

# 气象高性能计算系统的业务发展概述

■ 王彬 孙婧

中国气象部门从1978年建设第一台高性能计算机系统以来，逐步建立了支撑数值模式运行的高性能计算业务。高性能计算技术的发展日益呈现多态、异构和极大规模并行的趋势。高性能计算资源管理将逐渐从本地预先分配为主的方式过渡到本地远程统一调度、共享使用的方式。

基于并行运算的思想，高性能计算实现了对超级计算性能的需求。气象始终是高性能计算的重要应用领域之一。

## 1 高性能计算技术概述与现状

高性能计算（High Performance Computing, HPC）可定义为计算量巨大且快速高效的运算。承载支撑高性能计算运行的计算机系统即为高性能计算机系统，也被称作超级计算机。作为计算机科学与技术的一个分支学科，高性能计算HPC从体系架构、并行算法与应用开发等方面研发高性能计算机的技术。一般认为高性能计算、理论科学与实验科学三者共同构成了科学研究的支柱。

参考Flynn分类方法，HPC系统基本都属于多指令流多数据流类型计算机。不同厂商生产的HPC系统硬件组成部件大同小异，体系结构可分为Cluster与MPP，CPU包括RISC与X86等两类，内部互联技术有Infiniband（IB网）、OPA、高速以太网、其他专有网络等。作业调度管理软件有LSF、PBSpro、GridView、LoadLeveler及开源社区软件定制开发。并行编译开发采用Intel、GNU、PGI编译器以及MVAPICH、OpenMPI等并行运行环境。

目前HPC系统主要采用集群系统（cluster）架构。集群系统是指一组相互独立的服务器，包括计算、存储、服务、前后处理等用途，通过高速通信网络结构联结起来，构成统一的计算机系统，实现资源管理调度与高效并行处理，具有高可扩展性、可靠性与抗灾难性。

HPC系统的关键性能技术指标包括体系结构、总体规模、CPU（核）数量、峰值速度、Linpack测速、运行效率、访存速度、网络互联性能、每瓦特性能等。

超级计算机系统TOP500排行榜每年由国际HPC学术组织定期对全球高性能计算机系统评测，是业界普遍认可的衡量HPC技术水平的事实标准，成为了解

HPC发展现状和未来趋势的风向标。

最新一期世界超级计算机TOP500（2017年6月）数据显示：已经有138台系统超过千万亿次（PFLOPS），X86+Linux为代表的开放体系技术占据上风，CPU+加速部件（Accelerator Devices）混合架构的数量和比例不断增加。体系结构多为Cluster，内部互连网络以Infiniband技术为主，X86处理器芯片占据较大优势，Linux操作系统使用比较广泛。

尽管系统峰值性能有了很大的提升，但是HPC面临的计算、数据密集型应用需求也日益突出。HPC系统仍然面临着成本、能耗、并行程序支持的可用性以及有效沿用遗留代码（Legacy Code）等问题。

CPU性能的提升在过去几十年一直遵循着摩尔定律。利用不断提高的CPU主频，应用程序能够获得更好的性能。与此同时主频升高会加大能耗，导致散热、漏电等问题。在“后摩尔时代”，尤其是HPC峰值速度从万亿次发展到千万亿次量级过程中，仅依赖CPU芯片时钟频率的提高将无法满足应用要求，为此CPU+加速部件混合架构逐渐兴起，依托加速部件技术的快速发展和更好性价比加以支撑解决。

目前主流的加速部件技术包括GPU（Graphic Processing Unit）与众核。TOP500中混合架构构建系统的数量和比例不断提高，前10名系统已有5台采用了加速部件混合架构，排名第1的神威·太湖之光系统（Sunway TaihuLight）采用了国产申威众核技术，总计1064万个处理器核。

## 2 气象高性能计算业务发展

HPC技术的发展进步极大地推动了数值天气预报（NWP）的应用发展。从历史上看，NWP系统几乎在任一个时期都使用了同时代顶尖的HPC系统，NWP也只有使用最强大快速的HPC系统才能达到预报应用时效性要求。

基于计算技术的迅猛发展，气象HPC业务从无

到有地建立起来，有力地推动了NWP业务的发展。HPC系统成为中国气象局IT系统的重要组成部分，其计算峰值性能已成为气象部门现代化建设水平的标志之一。1978年11月，成功引进了日立M-170计算机，其计算能力每秒百万次、内存4 MB、存储达2.1 GB，在当时国内综合性能最强，用于气象数据处理和支撑MOS数值预报模式运行，结束了我国没有数值预报业务的历史。1980年7月，欧亚区域模式正式投入业务。1983年8月，亚洲区域模式开始业务运行，标志着我国NWP业务进入实用化阶段。

20世纪90年代中期以来，为适应气象NWP业务与科研工作的快速发展，气象HPC系统不断升级、能力显著增强。中国气象局国家级业务中心先后引进了银河II、克雷J90、克雷EL98、克雷C92、IBM SP2、IBM SP、曙光1000A、银河III、神威I、神威48I、神威32I、神威32P、IBM Cluster 1600、SGI Altix4700、神威4000A、IBM Flex P460等HPC系统，支撑了数值天气预报、数值气候预报业务运行与研究开发及卫星数据处理业务。

从20世纪90年代初至今，基本上每过5年，气象部门HPC计算能力就会跃升1个量级。随着IBM Flex P460系统在2014年正式业务运行，计算能力与“十一五”相比提高了将近30倍，比1978年更是提高了近10亿倍。

从机型架构演变来看，最初主要采用了通用巨型计算机，20世纪90年代中期起逐渐转向大规模并行架构计算机。从发展历程来看，2000年之前以进口HPC系统为主；2000年之后国产与进口HPC并驾齐驱，分别引进建设了3套进口机系统与3套国产机系统。

中国气象局国家级业务运行的主要HPC系统为IBM Flex P460。安装在国家气象信息中心二楼高性能计算机房，承担了国家级数值天气预报业务与准业务、短期气候预测业务以及研究开发等工作。此系统主要由P460服务器组成，共计37120个Power7 CPU核。系统总计算能力达1054.2 TFLOPS，内存总量163584 GB，存储物理容量超过4.2 PB。IBM Flex P460高性能计算机系统包括两个同样配置的子系统，每个子系统设计能力达527.1 TFLOPS，存储容量达2.1 PB。该系统内部采用无阻塞胖树结构的InfiniBand网络互连，构成了内部高速数据计算网络。该系统采用了冷却水制冷，所有机柜均配备了水冷背门，通过冷却水可以带走绝大部分的热能，大大提升制冷效率，节省能源消耗。

2000年以来，随着区域级气象数值预报工作的不

断发展，部分区域、省级气象部门根据自身业务发展需要，先后建设了不同规模的HPC系统，提升了本区域高性能计算能力。

HPC资源相对于气象部门需求来说始终是稀缺资源，由于国家区域之间，以及区域之间数值预报发展的不平衡，再加上计算资源在部门内分布的不均衡，需推进分布在国家局及各区域中心的HPC资源共享使用，统一分配管理、监控和资源共享，解决部门内国家级、区域的超算资源整合协同与共享管理。

制订异地HPC资源共享调度策略，需要考虑不同地点网络带宽、NWP模式输入数据环境、NWP模式输出产品的数据量和方式、其他系统使用模式产品的方式等因素。根据现有国内网络带宽租用成本、气象HPC系统互联网络条件以及模式数据环境现状，气象现有HPC资源管理应以本地使用优先为原则。针对国家级、区域级HPC系统，基于资源预分配应用方案，建立起资源使用统计、分配审计技术手段。结合每年用户、单位的资源需求与前一年资源实际使用情况，计算相应的年度资源分配额，协同调配系统资源，实现国家级、区域级HPC资源的合理使用与规划。通过建立气象HPC监控统一平台，实现国家、区域两级HPC系统资源使用与NWP业务运行状态的实时监控。

由于业务型和科研型数值模式作业运行具有不同的特点，相应地采取不同的管理方式。在了解各HPC系统资源负载状况的基础上，优先使用用户本地的计算资源，在本地资源不足的情况下考虑调配异地资源。在国家级中心建立数值模式统一运行管理平台，统筹规划，实现国家、区域两级NWP业务的运维管理和统一调度。由于业务型作业相对成熟稳定，输入、输出数据流程固定，因此此类作业将主要在本地运行。如区域中心本地资源无法满足数值预报业务模式，可以通过资源预分配管理调配至国家级资源运行。在平常时段，业务备份可采用冷备份的方式进行；一旦在重大气象服务或重大气象灾害应急等特殊时期，通过热备份的方式同时运行某一业务模式，保障业务的可靠运行。为了保持备份模式与原主业务系统的数据同步，利用云服务获取数值预报模式启动所需的初始场等数据并在备份系统启用后接收异地运行生成的模式产品，从而实现数据同步。

由于研发型模式作业具有应用多样、资源使用量不定的特点，按照本地优先、异地调配的原则分配使用资源。直接登录高性能计算机的方式使用资源，限于目前的网络性能及其他客观条件，远程用户不通过网络传输大数据，可直接在异地系统完成模式运行及

后处理, 在需要将结果回传到本地, 对必需的大容量数据或可采用硬盘拷贝方式。支持科研应用异地提交, 提供基于命令行登录和Web门户的两种资源访问方式, 通过浏览器方式实现可视化的配置试验运行控制和跨集群模式作业提交。

### 3 展望

随着多核与众核架构处理器、超大规模并行处理、异构编程语言、GPU计算、大数据、云计算等新技术的出现, 科学计算将经历思维和技术两方面的变革, 业界正处于计算技术发展的转折点, 未来的发展趋势必定是多方面技术的融合, HPC技术发展呈现出多态、异构和极大规模并行的趋势。国内外厂商正在加紧研制E级(百亿亿次)高性能计算机系统, 预计很可能在2018年左右出现。可以预见异构加速架构计算机将会迎来更大发展, 轻量核技术会推动处理器与加速设备一体化, 编程软件工具链日趋成熟。

按照气象信息化发展要求, 高性能计算机系统建设、资源管理、技术开发将会更好地支撑气象数值模式软件业务与科研的发展。因应技术发展潮流, 加大众核、GPU计算等新技术应用力度, 注重并行计算交叉型人才培养, 推进气象业务模式向新型并行技术架构平台的移植优化研发, 提高并行可扩展性。需要继续统筹集约气象部门内高性能计算资源的分布、建设与管理, 逐渐减少地理上分散的小规模系统。为了满足数值天气与气候预报模式业务运行和科研开发工作的需要, 依托工程项目, 将在2018年建成新一代

曙光国产系统, 缓解资源紧张局面, 支撑数值天气预报、气候预测与气候变化等业务与研究。高性能计算资源管理未来将逐渐从本地预先分配为主的方式过渡到本地远程统一调度、共享使用的方式。基于国家、区域中心统一的全国综合气象信息共享平台(CIMISS), 实现数值预报业务应用和模式研发等应用跨系统的动态统一调度, 实现业务应用系统可靠运行和异地备份。

致谢: 本文由中国气象局局校合作项目“气象模式众核平台移植开发”资助。

#### 深入阅读

- 李国杰, 2010. 21世纪上半叶信息科学技术展望. 中国科学院院刊, 25(1): 78-86.
- 沈文海, 2012. 从云计算看气象部门未来的信息化趋势. 气象科技进展, 2(2): 49-56.
- 沈文海, 2013. 对气象信息化的理解和再认识. 气象科技进展, 3(5): 56-62.
- 沈学顺, 苏勇, 胡江林, 等, 2017. GRAPES\_GFS全球中期预报系统的研发和业务化. 应用气象学报, 28(1): 1-10.
- 孙婧, 沈瑜, 2015. 气象应用的高性能计算机性能需求推算方法. 计算机技术与发展, 25(6): 206-210.
- 王彬, 2014. 高性能计算技术在气象部门的应用. 计算机工程与设计, 35(4): 1476-1479.
- 王彬, 肖文名, 李永生, 等, 2011. 华南区域中心计算资源管理系统的建立与应用. 气象, 37(6): 764-770.
- 赵立成, 2011. 气象信息系统. 北京: 气象出版社.
- 宗翔, 王彬, 2006. 国家级气象高性能计算机管理与应用网络平台设计. 应用气象学报, 17(5): 629-634.

(作者单位: 国家气象信息中心)

### 全球信息技术创新前10机构

本报告使用科睿唯安旗下知名数据库与信息分析平台进行数据的收集、整理与分析, 数据来源涵盖1.8万多种期刊、大量会议录文献、图书节选及其他文献材料, 以及超过7100多万件专利信息。

——摘自《2017年全球创新报告: 进无止境》, 科睿唯安, 2018年

图 50: 信息技术领域全球排名前十位的创新机构 (2016)

公司	国家/地区	发明数量
State Grid Corporation of China (国家电网公司)	中国	8231
IBM	美国	7433
Canon (佳能)	日本	5982
Samsung (三星)	韩国	5746
ZTE (中兴通讯)	中国	3317
Ricoh (理光)	日本	3108
LG	韩国	2746
Huawei (华为)	中国	2740
Google (谷歌)	美国	2706
Lenovo (联想)	中国	2595