

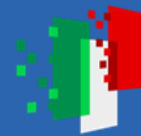


Introduction to High Performance and Data Intensive Computing: AN EXAMPLE

Dott. Costantino Zazza, costantino.zazza@unitus.it

IR0000032 – ITINERIS, Italian Integrated Environmental Research Infrastructures System
(D.D. n. 130/2022 - CUP B53C22002150006) Funded by EU - Next Generation EU PNRR-
Mission 4 "Education and Research" - Component 2: "From research to business" - Investment
3.1: "Fund for the realisation of an integrated system of research and innovation infrastructures"





HPC Technologies - DIBAF cluster

Agenda

► What about ?

HPC, Big Data, GFLOPS, PB, ...

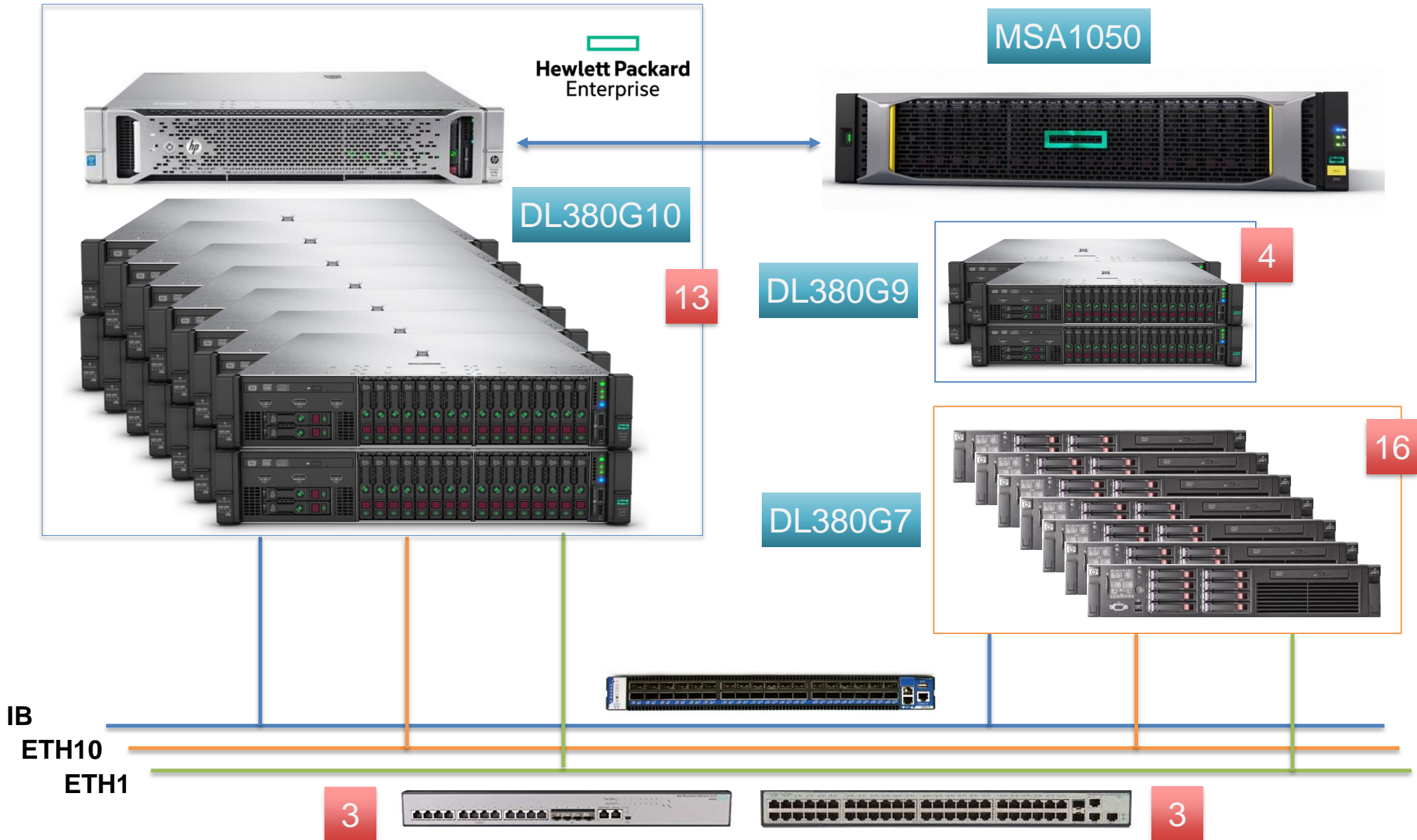
► How do they work ?

Calcolo numerico intensivo,
Data Processing, ...

► Useful for what ?

Chimica Computazionale, Bioinformatica,
IoT, Image Processing, Machine Learning, ...

DIBAF Cluster

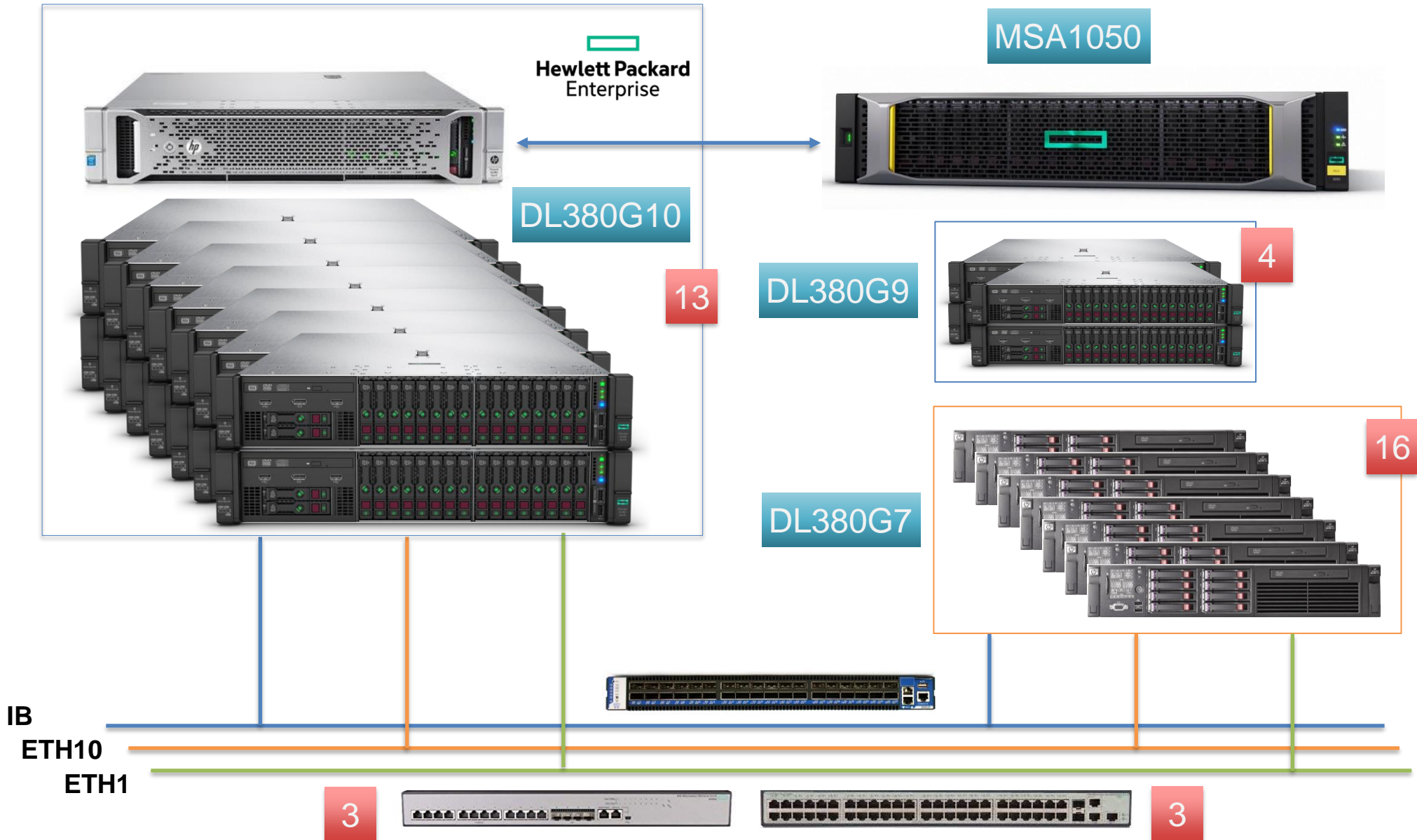




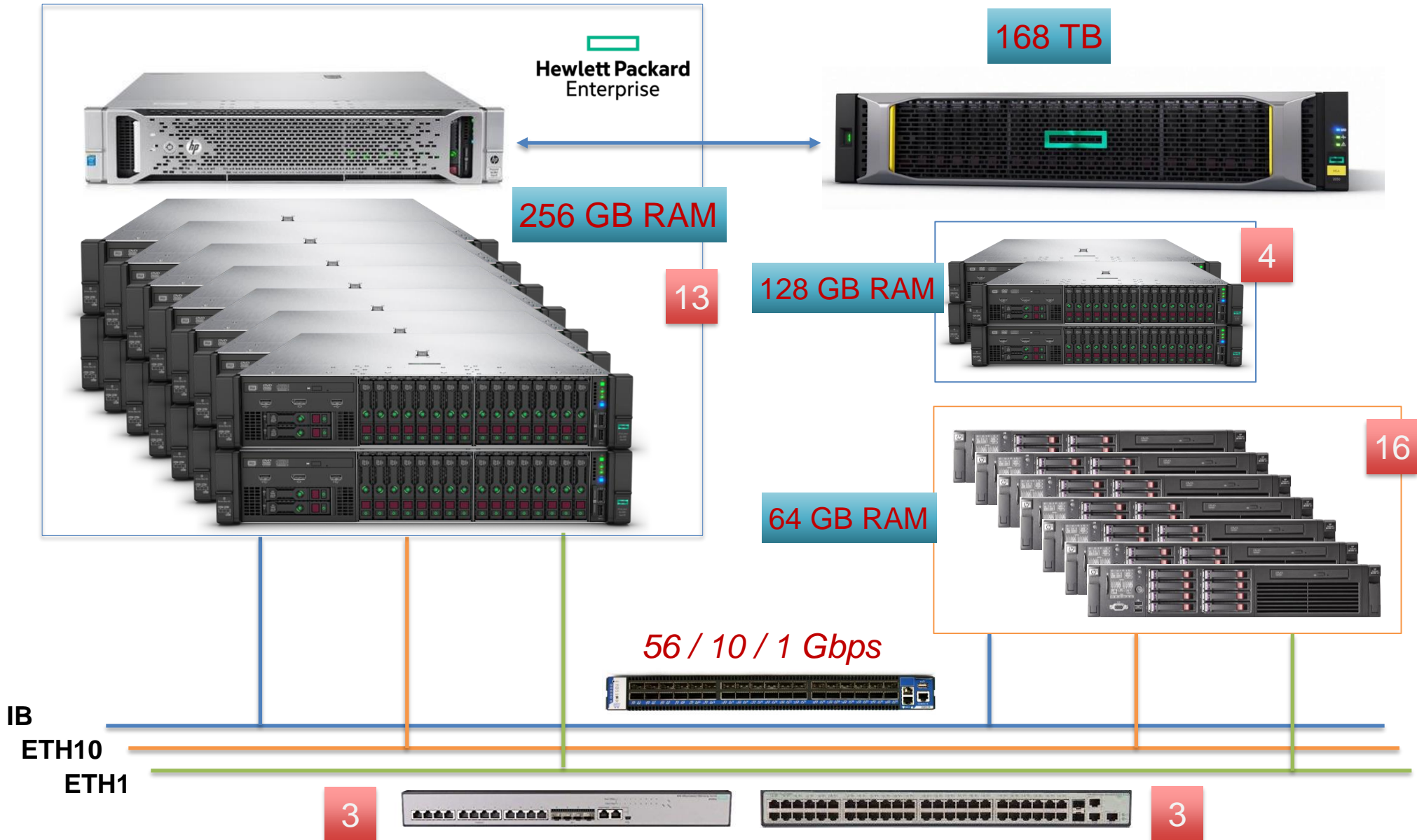
The DIBAF Cluster



DIBAF Cluster



DIBAF Cluster



DIBAF Cluster

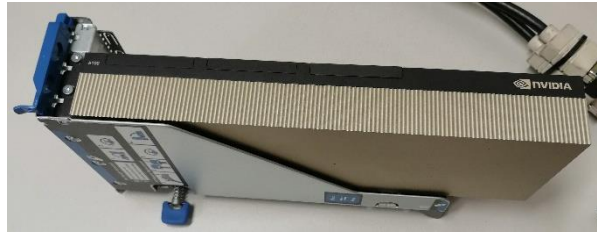


26



26

300 TFLOPS SP (HPC)



8

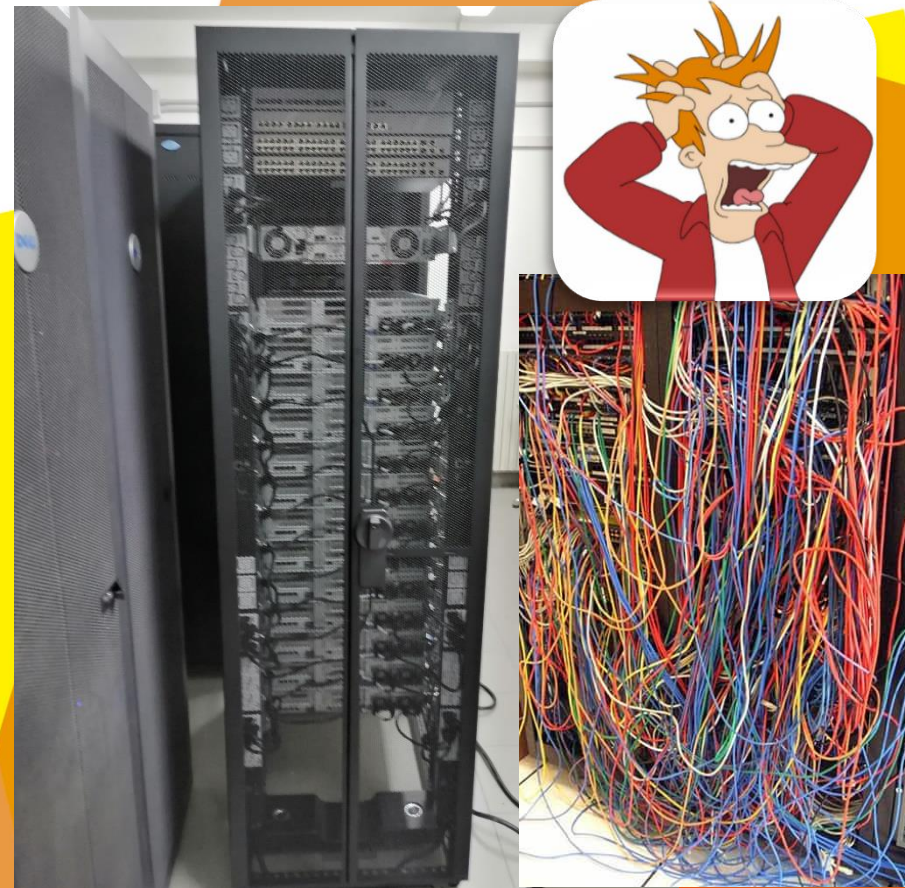
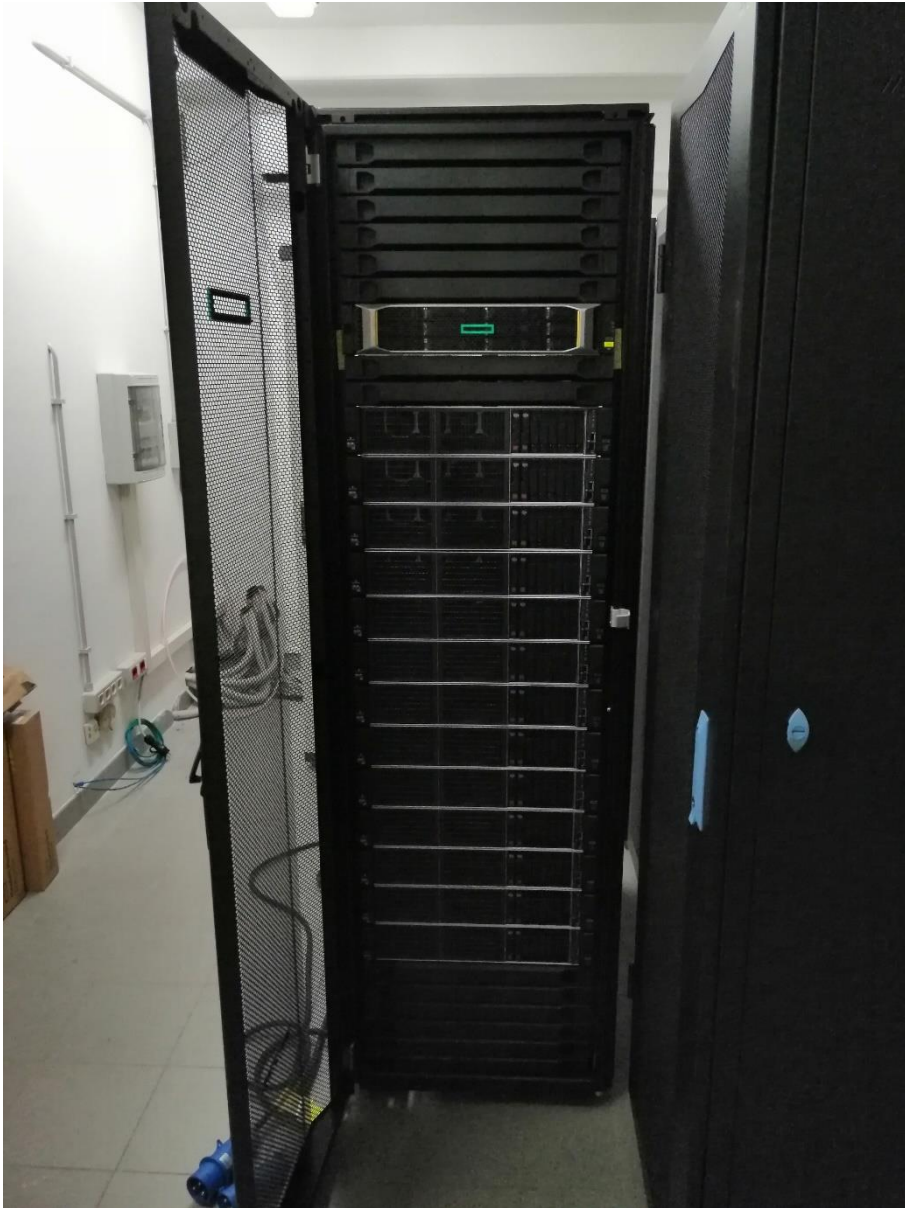


8

8 POPS INT4 (AI/ML)

1 PC = 0.01 TFLOPS

DIBAF Cluster



$26+8 = 34$ Ethernet cables
13 Infiniband cables
32 Power cables

DIBAF Cluster

Insight Cluster Management Utility v8.2.4 Enterprise

Using Insight CMU

- [Launch Insight Cluster Management Utility GUI](#)
- [Insight CMU User Guide](#)
- [Insight CMU Release Notes](#)
- [Insight CMU REST API Documentation](#)
- [Unable to launch Insight CMU GUI?](#)

Resources

- [Insight CMU Web Site](#) for up-to-date documentation, select Technical Support/Manuals.
- [Insight CMU public forum](#) for patch announcements and other discussions.

Cluster status

30 Nodes Total

- Up
- Unknown
- Down

Insight CMU v8.2.4 - cmu8 - root@cmu8

Resources

Filter

All Resources

192.168.100.150

Network Groups

Unassigned nodes

NARTEN

30 nodes (28 up, 2 down)

Instant View Bar Graph View Table View Time View Nodes Details Alerts

cpuload, max = 100.46 %

100.46 %

100.46 %

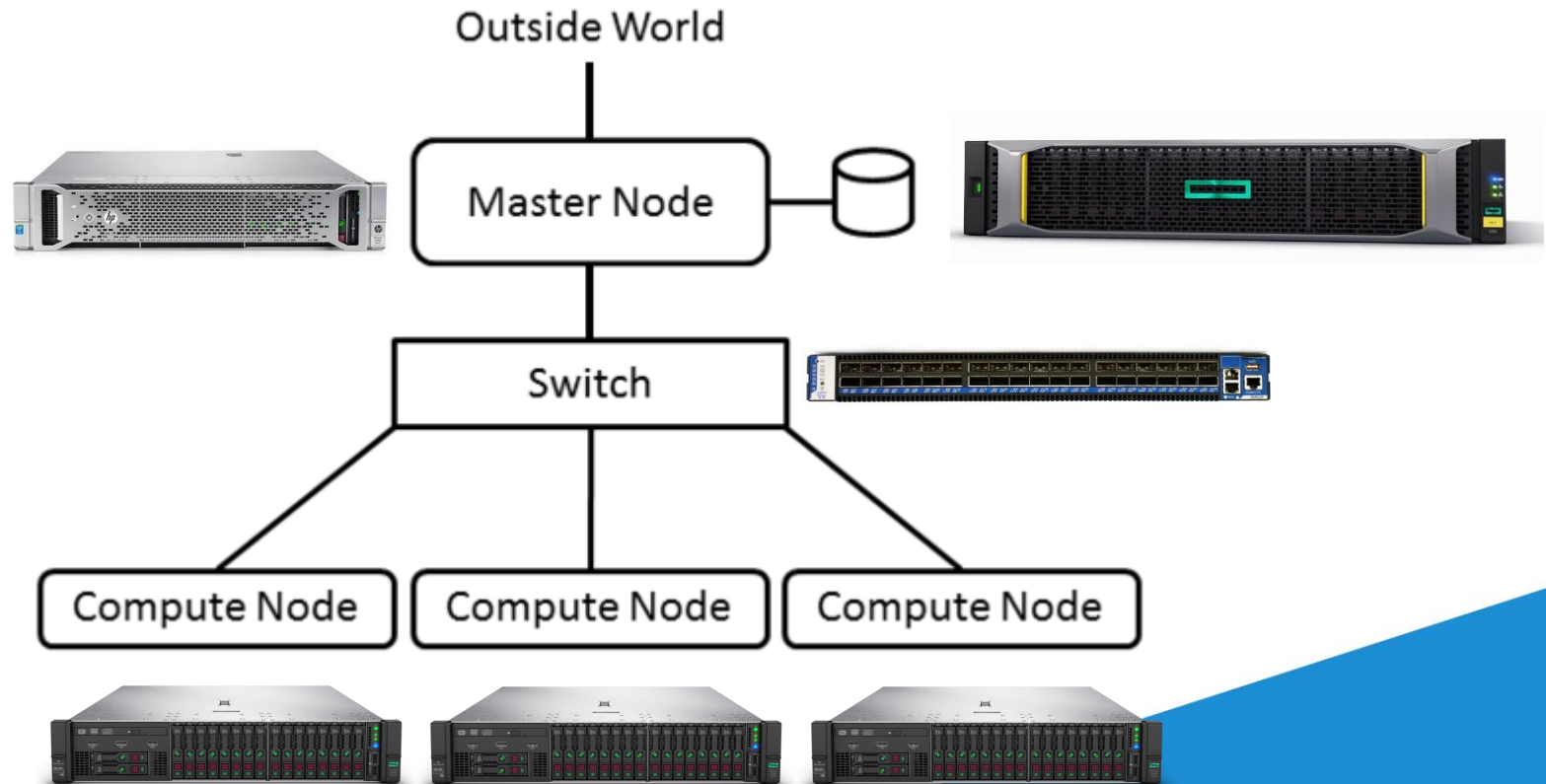
10.3 %

0 100.46

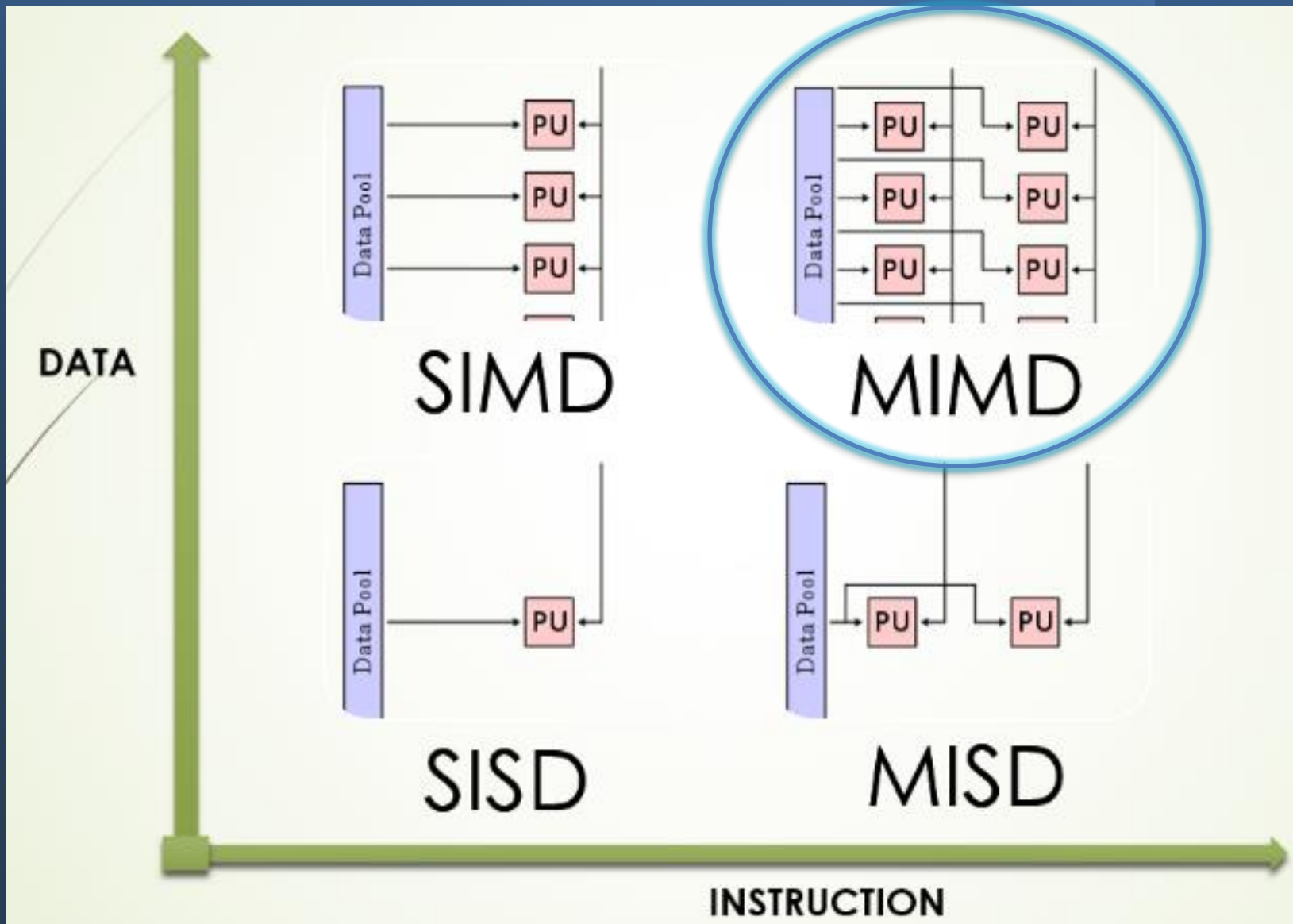
Information

Type	Date	Message
Custom menu configuration	Fri May 28 09:50:59 CEST 2021	Custom menu configuration updated
Remote connection	Fri May 28 09:50:58 CEST 2021	CMU server connection initialized

HPC – High Performance Computing



HPC Classification – Flynn Taxonomy

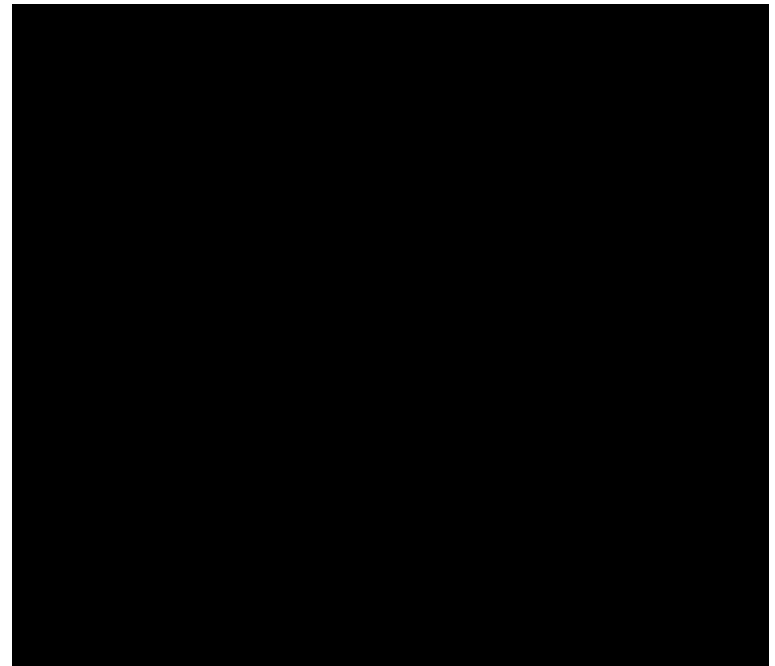
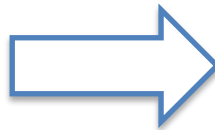
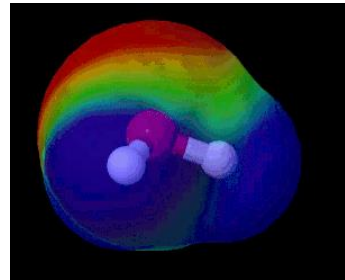


HPC – High Performance Computing

Brute force - Expected performance when scaling

Quantum Chemistry

	PC	Cluster
$O(N^3) / O(N^4)$	3-5 atomi(H ₂ O, NH ₃ , CH ₄)	> 50 atomi



HPC – High Performance Computing

Brute force - Expected performance when scaling

Molecular Dynamics

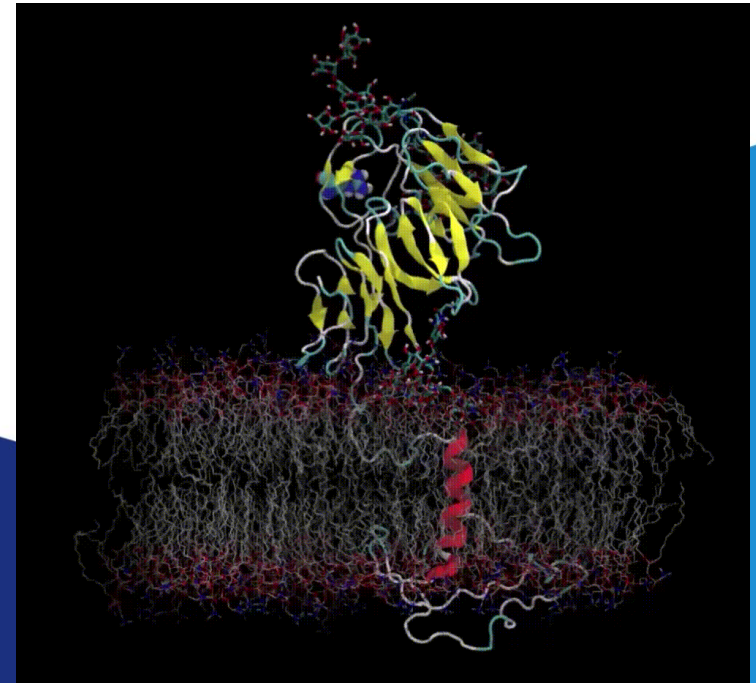
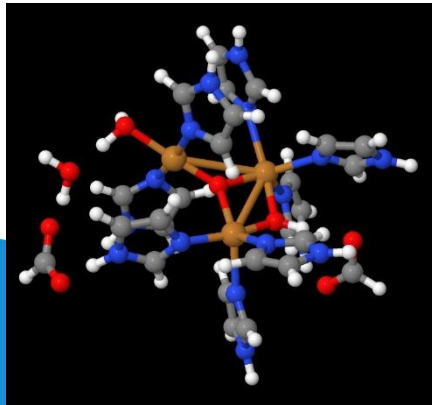
PC

Cluster

$O(N^2)$

20-50 atomi (peptide)

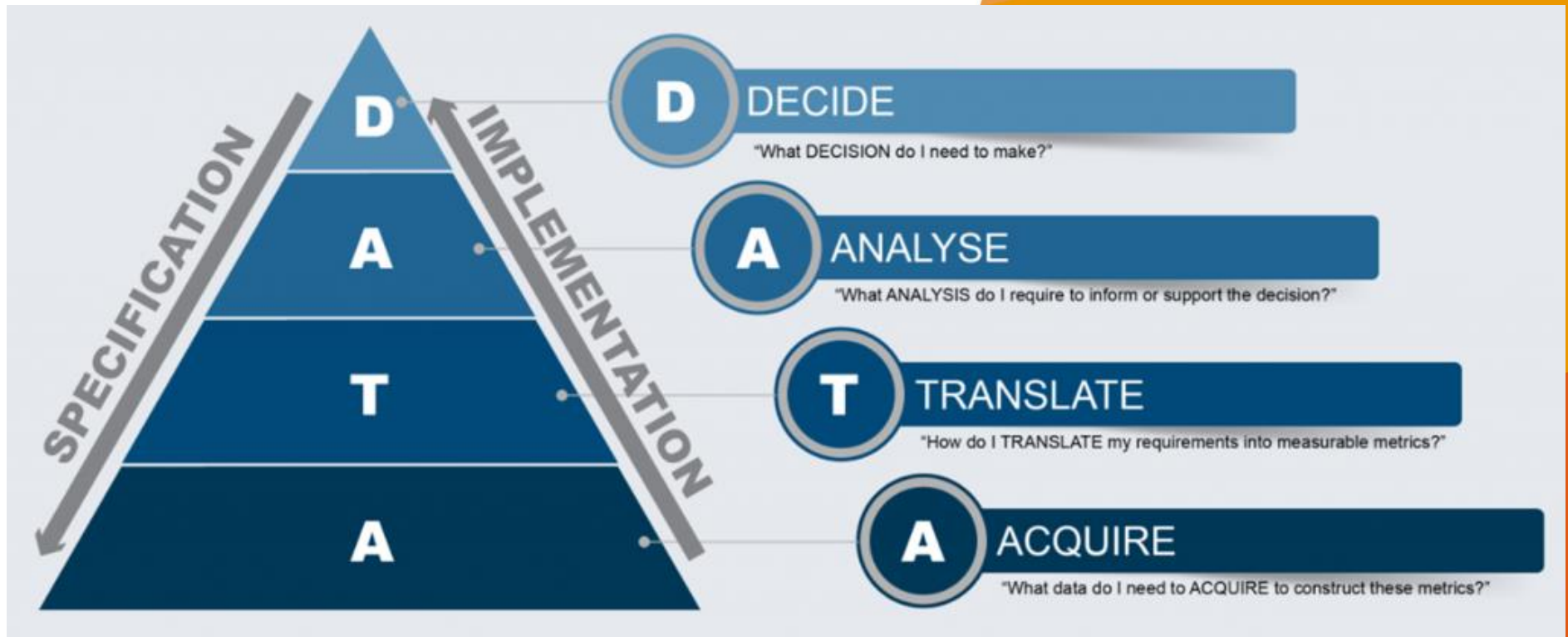
> 1 milione di atomi



Big Data & Analytics

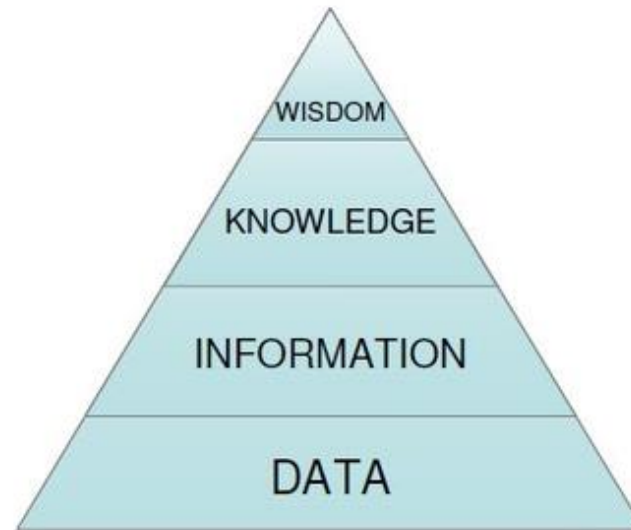
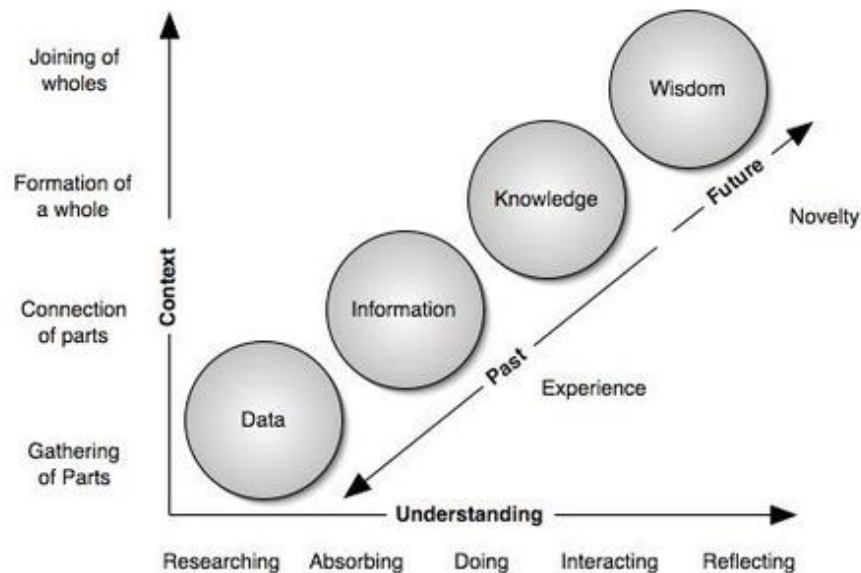
Data Driven model: make decision from data

Data-driven decision-making (**DDDM**) is an approach that emphasizes using data and analysis instead of intuition to inform “business” decisions.



Big Data & Analytics

DIKW «hierarchical» model



The cognitive pyramid

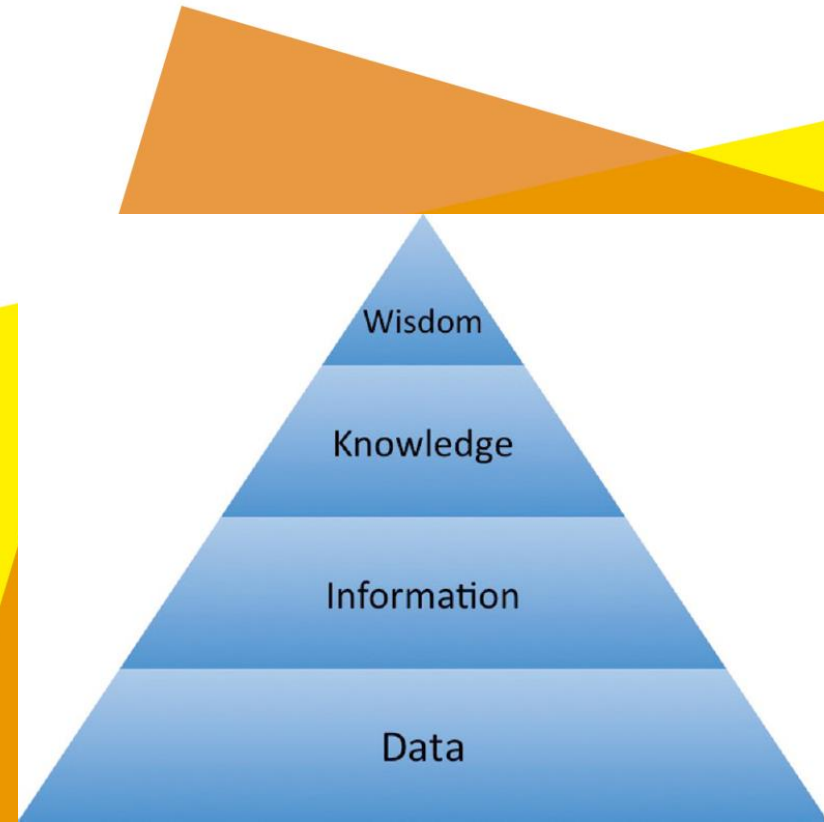
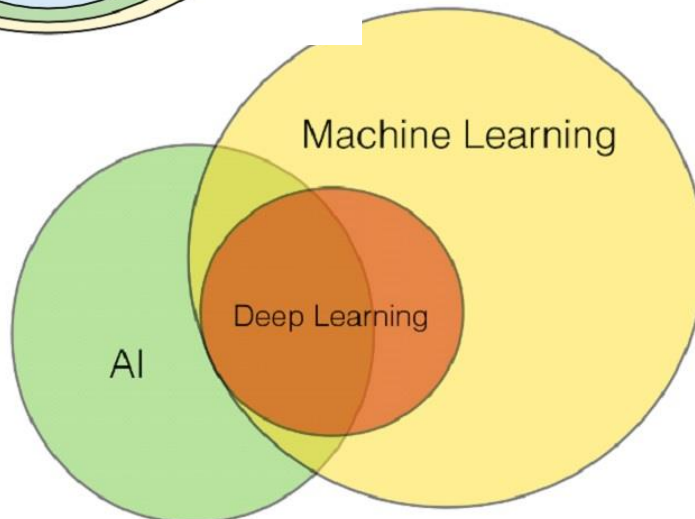
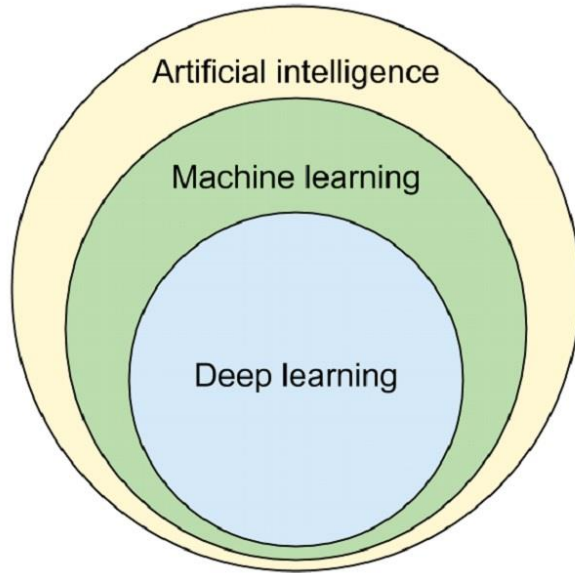
Pay attention on DATA QUALITY
at the base of the pyramid !

Reference(URL):

https://en.wikipedia.org/wiki/DIKW_pyramid

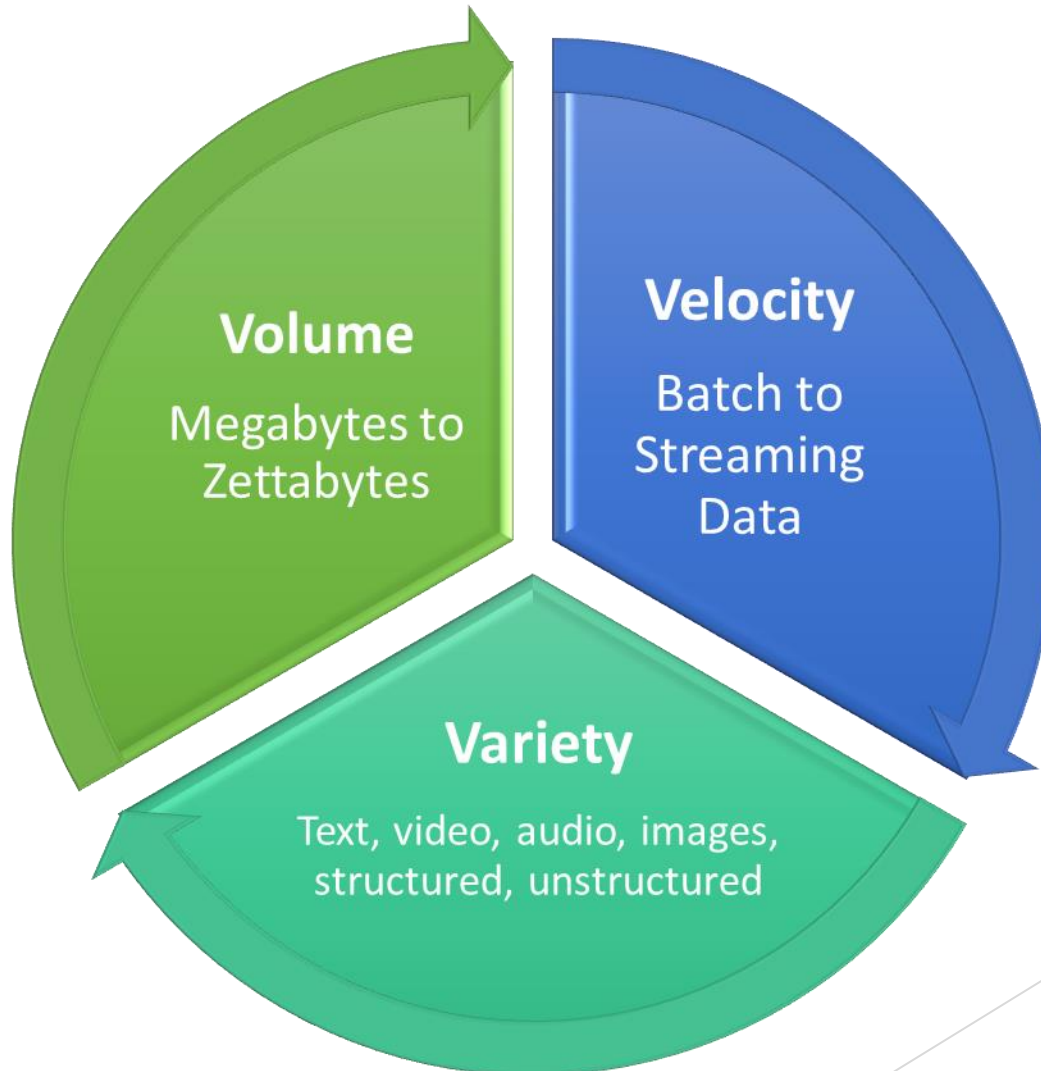
Big Data & Analytics

Theoretical model: AI, ML, DL



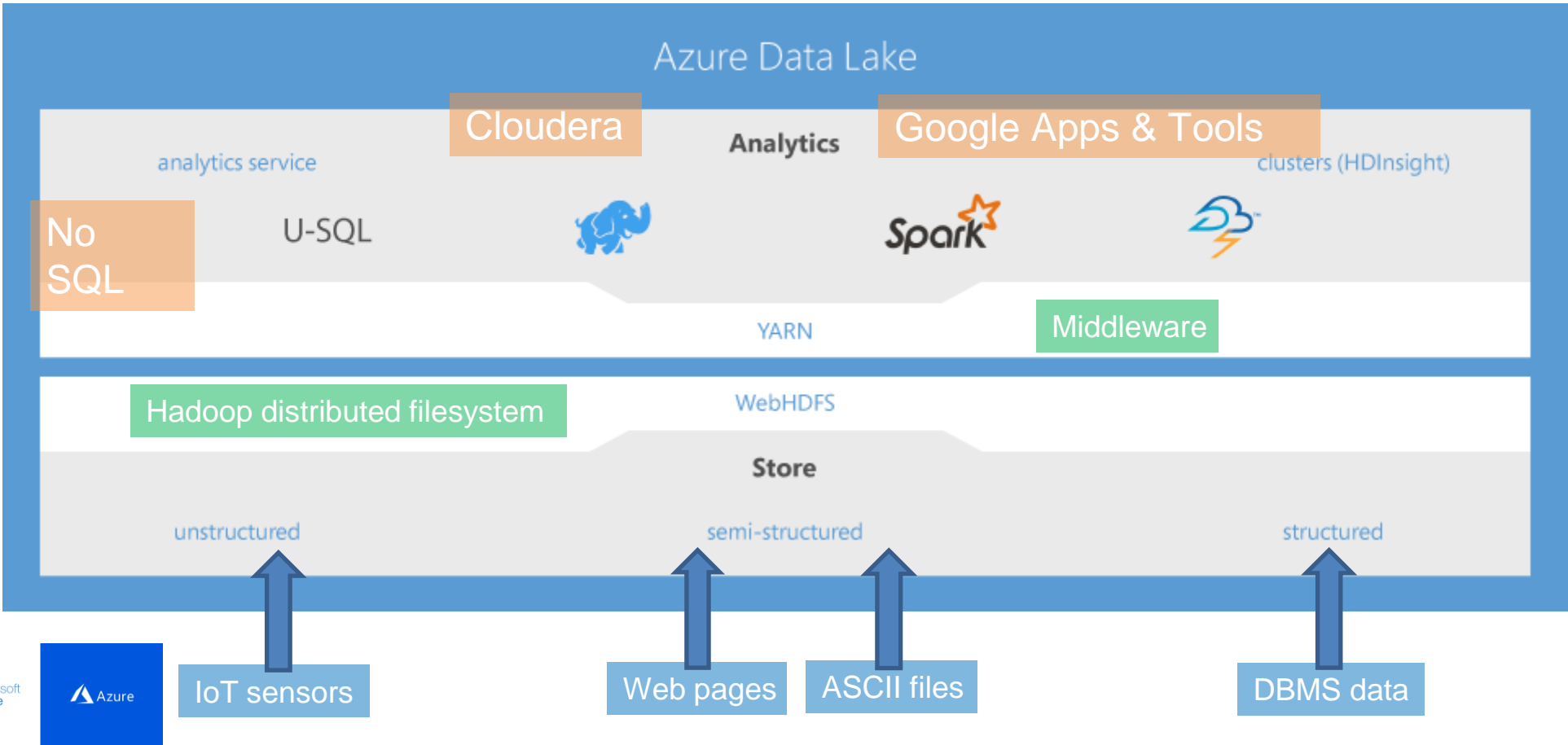
Big Data & Analytics

From Data to Big Data



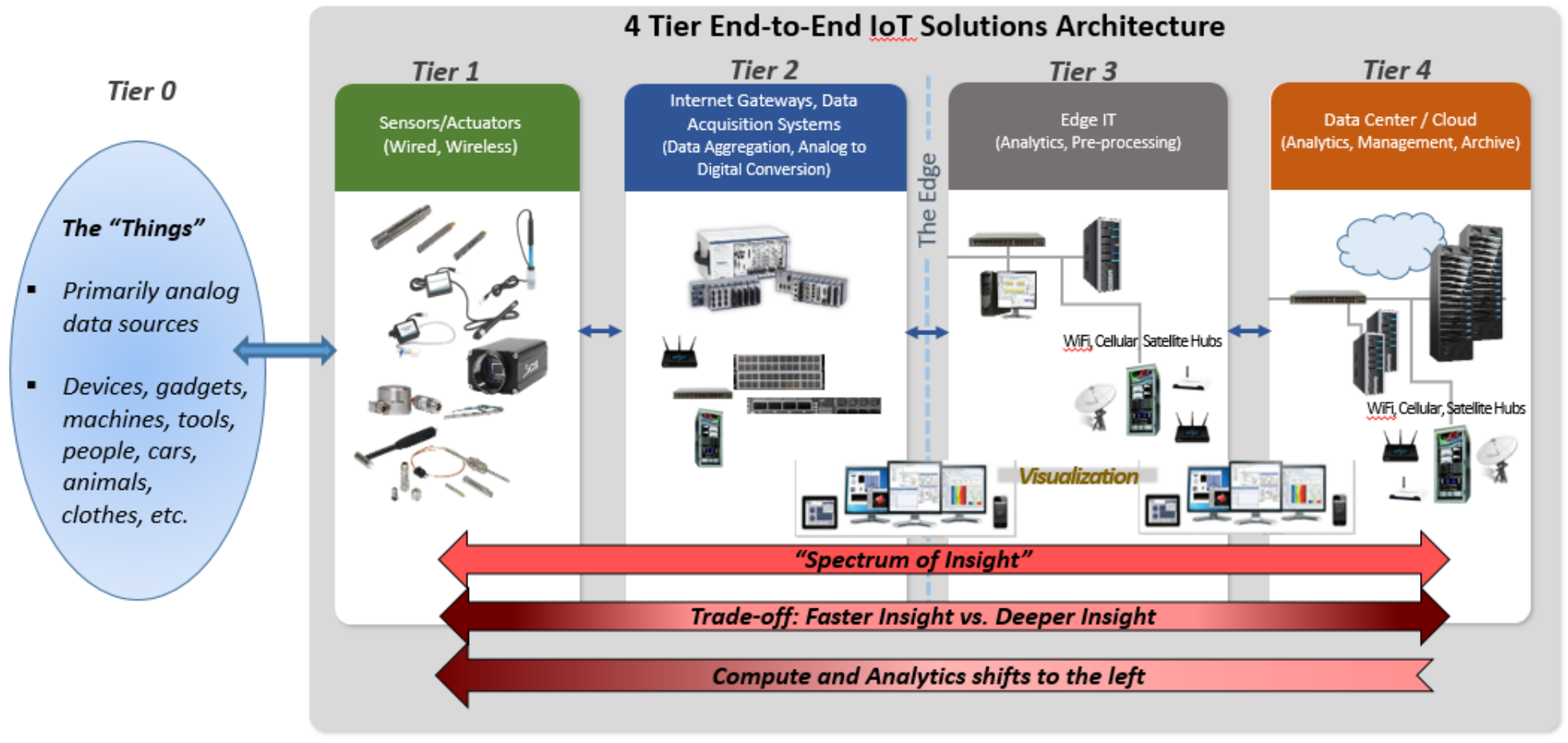
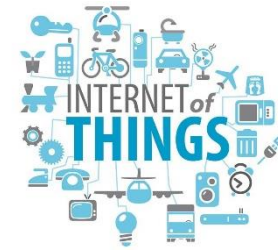
Big Data & Analytics

DIBAF BDA + AI model



Big Data & Analytics

DIBAF BDA + AI model

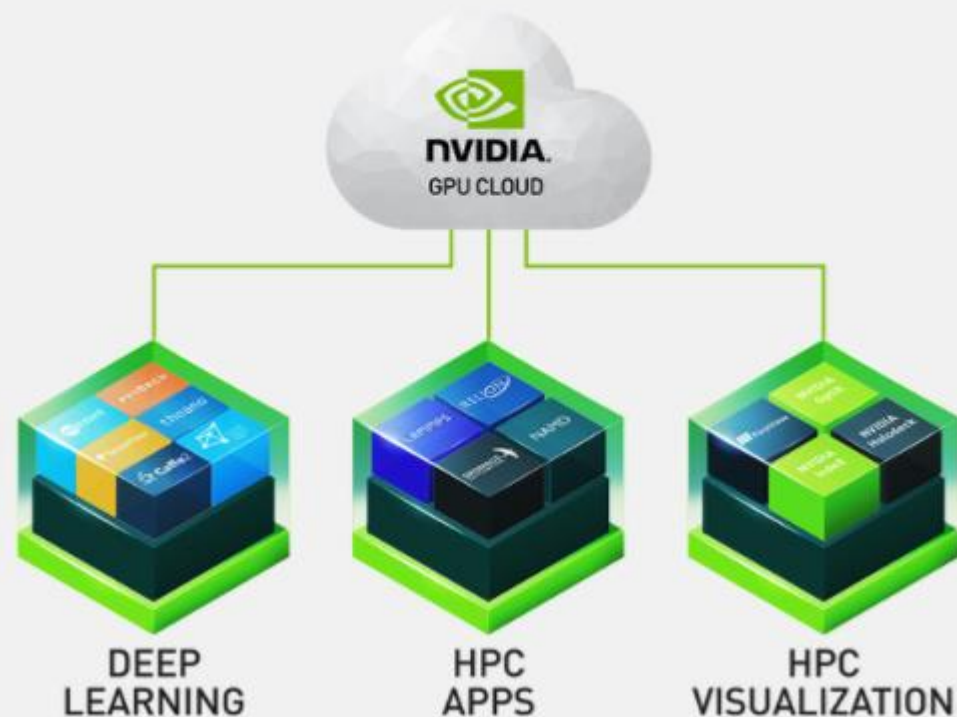


Big Data & Analytics

DIBAF BDA + AI model

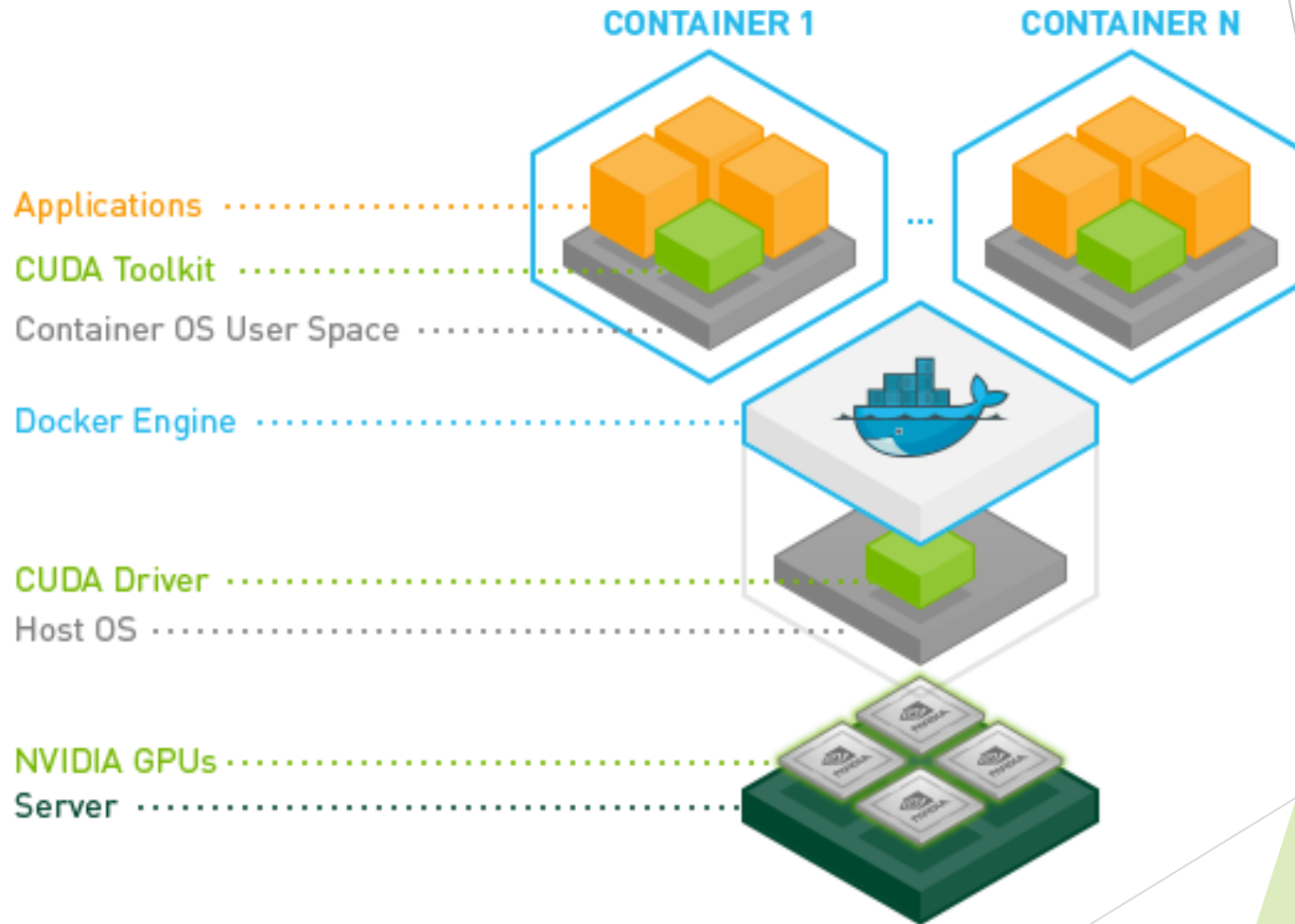


CONTAINER PER DEEP LEARNING ED
ELABORAZIONE AD ALTE PRESTAZIONI



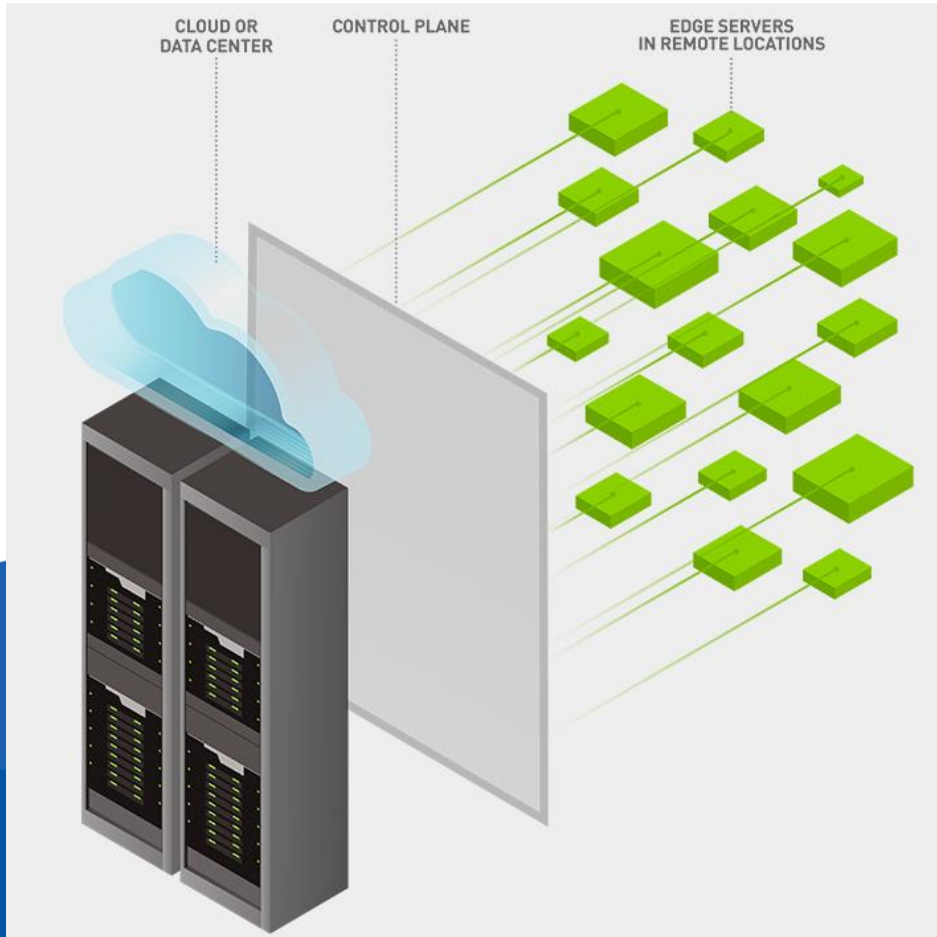
Big Data & Analytics

DIBAF BDA + AI model

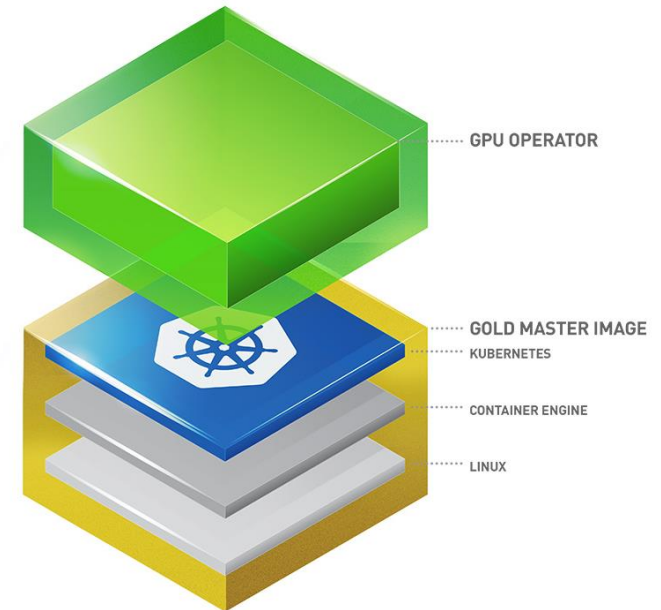


Big Data & Analytics

NVIDIA EGX platform for EDGE computing



Based on Kubernetes
From Datacenter to the EDGE



Big Data & Analytics

AI + IoT + 5G



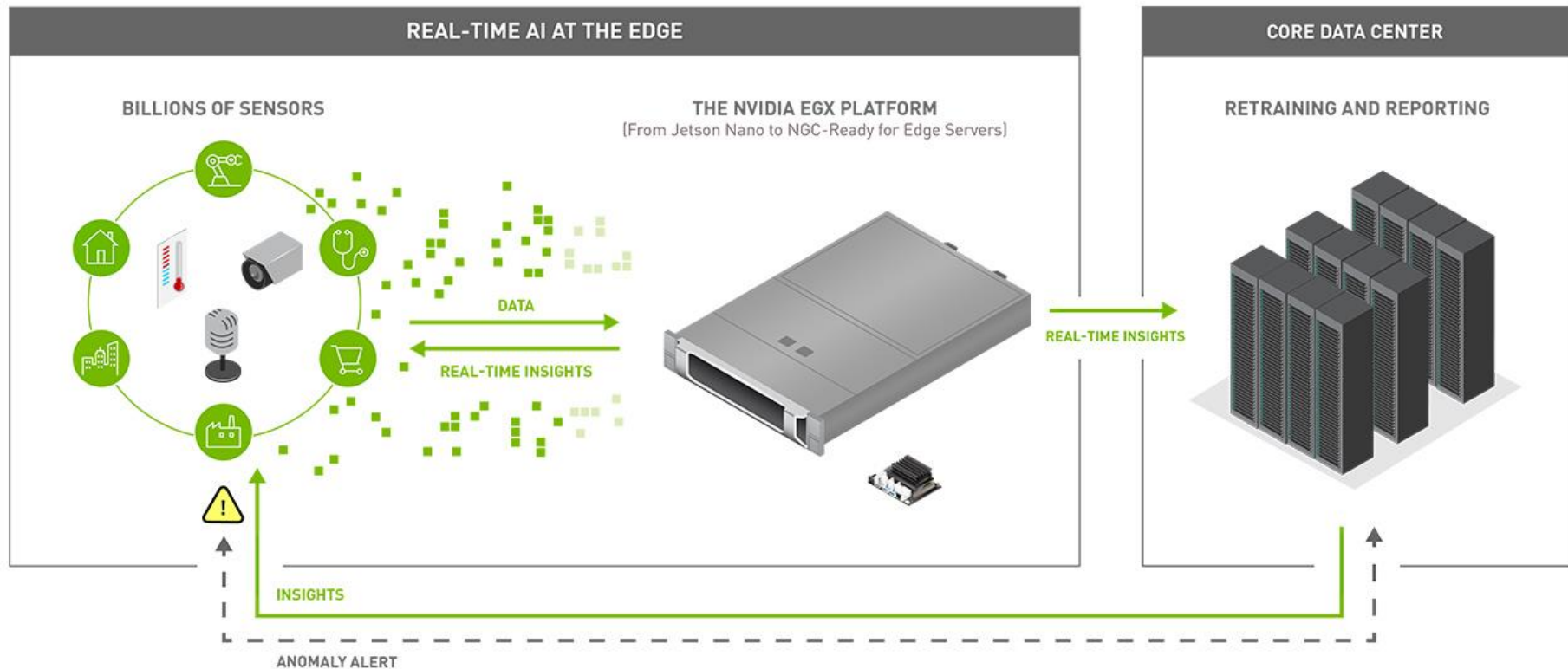
Ready for field test on Jetson Nano

100\$



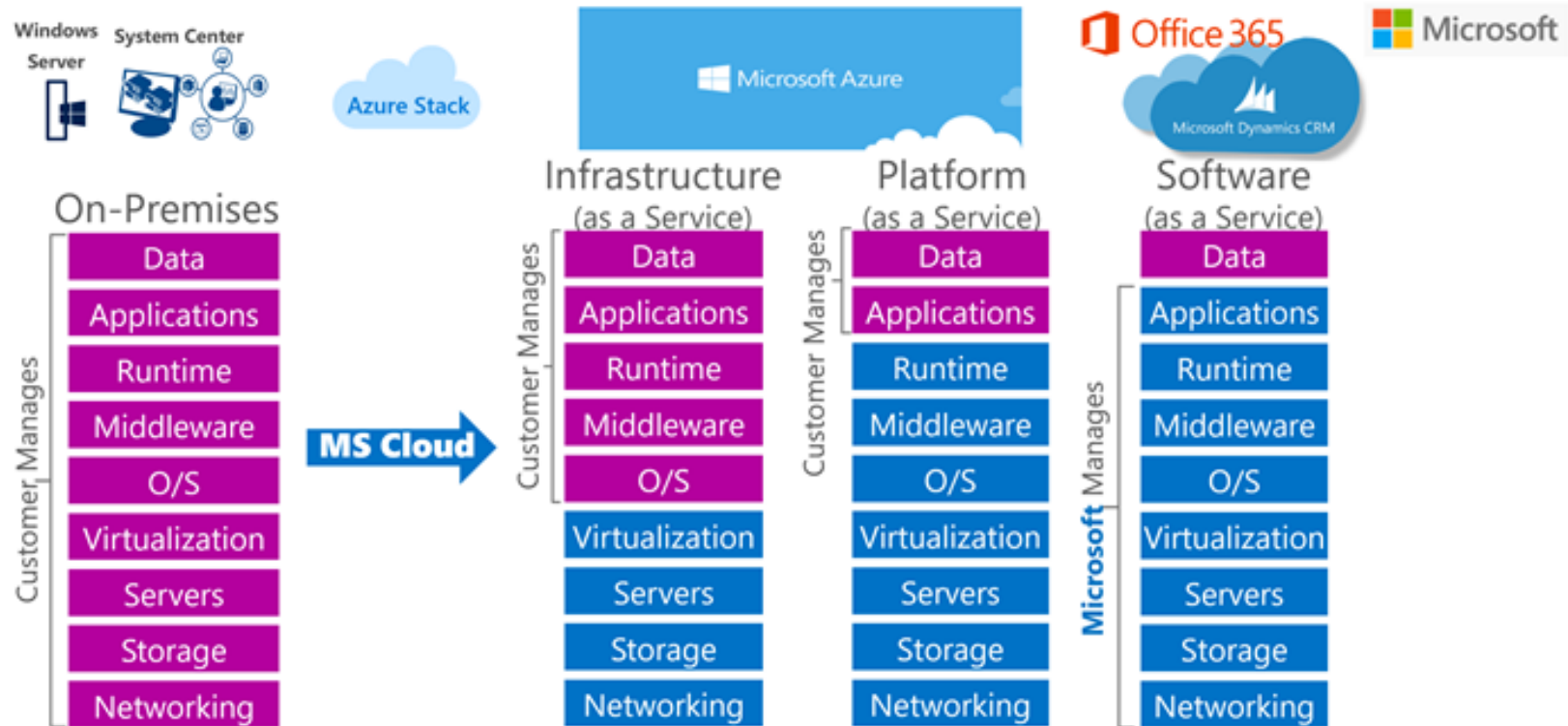
Big Data & Analytics

NVIDIA EGX platform for EDGE computing



Big Data & Analytics

DIBAF BDA + AI model



Big Data & Analytics

Supervised Machine Learning

Setting up our classifier

Field

Data in standard format					
case	Feature 1	Feature 2	...	Feature n	Class
1	xxx	x		xx	good
2	xxx	x		xx	good
3	xxx	x		xx	bad
...					...

Weather-Climate

Data in standard format					
case	Feature 1	Feature 2	...	Feature n	Class
1	xxx	x		xx	good
2	xxx	x		xx	good
3	xxx	x		xx	bad
...					...

Bio-NGS/Genome

Data in standard format					
case	Feature 1	Feature 2	...	Feature n	Class
1	xxx	x		xx	good
2	xxx	x		xx	good
3	xxx	x		xx	bad
...					...

Big Data & Analytics

Supervised Machine Learning

<i>Classifier</i>								
Feature	Feature	Feature	Feature	Feature	Feature	Feature	Feature	CLASS
A1	A2	A3	B1	B2	C1	C2	C3	
x	y	x	a	b	α	β	δ	GOOD
y	x	z	c	d	γ	π	α	BAD
...
...
y	x	x	a	f	γ	γ	δ	GOOD

