COPERNICUS MASTER IN DIGITAL EARTH

HPC FOR BIG DATA

1. Introduction

Big Data and HPC: Principles and Issues

Frédéric RAIMBAULT, Nicolas COURTY

University of South Brittany, France

IRISA laboratory, OBELIX team

Outlines

1. Big Data
2. High Performance Computing
3. Parallel Computing

1. Big Data

- No single and simple definition.
- Big Data is data that exceeds the processing capacity of conventional database systems. The data is too big, changes too fast, or doesn't satisfy the constraints of database architectures (e.g. Relational DBMS).
- Big Data is not just about size
- Within these data lie valuable patterns and information, previously hidden because of the amount of work required to extract them.

  → To gain value from this data, an alternative way (from RDBMS) to store and to process it must be applied.

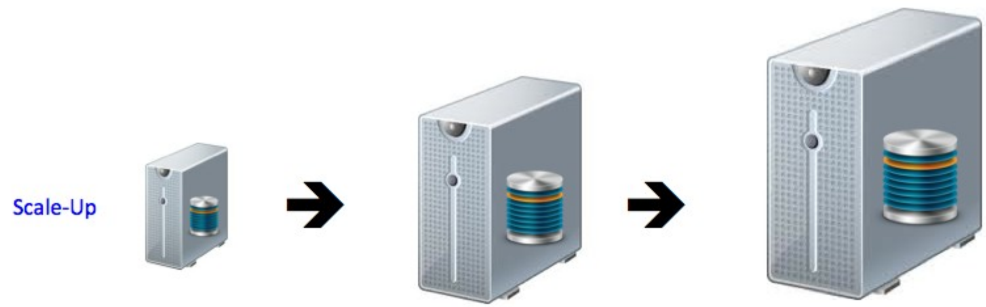
Key Concept: NoSQL vs. RDBMS

- **Relational Databases (aka SQL databases)**
  - e.g. : Oracle, MySQL, M$ SQL server, IBM Dn2, Azure SQL
  - Table-based data structure
  - With strict and predefined schema
  - Vertically scalable (by upgrading servers): high-end, costly materials
  - Closed source with licensing fees
- **Non-relational Databases (aka NoSQL databases)**
  - e.g.: Cassandra, HBase, MongoDB, Berkeley DB, Apache Giraph, Neo4J
  - Complex data structure (documents, graphs, key-value pairs, wide-column)
  - Don’t require predefined schema
    - allow to work with unstructured data and flexible data model
  - Horizontally scalable (by adding processing nodes): commodity (low-end) servers
  - Part of the open-source community

Key Concept: Scalability

- **Scalability is of most importance for Big Data**
  - Property of a system to handle a growing amount of work by adding hardware resources

- **Scale-up (or vertical scaling)**
  - Add (or replace) resources (CPU cores, memory) to a single node
  - Model: VM servers, Mainframes, Supercomputers

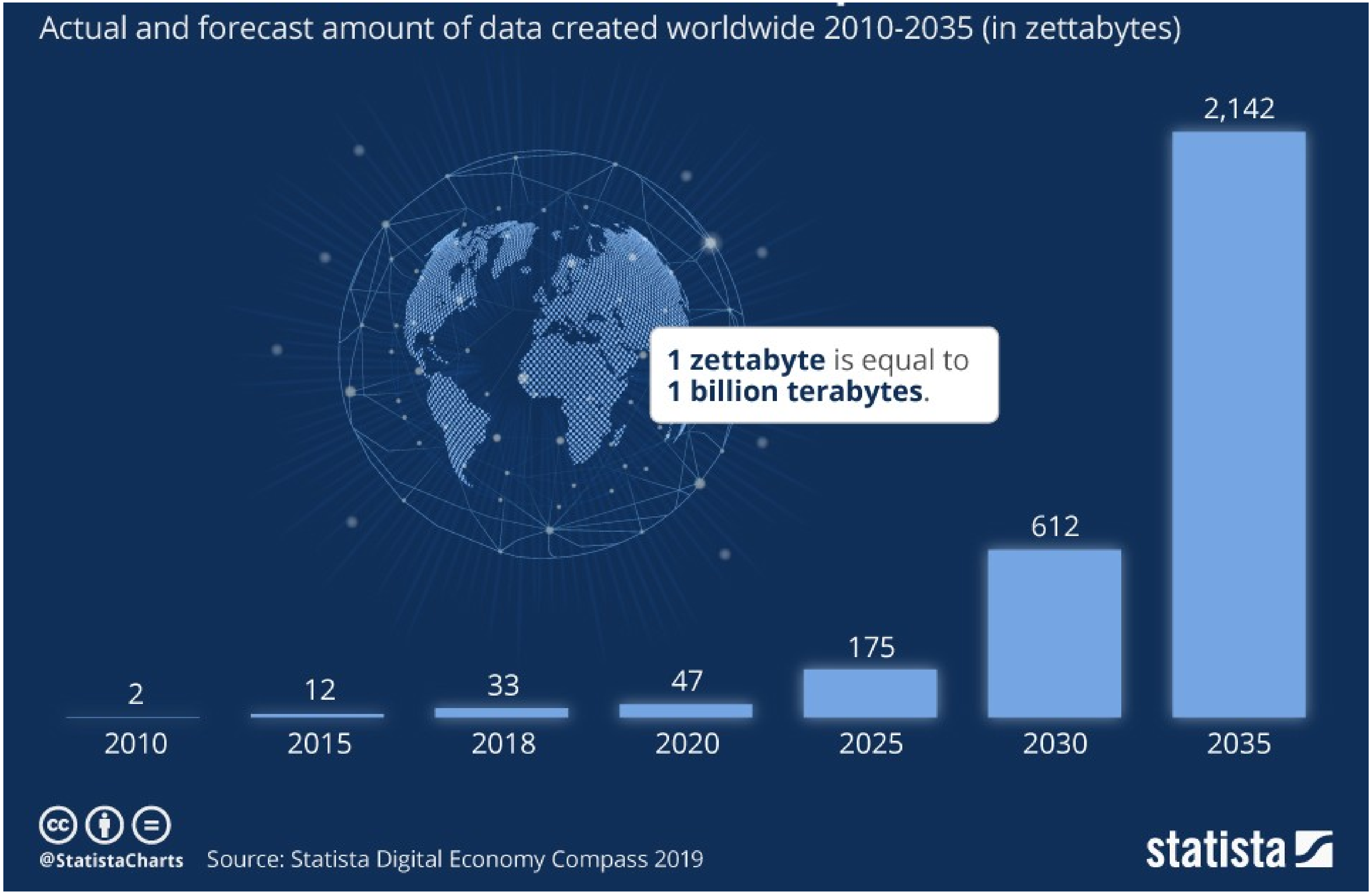


- **Scale-out (or horizontal scaling)**
  - Add more computing nodes to a system
  - Model: Data centers, Cloud infrastructures

Input Data

- **Data comes from**
  - Interactions
    - WWW, social networks, ...
  - Transactions
    - Online DB, online markets, stock echange, server logs, ...
  - Observations
    - earth, nature, particules, cosmos...
- **Provided by**
  - Humans,
  - Mobile phones,
  - Smart devices (IoT),
  - Geo-located devices,
  - Satellites and remote sensing instruments
  - ...

Global Data Creation is about to Explode

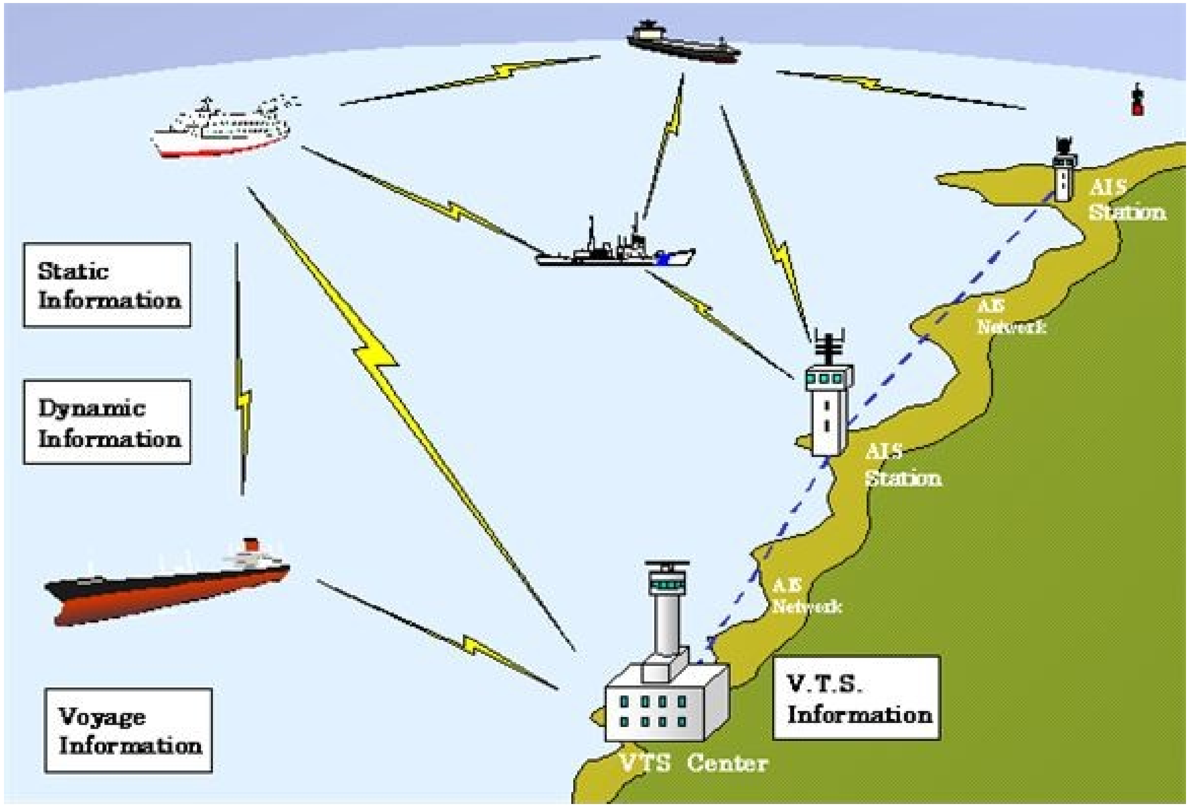


Data Example #1: AIS-UBS

- **Automatic Identification System**
  - An automated autonomous system for the exchange of navigational information between suitably equipped vessels and shore stations operating on designated marine VHF frequencies.
  - Avoiding ship collisions, monitoring traffic, assisting in navigation and Search And Rescue operation
- **Equipment classes**
  - Mandatory (class A) for regulated vessels
  - Optional (class B) for non-regulated vessels
  - AtoN for lighthouses, buoys

Navigational Information Broadcasts

- **Time Stamp**
- **MMSI number**
- **Name, type of ship**
- **Call sign, dimensions**
- **Latitude, Longitude**
- **Speed Over Ground**
- **Course Over Ground**
- **Class A only:**
  - IMO number, draught, destination, ETA
  - Rate of turn, navigational status



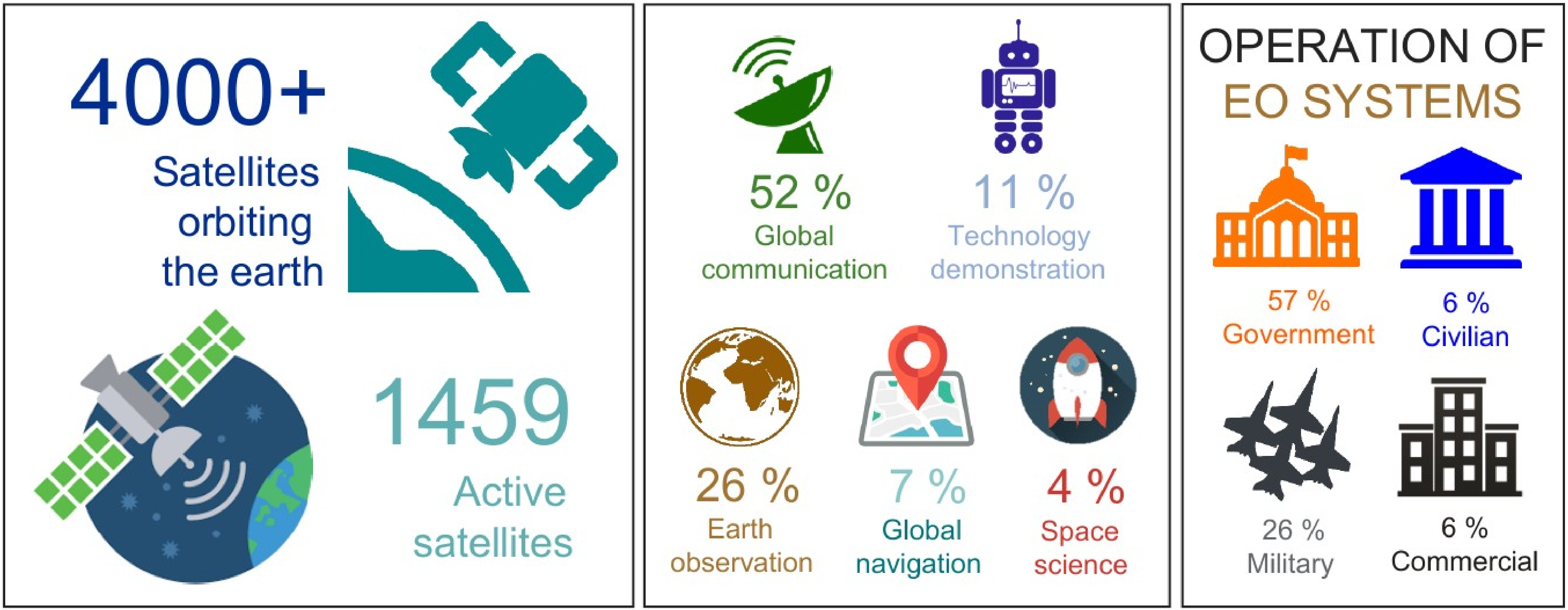
UBS Receiving Station

- **An antenna collecting AIS signals around the Morbihan Gulf and forwarding the raw data to the AISHUB community**
- **A website (ais.univ-ubs.fr) aggregating the AIS feed shared by the other members**
  - 800 stations around the world + satellites: Free AIS Ship Tracking of Marine Traffic
- **A storage accumulating worldwide AIS data since 2015**
  - compressed daily volume of 800 MB (250 GB / year).

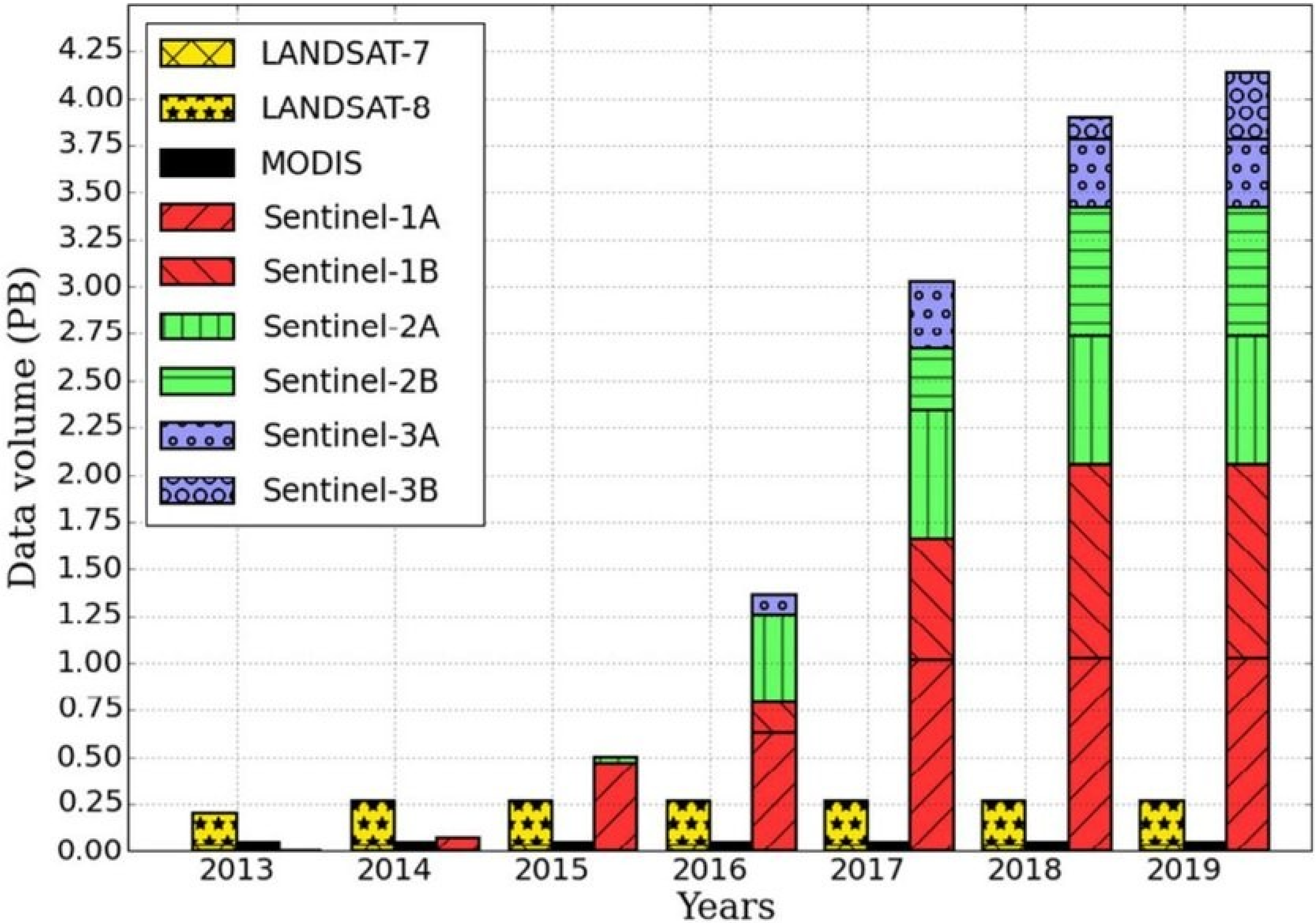
Practical Lab #1 (1)

- **Browse the ais.univ-ubs.fr website**
- **Connect to the `dmis` machine**
  - Read this: http://cluster-irisa.univ-ubs.fr/wiki/acces-ssh
- **Explore the worldwide dataset we own (files are located under `/share/projects/ais/world/20XX`)**
  - Compute its compressed size
  - Estimate its uncompressed size (use `bunzip2` command)
  - Look at the format and at the content
    - Fields are described here: http://www.aishub.net/api
- **Find some interesting statistics and the way to compute them**
  - And estimate the required compute time

Satellite Big Data



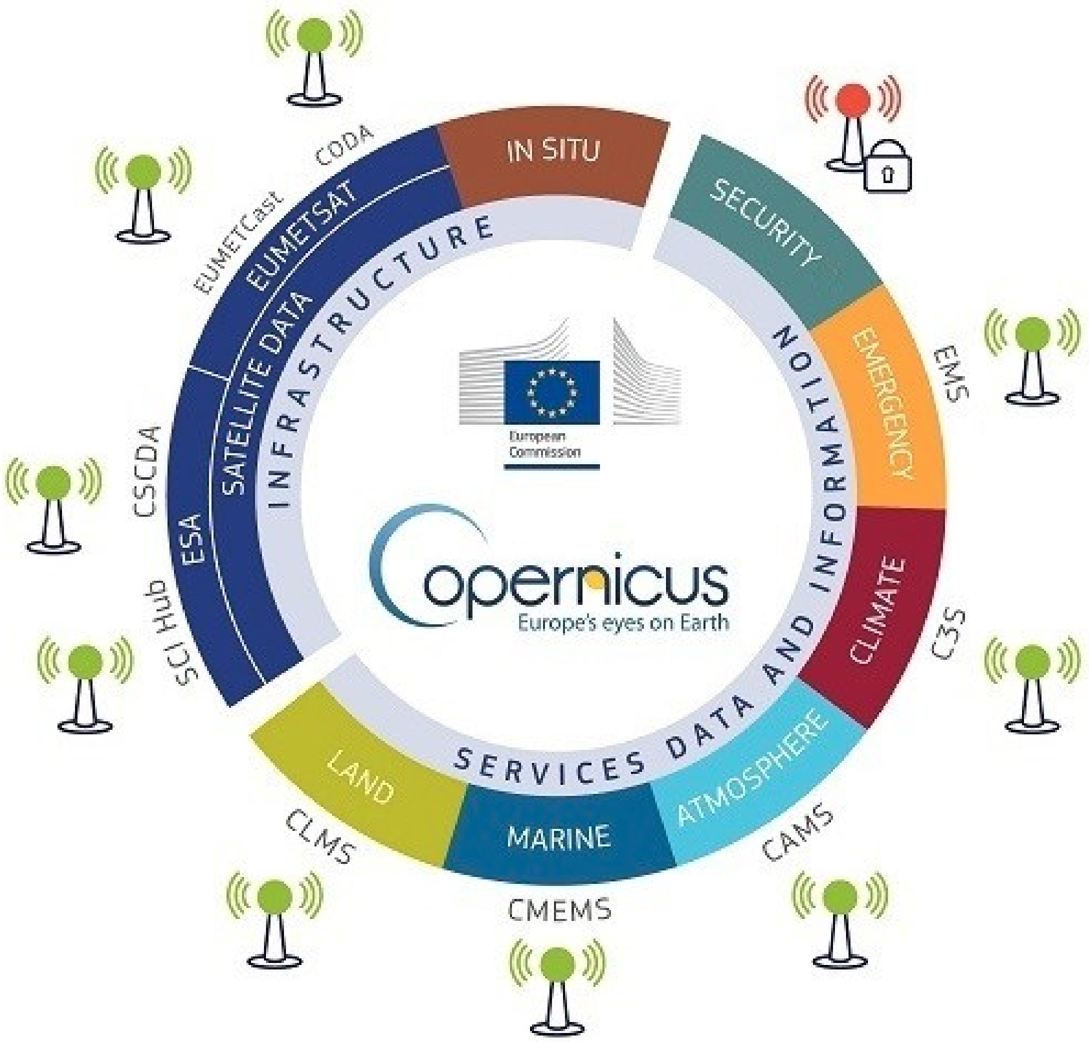
Big Data from Earth Observation Satellites



Data Example #2: Copernicus

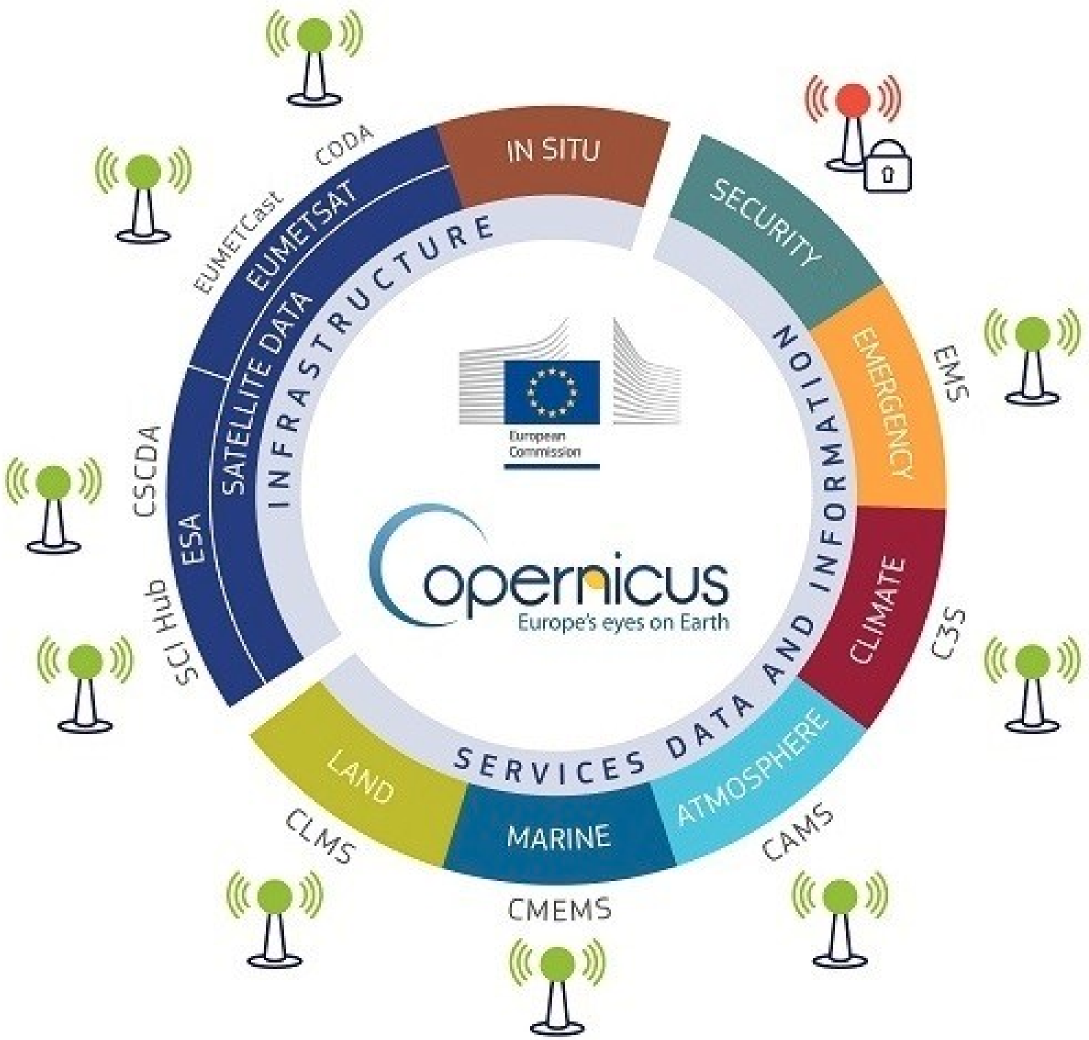
- **Earth Observation Program from EU builds on**
  - A constellations of satellites (Sentinel-1,2,3,5P)
  - A global network of sensors
- **Producing 12 TB per day**
  - Free, full and open access basis
  - Six main domains (Atmosphere, Marine, Land, Climate, Emergency and Security)
  - Accessed by several access points
    - Satellite data access hubs (free)
    - Online services data platforms (commercial services)

Copernicus: Data Access Hubs (1)



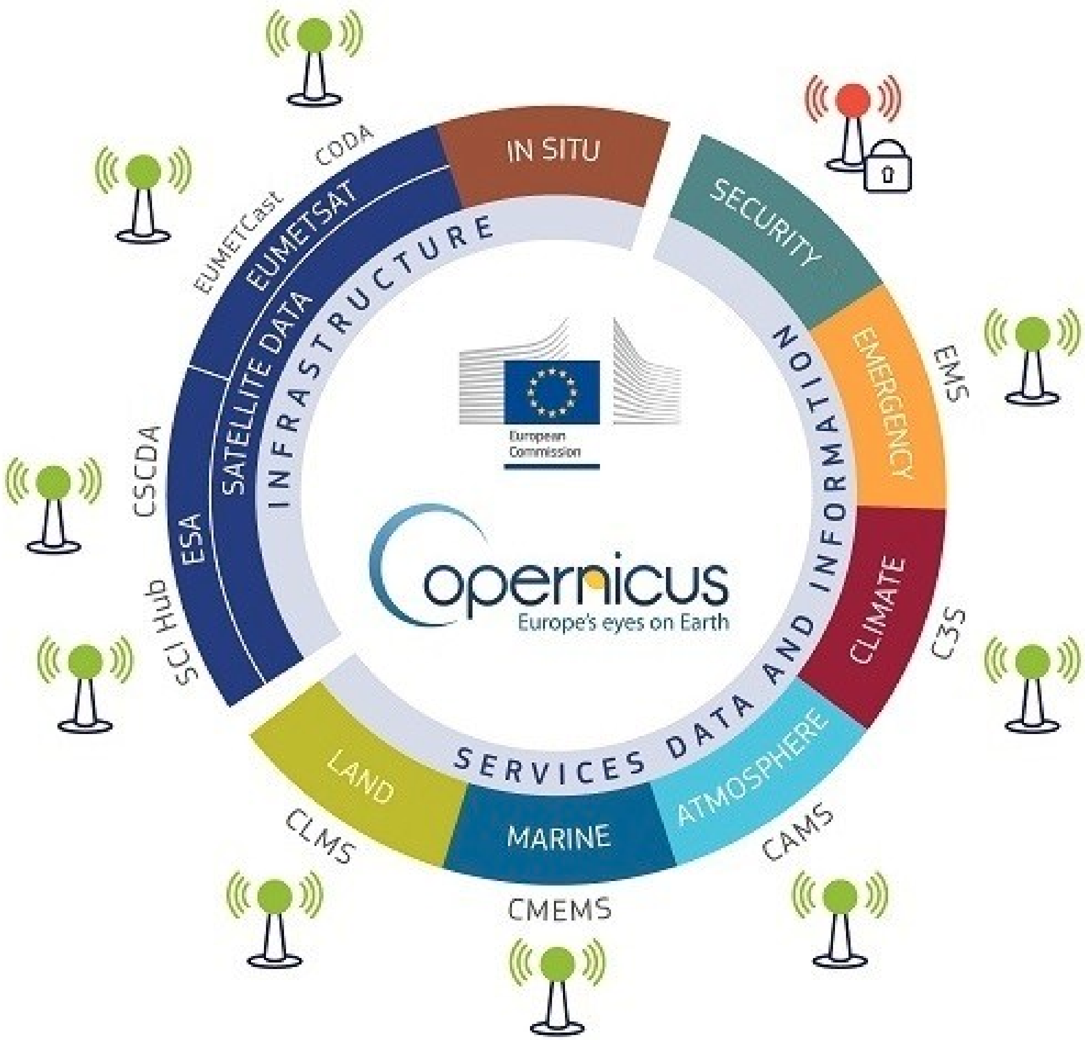
- **ESA: European Space Agency**
  - CSCDA: Copernicus Space Component Data Access portal
  - SCIHub: Copernicus Open Access Hub

Copernicus: Data Access Hubs (2)



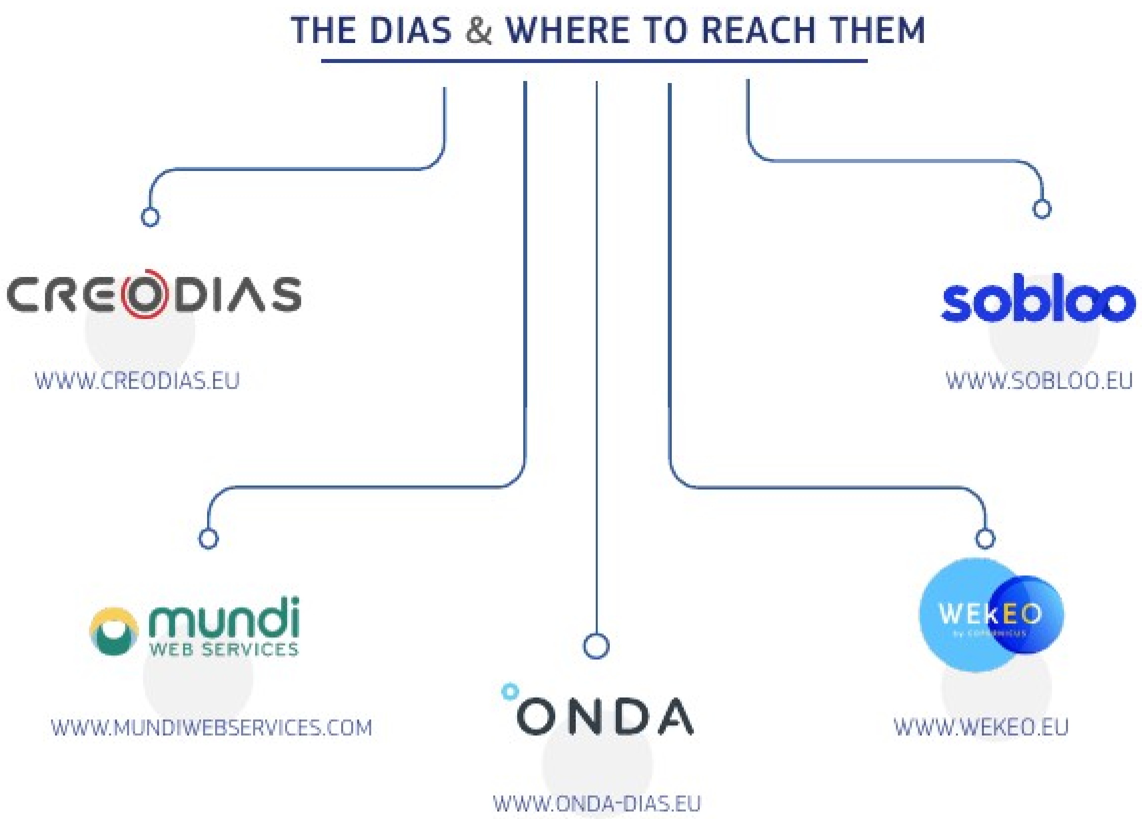
- **EUMETSAT: European Organisation for the Exploitation of Meteorological Satellites**
  - EUMETCast
    - near real-time delivery of satellite data
  - CODA: Copernicus Online Data Access

Copernicus: Data Access Hubs (3)



- **Service data and information in six thematic areas:**
  - CLMS: Copernicus Land Monitoring Service
  - CMEMS: Copernicus Marine Environment Monitoring Service
  - CAMS: Copernicus Atmosphere Monitoring Service
  - C3S: Copernicus Climate Change Service
  - Security
    - Border and Maritime surveillance ; EU external action
  - EMS: Copernicus Emergency Management Service

Copernicus: DIAS

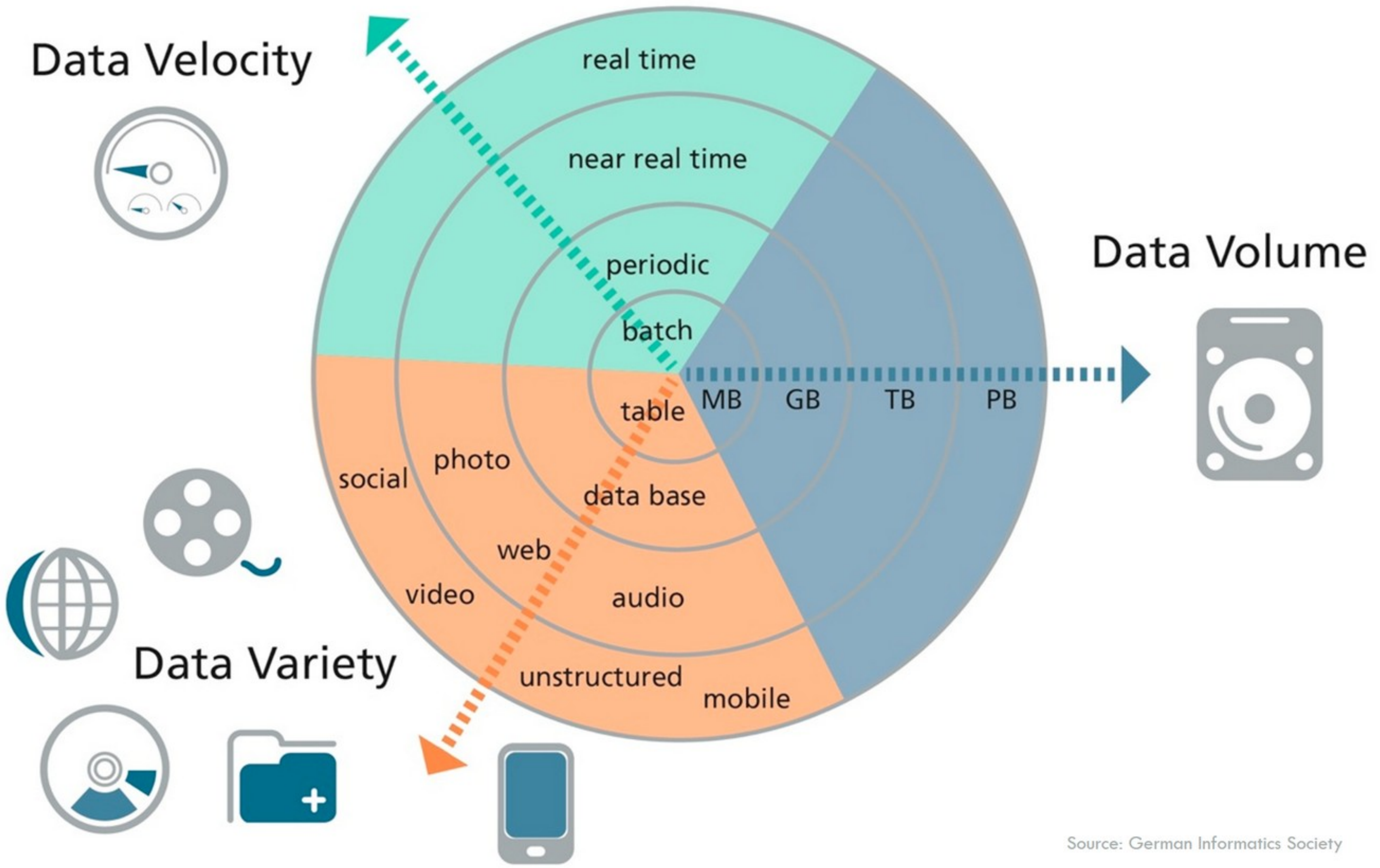


- **Data and Information Access Services**
  - 5 cloud-based platforms providing centralized access to Copernicus data and processing tools
  - Provides access to additional commercial satellite or non-space data sets

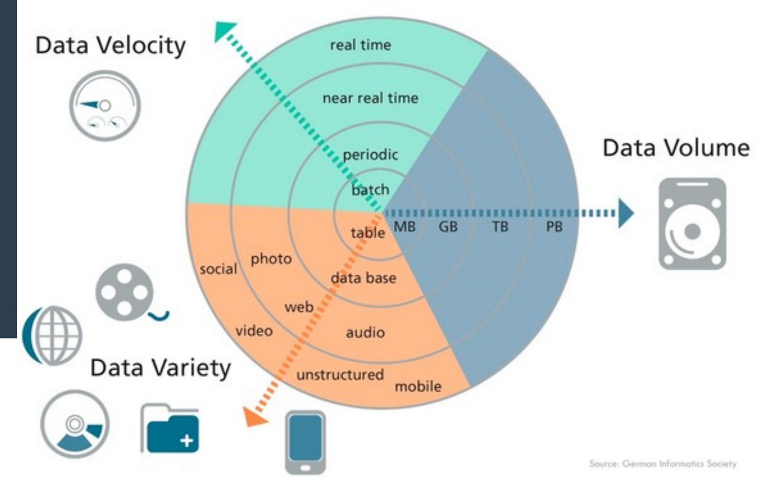
Practical Lab #1 (2)

- **Browse the Copernicus website from the URL https://www.copernicus.eu/en/access-data**
- **Perform self-registration on the Copernicus open access hub**
- **Explore the available datasets and ways to access them**
  - Try to download some data
  - Compare the format of the downloaded data with AIS data format
- **Look at the services (DIAS) and their added value**

Properties and Challenges (3V)



Data Volume

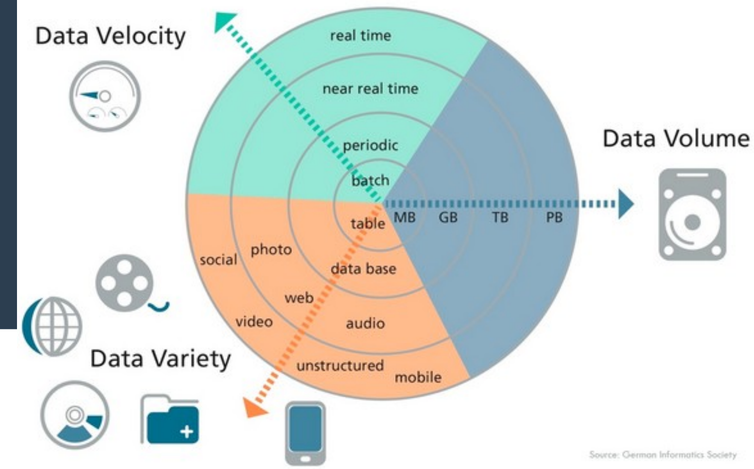


- **Having more data beats out having better models**
  - Keep everything, keep raw (source) data

- **90% of all data has been created in the last two years (source: IBM 2017)**
- **The digital universe is expected to reach 168 ZB by 2025**
  - It would take a person 181 million years to download all the data from the Internet (at 46 Mbps)
- **Statistics per day: A Day in Data (visualcapitalist.com 2019)**
  - In 2020, every person generates 1,7 MB per second (source: Domo US software company)
- **Storage requirements are huge**
  - Data management implies that storage centers constantly need capacity and transmission upgrades: distributed nodes, high-speed network, effective cooling systems, fault tolerance systems...

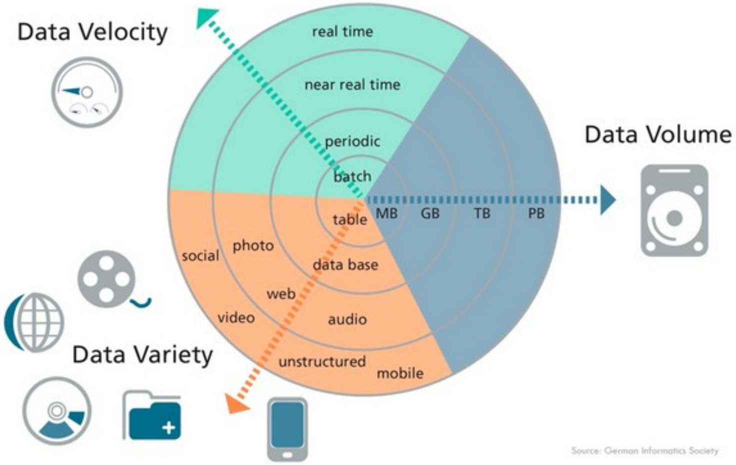
Memory unit	Size	Binary size
kilobyte (kB/KB)	10^3	2^{10}
megabyte (MB)	10^6	2^{20}
gigabyte (GB)	10^9	2^{30}
terabyte (TB)	10^{12}	2^{40}
petabyte (PB)	10^{15}	2^{50}
exabyte (EB)	10^{18}	2^{60}
zettabyte (ZB)	10^{21}	2^{70}
yottabyte (YB)	10^{24}	2^{80}

Data Velocity



- **Frequency of incoming data and speed at which it is generated, processed and transmitted**
- **Depending of the application, data can be analyzed in batch, e.g.**
  - Statistics on worldwide maritime traffic
  - Supervised classification of satellite image
- **or need to be analyzed in (near) real time, e.g.:**
  - Monitoring local maritime traffic with AIS
  - Supervising national border with satellite
- **Depending of the nature of data, it is said to age like wine (longer it's kept, the more insights you glean from it), e.g.**
  - Study of historic evolution
- **or it may has a limited-shelf life and erode quickly**
  - Decision to be taken quickly so as not to miss an opportunity

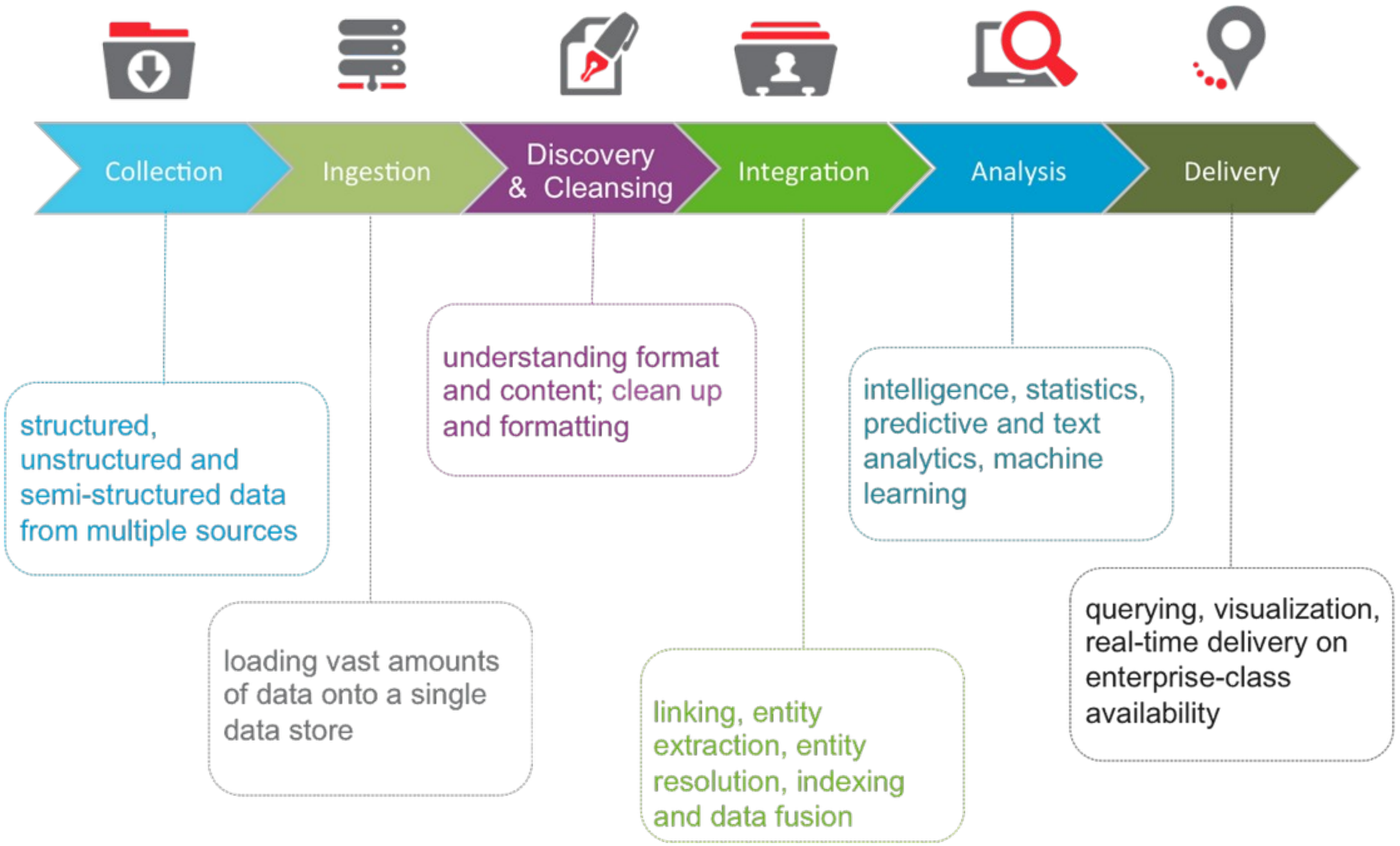
Data Variety



- **Variety refers to the number of types of data**
- **Structured**
  - From satellites sensors (images)
  - Variety on resolution (spectral, temporal, spatial, radiometric) depending on instruments purpose
- **Semi-structured or unstructured (raw)**
  - From remote sensors (mobile phones, IoT devices, scanners, cameras, radars...)

Big Data Typical Workflow

- Capturing data
- Ingesting data
- Storing the data
- Filtering and processing
- Making data available for analytics
- Analysis
- Vizualizing results

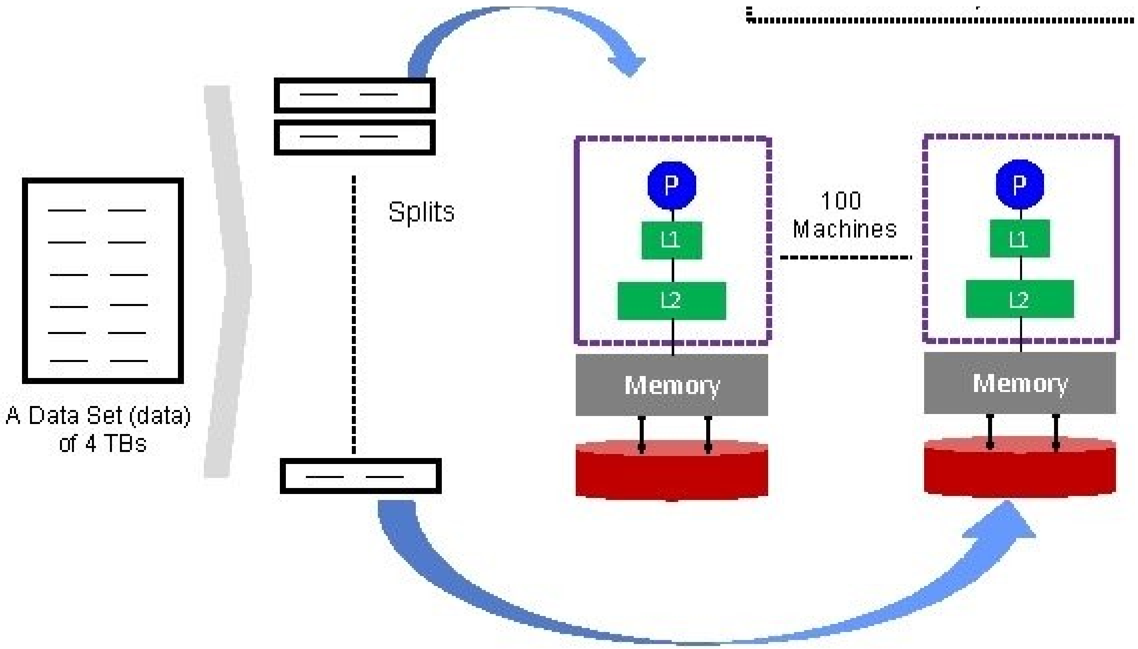


Some Issues of Big Data Analytics

- **Fast generated data are frequently corrupted, badly formatted, inconsistent, e.g.:**
  - Incorrect position in an AIS signal (VHF+compressed format) leads to the calculation of a outlier speed for a vessel.
  - Missing data for a certain period of time affect temporal statistics
  - Clouds that mask part of satellite images distorts count results
- **Data volume called for distributed storage**
  - The proliferation of storage devices leads to a high probability of failure when running a program, e.g.:
    - MTBF of 1 HDD is 3 years => MTBF of 1000 HDD is 1 day
    - MTBF of 1 SSD is 250 years => 80 days for 1000 SSD
  - Failed tasks are common and need to be managed
- **Distributed storage implies distributed processing**
  - Require complex frameworks and special programming skill

Key Concept: distributed systems

- **Distribute the data to several nodes memory**
  - Local file system => distributed file system



- **Distribute the processing**
  - Sequential running task => parallel running tasks

Big Data Technology Landcape

- **The Apache Hadoop stack**
  - makes large-scale computing flexible and affordable
  - Inspired by Google's internal tools
- **HDFS (Hadoop Distributed File System)**
  - Scalable storage of big data on commodity servers
  - Data files are broken into multiples blocks that are replicated and stored across a cluster of servers
- **MapReduce**
  - Batch-processing framework performing parallel processing tasks on data stored into HDFS
  - Processes and reduces raw big data into regular-size, tagged datasets that are easier to work with

Other tools related to Hadoop

- **Pig**
  - Dataflow language for building big data applications.
  - Generates MapReduce in the background
- **Hive**
  - Data warehouse software for querying data from a hadoop system using a SQL-like language
- **HBase**
  - A NoSQL database designed for storing and retreiving very large datasets
  - This wide column-store runs on top of HDFS and store sparse data
- **Spark**
  - In-memory data processing and analysis
  - enable near real-time big data analytics

Among other open-source frameworks

Alternative #1: MPP Databases

- **Teradata, ParAccel, Netezza (IBM) Vertica (HP), SQL PDW (Microsoft), Redshift (Amazon) … products**
- **Massively Parallel Processing databases are scalable, high performing, parallel processing alternatives to MapReduce**
  - Deployed on expensive, specialized hardware tuned for CPU, storage and network performance
  - Bounded by the high cost of the hardware and the software
  - Supports mostly structured data

Alternative #2: Cloud Computing

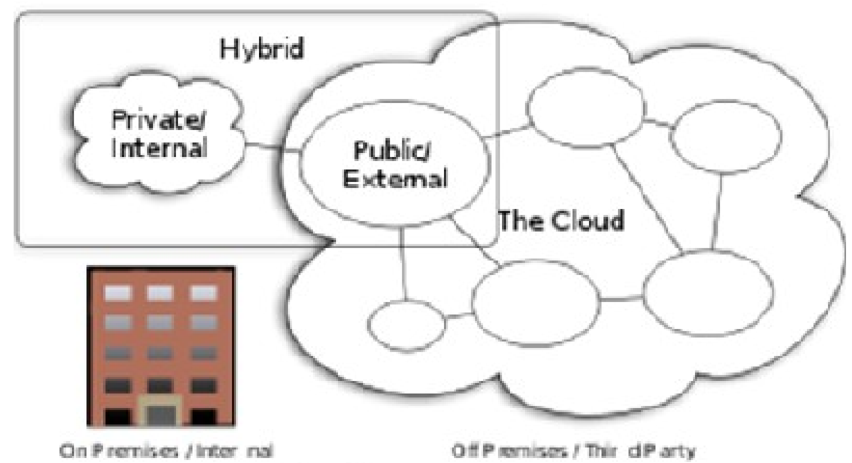
- **IT resources provided as a service**
  - Compute, storage, databases
- **Clouds leverage economies of scale of commodity hardware**
  - Cheap storage, high bandwidth networks & multi-core processors
  - Geographically distributed data centers
- **Offerings:**
  - Amazon Elastic Compute Cloud (EC2), Simple Storage Service (S3)
  - Google App Engine
  - Microsoft Azure Platform
  - GoGrid
  - AppNexus

Benefits of Cloud Computing

- **Cost & management**
  - Economies of scale, “out-sourced” resource management
- **Reduced Time to deployment**
  - Ease of assembly, works “out of the box”
- **Scaling**
  - On demand provisioning, co-locate data and compute
- **Reliability**
  - Massive, redundant, shared resources
- **Sustainability**
  - Hardware not owned

Types of Cloud Computing

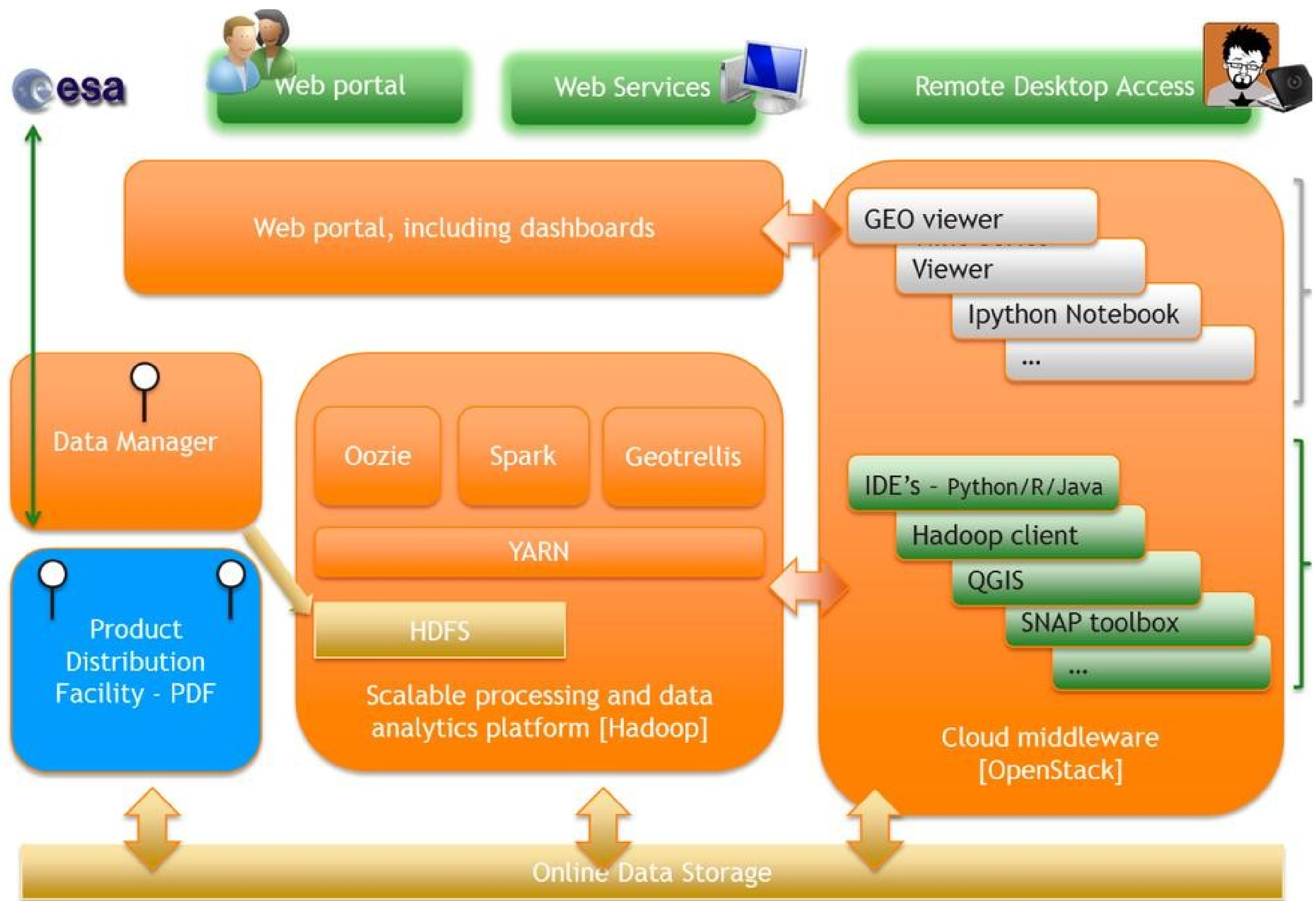
- **Public Cloud: Computing infrastructure is hosted at the vendor’s premises.**
- **Private Cloud: Computing architecture is dedicated to the customer and is not shared with other organizations**
- **Hybrid Cloud: Organizations host some critical, secure applications in private clouds. The not so critical applications are hosted in the public cloud**



  - Cloud bursting: the organization uses its own infrastructure for normal usage, but cloud is used for peak loads.
- **Community Cloud**

Example: Proba-V from ESA

- An operational MEP (Mission Exploitation Plateform)
- Improve the exploitation of the SPOT-Vegetation EO data archive (petabyte range)
- Private cloud + Hadoop-based processing environment

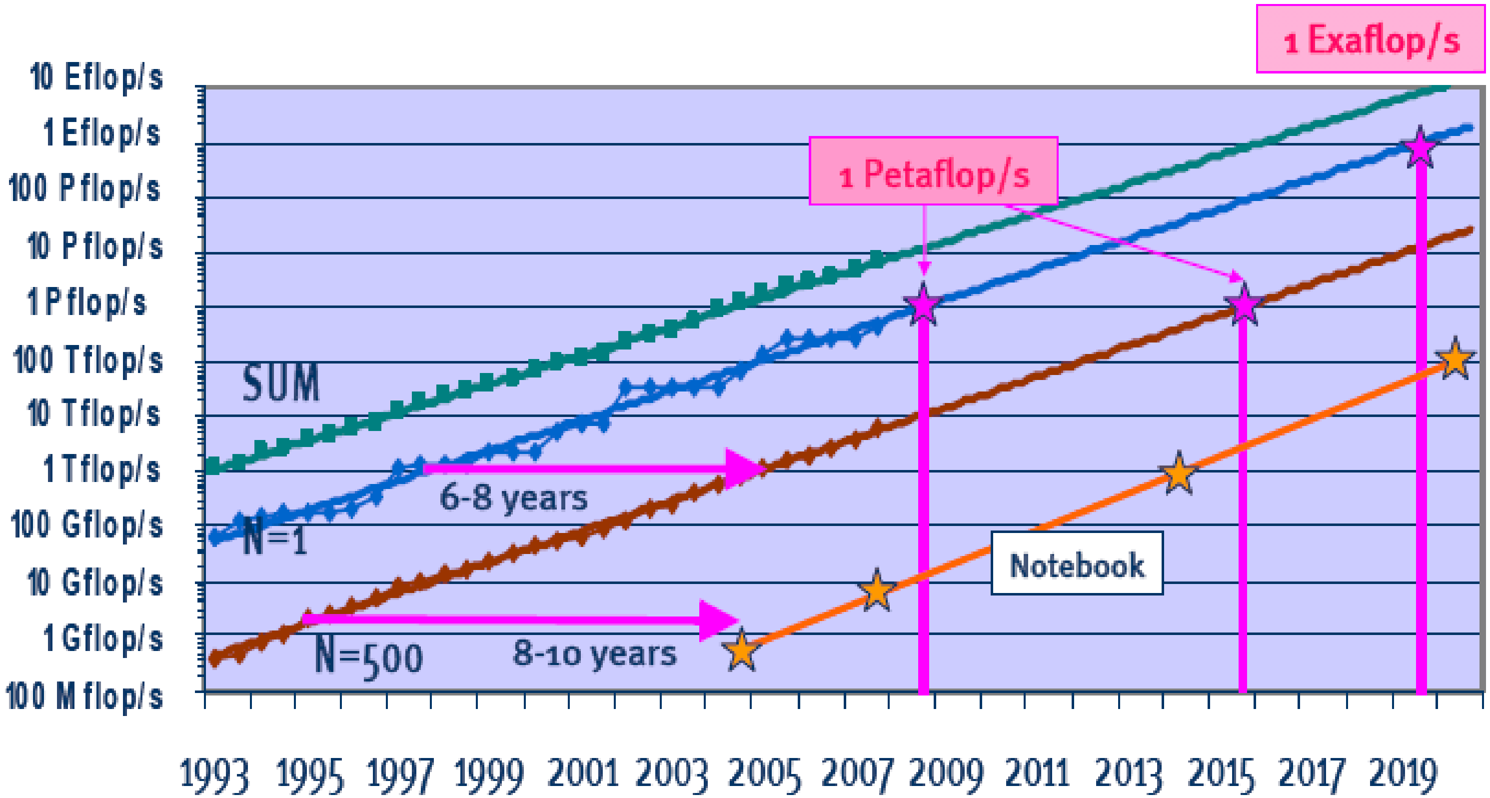


Outlines

1. Big Data
2. High Performance Computing
3. Parallel Computing

2. High Performance Computing

- **Refers to computing on High Performance Computers**
  - HPC computers deliver more than 1 PFlop/s (expected to reach soon 1 EFlop/s)



Key Concept: FLOP/S

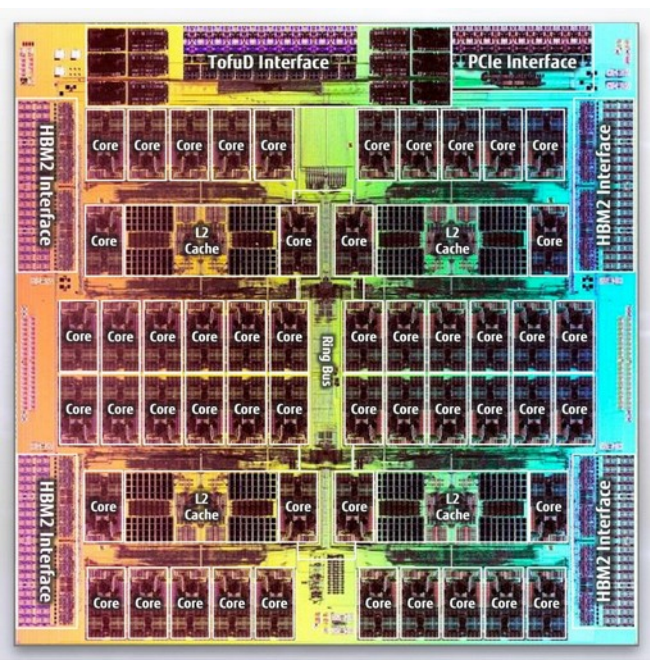
- **Floating Point Operation Per Second: measure of computational performance when operating on real numbers.**
- **On HPC system:** $\text{FLOPS} = \text{racks} \times \frac{\text{nodes}}{\text{rack}} \times \frac{\text{sockets}}{\text{node}} \times \frac{\text{cores}}{\text{socket}} \times \frac{\text{cycles}}{\text{second}} \times \frac{\text{FLOPs}}{\text{cycle}}$
  - Top1: 500 PFlop/s
- **On laptop or tabletop computer with 1 CPU:** $\text{FLOPS} = \text{cores} \times \frac{\text{cycles}}{\text{second}} \times \frac{\text{FLOPs}}{\text{cycle}}$
  - Current microprocessors core: 10 to 200 Gflop/s
  - Current GPUs: up to 10 TFlop/s
- **Whereas RDBMS refers to Transaction Per Second unit**
  - e.g. VISA claims its global payment system can handle 24,000 TPS.

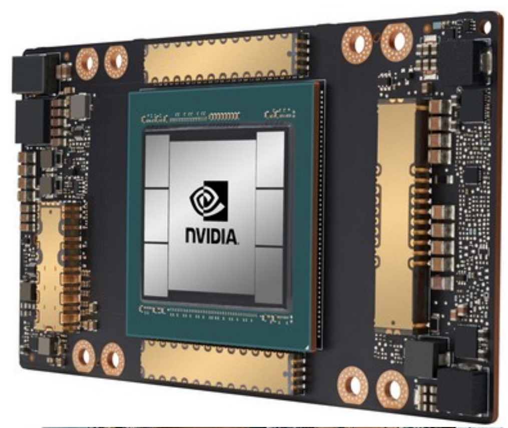
Use of HPC

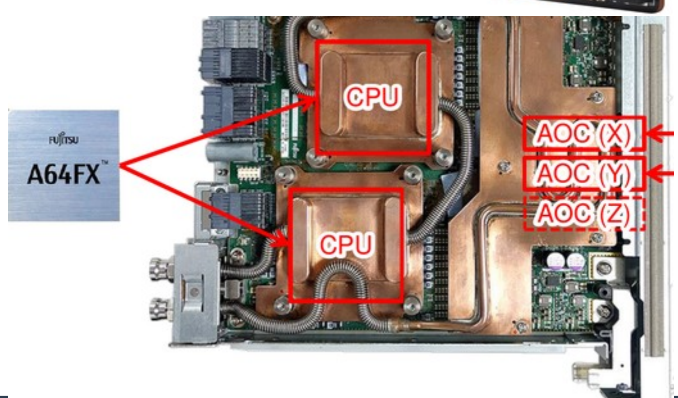
- **Solve scientific and engineering problems computationally complex in many domains**
  - **Societal challenges**: DNA sequencing, research and early detection and treatment of diseases, deciphering the human brain, forecasting climate evolution, space observation, preventing and managing large-scale natural disasters, designing renewable energy parks
  - **Science**: HPC and Big Data analysis provide scientist with deeper insights into previously unexplored areas and systems of the highest complexity, driving the innovation and discovery of almost all scientific disciplines, from earth sciences (climate modelling, weather forecast) to high-energy physics and to astrophysics and space exploration
  - **Industry**: significantly reduce R&D costs and development cycles, producing higher quality products and services, reducing the time of product development cycles thanks to numerical simulations.
  - **National security and defense**: increase safety and security, by simulating nuclear reactions, designing new aircraft fighters, developing strong encryption technologies or helping security authorities to access encrypted communications and perform large-scale suspicious pattern detection or face recognitions to secure sensitive areas

HPC Hardware

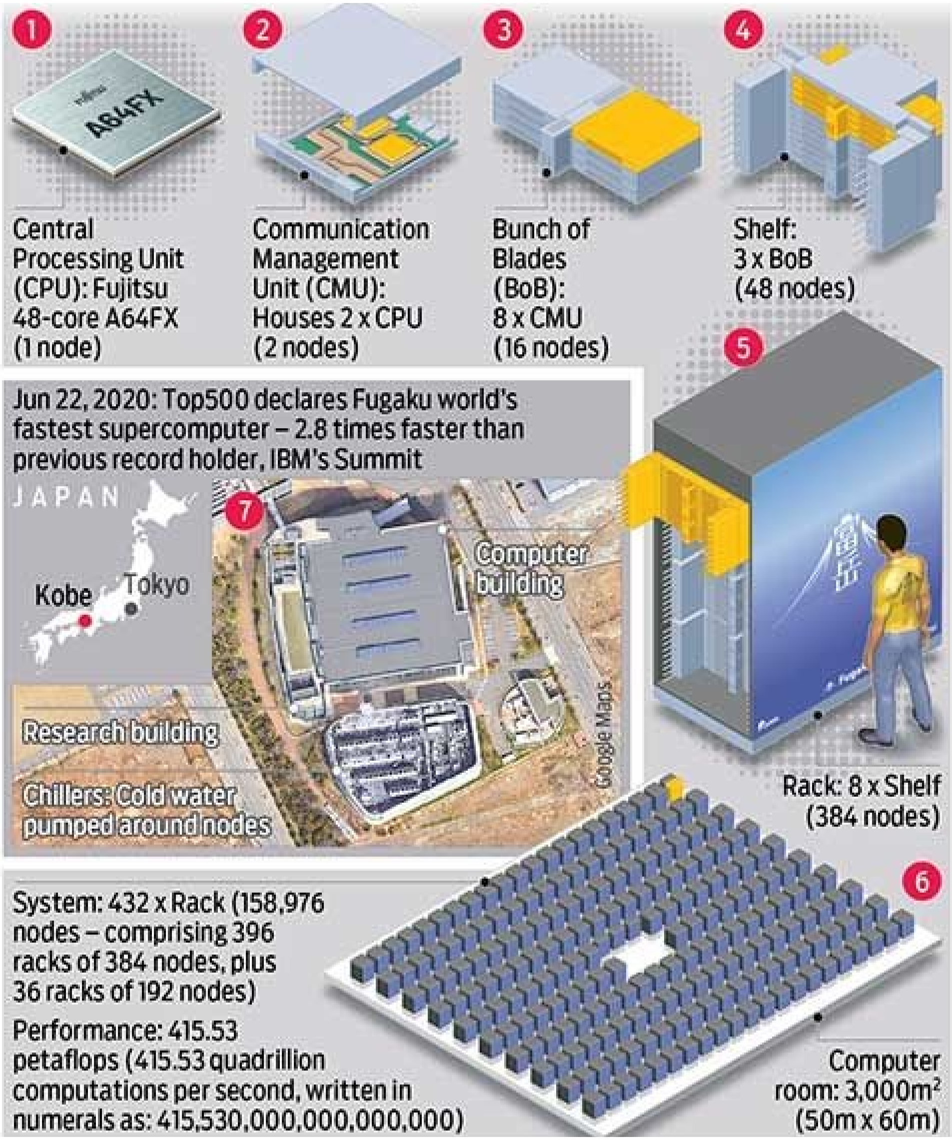
- **Evolved from specialized proprietary parallel supercomputers (too costly) to clusters of high-end processors and GPUs (more scalable)**
- **Key components: processors+memories+network**







Example: Fugaku



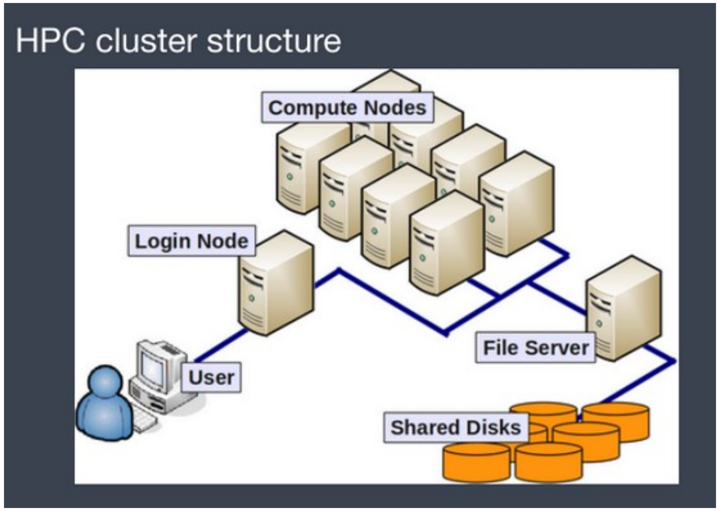
- Manufacturer: Fujitsu
- Rank Top500 in 2021: 1
- Processor: ARM A64FX
- Cores: 7,299,072
- Memory: 5 PB
- Interconnect: tofu 3D
- Power: 28 MW
- Cost: $1,213,000,000
- Linpack: 415 PFlop/s

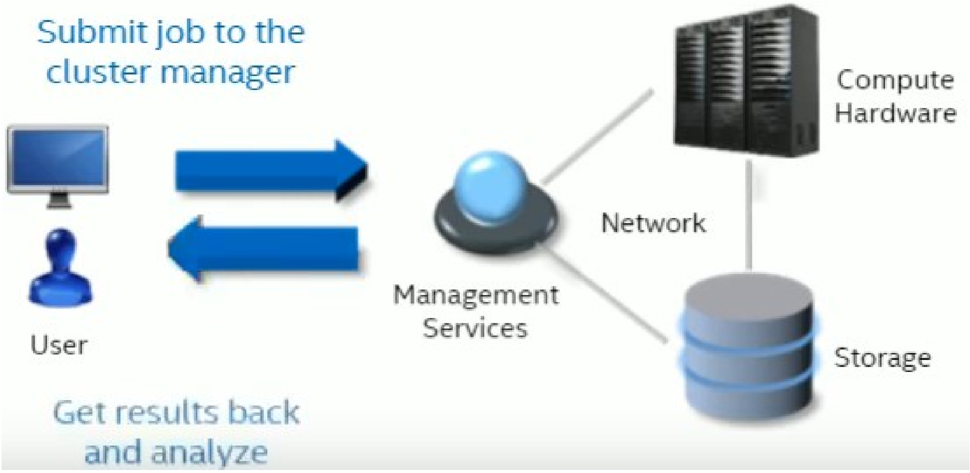
HPC Characteristics

- **Historically HPC systems has focus primarily to use the maximum computing power to solve a single large problem in the shortest amount of time**
  - As opposed to cost-effective computing aiming at solving few large problems or many small problems
- **Numerical simulation oriented**
  - Many numerical libraries (serving as benchmark, e.g. Linpack, LU decomposition of a large matrix)
- **Workflow: Model → Algorithm → Code**
  - Every step is optimized for exactly one supercomputer architecture
    - Hardware and software are highly intricate

Usage of a (HPC) Cluster

- **A cluster manager runs workload on distributed resources such as CPUs, GPUs and Disk drives, all interconnected via a network**
- **Jobs are submitted to the cluster manager using a web form, a GUI or a remote shell (ssh).**





Key point: (OS) Shell

- **A Command Line Interface to interact with a computer OS**
  - more efficient, more complete, more reproducible, more widespread (on clusters) and faster than a GUI
- **Prompt**
  - Indicating that the shell is waiting for input
  - May contain the current working directory name
- **Syntax: command [options argument*]***

```
[dmis:~]> ls -l -h /share/projects/ais/world/2021
```

- **Historic and Edition:** keystrokes ▲ ▼ ► ◄
- **Completion:** keystrokes **tab**
- **Scripting language**
  - File containing sequence of shell commands

Key point: Remote Secure Shell (ssh)

- **Keys based authentication**
  - Private key: should be keep secret and not changed

```
[avoranfix:~]> ssh-keygen -t ecdsa
```

  - Public key: should be saved on the remote machine

```
[avoranfix:~]> ssh-copyid -i ~/.ssh/id_ecdsa.pub raimbaul@dmis
```

- **CLI tools to gain remote access to a machine from a shell**
  - Without the need to type a password

```
[avoranfix:~]> ssh raimbaul@dmis
[dmis:~]>
```

  - Also used to launch a remote command or to transfer files

```
[avoranfix:~]> ssh dmis ls -l -h /share/projects/ais/world/2021
```

```
[avoranfix:~]> scp raimbaul@dmis:/share/projects/ais/world/2021/*.txt .
```

High Performance Data Analytics

- **HPC and Big Data evolved separately but share many concepts, such as parallelism.**
- **Big Data analytics is becoming more compute-intensive (thanks to deep learning), while data handling is becoming a major concern for HPC (scientific computing).**
- **HPC's technologies and know-how on parallelism meet the needs of Big Data and enable complex analyses on a large scale.**
  - e.g., a Hadoop cluster feeding data to a GPU-based HPC machine

Outlines

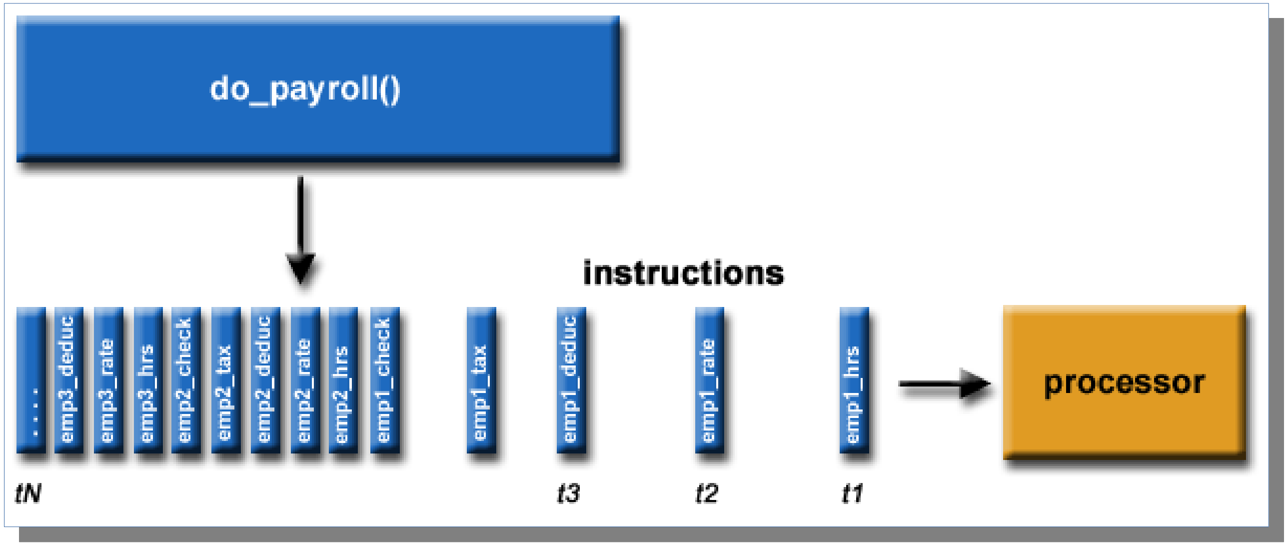
1. Big Data
2. High Performance Computing
3. Parallel Computing

3. Parallel Computing

- **Big Data and HPC share many concepts related to their implementation, one of them is parallel computing.**
- **What is parallel computing ?**
  - Simultaneous use of multiple compute resources to solve a computational problem

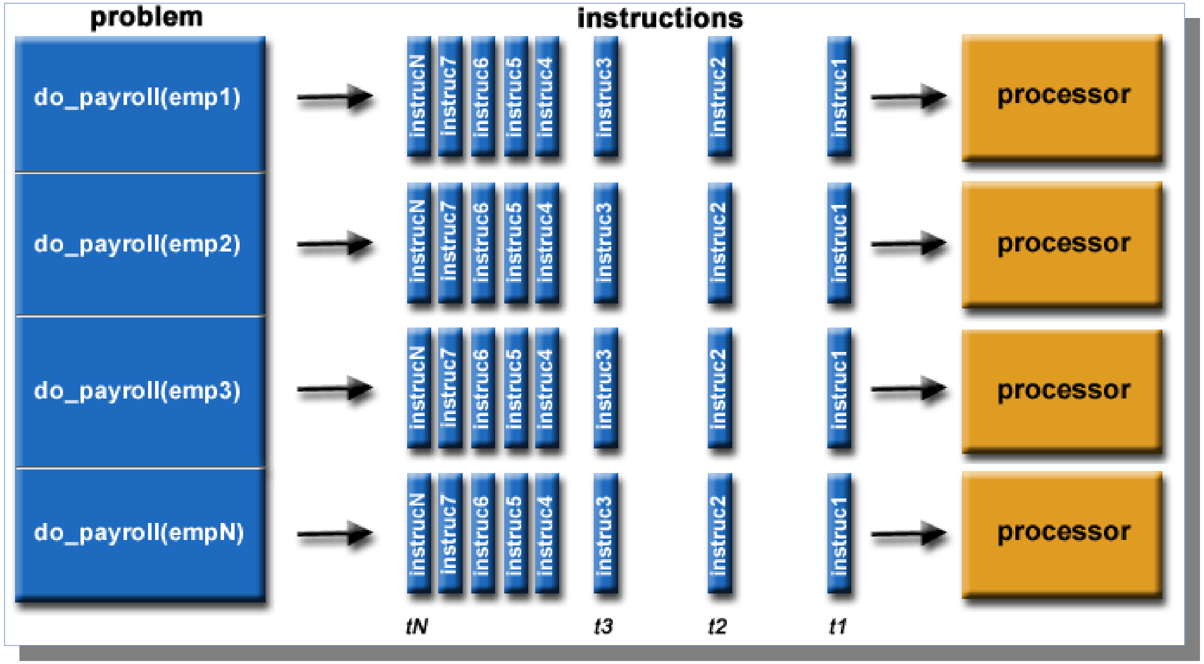
Serial Computing

- **Traditionally software have been written for serial computation**
  - The problem is broken into a series of instructions
  - Instructions are executed sequentially, one after another
  - Executed on a single processor
  - Only one instruction may execute at any moment



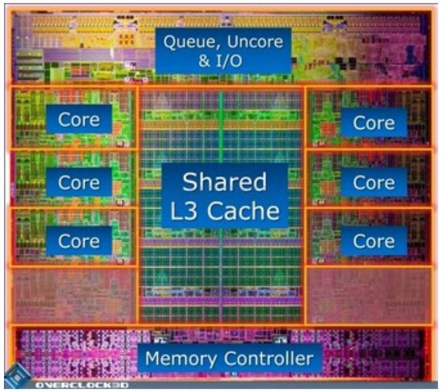
Parallel Computing

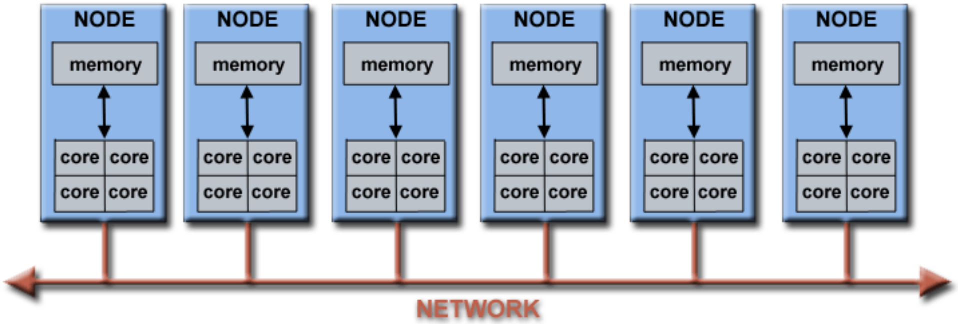
- A parallel problem is broken into discrete parts that can be solved concurrently
- Each part is further broken down to a series of instructions
- Instructions from each part execute simultaneously on different processors
- An overall control & coordination mechanism has to be employed



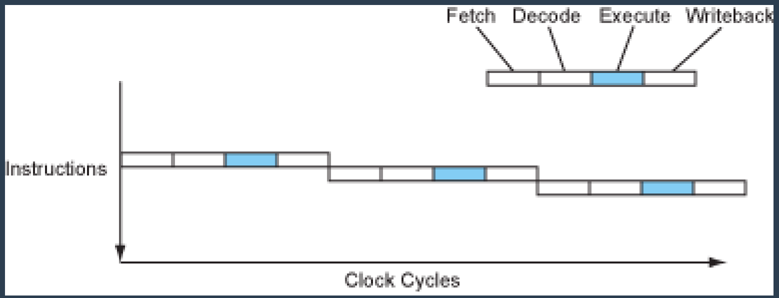
Parallel Computers

- **The compute resources are typically:**
  - A single computer with multiple processors/cores
  - An arbitrary number of such computers connected by a network
- **Virtually all stand-alone computers today are parallel from a hardware perspective:**
  - Multiple functional units
  - Multiple execution units/cores
  - Multiple hardware threads
- **Networks connect multiple stand-alone computers (nodes) to make computer clusters**





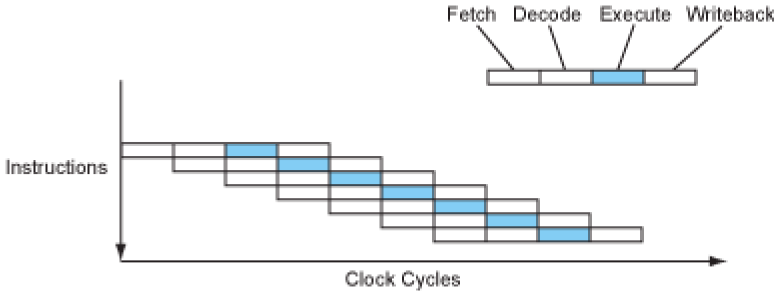
Instruction Parallelism

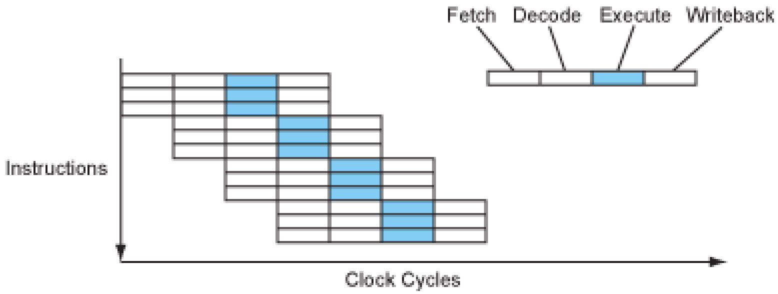


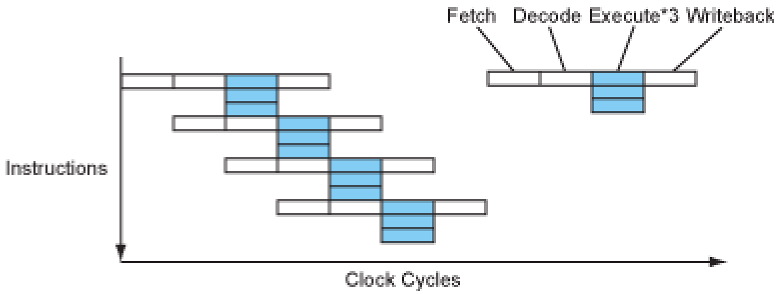
- **Parallel execution of a sequence of instructions**
  - Instruction pipeline
  - Multiple compute units
  - Out-of-order execution
    - Hardware (dynamic): Supercalar
    - Software (static): VLIW

    → 4 to 6 instructions / cycle
  - Speculative execution
  - Branch prediction

    → around 10 ins. / cycle

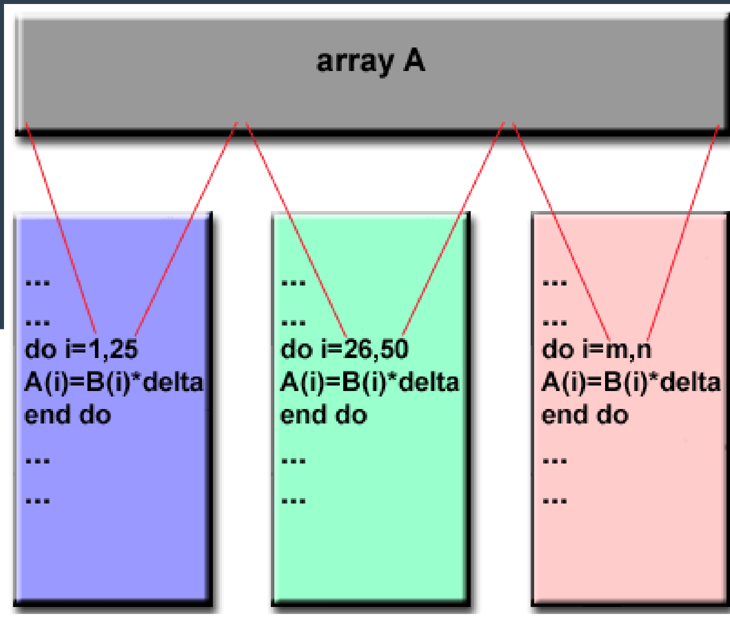


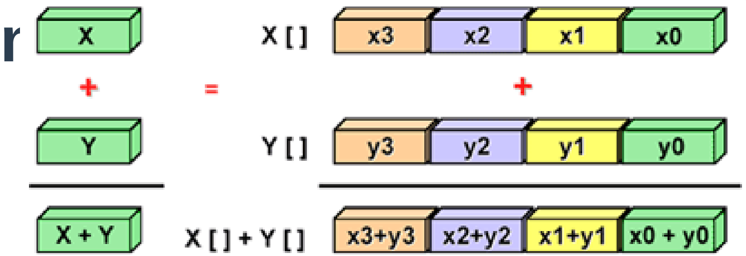


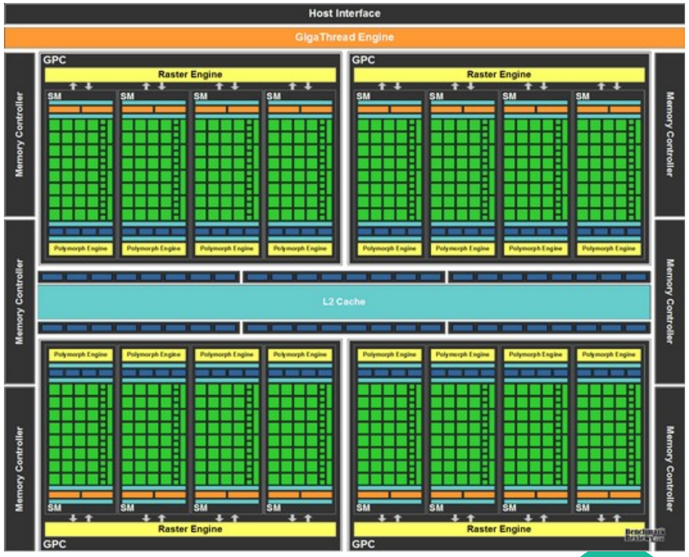


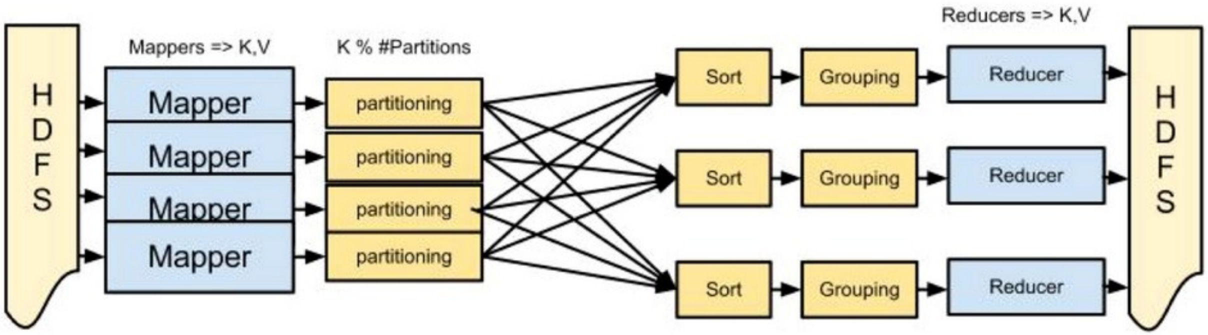
Data Parallelism

- The data set is organised into common structures (array, cube, …)
- And equally partitioned to tasks
- Each task performs <u>the same</u> operation on its own partition
- Applied to processor architecture: SIMD instructions in ISA, GPU card
- Applied to programming model: MapReduce or Spark framework



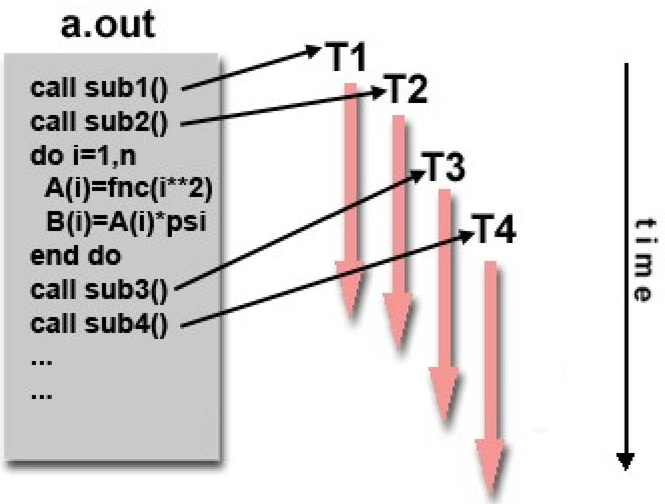


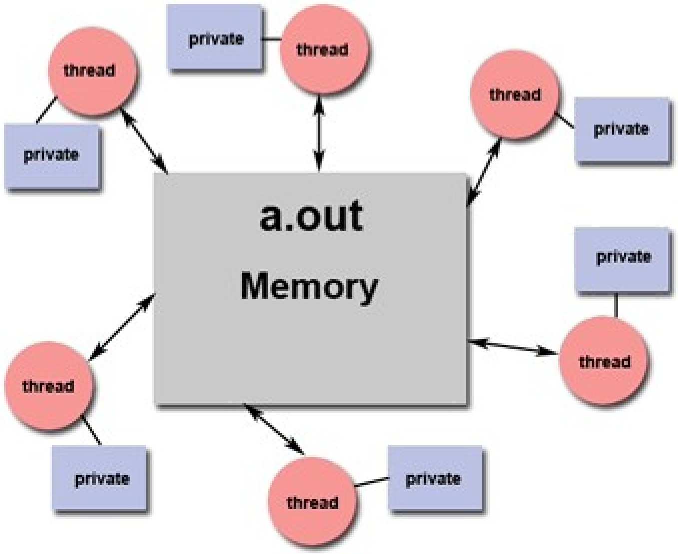


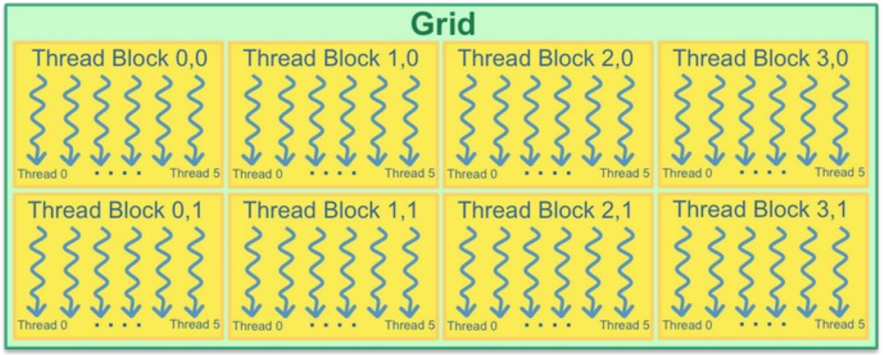


Control (Thread) Parallelism

- Focuses on distributing (different) tasks across different processors.
- Each task run independently
- Applied to processor architecture: multiple cores, GPU
- Applied to OS: processus, threads
- Applied to language: HPF, OpenMP
- Applied to programming model: Cuda







Designing Parallel Programs

- Partitions
- Communications
- Synchronizations
- Load Balancing
- Granularity
- I/O
- Debugging
- Performance analysis and tuning

See https://computing.llnl.gov/tutorials/parallel_comp/

→ Automatic parallelization tools and parallel programming frameworks can (in some situations) help !

Conclusion

- High Performance Computing (HPC) aims to run scientific simulations on supercomputers to address major scientific and engineering challenges.
- Data-intensive processing (Big Data) aims to analyze, extract knowledge and make decisions based on huge data corpora.
- HPC supercomputers are standalone machines with millions of computing cores aggregated through a dedicated network. They run highly-optimized parallel simulations, and performance is the major concern.
- Big Data analytics (data aggregation and mining) runs in the cloud, an elastic environment built on cost-effective commodity components. Data analytics relies on generic paradigms (map/reduce, neural networks) to process large static or streamed datasets, while trading performance for productivity and programming simplicity.