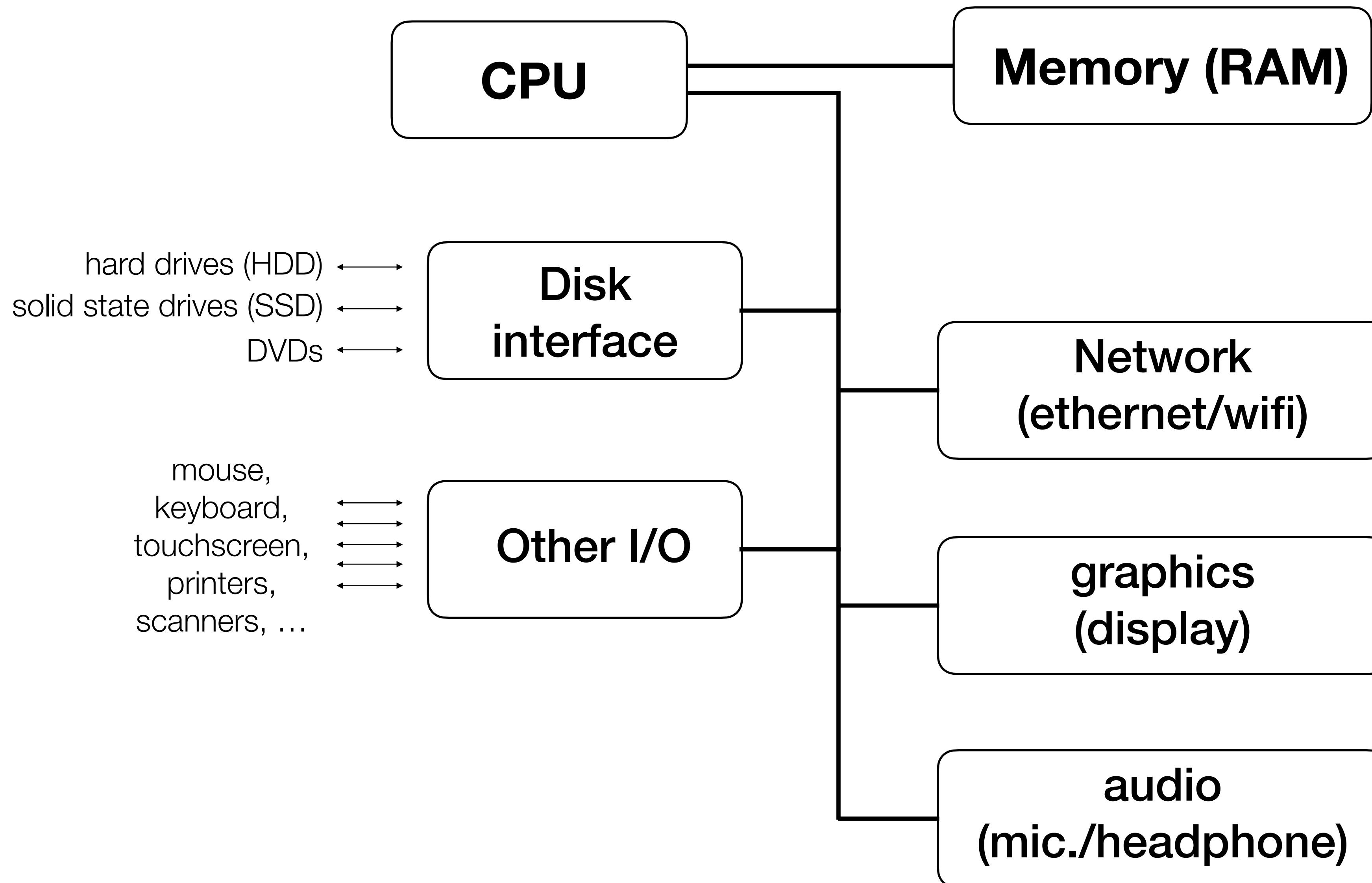


# Introduction to computers and computing environments



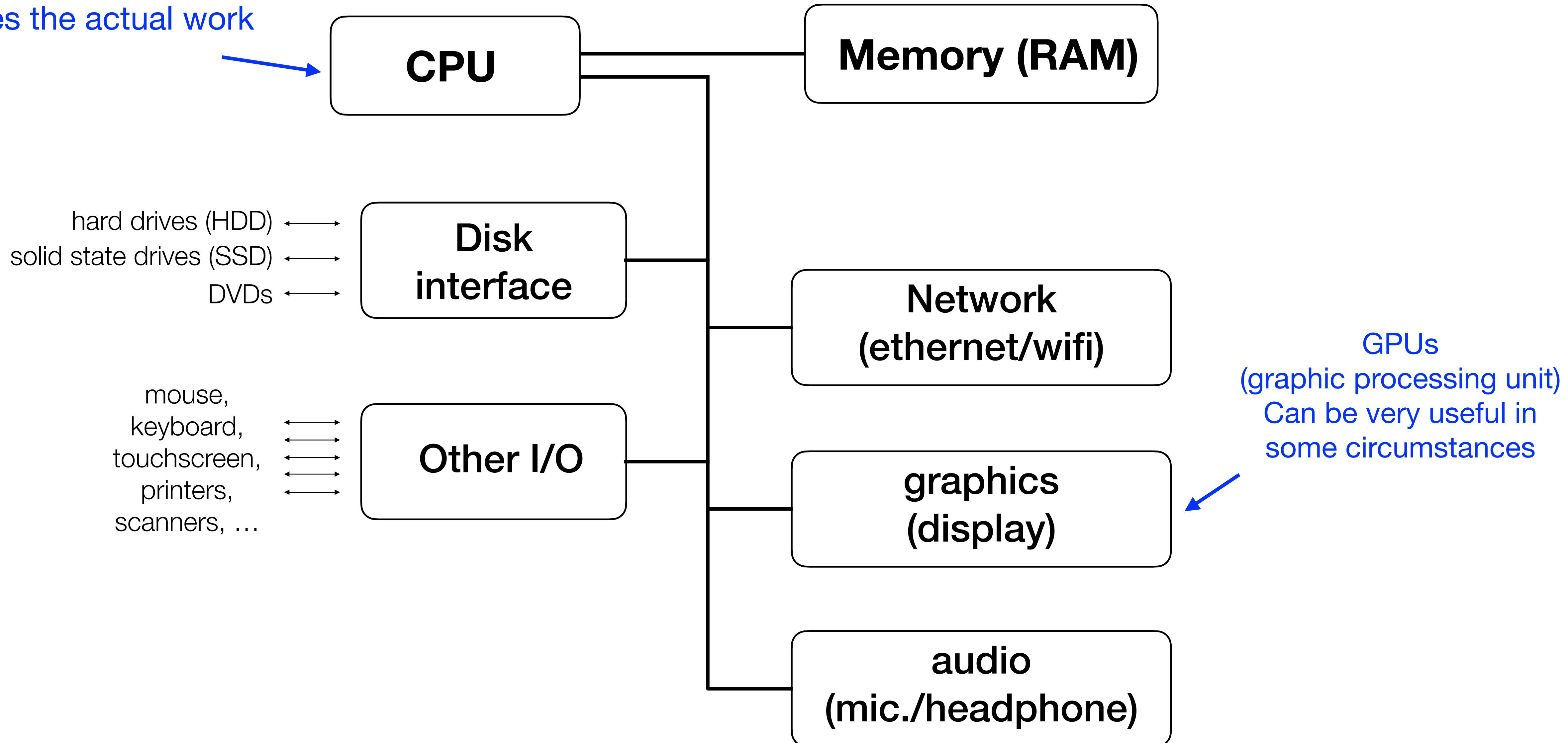
Mark Stenglein, MIP 280A4

# Basic computer architecture

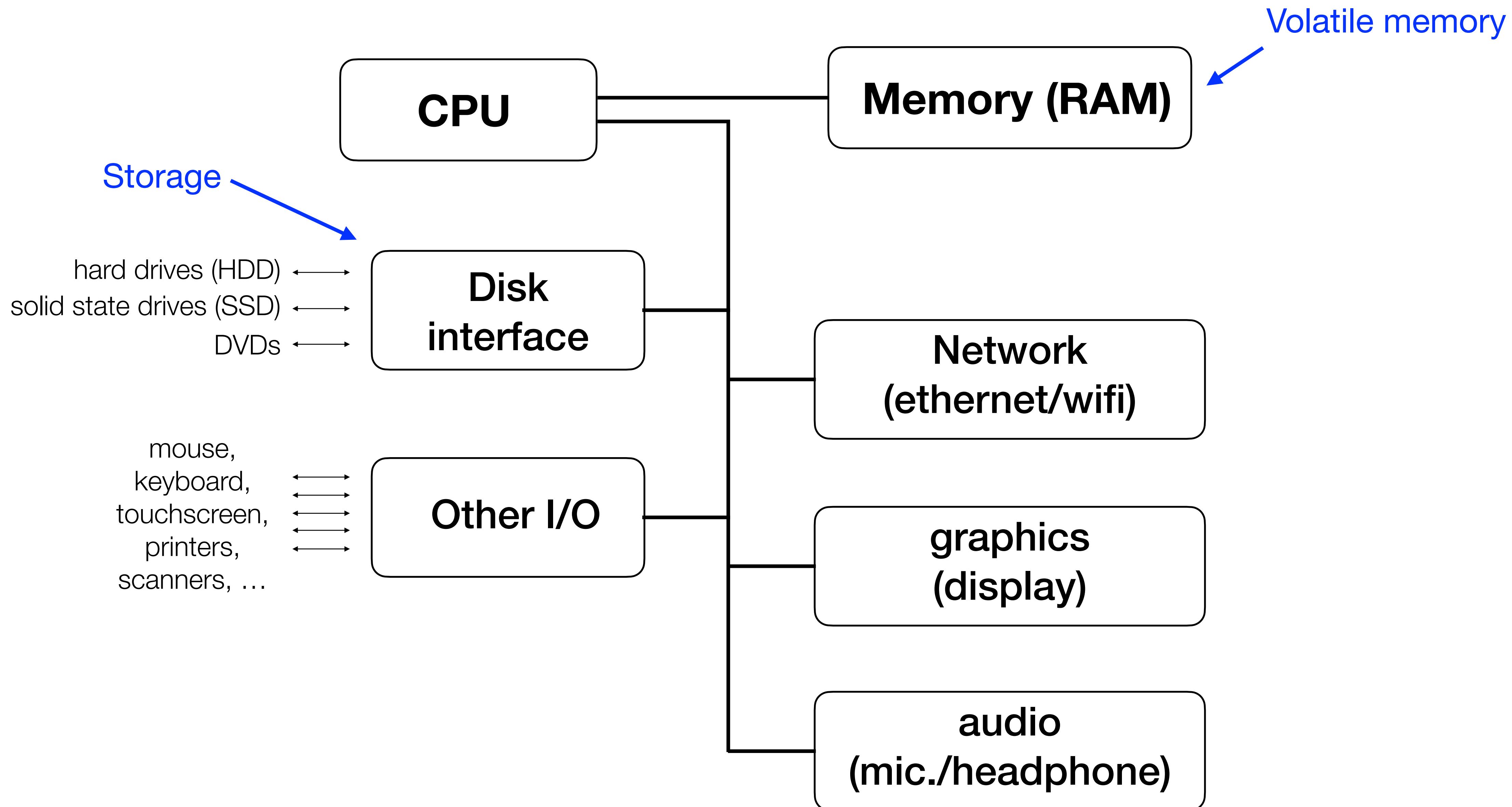


# Basic computer architecture

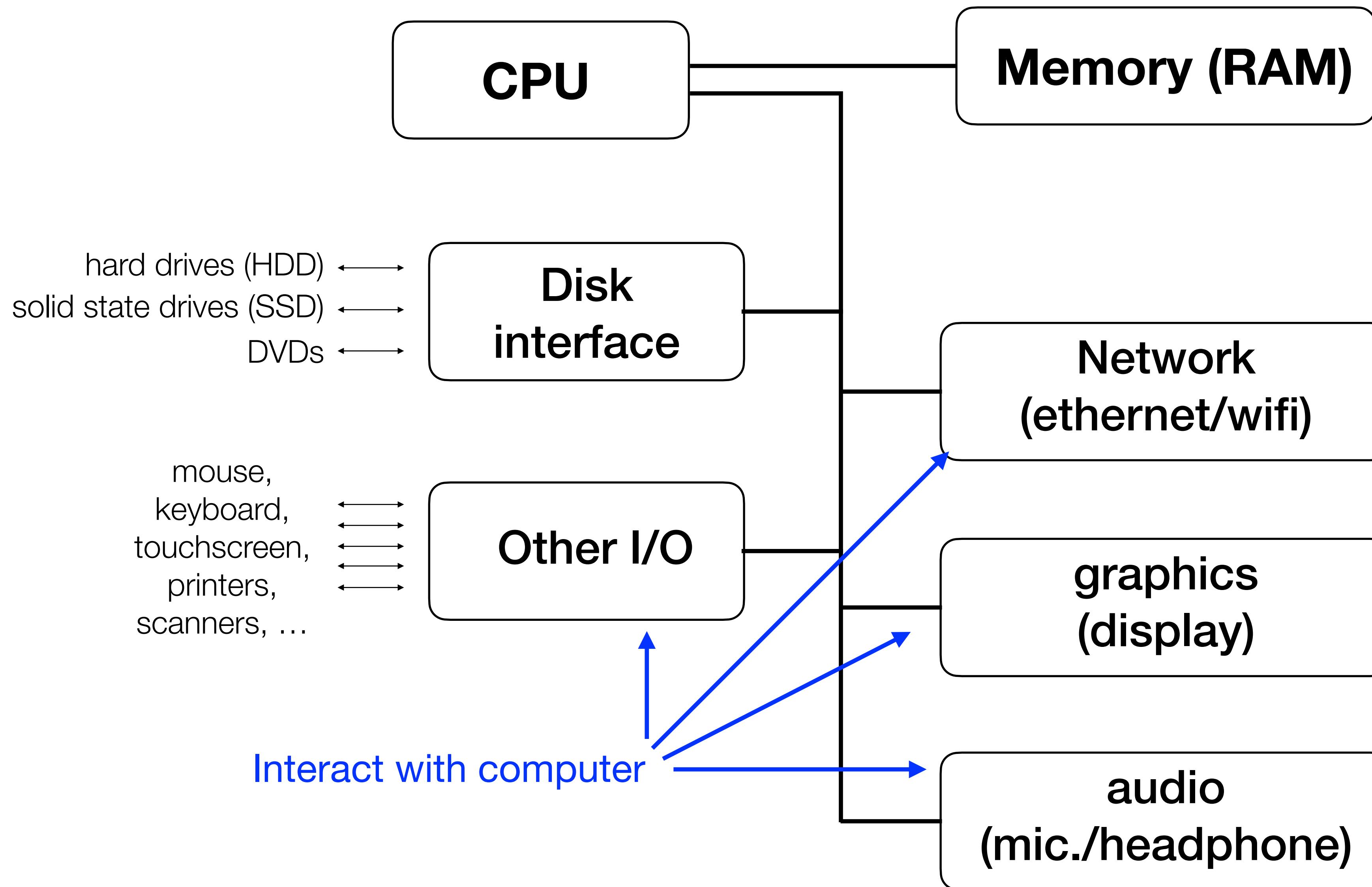
Central Processing Unit  
Does the actual work



# Basic computer architecture



# Basic computer architecture



Sometimes people conflate memory (RAM) and storage (hard drives)  
But they are different things with different purposes.



#### Core Components

CPU / Processors >

Memory > 

Motherboards >

Video Cards & Video Devices >

Computer Cases >

Power Supplies >

Fans & PC Cooling >

Barebone / Mini Computers >

Server Components >

Sound Cards >

#### Storage Devices

Hard Drives >

SSDs > 

Backup Devices & Media >

CD / DVD / Blu-Ray Burners & Media >



# Memory (RAM) vs. storage (SSD & HDD)

	<b>Hard drives (HDD)</b>	<b>Solid state drives (SSD)</b>	<b>Main memory (RAM)</b>
Speed	Slow	Much faster	Fastest
Cost per Gb	\$0.025	\$0.20	\$4
Volatile*	No	No	Yes
Uses	Long term storage of large datasets	Laptops, long term storage, OS boot drives	
Other comments	Has moving parts	Replacing HDDs	Some bioinformatic tasks, like genome assembly, require lots of RAM (>128 Gb)



\* volatile storage means that the contents of the storage will be lost when the computer is turned off

# Storage and memory capacity and computer file sizes are measured in bytes and use the same SI prefixes as genome size

SI prefixes used to quantify storage capacity and file sizes

Name	Symbol	Base 10*	Decimal	English Name
tera	T	$10^{12}$	1,000,000,000,000	trillion
giga	G	$10^9$	1,000,000,000	billion
mega	M	$10^6$	1,000,000	million
Kilo	k	$10^3$	1,000	thousand

1 byte = 8 bits (binary digit: 1 or 0)

byte 1	byte 2	byte 3
10110100	01110001	11110000

\* There is some disagreement about how these values are defined:  
(1 kilobyte also sometimes defined as 1024 bytes not 1000 bytes).  
Both definitions are used.

Memory (RAM) is like a “mise en place” station in a kitchen.  
Storage is like a pantry.

Limited capacity but quick to access



Larger capacity, slower to access



# How much RAM do you need?

Laptops usually have somewhere between **2Gb - 16 Gb** of RAM. That much RAM is plenty for some tasks.

Some bioinformatics tasks are processor (CPU) intensive, but use little RAM. For instance, making phylogenetic trees.

Other tasks, like assembly of large eukaryotic genomes or searching large databases with tools like BLAST, require lots of RAM (128 Gb or more). These types of analyses are better done on a dedicated server, a ‘super-computer’, or in the cloud.



# How much processing power do you need?

The CPUs in typical laptops (and cell phones even) are very fast compared to past computers.

Nevertheless, some computing tasks would take a long time to run on your laptops. These types of analyses are better done on a dedicated server, a ‘super-computer’, or in the cloud.

One way to know is to just try running a task: if it takes a long time, try moving to a faster computer (server or super-computer).



Parallelization makes life better by shortening the time it takes to accomplish tasks

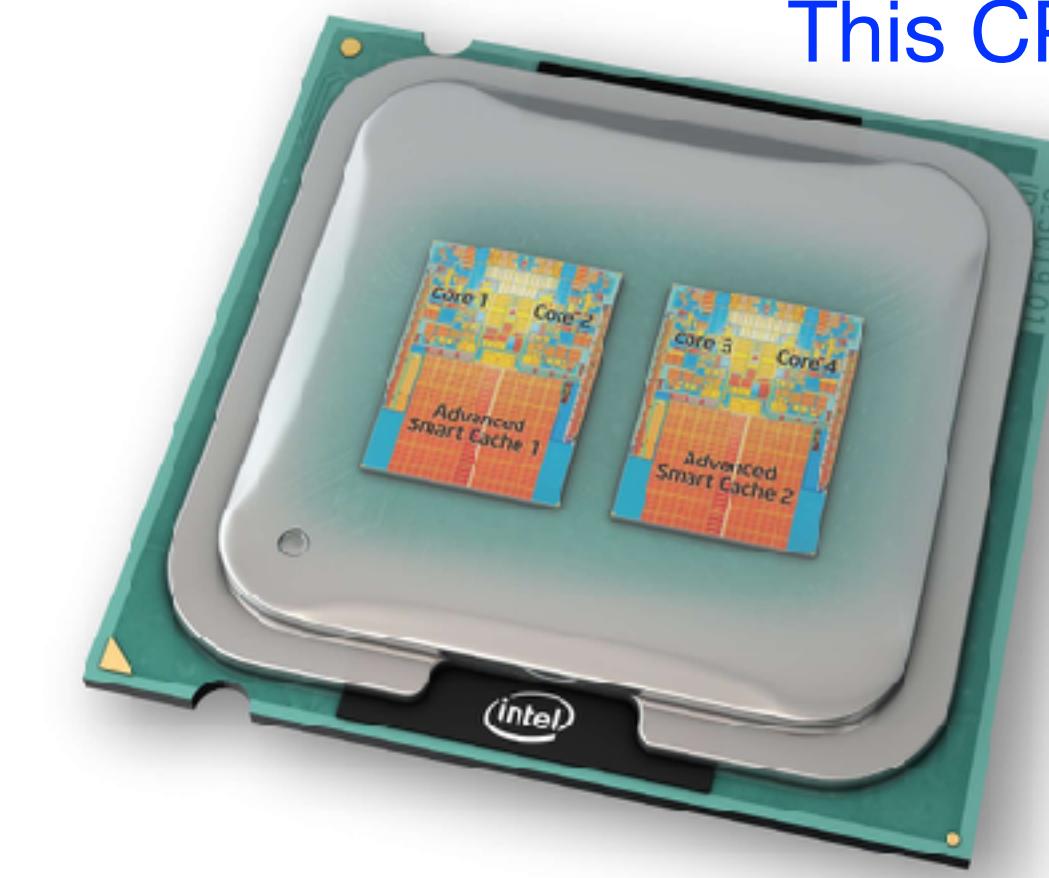


# Like airport security, computers can run in parallel



# Modern computers have more than one processor

Typically, these are referred to as “cores”.



This CPU has 4 cores (4 CPUs)

To add to the confusion: cores can have multiple “threads”, which are effectively additional cores. Intel cores are “hyper-threaded” with 2 threads, which means that the operating systems sees these 4 cores as 8 CPUs.

```
[MDSTENGL-M01:Desktop _mdstengl$ sysctl hw.physicalcpu hw.logicalcpu  
hw.physicalcpu: 4  
hw.logicalcpu: 8
```

# A typical bioinformatics server

CPUs  
 $(4 \times 16 \text{ cores} = 64)$

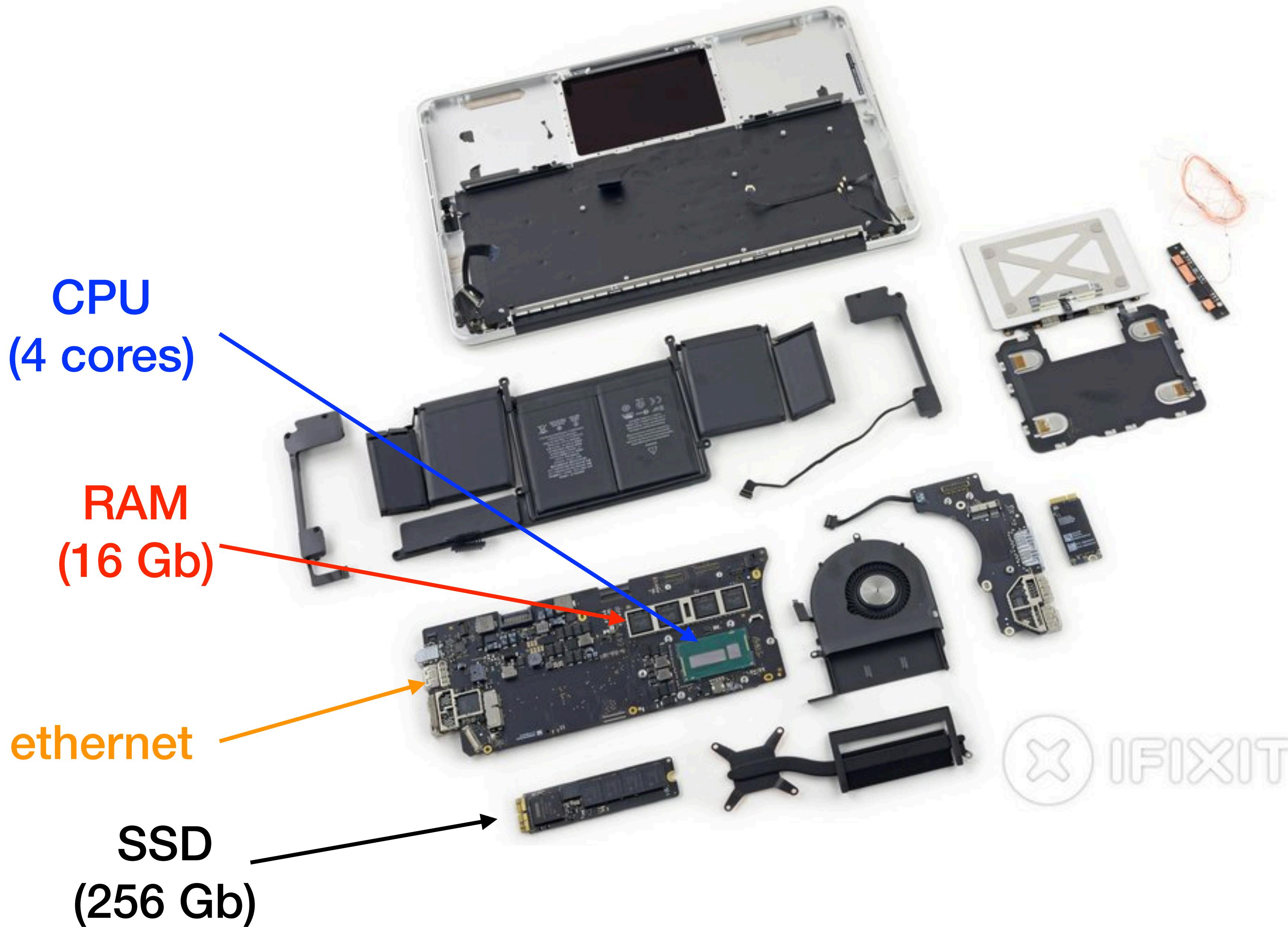
RAM  
 $(16 \times 32\text{Gb} = 512 \text{ Gb})$

ethernet  
(main user interface)

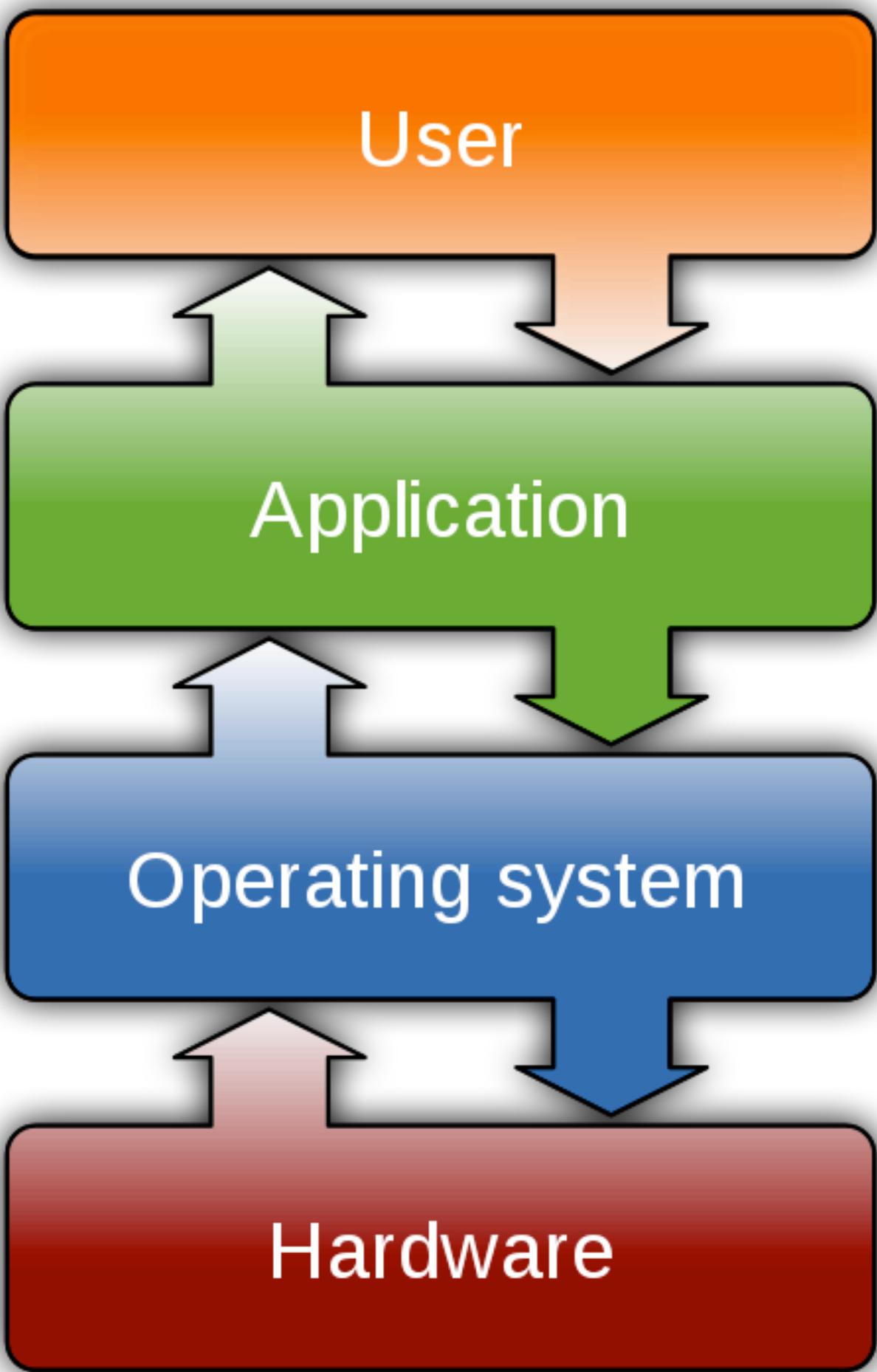


HDDs (6 x 8 Tb)  
& SSDs (2 x 128Gb)  
RAID configured

# The components of a Macbook



A computer's operating system (OS) is the software that provides an interface between the hardware and the applications and user



### Common Operating systems:

- Windows
- Mac OS
- Linux
- Chrome OS
- iOS
- Android

# How do you check resource usage and availability?

## **On your own laptop:**

- How much memory is available? How much is used?
- How much storage is available? What type is it? How much is used?
- How many cores and threads are there? What fraction of these are being used?



# Exercise: How do you check resource usage and availability?

## On your own laptop:

- How much memory is available? How much is used?
- How much storage is available? What type is it? How much is used?
- How many cores and threads are there? What fraction of these are being used?

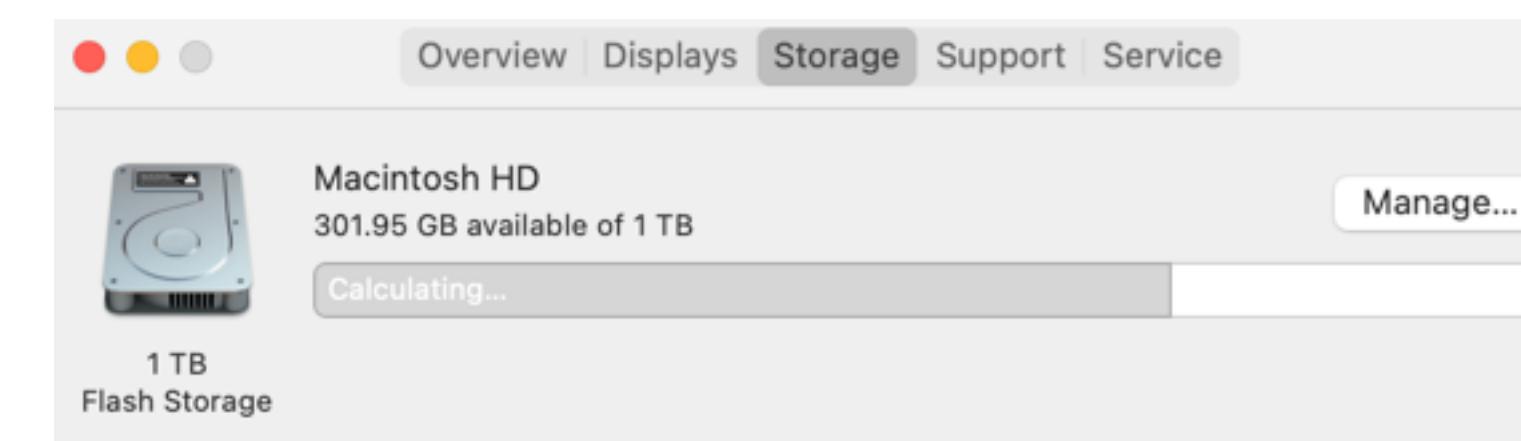
Check resources available on Mac:

Apple menu -> About This Mac



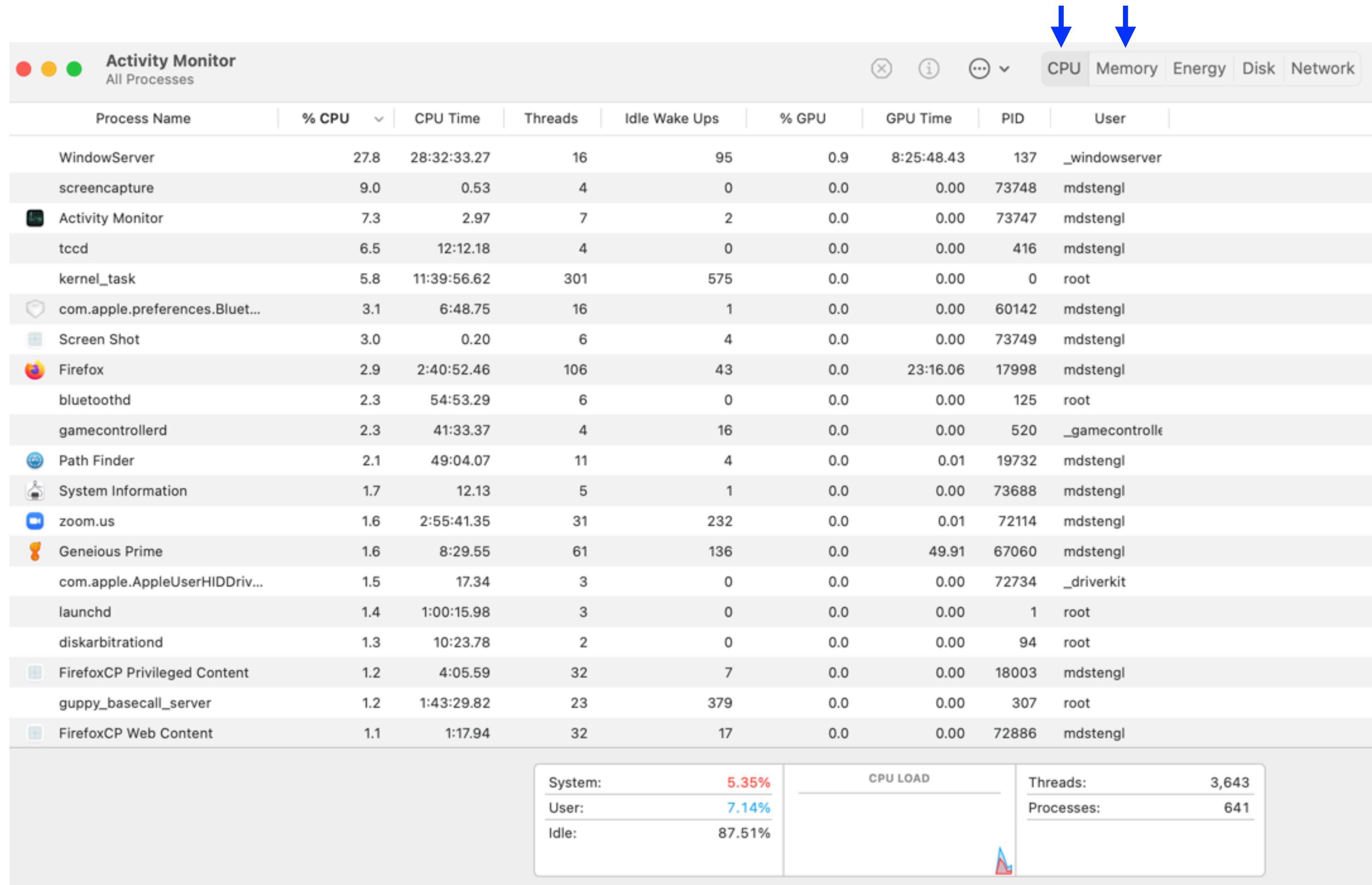
CPU: quad Intel core = 4 cores x 2 threads = effectively 8 CPUs

16 Gb of RAM



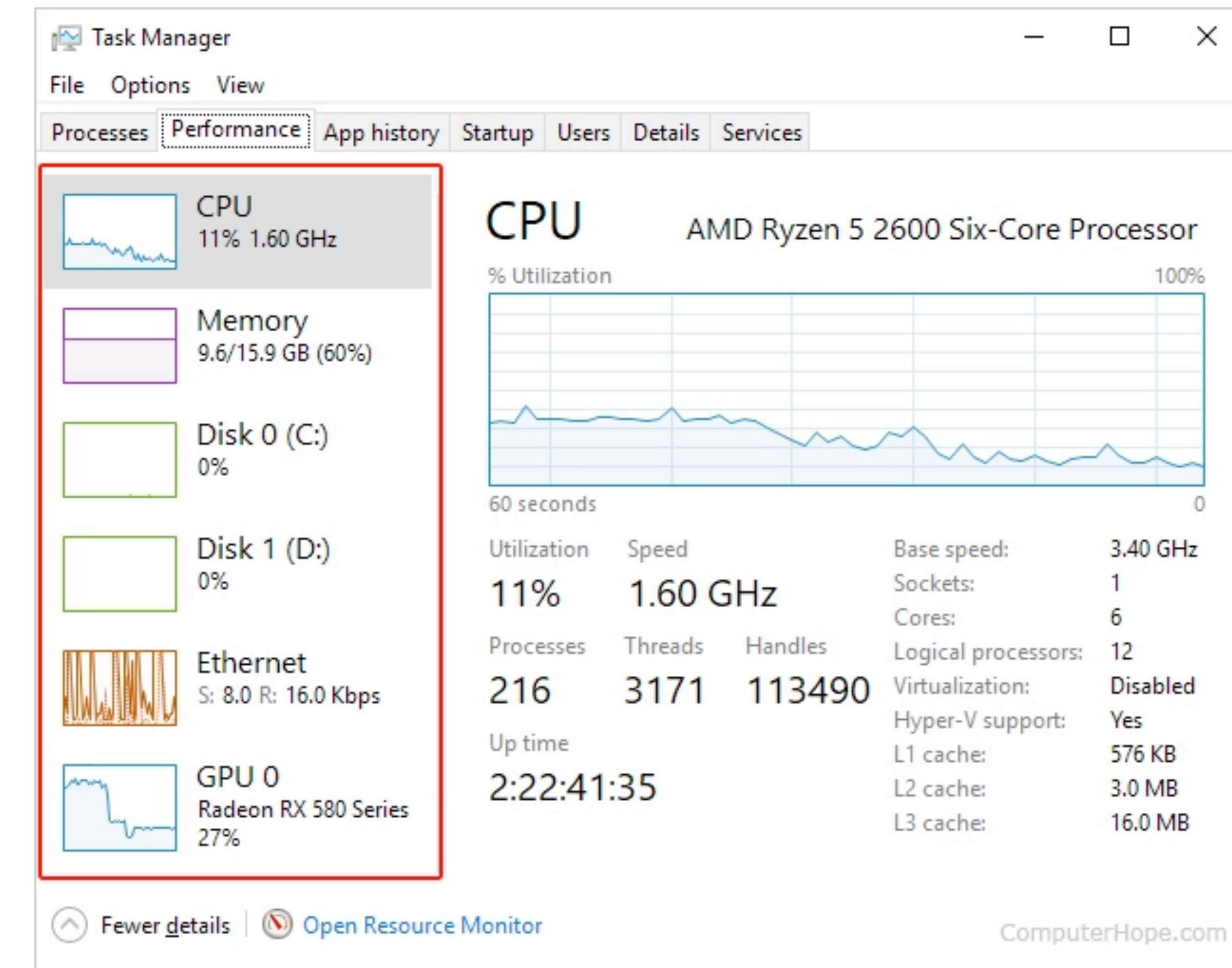
1 Tb SSD for storage  
(Flash storage = SSD)

# Activity Monitor: checks resource usage on a Mac



# Use Task Manager to check resource usage in Windows

- Windows tab (bottom left-hand corner)
- Type ‘task manager’
- Performance tab



# For chromebooks

*<https://www.lifewire.com/how-to-check-chromebook-specs-hardware-4782658>*

# Local vs. remote computing

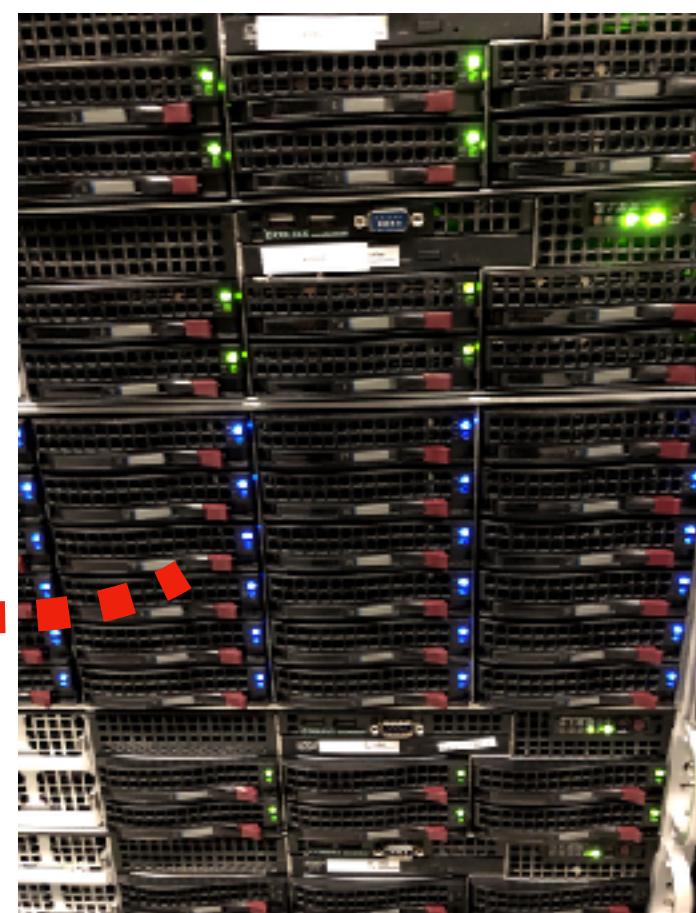
Local: you are using the resources on your computer



Remote: you are using the resources of a distant computer, but probably still connected through your own computer



a server somewhere



the internet

# Connecting to remote computers on the command line

Typically you connect to a remote server via the command line using a “secure shell” (ssh). For instance, to connect to the thoth01 server, I would open the Mac OS Terminal and run:

```
ssh mdstengl@thoth01.cvmbs.colostate.edu
```

*Replace mdstengl with your CSU eid*

- Mac OS: use the built-in Terminal app
- Windows: use MobaXterm: <https://mobaxterm.mobatek.net/>
- Chrome OS: use the ssh chrome extension:  
<https://chrome.google.com/webstore/detail/secure-shell/iodihamcpbpeioajjeobimgagajmlbd>

# Connecting to remote computers on the command line

Typically you connect to a remote server via the command line using a “secure shell” (ssh). For instance, to connect to the thoth01 server I would run:

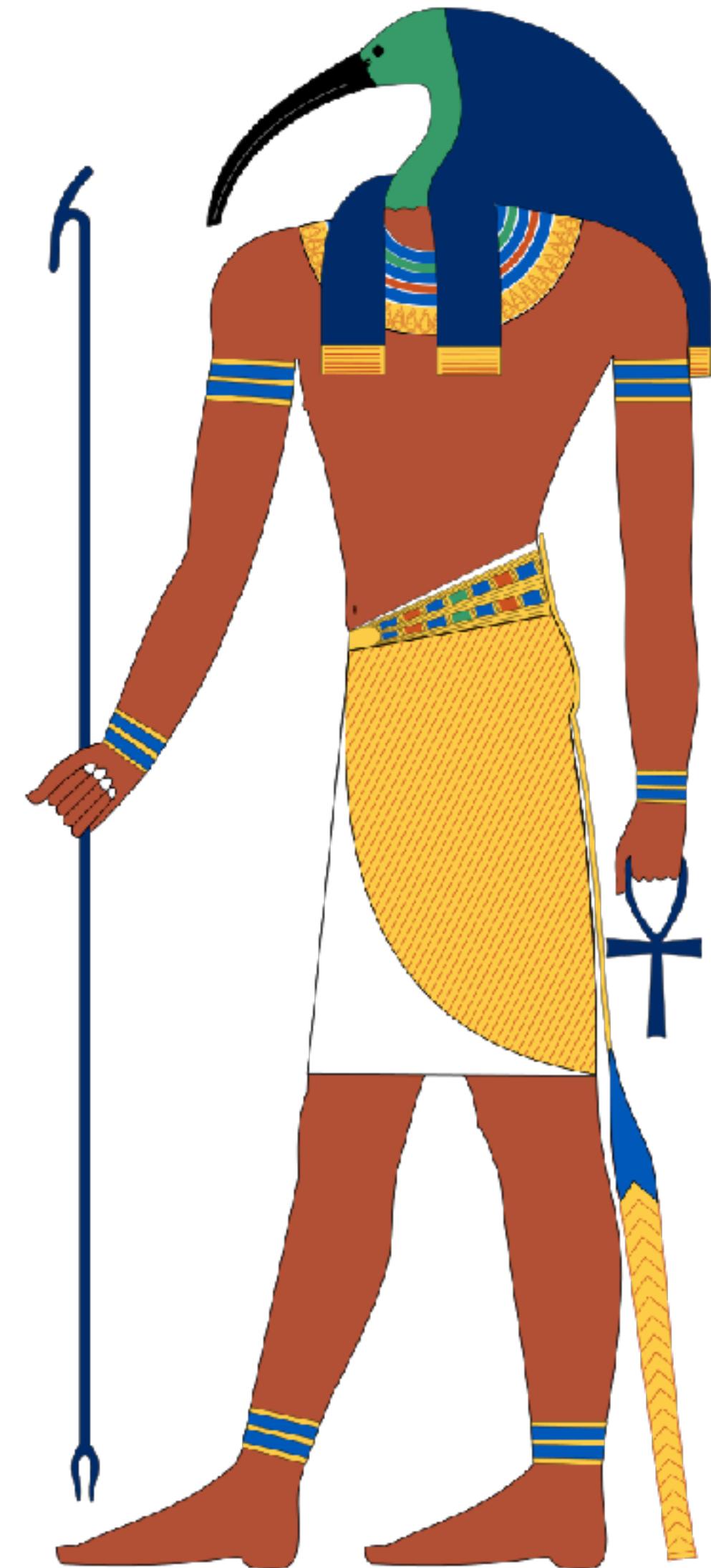
```
ssh mdstengl@thoth01.cvmbs.colostate.edu
```

Note that this server is inside the CSU firewall, so to connect to it you have to either:

1. Be on the CSU ethernet (ethernet cable plugged in to your computer)
2. Be on the csu-eid wifi network
3. Be connected through a virtual private network (VPN). See:

<https://www.acns.colostate.edu/security/#1478123291089-f3918698-6eec>

Thoth is the Egyptian god of  
“moon, wisdom, writing, hieroglyphs, science, magic, art, and judgment”



# Checking resource usage and availability on a remote server through the command line

command	better way to run it
df	reports disk usage and storage
top	reports CPU & memory usage

## Use these commands to answer the same questions:

- How much storage is available? How much is used?
- How much memory is available? How much is used?
- How many cores and threads are there? What fraction of these are being used?

# Cloud computing is an increasingly popular form of remote computing

## **Advantages:**

- Scalable and flexible
- Don't have to buy, host, or maintain servers
- Can take advantage of pre-existing images and containers

## **Disadvantages:**

- Can be expensive
- Have to pay to store and transfer data
- Can be slow to transfer data

Real physical computing resources are available as virtual computers through the internet



image: gigabitmagazine.com

# Cost of cloud computing depends on how many resources you need

	vCPU	ECU	Memory (GiB)	Instance Storage (GB)	Linux/UNIX Usage	
t2.large	2	Variable	8 GiB	EBS Only	\$0.1104 per Hour	Similar to typical laptops
t2.xlarge	4	Variable	16 GiB	EBS Only	\$0.2208 per Hour	
r5.12xlarge	48	173	384 GiB	EBS Only	\$3.36 per Hour	
r5.24xlarge	96	347	768 GiB	EBS Only	\$6.72 per Hour	Similar to the servers my lab uses

Plus ~\$0.05-\$0.15 per Gb-month for (short term) storage: \$50-\$150 per month per Tb

Plus ~\$0.09 per Gb to transfer from Amazon to anywhere else: \$90 per Tb

# An Amazon Web Services (AWS) linux environment

The screenshot shows the AWS EC2 console interface. On the left, the navigation menu includes options like EC2 Dashboard, Instances (selected), Launch Templates, Spot Requests, Reserved Instances, Dedicated Hosts, Scheduled Instances, Capacity Reservations, AMIs, and more. The main area displays a terminal session for an Ubuntu 18.04.2 LTS instance (ami-005bd005fb00e791). The terminal output shows system information, package updates, and a root login prompt.

Terminal Output:

```
[MDSTENGL-M01:Downloads _mdstengl$ ssh -i "mds_linux.pem" ubuntu@ec2-34-221-17-148.us-west-2.compute.amazonaws.com]$ Welcome to Ubuntu 18.04.2 LTS (GNU/Linux 4.15.0-1032-aws x86_64)

 * Documentation: https://help.ubuntu.com
 * Management: https://landscape.canonical.com
 * Support: https://ubuntu.com/advantage

System information as of Wed May 29 16:56:27 UTC 2019

System load: 0.0          Processes:      84
Usage of /: 13.7% of 7.69GB  Users logged in:  0
Memory usage: 14%          IP address for eth0: 172.31.49.140
Swap usage:  0%

Get cloud support with Ubuntu Advantage Cloud Guest:
http://www.ubuntu.com/business/services/cloud

0 packages can be updated.
0 updates are security updates.

Last login: Wed May 29 16:56:04 2019 from 129.82.26.66
To run a command as administrator (user "root"), use "sudo <command>".
See "man sudo_root" for details.

ubuntu@ip-172-31-49-140:~$ ]
```

Instance Details:

Description	Value
Instance ID	i-094d903b47d77bec
Instance state	running
Instance type	t2.micro
Elastic IPs	None
Availability zone	us-west-2a
Security groups	launch-wizard-2, view inbound rules, view outbound rules
Scheduled events	No scheduled events
AMI ID	ubuntu/images/hvm-ssd/ubuntu-bionic-18.04-amd64-server-20190212.1 (ami-005bd005fb00e791)

Network & Security:

Pv6 IPs	Key Name	Monitoring	Launch Time
mds_linux	disabled		May 29, 2019 at 10:37:
mds_linux	disabled		May 29, 2019 at 10:42:

Feedback: English (US) | © 2008 - 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved. | Privacy Policy | Terms of Use