Tune System Performance

Unit objectives

After completing this unit, you should be able to:

- Setting a `Tuning` Profile
- Influencing `Process Scheduling`

Understand performance optimization

- Workload's requirement
- Adjusting settings based on requirement
- The tuned daemon
- Tuning profiles

Type of Tuning

- Static Tuning
- Dynamic Tuning

Configuring Static Tuning

- When Service starts or
- Upon selection of new tuning profile
- Configures predefined kernel parameters
- Set for overall performance expectations
- Does not change over time

Configuring Dynamic Tuning

- The **tuned** daemon monitors system activity
- Adjust settings depending behavior change
- Continously monitor and adjust tuning fit workload's requirement
- Example of behavior changes:
  - Storage device experience high usage during startup, but minimal user activities.
  - Network device experience high usage during peak hours throughout the days.
- The tuned daemon capture these activities and adjust parameters to
  - allocate more CPU resource to storage during startup and
  - re-allocate CPU to Network device during the days

Tuning plug-ins

- `disk`: Sets different disk parameters, for example, the disk scheduler, the spin-down timeout, or the advanced power management.
- `net`: Configures the interface speed and the Wake-on-LAN (WoL) functionality.
- `cpu`: Sets different CPU parameters, for example, the CPU governor or the latency.
- Dynamic tuning is disabled.
- Enabled it by modifying /etc/tuned/tuned-main.conf
  - dynamic_tuning variable to 1

```
[root@host ~]$ cat /etc/tuned/tuned-main.conf
...output omitted...
# Dynamicaly tune devices, if disabled only static tuning will be used.
dynamic_tuning = 1
...output omitted...
# Update interval for dynamic tunings (in seconds).
# It must be multiply of the sleep_interval.
update_interval = 10
```

Installing and enabling tuned

```
# dnf install tuned

# systemctl enable --now tuned

# systemctl status tuned
```

Selecting a Tuning Profile 1/2

- with predefined settings
- Focus on
  - Low latency for storage and network
  - High throughput for storage and network
  - Virtual Machine performance
  - Virtualization host performance

Profiles distributed with RHEL v8/v9

Tuned Profile	Purpose
balanced	Ideal for systems that require a compromise between power saving and performance.
desktop	Derived from the **balanced** profile. Provides faster response of interactive applications.
throughput-performance	Tunes the system for maximum throughput.
latency-performance	Ideal for server systems that require low latency at the expense of power consumption.
network-latency	Derived from the **latency-performance** profile. It enables additional network tuning parameters to provide low network latency.
network-throughput	Derived from the **throughput-performance** profile. Additional network tuning parameters are applied for maximum network throughput.
powersave	Tunes the system for maximum power saving.
oracle	Optimized for Oracle database loads based on the **throughput-performance** profile.
virtual-guest	Tunes the system for maximum performance if it runs on a virtual machine.
virtual-host	Tunes the system for maximum performance if it acts as a host for virtual machines.

Managing profiles from command line

- List all available profiles

```
# tuned-adm list
```

- View current activated profile

```
# tuned-adm active
```

- Switch profile

```
# tuned-adm profile throughput-performance
```

- Ask for recommendation

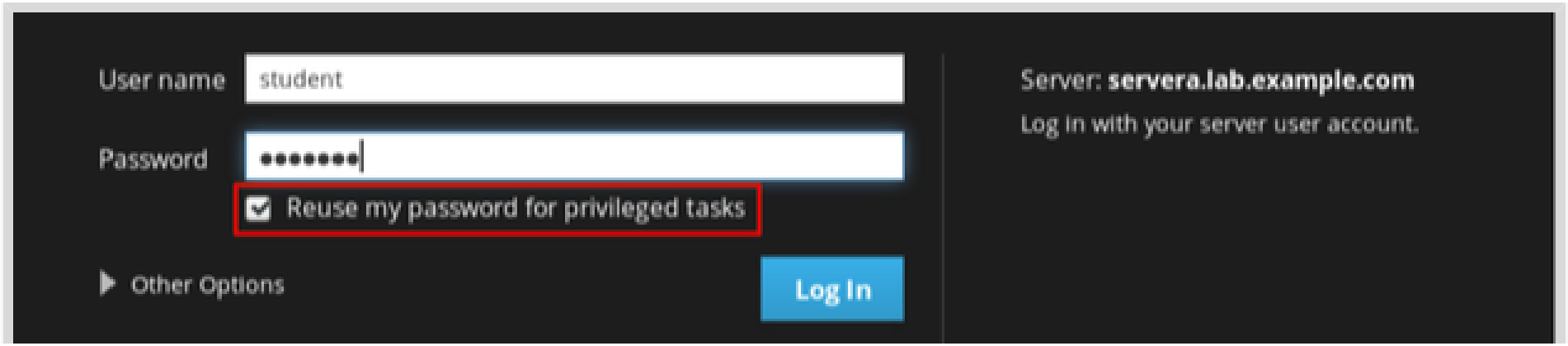
```
# tuned-adm recommend
```

- Disable tuning activity

```
# tuned-adm off
```

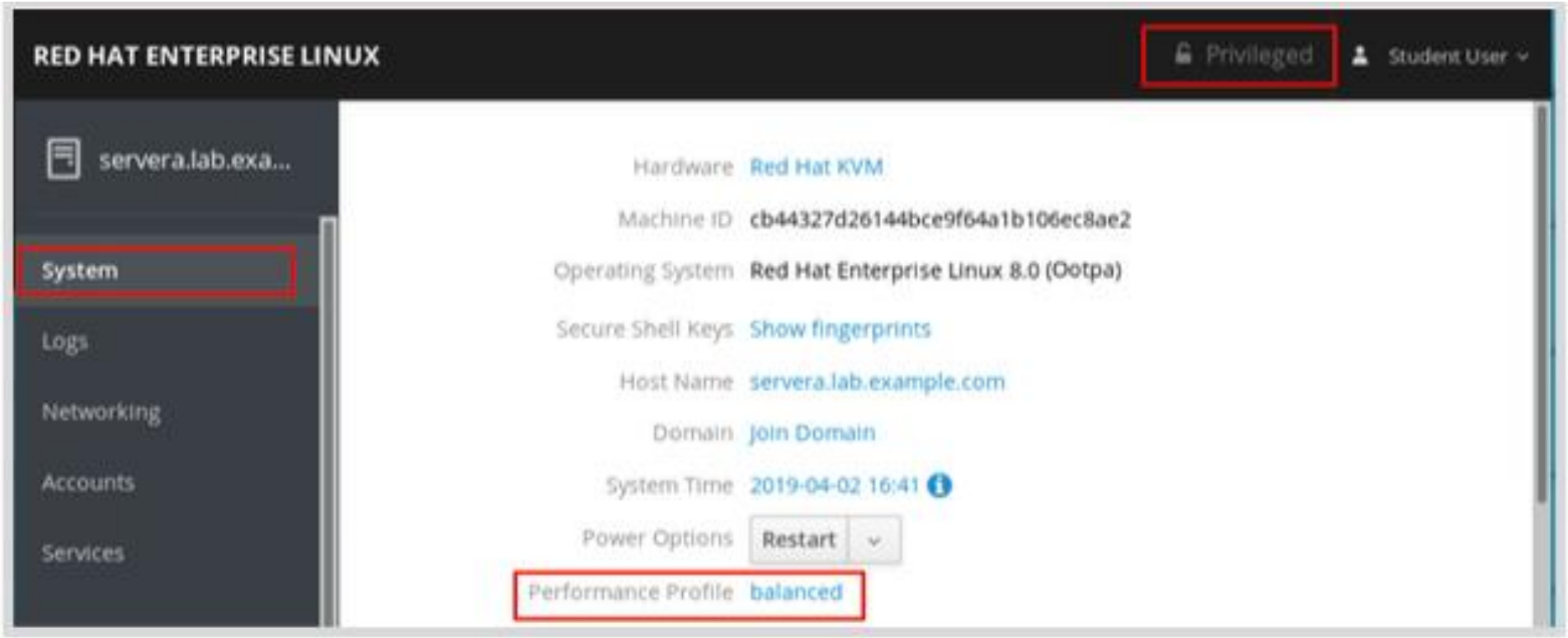
Managing profiles from Web Console

- For non-root, enable "Reuse my password for privileged tasks

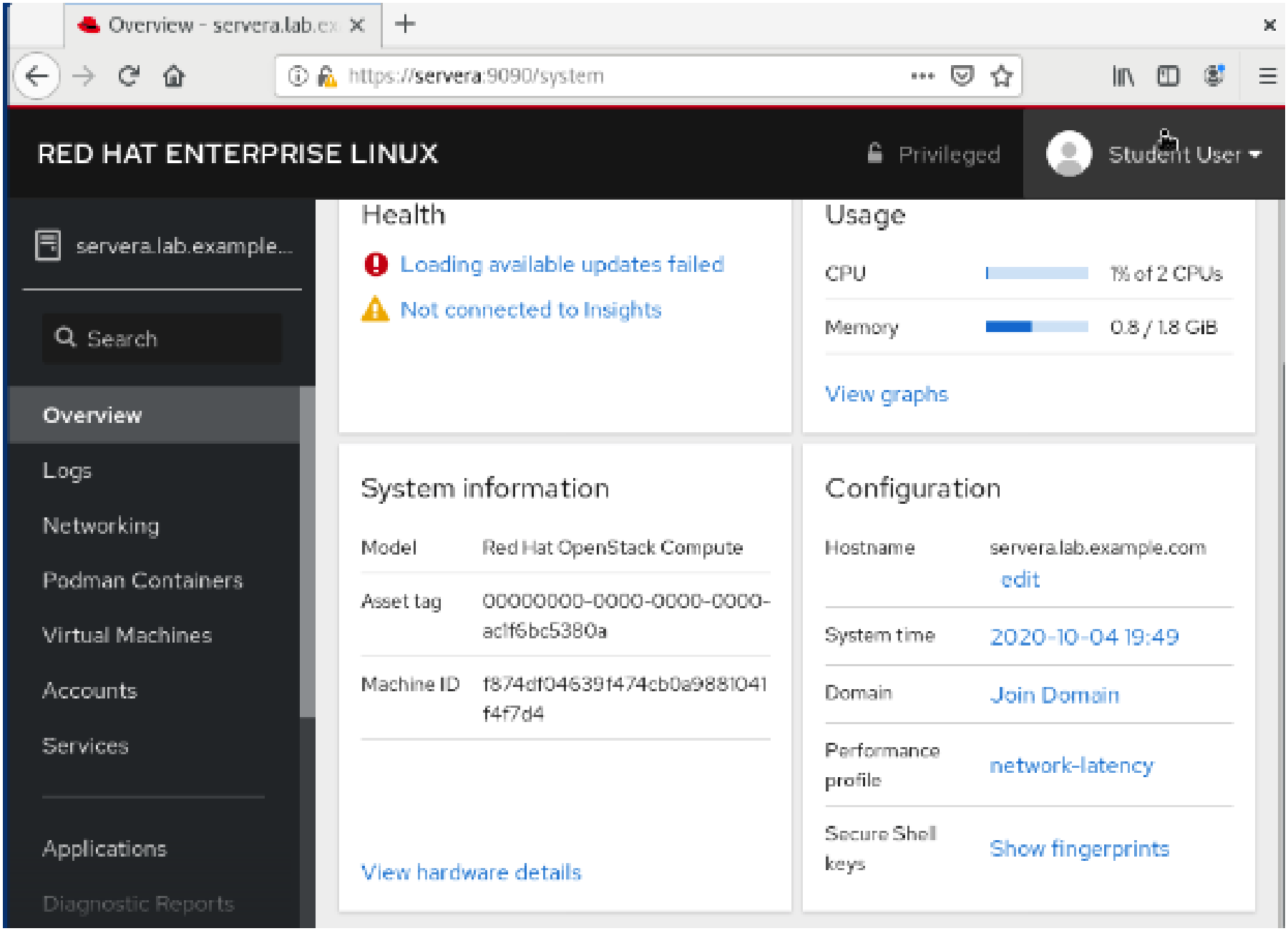


Managing profiles from Web Console

- Click System > Enter Performance Profile

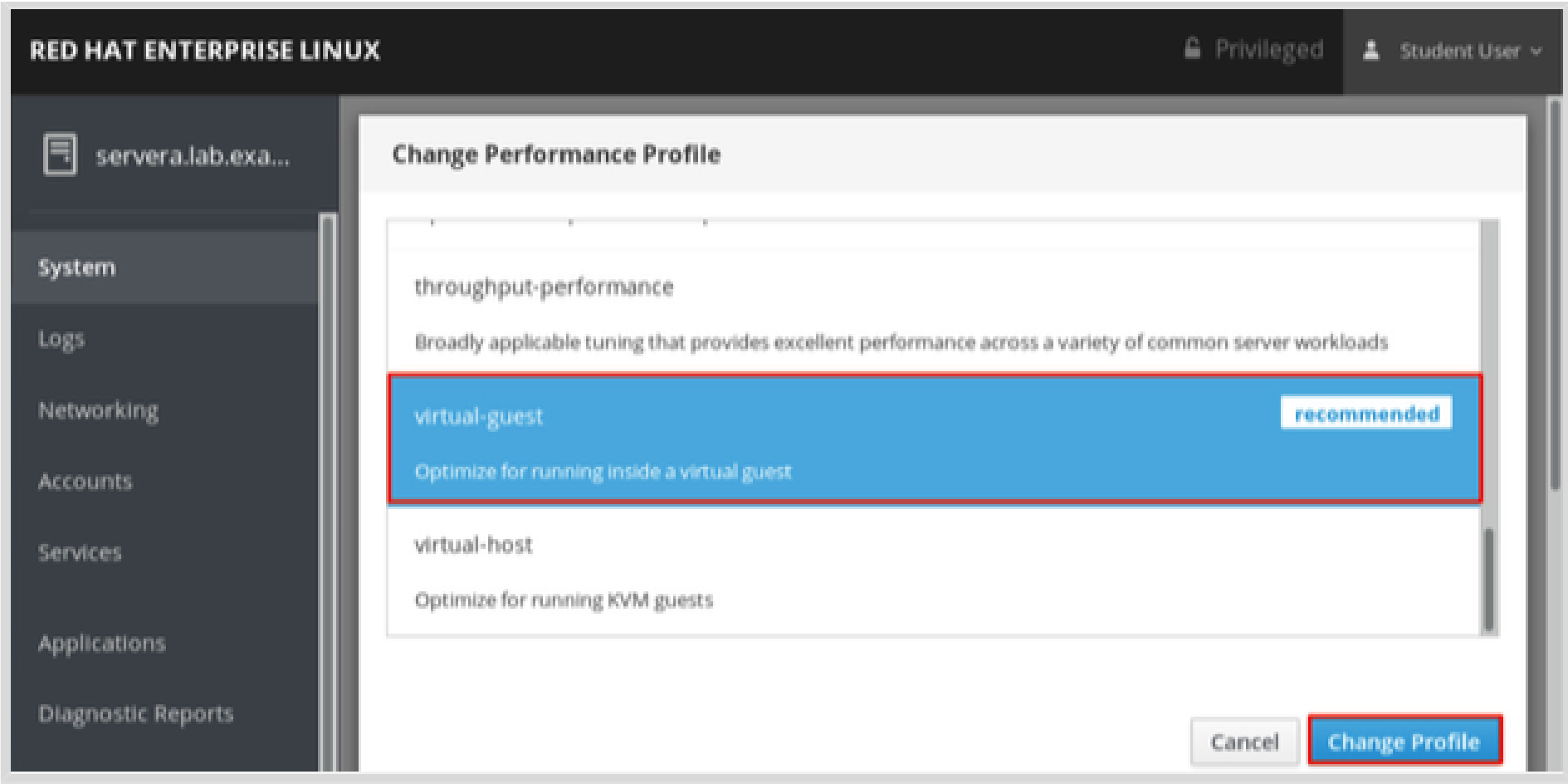


Managing profiles from Web Console v8.2



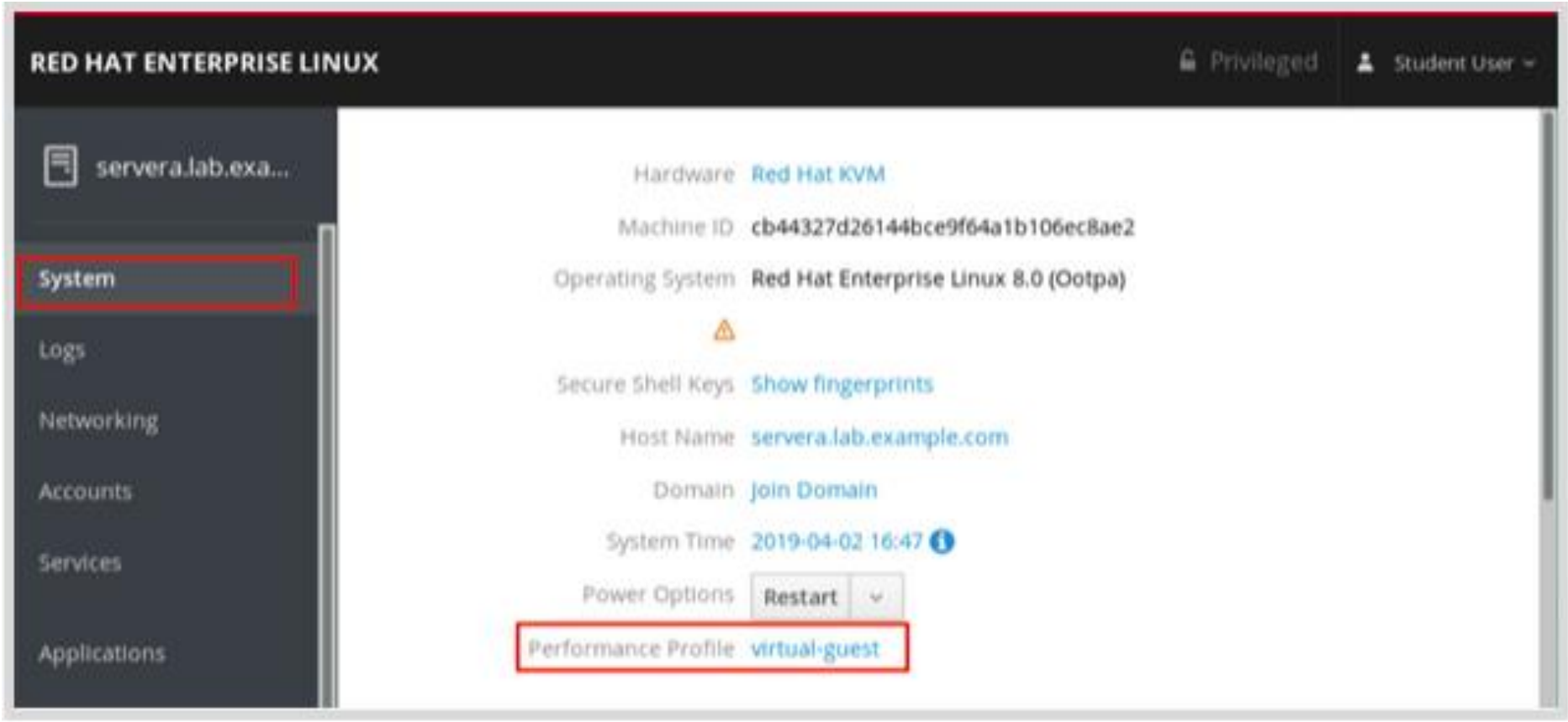
Managing profiles from Web Console

- Select a profile > Click Change Profile



Managing profiles from Web Console

- Verify changes



Influencing Process Scheduling

- Understand Linux Process Scheduling and Multitasking
- Relative Priorities
- Setting Nice Levels and Permissions

Understand Linux Process Scheduling and Multitasking

- Low end to High end Systems
- Single to Multiple CPUs
- CPU to execute instructions
  - dozens to millions
  - parallel
- Common traits: Requirement to run more process threads than CPU
- Linuxes uses time-slicing or multitasking
- Process Scheduler
  - Switch processes in/out single core
  - Multiple processes running simultaneously
  - Has multiple scheduling policies for different processes

Relative Priorities

- Each process given level of importance
- Process Scheduler
  - Policy : SCHED_NORMAL, SCHED_FIFO, SCHED_RR, and SCHED_OTHER
  - default to SCHED_OTHER
- Use PR and NICE to determine importance
- NICE range -20 (highest) to 19 (lowest)
- NICE value 19, often don't get CPU time when system is busy
- Default NICE 0
- Child process inherit nice from parent
- No contention for CPU, higher NICE process may use all CPU

Nice Levels and Permissions

- Settings NICE
  - Only root can adjust process NICE level
  - Non-root users permitted to increase process NICE level of their own processes

- Reporting Nice Levels
  - PR : Scheduled priority : Set by Kernel
  - NI : NICE level : Set by user

Displaying NICE levels using TOP command

- Result in real time

```
top - 08:10:36 up  5:28,  2 users,  load average: 0.38, 0.29, 0.24
Tasks: 300 total,   7 running, 293 sleeping,   0 stopped,   0 zombie
%Cpu(s):  0.1 us,  1.8 sy,  0.0 ni, 95.6 id,  0.0 wa,  2.4 hi,  0.2 si,  0.0 st
MiB Mem :   1800.6 total,    157.4 free,   1295.2 used,    348.1 buff/cache
MiB Swap:   2048.0 total,   2029.0 free,     19.0 used.    319.5 avail Mem
```

PID	USER	PR	NI	VIRT	RES	SHR	S	%CPU	%MEM	TIME+	COMMAND
7072	root	20	0	0	0	0	S	2.3	0.0	4:51.49	kvdo0:indexW
7073	root	20	0	0	0	0	S	1.3	0.0	4:31.10	kvdo0:indexW
5805	**root**	**20**	**0**	**0**	**0**	**0**	**R**	**1.0**	**0.0**	**2:54.92**	**callbackW**
5999	root	20	0	0	0	0	S	1.0	0.0	3:02.67	kvdo1:indexW
6000	root	20	0	0	0	0	S	1.0	0.0	4:44.57	kvdo1:indexW
5991	root	20	0	0	0	0	S	0.3	0.0	0:24.35	kvdo1:cpuQ1
1	root	20	0	246028	12080	7048	S	0.0	0.7	0:06.26	systemd
2	root	20	0	0	0	0	S	0.0	0.0	0:00.05	kthreadd
3	root	0	-20	0	0	0	I	0.0	0.0	0:00.00	rcu_gp
4	root	0	-20	0	0	0	I	0.0	0.0	0:00.00	rcu_par_gp
6	root	0	-20	0	0	0	I	0.0	0.0	0:00.00	kworker/0:0H
8	root	0	-20	0	0	0	I	0.0	0.0	0:00.00	mm_percpu_wq
9	root	20	0	0	0	0	S	0.0	0.0	0:01.10	ksoftirqd/0
10	**root**	**20**	**0**	**0**	**0**	**0**	**R**	**0.0**	**0.0**	**0:00.87**	**rcu_sched**
11	root	rt	0	0	0	0	S	0.0	0.0	0:00.00	migration/0
12	root	rt	0	0	0	0	S	0.0	0.0	0:00.00	watchdog/0
13	root	20	0	0	0	0	S	0.0	0.0	0:00.00	cpuhp/0
14	root	20	0	0	0	0	S	0.0	0.0	0:00.00	cpuhp/1
15	root	rt	0	0	0	0	S	0.0	0.0	0:00.00	watchdog/1
16	root	rt	0	0	0	0	S	0.0	0.0	0:00.00	migration/1
17	root	20	0	0	0	0	S	0.0	0.0	0:00.37	ksoftirqd/1

Displaying NICE levels using PS command

- Show unsorted process with extra information

```
# ps axo pid,comm,nice,cls
```

- Show process with extra information and sort by NICE ascendingly

```
# ps axo pid,comm,nice,cls --sort=nice
```

- Show process with extra information and sort by NICE descendingly

```
# ps axo pid,comm,nice,cls --sort=-nice
```

```
  PID COMMAND             NI CLS
    1 systemd              0  TS
    2 kthreadd             0  TS
    3 rcu_gp             -20  TS
    4 rcu_par_gp         -20  TS
    6 kworker/0:0H       -20  TS
    8 mm_percpu_wq       -20  TS
    9 ksoftirqd/0          0  TS
   10 rcu_sched            0  TS
   11 migration/0          -  FF
   12 watchdog/0           -  FF
   13 cpuhp/0              0  TS
   14 cpuhp/1              0  TS
   15 watchdog/1           -  FF
   16 migration/1          -  FF
   17 ksoftirqd/1          0  TS
   19 kworker/1:0H-kb   -20  TS
   20 cpuhp/2              0  TS
   21 watchdog/2           -  FF
   22 migration/2          -  FF
   23 ksoftirqd/2          0  TS
   25 kworker/2:0H-kb   -20  TS
   26 cpuhp/3              0  TS
   27 watchdog/3           -  FF
   28 migration/3          -  FF
   29 ksoftirqd/3          0  TS
```

Displaying NICE levels using HTOP command

```
  1  [|||                                                    4.8%]   Tasks: 59, 65 thr; 4 running
  2  [|||                                                    4.8%]   Load average: 0.11 0.28 0.28
  3  [|||                                                    5.4%]   Uptime: 05:56:04
  4  [||                                                     4.0%]
Mem[||||||||||||||||||||||||||||||||||||||||||||||||||||1.27G/1.76G]
Swp[|                                                  28.8M/2.00G]

    PID USER       PRI  NI  VIRT   RES   SHR S CPU% MEM%   TIME+  Command
  14384 root        20   0 1153M 17516 13004 S  0.0  0.9  0:00.00 /usr/libexec/udisks2/udisksd
  14280 root        20   0 1153M 17516 13004 S  0.0  0.9  0:00.00 /usr/libexec/udisks2/udisksd
  14267 root        20   0 1153M 17516 13004 S  0.0  0.9  0:00.00 /usr/libexec/udisks2/udisksd
  14265 root        20   0 1153M 17516 13004 S  0.0  0.9  0:00.00 /usr/libexec/udisks2/udisksd
   6845 root        20   0  459M  1640   840 S  0.0  0.1  0:04.53 /usr/bin/vmhgfs-fuse .host:/ /mnt/hgfs -o subtype=v
   6897 root        20   0  459M  1640   840 S  0.0  0.1  0:00.82 /usr/bin/vmhgfs-fuse .host:/ /mnt/hgfs -o subtype=v
   6896 root        20   0  459M  1640   840 S  0.0  0.1  0:00.87 /usr/bin/vmhgfs-fuse .host:/ /mnt/hgfs -o subtype=v
   6895 root        20   0  459M  1640   840 S  0.0  0.1  0:00.89 /usr/bin/vmhgfs-fuse .host:/ /mnt/hgfs -o subtype=v
   6852 root        20   0  459M  1640   840 S  0.0  0.1  0:00.00 /usr/bin/vmhgfs-fuse .host:/ /mnt/hgfs -o subtype=v
   6850 root        20   0  459M  1640   840 S  0.0  0.1  0:00.98 /usr/bin/vmhgfs-fuse .host:/ /mnt/hgfs -o subtype=v
   6849 root        20   0  459M  1640   840 S  0.0  0.1  0:00.92 /usr/bin/vmhgfs-fuse .host:/ /mnt/hgfs -o subtype=v
   5840 root        20   0  195M 28116 12188 S  0.0  1.5  0:01.92 /usr/sbin/dmeventd -f
   5998 root        20   0  195M 28116 12188 S  0.0  1.5  0:00.16 /usr/sbin/dmeventd -f
   5842 root        20   0  195M 28116 12188 S  0.0  1.5  0:01.03 /usr/sbin/dmeventd -f
   2334 root        20   0  186M 10988  9420 S  0.0  0.6  0:00.07 login -- root
  12365 root        20   0 26692  4820  2820 S  0.0  0.3  0:00.04 -bash
   1970 dnsmasq     20   0 71888  1448  1012 S  0.0  0.1  0:00.02 /usr/sbin/dnsmasq --conf-file=/var/lib/libvirt/dnsm
   1971 root        20   0 71860   428     0 S  0.0  0.0  0:00.00 /usr/sbin/dnsmasq --conf-file=/var/lib/libvirt/dnsm
   1662 rtkit       20   0  195M  3316  2980 S  0.0  0.2  0:00.10 /usr/libexec/rtkit-daemon
   1491 user        20   0 93988  5924  4292 S  0.0  0.3  0:00.28 /usr/lib/systemd/systemd --user
   1872 user        20   0 82592  4116  3628 S  0.0  0.2  0:00.01 /usr/bin/dbus-daemon --session --address=systemd: -
   1883 user        20   0 82592  4116  3628 S  0.0  0.2  0:00.00 /usr/bin/dbus-daemon --session --address=systemd: -
   1612 user        20   0  293M  6144  4960 S  0.0  0.3  0:00.08 /usr/bin/pulseaudio --daemonize=no
   1871 user        20   0  293M  6144  4960 S  0.0  0.3  0:00.00 /usr/bin/pulseaudio --daemonize=no
   1537 user        20   0  167M  3904     0 S  0.0  0.2  0:00.00 (sd-pam)
   1490 root        20   0 93968  5944  4328 S  0.0  0.3  0:00.26 /usr/lib/systemd/systemd --user
   1665 root        20   0 82592  4328  3844 S  0.0  0.2  0:00.02 /usr/bin/dbus-daemon --session --address=systemd: -
   1699 root        20   0 82592  4328  3844 S  0.0  0.2  0:00.00 /usr/bin/dbus-daemon --session --address=systemd: -
   1540 root        20   0  155M  4068     0 S  0.0  0.2  0:00.00 (sd-pam)
   1480 root        20   0 42624  1996  1792 S  0.0  0.1  0:00.01 /usr/sbin/atd -f
   1478 root        20   0 27876  3060  2200 S  0.0  0.2  0:00.20 /usr/sbin/crond -n
   1474 root        20   0  204M  6488  4388 S  0.0  0.4  0:00.94 /usr/sbin/rsyslogd -n
   1487 root        20   0  204M  6488  4388 S  0.0  0.4  0:00.02 /usr/sbin/rsyslogd -n
   1483 root        20   0  204M  6488  4388 S  0.0  0.4  0:00.88 /usr/sbin/rsyslogd -n
F1Help  F2Setup F3SearchF4FilterF5Tree  F6SortByF7Nice -F8Nice +F9Kill  F10Quit
```

Set NICE Levels – starting new process

- Default, child process start inherit from parent
- Process start from command line, inherit from shell
- Typically new processes runs with NICE value 0
- Example: Start a process with default NICE

```
[user@host ~]$ sha1sum /dev/zero &
[1] 3480
[user@host ~]$ ps -o pid,comm,nice 3480
  PID COMMAND          NI
 3480 sha1sum           0
```

- Example: Start a process with NICE value = 10

```
[user@host ~]$ nice sha1sum /dev/zero &
[1] 3517
[user@host ~]$ ps -o pid,comm,nice 3517
  PID COMMAND          NI
 3517 sha1sum          10
```

Set NICE Levels – starting new process

- Example: Start a process with custom NICE

```
[user@host ~]$ nice -n 15 sha1sum &
[1] 3521
[user@host ~]$ ps -o pid,comm,nice 3521
  PID COMMAND          NI
 3521 sha1sum          15
```

- Important:

  Unpriviledged users may only increase NICE value from current to maximum of 19.

  Root may reduce it to -20

Set NICE Levels – on existing process

- Example: Change process ID 3521 NICE value to 19

```
[user@host ~]$ renice -n 19 3521
3521 (process ID) old priority 15, new priority 19
```

- Example: Using htop (F7 & F8)

- Important:

  Unpriviledged users may only increase NICE value from current to maximum of 19.

  Root may reduce it to -20

Checkpoint

1. Change existing process id 1234 NICE to -15.
   a) chprio -pri -5 1234
   b) chprio 1234 –prio -5
   c) kill -15 1234
   d) renice –n -15 1234

2. Which are possible commands to see NICE value? [Choose all applies]
   a) ps aux
   b) ps axo pid,comm,ni
   c) top
   d) ps -ef

3. User james attempted in changing his own process NICE value # renice -5 25999 but got error saying permission denied. What is the possible reason?
   a) Underpriviledge user are not permitted to user renice command
   b) Underpriviledge user can only increase NICE value but never decrease it
   c) User james did not have appropriate ACL permission to the file
   d) User james did not have appropriate permission to the process

Checkpoint

1. Change existing process id 1234 NICE to -15.
   a) chprio -pri -5 1234
   b) chprio 1234 –prio -5
   c) kill -15 1234
   d) renice –n -15 1234

2. Which are possible commands to see NICE value? [Choose all applies]
   a) ps aux
   b) ps axo pid,comm,ni
   c) top
   d) ps -ef

3. User james attempted in changing his own process NICE value # renice -5 25999 but got error saying permission denied. What is the possible reason?
   a) Underpriviledge user are not permitted to user renice command
   b) Underpriviledge user can only increase NICE value but never decrease it
   c) User james did not have appropriate ACL permission to the file
   d) User james did not have appropriate permission to the process

Checkpoint

4. Devops team install and configure a database system . It was running fine until next reboot, the database system misbehave. You suspect this is due to dynamic tuning by the kernel. How do you stop the dynamic tuning?
   a) tuned --stop
   b) tuned --disable
   c) systemctl stop tuned
   d) tuned-adm off

5. True or False: Higher NICE value has more priority

Checkpoint

4. Devops team install and configure a database system . It was running fine until next reboot, the database system misbehave. You suspet this is due to dynamic tuning by the kernel. How do you stop the dynamic tuning?
   a) tuned --stop
   b) tuned --disable
   c) systemctl stop tuned
   d) tuned-adm off

5. True or False: Higher NICE value has more priority

Unit summary

Having completed this unit, you should be able to:

- Understand performance optimization
- Manage and Select Tuning Profiles
- Influence Process Scheduling