



# SIPEARL

SiPearl on HPC  
March 2022  
TREX Workshop

# SIPEARL CORPORATE OVERVIEW

## The European Server Processor Solution



- HQ: Maisons-Laffitte (Paris), France
- CEO and Founder, Philippe Notton
- Design centers:
  - France Maisons-Laffitte, Massy Palaiseau, Sophia Antipolis, Grenoble
  - Germany Duisburg (Düsseldorf)
  - Spain Barcelona
- Seed Money: 8.5M€ via Horizon 2020
- Key Personnel from ST, Intel, Atos, Marvell, Mstar-Mediatek, Nokia
- Architecture based on Arm Neoverse (V1 cores)



*Founded in 2019 as the production hand of the European Processor Initiative (EPI), SiPearl holds a central position in the EPI for Exascale Processor Development and EuroHPC Pilots.*





# SIPEARL OFFICES

Maisons-Laffitte



Massy-Palaiseau



Barcelona



Reims



Duisburg



Grenoble



Sophia Antipolis





FROM IP TO PRODUCTS  
FROM EPI TO SIPEARL

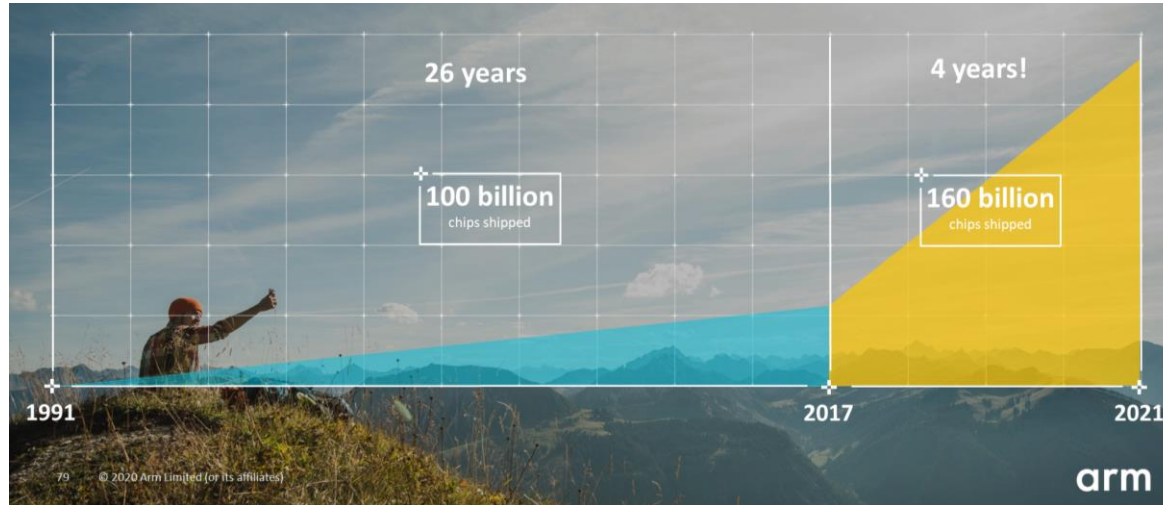
## European Processor Initiative

- High Performance General Purpose Processor for HPC
- High-performance RISC-V based accelerator
- Computing platform for edge and autonomous cars
- Will also target the AI, Big Data and other markets in order to be economically sustainable

### EPI Objective

Develop a complete EU designed high-end microprocessor, addressing Supercomputing and edge-HPC segments

# SIPEARL KEY PARTNERSHIPS



arm

- Most widely used ISA on the planet
- Low power high performance alternative to x86-64
- Arm design and modelling tools → time to market
- Arm toolchain and libraries
- Fully fleshed ecosystem is growing organically across HPC, Cloud; Edge, IOT, Automotive...

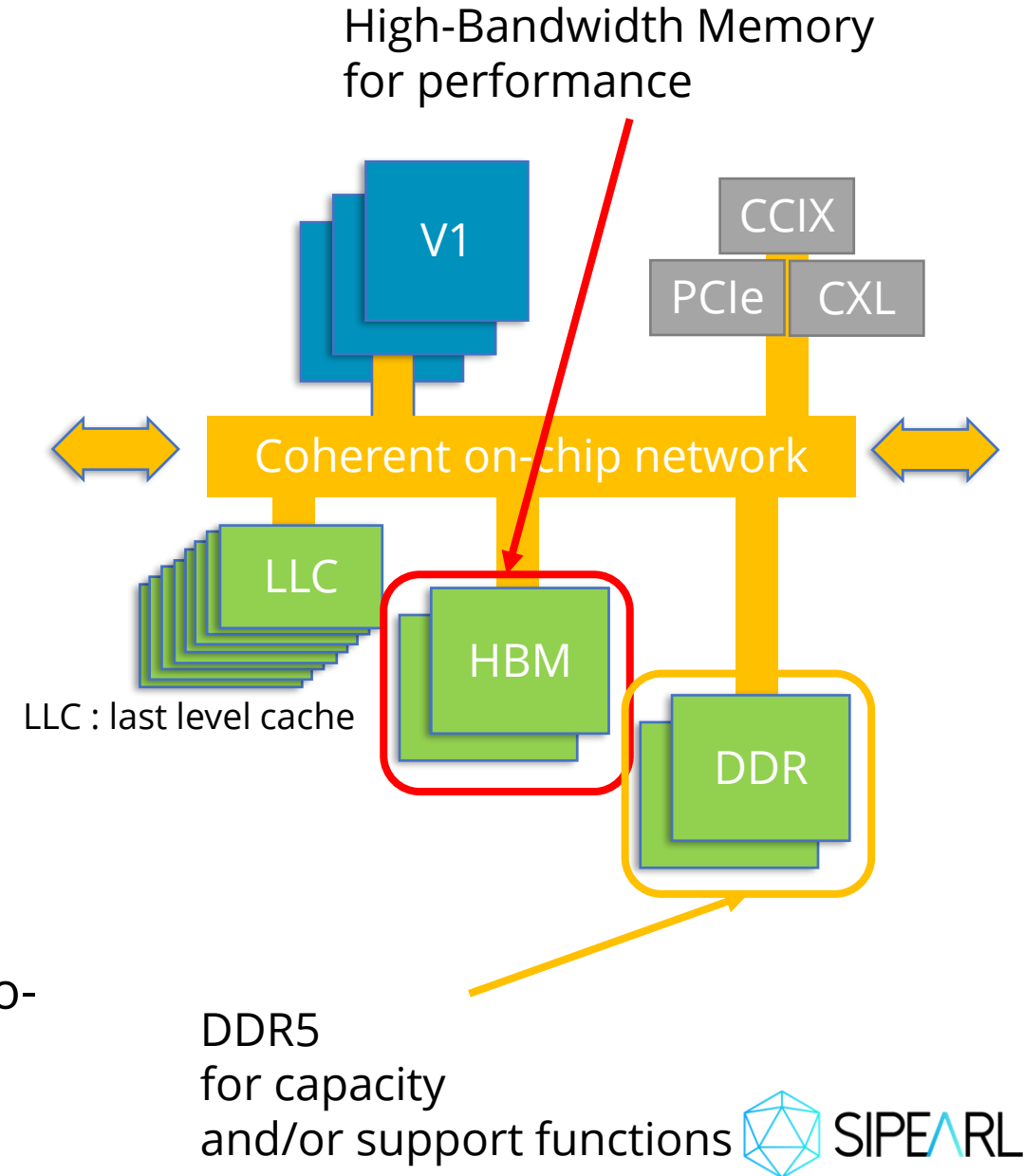


- Consortium funded by EU Government to foster a Sovereign EU based Server Processor & Ecosystem
- 27 funded partners across HPC Labs, System Integrators and OEMs, HW and SW companies



# SIPEARL RHEA HYPERSCALE CONFIGURATION

- V1 CPU core as compute unit building blocks
  - Performance computing and intelligent memory subsystem
  - General-purpose processor with rich software ecosystem
- Memory-coherent on-chip network
  - Topology-aware design for scalability and flexibility
  - Distributed last-level cache to memories with reconfigurable NUMA domains
  - Isolation between computing units
  - Coherent SMP between chip domains
- HBM and DDR for memory bandwidth + capacity
- Latest PCIe/CXL/CCIX links for interconnect and accelerators
- Low-power - low latency links for die-to-die or chip-to-chip connections

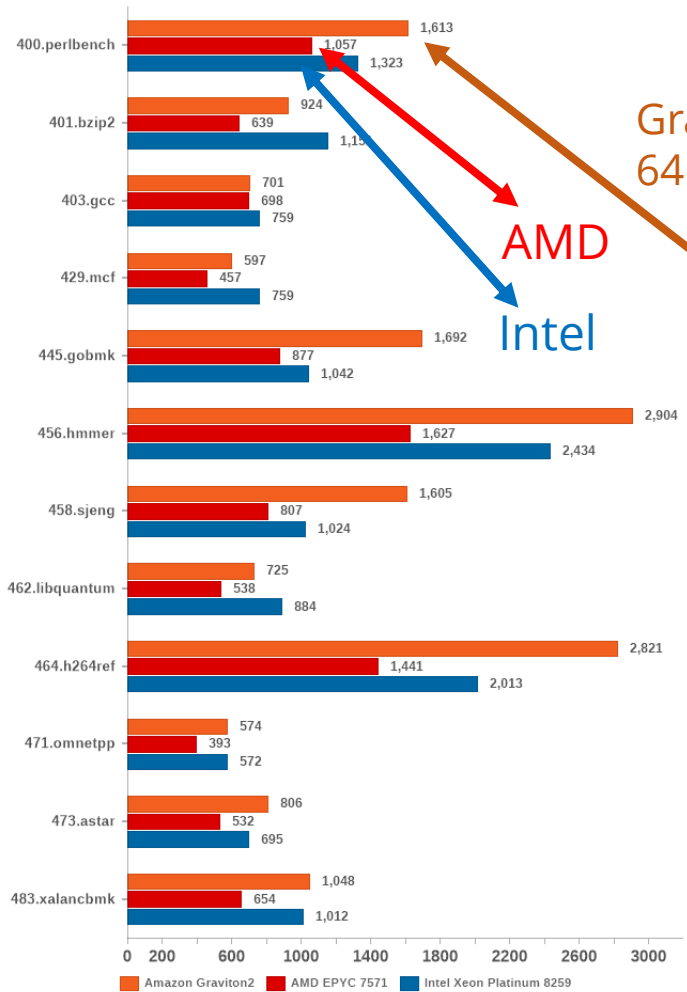




# CPU CORE PERFORMANCE

 **SPECint2006 Rate Estimated Scores (64 vCPU)**  
Score - Higher is Better

Source: Anandtech

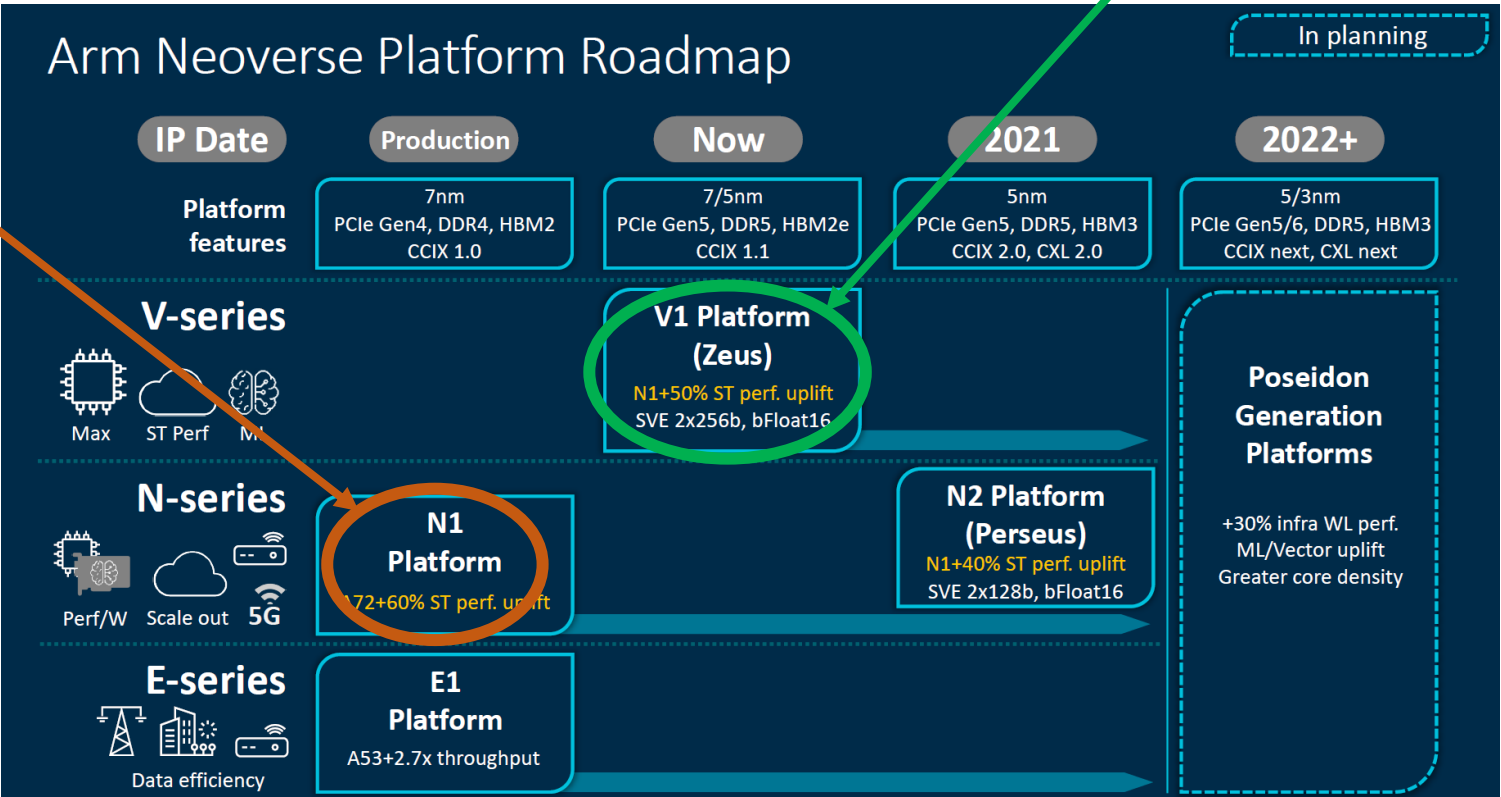


Graviton2: Existing N1 silicon,  
64 cores @2.5 GHz

AMD  
Intel

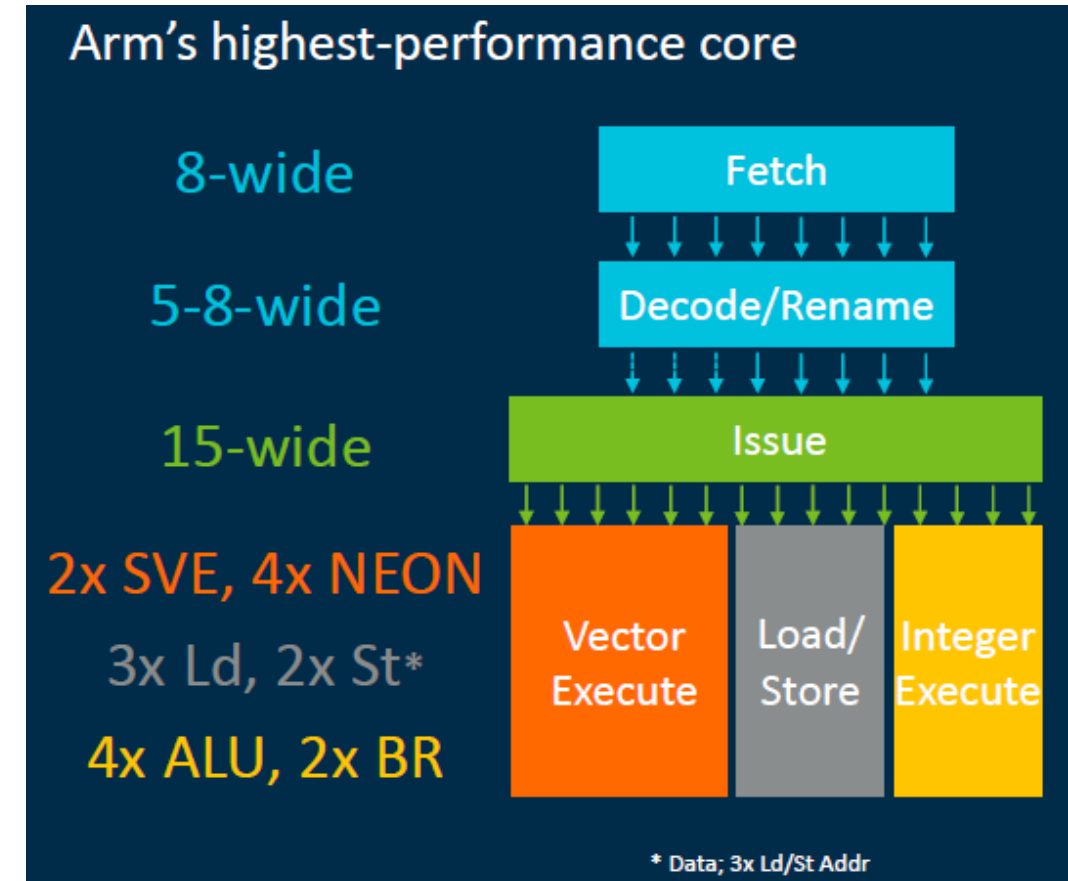
“Zeus” (V1) core used by SiPearl,  
64 cores @ 2.5+GHz planned

**V1 Platform  
(Zeus)**  
N1+50% ST perf. uplift  
SVE 2x256b, bFloat16



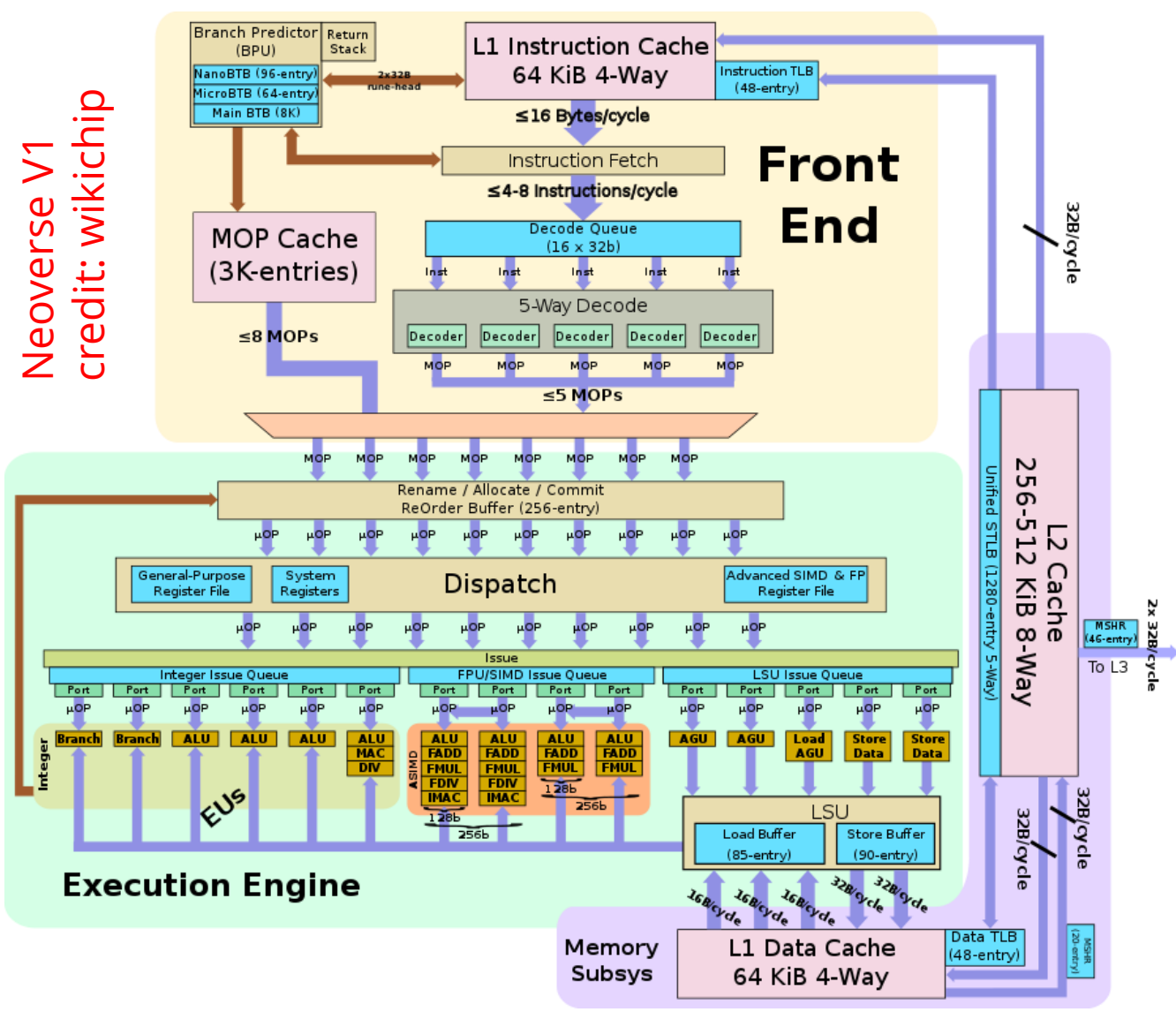
# ARM NEOVERSE V1

- Arm V8.4++
  - SIMD complex number support
  - bfloat/int8 matmul instructions
  - ...
- 2 x 256 SVE
  - Doubles as four 128-bit NEON engines for narrow vector / scalar code
  - DP/SP/BFloat16/Int8
  - Masking, scatter/gather, complex arithmetic, ...
- Ease of programming and software portability
  - augments the ability of the CPU for vector processing, AI and machine learning, and other jobs





# ARM NEOVERSE V1



Neoverse V1  
credit: wikipedia

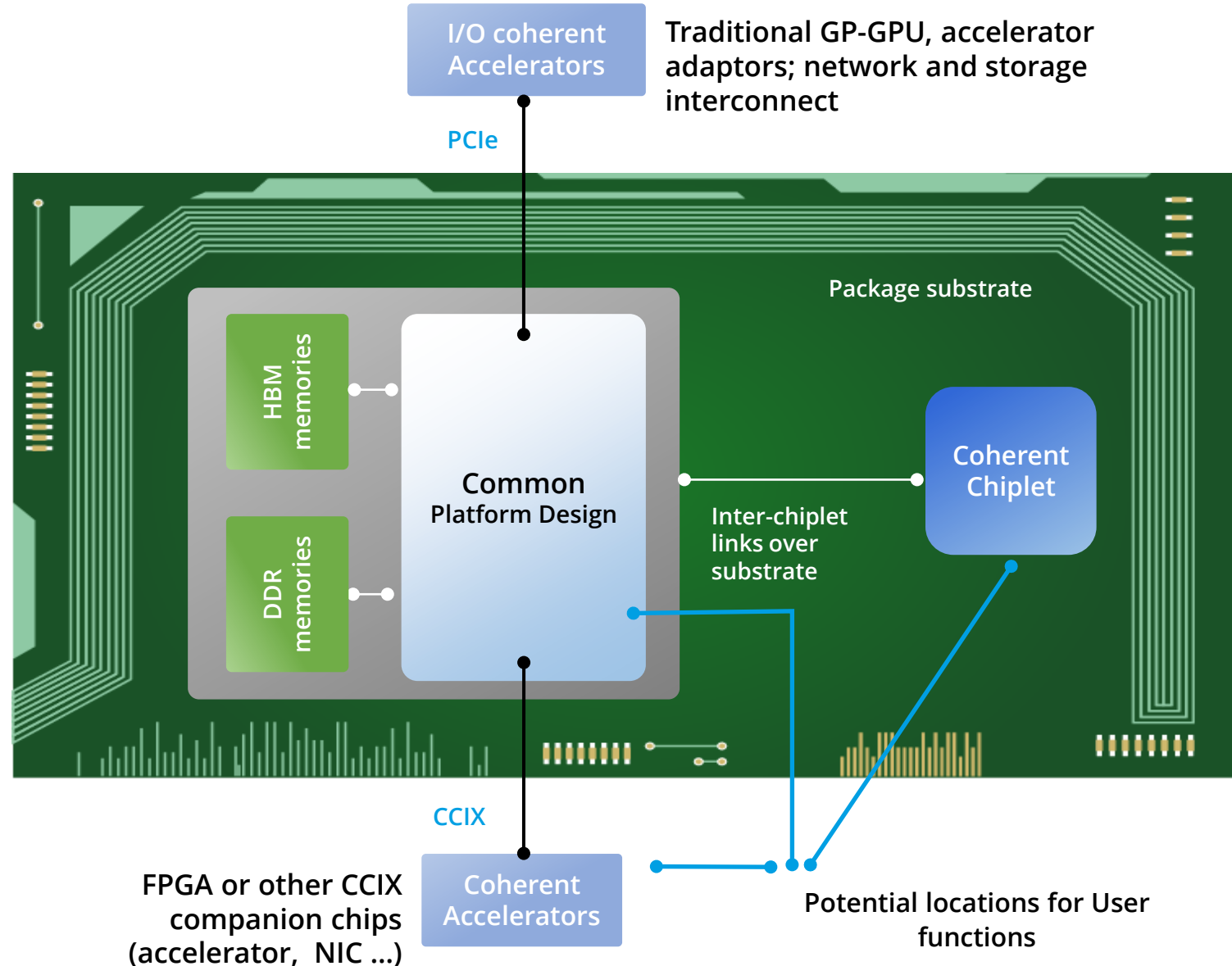
- Most sophisticated core Arm has ever done
- Optimization manual already available: <https://developer.arm.com/documentation/pjdoc466751330-9685/latest/>
- Narrower implementation of SVE than A64FX, but more flexible and with lower latencies instructions
- General-purpose, high-performance latency-oriented core

Table 3-43 SVE floating-point instructions

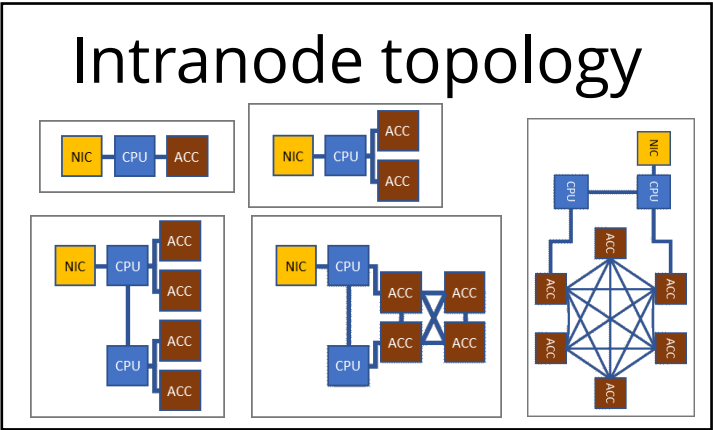
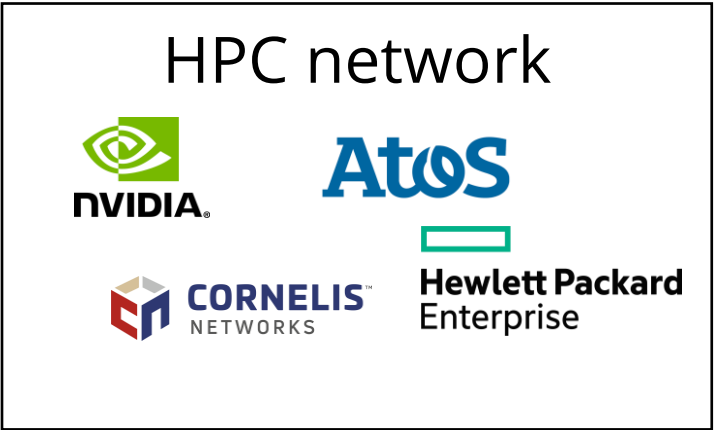
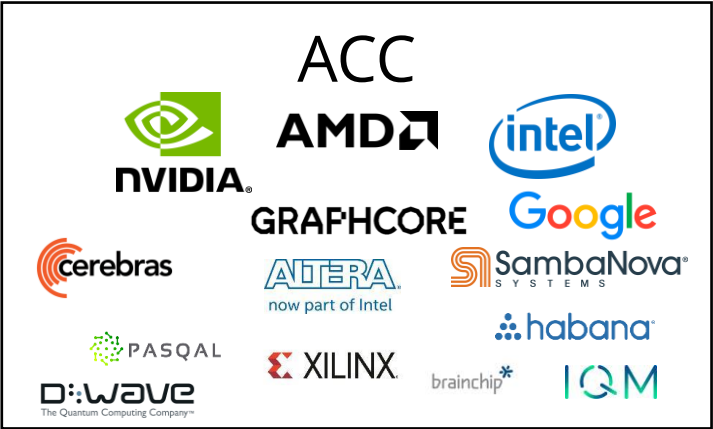
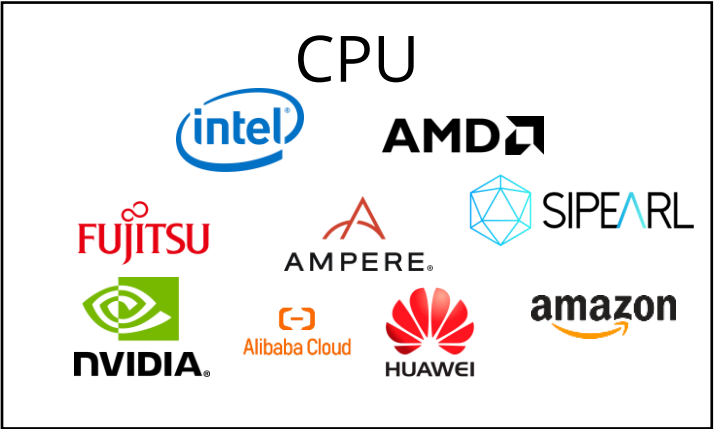
Instruction Group	SVE Instruction	Execution Latency	Execution Throughput	Utilized Pipelines	Notes
Floating point absolute value/difference	FABD, FABS	2	2	V01	-
Floating point arithmetic	FADD, FADDP, FNEG, FSUB, FSUBR	2	2	V01	-
Floating point associative add, F16	FADDA	19	1/18	V0	-
Floating point associative add, F32	FADDA	11	1/10	V0	-
Floating point associative add, F64	FADDA	8	2/3	V01	-

# RHEA IN THE EPI COMMON PLATFORM

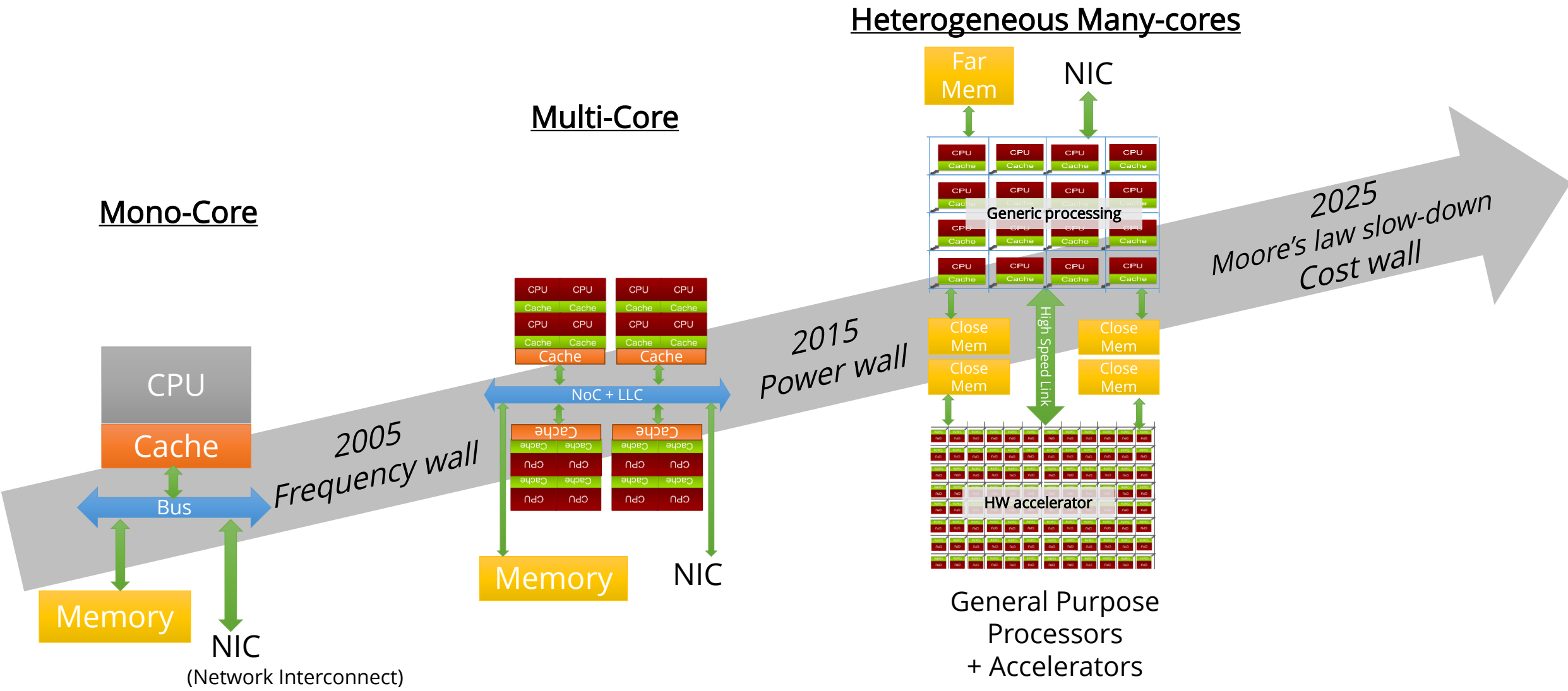
- Allows integration of customized functions in chip, in package, on board, or over PCIe or network link
- EPI Accelerators work in I/O coherent mode and share the same memory viewSingle or dual chiplet package for power efficient sizing
- Targeting high Byte/FLOP ratio
- HBM2e, DDR5 and PCIe gen5
- Coherent NoC with system level cache to keep data local
- D2D interface open to EPI (and beyond)



# HPC – THE HETEROGENEITY EXPLOSION

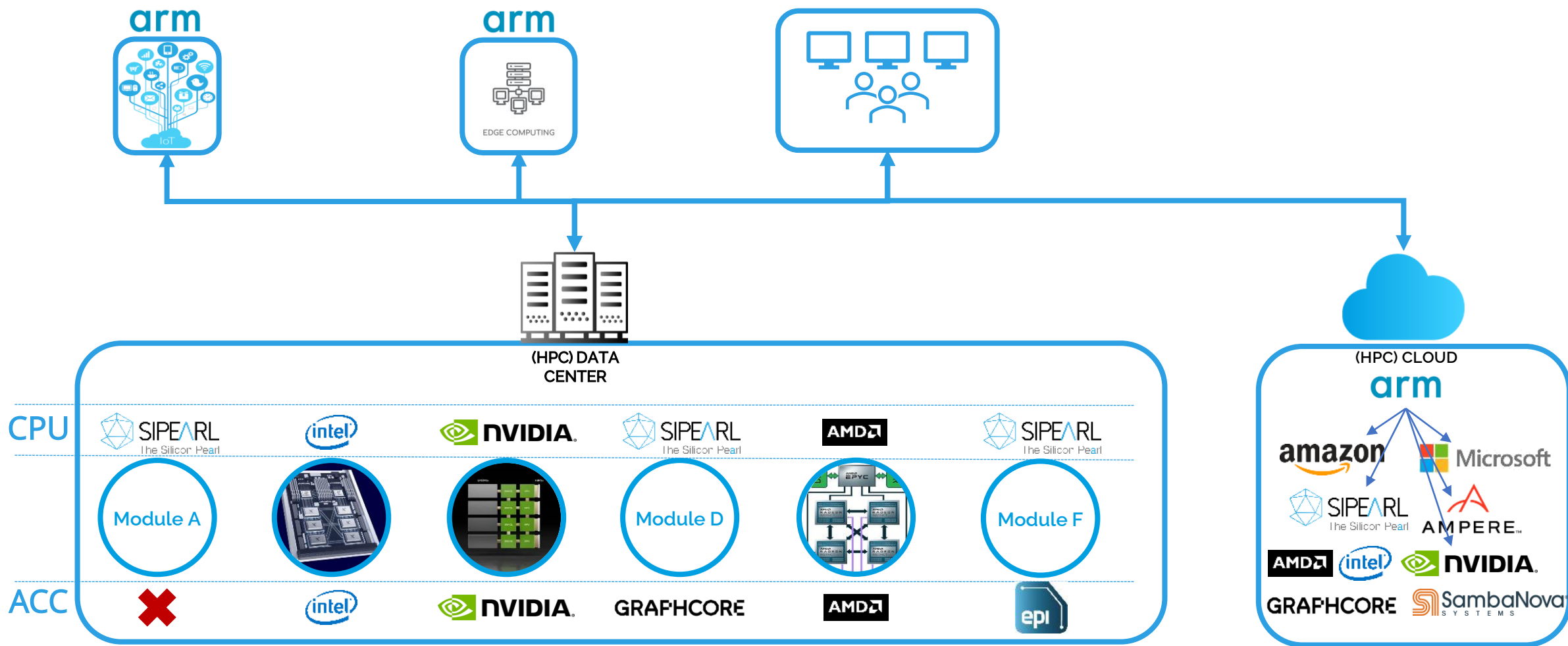


# COMPUTE NODE ARCHITECTURE EVOLUTION





# HPC MODULAR AND HYBRID ARCHITECTURE



# WHAT ABOUT CODES ?

- How do users deal with such complexity ?
  - Manually
  - Programming language
  - Programming model
  - Programming framework
  - Offline tools
  - Online tools
  - ...
- Traditional languages and parallelism
  - Fortran, C, C++, OpenMP, MPI, ...
- Data-parallel language/ model / framework
  - CUDA, OpenCL, SYCL, Kokkos, StarPU, SkePU, RAJA, OneAPI, ...
- Domain-specific languages
- ...
- Research projects and CoE: what to prioritize for support ?
- Co-design of future systems to ensure adequation between hardware capabilities and the *ability of software to exploit them*



# SIPEARL

SiPearl on HPC  
March 2022  
TREX Workshop

Thank You !

Romain Dolbeau  
[Romain.Dolbeau@sipearl.com](mailto:Romain.Dolbeau@sipearl.com)