL database which has been exported and loaded into our lab for further off-line analysis.

At the same time this exercise is used to check if our performance solution Vityl Monitor provides sufficient metrics/detail in order to investigate this issue.

We asked a date timed list of issues and VM system names for correlation purposes but did not receive anything until today.

So this document provides information in regards to the performance behaviour of the customer VMware infrastructure for the period 10th to 30th March/17 and some recommendations & remarks.

2. Recommendations & Remarks

2.1 In General

After looking at this VMware performance data for the period of 10th to 30th March/17 we can state that the LAN, I/O Subsystem and Memory do not introduce a performance issue into the current infrastructure due to the following elements:

- The ESX Host CPU resources are never used more than 22% so sufficient power. There are about 124 VM's with +/-42 per ESX Host.
- ESX-1/2/3 have the same chip type installed, 2 chips with each 12 cores and 2 threads if enabled. So ESX-1/2 have each 48 LCPU's while ESX-2 only 24 as the multi thread option has been disabled.
- The VM's show a medium vCPU load with peaks up to 60%.
- The VM's are not blocked due to CPU limitations.
- The co-vCPU time is very low meaning that the VM's are not blocked on getting their second vCPU.
- The%vcpu_ready activity is also small so no VM's blocked on getting a physical CPU (core).
- The LAN bandwidth utilization remains below the 1Gbit theoretical limitation per nic. Overall LAN activity is quite small.
- There where no LAN packets dropped nor errors nor frames with unknown protocols received.
- The I/O subsystem shows, apart from some sporadic peaks, a latency around 5 MSec which is excellent.
- Sufficient free memory as the maximum utilization is about 60% on the ESX Hosts. The total installed memory is about 727 Gbytes which is 255,87 GBytes per ESX Host.
- The VM's consume 30% to 40% of their memory with some peaks up to 90%.
- No "swapin", "swapout", "swapwait" (The amount of time in seconds the virtual machine was waiting for swap activity), was measured.

2.1 Attention Points

Although in general the infrastructure looks fine we saw some strange behaviour that must be investigated further:

- The I/O subsystem on ESX-01 shows an increased maxTotalLatency as of March the 20th. In general the maxTotalLatency is higher when comparing with to the two other hosts ESX-02 & ESX-03.
- The CPU latency, representing the percentage of time the virtual machine was ready to run but was not scheduled to run because of physical CPU resource contention, is quite high. In fact it contains "cpu ready", "swapwait" and "power regulation (C-state)". So If your "cpu ready" is ok, and "swap wait" is 0, then it is the power regulation. So we advise to do the following in order to try to decrease it:

ESXi 6.0 offers the following power policy options

- High performance
  - This power policy maximizes performance, using no power management features.
- Balanced
  - This power policy (the default in ESXi 6.0) is designed to reduce host power consumption while having little or no impact on performance.
- Low power
  - This power policy is designed to more aggressively reduce host power consumption at the risk of reduced performance.
- Custom
  - This power policy starts out the same as Balanced, but allows for the modification of individual parameters.

Choose High Performance to disable power management and adapt the Host BIOS so that the power setting shows "OS Control".

- The memory latency is also high but could be related to the CPU latency so see the impact on this metric after the power policy change.

3. Overall VMware Infrastructure

The infrastructure consists of a cluster in Wavre with ID “domain-c26” located in the datacenter “customer” with ID “datacenter-21”. There are eight data stores.

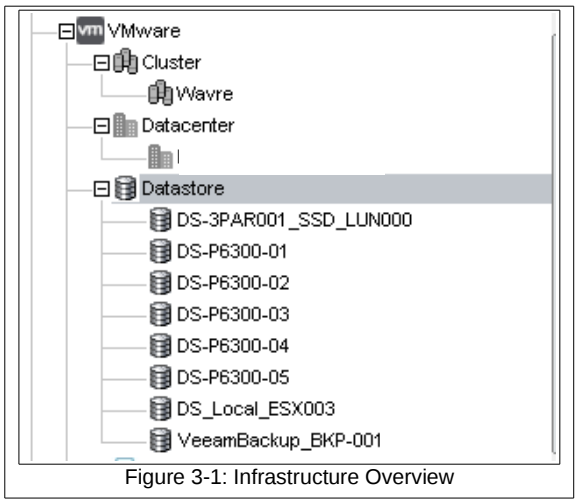

Figure 3-1: Infrastructure Overview

There are three ESX Hosts managed by a vCenter named “chimay”.

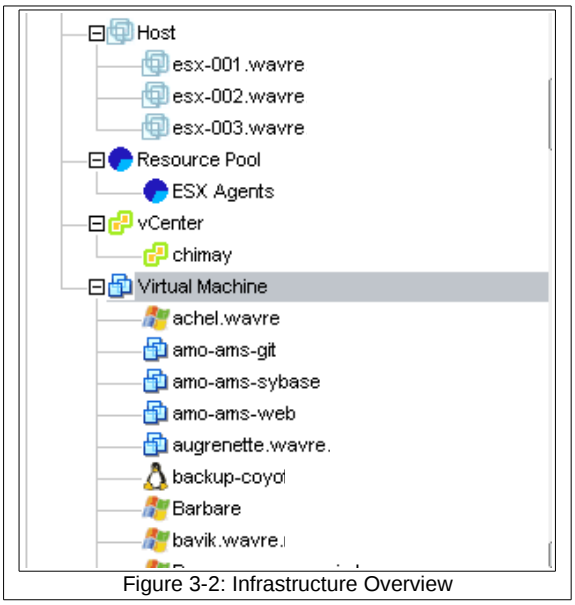

Figure 3-2: Infrastructure Overview

There are 3 active ESX Hosts and 124 VM’s.

Cluster Resource Allocation 10 mrt/17 00:00 - 31 mrt/17 00:00 (10 minute)

Time	vCenter	Cluster	ClusterId	Datacenter	DatacenterId	activeHosts	totalVMs
10 mrt/17 16:00	chimay	Wavre	domain-c26		datacenter-21	3	124
10 mrt/17 16:10	chimay	Wavre	domain-c26		datacenter-21	3	124

Figure 3-3: Cluster Resources

There is only one resource pool “ESX Agents” named “resgroup-10205”. But seems that this is not fully configured yet.

3.1 Cluster Level Changes

When checking the activity on VM's we see that there where no VM's moved from one ESX server to an other one.

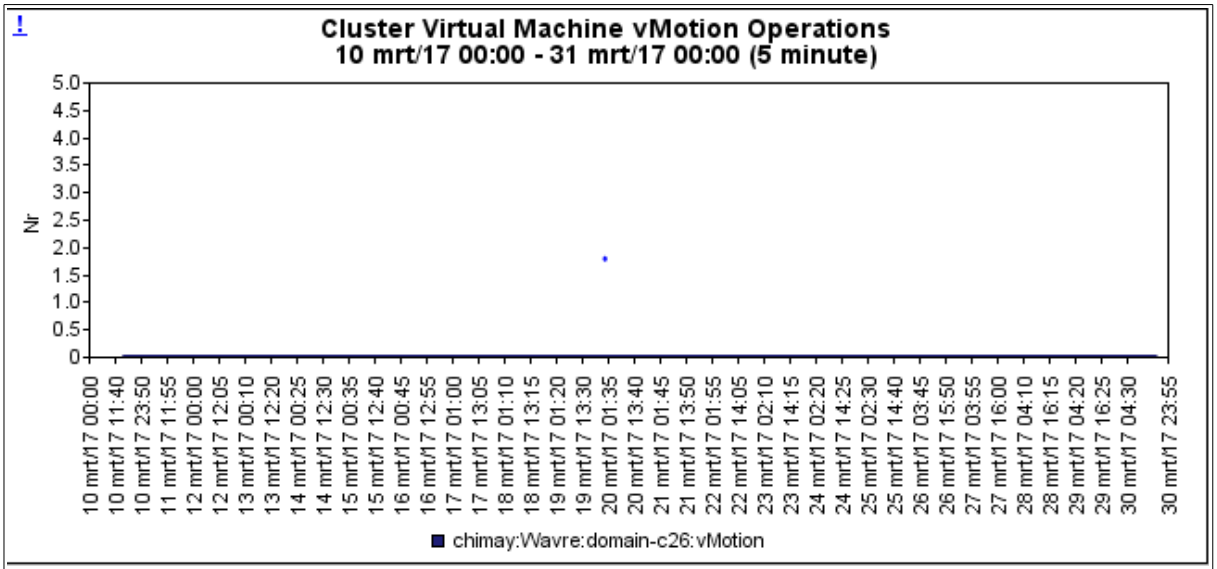

Figure 3-4: vMotion Activity

The number of VM's remained quite stable too.

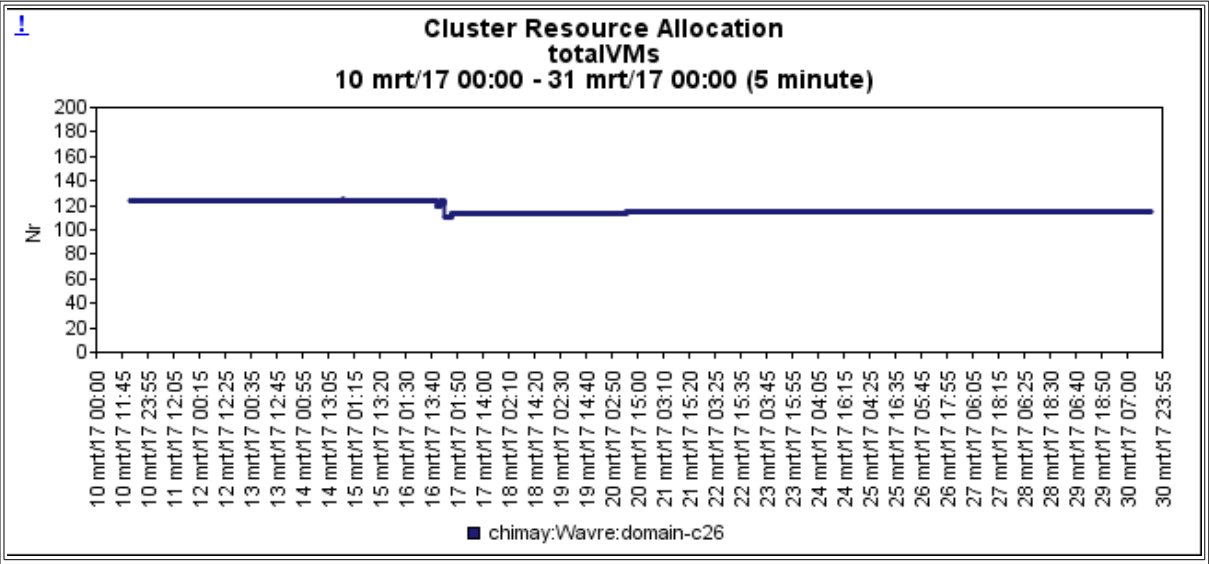

Figure 3-5: Total VM's

These VM's are distributed as following over the different ESX Hosts.

are.Host Configuration — 10 mrt/17 00:00 - 31 mrt/17 00:00

vCenter	Cluster	ClusterId	Host	Datacenter	DatacenterId	totalVMs
chimay	Wavre	domain-c26	esx-001.wavre		datacenter-21	42
chimay	Wavre	domain-c26	esx-002.wavre		datacenter-21	40
chimay	Wavre	domain-c26	esx-003.wavre		datacenter-21	42

Figure 3-6: VM Distribution Per ESX Host

4. CPU Resource

Here we look at the different levels at the configuration & consumption of the CPU resource.

4.1 Cluster Level

The effective CPU resources (in MHz) available to run virtual machines is shown below. This is the aggregated effective resource level from all running hosts. Hosts that are in maintenance mode or are unresponsive are not counted. Resources used by the VMware Service Console are not included in the aggregate.

This value represents the amount of resources available for the root resource pool for running virtual machines. We have a total of 147.441 MHz.

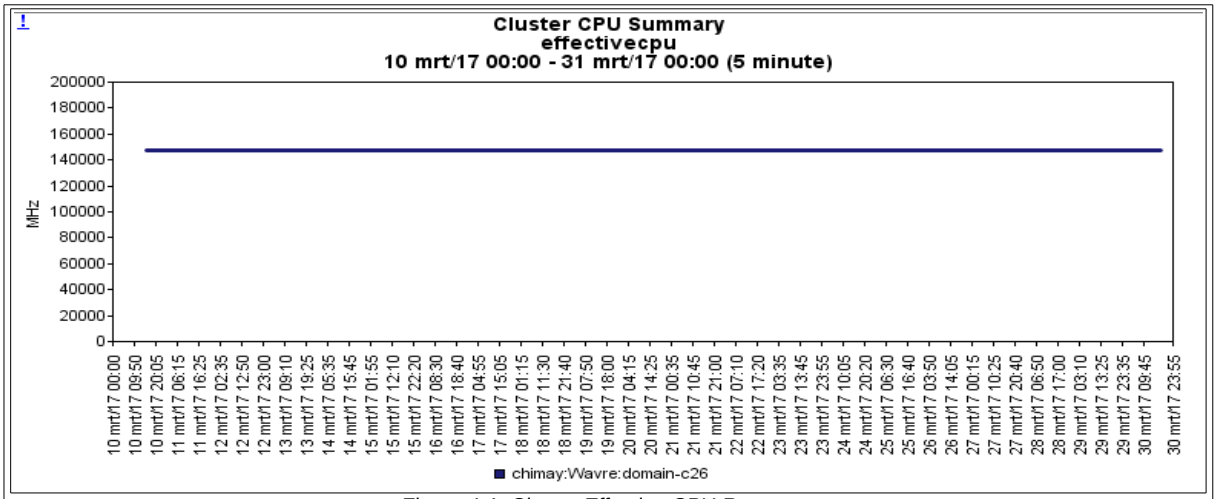

Figure 4-1: Cluster Effective CPU Resource

We do not have the percentage “%busy” of elapsed CPU time the processors were busy across all of the virtual machines in the Vmware cluster since the Vmware Distributed Resource Scheduler (DRS) is probably disabled for the VMware cluster.

But we will derive this activity a bit further in the Host Level section.

4.2 Resource Pool

The activity is very low since it is not fully configured, seems to be the root resource pool.

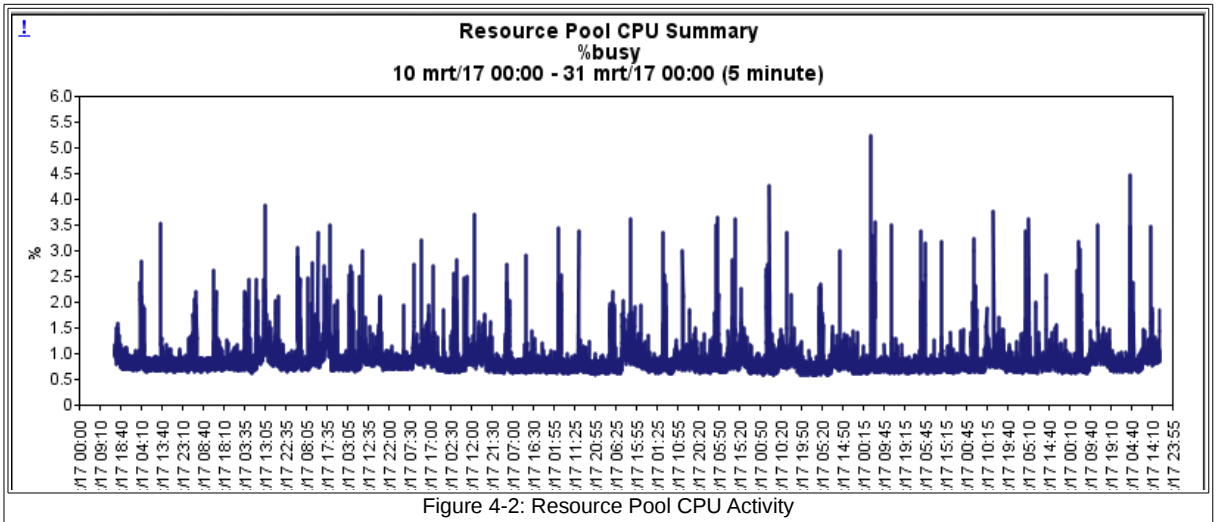

Figure 4-2: Resource Pool CPU Activity

4.1 ESX Host Level

These three servers have the same nr. of physical cores but not the same nr. of logical CPU's although these servers have the same CPU chip type.

Host CPU Summary			21 mrt/17 00:00 - 22 mrt/17 00:00 (24 hour)				
System	vCenter	Cluster	ClusterId	Datacenter	DatacenterId	online_cpus	online_cpus_phys
esx-001.wavre	chimay	Wavre	domain-c26		datacenter-21	48.000	24.000
esx-002.wavre	chimay	Wavre	domain-c26		datacenter-21	24.000	24.000
esx-003.wavre	chimay	Wavre	domain-c26		datacenter-21	48.000	24.000

Figure 4-3: ESX Server Cores & LCPU's

This is due to the fact that the Core Multi Thread has been disabled on server "ESX-002".

HINV.Summary			21 mrt/17 00:00 - 22 mrt/17 00:00		6/6		
System	CPU Chips	CPU Count	CPU Speed	CPU Type	Core Multi-Thread	Cores Per Chip	Logical CPU Count
esx-001.wa	2	24	2297.34 MHz	Intel(R) Xeon(R) CPU E5-2670 v3 @ 2.30GHz	Enabled	12	48
esx-002.wa	2	24	2297.34 MHz	Intel(R) Xeon(R) CPU E5-2670 v3 @ 2.30GHz	Disabled	12	24
esx-003.wa	2	24	2297.34 MHz	Intel(R) Xeon(R) CPU E5-2670 v3 @ 2.30GHz	Enabled	12	48

Figure 4-4: ESX Server Cores & LCPU's

The effective ESX host activity is quite small, the red line represents the cluster effective CPU resources available. So largely below the maximum.

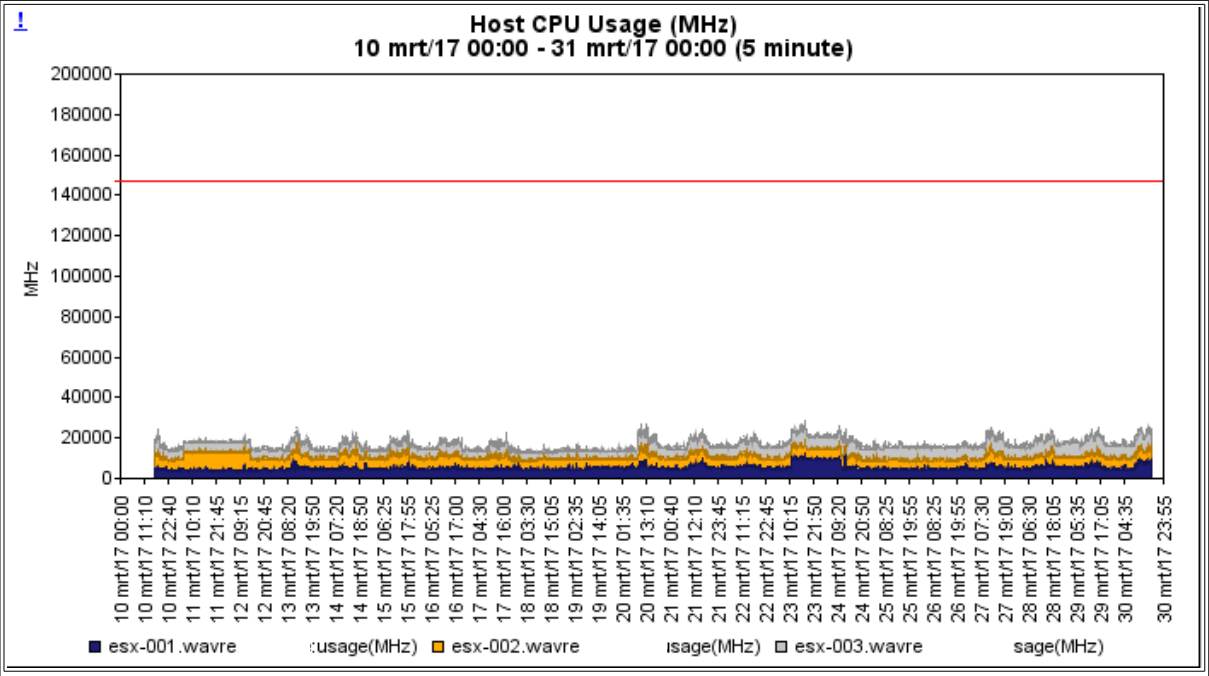

Figure 4-5: ESX Host CPU Activity

As such the load in percentage is quite low, CPU resource never used more than 22%.

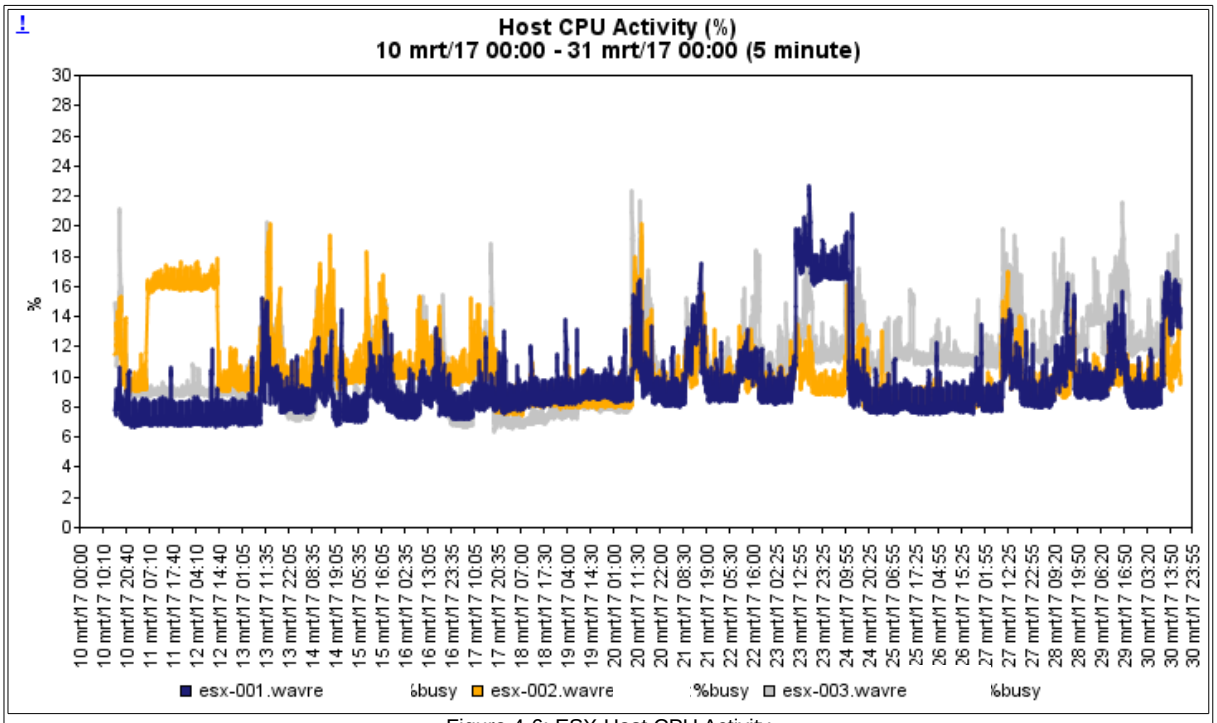

Figure 4-6: ESX Host CPU Activity

When verifying the load by LCPU for the whole period we see that it is evenly spread with few active days. On “ESX-001” we have higher activity between Thursday 23th 11h – Friday 24th 11h.

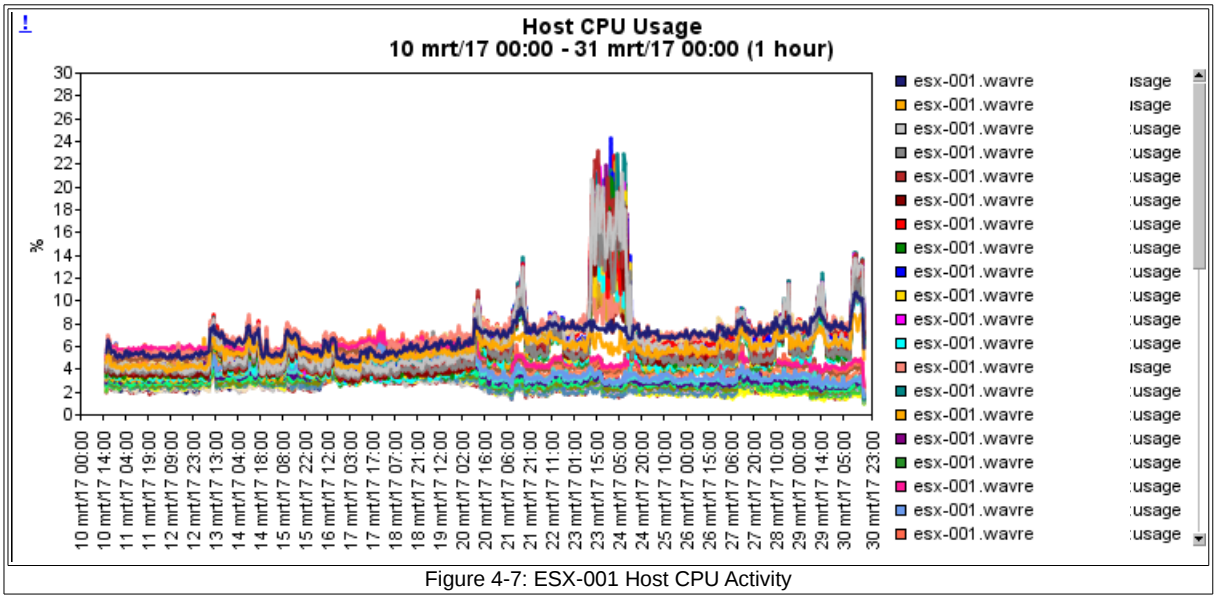

Figure 4-7: ESX-001 Host CPU Activity

On “ESX-002” we have higher activity between the Saturday 11th 5h – Sunday 12th 14h.

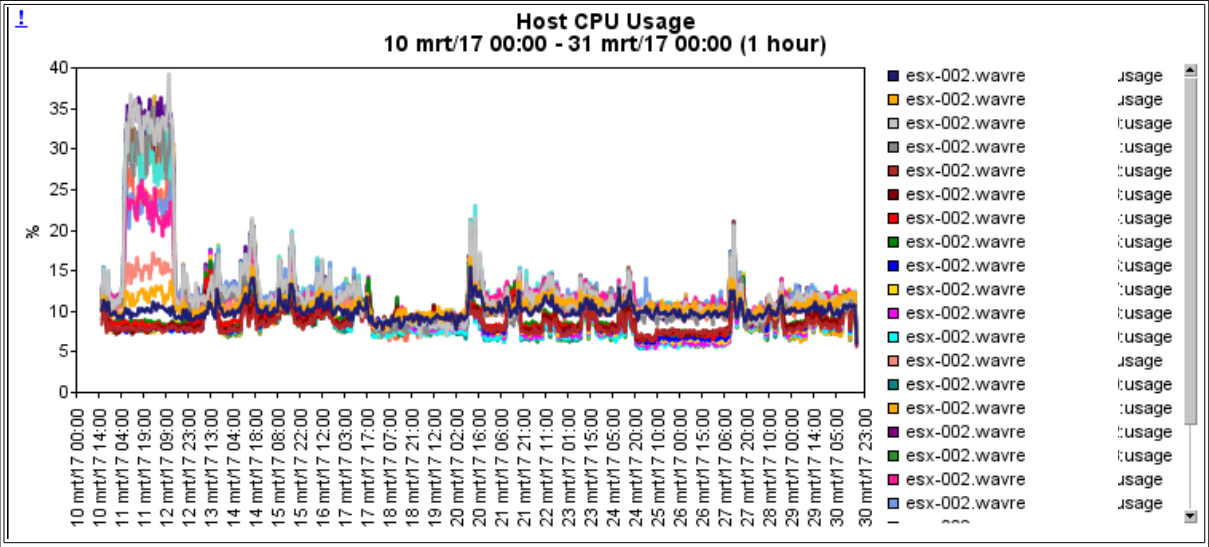

Figure 4-8: ESX-002 Host CPU Activity

On “ESX-003” no increased load was found.

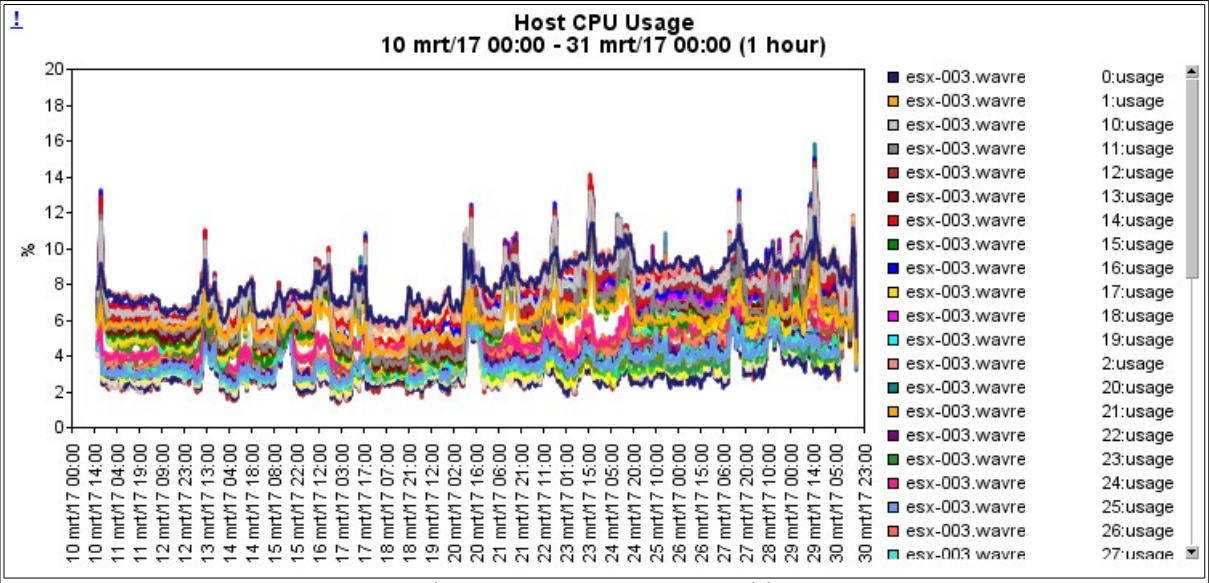

Figure 4-9: ESX-003 Host CPU Activity

The following chart shows “costop” representing the amount of time in seconds a Symmetric Multi-Processing (SMP) virtual machine was ready to run, but was delayed due to console virtual CPU (co-vCPU) scheduling contention.

In fact it is the additional time needed getting the next vCPU’s after receiving the first vCPU. So will only occur on VM’s with multiple vCPU’s configured.

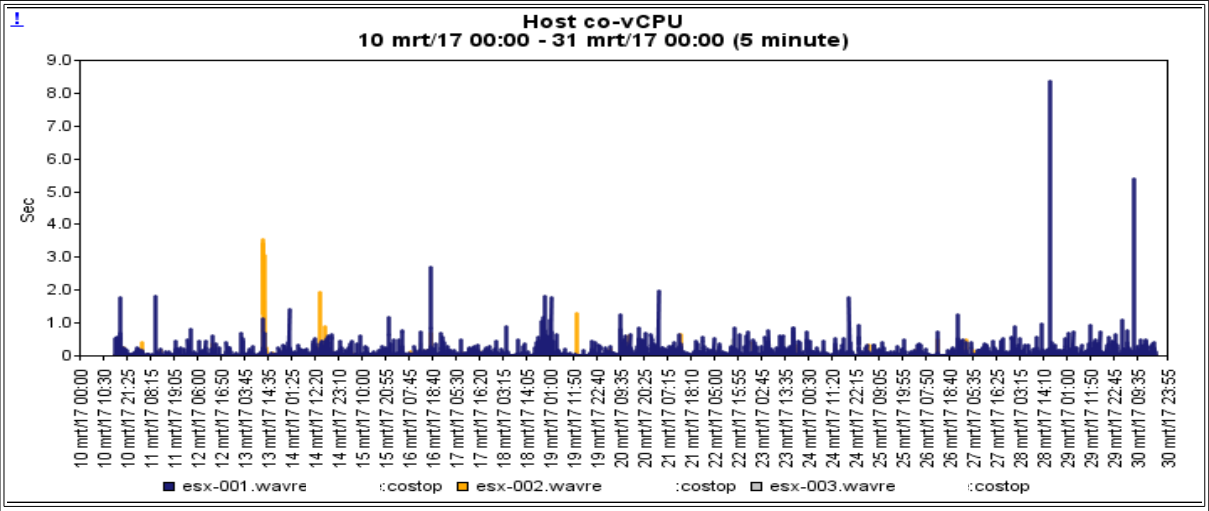

Figure 4-10: ESX Host CPU co-vCPU

Apart for some sporadic small peaks these values are very low which indicates that there is no contention in acquiring extra vCPU resources. The goal is to remain under the 3% which is the case here. 1 Sec of 5 Min is about 0,33%. The next three charts show each ESX Host separately.

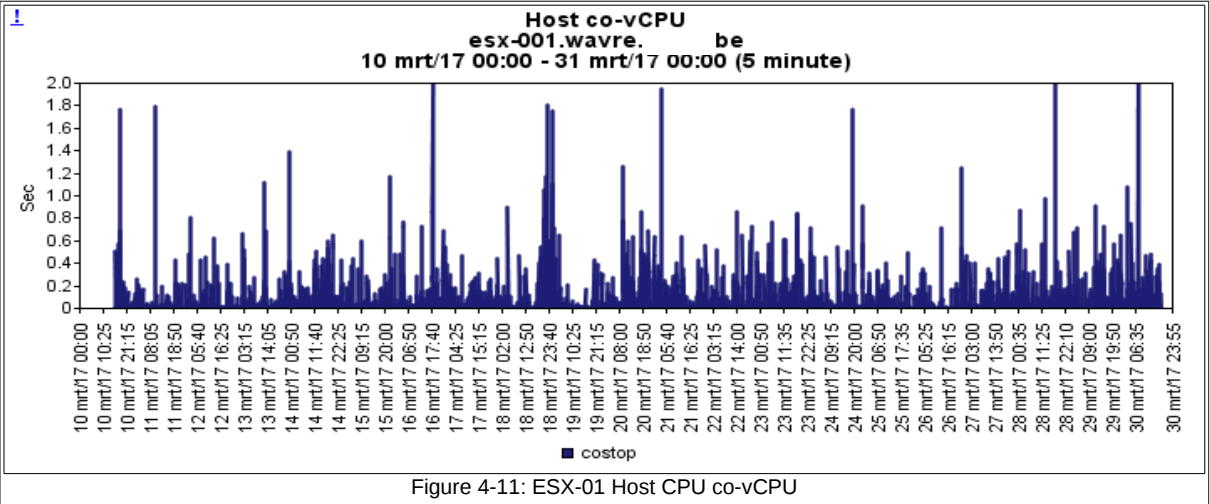

Figure 4-11: ESX-01 Host CPU co-vCPU

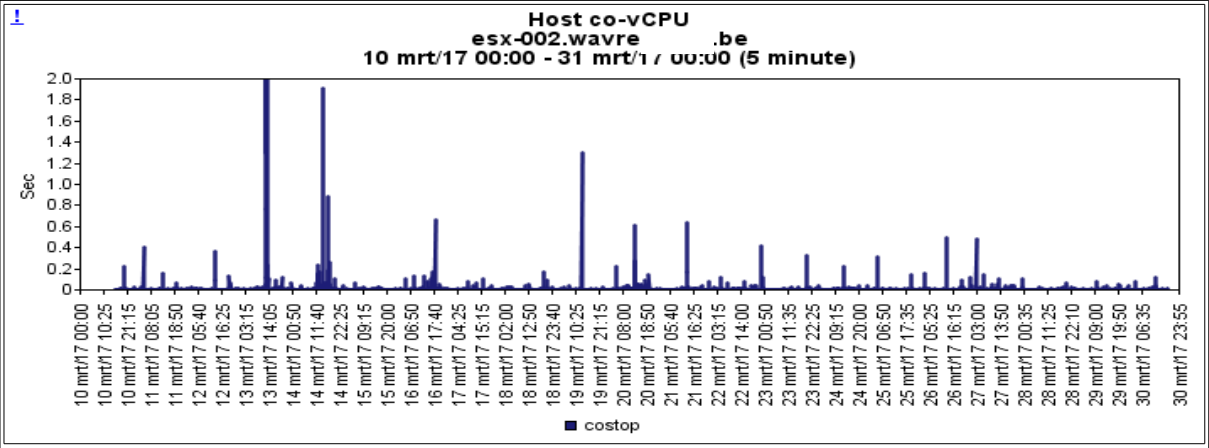

Figure 4-12: ESX-02 Host CPU co-vCPU

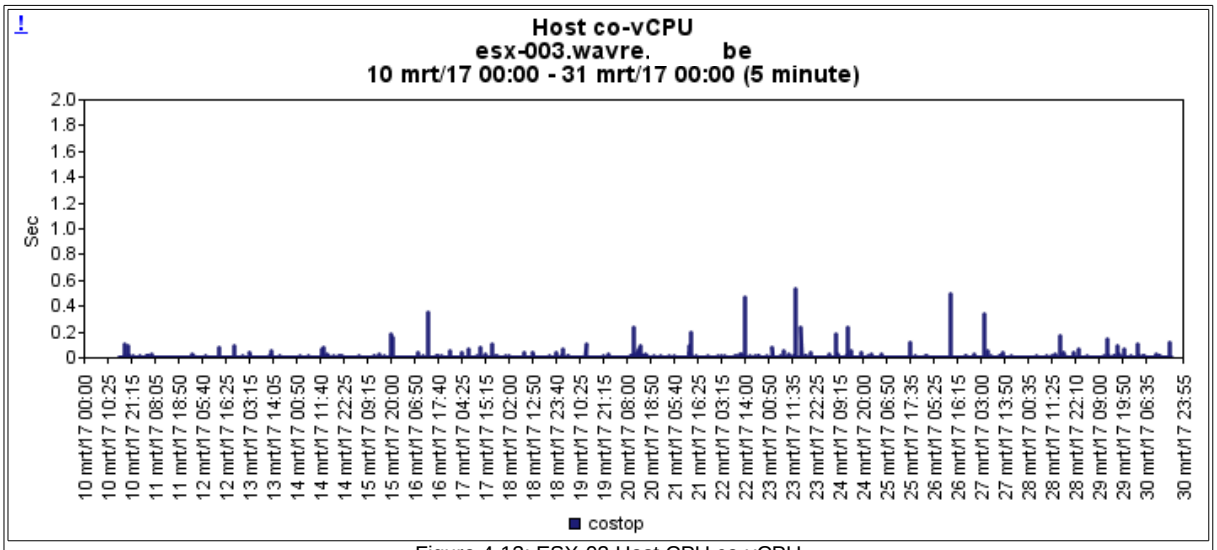

Figure 4-13: ESX-03 Host CPU co-vCPU

The “latency” activity, percentage of time the virtual machines where ready to run but not scheduled to run because of physical CPU resource contention, is shown below. We refer to the next section for more information.

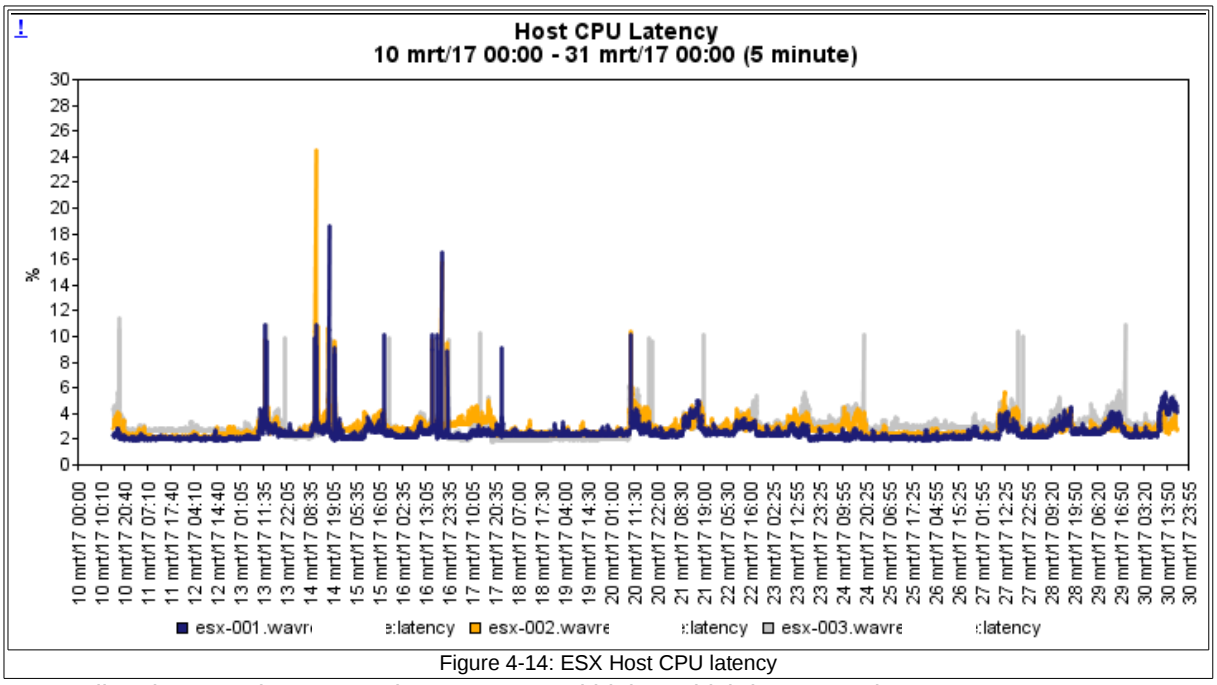

Figure 4-14: ESX Host CPU latency

Bottom line there are latency peaks up to 1% and higher which is not good.

4.1 ESX VM Level

As already mentioned there are about 124 VM’s configured. First we look at the “%vcpu_busy”, the percentage of the virtual machines virtual processors used for the period 13th – 24th March/17 without Sat/Sun.

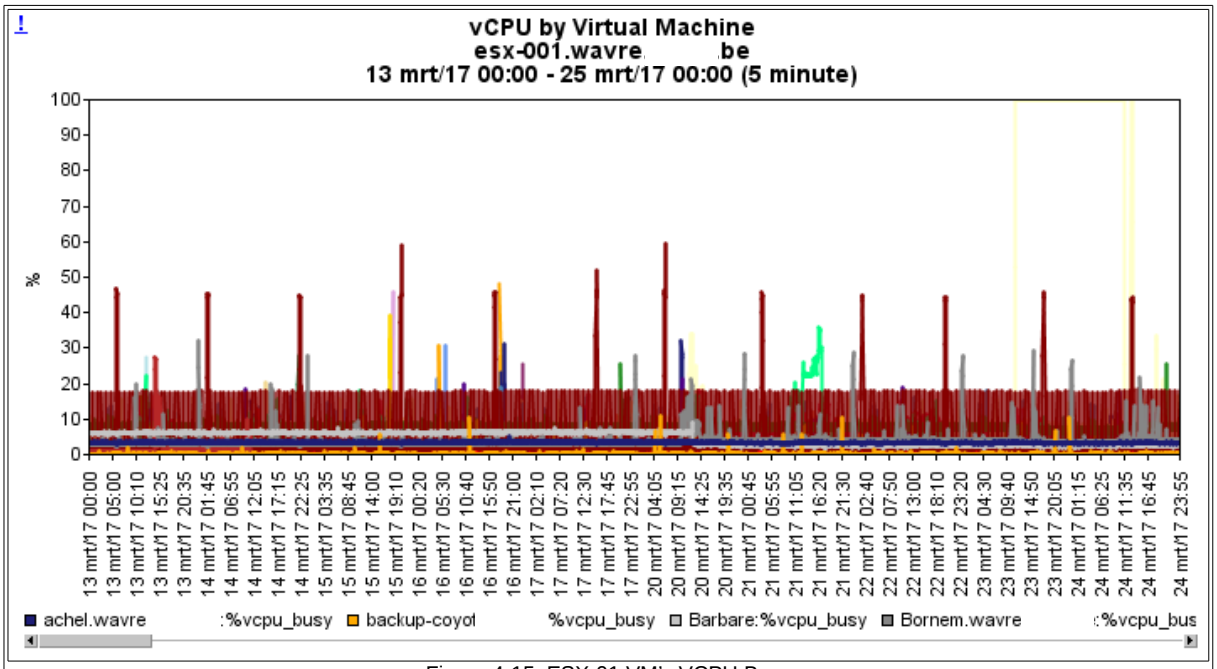

Figure 4-15: ESX-01 VM's VCPU Busy

Overall small load.

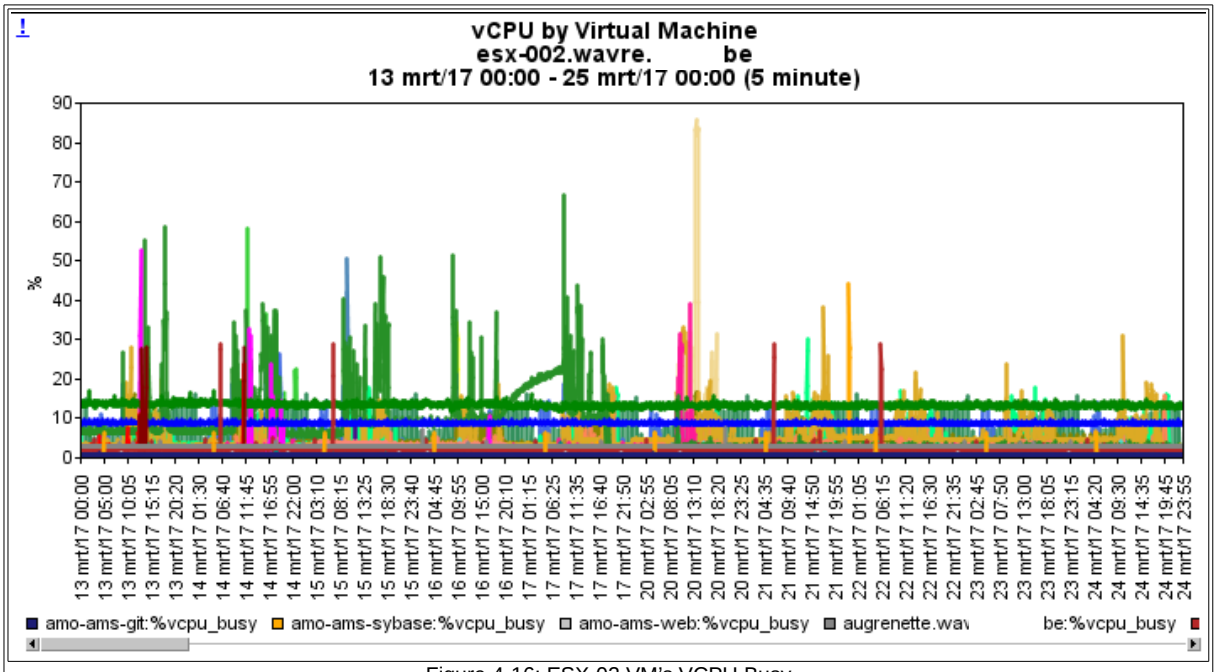

Figure 4-16: ESX-02 VM's VCPU Busy

First week is more active compared to the second one with medium load and few peaks up to 67% for VM "Monk-021" and VM "Cash-3" 86%.

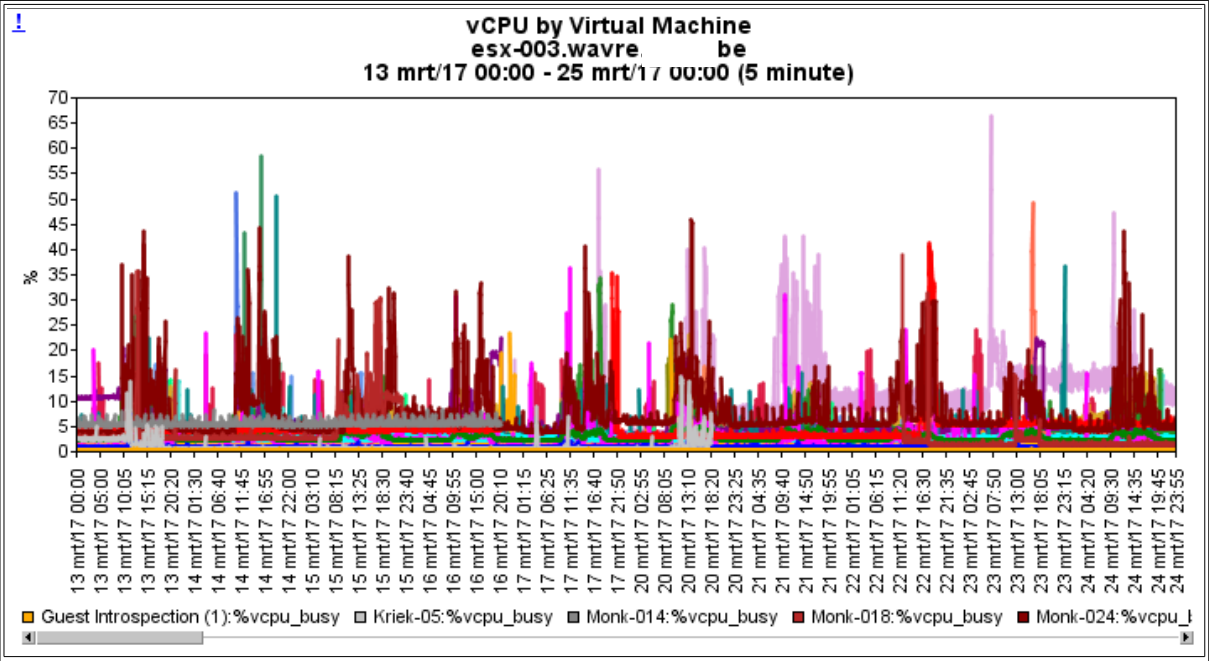

Figure 4-17: ESX-03 VM's VCPU Busy

Here also medium load with some peaks up to 60% for VM "Cash-03".

The "maxlimited", representing the amount of time in seconds the virtual machine was ready to run but did not run because it reached the maximum CPU limit setting, is shown below.

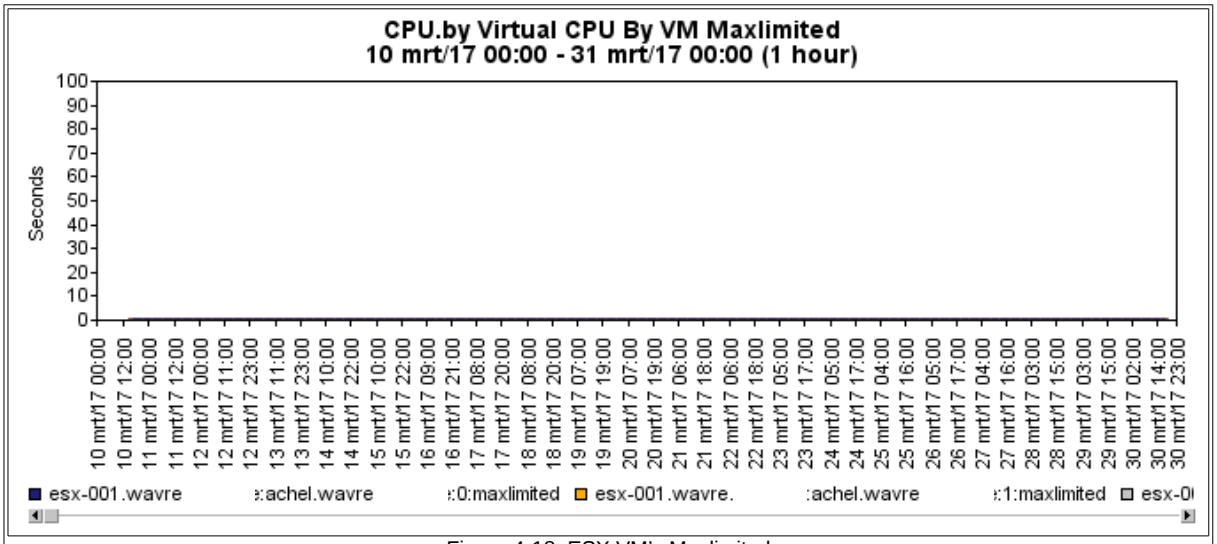

Figure 4-18: ESX VM's Maxlimited

It reveals that the VM's are not blocked due to a CPU limitation.

Next we look at the "costop" activity representing the amount of time in seconds a Symmetric Multi-Processing (SMP) virtual machine was ready to run, but was delayed due to console virtual CPU (co-vCPU) scheduling contention.

Remember it is the additional time needed getting the next vCPU's after receiving the first vCPU. So will only occur on VM's with multiple vCPU's configured.

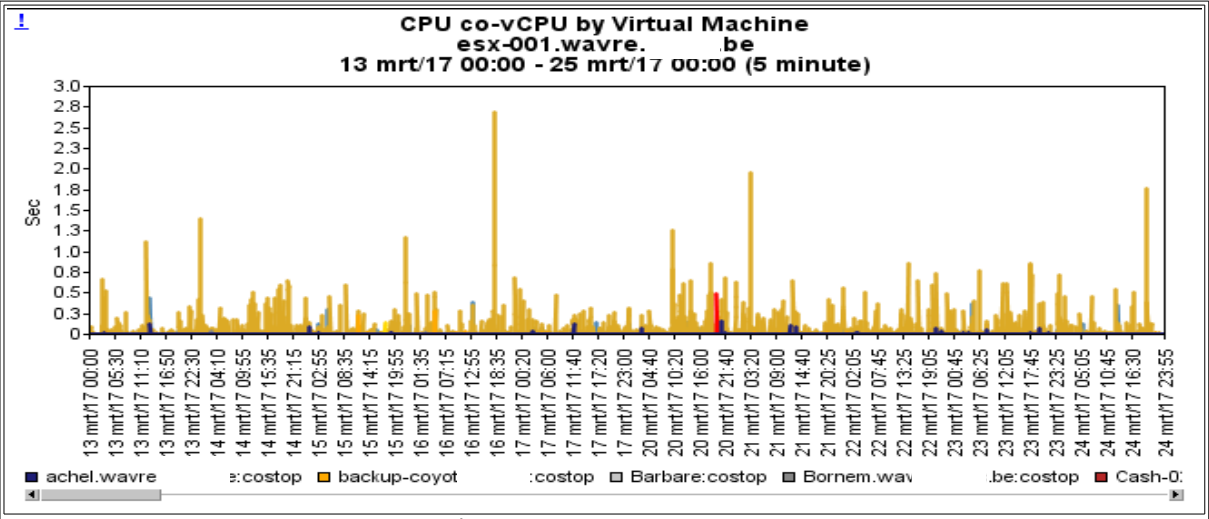

Figure 4-19: ESX-01 VM's costop

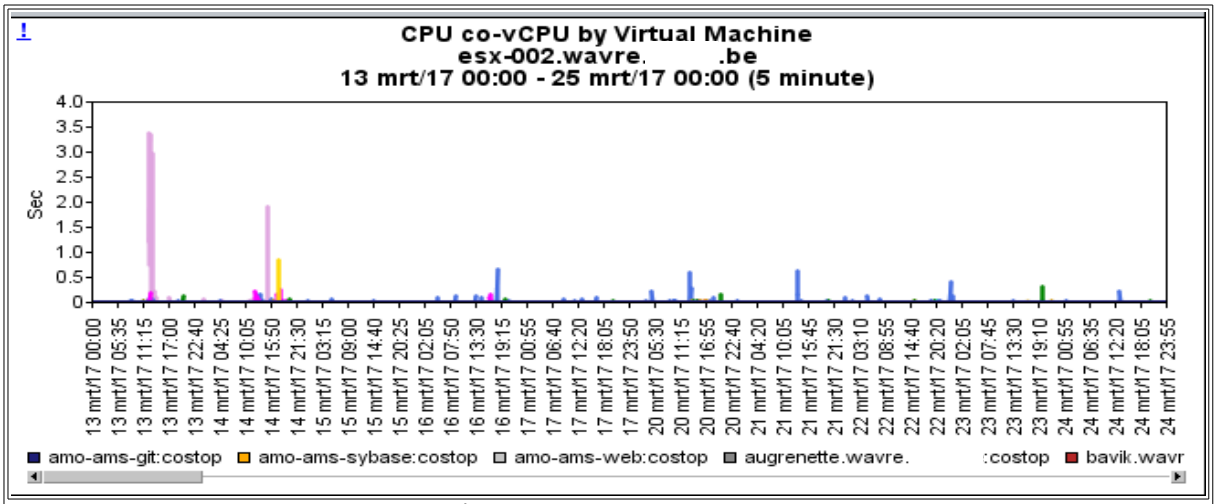

Figure 4-20: ESX-02 VM's costop

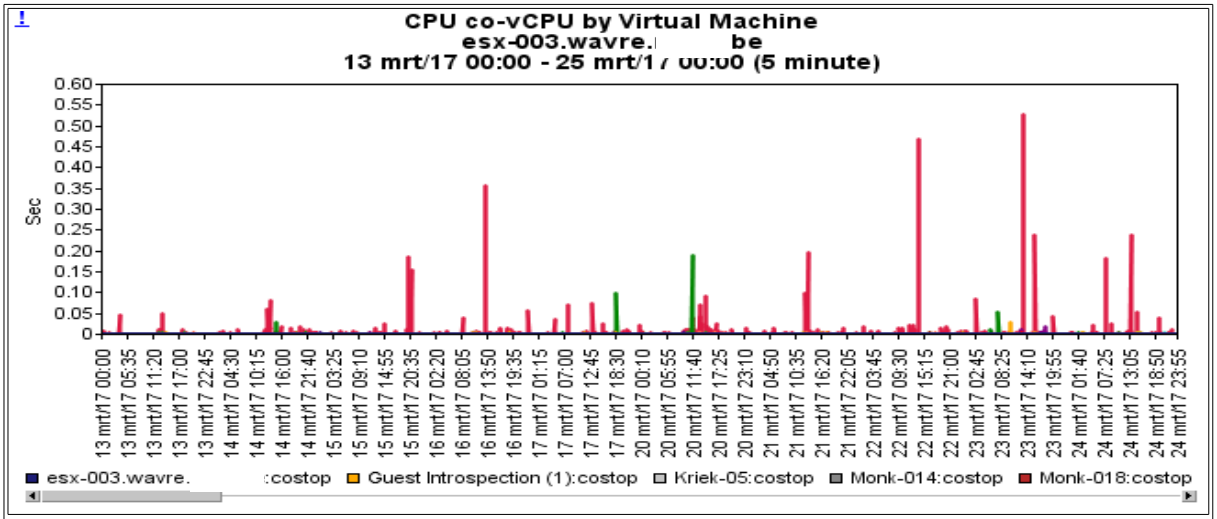

Figure 4-21: ESX-03 VM's costop

These are all quite low too so this excludes again CPU scheduling issues.

Now we look at the "%vcpu_ready" activity representing the percentage of time the virtual machine was ready to perform an operation but had to wait for a physical processor. In fact it is the time that the VM waits in a ready-to-run state (meaning it has work to do) to be scheduled on one or more of the physical CPUs by the hypervisor.

It is generally normal for VM's to have small values for CPU Ready Time accumulating even if the hypervisor is not over subscribed or under heavy activity, it's just the nature of shared scheduling in virtualization. For SMP VM's with multiple vCPU's the amount of ready time will generally be higher than for VM's with fewer vCPU's since it requires more resources to schedule/co-schedule the VM when necessary and each of the vCPU's accumulates the time separately.

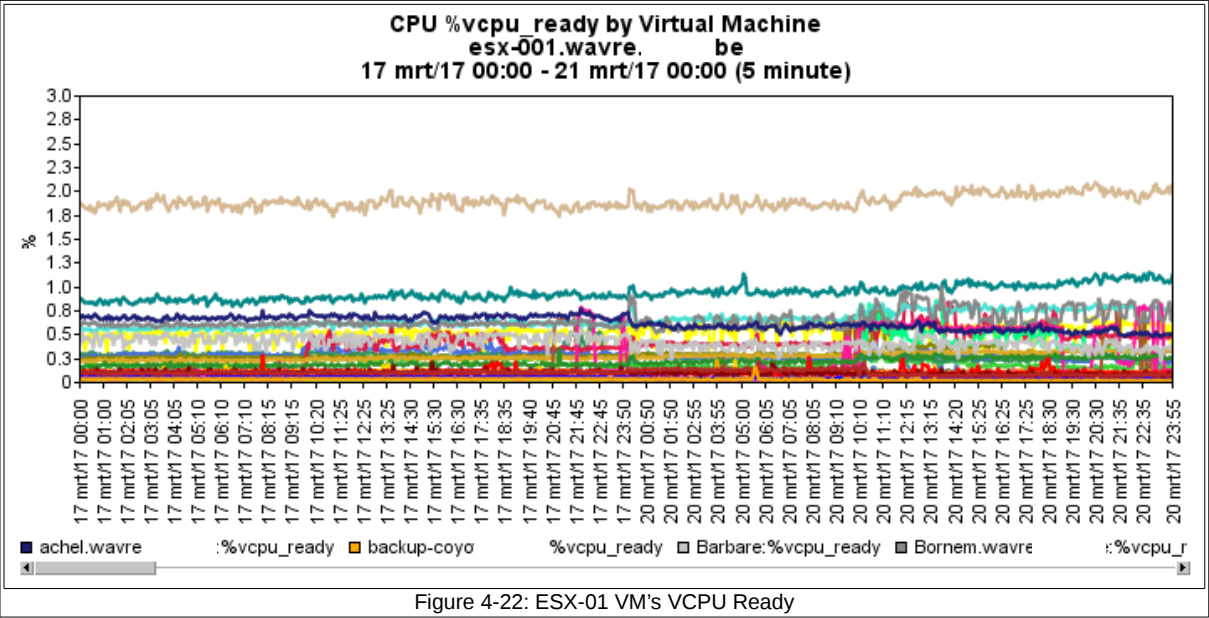

Figure 4-22: ESX-01 VM's VCPU Ready

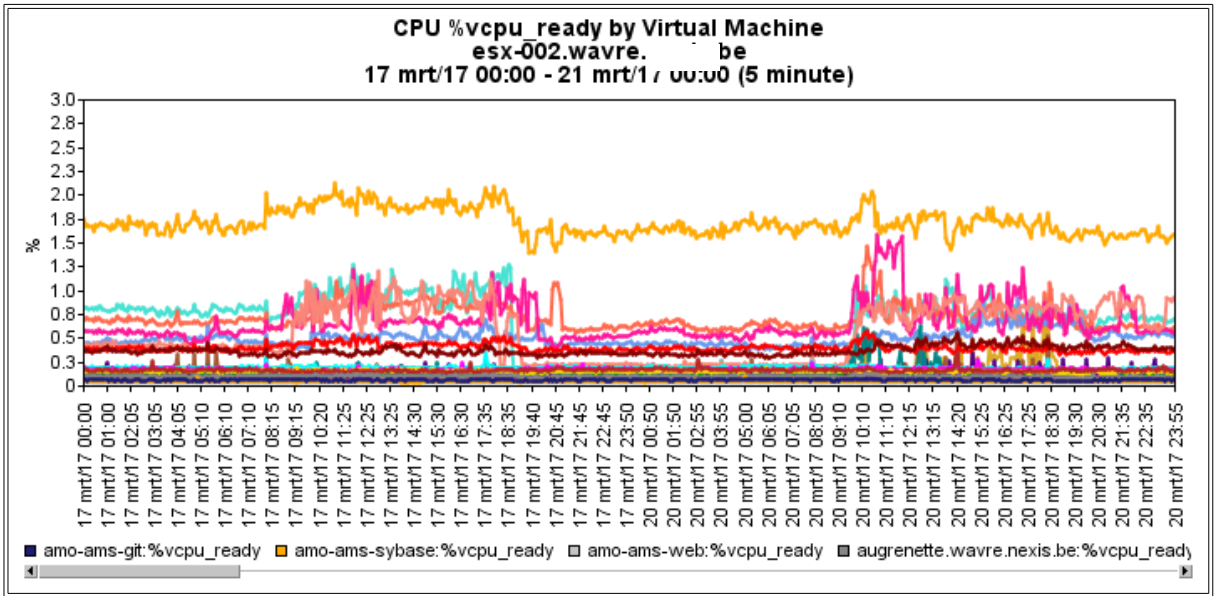

Figure 4-23: ESX-02 VM's VCPU Ready

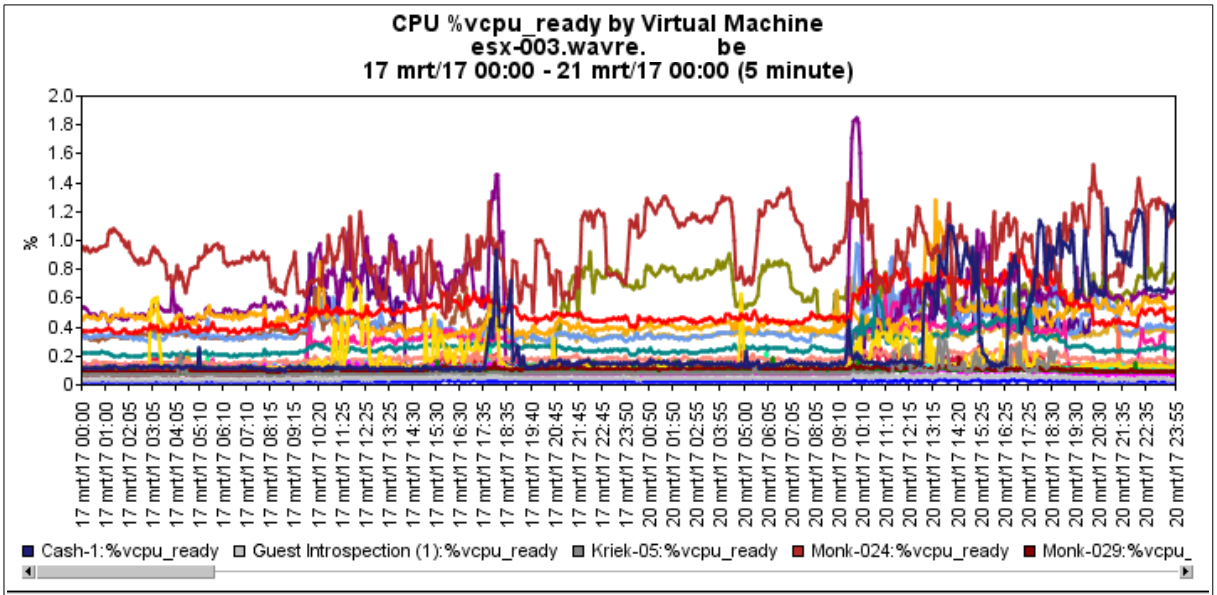

Figure 4-24: ESX-03 VM's VCPU Ready

These are all quite low so this excludes CPU scheduling issues.

Next we look at the "latency" representing the percentage of time the virtual machine was ready to run but was not scheduled to run because of physical CPU resource contention. In fact it contains "cpu ready", "swapwait" and "power regulation (C-state)". So If your "cpu ready" is ok, and "swap wait" is 0, then it is the power regulation.

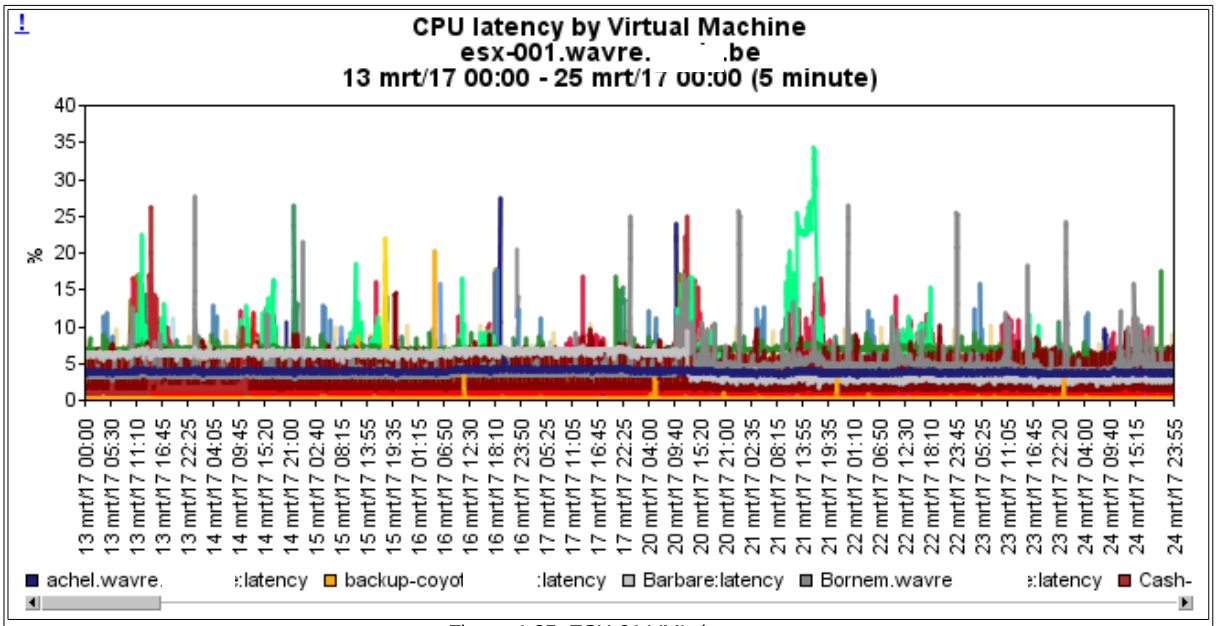

Figure 4-25: ESX-01 VM's latency

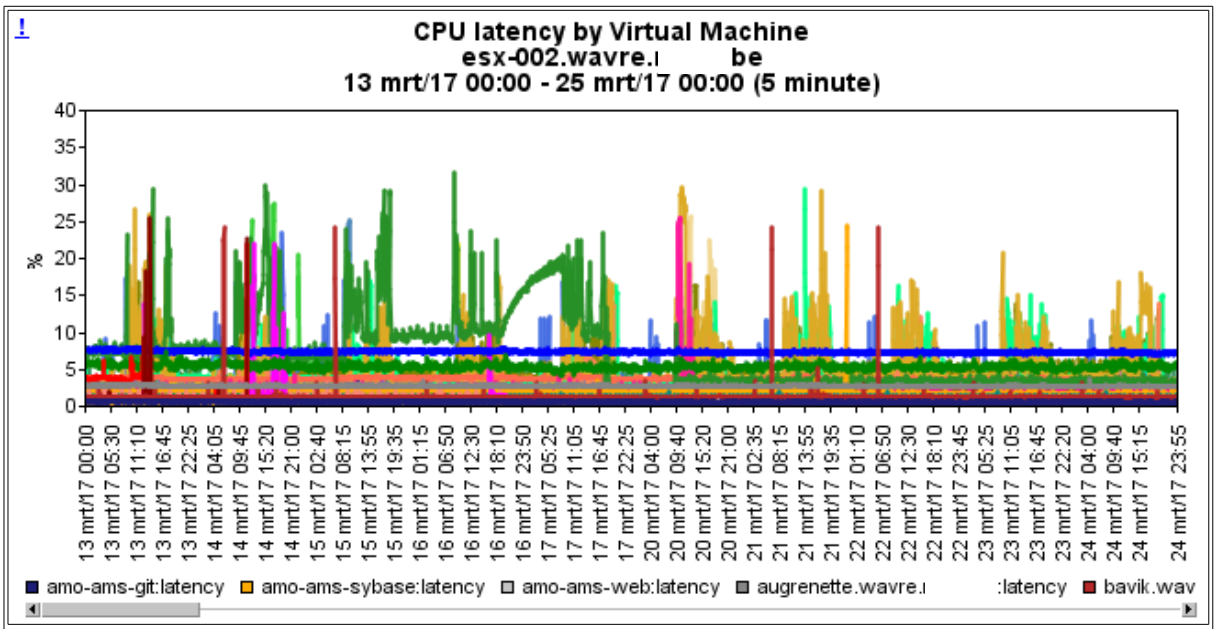

Figure 4-26: ESX-02 VM's latency

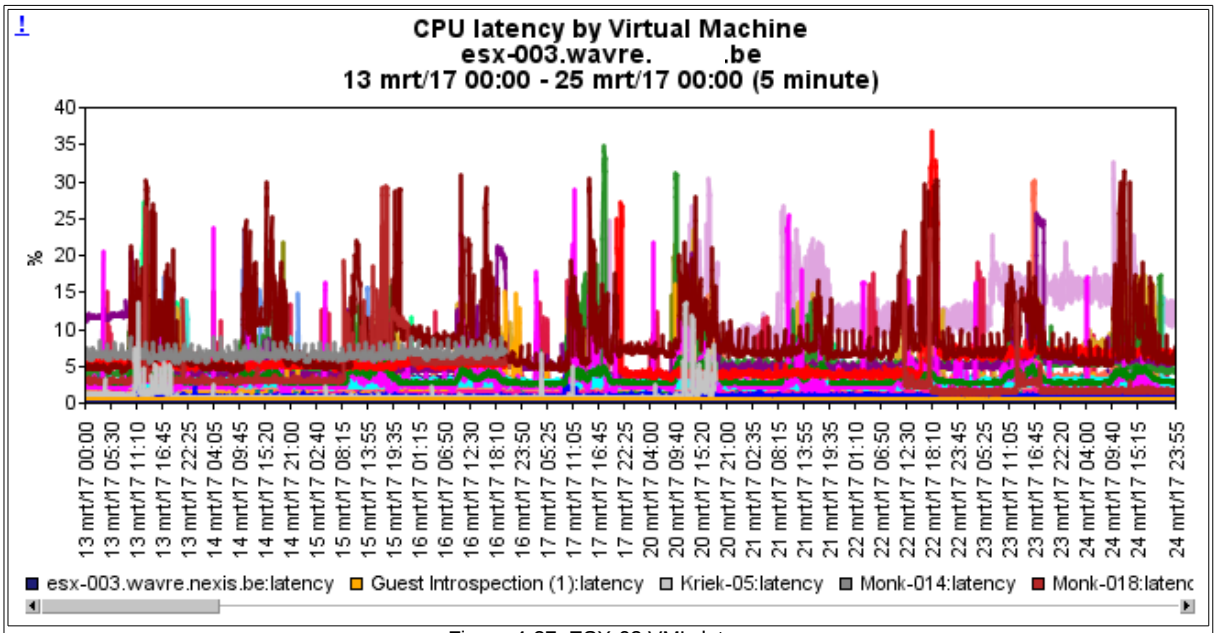

Figure 4-27: ESX-03 VM's latency

These latency figures are quite high and are in this case probably due to the processor power management implemented either by ESXi/ESX or by the server hardware. Certain applications that are very sensitive to processing speed latencies may show less than expected performance when processor power management features are enabled.

It may be necessary to turn off ESXi/ESX and server hardware power management features to achieve the best performance for such applications.

Disabling power management usually results in more power being consumed by the system, especially when it is lightly loaded. The majority of applications benefit from the power savings offered by power management, with little or no performance impact.

Therefore, if disabling power management does not realize any increased performance, It is recommended that power management be re-enabled to reduce power consumption:

ESXi 6.0 offers the following power policy options

- High performance
  - This power policy maximizes performance, using no power management features.
- Balanced
  - This power policy (the default in ESXi 6.0) is designed to reduce host power consumption while having little or no impact on performance.
- Low power
  - This power policy is designed to more aggressively reduce host power consumption at the risk of reduced performance.
- Custom
  - This power policy starts out the same as Balanced, but allows for the modification of individual parameters.

Choose High Performance to disable power management and adapt the Host BIOS so that the power setting shows "OS Control".

5. Memory Resource

Here we look at the different levels at the configuration & consumption of the Memory resource.

5.1 Cluster Level

The effective memory representing the total amount of memory in megabytes of all of the hosts within a VMware cluster that are available for the virtual machine memory and the virtual machine overhead memory. There is about 727 GBytes available.

Below we see the “memAvailable” representing the total amount of memory in megabytes that is available to satisfy the reservation for all of the virtual machines and resource pools in the cluster.

Second parameter is “memOverhead” representing the total amount of memory in megabytes that has been used to satisfy the reservation requirements of all of the descendants of the running virtual machines in the cluster.

The third one is “memReserved” representing the total amount of memory in megabytes that has been used to satisfy the reservation requirements of all of the descendants of the virtual machines and resource pools in the cluster.

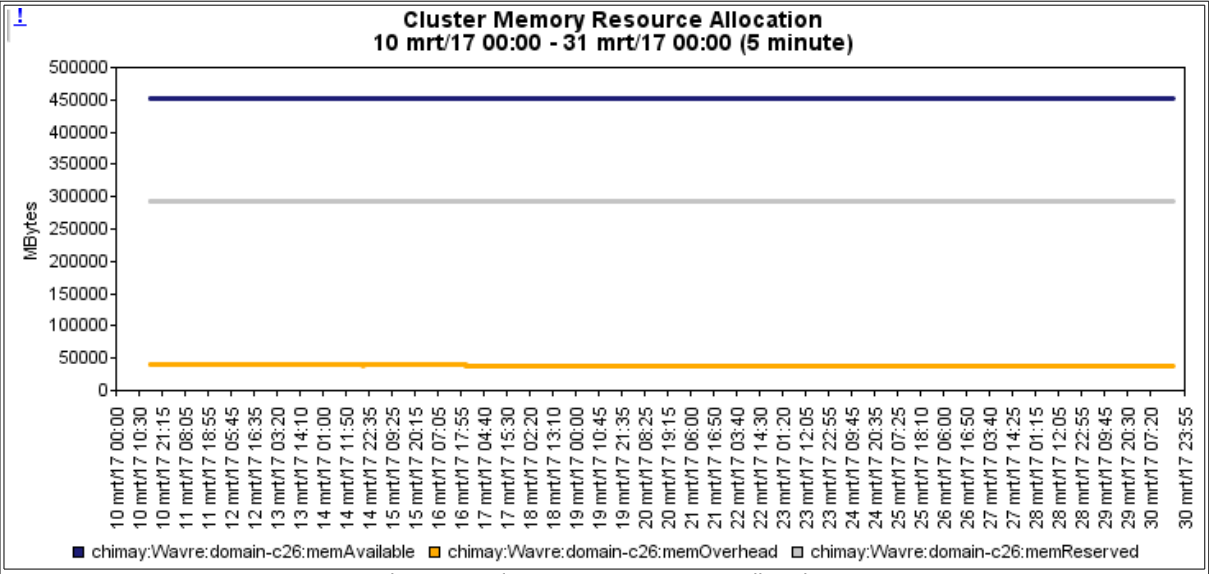

Figure 5-1: Cluster Memory Resource Allocation

There is sufficient available memory.

5.1 ESX Host Level

These three servers have the same amount of physical memory installed.

HINV.Summary 10 mrt/17 00:00 - 31 mrt/17 00:00

Time	System	Memory	Memory KB	Memory MB
10 mrt/17 16:10	esx-001.wavre.	255.87 GBytes	268302380	262014
10 mrt/17 16:10	esx-002.wavre.	255.87 GBytes	268302380	262014
10 mrt/17 16:10	esx-003.wavre.	255.87 GBytes	268302380	262014

Figure 5-2: ESX Server Memory

The %usage representing the percentage of memory usage over the collection interval looks as following.

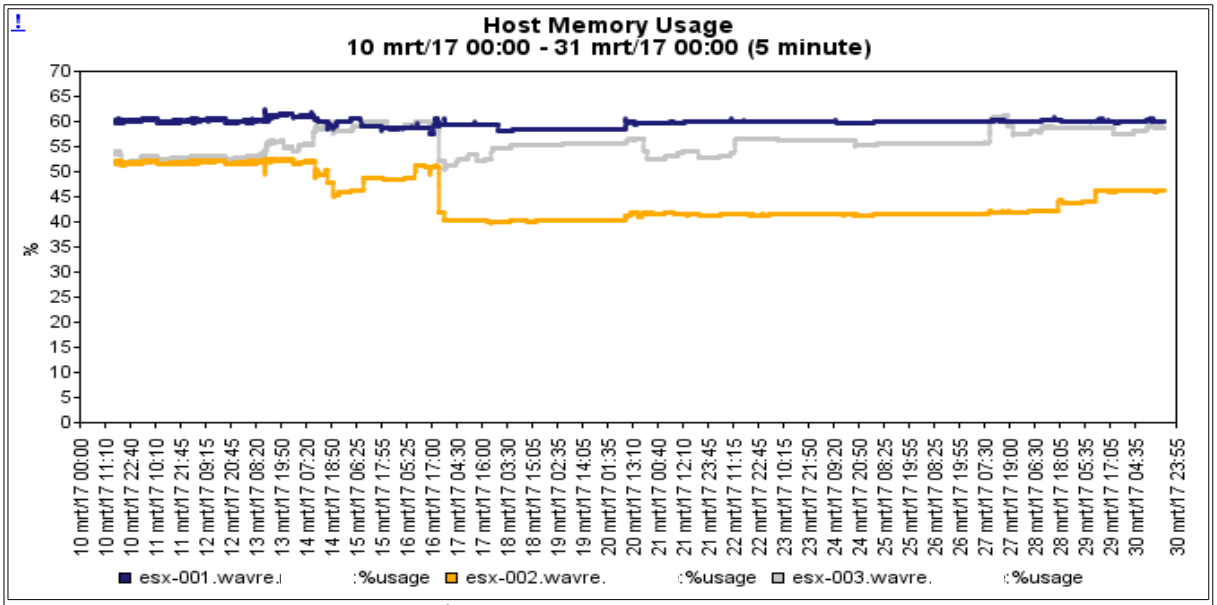

Figure 5-3: ESX Host Memory Usage

Remains below 60% leaving sufficient headroom for more load.

The memory latency representing the percentage of time the virtual machine was waiting to access swapped or compressed memory is un existing which is excellent.

We have the same figures for the “swapin”, “swapout” & “swapused” parameters.

The amount of time in seconds virtual machines where waiting for swap page-ins is shown below. CPU swap wait is included in CPU wait. So definitely no Memory issue here.

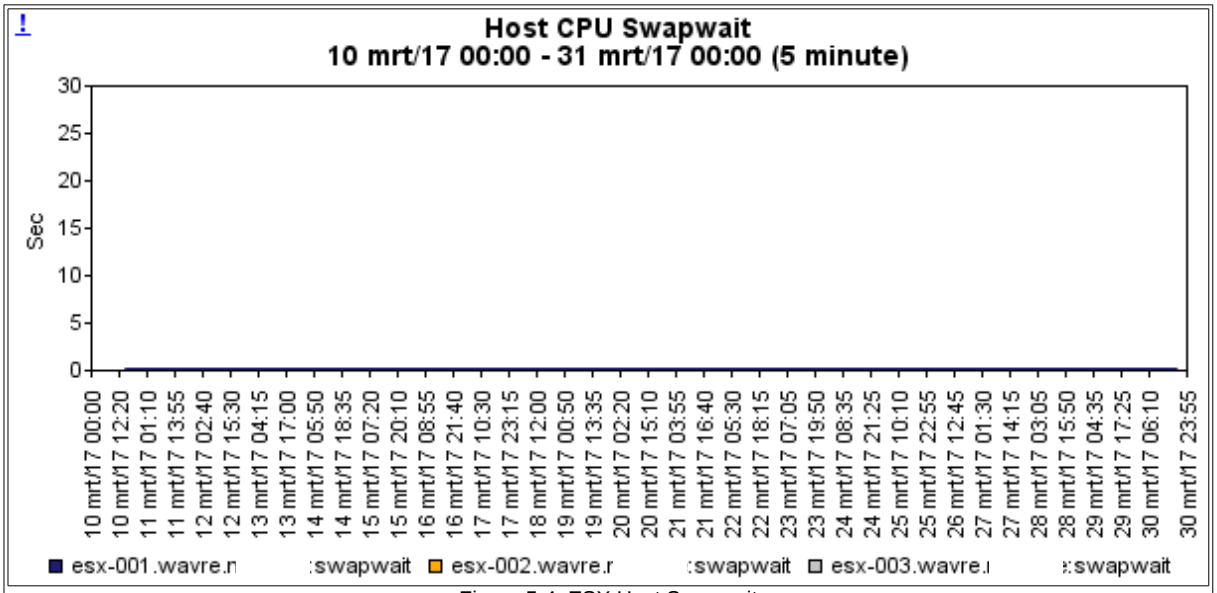

Figure 5-4: ESX Host Swapwait

5.1 ESX VM Level

The memory usage by VM is shown below.

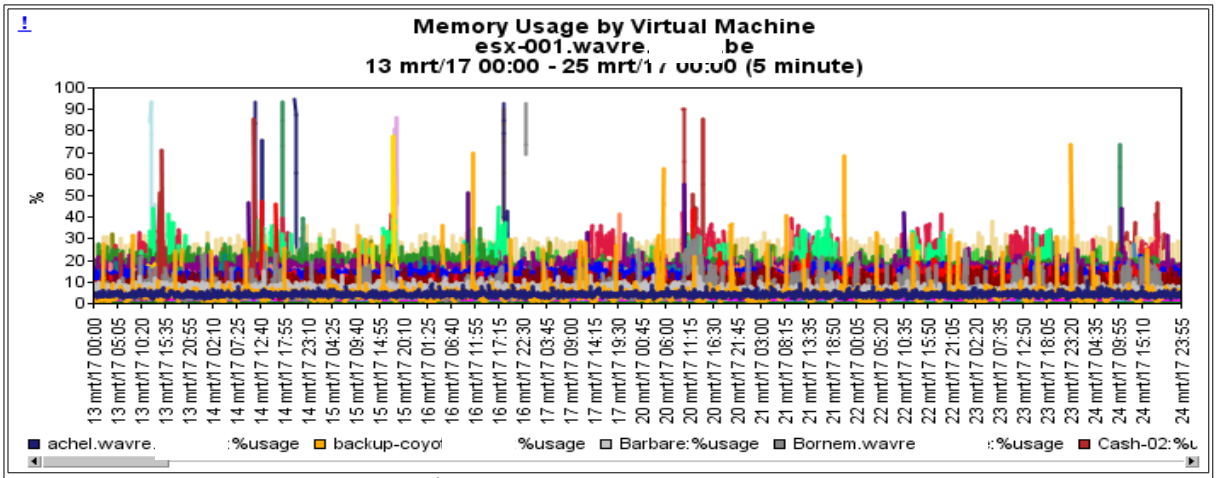

Figure 5-5: ESX-01 VM Memory Usage

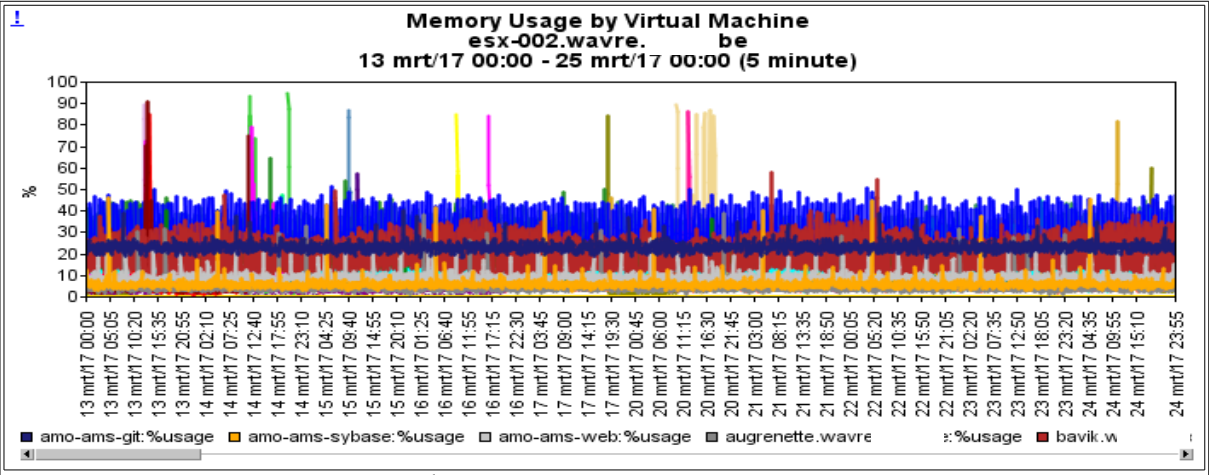

Figure 5-6: ESX-02 VM Memory Usage

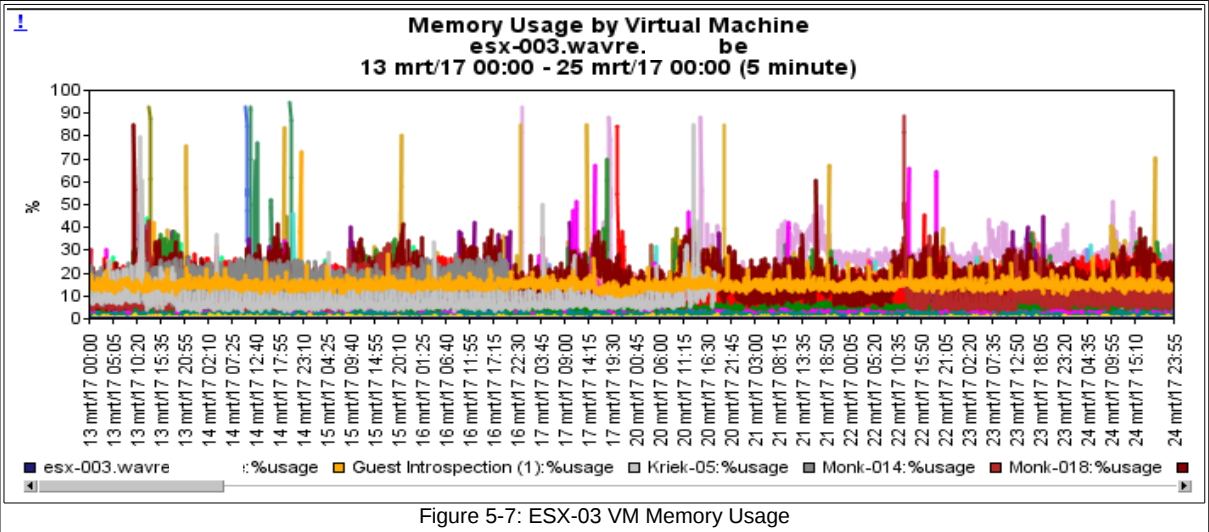

Figure 5-7: ESX-03 VM Memory Usage

The memory usage is in general around 30% to 40% with some small peaks up to 90% so no real issues. Next we look at the “latency” representing the percentage of time the virtual machine was waiting to access swapped or compressed memory.

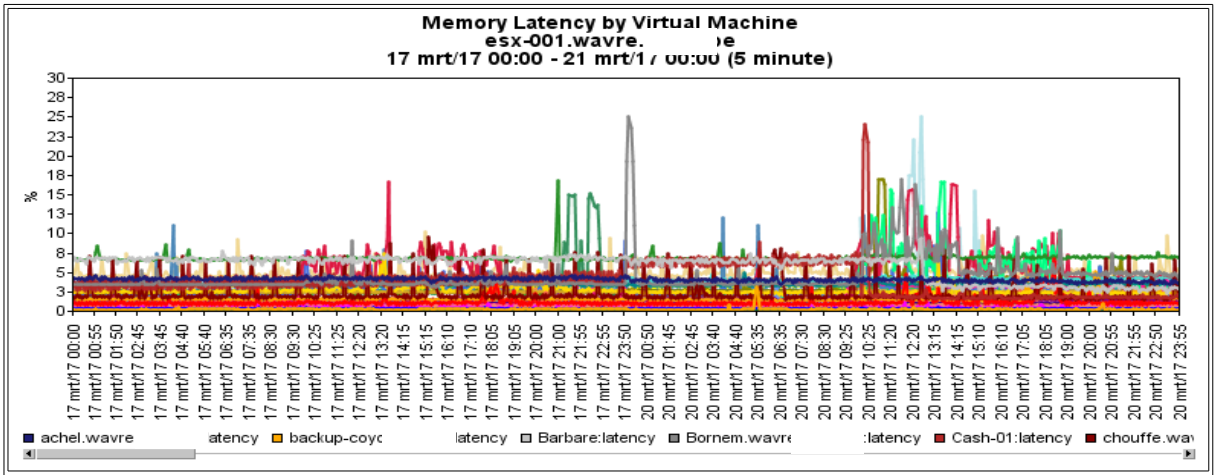

Figure 5-8: ESX-01 VM Memory Latency

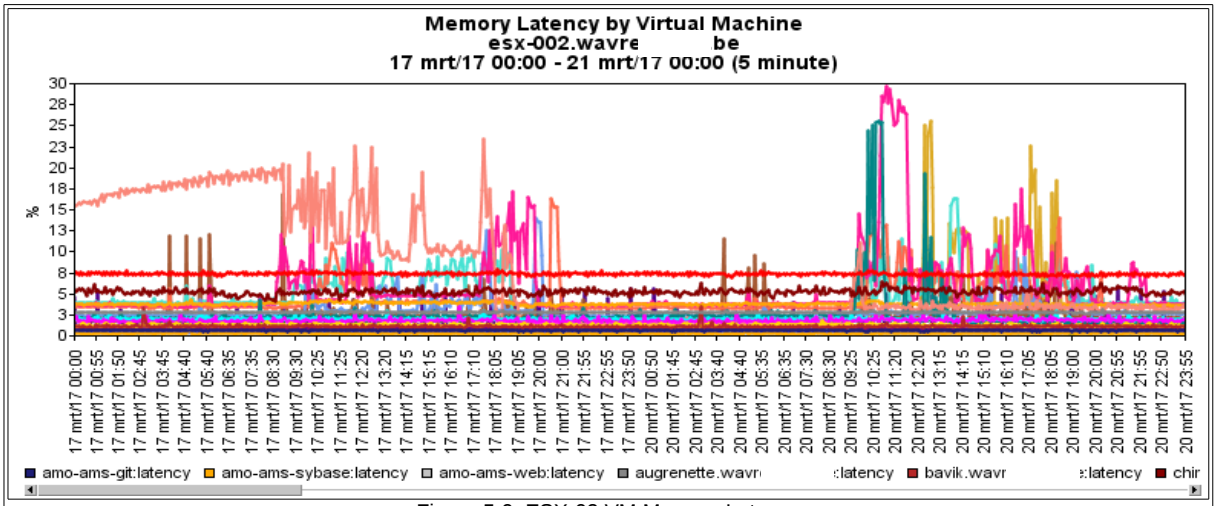

Figure 5-9: ESX-02 VM Memory Latency

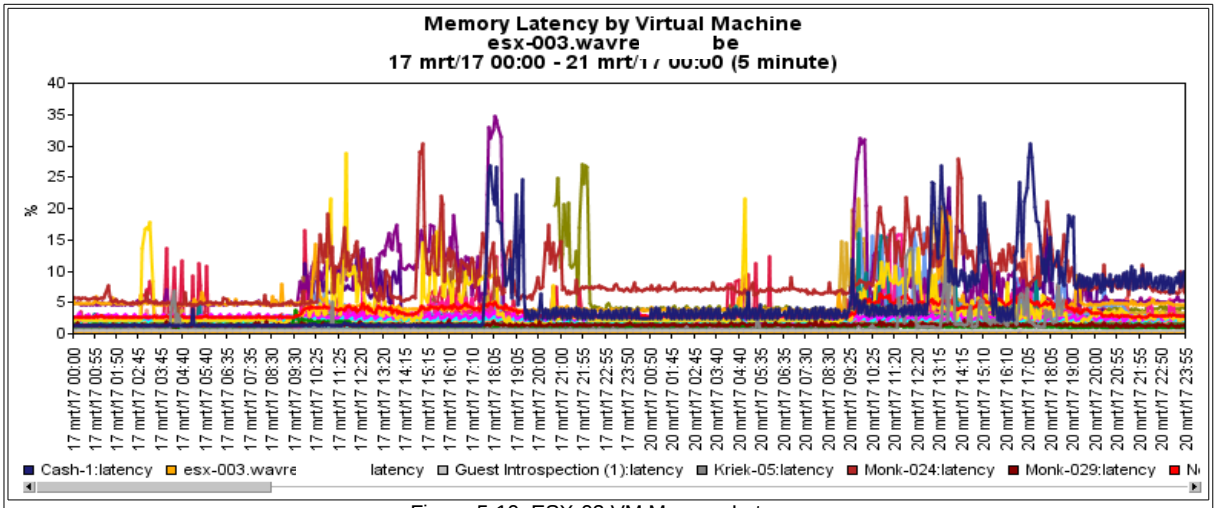

Figure 5-10: ESX-03 VM Memory Latency

This “latency” is too high and should be investigated further, see impact of changes for the CPU “latency”.

No “swapin”, ”swapout”, “swapwait” (The amount of time in seconds the virtual machine was waiting for swap page-ins.

CPU swap wait is included in CPU wait) nor “memctl” (the amount of memory in megabytes currently reclaimed using vmmemctl for the virtual machine) activity was measured.

6. I/O Resource

Here we look at the different levels at the configuration & consumption of the I/O subsystem.

6.1 ESX Host Level

We first look at the “maxTotalLatency” representing the highest latency value across all disks used by the host. Latency measures the time taken to process a SCSI command issued by the guest OS to the virtual machine.

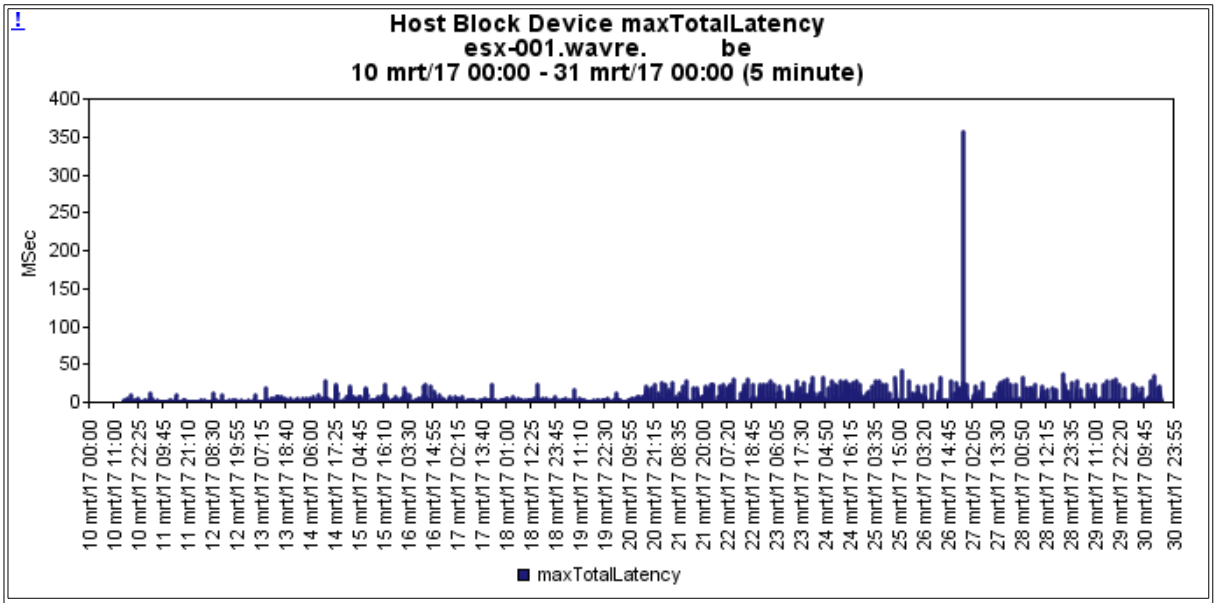

Figure 6-1: ESX-01 Host maxTotalLatency

Host “ESX-01” encountered one severe peak on the 26th 22h40. Apart from that peak we have a quite good latency of 5 to 10MSec up to the 20th afterwards it increased towards 25 MSec which is rather high for a SAN I/O subsystem, see a zoomed chart below.

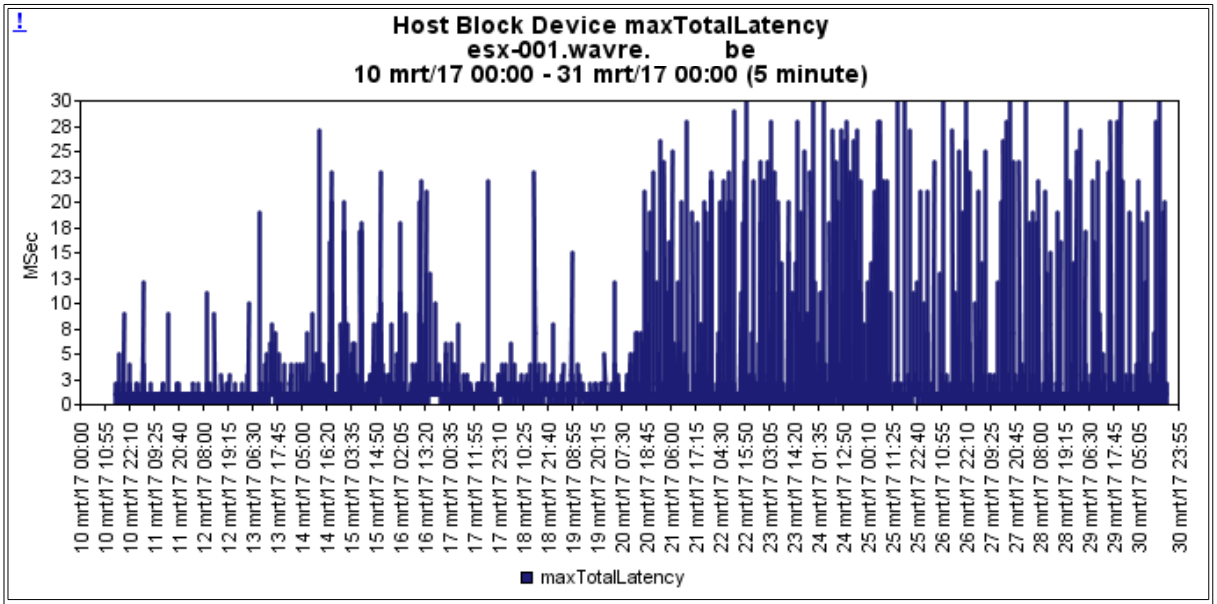

Figure 6-2: ESX-02 Host maxTotalLatency

Both the hosts “ESX-02/03” have an excellent low “maxTotalLatency” of about 5 MSec which is excellent.

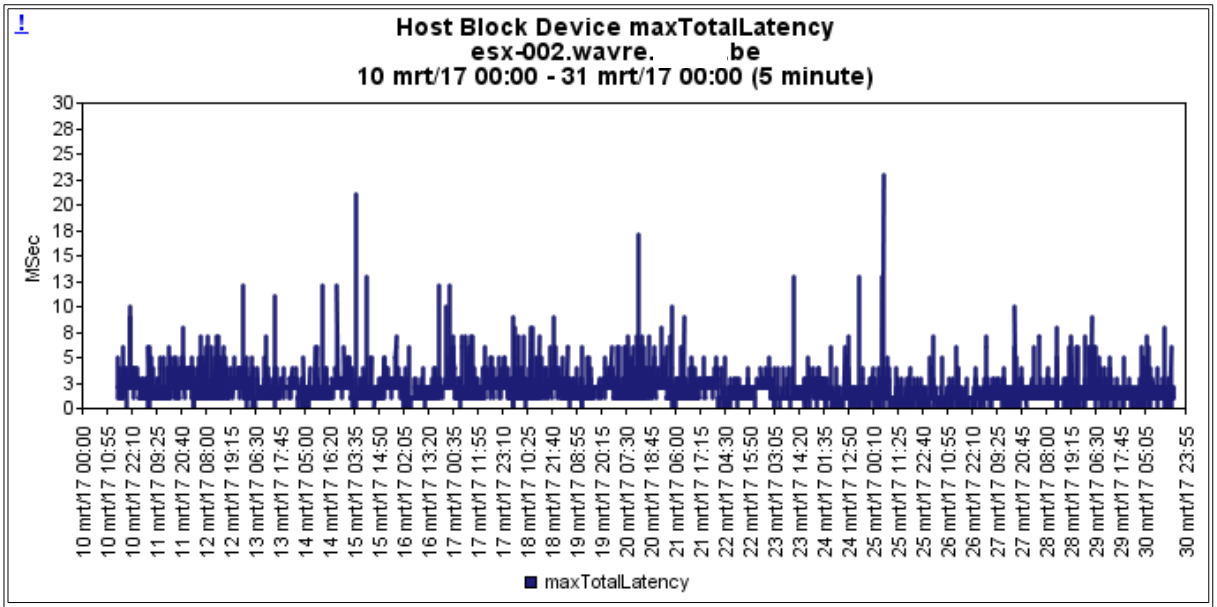

Figure 6-3: ESX-02 Host maxTotalLatency

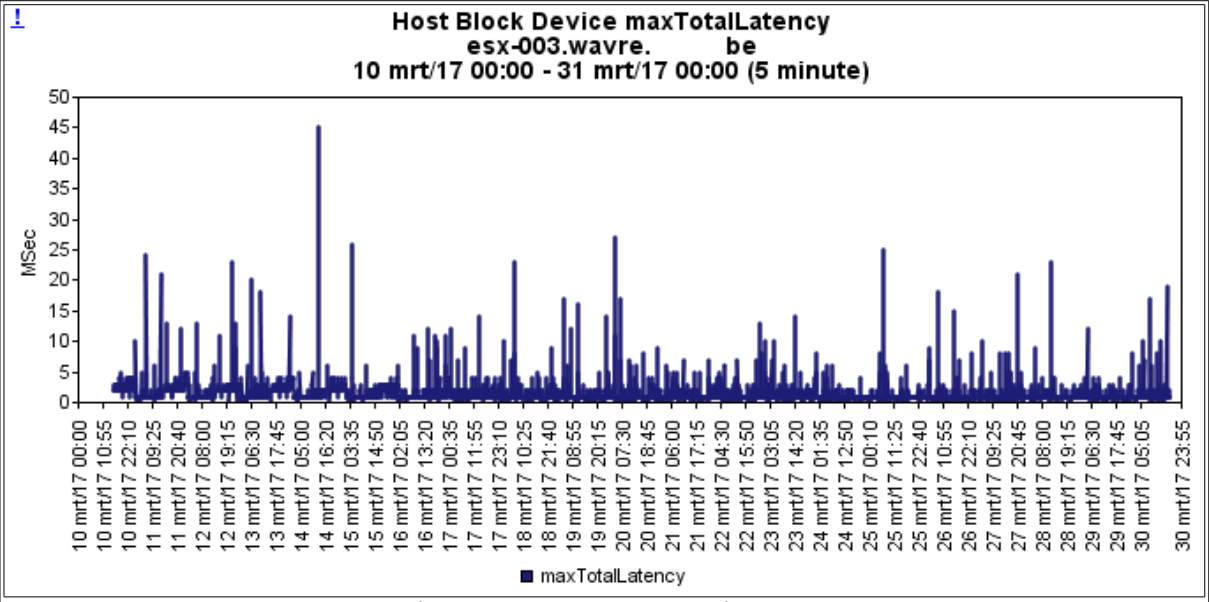

Figure 6-4: ESX-03 Host maxTotalLatency

When looking at the average latency per host & devices representing the average amount of time in milliseconds taken to complete a command request (queue and disk service time) by the host system disk, we see a very well performing I/O subsystem since it is around 5MSec.

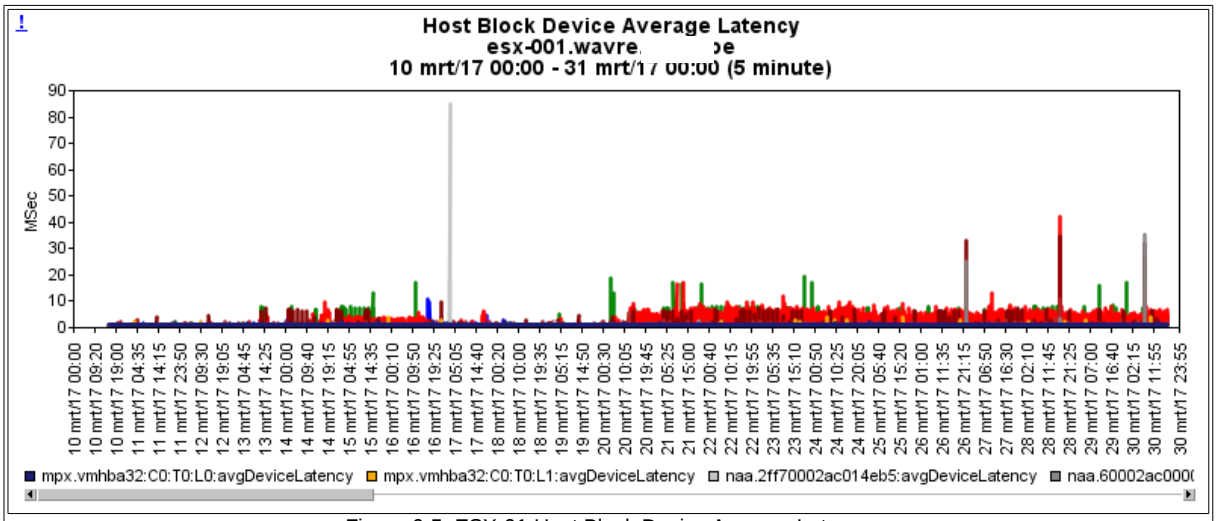

Figure 6-5: ESX-01 Host Block Device Average Latency

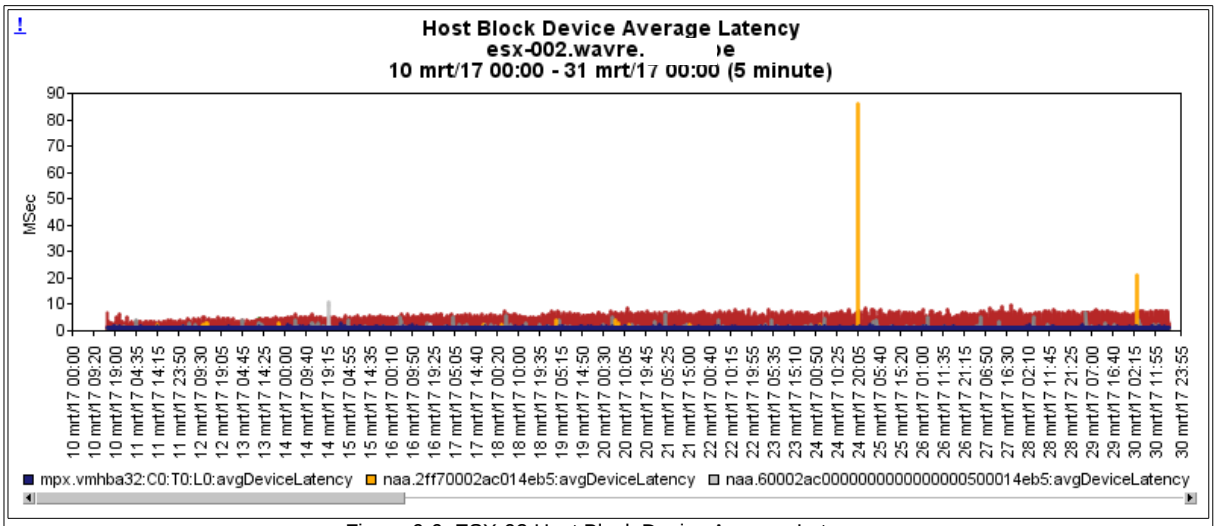

Figure 6-6: ESX-02 Host Block Device Average Latency

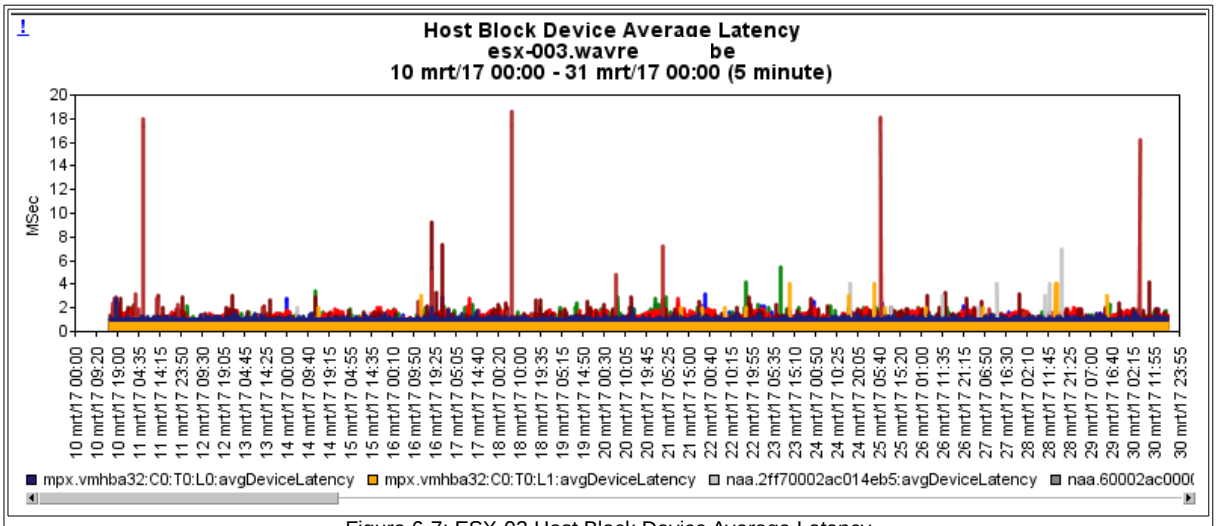

Figure 6-7: ESX-03 Host Block Device Average Latency

6.1 ESX VM Level

Here we look at the average latency per VM & block devices.

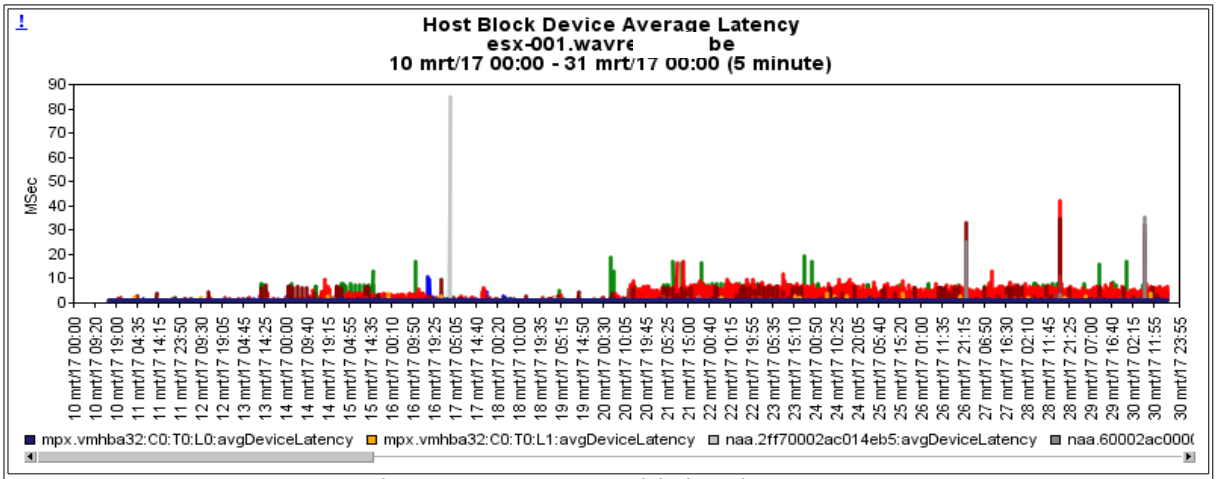

Figure 6-8: ESX-01 Host VM/Block Device Latency

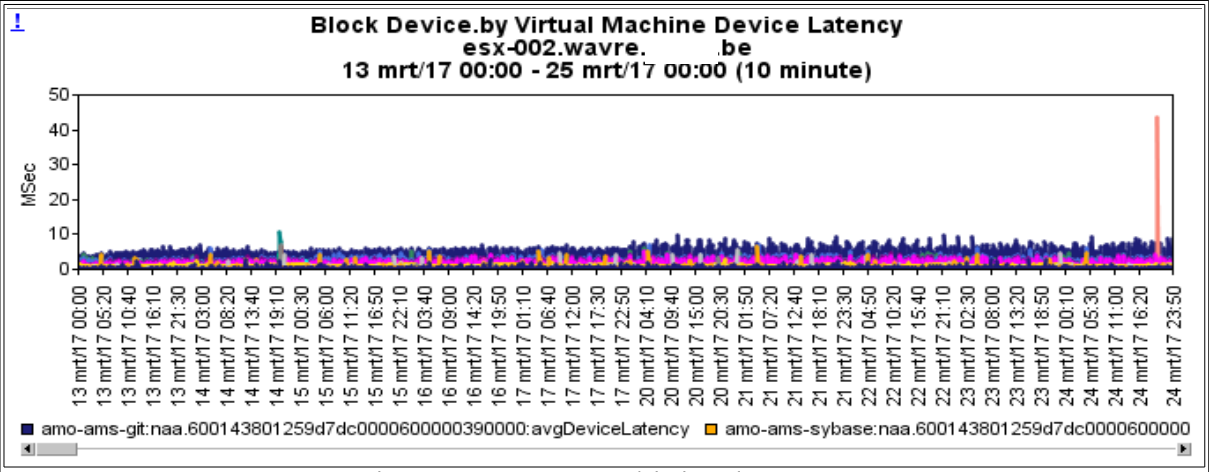

Figure 6-9: ESX-02 Host VM/Block Device Latency

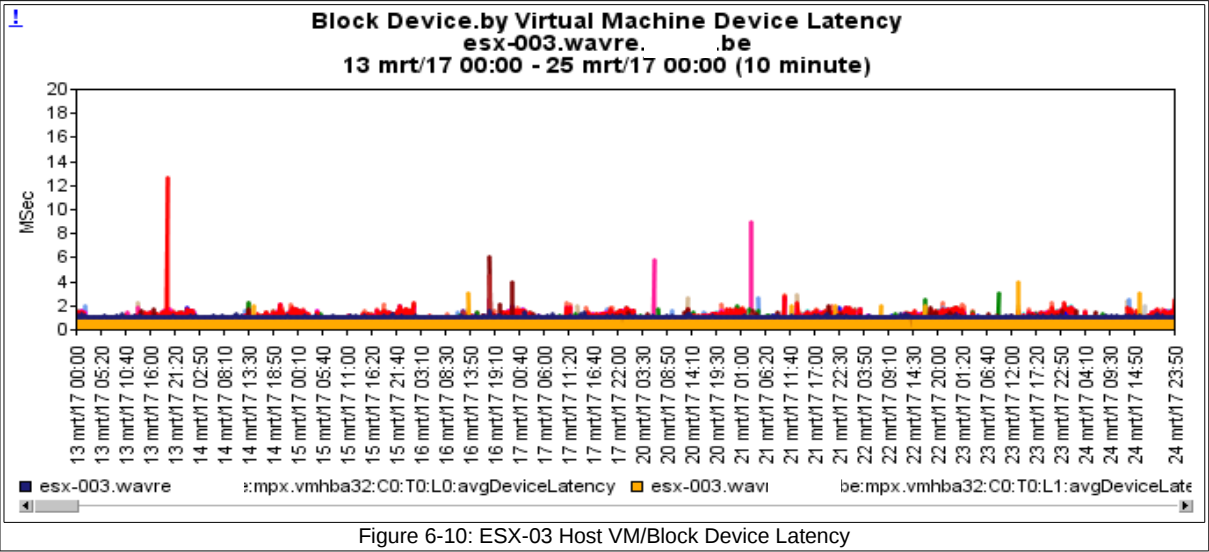

Figure 6-10: ESX-03 Host VM/Block Device Latency

7. Network Resource

Here we look at the LAN activity.

7.1 ESX Host Level

Looking at the throughput of all LAN cards per ESX Host we see that there is sufficient bandwidth left.

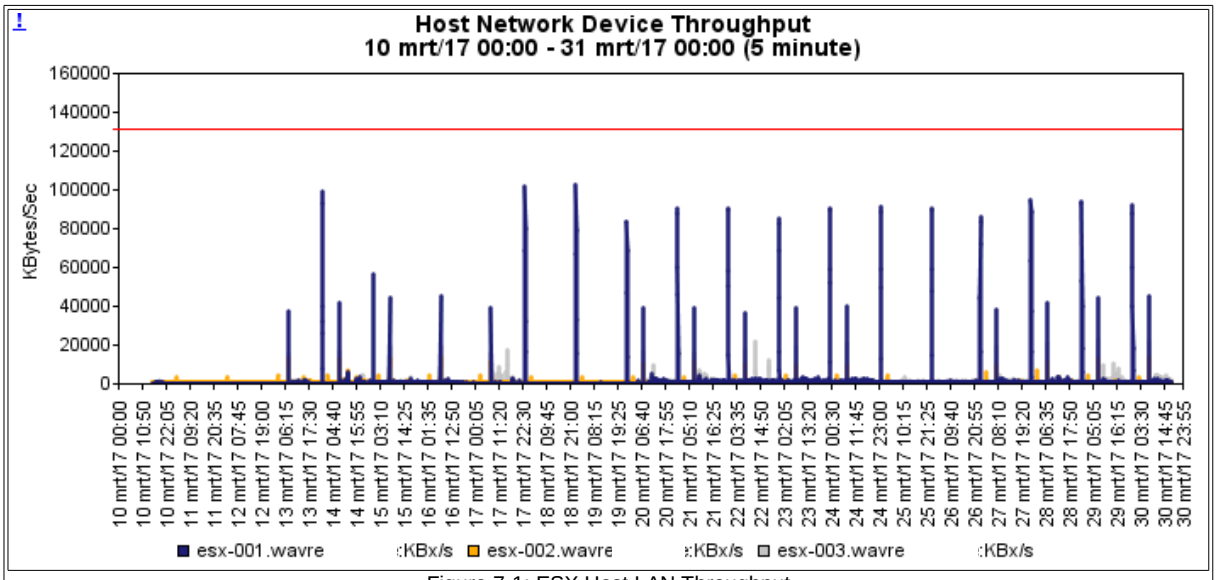

Figure 7-1: ESX Host LAN Throughput

We have a small activity with some peaks but still remaining below the theoretical limit of 1Gbit/Sec (red line).

There where no LAN packets dropped nor errors nor frames with unknown protocols received.

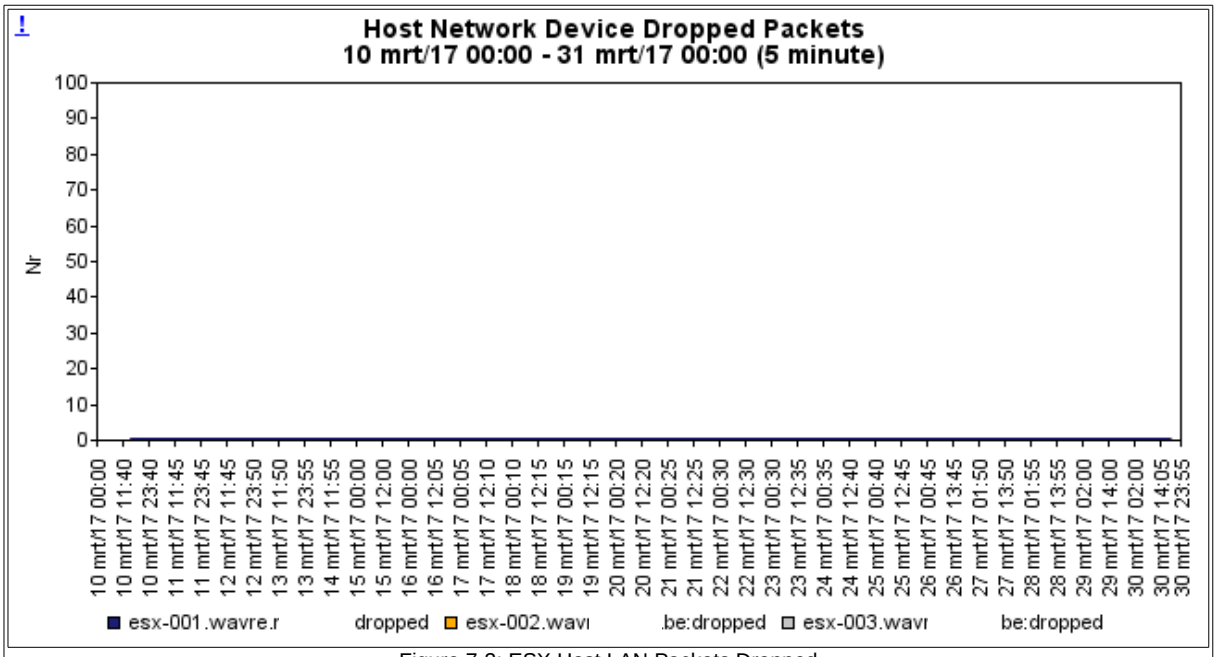

Figure 7-2: ESX Host LAN Packets Dropped

The following three graphs show the LAN throughput per ESX Host and nic's.

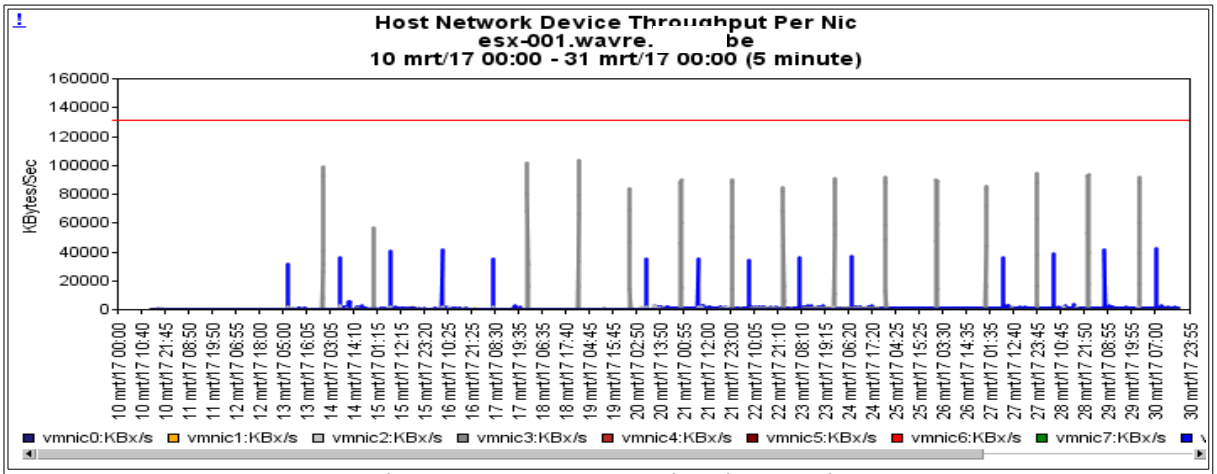

Figure 7-3: ESX-01 Host LAN Throughput Per Nic

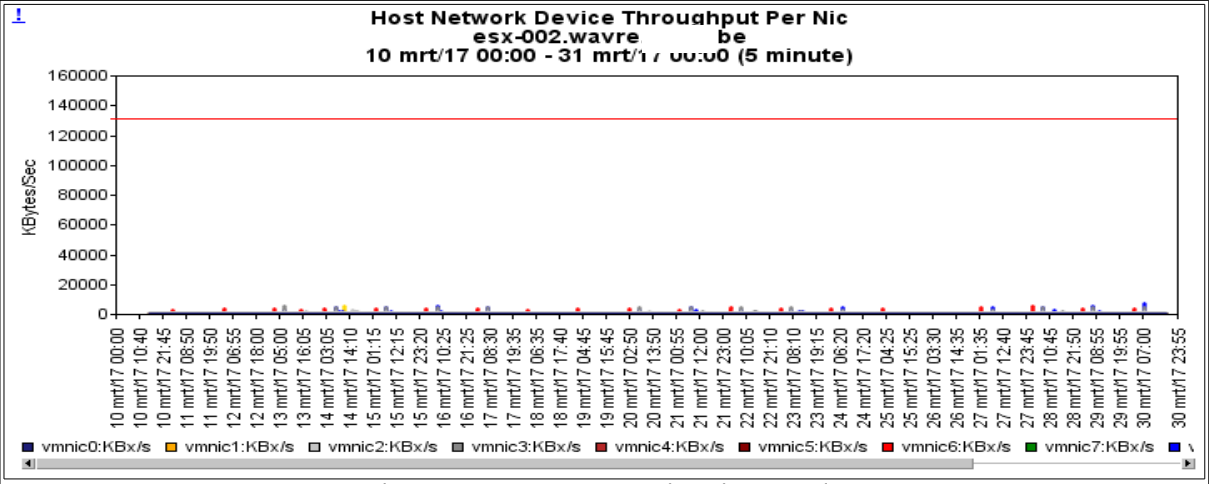

Figure 7-4: ESX-02 Host LAN Throughput Per Nic

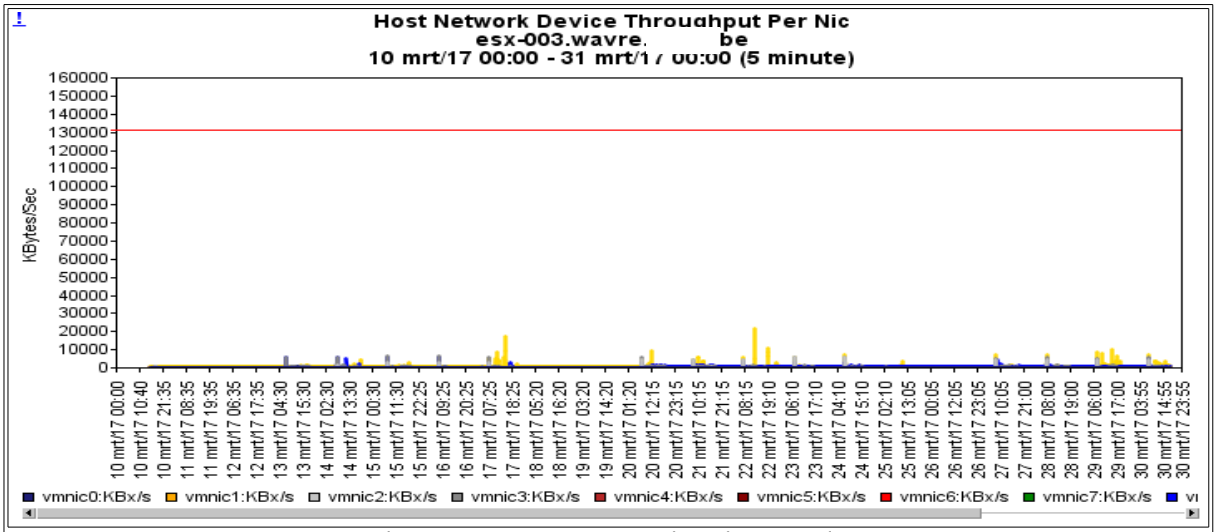

Figure 7-5: ESX-03 Host LAN Throughput Per Nic

So this resource is doing well.

7.1 ESX VM Level

We look at the bandwidth usage again but now for all VM's per ESX Host.

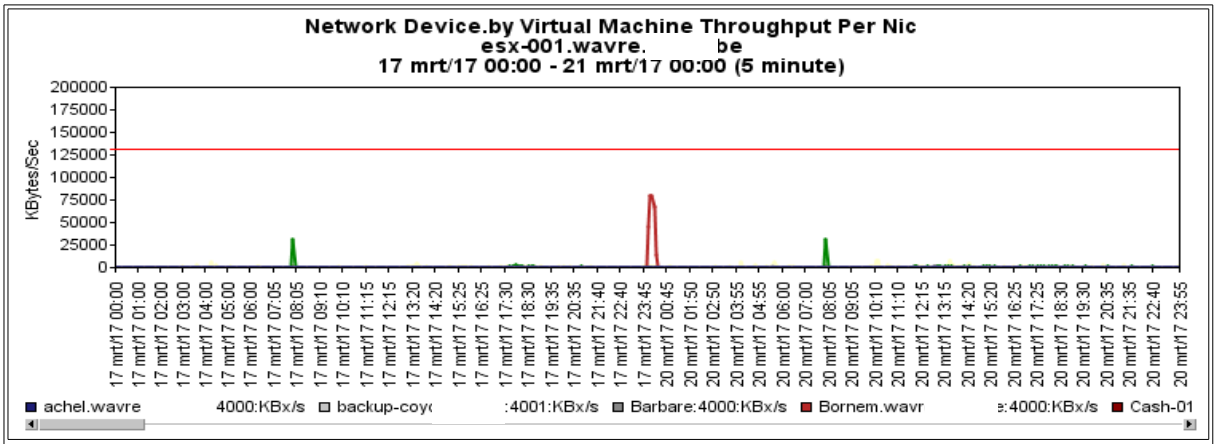

Figure 7-6: ESX-01 Host LAN Throughput Per VM Per Nic

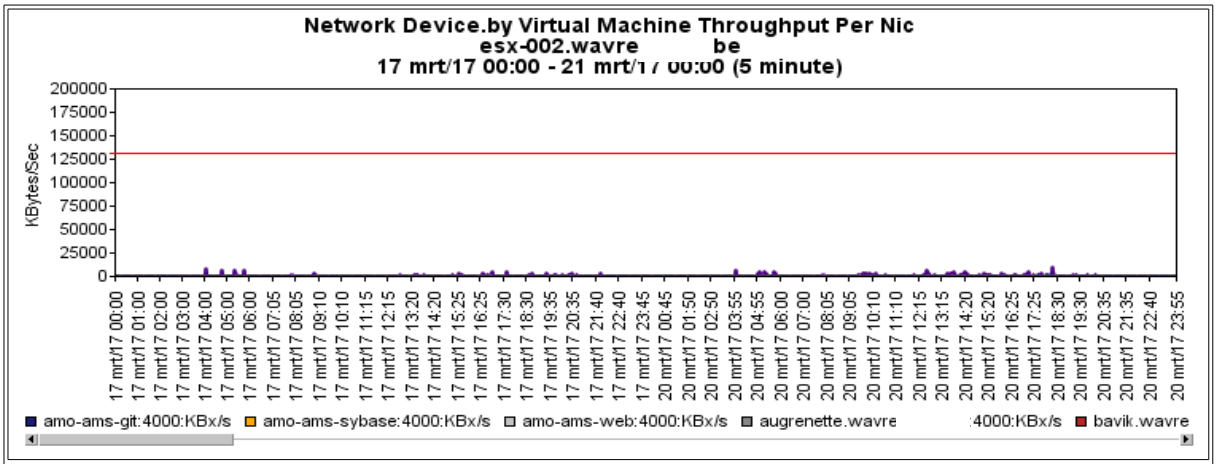

Figure 7-7: ESX-02 Host LAN Throughput Per VM Per Nic

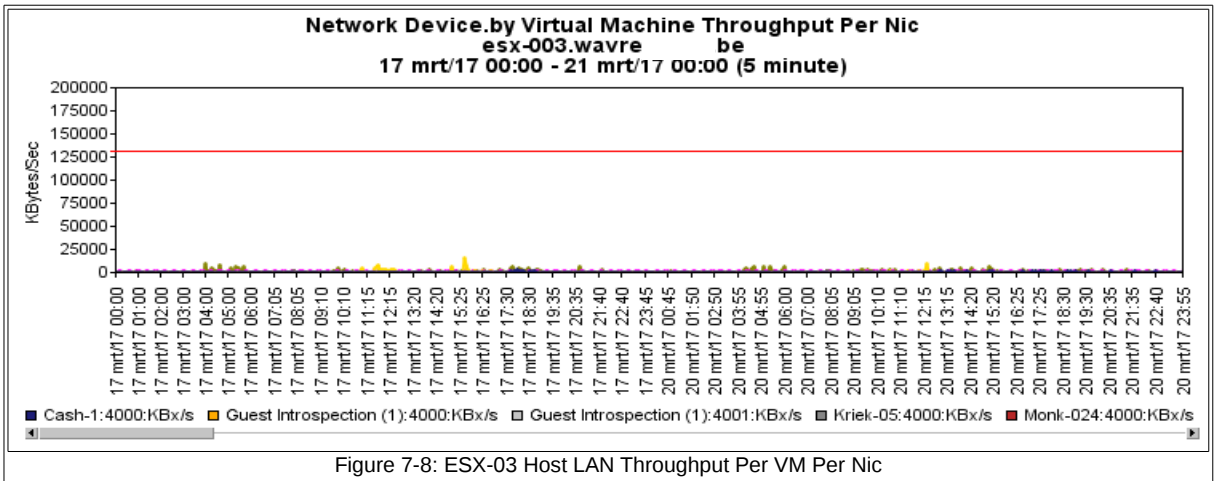

Figure 7-8: ESX-03 Host LAN Throughput Per VM Per Nic

As we already know, no issues with this resource.

-END-