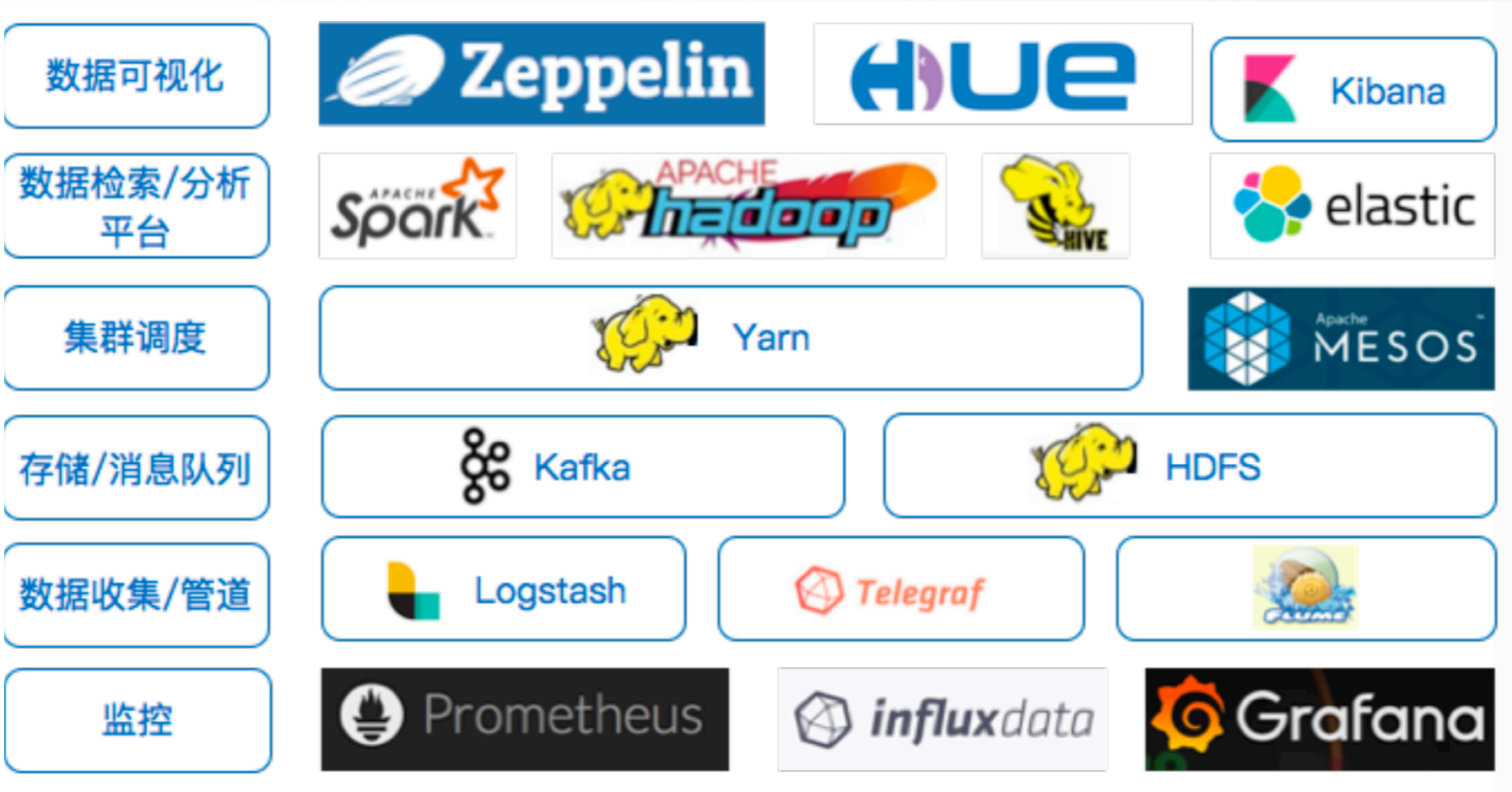
欢乐逛® 电商IT服务领导品牌

七牛云 QINIU | IT大咖说 知识分享平台

日志系统及计算任务平台

廖峻阳(欢乐逛*雨阳)

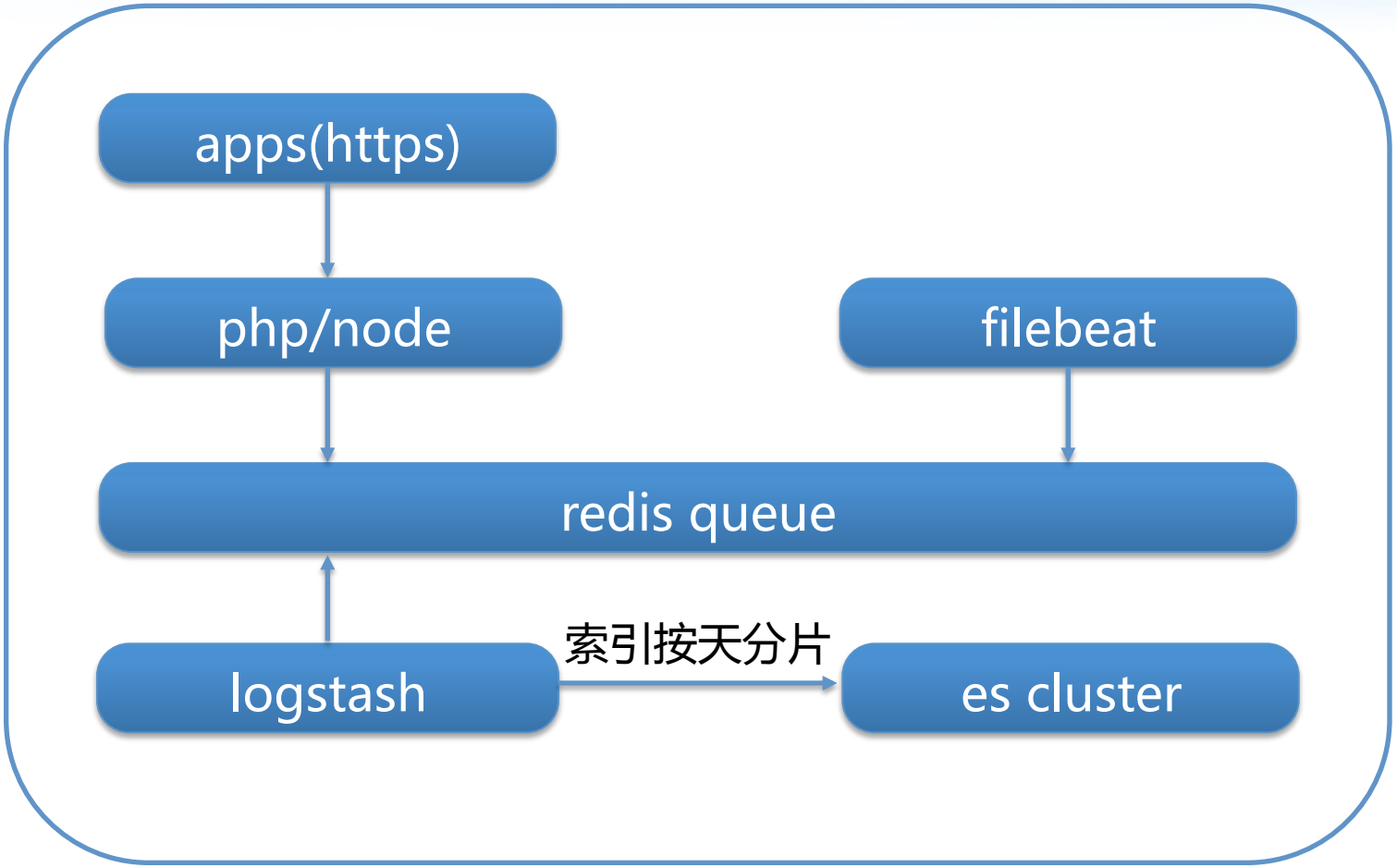
看到大数据想到什么


数据可视化
Zeppelin
HUE
Kibana
数据检索/分析平台
Spark
APACHE hadoop
HIVE
elastic
集群调度
Yarn
Apache MESOS
存储/消息队列
Kafka
HDFS
数据收集/管道
Logstash
Telegraf
监控
Prometheus
influxdata
Grafana

今天我们聊什么？

- logdb与logService
- tsdb选择与使用
- logdb <—> tsdb pipeline架构

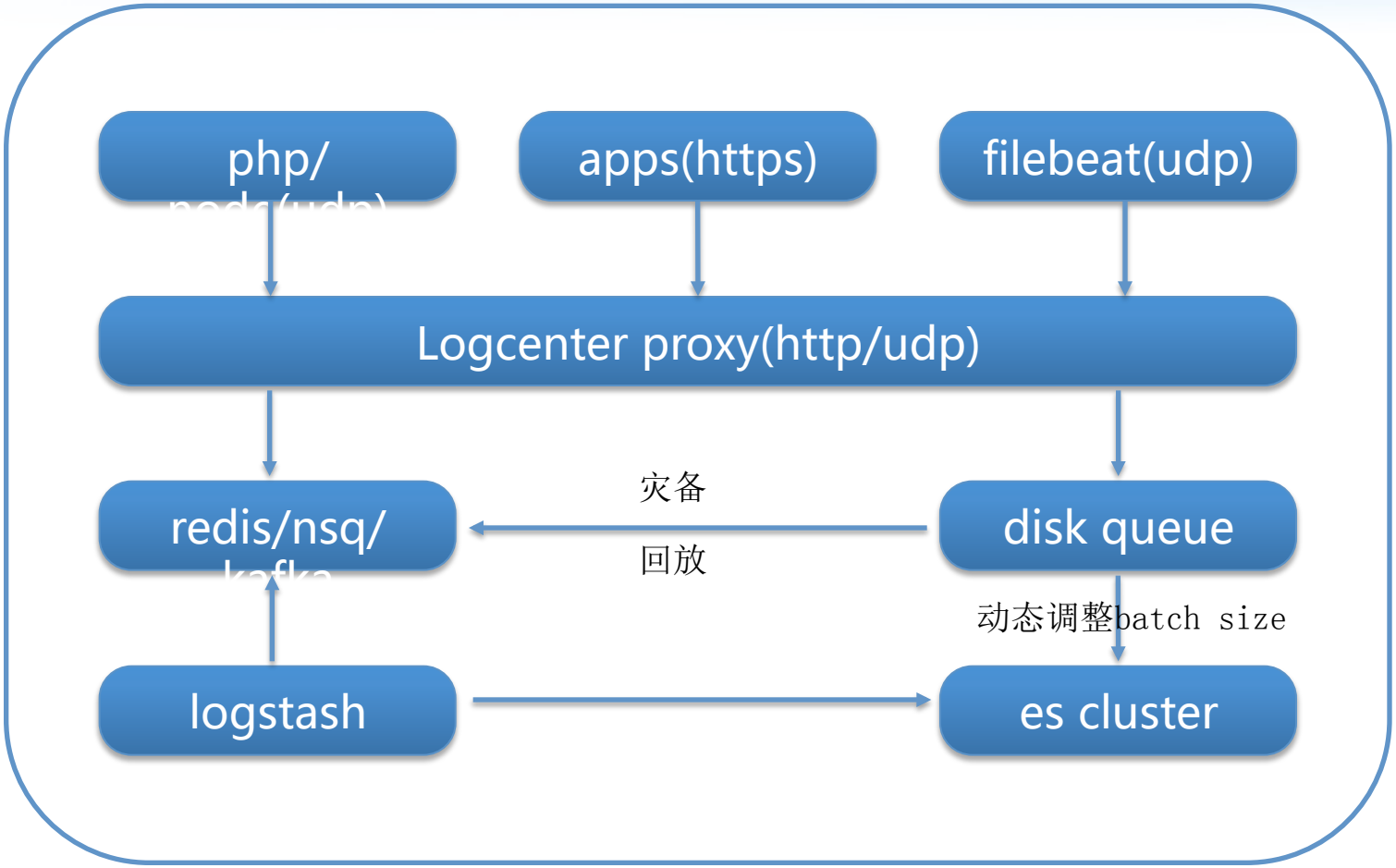
最基础的ELK架构



遇到了什么问题

- redis 直接暴露给应用(短链接应用无谓tcp消耗)
- redis单点,内存占用不可控
- logstash/es消费lag 引发连环灾难
- 接入点过于分散(客户端/前端日志接入不易)
- 索引统一按天分片导致索引shard数过多

架构调整



为什么增加proxy

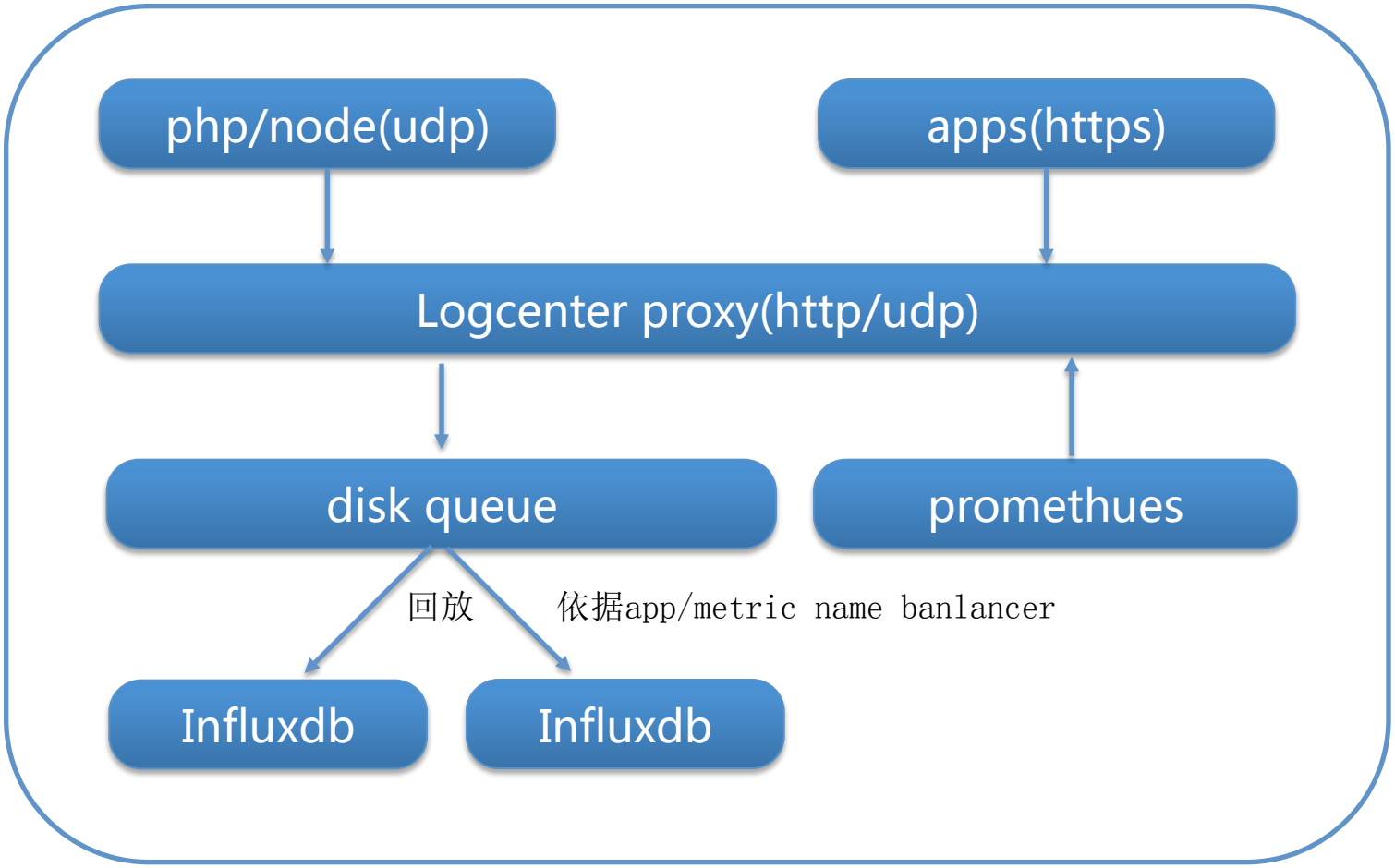
- 收拢接入点(规范化http/udp接入)
- 依赖disk queue作为队列灾备
- 数据规整,字段补齐,过滤清洗
- 预设值索引分片规则(容量预估)
- 可扩展性(多存储引擎)

tsdb选择：influxDb&Prometheus

- 推模型
- 类sql的查询语句(学习成本低)
- tsm/tsl 存储引擎
- 存储每一条doc(统计自由度高)
- 社区版本没有分布式支持
- 基础监控/较为复杂的时序统计

- 拉模型
- 自定义query dsl(较高学习成本)
- leveldb 存储
- 四种数据类型(扩展性低)
- 拉模型分布式需求相对较低
- 适合基础监控指标(docker兼容性)

如何接入


php/node(udp)
apps(https)
Logcenter proxy(http/udp)
disk queue
promethues
回放
依据app/metric name banlancer
Influxdb
Influxdb

这就结束了？

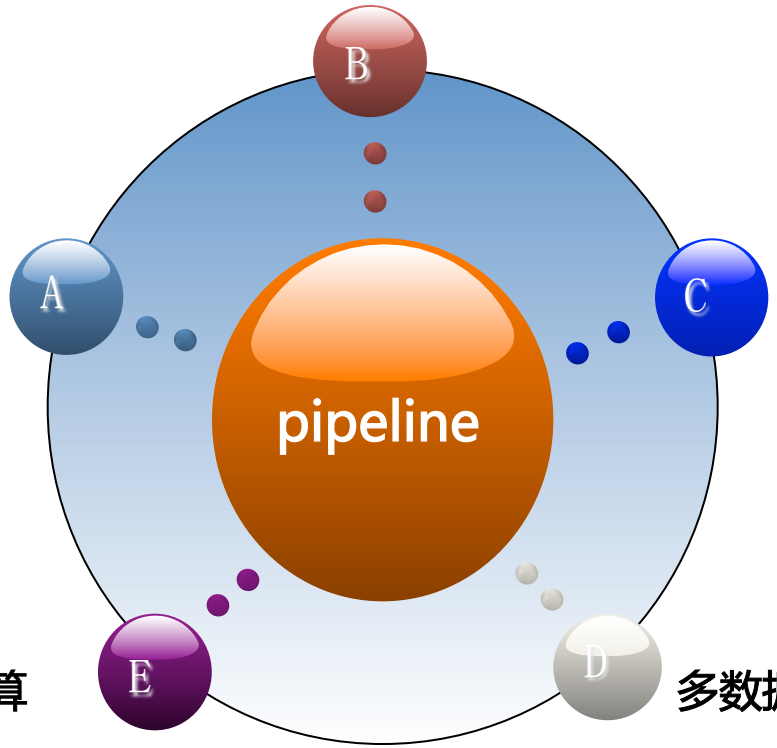
- grafana+influxdb只能做实时统计
- 无法一条sql完成统计需求(join/差集/自定义计算)
- 离线计算任务如何处理
- 有监控就有报警,自定义监控规则？
- 数据源导入导出(logdb->tsdb, mysql->tsdb, tsdb->mysql)

TICK架构

- influxdb长时计算cpu/memory消耗偏大
- Kapacitor 的tick脚本学习成本较高,编辑任务不方便
- Kapacitor只基于influxdb体系,输入源/输出源无法拓展
- 社区版本kapacitor没有分布式支持

自主实现pipeline

多数据源支持(es/influxdb/mysql)



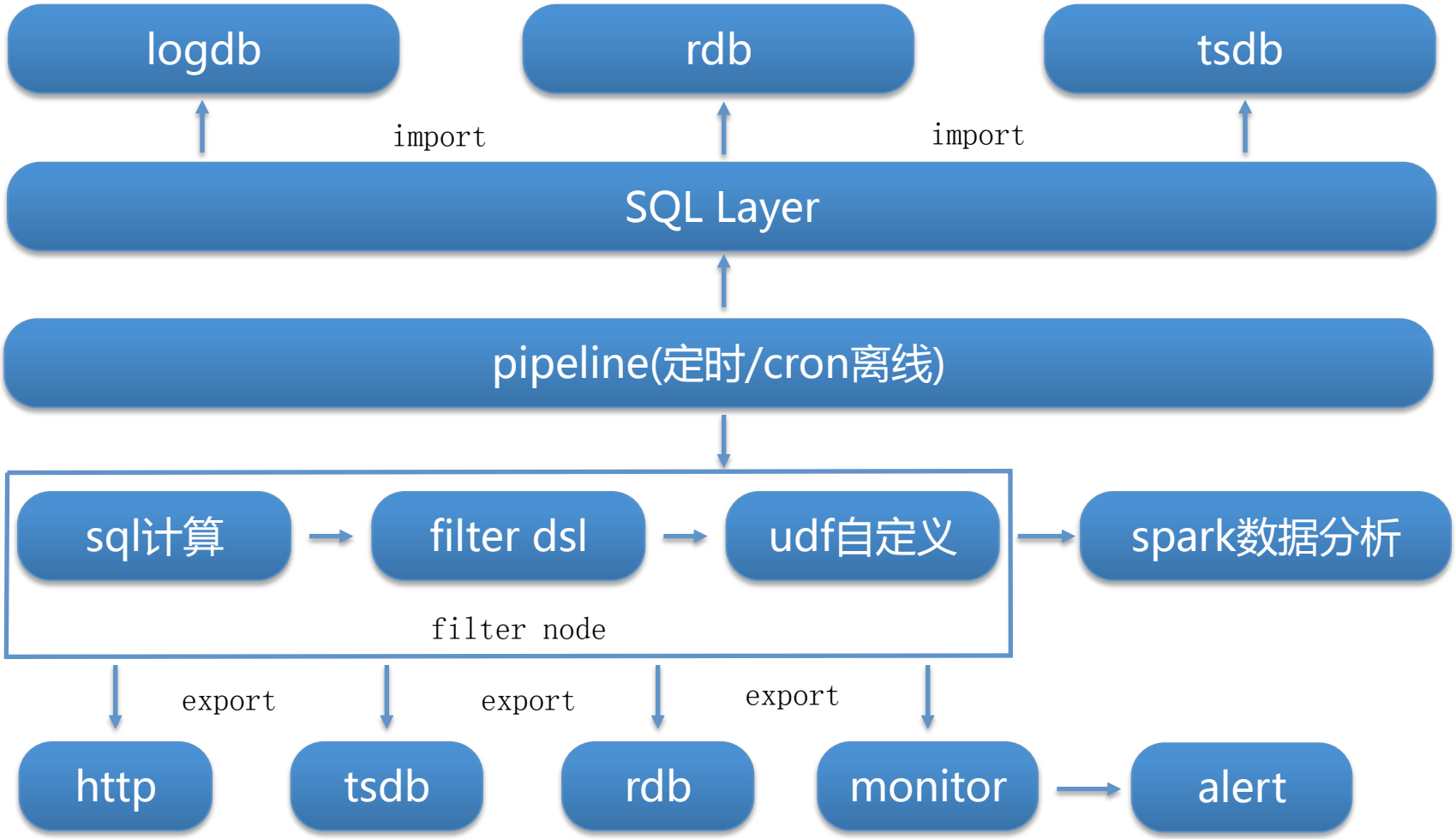
定时计算/cron离线任务

web管理,分布式支持

udf计算,多数据源聚合运算

多数据源统一监控报警支持

Pipeline架构体系



我们正在做的

- 优化pipeline分布式调度
- 基于spark实现多数据源聚合运算(交集/差集/filter)
- 基于mesos的整体资源调度
- 监控预警规则多样化,预警智能化

谢谢大家！