



十年架构 成长之路

SACC 第十届中国系统架构师大会

SYSTEM ARCHITECT CONFERENCE CHINA 2018

2018年10月17-10月21日 北京海淀永泰福朋喜来登酒店



小米大数据集成架构演化之路

夏军



SACC

第十届中国系统架构师大会
SYSTEM ARCHITECT CONFERENCE CHINA 2018

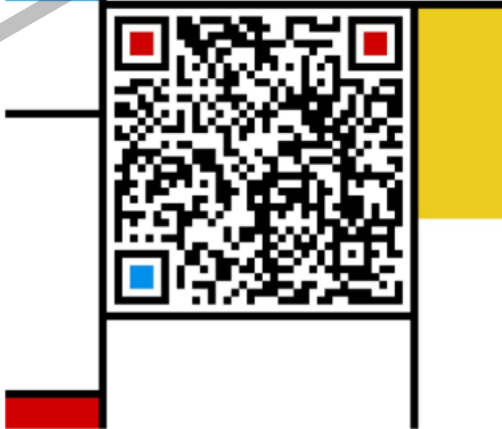


自我介绍

- 14年加入小米，来自于小米人工智能与云平台
- 关注于消息中间件，大数据集成，离线计算和流式计算等
- tedxia01@gmail.com



夏军 



扫一扫上面的二维码图案，加我微信



十年架构 成长之路



主要内容

- 问题与挑战
- 数据流整体框架
- 核心功能
- 应用场景解析



十年架构 成长之路



问题与挑战



十年架构 成长之路



问题与挑战

- 大数据系统众多，如何**高效集成**多个系统
- 如何保证数据**低延迟**传输
- 实时**发现/警报**数据流异常
- **量化**数据流整体链路情况
 - 数据延迟：3W1H (**Which** stream **When** and **Where** delay **How many** data)
 - 数据丢失：3W1H (**Which** stream **When** and **Where** drop **How many** data)



十年架构 成长之路



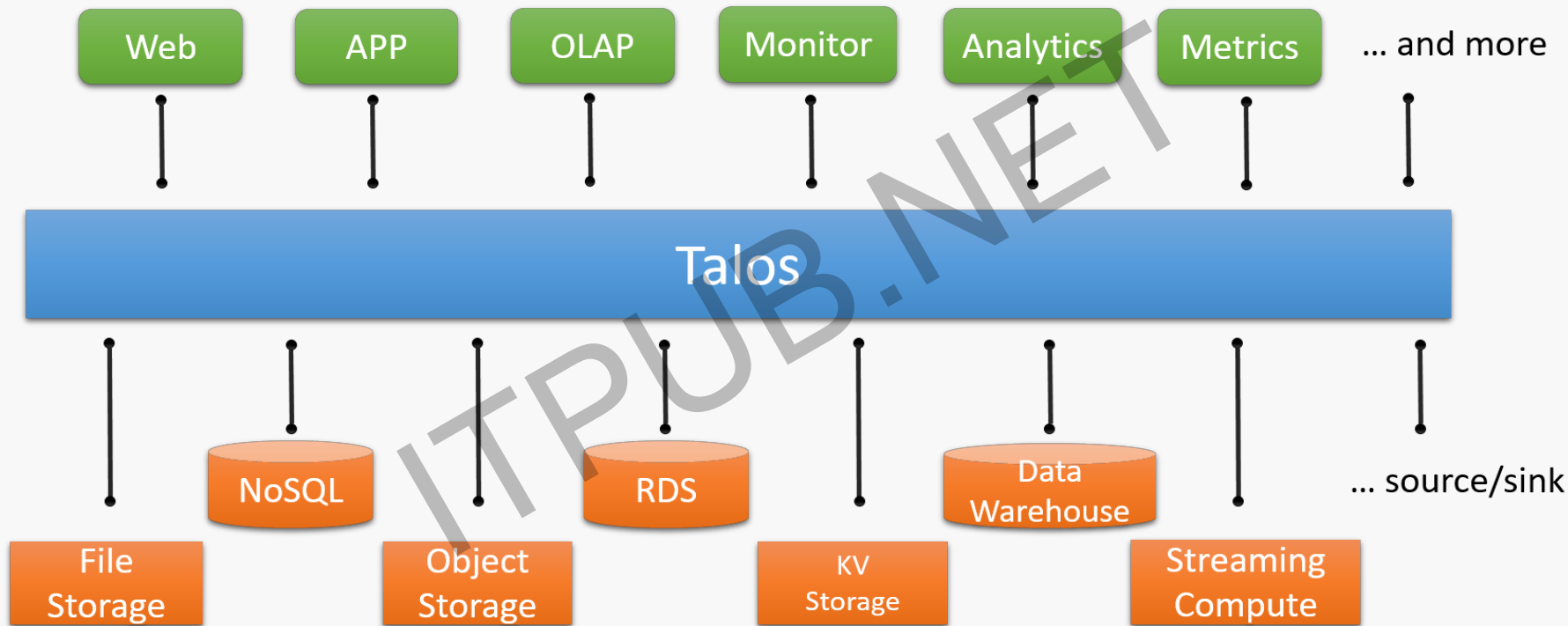
数据流整体框架



十年架构 成长之路



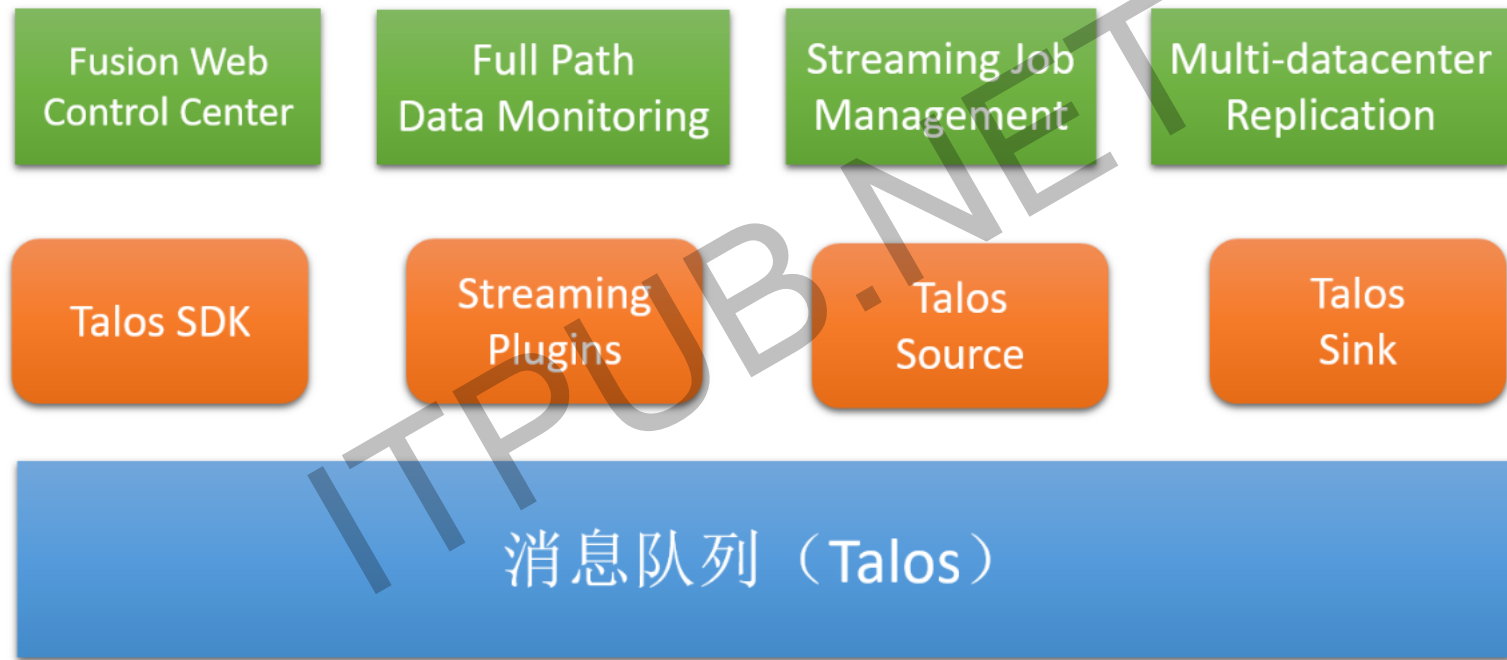
整体框架



十年架构 成长之路



核心模块



十年架构 成长之路



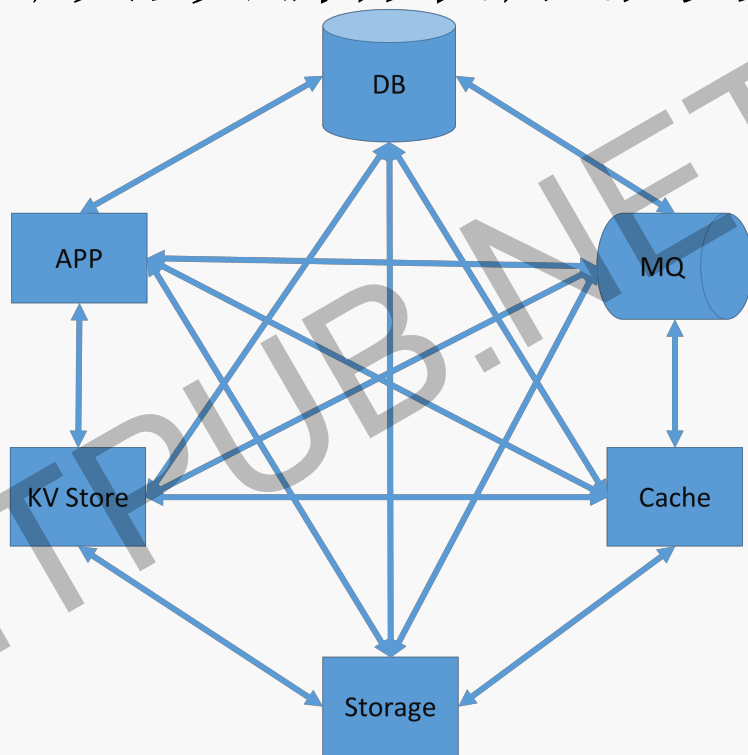
核心功能



十年架构 成长之路



一般数据集成方案



十年架构 成长之路

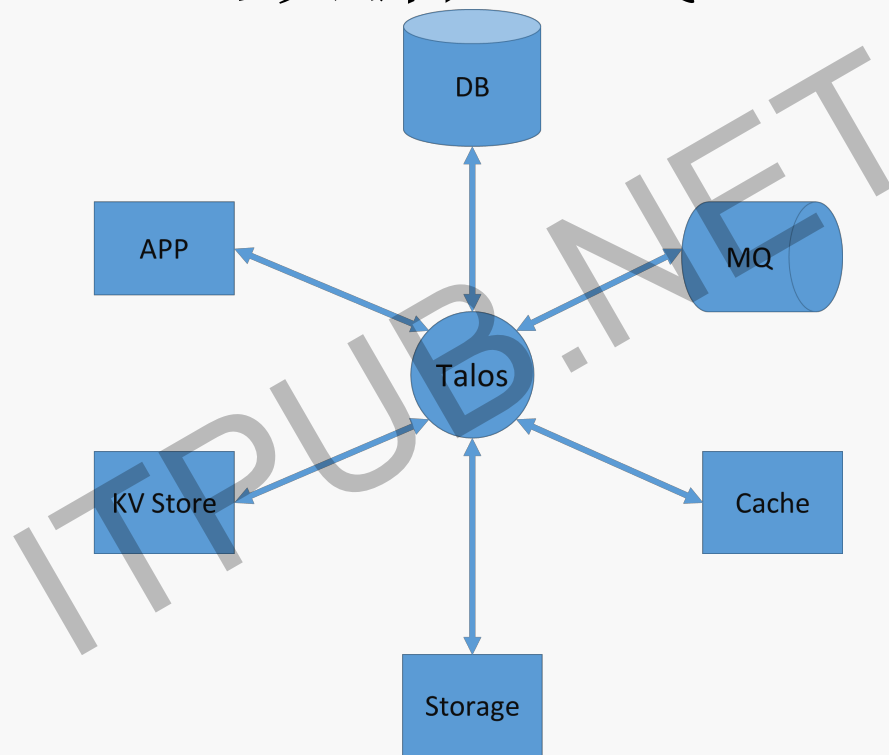


问题剖析

- 任意系统之间都需要进行交互，系统交互复杂度较高： $N*(N-1)$
- 当涉及的系统过多时，系统复杂度往往非常高，因此前期**研发成本**非常高，后续**功能添加**、**系统运维**、**系统重构**和**交接成本**会更高
- 各个业务均按照自身需求进行开发，但是由于彼此独立，因此**无法复用重复逻辑**
- 对于系统监控和数据质量，一般缺失或者很难做到很完备，因此无法保证**数据交互质量**
- 一般由业务独立部署，基本无法**抽象化**和**服务化**，进而也很难**积累经验**和**传承知识**



数据总线



十年架构 成长之路

Multi Source & Multi Sink

- 让数据在不同的系统之间**流动**起来
- 提供**产品化封装**，对所有业务团队屏蔽底层细节，大大降低了大数据应用的使用周期，避免业务**重复投入与研发**
- 所有数据均经过消息队列Talos中转，使得**流式计算**适用于所有数据
- Source & Sink组合的模式，系统集成复杂度降为**O(N)**
- 将众多的系统链接起来，不断适配新的系统，建立**数据流生态系统**



十年架构 成长之路



系统监控

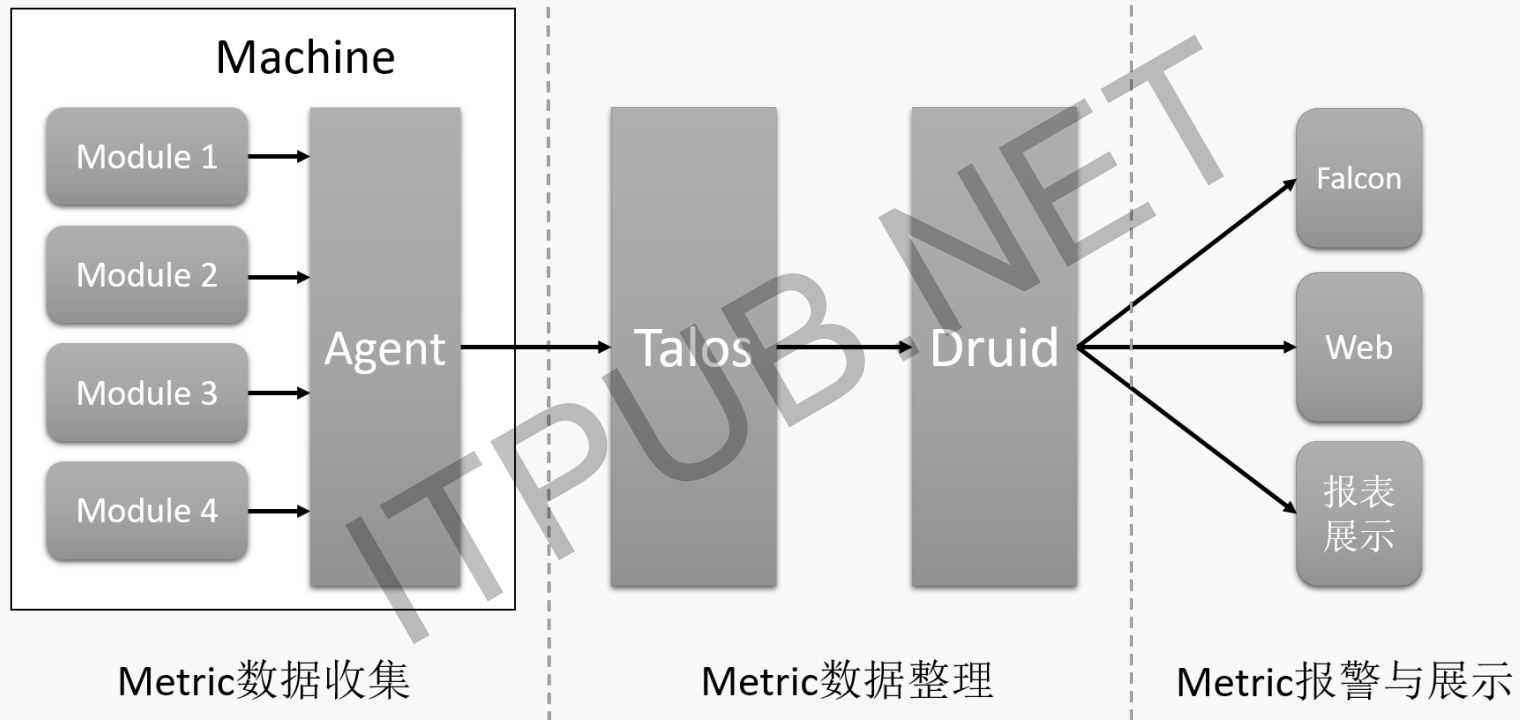
- 数据丢失监控
- 提数据延迟增加监控
- 服务进程异常监控
- 流量异常监控



十年架构 成长之路



系统监控



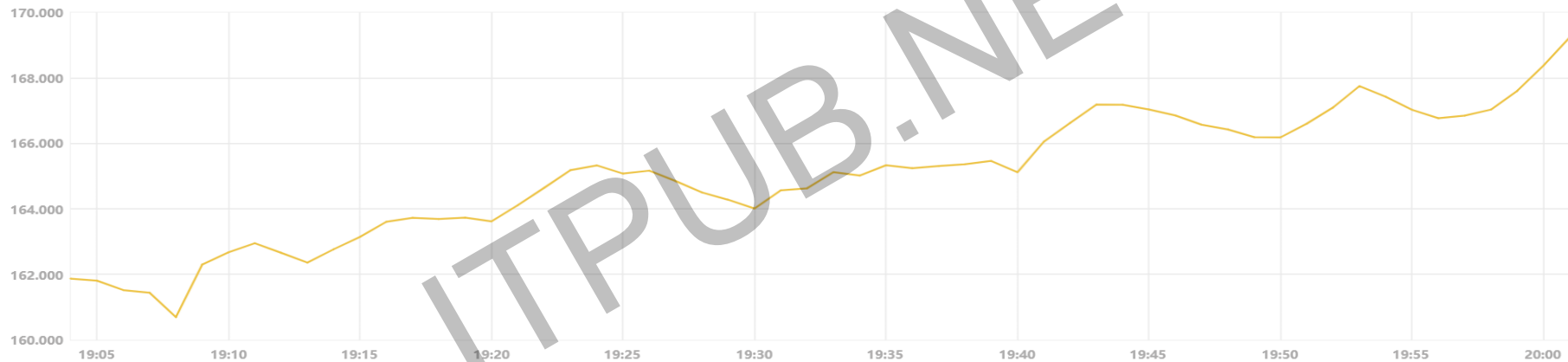
十年架构 成长之路



系统监控

时间跨度: -3600 采样方法: AVERAGE 求和: 原图

DropMessageNumber.300sRate



last ↓

169.245

min ↓

160.694

avg ↓

165.019

max ↓

169.245



刷新

输入字符过滤, 回车刷新



十年架构 成长之路



数据流端到端审计

- 量化数据流整体链路情况
 - 数据延迟：3W1H（**Which** stream **When** and **Where** delay **How many** data）
 - 数据丢失：3W1H（**Which** stream **When** and **Where** drop **How many** data）
- 展示形式：
 - Web页面：实时和历史数据展示（用户主动）
 - 报表推送：周期性汇总数据（用户被动）



十年架构 成长之路



数据流端到端审计

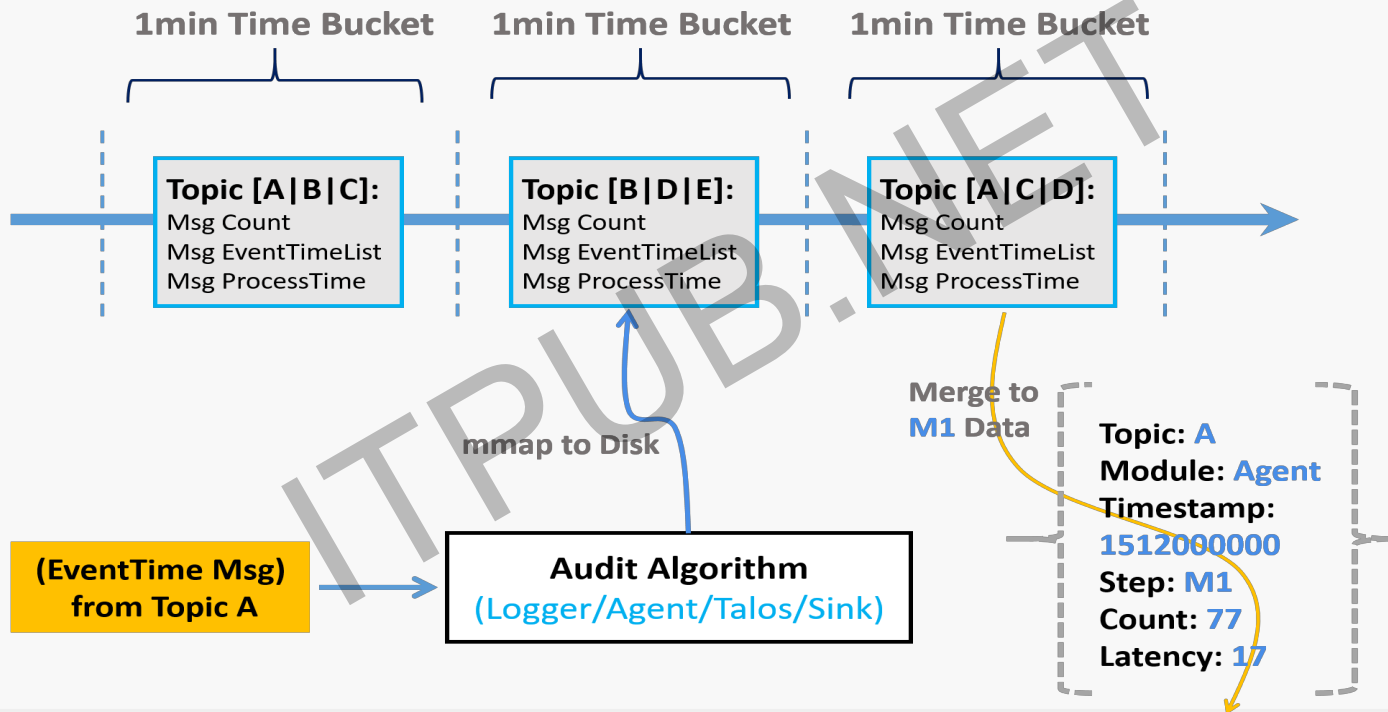
- **Event Time:** 消息实际产生的时间 (time at which messages actually **occurred**)
- **Processing Time:** 消息出现在某个模块的时间 (time at which messages are **observed** in the system)
- **Event Time:** 用于统计数据丢失的情况
- **Processing Time:** 用于统计数据延迟的情况;



十年架构 成长之路



数据流端到端审计



十年架构 成长之路



数据流端到端审计

开始时间

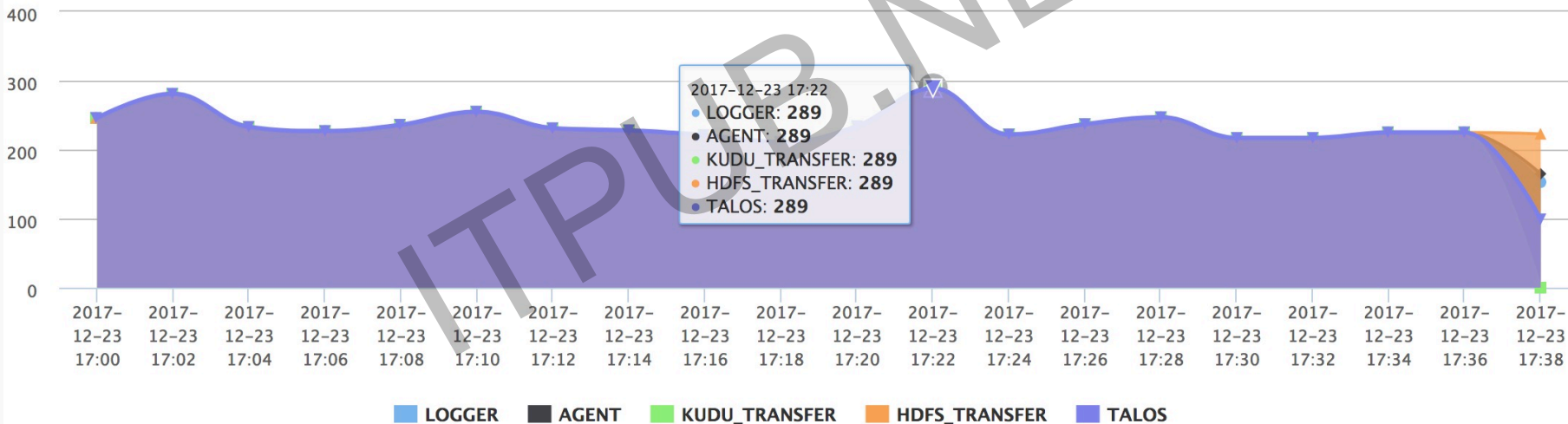
2017-12-23 16:40

结束时间

2017-12-23 17:40

查询

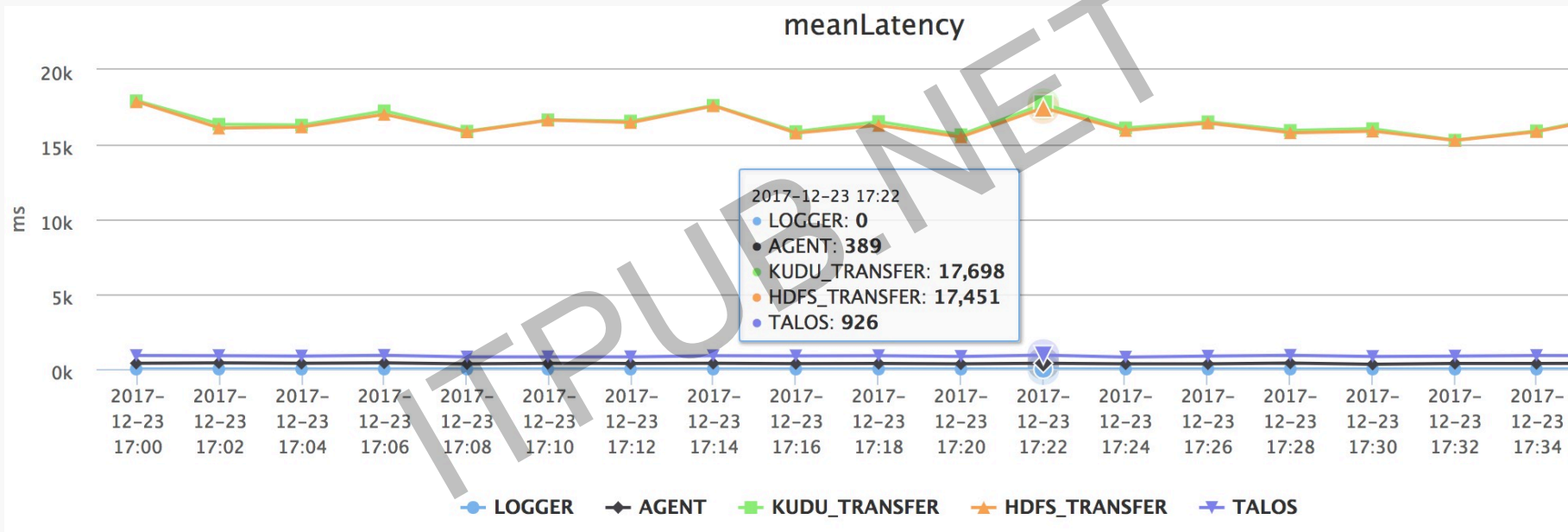
Message Received



十年架构 成长之路



数据流端到端审计



十年架构 成长之路



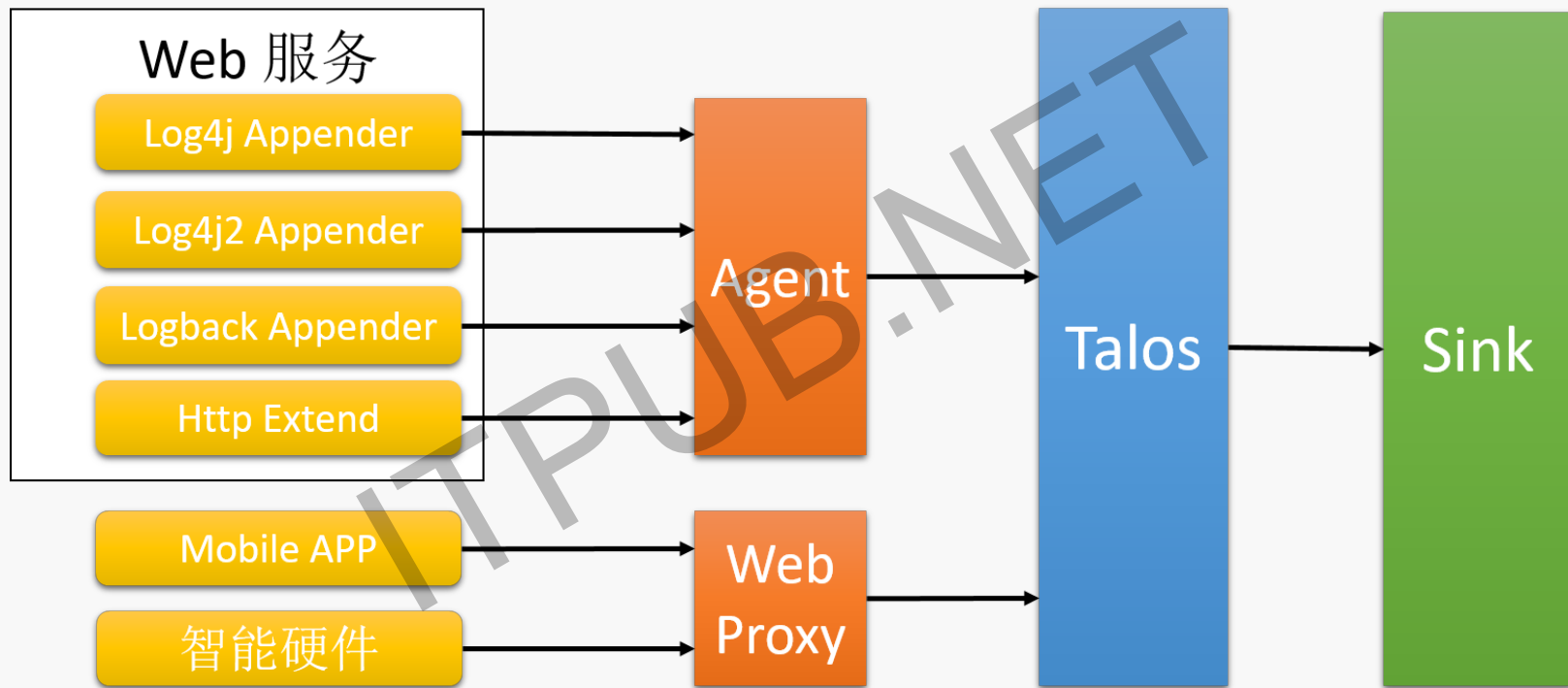
应用场景解析



十年架构 成长之路



埋点数据收集



十年架构 成长之路



埋点数据收集

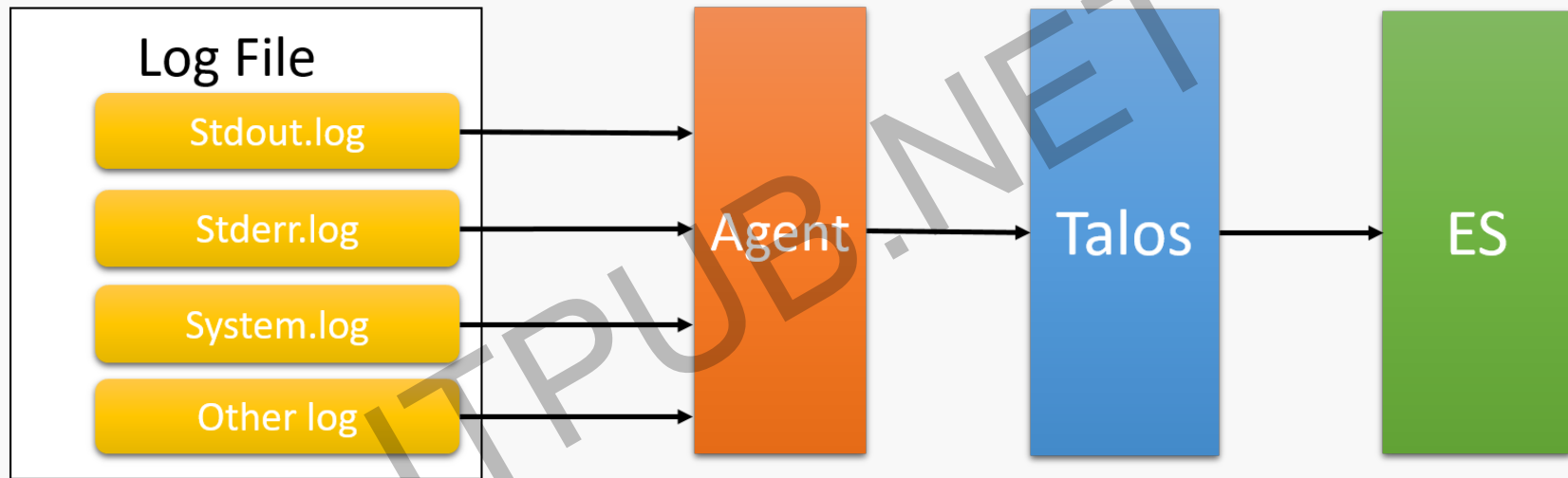
- 服务端埋点收集：
 - Java logger扩展
 - 其他语言http支持
- Web Proxy
 - App数据收集
 - 智能设备数据收集



十年架构 成长之路



实时日志分析

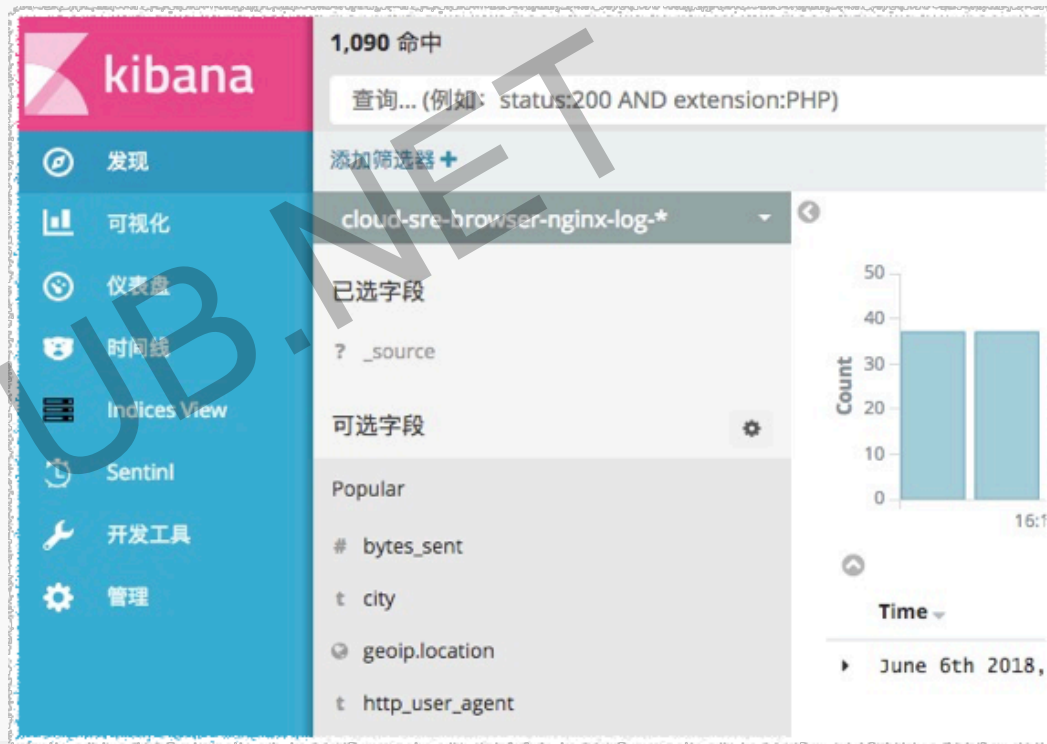


十年架构 成长之路



实时日志分析

- 日志准实时收集
- 日志文件动态更新
- Kibana可视化日志分析

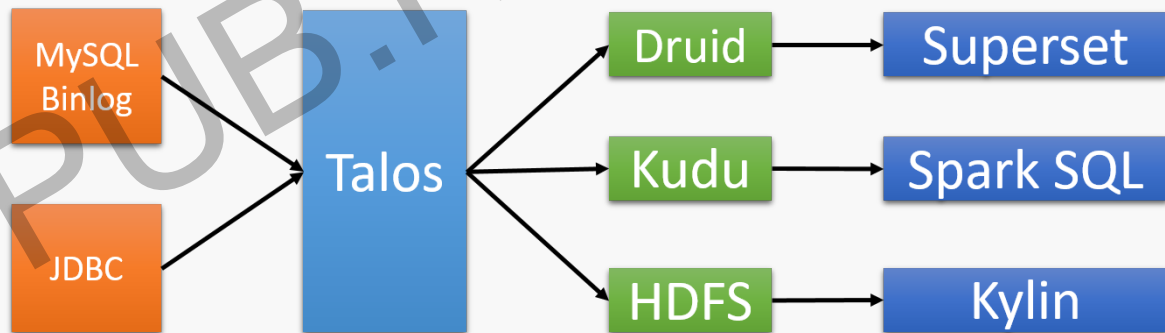


十年架构 成长之路



泛OLAP场景

- Druid: 多维度分析
- Kudu: 即时查询
- Kylin: T+1查询

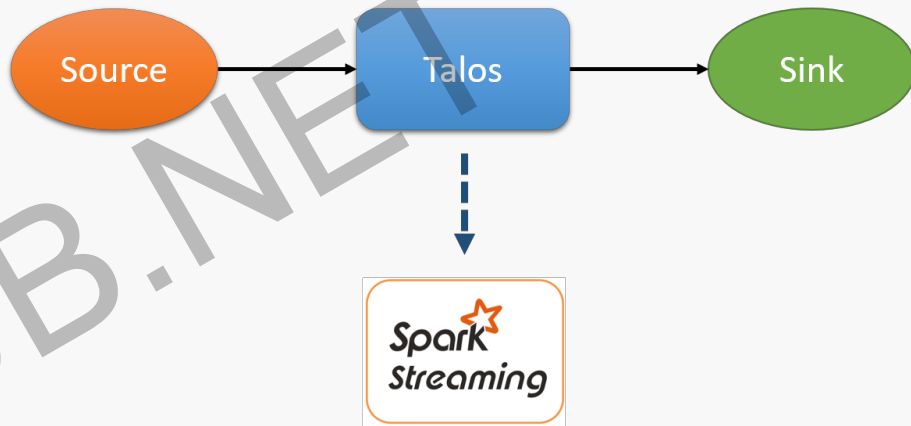


十年架构 成长之路



流式计算

- 秒级延迟
- 容错机制
- 支持exactly once语义



十年架构 成长之路





A network diagram consisting of several blue circular nodes connected by thin blue lines, forming a web-like structure across the upper half of the image.

THANKS



A large, light gray watermark text "ITPUGS.NET" is oriented diagonally across the center of the image, partially overlapping the "THANKS" text.



Abstract geometric shapes in the bottom right corner, including overlapping triangles and curved bands in shades of pink, orange, yellow, and light blue.