

 PULSAR ×  DORIS

Pulsar x Doris

实时分析+智能管控

打通“流”到“析”全链路

TIME 2025.08.14 19:30-21:00



ASCENTSTREAM

futeng@ascentstream.com



ASP 社区版

帮助

Apache Pulsar x Apache Doris
构建实时数仓新架构

CONTENTS

- 云原生消息队列 Apache Pulsar
- Pulsar x Doris 构建实时数仓新架构
- 谙流 ASP 社区版
- 部署和演示
- Q&A





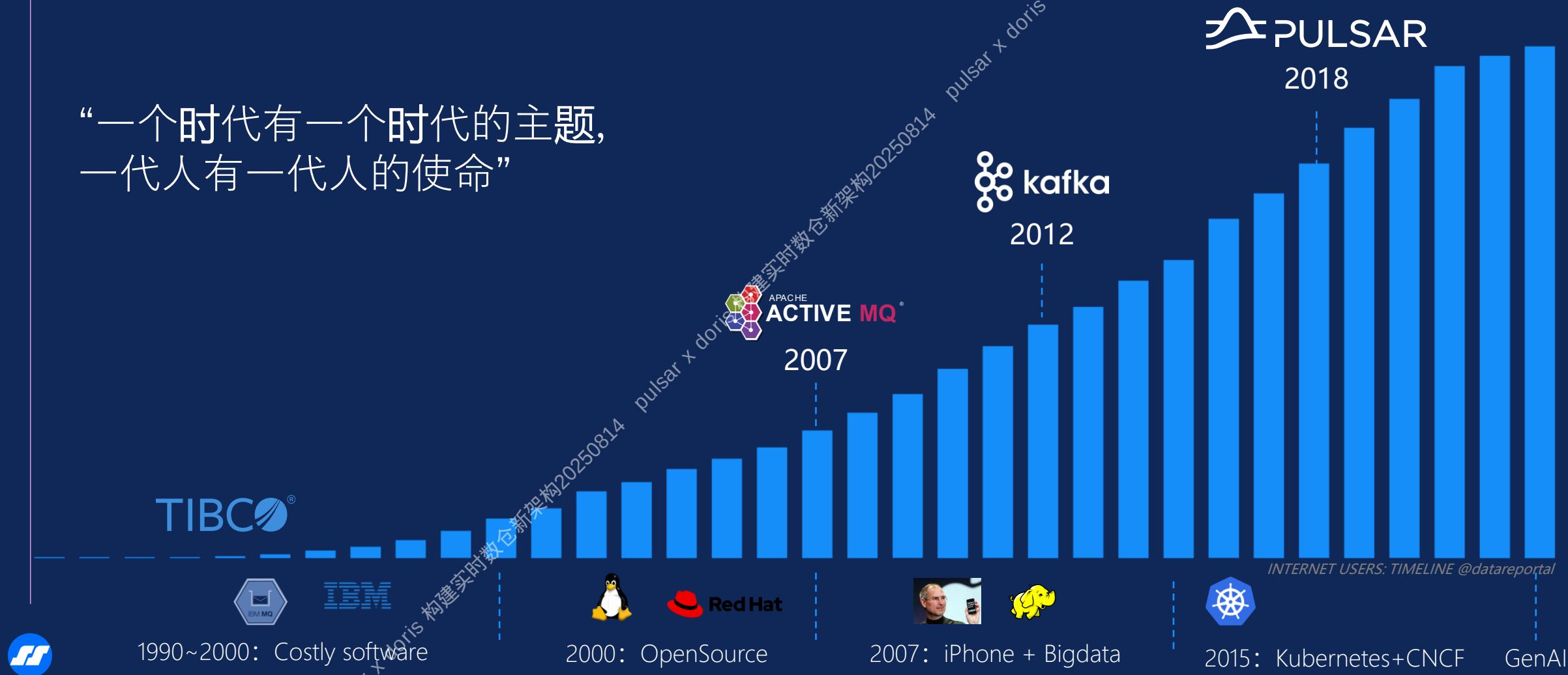
云原生消息流平台 Apache Pulsar

核心基础设施：消息中间件（队列 + 流）



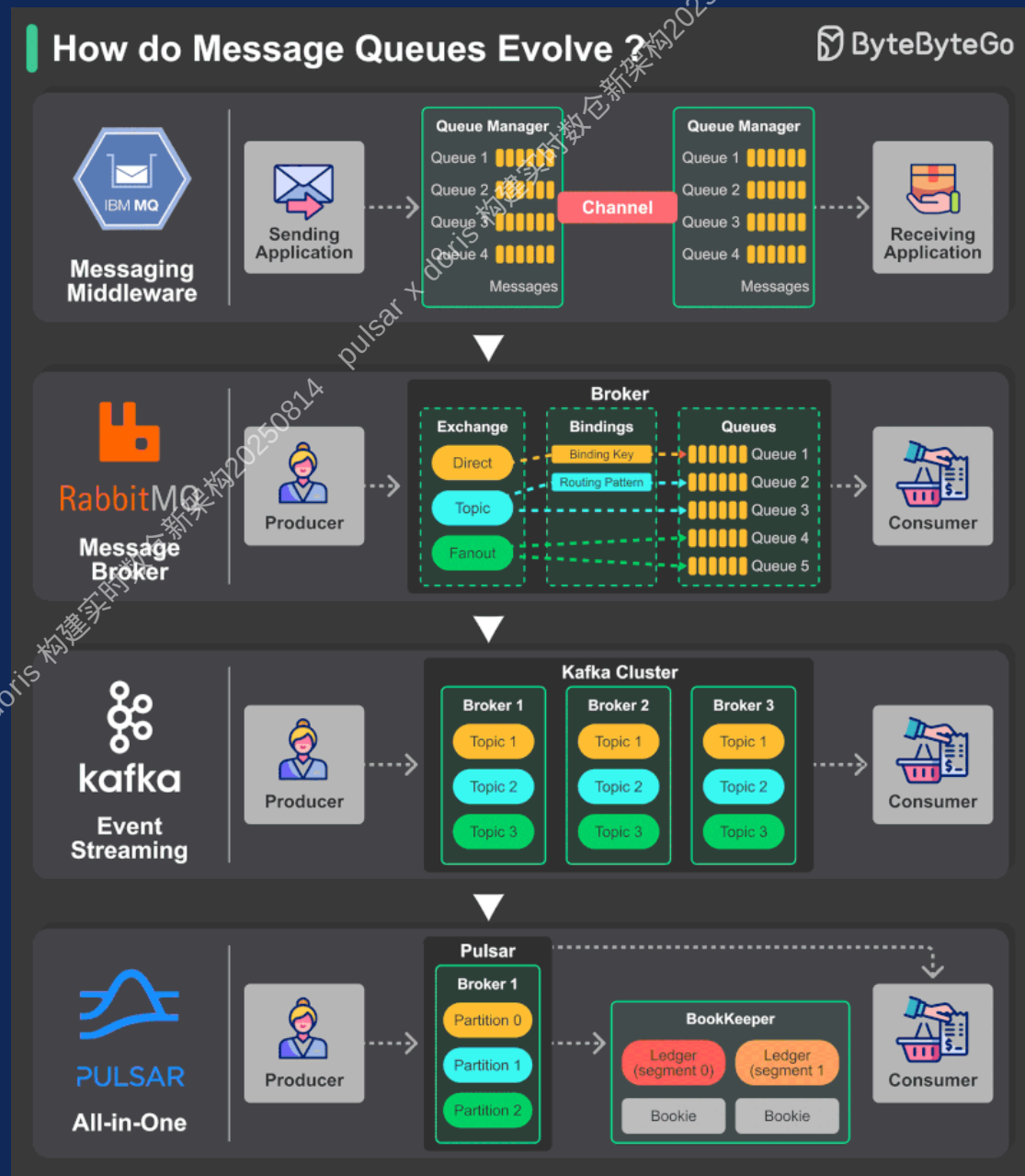
消息中间件发展历程 (APACHE TLP TIME)

“一个时代有一个时代的主题，
一代人有一代人的使命”



消息中间件 演进方向

存算分离、
走向消息和流的统一



APACHE PULSAR



下一代的云原生、高一致、高性能的消息流平台

<https://github.com/apache/pulsar>



PULSAR 的核心优势



统一 Queuing 和 Streaming 模型



多租户，企业级准备



云原生，弹性扩缩容



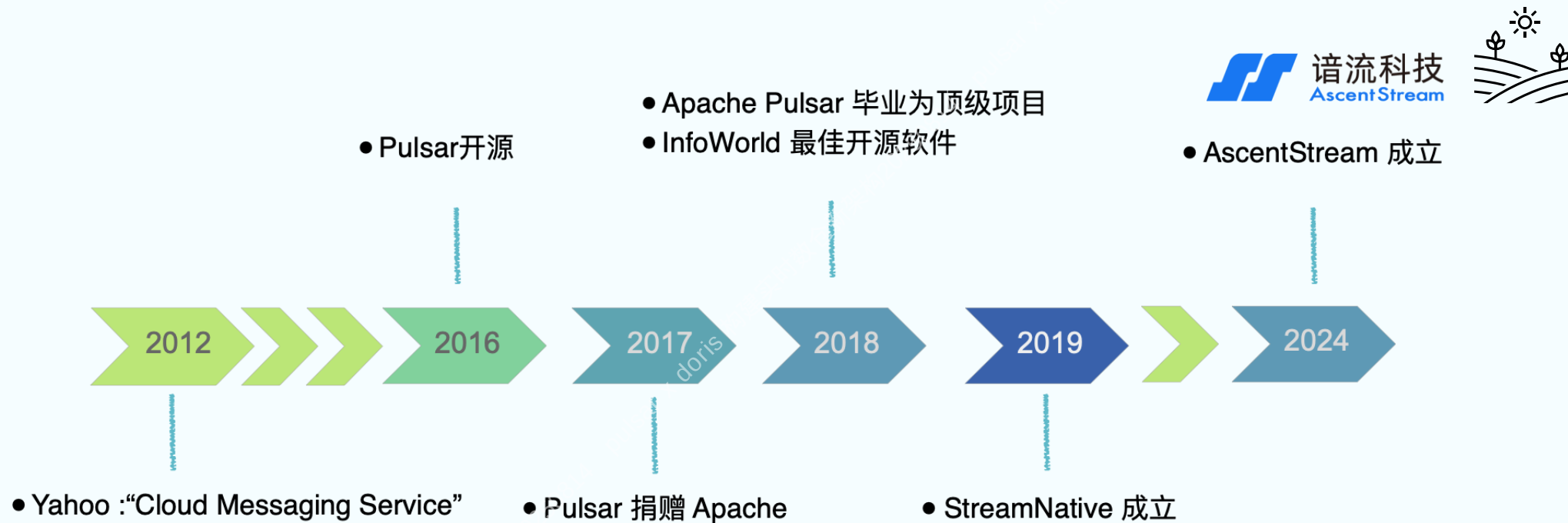
原生跨地域复制



冷热分离，层级存储



PULSAR 12 年：从创新、到稳定



活跃的开源生态

Top5
Commits



Projects by Number of Commits, 2021

Tomcat

1.4%

CloudStack

1.4%

Beam

1.5%

Shardingsphere

1.6%

NuttX

1.7%

Isis

1.9%

IOTDB

2.1%

Spark

2.2%

Pulsar

2.2%

Airflow

2.7%

Flink

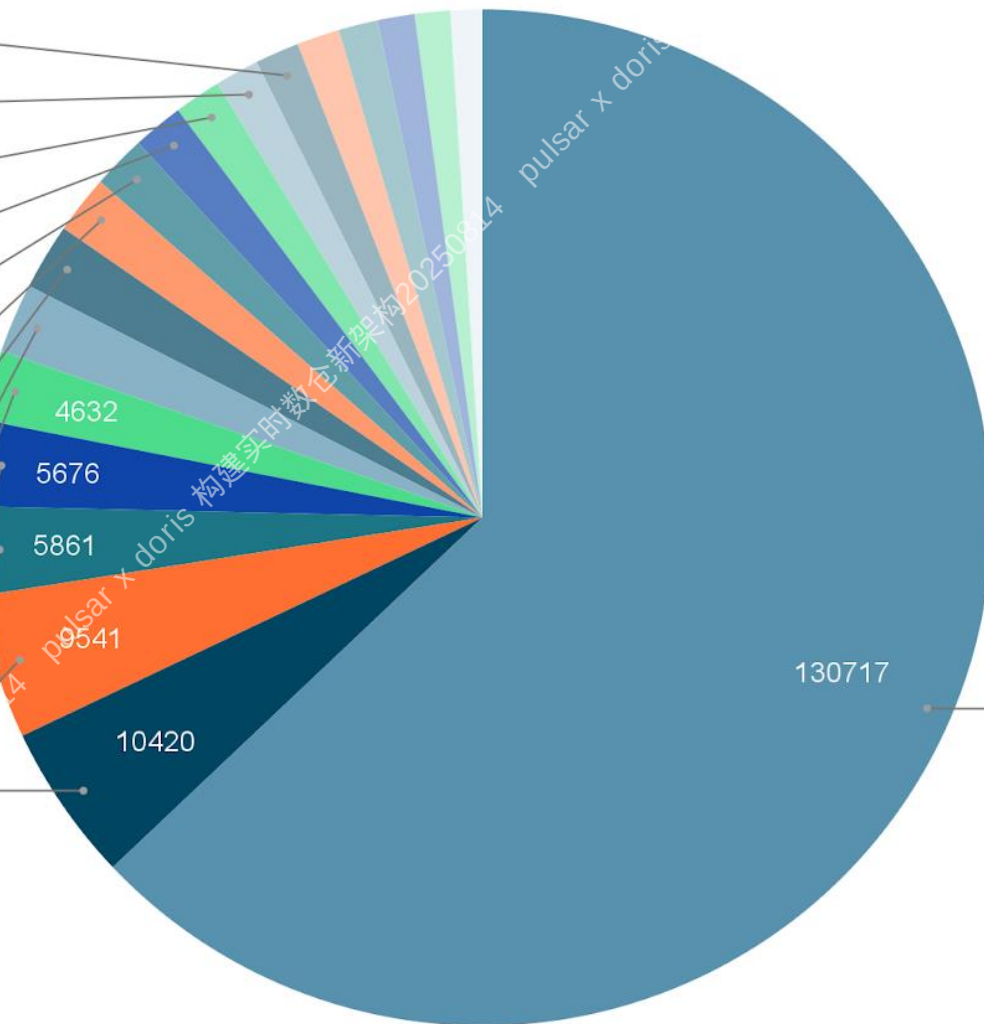
2.8%

Camel

4.6%

Superset

5.0%



INFOWORLD 两年最佳开源项目

The screenshot displays the Apache Pulsar website. At the top, the Pulsar logo is accompanied by the text "Apache Pulsar is an open-source distributed pub-sub messaging platform originally created at Yahoo and now part of the Apache Software Foundation". Below this, a diagram illustrates the Pulsar architecture, showing Producers (1 and 2) sending data to a central "Pulsar topic", which then routes it to various Subscriptions (A, B, and C) and finally to Consumers (A, B-1, B-2, C-1, and C-2). The website also features sections for "Pulsar Functions", "Proven in production", and "Horizontal scaling". At the bottom, the "InfoWorld BOSSIE 2018 AWARDS" banner is prominently displayed.

The screenshot shows the StreamNative website. The header includes the StreamNative logo and navigation links. The main heading reads "The unified messaging and streaming platform made by the creators of Apache Pulsar." Below this, a sub-header states "Built for Kubernetes. Made for the cloud. Enables multi-cloud and hybrid cloud architectures." A "Take A Tour" button is visible. The website also features a "StreamNative, Powered by Apache Pulsar" section with three sub-sections: "StreamNative Cloud", "StreamNative Platform", and "StreamNative Pro Services". On the right side, a detailed diagram of the StreamNative architecture is shown, including components like Pulsar IO, Pulsar Clients, Kafka Clients, Function Mesh, StreamNative Console, SLA Detector, Gradus, Prometheus, and AlertManager. At the bottom, the "InfoWorld BOSSIE 2021 AWARDS" banner is displayed.



APACHE PULSAR 周边生态

Protocol Handlers



Pulsar Functions (Lightweight Stream Processing)



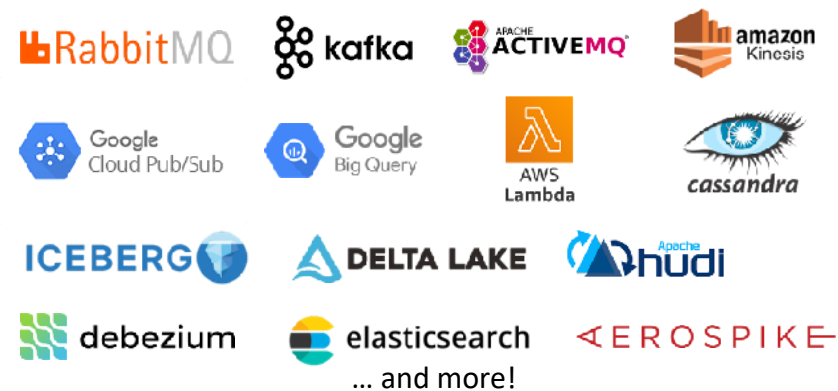
Processing Engines



Data Offloaders (Tiered Storage)



Connectors (Sources & Sinks)



Client Libraries



APACHE PULSAR 大量国际头部用户使用 在核心场景



APACHE PULSAR: 深度使用于中国头部企业核心业务中

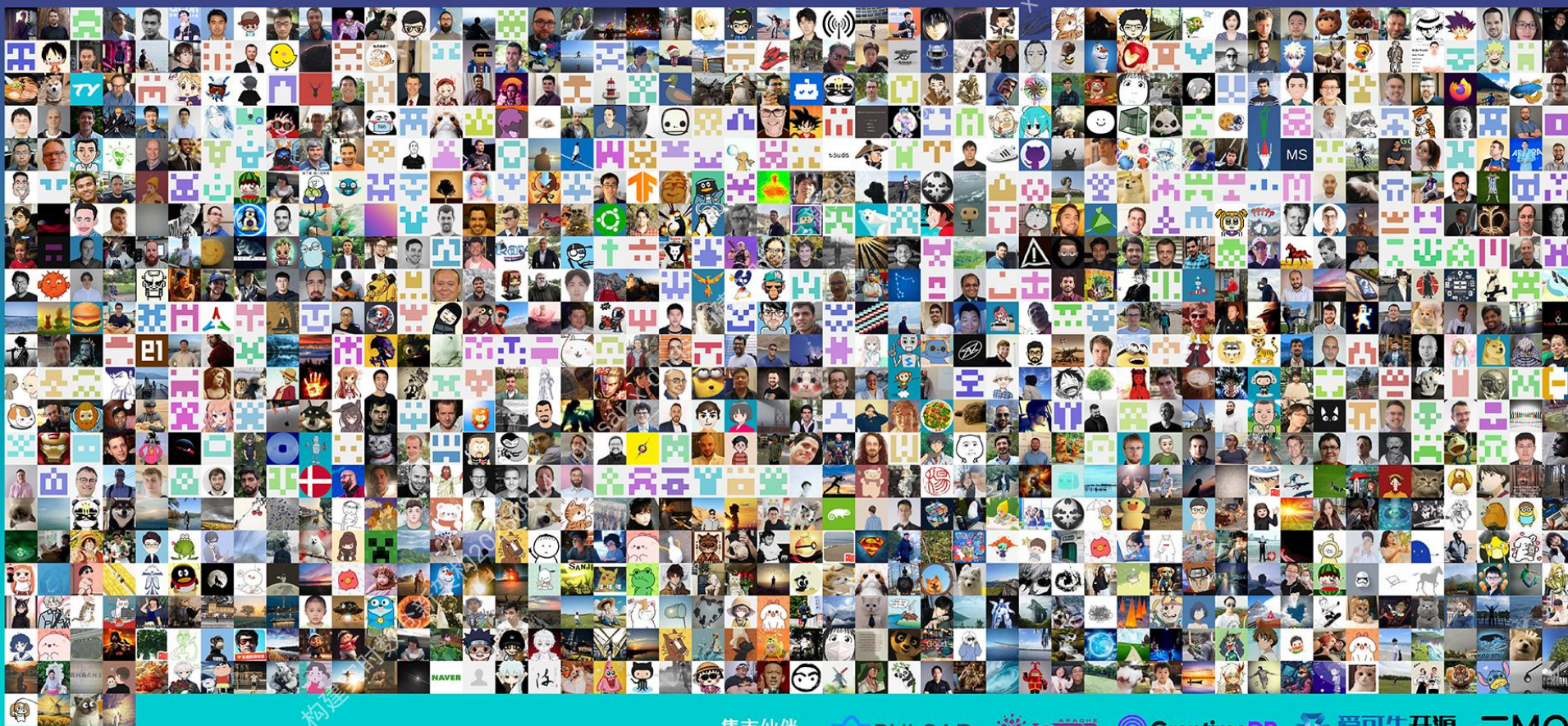




Pulsar Developer Day 2024

Pulsar 开发者大会

“允许让用 Pulsar 的人先富起来！”



集市伙伴





PULSAR X DORIS 构建实时数仓新架构

范式新升级：两者结合，实现了从数据采集、处理到分析的全链路实时化，为企业数字化转型提供了强有力的数据支撑

时代变迁：流数据处理的挑战与机遇

“90% 应用已完成迁移！LinkedIn 宣布砍掉 Kafka，全面切换自研消息系统 Northguard！”

> Introducing Northguard and Xinfra: scalable log storage at LinkedIn



数据规模指数增长

TB → PB

对传统分布式架构提出新的挑战



云原生技术普及

弹性 与 **成本** 成为关键

云原生架构要求系统具备自动扩缩容能力



实时性要求提高

秒级 甚至 **毫秒级** 成为标配

金融交易、物联网、实时分析等领域对延迟极度敏感



多租户场景增多

资源隔离 与 **共享** 需求并存

企业内多业务线需共享基础设施和部分核心数据，同时保证安全隔离



需求变更，KAFKA 跟不上了！

🔄 今时不同往日

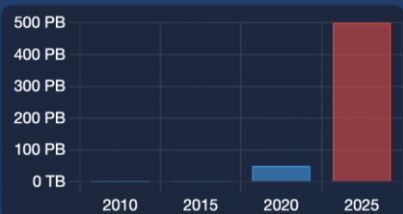


性能与成本 “既要又要”

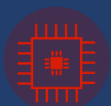
- 弹性扩缩容、秒级扩展、按需付费
- 高可用、跨区域、跨云部署
- 自动化运维、快速负载均衡
- 冷热数据隔离存储，按需提取



数据规模爆炸



从 GB 到 PB,
从简单传输到
海量 Topic



硬件升级

- SSD 普及，成本下降，容量提升
- 网络提升，从 1Gbps 到 100Gbps
- 计算资源池化，按需分配
- 对象存储价格实惠自带多副本



旧瓶难装新酒



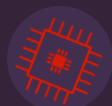
性能上不去，成本下不来

- 单体结构，扩容困难
- 云原生不兼容，高可用魔改
- 重平衡噩梦难以快速均衡
- 存储利用率低，也难以冷热分离



拆小集群，艰难应对数据量的暴增

- 分区数受限
- 为避免资源过载，要部署大量小集群
- 顺序写优势再高 IOPS 下退化



硬件升级的红利没享受到？

- SSD 性能无法被充分释放
- 一次分区迁移，网络流量剧增
- 存储与计算资源无法独立扩展
- 计算和存储资源不匹配，无法使用新技术



KAFKA单体架构硬伤之三座大山

分区绑定引发的I/O瓶颈

- 高并发时顺序写退化为随机写
- 触发PageCache污染，最终导致生产/消费延迟陡增



100%

磁盘利用率

扩容触发级联灾难

- 网络带宽打满
- 目标节点磁盘IO过载
- 分区级服务中断



3 X

问题叠加

存储耦合限制弹性

- 为扩容计算被迫扩容存储
- 磁盘整体利用率不高
- 运维碎片化，成本激增



<60%

磁盘利用率

“当Apache Kafka运维深陷扩容即崩溃的故障滚雪球的死亡循环：一次手动迁移触发全集群雪崩，凌晨三点的告警成为工程师的生理钟。”



时代变迁：推荐基于APACHE PULSAR 构建新一代实时数仓数据接入层



突破 IO 瓶颈

消除扩容灾难、利用最新 SSD 特性、极高吞吐、低延迟



弹性扩缩容能力

秒级扩容、业务无感知、高扩展性、高容灾能力



存算分离架构

分离架构，快速伸缩，资源最大化利用

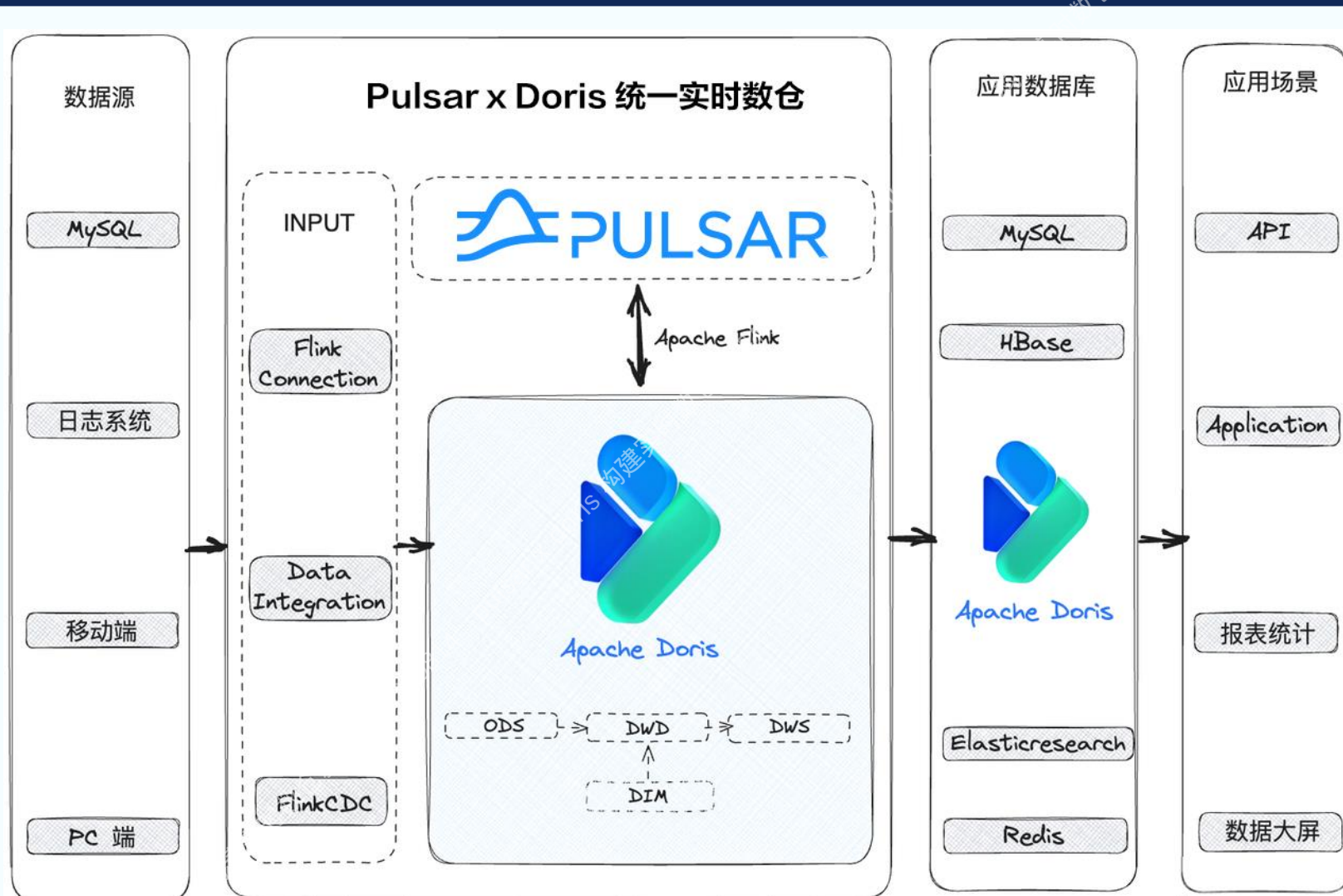


云原生特性

多租户治理、智能负载均衡、冷热数据分层进一步降低成本



Apache Pulsar x Apache Doris 实时数仓强强联合



Pulsar 与 Doris 结合的核心优势



毫秒级响应

- ✓ Pulsar 保证毫秒级延迟地将消息送达
- ✓ Doris 则能在数据写入后可立即对外提供查询服务
- ✓ 从数据产生到分析结果能达到“近实时”反馈



弹性伸缩

- ✓ Pulsar 存算分离可快速扩缩容使用流量波动
- ✓ Doris 的亦可快速横向扩容来提升并发查询能力
- 整体云原生弹性能力保证系统面对突发负载时依然稳定可靠



企业级准备

- ✓ Pulsar 的多租户、数据隔离、数据冷热分离等功能确保不同业务的数据流有序隔离
- ✓ Doris 拥有完善的权限管理和元数据管理机制，支持与 Iceberg/Hudi 等湖表联动，形成规范的数据湖仓架构



应用新场景

- ✓ Pulsar 消息流合并特性，可接入实时接收点击流、日志等事件，涵盖企业在线侧业务和离线分析侧业务等
- ✓ Doris 通过物化视图等丰富特性，涵盖实时数仓各种复杂数据查询场景
- ✓ Pulsar+Doris 架构能同时满足高吞吐的事件采集和低延迟的复杂查询需求，为企业创新解锁各类新场景





谙流 ASP 社区版

Apache Pulsar 现代化管理平台

Welcome to learn about ASP CE (AscentStream Community Edition)

谙流科技 (ASCENTSTREAM)

中国云原生消息流服务领导者

- 致力于打造**金融级云原生**消息流平台，助力企业挖掘实时数据价值。
- 由 Apache Pulsar 与 Apache BookKeeper **创始团队**成员组建。
- 专注中国商业化和社区，满足**信创要求**。



www.ascentstream.com



谱流科技核心产品



ASK(AscentStream for kafka)

基于 Apache Kafka 深度优化的新一代云原生国产化流数据平台。兼容生态，解决痛点。

- ✓ 100% 协议兼容
- ✓ 国产化全栈自主可控
- ✓ Kafka国产替代 可平滑迁移
- ✓ 专注私有化场景和行业赋能
- ✓ 存算分离 高扩展 高可操作性



ASP(AscentStream Platform)

基于 Apache Pulsar 打造的一站式企业级、金融级云原生消息流平台，满足高容灾、高性能、高可靠和高安全性要求。

- ✓ 企业级 Pulsar 核心发版
- ✓ 一站式安装、管控、完善的企业级功能
- ✓ Pulsar Operator，跨 K8S 容灾
- ✓ 消息追踪、多集群切换、全生态支持
- ✓ 商业公司支持，保障核心业务稳定可靠

- ASP 社区版：永久免费使用的 Pulsar 管控平台，替代 PulsarManager，长期维护支持。
- ASP 企业版：全功能商业级一站式解决方案，适合中大型企业。



ASP 社区版概述

○ 产品定义

- ✓ 谙流科技贡献给 Pulsar 生态的现代化管理平台
- ✓ 旨在降低 Pulsar 使用门槛和普及基础管理工具两大问题

○ 核心价值

- ✓ 降低使用门槛：帮助用户从 0 到 1 快速将 Pulsar 用起来
- ✓ 现代化管理工具：替代已不再积极维护的 Pulsar Manager
- ✓ 开放社区协作：持续收集用户反馈，定期发布更新

ASP社区版是 Pulsar Manager 的全面升级替代方案

CE 核心组件



MySQL

数据库，存储 ASP 配置信息、监控聚合指标等



Console

核心管控服务，后台 REST，前台提供 WEB 界面



Pulsar

样例 Pulsar 集群，建议初学者安装使用



Ping

样例程序，提供生产者、消费者代码样例和程序



ASK CE：核心功能和优势



核心功能支持

资源管控，安全无忧

- ✓ **集群管理**：支持示例和接管自由集群
- ✓ **资源管理**：便捷管理 Tenant、Namespace、Topic、Subscription、Produce、Consume 等资源
- ✓ **安全管理**：支持认证和授权机制
- ✓ **监控告警**：提供关键指标监控



让 Pulsar 开箱即用

优化体验，降低门槛

- ✓ **一键部署**：快速启动完整的开发环境
- ✓ **直观界面**：清晰易用的 web 控制台
- ✓ **统一视图**：集中管理各类核心资源
- ✓ **可视化监控**：实时监控集群性能



开放的社区协作

持续经营、永久免费

- ✓ **用户反馈**：持续收集和迭代社区需求
- ✓ **定期更新**：提供稳定可靠的版本迭代，提供完善的文档和最佳实践
- ✓ **社区参与**：后续欢迎开发者贡献代码
- ✓ **技术交流**：提供 Pulsar 专业问题解答

降门槛、提效率、可观测、保安全



ASP CE VS. PULSAR MANAGER

Feature	ASP CE	Pulsar Manager
维护状态	全职团队维护，定期更新，快速响应用户反馈	基本停止维护（2022年后几乎无更新），近200个未解决的问题
Pulsar兼容性	支持 Pulsar 2.10.x 及最新 3.0.x LTS 和4.0.x LTS 版本	最新 LTS 版本支持较弱
用户界面	现代化设计，响应式布局，直观易用	传统设计，界面简单但不直观
部署方式	一键部署，全自动化配置，简单快捷	支持 Docker 部署，但配置复杂
集群管理	支持示例集群和自有集群接管，稳定可靠	接管集群容易失败，多个用户报告无法显示集群
监控能力	全面的集群健康状态监控，可视化图表，支持实时生产消费曲线监控	基础的指标监控，有限的可视化能力，不支持生产消费曲线

ASP社区版承诺永久免费使用，作为谕流科技对 Apache Pulsar 社区的回馈，将持续提供稳定的维护和更新



快速开始：一键部署

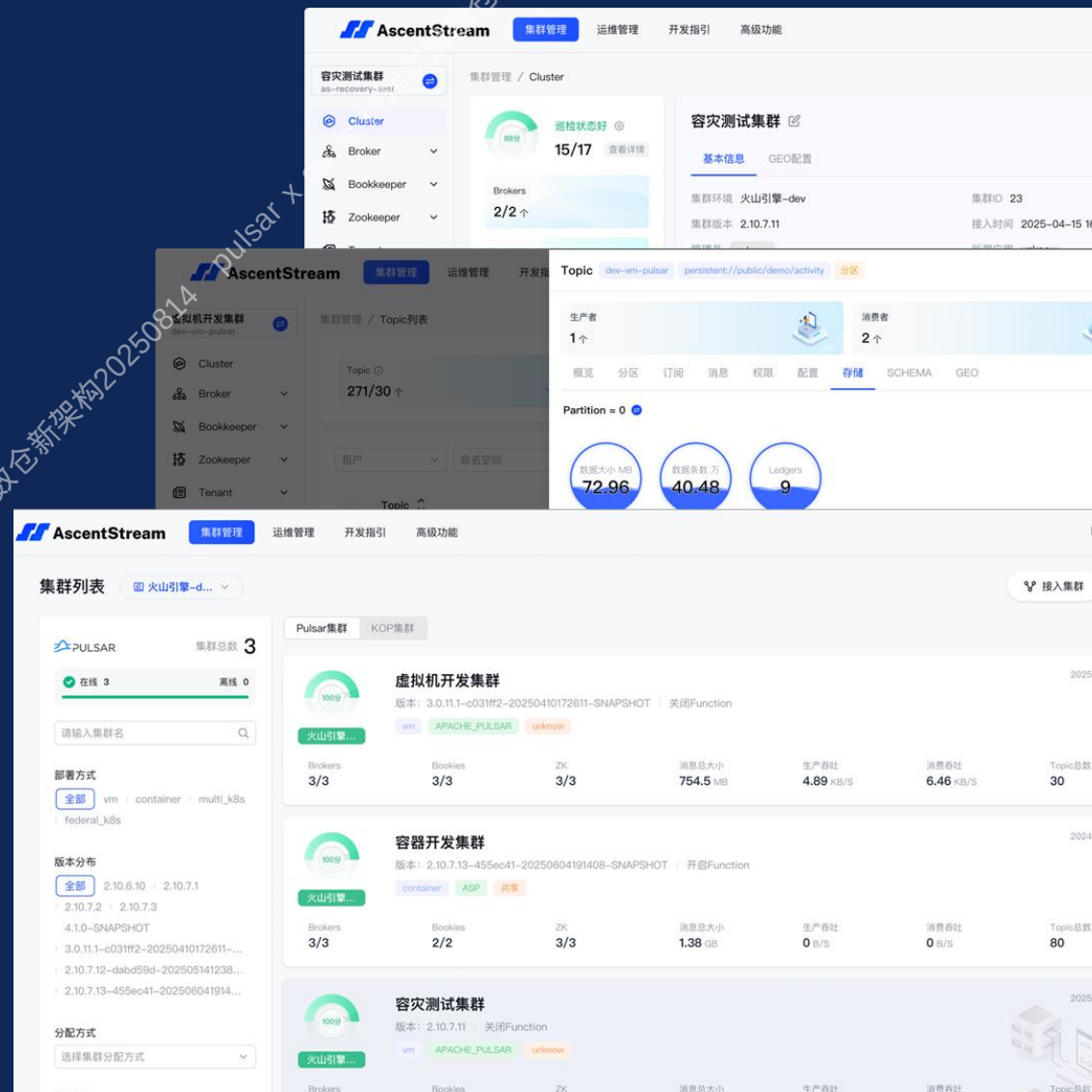
ASP CE

```
curl -s https://get.ascentstream.com/ce | bash
```



管控登录

HTTP://IP:8083



功能亮点：策略配置

AscentStream
Platform Community

集群管理 开发指引 系统管理

dev-as-platform-asp
dev-as-platform-asp

Cluster
Tenant
Namespace
Topic
订阅
生产者
消费者
权限

集群管理 / Topic

Topic 8/6 个

租户 命名空间

Topic persistent://public/default/fujun persistent://public/default/test

共 2 项数据

Topic dev-as-platform-asp persistent://public/default/fujun 分区

生产者 0 个

消费者 0 个

订阅 0/0 个

概览 分区 配置 存储 SCHEMA

消息保留策略 修改

保留时长 10080 (分) 保留大小 永久保留
过期时长 10080 (分) 过期大小 禁用

限流策略 编辑

生产限流 吞吐 -1 (Byte/s) 速率 -1 (条/s) 统计间隔 1 (秒)

消费限流 吞吐 -1 (Byte/s) 速率 -1 (条/s) 统计间隔 1 (秒)

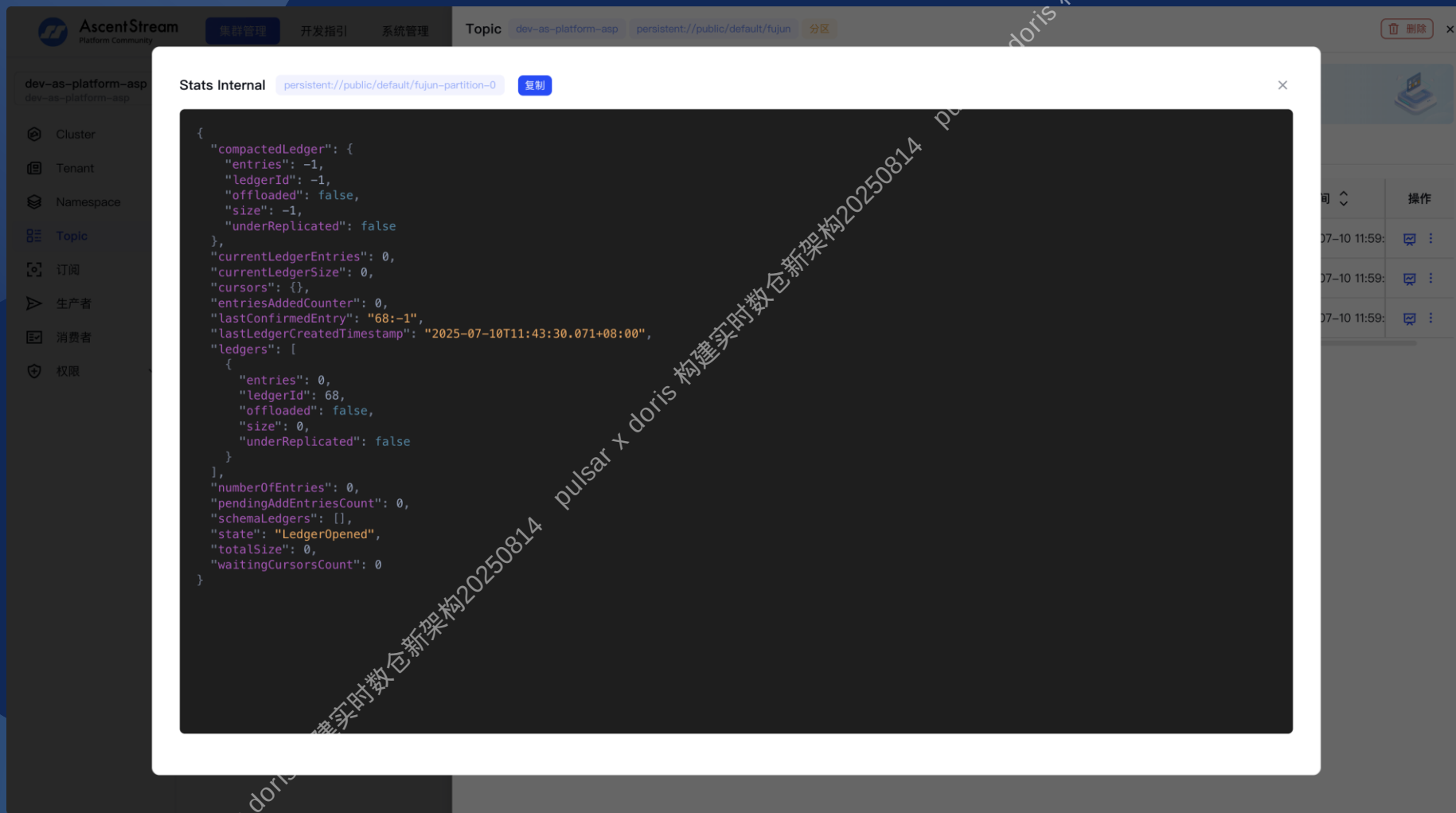
订阅速率 速率 0 (个/s) 统计间隔 30 (秒)

Schema 兼容策略 编辑

兼容策略 FULL

AscentStream: Stream Data for AI

功能亮点：细粒度状态统计



The screenshot displays the AscentStream web interface. A modal window titled 'Stats Internal' is open, showing the internal state of a topic partition. The background interface includes a sidebar with navigation options like Cluster, Tenant, Namespace, Topic, 订阅 (Subscription), 生产者 (Producer), 消费者 (Consumer), and 权限 (Permissions). The main area shows the topic 'dev-as-platform-asp' and its partitions. The 'Stats Internal' modal displays the following JSON data:

```
{
  "compactedLedger": {
    "entries": -1,
    "ledgerId": -1,
    "offloaded": false,
    "size": -1,
    "underReplicated": false
  },
  "currentLedgerEntries": 0,
  "currentLedgerSize": 0,
  "cursors": {},
  "entriesAddedCounter": 0,
  "lastConfirmedEntry": "68:-1",
  "lastLedgerCreatedTimestamp": "2025-07-10T11:43:30.071+08:00",
  "ledgers": [
    {
      "entries": 0,
      "ledgerId": 68,
      "offloaded": false,
      "size": 0,
      "underReplicated": false
    }
  ],
  "numberOfEntries": 0,
  "pendingAddEntriesCount": 0,
  "schemaLedgers": [],
  "state": "LedgerOpened",
  "totalSize": 0,
  "waitingCursorsCount": 0
}
```



功能亮点：活跃社区



部署和操作演示

一键部署



接管集群



基础操作



为什么现在应该切换到 ASPCE

Pulsar Manager 在当时无疑是一个非常有用的工具，但随着时间推移和 Pulsar 的快速发展，它已经无法满足现代 Pulsar 用户的需求。GitHub 上近 200 个未解决的问题清晰地说明了这一点。

ASP 社区版作为 Pulsar Manager 的现代化替代品，提供了以下显著优势：



全面兼容

- ✓ 支持 2.10.x
- ✓ 快速跟进 LTS 版本
- ✓ 支持虚拟化和容器化部署集群



基础管控

- ✓ 持续演进的认证和授权支持
- ✓ 基础核心的监控指标
- ✓ 多样的可视化面板



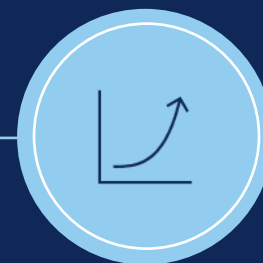
零侵入性

- ✓ 支持托管自有集群
- ✓ 对集群完全无侵入性



用户体验

- ✓ 现代化前后端分离架构，提供优秀、直观的界面
- ✓ 简化运维流程
- ✓ 增加开放支持



专业支持

- ✓ 由谡流科技全职团队维护，确保功能和安全性持续更新
- ✓ 逐步开放开源
- ✓ 永久免费承诺





加入社区

扫描下方二维码，加入 ASP 社区版交流群，与其他 Pulsar 爱好者和专家共同探讨、分享经验：



谙流公众号



谙流视频号



加我进ASP社区入群



THANK YOU

傅腾

futeng@ascentstream.com

www.ascentstream.com



直播回放/更多资料

欢迎关注 谙流科技AscentStream 及 ApachePulsar 公众号



关注谙流



关注Pulsar



加入社区

