Large scale distributed indexing with Kafka

Presented by Remi Dettai

Paris, France 2024

PERCONA UNIVERSITY

FerretDB

BlaBlaCar

Who am I?

Remi Dettai (@rdettai on Github)

Core dev on the Quickwit engine (Rust)

Lead on the Quickwit distribution for AWS Lambda

Data/Cloud engineer background

Agenda

1) The large scale logging problem

2) How Quickwit and its Kafka source solve the problem

- Architecture
- Indexing pipelines
- Exactly-once with the Kafka source
- Distributing indexing pipelines

3) Real world example

1) The large scale logging problem

1) The large scale logging problem

Some domains, like finance, have harsh logging requirements:

- Petabyte scale
- Huge traffic spikes that need to be queryable in seconds
- Cannot compromise on delivery reliability (exactly once)

1) The large scale logging problem

Query	Metric	Quickwit (indexed)	Loki (columnar only)
`queen`	Latency (s)	0.6s	9.3s (+1,425%)
	CPU Time (s)	2.7s	146s (+5,270%)
	# of GET Requests	206	14,821* (+7,095%)
`us-east-2` (label)	Latency (s)	0.6s	1.0s (+74%)
	CPU Time (s)	2.7s (+35%)	2s
	# of GET Requests	211* (+348%)	47

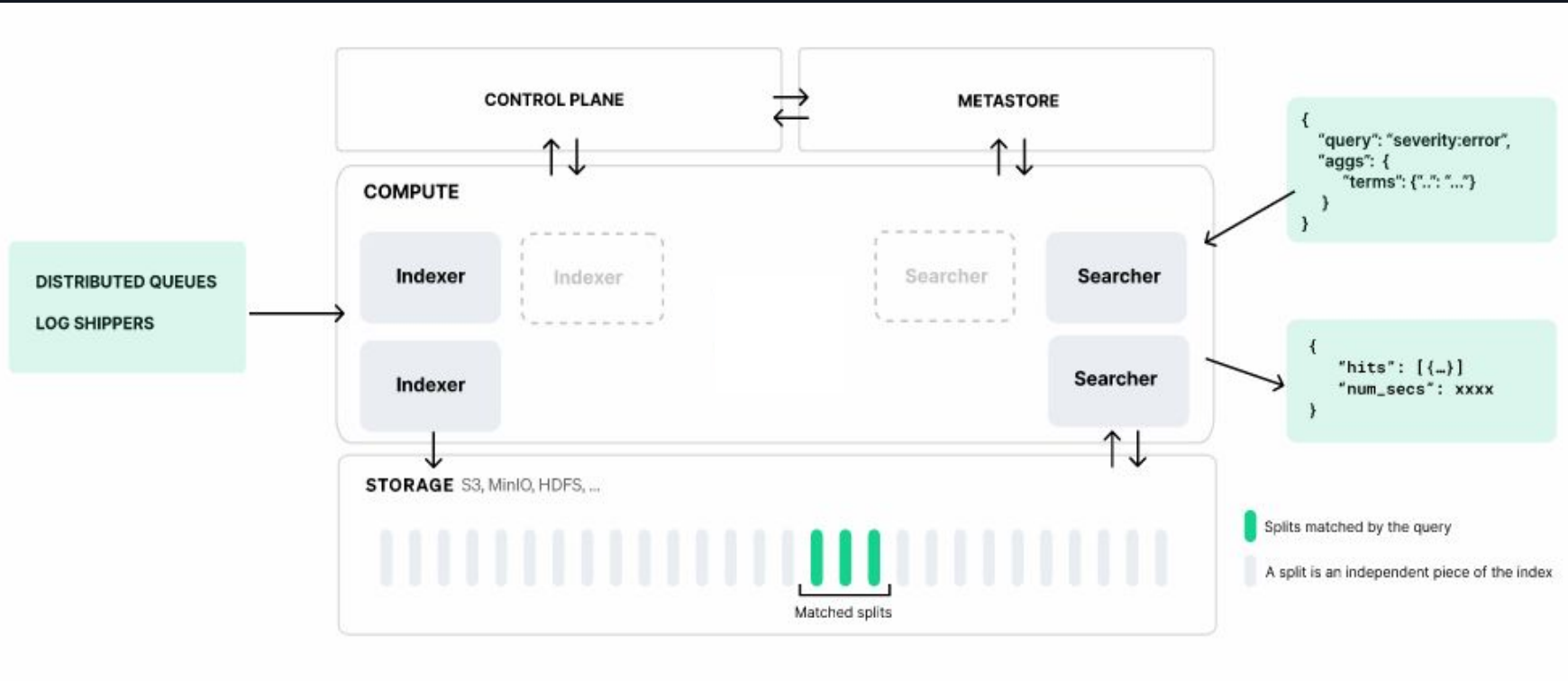
16 vCPUs / 64GB RAM

Elastic Integration Corpus Generator

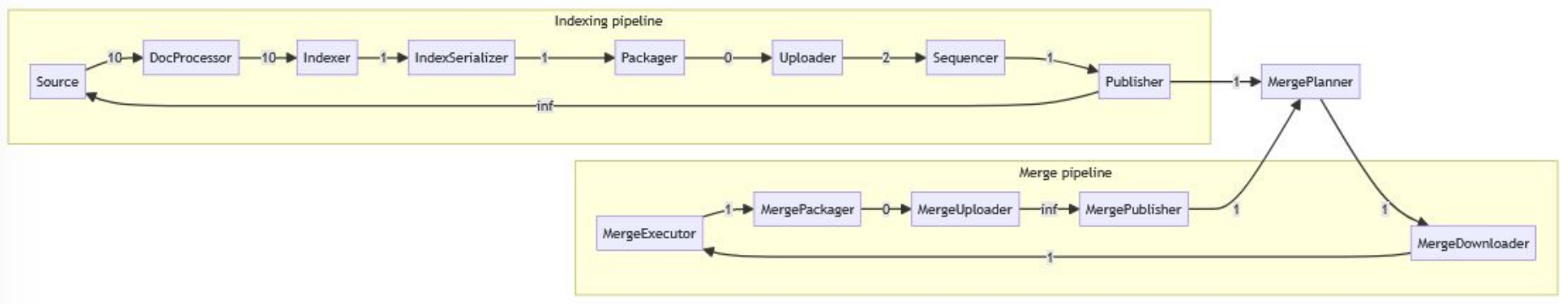
212.40 GB
243,527,673 logs

2) How Quickwit and its Kafka source solve the problem

2) Architecture

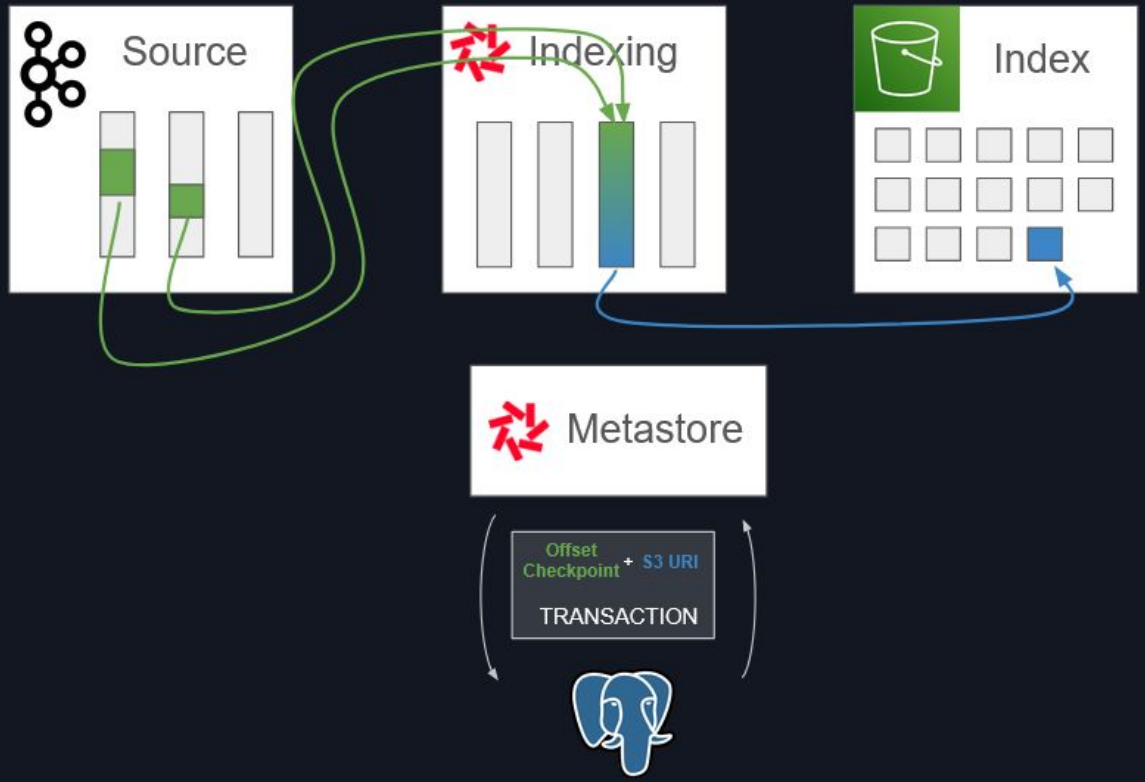


2) Indexing pipelines

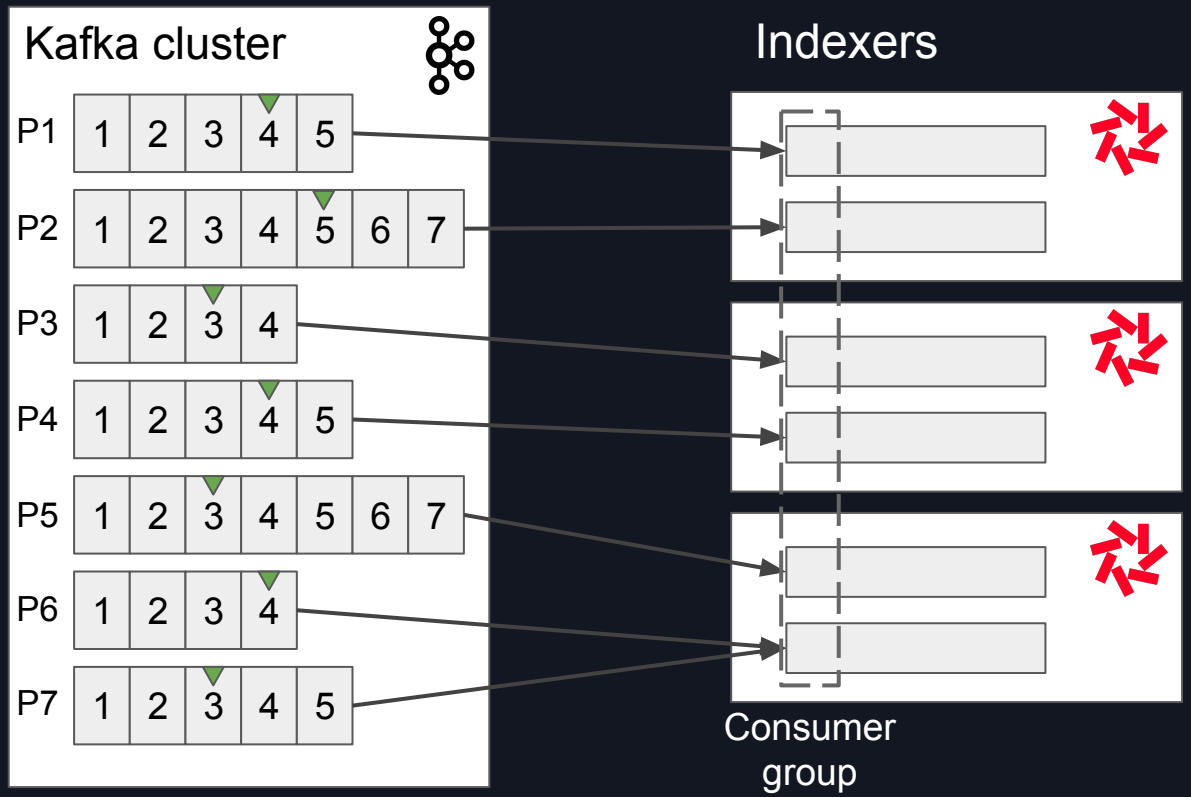


Engine	Quickwit	Loki
Ingestion time (min)	123 (+123%)	55
Mean vCPU	2.2	2.75
Total CPU time (min)	~270 (+80%)	~151
Number of files on GCS	25	145,756 (x5,829)
Bucket size (GiB)	53	55

2) Kafka source exactly-once semantics



2) Distributing indexing pipelines



Control plane

Target: 6 pipelines
Indexers: 3 x 8vCPU

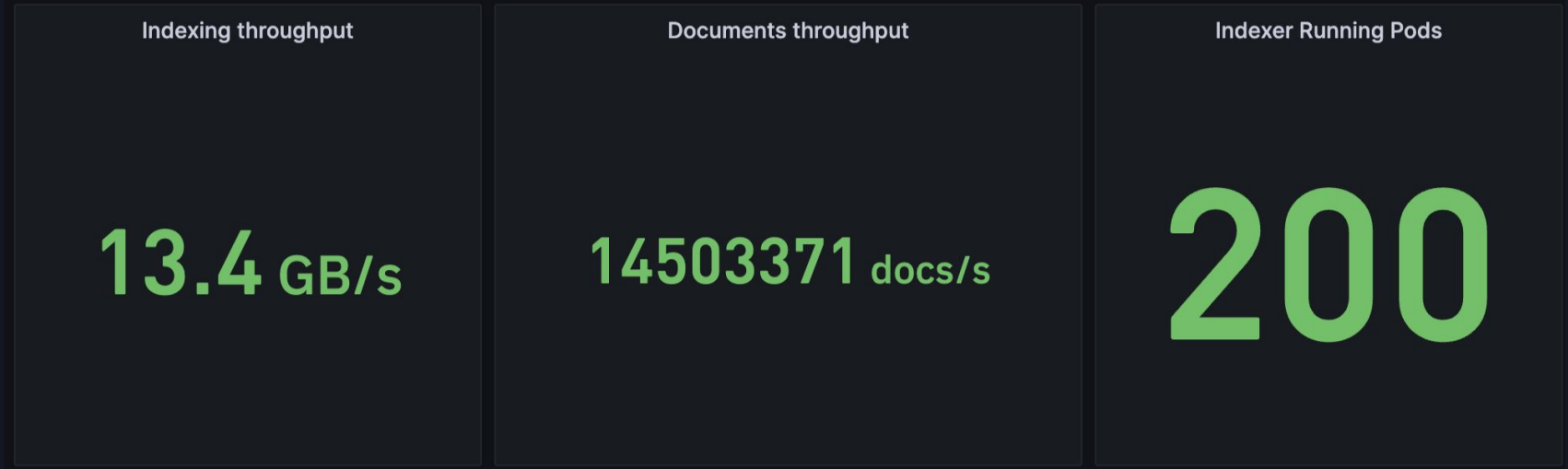
Metastore

P1 ▸ 4
P2 ▸ 5
P3 ▸ 3
P4 ▸ 4
P5 ▸ 3
P6 ▸ 4
P7 ▸ 3

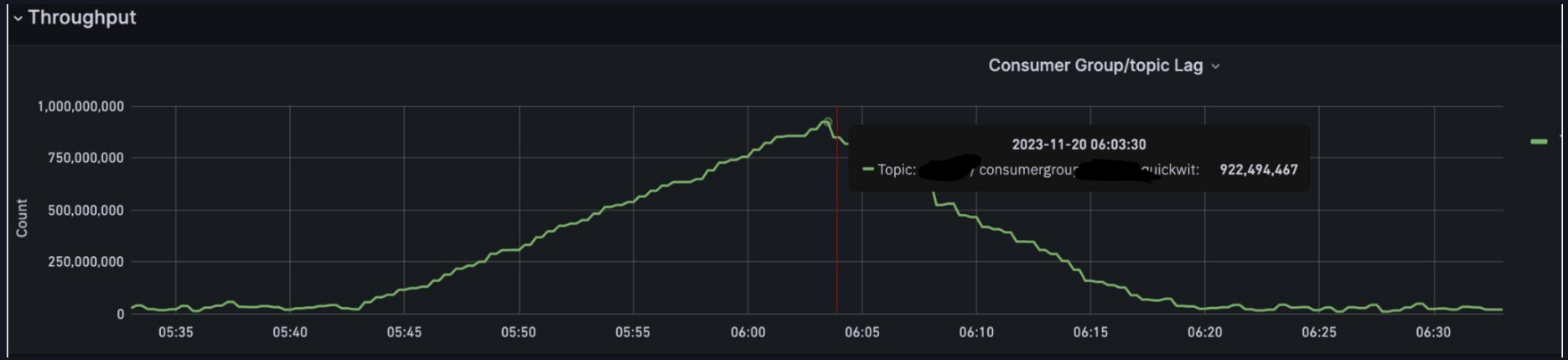
3) Real world example

3) Real world example

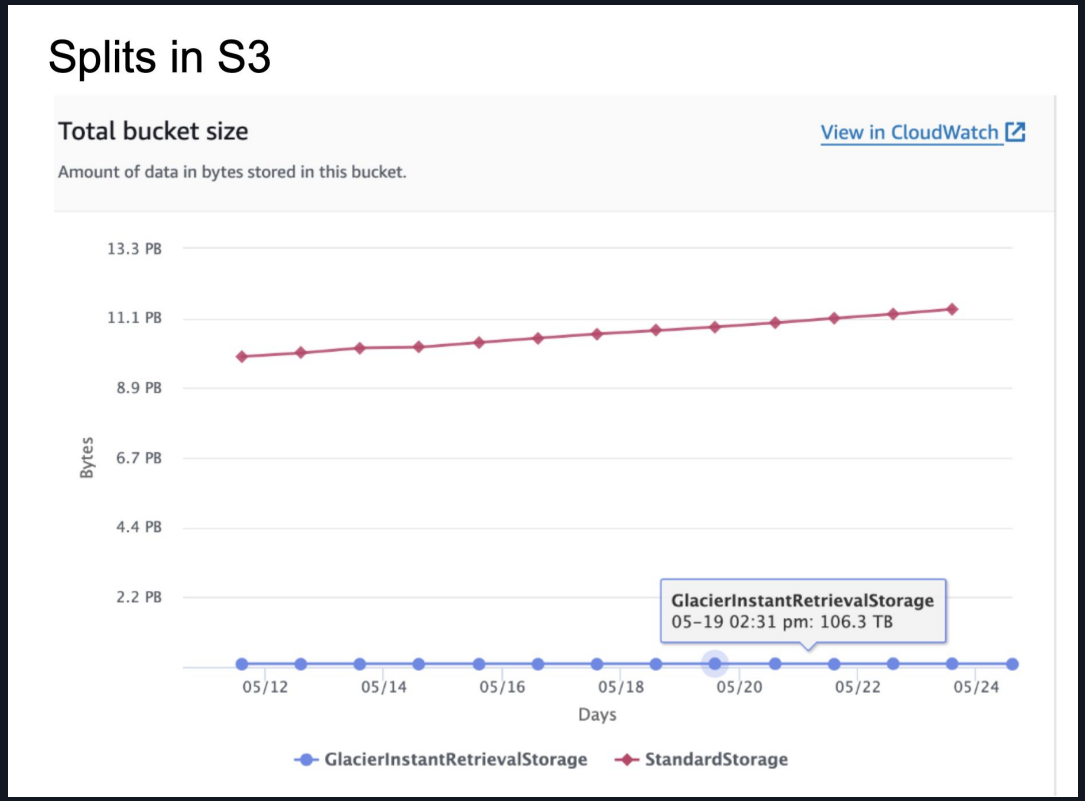
Indexing at Binance, leading crypto platform



3) Real world example



3) Real world example



Older data stored on Glacier:
- total is 20PB

With a compression ratio of 8:
- 160PB ingested

3) Real world example

Cost comparison for 1PB per day

Cloudwatch

- 1PB * $0.5/GB -> $500k / day

Quickwit

- 1000 indexers with 6CPU running on c7g
- $0.0361 (/cpu.h) * 1000 * 6 * 24 = $5200 / day

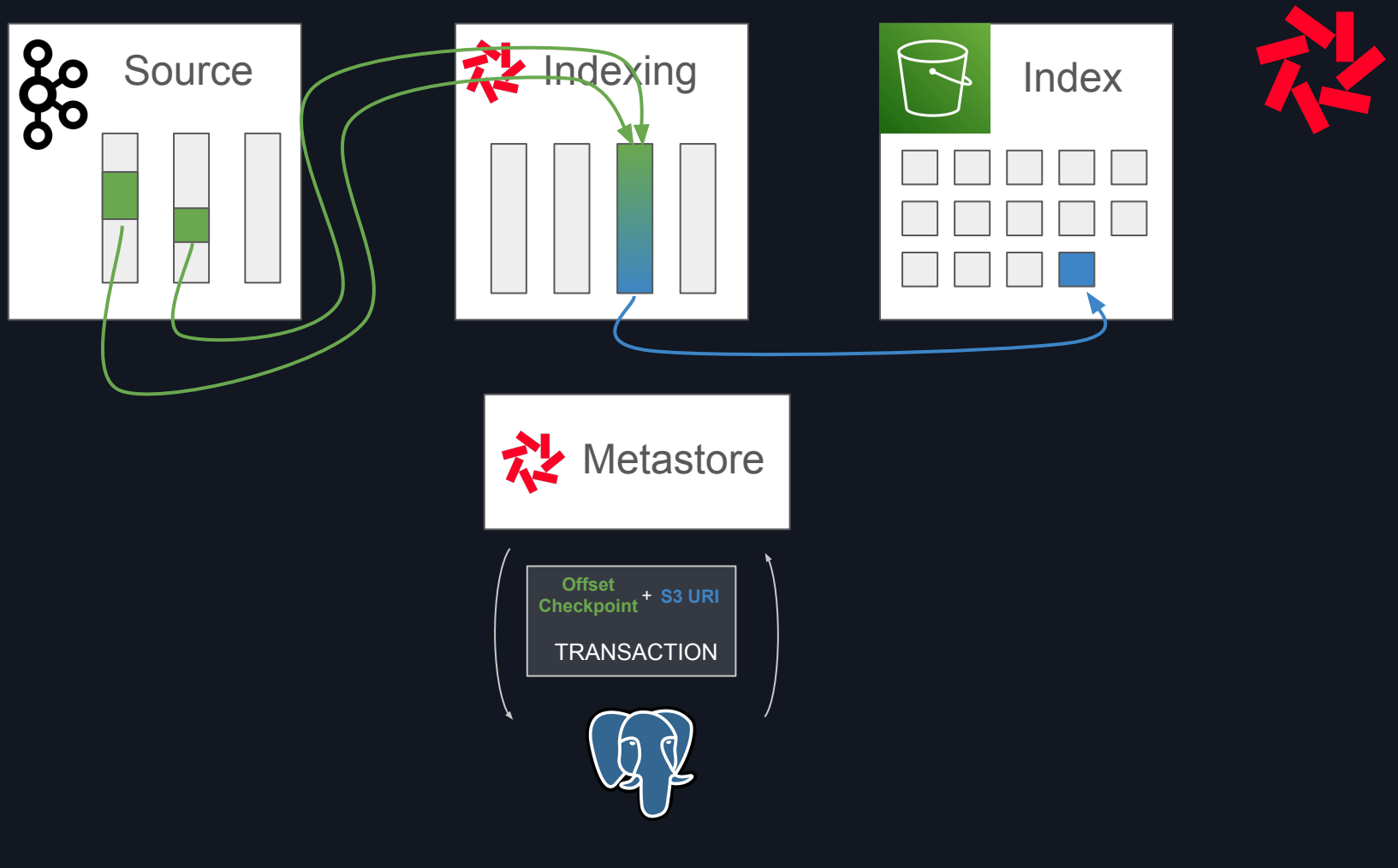
100X

References

Quickwit architecture	https://quickwit.io/docs/overview/architecture
Loki vs Quickwit	https://quickwit.io/blog/benchmarking-quickwit-loki
Kafka source tutorial	https://quickwit.io/docs/ingest-data/kafka
Kafka rebalancing	https://www.confluent.io/online-talks/everything-you-always-wanted-to-know-about-kafkas-rebalance-protocol-but-were-afraid-to-ask-on-demand/

Thank you

Contact: remi@quickwit.io


Source
Indexing
Index
Metastore
Offset Checkpoint
+
S3 URI
TRANSACTION