ZABBIX FORUM SHENZHEN 2021

ZABBIX对大流量监控数据的高效处理

鲍光亚

《深入理解Zabbix监控系统》作者，Zabbix开源社区专家

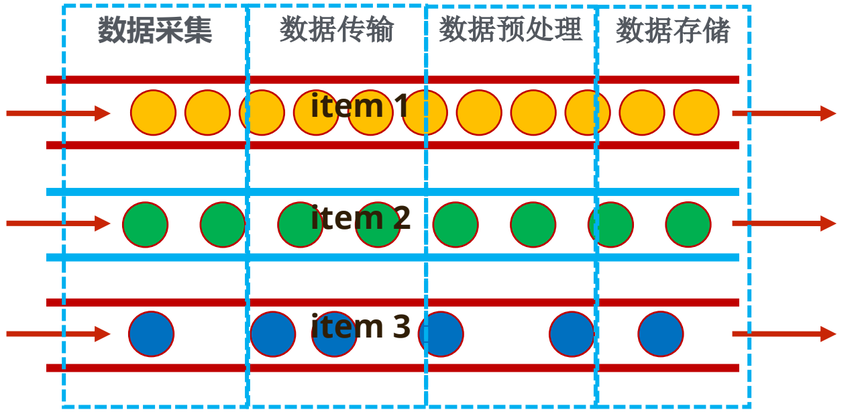
ZABBIX

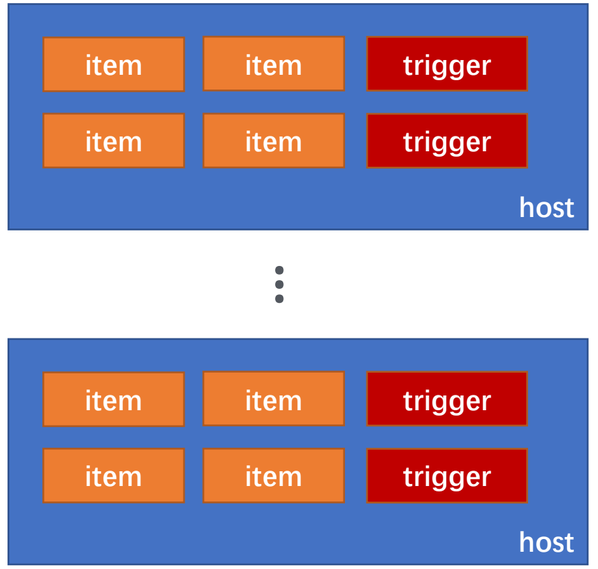
ZABBIX FORUM SHENZHEN 2021

监控数据的特点

监控数据的独立性——逻辑视角

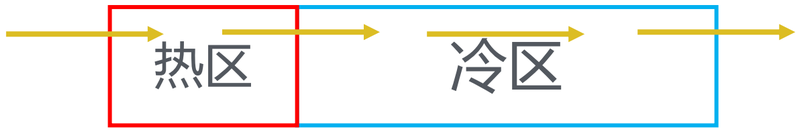
- 独立性的体现
  - 不同监控值（value）之间的独立性
  - 不同监控项（item）之间的监控数据
  - 不同监控对象（host）之间的监控数据
- 同一监控项的values构成时间序列，按照时间有序
- 元数据之间的**相对**独立性
- 为并行和可扩展性提供了基础
- 一切都指向……基于数据分组的微服务和并行化

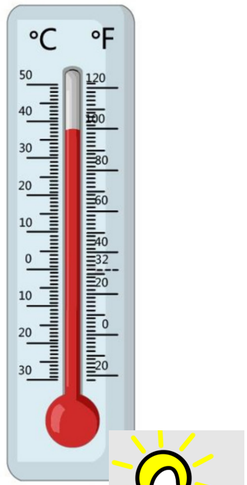




监控数据的热值以及热数据的规模

- 热值——访问频率（每分钟的访问次数）
- 热值与时间（age）高度相关
  - 数据写库之前
  - 触发器表达式计算对数据的访问范围（avg(60m) vs last(#2)）
  - 前端UI对数据的访问特点
- 冷热数据混存对热数据访问效率的影响
- 如果监控数据表拆分为热数据表和冷数据表
  - 如何实现热表的最小化？
- 元数据的热值
  - 数据规模和访问频率相对固定



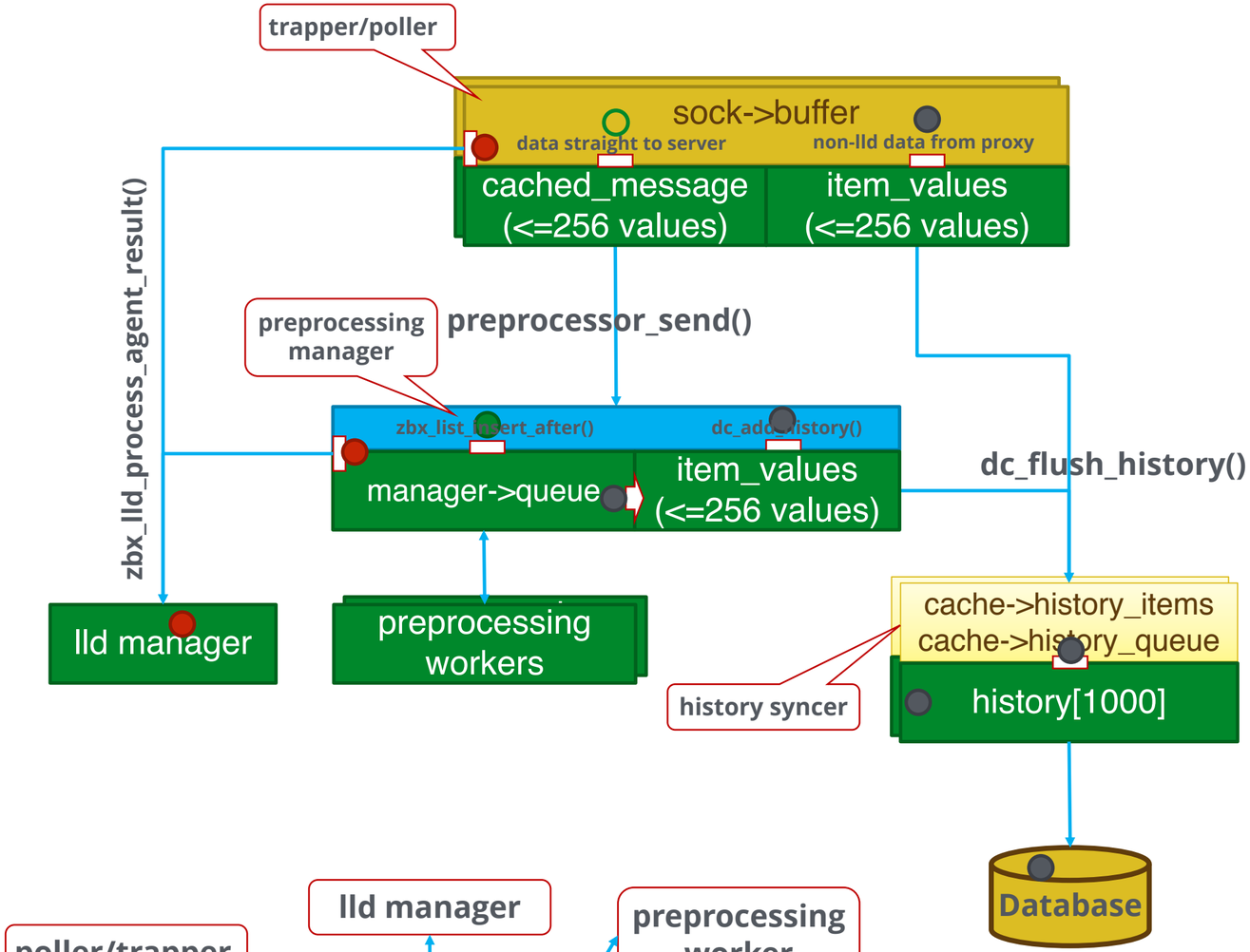


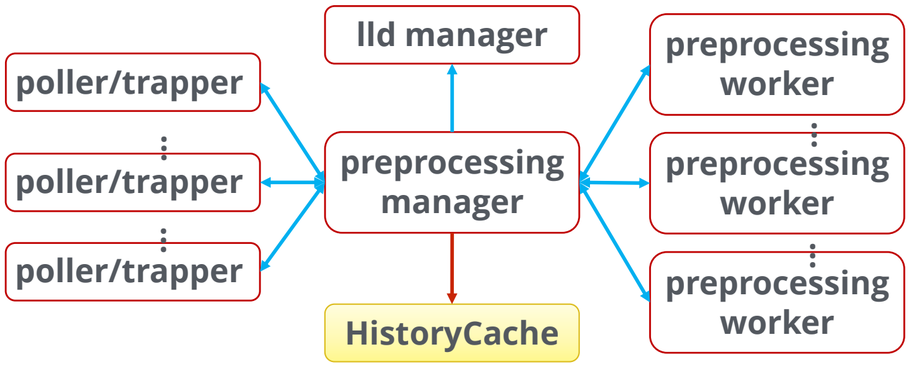
ZABBIX FORUM SHENZHEN 2021

ZABBIX对监控数据的高效处理

ZABBIX SERVER的数据IO

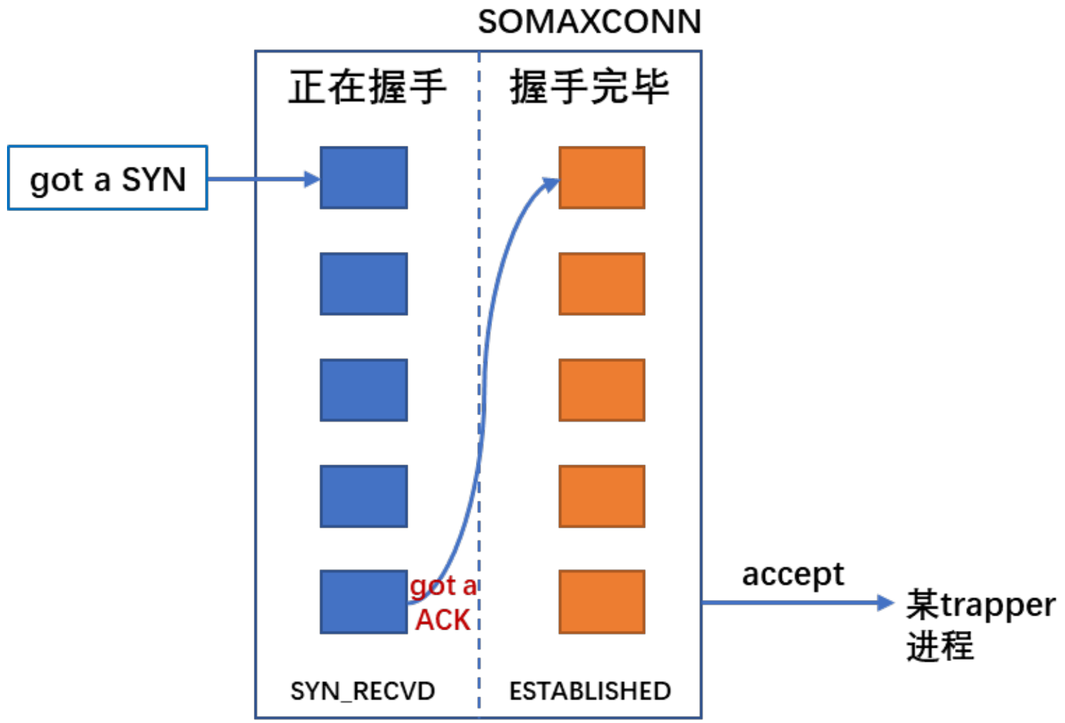
- trapper进程之间的数据分配均衡
  - trapper每个循环只处理一个连接
  - 每个trapper以相同几率accept连接
  - 如接受的连接包含相对较多的监控数据，trapper将花费更多时间处理数据
- poller进程之间的数据分配均衡
  - 每次请求一个值
  - 每个poller进程等概率获得任务
- 预处理进程的数据分配
  - manager与每个worker以及poller/trapper之间保持持久的unix domain socket连接
- history syncer进程的数据分配
  - 通过共享内存与上游进程通信
  - 输出数据写入数据库





大流量数据的接收

- 设计通信协议的考量
  - 消息中的数据量（单个值还是多个值）
  - 消息的格式：文本与二进制
- TCP连接用完即关闭
- 当监控规模很大时
  - 存在大量的对端主机
  - 需完成高频率的连接建立与关闭
  - poller与trapper
- 可扩展性——增加进程数量、监听多个地址



数据结构设计——解决处理速度问题

- 哈希集结构：解决快速查找问题
  - 哈希主键使用各种ID，itemid、hostid、triggerid、functionid ……
  - ConfigCache（对应CacheSize）——**item**、host、trigger等
  - HistoryCache和HistoryIndexCache——写库前的values
  - ValueCache——主要用于trigger计算
  - TrendCache——写库前的趋势数据
- 堆队列：解决快速排序问题
  - HistoryCache和HistoryIndexCache——按时间排序

元数据的访问与协调

- ConfigCache的内容与数据结构
- 元数据的作用
  - 其他进程对元数据的依赖
  - 元数据与数据库的同步
- 元数据的读写冲突与协调
  - 读写锁的粒度与性能
  - 读写锁对并行度的影响
- configuration syncer为什么采用单进程？
  - 维持数据之间的一致性
  - 锁的开销

```
zbx_hashset_t      items;
zbx_hashset_t      items_hk
zbx_hashset_t      template_items;
zbx_hashset_t      prototype_items;
zbx_hashset_t      numitems;
zbx_hashset_t      snmpitems;
zbx_hashset_t      ipmiitems;
zbx_hashset_t      trapitems;
zbx_hashset_t      dependentitems;
zbx_hashset_t      logitems;
zbx_hashset_t      dbitems;
zbx_hashset_t      sshitems;
zbx_hashset_t      telnetitems;
zbx_hashset_t      simpleitems;
zbx_hashset_t      jmxitems;
zbx_hashset_t      calcitems;
zbx_hashset_t      masteritems;
zbx_hashset_t      preprocitems;
zbx_hashset_t      httpitems;
zbx_hashset_t      functions;
zbx_hashset_t      triggers;
zbx_hashset_t      trigdeps;
zbx_hashset_t      hosts;
zbx_hashset_t      hosts_h;
zbx_hashset_t      hosts_p;
zbx_hashset_t      proxies;
zbx_hashset_t      host_inventories;
```

```
zbx_hashset_t      host_inventories_auto;
zbx_hashset_t      ipmihosts;
zbx_hashset_t      htmpls;
zbx_hashset_t      gmacros;
zbx_hashset_t      gmacros_m;
zbx_hashset_t      hmacros;
zbx_hashset_t      hmacros_hm;
zbx_hashset_t      interfaces;
zbx_hashset_t      interfaces_snmp;
zbx_hashset_t      interfaces_ht;
zbx_hashset_t      interface_snmpaddrs;
zbx_hashset_t      interface_snmpitems;
zbx_hashset_t      regexps;
zbx_hashset_t      expressions;
zbx_hashset_t      actions;
zbx_hashset_t      action_conditions;
zbx_hashset_t      trigger_tags;
zbx_hashset_t      host_tags;
zbx_hashset_t      host_tags_index;
zbx_hashset_t      correlations;
zbx_hashset_t      corr_conditions;
zbx_hashset_t      corr_operations;
zbx_hashset_t      hostgroups;
zbx_vector_ptr_t   hostgroups_name
zbx_hashset_t      preprocops;
zbx_hashset_t      maintenances;
zbx_hashset_t      maintenance_periods;
zbx_hashset_t      maintenance_tags;
```

ConfigCache中的部分数据结构（Zabbix 5.0）

大流量数据的存储

- history syncer进程的工作过程
  - 先写库后计算trigger
  - insert语句的构造与执行
- 数据存储工作的特点
  - History表数据存储（冷热数据）
  - Trends表数据存储
- Trigger表达式计算过程
  - 访问ValueCache（必要时访问数据库）
  - 生成event（内部事件、trigger事件等）
- 事件数据的存储
  - events表、problem表、escalations表、alerts表

ZABBIX FORUM SHENZHEN 2021

Q&A ?

ZABBIX

ZABBIX FORUM SHENZHEN 2021

THANK YOU!

ZABBIX

ZABBIX FORUM SHENZHEN 2021

本次深圳大会照片 https://live.aiyaopai.com/live/52886185

往期大会演讲视频 https://space.bilibili.com/476625813

微信交流群：17502189550

公众号干货：Zabbix开源社区