

#### 流数据优势：

对于持续生成动态新数据的大多数场景，采用流数据处理是有利的。这种处理方法适用于大多数行业和大数使用案例。通常情况下，各公司一开始都是从简单的应用程序开始，例如收集系统日志以及进行滚动计算最小值-最大值等初级处理。然后，这些应用程序逐渐发展为需要完成更加复杂的近实时处理。最初，应用程序可能通过处理数据流生成简单的报告，然后再执行一些简单的响应操作，例如在关键指标超出一定阈值时发出警报。最终，这些应用程序会执行形式更加复杂的数据分析，如应用机器学习算法，还会从数据中提取更深入的信息。经过一段时间后，开始应用复杂的流事件处理算法，如利用时间窗口衰减算法查找最近的热门电影，进一步丰富了信息内容。

#### 流数据示例：

1. 交通工具、工业设备和农业机械上的传感器将数据发送到流处理应用程序。该应用程序再监控性能，提前检测任何潜在缺陷，自动订购备用部件，从而防止设备停机。
2. 一家金融机构实时跟踪股市波动，计算风险价值，然后根据股票价格变动自动重新平衡投资组合。
3. 一家房地产网站跟踪客户移动设备中的一部分数据，然后根据其地理位置实时建议应走访的房产。
4. 一家太阳能发电公司实施了一个流数据应用程序，用以监控电力系统中的所有电池板，并实时调度服务，从而最大限度缩短了每个电池板的低产能期。
5. 一家媒体出版商对数十亿的在线内容点击流记录进行流处理，利用有关用户的人口统计信息汇总和丰富数据，并优化网站上的内容投放，从而实现关联性并为受众提供更佳体验。
6. 一家网络游戏公司收集关于玩家与游戏间互动的流数据，并将这些数据提供给游戏平台，然后再对这些数据进行实时分析，并提供各种激励措施和动态体验来吸引玩家。

#### 数据流特点：

将数据生产者与数据消费者相互解耦。当数据生产者将其数据简单写入数据流时，生产者不需要知道读取数据的消费者。消费者可以独立于生产者进行添加和删除。消费者可以在生产者不知情的情况下，启动/停止或暂停并恢复他们的消费。这种解耦简化了数据生产者和使用者的实现。

数据流是在大型分布式系统中存储和共享数据的一种非常有用的机制。如前所述，数据生产者只需将数据发送至数据流系统。生产者不需要知道任何关于消费者的事情。消费者可以在不影响生产者的情况下，上线、下线、添加或者移除自己。

#### ~~~~~ 数据流网络的应用场景分类：

1. 聊天与社交；
2. 多人游戏交流与协作；
3. 实时消息推送；
4. 金融行情数据流；
5. 系统日志采集；
6. 健康与安全领域；
7. 家居自动化；
8. 供应链；
9. 跨组织的实时协作、审计、跟踪等方面；
10. IoT/IoV/IoE等领域的设备监测、消息控制、持续性测试等方面；

#### 一) 聊天与社交

聊天是主流的应用，从简单的一对一到群组聊天，实时通信中发生越来越多的消息传递，包括聊天机器人（客服支持）、约会、地图、游戏、协作等。

由于人工智能的发展趋势，未来聊天机器人将会扮演重要角色，它可以与人类进行自然语言对话，其功能将会不断强大。从简单的基于规则的自动响应，到现在使用深度机器学习和AI，聊天机器人已经渗透到各个领域的聊天应用程序中，以多种方式帮助聊天者。他们可以基于语音和文本输入进行搜索，分析，过滤，增强和帮助，并且该功能正逐渐渗透到客户支持，零售和旅行等应用程序中。

需要涉及的服务提供商或技术，可分为：

- 消息协议（WebSocket、Long Polling, HTTP Streaming, Comet, MQTT）
- 开源服务和框架（socket.io, DeepstreamHub, Meter）
- IaaS云服务和云架构（AWS, DigitalOcean, Microsoft Azure, IBM Bluemix）
- PaaS 数据流网络，实时即服务（PubNub, Firebase, DeepstreamHub Cloud, Meter Galaxy）
- 聊天服务提供商，聊天API/聊天框架与SDK（ChatEngine, Layer, QuickBlox, Nexmo）
- SaaS 聊天应用（Intercom, zendesk chat）

隐私聊天的场景：

- 无限规模的1:1聊天，或多对多聊天。
- 通信的加密机制。
- 供第三方应用采用基础设施，如约会App等。

现场活动中的聊天场景：

- 允许用户彼此交互，可实时通信，补充视频或现场活动；
- 将聊天应用程序与实时流视频相结合，并显示当前有多少用户正在观看实时事件；
- 收集和汇总用户的地理位置数据，以了解世界各地有多少人在观看该事件；
- 允许用户对自定义民意调查投票，以增加用户参与度；
- 为现场和远程观众构建交互式且引人入胜的应用程序，例如允许用户举手或支持其他消息；
  - 实时发送各种来源的数据到应用界面；
  - 推送交互内容，如投票，调查等；
  - 向连接的用户传递推送通知和警报；
  - 实时监视所有订阅用户的在线/离线状态；

按需电子商务中的聊天场景：

- 提供完整的购物体验，从浏览到通过聊天履行订单；
- 利用具有自动和AI驱动的认知功能的聊天机器人来协助购物者通过消息传递界面浏览，订购和排查订单故障；
- 随着订单在执行完整的生命周期中的推进，流式传输订单更新和警报，例如推送通知和SMS；

聊天应用中的内容审核场景：

- 当今的聊天应用程序需要其中的审核工具来即时过滤消息，不仅用于文本，还用于图像；
- 即时发布消息时聊天，需要滤掉不适当的消息；
- 使用图像识别来实时过滤图像以滤除挑衅性图像；

地理位置与聊天协同的场景：

- 与其强迫用户在消息传递和地图应用之间切换，还不如直接将地理定位传递给聊天用户；
- 将具有实时地理位置跟踪的实时地图直接嵌入到消息流中。电子商务和按需服务公司可以利用此功能使客户实时监控其交付或服务；

提升用户体验的聊天应用场景：

- 实时发送任何格式的较大文件。高效的聊天应用程序不会发送整个文件，而是发送指向文件托管位置的引用，供最终用户下载；
- 用户应该能够对重要消息添加书签/加星标，并在以后轻松地将其检索；
- 邮件搜索和邮件标记/固定，并给予实时消息通知或传递等；

聊天机器人集成的应用场景：

- 当今市场上有大量强大的聊天机器人服务，可供继承到应用程序或微服务中；
- 在情况可预测时使用自动化发送响应。例如，用户开始与网站上的支持代理聊天。虽然可能无法立即提供座席，但您可以收集用户信息，例如电子邮件如自动回复所说：“我们的代理商会在不久后回覆。请提供您的电子邮件”。这使用户受益，因为他们无需等待很长时间即可收到代理商的回音，从而节省了时间。
- 在数据可用和可访问时使用自动化发送响应。在电子商务网站的示例中，用户可以在聊天框中键入“我想知道我的订单的退款状态”。这句话的目的是“退款状态”。可以使用迷你表单将自动回复发送给客户，要求他们“请提供您的订单ID和电子邮件”。这增强了用户数据收集体验。
- 通过询问问题并根据用户在网站上的位置进行上下文自动响应，我们可以改善用户的体验。例如，当他们在付款页面上时，您可以向他们显示他们可以使用的个性化优惠券。
- 将活动历史记录与客户输入相结合，以提出建议并实时回答问题。

驱动WebRTC和VoIP应用的场景：

- 利用实时信令，构建WebRTC和VoIP的连接和数据通道
  - 向用户传递应用内或推送通知和警报。
  - 将一对一或群聊添加到通信应用程序。
  - 呼叫等待，呼叫者ID，在线/离线等。
  - 强大的TLS和AES加密可确保网络上端到端的消息安全。

## 二) 多人游戏交流与协作

- 大型多人游戏聊天，成千上万的用户需要同时进行互动或协作，它需要可靠且稳定的用户体验。
- 拥有单独的分组讨论室或通道，这些讨论室或通道需要高速运行，且不会发生中断。
- 众多的用户群体，需要第三方的机器学习对接，如情感和语言分析，过滤不雅消息。也可以实时翻译代理，便于连接全球玩家。

多人游戏场景：

- 多人同步游戏状态：在所有连接设备之间，进行可靠的游戏状态、事件和输入的同步；
- 多人游戏中的聊天：轻松实现多人游戏体验的互动社交功能；
- 在线/离线的用户列表：创建在线和离线玩家的实时花名册；
- 多人游戏中警报和通知：提供推送通知和游戏内警报；
- 多人游戏实时统计和得分：游戏的得分和统计信息通过数据流的方式，显示到仪表板。

## 三) 实时消息推送

- 多方应用或系统集成，需要实时消息的推送，如警告、通知等；
- 通过HTTP，SMS，FCM，APNS2等上的实时通知，使每个人和所有内容保持实时。
- 当现实世界中发生某些事情时，需要立即与您的用户共享。无论是广播推送通知，跟踪地图上的位置还是更改股价，实时通知都可以提高受众的参与度，保留率和投资回报率。

地图通知的场景：

- 在不需要服务器或手动刷新的情况下，将任意不同来源的地理位置数据，通过流式传输，在对其处理后，将其显示在地图中。
- 添加实时动态映射功能，如地理围栏，地理编码，地理哈希，地理搜索，路线。可将交互式强大的地理位置功能构建到连接的应用程序中。

实时更新的场景：

- 实时更新，警报和推送通知；如从iOS，Android，Web的推送通知。也可以根据任何事件触发SMS；
- 广播实时通知，并做可视化，通过实时数据流使数据栩栩如生，可自定义并将任意数量的数据流发布到仪表板，比如提供实时更新的仪表板，用于各种数据可视化应用中。

## 四) 金融行情数据流（包括区块链数据流）

- 实时将金融数据更新流式传输到任意数量的设备或订阅者；
- 将推送通知和警报传递给已连接的用户；
- 实时触发金融交易和事件；
- 创建具有最新价格指数的实时财务仪表盘；

拍卖场景：

- 为了模拟实时拍卖或电子商务体验，需要确保在价格变化时以流方式传输和发布价格更新。在社交环境中，拍卖网站需要速度和可靠性来创建令人兴奋的快节奏用户体验。
  - 将财务数据实时更新给任意数量的订阅者；
  - 根据用户互动（即提交出价）触发价格变化；
  - 推送通知和警报给用户；
  - 使用户能够彼此实时通信。
  - 监视投标人和主持人的在线/离线状态。

## 五) 系统日志采集

- 各种服务器日志的实时采集，集中与存储；
- 各种应用程序日志的实时采集，集中与存储；

- 支撑流数据的分析，在各种实时Web服务中，如搜索引擎、购物网站的实时广告推荐，SNS社交类网站的实时个性化内容推荐，大型网站、网店的实时用户访问情况分析等。

#### 六) 健康与安全领域

- 在设备之间发送和接收医疗保健数据，或向任何数量的用户实时广播警报。
- 除加密数据外，需要符合医疗行业的标准要求。
- 触发设备操作并实时监视各个设备的状态。
- 需要与第三方服务的集成。
  - 医疗保健机构：PHI / EHR转移，符合HIPAA的协作，医院设施管理，按需紧急护理，医疗设备控制和监控；
  - IT和在线服务：符合HIPAA的在线聊天，调度和虚拟队列管理，紧急派遣和响应，公共卫生警报和通知服务，公共卫生地图；

#### 驱动可穿戴设备：

- 实时通信：发布/订阅模式；
- 双向信令：进行远程设备控制；
- 触发事件和动作：以基于接近程度发出信号，触发设备的事件和动作。
- 推送通知和警报：向移动和网络用户实时发送相关的推送通知和警报。
- 远程固件升级：使用安全的专用通道无需用户采取任何措施即可远程升级固件。
- 设备元数据监控：使用网络，持续跟踪设备元数据。
- 最小带宽使用：独立于协议的方法，使用监视和控制设备所需的最小带宽。

#### 七) 家具自动化

- 安全控制和监视智能家居中的连接设备，也是任何家庭自动化解决方案的核心。
- 让用户轻松地在几毫秒内触发设备交互，出门在外时监视其设备，并在知道设备安全的情况下轻松入睡。
  - 通过API进行远程设备控制。
  - 向移动和网络用户发送实时推送通知和警报。
  - 在客户端应用程序和家庭自动化中心之间无缝通信。
  - 即使通过家庭和企业防火墙，也可以使您的IoT设备即插即用。
  - 使用安全的专用通道无需用户采取任何措施即可远程升级固件。
  - 持续跟踪设备元数据。

#### 八) 供应链

- 伴随着新一代物联网技术的广泛应用，尤其是人工智能、信息技术、大数据、云计算等技术迅速发展，供应链领域的资金流、商流、信息流和物流“四流”得以高效连接，供应链发展到智慧供应链新阶段。
- 在各个环节，不同的系统模块之间，都需要实时的协作与通知；
  - 推送通知和警报给不同环节/系统的用户；
  - 使不同环节/系统的用户能够彼此实时通信。
  - 监视不同环节/系统中关键资源的各种状态。

#### 九) 跨组织的实时协作、审计、跟踪等方面

- 不同组织之间的应用、系统，需要通过第三方的异步流网络进行交互协作；

#### 十) IOV（车联网）

- 实时跟踪车辆的地理位置并触发调度事件；
- 使用任何运输或按需交付的应用程序，构建快速可靠的跟踪和调度功能，如能够跟踪位置，记录接送等事件，并根据命令或根据接近程度自动调度车辆。
  - 根据命令实时或根据接近程度自动调度车辆。
  - 在实时更新的地图上跟踪，监视和传输车辆位置。
  - 持续跟踪设备的元数据。
  - 在乘客，驾驶员和调度员之间建立紧密的实时通信回路。

#### 十一) IOT（物联网）

- 应用于传输在智能城市周围的传感器数据，用于智能工厂内传感器或者来自其他物联网设备传感器的流数据。

- 物联网设备控制：在连接的物联网设备之间安全地监视，控制，提供和传输数据。
- 触发设备操作，监视元数据，或流传输和处理传入和传出的数据。
- 可用于任何规模的IoT部署的通信基础设施和API。
- IOT设备长年累月的运作测试，以确认其设备稳定性。

构建安全的物联网基础设施的场景：

- 互联网上任何具有开放入站端口的设备都会受到攻击，这仅仅是时间问题而已；连接到安全物联网网络的设备应仅建立出站连接。这些连接不像打开入站端口那样容易受到攻击。仅出站设计模式消除了对物联网设备的主要威胁。为了支持这种设计模式，需要使用发布/订阅通信设计，以便设备可以双向发送数据。使用MQTT，CoAP，WebSockets和HTTP 2.0等协议的安全可靠的通信能够为没有开放端口的设备之间的发布/订阅通信提供支持。无论使用哪种协议，向外打开连接并使其保持打开状态都是最重要的，其次是使用发布/订阅作为该连接通信的范例。为了满足物联网规模的需求，发布/订阅连接应由分布在世界各地（具有数据存在点）的高性能服务器管理（数据流网络）。
- 虽然TLS / SSL适用于数据传输安全性，但从物联网设备生成的数据除非经过加密，否则在网络上仍然容易受到攻击。为了实现真正的端到端安全性，应使用高级加密标准（AES）加密规范对数据本身进行加密。对谁可以传输什么内容以及可以传输和接收数据的精细访问控制。由于可能有数以百万计的设备试图收听正确的频道和主题，因此要求终端设备过滤掉他们未订阅的主题是非常低效且不安全。相反，网络应能够处理大量任务。
- 在发布/订阅范例中，可以使用基于令牌的访问控制方法将令牌分发给设备，以授予对特定数据通道的访问权限。这种方法可以对创建哪些令牌，哪些设备接收这些令牌以及这些令牌授予访问哪些数据进行细粒度控制。它还可以集中控制撤销令牌的时间和方式，例如，切断非付费客户的数据流访问。这样，网络可以有效地充当交通警察，基于网络分发的令牌，授权设备访问并管理哪些设备可以在网络上发布和收听。
- 在消费和工业物联网中，主动监视设备的在线/离线状态（“存在”）至关重要。当诸如家庭安全监控器，油田传感器或家用电器之类的设备消失或停止发送和接收数据时，所有者或监视系统需要了解它。脱机设备可能意味着正在进行本地篡改，或者发生了更广泛的问题，例如电源或Internet中断。
- 物联网元数据跟踪需要一个独立的安全数据通道来流传输每个设备的状态数据，可以对其进行自定义以流传输在线/离线状态以及其他自定义状态，例如温度，加速度或地理位置。设备状态的每个方面都可以拥有自己的发布/订阅通道，以通过网络流式传输“心跳”，然后可以将其用于警报和其他操作触发器。例如，只有当门锁的手机不在前门20英尺范围内时，远程门锁才能向其所有者发出锁状态变化的警报。或者，如果太阳能发电厂的一系列传感器脱机，则网络可以立即派遣技术人员来调查问题。
- 使用标准出站端口80和443的发布/订阅范例可轻松安全地设置和配置IoT设备。插入电源后，设备会唤醒，订阅指定的“通知”频道，然后将自己通知数据流网络。然后，服务器返回一个专用通道，设备和服务器可以在该通道上进行通信。然后，服务器可以在通道上设置访问规则，并从那里进行设置。所有这些立即发生，从而为消费者带来了她所期望的即插即用体验。设置并配置设备后，实施一种方法来安全更新该设备的固件也很重要。如果消费者负责下载和安装固件更新，则他们可能不会下载关键更新，或者容易受到未经授权来源的恶意更新的攻击。为了应对这种情况，制造商应使用设备的安全发布/订阅通道来指导设备在可用时下载和安装固件更新。现场实时固件更新的设计模型始于服务器在所有设备都可以安全读取的频道上广播固件警报消息。然后，主服务器指示设备如何访问和安装更新。与存在状态监视配对时，IoT制造商可以确保如果单个设备处于脱机状态，则一旦重新启动，它将从网络接收到固件更新消息。在线设备会立即安装固件更新。