

# 超大规模计算的实现与应用

作者: 刘天海



国际商业机器公司（IBM）研发的超级计算机“Blue Gene/P”

人类自身的计算能力不适合重复、繁杂的计算任务，于是，人类创造了计算机来弥补不足，并一直锲而不舍地追求更强大的计算能力。现在的单核CPU可以稳定地进行每秒30多亿次的运算，但也到达了芯片制造工艺的极限。人多力量大，基于这个朴素的认知，为了获取更快的计算速度，我们一方面在单个CPU芯片里塞进更多的核心，另一方面把更多的计算机连在一起、组成一个超级大的计算机。为什么这样就能算得更快呢？我们来看一个例子。

老师要求学生计算从1到100的和，不用任何公式，就是简

单加法运算。老师指定了两种方法，第一种是让一个学生自己算，第二种是把任务平均分给50个学生，每个学生先算连续两个数的和，得到的50个中间结果再分配给25个学生，计算得到的25个中间结果再分给12个学生，以此类推得到最终结果。假设每次加法运算用时都是1分钟，则第一种方法用时99分钟，而第二种方法用时7分钟+7次分配任务所花的时间。只要分配任务所花时间足够小（小于13分钟），那么我们就可以通过这种方法取得计算速度的提升。这种方法在超大规模计算领域被称为并行可扩展，就是聚合计算资源的共同协作。



超级计算机可以通过并行处理来更快地工作，即把问题分解成多个部分，每个部分同时由单独的处理器处理，最后整合得出结果

这个例子描述了超大规模计算的一个简单应用，但是真实应用的计算工作量更大，达到几万亿亿次；计算也更为复杂，不再是100以内的加法，而是求解复杂的偏微分方程。比如，一个描述一根两端固定的铁链在重力作用下呈下垂形状的偏微分方程，即使微积分发明者之一的牛顿也要花好几天才能算出很低精度的结果，而一台超级计算机可能只需要几分钟就能算出高精度的结果（数字在 $\pm 1.79 \times 10^{308}$ 之间且精确到小数点后16位）。正因为有了这种高算力的加持，我们才能更快更准确地探索世界和预测未来。

看到这里，你可能开始好奇这个“怪物”到底是什么样的，接下来就让我们一起走进超级计算机。

## 天下武功，唯快不破——超级计算机

超级计算机（Super Computer，简称超算，也称高性能计算机）是超大规模计算的一种实现。在最基本的层面上，它是指聚合计算资源以提供更高的计算能力（简称算力）。这种算力通常用于解决国防、科学或工程领域中的大型或复杂问题，如对一些事故发生率较高、对生命安全造成极大威胁的高危

## 浮点运算

我们平时使用的小数在计算机中通常用浮点数来存储，通常以科学计数法表示，浮点计算就是它的四则运算，通常因无法精确表示而使用近似值，所以有单精度浮点数和双精度浮点数两种浮点格式。双精度浮点数使用64位来存储一个浮点数，表示的数字的绝对值范围为 $2^{-1024} \sim 2^{1024}$ 。每秒浮点运算次数（FLOPS）常用来测量运算速度。

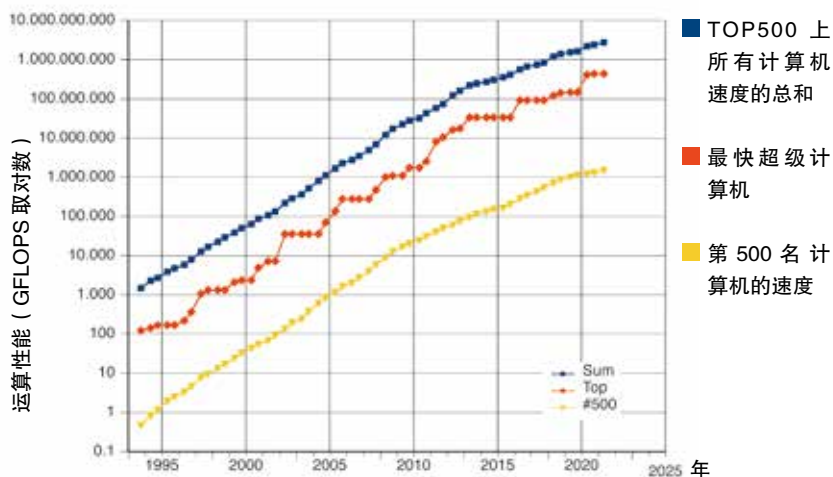
行业，以超级计算机代替人工进行作业，如地下采煤、高空作业、爆破工作和石油勘探等数据的处理和分析。

中国超算的起步不算晚，从1983年“银河一号”的诞生，在这40年里陆续研制出一系列超级计算机，如银河、天河、曙光、KD、神威和深腾，其中，“神威·太湖之光”是中国在2016年使用自主研发的芯片研制出来的，并且占据全球超算第一的位置达两年之久。“神威·太湖之光”坐落于无锡太湖旁边，由48个标准机柜（40个运算机柜和8个网络机柜）组成，每个机柜高2米、宽0.8米、深1.4米，和一个双开门冰箱差不多大，总占地605平方米，相当于一个网球场的面积。它采用28纳米制程的申威CPU 1.5 Ghz，每个CPU有260个核心，总共约1065万个核心，130万GB的内存，每秒可进行约9亿亿个双精度浮点运算，总造价18亿人民币，耗电量14兆瓦，运行一小时所需的电量足够中国一个普通家庭用10年。现在中国拥有世界最快的500个超级计算机中的173个，其次是美国，有149个。

2021年最快的超级计算机是日本神户的“富岳”（Fugaku），共432个标准机柜，占地2000平方米，总造价64亿人民币。“富岳”使用的是7纳米制程的AMD Fujitsu CPU 2.2 Ghz，每个CPU有48个核心，总共763万个核心，500万GB的内存，每秒可进行44亿亿个双精度浮点数运算，当然耗电量也达到了30兆瓦之巨。

接下来我们将介绍世界超算主流系统的组织结构和基本软硬件，对于如CPU内部的并行技术、内存并行、高速网络内连、内核优化和编译优化等具体细节先暂且不提。

国际TOP500组织是发布全球已安装的超级计算机系统排名的权威机构，由美国与德国的超算专家联合编制。从TOP500发布的数据中可以看出，超级计算机的综合性能正逐年快速增长（Credit: Top500.org）



## 集群

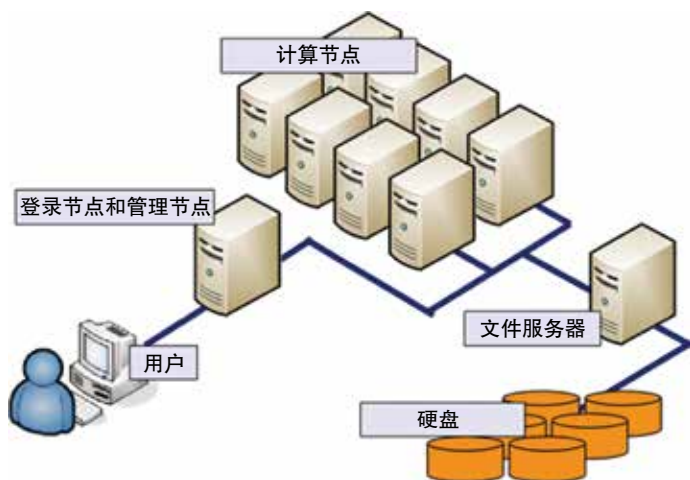
超级计算机500强中90%的计算机采用了集群体系结构。集群就是采用普通微型机或高端商用机作为计算节点，并采用高速网络互连的并行计算系统，节点之间通过发送和接收消息完成通信，如接收计算任务和返回计算结果。消息传递的延迟和分配计算任务的合理性从根本上决定了整个集群系统的计算性能。因为每个计算节点拥有独立的内存，这种分布式的内存结构使集群具有很高的可扩展性，即我们可以通过增加或减少计算节点的数量达到调整超级计算机算力的目的。

集群可分为五类设备。

(1) 登录节点。登录节点是普通用户访问超级计算机的门户。用户通常登录到此节点上编译、提交并开启总的计算任务。登录节点一般对算力要求不高，可能几台普通服务器即可。一个计算任务可能需要从用户提供的几百个硬盘中一点点地拷贝到登录节点。

(2) 管理节点。管理节点和集群内的计算机互连并且对其进行控制，如开关机、升级计算节点的操作系统等。管理节点对算力要求也不高。管理节点和登录节点也可以是同一设备，管理员和普通用户具有不同权限，他们只能执行其权限之内的操作。

(3) 计算节点。计算节点是整个集群的计算核心。总计算任务开启后，计算节点之间会互相通信完成任务分配，并且在计算过程中进一步调整下一步的任务分配。如果有的计算节点算得特别慢且有其他节点要等待它的计算结果，那么这个计算任务就会被同时分配给算得快的节点，不管哪个先算完，计算



超级计算机中的绝大部分都采用了集群体系结构。用户通过登录节点访问计算集群，提交任务给管理节点，管理节点将任务分解，合理分配给各个计算节点开始进行并行工作，然后收集计算结果，存入文件系统或硬盘中

结果将传给等待中的节点。

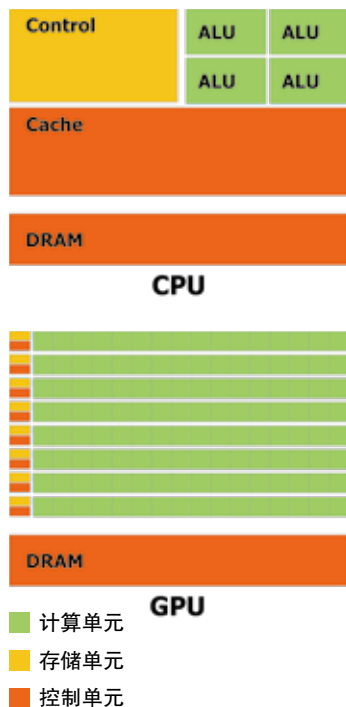
(4) 网络设备。网络设备包括高速交换机和用于节点互连的线路。当前大多超级计算机使用InfiniBand专用网络设备，一般的12X链路网络速度可达每秒290G比特，是5G网络的58倍。

(5) I/O设备和存储设备，即超级计算机的外设和基于光纤或采用SCSI接口的硬盘，用于存储计算数据和结果。

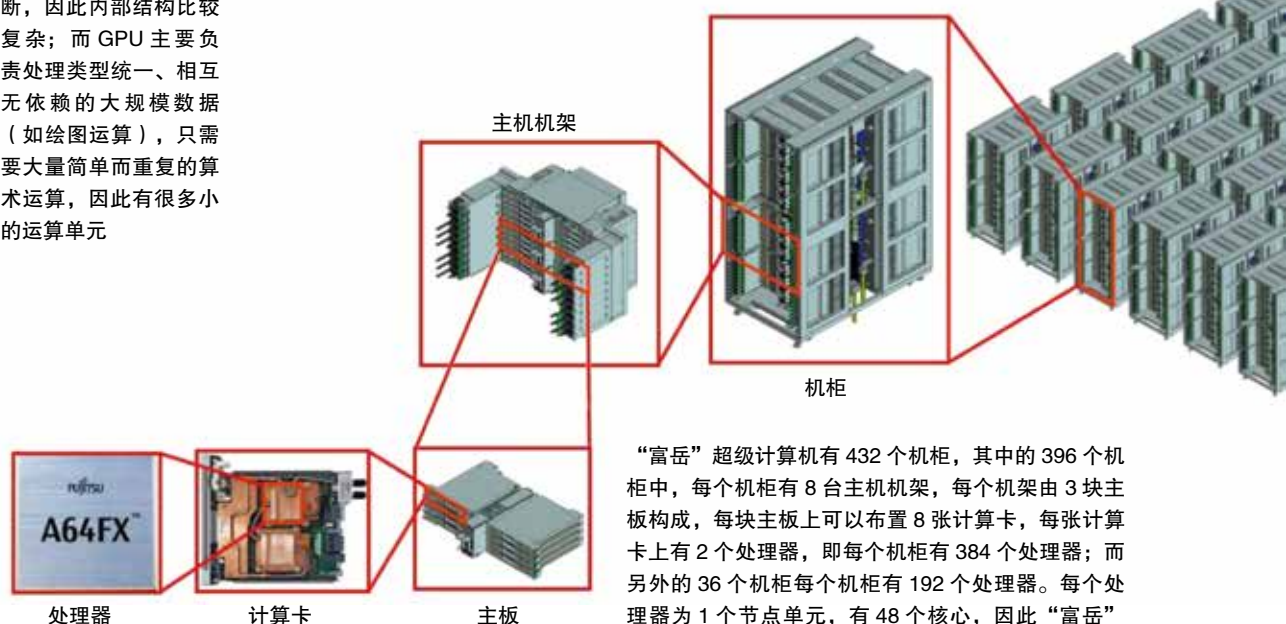
## 硬件

超级计算机计算节点的硬件组成与个人计算机的基本相同，都符合冯·诺伊曼结构。除了必要的CPU，超级计算机会在计算节点中使用GPU来大幅提升计算效率。

CPU是为大多数计算任务提供动力的通用处理器，而GPU最初是为渲染屏幕图形而开发的，渲染不需要逻辑运算，但需要大量简单而重复的算术运算，用很小的集成电路就能完美实现这种用途单一的运算，所以GPU的内核比CPU的要小很多。为了使渲染过程更高效，GPU里面被塞进了更多的内核。如Nvidia GeForce GTX 980 Ti拥有2816个核心，而同时期的CPU却只有32个核心。如果把CPU比作精通各种计算的博士，那GPU就是2000个只会加法的小学生。相比于博士，计算2000道简单的加法题肯定还是由2000个小学生同时计算会更快一些，这就是为什么超级计算机普遍采用



CPU 和 GPU 由于针对不同的应用场景，其架构也不相同。CPU 需要很强的通用性来处理各种不同的数据类型，且需要较强的逻辑判断，因此内部结构比较复杂；而 GPU 主要负责处理类型统一、相互无依赖的大规模数据（如绘图运算），只需要大量简单而重复的算术运算，因此有很多小的运算单元



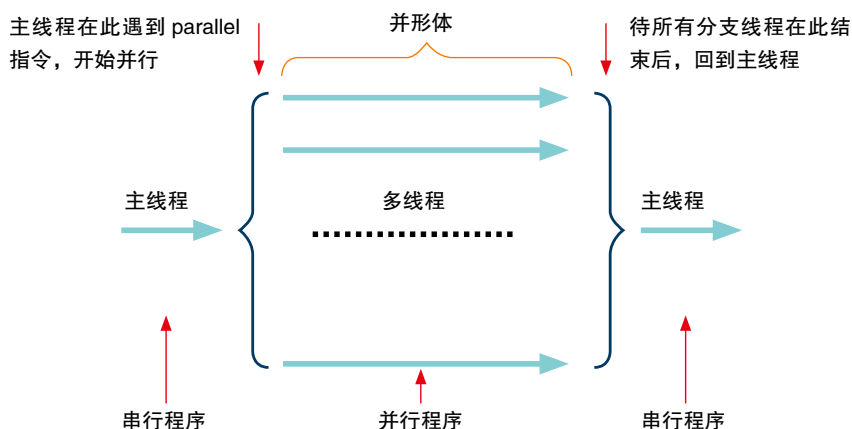
“富岳”超级计算机有 432 个机柜，其中的 396 个机柜中，每个机柜有 8 台主机机架，每个机架由 3 块主板构成，每块主板上可以布置 8 张计算卡，每张计算卡上有 2 个处理器，即每个机柜有 384 个处理器；而另外的 36 个机柜每个机柜有 192 个处理器。每个处理器为 1 个节点单元，有 48 个核心，因此“富岳”共有 158 976 个节点、763 万个核心，每秒可进行 44 亿亿个双精度浮点数运算

GPU作为加速器来加速大量简单的算术运算。

## 软件

高性能计算机在硬件方面是非常强大的，但同时为了发挥最大的性能，软件方面也很重要。首先在操作系统方面，几乎所有超级计算机都使用Linux操作系统，这得益于Linux与生俱来的开源策略和模块化设计。就像搭积木一样，人们可以很方便地往操作系统中添加新的驱动模块以支持新的硬件。其次，在编程语言方面，为了开发高性能的应用程序，C和C++是主流程序开发语言。最后，借助于并行化编程框架，人们可以专心地解决专业方面的问题和实现应用的功能，至于如何进行合理的计算任务分配，如何使计算节点之间协调合作，则由并行编程框架来完成。

我们将看一段利用并行编程框架OpenMP实现的100个数求和的C++程序。OpenMP是一套C++并行编程框架。它能使普通代码经过很小的改动自动转化成并行的程序。具体实现是通过分析OpenMP编译原语`#pragma`，将用原语定义的代码块自动转化成并行的线程去执行，最后再合并线程结果。子句`num_threads`控制创建的线程数，即分配几个线程来完成计算任务。`parallel for`用于生成一个并行域，并在多个线程之间分配计算任务，从而加快计算运行的速度。（更高级的控制字句请参考OpenMP官方文档。）例子中的计算任务被分配给10个线程，每个完成后把计算结果传给下一个计算节点，最终得到计算100个数的和。这段程序的执行时间是未使用OpenMP的1/5左右。



OpenMP 是一套 C++ 并行编程框架，它可以让普通代码经过很小的改动自动转化成并行的程序。即开始时只存在一个主线程，当需要进行并行计算的时候，派生出若干个分支线程来执行并行任务；并行代码执行完成后，分支线程会合，再并把控制流程交给单独的主线程。

当然，并不是线程数越多效率就越高，就像前文提到的，数据节点之间的低延迟和合理快速的任务分配也起着决定性作用。感兴趣的读者可以自己动手实验一下。

```
1  /*导入OpenMP头文件以便在代码中使用OpenMP指令*/
2  #include <omp.h>
3  /*程序运行始于main函数，终于最后一个大括号，并且大括
4   *号里面的语句被依次执行*/
5  int main()
6  {
7  /*定义一个变量sum来存储计算的结果，初始值为0*/
8      int sum = 0;
9  /*当编译器遇到编译原语#pragma时，将编译工作交付给
10   *OpenMP来处理for语句块，处理完后收回编译权*/
11      #pragma omp parallel for num_threads(10)reduction(+:sum)
12  /*OpenMP为每个线程创建并编译一个新的for语句块*/
13      for(int i=1; i<=100; i++)
14      {
15  /*求和*/
16          sum = sum + i;
17      }
18  /*在显示器上打印结果*/
19      printf("Sum of first 100 natural numbers is %d\n", sum);
20  /*程序终止，返回0代表程序运行至此没有错误*/
21      return 0;
22  }
```

一个真实超算应用的程序不需要做业务处理，在应对各种各样的交易时，它的代码行数可能不是很多，但其算法非常复杂，并且必须高度优化，从而最大化地发挥超级计算机的性能。

## 旧时王谢堂前燕，飞入寻常百姓家——超大规模数据中心

超级计算机主要面向国防、政府、研究机构等的运算任务，而用于私人或私企相关的工作会有一定阻碍。超算技术的进步推动了超大规模数据中心（HDC，Hyperscale Data Center）的建立，使个人或民间团体直接使用超大规模计算成为现实。

超大规模数据中心具有世界一流的计算性能和高效的资源

使用效率，包括容纳了大量数据的高性能计算机（又称为服务器）、网络设备和通信设施等，可以大规模高效地进行计算。当我们使用数字服务时，如微信付款，数据中心正在幕后执行许多繁重的工作，使数字服务无缝运行。

IT和云行业战略情报公司Synergy Research Group最近研究了世界各地的HDC。截至2021年底，HDC容量在四年内翻了一倍，数量增至659个，其中美国企业占据半壁江山，尤其是微软、亚马逊和谷歌占据了当前总数目的一半以上。当然，运营商的资本支出也不断地刷新纪录——过去四个季度已经达到1500亿美元。

超大规模数据中心也是超大规模计算的一种实现。它和超级计算机一样使用集群结构，但有几点不同。

（1）超算侧重计算节点的高算力，而不期望计算节点之间频繁大量的数据传输；HDC侧重海量数据的存储和传输，而相对忽视了算力。因此，超算的计算节点多，属于高端机，而HDC大都是普通的PC机并且拥有分布式文件系统和数据库提供海量的数据存储能力。

（2）超算面向国防、科技、工业等国家级计算任务，而HDC主要面向大数据业务。每天，我们创建约2.5万亿字节的数据。随着IOT（物联网）的日益普及，这种数据创建率将变得更大。另外，国家级的超算任务因为机密原因也不允许在商业级



随着每年数据信息量的快速增长，人们需要数据处理能力更强大的服务器。网格计算是一种造价低廉而数据处理能力超强的计算模式。它利用互联网把分散在不同地理位置的电脑组织成一个虚拟的超级计算机，其中每一台参与计算的计算机都是一个节点，而整个计算是由成千上万个节点组成的一张网格，协同解决靠本地资源无法解决的复杂问题

的HDC上执行。

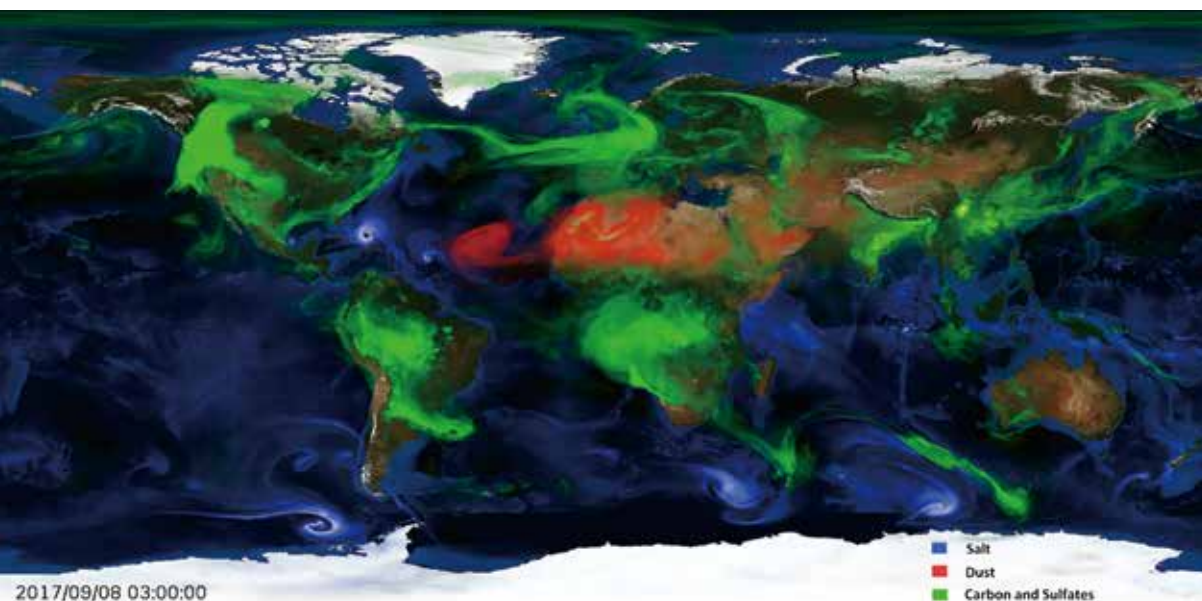
(3) 使用超算得经过层层审批，获取使用权后还要排队等候，并且计算资源的大小（如算力配置等）一经批准就无法随时更改，所以时间紧迫的计算任务一般处于很尴尬的境地。HDC就方便多了，如果有可用的资源，用户可以随时申请，自动获得使用权，也可以提前预订计算资源，使用方式灵活多变。另外，云计算赋予了HDC灵活的资源调整策略，用户可以按需分配计算资源。

## 天生我材必有用

### 数值天气预报

我们现在经常通过手机APP查看天气预报，尤其是准确率高的数值天气预报，它能告诉你几点、多大概率会下雨。这要归功于超级计算机对未来天气的推算。比如一个城市的数值天气预报：首先是数据收集，遍布城市各个角落的天气观察站将每隔几分钟收集一次天气数据，这些数据包括观察站所在地的风向、风速、温度和湿度等，同时，还有卫星、气象气球、海洋浮标等提供的各类气象数据；这些海量的天气数据被输入超级计算机中，计算机里已经有了气象学家根据大气运动变化规律总结的一系列偏微分方程组；超级计算机将根据已知的天气变化数据求解偏微分方程组，进而在规定的频次计算出短期、中期、长期不同气压高度的天气变化趋势，并以气象图的方式呈现，比如逐6小时的2小时累计降水分布图。

超级计算机能够根据海量的气候数据计算出短期、中期、长期不同气压高度的气候变化。图为戈达德地球观测系统正向处理（GEOS FP）模型生成的全球气溶胶预报，它每天每6小时运行一次，每次采集约400万次观测数据（Credit: NCCS）



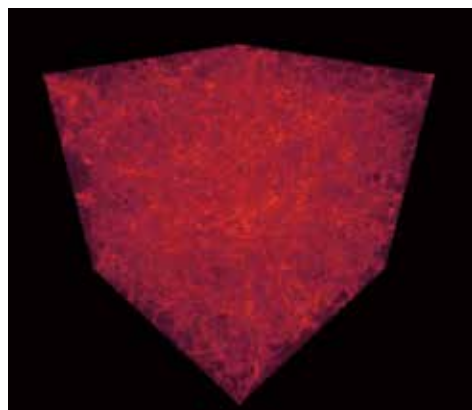
中国超级计算长沙中心就为中部某省气象局提供了数值天气预报计算的 platform 支持。业务化运营后，该气象局数值预报能力大幅提升，水平分辨率从20千米提升到4千米，暴雨水平分辨率从37千米提升到15千米，离预测到“隔河不下雨，十里不同风”又近了一步。

## 宇宙演变

暗物质大约占到宇宙所有物质的84%，余下的16%则为我们熟悉的普通物质。暗物质看不见摸不着，甚者电磁波也探不到，但它们在星系的形成中产生了巨大的影响。科学家发现暗物质只受引力作用。在忽略少量普通物质的其他作用力后，N体模拟是获得暗物质在宇宙里的运动轨迹及分布方式的一种非常理想的方法。为了分析暗物质在宇宙从混沌到形成某些特定结构的过程中所起的作用，国家天文台和中科院超算中心合作，在超级计算机深腾7000上进行了千核规模的模拟。他们用一个立体盒子里的 $N$ 个质量相同的点粒子代替宇宙中一个早期区域内的大量暗物质粒子，并根据宇宙学模型给予这些粒子原初位置和速度，这样通过这 $N$ 个粒子在引力下的运动，我们就能“看到”暗物质粒子最后形成的结构。在根据模拟结果绘制的视频中，每一帧代表某个时间点的宇宙状态，亮点代表此处已经形成了星系。这些巨大的亮点被很多丝状结构和片状结构连接，而这些丝状或片状结构围成的区域被称为空洞。如果所有暗物质都能被我们看见的话，那么宇宙将是图中这样的网络状结构，没有超级计算机，我们可能很难看清这一切。

## 工业工程模拟

当高速列车时速大于350千米时，气动噪声将成为其主要噪声源。为了提高高速列车乘坐的舒适性、缓解噪声对铁路沿线居民的影响，中科院超算中心和力学所合作，通过数值模拟手段，结合理论和实验研究，对高速列车的气动噪声进行了预测评估，提出了列车关键部件的降噪优化设计方案。



超算可以模拟暗物质主导下的宇宙演化过程。图为某个时间点的宇宙形态，亮点代表此处已经形成了星系。扫描视频二维码，可观看在暗物质的作用下，宇宙逐渐从混沌到形成某些特定结构的过程，这一可视化结果有助于天文学家观察到感兴趣的科学现象

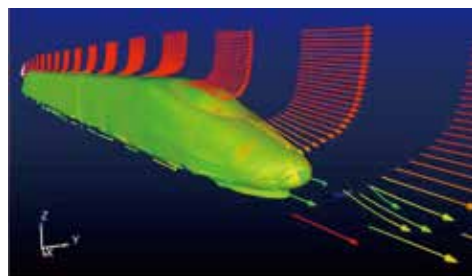


扫码观看视频



扫码观看视频

超算可预测评估高速列车的气动噪声，并提出针对列车关键部件的降噪优化设计方案。图为软件模拟的列车在高速行驶时周围的气流运动

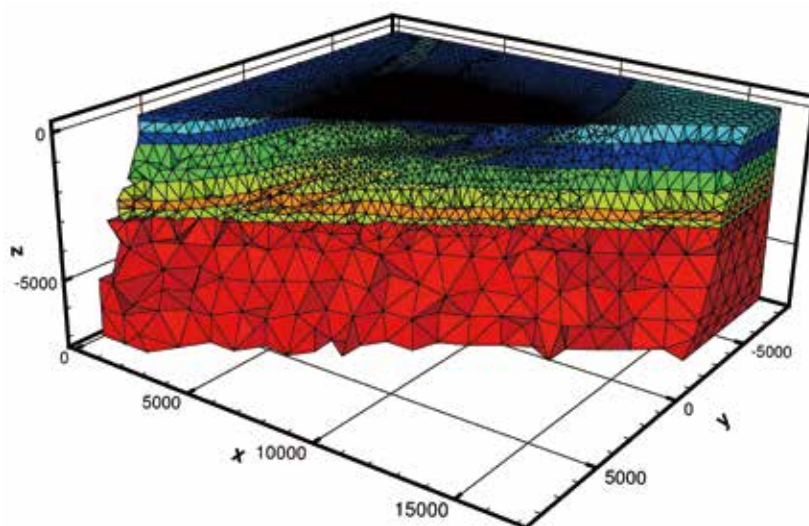


## 地质勘探

当今的地球物理学利用大量观测数据集和大规模计算模拟来提高我们对地球的认识。因此，固体地球模拟是高性能计算的最大应用之一，范围从地震模拟到勘探数据处理。比如，勘探一个 $20 \times 20 \times 20$ 立方千米（长 $\times$ 宽 $\times$ 深）地块的地下结构（实际勘探中地块会比这大得多），勘探队可能按4千米或其他大小的间距人工制造可控地震，然后地面接收装置收到地下反射的地震波并存入硬盘。地震波既有横波也有纵波，采集的数据量非常大，通常是几十TB，采集完后将数据硬盘带到超算中心进行建模。建模就是把这个地下结构分成很多网格，每个网格向周围网格反射或传递地震波，直至模型模拟的地震波与野外采集的匹配。为了使模型达到足够的精度，通常模型的网格要足够小，比如 $50 \times 50 \times 20$ 立方米的小区块，那么这个模型就有1.6亿个网格，计算机就要求解1.6亿个未知数，并且这些未知数互相依赖或独立。具体来说，就是求解有1.6亿个未知数的线性方程组。这只是对一次测量数据的建模，勘探人员经常需要从不同位置对同一地块多次制造人工地震并收集新的数据，这就要求一个模型需要多次计算才能满足石油勘探的要求，所涉及的计算工作量和难度可想而知。

石油勘探公司都有自己的超级计算机来满足勘探数据分析需求。早在2000年左右，中石油集团就开始采购曙光4000L超级计算机用于石油勘探，现在它是中国最大的地震数据处理中心。

超算可以模拟固体地球，范围从地震模拟到勘探数据处理。为了使模型达到足够的精度，地下结构被分成多个足够小的网格，如 $50 \times 50 \times 20$ 立方米的小区块有1.6亿个网格，即每次计算要求解上亿个未知数，而模型需要多次计算才能满足石油勘探的要求，计算工作量和难度可想而知



## 动画渲染

在动画电影制作中，人们先在电脑上绘出动画人物和场景环境，然后根据剧本设定动画人物表情动作和场景，每个人物的眼睛、眼睫毛、头发丝、衣服褶皱等都要独享一条视频轨道，并且场景中的每道光线、每扇窗户，甚至每片树叶也要独享一条轨道，最后动画渲染就是把所有视频轨道在同一时间点上的内容合在一处生成一帧二维或三维图片。如果导演不满意渲染出来的图片或动画片段，我们就只能“再来一条”，即调整人物表情动作或场景后进行再次渲染。一分钟的动画片段有1440帧（60秒×24帧），4K超清电影有800万个像素（4096×2160），如果每帧渲染需要30秒（对于4K电影来说已经很快了），那么我们将需要12小时的渲染时间。动画渲染是一项没有技术难度但工作量极大且很耗时的工作，而这正是超级计算机的专长。目前，动画渲染多采用超大规模数据中心实现。从《西游记之大圣归来》《大鱼海棠》到《哪吒之魔童降世》，再到《白蛇2：青蛇劫起》，国内大部分动漫制作公司已经接受这种Renderbus自助式渲染模式，总渲染时长已经超过3亿核小时。



动画渲染是把所有视频轨道在同一时间点上的内容合在一处生成一帧二维或三维图片，是一项没有技术难度但工作量极大且很耗时的工作，因此动画长片的渲染适合用超算进行处理。上图是 Blender 的动画渲染界面，这是一款开源的跨平台全能三维动画制作软件，提供从建模、动画、材质、渲染、到音频处理、视频剪辑等一系列动画短片制作解决方案（Credit: Blender）

## 潮平两岸阔，风正一帆悬

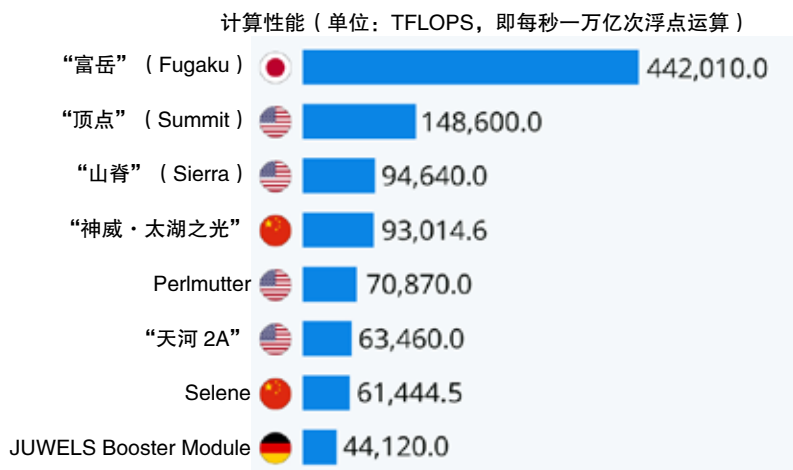
### 未来超算

根据专家研究和相关新闻报道，未来超大规模计算的发展趋势包括但不限于以下几点。

(1) 速度继续提升。算力仍是超算主要追求的目标。当前，世界TOP500的超级计算机基本都可以达到P级（千万亿次）运算速度，各个国家正在E级（十亿亿次）超级计算机的研制中激烈竞争。根据历年来超级计算机速度提升的情况，TOP500组织曾预计E级超级计算机将有望于2020年出现。但由于全球疫情对经济的冲击，半导体产业近3年来发展缓慢，甚至出现了芯片荒的情况。E级计算机的出现会晚一些到来。中国有报道称已经有两款系统达到了E级别的速度，现在正在部署之中。美国“国家战略计算项目（NSCI）”计划在2023年底开发完成2套E级超级计算机系统。

(2) 体型变小，能耗变少。超级计算机的体系在近10年没有显著变化，仍以集群为主流体系结构。在架构不变的基础上，比起单纯依靠增加计算资源来极大地提升算力，人们更愿意看到量子计算机和生物计算机带来的技术融合。量子计算机擅长解决几类传统计算机难以解决的问题。生物计算机具有生物体的一些特点，如发挥生物本身的调节机能，自动修复芯片上发生的故障，还能模仿人脑的机制等。生物计算机的运算速度要比当今最新一代计算机快10万倍，能量消耗仅相当于普通计算机的十亿分之一，还具有巨大的存储能力。

超级计算机的计算性能排行榜（数据截至2020年11月），日本的“富岳”是目前最快的超级计算机，每秒可进行44亿亿个双精度浮点数运算，几乎是世界上第二强大的超级计算机的三倍（Credit: Top500.org/Statista）



## “东数西算”工程

随着数字产业的飞速发展，中国数据中心规模已达500万个标准机柜，算力达到130EFLOPS（每秒一万三千亿亿次浮点运算）。随着数字技术向经济社会各领域全面持续的渗透，全社会对算力需求仍十分迫切，预计每年仍将以超过20%的速度快速增长。

目前，中国的数据中心大多数分布在东部地区。在寸土寸金、资源紧缺的东部地区继续发展大规模数据中心已是困难重重。然而，西部地区具有丰富的可再生资源，这就为在西部发展数据中心、承接东部的算力需求提供了机会。

2022年2月，国家“东数西算”工程正式启动，初步在京津冀、长三角、粤港澳大湾区、川、蒙、黔、陇、宁建设8个国家算力枢纽节点和10个国家数据中心集群。“东数西算”中的“数”指的是数据，“算”指的是算力，也就是对数据的处理能力。这项重大工程被称为数字经济时代的“南水北调”“西电东送”“西气东输”。这是为了优化中国算力供给结构，对算力进行供给方面的改革。

利用西部便宜的电力能源、通信带宽、土地资源等，让算力的成本更低，从而更有效地支持中国数字经济的发展。同时，这也能拉动西部经济，通过枢纽节点的建设带动西部数字经济的发展，并对相关产业产生拉动作用，发展优质数字产业。

这项工程在未来的5年内，每年将会拉动西部地区超千亿元的产业，包括数据中心运维、数据中心建设、服务器、芯片、系统和数据库相关软件、通信设施等，也包括东部未来发展的数据流通产业、数据基础产业、数据处理人工智能产业等。



- 可再生能源丰富，气候适宜，数据中心绿色发展潜力较大
- 用户规划较大，应用需求强烈

“东数西算”工程是通过构建数据中心、云计算、大数据一体化的新型算力网络体系，于2022年正式全面启动。它将东部算力需求有序引导到西部，利用西部便宜的电力能源、通信带宽、土地资源等，让算力的成本更低，从而优化数据中心建设布局，更有效地支持中国数字经济的发展

作者介绍：刘天海，德国卡尔斯鲁厄大学理论计算机系博士。曾在达姆施达特大学计算机系，德国伯布林根IBM研究所，德国 aicas 公司工作。从事领域主要有复杂软件的测试和验证，嵌入式实时虚拟机，遥测数据处理，智能电网 / 物流，工业物联网，云计算等。