

CDN 技术在广东省地震信息网中的应用研究*

陈晓发^{1,2,3}, 王挺^{1,2,3}

(1. 广东省地震局, 广东 广州 510070; 2. 中国地震局地震监测与减灾技术重点实验室, 广东 广州 510070;
3. 广东省地震预警与重大工程安全诊断重点实验室(筹), 广东 广州 510070)

摘要:在对 CDN 原理和广东省地震信息网配置情况进行深入分析的基础上, 结合智能 DNS 解析和内容缓存技术, 实现了广东省地震信息网 CDN 全站加速的架构, 并对该系统应用前后网站访问速度进行测试。结果表明该方案大幅提升了网站的访问速度和负载能力。最后, 结合河源 4.8 级地震对网站的访问情况进行了统计分析, 从网站流量统计、访问概况及 WEB 服务器负载 3 个方面进一步验证了系统的可靠性和实用性。

关键词:智能域名解析; 内容分发网络; 缓存技术; 全站加速

中图分类号: P315-392

文献标识码: A

文章编号: 1000-0666(2012)04-0583-06

0 引言

目前, 广东省地震信息网采用电信专线接入, 由于各网络运营商之间相互连接的可用带宽较小, 非电信接入用户访问广东省地震信息网的速度较慢, 且在省内发生有感地震等突发事件引发网站访问量激增的情况下, 容易造成网站源服务器负载过大而停止响应用户的请求, 从而影响对外的实时地震信息服务。日本“3·11”大地震后, 公众对即时地震信息的需求快速增加, 如何利用已有的资源又好又快地为公众提供即时的地震信息已经成为急需解决的问题。传统的双域名服务器镜像解决方案, 存在着需要 IP 资源多、镜像同步管理难、域名较多、不易于推广使用等问题。鉴于以上方案的不足, 笔者提出一种采用智能 DNS 解析加内容缓存技术的广东省地震信息网 CDN 全站加速解决方案, 充分利用已有的资源配置来解决网站的访问瓶颈。

1 CDN 技术概述

内容分发网络(Content Delivery Network, 简称 CDN)(Pallis, Vakali, 2006; Hofmann, Beaumont, 2005; 徐锐, 2008), 其目的是通过在现有的 Internet 中增加一层新的网络架构, 将网站的内容发布到最接近用户的网络“边缘”, 使用户可以

就近获得所需的内容, 从而改变 Internet 网络拥挤的状况, 减少用户访问网站的响应时间, 从技术上全面解决由于网络带宽小、用户访问量大、网点分布不均匀等原因所造成的用户访问网站响应时间长的问題。

一个最基本的 CDN 系统主要由以下 2 部分构成:

(1) 智能 DNS 解析系统

普通的 DNS 服务器只承担为用户解析出域名所对应的 IP 记录, 而不判定用户所在位置, 这样会造成所有访问用户都只能解析到一个固定的 IP 地址上, 从而造成网站源服务器负载较大, 影响网站的访问速度。智能 DNS 系统颠覆了这个概念, 该系统会自动判定用户的位置, 从而进行一系列的智能化处理, 然后就近指定相应的 IP 返回给用户。

例如访问者是电信用户, 智能 DNS 解析系统会就近将域名对应的电信 IP 地址解析给这个访问用户。如果访问者是移动用户, 则就近返回域名对应的移动 IP 地址给访问用户, 减少了网站的响应时间。

(2) 内容缓存服务器

内容缓存服务器承担了对网站源服务器的静态页面和部分动态页面的请求, 减轻了网站源服务器的负载压力。它位于本地源服务器和因特网之间, 处理所有对网站源服务器的请求, 组织了网站源服务器和因特网的直接通信。如果网络用

* 收稿日期: 2012-04-24.

基金项目: 地震行业科研专项(201108010)资助.

户请求的页面在缓存服务器上有缓存,缓存服务器直接将缓存内容发送给用户。如果没有缓存则先返回源服务器获取数据,本地缓存后再发送给用户。这种方式有效的减少了向网站源服务器的请求数,从而降低了网站源服务器的负载,提高了网站的访问速度。CDN 系统的工作流程如图 1 所示。

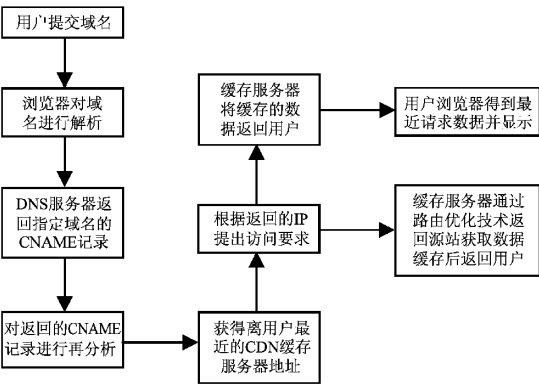


图 1 CDN 系统工作流程图

Fig. 1 The work flow chart of CDN system

2 广东省地震信息网 CDN 全站加速系统架构设计

根据现有的信息需求,采用 CDN 全站加速(WSA)方案(彭湘凯,2005;何波,2001;谢冬松,2006)对广东省地震信息网进行加速,将网站的静态内容长期缓存在边缘节点上,使访问者可以就近获取所需的内容,而动态内容因其具有不可缓存、内容敏感度高、容易受到攻击等特点,则通过网络路由优化、私有传输协议、数据压缩防攻击等优化技术来实现用户与源服务器快速的应用交互(熊明等,2005;赵问道,2004;牛从达,王峰,2004)。

智能 DNS 解析系统架构如图 2 所示,其具有以下 3 个特点:

(1) 系统能够为网站的访问者精准定位,细化粒度达到同一省份各个区域,根据最佳策略进行调度,为各个地区的访问者寻找最优节点,保证最优访问效果,提高整体访问质量;

(2) 系统以网络及设备状况的实时检测数据作为判断依据,而非静态数据,具有及时、准确的特点,通过对各分布式节点监控数据的实时采集及集中预警的方式,避免了错报、漏报的现象;

(3) 系统能够在节点或网络发生故障时自动采取措施,在 2~5 min 内完成最佳的调整方案,规避网络故障和波动,避免人工操作带来的风险和延时。

当用户访问广东省地震信息网时,智能 DNS 解析系统根据用户的 IP 就近指定相应的边缘服务节点来响应用户的请求,以达到分担用户的请求压力的效果,从而减轻网站源服务器的负载。

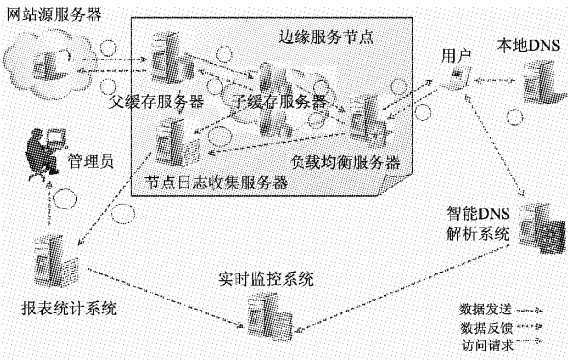


图 2 广东省地震信息网 CDN 系统架构

Fig. 2 The architecture of CDN system in Guangdong Earthquake Information Network

3 广东省地震信息网 CDN 全站加速系统的实现

根据 HTTPWATCH 抓包软件对广东地震信息网进行数据分析(图 3),可以发现 ASP 文件类型约占网站数据的 10%,HTML、CSS、JS 文件类型约占网站数据的 10%,GIF、JPG 文件类型约占

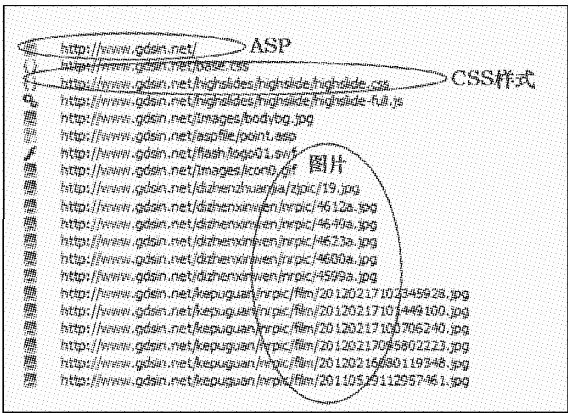


图 3 广东省地震信息网数据类型统计图

Fig. 3 Statistical charts data types in Guangdong Earthquake Information Network

网站数据的 80%，在这些数据类型中，通过 CDN 全站加速之后，使能够缓存的静态内容按规定的缓存策略进行缓存，例如 GIF、JPG 图片文件缓存 25 min，CSS 文件缓存 5 min，新闻页面文件缓存 1 min 等，而对于不能缓存即需要实时更新的内容（网站中包含最新地震事件的 ASP 页面文件）则以最优化的性能与源服务器进行交互，在提升网站响应速度的同时也保证了地震信息的时效性。

由于智能 DNS 解析系统是根据用户的接入点来就近指定相应的边缘节点响应用户请求，因此边缘节点的数量以及整体性能对于提升网站的响应速度就显得十分重要。本文所采用的 CDN 全站加速系统在广东省内边缘节点分布较为广泛，节点所属的网络运营商也涵盖了 5 大网络运营商且均匀分布，同时边缘节点通过负载均衡服务器对子缓存服务器进行了有效的负载均衡处理，能够在网站访问量激增时快速响应用户请求，为用户提供即时的地震信息。

CDN 全站加速系统在按照图 2 所示的架构配置完成后，还需进行相应的功能测试，检查网站的访问是否正常，内容是否正确等。通过设置本地机器 hosts 文件（文件目录为：C:\WINDOWS\system32\drivers\etc\），添加域名与 IP 的映射，跳过 DNS 解析步骤，使本地机器直接访问边缘节点。添加完成后通过 ping 命令来验证 host 设定是否生效。

测试完成后，如果网站访问正常，即可按照指定的 CNAME 记录进行别名切换操作，如果有 A 记录，还需在新增对应域名的 CNAME 记录后删除 A 记录。至此，广东省地震信息网的全站加速系统部署完成。

4 CDN 全站加速系统测试结果与分析

广东省地震信息网的全站加速系统采用国际上著名的网站性能检测服务商 Gomez 进行加速前后的对比测试。未加速前测试时间为一天，测试频率为每小时一次，加速后也采用同样的测试方式。Gomez 记录的网站加速前后的 Lastmile 测试（Lastmile 测试指模拟最终用户访问进行的测试）总体性能如图 4、图 5 所示。

对比图 4 和图 5 可以得出，加速后网站的访问

速度明显提升。通过表 1 可以得出，加速后网站的平均响应时间较加速前减少了 66.5%，访问成功率较加速前提升了 17.3%。因测试区域选定为全国，在省外节点分布较疏的地区由于网络故障等因素会出现数据丢包的情况，从而在一定程度上影响网站访问的成功率。另外，通过对加速后网站的访问概况统计图进行分析（图 6），可以得出只有 34.89% 的动态数据需要回源服务器进行实时响应，边缘节点有效地分流了源服务器的请求数，使得网站已有的 30 MB 带宽即可满足突发访问量的带宽需求。

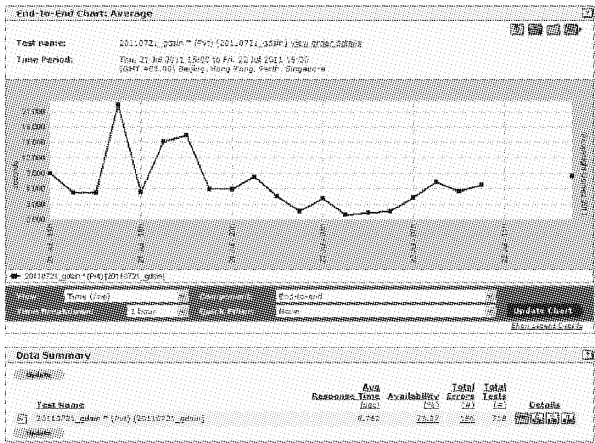


图 4 加速前 Lastmile 测试总体性能图

Fig. 4 Overall performance diagram of Lastmile test before the acceleration of Guangdong Earthquake Information Network

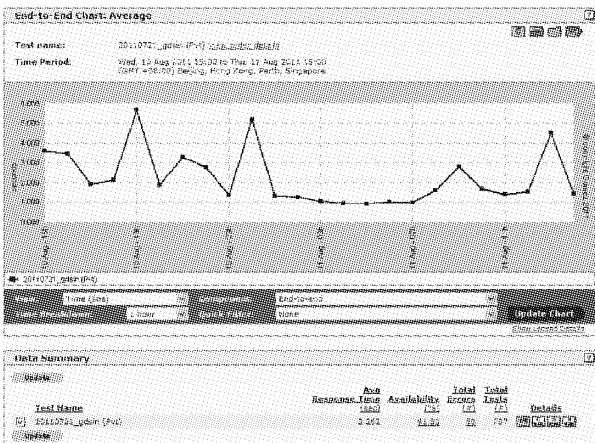


图 5 加速后 Lastmile 测试总体性能图

Fig. 5 Overall performance diagram of Lastmile test after the acceleration of Guangdong Earthquake Information Network

表 1 加速前后 Lastmile 测试结果对比

Tab. 1 Comparison of Lastmile test results before and after the acceleration of the system

	平均响应时间/s	访问成功率 (%)
加速前	6. 762	78. 27
加速后	2. 262	91. 80

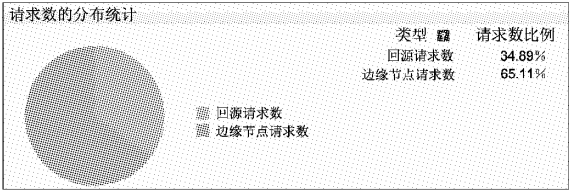


图 6 加速后网站的访问概况统计图

Fig. 6 Statistical charts of visiting the site after the acceleration of Guangdong Earthquake Information Network

5 案例分析

2012 年 2 月 16 日 2 时 34 分，广东省河源市东源县发生 4.8 级地震，地震发生后，广东省地震信息网实时发布相关地震信息，网站访问数也随之不断攀升，全天累计请求数超过十万人次。CDN 全站加速系统在负载激增的情况下，及时通过其庞大的边缘节点就近响应用户的请求，有效地分担了请求压力，极大程度上减轻了源服务器的负载，使得对外地震信息服务始终保持畅通。

5.1 网站流量统计分析

通过对网站全天访问量流量类型以及用户所使用的 ISP 类型进行统计分析，可以得出，边缘节点承担了大部分的流量，占总流量的 85%，仅有 15% 的流量指向服务器（图 7）；用户所使用的 ISP

中中国电信居多，约占 66.21%，其次是中国联通，约占 20.1%（表 2）。

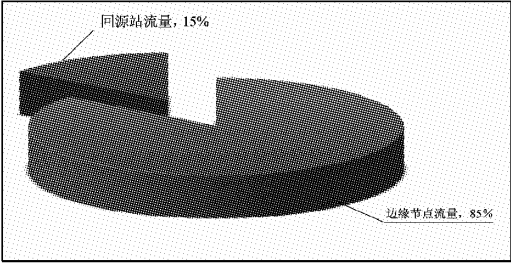


图 7 网站流量类型比例图

Fig. 7 Scale chart of the site throughput types

表 2 用户所使用的各 ISP 比例

Tab. 2 The proportion of different ISP by different users

ISP 类型	用户所使用各 ISP 比例 (%)
中国电信	66. 21
中国联通	20. 1
中国铁通	3. 39
中国移动	2. 63
中国教育网	2. 00
其它 ISP	5. 67

5.2 网站访问概况

通过对网站全天请求数进行统计分析（图 8），可以得出，边缘节点请求数占总请求数的 87.72%，回源的请求数仅占 12.28%，边缘节点承担了大部分用户的请求压力。分时段对网站请求数进行分析可以得出，网站请求数峰值出现在 12 时至 13 时间段内，累计总请求数为 100 885；对请求数按区域进行统计可以得出（表 3），省内用户的请求数占 92.75%，其中广州的请求数最多，占省内用户请求数的 58.94%，其次是河源，占 22.00%。

表 3 省内用户请求数按地市统计比例

Tab. 3 Statistical proportion of the users request number in Guangdong Province by the cities

地市	广州	河源	惠州	深圳	韶关	佛山	东莞	珠海	江门	阳江	中山
占省内请求数比例 (%)	58. 94	22. 00	3. 30	2. 34	1. 57	1. 27	0. 88	0. 59	0. 52	0. 23	0. 20
地市	清远	梅州	潮州	揭阳	汕尾	顺德	汕头	肇庆	湛江	茂名	云浮
占省内请求数比例 (%)	0. 18	0. 15	0. 14	0. 12	0. 09	0. 09	0. 07	0. 07	0	0	0

5.3 WEB 服务器负载情况分析

通过对 WEB 服务器 CPU、内存以及带宽等使用率的实时监控数据进行分析（表 4），可以得出，

WEB 服务器的 CPU 最大使用率仅为 11%，温度最高为 46°，内存最大使用率为 27%，带宽使用率也维持在较低水平，整体性能始终保持稳定。

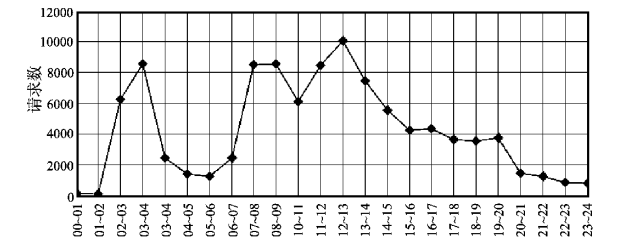


图8 2月16日全天网站请求数按时间段统计图
Fig. 3 Statistical charts of the site request number on Feb. 16 all day long by the time

表4 WEB 服务器 CPU、内存使用率以及 CPU 温度统计表
Tab. 4 The statistical table of the usage of CPU of WEB server, usage rate of memory and temperature of CPU

时间点	性能指标		
	CPU 使用率(%)	内存使用率(%)	CPU 温度/℃
0 点	1	14	36
2 点	1	16	37
4 点	11	27	46
6 点	6	19	38
8 点	7	20	39
10 点	11	26	44
12 点	10	25	42
14 点	9	22	41
16 点	8	20	40
18 点	8	20	40
20 点	8	21	39
22 点	6	18	38
24 点	6	18	37

在 CDN 全站加速系统的支撑下，广东省地震信息网在应对突发地震事件时的对外信息服务始终保持稳定，为公众提供了实时的地震信息服务，同时也间接的为内部开展应急处置工作提供了必要的网络带宽保障，确保了应急处置工作的顺利进行。

6 结语

结合 CDN 技术对广东省地震信息网进行全站加速，提升了网站的访问速度和负载能力，为公众提供即时稳定的地震信息服务。网站采用 CDN 全站加速后主要解决了以下 5 方面的突出问题：

(1) 当省内发生有感地震导致网站访问量激

增时，网站源服务器的负载被有效地分流到各个边缘节点，从而消除了源服务器负载过大而导致服务中断的影响。

(2) 极大地缓解了不同网络运营商之间互联的瓶颈造成网站访问较慢的影响，提高了非电信用户对网站的访问速度。

(3) 采用多源部署策略、离线工作模式等，避免单点故障导致网站无法正常服务。

(4) 网站源服务器没有直接暴露在公网，减少了被攻击的风险，采用了源站保护、防盗链、防篡改、防 d. d. o. s 攻击和 DNS 安全保护等多种技术（刘云华等，2011），提高了系统的安全性。

(5) 系统提供完整的访问日志，便于管理员分析访问者行为，并可实现统计报表的输出。本文所阐述的方案以网站点击数进行计费，计费单位为百万次，通过对网站带宽使用情况和访问量进行综合分析，采用基础配置（应对一般的访问需求）加预备配置（应对震时激增的访问需求）的方式即可确保网站的稳定运行，整体投入不大，实用性较高，便于在行业内部应用。该方案自 2011 年 8 月 12 日部署于广东省地震信息网以来，确保了多次地震信息发布工作以及网络知识竞赛等活动的顺利进行，达到了预期的效果。

参考文献：

何波. 2001. CDS 技术及其应用[J]. 电信技术,(11):7-8.
刘云华,六治,单新建,等. 2011. 地震信息网自适应网络安全技术模型研究[J]. 地震研究,34(1):97-99.
牛从达,王峰. 2004. VPN 技术及其在地震信息广域网建设中的运用[J]. 地震研究,27(4):380-383.
彭湘凯. 2005. CDN 网络及其应用[J]. 微计算机信息,(2):150-151.
谢冬松. 2006. 关于应用 CDN 技术改善远程教育网络服务的探讨[J]. 教育信息化,(1):76-77.
熊明,赵政,赵怪娃. 2005. 城域网上 CDN 技术的应用[J]. 计算机应用,25(1):196-198.
徐悦. 2008. 构建大型 CDN 网络的关键技术研究[D]. 北京:北京邮电大学.
赵问道. 2004. 基于网络拓扑的 CDN 内容路由技术研究[J]. 浙江大学学报,38(4):414-419.
Hofmann M, Beaumont L R. 2005. Content networking: architecture, protocols, and practice [M]. San Francisco, CA, USA: Morgan Kaufmann Publishers.
Pallis G, Vakali A. 2006. Insight and perspectives for content delivery networks[J]. Commun ACM,49(1):101-106.

Application Research of CDN Technology in Guangdong Seismic Information Center

CHEN Xiao-fa^{1,2,3}, WANG Ting^{1,2,3}

(1. *Earthquake Administration of Guangdong Province, Guangzhou 510070, Guangdong, China*)

(2. *Key Laboratory of Earthquake Monitoring and Disaster Mitigation Technology, CEA, Guangzhou 510070, Guangdong, China*)

(3. *Key Laboratory of Guangdong Province Earthquake Early Warning and Safety Diagnosis
of Major Projects, Guangzhou 510070, Guangdong, China*)

Abstract

Based on in-depth analyzing principle of Content Delivery Network (CDN) technology and figuration of Guangdong Seismic Information Network (GSIN), we build the acceleration architecture of whole site in CDN for GSIN by techniques of intelligent DNS resolution and content cache. Through the speeds test of access to the web site before and after the application of the system, and the test results show that it significantly improved the speed of access to the site and site load capacity. Finally, we make a statistical analysis of access to the GSIN combined with Heyuan *M*4.8 earthquake, which verify the rialable and realize of system from the three aspects such as traffic statistics, access overview and load of website server.

Key words: intelligent DNS resolution; CDN; cache technology; whole site acceleration