



ZOMI

# 高性能计算 发展趋势

# Content



# AI Infra Architecture

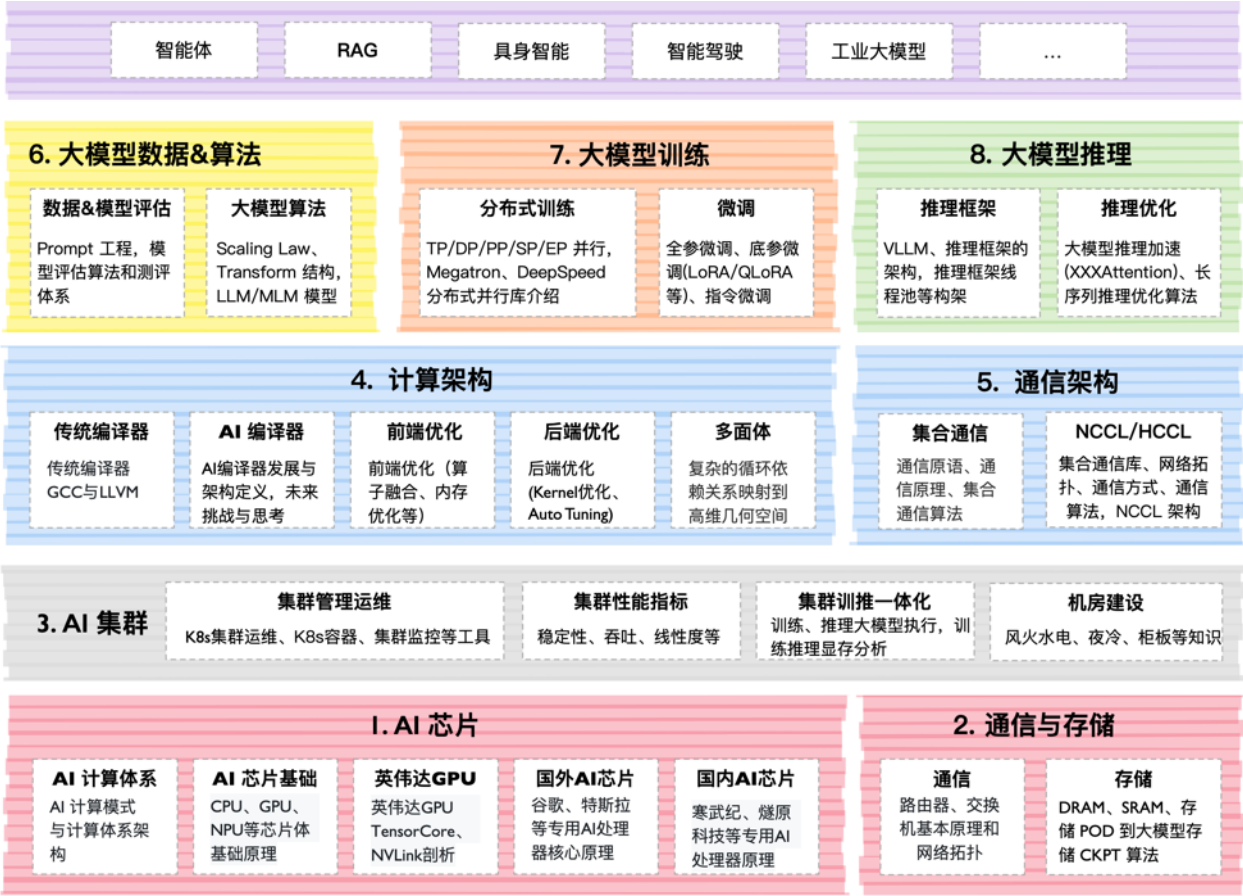


时事  
热点

大模型  
训推

编译  
计算  
架构

硬件  
体系  
结构



# 三类计算集群的主要区别

集群类型	应用目的	数据类型	计算特征	网络特征	存储特征
HPC高性能计算	国防科研	FP64	计算密集，高并行	密集通信（每个应用流量特征不同，部分可隐藏）	密集&复杂 IO
AI人工智能计算	AI训练推理	FP16/BF16/FP8/FP4		密集通信（部分可隐藏）	密集 IO（按节奏迭代）
云数据中心	互联网云计算	INT32 FP32	通用计算，高并发	分散通信	分散/密集 IO





# Content

1. 核心硬件（高性能-处理器、存储、网络、服务器）
2. 基础软件（编译器与运行时、计算库、通信中间件、存储系统、调度系统）
3. 应用软件（发展历程、行业应用趋势）



02

基础软件

Software



# 基础软件

1. HPC 基础软件是连接硬件与应用的核心桥梁
2. 发展历程围绕资源高效调度、数据高速流动、计算极致优化
3. 主要有调度、存储、通信、编译、计算库及基础算法六大维度



# 01. 编译器与运行时

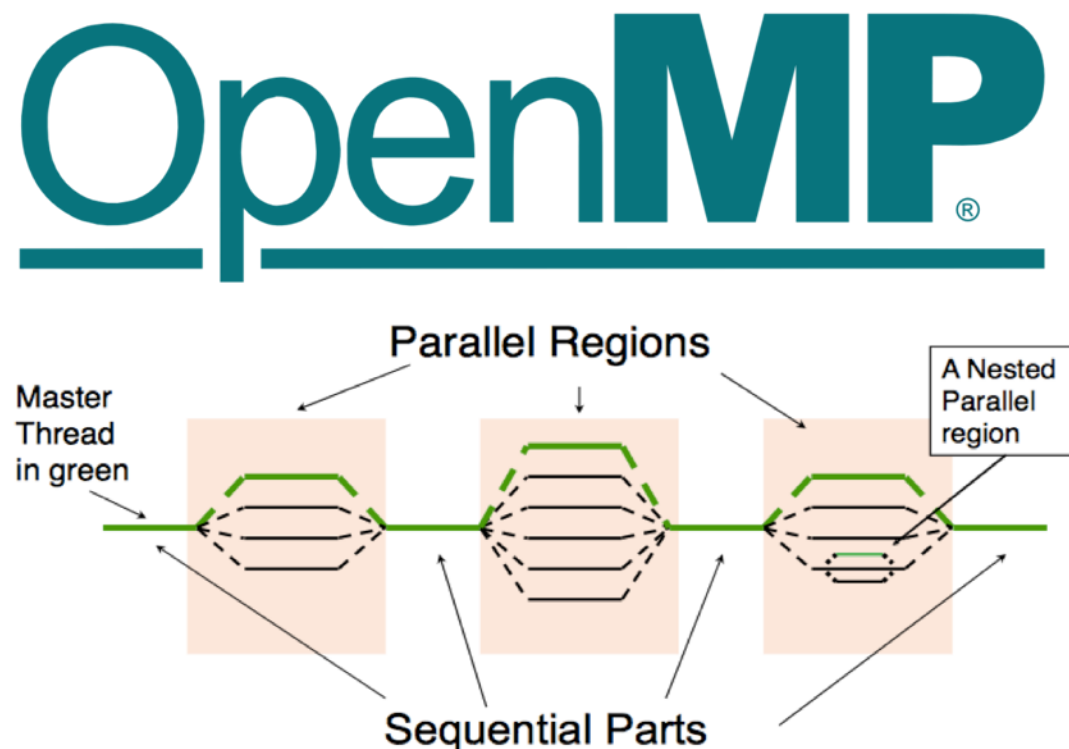
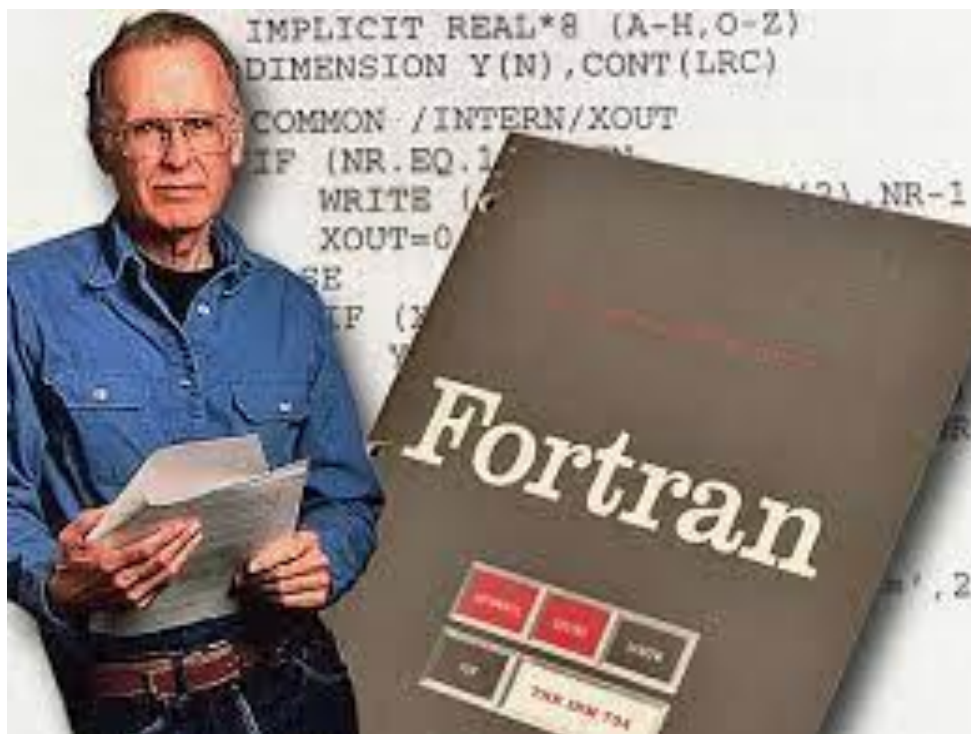
## 从手动优化到自动代码生成





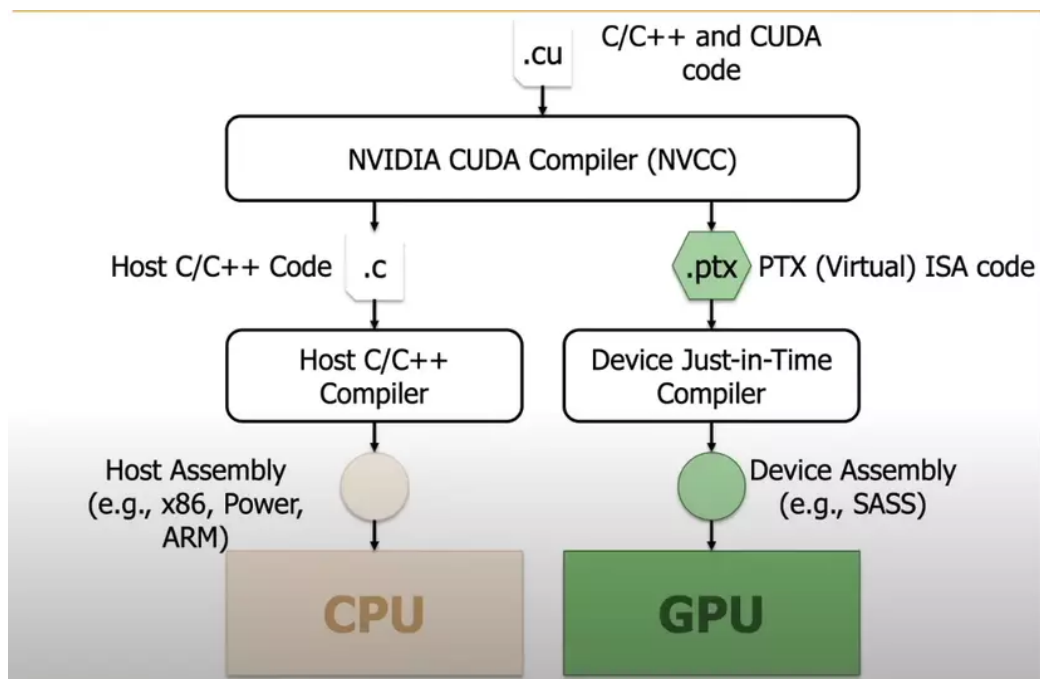
# 发展历程：传统编译器（1980s-2000s）

- Fortran 主导：Intel Fortran Compiler、PGI Compiler 通过向量化和循环展开提升性能。
- OpenMP 初现：1997 年推出共享内存并行编程模型，支持多线程优化。



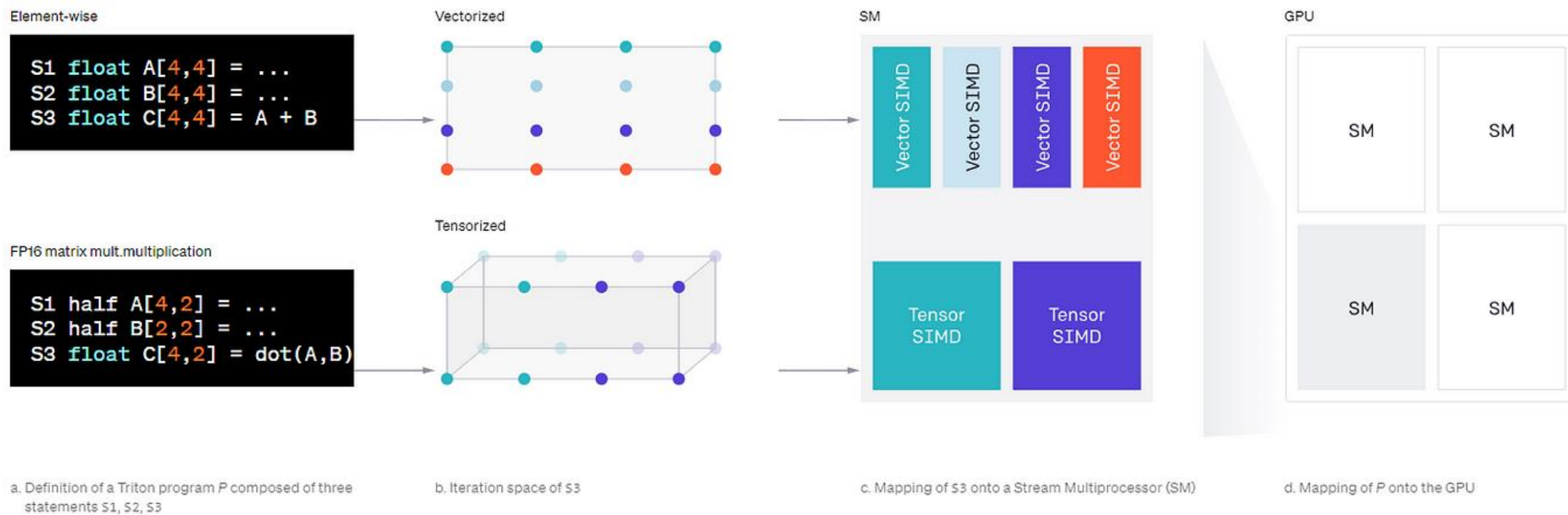
# 发展历程：异构编译时代（2010s~至今）

- CUDA & OpenCL: NVIDIA NVCC 实现 CUDA 代码生成, OpenCL支持跨厂商异构设备。
- SYCL & oneAPI: 基于 C++ 的 SYCL编译器实现一次编写, 多架构运行。
- 国产突破: 华为毕昇编译器优化鲲鹏 CPU 与昇腾 GPU 异构计算。



# 未来趋势

1. **AI 驱动自动调优**：TVM、Ansor 等通过机器学习算法自动生成最优计算内核。
2. **多级中间表示**：LLVM MLIR 项目实现跨架构代码转换，降低编译器开发成本。
3. **面向新编程方式**：e.g. Triton 编译器一种基于块级数据流分析的 AI 编译器。



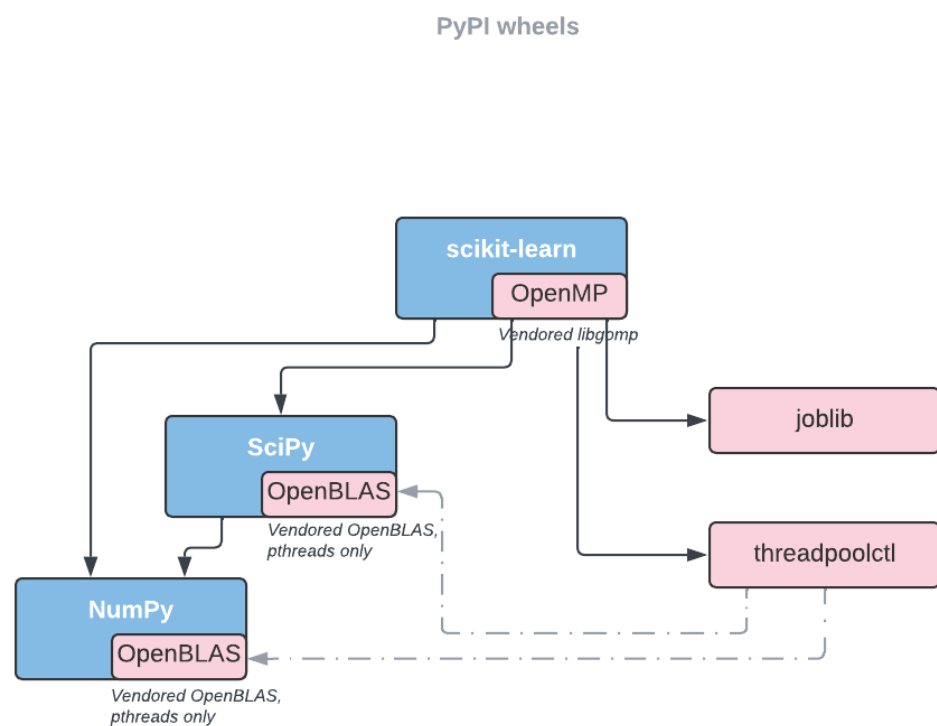
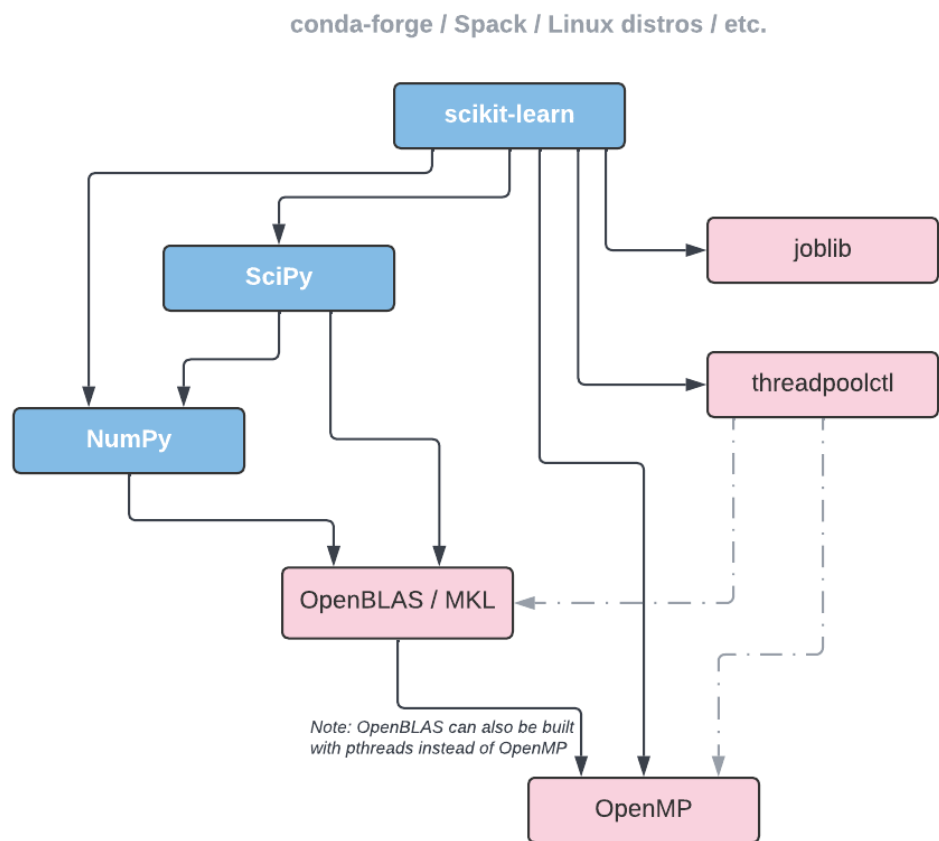
# 02. 计算库

从基础数学到领域专用



# 发展历程：经典数学库（1970s-2000s）

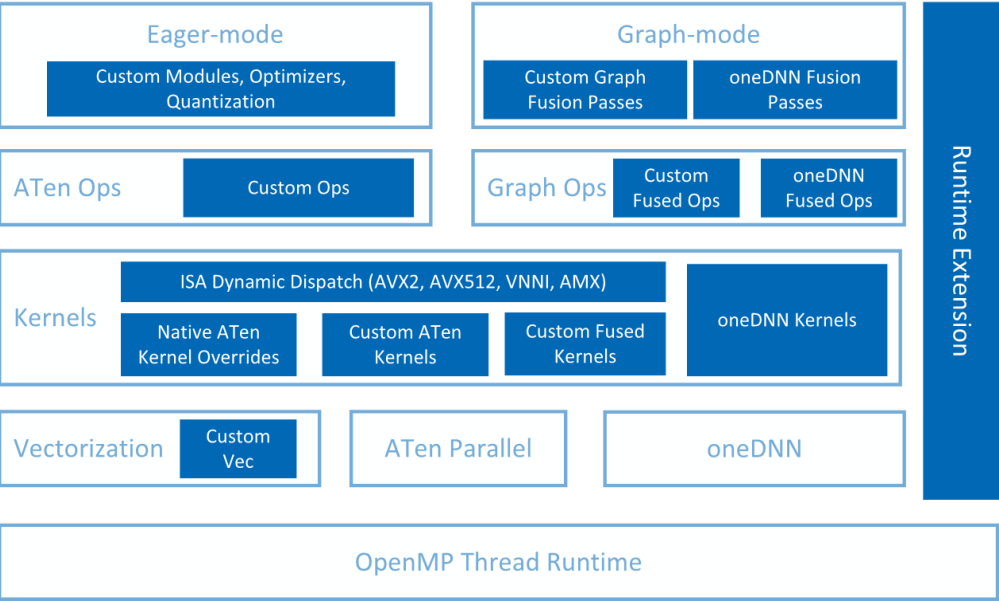
- BLAS/LAPACK：基础线性代数子程序库（e.g. Intel MKL、OpenBLAS）优化矩阵运算性能。
- FFTW：快速傅里叶变换库支持千万级数据点计算。



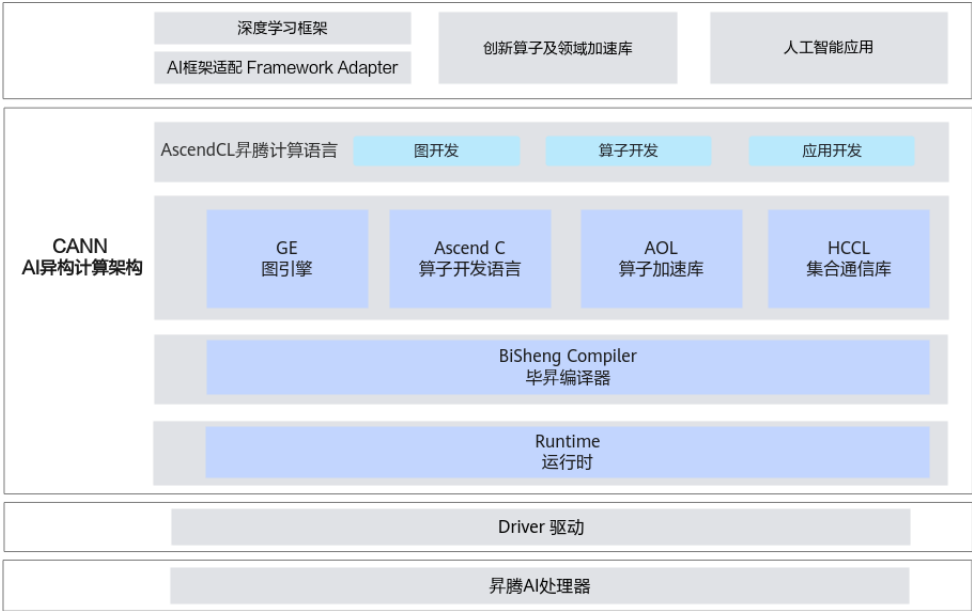
# 发展历程：领域专用库（2010s至今）

- **稀疏矩阵库**：SuiteSparse、PETSc 支持非结构化网格求解。
- **AI加速库**：cuDNN（NVIDIA）、OneDNN（Intel）优化 AI 算子。
- **国产库突破**：华为CANN异构计算架构提供 AI算子库。

## OneDNN



## CANN

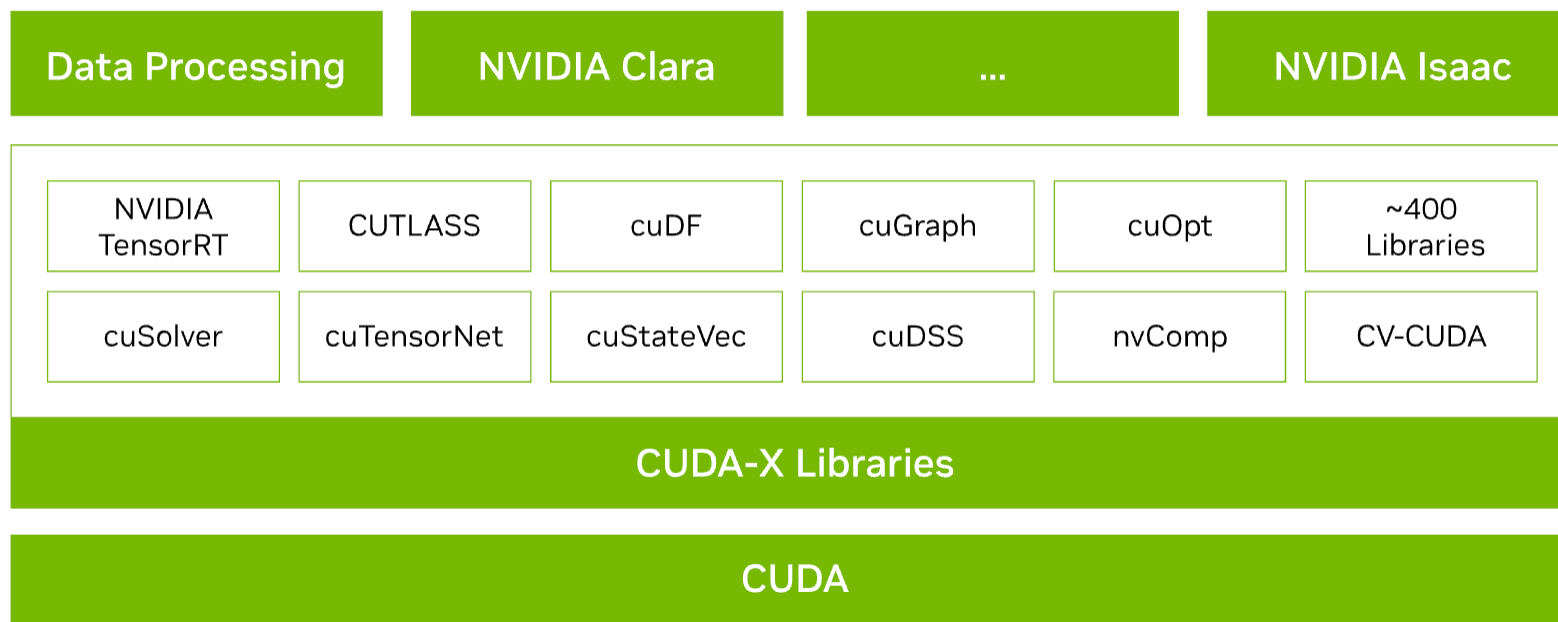




# NVIDIA CUDA 计算库

- **AI与机器学习库:** cuDNN、 cuML、 cuGraph、 cuOpt
- **数学与科学计算库:** cuBLAS、 cuFFT、 cuSPARSE、 cuSolver、 cuRAND、 cuTENSOR
- **数据处理与分析库:** cuDF、 Thrust; **图像与多媒体处理库:** NPP

## CUDA-X Microservices



# 未来趋势

- 自动微分库：JAX、Enzyme 支持反向传播与 Hessian 矩阵计算。
- 量子算法库：Qiskit Aer 模拟器支持量子线路优化，加速量子化学模拟。
- 领域特定语言 DSL：Julia语言 SciML 生态实现微分方程求解自动化。



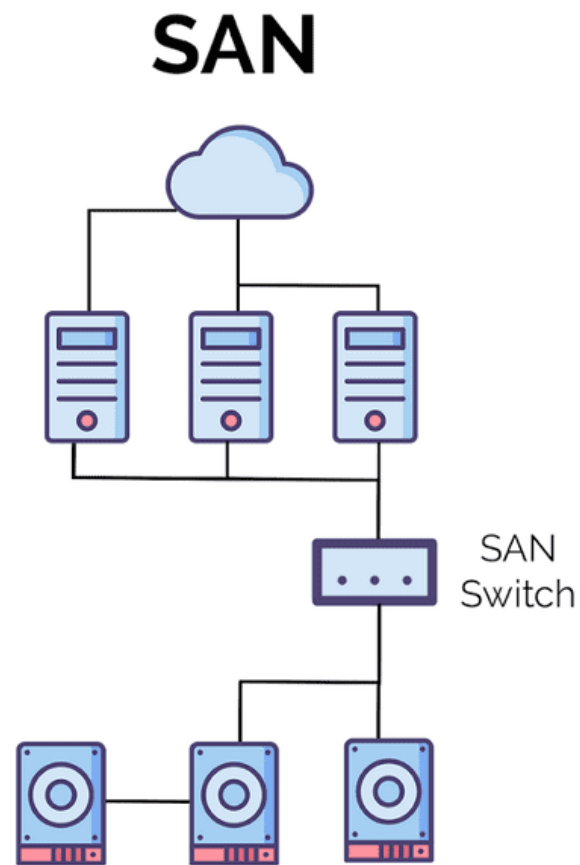
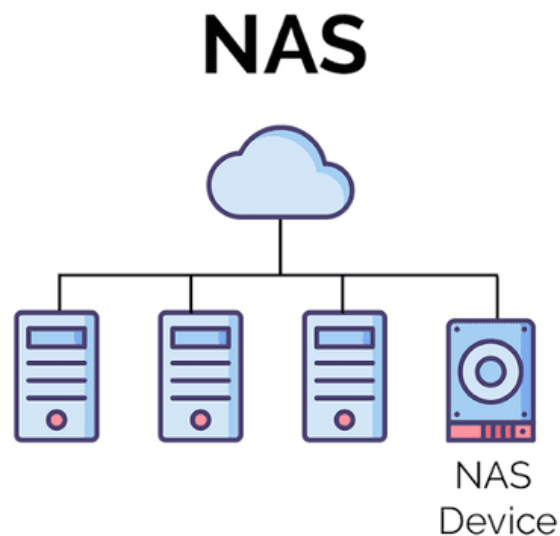
# 03. 存储系统

## 本地IO到分布式高速访问



# 发展历程 1：本地存储时代（1980s ~ 2000s）

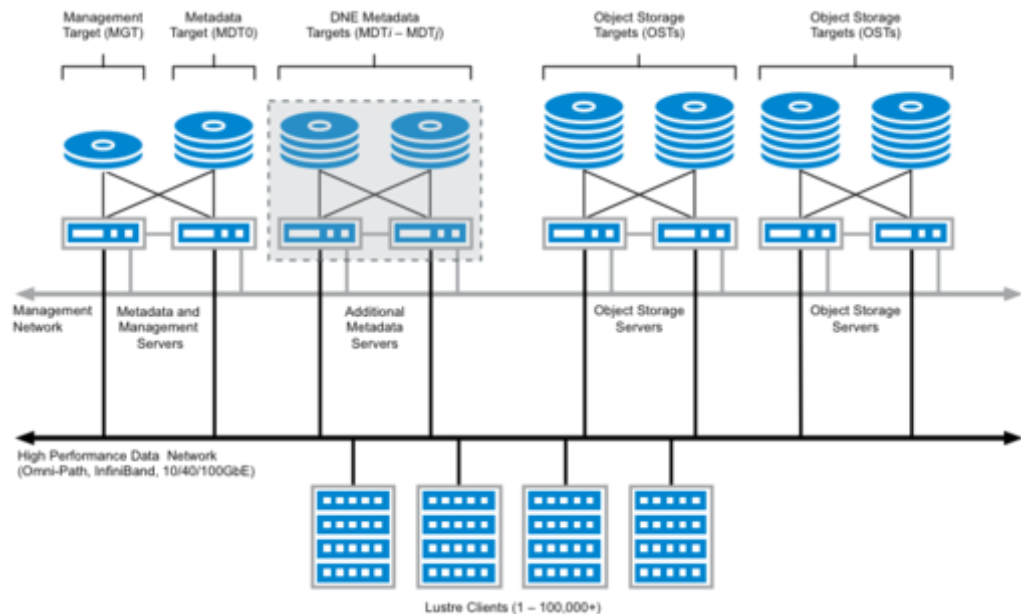
- NFS & SAN：网络文件系统 NFS 和存储区域网络 SAN 解决局部共享问题，但带宽瓶颈明显（<1G B/s）。



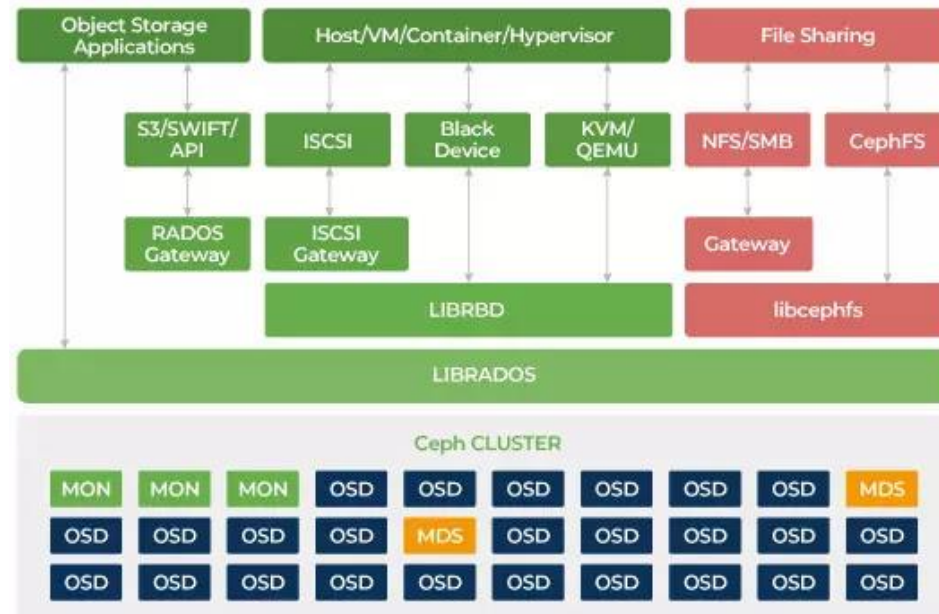
## 发展历程 2：并行文件系统（2010s ~ Today）

- Lustre：开源并行文件系统，支撑全球 Top 500 超算中 70% 的存储
- Ceph & Gluster：对象存储 Ceph 与分布式卷存储 Gluster 支持 EB 级数据管理
- 国产突破：DeepSeek 3FS 通过优化数据存储和访问流程降低 MoE 训推成本

### Lustre



### Ceph



# 未来趋势

1. **存储计算融合**: CXL 协议实现存储池化, 支持 GPU 直接访问远程持久内存。
2. **智能数据预取**: 基于机器学习的 IO 预测减少数据读写延迟。
3. **冷热数据分级**: 结合 SSD/NVMe 与 HDD 混合存储架构, 成本降低 30% 同时保障性能。





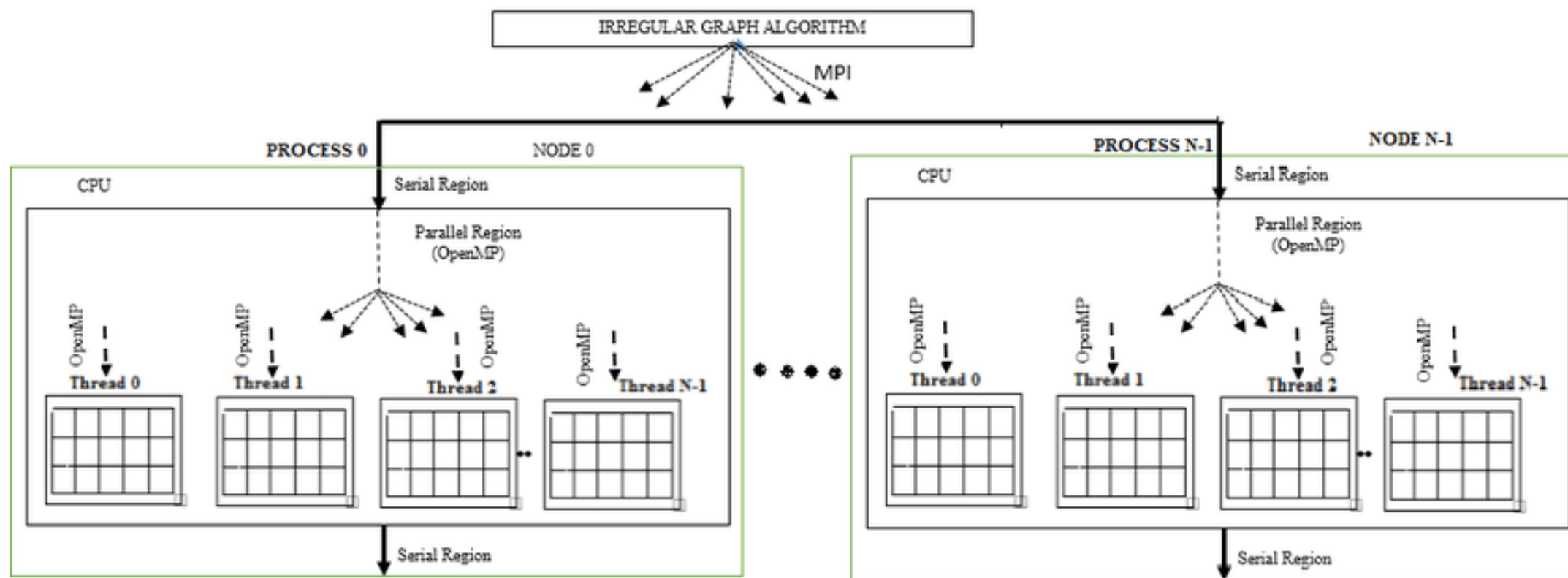
# 04. 通信中间件

## 从 MPI 到跨架构互联



# 发展历程：MPI 统治时代（1990s-2010s）

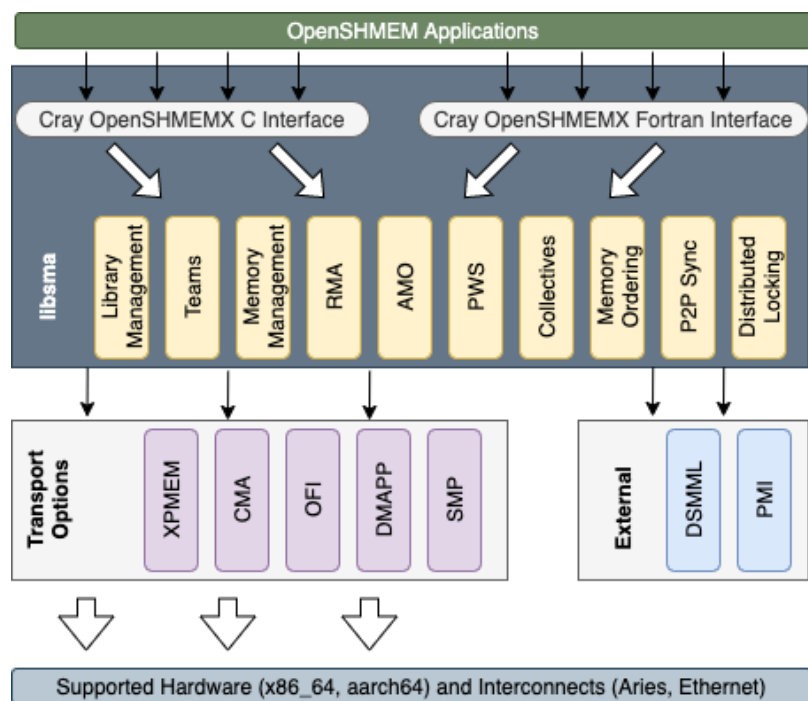
- 消息传递接口 MPI：HPC 通信标准（Open MPI/Intel MPI），支持P2P与集合通信。
- 性能优化：RDMA 技术将延迟降至  $1\mu\text{s}$  级（e.g. InfiniBand Verbs API）。



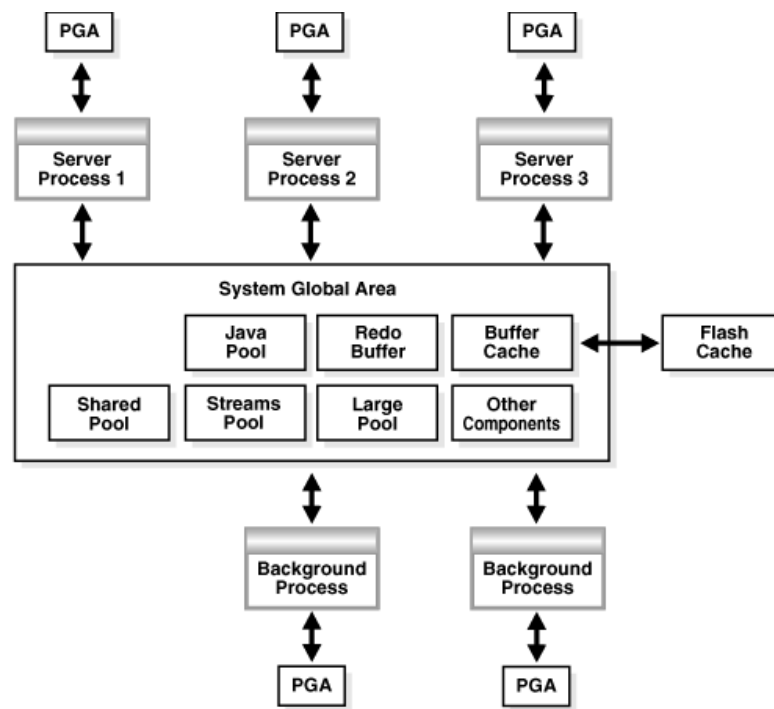
# 发展历程：多协议协同（2020s至今）

- UCX（Unified Communication X）：支持 MPI/OpenSHMEM/PGAS 等多编程模型。
- 云原生通信：gRPC 与 RDMA 结合降低跨容器通信开销。

## OpenSHMEM



## PGAS



# 未来趋势

1. 光网络抽象层：硅光芯片驱动的零拷贝通信协议（e.g. Intel A photonics API）。
2. 自适应路由：基于拓扑感知的动态路径选择（e.g. SLURM Gang Scheduling）。



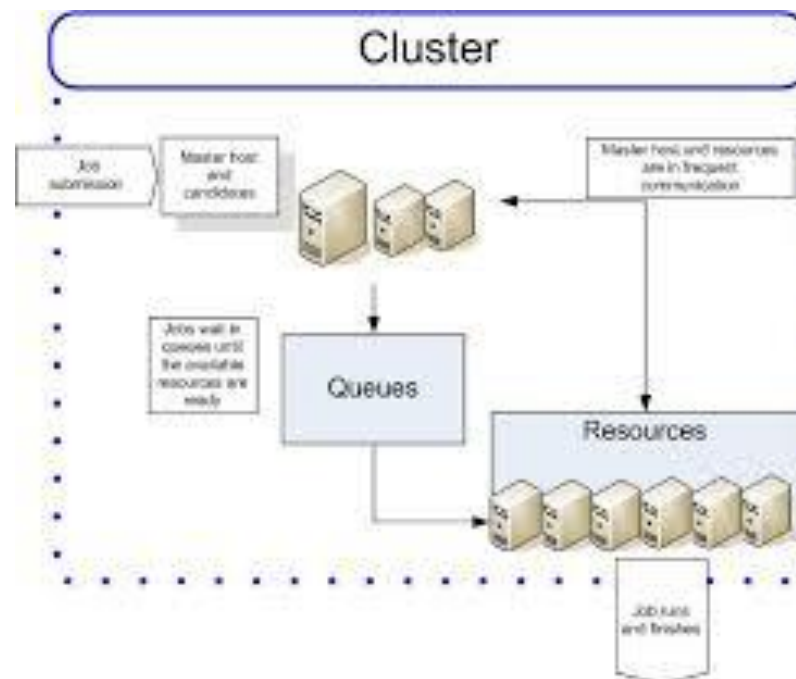
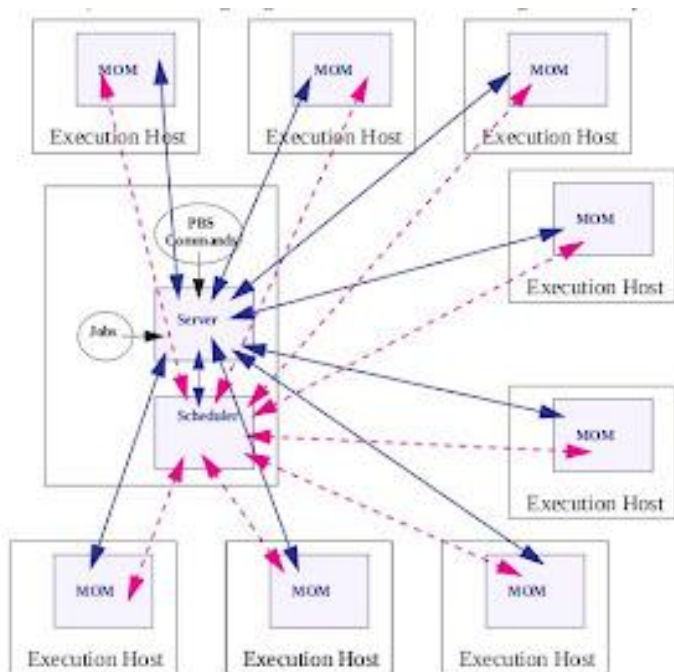
# 05. 调度系统

静态分配 to 智能动态管理



# 发展历程：早期阶段（1990s-2000s）

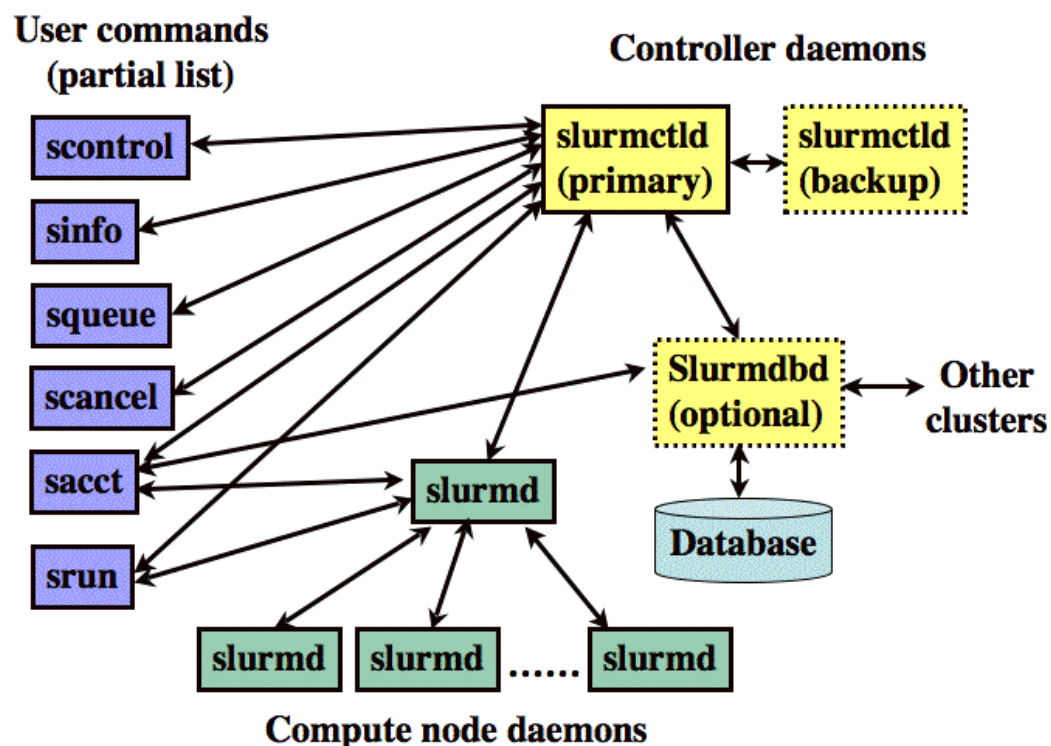
- **单机调度工具**：e.g. Unix 的 at 和 cron 仅支持本地任务调度。
- **集群调度系统**：Portable Batch System（PBS）、Load Sharing Facility（LSF）作业队列管理，支持多节点资源分配。



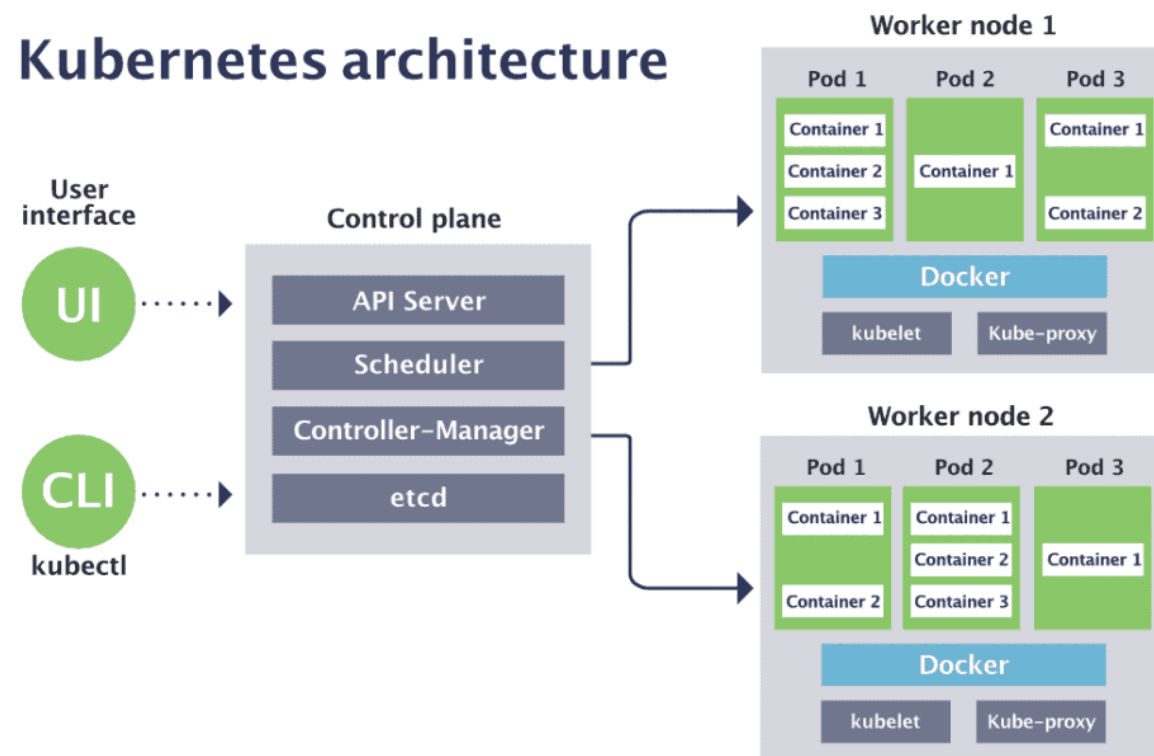


# 发展历程：现代调度系统（2010s至今）

- Slurm：开源社区主导，支持千万级核心调度，提供抢占式调度、NPU 资源细粒度分配功能。
- 云原生调度**：Kubernetes通过插件扩展支持HPC混合负载，实现容器化任务弹性伸缩。



## Kubernetes architecture



# 未来趋势

- **AI驱动预测调度**：基于 RL 调度算法，动态优化任务优先级。
- **异构资源协同**：CPU/GPU/NPU 混合调度。
- **绿色节能调度**：结合硬件功耗监控、能耗感知调度策略，降低 PUE 10%-20%。



# 基础软件面向未来挑战

- **硬件碎片化**: 开放标准 (oneAPI/UXL) + 统一中间表示 (MLIR)
- **存储与计算失衡**: 存算一体架构 (Near-Data Processing) + CXL 内存池化
- **能源效率优化**: 运行时动态调频 + 作业功耗感知调度, 算法硬件协同设计降低每 FLOP 能耗。
- **HPC 基础软件将向智能化、一体化、开放生态演进**

03

应用软件

Application



# 阶段1：专用代码时代（1990s-2000s）

- **特征：**

- 各领域自研封闭式代码，强依赖底层硬件（e.g. 核爆模拟 LASNEX、气象 WRF MM5）

- **瓶颈：**

- 并行扩展性差（MPI 实现粗糙）
- 维护成本高（Fortran/C 主导，代码动辄百万行）
- 移植困难（e.g. 从向量机转向集群）



# 阶段2：社区软件包崛起（2000s-2010s）

- 领域标准化：出现可复用开源框架，行业专家与 HPC 开发者协作优化

领域	代表软件	基础技术支撑
流体力学	OpenFOAM, SU2	MPI并行 + 非结构化网格库
分子模拟	GROMACS, LAMMPS, NAMD	粒子邻居列表算法 + GPU加速
天气预报	WRF, MPAS	嵌套网格并行 + 高效通讯库
宇宙学	ENZO, GADGET	多尺度树形算法（Barnes-Hut）
结构分析	CalculiX, Code_Aster	稀疏线性求解器（MUMPS, PETSc）





## 阶段3：多学科融合与AI赋能（2020s-至今）

- **新需求驱动：**

- 精度提升（e.g. 飞机制造中亿级网格 CFD 仿真）
- 实时性要求（e.g. 量化交易毫秒级响应）
- 数据爆炸（e.g. 基因测序 PB 级数据处理）

- **技术革命：**

- AI/ML融入：传统模拟结合深度学习代理模型（e.g. 气候预测中 Diffusion Model）
- 云原生架构：通过 Kubernetes 管理混合 HPC+AI workflow



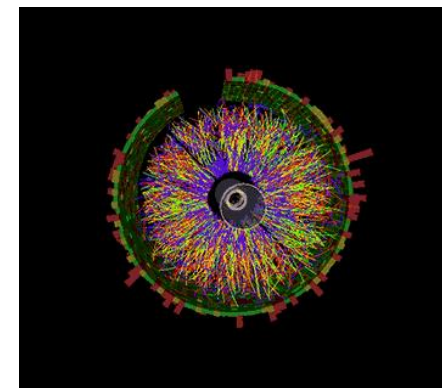
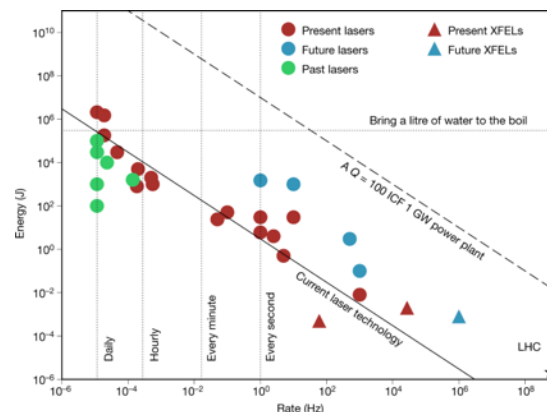
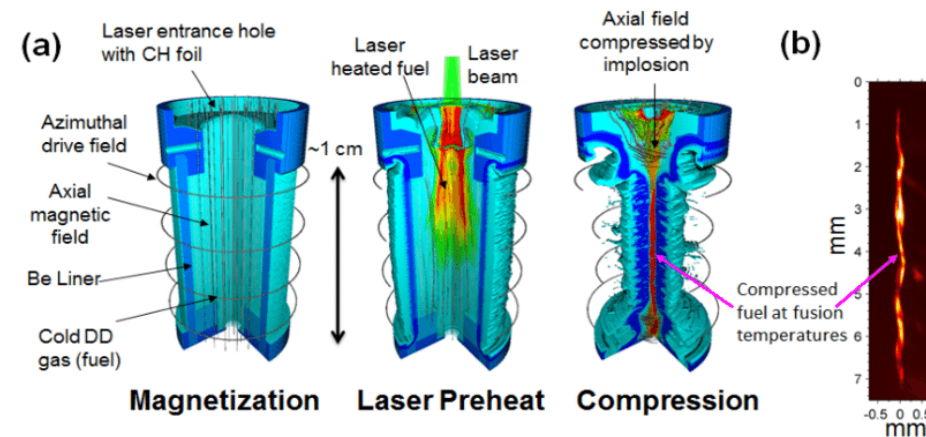
# 行业应用1：基础科学领域

- 核爆模拟：

- 趋势：多物理场强耦合（辐射流体+材料相变）
- 技术：自适应网格加密（AMR）+ 异步通信（减少全局同步）
- 案例：美国 LA-UR 代码在 Frontier 超算上实现 exascale 扩展

- 高能物理：

- 趋势：边缘计算预处理（LHC 实验中先过滤 99.99% 数据）
- 技术：流式处理框架（Apache Kafka）+ GPU 加速分析



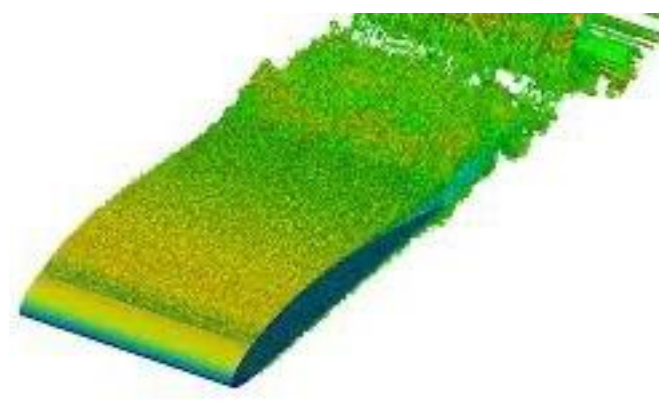
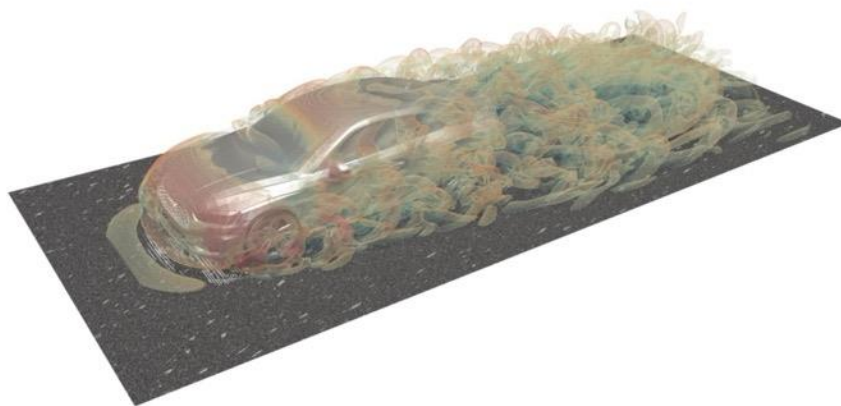
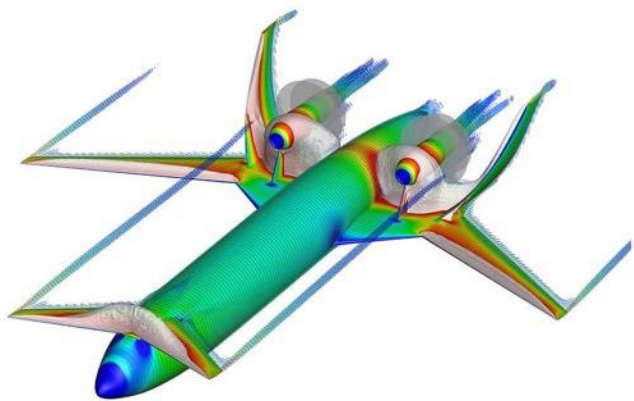
## 行业应用2：工程制造领域

- **飞机制造：**

- 趋势：数字孪生实时校准（基于传感器数据更新模型）
- 技术：非结构化网格求解器（SU2 的 GPU 版）、多目标优化库（DAKOTA + ML 代理模型）

- **汽车碰撞测试：**

- 趋势：混合精度计算（FP64 验证 + FP16 训练安全 AI 模型）
- 工具链：LS-DYNA + NVIDIA Modulus 物理 AI 框架



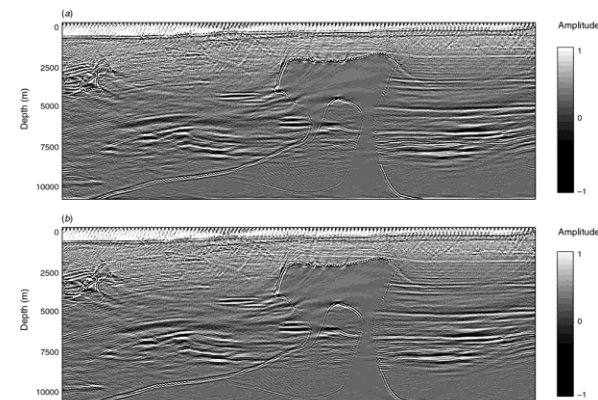
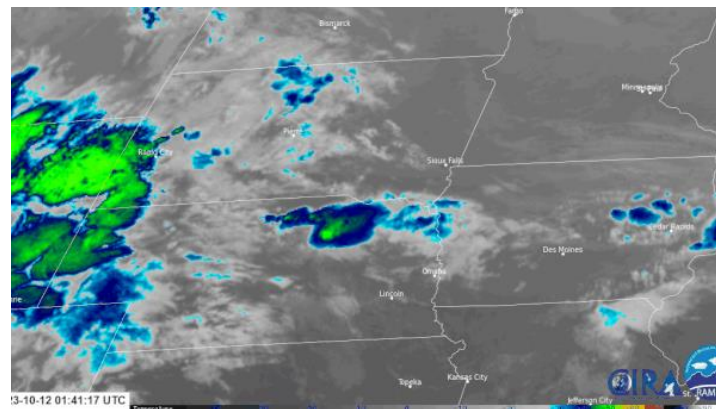
# 行业应用3：地球科学领域

- 天气预报：

- 趋势：公里级分辨率全球模拟
- 软件演进：WRF → MPAS（非结构化网格） → FVM；四维变分同化（4D-Var）算法优化

- 油气勘探：

- 趋势：全波形反演（FWI）
- 性能瓶颈突破：通信避免算法（CA-FWI）；波场重建用Tensor Core加速



# 总结与思考





# Question? HPC vs AI

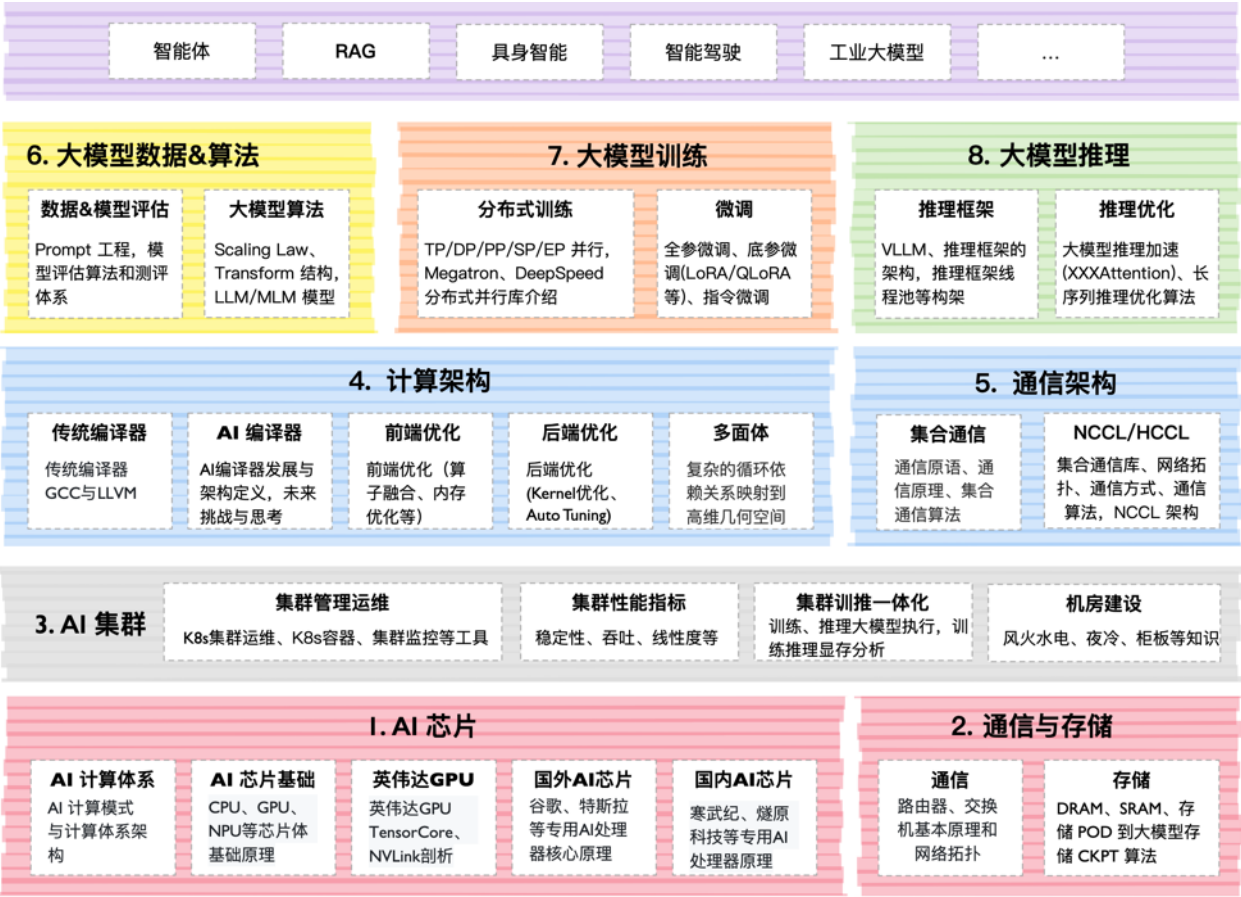


时事  
热点

大模型  
训推

编译  
计算  
架构

硬件  
体系  
结构





# Thank you

把AI系统带入每个开发者、每个家庭、  
每个组织，构建万物互联的智能世界

Bring AI System to every person, home and  
organization for a fully connected,  
intelligent world.

Copyright © 2024 XXX Technologies Co., Ltd.  
All Rights Reserved.

The information in this document may contain predictive statements including, without limitation, statements regarding the future financial and operating results, future product portfolio, new technology, etc. There are a number of factors that could cause actual results and developments to differ materially from those expressed or implied in the predictive statements. Therefore, such information is provided for reference purpose only and constitutes neither an offer nor an acceptance. XXX may change the information at any time without notice.



**ZOMI**

GitHub <https://github.com/chenzomi12/AllInfra>



ZOMI

# 引用与参考

- <https://zhuanlan.zhihu.com/p/683671511>
- PPT 开源在: <https://github.com/chenzomi12/AllInfra>

