

基于 MDA-FHMIPv6 的云计算虚拟机迁移系统

陈 俊, 张伟云, 杨建军

(贵州师范大学 图书馆, 贵州 贵阳 550001)

摘 要: 为了充分利用 IPv6 网络资源建设云计算基础平台, 设计了一种基于改进 FHMIPv6 的 MDA-FHMIPv6 协议 IPv6 虚拟机在线迁移系统. 利用设计的云计算控制引擎为核心转换连接 IPv6 网络完成虚拟机在线迁移, 并向客户端提供云计算服务. 该云计算控制引擎联合特权虚拟机实现了基于 FHMIPv6 的 MAP 动态自适应功能. 经实验验证该系统可应用于 IPv6 网络云计算基础平台建设, 并提高云计算环境中的虚拟机迁移效率.

关键词: 云计算; 微移动; FHMIPv6; MDA-FHMIPv6

中图分类号: TP393

文献标识码: A

文章编号: 1000-7180(2014)03-0102-02

Cloud Computing Virtual Machine Migration System Based on MDA-FHMIPv6

CHEN Jun, ZHANG Wei-yun, YANG Jian-jun

(Library of Guizhou Normal University, Guiyang 550001, China)

Abstract: In order to fully use IPv6 network resources and provide resource requirement for cloud computing platform, this paper designed an IPv6 virtual machine live migration transition framework based on the improved FHMIPv6, named MDA-FHMIPv6 protocol. To use the paper designed cloud computing control engine as a core connect IPv6 network completion virtual machine migration, and provides cloud computing service for client. The cloud computing control engine can realize a dynamic adaptive MAP function based on FHMIPv6. Experiment shows the framework can support IPv6 network cloud computing platform construction, and provides improve virtual machine migration efficiency in cloud computing environment.

Key words: cloud computing; micro-mobility; FHMIPv6; MDA-FHMIPv6

1 引言

虚拟机在线迁移技术为云计算平台资源管理, 从而实现服务器之间负载动态平衡, 避免产生热点提供了解决方案. 现有虚拟机在线迁移技术, 如 VMware VMotion 与 Xen Live Migration 等, 将网络重定向局限在局域网内^[1]. Xen 采用主动地址解析协议技术, 广播更新交换机 MAC 地址表中 MAC 地址与端口号映射以实现虚拟机迁移后网络重定国内外学者对虚拟机跨网段迁移提出了一些解决方案: 文献[2]中提出跨局域网进行虚拟机在线迁移需

针对虚拟机迁移中和迁移后已提供服务维持. 到其他网络后, 依据网络要求重新进行网络配置, 然后启动虚拟机.

文献[3]中提出建立双向 IP 隧道, 通过隧道封装、解封保持源 IP 地址不变. 新通信到来时, 数据包经源节点转发至迁移后目的虚拟机.

文献[4]中提出利用虚拟专用局域网业务桥接分布在 WAN 范围内的多个物理节点以形成一个虚拟局域网. 在该虚拟局域网内, 虚拟机可完成类似在局域网内虚拟机迁移而实现的网络重定向技术, 虚拟机无需改变 IP 地址即可实现网络重定向.

收稿日期: 2013-04-29; 修回日期: 2013-06-20

基金项目: 贵州师范大学资助博士科研项目(基于 MDT-FHMIPv6 协议 IPv4/IPv6 虚拟机微移动迁移系统研究); 贵州省高校人文社科研究基地招标项目(12JD059); 贵州省科学技术基金项目(黔科合 J 字 LKS[2012]40 号)

以上研究成果在 IPv4 网络环境下实现了局域网内或跨网段的虚拟机迁移。文献[2]中进行网络重定向需改变源 IP 地址,文献[3-4]可保持原有 IP 地址不变。

本文针对 IPv6 网络环境跨网段虚拟机迁移进行研究,设计的迁移系统,将 FHMIPv6 协议进行 MAP 域动态自适应改进后运用于云计算虚拟机迁移系统,并将其称为 MDA-FHMIPv6 (Map Dynamic Adaptive-FHMIPv6)。

2 针对 FHMIPv6 分析

2.1 FHMIPv6 运行原理 MDA-FHMIPv6 迁移系统设计

对比 FMIPv6 不同的是,FHMIPv6 在 MAP 和 NAR 之间(FMIPv6 是 PAR 和 NAR 之间)建立快速切换的分组转发隧道。为此,移动节点与 MAP (FMIPv6 是移动节点与 PAR)交换快速切换消息。FHMIPv6 仍然采用 FMIPv6 中的消息完成快速切换,不必定义任何新的消息^[6]。现针对地址称谓做如下定义:

(1)先前链路转交地址(Previous On-link CoA, PLCoA):移动节点在 PAR 链路上采用的链路转交地址。它与 FMIPv6 协议中的 PCoA 相对应。

(2)新链路转交地址(New On-link CoA, NL-CoA):移动节点在 NAR 链路上采用的链路转交地址,它与 FMIPv6 协议中的 NCoA 相对应。

2.2 FHMIPv6 虚拟机迁移应用分析

FHMIPv6 可大幅缩短切换时延,且可明显减少分组丢失。该运行原理可应用于云计算虚拟机迁移,在实验实现中使用的 IPv6 网络环境见图 1。针对 FHMIPv6 中 HMIPv6 原理部分进行进一步分析可发现,MAP 区域的设定对于迁移效率的提高有直接关联。如区域过大,因发给区域下迁移虚拟机都需要经过 MAP 转发,则可能造成 MAP 转发数据包的瓶颈效应,极大的降低迁移后相互之间的通信效率。且针对区域内的迁移亦需要向 MAP 注册 LCoA,可能会造成 MAP 负担过重极大降低 MAP 的工作效率。如区域过小,会因为大部分迁移都定义为跨区域迁移,则期望达到减少注册花费时间的效果不能完成。

下文将针对 MAP 间微移动频繁情况,使用 MAP 动态自适应方式来进行 MAP 区域的设定,从而提高虚拟机迁移效率

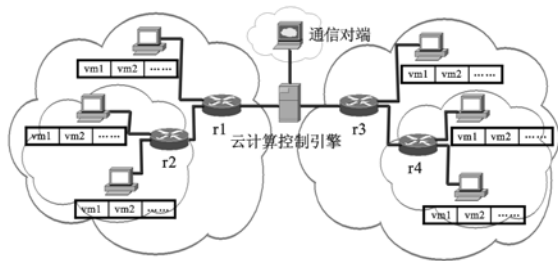


图1 FHMIPv6 协议在虚拟机微移动迁移中的运用

3 MDA-FHMIPv6 迁移系统设计

3.1 FHMIPv6 运行原理 MDA-FHMIPv6 迁移系统设计

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情况,依据负载均衡策略对通信对端指定虚拟机服务并帮助完成虚拟机迁移决策。云计算控制引擎模型见图 2。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。

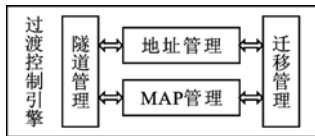


图2 云计算控制引擎模型

3.2 FHMIPv6 运行原理 MDA-FHMIPv6 迁移系统设计

因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^[7]。该特权虚拟机主要功能如下:

- (1)帮助生成临时转交地址。
- (2)提供隧道通信支持。
- (3)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。
- (4)记录阶段迁移行为记录。

此处,开发的特权虚拟机,因将应用于采用 MAP 动态自适应技术的云计算虚拟机迁移系统,因此,将比仅使用 FHMIPv6 技术的云计算虚拟机迁移系统增加(4)的功能模块。

3.3 基于 FHMIPv6 MAP 动态自适应技术

基于以上分析,区域的设定与后期能否达到迁移效果关联非常紧密。如采用传统方式进行固定区域范围设定,则不能满足变化繁多的云计算迁移系

统效果要求. 我们采用动态自适应方式进行 MAP 区域的设定.

网络物理架构和上节讨论的完全相同, 见图 1. 特别注意的是, 针对 MAP 域间虚拟机迁移的 MAP 域调整, 即为 MAP 域扩大. 如果仅仅是 MAP 域间低频率迁移 (实现实验中我们在 MAP 中设定相邻 MAP 关联权重分值数组变量, 用以判断是否达到 MAP 区域的改变条件), 则 MAP 区域不需要改变. 针对 MAP 域内虚拟机迁移的 MAP 域调整, 即为 MAP 域缩小. 如果仅仅是 MAP 域内虚拟机短时间未发生迁移行为, 则 MAP 区域亦不需要发生改变. MAP 区域的频繁改变, 将会对整个原 MAP 域乃至扩展 MAP 域内的所有虚拟机运行效率产生较大影响.

我们采用虚拟机迁移方式来实现 MAP 区域的改变. 针对 MAP 区域的扩大调整, 此处设定 MAP 中相邻 MAP 关联权重分值变量为 5 时, 进行 MAP 与该关联 MAP 的区域迁移, 其所有针对关联 MAP 区域权重分值初始设定为 0, 10 分钟归 0 一次. 每当不同 MAP 区域的两台物理机在 1 分钟内进行两次相互虚拟机迁移, 则针对该关联 MAP 的权重分值加 1, 同一物理机无论迁移的次数, 至多只对关联 MAP 权重分值贡献 1 分值. 这样, 可避免因某一物理机中虚拟机的乒乓迁移行为过多而影响到 MAP 区域内所有虚拟机的 MAP 认定.

针对 MAP 区域的缩小调整, 此处设定 10 分钟内, 如针对本 MAP 的域内迁移权重分值未达到 2 时, 进行 MAP 区域分离, 其所有针对本 MAP 的域内迁移权重分值初始设定为 0, 10 分钟归 0 一次. 同样, 每当在本 MAP 区域内的两台物理机在 1 分钟内进行两次相互虚拟机迁移, 则针对本 MAP 的域内迁移权重分值加 1, 同一物理机无论迁移的次数, 至多只对本 MAP 的域内迁移权重分值贡献 1 分值.

特别注意的是, 因为我们需要在虚拟机迁移行为中定性频繁迁移行为, 则需要对迁移行为做记时迁移记录, 否则无法判断. 如将迁移行为记录工作交给云计算控制引擎, 任务可以完成, 但因为考虑到云计算为巨大的网络环境, 其平台内拥有的机器数目众多, 如皆实现自适应 MAP 区域功能, 则需要云计算控制引擎记录平台内所有虚拟机的阶段迁移行为. 这样对于云计算控制引擎性能要求过高, 将可能造成云计算控制引擎性能瓶颈, 对云计算的整体效能产生影响. 因此, 我们考虑采用确定的 MAP 设备

与物理机器上运行的特权虚拟机进行交互从而实现这一记录阶段内虚拟机迁移行为功能的要求. 因为任务分布式完成, 则可消除该任务由于平台内物理机器数量众多而造成的瓶颈效应.

现假定云计算控制引擎依据负载均衡系统作出, 令 $r1$ 连接物理主机上运行的虚拟机 sv , 向 $r2$ 连接的物理主机进行虚拟机迁移的决策 (域间迁移), 我们令迁移后目的虚拟机为 dv . 此时我们再假设 sv 的迁移行为, 可为相邻 MAP 关联权重分值贡献 1 (即该虚拟机所在物理机已曾在 1 分钟内与目的物理主机进行过一次虚拟机迁移), 且相邻 MAP 关联权重分值在加后值为 5. 则根据我们的实验假设相邻 MAP 关联权重分值到达 5 已构成 MAP 域间频繁迁移, 应进行 MAP 区域扩大.

迁移设计步骤如下:

(1) 云计算控制引擎作出迁移决策后将其迁移决策发送给 sv 所在物理机的特权虚拟机.

(2) 该特权虚拟机向 MAP 发送路由请求代理信息请求获得 NAR 相关信息.

(3) 在收到 $Rtsolpr$ 后 MAP 发现其为域间迁移则向云计算控制引擎请求目的区域 MAP 信息与路由器路由地址.

(4) 云计算控制引擎向迁移目的主机所在特权虚拟机请求区域 MAP 信息与路由器路由地址, 并将其发送给源请求 MAP.

(5) 源 MAP 即发送一路由通告数据报 $Prtradv$ (Proxy Router Advertisement) 给 sv 所在物理机的特权虚拟机, 该特权虚拟机可据此获取 MAP 信息与路由器路由地址.

(6) 特权虚拟机可用选定获取的 MAP 信息构造 $RCoA$, $RCoA$ 与 MAP 具有相同的 Prefix; 并用获取的路由器路由地址构造 $LCoA$, $LCoA$ 与路由器路由地址具有相同的 Prefix.

(7) 源 MAP 此时会将需传送给 sv 的数据报暂存, 同时将配置地址信息以切换发起数据报 Hi (Handover Initiate) 数据报送给 NAR, 要求其对地址配置进行验证. 收到 Hi 数据报后 NAR 会进行 DAD 检测, 验证其 IP 合法性.

(8) NAR 重复地址检测 DAD 确认地址有效性并发送切换确认 $HACK$ (Handover ACKnowledgement) 作为对 Hi 信息的确认同时用来建立源 MAP 和 NAR 之间的隧道, 使得缓存下来的发向 PAR 的分组发向 NAR.

(9) 随后 sv 开始实际虚拟机迁移, 当 sv 迁移

至新网络时即为 dv, dv 即发送 BU 包给目的 MAP, 此 BU 包包含 RCoA 与 LCoA 信息. MAP 会记录两者相互映射信息, 并将其插入 Binding Table 中, 完成后回传 BA 给 dv.

(10) dv 同时传送 BU 包给源 MAP 与通信节点, 此 BU 包包含 Home Address 与 RCoA.

迁移完成后, 云计算控制引擎开始启动 MAP 动态自适应行为, 设计步骤如下:

(1) sv 上特权虚拟机根据本次迁移行为, 计算在一分钟内是曾否出现本次迁移中源物理机与目的物理机之间的虚拟机迁移, 如曾出现, 则向其 PAR、NAR 递交关联权重分值加 1 的请求.

(2) PAR、NAR 接到请求后, 验证其配对主机是否已对该关联权重分值进行过加分行为, 如有则不予加分, 否则关联权重分值加 1 分.

(3) 云计算控制引擎向 MAP 中特权虚拟机针对迁移至目的 MAP 请求阶段迁移行为记录.

(4) 该特权虚拟机则返回其阶段迁移行为记录给云计算控制引擎, 云计算控制引擎随后根据希望迁移的具体位置向具体物理机上运行的特权虚拟机请求区域 MAP 信息 (其 MAP 信息应包括相邻 MAP 地址信息、本 MAP 地址信息和 MAP 权重分值).

(5) 云计算控制引擎判断 MAP 权重分值, 如果已大于等于 5, 则认为应该进行 MAP 区域扩大改变.

(6) 此次实验我们假定此刻 MAP 权重分值已经等于 5, 则开始进行 MAP 区域改变.

(7) 该 MAP 亦按照虚拟机迁移方式进行, 此刻, 我们依据其迁移位置的变化为相邻 MAP 的性质, 作出 r2 MAP 作为虚拟机迁移合并至 r1.

(8) MAP 迁移完成后, 即云计算控制引擎的左端网络皆定义为 r1 的 MAP 区域. 类似 sv 迁移至 dv 行为已由域间迁移变为了域内迁移.

现进行 sv 迁移至 dv 的虚拟机域内迁移设计步骤讨论:

(1) 云计算控制引擎作出迁移决策后将其迁移决策发送给 sv 所在物理机的特权虚拟机. 该特权虚拟机向 MAP 发送路由请求代理信息请求获得 NAR 相关信息.

(2) 在收到 Rtsolpr 后 MAP 即发送一路由通告数据报 Prtradv (Proxy Router Advertisement), 该特权虚拟机可据此获取 NAR 前缀地址信息.

(3) 特权虚拟机可以将 NARPre-fix 和本物理

机 MACAddress 组成 NLCoa, 再将此地址以 FBU 传给 MAP.

(4) MAP 此时会将需传送给 sv 的数据报暂存, 同时将 NLCoa 以切换发起数据报 Hi (Handover Initiate) 数据报送给 NAR, 要求其对 NLCoa 进行验证. 收到 Hi 数据报后 NAR 会将 NLCoa 用 DAD 检测其 IP 合法性.

(5) NAR 重复地址检测 DAD 确认地址有效性并发送切换确认 HACK (Handover Acknowledgement) 作为对 HI 信息的确认同时用来建立 MAP 和 NAR 之间的隧道, 使得缓存下来的发向 PAR 的分组发向 NAR.

(6) MAP 收到 HACK 之后发送快速绑定更新确认消息给 sv. 此时, sv 通过云计算控制引擎开始实际迁移. 当 sv 迁移至新网络时即发送快速邻居通告消息给 NAR. NAR 收到请求后即转发分组给 dv.

针对该次域内迁移, 云计算控制引擎亦会启动 MAP 动态自适应行为, 设计步骤如下:

(1) sv 上特权虚拟机根据本次迁移行为, 计算在一分钟内是曾否出现本次迁移中源物理机与目的物理机之间的虚拟机迁移, 如曾出现, 则向其 MAP 递交关联权重分值加 1 的请求.

(2) MAP 接到请求后, 验证其配对主机是否已对该关联权重分值进行过加分行为, 如有则不予加分, 否则关联权重分值加 1 分.

(3) 云计算控制引擎向 MAP 中特权虚拟机针对 MAP 请求阶段迁移行为记录.

(4) 特权虚拟机则返回其阶段迁移行为记录给云计算控制引擎.

(5) 云计算控制引擎判断 MAP 权重分值, 如果已大于等于 3, 则认为不用进行 MAP 区域缩小改变.

(6) 此次实验我们假定此刻 MAP 权重分值已经等于 2, 则开始进行 MAP 区域改变.

(7) 该 MAP 亦按照虚拟机分离迁移方式进行.

(8) MAP 迁移完成后, 类似 sv 迁移至 dv 行为已由域内迁移变为了域间迁移.

4 实验验证

现对采用 MDA-FHMIPv6 技术云计算 IPv6 网络进行虚拟机迁移实验验证. 该实验进行虚拟机迁移针对无负载虚拟化系统迁移, 虚拟机未安装任何应用软件. 实验对比模拟了域间迁移与域内迁移两

种情况,迁移花费时间见表 1.

表 1 MDA-FHMIPv6 虚拟机迁移

内存/MB	域内迁移	域间迁移
128	12 331	11 261
256	26 522	23 523
512	48 712	47 238

实验验证系统可完成 MDA-FHMIPv6 虚拟机域间、域内迁移.域间迁移因涉及到云计算控制引擎的较多参与,较域内迁移耗费更多迁移时间.

本文还针对以虚拟机迁移为方法的 MAP 域扩大、缩小调整做了实验验证.耗费时间表见表 2.

表 2 MDA-FHMIPv6 虚拟机 MAP 动态调整

内存/MB	MAP 域缩小	MAP 域扩大
128	10 211	11 021
256	21 322	22 512
512	40 131	39 231

5 结束语

本文描述了一种面向云计算的 IPv6 子网间虚拟机在线迁移系统,该系统基于改进的 FHMIPv6 实现,文中称为 MDA-FHMIPv6.本文中详细叙述了针对 FHMIPv6 改进的 MAP 动态自适应技术,从一定程度上解决了云计算环境虚拟机迁移频繁微移动情况下的效率提升问题.但从上节实验数据发现,当系统进行 MAP 域调整时存在较大时间开销,且该技术的使用增加了系统复杂度,从一定程度上增加了系统耗费.因此,该系统适用迁移行为较为频繁,且迁移呈现定期规律变化的系统.

参考文献:

- [1] Alex C, Snoeren, Hari Balakrishnan. An end to end approach to host mobility[C]//Proc of the 6th Annual International Conference on Mobile Computing and Networking. NY, USA: ACM, 2000: 155-166.
- [2] Bradford R, Kotsovinos E, Feldmann A, et al. Live

wide-area migration of virtual machines including local persistent state [C]//Proc of the 3rd International Conference on Virtual Execution Environments. NY, USA: ACM, 2007: 169-179.

- [3] 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.
- [4] Wood T, Ramakrishnan K, Shenoy P, et al. CloudNet: dynamic pooling of cloud resources by live WAN migration of virtual machines[C]//Proc of the 7th ACM SIGPLAN/SIGOPS International Conference on Virtual Execution Environments. New York: ACM Press, 2011: 121-132.
- [5] Robert Hsieh, Aruna Seneviratne. A comparison of mechanisms for improving mobile IP handoff latency for end-to-end TCP[J]. Proceedings of the Annual International Conference on Mobile Computing and Networking. NY, USA: ACM, 2003: 29-41.
- [6] Shariq Haseeb, Ahmad Faris Ismail. Handoff latency analysis of mobile IPv6 protocol variations [J]. Computer Communications, Nature - Inspired Distributed Computing, 2007(30): 849-855.
- [7] Kong K, Lee W, Han Y, et al. Mobility management for All-IP mobile networks: mobile IPv6 vs Proxy Mobile IPv6[J]. IEEE Wireless Commun, 2008, 15(2): 36-45.

作者简介:



陈俊男, (1979—), 博士, 副教授. 研究方向为下一代互联网技术.

张伟云 女, (1955—), 研究馆员. 研究方向为网络技术.
杨建军 男, (1962—), 副教授. 研究方向为网络技术.