

文章编号：1005-8451（2016）09-0063-05

# 大数据背景下的IT基础架构研究

解辰辉，刘承亮

（中国铁道科学研究院 电子计算技术研究所，北京100081）

**摘要：**大数据技术的发展对于企业来讲是新的挑战和机遇，企业的信息化建设需要建立以云计算为基础的新型数据中心，传统IT基础架构已无法满足新的要求。为适应大数据战略发展，企业需要从存储、网络等多方面进行适应性改造。本文主要探讨了在大数据技术背景下，企业在IT基础架构的技术策略。

**关键词：**大数据；存储架构；软件定义网络；虚拟化

**中图分类号：**U29-39      **文献标识码：**A

## IT Infrastructure based on big data

XIE Chenhui, LIU Chengliang

(Institute of Computing Technologies, China Academy of Railway Sciences, Beijing 100081, China)

**Abstract:** The development of big data technology is a new challenge and opportunity for enterprises. Enterprise information needs to establish a new cloud-based data center and traditional IT infrastructure has been unable to meet new requirements. In order to meet the strategic development of big data, enterprises need adaptive transform from storage, networking and other aspects. This article mainly discussed the technical strategy of enterprises IT infrastructure under the background of big data technology.

**Key words:** big data; storage architecture; software definition network (SDN); virtualization

目前，大数据的概念已经成为IT界的热门话题，如何在大数据的背景下合理选择满足企业发展的技术架构，已经成为人们研究的重点。

企业IT架构已经进入了大数据阶段，企业级的大数据架构需要满足以下业务需求：

(1) 能够满足基于数据容量大、数据类型多、数据流转快的大数据基本处理需求，同时支持大数据的采集、存储、处理以及分析；(2) 能够满足企业级应用可用性、可靠性、可扩展性、容错性、安全性和数据保护方面的基本准则；(3) 能够满足用原始技术和格式实现数据分析的基本要求。

本文主要从存储、网络以及虚拟化技术几个方面，对大数据环境下企业IT基础架构进行了探讨。

## 1 大数据与存储结构的整合

### 1.1 传统存储架构对比分析

传统存储架构有直接连接存储(DAS)、存储区域网络(SAN)、网络连接存储(NAS)以及Internet

小型计算机系统接口(iSCSI)等方式，如表1所示。在Hadoop集群中，最普遍的存储方式就是各个节点的本地DAS，但大数据中需要更好地解决数据保护的问题，需要有效地整合当前的存储架构，并对存储架构提出了新的要求。FC-SAN和iSCSI架构提供块存储的外置设备，比较适合关系型数据库的存储，而NAS是提供文件及共享的外置存储设备，更适合基于文件的非结构化数据存储。

表1 常用存储架构比较

	DAS	SAN	iSCSI	NAS
存储网络	无	光纤+FC交换机	以太网	以太网
存储协议	SCSI	SCSI	SCSI	NFS、CIFS
存储空间	服务器自身硬盘，只能依靠纵向扩展	多台服务器共用后台存储设备，存储空间以LUN的形式划分给各服务器，存储空间独立	多台服务器共用后台存储设备，存储空间以LUN的形式划分给各服务器，存储空间独立	多台服务器共享后台存储设备，后台NAS设备的存储空间通过CIFS/NFS协议共享给服务器
数据共享	无	无	无	可同时对同一目录文件进行并发读写
可扩展性	差	一般	一般	较好
速率	无	8 G bps	1 G bps、10 G bps	1 G bps、10 G bps
成本	低	高	中	中

收稿日期：2016-06-15

作者简介：解辰辉，工程师；刘承亮，高级工程师。

## 1.2 分布式存储与软件定义存储

随着非结构化数据的海量增长，现有的存储架构已经越来越无法满足需求。比如，企业中有多台独立运行的NAS存储设备以及SAN存储设备，形成诸多“数据孤岛”，各个使用部门无法实现有效的信息共享，如果为不同的应用配置不同的集中存储，企业将需要购置大批分散的集中存储，从而无法有效地管理。此外，集中存储方式存在扩展性问题，存储磁盘容量只能进行纵向扩展，不仅扩充能力有限，并且成本较高，无法从根本上实现存储容量的水平扩展。由于数据中不仅存在着大量的频繁被处理的数据，也有很多对实时性要求不高，访问频率较低的数据，因此，如何有效地进行数据的存储及管理，是大数据存储的一个关键问题。

### 1.2.1 分布式存储

与以往企业中常用的集中式存储不同，分布式存储技术并不是将数据保存在某台存储介质（如存储磁盘阵列）中，而是通过网络（往往需要万兆环境）充分利用企业中的每台设备（主要为服务器）的存储空间，将分散的资源组成一台逻辑上统一的存储设备。

分布式存储的架构如图1，其特点如下。

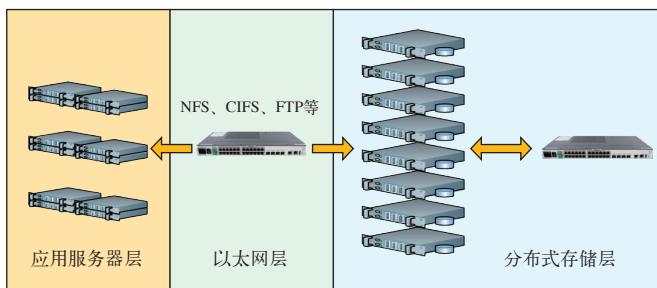


图1 分布式存储架构

#### 1.2.1.1 高扩展性架构

一般的分布式存储系统由3部分组成，包括应用访问层、存储节点、存储网络，每个部分都可以采用适当的技术而不用改变存储的架构，易于扩展。

#### 1.2.1.2 分布式操作系统

所有对存储系统的操作都由分布式操作系统统一的调度控制，并对各种资源进行动态的分派，协调各分布式任务并行完成。

#### 1.2.1.3 易于管理

分布式存储提供一种集中、易于管理的方式，对客户端没有任何的硬性影响，可以采用标准的访问协议（NFS、CIFS）访问分布式存储系统。

#### 1.2.1.4 负载均衡

分布式存储通过软件功能实现了前端网络及后端网络的负载均衡。应用访问存储系统的行为会通过相应的负载均衡策略，将访问的请求根据一定的原则分配到各个存储节点上。

### 1.2.2 软件定义存储

软件定义存储（SDS）是指通过软件对存储设备进行定义，并能够根据用户需求分配这些存储资源，SDS是软件定义数据中心不可缺少的一部分，其以虚拟化为基础，但又超脱了虚拟化范围。分布式存储的实现依赖于软件定义存储的能力，它将所有的控制工作都集中在统一存储服务平台上，这套平台会单独部署在服务器或虚拟机中。

在引入SDS以前，存储都是在项目开始阶段配置和部署的，在其生命周期中不再更改。如果要求更改虚拟机所利用的逻辑单元号（LUN）或卷的某些方面或功能，则在许多情况下，需要删除原始LUN或卷并创建具有所需功能的新卷。这是一项干扰性很强且非常耗时的操作，可能需要花费数周的时间进行协调。

SDS把应用于服务器的先进技术运用于存储领域，对异构存储资源进行抽象化处理，以支持存储在逻辑上的池化、复制和按需分发。这使存储层具有了聚合、灵活、高效和弹性扩展的特点，全面降低了存储基础架构的成本和复杂性。

综合来看，SDS具备如下3个特征。

#### 1.2.2.1 以应用为中心，实现存储使用自动化

SDS支持对异构存储池中的所有资源实施一致的策略，使存储的使用像为每个应用或虚拟机指定容量、性能和可用性要求那样简单。这种基于策略的自动化最大限度地利用了底层存储资源，同时将管理开销降至最低。

#### 1.2.2.2 与硬件无关的数据服务

SDS的提出让人们关注的焦点由存储硬件转变为了存储服务。独立于底层存储硬件的特点使得数据服务（如快照、克隆和复制）的分配极其敏捷和灵活。

### 1.2.2.3 通过硬盘和固态磁盘虚拟化确保数据持久性

随着现如今服务器功能的增多，软件定义存储的方案可让企业利用廉价的标准硬件来扩大其存储资源。利用固态磁盘和硬盘作为虚拟机的共享存储，可获得高性能、内置的恢复能力和动态可扩展性，并极大地降低存储的总体拥有成本。

### 1.2.3 Ceph技术

分布式存储的目标是利用软件定义的手段简化存储系统实现，满足大数据存储和分析过程中多样化的数据访问要求<sup>[1]</sup>。

目前，该领域最受关注的技术之一是 Ceph，其更符合软件定义存储的特性，Ceph 的存储类型是可以通过软件方式定义的。同样的一个分布式对象存储集群，可以通过安装不同的上层软件和对应的客户端程序，实现块存储、对象存储和文件系统存储功能，这一特性对于传统的存储系统难以想象。除此之外，Ceph 的存储策略，如备份策略、后台数据处理功能等，也都可以方便地通过软件方式加以定义或扩展。Ceph 主要特点为：

(1) 高扩展性：使用普通 x86 服务器，支持 10~1000 台服务器，支持 TB 到 PB 级的扩展；

(2) 高可靠性：没有单点故障，多数据副本，自动管理，自动修复；

(3) 高性能：数据分布均衡，并行化度高。

Ceph 将文件分割后均匀随机地分散在各个节点上，采用了 CRUSH 算法来确定对象的存储位置，只要有当前集群的拓扑结构，Ceph 客户端就能直接计算出文件的存储位置，直接跟对象存储设备（OSD）节点通信获取文件而不需要询问中心节点获得文件位置，这样就避免了单点风险。

## 2 大数据网络架构

### 2.1 大数据对网络架构的要求

在大数据中，其高速度的特征增加了网络中对实时性和大负载事务的要求。这对企业现有的网络构架提出了新的要求。

(1) 网络连接必须是健壮的，要足以保证数据可以快速、高效地传输；

(2) 大数据是实时变化的，在一定时间内的网

络流量可能会时高时低。当数据加载任务被请求时，数据必须通过高速网络传输到各集群节点中，为防止在传输过程中出现延时等问题，必须拥有足够的网络资源池；

(3) 网络的一致性问题，大数据应用一般具有较高的同步性，事务作业往往并行计算，而各个作业之间较大的性能差异可能会引发应用程序的故障。因此，网络不仅要能足够高效，而且要在空间和时间上具有一致的性能；

(4) 灵活的交换机配置能力是高效网络的要点之一，企业的大数据部署规模可能是从小到大，这就需要考虑如何将目前的网络结构平衡地扩展以支持未来的部署规模。

### 2.2 软件定义网络

传统企业网络的建设需要依赖于大量的网络设备（如交换机、路由器、防火墙等），以及各种复杂的网络协议（软件）。网络工程师负责配置各种策略，以应对各种各样的网络事件和应用场景。他们需要手工地将这些策略转换为低层的配置命令，这些繁杂的任务通常只能通过有限的工具完成，使得网络管理控制和性能调优任务总是带有挑战性并易于出错。

软件定义网络（SDN）能够将网络控制从物理基础设施中解耦出来。通过软件和虚拟化技术，从更加全面的角度上对网络设备进行控制，它将网络管理的功能从网络设备中分离出来，使得用户可以构建简单、灵活且高度可扩展的网络。不同的网络设备通过开放的接口进行整合，这已经成为简化或自动化网络配置的一种主流趋势，用户不需要手动操作网络中的任何交换设备，就可以快速地配置更多的网络功能。

SDN 的网络架构从原理上可分为 2 个平面，数据平面和控制平面，如图 2 所示，其中，数据平面处理和交付网络上传输的数据包，控制平面则是数据网络中做出转发决定的部分，如路由协议，策略以及网络上运行这些协议的软硬件资源等。

控制器是 SDN 的核心，负责对底层转发设备的集中统一控制，同时向上层业务提供网络能力调用的接口，在 SDN 架构中具有举足轻重的作用。控制器会从与设备一起工作的 SDN 软件中提取物理网

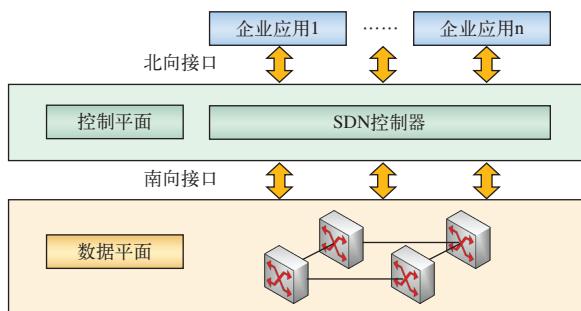


图2 SDN网络架构

络设备。控制器与网络设备之间是高度融合的，在OpenFlow环境中，控制器会使用OpenFlow协议和NETCONF协议与交换机进行联系。OpenFlow是将流数据发送到交换机的API，而NETCONF是网络配置API。

SDN控制器对网络的控制主要是通过南向接口协议实现，包括链路发现、拓扑管理、策略制定、表项下发等，其中，链路发现和拓扑管理主要是控制其利用南向接口的上行通道对底层交换设备上报信息进行统一监控和统计，而策略制定和表项下发则是控制器利用南向接口的下行通道对网络设备进行统一控制。

SDN北向接口是通过控制器向上层业务应用开放的接口，目标是使得业务应用能够便利地调用底层的网络资源和能力。通过北向接口，网络业务的开发者能以软件编程的形式调用各种网络资源，同时上层的网络资源管理系统可以通过控制器的北向接口调控整个网络的资源状态，并对资源进行统一调度。北向接口是直接为业务应用服务的，因此其设计需要密切联系业务应用需求，具有多样化的特征。

### 3 大数据与虚拟化技术的整合

虚拟化是指计算组件在虚拟的平台上而不是在真实的物理基础上运行，虚拟化技术的提出是为了简化管理和优化资源，图3所示是传统技术架构与虚拟化技术架构的比较。虚拟化技术提供给应用一个集中的、共享的资源池，同时屏蔽了底层计算资源的异构，对于用户而言，在计算资源池中的所有计算资源都是标准化的桌面、存储、服务器和网络虚拟设备，计算能力是能够被灵活定义的。这使得计算资源的可用性极大地提升，所有资源可以被统一部署、管

理、动态调配以及灵活收缩，避免了计算资源的浪费。因为采用标准化的计算能力，可以实现管理自动化和按需分配，降低了计算资源的投资成本，因此，虚拟化技术已经成为支撑大数据应用的主流技术。

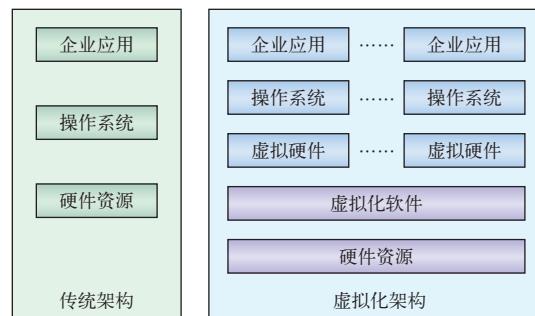


图3 传统架构与虚拟化架构对比

主机层面的虚拟化技术，可以实现“一分多”，即将一台服务器虚拟成多台虚拟机，在进行Hadoop平台的安装和实验阶段，可以采取这样的方法来进行。主机虚拟化技术也可以实现“多合一”，将多台服务器虚拟成一台虚拟服务器，实现不同大数据计算集群的统一关联和资源共享。图4所示为虚拟化技术统一大数据平台，可以使得计算资源池能够按需快速、容易地提供新的数据集群，允许工作负载的混合，利用虚拟机来提供隔离，基于虚拟的拓扑来优化数据性能，给予的虚拟拓扑使得系统更加可靠。

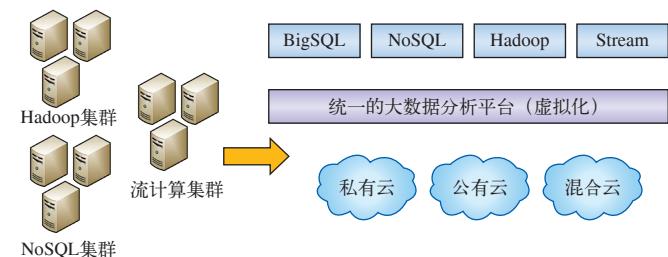


图4 大数据分析平台

### 4 结束语

大数据技术的发展促进了传统IT架构的转型，使之变得更加灵活及高效。本文主要提出了传统IT架构存在的一些问题，并探讨了软件定义存储、软件定义网络以及虚拟化的一些技术特点，这些都是软件定义数据中心的重要组成部分，其解决的核心的问题是如何让企业以更小的投入（人力，财力）来获得更加灵活、快速的业务部署和管理。在大数据时代人们关注的焦点已不仅仅是上层的应用，对于

(下转 P71)

析、NoSQL 和交互式分析（SQL）。该基准提供公开的数据集和负载。

铁路行业用户可通过研究以上3种基准中各负载与各种铁路大数据应用之间的近似映射关系，以尽可能利用公开测试基准的负载和数据集，减少铁路业务建模和数据收集清洗的工作量。对于映射负载集中可获得部分铁路业务数据的负载，可利用大数据生成工具将这些数据生成大数据集来替代原有的公开数据集，以取得更加符合铁路数据特点的测试效果。经过多年的发展，铁路行业已积累了丰富的业务模型，其中有些甚至做了并行化处理，对此可选取部分模型加入到映射负载集合中，以构建尽可能反映铁路业务特点的大数据服务平台测试基准。

#### 4 结束语

大数据服务平台对于当前铁路智能化发展具有非常重要的作用。本文针对大数据服务平台在铁路行业应用中面临的选型问题，通过对目前主流的大数据平台技术架构进行对比分析，为智能铁路大数据服务平台技术架构选型提供了一定的支撑；同时，结合铁路领域大数据应用的特点，提出了智能铁路大数据服务平台选型中的测试指标体系，给出测试基准的确定方法，为智能铁路大数据服务平台相关产品选型提供一定的技术参考。

#### 参考文献：

- [1] 马小宁. 智能铁路总体框架研究 [D]. 北京：中国铁道科学研

责任编辑 徐侃春

(上接 P66)

基础架构层面也在不断地进行技术创新，这也将是我们的研究重点。

#### 参考文献：

- [1] 赵刚. 大数据技术与应用实践指南 [M]. 北京：电子工业出版社，2013.
- [2] 郎羽生. SDN 应用编排和资源管理技术的发展 [N]. 人民邮电报, 2015-12-10.
- [3] 黄韬, 刘江, 魏亮, 等. 软件定义网络核心原理与应用实践 [M]. 北京：人民邮电出版社，2014.
- [4] 傅帅. 软件定义网络在云计算领域的实施部署 [J]. 信息通信技术, 2015 (2) : 41-45.

- 究院, 2013.
- [2] 史天运, 郭歌, 李平. 智能铁路总体框架的研究 [C]. 第六届中国智能交通年会暨第七届国际节能与新能源汽车创新发展论坛优秀论文集（上册）—智能交通. 2011.
- [3] 于富东. 大数据平台的关键技术及组网方案 [J]. 电信科学, 2015, 31 (7) : 158-163.
- [4] 王晖, 唐向京. 共享开放的运营商大数据平台架构研究 [J]. 信息通信技术, 2014 (6) : 52-58.
- [5] 2012 Ghazal A, Rabl T, Hu M, et al. BigBench: towards an industry standard benchmark for big data analytics[C]. ACM SIGMOD International Conference on Management of Data. ACM, 2013:1197-1208.
- [6] Cooper B, Silberstein A, Tam E, et al. Banchmarking cloud serving systems with YCSB[C]. Proceedings of the 1st ACM Symposium on Cloud Computing. Indianapolis, Indiana, USA, 2010: 143-154.
- [7] 詹剑锋, 高婉铃, 王磊, 等. BigDataBench: 开源的大数据系统评测基准 [J]. 计算机学报, 2016 (1) : 196-211.
- [8] Paulson, A. Rasin, D. J. Abadi, D. J. DeWitt, S. Madden, and M. Stonebraker, A comparison of approaches to large-scale data analysis[C]. SIGMOD Conference, pp. 165-178, 2009.
- [9] Ferdman M, Adileh A, Kocberber O, et al. Clearing the clouds: a study of emerging scale-out workloads on modern hardware[J]. ACM SIGPLAN Notices, 2012, 47(4):37-48.
- [10] 金澈清, 钱卫宁, 周敏奇, 等. 数据管理系统评测基准：从传统数据库到新兴大数据 [J]. 计算机学报, 2015, 38 (1) : 18-34.

[5] 赵为民. 温故知新：详解存储集群的优势和应用 [J]. 网络与信息, 2012, 26 (5) : 34-35.

[6] 刘圆, 王峰, 杨明川. 面向大数据的分布式存储技术研究 [J]. 电信技术, 2015 (6) : 33-36

[7] 刘露, 陈清金, 张云勇. 对 SDN 技术的研究与思考 [J]. 互联网天地, 2013 (3) : 1-3.

[8] 房秉毅, 张云勇, 陈清金, 等. 云计算网络虚拟化技术 [J]. 信息通信技术, 2011 (5) : 50-53.

[9] 陈熹, RickySun. 软件定义数据中心：技术与实践 [M]. 北京：机械工业出版社, 2014, 12.

[10] 詹明非. 软件定义存储技术及其应用研究 [J]. 电信技术, 2014 (12) : 30-32.

责任编辑 徐侃春