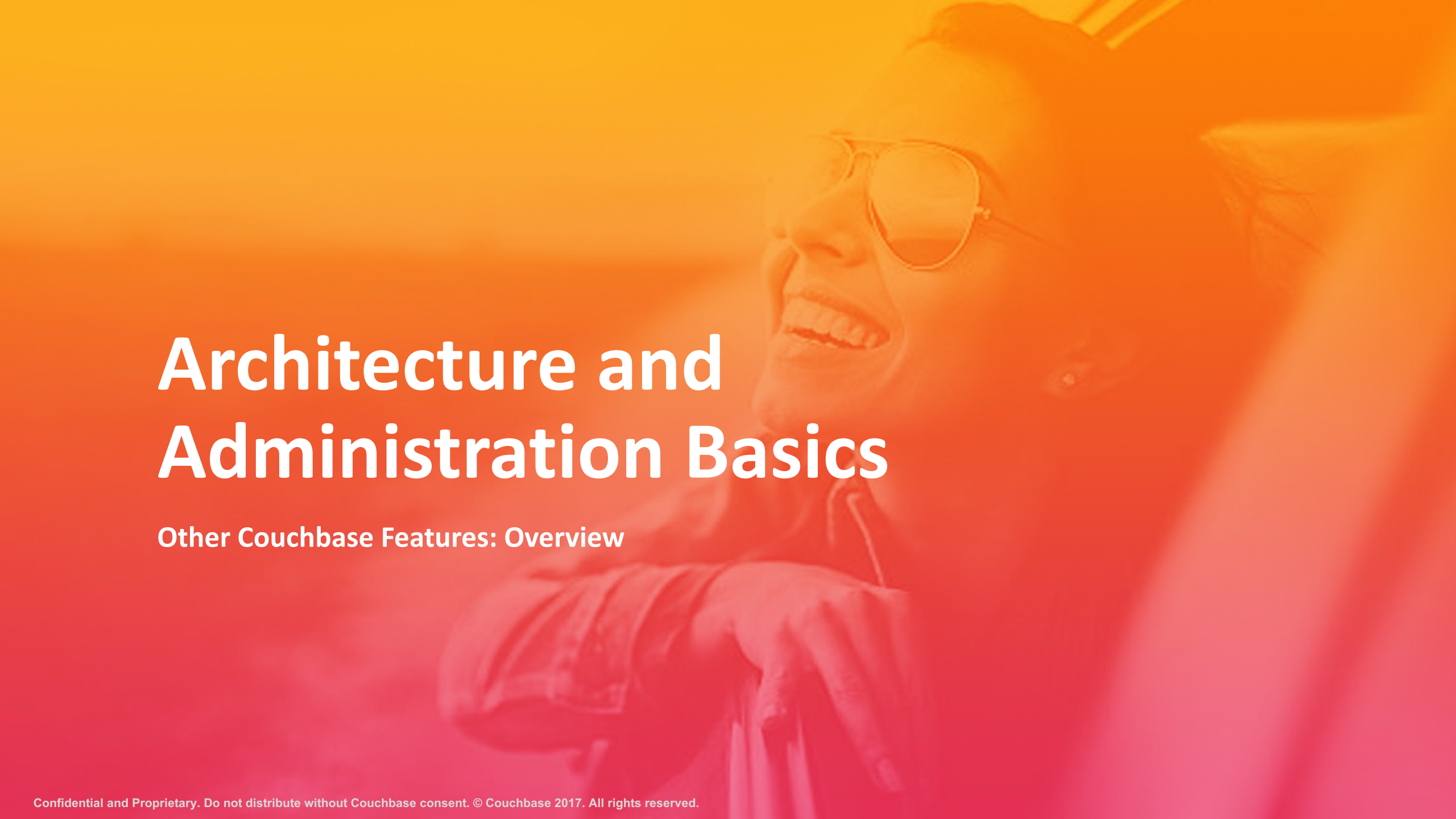


Architecture and Administration Basics

Other Couchbase Features: Overview

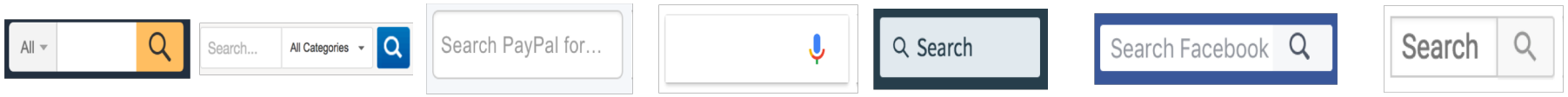
1 | Full-Text Search (FTS)

Search is pervasive...

Born-digital businesses are relentlessly innovating to provide better experiences

amazon ebay PayPal Google LinkedIn facebook YouTube

"Search" is foundational to all born-digital businesses...



Full Text Search

Basic document search use case

- Query String

search couchbase docs...

SEARCH

Search, Score & Snippet

Basic document search use case

- Query String

- Term matching
- Scoring
- Context snippet

best hotel location

SEARCH

Search results

Scoring	Document ID	Description Matches
1.88	hotel_1234	best location
1.82	hotel_2345	loved hotel location
1.37	hotel_3456	location is awesome
1.25	hotel_4557	hard to locate

Index, Analyze & Search

1. Index fields of a document
   "...located in the heart of the new City Quay development. The hotel has views of ..."

2. Analyze terms for index
   located ... heart .. new City Quay development. .. hotel has views

3. Query the index
   description: *location*

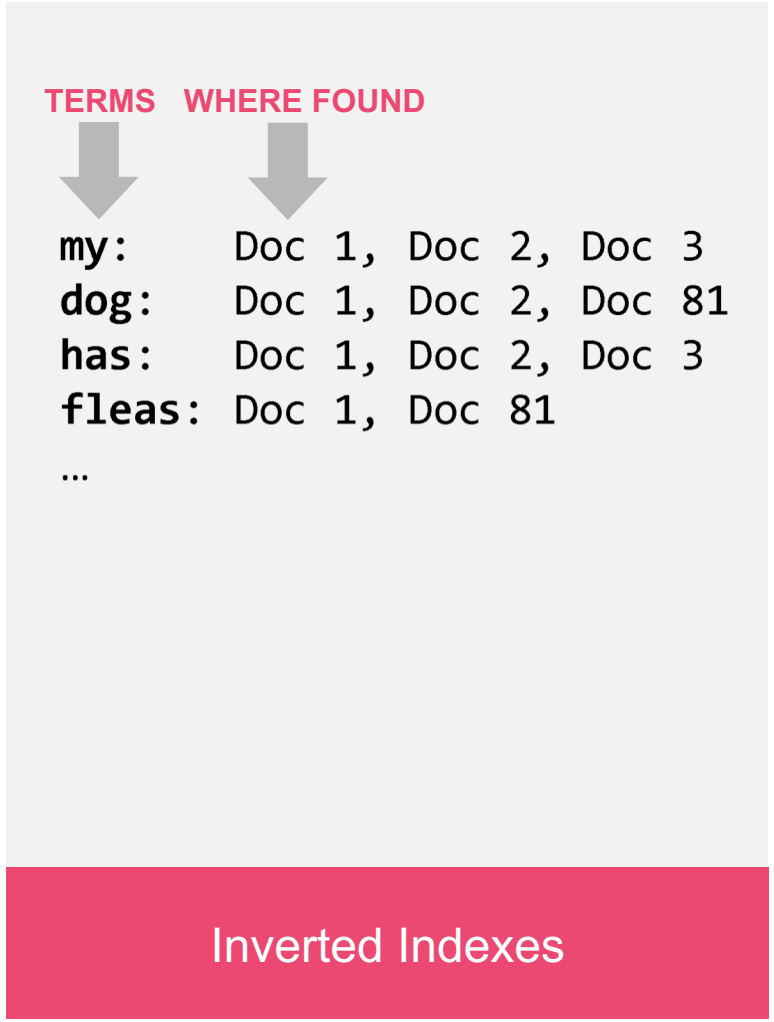
best hotel location

SEARCH

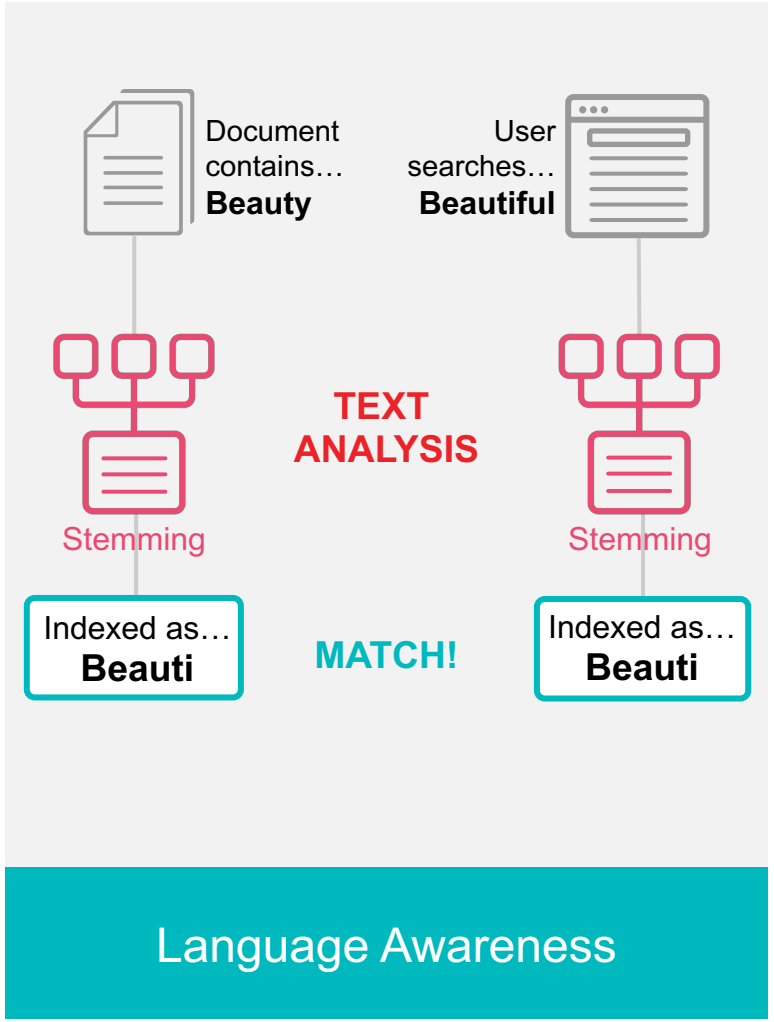
Return scored documents list

Scoring	Document ID	Description Matches
1.88	hotel_1234	best location
1.82	hotel_2345	loved hotel location
1.37	hotel_3456	location is awesome
1.25	hotel_4557	hard to locate

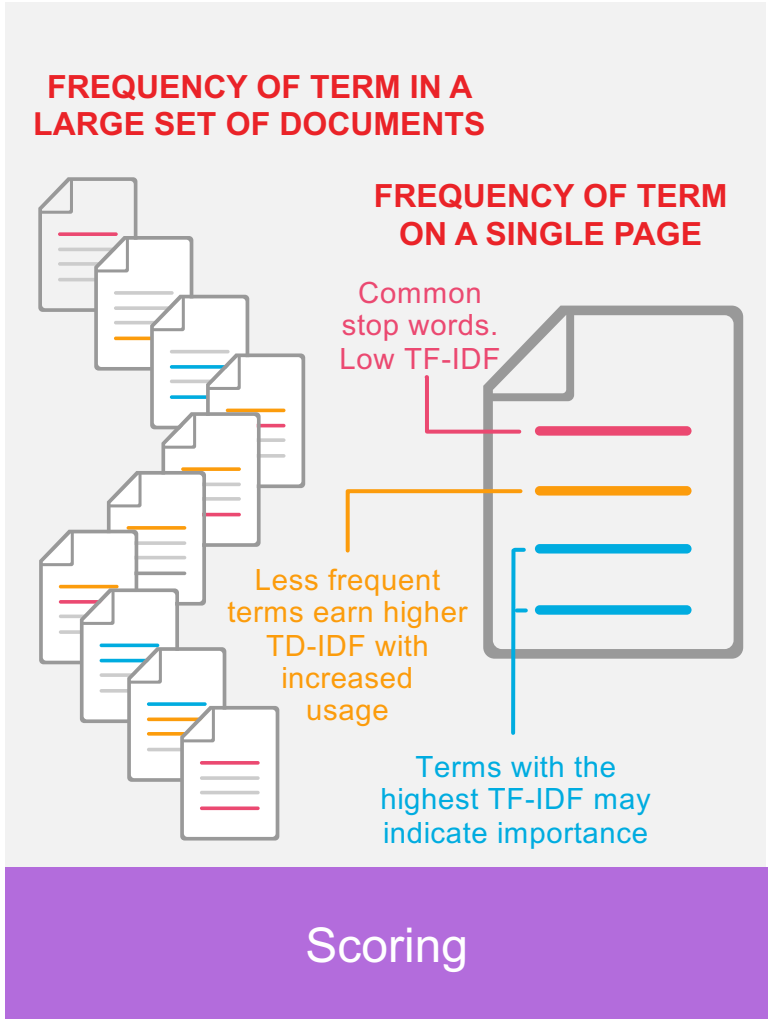
Underlying Concepts



Inverted Indexes



Language Awareness



Scoring

1 - 2 - 3 Step Approach to FTS

Index

- Identify document type
- Exclude fields / sub-sections
- Configure index behavior per field

Analyze

- Remove undesirable characters in input using **character filters**
- Split Input String to token streams using **Tokenizers**
- Process token streams by chaining **Token filters**

Query

- Query using SDK, REST API or web console
- Optionally query using Field Scoping(:), Boosting(^)
- Sort query results by any indexed field or score

Full Text Search - Capabilities

Query

- **Basic**: Match, Match Phrase, Fuzzy, Prefix, Regexp, Wildcard, Boolean Field
- **Compound**: QueryString, Boolean, Conjunction, Disjunction
- **Range**: DateRange, NumericRange
- **Special Purpose**: DocID, MatchAll, MatchNone, Phrase, Term, Geospatial
- **Scoring** (TF/IDF), boosting, field scoping
- **New/DP:** TermRange, Geospatial

Indexing

- Real time indexing (inverted index, auto-updated upon mutation)
- Default map and map by document type
- Dynamic mapping
- Stored fields, Term vectors
- Analyzers: Tokenization, Token Filtering (stop word removal, stemming – language specific)
- Aliasing

Full Text Search – SDK

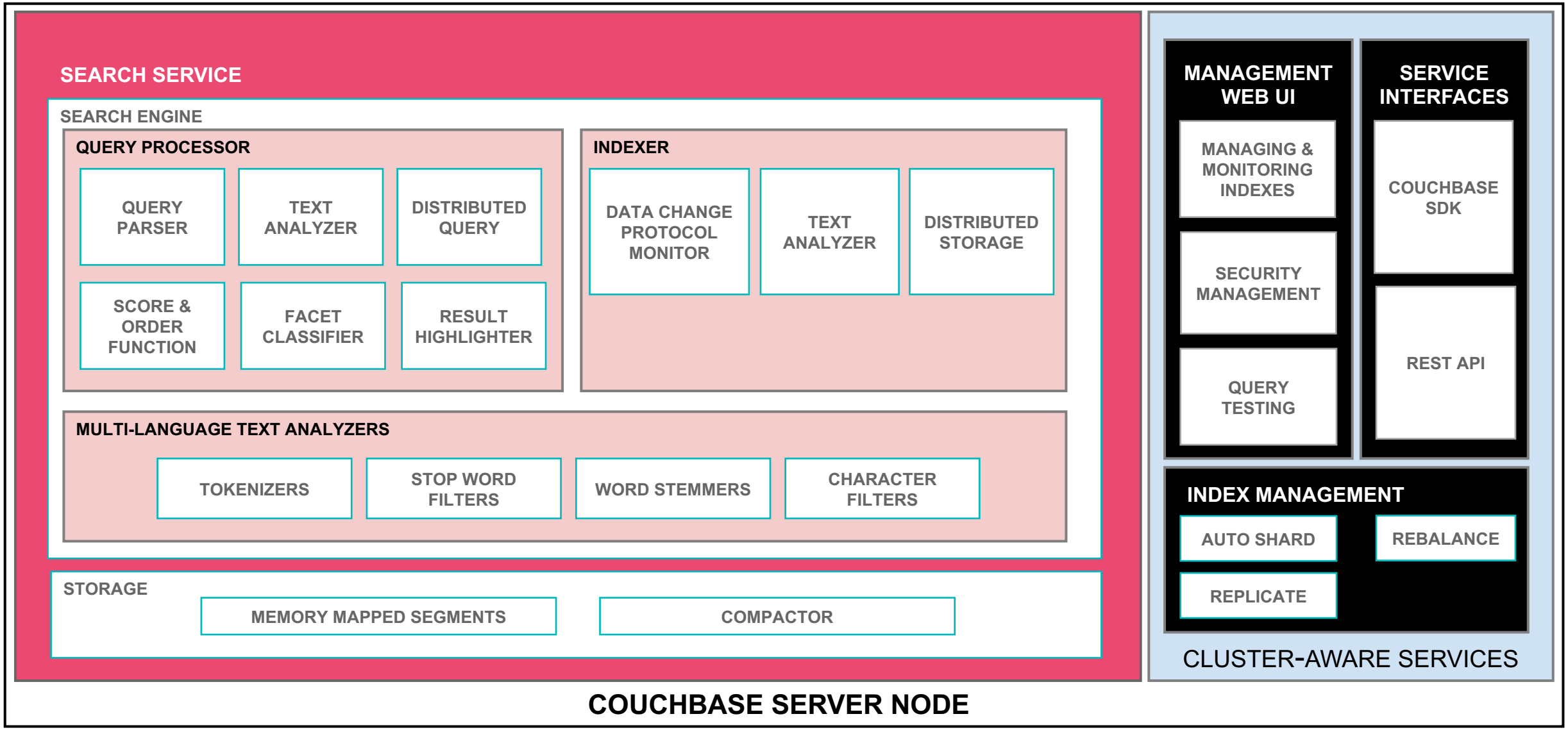
Python

```
termquery = fulltext.TermQuery('office')

results = bucket.search('travel-search', termquery, limit=25)
for result in results:
  ...
```

Java

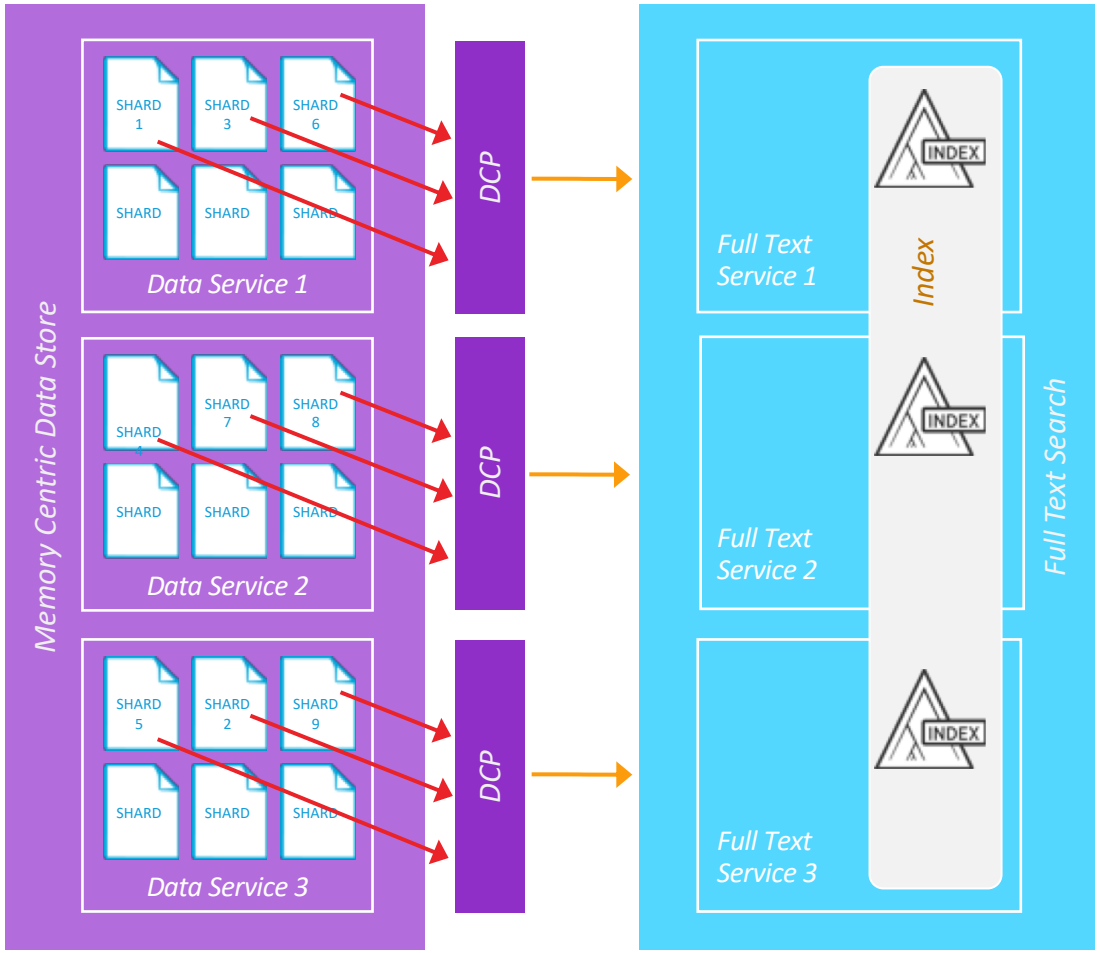
```
TermQuery termquery = SearchQuery.term("office");

SearchQueryResult results = bucket.query(
        new SearchQuery("travel-search", termquery)
);

for (SearchQueryRow row : results) { ... }
```

Full-Text Search Service



Full-Text Search Architecture

- Each FTS node a DCP stream subscriber
- Indexing distributed across FTS nodes
- Any FTS service can receive queries
  - “scatters” to other nodes
  - “gathers” response
- Application sees single logical Index

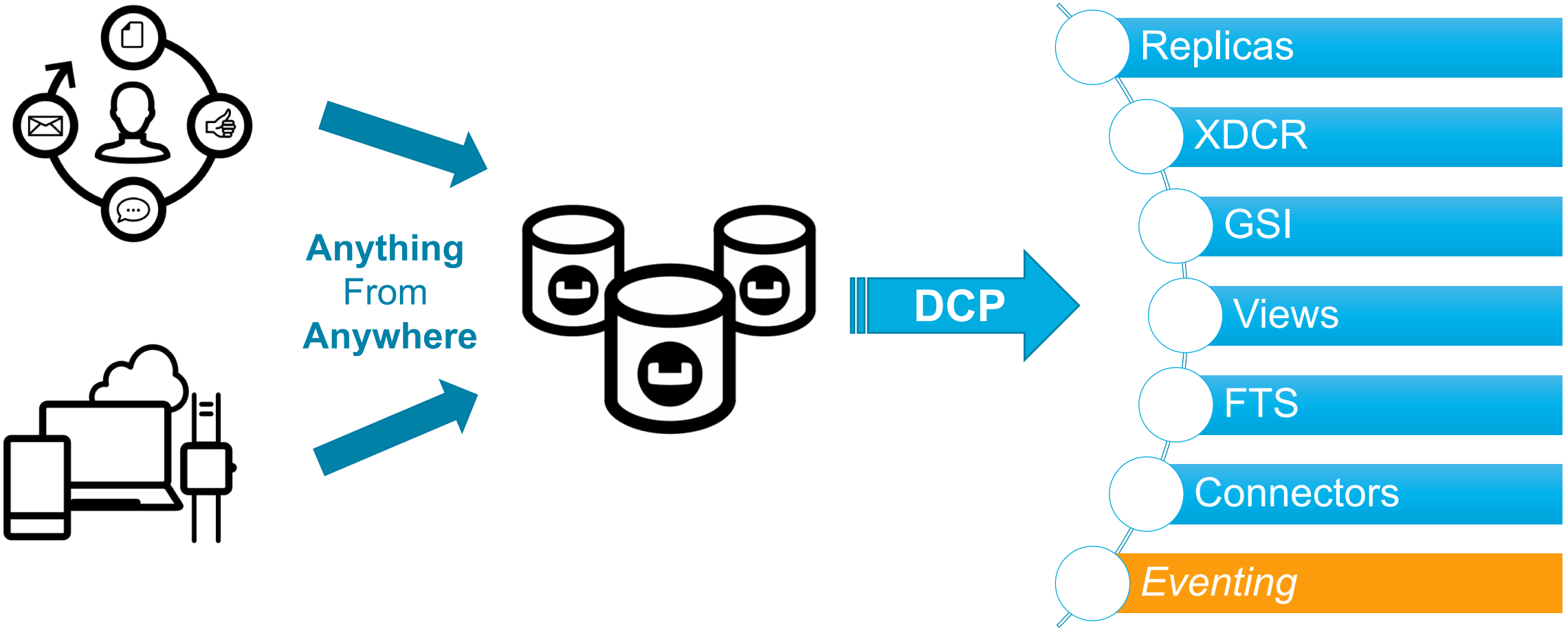


2 | Eventing

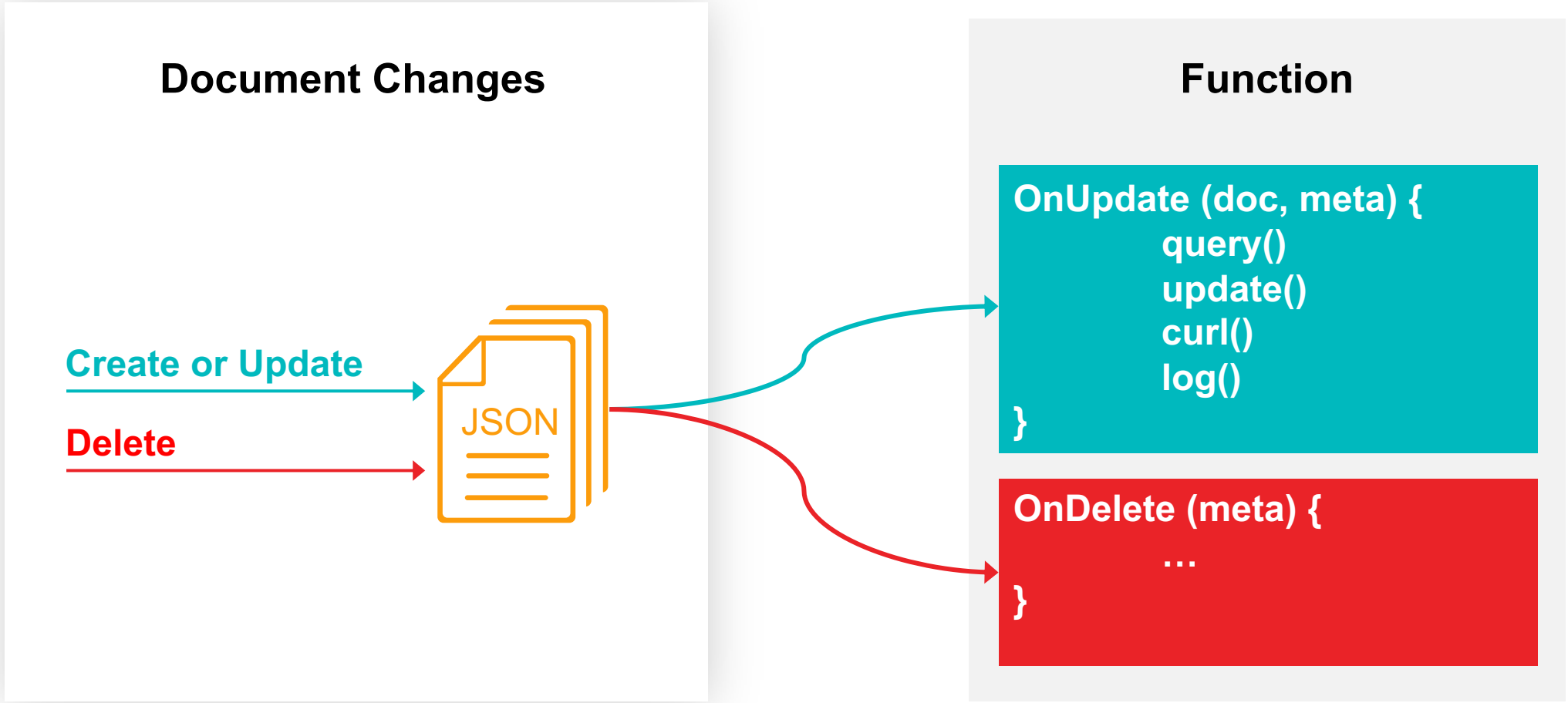
Eventing: Couchbase Functions Service

- An easy-to-use, easy-to-manage, highly-performant service
- Stateless Compute for Time Bracketed Workloads
- Based on Event-Condition-Action model
- Manage Data Operations in near Real-Time than at Query Time
- Simple Programming Model
- Utilizes the trends in Compute (multi-core CPUs)
- a.k.a Post Trigger

Eventing – A DCP Subscriber

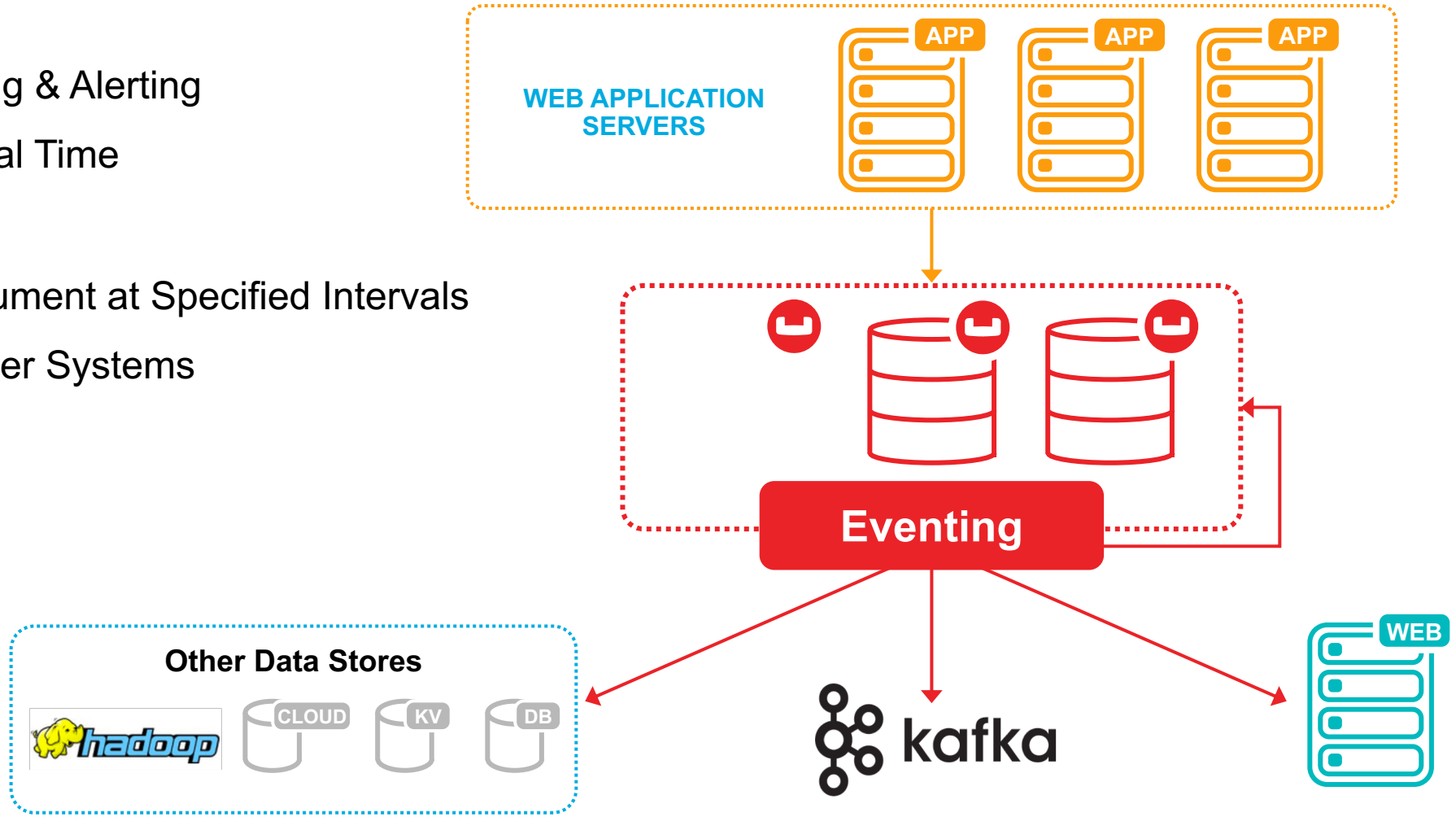


Eventing Functions: Custom Handlers of Change Events

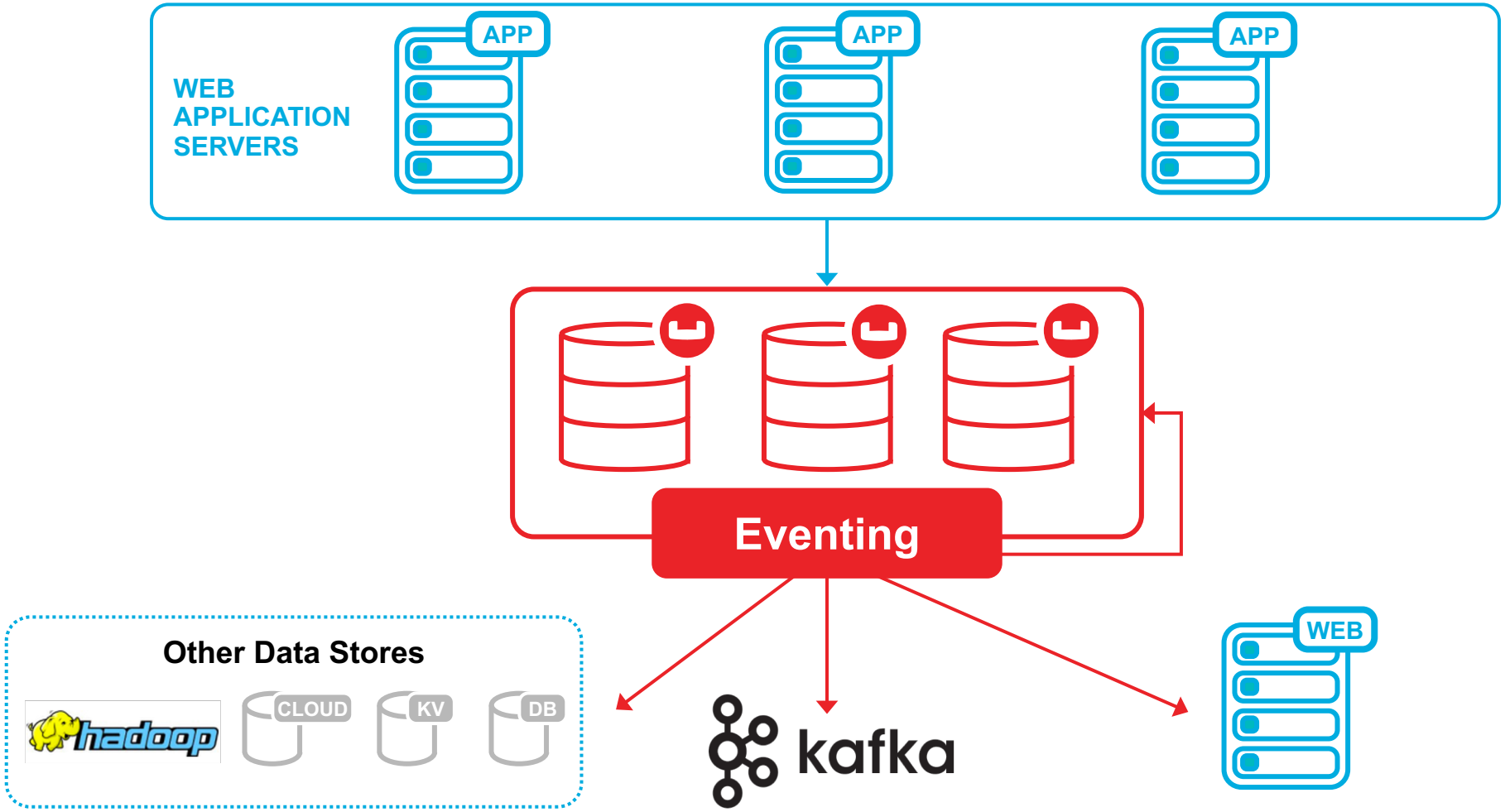


What can you use Eventing for?

- Notifications Before Expiry
- Threshold Based Monitoring & Alerting
- Enrich Content in Near-Real Time
- Cascade Deletes
- Trigger a routine on a Document at Specified Intervals
- Propagate Changes to Other Systems

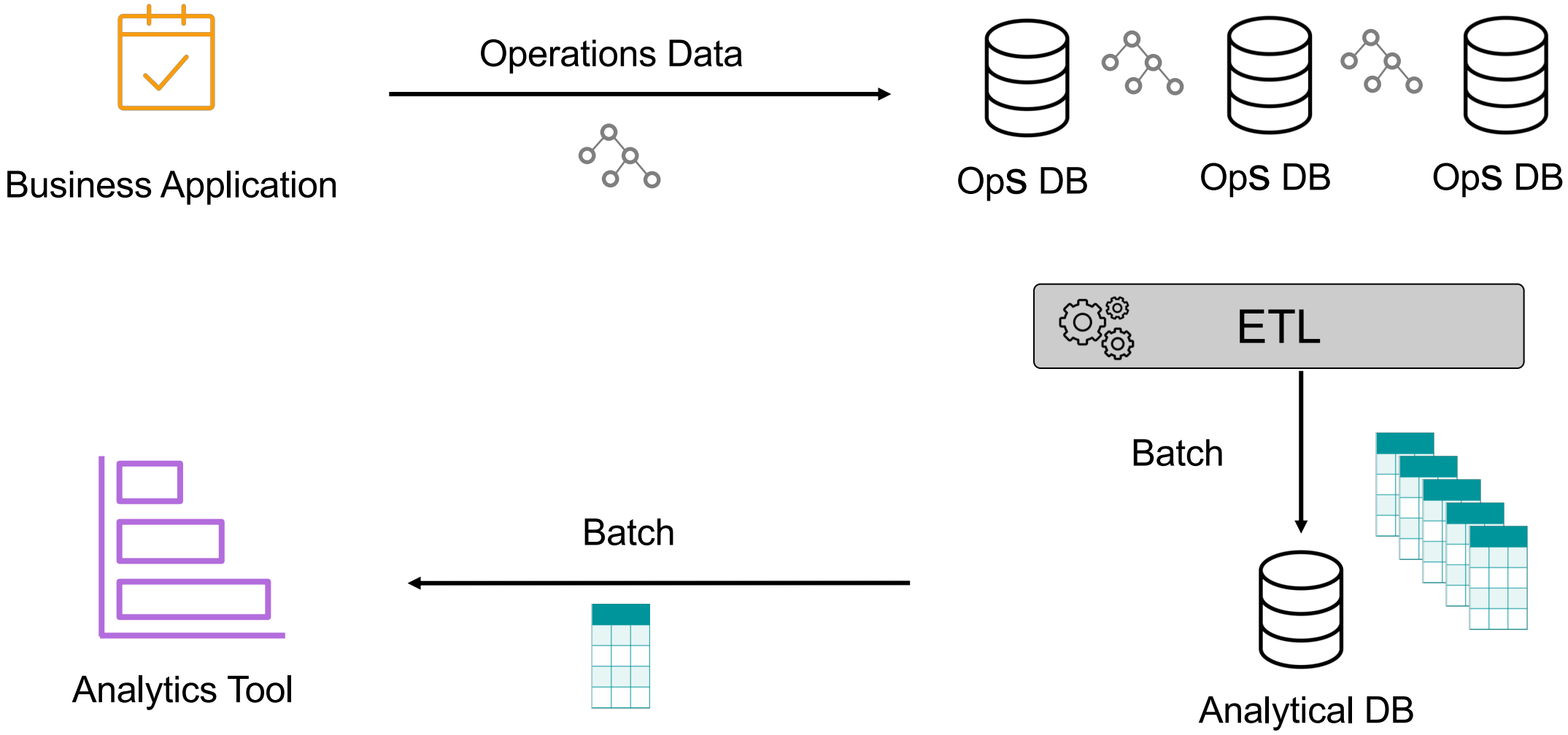


Couchbase 7.x: Eventing for Propagating Changes



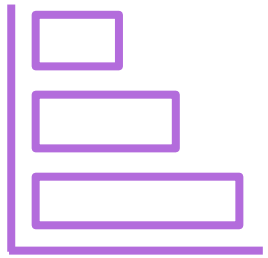
3 | Analytics

Traditional Analytics Solutions

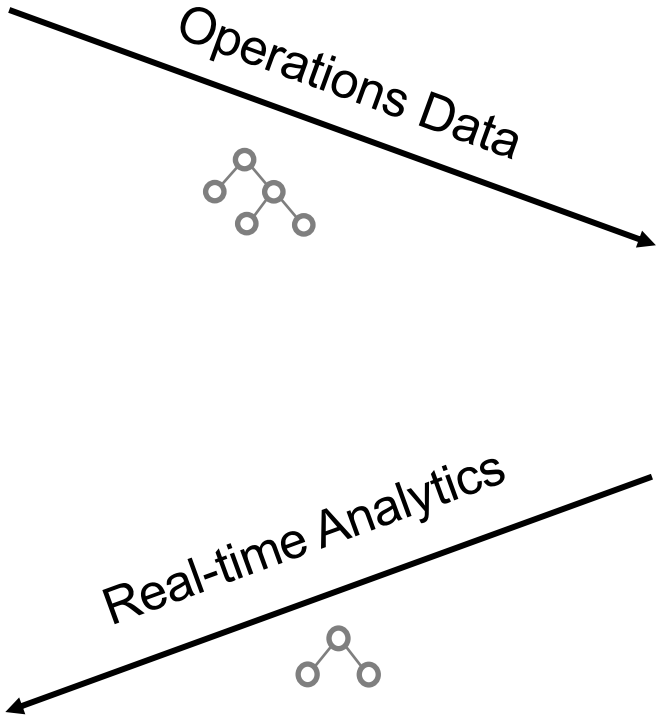


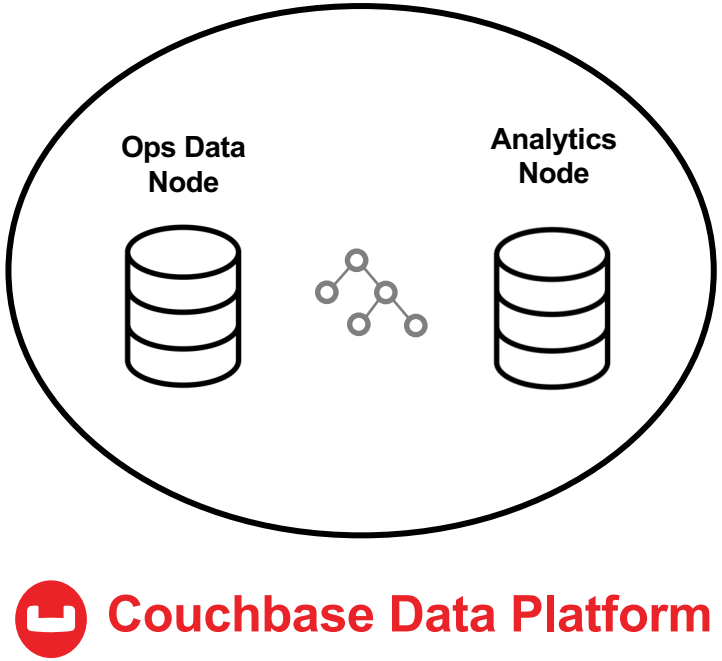
Couchbase Analytics – Bringing NoETL to NoSQL

Business Application



Analytics Tool





Advantages

- Support OLTP and OLAP processing in a single platform
- Eliminate the need for a separate OLAP system
  - Eliminate ETL
  - Reduces latency
  - Reduces complexity
  - Enables more intelligent applications
- Enable data exploration and ad hoc analytics

Analytics Service vs. N1QL

The Couchbase Analytics Service:

- Supports large join, set, aggregation, and grouping operations
- ...which often employ large amounts of data
- ...which can be highly consumptive of processor, networking, and memory resources
- ...and/or highly demanding in terms of cross-node coordination
- ...and/or of extensive duration

Analytics Service vs. N1QL

Analytic queries can be:

- Ad hoc
- Predetermined (often providing greater efficiency)

At its core, performance is enhanced by:

- Supporting parallel query-processing and bulk data-handling
- Allowing analytic queries to be run on local, automatically sharded indexes

Query and Analytics Services

QUERY SERVICE

- Online search and booking, reviews and ratings
  - Property and room detail pages
  - Cross-sell links, up-sell links
  - Stars & likes & associated reviews
  - Their booking history

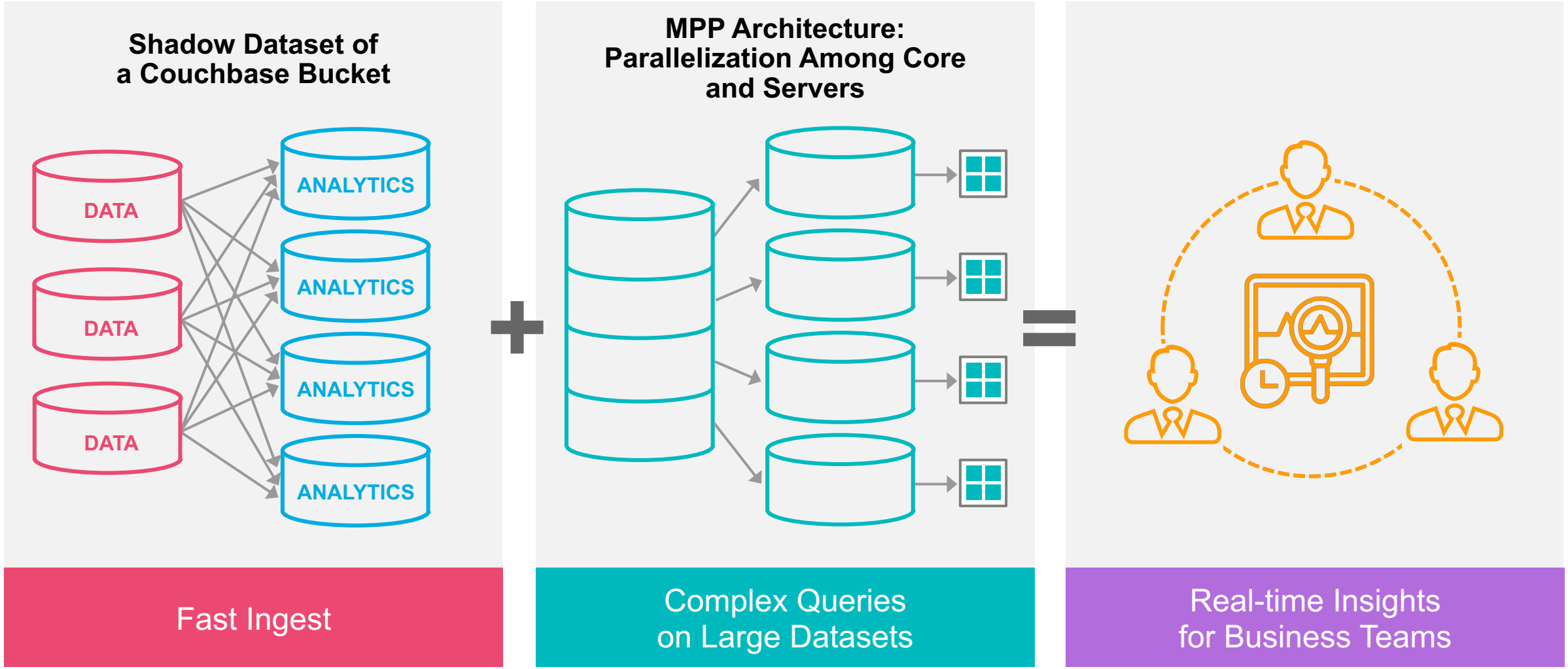
N1QL queries behind every page display and click/navigation

ANALYTICS SERVICE

- Reporting, Trend Analysis, Data Exploration
  - Daily discount availability report
  - Cities with highest room occupancy rates
  - Hotels with biggest single day drops
  - How many searches turn into bookings grouped by property rating? grouped by family size?

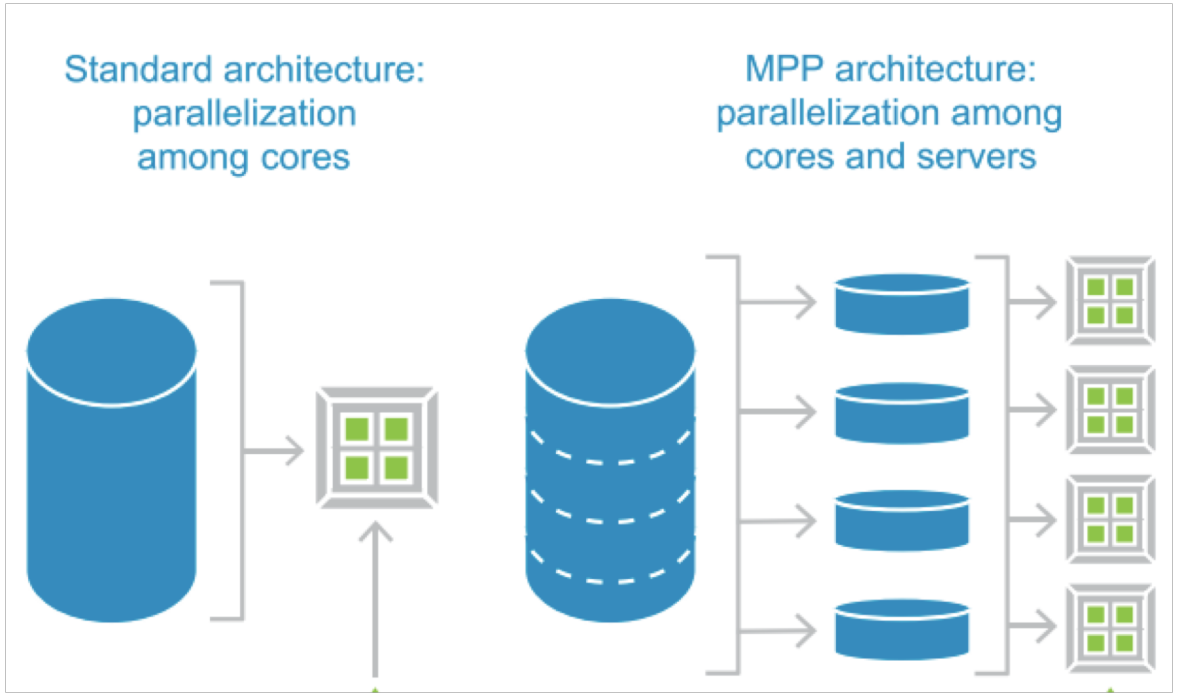
Business Analysts ask these questions without knowing in advance every aspect of the question

Underlying Architecture



"Secret" Sauce: Query Parallelism

- Massively Parallel Query Processor (MPP) executes complex queries on large datasets
- Comprehensive query language

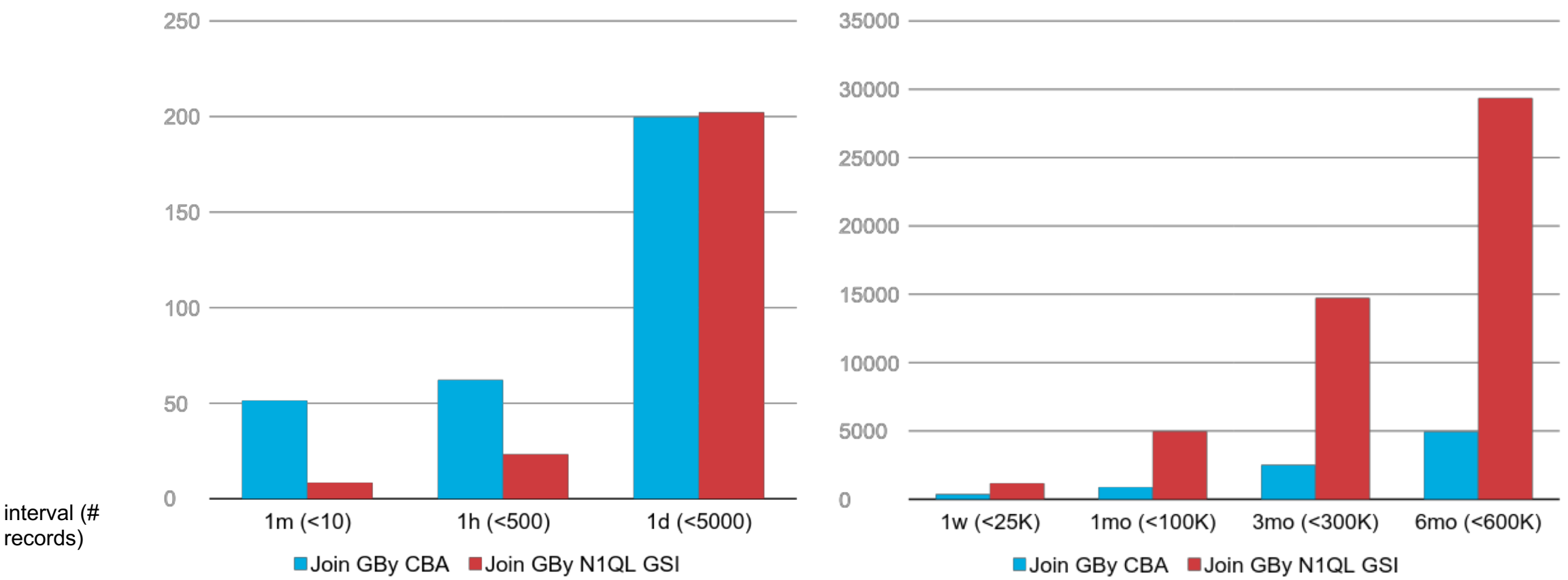


Example: Join, Grouping, and Aggregation

"Get the 10 chattiest users in a timeframe"

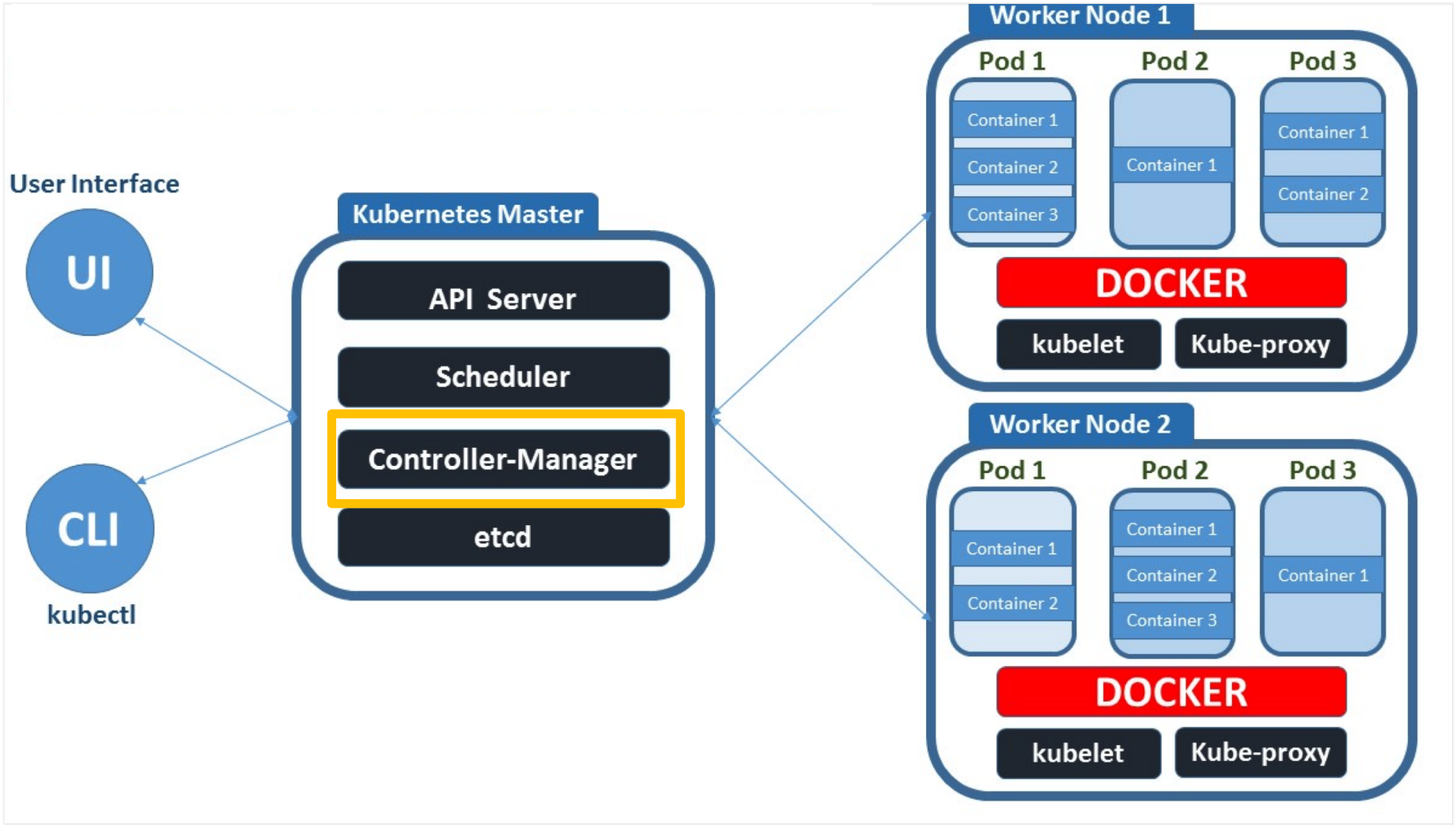
```
SELECT user.id, COUNT(message) AS count
FROM gbook_messages AS message, gbook_users AS user
WHERE message.author_id = user.id
  AND message.send_time BETWEEN "2001-11-28T09:57:13" AND "2001-11-29T09:57:13"
GROUP BY user.id
ORDER BY count DESC
LIMIT 10;
```

Couchbase Query and Analytics – Performance Tradeoff

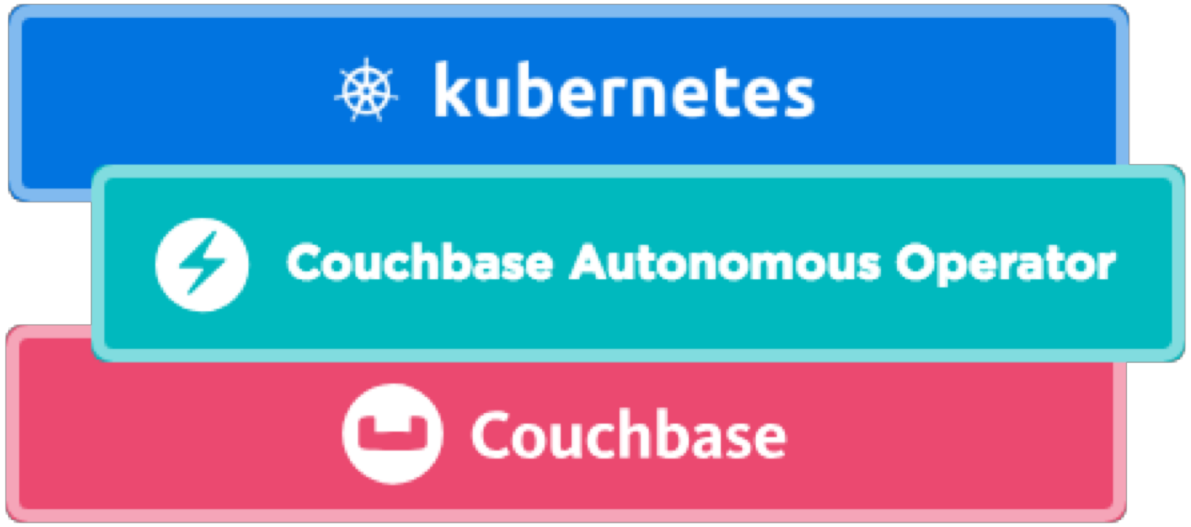


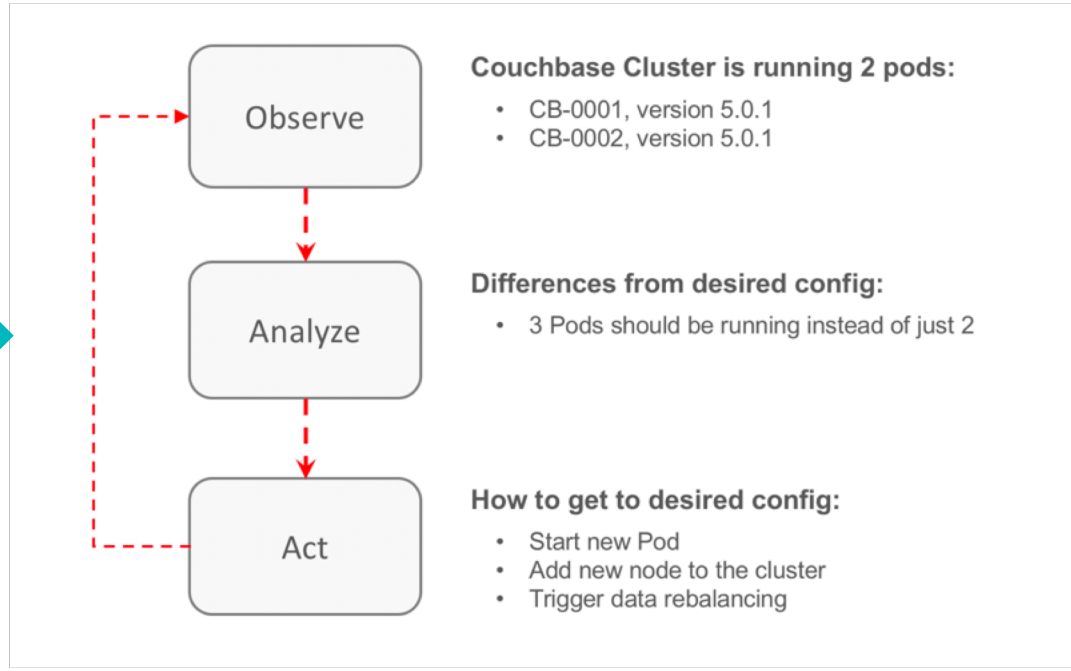
4 | Kubernetes Operator

KUBERNETES ARCHITECTURE



INTRODUCING COUCHBASE AUTONOMOUS OPERATOR

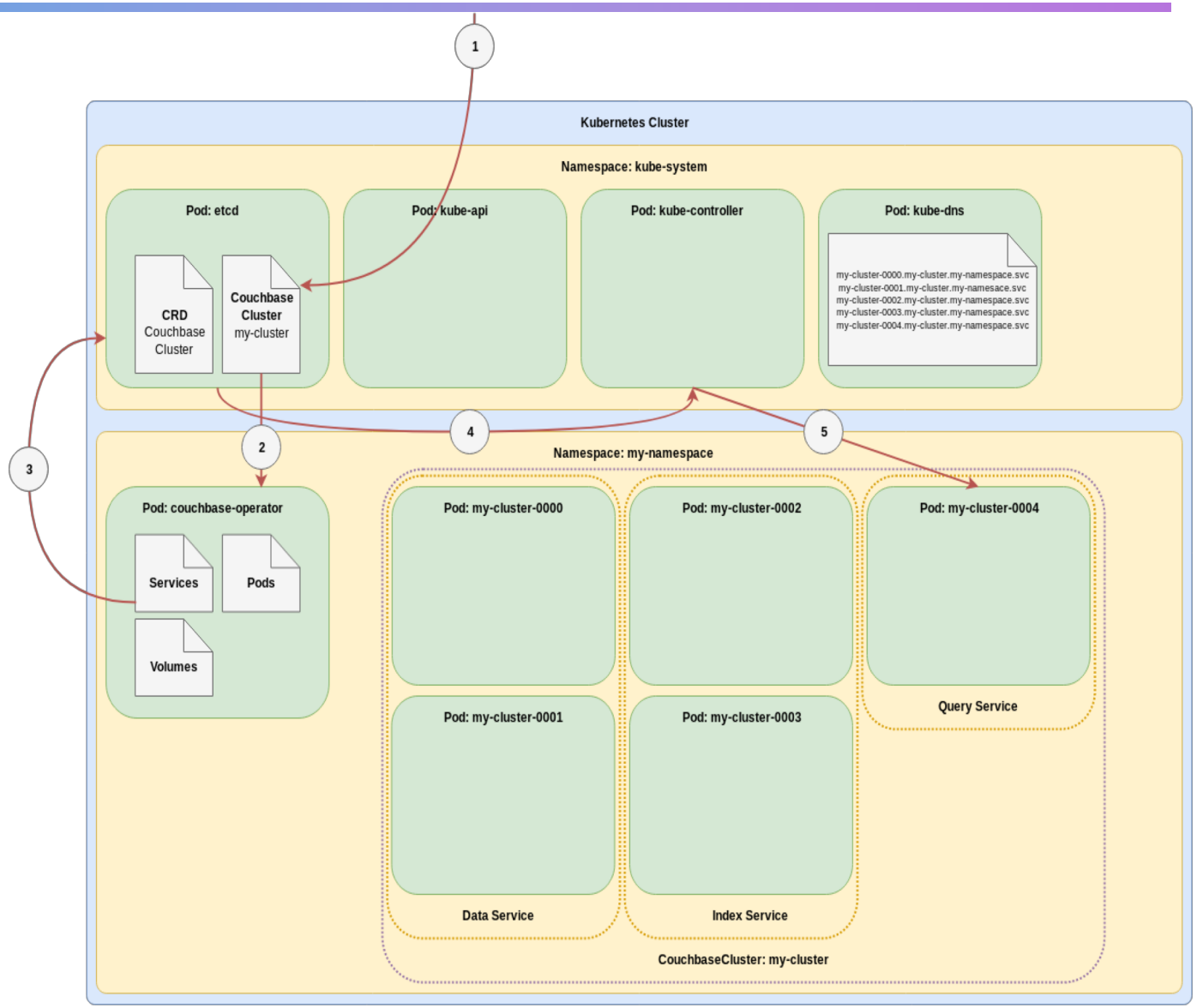




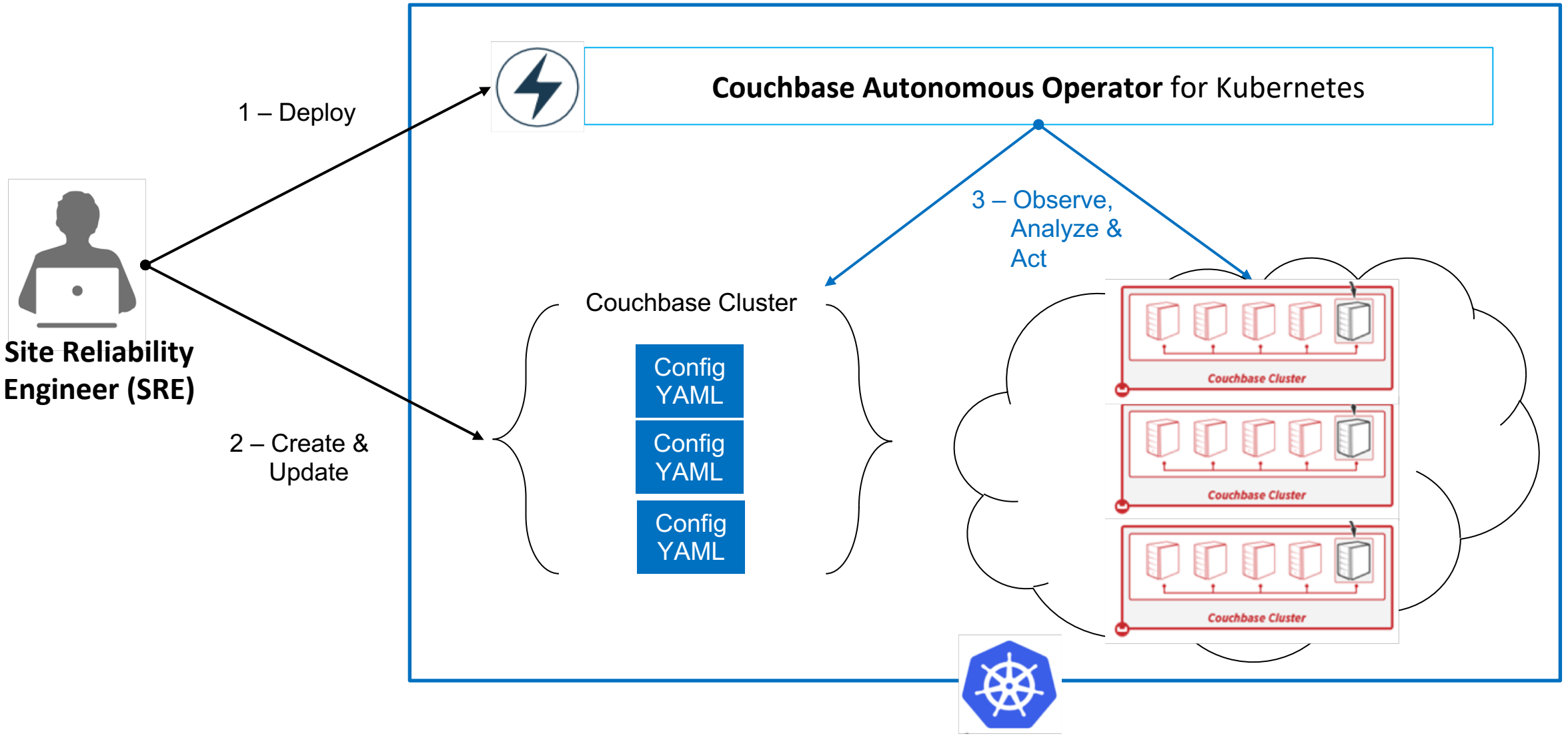
Couchbase Autonomous Operator is an **application-specific controller** that **extends** the **Kubernetes API** to create, configure and manage instances of complex **stateful** applications on behalf of a Kubernetes user. It builds upon the basic Kubernetes resource and controller concepts, but also includes domain or **application-specific knowledge** to **automate common tasks** better managed by computers.

Couchbase Operator Architecture

1. Register Operator yaml to Kubernetes etcd.
2. Operator does the RBAC validation and creates Services and Pods resources.
3. Operator starts a controller that creates CouchbaseCluster objects.
4. Operator sits listening to the API for updates to a CouchbaseCluster resources
5. Operator ensures the actual cluster state matches the desired state



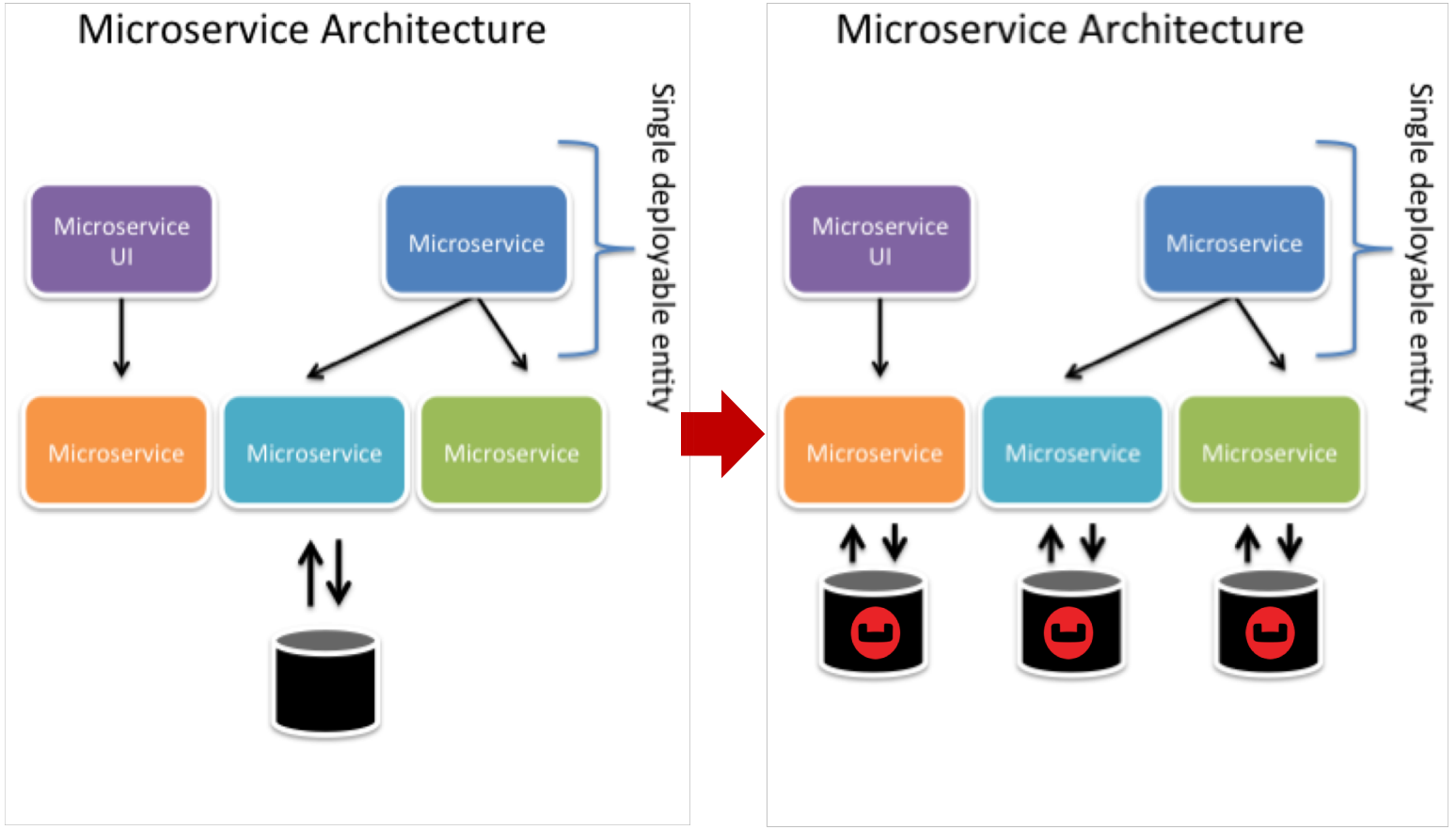
Key Benefit: EASE OF MANAGEMENT



Key Benefit: FASTER INNOVATION

Run and Manage Couchbase on Microservice Architecture

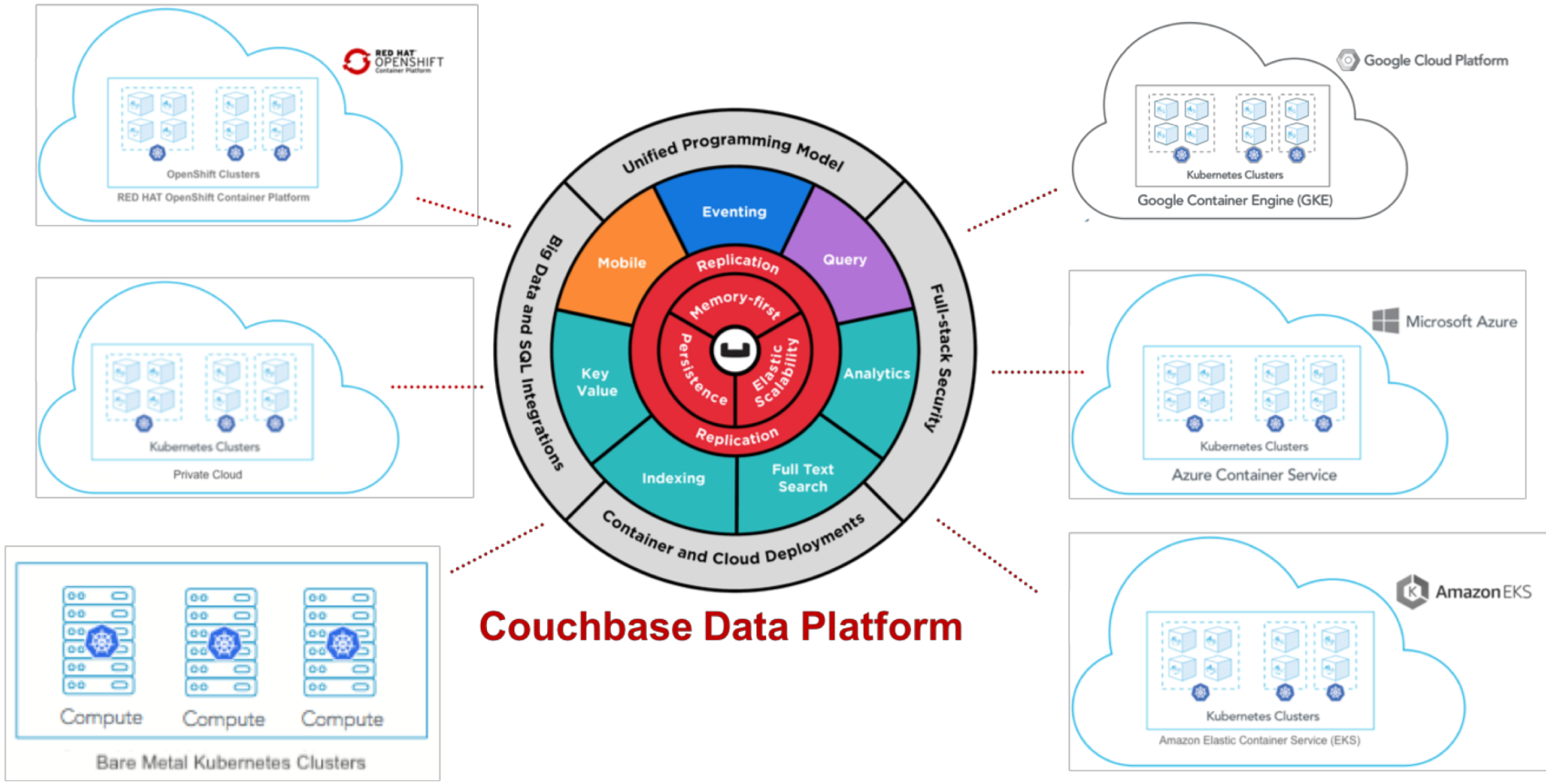
Reduce your DevOps workload by running the Couchbase Data Platform as an autonomous, fully managed stateful database application next to your microservices applications on the same Kubernetes platform.



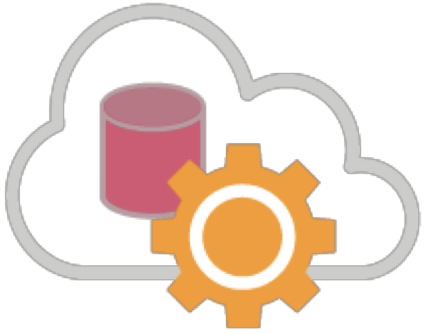
Key Benefit: INFRASTRUCTURE AND CLOUD AGNOSTIC

Hybrid Cloud and Multi-Cloud Strategy

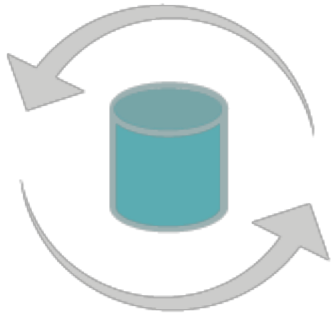
Provides a cloud-agnostic application deployment and management platform so we treat cloud providers almost like a commodity, as you will be able to deploy and migrate freely between them.



Autonomous Operator – Capabilities



Auto Provisioning



Auto-Recovery

Geo-Distribution

Centralized Management

Persistent Storage

Supportability Features

Thank you

Couchbase