

Contributions prévues au sein du PEPR NumPEX

13 juin 2024

Jussieu, Paris

Samuel Thibault



Inria

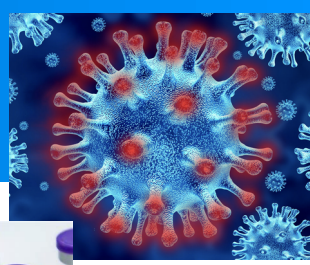
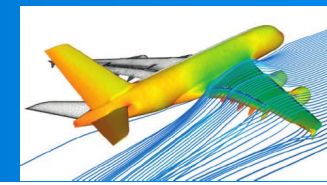
anr®



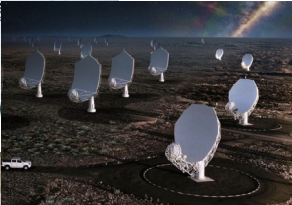
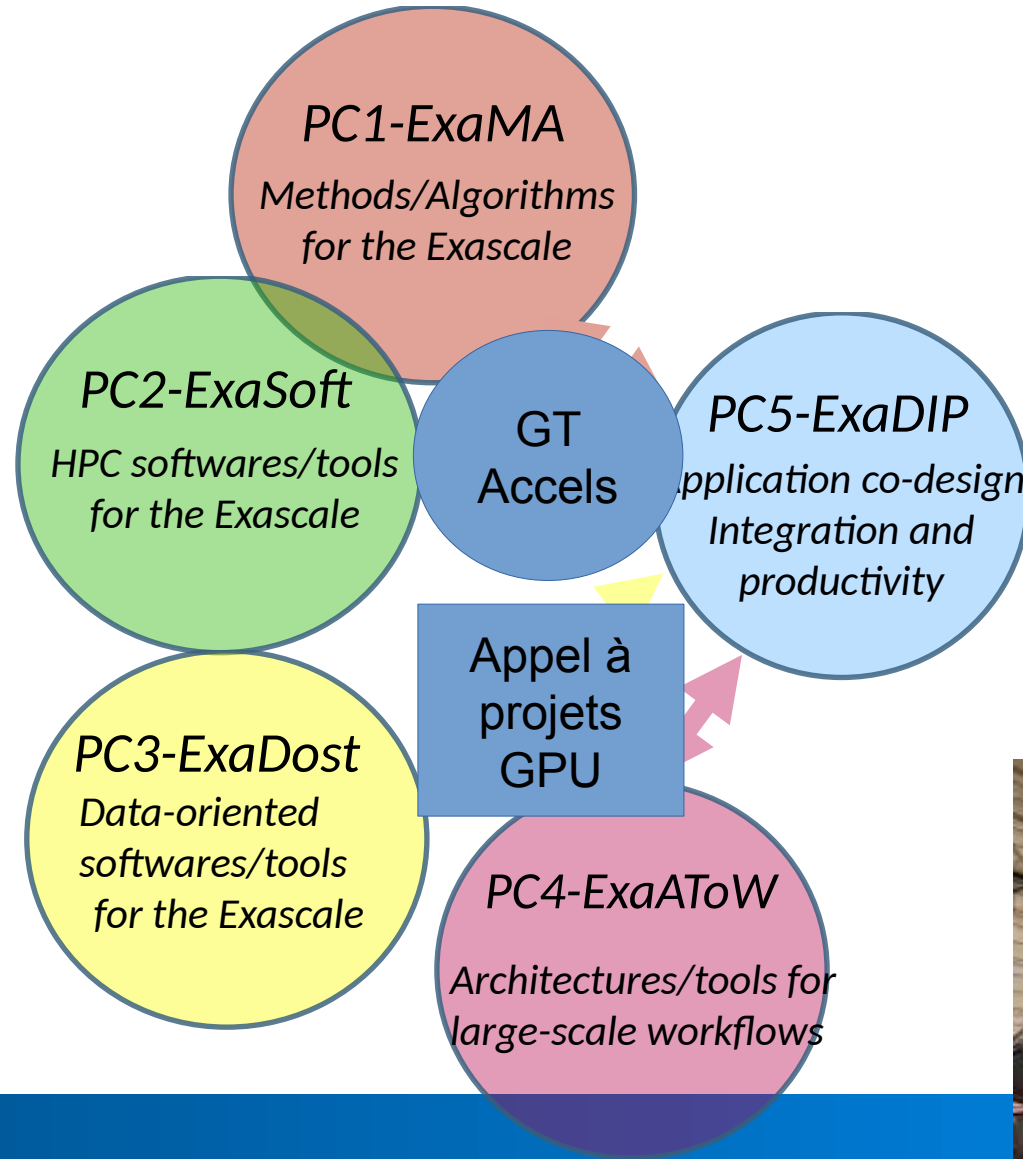
Machines

- Jean Zay (GENCI)
 - 37 PFlop/s
 - 3000 GPUs
- Jupiter (Jülich, Allemagne)
 - > 1EFlop/s
 - 24,000 GH200 CPU+GPUs
- Jules Verne (TGCC)
 - > 1EFlop/s





- Briques logicielles pour l'ExaScale
- 5 ans depuis 2023



Applicative PCs: GPU support

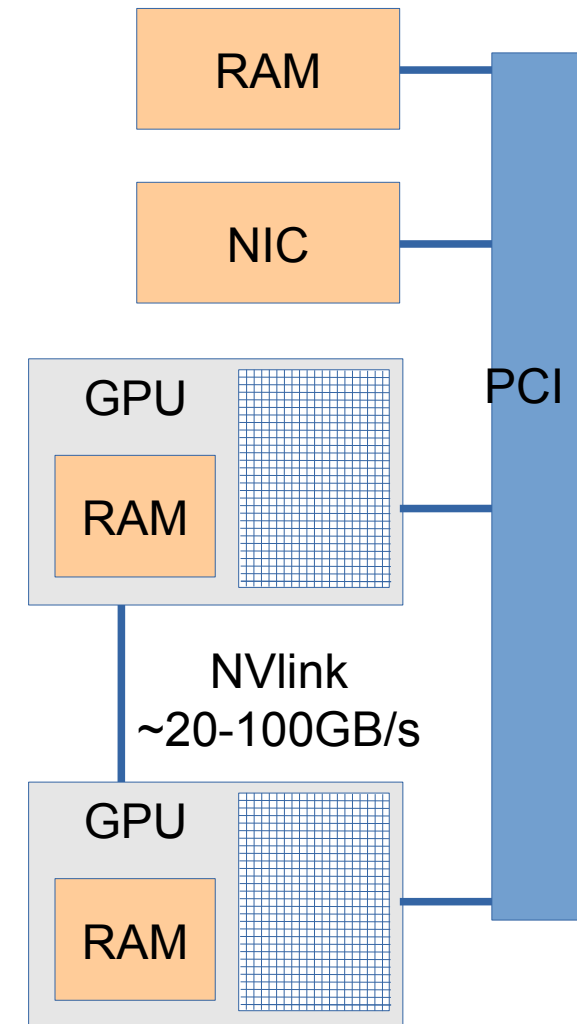
- PC1 Exa-MA Methods - Algorithms
- PC3 Exa-Dost analysis, statistical, ML, DL libraries
- PC4 Exa-AToW Architectures/tools for large-scale workflows
- PC2 Exa-SofT GPU-ready HPC Software stack
- PC5 Exa-DI Integration and productivity

PC2 Exa-Soft GPU-ready HPC Software stack

- WP1: composition model for GPUs, GPU kernel development
- WP2: generating CPU/GPU versions for tasks
- WP3: Runtime optimizations for GPUs
- WP4: Task-based numerical building blocks
- WP5: Energy profiling and scheduling

WP3: Runtime optimizations for GPUs

- GPU - NIC optimization
 - GPU-Direct
 - Send from GPU
 - Receive to GPU
 - Allocation on GPU
 - *On time*
 - Thanks to knowledge of future from tasks

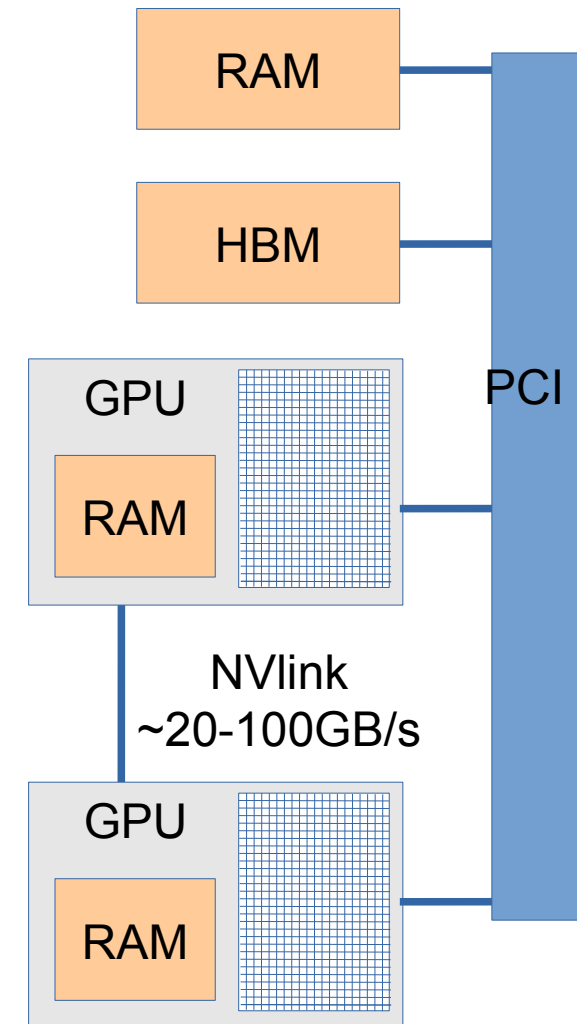


WP3: Runtime optimizations for GPUs

- Automatic partitioning for CPU/GPU distribution
 - Heterogeneous tile size
 - Need to be large for GPUs, can be small for CPUs
 - Heterogeneous acceleration
 - GEMMs-like on GPUs, Factorization-like on CPUs
 - Critical path acceleration
 - When required

WP3: Runtime optimizations for GPUs

- Data placement for CPU/GPU distribution
 - RAM / HBM choice
 - GPU choice
 - Copying data vs remote access



Conclusion

- Applicative component porting
- Software stack optimization
- Software Integration