

Databases on Kubernetes Using a Custom Operator

Day 1, Day 2, and Beyond

@unterstein @neo4j @KubeCon_ #kubecon





Johannes Unterstein

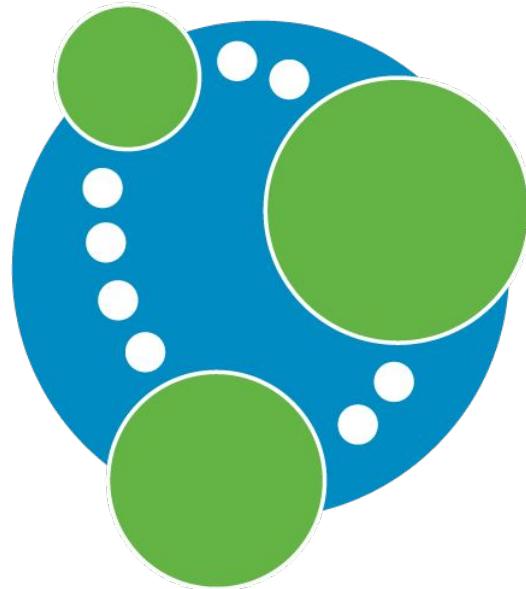
Software Engineer
[Neo4j.com/Cloud](https://neo4j.com/cloud)



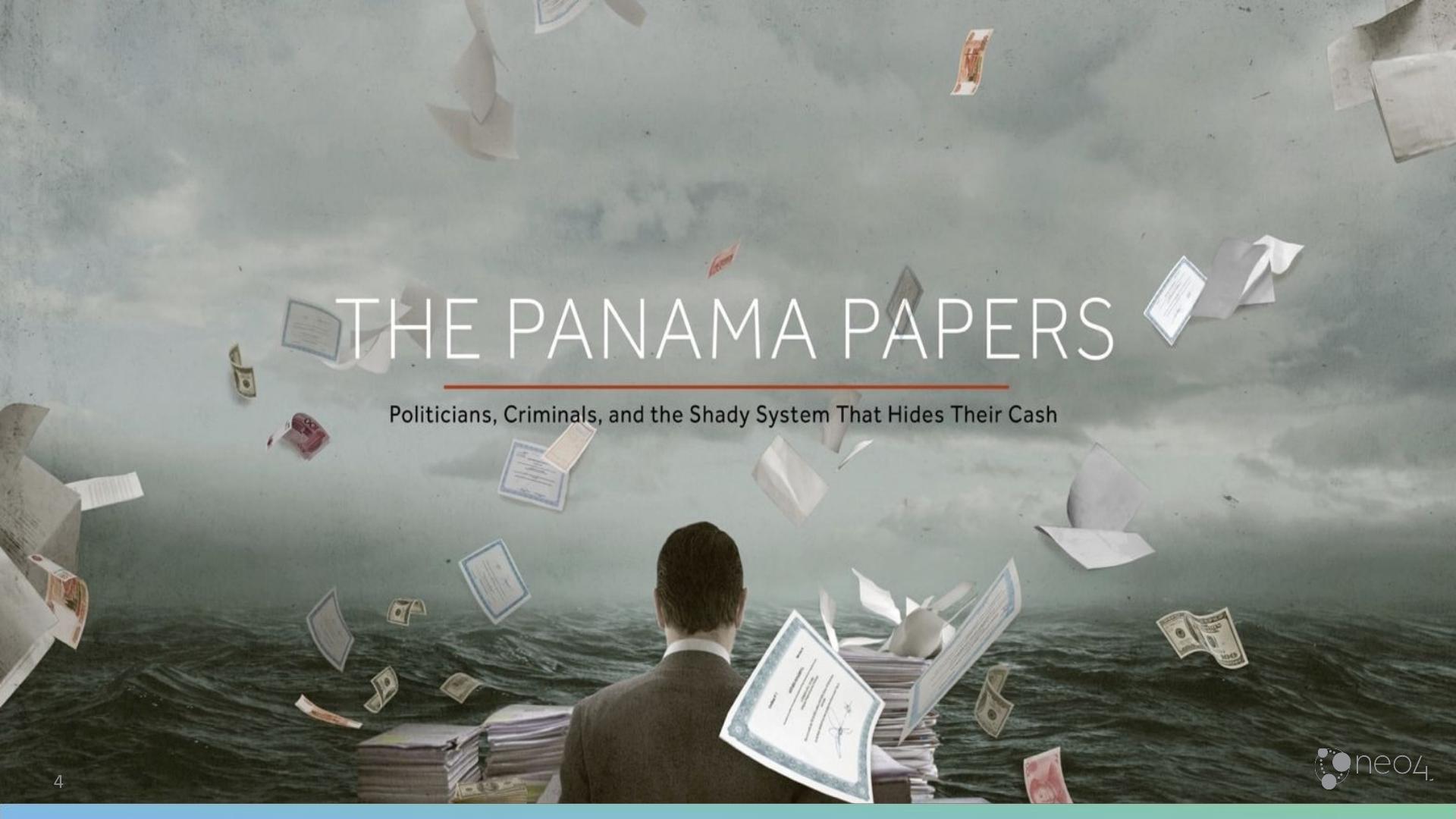
@unterstein



Neo for what?



neo4j



THE PANAMA PAPERS

Politicians, Criminals, and the Shady System That Hides Their Cash

Law firm based in panama was leaked



11.5 million documents

Emails, Scanned Documents,
Bank Statements etc...



2.6 TB

They tried to analyze all that with excel

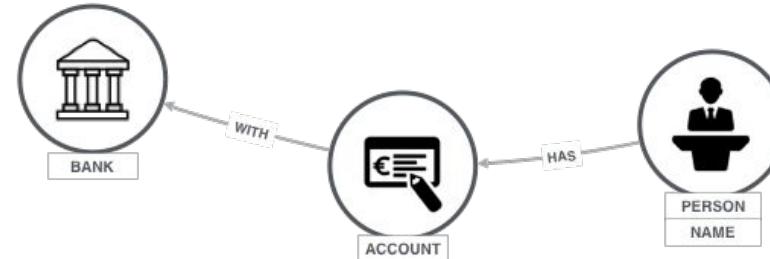


11.5 million documents

Emails, Scanned Documents,
Bank Statements etc...



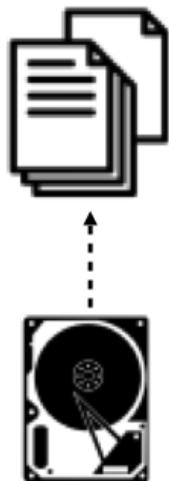
2.6 TB



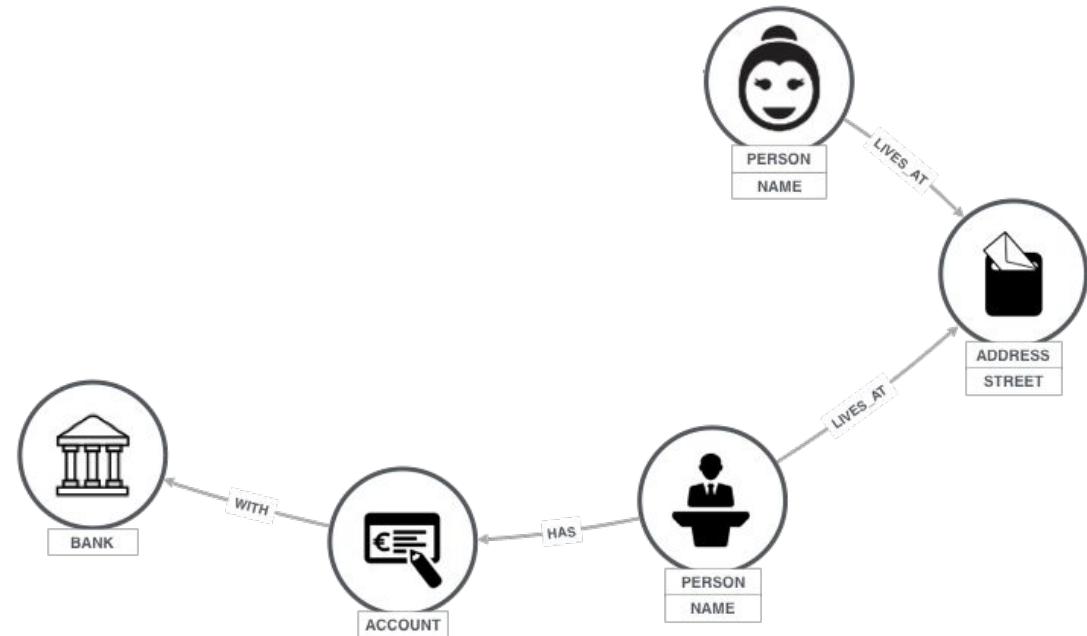
But then they found graphs ...

11.5 million documents

Emails, Scanned Documents,
Bank Statements etc...



2.6 TB



... and the pattern!

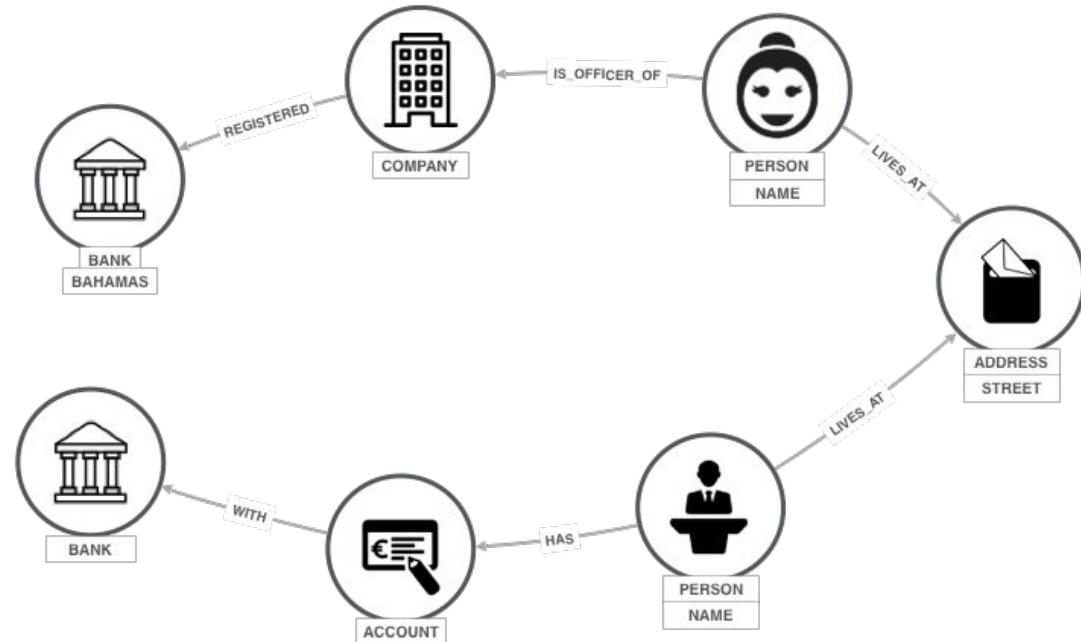


11.5 million documents

Emails, Scanned Documents,
Bank Statements etc...



2.6 TB



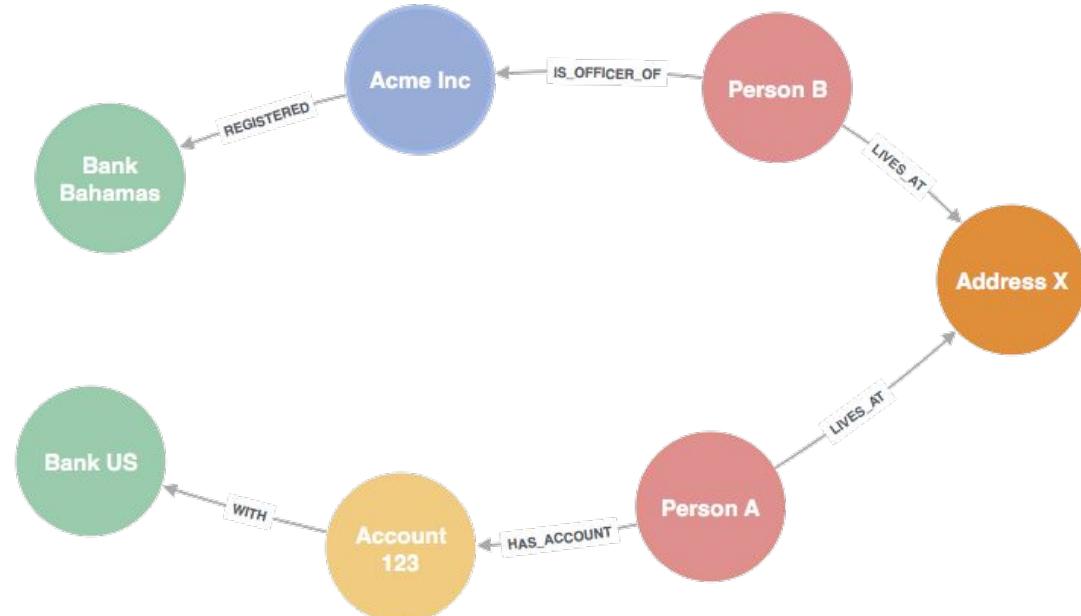
So, this is what we do.

11.5 million documents

Emails, Scanned Documents,
Bank Statements etc...



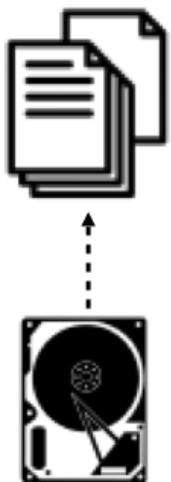
2.6 TB



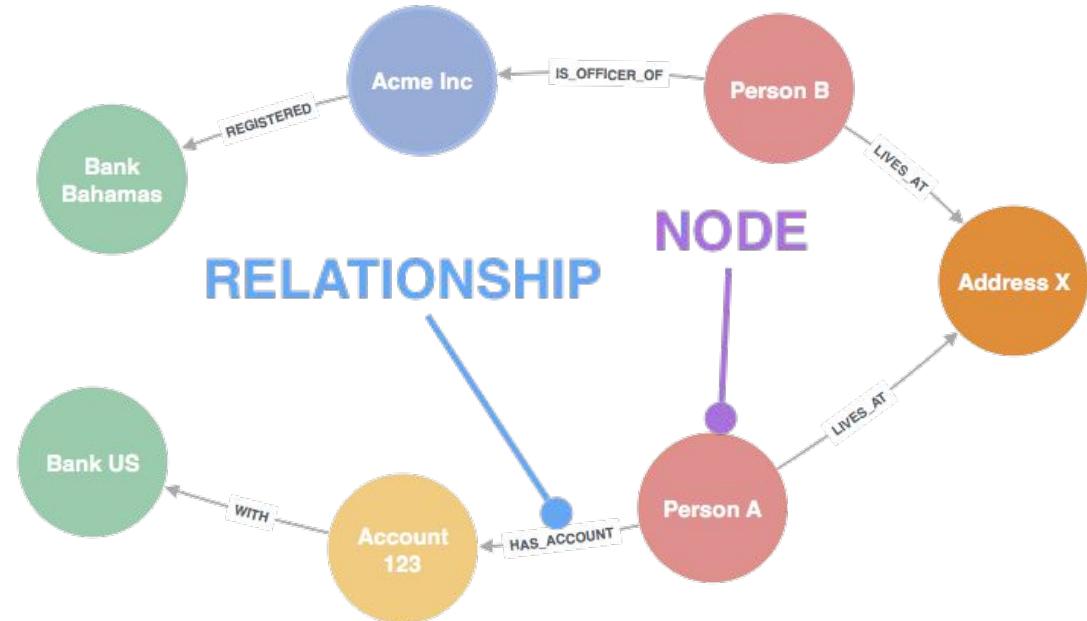
That's the data model!

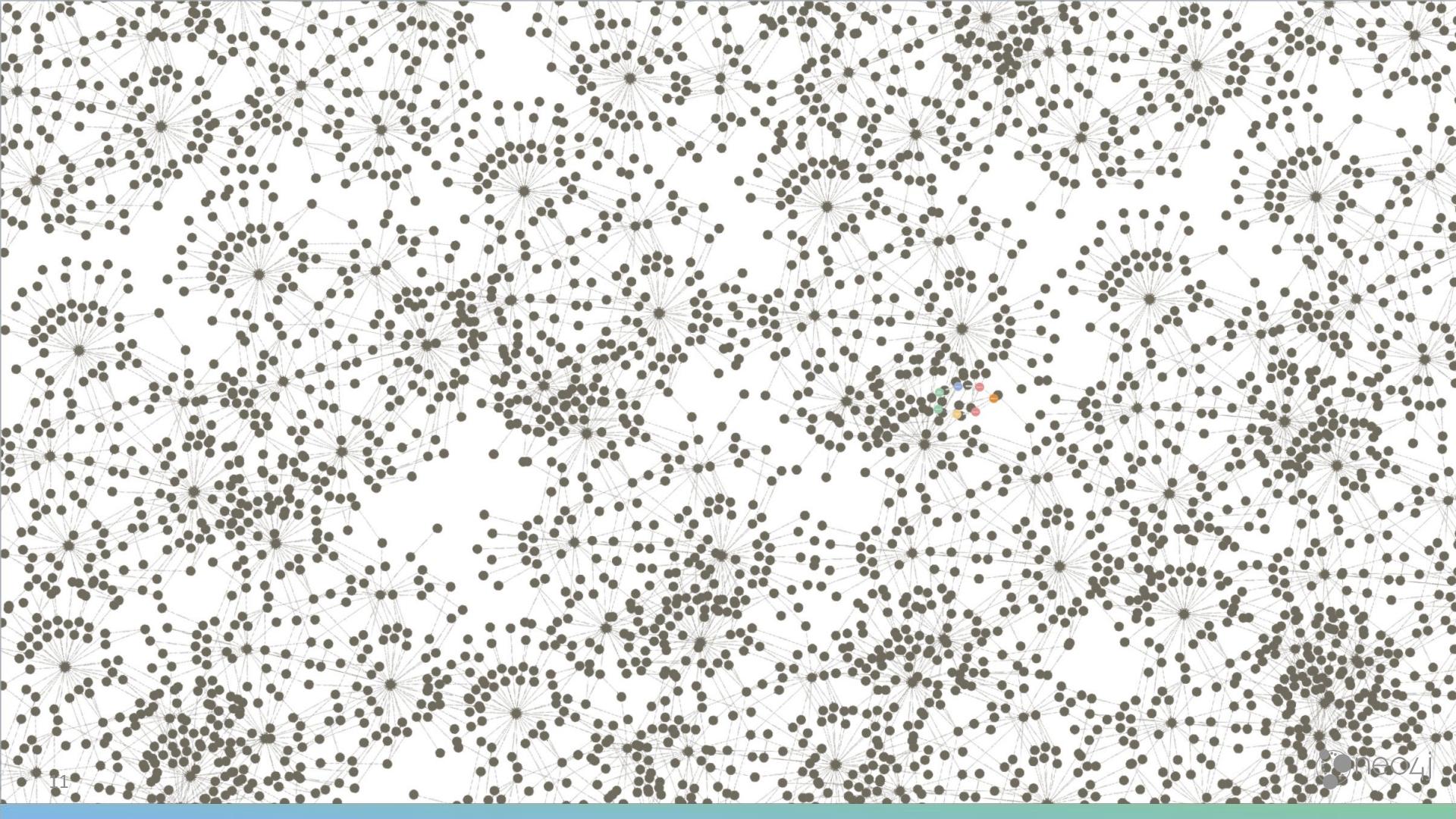
11.5 million documents

Emails, Scanned Documents,
Bank Statements etc...



2.6 TB







Finding a cure for cancer



Johnson Space Center

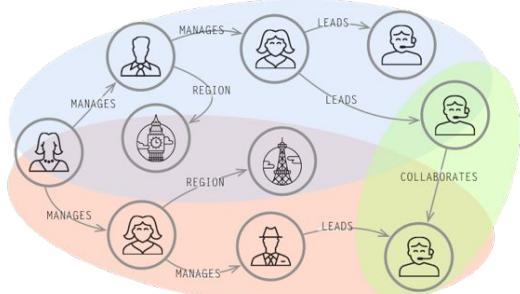
Houston, Texas

Helping to bring people to the mars

More traditional use cases

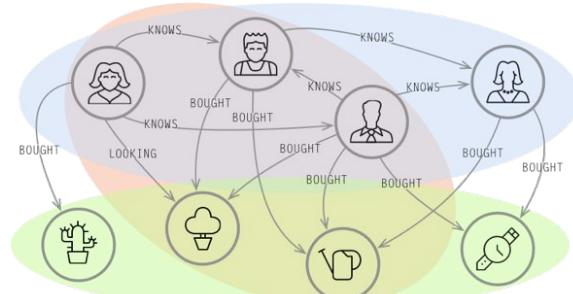
Internal Applications

- Master Data Management
- Privacy, Risk and Compliance
- Fraud Detection

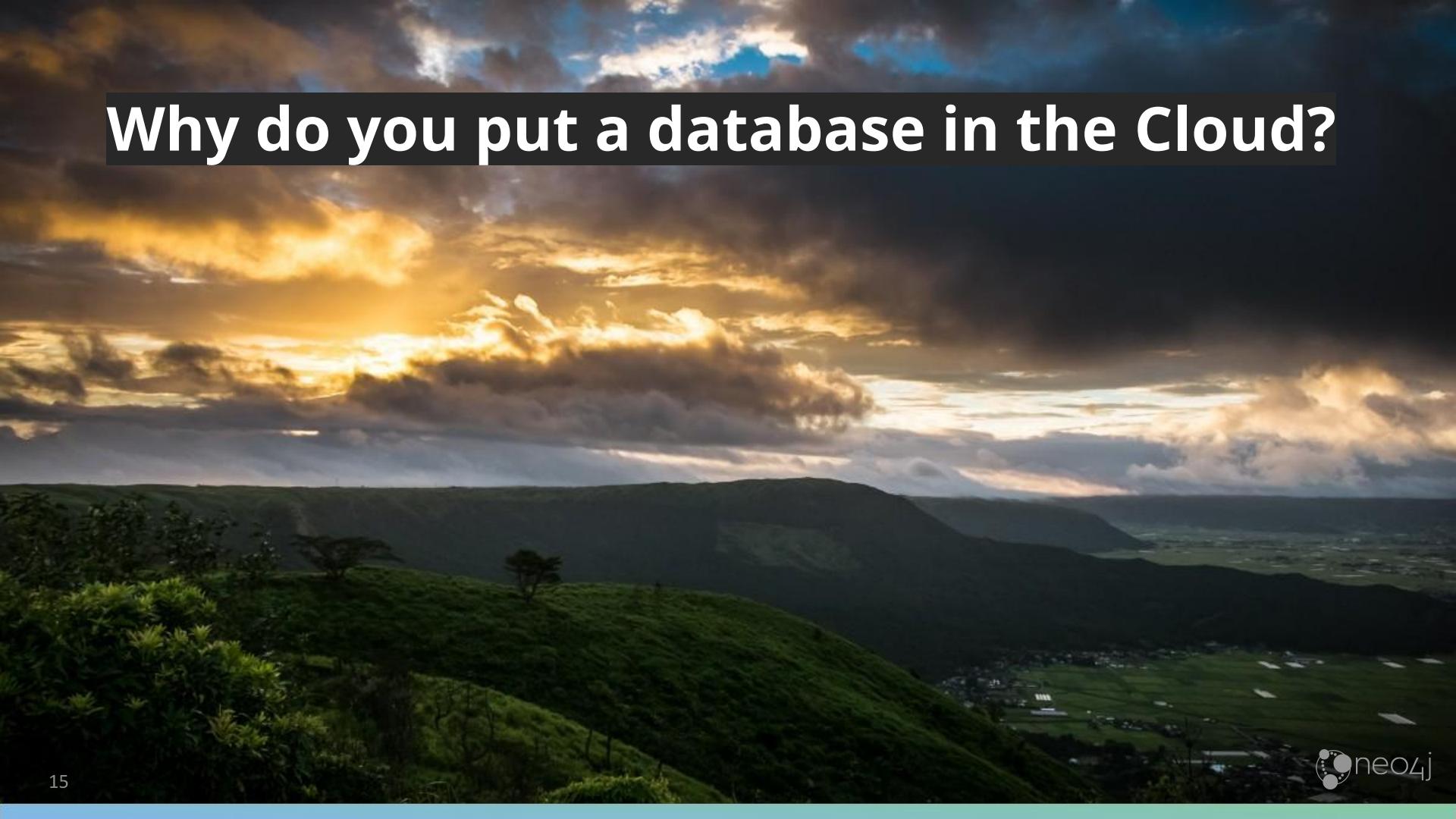


Customer-Facing Applications

- Real-Time Recommendations
- Graph-Based Search
- Identity and Access Management



Why do you put a database in the Cloud?

A wide-angle photograph of a rural landscape at sunset or sunrise. The foreground is dominated by a lush green hillside with some low-lying bushes in the bottom left. In the middle ground, there are more hills and a small cluster of buildings in a valley. The background features a vast sky filled with dramatic, layered clouds. The colors range from deep blues and blacks at the top to bright oranges and yellows near the horizon, suggesting a sunrise or sunset.

Why do you put a database in the Cloud?



Kelsey Hightower A blue circular badge with a white checkmark.

@kelseyhightower

Replies to [@gabidavila](#)

I also prefer a managed database because of all the other benefits such as backups, upgrades, and high availability, none of which Kubernetes can provide for all databases out of the box.

10:07 PM · Mar 25, 2018 · Twitter Web Client

The story so far



Why change it then?





Running clusters on Kubernetes!



HashiConf 2016

Running clusters on Kubernetes!



Kelsey Hightower

@kelseyhightower

Most people get really excited about running a database inside [Kubernetes].

This is going to make you lose your job — guaranteed.

HashiConf 2016



Kelsey Hightower @kelseyhightower

Kubernetes has made huge improvements in the ability to run stateful workloads including databases and message queues, but I still prefer not to run them on Kubernetes.

3:04 PM · Feb 13, 2018 · [Twitter Web Client](#)



Kelsey Hightower

@kelseyhightower

Replying to [@clintkitson](#)

I think it's important to remember that Kubernetes only solves part of the problem. The other parts must be solved by the stateful service and through operational expertise.

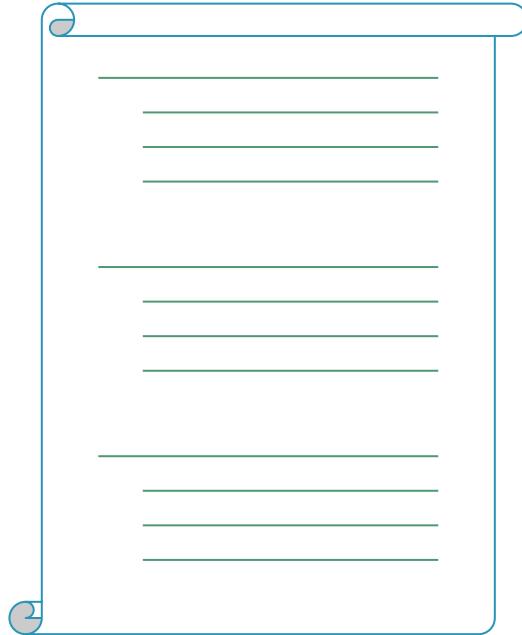
6:09 PM · Feb 13, 2018 · [Twitter Web Client](#)

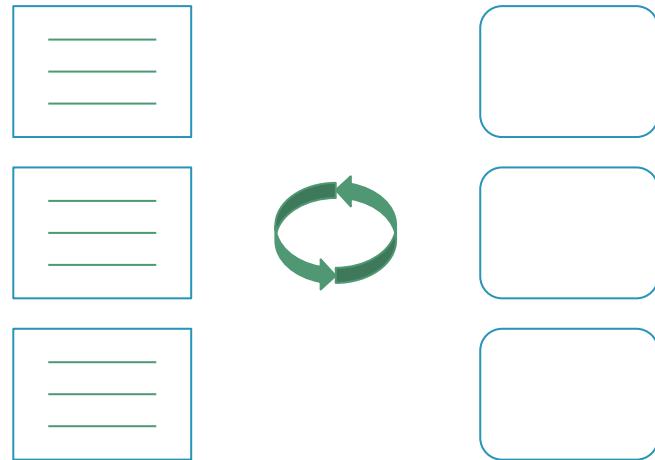
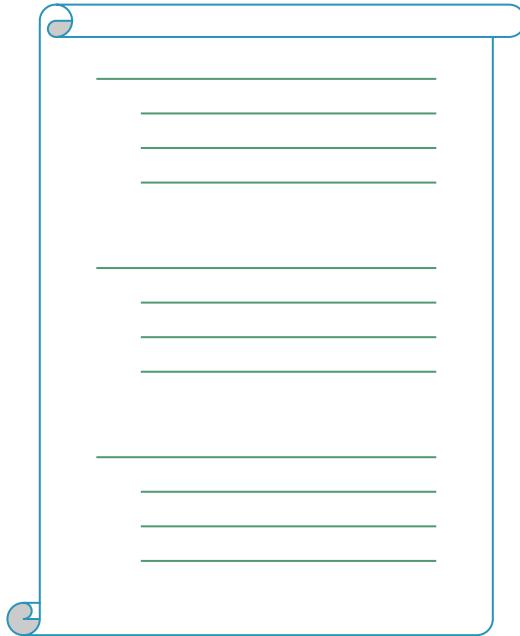
Running databases on Kubernetes!

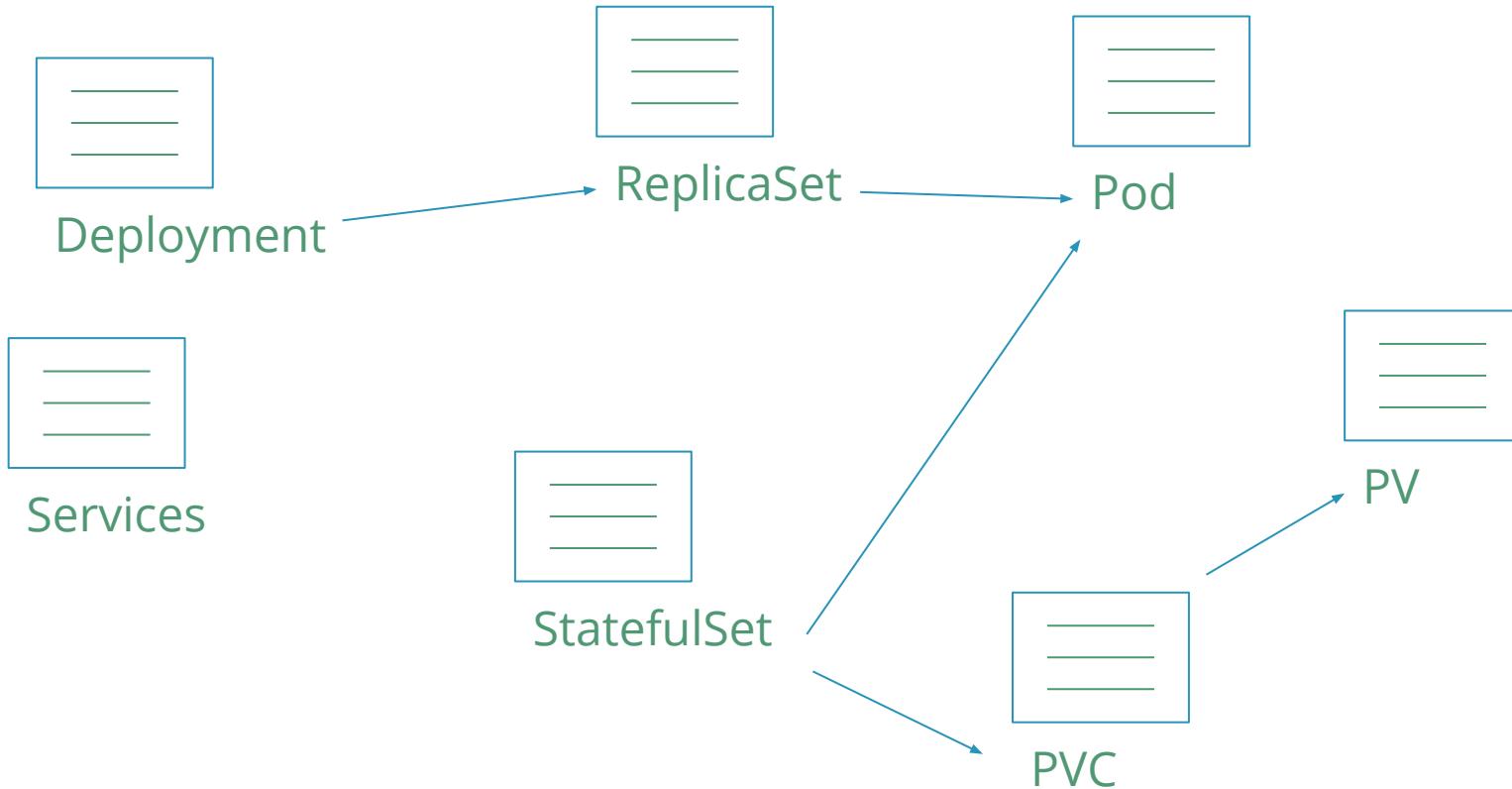


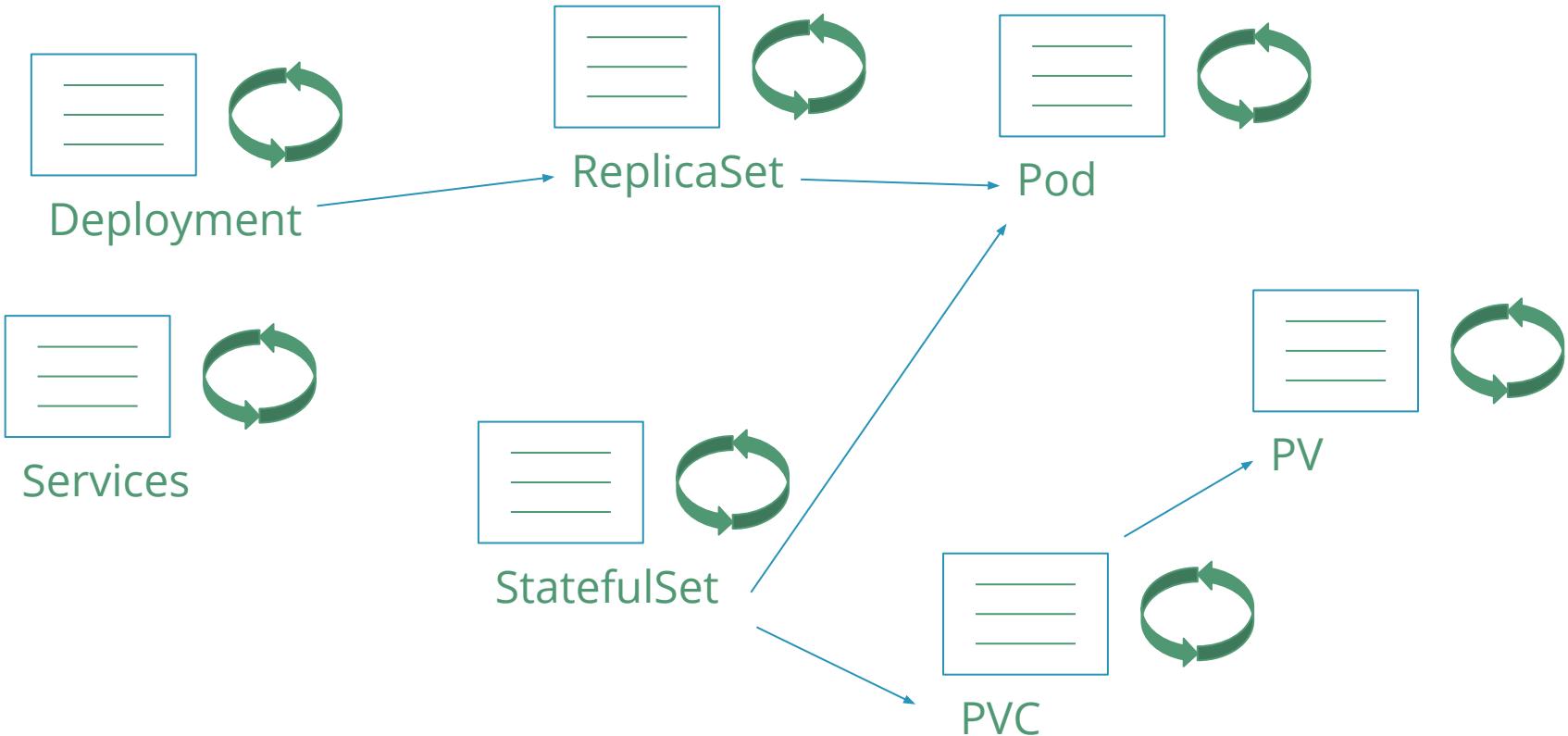
Deep dive Kubernetes

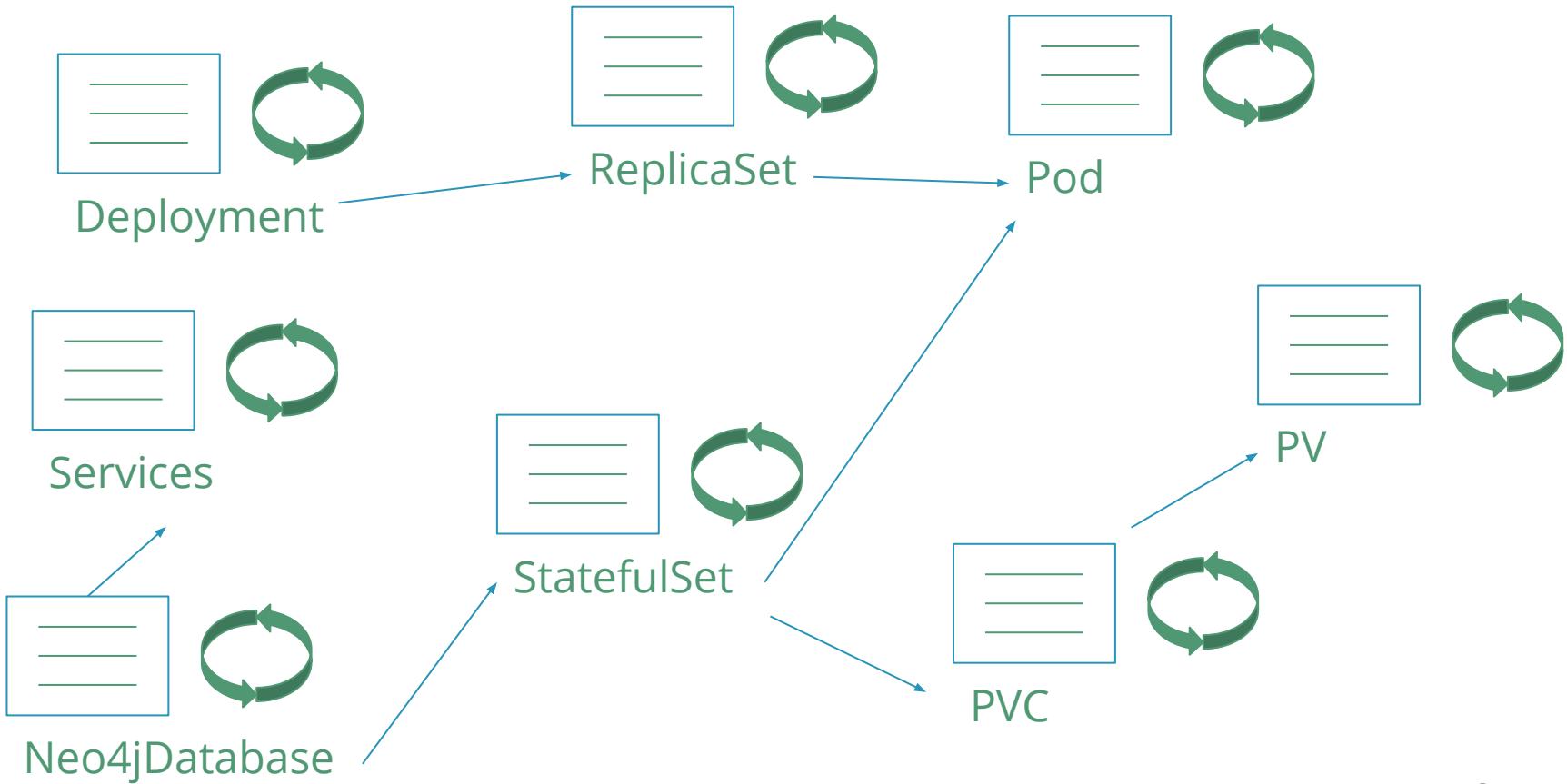








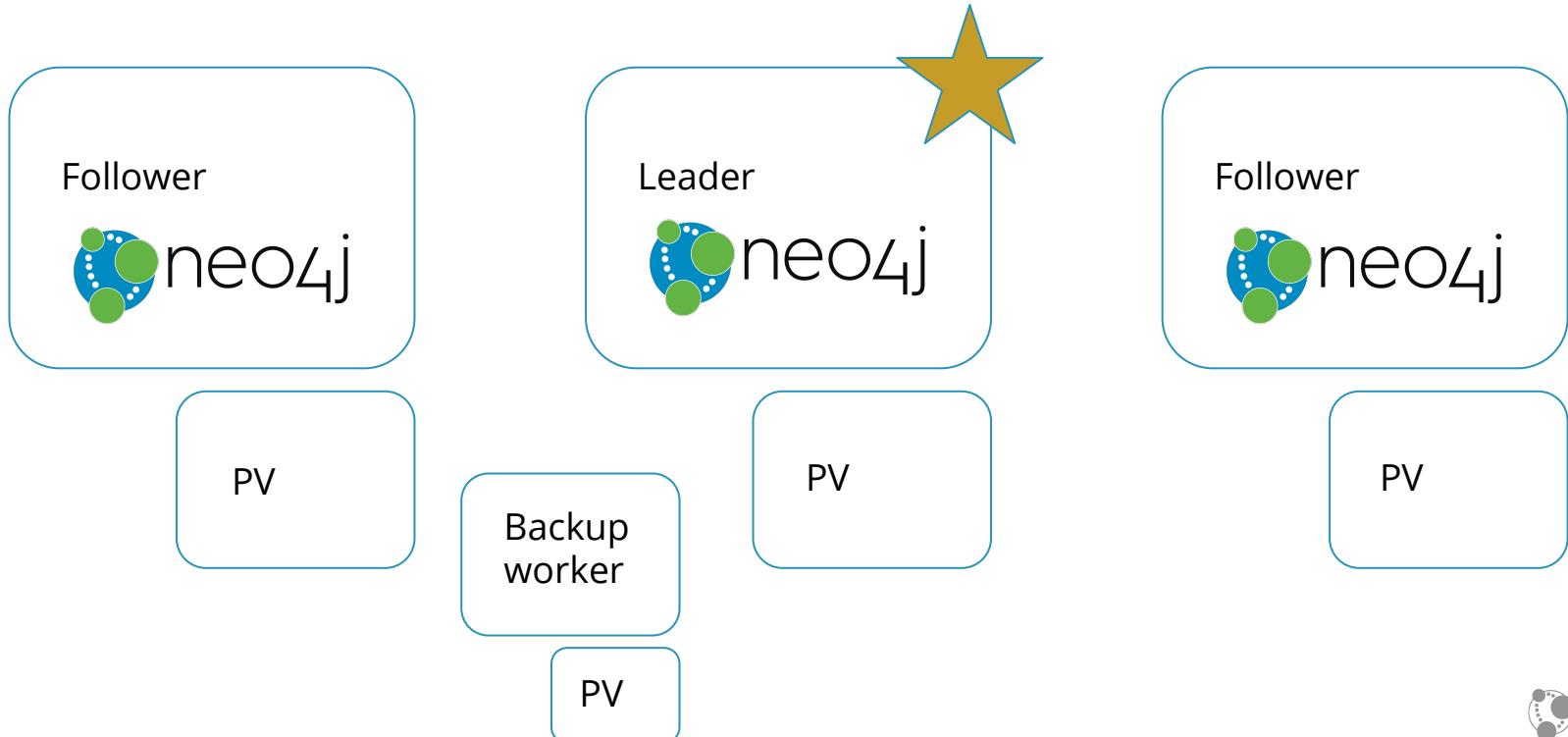




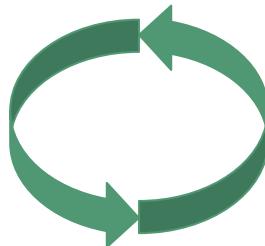
The operator



Neo4j clustering



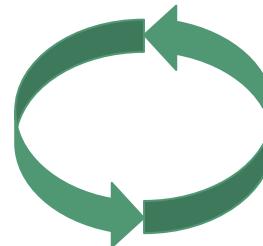
How we structured the reconciler



How we structured the reconciler

backup
load
restore

configs
certs
dns



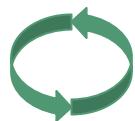
instances

services

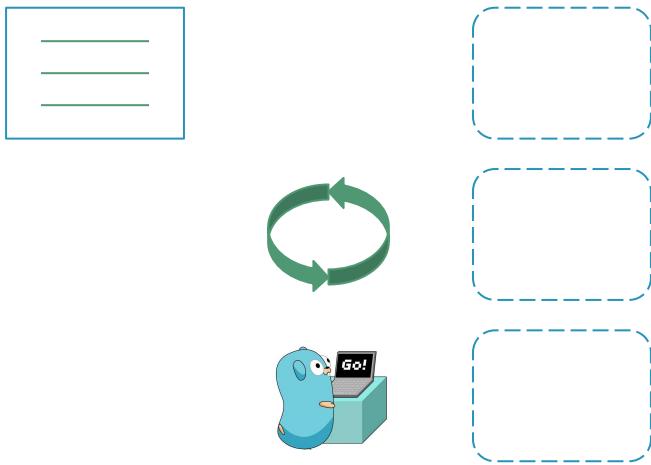
The desired state calculation

- Main reconciler builds actual current and desired state
- Then it loops over sub reconcilers until:
 - all finished
 - or an error happened
 - or a sub reconciler requests changes

Create



Create



Create



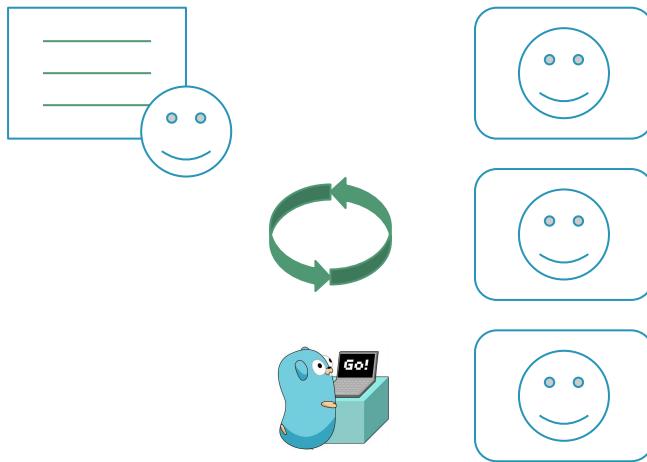
Create



Create



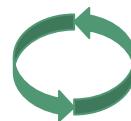
Create



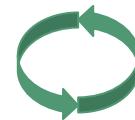
Heal



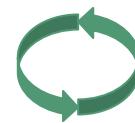
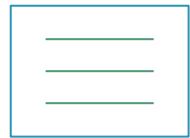
Heal



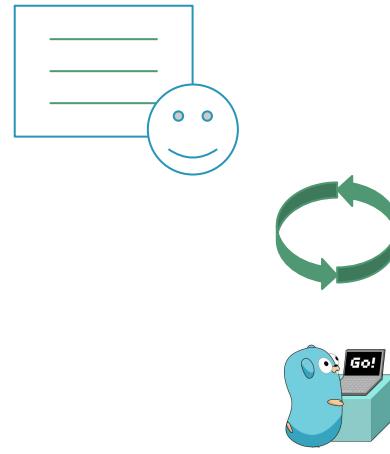
Heal



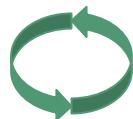
Heal



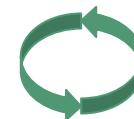
Heal



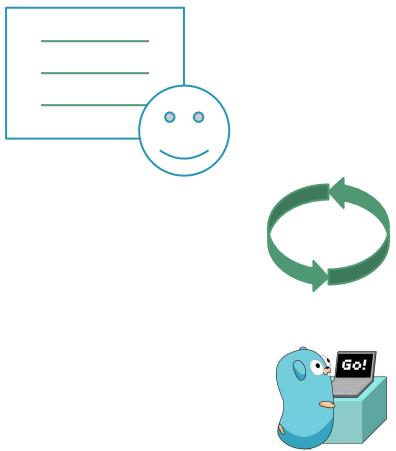
Create



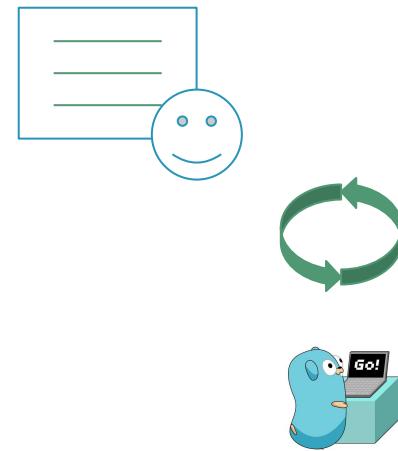
Heal



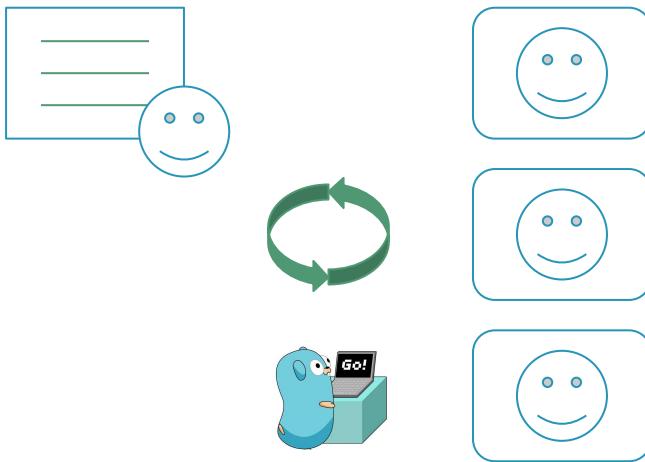
Create



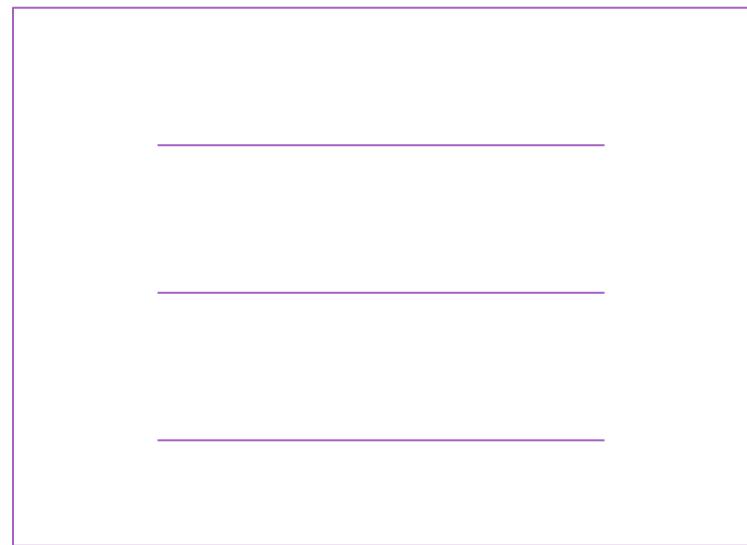
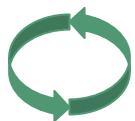
Heal



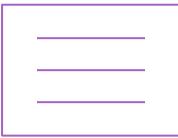
Update



Update



Update

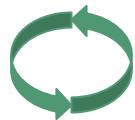
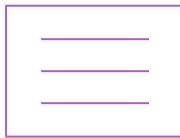


New Neo4j version?

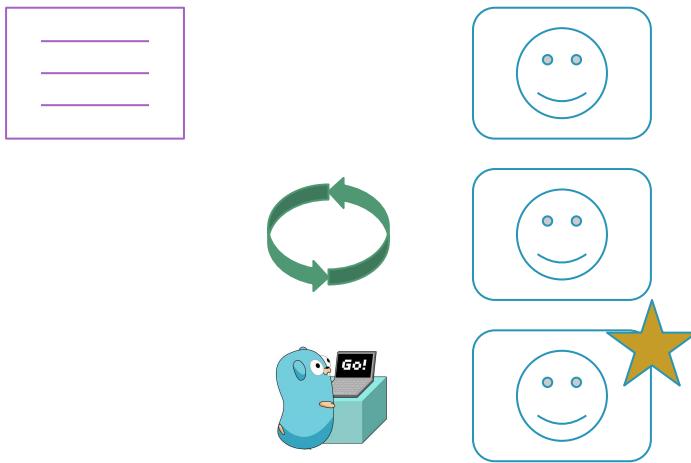
Resource changes?

Password reset?

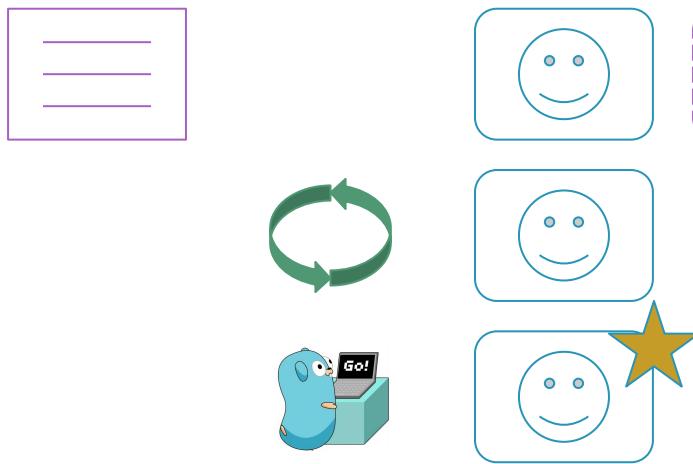
Update



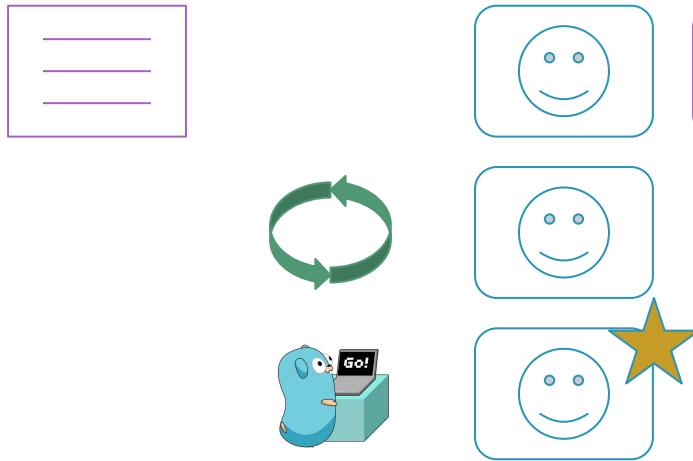
Update



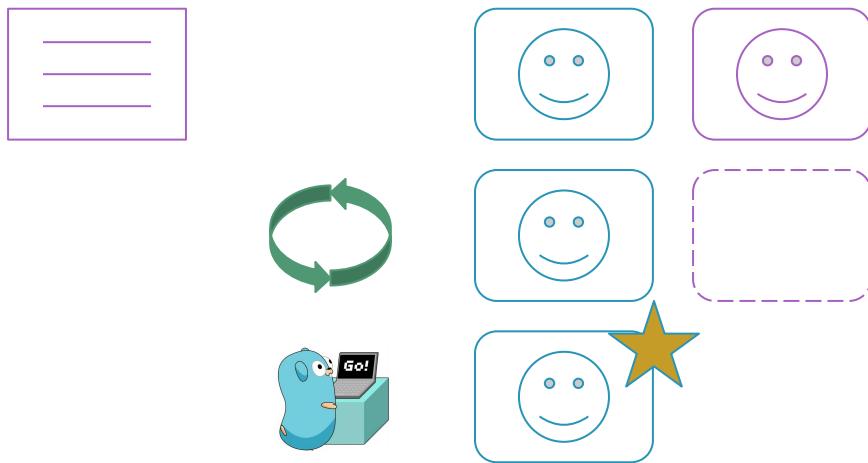
Update



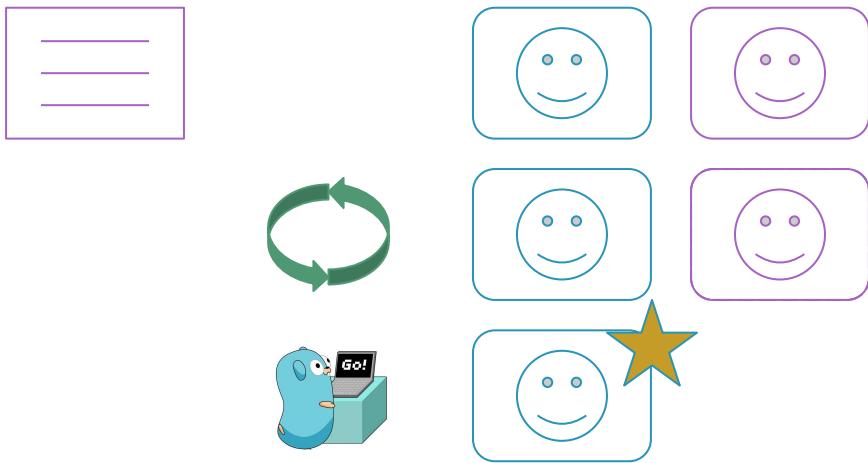
Update



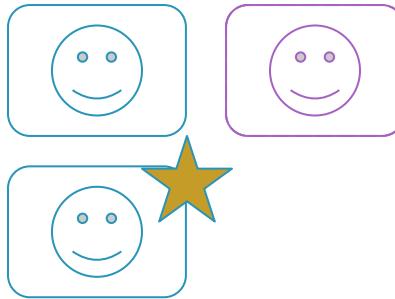
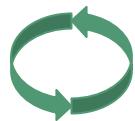
Update



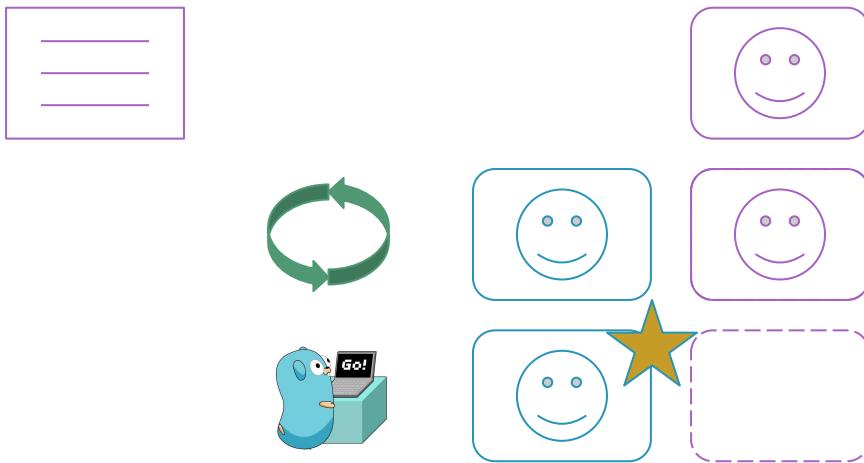
Update



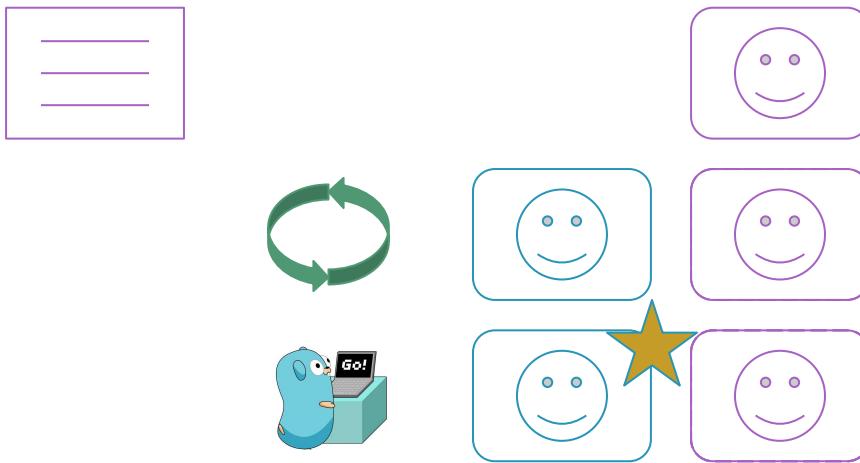
Update



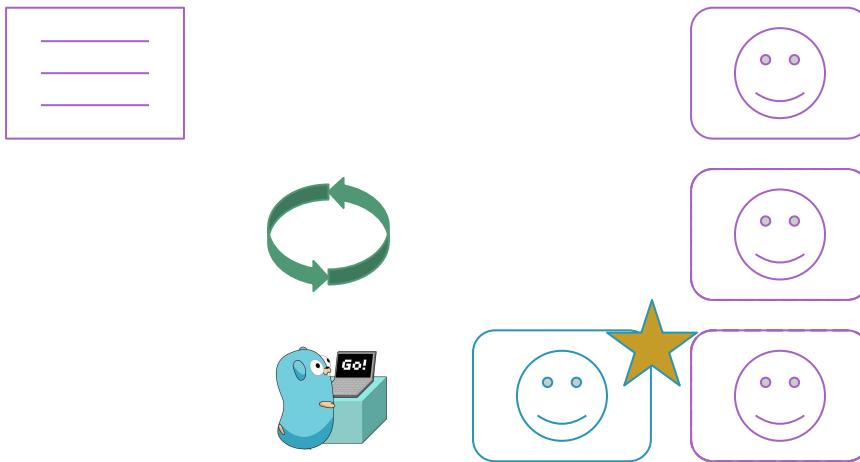
Update



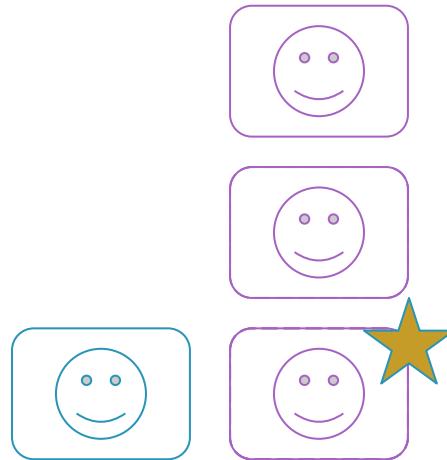
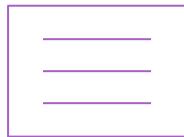
Update



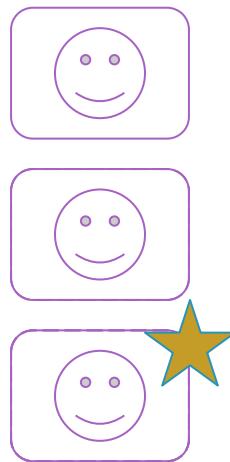
Update



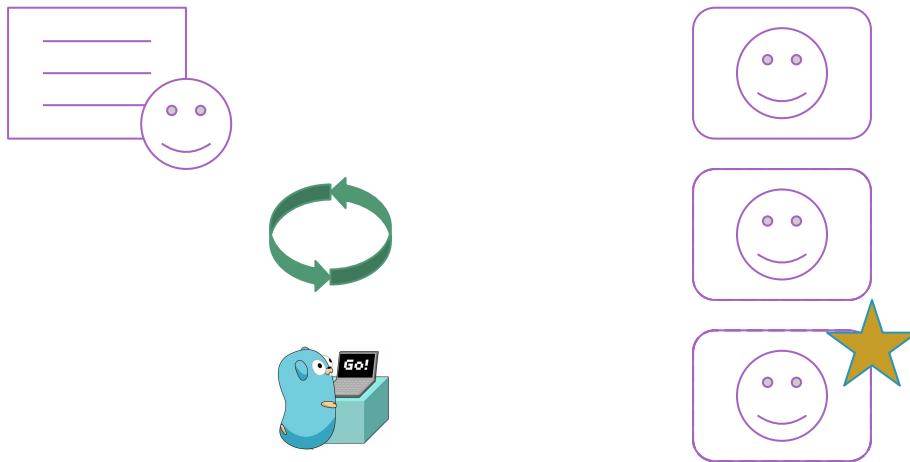
Update



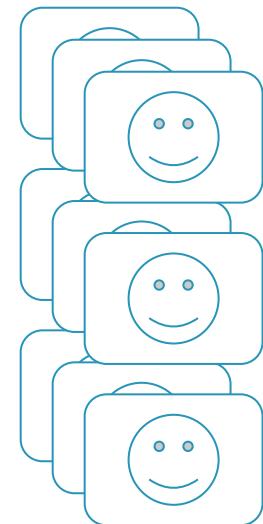
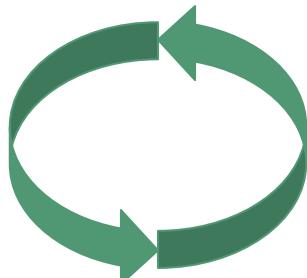
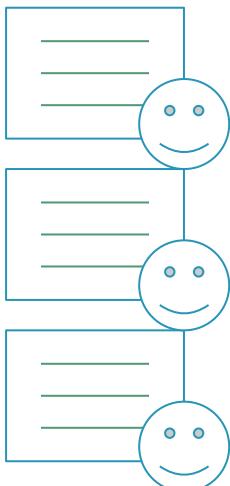
Update



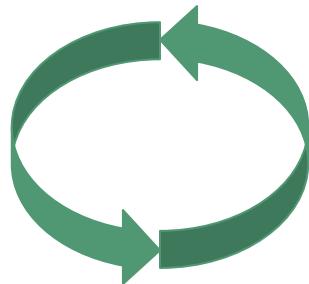
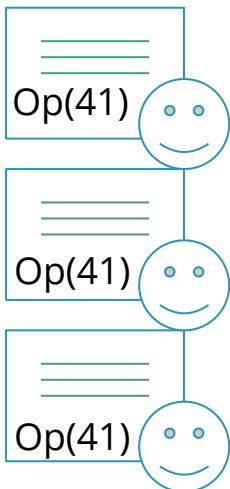
Update



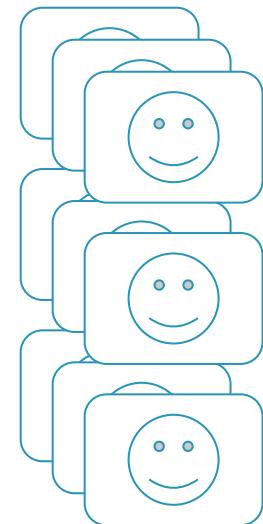
Update the operator itself



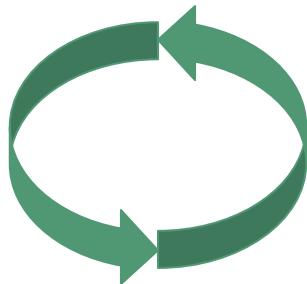
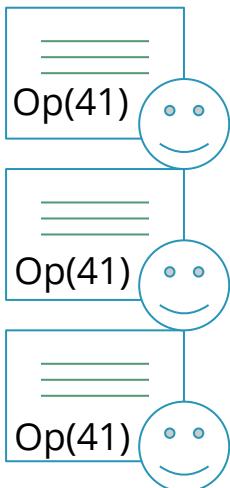
Update the operator itself



Operator 41



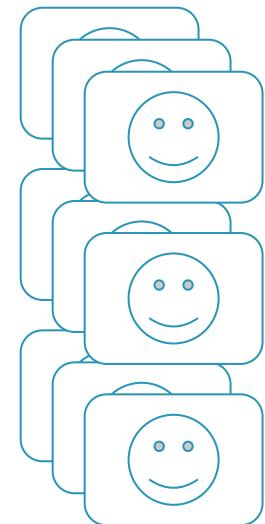
Update the operator itself



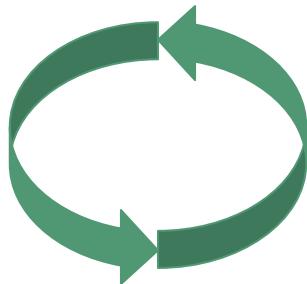
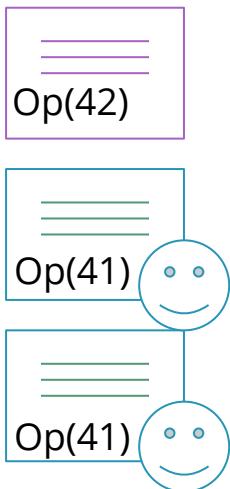
Operator 41



Operator 42

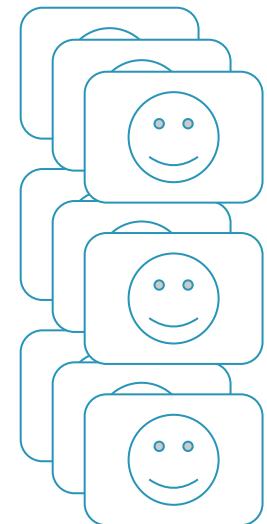


Update the operator itself

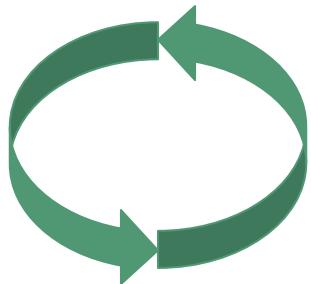
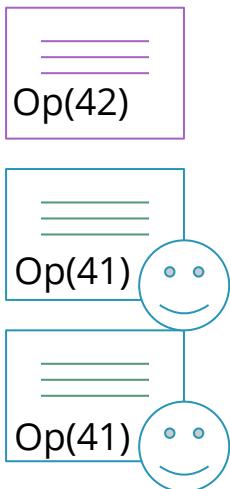


Operator 41

Operator 42

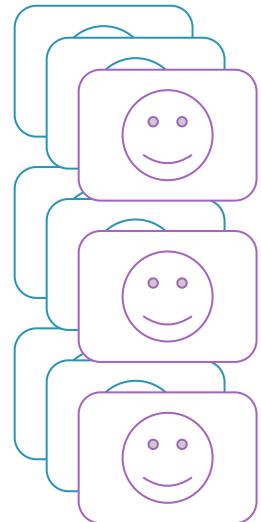


Update the operator itself

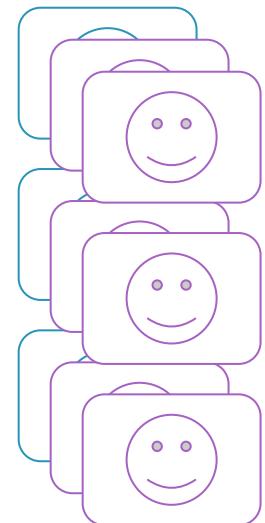
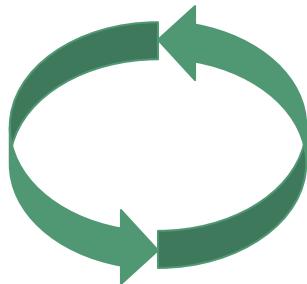
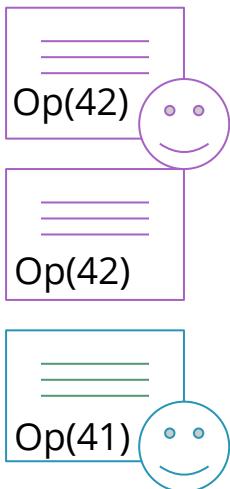


Operator 41

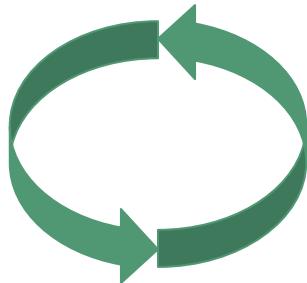
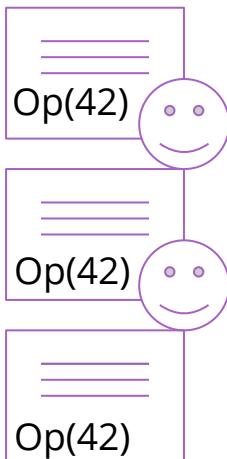
Operator 42



Update the operator itself



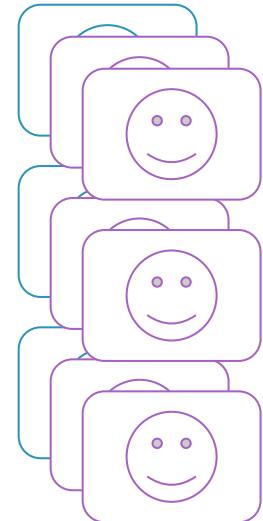
Update the operator itself



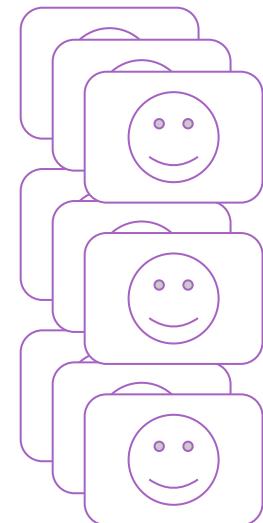
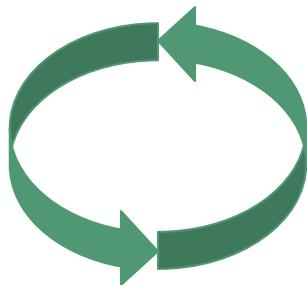
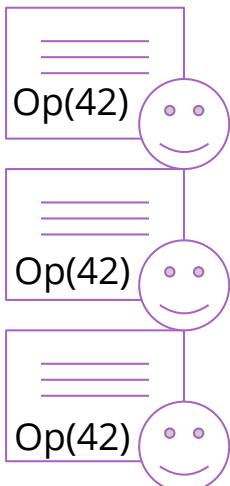
Operator 41



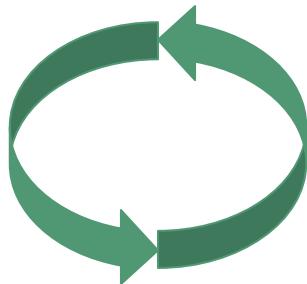
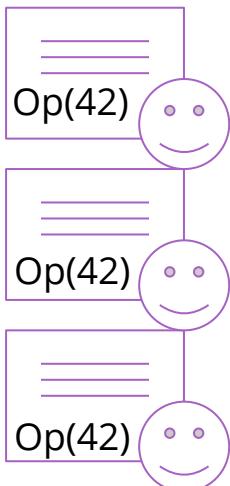
Operator 42



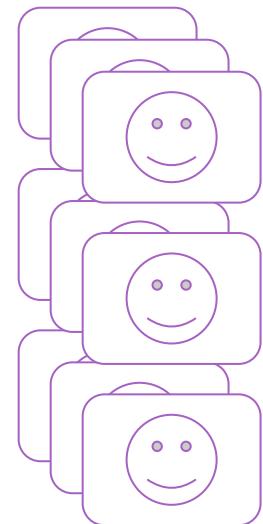
Update the operator itself



Update the operator itself



Operator 42



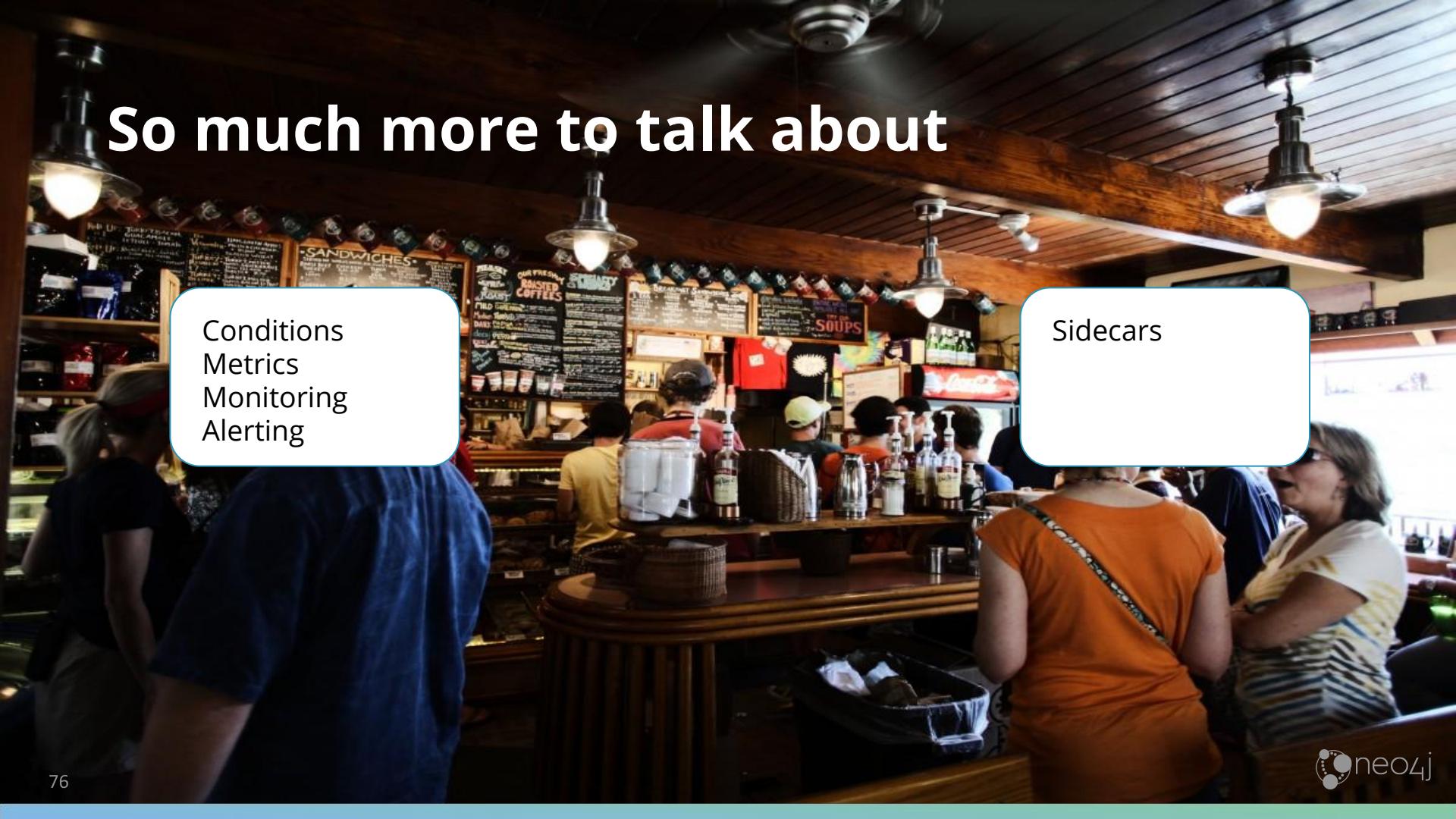
So much more to talk about



So much more to talk about

Conditions
Metrics
Monitoring
Alerting

So much more to talk about

A photograph of a busy restaurant interior. In the foreground, several people are seen from behind, looking towards the counter. The counter is well-stocked with various items, including a large container of ice, glasses, and bottles. Above the counter, there are several chalkboards displaying menus for "SANDWICHES", "ROASTED COFFEE", and "SOUPS". The ceiling is made of dark wood planks, and there are several pendant lights hanging down. The overall atmosphere is casual and bustling.

- Conditions
- Metrics
- Monitoring
- Alerting

- Sidecars

So much more to talk about



Conditions
Metrics
Monitoring
Alerting

Sidecars

Debuggability
'Do not touch' mode



Is everything happy then?



What about testing?



Unit tests

```
tests := map[string]struct {
    states      []*neo4j_status.StatusEndpointResponse
    expected    bool
    boltResponses int
}{}
"expect true when 2 instance have same voting set as leader and 1 has different": {
    states: []*neo4j_status.StatusEndpointResponse{
        healthyResponse(id_2),
        healthyResponse(id_3),
        unhealthyResponse(id_1, t: 0, id_1),
    },
    expected: true,
    boltResponses: 2,
},
"expect false when 2 leaders": {
    states: []*neo4j_status.StatusEndpointResponse{
        healthyResponse(id_2),
        unhealthyResponse(id_1, typeTwoLeaders),
    },
    expected: false,
    boltResponses: 2,
},
}
for name, tc := range tests {
    t.Run(name, func(t *testing.T) {
        actual := AvailableCluster(tc.states, boltResponses, logger)
        assert.Equal(t, tc.expected, actual)
    })
}
```

Integration tests

```
func TestCreateConfigMap(t *testing.T) {
    g := setUpConfigMapReconcilerTests(t)

    mockWriteFacade := &context.MockWriteFacade{}

    cluster := newFakeCluster(neo4japi.DBID("cluster-tcmu"))
    neo4jContext := context.NewFakeNeo4jContext(t, mockWriteFacade)

    actual := NoConfigMap
    desired := NewConfigMapState(NewConfigMap(cluster))

    mockWriteFacade.On( methodName: "CreateConfigMap", mock.Anything, mock.Anything).
        | Return(noError)

    reconciler := NewConfigMapReconciler(actual, desired, neo4jContext)

    g.Expect(reconciler.Reconcile(cluster)).
        | To(BeTrue(), optionalDescription: "Expected reconciliation to be complete")

    mockWriteFacade.AssertCalled(t, methodName: "CreateConfigMap", cluster, desired.ConfigMap())
}
```

System integration tests

A dramatic, low-angle photograph of a fallen knight in full armor. The knight's helmet, which features a prominent plume of dark feathers, lies on a rough, light-colored stone ledge. The armor is highly reflective, showing bright highlights and deep shadows. The knight's body is slumped over, with one arm resting on the ground and the other bent at the elbow. The background is blurred, showing more of the stone wall and the ground, creating a sense of depth and focus on the fallen figure.

Automated E2E tests

The screenshot shows the neo4j Cloud interface for creating a new database. On the left, there's a sidebar with a logo and two options: "neo4j Cloud" and "Dashboard". The main area has a title "Create a new database" and a "Database Name" input field containing "fancy e2e database". Below it is a "RAM Capacity" slider with options from 1 GiB to 64 GiB, with 8 GiB selected. At the bottom are "Cancel" and "Create Database" buttons.

neo4j Cloud

Dashboard

Create a new database

Database Name

fancy e2e database

RAM Capacity

1 GiB 2 GiB 4 GiB **8 GiB** 16 GiB 32 GiB 64 GiB

Cancel Create Database

Canaries and chaos monkeys



Thank you!

We are hiring SREs and DBaaS K8s Devs

👉 l.neo4j.org/dbaas-k8s 👈



@neo4j, @unterstein

