



How rejecting MAU pricing can unlock innovation in fan engagement



Matthew O'Riordan

CEO and Co-founder
Ably





What it means to be a fan is changing. Whether the fandom focuses on a sports team, a streaming drama series, or a musical artist, fans no longer want to be passive consumers. Instead, they're looking to feel as though they are a part of the action through channels such as live chat, realtime stats, and behind the scenes insights.

But there's a tension in delivering those fan engagement experiences. Without [the right strategy](#), technology costs can quickly make them economically unviable. And, arguably, many of the economic assumptions we take for granted elsewhere are turned on their head when it comes to realtime fan engagement.

The challenges of driving a return on investment from fan experiences falls into three broad categories:

01

Fan engagement doesn't scale linearly

As more fans join an experience, both the number of messages and interactions and the data to be distributed will grow in line with the user volumes, causing the system load to grow exponentially. This means that you have to solve the N-squared problem.

02

Fan engagement is spiky

A goal, the season finale, or the concert trigger intense bursts of activity surrounded by longer periods of much lower usage.

03

Fan engagement is cost sensitive

With extreme peaks and troughs of engagement, pricing models that may work in other contexts, are a poor fit for realtime fan experiences.

There are architectural and technology choices you can make to help manage some of these challenges. For example, serverless and scale-out architectures give you the elasticity to handle surges in demand without paying for resources to sit idle during quieter periods. Similarly, you can make your data delivery more efficient by batching messages and partitioning fans into smaller groups.

However, pricing is harder for you to influence. And the reality is that the peak pricing model, particularly Monthly Average User (MAU) pricing, is the approach taken by most realtime (e.g. pub/sub, chat) platforms. This pricing model is a poor fit for companies delivering fan experiences, because it results in them having to pay for a month's worth of use, even if the user is on the platform for a matter of minutes. And paying high fees based on peak usage could mean you have to compromise on functionality or shelve projects altogether because the numbers don't add up.

But there is a better way. In this article, we're going to delve into what makes the peak pricing model unsuitable for companies providing fan engagement experiences and introduce an alternative model that gives you the room to innovate while driving a return on your investment.

Fantasy league

LIVE LEADERBOARD

	1	FootyFanatic22	3291pts ▲
	2	GoalGuru96	3018pts ▲
	3	PitchPundit23	2947pts ▲
	4	StrikerSensation	2832pts ▲
	5	KeeperSweeper	2711pts ▲
	6	RedCardRebel	2605pts ▲
	7	PassPlayPat	2500pts ▲





The promise of monthly active user pricing

Let's take a moment to remind ourselves of why peak pricing models are so popular with customers and vendors by looking at MAU in particular.

Say you are the customer of a live chat platform that lets your fans connect in realtime. The billable features of that platform are bundled in different tiers typically including:

Number of users.



Peak concurrent connections.



Storage for chat history, rich media, and other data.



At first glance, this seems like an easier way of doing things. True, you might not always use your full allowance but you know that you'll get the same bill each month so long as you stay within those limits.

And, of course the vendor likes this model because it simplifies their billing, makes their income more predictable, and it lets them structure plans to maximize margins.

So, what's the downside?

The problem is that MAU pricing really only benefits those who have relatively consistent usage throughout the month. But fan engagement is all about the peaks and troughs of usage that are common with events. It's a bad deal on a number of levels.

The key issues with MAU pricing

PROBLEM 01

You pay an entire month's fee even for users who connect for just a few minutes within that month.

PROBLEM 02

You get punished for your success because anything that takes you outside your plan's limits will usually lead to outsized overage charges.

PROBLEM 03

You have to consider delaying or shelving new fan experience functionality because it would be economically unviable.

Let's look at how each of these problems come about in more detail.

PROBLEM 01

Paying for fans who aren't there

You could argue that “monthly active users” is a carefully chosen name that hides the fact that you are actually paying for usage peaks. Here's why.

Imagine you're building fan experiences for a Premier League football team. From August to May, your team plays 38 league games. Each match, including pre-game build-up and injury time, drives intense fan engagement for about two and a half hours.

Add domestic cup competitions, European matches, and pre-season friendlies, and you're looking at 45 to 60 matches across the calendar year; each one driving a brief but massive surge in fan activity.

At the upper end, [Manchester United claims 1.1 billion fans globally](#). Even smaller teams, [such as Wolves](#), have millions of fans across the world. That translates to millions of people who want to capture a piece of the game-day camaraderie and excitement enjoyed by fans in the stadium. But after the game? They're back to the rest of their lives and your fan activity reduces to a quiet baseline.

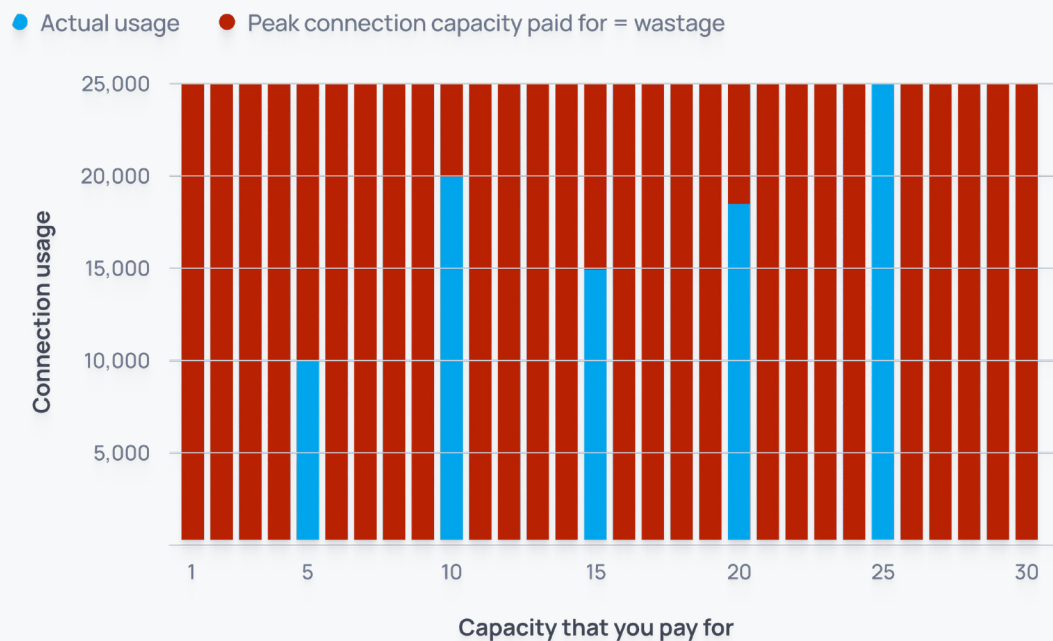
Monthly active user (MAU) pricing means you're charged the same whether someone engages for a single second or every hour of every day.

EXAMPLE

Cost wastage of MAU model

Imagine a scenario with five game days in a month. On one of those days, 215,000 people connect to the fan experience. With the MAU pricing model, you're billed for 25,000 connections for the entire month.

The imbalance of this model becomes clear when represented in the chart below.



Here you can see that on the 25th of the month 25,000 fans connect to the service. The next highest peak is 20,000 connections on day 10. Across the month, there are only five days with any activity at all. Yet, under MAU pricing the bill is the same for having 25,000 connections active every single day. The red areas in the chart highlight the unused capacity you're still paying for—connections billed but never used.

PROBLEM 02

MAU makes you the victim of your own success

Let's say you're locked into an annually recurring plan. You've done your calculations based on regular season games and typical fan engagement patterns. But then the team has unexpected success. Perhaps it's a playoff run, qualifying for European competition, or becoming the subject of a viral documentary series like Wrexham AFC.

Whatever the cause, you find that existing fans are more excited and new fans from around the world want to get involved. You burst through your plan's limits and now you're into overage territory. With some vendors, that can mean punitive additional charges that outstrip the cost of the original contract.

So you're hit twice: paying for unused capacity most of the month, then paying premium rates exactly when your platform needs to deliver.

Even if the platform vendor upgrades you to the next plan, that just means that you're paying a higher fee for each user who is sitting idle for most of the month. **In other words, MAU pricing punishes you for being successful.**

"MAU draws you in by appearing predictable. But in reality, it places penalties on unexpected success with premium rates. And, you consistently pay for those peaks."

Matthew O'Riordan
CEO and Co-founder, Ably



PROBLEM 03

Peak pricing leads to waste - and that makes new functionality harder to justify

When the pricing model locks you into overpaying for your infrastructure, it means that every new feature has to fight twice as hard to show a return on your investment. Even before you consider new functionality, peak pricing means you're already overpaying. Additional usage driven by new features could wipe out any potential return.

Let's use our football example to illustrate the issue. You introduce a new way for fans to comment specifically on live action replays. During the first game after launch, your team scores the match winning goal.

Ten thousand fans open the live action replay on your app to relive the excitement. At first, one thousand of those fans each sends a celebratory message. That alone generates around ten million message deliveries, as all 1,000 messages must be delivered to 10,000 other fans.

With peak pricing, you're now left with an ongoing bill based on ten million deliveries even though the moment lasted a handful of minutes. That's even more painful because usage went to zero after the team's fans had finished watching the replay.

But ten million deliveries represents just the first wave of reactions. If the live action replay inspires one thousand fans to send a celebration message, it's likely that other fans will want to reply. Maybe one hundred messages are funny or insightful enough to generate a reply from a thousand people. Now the platform must deliver one billion messages because those replies all need to be delivered to 10,000 people.

This is the N-squared problem we touched on at the top of the article. Any functionality you introduce that is billed using peak pricing leads to waste because you pay for capacity you won't use. But this is especially true for fan engagement as the numbers scale exponentially rather than linearly. While there are [techniques to reduce the impact of the N-squared problem](#), the issue remains that peak pricing can make new functionality economically unviable.

EXAMPLE

The N-squared problem in numbers

No. of events generated	10,000 fans	50,000 fans	100,000 fans	500,000 fans	1 million fans
Original posts per hour	20,000	100,000	200,000	1 million	2 million
Responses per hour	200,000	1 million	2 million	10 million	20 million
Total deliveries per hour (rounded)	2.2 billion	55 billion	220 billion	5.5 trillion	22 trillion



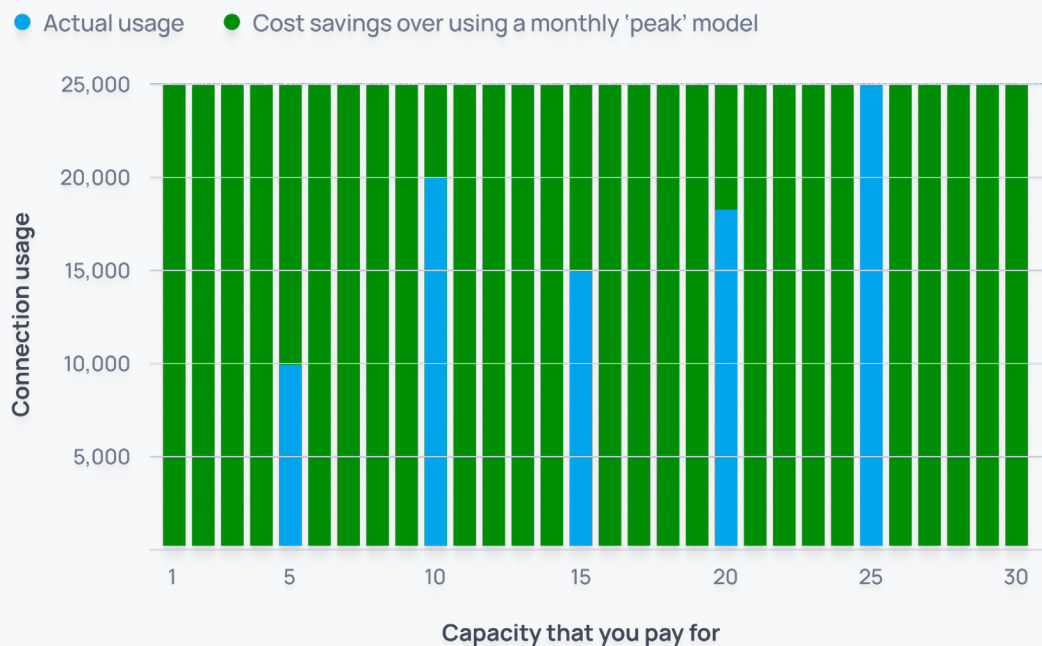
How true consumption-based pricing is fairer and more sustainable

Companies like Slack and Amazon have shown there's a better way than the waste and unpredictability of MAU pricing. They've proven you can charge customers only for what they actually use—down to active workspace members or computing resources by the second. This is at the heart of our new billing model at Aby.

EXAMPLE

Cost savings of a per-minute model

Let's go back to our example of those 25,000 match-day fans dropping to 5,000 during quiet periods. With per-minute billing, you only pay for the time fans are actually connected. The green bars in the graph show just how much this matters: **what would be a hefty monthly fee becomes a fraction of the cost**, because you're only paying for real engagement rather than theoretical capacity.





Consumption-based pricing at Ably

Recognizing the limitations of the peak pricing models such as MAU, we've rebuilt our pricing at Ably around the idea of billing customers for what they actually use. We've taken inspiration from companies like AWS and Snowflake to create a pricing model that matches the natural rhythms of fan engagement.

This means you can focus on building experiences that excite and unite fans, rather than worrying about hitting usage limits or costly overages.

Our approach is built on four key principles:

01

Fan-first pricing

We align costs with actual fan behavior, supporting those moments of peak excitement rather than penalizing them.

02

Pay only for what you use

Just as AWS revolutionized server pricing, we charge based on actual usage—the minutes fans are connected, the messages they share during key moments, and the channels they engage in.

03

Success should reduce costs

As your fan engagement grows, your unit costs decrease through volume discounts, making it more affordable to scale.

04

Smart cost management

Features like server-side message batching help you optimize costs without compromising the fan experience. During high-intensity moments like goals or dramatic plays, these tools can reduce costs by up to 100x compared to MAU pricing.


Get a price estimate


You can see how this works in practice with [our pricing estimator tool](#).


Ready to break free from the MAU pricing trap and build fan experiences without limits?


Start with a [free Ably account](#) today.


Select a use case:


 Live streaming chat

 Large audiences posting & interacting with each other in short periods, during large events

 Data broadcast

 Support chat

 Notifications

 Custom

Adjust to reflect your usage:

How many events per month?

1

How many people attend each event on average?

25,000

What is the average duration of these events?

1

Hrs

00

Mins



The definitive realtime experience platform.

Built for scale.

Aby is the definitive realtime experience platform of the internet. We serve more WebSocket connections than any other pub/sub platform on earth. Businesses like HubSpot, NASCAR and Split trust us to power their critical applications - reliably, securely and at serious scale.

The Aby product suite places composable realtime in the hands of developers. Simple APIs, use case specific SDKs and pre-built integrations, enables the creation of a host of live experiences - including live updates, chat, collaboration, notifications and fan engagement - with no infrastructure to build.

Talk to our technical team

Reach out to learn more about Aby and how we can help you easily build scalable and reliable apps.

[Get in touch](#)

