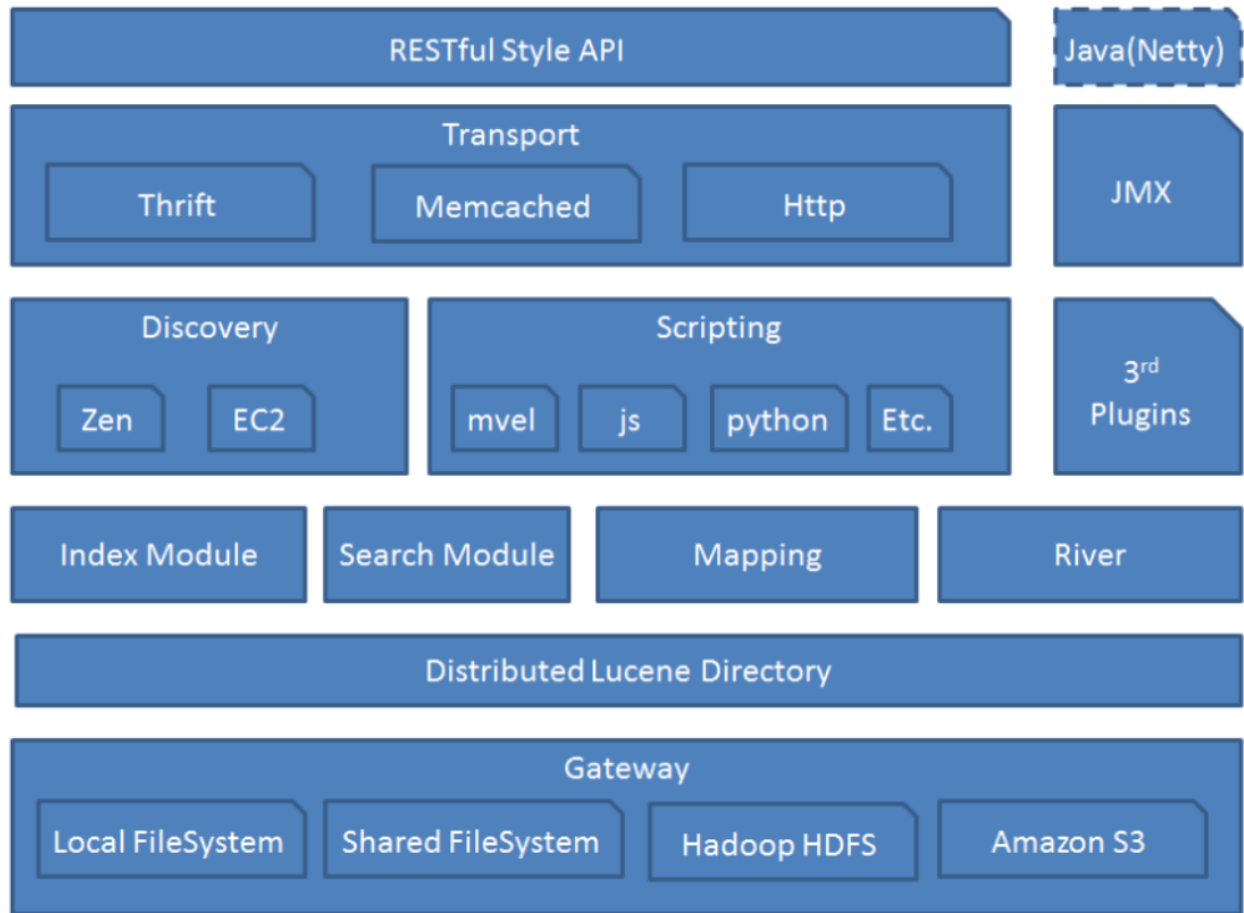


技术架构图



Gateway

gateway 层主要用于 ES 的持久化存储，支持 **本地文件系统、共享文件系统、HDFS、Amazon S3**，默认使用本地文件系统，没有网络延迟效率最高，但是扩展相对麻烦。

Distributed Lucence Directory

ES 底层基于 Lucence 进行搜索，**Distributed Lucence Directory** 由 Lucence 的索引文件组成，用于管理索引文件的添加、删除、合并等。

Index Module/Search Module

主要是 Lucence 的 索引和搜索模块。lucence 是一个[开放源代码](#)的全文检索引擎工具包，但它不是一个完整的全文检索引擎，而是一个全文检索引擎的架构，提供了完整的查询引擎和索引引擎，部分[文本分析引擎](#)

Mapping

ES 处理 index 结构的模块

River

ES 数据源模块，用于同步其他数据源到 ES，如Kafka、redis、mysql等，以插件的形式存在 ES 服务中。

Discovery

该模块主要负责集群中节点的自动发现和Master节点的选举。节点之间使用p2p的方式进行直接通信，不存在单点故障的问题。Elasticsearch中，Master节点维护集群的全局状态，比如节点加入和离开时进行shard的重新分配。默认使用的是 Zen 进行广播，其他方式以插件的形式存在。

Scripting

脚本模块，ES 查询API 中可以嵌入脚本对查询数据进行操作，支持 mvel、js、python等语言。

Transport

ES 通信模块，Thrift 是RPC框架，Memcached是分布式内存对象缓存系统，http 为网络请求。

JMX

JMX在Java编程语言中定义了应用程序以及网络管理和监控的[体系结构](#)、设计模式、[应用程序接口](#)以及服务。通常使用JMX来[监控系统的运行状态](#)或管理系统的某些方面，比如清空缓存、重新加载配置文件等。

Netty

Netty 是一个基于NIO的客户、服务器端编程框架