

面向海量数据处理的数据密集型计算发展研究

李亢, 刘东

(装备学院 复杂电子系统仿真重点实验室, 北京 101416)

摘要:随着信息技术的发展和普及,当今社会的信息量呈现出爆炸式增长的趋势,数据信息在无论是在量上还是在增长速率上均相当惊人,数据密集型计算应运而生。数据密集型计算是指获取、存储和管理大规模数据集,并基于这些数据集进行计算和分析。概述了数据密集型计算的概念和特点,研究了数据密集型计算的数据并行和数据本地性的两个基本原理及应用情况,分析了数据密集型计算最新技术的国内外研究现状与发展趋势。

关键词:数据密集型计算;海量数据;数据探索模式

本文引用格式:李亢,刘东. 面向海量数据处理的数据密集型计算发展研究[J]. 四川兵工学报,2015(7):93-96.

Citation format: LI Kang, LIU Dong. Development Research on Data-Intensive Computing Towards Massive Data Processing [J]. Journal of Sichuan Ordnance, 2015(7):93-96.

中图分类号: TP391

文献标识码: A

文章编号: 1006-0707(2015)07-0093-04

Development Research on Data-Intensive Computing Towards Massive Data Processing

LI Kang, LIU Dong

(Science and Technology on Complex Electronic System Simulation Laboratory,
Equipment Academy, Beijing 101416, China)

Abstract: With the development and popularization of information technology, the amount of information in today's society has shown explosive growth trend. It is quite amazing in the data amount either or both in the increase in the rate, and then the data-intensive computing emerged. Data-intensive computing is a data center, and the system is responsible for obtaining, maintaining continuous change data sets, and while large-scale computing and processing on these data sets. This paper showed a brief overview of the concepts and features data-intensive computing and introduced the two basic principles of data parallelism, data locality and applications of data-intensive computing, and focused on the development status of data-intensive computing and the prospect of the development trend of its technology.

Key words: data-intensive computing; massive data; development research

随着信息技术的发展和普及,当今现代电子战争中的信息量呈现出爆炸式增长的趋势^[1]。海量数据的处理问题引起了世界各国军事领域专家的广泛关注^[2-6]。2014年,互联网数据中心发布的研究报告(Digital universe of opportunities)中指出,到2020年全球的信息总量将从现有的4.4亿GB增长到44亿GB。数据是信息化战争中的重要战略资源信息,谁能够掌握了信息优势谁就夺取了未来战争的至高点。在

信息化战争中,制胜的关键不仅取决于装备的硬件和软件水平,还包括海量数据处理的计算能力。

面对海量数据的高速增长, Tony Hey 等在其新书《The Fourth Paradigm》^[7]中,从实验实证、理论分析和计算模拟3种研究模式基础上诞生了第4种研究模式——数据探索模式(data exploration),即数据密集型科学发现。越来越多的领域专家或科学家不再根据已知的规则编写程序解决问题,

收稿日期: 2015-02-16

基金项目: 国家自然科学基金资助项目(60904082)

作者简介: 李亢(1987—),男,博士研究生,主要从事大数据处理、云计算研究。

而是去分析海量数据,从浩如烟海的数据中寻找问题的答案,获得理论上的突破^[8-9]。数据密集型计算也就是在这个应用背景下提出来的。

1 数据密集型计算概念和特点

在有关数据密集型应用的学术讨论中,《自然》和《科学》杂志分别于2008年和2011年出版了《大数据(Big Data)》和《数据处理(Dealing with Data)》专刊,从互联网技术、互联网经济学、超级计算、环境科学、生物医药等多个方面讨论海量数据处理和应用专题^[8-9]。在2007年,卡耐基·梅隆大学等相继提出了“数据密集型超级计算(data intensive supercomputing)”和“数据密集型可扩展计算(data intensive scalable computing)”的概念。数据密集型计算是指获取、存储和管理大规模数据集,并基于这些数据集进行计算和分析。它通过网络构建而成的大规模计算机系统,能够有效实现数据的并行计算和分布式计算,更加注重用户对数据快速存储和读取、交互式编程模型的高效便捷访问以及灵活可靠性等方面的需求。

数据密集型计算研究的目标是解决大规模数据集和数据流的计算问题,它具有如下几个特点^[10]:

- 1) 通过 MapReduce 等分布式编程框架,支持对大型计算机系统的分布式计算,简化开发应用过程中的分布式设计,能够使集群具有很好的可扩展性。
- 2) 能够有效支持多种类型的计算任务,可以利用大规模分布式集群对大规模数据进行遍历分析,提取出有价值的信息。
- 3) 通过据复制、自动选择计算资源等容错机制提高集群的容错能力,可以自动地对多种故障进行容错,降低了大规模集群开发和维护的成本。

2 数据密集型计算的基本原理

数据密集型计算普遍应用了数据并行性原理和数据本地性原理。

2.1 数据并行性原理

数据并行性原理^[11]的基本思想是将数据切分成多个片段,同时对每个片段执行相同的操作。数据并行性原理在于计算机领域中有非常广泛的运用。例如 Intel 的 SSE(StreamingSIMD Extensions,单指令多数据流扩展)指令集和 GPU(Graphics Processing Unit,图形处理器)分别在指令级和线程级利用了数据并行性。下面结合一个简单的实例阐述数据并行原理对海量数据处理的重要意义。

假设磁盘上存储的数据为 1TB,磁盘读取带宽为 100 MB/s。从 1 个磁盘上读取,需要花费 10 000 s 才能读取完 1TB 的数据集。然而,如果将 1TB 的数据集平均分布到互联网的 100 个磁盘上,每个磁盘只需要存储 10GB 数据,在这种条件下同时驱动 100 个磁盘通过并行计算的方式读取数据只

需要 100 s,也就是说读取数据的效率能够提高 100 倍。

2.2 数据本地性原理

在数据密集型计算中充分利用数据本地性原理^[12]可以大幅提高系统的吞吐量。数据本地性原理可以简单地概括为:尽量将任务放置在离输入数据最近的节点运行。数据中心的网络带宽是稀缺资源,远远低于计算节点本地磁盘的带宽总和。因此,应尽可能地把计算任务调度到数据所在的计算节点上运行,这样就可以避免消耗网络带宽从而直接处理本地磁盘中的数据。其次,机架内两个节点之间的带宽远大于跨机架的两个节点之间带宽,因此将任务调度到同一个机架内进行运算是次优的选择。总而言之,在调度任务时应当“移动代价很小的计算而不是移动代价昂贵的数据”^[13],从而避免有限的网络带宽处于饱和状态。

3 数据密集型计算的应用

数据密集型计算使得以前无法实现和想象的问题变成了现实,深刻地改变了人类的生产生活方式。如何将面对海量的数据转换成有用的信息、知识、规律或趋势,是数据密集型计算要面临的重要挑战。目前,在基因工程、天文计算、商业计算等众多领域,数据密集型计算已经得到了较为广泛的应用。

3.1 基因工程

随着基因测序技术的迅猛发展,推动了当前以数据为驱动的生物信息学研究所引导的新一代生命科学技术变革。新的基因序列不断被发现导致基因数据库也变得越来越大,美国国家生物技术创新中心(NCBI)维护的核苷酸序列信息的 GenBank 数据库,其数据量以每 10 个月两倍的速度递增^[14]。到 2006 年 8 月为止,GenBank 数据库已经存储了超过 100 000 个不同器官的 650 亿个核苷酸序列信息,通过传统计算的方式进行处理和计算如此巨大的数据集是难以实现的^[15]。许多生物企业和科研机构结合数据密集型计算的优势开展研究,例如华大基因研究中心王丙强等人对高通量测序所产生的大规模数据集的维护和处理,利用集群计算、云计算、网格计算和异构计算等平台进行了深入分析^[16]。

3.2 天文计算

在天文研究领域,需要对实时的存储和处理由天文望远镜等传感器源源不断地获取图像数据。例如大口径全景巡视望远镜 LSST 项目每年将产生图片和目录数据达 PB 级。美国射电望远镜阵列系统平方公里阵列 SKA 大约每秒产生 200 GB 的数据量,并且需要以千万亿次/每秒的计算速度来处理这些数据以获得宇宙的射线图谱。目前世界各国天文台广泛研究的虚拟天文台^[17]即是通过数据密集型计算的方法将全球的天文信息资源集合到一起,在此基础上开展合作研究。

3.3 商业计算

Google 公司每天要接收近 30 亿次的搜索请求,百度公司在每次搜索过程中要处理 3 000 亿左右的中文网页,数据

量达到 10 ~ 50 PB^[18]。著名的社交网站 Facebook 在 2011 年 10 ~ 12 月期间,平均每天有超过 2.5 亿张的照片上传,照片所占用的空间超过 100PB。以 Amazon 为代表的电子商务公司提供的系统,如 EC2、S3 等,已成为数据密集型相关研究的基础平台^[19]。Yahoo、阿里巴巴、腾讯等互联网巨头也涉及数据密集型研究。由于商业计算的用户规模非常庞大,因此这需要处理海量、异构、多变的数据,需要具有通用的技术,具有很大的技术挑战性。

此外,数据密集型计算对其他行业也产生了深远的影响,在互联网技术、超级计算、环境科学、生物医药、经济学等多个领域都已经展开了海量数据处理和应用的研究专题,数据密集型计算得到了越来越广泛的应用^[20-21]。

4 数据密集型计算的发展现状

信息化时代的人类社会数据规模呈现出爆炸式的增长态势是,既给人们的人工作和生活带来了便利,同时也为信息技术的发展带来了巨大挑战。为了应对海量数据处理的挑战,Google、微软、雅虎等公司根据自身的业务特点,研发了一系列数据密集型计算的核心技术,极大地推进了数据密集型计算的发展,包括 IBM 和 Google 联合开展的“蓝云(Blue Cloud)计划”、Google 提供的 App Engine 服务,以及微软推出的典型“软件+服务”的平台 Live Mesh 和操作系统 Azure 等。美国的 Carnegie Mellon 大学开展的对数据密集型的超级计算方面的研究,从科学和理论上推动了数据密集型计算相关技术的发展^[22]。

Google 设计并实现了面向大规模数据处理的 MapReduce 编程模型,满足了 Google 迅速增长的数据处理需求,能够同时让内部员工在大规模集群上开发各种数据密集型应用。目前 MapReduce 每天处理的数据量超过 20PB,广泛应用于 Web 网页索引、文本处理和数据挖掘等领域。

微软硅谷研究院开发了 DryadLINQ、Dryad 和 SCOPE 等产品,提供了一个可供分布式并行计算的平台。DryadLINQ 是一种高级程序语言的编程接口,它能够使普通程序员通过利用这个接口方便地对大规模数据集进行分布式计算,该平台涉及了微软公司研发的 Dryad 技术和 LINQ 技术。Dryad 与 MapReduce 类似,它不仅仅是一种编程模型,同时能够对完成对计算任务的调度。Dryad 并不局限于在云计算上的应用,它同样可以应用于由多核和多处理器以及异构机群组成的数据密集型计算平台,并且有比较好的效果。

雅虎公司研发了 Pig Latin 数据流语言,运行在 Pig 系统上。Pig 是介于高层次的、声明式的 SQL 语言与低层次、过程式的 MapReduce 编程之间的折中解决方案。Pig 系统将 Pig Latin 程序(由一系列的操作或转换组成)编译成一组 Hadoop 作业。Pig 提供了常见的数据处理操作,包括加载 load、存储 store、过滤 filter、分组 groupby、排序 sort 和连接 join 等。Pig 提供了丰富的数据模型,支持原子类型、字典、元组和 bag(集合)等数据结构,并支持嵌套。

Facebook 公司在开源分布式处理平台 Hadoop 基础上开

发了数据仓库系统 Hive,提供了类 SQL 的检索方式处理存储在 HDFS (Hadoop Distributed File System) 中的海量数据。Hive 数据管理系统按照表、分区和桶的方式组织数据。每个表对应 HDFS 的一个目录,表的数据以序列化的方式存储在 HDFS 目录下。

除 Google 等 IT 公司,科研学术界也进行了深入研究^[23],并提出了许多相关的技术。针对数据密集型应用中数据规模庞大、难以管理的问题,学术界提出了 CStore、MonetDB 等基于列存储的分析型数据库管理系统,和一定的缓存与索引技术相结合,CStore、MonetDB 可以在不影响其他查询效率的前提下,使得分析型查询的效率提高 10 ~ 100 倍。另外,学术界分别基于开源的 Hadoop 技术和数据库管理系统的技术,从两个方向研究数据密集型应用中的大规模数据处理问题,代表性的成果有 HadoopDB、Amazon S3 的数据库系统、Greenplum、Aster Data 等。

5 结束语

目前,数据密集型计算已经成为深受国内外各行各业关注的研究热点。在国内外针对数据密集应用的数据处理技术研究中,基于 MapReduce 编程模型进行再次开发利用成为最主要的研究思路。MapReduce 编程模型首先由 Google 为其搜索相关业务提供支持而开发,目前已经成为大多数进行海量数据处理的单位或科研机构采用的计算基础,这些领域涉及数据库管理系统、机器学习、科学数据分析、在线聚合和分布式主动处理等。

自从以 Hadoop 为代表的数据库密集型计算平台广为流行之后,许多军队科研单位也着手在 Hadoop 基础上开展研究,不断地推动着数据密集型计算在军事领域的应用发展,极大促进了数据密集型计算在信息化战争中的推广和运用。

参考文献:

- [1] Winter R. Why Are Data Warehouses Growing So Fast[EB/OL]. [2008-08-23]. <http://www.b-eye-network.com/view/7188>.
- [2] Jeffrey Dean, Sanjay Ghemawat. MapReduce: Simplified data processing on large clusters. In Proceedings of the USENIX Symposium on Operating Systems Design and Implementation[Z]. 2004.
- [3] Michael Isard, Mihai Budiu, Yuan Yu, et al. Dryad: Distributed data-parallel programs from sequential building blocks [C]// In Proceedings of the ACM SIGOPS/EuroSys European Conference on Computer Systems, [S. l.]: [s. n.], 2007.
- [4] Ronnie Chaiken, Bob Jenkins, Per-Ake Larson, et al. SCOPE: Easy and efficient parallel processing of massive data sets[J]. Proceedings of the VLDB Endowment (PVLDB), 2008, 1(2): 1265-1276.

- [5] Alan Gates, Olga Natkovich, Shubham Chopra, et al. Building a highlevel dataflow system on top of MapReduce: The Pig experience[J]. Proceedings of the Very Large Database Endowment, 2009, 2(2): 1414 - 1425.
- [6] Ashish Thusoo, Joydeep Sen Sarma, Namit Jain, et al. Hive-A petabyte scale data warehouse using hadoop[C]//In Proceedings of the IEEE International Conference on Data Engineering. [S. l.]: [s. n.], 2010.
- [7] Tong hey. The Fourth Paradigm: Data-Intensive Scientific Discovery[M]. Redmond, Washington.
- [8] 覃雄派, 王会举, 杜小勇, 等. 大数据分析——RDBMS 与 MapReduce 的竞争与共生[J]. 软件学报, 2012, 23(1): 32? 45.
- [9] R. E. Bryant. Data-intensive supercomputing: The case for DISC[Z]. Carnegie Mellon, 2007.
- [10] 周松. 面向数据密集型超级计算的基于纠错码的容错存储技术研究[D]. 长沙: 国防科学技术大学, 2010.
- [11] W. Daniel Hillis, Jr. Guy L. Steele. Data parallel algorithms [J]. Communications of the ACM, 1986, 29(12): 1170 - 1183.
- [12] Jeffrey Dean, Sanjay Ghemawat. MapReduce: Simplified data processing on large clusters [C]//In Proceedings of the USENIX Symposium on Operating Systems Design and Implementation[S. l.]: [s. n.], 2004.
- [13] 王鹏. 海量数据处理编程模型的研究[D]. 北京: 中国科学院计算技术研究所, 2011.
- [14] Clifford Lynch. How do your data grow? [J]. Nature, 2008, 455(9): 28 - 29.
- [15] RANDAL E. Bryant. Data-Intensive supercomputing: the case for DISC[Z]. CMU Technical Report, 2007.
- [16] 王丙强. 应对生命科学的大数据挑战[J]. 中国计算机学会通讯, 2012, 8(6): 29 - 33.
- [17] 赵永恒, 崔辰州. 从虚拟天文台到天文信息学[J]. 科研信息化技术与应用, 2011, 2(3): 3 - 12.
- [18] Matthias Brantner, Daniela Florescu, David A. Graf, et al. Building a database on S3 [C]//SIGMOD Conference [S. l.]: [s. n.], 2008: 251 - 264.
- [19] LSST. Large Synoptic Survey Telescope[EB/OL]. [2012 - 08 - 12]. <http://www.lsst.org/lsst/faq#q1>. 2012.
- [20] Ludaseher B, Altintas I, Berkley C, et al. Scientific workflow management and the Kepler system [J]. Concurrency and Computation: Practice and Experience, 2005, 18(10): 1039 - 1065.
- [21] Oinn T, Addis M, Ferris J, et al. Taverna: A tool for the composition and enactment of bioinformatics workflows[J]. Bioinformatics, 2004, 20(17): 3045 - 3054.
- [22] Nurmi D, Wolski R, Grzegorzczak C, et al. Eucalyptus: A Technical Report on an Elastic Utility Computing Architecture Linking Your Programs to Useful Systems[Z]. Technical Report 2008 - 10, UCSB Computer Science, 2008.
- [23] Navin S, David I, Prashant S. Multisense: Fine-grained multiplexing for steerable sensor networks[Z]. Technical Report UM-CS-2009-033, University of Massachusetts at Amherst, April 2009.

(责任编辑 杨继森)

(上接第 44 页)

- [2] 吴正龙, 赵忠实. 基于自适应模糊神经网络系统的射击毁伤评估模型研究[J]. 兵工学报, 2012(11): 1352 - 1357.
- [3] 吴大梁. 射击效果的评定及最优化方法[M]. 南京炮兵学院出版社, 1986.
- [4] 刘怡昕. 炮兵射击学[M]. 北京: 海军出版社, 2007.
- [5] 江华, 许梅生. 炮兵射击基本理论分析[M]. 北京: 兵器工业出版社, 2004.
- [6] 程云门. 评定射击效率原理[M]. 北京: 解放军出版社, 1986.
- [7] 聂鹏, 曹兵. 破片式战斗部对空中飞机的毁伤建模与仿真[J]. 指挥控制与仿真, 2012(6): 80 - 83.
- [8] 曾前腾, 吴慧中. 基于 C-4ISR 的空炸射击高炮拦阻射击效力仿真研究[J]. 兵工学报, 2007(10): 1164 - 1168.
- [9] 单显明, 肖俊岭. 炮兵雷达组网校射评估系统的设计与实现[J]. 指挥控制与仿真, 2010(4): 94 - 96.
- [10] 王亮宽, 舒立鹏, 程广伟. 配用 AHEAD 弹高炮武器系统射击效率模型[J]. 火炮发射与控制学报, 2010(2): 5 - 8.

(责任编辑 周江川)