PEPR NumPEx

PC3: Exa-DoST

Data-oriented Software and Tools

Thomas Moreau

NumPEx - Software for the Exascale computer

HPC: High Performance Computer

Exascale machine: throughput of 1 Exaflops

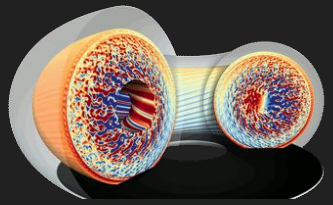
1 billion of billion of floating point operation/sec

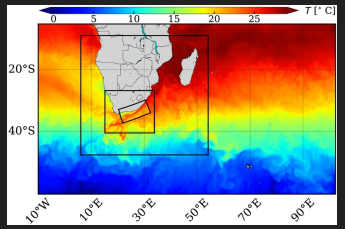
Delivered in 2026 in France

NumPEx: build software to run on the machine

Applications:

Complex system simulations





Analysis of large data-stream

ExaDoST - Data at exascale: a challenge in hardware

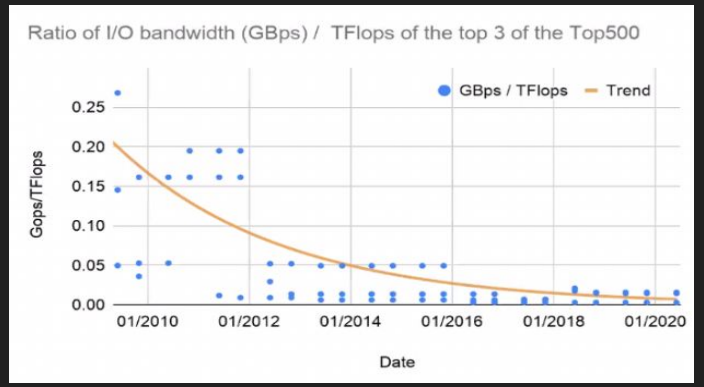
Increasing **gap between compute and I/O**

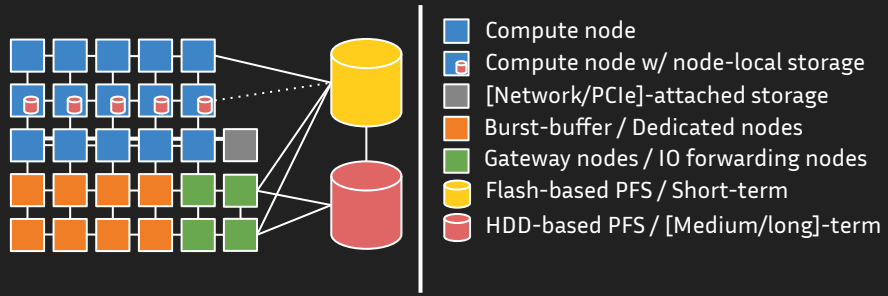
Ratio of I/O to computing power divided by ~10 over the last 10 years on the top 3 supercomputers

=> Data deluge!

New hardware try to mitigate this but this does not go fast enough.

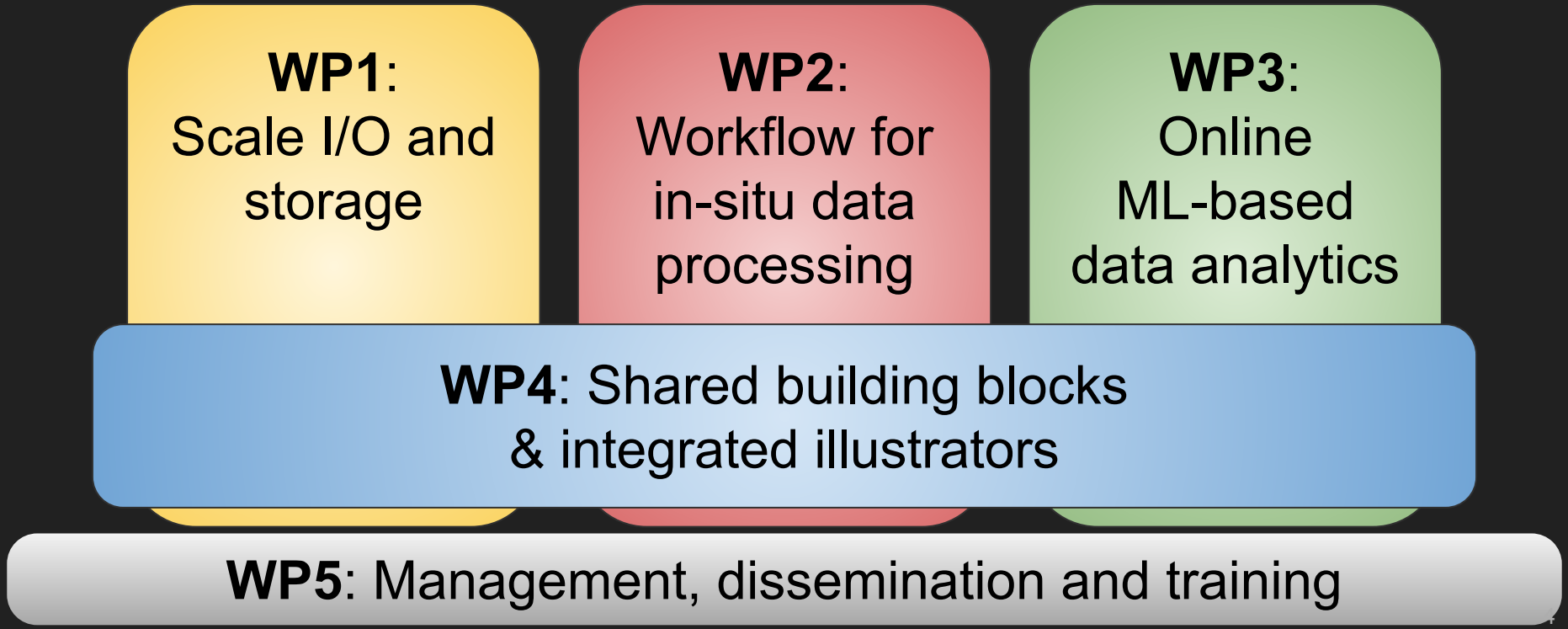
ExaDoST: software approach to the I/O issue





Trend in storage technologies available on extreme-scale systems

Work Packages in ExaDoST



WP3: ML-based data analytics

WP-leaders

B. Raffin
Inria
T. Moreau
Inria

Main objective: reduce need for storage by extracting information online

=> Why MIND?

- **Statistical analysis of large signals:** processing fMRI requires expertise in large signal processing.

- **joblib/scikit-learn:** 1 engineer to work on scaling the packages so they can be used on the exascale machine.

In collaboration with **Statify**, **Datamove** and **Thot** (3 Inria teams in Grenoble).