

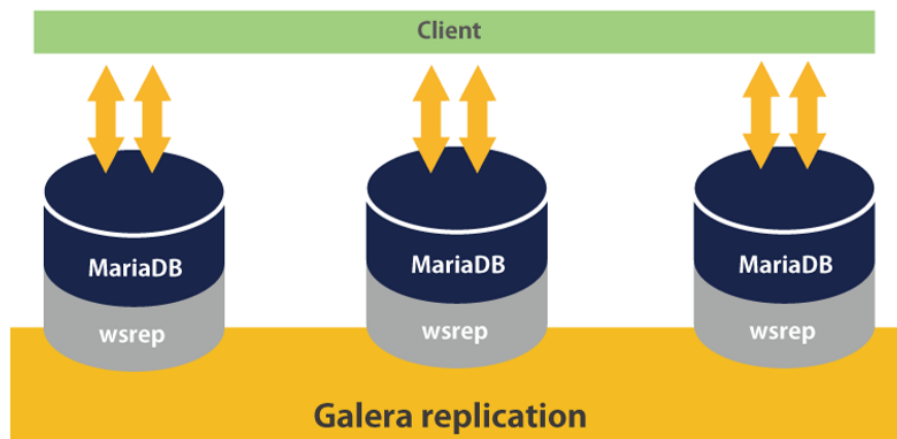


Galeraccluster & Maxscale



Galera Cluster

- **Mysql & MariaDB** 지원
- **Multi Master** : 모든 노드에서 일기 쓰기가 가능합니다.
- **동기적 복제** : 슬레이브 지연이 없고 노드 충돌시에 데이터 손실이 없습니다.
- **일관적인 데이터**: 모든 노드는 같은 상태를 유지합니다.
- **Multi-thread slave** : 어떠한 워크로드에서도 더 나은 성능을 가능하게 합니다.
- **Hot standby** : 장애 복구시 down-time이 없습니다.
- **read/write split** 이 필요 없습니다
- **InnoDB** 엔진만 완벽히 지원한다.
- 동기적 복제이기에 성능이 낮은 노드에 의해 전체 성능이 결정되게 된다.
- 별도의 vip는 필요없다.
- **binarylog** 로그를 전송하는 전송하고 slave에서 릴레이로그(bin 로그를 전송 받는)를 받는 형식으로 동작한다.
- **비동기** : 다른 노드에서 데이터를 받았는지 확인안하고 바로 응답.
- **반동기**: 한개의 복제본이라도 데이터를 받았는지는 확인함.
- **동기** : 모든노드가 트랜잭션을 받았는지를 확인

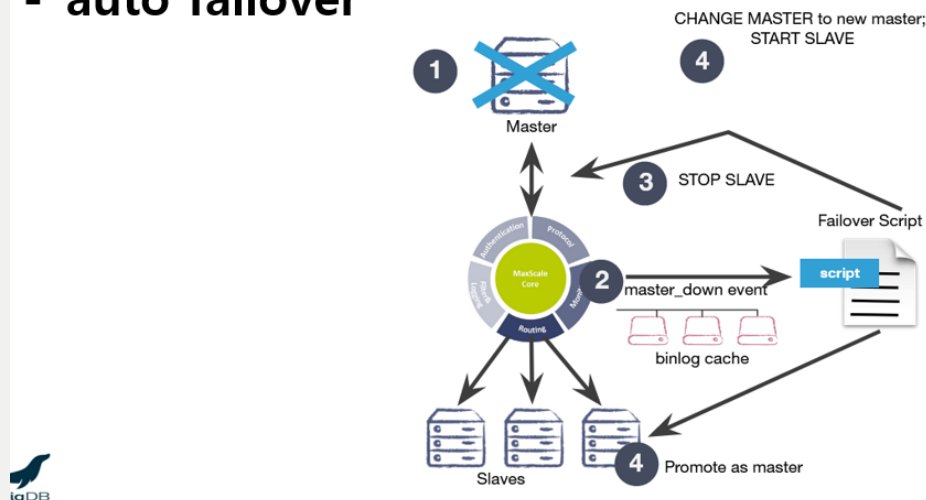




Maxscale

- 로드밸런싱: 각각의 노드에 읽기(Read)와 쓰기(Write) 역할을 나눠 트랜잭션을 분산 처리할 수 있습니다. 이를 통해 시스템의 효율성과 성능을 극대화할 수 있습니다.
- 고가용성: 자동 장애 복구 기능으로 시스템 다운타임을 최소화하고 지속적인 서비스 가용성을 보장
- 운영: 웹 ui로 관리할 수 있기에 유지/관리에 용이하다.
- 설치 install 해야되는 패키지는 두개이다. **Mysql-client , Maxscale**

■ auto failover



설치

설치 (centos7) Galera&Maxscale

갈레라 only

장애 발생시 문의처

[MariaDB]
<https://mariadb.com/contact/>

커뮤니티

[galeracluster]
<https://galeracluster.com/community/>
<https://mariadb.com/kb/en/galera-cluster/>

[Maxscale]
<https://mariadb.com/kb/en/maxscale/>

Maria

[galeracluster]
<https://mariadb.com/kb/en/galera-cluster/>

[maxscale]
<https://mariadb.com/kb/en/maxscale/>

유튜브

[maxscale]
<https://www.youtube.com/watch?v=rhtL0mtIGP0&t=2390s>

[설치] 갈레라 && 맥스스케일
<https://www.youtube.com/watch?v=3bXVdpU7BLM&t=84s>
<https://www.youtube.com/watch?v=mzFKWfsBhtQ&feature=youtu.be>

파일

Maira && galeracluster && Maxscale
<http://naver.me/F1reYRzw>

Mysql
<http://naver.me/xbA73b5M>

강사님 자료:

<http://naver.me/FaffQmuI>

유데미

- 아래 방법으로 한글자막으로 시청 가능합니다.

스페인어 이기에 자막 사용을 위해 아래 url 참고해서 플러그인 추가필요.

<https://wendy308games.tistory.com/20>

<https://www.udemy.com/course/galera-cluster-para-mariadb-y-my>

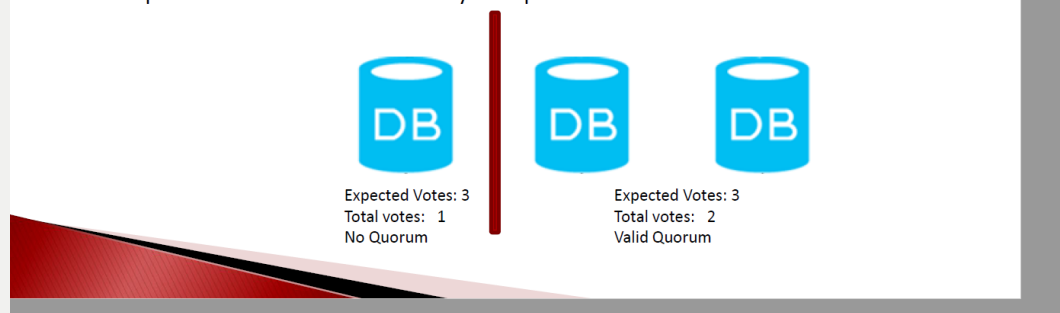
구성



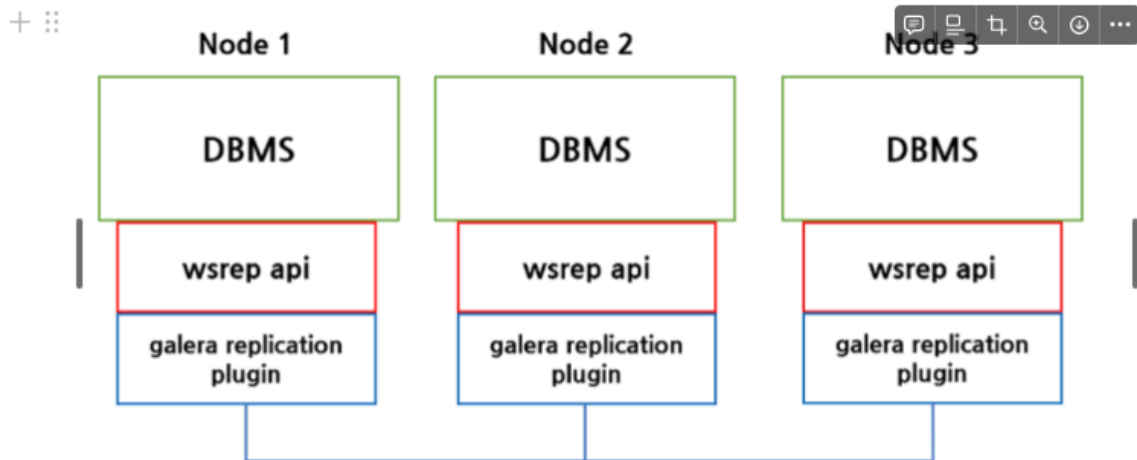
갈레라의 경우 3개의 노드 구성을 권장한다.

- 각 노드는 자신이 우선권이 있다고 판단할때 최종 투표권을 가지고 있는게
- 3번째 노드이다.

- ▶ Galera Cluster takes a quorum vote whenever a node does not respond and is suspected of no longer being a part of the cluster.
- ▶ The component that has quorum alone continues to operate as the Primary Component, while those without quorum enter the non-primary state and begin attempt to connect with the Primary Component.



동기 복제 방식



- Write-Set Replication API : dbms 간 복제를 위한 api
- 각 노드의 업데이트 내용을 임시 gcache에 저장하게 됨.

포트

기본 : 3306

복제 트래픽을 위한 포트 : 4567

ist : 트랜잭션만을 다른 노드에 전송 4568

sst : 상태 그대로를 다른 노드에 전송 4444

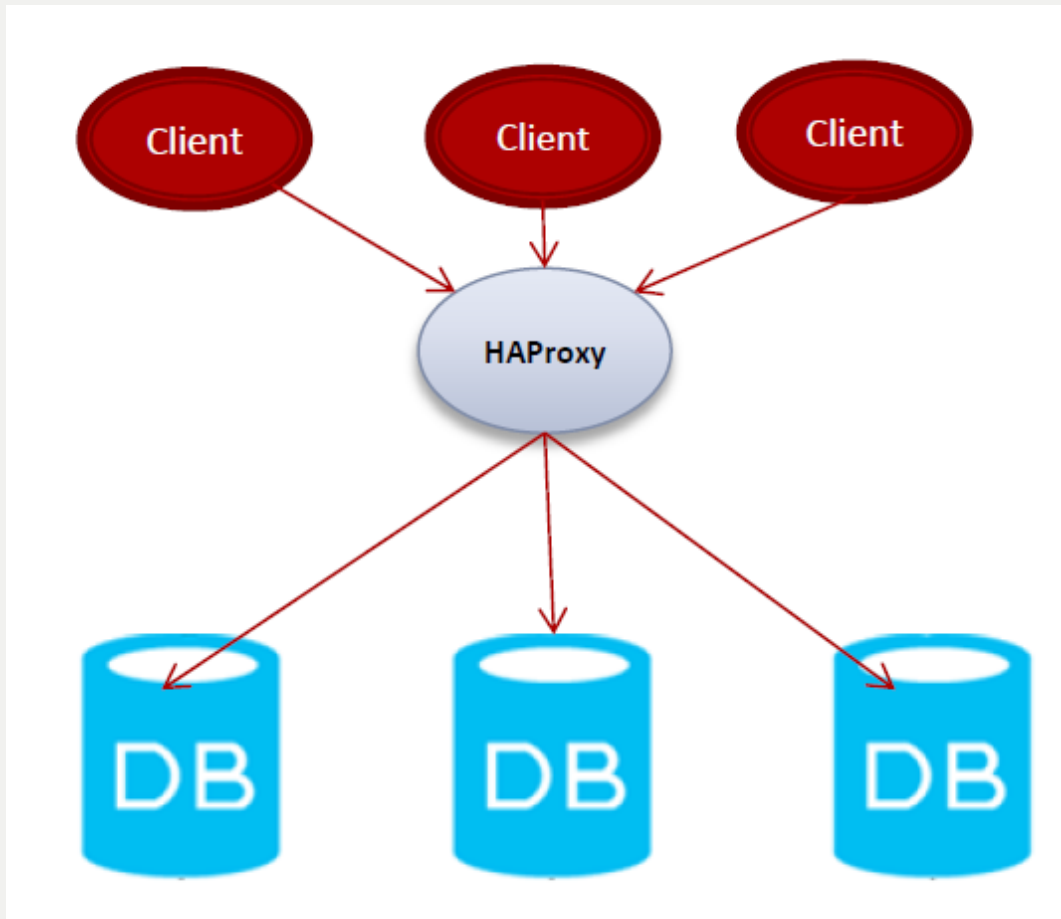
각각의 전송받은 데이터를 Gcache 경로를 지정하고 임시로 지정하게 된다.

로드 밸런싱

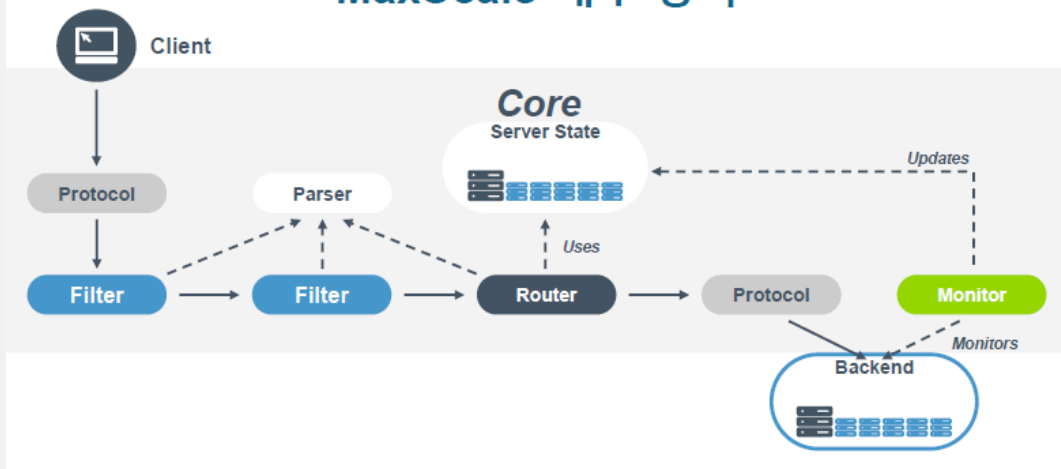


여러 사용자가 요청시 분산 시키는 개념 이미지에서 HAproxy 역할을 하는 것중 하나가 maxscale로 볼 수 있다.

즉 연결되어있는 maxscale이 다운시 연결되어있는 노드들이 연결이 안되기에 maxscale 자체도 이중화로 운영된다.

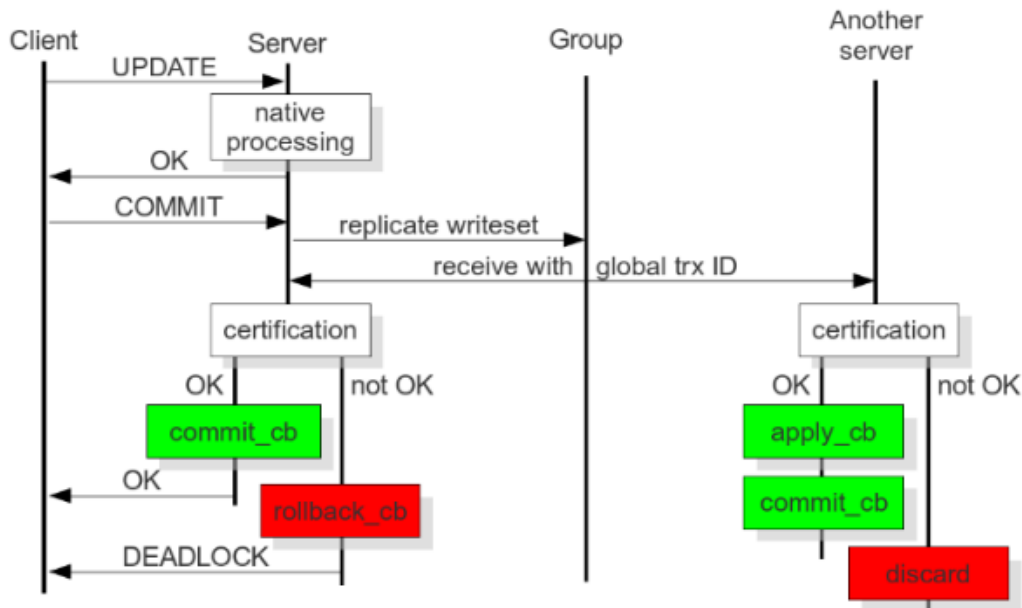


MaxScale 내부 동작



인증기반 복제

- 인증복제



- a\ a 구조라 갈레라에서는 양방향으로 복제가 이루어진다.
- 사용자의 커밋
- 각 노드로 복제
- writeset은 복제할 데이터베이스 행과 트랜잭션 중 보유한 lock에 대한 정보를 담고 있다
- 모든 노드에서 인증을 진행
- 인증에는 wsrep replication plugin(Galera Replication Plugin)과 writeset이 사용됩니다. (libgalera_smm.so)

Write-Set Replication API

- wsrep API - 갈레라 클러스터의 작동 인터페이스를 담당한다.
- wsrep hooks - DB 엔진 안에서 작동하는 wsrep 구조다.

- **Galera provider** - 갈레라 클러스터 라이브러리에서 **wsrep API** 가 작동하도록 한다.
- **certification** - DB 내에서 일련의 쓰기 세트 및 이에 대한 검증 수행을 담당한다.
- **replication** - 복제 프로토콜 및 클러스터에 대한 통합 명령 기능을 제공한다.
- **GCS framework** - 클러스터 그룹간의 통신 시스템에 대한 플러그인 구조를 제공한다.
(갈레라 클러스터는 **vsbes**, **gemini** 를 **base** 로 한 프레임워크를 수용했으나 이 이외에 다른 프레임워크 형식을 사용 할 수 있다.)

Gcache

갈레라 에서 데이터를 전송받는 방법은 여러가지이지만

- 행 단위로 복제하는 방식과
- 메모리에 넣는 방식이 있다.
- GCACHE 방식이 있다. 기본적으로는 128M을 지원한다.

```
wsrep_provider_options="gcache.size=512M; gcache.name=/tmp/ga
```

파라미터

파라미터 정리

노드 3개 이유

- 갈레라 클러스터에서는 최소 3개의 인스턴스를 권장한다.
- 홀수개의 노드를 두어 스플릿 브레인 상황을 방지하기 위함이다.

Maxscale

Maxscale

크래쉬 상황

```
# GALERA saved state
version: 2.1
uuid: 9bd8af8a-68f5-11ef-a77e-a3ce94b1771e
seqno: 28
safe_to_bootstrap: 0
```

- seqno 순이 높은 게 가장 최신의 데이터를 가지고 있는 것이다,.
- 그러므로 safe_to_bootstrap 레벨을 높여 기준점이 되게 하여
- 데이터베이스를 다시 올려야 한다.
- 기본적으로는 1로 활성화 시킨다.

시나리오

- 1: 3개의 노드가 전부죽어 가장 seqno가 높은 노드를 토대로 새로운 클러스터로 기동
- 2: 1개의 노드가 죽은 상태에서 트랜잭션에 클러스터 그룹에서 트랜잭션 정보를 얻어 올림.
- 3: 장애 조치 - Galera Cluster pk없는 테이블 삭제로 인한 동기화 전송 지연으로 노드 분리후 다시 갈레라 클러스터 연결

시나리오 1

시나리오 2

시나리오 3