

# 大数据与高性能计算

陈文光

清华大学计算机科学与技术系 北京 100084

## 摘要

大数据和高性能计算都是计算机技术发展的产物。高性能计算主要采用模拟方法,被称作科学发现的第三范式;大数据主要从数据中总结规律,即使在对研究对象缺乏深入理解的情况下也可以发现一定的相关性,被称作科学发现的第四范式。从研究范式、主要应用类型以及计算机软硬件系统的角度对大数据与高性能计算的关系展开阐述。

## 关键词

大数据;高性能计算;并行与分布式计算;相关性

## *Big Data and High Performance Computing*

Chen Wenguang

Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China

### *Abstract*

Both big data and high performance computing (HPC) are based on the computer technologies. The main methodology of HPC is simulation, which is called the third paradigm of scientific discovery. Big data explore data for correlations even without much knowledge on the object of study, which is called the fourth paradigm of scientific discovery. Big data and HPC with several aspects were compared, such as the research paradigm, main application domain and underlying hardware/software systems.

### *Key words*

big data, high performance computing, parallel and distributed computing, correlation

## 1 引言

科学发现是改善人类生活最重要的基础,例如对动植物的驯化和耕作技术使人类能够获得稳定的食物来源,数学和力学的发展使人类可以构建房屋,抗生素的发现和双盲测试使得人类摆脱了传统医学,DNA的发现更标志着人类对自身的了解达到了新的高度。由重大科学发现引发或推动的现代科学技术发展,正以超出想象的速度改善着人类的生存条件,促进社会进步。例如,平均预期寿命是人类社会发展程度的一个标志性指标,如图1所示<sup>[1]</sup>。可以看出,20世纪初,全球平均预期寿命仅有约30岁;2010年,全球预期寿命已达到67.2岁,许多发达国家的平均预期寿命已超过80岁<sup>1</sup>。

科学发现对人类如此重要,那么是什么因素导致了这些科学发现,科学发现是否有“模式”?数千年来,人们逐渐总结出科学发现的若干范式:第一范式是实验方法,即通过实验验证假说是否成立的科学发现方法,实验方法的一个非常重要的地方是控制实验条件,以排除各种非实验因

子的干扰;第二范式是理论方法,即通过数学方法进行分析得出结论;随着计算机技术的出现,人们开始通过计算对复杂系统进行模拟,从而产生了科学发现的第三范式,对大规模科学与工程问题的模拟催生了高性能计算;大数据则提供了进一步的科学发现机会,Jim Gray将直接从数据中总结规律的方式,称作科学发现的第四范式<sup>[2]</sup>。

高性能计算和大数据都是计算机技术发展的产物,它们之间既有区别又存在紧密的联系,本文将从研究范式、主要应用类型以及计算机软硬件系统的角度对大数据与高性能计算的关系展开阐述。

## 2 高性能计算

高性能计算主要面向挑战性的科学与工程问题,例如飞行器设计、气象预报、全球气候变化模拟、核聚变模拟、新材料设计、药物设计以及人类基因组等<sup>[3]</sup>。

高性能计算主要采用数值模拟的方法,即科学发现的第三范式。以气象预报为例,首先气象科学家将气象预报问题抽象为地球表面、云和太阳等实体之间的物理过程和相互关系(如图2所示),然后将

1  
http://en.  
wikipedia.org/  
wiki/Life\_  
expectancy

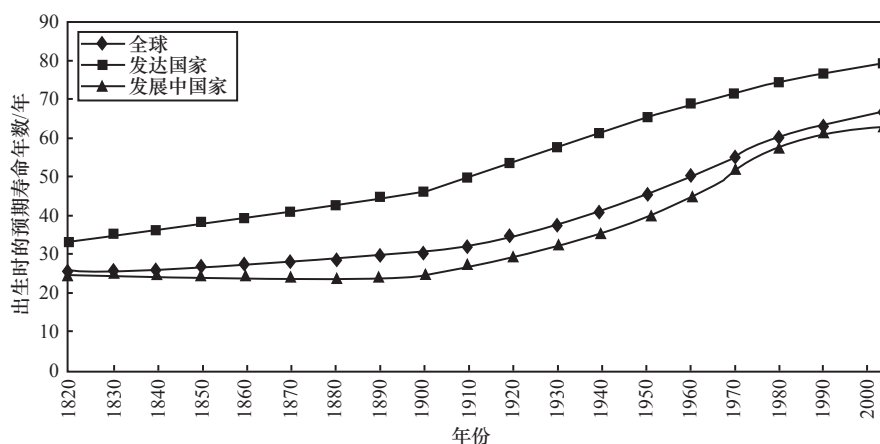


图1 1820-2003年的人类预期寿命

问题描述为流体力学和热力学方程组,利用各种气象观测数据取得初始条件(经过数据同化),再使用计算机用数值模拟的方法求解方程组,获得预报值。药物设计则是使用分子动力学方法对大量的药物分子进行筛选,计算候选药物对特定靶点的活性,筛选其中活性较高的药物再进入实验筛选。

可以看出,面向科学与工程的高性能计算是在对所求解的科学或工程问题已经有相当了解的情况下展开的,即已经能够为研究对象建立数学模型,并了解相应的物理、化学过程的原理。由于系统的复杂性,这些方程通常无法得到解析解,因此需要数值模拟方法求解方程组。

数值模拟方法的重要性显而易见,对很多无法进行实验的复杂问题,数值模拟方法提供了一种研究问题的新手段,例如天气预报和气候变化研究,科学家无法在地球上进行控制条件的实验来验证相关猜想。对新药筛选这类问题,虽然可以通过实验的方法得到候选药物对靶点的活性,但大量筛选药物的成本非常高,时间也非常长,数值模拟可以有效地降低药物筛选的成本,加快筛选速度。

解决重大挑战问题所需数值模拟的计算量和内存量通常非常大。以气候模拟为例,云层模型的网格粒度需要精细到1.5 km以下,模拟时间达到真实时间的1/1 000以下,才能满足气候模拟研究者对精度和速度的要求。要达到这一要求,需要200 PFLOPS(1 P =  $10^{15}$ , FLOPS为每秒浮点运算次数)的峰值性能和10 TB以上内存,这远远超出了现有单台计算机的性能和内存容量。

因此,聚合多台计算机能力的并行计算成为高性能计算的基本模式。相应地,高性能计算需要解决并行性引入的一系列问题:算法收敛性、执行的不确定性、负载

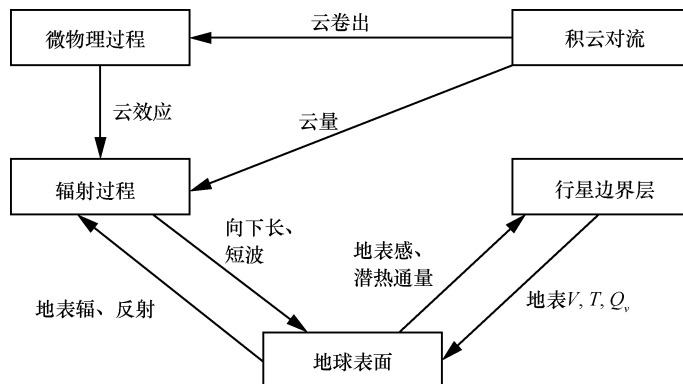


图2 中尺度天气预报GRAPES模式所描述的复杂物理过程及其相互作用关系

平衡、容错能力、功耗、编程性以及通信开销等。高性能计算机一般采用高质量服务器结点和高速专用网络,对系统的绝对性能有很高要求,有代表性的软件包括并行编程模型MPI和OpenMP、用于GPU加速器编程的CUDA和OpenCL、并行文件系统Lustre和PVFS等。

截至2015年4月,世界上最快的高性能计算机是中国的天河2号。该系统由国防科学技术大学研制,目前安装在广州超级计算中心。系统由16 000个结点组成,每个结点有2个通用处理器和3块加速卡,共有1.4 PB内存和12.4 PB外存系统,互连网络采用国防科学技术大学自主研制的高速专用网络TH Express-2,系统峰值速度达到了5.49 PFLOPS,整机功耗达到24 MW(包括冷却部分)。可以看出,即使是天河2号计算机这样的世界最大的超级计算机,其运算能力与高精度全球气候模拟的要求仍然存在很大差距。

### 3 大数据

大数据近年来得到了各行各业的广泛关注,其定义多种多样,其中Gartner的3V

定义是最为流行的定义之一，即大数据大在数据量 (volume)、产生速度 (velocity) 以及多样性 (variety)。对大数据定义的讨论已有较多，本文不再赘述。

当前大数据的应用主要以数据挖掘算法和机器学习算法 (如关联规则挖掘、LR 回归、决策树、神经网络等) 为主，广泛应用于政府、商业、金融等领域，并正在向工业、农业领域扩散。

本文主要讨论大数据的一个重要特征，即基于数据的规律发现，也就是科学发现的第四范式。需要说明的是，尽管大数据强调“大”，但基于数据的规律发现并不一定要求数据量非常大，而且这一方法甚至不一定需要计算机。例如，开普勒从第谷对行星的观测数据中总结出了开普勒三定律，为几十年后牛顿发现万有引力定律提供了基础，这是从数据中发现规律的经典案例。随着信息技术的发展，特别是近年来互联网和物联网的飞速发展，大量的数据被产生、收集、存储，亟待有效的分析方法从数据中挖掘有意义的规律，从数据中发现规律的可能性和重要性大大增加了，这也是大数据在近年出现并流行的重要原因。

这种从数据中挖掘规律的方法有两个明显特点。

(1) 观测到的规律可能是概率的而非决定性的。例如，通过对基因和乳腺癌之间的关系进行研究，从数据中发现“BRCA1基因突变的人发生乳腺癌的风险是55%~65%”，这是个非常有意义的规律，因为没有相关基因突变的人患乳腺癌的风险要低得多，实际上著名影星安吉丽娜·朱莉检测出自己的BRCA1基因突变后，通过手术切除了自己的乳房。但是，这一规律并非是决定性的，即使携带了BRCA1突变基因的人，仍然有40%左右的概率不会发病。

(2) 从数据中总结的规律很多时候仅

是相关性规律，而非因果性。例如，通过分析历史数据发现，全球平均气温与当时的二氧化碳浓度成正相关，即二氧化碳浓度高的时候，全球气温也高。但是仅凭气温与二氧化碳浓度的数据分析，不能得出二氧化碳导致气温升高的结论，因为还有一种可能性是气温升高导致二氧化碳浓度升高。如果没有进一步的研究工作说明二氧化碳导致温度升高的机理，仅能得出相关性，而不能得出因果性。当然，很多时候仅有相关性就够了，比如超市通过数据分析发现，购买电筒的人经常会购买蛋挞，虽然不知道其原因，但仍然可以将这两种商品放在一起，增加销售额。因此，关注相关性而非因果性是大数据的一个显著特征。然而，在具体使用大数据分析得出的相关性时，需要注意相关性成立的条件。例如，如果数据都是春天的数据，那么得到的相关性是否适用于秋天？

可以看出，基于大数据的第四范式与基于高性能计算的第三范式有明显区别。第三范式通常对研究对象已经有了深入理解，可以用方程组描述其行为，模拟仅是为了求解复杂方程组；第四范式则可能对研究对象不是非常了解，仅有一些观测数据，但希望从这些数据中能够发现相关性规律。

从大数据定义中的3V可以看出，单台计算机一般也是很难支撑大数据分析所需要的计算能力、内存容量和存储容量的，因此大数据分析平台也很自然地利用分布式系统进行并行计算，同样会遇到高性能计算也会遇到的算法收敛性、执行的不确定性、负载平衡、容错能力、功耗、编程性以及通信开销等问题。与高性能计算不同的是，大数据最初是由Google、Yahoo等互联网公司主要推动的，其硬件平台多由廉价的服务器通过普通以太网连接起来，而不像高性能计算机采用高质量的服务器和高速专用网络连接。因此与高性能计算相比，

大数据处理软件更注重系统的扩展性和容错性，对系统的绝对性能关注相对较少。其代表性的软件平台是Google的GFS和MapReduce/BigTable/Spanner等以及开源的Hadoop、Spark系统等。

## 4 大数据与高性能计算的相互借鉴与融合

表1从研究范式、应用领域、硬件平台、软件平台的角度比较了大数据与高性能计算。

尽管表1列出了大数据与高性能计算在多个方面的显著区别，但这两个领域也存在许多共同点，存在着相互借鉴与融合的趋势。

从研究范式来讲，科学发现往往是多范式的结合，大数据分析发现的相关性尽管不包含因果性，却为进一步发现因果性提供了基础。例如，在BRCA1基因突变可能导致乳腺癌后，进一步的研究就可以集中在BRCA1基因突变所导致的生物过程上，为研究乳腺癌的机理提供了更为明确的途径。另一方面，在Jim Gray关于第四范式的介绍中，认为第四范式实际上融合了实验、理论和模拟这前3个范式，即数据可以通过实验，也可以是通过模拟得到，大数据算法本身就会用到理论，特别是统

计学<sup>[2]</sup>。

从硬件平台上来看，现有普通服务器和网络在处理通信较为频繁的大数据问题时效率不高，而高性能计算机由于其高性能计算结点和高速专用网络，对这类问题的处理更为高效。例如，大数据平台仍然广泛使用吉比特网，只有少数使用万兆网，而高性能计算机早已开始使用带宽为40~56 Gbit/s的InfiniBand，并在广播、多播、规约等操作上提供了硬件优化，还提供了基于RDMA（远程直接内存访问）的快速通信机制。因此，大数据处理也开始借鉴高性能计算机的硬件平台技术。例如，Oracle推出了大数据一体机Exadata，其内部采用了高性能结点、高速专用网络InfiniBand和高速存储<sup>2</sup>。

在软件层面，许多大数据算法可以表达为稀疏矩阵运算，并通过GPU等加速器进行加速，而高性能计算在稀疏矩阵的CPU和GPU加速方面都有很好的软件库，可以用来加速大数据算法。许多研究者发现，基于MapReduce和Spark的大数据算法实现效率过低，某些情况下甚至不如经过良好优化的单机并行程序，采用高性能计算的思路优化大数据算法，也是一个重要的研究方向<sup>[4]</sup>。

另一方面，MapReduce、Spark等大数据编程系统所具有的良好容错性也为解决极大规模高性能计算的容错问题提供了新

2  
<https://www.oracle.com/engineered-systems/exadata/index.html>

表1 大数据与高性能计算的比较

	大数据	高性能计算
研究范式	从数据中总结规律（第四范式）	数值模拟（第三范式）
主要应用领域	政府、商业、金融	科学与工程
主要硬件平台	廉价的商用服务器和普通以太网，网络延迟100 μs，带宽100 Mbit/s~1 Gbit/s	高性能服务器结点和高速专用网络，网络延迟1 μs，带宽40~56 Gbit/s
主要软件平台	着重容错与扩展性Hadoop、Spark等	着重执行效率MPI、OpenMP、Lustre等

的思路。在天河2号这样的P级系统中,全系统的平均无故障时间一般不超过10 h,但传统高性能计算(MPI)编程模型的容错代价太大,常用的保存检查点方法通常会带来巨大的I/O量,不仅开销大,还影响了系统稳定性,提高了系统故障率。利用编程系统和算法进行更高效的容错,是高性能计算发展的重要方向。更进一步,利用大数据的方法分析高性能计算系统运行时产生的事件记录,可以有效预测系统中可能发生故障的部件,从而采取主动容错的方式,在故障还未发生时就采取措施,降低故障给程序运行带来的开销<sup>[5]</sup>。

## 5 结束语

大数据与高性能计算虽然起源于不同的研究范式,但都是利用计算进行规律发现和预测的方法,尽管它们在研究范式、应用领域、硬件平台、软件平台上有所区别,但面临类似的技术挑战,两个领域也存在相互借鉴、共同发展和融合的趋势。大数据和高性能计算的融合可望为人类提供更加强有力的科学发现工具,改善人类的生活,促进社会的发展。

## 参考文献

- [1] Goklany I M. The Improving State of the World: Why We're Living Longer, Healthier, More Comfortable Lives on a Cleaner Planet. Washington: Cato Institute, 2006
- [2] Hey H, Tansley S, Tolle K. The Fourth Paradigm: Data-Intensive Scientific Discovery. Washington: Microsoft Research, 2009
- [3] National Science Foundation Advisory Committee for Cyber infrastructure Task Force on Grand Challenges Final Report. [https://www.nsf.gov/cise/aci/taskforces/TaskForceReport\\_randChallenges.pdf](https://www.nsf.gov/cise/aci/taskforces/TaskForceReport_randChallenges.pdf), 2011
- [4] Kyrola A, Blelloch G, Guestrin C. GraphChi: large-scale graph computation on just a PC. Proceedings of the 10th USENIX Symposium on Operating Systems Design and Implementation, Hollywood, CA, USA, 2012
- [5] Xu W, Huang L, Fox A, *et al.* Detecting large-scale system problems by mining console logs. Proceedings of the 22nd ACM Symposium on Operating Systems Principles, Big Sky, Montana, USA, 2009

### 作者简介



陈文光,清华大学计算机科学与技术系教授,ACM中国理事会副主席,中国计算机学会杰出会员和杰出讲者、副秘书长、YOCSEF荣誉委员,《ACM China Magazine》主编,主要研究领域为操作系统、编译器与并行计算。

收稿日期: 2015-05-06; 修回日期: 2015-05-08

论文引用格式: 陈文光. 大数据与高性能计算. 大数据, 2015003

Chen W G. Big data and high performance computing. Big Data Research, 2015003