

1 Engineer, 300,000 Users

@adriangohjw | Red Dot Ruby Conference 2024

400,000

1 Engineer, ~~300,000~~ Users

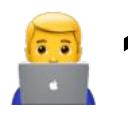
@adriangohjw | Red Dot Ruby Conference 2024



Adrian Goh
@adriangohjw

Co-Founder of NodeFlair

About NodeFlair (People)



1 Full-Time Engineer (Me)

Backend + Frontend + DevOps + Database
Manage (Hire / Onboard)



Many Goals

Ship + Scale + Grow + Keep cost low

About NodeFlair (People)



1 Full-Time Engineer (Me)

Backend + Frontend + DevOps + Database
Manage (Hire / Onboard)



Many Goals

Ship + Scale + Grow + Keep cost low

Approach: Pragmatism

“Boring” is a feature, not a bug

Many Web Pages (> 2 Million)

 Many crawlers and bots - some useful, some not

About NodeFlair (Product + Engineering)

Many Web Pages (> 2 Million)

 Many crawlers and bots - some useful, some not

Read-Heavy Application

 Egress is expensive!

About NodeFlair (Product + Engineering)

About NodeFlair (Product + Engineering)

Many Web Pages (> 2 Million)

🤖 Many crawlers and bots - some useful, some not

Read-Heavy Application

💸 Egress is expensive!

B2C - Users Expect It To Be Fast

🏃 53% of mobile site visits abandoned if load time >3 seconds

About NodeFlair (Product + Engineering)

Many Web Pages (> 2 Million)

 Many crawlers and bots - some useful, some not

Read-Heavy Application

 Egress is expensive!

B2C - Users Expect It To Be Fast

 53% of mobile site visits abandoned if load time >3 seconds

Global Users

 Only as fast (or slow) as your network latencies

About NodeFlair (Product + Engineering)

Many Web Pages (> 2 Million)

 Many crawlers and bots - some useful, some not

Read-Heavy Application

 Egress is expensive!

B2C - Users Expect It To Be Fast

 53% of mobile site visits abandoned if load time >3 seconds

Global Users

 Only as fast (or slow) as your network latencies

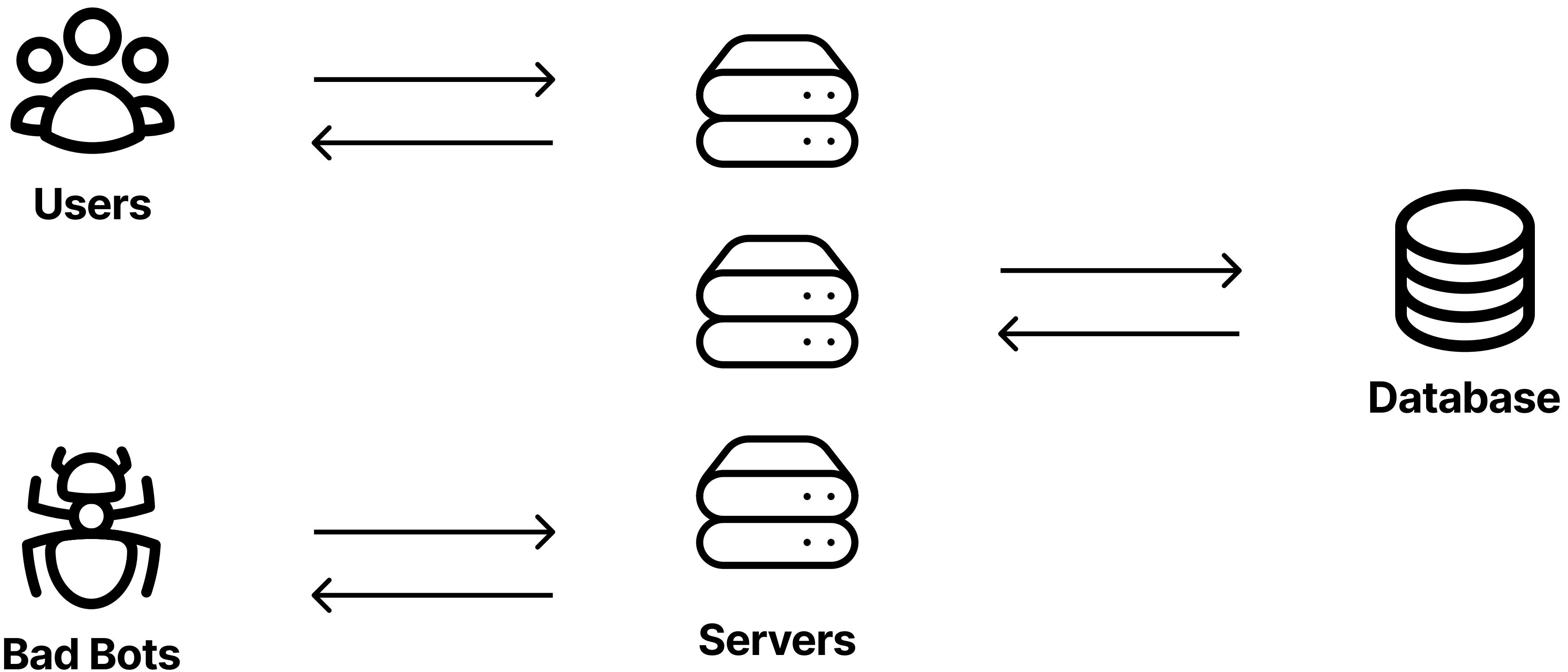
Data Intensive

 NodeFlair Salaries have ~6 million data point

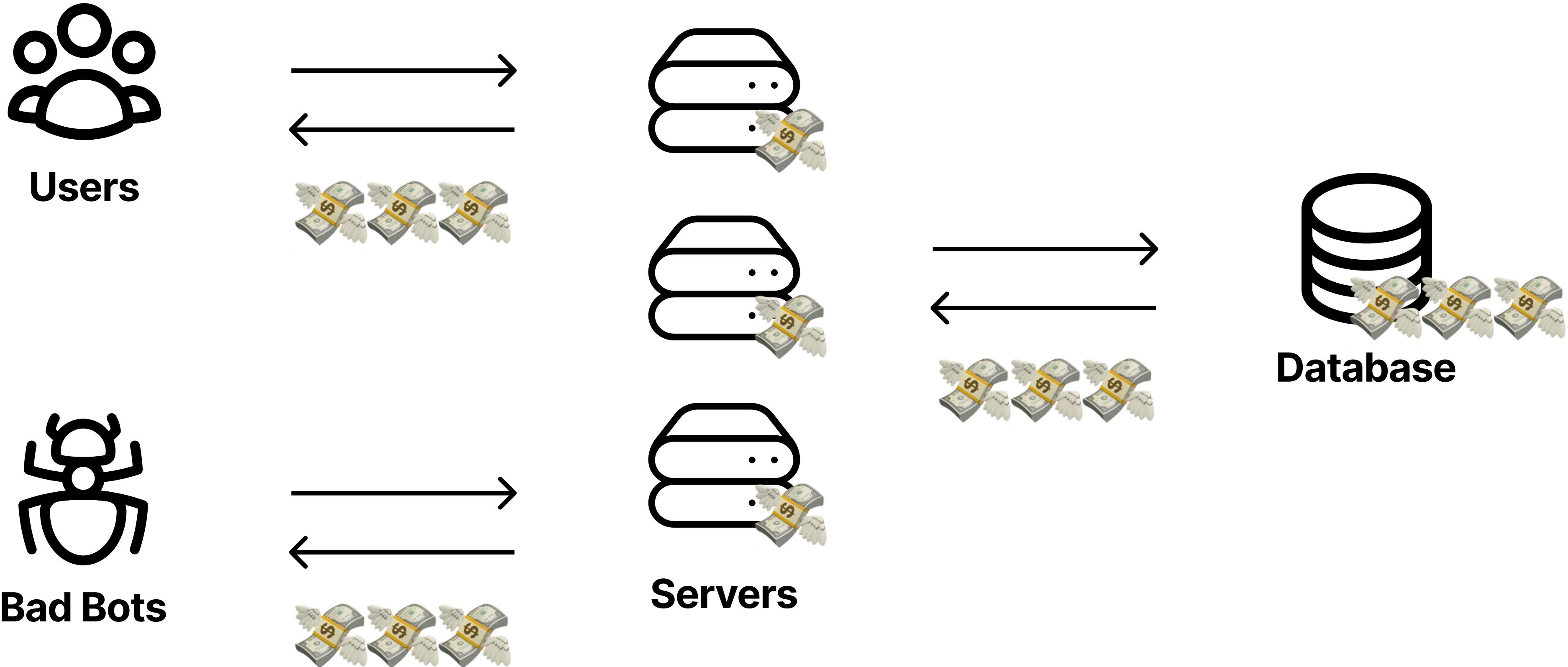
Heads Up

One size doesn't fit all

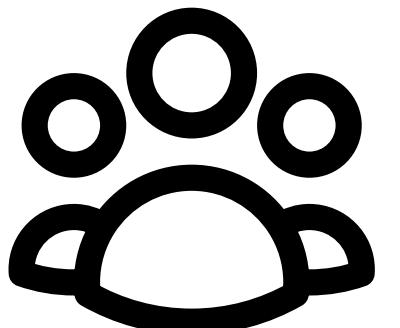
How it started...



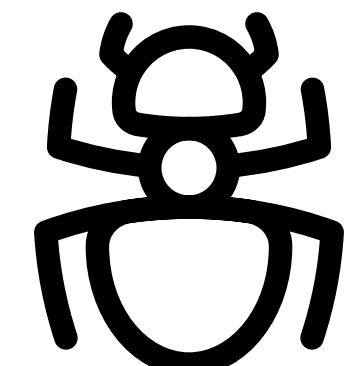
Hardware & Egress are Expensive!!!



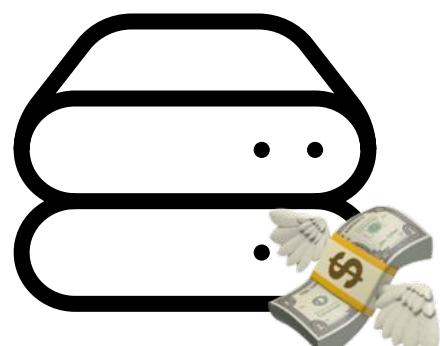
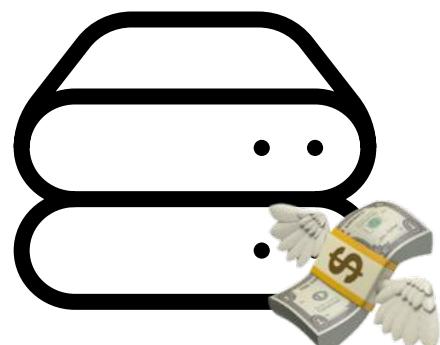
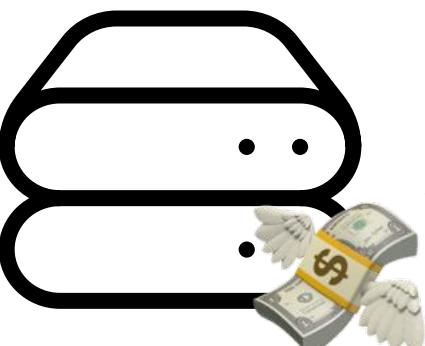
Hardware & Egress are Expensive!!!



Users



Bad Bots

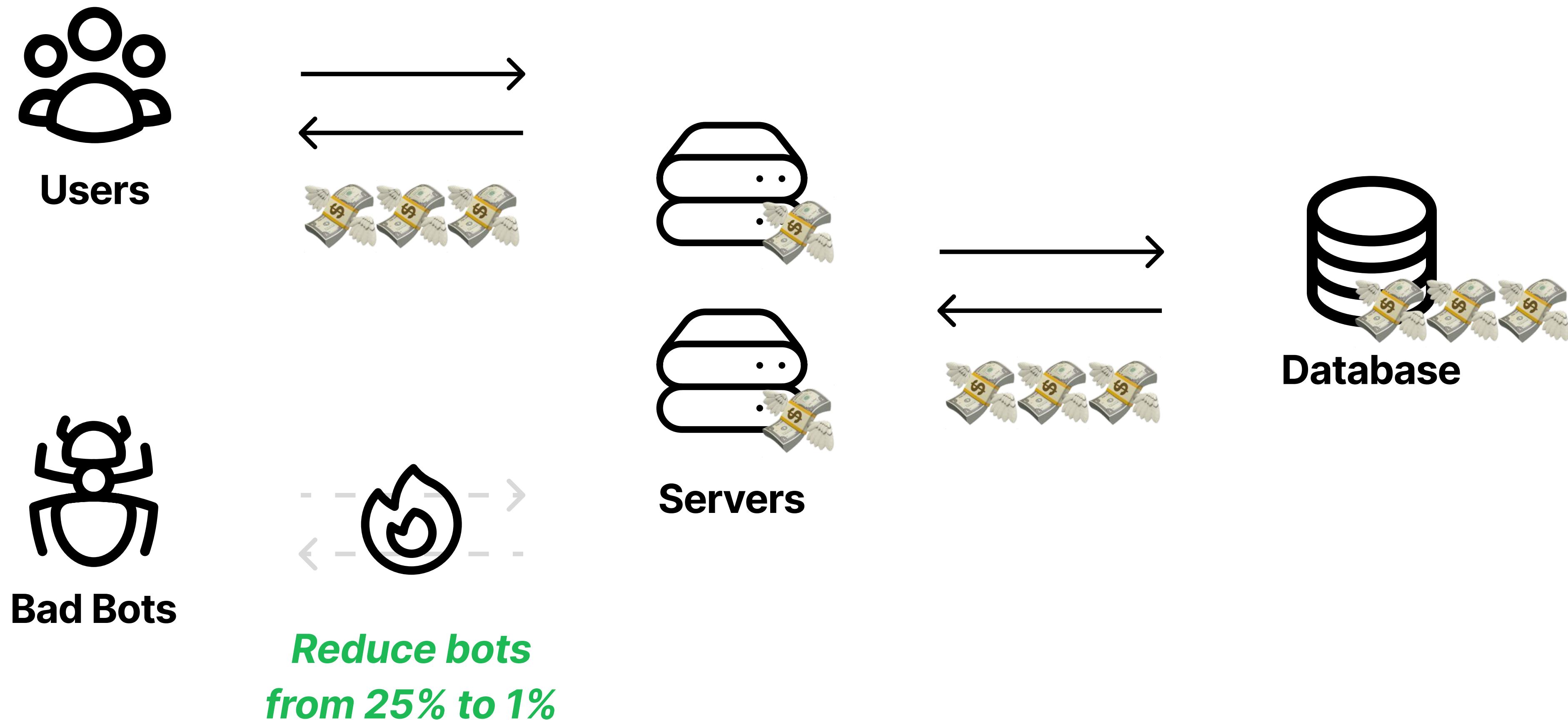


Servers



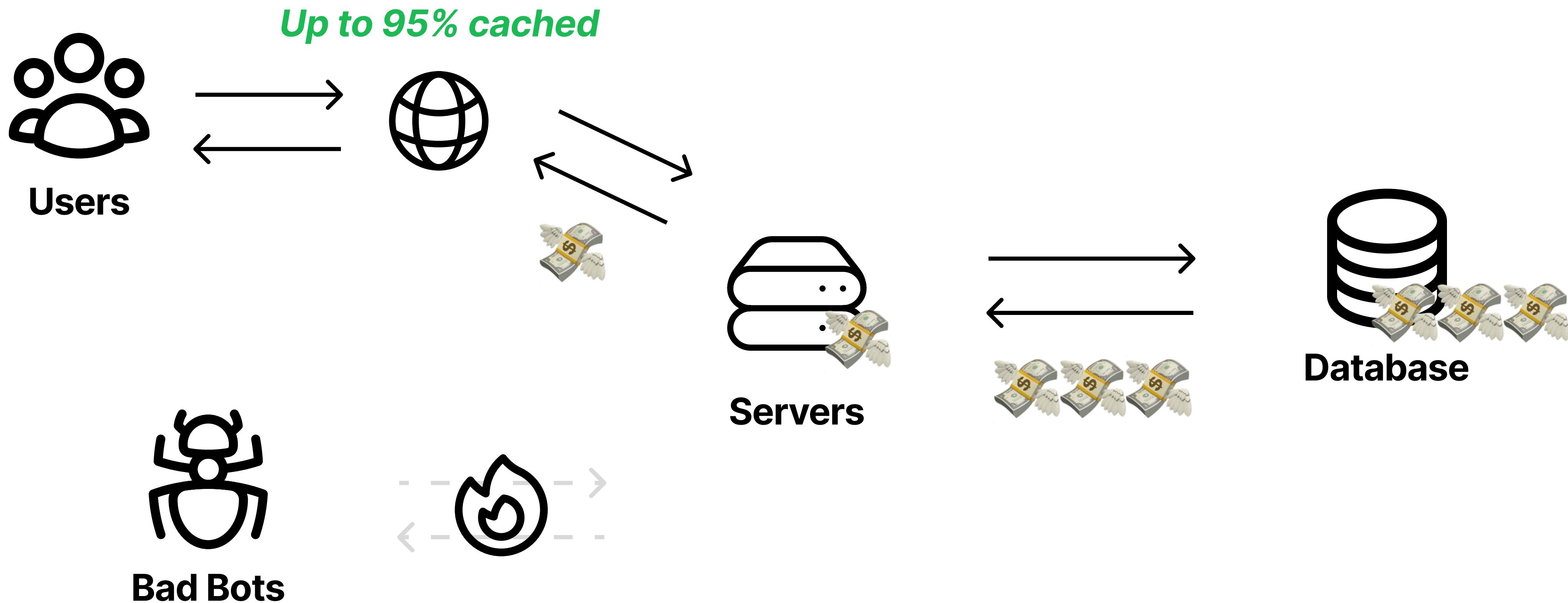
Tip #1

DDoS Protection + WAF (Web Application Firewall)



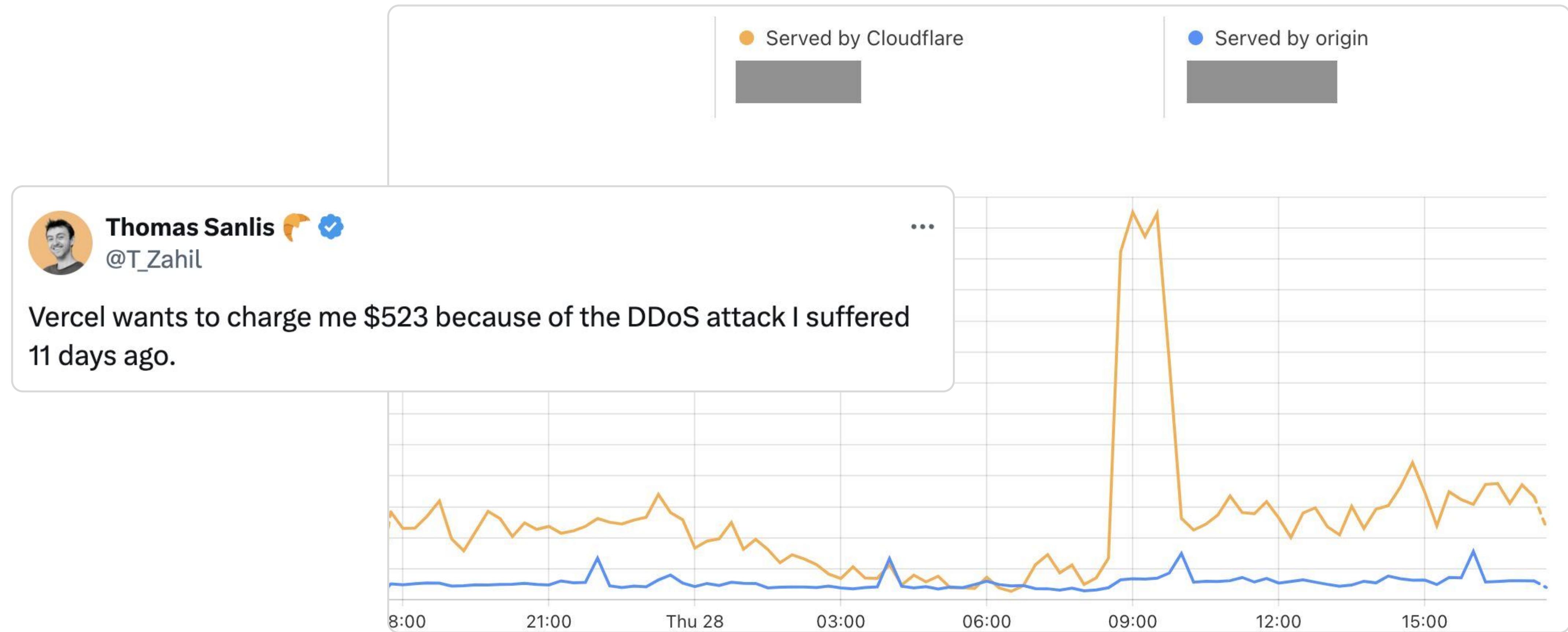
Tip #2

CDN to Cache Everything - Cheaper & Faster



Tip #2

CDN to Cache Everything - Further DDoS Mitigation



Tip #3

Optimizing DB: Choose the Most Suitable Option

HETZNER

fly-apps/postgres

Deploy a Postgres database on Fly, ready for your Fly apps.



Self-Managed

Cheaper + More work + Higher Risk

Managed

More Expensive + Less work + Lower Risk

 Google Cloud

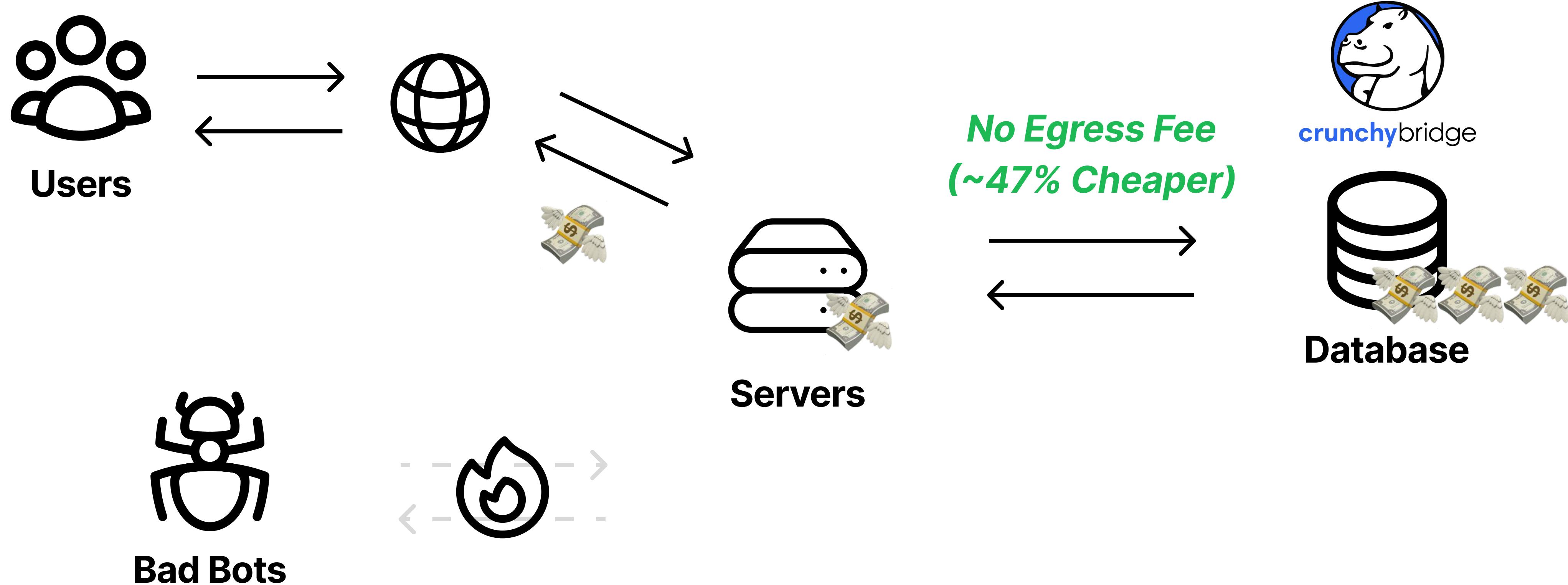
 supabase

 DigitalOcean

 NEON
COMMUNITY

Tip #3

Optimizing DB: Choose the Most Suitable Option



Tip #4

Optimizing DB: Adding Indexes



Tip #5

Optimizing DB: Low-level Caching of Queries

The screenshot shows the Glassdoor company profile for Grab. The top navigation bar includes the Grab logo, a 'Top Companies' badge, and buttons for 'Follow' and '+ Contribute'. It also features a message for employees: 'You work here? Claim this page for FREE!'. Below the header, there are tabs for 'Overview', 'Salaries 1073', 'Reviews 556', 'Jobs 235', and 'Interviews 791'. The 'Salaries' tab is currently selected.

1,073 Grab Salaries

Filter: Singapore

8 Job Openings!

Product Manager

Level	Salary Range
Intern	\$2,500 ✓
Junior	\$7,143 ✓
Mid	\$8,244 ✓
Senior	\$10,390 ✓
Lead	\$13,497 ✓
Manager	\$16,700 ✓
Principal	\$18,125 ✓

Grab Interview Questions 791

Technical Jul 6, 2024
Data Scientist
How do you estimate the size of the sample in AB testing?

Background Jul 5, 2024
Software Engineer
Can you describe your experience in the field?

Tip #5

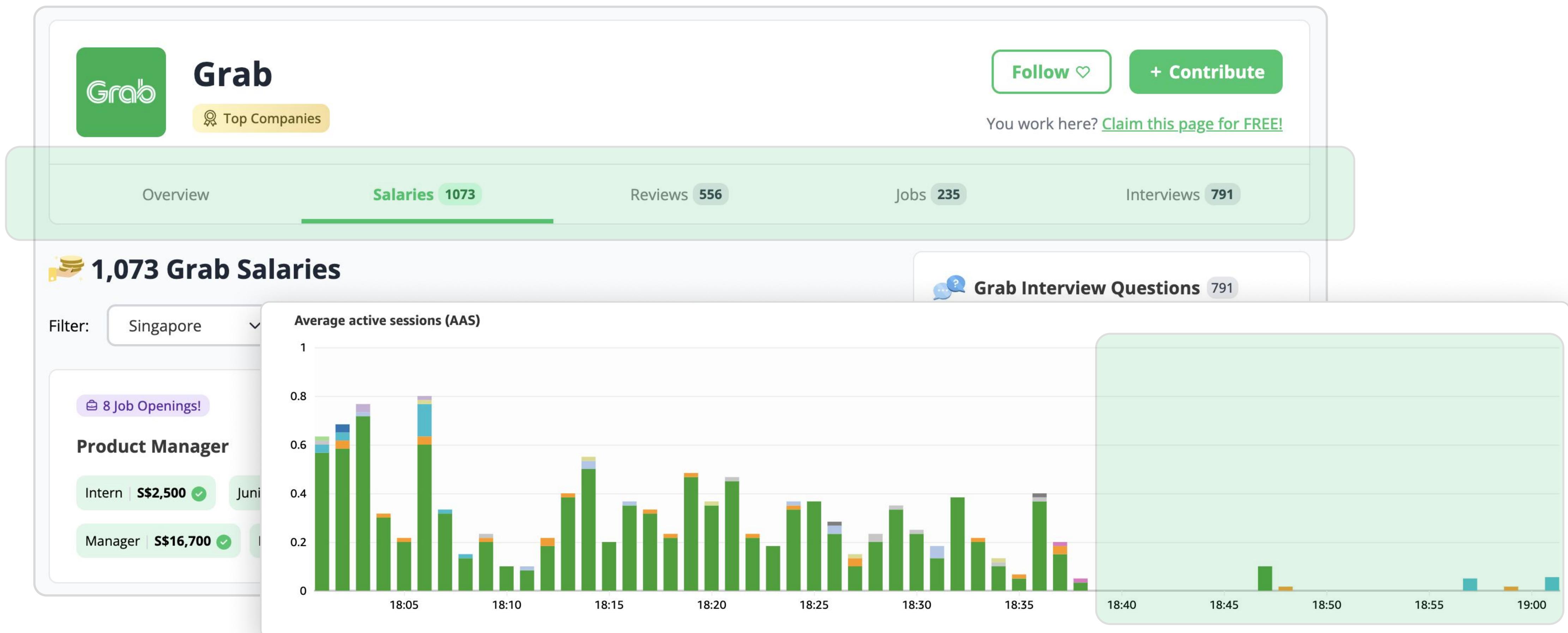
Optimizing DB: Low-level Caching of Queries

The screenshot shows a company profile page for Grab on Glassdoor. At the top, there's a green Grab logo, the word "Grab", a "Top Companies" badge, and buttons for "Follow" and "+ Contribute". Below the header, there are five main navigation links: "Overview", "Salaries 1073" (which is underlined in green), "Reviews 556", "Jobs 235", and "Interviews 791". A callout box highlights the "Interviews" link. The main content area displays "1,073 Grab Salaries" with a filter dropdown set to "Singapore". It shows salary ranges for various job levels: Intern (\$\$2,500), Junior (\$\$7,143), Mid (\$\$8,244), Senior (partially visible), Manager (\$\$16,700), and Principal (\$\$18,125). To the right, a box titled "Grab Interview Questions 791" is shown. A large green callout box on the right side of the page contains the following Ruby code:

```
Rails.cache.fetch(
  "CompanyNavbarCount-#{@company.id}",
  expires_in: 1.day
) do
  { salaries:, reviews:, jobs:, interviews: }
end
```

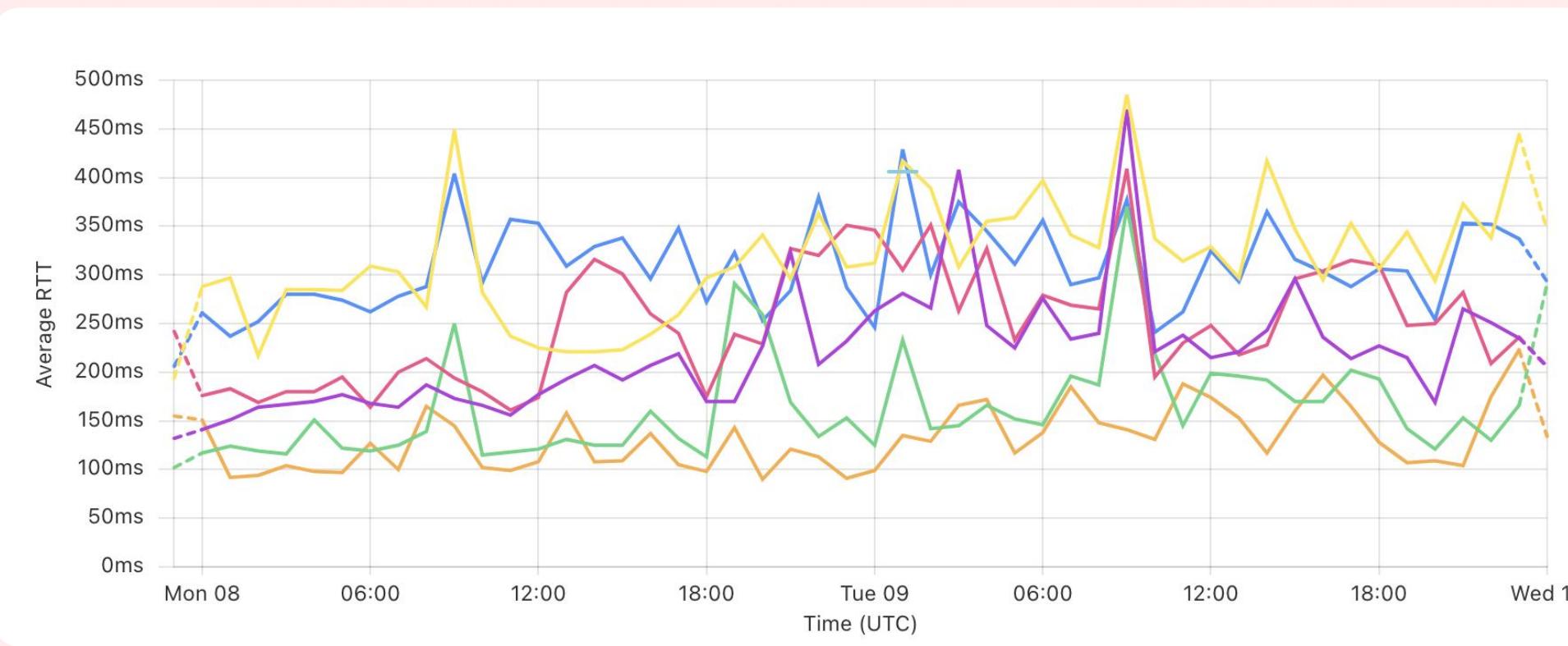
Tip #5

Optimizing DB: Low-level Caching of Queries

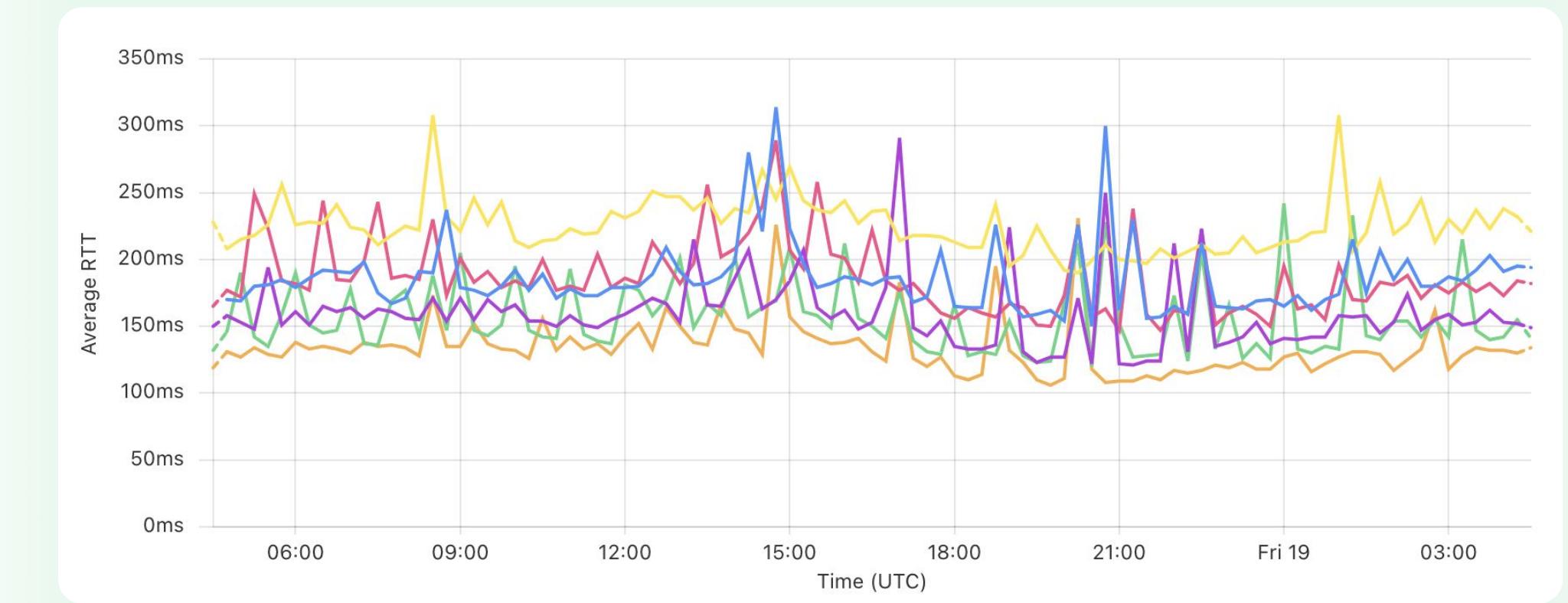


Tip #4, #5

Optimizing DB: Adding Index + Low-level Caching of Queries



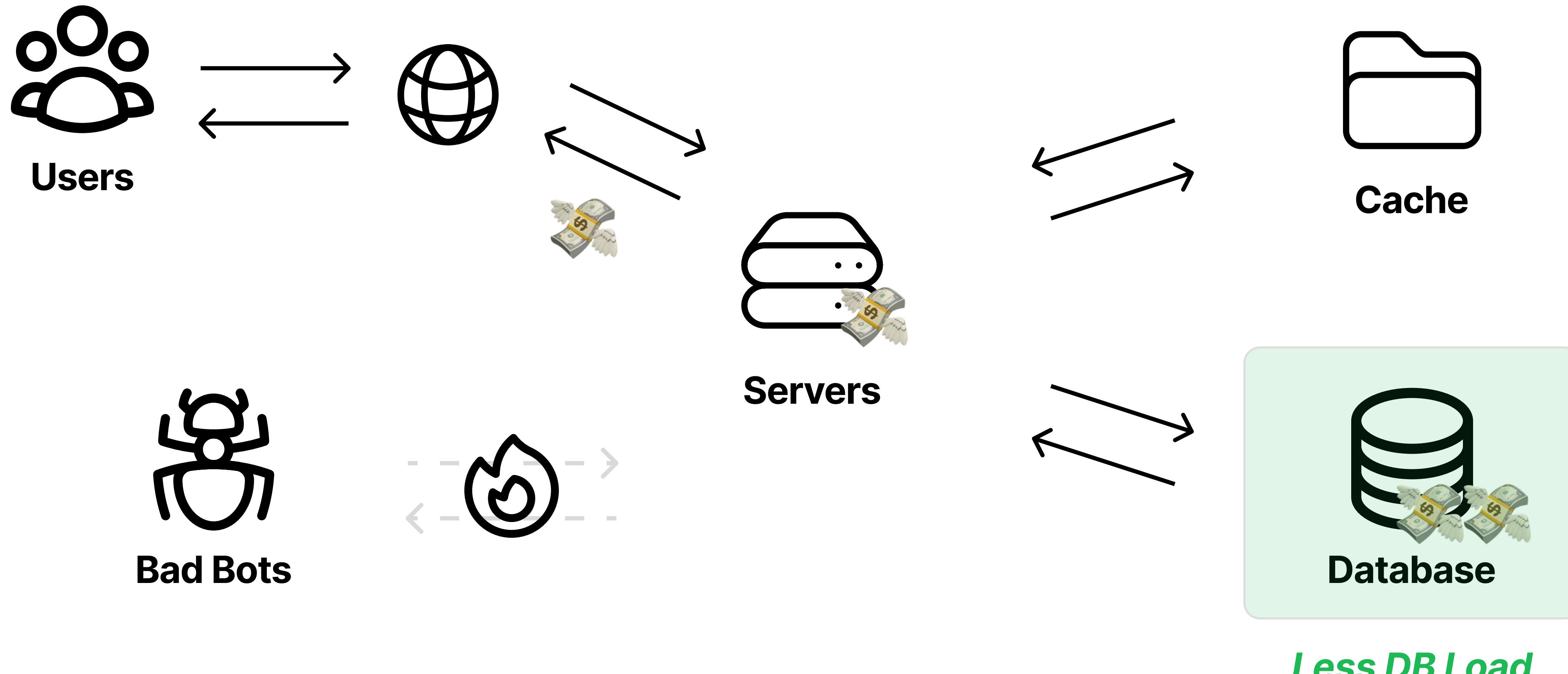
Before 100ms to 400ms



After 100ms to 250ms

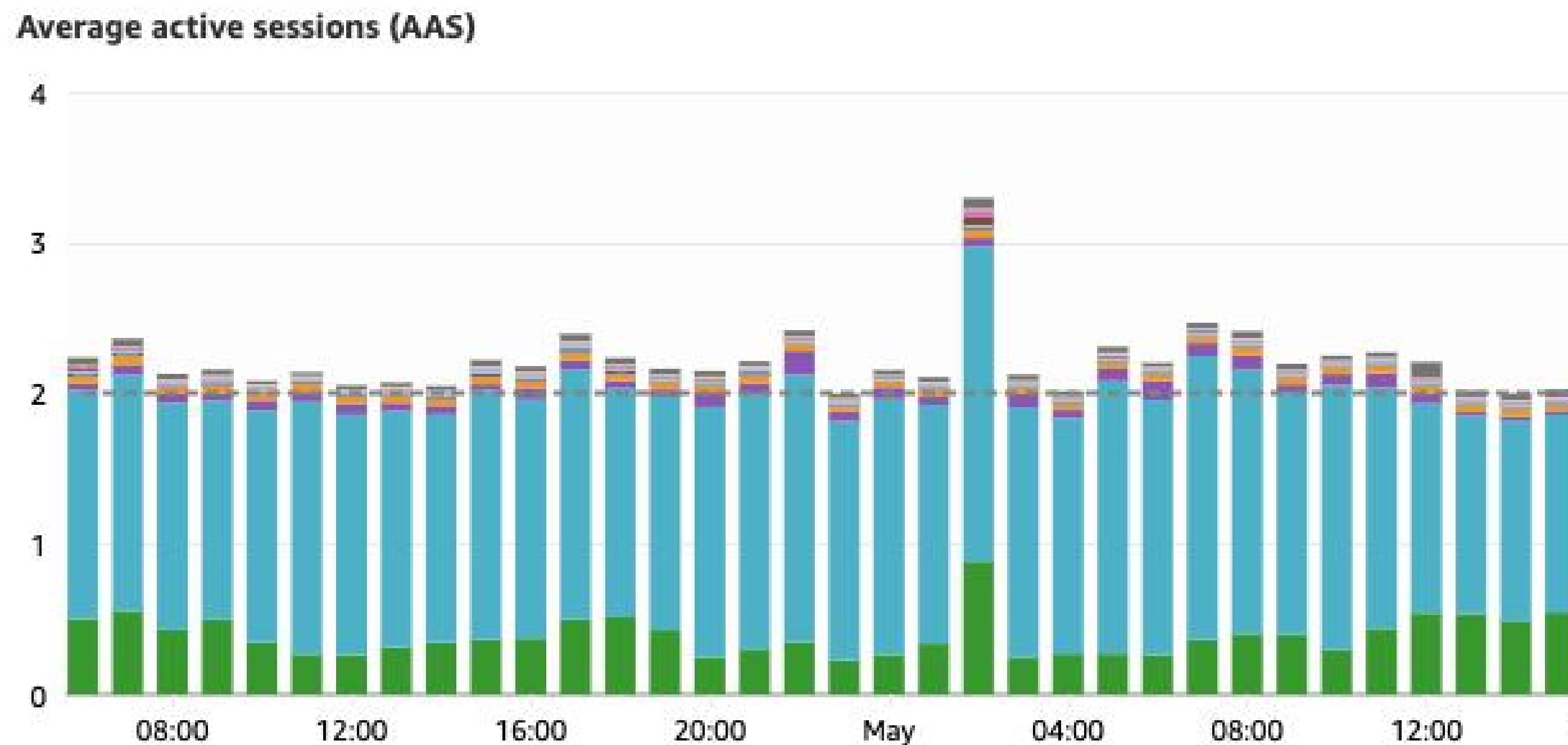
Tip #4, #5

Optimizing DB: Adding Index + Low-level Caching of Queries



Tip #6

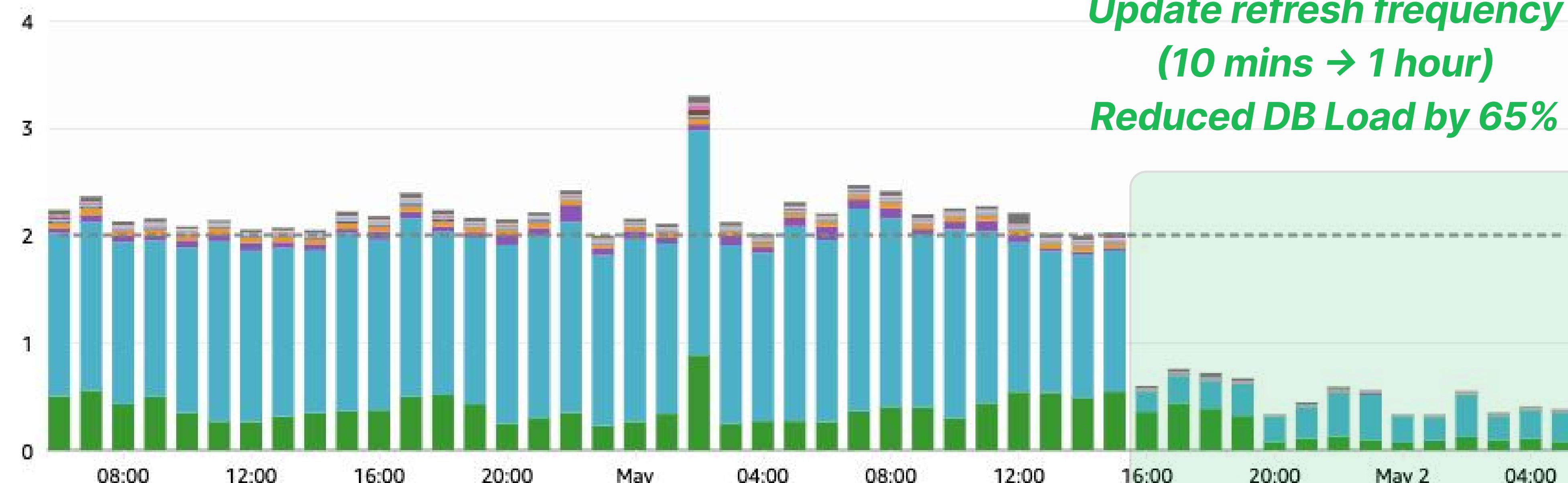
Optimizing DB: Materialized View



Tip #6

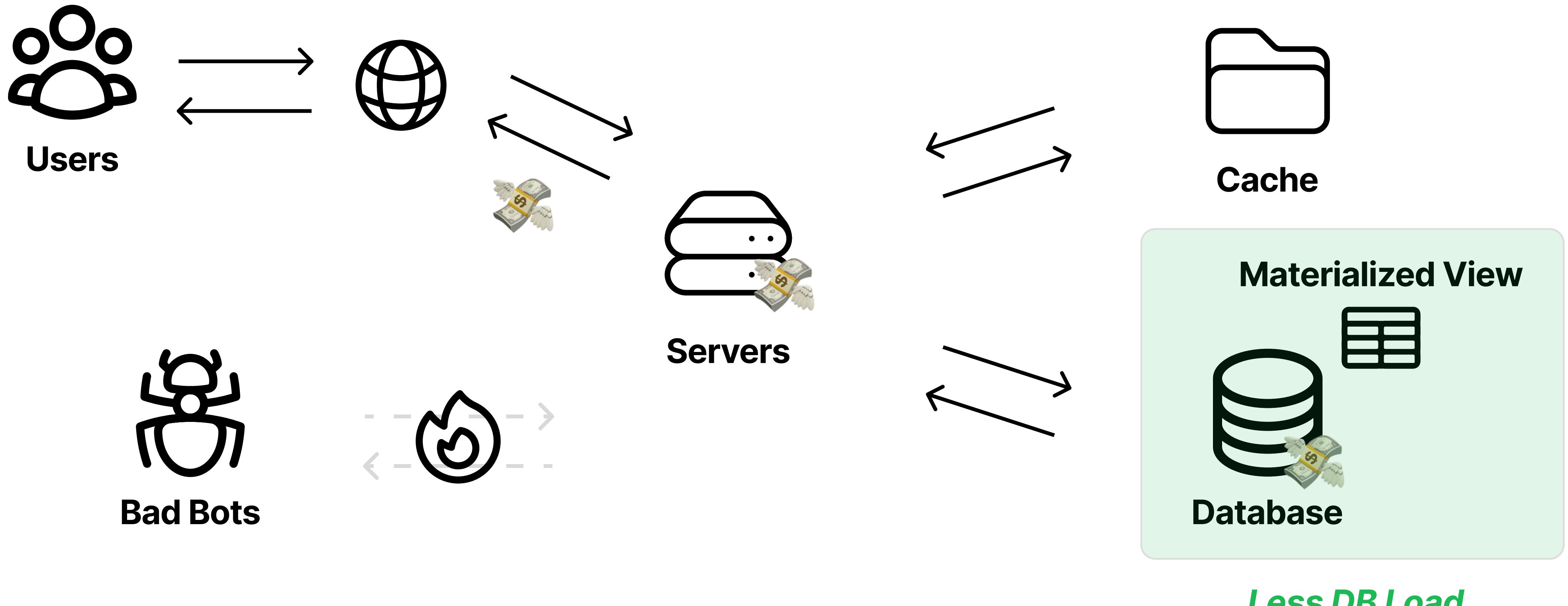
Optimizing DB: Materialized View

Average active sessions (AAS)

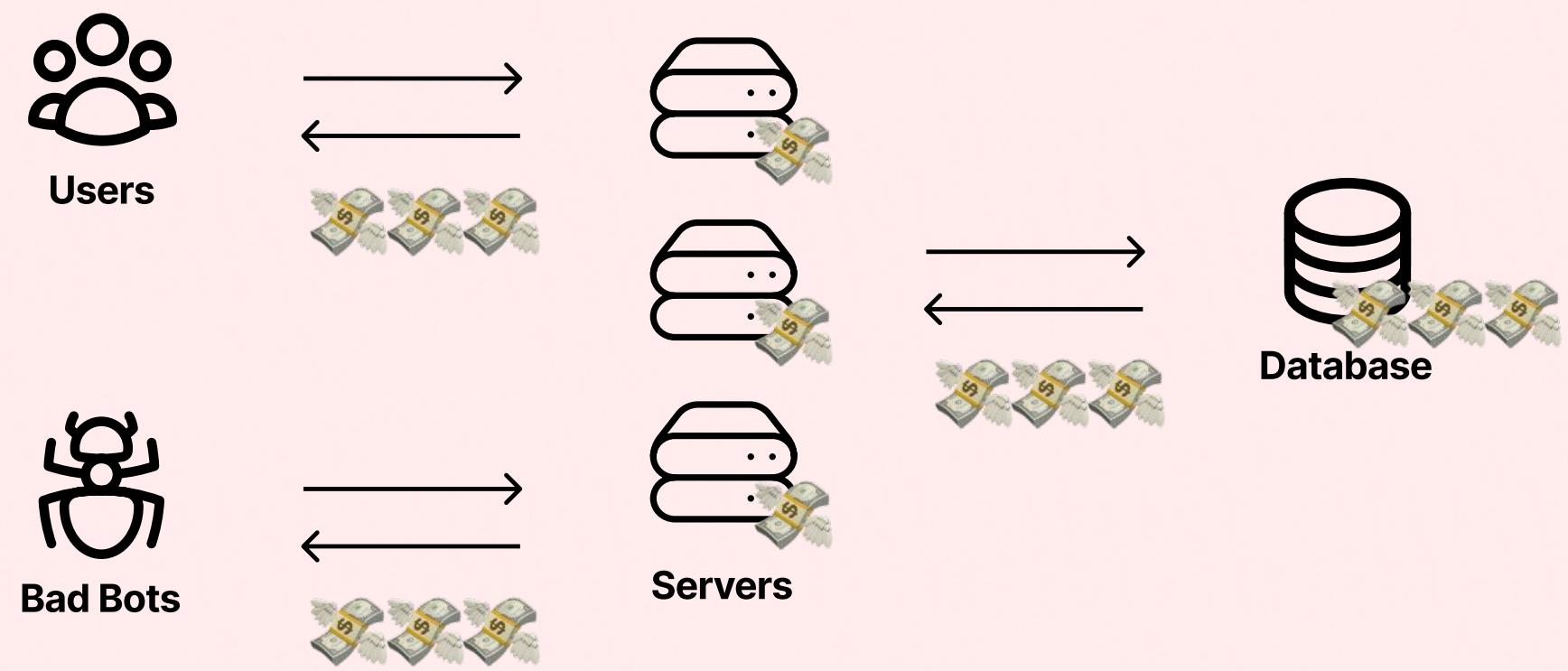


Tip #6

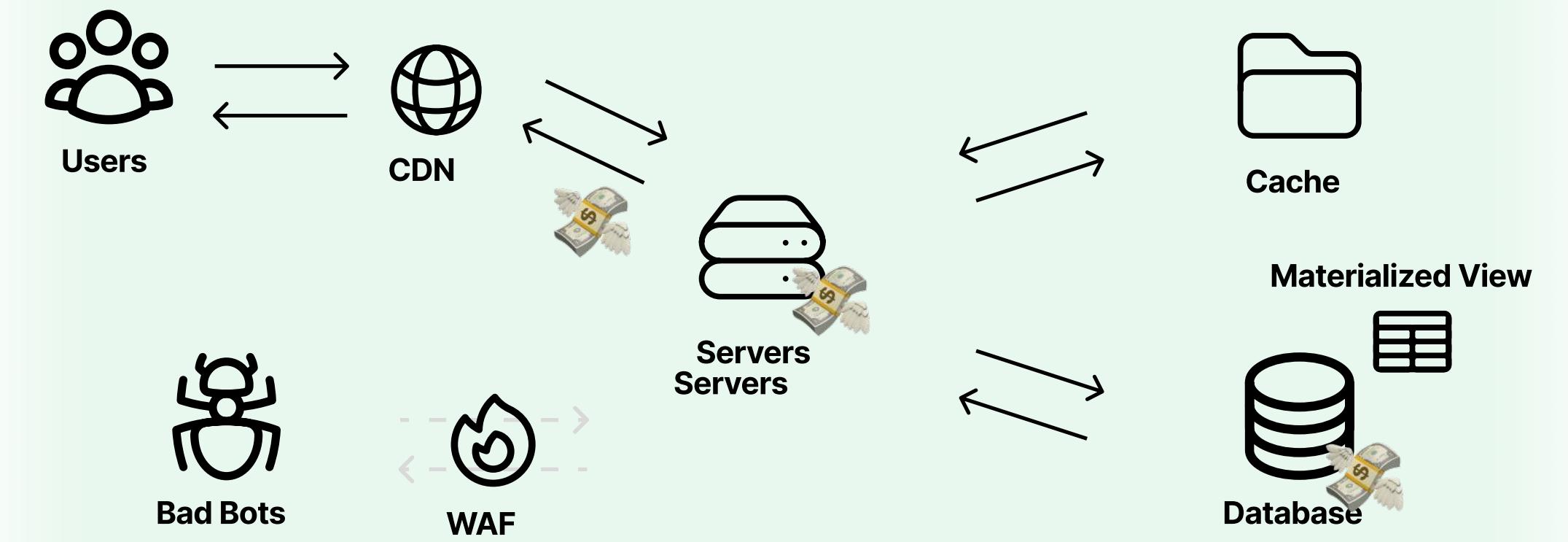
Optimizing DB: Materialized View



Cheaper + Faster



Before



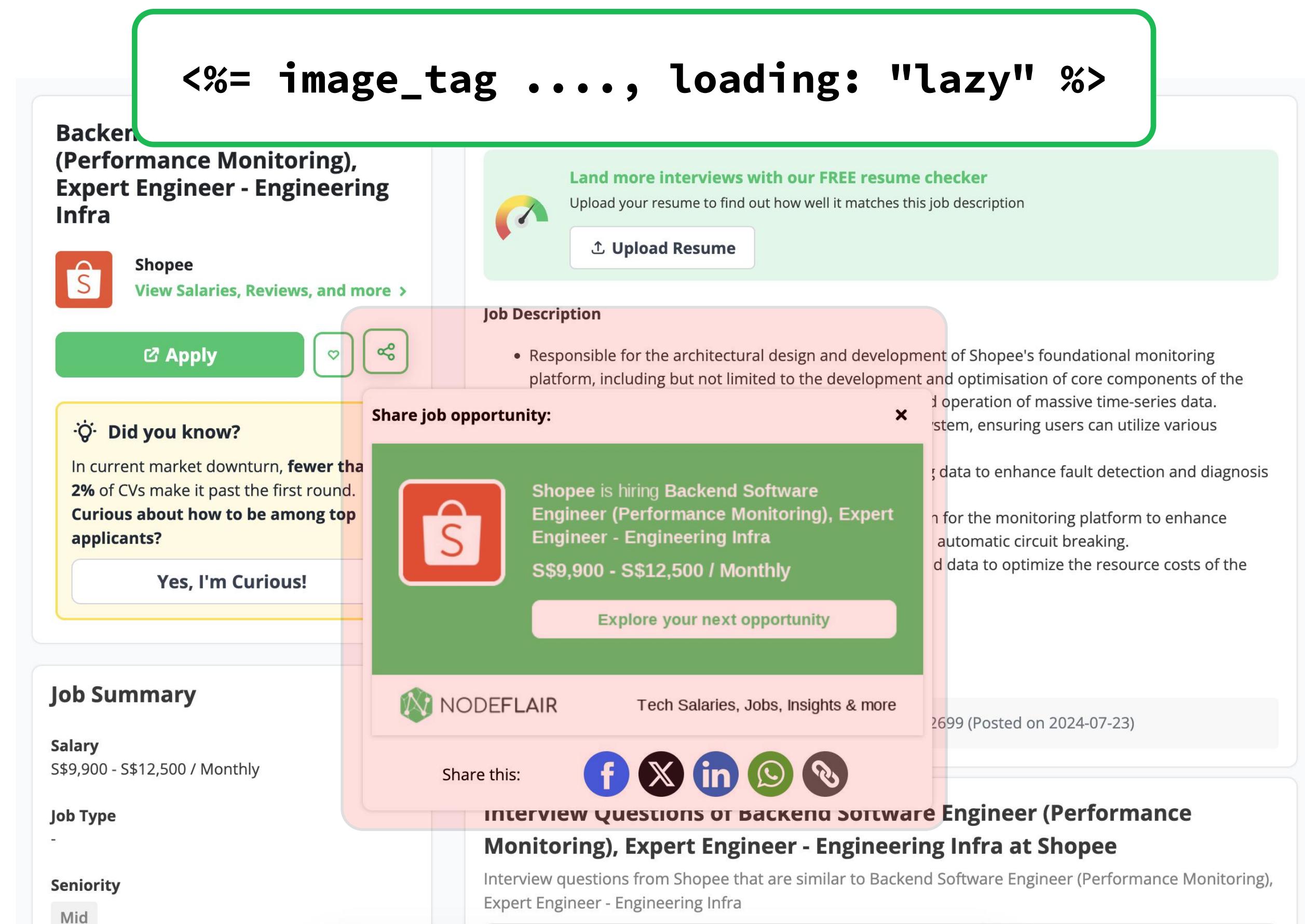
After

Optimizing Database

No, you (likely) don't need more vCPUs and RAM

Tip #7

Speed (Perceived): Lazy Loading



Tip #8

Speed (Actual): Reduce Blockers for Page Rendering

The screenshot shows a web page for Grab, specifically a section titled "791 Grab Interview Questions". The page is structured with several parallel rendering components:

- Top Bar:** Features the Grab logo and navigation links for "Overview", "Salaries 1073", and "Reviews 5".
- Profile Overlay:** A modal-like overlay for "nateberkopec" from 4 months ago, containing a message about retrofitting ActiveAdmin with turbo frames for speed.
- Interview Questions List:** A grid of three interview questions for Software Engineers, each with a "Generate Answer" button and a "Share" link.
 - Question 1: "Can you reverse a sentence of words?" (Technical, Feb 24, 2016)
 - Question 2: "Can you find the only non-duplicated number from an array?" (Technical, Feb 24, 2016)
 - Question 3: "Can you find the last nth node in a linked list?" (Technical, Feb 24, 2016)
- Jobs Section:** A sidebar listing "Jobs at Grab" with 236 results, including a "Senior Software Engineer, Backend" position requiring Kubernetes and Azure.
- Salaries Section:** A sidebar titled "Grab Salaries" showing salary data for "Software Engineer" roles, with a highlighted median salary of "\$10,060".

Many Web Pages (> 2 Million)

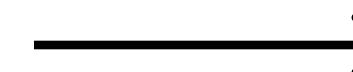
🤖 Many crawlers and bots - some useful, some not



DDoS Protection + WAF

Read-Heavy Application

💰 Egress is expensive!



CDN (Caching)

Migrate to Free-Egress Provider

B2C - Users Expect It To Be Fast

🏃 53% of mobile site visits abandoned if load time >3 seconds



Lazy Loading

Reduce Blockers for Page Rendering

Global Users

🌐 Only as fast (or slow) as your network latencies



CDN (Route Optimization)

Data Intensive

📁 NodeFlair Salaries have ~6 million data point



Materialized View

😔 Mistake

Setting High `RAILS_MAX_THREADS`



😔 Mistake

Setting High `RAILS_MAX_THREADS`



Free Gems

They worked great for us

Split

A/B Testing Framework

<https://github.com/splirb/split>

Scenic

Bringing Database View into ActiveRecord

<https://github.com/scenic-views/scenic>

Rails Performance

Simple and free APM alternative to New Relic, Datadog etc.

https://github.com/igorkasyanchuk/rails_performance

PgHero

Performance Dashboard for Postgres

<https://github.com/ankane/pghero>

Free Gems

1. Split (A/B Testing)

```
def set_landing_page_variation
  @landing_page_variation = ab_test(
    :landing_page,
    {'Old' => 9},
    {'New' => 1}
  )
end
```

Free Gems

2. Scenic (Database View)

```
rails generate scenic:view salaries
```

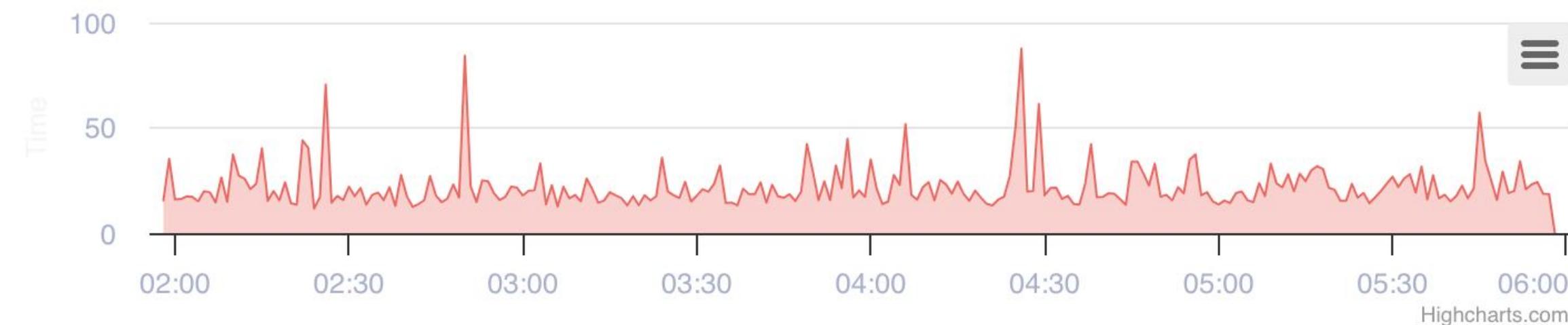
Free Gems

3. Rails Performance (APM)

Throughput



Average Response Time

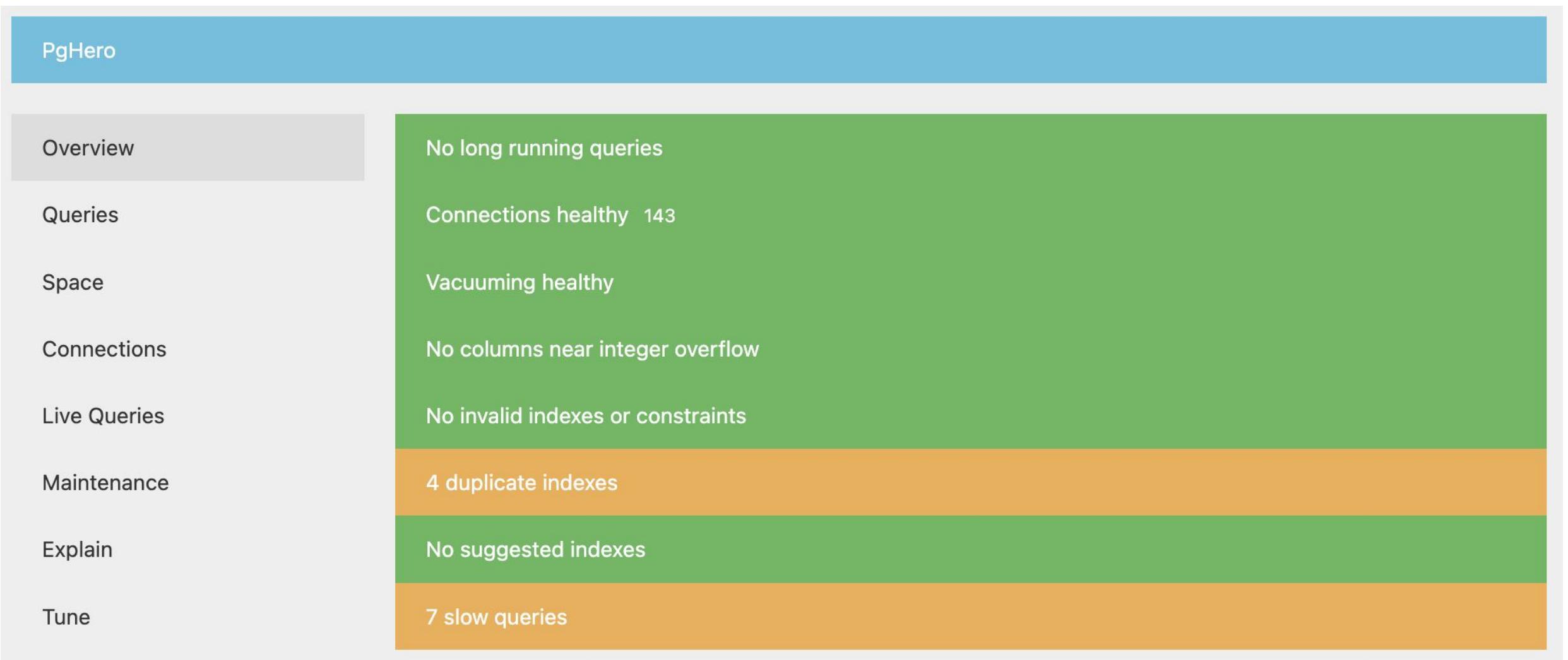


Requests

Datetime	Method	Path	Format	Status	Duration	Views	DB
2024-07-25 05:57:58	GET	/api/v2/interviews?page=1&company_names%5B%5D=Global...	json	200	16.4 ms	5.7 ms	8.5 ms
2024-07-25 05:57:54	GET	/api/v2/interviews?page=1&company_names%5B%5D=GovTec...	json	200	12.7 ms	4.0 ms	6.1 ms
2024-07-25 05:57:51	GET	/api/v2/interviews?page=1&company_names%5B%5D=+Newge...	json	200	15.9 ms	4.7 ms	8.5 ms
2024-07-25 05:57:45	GET	/api/v2/interviews?page=3&company_names%5B%5D=Accent...	json	200	48.8 ms	4.8 ms	42.1 ms
2024-07-25 05:57:44	GET	/api/v2/interviews?page=1&company_names%5B%5D=Shopee...	json	200	29.8 ms	3.2 ms	24.4 ms
2024-07-25 05:57:43	GET	/api/v2/interviews?page=1&company_names%5B%5D=Global...	json	200	15.1 ms	4.2 ms	8.8 ms
2024-07-25 05:57:33	GET	/api/v2/interviews?page=1&company_names%5B%5D=Micros...	json	200	30.9 ms	5.9 ms	23.0 ms

Free Gems

4. PgHero (Postgres Perf. Dashboard)



400,000

1 Engineer, ~~300,000~~ Users

@adriangohjw | Red Dot Ruby Conference 2024