



EDB™

Cluster your application in
40 minutes

Raphael Salguero
PGConf.de 2023 - Essen

Internal Server Error

500

The server encountered an internal error or misconfiguration
and was unable to complete your request

#whoami

Raphael Salguero

Senior Migrations Consultant

- With EDB since 04/2023
- DBA, Trainer & Consulting
- Oracle, PostgreSQL, MySQL

- DOAG delegate

- First time PGConf

- Husband and proud dad
- football, enjoying food, spending time with family



Disclaimer

- Different audiences, different focus.
- My 'Patroni Stack' is not the ideal solution – it's just an idea growing through the talk
- My focus is not about Patroni itself, so there won't be a focus on all different configuration possibilities

AGENDA

01

Basic

AGENDA

01

Basic

02

Patroni
Stack

AGENDA

01

Basic

02

Patroni
Stack

03

PGBouncer

AGENDA

01

Basic

02

Patroni
Stack

03

PGBouncer

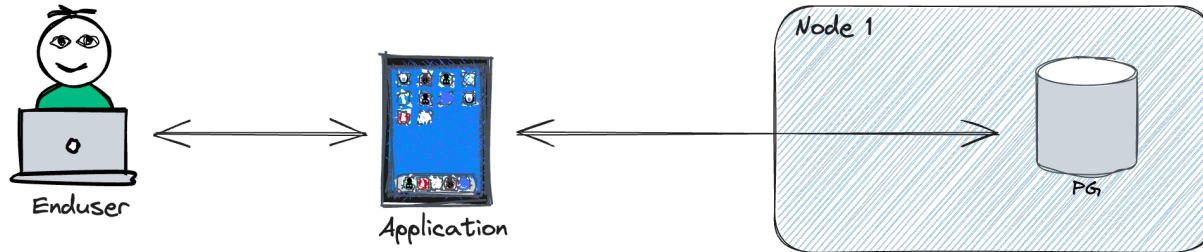
04

Demo

Basic

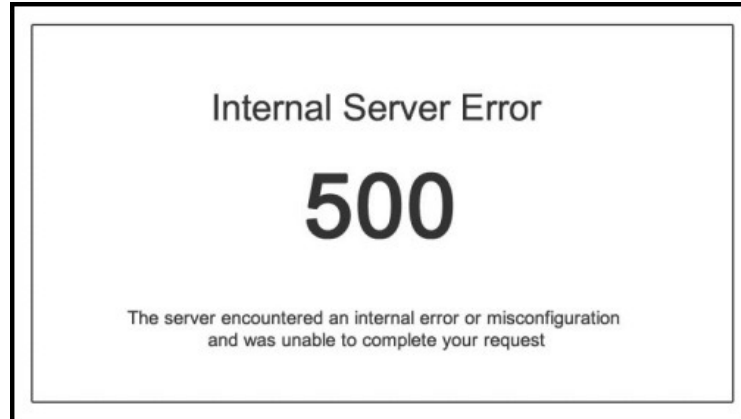
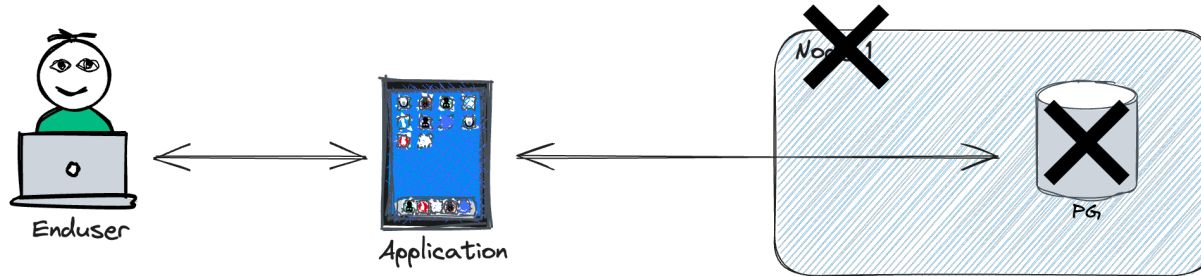
www.app.de/app

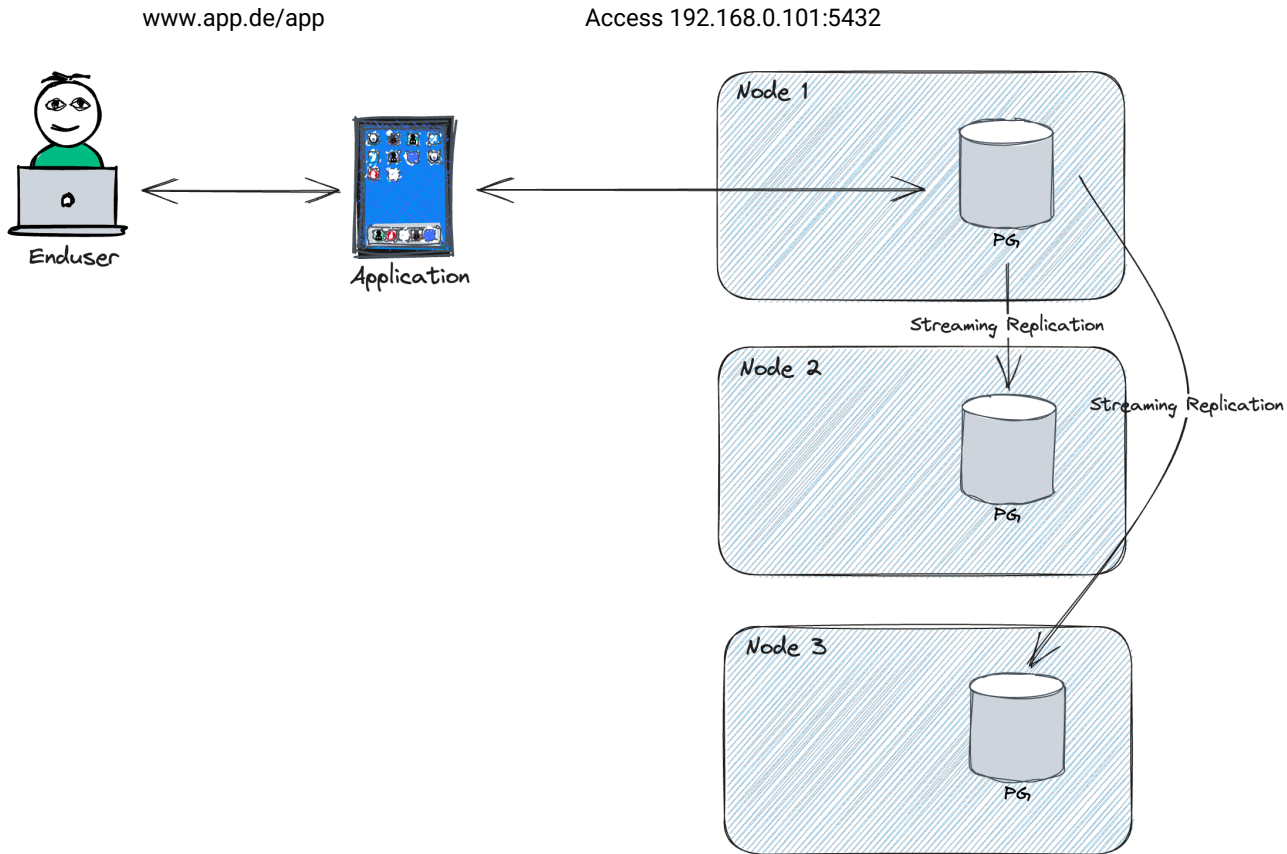
Access 192.168.0.101:5432

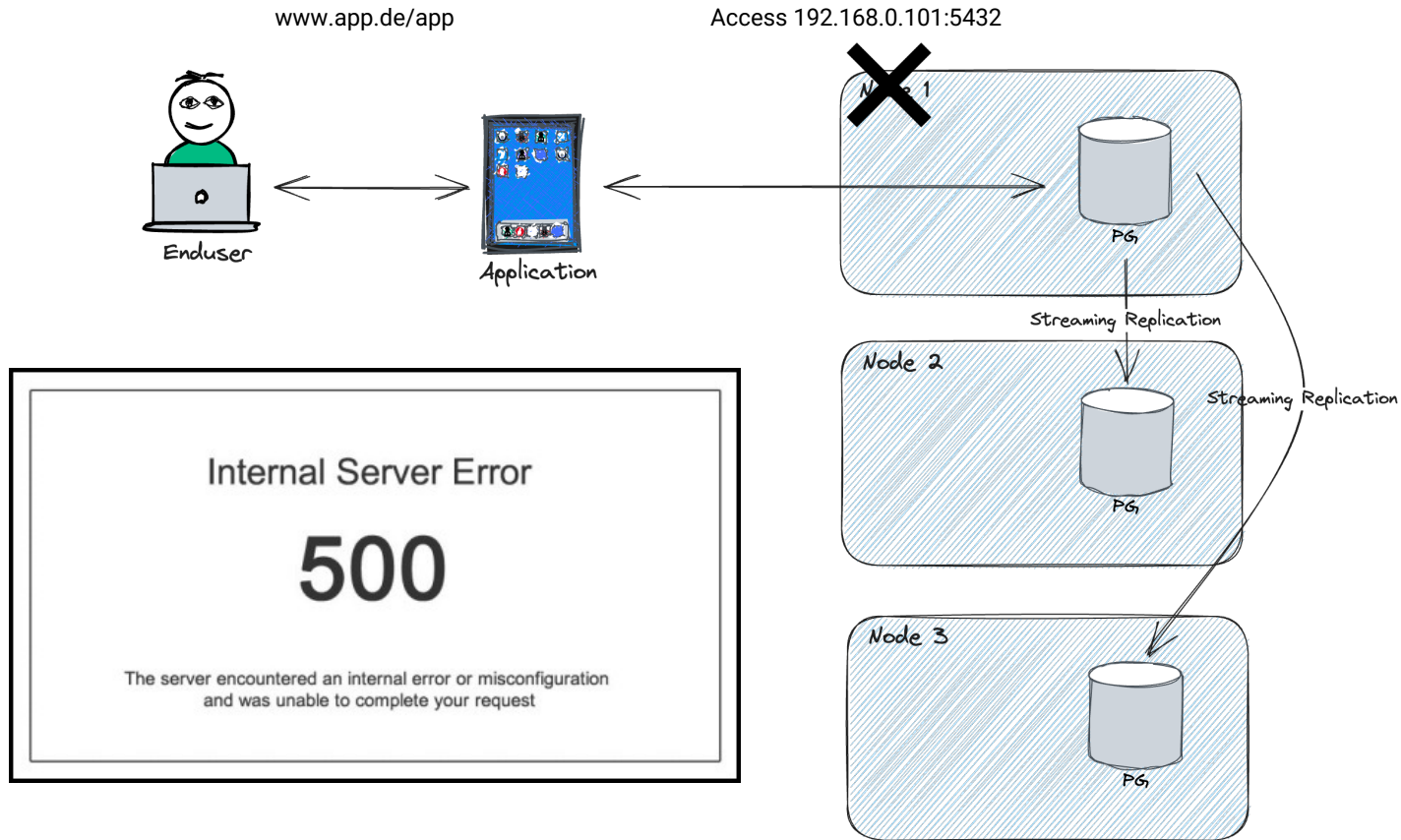


www.app.de/app

Access 192.168.0.101:5432

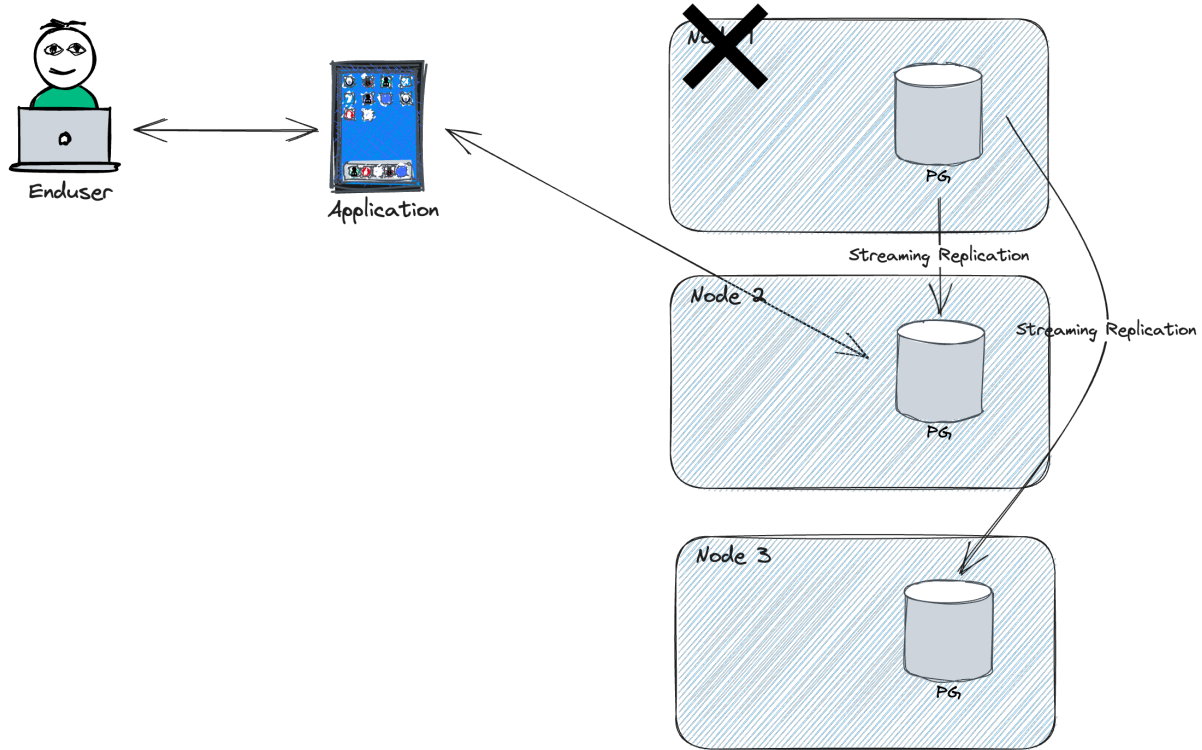


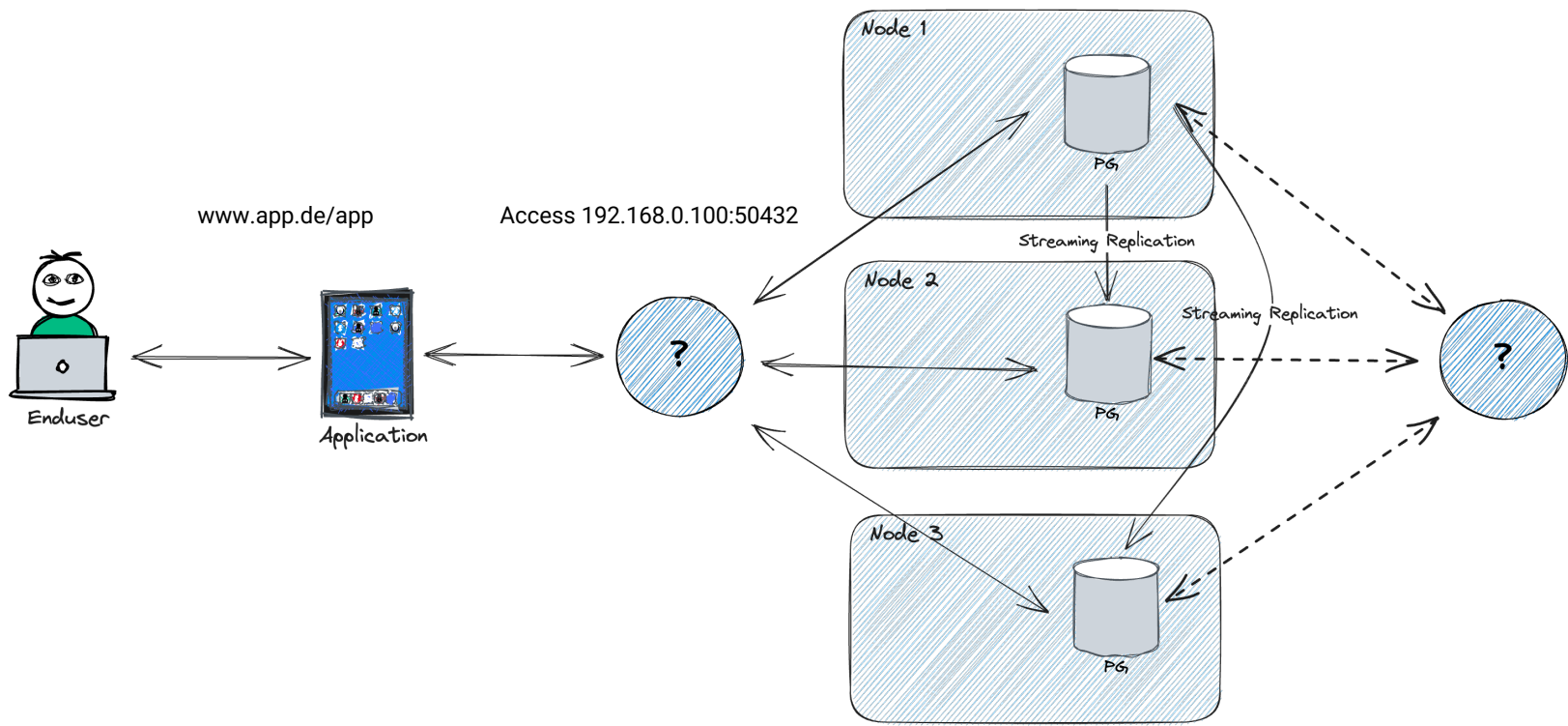


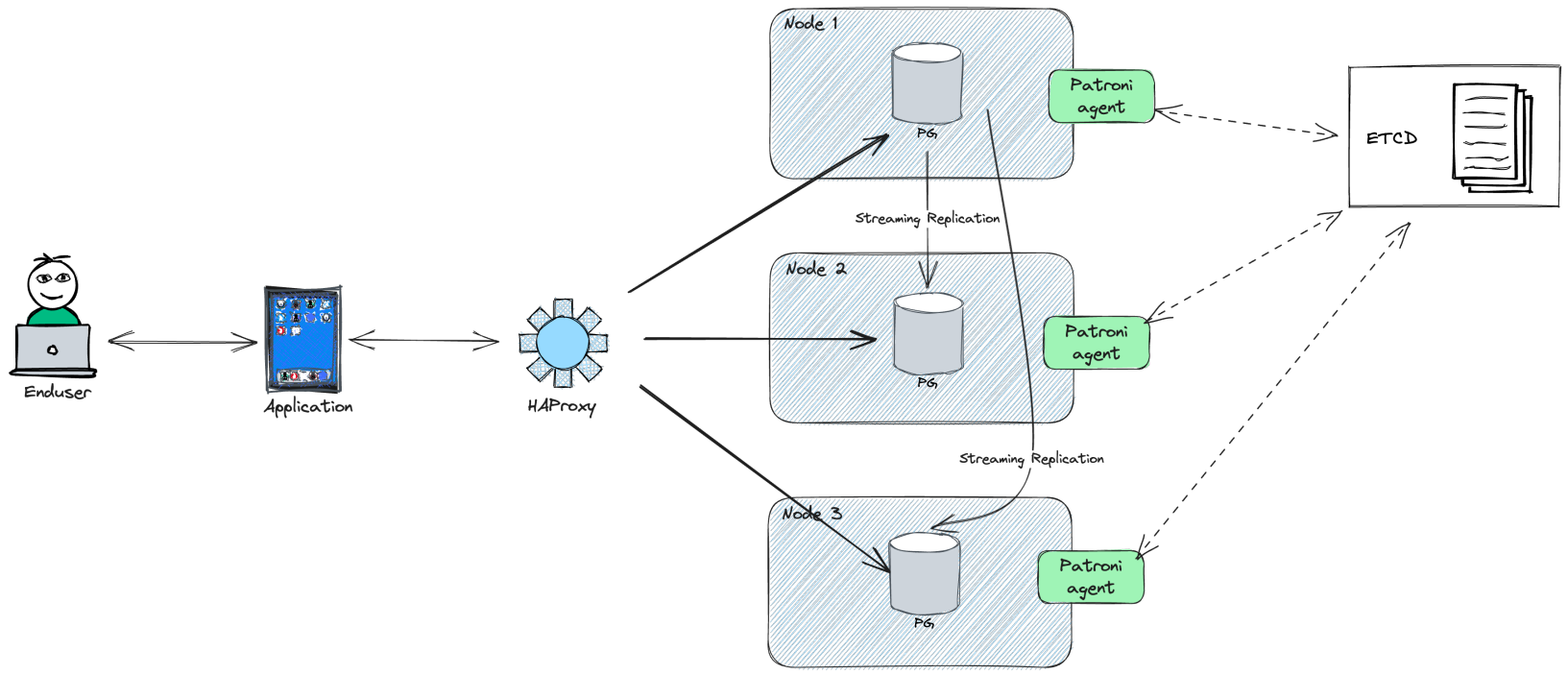


www.app.de/app

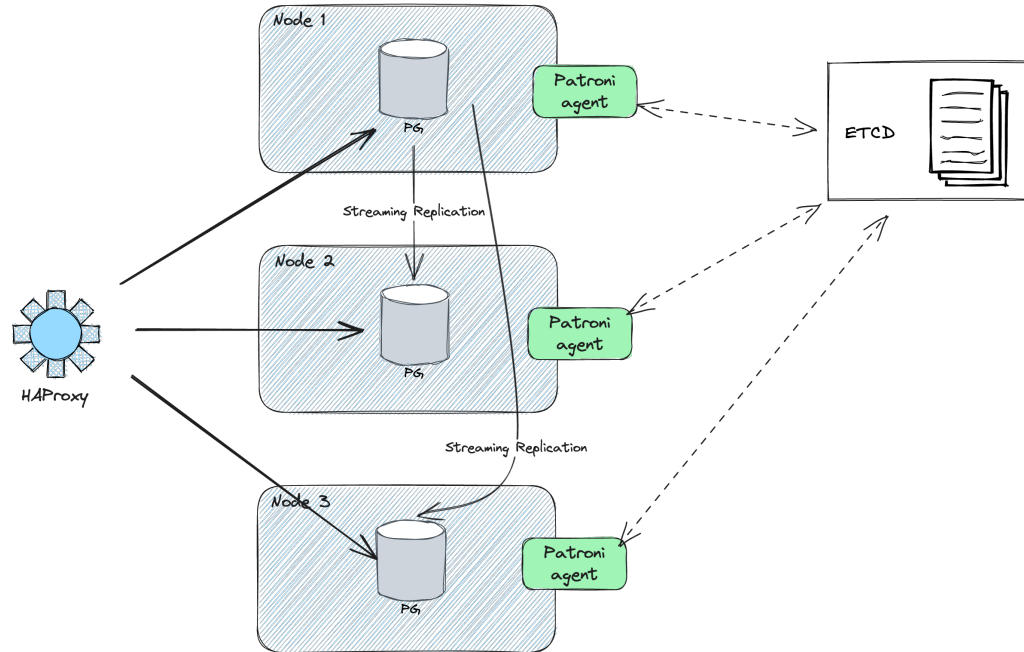
Access 192.168.0.102:5432





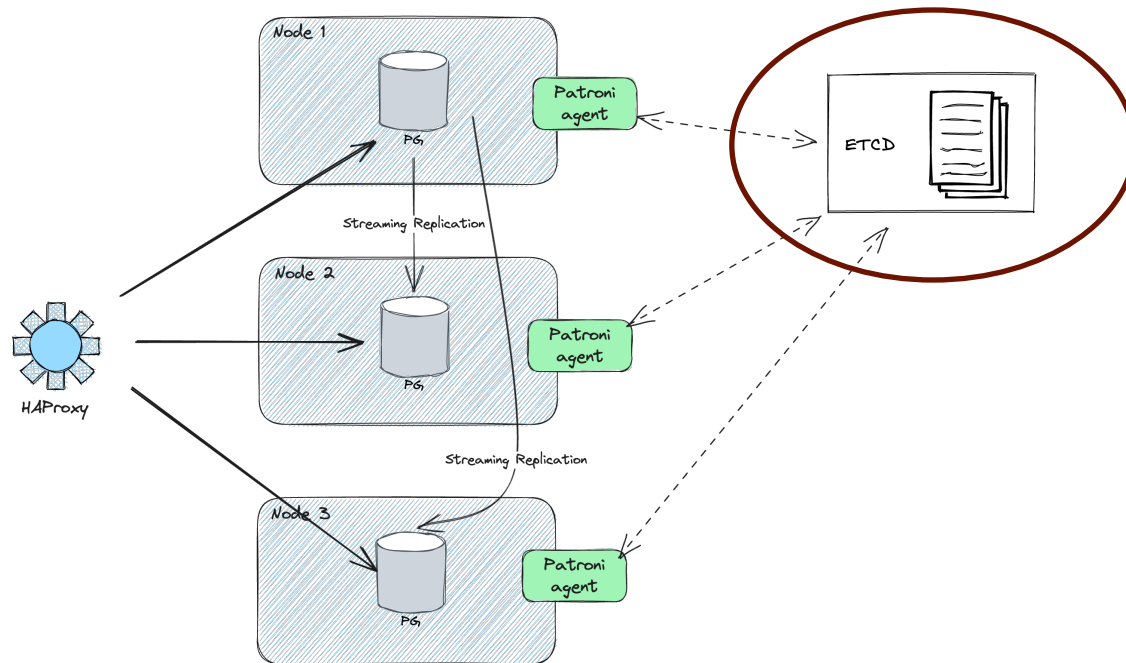


Let's call it my "Patroni Stack" ..



Patroni Stack

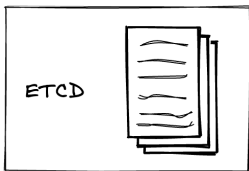
“Patroni Stack”





Role of ETCD

- Cluster Coordination
- Leader Election
- Failover and High Availability
- Configuration Management
- Consistency and Synchronization
- Monitoring and Health Checking



Basic Configuration

```
ETCD_LISTEN_PEER_URLS=http://127.0.0.1:2380,http://127.0.0.1:7001
```

```
ETCD_LISTEN_CLIENT_URLS=http://127.0.0.1:2379, http://127.0.0.1:2379
```

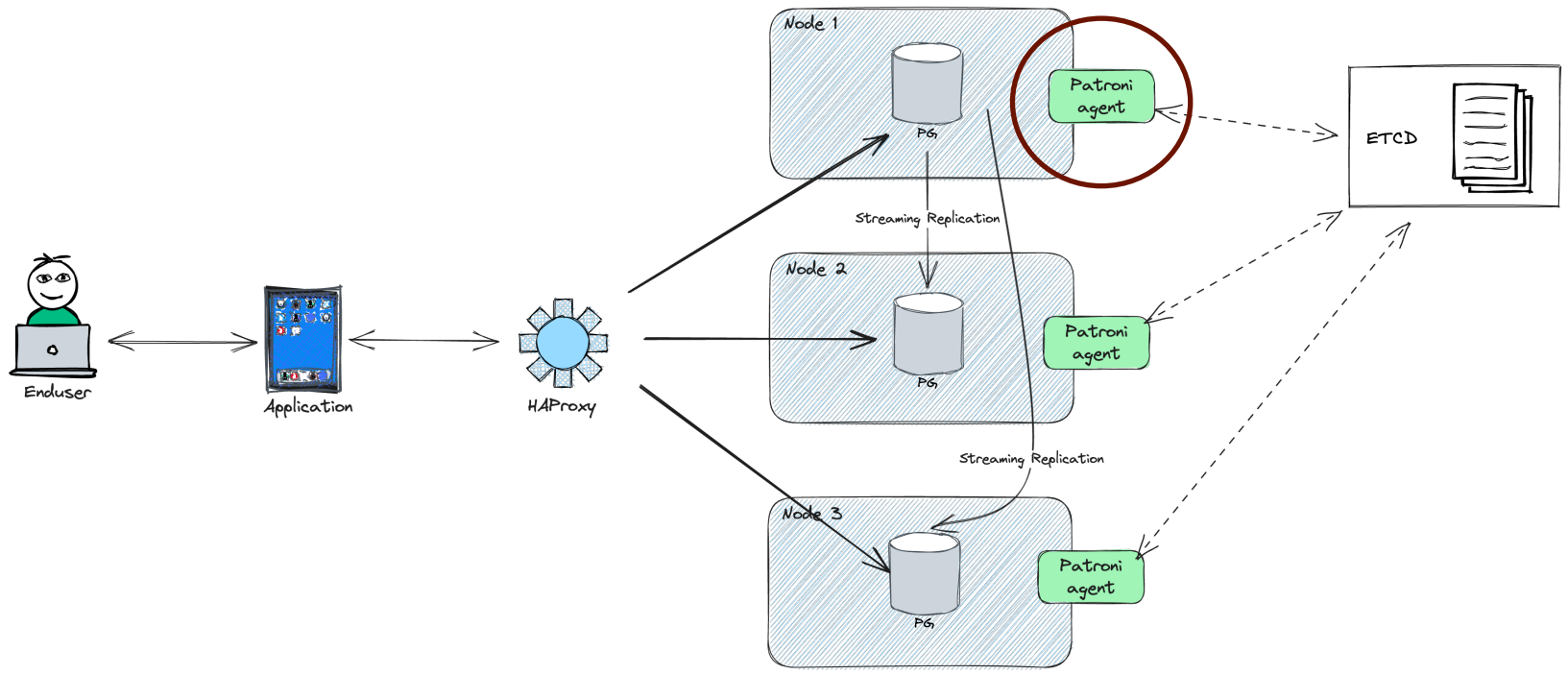
```
ETCD_INITIAL_ADVERTISE_PEER_URLS=http://127.0.0.1:2380
```

```
ETCD_INITIAL_CLUSTER="etcd0=http://127.0.0.1:2380"
```

```
ETCD_ADVERTISE_CLIENT_URLS=http://127.0.0.1:2379
```

```
ETCD_INITIAL_CLUSTER_TOKEN="cluster1"
```

```
ETCD_INITIAL_CLUSTER_STATE="new"
```





A diagram consisting of a large blue rounded rectangle. Inside it, on the left side, is a smaller blue rounded rectangle. Inside that smaller rectangle is a green rounded rectangle with the text "Patroni agent" written in black, handwritten-style font.

Patroni
agent

Role of Patroni Agents

- Node Management
- Automatic Leader Election
- Failover Handling
- Configuration Management
- Promotion and Demotion
- Health Monitoring and Reporting
- Integration with External Tools

Patroni
agent

patroni.yml

```
scope: patroni15
name: pg1
restapi:
  listen: localhost:8008
  connect_address: localhost:8008
etcd3:
  hosts: localhost:2379
bootstrap:
  initdb:
    - encoding: UTF8
    - data-checksums

  pg_hba:
    - host replication postgres 0.0.0.0/0 scram-sha-256
    - host all all 0.0.0.0/0 scram-sha-256
```


Patroni
agent

```
postgresql:  
  listen: localhost:5432  
  bin_dir: /usr/lib/postgresql/15/bin  
  connect_address: localhost:5432  
  proxy_address: localhost:5432  
  data_dir: /var/lib/postgresql/15/main  
  pgpass: /tmp/pgpass  
  authentication:  
    replication:  
      username: postgres  
      password: postgres  
  superuser:  
    username: postgres  
    password: postgres  
  rewind:  
    username: postgres  
    password: postgres  
  parameters:  
    cluster_name: patroni15
```

Patroni
agent

patroni.service

```
[Unit]
Description=agent for patroni HA cluster
After=syslog.target network.target

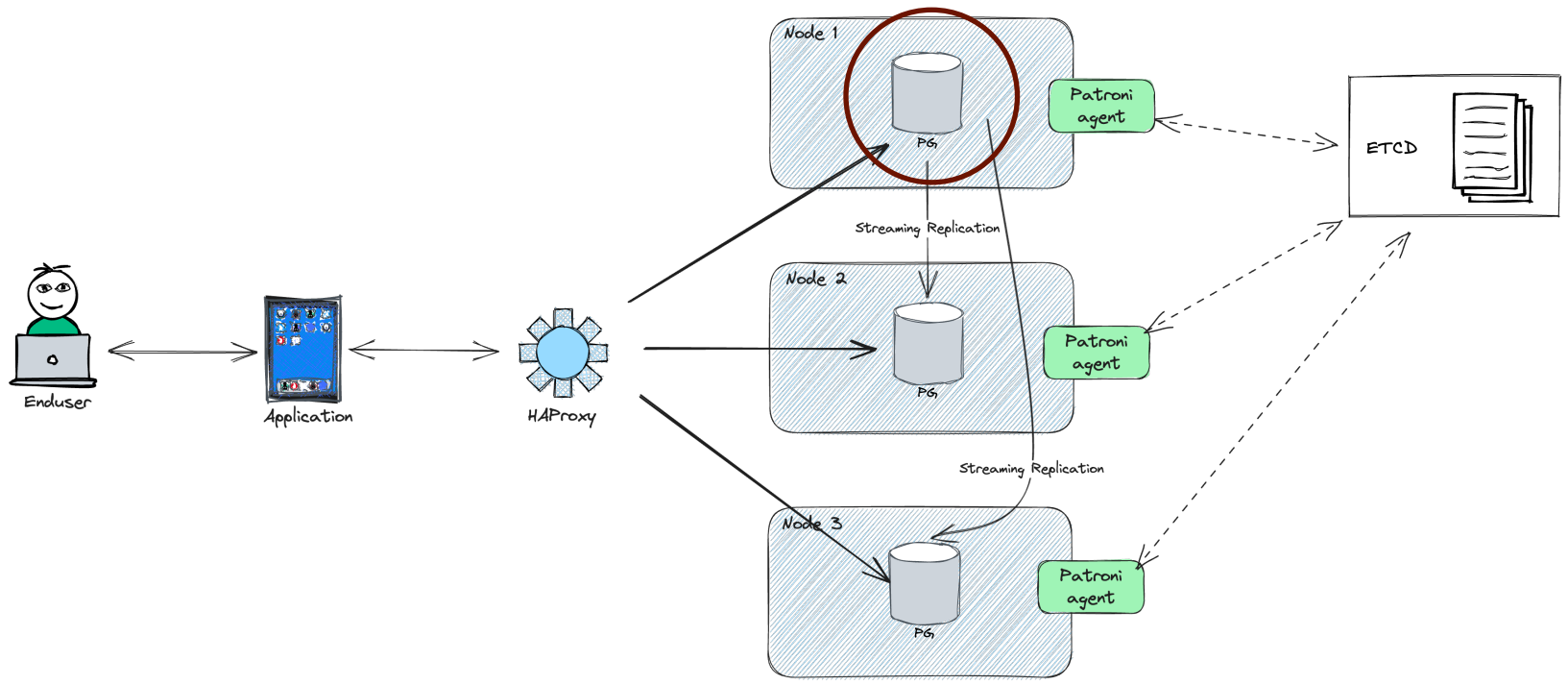
[Service]
Type=simple
User=postgres
Group=postgres
ExecStart=/var/lib/postgresql/.local/bin/patroni /etc/patroni1.yml
KillMode=process
TimeoutSec=30
Restart=no

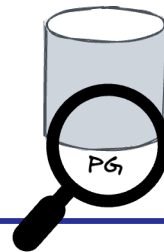
[Install]
WantedBy=multi-user.target
```

Patroni
agent

```
systemctl daemon-reload
systemctl start patroni
systemctl status patroni
```

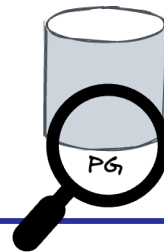
```
...
Jun 13 11:15:37 ip-172-31-2-106 patroni[34284]: 2023-06-13
11:15:37,530 INFO: no action. I am (pg1), the leader with the lock
...
```



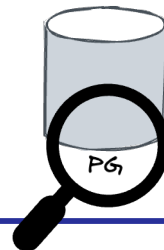


PG Result

- Patroni will
 - a) setup a new cluster
 - b) configure on top of an existing cluster
- Configuration based on parameters
- Streaming replication will be configured and enabled
- Administration users will be created
- pg_hba.conf will be enhanced



```
postgres=# select * from pg_stat_replication;
-[ RECORD 1 ]-----+-----
pid            | 34595
usesysid      | 10
username       | postgres
application_name | pg2
client_addr    | 127.0.0.1
client_hostname |
client_port    | 52686
backend_start  | 2023-06-13 08:30:38.166899+00
backend_xmin   |
state          | streaming
sent_lsn       | 0/7478CC0
write_lsn      | 0/7478CC0
flush_lsn      | 0/7478CC0
replay_lsn     | 0/7478CC0
write_lag      |
flush_lag      |
replay_lag     |
sync_priority  | 0
sync_state     | async
reply_time     | 2023-06-13 11:03:59.378867+00
```

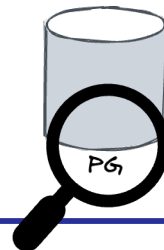


```
patronictl -c /etc/patroni1.yml list
```

```
+ Cluster: patroni15 -----+-----+-----+-----+
| Member | Host           | Role   | State  | TL | Lag in MB |
+-----+-----+-----+-----+-----+
| pg1    | localhost:5432 | Leader | running | 1 |           |
| pg2    | localhost:5433 | Replica | running | 1 |           0 |
| pg3    | localhost:5435 | Replica | running | 1 |           0 |
+-----+-----+-----+-----+-----+
```

```
patronictl -c /etc/patroni1.yml topology
```

```
+ Cluster: patroni15 -----+-----+-----+-----+
| Member | Host           | Role   | State  | TL | Lag in MB |
+-----+-----+-----+-----+-----+
| pg1    | localhost:5432 | Leader | running | 1 |           |
| + pg2  | localhost:5433 | Replica | running | 1 |           0 |
| + pg3  | localhost:5435 | Replica | running | 1 |           0 |
+-----+-----+-----+-----+-----+
```



```
patronictl -c /etc/patroni1.yml switchover
```

```
Current cluster topology
```

```
+ Cluster: patroni15 -----+
| Member | Host           | Role   | State  | TL | Lag in MB |
+-----+-----+-----+-----+-----+-----+
| pg1    | localhost:5432 | Leader | running | 3 |           |
| pg2    | localhost:5433 | Replica | running | 3 |           0 |
| pg3    | localhost:5435 | Replica | running | 3 |           0 |
+-----+-----+-----+-----+-----+-----+
```

```
Primary [pg1]:
```

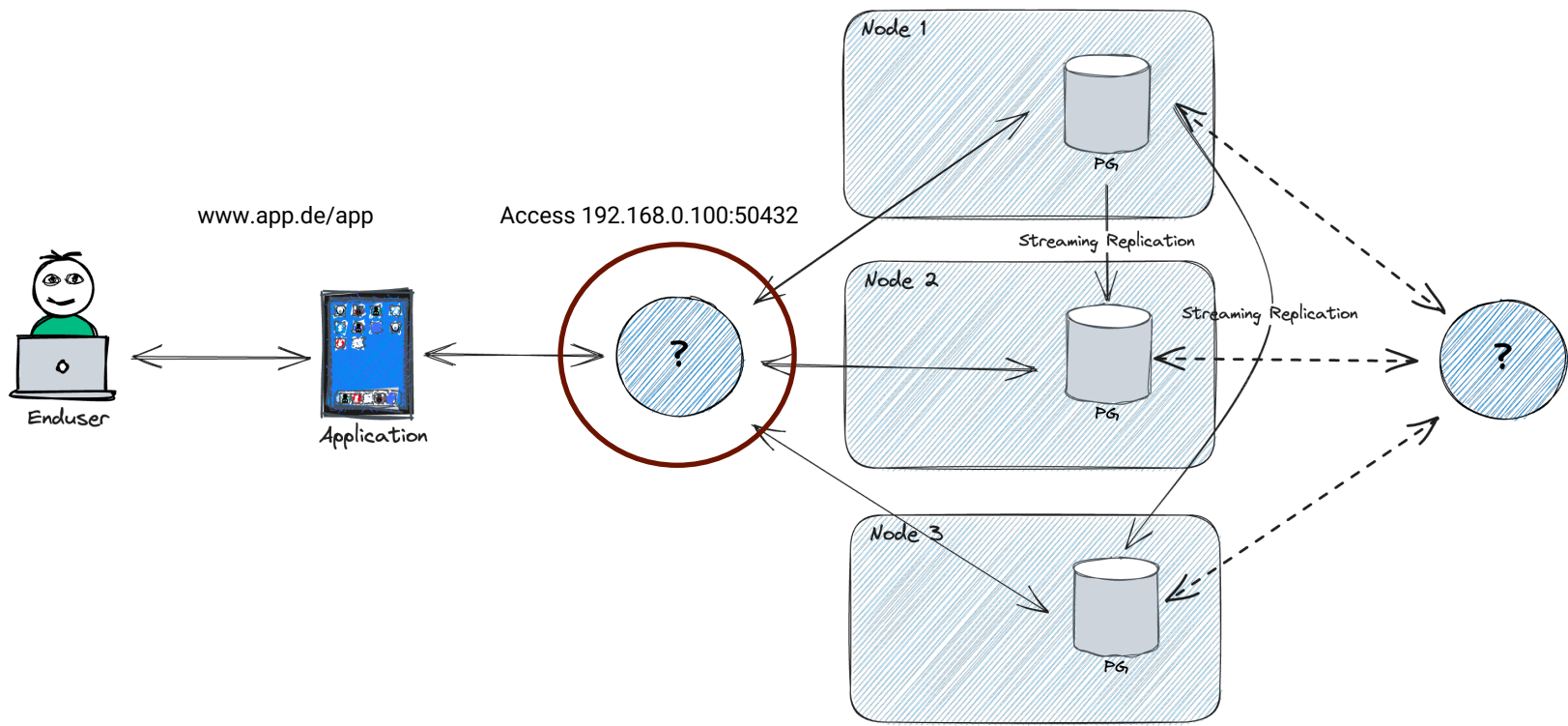
```
Candidate ['pg2', 'pg3'] []: pg3
```

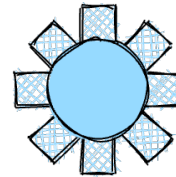
```
When should the switchover take place (e.g. 2023-06-14T13:28 ) [now]:
```

```
Are you sure you want to switchover cluster patroni15, demoting current leader pg1? [y/N]: y
```

```
patronictl -c /etc/patroni1.yml history
```

```
+-----+-----+-----+-----+-----+-----+
| TL | LSN | Reason | Timestamp | New Leader |
+-----+-----+-----+-----+-----+-----+
| 1 | 122129952 | no recovery target specified | 2023-06-14T12:25:55.320948+00:00 | pg2 |
| 2 | 122131304 | no recovery target specified | 2023-06-14T12:27:48.816808+00:00 | pg1 |
| 3 | 122131936 | no recovery target specified | 2023-06-14T12:28:52.652736+00:00 | pg3 |
+-----+-----+-----+-----+-----+-----+
```



HAProxy

haproxy.cfg

```
listen stats
  mode http
  bind *:7000
  stats enable
  stats uri /
```

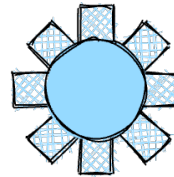
```
listen primary
  bind *:5000
  option httpchk OPTIONS /master
  http-check expect status 200
  default-server inter 3s fall 3 rise 2 on-marked-down shutdown-sessions
  server pg1 localhost:5432 maxconn 100 check port 8008
  server pg2 localhost:5433 maxconn 100 check port 8009
  server pg3 localhost:5435 maxconn 100 check port 8010
```

checking for patroni „role“

access via patroni restapi

restapi:

```
listen: localhost:8008
connect_address: localhost:8008
```



HAProxy

HAProxy version 2.4.22-0ubuntu0.22.04.1, released 2023/03/22

Statistics Report for pid 72853

> General process information

pid = 72853 (process #1, nbproc = 1, nbthread = 2)
 uptime = 5d 22h02m36s
 system limits: memmax = unlimited; ulimit-n = 253
 maxsock = 253; maxconn = 100; maxpipes = 0
 current conns = 1; current pipes = 0/0; conn rate = 1/sec; bit rate = 0.000 kbps
 Running tasks: 0/20; idle = 100 %

■ active UP
■ active UP, going down
■ active DOWN, going up
■ active or backup DOWN
■ active or backup DOWN for maintenance (MAINT)
■ active or backup SOFT STOPPED for maintenance
■ not checked
■ backup UP
■ backup UP, going down
■ backup DOWN, going up
■ not checked
 Note: "NOLB"/"DRAIN" = UP with load-balancing disabled.

Display option:

- Scope:
- Hide DOWN servers
- Refresh now
- CSV export
- JSON export (schema)

External resources:

- [Primary site](#)
- [Updates \(v2.4\)](#)
- [Online manual](#)

	Queue			Session rate			Sessions					Bytes				Denied		Errors		Warnings		Server									
	Cur	Max	Limit	Cur	Max	Limit	Cur	Max	Limit	Total	LbTot	Last	In	Out	Req	Resp	Req	Conn	Resp	Retr	Redis	Status	LastChk	Wght	Act	Bck	Chk	Dwn	Dwntme	Thrtle	
Frontend				1	2	-	1	3	100	11			992	53 491	0	0	8					OPEN									
Backend	0	0		0	0		0	0	10	0	0s		992	53 491	0	0			0	0	0	5d22h UP		0/0	0	0			0		

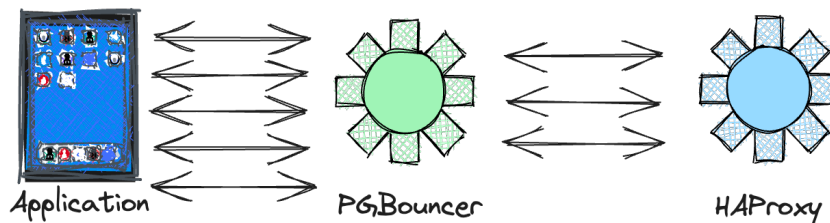
	Queue			Session rate			Sessions					Bytes				Denied		Errors		Warnings		Server									
	Cur	Max	Limit	Cur	Max	Limit	Cur	Max	Limit	Total	LbTot	Last	In	Out	Req	Resp	Req	Conn	Resp	Retr	Redis	Status	LastChk	Wght	Act	Bck	Chk	Dwn	Dwntme	Thrtle	
Frontend				0	1	-	0	1	100	1			232	488	0	0	0	0	0	0	0	OPEN									
pg1	0	0	-	0	0	-	0	0	100	0	0	?	0	0	0	0	0	0	0	0	0	5d22h DOWN	L7STS/503 in 1ms	1/1	Y	-	1	1	5d22h	-	
pg2	0	0	-	0	0	-	0	0	100	0	0	?	0	0	0	0	0	0	0	0	0	5d22h DOWN	L7STS/503 in 3ms	1/1	Y	-	1	1	5d22h	-	
pg3	0	0	-	0	1	-	0	1	100	1	1	5d22h	232	488	0	0	0	0	0	0	0	5d22h UP	L7OK/200 in 4ms	1/1	Y	-	0	0	0s	-	
Backend	0	0		0	1		0	1	10	1	1	5d22h	232	488	0	0			0	0	0	5d22h UP		1/1	1	0		0	0s		

	Queue			Session rate			Sessions					Bytes				Denied		Errors		Warnings		Server									
	Cur	Max	Limit	Cur	Max	Limit	Cur	Max	Limit	Total	LbTot	Last	In	Out	Req	Resp	Req	Conn	Resp	Retr	Redis	Status	LastChk	Wght	Act	Bck	Chk	Dwn	Dwntme	Thrtle	
Frontend				0	0	-	0	0	100	0			0	0	0	0	0	0	0	0	0	OPEN									
pg1	0	0	-	0	0	-	0	0	100	0	0	?	0	0	0	0	0	0	0	0	0	5d22h UP	L7OK/200 in 3ms	1/1	Y	-	0	0	0s	-	
pg2	0	0	-	0	0	-	0	0	100	0	0	?	0	0	0	0	0	0	0	0	0	5d22h UP	L7OK/200 in 2ms	1/1	Y	-	0	0	0s	-	
pg3	0	0	-	0	0	-	0	0	100	0	0	?	0	0	0	0	0	0	0	0	0	5d22h DOWN	L7STS/503 in 2ms	1/1	Y	-	1	1	5d22h	-	
Backend	0	0		0	0		0	0	10	0	0	?	0	0	0	0			0	0	0	5d22h UP		2/2	2	0		0	0s		

PGBouncer

PGBouncer

- Lightweight connection pooler
- Maintains a pool of connections for each unique user <-> database pair
- You won't lose your connection when the backend changed

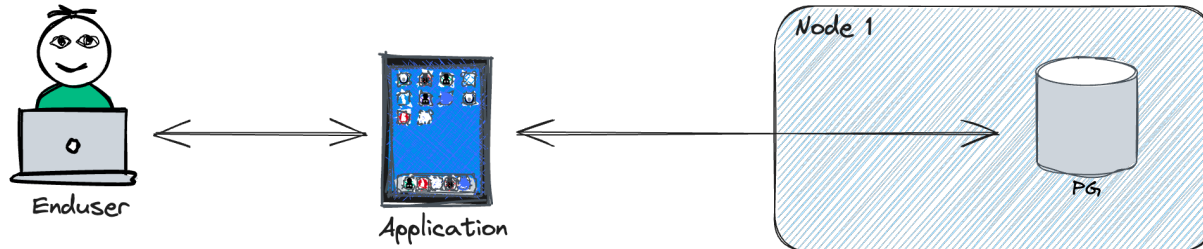


Demo

Conclusion and Final Words

www.app.de/app

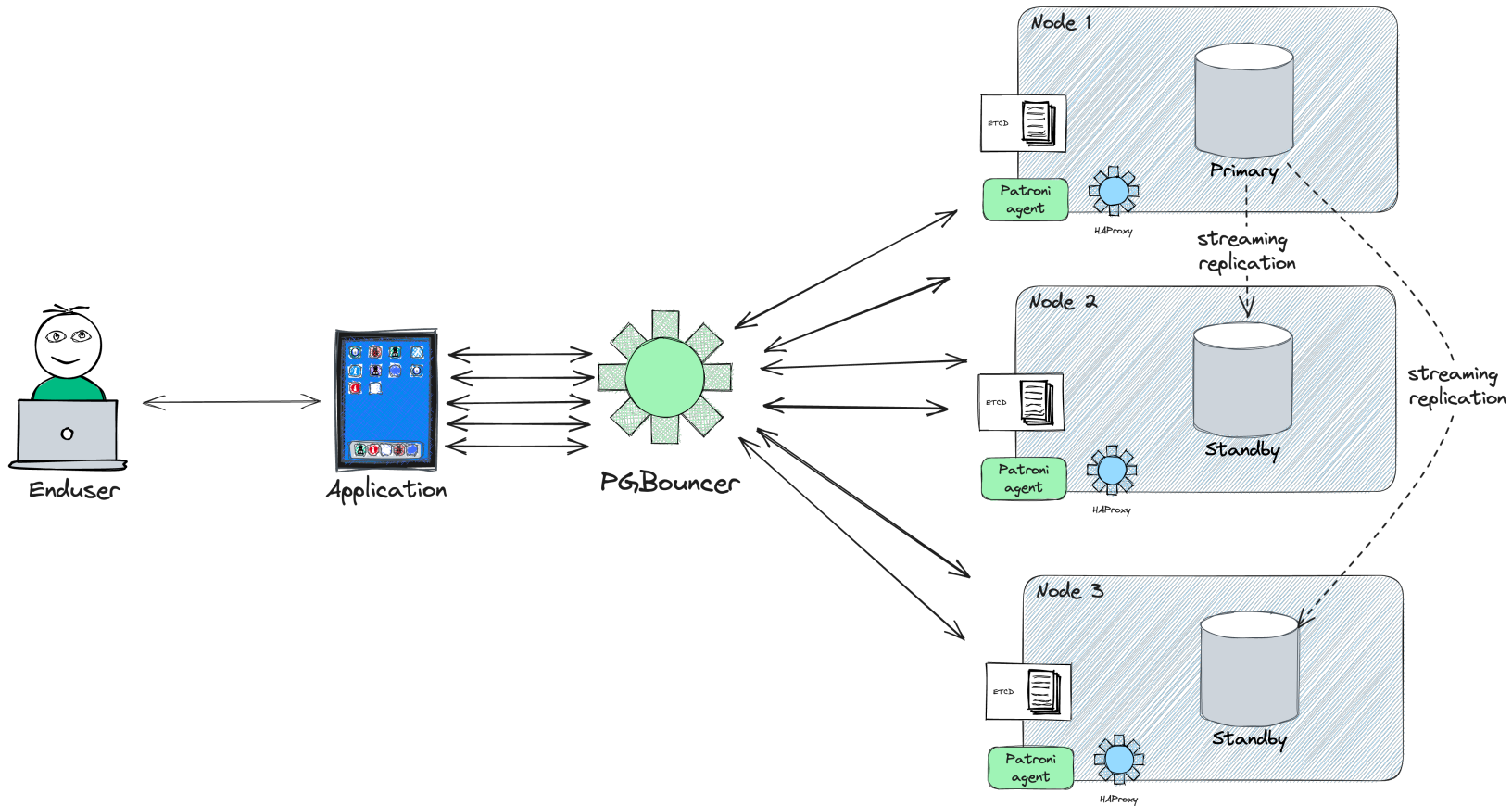
Access 192.168.0.101:5432

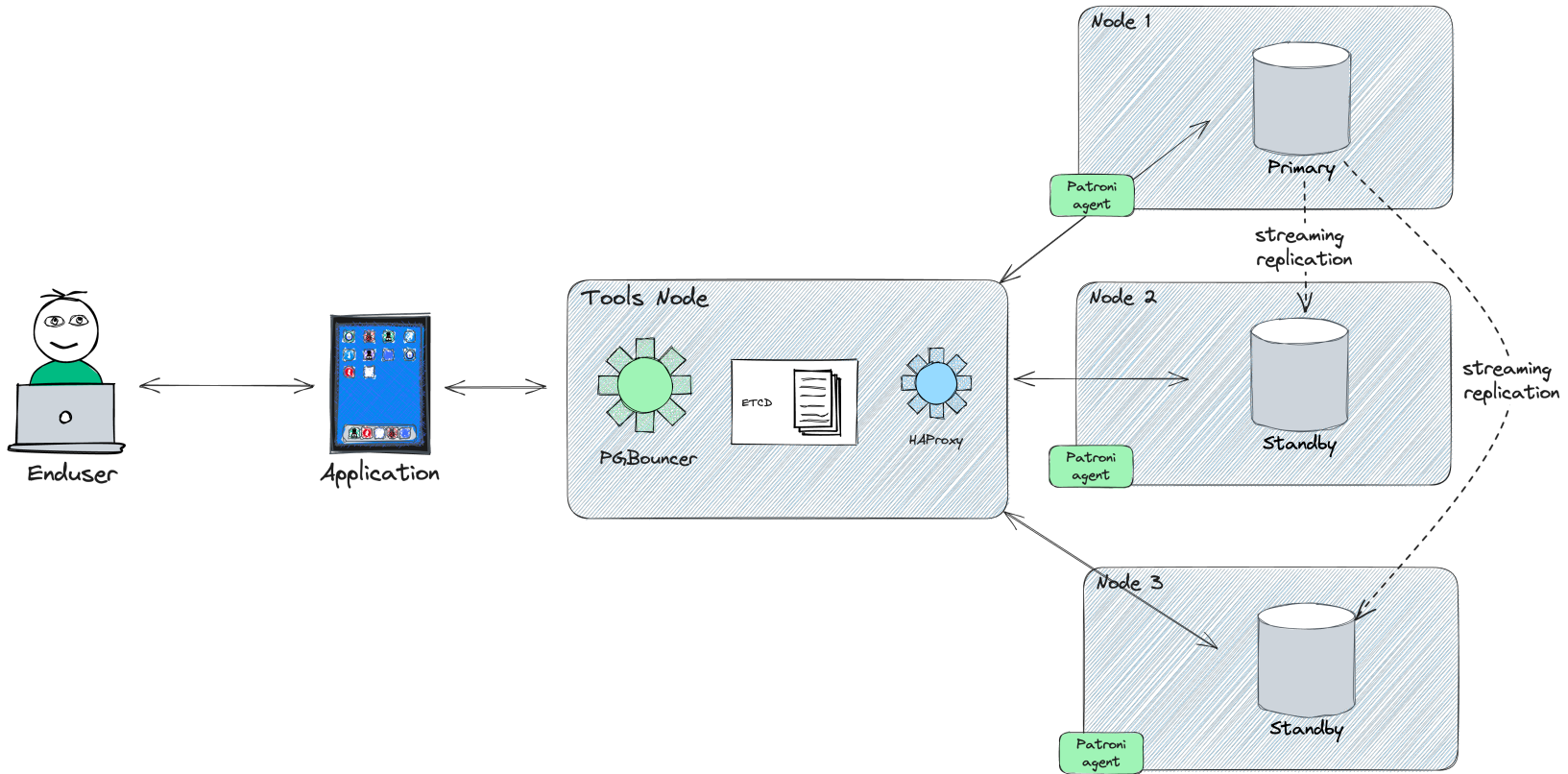


Internal Server Error

500

The server encountered an internal error or misconfiguration
and was unable to complete your request





ENTER TO WIN

Take part in our prize draw to win a collectible LEGO® Star Wars™ set with EDB.



Scan the QR code and fill in the form to enter.



EDB™

THANK YOU!