PGConf.Russia 2023

PostgresPro

Кластер Corosync-Pacemaker
Работа над ошибками

Игорь Косенков
Постгрес Профессиональный
инженер

PostgresPro

Последствия неправильной настройки кластера

- ✓ Реплика не стартует
- ✓ Длительное восстановление отказавшего узла
- ✓ Отказ в обслуживании
- ✓ Split-brain

PostgresPro

Последствия неправильной настройки кластера

- ✓ Реплика не стартует
- ✓ Длительное восстановление бывшего мастера
- ✓ Отказ в обслуживании
- ✓ Split-brain

Реплика не стартует

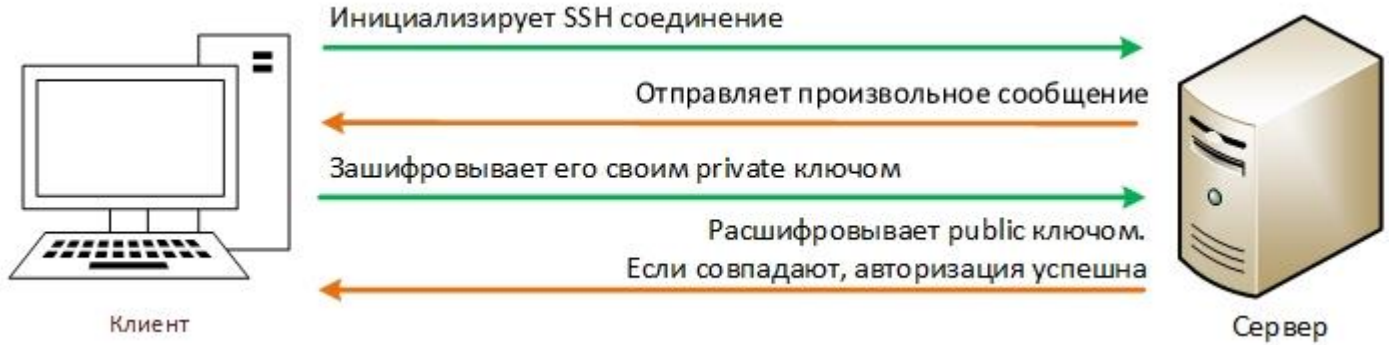
archive_command = “scp -i key %p user@host:/archive-wals/%f”

restore_command = “scp -i key user@host:/archive-wals/%f %p”

PostgresPro

Реплика не стартует

Сессия ssh/scp




```
$ scp –v user@host:/archive-wals/000000010000000000000002 ./

…
Transferred: sent 4904, received 16800044 bytes, in 0.5 seconds
Bytes per second: sent 24499.6, received 83930261.6
Exit status 0
…
```

config pgsql

Реплика не стартует

PostgresPro

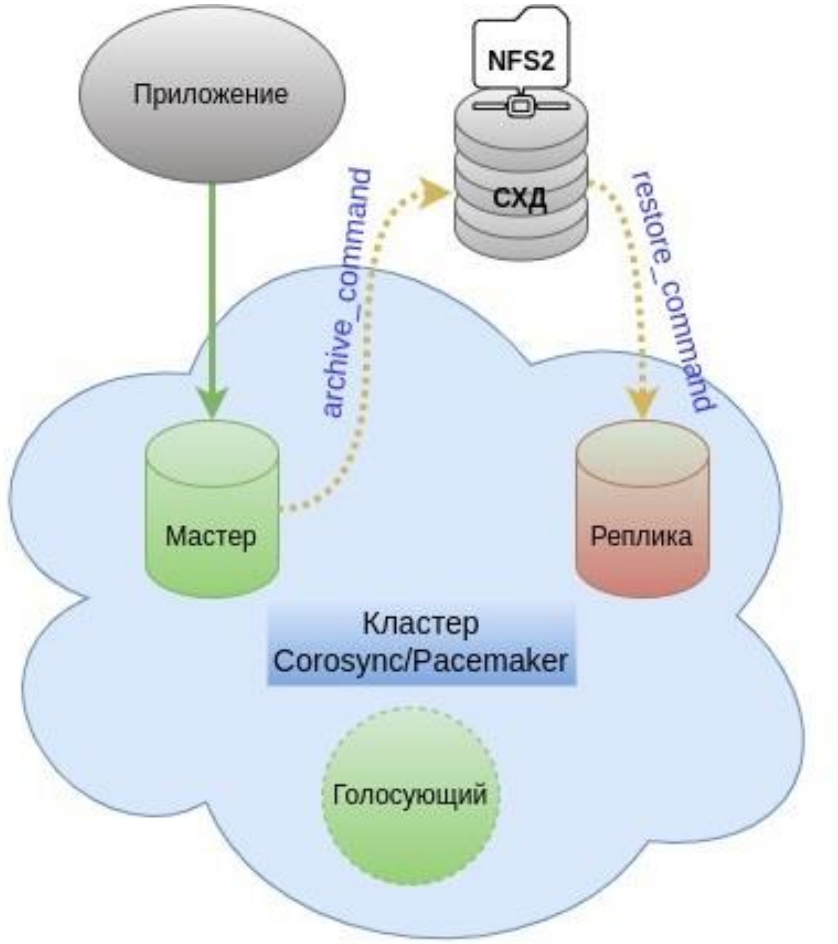
…
Resource: pgsql (class=ocf provider=heartbeat type=pgsql)
 Operations: **start interval=0s timeout=120s (pgsql-start-interval-0s)**
…

При wal > 240шт
реплика не запустится

Использование NFS или CIFS

Сессия rsync

Реплика не стартует



$ mount.cifs //host/archive-wals /archive-wals -o user=user

$ rsync --stats /archive-wals/000000010000000000000002 ./

File list transfer time: 0,001 seconds

Реплика не стартует

PostgresPro

config pgsql

```
Resource: pgsql (class=ocf provider=heartbeat type=pgsql)
 Operations: demote interval=0s timeout=120s (pgsql-demote-interval-0s)
             methods interval=0s timeout=5s (pgsql-methods-interval-0s)
             monitor interval=30s timeout=30s (pgsql-monitor-interval-30s)
             monitor interval=29s role=Master timeout=30s (pgsql-monitor-interval-29s)
             notify interval=0s timeout=90s (pgsql-notify-interval-0s)
             promote interval=0s timeout=120s (pgsql-promote-interval-0s)
             start interval=0s timeout=120s (pgsql-start-interval-0s)
             stop interval=0s timeout=120s (pgsql-stop-interval-0s)
```

Тюнинг start’а ресурса pgsql

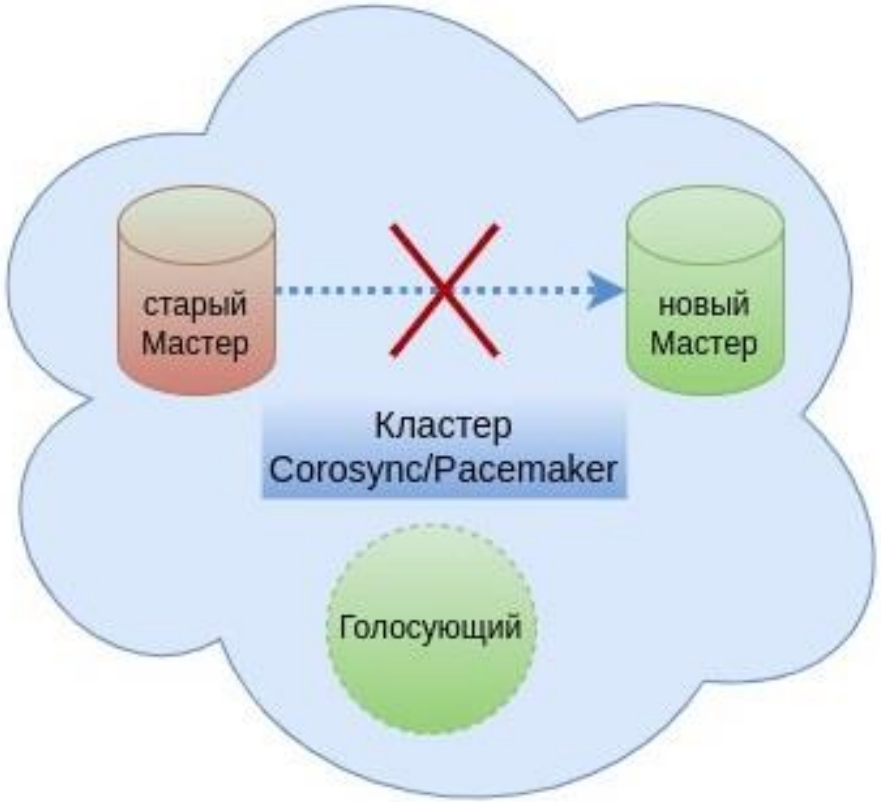
```
pcs resource update pgsql op start timeout=300s
```

PostgresPro

Последствия неправильной настройки кластера

- ✓ Реплика не стартует
- ✓ Длительное восстановление бывшего мастера
- ✓ Отказ в обслуживании
- ✓ Split-brain

Длительное восстановление бывшего мастера



config pgsql:

Resource: pgsql (class=ocf provider=heartbeat type=pgsql)

Attributes: …restart_on_promote=true…

Длительное восстановление бывшего мастера

restart_on_promote=true или false???

значение	метод превращения	pg_rewind	возврат узла	Failover
true	rm standby.signal pg_ctl restart	нет	медленный	медленный
false	pg_ctl promote	да	быстрый	быстрый

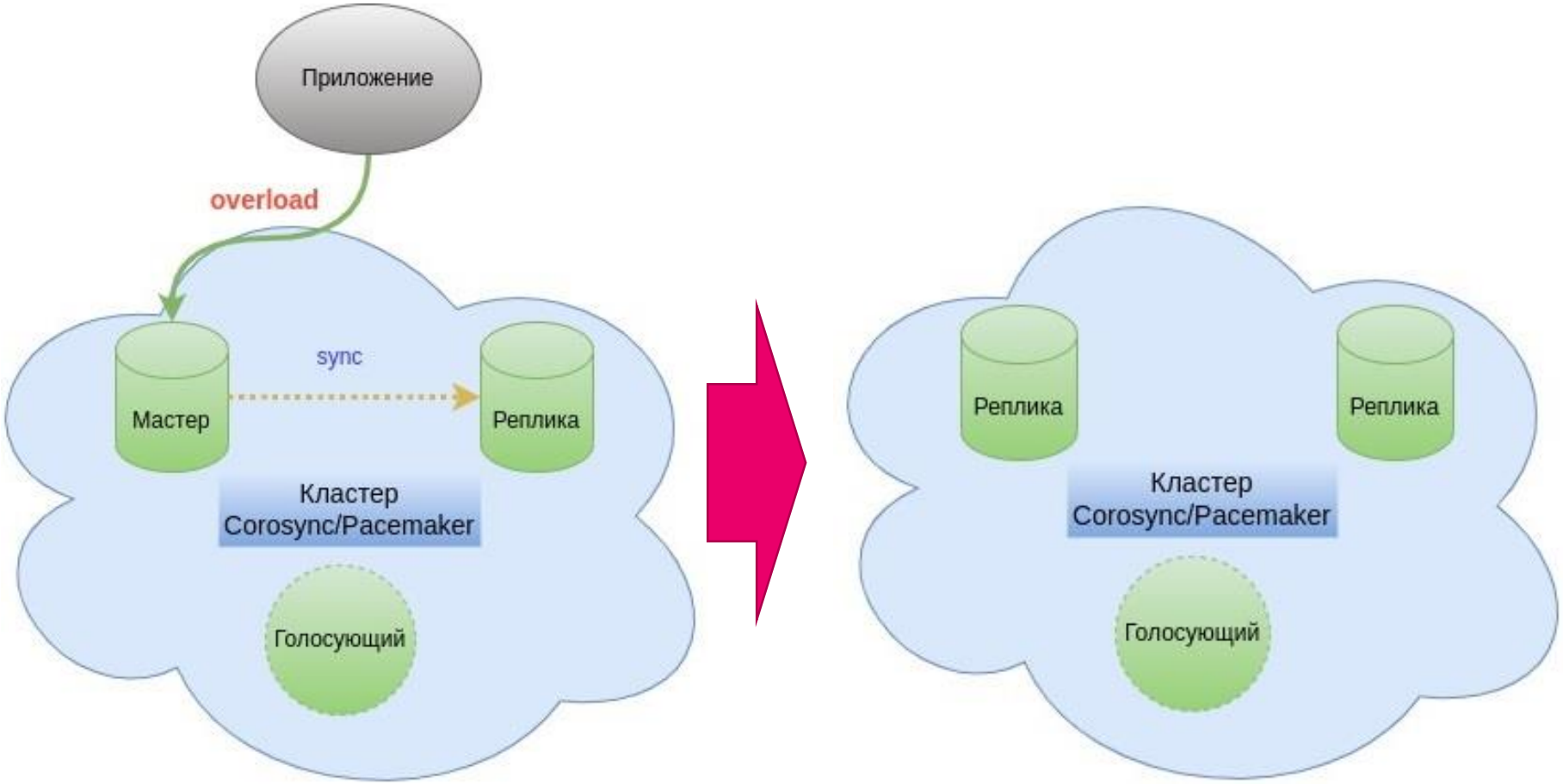
PostgresPro

Последствия неправильной настройки кластера

- ✓ Реплика не стартует
- ✓ Длительное восстановление бывшего мастера
- ✓ Отказ в обслуживании
- ✓ Split-brain

Кластер без фенсинга

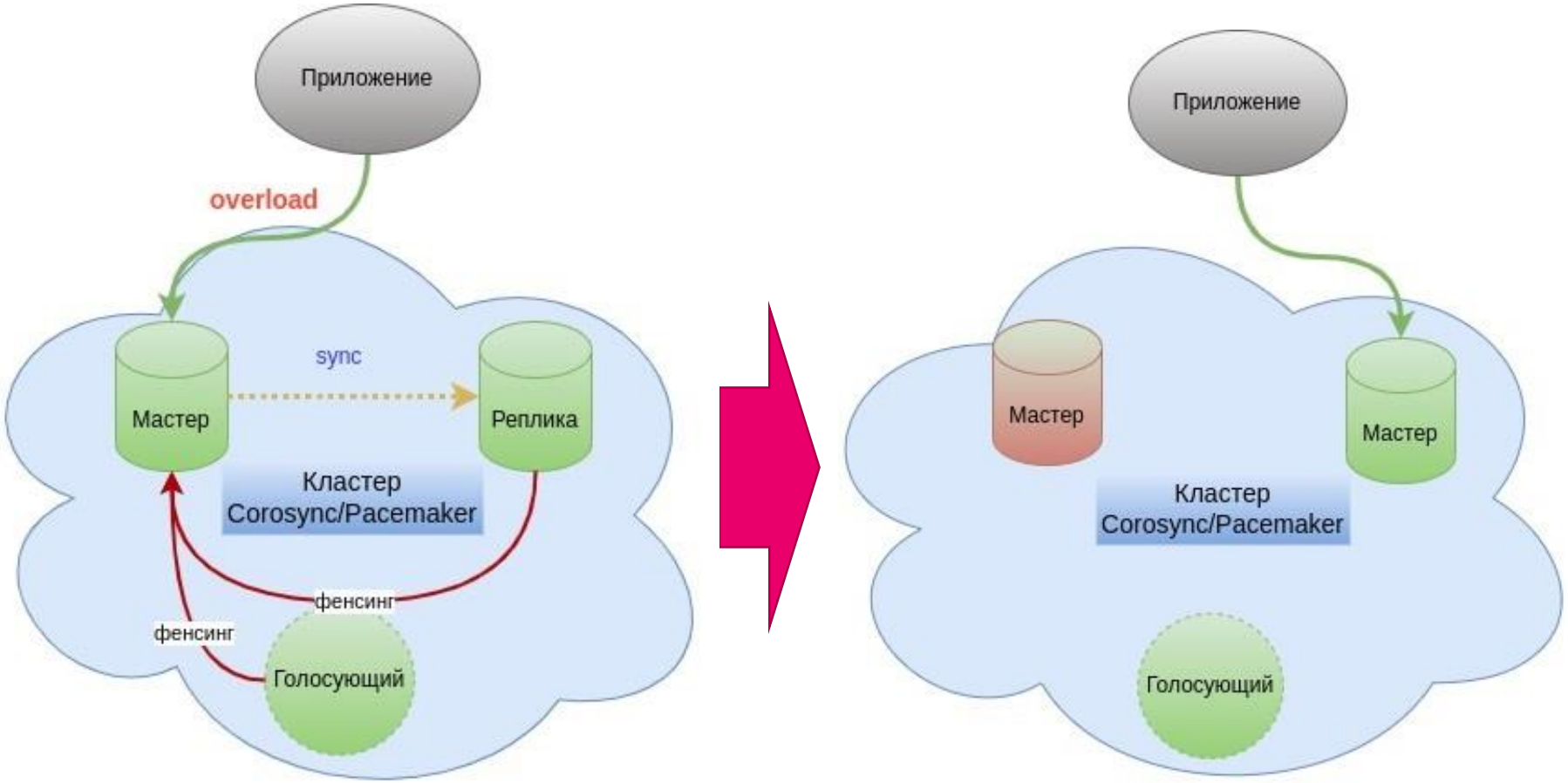
Отказ в обсуживании



1. Кластер попытается остановить зависший мастер
2. Неудачные попытки и таймаут переведут ресурс в FAILED
3. Промоута реплики не произойдет

Кластер с фенсингом

Отказ в обсуживании



1. Кластер отправит узел в «жесткую» перезагрузку (hard reset)
2. После успешной перезагрузки будет промоут реплики
3. Сервис СУБД продолжит функционировать
4. Ситуация split-brain исключена

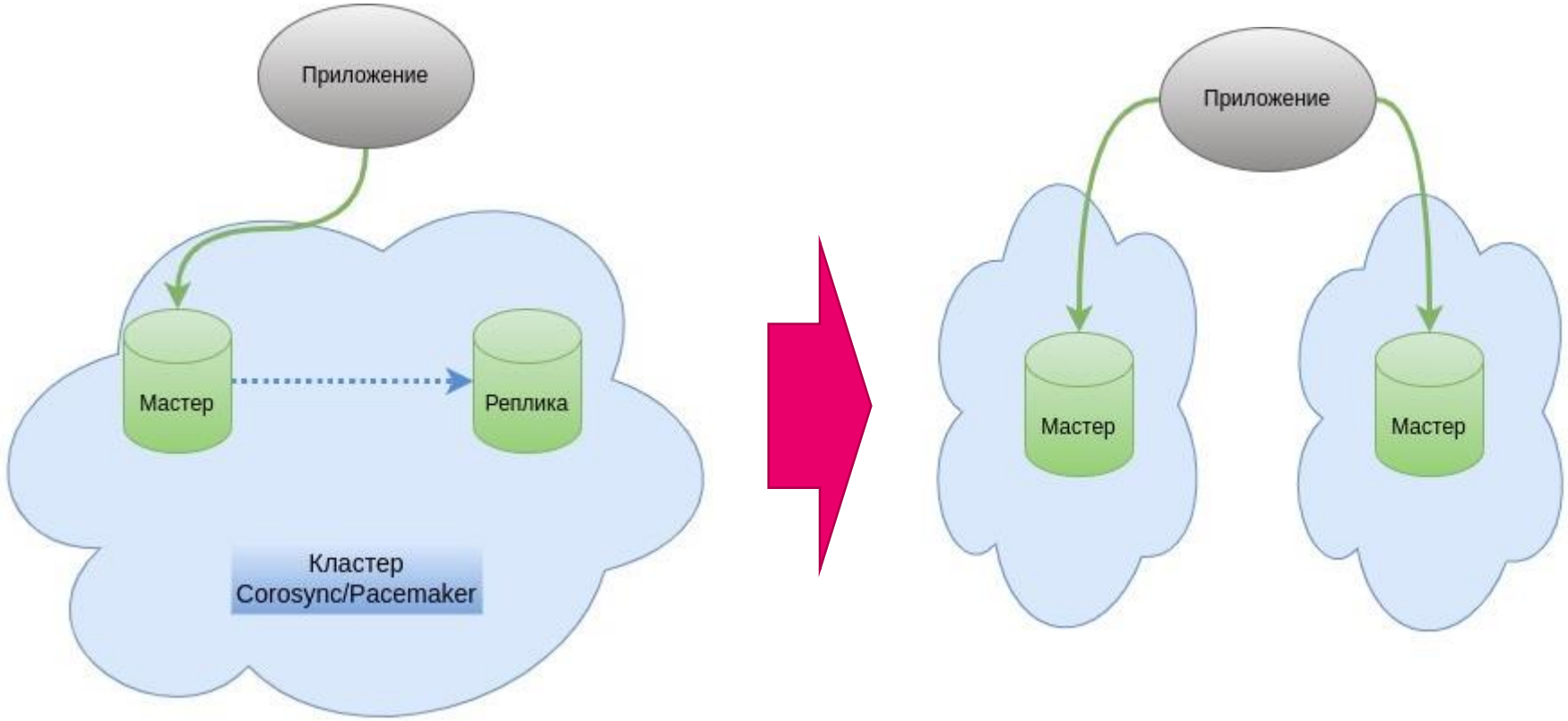
PostgresPro

Последствия неправильной настройки кластера

- ✓ Реплика не стартует
- ✓ Длительное восстановление бывшего мастера
- ✓ Отказ в обслуживании
- ✓ Split-brain

Кластер без фенсинга и кворума

SPLIT-BRAIN

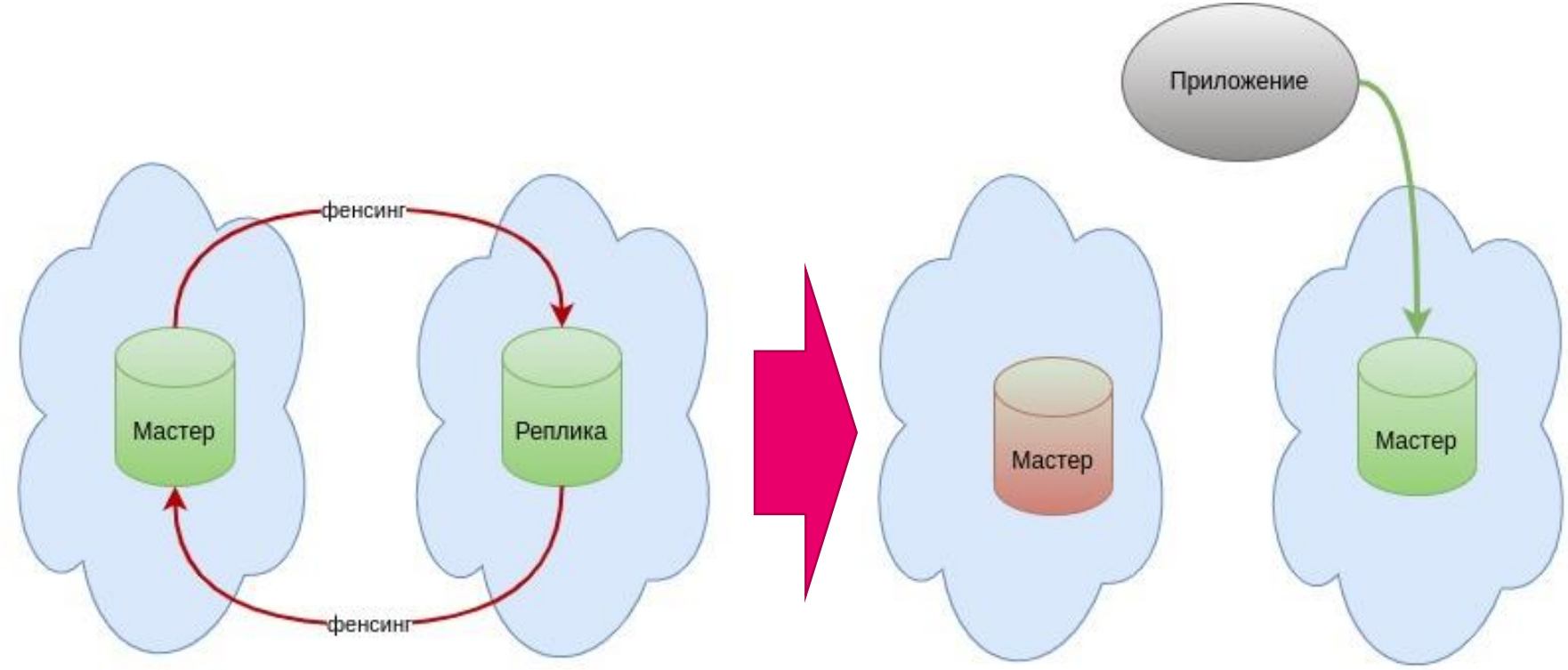


1. Сетевое разделение - каждый узел считает себя «живым»
2. Выполняется промоут реплики
3. split-brain: 2 мастера + 2 виртуальных IP

SPLIT-BRAIN

PostgresPro

Кластер с фенсингом



Варианты фенсинга

1. Железо: fence_ipmilan, fence_ilo
2. Виртуальные среды: fence_vmware, fence_virsh
3. Облака: fence_sbd, fence_scsi

PostgresPro

Атрибуты правильного кластера

- ✓ Фенсинг
- ✓ Кворум
- ✓ Архив wal-ов
- ✓ Настроенные таймауты

Q&A

PostgresPro

Спасибо!