



# HPC Workshops

**Introduction to High Performance Computing**

**Guillermo Avendaño-Franco**

**Nathaniel Garver-Daniels**

**Daniel Turpen**

**Nicolas Fazee**

# High Performance Computing

## SIZE

Supercomputers are big machines of the size of entire rooms and with special requirements in terms of power and cooling.

HPC Clusters are build from tens, hundreds or even thousands of relatively normal computers that are link together with special networks.

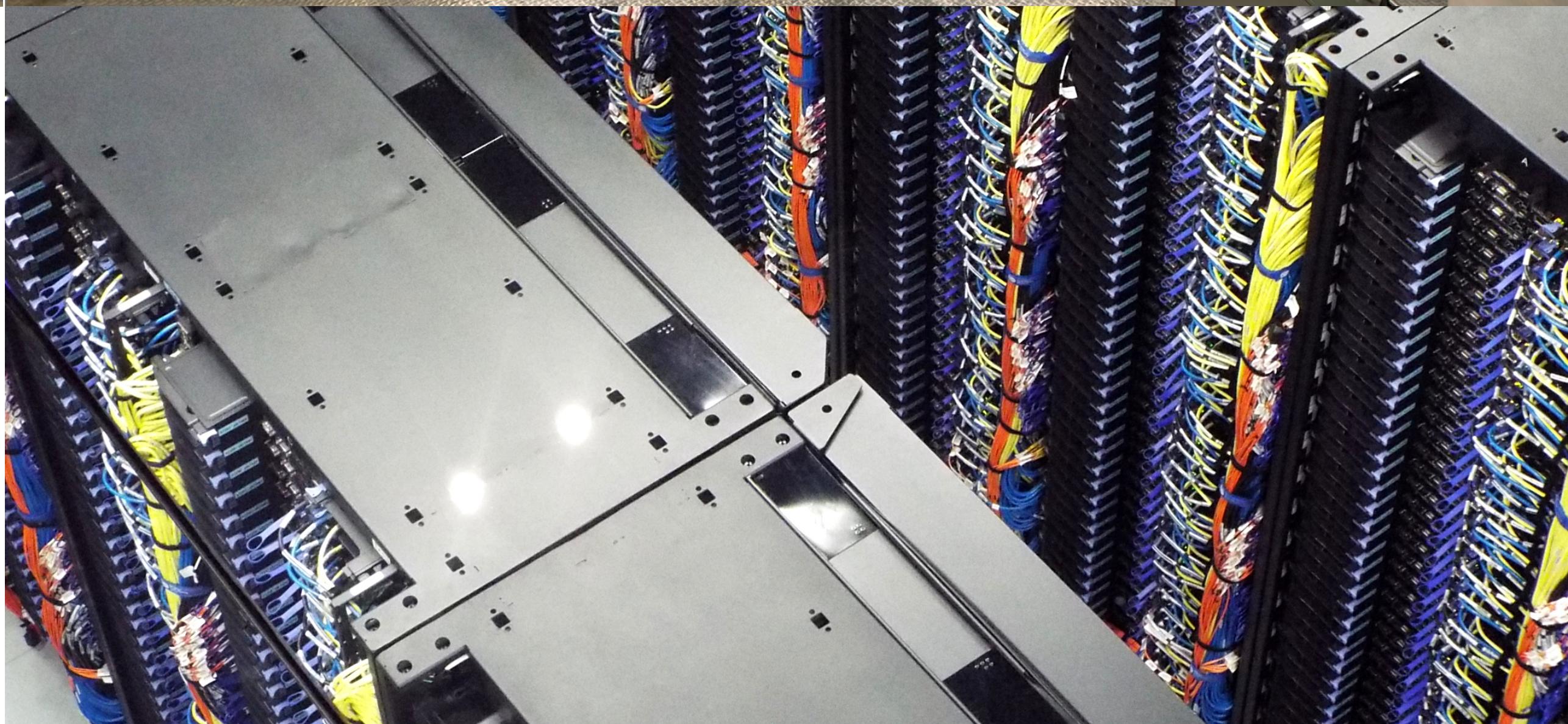
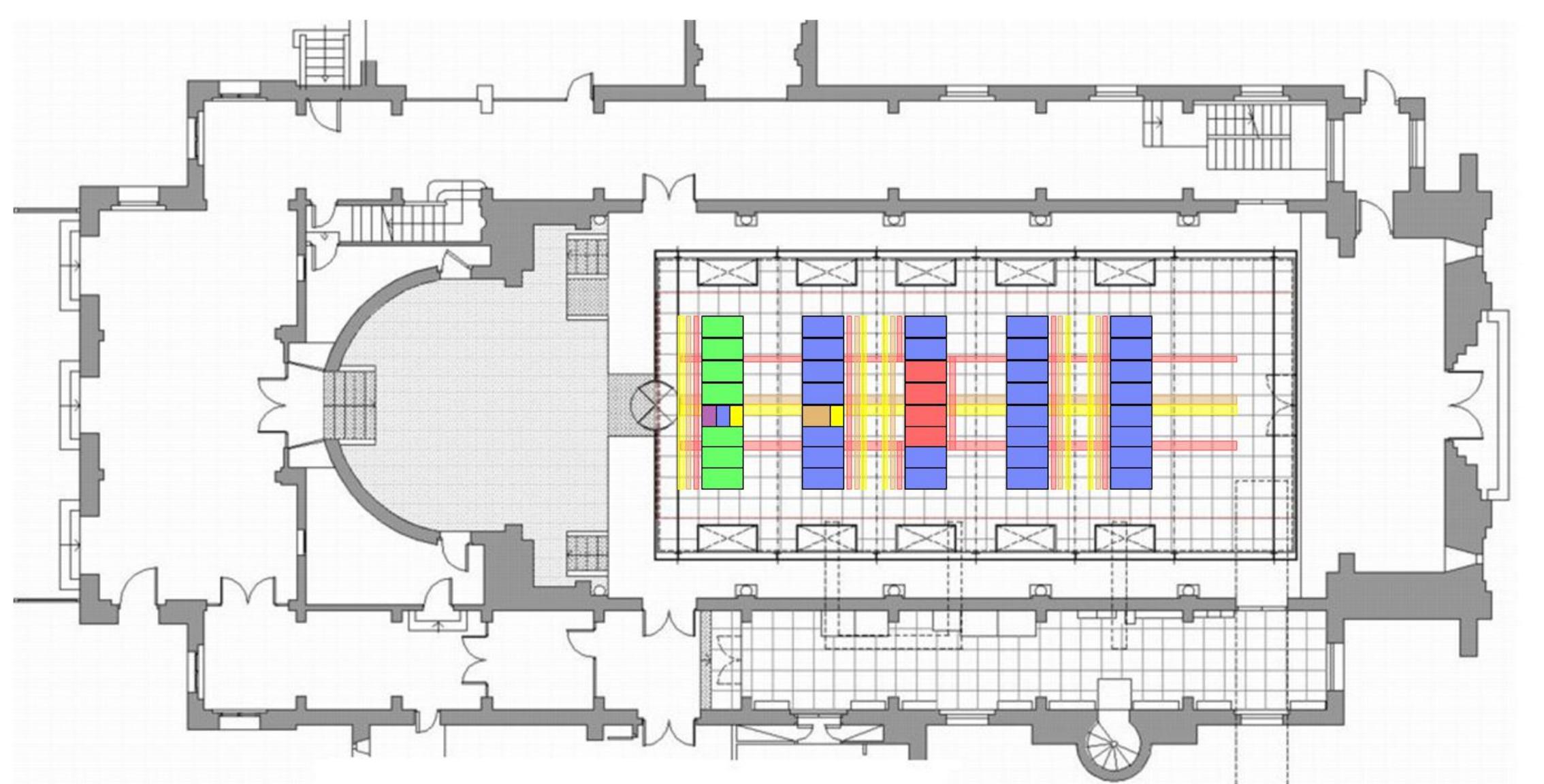
## SPEED

Each individual compute node is more powerful than a normal desktop computer and often even more capable than high-end Workstations.



Summit or OLCF-4 is a supercomputer developed by IBM for use at [Oak Ridge National Laboratory](#), Summit is estimated to cover the space of two basketball courts and require 136 miles of cabling.

# Marenostrum



# HPC: Speed

## FLOPS

Floating Operations Per Second

It is a measure of the computational capabilities of a CPU



## Benchmarks

**HPL Linpack Benchmark**  
Dense Linear Algebra

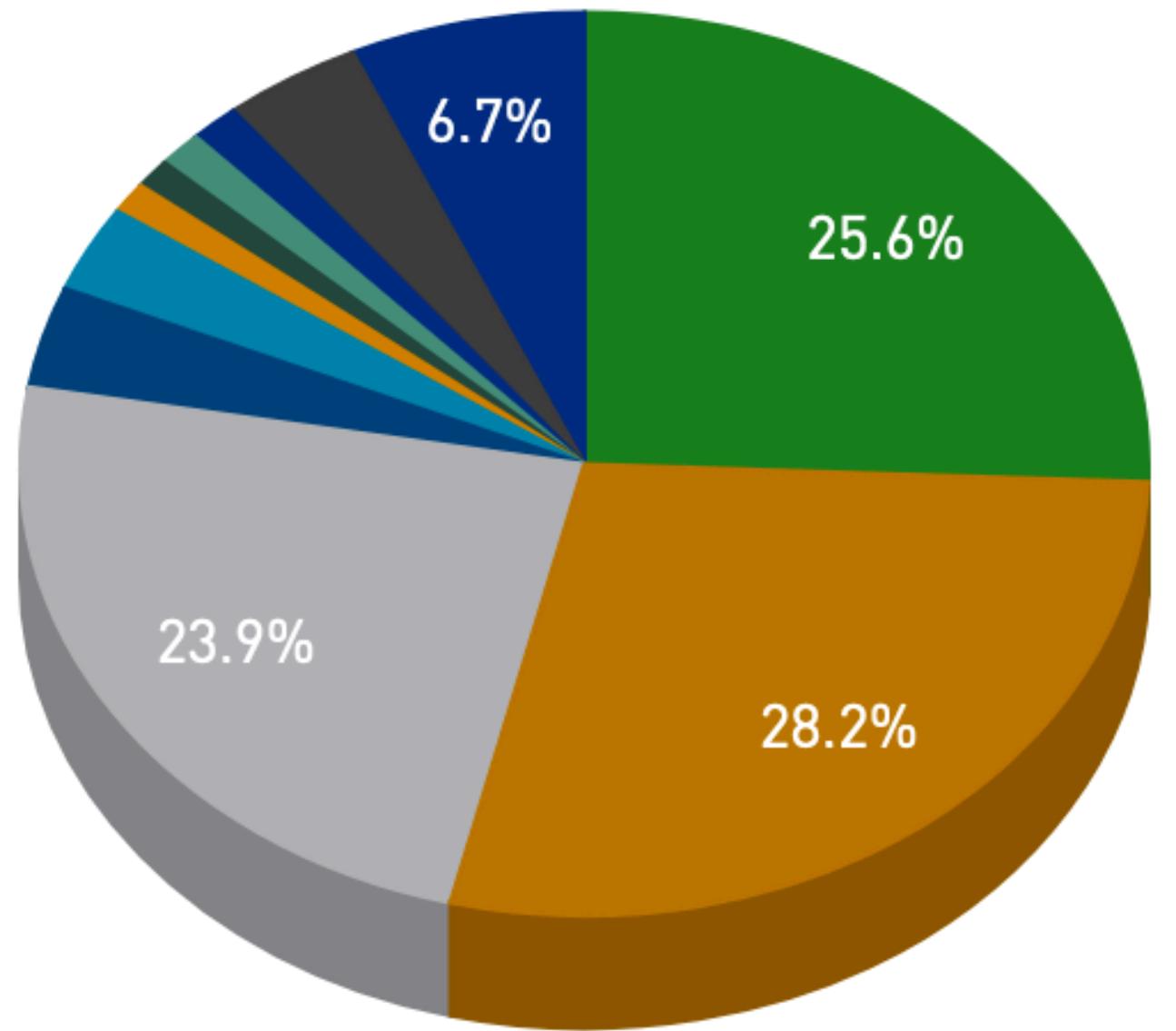
**High-Performance Conjugate Gradient (HPCG) Benchmark**  
Sparse matrix-vector multiplication

**HPL-AI Mixed-Precision Benchmark**  
Artificial Intelligence - Deep Learning

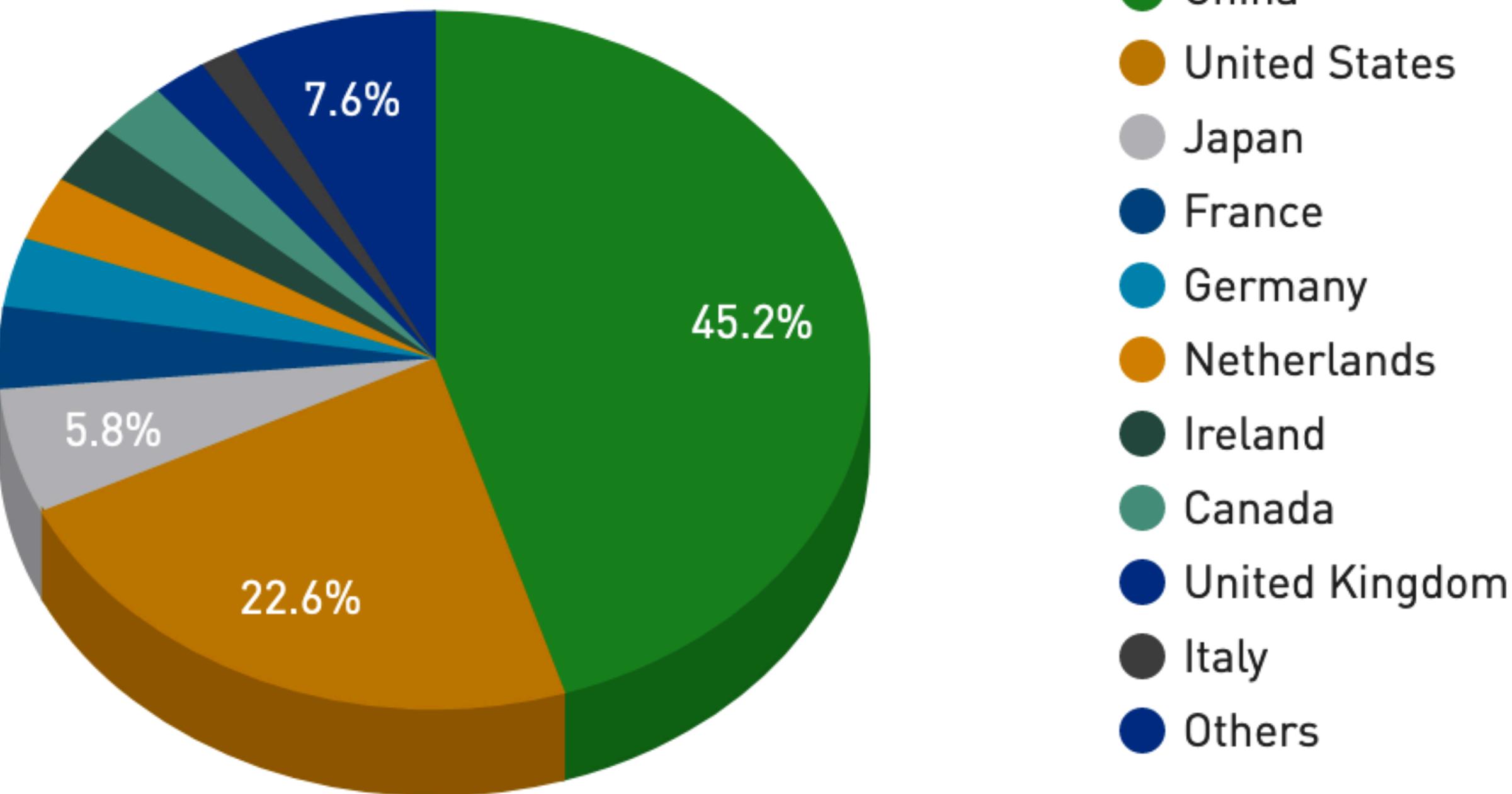
# Top 500 Statistics

## June 2020

Countries Performance Share



Countries System Share



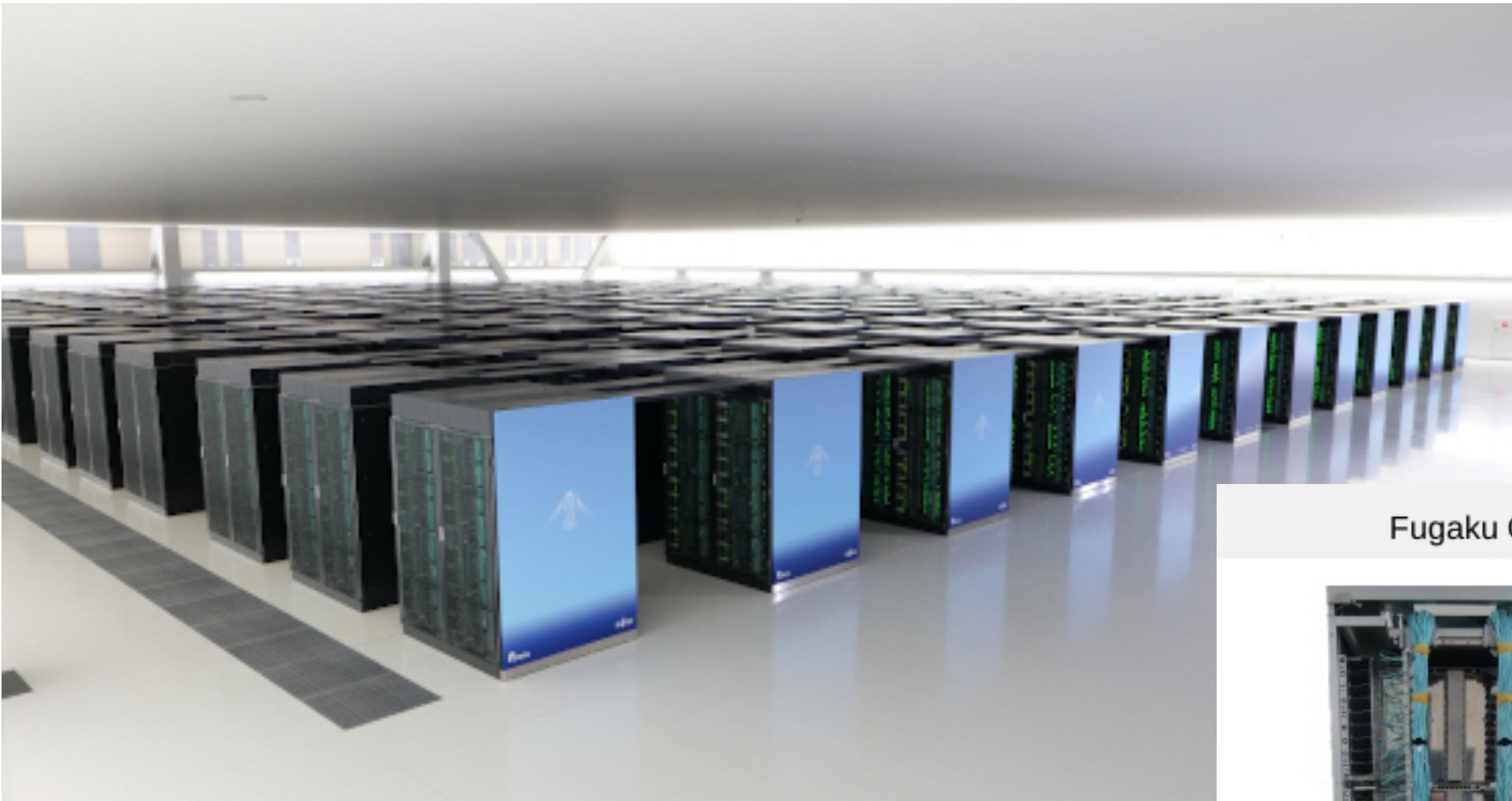
- China
- United States
- Japan
- France
- Germany
- Netherlands
- Ireland
- Canada
- United Kingdom
- Italy
- Others

# Data Storage Terminology

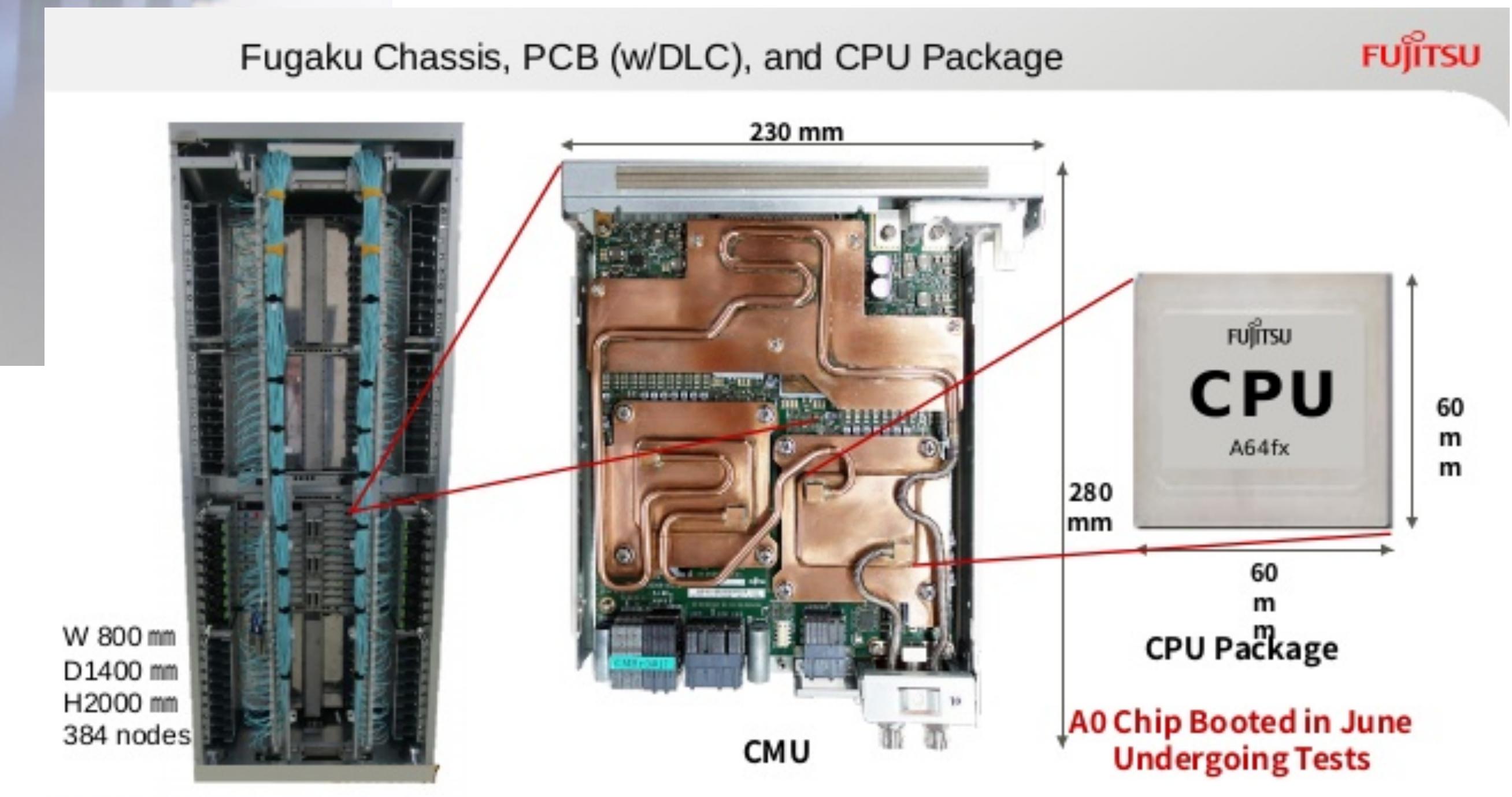
Size	In terms of the unit above...	In human terms...
Byte		Can represent all positive integers from 0 to 255
Kilobytes (KB)	1,000 bytes ( $10^3$ Bytes)	Like a paragraph of a text document.
Megabytes (MB)	1,000 Kilobytes ( $10^6$ Bytes)	A small novel, around 300 pages, no compression.
Gigabytes (GB)	1,000 Megabytes ( $10^9$ Bytes)	Beethoven's 5th Symphony (lossless FLAC format)
Terabytes (TB)	1,000 Gigabytes ( $10^{12}$ Bytes)	Images from all X-rays in a large hospital
Petabytes (PB)	1,000 Terabytes ( $10^{15}$ Bytes)	The contents of all US academic libraries
Exabytes (EB)	1,000 Petabytes ( $10^{18}$ Bytes)	1/5 of the words people have ever spoken
Zettabytes (ZB)	1,000 Exabytes ( $10^{21}$ Bytes)	Data stored by Google, Amazon, Facebook, Apple, and Microsoft)
Yottabytes (YB)	1,000 Zettabytes ( $10^{24}$ Bytes)	Number of atoms in 4 100g bouillons of Gold



# Fugaku, the fastest supercomputer in 2020



- 995,328 CPU cores
- 158,976 nodes
- 384 nodes × 396 racks = 152,064
- 192 nodes × 36 racks = 6,912
- 4.85 PB of total memory
- 163 PB/s memory bandwidth
- 28 MW of energy consumption



The Fugaku compute system was designed and built by Fujitsu and RIKEN. Fugaku 富岳, is another name for Mount Fuji. The system is installed at the RIKENCenter for Computational Science(R-CCS) in Kobe, Japan.

# High Performance Computing

Specifications	Desktop / Workstation	HPC Cluster (T)
<b>Number of Nodes</b>	1	10s 100s or 1000s Thorny (~110 nodes)
<b>Total RAM per node</b>	[4GB, 8GB ... 24 GB..]	[100s GB] Thorny (96, ..., 768GB)
<b>Number of Cores per node</b>	[4, 8, ..., 16]	[16, 32, ..., 128] Thorny (40 cores)
<b>Storage</b>	[2 - 10 TB]	100s TB ... 10s PB Thorny (~750 TB)
<b>GPUs</b>	1	Several GPUs Thorny (3 GPUs Nvidia P6000)
<b>CPU speed</b>	2 GHz	2 GHz

# **Linux/Unix Commands**

# echo

**cat**

# date

**pwd**

cd

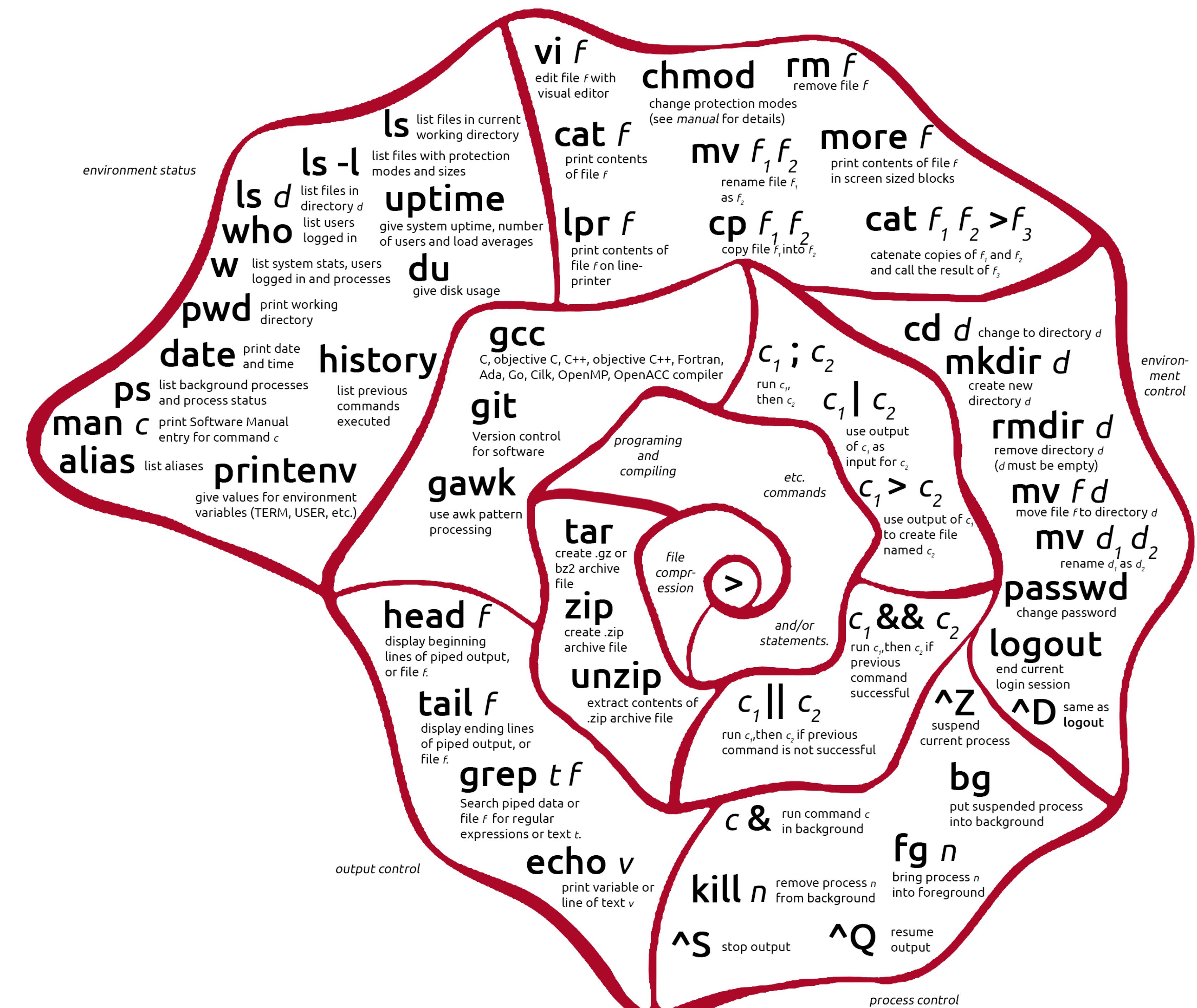
# **mkdir**

# touch

cp

mv

rm



# Text Editors

## nano

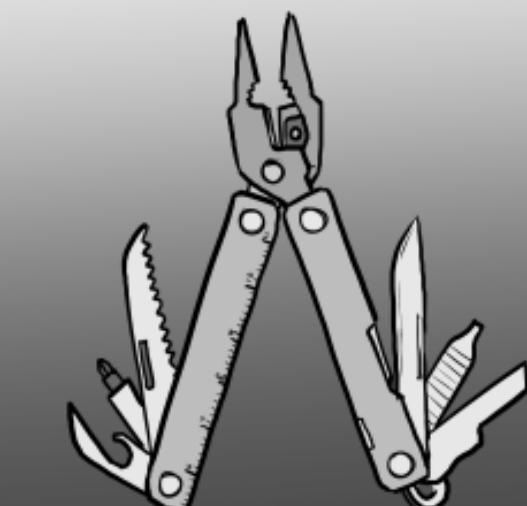
- No learning curve.
- Easy to use.
- Good for simple edits.
- Commands on the bottom.

VIM  
usable in just about  
any environment.

does one thing, well.



EMACS  
flexible, customizable, and  
packed with every feature  
known to man.



## emacs

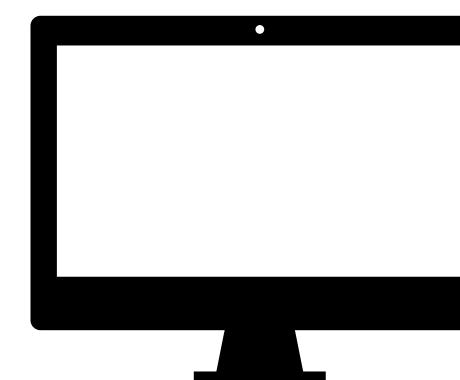
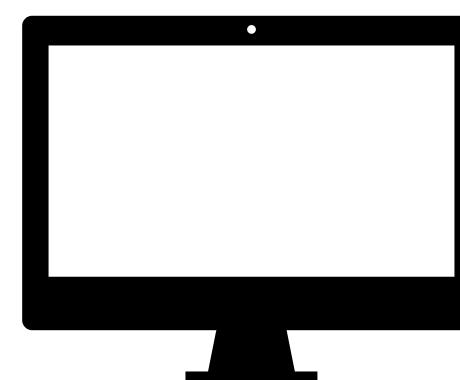
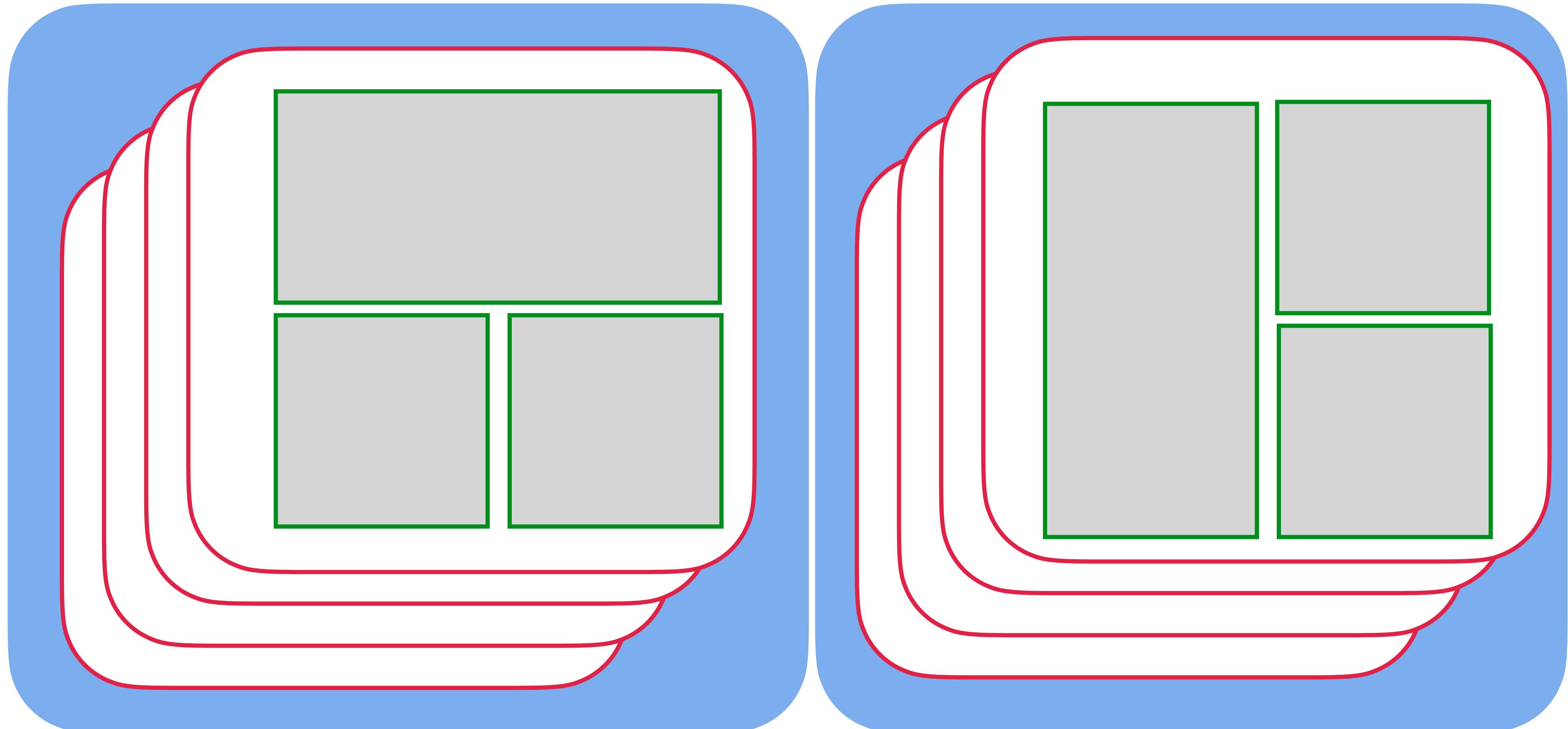
- Customizable and extensible.
- Powerful editing capabilities.
- Syntax highlighting for many languages
- Integration with debuggers
- Capable beyond edit files.

## vi/vim

- Modal editor
- Complex edits can be quick and easy.
- Multiple file/window support
- Keyboard shortcuts, macros and registers
- Quick command repetition, auto-completion.

# TMUX Concepts

- Clients
- Sessions
- Windows
- Panes





# ENVIRONMENT **MODULES**

**Local Variables** – It is present within the current instance of the shell. It is not available to programs that are started by the shell. They are set at the command prompt.

**Environment Variables** – It is available to any child process of the shell. Some programs need environment variables in order to function correctly.

**Shell Variables** – A shell variable is a special variable that is set by the shell and is required by the shell in order to function correctly. Some of these variables are environment variables whereas others are local variables.

The **Environment Modules** package is a tool that simplify shell initialization and lets users easily modify their environment during the session with **modulefiles**.

Each **modulefile** contains the information needed to configure the shell for an application. Once the Modules package is initialized, the environment can be modified on a per-module basis using the module command which interprets **modulefiles**.

Typically **modulefiles** instruct the module command to alter or set shell environment variables such as **PATH**, **LD\_LIBRARY\_PATH**, and other. **modulefiles** may be shared by users on a system and users may have their own collection to supplement or replace the shared **modulefiles**.

# Torque/Moab

## TORQUE

Resource Manager, a computer application for controlling unattended background program execution of jobs and interactive jobs.

## MOAB

Cluster workload management package, available from Adaptive Computing, Inc. that integrates the scheduling, managing, monitoring and reporting of cluster workloads.



# Torque/Moab

Torque

**qsub**

**qstat**

Moab

**checkjob**

**mjobctl**

**showq**

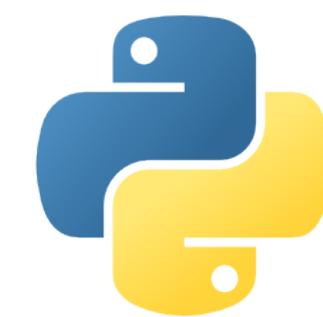
**showstart**

# Scientific Programming Languages

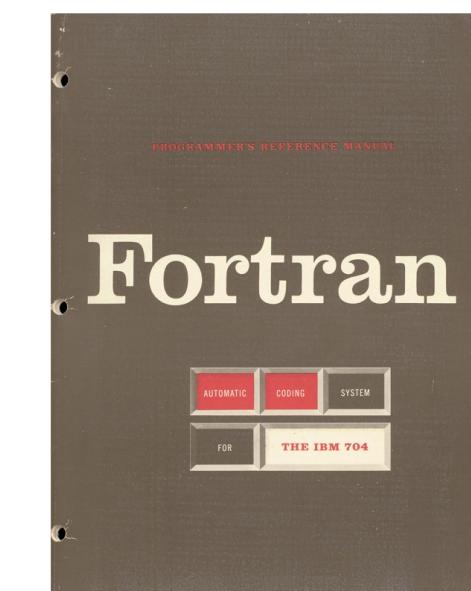


julia

UPPER



python™

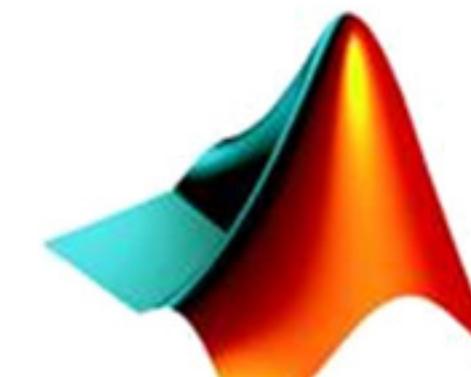


LOWER



Maple™

COM



MATLAB®



Wolfram  
Mathematica®



# Parallel Computing

MPI

Distributed  
Parallel  
Computing

POSIX Threads

Shared Memory  
Parallel  
Computing

OpenMP

Annotated Programming

CUDA

GPU  
Computing

OpenACC

# Parallel Algorithmic methods

- Linear Algebra
  - Dense linear algebra
  - Sparse linear algebra
- Spectral methods
  - FFT
- N-body problems
- Grid Methods
  - Structured
  - Unstructured
- Monte Carlo method
- Combinational logic
- Graph traversal
- Dynamic programming
- Branch and bound methods
- Graphical models
- Finite-state machine simulation

# Compiling Codes: Libraries

Library	Purpose	Linking FLAG
BLAS/LAPACK	Dense Linear Algebra	gcc code.c -lblas -llapack <b>OR</b> gcc code.c -lopenblas <b>OR</b> ifort code.f90 -L\${MKLROOT}/lib/intel64 -Wl,--no-as-needed -lmkl_intel_lp64 -lmkl_sequential -lmkl_core -lpthread -lm -ldl
FFTW	Fast Fourier Transform	gcc code.c -lfftw3
HDF5	Hierarchical Data Format	gcc code.c -lhdf5
MPI	Message Passing Interface	mpicc code.c mpif90 code.f90
GSL	GNU Scientific Library	gcc code.c -lgsl -lgslcblas -lm
Boost	Portable C++ source libraries	NONE for header only libraries
OpenMP	Shared Memory Parallel Programming	gcc -fopenmp code.c icc -qopenmp example_omp.c

# Git

