

移动 IPv4/IPv6 虚拟机迁移系统设计

陈 俊^{1,2}, 陈孝威²

(1 贵州师范大学 教育科学学院, 贵州 贵阳 550001; 2 贵州大学 计算机科学与信息学院, 贵州 贵阳 550025)

摘 要: 为充分利用异构网络资源建设云计算基础平台, 满足云计算平台的资源需求, 文中设计了一种基于隧道技术、前缀管理、地址池管理协作及移动 IP 的 IPv4/IPv6 虚拟机迁移过渡框架。框架将传统 IPv4/IPv6 过渡技术与移动 IP 技术应用于云计算平台的虚拟机迁移, 利用过渡控制引擎作为核心与开发的 IPv4/IPv6 插件交互完成虚拟机迁移。经实验验证框架建立的网络结构可向客户端跨 IPv4/IPv6 网络提供云计算服务适用于 IPv4 虚拟机与 IPv6 虚拟机之间的无缝迁移。该框架可应用于 IPv4/IPv6 过渡期间云计算基础平台建设。

关键词: 云计算; 移动 IP; 虚拟机; 隧道技术

中图分类号: TP393

文献标识码: A

文章编号: 1000-7180(2011)09-0051-04

Virtual Machine Migration System Design Based on Mobile IPv4/IPv6

CHEN Jun^{1,2}, CHEN Xiao-wei²

(1 School of Education Science, Guizhou Normal University, Guiyang 550001, China;

2 College of Computer Science & Information, Guizhou University, Guiyang 550025, China)

Abstract: It can fully use heterogeneous network resources, provide resource requirement for cloud computing platform, this paper designed an IPv4/IPv6 virtual machine migration transition framework based on tunnel technology, prefix management, address pool management and mobile IP. The framework application tradition IPv4/IPv6 transition technology and mobile IP technology for virtual machine migration in cloud computing. It uses the paper designed cloud computing control engine as a core and IPv4/IPv6 plug-in collaboration completion virtual machine migration. Experiment shows the framework can provides IPv4/IPv6 cloud computing service for client and support IPv4/IPv6 virtual machine seamless migration. The framework can be applied to construction cloud computing platform in the IPv4/IPv6 transition period.

Key words: cloud computing; mobile IP; virtual machine; tunnel technology

1 引言

云计算平台可利用虚拟机在线迁移技术^[1]在线实现虚拟机和物理资源的重映射, 从而动态实现整个系统的负载均衡。IPv6 海量的地址空间为继分布式计算、并行计算和网格计算等模型之后, 学术界提出的云计算模型提供了发展空间。云计算平台需要海量的廉价设备共同构筑, IPv6 的地址空间满足了云计算平台的规模需求, 移动 IPv6 的提出^[2-3]对比移动 IPv4, 移动节点具备了更好的效率与可操作

性。IPv6 全面取代 IPv4 已是大势所趋, 但 IPv4 到 IPv6 的过渡不是在短时间内能完成的, 网络将会呈现出大量纯 IPv4 网络和纯 IPv6 网络互相交错的局面。

文中针对云计算平台设计提出了一种基于移动 IP 的虚拟机在线迁移系统。该迁移系统通过云计算控制引擎与特权虚拟机的交互并可跨 IPv4/IPv6 网络实现虚拟机透明在线迁移并可跨 IPv4/IPv6 网络给服务申请者通信对端提供云计算服务。该方案实现了 IPv4 向 IPv6 演进的情况。

收稿日期: 2010-12-07; **修回日期:** 2011-01-15

基金项目: 贵州大学研究生创新基金项目(省研理工 2010006)

2 相关技术介绍

2.1 MIPv4

如图 1 所示, 移动 IPv4 主要由移动节点(MN)、家乡代理(HA)、外地代理(FA)和通信伙伴节点(CN) 4 个实体组成. MN 在家乡网络拥有一个永久性的 IP 地址(HoA), 当 MN 移动到外地网络后将获得一个临时的关照地址(CoA). MN 随后需要同 HA 完成移动注册过程, MN 发送绑定更新(binding update, BU)消息向其 HA 通告 CoA, HA 则通过绑定确认(binding acknowledge, BA)消息响应此前的 BU. 当有 CN 与 MN 通信时, 由于此前不知道 MN 产生移动, 仍以 HoA 为目标地址发送数据包, 数据包被 MN 的 HA 截获, HA 以隧道方式把数据包转发给外地网络的外地代理 FA(第一种 CoA 地址)或移动节点自身(第二种 CoA 地址). 在第一种情况下, 外地代理再把数据包转发给移动节点. 数据包在不同子网间传送成功. 移动节点发往外地的数据包按一般的 IP 寻径方法送出, 不必通过家乡代理.

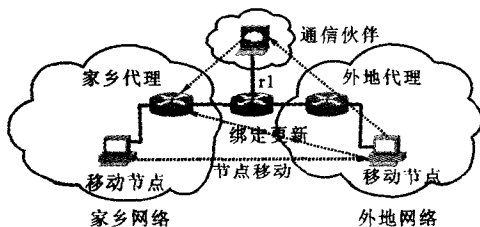


图 1 移动 IPv4

2.2 MIPv6

如图 2 所示, 移动 IPv6 不需要外地代理主要由移动节点(MN)、家乡代理(HA)和通信伙伴节点(CN) 3 个实体组成. 通信过程可参照移动 IPv4, 不同的是当 MN 移动到外地网络后通过邻居发现协议获得一个临时关照地址(care of address, CoA), 因移动 IPv6 无外地代理, 此 CoA 类似于移动 IPv4 的第二种 CoA. 且 MN 如发现了 CN 发来数据包经过了 HA 的中转, 则会向 CN 发送 BU 消息通告当前 CoA, 后续发往 MN 的数据包将会直接以 CoA 作为目标地址, 移动 IPv4 通信过程中三角路由得以消除.

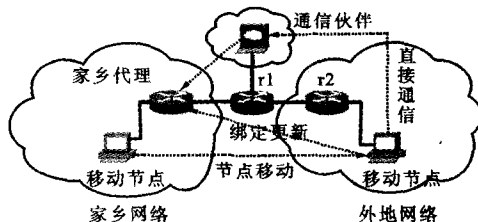


图 2 移动 IPv6

情况, 并依据负载均衡策略对通信对端指定虚拟机服务自适应的作出并帮助完成虚拟机迁移决策. 云计算控制引擎模型如图 3 所示. 在实际使用中为防止云计算控制引擎单点故障所造成的系统崩溃, 可增加一后备冗余云计算控制引擎. 当特权虚拟机多次与主云计算控制引擎交互失败后则激活冗余云计算控制引擎进行服务接管.

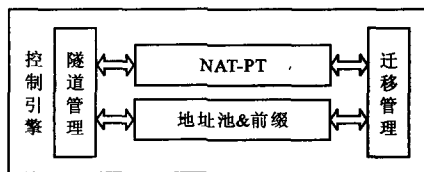


图 3 云计算控制引擎模型

3.2 方案设计与实现

适合演进前期、后期的迁移类型如图 4 所示. 首先建立一个 IPv4 网络和一个 IPv6 网络. 因为系统使用到隧道^[4]传送, r2、r4 需配置为双栈路由器. 然后, 在 IPv4 路由器与 IPv6 路由器之间部署文中设计的云计算控制引擎, 且通过该云计算控制引擎连接客户机通信对端. 当通信对端向云计算控制引擎申请服务时, 云计算控制引擎会依据策略指定相应虚拟机提供服务. 且云计算控制引擎还肩负着依据与特权虚拟机交互构造的资源使用视图作出虚拟机迁移决策并帮助完成在线迁移的功能.

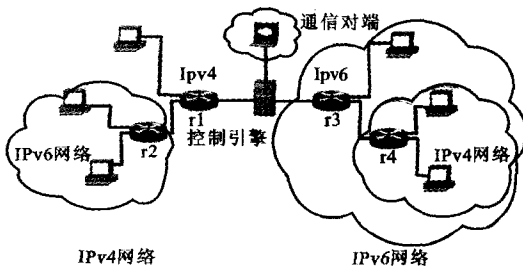


图 4 基于云计算的 IPv4/IPv6 虚拟机在线迁移系统

文中选择一个较为极端的情况进行讨论. 现今左端的被 IPv4 网络包围的 IPv6 网络中 IPv6 虚拟机向被 IPv6 网络包围的 IPv4 网络中的 IPv4 物理机进行虚拟机迁移.

现今被 IPv4 网络包围的 IPv6 网络中的待迁

3 系统设计

3.1 云计算控制引擎模型

云计算控制引擎通过和下文设计的特权虚拟机通信, 从整体上掌握云计算平台中物理资源的使用

IPv6 虚拟机为 sv , 被 IPv6 网络包围的 IPv4 网络中的目的 IPv4 物理机中迁移而来的 IPv6 虚拟机为 dv .

(1) sv 迁移后据移动 IPv6 协议请求地址, 该物理 IPv4 主机中运行的特权虚拟机(支持移动 IP 协议迁移的物理主机皆需运行该特权虚拟机)协议据此判断迁移而来的为一 IPv6 虚拟机.

(2) 特权虚拟机用已收到 $r4$ (该路由器已配置为双栈路由器)发出的代理通告可确定 $r4$ 为 dv 的 IPv4 外区代理, 并在其维护的地址池中确定一 IPv4 地址作为临时关照 IPv4 地址.

(3) 特权虚拟机发送 IPv6 报文向 IPv6 路由器请求本 IPv4 网段的通告前缀. 因该物理主机可能位于该 IPv4 网络内的多层 IPv4 路由器下, 为使该框架具备普适性该请求前缀 IPv4 包需被特权虚拟机进行 IPv6-in-IPv4 封装, 发送至 IPv4 边缘 IPv4 路由器即 $r4$. $r4$ 再将其进行解封装并生成前缀回应, 且对回应数据包进行 IPv6-in-IPv4 封装传回到特权虚拟机, 特权虚拟机解封装后获得前缀. 特权虚拟机对 dv 进行请求回应. dv 则获取临时关照 IPv6 地址, 其由已获得的通告前缀与临时关照 IPv4 地址(不足位补 0)组成.

(4) 据移动 IPv6 协议, dv 向 $r2$ 家乡代理(为双栈路由器)发送绑定注册报文. 绑定地址为原 sv 的 IPv6 地址与获取临时关照 IPv6 地址. 该报文由特权虚拟机截获并进行 IPv6-in-IPv4 封装, 隧道出口为 $r4$.

(5) $r4$ 解封装得到原始 IPv6 报文并进行 IPv6-in-IPv6 封装, 出口为云计算控制引擎. 云计算控制引擎解封装后再进行 IPv6-in-IPv4 封装传送到 $r2$ 后解封装. 如 sv 位于的物理主机位于其下层 IPv6 路由器链路, 则向下传递至家乡代理建立绑定.

(6) 下一时间通信伙伴则为 IPv6 主机, 欲获取云计算服务, 云计算控制引擎作出由原 sv 虚拟机提供服务的策略.

(7) 云计算控制引擎将报文进行 IPv6-in-IPv4 封装发送至 $r4$, 如原 sv 位于 $r4$ 的下层路由器则 $r4$ 将其解封装后再发送给其家乡代理, 否则 $r4$ 为家乡代理则接受报文.

(8) 家乡代理将报文据移动 IPv6 协议进行 IPv6-in-IPv6 封装, 该报文再经由 $r4$ 解封装还原其 IPv6 报文并进行 IPv6-in-IPv4 封装出口为云计算控制引擎.

(9) 云计算的控制引擎将其封装后进行 IPv6-in-IPv6 封装发送至 $r2$, 其解封装后再进行 IPv6-in-IPv4 封装发往 dv 临时关照 IPv4 地址. 则特权虚拟机将其捕获解封装发送至 dv 接收.

(10) dv 发现其经由家乡代理中转, 则向通信伙伴发送地址绑定信息. 该绑定报文经特权虚拟机截获进行 IPv6-in-IPv4 封装, 出口为 $r4$, $r4$ 再将其解封装发送至 IPv6 路由器最终转发至通信伙伴.

(11) 通信伙伴收到报文建立地址绑定, 后续发往原 sv 的数据包可直接发往其临时关照 IPv6 地址.

如通信伙伴为 IPv4 主机, 则需云计算控制引擎首先对 IPv4 报文进行 NAT-PT^[5] 转换(通信伙伴 IPv6 地址由其 IPv4 地址前加上一分配前缀组成), 然后可将该请求报文进行 IPv6-in-IPv4 封装发至家乡代理网络. 报文可从家乡代理出发经过同样的步骤来到 dv . dv 与通信伙伴的通信同样经相同步骤来到云计算控制引擎, 云计算控制引擎再进行 NAT-PT 转换发送至通信伙伴.

现再令被 IPv6 网络包围 IPv4 网络中的待迁 IPv4 虚拟机为 sv , 被 IPv4 网络包围的 IPv6 网络中的目的 IPv6 物理机中迁移而来的 IPv4 虚拟机为 dv .

(1) sv 迁移后据移动 IPv4 协议申请外区代理, 该物理 IPv6 主机中运行的特权虚拟机据此判断迁移而来的为一 IPv4 虚拟机.

(2) 特权虚拟机以本物理 IPv6 主机前缀为前缀, 并以后 32 位为原 sv 的 IPv4 地址加上前缀(不足位补 0)组成 dv 对应的临时关照 IPv6 地址.

(3) 特权虚拟机在管理的地址池中选择一 IPv4 地址作为 dv 的临时关照 IPv4 地址. 并对 dv 进行临时关照 IPv4 地址的回应.

(4) dv 据移动 IPv4 协议发出 IPv4 报文, 目的 IPv4 地址为 $r4$ (家乡代理)IPv4 地址, 进行原 IPv4 地址与获取的临时关照 IPv4 地址的绑定注册.

(5) 绑定注册报文被特权虚拟机截获并做 IPv4-in-IPv6 封装出口为 $r2$, $r2$ 解封装后再进行 IPv4-in-IPv4 封装出口为云计算控制引擎. 云计算控制引擎解封装后再进行 IPv4-in-IPv6 封装来到 $r4$ 解封装后被家乡代理接收.

(6) 家乡代理建立绑定.

(7) 下一时间, 通信伙伴(IPv4 主机)欲获取云计算服务, 云计算控制引擎作出由原 sv 进行服务提供的策略.

(8) 该报文被云计算控制引擎进行 IPv4-in-IPv6 封装经 r4 解封装,并被家乡代理截获且根据移动 IPv4 协议封装 IPv4-in-IPv4 发送该包去往临时关照 IPv4 地址. 该包将分别经过 r4、云计算控制引擎、r2、特权虚拟机的封装解封装来到 dv.

(9) dv 回应通信伙伴,该 IPv4 报文经特权虚拟机进行 IPv4-in-IPv6 封装,隧道出口为 r2. r2 将其解封装后发往云计算控制引擎最终转发给通信伙伴.

如通信伙伴为 IPv6 主机,则需云计算控制引擎首先对 IPv6 报文进行 NAT-PT 的转换,然后将该请求报文进行 IPv4-in-IPv6 封装发至家乡代理网络. 报文从家乡代理出发经过同样的步骤来到 dv. dv 与通信伙伴的通信同样经相同步骤来到云计算控制引擎,云计算控制引擎再进行 NAT-PT 转换发送至通信伙伴.

IPv4/IPv6 过渡中期实现的步骤较前讨论的 IPv4/IPv6 演进的前期和后期迁移较为简单,结构的简化步骤和上相同.

3.3 特权虚拟机

因跨异构网络进行虚拟机迁移移动通信的需要,每一欲具备支持移动功能实现虚拟机迁移的物理主机需运行文中开发的特权虚拟机^[6]. 该特权虚拟机主要功能为:

(1) 判断外区网络协议环境.

此处如多次申请代理无响应则可判断来到异构网络.

(2) IPv4 或 IPv6 虚拟机迁移至 IPv6 或 IPv4 网络时,帮助生成临时关照地址.

(3) 提供隧道通信支持.

需进行相应的隧道配置.

(4) 与云计算控制引擎资源监控模块交互帮助云计算模块决定迁移策略^[6].

此处根据对虚拟机 CPU、内存、网络流量的使用情况监控,判断资源使用热点从而进行虚拟机迁移判断.

4 实验与结果分析

4.1 框架实现

实验架构见图 4,可支持基于移动通信的虚拟机迁移物理主机运行 Windows xp 系统,云计算控制引擎操作系统为 fedora8.0. 系统总体框架实现基于 Libpcap 和 Libnet 函数库,进行数据捕获读取并生成发送数据包.

定义 IPv4/IPv6 基本报头的回调函数:

```
void ip_protocol_packet_callback(u_char * argument, const struct pcap_pkthdr * packet_header, const u_char * packet_content)
```

根据以上 IPv4/IPv6 基本报头的格式定义相应的结构体然后在回调函数中定义指向相应结构体的指针.

```
struct ipv4_header * ip4_protocol; //指向 IPv4 基本报头结构体的指针
```

```
struct ipv6_header * ip6_protocol; //指向 IPv6 基本报头结构体的指针
```

```
ip4_protocol = (struct ipv4_header *) (packet_content + 14);
```

```
ip6_protocol = (struct ipv6_header *) (packet_content + 14);
```

读取通信对端发来的报文,再根据其和服务虚拟机的协议类型生成新报文并通过 Libnet 函数库发送报文.

过渡控制引擎功能模块使用 Linux 的 Netfilter 框架. 文中框架在 NF-IP-PRE-ROUTING 钩子点挂载过渡控制引擎功能模块.

IPv4 主机和 IPv6 主机的地址映射信息和通信状态存储在 ip_contrack 这个数据结构中.

```
struct ip_contrack{
    struct nf_contrack ct_general;
    struct ip_contrack_tuple_hash tuplehash[IP_CT_DIR_MAX];
    unsigned long status;
    struct timer_list timeout;
    struct nf_ct_info infos[IP_CT_NUMBER];
    union ip_contrack_proto proto;
};
```

IPv4 通信主机移动至 IPv6 网络后需通过 IPv4/IPv6 插件向过渡控制引擎的地址池管理模块申请一 IPv4 地址作为临时关照 IPv4 地址. 地址池管理模块需要为 IPv6 临时关照地址与 IPv4 临时关照地址建立一个映射. 这个地址映射关系存储在 ip_contrack 结构中.

4.2 实验分析

该框架的不足主要在设计的云计算控制引擎很容易成为云计算平台虚拟机迁移和服务提供过程中的瓶颈,且因针对演进的前期、后期隧道机制步骤较为复杂,这些因素都会影响到云计算平台的整体性

(下转第 58 页)

小. 基于新颖电流模式 CCCII+ 设计了一种电流采样保持电路, 直接处理电流信号, 无需传统的 V-I 变换级. 采样精度和采样速率能满足电流模式高精度流水线模数转换的采样保持要求, 适应于标准 CMOS 集成工艺下低成本的大规模电路的应用.

参考文献:

- [1] Wilson B. Recent developments in current conveyors and current-mode circuits[J]. IEE Proceedings G: Circuits, Devices and Systems, 1990, 137(2): 63-77.
- [2] 彭建华, 陈向东, 杜广涛. 一种轨对轨输入的低压低功耗 CMOS 电流传输器[J]. 微电子学与计算机, 2008, 25(11): 72-76.
- [3] 董招辉, 彭良玉, 欧增强. 新颖的基于 MOCCII 的多相位正弦振荡器[J]. 微电子学与计算机, 2008, 25(7): 222-224.
- [4] 罗秀龙, 王春华, 谢莉, 等. 新型 CMOS 电流控制电流差分缓冲放大器(CCCDBA)[J]. 微电子学与计算机, 2010, 27(5): 106-110.
- [5] Fabre A, saaid O, Boucheron C. High frequency applications based on a new current controlled conveyor[J]. IEEE Transactions On Circuits And System I: Fundamental Theory And Applications, 1996, 43(2): 82-91.

- [6] Ronak Trivedi. Low power and high speed sample-and-hold circuit[C]//IEEE International Midwest Symposium on Circuits and Systems. USA: Puerto Rico, 2006: 453-456.
- [7] Rajaee O, Jahanian A, A Bakhtiar M S. A low voltage, high speed, high resolution class AB switched current sample and hold[C]//IEEE International Symposium on Circuits and Systems. Greece: Kos, 2006: 1039-1042.
- [8] Yuh-Shyan Hwang, Lu-Po Liao, Chia-Chun Tsai, et al. A new CCII-based pipelined analog to digital converter[C]//IEEE International Symposium on Circuits and Systems. Japan: Kobe, 2005: 6170-6173.

作者简介:

- 李湛 男, (1986-), 硕士研究生. 研究方向为电流模式模数转换器设计.
- 王春华 男, (1963-), 教授, 博士生导师. 研究方向为电流模式电路和射频集成电路设计.
- 彭关超 男, (1984-), 硕士研究生. 研究方向为 CMOS 模拟集成电路.
- 王会杰 男, (1984-), 硕士研究生. 研究方向为 CMOS 模拟集成电路.

(上接第 54 页)

能与服务质量. 本实验进行了无负载基础系统 IPv4 虚拟机迁移至 IPv6 网络和无负载基础系统 IPv6 虚拟机迁移至 IPv4 网络验证^[7]. 实验结果发现该框架可实现云计算的移动 IPv4/IPv6 过渡.

5 结束语

文中针对云计算平台描述了一种自适应的虚拟机跨 IPv4/IPv6 子网间的在线迁移系统, 该框架从一定程度上解决了云计算环境中网络协议异构所带来的资源使用问题. 方案不足之处为单核模式下系统及系统本身复杂性带来的平台整体性能和服务质量的下降, 这将是下一步研究工作的重点.

参考文献:

- [1] Clark C, Fraser K, Hand S, et al. Live migration of virtual machines[C]//Proc of the 2nd ACM/USENIX Symposium on Networked Systems Design and Implementation. 2005: 273-286.
- [2] Johnson D, Perkins C. Mobility support in IPv6[S]. Rice

University, RFC 3775, 2004.

- [3] LanCaster University. MIPL mobile IPv6 for Linux[EB/OL]. <http://mobile-ipv6.org/>. 2006.
- [4] Durand A, Fasano P, Guardini I, et al. IPv6 tunnel broker [S]. California: Sun Microsystems Inc, RFC3053, 2001.
- [5] Tsirtsis G, Srisuresh P. Network address translation-protocol translation (NAT-PT)[S]. Campio Communication, RFC276-6, 2000.
- [6] 刘鹏程, 陈榕. 面向云计算的虚拟机动态迁移框架[J]. 计算机工程, 2010(3): 37-39.
- [7] 江雪, 李小勇. 虚拟机动态迁移的研究[J]. 计算机应用, 2008, 28(9): 2375-2377.

作者简介:

- 陈俊 男, (1979-), 博士, 讲师. 研究方向为下一代互联网技术.
- 陈孝威 男, (1945-), 教授, 博士生导师. 研究方向为网络技术.