一种基于多链路的服务加速方案

钱 杰, 崔 建, 张 蓓
(北京大学计算中心, 北京 100871)

摘要: 提出了一种多链路提供服务的方案, 在不改变服务器配置、已有路由表及内网物理拓扑的基础上, 通过 DNS把慢速客户端定向到服务器在补充链路的虚地址上, 再经过 NAT设备对虚地址进行转换, 实现了内部服务器的多链路访问, 解决了单链路服务拥塞的问题.

关键词: 多链路; DNAT; 服务器; DNS
中图分类号: TP 393 文献标识码: A 文章编号: 0438-0479(2007)S2-0129-03

由于用户对带宽、速度要求的提高, 企业网单一链路接入方式有时无法满足业务的需求, 有些企业采取了多链路接入的方式, 通过租用不同 ISP的链路, 用路由和 NAT技术对流量分流, 很好地解决了本地访问 INTERNET速度的瓶颈, 也就是解决了从里到外访问速度的问题. 本文则是在上述多链路的基础上, 通过利用 DNS和 DNAT技术, 让企业网外慢速客户端通过选择不同链路的方式, 来加速内网对外服务的速度, 即解决从外到里访问速度的问题.

1 方案设计

1.1 多链路接入拓扑结构

多链路接入一般是在企业网原来的边界即总出入口路由器的前端, 再部署一条或多条链路, 原来链路为本地地址全球路由链路, 为主链路; 而其它多条链路则需中间的 NAT地址转换设备, 把数据流中本地地址与其它链路地址相互转换后才能完成通信, 称为补充链路. 网络拓扑大致如图 1所示.

企业提高内网访问外网速度一般的做法是在企业边界路由器 R1上对目的地址做策略路由, 把主链路访问慢的地址块路由重新定向到补充链路上, 经 NAT转换设备对出的 Request数据报文做源地址替换和对进的 Reply数据报文做目的地址转换, 再结合对应的端口进行数据包的转发, 即 SNAT工作方式, 来达到本地网加速的目的. 但这只是单向的加速, 而对那些向慢速外网客户提供服务的内网服务系统来讲, 实现同样加速的目的, 则需用到下面介绍的 DNAT技术.

收稿日期: 2007-08-20
作者简介: 钱杰, 男, 工程师.
Email: qj@pku.edu.cn

1.2 DNAT工作机理

外网客户端从补充链路访问内网服务器的必要条件是服务主机需具备补充链路上的一个虚拟路由地址, 而中间的 NAT地址转换设备实质就是一代理服务器, 有两个逻辑地址, 对外的虚拟服务器地址和对内的虚拟客户端本地地址. 外网客户端向服务器的虚拟地址发 Request数据报文, 路由到 NAT转换设备时, 数据报文目的地址被修改为服务器本地真实地址, 报文源地址被替换为 NAT设备本地地址来虚拟外网客户端, 然后结合 NAT内部端口对应关系, 数据包被转发到内网服务器上, 这个过程相当于 NAT转换设备代理外部客户端向内网服务器发出请求; 同理, 服务器回应的 Reply数据报文必定被先发到 NAT转换设备上, 数据报文目的地址修改为外部客户端地址, 报文源地址则被替换为服务器虚拟地址, 再结合 NAT内部保存的端口对应关系, Reply数据包最终返回到实际的客户端, 此过程相当于 NAT转换设备代理内部服务器回应外网客户端的请求.

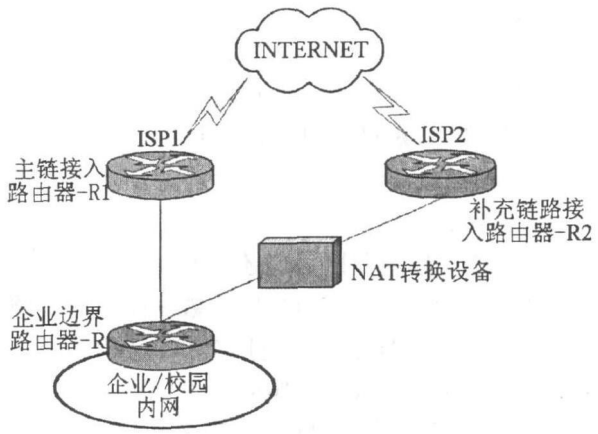

图 1 多链路接入网络拓扑结构

1.3 基于DNS的分流策略

基于多链路的内网服务器需在补充链路上具有一个虚拟路由地址,而客户端是通过名字解析DNS得到服务器地址的,因此对客户端分流的控制完全可以基于DNS来实现.

由于Bind的DNS被普遍使用,以下是基于其版本9.4来讨论的.Bind 9支持View视图的功能,也就是通过划分不同的视图,使得不同视图里的客户端看到的域名空间是不同的,例如,

主机名www.pku.edu.cn,子网162.105/16用户解析到的IP是162.105.129.12,而其它非162.105/16子网的用户看到的IP却是202.112.7.40,即子域pku.edu.cn存在两套域名空间,分别对应不同View视图的客户端,类似如下的配置:

```
named.conf
  view "internal" {
    matchr-clients { 162.105/16; };
    //对应子网 162.105/16集合
    zone "pku.edu.cn" in {
      type master;
      file "db.internal";
      //视图 internal域名空间
    };
};
view "external" {
    matchr-clients { any; }
    //所有地址空间
    zone "pku.edu.cn" in {
      type master;
      file "db.external";
      //视图 external域名空间
    };
};
db.internal
    www.pku.edu.cn IN A 162.105.129.12
db.external
    www.pku.edu.cn IN A 202.112.7.40
```

其中,子域pku.edu.cn被划分为internal和external两个视图,分别对应db.internal和db.external数据库;View中的match-clients是专门定义本视图IP地址集合的指令.

通常中大型企业网可能还有Slave辅助名字服务器,由于存在多个名字空间Zone数据同步的问题,所以主辅名字服务器之间需使用TSIG区分不同View视图Zone数据的交换,以保证不同View视图数据库正确的同步,下面是辅助名字服务系统named.conf的配置:

```
主服务器 Master 202.112.7.13
key "external" {
      //定义一个 key
      algorithm hmac-md5;
      //密钥生成算法
      secret "xxxxxxxx";
      //密钥 key
};
view "internal" {
    matchr-clients { ! key external; 162.105/16; };
    zone "pku.edu.cn" in {
      type master;
      file "db.internal";
    };
};
view "external" {
    matchr-clients { key external; any; };
    server 162.105.129.27 { keys external; };
    //辅服务器的 IP和使用的 key
    zone "pku.edu.cn" in {
      type master;
      file "db.external";
    };
};
辅服务器 Slave 162.105.129.27
key "external" {          //定义同样的 key
    algorithm hmac-md5;
    //密钥生成算法
    secret "xxxxxxxx";
    //密钥 key
};
view "internal" {
    match-clients { ! key external; 202.112.7.13; 162.
105/16; };
    zone "pku.edu.cn" in {
      type slave;
      file "db.internal";
      masters { 202.112.7.13; };
    //主名字服务器 IP
    };
};
view "external" {
    matchr-clients { key external; any; };
    server 202.112.7.13 { keys external; };
    //主服务器的 IP和使用的 key
    zone "pku.edu.cn" in {
      type slave;
      file "db.external";
      masters { 202.112.7.13; };
```

```
        //主名字服务器 IP
        };
    };
```

其中,关键部分是定义了一个主辅服务器通信的密钥 key "external",缺省无密钥同步的是 view internal 视图的 ZONE 区数据,带密钥同步的是 view external 视图的 ZONE 区数据,其实用了两个密钥,只是其中一个 key是无密钥的,完全可以使用两个有密钥的 key 进行不同视图 ZONE区数据的同步.

实际使用时还存在一个问题,例如子域 pku edu cn下登记有上百台主机名,而做多链路服务的主机毕竟只是少数的服务器,也就是 db internal与 db. external里的数据绝大多数是重复的,而且每次增删数据时,两个文件都要做修改,极容易出错,不利于维护,本文用了下述方法解决了此问题.

在本地根域下不做多视图的划分,而是另建一个子域来做,例如在根域 pku. edu cn下新建子域 sub pku edu cn专门登记那些做多链路服务的主机,这样子域数据的改动不会影响根域下全局的数据,但新问题是原来在根域下注册的名字无法沿袭使用,如 www. pku edu cn得改称 www. sub. pku edu cn,显然这是用户无法接受的,可通过修改根域数据库原来 A 记录条目为别名 CNAME 记录后就能巧妙解决此问题,例如:

原来在根域下的条目:

www. pku edu cn IN A 162. 105. 129. 12

修改为:

www. pku edu cn IN CNAME www. sub. pku edu cn

经这样处理后,付出的代价仅是客户端后台 DNS 程序解析 www. pku edu cn主机名时比原来多了一跳而已,且解析过程对用户也完全是透明的,优点是原来主机名得到了沿用,绝大多数服务器主机也无须改动配置.

2 结　论

利用 DNS在多链路上提供服务被业界广泛使用,但大多数为服务器多点部署的结构,目的是实现负载均衡,解决服务器负担重的问题. 但象校园或中小型企业网,服务器负载则相对较小,主要还是速度瓶颈问题,如外出的员工或放假的学生,当接入网络的地理位置发生变化时,由于速度慢引发访问校园/企业内网服务器故障的问题,本文就是针对此问题而设计的一种解决方案,可以基于已有多链路的基础上,无须更改已有网络拓扑和服务器配置,仅通过 DNS软件配置就能实现,具有简单可靠、易操作、投资小等诸多优点.

参考文献:

[1] Internet Systems Consortium. BIND9 Administrator Reference Manual. http://www.isc.org/.

[2] Martin A. Brown. Network Address Translation (NAT). http://linux-ip.net/html/ch-nat.html

[3] FAQ about BIND 9. http://www.isc.org/index.pl?/sw/bind/FAQ.php.

[4] RFC 1034-Domain names-concepts and facilities. http://www.faqs.org/rfcs/rfc1034.html

A Solution of Service Speedup Based on Multi-links Access

QIAN Jie, CUI Jian, ZHANG Bei

(Computer Center, Peking University, Beijing 100871, China)

Abstract A solution is introduced to speed service access based on mulitl-links. It neither need to alter the server config nor change route table or inner network topology. The principal technic is that, with DNS, the slow clients are redirected to Intranet server's supplementary link virtual IP, and then by NAT system, the packets are regulated to the right route way.

Key words multi-links; DNAT; Intranet server; DNS