

数据交换平台简介

东师理想·大数据研发部

目 录

- 平台建设目标
- 平台核心流程
- 平台核心功能
- 平台部署架构
- 平台技术架构

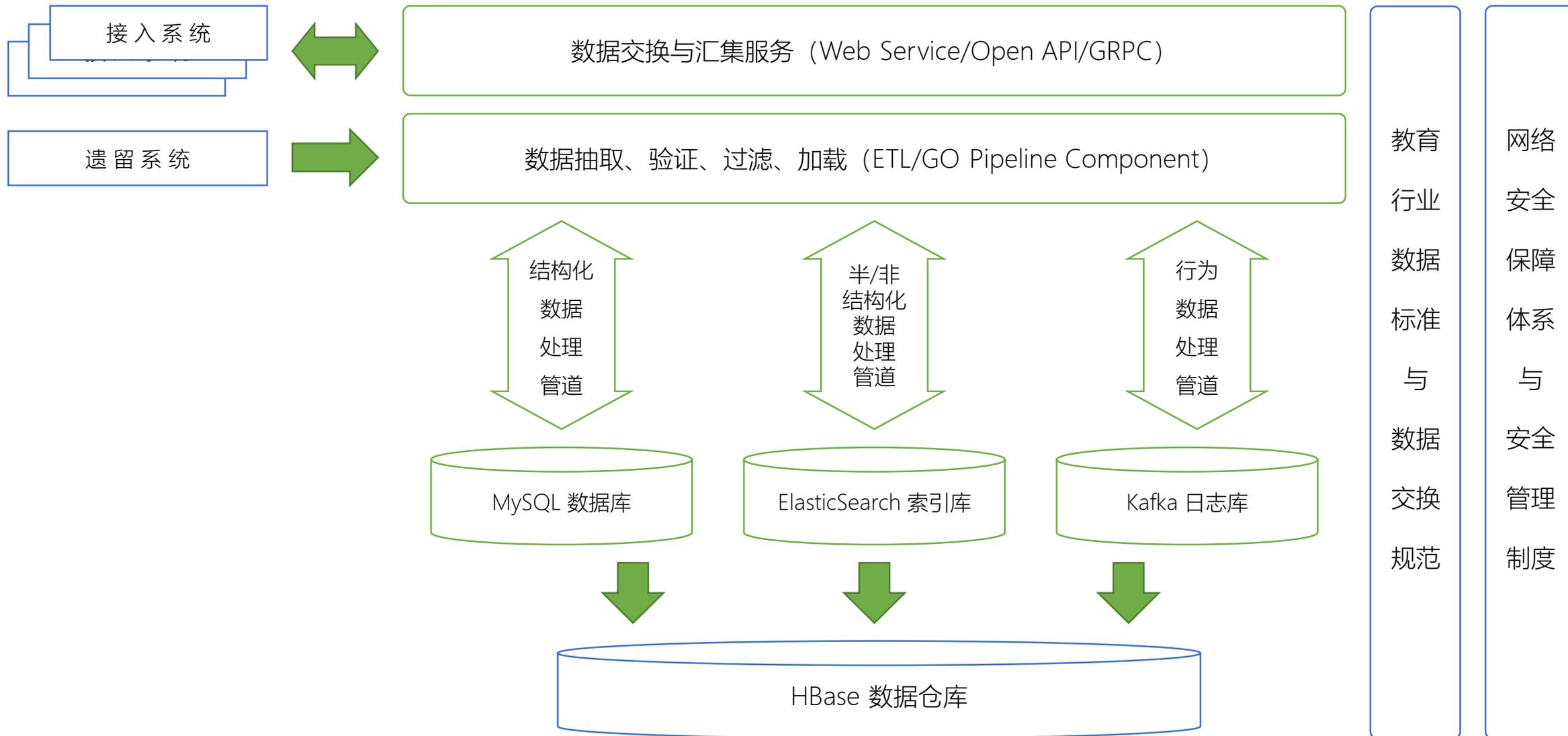
平台建设目标-1

- 东师理想大数据中心 数据交换平台（以下简称“数据交换平台”），是一套面向云计算、大数据、人工智能与物联网等典型应用场景，结合教育行业规范与数据标准，支持基础数据规范管理、业务数据互联互通、用户行为与日志数据实时汇集的后台支撑系统。
- 数据交换平台支持数据的纵向（多个中心、分布式部署）和横向（集中式部署）的交换与汇集；
- 支持平台基础数据、系统业务数据、系统日志数据和用户行为数据的交换与汇集；
- 支持结构化数据、半/非结构化数据和一般文件的交换与汇集，并支持使用ETL方式集成遗留系统数据（遗留系统需要开放数据库）。
- 数据交换平台能够与东师理想基础数据管理系统、大数据分析呈现系统实现无缝对接，高效的实现端到端大数据应用，为用户实施大数据项目节约成本、缩短周期、降低风险。

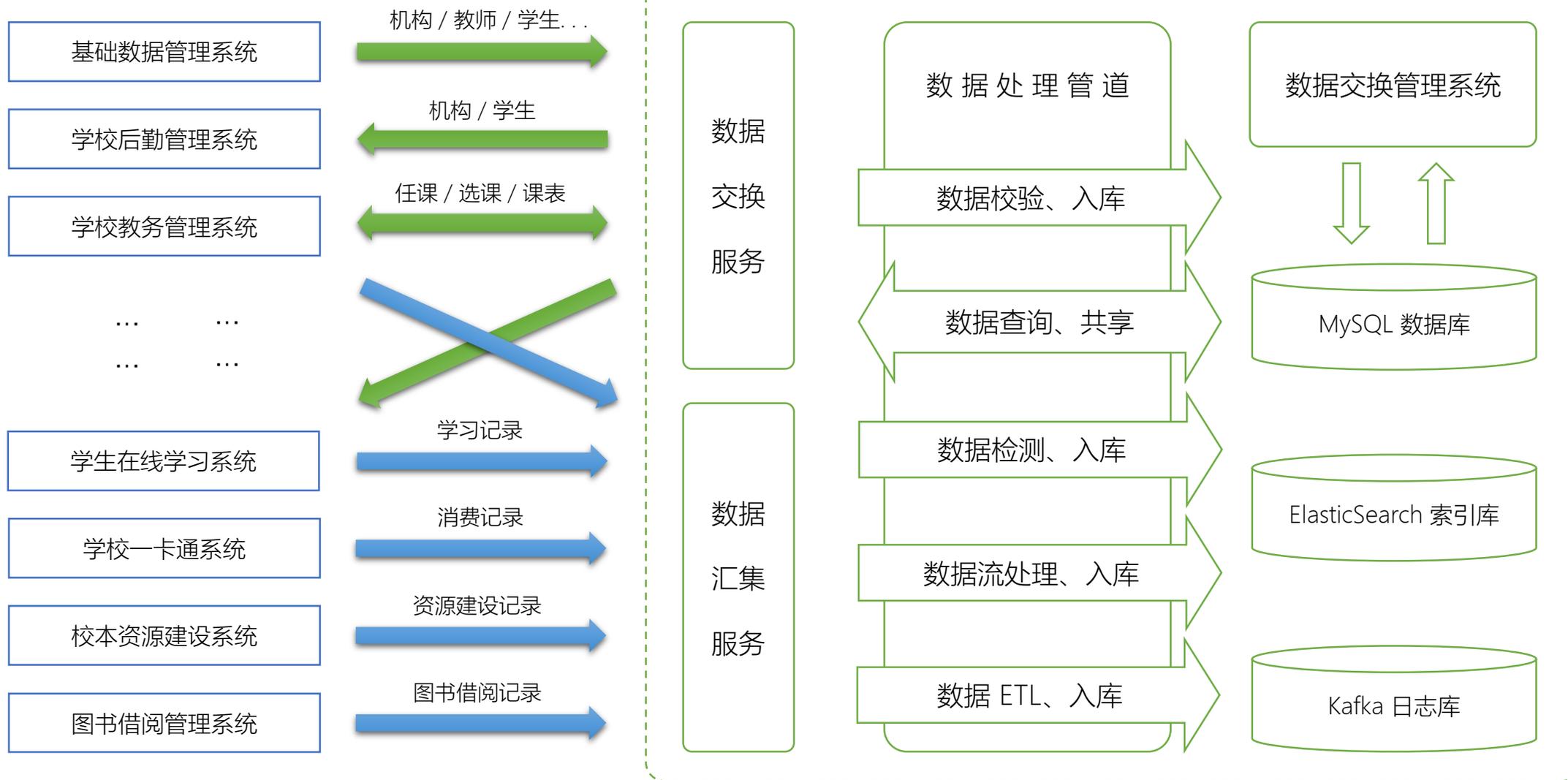
平台建设目标-2

- 应用数据交换平台，能够有效帮助教育局与学校用户解决以下问题：
 - 一、数据不规范、不一致：完全参照《JY/T 1001 - 1007 2012》系列教育行业数据标准，实现数据校验和数据过滤，保证交换与汇集数据的规范性、一致性和有效性，提高数据质量。
 - 二、系统信息孤岛难题：提供跨平台、多语言支持的通用数据交换接口，高效、快速整合各个接入系统，实现数据交换与数据汇集，打通信息孤岛，实现数据的互联互通、有效治理。
 - 三、“大数据”汇集难题：提供半结构化/非结构化数据、用户行为数据与系统日志数据的汇集接口，实现对大数据的高并发、高性能汇集、处理和存储，为大数据分析和AI智能决策提供数据支撑。
 - 四、遗留系统整合难题：提供成熟的ETL整合框架，支持全部主流数据库的点对点数据集成（遗留系统需要开放数据库），保证用户数据资产的充分利用、全面治理。

平台核心流程



交换流程示例



平台核心功能

数据交换、汇集与管理服务接口 (Web Service / Open API / GRPC)

数据交换管理

接入系统管理

数据源配置管理

元数据/规则管理

数据订阅/发布管理

数据访问权限管理

组织机构目录管理

交换/汇集数据查询

平台监控数据统计

数据交换服务

接入系统鉴权服务

数据共享出库服务

数据交换入库服务

数据校验过滤服务

数据映射转换服务

数据订阅发布服务

数据交换监控服务

数据源持久化服务

数据汇集服务

数据流入库服务

数据流处理服务

数据汇集监控服务

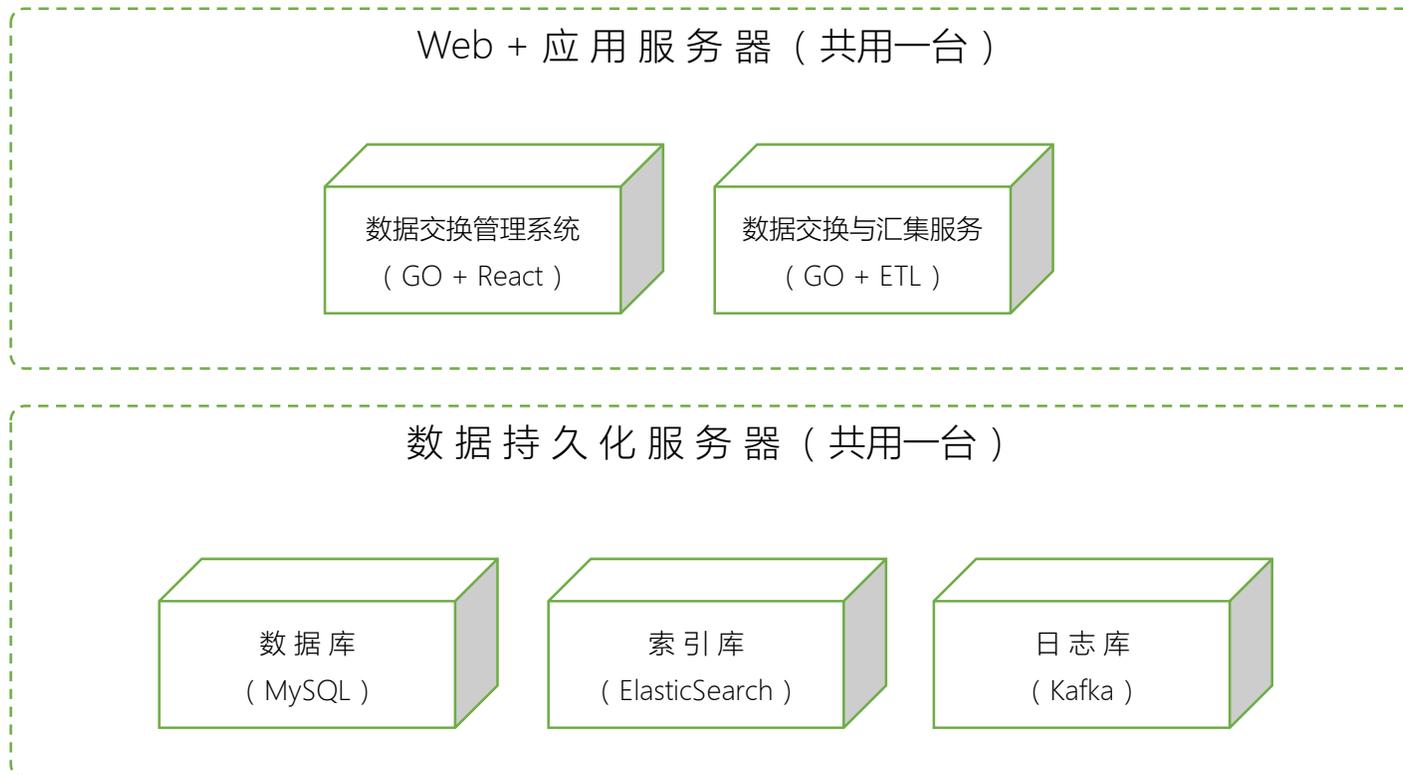
数据 ETL 服务

数据 ETL 工作任务

ETL 任务调度服务

ETL 进度监控服务

双机部署架构

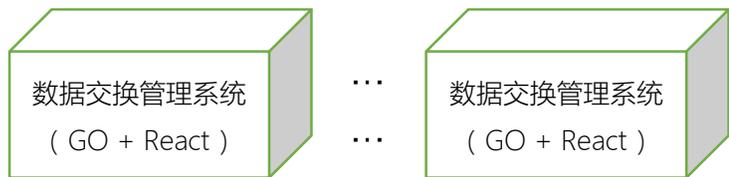


集群部署架构

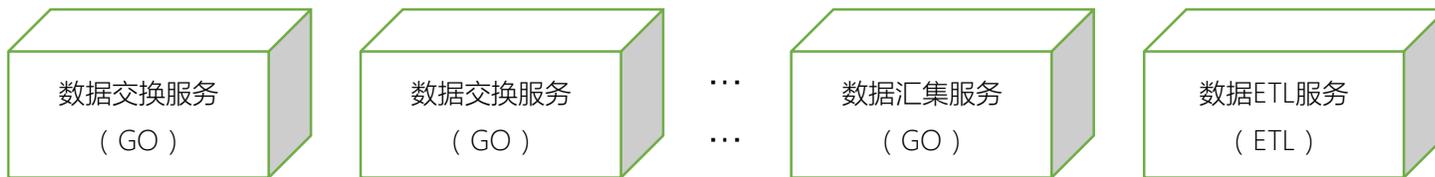
Web 负载均衡服务器
(Nginx)

应用负载均衡服务器
(Nginx)

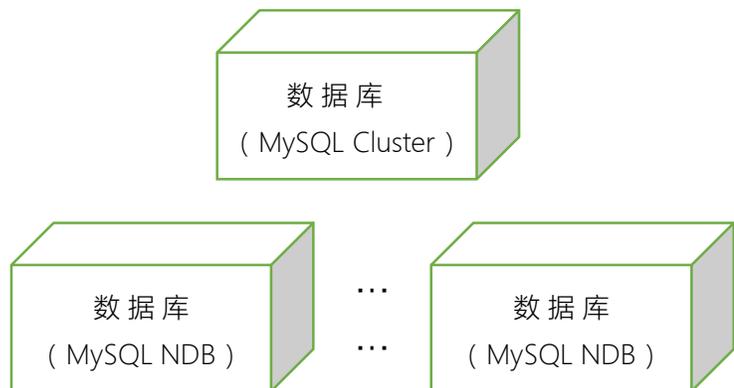
Web 服务器集群



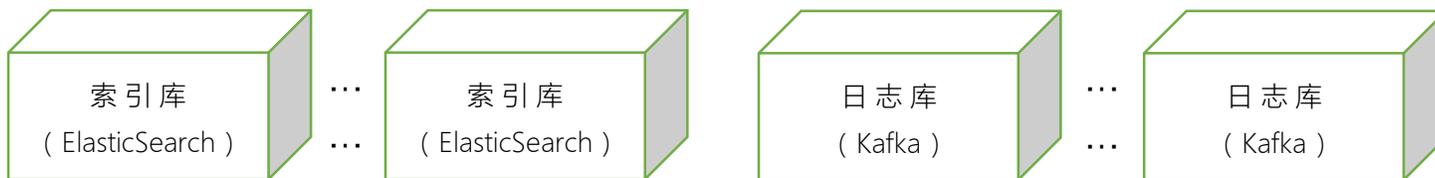
应用服务器集群



数据库服务器集群



数据持久化服务器集群



网络磁盘 / 光纤存储阵列

平台技术架构

