

面向智能电网的云数据处理系统评价方法

刘 越^{1,2}, 王晶华³, 虎嵩林¹

(1 中国科学院 计算技术研究所, 北京 100190; 2 中国科学院 研究生院, 北京 100190;

3 国网信息通信有限公司, 北京 100761)

摘 要: 智能电网数据处理系统需要存储从巨量设备端点周期性采集到的海量数据, 处理高度并发的读写请求, 并要求系统具有良好的可扩展性. 新兴的云数据服务系统为我们提供了很好的选择. 但是现在云数据服务系统种类繁多, 各方面的性能存在差异, 如何选择合适的云数据服务系统成为关键问题. 基于智能用电场景, 提出了一种对云数据服务系统的性能评价和分析方法, 并利用该方法评测 HBase、Voldemort 和 Cassandra 三者的性能.

关键词: 智能用电; 云数据服务系统; 评测; 性能

中图分类号: TP303

文献标识码: A

文章编号: 1000-7180(2011)08-0035-04

Performance Evaluation and Analysis Method of Cloud-based Data Processing System for Smart Grid

LIU Yue^{1,2}, WANG Jing-hua³, HU Song-lin¹

(1 Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190, China;

2 Graduate University of Chinese Academy of Sciences, Beijing 100190, China;

3 State Grid Information & Telecommunication Co. Ltd., Beijing 100761, China)

Abstract: Smart Grid data processing system needs to store the large-scale data collected from the mass device end-point periodically, handle the high-concurrent read and write requests, and the high-scalability is more important. Cloud-based data serving systems provide a good choice for us. However the large variety of cloud-based data serving systems and the various performance make it difficult for developers to choose an appropriate system. Here we propose a method to benchmark the performance of cloud-based data serving systems based on the special scene of smart grid. Finally we benchmark the performance of HBase, Voldemort and Cassandra through our method.

Key words: Smart Grid; Cloud-based data serving systems; benchmark; performance

1 引言

随着云计算技术的进步和发展, 出现了很多云数据服务系统, 每种云数据服务系统都是根据特殊的应用场景开发出来的, 它们在性能优化、副本策略和一致性等方面的决策不同, 这导致了它们在各方面性能上存在差异. 除此之外, 每种云数据服务系统公布的性能数据都是在最合适的工作负载下测试得到的最好性能数据, 参考价值有限.

为了选择合适的云数据服务系统, 为智能用电

提供稳定、高效的存储服务, 我们必须研究符合智能电网数据处理场景的方法, 在真实用电数据负载和访问规律下进行测试, 从而分析评价各种云数据服务系统的性能.

由雅虎研究院开发的 YCSB 系统^[1]主要关注一定系统规模下, 在某种访问分布规律下云数据服务系统的读写性能, 而在智能用电场景中, 除了读写性能, 我们还关注一些系统方面的特性, 如内存占用率、数据压缩比和数据负载均衡等. Chris Bunch 等^[2]人利用 Google App Engine 提供的 API 对不同

收稿日期: 2011-05-13; 修回日期: 2011-06-26

基金项目: 国网信息通信有限公司科技项目(SGIT[2010]449)

的云数据处理系统进行评测,主要关注读写性能,没有考虑访问分布和系统方面的特性。

2 云数据服务系统概览

云数据服务系统在海量数据存取上具备关系型数据库无法比拟的性能优势。由于每种云数据服务系统在性能优化、CAP(在分布式环境中,一致性、可用性和分区容忍性只能同时实现两者)选择和数据负载等方面的决策不同,造成了它们各方面的性能存在差异。所以我们在测试的时候应该选取一些在决策方面具有代表性的系统。

在智能用电场景中,对写性能要求比较高,周期性采集的海量数据必须在有限的时间内写入,否则会被下一批数据覆盖。不仅如此,智能用电对系统的可用性也有较高的要求,数据服务系统不能出现拒绝服务的现象,同时,数据需保证最终一致性。除此之外,智能用电要求数据自动负载均衡,无需人为控制,这样可以平衡访问压力、减少热点问题。结合以上的理论依据和需求,选取了三种具有代表性的云数据服务系统,利用我们的性能评价和分析方法进行评测:

HBase, HBase 是 Apache 的 Hadoop 项目的子项目,采用了 Google BigTable^[3] 的思想。HBase 对写性能进行了优化,相对读性能较差;CAP 中选择了 CP,提供了较好的一致性保证;采用了随机的放置策略。

Voldemort, Voldemort 由 LinkedIn 开发,采用了 Amazon Dynamo^[4] 的思想。Voldemort 没有像 HBase 那样对写性能进行优化,这样其读性能相对较好。CAP 中选择了 AP,提供了很好的可用性;采用了一致性哈希的放置策略。

Cassandra, Cassandra 由 Facebook 开发,采用了 Google BigTable 和 Amazon Dynamo 的思想,很多方面兼有 HBase 和 Voldemort 的特点。Cassandra 对写性能进行了优化;CAP 中选择了 AP,不过在 Cassandra 中每次的访问请求都可以指定一致性级别,这相当于 CA 是可调节的;采用了一致性哈希放置策略。

3 性能测试指标体系与 SUDE 系统

3.1 智能用电数据操作的特征

与面向互联网的应用相比,智能电网的数据操作具有明显的特征。数据操作总体上可以分为以下四类。(1)周期性数据采集:周期性地从用电终端采

集电能、电量等数据信息,并将数据写入到数据服务系统,属于大并发的写操作。该操作一般频率较高,通常 15 分钟之内完成一次管理域内所有终端的用电采集和写入操作。(2)数据统计分析:每小时或者每日进行三相用电分析等统计和分析数据处理,既有较大规模的读操作,也可能包含统计结果的写操作。(3)用户交互查询:指用户交互查询用电信息,为随机性较强的读操作。用户通常会针对一个家庭进行多时间点的查询,以满足用电曲线绘制等需求。(4)数据压缩备份:按照电网业务要求,对历史电能、电量信息进行压缩和备份,体现为密集的、批量的读操作。

3.2 性能测试指标体系

为了描述云数据服务系统的性能特征,需要定义一系列性能指标。由于智能用电数据处理系统需要及时存储周期性采集的海量数据,要求系统具有较高吞吐量和较小的读写延迟。所以在实验中我们更为关注读写性能,包括读写延迟、读写吞吐量以及改写延迟三个因素。为支持服务系统内存、外存的配置策略设计,也需要对内存占用情况、数据存储的压缩比、系统的负载均衡等情况进行合理评估。在实验中,我们将调整数据处理规模以及系统本身的规模,以便评价在不同变量条件下的性能效果。(1)数据处理规模:变化单位时间总的数据处理量以反映一个云数据服务系统可能支撑的智能用电设备终端的规模;(2)系统规模:变化服务系统的节点数量以反映服务系统在节点规模扩展情况下的性能提升效果。

3.3 SUDE(Smart utility data service evaluator) 测试系统

为了对云数据服务系统进行全面测试,我们开发了 SUDE 测试系统,系统结构如图 1 所示,图中操作模拟器根据操作配置进行读写操作模拟;采集数据写进程读取智能用电数据作为测试集;测试执行器根据定时器的频率执行操作模拟器中的操作,并通过数据读写代理对具体的云数据服务系统

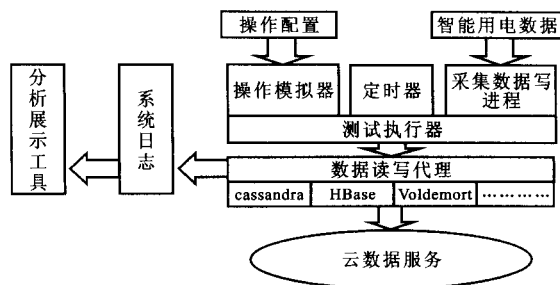


图 1 SUDE 测试系统结构图

进行操作;操作生成的日志文件和统计信息记录在系统日志中,并通过分析展示工具展示。

4 实验结果与分析

4.1 独立读写测试

这里用 3 节点集群和 19 节点集群进行测试,向三种云数据服务系统中分别写入读取 50 万、100 万、300 万、500 万、800 万和 1 000 万条数据,记录平均读写延迟和读写吞吐量。实验结果表明,随着集群规模的增加,三者的读写性能都得到了有一定的提升,其中 Voldemort 的性能提升最大。HBase 的写性能很好,但是会随着数据量的增加略有下降,读性能在数据量较小的时候比较好;Voldemort 的读性能很好,并且读写性能比较稳定,随着集群规模的增加,读写性能显著提升;Cassandra 在集群规模较小的时候,写性能很好,但是随着集群规模的增加写性能下降,Cassandra 的读性能比较差。

4.2 混合读写模拟测试

这里用 3 节点集群和 19 节点集群进行测试,该测试模拟实际应用场景,这里模拟数据采集(write 进程)、数据统计分析(analysis 进程)、用户查询(user 进程)和数据备份(backup 进程)四种操作,测试三种云数据服务系统在读、写、随机读取等方面的性能。为了尽可能反映在读写高峰时的情况,并减少实验的运行时间,我们对模拟处理的时间间隔进行了压缩。这里 write 进程每隔一定的时间写入 100 万的数据量,analysis 进程每隔一定的时间随机读取已写入多个版本中的一个版本的全部数据进行统计分析,user 进程每隔一定的时间发起一组随机的查询请求,读取随机的 Key 对应的全部 Value 值;backup 进程每隔一定的时间备份一次数据,读取从上次备份的版本到现在最新版本之前的全部数据,以上各个进程都采用 5 个线程。结果如图 2 和图 3 所示。

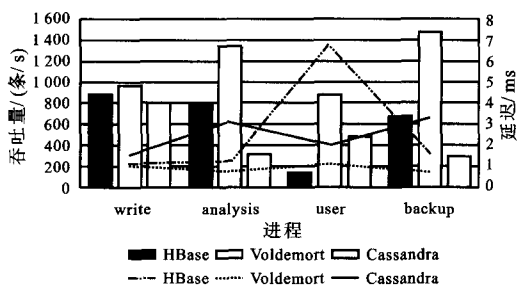


图 2 3 个节点的吞吐量和延迟

从图中我们发现,无论集群规模大小,Voldemort 的各个模拟进程的性能都好于其他两者。同时,我们发现 HBase 的 user 进程延迟较大,这说明

HBase 在多线程随机读取方面的能力比较差。

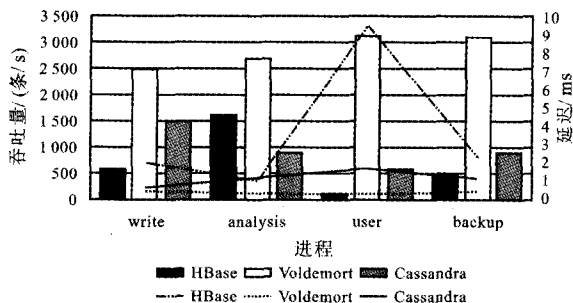


图 3 19 个节点的吞吐量和延迟

4.3 改写测试

这里采用 19 个节点的集群,先向三种云数据服务系统中写入 500 万条数据,然后每次修改 10% 的数据,修改两次,记录每种云数据服务系统的改写延迟。三种云数据服务系统的改写延迟如表 1 所示。

表 1 改写延迟 ms

系统	HBase	Voldemort	Cassandra
延迟	0.15	0.46	0.33

在 HBase 和 Cassandra 中,改写操作即为覆盖写操作,本质即为写操作,用时很短。而在 Voldemort 中,没有内嵌的改写操作,需要先将数据读出,修改,再写回,所以用时较长。

4.4 数据压缩比测试

采用 19 个节点的集群,向三种云数据服务系统中分别写入 100 万、300 万和 500 万的数据,统计各个节点上的数据量,然后求出每个副本数据大小与原数据大小之比。实验结果表明,当数据量较小的时候,写入相同数据量的数据,Cassandra 占用的磁盘空间最小,而当写入数据量较大的时候,Voldemort 占用的磁盘空间较小。

4.5 内存占用测试

这里采用 19 个节点的集群,向三种云数据服务系统写入 500 万条数据,每 30 s 记录云数据服务系统的内存占用情况,对比三种云数据服务系统的内存占用差别。结果如图 4 所示。

从图中我们发现,写入相同数据量的数据时,HBase 和 Voldemort 占用的内存较小,而 Cassandra 占用的内存较大,并且随着程序的运行,内存占用率升高很快。

4.6 数据负载均衡测试

采用 19 个节点的集群,向三种云数据服务系统中分别写入 100 万和 500 万的数据,然后统计每个节点上数据量的大小,发现 Voldemort 的数据负载

很均衡, HBase 和 Cassandra 的数据负载不均衡.



图 4 内存占用

5 结束语

综合以上的理论分析和测试结果, HBase 的单独写性能很好, Voldemort 综合性能最好, 而且非常稳定, Cassandra^[5] 的性能随数据量的变化波动比较大. 最终, Voldemort 以其稳定的性能表现使之更适宜于智能用电的数据处理需求.

为了使 Voldemort 能更加稳定高效的为智能用电提供存储服务, 今后我们会对它的源代码进行改进, 或者增加一些中间件来进一步提高它的吞吐量、稳定性等方面的性能, 使其更加适用于智能用电应用场景.

参考文献:

[1] Brian F Cooper, Adam Silberstein, Erwin Tam, et al.

Benchmarking cloud serving systems with YCSB[C]// SoC-C'10. Indianapolis, Indiana, USA, 2010.

[2] Chris Bunch, Navraj Chohan, Chandra Krintz, et al. An evaluation of distributed datastores using the appscale cloud platform. cloud[C]// IEEE 3rd International Conference on Cloud Computing. USA: Florida, Miami, 2010.

[3] Fay Chang, Jeffrey Dean, Sanjay Ghemawat, et al. Bigtable: a distributed storage system for structured data[J]. ACM Trans. Comput. Syst. 2008, 26(2):26.

[4] Giuseppe DeCandia, Deniz Hastorun, Madan Janpani, et al. Amazon's highly available key-value store[C]// SOSPP'07. Stevenson, Washington, USA, 2007.

[5] Avinash Lakshman, Prashant Malik. Cassandra - a decentralized structured storage system [C]// ACM SIGOPS Operating Systems Review archive. ACM New York, NY, USA, 2010, 44(2):35-40.

作者简介:

刘 越 男, (1988-), 硕士. 研究方向为大规模分布式系统.

王晶华 女, (1962-). 研究方向为智能电网中的数据处理技术.

虎嵩林 男, (1973-), 博士, 副研究员. 研究方向为服务计算、分布式事件系统、中间件等.

(上接第 34 页)

[2] Borkar S. Thousand core chips - A technology perspective[C]// Proc ACM/IEEE 44th Design Automation Conf. San Diego, California, USA: IEEE, 2007: 746-749.

[3] Sun X H, Chen Y. Reevaluating amdahl's law in the multicore era[J]. Parallel Distrib Comput, 2010, 70(2):183-188.

[4] Gustafson J L. Reevaluating amdahl's law[J]. Communications of the ACM, 1988(31):532-533.

[5] Li J, Martinez J F. Power-performance implications of thread-level parallelism on chip multiprocessors[J]. Performance Analysis of Systems and Software, 2005

(3):124-134.

[6] Suleman M A, Mutlu O, Qureshi M K, et al. An asymmetric multi-core architecture for accelerating Critical Sections[J]. Micro, IEEE, 2010, 30(1):60-70.

作者简介:

冯 叶 女, (1986-), 工学硕士. 研究方向为高性能计算及其应用.

邓倩妮 女, (1973-), 博士, 副教授, 硕士生导师. 研究方向为高性能计算及其应用、数据中心网络.