

Min HW Req for 100G PS

- Goal
- Method
- Ref
 - Spec of 100G PS using VM
 - Change CPU (num of cpu, on/off multithreading, freq)
 - Ref spec for 100G PS

Goal

Find minimum hardware requirement of 100G perfSonar

Method

Change num of CPU, freq of CPU → test

Ref

Spec of 100G PS using VM

3.5GHz 2 core, 16GB RAM

<http://ps-100g-hpr01.stanford.edu/>

RAM	16 GB
More Info	Details ▾
perfSONAR Privacy Policy Close ✕	
Host Details	
Vendor	QEMU
VM Host	Standard PC (Q35 + ICH9, 2009)
CPU / Cores	2 / 2
CPU Speed	3500 MHz
OS	CentOS Linux 7.9.2009 (Core)
perfSONAR version	4.4.4-1.el7
TCP Tuning Info:	
Algorithm	htcp
Max Autotuning (Receive)	2147483647
Max Autotuning (Send)	2147483647
TCP Max Backlog	1000
Max Buffer (Send)	2147483647
Max Buffer (Receive)	2147483647

Change CPU (num of cpu, on/off multithreading, freq)

```
# lscpu
CPU(s):                24
On-line CPU(s) list:   0-23
Model name:            Intel(R) Xeon(R) CPU E5-2643 v4 @ 3.40GHz

# cat /proc/cpuinfo | grep processor | wc -l
24

# echo 0 > /sys/devices/system/cpu/cpu{20,21,22,23}/online

# lscpu
CPU(s):                24
On-line CPU(s) list:   0-19
Off-line CPU(s) list:  20-23

# cat /proc/cpuinfo | grep processor | wc -l
20

# echo 1 > /sys/devices/system/cpu/cpu{20,21,22,23}/online
```

```
# lscpu
CPU(s):                24
On-line CPU(s) list:   0-23

# cat /sys/devices/system/cpu/smt/active
1

# echo off > /sys/devices/system/cpu/smt/control

# lscpu
CPU(s):                24
On-line CPU(s) list:   0-11
Off-line CPU(s) list:  12-23

# cat /proc/cpuinfo | grep processor | wc -l
12
```

```
# echo off > /sys/devices/system/cpu/smt/control

# lscpu
CPU(s):                16
On-line CPU(s) list:   0-15
Off-line CPU(s) list:  8-15
Model name:            Intel(R) Xeon(R) CPU E5-2670 0 @ 2.60GHz
CPU MHz:               1243.188
CPU min MHz:           1200.0000
CPU max MHz:           3300.0000

# cpupower frequency-info
analyzing CPU 0:
driver: intel_pstate
hardware limits: 1.20 GHz - 3.30 GHz
available cpufreq governors: performance powersave
current policy: frequency should be within 1.20 GHz and 3.30 GHz.
                    The governor "powersave" may decide which speed to use within this range.
current CPU frequency: 1.20 GHz (asserted by call to hardware)
boost state support:
Supported: yes
Active: yes
3200 MHz max turbo 4 active cores
3200 MHz max turbo 3 active cores
3300 MHz max turbo 2 active cores
3300 MHz max turbo 1 active cores

# cpupower frequency-set --governor performance
Setting cpu: 0
Setting cpu: 1
...
Setting cpu: 7

# cat /sys/devices/system/cpu/cpu*/cpufreq/scaling_governor | sort -u
performance
```

```
# vi /etc/default/grub
# Append "intel_pstate=disable" to GRUB_CMDLINE_LINUX= option
# update grub via some commands
# reboot

# modprobe cpufreq_userspace
# cpupower frequency-set --governor userspace
# cpupower --cpu all frequency-set --freq 1.5GHz
```

```
# ethtool p6p1
Settings for p6p1:
    Supported ports: [ FIBRE ]
    Supported link modes: ... 100000baseCR4/Full

## check NUMA locality
# cat /sys/class/net/p6p1/device/numa_node
0

## if single socket system, no need to be set
# numactl -N 0 iperf3 -s
```

```
# watch -n 1 grep MHz /proc/cpuinfo
cpu          MHz                :          1199.960
...
cpu          MHz                :          3700.000
...
```

<https://fasterdata.es.net/assets/Papers-and-Publications/100G-Tuning-TechEx2016.tierney.pdf>

Disable hyperthreading from within Linux (no access to BIOS), <https://serverfault.com/questions/235825>

Programmatically disable CPU core, <https://stackoverflow.com/questions/44907731>

How to on/offline CPU on RHEL(CentOS), <https://blog.seabow.pe.kr/?p=7303>

★ 100g Network Adapter Tuning - <https://src.stanford.edu/100g-network-adapter-tuning>

★ (2016) Recent Linux TCP Updates, and how to tune your 100G host - <https://fasterdata.es.net/assets/Papers-and-Publications/100G-Tuning-TechEx2016.tierney.pdf>

★ (2016) Network Performance Tuning, Senior Software Maintenance Engineer Red Hat, Inc. - https://sysadmin.miniconf.org/2016/lca2016-jamie_bainbridge-network_performance_tuning.html#1

Red Hat > Monitoring and managing system status and performance Chapter > 17. Tuning CPU frequency to optimize energy consumption - https://access.redhat.com/documentation/enus/red_hat_enterprise_linux/8/html/monitoring_and_managing_system_status_and_performance/tuning-cpu-frequency-to-optimize-energy-consumption_monitoring-and-managing-system-status-and-performance

How to change and check cpu frequency governors - <https://docs.altibase.com/pages/viewpage.action?pageId=13436485>

CPU Frequency Scaling On Linux With cpupower - <https://krimson8.github.io/linux/2019/03/13/Configuring-CPU-frequency.html>

How to solve the cpufreq errors - <https://stackoverflow.com/questions/23526671>

How to set init.d (rc.d) for permanent settings - <https://docs.altibase.com/pages/viewpage.action?pageId=13436485> / <https://kwonnam.pe.kr/wiki/linux/cpufreq>

Optimize iperf Performance of External Loopback in Dual-port Design - <https://github.com/corundum/corundum/issues/33>

Ref spec for 100G PS

perfSONAR Deployment Examples > 100G Configuration - https://docs.perfsonar.net/deployment_examples.html

- **CPU:** 2 x Intel Xeon Haswell **E5-2643V3 3.4GHz** 6 Cores (Total 12 Threads each)
- **Memory:** 8 x 16GB DDR4-2133MHz RAM ECC/REG (**128GB** total)
- **Network:** Mellanox MCX455A-ECAT **100Gbps NIC**

perfSONAR Node Reference Implementation - <https://fasterdata.es.net/performance-testing/perfsonar/perfsonar-node-reference-implementation>

- Current (~2021/22) 40/50/100Gbps Capable Design 1
 - The total cost of this server was around \$4K in mid 2021.
 - **CPU:** AMD EPYC 7302P 16-Core Processor (**3.0GHz base clock**, 3.3GHz max boost)
 - **Memory:** 8 x 8GB DDR4 ECC/REG RAM (**64G Total**)
 - **Network:** Mellanox MCX556A-ECAT Dual Port Connect X5
- Current (~2021/22) 40/50/100Gbps Capable Design 2
 - We also tested and liked the following system, but due to supply chain issues, went with the above system.
 - **CPU:** Intel Xeon Silver 4215R **8C 3.2Ghz**
 - **Memory:** 6 x 8GB DDR4 ECC/REG RAM (**48G Total**)
 - **Network:** Mellanox MCX556A-ECAT Dual Port Connect X5