

基于移动 IPv4/IPv6 演进云计算框架设计*

陈俊^{1,2}, 陈孝威¹

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

摘要: 为满足云计算平台的资源需求,充分利用异构网络资源建设云计算基础平台,设计了一种基于移动 IP 技术的 IPv4/IPv6 虚拟机在线迁移框架。该框架的实现需利用设计的云计算控制引擎为核心转换连接异构网络并在 NAT-PT 与隧道技术协作下完成。框架建立的网络结构适用于 IPv4 与 IPv6 虚拟机之间的无缝迁移,并可向客户端跨 IPv4/IPv6 网络提供云计算服务。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 IPv4; 移动 IPv6; 隧道技术; NAT-PT

中图分类号: TP393 **文献标志码:** A **文章编号:** 1001-3695(2011)06-2321-03

doi:10.3969/j.issn.1001-3695.2011.06.087

Mobile IPv4/IPv6 transition framework design for cloud computing

CHEN Jun^{1,2}, CHEN Xiao-wei¹

(1. College of Computer Science & Information, Guizhou University, Guiyang 550025, China; 2. School of Education Science, Guizhou Normal University, Guiyang 550001, China)

Abstract: In order to provide resource requirement for cloud computing platform, and fully used heterogeneous network resources, designed an IPv4/IPv6 live migration framework for virtual machine based on mobile IP technology. The framework used the designed cloud computing control engine as a core translates and link heterogeneous network, and needed NAT-PT and tunnel technology collaboration. The framework established for IPv4/IPv6 virtual machine seamless live migration, and provided IPv4/IPv6 cloud computing service for client. The framework can be applied to construction cloud computing platform in the IPv4/IPv6 transition period.

Key words: cloud computing; MIPv4; MIPv6; tunnel technology; NAT-PT

继分布式计算、并行计算和网格计算模型后,学术界提出了云计算模型。云计算平台利用虚拟机在线迁移技术^[1]可在线实现虚拟机和物理资源的重映射,从而动态地实现整个系统的负载均衡,为避免产生热点提供了解决方案。随着普适计算和无线移动通信技术的快速发展,互联网正向支持大范围移动性方向发展。目前移动 IP 技术有移动 IPv4 和移动 IPv6 两种。移动 IPv6 的提出^[1,2]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 和纯 IPv6 网络互相交错的局面。

云计算平台需要海量的廉价设备构建,为充分利用异构网络协议的计算资源,需要解决 IPv4 和 IPv6 虚拟机之间网络连接重定向的过渡问题。目前国内外一些关于虚拟机迁移过渡技术方案虽然能够解决虚拟机跨子网的在线迁移,但没有考虑到 IPv4/IPv6 子网间的在线迁移。例如, Snoeren 等人^[3]提出以源虚拟机和目标虚拟机共同运行为核心的网络连接重定向技术; Travostino 等人^[4]提出基于隧道建立的动态 IP 隧道配置网络连接重定向技术。为了更好地利用现有网络资源服务于云计算基础平台建设,必须解决虚拟机在 IPv4/IPv6 子网间的在线迁移。

鉴于此,本文针对移动通信环境设计了基于云计算的移动 IPv4 虚拟机与 IPv6 虚拟机过渡框架。

1 相关技术

1.1 移动 IPv4

如图 1 所示,移动 IPv4 主要由四个实体组成:移动节点(MN)、家乡代理(HA)、外地代理(FA)和通信伙伴节点(CN)。MN 在家乡网络拥有一个永久性的 IP 地址(home of address, HoA),当 MN 移动到外地网络后将获得一个临时的关照地址(care of address, CoA),此处关照地址分为两种,即外地代理 IP 地址和通过某种机制与移动节点暂时对应起来的网络地址。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。

1.2 移动 IPv6

如图 2 所示,移动 IPv6 不需要外地代理,主要由三个实体组成:移动节点(MN)、家乡代理(HA)和通信伙伴节点(CN)。

收稿日期: 2010-10-28; 修回日期: 2010-12-03 基金项目: 贵州大学研究生创新基金资助项目(省研理工 2010006)

作者简介: 陈俊(1979-),男,贵州贵阳人,讲师,博士,主要研究方向为分布式计算、下一代互联网技术(starcraft_cj@163.com); 陈孝威(1945-),男,教授,博导,主要研究方向为网络通信技术。

通过程可参照移动 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。并且 MN 若发现 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 的数据包将直接以 CoA 作为目标地址,移动 IPv4 通过程中的三角路由得以消除。



2 系统框架

2.1 网络架构

如图 3 所示,首先建立一个 IPv4 和一个 IPv6 网络,然后在 IPv4 与 IPv6 路由器之间部署文中设计的云计算控制引擎,并通过该云计算控制引擎连接通信伙伴。该云计算控制引擎会与开发运行在物理主机上的特权虚拟机交互完成 IPv4/IPv6 虚拟机跨异构网络迁移的透明过渡。

2.2 云计算控制引擎

云计算控制引擎为双栈路由器,主要由下列功能模块实现:地址池前缀管理、隧道管理、迁移管理和 NAT-PT(图 4)。



1)地址池管理 该模块将在过渡控制引擎中建立 IPv4 地址池。其用于 IPv4 虚拟机迁移到 IPv6 网络后从其动态分配一 IPv4 地址作为关照地址,或用于 IPv6 服务获取主机和 IPv6 虚拟机进行动态 IPv4 地址分配。

2)前缀管理 该模块将针对每一连接至云计算控制引擎的 IPv4 网段进行 IPv6 前缀指定。其用于帮助 IPv6 虚拟机迁移到 IPv4 网络用指定前缀结合申请得到的 IPv4 关照地址生成 IPv6 关照地址与原 IPv6 地址绑定告知家乡代理。

3)隧道管理^[5] 该模块利用隧道封装、解封装技术为家乡代理利用隧道转发数据包;跨 IPv4 网络实现 IPv6 通信和跨 IPv6 网络实现 IPv4 通信提供支持。其封装方式有 IPv4 in IPv4、IPv6 in IPv6、IPv4 in IPv6、IPv6 in IPv4。

4)迁移管理 该模块与支持移动功能的虚拟机迁移物理主机上运行的特权虚拟机交互寻找资源热点,作出迁移决策。

5)NAT-PT^[6] 该模块用于服务获取主机与服务提供虚拟机为异构网络协议情况下进行数据包协议转换,从而可实现异构网络协议之间的服务提供。

2.3 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。该特权虚拟机主要功能如下:

a)判断外区网络协议环境。

b)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。

c)提供隧道通信支持。

d)与云计算控制引擎的资源监控模块交互,帮助云计算模块决定迁移策略^[7]。

3 过渡方案原理

如图 3 所示,现今云计算控制引擎 IPv6 地址为 2::1,IPv4 地址为 172.168.2.1。 r_1 的 IPv4 地址为 192.168.1.2, r_2 的 IPv6 地址为 1::1。

3.1 IPv4 虚拟机迁移至 IPv6 网络

现今 r_1 连接的 IPv4 网络中待迁移虚拟机为 sv ,IPv6 网络中完成 sv 迁移的目的虚拟机为 dv 。迁移过程的数据包传送需要经过云计算控制引擎进行 NAT-PT 转换。

a) 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。

b)特权虚拟机以本物理 IPv6 主机的前缀为前缀,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。此处为 1::193.168.1.1。

c)特权虚拟机发送 IPv6 报文向云计算控制引擎申请一 IPv4 地址作为临时关照 IPv4 地址。现今云计算控制引擎从 IPv4 地址池中取出 172.168.2.2 作为该 IPv4 地址进行回应。

d)特权虚拟机对 dv 进行临时关照 IPv4 地址 172.168.2.2 的回应。

e) dv 根据移动 IPv4 协议发出 IPv4 报文,目的 IPv4 地址为 r_1 (家乡代理)IPv4 地址,进行 192.168.1.1 原 IPv4 地址与 172.168.2.2 临时关照 IPv4 地址的绑定注册。

f)特权虚拟机对该报文进行 IPv4-in-IPv6 隧道封装,隧道出口为云计算控制引擎。在出口处由云计算控制引擎的隧道管理模块解封装后发往 r_1 。

g) r_1 收到注册包后建立地址绑定。

h)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。

i)该报文被 r_1 (家乡代理)截获并根据移动 IPv4 协议封装 IPv4-in-IPv4 发送该包去往临时关照 IPv4 地址 172.168.2.2。

j)报文经由云计算控制引擎发现,目的 IPv4 地址 172.168.2.2 为自己地址池中已分配 IPv4 地址,则解去 IPv4 包头封装,对原 IPv4 报文进行 IPv4-in-IPv6 重封装。目的 IPv6 地址则为 dv 临时关照 IPv6 地址 1::192.168.1.1。

k)特权虚拟机截获该报文解封装后发送给 dv 完成通信。

l) dv 回应通信伙伴,该 IPv4 报文经特权虚拟机进行 IPv4-in-IPv6 封装,隧道出口为云计算控制引擎。

m)云计算控制引擎收到该报文解封装后转发给通信伙伴完成通信。

以上步骤 h)~m)若通信伙伴为 IPv6 主机,则需云计算控制引擎首先进行 NAT-PT 转换再将该请求报文发至家乡代理网络;同样,步骤 m)的云计算控制引擎收到回应报文后,应先解封装得到原始 IPv4 报文后再进行 NAT-PT 转换为 IPv6 报文转发给 IPv6 通信伙伴完成通信。

3.2 IPv6 虚拟机迁移至 IPv4 网络

现今 r_2 连接的 IPv6 网络中待迁移虚拟机为 sv ,IPv4 网络

中完成 sv 迁移的目的虚拟机为 dv 。同样迁移过程的数据包传递需要经过云计算控制引擎进行 NAT-PT 转换。

a) sv 迁移后根据移动 IPv6 协议请求地址,该物理 IPv4 主机中运行的特权虚拟机据此判断迁移而来的为 IPv6 虚拟机。

b) 特权虚拟机用已收到 r_1 发出的代理通告确定 r_1 为 dv 的 IPv4 外区代理,并与 DHCP 通信,为 dv 确定一 IPv4 地址作为临时关照 IPv4 地址。现令为 192.168.1.3。

c) 特权虚拟机发送 IPv4 报文向云计算控制引擎请求本 IPv4 网段的通告前缀。现令为 3::。特权虚拟机对 dv 进行请求回应。 dv 获取临时关照 IPv6 地址,它由与获取的通告前缀加临时关照 IPv4 地址(不足位补 0)组成。

d) 根据移动 IPv6 协议, dv 向 r_2 (家乡代理)1::1 发送绑定注册报文。绑定地址为原 sv 的 IPv6 地址 1::2 与临时关照 IPv6 地址 3::192.168.1.3。该报文由特权虚拟机截获进行 IPv6-in-IPv4 封装。隧道出口为云计算控制引擎。

e) 云计算控制引擎收到报文解封装后发往家乡代理 1::1。

f) r_2 建立地址绑定。

g) 下一时间通信伙伴 1::2(IPv6 主机)欲获取云计算服务,云计算控制引擎作出由 IPv6 地址为 1::2 的原 sv 进行服务提供策略。

h) r_2 截获该报文并进行 IPv6-in-IPv6 封装。外部 IPv6 包头的目的 IPv6 地址为 3::192.168.1.3,其为 dv 的临时关照 IPv6 地址。

i) 报文经由云计算控制引擎,云计算控制引擎发现外部 IPv6 包头的目的 IPv6 地址为自己通告前缀,则解包并进行 IPv6-in-IPv4 重封装,目的 IPv4 地址为 192.168.1.3。

j) 特权虚拟机收到该包解封装后发送给 dv 完成通信。

k) dv 发现该报文经由家乡代理转发,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。隧道出口为云计算控制引擎。绑定地址为原 sv 的 IPv6 地址 1::2 与获取的临时关照 IPv6 地址,此处为 3::192.168.1.3。

l) 云计算控制引擎收到该包解封装后转发给通信伙伴。

m) 通信伙伴收到报文建立地址绑定。后续发往 dv 的数据包可直接使用 3::192.168.1.3 地址发送。需经云计算控制引擎利用 IPv6-in-IPv4 封装进行隧道发送至特权虚拟机,特权虚拟机解封装后发送给 dv 完成通信。

以上步骤 g) ~ m) 若通信伙伴为 IPv4 主机,则需云计算控制引擎首先对绑定的 IPv6 虚拟机在 IPv4 地址池中指定一 IPv4 地址,再对此 IPv4 地址为目的 IPv4 地址的报文进行 NAT-PT 转换(通信伙伴的 IPv6 地址由其 IPv4 地址前加上一分配前缀组成),然后将该请求报文发至家乡代理网络。同样,步骤 k) dv 发向通信伙伴的报文经特权虚拟机截获进行 IPv6-in-IPv4 封装来到云计算控制引擎,云计算控制引擎应先解封装后再进行 NAT-PT 转换。若该数据包为绑定地址的报文(步骤 k)即是如此),则应将绑定内容转变为云计算控制引擎作出指定 IPv6 虚拟机策略时指定的 IPv4 地址与 dv 的临时关照 IPv4 地址之间的绑定。

4 系统实现与分析

4.1 系统实现

系统总体框架实现基于 Libpcap 和 Libnet 函数库^[8],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
```

无论 IPv4 或 IPv6,其以太网首部都完全一样,所以此处分析 IP 头部皆从数据包首部跳跃 14 Byte(以太网首部的长度)找到 IP 头部。

4.2 实验分析

实验架构如图 3 所示,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 fedora 8.0。该实验目的为验证评估云计算的移动 IPv4/IPv6 过渡框架的有效性。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 网络的验证。实验结果发现,该框架可实现云计算的移动 IPv4/IPv6 过渡。

该框架的不足主要在于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,进而影响云计算平台的整体性能和服务质量。因为云计算平台需要大量的计算资源构建,迁移行为有可能较为频繁。这样需考虑平滑切换与快速切换的问题。为消除瓶颈效应,可考虑设置多核并行工作或分工协作的云计算控制引擎,并应适当考虑云计算控制引擎的备份冗余以增加框架的整体可靠性。另增加微移动机制解决平滑切换与快速切换问题。

5 结束语

本文描述了一种云计算的移动 IPv4/IPv6 过渡框架方案。该框架在一定程度上解决了云计算环境中网络协议异构性所带来的资源使用问题。下一步工作的研究重点将考虑云计算控制引擎的并行、备份与物理主机对过渡透明性的要求。

参考文献:

- [1] JOHNSON D, PERKINS C. RFC 3775, Mobility support in IPv6[S]. 2004.
- [2] MIPL mobile IPv6 for Linux[EB/OL]. (2006). <http://mobile-ipv6.org/>.
- [3] SNOEREN A C, BALAKRISHNAN H. An end to end approach to host mobility[C]//Proc of the 6th Annual International Conference on Mobile Computing and Networking. 2000:155-166.
- [4] TRAVOSTINO F, DASPIT P, GOMMANS L, et al. Seamless live migration of virtual machines over the MAN/WAN[J]. Elsevier Future Generation Computer Systems, 2006, 22(8):901-907.
- [5] DURAND A, FASANO P, GUARDINI I, et al. RFC 3053, IPv6 tunnel broker[S]. 2001.
- [6] TSIRTISIS G, SRISURESH P. RFC 2766, Network address translation-protocol translation (NAT-PT)[S]. 2000.
- [7] 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, 36(57):37-39.
- [8] STEVEN M. The BSD packet filter: a new architecture for user-level packet capture[C]//Proc of the Winter USENIX Conference. 2003: 259-269.