

基于网格技术的密码学计算平台的设计*

任俊伟, 林东岱, 姜中华

(中国科学院 信息安全国家重点实验室; 中国科学院 研究生院, 北京 100080)

摘要: 结合现有的网格技术、思想和密码计算的特点, 分析了利用网格技术实现密码学计算的可行性, 并结合 J2EE 开发方式的优势实现系统支持, 最后提出了一种有效的、可行的密码计算网格平台构架。

关键词: 网格; 网格计算; 密码计算; 分布式

中图分类号: TP309.7 **文献标识码:** A **文章编号:** 1001-3695(2005)08-0030-03

Design of Cipher Computing Platform Based on Grid Technology

REN Jun-wei, LIN Dong-dai, JIANG Zhong-hua

(State Key Laboratory of Information Security of Software; Graduate School, Chinese Academy of Sciences, Beijing 100080, China)

Abstract: Combining recent idea and technology of grid computing, also by considering the characteristic of cipher computing, this paper analyzes the feasibility of using grid technology to achieve efficient cipher computing. Last an efficient and feasible construction of grid computing platform is given for cipher computing by using J2EE as the system support.

Key words: Grid; Grid Computing; Cipher Computing; Distributed

随着网络技术和信息技术的飞速发展, 信息安全的重要性日趋突出, 而作为安全技术基础的密码学算法, 对其强度的要求也在不断提高。大计算量、高复杂性已经成为密码学算法实现有效与安全的不争前提。比如现行的对称密码体制和公钥密码体制, 前者的算法要求较大的密钥空间(DES 的密钥空间为 2^{56}), 后者的算法大都基于大整数分解和离散对数等复杂的数学问题, 算法的安全性基于计算的复杂性。相应的, 与密码学应用相关的密码分析, 密码算法强度测试以及密码攻击, 都会牵涉到大计算量的问题。因此, 如何在现有条件下为应用提高计算速度成为密码学研究的重要课题。

在此, 基于大计算量这样一个共性, 可以把密码计算理解为与密码学应用相关的所有数学计算问题。国内外目前大都采用大型机、并行计算和分布式计算来提高密码计算的速度。它们存在以下特点和不足: 大型机价格昂贵, 一般机构无力承担如此高昂的费用; 并行计算一般局限于一个或几个实验室, 能利用的资源有限; 分布式计算一般侧重利用企业或研究机构内部的局域网计算资源。这些技术都存在可利用计算资源有限, 难以扩充的缺点。新兴的网格技术的提出为解决这些不足, 实现有效的密码计算提供了一种可行的方案。

1 网格及密码计算简介

1.1 网格的含义和特性

网格是伴随着互联网技术而迅速发展起来的, 最初是专门针对复杂科学计算应用的一种新型计算模式。这种计算模式是把整个网络整合成一台巨大的超级计算机, 利用大量分布的

存储资源、计算资源完成分布式超级计算、地球观测、DNA 分析、分布式数据挖掘等高复杂应用。有别于传统的客户机/服务器系统, 网格应用使用大量的网格资源, 要求资源能动态请求和分配, 协同资源共享管理和复杂通信, 具有可扩充性、动态性和异构性。随着网格技术的发展和应用的扩展, 网格逐步进入到商业使用中。人们预言, 网格技术和应用将成为 Internet 信息技术的下一个浪潮, 网格将成为具有高性能处理、海量数据存储和大量仪器设备等特征的 21 世纪人类社会的信息处理基础设施。

实际上, 可以从三个方面来理解网格: 网格是构建信息系统新的思维方式和新的概念; 网格计算技术是解决广泛的分布资源共享和协同工作的新技术, 可以实现更广范围的分布式计算; 网格是支撑各类应用的新的国家信息基础设施。

现有的网格计算环境多种多样, 其中以 Globus 技术最为成熟, 使用最为广泛。它是一个基于社团的开放源码集合, 最新的 Globus Toolkit 3.0 采用了以服务为中心的开放网格服务体系结构(Open Grid Service Architecture, OGSA)。而 OGSA 体系结构结合了 Web Services 的大量成熟技术, 包括 XML, SOAP 和 WSDL 等网络服务标准的使用, 同时支持服务的动态创建、服务状态、生命期管理、通知、日志和安全等特性, 是实现网格计算平台构建的理想体系结构。

1.2 密码计算的特点

密码计算的基本特点是计算量很大, 但是需要的存储量相对较小。各种密码计算(如加密和解密)都是对一些基本计算(如大整数运算、代换和置换等)的多次重复。例如, 对 DES 算法已知明文的穷举攻击, 需要穷举 DES 的密钥空间, DES 的密钥空间为 $2^{56} = 72\ 057\ 594\ 037\ 927\ 936 \approx 7 \times 10^{16}$, 因此平均需要 3.5×10^{16} 次加密运算。其中的每一次加密运算与其他的加密运算间没有依赖关系, 可以分布在网格的任一个节点上进

收稿日期: 2004-07-16; 修返日期: 2004-08-31

基金项目: 国家自然科学基金资助项目(NSFC90204016); 国家“863”计划资助项目(2003AA144030)

行。同时可以将大小为 2^{56} 密钥空间分成若干个小的区间(如每个区间为 1 000 个密钥), 这样将 DES 密钥空间分成了 7×10^{13} 个小区间, 即分成了 7×10^{13} 个小任务, 分布到 1 000 个计算节点上(实际上网格计算中的节点数将远远大于 1 000), 每个计算节点平均需要进行 7×10^{10} 次小任务的计算。这样计算时间将会是原来的 1/1000。从而极大地提高了密码计算速度。

对于所有的分组密码算法, 都采用了代换和置换两种基本操作, 其区别主要是密钥的长度不同, 数据分组长度不同, 迭代的轮次不同, S 盒的设计不同, 因此都可以用类似的方法处理。如 Rijindael 算法的最小密钥长度是 128 位, 分组长度是 128 比特, 最小轮数为 10 轮, 并且将其 S 盒映射到 8 比特。

对于公钥密码体制, 大都利用了一些数学难题, 主要基于大整数的因子分解和离散对数在计算上的困难性。大整数的因子分解有一些方法, 其中三种最有效的算法是二次筛(Quadratic Sieve)、椭圆曲线分解算法(Elliptic Curve Factoring) 和数域筛法(Number Field Sieve)。二次筛对于小于 125 130 比特的十进制数的分解是一个较好的算法。该算法可以分布式分解大整数。资料显示椭圆曲线密码体制也能使用分布式计算方法。

对于一些流密码也可以使用分布式计算缩短加解密运算时间。对于一些比较费时的加密和解密运算也可以在密码计算网格上进行分布式计算。

2 结合网格技术实现密码计算平台的设计

2.1 系统需求

(1) 可利用的资源。一类是平台开发者提供的专用于密码计算的 PC 机, 它们的操作系统为 Linux; 另外一类是 Internet 上具有空闲计算资源的桌面办公用计算机, 它们的操作系统绝大部分为 Windows 系统。对于办公用计算机, 我们可以使用 C 或 C++ 编写计算代码来提高计算速度, 这些计算机通过互联网互连。

(2) 密码计算的需求。所进行的密码计算对资源需求大, 对存储空间需求小, 分布计算的结果可以合并, 分布的小计算任务之间没有依赖关系。这些特点使其能方便地进行分布计算。除了系统提供的常用密码计算服务外, 用户或第三方也可以自己根据需要设计新的计算任务, 然后动态地部署到网格中。

(3) 利用现有的 GT3(Globus Toolkit 3.0) 网格计算开发环境。它解决了分布式计算的安全性和动态的资源共享和分配能力, 利用它来实现异构的跨平台、跨组织的资源(包括计算资源) 共享和管理。

(4) 系统平台的参与者, 即服务使用者(客户或用户)。

使用 Web 接口的用户。他们使用系统提供的 Web 界面完成常见的密码计算任务。

下载或编写客户端代码, 启动客户端程序来访问服务的用户。这些用户利用系统提供的 API, 通过一个在客户端运行的程序来实现服务访问的交互。

他们的对比是 B/S 系统与 C/S 系统的对比, 我们选择对用户更为友好, 更容易为用户接受的 B/S 体系结构, 即客户以

Web 方式访问服务。

贡献 Windows 计算节点的人, 他们愿意把自己的计算资源贡献给密码计算网格, 他们是 Internet 资源的贡献者。

新计算服务代码提供者。这些服务代码可能为实现新的密码算法、对某种密码算法进行攻击或测试而设计。它需要通过管理员的协助加入到密码计算网格中。

系统管理员。其职责为复杂系统维护和灾难恢复, 用户管理, 新计算任务的部署, 日志审查等。

密码计算网格系统总体构架如图 1 所示。

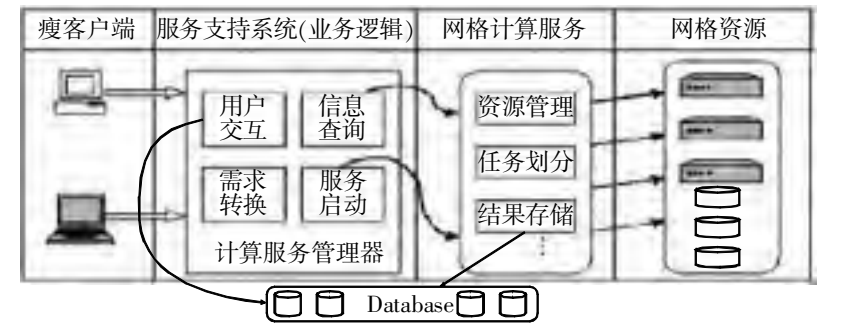


图 1 密码计算网格系统总体构架

2.2 计算服务流程

采用 N 层结构设计方法: 瘦客户端 + 运行支持环境 + 网格服务和网络资源 + 数据库。用户登录服务支持系统的服务器网页, 获知可用计算服务, 选择需要的一个。

- (1) 根据选中计算服务要求填写信息(提供服务所需参数、用户信息等), 提交;
- (2) 服务支持系统完成必要业务逻辑处理、数据转换、服务启动工作并通知用户计算已开始执行;
- (3) 网络计算服务子系统独立地进行资源管理、任务划分和结果存储(到数据库系统);
- (4) 网络计算子系统通知服务支持系统任务完成, 服务支持系统通过查询数据库将结果反馈给用户。

2.3 实现方法

- (1) 以 Web 浏览器作为用户交互界面, 以瘦客户方式获取服务;
- (2) 利用 J2EE 技术实现服务支持系统, 实现业务逻辑处理和基本的数据库访问;
- (3) 利用网格技术实现计算服务子系统, 实现资源管理, 服务部署与发布、子任务划分、计算结果存储。

2.4 组成模块

- (1) 瘦客户端(IE)。将用户的负担降到最小, 一切工作都在服务器端进行, 用户只需简单地利用网页实现交互。
- (2) 服务支持子系统(结合 J2EE 技术和 GT3 提供的部分服务实现): 用户交互子模块与计算用户交互, 包括请求任务时的交互以及任务执行过程中、任务执行完毕后计算状态和计算结果的查询; 需求转换子模块, 对用户输入进行逻辑匹配(业务逻辑处理), 弥补不足信息, 解决用户信息与网格信息差异; 信息查询子模块, 对 MDS(Monitoring and Discovery Service) 基本信息、计算服务接口信息、计算服务部署信息进行查询; 运行管理子模块, 计算环境的准备和恢复, 计算程序的远程启动、运行情况的监视和控制。
- (3) 网络计算服务、网络资源子系统(GT3 网格开发环境实现): 计算服务的提交和添加。计算服务包括密码计算中

的加密和解密,常见的密码攻击,常见的密码基本运算(大整数分解、离散对数求解、通用的代换和置换运算等),以及基于常见密码基本运算的其他密码计算。计算状态和计算结果的存储。该功能主要存储计算的完成情况、计算的状态、估计计算完成的时间和计算结果,供支持系统查询并反馈给用户。

系统管理功能,如用户管理、证书的管理等任务。功能的描述和系统说明。这个功能被网格的所有参与者使用,为用户规范使用服务提供依据。新的计算节点的添加。Windows 桌面计算机可以被动态地加入到网格中,为网格提供计算资源,参与密码计算。

(4) 数据库子系统。为支持子系统和网格计算服务子系统所共享,为用户管理信息和服务结果提供可靠存储,用于支持数据存储和信息反馈。为提高安全性和鲁棒性可以实现多个数据库映射,以实现数据的多备份。

2.5 计算服务实现特点

- (1) 对计算程序的抽象;
- (2) 提供计算服务接口作为访问调用基础;
- (3) 屏蔽计算程序实现;
- (4) 计算服务实现(程序)由服务提供者在一个或多个计算节点上事先部署、调试、试运行;
- (5) 部署好的计算程序可以随时启动。

部署结构图如图 2 所示。

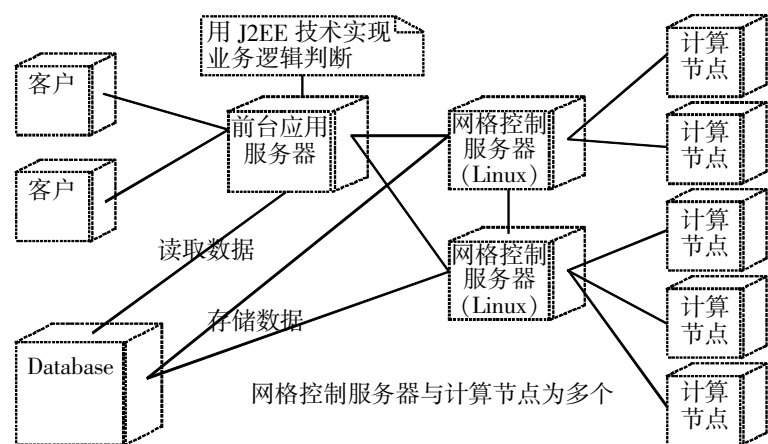


图 2 部署结构简图

2.6 系统的评价及总结

系统的出发点有三个:

(1) 尽量使用可以使用的计算机资源,它包括自己拥有的计算机资源和 Internet 上空闲的计算机资源,从而提高密码计算的速度。

(2) 系统尽量方便用户的使用。提供基于 Web 的人机接口,使得密码计算用户可以在各种操作系统和各种硬件平台下都可以使用密码计算网格,同时简化了我们对客户端的重复开发。

(3) 尽力提高系统的安全性。尽量利用 GT3 提供的安全性来保证系统的安全。另外,对于 Web 用户与应用服务器之间可以使用 SSL 与其他的多种安全措施来保证认证和授权。在应用服务器与数据库之间可以采用 SSL 和数据本身提供的安全措施保证它们之间的安全通信。在网格节点之间的认证采用了 GT3 的证书机制。待进一步解决的问题如下:

(1) 如何将新的计算服务对应的代码透明地加入到各个网格节点上,其中的一个问题是如何将代码传输到网格节点

上,其二是在不中止计算任务的情况下把新的计算任务部署到计算机节点上去。

(2) 如何使用通用的方法将各个计算任务用一种统一的方式组织起来,主要解决应用服务器的数据接收的统一性。

(3) Windows 计算节点的组织问题,采用什么机制完成任务接收和软件更新。

(4) 如何将 Linux 计算节点分布到互联网上,以及对这些节点的管理问题。

(5) 网格系统用户的授权问题。因为系统设计多种用户,来自多个组织,如何设计有效的授权策略。

(6) 实现后的系统的性能测试以及评价网格性能的方法。

系统的关键点,即各种密码算法的实现和采用一种统一的可扩展的数据表示方法,实现系统内的数据传输。

3 结束语

业内人士普遍认为网格技术将是下一代互联网技术的核心,它利用高速互联网把分布于不同地理位置的计算机、数据库、存储器和软件等资源连成整体,就像一台超级计算机一样为用户提供一体化信息服务,充分实现了资源共享,具有成本低、效率高、使用更加方便等优点。与此同时,作为信息安全技术基础的密码学技术在信息化时代的作用日渐突出,如何有效地实现密码学相关的计算问题已成为密码学与网络安全发展的关键。

结合网格技术实现密码计算是用最前沿的理念实现最关键的需求,具有重大的研究价值和现实意义。系统结构的分析和实现将为实际的应用与开发指明方向,其体系结构所具有的通用性也将为其他基于网格技术的计算平台的搭建提供有价值的参考。

参考文献:

- [1] Bruce Schneier. 应用密码学[M]. 北京:机械工业出版社,2000. 158-349.
- [2] William Stallings. Cryptography and Network Security: Principles and Practice (2nd ed.) [M]. Printice-Hall, Inc., 1999. 58-169.
- [3] 陈国良. 并行计算[M]. 北京:高等教育出版社,1999.
- [4] 都志辉, 陈渝, 刘鹏. 网格计算[M]. 北京:清华大学出版社, 2002. 8-59.
- [5] Justin Couch. J2EE 宝典[M]. 北京:电子工业出版社,2002. 337-435.
- [6] Joe Zuffoletto. BEA WebLogic Server 宝典[M]. 北京:电子工业出版社,2003. 406-446.
- [7] 陈国良. 并行算法的设计与分析(修订版)[M]. 北京:高等教育出版社,2002. 126-169.
- [8] 黄铠, 徐志伟. 可扩展并行计算——技术、结构与编程[M]. 北京:机械工业出版社,2000. 357-398.

作者简介:

任俊伟(1979-),男,湖南长沙人,硕士研究生,主要研究方向为分布式密码算法及并行化技术、网格技术;林东岱(1964-),男,山东人,研究员,博士生导师,主要研究方向为密码算法试验平台开发、分布式密码算法及并行化技术;姜中华(1972-),男,湖北人,博士研究生,主要研究方向为分布式密码算法及并行化技术。