



SDCC 2016

中国软件开发者大会

SOFTWARE DEVELOPER CONFERENCE CHINA

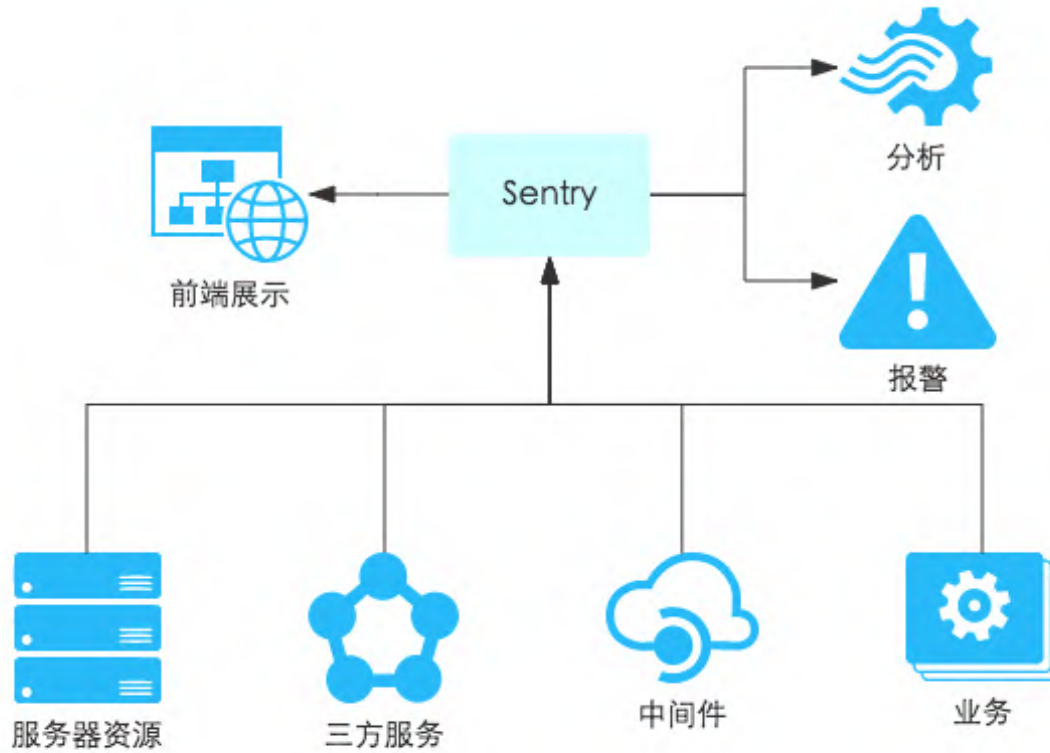


蘑菇街监控系统架构演进

嘉宾：卢鹏程



- 监控系统概述
- 需求/设计目标
- 架构及其演进
- 持续演进
- 经验总结



蘑菇街监控系统概述

应用场景

- 监控系统概述
- 需求/设计目标
- 架构及其演进
- 持续演进
- 经验总结

设计目标

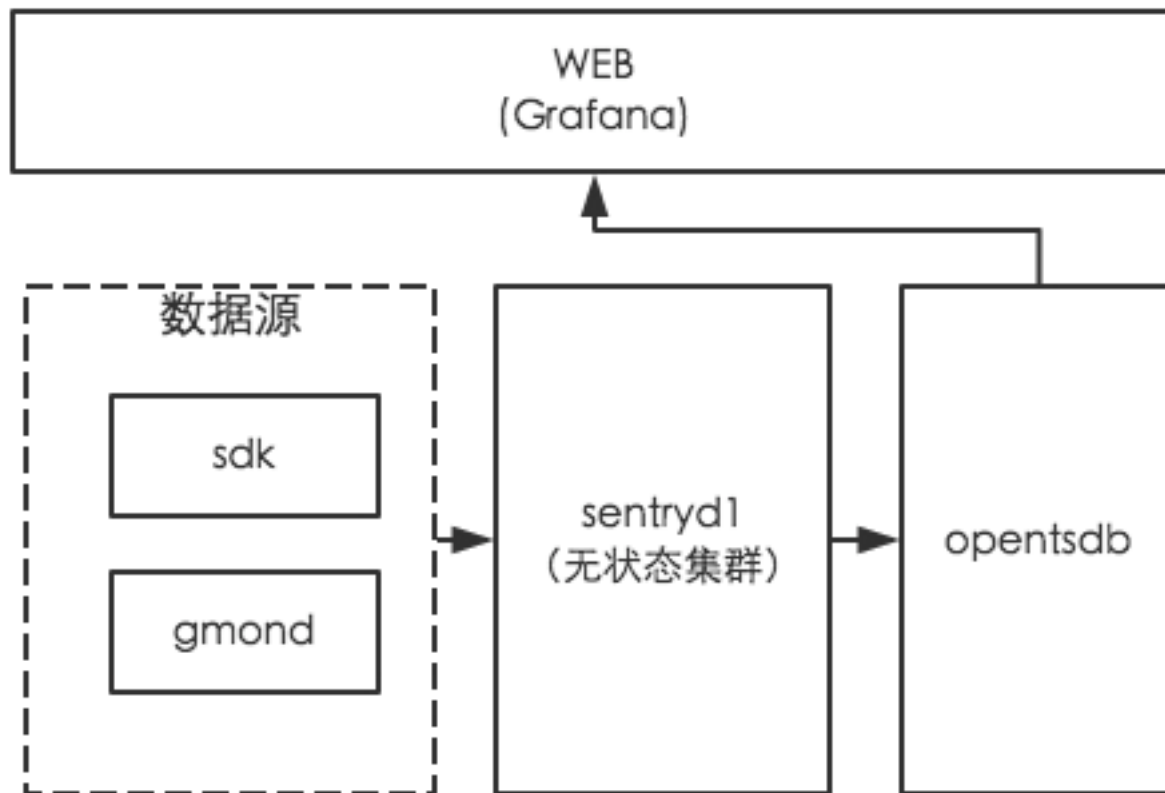
- Opentsdb兼容格式 (metric、tags、time、value)
- 查询时聚合能力 (单次万条曲线一秒内返回)
- 秒级数据
- 单机40万 tps
- 单机2万 qps
- 单机千万级活跃曲线
- 高可用，多机房

- 监控系统概述
- 需求/设计目标
- 架构及其演进
- 持续演进
- 经验总结

V1.0——从无到有

- 版本要求
 - 尽快提供服务
 - 兼容已有监控数据
 - 可平滑过渡到后续版本
- 选型
 - Gmond
 - Sentryd1 (proxy)
 - Opentsdb

V1.0——架构总览



V1.0——功能与缺陷

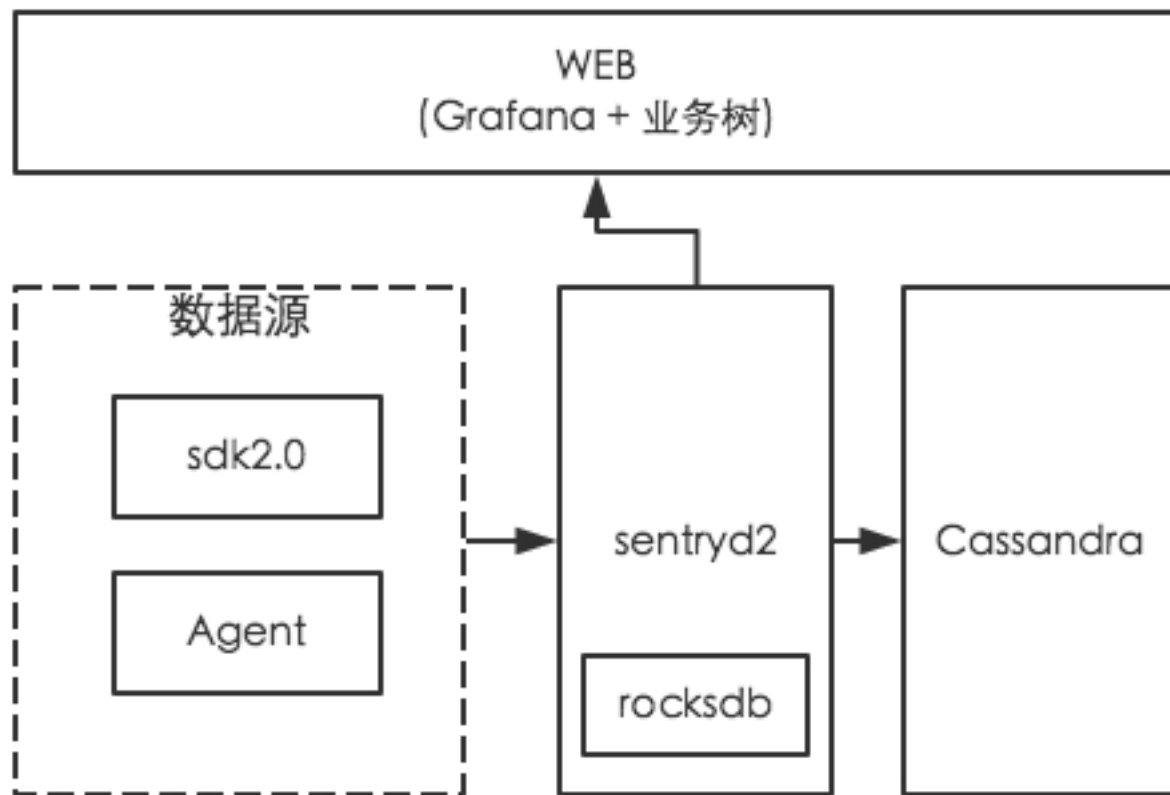
- Sentry1 (无状态)
 - 转发
 - 白名单
 - 元信息查询
 - 转码/fix
- 坑
 - 查询慢——聚合慢
 - 中文及特殊字符支持
 - 元信息
 - HTTP上报
 - 采样/聚合方式不完善，无法扩展

V2.0——优化试探

- 版本要求
 - 弥补1.0时代的几个不足
 - 兼容已有监控格式
 - 优化效率
 - 实验各种想法与方案
- 选型
 - 无依赖
 - Sentryd2去中心 / 本地存储 (rocksdb)
 - Agent
 - Cassandra

V2.0——架构总览

- Sdk封装优化，更易用
- Agent，可控/高效/反馈
 - 与sentryd2 TCP长连
 - 采集脚本针对性下发
- Sentryd2，无中心
 - rocksdb本地存储
 - gossip无中心分布式
- 历史数据存cassandra



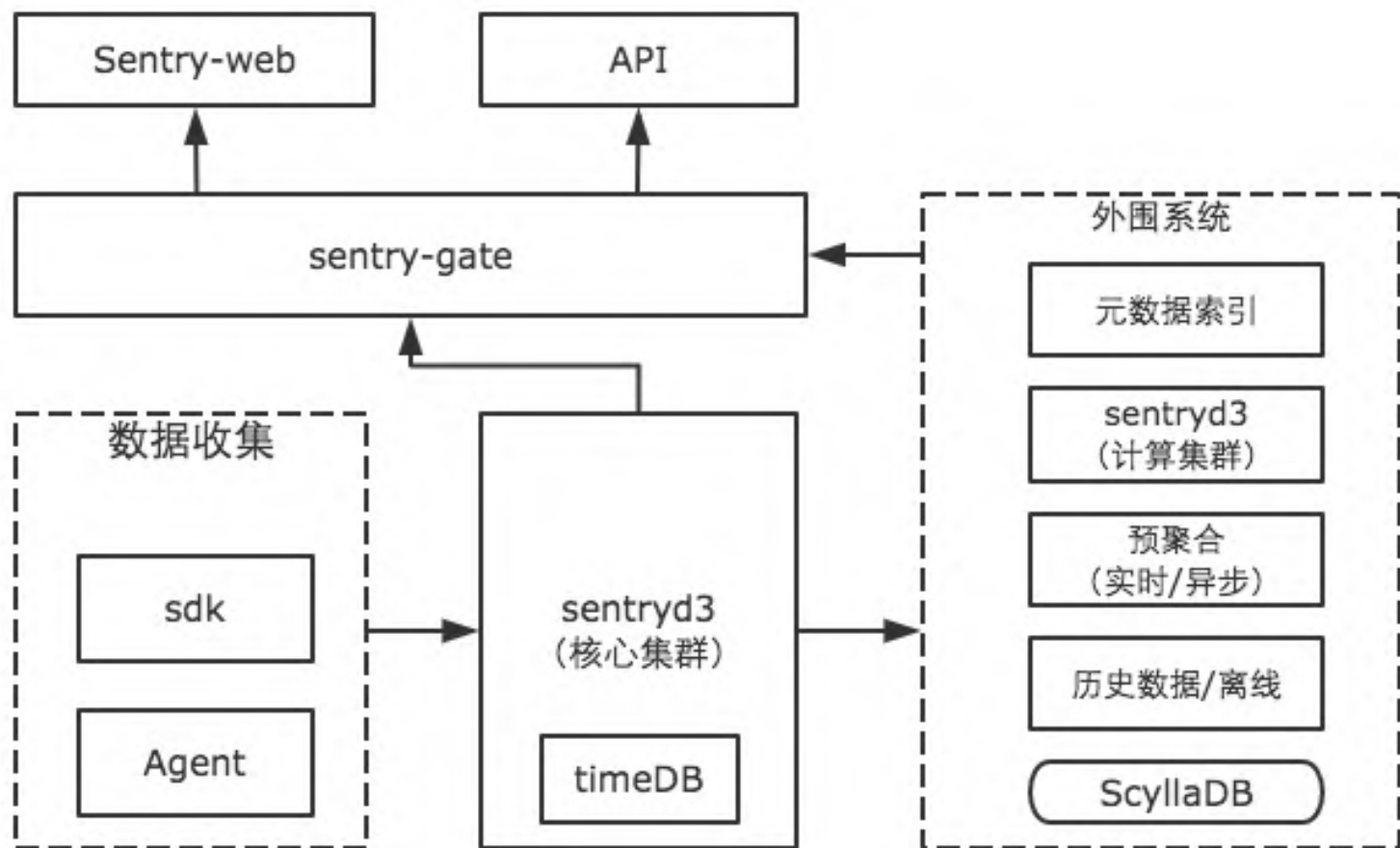
V2.0——不足

- Rocksdb这个轮子不太合适，时间序列数据有其特点
 - 无追加模式？Merge？
 - compact性能
 - 写放大
 - TTL问题
 - 数据迁移复杂
- 大促临时缩扩容问题
- 大查询阻塞

V3.0——曙光

- 版本要求
 - 全新架构。第一个完善版本。
 - 可用（查询、稳定、）
 - 兼容V2 API
 - 优化效率，提升4倍写入能力
 - 具备快速缩扩容能力
- 选型
 - 造轮子，自研TimeDB
 - Scylladb
 - 去Grafana，重写Sentry-web

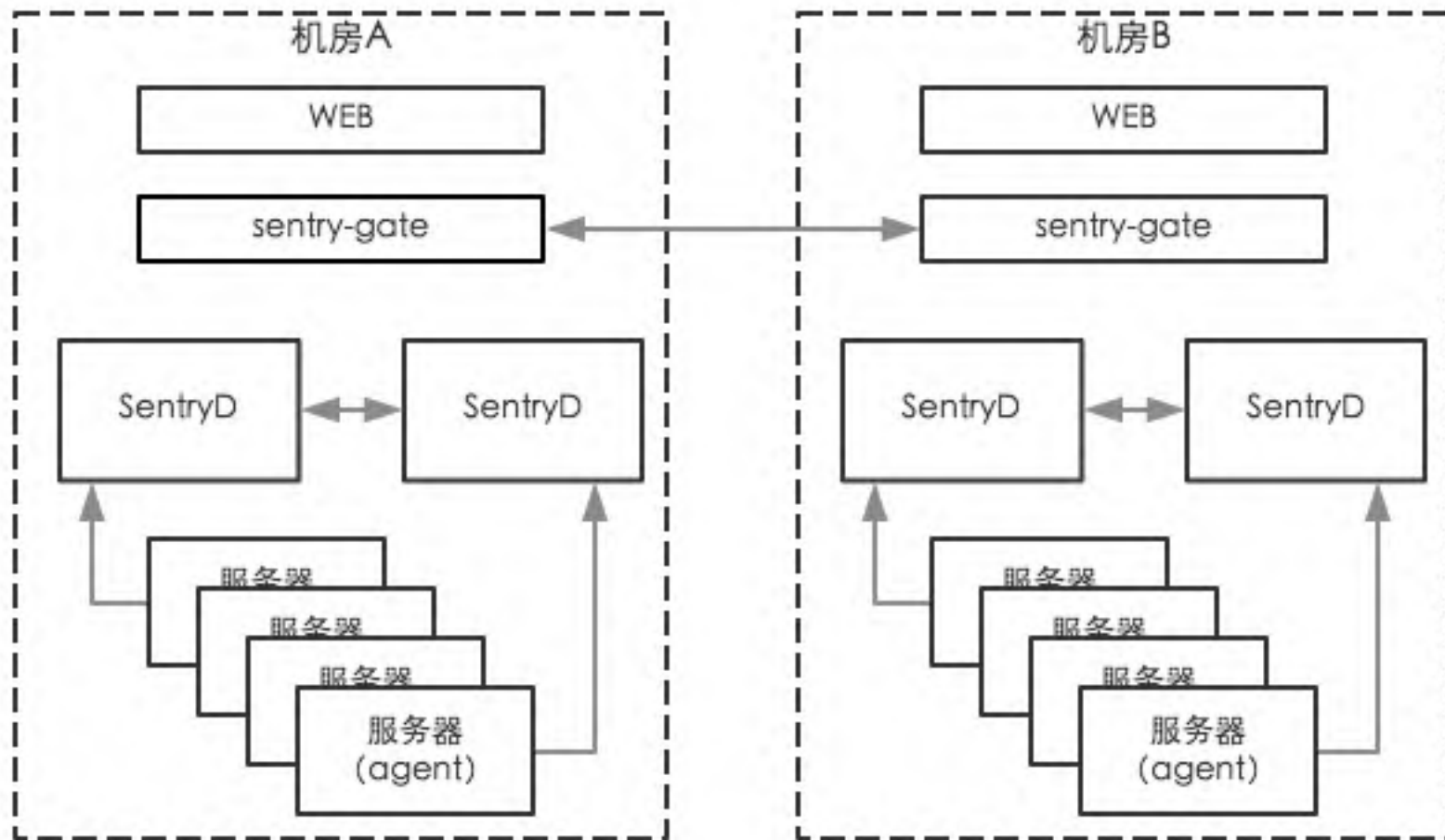
V3.0——架构总览



V3.0——架构特点

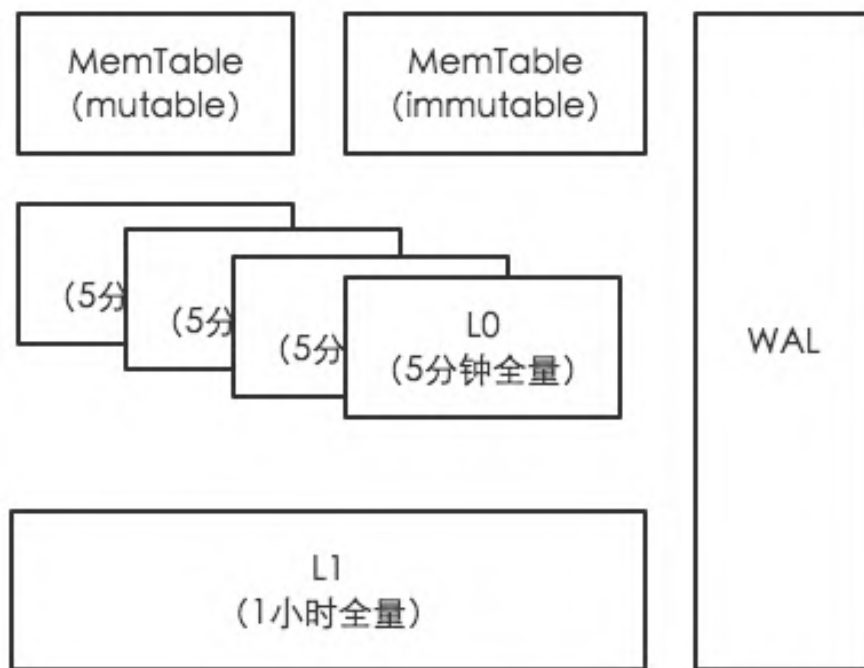
- 去中心
 - gossip + SWIM
 - 易部署、易维护、高可用、低风险
- 存角色
 - 核心、外围分离，
 - 大查询、热查询分离
 - 不同集群不同策略，兼顾CAP
- 无依赖
- 分场景

V3——分布式架构



V3——TimeDB

- 针对时间序列数据
- 写放大少
- Compact可控，每小时1次
- 读取速度快
- L0\L1自描述，便携易迁移
- L0存SSD，L1存SATA
- 过期数据，直接文件删除



V3——核心集群/查询集群

- 相同功能，不同角色
 - 核心集群：保证长期数据服务，降级查询。L0@SSD，L1@SATA
 - 查询集群：1~4小时，短期数据服务。L0&L1@SSD
- 为什么？
 - 快速临时扩容，应对大促
 - 针对性扩容，应对特定大查询
 - 辅助系统维护

V3——其它

- 数据格式
 - Sum、count、count、min、max
- 预聚合
 - 进一步增强大查询效率
 - 增长查询区间
 - 降低核心系统压力

- 监控系统概述
- 需求/设计目标
- 架构及其演进
- 持续演进
- 经验总结

持续演进

- 两端进化
 - 数据采集：Height-level SDK
 - Web展示：控件模板
- 场景化
 - 上报简单
 - web端自动配置
 - 聚合？采样？Sum？Avg？Zim？Sum-sum？Sum-avg？Avg-sum？
 - 自动数据处理：抽象、采样、聚合。
 - 分级采样
- 智能化

- 监控系统概述
- 需求/设计目标
- 架构及其演进
- 持续演进
- 经验总结

经验总结

- Opentsdb
 - 不支持特殊字符
 - tags数限制
- RocksDB
 - Merge , TTL
 - key与写放大
- 插值 / 平移 / 补0
- 数据结构
 - 整形/浮点
 - Time/Value,SCCMM



SDCC 2016

中国软件开发者大会

SOFTWARE DEVELOPER CONFERENCE CHINA

谢谢！