



一种基于分布式服务器集群的可扩展负载均衡策略技术

孙乔¹, 邓卜侨¹, 王志强², 裴旭斌²

(1. 北京国电通网络技术有限公司, 北京 100070;

2. 国网浙江省电力公司信息通信分公司, 浙江 杭州 310007)

摘要: 提出了一种基于软件定义网络的分布式数据库负载均衡算法, 将数据、控制、应用分离的同时计算服务器集群中单个服务器的实际负载。通过查询流量采样记录来决策最少连接的服务器路径, 减少了访问请求的响应时间, 提高了系统吞吐量和容错能力, 实现了充分利用服务器资源的目的。内网中的分布式数据库实验对通用负载均衡技术和基于软件定义网络的负载均衡技术进行了比较, 在不同服务器集群的负载状态下, 后者的平均响应时间小于前者, 并得到了更好的负载均衡效果。

关键词: 可扩展负载均衡策略; 软件定义网络; 自适应流量采样; 分布式服务器集群

中图分类号: TN929.52

文献标识码: A

doi: 10.11959/j.issn.1000-0801.2017264

A scalable load balancing strategy based on distributed server cluster

SUN Qiao¹, DENG Buqiao¹, WANG Zhiqiang², PEI Xubin²

1. Beijing GuoDianTong Network Technology Co., Ltd., Beijing 100070, China

2. Information&Telecommunication Branch of State Grid Zhejiang Electric Power Company, Hangzhou 310007, China

Abstract: A distributed database load balancing algorithm based on software defined networking was proposed, which separated the data, control and application to calculate the actual load of a single server in the server cluster. By querying the streams sampling record, the minimum number of server paths were determined, the request response time was reduced, the system throughput and fault tolerance were improved, and the full use of the purpose of server resources was achieved. The general load balancing technology was compared with the load balancing technology based on the software defined networking in the internal network of the distributed database experiment, the result shows that in the load state of the different server clusters, the average response time of the latter is smaller than that of the former, and can get better load balancing effect.

Key words: extensible load balancing strategy, software defined networking, adaptive stream sampling, distributed server cluster

收稿日期: 2017-06-07; 修回日期: 2017-08-31

基金项目: 国家电网公司科技基金资助项目 (No.SGBJDK00KJJS1500180); 国网信息通信产业集团有限公司科技基金资助项目 (No.SGITG-KJ-JSKF[2015]0010)

Foundation Items: Science and Technology Project of State Grid (No.SGBJDK00KJJS1500180), Science and Technology Project of State Grid Information& Telecommunication Group Co., Ltd. (No.SGITG-KJJSKF[2015]0010)

1 引言

互联网技术的快速发展使大型互联网服务提供商的服务器集群面临更严重的挑战。随着用户和网络带宽的增长,服务器需要在短时间内处理大量访问请求。如果服务器不能及时处理用户访问请求,增加用户的等待时间,则会影响用户体验,大大降低服务质量。这使服务器成为网络的新瓶颈,从而迫使研究人员开始研究如何提高服务器性能。对于这些通过网络连接起来的数据库而言,必须能够实时处理大量的用户请求,而且必须能向用户提供高质量服务,这就要求系统能使每个数据库高效运行,充分发挥它们的作用。

为了提高服务器性能,企业采取了一系列措施,如提高CPU处理速度、增加服务器缓存容量、使用高速磁盘阵列以及构建服务器集群。简单地升级硬件系统不仅会导致现有资源闲置,而且当业务继续扩大时,企业同样将面临困难局面。通过建立服务器集群,企业可以将访问请求转发到多个服务器。这种方法在一定程度上提高了服务器性能。然而,该解决方案有一个新问题,当服务器集群接收到一个服务器应该响应的访问请求时,如果控制系统不能合理分配访问请求,会出现服务器之间的负载不均衡情况。如何才能提高系统的响应速度、稳定性和扩展性,并且保护最初的硬件投资,避免大量用户请求对系统带来的冲击,一个解决方案就是负载均衡技术。

负载均衡的主要目标是通过硬件或者软件的策略来决定由哪个数据库执行用户发出的请求。负载均衡机制在后端数据库间分发用户请求,以达到减少系统瓶颈、增强系统响应能力的目的。因此,负载均衡对网络环境下的异构数据库系统提升性能与效率具有重要意义。

负载均衡常见的方法是在服务器集群前部署一个负载均衡器,因此它可以为几个服务器合理分配负载以充分利用服务器资源。负载均衡技术可以

有效解决服务器之间的负载不均衡问题,从而减少了访问请求的响应时间,提高系统吞吐量和容错能力。经典的负载均衡技术需要使用昂贵的硬件设备,而且不能实现流量负载的精确控制。这些限制使传统负载均衡技术不适合大规模应用。

通常的负载均衡技术需要较多硬件费用的投入,且无法进行负载定量控制,因此无法实现在较大规模的分布式数据库应用场景中。软件定义网络(software defined networking, SDN)技术的出现给管理者提供了具有低成本和灵活操作优势的流量管理技术。软件定义网络是计算机网络管理员通过提取低级功能来管理网络服务的新方法。它把传统的网络体系结构中的控制平台和转发平台分开,运行在控制平台的控制器可以通过管理开关流程表实现数据转发的控制^[1-3]。

本文提出了一种基于软件定义网络的分布式数据库负载均衡算法,将数据、控制、应用分离的同时计算服务器集群中单个服务器的实际负载,通过查询流量采样记录来决策最少连接的服务器路径,减少了访问请求的响应时间,提高了系统吞吐量和容错能力,实现了充分利用服务器资源的目的。

2 基本原理

由于互联网资源及可用性的迅速发展,分布式数据库计算应用一直呈现指数级增长,在业界被广泛构建和使用。随着分布式数据库规模的扩大,其服务器集群需要处理更多、更大体量的查询等请求,如何保持性能的稳定性是十分重要的问题。负载均衡方法被认为是利用并行和分布式系统的全部资源的先决条件之一。

负载均衡允许在一个或多个服务器、数据中心、硬盘驱动器或其他计算资源之间分配工作负载,从而可提供一种机制,用于在位于数据中心的任意数量的应用程序部署中分发应用程序请求。负载均衡机制可以大致分为集中式或分散式、动态或



静态、周期性或非周期性。负载均衡技术的研究包括很多内容,如在最小执行时间中,如何在节点不同负载条件下将命令分配给预期执行速度最快的节点,如何提高如分布式数据挖掘等数据密集型应用程序中的计算性能等现实应用问题^[4-6]。

在如图 1 所示的服务器负载均衡模型的计算方式中,负载均衡可以改善多个计算资源之间的工作负载分配。负载均衡旨在优化资源使用、最大化吞吐量、最小化响应时间,并避免任何单一资源的过载。用户的访问请求首先通过互联网发送到负载均衡器,而不是直接发送到服务器集群。负载均衡器通过使用指定算法选择最优服务器并转发请求到服务器。使用具有负载均衡而不是单个组件的多个组件可以通过冗余来增加可靠性和可用性。负载均衡通常涉及专用软件或硬件,如多层交换机或域名系统服务器进程。负载均衡与通道绑定不同之处在于负载均衡在网络套接字(OSI 模型层 4)基础上划分网络接口之间的流量,而通道绑定意味着在每个数据分组较低级别的物理接口之间或数据链路的基础上,使用最短路径桥接协议^[7-9]。

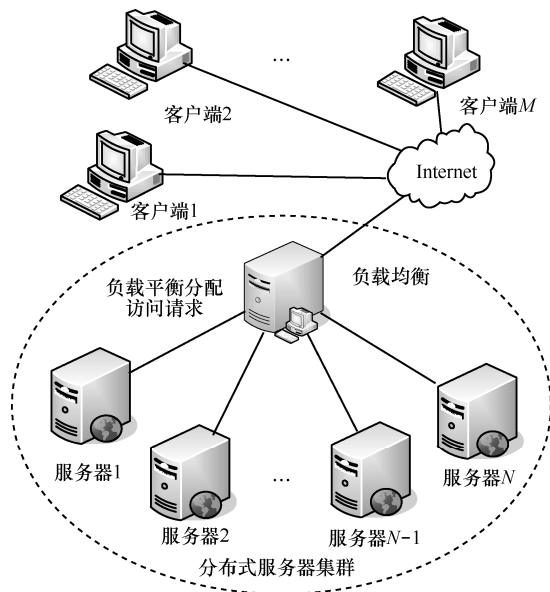


图 1 服务器负载均衡模型

负载均衡算法通常可以分为静态算法和动态算法。静态负载均衡算法不考虑服务器的负载状

态,所以它适合可以提前预测负荷状态的情况。静态算法主要包括 Round-Robin 算法、比率算法和优先级算法。Round-Robin 算法中,访问请求根据到来序列顺序分配给每个服务器;比率算法中,根据预设比设定每个服务器相应的负载比例;优先级算法中,当负载不均衡时,备份服务器将被启用。

动态负载均衡算法是指负载均衡器可以动态分配负载,以根据服务器的负载状态优化服务器。典型动态算法主要包括最少连接算法、预测算法和响应时间算法。最少连接数算法的负载均衡器定期检测服务器和用户之间已建立的连接数,当一个新的访问请求到达时,负载均衡器将该请求转发至少连接到服务器;预测算法的控制器可以通过预测模型预测下一个周期的服务器负载状态;响应时间算法的负载均衡器通过发送探测指令估计每个服务器的负载状态。

由于传统网络的局限性,负载均衡技术的发展逐渐遇到瓶颈。大多数负载均衡器为特定应用或产品定制,但其代码的可重用性和不同架构的适用性不佳。此外,用于负载均衡的复杂软件或硬件系统会增加企业的运营成本。

3 软件定义网络的分布式数据库负载均衡算法

软件定义网络起源于 2006 年美国斯坦福大学的研究课题,并于 2009 年被正式提出。软件定义网络的概念是:软件定义网络利用分层的思想将数据与控制相分离,在控制层包括具有逻辑中心化和可编程的控制器,可掌握全局网络信息,方便管理配置网络和部署新协议等;在数据层包括 dumb 交换机,交换机仅提供简单的数据转发功能,可以快速处理匹配的数据分组,适应流量日益增长的需求。两层之间采用开放的统一接口进行交互,控制器通过标准接口向交换机下发统一标准规则,交换机仅需按照这些规则执行相应的动作即可。软件定义网络技术具有控制与传输分离、集中式控制和软

件可定义的特征^[10]。因此软件定义网络技术的基本特性增加了网络部署的灵活性和网络的可编程性,因此可以满足快速高效的网络需求^[11]。

由于传统网络负载的高成本使得负载均衡技术很难得到广泛应用,如图2所示的OpenFlow技术提供了一个具有低成本、操作灵活等优点的流通管理技术。已经有一些研究关注OpenFlow网络中的负载均衡问题,另一些研究则提出了另一种负载均衡解决方案:通过使用多个OpenFlow控制器来处理多服务的负载,但是这些方法仍然无法实现真正的负载均衡^[12,13]。

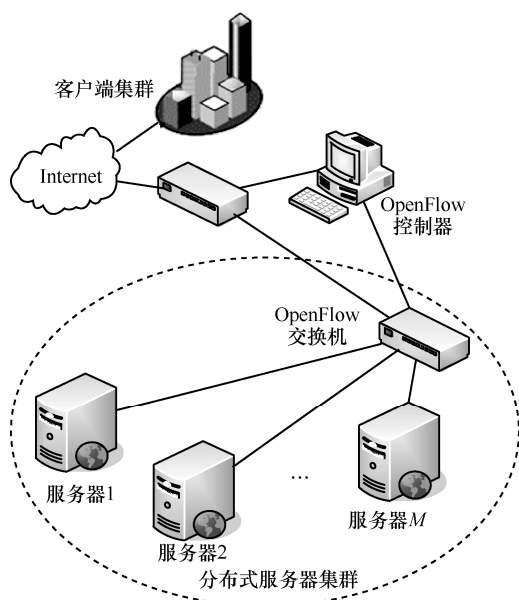


图2 OpenFlow 负载均衡结构

为了获得每个服务器的负载 L 、计算能力 C 和优先级 P , 需要对计算服务器集群中单个服务器的实际负载进行统计。本文采用的统计算法是每 0.1 s 对每个服务器计算一次上述的 L 、 C 和 P 值, 并将其定义为 $S=L \cdot C \cdot P$ 的服务优先级统计结果存储在服务器优先级数据库中。当用户向服务器发送请求时, 此服务需求同时进入负载统计的计算路径, 并与服务器优先级数据库中的所有 S 值进行比对, 找到 S 值大于用户服务需求的服务器, 然后使用轮询方法将用户的请求发送到这些服务器上。其流程如图3所示。

为了应对软件定义网络通常存在的可扩展性问题, 还需要采用基于软件定义网络的流量采样方法, 对流量进行部分采样并对数据流进行记录, 便于通过查询流量采样记录决策最少连接的服务器路径。其具体步骤如下。

步骤 1 检查新数据分组是否已经到达交换机并处于流量监控操作模式下。

步骤 2 如果符合任何一个操作模式, 则对数据分组的计数器进行更新, 并将数据分组导向下一个目标路径。

步骤 3 如果不符合, 则查询是否符合被采样的规则, 如果与其中的某个规则匹配, 则将其引导至下一个路径并同时发给控制器, 且在记录表中增加信息以对后续数据流进行采样。

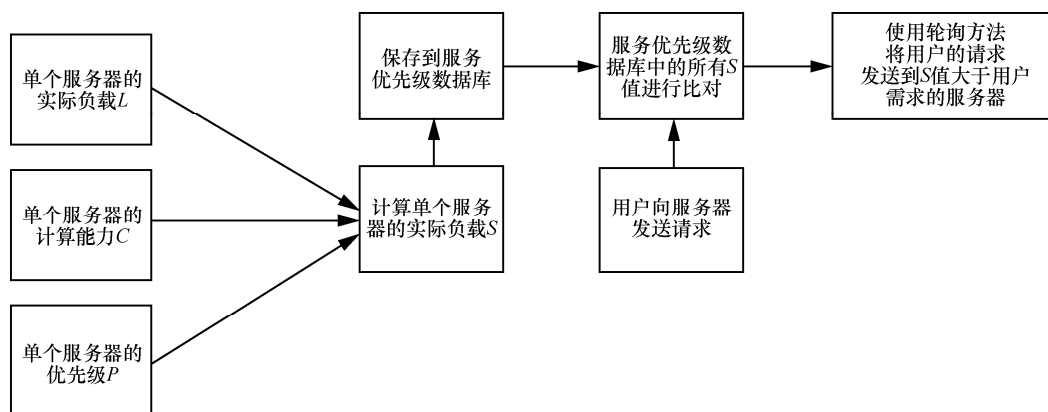


图3 单个服务器实际负载统计方法



一般来说,通过采用 OpenFlow 控制器方式,可以将管理决策都发送至交换机来存储和收集流量统计信息,并在每个交换机上以计数器形式存储全部流信息的记录。但现有 OpenFlow 技术的 TCAM 条目数量和处理能力无法支持过大的流量统计记录,因此在采集流量测量时实施流量采样记录,以达到获得较细粒度的数据并降低 OpenFlow 控制器操作次数的目的。当交换机空闲或超时,系统就进行测量并实施流统计,其流程如图 4 所示。

事实上,上述统计方式对于较大时间跨度的流量统计数据来说是非实时的。当需要在具有高度波动流量的动态环境中获得准确的测量结果时,需要增加交换机端口自适应函数,即通过轮流监听 TCP 或 UDP 等端口,在源端口发出数据后,在某些流量超过阈值的目标端口,将其数据分组自适应调整到其他流量较低的目标端口,可以避免由于源端口和目标端口冲突而增加的计算负担。

4 实验与测试结果

为了评价上述方法的性能,在现有网络环境中进行了评估与测试。实验具体配置为:64 台服务器 IBM-x3650M4-2U(2CPU,6 核 12 线程 Xeon E5-2620,48 GB 内存,2 TB 硬盘)分别作为 DataNode,操作系统为 Ubuntu11.10 及 Windows7,

均预装 FTP 服务。将其中一台服务器定义为主服务器并部署了内网企业网站,网站的文本、图像、视频等数据资源分别存放在多个服务器上,通过改变访问网站的访问量来模拟负载均衡的效果。

服务器负载取决于负载指标,负载指标是衡量服务器负载状态的特征值,用于描述服务器的忙闲程度,记录服务器处理服务的能力,一般有 CPU 处理能力、CPU 利用率、CPU 就绪队列长度、内存、磁盘容量、网络带宽、当前连接数等^[13]。本文选取有代表性的 CPU 利用率和网络上行速度作为负载指标,通过不同的采样率下的指标来比较一般负载均衡方法与无端口自适应/有端口自适应流量采样负载均衡方法的效果。在 24 h 内对该网络环境中的性能进行统计,其结果如图 5 和图 6 所示。

从图 5 可以看出,在 CPU 利用率最小值方面,提出的流量采样负载均衡方法比一般的负载均衡方法存在一定优势,尤其是采样率相对较高的情况下;而有端口自适应流量采样负载均衡方法在 CPU 利用率最大值这一指标上体现出的负载均衡程度存在明显优势。

图 6 则显示出,采样负载均衡方法与一般的负载均衡方法相比,在采样率较低时差异不大;随着采样率的增加,有端口自适应流量采样负载均衡方法比无端口自适应流量采样负载均衡方法和一般负载均衡方法在各项指标上都表现出网络

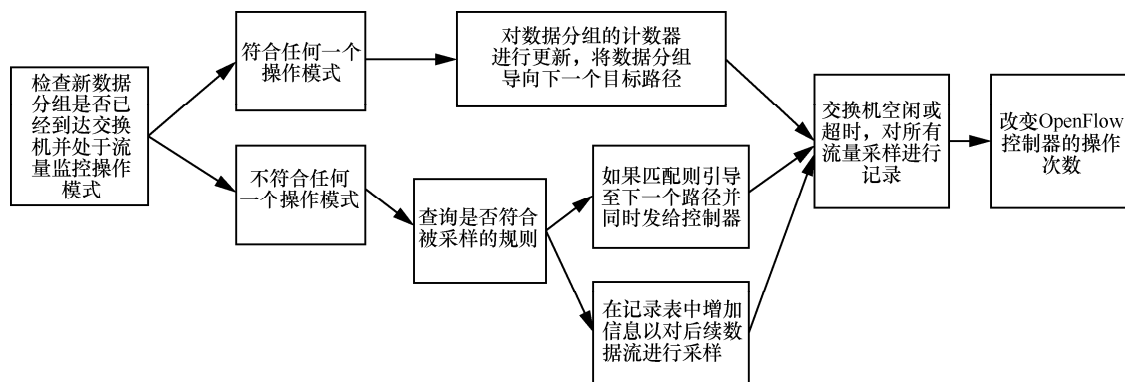


图 4 基于软件定义网络的流量采样流程

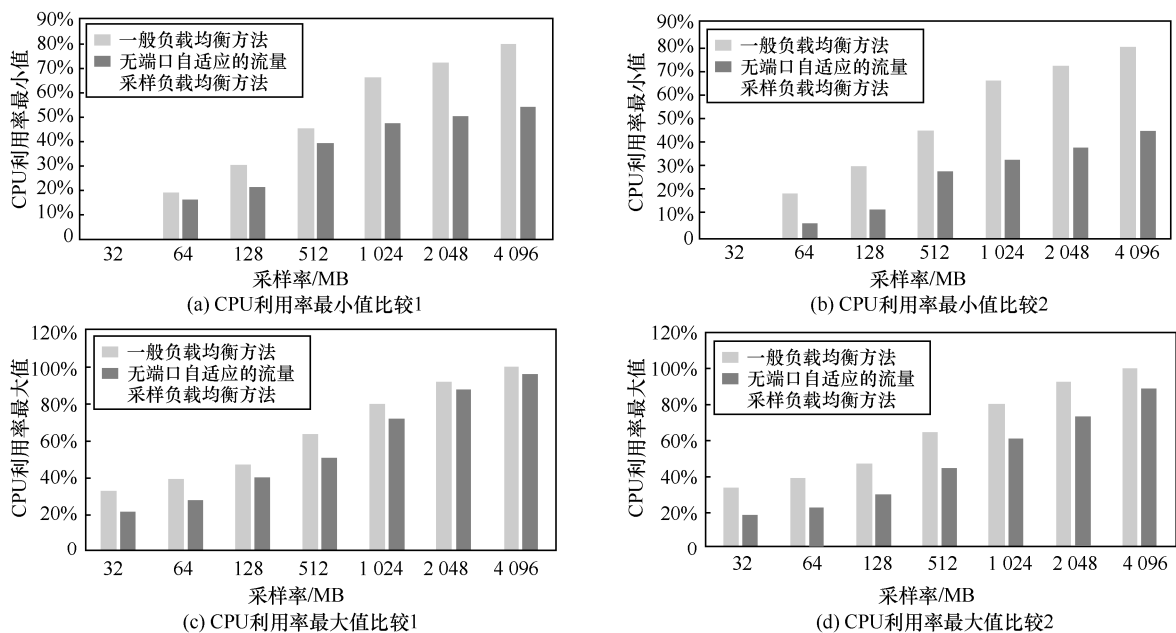


图5 CPU 利用率最小值和最大值比较

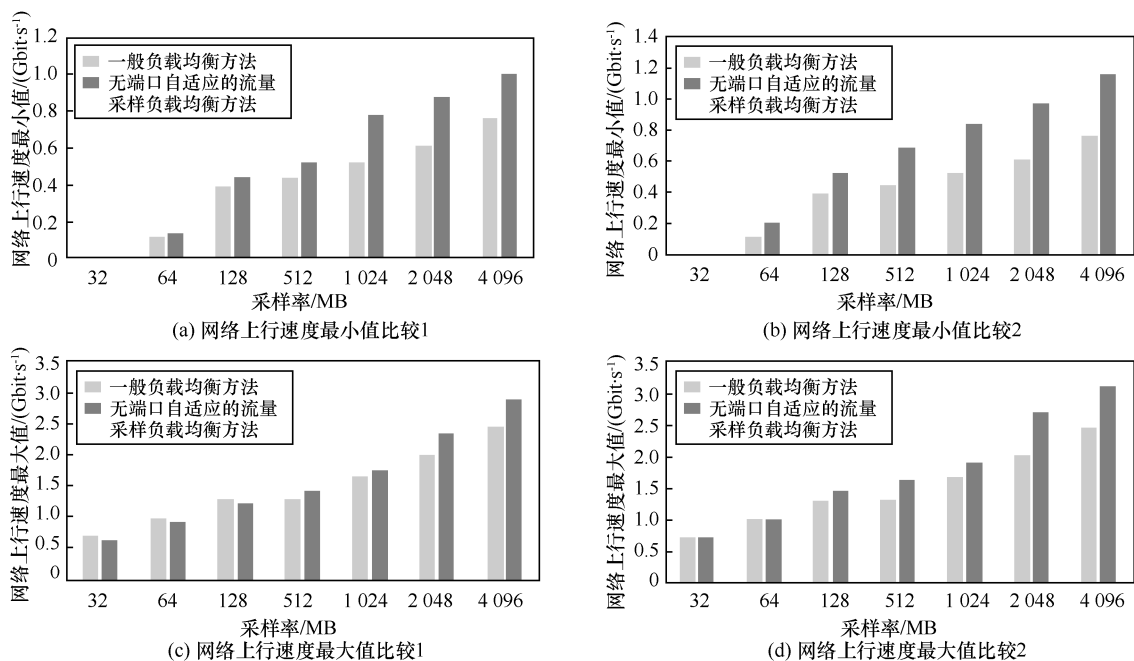


图6 网络上行速度最小值和最大值比较

上行速度指标有增加的趋势。

5 结束语

提出了一种基于软件定义网络的分布式数据库负载均衡解决方案,为了减少控制器的开销和交换机的计算代价,提出了一种基于 OpenFlow 的受限流量采样方法。实验结果表明,与一般的负

载均衡方案相比,本文提出的方法在降低计算代价和提高负载均衡效果上有一定优势。下一步将研究如何进一步提高检索统计数据的精确性。

参考文献:

- [1] ZHANG H, GUO X. SDN-based load balancing strategy for server cluster[C]//2014 IEEE 3rd International Conference on Cloud Computing and Intelligence Systems (CCIS), Nov 27-29,



- 2014, Shenzhen, China. New Jersey: IEEE Jersey, 2014: 662-667.
- [2] 张朝昆, 崔勇, 唐嵩伟, 等. 软件定义网络 (SDN) 研究进展[J]. 软件学报, 2015, 26(1): 62-81.
ZHANG C K, CUI Y, TANG H Y, et al. State-of-the-art survey on software-defined networking(SDN)[J]. Journal of Software, 2015, 26(1): 62-81.
- [3] 刘同. 负载均衡技术在数据库集群系统中的应用与实现[D]. 长沙: 国防科学技术大学, 2009.
LIU T. Application and realization of load balancing technology in database cluster system[D]. Changsha: National University of Defense Technology, 2009.
- [4] DAS S K, SEN S K, JAYARAM R. A dynamic load balancing strategy for channel assignment using selective borrowing in cellular mobile environment[J]. Wireless Networks, 1997, 3(5): 333-347.
- [5] DASGUPTA K, MANDAL B, DUTTA P, et al. A genetic algorithm based load balancing strategy for cloud computing [J]. Procedia Technology, 2013(10): 340-347.
- [6] DAM S, MANDAL G, DASGUPTA K, et al. An ant colony based load balancing strategy in cloud computing [M]. Berlin: Springer International Publishing, 2014.
- [7] MONDAL B, CHOUDHURY A. Simulated annealing (SA) based load balancing strategy for cloud computing[J]. International Journal of Computer Science and Information Technologies, 2015, 6(4): 3307-3312.
- [8] IZAGIRRE A, MAKOWSKI A M. Light traffic behavior under the power-of-two load balancing strategy: the case of heterogeneous servers[J]. arXiv preprint arXiv:1701.06004, 2017.
- [9] Load_balancing [EB/OL]. (2017-03-01)[2017-06-07]. https://en.wikipedia.org/wiki/Load_balancing.
- [10] 王力, 赵季红, 曲桦, 等. 软件定义承载网中基于负载均衡的虚拟网络资源分配算法[J]. 电信科学, 2015, 31(11): 32-38.
WANG L, ZHAO J H, QU H, et al. Virtual network resource allocation algorithm based on load balance in carrier-SDN[J]. Telecommunications Science, 2015, 31(11): 32-38.
- [11] LIU J, LAI Y, DIAO Z, et al. A trusted access method in software-defined network[J]. Simulation Modelling Practice and Theory, 2017(74): 28-45.
- [12] TAHA EI H, SALLEH R, KHAN S, et al. A multi-objective software defined network traffic measurement[J]. Measurement, 2017(95): 317-327.
- [13] 杨伟. 基于时间序列预测的 IP 控制网关集群的设计与实现[D]. 苏州: 苏州大学, 2006.
YANG W. Design and implementation of IP control gateway cluster based on time series prediction[D]. Suzhou: Soochow University, 2006.

[作者简介]



孙乔 (1978-), 男, 北京国电通网络技术有限公司中心副总监, 主要研究方向为实时库及自主关系数据库。



邓卜乔 (1985-), 男, 北京国电通网络技术有限公司技术总监, 主要从事关系数据库及分布式关系数据库研发工作。



王志强 (1966-), 男, 国网浙江省电力信息通信分公司副总工程师, 主要研究方向为电力信息新技术应用。



裴旭斌 (1973-), 男, 国网浙江省电力信息通信分公司数据库运维专职, 主要研究方向为数据库应用技术。