



BUILDING FOR THE ROAD AHEAD

DETROIT 2022

Cortex – Now and Future

Alvin Lin, AWS
Alan Protasio, AWS

The New Maintainers



The releaser

Alvin Lin
Software Dev Manager
AWS



The New Maintainers



Alan Protasio
Software Engineer
AWS

The coder and gatekeeper



The New Maintainers



Friedrich Gozalez
Software Engineer
Adobe

The community manager
and user advocate



Agenda

- Quick Introduction to Cortex
- Coming Soon
- One Operational Tip
- Looking Back
- Voice of the Community
- Call to Actions
- Q&A



What is Cortex?



Horizontally scalable, highly available, multi-tenant, long term storage for Prometheus.



What is Cortex

- CNCF incubating project
- Apache 2.0 License

Contributors 243

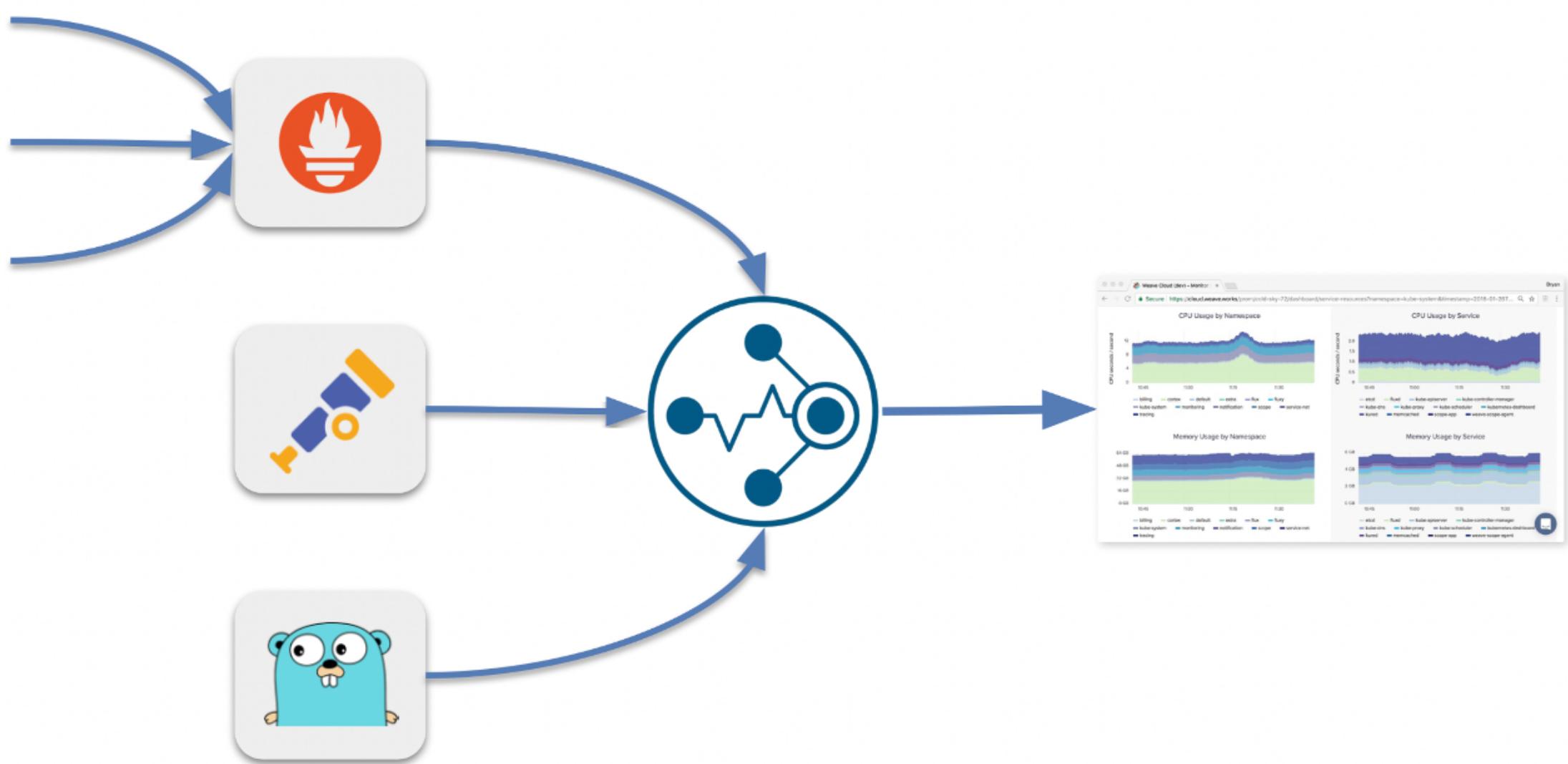
 Watch 113 ▾

 Star 4.9k ▾

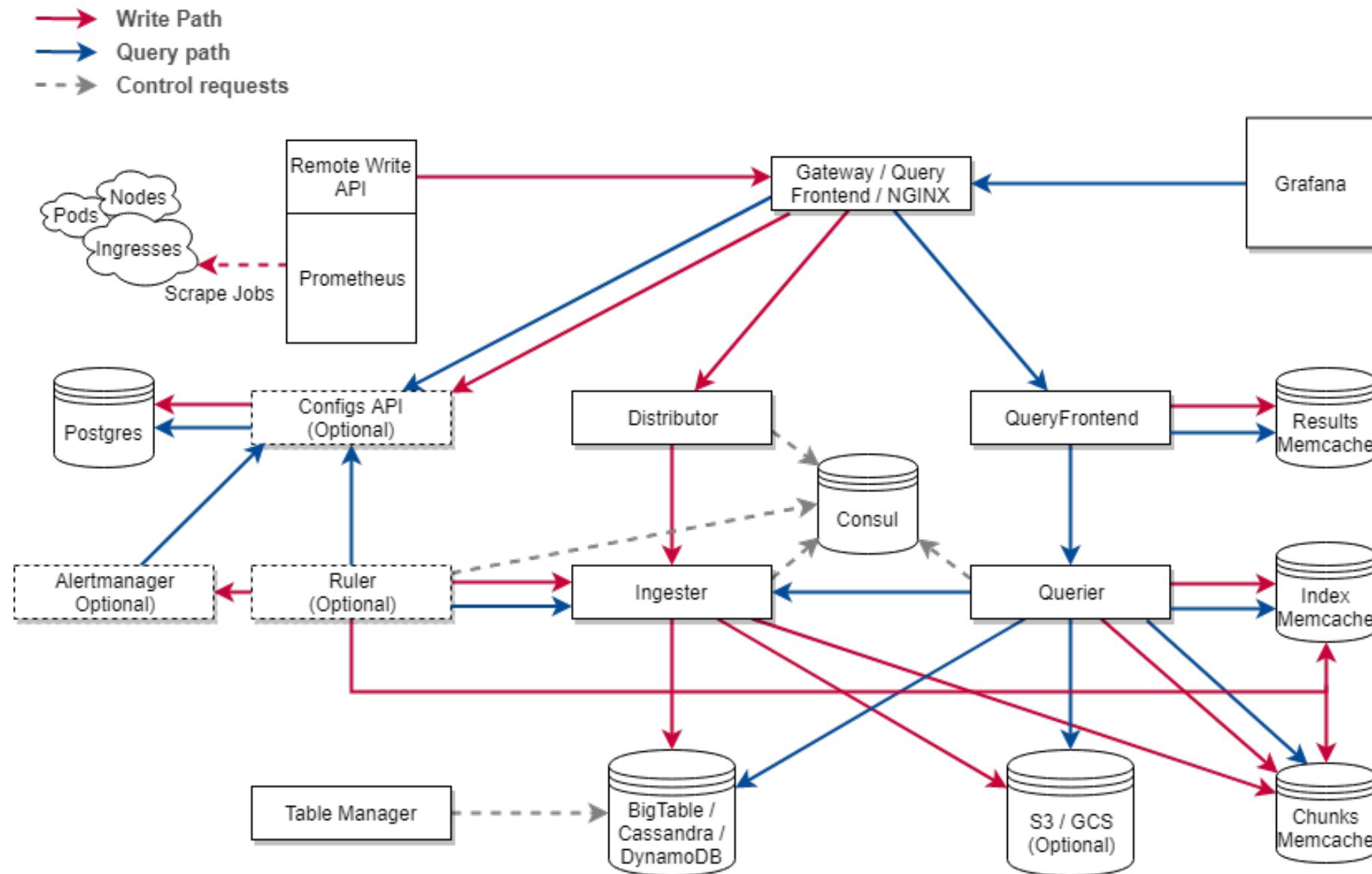
 4,484 commits



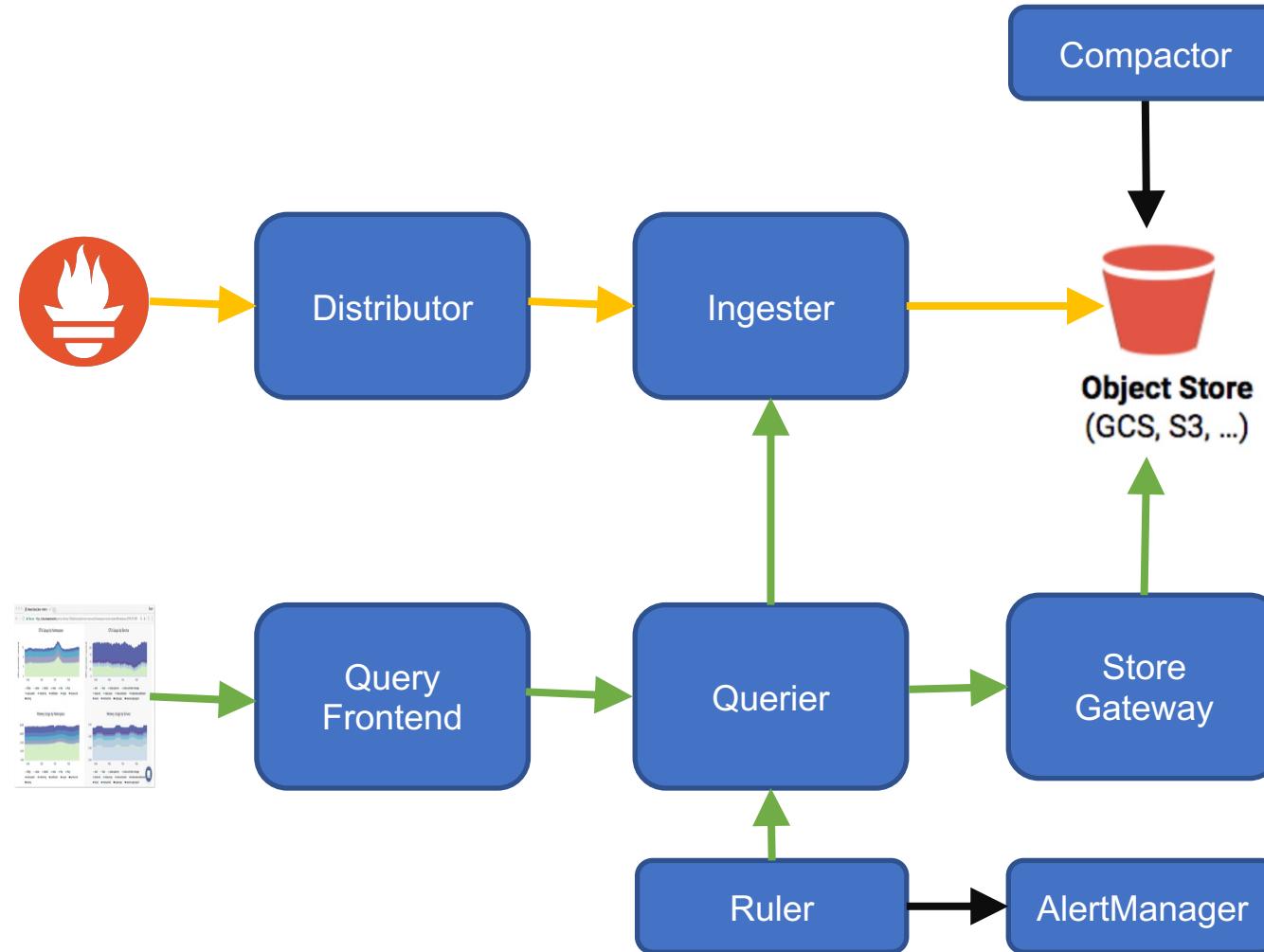
What is Cortex?



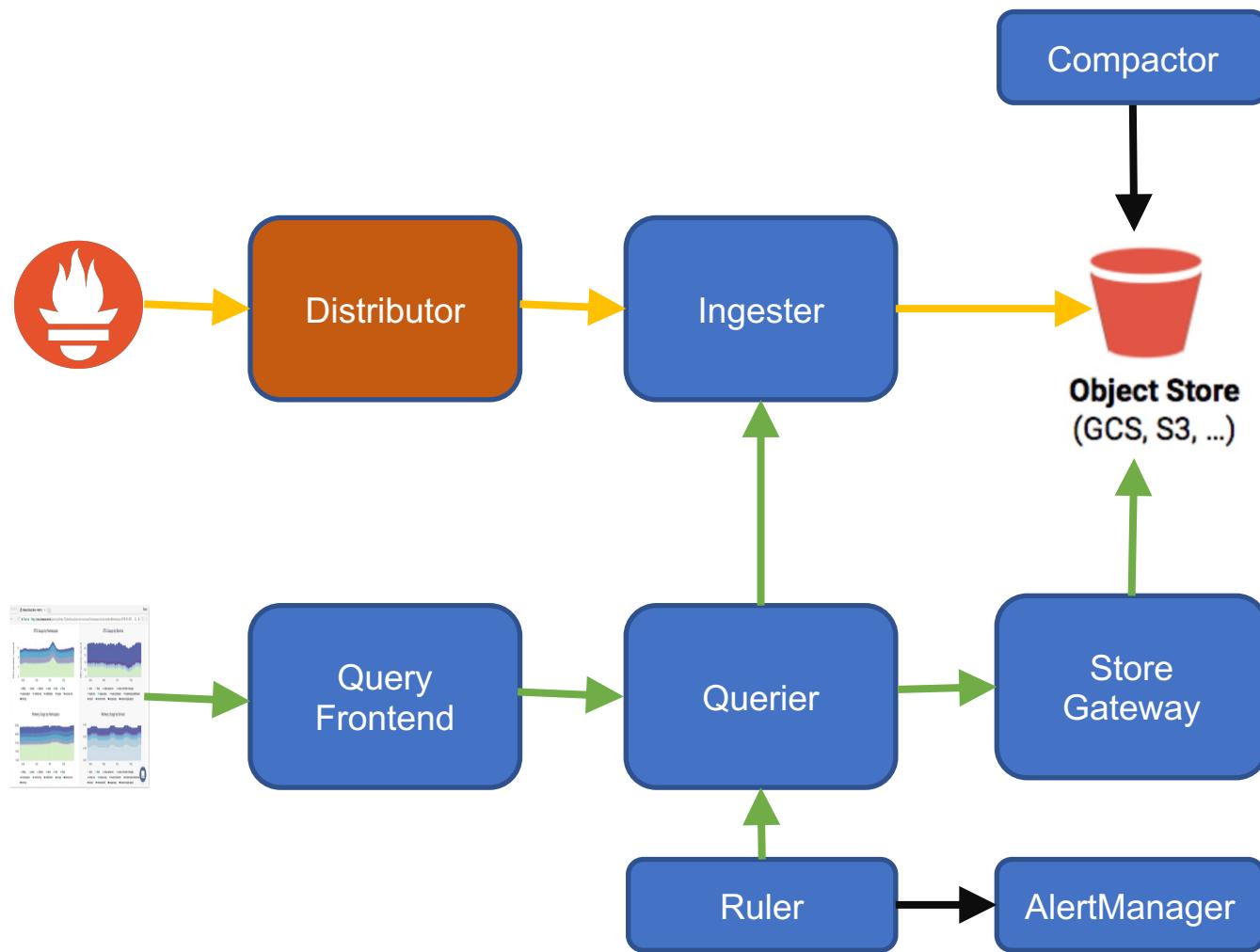
What is Cortex?



Cortex Components



Cortex Components

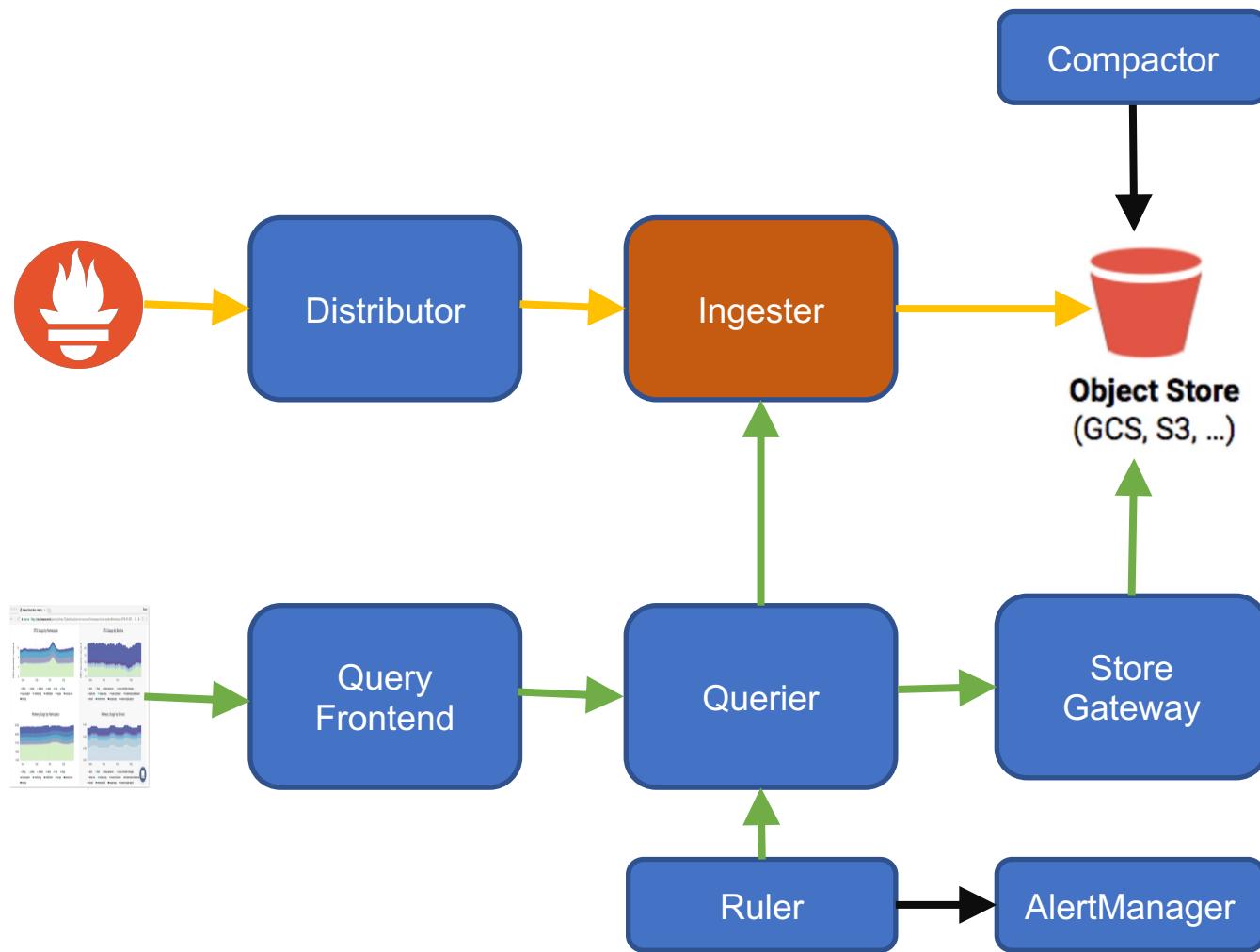


Distributors Key Responsibilities:

- RemoteWrite entrypoint
- Shard and replicate series between ingestors
- **Shuffle sharding (tenant isolation)**
- **Zone Awareness Replication**
- Rate Limiting
- HA Dedup



Cortex Components

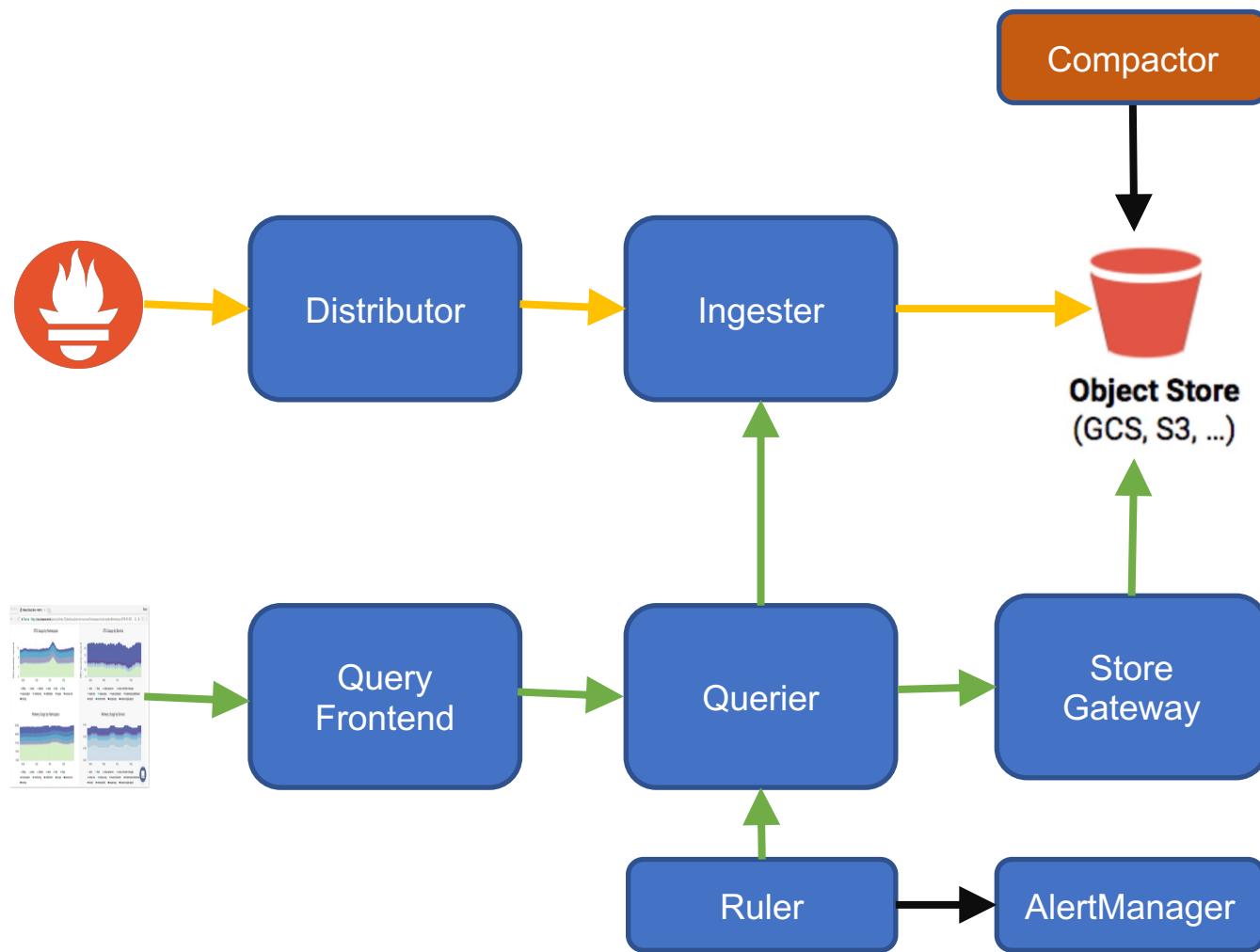


Ingestors Key Responsibilities:

- Manage and store multiple TSDBs instances on local disk.
- Receive samples from distributors and append them to the respective tenant TSDB instance.
- Periodically upload the TSDB blocks to the long term storage (S3, GCS).



Cortex Components

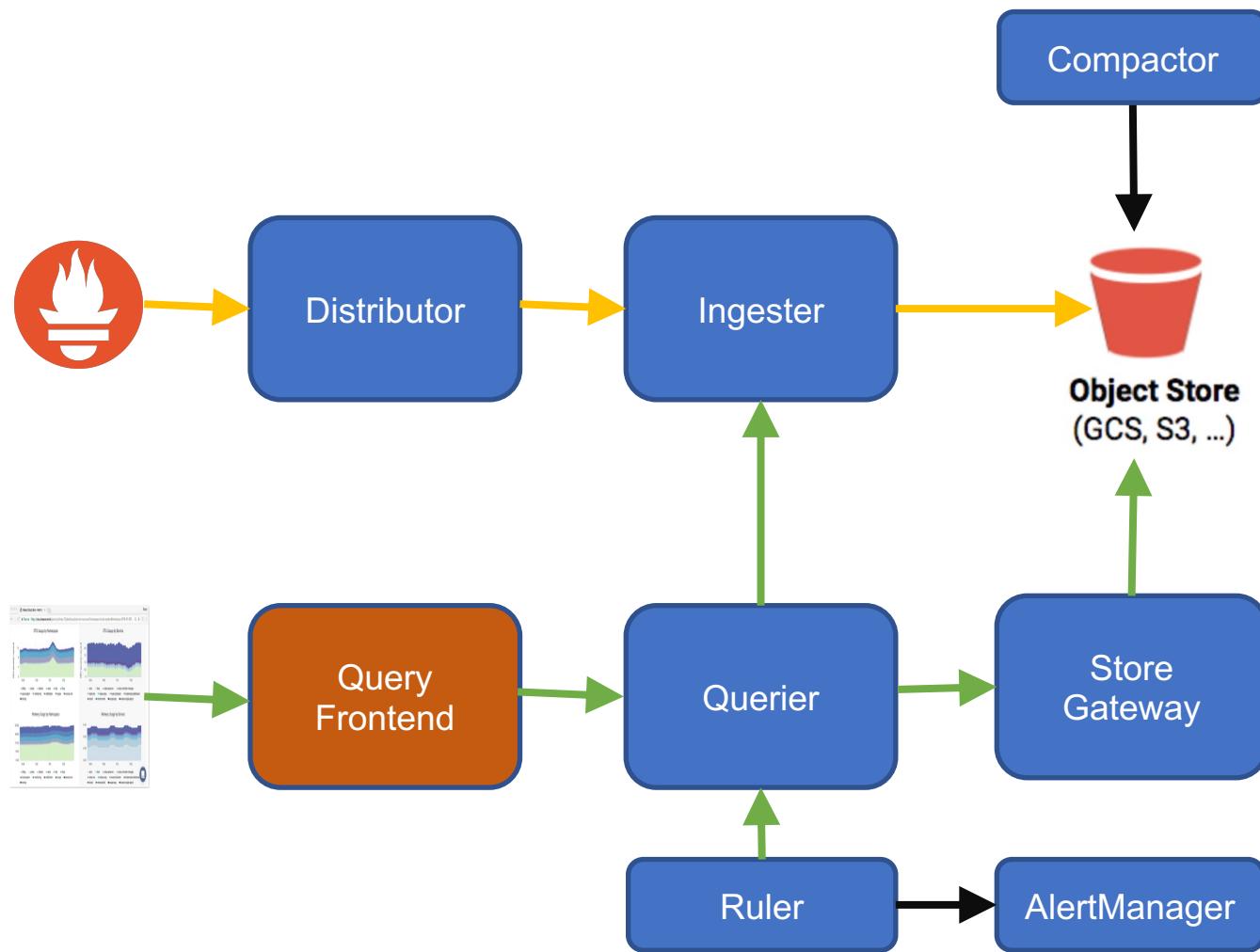


Compactor:

- Merge and deduplicate smaller blocks into larger ones.
- Reduce the number of blocks on the long term storage.
- Improve Query Efficiency



Cortex Components

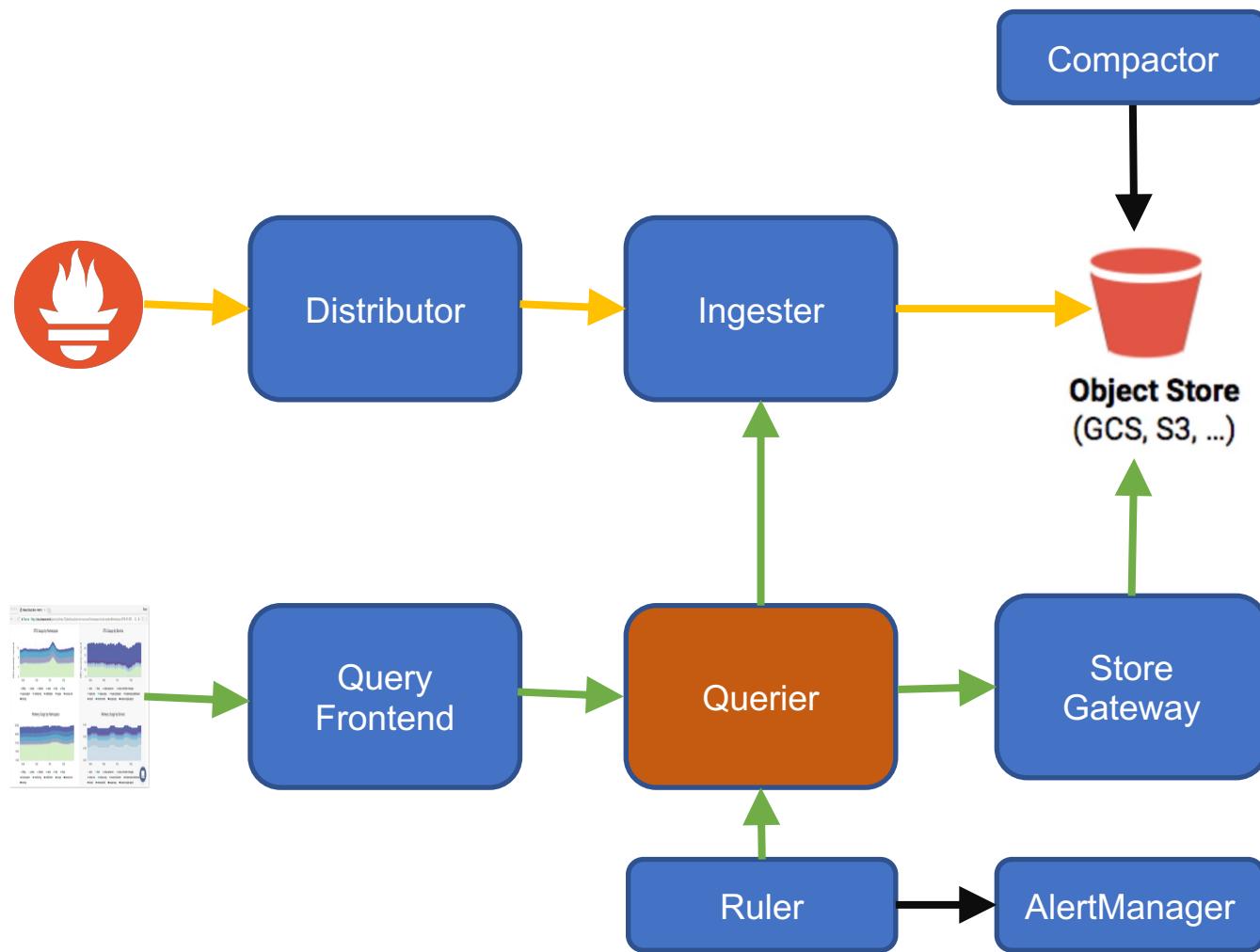


Query Frontend Key Responsibilities:

- QoS.
- Shuffle sharding / Zone Awareness
- Cache query results.
- Implements vertical and horizontal query sharding.



Cortex Components

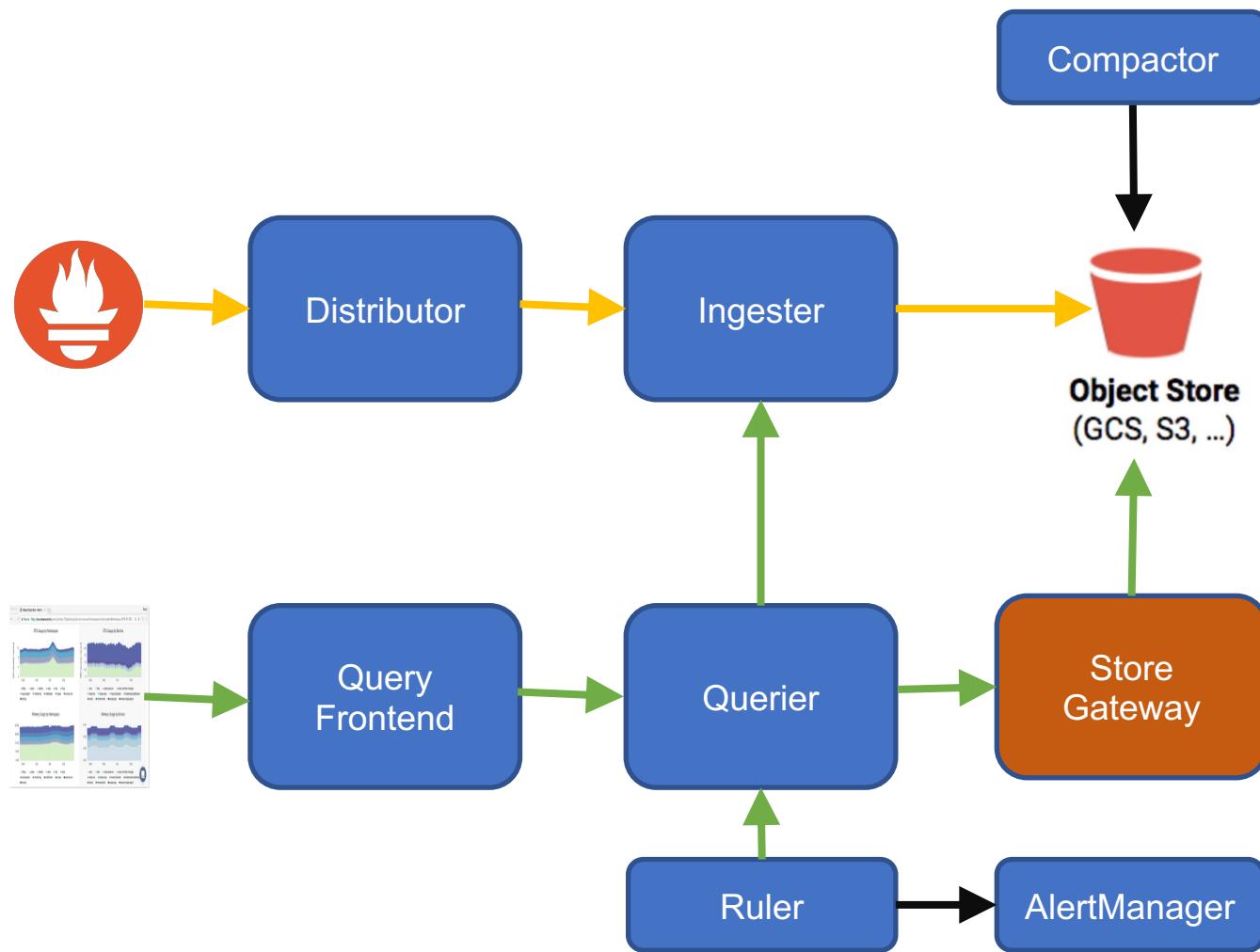


Querier Key Responsibilities:

- Execute query requests from the Query Frontend.
- Fetch data from ingestors (recent) or store-gateways (historical data) according to the query parameters.
- Apply query limits.



Cortex Components

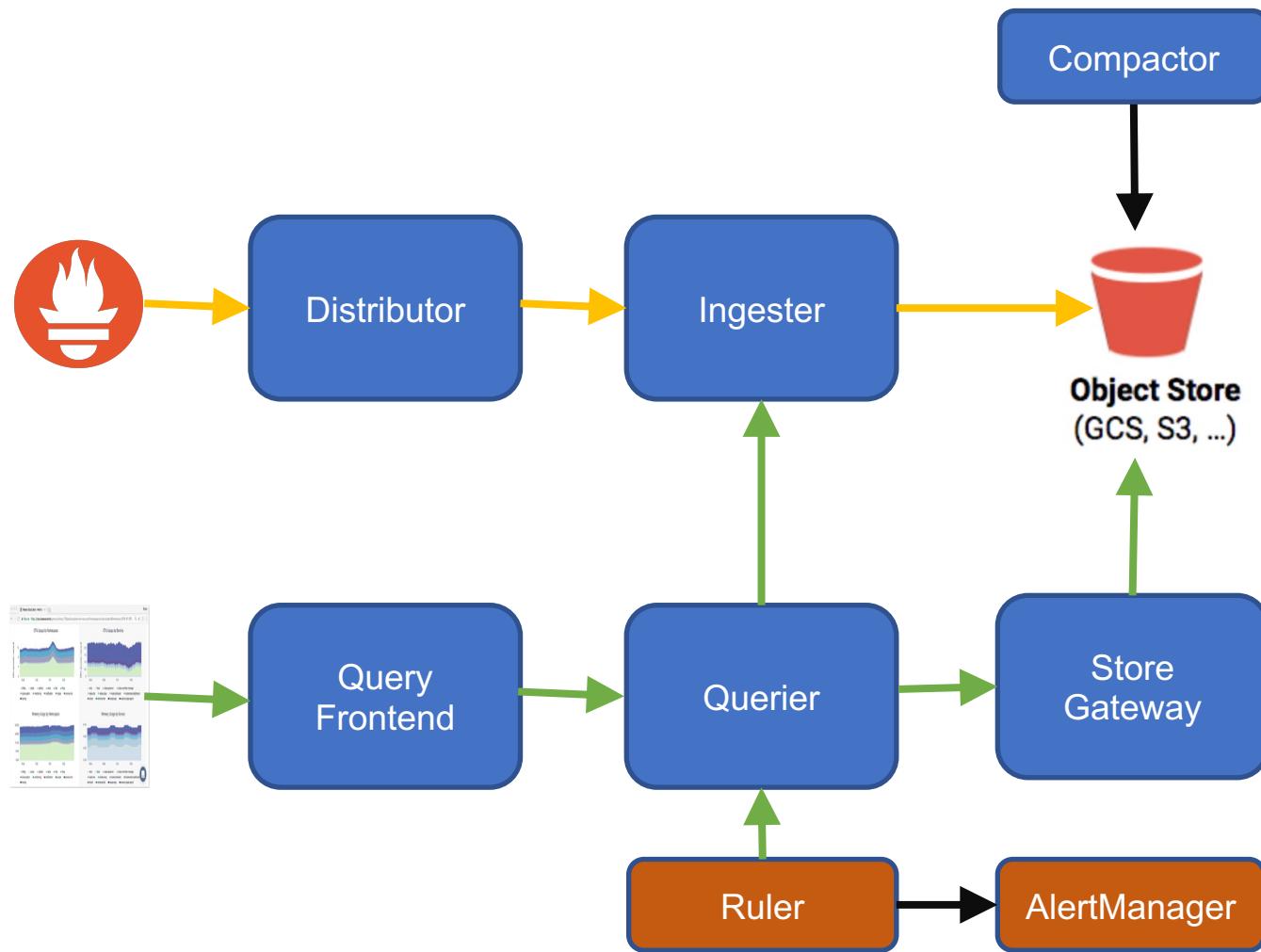


Store Gateway Key Responsibilities:

- Gateway between the querier and the long term storage.
- Keep an up-to-date view of the blocks on the storage bucket.
- Keep specific pieces of blocks' index on local disk to speed up the query execution.



Cortex Components



Ruler Key Responsibilities:

- Evaluate recording and alerting rules.
- Generate Alerts notifications to AlertManager

AlertManager Key Responsibilities:

- Receives alert notifications from the ruler.
- Groups and dedups alerts
- Route alerts to the correct notification channel, such as email, PagerDuty or OpsGenie.



Who Uses Cortex



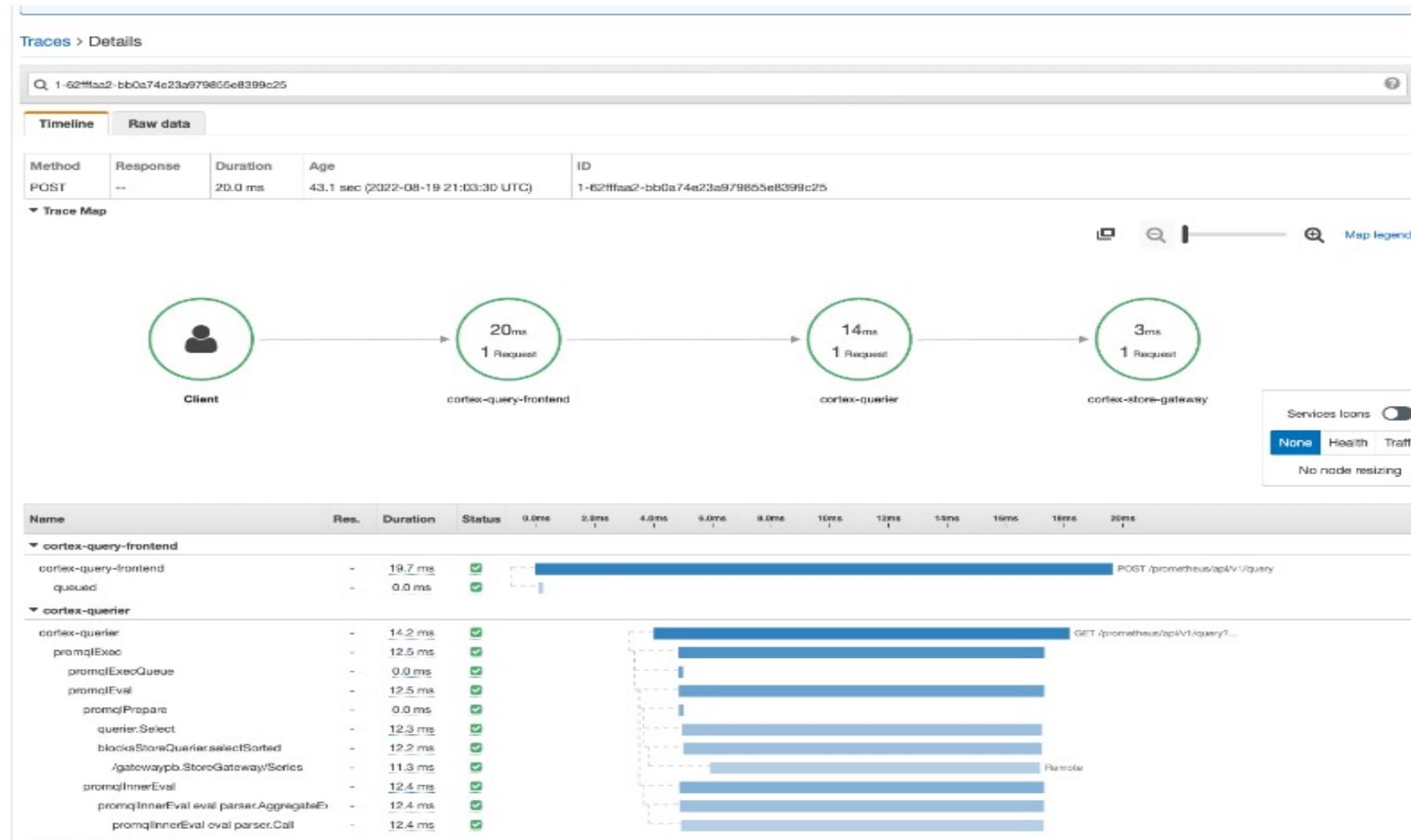
Etsy



Coming Soon – OTEL Bridge for Tracing



cortex



Doc: https://cortexmetrics.io/docs/configuration/configuration-file/#tracing_config

<https://cortexmetrics.io/>



Coming Soon - OTEL Bridge for Tracing

- Available in main branch
- Gives an user more choices for their trace destinations like Jaeger, Zipkin, Kafka, and AWS X-Ray
- Open standard; CNCF project



Coming Soon – Partitioning Compactor



Coming Soon – Partitioning Compactor

- Work in progress; not yet in main branch (almost done ☺)
- No more block index is bigger than 64GB issue¹
- Compactor downloads less source blocks when doing compaction
- 50% compaction time reduction for 200 millions time series

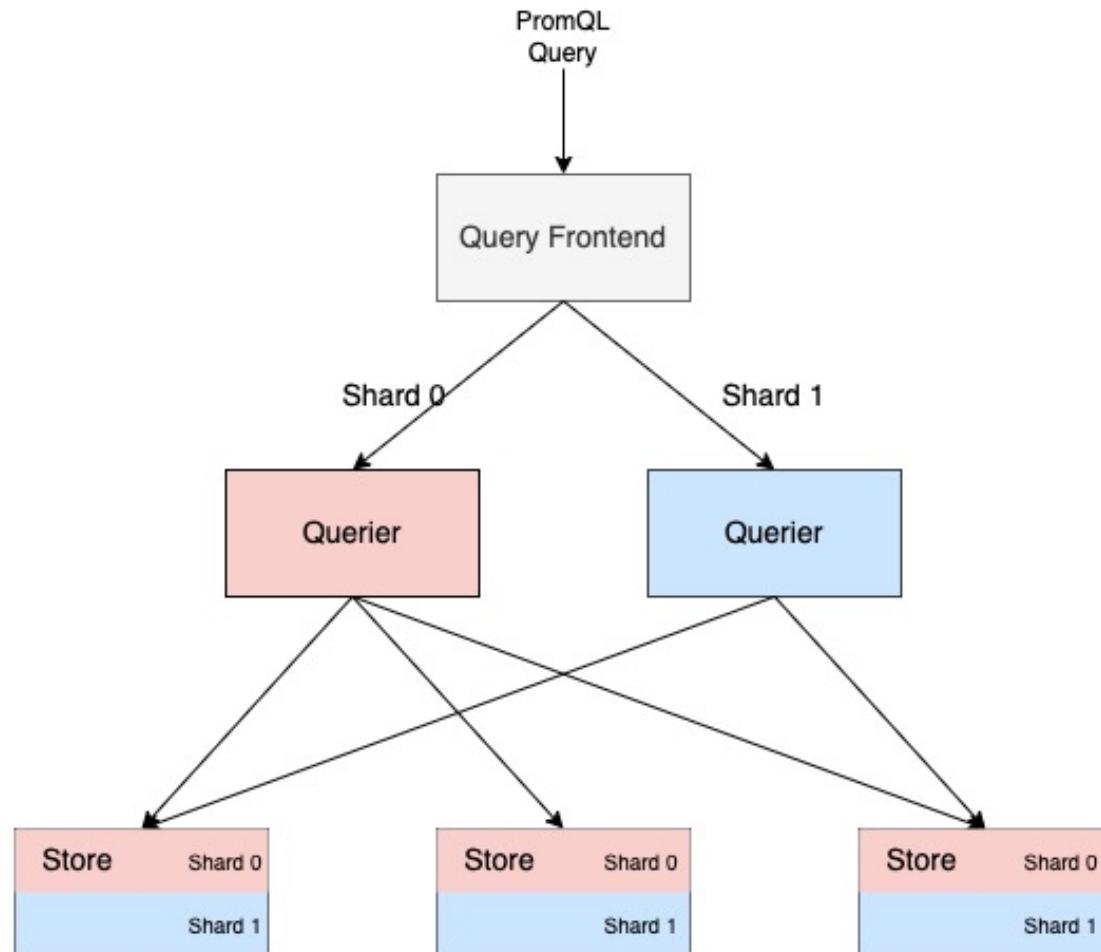
1. <https://github.com/prometheus/prometheus/issues/5868>



Coming Soon - Vertical Query Sharding



cortex



Credit: <https://thanos.io/tip/proposals-accepted/202205-vertical-query-sharding.md>
<https://cortexmetrics.io/>



Coming Soon - Vertical Query Sharding

- Experimental feature
- Available in main branch
- Shardable query's speed improved 30%+
- Enable through `-frontend.query-vertical-shard-size` or `query_vertical_shard_size`
- Doc: <https://cortexmetrics.io/docs/configuration/v1guarantees/#experimental-features>



Deep Dive – Vertical Query Sharding

We have following metrics:

```
number_of_users{region="NA", pod="1"}  
number_of_users{region="NA", pod="2"}  
number_of_users{region="NA", pod="3"}  
number_of_users{region="EU", pod="1"}  
number_of_users{region="EU", pod="2"}  
number_of_users{region="EU", pod="3"}
```



Deep Dive – Vertical Query Sharding

sum by (region) (number_of_users)

Time	region	Value
2022-10-28 11:55:00	NA	100
2022-10-28 11:55:00	EU	95



Deep Dive – Vertical Query Sharding

Without vertical sharding, a single pod executes the query. With vertical sharding, multiple pods can execute the query.

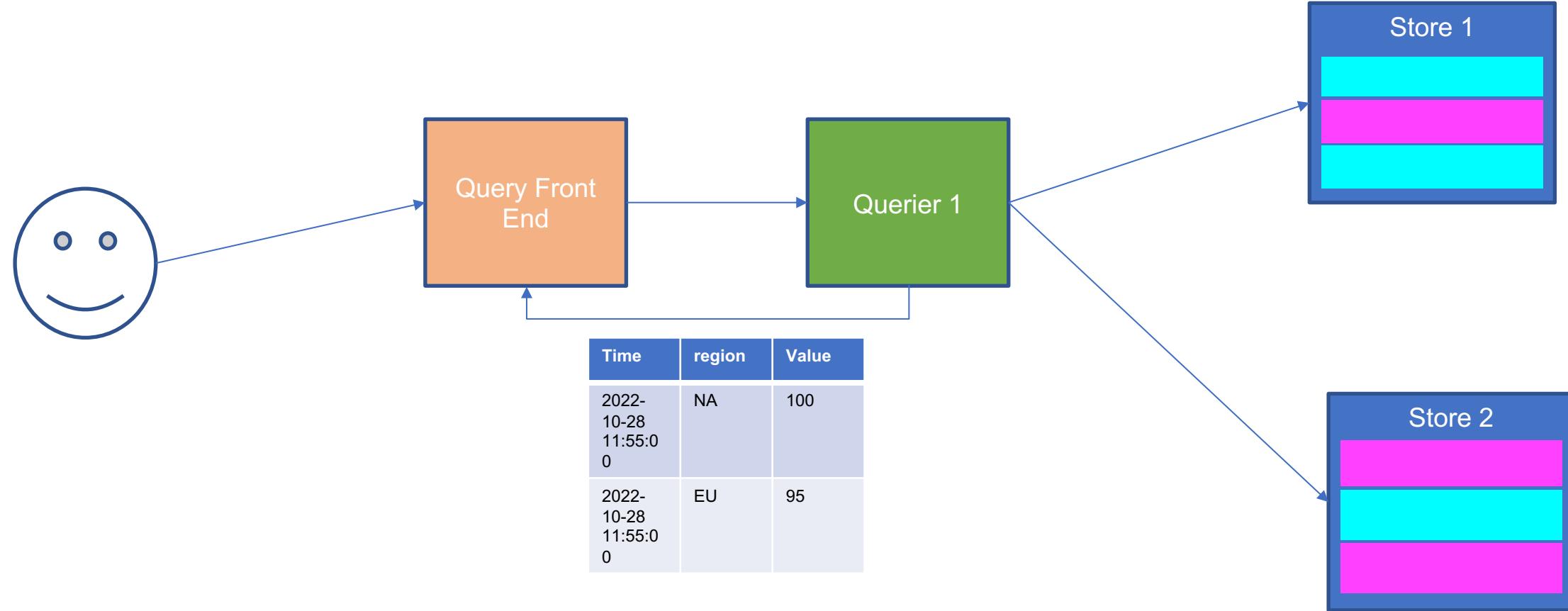
```
number_of_users{region="NA", pod="1"}  
number_of_users{region="NA", pod="2"}  
number_of_users{region="NA", pod="3"}
```

```
number_of_users{region="EU", pod="1"}  
number_of_users{region="EU", pod="2"}  
number_of_users{region="EU", pod="3"}
```



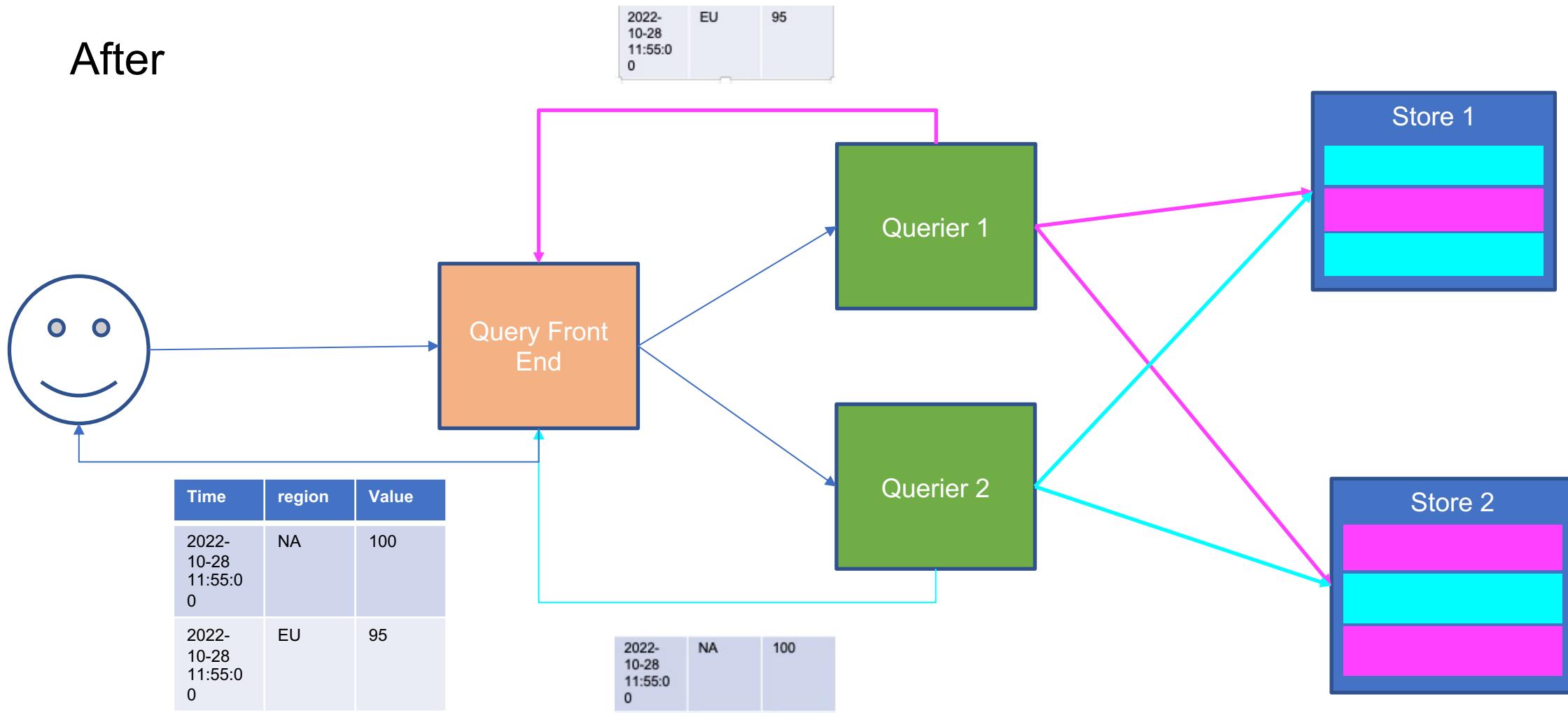
Deep Dive – Vertical Query Sharding

Before



Deep Dive – Vertical Query Sharding

After



One Operational Tip

- Enable streaming
- querier.ingester-streaming
- querier.ingester-metadata-streaming
- Reduces memory footprint



Looking Back

- Release 1.13 had many enhancement on July 2022 with memory and performances improvements.
- Forked Cortex related repos (e.g. cortex-tools) from Grafana's GitHub org to Cortex's org for ease of future development
- Shout out to Alan and Friedrich for stepping up when Cortex needs them the most



From a Long Time Cortex User ...

- At least 99.9% of the remote write requests are less than 1 second
- At least 99.9% of the query requests are less than 1.3 seconds
- Upgrade from version 0.6.1 to 1.13.0 was easy
- With 1.13.0 release, 150 millions active time series was piece of cake (32 millions were challenge before)



Voice of the Community

- Out of order samples ingestion
- Down sampling
- Per-tenant encryption
- Delete time series
- Much more! Please tell us what you want!



Voice of the Community

- Join the [Cortex's Slack Channel](#) and let's chat
- Upvote issues that are important to you with “”



Interested In Contributing?

Filters ▾ Q is:open is:issue label:"good first issue" Labels 61 Milestones 1 New issue

Clear current search query, filters, and sorts

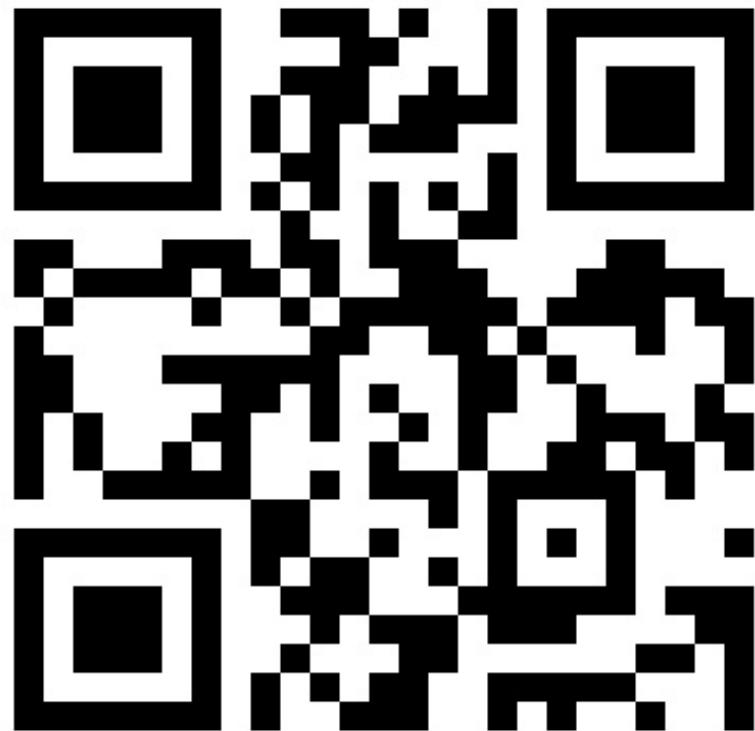
Author ▾	Label ▾	Projects ▾	Milestones ▾	Assignee ▾	Sort ▾
<input type="checkbox"/> 7 Open	<input checked="" type="checkbox"/> 24 Closed				
<p><input type="checkbox"/> Publish and test arm docker images for every PR good first issue keepalive 6</p> <p>#4733 opened on May 9 by friedrich-at-adobe</p>					
<p><input type="checkbox"/> Unit tests leak goroutines good first issue help wanted type/tests 7</p> <p>#3962 opened on Mar 17, 2021 by stevesg</p>					
<p><input type="checkbox"/> Slow unit tests good first issue help wanted type/tests 2</p> <p>#3922 opened on Mar 8, 2021 by pstibrany</p>					
<p><input type="checkbox"/> Improve HA tracker documentation component/documentation good first issue hacktoberfest help wanted 1</p> <p>#3011 opened on Aug 11, 2020 by pracucci</p>					
<p><input type="checkbox"/> Update remote read to the new streaming interface. good first issue help wanted 4</p> <p>#2495 opened on Apr 21, 2020 by tomwilkie</p>					
<p><input type="checkbox"/> Don't log "repeated timestamp but different value" message as error/warning good first issue hacktoberfest 11</p> <p>help wanted type/chore</p> <p>#1478 opened on Jun 27, 2019 by weeco</p>					
<p><input type="checkbox"/> Worry more about fingerprint clashes component/querier good first issue help wanted type/bug 4</p> <p>#717 opened on Feb 19, 2018 by bboreham</p>					



Call To Actions

- More contributors and maintainers
- Help to maintain Helm Chart
- Help to maintain <https://cortexmetrics.io>
- Help to improve Cortex documentation
- Engage with <https://github.com/thanos-community/promql-engine>
- Automated performance benchmarking framework that can be kicked off through PRs





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