

基于 RIA 的在线实验室数据管理系统^①

李 凯 南 凯 郑依华 (中国科学院计算机网络信息中心 北京 100190)

摘 要: 介绍了西安加速器质谱中心实验室数据管理现状, 针对实验数据结构和关系复杂、数据表现要求高、异地用户多等特点, 结合 Flex 和 J2EE 等技术, 提出了基于 RIA 的在线实验室数据管理系统。基于 RIA 的在线实验室数据管理系统构建了符合实验流程的数据管理方法, 提供了更友好的交互界面和数据可视化能力。在实现该系统的过程中, 还总结和提出了一套基于 PureMVC 框架开发 Flex 应用的过程和方法。

关键词: 富互联网应用; 实验室数据管理系统; 科研信息化; PureMVC; Flex

RIA-Based Online Laboratory Data Management System

LI Ji, NAN Kai, ZHENG Yi-Hua

(Computer Network Information Center, Chinese Academy of Sciences, Beijing 100190, China)

Abstract: This paper introduces the topic of data management in Xi'an Accelerator Mass Spectrometry and proposes a solution for building RIA-based online laboratory data management system with Flex and J2EE technology. This system provides a more effective data management method, a more friendly interface, and a stronger data visualization. Moreover, the paper summarizes the processes and methods for developing Flex application with PureMVC framework.

Keywords: RIA; laboratory data management system; e-Science; PureMVC; Flex

随着 e-Science 的推广, 协同工作环境软件受到越来越多的重视和关注。而基于协同工作环境平台上, 可以搭建形式多样内容丰富的 e-Science 应用。实验室数据管理系统就是这样的应用之一。为了提高科研效率进一步探索 e-Science 的应用和推广模式, 中科院网络中心协同工作环境研究中心与西安加速器质谱中心合作开发了“西安加速器质谱中心智能化数据管理系统”。该系统围绕 Be、C、Al、I 四种元素的实验处理流程, 设计并实现了数据采集、数据分析、数据统计、数据展现等功能模块。这些模块将实验数据有效的管理起来, 紧密结合实验工作流程, 提高了科研人员的工作效率。

1 引言

“西安加速器质谱中心”(Xi'an Accelerator

Mass Spectrometry 简称 XAAM 中心)由中国科学院地球环境研究所与西安交通大学于 2004 年 7 月联合共建。中心所拥有的加速器质谱仪是精确探测微量的长寿命放射性同位素的前沿大型仪器设备, 具有快速、超高灵敏度分析和微量分析三大优点。

加速器质谱仪作为面向国内国际的公共科研设备, 由多家单位分时占用, 每次实验都会有大量数据生成和分析。在本文系统实现前, 该中心使用基于 C/S 的数据管理客户端。此客户端实现了数据管理基本功能, 但还存在如获取样品元数据和样品测量结果数据困难, 样品数据组织关系松散, 不利于数据比较、分析与计算等不足之处。另外作为 C/S 结构的程序都需要在用户本地进行单独安装并且需要运行在一定类库基础之上。为了避免客户重复安装, 更有效的管理和整合实验数据, 因此我们考虑使用 B/S 结构设计实验

^① 基金项目: 国家高技术研究发展计划(863)(2006AA01A120); 中科院信息化专项(INFO-115-D01)

收稿时间: 2010-02-08; 收到修改稿时间: 2010-04-01

室数据管理系统^[1]。

另一方面，科学数据在可视化方面往往有比较高的要求，经常需要对一些科学数据进行计算并生成统计图表。而且实验室数据的录入、分析和处理等过程都有严格的先后顺序以及较为频繁的人机交互^[2]。因此使用常规的动态网页技术无法满足这两个方面的需求，因此我们提出基于 RIA 的在线实验室数据管理系统。

2 相关技术与工具

Duckling 软件是一套由中国科学院计算机网络信息中心协同工作环境研究中心开发的基于互联网的支持 e-Science 的工具软件包，是一个专为科研团队提供的综合性资源共享和协作平台。该平台能有效集成学科应用插件，聚合网络环境中的硬件、软件、数据、信息等各类资源，为科研人员提供先进的信息化科研平台。西安加速器实验室数据管理系统就是建立在 Duckling 平台基础之上的面向特定领域的科研信息化应用^[3]。

所谓富互联网应用程序 (Rich Internet Application, 简称 RIA) 是一种具有近似于传统桌面应用软件系统功能和特性的网络应用系统。RIA 系统的特性是运行于浏览器中，不需要额外安装软件，具有丰富的表现力，将大部分处理任务都移植到客户端，仅保留一些必要数据通讯。

本系统采用 RIA 中的 Flex 技术框架实现。Flex 是 Adobe 公司发布的开源 RIA 开发框架。通过 Flex 技术和相关服务器端技术，开发者能够赋予 Flash 与后端服务器交互的能力。Flex 以 ActionScript、MXML 和扩展类库作为基础，编译生成 SWF 格式文件 (Flash 文件格式)。Flex 应用将 SWF 文件与 HTML 页面一起部署到 Web 服务器上。用户访问 Flex RIA 应用时需要预先安装 Flash Player 作为 ActionScript 的解析引擎。因此，Flex 框架生成的 SWF 格式文件，能运行于各种安装了 Flash Player 的操作系统上，具有良好的通用性和跨平台性。同时，Flash Player 的强大解析能力为 Flex 应用程序提供了更强大的交互和动画能力^[4]。

3 系统功能结构

本系统的采用 B/S 结构，浏览器端使用 Flex 以及

J2EE 混合结构，其中较为复杂的需要图表展示的数据计算和数据统计部分采用 Flex 实现，而数据录入部分则使用 J2EE 实现。整个系统的结构框图如图 1 所示。

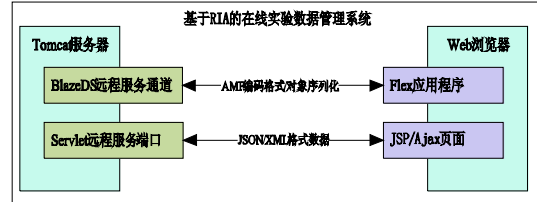


图 1 系统结构框图

其中 Flex 与服务器进行交互的部分采用的 AMF 编码格式对数据对象进行序列化，并使用 BlazeDS 开源包所提供的远程服务完成过程调用^[5]。

系统根据西安加速器质谱中心所提供需求，整理和设计出数据管理、日常管理、日常统计以及文档库四大部分。其系统架构图如图 2 所示。而这些四大模块都是建立在 Duckling 协同软件套装基础之上的，Duckling 软件为系统提供了完备的用户模块和文档管理模块。

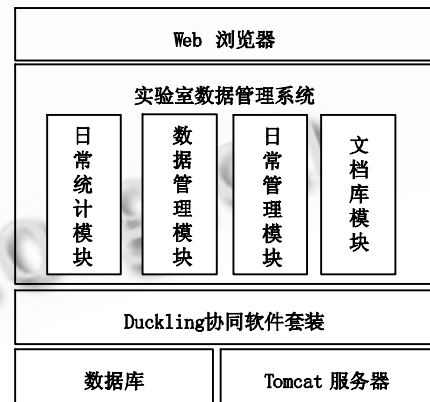


图 2 实验室管理系统架构图

用户模块可以实现用户的注册、登录、权限控制、建立虚拟组织等功能，由 Duckling 中的 UMT (User Management Tool) 完成。文档管理模块则负责系统内的文档库的建立和管理，主要功能有文档的上传、下载、更新、删除、授予权限等功能，由 Duckling 中的 DCT (Document Collaborative Tool) 完成。因此，在 Duckling 的支持下，我们将更多的时间投入到与客户业务密切相关的数据管理、日常管理和日常统计三大模块的开发过程中。

4 核心模块详细设计

系统的数据管理部分围绕加速器质谱仪工作流程,包括样品测试数据管理、日常管理、统计报表等三大部分。

4.1 数据管理模块

数据管理模块负责实验数据的采样、收集、分析、展示、保存和下载等功能,该部分与具体业务联系紧密也是整个系统中最核心的部分。本系统各过程中的具体功能说明,如图3所示。

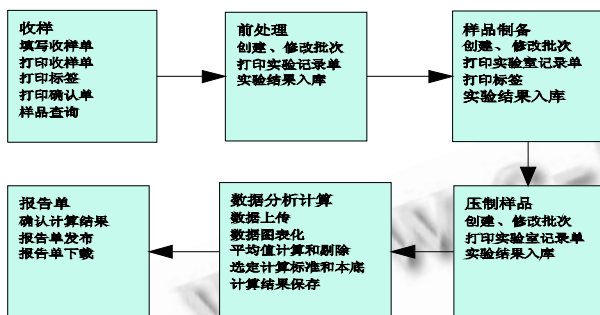


图3 数据管理部分工作流程图

在这些子模块当中,前四部分都是以表单操作为基本元素其主要的难点在于各个阶段数据状态的同步和验证。所有表单都能以PDF格式导出和打印,方便对数据进行归档。

由于数据分析计算部分经常需要对数据进行上卷、下钻操作等操作并且需要对不同粒度下的数据进行画图处理以便剔除一些无效数据。因此我们提出用Flex+J2EE开发数据分析与计算子模块。下面着重介绍数据分析计算子模块的实现。该模块流程如图4所示。

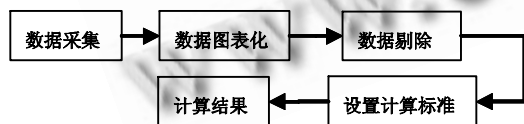


图4 数据管理子模块流程图

1)数据采集:该部分内容包括数据的上传、下载,对上传的实验数据进行编号,存入数据库。

2)数据图表化:主要完成抽取样品元数据信息,抽取数据测量结果,信息入库,并进行数据的归类与组织,实现对实验数据进行列表化显示,通过点击操作下钻显示粒度更细的数据。对不同粒度下的数据进

行绘图。

3)数据剔除:在数据图表化的基础之上,计算平均值,人工干预剔除无效数据,并保存入数据库。

4)设置标准:为最后计算步骤选定参照系和相关参数,如样本标准、本底等。

5)计算结果:根据不同元素进行计算,计算结果按照列表显示,用户可选择是否保存最终结果,并可以下载最终计算结果到本地。

4.2 日常管理模块

日常管理模块包括实验室运行状态记录、设备运行状态记录、大厅环境检测记录、设备真空度监测记录、用户测样合同管理、化学药品管理、大厅支持系统参数记录等功能。这些功能在形式上都采用表格结构,表格中分别记录日期、时间、描述等字段。表格中的每一条记录均可以进行添加、删除、编辑等操作,并且可以按照时间范围进行相关记录的查询。

4.3 日常统计模块

此部分主要是完成对实验室使用情况的一个统计功能,主要包括人员工作量统计、样品状态统计、样品来源统计以及标准本底测量数统计等。这些统计均能够以不同的分类标准对统计目标进行绘图。该部分的开发也是采用的Flex进行开发。

5 基于PureMVC的开发模型

在开发Flex程序时,难免会有功能添加和修改的地方。又因为Flex程序采用的事件监听响应的方式完成数据交互。而事件触发、事件监听、事件响应、数据更新等步骤往往涉及到多个类和方法,代码管理起来相当不便。因此我们选用PureMVC框架来规范开发步骤,降低系统耦合。

PureMVC框架是一款严格遵循模型-视图-控制器(Model-View-Controller,简称MVC)编程模型的开发框架,可用于多种开发语言。其体系框架图如图5所示。

图5中心的四个圆球分别代表这Model层、View层、Controller层和Facade层。Model层包含两个部分Proxies和Data Objects。Proxy即代理类,它负责对数据进行远程访问和更新。View层由若干个Mediators组成,可以理解为中介者类,它对应一个特定Flex Component并监听此Component中的所有事件。Mediator会发送一种名叫Note的消息,每个Note都有自己的名字,而且每个Note还可以携带

任意多的数据内容，称为消息体(Note Body)。Controller层则由若干个Commands组成，每个Command(命令)类都会监听一个或多个从View层发送过来的消息并启动对应Proxy中的方法。

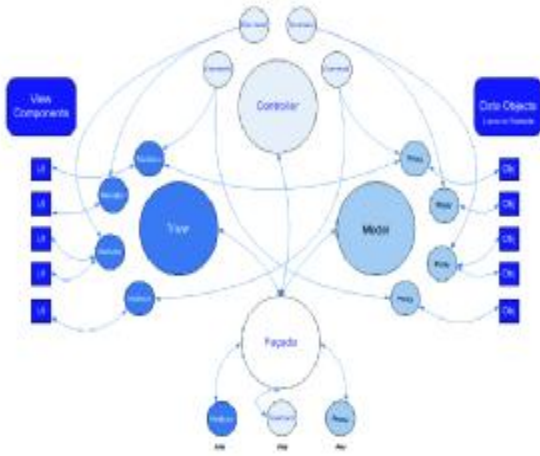


图5 PureMVC体系框架图[6]

这里比较特殊的是Facade层，它包含一个Facade类。这个类相当于整个系统的注册表，所有的消息名字都在此赋值。它清楚整个程序的细节，也是程序的启动的入口。

在PureMVC框架下Flex程序的运行流程如图6所示，该图描述的是一个具有服务器通讯的Flex程序的完整周期。



图6 PureMVC框架下的Flex运行流程图

添加任何一个新的功能模块需要做如图7所示的规范步骤。每一个新功能的添加都是将上述过程重新迭代一遍。通过此流程能够在最大程度上降低系统耦合，明晰系统开发步骤，提高开发效率。

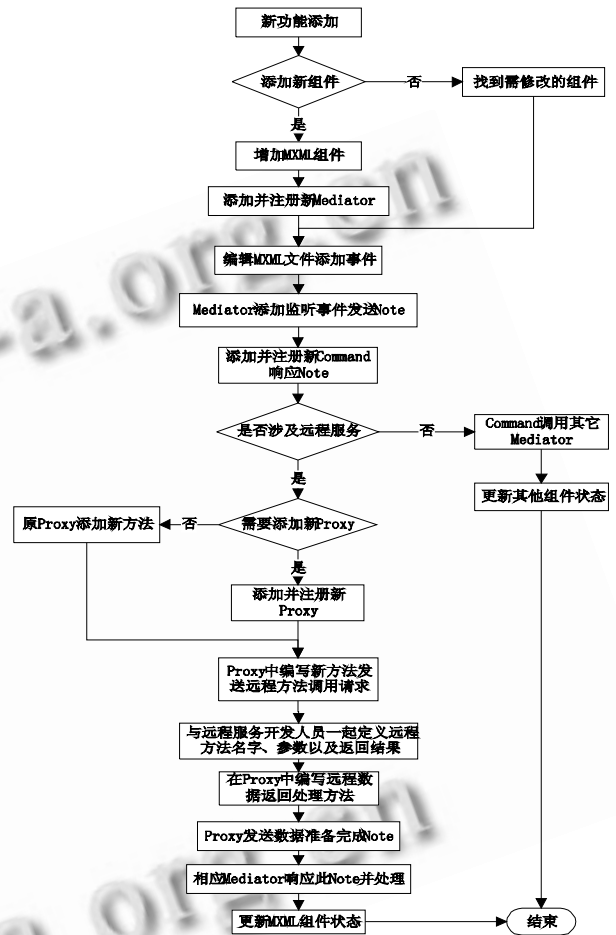


图7 PureMVC开发过程示意图

6 相关工作的比较

图8和图9所示的为我们系统中数据计算部分的运行效果图。



图8 数据计算中能谱图运行效果图



图9 系统框架及样品制备运行示意图

我们的系统与之前的 C/S 客户端比较而言，具有如下几个方面的优势：

1) 统一的用户和实验数据管理，用户无须手动维护本地端的实验数据；

2) 更为丰富的数据可视化能力和交互能力，该系统可以绘制具有 2048 个点的能谱图，并且能够对其进行缩放等操作。数据剔除方面采取绘制误差标准线的辅助科研人员进行剔除，无须重复计算。

3) 进一步明确实验室数据之间的逻辑关系，提供良好的数据下钻和上卷的能力，总结提炼更为清晰的数据分析与计算流程。

4) 无须用户安装客户端软件。即时查询历史信息，提供实验室日常管理功能和日常数据统计功能，方便实验室管理。

7 总结

本文为西安加速器质谱中心提供了一套基于RIA

的在线实验室数据管理系统。该系统采用 B/S 架构，结合 Flex 与 J2EE 技术，将实验室数据和业务流程有效的组织和管理起来，提供更为丰富的数据可视化能力和更为一致的数据状态，很好的提高了实验室科研效率。本文在实现系统的过程中采用 PureMVC 框架开发 Flex 应用程序，并总结一套通用的开发过程，对开发同类软件有很好的参考价值。

参考文献

- 1 Thomas N, Katy B. Representing, Analyzing and Visualizing Scholarly Data in Support of Research Management. Proceedings of the 11th International Conference Information Visualization. 2007. Pages: 124 – 129.
- 2 Kaletas EC, Hertzbege LO, Afsarmanesh H. A collaborative experimentation environment for biosciences. International Journal of Networking and Virtual Organizations archive, 2004,2(3):209 – 231.
- 3 南凯,董科军,马永征,杨德婷,李华飏.支持 e-Science 的协同工作环境.科研信息化技术与应用, 2008,1(1): 35 – 40.
- 4 陈一冰. Flex 第一步.北京:电子工业出版社, 2006.
- 5 Chafic Kazoun, Joey Lott. Programming Flex 3:The Comprehensive Guide to Creating Rich Internet Applications with Adobe Flex, O'Reilly. 2008.
- 6 Hall C. PureMVC Implementation Idioms and Best Practices[2009-05-30].http://puremvc.org/component/option,com_wrapper/Itemid,174/