**ESAgent系统设计文档**

V0.0.1

常官清 @ Enlink 数据中心

**目录**

[1. 概述 3](#_Toc529279821)

[1.1. 说明 3](#_Toc529279822)

[1.2. 系统特点 3](#_Toc529279823)

[1.3. 技术选型 3](#_Toc529279824)

[1.4. 其他因素 4](#_Toc529279825)

[2. 系统功能分析 5](#_Toc529279826)

[2.1. 配置管理 5](#_Toc529279827)

[2.2. 查询代理 5](#_Toc529279828)

[2.3. 内部测试 5](#_Toc529279829)

[3. 系统模块设计 6](#_Toc529279830)

[3.1. 访问控制 6](#_Toc529279831)

[4. 系统API设计 7](#_Toc529279832)

[4.1. 静态资源：/static 7](#_Toc529279833)

[4.2. 配置管理：/cfg 7](#_Toc529279834)

[4.3. 数据查询 8](#_Toc529279835)

[4.3.1. SQL查询：/sql 8](#_Toc529279836)

[4.3.2. DSL查询：/dsl 8](#_Toc529279837)

[5. 错误信息 9](#_Toc529279838)

# 概述

本文档描述ESAgent（ElasticSearch查询代理）模块的系统功能设计，旨在描述本模块的整体设计思路，帮助研发人员对该产品进行开发。

## 说明

ESAgent是数据中心产品中与数据查询相关的模块，其他所有模块对ES（Elasticsearch）的查询，必需通过本模块进行代理。本模块主要实现以下功能:

* **查询代理**  
  为其他模块查询ES提供代理，即提供查询接口；
* **SQL支持**  
  支持其他模块通过SQL语句查询，其中SQL为“内部”自定义格式；
* **访问控制**  
  对时间范围过大、维度过多、数据量较大等查询执行禁止操作，以保护ES集群的可用性；
* **研发测试**  
  为内部研发、学习等提供测试页面，使团队其他成员快速掌握SQL的使用。本功能只对内提供，不面向真实用户。

本文档参照《数据中心-功能需求文档-V1.0.0》编写，数据中心产品的需求及总体架构请参阅相关文档。

## 系统特点

根据ESAgent的功能需求，本模块有如下特点：

* **中量并发**  
  本模块主要的使用者为可视化、统计、分析、报警等模块，并发量属于中等，预计在“千”级别左右；
* **少量计算**  
  数据的统计、聚合、过滤等操作均由ES集群承担，故ESAgent不会涉及大的计算操作；
* **逻辑简单**

除SQL解析、访问控制以外，无其他复杂的应用场景。

所以本模块在设计时，遵循轻巧、实用、简单等原则，避免出现复杂的设计构架，而实现时亦避免出现复杂的实现逻辑。

## 技术选型

根据ESAgent的功能需求、系统特点等，对ESAgent的技术选型如表1-1所示：

表格 1-1 ESAgent技术选型

|  |  |  |
| --- | --- | --- |
| **功能** | **选型** | **说明** |
| 开发语言 | Golang |  |
| HTTP服务框架 | Go Echo | 高性能、极简 |
| SQL解析实现 | Go Yacc | 易于实现（学习成本较高） |
| IDE | GoLand |  |

其中，选择Go语言是出于“单机并发”的需要考虑，由于Go语言中“协程”的实现，可轻松获得高并发性能。Go Yacc 为yacc工具的Go版本，与yacc本身的兼容性较高，可快速实现语法的解析功能。

## 其他因素

根据数据中心产品和ESAgent模块的定位，以及ESAgent模块的系统特点，除并发性能以外，暂时不需要考虑扩展性、用户控制等其他因素。

# 系统功能分析

本章节描述ESAgent详细的系统功能。本章节参考功能需求文档编写，并对需求文档进行了详细的补充。

## 配置管理

ESAgent系统初始化时，需加载一些必要的配置信息（例如ES的集群地址等）。为方便管理及动态调整，该功能需满足以下需求：

* **本地存储**  
  由于配置内容不多，所以使用JSON格式将配置信息保存在本地。为了避免配置修改时重启程序，故ESAgent应对配置文件定时读取、更新；
* **配置页面**  
  为方便“非专业人员”操作，ESAgent应提供配置管理页面，以修改配置信息；
* **第三方接入**  
  为方便其他模块（例如集中式管理工具）对本模块的管理，本模块应向外部提供配置管理API接口。

ESAgent应提供必需的日志功能，记录所有配置修改的操作。

## 查询代理

查询代理是ESAgent模块的主要功能，对应的API接受SQL与DSL（ES原生查询语言）两种格式的查询请求。本模块通过将SQL转换成DSL的方式，来实现对SQL的支持。

出于特殊场景的考虑，ESAgent模块仅支持SQL的SELECT语句，同时对SELECT子句进行的扩展和重定义，请阅读《ESAgent SQL用户手册》学习更多详细内容。

SQL基本格式为：

|  |
| --- |
| SELECT … FROM … WHERE … GROUP BY … ORDER BY … LIMIT … |

查询代理功能应满足以下需求：

* **同时支持SQL与DSL**  
  考虑到SQL功能的局限性，提供DSL格式的请求以弥补SQL的不足；
* **预估查询请求复杂度**  
  包括时间复杂度、计算复杂度等，ESAgent应拦截复杂度较高的请求，以保障ES集群的正常运行；
* **访问频率限制**  
  统计查询请求的频率并进行限制，以防止ES集群“过载”。

## 内部测试

针对SQL功能，提供内部测试页面方便内部对ESAgent支持的SQL语法、功能进行学习和交流。

# 系统模块设计

本章节描述ESAgent系统中所涉及的各个模块，以及各模块之间的关系。模块的设计依据[章节2. 系统功能分析](#_系统功能分析)所描述的内容，如图3-1所示：

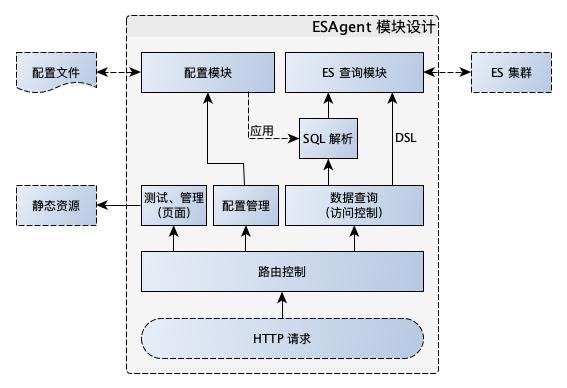


图 3‑1 ESAgent 系统模块设计

其中“HTTP请求”、“静态资源”、和“ES集群”不属于ESAgent模块设计范围之内，图中的箭头表示“请求”传递的方向。

## 访问控制

访问控制功能主要实现对访问频率的限制，以及对较高复杂度请求的拦截，暂时不需要考虑用户权限的控制。

# 系统API设计

本章描述ESAgent对外提供的API设计原则，详细的说明请参阅API说明文档。根据图3-1所示的系统模块，将URL按层级划分，方便根据URL进行路由。

## 静态资源：/static

静态资源API使用static进行标识，其中静态资源包括配置管理、内部测试页面使用的html、CSS、JS等资源。

由于静态资源（页面）分为管理与测试两个部分，所以使用URL中的二级目录进行标识或划分，如表4-1所示：

表格 4-1 静态页面API设计

|  |  |
| --- | --- |
| **URL** | **说明** |
| / | 主页，重定向到：/static/cfg/index.html（可配置）。 |
| /static/\* | 公共静态资源。 |
| /static/cfg/\* | 配置管理静态资源。 |
| /static/dev/\* | 开发测试静态资源。 |

其他未涉及（或新增）功能，请参阅以上描述进行实现。

## 配置管理：/cfg

配置管理API使用cfg进行标识，提供配置信息的查询和更改操作。 该功能API接口设计如表4-1所示：

表格 4-2 配置管理API设计

|  |  |
| --- | --- |
| **URL** | **说明** |
| /cfg/query | 配置信息查询接口。 |
| /cfg/update | 配置信息更新接口。 |

其中query接口的返回结果（配置项）应包含名称（或ID）、类型、默认值以及说明等信息，如表4-3所示：

表格 4-3 配置项内容

|  |  |  |
| --- | --- | --- |
| **字段** | **名称** | **说明** |
| name | 名称 | 配置项（或字段）的名称，唯一标识配置字段。 |
| type | 类型（int，string等） | 配置项的数据类型。 |
| range | 范围（或列表） | 配置项的取值范围或取值列表。 |
| default | 默认值 | 配置项的默认取值（当前取值）。 |
| desc | 描述信息 | 配置项的“用户可读”性描述。 |

用户可通过配置查询接口（/cfg/query）查询所有配置项，或者查询指定的某一个或多个配置项，查询接口实现时不应该限制其查询的内容。

用户可通过配置修改接口（/cfg/update）修改一个或多个配置项，该接口实现时不应该限制其修改内容和配置项个数。

## 数据查询

数据查询包包括两个功能：（1）SQL查询；（2）DSL查询。ESAgent使用sql和dsl分别标识对应的两类API接口。

### SQL查询：/sql

SQL查询功能除了支持数据查询功能以外，还支持其他与SQL相关的必要功能，其API接口设计如表4-4所示：

表格 4-4 SQL查询API设计

|  |  |
| --- | --- |
| **URL** | **说明** |
| /sql/query | SQL数据查询接口。 |
| /sql/translate | SQL翻译接口（将SQL翻译成DSL）。 |
| /sql/format | SQL格式化接口。 |

其中，数据查询接口（/sql/query）返回的结果如表4-5所示：

表格 4-5 SQL数据查询返回结果

|  |
| --- |
| [      { "name": "zhang", "age": 18, "sex": "male" },      { "name": "wangl", "age": 19, "sex": "male" },      { "name": "hanwu", "age": 22, "sex": "male" },      ...  ] |

其他接口的返回结果，请阅读API说明文档。

### DSL查询：/dsl

DSL查询功能只是对ES查询的代理，本文档不对其进行限制，实现时请参考ES的API实现，应尽量提供较全面的实现。

### Elasticsearch反向代理：/elasticsearch

Elasticsearch反向代理功能针对ES集群的所有Http协议方式的API，均可通过/ elasticsearch/{url}访问路径为url的ES接口。

# 错误信息

ESAgent模块处理HTTP请求时，除返回标准的HTTP状态码以后，需要定义一系列的错误码及错误信息。本章节仅提供参考作用，具体以实现为准。

参考示例如表5-1所示：

表格 5-1 错误信息

|  |  |
| --- | --- |
| **错误码** | **错误信息** |
| 1001 | 请求参数错误。 |
| 1002 | 数据查询超时。 |
| 1003 | 请求被终止。 |
| ... |  |
|  |  |