



中国科学院软件研究所

杰出青年人才发展专项计划年度进展报告书

(2011 年度)

姓名	钟华	课题名称	面向可信和服务质量的中间件平台模型及关键技术
资助金额	175 万	课题起止时间	2009.08~2013.07
资助类别	应用基础研究	所在部门	软件工程技术研发中心

研究工作主要进展和阶段性成果（含论文、研究生培养等）

课题从基于网络的复杂软件的可信和服务质量等目标着手，研究网构软件中间件平台伺服模型、平台支撑机制和技术，以及开发方法与工具，课题的研究路线如下图所示：

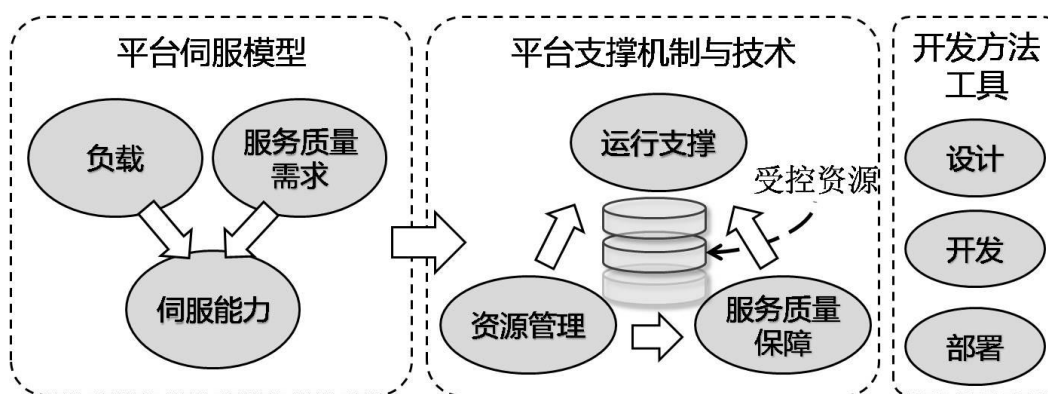


图 1 网构软件支撑平台研究内容

本年度的具体研究进展如下：

(1) 网构软件中间件动态建模技术

网构软件是一类开放、动态和难控的软件形态，它的性能受到多个方面的影响，包括用户的使用、应用的实现、资源管理与调度的方式，以及底层硬件资源的计算能力。同时，网构软件对软件的质量以及资源的充分利用提出了新的要求。因而，如何在多变环境下，持续保障系统性能，按需供给资源，就成了一个研究重点。本研究旨在网构软件及其运行平台构造一个可控的仿真环境，定量分析各种影响因素及资源配置策略对系统带来的影响，诊断分析性能瓶颈的原因，搜索最优的资源配置方案。

性能模型需要获得运行时影响应用性能的若干属性，但是在应用运行时难

以获得相关属性，传统的建模方法大多依据模型设计人员的经验或者在测试环境中获取这些属性。这类方法的主要弊端是难以真实反映应用实际运行特征，并且也不能满足网构软件开放、动态的需求。

针对上述问题，本研究重点研究如何在运行时，通过分析系统日志信息，动态构造性能模型。可分为应用组件、中间件容器和物理资源消耗三个层次。在应用组件层主要研究如何获取“执行流程”，“请求类型”和“参数依赖”这三个与应用组件逻辑相关的性能要素；在中间件容器层主要研究如何获取与中间件平台相关的“逻辑资源”和“内部状态”，并与应用层性能要素结合，形成与平台相关的应用模型；最后分析各组件的“物理消耗”，本文主要关注 CPU 资源的消耗情况。最终，将形成一个与应用和具体部署和配置环境相关性能模型。

（2）网构软件中间件极限事务处理

网构化软件的用户规模呈现出动态、开放、不可预计等特性，许多企业或组织时常会面对“极限”场景的应用需求：百万级的并发用户访问、每秒处理数以千计的并发事务、灵活的弹性与可伸缩性、低延时及 7*24*365 可用性等。此类应用被称为极限事务处理型应用。以互联网企业 Twitter 为例，Twitter 每秒接受近 40,000 次用户请求，处理超过 4M 的用户数据，注册用户数超过 1,000,000。Gartner 将极限事务处理定义为一种应用模式，旨在对事务型应用的开发、部署、管理和维护提供支持，特别是对性能、可扩展性、可用性、可管理性等方面的极限需求。Wikipedia 认为极限事务处理是每秒多于 500 次事务或大于 10,000 并发访问的事务处理。Gartner 在其技术报告中指出，极限事务处理型应用的规模已由 2005 年的 10% 提升到 2010 年的 20%，极限事务处理技术是未来 5-10 年的热点技术。极限事务处理技术已成为大型金融、电信、零售类企业应用以及社交网站的核心技术。

传统基于层次的应用引擎架构（分为展现层、业务逻辑层、数据层）主要利用集群技术进行扩展，难以达到线性扩展效果，无法满足极限事务处理场景的需要，具体体现为以下几点：1) 每层都由 1 个集群组成，各层具有异构的集群模型和 H/A 模型，难于统一监控、管理和部署；2) 集群同步及层间交互开销大，网络易成为主要瓶颈；3) 高可用/高可靠实现机制复杂；4) 各层独立购买授权 (licence)，实施成本高。极限事务处理的需求决定了我们需要提供具有如下特

点的新型应用引擎：1) 低延时；2) 线性可扩展；3) 高可用高可靠；4) 易于维护和管理。

打破旧有的层次架构，建立适用于极限事务处理的新型应用引擎架构，成为该领域发展的主要趋势。例如，Gigaspace XAP 将业务逻辑、消息与数据组合为一个高内聚的处理单元 (processing unit)，处理单元成为集群管理和伸缩的基本要素，提供线性扩展的事务处理能力且易于维护；所有事务/计算处理均在内存中执行，最大限度的降低了层间通讯开销，提高了事务处理的效率；高可用高可靠的实现以处理单元为粒度，实现机制简单。

(3) 获益驱动的网络软件中间件资源管理技术

Web 系统 QoS 保障是学术界研究的热点。随着虚拟化技术的发展和云计算的兴起，虚拟机 (Virtual Machine) 作为 Web 系统宿主运行环境逐渐成为一种典型架构模式，而基于虚拟机粒度的动态资源分配则成为保障 Web 系统 QoS (Quality of Service) 的一种有效机制。然而基于虚拟机的重配置操作具有多样性，包括容量控制、虚拟机复制和在线迁移三种；基于虚拟机的重配置具有耗时长并导致 Web 系统性能衰减的特点，表 1 描述了测试基础 TPC-W 执行在线迁移操作的性能衰减和耗时情况；虚拟机配置具有异构性，表 2 描述了 Amazon EC2 云平台提供的虚拟机资源配置类型。上述特点使得当 Web 系统负载变化时，执行哪种资源重配置操作和选择哪种虚拟机配置资源才能同时达到最小化资源成本和保障 Web 系统 QoS 成为挑战。

针对资源成本和资源重配代价约束下的资源最优分配选择问题，特别针对已有成本敏感 (Cost-aware) 方法因忽略了 Web 系统在新配置下的稳定周期而导致在负载突变 (Flash Crowds) 模式下频繁违反 QoS 的不足，提出了一种收益敏感 (Benefit-aware) 的资源分配方法。该方法首先将 Web 系统在新配置下的稳定周期描述成收益，因为计算能力的不足是最大社交网站 Friendster 衰落的重要原因之一，这说明 Web 系统在资源重配置过程中付出代价的同时也隐含着潜在的收益；然后将 Web 系统重配置操作引起的资源成本的变化描述成设施成本，将 Web 系统重配置操作引起的 Web 系统性能衰减并违反 QoS 约束描述成操作成本，并根据加权原则将设施成本和操作成本描述成 Web 系统执行某资源重配置操作的总成本；最后根据将“收益/总成本”最大化原则作为资源最优分配的判定标准。

基于收益敏感(Benefit-aware)的资源分配方法主要包括负载监测、负载预测、资源规划和资源调整 4 个功能模块和性能模型、性能衰减模型 2 个模型。Web 系统的用于估算当前负载下的响应时间；Web 系统性能衰减模型用于估算当 Web 系统在某负载下执行指定资源重配置操作过程的相应时间变化。当 Web 系统对外提供服务的过程中，负载监测模块周期性的检测 Web 系统当前负载；当负载超出期望负载波动范围时，采用负载预测模块预测新的期望负载范围；资源规划模块首先将期望负载输入到性能模型，并将估算的响应时间与 QoS 约束作对比，当 QoS 不满足时，资源规划模块会评估每种资源重配置操作的耗时及其导致的性能衰减情况，然后根据将“收益/总成本”最大化原则进行资源分配决策；资源调整模块将资源规划模块决策结果作用到 Web 系统的运行载体虚拟机上。

本方法针对 Web 系统具有多层架构的客观事实，采用 LQN 性能模型作为 Web 系统的性能模型，LQN 模型具有刻画 Web 系统不同组件服务时间和等待队列的能力，并能根据队列长度估算出 Web 系统在某一负载下的响应时间；Web 系统性能衰减模型采用性能衰减表来描述，通过离线对 Web 系统不同组件在典型负载下执行不同资源重配置操作的方式，将资源重配置操作过程中响应时间的变化记录到性能衰减表，并根据负载近似匹配原则估算 Web 系统在某负载下执行指定资源重配置操作的响应时间变化；基于 ARMA 模型来预测某种负载的有效周期，Web 系统负载具有相关性、时变性、随机性、平稳性，可以看作是平稳的随机序列，而 ARMA 模型则是描述平稳随机序列的最常用的一种模型；收益函数与 Web 系统在新配置下的稳定周期及其在稳定周期的单位收益有关，用于刻画 Web 系统执行某资源重配置操作后租户的满意程度。

(4) 网构软件中间件多租户支撑技术

Internet 的开放性导致网构化软件的用户负载规模难以预计，80%用户日常访问规模往往远低于峰值规模，这些用户组成了数量庞大的“长尾用户”。为满足“长尾用户”的服务需求而配备独立的基础设施，已经造成了极大的能源、设施的浪费，并带来巨大的维护成本，因此产生了多租户共享的应用系统和服务模式。多租户(multi-tenancy)将来自不同组织(租户)的并发用户共享同一基础设施资源，有利于成本降低和收益提高。

目前存在四种典型的多租户模式：硬件级共享，操作系统级共享，中间件级

共享以及应用级共享. 应用级共享的多租户模式提供了更高层次的基础设施共享, 性能开销更低, 可以承载更多的租户, 即租户扩展性更高, 因此是一种高效能的多租户模式, 已成为下一代企业计算的发展方向. 然而, 开放环境下的租户难以完全信赖, 即使同一组织内部、相互信赖的多个租户, 仍有可能由于误操作或过载而导致个别租户侵占大量资源, 进而影响其它租户的访问性能, 需要通过性能隔离 (Performance Isolation) 技术来避免此类性能干扰 (Performance Interference) 现象. 解决这一问题, 仍面临诸多技术挑战.

首先, 下一代的企业计算设施面临细粒度的应用级资源管理需求. 租户性能的隔离取决于租户资源的隔离. 在已有工作中, 一些基于事件驱动的资源管理模型被提出, 其中的代表工作是基于阶段的事件处理架构 SEDA. 由于采用白盒式的资源管理架构, SEDA 支持细粒度的应用级资源管理. 但是, 此类资源管理模型并不适用于典型的、基于多层架构的事务型应用, 这是因为此类应用的事务处理过程涉及的应用组件或服务调用仍以同步调用为主, 而采用事件驱动的架构需重构交互机制, 代价巨大.

另外, 实现租户性能隔离的关键是控制租户对系统资源的使用, 避免个别租户侵占过多资源. 对于应用级共享的多租户模式而言, 性能干扰将导致所有租户性能下降, 因此, 难以通过租户的性能表象 (如响应时间) 检测个别租户的资源侵占行为, 需要在线获取并分析不同租户的资源使用情况. 然而系统资源的在线度量, 尤其是 CPU 资源的在线度量, 仍存在技术挑战.

针对上述问题, 课题组提出一种细粒度的应用级资源管理模型 (Application-level Resource Management Model, 简称 ARMM), 依据多租户应用的上下文信息构造事务处理链, 对并发事务处理进行租户区分和阶段划分, 并基于信号量机制给出了 ARMM 模型的并发控制算法, 在一次事务处理的各阶段实现线程复用, 避免已有方法易产生的交互协议重构问题. 上述工作为进一步实现租户性能隔离提供细粒度的操控行为.

接着, 课题组针对租户资源使用情况的在线获取问题, 提出一种基于统计分析的 CPU 资源在线评估方法. 方法基于生产环境中常用的监测数据, 利用 Kalman 滤波在线评估多租户的 CPU 资源使用情况, 具有无性能开销、非侵入和操作系统独立等优点. 实验结果表明, 方法可动态适应持续变化的负载环境, 并且降低了事

务型负载普遍存在的多元共线性现象的影响,与直接度量方法相比,具有可接受的评估误差.上述工作为租户性能隔离提供了依据.

(5) 基于负载模式识别的异常诊断方法

以 JEE 应用为代表的事务型 Web 系统已经成为网络服务主要的提供方式.然而,面对复杂、难控的 Internet 环境,以及 Web 系统本身的复杂性,特定环境下出现的上下文异常难以避免,例如,某个请求引发未经测试的内存泄露,或是在某种访问序列下并发线程竞争共享资源造成的死锁等.这些异常往往会在一段时间后引发系统失效,而可能引起巨大的经济损失.据 Tellme Networks 报告指出,失效恢复时间的 75%用在检测错误,18%用在诊断错误.另有研究表明,早期检测到失败能够缓解或阻止 65%的错误发生.因此,Web 系统的异常检测是消除或缓解服务失效的有效手段,对于提高网络服务质量非常关键.

然而,已有的 Web 系统的在线异常诊断方法还难以适应开放的 Internet 环境.工业产品中通常采用基于规则的方法通过设置各度量的阈值,当监测的度量超出阈值时根据规则执行预定动作,如 IBM Tivoli、HP OpenView、EMC Smarts 等,但是,系统度量数量繁多,人工设置困难.学术界采用的典型方法是建立一个模型刻画系统状态,当检测与预期状态发生偏差则认为可能有错误出现.现有的方法利用特定的分析形式,如线性,建立度量间关联关系.但度量间存在各种关系难于用统一的模型刻画,同时,负载的变化使得建立的关联关系改变.

值得关注的是,负载具有较强的动态性和周期性,呈现出 time-of-day 或 month-of-year 效应.我们提出一种基于负载模式识别的 Web 系统在线异常检测方法.该方法通过在线增量式聚类建立负载模式及其对应的度量空间,通过对当前负载进行模式识别,基于其对应的度量空间进行异常检测,这就解决了负载变化造成的异常难于准确检测的问题.并且,利用局部异常因数检测系统异常状态和量化异常程度,通过监测数据点的相对位置来发现问题,解决了度量间多种关联难于刻画的问题.而后,对于检测到的系统异常状态,通过学生 t 测试方法计算各度量的异常值以定位问题原因.具体而言,可以分为负载模式识别和异常检测两部分.

工作负载通常由会话组成,根据会话特征可以分为不同的访问模式,将每个请求作为状态用节点来表示,状态间的转移概率表示访问模式.我们引入负载

向量，即周期性统计组件间调用次数，以表示系统的负载状况，不仅反映了访问模式，而且表现了请求密度。具有相似资源需求的负载划分为一种负载模式，可以通过负载向量间的相似度来衡量。负载模式具有动态性难以预知，我们利用增量式聚类在线学习负载模式。

Web 应用组件的调用序列和请求密度反映了工作负载，而 Web 应用组件执行造成物理资源的消耗，表现为度量的变化，如 CPU、内存、网络利用率和磁盘读写频率等。因此，工作负载与系统度量之间存在着对应关系，我们引入运行时状态由工作负载和系统度量共同刻画。在运行过程中，我们将负载向量与已知负载模式进行匹配，从而获得其对应的度量空间，利用局部异常因数根据度量空间计算当前度量向量的异常分数；当判定异常时，则利用学生 t 测试对各个度量分别考察以定位异常度量。我们通过监测度量的变化检测 Web 系统运行过程中出现的异常状态，同时可以通过负载模式识别检测出异常负载变化。

本年度主要研究成果包括：

- ✓ 获国家科技进步奖二等奖 1 项（个人排名第 5）。
- ✓ 获自然科学基金面上项目一项（60 万）。
- ✓ 在国内核心学术刊物和国际会议上发表相关论文 12 篇，其中 SCI/EI 收录 7 篇。
- ✓ 获得专利授权 3 项，申请专利 8 项。
- ✓ 毕业硕士研究生 2 名。

本年度国际交流合作情况：

- ✓ 赴德国慕尼黑参加 COMPSAC 2011 国际会议。
- ✓ 邀请海外学者来中心访问 1 人次（秦锋，俄亥俄州立大学计算机系助理教授，2011 年 6 月 10 日到 7 月 9 日在中心短期访问）。

文章列表：

- 1 Zhezhe Chen, Xinyu Li, Jau-Yuan Chen, Hua Zhong, Feng Qin, SyncChecker: Detecting Synchronization Errors Between MPI Applications and Libraries, IPDPS 2012, accepted.

- 2 Xiulei Qin, Wenbo Zhang, Wei Wang, Jun Wei, Hua Zhong, Tao Huang, A Comparative Evaluation of Cache Strategies for Elastic Caching Platforms, In Proceedings of the 11th International Conference on Quality Software (QSIC 2011), 13-14 July 2011 Madrid, Spain, pp.166~175.
- 3 Xiulei Qin, Wenbo Zhang, Wei Wang, Jun Wei, Hua Zhong, Tao Huang, On-line Cache Strategy Reconfiguration for Elastic Caching Platform: A Machine Learning Approach, In Proceedings of the The 35th International Computer Software and Applications Conference (COMPSAC 2011), July 18-21, 2011 Munich, Germany.
- 4 Wei Wang, Xiang Huang, Yunkui Song, Wenbo Zhang, Jun Wei, Hua Zhong, Tao Huang, An Statistical Approach for Estimating CPU Consumption in Shared Java Middleware Server, In Proceedings of the The 35th International Computer Software and Applications Conference (COMPSAC 2011), July 18-21, 2011 Munich, Germany.
- 5 Wenbo Zhang, Sa Wang, Wei Wang, Hua Zhong, Bench4Q: A QoS-Oriented E-commerce Benchmark, In Proceedings of the 35th Annual IEEE International Computer Software and Applications Conference (COMPSAC 2011). July 18-22, 2011, Munich, Germany.
- 6 Guoquan Wu, Jun Wei, Chunyang Ye, Xiaozhe Shao, Hua Zhong, Tao Huang, Runtime Monitoring of Data-centric Temporal Properties for Web Services, In Proceedings of IEEE International Conference on Web Services(ICWS)2011, Jul 5-10, 2011, Washington, USA.
- 7 Guoquan Wu, Jun Wei, Chunyang Ye, Xiaozhe Shao, Hua Zhong and Tao Huang, Runtime Verification of Data-Centric Properties in Service based Systems, In Proceedings of the 2 Runtime Verification 2011.
- 8 朱鑫 秦秀磊 王联华 张文博 钟华. 弹性分布式缓存动态扩展方法研究. 2011 年全国软件与应用学术会议 (NASAC), 2011 年 10 月 28-30 日, 中国长春.
- 9 孙国洋 杨燕 钟华. 支持 JSR286 和 WSRP2.0 规范的 Portlet 协作的设计. 计算机应用与软件, 2011, 28(5): ~.
- 10 尼建 王帅 杨燕 钟华. 一种多门户环境下服务管理框架的设计. 计算机应用与软件, 2011, 28(6):93~95.
- 11 詹孟粮 马志柔 钟华 王鹏. 一种高可靠性的组播树恢复方法. 计算机应用与软件.
- 12 李辉 杨燕 王帅 白琳 钟华. 一种支持声明式表示层集成的组件模型. 计算机系统应用

授权专利列表:

- 1 一种 Java EE 应用服务器并发处理方法, ZL200810117820.3
- 2 线程池容量自适应调节方法及应用服务器并发控制方法, ZL200810119285.5
- 3 一种面向应用服务器的资源敏感性能优化方法及其系统, ZL200810119278.5

申请专利列表:

- 1 一种 Web 应用细粒度性能建模方法及其系统, 201010275216.0
- 2 一种软件生产线构造方法及系统, 201010279066.0
- 3 基于 OSGi 的软件构件监测方法与系统, 201010523324.5
- 4 一种基于 J2EE 的数据持久化方法及系统, 201010562533.0
- 5 一种面向 Web 应用宿主平台的资源供给方法, 201010578793.7
- 6 支持 Portlet 协作构建复杂业务逻辑的方法和系统, 201010578929.4
- 7 一种支持负载均衡的分布式缓存动态伸缩方法及系统, 20110230333.X
- 8 一种基于机器学习的分布式缓存策略自适应切换方法及系统, 201110167018.7

下一年度工作计划, 包括国内外合作与交流计划

下一年度研究内容包括:

- ✓ 网构软件中间件服务动态协同配置机制研究。
- ✓ 网构软件中间件部署单元动态迁移机制研究。
- ✓ 网构软件中间件监控管理工具研发。
- ✓ 面向特定领域的网构软件中间件可信与服务质量保障测试规范研究。

进度计划安排如下:

- ✓ 2012 年 11 月~2012 年 5 月研究网构软件中间件服务动态协同配置机制与动态迁移机制。
- ✓ 2012 年 3 月~2012 年 10 月研制完成网构软件中间件监控管理工具。

✓ 2012 年 7 月～2012 年 12 月面向特定领域的网构软件中间件可信与服务 质量保障测试规范。 国际合作交流安排如下： ✓ 赴欧美参加高水平国际会议 1—2 次。 ✓ 邀请欧美或港澳台学者来访 1—2 次。
年度经费使用情况概要
本年度使用经费 112328.9 元，主要用于资料费、会议费、差旅费等。研发 所需的设备费、文献知识产权、劳务费等主要在相关的其他科研项目中进行了支 出。
存在的问题、建议及其他需要说明的情况
Empty box for content

备注：1.杰出青年人才入选者应实事求是地填写此报告，禁止弄虚作假；
2.此报告作为专项计划跟踪、管理的主要依据，每年报送人力资源处；
3. 此报告将在所网站对外发布。