

GOTC

全球开源技术峰会

THE GLOBAL OPENSOURCE TECHNOLOGY CONFERENCE

OPEN SOURCE , OPEN WORLD

「CNCF云原生」专场

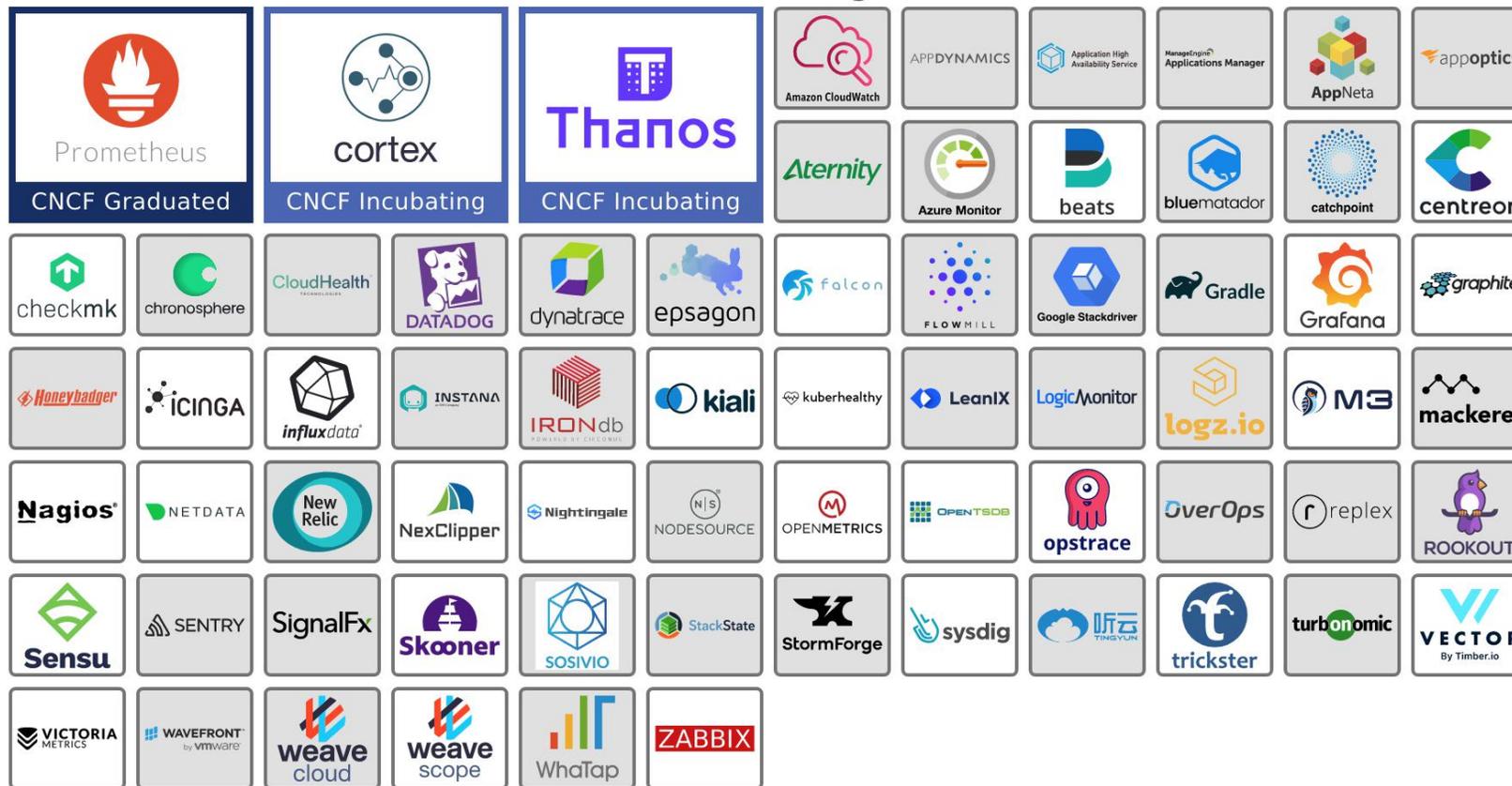
本期议题：如何进行多集群统一监控

李平辉 2021年08月01日

1. 云原生监控概览
2. 多集群场景的监控需求
3. 方案的对比选型
4. 实践总结

Prometheus是Kubernetes监控的事实标准

Monitoring



为什么是Prometheus

- 多维度数据模型

基于键值对，类似于Kubernetes用label组织数据，可以提供灵活的时序数据存储查询。

- 简单直接的指标格式和接口

Prometheus指标由HTTP协议暴露，是自描述、可读的格式。可以简单地用浏览器作检查。

- 服务发现

可以自动发现和更新指标采集的目标，从而适应容器平台易变的应用负载。

- 模块化组件

指标采集、告警、图形化交给不同组件去做，各组件支持高可用。

Prometheus does one thing, and it does it well.
—— Brian Brazil

Prometheus的局限

- 基于指标(Metrics)的拉取(Pull)模型
非事件驱动。
统计误差, 如速率函数推断, 降采样降低数据精度。
- HA的数据一致性
多副本普罗米修斯不保证数据一致性。
- 横向扩缩容
不能支持灵活的横向扩缩容。

管理多集群是普遍的需求： 77.8%的受访者表示在使用多云或混合云
—— K8s和云原生运维报告, Canonical, 2021

多集群场景的监控需求

- 全局视图
 - 能聚合查询多个集群的基础设施和应用服务的监控指标
- 提供良好的扩容能力
 - 避免全局监控系统随集群/服务数量的增加成为瓶颈。
- 监控系统与受监控服务之间的可达性
 - 监控系统需要能采集到不同集群的服务指标。
- 可用性
 - 监控系统的高可用。

多集群监控方案对比

GOTC

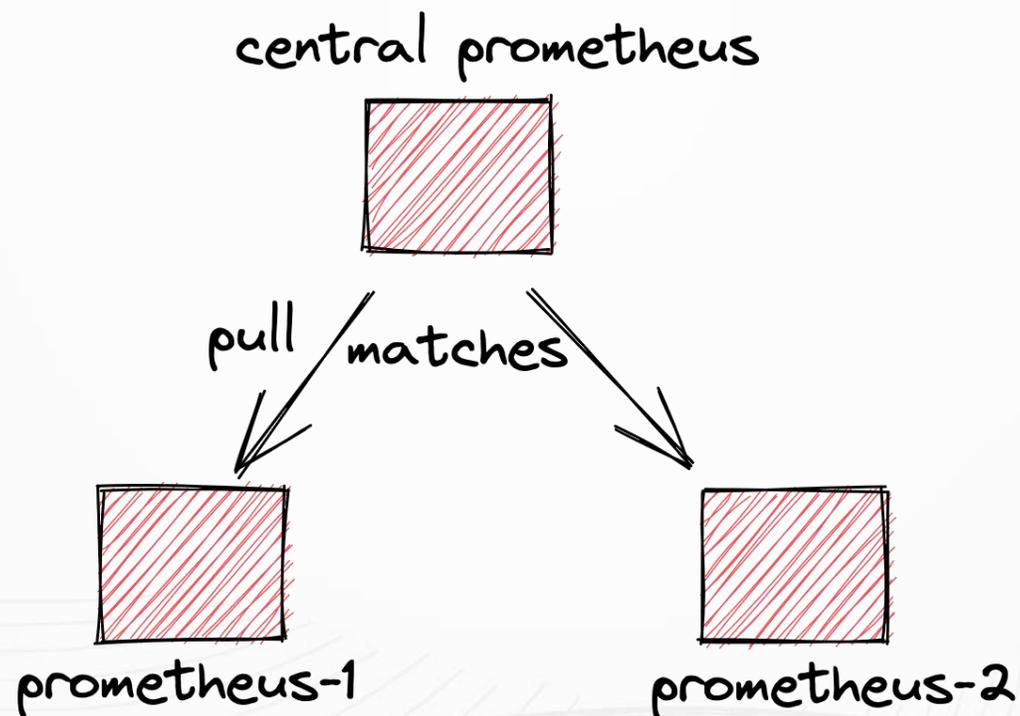
1. Prometheus
2. Cortex
3. Thanos

全球开源技术峰会

THE GLOBAL OPENSOURCE TECHNOLOGY CONFERENCE

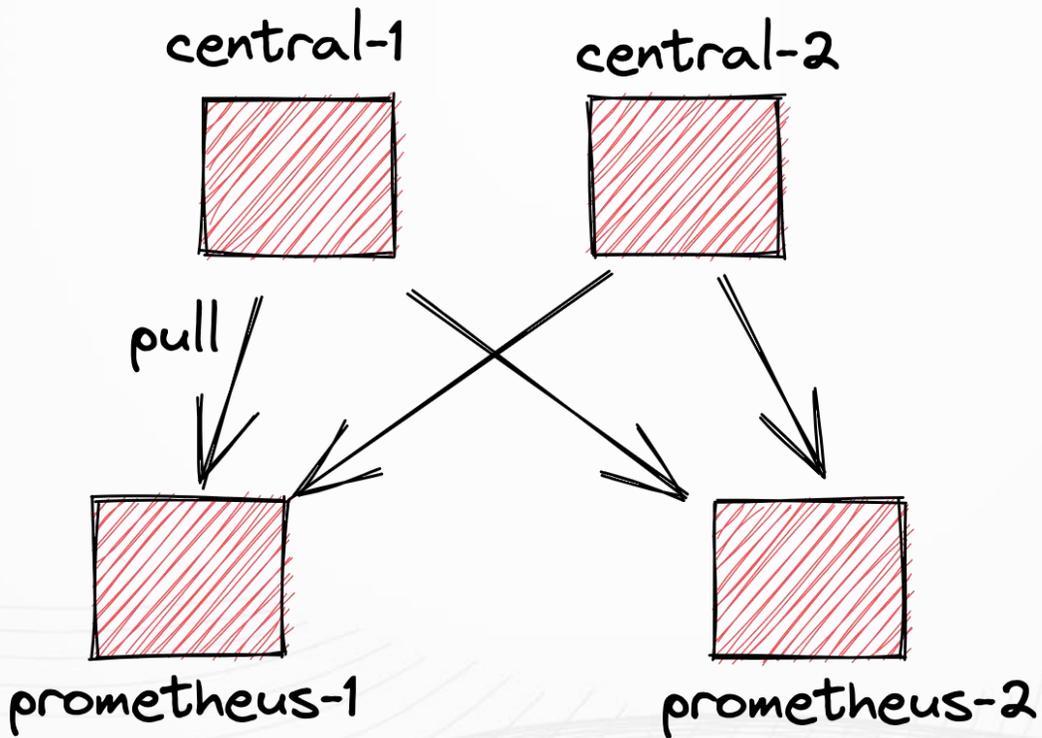
Prometheus Federation

示意图



Prometheus Federation

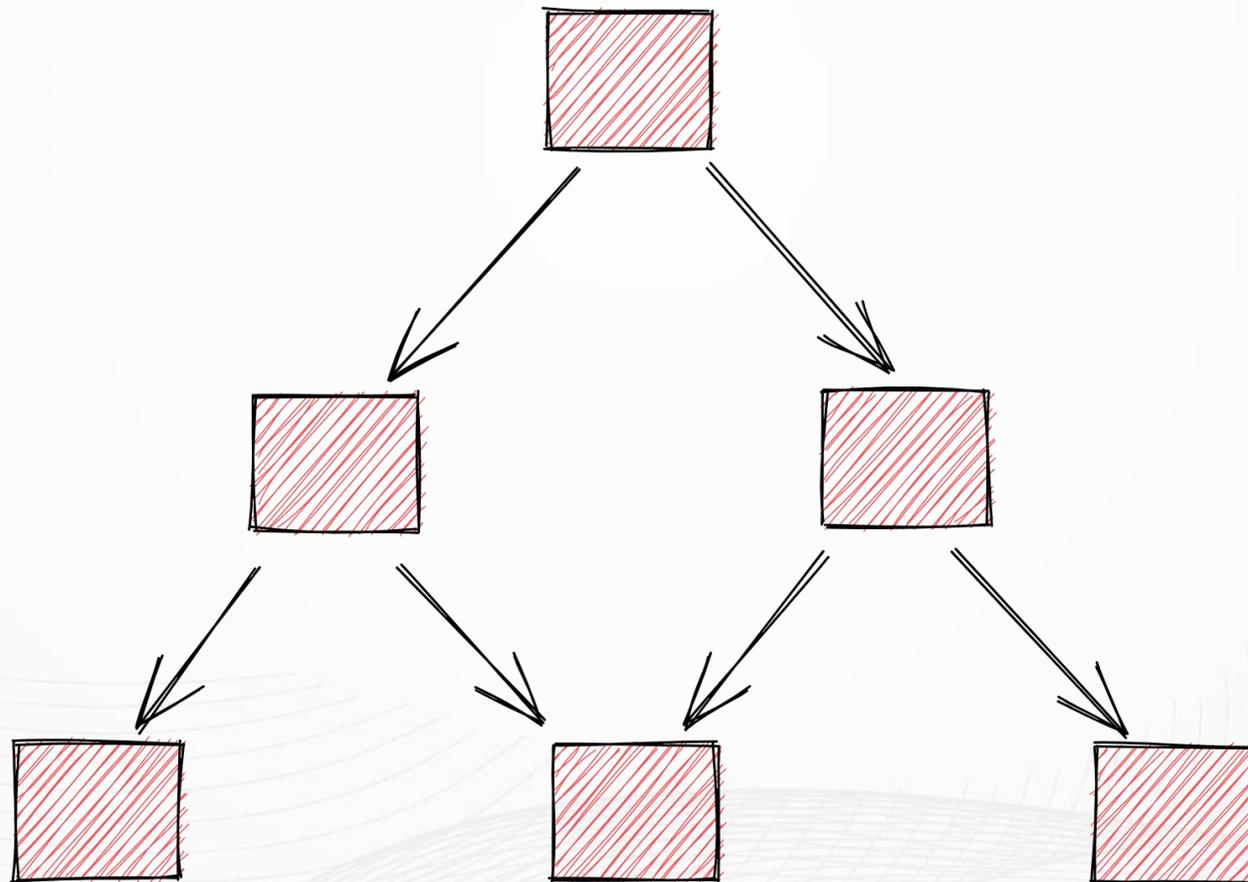
示意图



Prometheus Federation

GOTC

示意图



全球开源技术峰会

THE GLOBAL OPENSOURCE TECHNOLOGY CONFERENCE

► Prometheus Federation

GOTC

总结:

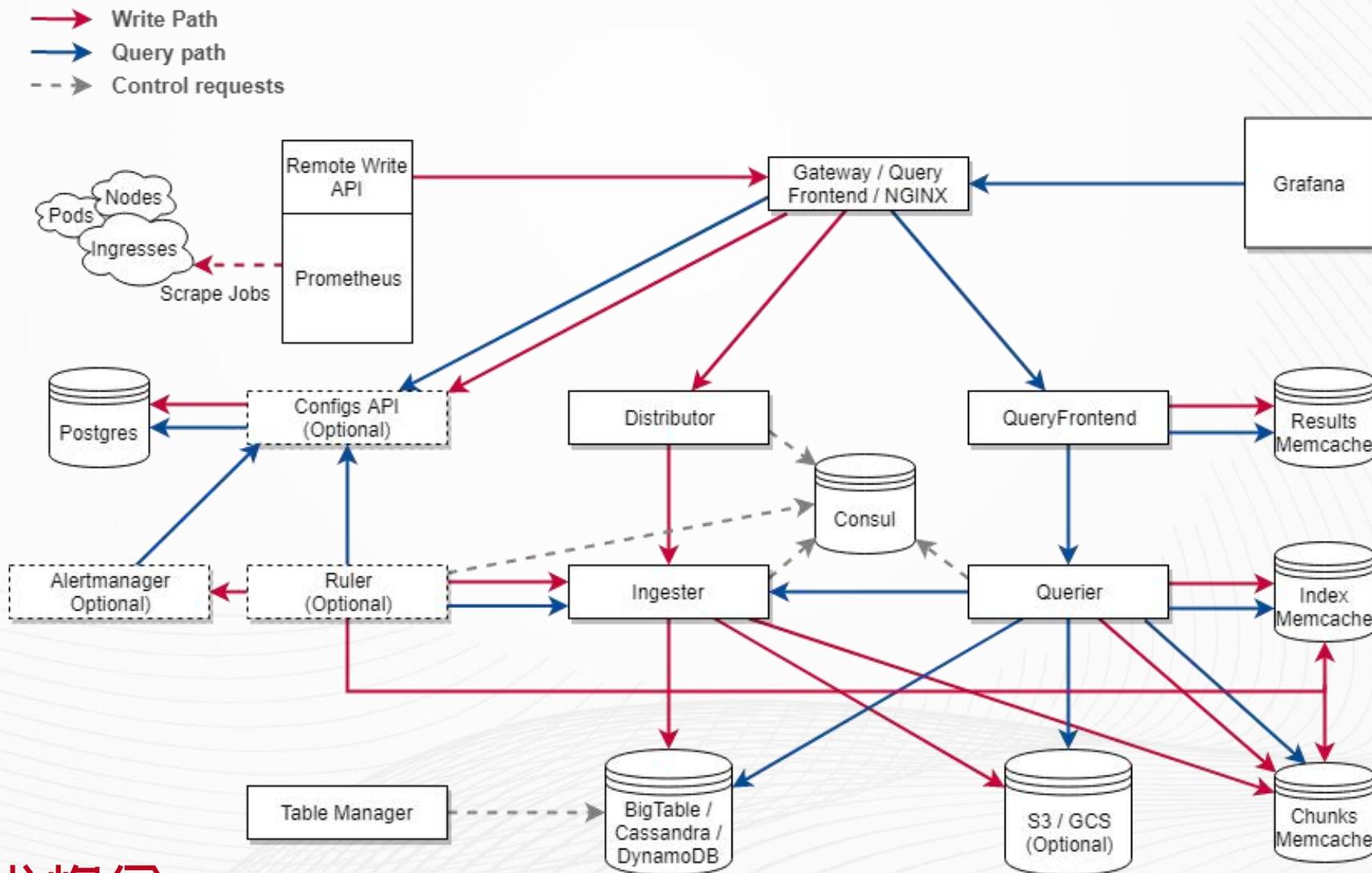
- 架构简单
- 横向扩容

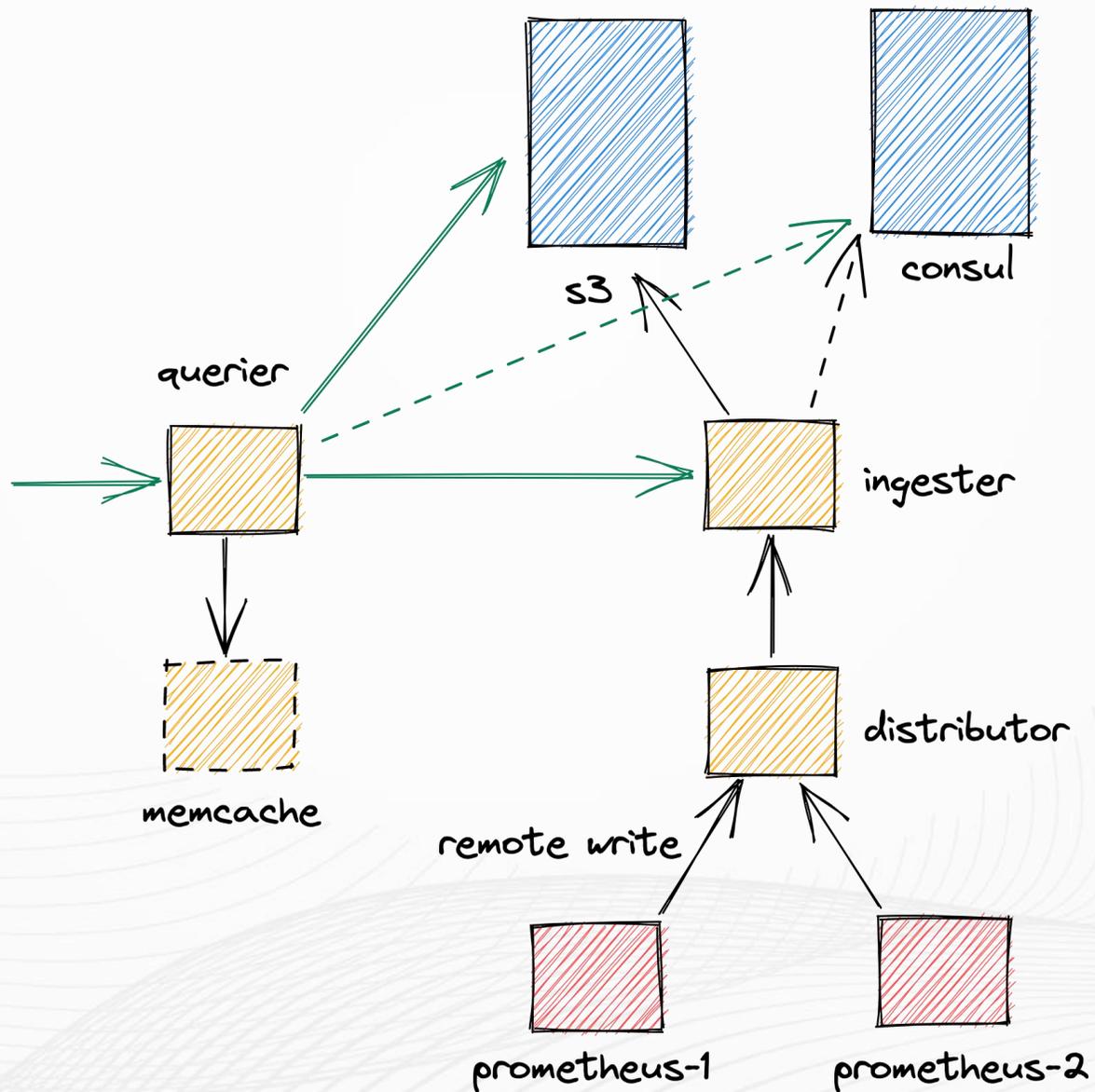
- 数据子集
- 重复采集
- 多副本数据去重

全球开源技术峰会

THE GLOBAL OPENSOURCE TECHNOLOGY CONFERENCE

示意图

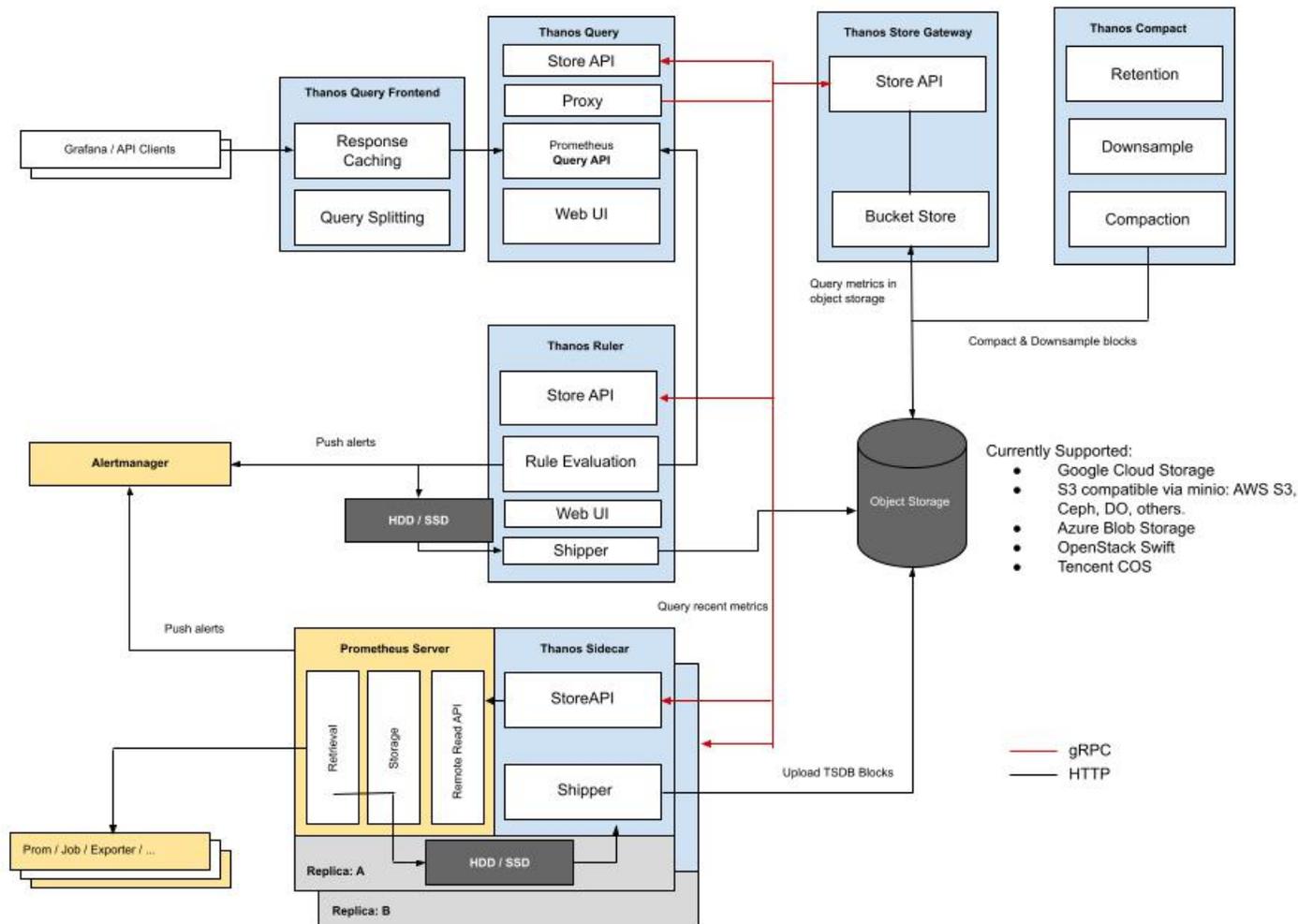




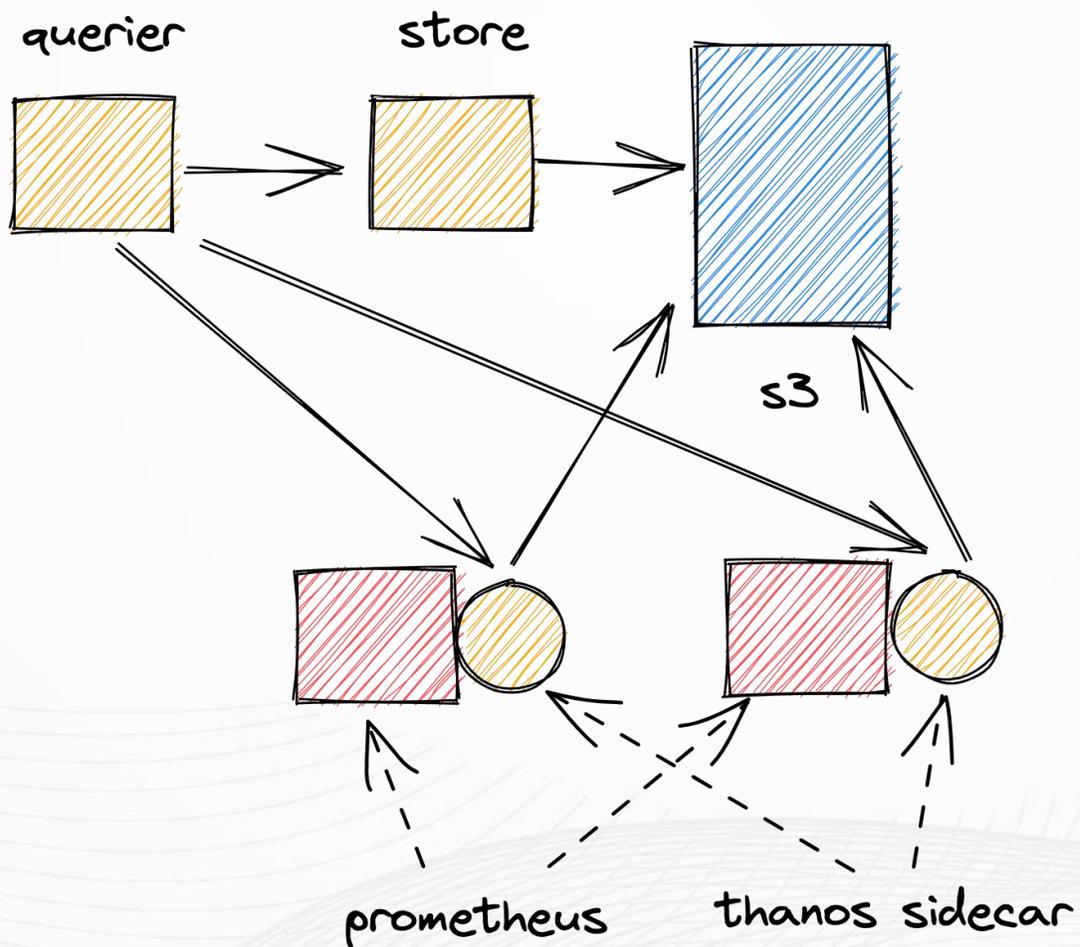
总结:

- 提供鉴权和访问控制
 - 提供整合的全局视图
 - 高可用
-
- 架构较复杂
 - 要调整Prometheus配置

示意图



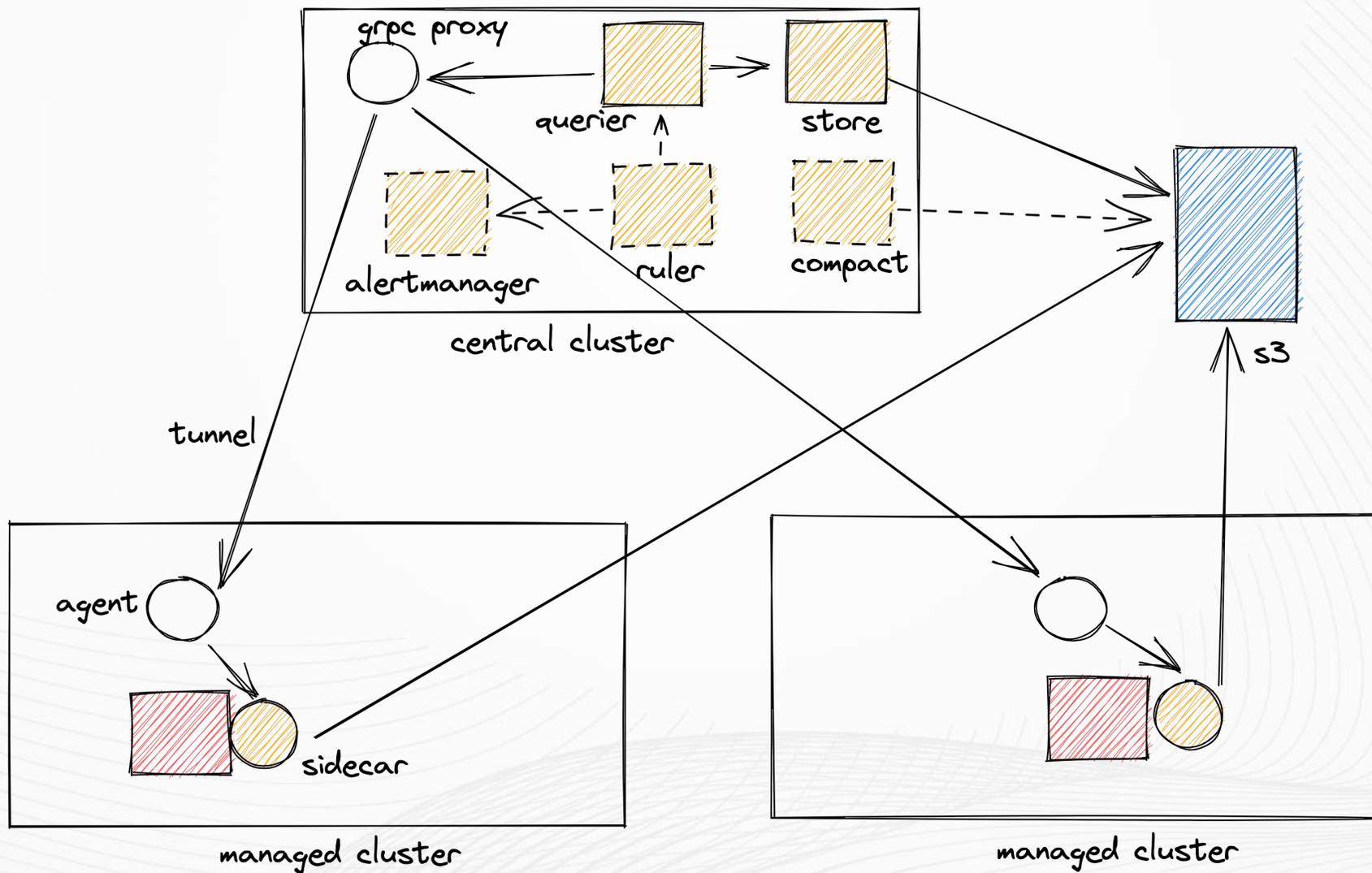
示意图



总结:

- 架构简单
 - 无侵入配置
 - 提供整合的全局视图
 - 高可用
-
- 没有鉴权和访问控制

特性	Cortex	Thanos
多租户	✓	
可用性	✓	✓
长期持久存储	✓	✓
压测性能		✓
架构复杂度		✓
社区活跃度		✓



1. 监控系统本身的稳定性
 - 配置监控系统本身指标相关的告警
 - 失能开关
2. Prometheus性能主要受样本量影响，三百万以上规模考虑分片
3. 容量规划
 - 大规模场景Prometheus内存占用较大
 - 指标存储用量： $\sim \text{size_per_sample} * \text{scrape_rate} * \text{series_number}$
 - 远端存储轮换
4. 减少不必要的指标采集，避免告警泛滥

GOTC

THANKS

全球开源技术峰会

THE GLOBAL OPENSOURCE TECHNOLOGY CONFERENCE