

大数据平台的基础能力和性能测试

姜春宇^{1,2}, 魏凯^{1,2}

1. 中国信息通信研究院移动互联网与大数据部, 北京 100191
2. 数据中心联盟大数据发展促进委员会, 北京 100045

摘要

目前整个大数据技术还处于以开源方式为主导、多种技术并存的阶段。开源技术催生了大量的商业发行版大数据平台软件, 大数据企业级市场竞争加剧, 如何测试和评估这些大数据平台软件成为新的研究主题。简要地介绍了大数据技术发展的背景以及大数据技术标准的需求, 综述了国际大数据平台标准化和评测的现状, 详细介绍了数据中心联盟在大数据平台技术标准化和测评方面的实践, 最后总结了当前工作的问题, 并展望了下一步大数据技术和评测的发展方向。

关键词

大数据; 大数据技术标准化; 大数据产品评测; 数据; 负载

中图分类号: TP311

文献标识码: A

doi: 10.11959/j.issn.2096-0271.2017040

Basic capability and performance test of big data platform

JIANG Chunyu, WEI Kai

1. Department of Mobile Internet and Big Data, China Academy of Information and Communications Technology, Beijing 100191, China
2. Council for the Promotion of Big Data Development, Beijing 100045, China

Abstract

The whole big data technology is now led by open source society which results in coexist of many competing technologies. Open sources also help to cultivate a great number of big data commercial software. The enterprise market is now crowded by various providers. How to evaluate these softwares becomes a new research topic. At the beginning, the development of big data system was briefly reviewed. Then the requirement of big data technology standardization was illustrated. After reviewing the progress of international big data technology standardization, the standardization and test practices in big data products under the organization of Data Center Alliance was introduced. Finally, the drawbacks of current practices were discussed, and the future direction of standardization and test for big data products was summarized.

Key words

big data, big data technology standardization, big data products evaluation, data, workload

1 引言

大数据的应用和技术起源于互联网,首先是网站和网页的爆发式增长,搜索引擎公司最早感受到了海量数据带来的技术上的挑战,随后兴起的社交网络、视频网站、移动互联网的浪潮加剧了这一挑战。互联网企业发现新数据的增长量、多样性和对处理时效的要求是传统数据库、商业智能纵向扩展架构无法应对的。在此背景下,谷歌公司率先于2004年提出一套分布式数据处理的技术体系,即分布式文件系统——谷歌文件系统(Google file system, GFS)、分布式计算系统(MapReduce)和分布式数据库(BigTable),以较低成本很好地解决了大数据面临的困境,奠定了大数据技术的基础。受谷歌公司论文启发,Apache Hadoop实现了自己的分布式文件系统——Hadoop分布式文件系统(Hadoop distribute file system, HDFS)、分布式计算系统(MapReduce)和分布式数据库(HBase),并将其进行开源,这是大数据技术开源生态体系的起点。

经过10年左右的发展,大数据技术形成了以开源为主导、多种技术和架构并存的特点。从数据在信息系统中的生命周期看,大数据技术生态主要有5个发展方向,包括数据采集与传输、数据存储与管理、计算处理、查询与分析、可视化展现。在数据采集与传输领域渐渐形成了Sqoop、Flume、Kafka等一系列开源技术,兼顾离线和实时数据的采集和传输。在存储层,HDFS已经成为了大数据磁盘存储的事实标准,针对关系型以外的数据模型,开源社区形成了K-V(key-value)、列式、文档、图4类NoSQL数据库体系,

HBase、Cassandra、MongoDB、Neo4j、Redis等数据库百花齐放。计算处理引擎方面慢慢覆盖了离线批量计算、实时计算、流计算等场景,诞生了MapReduce、Spark、Flink、Storm等计算框架。在数据查询和分析领域形成了丰富的SQL on Hadoop的解决方案,Hive、HAWQ、Impala、Presto、Drill等技术与传统的大规模并行处理(massively parallel processor, MPP)数据库竞争激烈。在数据可视化领域,敏捷商业智能(business intelligence, BI)分析工具Tableau、QlikView通过简单的拖拽来实现数据的复杂展示,是目前最受欢迎的可视化展现方式,开源社区也形成了一批图表工具,例如Chart.js、Chated、D3、Raw等。

大数据技术通过开源发展壮大,相比闭源软件,开源技术带来了更多的灵活性、自由度,而且通过所谓的“创新民主化”来孕育创新,与此同时开源也给大数据技术带来诸多挑战。首先,开源技术并不等于成熟的产品,不够稳定,易用性差,缺少服务支持。其次,开源的技术门槛较高,需要大量的技术人员实现技术的定制化。最后,开源技术种类太多,技术选型比较困难。互联网企业由于其业务需求的驱动以及丰富的技术人员储备,能够从容应对开源技术的挑战,自如使用乃至贡献开源技术,形成良性循环。而传统企业由于自身技术力量储备不足,业务与技术组织架构的耦合度低,无法及时将业务需求映射到技术需求上,并依赖于第三方IT公司的帮助来实现开源技术的产品化。

在Hadoop发行版方面,国际上主要包括3个独立的专业供应商——Cloudera Inc.、Hortonworks Inc.、MapR Technologies Inc.。国内自2013年以来,以银行、电信行业为代表的企业开始逐步采用Hadoop,早期的采购以实验性质为主,

涌现出大量做Hadoop发行版的企业,市场持续繁荣的同时也变得混乱,产品供应商和用户之间存在巨大的信息鸿沟。对于用户来说,初期的大数据需求比较模糊,产品的技术性语言难以和业务需求相对应,面对众多良莠不齐的发行版厂商,技术和产品选型困难,产品评估的过程成本很高。对于大数据产品供应商,同类的产品很多,缺少技术性标准来规范市场,往往陷入价格战的泥潭中,此外重复参加各个用户组织的POC (proof of concept) 测试,测试花费太多。

大数据技术的标准化旨在通过共性的评估体系将技术与应用场景相对应,将复杂的产品转化为容易理解的指标,为行业设立技术门槛。对于产品供应商,大数据技术标准通过客观的指标体系保证厂商之间的有序竞争;对技术难点集中进行攻关,专注于性能的改进;定义合作的用户场景,引导产品的研发方向。对于产品用户,大数据技术标准帮助用户更好地判断大数据产品的优劣,经济有效地进行大数据产品的选型。对于公众来说,大数据技术标准是一种沟通的工具,其指标易于理解,测量的方法公开公平。

2 大数据平台技术评测现状

目前大数据相关标准的制定处于起步阶段,国际上主要有ISO/IEC、国际电信联盟(International Telecommunications Union, ITU)、美国国家标准与技术研究院(National Institute of Standards and Technology, NIST)、事务处理性能委员会(Transaction Processing Performance Council, TPC)、标准性能评估机构(Standard Performance Evaluation Corporation, SPEC)等组织进行大数据标准化工作。其中ISO/IEC中

JTC1 WG9主要开发大数据基础性标准,包括参考架构和术语;JTC1 SC32 “数据管理和交换”分技术委员会负责制定信息系统内、信息系统之间的数据管理与交换标准。NIST的大数据公共工作组(NBD-PWG)^①负责制定大数据的定义、术语、安全参考体系和技术路线图,提出大数据分析、存储、管理和基础设施的参考架构。TPC和SPEC两个组织关注大数据技术平台的基准测试,TPC是数据库基准测试领域主导者,早在2012年TPC成立了大数据基准测试社区(BDBC),开始研究制定大数据场景下的基准测试标准。它在2014年8月发布了针对Hadoop系统的TPCx-HS标准^②,以TeraSort为负载,从性能、性价比和能耗等指标衡量底层硬件的能力,目前为止有美国思科公司、华为技术有限公司、美国戴尔公司3家公司的服务器参与了测试,并发布了测试结果。2016年TPC发布了新的针对大数据平台的性能基准测试标准TPCx-BB^③,TPCx-BB测试通过模拟零售商的30个应用场景,执行30个查询算法来衡量基于Hadoop大数据系统的服务器软硬件性能。SPEC于2014年成立了大数据的研究组(Research Group Big Data)^④,在大数据基准测试领域陆续举办了十多次的见面会和分享,最后与TPC进行合作,将TPCx-BB引入SPEC作为其测试的标准。

3 大数据平台性能标准化实践

中国信息通信研究院于2014年6月开始大数据技术标准化的尝试,最早以Hadoop平台为起点,在数据中心联盟联合20多家企业,历时5个月制定了《大数据平台基准测试 第一部分:技术要求》《大数据平台基准测试 第二部分:测试方

① <https://bigdatawg.nist.gov/>

② <http://www.tpc.org/tpcx-hs/default.asp>

③ <http://www.tpc.org/tpcx-bb/default.asp>

④ <https://research.spec.org/working-groups/big-data-working-group.html>

法》两个标准，分别梳理出了大数据平台（Hadoop平台）基准测试的方法论和具体实施方案。基准测试主要从性能的角度衡量大数据平台，主要考虑数据生成、负载选择和明确测试指标等内容。根据Hadoop平台的特点，从NoSQL任务、机器学习、SQL任务、批处理四大类任务中选择了10个测试用例作为负载，如图1所示，明确了每个测试负载的数据规模和数据生成方式，这是大数据技术标准化的第一阶段。

2015年10月到2016年6月，在数据中心联盟大数据技术与产品工作组内，中国信息通信研究院和20多家企业经过20多次线上线下的讨论，形成了《Hadoop平台基础能力测试方法》《Hadoop平台性能测试方法》两套技术标准。性能测试在基准测试的基础上进行了升级，如图2所示，增加了SQL用例的比重，按照SQL任务的类型，从两个维度确立了5类任务，分别是

CPU密集型任务、I/O密集型任务、报表任务、分析型任务、交互式查询。HBase的优势在于并发检索和读的性能，在负载方面选择了批量写入数据、并发读任务占多、并发更新任务较重、读取更改然后写回4类场景，充分检验了HBase数据库的并发执行能力。在机器学习和批处理方面相对基准测试方法减少了2个用例。

4 大数据平台基础能力标准化实践

大数据平台（Hadoop平台）基础能力主要考虑：Hadoop平台的基本功能及数据的导入导出对SQL任务、NoSQL任务、机器学习、批处理任务的支持；Hadoop产品是否能够通过界面的形式方便用户进行非运行维护，主要包括集群的安装、监控、配置、操作等；Hadoop产品是否能够提供基

NoSQL任务	机器学习	SQL任务	批处理
数据并发导入	K-means无监督聚类	join	Terasort MR I/O密集
并发读	贝叶斯有监督分类	聚合	WordCount
并发扫描	PageRank		

图 1 基准测试用例分布

SQL任务	NoSQL任务	机器学习	批处理
I/O密集型任务	数据并发导入	K-means无监督聚类	Terasort MR I/O密集
CPU密集型	95%的读，5%的写	贝叶斯有监督分类	
报表任务	50%的读和50%的写		
分析型任务	读—更改—写		
交互式查询			

图 2 Hadoop 平台性能测试用例分布

本的安全方案,包括是否具备认证功能以防止恶意访问和攻击、是否能够进行细粒度的权限管理、是否能够提供审计和数据加密功能等;Hadoop产品是否具备高可用的机制,防止机器的失效带来的任务失败以及数据丢失;是否能够支持机器快速平滑地扩展和缩容以及扩展是否带来线性的计算能力;是否能够支持多个调用接口以及对SQL语法的支持情况;是否能够根据队列、用户的权重来细粒度地分配计算资源。基于这7方面的考虑,Hadoop平台基础能力涵盖了运维管理、可用性、功能、兼容性、安全、多租户和扩展性等指标,总共38项测试用例,如图3所示。

除Hadoop平台的标准化外,针对MPP数据库(面向在线分析处理(on-line analytical processing, OLAP))的标准化也提上日程,经过4个多月的研讨,完成了《MPP数据库基础能力测试方法》,通过六大指标、48个测试用例来衡量MPP数据库的功能完备性。经过两年多的标准化历程,针对底层的大数据基础平台,数据中心

联盟形成了基础能力和性能两套测试标准体系,覆盖了Hadoop平台和MPP数据库两类产品,未来标准化工作将聚焦于数据治理工具、分析和可视化工具。

5 大数据产品评测实践

大数据技术的标准化过程中,测评认证是标准化的落地形式,只有通过评测的方式才能全方位对比不同的大数据产品,发现当前产品的问题,优化性能,最终帮助用户选购合适的产品。2016年有5家厂商的Hadoop产品参与了性能测试,包括新华三集团、腾讯、星环信息科技有限公司、北京东方金信科技有限公司和百分点集团。测试集群由32台x86服务器构成,配置见表1。

数据规模方面,SQL任务规定生成10 TB的TPC-DS数据,TPC-DS由17个维度表、7个事实表构成。在99个查询语句中,选出了5条最有代表性的语句,这5条语句

运维管理	可用性	功能	兼容性	安全	多租户	扩展性
自动化部署	NameNode主节点失效恢复	数据导入	ODBC兼容性	认证	租房管理	集群动态扩展
资源监控	NameNode备节点失效恢复	SQL任务能力	JDBC兼容性	授权	资源管理	集群动态收缩
作业监控	DateNode节点失效恢复	NoSQL数据库	SQL支持度	加密	资源隔离	
集群操作	HMaster节点失效恢复	机器学习	传统数据库同步	审计	权限管理	
故障管理	RegopmServer节点失效恢复	流处理能力	跨不同数据库表关联操作			
日志管理	HDFS备份恢复					
配置管理	HBase备份恢复					
权限管理	运维管理节点失效及恢复					
用户管理						
无宕机升级						

图3 Hadoop平台基础能力测试用例分布

表 1 32 台集群配置

组件	配置	台数/台
CPU	2×英特尔至强E5-2 620 v3 2.4 GHz, 15 MB缓存	32
内存	4×16 GB RDIMM, 2 133 MT/s	31
	8×16 GB RDIMM, 2 133 MT/s	1
硬盘	10×1.2 TB 10K RPM SAS 6 Gbps 2.5英寸 热插拔硬盘	22
	10×1.2 TB 10K RPM SAS 12 Gbps 2.5英寸 热插拔硬盘	10
网卡	单口万兆网卡	32
交换机	锐捷RG-S6 220-48XS4QXS万兆交换机	1

涉及13个表, 其中表Store_sales的记录数超过了280亿行, 表Caltalog_sales的记录数超过143亿条。HBase的操作数规模与之前保持一致, 是20亿个操作, 10个客户端, 每个客户端2亿个操作, 每个操作执行大概1 KB大小的数据, HBase最后一个负载因为要执行读取、更改、写入等操作, 实际每个客户端执行3亿个操作。批处理选择Terasort作为负载, 数据规模是29.1 TB, 机器学习采用中国科学院计算技术研究所的BigDataBench来生成数据, 具体数据规模见表2。

图4展示了5家厂商在5个SQL任务上测试的最优、中位和标准方差, 以执行时间来衡量, 时间越短越好。10 TB级别的SQL任务大致执行时间在几百秒以内, 其中第二个SQL (Query47) 的执行时间最长,

最优和中位的差别很大, 标准方差值也较高, 说明在该项负载上厂商的差别较大, 用例本身耗时也较长。Query71标准方差相对较高, 说明该用例上执行的差别较大, 还有一定的调优空间。其他任务执行情况较为接近。

HBase性能由YCSB (Yahoo! cloud serving benchmark) 工具来测试, 分别模拟了数据并发导入、95%读5%写、50%读50%写和读一更改一写等场景。测试的指标是吞吐量, 即单位时间内执行的操作数, 操作数越大越好。如图5所示, 总体上前3个用例的最优吞吐率都在百万级别, 最后一个负载由于需要更多的操作数, 所以显得更慢。从标准方差可以看出来, 在HBase任务上各厂商的差别普遍比较大,

表 2 测试数据规模

	负载	数据量	数据描述
SQL任务	I/O密集型任务	10 TB	由7个事实表和17个维度表构成, 本次5个语句涉及13个表, 其中表Store_sales的记录数超过280亿条, 表Caltalog_sales记录数超过143亿条
	CPU密集型	TPC-DS数据	
	报表任务		
	分析型任务		
	交互式查询		
HBase任务	数据并发导入	20亿条操作数	每个客房执行2亿个操作, 10个客房端一共执行20亿个操作, 每个操作执行大概1 KB大小的数据
	95%的读, 5%的写	20亿条操作数	
	50%的读和50%的写	20亿条操作数	
	读一更改一写	20亿条	
MR任务	TeraSort	29.1 TB	实际读操作数为20亿条, 更改和写回10亿条操作 利用TeraGen生成29.1 TB数据
机器学习	K-means	255.3 GB	利用Facebook社交数据进行扩展 (BigDataBench)
	贝叶斯	248.3 GB	利用亚马逊电影评论数据扩展生成 (BigDataBench)
		99.3 GB	利用亚马逊电影评论数据扩展生成 (BigDataBench)

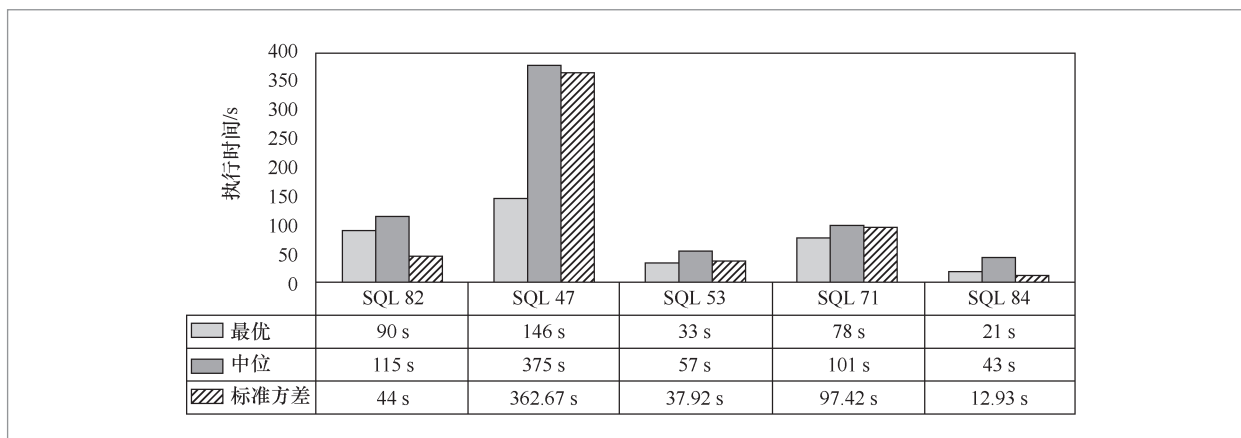


图4 SQL任务测试结果

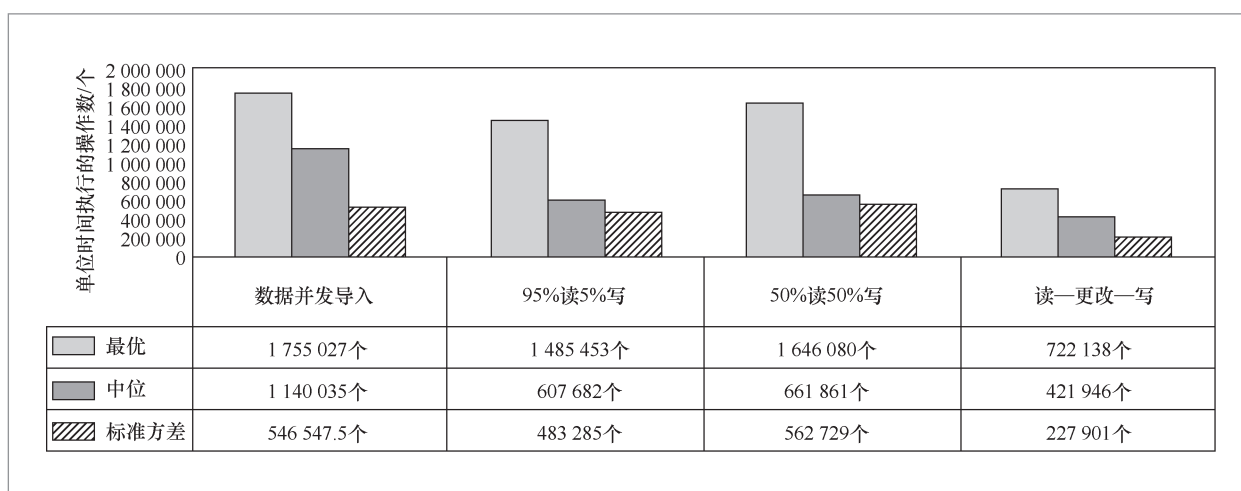


图5 HBase性能测试结果

说明参测厂商之间在产品本身和调优能力方面都有一定的差距。

2015年基准测试是用15个计算节点来对13.6 TB的数据进行排序,见表3,最优执行时间是8 374 s。2016年的性能测试是采用32个计算节点对29.1 TB的数据进行排序,最优执行时间是10 083 s,节点数扩了一倍多,数据规模也扩张一倍多,时间只

稍微多了一些。15个计算节点的最优吞吐量是1 703 MB/s, 32台计算节点最优吞吐量是3 026 MB/s,从图6来看,并没有呈现一个完全线性的扩展,或者是一个线性可比性。有两种解释,一种是本身Terasort或者是Hadoop数据扩展就不会出现线性的扩展。第二种可能是32节点的最优情况不如15节点更优。Terasort最优执行时间

表3 Terasort 2015年和2016年执行情况对比

年份	类型	负载	节点/台	数据量/TB	最优执行时间/s	中位执行时间/s	最优吞吐率/(MB·s ⁻¹)	最优单机吞吐率/(MB·s ⁻¹)
2016年	批处理	Terasort	32	29.1	10 083	11 985	3 026	94.5
2015年			15	13.6	8 374	—	1 703	113.5

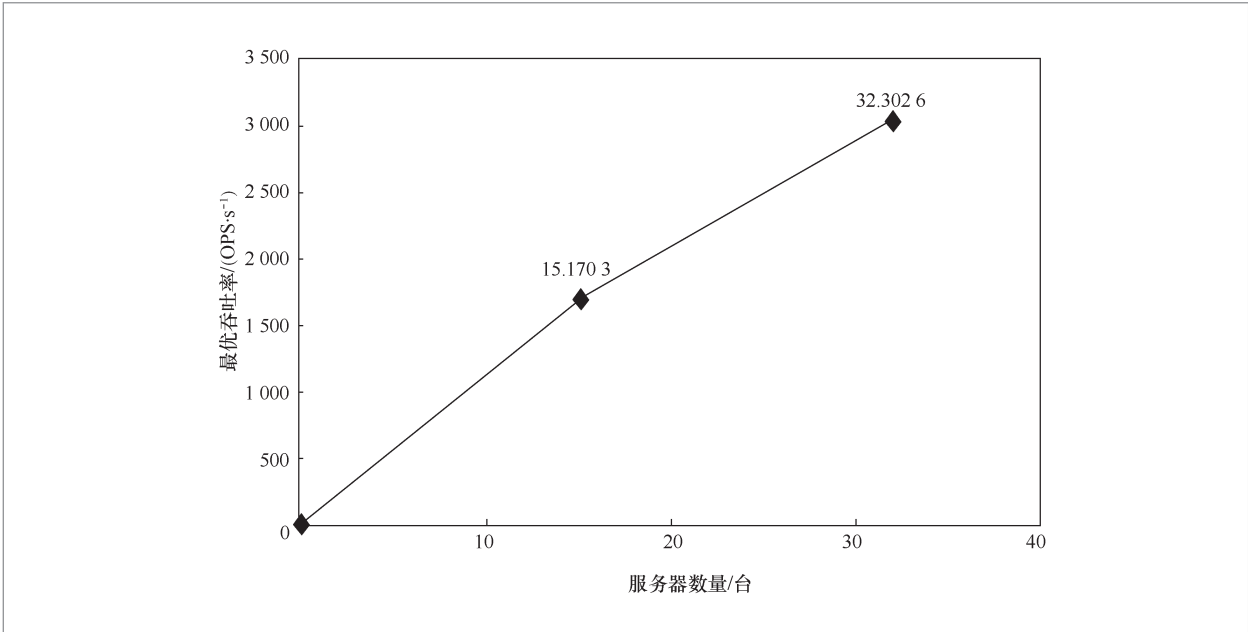


图 6 TeraSort 2015 年与 2016 年执行情况

是10 083 s左右,大概是2 h 48 min,中位值是3 h 19 min,厂商之间的差距并不大。

性能的测评实践进一步考察了产品本身的功能完备性和性能,尤其是性能评测,采用业界领先的测试集群规模和配置,考察产品的本身性能、易部署性、稳定和易运维性,同时也考核了参测团队综合使用大数据平台的能力,包括环境部署与集群规划、测试工具熟悉程度、多任务调优能力、时间进度安排能力、集群故障处理、运行维护等能力。

6 结束语

本文首先阐述了大数据技术发展的特点,即开源引领,百花齐放。开源技术促进了大数据产业的繁荣,同时也给传统企业级市场带来了一些挑战,一是开源软件本身的不稳定,与成熟产品有一定差距;二是大数据开源技术种类繁多,选择成本较高;三是传统企业直接使用大数据

开源技术的技术成本较高;四是大数据开源技术降低了IT产品供应商研发的门槛,市场上发行版产品鱼龙混杂。大数据技术的标准化和评测旨在建立通用的评价质检体系,为行业设立技术门槛,帮助用户更高效地挑选优质产品。国际和国内都有部分行业标准化和评测组织在从事大数据技术标准化工作,主要集中在大数据平台的评测领域,本文介绍了数据中心联盟在大数据平台(Hadoop平台)技术的标准化与测评历程,经过3年多的实践,数据中心联盟形成了国内领先的大数据产品标准化和测评体系,从基础能力和性能两方面来评估大数据平台类产品的能力。目前已经完成了4批评测,共有24家厂商的25个大数据产品通过评测,为国内大数据技术的发展保驾护航。

目前标准和测评还存在一些局限性,一是性能目前是单项极限值测试,缺少对混合负载的考量。二是现场测试人员的前期准备和现场发挥对性能测试结果有一定的影响。三是测评主要采用开源的测试工

具,覆盖的实际场景较少。四是由于时间限制,无法进一步对稳定性做出考量。

未来基础能力会引入工作流和可视化测试用例,详细考察数据的导入导出,统一考虑权限管理,引入数据管理功能,增加流计算和图计算。性能以任务为导向,寻找

行业数据,考虑混合任务、高并发等场景。

目前正在进行的第四批测试除Hadoop平台外,还对MPP数据库基础能力进行了评测,已经在2017年3月27日发布。除大数据平台类产品外,还计划覆盖BI工具、数据挖掘平台和其他大数据技术产品。

作者简介



姜春宇 (1987-), 男, 中国信息通信研究院移动互联网与大数据部工程师, 数据中心联盟大数据技术与产品工作组组长。主要研究方向为大数据技术体系, 从事大数据技术和产品的标准化工作, 牵头建立了国内首个大数据产品能力评测体系。参与起草《促进国家大数据发展行动纲要》, 参与编写工业和信息化部《大数据产业“十三五”发展规划》。



魏凯 (1981-), 男, 中国信息通信研究院移动互联网与大数据部高级工程师、副主任, 数据中心联盟大数据发展促进委员会常务副主任委员。主要研究方向为云计算和大数据相关技术、标准和政策, 参与《促进大数据发展行动纲要》《大数据产业“十三五”发展规划》等文件的研究起草。

收稿日期: 2017-03-22