

22 | 答疑解惑：不同即时消息场景下架构实现上的异同

2019-10-16 袁武林

即时消息技术剖析与实战

[进入课程 >](#)



讲述：袁武林

时长 09:56 大小 7.97M



你好，我是袁武林。

随着专栏最后一个进阶篇模块的更新后，咱们的即时消息专栏课程，到这里就要告一段落了。首先，感谢你在这段时间里对专栏的持续关注，也非常高兴看到你一直在积极地思考和学习。

在专栏的讨论区，同学们也都十分活跃，都在热情地留言和互动讨论，留下了很多比较典型和有意义的问题。

有一些问题可能限于篇幅的原因，我没有详细展开，那么今天，我就摘录出比较有代表性的 5 个问题，来做一下集中的整理和回复。也希望通过这种方式，能够帮助你对这些知识点有更清晰的理解和认识。

第一问：消息的服务端存储和本地存储

在[第 1 讲 “架构与特性：一个完整的 IM 系统是怎样的？”](#) 中，有很多同学都问到了这个问题：**即时消息系统实现中，消息一定需要在服务端的存储服务里进行存储吗？**

首先呢，关于服务端存储的问题，我们需要更多地考虑存储成本和数据安全。

一方面，如果消息不在服务端存储，服务器只是作为消息的中转路由服务，那么，相应的消息存储成本就会低很多，在数据安全性方面也更好一些。

另外，这个问题也和产品定位有关。

比如，如果你的产品定位上不需要支持消息多终端同步（比如微信），那么像核心的消息等这些数据，就可以不在服务端进行存储。不过，用户的好友关系等数据信息还是需要存储在服务端的。

你还需要考虑的一点是，即使是不需要支持多终端消息同步的产品，大部分即时消息系统也是支持离线消息的，这种情况其实也需要在服务端对离线消息进行暂存，虽然可能只是暂存较短时间。

除此之外，消息是否需要在服务端存储，你还需要考虑国内监管机制是否允许的问题。如果监管有要求，那么我们所有的消息数据，都需要在服务端存储一定的时间供监管调看，只是这里存储的使用方不是普通用户，而是监管部门。

第二问：长连接消息推送的实现

在[第 3 讲 “轮询与长连接：如何解决消息的实时到达问题？”](#) 中，我留下了一个思考题：**TCP 长连接的方式是怎么实现 “当有消息需要发送给某个用户时，能够准确找到这个用户对应的网络连接” ？**

对于这个问题，不少同学给出的答案都很棒，比如 @王棕生、@小可等几位同学。这个问题由于本身比较重要，提问的同学也不少，所以我这里也专门来回答一下。

首先，用户在长连建立后，需要再执行一个“**上线**”操作。

这个“上线”操作主要的工作就是：

将当前登录用户的信息提供到服务端，服务端长连接收到这个用户信息后，在服务端维护一个“用户” -> “连接”维度的映射，这个映射可以存到中央资源里或者网关机的本地内存中；

当有消息需要推送给这个用户时，负责消息下推的服务查询这个中央的“用户” -> “连接”维度的映射，来获取该用户的连接，通过这个连接将消息进行下推；

或者将消息下发给所有网关机，由网关机来查询本地维护的这个映射，是否有该用户的连接在本机，如果有，就通过当前网关机维护的这个连接来进行消息下推；

当用户断连下线的时候，再从这个中央的“用户” -> “连接”维度的映射或者网关机本地删除掉这个映射。

第三问：应用层 ACK 的必要性

在课程的[第 4 讲 “ACK 机制：如何保证消息的可靠投递？”](#)中，我留下的思考题是：**有了 TCP 协议本身的 ACK 机制，为什么还需要业务层的 ACK 机制？**

这个问题大家讨论得也比较多，有几位同学也都比较正确地讲出了这两种 ACK 的差异性，比如 @小伟、@恰同学少年、@阳仔、@王棕生等同学都回答得很棒。我这里也简单说一下这个问题的答案。

这是因为，虽然 TCP 协议本身的 ACK 机制，能够保证消息在正常情况下传输层数据收发的可靠性，但在连接异常断开等场景下，也可能存在数据丢失的风险。比如，TCP 的发送缓冲区和接收缓冲区里的数据都可能会丢失。

另外，即使消息从 TCP 传输层成功给到了应用层，也并不能保证应用层数据收发的可靠性，因为应用层在接收到传输层的数据后，也可能发生其他异常。

比如，手机 Crash 了或者突然没电关机了，又或者客户端在将消息写入本地数据库时，发生异常失败了，这些情况都可能会导致消息即使成功在 TCP 层被 ACK 了，但在业务层上仍然会被丢失。

但是业务层的 ACK 机制，是在应用层接收到数据，并且成功执行完必要的存储等逻辑后，再 ACK 给服务端，所以能够更好地保障消息收发在业务层的真正可靠性。

第四问：离线消息下推的优化

在[第 9 讲 “分布式一致性：让你的消息支持多终端漫游？”](#) 中，有不少同学还问到：**在离线消息如果特别多的情况下，我们应该采取什么方式来下发这些消息呢？**

对于长时间没有登录或者消息接收频繁的用户来说，当这些用户下线一段时间后再次上线，经常会面临大量的离线消息下推的问题。

一种比较好的优化方案是：对多条离线消息采取批量打包 + 压缩的方式进行下推。

通过这种方式，能够大幅降低传输的数据大小，也能有效减少大量离线消息下推后的 ACK 包数量。

但是面对上万条的离线消息，只单纯地采用批量 + 压缩的方式还是不够的，这种情况下，我们还可以采用 **“推拉结合”** 的方式进行优化。

用户上线后，对于服务端离线消息的下推，可以只推送用户最近 N 条消息；而对于后续要推送的消息，我们可以让用户通过向上翻页来触发自动拉取，客户端再从服务端来获取当前聊天会话剩下的消息。

另外，在实际的离线消息场景里，我们还需要考虑到离线消息存储成本的问题。

绝大部分 IM 系统并不会一直缓存用户的离线消息，而是一般会按照条数或者时间周期，来保留用户最近 N 条离线消息，或者最近多长时间内的离线消息。

第五问：群聊和直播互动消息下推的区别

在[第 10 讲 “自动智能扩缩容：直播互动场景中峰值流量的应对”](#) 一课中，有同学对这个问题感到疑惑：**群聊场景和直播互动场景在消息推送时，实现上不一样的地方在哪？**

@淡蓝小黑同学也在留言中问到：

文中提到【通过这个优化，相当于是把直播消息的扇出从业务逻辑处理层推迟到网关层】和您回复我的【业务层在把消息扇出给到网关机时，就已经具体到接收人了，所以不需要通知网关机来进行变更。】这两条不冲突吗？(我看唯一的区别是文中说的是直播间场景，您回复我的是群聊场景，还是说群聊和直播间是用不同的模式扇出消息的？)

其实我想问的是，用户加入直播间，网关机本地维护【本机某个直播间有哪

些用户】这个数据的话，当用户离开直播间，加入另一个直播间时，业务处理层还要通知网关层更新本地维护的那个数据，有可能会出现数据不一致的情况，导致用户加入新直播间了，但由于网关层数据没有更新，用户收到不到新直播间的消息。

这是个好问题！对此，我们可以先想一想群聊和直播的使用场景。

对于直播场景来说，房间和房间之间是隔离状态。也就是说，每次用户进入到一个新房间，消息推送层面只和用户当前进入的房间有关联，与其他房间就没有关系了。

所以，每次用户进入新房间时，都需要通过“加入房间”或者“切换房间”等这些携带用户信息和房间信息的信令，来告诉网关机自己当前连接的是哪个房间。

正是通过这些“加入房间”和“切换房间”的信令，网关机才能够在服务端标记这个“用户 -> 房间 -> 连接”的映射关系。当这些网关机收到某一条直播间的消息时，就能够根据这个映射关系，找到对应的房间用户的连接。

换言之，对于直播间消息来说，当我们的消息从业务层给到网关机时，服务端只需要按房间维度来下发消息就可以了。所以在直播互动场景中，消息的扇出是可以推迟到网关机层的。

但是对于群聊来说，群和群之间并不是隔离状态。

对于任何群的消息，服务端都需要通过一条长连接通道来进行消息的下推，并没有一个所谓的“进入群聊”来切换多个群的行为。

所以，在 App 打开建立好长连后，客户端发出的“上线”这个操作，只是会上报一个当前用户信息，不需要、也没有办法再告知自己当前进入了哪个群，因而在服务端网关机也只会建立一个“用户” -> “连接”的映射。

因此，在群聊场景中，当某一个群有一条消息发出时，我们需要在业务层将这条消息从群维度扇出成 UID 维度，再下发给网关机。

那么，我们为什么不能在网关机本身查询了群成员的相关信息后，再扇出消息呢？

这个操作在实现上当然也是可以的。不过咱们前面提到，为了解耦网关机和业务层逻辑，我们尽量不在网关机做业务维度的逻辑。

因此这里建议对于群聊消息来说，我们还是在业务层扇出后，再提交到网关机进行处理。

小结

正如我在开篇词里讲到，即时消息技术实际上是众多前后端技术的紧密结合，涉及到的知识面也是非常广泛的，而且随着业务形态的不断变化和升级，相应的架构设计和技术侧重点上也各有异同。

因此，形而上学的方式往往很难真正做到有的放矢。但我相信，只要我们从问题和业务形态的本质出发，经过不断的思考和实践，就能够在掌握现有即时消息知识的基础上，找到最适合自身业务的架构和方案。

好了，今天的答疑就到这里，如果你还有问题，可以继续给我留言，我们一起讨论。感谢你的收听，我们下期再见。



即时消息技术剖析与实战

10 周精通 IM 后端架构技术点

袁武林

微博研发中心技术专家



新版升级：点击「 请朋友读」，20位好友免费读，邀请订阅更有**现金**奖励。

© 版权归极客邦科技所有，未经许可不得传播售卖。页面已增加防盗追踪，如有侵权极客邦将依法追究其法律责任。

精选留言 (5)

写留言



小可

2019-10-18

直播房间和群聊消息扇出这个问题特别好，两者确实有区别，老师的解答也很清楚😊

作者回复: 谢谢支持😊



2



CountingStars

2019-10-16

老师我有一个问题，这个用户--> 连接的映射是如何存储在中央资源里的，对于像 java 这样的语言，其实这个所谓的连接其实是一个 java 对象，对于更底层的实现，这个连接其实就是一个表示文件描述符的数字而已，而这个数字不同的网关机肯定会重复的，所以我想问一下，这个用户到连接的映射到底是怎么存储在中央资源里的。对象或者一个可能会重复的数字，存储起来也没有意义吧。是不是用户到连接的映射实际上只存储在网关机...

展开

作者回复: 是的，用户上线的时候把用户和连接的网关ip作为映射存在中央存储，同时网关机本机内存也存储一个uid到连接的映射（这个映射可以直接把连接这个对象放在Map里），然后消息推送时读取中央全局的映射，查询待推送消息的接收人所在的网关机，再通过rpc方式把这条消息发给这台网关机就可以了。



1



煜

2019-10-19

老师，websocket网关和im服务能否用dubbo通信？或者老师对他们之间的rpc通信有什么好的建议



mong

2019-10-17

老师你好，有一个问题，想和你交流一下，关于离线群消息的存储的问题，我想问一下，

离线的群消息你们是为群里的每个人都存储他的离线消息，还是消息只是存储一份；，如果选择存储一份（大家进行共用），你是如何他们在拉离线消息的时候，进行标记哪个用户拉取到群的离线消息的哪一条了？

展开 ▾

作者回复: 一般会每个用户存储一份，不过只会存储消息id，内容真正下推前再获取。



忘

2019-10-17

谢谢老师，以后有问题了，还可以和你讨论吗？

展开 ▾

作者回复: 没问题的呀，随时欢迎大家来交流讨论。

