www.emqx.io

EMQ X Kuiper

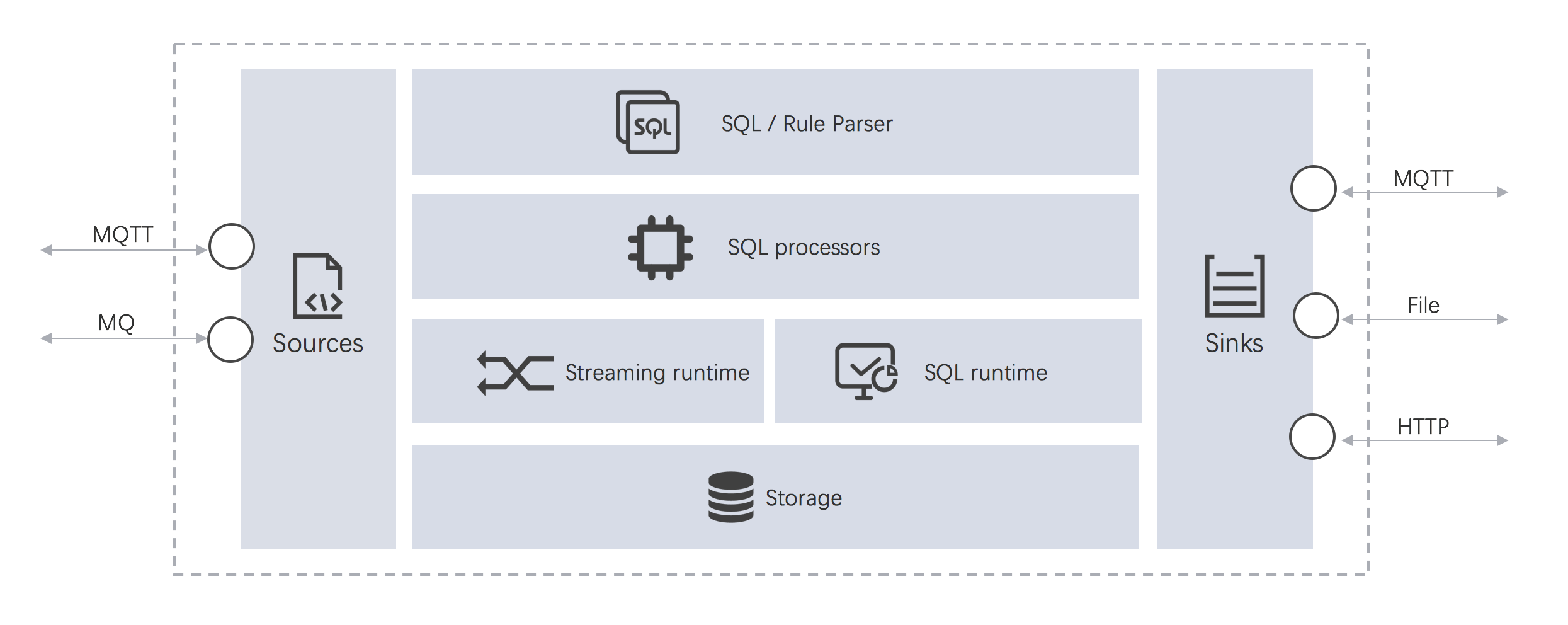
**白 皮 书**

**产品概述**

2019 年是 5G 商用元年，随着 5G 部署的深入，边缘计算日渐流行，应用范围越来越广。边缘计算通过在靠近接入网的机房增加计算能力，可以，

* 大幅降低业务时延
* 减少对网络的带宽压力，降低传输成本
* 减少在云端的数据存储成本
* 提高内容分发效率提升用户体验

物联网边缘计算很大一部分指的是流数据的处理，所谓流数据是一组顺序、大量、快速、连续到达的数据序列，一般情况下，流数据可被视为一个随时间延续而无限增长的动态数据集合。流式数据处理可以帮助用户实时了解系统设备的状态，并对异常情况作出快速响应。边缘端的计算资源（CPU，内存等）不像在云端丰富，传统的流式数据处理框架类似于 Apache Storm 或者 Apache Flink 等，由于其安装包过大，或者部署结构与过程过于复杂，运行时的高消耗等原因，并不适合于在这些资源受限的边缘设备（工控机、网关，或者配置不高的 X86 服务器等设备）上运行。



Kuiper 架构

EMQ X Kuiper（<https://github.com/emqx/kuiper>） 是使用 Golang 实现的轻量级物联网边缘分析、流式处理开源软件，可以运行在各类资源受限的边缘设备上。Kuiper 设计的主要目标之一就是将在云端运行的实时流式计算框架（比如 Apache Spark，Apache Storm 和 Apache Flink 等）迁移到边缘端。Kuiper 参考了上述云端流式处理项目的架构与实现，结合边缘流式数据处理的特点，采用了编写基于源 (Source)，SQL (业务逻辑处理), 目标 (Sink) 的规则引擎来实现边缘端的流式数据处理。

**产品特点**

* 完全开源：

基于 Apache 2.0 开源协议，对商业化非常友好。

* 轻量级：

安装包约 10MB左右；不依赖于第三方的库与运行时，下载解压后即可运行，安装和使用非常方便

* 跨操作系统支持：

目前可以运行在各类 Linux 发行版和 Mac 上，后续版本已规划会支持 Windows 操作系统

* 高性能：

在树莓派 3B+ 上，达到12k/秒的消息吞吐量；AWS t2.micro( 1 Core \* 1 GB)上，单实例支持多达 8000 条规则的并发运行；

* 使用简单，快速应对敏捷业务变化：

用户可以通过写 SQL 来实现业务处理；如果业务有变化，只需更新相关的 SQL 语句即可，而无需进行复杂的应用开发和部署

* 内置支持 MQTT 消息的处理：

作为物联网领域的事实协议标准，Kuiper 内置了对接 EMQ X 等 MQTT 消息服务器，实现对来自于物联网消息处理的无缝对接；另外还支持接入 EdgeX 消息总线，以及从 HTTP 服务器定时获取数据进行分析等支持。

* 可扩展性：

定义了 Sources、Functions和Sinks的扩展接口，实现自定义的数据分析和处理，快速实现自定义的业务逻辑处理

* 可管理性：

Kuiper 定义了管理接口和控制台界面，用户可以通过 REST API 或者管理界面来实现对插件、业务规则的定义和管理等操作

* 易集成性

Kuiper 非常轻量、易于使用，非常容易和别的边缘计算框架进行快速集成。Kuiper已经与开源社区别的著名的开源计算框架进行深入合作，包括在 2020 年 5月份发布的 Linux 基金会旗下的 EdgeX Geneva 版本中， Kuiper 作为核心的规则引擎服务，正式集成到 EdgeX 项目中；另外，Kuiper 与华为KubeEdge 项目合作，实现云边协同的边缘计算。

**支持平台与安装环境**

* 硬件支持：

工控机（支持各类 x86、ARM 架构）

x86 服务器

树莓派

* 安装包大小： 10MB
* 最小运行时环境：

CPU：1GHz

内存：64MB

磁盘：1GB

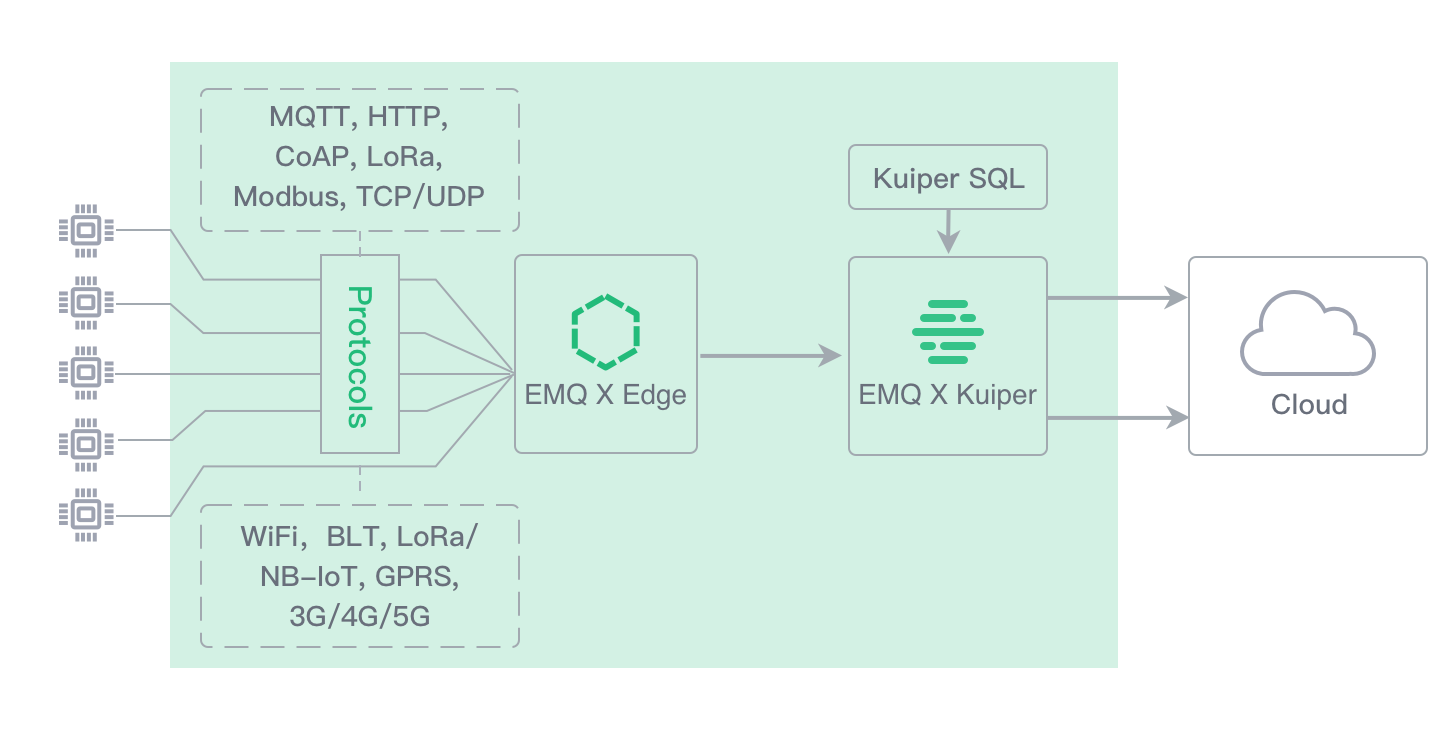
* 操作系统支持：

CentOS、Alpine、Raspbian、Debian、Ubuntu以及Docker镜像安装。

**产品功能**

1. 与 EMQ X Edge 的内置集成能力：

Kuiper 通过 “MQTT 源”可以直接和 EMQ X Edge 进行无缝集成。对通过 EMQ X Edge 接入的数据进行实时分析后，通过 EMQ X Kuiper 提供的 Sink 组件将数据发送到云端。



1. 管理接口：

提供了命令行和REST API 管理接口，用户可以实现流的创建、删除，业务规则的创建、删除、启动、停止和重启等操作；通过命令行工具可以测试 SQL 语句。

1. 完整的SQL分析能力：

提供了基于 SQL 的完整分析能力，包括 SELECT、FROM、JOIN（LEFT、RIGHT、FULL与CROSS连接）、WHERE、ORDER、HAVING和GROUP关键字，实现对于大部分数据的分析能力

1. 丰富的函数支持：

内置提供了 60+ 函数，快速实现对数据的处理：

* 1. 聚合函数：平均、计数、最大、最小、求和
  2. 数学运算函数：绝对值、三角函数等
  3. 字符处理函数：字符连接、长度、子字符串判断等
  4. 转化函数：cast、encode 等
  5. 哈希函数：md5、各类sha函数

1. 窗口：

内置的时间窗口，可以实现对数据的时段分析：

* 1. Tumbling 时间窗口：使用固定的时间长度对流式数据进行分割。每个窗口不会重叠，并且每个数据点只能属于一个时间窗口。
  2. Hopping 时间窗口：以固定的时间段往前跳跃的窗口。该类窗口可以认为是可以重叠的 Tumbling 窗口。
  3. Siding 窗口：由事件触发的时间窗口，每个窗口至少包含一个数据点。
  4. Session 窗口：将在差不多时间产生的数据点进行分组，过滤掉那些不产生任何数据的时间段。

计数窗口：可以实现对实现某特定事件的计数，达到某条件后触发该事件

1. 扩展能力：

更加灵活的可扩展能力，支持不同应用场景下的数据处理。

* 1. 源扩展（source）：扩展支持接入来自于不同的源的数据
  2. 目标（sinks）：扩展支持向不同目标写结果数据
  3. 函数（用户自定义函数）：扩展支持不同类型的函数

**应用场景**

随着 5G 技术的商用落地，其增强移动带宽、海量物联、高可靠超低延时通信等特性将进一步提升物联网技术应用价值， 以物联网为代表的新一代信息技术成为构建工业基础性行业竞争优势的主要推动力。

**车间的工业网关**

实时分析来自于生产线的高频数据，并将分析结果推送至云端。车间本地可视化系统，或者远程监控系统可以查看生产线的实时状态。使工厂拥有更高的响应度，更具前瞻性和预测性，并助企业免受运营停工及其他生产难题的困扰，并且可以使客户交付时间缩短，整体成本下降，以及产能提升，残次品减少。某客户利用 Kuiper 扩展了自定义函数，实现对来自于车间生产线上的数据进行傅立叶分析，实时、快速地实现对生产线的监控和分析。

**车联网车机**

实时分析汽车总线的大量数据，经过分析将有价值的数据推送至云端或者本地存储，车载系统或者用户的手机应用可以实时查看汽车的状态。更好地支持车联网广泛的地理分布，移动性，位置感知，低延迟，支持实时交互等需求，以满足日益增长的车联网业务发展。

**智能家居网关**

通过实时分析家居采集的各类数据，将重要的结果通过本地显示设备，或者通过云端发送给用户的手机应用，实现对家庭设备的即时状态管理与控制，保证设备有条不紊、综合协调、科学地运行，从而达到有效地保证建筑物内有舒适的工作环境、实现节能、节省维护管理工作量和运行费用的目的。某方案通过在家庭网关中部署了 Kuiper，对于经过户外的人进行本地统计和分析，并将分析结果推送到用户的手机进行实时展示。

**联系我们**

公司: 杭州映云科技有限公司

官网: [https://www.emqx.io](http://www.emqx.io)

电话: 400-696-5502

博客: [https://www.emqx.io/b](https://www.jianshu.com/u/9cbcdf094d33)log

邮箱: [contact@emqx.io](mailto:contact@emqx.io)

微博: [https://weibo.com/emqtt](http://weibo.com/emqtt)

Twitter: @emqtt

微信:



[版权申明]

©2013-2020 杭州映云科技有限公司版权所有。