

Give Your Coding Agent an (Elastic) Memory

Decisions and session history that survive the context window.

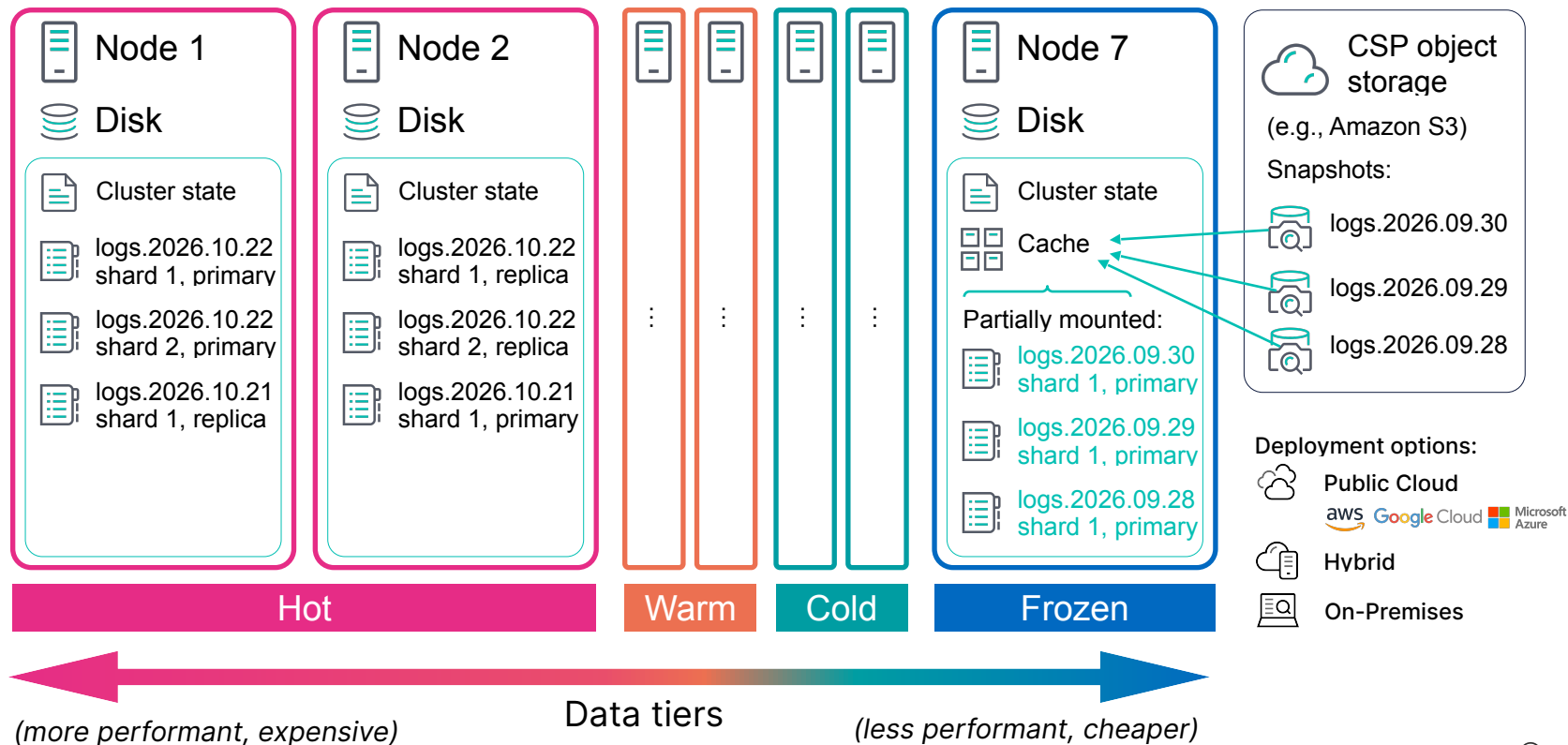
Engin Diri · Principal Solutions Architect, Pulumi

David Pilato · Developer | Evangelist, Elastic



- Cluster
- Nodes
- Index
- Shards
- Segments
- Cluster state
- Data streams
- Index Lifecycle Management (ILM)

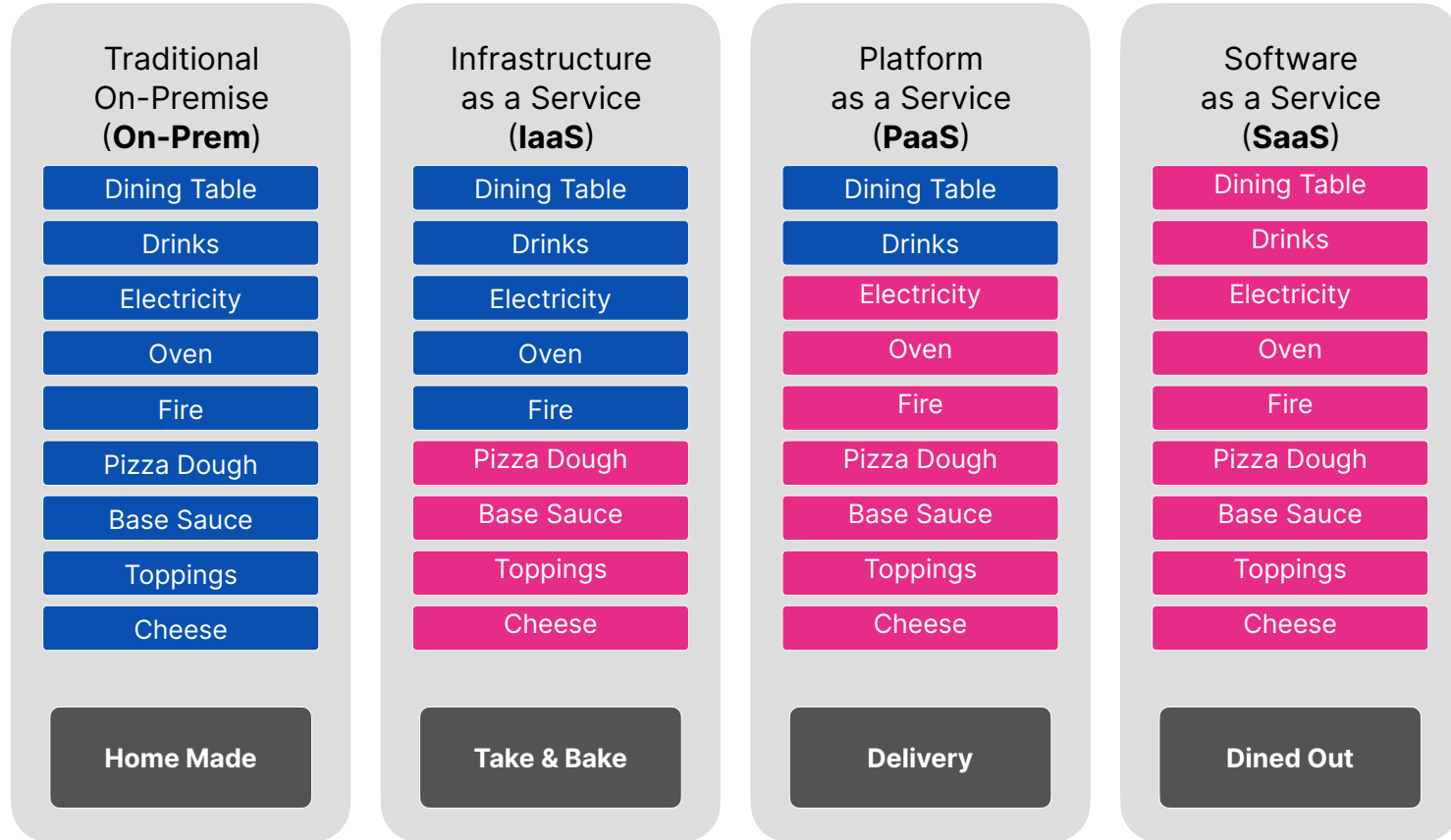
Shared-nothing stateful architecture



Disadvantages of stateful

- **User** defines cluster (RAM, CPU, disk)
- CPU & RAM **coupled** with storage
- CPU **shared** by ingestion & search
- Needs to think of **data tiers**
- **Adding/removing nodes** moves data

Pizza as a Service



 Self Managed

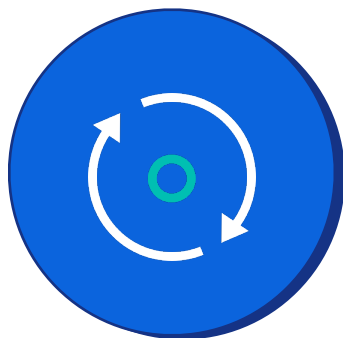
 Managed by Vendor

Introducing **Elastic Cloud Serverless**



Elastic Self Managed

Download and
run it yourself



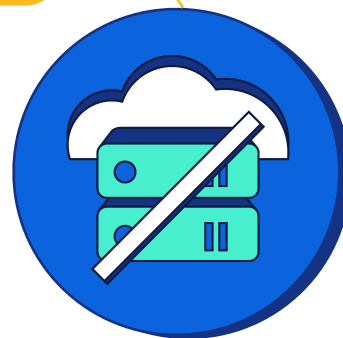
Elastic Cloud on K8s

We give the
orchestration to you



Elastic Cloud Hosted

We orchestrate
the infrastructure



Elastic Cloud Serverless

We manage
everything

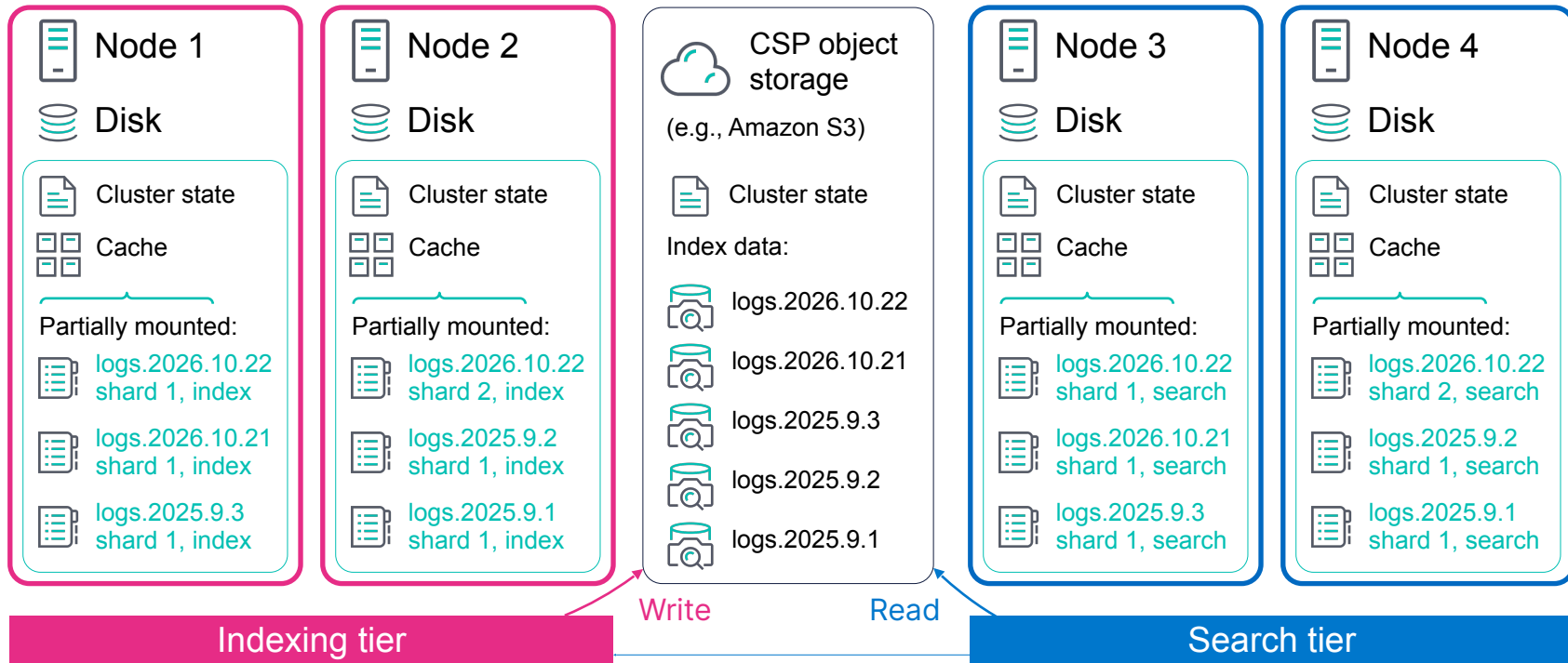
Elastic Cloud



Serverless Elasticsearch

- SaaS
- Exploits **Cloud Service Provider**
- Stateless: **decouples compute from storage**
- **Same API** as stateful Elasticsearch
- **Pay-as-you-go** model (retention, indexing, searching)
- **No more cluster specifics**, sizing, upgrading, HA
- 2 simplified data tiers: **indexing, search**

New stateless architecture



Stateless example

*offloading shard data to an
object store*

Offloading index data to the object store (stateless)

Node 1 (**indexing**)

New docs [1, 5]

Index “logs.2026.10.22”, shard 1

Lucene segments
generation:

Docs:



Flushed
commits:



Translog
generation:

Docs:



1

[1, 5]

Lucene in-memory
indexing buffers



CSP object
storage

(e.g., Amazon S3)

Offloading index data to the object store (stateless)

Node 1 (**indexing**)

New docs [6, 12]

Index “logs.2026.10.22”, shard 1

Lucene segments
generation:

Docs:



Flushed
commits:



Translog
generation:

Docs:



1

[1, 5]



2

[6, 12]

Lucene in-memory
indexing buffers



CSP object
storage

(e.g., Amazon S3)

Offloading index data to the object store (stateless)

Node 1 (**indexing**)

New docs [6, 12]

Index “logs.2026.10.22”, shard 1

Lucene segments
generation:

Docs:



Flushed
commits:



Translog
generation:

Docs:



1

[1, 5]



2

[6, 12]

Lucene in-memory
indexing buffers

Upload (< 200ms)



CSP object
storage

(e.g., Amazon S3)

Translog data:



Node 1, file 1

Offloading index data to the object store (stateless)

Node 1 (**indexing**)

Acknowledged writes

Index “logs.2026.10.22”, shard 1

Lucene segments
generation:

Docs:



Flushed
commits:



Translog
generation:

Docs:



1

[1, 5]



2

[6, 12]

Lucene in-memory
indexing buffers



CSP object
storage

(e.g., Amazon S3)

Translog data:

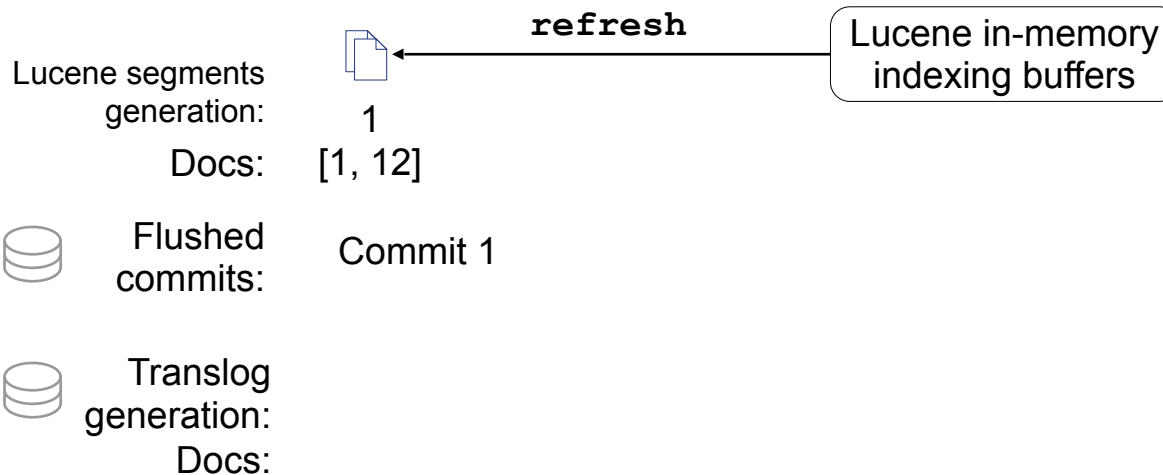


Node 1, file 1

Offloading index data to the object store (stateless)

Node 1 (**indexing**)

Index “logs.2026.10.22”, shard 1



 **CSP object storage**
(e.g., Amazon S3)

Translog data:
 Node 1, file 1

Offloading index data to the object store (stateless)

Node 1 (**indexing**)

New docs [13, 15]

Index “logs.2026.10.22”, shard 1

Lucene segments
generation: 1
Docs: [1, 12]

 Flushed
commits: Commit 1

 Translog
generation:
Docs:

Lucene in-memory
indexing buffers


3
[13, 15]



CSP object
storage

(e.g., Amazon S3)

Translog data:



Node 1, file 1

Offloading index data to the object store (stateless)

Node 1 (**indexing**)

New docs [13, 15]

Index “logs.2026.10.22”, shard 1

Lucene segments
generation: 1

Docs: [1, 12]



Flushed
commits: Commit 1



Translog
generation:
Docs:

Lucene in-memory
indexing buffers

Upload (< 200ms)



3
[13, 15]



CSP object
storage

(e.g., Amazon S3)

Translog data:



Node 1, file 1



Node 1, file 2

Offloading index data to the object store (stateless)

Node 1 (**indexing**)

Acknowledged writes

Index “logs.2026.10.22”, shard 1

Lucene segments
generation: 1
Docs: [1, 12]

 Flushed
commits: Commit 1

 Translog
generation: 3
Docs: [13, 15]

Lucene in-memory
indexing buffers



CSP object
storage

(e.g., Amazon S3)

Translog data:



Node 1, file 1

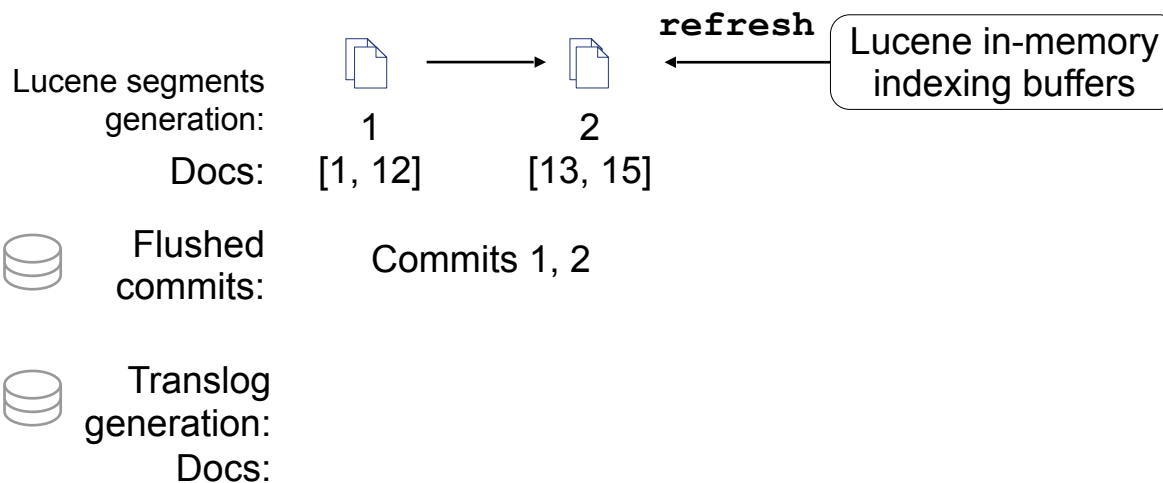


Node 1, file 2

Offloading index data to the object store (stateless)

Node 1 (**indexing**)

Index “logs.2026.10.22”, shard 1



 CSP object storage
(e.g., Amazon S3)

Translog data:

 Node 1, file 1
 Node 1, file 2

Offloading index data to the object store (stateless)

Node 1 (**indexing**)

Index “logs.2026.10.22”, shard 1

Lucene segments
generation: 1 → 2
Docs: [1, 12] [13, 15]

Lucene in-memory
indexing buffers

 Flushed
commits:

Commits 1, 2

Index **flush**

 Translog
generation:
Docs:

Upload

 CSP object
storage
(e.g., Amazon S3)

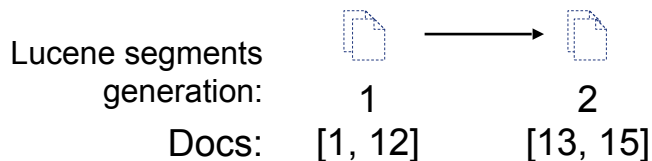
Translog data:
(deleted)

Index data:
 logs.2026.10.22
shard 1, BCC1
(Batched compound commit)

Offloading index data to the object store (stateless)

Node 2 ([search](#))

Index “logs.2026.10.22”, shard 1



Flushed
commits:

Commits 1, 2



Translog
generation:
Docs:

Lucene in-memory
indexing buffers



Node Cache

Partially mounted:



CSP object
storage

(e.g., Amazon S3)

Translog data:

Index data:



logs.2026.10.22
shard 1, BCC1
(Batched compound commit)

Offloading index data to the object store (stateless)

Node 2 (**search**)

Search documents 2 & 14

Index “logs.2026.10.22”, shard 1

Lucene segments generation: 1 → 2
Docs: [1, 12] [13, 15]



Flushed commits:

Commits 1, 2



Translog generation:
Docs:

Lucene in-memory indexing buffers



Node Cache

Partially mounted:



Files of CC1



Files of CC2



CSP object storage
(e.g., Amazon S3)

Translog data:

Index data:



logs.2026.10.22
shard 1, BCC1
(Batched compound commit)