JUST DON'T DO IT, POSTGRESQL IN KUBERNETES

JAN KARREMANS

CHIEF DATA ARCHITECT

CYBERTEC

POSTGRESQL SERVICES & SUPPORT

“THE MOST TRANSFORMATIVE TECH SINCE LINUX.”

MARC LINSTER,
POSTGRESQL FELLOW

PostgreSQL Cloud

TOO MUCH TO SAY, TOO LITTLE TIME

Jan Karremans

Techie in Sales

30 years in databases

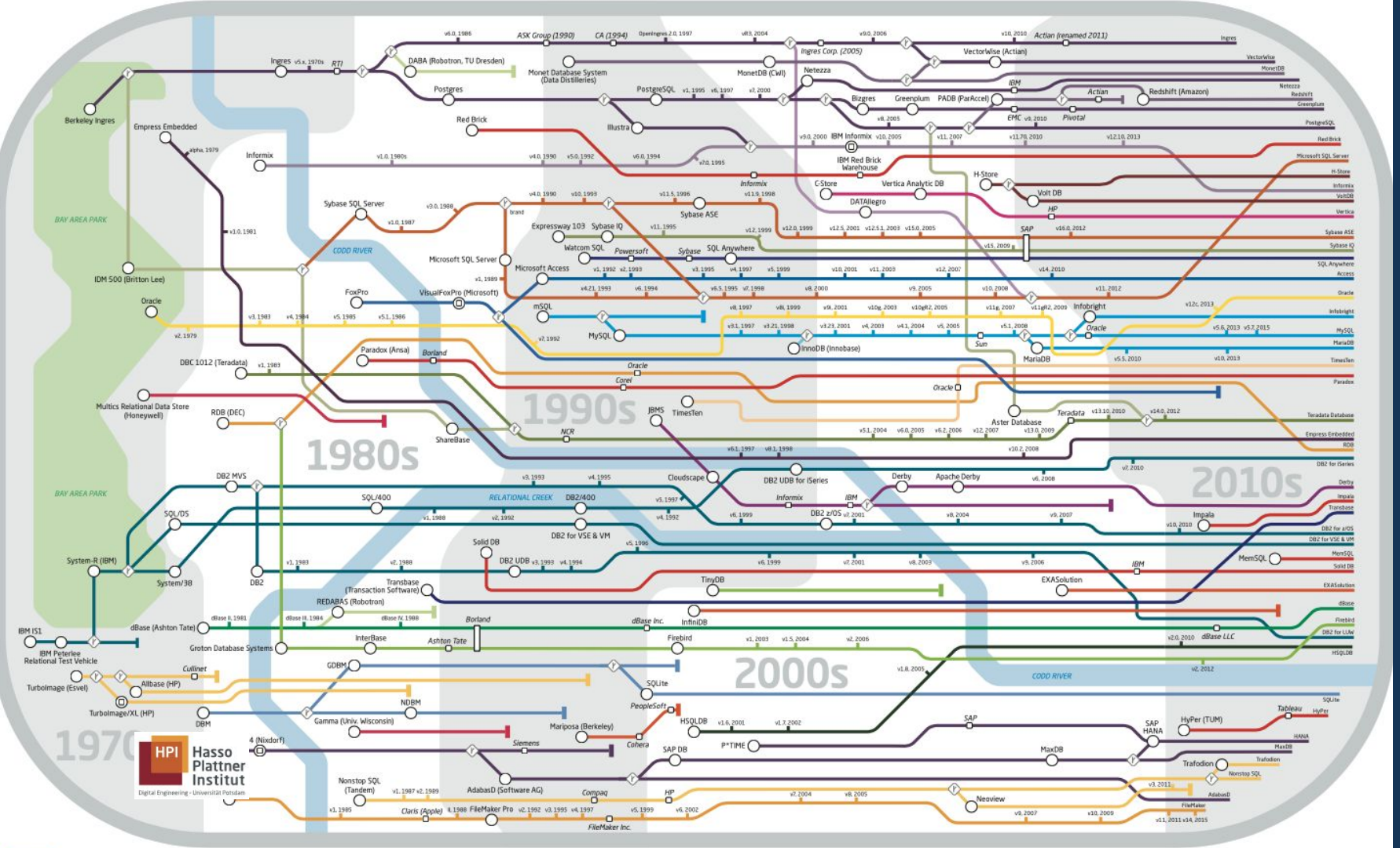
3 years in Kubernetes

On a mission

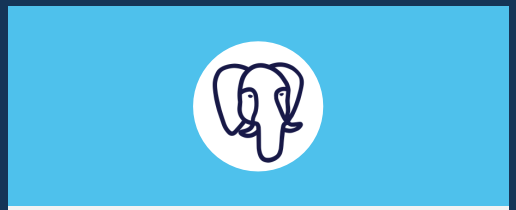
ORACLE ACE Alumni

CERTIFIED kubernetes ADMINISTRATOR -Trained

PostgreSQL Professional Certification


1970s
1980s
1990s
2000s
2010s
BAY AREA PARK
CODD RIVER
RELATIONAL CREEK
Berkeley Ingres
Ingres
Empress Embedded
Informix
Sybase SQL Server
IDM 500 (Britton Lee)
Oracle
FoxPro
VisualFoxPro (Microsoft)
Microsoft SQL Server
Microsoft Access
mSQL
MySQL
Paradox (Ansa)
DBC 1012 (Teradata)
Multics Relational Data Store (Honeywell)
RDB (DEC)
ShareBase
DB2 MVS
SQL/400
SQL/DS
DB2/400
DB2 for VSE & VM
System R (IBM)
System/38
DB2
DB2 UDB v3.1993 v4.1994
Solid DB
Transbase (Transaction Software)
REDABAS (Robotron)
dBase (Ashton Tate)
IBM IS1
IBM Peterlee Relational Test Vehicle
InterBase
Groton Database Systems
Cullinet
Allbase (HP)
TurboImage (Esvel)
TurboImage/XL (HP)
GDBM
NDBM
DBM
Gamma (Univ. Wisconsin)
4 (Nixdorf)
Nonstop SQL (Tandem)
Claris (Apple)
FileMaker Pro
FileMaker Inc.
AdabasD (Software AG)
Siemens
Mariposa (Berkeley)
Cohera
Compaq
HP
SAP DB
HSQLDB
P*TIME
PeopleSoft
SQLite
Firebird
InfiniDB
dBase Inc.
Borland
Ashton Tate
dBase LLC
TinyDB
Transbase
Cloudscape
DB2 UDB for iSeries
DB2 z/OS
Informix
IBM
Derby
Apache Derby
Impala
MemSQL
EXASolution
Aster Database
Teradata
NCR
JBMS
TimesTen
Corel
Oracle
Sun
InnoDB (Innobase)
MariaDB
Infobright
Netezza
Ingres Corp. (2005)
VectorWise (Actian)
Actian (renamed 2011)
Actian
Redshift (Amazon)
Bizgres
Greenplum
PADB (ParAccel)
EMC
Pivotal
Illustra
PostgreSQL
Red Brick
IBM Informix
IBM Red Brick Warehouse
C-Store
Vertica Analytic DB
DATAllegro
H-Store
Volt DB
HP
Sybase ASE
Sybase IQ
Watcom SQL
Powersoft
Sybase
SQL Anywhere
Expressway 103
SAP
HyPer (TUM)
Tableau
SAP HANA
MaxDB
Trafodion
Neoview
HPI
Hasso Plattner Institut
Digital Engineering · Universität Potsdam

LAY OF THE LAND



APPLICATION DEVELOPMENT

MONOLITHIC TO AGILE

Cloud changes much more than just your deployment method

ENABLING SPEED AND INNOVATION

FROM BUZZWORD BINGO TO BUSINESS BENEFIT

Agile

Ensure development and deployment teams can (re)deploy and test quickly and seamlessly

Microservices

Transform traditional monolithic applications to cloud native, microservices based solutions

DevOps

Development and deployment of applications are no longer disjointed operations but fully integrated

CI/CD

New value for your solution, released quickly and securely in short and safe deployment cycles

WHY
DEVOPS

INTRODUCE DORA

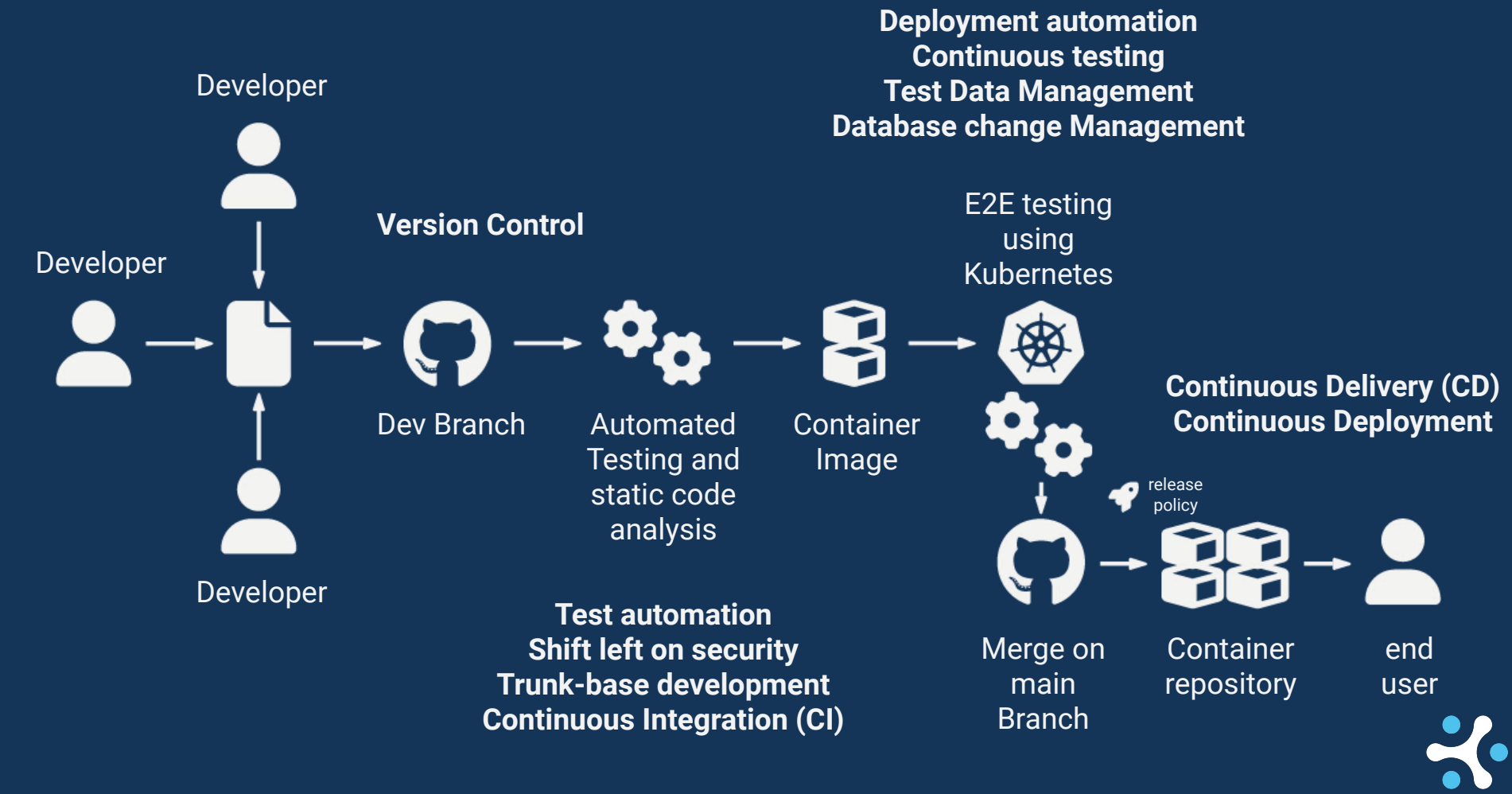
DevOps Research and Assessment

The longest running academically rigorous research investigation of its kind

Providing an independent view into the practices and capabilities

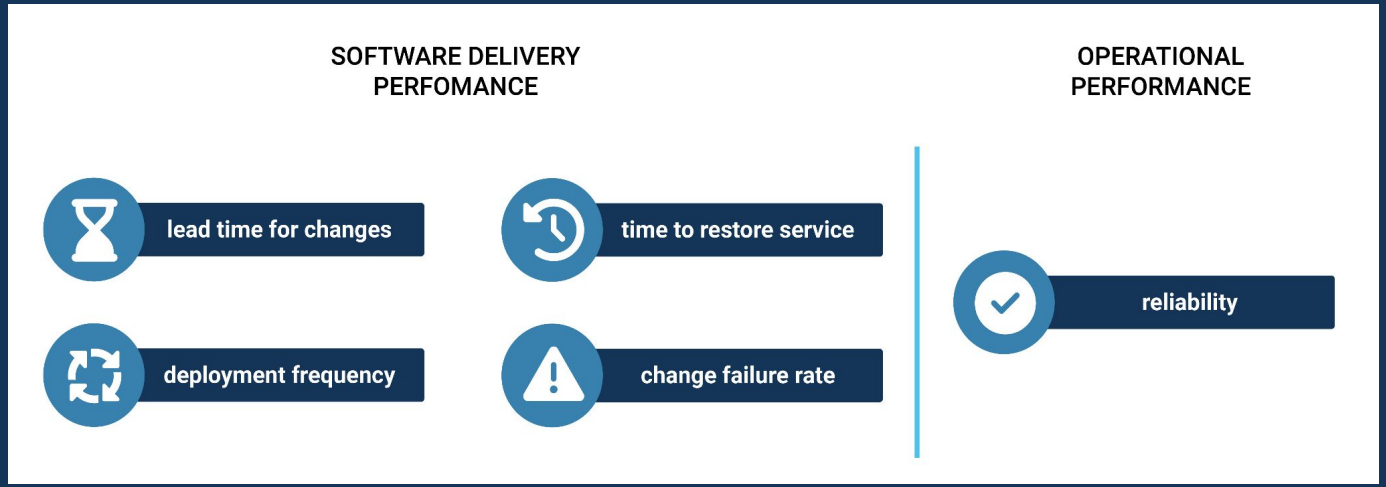
That drive high performance in technology delivery and organizational outcomes

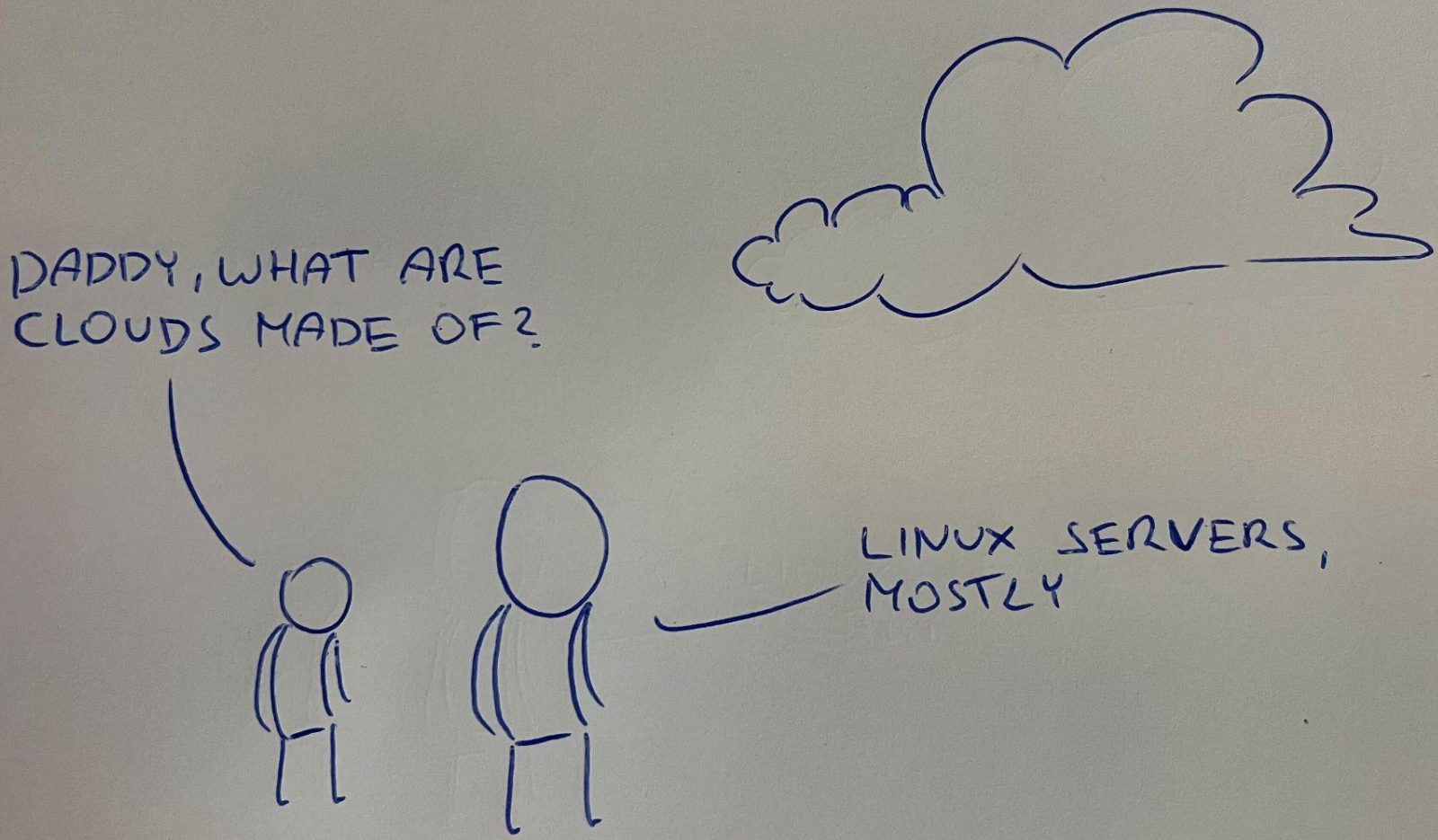
https://www.devops-research.com/research.html


Deployment automation
Continuous testing
Test Data Management
Database change Management
Developer
Developer
Developer
Version Control
E2E testing
using
Kubernetes
Dev Branch
Automated
Testing and
static code
analysis
Container
Image
Continuous Delivery (CD)
Continuous Deployment
release
policy
Test automation
Shift left on security
Trunk-base development
Continuous Integration (CI)
Merge on
main
Branch
Container
repository
end
user

FROM AVAILABILITY TO RELIABILITY

It is all about the metrics

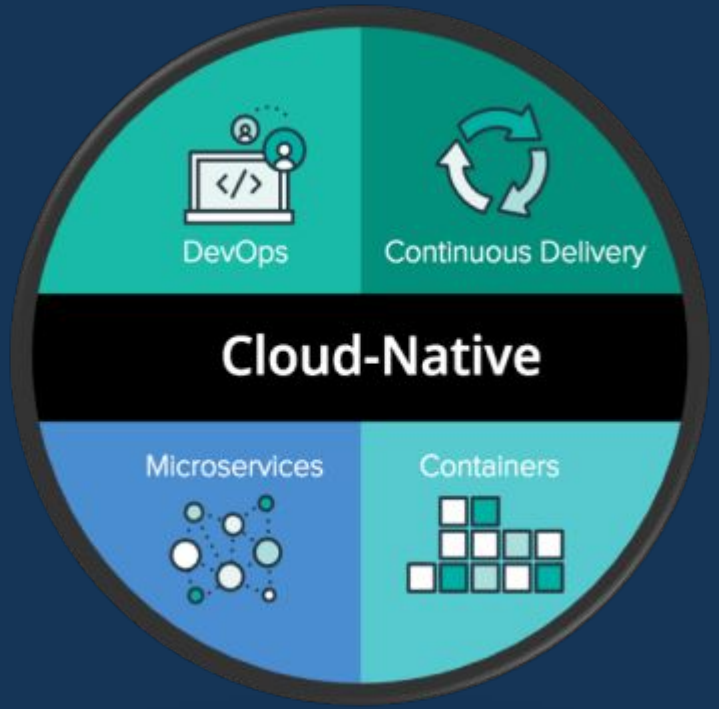



DADDY, WHAT ARE CLOUDS MADE OF?
LINUX SERVERS, MOSTLY

Cloud native technologies empower organizations to **build** and **run scalable applications** in modern, **dynamic environments** such as public, private, and hybrid **clouds**. Containers, service meshes, microservices, immutable infrastructure, and declarative APIs exemplify this approach.

These techniques enable **loosely coupled systems** that are **resilient, manageable, and observable**. Combined with robust automation, they allow engineers to make **high-impact changes frequently** and **predictably with minimal toil**.

The Cloud Native Computing Foundation seeks to drive adoption of this paradigm by fostering and sustaining **an ecosystem of open source**, **vendor-neutral projects**. We democratize state-of-the-art patterns to make these innovations **accessible for everyone.**



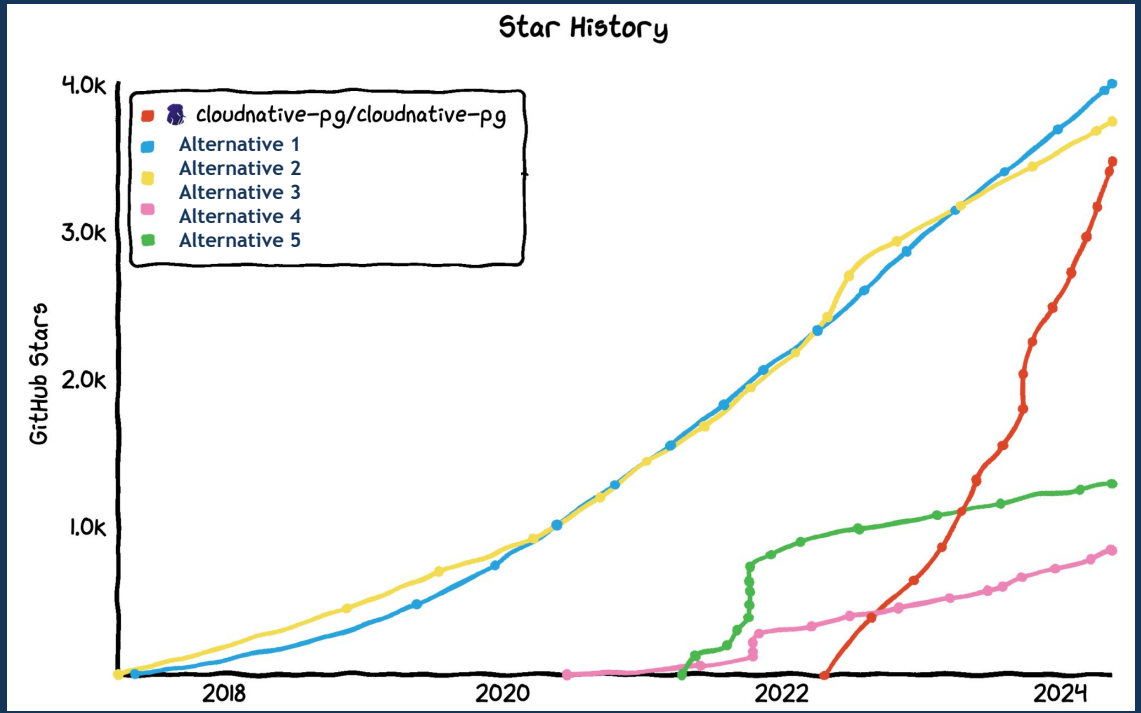
BRINGING IT TOGETHER

VERY POWERFUL COMMUNITIES TOGETHER

&

RUN POSTGRES, THE KUBERNETES WAY

THE LAY OF THE LAND



LAY OF THE LAND

Deploy anywhere
Lightweight, immutable Postgres containers

Automate DBA Tasks
Failover, switchover, backup, recovery, and rolling updates

Avoid lock-in Operator and images are portable to any cloud

ABOUT CLOUD NATIVE PG

- Kubernetes operator
- Day 1 & 2 operations of a PostgreSQL database
  - In traditional environments usually reserved to humans
- Open source
  - Originally created and developed by EDB
  - Vendor neutral/openly governed community
  - Apache 2.0 license
  - Submitted to the CNCF Sandbox
- Production ready
  - BigAnimal - EDB's DBaaS
  - Several EDB customers
- Latest minor version is 1.15

CloudNativePG

Documentation Blog Support End Users

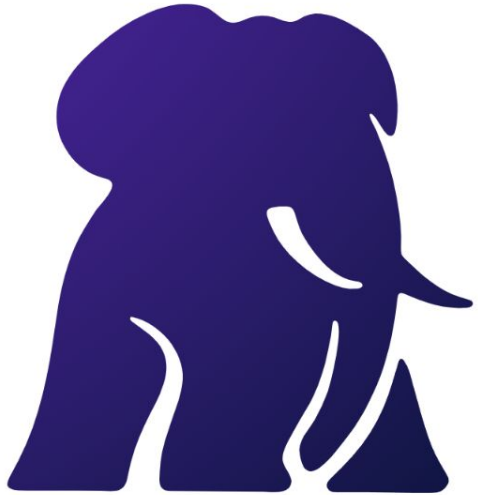
Star 3,458

Run PostgreSQL.

The Kubernetes way.

CloudNativePG is the Kubernetes operator that covers the full lifecycle of a highly available PostgreSQL database cluster with a primary/standby architecture, using native streaming replication.

View on GitHub



lot

he steps that a
r would do to deploy
a Postgres database

Data persistence

It doesn't rely on statefulsets and uses its own way to manage persistent volume claims where the

Designed for Kubernetes

It's entirely declarative, and directly integrates with the Kubernetes API

cloudnative-pg.io

DAY 0
OPERATIONS

PLAN YOUR K8S INFRASTRUCTURE FOR POSTGRESQL WORKLOADS

- **First impressions last**
  - K8s infrastructure often planned for stateless-only workloads
  - Common choice: database outside Kubernetes - DBaaS
- You can run databases inside Kubernetes
  - Fully leverage devops
  - Shared/Shared nothing architectures
  - Storage sector in K8s is growing fast
- Choose your storage wisely
  - Like you are used to in VMs and bare metal

INSTALLING CLOUD NATIVE PG

```
kubectl apply -f \
https://raw.githubusercontent.com/cloudnative-pg/cloudnative-pg/main/releases/cnpg-1.19.1.yaml
```

Declarative configuration via YAML manifest

https://github.com/sergioenterprisedb/kubecon2022-demo

DAY 1
OPERATIONS

OBJECTIVE FOR DAY 0 IS A 3 NODE POSTGRES CLUSTER

- Install the latest minor version of PostgreSQL
- Create a new PostgreSQL Cluster
- One primary and two standby servers
- mTLS authentication with replicas
- 4GB of RAM, 8 cores, 50Gb of storage
- 1GB of shared buffers
- A way to access the primary via network
- A user for the application
- A database for the application

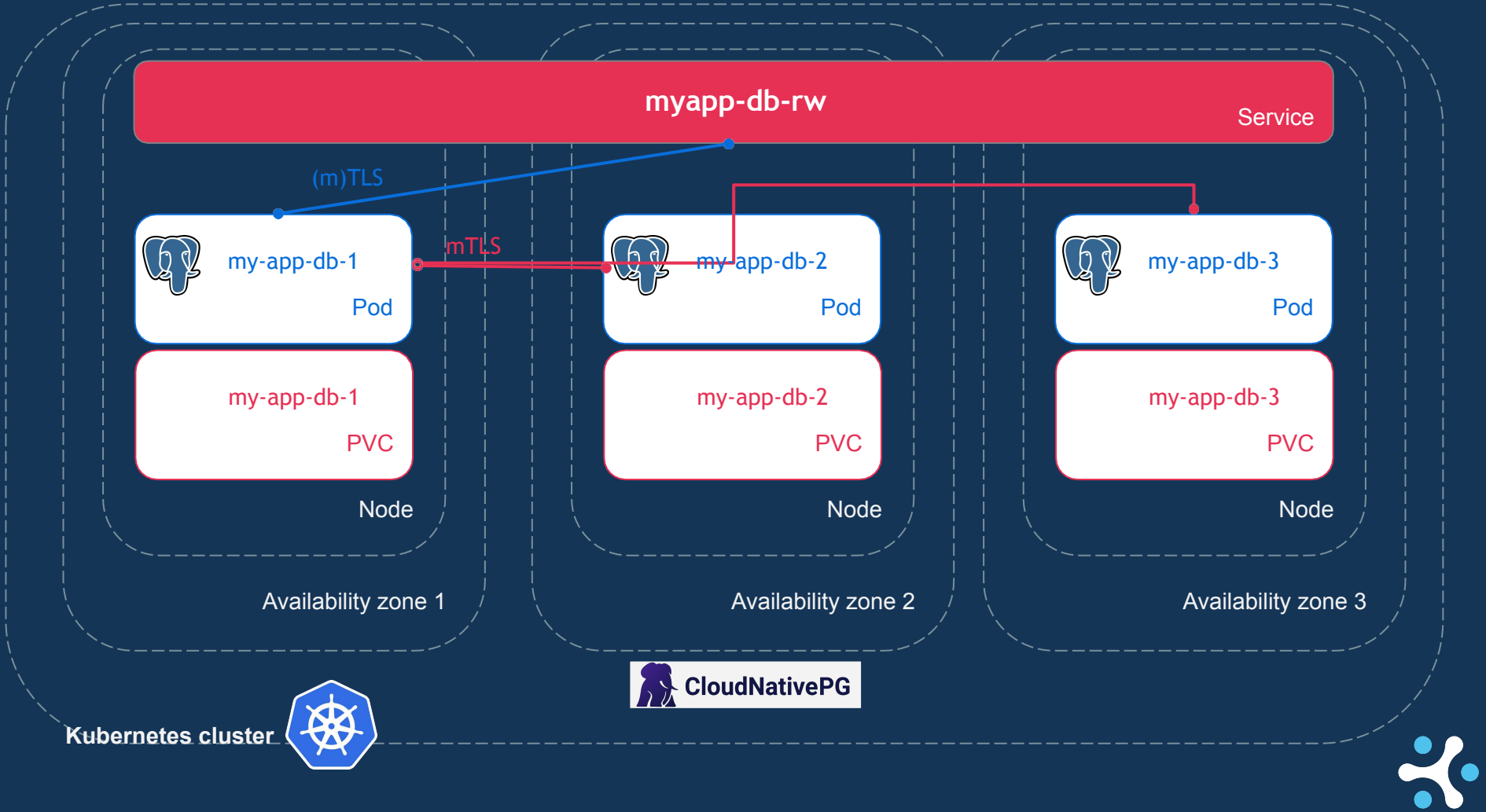
MYAPP-DB.YAML

apiVersion: postgresql.cnpg.io/v1

```yaml
Kind: Cluster
metadata:
  name: myapp-db
spec:
  instances: 3
  postgresql:
    parameters:
      Shared_buffers: "1GB"
resources:
  requests:
    memory: "4Gi"
    cpu: 8
  limits: memory: "4Gi" cpu: 8
storage:
  size: 50Gi
```

HOW TO DEPLOY THE CLUSTER

```
1| kubectl apply   -f myapp-db.yaml
```


myapp-db-rw
Service
(m)TLS
mTLS
my-app-db-1
Pod
my-app-db-1
PVC
Node
my-app-db-2
Pod
my-app-db-2
PVC
Node
my-app-db-3
Pod
my-app-db-3
PVC
Node
Availability zone 1
Availability zone 2
Availability zone 3
CloudNativePG
Kubernetes cluster

THERE'S MORE

- A service to access read-only replicas (myapp-db-ro)
- A service to access any instance for reads (myapp-db-r)
- Many other Kubernetes objects are created:
  - Secrets
  - ConfigMaps
  - Roles
  - RoleBindings
  - ServiceAccounts
  - …
- Convention over configuration

POSTGRESQL CONFIGURATION

- Most GUCs are configurable
  - `.postgresql.parameters` section
  - Some cannot be changed (e.g. log_destination)
  - Some have defaults
- Host-Based Authentication can be configured
  - `.postgresql.pg_hba` section
  - By default:
    - Requires TLS authentication for streaming replicas
    - Fallback sets sha-256/md5 authentication

POSTGRESQL CONFIGURATION

- CloudNativePG supports changes of configuration
  - Reload
  - Rolling updates if restart is required
  - **Update of standby sensitive parameters**

DAY 2
OPERATIONS

THE ROLE OF A KUBERNETES OPERATOR FOR POSTGRES

- Simulate the work of a human DBA
- Do it in a programmatic and automated way
- Extend the Kubernetes API server
  - The only authority for the whole infrastructure
  - Single source of truth of the status of the infrastructure
    - Current status
    - Desired status

ROLLING UPDATES

- Update of a deployment with ~zero downtime
  - Standby servers are updated first
  - Then the primary:
    - supervised / unsupervised
    - switchover / restart
- When they are triggered:
  - Security update of Postgres images
  - Minor update of PostgreSQL
  - Configuration changes when restart is required
  - Update of the operator
    - Unless in-place upgrade is enabled

BACKUP AND RECOVERY

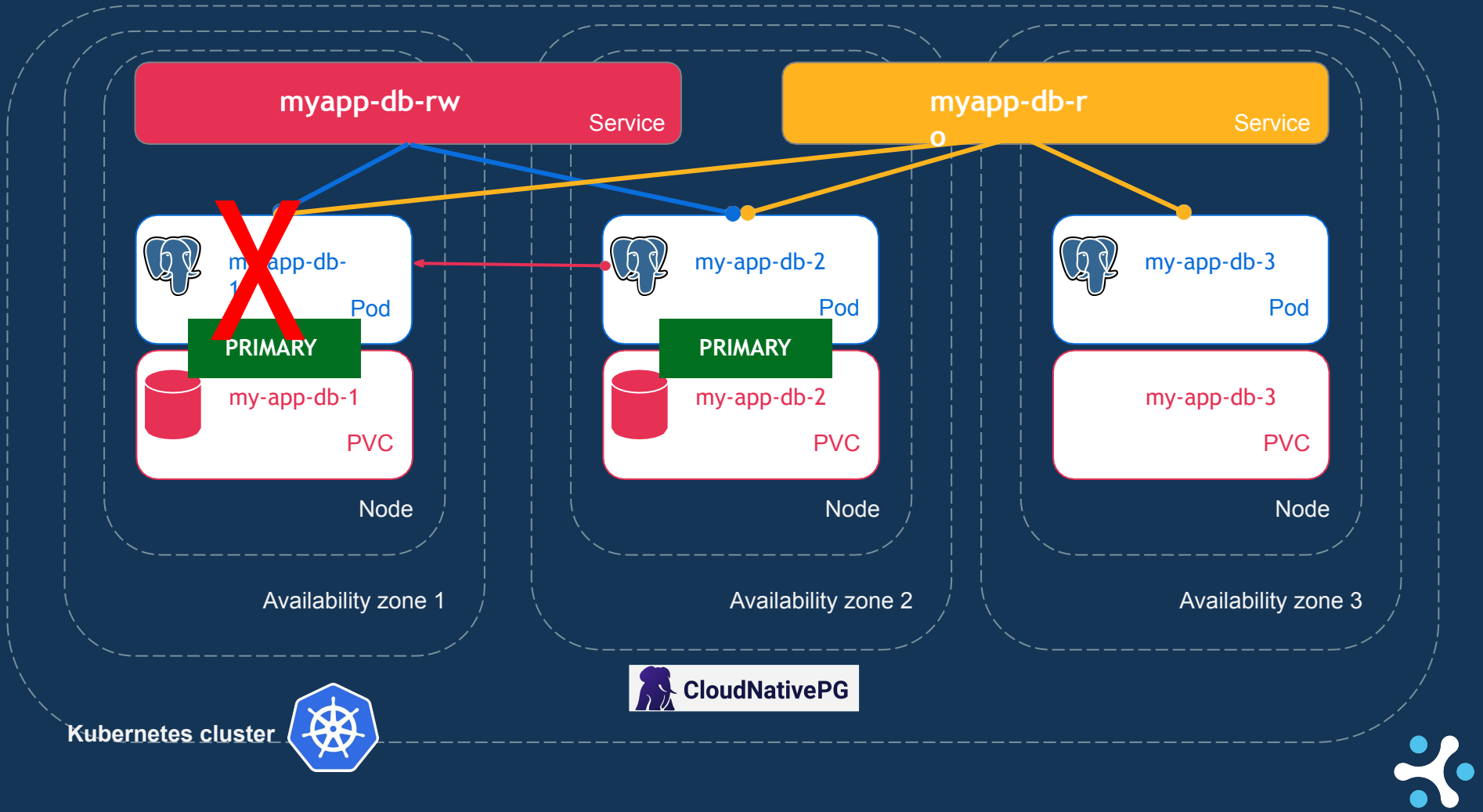
- Continuous physical backup on “backup object stores”
  - Scheduled and on-demand base backups
  - Continuous WAL archiving (including parallel)
- Support for recovery window retention policies (e.g. 30 days)
- Recovery means creating a new cluster starting from a “recovery object store”
  - Then pull WAL files (including in parallel) and replay them
  - Full (End of the WAL) or PITR
- Both rely on Barman Cloud technology
  - AWS S3
  - Azure Storage compatible
  - Google Cloud Storage

SYNCHRONOUS REPLICATION

- Quorum-based synchronous streaming replication
- Controlled by two options:
  - `minSyncReplicas`
  - `maxSyncReplicas`
- CloudNativePG takes care of `synchronous_standby_names`
  - `ANY q (pod1, pod2, ...)`
  - Where:
    - `1 <= minSyncReplicas <= q <= maxSyncReplicas <= readyReplicas`
    - pod1, pod2, ... is the list of all PostgreSQL pods in the cluster

MONITORING

- Native support for Prometheus
- Built-in metrics at the operator level
- Built-in metrics at the Postgres instance level
- Customizable metrics at the Postgres instance level
  - Via ConfigMap(s) and/or Secret(s)
  - Syntax compatible with the PostgreSQL Prometheus Exporter
  - Auto-discovery of databases
  - Queries are:
    - Transactionally atomic and read-only
    - Executed with the `pg_monitor` role
    - Executed with `application_name set to cnp_metrics_exporter`
- Support for `pg_stat_statements and auto_explain`


myapp-db-rw
Service
myapp-db-r
o
Service
my-app-db-2
Pod
my-app-db-3
Pod
PRIMARY
PRIMARY
my-app-db-1
PVC
my-app-db-2
PVC
my-app-db-3
PVC
Node
Node
Node
Availability zone 1
Availability zone 2
Availability zone 3
CloudNativePG
Kubernetes cluster

FINAL REMARKS

RESHAPING THE DBA ROLE

- Most infrastructure related problems are automated
- You as a DBA are crucial in the organization
  - Leverage skills and experience from traditional environments
  - Subject Matter Expert of PostgreSQL in DevOps teams
- Unlearn to learn
- Protect Postgres, from Day 0:
  - Infrastructure: choose the right storage!
  - Application: model the database with developers!
- Examples of day 2 operations:
  - Infrastructure: monitoring, alerting, backup verification
  - Application: query optimization, index optimization, data modeling

JOIN US!

- We adopt the CNCF code of conduct
- Simple governance model based on maintainers for the initial phase
- Public roadmap using GitHub Projects beta
- Start from the CONTRIBUTING.md file
  - GitHub issues and discussions primarily
  - Slack channel
  - Participate to the biweekly developer meetings
- Special instructions for source code contributions
  - Work in progress
  - Setup of the dev environment
  - Setup of the test environment to run E2E tests with kind and k3d
  - Developer Certificate of Origin (DCO) required

WE ARE HIRING!

CHECK OUT OUR JOB OPENINGS:

https://www.cybertec-postgresql.com/en/jobs-and-opportunities/

CYBERTEC
POSTGRESQL SERVICES & SUPPORT

QUESTIONS, ANYONE?

GET IN TOUCH

WE'D LOVE TO HEAR YOUR THOUGHTS

EMAIL

jan.karremans@cybertec-postgresql.com

PHONE

+31 6 1638 9607

WEB

www.cybertec-postgresql.com