## **《大数据架构与技术》研读讨论作业**

## **Main Ideas and Major Contributions of Readings Discussion 2 (in Chinese)**

**1 **引言（Introduction）****

随着架构、系统和数据管理社区越来越关注创新型大数据系统和架构，对评估这些系统的压力也随之增加。大数据系统的广泛应用，对其性能进行评估的需求日益增长。然而，大数据系统的复杂性、多样性、工作负载的频繁变化和快速演进，给大数据基准测试带来了巨大挑战:

首先，许多大数据应用程序类没有进行全面的表征。其次，大多数大数据应用程序都基于复杂的系统软件堆栈构建，例如广泛使用的 Hadoop 系统。第三，即使一些大数据应用程序在业务和技术方面已经成熟，因为大多数互联网服务提供商将数据、应用程序和网页访问日志视为商业机密，客户、供应商或学术界或甚至不同行业领域的研究人员对彼此了解也不够。

为了公平起见，大数据基准测试必须包括多样化和代表性的数据集，这是评估大数据系统和架构的先决条件。但现有的基准测试套件往往只针对特定类型的应用或系统软件堆栈，无法涵盖多样化和真实世界的数据集。因此，本文提出了一个由工业界合作伙伴共同研发的大数据基准测试套件 BigDataBench。

### 2 ****核心思想（Core Ideas）****

本文提出了 BigDataBench，一个由工业界合作伙伴共同研发的大数据基准测试套件。该套件旨在公平地衡量和评估大数据系统和架构，涵盖了广泛的场景，并包含多样化和代表性的数据集，以满足不同应用领域的需求。

从应用场景、操作/算法、数据类型、数据来源、软件堆栈和应用类型等维度选择了 19 个大数据基准测试并使用不同数据输入全面描述了 BigDataBench 中的 19 个大数据工作负载，它们能够全面地衡量和评估大数据系统和架构。最后在一款典型的商用处理器 Intel Xeon E5645 上测试并得到积极的观察结果。

### 3 ****主要过程（Key Methodology）****

****3.1**衡量和比较大数据系统和架构**

**3.1.1衡量和比较大数据系统和架构：**

大数据基准测试的目的是根据用户关注点来衡量、评估和比较大数据系统和架构。考虑到大数据系统的广泛用例，为了公平起见，大数据基准测试套件候选者不仅必须涵盖广泛的应用场景，还必须涵盖多样化且具有代表性的真实数据集。

**3.1.2以数据为中心**

基准测试应考虑大数据的“4V”特性：大量、多样性、速度和真实性。

**3.1.3多样化的代表性工作负载**

基准测试应反映不同应用场景的多样性，并包括不同类型的工作负载，以便系统和架构研究人员能够获得大数据的全面工作负载特征。为系统设计和优化提供了有用的指导。

**3.1.4涵盖代表性的软件堆栈**

基准测试应涵盖不同应用场景下广泛使用和最先进的软件堆栈，针对特定用户关注点开发创新的软件堆栈。

**3.1.5先进的技术**

大数据系统的快速发展为新兴技术带来了巨大的机会，大数据基准测试套件候选者应跟上底层系统的改进。因此，基准测试应包含不同领域的新兴技术，并易于扩展以适应未来的变化。

**3.1.6可用性**

大数据系统在应用场景、数据集、工作负载和软件堆栈方面的复杂性使普通用户无法轻松使用大数据基准，因此其可用性非常重要。要求基准测试应易于部署、配置和运行，并且性能数据应易于获取。

**3.2 现有的基准测试方法**

大多数最先进的大数据基准测试工作的重点是评估特定类型的应用程序或系统软件堆栈，因此无法测量广泛用于广泛应用场景的大数据系统和架构。现有的一些评估特定类型应用程序或系统软件堆栈的基准测试：如HiBench、YCSB、LinkBench、BigBench、CloudSuite、DCBench。这些基准测试无法全面衡量大数据系统和架构。

**3.3本文的基准测试方法和决策**

**3.3.1方法论**

从真实系统出发，采用增量迭代的方法，本文调查了互联网服务的主要应用领域，根据广泛接受的指标页面浏览量和每日访问者的数量，关注互联网服务的主要应用领域：搜索引擎、社交网络和电子商务。

从数据类型和数据源的角度考虑数据多样性，关注结构化、半结构化和非结构化数据。此外，本文挑选了互联网服务主要应用领域中的三个重要数据源，包括文本数据（在搜索引擎中执行分析和查询量最大）、图形数据（社交网络中的最大量）和表数据（电子商务中的最大量）。以及其他重要的数据源将不断添加。本文还提出了满足数据量、种类、速度和准确性要求的新型数据生成工具。

为了覆盖多样化和具有代表性的工作负载，本文从用户的角度将大数据应用程序分为三种类型：在线服务、离线分析和实时分析。在线服务对延迟非常敏感，对于每个请求，都会执行相对简单的操作以立即向最终用户提供响应。对于离线分析，对延迟较长的大数据执行复杂的计算。而对于实时分析，最终用户希望以交互方式获取分析结果。我们同样关注三种应用程序类型。本文还从两个维度选择典型工作负载：典型应用场景中的代表性操作和算法，分别针对三种应用类型广泛使用和最先进的软件堆栈。

**3.3.2数据集**

数据集应该是多样化的，并且在数据类型和来源方面都具有代表性。在调查了三个应用领域之后，本文收集六种具有代表性的真实世界数据集，涵盖结构化、半结构化和非结构化数据，以及文本、图形和表格数据，并在在保持其准确性的同时扩展数据集的数量。

**3.3.3工作负载**

本文在选择 BigDataBench 工作负载时考虑了以下因素：

1. 对不同类型的应用程序给予同等关注：在线服务、实时分析和离线分析;
2. 覆盖多样化且具有代表性的应用场景的工作负载;
3. 包括不同的数据源：文本、图形和表格数据;
4. 覆盖代表性的大数据软件栈。选择19种大数据基准测试，涵盖不同的应用场景、操作/算法、数据类型、数据源、软件堆栈和应用类型。

**3.4合成数据生成方法和工具**

针对如何获取大数据这一与大数据对标的重要问题，一个自然想法是生成合成数据，同时保留真实数据的重要特征。本文的解决方案是通过利用基于特定应用程序的几个具有代表性的数据集，使用真实世界的数据估计数据模型的参数，在保持数据特征的同时，根据从真实世界数据中获得的数据模型和参数生成合成真实世界数据集。

本文开发了一个全面的工具Big Data Generator Suite（简称 BDGS），用于生成具有“4V”特性的合成大数据。数据生成器专为各种应用程序域（搜索引擎、estart 和社交网络）而设计，并将扩展到其他应用程序域。通过基于六个真实数据集开发数据生成器来证明其有效性，数据集涵盖三种代表性数据类型（结构化、半结构化和非结构化数据）、三种数据源（文本、图形和表格）。

**3.5工作负载特征化实验**

本文选择Hadoop作为基础软件堆栈，部署MPICH2和Rubis来了解不同的工作负载，在具有匹配规模（14节点）的集群上进行实验。每个节点都有两个 Xeon E5645 处理器，配备 16 GB 内存和 8 TB 磁盘。测试台上启用了超线程。操作系统为 Centos 5.5 和 Linux 内核 2.6.34。对于输入数据，将分析工作负载的大小从 32GB 到 1TB 不等。由于具有较大的数据复杂性，因此 BFS、CC 和 CF 工作负载等图形相关工作负载的输入数据以顶点集来衡量，而 Index 和 PageRank 工作负载的输入数据则以网页来衡量。并将请求数量从每秒 100 个请求更改为每秒 3200 个请求。

**3.5.1实验方法和指标**

使用Perf工具收集微架构级性能计数器，并访问proc文件系统收集操作系统级性能数据。评估使用两类指标。第一类指标是用户感知的指标，可以方便地被用户观察和理解。第二个是架构指标，主要由架构研究利用。在第一类指标中，本文分别为不同的工作负载选择三个测量单位。选择第二个类别是为了比较不同工作负载下的性能。尽管用户可感知的指标可以帮助评估同一类别中的不同工作负载，但无法比较不同类别的工作负载（例如，数据库服务器和 MapReduce 工作负载）的性能。因此，统一的架构指标是必要的。由于实验不涉及异构平台，因此我们在架构研究中选择了广泛接受的性能指标，例如 MIPS 和缓存 MPKI。即使用用户可感知的度量（例如RPS、OPS和DPS）和架构度量（例如MIPS和缓存MPKI）进行评估。

**3.5.2数据量对架构研究的影响**

1.数据量对微架构事件有显著影响： 对于不同的工作负载，随着数据量的增加，指令执行行为和缓存行为表现出不同的趋势。例如，Grep 和 WordCount 的 MIPS 数量在数据量增加到基线的 16 倍之后增加，而其他一些工作负载在数据量增加到一定程度后趋于稳定。所以，大数据工作负载的特征化需要考虑数据量，不同的数据量会导致不同的微架构事件。

2.数据量对性能影响显著： 对于 Sort 工作负载，当内存无法容纳所有输入数据时，它会变成 I/O 密集型工作负载，导致性能随数据量增加而下降。这表明数据量对性能有显著影响，在进行架构研究时需要考虑这一点。所以，大数据架构研究需要使用更大的数据集，传统的架构研究方法可能不足以反映大数据工作负载的真实性能特征，不能准确地评估其性能。

3.大数据工作负载需要更大的数据集： 传统的架构研究通常使用小型数据集进行模拟，但对于大数据工作负载来说，这种方法可能不足以反映其真实的性能特征。因此，大数据架构研究需要使用更大的数据集进行模拟，才能更准确地评估其性能。所以，大数据架构设计需要考虑缓存效率，大数据工作负载对缓存效率有更高的要求。

**3.5.3工作负载特征描述**

本文重点介绍大数据工作负载的操作强度和缓存行为特征方面进行分析和描述：

（1）.操作强度低： 与传统基准测试相比，BigDataBench 中的大数据工作负载的操作强度（计算与内存访问的比率）较低。这表明大数据工作负载更依赖于数据移动而不是指令执行，对浮点单元的需求较低。

（2）.L1 指令缓存未命中率 (MPKI) 高： BigDataBench 中的大数据工作负载的 L1 指令缓存 MPKI 比传统基准测试高得多。这可能是由于大数据工作负载具有巨大的代码量和复杂的软件堆栈导致的。

（3）.L2 缓存 MPKI 高： BigDataBench 中的大数据工作负载的 L2 缓存 MPKI 也比传统基准测试高。这表明大数据工作负载对 L2 缓存的需求较高，需要进一步提高 L2 缓存的效率。

（4）.L3 缓存对大数据工作负载有效： BigDataBench 中的大数据工作负载的 L3 缓存 MPKI 较低，表明 L3 缓存对大数据工作负载是有效的。

（5）.ITLB 和 DTLB MPKI 高： BigDataBench 中的大数据工作负载的 ITLB 和 DTLB MPKI 也比传统基准测试高。这表明大数据工作负载对 TLB 的需求较高，需要进一步提高 TLB 的效率。

所以，大数据工作负载对缓存的需求较高，需要进一步提高缓存的效率，例如通过增加缓存容量、提高缓存命中率等方式。大数据工作负载对 TLB 的需求较高，需要进一步提高 TLB 的效率，例如通过增加 TLB 容量、提高 TLB 命中率等方式。传统的架构设计可能无法满足大数据工作负载的需求，需要针对大数据工作负载进行专门的架构设计，以提高其性能和效率。与传统基准测试相比，BigDataBench 中的大数据工作负载的操作强度（计算与内存访问的比率）较低。这表明大数据工作负载更依赖于数据移动而不是指令执行，对浮点单元的需求较低。

### ****4主要贡献（Main Contributions）****

1. 全面的大数据基准测试套件BigDataBench：本文深入分析了大数据系统的基准测试需求，包括广泛的应用场景、多样的工作负载、真实的数据集、代表性的软件堆栈、先进的技术和易用性等方面。 开发了涵盖了广泛的场景和数据集的工具BigDataBench ，能够公平地衡量和评估大数据系统和架构。

2. 创新的合成数据生成方法： BDGS 工具能够生成可扩展的合成数据，，保留真实数据集的特征，方便用户进行测试。利用BDGS工具生成具有“4V”特性的合成大数据。数据生成器专为各种应用程序域（搜索引擎、estart 和社交网络）而设计，并将扩展到其他应用程序域。通过基于六个真实数据集开发数据生成器来证明其有效性，数据集涵盖三种代表性数据类型（结构化、半结构化和非结构化数据）、三种数据源（文本、图形和表格）。

3. 大数据工作负载特征分析： 通过对 BigDataBench 的工作负载进行特征分析，揭示了大数据应用在操作强度、缓存行为等方面的特点，为系统设计和优化提供了指导。特点如下：

（1）L1 指令缓存未命中率 (MPKI) 高： BigDataBench 中的大数据工作负载的 L1 指令缓存 MPKI 比传统基准测试高得多。

（2）L2 缓存 MPKI 高： BigDataBench 中的大数据工作负载的 L2 缓存 MPKI 也比传统基准测试高。

（3）L3 缓存对大数据工作负载有效： BigDataBench 中的大数据工作负载的 L3 缓存 MPKI 较低，表明 L3 缓存对大数据工作负载是有效的。

（4）ITLB 和 DTLB MPKI 高： BigDataBench 中的大数据工作负载的 ITLB 和 DTLB MPKI 也比传统基准测试高。

### ****5 结论（Conclusion）****

BigDataBench 是一个全面、创新且易于使用的大数据基准测试套件，能够有效评估大数据系统和架构的性能，并为系统设计和优化提供参考。可以用于衡量和比较大数据系统和架构。本文还提出了一种创新的数据生成方法和工具，在生成可扩展的大数据量的同时保持 4V 特性。此外，本文还报告了大数据的工作负载特征分析结果，得出与传统基准相比，大数据工作负载的操作强度更低，但对数据移动的需求高，数据输入量对微架构特征有显著影响，因此仅使用简单应用程序和小数据集的架构研究不足以满足大数据场景的需求。最后，在典型的商用处理器Intel Xeon E5645 上，本文发现对于大数据工作负载，L3缓存对大数据工作负载有效，而L1I缓存需要改进，因为大数据工作负载遭受高 L1I 缓存 MPKI需要更好的 L1I 缓存性能。

**6 参考文献（References）**

"BigDataBench: a Big Data Benchmark Suite from Internet Services." IEEE (2014):488-499