**云计算环境下的在线评判系统**

**【摘要】** 云计算是下一代互联网的发展趋势。在教育领域，云计算将是基础环境和平台。本文主要提出了在云计算环境下的在线评判系统。本文所提出的在线评判系统是基于Google App Engine( 简称GAE)， 可以在线上评判问题，分享问题，自动生成问题的主题和用户排行榜。它可以很容易地处理大规模数据和挖掘更多有趣的，对信息化教育非常有用的用户和问题。

**【关键词】** 在线评判，云计算， 谷歌应用程序引擎

# 1简介

在线评测系统是在程序设计竞赛中的在线评测程序的系统。该系统可以编译和执行代码，并且用预先准备的测试数据进行测试。用户提交的代码会在一定限制下执行，包括时间限制，内存限制，安全问题控制等。系统运行用户提交代码后的输出将与标准的测试输出进行对比。接着系统将返回评判结果。在大多数评测系统中，一些分析性的信息将会被展示，比如最大的通过记录数，用户的排行榜等。目前在互联网上存在很多的在线评测系统，比如UVa Online Judge，北京大学Online Judge，浙江大学Online Judge。在线评测一般来说是一个包括不同比赛的问题描述及评测参赛用户提交问题解决程序的数据集的服务器。在世界上任何地方的用户都可以免费注册在线评测系统的账号，解决他想要的问题。他可以提交很多他想要获取评判分析结果的解决方案。不仅是关于评判结果，而且是关于程序运行的时间，以便通过优化程序或使用的算法来解决选择的挑战问题。在线测评系统其中一个主要的特点是允许用户通过了解信息包括算法和程序来自我竞争。

但是在传统的在线评测系统上有很多缺陷，比如用户不可以测试算法在一个大规模数据的情况下，用户不能分享问题，用户不能从其学习过程收集和生成有用的信息。因此越来越多的在线评判系统变得与网络无关。在本文中，一种新的服务计算机模型—云计算平台提供了我们所需要的信息技术，并且云计算平台也提供了集成平台和技术环境的必要基础，适合各种零散的编程教学资源，在现有条件下尽可能满足教学和研究的资源。目前，国际知名IT公司，像谷歌，亚马逊，雅虎，微软，IBM等，非常重视云计算，并积极推动云计算的研究和应用。

本文安排如下：第二部分调查背景及相关研究，第三节所提出的系统的细节，第四节提供结论以及未来的研究工作。

# 背景和相关研究

## 2.1 云计算的概念

云计算是基于互联网的计算，从而共享资源，软件和信息提供给计算机和其它设备上的需求，比如电网。

在云计算中，用户需要的应用程序不是运行在用户的个人电脑，移动手机和其他终端设备上，而是在一个大规模的运行服务器集群。用户不需要处理数据的本地分配，仅仅需要存储在互联网数据中心里即可。管理和维护这些数据中心的正常运转这种寻常操作是由企业提供的。这些企业提供了云计算服务和确保有足够的计算能力及用户存储空间。在任何时候任何地方，任何用户可以连接到互联网终端。因此，无论是企业还是个人，他们可以按需实现云服务的使用。

## 2.2 云计算的特点

按需分配：所有资源将被分配，并只是按照用户的请求来分配。

超大规模：云计算的规模是非常巨大，他们有超过数十万台服务器。

高度抽象：用户只需专注于自己的个人东西。

高效率：收集所有资源，并通过软件分发管理计算IT资源管理的效率。

高可靠性：云使用的数据多转录容错，计算节点同构可互换等，以保证服务的高可靠性。这让使用云计算比本地计算机更加可靠。

## 2.3 云计算平台的系统架构

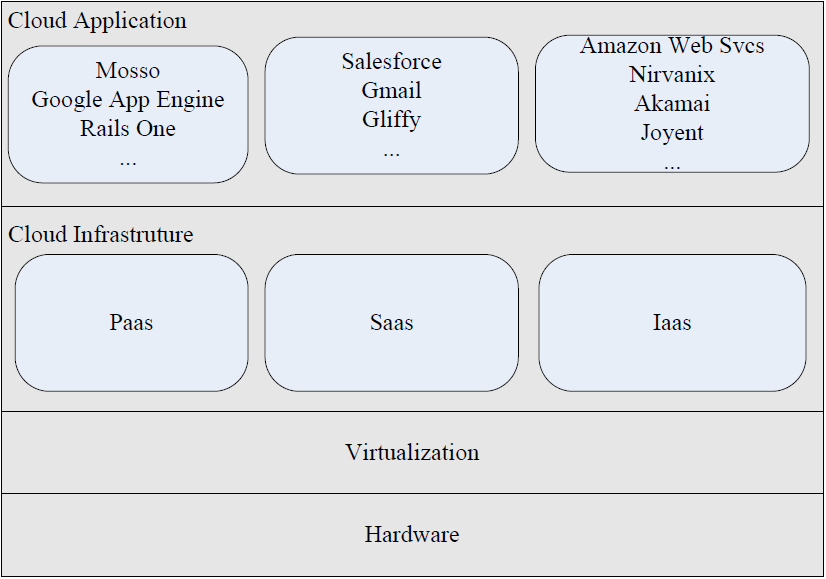


图1. 云计算平台的系统架构模型

云计算体系结构的最低层是被虚拟化，以提供一个灵活的自适应平台，以提高资源利用硬件层。

在虚拟化层，物理硬件被切成虚拟机，每个都有自己安装的小型操作系统。然后虚拟机基于他们所拥有的资源被集中在资源池中。

IaaS（一种基础设施服务，或IaaS的层）通过提供机制规定扩展了虚拟化层和效用计算的方式控制虚拟机。 IaaS的是基础设施租赁（计算资源和存储）服务。这不仅意味着虚拟化的电脑有保证的处理能力，而且保留带宽用于存储和互联网接入。从本质上说，它是租赁的计算机或者数据中心的服务质量的特定具有执行任意操作系统和软件。

Saas(一种软件服务，或Saas层) 是一个从集中系统分离出来的部署软件，可以运行在本地计算机上的（或者是远离云服务的）。作为一种计量服务，SaaS允许你租用一个应用程序，并且只需要支付在你使用时的费用。

# 云计算环境下的在线评判系统

## 3.1 使用样例

UC1: 在线评测系统

该系统可以编译和执行用户代码和测试代码预构建的数据。在一定条件限制下运行提交的代码，包括时间限制，内存限制，安全限制等。用户代码的输出将被系统捕获，并且与标准输出进行比较。

UC2: 问题分享

用户可以与其他人分享他们的问题。

UC3: 问题主题的自动生成

系统可以从很多问题中自动生成它们的主题，因此用户可以很容易地选择他们感兴趣的主题。

UC4: 排行榜

系统将给出根据用户做题记录和问题难易程度对比的排行榜。

* 1. 系统架构

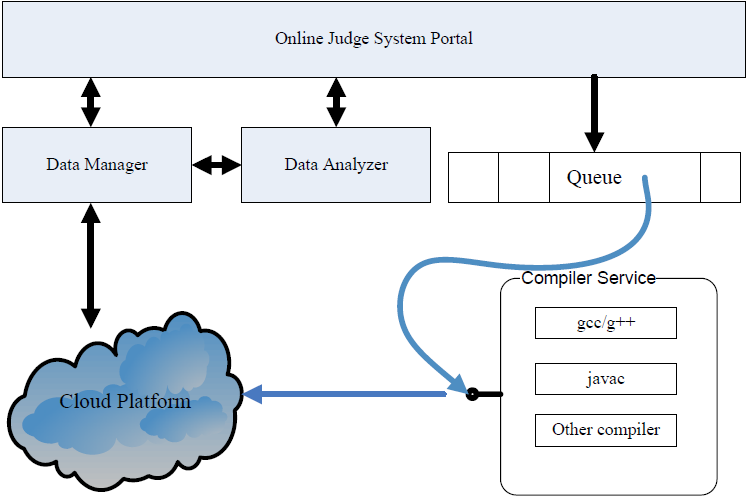


图2 系统架构

基于云计算的在线测评系统的系统架构图如上图2.

本文实现的系统包括在线评测系统的网站，数据管理服务，数据分析服务和编译系统服务。

## 3.2 问题主题的自动生成

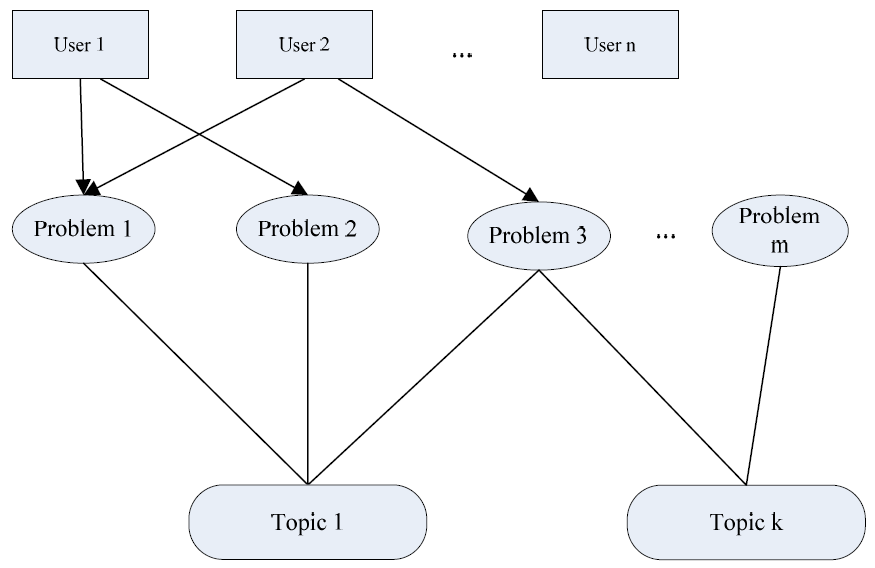


图3 概要模型

如何找到的文件的主题是一个在数据挖掘领域的普通任务。在在线评测系统中，每个用户都可以解决多个问题，并且每一个主题是可以由多个问题中生成（如图3）。因此，我们可以通过使用文本挖掘算法自动生成问题的主题。在本文中，我们采用文献[10]提出的算法，它被命名为潜在狄利克雷分配（LDA）。

潜在狄利克雷分配（LDA）是一种生成模型，它允许通过解释为什么数据的某些部分是相似的不可观测组进行说明集的归类。例如，如果观测是的文档中的单词的话，它假定，每个文档是一个小数目的主题的混合物，并且每个单词的出现是归属于该文档的主题之一。

## 3.3 系统的具体实现

我们使用Google App Engine（GAE）为原型实现我们提出的系统。 GAE，作为新一代的基于云计算的Web应用程序开发平台，可以让用户开发和运行在谷歌基础架构内Web应用程序。 GAE是容易建立和维护，并能根据用户的访问流量和数据存储需求进行轻松扩展。 GAE是用户不需要维护服务器：它可以立即为用户提供服务，只要开发者上传他们的应用程序，并且还可以将它们限制到授权用户。

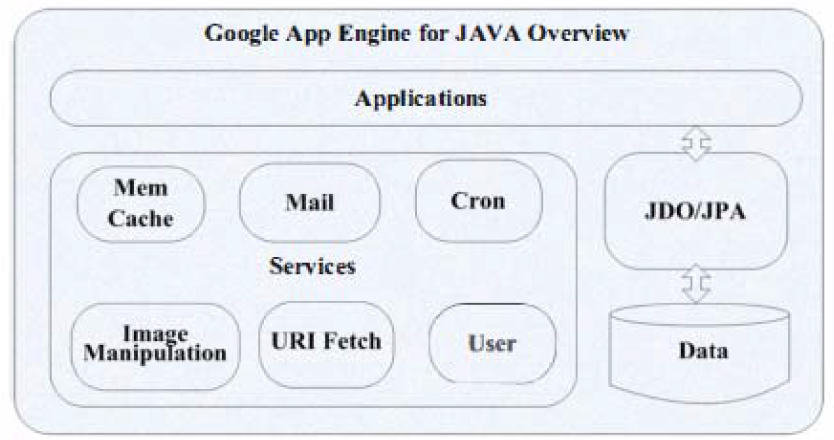


图4 在Java下的GAE预览图



图5 基于GAE开发的在线测评系统界面图

早在2008年四月GAE首次发布，它提供了一个运行环境，使用Python编程语言。 2009年4月，已经GAE宣布Java语言支持。图 3.是谷歌应用程序引擎的Java的概述。在Web应用程序的开发中，GAE支持JSP和Servlet技术。在数据存储上，GAE支持两套机制--- JDO和JPA，而开发者只需要考虑在面向对象的层面的数据持久，不需要关注关系数据库。此外，GAE还提供了一系列常用的服务，例如发送和接收电子邮件，图像处理和缓存。

图5是系统实现的原型截图

# 总结和展望

云计算是信息领域的一个重要趋势，也是研究教育信息化的热点话题。在本文中，我们提出了一个在云计算环境下的在线评测系统的设计与实现。

所提出的系统提供在线评测功能，问题分享功能，自动生成主题功能，以及排名列表的功能。该系统是基于谷歌的App Engine（GAE），它可以用来构建可伸缩系统的Web应用程序。 GAE两个优势是比其他的云计算解决方案的用户友好性和工具的支持。

在未来的研究工作中，我们将重点挖掘的数据将显示用户的学习行为。

致谢

这项研究是由宁波市教育服务重点专业建设基金（项目编号JSJZYQ-201035）的支持。

**【参考文献】**

[1] Zhang Guoli , and Liu Wanjun, “The applied research of cloud computing platform architecture in the E-Learning area,” 3rd International Conference on Computer and Automation Engineering, vol. 3, pp. 356- 359, Feb 2010.

[2] Hemant Kumar Mehta, Manohar Chandwani, and Priyesh Kanungo, “Towards development of a distributed e-Learning ecosystem,” The 2nd Annual Intl Conference on Technology for Education, pp. 68-71, July 2010.

[3] Bo Dong, Qinghua Zheng, Jie Yang, Haifei Li, and Mu Qiao, “An E-learning ecosystem based on cloud computing infrastructure,” 9th IEEE International Conference on Advanced Learning Technologies, pp. 125- 127, July 2009.

[4] Huibin Yin, Jun Han, Jing Liu ,and Xu Hongyun, “Development and research of multimedia courseware sharing platform based on GAE,” 1st International Conference on e-Health Networking, Digital Ecosystems and Technologies, pp. 409-412, April 2010.

[5] Miguel A. Revilla, “Competitive learning in informatics: the UVa online judge experience,” Olympiads in informatics, vol. 2, pp. 131-148, 2008.

[6] Andy Kurnia, Andrew Lim and Brenda Cheang, “Online judge,” Computers & Education, vol. 36, pp. 299-315, March 2001.

[7] SeungGwan Lee, Daeho Lee, and Sungwon, “Personalized DTV program recommendation system under a cloud computing environment,” IEEE Transaction on Consumer Electronics, vol. 56, pp. 1034-1042, May 2010.

[8] Chengxiang Zhai, Atulya Velivelli, and Bei Yu, “A Cross-Collection mixture model for comparative text mining,” The 10nd International Conference on Knowledge Discovery and Data Mining(KDD), pp. 271- 277, Aug 2004.

[9] Daniel Ramage,David Hall, Ramesh Nallapati, and Christopher D.Manning, “Labeled LDA: A supervised topic model for credit attribution in multi-labeled corpora,” Proceeding of the 2009 Conference on Empirical Methods in Natural Language Processing, pp. 248-256, Aug 2009.

[10]David M.Blei, Andrew Y.Ng, and Michael I.Jordan, “Latent Dirichlet Allocation,” Journal of Machine Learning Research, vol. 3,pp. 993-1022, 2003.