

融合网格的 NGN 增值服务平台及建模分析

李立 杨放春

(北京邮电大学网络与交换技术国家重点实验室 北京 100876)

摘要: 该文提出一种融合网格的下一代网络增值服务平台以解决下一代网络应用服务器的性能问题。该平台将下一代网络应用服务器中的增值业务以网格服务的形式进行封装,并为封装后的增值服务提供运行环境。平台采用一种基于网格技术的任务调度算法,通过利用现有网络上的处理能力以提高系统整体处理能力和性能。采用随机 Petri 网对平台进行建模,并进行性能仿真实验。实验结果说明本平台能较好地达到利用网络中空闲处理能力以提高系统整体性能的目标。

关键词: 下一代网络; 网格; 增值服务平台; 随机 Petri 网

中图分类号: TN915.5

文献标识码: A

文章编号: 1009-5896(2007)11-2545-04

A Grid-Enabled Value-Added Service Platform in NGN and Model

Li Li Yang Fang-chun

(State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876 China)

Abstract: This paper introduces a value-added service platform in Next Generation Network (NGN) enabled by grid technology to solve the performance issue with traditional application server in NGN. All of the value-added services hosted by this platform are wrapped into grid service. This platform provides running environment to value-added service in NGN. A grid-based task scheduling algorithm is also proposed with this platform. This grid-based task scheduling algorithm can solve the system performance bottleneck problem when system gets overload by leveraging existing resources in the same virtual organization with the help of grid technology. This paper models the service platform using Stochastic Petri Net (SPN), and the simulation experiment is done with SPNP (Stochastic Petri Net Package) to test the performance of the service platform according to this model. The simulation result verifies the platform can achieve the goal of improving the system performance.

Key words: NGN; Grid; Value-added service platform; Stochastic Petri net

1 引言

下一代网络(NGN)是业务驱动的网络,它所提供的业务数量和种类远远多于传统电信网络。随着 NGN 中业务的大量涌现,对运营业务的应用服务器的计算能力和性能提出了更高的要求。为提高系统性能,现有应用服务器系统多采用分布式设计,并在此基础上设计了多种负载平衡和过载控制算法。这些方法的出发点是基于应用服务器分布式系统,充分利用系统本身现有资源以提高性能,在某种程度上缓解了系统性能问题。但从全网角度来看,由于业务流量分布的不均衡,运营商仍有大量闲置处理能力未得到充分利用。这就出现了处理能力空闲与处理能力不足之间的矛盾,现有应用服务器并无法很好的解决这一矛盾。

网格技术的出现使得解决上述矛盾成为可能。网格计

算^[1,2]是在动态变化的拥有多个部门或团体的复杂虚拟组织(Virtual Organization, VO)内,灵活、安全地协同资源共享与问题求解。开放式网格服务架构(Open Grid Services Architecture, OGSA)以网格服务的方式对外提供服务接口,将一切网络资源都抽象为“服务”,实现“服务”的共享和协作。从技术角度,OGSA通过提供标准接口使有状态服务和异步交互成为可能,同时将底层跨越组织(或本组织内部)异构、动态的资源抽象为“虚拟”服务,进一步可实现多层次“虚拟”服务集成。

本文提出一种融合网格技术的下一代网络增值服务平台 GSEE(Grid-enabled value-added Service Execution Environment)。GSEE 将传统 NGN 应用服务器中的增值业务以网格服务的形式实现,通过 Web Service 这一标准开放的技术,实现 NGN 业务能力的开放。而且借助于网格对资源协同和共享的支持,GSEE 将运营商的各种资源聚合起来,为 NGN 增值服务平台提供强大、可伸缩的资源,从而解决

2006-04-26 收到, 2006-11-07 改回

国家 973 重点基础研究发展规划项目基金(2003CB314806), 长江学者和创新团队发展计划资助课题

传统应用服务器在处理能力和性能方面的瓶颈问题。

2 融合网格的 NGN 增值服务平台

下一代网络通过开放式API(例如Parlay^[3])实现网络能力的开放, NGN的增值业务(例如点击拨号业务等)通过这些开放接口访问底层网络能力实现自身功能。GSEE将NGN中的增值业务封装成网格服务, 支持服务的组合以完成更高级更复杂的功能。GSEE中定义了基本增值服务(Basic Value-added Service, BVS)和将若干基本增值服务组合而成的高级增值服务(Advanced Value-added Service, AVS), 并提供BVS、AVS的运行环境。GSEE系统的体系结构如图 1 所示。

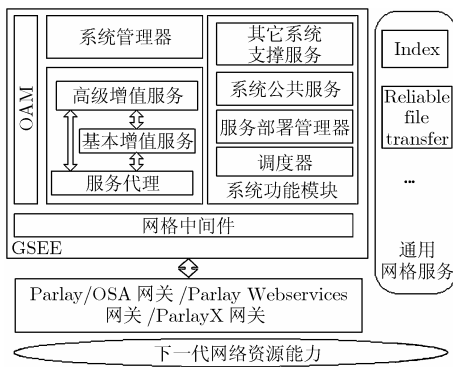


图 1 GSEE 系统体系结构

GSEE 体系结构图中各部件的功能简述如下:

- (1)网格中间件: 为使用网格技术, GSEE 基于网格中间件构建, 例如 Globus。
- (2)服务代理是 GSEE 中增值服务(包括 BVS 和 AVS)的高级代理, 发送给增值服务的请求首先被相应的服务代理获得并进行后续处理。
- (3)调度器是系统任务调度的核心模块。
- (4)系统公共服务提供 GSEE 系统正常工作所需的公共基本功能, 包括日志服务、告警服务等。
- (5)服务部署管理器负责增值服务的部署工作, 包括加载、迁移、更新、停止等。
- (6)系统管理器提供 GSEE 系统工作所需后台管理功能, 包括服务管理、统计管理、计费管理等。
- (7)运行管理与维护模块(Operation, Administration, Maintenance, OAM)提供管理接口方便对 GSEE 的运行管理和维护, 所提供管理接入方式包括 Web, 命令行和 GUI(图形化用户接口)方式。

GSEE采用网格技术, 通过利用各种已有空闲资源实现处理能力的共享和协同以解决传统NGN应用服务器处理能力瓶颈问题。为此, GSEE中定义了服务代理(Service Agent)、调度器和服务部署管理等系统组件。在GSEE中,

通过服务代理搭建灵活的虚拟运行环境^[2]。服务代理作为GSEE中增值服务的高级代理, 发送给增值服务的请求首先被相应的服务代理获得, 服务代理通过与调度器以及服务部署管理器的配合工作, 并借助于通用网格服务所提供的诸如索引服务, 资源分配和管理服务以及可靠文件传输服务, 为请求创建合适的服务实例, 实现资源的共享, 使系统在处理能力出现瓶颈的情况下也能提供有保证的服务质量, 进而解决传统NGN中应用服务器处理能力不足与网络大量处理能力空闲之间的矛盾。

网格环境中, 不同的 GSEE 系统按照特定的共享策略组成虚拟组织, 虚拟组织内的 GSEE 服务器之间可以共享各自的能力。如图 2 所示, GSEE 作为分布式服务平台由若干个独立的物理节点(此处物理节点是指具有处理能力的节点, 例如部署了 GSEE 软件, 处理服务请求的工作站或服务器)组成, 本文将这些 GSEE 系统内部的物理节点称为 Gis(GSEE inner server), 将与本 GSEE 系统同属一个 VO 并提供能力共享的外部服务器称为 Gos(GSEE outer server)。

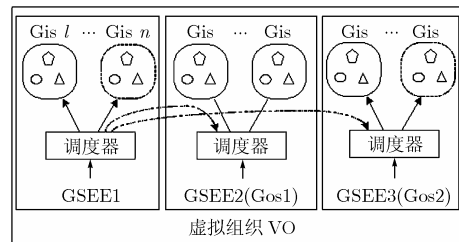


图 2 系统资源共享和任务调度示意图

调度器是 GSEE 任务调度的核心模块, 其对服务请求的调度算法如下:

```

定义(loadlevel == 3)为过载状态
定义(loadlevel < 3)为未过载状态
while(有服务请求到){
  if (存在一个 Gis, Gis.loadlevel < 3){
    选择该 Gis 响应该服务请求;
    break;
  }
  If (对于所有的 Gis, Gis.loadlevel == 3)&&
  (存在一个 Gos, Gos.loadlevel < 3){
    选择该 Gos 响应该服务请求;
    break;
  }
  If (对所有的 Gos, Gos.loadlevel == 3){
    按照预先设置策略过滤该服务请求;
  }
} //end of while

```

该调度算法中 Gis 的负载级别(loadlevel)由 CPU 利用

率、内存占用率等系统性能参数计算而得, Gos 的负载级别定义为 Gos 内所有内部服务节点(Gis)负载级别的最小值。Gos 的负载级别等信息可通过向网格索引服务(Index)检索获得。当 GSEE 过载时,调度器根据调度算法将服务调度到具有空闲处理能力的节点上运行,从而解决系统处理能力不足的问题。可见, GSEE 利用网格技术对资源共享和协同的支持,在系统过载时采用“疏导为主”的方法,将服务请求调度到有空闲处理能力的外部服务器中执行,这相当于为 GSEE 系统进行“动态扩容”,因此能有效提高系统处理能力。

3 系统 SPN 建模

随机Petri网(Stochastic Petri Net, SPN)模型技术的并行、并发、资源共享描述特性和模型分解、压缩技术适合于系统资源管理、请求调度等模型的研究^[4,5]。而这恰好适合 GSEE所具有的特征,因此本文采用SPN对GSEE进行请求调度建模和模型求解,本文利用美国Duke 大学提供的SPNP 软件包作为求解Petri 网的工具^[6]。

在本文的 GSEE 请求调度模型中,请求的产生和服务过程由时间变迁(timed transition)表示,请求进入缓冲队列和共享互斥区由瞬时变迁(immediate transition)表示。瞬时变迁不占用处理时间,但联系随机开关(实施概率)。缓冲队列用位置(place)表示,位置的占有情况由位置中的标识(marking)表示。瞬时变迁比时间变迁拥有更高的实施优先级。允许变迁实施的条件用可实施谓词规定。当满足变迁谓词条件时,变迁可以实施。模型中的请求用位置中的标记(token)来表示。

GSEE 是一个分布式系统。不失一般性,设 GSEE 可以接受 n 类业务请求并包含 m 个内部服务器节点(Gis),与其同属一个虚拟组织的还有 w 个可以共享处理能力的外部 GSEE 服务器(Gos)。对 GSEE 系统模型描述做如下约定 ($1 \leq j \leq m, 1 \leq i \leq n, 1 \leq l \leq w$):

- (1)有 m 个 Gis, 不同的服务请求可有不同的服务时间,使用标识符 s_j 表示第 j 个服务器,服务速率为 μ_j 。服务速率是独立的、依指数分布的。
- (2)系统包含 m 个内部队列,队列使用标识符 q_j 表示,使用下标区别。
- (3)每个内部队列的缓冲空间是有限的,队列 q_j 的容量为 $buff_j$ 。
- (4)有 w 个 Gos, 不同的服务请求可有不同的服务时间,使用标识符 t_l 表示第 l 个 Gos, 服务速率为 v_l 。服务速率是独立的、依指数分布的。
- (5)每个 Gos 具有缓冲队列,该缓冲队列使用标识符 r_l 表示,使用下标区别。
- (6)每个 Gos 的缓冲队列是有限的,队列 r_l 的容量为 $length_l$ 。

(7)有 n 类业务请求,任意一类请求的到达事件为泊松(Poisson)过程,第 i 类请求到达速率为 λ_i 。第 i 类请求可以由本 GSEE 系统的 Gis 节点处理,或调度到 Gos 上处理。

图 3 给出了 GSEE 的随机 Petri 网模型,模型中位置和变迁的含义描述如下:

c_i : 表示 i 类服务请求到达的时间变迁,到达过程符合泊松分布,速率为 λ_i 。

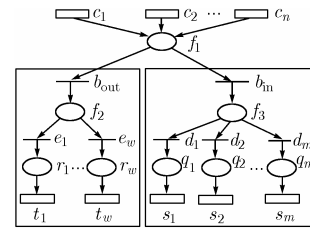


图 3 GSEE 的 SPN 模型

f_1 : 表示判断的位置,它瞬时保留到来的请求,根据 b_m 和 b_{out} 联系的可实施谓词和/或随机开关以决定到来的请求放到哪一个队列。

f_2 : 表示判断的位置,它瞬时保留到来的请求,根据 e_l 联系的可实施谓词和/或随机开关以决定到来的请求放到哪一个队列。

f_3 : 表示判断的位置,它瞬时保留到来的请求,根据 d_j 联系的可实施谓词和/或随机开关以决定到来的请求放到哪一个队列。

b_m, b_{out}, e_l, d_j : 表示分配执行的瞬时变迁,分配方案可由其所联系的可实施谓词和随机开关表达。

q_j : 表示 Gis 服务器 j (Gis _{j})接受请求的队列,其容量限定为 $buff_j$ 。

s_j : 表示 Gis 服务器 j (Gis _{j})的变迁,其实施速率为 μ_j ,实施速率表现为服务器的处理能力。

r_l : 表示 Gos 服务器 l (Gos _{l})接受请求的队列,其容量限定为 $length_l$ 。

t_l : 表示 Gos 服务器 l (Gos _{l})的变迁,其实施速率为 v_l ,实施速率表现为服务器的处理能力。

4 系统性能仿真与分析

为验证 GSEE 使用网格技术能使 GSEE 相较于传统应用服务器(Application Server, AS)有较大的性能提高,仿真实验根据本文所提出的 GSEE 的 SPN 模型,测量 GSEE 的系统吞吐量和 Gos 的系统吞吐量;对 AS 建立 SPN 模型,测量其系统吞吐量,并与 GSEE 系统吞吐量进行比较。

4.1 仿真条件

利用 SPNP 软件包对 GSEE 的 SPN 模型进行计算。与传统 AS 相比, GSEE 能够利用同属于一个 VO 中的资源为自己服务,以提高整体处理能力,因此本试验设 Gis 的数量

为 $2(m=2)$, 并有 1 个 Gos 共享其空闲处理能力($w=1$), 具体模型参数的选择列于表 1。表中位置参数的单位为 call, 变迁实施速率的单位为 call/s。

按照调度算法将变迁 d_j , e_l , b_{in} 和 b_{out} 的可实施谓词和随机开关设置如下, 其中设变迁 d_j 和 e_l 采用随机调度策略。

表 1 GSEE 模型仿真实验参数

参数	m	w	μ_1	μ_2	v_1
	2	1	10	10	8
参数	$buff_1$	$buff_2$	$length_l$		
	5	5	5		

(1) 变迁 d_j 的可实施谓词为 $y_j : M(q_j) < buff_j$ 。

变迁 d_j 的随机开关 g_j 为

$$g_j(M) = \begin{cases} \frac{1}{\|RS_1(M)\|}, & j \in RS_1(M) \\ 0, & \text{其他} \end{cases}$$

其中 $RS_1(M) = \{k \mid M(q_k) < buff_k\}$ 。

(2) 变迁 e_l 的可实施谓词为 $y_l : M(r_l) < length_l$ 。

变迁 e_l 的随机开关 g_l 为

$$g_l(M) = \begin{cases} \frac{1}{\|RS_2(M)\|}, & l \in RS_2(M) \\ 0, & \text{其他} \end{cases}$$

其中 $RS_2(M) = \{k \mid M(r_k) < length_k\}$ 。

(3) 变迁 b_{in} 的可实施谓词为 $y_{b_{in}} : \exists j, 1 \leq j \leq m,$

$M(q_j) < buff_j$ 。

变迁 b_{in} 的随机开关 g_{in} 为

$$g_{in}(M) = \begin{cases} 1, & \exists j, 1 \leq j \leq m, M(q_j) < buff_j \\ 0, & \text{其他} \end{cases}$$

(4) 变迁 b_{out} 的可实施谓词为

$$y_{b_{out}} : \forall j, 1 \leq j \leq m, \exists l, 1 \leq l \leq w, (M(q_j) \geq buff_j) \wedge (M(r_l) < length_l)$$

变迁 b_{out} 的随机开关 g_{out} 为

$$g_{out}(M) = \begin{cases} 1, & \forall j, 1 \leq j \leq m, \exists l, 1 \leq l \leq w, ((M(q_j) \geq buff_j) \wedge (M(r_l) < length_l)) \\ 0, & \text{其他} \end{cases}$$

实验所建立的 AS 的 SPN 模型采用与 GSEE 的 SPN 模型相同的实验参数。该 AS 包括 2 个物理节点负责处理服务请求($m=2$), 每个节点的处理能力均为 10call/s ($\mu_1 = \mu_2 = 10$), 每个节点具有 5 个服务请求的缓冲队列($buff_1 = buff_2 = 5$)。限于篇幅, 不再对 AS 的 SPN 模型进行详述。

4.2 仿真结果及分析

将仿真实验计算的结果绘制成曲线图, 如图 4 所示。图 4 对 GSEE 系统与 AS 系统在相同呼叫条件下的系统吞吐量进行比较, 可以看到 GSEE 的系统吞吐量比同样配置的传统

AS 有了较大提高。例如当呼叫速率为 20call/s 时, GSEE 吞吐量是 AS 的 1.11 倍; 当呼叫速率为 30call/s 时, GSEE 吞吐量是 AS 的 1.32 倍。这是由于采用网格技术的 GSEE 系统在其内部服务器节点 Gis 达到过载情况时, GSEE 系统利用同属于一个 VO 的空闲 Gos 服务器处理呼叫, 从而提高了空闲资源的利用率和系统整体性能, 而利用系统外部空闲资源为自己服务的能力却是传统 AS 所不具有的。

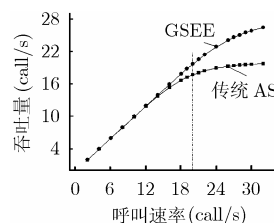


图 4 GSEE 与 AS 吞吐量比较图

5 结束语

本文在充分借鉴网格技术优点的基础上, 提出一种融合网格技术的下一代网络增值服务平台 GSEE。GSEE 将 NGN 中增值服务封装为网格服务并为其提供运行环境。GSEE 利用网格技术对资源共享和协同的支持, 在系统过载时采用“疏导为主”的方法, 将服务请求调度到同属一个虚拟组织的有空闲处理能力的外部服务器中执行, 从而能有效提高系统处理能力和性能。本文基于随机 Petri 网理论对 GSEE 系统进行了系统建模, 并采用 SPNP 软件包对 GSEE 的 SPN 模型进行了仿真实验, 实验结果说明 GSEE 采用网格技术能较好地提高服务平台的整体性能。

参考文献

- [1] Foster I, Kesselman C, and Nick M. The physiology of the grid [J/OL]. www.globus.org, 2002-6-22.
- [2] Foster I, Kesselman C, and Tuecke S. The Anatomy of the grid: enabling scalable virtual organizations [J]. *International J. Supercomputer Applications*, 2001, 15(3): 200-222.
- [3] Moerdijk A J and Klostermann L. Opening the networks with Parlay/OSA: standards and aspects behind the APIs [J]. *IEEE Network*, 2003, 17(3): 58-64.
- [4] 林闯. Web 服务器集群请求分配和选择的性能分析[J]. *计算机学报*, 2002, 23(5): 500-508.
- [5] 林闯, 李雅娟, 单志广. 基于随机 Petri 网的系统性能评价[J]. *清华大学学报*, 2003, 43(4): 475-479.
- [6] Ciardo G, Muppala J, and Trivedi K. SPNP: Stochastic Petri Net Package [A]. *Proceedings of the Petri Nets and Performance Models[C]*. Kyoto: IEEE Computer Society, 1989: 142-151.

李立: 男, 1979 年生, 博士, 研究方向为通信软件与下一代网络、网格计算。

杨放春: 男, 1957 年生, 博士, 教授, 博士生导师, 研究领域为

通信软件、智能网、下一代网络.