



ClickHouse + PeerDB Bangalore Meetup

**September
2024**

ClickHouse

Thank you to our partner!



Hosts and Speakers



Sai Srirampur

PeerDB
(now part of ClickHouse)



Kunal Gupta

PeerDB
(now part of ClickHouse)



Amogh Bharadwaj

PeerDB
(now part of ClickHouse)



Kevin Biju

PeerDB
(now part of ClickHouse)



Aditya Patadia

Co-Founder & CEO
Gumlet



Navaneet Krishna

Lead - Business Intelligence
Increff



Tech Talks



ClickHouse + PeerDB: The Tale that Connects Us

Sai Srirampur and Kunal Gupta from PeerDB Team @ ClickHouse



Tracking 1 Billion+ Image/Video Views using ClickHouse

Aditya Patadia, CEO @ Gumlet



Unlocking the Power of Big Data: Scalable BI with ClickHouse at Increff

Navaneet Krishna, Lead - Business Intelligence @ Increff



GUMLET



INCREFF



ClickHouse + PeerDB: The Tale that Connects Us

**September 18 ,
2024**

 ClickHouse

Company and culture

ClickHouse acquires PeerDB to boost real-time analytics with Postgres CDC integration

ClickHouse, Inc., the company behind the world's fastest and most popular real-time analytical database, is thrilled to announce the acquisition of PeerDB, a leading provider of change data capture (CDC) solutions



ClickHouse Team
Jul 30, 2024



ClickHouse acquires PeerDB to boost real-time analytics with Postgres CDC integration

 ClickHouse

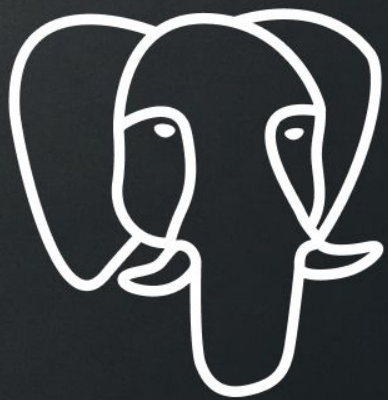


Transactional
Row Oriented
OSS
Widely Adopted



Analytical
Column Oriented
OSS
Widely adopted



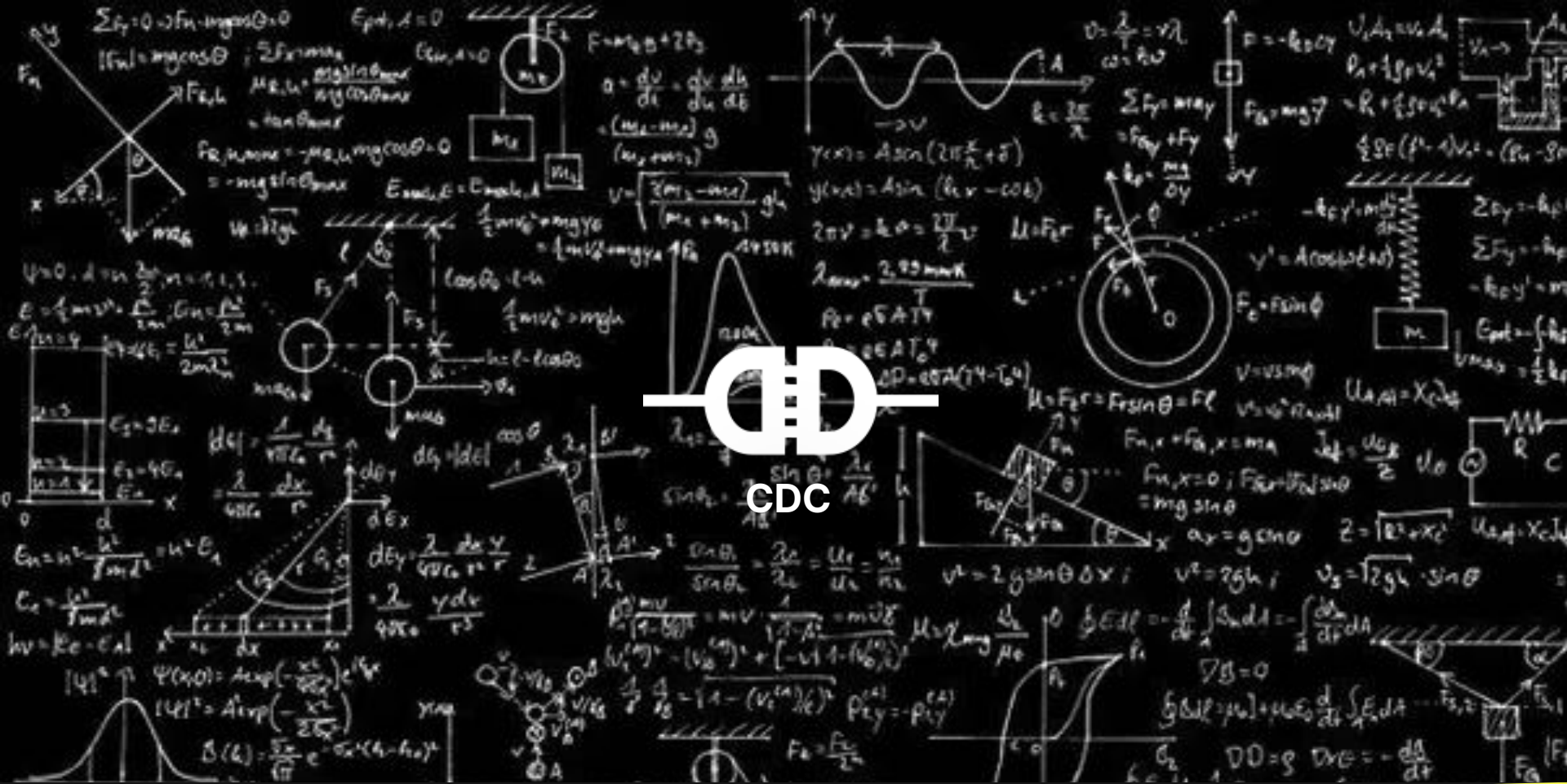


Transactional
Row Oriented
OSS
Widely Adopted



Analytical
Column Oriented
OSS
Widely adopted





CDC is hard. It should ensure:

- ClickHouse is always upto date irrespective of load on Postgres
- Minimal load on Postgres and is crash safe
- Data Replication in the most native format on ClickHouse
- Resilient to any failure
- and so on..





**Fast, simple, and cost effective
Postgres replication**



Company and culture

ClickHouse acquires PeerDB to boost real-time analytics with Postgres CDC integration

ClickHouse, Inc., the company behind the world's fastest and most popular real-time analytical database, is thrilled to announce the acquisition of PeerDB, a leading provider of change data capture (CDC) solutions



ClickHouse Team
Jul 30, 2024



ClickHouse acquires PeerDB to boost real-time analytics with Postgres CDC integration

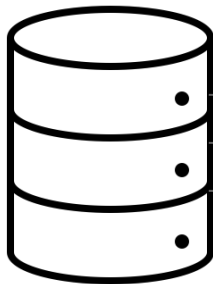
 ClickHouse



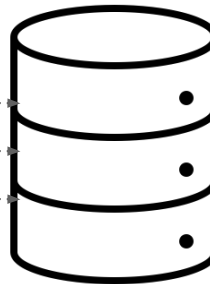
PeerDB Deep Dive



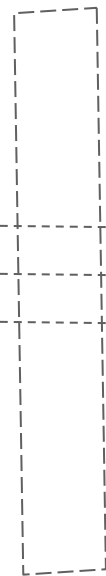
Simple UI or a Powerful SQL Interface



Source PEER - Postgres



Target PEER - ClickHouse



MIRROR

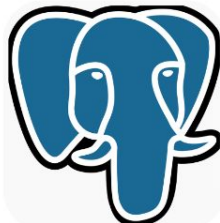
**Binary Compressed Data Transfer
w/ Avro**

PeerDB is Built For Postgres

Native Monitoring and metrics

No Kafka or other
external deps

**Parallel reading of CDC
replication slots**



CTID based initial load



Resilient Orchestration

**Binary Copy For
Homogeneous Transfer**

Streaming Server Side Cursors

Query Based Streaming

Parallel Snapshotting



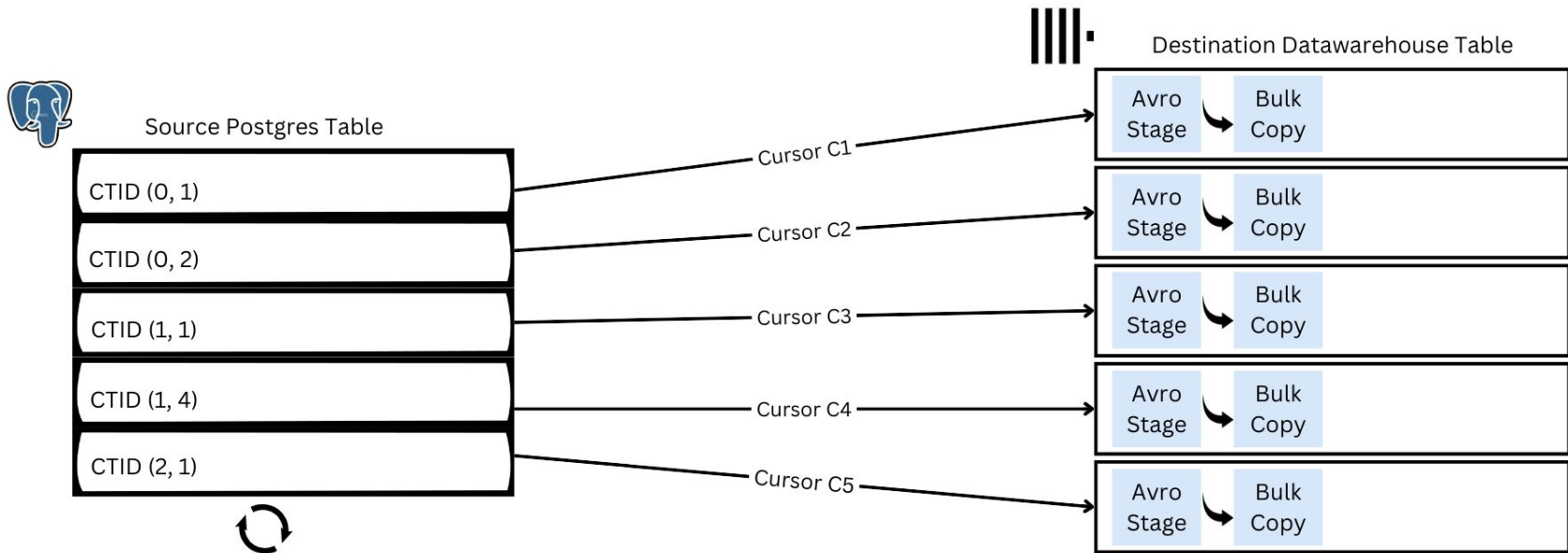
**DBT
Integration**

Toast Column Support

**Lossless Unified
Datatypes**

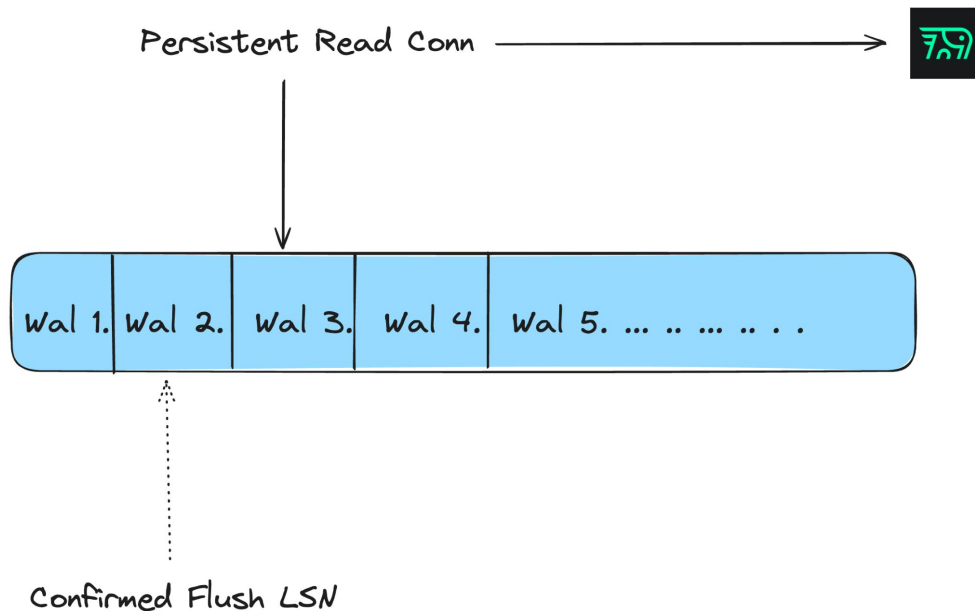
Parallel Snapshotting

..... Snapshot Connection



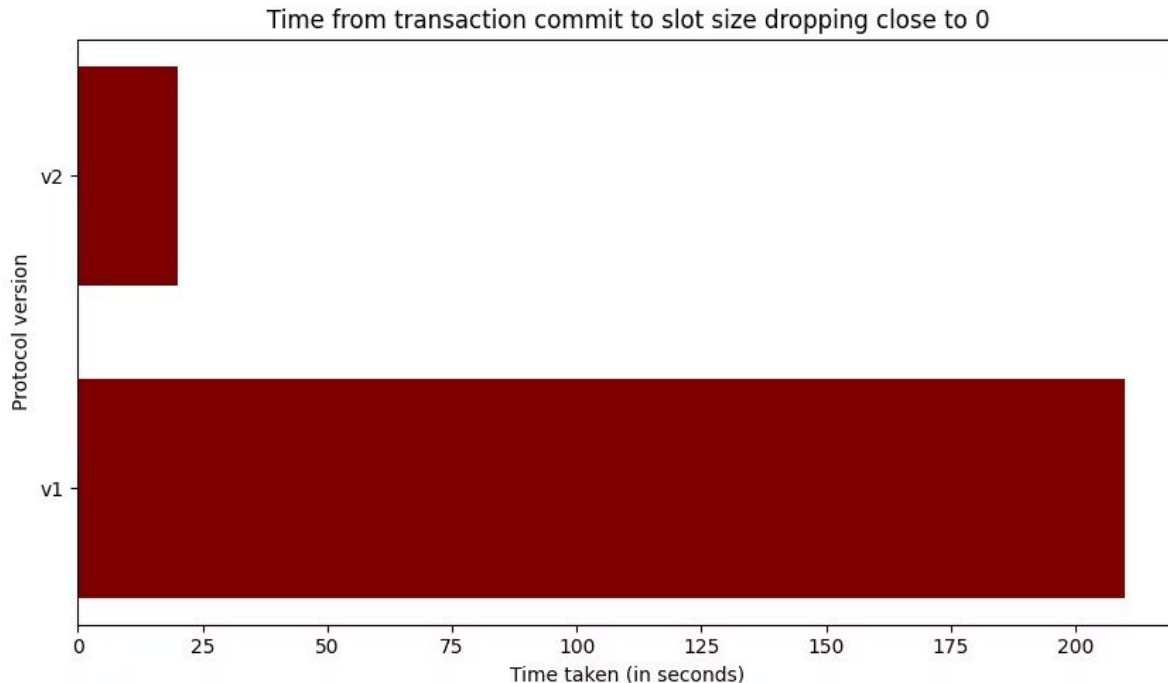
Move Terabytes within Hours instead of Days

Always Consume the replication slot



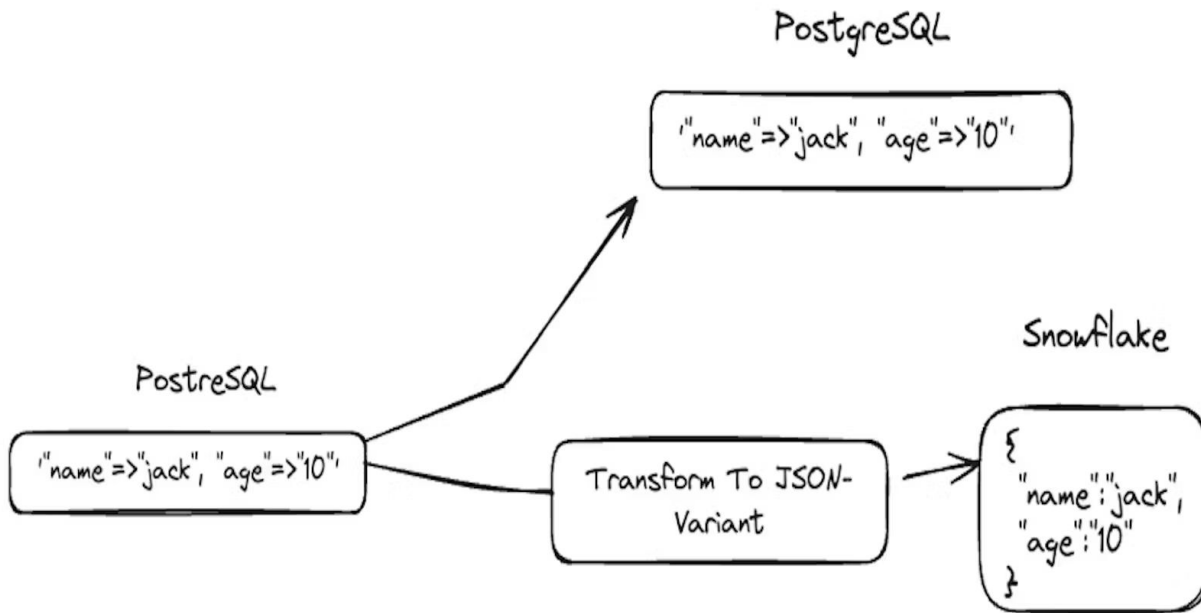
Your Postgres Database is Crash Safe

Consume in-flight transactions using version 2



Drastically reduced slot growth and better perf

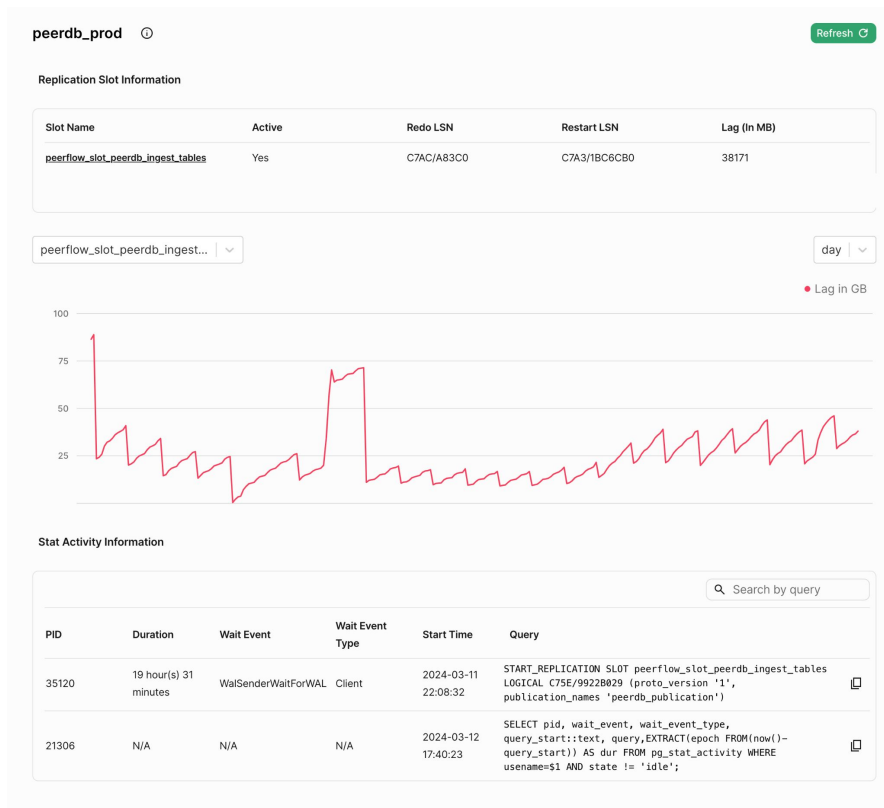
Replicating Data Types Natively



Avoids transformations to reduce tech debt and costs

Postgres Native Monitoring

1. Replication Slot Growth / Lag
2. Number of connections
3. Activity on Postgres - Wait Events



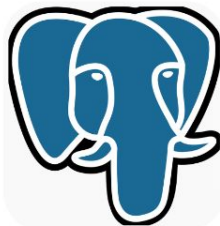
**Binary Compressed Data Transfer
w/ Avro**

PeerDB is Built For Postgres

Native Monitoring and metrics

No Kafka or other
external deps

**Parallel reading of CDC
replication slots**



CTID based initial load



Resilient Orchestration

**Binary Copy For
Homogeneous Transfer**

Streaming Server Side Cursors

Query Based Streaming

Parallel Snapshotting



**DBT
Integration**

Toast Column Support

**Lossless Unified
Datatypes**

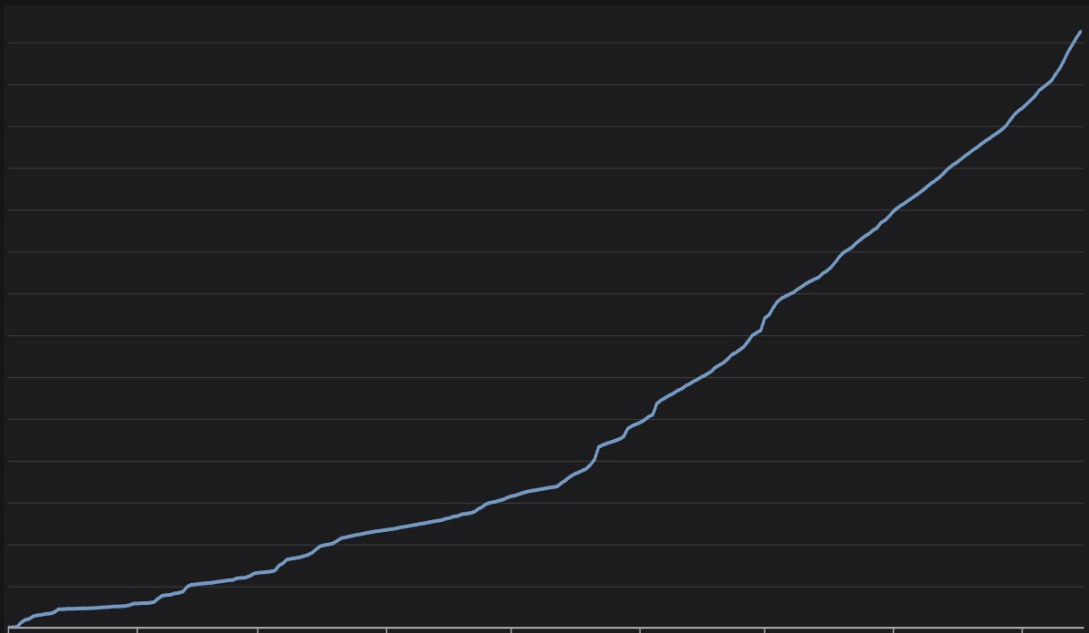
How to use PeerDB ?

- PeerDB Open Source - <https://github.com/PeerDB-io/peerdb>
- PeerDB Cloud - <https://peerdb.cloud/>
- Active work on integrating PeerDB Cloud into ClickPipes, native ingestion service of ClickHouse.

How we Built PeerDB Cloud

Scale to Date

- Multiple Petabytes of data transferred
- 100+ active mirrors



SYNC STATUS			
<input type="checkbox"/>	Unknown	0	
<input type="checkbox"/>	Synced	145	
<input type="checkbox"/>	OutOfSync	0	
PROJECTS			<button>CLEAR</button>
<input checked="" type="checkbox"/>	peerdb-cloud		

Built for Resiliency

- AWS
- Kubernetes
 - Autoscaling, HA



Built for Resiliency

- AWS
- Kubernetes
 - Autoscaling, HA
- GitOps
- Infrastructure as Code

```
> ls eks --tree -L2
eks
├── ap-south-1
│   └── prod-2
├── ap-southeast-2
│   └── prod-3
├── eu-central-1
│   └── prod-4
├── us-east-1
│   └── prod-5
├── us-east-2
│   └── prod-6
├── us-west-2
│   └── prod-cloud-1
```



git



Built for Resiliency

- AWS
- Kubernetes
 - Autoscaling, HA
- GitOps
- Infrastructure as Code
- ArgoCD
 - Dependency Management
 - State reconciliation

```
> ls eks --tree -L2
eks
├── ap-south-1
│   └── prod-2
├── ap-southeast-2
│   └── prod-3
├── eu-central-1
│   └── prod-4
├── us-east-1
│   └── prod-5
├── us-east-2
│   └── prod-6
├── us-west-2
│   └── prod-cloud-1
```



git



argo

Built for Resiliency

- AWS
- Kubernetes
 - Autoscaling, HA
- GitOps
- Infrastructure as Code
- ArgoCD
 - Dependency Management
 - State reconciliation
- Temporal
 - For Data Synchronization State Management

```
> ls eks --tree -L2
eks
├── ap-south-1
│   └── prod-2
├── ap-southeast-2
│   └── prod-3
├── eu-central-1
│   └── prod-4
├── us-east-1
│   └── prod-5
├── us-east-2
│   └── prod-6
├── us-west-2
│   └── prod-cloud-1
```



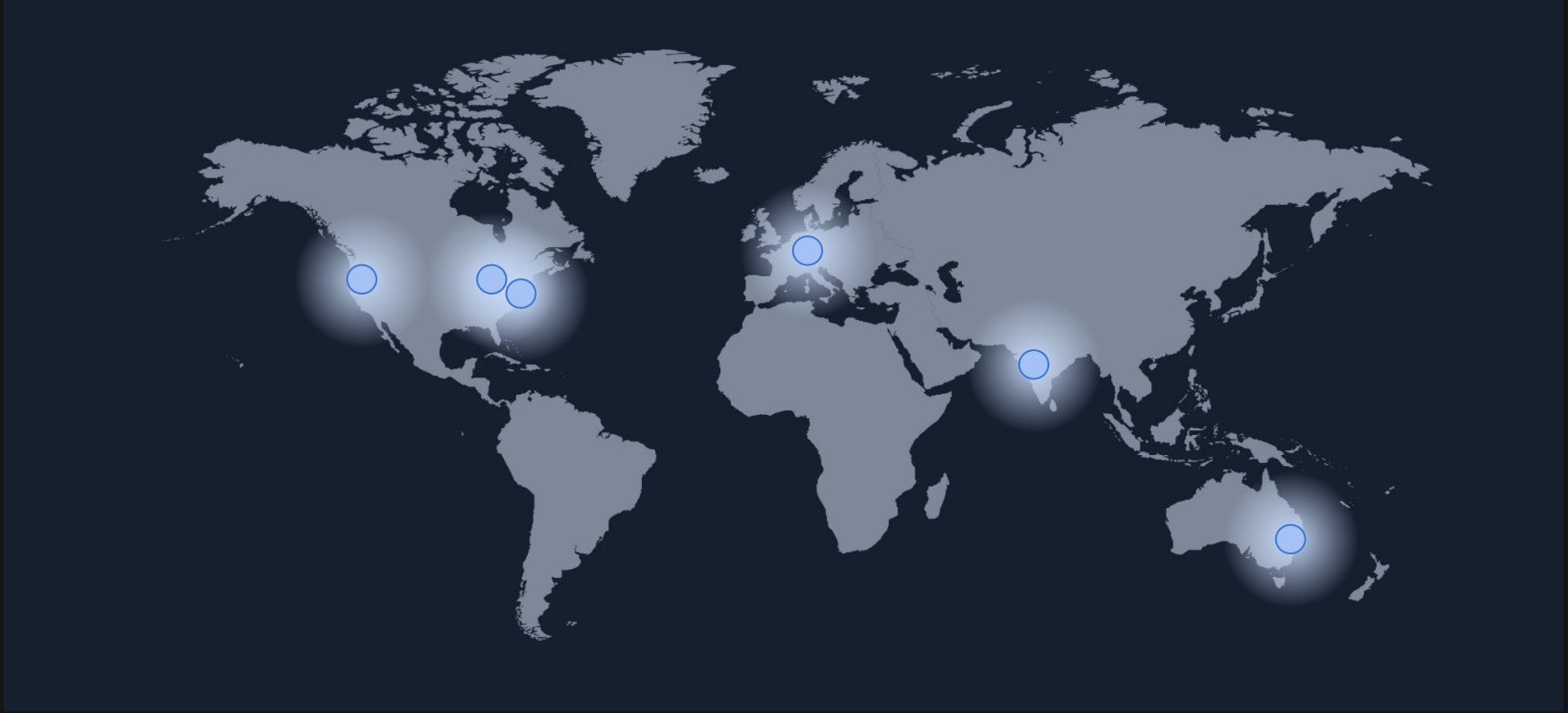
Built for Resiliency

- AWS
- Kubernetes
 - Autoscaling, HA
- GitOps
- Infrastructure as Code
- ArgoCD
 - Dependency Management
 - State reconciliation
- Temporal
 - For Data Synchronization State Management
- Detailed Internal documentation

```
> ls eks --tree -L2
eks
├── ap-south-1
│   └── prod-2
├── ap-southeast-2
│   └── prod-3
├── eu-central-1
│   └── prod-4
├── us-east-1
│   └── prod-5
├── us-east-2
│   └── prod-6
├── us-west-2
│   └── prod-cloud-1
```

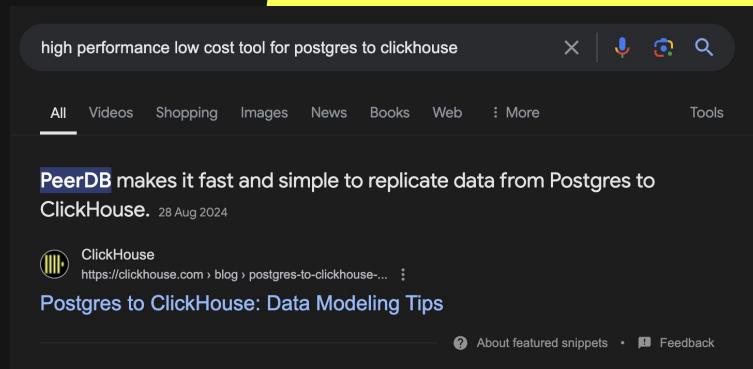


Global Availability across 6 Regions



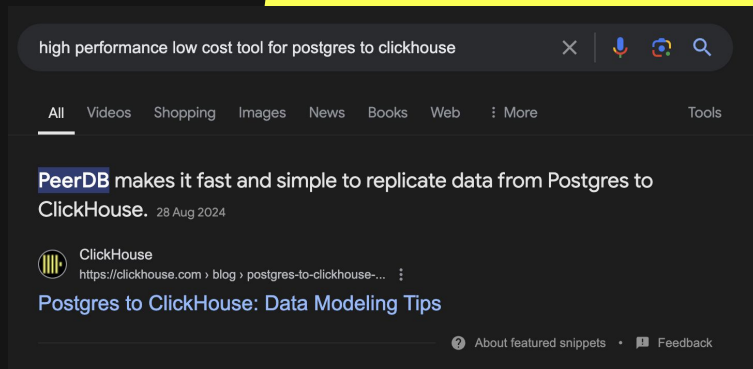
Features and Design Benefits

- Cost Efficiency



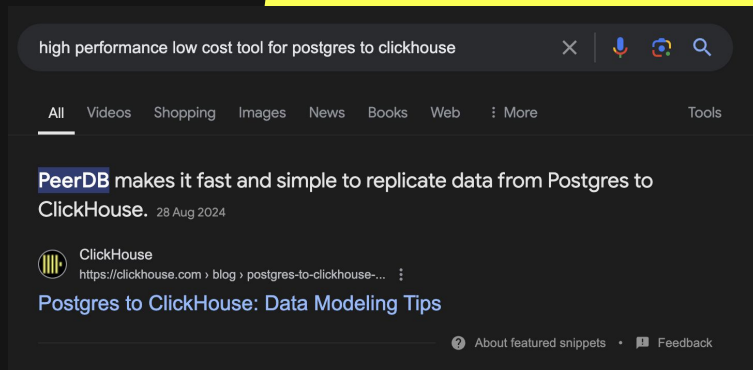
Features and Design Benefits

- Cost Efficiency
- Observability and Monitoring
 - All alerts routed internally
 - 24×7 on-call monitoring
 - OTel integration for custom metrics



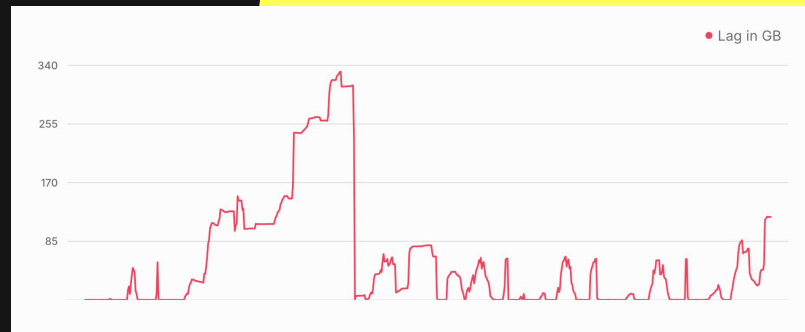
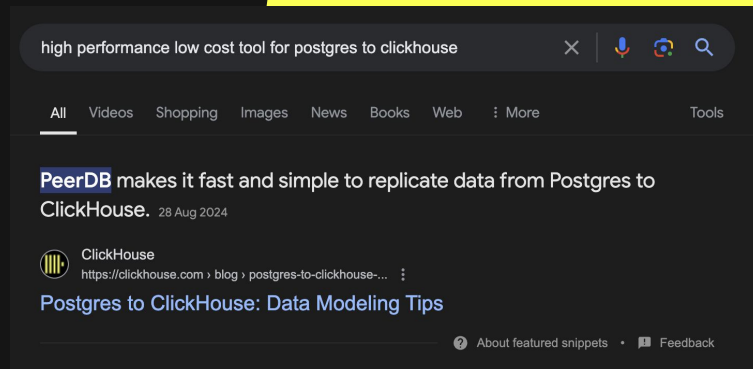
Features and Design Benefits

- Cost Efficiency
- Observability and Monitoring
 - All alerts routed internally
 - 24×7 on-call monitoring
 - OTel integration for custom metrics
- Customer visibility
 - Alerts for slot lag and open connections



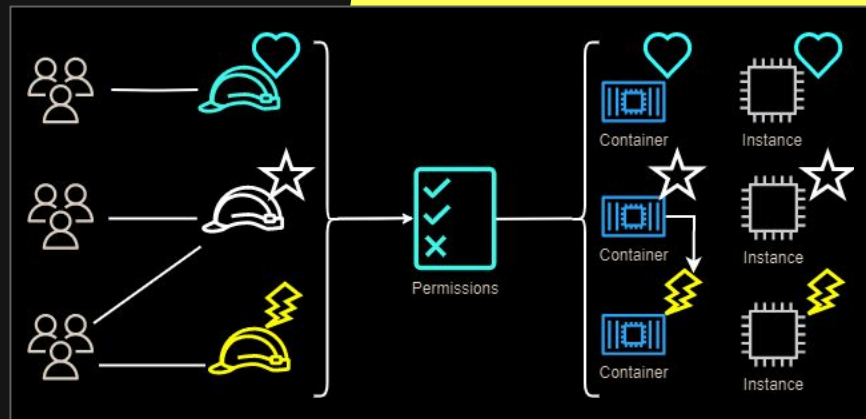
Features and Design Benefits

- Cost Efficiency
- Observability and Monitoring
 - All alerts routed internally
 - 24×7 on-call monitoring
 - OTel integration for custom metrics
- Customer visibility
 - Alerts for slot lag and open connections
- Transparent and seamless upgrades



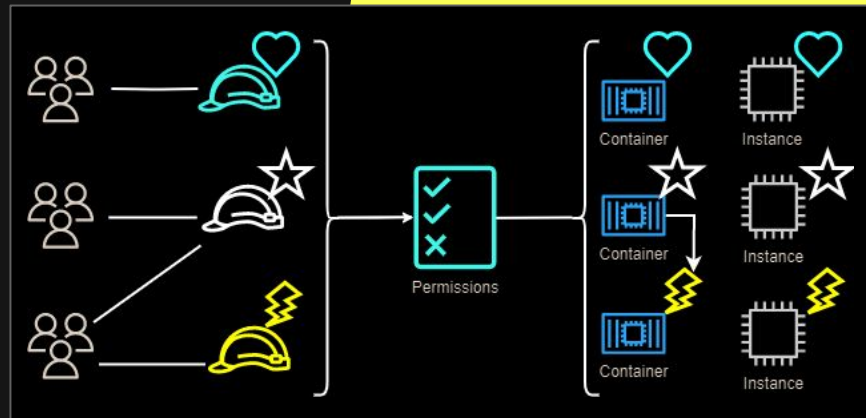
Built with Security

- Principle of least privilege



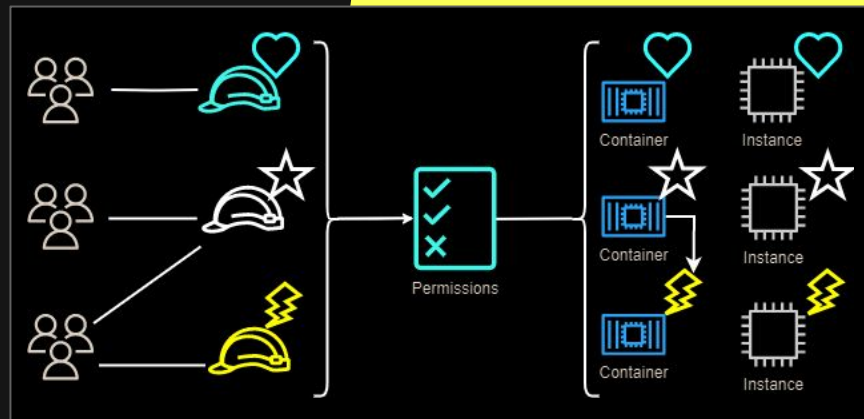
Built with Security

- Principle of least privilege
- Data Retention



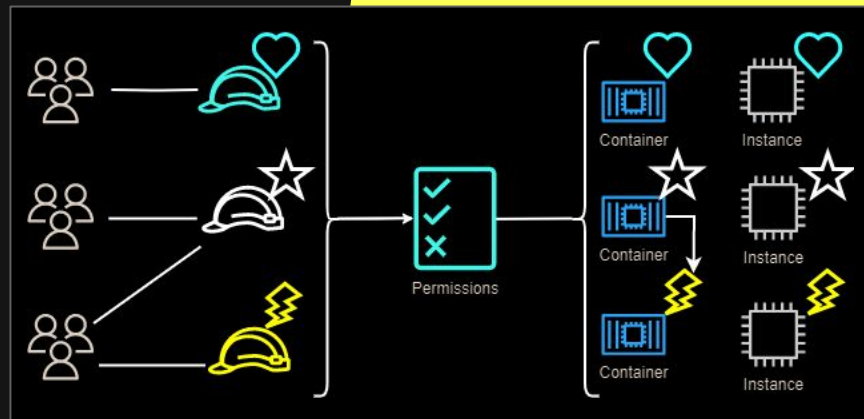
Built with Security

- Principle of least privilege
- Data Retention
- Encryption everywhere



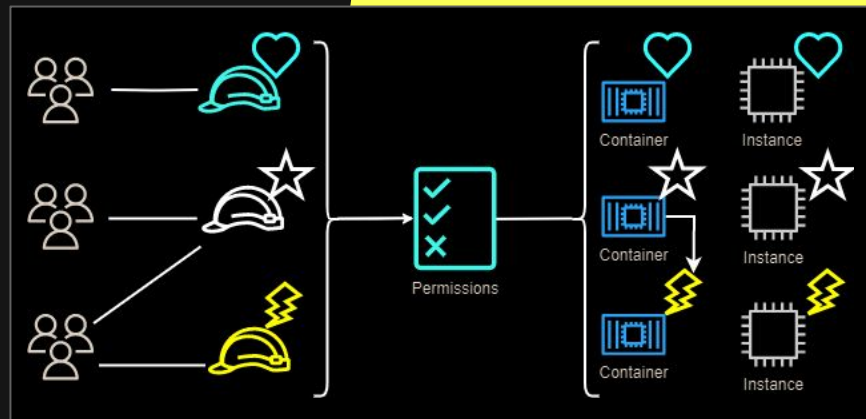
Built with Security

- Principle of least privilege
- Data Retention
- Encryption everywhere
- Additional features for customers
 - SSH Tunnelling
 - Private Link
 - IP Whitelisting



Built with Security

- Principle of least privilege
- Data Retention
- Encryption everywhere
- Additional features for customers
 - SSH Tunnelling
 - Private Link
 - IP Whitelisting
- Compliance Frameworks
 - SOC2
 - GDPR



PeerDB Cloud ↔ ClickHouse Cloud

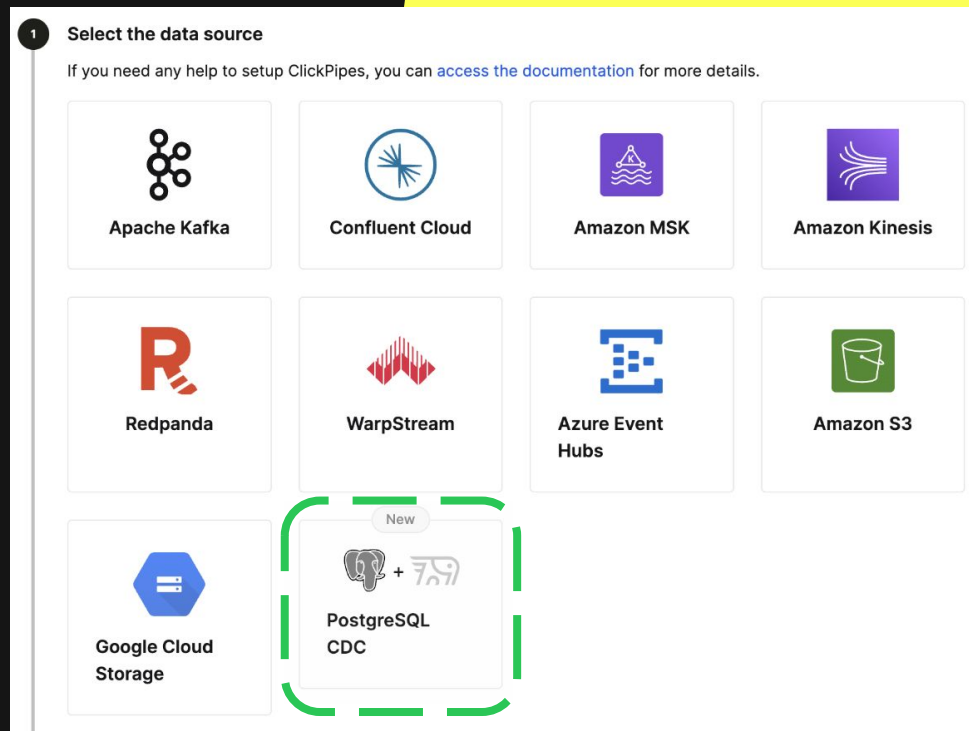
- Scalability
- High Availability
- Observability and Monitoring
- Security



PeerDB Cloud ↔ ClickHouse Cloud

- Scalability
- High Availability
- Observability and Monitoring
- Security

to be available in **ClickPipes** soon...



Demo of PeerDB

PeerDB Cloud has a 4-month free trial

Sign up by visiting

<https://peerdb.cloud>



Create a new instance

Instance Name

Password

Region

Instance Offerings

We provide four offerings of PeerDB instances.

PeerDB Micro

Compute: 2 vcores



Recommended for up to 50 million monthly active rows.

This offering is free to use for 120 days, after which you will be invoiced \$250 per month.

PeerDB Small



Compute: 4 vcores

Recommended for 50 - 500 million monthly active rows.

This offering is priced at \$1,000 invoiced every month.

PeerDB Medium



Compute: 8 vcores

Recommended for 500 million - 1 billion monthly active rows.

This offering is priced at \$2,000 invoiced every month.

PeerDB Large



Compute: 16 vcores

Recommended for 1 billion+ monthly active rows.

This offering is priced at \$4,000 invoiced every month.

Continue



Thank you!

Keep in touch!



clickhouse.com/slack



clickhouse-bangalore-user-group



#clickhouseDB @clickhouseinc



clickhouse



Q&A