



# 开放架构HPC技术与生态白皮书



# 开放架构 HPC 技术与生态白皮书

## 编写组委会

顾问：陈国良 陈润生 郑纬民 袁国兴  
主编：张云泉 卢凯 张熙伟 安静  
执行主编：刘杰 贾海鹏 袁良 谢海波  
编委：龚春叶 方建滨 张先轶 张华 陈超 张超  
秦品乐 黄典 林新华 李会民 陈文波 张洋  
孟祥飞 康波 霍吉东 王继彬 黎宇翔 周文韬  
毛峰 金天骄 田洪镇 王鹤潼 张群 代民果  
王军辉 崔向阳 李若 肖晟 易大江 陈志广  
强胜龙 赵磊 戴中洋 柳朝晖 张茂强 杨明林  
田群 戴光丰 (排名不分先后)

指导单位：中国计算机学会高性能计算专委会

联合发布单位：

国防科技大学	国家超级计算天津中心
清华大学	国家超级计算长沙中心
上海交通大学	国家超级计算深圳中心
中国科学技术大学	国家超级计算济南中心
华中科技大学	国家超级计算广州中心
兰州大学	上海超级计算中心
西安电子科技大学	武汉超算中心
西北工业大学	西安未来计算中心
北京理工大学	
中北大学	华为技术有限公司
中国科学院计算技术研究所	飞腾信息技术有限公司
中国科学院计算机网络信息中心	中国移动通信集团浙江有限公司
中国科学院过程工程所	深圳华大生命研究院
中国科学院国家天文台	彭峰（北京）科技有限公司
自然资源部第一海洋研究所	湖南迈曦软件有限责任公司
中国空气动力研究与发展中心	英特工程仿真技术（大连）有限公司
航空工业西安航空计算技术研究所	深圳十洋科技有限公司
中国核动力研究设计院	北京龙讯旷腾科技有限公司
广东省气象局	湖南迈曦软件有限责任公司
	北太振寰（重庆）科技有限公司

(排名不分先后)

## 目录

<b>1 引言</b>	01
<b>2 技术篇</b>	03
2.1 从国家超算战略看 ARM 指令集兼容架构在 HPC 领域的发展	04
2.2 从 TOP 500 看 ARM 指令集兼容架构在 HPC 领域的发展	04
2.3 ARM 指令集兼容架构在商用 HPC 领域的崛起	07
2.4 ARM 指令集兼容架构针对 HPC 的改进及演化	10
2.4.1 从终端走向 HPC	10
2.4.2 面向 HPC 的 ARM 指令集兼容架构演进	12
2.4.3 国内 ARM 指令集兼容架构处理器针对 HPC 领域的进展	12
2.5 ARM 指令集兼容架构在 HPC 的未来发展趋势	14
<b>3 生态篇</b>	15
3.1 ARM 指令集兼容架构面向 HPC 领域的生态构成	16
3.2 操作系统生态	16
3.2.1 概述	16
3.2.2 OpenEuler	17
3.2.3 银河麒麟	17
3.2.4 Red Hat	18
3.2.5 SUSE	18
3.2.6 Ubuntu	18
3.2.7 Arch Linux	19
3.2.8 Debian	19
3.2.9 Fedora	19
3.3 编译器生态	20
3.3.1 ARM 官方编译器	20

3.3.2 亚马逊编译器	20
3.3.3 Ampere 编译器	21
3.3.4 富士通编译器	21
3.3.5 Cray 编译器	22
3.3.6 毕昇编译器	23
3.3.7 天鹰编译器	23
<b>3.4 平台层生态</b>	<b>24</b>
3.4.1 概述	24
3.4.2 OpenHPC 框架	24
3.4.3 华为鲲鹏 DevKit 与 BoostKit	25
3.4.4 ARM 性能库 ARMPL	25
3.4.5 Arm Compute Library 开源数学库	26
3.4.6 华为 KML 数学库	27
3.4.7 OpenBLAS 开源数学库	27
3.4.8 LIBXSMM 开源数学库	28
3.4.9 LibShalom 开源数学库	28
3.4.10 BLASFEO 开源数学库	29
3.4.11 BLIS 开源数学库	30
3.4.12 ATLAS 开源数学库	30
3.4.13 稠密线性系统解法器	31
3.4.14 稀疏线性系统解法器	31
3.4.15 ELPA 社区	32
<b>3.5 面向 ARM 兼容指令集架构的应用软件生态</b>	<b>32</b>
3.5.1 概述	32
3.5.2 基础科研	34
3.5.3 气候海洋	37



3.5.4 制造仿真 .....	41
3.5.5 材料计算 .....	43
3.5.6 生命科学 .....	44
3.5.7 地球物理 .....	46
<b>3.6 面向 ARM 指令集兼容架构的国产科学 / 工业软件</b> .....	<b>47</b>
3.6.1 概述.....	47
3.6.2 中国科学院过程工程研究所芯片后端验证工具 ChEDA .....	47
3.6.3 中国科学院过程工程研究所离散颗粒运动模拟软件 DPS .....	47
3.6.4 中国科学院过程工程研究所拟颗粒模拟 PPM+ .....	48
3.6.5 中国科学院国家天文台天体物理模拟软件 PHoToNs .....	48
3.6.6 中国科学院计算机网络信息中心相场模拟软件 ScLETD.....	48
3.6.7 中国科学院计算技术研究所分子动力学模拟软件 DeePMD-kit .....	49
3.6.8 中国科学院格点量子色动力学 CLQCD .....	49
3.6.9 中国科学院海洋环流模式 LICOM .....	49
3.6.10 中国空气动力研究与发展中心 NNW 系列软件 .....	50
3.6.11 国防科技大学银河矩阵计算软件 YHMatrix .....	54
3.6.12 国防科技大学不可压缩流体计算软件 YHACT.....	55
3.6.13 国防科技大学银河通用网格生成软件 YHGRID .....	55
3.6.14 西安电子科技大学计算电磁软件 LASPCEM .....	56
3.6.15 自然资源部第一海洋研究所海浪模式 MASNUM .....	56
3.6.16 航空工业西安航空计算技术研究所外流仿真分析软件 WiseCFD-UG .....	57
3.6.17 中国核动力研究设计院 MOI-X 与 MORE .....	58
3.6.18 西北工业大学航发仿真应用计算软件 AES .....	59
3.6.19 北京理工大学全波数值电磁仿真软件 SINOCOM .....	60

<b>4 用户篇</b>	61
4.1 上海交通大学	62
4.2 兰州大学	63
4.3 中国科学技术大学	63
4.4 国家超级计算天津中心	65
4.5 国家超级计算深圳中心	68
4.6 国家超级计算长沙中心	71
4.7 国家超级计算广州中心	77
4.8 国家超级计算济南中心	82
4.9 武汉超算中心	83
4.10 广东气象局	84
4.11 深圳华大生命科学研究院	85
4.12 飞腾信息技术有限公司	86
4.13 彭峰（北京）科技有限公司	87
4.14 中国移动通信集团浙江有限公司	89
4.15 英特工程仿真技术（大连）有限公司	91
4.16 深圳十沣科技有限公司	94
4.17 湖南迈曦软件有限责任公司	96
4.18 北京龙讯旷腾科技有限公司	97
4.19 北太振寰（重庆）科技有限公司	98
<b>5 结语</b>	100

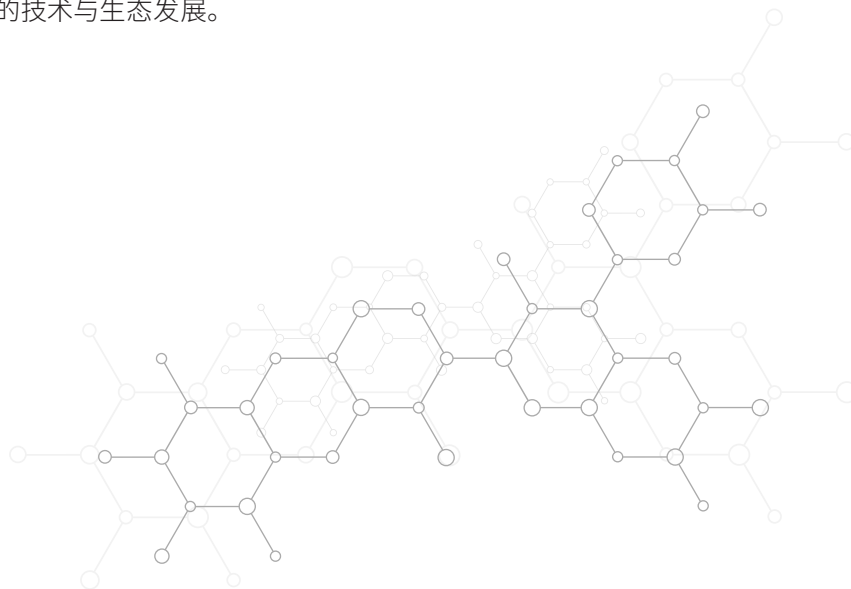


# 1 引言

HPC已进入E级计算时代。E级算力时代有个显著特征，包括美日中欧等HPC算力高地都开始选择ARM指令集兼容芯片作为自己E级机的主要技术路线。包括美国在内的各国陆续推出多款ARM指令集兼容的芯片产品，用于超算、商用HPC系统甚至HPC Cloud云服务。这是个巨大的变化，HPC计算芯片不再由一种或几种完全封闭的指令集所统治。这些ARM指令集兼容的产品形成一种有趣的现象，就像HPC领域的5G标准，参与者遵循相同的标准，但又有机会做出差异化的实现。这种指令集兼容、差异化实现、可自由演进的自主架构被称之为“开放架构”。这个变化不是一日之功，背后是“开放架构”厂家、生态及产业链伙伴持续多年的战略定位、技术与投资而产生的结果。

“开放架构HPC技术与生态白皮书”，旨在介绍ARM指令集兼容架构的技术发展、演进与成果；并通过HPC生态特点分析，介绍ARM指令集兼容架构的生态完备度、易用性特征。最后介绍产业界伙伴，包括系统建设者、系统使用者与生态构建者的具体案例。

希望本白皮书的推出能够形成HPC产业界上下游的伙伴、用户对“开放架构HPC”技术与生态的认知与共识，共同推动和建设相关的技术与生态发展。







开 放 架 构 H P C 技 术 与 生 态 白 皮 书

技术篇

TECHNOLOGY

## 2 技术篇

ARM 指令集兼容架构已成为 HPC 主流技术与未来发展的重要趋势，可满足大型超算系统与商用 HPC 系统的技术需求。

### 2.1 从国家超算战略看 ARM 指令集兼容架构在 HPC 领域的发展

中欧美日等算力高地都通过国家战略牵引各自的超算建设。如中国的科技部高性能计算专项、美国 ECP 计划、欧洲 EuroHPC 计划、日本 FLAGSHIP 2020 计划进行牵引。其中中国天河新一代超算依托 ARM 指令集兼容芯片进行建设，欧洲 EuroHPC 计划中设置 ECP 计划开发欧洲独立自主的 ARM 指令集兼容架构 CPU，日本富士通采用 ARM 指令集兼容架构开发 A64FX 芯片，打造“富岳”超算，美国将采用基于 ARM 指令集兼容架构研发的 NVIDIA Grace CPU 构建一台名为 Venado 的超级计算机。

各个国家超算战略的技术选择，为 ARM 指令集兼容架构这种开放架构技术演进与生态发展提供了重要支撑。

### 2.2 从 TOP 500 看 ARM 指令集兼容架构在 HPC 领域的发展

2022 年 11 月，全球超级计算机 TOP500 排行榜中，共有 5 台基于 ARM 指令集兼容架构处理器构建的超级计算机入围。同时，美国、日本、欧洲也都发布了多台基于 ARM 指令集兼容架构处理器的超级计算机建设计划，ARM 指令集兼容架构正在成为未来 HPC 的主流技术和发展趋势。

从表 2-1 中可以看出，日本在基于 ARM 指令集兼容架构处理器的超级计算机建设方面走在了世界前列：入围的 5 台超级计算机中有 4 台为日本建造，都采用了富士通（Fujitsu）基于 ARM 指令集兼容架构的 A64FX CPU，其中“Fugaku”超级计算机更是夺得了 2020 年 6 月发布的全球超级计算机 TOP500 的榜首。美国在 2018 年也推出了 Astra 超级计算机，该超级计算机采用 Marvell 推出的基于 ARM 指令集兼容架构的 ThunderX2 处理器，整个系统峰值性能也达到了 2.3PFlops。基于 ARM 指令集兼容架构处理器的超级计算机进入全球超级计算机 TOP500 排行榜，在很大程度上彰显了 ARM 指令集兼容架构在高性能计算领域的潜力。

在 TOP500 榜单之外，由 GW4 联盟、Cray 公司、英国气象局在 2018 推出的欧洲第一台基于 ARM 指令集兼容架构处理器的超级计算机：“Isambard”超级计算机；中国天河新一代超级计算机同样采用了基于 ARM 指令集兼容架构的异构融合处理器。这充分说明 ARM 指令集兼容架构在未来超级计算机领域必将占领一席之地。



表 2-1 2022-11 全球高性能计算机 TOP500 排行榜中基于 ARM 指令集兼容架构处理器的超级计算机

超算名称	SC2022 TOP500 排名	峰值性能	CPU 数量 / 核数	处理器型号	处理器架构	部署地	部署 年份
Fugaku	2	537.21 PFlop/s	158976/48	A64FX 48C 2.2GHz	Armv8.2-A SVE 512 位	日本理研计 算科学中心	2020
Wisteria	23	25.95 PFlop/s	7680/48	A64FX 48C 2.2GHz	Armv8.2-A SVE 512 位	日本东京大 学信息技术 中心	2021
TOKI-SORA	39	19.46 PFlop/s	5760/48	A64FX 48C 2.2GHz	Armv8.2-A SVE 512 位	日本宇宙航 空工业振兴 机构	2020
Flow	87	7.79 PFlop/s	2304/48	A64FX 48C 2.2GHz	Armv8.2-A SVE 512 位	日本名古屋 大学信息技 术中心	2020
Astra	467	2.30 PFlop/s	3990/36	Marvell ThunderX2 CN9975-2000 28C 2GHz	Armv8.1	美国桑迪亚 国家实验室	2018

目前采用 ARM 指令集兼容架构的超级计算机介绍如下：

### 1) “Fugaku” 超级计算机

2020 年 6 月 22 日，新一期全球超级计算机 TOP 500 榜单公布，日本 Fugaku（富岳）超级计算机夺得榜首。Fugaku 超级计算机采用了 Fujitsu（富士通）基于 ARM 指令集兼容架构的 A64FX CPU，整个系统理论峰值达到 537.21 PFlop/s，这是基于 ARM 指令集兼容架构处理器的超级计算机首次登顶全球高性能计算机 TOP500 榜单。Fugaku 超级计算机部署在日本神户理研计算科学中心，主要应用于药物研发、预防医学、模拟自然灾害、天气预报、清洁能源发展，开发 AI 模型等领域。

### 2) “Wisteria” 超级计算机

Wisteria 超级计算机是由 Fujitsu（富士通）公司研发、部署在日本东京大学的一台集成 HPC 和 AI 的超级计算机，主要用于仿真、大数据、人工智能等领域。Wisteria 采用节点间异构形式，包括两个部分：Odyssey（仿真节点组）和 Aquarius（数据分析节点），其中 Odyssey 包含了 7680 个节点，每个节点包括 1 路基于 ARM 指令集兼容架构的 48 核 A64FX CPU，集群总峰值性能为 25.9 PFlops，主要负责仿真任务；Aquarius 为 GPU 计算集群，双精度峰值总性能达 7.2 PFlops，主要负责数据分析与人工智能。

### 3) “Flow” 超级计算机

Flow 超级计算机采用了与 Fugaku 相同的 CPU，部署在名古屋大学的信息技术中心，峰值性能达到 7.79PFlops，主要应用于新冠肺炎解药的开发和解析、超级台风等异常气候检测、以及材料研究等领域。

### 4) “Astra” 超级计算机

Astra 超级计算机由惠普公司打造，部署在美国新墨西哥州阿尔伯克基的桑迪亚国家实验室。Astra 超级计算机采用 Cavium 推出的基于 ARM 指令集兼容架构的 ThunderX2 处理器，每颗处理器拥有 36 个核心，全系统共 2592 个计算节点，峰值计算性能可达到 2.3 petaflops，主要用于核武器安全性和可靠性方面的研究。

### 5) “Isambard” 超级计算机

Isambard 由 GW4 联盟、Cray 公司、英国气象局于 2018 年联合建造，是欧洲第一台基于 ARM 指令集兼容架构处理器的超级计算机，采用 Marvell ThunderX2 处理器，包含 10000 个 ARM 指令集兼容架构的处理器核心，理论峰值为 172TFlop/s。其升级版 Isambard 2 超级计算机将其规模扩展了两倍，处理器核心数达到了 21504 个。Isambard 部署在英国的布里斯托尔大学，是英国新的 HPC 服务的基础，

作为支持科学研究和评估的高性能计算技术的平台。

同时，美国和欧洲还有多台基于 ARM 指令集兼容架构处理器的超级计算机建造计划，主要如下：

### 1) “Mont-Blanc” 超级计算机

Mont-Blanc 项目旨在为欧洲下一代百亿亿次超级计算机进行验证，将采用 Cavium 的 Thudner X2 64 位 ARM 指令集兼容架构处理器进行建造，部署地以及算力等具体参数尚未公布。现阶段 Mont-Blanc 2020 演示器已经完成，对下一阶段百亿亿次级完整超级计算系统的研制有重要参考作用。

### 2) “阿尔卑斯山” 超级计算机

2021 年 4 月 12 日，瑞士国家超级计算中心、惠普和 NVIDIA 宣布，他们正在基于 ARM 指令集兼容架构处理器架构打造全球最强大的 AI 超级计算机：“阿尔卑斯山” 超级计算机，该超级计算机计划于 2023 年上线，将取代瑞士国家超级计算中心现有的 Piz Daint 超级计算机。“阿尔卑斯山” 将使用全新基于 ARM 指令集兼容架构的 NVIDIA Grace CPU，将应用在包括气候和天气，材料科学，天体物理学，计算流体动力学，生命科学，分子动力学，量子化学和粒子物理学，以及经济学和社会科学等领域。

### 3) “Venado” 超级计算机

NVIDIA 将使用 Grace CPU 和 Grace Hopper GPU 为美国洛斯阿拉莫斯国家实验室（Los Alamos National Laboratory, LANL）构建一台名为 Venado 的超级计算机，预计 AI 峰值性能将超过 10 ExaFlops。Venado 将成为美国第一个采用英伟达 Grace CPU 的超级计算机，将主要应用于材料科学和可再生能源等领域。NVIDIA Grace CPU 基于最新的 ARM V9 架构，专为 AI、HPC、云计算和超大规模应用而设计。

## 2.3 ARM 指令集兼容架构在商用 HPC 领域的崛起

除在超级计算机领域发展的如火如荼之外，ARM 指令集兼容架构在商用 HPC 领域也异军突起，基于 ARM 指令集兼容架构处理器的商用 HPC 系统不断推出，并在不同行业得到成功应用。

国内基于 ARM 指令集兼容架构的服务器芯片厂商主要有三家：华为、飞腾以及阿里平头哥。

**基于华为鲲鹏 920 CPU 的 TaiShan 服务器** 鲲鹏 920 CPU 是华为于 2019 年初推出的一款基于 ARM 指令集兼容架构的高性能处理器，具有高性能、高吞吐、高集成和高效能等特点。基于鲲鹏 920 CPU 的 TaiShan 系列服务器，推出高性能型、均衡型、存储型和高密型等不同机型，分别面向高性能计

算、大数据、分布式存储和 ARM 原生应用等场景，能够充分发挥 ARM 指令集兼容架构在多核、高效等方面的优势，为企业构建高性能、低功耗的新型计算平台。目前 TaiShan 系列服务器已经应用在制造、交通、金融等不同领域，全球已经有 11 家整机厂商基于鲲鹏 920 CPU 推出自主品牌服务器。

**基于飞腾系列 CPU 的商用服务器** 飞腾系列 CPU 是基于 ARM 指令集兼容架构设计的处理器，共推出高性能服务器 CPU、高效能桌面 CPU 和高端嵌入式 CPU 等多个系列，其中服务器芯片主要有飞腾腾云 S2500、FT-2000+/64 和 FT-1500A/16 三款不同型号。截至 2022 年 9 月底，飞腾的生态伙伴数量已经突破 4600 家，包括集成商合作伙伴 300 余家、硬件合作伙伴 984 家、软件合作伙伴 3413 家。同时飞腾已联合千余家国内软硬件厂商，支撑了 2600 多款飞腾平台设备上市，已经和正在适配的软件和外设超过 22600 款，分布在操作系统、应用软件、安全、云产品、数据库、中间件等各个领域。

**基于倚天 710 CPU 的商用服务器** 2021 年 10 月 19 日云栖大会，阿里巴巴旗下半导体公司平头哥发布自研高性能倚天 710 CPU。倚天 710 CPU 采用最新的 ARMv9 架构，是阿里首款全栈自研的通用 CPU，也是全球首款采用 5nm 工艺的服务器芯片。目前，基于倚天 710 CPU 的商用服务器已部署于阿里云数据中心并对外提供服务。

国外基于 ARM 指令集兼容架构的服务器芯片厂商主要有五家：NVIDIA、MARVELL、Ampere、亚马逊和富士通。

**基于 NVIDIA Grace 超级芯片的商用服务器** NVIDIA Grace CPU 基于最新的 ARMv9 架构，专为 AI、HPC、云计算和超大规模应用而设计，部署的 144 个核心和 1TB/s 的内存带宽将为基于 CPU 的高性能计算应用提供前所未有的性能。同时，基于 NVIDIA NVLink-C2C 技术的 NVIDIA Grace Hopper 超级芯片将 Grace CPU 和新一代 Hopper GPU 高速相连，并能够为加速 AI 和 HPC 计算提供 CPU+GPU 的一致内存模型。更为重要的是，Grace CPU 可以运行所有的 NVIDIA 计算软件栈。

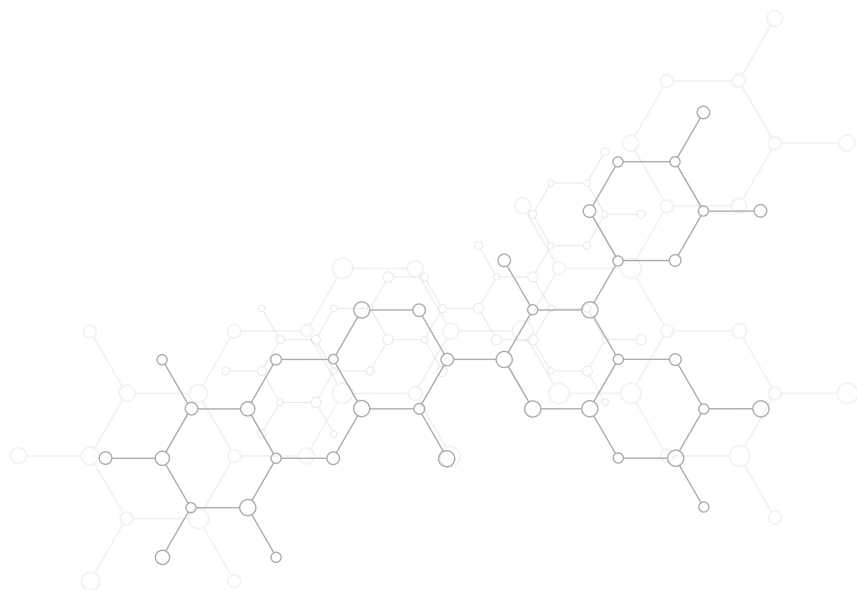
**基于 MARVELL ThunderX 系列 CPU 的商用服务器** Marvell ThunderX 系列 CPU 是基于 ARM 指令集兼容架构设计的处理器，先后推出 ThunderX1、ThunderX2 和 ThunderX3 三款高性能服务器 CPU。ThunderX 系列 CPU 可应用于高性能计算、云计算、边缘计算等诸多领域，满足计算、存储、联网工作负载的最高要求，并通过提供更高级别的计算性能和生态系统，加速 ARM 服务器在主流云、高性能计算数据中心内的采用和部署。ThunderX 系列 CPU 应用于超算领域外，在商业领域也应用广泛，包括微软 Azure 富士康、技嘉、和硕、惠普和 Atos 等公司均使用该系列服务器芯片。



**基于 Ampere 系列 CPU 的商用服务器** Ampere Computing（安晟培半导体）是英特尔前总裁 Renee James 创立的服务器芯片独角兽公司。安晟培半导体先后发布了多款基于 ARM 指令集兼容架构的服务器芯片，如 80 核 Ampere Altra CPU、128 核 Ampere Altra Max CPU、以及新推出的以 ARM 指令集兼容架构为基础的 Ampere One CPU。目前 Azure、甲骨文云、腾讯云、阿里云、优刻得 UCloud、京东、Equinix、CloudFlare、字节跳动等多个超大规模数据中心正在部署 Ampere Computing 产品。

**基于亚马逊（AWS）的 Graviton 系列服务器** ARM 与亚马逊（AWS）合作，在亚马逊云中成功应用基于 ARM 指令集兼容架构的 CPU 芯片。2018 年，AWS 首次推出基于 ARM 指令集兼容架构的自研服务器芯片 Amazon Graviton，将部分工作负载能耗降低至 45%；2019 年底推出 Graviton2，将针对各种工作负载又提升了 40% 的性价比。接着，AWS 在 2021 年 12 月推出了 Graviton3，在相同性能下，Graviton3 处理器与同类型基于 x86 的实例对比，可节省高达 60% 的能耗。同时 AWS 还发布了采用了 Graviton 3 的 C7g 应用实例，成为业界首款采用 ARM Neoverse 架构并支援 DDR5 的云端应用实例。鉴于其优异的表现，AWS 揽入了 Domo、F1 方程式赛车、Snap 等一干重量级客户。

**富士通的 HPC 商业化** 富士通制造的超级计算机 Fugaku 是现在世界上最快的基于 ARM 指令集兼容架构处理器的超级计算机，同时富士通在基于 ARM 指令集兼容架构的 HPC 商业化方面也很出色。2020 年 2 月 3 日，富士通为名古屋大学信息技术中心打造的超级计算机将成为世界上第一台采用富岳技术的商用超级计算机。2022 年 4 月 13 日，冰岛初创公司 Atmonia ehf. 与富士通合作，利用高性能计算（HPC）和人工智能技术，开展联合研究，加速用于清洁生产氨的催化剂开发。2021 年 2 月 11 日，富士通被选为主承包商为欧盟和葡萄牙政府提供高性能计算千亿级超级计算机的采购、交付、安装以及硬件和软件维护。2020 年 7 月 15 日德国雷根斯堡大学为了扩展他们对宇宙起源的探索，使用富士通新型超级计算机 PRIMEHPC FX700 和基于 ARM 指令集兼容架构的新型富士通 A64FX 处理器进行 QPACE4 的建造，为后续研究提供支持。2020 年 09 月 23 日，富士通为佳能提供新的超级计算机系统，将在促进佳能“无原型”产品开发计划方面发挥关键作用，在佳能产品开发过程中提供增强的功能和分析适用范围。



## 2.4 ARM 指令集兼容架构针对 HPC 的改进及演化

在 HPC 领域，X86 CPU 作为主处理器一直“一统江湖”，虽然期间受到 Power、MIPS 以及国内申威架构的挑战，但谁都无法撼动 X86 架构在 HPC 领域的统治地位。近年来，随着移动领域的霸主 ARM 架构处理器向服务器芯片市场进行拓展，特别是 2020 年搭载基于 ARM 指令集兼容架构处理器的日本 Fugak 超级计算机荣登全球超算 TOP500 榜首，以及中国、美国、欧洲多台基于 ARM 指令集兼容架构处理器的超级计算机的建成，ARM 指令集兼容架构在高性能计算领域异军突起，形成了对 X86 架构的持续性挑战。

### 2.4.1 从终端走向 HPC

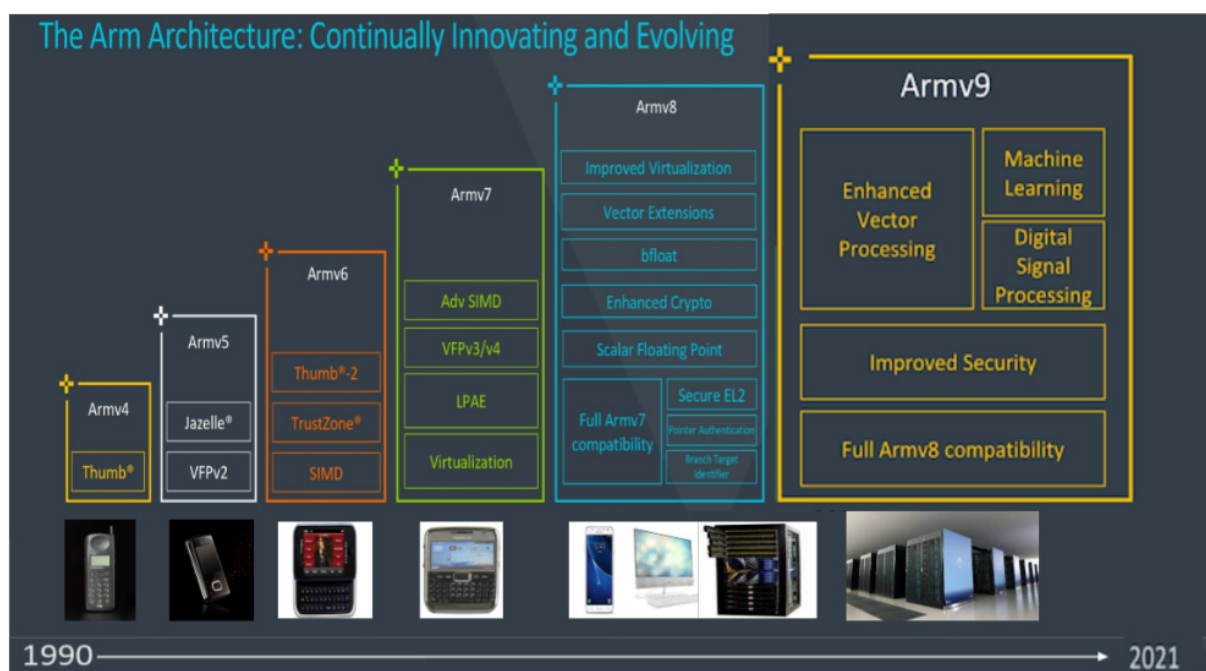


图 2-1 ARM 架构从终端到服务器的发展

传统上，ARM 指令集兼容架构处理器主要用于智能手机、终端等移动设备，苹果的 M1 和 M2 处理器的推出使得 ARM 指令集兼容架构处理器开始进入 PC 市场。随着 ARMv8 和 ARMv9 架构的发布，ARM 架构无论从体系结构设计，还是从指令集的丰富程度上，已经完全可以适应 HPC 领域的工作负载，ARM 指令集兼容架构处理器开始在 HPC 领域异军突起，真正展示出了与 x86 争夺 HPC 市场的能力。

总体来看，ARM 指令集兼容架构进入高性能计算领域分为三个阶段：

**第一阶段“启程”** 早在 2008 年，ARM 公司内部就开始酝酿进军服务器市场的计划，并参与投资

了名为 Calxeda 的初创公司。Calxeda 公司的目标是降低能耗以及提升相同空间的计算密度，并在 2011 年发布了基于 ARM A9 架构的芯片 EnergyCore ECX-1000 CPU，并由 30 块集成 4 颗芯片的板卡组成了 490 个核的服务器系统。不过很可惜，Calxeda 公司倒在了第三轮融资的路上，于 2013 年破产倒闭。Calxeda 成为 ARM 指令集兼容架构在服务器领域第一个“吃螃蟹的人”，也吹响了 ARM 指令集兼容架构向服务器领域进军的号角。除 Calxeda 公司外，在这一阶段基于 ARMv7 架构的 Marvell 四核系列处理器在 ARM 指令集兼容架构进军服务器领域留下了浓重的一笔，Dell 公司以此为核心推出的“Copper”服务器系统在百度部署过，成为 ARM 指令集兼容架构服务器在互联网领域的第一个案例。ARM 指令集兼容架构进军服务器领域的第一阶段“始于 Calxeda，终于 Calxeda”，32bit 的服务器和软件生态的不完善成为 ARM 指令集兼容架构走向 HPC 领域的绊脚石。但确实如 Calxed 最后的结业邮件中提到的，因为基于 ARM 指令集兼容架构服务器的出现，“the industry will be transformed forever”。

**第二阶段 “觉醒”** 2011 年 10 月，随着 ARM 公司第一款 64 位 ARMv8 架构的发布，标志这 ARM 在 HPC 领域的“觉醒”。相比 ARMv7，ARMv8 是一个巨大的跨越：64 位通用计算器、128 位向量寄存器、专为高性能计算新开发的向量指令集 NEON，使 ARM 指令集兼容架构芯片有了更为广泛的应用场景，可以说 ARMv8 让 ARM 指令集兼容架构芯片从嵌入式或移动设备，进入了高性能计算领域。

特性	ARMV8	ARMV7
指令集	64位指令集 AArch64，并且兼容32位指令集 AArch32	32位指令集 A32 和16位指令集 T16
支持地址长度	64位	32位
通用寄存器	31个 x0-x30 (64位) 或者 w0-w30 (32位)	15个, r0-r14 (32位)
异常模式	4层结构 EL0-EL3	2层结构vector table
NEON	默认支持	可选支持
LAPE	默认支持	可选支持
Virtualization	默认支持	可选支持
big.LITTLE	支持	支持
TrustZone	默认支持	默认支持
SIMD寄存器	32个 X 128位	32个 X 64位

图 2-2 ARMv8 架构与 ARMv7 架构的差别

从这一阶段开始，采用 ARM 指令集兼容架构的服务器芯片开始对标主流服务器芯片性能。Marvell 的 ThunderX 和 ThunderX2 CPU 是这一阶段的代表性芯片，这两款芯片已经具备了服务器芯片的主要特征：超多核设计（ThunderX CPU 48 个核，ThunderX2 CPU 32 个核）、高性能、高吞吐，处理器整体性能已经可以对标 X86 架构。美国更是在桑迪亚国家实验室建设了一台基于 ThunderX2 CPU 的 Astra 超级计算机，峰值性能达到 2.3PFlops，ARM 指令集兼容架构服务器芯片开始在超算领域初露峥嵘。同时，我国的华为和飞腾也先后推出了鲲鹏 916 和飞腾的 1500A 两款基于 ARM 指令集兼容架构的处理器，也成为这一时期的代表性芯片。

**第三阶段 “成熟”** 2019年，ARM公司推出了2个瞄准高性能计算的全新平台——Neoverse N1和Neoverse E1，两者都是基于ARMv8.2-A架构。其中，Neoverse N1专为服务器和基础设施市场设计，Neoverse E1专为高吞吐量运算工作负载设计。随后，ARM公司推出两个新的升级平台：ARM Neoverse V1和Neoverse N2，其中，V1是一个全新的产品，基于ARMv8.4-A架构，适用于繁重的HPC、云端运算及AI/ML辅助工作负载；N2是N1的升级版，是首个基于ARMv9架构且支持SVE的处理器IP，可为多样工作负载提供40%的性能提升。

Neoverse架构的发布，标志这ARM架构在服务器领域的成熟。这一阶段基于ARM指令集兼容架构的服务器芯片百花齐放，国内如华为鲲鹏920、飞腾、阿里倚天710等CPU；国外如Marvel的ThunderX3、富士通的A64FX、Ampere Altra，AWS Graviton2等CPU。除此之外，SiPearl、NUVIA等一批初创企业的加入也为ARM指令集兼容架构服务器芯片市场带来了新的助力。总之，随着硬件架构的逐渐成熟和软件生态的日益完善，ARM指令集兼容架构服务器芯片终于从这一阶段开始在超级计算机和商用HPC领域展露锋芒，在竞争日益激烈的服务器市场最终占据了一席之地。

## 2.4.2 面向 HPC 的 ARM 指令集兼容架构演进

ARMv8架构是ARM公司在2011年推出的第一款64位架构，相比ARMv7架构是一个巨大的跨越。ARMv8架构为高性能计算而新开发的向量指令集NEON，使ARMv8架构可用于服务器和超算领域。同时，针对高性能计算、数据中心的任务负载情况，ARM公司设计和引入了新的指令集：可伸缩向量扩展(Scalable Vector Extensions/SVE)指令集，作为ARMv8架构的可选扩展指令集，SVE支持最低128-bit、最高2048-bit的向量化处理。

2021年，ARM公司推出的全新ARMv9架构，针对高性能计算的计算负载特征进行了一系列改进，ARMv9相比于ARMv8最主要的区别有三点：AI、改进向量扩展指令集SVE2以及安全。机器学习成为ARMv9架构的重要组成部分，ARM公司认为未来ML工作负载将会越来越普遍，同时，矩阵乘法指令成为ARMv9架构的基础功能，也标志着ARM架构对AI工作负载支持的强化。SVE2指令集一方面完善了SVE指令集的不足，另一方面基本做到了原来Neon指令集所要做的事情，而且有着更高的效率和更好的灵活性，可变长度让其不但适应使用长向量寄存器的工作，还能胜任使用短向量寄存器的任务。安全是ARMv9架构的重要特征，ARM公司希望通过引入机密计算体系结构来重新设计安全应用程序的工作方式，以应对越来越严峻的安全问题。这也是ARM架构针对服务器领域的一次重大改进。

## 2.4.3 国内 ARM 指令集兼容架构处理器针对 HPC 领域的进展

ARM公司对其架构的授权方式有如下几种：



- 处理器优化包 / 物理 IP 包授权（硬核）：按照 ARM 公司设计好的处理器规格、在指定的代工厂和工艺进行生产。
- 内核层级授权：用户可以将 IP 核应用到其自行设计的芯片中，但用户不能对 IP 核进行修改。拥有内核级授权的厂商多如牛毛，包括德州仪器、博通、以及富士通等。
- 处理器核授权（软核）：拥有 RTL 代码，可进行代码级的单元测试，可以自行完成逻辑设计和物理设计。
- 架构层级授权：是指用户拥有处理器架构级的设计和制造许可，用户甚至可以从自己的具体需求出发，对 ARM 架构进行大幅度改造或者对指令集进行扩展或缩减，以便达到更高性能、更低功耗或更低成本的目标。拥有架构授权的典型厂商包括高通、苹果、微软等。中国的华为、飞腾等芯片厂商同样拥有 ARM 公司的架构级授权，这是高度自主研发服务器芯片的设计根基。目前，基于 ARM 指令集兼容架构的服务器 CPU 在国内应用广泛，已经成为在自主创新、安全可控、产业生态等领域的最佳选择。

国内采用 ARM 指令集兼容架构处理器的服务器，目前大体分为华为的鲲鹏系列、中国电子的飞腾系列和阿里的倚天系列：鲲鹏系列主要是以华为 TaiShan 系列服务器为核心、飞腾系聚焦于国产芯片研发，直接向服务器合作厂商出售服务器芯片、阿里倚天系列目前主要满足其自己数据中心的使用。

飞腾是国内第一家获得 ARMv8 架构级永久授权的公司，并以 2014 年基于 ARM 指令集兼容架构 V8 架构的 16 核 FT-1500A/16 CPU 问世为标志，开启了飞腾 CPU 快速发展的新篇章，标志着飞腾基于 ARM 指令集兼容架构的 CPU 进入到 HPC 领域。随后，飞腾分别于 2016 年推出第一款 64 核处理器芯片 FT-2000/64、2017 年推出第二代 64 核处理器芯片 FT-2000/64、以及随后推出的 S2500 服务器级 CPU。同时，在生态合作方面，飞腾迅猛暴发，从 OEM 厂商到硬件产品设计，再到软件兼容认证，飞腾在服务器领域形成了全栈解决方案和生态图谱。

华为自研的鲲鹏 920 芯片是基于华为自研 TaiShan 核的高性能处理器，整个芯片代码约 300 多万行，相关代码全部都是华为公司自研，是华为公司多年技术积累的结果，华为拥有完全的知识产权。

## 2.5 ARM 指令集兼容架构在 HPC 的未来发展趋势

**1. 单核性能更高，超多核 / 众核趋势明显** 为了应对 HPC 领域越来越高的工作负载，ARM 指令集兼容架构服务器芯片频率越来越高，单核性能越来越高，逐渐可以对标 X86 单核性能。除此之外，从富士通 A64FX CPU 的 48 核，到鲲鹏 920 CPU 的 64 核，再到 Ampere Altra Max CPU 的 128 核，ARM 指令集兼容架构服务器芯片的超多核 / 众核的趋势越来越明显，以实现在相同体积内能够部署更高的算力。

**2. 更完善的向量化指令集** 针对 HPC 领域工作负载的多样性，SVE2 指令集更加完善，已经可以初步对标 X86 AVX512 指令集；同时，SVE2 可变长向量功能，能够适应更加灵活的数据访问和计算。除了对通用指令继续完善外，还添加了对矩阵乘指令的支持，加强了对 AI 工作负载的支持。

**3. 更高的访存性能** 使用 HBM2e/HBM3 高带宽内存，提高访存性能，提高对数据密集型应用的支持。包括华为鲲鹏 920 在内的多款 ARM 指令集兼容架构 CPU 的访存带宽已经整体上超越了相应主流 CPU。

**4. 更安全** 提高安全性是 ARM 指令集兼容架构的一个重要特征，为此，许多 ARM 指令集兼容架构处理器特意增加了专门的安全模块，以此来应对越来越严峻的安全威胁。这是 ARM 指令集兼容架构进军服务器和 HPC 市场的一个重要进步。

**5. 自研芯片会越来越多，ARM 指令集兼容架构服务器芯片越来越“百花齐放”** ARM 指令集兼容架构芯片强大的定制能力，使得越来越多的芯片厂商在授权的 ARM IP 基础上，开始根据市场的判断对芯片进行重新设计和改进，以适应不同的应用场景。

**6. “终端”用户“下海”** 除传统芯片厂商外，如亚马逊、阿里巴巴等终端用户也开始研制 ARM 指令集兼容架构服务器芯片，部署在自己的大数据中心或者云服务中心。

**7. 应用迁移难度不断降低** ARM 指令集兼容架构在 HPC 领域的产业链成员越来越丰富，大量的开发者不断参与到 ARM 指令集兼容架构 HPC 生态的构建中，而随着生态的不断完善和技术的不断成熟，典型应用到 ARM 指令集兼容架构 HPC 上的迁移难度越来越低，这会有力推动 ARM 指令集兼容架构在 HPC 领域的发展。同时，越来越多的 ARM 指令集兼容芯片加入了“二进制转码”功能，例如苹果公司等。这类技术可以动态的进行其他指令集到 ARM 指令集的翻译，在无需进行代码重新迁移的情况下，直接运行其他指令集编译的软件。

开 放 架 构 H P C 技 术 与 生 态 白 皮 书

生态篇  
ECOLOGY



## 3 生态篇

ARM 指令集兼容架构的生态已具规模。其中基础软件层完备，应用软件层开源生态快速发展，商业软件生态获得突破，具备大规模进入生产系统的条件。

### 3.1 ARM 指令集兼容架构面向 HPC 领域的生态构成

ARM 指令集兼容架构面向 HPC 领域的基础生态完整，在系统层、编译层、平台层和应用层具有完备易用的系统实现和工具链。

**(1) 系统层** ARM 指令集兼容架构处理器操作系统一般采用 Linux 操作系统。Linux 操作系统是开源软件，国内也有基于 Linux 开发的麒麟操作系统等，并对华为鲲鹏和飞腾芯片都适配得较好，在行业也得到广泛应用。国外，linux 系操作系统都提供了针对 ARM 指令集兼容架构的商用发行版。

**(2) 编译层** 编译层主要包括编译器、调测试工具和并行编程模型等。编译器通过基于 GCC 或者 LLVM 二次开发和优化得到。调测试工具主要针对具体芯片满足断点调测试和并行调测试的需求。并行编程模型主要包括 MPI、OpenMP、OpenCL 及自定义并行编程模型等。上述编译层的工具例如 GCC 和 MPI 等均为或具有开源实现。

**(3) 平台层** 平台层主要包括求解器、算法库和前后处理。求解器主要包括稠密和稀疏线性代数求解器，算法库包括领域算法库、BLAS、FFT、信号处理、图像处理、超越函数等。前处理主要包括网格生成和数据预处理等，后处理主要指数据可视化。

**(4) 应用层。**应用层主要包括具体应用软件，应用软件可能依赖若干求解器和算法库，也可能是独立的代码集合，分别开发不同行业的应用软件，如粒子物理、CFD 计算、信息安全、结构力学、电磁仿真、材料计算、分子动力学、生物信息、计算物理、计算化学、气候气象、金融分析、石油勘探、人工智能和大数据处理等。

### 3.2 操作系统生态

#### 3.2.1 概述

本篇仅介绍 ARM 指令集兼容架构上 Linux 操作系统生态，而忽略 Windows 操作系统生态，原因主要有两点：首先，虽然 Windows 操作系统在 PC 市场占据统治地位，但 HPC 领域主流仍为 Linux 操作



系统，世界 TOP500 超级计算机全部采用 Linux 操作系统；其次，微软在 2012 年首次推出基于 ARM 的 Windows 8 RT、2016 年的“Windows on ARM”、2018 年的 windows10 for ARM 以及最近的 Windows 11 均存在一定缺陷，包括不能运行 64 位程序、模拟运行等，对 ARM 处理器支持有限。

Linux 操作系统主要指 Linux 发行版（Linux distribution）是基于 Linux 内核构建的可安装操作系统，包含常用用户程序和相关库，Linux 发行版可分为由商业公司维护的商业版本和由开源社区维护的免费发行版本两大类。国外主流的商业版包括 Red Hat 和 SUSE，Red Hat 主要的市场是美国而 SUSE 主要面向欧洲用户，国内的主要商业版本包括银河麒麟和统信。相比而言，开源操作系统百花齐放，Linux 开源版本主要有相应社区维护，包括 CentOS、Fedora、Debian、Ubuntu、OpenEuler 等。需要指出，主流商业版包括 SUSE、Red Hat 和麒麟均存在相应的社区版本，例如 Fedora、OpenSUSE 和 OpenKylin，事实上这些开源社区已成为其对应的商业版本的上游社区，商业版吸收社区版本进展并开展性能、安全等方面的完整测试，出售专业技术服务。总体而言，几乎所有主流 Linux 商业版和开源版均支持 ARM 指令集兼容架构处理器，本篇介绍若干 ARM 指令集兼容架构上重要的 Linux 发行版。

### 3.2.2 OpenEuler

OpenEuler 是由开放原子开源基金会（OpenAtom Foundation）孵化及运营的开源项目，OpenEuler 是面向数字基础设施的开源操作系统。可广泛部署于服务器、云计算、边缘计算、嵌入式等各种形态设备，应用场景覆盖 IT（Information Technology）、CT（Communication Technology）和 OT（Operational Technology），实现了统一操作系统支持多设备，应用一次开发覆盖全场景。通过不断的技术创新与技术升级，OpenEuler 不仅具备高性能、高安全、易运维基础能力，还具备全场景协同领先优势。

### 3.2.3 银河麒麟

目前国内的 ARM 指令集兼容架构服务器都能够支持 HPC 麒麟操作系统，银河麒麟操作系统适应虚拟化、云计算、大数据、工业互联网时代对主机系统可靠性、安全性、性能、扩展性和实时性等需求，依据 CMMI5 级标准研制的提供内生本质安全、云原生支持、自主平台深入优化、高性能、易管理的新一代自主服务器操作系统，同源支持飞腾、鲲鹏等自主平台；应用于政府、金融、教育、财税、公安、审计、交通、医疗、制造等领域。基于银河麒麟高级服务器操作系统，用户可轻松构建数据中心、高可用集群和负载均衡集群、虚拟化应用服务、分布式文件系统等，并实现对虚拟数据中心的跨物理系统、虚拟机集群进行统一的监控和管理。

由麒麟软件等数十家基础软硬件企业、非营利性组织、社团组织、高等院校、科研机构和个人开

发者共同创立的开源社区 OpenKyling 在开源、自愿、平等和协作的基础上，通过开源、开放的社区合作构建系统开源社区，推动 Linux 开源技术及其软硬件生态繁荣发展。目前，OpenKyling 正在推动其 ARM 指令集兼容架构上的开放操作系统发布。

### 3.2.4 Red Hat

Red Hat 于 1995 年发布其第一个 Linux 发行版，也是最早使用 RPM 的操作系统。2003 年，Red Hat 发布其企业版 Red Hat Enterprise Linux (RHEL)，并将支持开源社区继续开发其家用版 Fedora Linux。RHEL 是一个开源的商业 Linux 发行版，将 Fedora 作为其上游。Red Hat 于 2015 年发布了一个预览版本支持 64 位 ARM 架构，主要用于硬件合作商研发验证。2017 年，Red Hat 在其企业版 Red Hat Enterprise Linux 7.4 中加入了对 ARM 架构的支持，其目标是构建一个单一的操作系统平台覆盖不同厂商基于 ARMv8 架构的 64 位服务器级处理器。与 x86 处理器不同，ARM 系统硬件设计与实现具有较强独立性，因此与 Linux 社区特征配合，而 RHEL 拥有包含上百家硬件厂商的庞大社区组织，可以即使获取相关软硬件程序，共同促进 ARM 新技术的迭代更新。

### 3.2.5 SUSE

SUSE Linux 是德国 SUSE Linux AG 公司发行维护的 Linux 发行版，1994 年年初发布第一个版本，2004 年 SUSE Linux AG 被 Novell 公司收购。SUSE Linux 企业版 (SUSE Linux Enterprise Server, SLES) 存在 ARM 版本支持 AArch64 架构 ARM 服务器并开展了相应的性能优化、企业级安全、高可靠性以及相应的技术支持。与其他 SUSE 一样，SUSE Linux 企业服务器 ARM 版本也支持免费下载，只是对服务及附件软件收费。

OpenSUSE 是在 Novell 收购 SUSE Linux AG 后创建的社区项目，目前是 SLES 的上游社区，即 SLES 来源于 OpenSUSE 但具有更长的发行期。OpenSUSE 存在两个版本，OpenSUSE Leap 是传统的发行版，在发行间隔期间提供安全更新与错误修复，并且与面向企业的 SUSE Linux Enterprise (SLE) 共享代码库；OpenSUSE Tumbleweed 则替代了传统的数字版本号且周期性更新的版本，使用滚动发布体系：软件更新持续不断地推出。OpenSUSE 与 SLE 的关系类似 Fedora 与 RHEL 关系。同样，OpenSUSE 也支持 AArch64 架构 ARM 处理器。

### 3.2.6 Ubuntu

Ubuntu 是由 Mark Shuttleworth 于 2004 年开发的基于 Debian 的发行版，与 Debian 哲学相反，Ubuntu 非常易于使用。Ubuntu 最新 22.04.1 版支持最新 64 位 ARM 架构服务器，支持超过 5 万种软

件和运行时，包括 Go，Java，Javascript，PHP，Python 和 Ruby，Ubuntu Server for ARM 提供服务器级别的性能，并保持高可靠性和与其他 Ubuntu 系统一样的使用体验。Ubuntu ARM 服务器版本支持 LXD 容器监视器，允许实时查看基础性能参数，在网络服务、数据库服务等方面也具有强大的支持和部署功能，此外，Ubuntu 支持企业级全天候服务。

### 3.2.7 Arch Linux

Arch Linux ARM 是 Arch Linux 的移植版，而 Arch linux 是一种以轻量简洁为设计理念的 Linux 发行版，基本理念是快速、轻巧、弹性与简单。其提供的安装盘只包含安装所需的基本系统，用户可以根据自己的特定需求选择配置安装相应的软件。Arch linux 采用滚动升级方式，提供多数软件的最新稳定版本。Arch Linux ARM 为各种消费类设备和开发平台上的软浮动 ARMv5te，硬浮动 ARMv6 和 ARMv7 以及 ARMv8 AArch64 指令集提供有针对性的内核和软件支持，旨在为最终用户提供简单性和完全控制。

### 3.2.8 Debian

Debian 由 Ian Murdock 于 1993 年创建，目前有稳定、测试和非稳定三个版本，长期支持的稳定版大约每两年更新一版并被支持 5 年。Debian 以精简、轻量、快速著称，广泛用于操作系统高级用户，在近 300 个 Linux 活跃发行版中，有近一半源自 Debian。Debian 提供了三种 ARM 移植，可以为各种不同的机器提供最佳支持，这是由于 ARM 架构是随着时间的推移而发展，现代 ARM 处理器提供了旧型号上不可用的功能。Debian/armel 针对旧的 32 位 ARM 处理器，而不支持硬件浮点单元（FPU）；Debian/armhf 仅适用于较新的 32 位 ARM 处理器，其至少实现了 ARMv7 架构，且支持 ARM 矢量浮点规范（VFPv3）第 3 版。此移植可利用这些型号上可用的扩展和性能增强功能；Debian/arm64 适用于 64 位 ARM 处理器，其至少实现了 ARMv8 架构。Debian 现在仅支持小尾端 ARM 系统，也自从 Debian 11 开始内核映像编译时加入 SMP 支持。

### 3.2.9 Fedora

Fedora Linux 是由开源社区维护的 Linux 发行版，为 Red Hat Enterprise Linux 的上游。自从 Fedora 35 后，Fedora 存在 6 个版本，PC、工作站、服务器、云计算、容器和物联网，Fedora 的特点是更新速度，其与各种新技术预计上游 Linux 社区保持紧密合作，每 6 个月更新一次版本。Fedora ARM 支持 ARMv7 (armhfp, aarch32, and arm-32) 和 aarch64 (ARMv8 and arm64)，覆盖的硬件范围包括服务器、工作站以及 IoT 等。对于 32 和 64 位 ARM 架构而言，Fedora 软件栈是一致的。目前 Fedora 已经停止对 32 位 ARM 架构的继续开发，并鼓励用户使用其 aarch64 版本。

## 3.3 编译器生态

### 3.3.1 ARM 官方编译器

ARM Allinea Studio 是一整套高性能工具，可用于开发面向 ARM 服务器的应用。ARM Compiler for Linux 22.1 编译器套件是其中的编译开发工具。该套件提供了一个完整的编译环境，为 ARMV8-A 指令集的最新体系结构功能和扩展提供了支持，用于在基于 ARM 的平台上进行应用开发与程序性能调优，特别针对基于 NEOVERSE V1 和 NEOVERSE N1 处理器的平台（如 Amazon Graviton 2 和 Ampere Altra）进行了优化。该开发套件主要包括 ARM C/C+/Fortran Compiler 22.1，ARM Performance Libraries 22.1.0 和 GCC 11.2.0 三部分。

ARM 编译器 22.1 是一个 Linux 用户空间 C/C++ 和 Fortran 编译器，为科学计算、HPC 等量身定制；该编译器是建立在开源的 Clang 前端和基于 LLVM13.0.1 的优化和代码生成后端上。ARM 编译器支持现代 C/C++、Fortran、OpenMP 4.0 和 OpenMP 4.5 标准，具有内置的自动向量化模块（SVE 和 NEON），并针对基于 ARM v8-A 的处理器进行了修改，支持对 SVE 和 SVE2 目标平台的编译。

ARM 性能库是一组专门为 ARM 处理器定制的数学库，它主要部署在高性能计算和云计算场景中，作为矢量和矩阵计算的性能解决方案；它主要针对密集数据型计算负载，提供了稀疏线性代数、FFT 和 libm 函数的解决方案。具体例程包括：BLAS，LAPACK，FFT，Sparse，libamath，libastring 等。

此外，ARM 官方还提供了服务于 HPC 的开发工具套件 Arm Forge Ultimate，它提供一套全面的并行调试和分析功能，还具备高阶指标和性能分析，包括性能报告生成和应用效率建议。该套件中包含的 Arm DDT 调试器可用于在 CPU 和加速器上开发 C++、C 或 Fortran 并行和线程应用。Arm MAP 是该套件的性能分析器，能够提供宝贵的优化建议。该工具开箱即用，能够适应于从单个进程到数万个进程的并行应用程序。Arm 性能报告是一款低开销工具，可生成单页文本和 HTML 报告，总结并描述标量和 MPI 应用性能。此工具有助于有效地描述和了解 HPC 应用的运行性能。

### 3.3.2 亚马逊编译器

亚马逊云科技 (AWS) 基于 ARM Neoverse 创建了 AWS Graviton 系列 CPU，包括 2019 年发布的基于 Neoverse N1 的 Graviton2 和 2021 年发布的基于 Neoverse V1 的 Graviton3。针对 Graviton 系列 CPU 的编译器是通过扩展 GCC 或 clang 开源编译器实现，并通过 -mcpu 选项指定适当的架构和优化。Graviton2 和 Graviton3 处理器支持在 ARMv8.1 中首次引入的大型系统扩展 (Large-System Extensions, LSE) 相关指令功能，能够提供低成本的原子操作，提高 CPU 到 CPU 间通信、锁和互斥锁的系统吞吐量。



### 3.3.3 Ampere 编译器

Ampere 公司当前一代的 Ampere Altra 和 Ampere Altra Max 处理器都采用 ARM Neoverse-N1 内核设计。腾讯云、UCloud，谷歌云等都推出了基于 Ampere Altra 的服务器。2022 年，Ampere 公司推出其下一代 AmperOne 处理器，不再使用现有的 ARM Neoverse 内核，转而采用完全定制的 ARM 内核。在编译器方面，2021 年 12 月公司发布了 Ampere GCC 10.3.1 compilers，目标是在 Ampere Aarch64 系统上构建高性能应用，并指出该编译器开始支持下一代 Ampere 处理器。2022 年，Ampere 公司在 llvm15.0 中添加了对“amper1”后端的初始编译器支持，该后端确认对 ARMv8.6-A 与 FP16 和 MTE（内存标签）扩展，以及启用投机障碍（SB）和投机存储旁路安全（SSBS）选项的支持。

### 3.3.4 富士通编译器

富士通公司的 A64FX 是世界上第一个实现可伸缩向量扩展（SVE）的处理器，是基于 ARMv8.2 指令集的扩展。富士通公司的超级计算机 Fugaku，PRIMEHPC FX700，FX1000，HPE 的 Apollo 80 都基于该处理器进行构建。

PRIMEHPC FX1000 安装有富士通公司开发的 HPC 中间件 FUJITSU Software Technical Computing Suite。

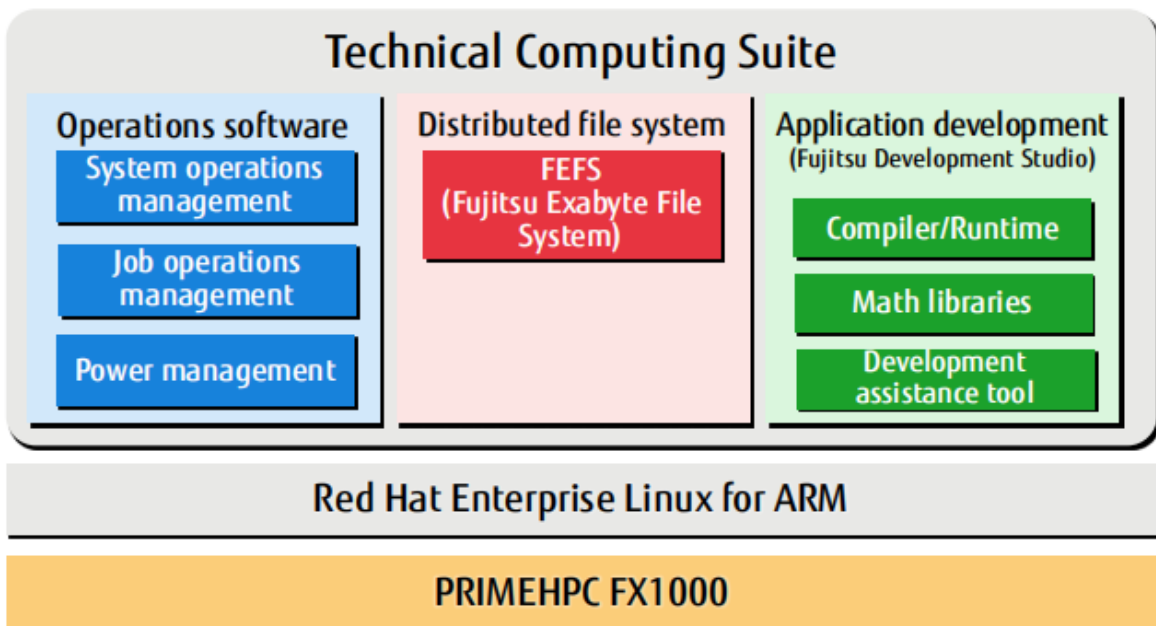


图 3-1 富士通 HPC 中间件

其中，Application development (Fujitsu Development Studio) 是涉及编译的集成软件套件，该套件用于编译、调试、调优和执行用 Fortran、C 或 C++ 编写的科学计算程序，并支持自动并行化、OpenMP、MPI 等并行化技术其中编译器和数学库提供对 ARM 的新 HPC 扩展、SVE 功能以及新的语言标准的指令支持。Development assistance tool 则是一个应用程序综合开发环境。其通过三级分析（基本分析器、详细分析器以及 CPU 性能分析报告）支持应用程序的分析，可以以图形化的方式显示从应用程序层收集到的信息，帮助用户了解程序运行状态和瓶颈。在调试方面，Development assistance tool 包含三类调试功能，即异常调查功能、死锁调查功能以及使用命令文件的调试功能，从而支持对大规模的并行处理中容易出现的异常、死锁等问题的调试。

### 3.3.5 Cray 编译器

英国气象局的 Isambard 超级计算机和美国桑迪亚实验室 (Sandia National Lab) 的 Astra 超级计算机都是基于 ThunderX2 处理器的系统。Isambard2 还增加了基于 A64FX 处理器的 Cray 系统机柜。支持的编译模块包括 GCC 编译器，ARM 编译器，富士通编译器以及 Cray 编译器 (The Cray Programming Environment for ARM)。

The Cray Programming Environment (CPE) 提供了诸如编译器、优化库、分析器和调试器等工具。Cray 编译器 (Cray Compiling Environment, CCE) 支持的语言包括 Fortran、C 和 C++，以及 UPC (Unified Parallel C) 等。CPE for ARM 包括两类编译器套件和相关模块，即支持 SVE 功能的 cce-sve 和不支持 SVE 功能的 cce。其中，cce-sve 必须与 craype-arm-nsp1 目标模块配对，以生成 ARM SVE 代码，支持对 A64FX 处理器的编译；而 cce 则可以与 craype-arm-thunderx2 配对，编译生成 ARM Neon 代码，支持对 ThunderX2 处理器的编译。优化库 (Cray Scientific and Math Libraries, CSML) 实现了包括稀疏和密集线性代数 (BLAS, LAPACK, ScaLAPACK) 以及快速傅里叶变换 (FFTW) 在内的高性能库。在 MPI 通信上，支持 OpenMPI 和 HPE Cray MPI，默认的 MPI 库为 CRAY MVAPICH，可以和 GCC 和 Cray 编译器一起使用。

分析器 (Cray Performance Measurement & Analysis Tools, CPMAT) 供了一组工具来分析运行程序的性能和行为，为计算、通信、I/O 和内存利用率的测量、分析和可视化提供了一个集成的基础设施，以帮助用户优化程序，以实现更快的执行和更高效的计算资源使用。调试器 (Cray Debugging Support Tools, CDST) 提供了包括 gdb4hpc、Valgrind4hpc、异常终止处理 (Abnormal Termination Processing, ATP) 以及 Cray 比较调试器 (Cray Comparative Debugger, CCDB) 在内的调试工具。

### 3.3.6 毕昇编译器

华为公司面向鲲鹏处理器开发了毕昇编译器。该编译器是基于开源 LLVM 10.0.1 版本开发，并进行了优化和改进，同时将 flang 作为默认的 Fortran 语言前端编译器，是一种 Linux 下针对鲲鹏 920 的高性能编译器，其针对鲲鹏平台进行了深度优化的高性能编译器。除支持 LLVM 通用功能和优化外之外，对编译中端及后端进行了深度优化，具体结构如图 X 所示。

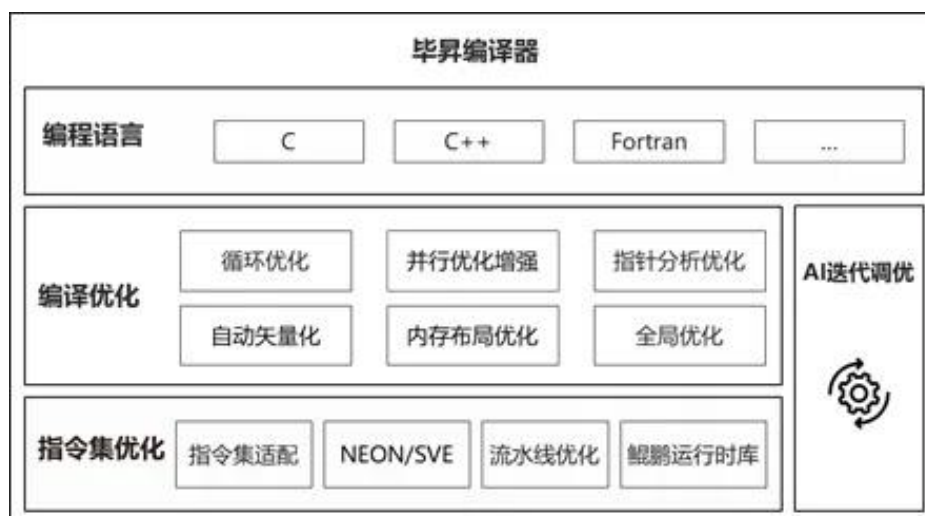


图 3-2 毕昇编译器架构图

毕昇编译器关键特性包括：支持鲲鹏微架构芯片及指令优化；通过软硬协同提供相较开源 LLVM 更高的性能；集成 Auto-tuner 特性支持编译器自动调优。内置的典型优化包括：结构体内存布局优化，大幅提升缓存命中率，突破访存瓶颈；结构体指针压缩优化，大幅降低内存使用，提升缓存命中率；软件预取，大幅提高程序性能，提升缓存命中率；自动向量化；循环优化；自研的基于 ML 的自动搜索技术 (ML-based Search) 的 Autotuner 工具，快速获取最优编译配置。

### 3.3.7 天鹰编译器

国防科大编译团队研制了一套面向国产 ARM 指令集兼容架构处理器的“天鹰编译系统”，包含基础编译、OpenMP、OpenCL 并行编译、数学库、安全编译等。该编译系统提供 C、C++、Fortran、Java 等基础语言在飞腾处理器平台上的编译和运行时环境，包含前端、中端优化及后端代码生成等模块，并为

各种编程语言提供调试和性能分析的支持；提供 OpenMP、OpenCL、hthreads 等并行编程接口的编译和运行时环境支撑；其数学库包含基础数学库、基础线性代数库、快速傅里叶变换库、深度学习库等，提供了标准、丰富的数学库接口，并面向飞腾处理器进行深度性能优化；在安全编译方面，天鹰编译系统实现了安全编译验证检查，提供 C 编译器自动测试、C 语言安全子集支持与测试等模块，为编译器的编译过程及生成代码提供安全保障。天鹰编译系统还提供了对应的调试与性能分析工具，支持串行程序与并行程序的调试与性能统计分析。面向飞腾处理器的编译系统支持 FT-1500A、FT-2000+ 等处理器，以及麒麟、深度等操作系统。

## 3.4 平台层生态

### 3.4.1 概述

对于高性能计算而言，上层科学和工程计算领域包罗万象，每种领域内的应用软件也是百家争鸣，然而归约到计算机科学而言，其主要显式参数为计算性能，例如并行规模、并行效率等。为了发挥高性能计算机硬件的计算能力并保证一定的性能可移植性，著名的 Berkeley View 明确了高性能计算硬件和高性能计算应用之间的桥梁，即共性计算函数概念的重要性。对高性能计算定义了七种共性计算模版，包括稠密矩阵计算、稀疏矩阵计算、FFT 等。对于这些共性函数需要进行极致的性能优化，而操作系统和编译器优化能力均无法满足要求，因此催生了众多包含高性能数学库在内的平台层软件。实际上，对包含 x86 架构在内的主流 CPU 甚至众核处理器厂商而言，其数学库的性能与编译器的优化对于高性能计算应用的重要性并驾齐驱。本篇介绍若干 ARM 处理器上平台层软件。

### 3.4.2 OpenHPC 框架

OpenHPC 是 Linux 基金开源的超算项目，致力于为高性能计算构建一个开源框架，适应学术研究的需求，为 HPC 环境创建一个开源框架，降低成本，目前 OpenHPC 仅支持两个架构即 X86 与 ARM，OpenHPC 对 ARM 提供完整支持。OpenHPC 于 2021 年 11 月发布 V2.4 版。创始成员包括 Altair、阿贡国家实验室、ARM、巴塞罗那超级计算中心、克雷公司、戴尔、富士通、惠普、英特尔、劳伦斯伯克利国家实验室 (LBNL)、劳伦斯利福摩尔国家实验室 (LLNL)、莱布尼茨超级计算中心 (LRZ)、联想、洛斯阿拉莫斯国家安全公司 (LANS)、匹兹堡超级计算中心、日本理化学研究所、桑迪亚国家实验室 (SNL) 等。ARM 公司的 ARM 编译器就是面向 OpenHPC 开发。



### 3.4.3 华为鲲鹏 DevKit 与 BoostKit

面向多样性计算时代，华为全面开放鲲鹏全栈能力，分享多样性计算工具套件：鲲鹏应用使能套件 Kunpeng BoostKit 和鲲鹏开发套件 Kunpeng DevKit，加速产业创新，使能极简开发，携手伙伴一起构建鲲鹏计算产业生态。

鲲鹏开发套件 DevKit 提供涵盖代码开发、编译调试、云测服务、性能分析及系统诊断各环节的开发使能工具，方便开发者快速开发出鲲鹏亲和的高性能软件，帮助开发者加速应用迁移和算力升级。同时面向全研发作业流程，提升应用迁移和调优效率，加速原生开发。在开发阶段，DevKit 能够帮助开发者便捷学习鲲鹏架构知识，充分应用鲲鹏架构优势，高效开发高性能应用。迁移阶段，DevKit 的代码迁移工具在原有的软件迁移评估、源代码迁移、软件包重构、专项软件迁移基础之上进一步能力增强，丰富了专项软件迁移和 HPC 应用迁移的能力，同时新增鲲鹏亲和检查等能力，提升软件在鲲鹏平台的运行效率。在编译调试阶段，提供基于鲲鹏平台优化后的 GCC for openEuler、毕昇编译器、毕昇 JDK 等全套编译工具，屏蔽硬件架构差异，充分发挥鲲鹏架构优势。在调优 & 诊断阶段，性能调优工具在原有的系统性能分析、Java 性能分析基础之上，新增调优助手功能，指导用户快速掌握调优方法、降低调优难度。

2020 年发布鲲鹏 BoostKit 应用使能套件，全方位使能开源、贡献开源，繁荣鲲鹏的软件生态。目前，鲲鹏 BoostKit 已经使能超过 90% 主流开源软件支持鲲鹏平台，覆盖大数据、分布式存储、数据库、虚拟化等主流场景。鲲鹏 BoostKit 将聚焦使能主流开源软件支持鲲鹏平台发挥高性能，提供全栈优化的应用加速能力。鲲鹏 BoostKit 虚拟化总体架构主要由硬件基础设施、操作系统、云平台、云管理集群平台构成，其中云平台支持华为自研的 HCS (HUAWEI CLOUD Stack) 私有云平台以及开源 QEMU-KVM、开源 Docker 容器平台。云管理集群平台包括开源 OpenStack 平台、开源 oVirt 平台和开源 Kubernetes 平台。BoostKit 支持包括 OpenStack+kVM、Kubernetes+Docker、Ovirt+KVM 的虚拟化，并实现了 OpenStack 虚拟机与 X86 平台的混合部署，从而能够为用户提供丰富的计算资源。

### 3.4.4 ARM 性能库 ARMPL

ArmPL 作为标准核心数学库，为基于 Arm 的 64 位处理器上的高性能计算应用程序提供优化。ArmPL 是 ARM 官方性能库，可以通过 Fortran 和 C 接口使用，并可以为 ARM 全平台提供加速。其子程序包括基本线性代数子程序 BLAS，综合的高级线性代数程序包 LAPACK，使用 FFTW 接口的实数和复数数据的快速傅里叶变换子程序，一套优化的数学函数 libamath，以及一套优化的字符串函数 libastring。基本线性代数子程序 (BLAS) 是一组定义好的基本线性代数操作。其操作分为三级，第 1 级：

仅作用于向量的操作（例如，点乘），第 2 级：矩阵 - 向量操作（例如，矩阵 - 向量乘法），第 3 级：矩阵 - 矩阵运算（如矩阵 - 矩阵乘法）。许多现代高性能计算机都有 BLAS 的高效特定机器实现。在这些系统上实现更高层次的线性代数算法，关键在于使用 BLAS 作为构建模块。对于 C 语言用户，Arm Performance Libraries 包括 CBLAS，它们是 Fortran BLAS 的 C 语言接口。在 CBLAS 接口中，标量输入参数是通过值传递的。Fortran 接口通过引用传递所有参数（除了通常对 Fortran 程序员隐藏的字符串长度参数）。LAPACK 是一个用于解决数值线性代数问题的 FORTRAN 77 子程序库。LAPACK 组件可以解决线性方程组、线性最小二乘法问题、特征值问题和奇异值问题。它提供了密集矩阵和带状矩阵的处理方法，但没有面向稀疏矩阵。LAPACK 子程序是通过尽可能多地调用 BLAS 来完成的。LAPACK 的效率在很大程度上取决于被调用的 BLAS 的效率。此外，关键的 LAPACK 子程序已经用 OpenMP 进行了处理，以便在 SMP 机器上运行时利用多处理器的性能优势。对于快速傅里叶变换，ArmPL 使用与 FFTW3 相同的接口。

ArmPL 在许多 BLAS、LAPACK、FFT 程序中使用 OpenMP 构建，以便最大限度地提高多处理器性能。但是 ArmPL 有些时候并不能提供极致的性能，例如在鲲鹏 920 处理器上，KML 的性能要好于 ARMPL。

### 3.4.5 Arm Compute Library 开源数学库

是针对 Cortex-A CPU 和 Mali GPU 架构优化的底层机器学习函数的集合，使用 SIMD 技术为 Arm CPU 和 GPU 进行了优化。该库是在许可的 MIT 许可下提供的开源软件。Arm Compute Library 计算库提供优于其他开源替代方案的卓越性能，并支持新的 Arm 技术，例如 SVE2。

它包含如下特性：

超过 100 种用于 CPU 和 GPU 的机器学习功能

- 多种卷积算法（GEMM、Winograd、FFT 和 Direct）
- 支持多种数据类型：FP32、FP16、int8、uint8、BFloat16
- 关键机器学习原语的微架构优化
- 高度可配置的构建选项，支持轻量级二进制文件
- 高级优化技术，例如内核融合、快速数学启用和纹理利用
- 使用 OpenCL Tuner 和 GEMM 优化的启发式方法进行设备和工作负载特定的调整

Arm Compute Library 包含核心库和运行库两个概念。核心库是一个算法实现的底层操作集合，它被嵌入到现有的项目和应用程序中，具有以下特点：不分配任何内存（所有的内存分配 / 映射都要由调用者处理）；不执行任何形式的多线程（但向调用者提供关于如何分割工作负载的信息）。

运行时库是核心库的一个非常基本的包装，可用于快速原型设计，它的特点在于：它通过使用标准的 malloc() 来分配图像和张量；使用一个非常简单的线程池，以非常基本的方式对 Arm Neon 代码进行多线程处理；对于 OpenCL，它使用默认的 CLScheduler 命令队列进行所有映射操作和内核。

Arm Compute Library 支持不同类型的卷积方法，fast-math 标志只用于 Winograd 算法：当启用快速数学标志时，Arm Neon 和 CL 卷积层都会尝试调度可用的最快实现，这可能也会带来精度的下降。虽然该库在工作负载调度过程中支持多线程，从而使工作负载在多个线程上并行执行，但目前的运行时模块实现并不是线程安全的，这是因为提供的调度机制在设计时并没有考虑到线程安全。

### 3.4.6 华为 KML 数学库

华为开发的鲲鹏数学库 (Kunpeng Math Library) 提供了基于鲲鹏平台优化的高性能数学函数，所有接口由 C/C++、汇编语言实现，部分接口提供 Java 语言封装的接口。鲲鹏数学库主要组成部分有：KML\_BLAS 基础线性代数运算数学库 (Basic Linear Algebra Subprograms)，基于鲲鹏架构，通过向量化、数据预取、编译优化、数据重排等手段，对 BLAS 的计算效率进行了深度挖掘，使得 BLAS 接口函数的性能逼近理论峰值；KML\_SPBLAS 稀疏基础线性代数运算库 (Sparse Basic Linear Algebra Subprograms)，基于鲲鹏架构为压缩格式的稀疏矩阵提供了高性能向量、矩阵运算；KML\_VML 向量数学库 (Vector Mathematical Library)，通过 Neon 指令优化、内联汇编等方法，对输入数据进行向量化处理，充分利用了鲲鹏架构下的寄存器特点，实现了在鲲鹏服务器上的性能提升，应用于向量的数学运算，包括幂函数、三角函数、指数函数、双曲函数、对数函数等；KML\_MATH 基础运算数学库 (KML\_MATH)，通过周期函数规约、算法改进等手段，提供了基于鲲鹏处理器性能提升较大的函数实现，适用于基础数学运算，包括三角函数、对数函数、指数函数、幂函数运算；KML\_FFT 快速傅里叶变换 (Fast Fourier Transform) 运算库，适用于 FFT 的 C2C、C2R、R2C、R2R 变换；KML\_LAPACK 线性代数运算库 (Linear Algebra PACKage)，通过分块、求解算法组合、多线程、BLAS 接口优化等手段，基于鲲鹏架构对 LAPACK 的计算效率进行了优化，应用于线性方程组运算，包括求解、特征值和奇异值问题求解等。

### 3.4.7 OpenBLAS 开源数学库

OpenBLAS 是基于 GotoBLAS2 1.13 BSD 版本优化的基本线性代数子程序 (BLAS) 库。其中著名 BLAS 库 GotoBLAS 的相关工作就是针对 GEMM 的相关特性提出了几条基本的分块原则。这些原则通过

调整和限定矩阵的分块策略来充分利用硬件多级存储结构，提高数据局部性利用率。通过这些原则可以构建高性能的 GEMM 实现。目前 GotoBLAS 已经停止维护。而 OpenBLAS 继承于 GotoBLAS 库，并创新性将模板定义和自适应优化等一系列技术结合起来，并使得 Level 3 BLAS 的实现性能提升巨大。OpenBLAS 被应用于科学计算、数据分析、深度学习算法、人工智能等领域，被 Caffe、MXNet、julia、Ubuntu、debian、OpenSuse、GNU Octave 等知名项目所集成。目前 OpenBLAS 已经支持在飞腾、鲲鹏 920 等国产处理器。

### 3.4.8 LIBXSMM 开源数学库

LIBXSMM 是一个专门用于密集和稀疏矩阵运算库，同时该库也被用于诸如深度学习领域中的卷积计算等领域，LIBXSMM 提供了对包括 X86 和 ARM 在内的多个计算平台的支持。该库率先提出并主要使用的代码生成技术基于 JIT(Just-in-time) 编译，以提供与编译器无关的性能，即不需要特殊的编译标志来挖掘可用的性能。LIBXSMM 适用于“一次构建，到处部署”。目前该库支持的 GEMM 数据类型有：FP64、FP32、bfloat16、int16 和 int8。在 LIBXSMM 的工作中，它根据 L2 cache 规模和实际过程中可能获取的性能优势来定义小规模 GEMM 的大小为  $3\sqrt{n} \leq 80$ 。LibXSMM 对于小规模 GEMM 的工作分为前后两端，它的前端同著名的工业标准接口 LAPACK 的 BLAS level-3 函数兼容，它的后端有内部的 API 驱动来直接汇编机器代码。对于它的应用程序编程接口而言，它首先会检查输入参数是否满足小规模 GEMM 的定义。如果输入参数满足小规模 GEMM 的定义，LibXSMM 的后端直接使用内部的接口来构建机器代码以提供用户使用。如果输入参数不满足小规模 GEMM 的定义，则 LibXSMM 调用通用 GEMM 代码。并且前端还提供有代码分发机制：(1) 在线程安全的情况下，快速确定指定规模的代码是否生成，(2) 如果该代码仍然没有被汇编，则进入锁定区域，(3) 分配用来保存机器指令的可执行缓冲区，并调用后端实际生成代码，(4) 保存该代码的函数指针用于之后的代码调用。该机制可以保证指定规模 GEMM 代码生成一次，并在之后重复利用，无需多次生成，并且可以快速调用其函数指针。它的后端会生成内联汇编代码，纯汇编代码或者在运行时生成代码。

### 3.4.9 LibShalom 开源数学库

LibShalom 是一个基于 ARMv8 处理器，针对小规模和不规则形状 GEMM 的开源算法库。它通过改进现有 BLAS 库的缺点，如 Pack 操作占用了大量运行时间、低效的边角处理和不合理的并行方法，来提升基于 ARMv8 处理器的，小规模和不规则形状 GEMM 的性能。LibShalom 提出将 GEMM 的打包步骤融合到计算中，以实现隐藏打包开销的目的。在 GEMM 计算过程中，处理器会执行许多乘加融合指令 (FMA)，此时访存部件处于空闲状态，将打包操作所需的访存指令插入 FMA 指令中来使访存部件保持忙碌，这样可以有效重叠计算与打包。为了处理 GEMM 的边缘，本文还设计了高效的边缘微内核，来隐藏访存取数据的开销。现有的 BLAS 库，如 OpenBLAS，其使用的边缘微内核中的指令尚未进行充分优化调度，



有写后读依赖的指令之间距离很近，不能隐藏取数指令的延迟。LibShalom 对边缘微内核中存在依赖的指令进行了重排，把取数指令排在尽量靠前的位置，确保数据在使用时已经被取到了向量寄存器。此外，本文采用两级并行化策略来并行不规则形状的 GEMM，即对于任务矩阵 C 的 M 和 N 维度进行划分，每个维度使用多少个线程是根据计算访存比这一指标确定。

### 3.4.10 BLASFEO 开源数学库

BLASFEO 是一个开源稠密线性代数库，它的目标是提供 BLAS-like 和 LAPACK-like 的高性能函数实现，并应用于嵌入式优化和一般的小规模高性能计算领域。BLASFEO 最明显的一个特点是为稠密线性代数函数提供了一个全新的接口 BLASFEO API。该 BLASFEO API 向用户开放高性能实现的矩阵计算函数，矩阵 Pack 操作和解 Pack 操作。另外该库还针对可以存入 L2 data cache 的矩阵（即每个维度为 200 左右）进行了性能优化。除此，BLASFEO 还提供了 BLAS API。因为该接口仅针对标准 BLAS API 进行优化，因此并没有考虑代码生成，C++ 模板和专用编译器等方法。而 BLASFEO 的 BLAS API 相比较 BLASFEO API，其性能损失浮动在可以忽略到 20% 之间，其平均值在 10-15%。

BLASFEO 提出 5 种 GEMM 算法。算法 A 为利用 BLASFEO API 执行 GEMM 相关的全部操作，首先为矩阵 Ab、Bb 和 Cb 动态分配内存并对齐 cache 块，之后 Pack 输入矩阵 A、B 和 C 到 Ab、Bb 和 Cb 中，之后计算矩阵并将结果返回。该算法会通过 Pack 操作处理所有的矩阵转置情况。算法 A 的优势是易于实现，不用编写新的 kernel 算法，缺点是有大量无用的动态内存分配影响性能。算法 B 不再 Pack 矩阵 C，而这需要新 kernel 实现。不仅如此这个算法还是有过多的 Pack 操作和动态内存分配。算法 C 为进一步减少 Pack 操作和动态内存分配，该算法只 Pack 矩阵 A。而该算法的缺点是当矩阵 B 为行主序时，性能会有较差表现。算法 C 与算法 C 相似只 Pack 矩阵 A。而算法 D 不 Pack 矩阵而将所有矩阵视为列主序，而该算法无法有效处理所有的转置类型，所以需要为 4 个转置类型 NN、NT、TN、TT 提供 kernel。其中 BLASFEO 认为算法 D 中 TN、TT 的 kernel 不利于并行计算，应利用算法 C 或 C 中 Pack 操作的矩阵转置操作转置矩阵，之后再行计算。

BLASFEO 通过性能对比，针对不同的矩阵数据类型，转置类型，规模和形状，选择最优的算法进行实现。以双精度浮点数的 DGEMM 为例，BLASFEO 首先在 BLASFEO API 中实现算法 A 的 NN 和 NT 转置的 kernel，之后又实现算法 B 的 NT 转置的 kernel，算法 C 的 NN 和 NT 转置的 kernel，算法 C 的 NT 和 TT 转置的 kernel（该实现内部还基于算法 C 的实现），算法 D 的 NN，NT 和 TT 转置的 kernel（其中 TT 转置实现内部基于算法 D 的 NN 转置）。以上这些 GEMM kernel 都是利用模块化的汇编程序编写。BLASFEO 会将 kernel 算法分成基本的操作并封装为函数，之后 kernel 的实现都是通过调用这些函数实现的，该方法可以大副提升代码复用。

### 3.4.11 BLIS 开源数学库

BLIS 是一个用于实例化高性能 BLAS（如密集线性代数库）的可移植软件框架，已经提供了对 ARM 计算平台的完善支持。该框架的设计目的是除去计算用的必要的 kernel，这些 kernel 经过优化后，可以立即优化大多数常用的计算密集型操作的实现。虽然 BLIS 拥有一个新的类似 BLAS 的 API，但它还包括一个 BLAS 兼容层，该层允许应用程序开发人员通过传统的 BLAS 例程调用访问 BLIS 实现，并且 BLIS 独有的基于对象的 API 也可用。

BLIS 将小规模 GEMM 视为瘦 GEMM(skinny GEMM) 的一部分。BLIS 定义瘦 GEMM 为输入矩阵中只要有一维较小，则该 GEMM 即为瘦 GEMM。BLIS 总结出瘦 GEMM 应支持 4 种转置情况：NN、NT、TN、TT，和 8 种存储方式：RRR、RRC、RCR、RCC、CRR、CRC、CCR、CCC（R 为行主序、C 为列主序，三个字母分别对应 CAB 矩阵的存储方式）共 32 种问题情况。因为矩阵 A 和 B 的转置可以通过存储方式进行转换，如规模为  $m \times n$  行主序的转置矩阵等价于规模为  $n \times m$  列主序的不转置的矩阵。因此可以减少为 8 种问题情况：RRR、RRC、RCR、RCC、CRR、CRC、CCR、CCC。

传统 BLIS 会将矩阵分块，之后用对应规模的 microkernel 计算。这些 microkernel 的实现在瘦 GEMM 中也有所不同。首先 microkernel 都是对 k 维循环。而针对 m 和 n 维，有两种实现方式，以存储方式 CCR 为例：（1）加载矩阵 A 的数据进连续的向量寄存器和加载并广播矩阵 B 的向量寄存器相乘，（2）加载矩阵 B 的数据进连续的向量寄存器和加载并广播矩阵 A 的向量寄存器相乘，该方法可能需要对矩阵 C 的数据进行寄存器内矩阵转置操作。BLIS 为尽量避免寄存器内矩阵转置操作，共总结 3 种方法来实现上述 8 种存储方式：（1）矩阵 B 为行主序，其数据按行并按顺序加载，矩阵 A 不论存储方式都按列加载并广播其数据，针对矩阵 C 的存储方式判断是否进行寄存器内矩阵转置操作，该方法实现 RCR、RRR、CRR 的情况。（2）矩阵 A 为列主序，其数据按行并按顺序加载，矩阵 B 不论存储方式都按行加载并广播其数据，针对矩阵 C 的存储方式判断是否进行寄存器内矩阵转置操作，该方法实现 CCR、CCC、RCC 的情况。（3）矩阵 A 为行主序，其数据按行并按顺序加载，矩阵 B 为列主序，其数据按顺序加载，针对矩阵 C 的存储方式存储数据，无需进行寄存器内矩阵转置操作，该方法实现 CRC、RRC 的情况。BLIS 之后会利用 microkernel 进行边角处理操作。其选择实现边长为 2 的倍数的 microkernel。最后这些 kernel 按照经典的 GotoBLAS 算法循环执行。然而由于 BLIS 的分块方法和边角处理并不完善，并且小规模矩阵所以 BLIS 无法获得近似最优的小规模 GEMM 的性能。

### 3.4.12 ATLAS 开源数学库

ATLAS (Automatically Tuned Linear Algebra Software) 是自动调优方向上一个具有代表性的 BLAS 库，支持包括 X86、ARM 在内的多个平台。为了能有效减少开发人员设计高性能计算算法的工作量，

ATLAS 及其前身 PHiPAC 将自动调优的思想引入高性能计算领域。ATLAS 的作者认为针对不同体系结构优化 BLAS 库函数是一项工作量极大及存在较大难度的工作。而根据体系结构自动调优能有效解决该问题。但是 ATLAS 库也存在着自己的问题。首先 ATLAS 实际上也是需要手动优化的内核，然后基于这个给定的内核，ATLAS 会自动调整计算参数来试图优化程序的性能。其次 ATLAS 没有优化 BLAS 库中所有的函数，只优化了其中的一部分函数。并且 ATLAS 对于比较 BLAS 库中比较复杂但在实际开发过程中使用率不高的函数并没有达到很好的优化效果，只是使用最原始的方法实现。

### 3.4.13 稠密线性系统解法器

ScaLAPACK (Scalable LAPACK) 是由 Tennessee 大学、Berkeley 大学和 Denver 大学等负责开发，主要针对密集和带状线性代数系统，提供线性代数求解功能，包括各种矩阵运算，矩阵分解，线性方程组求解，最小二乘问题，本征值问题，奇异值问题等，具有高效、可移植、可伸缩、高可靠性等优点，基于 MPI 实现大规模分布式并行，利用 ScaLAPACK 可以开发出基于线性代数运算的并行应用程序，在电磁仿真、材料计算、分数阶分析等领域应用较多。

SLATE (Software for Linear Algebra Targeting Exascale) 是 Tennessee 大学等开发的稠密线性代数解法器，目标是 ScaLAPACK 的替代，能够提供诸如 GPU 上的稠密线性代数加速计算能力。

国防科大针对天河系统开发的 YHSCALAPACK 是基于稠密线性系统开源软件 SCALAPACK 开发了异构融合 HU-SCALAPACK，提供异构计算接口，支持 GPU、MIC、MATRIX2000 等加速卡。

### 3.4.14 稀疏线性系统解法器

PETSc (Portable, Extensible Toolkit for Scientific Computation)，是美国能源部 ODE2000 支持开发的 20 多个 ACTS 工具箱之一，由 Argonne 国家实验室开发的移植可扩展科学计算工具箱，主要用于在分布式存储环境基于 MPI 高效求解偏微分方程组及相关问题，是应用最广泛的稀疏线性系统解法器。PETSc 主要包含基础数学对象 Mat、Vec 和 Index，几何拓扑对象管理 MDA 和 DMPLex，稀疏矩阵迭代求解器 KSP 和 PC，非线性方程求解器 SNES，最优化求解器 TAO，时间域求解器 TS。PETSc 提供了大量基于 Krylov 子空间方法 KSP 和各种预条件子 PC 的成熟而有效的迭代方法，是其最大的优势之一。

HYPRE 是美国 Livermore 国家实验室开发的针对结构和非结构网格的多重网格线性代数解法器，在流体和电磁等领域应用较多。

国防科大针对天河系统开发的 YH-PARANUT 并行数值算法库工具集，支持国产 ARM 指令集兼容架构、迈创等处理器架构，包括求解稠密线性方程组、特征值问题、稀疏线性方程组、几何多重网格网格和代数多重网格等功能。

YHAMG 是由国防科技大学研发的开源并行代数多重网格算法库，主要用于求解形如  $Ax = b$  的稀疏线性系统。YHAMG 的目的是在大规模并行计算机上为用户提供高性能可扩展的并行线性求解器，在天河超算上扩展到五万节点，并行效率 70% 以上。求解方案是预条件子和子空间迭代法，包含了以下算法：稀疏线性代数基本运算；子空间迭代法，包括 CG 和 GMRES 等；经典代数多重网格算法；Jacobi、GS、SOR 等基本迭代法。

### 3.4.15 ELPA 社区

ELPA 社区是面向于 E 级系统的一种高效特征值直接求解器，主要用于材料领域的应用软件。目前有包括 ABINIT、CP2K、VASP、QuantumEsspresso、NWChem 等 18 款软件基于 ELPA 软件包进行开发。ELPA 社区已经支持 ARM 指令集。

## 3.5 面向 ARM 兼容指令集架构的应用软件生态

### 3.5.1 概述

HPC 应用生态的行业特征明显，例如油气、制造 CAE 行业以商业软件为主，但教育、气象、生命等行业以开源软件为主；不同学科的软件应用特征差异很大。

HPC 应用生态的另一个特点是算力与软件占比均遵循 2-8 原则：

首先，对于算力分布而言，国家超算中心占据公开超算算力平台 80% 的算力。如国家超级计算天津中心的天河新一代、国家超级计算长沙中心占据了大部分算力资源。随着中美科技竞争的加剧，超算系统国产化进程加速前行，2022 年 10 月 9 日，长沙中心“天河”新一代超级计算机的双精度浮点峰值计算性能达 200P Flops、数据存储能力不低于 20PB、峰值功耗不高于 8 兆瓦，算力水平国际先进、国内领先。

其次，对于应用软件分布而言，TOP10 和 TOP20 软件消耗 50% 和 80% 以上算力。行业超算中心一般聚焦某几个单一场景。对于粒子物理领域，聚集在非确定性粒子输运、确定性粒子输运及辐射流体力学模拟等场景。对于航空航天飞行器设计、新型汽车设计等领域，聚集在结构非结构 CFD 计算，兼顾结构力学分析、热分析、通信信号分析、总体性能仿真等。对于信息安全领域，聚集于口令恢复、大整数分解离散对数求解等场景。对于水下航行器，聚集于水声模拟仿真。对于天气预报领域，聚集于中小尺度大气环流模拟等。对于天体物理领域，聚集于行星演化模拟、射电数据处理等。对于行业超算中心，TOP10 软件消耗 80% 以上算力。如对于航空航天飞行器设计，CFD 计算占据 80% 的计算任务数，90%

的计算时间，系统使用率达到 95% 以上，甚至经常任务出现排队情况。对于粒子物理领域，50% 的计算时间被非确定性粒子输运模拟消耗。对于信息安全领域，90% 的计算时间被口令恢复、大整数分解等应用软件消耗。对于公开算力平台，VASP、OpenFoam、WRF、Gromacs 等应用软件消耗了 60% 以上的计算资源。

同时，HPC 应用生态具有明显的多场景模式，特别是高校与政府超算中心多作为算力平台，生态较为复杂。以天津超算为例，需要完成缓解能源危机进程的百万催化剂筛选、工业和信息化部第二批“新一代人工智能产业创新重点任务揭榜项目”、风电行业应用创新服务、新款压缩机的关键设计参数优化、新冠药物筛选、数字数值融合引领环保装备先进工艺优化、高性能数字仿真设计助力高端装备设计、揭示新冠病毒 Omicron 变异株特性、基于自主创新的石油地震勘探行业应用平台等项目以及其他领域计算任务，从而需要比较完善的生态环境。

ARM 应用生态也存在与其他架构应用生态类似的几个特点：

(1) 大型软件与小型软件并存。传统高性能计算领域，如 CFD 计算、结构力学、气候气象等，主要依赖于若干大型应用软件一次性完成仿真计算工作。生物信息等领域，依赖一系列软件，分步骤协同完成仿真计算工作。

(2) 计算密集型与访存密集型并存。信息安全领域中的口令恢复和分子动力学计算，计算量大，访存和内存要求低。而对于离散成稀疏线性系统的科学计算，计算访存比低，访存速度是限制实际应用计算性能的首要因素。

(3) 通信自由型和通信受限型并存。对于线性系统求解类应用，由于只需要对残差等少量数据进行全局通信，可扩展性好，并行效率高。对于大量数据全局操作，如矩阵转置等，可扩展性差，有些只能扩展到几十个节点。

(4) 进程级并行和任务级并行并存。能力型应用和大型问题采用进程级并行，通过分布式多核并行实现极大规模可扩展并行计算。普适型应用和海量问题通常采用任务级并行，一次性提交成百上千个任务，一个任务使用 1-100 个计算节点，通过高并发实现并行计算。

(5) 单一并行与混合并行并存。大部分软件采用单一并行模式，85% 以上支持 MPI 并行。小部分仅支持 OpenMP 并行，极少数支持多级混合并行，甚至异构并行。

(6) 核心算法简洁与流程复杂并存。科学计算应用聚集到稀疏线性系统求解，口令恢复主要包括 AES 等核心算法，非确定性粒子输运模拟主要是粒子追踪。应用核心算法清楚了，但是实际应用代码



动辄百万行，用于复杂情况和异常情况处理。

总体而言，ARM 指令集兼容架构在 HPC 领域的应用生态是相对完备的，在重要的科学计算领域中无论是商业软件还是开源软件均有 ARM 指令集版本。以国家超级计算长沙中心的天河新一代系统为例，已成功部署了 Lammmps、Abinit、VASP、Gromacs、CP2K 等基础科研软件，WRF 等气象海洋软件，OpenFoam、NekTar++ 等制造仿真开源软件和 YH-ACT、YH-Aero、YH-SAS、YH-MAX 等天河生态工程计算软件，tensorflow、tensorflow-lite 等人工智能软件，opencv、graph500 等图计算软件，以及 beagle、bookleaf、fftw、Laghos、numpy、atlas 等数学库，与基于 x86 架构的超算系统对科学计算应用的支持覆盖率相当。

本篇介绍 ARM 上若干重要科学和工程计算领域应用软件。

### 3.5.2 基础科研

理论推导、实验测试、计算模拟是当今科技创新的三大主要手段。超级计算机作为计算模拟的主要工具，是不可或缺的科研基础设施，也是国家科技创新体系的重要组成部分，已成为世界各国竞相争夺的战略制高点。E 级超算所拥有的强大算力，将有力地支撑基础科研领域的原始创新，促进重大成果的产出和关键问题的解决。通过建设基础科研领域的软件生态，可以方便研究人员的使用，加速科技创新，推动科学进步，产出具有国际影响力的科研成果。基础科研涵盖的领域范围广，涉及分子动力学、计算化学、地球物理、石油地质、高能物理、天体物理、量子力学等，相关软件种类繁多，其中相当一部分软件为用户自研。

分子动力学是一套分子模拟方法，该方法主要是依靠计算机来模拟分子、原子体系的运动，是一种多体模拟方法。通过对分子、原子在一定时间内运动状态的模拟，从而以动态观点考察系统随时间演化的行为。通常，分子、原子的轨迹是通过数值求解牛顿运动方程得到，势能通常可以由分子间相互作用势能函数、分子力学力场、全始计算给出。常用软件包括 NAMD、Lammmps、Amber、ABinit、CPMD、VASP、QE、DL POLY、Siesta、Gromacs 等。其中 Lammmps、Abinit、VASP、Gromacs 已在国家超级计算长沙中心已在天河新一代系统上部署。

计算化学是理论化学的一个分支，其主要目的是利用数学近似和电脑程序计算分子性质，例如总能量、偶极矩、四极矩、振动频率、反应活性等，并用以解释一些具体的化学问题。常用软件包括 Wannier90、PSI、GAMESS、CP2K、NWchem、Octopus 等。，其中，CP2K 已在国家超级计算长沙中心已在天河新一代系统上部署。

地球物理是透过定量物理方法研究地球的自然科学学科，研究范围包括地球的地壳、地幔、地核和大气层，通常使用地震波、重力、电磁、地热和放射能等定量的物理方法。传统地球物理学主要指固体地球物理学，现代地球物理学的研究延伸到地球大气层外部的现象，例如电离层电机效应、极光放电和磁层顶电流系统，甚至延伸到其他行星及其卫星的物理性质。常用软件包括 Landmark、Petrel、tesseral、Madagascar 等。

石油地质学是研究石油和天然气在地壳中生成、运移和聚集规律的地质学分支学科，是石油和天然气地质学的简称。主要的勘探方法包括四类：地质法、地球物理法、地球化学法和钻探法。常用软件包括 KronosFlow、basinmod、ChemStat、CMG Suite、FlowSolv 等。

高能物理学是研究组成物质和射线的基本粒子以及它们之间相互作用的一个物理学分支。由于许多基本粒子在大自然的一般条件下不存在或不单独出现，物理学家只有使用粒子加速器在高能相撞的条件下才能生产和研究它们。常用软件包括 Geant4、ROOT、Moose Framework 等。

天体物理学，又称天文物理学，是研究宇宙的物理学，这包括星体的物理性质（光度，密度，温度，化学成分等等）和星体与星体彼此之间的相互作用。应用物理理论与方法，天体物理学探讨恒星演化、恒星结构、星际物质、宇宙微波背景、太阳系的起源和许多跟宇宙学相关的问题。常用软件包括 CASA、DS9、Starfinder 等。

量子力学是研究物质世界微观粒子运动规律的物理学分支，主要研究原子、分子、凝聚态物质，以及原子核和基本粒子的结构、性质的基础理论。它与相对论一起被认为是现代物理学的两大基本支柱。许多物理学理论，如原子物理学、固体物理学、核物理学和粒子物理学以及其它相关的学科，都是以其为基础。量子理论的重要应用包括量子化学、量子光学、量子计算、超导磁体、发光二极管、激光器、晶体管和半导体如微处理器等。常用软件包括 VASP、CASTEP、MS、ELK、ASW、abinit、CP2K、QE、flapw 等。其中 VASP、abinit、CP2K 已在国家超级计算长沙中心已在天河新一代系统上部署。鲲鹏平台也已移植适配基础科研领域大多数常见软件，形成对教育科研的有利支撑。

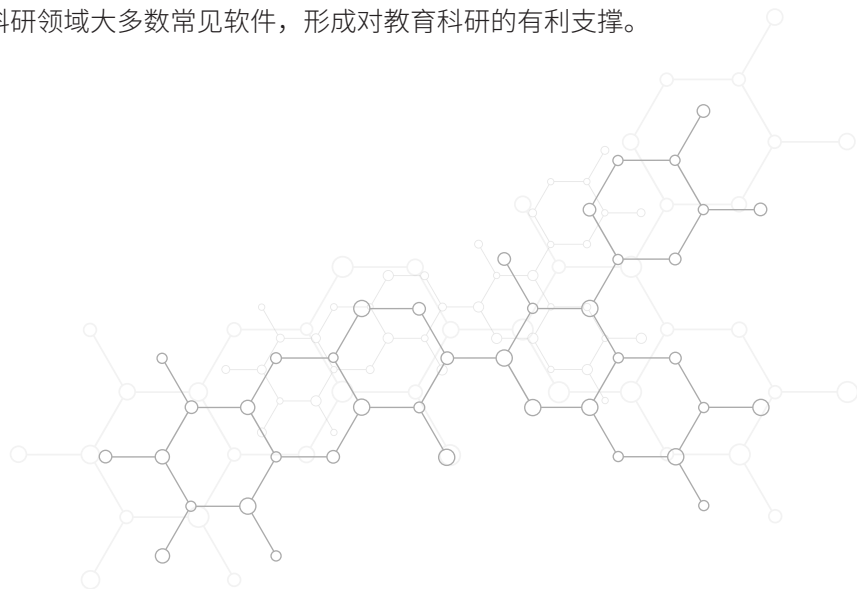


表 3-1 基础科研领域软件清单

序号	名称	类型	领域分类	功能特点
1	NAMD	开源	分子动力学	分子动力学模拟程序
2	Lammps	开源	分子动力学	分子动力学模拟程序
3	Amber	商业	分子动力学	分子力学和动力学模拟程序
4	Abinit	开源	分子动力学	材料光学 / 机械等属性计算软件
5	CPMD	开源	分子动力学	分子动力学模拟软件
6	VASP	商业	分子动力学	分析动力学模拟软件
7	QE	开源	分子动力学	电子结构计算和材料建模
8	DL_POLY	开源	分子动力学	分子动力学模拟软件
9	siesta	开源	分子动力学	材料计算模拟软件
10	gromacs	开源	分子动力学	分子动力学模拟软件
11	Geant4	开源	高能物理	蒙特卡罗应用软件包
12	ROOT	开源	高能物理	模块化科学软件工具包
13	Moose Framework	开源	高能物理	有限元多物理场环境模拟软件
14	Wannier90	开源	量子化学	计算能带结构、介电特性等
15	PSI	开源	量子化学	模拟量子化学
16	GAMESS	开源	量子化学	分子量子化学计算
17	CP2K	开源	量子化学	量子化学和固态物理软件包
18	NWchem	开源	量子化学	可扩展的计算化学工具
19	Octopus	开源	量子力学	量子力学模拟软件
20	Want	开源	前后处理	标准电子结构计算后处理工具
21	udunits	开源	前后处理	物理量单位数据库
22	miniGhost	开源	功能组件	Halo Exchange 迷你应用程序
23	Elk code	开源	量子力学	全电位线性增强平面波代码
24	MOPAC	开源	量子力学	量子化学程序

25	MuSic	开源	前后处理	一种反卷积方法
26	Q-Chem	商业	量子力学	高进度量子化学计算专业工具
28	Quantum Package	开源	量子力学	量子化学开源编程环境
29	vaspsol	开源	分子动力学	vasp 代码补丁
30	vaspvtst	开源	分子动力学	vasp 的过渡态工具
31	vaspkit	开源	分子动力学	vasp 代码的后处理工具
32	Yambo	开源	量子力学	用于固态和分子物理中的多体计算
33	AMBERTools	开源	分子动力学	分子力学和动力学模拟程序
34	ESPResSOmd	开源	分子动力学	分析粗粒度原子模型或珠簧模型
35	Towhee	开源	分子动力学	分子模拟代码
36	QMCPACK	开源	材料模拟	计算原子、分子和固体的电子结构
37	QDK	开源	量子力学	微软的量子开发工具包
38	Gaussian	商业	量子力学	电子结构软件
39	phonopy	开源	量子力学	用于谐波和准谐波级别的声子计算
40	phono3py	开源	量子力学	使用超级单元方法计算声子

### 3.5.3 气候海洋

数值天气预报需要海量的地理信息、实时的气象信息以及复杂的计算逻辑，因此预报模式的发展是与计算机技术发展紧密同步的。在 20 世纪 50 年代后的几十年内，各种新方法和新技术层出不穷，全球各国也相继研发了各具特色的气象模式软件，然而这些软件在数据交换等方面缺乏交互性，对气象研究发展不利。针对这一问题，在美国国家大气研究中心 NCAER(National Center for Atmospheric Research)、美国国家环境预测中心 NCEP(National Centers for Environmental Prediction) 等部门的联合合作下，从 20 世纪 90 年代开始研发下一代中尺度天气预报系统 WRF。

WRF (Weather Research and Forecasting Model) 是当前最为先进和广泛使用的数值天气预报系统，即利用计算机模型模拟和预测大气运动。WRF 可以用于大气科学研究和业务预报，WRF 包括动力框架、数据

同化功能，其软件框架设计也支持高可扩展并行计算。WRF 由计算机和气象领域专业的专家团队开发和维护，公开源代码并支持和鼓励科研和业务单位用户完善已有功能并开展新型模式的创新，因此 WRF 的发展十分迅速，目前有超过 160 多个多家的近 6 万多个注册用户，其功能也得到持续拓展。目前，除了应用于中尺度气象模拟，WRF 也可用于区域气候研究 (Climate-WRF)、行星大气研究 (Planet-WRF)、全球气象模拟 (Global-WRF)。经过多年的更新迭代，WRF 已经成为领域内最为成熟应用最为广泛的大气数值模式之一。

WRF 为国外开源软件，采用 Fortran 语言开发，支持并行计算，具有十分丰富使用手册等资料。WRF 模式可以读取处理气象场数据以及下垫面数据，然后求解一系列大气动力学热力学的参数化方程，最终得到目标区域和范围内的如温度、湿度、气压等气象要素，通过这一功能可以实现天气预报以及历史天气模拟。大气科学研究人员也能够通过更改一系列参数或者程序代码，模拟不同参数条件下天气情况演变的差异，以此来进行科学研究。WRF 的模块化发展非常完善，里面囊括了多种研究情景案例以供用户按需选择，比如大涡模拟、理想大气情景模拟等。除此之外，相关开发团队也将功能向其他交叉领域进行了拓展，比如将化学模块 Chem 也耦合到了程序中，一般称为 WRF-Chem 模式，用户能够自主选择是否使用化学模块，进行大气化学领域的相关研究。与此相似的，还有大气 - 野火耦合模式 (WRF-Fire)、大气 - 温室气体耦合模式 (WRF-GHG) 等多种多功能多领域的耦合模块，这些模块都是围绕 WRF 开发的模块程序，即插即用，因此 WRF 在一些和大气有所交叉的领域中也有广泛的使用。

WRF 是基于 Fortran 开发的开源软件，已经在基于 ARM 指令集兼容架构处理器的天河超算系统、国家超级计算长沙中心的天河新一代系统和基于鲲鹏 920 处理器的超算系统上成功移植、部署和运行。

气象海洋始终是高性能计算的重要应用领域，计算量大、时效性高。随着地球系统模型向着更高分辨率、更复杂物理过程、集合预报、多圈层耦合方向发展，对算力和高性能计算技术的要求也越来越高。粤港澳大湾区面临着巨大的气象灾害风险，气象服务每提升一个数量级，要求算力需要增加四个数量级。气象海洋领域包含天气预报、气候预测、海洋模拟、空气质量预报等多个方面，可大大提升对台风、暴雨、雷电等极端天气的预测能力，有效减少人民的生命财产损失。

天气预报即综合气象站、卫星等采集的信息数据，对一定区域范围内未来一段时间的天气情况进行预测，时间尺度一般是 1 小时~30 天。天气预报可以指导人民生产生活，防范极端天气，方便人们出行，让老百姓直接受益。天气预报的常用开源软件有 WRF、GFS，由美国主导开发，目前均已在鲲鹏平台完成适配。随着国内气象领域的发展，涌现出一批优秀的国产自研软件，如：GRAPES、BCC，目前也已在鲲鹏平台完成适配。天气预报一方面需要对大量的观测数据进行质量检查和同化分析才能获得初值条件，处理过程计算量很大；另一方面，数值求解上百万个自由度的控制偏微分方程本身也需要大量的计算，对时效性要求高，这就要求有大算力支撑。



气候预测与天气预报不同，除了依赖于大气初值以外，还需要考虑下垫面状况包括陆面、海洋、海冰的影响，以及它们之间的相互作用，甚至包括人类活动的影响。气候预测研究范围包括冰川融化、海平面上升速度，碳循环的过程、大气成分演变过程，气候变化、厄尔尼诺现象的产生等。气候预测对 E 级计算能力的要求是迫切的，E 级计算将为更复杂的模型模拟提供更高的精度并有效缩短计算时间。气候预测的常用软件为美国国家大气研究中心提出的通用地球系统模式（The Community Earth System Model），已在鲲鹏平台完成适配。

海洋模拟可用于全球海洋洋流变化、海洋温度、波浪变化、河口海口沉积物演变等研究。海洋领域模式众多，且大多为开源。比较常见的软件包括：NEMO、SWAN、FVCOM、ROMS、HYCOM、WaveWatch3、ECOMSED、ECOM、POP、POM、MOM4、MITgcm 等，均已在鲲鹏平台完成适配。其中，NEMO 用于海洋环流、气候研究，ROMS 用于三维海洋建模，SWAN 用于波浪仿真。海洋模拟需要对海量数据、众多的数据源的数据集进行调查、观测、计算，并对海洋数据进行快速交互与分析处理。随着 E 级超算平台的研发，全球海洋模式的水平分辨率必然也会向公里级、百米级发展，可分辨种类更齐全的海洋运动形式，如次中尺度涡、内波等。国产 E 级超算平台的建设是发展我国自主研制超高分辨率海洋模式的绝佳契机。

随着“金山银山不如绿水青山”、“碳中和”、“碳达峰”等理念的提出，空气质量预报研究越来越受到人们的重视。空气质量预报涉及的程序模式众多，但都有计算量大、通信频繁、I/O 需求大等共同特点，对计算机的性能要求非常高，所以一个具备高性能、高可用性、高可靠性的高性能计算系统平台成为空气质量数值预报领域研究和应用发展的关键。鲲鹏平台已适配常用的空气质量模拟软件，如 WRF-Chem、CMAQ、CAMx 等，对于研究空气污染浓度预报及污染物跨区域传输机制具有非常重要的意义。

目前，鲲鹏平台已移植适配气象海洋领域软件共 25 款，涵盖了该领域几乎所有常用软件，满足绝大部分用户关于天气预报、气候预测、海洋模拟及空气质量等方面业务需求。下一步，围绕 E 级超算平台，将进一步建设并完善气象海洋领域的软件生态。

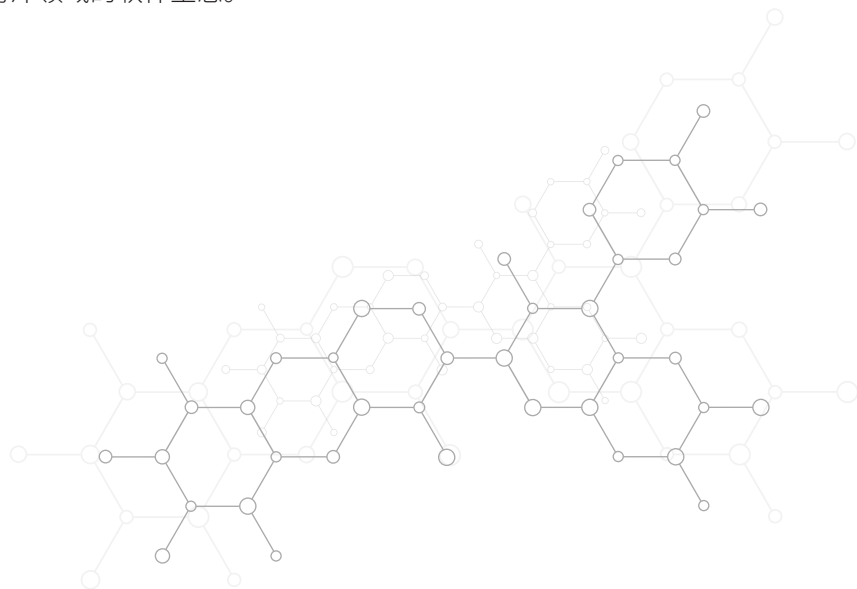


表 3-2 气象海洋领域软件清单

序号	名称	类型	领域分类	功能特点
1	WRF	开源	气象预测	中尺度天气预报模式
2	GFS	开源	气象预测	全球天气预报系统
3	CESM	开源	气候预测	全球气候模拟系统
4	NEMO	开源	海洋模拟	海洋环流、气候研究建模框架
5	ROMS	开源	海洋模拟	区域海洋三维模拟系统
6	WAVEWATCH	开源	海洋模拟	海洋、波浪模拟系统
7	MITGCM	开源	海洋模拟	模拟大尺度海洋环流的模型软件
8	MOM	开源	海洋模拟	模块化海洋模型
9	HYCOM	开源	海洋模拟	三维海洋模型
10	SWAN	开源	海洋模拟	第三代波浪仿真模型
11	CMAQ	开源	空气质量	通用多尺度空气质量模拟
12	CAMx	开源	空气质量	气体和颗粒污染的多尺度光化学模拟系统
13	CDO	开源	结果处理	气象类数据处理运算
14	eccodes	开源	结果处理	用于 WMO GRIB 和 BUFR 格式消息编码解码
15	ecflow	开源	结果处理	工作流管理器
16	NCEPLIBS	开源	结果处理	气象领域常用依赖库
17	NCL	开源	结果处理	科学数据分析和可视化语言软件
18	PyFerret	开源	结果处理	可视化和分析环境
19	FVCOM	开源	海洋模拟	沿海海洋环流模型
20	ARPS	开源	气象预测	风暴尺度预报系统
21	GADGET-2	开源	天文物理	用于解决广泛的天体物理学问题
22	MM5	开源	气象预测	中尺度大气环流预测模型
23	GRAPES	开源	气象预测	中国气象局自主研发的天气预报系统
24	CosmoMC	开源	天文物理	用于探索宇宙参数空间
25	WPS	开源	气象预测	WRF 模式前处理系统

### 3.5.4 制造仿真

伴随中国制造业转型升级进程加速，工业产品设计研发过程中正在大量使用计算机辅助创新技术，以计算机辅助设计（CAD）、工程分析仿真验证（CAE）、电子设计自动化（EDA）、计算机辅助工艺设计（CAPP）为代表的 CAX 研发软件已经成为先进制造业及相关科研院所、设计单位不可或缺的重要研发工具。而这些仿真软件所计算的模型往往较为庞大，对算力规模有很高要求。通过 E 级超算建设，加快制造仿真计算，加速工业 4.0 推进，对于粤港澳大湾区工业制造转型有重大意义。制造仿真包含计算流体力学仿真、结构力学仿真、电磁仿真、计算电磁学等多个子领域。

计算流体力学（Computational Fluid Dynamics, CFD）仿真包含流动、湍流、热传递和反应等广泛物理现象，其应用范围广，从飞机气动布局设计到汽车车身设计、从发动机燃烧模拟到人体血液模拟都可用到，是现代工业发展的基础之一。欧美国家已在计算流体力学领域深耕多年，开发出了以 Ansys、Abaqus、ESI VPS 等为代表的典型应用在行业具有广泛的使用，这其中 ESI VPS 已有鲲鹏版本，可直接运行在 E 级超算平台，其余软件可通过转码工具运行在鲲鹏平台。开源软件如 OpenFOAM、SU2、PALABOS、Deal.II、Code Saturne 等已在鲲鹏平台完成适配。流体力学仿真问题模型较大，对算力需求较高，以大型舰船的模拟仿真为例，往往需要千万核计算资源，E 级超算平台建设势在必行。其中，OpenFOAM、NekTar++ 等软件已在国家超级计算长沙中心的天河新一代系统上部署运行。

结构力学仿真研究的主要内容包括计算固体材料的形变、应力和应变，通常用来确定结构（例如桥梁）的强度，以防止发生损坏或事故。结构力学应用范围包括建筑桥梁设计、器件模块设计，车辆、舰船设计等。当前结构力学领域典型软件如 Elmer、Code Aster 已在鲲鹏平台完成适配。

电磁仿真领域通常采用有限元、有限差分、矩量法等方法对物理场进行网格划分，可用于手机天线仿真、部件电磁兼容计算等民用计算场景，也可用于雷达仿真、飞机 RCS 计算等重要军事领域。随着计算电磁波频率的升高，要求将计算网格划分的更细，这就导致算力需求呈指数增长，要求更大的计算算力。在计算大型飞机、舰船的高频电磁散射时，往往需要上千万核才能快速完成计算，这就依赖 E 级超算平台的建设。鲲鹏平台已适配 MEEP、MPB 等开源电磁仿真软件，可满足常用电磁仿真需求。

计算电磁学是现代电磁理论研究的主流，其建立在电磁场理论上，以高性能计算机技术为工具，运用计算数学方法，专门解决复杂电磁场与微波工程问题。仿真计算是认知电磁特性不可或缺的现代化技术手段，电磁工业软件作为仿真计算的工具体，如 ANSYS HFSS、FEKO、ANSYS Maxwell，已被广泛应用于电子装备的电磁设计中。

基于 ARM 指令集兼容架构处理器，国防科技大学研制了一系列面向制造仿真领域的工程计算软件，目前已在国家超级计算长沙中心的天河新一代系统移植适配。

表 3-3 YH-CANGLONG 系列软件

序号	名称	领域分类	应用场景
1	YH-ACT	流体力学	反应堆热工水力、船舶海洋等领域水动力学 CFD 计算，流固耦合传热计算
2	YH-Aero	流体力学	航空航天、地面装备设计等领域空气动力学计算
3	YH-SAS	结构力学	典型建筑物构件 / 结构静力变形与应力分布仿真。
4	YH-MAX	电磁学	地球介质电磁场计算、RCS 计算
5	YH-PARTICLE	粒子物理	反应堆物理堆芯数值模拟
6	YH-GRID	网格生成	面向多领域的网格生成
7	YH-VIEW	可视化	流场可视化数据分析
8	YH-MULTIP	多物理耦合	结构力学仿真、CFD 仿真、流固耦合计算

鲲鹏也已适配制造领域软件 15 款。下一步，基于 E 级超算平台建设制造领域软件生态，将更大范围覆盖制造仿真业务的需求。

表 3-4 制造仿真领域软件清单

序号	名称	类型	领域分类	功能特点
1	ESI VPS	商业	结构力学	碰撞、爆炸、穿甲仿真
2	Elmer	开源	结构力学	多物理场仿真软件
3	Code Aster	开源	结构力学	基于有限元方法的力学求解器
4	OpenFOAM	开源	流体力学	计算流体力学分析
5	SU2	开源	流体力学	高精度偏微分方程求解器
6	PALABOS	开源	流体力学	LBM 方法 ,CFD
7	Deal.II	开源	流体力学	求解偏微分方程开源有限元软件
8	Code Saturne	开源	流体力学	计算流体动力学
9	FDS	开源	流体力学	火灾模拟工具
10	MEEP	开源	电磁仿真	时域有限差分法模拟
11	MPB	开源	电磁仿真	有限时域电磁仿真
12	AMG	开源	前后处理	网格划分工具
13	mfem	开源	电磁仿真	有限元离散化库
14	OpenLB	开源	流体力学	用于实现格子 boltzmann 方法
15	paraFEM	开源	结构力学	结构力学有限元计算软件

### 3.5.5 材料计算

VASP 是维也纳大学 Hafner 小组开发的用于电子结构计算和量子力学 - 分子动力学模拟软件包。VASP 全称 Vienna Ab-initio Simulation Package, 用 Fortran 语言编写, 支持并行模式 (MPI/OpenMP), 是开源的商业软件。VASP 是目前材料模拟和计算物质科学研究中最流行的商用软件之一, VASP 基于密度泛函理论、投影增强波法、分子动力学、GW 近似等理论, 实现了结构优化、电子基态特性、选择自由度、第一性原理分子动力学、分子动力学采样、机器学习力场训练、声子计算、静态和频率相关的介电特性、Berry 相位、光谱学 (UV、VIS、X 射线)、声子等功能。VASP 是超算中心机时占比很高的一个应用软件, 已在国产环境下部署, 计算效率可以充分发挥国产 CPU 的计算性能, 支持 MPI 和 OpenMP, 在单节点计算、跨节点计算均具有较好的性能。VASP 通过组合多个并行化选项来产生最佳性能, 因为每个级别的并行效率都会下降到接近其极限。默认情况下, VASP 在可用的 MPI 等级上分配波段数 (NBANDS), 但添加 FFT 的并行化 (NCORE)、k 点的并行化 (KPAR) 和单独计算的并行化 (IMAGES) 通常可以提高并行效率。已在国家超级计算天津中心和长沙中心的天河新一代系统部署适配。

AMBER 是一套用于生物分子动力学模拟的国外商业软件, 采用 Fortran90 和 C++ 语言编写, 支持 MPI 并行模式, AMBER 最早问世于 20 世纪 70 年代, 它利用公开已知的分子力场对生物分子进行动力学模拟。目前最新的 AMBER 软件为需要取得付费授权使用的分子动力学模拟软件 AMBER22 和可以免费使用的生物分子工具集 AMBERTOOLS22, 两款软件均可独立使用。

Quantum-ESPRESSO (op(E)n (S)ource (P)ackagefor (R)esearchin (E)lectronic (S)tructure, (S)imulation, and (O)ptimization) 软件包的开发遵守 GNU 自由软件的协议, 是基于密度泛函理论, 应用平面波基组和赝势方法的第一性原理计算软件。先前由于计算软件的落后而使得一些有用的方法如线性响应、超软赝势 CP 分子动力学 (MD) 方法等, 受到了应用上的阻碍, 这个软件的发布正是基于这种情况, 从而提供了应用这些方法的一个平台。国内外使用 QE 程序的团队目前正在逐年增多, 该软件开源, 有类似 VASP 一样的计算效果。

CP2K 是运行最快的开源第一性原理材料计算和模拟软件, 可研究上千个原子的大体系, 广泛用于固体、液体、分子、周期、材料、晶体和生物系统的模拟。CP2K 是基于 Fortran 的国外开源软件, 同时它也依赖 C/C++。CP2K 支持 MPI 方式并行也支持 OpenMP 方式并行。CP2K 可以使用 NEB 或二聚体方法进行分子动力学, 蒙特卡洛, 振动分析, 能量最小化和过渡态优化的模拟。适用于第一原理相图, 原位光谱电化学探测, 溶剂化电子的浅层和深层阱态及其形成、电子激发和弛豫动力学, 纳米流体中的渗透运输等。CP2K 的最大优势在于计算速度很快尤其是 quickstep 模块。AIMD 功能是最常用的, 适合计算大体系 (> 200 个原子), 可以在国产环境下实现大规模并行。目前, 主要用于材料的快速相变、团簇特性分析、固体结构分析等。已在国家超级计算长沙中心的天河新一代系统部署适配。



SIESTA (Spanish Initiative for Electronic Simulations with Thousands of Atoms) 是一种实现电子结构计算和第一性原理分子动力学模拟的程序，用于分子和固体的电子结构计算和分子动力学模拟。SIESTA用C/C++语言编写，支持的并行模式MPI，是国外开源软件。SIESTA程序的一个非常重要的特点是，它的精度和成本可以在很大范围内调整，从快速探索性计算到与其他方法（如平面波方法）质量相匹配的高精度模拟，SIESTA支持功能总能量和部分能量、原子作用力、应力张量、电偶极矩、电子密度、电介质极化、能带结构、声子振动等计算。SIESTA在国产环境下均已部署，在国产处理器上性能可以得到很好的释放。多个科研团队使用国产系统通过 siesta 软件进行结构弛豫、电子性质、声子性质、分子动力学的计算以及生成赝势文件。

PWmat 是一款通用的平面波密度泛函理论计算软件。相较于业界同类软件，PWmat 是世界上最早利用 GPU 加速 DFT 计算的软件代码，对比 CPU 版本程序可实现超 30 倍以上的性能加速。基于这一独特的性能优势，PWmat 非常适合于大体系计算，可以在 GPU 集群上以第一性原理计算精度实现 10,000 原子规模体系的计算模拟。即使在单节点 GPU 服务器上也可轻松计算 1000 原子体系。PWmat 具有独特的算法优势，可快速实现 HSE、TDDFT、NAMD、Solvent model、电声耦合、缺陷体系声子计算、非辐射衰减、电荷俘获计算、带电缺陷计算、VCA 计算等多种类计算。目前，PWmat 软件代码已超过 100 万行，并仍然在通过集成开源代码及自主研发等方式补充完善新的计算功能模块，目前已有超过 70 个计算模块（module）可满足多个不同领域的计算需求。PWmat 已在武汉超算中心的国产化超算平台部署适配并对外提供计算服务。

### 3.5.6 生命科学

生命科学是高性能计算的重要应用领域，数据量大、数据种类多样。随着基因组、转录组和蛋白质组、医疗影像等生物大数据井喷式增长，对数据存储的要求越来越高，同时，数据分析和挖掘过程中对算力和高性能计算技术的要求也越来越高。生命科学研究很多时候需要与时间赛跑，更快的药物研发和疫苗开发速度，意味着更少的病人死于新冠或其它疾病，而算力是助跑剂，更大的计算量代表着更短的虚拟筛选或模拟时间。生命科学领域包括基因序列比对、虚拟药物筛选、疫苗研发和医学影像研究等多个方面，借助 E 级超级计算机，可大大加快对药物和疫苗等研发的速度，最大限度保障人民群众生命安全、减少财产损失。

序列比对技术是生物信息学对人类社会最为重要的贡献之一，无论是亲缘鉴定、新物种识别还是在遗传病分析，都离不开序列比对。目前众多医疗机构在超算能力的加持下，将数天的诊断时间大幅缩短至几个小时，同时透过人体基因资料库进行比对及分析，从中找出最适合病患的治疗方法。常用软件包括：BLAST、BWA、CNVnator、HISAT2、Bowtie、Minimap2、diamond、hisat 等已在鲲鹏平台完

成适配。基于 E 级超算平台赋能的基因序列比对，能够在极短时间内找到序列间的公共部分，为以往基因筛选难、检测时间长所导致的问题提供解决新模式。

虚拟药物筛选是药物筛选技术发展的另一个方向，是将药物筛选的过程在计算机上模拟，对化合物可能的活性作出预测。生成新分子及优化分子的关键属性，如生物活性、成药性、安全性和选择性等，是影响药物分子设计成败的两个关键问题。比较常见的软件包括：PyRx、DOCK、DSMS、Sybyl、QSAR、octave 等。新一代 E 级超级计算机提供的强大算力能够支撑百亿级别药物小分子的快速筛选，辅以更先进的算法，可以实现虚拟药物筛选效能几十乃至上百倍的提升；同时高效的药物筛选，还可以与中药有效组分发现结合，推动中药研发的现代化。

医学影像提供了功能性和结构性的临床信息，如何从影像中分析疾病相关信息，并加快其分析速度以及提高医学影像的图像质量是医学影像领域亟需解决的问题。医学影像研究范围包括 MRI 核磁共振、CT 断层扫描、X 光扫描图像等。鲲鹏平台已适配常用图像处理软件，如 Relion、GDS3D、eman2 等。随着医学影像库的数据规模达到 TB 级，计算过程中对缓存及存储的需求大幅上升，为影像数据提供胖节点满足大内存需要，提供高性能的存储池来保障大 IO 带宽，有助于推动实现影像科室数字化，降低诊断失误和手术风险。

目前，生物医药筛选的 Gromacs 软件和国防科技大学自研的智能药物筛选软件 YH-IDVS 已在国家超级计算天津中心和长沙中心的天河新一代系统部署适配，基于该软件的研究成果获得戈登贝尔奖新冠特别奖提名。

目前鲲鹏平台已经适配了 170 款生命科学领域的开源软件，为生命科学领域提供超大规模的算力支撑，涵盖基因序列比对、虚拟药物筛选、疫苗研发和医学影像研究等领域的主要过程，并进一步完善生命科学领域的软件生态。

表 3-5 生命科学领域常见软件清单

序号	名称	类型	领域分类	功能特点
1	Canu	开源	生物基因	高噪声单分子测序工具
2	BLAST	开源	生物基因	基因比对
3	BWA	开源	生物基因	基因比对
4	GATK	开源	生物基因	分析基因组的工具箱
5	Bowtie2	开源	生物基因	基因比对
6	FastQC	开源	生物基因	高通量序列数据质量检控工具
7	Bowtie	开源	生物基因	基因比对

8	Abyss	开源	生物基因	基因序列汇编器
9	FASTA	开源	生物基因	基因搜索和比对程序
10	Minimap2	开源	生物基因	基因比对
11	beast2	开源	生物基因	分子序列分析
12	DESeq2	开源	转录组学	基于负二项分布的差异基因表达分析
13	Blastz	开源	生物基因	核苷酸局部比对程序
14	CAFE	开源	生物基因	基因家族大小分析工具
15	fastp	开源	生物基因	多合一 FASTQ 预处理器
16	Trinity	开源	生物基因	基因组组装
17	HMMER	开源	生物基因	基因比对
18	ABSOLUTE	开源	生物基因	癌症 CNV 分析软件
19	Bioconda	开源	生物基因	生物信息学 conda 软件包管理
20	eman2	开源	前后处理	科学图像处理软件套件
21	vmd	开源	前后处理	分子建模与可视化软件
22	cisTEM	开源	冷冻电镜	大分子复合物冷冻 EM 图处理
23	Relion	开源	冷冻电镜	低温电镜 (cryo-EM) 图像处理
24	spiderweb	开源	冷冻电镜	电镜单粒子重建系统图像处理
25	yem	开源	冷冻电镜	生物电镜相关功能集合

### 3.5.7 地球物理

CitcomS (for California Institute of Technology Convection in the Mantle) 是一种有限元代码，用于解决与地幔有关的可压缩热化学对流问题。软件多数版本 C 语言与 python (基于 Python 的建模框架 Pyre) 进行编写，最新的 3.3.1 版本只采用 C 语言进行编写，可以在多种并行处理计算机上运行包括共享与分布式内存平台，支持 MPI 环境并行。软件为开源软件遵循 GPL-2.0 协议。主要开发者为 Louis Moresi, Shijie Zong, Lijie Han 等。目前已在国家超级计算天津中心和长沙中心天河新一代系统移植，并实现异构优化。

ASPECT (Advanced Solver for Problems in Earth's ConvecTion) 是由 CIG (Computational Infrastructure in Geodynamics initiative, 地球动力学计算基础设施计划) 组织开发的一套开源地球动力学热对流模拟代码，采用 C++ 语言开发，基于开源有限元求解库 deal.II 开发，支持 MPI/OpenMP 并行，旨在为地学工作者提供一套具备高可用性、高可扩展性、数值求解算法先进、文档完备的开源软件。目前该软件主要用于模拟地幔对流问题，但后续将支持更广泛的地球动力学热对流模拟问题。

## 3.6 面向 ARM 指令集兼容架构的国产科学 / 工业软件

### 3.6.1 概述

过去十多年，我国超级计算机屡次夺得世界第一，表明我国在高性能计算机硬件方面的发展已成为世界第一梯队。相比而言，虽然大规模应用软件的研制整体情况仍落后发达国家，但是随着 3 次获得高性能计算应用戈登贝尔奖及多次提名，我国在高性能计算应用方面的研制也呈百花齐放局面，应用领域覆盖面较广，部分应用软件并行规模和并行效率均达到国际领先水平。科学和工程应用作为上层软件直接面向领域用户，通常需要领域科学家和计算机科学家协同配合完成大规模应用开发和优化，本篇介绍我国自主研制的若干科学和工程应用软件。

### 3.6.2 中国科学院过程工程研究所芯片后端验证工具 ChEDA

芯片的设计与制造是我们国家的战略方向，半导体材料、器件与芯片电路的计算模拟对后摩尔时代高端微处理器电路的设计与开发具有重要意义。中国科学院过程工程研究所研制的电子设计自动化软件 ChEDA 是芯片电子设计自动化 EDA 软件的后端验证工具，主要面向集成电路的热效应校验和功耗进行芯片的设计与开发。ChEDA 软件支持半导体材料纳微结构、不同晶体管器件、以及大量晶体管组成的集成电路及其组成功能模块的模拟。尤其对于 10nm 以下线宽的芯片工艺，能够考虑纳微尺度多种微观因素的影响，相比传统 EDA 软件，具有更强的预测准确性。软件已经在硅纳微结构、芯片电路中的不同功能模块等方面得到实例验证与初步应用。软件已完成鲲鹏百节点环境的适配和调优，百节点强扩展性达 60%，并行效率优化与性能优化持续进行中。

### 3.6.3 中国科学院过程工程研究所离散颗粒运动模拟软件 DPS

颗粒广泛存在于化工、材料、能源、和食品等过程工业中。超过 50% 的工业产品为颗粒类物质，市场规模超千亿，但现有宏观模型与连续模型预测精度不高。基于离散单元法的离散颗粒模型，直接在颗粒尺度上计算颗粒与颗粒之间的作用力，跟踪颗粒的运动，可以方便地研究多粒径体系、颗粒间复杂作用力、颗粒停留时间分布、流动 - 传热 - 传质和反应过程的耦合等，是精确预测颗粒系统特性的有效工具。但是，实际工业设备运行与颗粒运动存在多个时空尺度的差别，颗粒数量可达十万亿，并且还存在着分布不均匀，时间步长小的问题，巨大的计算量制约着离散颗粒模型的应用规模。因此，采用大规模并行计算，是解决这一问题的关键，在 P 级算力下，颗粒模拟规模仅为十亿量级，只能模拟实验尺度的设备，模拟工业设备时需要采用较高的粗粒化率，降低了模拟的精度。使用 E 级算力或更高算力，颗粒模拟规模可达到万亿量级，可以直接模拟工业设备中颗粒的运动，预测颗粒系统的特性。更精确的颗粒信息可以为设备选型、催化剂设计等提供更精确的指导。中国科学院过程工程研究所研制的 Discrete

Particle Simulator 软件已完成鲲鹏百节点环境的适配和调优，百节点强扩展性达 67%，并行效率优化与性能优化持续进行中。

### 3.6.4 中国科学院过程工程研究所拟颗粒模拟软件 PPM+

超过 80% 的化工过程使用催化剂，其高效、稳定、绿色对节能环保低碳发展意义重大。催化剂颗粒常具有复杂的多级孔道，显著影响反应物和产物扩散及表观反应动力学，但目前缺乏准确量化方法。中国科学院过程工程研究所研制的 PPM+ 软件耦合拟颗粒模拟（Pseudo-Particle Modeling, PPM）及多种离散模型，计算效率高、可扩展性好，是过程工业和航空航天等领域中纳微多相系统、气固表面作用和复杂孔道内反应 - 传递耦合过程的有力研发工具。软件已完成鲲鹏百节点环境的适配和调优，百节点强扩展性超过 40%，并行效率优化与性能优化持续进行中。

### 3.6.5 中国科学院国家天文台天体物理模拟软件 PHoToNs

天体物理模拟对于天文学家研究非线性结构的形成以及暗物质、暗能量等的课题具有重要意义。未来的高精度天体物理模拟包含数千亿甚至数万亿粒子，因此需要巨大的计算能力和高效的算法。超级计算机作为大规模高分辨率天体物理模拟的最重要工具，在过去几十年里发展迅速。20 世纪以来，宇宙学研究有了很大的进步，许多著名的理论是基于观测天文学提出的，例如宇宙膨胀和大爆炸理论。由于暗物质对整个电磁光谱都是不可见的，因此模拟暗物质演化就变得及其重要。这些天体物理模拟给予天文学家了解诸如星系、暗物质、暗能量等物质的非线性结构形成理论给予很大的帮助。宇宙 N 体模拟已经成为超级计算机的主要应用之一。

中国科学院国家天文台研制的 PHoToNs 支持 N 体模拟相关算法。为了克服超大规模模拟带来的挑战，软件选用最合适的快速算法：FMM-PM。快速多极子方法（FMM）被列为 20 世纪十大算法之一，其计算复杂度为  $O(N)$ ，但编程实现比其他 N 体方法更为复杂。为了解耦不同尺度上的计算，PHoToNs 使用 FMM 与谱方法相结合的策略，这一策略还有利于构造非本地作用计算列表，减少并行通信量。PHoToNs 支持并行异构平台。软件已完成鲲鹏百节点环境的适配和调优，百节点强扩展性高于 80%。

### 3.6.6 中国科学院计算机网络信息中心相场模拟软件 ScLETD

材料科学在我国非常活跃。由于计算方法在材料发展中起着催化作用，自 1990 年以来，计算材料科学在中国一直受到重视。材料设计和性能预测一直是研究的重点领域。近年来，随着新计算方法的发展和更大的计算能力，我国研究人员已经能够应用量子力学来设计具有规定性质的复杂材料。相场模拟方法是针对组织相变过程而提出的一种模拟方法，目前已经在材料领域的各类组织演化问题中得到了广



泛应用。与传统的基于局部能量的均匀体热力学理论相比，相场模拟方法的优点在于总自由能中同时包含局部自由能和非局部的梯度能。非局部能量用来描述非均匀组织形态中的扩散型界面，是所有相场模型的核心。中国科学院计算机网络信息中心研制的 ScLETD (Scalable compact Localized Exponential Time Differencing) 软件使我们能够在现代高性能计算机上进行高效的大规模长时间仿真。软件已完成鲲鹏百节点环境的适配和调优，强扩展性可达 70%，持续优化各项性能。

### 3.6.7 中国科学院计算技术研究所分子动力学模拟软件 DeePMD-kit

分子动力学的仿真模仿广泛应用于医药、化学、生物、资料等畛域，钻研模仿物质的宏观构造，能解物质的宏观性质，甚至对物质的宏观性质做出预测，物质的宏观构造是由原子间的相互作用决定的，所以分子动力学仿真模仿的重点就在于如何建设模型形容分子间的相互作用。中国科学院计算技术研究所研制的 DeePMD-kit 是一款用深度学习做分子动力学模拟的开源软件包，通过神经网络拟合第一性原理的计算结果来完成精确的量子化学计算，并可以配合分子动力学软件来完成最终的分子动力学模拟。DeePMD-kit 使用的神经网络比较小，计算速度非常快，并且可以保持第一性原理的计算精度，此外模型的设计使其能更好地进行分布式计算，因此可以完成更大规模的分子动力学模拟。软件已完成鲲鹏百节点环境的适配和调优，百节点强扩展性 >70%+；达到相比 2020 年 GB 奖水平，计算能效提高 5 倍，并行效率优化与性能优化持续进行中。

### 3.6.8 中国科学院格点量子色动力学 CLQCD

格点量子色动力学 (Lattice Quantum Chromo Dynamics, LQCD) 是解决粒子物理、核物理各种重大物理科学问题的关键理论支撑之一。LQCD 研究最早是从 20 世纪 80 年代初开始的，随着国内计算条件的改善，本世纪初中国逐渐开始开展 LQCD 的数值模拟研究。中国格点合作组 (China Lattice QCD Collaboration, CLQCD) 由中科院理论所、北京大学、南开大学、浙江大学和中科院高能所等发起，研究兴趣涵盖强子谱学、强子结构、QCD 相变、高精度前沿、软件架构研发等方面，并与国际上重要的格点研究团队有密切的合作关系。本应用软件是从零开发的国产格点 QCD 计算软件，功能是格点 QCD 中 PDE (Dirac 方程) 的求解，即传播子的求解器。已实现 CG、BiCGSTAB 以及基于奇偶区域分解预处理的 BiCGSTAB 等求解算法。软件已完成鲲鹏百节点环境的适配和调优，强弱扩展性可达 90%, 80%，持续优化各项性能。

### 3.6.9 中国科学院海洋环流模式 LICOM

海洋环流模式 (Ocean general circulation models, OGCMs) 是定量描述海洋现象及其变化的数值模型，通过求解复杂的方程对海洋系统进行建模和模拟，预测在未来某个时刻海洋的括温度、盐度、密度、流速、海表高度、混合层深度等重要因素。

中国科学院研制的海洋环流 LICOM (LASG/IAP Climate system Ocean Model) 为国内唯一自主研发的全球气候海洋环流模式。LICOM 的最高分辨率 (3 公里) 和模拟性能与国际先进水平相当, 参与历届国际模式比较计划 CMIP (1-6)。我们初步实现了自主海洋模式在国产计算平台上的异构并行, 在“先导 1 号”上实现全机模拟 LICOM3-HIP 是入围 2020 年 CCF 先进计算最佳应用奖的 4 个应用之一。软件已完成鲲鹏百节点环境的适配和调优, 百节点强扩展性高于 91%, 并行效率优化与性能优化持续进行中。

### 3.6.10 中国空气动力研究与发展中心 NNW 系列软件

#### 1) 流动之星 NNW-FlowStar

空气动力学被誉为飞行器设计的先行官, 型号研制, 气动先行, 而计算流体力学 (CFD) 是空气动力学研究的重要手段。计算流体力学仿真软件能够模拟飞行器真实飞行状态、提供详细的流场细节, 是飞行器设计的重要工具。流动之星 NNW-FlowStar 是面向航空 / 航天 / 兵器领域飞行器复杂流动模拟的 CFD 工业软件, 能够准确开展飞机、直升机、导弹、再入飞行器及其他飞行器低、亚、跨、超和高超声速定常 / 非定常气动力、动态特性、多体运动、进气道内流、喷流干扰等场景的数值模拟。软件以多体运动模拟为特色, 凝聚了航空 / 航天 / 兵器领域专业 CFD 工程师近 20 年的软件开发和应用经验。软件多次参加国际先进升阻力预测对比会议, 计算精度达到国际先进水平; 软件在国内武器装备的各类真实复杂型号应用中历经考验, 计算鲁棒, 精度可靠, 效率优良。目前国内 40 多家工业部门和科研机构已安装并使用该软件, 同时在天津超算、广州超算、成都超算等国家超算中心部署安装, 提供仿真计算服务。

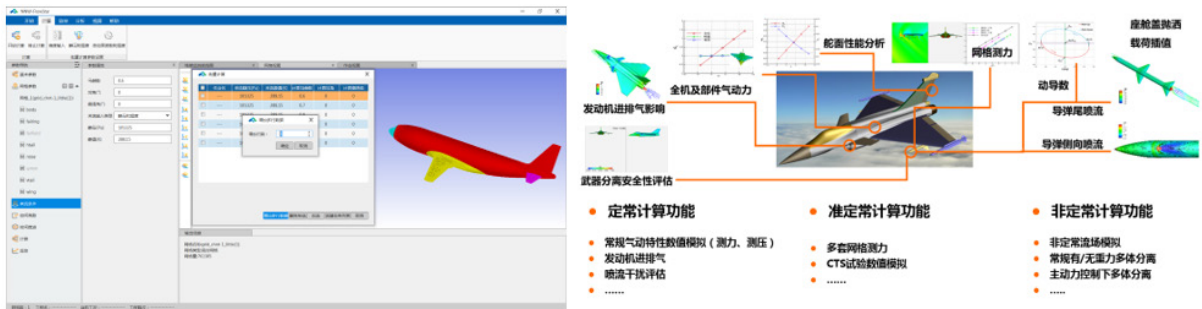


图 3-3 NNW-FlowStar 操作界面及应用

#### 2) CFD 仿真软件 NNW-HyFLOW

高超声速飞行器具有巨大的军民两用价值, 是当今世界各国在航空航天领域的重要发展方向。高性能计算流体力学 (CFD) 模拟与高超飞行试验、地面高焓设备试验研究相互印证, 是高超飞行研究的主要手段之一, 在下一代新型高超飞行器研制中将发挥主导作用和更大价值。NNW-HyFLOW 软件是国内



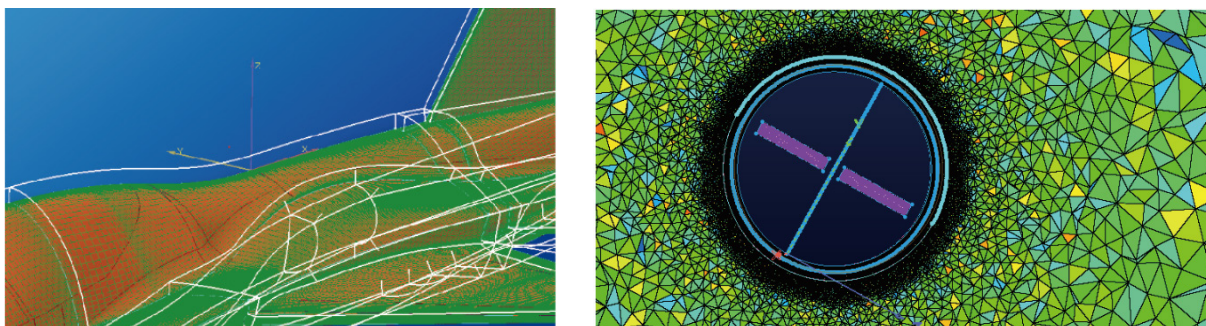


图 3-6 网格之星 NNW-GridStar 示意图

#### 4) 工业级可视化软件 NNW-TopViz

基于计算空气动力学（CFD）的流体仿真在航空航天、交通运输、能源动力、桥梁建筑等领域发挥着重要基础作用。随着高性能计算发展，CFD 模拟能力不断提升，解决的问题规模越来越大，模拟精细度越来越高，需提供与 CFD 能力相匹配的数据分析能力和可视化能力。NNW-TopViz 是由中国空气动力研究与发展中心研制，可替代国外主流软件的工业级可视化软件。软件支持多种主流的流场数据格式，集成流线、等值面、纹理、体绘制等常用标量场 / 矢量场可视化功能，提供丰富的流场数据表达方法，具备图表、多视口、数据探针等数据分析能力，采用线程并行、进程并行、数据区域分解等技术，支持海量流场数据并行处理。NNW-TopViz 能够对 CFD 数值模拟生成的流场数据按需进行可视化计算和分析，为 CFD 研究人员提供对流场数据及其几何外形的显示控制和交互功能，可以直观、有效地洞察流场数据的内涵与本质，快速、准确、全方位地把握流场要素特征信息、流动机理及变化规律。软件已完成 ARM 指令集兼容架构处理器高性能计算平台的部署适配，实现了百亿网格流场数据并行可视化和交互式分析，万核并行效率达到 30%，并持续优化性能和效率。

#### NNW-TopViz 交互界面及部分功能展示

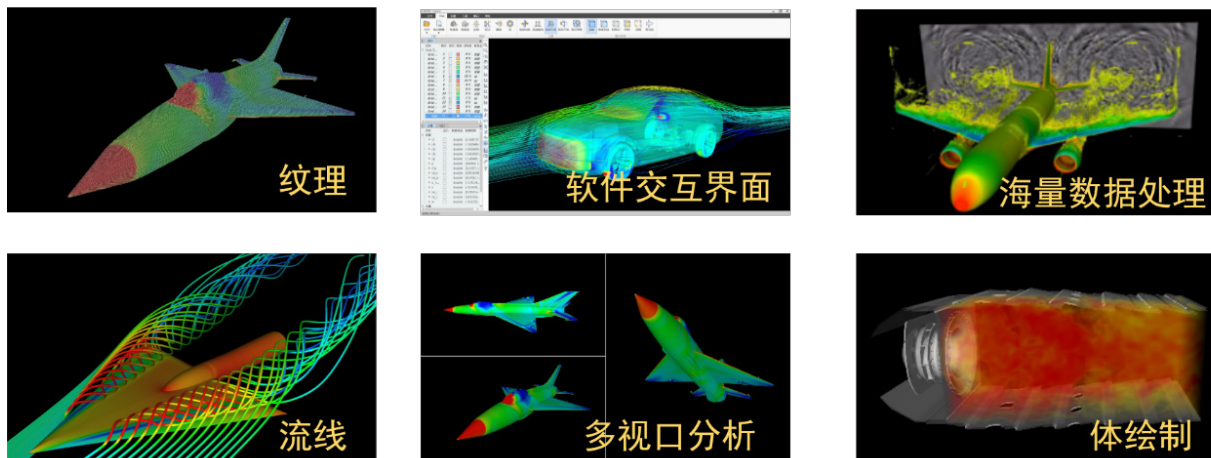


图 3-7 NNW-TopViz 交互界面及部分功能展示



### 5) 通用 CFD 软件 NNW-PHengLEI

NNW-PHengLEI 软件是中国空气动力研究与发展中心计算空气动力研究所研发的一款同时适合于结构网格、非结构网格，具备多学科计算扩展能力和超大规模并行能力，应用范围涵盖低速、亚跨超、高超声速的通用 CFD 软件。软件具备结构 / 非结构求解器混合计算、重叠网格计算、大涡模拟、多组分气体扩散模拟等特色功能，具备百亿网格百万核并行、200 亿混合网格自动并行生成等能力，2021 年风雷软件勇夺第九届全国并行应用挑战赛冠军。在国家数值风洞工程的支持下，风雷软件于 2020 年 12 月面向全国开源，在互联网“红山开源平台”开放代码，是国内首款工业级 CFD 开源软件，在天河二号、天河三号 E 级原型机、天河新一代上完成适配与优化，取得良好的并行可扩展性。目前共 100 余家单位申请使用风雷软件，社区用户数 1500+，已有 21 个单位基于风雷产生 36 项应用与二次开发成果，用户基于风雷软件开展学术研究发表数十篇学术论文，同时孵化形成了多款工业定制化软件，构建起了蓬勃发展的生态，促进了国内 CFD 软件的整体发展。

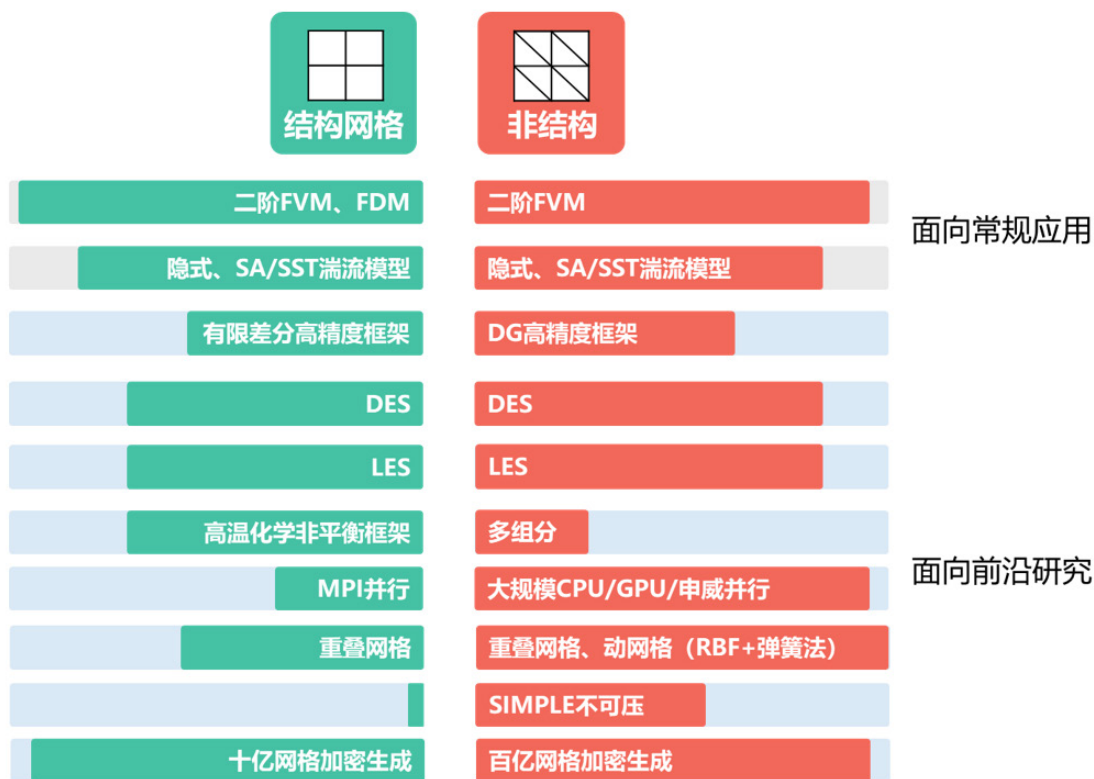


图 3-8 NNW-PHengLEI 软件主要功能



### 3.6.11 国防科技大学银河矩阵计算软件 YHMatrix

矩阵计算软件广泛用于设计建模、数值分析、信号处理等科学研究与工程计算领域，是快速实现算法验证和生成原型软件的重要编程工具。国防科技大学研制的银河矩阵计算软件（YHMatrix）是一款适用于科学研究、工程设计等众多应用场景的矩阵计算软件，基于天河超算应用支撑环境、Octave、SCILIB 开发，采用开放、开源、众包和在线的开源 2.0 模式进行迭代开发。YHMatrix 软件旨在解决 Matlab 禁运和技术封锁问题，满足 Matlab 遗留代码快速适配需求，弥补工业软件水平差距；同时，缓解现有国产矩阵计算软件并行计算能力弱、性能效率不足问题，发展基于国产超算的科学计算、可视化以及交互式程序设计软件。YHMatrix 软件已提供基础 GUI 界面，用于求解典型线性和非线性数学问题。软件允许使用接近 Matlab 语言的语法进行数值计算实验，具有较好的易使用性、高效性、可维护性和可一致性，能够为各领域科学家与工程师提供优质、可靠的科学计算环境，软件的在线版、计算核心性能优化及工具箱正持续进行中。

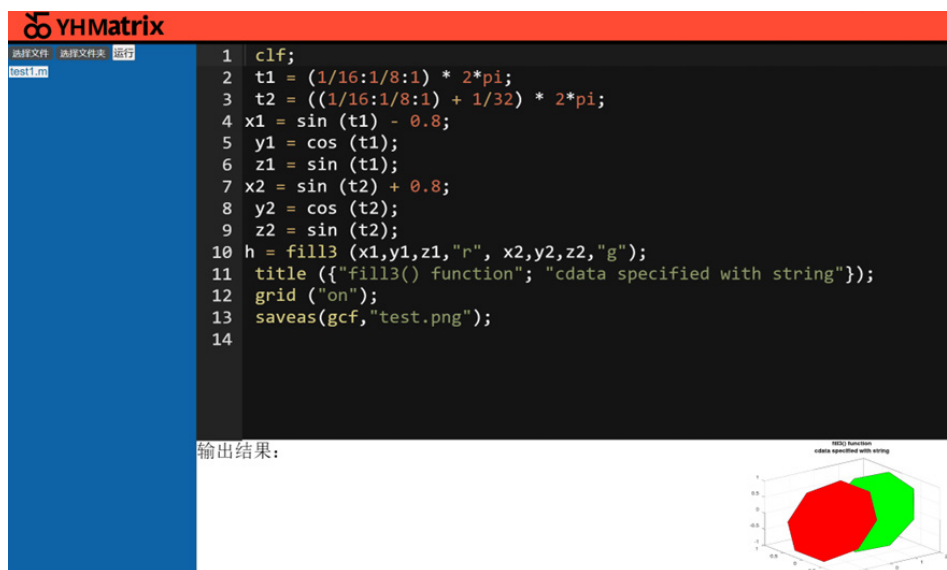
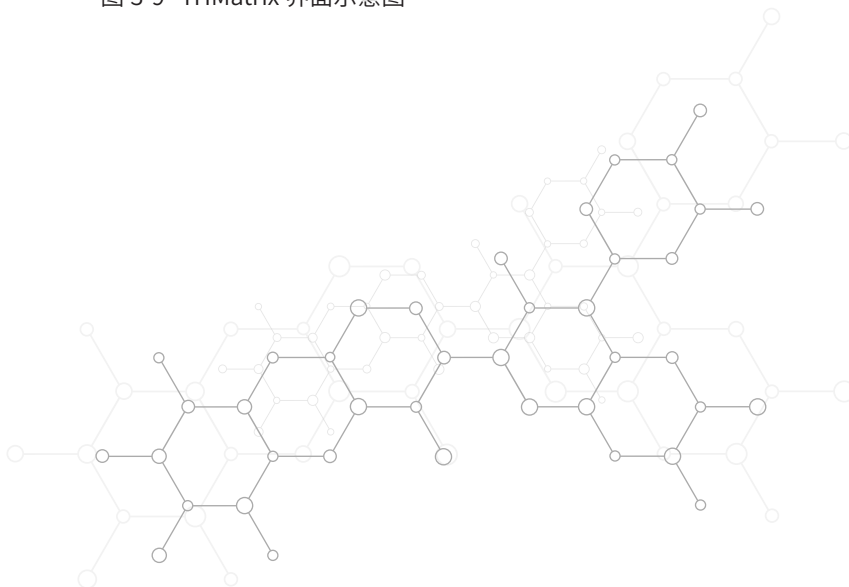


图 3-9 YHMatrix 界面示意图



### 3.6.12 国防科技大学不可压缩流体计算软件 YHACT

以精细化计算流体力学模拟为核心的热工水力分析在核反应堆设计、安全风险评估等领域具有重要意义。国防科技大学研制的通用计算流体力学软件 YHACT 主要面向复杂工程应用中的流动、传热模拟，通过新型的面向对象软件架构，提供高可扩展的数值离散格式与物理模型的扩展。YHACT v1.20 已具备多种 RANS 湍流模型与大涡模拟模块，支持高精度传热以及流体 - 固体共轭传热模拟，可选压力速度分离式求解与高效的全耦合求解算法。在并行计算方面，具有相对主流商业软件更好的并行效率，可实现亿级以上网格规模的复杂工程案例。该软件目前已在中国核动力研究设计院等单位进行应用验证，并适配天河超算等高性能计算平台，在 2048 核以上规模的并行测试中可扩展性良好。

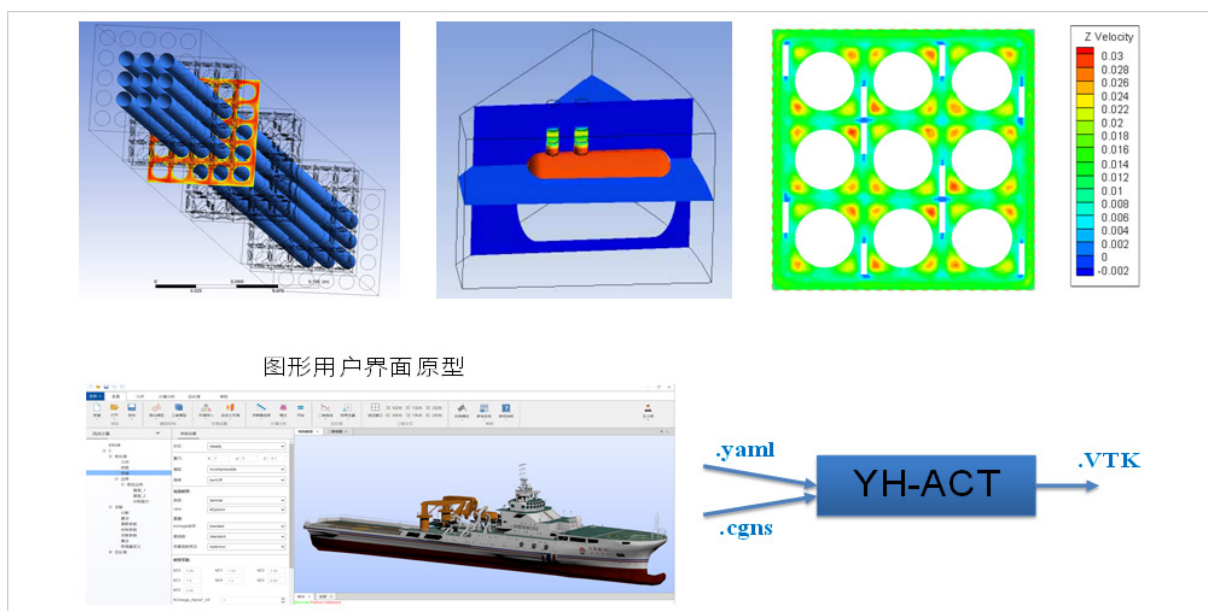


图 3-10 YHACT 效果示意图

### 3.6.13 国防科技大学银河通用网格生成软件 YHGRID

网格生成是计算流体力学、结构力学仿真、电磁仿真等数值模拟领域必需的前处理步骤，它将仿真对象的计算域按要求离散成基本的几何单元集合，才能在离散单元组成的网格上进行有限差分、有限体积和有限元等数值计算。因此，网格生成是进行 CAE 仿真的基础。国防科技大学研制的 YHGRID 是一款自主研发的通用网格生成软件，致力于为工程计算提供高品质的网格生成功能，尽可能降低工程师和科研人员在前处理的时间消耗。YHGRID 支持 IGES、STEP、STL、CGNS 等通用格式的数模和网格导入导出，支持数模的一键式修复和基本几何的创建，支持三角形 / 四边形单元的表面网格和四面体单元的体网格快速并行生成，提供附面层网格生成、网格质量评价等功能，目前正在开发笛卡尔网格生成、结构网格生成等功能。软件已完成天河超算环境的适配和调优，能轻松生成数亿规模网格，并行效率优化与性能优化持续进行中。

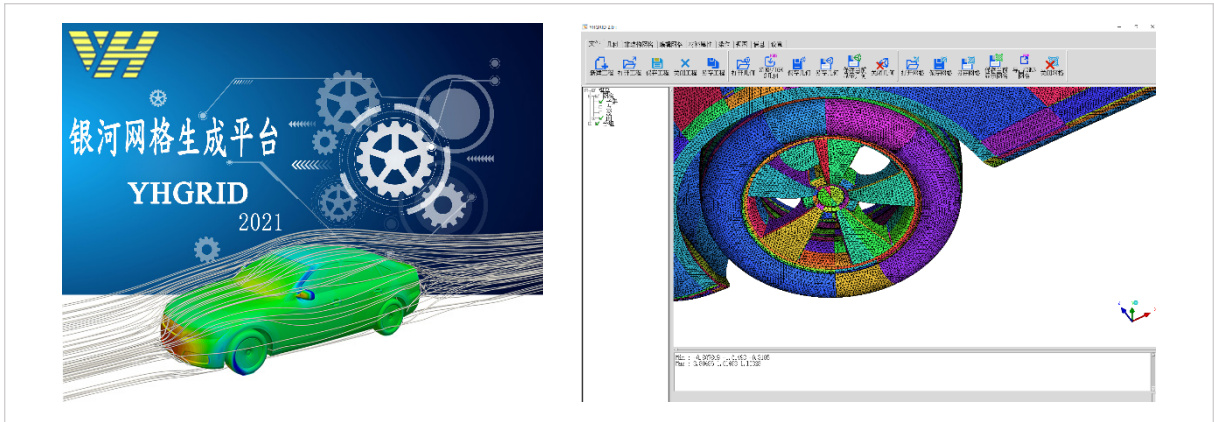


图 3-11 YHGRID 效果示意图

### 3.6.14 西安电子科技大学计算电磁软件 LASCPEM

在电磁设计领域中，电磁计算，尤其是电磁精确计算，是当前及未来信息化系统电磁分析与设计的共性基础技术。西安电子科技大学研制的 LASCPEM (Large Scale Parallel Computational Electromagnetics) 软件针对电磁算法中矩量法产生的复数稠密矩阵方程和有限元法产生的复数稀疏矩阵方程，采用单向通信 CALU 和超节点法，基于鲲鹏 920 高性能计算平台研制了复数稠密 / 稀疏线性系统并行直接求解器，突破直接求解算法的负载均衡、并行选主元策略、细粒度任务映射、通信隐藏等关键技术，满足高精度电磁算法对大型矩阵方程求解的需求，确保其在大规模并行计算时稳定可靠。未来可以将自主研发的电磁计算软件、矩阵方程求解器、国产超级计算机的结合，能够为解决我国电磁设计领域卡脖子难题提供重要的高性能计算技术支撑作用，有效打破国外电磁商业软件的市场垄断。除用于电磁计算软件调用外，本软件还能用于其它需要处理稠密 / 稀疏矩阵方程直接求解的计算领域，为国产高性能计算平台多元化生态应用建设提供重要的数学基础工具。软件已完成鲲鹏百节点环境的适配和调优，百节点强扩展性达 64%，并行效率优化与性能优化持续进行中。

### 3.6.15 自然资源部第一海洋研究所海浪模式 MASNUM

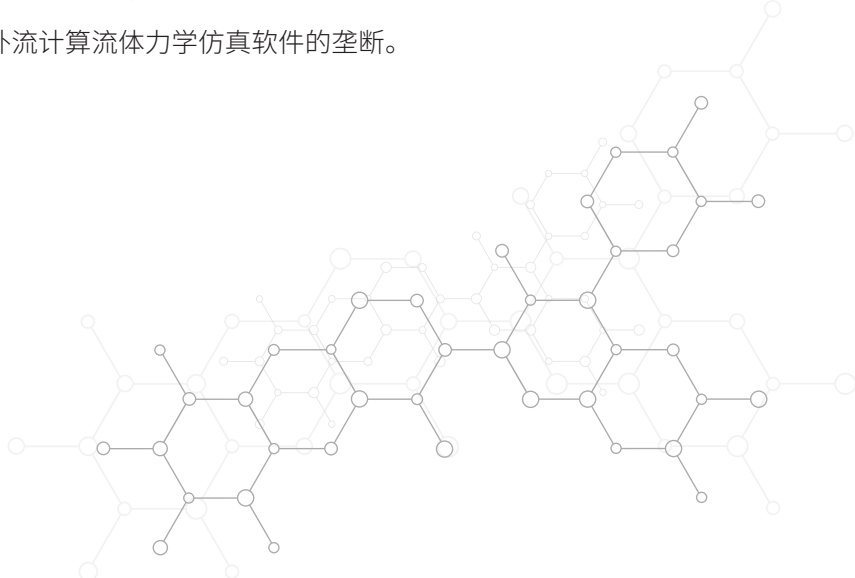
气象海浪领域的科学研究对城市发展、人民生活、社会经济、国防安全都影响深远。特别是近年来频繁发生的极端天气和海洋灾害，不仅给对生产生活造成严重影响，甚至会有不可估量的生命和财产损失。同时，国家重要赛事、政治经济活动以及社会公众都对天气情况及时性、准确性的需求日益增长，这也给气象海洋预报提出了更精细、更准确、更及时的要求。现在的气象数值预报模式是根据大气当前状态，给定初边值条件，基于大气状态方程、流体力学、热力学方程等，利用超级计算机进行数值方法的求解。

自然资源部第一海洋研究所研制的 MASNUM (Marine science and numerical modeling) 海浪模式是我国自主发展的第三代海浪数值模式,通过特征线求解波数谱平衡方程,并考虑风输入、波浪破碎耗散、底摩擦耗散、波波非线性相互作用和波流相互租用等源函数,模拟波数空间下的海浪方向谱,获得海浪的波高、周期和波向等要素。

海洋防灾减灾、海洋生态保护等都需要精准的海洋环境模拟与预报。对预报而言,海洋数值模式相当于工业的“芯片”,是国家海洋科技实力的综合体现。海洋数值模式的研究水平对于促进海洋经济发展、应对海上突发事件、加强海洋防灾减灾具有十分重要的意义。但半个多世纪以来,海洋数值模式的发展一直由欧美等海洋科技发达国家主导。MASNUM 海洋环流模式对提升我国海洋自主创新科技实力具有重要意义。目前软件已完成鲲鹏百节点环境的适配和调优,百节点强扩展性高于 80%。

### 3.6.16 航空工业西安航空计算技术研究所外流仿真分析软件 WiseCFD-UG

计算流体力学 (CFD) 仿真技术是航空、航天、兵器、高铁等军民制造业装备研制过程中不可或缺的关键共性支撑技术。尤其随着数字化设计认证时代的到来,CFD 在上述装备从设计到仿真认证全生命周期均将发挥至关重要作用。本软件瞄准支撑航空领域的动态气动力仿真、多学科优化设计等仿真需求,依托西安航空工业计算技术研究所长期的技术积累,历经 10 余年技术积累和持续开发。结合无人值守式嵌套网格预处理工具软件的开发,打造了一款高保真、高效率,面向从低速到高速的全流域定常/非定常气动性能分析、多体分离、动态特性和气动弹性分析等应用场景的外流仿真分析软件 WiseCFD-UG,并提供基于伴随方程的气动及以气动为核心(如气动-结构)的多学科优化设计能力。核心求解器采用兼容任意多面体网格和嵌套网格的非结构混合网格技术和有限体积方法,支持基于 MPI 的超大规模并行计算和基于 MPI-CUDA 的异构并行计算。软件已在海光 x86、ARM 指令集兼容架构处理器等国产系统完成适配,并在天河 2 号安装试用,核心模块成果有效支撑了“十二五”自主可控万核级混合网格并行 CFD 软件的研制。未来可将该自主可控的核心求解器软件、嵌套网格预处理工具软件与国产超级计算机系统结合,在解决动态仿真和以气动为核心的多学科仿真等“难、卡”特色场景应用中发挥关键替代作用,有效打破国外外流计算流体力学仿真软件的垄断。



### 3.6.17 中国核动力研究设计院 MOI-X 与 MORE

通用堆芯蒙特卡罗运行集成工具主要用于反应堆全堆芯蒙特卡罗模拟计算，通过开发临界 - 燃耗、临界搜索、控制体微积分价值、堆芯装载等多个工程功能，为核电领域提供强有力的核设计与验证工具。中国核动力研究院研发的通用堆芯蒙特卡罗运行集成工具 MOI-X，基于全国产天河超算平台 ARM 兼容处理器上采用 MPI、OpenMP 混合同行对大规模粒子输运并行的优化，计算性能提升 10 倍以上，大规模提高设计和验证效率。目前已应用于多个型号工程堆芯的设计或验证中。

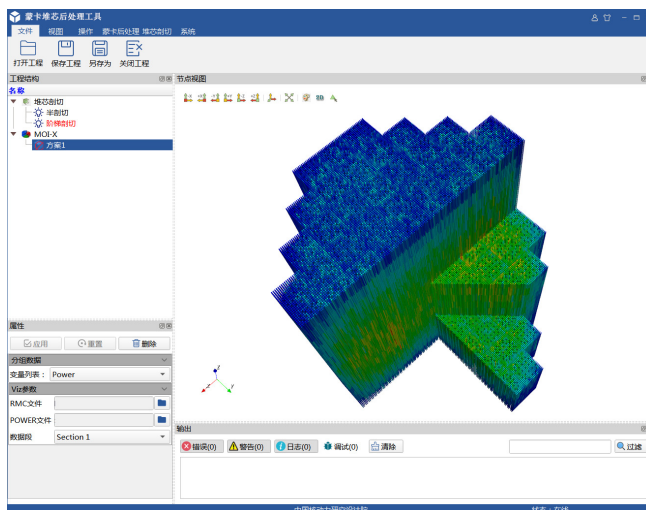


图 3-12 MOI-X 效果图

反应堆多物理多尺度耦合框架主要用于核反应堆领域，通过对耦合计算的共性技术进行抽象，形成以耦合流程控制、网格映射、迭代算法库、程序封装为基础的耦合框架 MORE，搭配良好的人机交互界面，能够针对多种需求的反应堆多专业耦合计算高效地实现耦合流程的搭建和计算。该框架基于天河超算 ARM 兼容处理器上进行深度定制优化，实际计算性能优于同类 X86 平台，广泛应用于堆芯级耦合计算和系统级耦合计算，大幅度提高了实现耦合计算软件的开发效率和耦合功能拓展能力。

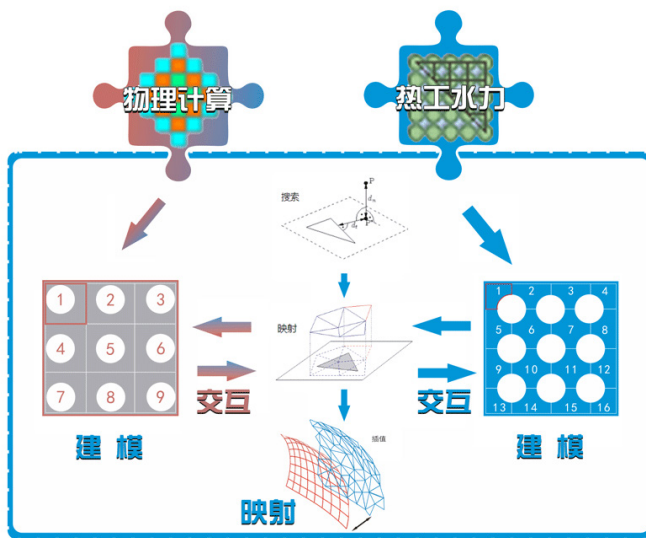
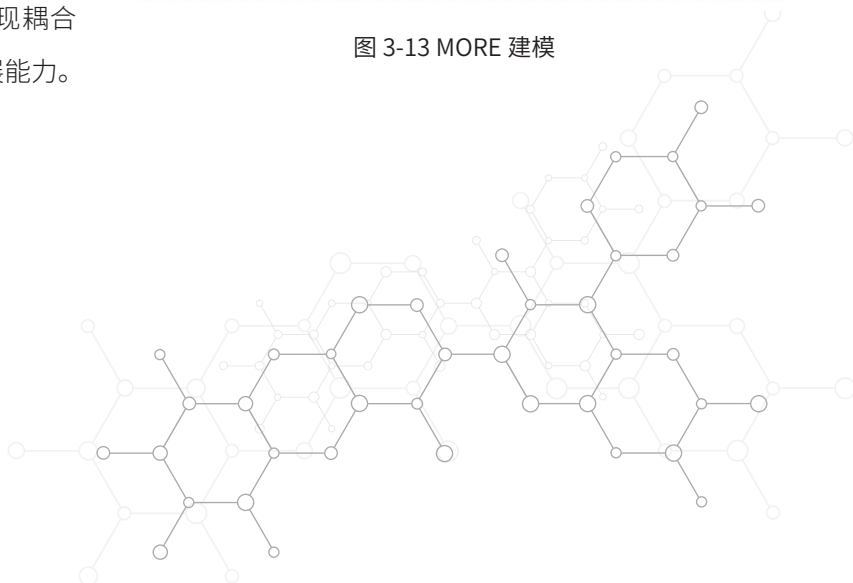


图 3-13 MORE 建模





### 3.6.18 西北工业大学航发仿真应用计算软件 AES

航空发动机 (Aero-engine) 作为航空器的“心脏”设备，被誉为“工业皇冠上的明珠”，是世界上公认的总体技术水平最高、核心技术封锁最严、结构最复杂的工业产品。高性能航空发动机的设计研发不仅与整个工业体系的发展紧密相连，同时与数字及计算技术的发展密不可分。结合高性能计算技术、计算流体力学、计算传热学、计算燃烧学等多学科耦合形成的数值发动机仿真技术，在航空发动机设计研发、产品定型验证、安全稳定运行、健康管理、故障排故维修等发动机全生命周期中可以起到缩短产品研发周期、加速技术迭代更新、加快产品换代升级速度等效果。AES (Aero-Engine numerical Simulation) 是从零开发的国产航发仿真应用计算软件，所有代码自主可控。软件基于全新的迭代式耦合 FVM 算法、区域级组合重构算法，融合实现了多学科多物理场耦合计算；同时面向超大规模异构计算平台，发展了多层级可扩展并行拓扑、隐式高效通信模型，已实现了对航发内流压气机、涡轮等部件的全环非定常场景以及百亿非结构网格规模、百万计算核心的并行加速应用。软件已完成天河新一代数千节点环境的适配和调优，弱可扩展性可达 80%；同时针对新型多样性融合算力架构，扩展有针对性稀疏求解系统的混精算法以及 AI+HPC 的湍流模拟场景，相应性能持续优化中。

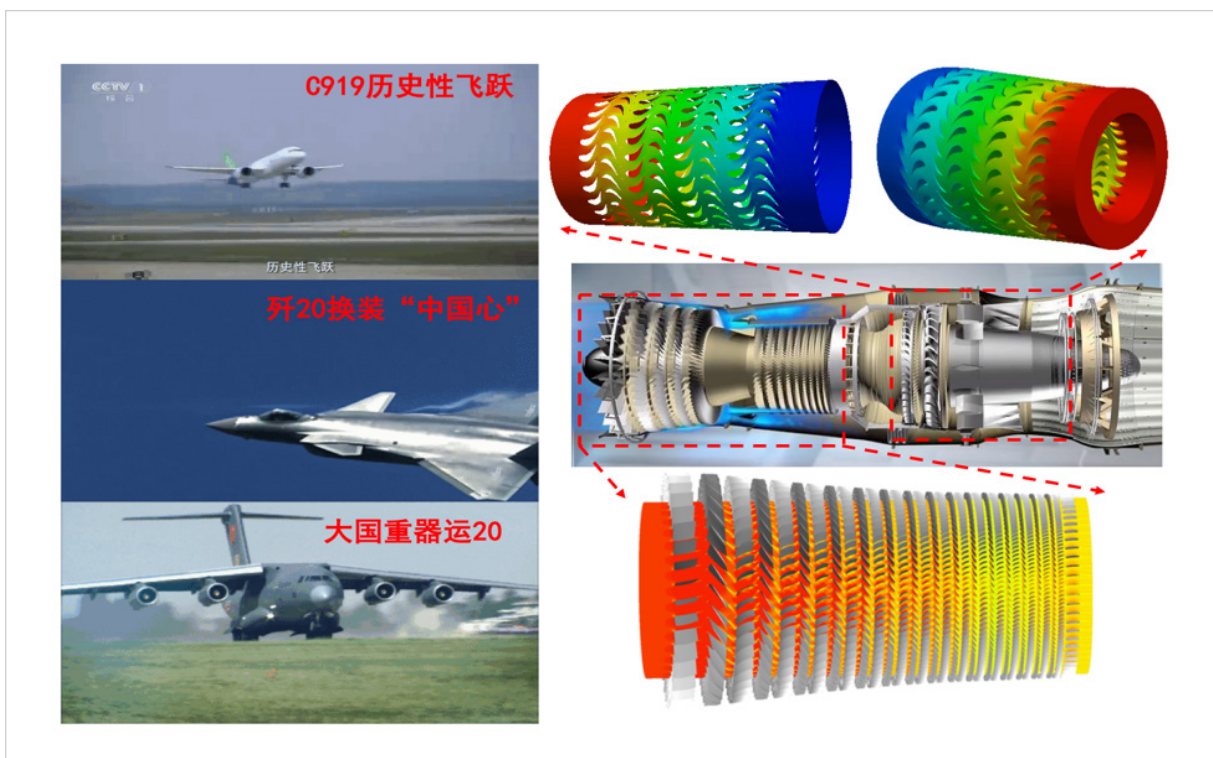


图 3-14 国产航发仿真应用计算软件 AES 应用

### 3.6.19 北京理工大学全波数值电磁仿真软件 SINOCOM

中算电磁仿真软件是一款国内完全自主可控的全波数值电磁仿真软件，主要服务于超电大复杂目标特性和隐身设计、大型天线阵和天线罩分析和设计、超电大平台电磁环境精确预测等应用。该软件在国产三大类型高性能平台上都成功开发了并行版本。该软件基于非共形区域分解合元极方法。该方法有机结合了有限元、边界元、多层快速多极子方法，兼有通用、精确、高效特点。尤其是并行多层快速多极子（multilevel fast multipole algorithm, MLFMA）算法是解决超电大目标电磁散射快速精确计算的核心技术。其并行尤为复杂，极具挑战。针对大规模、万核以上并行应用，中算研发团队提出一种新型多层快速多极子三元并行方案，并针对百亿未知数模拟的超大规模问题的网格产生难题与多极子树结构存储瓶颈，提出并实现了基于辅助树的多层快速多极子百亿网格并行细剖生成技术以及多层快速多极子树结构混合存储，在国际上首次实现了百亿、万波长电超大金属目标的精确计算。相关技术被集成为“中算”电磁仿真软件电大金属目标电磁散射计算模块和薄涂覆阻抗边界目标电磁散射计算核心模块。本软件能够为解决我国电磁设计领域卡脖子难题提供高性能电磁仿真软件支撑作用，有效打破国外电磁商业软件的市场垄断，支撑飞行器隐身设计与目标特性分析重大国防应用。软件已完成天河超算环境的适配和调优，在天河新一代上 16384 节点强扩展性高于 60%。

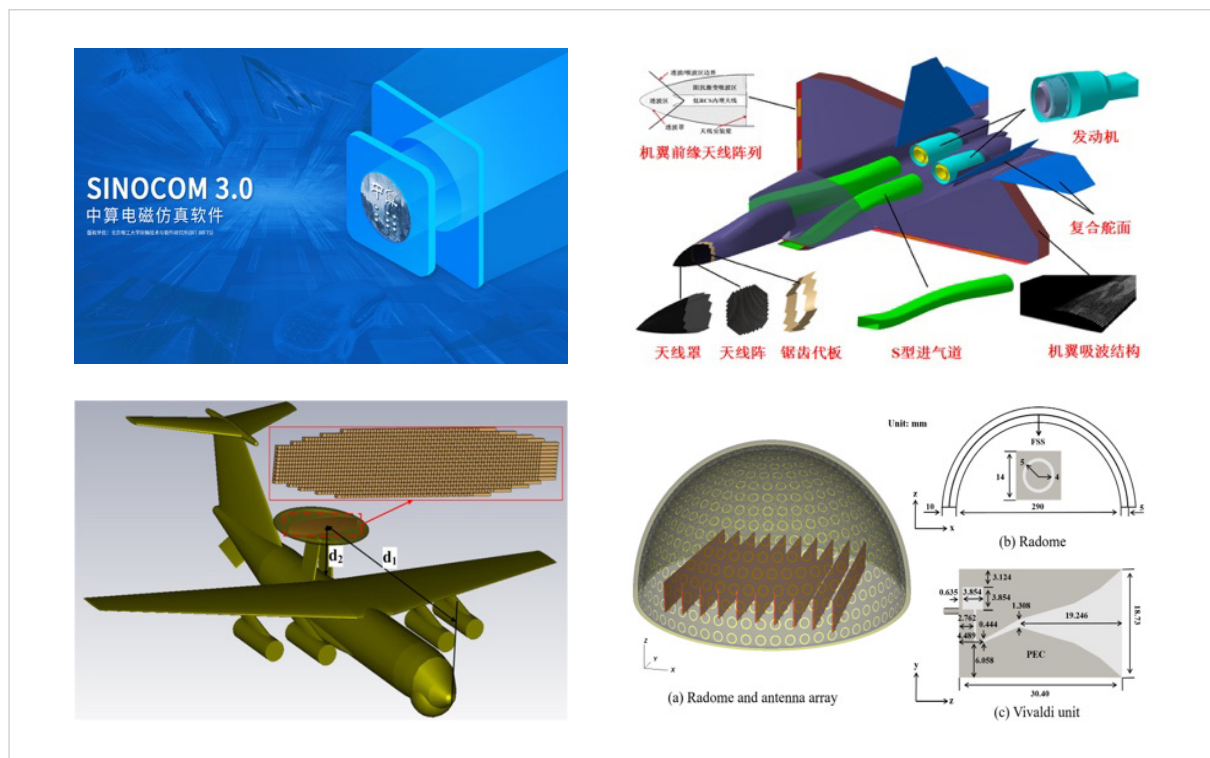


图 3-15 全波数值电磁仿真软件 SINOCOM 示意图



开 放 架 构 H P C 技 术 与 生 态 白 皮 书

# 用户篇 USER

## 4 用户篇

ARM 指令集兼容架构已在多种类型的头部客户获得深入成熟使用。

### 4.1 上海交通大学

上海交通大学“交我算”校级计算平台，于 2020 年 12 月在上海交通大学闵行校区的网络信息中心上线了国内高校首台基于鲲鹏处理器的校级超算系统，在 2021 年 4 月完成性能测试后面面向校内开放使用。

“交我算”鲲鹏超算共 100 个计算节点，节点采用双路华为鲲鹏 920 处理器（64 核），每个计算节点拥有 128 核处理器和 256GB 内存，总计 12800 核，系统的理论双精度峰值性能达 133TFLOPS。

“交我算”鲲鹏超算目前面向用户提供 18 款常用的高性能计算软件开放使用，覆盖了材料科学、生命科学、大气科学和流体力学 4 大高性能计算应用领域。

表 4-1 “交我算”鲲鹏超算平台上部署的软件

应用领域	软件名	软件简介
材料科学	GROMACS	研究生物分子体系的分子动力学程序包
	VASP	量化计算软件
	Amber	分子动力学模拟软件
	LAMMPS	桑迪亚实验室开发的分子动力学模拟软件，用于原子分子大规模并行模拟
	LAMMPS-RBE	上海交通大学徐振礼课题组基于 Lammps 二次开发的大规模并行版本
	Quantum ESPRESSO	纳米尺度上进行电子结构计算和材料建模的软件
	CP2k	基于密度泛函理论的从头算分子动力学软件
生命科学	BCFtools	VCF 和 BCF 文件的工具合集
	Blast-plus	基于局部比对算法的搜索工具
	DELLY	一种集成的结构变异（SV）预测方法
	HISAT2	快速、灵敏的比对软件
	HMMER	序列比对软件
	Manta	检测单个样本的 germline 变异和 tumor/normal 配对样本的 somatic 变异
	GATK	二代重测序数据分析软件
大气科学	BWA	一款将序列比对到参考基因组的软件
	WRF	中尺度天气预报模式
流体力学	CMAQ	中尺度天气预报模式
	OpenFOAM	计算流体力学类库

2021年4月开放使用至今，鲲鹏集群的平均利用率约为41.2%，共服务了校内74个课题组，计算作业数量超100万个。重点用户与“交我算”提供的软件支持情况紧耦合，主要来自于交大的材料学院、生命科学学院、航空航天学院、机动学院、船建学院、物理学院和化工学院。其中，生命科学学院韦朝春教授，利用GATK在平台上开展了大量面向水稻的基因组学计算；材料学院孔令体和朱虹教授，利用LAMMPS在平台上开展了大量材料基因组高通量计算；航空航天学院印子斐副教授，利用OpenFOAM在平台上开展了大量面向湍流建模的流体力学计算。

## 4.2 兰州大学

兰州大学超算中心鲲鹏HPC集群2020年6月投入使用，一共20个节点，每个节点采用两颗华为鲲鹏920处理器，每颗主频2.6GHz，48核心，512GB内存，总核数1960，节点之间为100Gb Infiniband计算网络。

目前集群上部署过气象预测模式软件WRF、CESM、NEMO，分子动力学软件GROMACS、LAMMPS、AmberTools、NAMD、CP2k，第一性原理软件Quantum Espresso，生物信息软件BUSCO、HMMER、Bamtools、Trimmomatic、Mmseqs、Repeatmasker、Bwa、Griaaffe、Gemoma、Samtools、PASA、vg pack等。该集群主要用户来自大气科学学院、资源环境学院和生态创新学院。其中在比较基因组学和群体遗传学研究领域，基因对比软件Bwa在24G的基因组序列比对3.2G左右的参考基因组中计算性能较Intel 6248有30%的提升。运用WRF在降雨量预测研究中，鲲鹏920单节点计算性能比Intel 6248 CPU高出近40%的计算效率。

### 典型案例如下：

**案例 1：**大气科学学院刘鹏博士在鲲鹏架构集群下利用中尺度模式WRF对气象要素进行模拟，再利用机器学习对风电功率进行准确预报，每24小时预报一次，每次预报时长72小时。

**案例 2：**生态创新学院郑泽宇博士研究构建基于泛基因组的重测序分析工具，极大地提高了群体SV的分析速度和准确性。1139个拟南芥的重测序分析（3T数据量），使用7个鲲鹏920节点仅需43小时，资源利用率82.2%。

## 4.3 中国科学技术大学

中国科学技术大学超级计算中心瀚海20超级计算系统中建有20台华为Taishan 2280V2服务器，每台含有2颗鲲鹏920 CPU（48核，2.6GHz），256GB DDR4 2666MHz内存，采用100GE（支持RoCE）高速互联，共1920颗核心，峰值性能40万亿次/秒。该套系统于2019年年底建成。



• 1\* 25GE Switches (24 ports)

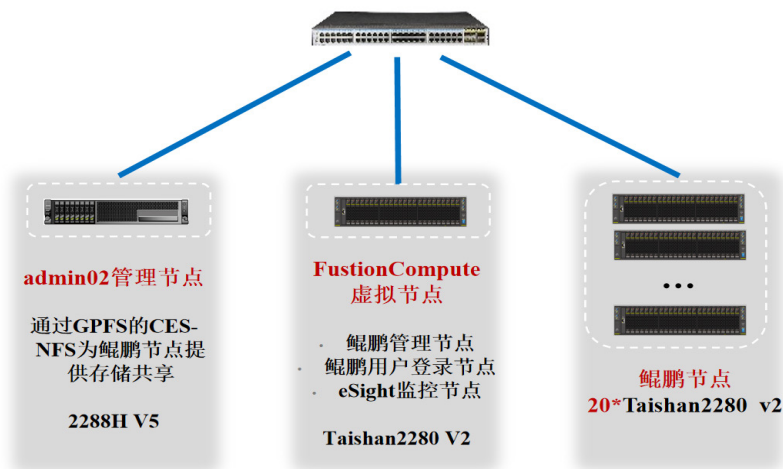


图 4-1 瀚海超级计算机网络拓扑图

目前在用账户 102 个，涉及物理学、材料科学、核科学、信息科学、地学等，从 2020 年至今，共完成作业 35 万个，累计 860 万 CPU 核小时。

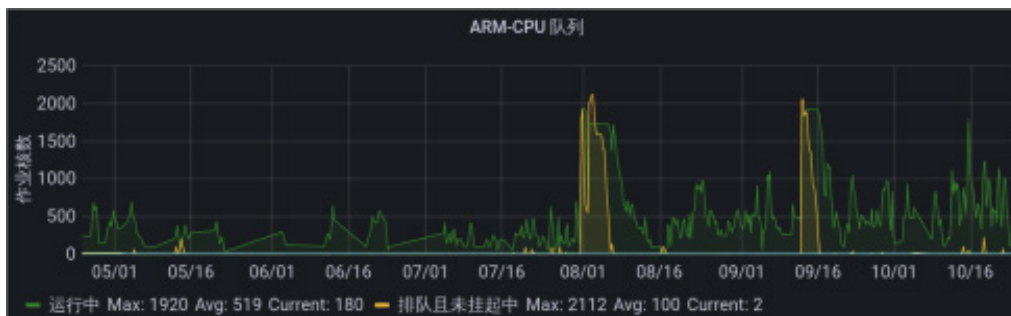
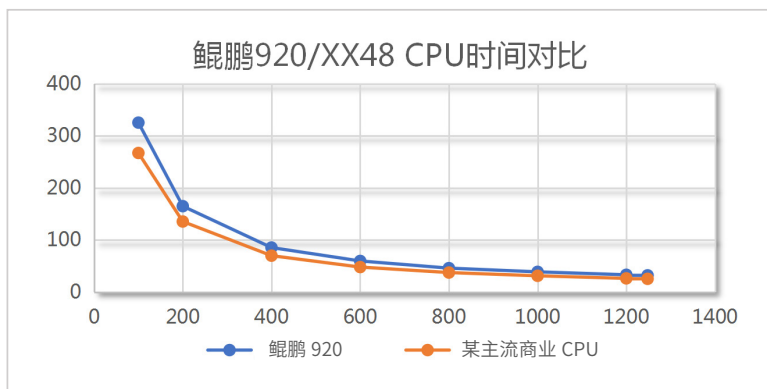


图 4-2 瀚海超级计算机任务提交情况

应用 1: 自编邮件弱密码探测，比对 8 万多个加密后密码是否在于 170 万条的弱密码库中，并行度可以到千核。虽然单核性能比某主流商业 CPU 低，约为 82%，单颗 CPU 性能为主流商业 CPU 的两倍。



横坐标为核数  
纵坐标为运行时间 (秒)

图 4-3 鲲鹏 920 与 XX48 CPU 的性能对比图

表 4-2 投影纠缠对态 PEPS(Projected Entangled Pair States) 计算模拟

应用	算例名称	规模	鲲鹏 920 CPU (秒)	XX48 CPU (秒)	鲲鹏 920 VS XX48
PEPS(Projected Entangled Pair States)	Running Job 2	8 核	12279	20416	166.27%
		32 核	11395	17641	154.81%

表 4-3 流体力学里面的层次分解

OpenFoam	Hierarchical Decompose	单节点	7838	9797	124.99%
		双节点	3788	4714	124.45%

## 4.4 国家超级计算天津中心

国家超级计算天津中心与多个国内单位合作，在天河新一代超级计算机上取得了一系列代表性成果。

国家超级计算天津中心联合国防科技大学计算机学院、海南大学药学院、中山大学药学院、中国海洋大学医药学院、美国肯塔基大学药学院等单位，在天河新一代超级计算机上完成的“基于自由能微扰-绝对结合自由能方法的大规模新冠药物虚拟筛选”工作成功入围 2021 年度的戈登贝尔新冠特别奖，这是我国首次入围该特别奖奖项。根据实际测算，在天河新一代超级计算机的支撑下，基于自由能微扰-绝对结合自由能方法的药物筛选效能提升了 200 倍。该项成果对于实现针对突发疫情的快速药物响应，具有极为重要的现实意义。

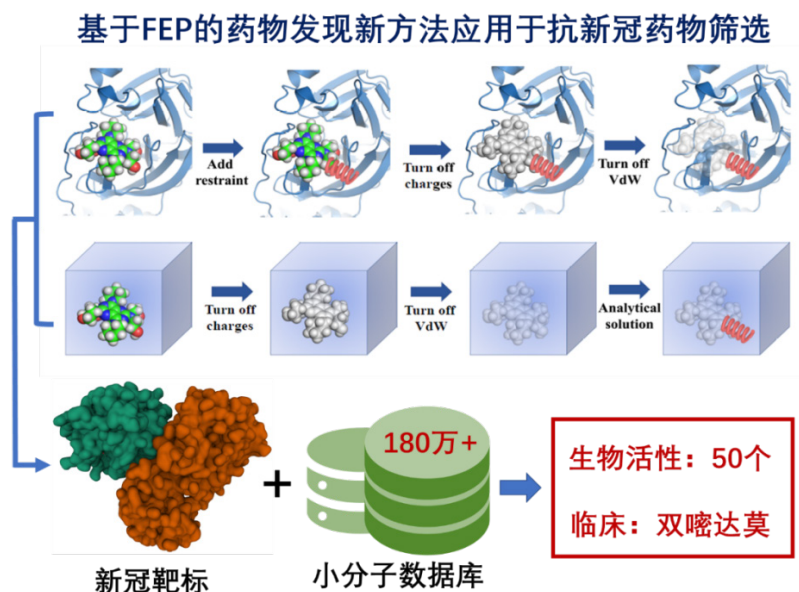


图 4-4 基于 FEP 的药物发现新方法应用于抗新冠药物筛选

基于国家超级计算天津中心的天河新一代超级计算机，北京大学开展格点 QCD 模拟并发现，在现有格点能够模拟的参数条件下，软函数的提取会受到很强的高扭度效应的干扰，计算结果甚至有可能会改变符号。为此，基于此前格点 QCD 合作组的计算结果，发展了新的计算方案和重整化方案，有效降低了高扭度效应的影响，为后续更精确的计算奠定了基础。

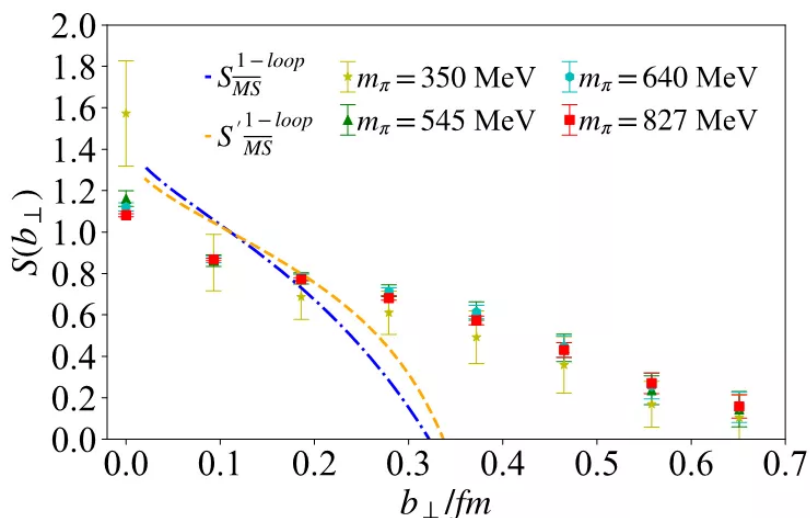


图 4-5 通过格点 QCD 计算得到的核子横向动量依赖软函数

基于国家超级计算天津中心的天河新一代超级计算机，国防科技大学、国家超级计算天津中心、湖南大学等组成的联合团队，在天河新一代超级计算机上完成了缓解能源危机进程的百万催化剂筛选。基于 HPC+AI 的材料筛选规模相比天河一号超级计算机提升 200 倍、整体效能提升 144 倍。基于 AI 赋能高通量材料计算的高效电催化筛选，最佳单日高效催化材料表面筛选数为 569 个，高于之前卡内基梅隆大学研究者一年时间筛选出的 389 个高效催化材料表面规模。

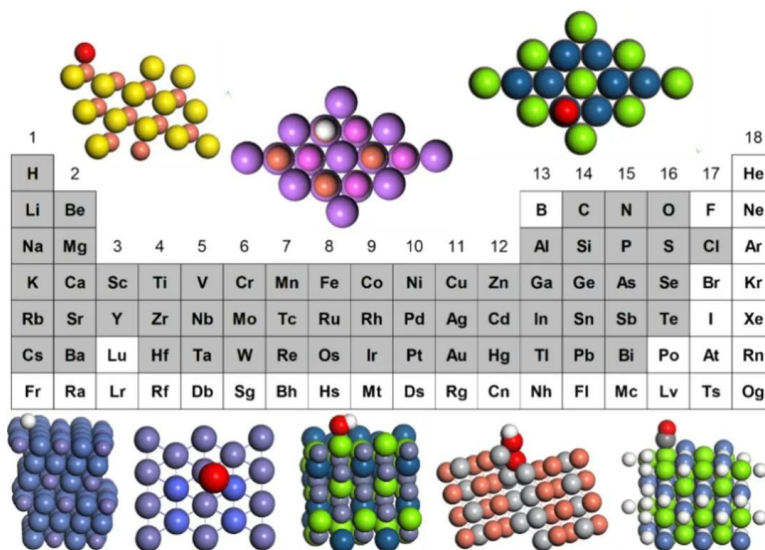


图 4-6 催化材料的元素分布及构成示意图

中国航空工业空气动力研究院（空气动力研究院）自主研发的 CFD 软件——“飞廉”在国家超级计算天津中心的天河新一代超级计算机上完成了软件移植和大规模并行测试。结果表明，飞廉软件与国产高性能计算硬件具有较好的适配性，软件移植方便；同时实现了 10 亿量级网格的万核并行测试，具备了万核并行效率 90% 以上以及 3 万核并行效率 80% 左右的并行计算能力，可以满足飞行器设计工程应用中超大规模网格高效并行计算需求。



图 4-7 飞廉软件架构设计

天津大学合成生物学团队联合国家超级计算天津中心团队基于天河新一代超级计算机在探究德布莱英图在 DNA 存储中的应用取得了突破性进展，建立了一种针对 DNA 链式存储的错误纠错解码方案，与之前发表的同类算法相比，在纠错能力和解码速度上有了突破性的提升。

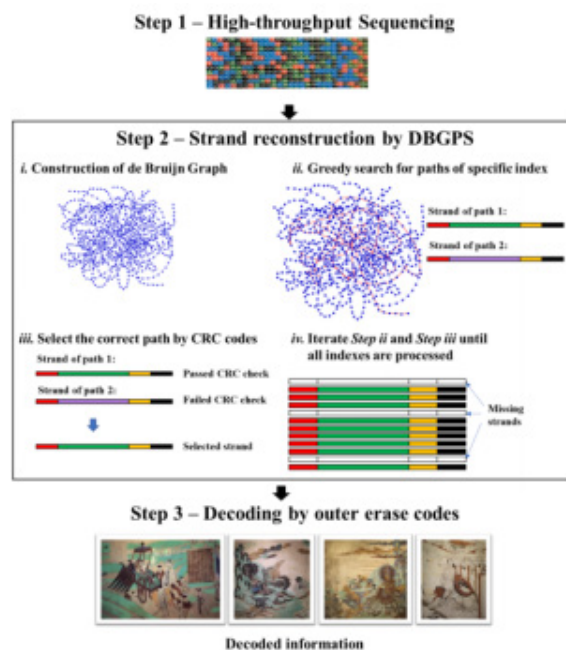


图 4-8 DNA 链式存储的错误纠错解码方案

哈尔滨工业大学（深圳）、浙江大学以及浙江大学-上海高等研究院团队基于天河新一代超级计算机系统开展真实航空发动机燃烧室内雾化燃烧过程的高精度模拟。雾化燃烧的模拟结果与实验测量进行了对比分析，包括气相场在不同截面处的速度分布、温度分布，以及液滴颗粒在不同轴向位置处的速度分布和粒径大小分布。对比结果表明模拟与实验达到了很好的吻合，但是在近壁面区域和下游区域与实验测量仍有一定误差。

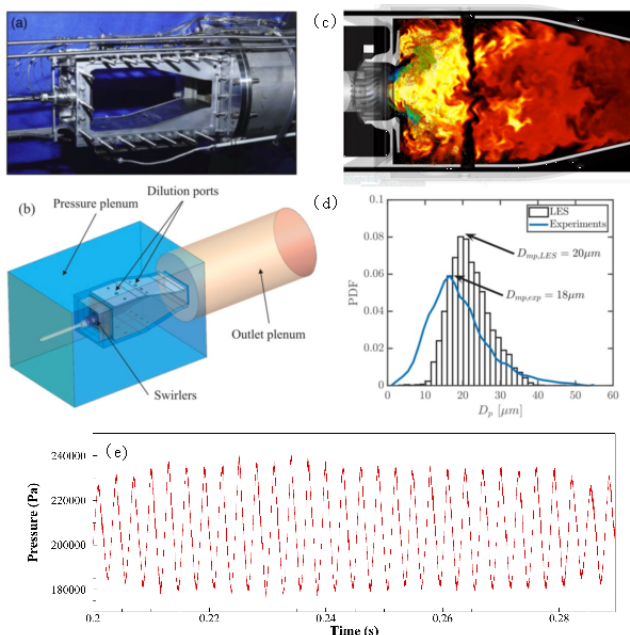


图 4-9 真实航空发动机燃烧室内雾化燃烧过程的高精度模拟

## 4.5 国家超级计算深圳中心

国家超级计算深圳中心（深圳云计算中心）（以下简称“深圳超算”）是 2009 年获国家科技部批准成立的两家国家级超算中心之一，是深圳建市以来规模最大的国家级重大科技创新基础设施。深圳超算总投资 12.3 亿元，一期建设用地面积 1.2 万平方米，总建筑面积 4.3 万平方米。深圳超算配置有国产曙光 6000 超级计算机系统，2010 年 6 月世界超算 TOP500 排名第二，运算速度达每秒 1271 万亿次。当前，深圳超算已经启动二期建设，落户光明科学城大科学装置区，计划建成全球领先的超级计算机系统。深圳超算是粤港澳大湾区国际科技创新中心和综合性国家科学中心的重要基础设施，致力于超算技术与应用创新，支撑基础科研、民生工程与产业创新等各个领域。

深圳超算是国家高效能计算机及网格服务环境节点单位，拥有电子政务云计算应用技术国家工程实验室、深圳市大数据研究院、深圳市材料云服务公共技术服务平台、深圳市工业设计云服务平台等多个创新技术研发机构。近十年来，累计服务三万个以上用户团队，完成各类计算任务逾千万个，完成 15 亿核小时计算，服务企事业单位和科研院所过万家，其中已经在深圳市科技、气象、医疗等多个应用领域形成显著特色和优势，已成为计算机资源服务形式最丰富、资源利用率最高的国家级超算中心。

深圳超算近年来承担国家、省、市科研项目二十余项，包括国家重点研发计划“国家高性能计算环境服务化机制与支撑体系研究”、“面向航天行业的工程力学优化设计平台研发与应用示范”、“基于



材料基因组技术的全固态锂电池及关键材料研发”，国家科技部 863 项目“曙光高效能计算机系统关键技术及应用”、“高性能计算环境应用服务优化关键技术研究”，国家发改委云计算重点示范工程“深圳市云计算公共服务平台”，以及深圳市技术攻关项目“大数据的组织模式和高效检索”、“基于云计算的跨平台应用关键技术研究”、“深圳市材料计算云服务公共技术服务平台”、“深圳市工业设计云服务平台”、“广东省重点领域研发计划项目”等，多项已完成并通过项目验收，其中“曙光高效能计算机系统关键技术及应用”项目获得 2013 年度国家科学技术进步奖二等奖。

深圳超算二期将建设世界一流、生态完备、稳定可靠的 E 级通用超级计算机平台和高水平的人才队伍，支撑深圳、大湾区乃至国家基础科研、社会发展和产业创新。通过建设以国产鲲鹏高性能通用 CPU 为核心的性能先进、结构合理、功能完善的通用超级计算机系统，实现大规模科学计算、工业计算和专业大数据处理等超算功能，支撑基础科研、能源材料、防灾减灾、城市安全、生命健康等领域的创新活动，为深圳市、大湾区乃至全国的科技创新提供强劲动力。

深圳超算二期 E 级超级计算机拟建设开放的公共通用算力平台，支撑数值计算、数据处理等多种类型计算，在基础科研、社会发展和产业创新等多应用场景中发挥作用，为深圳建设综合型国家科学中心提供重要支撑。为实现该建设目标，深圳超算开展了面向大规模 E 级应用场景、面向大湾区重大应用场景和面向大范围基础应用场景的三部分工作。大范围基础应用主要围绕基础软件开展适配与性能优化工作，要求基础软件覆盖范围广、计算性能高。大湾区重大应用主要面向国民经济主战场，启动十大应用场景建设工作，优先保障本地需求。大规模 E 级应用主要针对重大科技创新问题，发挥 E 级超算全系统优势，进一步加速科技前沿、国家重大需求等问题取得突破。



图 4-10 深超二期生态及应用架构图

部分鲲鹏 HPC 应用案例介绍如下：

### 1) 聚集诱导发光分子数据库

深圳超算高性能计算团队与香港中文大学（深圳）唐本忠院士团队合作，构建聚集诱导发光分子数据库。基于鲲鹏 CPU 集群建立聚集诱导分子光学性质预测模型，在使用远少于现有文献的特征维数的条件下，实现了聚集诱导发光分子材料的快速精准识别。该成果将用于共同开发的聚集诱导发光分子数据库中，成果已发表于国际知名期刊《Aggregate》。

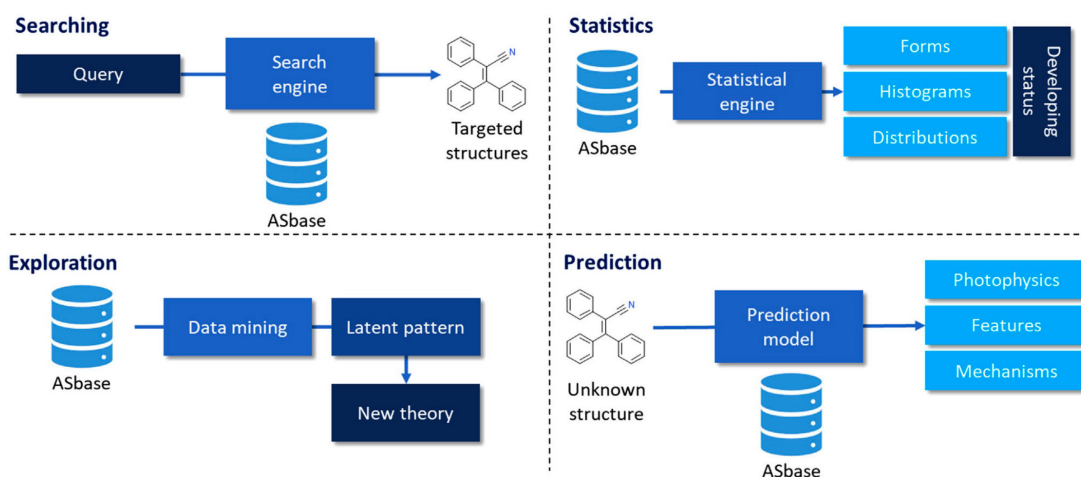


图 4-11 聚集诱导发光分子数据库总体框架

### 2) 石墨烯表面分子聚集结构研究

深圳超算高性能计算团队与南京理工大学 Herbert Gleiter 研究院合作，开展石墨烯表面分子聚集结构研究。基于鲲鹏 CPU+GPU 异构 HPC 集群研究了石墨烯表面的分子聚集结构，使用二十张 GPU 计算卡在一天内解析出具有二十万个原子的石墨烯材料表面致密聚集形态，解释了分子涂膜的抗摩擦机制，相关成果发表于材料领域 Top 期刊《Friction》。

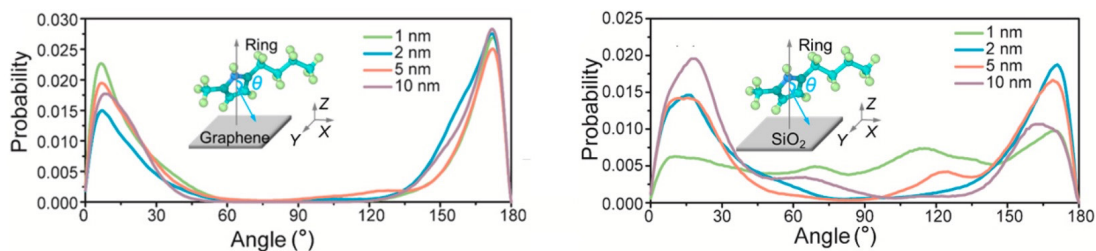


图 4-12 聚集诱导发光分子数据库总体框架

### 3) MOF 材料表面的液体吸附行为研究

深圳超算高性能计算团队与江西师范大学先进材料研究院合作，基于鲲鹏 CPU+GPU 异构 HPC 集群研究了 MOF 材料表面的液体吸附行为，对静电作用的求解过程开展性能优化工作，精确解析出电解液分子的界面相结构，为设计高性能的电极材料提供了理论指导，相关成果发表于电化学领域内 Top 期刊《Langmuir》。

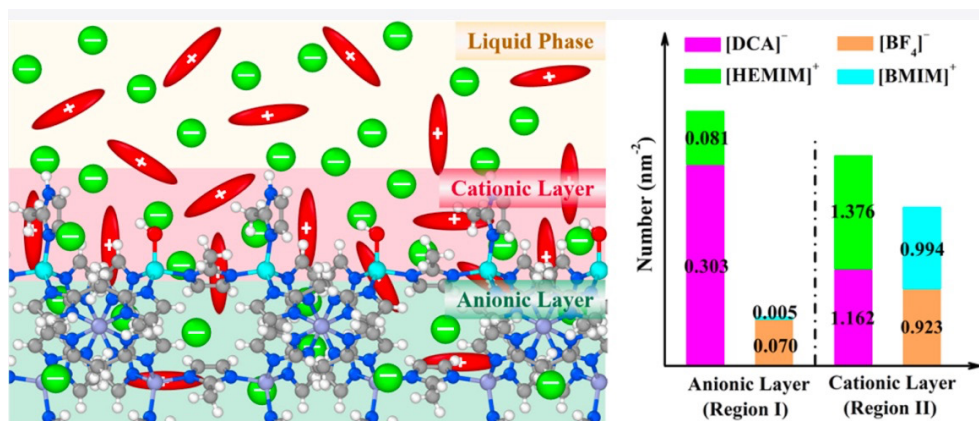


图 4-13 电极附近电解液的双层模型

## 4.6 国家超级计算长沙中心

长沙超算是第三家国家级超算中心，目前已经完成天河新一代的升级工作。中心基于国内新一代超算先进技术，采用可配置柔性体系结构和高速互连等关键技术，部署了“天河”新一代主机系统。计算部件采用国产 ARM 指令集兼容架构 CPU 和迈创 -3000 加速器。系统包含计算处理、服务处理、互连通信、全局存储、监控诊断、基础架构和辅助算力系统等部分。系统 64 位通用超算精度不低于 200P Flops，系统磁盘总容量不少于 20PB，点点双向通信带宽不低于 400Gbps。

中心部署的应用软件主要可分为开源软件和商用软件两大类。开源软件和商业软件涵盖科学计算、地理信息、三维建模、人工智能、CAD/CAE 建模及分析、流体力学仿真、动力学仿真、电磁仿真、工程仿真等领域。已经部署的开源软件如下表所示。

表 4-4 国家超级计算长沙中心已部署的软件清单

序号	软件名称	应用领域	应用场景
1	blender	3D 结构生成	建模、动画、材质、渲染、到音频处理、视频剪辑
2	Code_Saturne	CFD	大气模拟、煤粉、重质燃料及生物质的燃烧模块、电弧与焦耳效应模块、颗粒追踪模块、流体机械转子-定子互动模块等
3	Odin	Lagrangian-Eulerian	拉格朗日 - 欧拉分析
4	EPOCH	PIC	可扩展的 PIC 开放协作
5	cloverleaf	安全可靠	使用明确的二阶精确方法求解笛卡尔网格上的可压缩欧拉方程
6	cloverleaf3d	安全可靠	使用明确的二阶精确方法求解笛卡尔网格上的可压缩欧拉方程
7	beagle	贝叶斯与最大似然计算	相基因型和估算未基因型标记
8	charm	并行编程框架	用 Python 语言开发时提高其效率的工具
9	horovod	并行神经网络	TensorFlow、Keras、PyTorch 和 Apache MXNet 的分布式深度学习训练框架
10	CONQUEST	材料科学	进行对角化和线性缩放，或 O(N) 计算
11	abinit	材料科学	计算材料的光学、机械、振动和其他可观察特性
12	bigdft	材料科学	计算表面和周期性体系、带约束结构优化、Born-Oppenheimer 分子动力学
13	castep	材料科学	表面化学，键结构，态密度和光学性质等研究
14	dgemm	稠密矩阵乘性能分析	稠密矩阵乘性能分析
15	ccp4	大分子射线晶体学	通过 X 射线晶体学确定大分子结构

16	lammmps	大规模分子动力学模拟	模拟液体中的粒子，固体和气体的系综、模拟全原子，聚合物，生物，金属，粒状和粗料化体系。
17	IDL	大数据分析	数据分析与图像处理
18	cesm	地球科学	利用耦合器协同大气、海洋、陆面、海冰等分量模式进行气候模拟
19	AWP-ODC	地震模拟	模拟 3D 粘弹性或弹性固体中的波传播
20	comsol	电气科学数值模拟	力学、电磁场、流体、传热、化工、MEMS、声学物理场模拟。
21	Quicksilver	动态蒙特卡洛	解决一个简化的动态蒙特卡罗粒子传输问题来表示 Mercury 工作负载的一些元素。
22	clustal	多序列比对	进行蛋白分类，证明序列间的同源性，帮助预测新序列二级结构与三级结构，确定 PCR 引物，以及在分子进化分析
23	clustershell	多序列比对	进行蛋白分类，证明序列间的同源性，帮助预测新序列二级结构与三级结构，确定 PCR 引物，以及在分子进化分析
24	bookleaf	非结构化拉格朗日分析	非规则区域多相流动模拟的物理模型和数值方法
25	ceph	分布式对象存储和文件系统	存储数据、维护着展示集群状态的各种图表、执行诸如 ls、find 等基本命令
26	VPIC	分子动力学	分子动力学模拟
27	bqcd	分子动力学	动力学威尔逊费米子模拟晶格 QCD
28	ccs-qcd	分子动力学	求解位夸克传播子
29	comd	分子动力学	分子动力模拟
30	gromacs	分子动力学	分子模拟软件
31	chimera	分子结构交互式分析与可视化工具	蛋白分子可视化、包括密度图，分子组装，序列比对，对接结果，轨迹和构象整合



32	VMD	分子结构可视化	变分模态分解
33	beast2	分子序列贝叶斯分析	使用 MCMC 对分子序列进行贝叶斯分析
34	bioconductor	高通量基因数据分析	分析和理解高通量基因组数据
35	blis	高性能 BLAS 库应用框架	矩阵操作的接口
36	oasis	高性能纳维 - 斯托克斯解法器	高性能纳维 - 斯托克斯解法器
37	HYCOM	海洋模拟	开发和评估数据同化混合等密度 - 西格玛 - 压力 (广义) 坐标海洋模型
38	HACCKernels	核模拟	核设计与辐射安全评价软件
39	gamera	核物理	应用于东京的基于物理的城市地震响应分析
40	adventure	机械设计与制造	用于大规模分析和设计的计算力学系统 (非结构化网格)
41	bowtie	基因比对	适用于将长度大约为 50 到 100 或 1000 字符的 reads 与相对较长的基因组 (如哺乳动物) 进行比对
42	bowtie2	基因比对	适用于将长度大约为 50 到 100 或 1000 字符的 reads 与相对较长的基因组 (如哺乳动物) 进行比对
43	bwa	基因比对	用于将低差异序列与大型参考基因组 (例如人类基因组) 进行映射
44	RAJA-performance-suite	基于循环的计算内核性能测试	旨在探索 HPC 应用程序中基于循环的计算内核的性能
45	opencv	计算机图形学	1、人机互动 2、物体识别 3、图像分割 4、人脸识别 5、动作识别 6、运动跟踪 7、机器人 8、运动分析 9、机器视觉 10、结构分析 11、汽车安全驾驶
46	openfoam	计算流体力学	每一种流体流动都可以用一系列的偏微分方程表示, 求解这种运动的偏微分方程的代码, 即为 OpenFOAM 的一个求解器

47	code-aster	结构与热力学分析	结构与热动力学研究分析
48	TyphonIO	科学计算问题中数据高速 IO 库	科学计算问题中数据高速 IO 库
49	SDF	科学数据文件存储格式	文件系统
50	paraview	可视化	二维和三维数据进行分析和可视化的程序
51	fftw	离散傅里叶变换	用于计算一个或多个维度、任意输入大小、实数和复数数据（以及偶数 / 奇数数据，即离散余弦 / 正弦变换或 DCT/DST）
52	cp2k	量子化学与固态物理科学	可研究上千个原子的大体系，广泛用于固体、液体、分子、周期、材料、晶体和生物系统的模拟
53	eulag	流体力学	用于全尺度地球物理流的数值求解器
54	isaac	密码科学	密码学研究
55	XIOS	气象模拟中 IO 管理	气象模拟
56	caffe	人工智能	人工智能深度学习框架
57	caffe2	人工智能	人工智能深度学习框架
58	mxnet	人工智能	人工智能深度学习框架
59	pytorch	人工智能	人工智能深度学习框架
60	tensorflow	人工智能	人工智能深度学习框架
61	tensorflow-lite	人工智能	人工智能深度学习框架
62	asearch-blas	三角稠密矩阵解法器	三角稠密矩阵
63	NEST	神经网络模型	脉冲神经仿真平台

64	BLAST	生物信息	核苷酸或蛋白质序列与序列数据库进行比较，并计算匹配的统计显著性
65	crass-crisp r	生物信息	生物信息模拟
66	mauve	生物信息	基因组共线性图
67	plasma	生物信息	等离子研究
68	Laghos	时间依赖欧拉解 法器	它使用非结构化高阶有限元空间离散化和显式高阶时间步长求解移动拉格朗日坐标系中可压缩气体动力学的瞬态欧拉方程
69	numpy	数学库	一个强大的 N 维数组对象 Array; 2、比较成熟的 (广播) 函数库; 3、用于整合 C/C++ 和 Fortran 代码的工具包; 4、实用的线性
70	wrf-modeler	天气预报	天气研究与预报建模
71	Graph500	图计算	云计算
72	BeeGFS	文件系统	云计算
73	alya	物理科学	物理科学研究
74	cactus	物理科学	物理科学研究
75	NHC	系统监控	系统监控软件
76	NekTar++	有限元分析	支持使用谱 /hp 元素方法开发高性能的可扩展偏微分方程求解器
77	libmesh	有限元分析	支持对多种流行的几何和有限元类型进行 1D、2D 和 3D 稳态和瞬态模拟
78	atlas	自调优线性代数 库	数学库

## 4.7 国家超级计算广州中心

国家超级计算广州中心部署了一定规模的基于 ARM 指令集兼容架构处理器的集群，并围绕高性能计算、大数据处理展开了一系列的应用移植和优化工作。在应用移植方面，选取超算中心常用的 30 余款软件移植到 ARM 指令集兼容架构处理器上，并与峰值性能相当的 X86 处理器进行对比，分析 ARM 指令集兼容架构处理器在计算性能、访存性能等方面的优劣。在应用研发与优化方面，开展的典型研究工作如下。

### (1) 面向基于 ARM 指令集兼容架构处理器的超算系统的 Spark 优化

Spark 是当前应用最广泛的大数据批处理框架之一，各行业的大量业务依托 Spark 展开，围绕 Spark 已经形成完善的应用生态。随着近几年美国对我国在高端芯片上的技术封锁，以鲲鹏为代表的国产 ARM 指令集兼容架构处理器逐步成熟，并在各行业得到大量的应用。鉴于 Spark 在大数据处理领域的重要地位，围绕 ARM 指令集兼容架构处理器移植优化 Spark 框架具有重要意义。然而，由于 Intel 的 X86 架构处理器仍然是当前应用最广泛的处理器，目前的 Spark 框架主要面向 X86 处理器优化设计。不同于 Intel 的 X86 架构处理器，ARM 指令集兼容架构处理器为了在单个芯片内集成更多的核心，一般采用 NUMA 结构。例如，64 核心的国产 ARM 指令集兼容架构处理器分成 8 个 NUMA 组，每个组含有 8 个处理器核。系统管理员将内存条均匀地配置到各个 NUMA 组上。同一 NUMA 组内部的处理器核访问属于本组的内存时延迟一致，且相对较低；一个处理器核访问其它 NUMA 组的内存时延迟相对较高。由于多个 NUMA 组采用共享内存的方式，每个处理器核都能访问所有的内存空间，这样就不可避免地会导致远程内存访问的问题，即：一个处理器核访问其它 NUMA 组的内存空间。实际上，由于当前的 Spark 框架主要面向 Intel 的 X86 架构处理器，并没有考虑 ARM 指令集兼容架构处理器的 NUMA 结构，由此带来的远程内存访问显著降低了 Spark 在 ARM 指令集兼容架构处理器上的性能。

针对以上现状，国家超级计算广州中心面向 ARM 指令集兼容架构处理器的 NUMA 结构优化了当前的 Spark 框架，显著减少远程内存访问从而提升大数据处理性能。具体地，在当前 Spark 框架面向集群的资源管理器的基础上，在计算节点内部增加了一级面向 NUMA 组的资源管理组件。该组件根据数据在内存中的位置，将计算任务 (Task) 调度到数据所在的 NUMA 组中，从而避免远程内存访问。

任务调度到计算节点后，由操作系统向各个处理器核派发任务。然而，操作系统难以感知计算任务与数据之间的对应关系。在计算节点内部调度任务时不可避免地导致计算任务和其对应的数据位于不同的 NUMA 组内，最终造成远程数据访问。相比之下，资源管理器能够感知计算任务与数据之间的对应关系。为此，国家超级计算广州中心在资源管理器中引入一层 NUMA 结构感知的组件。该组件通过操作系统探知处理器核之间的逻辑关系，同时从分布式资源管理器获取计算任务和数据之间的对应关系，综合以上两种信息将计算任务和对应的数据调度到相同的 NUMA 组中，从而避免远程数据访问。实验表明，

我们提出的优化方法能够将 TeraSort、PageRank 等主流 Benchmark 在国产 ARM 指令集兼容架构处理器上的性能提升 30% 以上。

## (2) 面向 ARM 指令集兼容架构处理器的大数据并行算法库

对高性能计算应用产生的数值模拟结果实施大数据分析是科学发现的有效手段。为了保证数值模拟与后续的大数据分析在同一套计算系统中进行（从而减少数据移动的开销），国家超级计算广州中心研发了面向国产超算的大数据并行算法库，为高性能大数据处理应用提供有效支撑。针对机器学习、数据挖掘等领域的大数据处理算法展开并行优化研究，已经在新一代天河系统上研发了包含 17 个算法的大数据并行算法库，形成了比较完备的大数据分析处理生态环境。具体成果体现在以下几方面。

首先，确立了大数据并行算法库架构体系。传统的高性能计算并行算法库主要强调算法的并行，着力点在数据划分、计算并行、通信优化等方面。相比之下，大数据并行算法库面向的应用一般数据量较大，且主要是高维的数据，如何对这些数据实施统一的表示是需要考虑的问题。除此之外，大数据算法大多依赖距离的计算，而距离计算与数据维度、度量等因素紧密相关。为了赋予算法库使用者充分的灵活性，我们设计了大数据并行算法库结构体系，针对每个算法，主要从算法并行和 IO 并行两个方面对算法做大规模的并行优化，并通过动态库的形式提供给应用程序员调用。由于大数据算法处理的数据具有不同的形态，算法库依赖于用户自定义表示数据样本的数据结构，如数据维度、每个维度的数据类型等，并定义样本间距离的计算方法。如有必要，应用程序员还必须指定一些个性化的信息。本课题研发的算法库为以上自定义的数据或函数提供必要的接口，如：样本数据结构定义、样本距离计算方法、一些个性化描述等。

基于以上的架构体系，国家超级计算广州中心重点突破了大数据并行算法库优化技术。具体地，根据并行优化时面临的挑战将大数据算法分为三类，并分别展开有针对性的优化。

第一类是“密度聚类”类算法。这类算法以 DBSCAN、KNN、图算法为代表，它们的计算开销在很大程度上取决于样本的分布情况和数据在计算节点上的划分情况，且计算样本点的距离时存在不规则的访存、通信操作，在大规模并行是一般并行效率较低。针对这类算法，采用的优化措施包括：通过数据预处理实施优化的数据划分，减少跨节点通信；通过消息聚合，将小消息变为长消息；合理安排各节点间通信的时机，减少网络拥塞。

第二类是“Scan”类型算法。这类算法以 K-means、Naive Bayes 为代表，它们的计算逻辑比较简单，没有复杂的通信。但是由于计算较为简单，往往出现 CPU 闲置等待内存或存储系统中数据的情况。为此，采用的优化措施包括：设置专门线程负责从存储系统预取数据；设计新的数据结构，增加内存的读写粒度；优化多核之间的访存时序，减少多核对内存带宽的竞争。



第三类是“任务异构”类算法。这类算法以决策树、随机森林为代表，与大部分高性能计算应用的核心算法不同，这类算法由大量的异构任务组成，即：整个算法由大量的子任务组成，这些子任务的计算量不尽相同。资源管理器针对这类算法做任务调度时不可避免的会存在负载不均衡的情况，从而影响计算效率。针对这类问题，采用的优化措施包括：将各个任务进一步划分为更细粒度的子任务；同一计算节点上同时承担不同的子任务；计算开销大的任务使用向量指令，开销小的使用普通 CPU 指令。

### (3) 面向 ARM 指令集兼容架构处理器的典型应用研发

关于典型大数据应用，国家超级计算广州中心研发了面向 E 级超算系统的高精度基因组组装软件，能够支撑宏基因组、单细胞基因组等数据量大、重复度低的数据集的高效组装；研发了蛋白质结构相似性分析软件，与美国 ECP 计划的水平相当；研发了面向 ARM 指令集兼容架构处理器的大规模可扩展分子对接软件，支持虚拟药物筛选、新材料设计等应用；研发了基于 FAST（平方公里望远镜）数据处理的中子星搜索软件，实现 TB 量级的数据处理；研发了高精度大气正演软件，实现高性能计算与大数据处理的融合。主要工作介绍如下：

#### A) 基因分析

基因分析是生物信息学领域科学研究和产业应用的重要基础。对于一个全新的物种，获取基因分析的参考组是后续深入研究的先决条件。大量基因组测序数据需要将基因片段组装出参考基因组，尤其是随着宏基因组测序、三代测序兴起，测序数据的增长速度高于摩尔定律，基因组组装算法在性能、规模、精度和适应性等方面面临巨大的挑战。针对当前基因组组装算法在内存开销和计算精度等方面的不足，国家超级计算广州中心结合新一代天河的体系结构特征研发了面向海量宏基因组测序数据的大规模并行基因组组装软件 Exassembler，利用超算系统实现大规模基因组组装，依托超算系统的超大内存容量和超级并行处理能力提升可解决问题的规模和基因组组装的准确度。设计分块读取 Fasta 和 Fastq 格式文件的函数接口实现并行读取 Fasta/Fastq 文件，支持大规模宏基因组组装的海量原始数据读写，显著缩短算法启动时间。设计有序数组提升通信效率，提出基于节点内局部归并排序和 MPI 进程间并行奇偶交换排序相结合的方法，实现海量 K-mers 的高效排序，突破内存限制、具有良好的可扩展性。

#### B) 生物网络

生物网络需表达出相关实体的结构或功能特性，如分子、蛋白质或基因。典型的例子是基因表达网络或蛋白质 - 蛋白质相互作用网络，它们包含有关功能亲和力或结构相似性的信息。生物数据的不断增长，生物网络的不断扩大，给聚类算法带来了巨大的挑战，导致大规模生物网络的聚类和分析成为了一个难题。虽然马尔可夫聚类（MCL）在聚类序列相似性或表达式网络上已经取得了一定成功。但由于运

行时间较长和内存需求较高，MCL 对大规模数据集分析的可扩展性仍然是一个瓶颈，且现有的 MCL 算法无法处理超大规模数据，比如在数十亿个蛋白质序列中有数万亿个非零相似性的数据集。所以才有了 HipMCL，即高性能马尔可夫聚类算法，是采用并行的方式实现了 MCL 算法，充分利用了超算的分布式大内存和计算能力。

在 MCL 算法执行过程中，迭代地执行两个操作，分别是 expansion 和 inflation。Expansion 操作是对矩阵进行  $e$  次幂，inflation 操作是计算矩阵中每个元素的  $n$  次幂。将 MCL 或者 HipMCL 算法应用到生物网络中，都需要将生物网络转化成稀疏的邻接矩阵。在 HipMCL 算法中，还需对这邻接矩阵依据进程数  $p$  进行切分成  $p$  个子矩阵，逻辑上表示为二维的进程网格，每个进程在内存中存放一个子矩阵。由于 expansion 操作是计算密集型和内存密集型，HipMCL 对 expansion 操作采用了 Distributed-memory parallel SpGEMM 算法，SpGEMM 根据进程数  $p$  进行步迭代，每一步都执行数据分发（按照进程网格的行和列广播数据）和矩阵乘法计算，最终完成 expansion 操作。Inflation 操作则是各个进程独立完成对应子矩阵中元素的  $n$  次幂，无需和其他进程通信。另外，根据数据特性和稀疏性，以及内存使用情况，HipMCL 也可能执行 pruning, selection 或 recovery 操作。

在新一代天河系统中，国家超级计算广州中心对 HipMCL 中 inflation 和 expansion 操作进行进一步优化，将部分计算移植到加速器上，采用了向量指令和标量指令加快计算。在处理 160 万蛋白质数据时，能够在 10 分钟内完成分析。

### C) 分子对接

分子对接是研究蛋白质 - 配体分子间相互作用与识别的有效方法，基本原理是把配体分子放在受体活性位点的位置，按照几何互补和能量互补的原则来评价配体与受体相互作用的优劣，以找到两个分子之间最佳的结合模式，近年来主要应用于基于受体和配体结构的药物设计和筛选。其基本流程依次是解析分子结构进行同源建模，利用格点计算方法确定对接活性点位置，利用蒙特卡洛方法进行构象搜索，采用打分函数评价构象，调整构象优化残基等的柔性结构，输出最优构象得到稳定的复合物结构。

为充分利用计算资源，本程序采用层次化并行技术。在任务层采用 MPI 弱扩展方式，为不同进程独立分配对接任务，各进程独立计算相应的最佳构象，进程之间互不干扰。在构象搜索部分，使用工具分析得知热点程序为插值计算部分。针对此部分代码，不仅从算法上进行优化，采用 TDMA 算法利用三对角矩阵的特性简化计算过程降低时间和空间复杂度，在代码上也采用了 SIMD 模型进行优化，将原程序中的数据结构解封、多重循环展开，实现数据切分与分发。此外，考虑到处理器访问缓存与内存的延迟差异以及缓存空间限制，优化了缓存管理，在使用加速器计算时，将重复使用的小数据常驻缓存，其他只使用一次的数据分段加载到缓存上。

## D) 中子星搜索

脉冲星是一种具有高密度、强引力场、强磁场并快速旋转的中子星，其具有稳定的自转周期。对脉冲星的观测研究，极大地推动了核物理、天体物理、粒子物理、等离子体物理、广义相对论、引力波和导航等众多领域的发展。目前 FAST 天文望远镜对脉冲星漂移扫描巡天采集数据量已达数 PB，每天产生巨量的数据文件，巡天数据已进入 PB 时代。Presto 是一款常用的脉冲星数据分析处理软件，用于探索和搜寻脉冲星。

Presto 搜索脉冲星流程包含以下几个步骤：读取数据文件、消除干扰、消除色散、傅利叶变换、周期搜索、数据折叠。整个流程的主要核心在于消除色散、傅利叶变换和周期搜索。消除色散操作需要根据色散值的范围和精度进行尝试，通常需要进行成百上千个不同色散值的尝试，每个色散值产生对应的结果文件，继而对每个文件进行傅利叶变换将时域转化成频域，再在频域内进行周期搜索即得到候选脉冲星，最后进行候选过滤和数据折叠。

Presto 本身是一个单节点工具，并不提供并行方法，对 Presto 进行并行化主要包含以下几方面：

- ①任务并行，不同天文数据文件相互独立，各个数据文件作为独立任务并行执行；
- ②消除色散操作多 CPU 进程并行，消色散流程多个色散值对应计算和结果相互独立，将色散值范围分割多进程并行执行；
- ③傅利叶变换和周期搜索多 CPU 进程并行，针对消色散产生的大量结果文件，这些文件相互独立，多进程并行同时对多个文件执行操作；
- ④傅利叶变换操作在 CPU 进程内实现多加速器核并行。

国家超级计算广州中心将并行化的 Presto 从 200 节点到 12800 节点进行测试，数据文件数从 100 到 6400 进行递增，在小于 6400 个节点的测试中，随着测试节点和数据量不断增加，执行时间基本保持相同，说明并行化取得很好的线性扩展性能。

火山喷发污染物的模型模拟和预测在航空安全、全球气候变化研究、大气动力学等许多领域具有重要意义，但受制于当前观测技术（如卫星测量）的自身局限性，无法提供详细和完整的时空信息。在给定适当的初始化条件下，数值模拟可以提供完整且高分辨率的信息，而模型预测可以为空中交通管制提供预警信息。得益于超级计算机以及大规模并行计算技术的出现，大气传输实时预测的分辨率和准确性越来越高。MPTRAC 是一个大规模并行轨迹计算软件，用于模拟和预测火山喷发后污染物（SO<sub>2</sub>）的传输扩散轨迹。

MPTRAC 正演过程包含以下步骤: 读取气象数据文件、大气平流、大气扩散、大气沉降、质量指数衰减、高精度结果记录。其中核心过程为大气平流、大气扩散和大气沉降。大气平流利用气象数据进行插值模拟, 根据气流速度等信息计算当前时刻 SO<sub>2</sub> 轨迹的三维地理位置 (经纬度、气压) 估计值; 接着利用高斯扩散模型模拟 SO<sub>2</sub> 在水平和垂直方向上的扩散过程; 大气沉降模拟 SO<sub>2</sub> 沉降直至到地面, 主要涉及气压 (海拔高度) 改变; 利用拉格朗日粒子扩散模型计算 SO<sub>2</sub> 分子的质量衰减变化。最后根据设定的时间、空间精度记录扩散轨迹结果。

国家超级计算广州中心采取数据并行的方式进行并行化, 每个 MPI 进程读取自己的输入数据, 并将数据进一步划分给所有核, 每个核计算各自的 SO<sub>2</sub> 扩散轨迹。

## 4.8 国家超级计算济南中心

济南超算致力于国产超算生态建设, 开展围绕国产基础操作系统、并行调度管理、应用系统移植优化等方面的科研及服务工作, 其中“山河”超级计算平台于 2021 年 12 月上线运行, 围绕山河超级计算平台建设有国内基于鲲鹏 920 的超算系统用于支撑高性能计算和云计算服务。当前, 济南超算已部署 1000 台鲲鹏计算节点, 采用 100GE (支持 RoCE) 高速互联, 节点采用双路华为鲲鹏 920 处理器 (32 核), 每个计算节点拥有 64 核处理器和 512GB 内存, 总计 6.4 万核, 系统的理论双精度峰值性能达 665TFLOPS, 其中 900 计算节点用于云计算服务, 100 个计算节点部署 HPC 典型应用提供算力服务。

典型应用如下表所示:

表 4-5 “山河”超级计算机部署的应用软件

应用领域	软件名	软件简介
材料科学	GROMACS	分子动力学
	VASP	量化计算
	Gaussian	量子化学
	GAMESS	量子化学
	LAMMPS	分子动力学
	NAMD	分子动力学
	Amber	分子动力学
	Quantum ESPRESSO	第一性原理
	Conquest	第一性原理
	CP2k	第一性原理
生命科学	Autodock	分子建模
大气科学	WRF	中尺度天气预报模式
流体力学	OpenFOAM	流体力学
	Palabos	流体力学

## 4.9 武汉超算中心

武汉超算中心是省市重大科技项目，匹配武汉“965”产业布局，聚焦打造智能制造、生命科学、材料科学和航空航天等重要领域的产业应用服务平台，打通“产学研用”，以普惠公共算力服务平台赋能产业科研创新、产业汇聚发展、产业人才培养，将成为湖北数字经济发展、产业升级转型的新引擎。

武汉超算中心包含鲲鹏 CPU 及 GPU 两种算力，总规模 50P。当前已吸引 48+ 单位使用，包括科学计算大应用、气象、生物信息、物理、化学等用户，所使用的应用包括：WRF,Grapes,Relion,GROMACS,LAMMPS,SAW, DeepMP, CP2,openForm,PFlows,PWMat,WESP,RHMC,Simdroid,Cistem,SMMD,GSIMS,SMMD,GIMS,PARTICLE,QCD,LICOM,ScETD,MOM,PhotoNs,MASNUM 等。

其中：PFlows，RHMC，WESP 是武汉本地的典型大应用：

### 1) PFlows

Pflows 计算流体力学仿真软件，基于格子玻尔兹曼模型（LBM）自主研发，计算效率业界领先。该软件支持自主一步索引 LBM 高效算法、流固耦合 IBM 高精度算法；正交网格算法，高效处理复杂运动结构；具备线性并行加速性能，可支持万核 / 千卡级异构高效并行；支持多物理场，包括：颗粒两相流、非牛顿流体、非等温传热、化学反应流。计算网格规模可达 1000 亿，大规模并行计算效率大于 85%。

近期基于鲲鹏 +GPU 平台进行了验证测试，最高测试网格数达到 680 亿，512 卡并行加速效率达到 98%

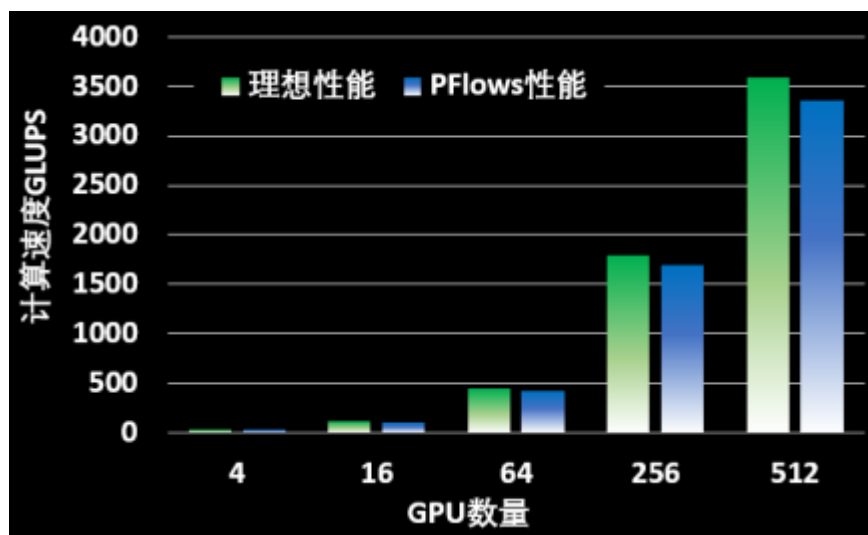


图 4-14 鲲鹏 +GPU 并行效率



## 2) WESP

WESP 软件由中国科学院精密测量科学与技术创新研究院自主研发的量子化学软件。该软件面向化学、材料、药物等领域的并发式高通量计算，国内同行能够方便利用该软件发展新理论方法。该软件目前已完成向鲲鹏平台的移植，可以基于武超平台运行，同时该软件可支持鲲鹏 +GPU 的异构并行计算并支持大规模扩展，当前可支持 P 级计算规模。

该软件针对某主流商业 CPU 和鲲鹏进行了对比测试，经过优化，部分场景，如：CPU 自治场过程平均时间优于该主流商业 CPU 表现：

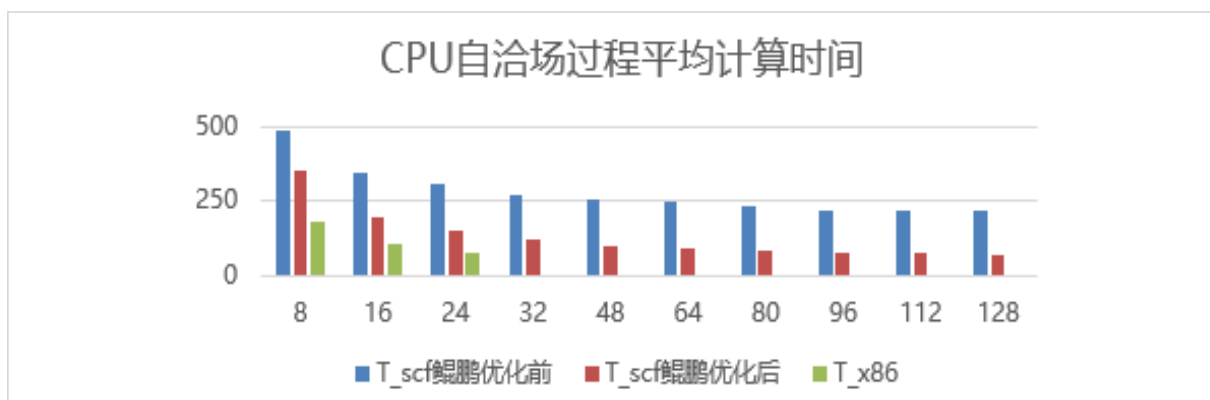


图 4-15 鲲鹏与某某主流商业 CPU 的性能对比

## 3) RHMC

RHMC 是华中师范大学自研的超算应用软件，该软件用于格点量子色动力学研究，包括：研究宇宙、物质界强相互作用力等。该软件已迁移支持鲲鹏平台，且支持鲲鹏 +GPU 异构并行及横向扩展，可支持 P 级并行扩展。

## 4.10 广东气象局

全球气候变暖，极端天气、气候事件频发。面对广东省多台风、多暴雨等气候特征，广东省气象局瞄准气象科技前沿，自主研发数值预报模式，联合华为公司建设完成首个全国产化的气象高性能计算系统，实现了历史性的突破。

广东气象局传统高性能计算系统，面对中外日益加剧的“卡脖子”科技威胁，以及气象行业逐步向碳源汇模式、环境气象模式、1km 及 500m 区域的预报的数值预报模式发展，计算性能需持续提升，计算平台亟需由 x86 系统向国产化平台移植与优化，兼容 GRAPES 数学库、工具库、第三方 I/O 库等气象软件和工具。

广东气象局引进“鲲鹏”高性能计算机系统，基于鲲鹏芯片构建算力底座，包括近 1 万个运算核心，72 个节点，峰值计算速度为 100 TFLOPS（万亿次）。全栈平台包括国产鲲鹏 CPU、RoCE 高速网络、存储硬件设备，集群管理调度软件、基于鲲鹏的毕昇编译器和 HCC 编译器、鲲鹏数学库、MPI 高速计算网通信库等软件。广东气象局联合华为公司共同推出了全面国产化、极致有效算力的数值预报模式气象平台。

相比原有计算平台，广东气象局全“芯”国产化鲲鹏高性能气象系统，业务应用和鲲鹏 HPC 完成 100% 适配，结果表明“鲲鹏”与“天河二号”高性能计算机运行的模式结果精度接近，但单个节点计算能力提高了 1.7 倍，模式的运行速度更快。数值预报模式国产化研究工作的进步，将更好地服务于广东省经济社会发展和防灾减灾需求。

### 4.11 深圳华大生命科学研究院

作为华大集团的核心研发机构，深圳华大生命科学研究院多年以来面向国家重大战略需求，聚焦基础研究与应用转化，深耕基因组学领域，已经实现了低成本可扩展测序平台、新型合成系统等拥有自主知识产权的世界领先核心技术突破，迅速发展为世界领先的基因组学研究中心。时空组学技术由华大自主研发，基于 DNA 纳米球（DNA Nano Ball, DNB）技术开发，是具有高通量、超高分辨率、大视场的原位全景式技术，可以实现同一样本在组织、细胞、亚细胞、分子“四尺度”同时进行空间转录组分析。该技术通过时空芯片捕获组织中的 mRNA，并通过时空条形码（Coordinate ID, CID）还原回空间位置，实现组织原位测序，为深入地了解细胞的基因表达及形态与局部环境之间的关系建立强大的研究基础。基于时空组学技术的小鼠胚胎发育，蝾螈脑再生，拟南芥、果蝇、斑马鱼发育时空图谱先后以封面形式发表在 Cell、Science、Developmental Cell 期刊上。

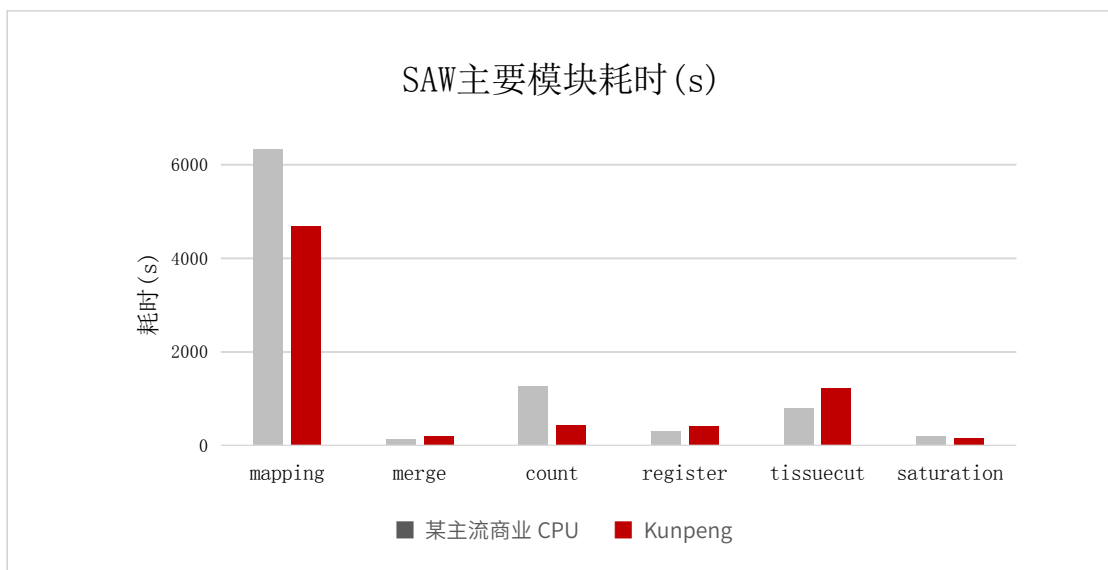


图 4-16 SAW 主要模块在不同计算平台上的耗时分析

表 4-6 SAW 主要模块在鲲鹏 920 和某主流商业 CPU 8xxx 上的性能分析

测试项	主流商用芯片 8xxx	鲲鹏 优化 V1	鲲鹏 优化 V2
CIDcount	22.16	30.07	29
mapping	6338.01	5023.57	4441.85
merge	142.87	204.93	203.8
count	1266.84	615.67	446.21
register	312.8	461.94	414.14
tissuecut	795.51	1234.27	582.2
spatialCluster	46.1	80.97	82.87
saturation	203.06	166.88	166.88
report	3.7	3	2.96
gef2gem	27.67	38.41	38.81
总耗时	9158.72	7859.71	6408.72
性能比例	1	+14.18%	+30.03%

时空组学分析完整流程实现了鲲鹏平台的适配及优化，通过采用：1) 基于毕昇编译器完成 C/C++ 部分的编译与优化；2) 结合 NEON/SVE 等内嵌指令技术，深度优化指令编译和运行时库；3) 使用鲲鹏架构深度优化的 Python 及数学库；4) 利用鲲鹏软加速库优化 Gzip、ZSTD 等算法的运行效率；5) 通过 NUMA 绑定减少线程数据存取开销等优化措施，流程中非常关键的 Barcode 的对比，以及 RNA 序列比对提升效率 40%-50%，整个流程节约时间近 30%，对时空组学批量数据分析起到了重要的推动作用。

## 4.12 飞腾信息技术有限公司

飞腾集群部署于某金融机构，具有高性能、高可用、可伸缩等特征。截止目前，某金融机构已经部署超过 3000 个飞腾服务器节点，到 2022 年底将会扩展到 5000 个。飞腾集群目前已经部署包括 Hadoop、金融业务应用、易捷行云云平台、易捷行云操作系统、红旗操作系统在内的多个软件平台，主要应用于金融，部署在会计业务和风控业务等领域。目前集群稳定运行，已经满足了银行去 IOE 的需要，并采用易捷行云 EasyStack 提供的开源云平台解决方案，构建了中国最大的基于 OpenSack 的金融生产云。

飞腾集群还部署在某互联网大厂数据中心，目前已经部署和稳定运行了包括阿里云 OS，阿里云云平台 mysql 数据库在内的过个软件，在地方政务云方面发挥重要作用。同时，飞腾集群还中标某集团项目，

节点数量超过 1400 个，目前已经部署和稳定运行多少 IaaS 和 SaaS 产品，包括数据库产品 Tbase 和 TDSql、大数据产品 TBDS 和安全治理产品 PSG、私有云管理引擎 DCOS、网关产品 NATGW、均衡负载 CLB；虚拟机、容器和裸金属产品等。该集群的主要应用领域为票据业务，并于 2021 年 12 月 31 日开出了第一张票据。

### 4.13 澎峰（北京）科技有限公司

澎峰（北京）科技有限公司是国内知名的基础计算软件栈企业，由 OpenBLAS 项目创建者张先轶博士建立。澎峰科技是一家致力于以软件能力开启算力时代的加速计算基础软件企业。公司产品为异构计算软件栈，包括 PerfMPL 高性能计算库，PerfXAPI 跨平台异构计算框架，PerfXPy 新一代科学计算平台等。公司客户包括算力芯片公司，以及算力需求行业（算法企业、工业软件企业、算力中心）等。此外，公司积极参与开源计算软件生态建设，领导国际知名开源矩阵计算库 OpenBLAS，以及开源指令集 RISC-V 的计算库项目 RVCL。目前澎峰科技在北京、长沙和青岛都建立了分公司和研发基地。针对 ARM 开放指令集兼容建构，澎峰科技的主要工作如下：

PerfMPL<sup>®</sup> (Math Performance Library) 是一套针对处理器各类指令集和各领域应用程序经过高度优化的数学计算库。PerfMPL<sup>®</sup> 自带强大的包管理能力，所以它既可以作为 x86、ARM、RISC-V 指令集的 CPU 平台和各类 DSA 加速硬件的数学函数库基础软件，也可以作为领域应用程序（如 AI、CAE、EDA、信号处理等）的依赖数学计算库。PerfMPL 是以 OpenBLAS 为发展基础，经过多年发展逐步覆盖至 FFT（快速傅立叶变换）、SPARSE（稀疏矩阵解法器）、MATH（基础数学库）、VML（Vector Math Library 向量数学库）、DNN（Deep Neural Networks 深度神经网络）等计算库等。

PerfXPy 是澎峰自主研发的支持异构计算的高性能 Python 集成开发环境，重点支持国产计算芯片的适配，并面向科学计算、AI 计算、大数据分析等领域提供各类组件。PerfXPy 兼顾易用性与高性能，针对国产 ARM 指令集兼容架构处理器进行加速计算优化，加强了对异构加速设备 GPU/NPU 的支持，并不断研发集成各类工具，赋能科研、教学和 AI 算法的开发。目前澎峰科技与飞腾已经完成了 PerfXPy 在 FT-2000/4、FT-2000+/64 处理器的产品兼容性相互认证测试。测试结果表明 PerfXPy 软件与飞腾系列处理器上实现了全面兼容，并且在功能和性能上实现对原生 Python 的超越。这标志着国产硬件和国产软件在技术上的深度融合，和实现性能超越。国产化计算机软硬件生态融合将携手助力关键信息基础设施国产化信息化全建设。

PerfXPY 相较于原始 Python 安装包的核心优势如下：

1) 性能提升。PerfXPY 集成了澎峰科技的高性能数学库 PerfMPL，使得飞腾系列处理器在数值计算、科学计算、AI 计算的效率提升显著。

2) 支持异构计算。PerfXPY 集成了澎峰科技的 PerfXAPI 异构计算软件栈，使得飞腾处理器可以很容易支持各类 NPU/DSA, GPGPU 计算加速卡，助力飞腾处理器实现与各国产加速卡的生态融合。

3) 功能丰富和使用方便。PerfXPY 是一个支持异构计算的 Python 集成开发环境，提供了 Python 相关的丰富科学包以及依赖项等，完美的代码编辑，调试和运行开发环境。PerfXPY 集成了数据分析、科学计算、AI 所需的软件包，实现开箱即用。PerfXPY 强大的包管理功能，支持用户下载丰富的 Python 工具包和自定义。

4. PerfXPY 支持各类 AI 算法，与国内算法公司广泛合作，服务于各类应用场景。

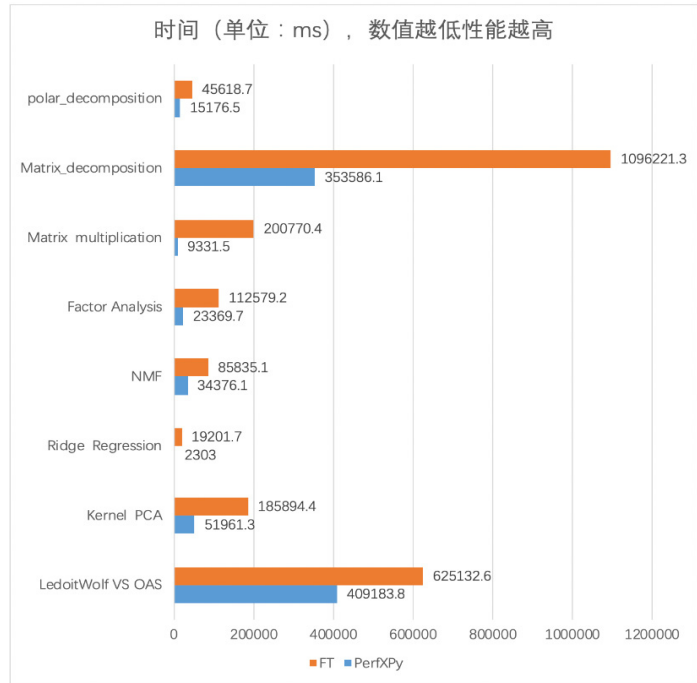


图 4-17 FT-2000/4 处理器平台上的性能比较

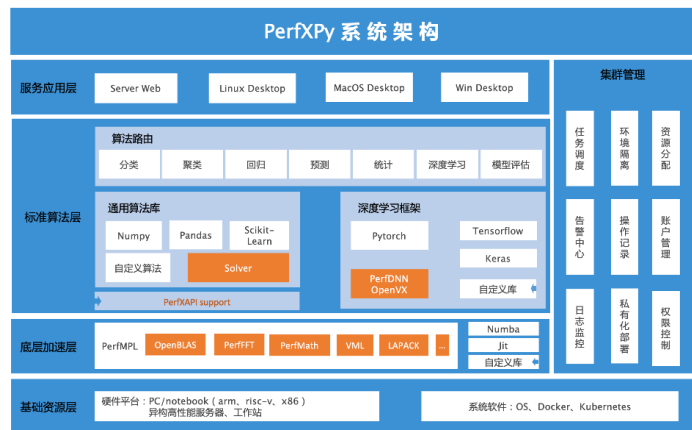


图 4-18 PerfXPY 系统架构图

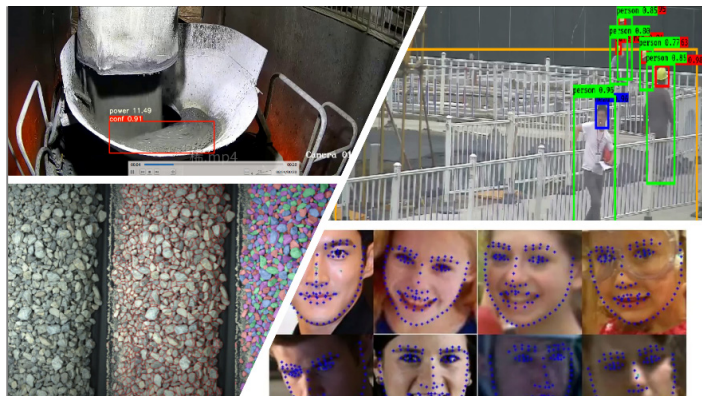


图 4-19 PerfXPY 应用示意图



## 4.14 中国移动通信集团浙江有限公司

### (1) 浙江移动打造全球首个基于鲲鹏的 ICT 全场景样板点

浙江移动 4G/5G 核心网、大数据、CRM、计费、CDN、DPI、OSS、政务云、分布式存储、数据库等 ICT 全场景使用，鲲鹏商用 4000+ 台，占新增加服务器的 35% 以上。



图 4-20 基于鲲鹏的 ICT 全场景样本点

### (2) 大数据场景采用鲲鹏服务器打破数据分析耗时瓶颈, 提升整体性能 21.6%。

浙江移动大数据平台汇聚内外数据, 对内外部开放数据分析能力。原有大数据集群整个业务数据分析过程接近 2 小时, 成为话费、商城等业务的瓶颈。我们的大数据平台针对基于 ARM 指令集兼容架构的鲲鹏 920 做适配和内核改造, 充分发挥了 ARM 指令集兼容架构在多核和硬件加速器方面的优势。

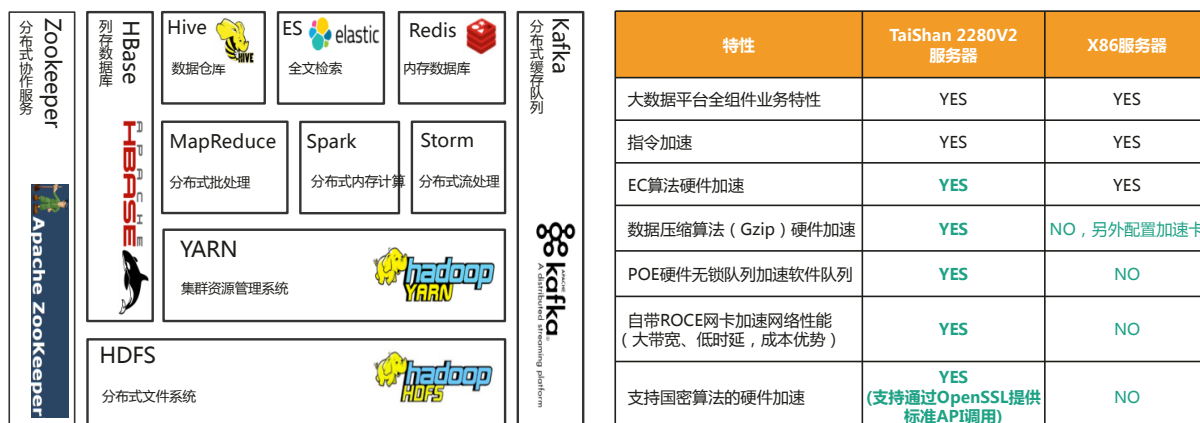


图 4-21 基于鲲鹏 920 的大数据平台

由于开源 Hadoop 的内核是围绕 X86 的生态设计的，无法充分发挥 ARM 指令集兼容架构的架构优势。我们切换到 ARM 指令集兼容架构生态，针对 ARM 指令集兼容架构做了适配和内核改造，从而发挥 ARM 指令集兼容架构多核和硬件加速器的优势。

1) 业务适配工作包括：

- (a) 适配鲲鹏 920 加速器驱动，使能压缩、加密、EC、RoCE 加速；
- (b) 使用 HUAWEI JDK/OpenJDK(ARM 指令集兼容架构)；
- (c) 使用 Euler OS；
- (d) 组件基于 ARM 指令集兼容架构参数优化，发挥 ARM 指令集兼容架构的多核优势；

2) 内核改造工作包括：

(a) 平台增加对国密算法 OS 库函数的调用流程，实现加密算法切换；Hive、HBase、HDFS 等组件增加对接 KMS 功能；配置加密套件，支持 SM2 传输加密；改造认证流程，在 Kerberos 中支持新的国密算法。

- (b) 增加华为自研 FlexEC 函数库及调用流程。
- (c) 开源包没有 ARM 指令集兼容架构版本依赖，需要重新适配 ARM 指令集兼容架构。

大数据是分布式并发计算框架，将大数据量切分成多个独立的数据集同步进行处理。鲲鹏 920 的多核架构能够大幅提升数据处理的并发能力，有效提升大数据处理效率。

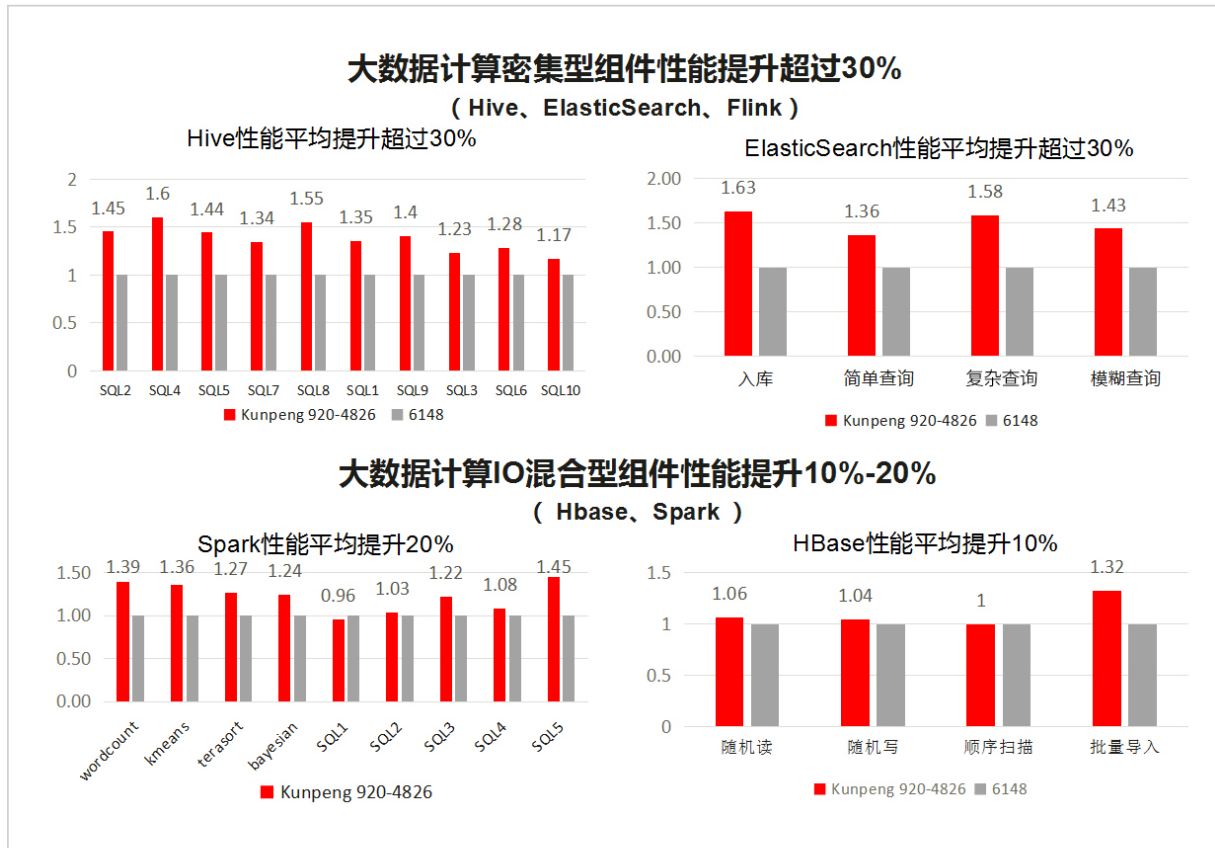


图 4-22 采用鲲鹏 920 后的性能提升情况

### 4.15 英特工程仿真技术（大连）有限公司

现有 windows-intel 环境开发的国产高性能计算软件 INTESIM-MultisimV5.0 已拥有解决工程实际应用问题的大部分功能，已有的建模与可视化软件是国内领先的一体化通用前后处理软件，功能完备。INTESIM-MultisimV5.0 具有完全自主可控代码，具备代码级跨平台能力。从图形显示、块设备 I/O、内存管理、多核结构突破多种国产 CPU/OS 计算机平台环境适配，完善国产 PC、服务器、超算集群高端研发设计应用生态，解决航天、航空、船舶、轨道交通、汽车、电子、电器等领域的高端研发设计与国外操作系统及芯片的紧耦合性。促进信息安全自主可控，具有较高的工程应用价值。INTESIM-MultisimV5.0 软件适配国产 CPU/OS 已具有一定基础，核心求解器已对 ARM 指令集兼容架构 HPC 体系进行适配。

英特仿真针对华为鲲鹏系列 aarch64 架构处理器将 INTESIM 进行了移植。华为对鲲鹏处理器配套研发了麒麟（针对计算机集群）和 OpenEuler（针对个人计算机）操作系统。华为研发团队发布了适配的 GCC 编译器和 HyperMPI，我们在此基础上编译了 MultiSim，以及矩阵库 OpenBLAS 和稀疏系统求解器等第三方组件。编译所得程序可在鲲鹏计算平台上以命令行及脚本的方式多节点启动，并通过了案例库测试。

MultiSim 本身采用基于网格分解的分布式多级混合并行计算，跨多种软硬件平台，适合大型计算机集群；支持单场、多场强耦合、多场弱耦合等的计算模型，以及支持多波前、直接 LU、MG 预调共轭梯度、Jacobi（雅可比）预调双共轭梯度等多种高效的大型稀疏系统求解方法。

华为推出了针对鲲鹏平台优化的数学库 KML，未来将逐步加入其矩阵基础运算功能和稀疏矩阵求解功能，从而使鲲鹏平台的 MultiSim 性能进一步提升。

使用英特仿真的耦合求解器对航空飞行员弹射救生过程所佩戴的头盔面罩组合体进行热流固耦合仿真分析，计算获得头盔面罩组合体的压力云图以及结构位移云图，为头盔面罩组合体的设计提供依据。如下图所示为头盔面罩组合体的分析模型。远场边界条件设置界面如右图所示，马赫数为 0.8，攻角为  $0^\circ$ ，侧滑角为  $0^\circ$ ，静温为 300K。

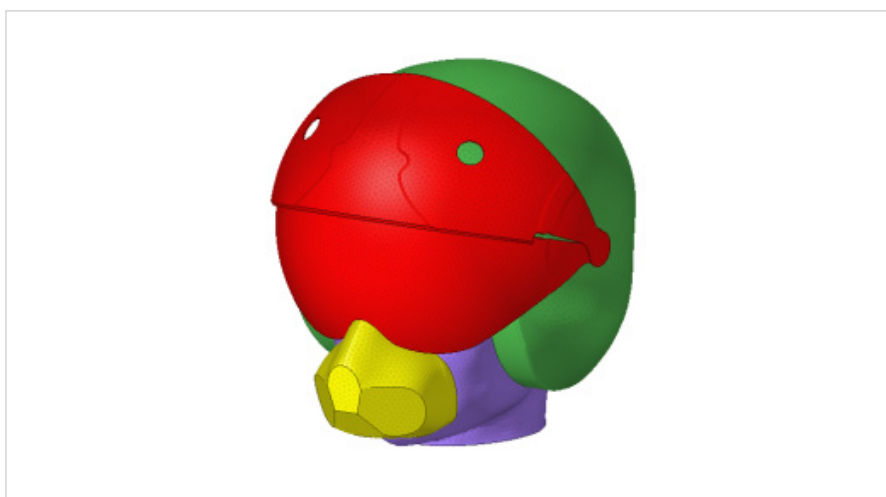


图 4-24 头盔面罩组合体分析模型

耦合分析计算获得的头盔面罩组合体表面压力云图以及与商软对比结果如下图所示。

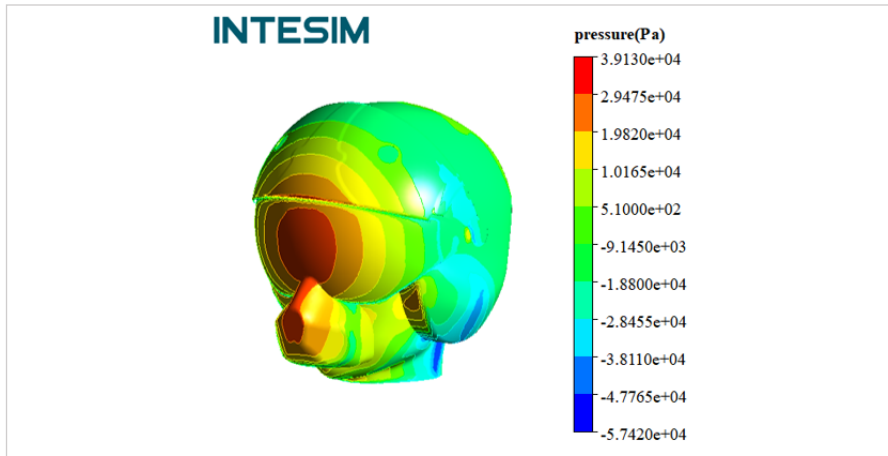


图 4-25 热流固耦合分析工具计算头盔面罩组合体表面压力云图

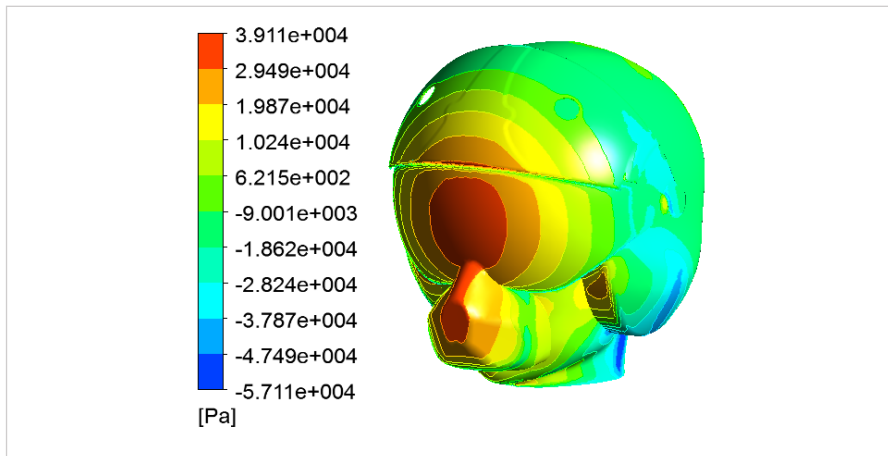


图 4-26 某对标软件计算头盔面罩组合体表面压力云图

热流固耦合分析计算头盔面罩组合体位移云图以及与商软对比结果如下图所示。

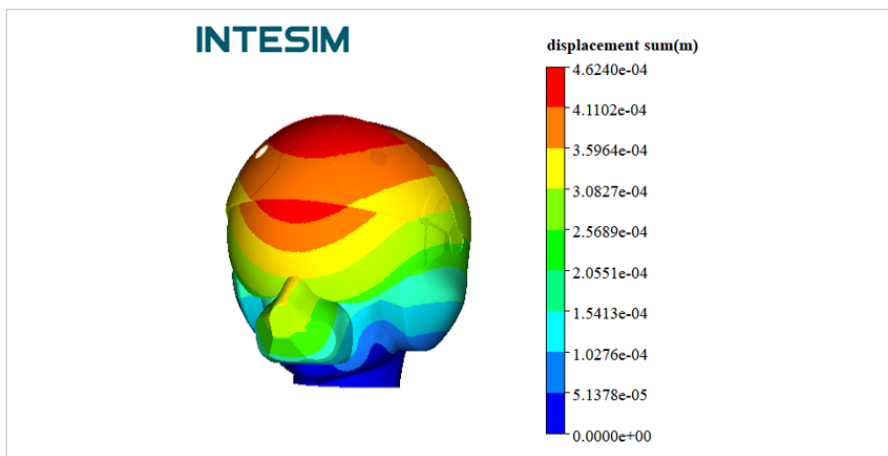


图 4-27 热流固耦合分析工具计算头盔面罩组合体结果位移云图



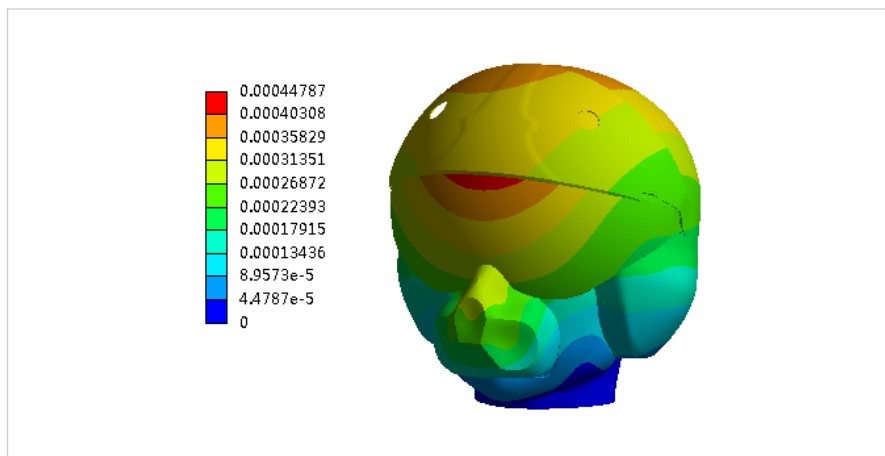


图 4-28 某对标软件计算头盔面罩组合体结果位移云图

结果表明，使用英特仿真的耦合分析工具对某型航空飞行员用头盔面罩组合体模型进行热流固耦合仿真分析，计算获得头盔面罩组合体表面压力云图以及头盔面罩组合体位移云图。并使用某对标软件采用相同的设置条件进行计算对标，误差在 5% 以内，证明热流固耦合分析工具能够满足航空工程应用仿真分析计算要求。

此外，还有轴流压气机单向热流固耦合案例、48 槽 8 极表贴永磁同步电机案例等典型案例，不再赘述。

## 4.16 深圳十沣科技有限公司

深圳十沣科技有限公司出品的计算流体力学软件 QFLUX 的求解器遵循 C++17 语言规范，采用 CMake 管理编译安装过程，支持 GCC 编译器和 CLANG 编译器。华为鲲鹏平台提供了完善的功能支持，并且提供了针对 ARM 指令集兼容架构优化的 gcc 10.3.1 编译器和毕升编译器。由于 QFLUX 依赖的第三方库都有完整的源代码，QFLUX 向鲲鹏平台移植过程中，将 QFLUX 依赖的第三方库在鲲鹏平台重新进行了编译安装，同时将 QFLUX 求解器代码在鲲鹏平台进行了重新编译安装，即完成了移植过程，移植过程非常方便。利用方腔测试算例在鲲鹏平台上和主流商业 CPU 平台对 QFLUX 进行对比测试，测试结果表明鲲鹏平台的算例残差数据前 500 步与在主流商业 CPU 上的数据完全吻合（即表格中的前 4 列）。对于方腔算例，单元数 1600 万，迭代 1000 步，采用 20 核计算，某主流商业 CPU 上耗时 10108s，鲲鹏平台上耗时 7265s，某主流商业 CPU 上耗时比鲲鹏平台上耗时多 39%。

表 4-7 QFLUX 在某主流商业 CPU 和鲲鹏 920 上的精度对比

	#iter	continuity	U.x	U.y	U.z	walltime	lefttime
某主流商业 CPU	491	8.4870e-06	5.4682e-04	1.2700e-03	4.0710e-03	4.2210e+03	4.3758e+03
	492	7.3313e-06	5.4622e-04	1.2679e-03	4.0669e-03	4.2322e+03	4.3699e+03
	493	7.9500e-06	5.4563e-04	1.2658e-03	4.0627e-03	4.2434e+03	4.3639e+03
	494	7.8968e-06	5.4506e-04	1.2639e-03	4.0585e-03	4.2546e+03	4.3580e+03
	495	1.2674e-05	5.4450e-04	1.2619e-03	4.0544e-03	4.2651e+03	4.3513e+03
	496	7.9273e-06	5.4393e-04	1.2599e-03	4.0501e-03	4.2764e+03	4.3453e+03
	497	1.2632e-05	5.4336e-04	1.2580e-03	4.0459e-03	4.2867e+03	4.3385e+03
	498	7.3188e-06	5.4276e-04	1.2559e-03	4.0417e-03	4.2979e+03	4.3324e+03
	499	7.4532e-06	5.4219e-04	1.2538e-03	4.0376e-03	4.3091e+03	4.3264e+03
	500	8.6326e-06	5.4163e-04	1.2518e-03	4.0334e-03	4.3199e+03	4.3199e+03
	#iter	continuity	U.x	U.y	U.z	walltime	lefttime
鲲鹏处理器	491	8.4870e-06	5.4682e-04	1.2700e-03	4.0710e-03	3.3506e+03	3.4734e+03
	492	7.3313e-06	5.4622e-04	1.2679e-03	4.0669e-03	3.3579e+03	3.4671e+03
	493	7.9500e-06	5.4563e-04	1.2658e-03	4.0627e-03	3.3651e+03	3.4607e+03
	494	7.8968e-06	5.4506e-04	1.2639e-03	4.0585e-03	3.3723e+03	3.4542e+03
	495	1.2674e-05	5.4450e-04	1.2619e-03	4.0544e-03	3.3788e+03	3.4471e+03
	496	7.9273e-06	5.4393e-04	1.2599e-03	4.0501e-03	3.3859e+03	3.4405e+03
	497	1.2632e-05	5.4336e-04	1.2580e-03	4.0459e-03	3.3925e+03	3.4335e+03
	498	7.3188e-06	5.4276e-04	1.2559e-03	4.0417e-03	3.3999e+03	3.4272e+03
	499	7.4532e-06	5.4219e-04	1.2538e-03	4.0376e-03	3.4072e+03	3.4208e+03
	500	8.6326e-06	5.4163e-04	1.2518e-03	4.0334e-03	3.4140e+03	3.4140e+03

此外，分别采用了系统自带的 gcc 7.5.0, 华为提供的 gcc 10.3.1 以及华为提供的毕升编译器对 QFLUX 的求解器源代码进行编译，采用方腔算例测试了鲲鹏平台上不同编译器的运行效率。测试表明 QFLUX 和鲲鹏平台能够很好的适配，且在 8 核前保持线性加速，此后由于访存成为瓶颈导致并行效率下降，这个行为与我们在主流商业 CPU 平台上测试的行为一致，但是鲲鹏平台上 80 核依然能够保持 50% 以上的效率，优于我们在主流商业 CPU 平台上测试的结果。

#### 4.17 湖南迈曦软件有限责任公司

湖南迈曦软件有限责任公司自主开发了工程 CAE 仿真分析软件 MxSim，包含有完整的通用结构分析、显式动力学分析、流体仿真分析以及多物理场分析模块，以及电磁、声场、热场等求解器，底层求解器 MxDSV 完全自主开发，包含直接法方程组求解器、迭代法方程组求解器以及特征方程组求解器。MxSim 深度适配国产操作系统以及 ARM 指令集兼容架构与 x86/x64 多架构国产处理器，采用 cmake 进行跨平台编译，支持多核、多节点以及 CPU/GPU 异构等多类型并行计算环境，致力于营造工业 CAE 软件国产生态。目前，MxSim 已经适配于华为鲲鹏云计算环境，以及飞腾系列 CPU、麒麟等国产软硬件平台。MxSim 的异构并行计算支持 OpenCL、HIP、CUDA 等多种架构，MxSim 显式动力学分析模块已经适配于海光 DCU 计算平台，计算效率相对于 CPU 并行计算提升明显。同时，已经完成了在景嘉微 JM92 系列图形处理器上的单精度计算能力适配。MxSim 所包含的底层求解器采用独创的异构自适应算法，实现对 CPU 和 GPU 计算资源的充分调度，峰值性能可以达到硬件计算能力的 95% 以上；自研基于数学规划的异构多设备负载均衡策略，可在任意设备组合中实现高效精准的任务分配，在复杂计算平台中计算性能最高提升 150%。

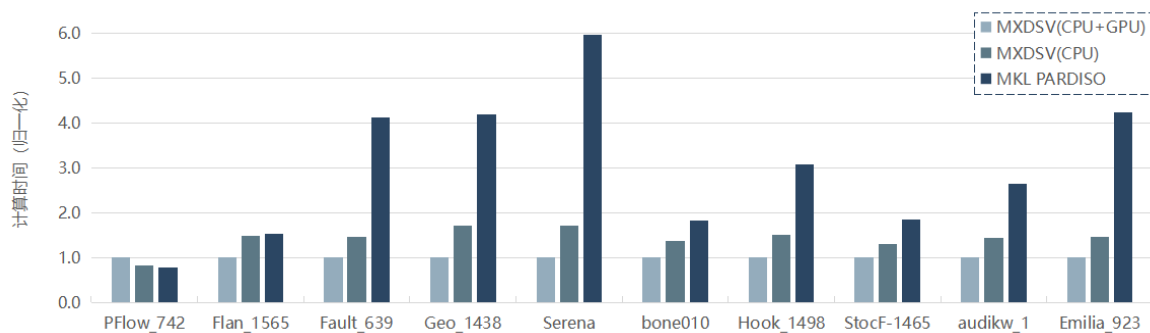


图 4-29 MxDSV 与国外主流求解器的计算能力对比

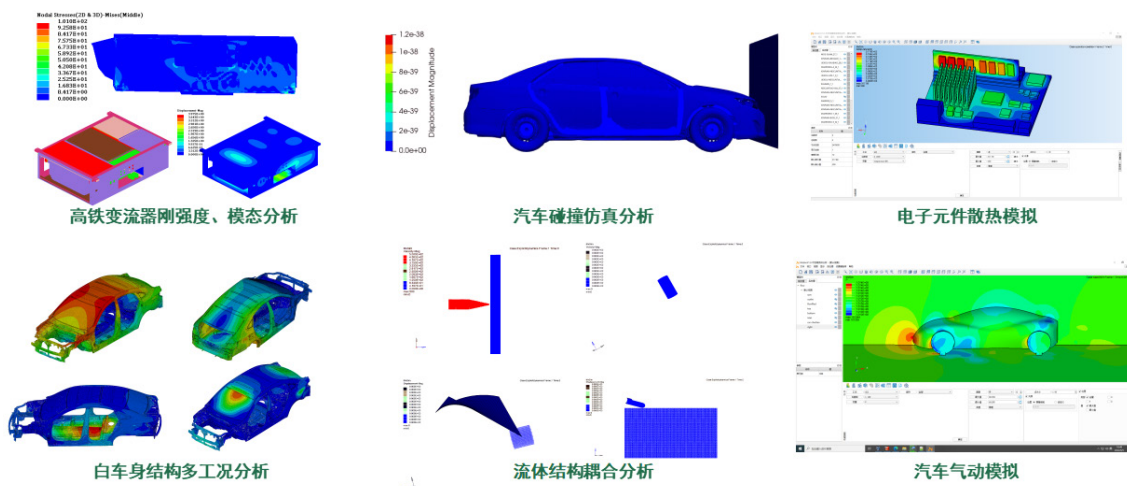


图 4-30 MxSim 典型应用场景

### 4.18 北京龙讯旷腾科技有限公司

龙讯旷腾成立于 2015 年的北京龙讯旷腾科技有限公司是国内材料计算模拟工具软件研发创新的领导者，其具有完全自主知识产权的国产第一性原理材料计算软件 PWmat 在求解算法、异构加速及大规模并行优化等方面具有国际领先的技术优势，在保证计算精度的同时大幅度提升计算规模和计算效率，可以将基于量子力学的原子级别计算应用到新材料设计、加工、优化等一系列研发生产环节，为微电子、化工、新能源、生物医药等行业带来革命性的创新驱动动力。

龙讯旷腾团队与华为 HPC 优化团队深度合作，基于鲲鹏多核优势，对 PWmat 代码进行了完整的平台移植，并针对编译、数学库、MPI 库等进行了初步优化。现有鲲鹏 920 程序效率对比主流商业 CPU 并不逊色，并且后续还会有很大的优化空间。

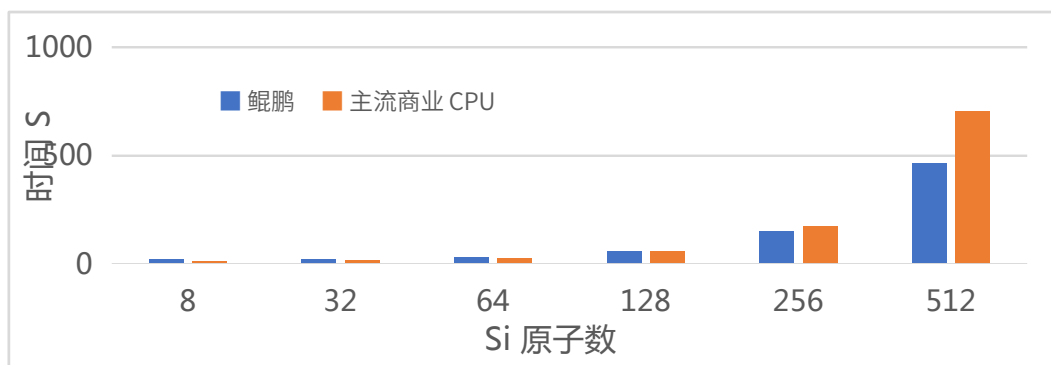


图 4-31 不同原子数的计算效率对比

此外，鲲鹏 920 HPC 计算平台采用独立 100Gb RoCE 网络，可以充分保证并行计算和后端存储数据交换性能。加上优化的 MPI 通道，降低应用分布式任务之间通信开销，可有助于程序轻松完成大尺度材料体系的科学研究。

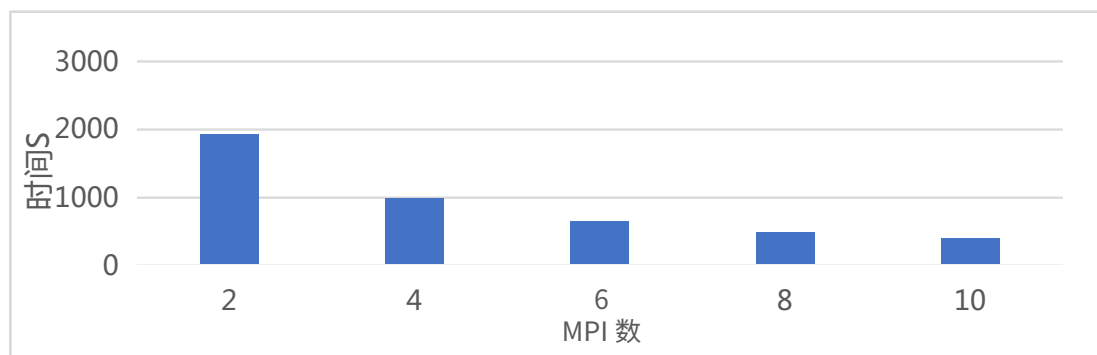


图 4-32 1024 个 Si 原子的自治计算可扩展性测试

## 4.19 北太振寰（重庆）科技有限公司

北太振寰（重庆）科技有限公司是北京大学重庆大数据研究院基础软件科学研究中心孵化的高科技企业，自主开发了北太天元数值计算通用软件 Baltamatica 对标 Matlab 以提供国产化替代。

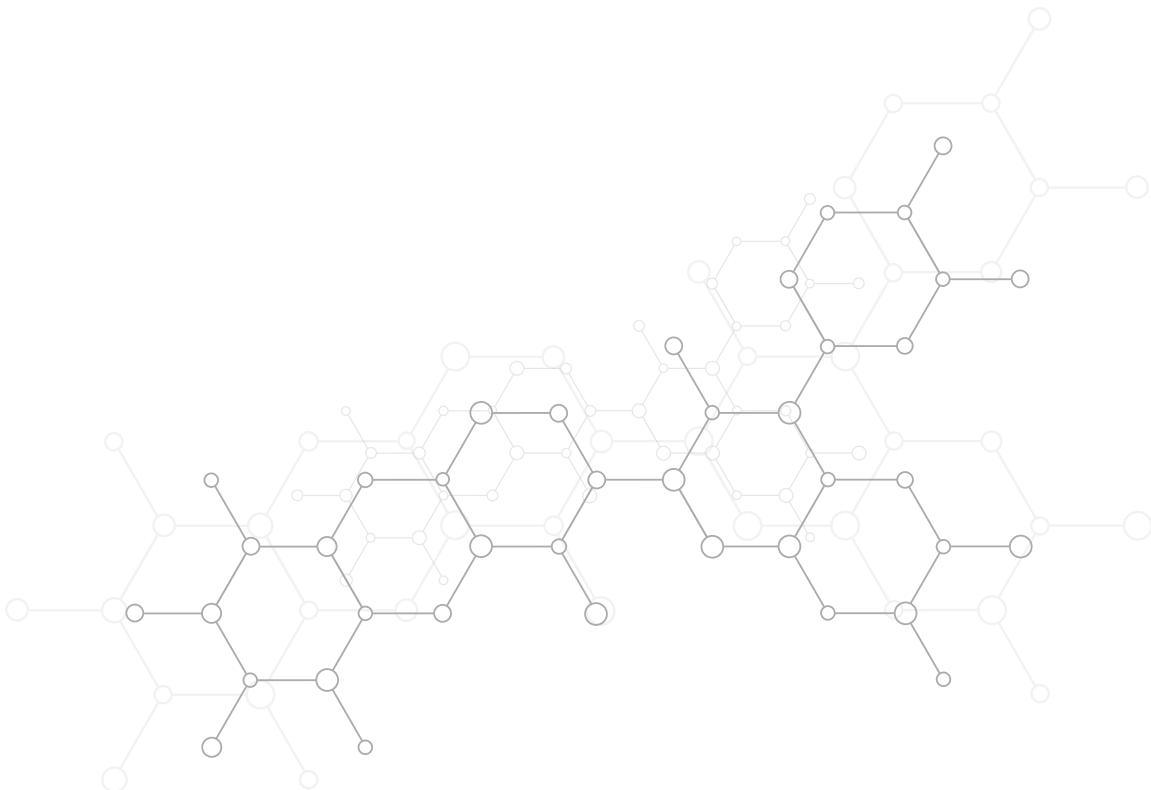
北太天元数值计算通用软件 Baltamatica（是面向科学计算与工程计算的国产通用型科学计算软件，融合数学、物理、计算机、工业工程等多学科理论知识与实践经验，围绕自主知识产权的科学计算软件底层根技术，发展国产数学软件、工业软件、专业软件等核心关键软件，构建以应用数学为基础的技术生态、应用生态与产业生态。本软件具有自主知识产权，提供科学计算、可视化、交互式程序设计，具备丰富的底层数学函数库，支持数值计算、数据分析、数据可视化、数据优化、算法开发等工作，并通过 SDK 与 API 接口，扩展支持各类学科与行业场景。

北太天元数值计算通用软件 Baltamatica 深度适配国产操作系统以及 ARM 架构国产处理器。北太天元软件的研发遵循 C++17 语言标准，可采用 CMake 工具进行跨平台编译。目前，软件已完成对飞腾等国产 CPU 及麒麟等国产操作系统的适配。在超算平台部署方面，北太天元已部署至天河新一代超算，在进一步实现北太天元的超算能力的同时，为用户提供更高算力要求的科学计算服务。北太天元软件所依赖的库拥有完整代码，或已存在适配麒麟平台的二进制库文件，在移植过程中，只需获取相应二进制库文件并编译相关源码即可完成移植，迁移成本较低。





图 4-33 北太天元数值计算通用软件 架构图



## 5 结语

随着 ARM 指令集兼容架构技术的发展，特别是 ARM 指令集兼容架构向服务器处理器领域进军步伐的加快，ARM 指令集兼容架构在高性能计算领域异军突起。特别是随着 2020 年搭载 ARM 指令集兼容架构处理器的日本 Fugaku 超级计算机荣登全球超算 TOP500 榜首，以及中国、美国、欧洲多台基于 ARM 指令集兼容架构超级计算机的建成，ARM 指令集兼容架构将在 HPC 领域的发展如火如荼，大放异彩。

同时，ARM 指令集兼容架构在 HPC 领域的基础软件生态已经基本完备，包括操作系统、编译器以及高性能支撑层软件生态已无明显短板；在此基础上，应用软件生态不断完善，重点应用领域不断扩展。目前，基于 ARM 指令集兼容架构的高性能计算机在高校、科研院所、超算、企业的应用越来越广泛，ARM 指令集兼容架构在 HPC 领域的市场化已经取得了巨大进展。相信 ARM 指令集兼容架构在 HPC 领域的发展会越来越好。

希望本白皮书的发布能够形成 HPC 产业界各个领域伙伴对 ARM 指令集兼容架构在技术、生态上的进一步认知，一起加入到这种“指令集兼容、生态开放、技术自主”的开放架构工作中来。

