**<A政府大数据应用系统>**

**软件体系结构文档**

**版本 <1.0>**

**修订历史记录**

|  |  |  |  |
| --- | --- | --- | --- |
| **日期** | **版本** | **描述** | **作者** |
| <30/11/2018> | <1.0> | 本文档目的在于提供关于A政府大数据应用系统的架构设计和物理部署。 | <汤晓岚，解坤> |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

目录

[I. 软件体系结构文档 4](#_Toc531860962)

[A. 简介 4](#_Toc531860963)

[1. 目的 4](#_Toc531860964)

[2. 范围 4](#_Toc531860965)

[3. 定义、首字母缩写和缩写 4](#_Toc531860966)

[4. 参考资料 4](#_Toc531860967)

[5. 概述 4](#_Toc531860968)

[B. 体系结构表示 5](#_Toc531860969)

[1. 架构视图阅读指南 5](#_Toc531860970)

[2. 图表与模型阅读指南 5](#_Toc531860971)

[C. 体系结构目标和约束 5](#_Toc531860972)

[1. 关键功能 5](#_Toc531860973)

[2. 关键质量属性 6](#_Toc531860974)

[3. 业务需求和约束因素 7](#_Toc531860975)

[D. 用例视图 8](#_Toc531860976)

[E. 逻辑视图 9](#_Toc531860977)

[1. 概述 9](#_Toc531860978)

[2. 重要体系结构设计包 10](#_Toc531860979)

[3. 用例实现 10](#_Toc531860980)

[F. 流程视图 11](#_Toc531860981)

[G. 部署视图 12](#_Toc531860982)

[1. 控制流组织 12](#_Toc531860983)

[H. 实施视图 13](#_Toc531860984)

[1. 概述 13](#_Toc531860985)

[2. 层 13](#_Toc531860986)

[I. 数据视图 15](#_Toc531860987)

[1. 持久化机制的选择 15](#_Toc531860988)

[2. 持久化储存方案 15](#_Toc531860989)

[3. 数据同步与复制策略 15](#_Toc531860990)

[J. 大小和性能 15](#_Toc531860991)

[1. 质量属性 15](#_Toc531860992)

[2. 约束属性 16](#_Toc531860993)

[K. 质量 16](#_Toc531860994)

# 软件体系结构文档

## 简介

### 目的

本文档将从架构方面对系统进行综合概述，其中会使用多种不同的架构视图来描述系统的架构设计和物理部署，用于记录并表述对系统架构方面作出的重要决策。该架构文档可提供给最终用户、研发人员、质保人员等相关人员阅读。

### 范围

本架构文档的主要内容有：文档简介、架构描述方式、架构设计目标、架构设计原则、逻辑，开发，运行，数据，物理架构视图、关键质量属性的设计原理。

### 定义、首字母缩写和缩写

### 参考资料

[1]. GB 24463.1-2009-T 交互式电子技术手册 第1部分：互操作性体系结构

[2]. IEEE1471-2000

[3]. ISO\_IEC\_IEEE 42010-2011

[4]. http://p.primeton.com/articles/53c8c7bfe138233ff8000013 基于服务的多源异构数据整合平台解决方案

[5]. 朱好好. 异构数据融合中间件的研究与应用[D]. 华中科技大学, 2012.

[6]. 佟明川, 陈明锐, 许斌,等. 基于多源异构数据的市政管理知识库构建[J]. 海南大学学报(自然科学版)自然科学版, 2016, 34(3):228-236.

[7]. 李广建, 汪语宇, 张丽. 数字资源整合的实现机制及关键技术——对国外数字资源整合系统的实证研究[J]. 中国图书馆学报, 2007, 33(2):75-80.

[8]. 赵馨. 数字化校园建设中实现异构数据库信息共享[J]. 科技信息, 2006(S3):25-25.

### 概述

本文档一共分为十一个小节。第二节描述用于当前系统的是何种软件体系结构，以及它是如何表示的；第三节描述对于体系结构具有重大影响的软件需求和目标；第四节列出用例模型中满足的用例或场景；第五节描述设计模型在体系结构上很重要的部分，第六节描述将系统分解成轻量级进程（单个的控制线程）和重量级进程（轻量级进程的组合）；第七节描述用以部署和运行软件的一个或多个物理网络（硬件）配置；第八节描述实施模型的整体结构，实施模型中从软件到层和子系统的分解，以及任何重要体系结构组件；第九节描述系统持久数据存储透视图；第十节描述影响体系结构的软件主要规模特征以及目标性能约束；第十一节描述软件体系结构如何对系统的所有能力起作用。

## 体系结构表示

### 架构视图阅读指南

### 图表与模型阅读指南

后续图表均根据UML建模语言绘制。

## 体系结构目标和约束

### 关键功能

核心功能： （1）用户成功登录系统；

（2）综合下属机构的系统，包括资金（基金）综合管理系统、社会保险管理系统人力资源、社会保障决策分析系统、劳动保障网上办事系统等，使得数据同步；

（3）简化数据库管理系统的类型，方便对需要的信息进行增删改查；

（4）减少重复填写个人信息，建立统一平台，且开发移动端平台；

（5）成功从多源数据库进行数据的读写，从而进行数据交互；

（6）生成数据报表与数据分析结果；

必做功能： （1）对数据进行增删改查；

（2）存储数据；

（3）导入导出数据；

（4）共享数据；

（5）数据集成；

（6）用户操作结果反馈；

（7）数据异常报错机制；

（8）移动端平台；

高风险功能： 数据安全，数据修复

独特功能： 数据图形化智能化处理

### 关键质量属性

根据上述背景，目标系统应该具有的质量属性有：

a)Avaliability（可用性）

作为提供给政府使用，登记市民相关信息的大数据应用系统，该系统必须具有比较强的可用性。在工作日政府工作时间不能出现系统崩溃的情况。在子系统出错的情况下，整个系统也不能出现崩溃。系统应该在政府下班或周末的休息时间进行养护，检查潜在的隐患，以避免工作时间因程序错误引起系统崩溃。

改：政府工作时间不能出现崩溃，休息时间可以停机养护，可用性要求高

b)Interoperability（互操作性）

作为统一登记市民信息的大数据应用系统，该系统需要具有较强的互操作性。一方面，系统需要与子系统建立合适的连接，快速得到系统需要得到的数据；另一方面，系统应对各个子系统的共性信息进行集中存储，解决“数据冗余”问题；用户在系统中注册后，所有子系统应同步用户信息，解决“数据孤岛”问题。

改：子系统标准不一，数据存储分散冗余，数据冗余，数据孤岛问题严重，构建数据互通的统一数据平台，互操作性要求高。

c)Modifiability(可修改性)

作为一个大数据应用平台，设计系统时应兼顾未来应用需求，为未来10年的政府业务流程变化做好准备。在政府业务流程，部门组织变化时，系统应该可以进行修改，以响应变化后的政府业务需求。

改：未来政府组织结构，业务流程可能发生变化，有一定可修改性要求

d)Performance（性能）

系统应该具有良好的性能，以满足该市各个部门的业务需要。系统应具备应对来自各个子系统较高查询，修改任务的能力，对每条请求的响应时间不超过10秒钟。

改：系统涉及使用部门多，请求集中发生在工作时段，需要处理大量添加，查询，修改，删除操作，性能要求高。

e)Security(安全)

由于系统存储市民各方面敏感信息，因此系统应具有较强安全性。系统应保证信息不被非授权用户获取，篡改；系统还应保证在遭受攻击时依然具有较强的可用性。

改：系统存储的用户数据不能被外界用户轻易窃取，修改，在系统遭受攻击时依然需要较强可用性，安全性要求高。

f)Tersability（可测试性）

由于系统运行关键业务，因此系统应具有良好可测试性。系统在开发，评审，部署阶段都需要经过测试才能投入运营。系统应具有详细的系统日志，以记录系统运行时状态；系统应有完善的异常机制，以明确系统出现错误的原因。

改：系统涉及功能多，内容复杂，模块开发，集成，部署阶段都需要进行测试，有一定可测试性要求。

g)Usability（易用性）

系统应具有一定易用性，经过简单培训的业务员可以对其部门相关业务进行操作；供高层查阅的数据分析界面要具有较强易用性，并具有友善的操作逻辑，给高层留下良好印象。统一办理全部业务的市民平台应具有友好的用户界面，让市民可以在4次点击以内找到需要找到的部门业务内容，并提供移动端平台访问。

改：系统要具有供市民直接使用的app客户端和供高层查看审阅的数据端，易用性要求高。

### 业务需求和约束因素

约束需求包括业务环境因素，使用环境因素，构建环境因素，技术环境因素：

（1）业务环境需求中：

考虑到市民对一个完善便利的大数据应用系统的需求，此系统预计在9个月内完成，预算2个亿，集成需要等非功能需要。大数据应用系统为了市民和政府工作人员开发，涉及到如下业务：资金（基金）综合管理系统，社会保险管理系统，人力资源与社会保障决策分析系统，劳动保障网上办事系统，退休人员管理系统。

（2）使用环境因素中：

此系统将提供给A市市民及政府相关工作人员使用。

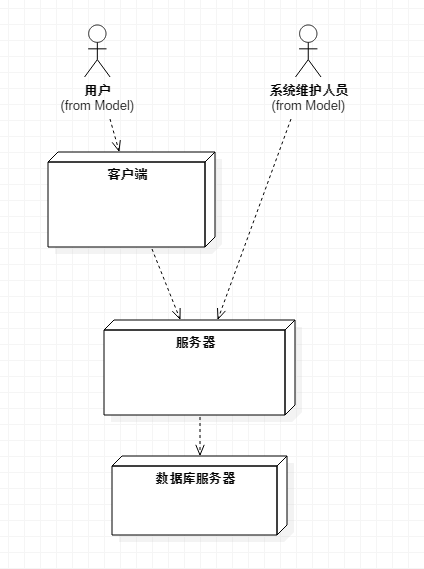
（3）构建环境因素中：

首先确保组建一个专业的开发团队进行合作，在开发管理方面，源代码保密需要顾及。

（4）技术环境因素中：

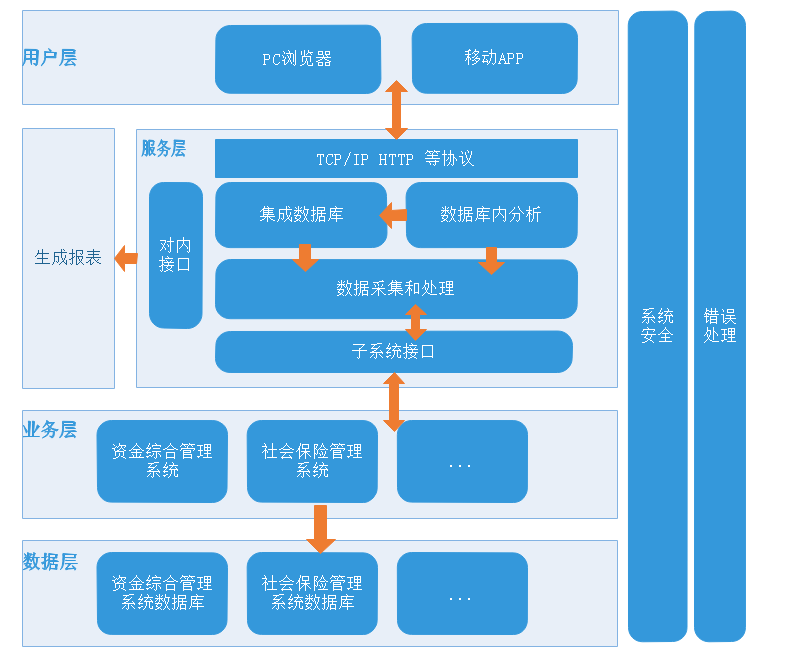
确保统一明确所有开发人员的技术平台，中间件以及编程语言。大数据应用具有非常重要的意义，能够解决数据孤岛的问题，对海量数据进行综合运用。

## 用例视图



## 逻辑视图

### 概述



*注：表示由访问方指向被访问方*

*表示可以互相调用访问*

系统分为用户层，服务层，业务层和数据层共计四层：

用户层

为大数据应用系统与用户交互的界面应用，即系统手机端APP，和web端界面。

服务层

为政府服务系统实时监控客户端,它负责接收用户下发的服务请求，通过子系统接口将服务请求下发给业务层，并将下层返回的结果显示在客户端界面上。也可以通过子系统接口对各子系统数据进行采集、处理，以便各子系统之间的数据交互和生成报表。

向上： 1.提供远程通道服务管理，包括通道连接用户、用户授权及访问IP过滤等安全功能。

 2.提供数据的标准管理及标准的加载。

 3.采用完全的B/S图形界面配置，简单易用。

向下： 1.支持前置库交换模式。服务层提供部门应用系统异构数据库与前置库的对接，包括数据抽取、数据转换、数据过滤、数据加载及数据实时同步交换等功能。

 2.提供读取数据库在线日志的非入侵式(CDC)数据增量、Trigger触发器增量、标识位增量等模式获取应用系统的实时增量数据; 支持全量(FullLoad)+增量(CDC)模式。

 3.支持对交换数据的过滤(Filter)、转换(Transformation)和函数处理。

 4.支持数据远程传输的可靠性、断点续传和"传且传一次"的特性。

 5.支持按应用系统处理事务的时间次序传输数据，保证业务数据的完整性和一致性。

 6.提供图形化的本地发送任务监控和本地日志管理。

业务层

为原各系统，负责合理安排调度通过通信层接收到的服务请求。由于本系统实现的是并发控制，因此同 一时间可能会接收到来自不同界面的任务请求，或者同一界面有不同的服务请求，这样任务管理层会通过一个任务队列以及多线程处理机制来管理这些服务请求，一方面保证接收到的服务请求不会丢失,另一方面保证将返回结果发送至正确的客户端界面,并且能做到效率高、响应快。

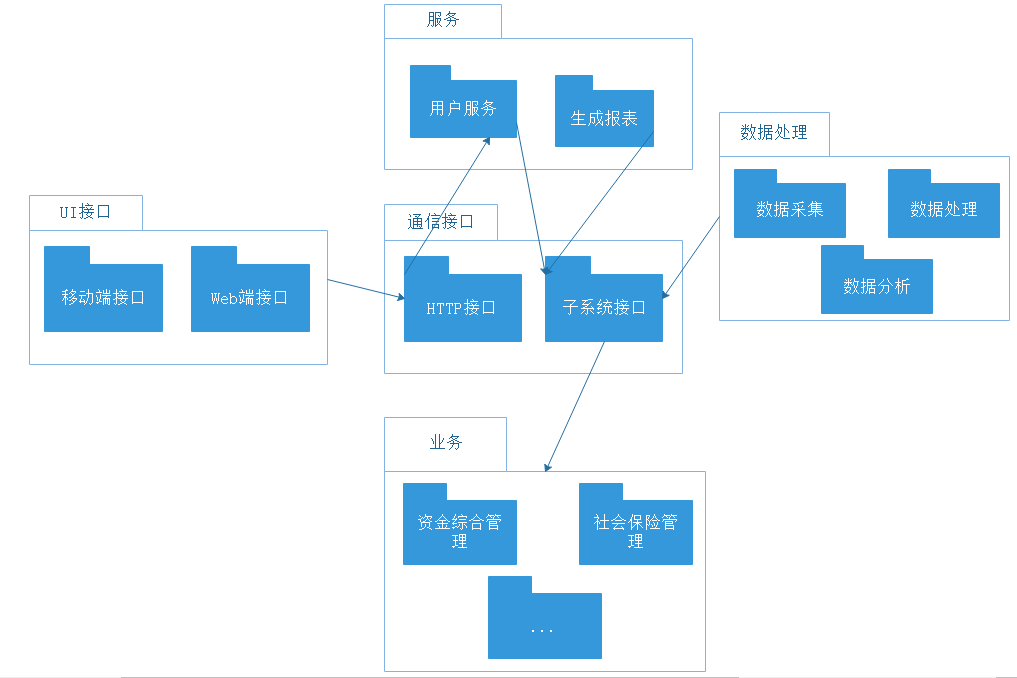
数据层

为各子系统的数据库。

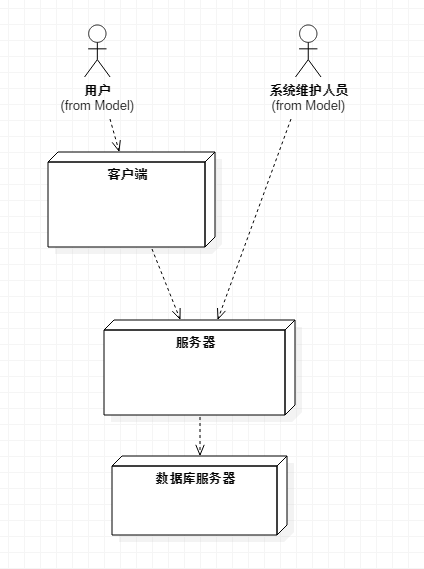
CAP权衡

上述架构在一致性和分区容错性上有较多的重视，在信息的调用过程中需要经过服务调度、数据检索、数据融合（转化）、数据同步等多个关键步骤，但也因此在性能的保障上有所欠缺。

### 重要体系结构设计包

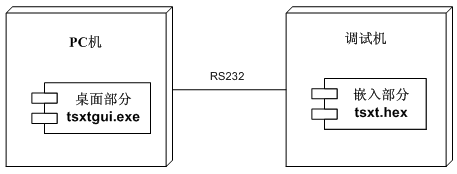


### 用例实现

政府服务系统实时监控客户端,它负责接收用户下发的服务请求，通过子系统接口将服务请求下发给业务层，并将下层返回的结果显示在客户端界面上。也可以通过子系统接口对各子系统数据进行采集、处理，以便各子系统之间的数据交互和生成报表。

## 流程视图

前端实现，通信协议和接口，服务请求任务调度，中间缓存数据库建立与读写接口实现，异常检查和处理，数据同步。



## 部署视图

### 控制流组织

(1)进程控制流

进程有三种状态：

*运行：* 进程在CPU上执行，或等待被执行（会被调度）。

*停止*： 进程被挂起（不会被调度）。收到SIGSTOP、SIGTSTP、SIDTTIN、SIGTTOU信号，进程停止；收到SIGCONT信号，进程再次开始运行。

*终止*： 进程永远停止。原因可能是：收到终止进程的信号，从主程序返回，调用exit 函数。

进程控制流的逻辑控制：

(2) 中断控制流

异常控制流:

将控制返回给当前指令（事件发生时正在执行的）。

将控制返回给下一条指令（没有异常将会执行的）。

终止被中断的程序。

## 实施视图

### 概述

用户使用移动设备或用户PC机通过HTTP连接对政府服务系统进行访问，获得服务。

管理员使用管理员PC机对系统进行维护。

政府服务系统服务器通过接口对各子系统进行访问，采集、处理数据。

### 层

## 数据视图

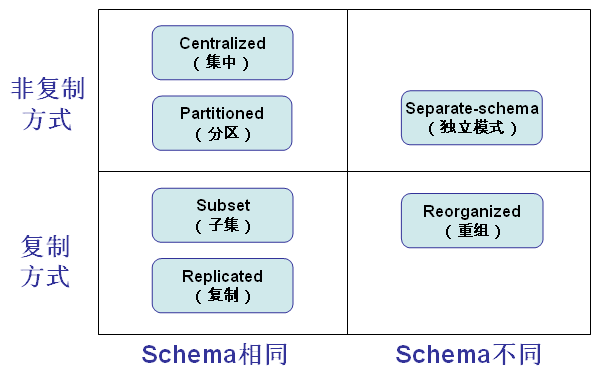
### 持久化机制的选择

关系数据库

### 持久化储存方案

表：数据名称 唯一ID 字长 安全等级 所属数据库 外键

### 数据同步与复制策略

强同步复制

## 大小和性能

### 质量属性

a)Avaliability（可用性）

b)Interoperability（互操作性）

c)Modifiability(可修改性)

d)Performance（性能）

e)Security(安全)

f)Tersability（可测试性）

g)Usability（易用性）

### 约束属性

约束需求包括业务环境因素，使用环境因素，构建环境因素，技术环境因素：

## 质量

|  |  |  |
| --- | --- | --- |
| 目标 | 场景 | 决策 |
| 互操作性 | 各子系统，使用的数据库类型不同，无法进行数据交换 | 数据处理接口 |
| 子系统数据和缓存数据库中的数据不一致 | 数据离线同步和增量同步 |
| 用户，重复输入多次个人信息 | 子系统数据交换 |
| 数据库异常但检查维护成本高 | 任务出错报警机制 |
| 安全性 | 黑客，试图访问、修改系统数据 | 数据备份、回滚，公私钥机制保护数据安全 |
| 用户，试图访问、修改系统数据 | 设置权限 |
| 用户，用户误删信息 | 数据回滚 |
| 数据库，数据库损坏 | 定义数据安全级别，数据备份，回滚 |
| 可用性 | 系统，出错，无法访问信息 | 维护、修复系统 |
| 系统，系统维护无法进行操作 | 拒绝访问系统，提醒用户系统维护 |
| 性能 | 客户端，重复请求页面，Web服务器请求数多负载压力大 | 代理服务器 |
| 客户端，重复请求页面，页面生成逻辑重复执行 | HTML静态化 |
| 客户请求，来自不同ISP，页面跨网络传递慢 | 内容分发网络 |
| 程序，大量申请数据，硬盘IO压力大  程序，申请不同数据，DBMS缓存低效 | 数据库拆分 |
| 易用性 | 用户，修改信息，获得服务 | 优化界面 |
| 界面操作复杂 | 提供易用的图形化管理控制台 |