文章编号:1001 - 9081(2005)02 - 0245 - 03

内容配送网路由请求系统的研究

胡 鹏,洪佩琳,李津生,杨海松
(中国科学技术大学 电子工程与信息科学系,安徽 合肥 230027)
(hupeng@mail.ustc.edu.cn)

摘 要:分析了内容配送网现有的 DNS 路由请求机制,提出一种基于策略的路由请求系统 PBR²S。该系统能够根据网络的当前状态,动态地更新 DNS 名字服务器,可以较为安全地处理突发数据。对 PBR²S 的架构、协议和功能模块做了详细的说明,最后分析了系统的特性。

关键词:路由请求;策略;判决;动态更新;安全;可扩展性
中图分类号: TP393.04 **文献标识码**:A

Research on request-routing system in CDN

HU Peng, HONG Pei-lin, LI Jin-sheng, YANG Hai-song
(Department of Electronic Engineering and Information Science, University of Science and Technology of China, Hefei Anhui 230027, China)

Abstract: After DNS-based request-routing in CDN(Content Delivery Network) was analyzed firstly, the PBR²S(Policy-Based Request Routing System) architecture with policy-based of request-routing system was proposed in this article. This system could update DNS Name Server dynamically according to the state of network, it could process flash crowd more securely. Architecture, protocol and functional module were described in detail, at last the character of the system was summarized.

Key words: request-routing; policy; decision; dynamic update; security; extendiblity

0 引言

随着网络规模的日趋增长和新业务的不断涌现,网络负载呈现加重的趋势。为了优化 Internet 的数据流动和提高服务质量(QoS),近几年来内容配送网(Content Delivery Network,CDN)技术应运而生。CDN 的基本思想是在距离用户较近的位置分散地放置多个边缘服务器,内容提供商将所发布的内容复制到这些边缘服务器中,并由它们代表源服务器向用户提供内容服务。研究表明,CDN 的应用能够有效地降低服务器端的负载,缓解网络拥塞,降低业务在网络中传输的延迟[1]。

根据 IETF 定义[2],一个完整的 CDN 架构应包括如下四个组件:路由请求系统、分发系统、计费系统和边缘服务器。路由请求系统将用户的请求重定向到一个"最佳"的边缘服务器上(最佳的含义可以是地理位置最近或者服务器负载最轻);分发系统将源服务器的内容发布到边缘服务器上,使得内容更加靠近用户;计费系统负责跟踪、记录数据在 CDN 各组件间的流动信息,并以此为依据实施计费;边缘服务器则是内容分发的目的地,也是向用户直接提供服务的实体。

1 路由请求相关问题

路由请求系统是 CDN 架构中至关重要的组成部分。因为它需要根据用户的请求以及边缘服务器的状态信息,实时、智能地决策判断,从而选取最佳的边缘服务器。由此可见,设计有效的路由请求系统是提升 CDN 整体性能的关键。为了便于集中讨论路由请求机制,这里假设源服务器已经向所有边缘服务器分发了内容。

1.1 基于 DNS 的路由请求机制

路由请求机制具体可分为基于域名服务、基于传输层和基于应用层等几类[3]。鉴于域名服务(DNS)在 Internet 中的广泛应用,因此基于 DNS 的路由请求比传输层和应用层路由请求具有更高的可操作性。其基本规程是 DNS 服务器对用户的域名查询请求进行重定向,最终返回某一个边缘服务器的 IP 地址。具体过程如下:用户首先向本地 DNS 服务器发起一个域名查询,如果查询在 DNS Cache 中命中,则直接返回相应的 IP 地址。否则,本地 DNS 会通过迭代或者递归的方式,向上级被授权的 DNS 服务器继续提交查询,请求最终会被提交到此内容所在域授权的某一个名字服务器,此服务器基于查询请求的来源进行决策,并返回所选取的边缘服务器 IP 地址。

1.2 边缘服务器的选取问题

大多数 DNS 服务器的实现都缺省支持轮询(Round-Robin)机制,接收查询时,在 IP 地址集合中随机选取一个条目应答,从而可被用作负载均衡,但这种方式达不到路由请求

收稿日期:2004 - 07 - 29;**修订日期**:2004 - 10 - 09 **基金项目**:国家自然科学基金资助项目(90104011)

作者简介:胡鹏(1981 -),男,安徽固镇人,博士研究生,主要研究方向:无线 Ad Hoc 网络、网络安全、信息通信网; 洪佩琳(1961 -),女,浙江宁波人,教授,博士生导师,主要研究方向:网络通信和信息安全; 李津生(1937 -),男,福建厦门人,教授,博士生导师,主要研究方向:下一代网络体系结构; 杨海松(1976 -),男,河南内黄人,博士研究生,主要研究方向:网络安全、策略管理、服务质量.

系统所预期的目标。有些 CDN 通过构建特殊的 DNS 服务器,在域名查询到达的时候,或者定期对各个边缘服务器的状态进行探测,根据某些判别规则,例如连接数目或包丢失率最小,选取最佳的边缘服务器[4]。这种方法综合考虑所有边缘服务器的状态,在一定程度上提高了 CDN 的性能,但是依然存在缺点:DNS 服务器查询各个服务器状态采取的是请求应答模式,边缘服务器不具有主动性,例如遭受突发业务流的时候,不能及时通知 DNS 服务器;可扩展性不好,仅使用一个特殊的 DNS 服务器负责收集和处理众多信息,还要应答查询请求,尤其随着规模的扩大,DNS 服务器本身就成为性能的瓶颈。同时,功能过分集中给安全带来威胁,不利于有效管理。

2 基于策略的路由请求系统

为了解决上述问题,使 CDN 路由请求系统更具智能化,下面针对 DNS 服务器引入策略管理的思想,提出基于策略控制的路由请求系统——PBR²S。

2.1 策略控制

策略是操作、管理和控制网络资源权限的一套规则[5]。根据 IETF 策略框架工作组(Policy Framework Working Group)的规定,策略框架包括 4 个组件,如图 1 所示。其中,策略管理工具(Policy Management Tool,PMT)负责定义、更新和选择性地监测策略规则的部署情况。策略库(Policy Repository,PR)被用来存取策略规则。策略实施点(Policy Enforcement Point,PEP)表示受策略规则控制和实施策略的实体。策略决策点(Policy Decision Point,PDP)负责维护一致性,从策略库取出策略规则,并对其进行翻译,以及判决对策略实施点采取何种策略集合。

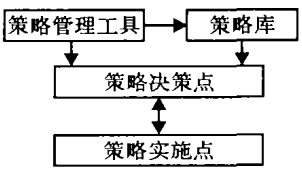


图 1 IETF 定义的策略框架

策略的管理主要优点在于提高管理系统的可扩展性和灵活性。经过配置的策略可以动态更新,这使得在不中断系统操作的前提之下,能够改变系统的行为。

2.2 架构描述

CDN 内容提供商可以将整套内容植入每个边缘服务器,也可以选择复制部分内容到不同的边缘服务器。研究表明,将同一套内容分为几个部分,从而每个边缘服务器传送内容的一个子集给用户所获得的效率并不好[6],所以 PBR²S 假设前提是内容已经预先被完整分发(full distributed)到各个边缘服务器。

假设用户网络由多个自治系统组成,PBR²S 预先以一个或者几个相近的自治系统为单位,在其内部部署一组边缘服务器和分布 DNS 服务器,它们负责本自治系统内部的内容服务。这样,用户网络中广泛地分布着这些 CDN 的组件,如图 2 的下半部分所示。

分布 DNS 服务器需要预先被设置为服务内容的域名所授权的名字服务器。这样通过对自治系统内 DNS 服务器进行适当的配置,可以保证用户对服务内容的 DNS 查询请求最终都会被提交到对应的分布 DNS 服务器。由于每个分布 DNS 服务器的作用范围都覆盖一个或者多个距离较近的自治系统,用户网络中对服务内容的所有请求都会被提交到 PBR²S 架构的分布 DNS 服务器。再由其重定向请求到某个边缘服务器(Edge Server,ES),缺省是与分布 DNS 服务器一组的边缘服务器,实施的时候通过 PDP 动态选取。

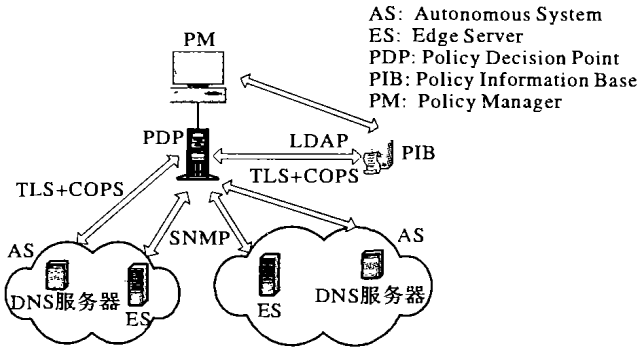


图 2 PBR²S 的基本架构

如图 2 所示,PBR²S 引入的策略组件包括全局唯一的策略管理平台(PM)、策略决策点(PDP),以及分布在各个相对独立的自治系统中的策略实施点(PEP)。策略框架的标准组件在 PBR²S 中被赋予了具体化的功能。策略管理平台负责编辑和向下分发分布 DNS 服务器的配置策略,它允许管理员预先定义初始化的策略,存储在策略信息库中。分布 DNS 服务器的配置策略就是按照系统的分布情况,为每个分布 DNS 服务器配置和更新资源记录。策略决策点是整个 PBR²S 架构的关键部件,主要功能有两个:1)如同传统的策略决策点一样,负责策略自上而下的翻译和分发、一致性维护等相关工作;2)根据扩展的 policy-extension-by-policy[7],结合边缘服务器汇报上来的系统和网络监测信息,实时、动态地生成策略,对 DNS 服务器的资源记录进行适应性的更新。这个功能与传统 CDN 网络有着本质的区别,其显著优点是有效地利用了反馈的机制,优化了用户对路由信息的自动化获取。策略实施点,即分布 DNS 服务器,是实施配置更新策略的最终实体。实体中包含策略代理(Policy Agent,PA),并由它调用 DNS 引擎对外提供的接口方法(Interface Method)更新资源记录。

2.3 协议体系

PBR²S 所包含的协议都是标准化的。这一系列协议包括传输层安全 TLS(Transport Layer Security),在 PBR²S 中它被用作策略实施点与策略决策点进行 COPS(Common Open Policy Service)协议通信的承载协议,为其提供安全可靠的信道。另外,系统还使用简单网络管理协议(SNMP)在边缘服务器和策略决策点之间进行信息交互。策略信息库与策略决策点按照标准通过 LDAP(Lightweight Directory Access Protocol)协议进行读写操作。

1) TLS 协议在两个应用的通信过程中提供私密性和数据完整性保证。

2) SNMPv3 是由 IETF SNMP Version3 工作组标准化的协议体系,是简单网络管理协议的第三个版本,在安全性和管理能力方面对前两个版本进行了扩展,提供了鉴别、加密和访问控制规则,使其更具有实际应用价值。在 PBR²S 中使用

SNMP 的目的在于通过对管理信息库(MIB)的读写操作进行网络监测和管理。并同时监测系统负载信息，为下一步动态生成策略提供参考依据。

3) COPS 协议描述了一个支持策略控制的客户/服务器模型，用于在策略决策点及其策略实施点之间传输策略信息，具有可扩展性。COPS 协议本身已经具有相当的安全性，可以提供消息级别的认证、抗重播和完整性校验。为了保证下达的策略不被窃取和维护网络整体的安全可靠，PBR²S 使用 TLS 协议对 COPS 协议进行加密。

4) LDAP 协议按照标准规定用于策略决策点和策略信息库之间的通信。

5) 相关标准规定了 DNS 动态更新的方法[8]，在此基础上，可以动态修改 DNS 的资源记录(Resource Record)。需要注意的是，策略决策点要具有修改各 DNS 的权限。为了防止 DNS Cache Poisoning 等攻击的发生，在策略代理部署策略的时候，应按照安全规范严格实施[9]。

2.4　功能模块分析

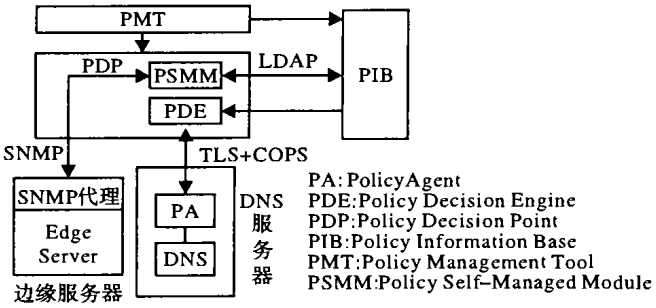
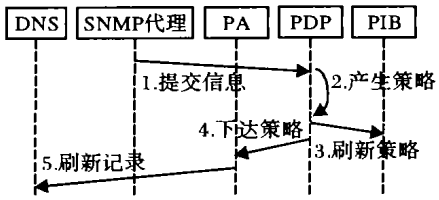


图 3　组件功能和模块



1.SNMP代理向PDP提交网络信息　2.PDP动态生成策略
3.刷新策略信息库PIB
4.PDP向策略代理下达策略，用TLS加载COPS协议，保证安全
5.策略代理通过DNS更新机制刷新分布DNS服务器的资源记录

图 4　组件间通信的规程

图 3 给出了各组件的功能描述，边缘服务器中的 SNMP 代理的主要作用是通过简单网管协议与 PDP 中的 PSMM 通信。DNS 中的 PA 是一个策略代理，实际部署时集成在分布 DNS 设备上，这样有效地利用了负载较轻的分布 DNS 服务器，并且提高安全性。PDP 中的策略决策引擎(PDE)向 PA 下达策略，该策略可以使用 TLS 直接加载 COPS 协议来传送。此处 PDP 不仅负责从 PIB 处获取新的策略，也根据事件参与产生新的策略。策略自管理模块(Policy Self-Managed Module，PSMM)负责接收从 SNMP 代理递交上来的网络和系统状态信息，并依据 policy-extension-by-policy 的机制，由 PIB 中定义的特殊策略，即策略之策略，动态地产生新的策略，及时更新策略信息库 PIB。此后通过策略决策引擎下达到策略代理，最后作用于分布 DNS 服务器。策略管理工具 PMT 是管理员进行管理的操作界面。

图 4 给出了 PBR²S 组件间通信的规程举例，在该例子中，PDP 依据 SNMP 代理实时汇报的负载和网络信息，动态生成策略并且及时刷新(保持一致性)。

2.5　安全分析

系统本身的安全性保障在上面已经论述，这里主要分析 PBR²S 应对异常业务的处理能力。根据其设计框架，PBR²S 引入了实时监控的机制，所以能够对网络状况做出定制的动态响应。尤其是某一局部出现大规模突发业务的时候，系统的 PSMM 模块能够适应性地产生新的策略，通过动态域名更新方法[8,9]，系统动态修改分布 DNS 域名服务器的 A 记录，把域名与其他边缘服务器 IP 地址关联起来，从而将业务导向其他合适的边缘服务器。

PSMM 在其具体实施时可以引入或改进各种成熟的调度策略，这些策略与 PBR²S 平台独立，且可以根据网络环境变化而修改生成策略的规则。在大规模突发业务出现时，可以采用文献[5]的方法，将边缘服务器的负载划分高低两个门限，当负载超出高门限时，采取适当调整算法，动态修改为其他某个边缘服务器的 IP 地址，在负载下降到低门限之下时，恢复其缺省 IP 地址。其他边缘服务器的选取可以基于响应时间、负载等因素灵活考虑，这也是 PBR²S 架构的灵活性的体现。

2.6　仿真分析

这里采用仿真度量系统性能。我们将边缘服务器放置在 7 个点，点与点之间的网络延迟通过网络业务实测得出。同时采用高负载的随机分布模型对网络业务建模，在此基础上，将调度策略分别设置为路径优先、负载优先和全局优先(路径与负载结合)。比较这三种情况下的服务平均响应时间，如图 5 所示。

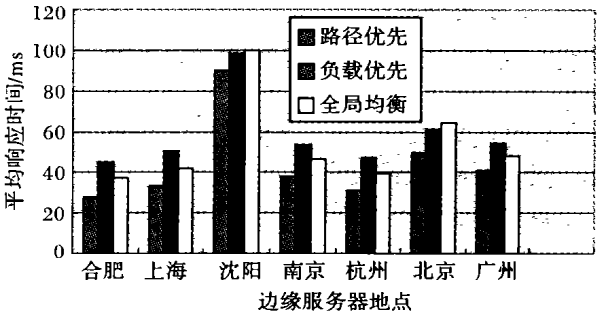


图 5　各个边缘服务器的平均响应时间

3　性能评估

首先，PBR²S 具有良好的可操作性和有效性，它在现有的 CDN 架构上，根据管理的需要添加了策略框架层。在架构中，引入的分布 DNS 服务器，相比一般 DNS 重定向而言，减少了 DNS 查询的轮回，降低了延迟。而且，随着策略框架的引入，PBR²S 实现了将最合适的边缘服务器地址动态地“推”到分布 DNS 服务器，以备随时查询，提高了响应速度。

其次，PBR²S 是一个安全的系统。从支撑系统的各个安全协议，到系统中某些组件的集成化，都有效地增加了系统的安全性。边缘服务器在全局范畴上的联合协作，增强了其稳健性，具有一定的抗 DDoS 攻击的能力。

(下转第 269 页)

4 结语

TCP 协议是针对固定有线网络设计的,在移动 Ad Hoc 网络中,经常发生信道误码或网络分割等情况。TCP 协议不能正确区分和处理这些情况造成的丢包,因此传输性能急剧下降。为提高 TCP 的传输性能,近年来提出了一些 Ad Hoc 网络中 TCP 的改进方案,主要分为两类:基于网络反馈的方案和"端到端"的方案。

Ad Hoc 网络未来将广泛应用于个人通信领域,需要使一台普通电脑(如笔记本)能够随心所欲的加入到 Ad Hoc 网络中。"端到端"的改进方案不需要中间节点的支持,将更有利于实现这一点。

从另一个角度考虑,在军事领域或其他特殊领域,Ad Hoc 网络是作为边缘网络存在,网络中的主机要面临能源和带宽受限等诸多问题。此时更准确的获取网络状况,提高 TCP 协议的传输性能更为关键,因此网络反馈的方案更适合这些领域。

需要指出的是,在移动 Ad Hoc 网络中,TCP 传输性能的提高还不仅仅依赖于对 TCP 协议本身的改进。例如,在 MAC 层的机制中,隐终端和暴露终端问题,以及带宽不对称问题,都在很大程度上影响了 TCP 协议的性能;路由层中,研究设计出稳定且健壮的 Ad Hoc 网络路由协议,也对 TCP 的性能起着至关重要的作用。

参考文献:

[1] CORSON S, MACKER J. Mobile Ad hoc Networking (MANET): Routing Protocol Performance Issues and Evaluation Considerations [S]. RFC2501, 1999.

[2] JOHNSON DB, MALTZ DA. The Dynamic Source Routing Protocol for Mobile Ad Hoc Networks [EB/OL]. Internet-Draft, draft-ietf-manet-dsr-10.txt, 2004.

[3] PERKINS C, BELDING-ROYER E, DAS S. Ad hoc On-Demand Distance Vector (AODV) Routing[S]. RFC3561, 2003.

[4] SEMERIA C. Supporting Differentiated Service Classes: TCP congestion Control Mechanisms [Z]. Sunnyvale: Juniper Networks Inc, 2002.

[5] LEINEN S. ECN (Explicit Congestion Notification) in TCP / IP [EB/OL]. http://www.icir.org/floyd/ecn.html, 2004.

[6] HOLLAND G, VAIDYA N. Analysis of TCP Performance over Mobile Ad Hoc Networks [Z]. MOBICOM'99, 1999.

[7] LIU J, SINGH S. ATCP: TCP for Mobile Ad Hoc Networks [J]. IEEE Journal, 2001, 19(7): 1300-1315.

[8] WANG F, ZHANG YG. Improving TCP Performance over Mobile Ad Hoc Networks with Out-of-Order Detection and Response [Z]. Mobihoc'02, 2002.

[9] ZHENG HF, GREENSTEIN B, MENG XQ. Design and Implementation of a TCP-Friendly Transport Protocol for Ad Hoc Wireless Networks [Z]. ICNP'02, 2002.

(上接第 247 页)

PBR²S 的优点还体现在其灵活性方面。一方面,基于策略管理的特点就是灵活性;另外,其分布 DNS 服务器具有可重用性,可以被其他 CDN 重复使用,从技术角度来看,这样有利于灵活地组建更大规模的 CDI。

最后,PBR²S 还具有可扩展性。这里的可扩展性体现在三个方面:1)规模的可扩展性,采取 PBR²S 的 CDN 系统之间可以很方便地建立互联关系,通过共用的分布 DNS 服务器群,简化了将多个孤立的 CDN 扩展成 CDI 的过程。2)策略框架中层次结构的可扩展性,上述内容所提及的只有一个全局的 PDP,在实际的操作中可以设计多个 PDP,从而增强系统的层次感和规模。3)性能的可扩展性,若是将各边缘服务器进行集群化,并且在前端放置第 4/7 层交换机,则系统的稳定性、安全性以及服务的粒度将会得到显著的改善。

4 结语

我们对现有 CDN 路由请求机制进行分析的同时,提出了 PBR²S 架构。PBR²S 是一个基于 DNS 机制并包含策略框架的路由请求系统,该系统架构采取分布式的结构,标准化的协议,具有良好的安全性、灵活性和可扩展性,能够对 CDN 的路由请求进行合理有效的规划,最大限度地满足用户对内容的需求。

参考文献:

[1] JOHNSON KL, CARR JF, DAY M, *et al.* The measured performance of content distribution networks[J]. Computer Communications, 2001, 24: 202-206.

[2] DAY M, CAIN B, TOMLINSON G, *et al.* RFC 3466, A Model for Content Internetworking (CDI)[S], February 2003.

[3] BARBIR A, CAIN B, DOUGLIS F, *et al.* RFC 3568, Known CN Request-Routing Mechanisms[S], July 2003.

[4] BILIRIS A, CRANOR C, DOUGLIS F, *et al.* CDN Brokering[A]. AT&T, In Proceedings of the 6th International Workshop on Web Caching and Content Distribution[C]. Boston, MA, June, 2001.

[5] WESTERINEN A. RFC 3198, Terminology for Policy-Based Management[S], November 2001.

[6] KANGASHARJU J, ROSS KW, ROBERTS JW. Performance Evaluation of Redirection Schemes in Content Distribution Networks[A]. 5th Web Caching Workshop[C]. Lisbon, Portugal, May 2000.

[7] KANADA Y. Dynamically Extensible Policy Server and Agent[A]. 3rd International Workshop on Policies for Distributed Systems and Networks (Policy 2002)[C], June 2002.

[8] VIXIE P, THOMSON S, REKHTER Y, *et al.* RFC 2136, Dynamic Updates in the Domain Name System (DNS UPDATE)[S], April 1997.

[9] WELLINGTON B. RFC 3007, Secure Domain Name System (DNS) Dynamic Update[S], November 2000.

[10] JUNG J, KRISHNAMURTHY B, RABINOVICH M, *et al.* Flash Crowds and Denial of Service Attacks: Characterization and Implications for CDNs and Web Sites[A]. In Proceedings of the International World Wide Web Conference[C], May 2002.