数据平台

种子服务解决方案

版本号：V1.0

|  |  |  |  |
| --- | --- | --- | --- |
| 拟制: | 王尧 | 日期： | 2016/5/10 |
| 审核: |  | 日期： |  |
| 批准: |  | 日期： |  |

**文档修订记录**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **修订版本** | 日期 | **作者** | 修改内容描述 | **备注** |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

目录

1 引言 4

1.1 目的 4

1.2 预期读者 4

2 现有问题和解决思路 4

2.1 种子服务 4

2.1.1 概述 4

2.1.2 技术解决思路及措施 4

# 引言

## 目的

通过编写方案，解决现有种子服务的推送、查询问题，指导项目实施。

## 预期读者

(1).数据平台开发人员、爬虫开发人员、运维开发人员、产品经理、项目管理人员。

# 现有问题和解决思路

## 种子服务

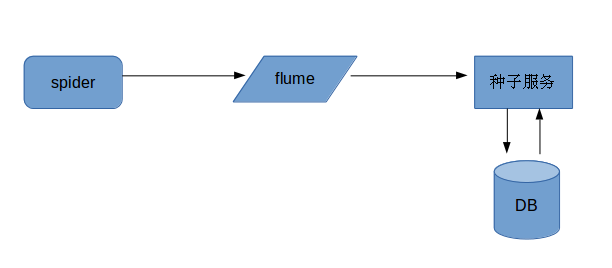
概述见《**数据平台种子需求**》

**2.2 技术解决思路及措施**

数据平台搭建种子服务，提供HTTP和thrift接口，提供搜索和查看功能。

数据平台提供爬虫端python版种子日志结构，爬虫需要实时将日志发往运维端flume接收。

运维flume接收日志后，统一发往种子服务接口，由种子服务解析入库，以便查询。



**2.3 种子数据结构**

json格式，字段如下

