GDG

Why You Need to Understand Database Structures for Optimal Performance

Speaker

Armando Ferrara

Event

Google Developer Group - Zürich 2023

We.
Have.
A.
Plan.

#DatabaseInternals

#WhyShouldIKnowThis?

#DatabaseInternals

The most dangerous zodiac signs

Source: New York Post

	Sign	Rank(PK)	Polarity	Modern Ruler
First Row	Virgo	1	Negative	Mercury
Second Row	Pisces	2	Negative	Neptune
Third Row	Gemini	3	Positive	Mercury
Fourth Row	Aquarius	4	Positive	Uranus
Fifth Row	Scorpio	5	Negative	Pluto
Sixth Row	Sagittarius	6	Positive	Jupiter
Seventh Row	Capricorn	7	Positive	Saturn

Learn more: nypost.com/2021/12/23/most-dangerous-zodiac-signs

Row-Oriented Database

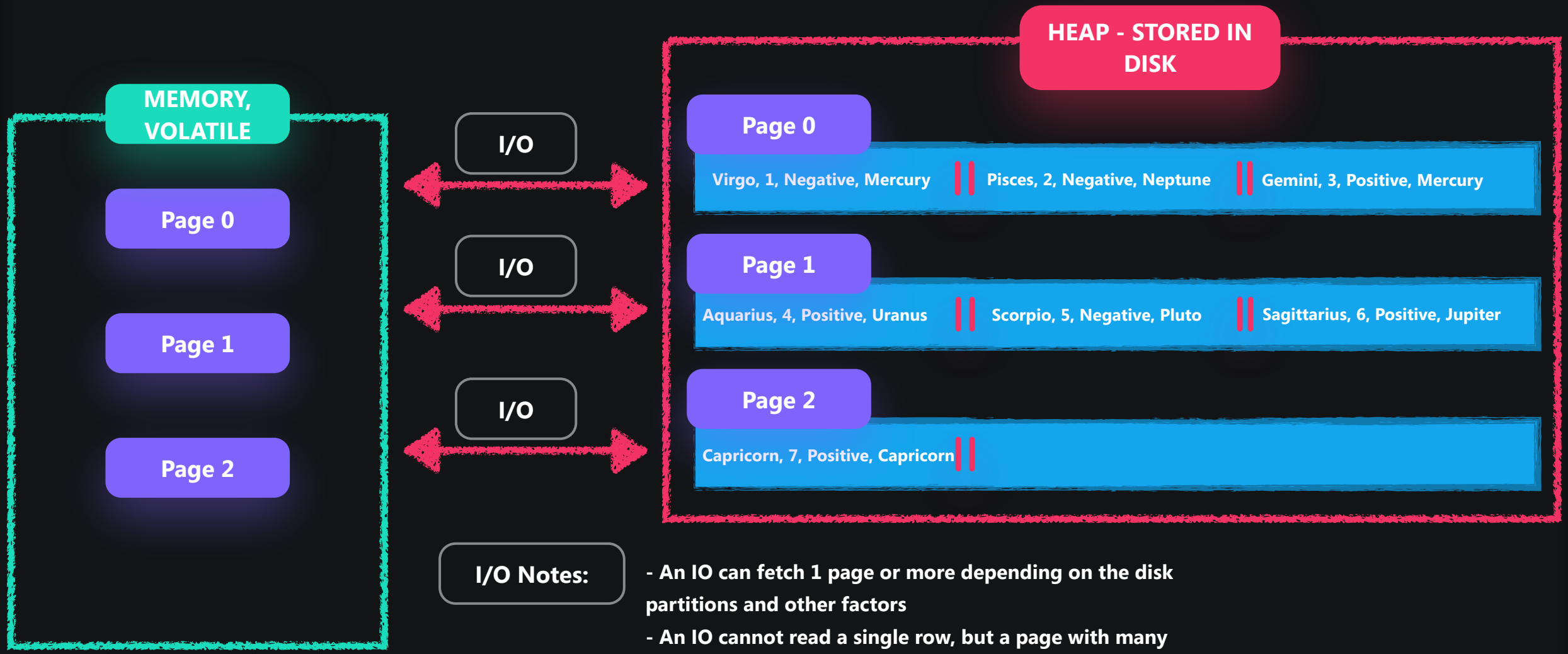
The database doesn't read a single row, it reads a page or more in a single IO and we get a lot of rows in that IO.

Each page has a size (e.g. 8KB in Postgres, 16KB in MySQL).

Assume each page holds 3 rows in this example, with 7 rows, we will have 3 rows (7/3)

The pages are stored one after another in a data structure called Heap.

No Index -
SELECT * FROM ZODIAC WHERE Sign="Capricorn";



I/O Notes:

- An IO can fetch 1 page or more depending on the disk partitions and other factors
- An IO cannot read a single row, but a page with many rows in them, you get more rows for free!

F-18 Hornet vs Banana slug

MEMORY = RAM = F-18 Hornet

RAM Latency = 83 nanoseconds
F-18 Hornet = 1.915 km/h

By scoutapm.com

DISK = Banana Slug

Disk Latency = 13 milliseconds
Banana Slug = 8 centimeters per minute

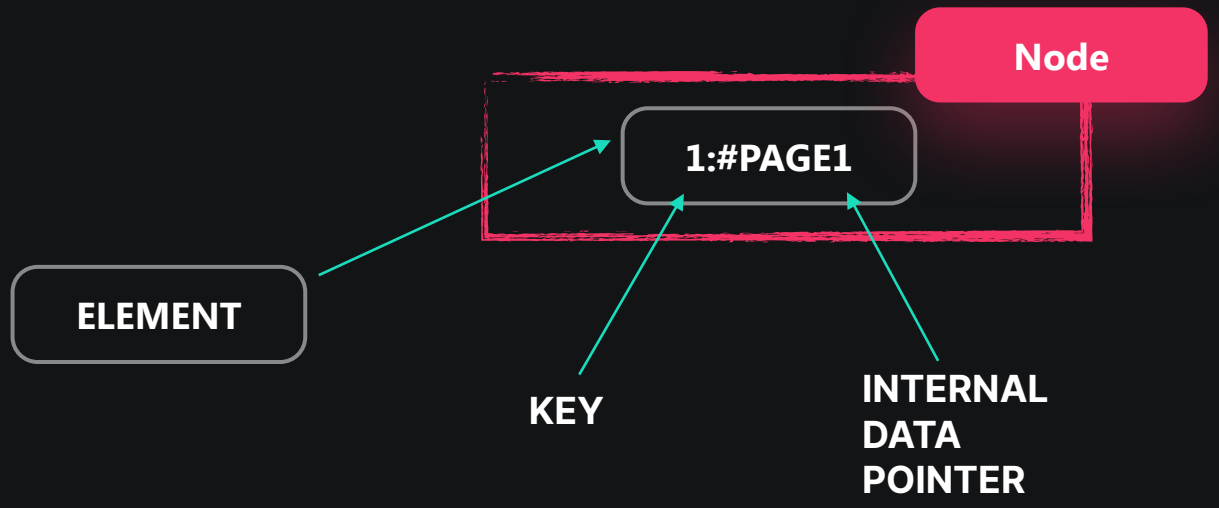
By scoutapm.com

B-TREE

Goal: minimise the searchable space, in this way I reduce the I/O to find my rows.

	Sign	Rank(PK)	Polarity	Modern Ruler
#PAGE1	Virgo	1	Negative	Mercury
#PAGE1	Pisces	2	Negative	Neptune
#PAGE1	Gemini	3	Positive	Mercury
#PAGE2	Aquarius	4	Positive	Uranus
#PAGE2	Scorpio	5	Negative	Pluto
#PAGE2	Sagittarius	6	Positive	Jupiter
#PAGE3	Capricorn	7	Positive	Saturn

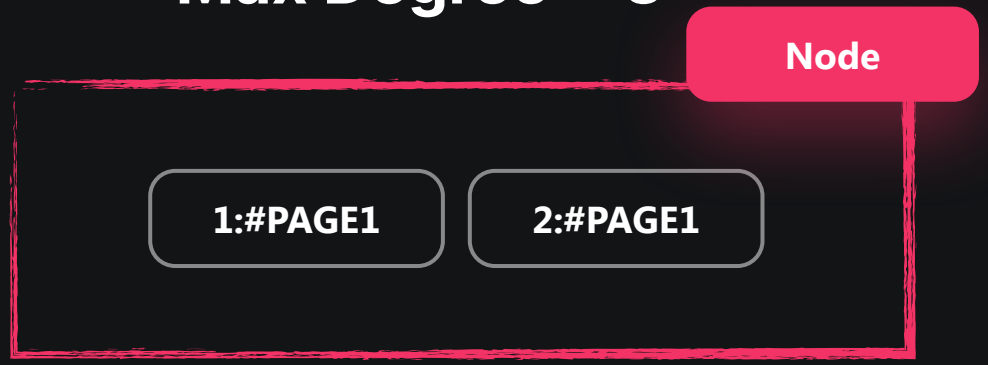
INSERT



Max Degree = 3
Max Number Elements Inside a Node (3-1)
Max Number Children (3)

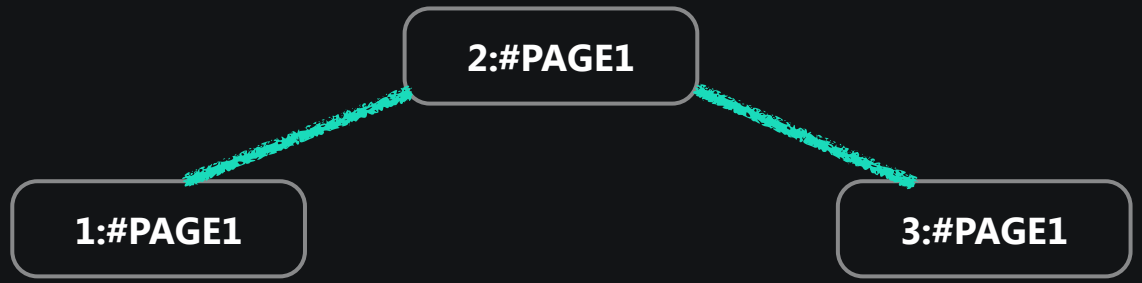
INSERT

Max Degree = 3



INSERT

Max Degree = 3



INSERT

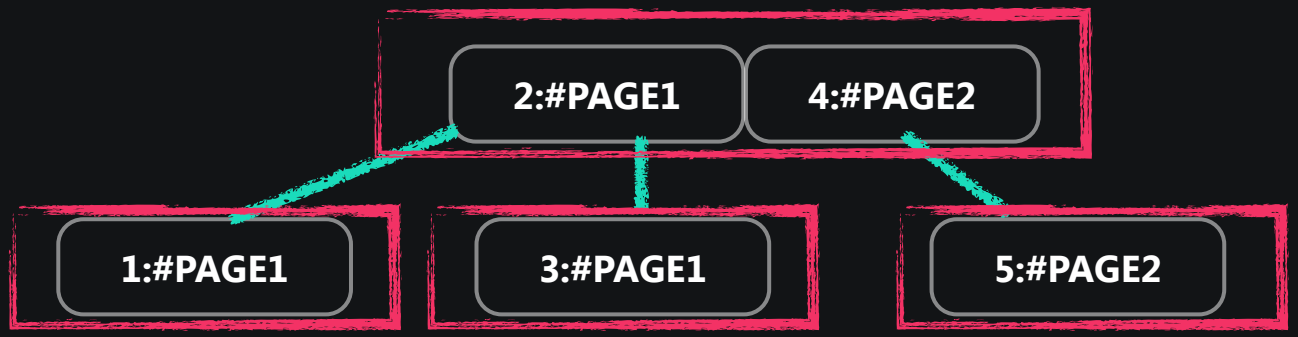
Max Degree = 3

INSERT

Max Degree = 3

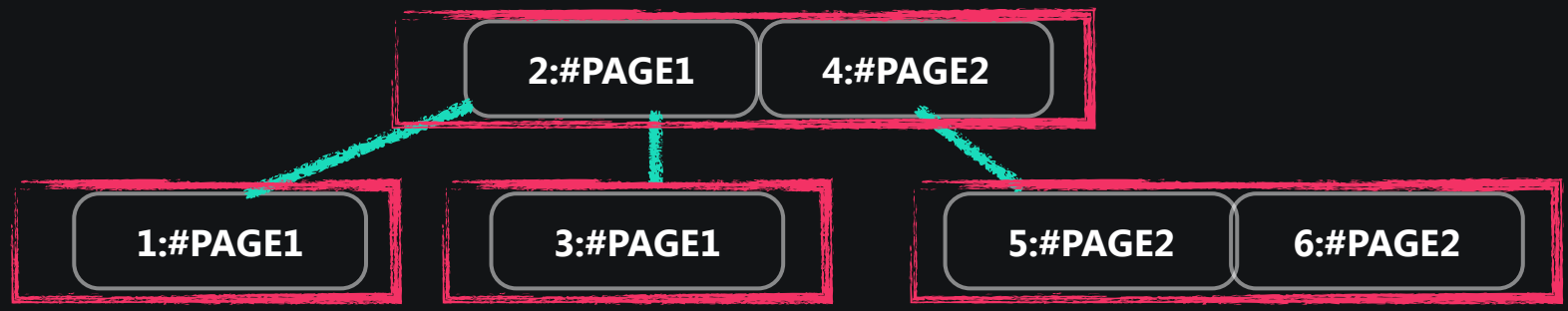
INSERT

Max Degree = 3



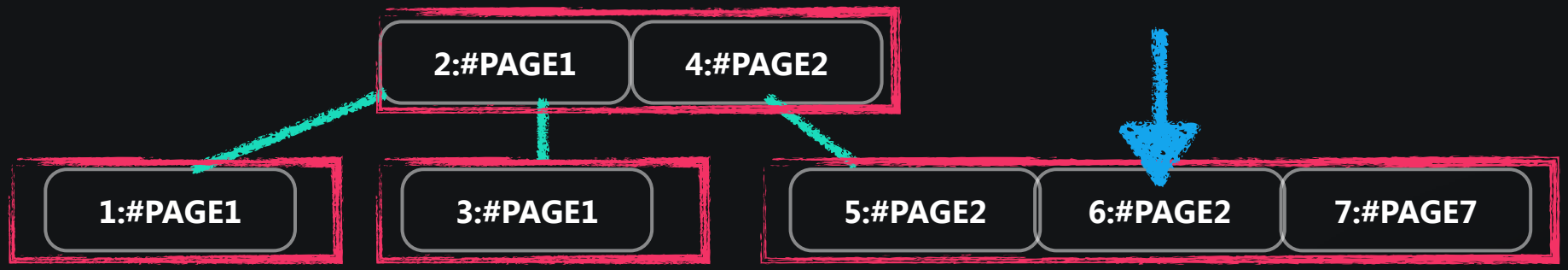
INSERT

Max Degree = 3



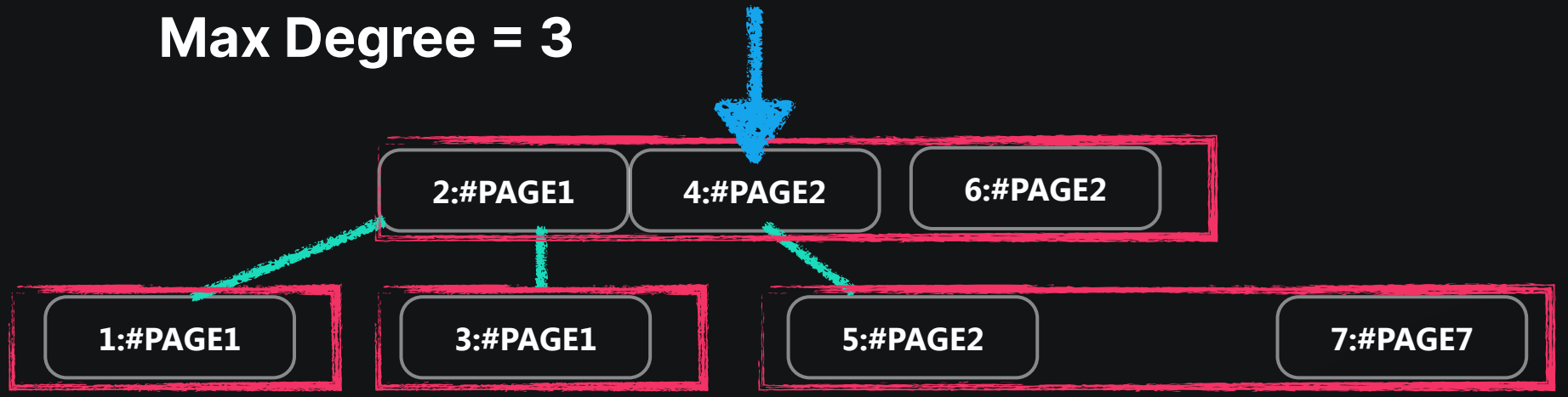
INSERT

Max Degree = 3



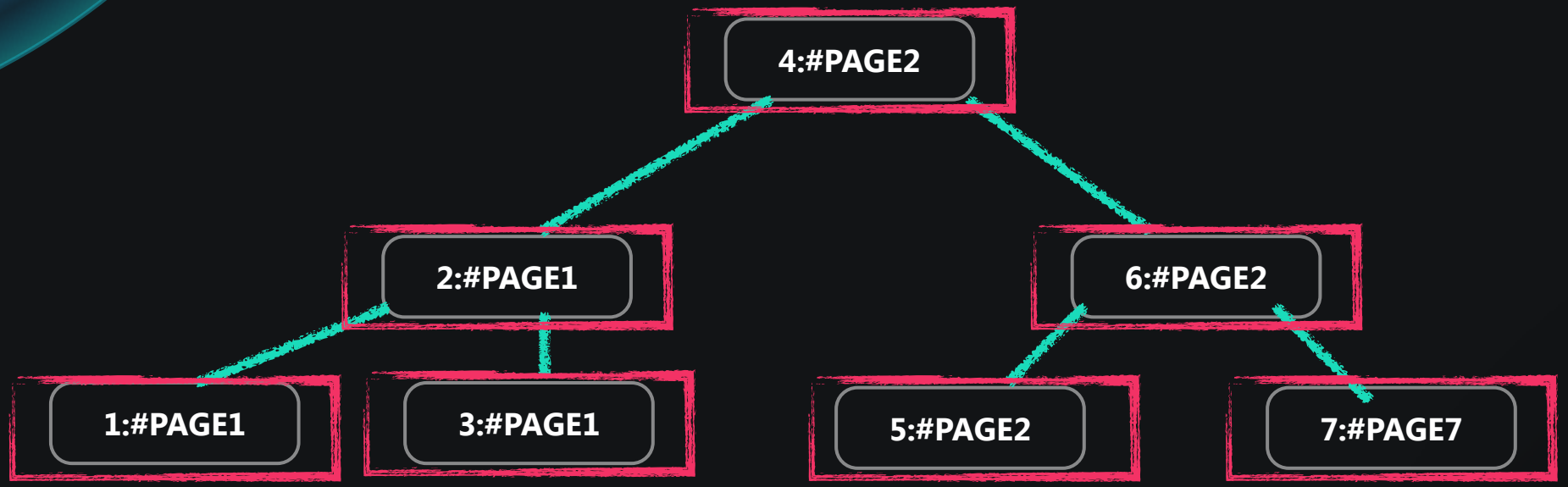
INSERT

Max Degree = 3



INSERT

Max Degree = 3



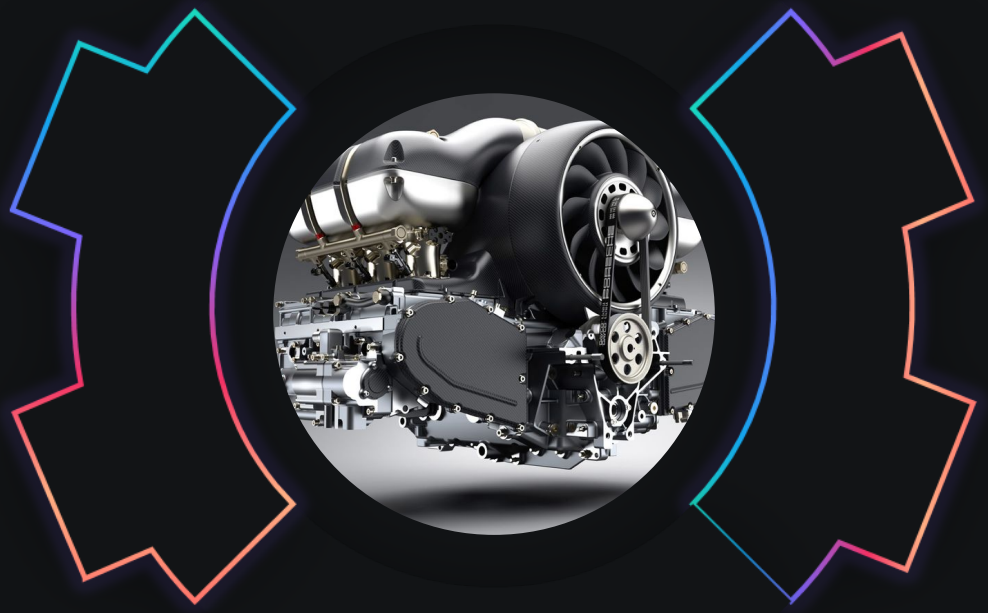
FIND RANK 5

Max Degree = 3

#WhyShouldIKnowThis?

DATABASE ENGINE

- Library that takes care of the on disk storage and CRUD. Can be simple or complex as full support for transactions.
- DBMS can use the database engine and build features on top (server replication, isolation …)
- Want to write a new db? Don’t start from scratch, use an engine.

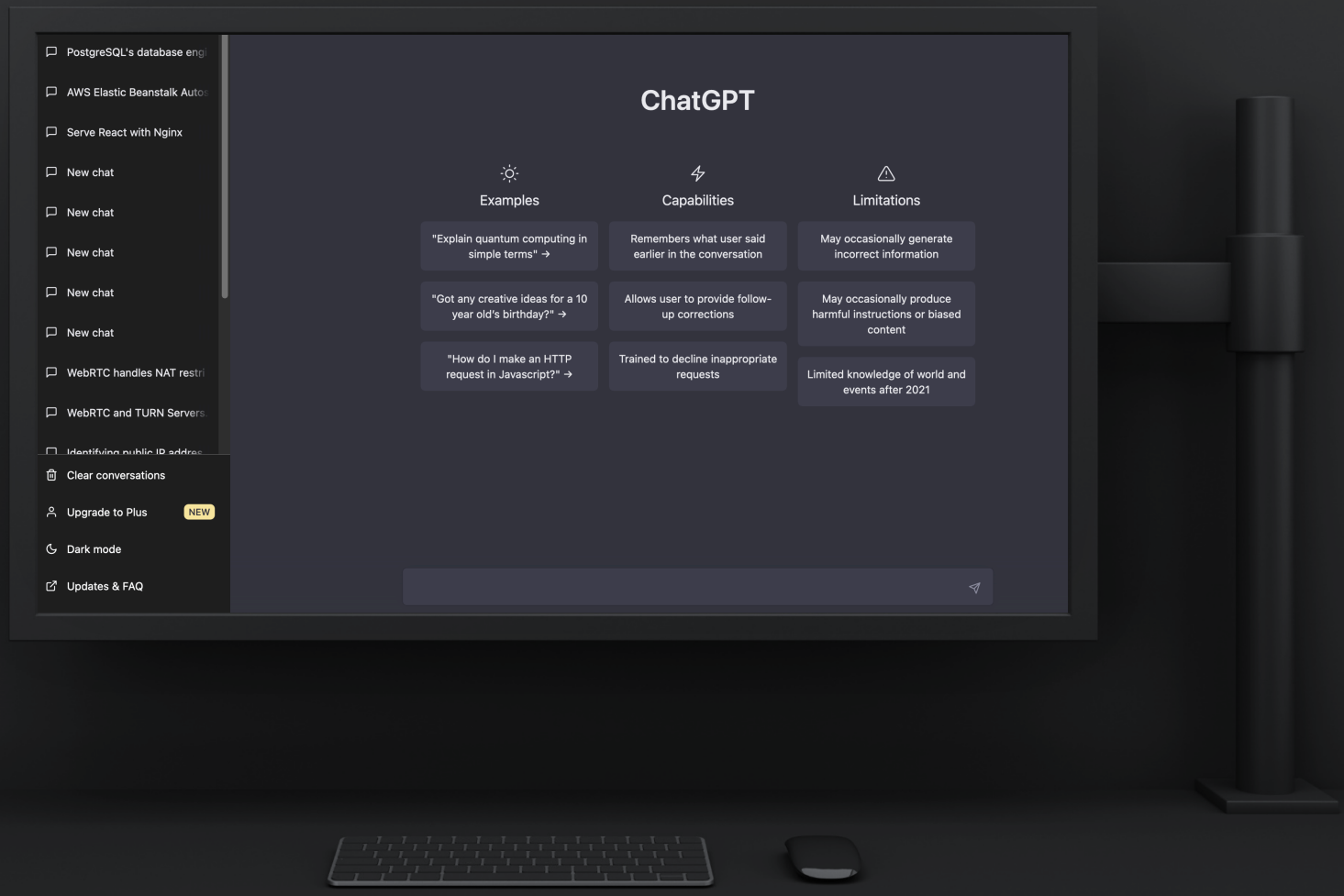


E.G.

- InnoDB, uses B+Tree, adopted by MySQL, OrackeDB
- SQLite, uses B-tree (LSM as extension)
- RocksDB
- LevelDB

…

Case Study


PostgreSQL's database engi
AWS Elastic Beanstalk Auto
Serve React with Nginx
New chat
New chat
New chat
New chat
New chat
WebRTC handles NAT restri
WebRTC and TURN Servers
Clear conversations
Upgrade to Plus
NEW
Dark mode
Updates & FAQ
ChatGPT
Examples
"Explain quantum computing in simple terms" →
"Got any creative ideas for a 10 year old's birthday?" →
"How do I make an HTTP request in Javascript?" →
Capabilities
Remembers what user said earlier in the conversation
Allows user to provide follow-up corrections
Trained to decline inappropriate requests
Limitations
May occasionally generate incorrect information
May occasionally produce harmful instructions or biased content
Limited knowledge of world and events after 2021

Create Conversation Insert Sta
MySQL to PostgreSQL Conver
MySQL Editor Syntax Error.
"Create ChatGTP schema"
MySQL UUID Function Conver
New chat
Btree database page splitting
New chat
OpenAI chat data storage.
PostgreSQL's database engin
AWS Elastic Beanstalk Autosca
Clear conversations
Upgrade to Plus NEW
Dark mode

The speed difference between SSDs and HDDs can vary depending on several factors, such as the specific models of the drives and the types of tasks being performed. However, in general, SSDs are significantly faster than HDDs, with read and write speeds that can be up to 5 to 10 times faster or more.

For example, a typical SATA III SSD can achieve read speeds of up to 550 MB/s and write speeds of up to 520 MB/s, while a typical 7200 RPM HDD may only achieve read speeds of around 100-200 MB/s and write speeds of around 80-160 MB/s. This means that SSDs can be up to 3-5 times faster than HDDs in terms of read and write speeds.

It's important to note that the actual speed difference between SSDs and HDDs will depend on the specific models and configurations being used, as well as the types of tasks being performed. However, in general, SSDs provide a significant performance advantage over HDDs when it comes to read and write speeds.

Regenerate response

https://chat.openai.com/chat/ef408033-082c-44a7-bb60-339dd506b993

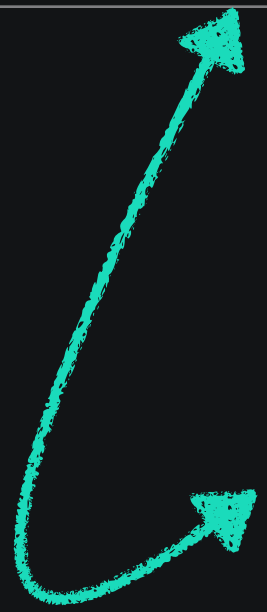
Yes, UUIDv4


```
"e7298e07-e65a-4c10-b0d0-dc5e6d047442": {
  "id": "e7298e07-e65a-4c10-b0d0-dc5e6d047442",
  "message": {
    "id": "e7298e07-e65a-4c10-b0d0-dc5e6d047442",
    "author": {
      "role": "user",
      "metadata": {}
    },
    "create_time": 1679755657.737997,
    "content": {
      "content_type": "text",
      "parts": ["example of how declaring my own nginx"]
    },
    "weight": 1.0,
    "metadata": {
      "timestamp_": "absolute"
    },
    "recipient": "all"
  },
  "parent": "fd92b53a-b98d-4783-a82e-3c4997a6bcf4",
  "children": ["fea1df3d-0086-4f78-b8dc-b3f5f14d01db"]
},
"fea1df3d-0086-4f78-b8dc-b3f5f14d01db": {
  "id": "fea1df3d-0086-4f78-b8dc-b3f5f14d01db",
  "message": {
    "id": "fea1df3d-0086-4f78-b8dc-b3f5f14d01db",
    "author": {
      "role": "assistant",
      "metadata": {}
    },
    "create_time": 1679755690.591684,
    "content": {
      "content_type": "text",
      "parts": ["Sure, here's an example of how you could declare your own
```

Conversations Table:

Column Name	Data Type	Description
id	Integer	Unique identifier for the conversation
user_id	Integer	Foreign key to the Users table
started_at	DateTime	Timestamp for when the conversation started
ended_at	DateTime	Timestamp for when the conversation ended
status	String	Status of the conversation (e.g. ongoing, completed, cancelled)

What would you choose as an id for the conversation table?



Messages Table:

Column Name	Data Type	Description
id	Integer	Unique identifier for the message
conversation_id	Integer	Foreign key to the Conversations table
user_id	Integer	Foreign key to the Users table
created_at	DateTime	Timestamp for when the message was sent
message	Text	Content of the message

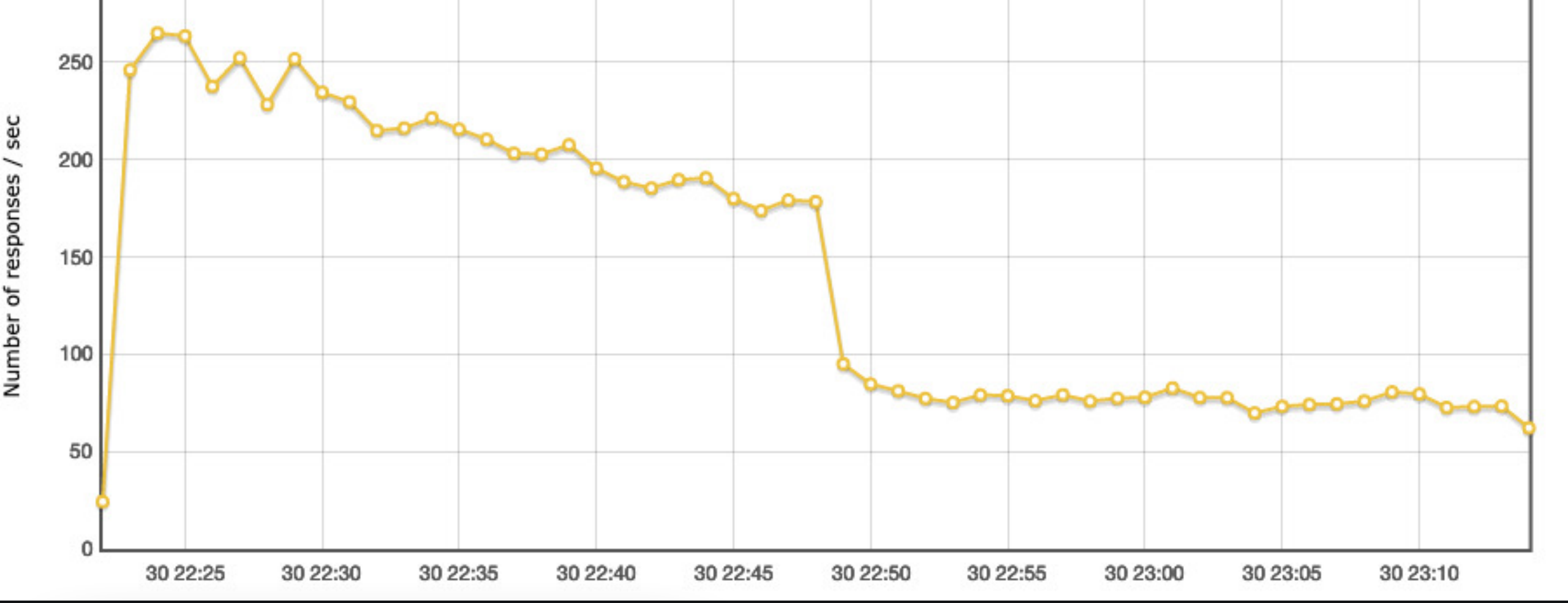
INSTANCE ON GOOGLE CLOUD

If you want to replicate the benchmarks:

https://github.com/Armando1514/Google-Developer-Groups-Zurich-Talk-2023

Summary

Region	europe-west6 (Zurich)
DB Version	MySQL 8.0
vCPUs	1 vCPU
Memory	628.74 MB
Storage	100 GB
Network throughput (MB/s)	125 of 125
Disk throughput (MB/s)	Read: 12.0 of 125.0 Write: 12.0 of 37.9
IOPS	Read: 75 of 3,000 Write: 150 of 15,000
Connections	Public IP
Backup	Manual
Availability	Single zone
Point-in-time recovery	Disabled



InnoDB (MySQL, OracleDB)

INSERTS WITH UUIDv4

https://ferrara.link/

Clustered Index on Disk

INSERT ROW WITH ID 8

INSERT ROW WITH ID 9

Sequential Writes

MEMORY, VOLATILE

ROW ID: 7
ROW ID: 8
ROW ID: 9

PAGE2

READ

WRITE

PAGE2

HEAP - STORED IN DISK

Page 0 - CONTAINS IDs FROM 1 to 3

Virgo, 1, Negative, Mercury || Pisces, 2, Negative, Neptune || Gemini, 3, Positive, Mercury

Page 1 - CONTAINS IDs FROM 4 to 6

Aquarius, 4, Positive, Uranus || Scorpio, 5, Negative, Pluto || Sagittarius, 6, Positive, Jupiter

Page 2 - CONTAINS IDs FROM 7 to 9

Capricorn, 7, Positive, Saturn ||

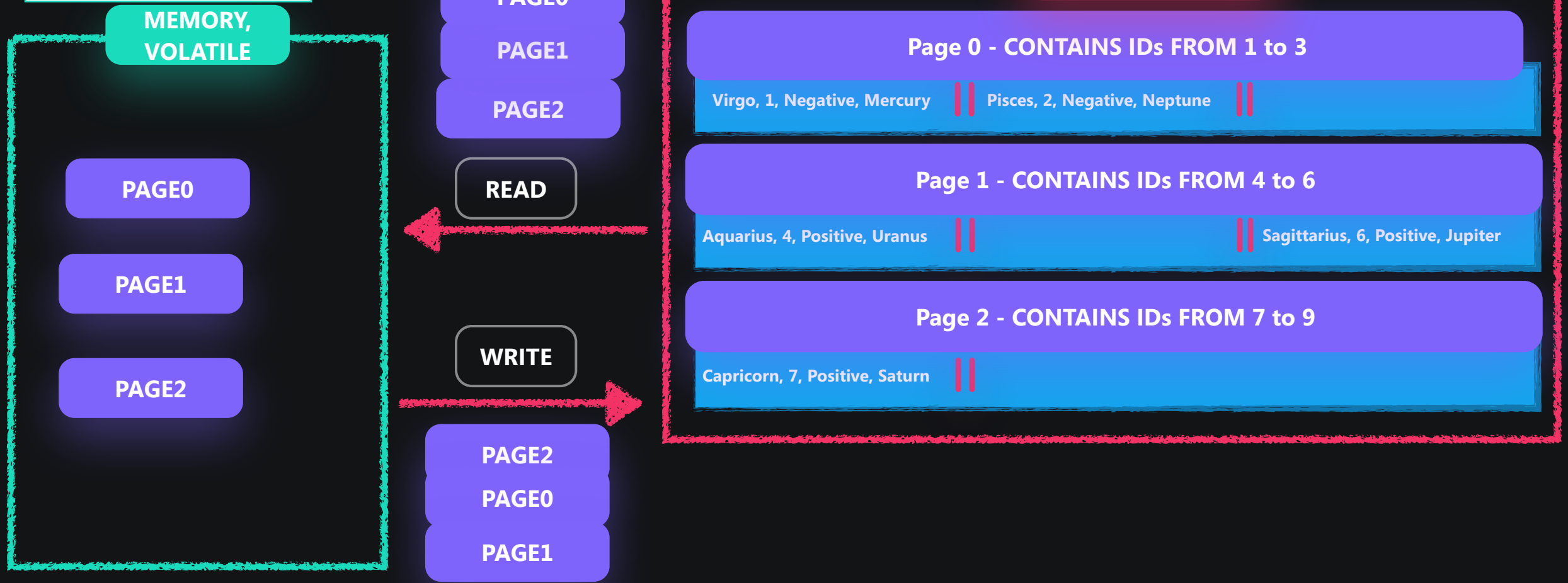
Clustered Index on Disk

RANDOM ID - NO HOT PAGES

INSERT ROW WITH ID 5

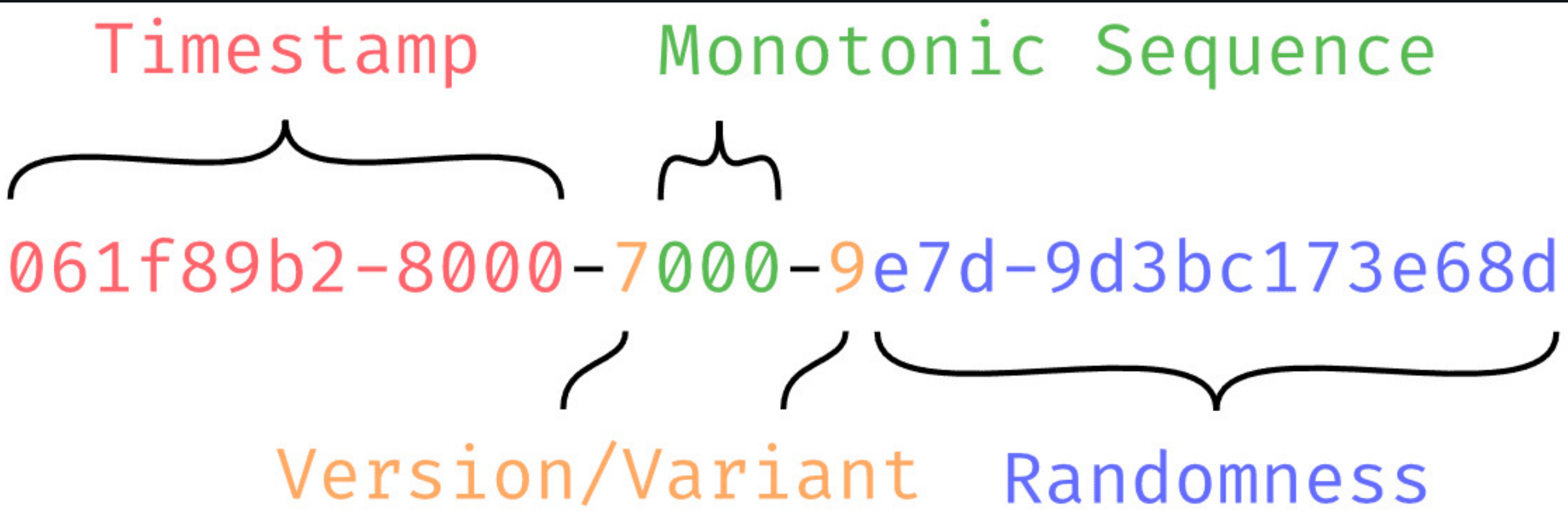
INSERT ROW WITH ID 3

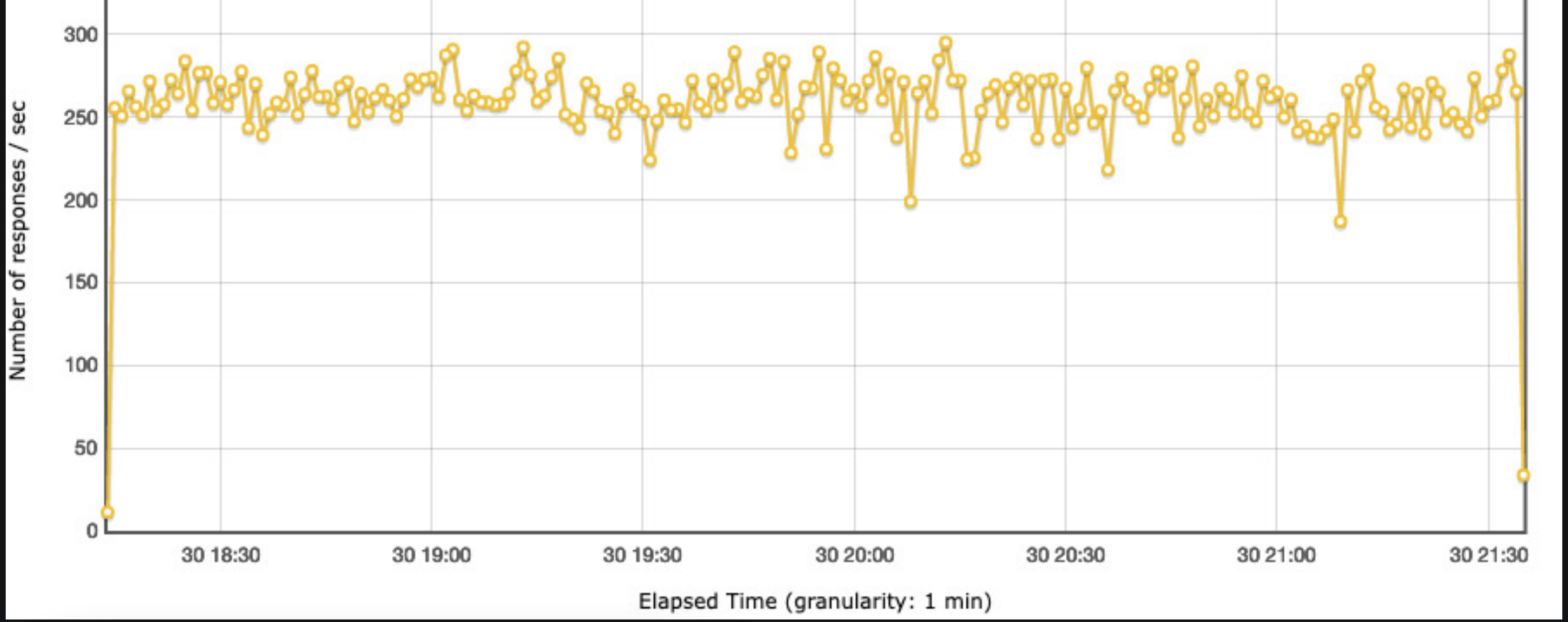
INSERT ROW WITH ID 9



~~UUIDv4~~

UUIDv7

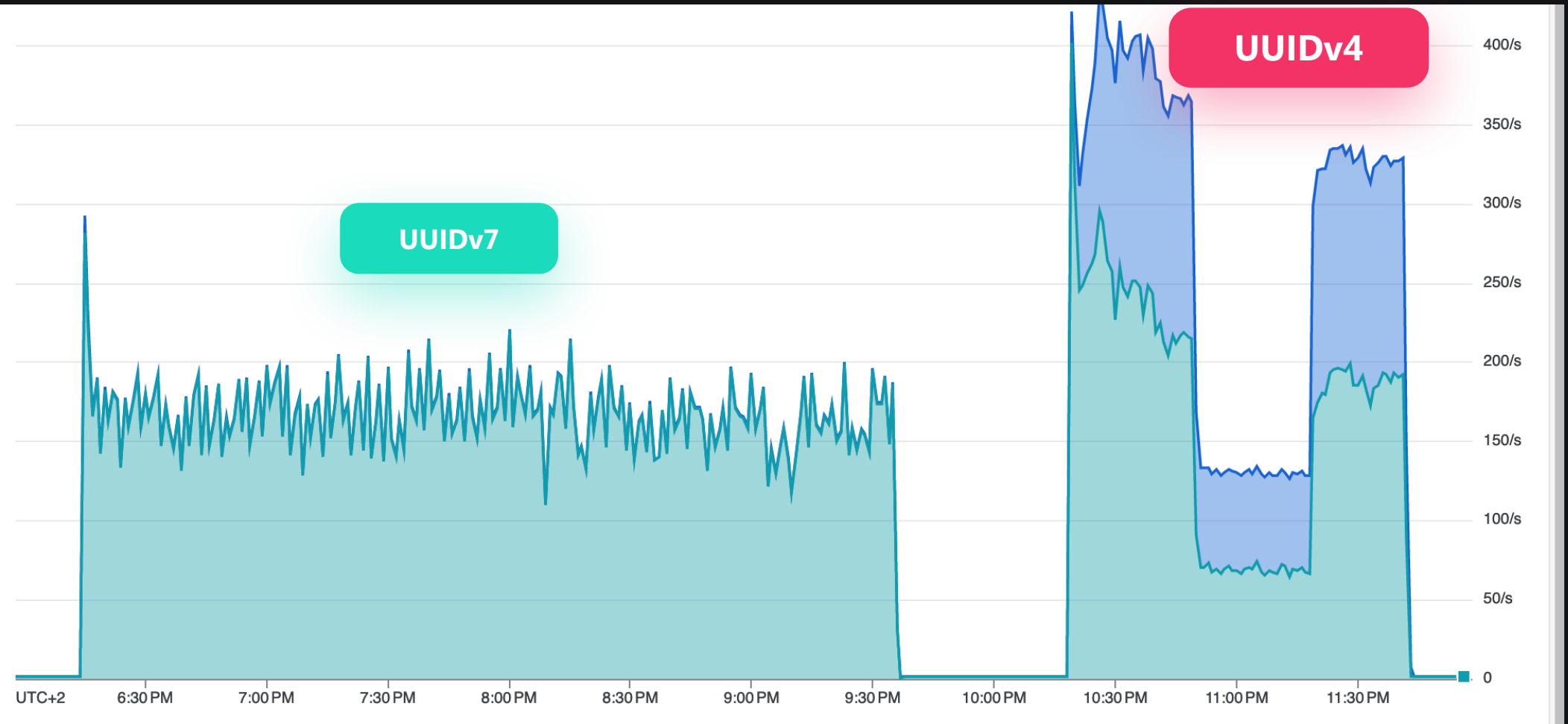




InnoDB (MySQL, OracleDB)

INSERTS WITH UUIDv7

https://ferrara.link/



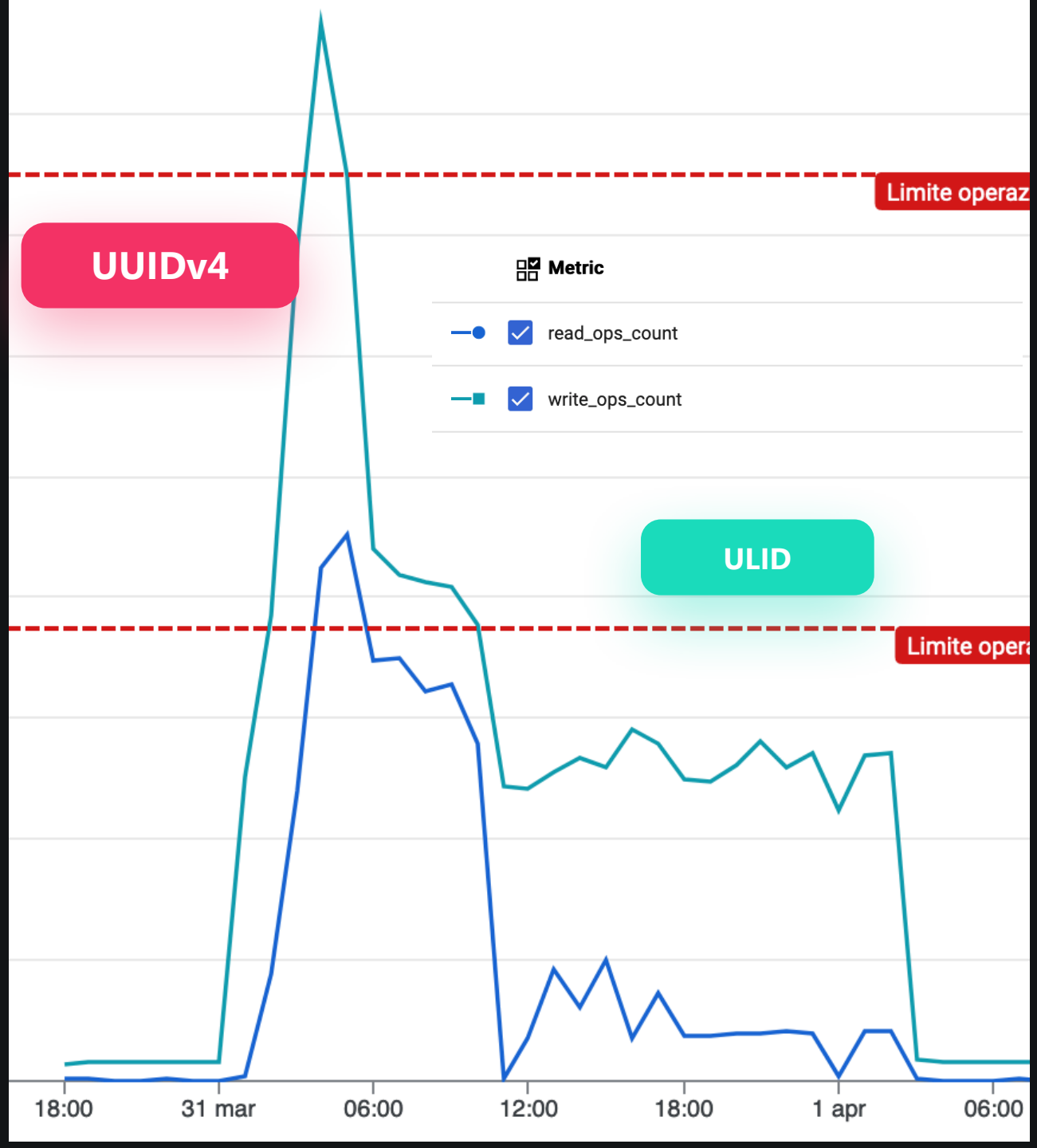
Metric	database_id	Value
innodb_pages_read		0
innodb_pages_written		1.15/s

UUIDv4 Vs UUIDv7

IF PostgreSQL

By default, pages on disk do not have a clustered index. However, having an index on a primary key with random IDs can cause many page splits. While you may not experience this immediately, as with InnoDB, the same effect will occur at a later time.

When you perform an insert with an index in PostgreSQL, the database first inserts the row into the table and then synchronously updates the index B-tree. This means that the insert operation does not return until both the table and the index have been updated.



PostgresDB

UUIDv4

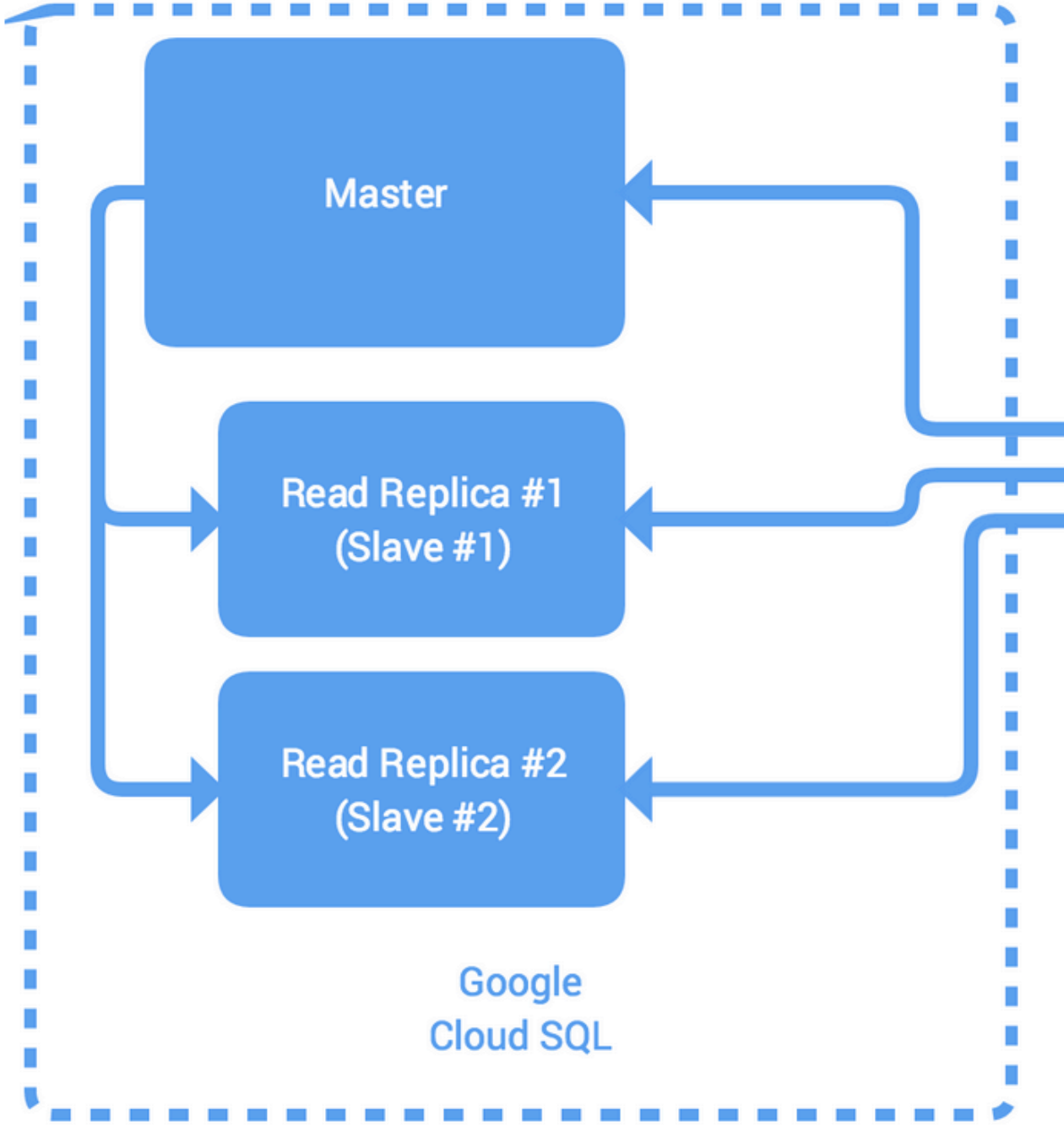
https://ferrara.link/

PostgresDB

ULID

https://ferrara.link/

WHY INSERTS ARE IMPORTANT



WHAT ABOUT READS ?

SELECT MESSAGES IN THE LAST 5 MINUTES

MEMORY, VOLATILE

PAGE 0

PAGE0

READ

HEAP - STORED IN DISK

Page 0 - CONTAINS MESSAGES WITH TIMESTAMPS THAT LIKELY COVER THE LAST 5 MINUTES

Page 1 - CONTAINS MESSAGES WITH TIMESTAMPS THAT LIKELY COVER THE LAST 10 MINUTES

Page 2 - CONTAINS MESSAGES WITH TIMESTAMPS THAT LIKELY COVER THE LAST 15 MINUTES

use for indexing. In one high-throughput system at Shopify we've seen a 50 percent decrease in INSERT statement duration by switching from UUIDv4 to ULID for idempotency keys.

AND IS SOMEONE AWARE OF THAT IN PRODUCTION?

shopify

https://shopify.engineering/building-resilient-payment-systems

Any Questions?

Yes

No!

Bye See You

Let’s Connect!

https://www.linkedin.com/in/ferrara-armando/