2ndQuadrant®
PostgreSQL

Multi-Master PostgreSQL Architectures in Cloud

Gülçin Yıldırım Jelínek

select * from me;

Board of Directors @ PostgreSQL Europe

Cloud Services Manager @ 2ndQuadrant

Main Organizer @ Prague PostgreSQL Meetup

Member @ Postgres Women

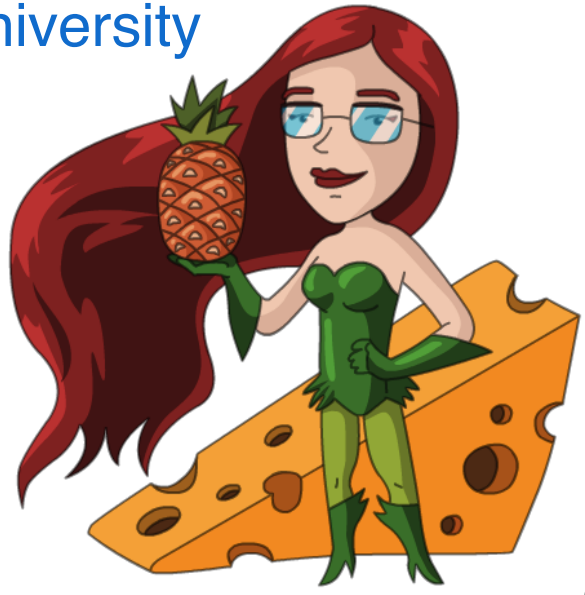
MSc, Computer & Systems Eng. @ TalTech

BSc, Applied Mathematics @ Yildiz Technical University

Writes on 2ndQuadrant blog

From Turkey

Lives in Prague



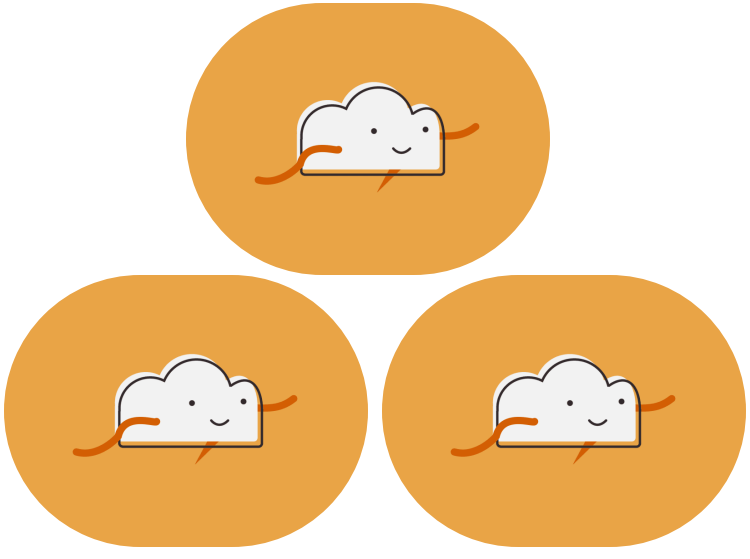
@apatheticmagpie

Github: gulcin

Agenda

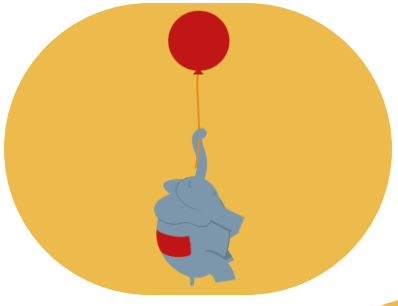
- What is Multi-Master?
- Physical vs Logical Replication
- Postgres-BDR
- Multi-Master Architectures
- Multi-Master Postgres in Cloud (GDS)
- Our Experience with Cloud Postgres
- Conclusion
- Questions

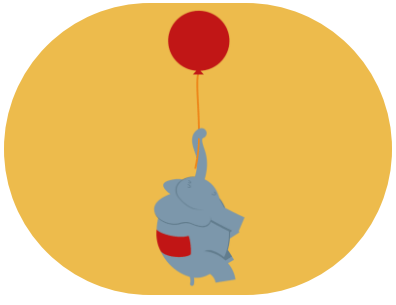
What is Multi-Master?



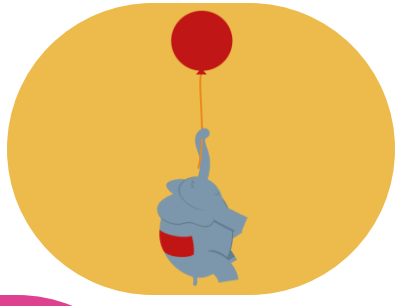
- Replicate writes between multiple masters
  - Asynchronous with conflicts
  - Conflict-free (consensus)

Physical vs Logical Replication



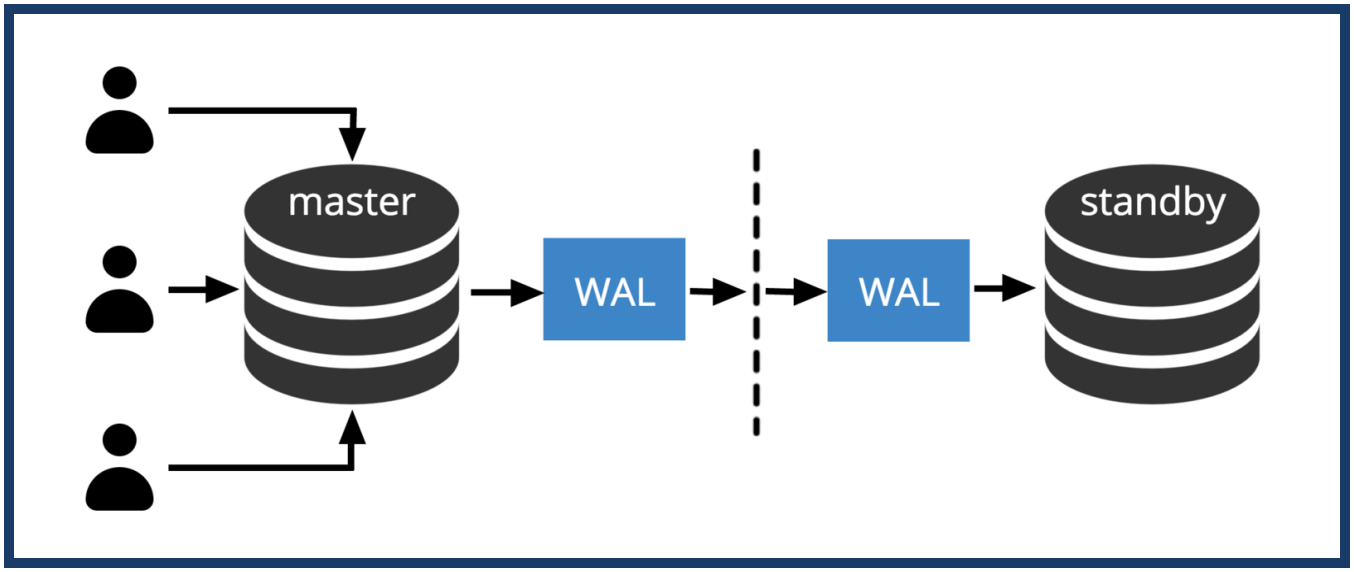


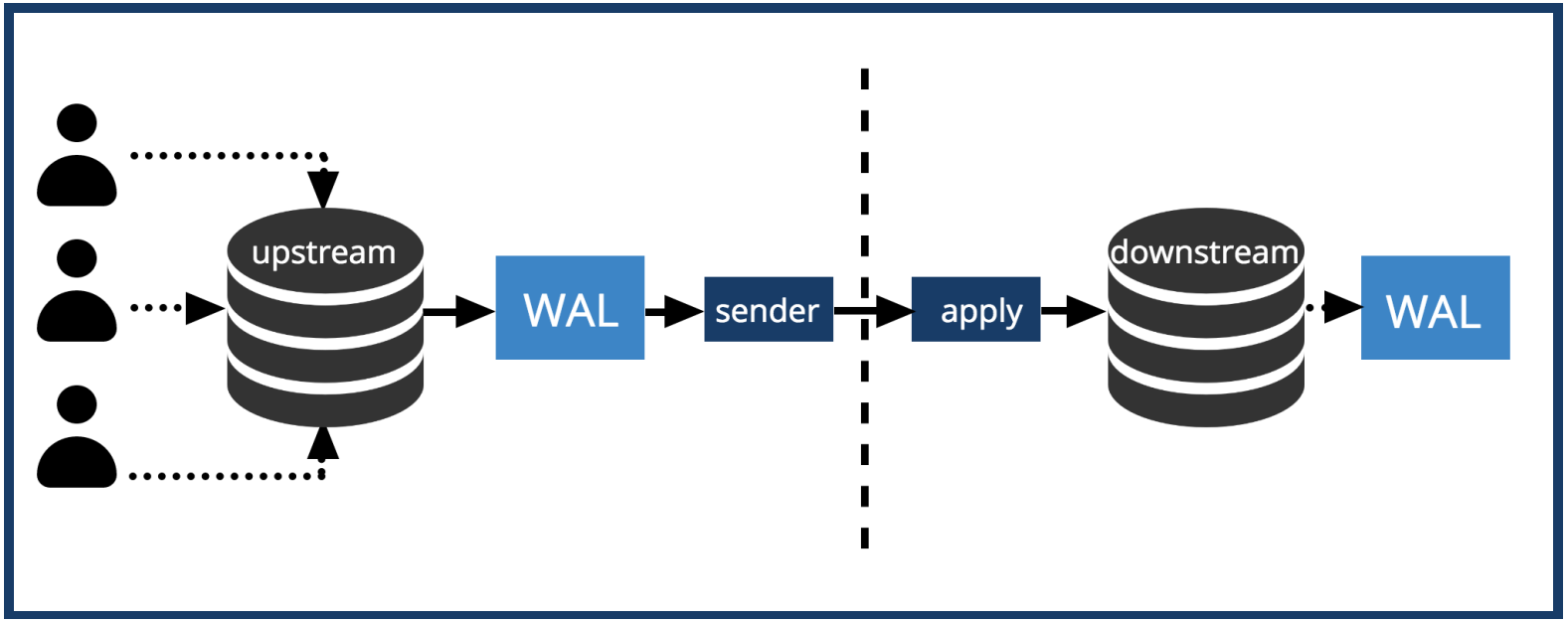
- Replay WAL
  - WAL cannot be changed
- Bit copy of upstream
  - All or nothing
  - Same major version
- Can't write to downstream



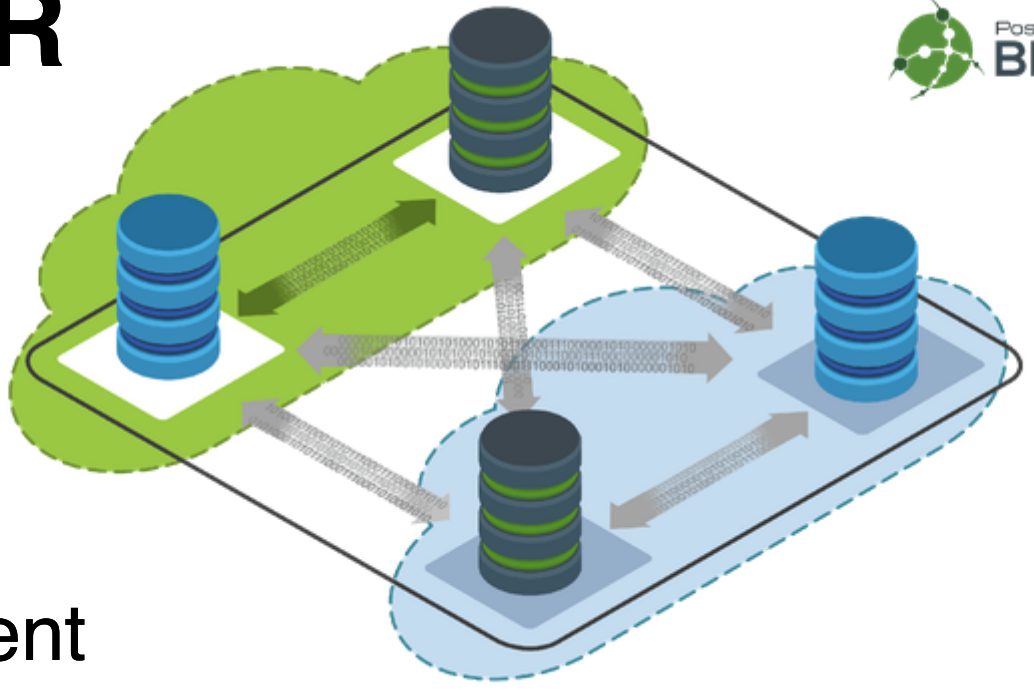
- Logical decoding
  - Extension API
  - Row changes
  - Committed changes
- Selective replication
- Cross-version
- Can write to all nodes

Physical vs Logical Replication



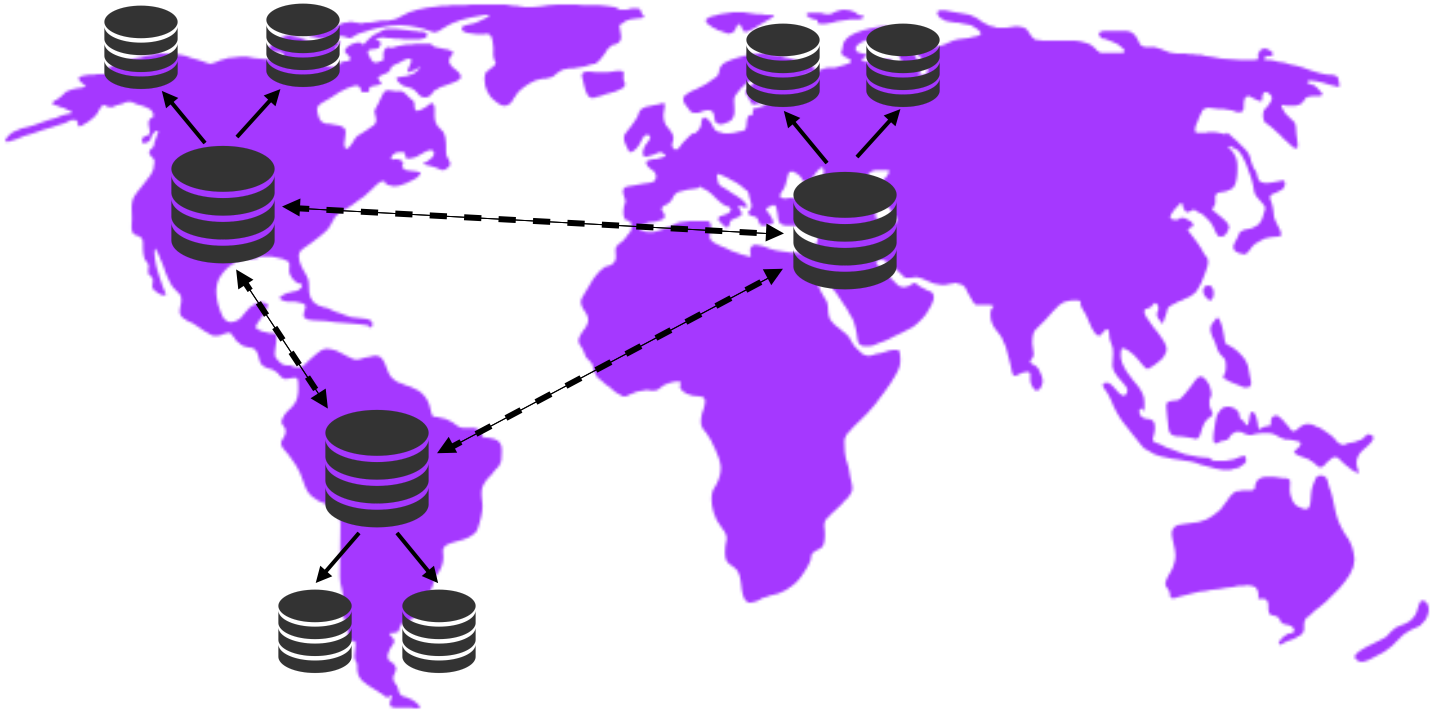


Postgres-BDR



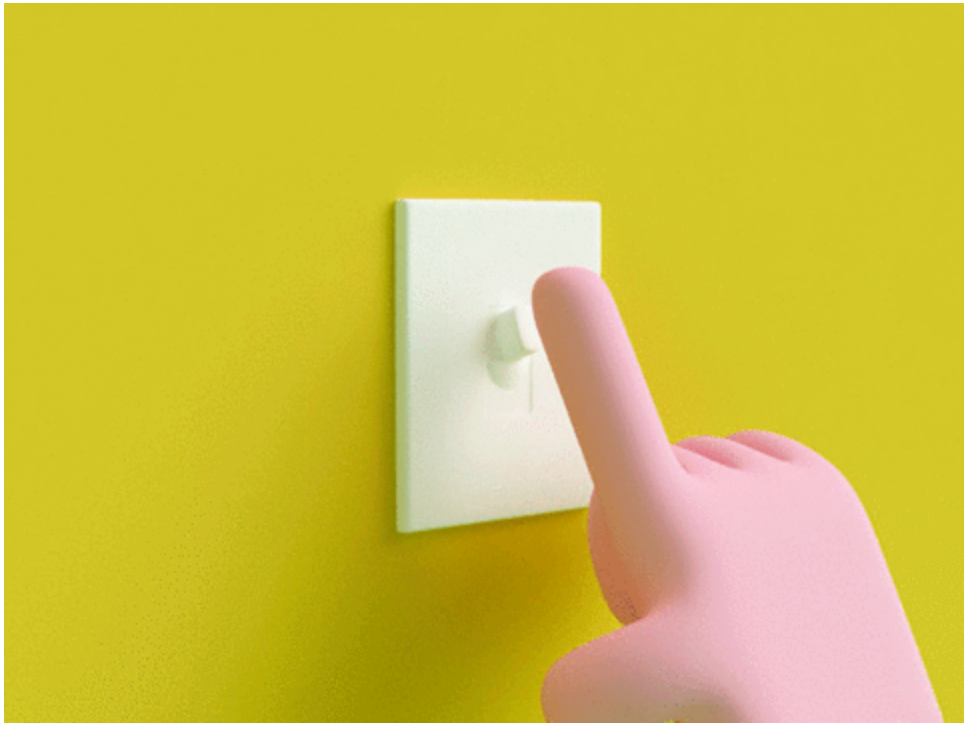
- Multi-master
- Asynchronous
- Eventually consistent
  - Does not prevent concurrent writes
  - Optimistic conflict detection (after commit)
  - Automatic conflict resolution
- Used for Logical Replication development in PG

Multi-Master Architectures - I



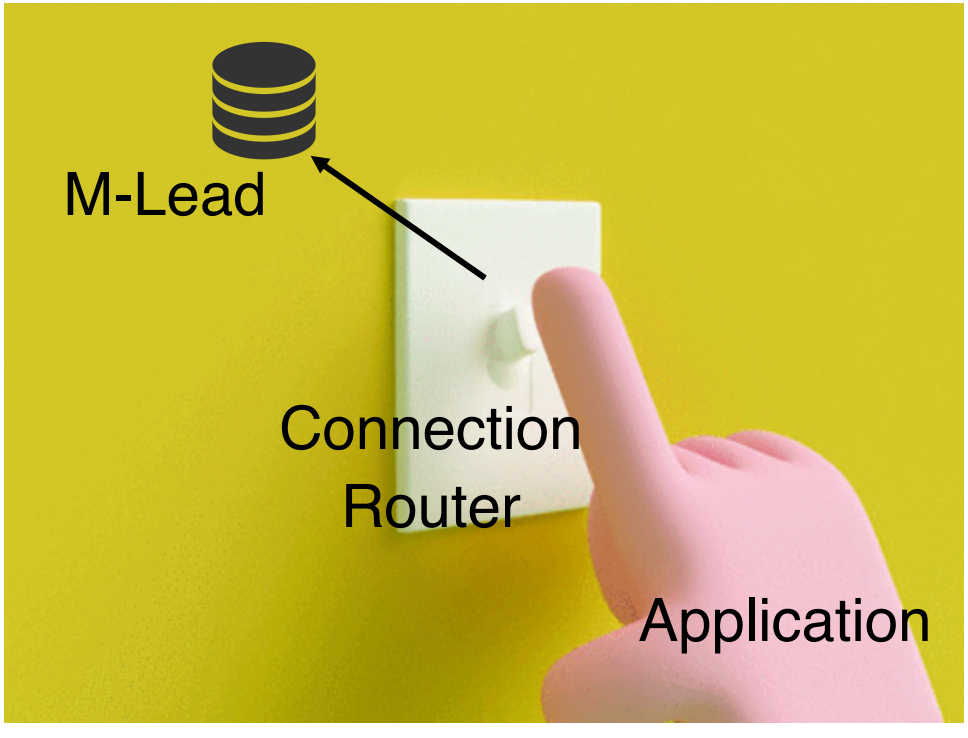
Geographically distributed cluster

Multi-Master Architectures - II



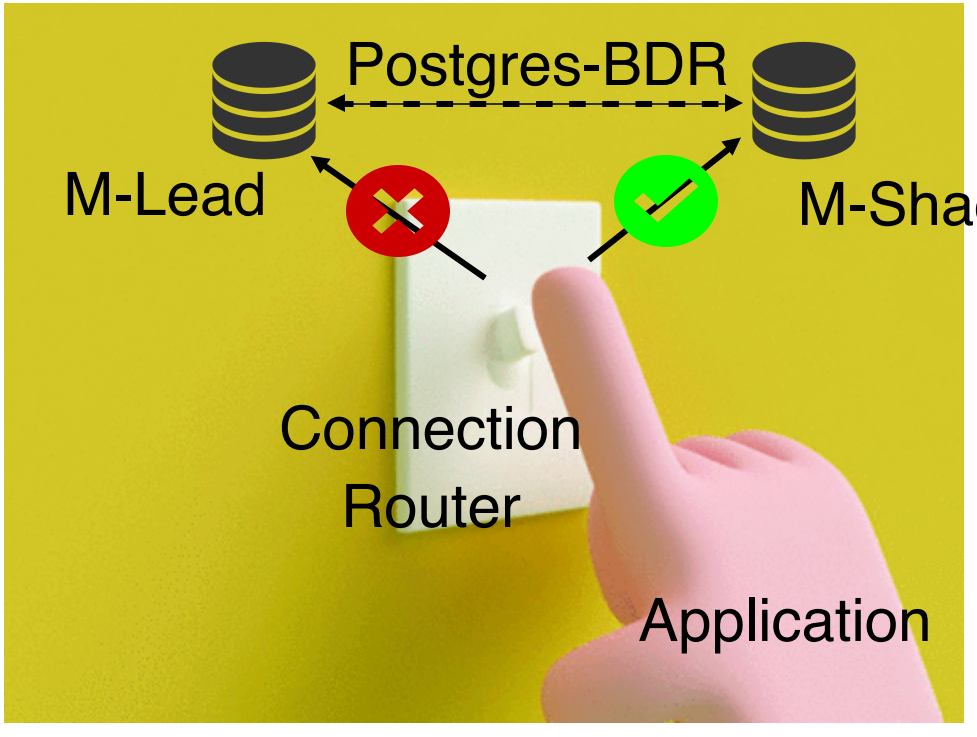
AlwaysOn Architecture

Multi-Master Architectures - II



AlwaysOn Architecture

Multi-Master Architectures - II

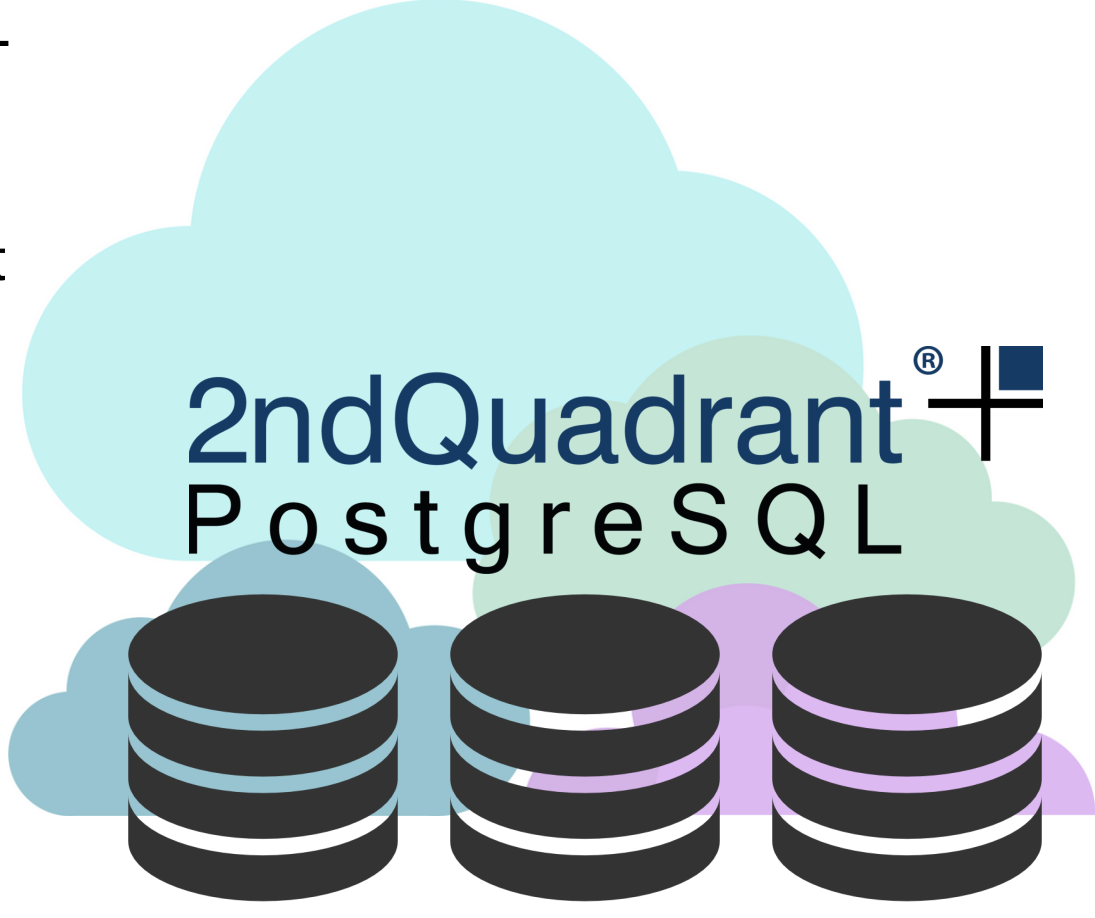


AlwaysOn Architecture

Global Database as a Service

Bundled up best PostgreSQL practices into our cloud service, with high-availability, and 24x7 support by the best PostgreSQL engineers!

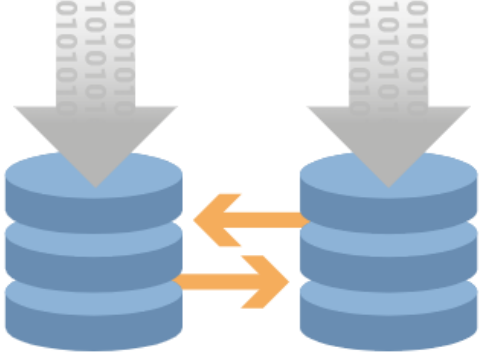
- We develop
- We host
- We manage
- We support
- **YOU USE!**



GDS Flavours

Single Master

Single master in region of your choice. Best option to start with.



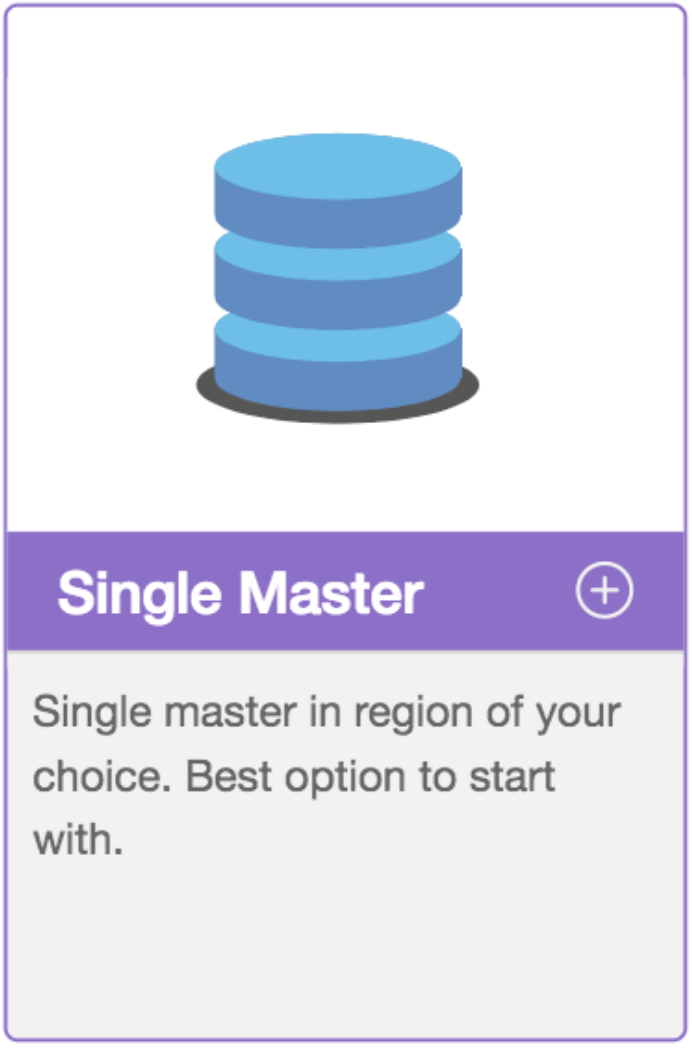
Multi Master

3rd generation of BDR achieves efficiency and accuracy, ensuring very high availability of all nodes in a geographically distributed cluster.

AlwaysOn

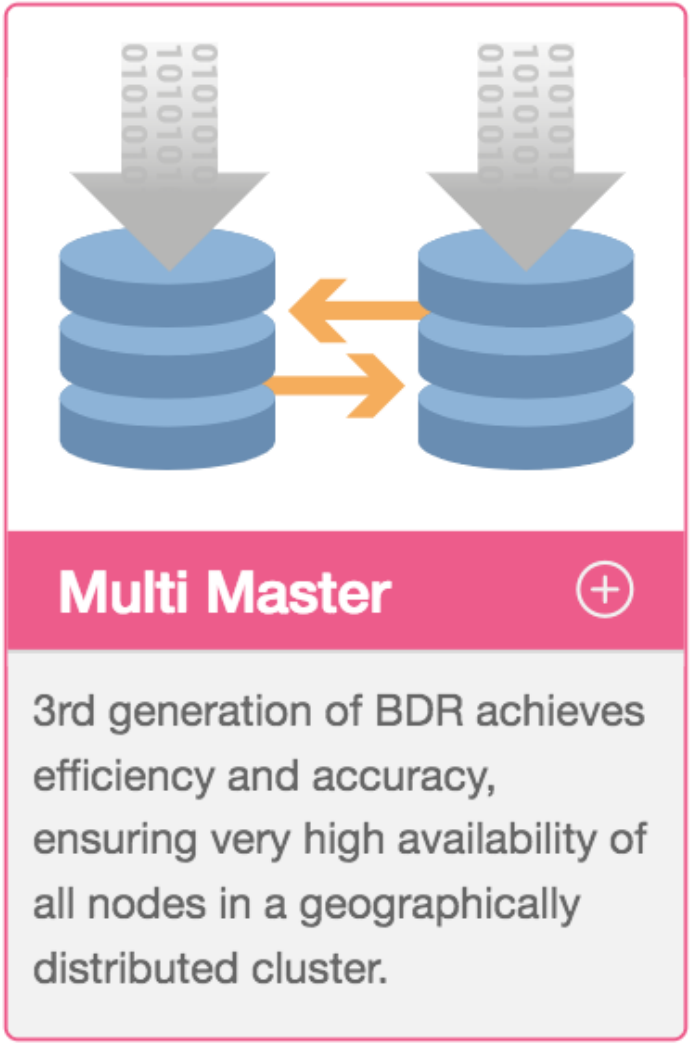
Shadow masters / standby nodes with auto failover allow a highly available environment catering to all possible issues impacting availability.

Single Master



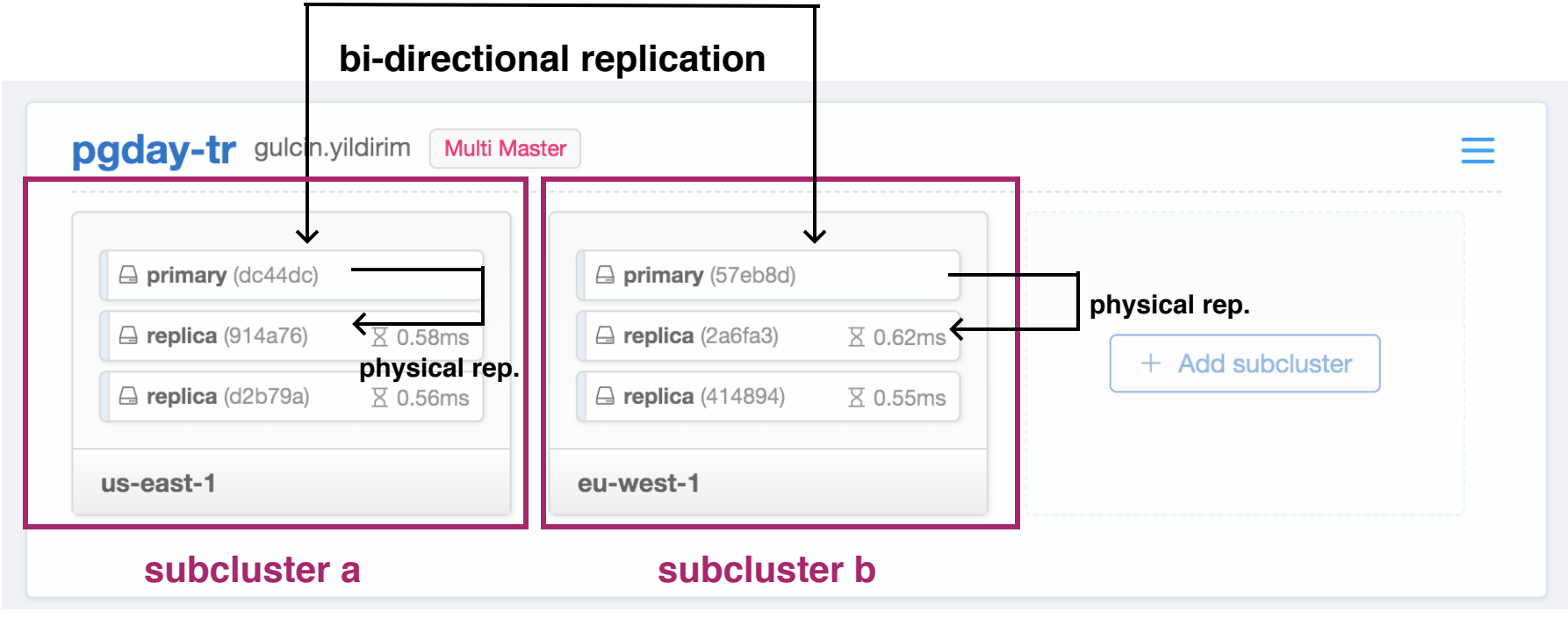
- Development
  - 1 primary database
  - No support
- Test
  - 1 primary database
  - 1 replica database
  - Gold Support
- Production
  - 1 primary database
  - 2 replica databases
  - Backups
  - Platinum Support

Multi Master



- Bi-Directional Replication with Postgres-BDR 3.0
- High availability
- Geographically distributed cluster
- Zero Downtime Upgrades
  - Different Postgres versions
- Flexibility
  - Different configuration
  - Different instance sizes (CPU, disk, memory etc)

Multi Master Architecture

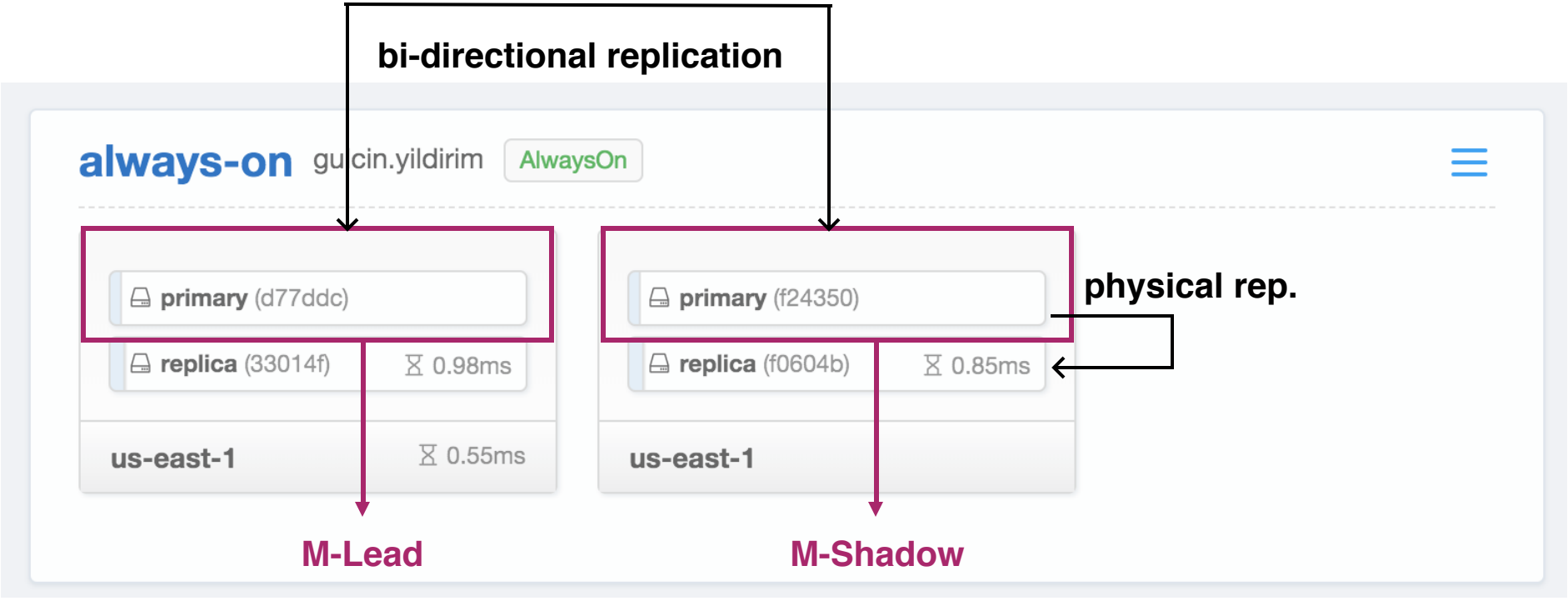


geo cluster

AlwaysOn

- Provides at least 4 Nines (99.99%) availability
- Deployed in single region with two different availability zones:
  - 2 BDR primary
  - 2 physical replica
  - Backup
- “Very High Availability” described in the BDR whitepaper.

AlwaysOn Architecture



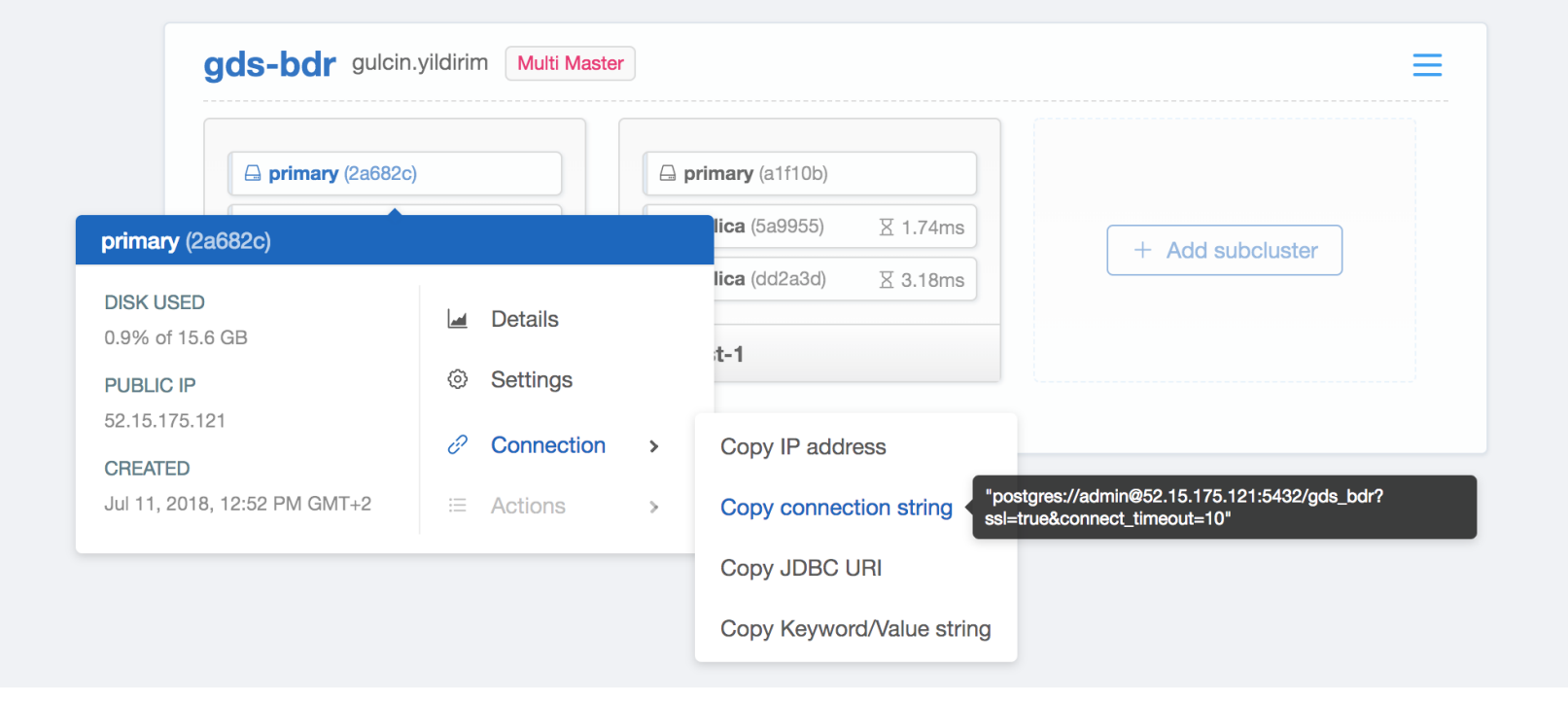
How to Connect to your DB?

How to Connect to your DB?

1. Set your database password first
2. Choose the instance you want to connect
3. Choose one of the provided connection string types
4. Connect with your db client

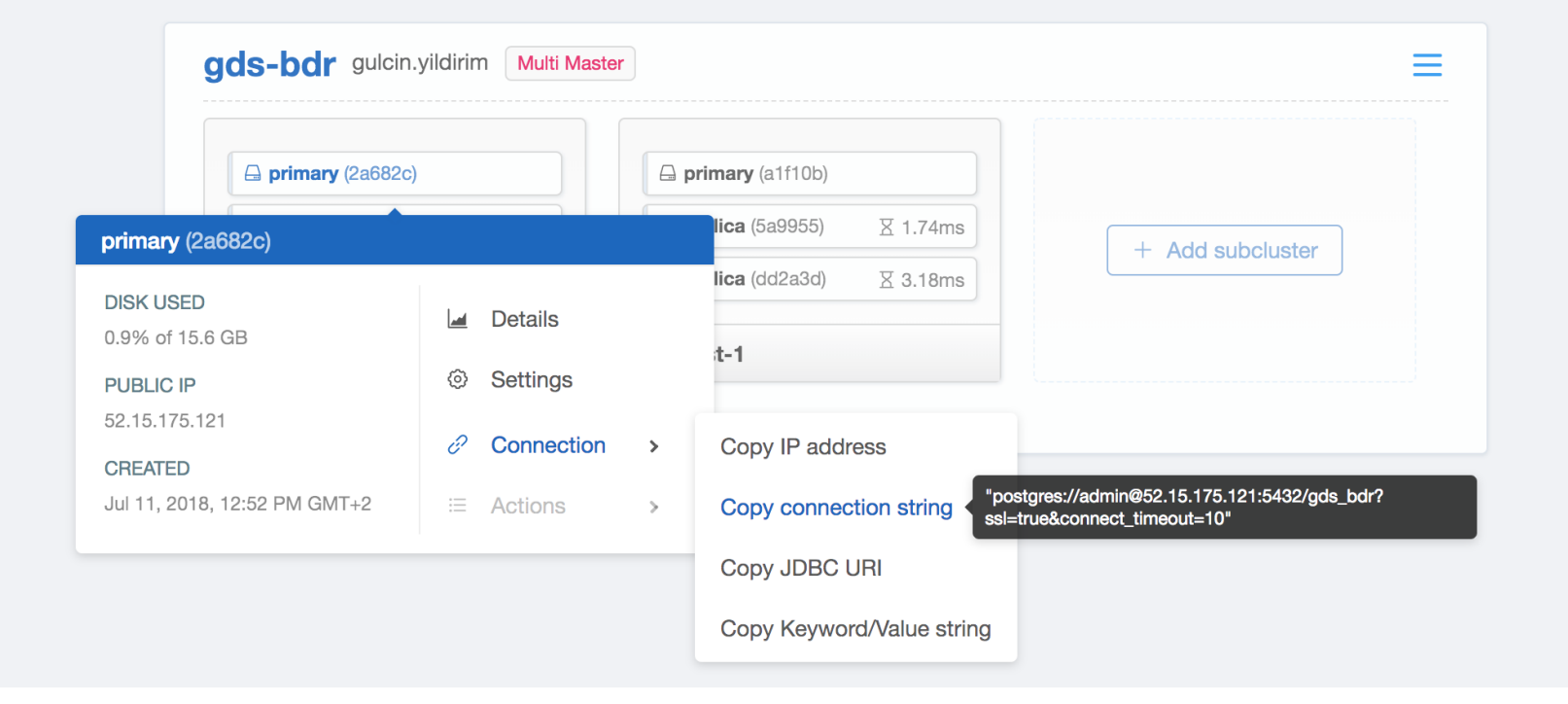
How to Connect to your DB?

1. Set your database password first
2. Choose the instance you want to connect
3. Choose one of the provided connection string types
4. Connect with your db client



How to Connect to your DB?

1. Set your database password first
2. Choose the instance you want to connect
3. Choose one of the provided connection string types
4. Connect with your db client

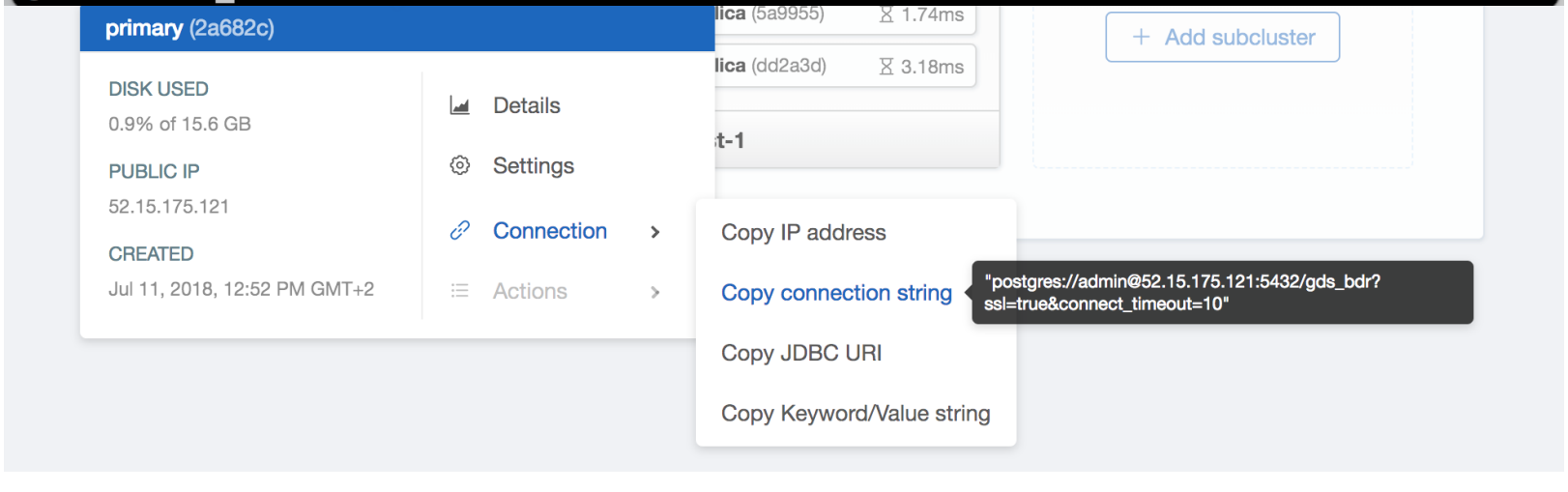


How to Connect to your DB?

1. Set your database password first
2. Choose the instance you want to connect

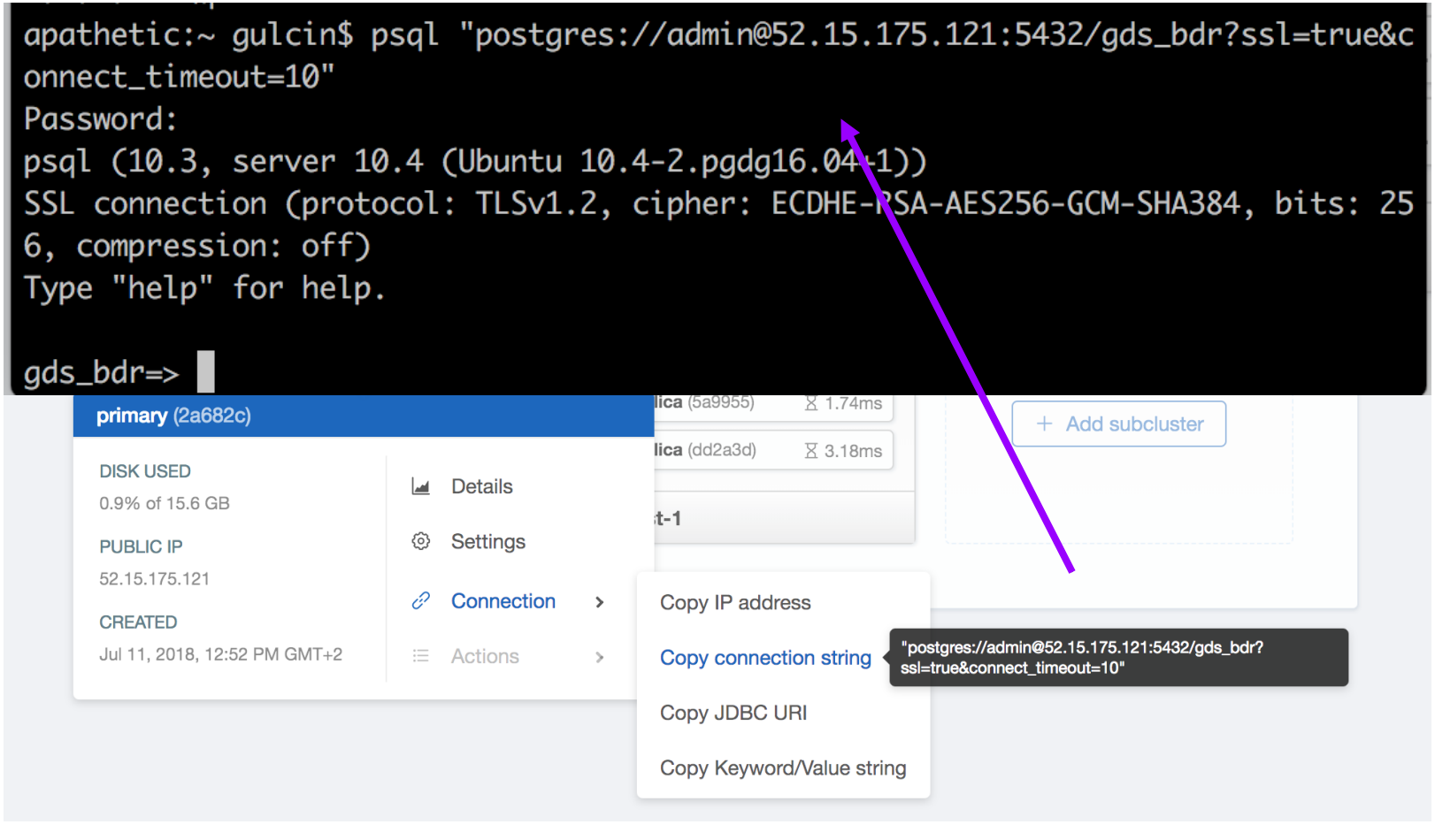
```
apathetic:~ gulcin$ psql "postgres://admin@52.15.175.121:5432/gds_bdr?ssl=true&c
onnect_timeout=10"
Password:
psql (10.3, server 10.4 (Ubuntu 10.4-2.pgdg16.04+1))
SSL connection (protocol: TLSv1.2, cipher: ECDHE-RSA-AES256-GCM-SHA384, bits: 25
6, compression: off)
Type "help" for help.

gds_bdr=>
```



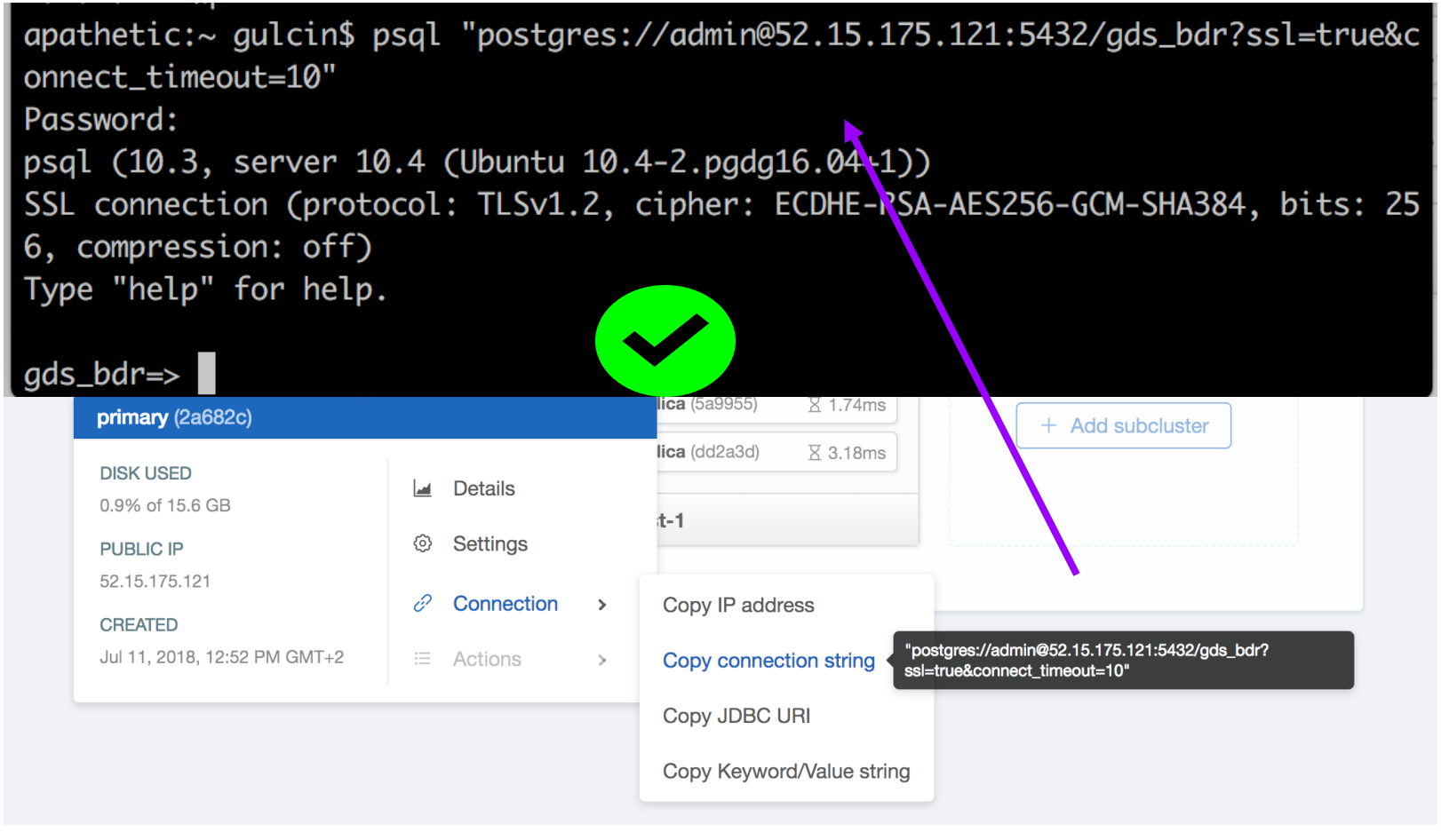
How to Connect to your DB?

1. Set your database password first
2. Choose the instance you want to connect

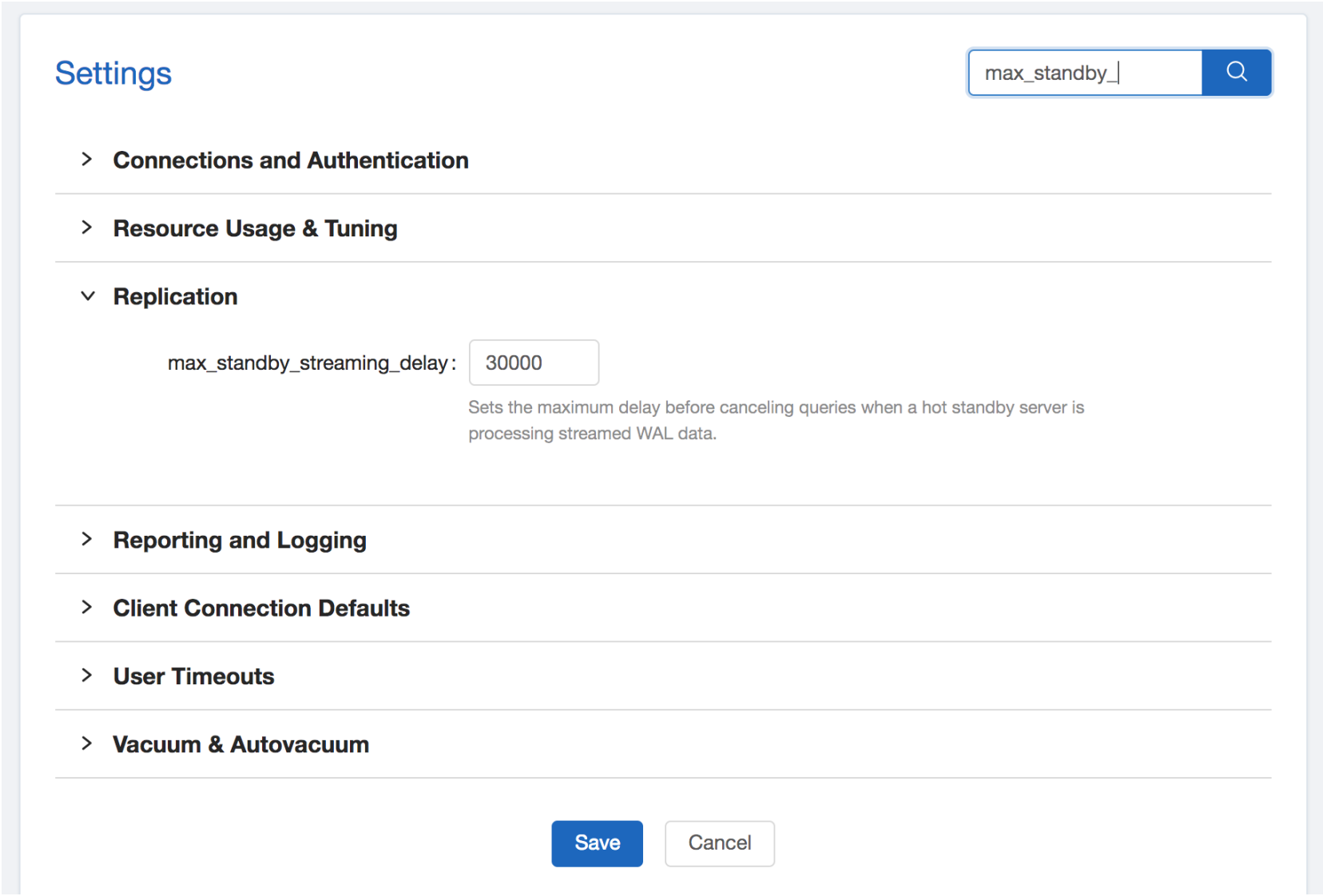


How to Connect to your DB?

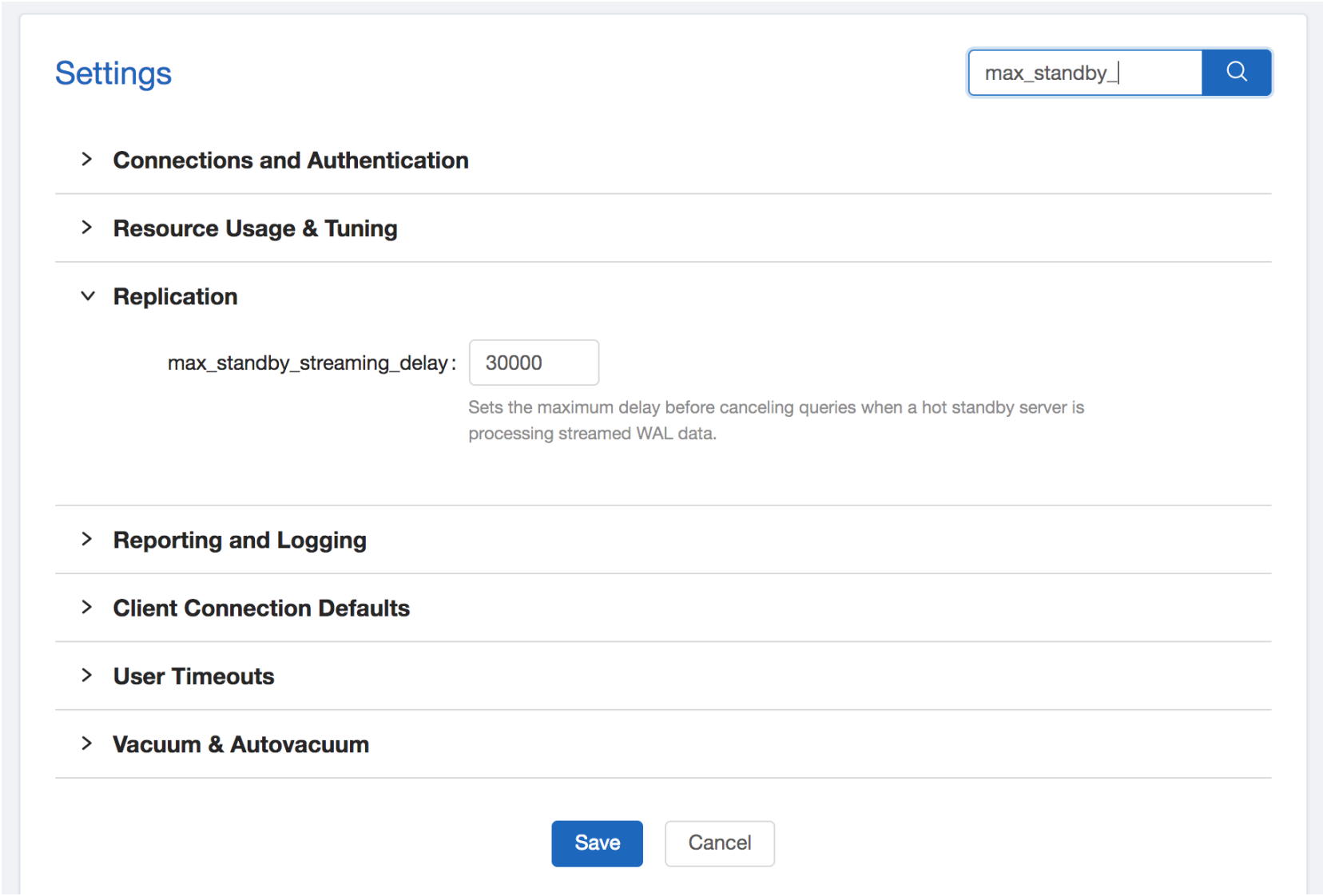
1. Set your database password first
2. Choose the instance you want to connect



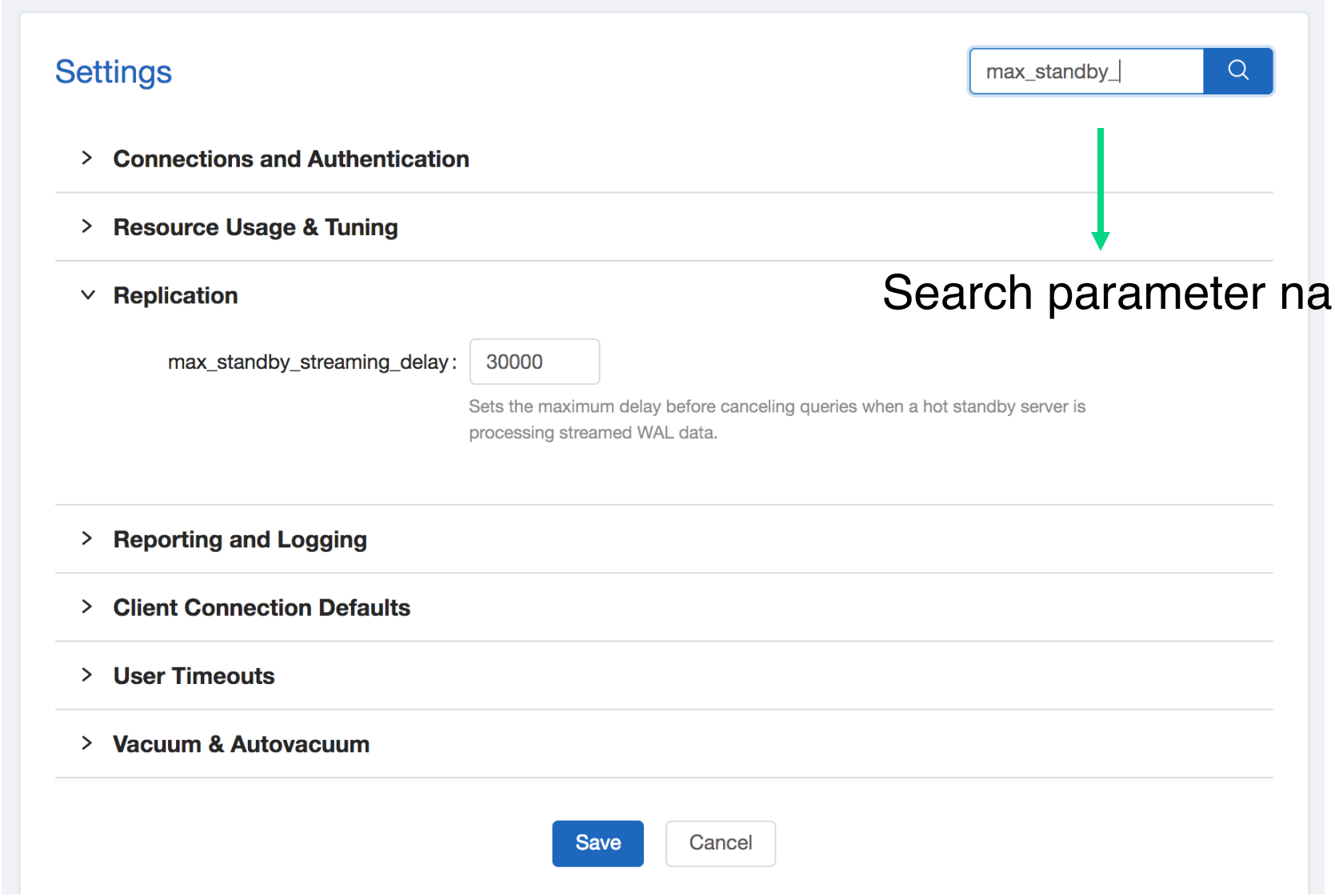
Database Configuration



Database Configuration



Database Configuration



Search parameter name

Monitoring Dashboard

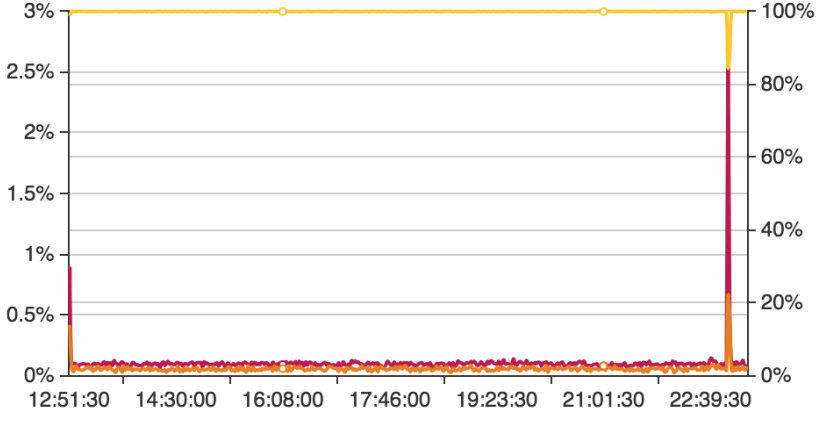
Clusters \ gds-bdr \ us-east-2 \ primary

Overview | Database | Queries | Replication | System

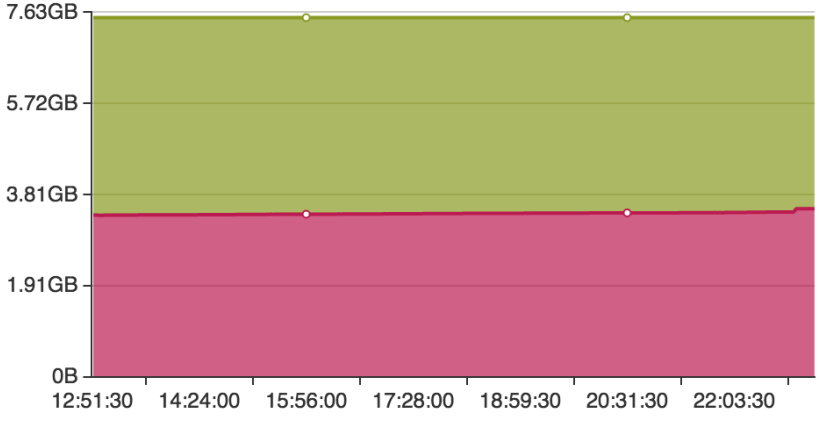
Last data less than a minute ago.

99.91%	4	0	...	0.1%	0.07%
Cache hit ratio	Active sessions	Deadlocks	Disk used	CPU user	CPU system

CPU

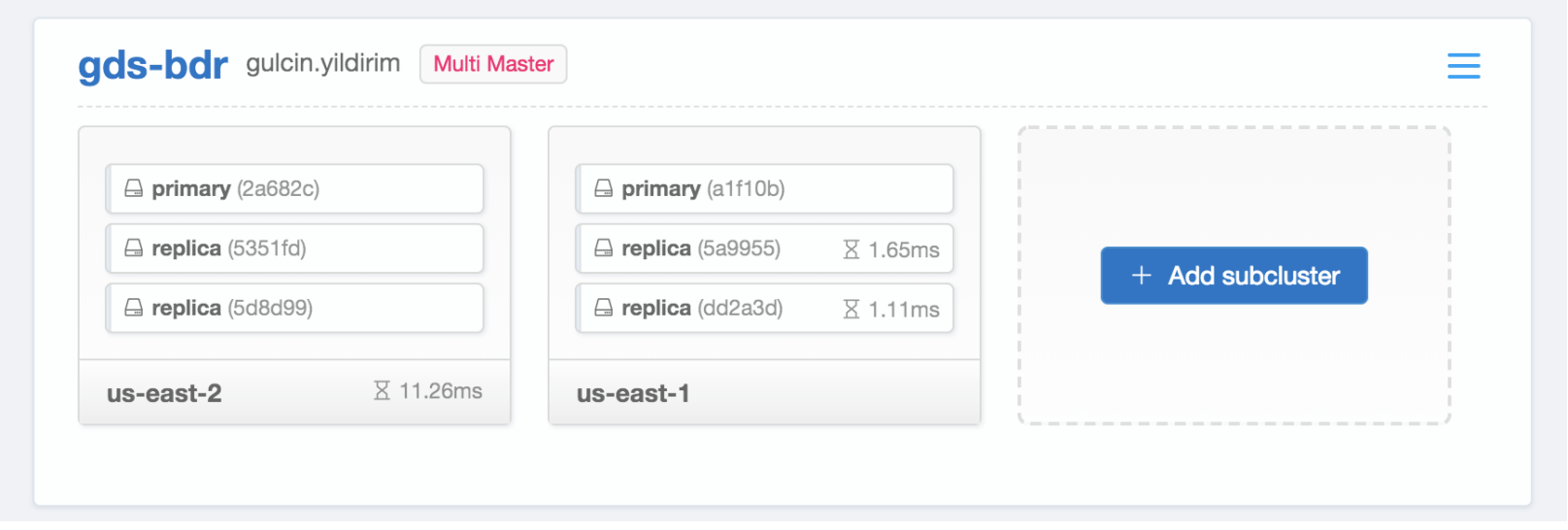


Memory

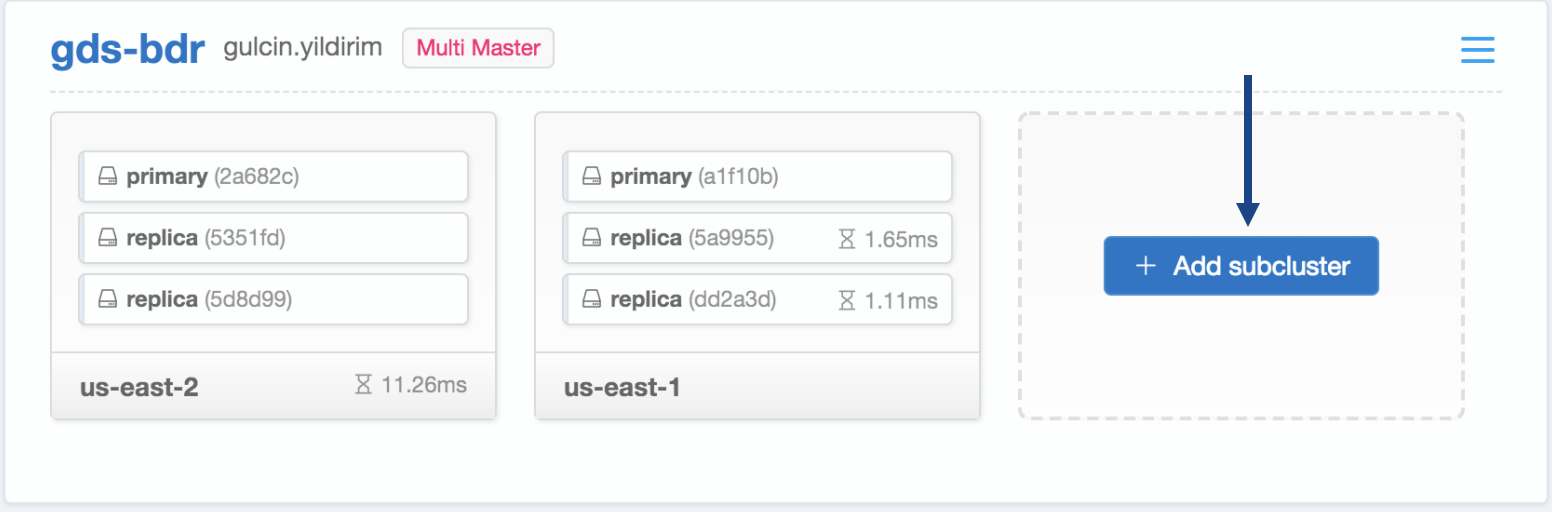


Dynamic Scaling

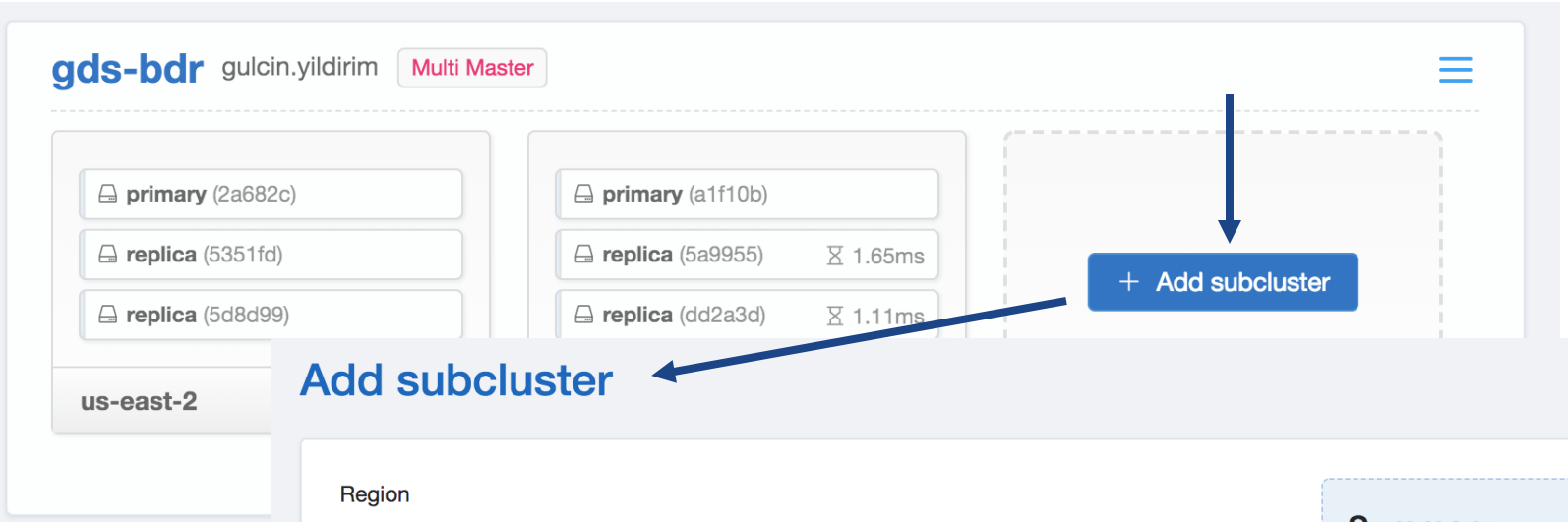
Dynamic Scaling



Dynamic Scaling



Dynamic Scaling



Add subcluster

Region

us-east-2 | us-east-1 | eu-west-1 | us-west-2

vCPU and RAM

vCPUs	Memory	Max IOPs
1	1 GB	3,000
2	8 GB	16,000
2	15 GB	3,000
4	16 GB	16,000
4	31 GB	6,000

Summary

PostgreSQL 10 running at **AWS** as **Multi Master** used for **Production**.

Region:

- **1** subcluster in **eu-west-1**

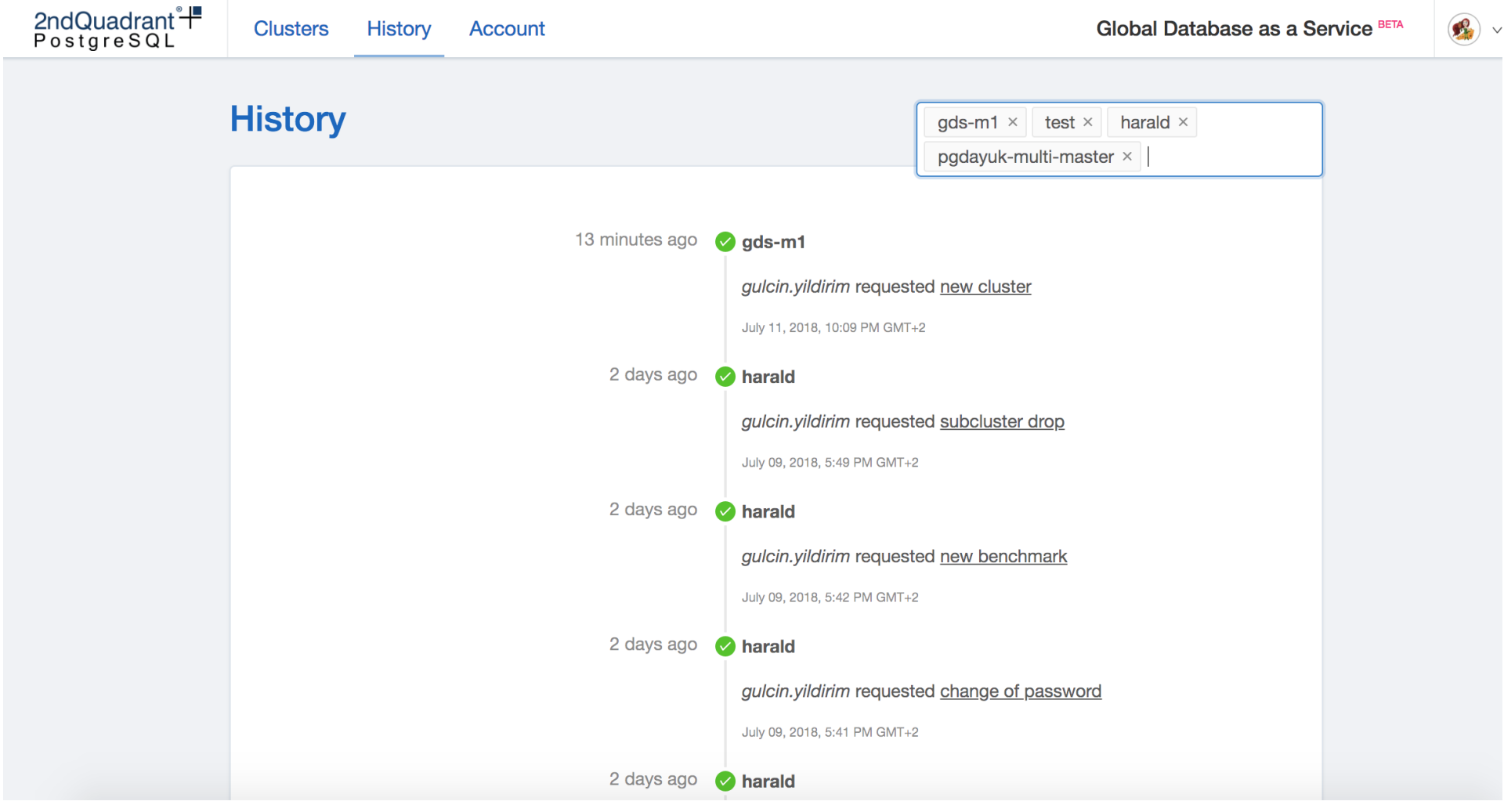
Each subcluster will have:

- 1 primary database
- 2 replica databases
- backup

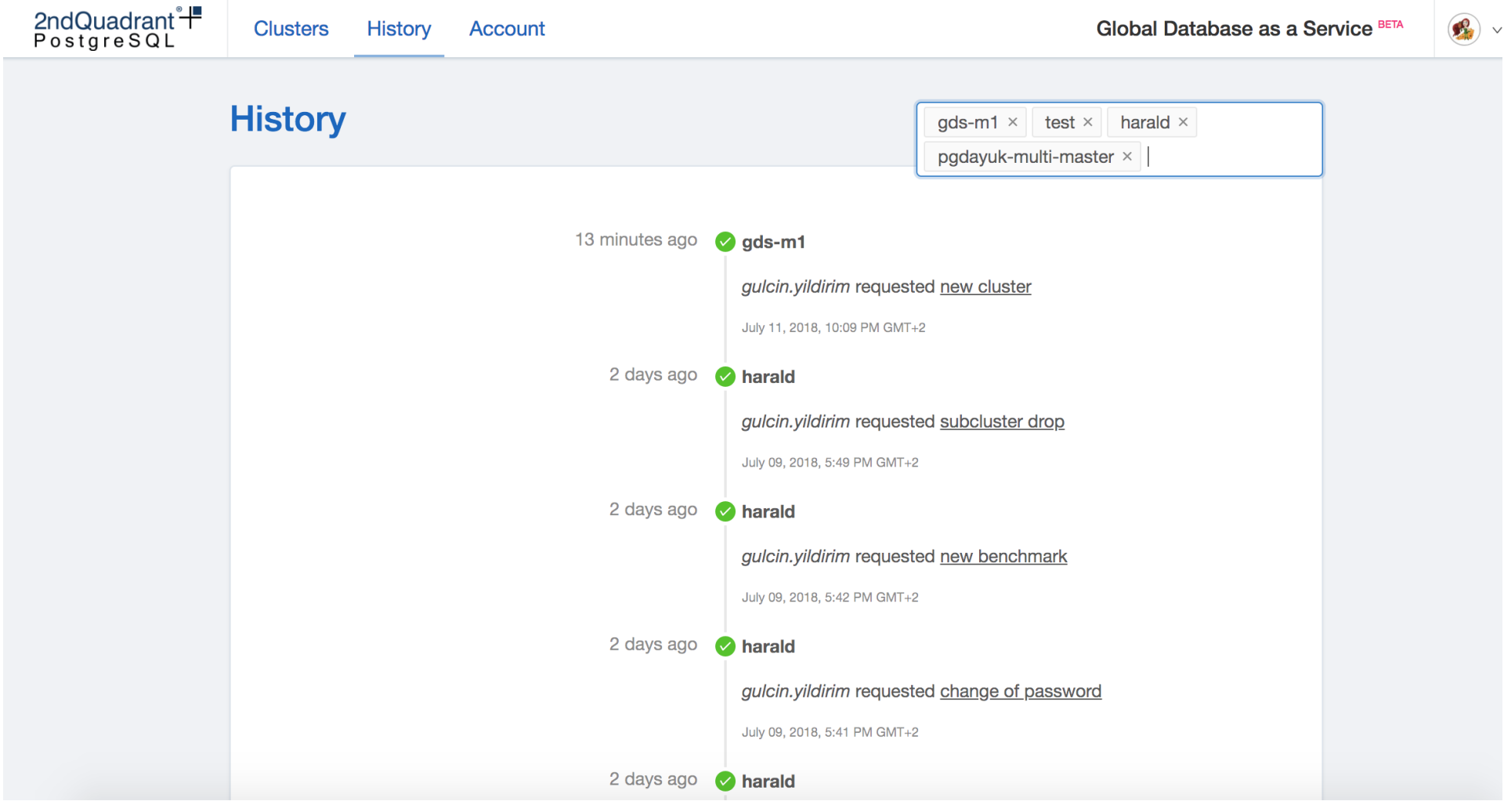
Instance specification:

- 2 vCPUs
- 8 GB RAM
- 100 IOPS
- 4 GB STORAGE

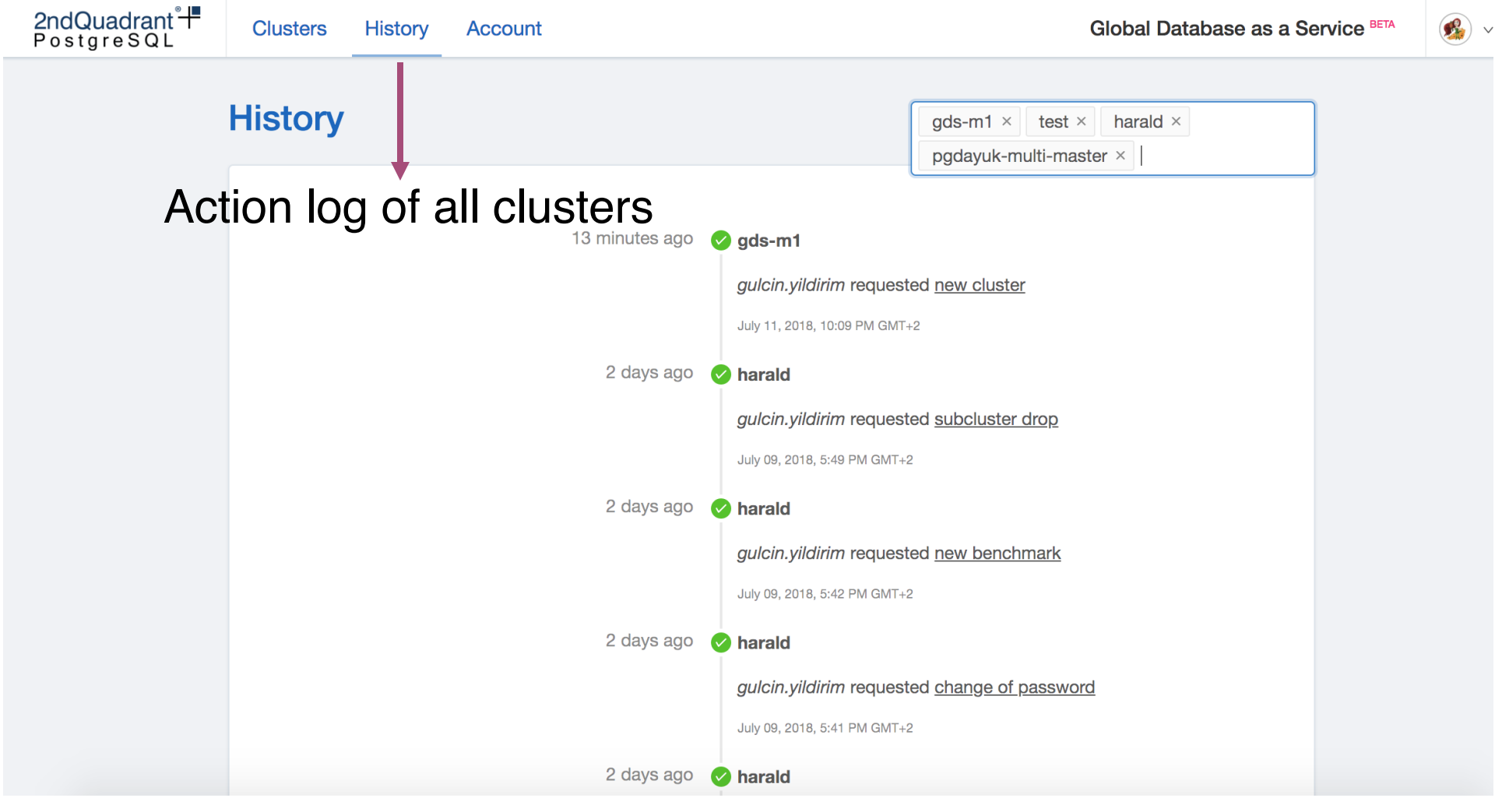
Logging


2ndQuadrant PostgreSQL
Clusters
History
Account
Global Database as a Service BETA
History
gds-m1 × test × harald ×
pgdayuk-multi-master ×
13 minutes ago
gds-m1
gulcin.yildirim requested new cluster
July 11, 2018, 10:09 PM GMT+2
2 days ago
harald
gulcin.yildirim requested subcluster drop
July 09, 2018, 5:49 PM GMT+2
2 days ago
harald
gulcin.yildirim requested new benchmark
July 09, 2018, 5:42 PM GMT+2
2 days ago
harald
gulcin.yildirim requested change of password
July 09, 2018, 5:41 PM GMT+2
2 days ago
harald

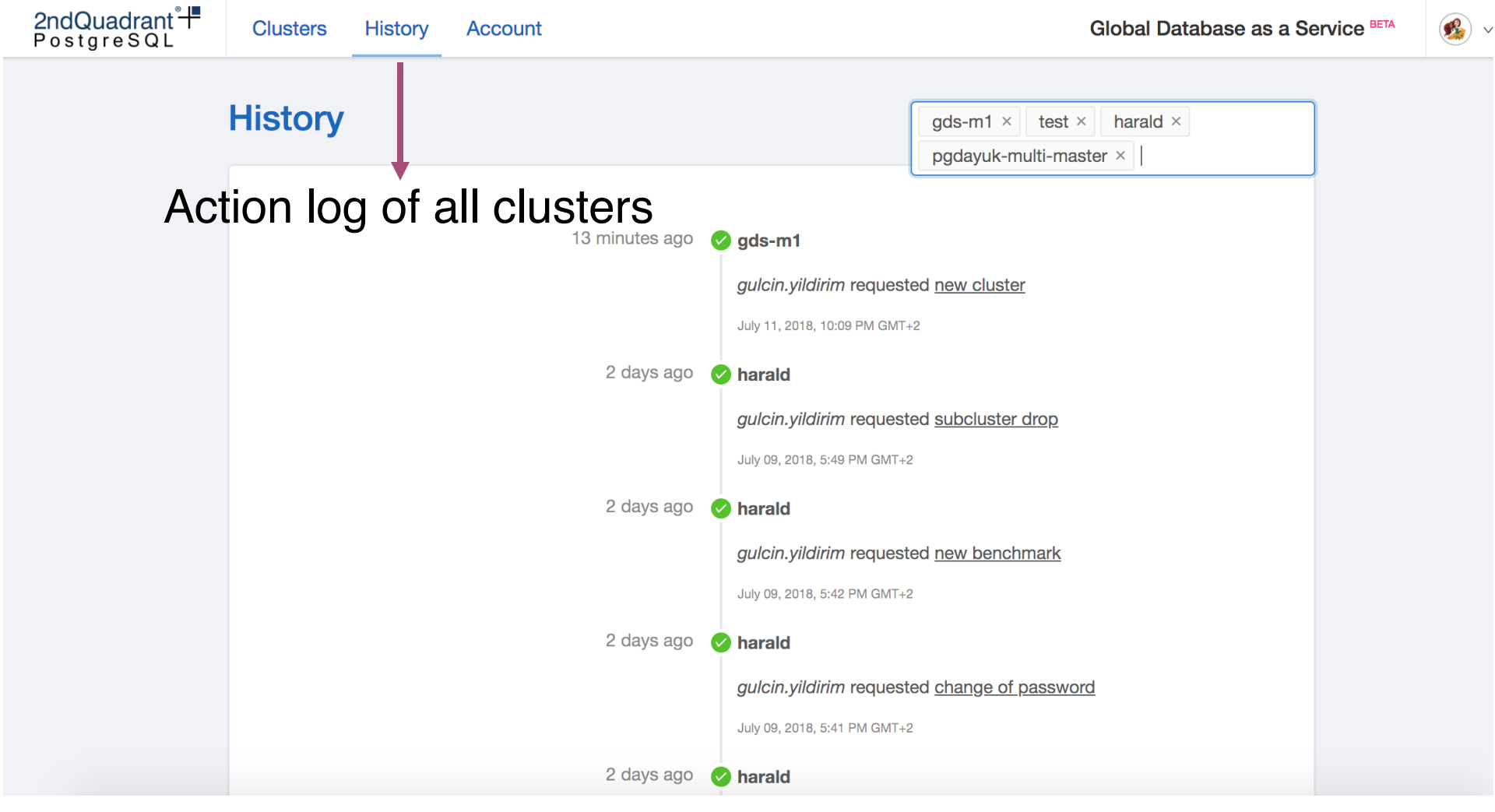
Logging


2ndQuadrant PostgreSQL
Clusters
History
Account
Global Database as a Service BETA
History
gds-m1 × test × harald ×
pgdayuk-multi-master ×
13 minutes ago
gds-m1
gulcin.yildirim requested new cluster
July 11, 2018, 10:09 PM GMT+2
2 days ago
harald
gulcin.yildirim requested subcluster drop
July 09, 2018, 5:49 PM GMT+2
2 days ago
harald
gulcin.yildirim requested new benchmark
July 09, 2018, 5:42 PM GMT+2
2 days ago
harald
gulcin.yildirim requested change of password
July 09, 2018, 5:41 PM GMT+2
2 days ago
harald

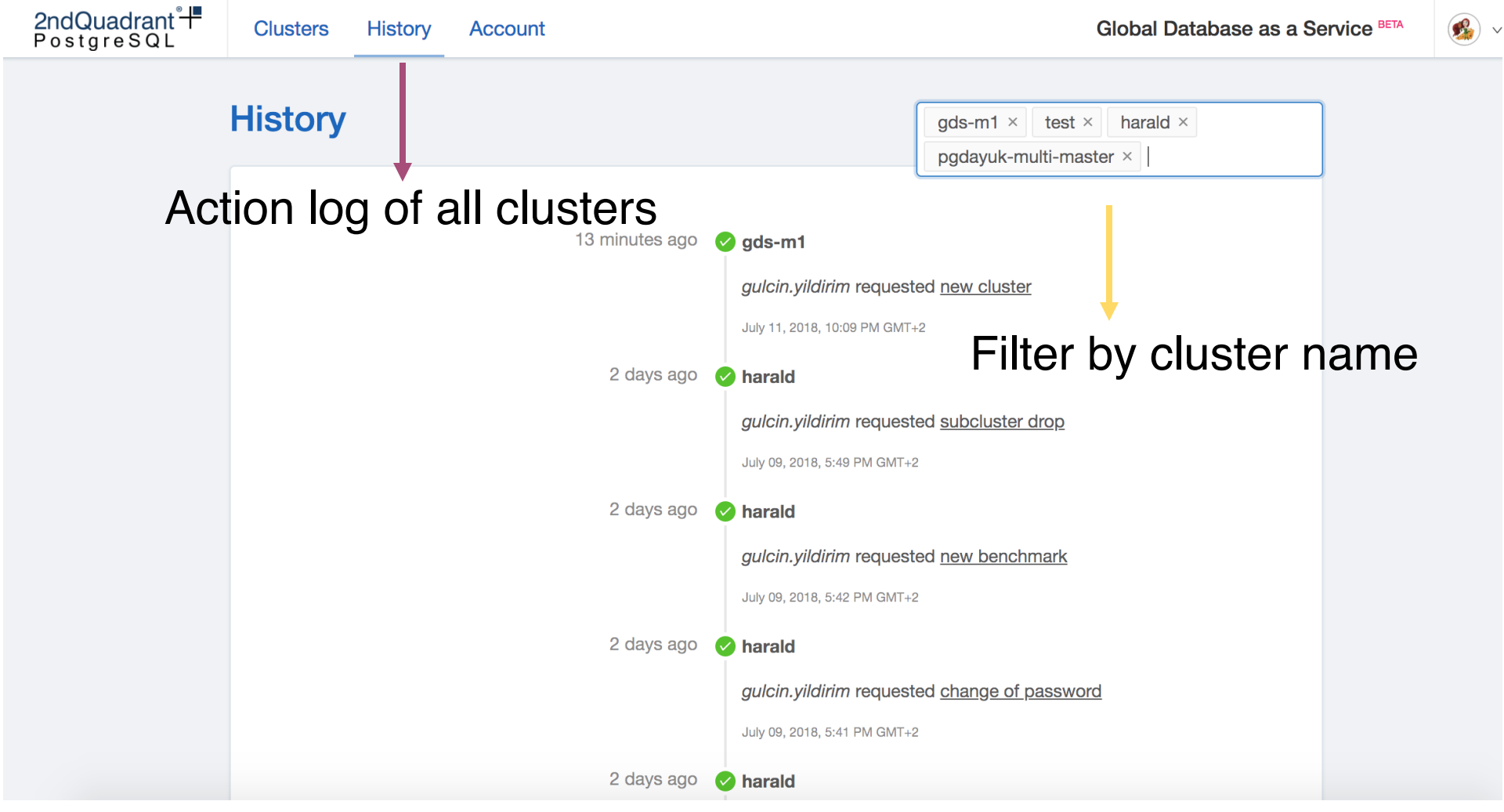
Logging


2ndQuadrant PostgreSQL
Clusters
History
Account
Global Database as a Service BETA
History
gds-m1 × test × harald ×
pgdayuk-multi-master ×
Action log of all clusters
13 minutes ago
gds-m1
gulcin.yildirim requested new cluster
July 11, 2018, 10:09 PM GMT+2
2 days ago
harald
gulcin.yildirim requested subcluster drop
July 09, 2018, 5:49 PM GMT+2
2 days ago
harald
gulcin.yildirim requested new benchmark
July 09, 2018, 5:42 PM GMT+2
2 days ago
harald
gulcin.yildirim requested change of password
July 09, 2018, 5:41 PM GMT+2
2 days ago
harald

Logging

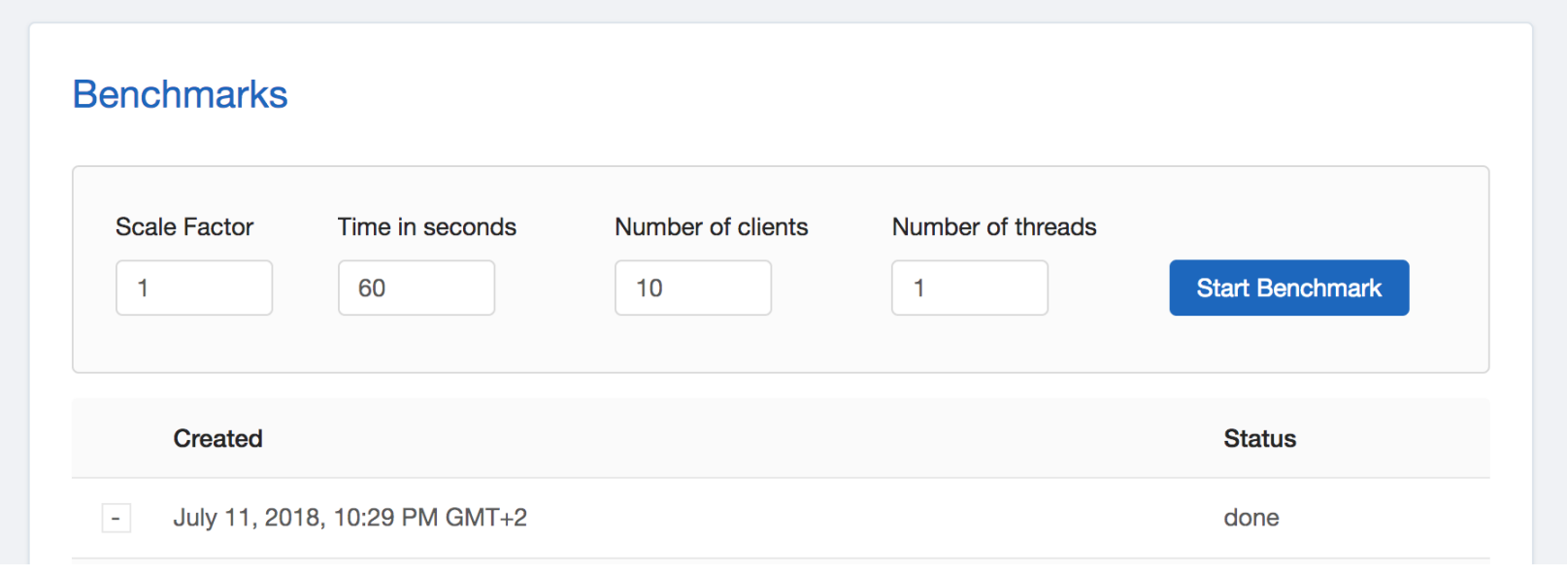

2ndQuadrant PostgreSQL
Clusters
History
Account
Global Database as a Service BETA
History
gds-m1 × test × harald ×
pgdayuk-multi-master ×
Action log of all clusters
13 minutes ago
gds-m1
gulcin.yildirim requested new cluster
July 11, 2018, 10:09 PM GMT+2
2 days ago
harald
gulcin.yildirim requested subcluster drop
July 09, 2018, 5:49 PM GMT+2
2 days ago
harald
gulcin.yildirim requested new benchmark
July 09, 2018, 5:42 PM GMT+2
2 days ago
harald
gulcin.yildirim requested change of password
July 09, 2018, 5:41 PM GMT+2
2 days ago
harald

Logging


2ndQuadrant PostgreSQL
Clusters
History
Account
Global Database as a Service BETA
History
Action log of all clusters
gds-m1 × test × harald ×
pgdayuk-multi-master ×
Filter by cluster name
13 minutes ago
gds-m1
gulcin.yildirim requested new cluster
July 11, 2018, 10:09 PM GMT+2
2 days ago
harald
gulcin.yildirim requested subcluster drop
July 09, 2018, 5:49 PM GMT+2
2 days ago
harald
gulcin.yildirim requested new benchmark
July 09, 2018, 5:42 PM GMT+2
2 days ago
harald
gulcin.yildirim requested change of password
July 09, 2018, 5:41 PM GMT+2
2 days ago
harald

Benchmarks

Benchmarks


Benchmarks
Scale Factor
Time in seconds
Number of clients
Number of threads
1
60
10
1
Start Benchmark
Created
Status
-
July 11, 2018, 10:29 PM GMT+2
done

Benchmarks

Benchmarks

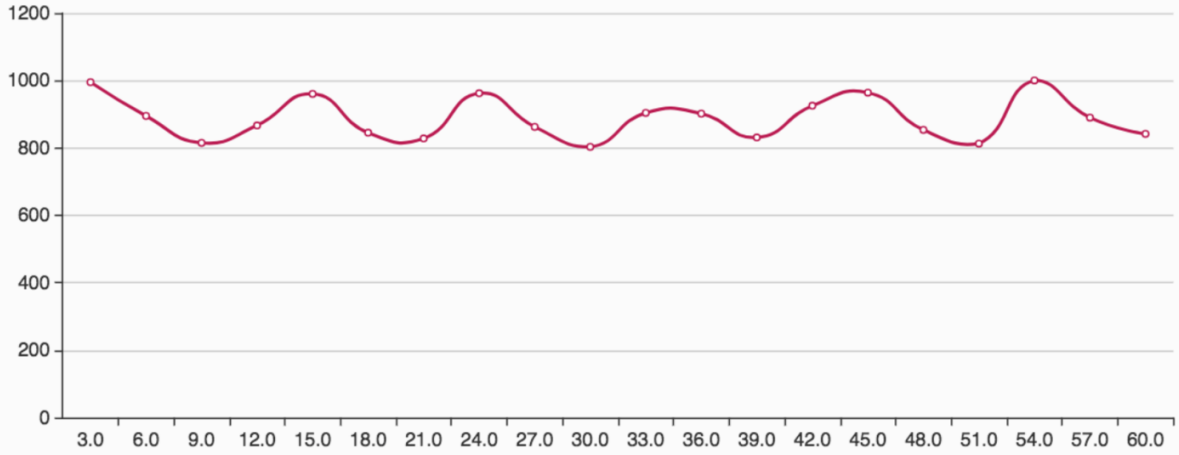
Scale Factor: 1

Time in seco: 60

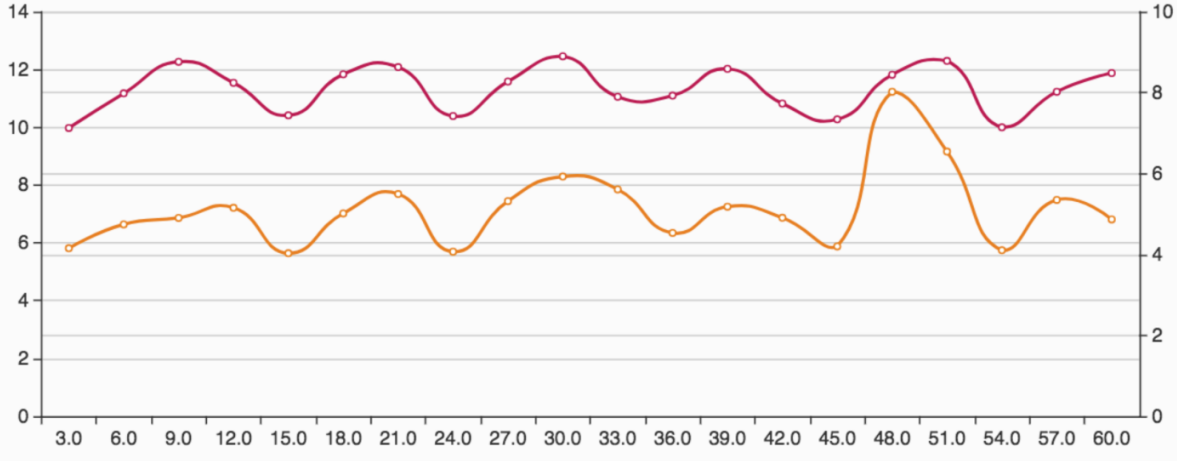
Created
- July 11, 2018, 10:29 PM GM

Results

Transactions per second 887.827543



Latency 11.259 ms



Benchmarks

Benchmarks

TPS

Scale Factor | Time in seco

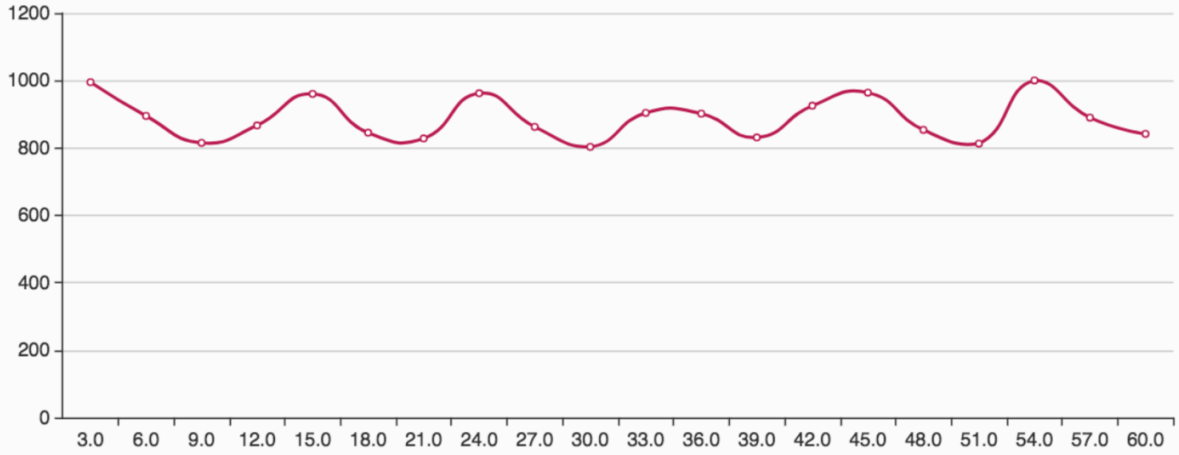
1 | 60

Created

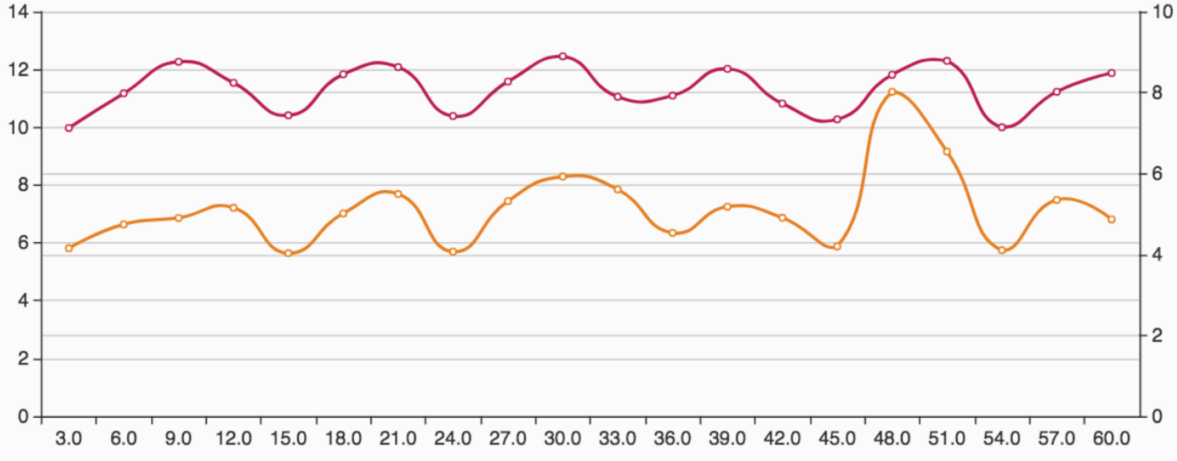
- July 11, 2018, 10:29 PM GM

Results

Transactions per second 887.827543



Latency 11.259 ms



Benchmarks

Benchmarks

TPS

Scale Factor | Time in seco

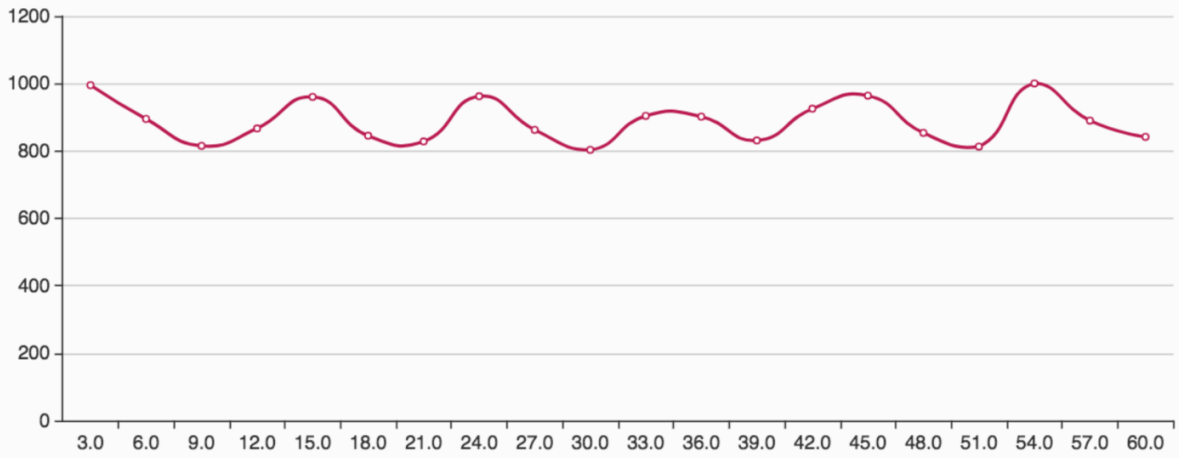
1 | 60

Created

- July 11, 2018, 10:29 PM GM

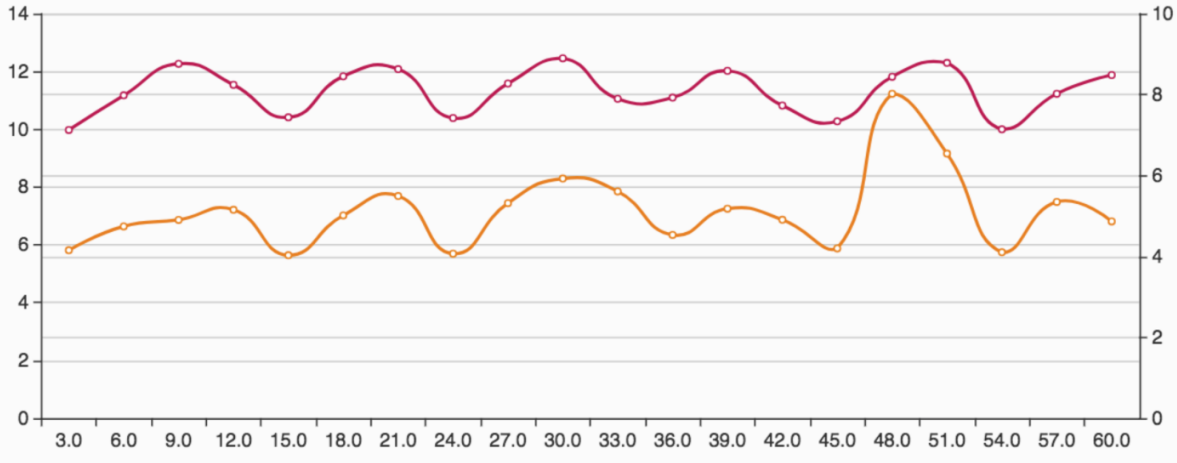
Results

Transactions per second 887.827543



Latency

Latency 11.259 ms



Security

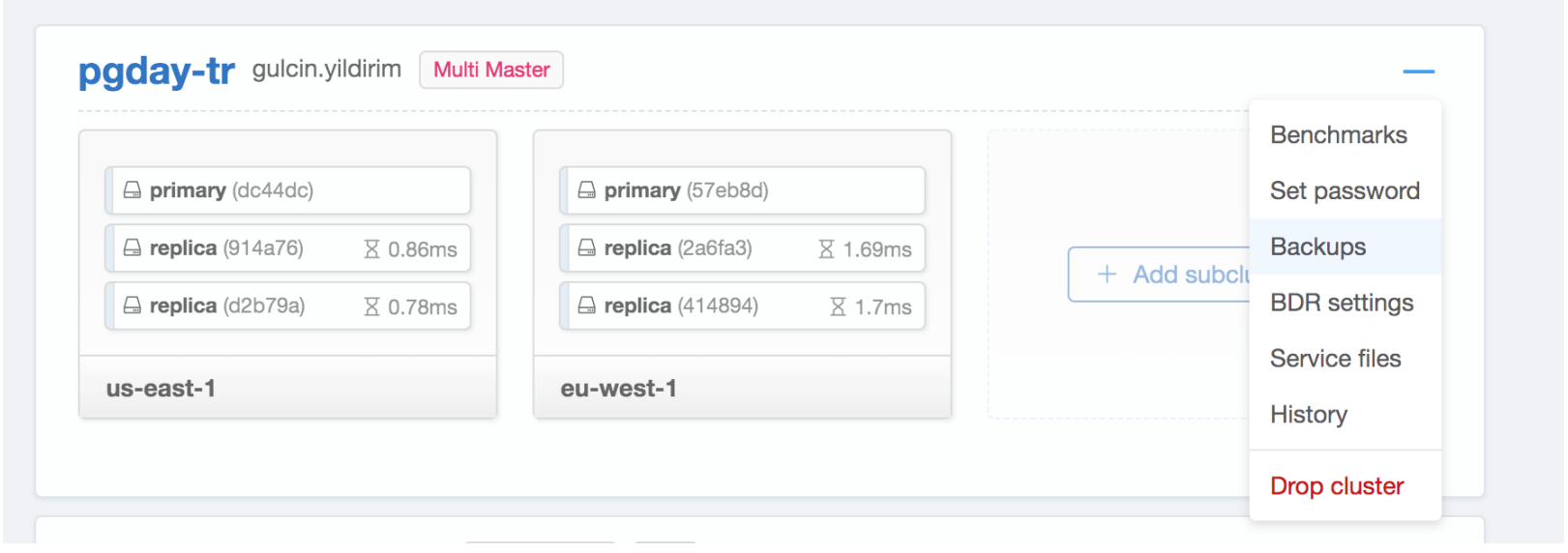
- Authenticate with 2ndQuadrant SSO
- SSL connections only

```
"postgres://admin@52.15.175.121:5432/gds_bdr?
ssl=true&connect_timeout=10"
```

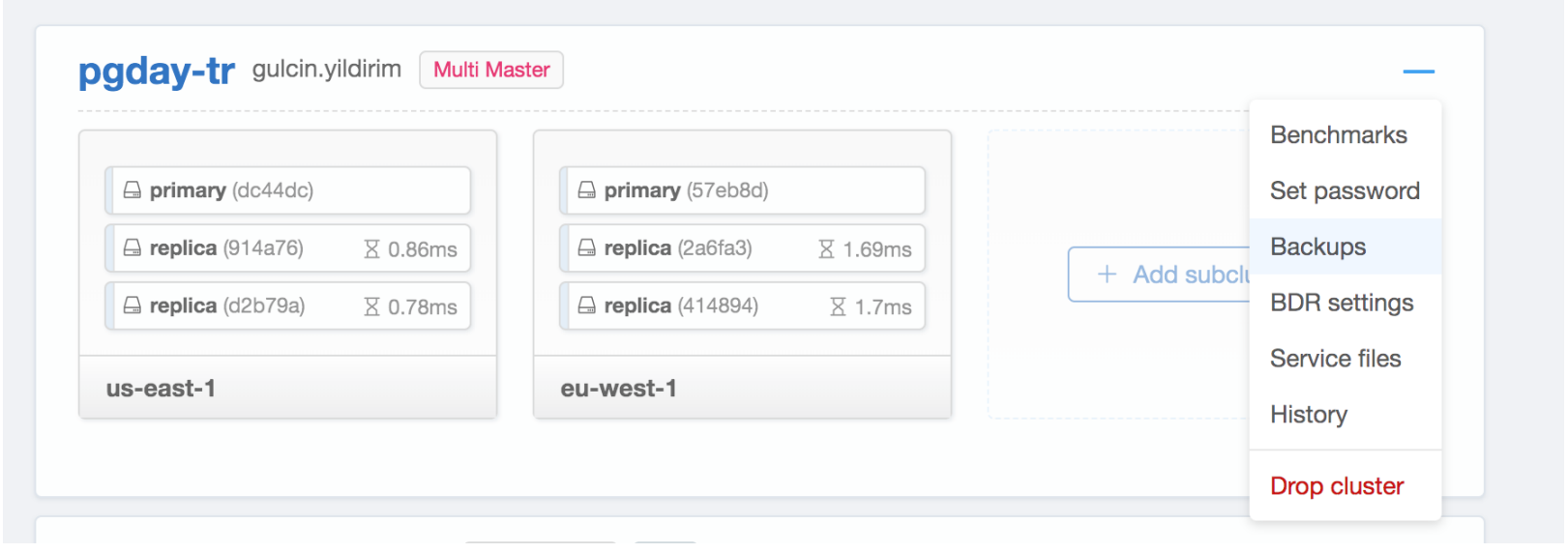
- Encrypted Data at Rest
- No SUPERUSER
- VPC Peering
- IP Whitelisting

Automated Backups

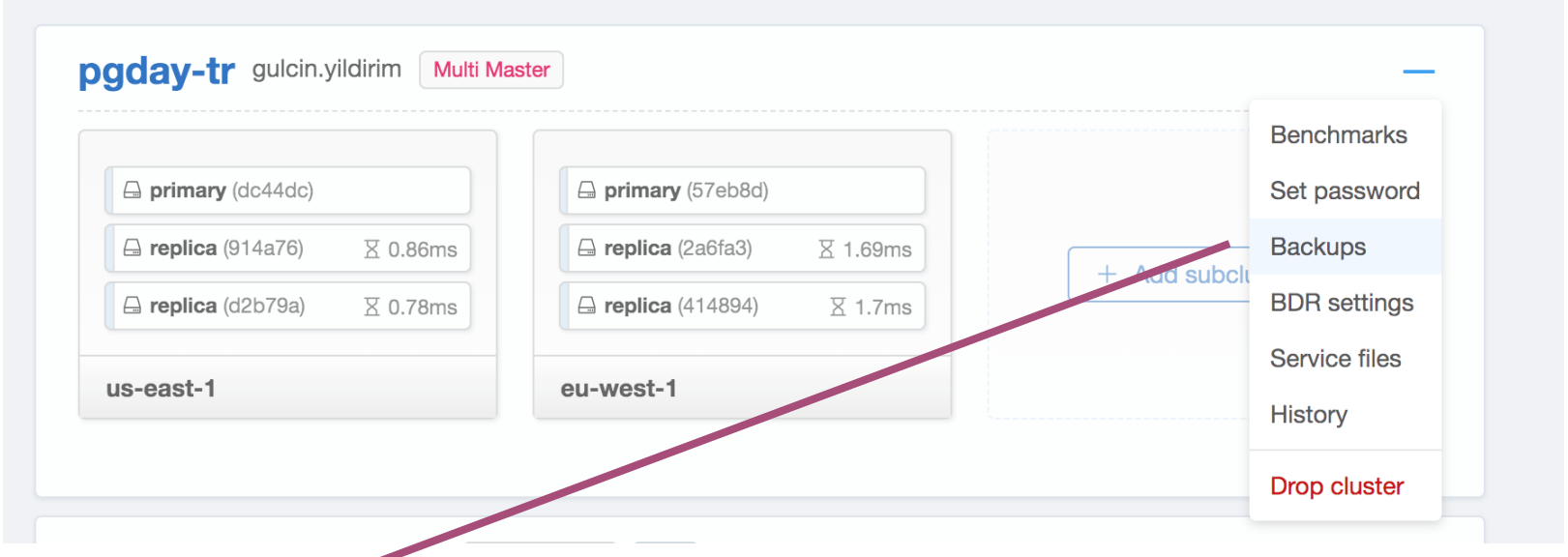
Automated Backups

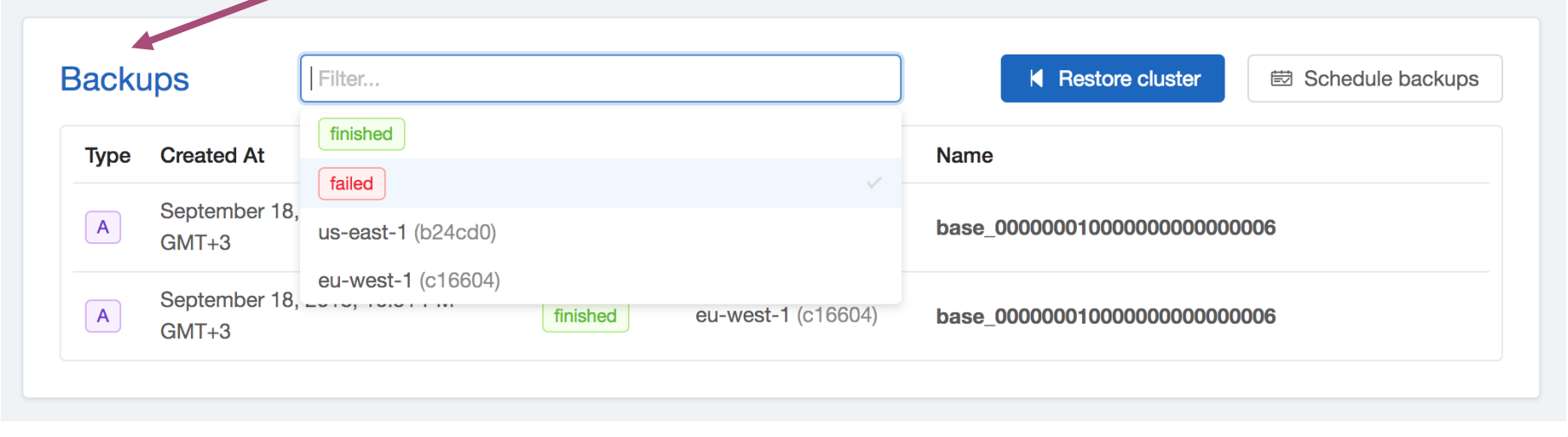


Automated Backups

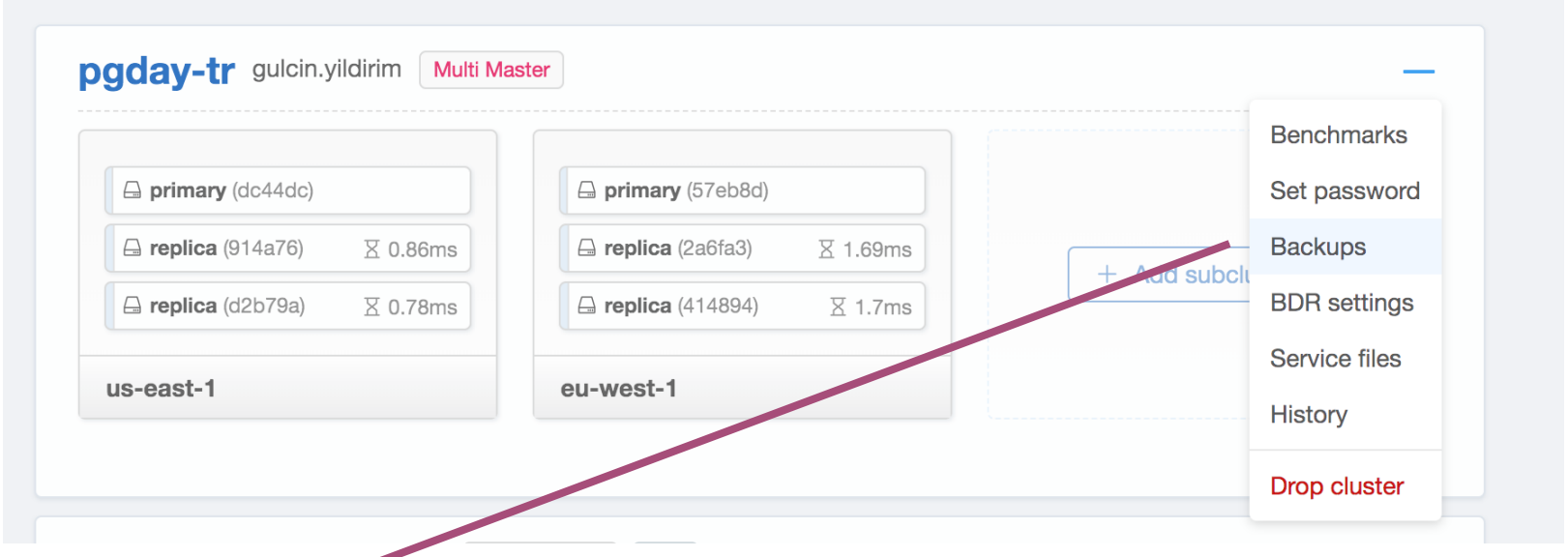


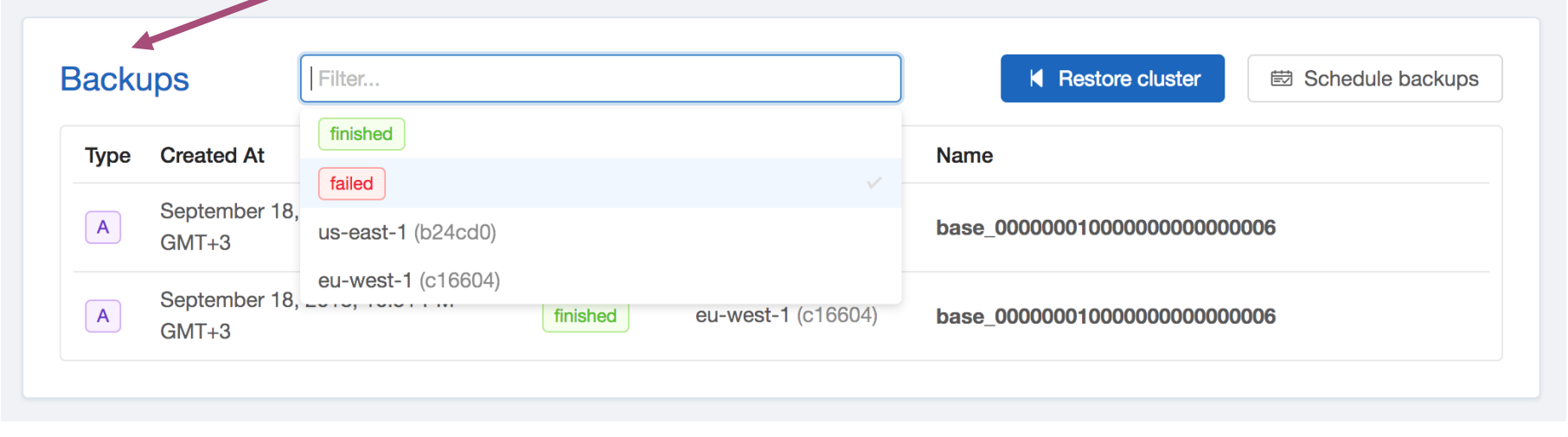
Automated Backups



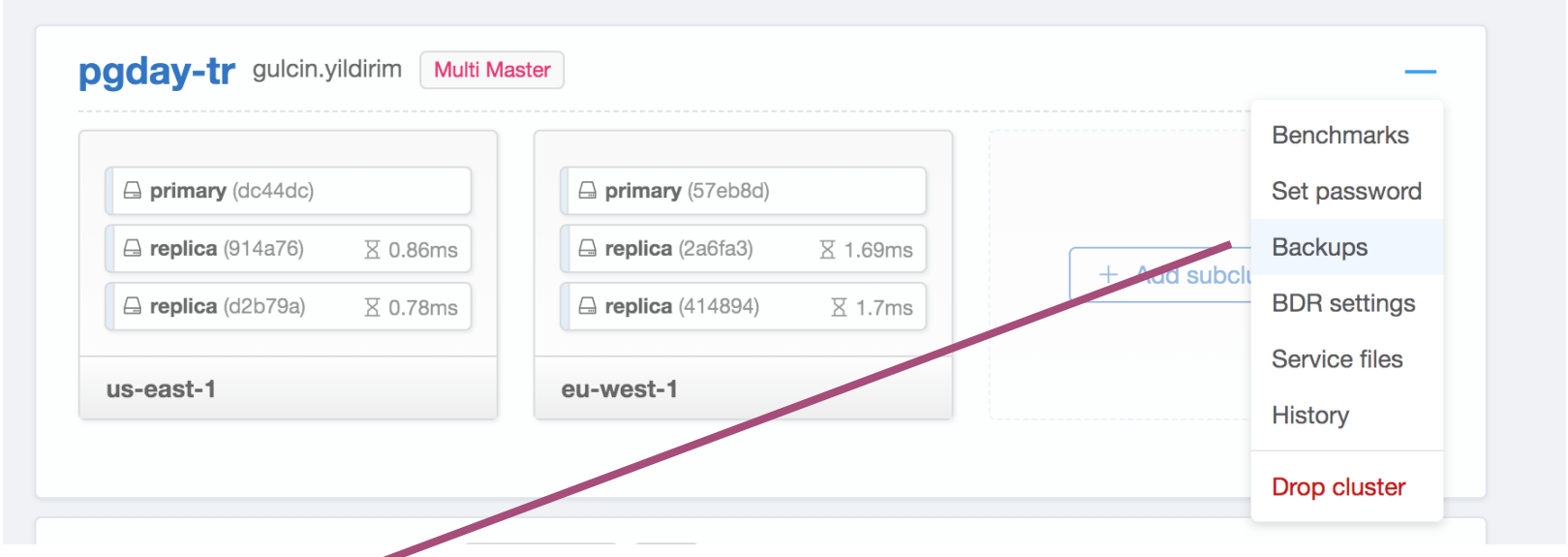


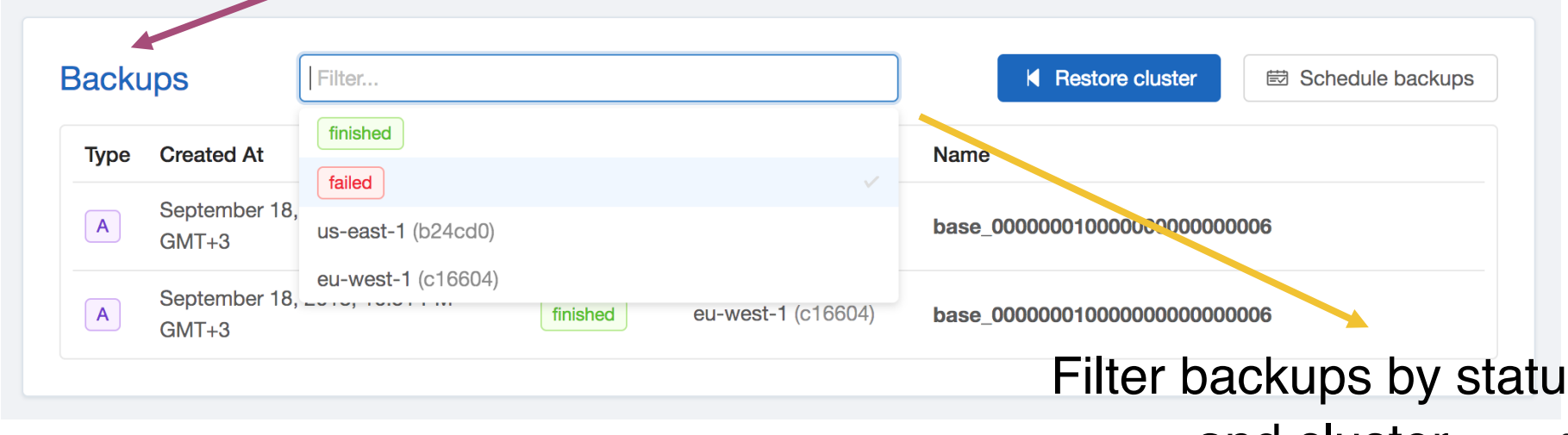
Automated Backups





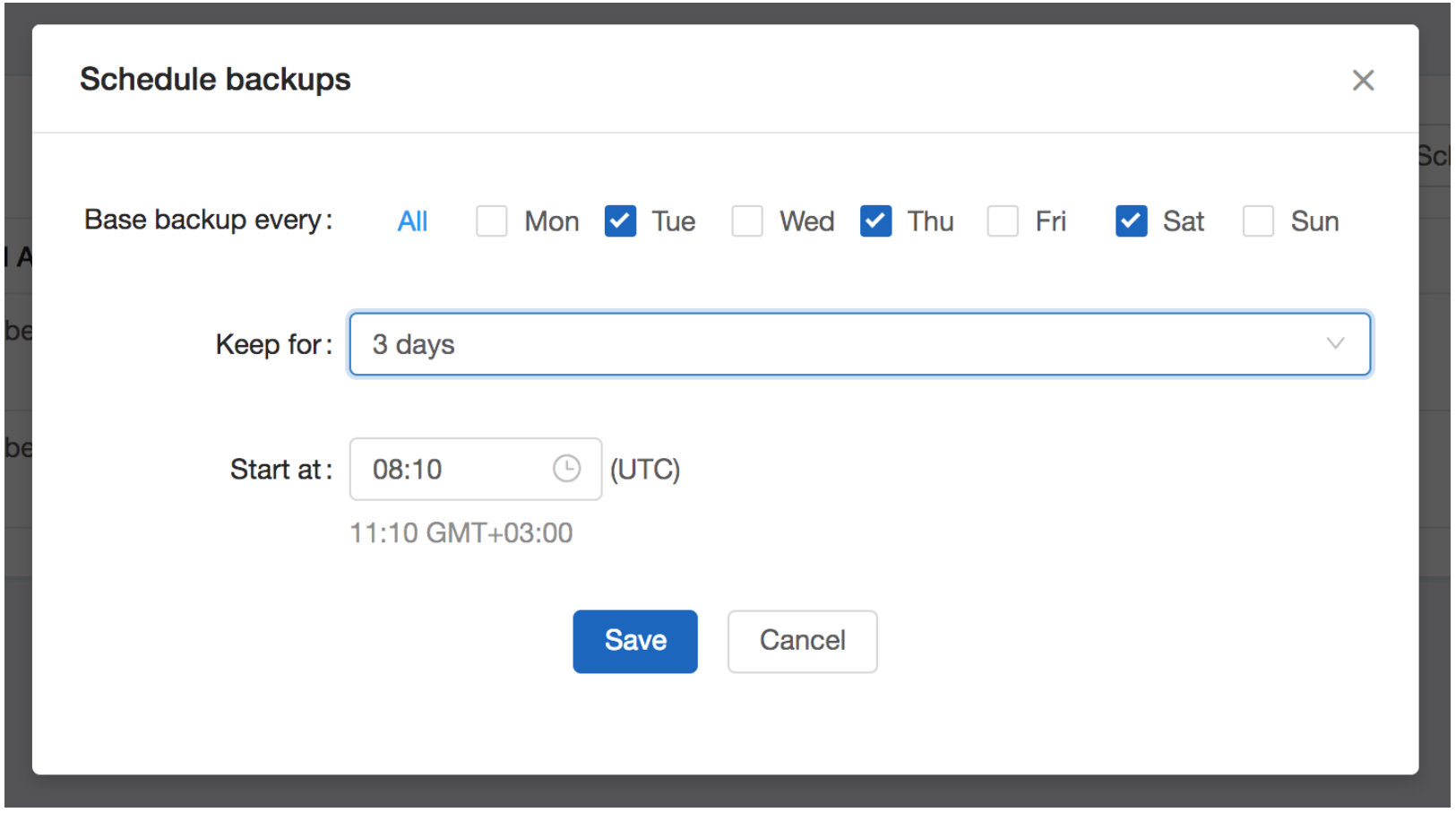
Automated Backups



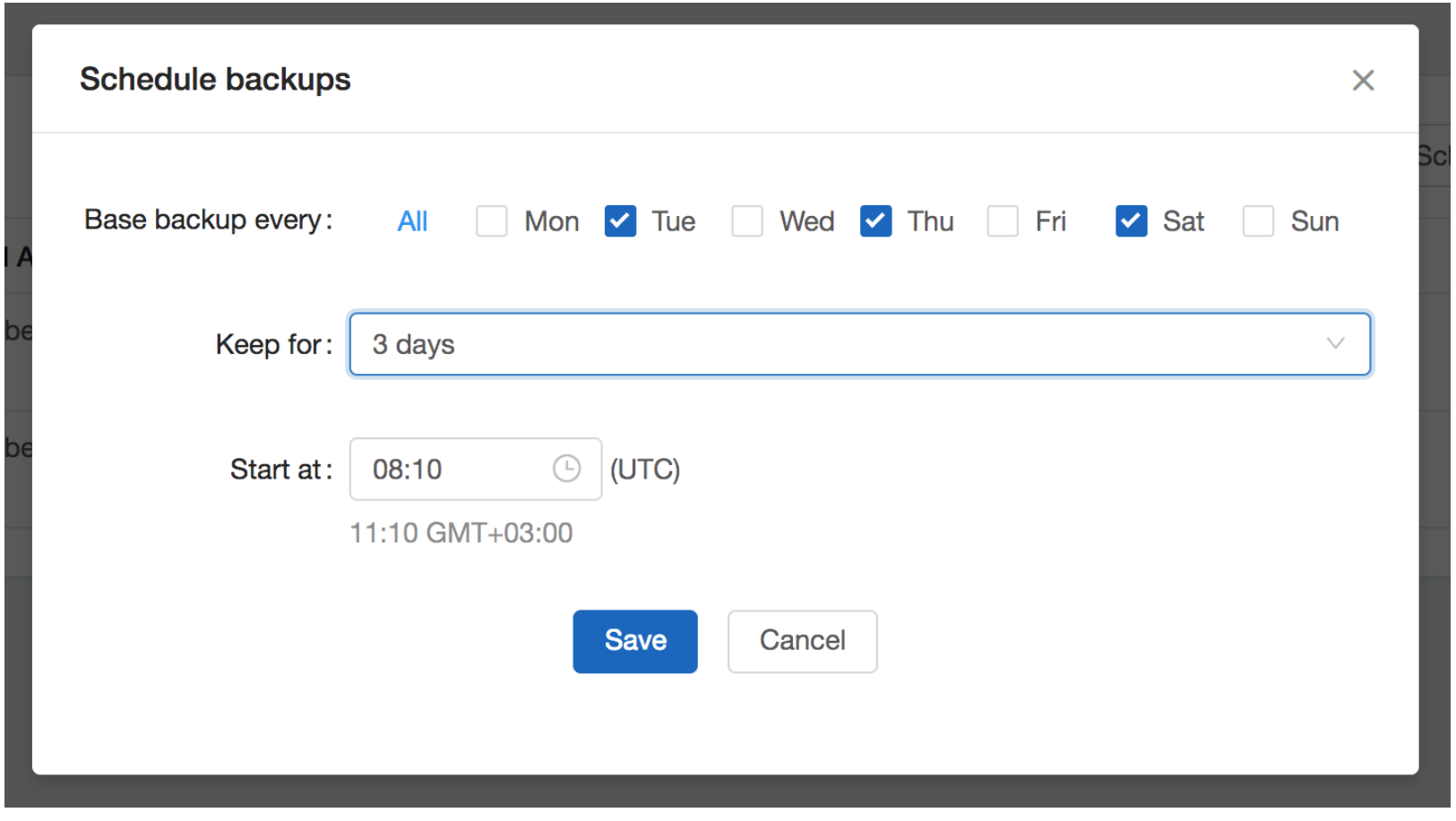


Filter backups by status and cluster

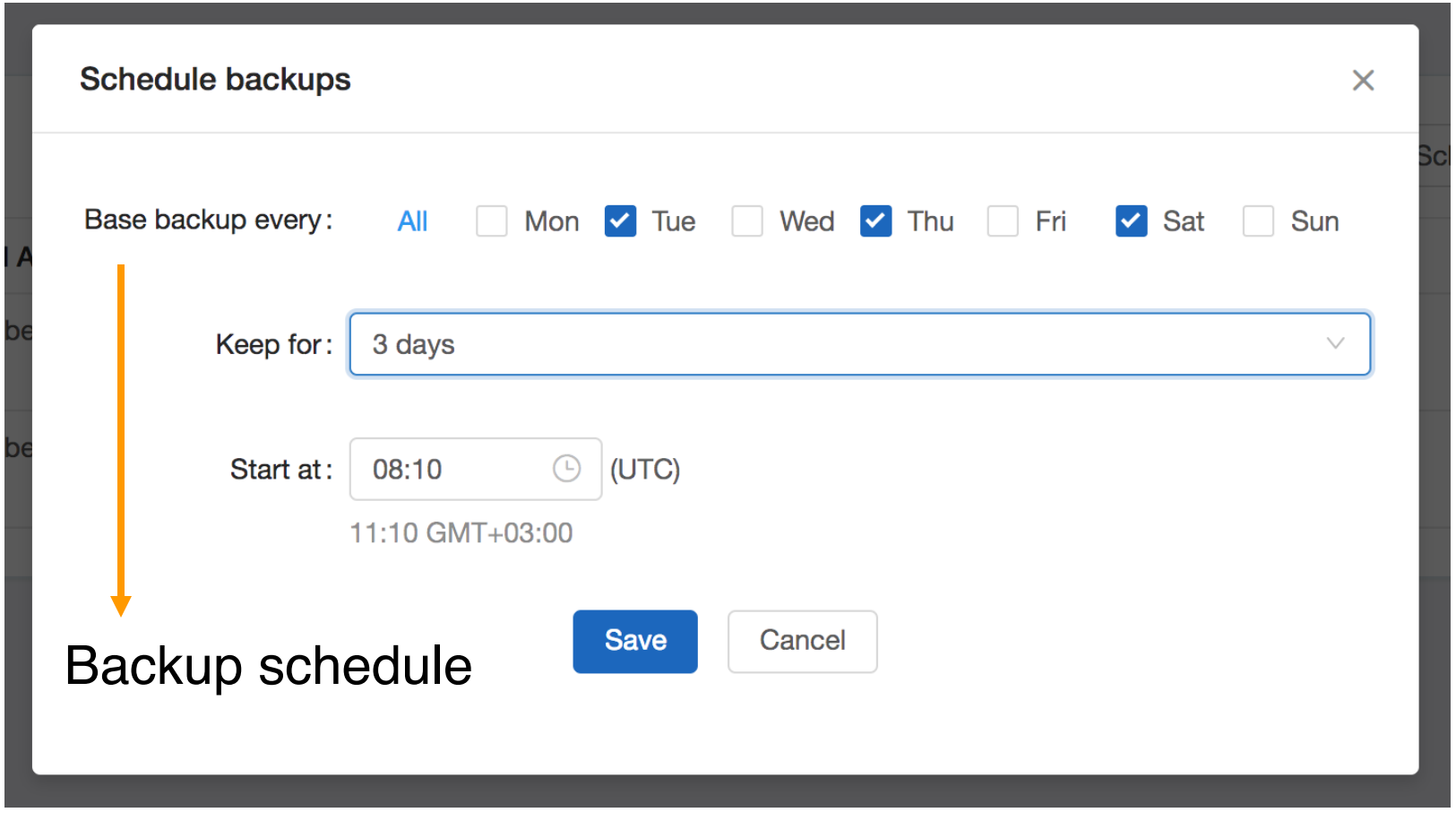
Scheduled Backups



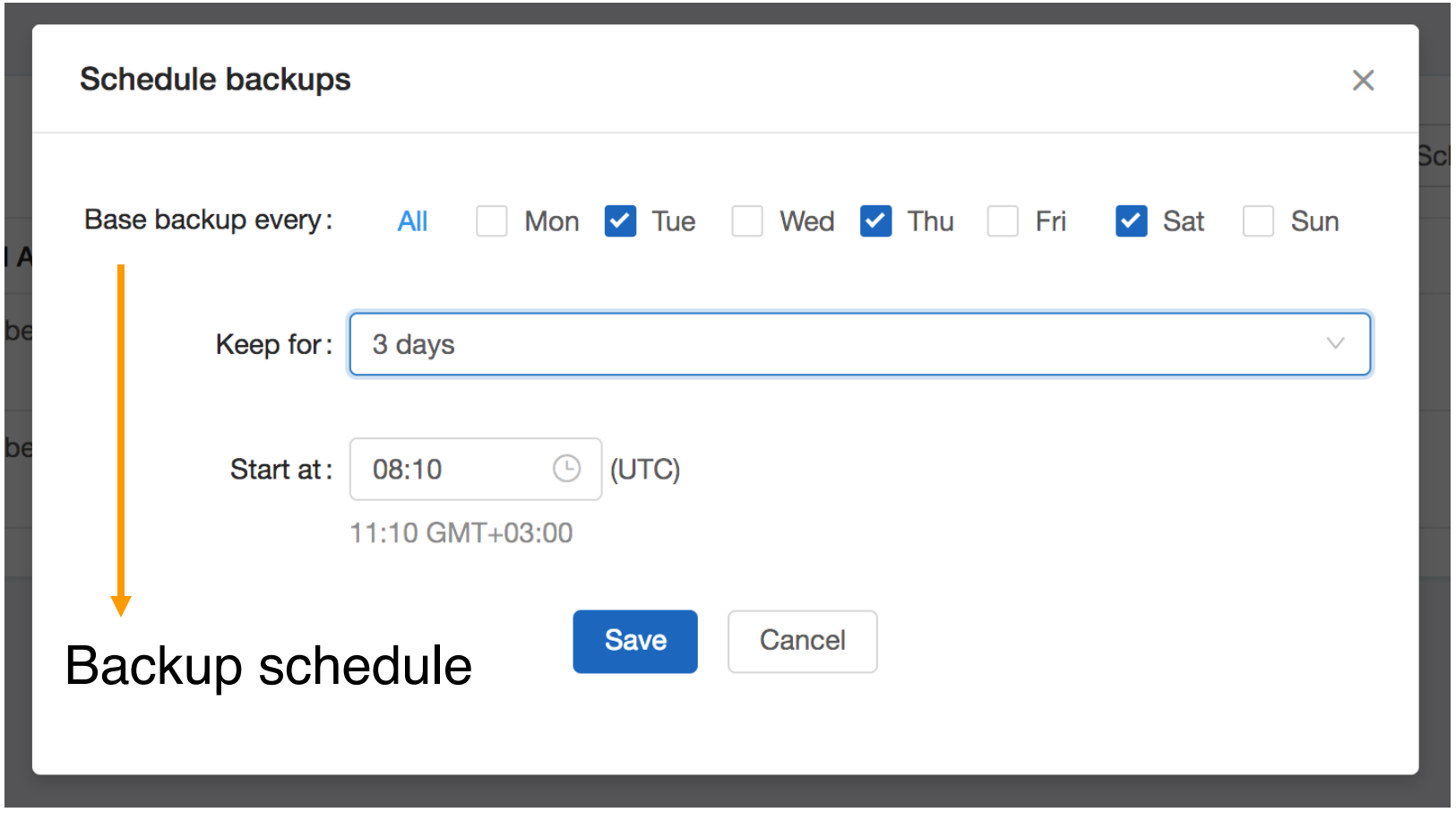
Scheduled Backups



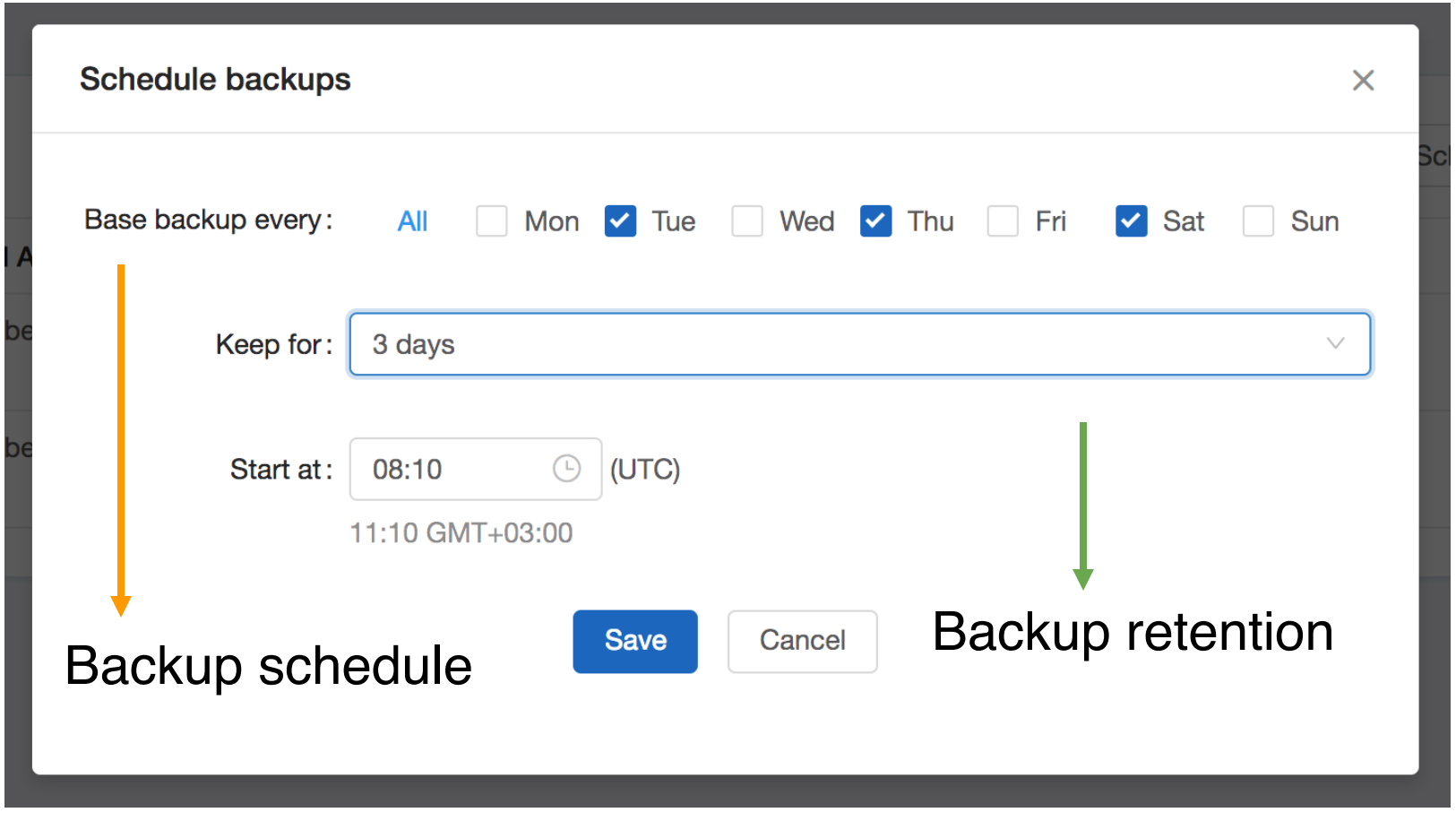
Scheduled Backups



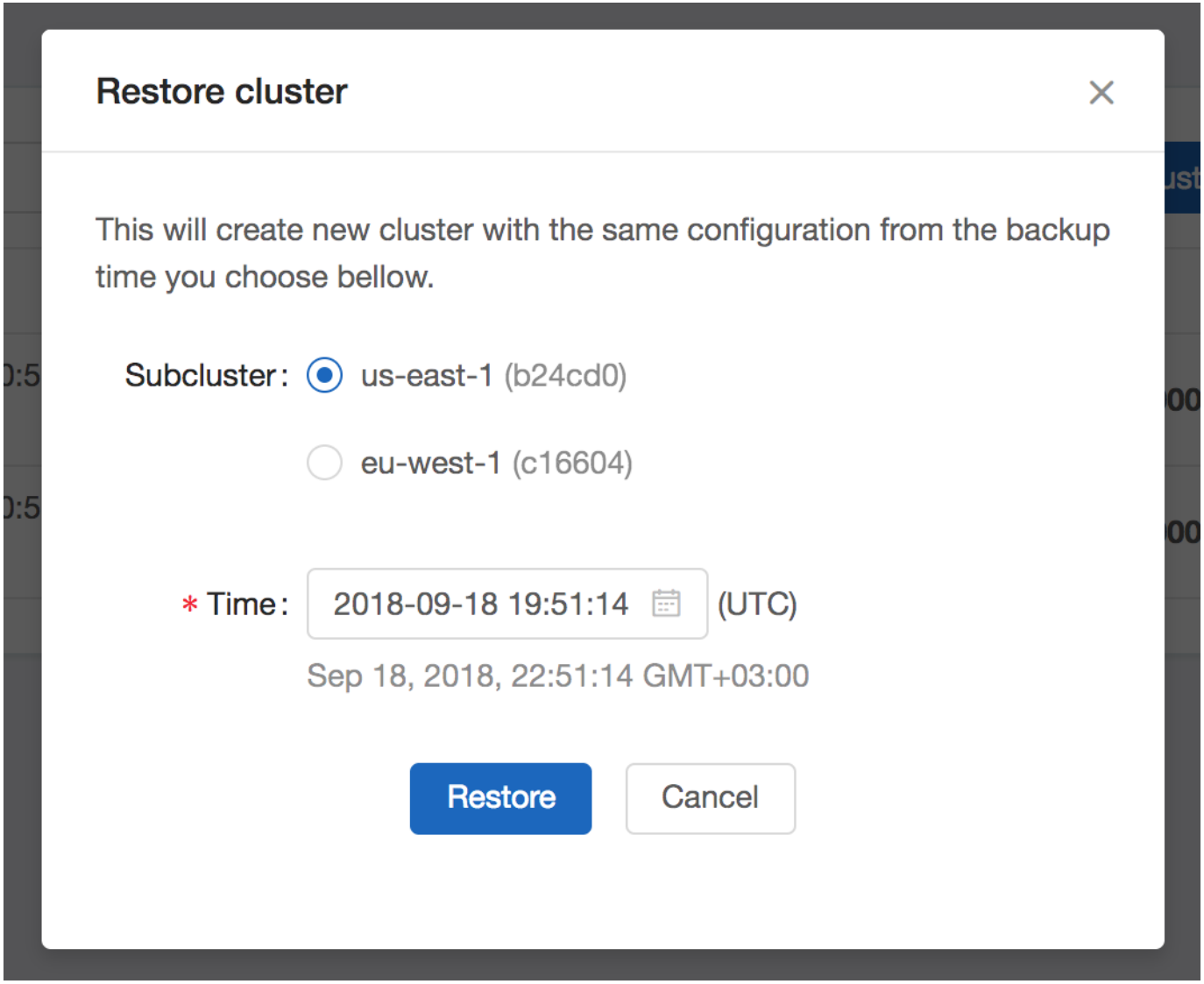
Scheduled Backups



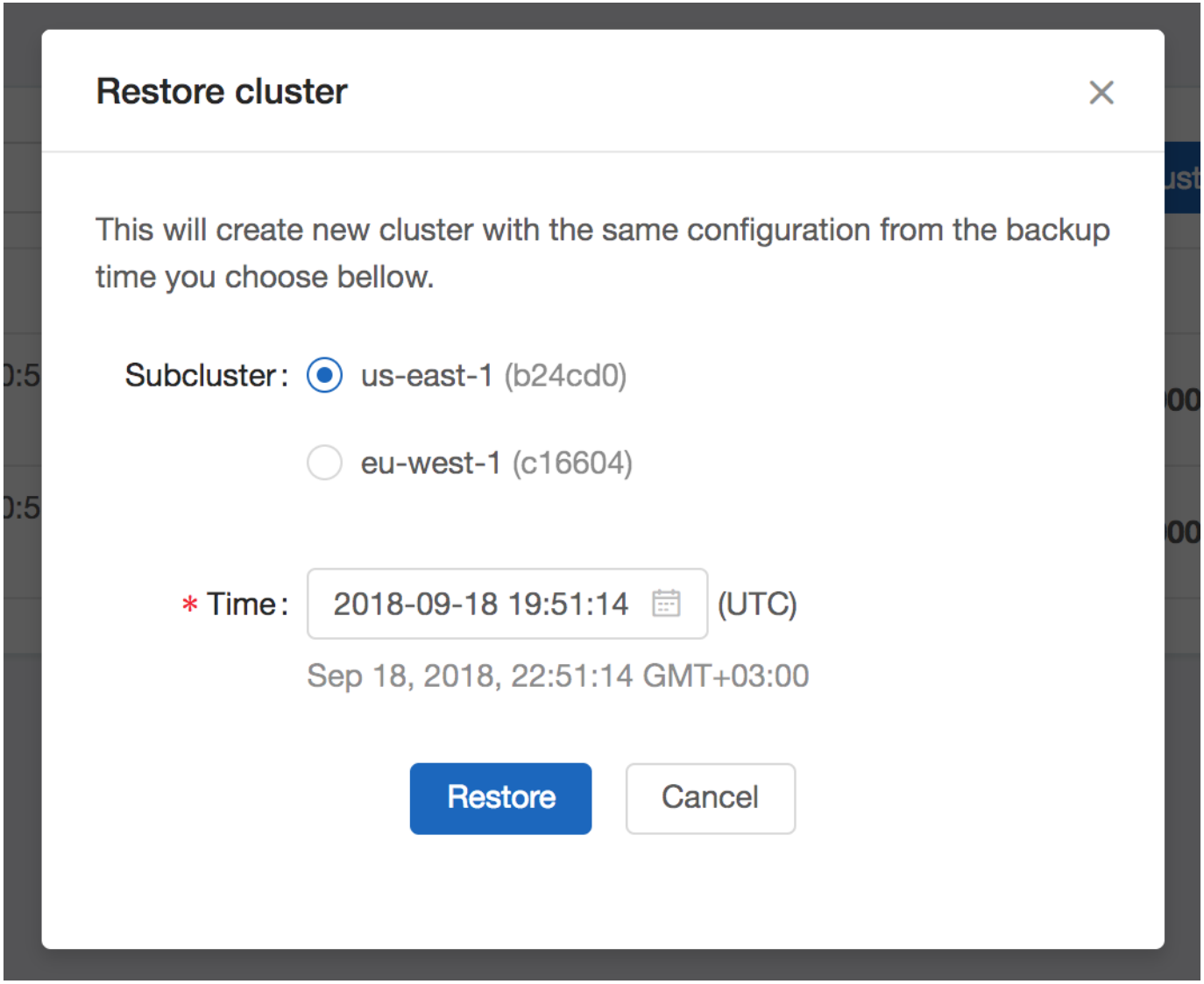
Scheduled Backups



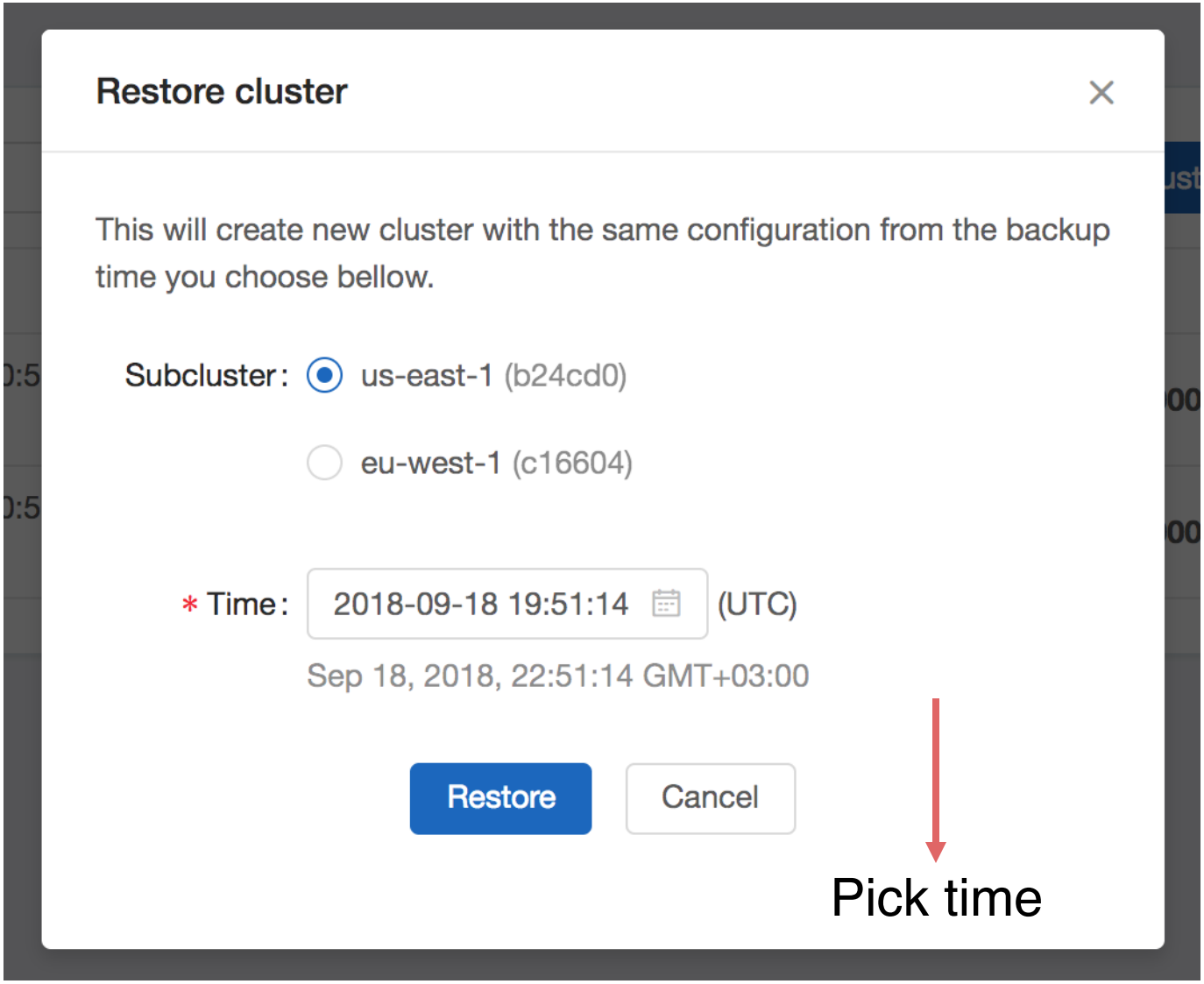
Point-in-Time Recovery



Point-in-Time Recovery



Point-in-Time Recovery



Documentation

Documentation

GDS comes in three flavours:

- **Single Master** (offers physical streaming replication)
- **Multi-Master** (offers BDR and physical streaming replication)
- **Always On** (offers BDR in a very highly available setup)

GDS Deployment Types

GDS Single Master has 3 deployment types:

- **Development:** Useful for development purposes
- **Test:** Useful for testing/staging your setup before going production
- **Production:** Best fit for production use, comes with Platinum Support and HA

GDS Multi-Master has one type and you can use it for all deployments. You will have 2 standby servers for each BDR primary server you created. The design enforces high availability starting with the initial configuration. You will also get 2ndQuadrant Diamond Support (the highest level) by creating a Multi-Master.

GDS Always On is designed to provide at least 4 Nines Availability (%99.99) on BDR architectures.

How to Start?

You can login to GDS through our Customer Portal, once you are authorized you will be able to create clusters. Start by clicking “Create new cluster” at the top right of the main page.

Documentation

GDS comes in three flavours:

- **Single Master** (offers physical streaming replication)
- **Multi-Master** (offers BDR and physical streaming replication)
- **Always On** (offers BDR in a very highly available setup)

GDS Deployment Types

GDS Single Master has 3 deployment types:

- **Development:** Useful for development purposes
- **Test:** Useful for testing/staging your setup before going production
- **Production:** Best fit for production use, comes with Platinum Support and HA

GDS Multi-Master has one type and you can use it for all deployments. You will have 2 standby servers for each BDR primary server you created. The design enforces high availability starting with the initial configuration. You will also get 2ndQuadrant Diamond Support (the highest level) by creating a Multi-Master.

GDS Always On is designed to provide at least 4 Nines Availability (%99.99) on BDR architectures.

How to Start?

You can login to GDS through our Customer Portal, once you are authorized you will be able to create clusters. Start by clicking “Create new cluster” at the top right of the main page.

Docs (WiP)

Documentation

GDS comes in three flavours:

- **Single Master** (offers physical streaming replication)
- **Multi-Master** (offers BDR and physical streaming replication)
- **Always On** (offers BDR in a very highly available setup)

GDS Deployment Types

GDS Single Master has 3 deployment types:

- **Development:** Useful for development purposes
- **Test:** Useful for testing/staging your setup before going production
- **Production:** Best fit for production use, comes with Platinum Support and HA

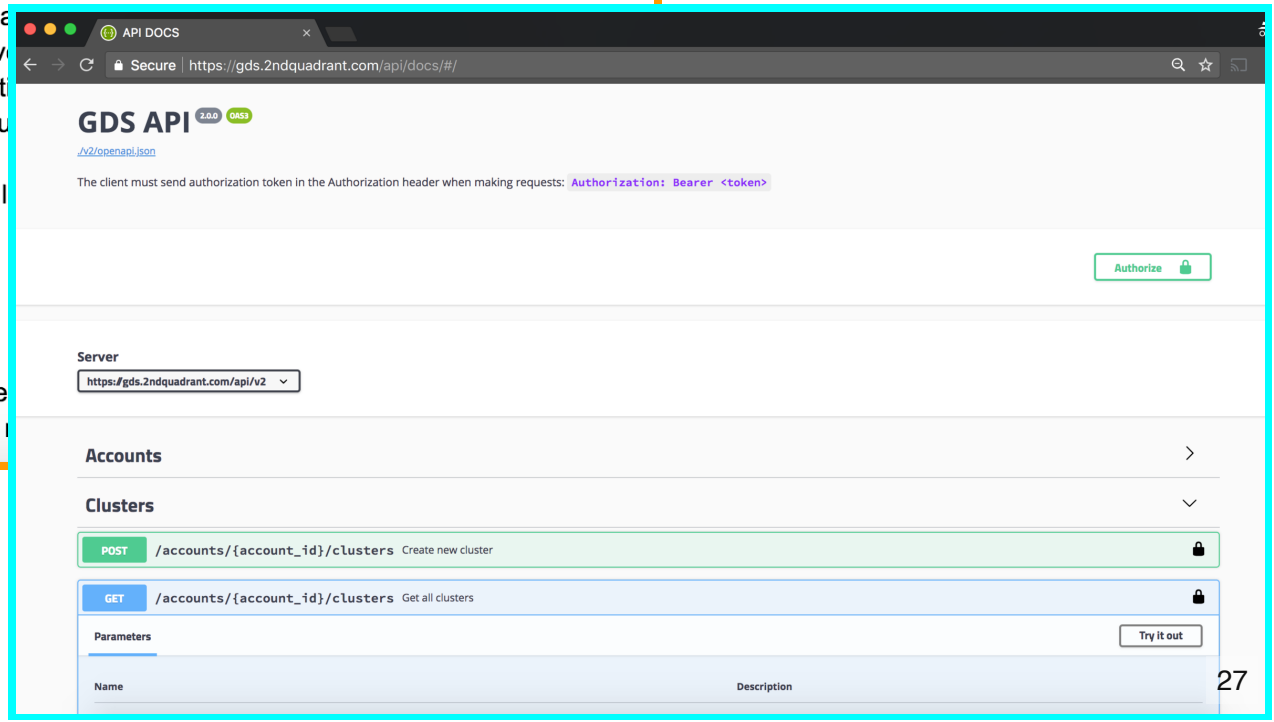
GDS Multi-Master has one type and you ca
standby servers for each BDR primary serv
availability starting with the initial configurat
Support (the highest level) by creating a Mu

GDS Always On is designed to provide at l
architectures.

How to Start?

You can login to GDS through our Custome
to create clusters. Start by clicking “Create

Docs (WiP)



Documentation

GDS comes in three flavours:

- **Single Master** (offers physical streaming replication)
- **Multi-Master** (offers BDR and physical streaming replication)
- **Always On** (offers BDR in a very highly available setup)

GDS Deployment Types

GDS Single Master has 3 deployment types:

- **Development:** Useful for development purposes
- **Test:** Useful for testing/staging your setup before going production
- **Production:** Best fit for production use, comes with Platinum Support and HA

GDS Multi-Master has one type and you ca
standby servers for each BDR primary serv
availability starting with the initial configurat
Support (the highest level) by creating a Mu

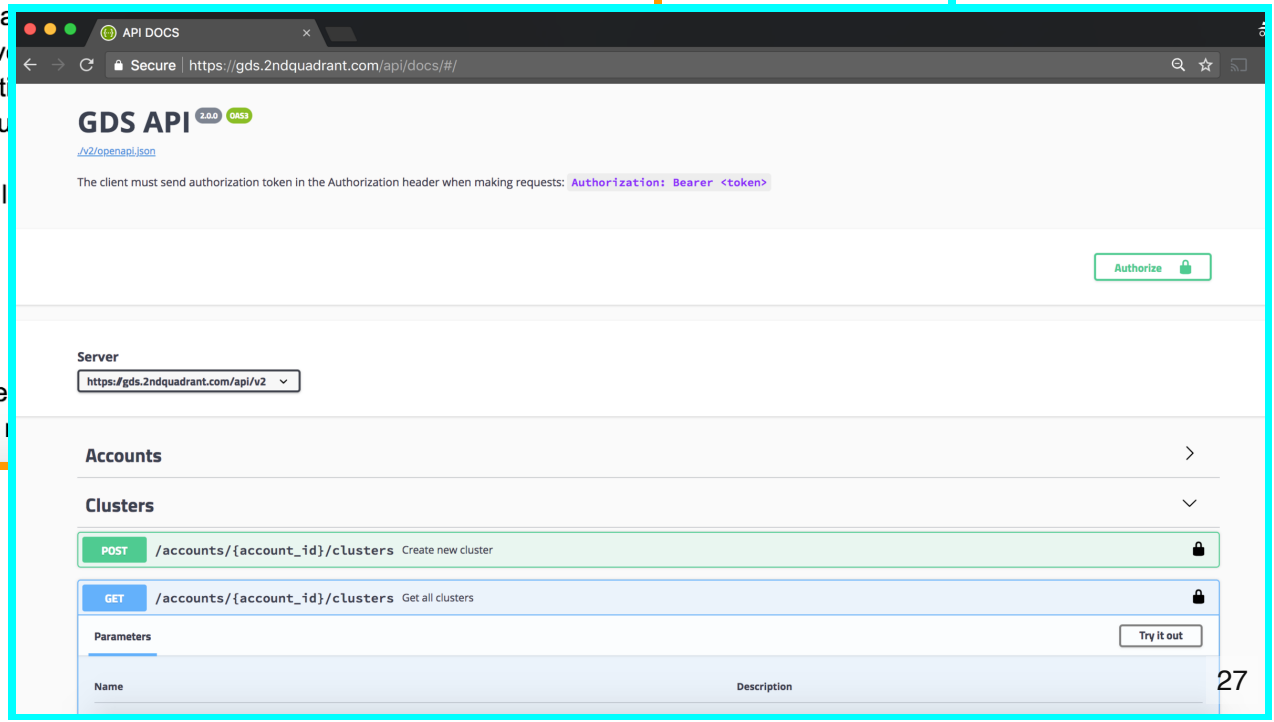
GDS Always On is designed to provide at l
architectures.

How to Start?

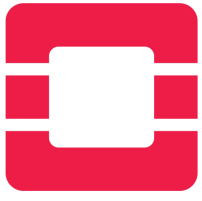
You can login to GDS through our Custome
to create clusters. Start by clicking “Create

Docs (WiP)

Interactive API docs



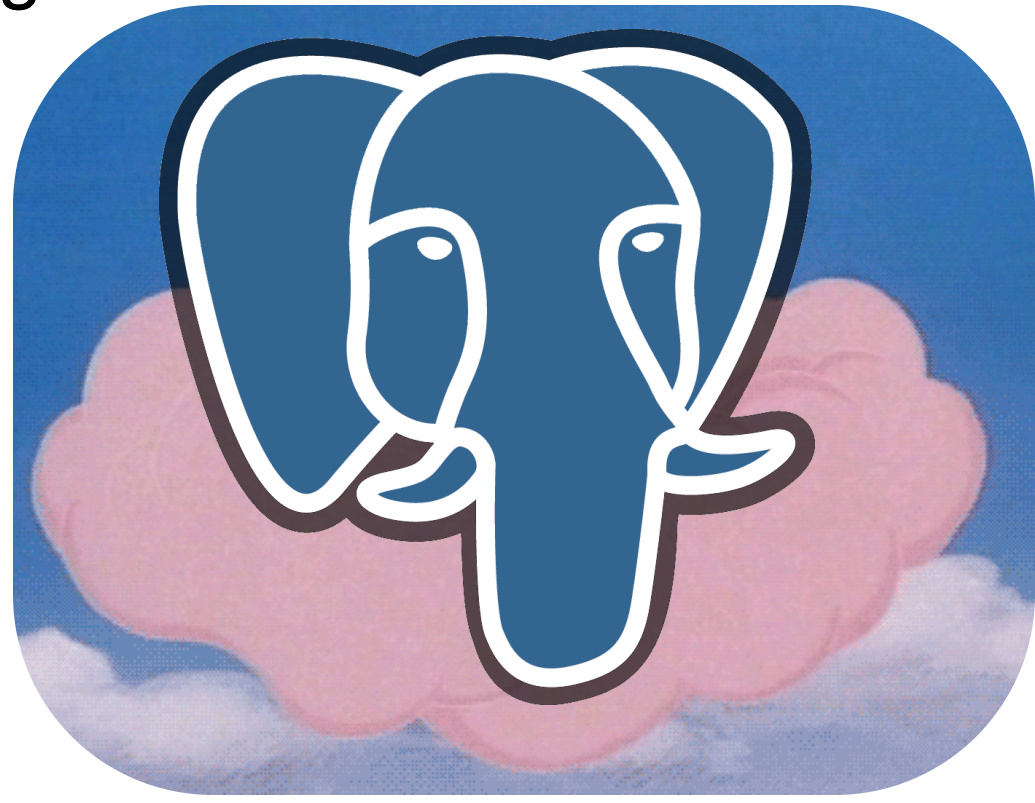
Postgres Cloud Manager



openstack®

Conclusion

- Postgres in Cloud is a go!
- Multi-Master adds new possibilities
- It's possible to create your own cloud independent of the platform
- Automation enables easy management
- Postgres Support applies to Cloud!



2ndQuadrant® PostgreSQL

Tešekkürler!

Questions?