Understand Systems with OpenTelemetry: A Hybrid Telemetry Data Backend

Ran Xu, Huawei Cloud & Xiaochun Yang, Northeastern University China

Xu Ran

A database expert in Huawei Cloud Database Innovation Lab and openGemini time-series database architect
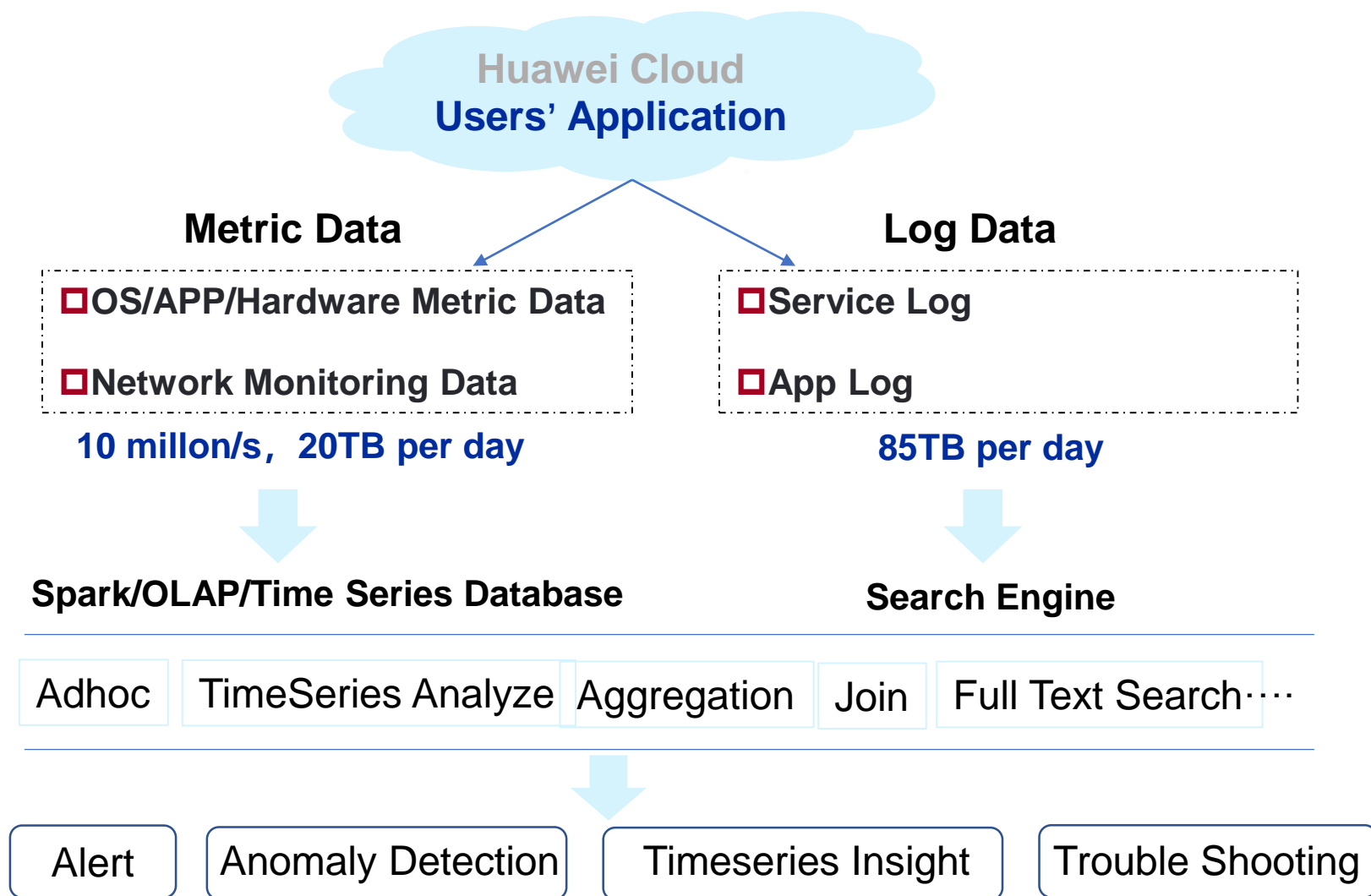
Xiaochun Yang

A professor of Northeastern University in China, has been engaged in teaching and research in the field of data management and analysis for a long time.

Contents

- **Challenge of Observability**

- **openGemini as the Backend Storage Solution of OpenTelemetry**

- **How openGemini Works for Storing Logs, Metrics, and Traces**

Challenge: Massive Telemetry Data with Different Technologies

Huawei Cloud
Users' Application

Metric Data

☐ **OS/APP/Hardware Metric Data**

☐ **Network Monitoring Data**

10 millon/s，20TB per day

Log Data

☐ **Service Log**

☐ **App Log**

85TB per day

Spark/OLAP/Time Series Database

Search Engine

| Adhoc | TimeSeries Analyze | Aggregation | Join | Full Text Search···· |

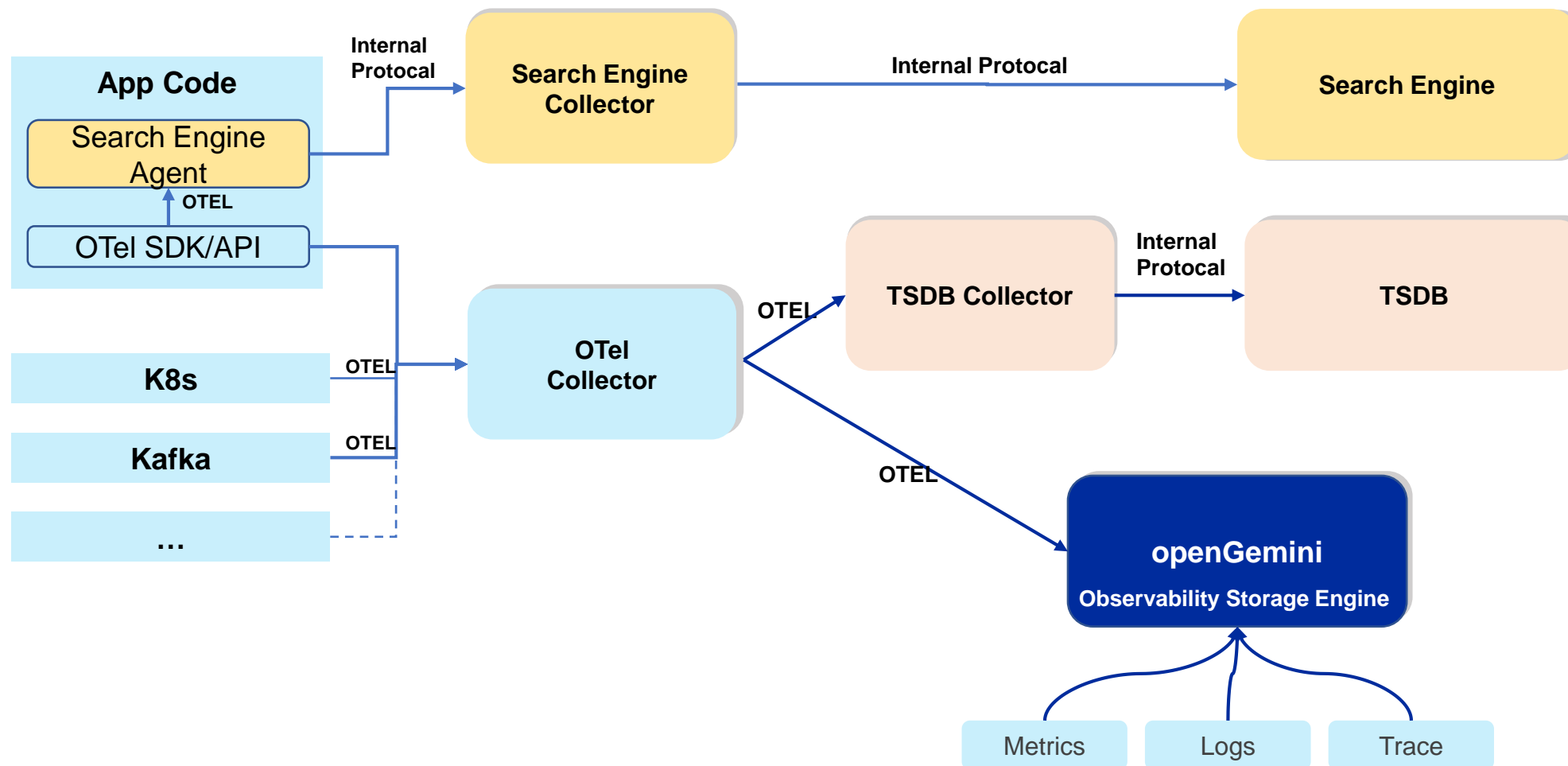| Alert | Anomaly Detection | Timeseries Insight | Trouble Shooting |

····

Technical Challenges

- **Massive Data**

- **Cost Sensitive**

- **Operation**

- **Correlation Analysis**

OTel: Unified Standards, Correlation Between Metrics, Logs, and Traces
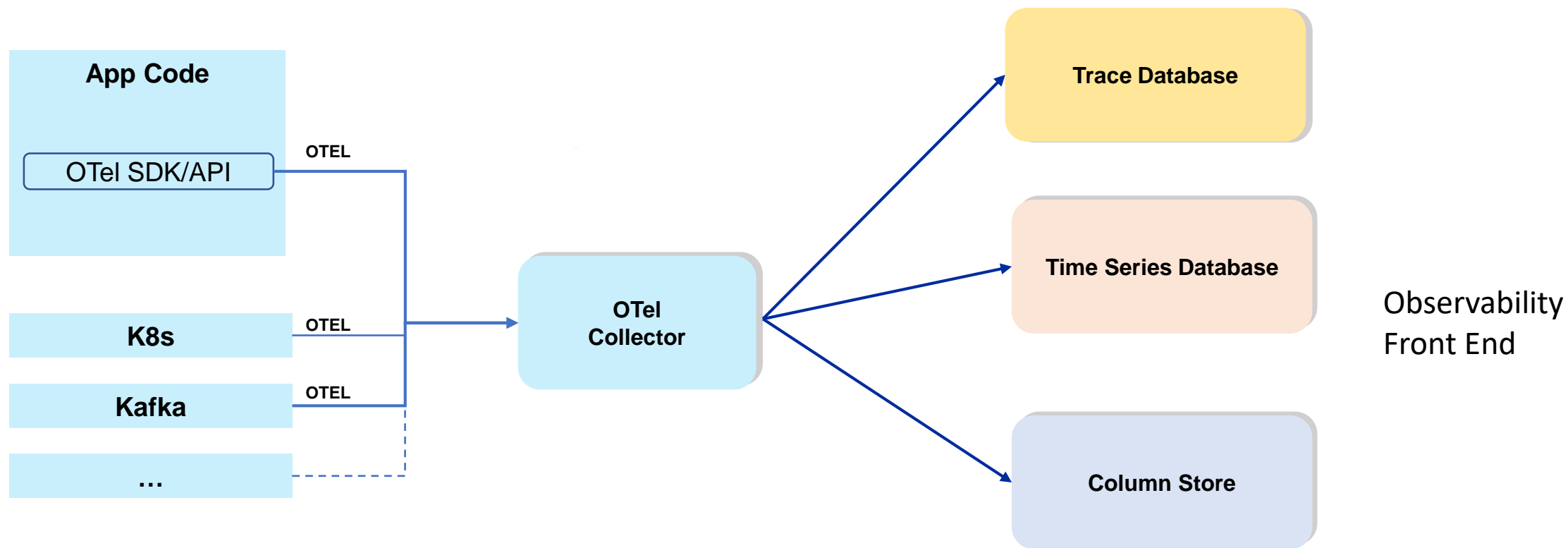
- ➢ **Unified Access Protocols：** metrics, traces, and logs.

- ➢ **Standard Data models**

- ➢ **Various SDKs:** 11 common languages, such as C++ and Go, are supported.

- ➢ **Open Source**

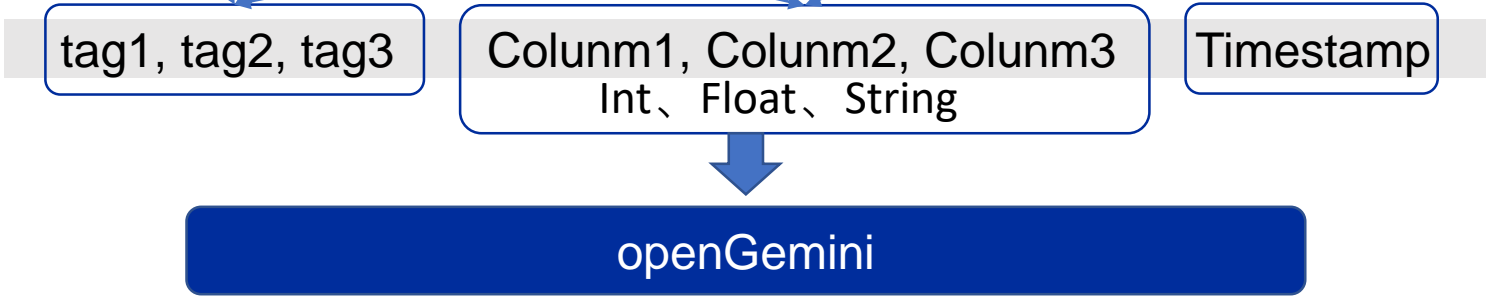# openGemini: High-performance Backend Storage of openTelemetry
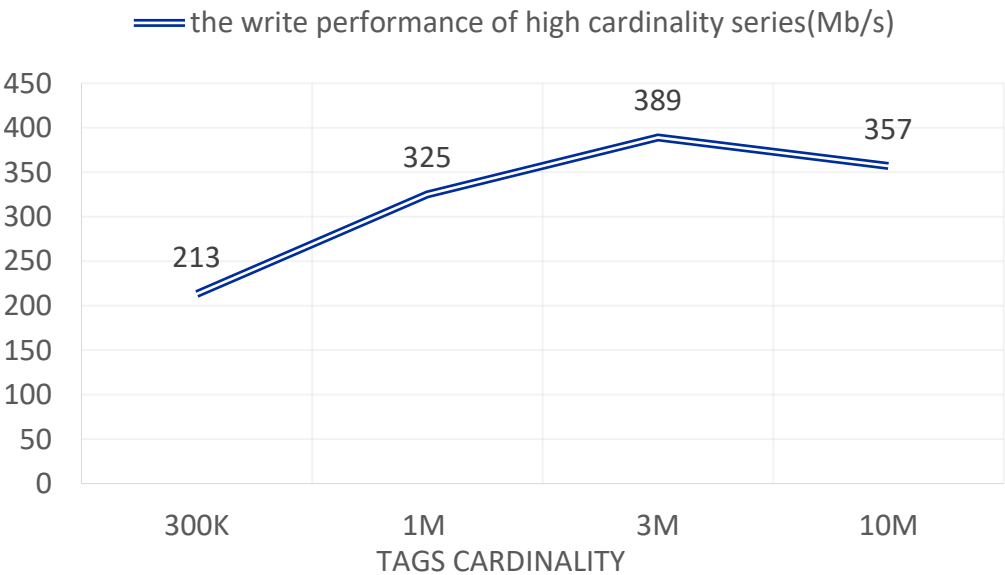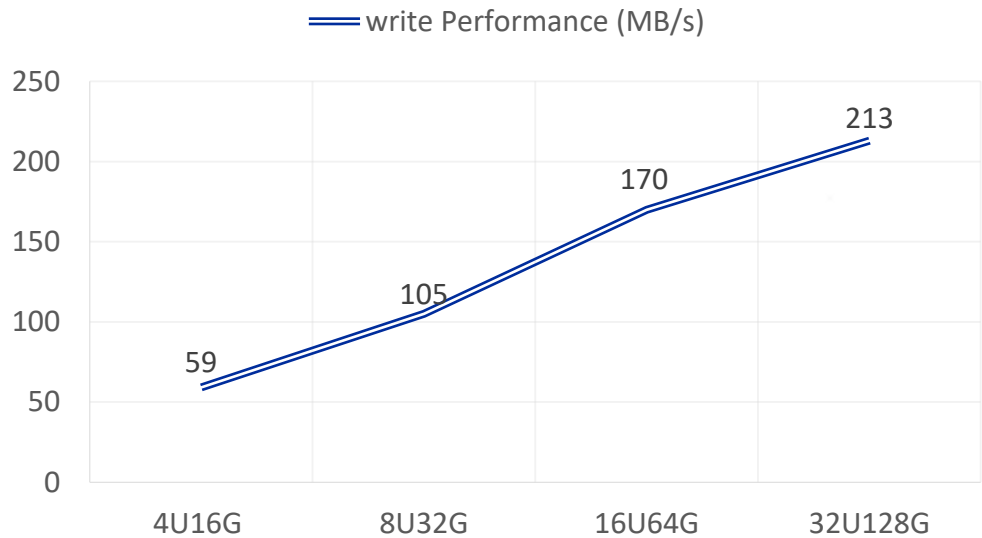
Converts OTEL Format to Storage Schema

OTEL data format sample

Metrics	Traces	Logs

Metrics

```
Resource Attributes
    host.name = "Str(servername)"
    process.command_args = Slice(["./main"])
    process.pid: Int(30653)
    …

Data point attributes
    …
    http.status_code: Int(200)
    http.user_agent: Str(Go-http-client/1.1)
…

Value : 26
```

Traces

```
Span #0 {
    trace_id : 467e3765765a170e881f9c765fb94
    span_id : bf611bd6810518f2
    parentSpanId: "7ab66566adaca823",
    name: "HTTP GET",
    kind: 3,
    startTimeUnixNano: "1679552089108338000",
    endTimeUnixNano: "1679552089653340238",
    Attributes: -> [json]
        -> http.method: Str(GET)
        -> http.flavor: Str(1.1)
        -> http.url: Str(http://0.0.0.0:7080/hello)
        -> net.peer.name: Str(0.0.0.0)
        -> …
```

Logs

```
Resource attributes
    -> host.name: Str(xxx-Pro.local)
    -> process.command_args: Slice(["./main"])
    -> process.executable.name: Str(main)
    …
    -> process.owner: Str(xiangyu)
    -> service.name: Str(demo-server)
    -> process.pid: Int(30653)
    -> telemetry.sdk.language: Str(go)
    -> telemetry.sdk.name: Str(opentelemetry)
    -> telemetry.sdk.version: Str(1.13.0)
trace_id : 467e3a170e881f9d89639bb5c765fb94
span_id : bf611bd6810518f2
Body："Error: connect time out"
```
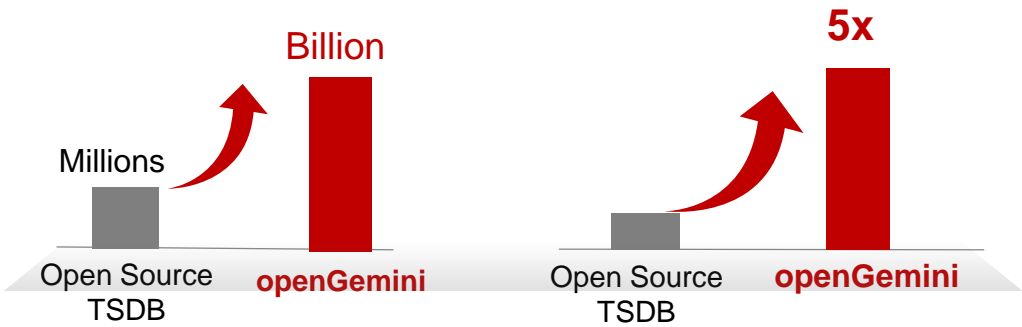
unified data schema

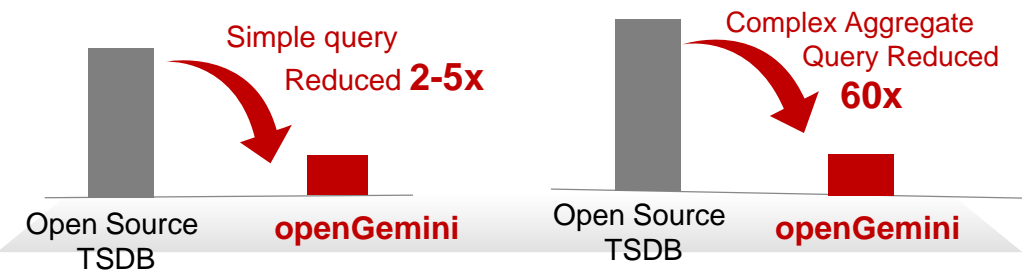| tag1, tag2, tag3	Colunm1, Colunm2, Colunm3 — Int、Float、String	Timestamp

openGemini

Metric Store Evaluation

write Performance (MB/s)

	4U16G	8U32G	16U64G	32U128G
	59	105	170	213

the write performance of high cardinality series(Mb/s)

	300K	1M	3M	10M
	213	325	389	357

TAGS CARDINALITY

Performance

Write

Cardinality

Millions → Billion

Open Source TSDB → openGemini

Write

5x

Open Source TSDB → openGemini

Simple Query latency

Simple query Reduced **2-5x**

Open Source TSDB → openGemini

Complex Query latency

Complex Aggregate Query Reduced **60x**

Open Source TSDB → openGemini

Log Store Evaluation

Index Size

Search Engine vs. openGemini

64.9GB

22.5GB

65%

Memory Space

■ Search Engine ■ openGemini

Index Build Time

Search Engine vs. openGemini

18547S

7390S

60%

Index Time

■ Search Engine ■ openGemini
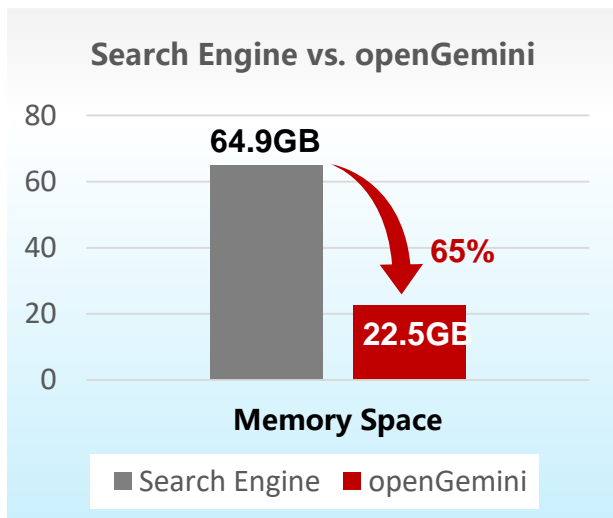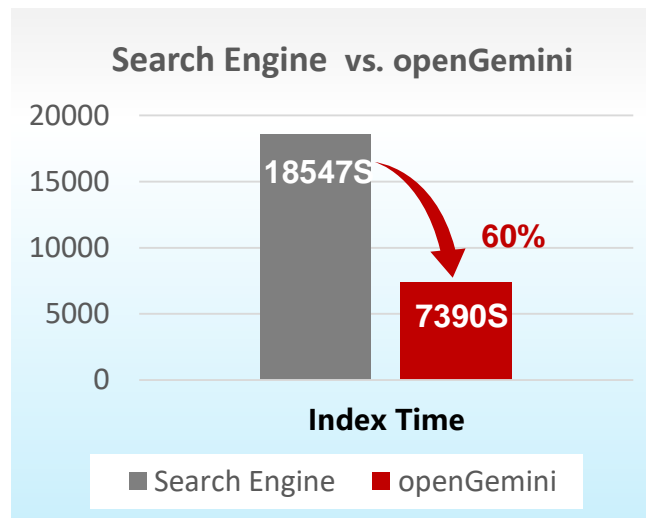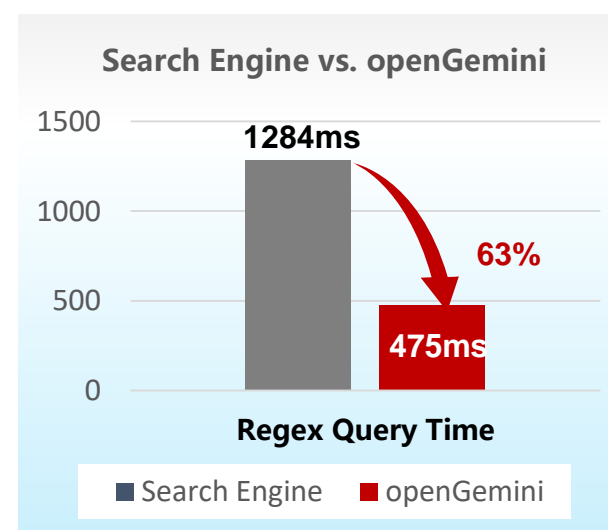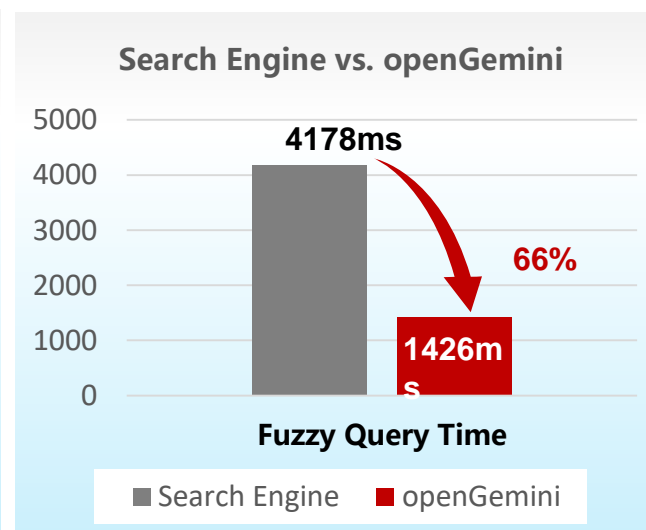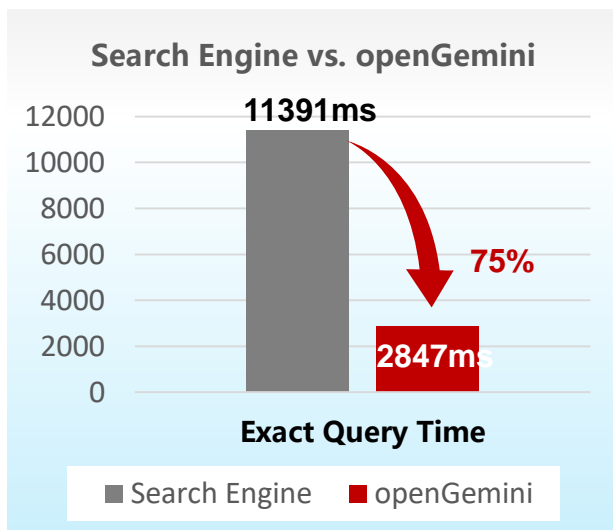
1 billion data test results

Reduces memory index space by **65%** and index build time by **60%**

Under end-to-end scenarios, for average per query, exact query time reduce by **75%** and fuzzy query time reduce by **66%** and regex query time reduce by **63%**

Search Engine vs. openGemini

11391ms

2847ms

75%

Exact Query Time

■ Search Engine ■ openGemini

Search Engine vs. openGemini

4178ms

1426ms

66%

Fuzzy Query Time

■ Search Engine ■ openGemini

Search Engine vs. openGemini

1284ms

475ms

63%

Regex Query Time

■ Search Engine ■ openGemini

Query Types For Logs

Exact query

Fuzzy query

Regex

Prefix/ suffix

score

Average Per Query Time (End-to-end Performance)

Data Format

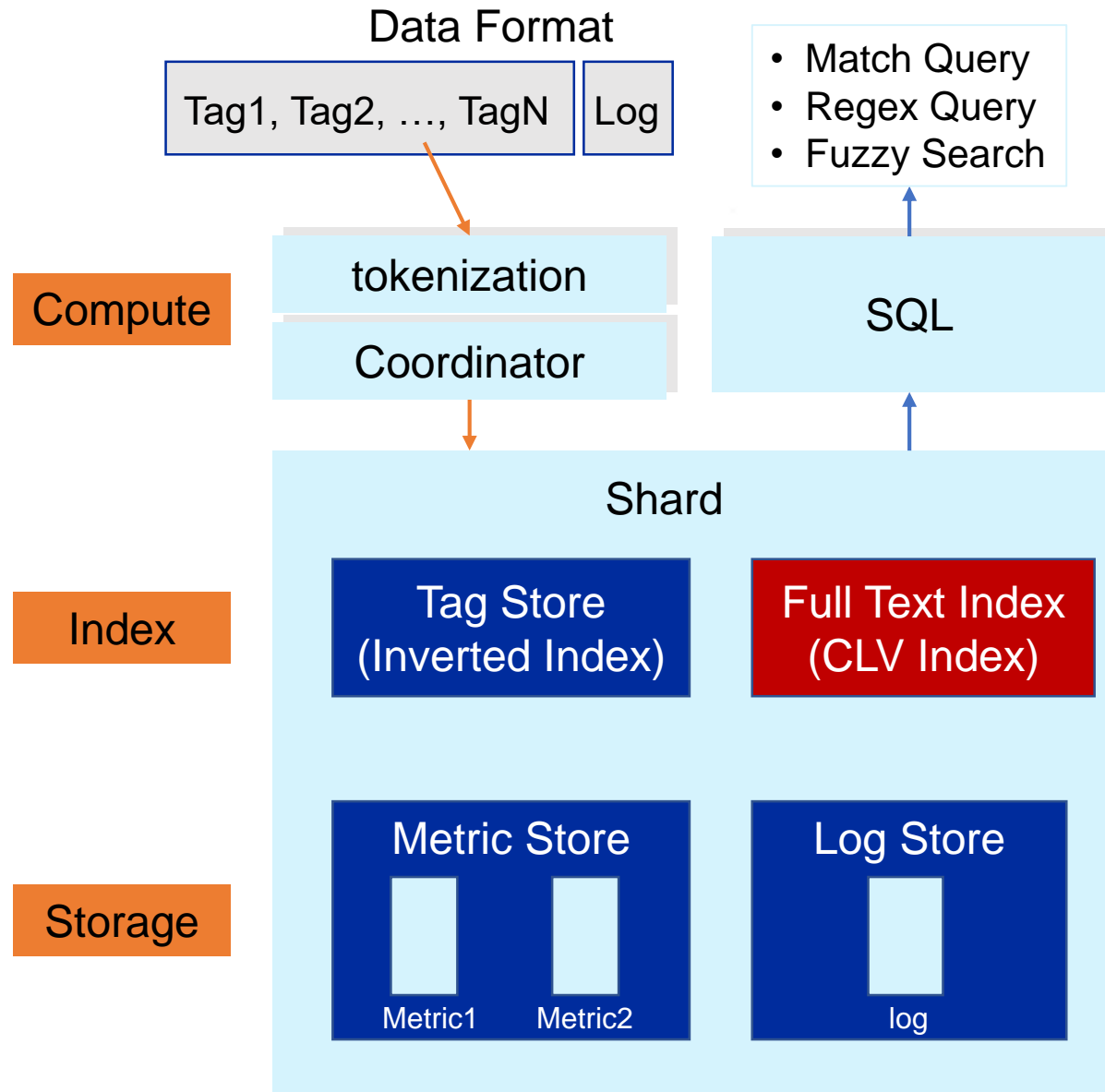| Tag1, Tag2, …, TagN	Trace1, Trace2
Tag1, Tag2, …, TagN	Metric1, Metric2

Indexed   Not Indexed

- Adhoc
- Aggregation
- Timeseries Analyze

Compute

Coordinator

SQL

Shard

Index

Tag Store(Inverted Index)

Term1： Position-list

Term2： Position-list

Storage

Metric Store

Metric1   Metric2
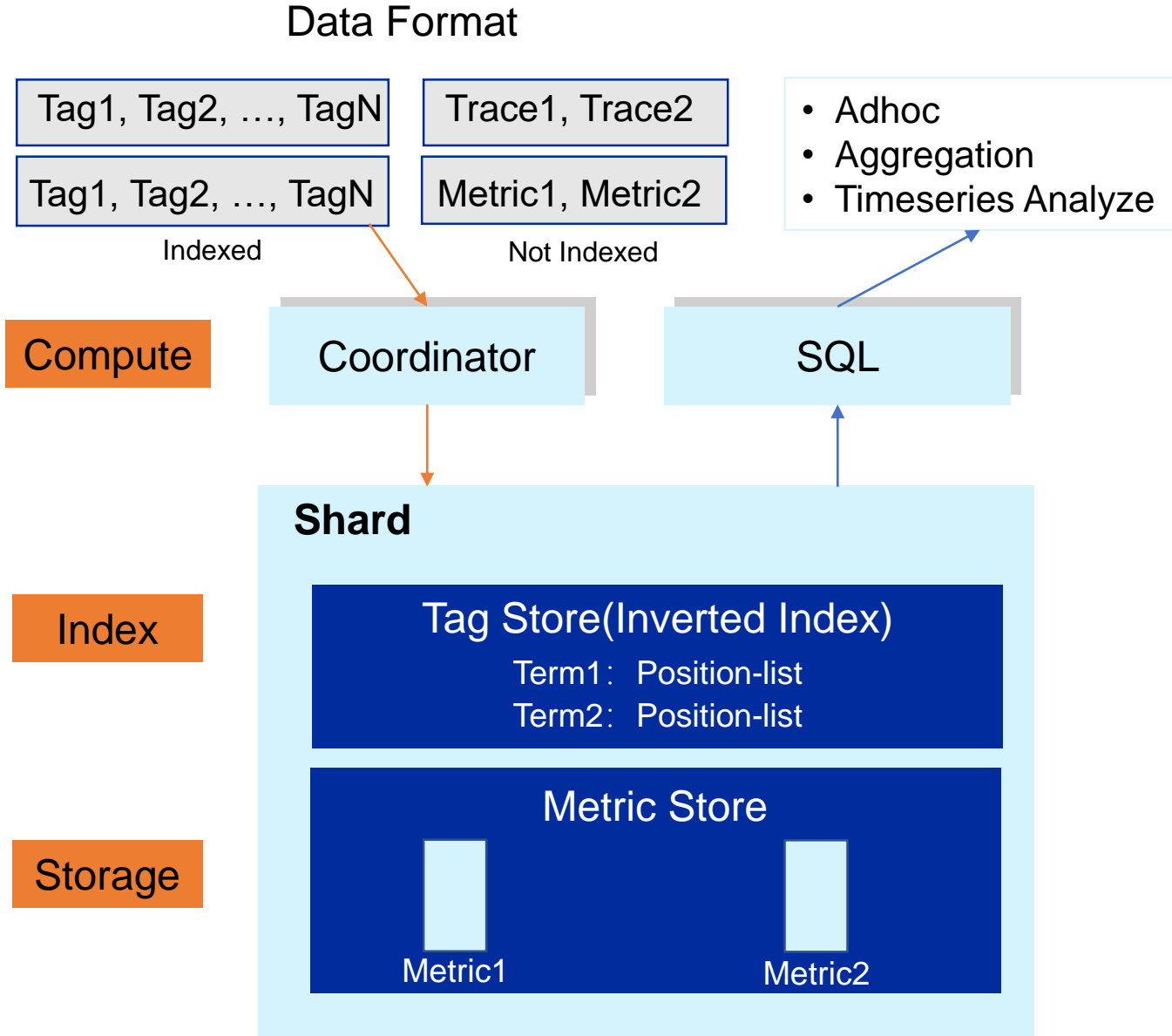
Pros

- Powerful Selectivity Capability

- Easy data compression

Cons

- High Cardinality Tag

- Capability for Logs

Data Format

Tag1, Tag2, …, TagN | Log

- Match Query
- Regex Query
- Fuzzy Search

Compute

tokenization

Coordinator

SQL

Index

Shard

Tag Store (Inverted Index)

Full Text Index (CLV Index)

Storage

Metric Store

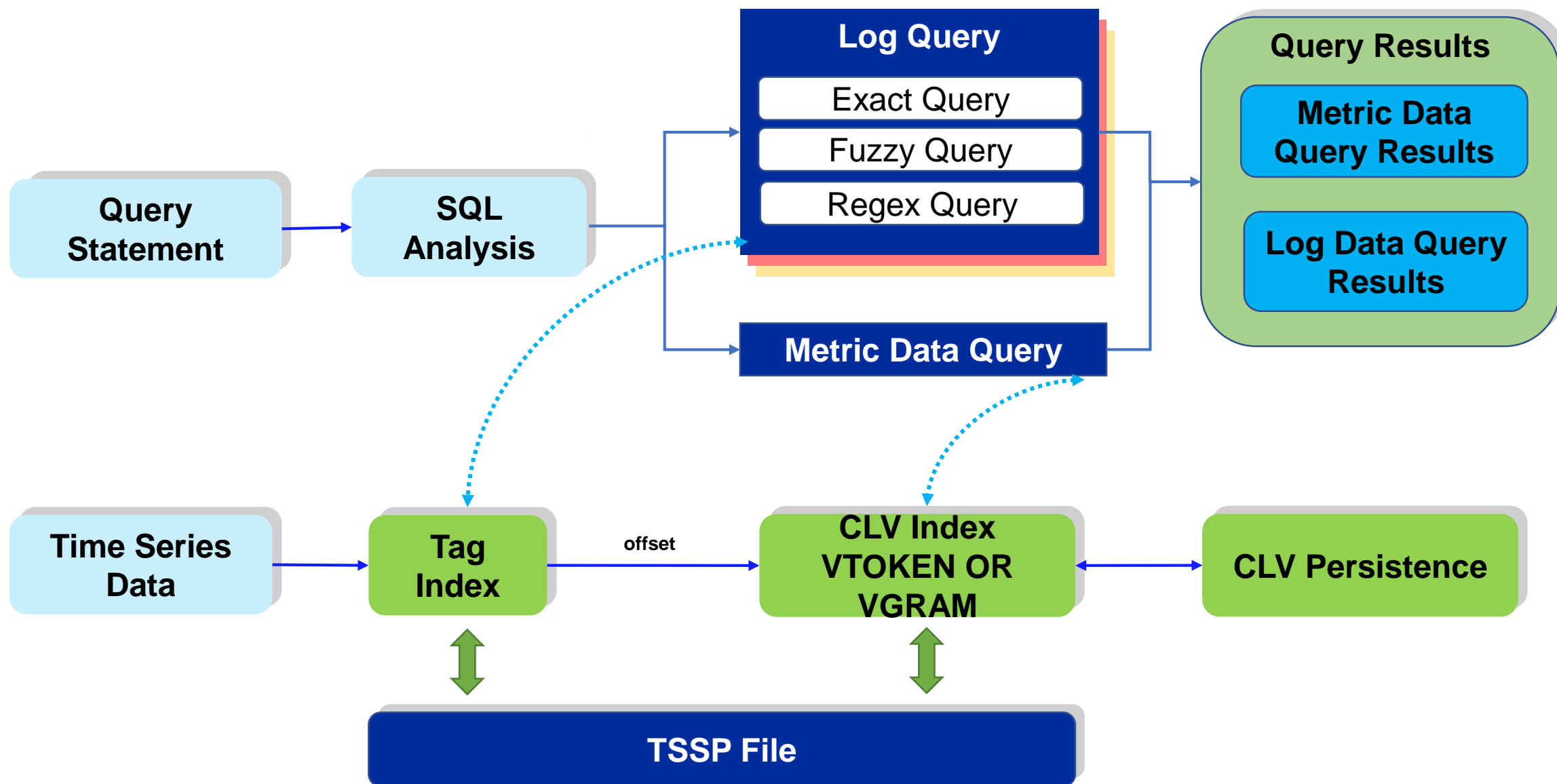Metric1   Metric2

Log Store

log

Tag Index + Full Text Index

Column Store
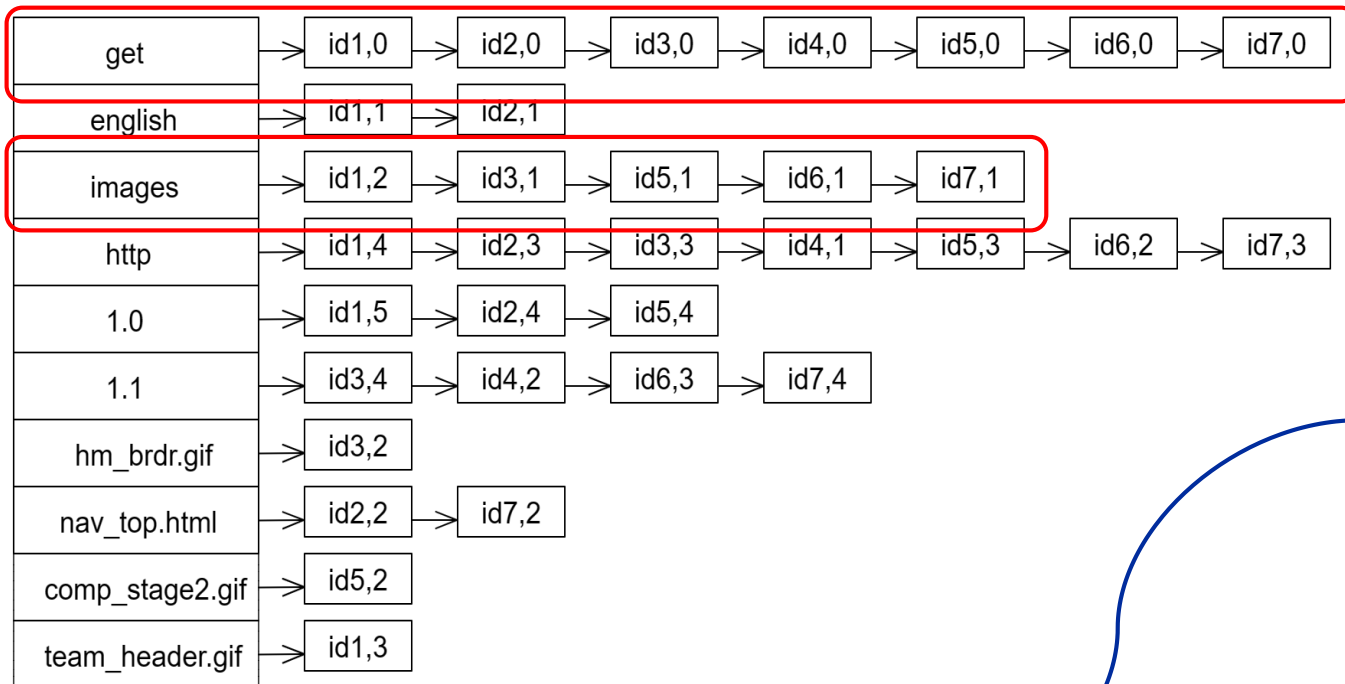
One Engine For Metrics And Logs

CLV Index: Full Text Search Index For Log

CLV Index: Compressed Indices for Efficient Search

Inverted index

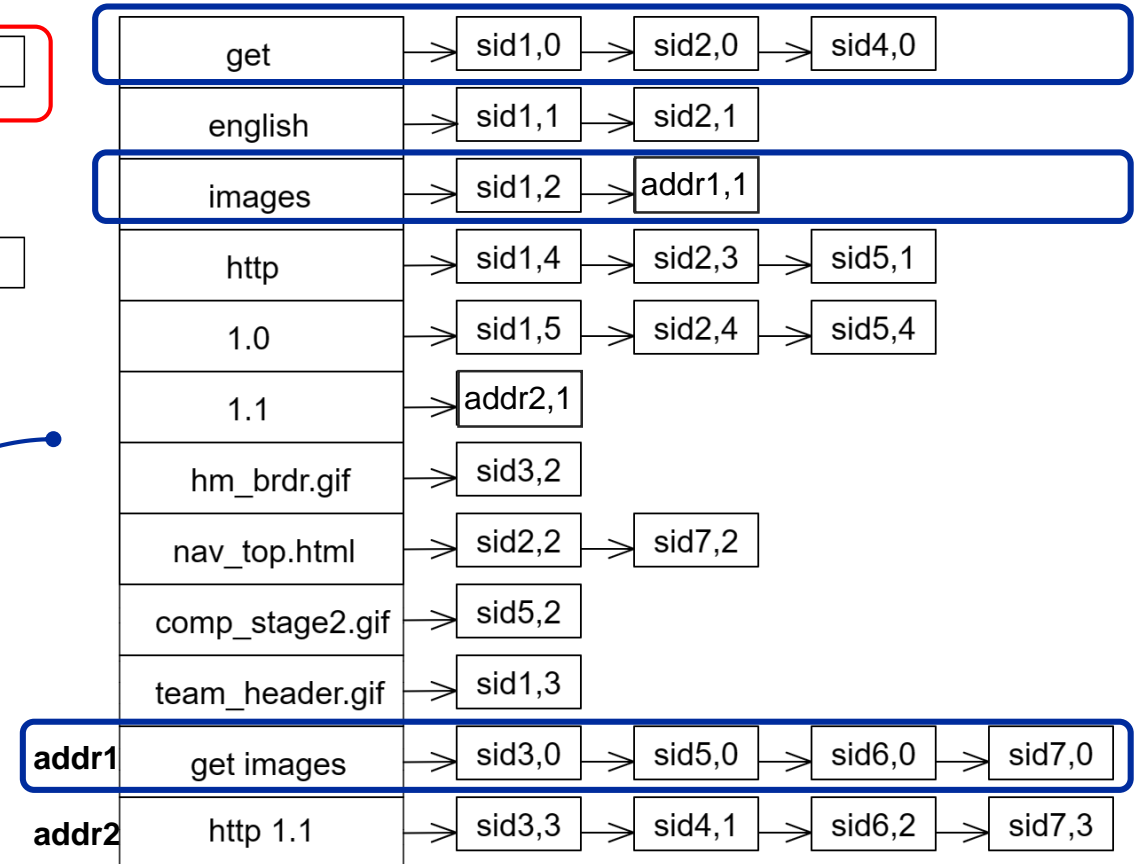Number of keys: **10**   Index list size: **33**

Compressed inverted index

Number of keys: **12**   Index list size: **25** + **2** pointers

Compressed Inverted Index
- Less index space
- Efficient query time

OpenGemini Evolution Roadmap

Observability Backend Store

- Tracing data storage
- High cardinality optimization
- Metric、Log、Tracing Correlation Analysis

3

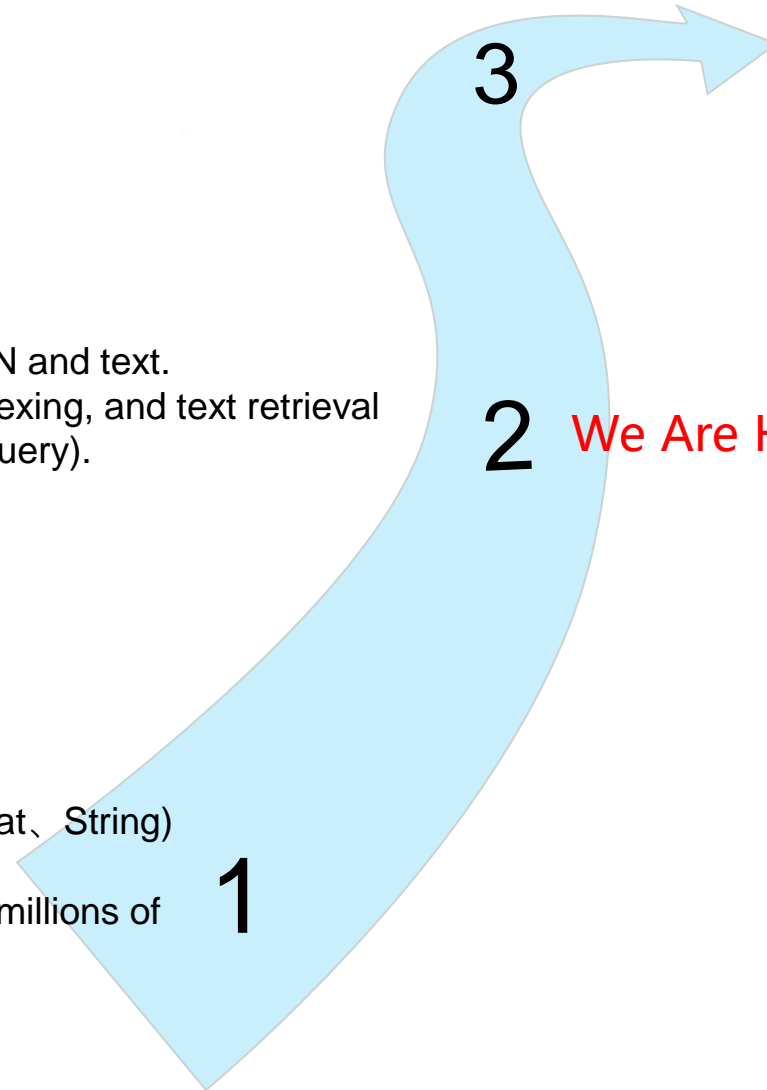**Log Store**

- More Data types such as JSON and text.
- Supports tokenization, text indexing, and text retrieval (exact match 、 fuzzy match query).
- Supports the OTEL protocol.

2 We Are Here

Metric Store

- Metric data storage and query.(Int、Float、String)
- SQL-Like.
- PB-level data storage and hundreds of millions of time series management.

1

Welcome to try and Give Feedback

```
docker run -d -p 8086:8086 --name openGemini-dev opengeminidb/opengemini-server
```

https://github.com/openGemini

https://join.slack.com/t/opengemini/shared_invite/zt-1sncd7nwn-9Mmq2bPVt3b0DP3ng2NOHA

https://twitter.com/openGemini   http://www.openGemini.org

Learn more about openTelemetry and openGemini

Welcome to Booth：D7

Please scan the QR Code above
to leave feedback on this session