Red Hat Ceph Storage 5

구성 가이드

Red Hat Ceph Storage 구성 설정

Last Updated: 2022-08-14

Red Hat Ceph Storage 5 구성 가이드

Red Hat Ceph Storage 구성 설정

Enter your first name here. Enter your surname here.
Enter your organisation's name here. Enter your organisational division here.
Enter your email address here.

법적 공지

초록

이 문서에서는 부팅 시 및 런타임 시 Red Hat Ceph Storage를 구성하는 방법을 설명합니다. 또한 구성 참조 정보를 제공합니다. Red Hat은 코드, 문서, 웹 속성에서 문제가 있는 용어를 교체하기 위해 최선을 다하고 있습니다. 먼저 마스터(master), 슬레이브(slave), 블랙리스트(blacklist), 화이트리스트(whitelist) 등 네 가지 용어를 교체하고 있습니다. 이러한 변경 작업은 작업 범위가 크므로 향후 여러 릴리스에 걸쳐 점차 구현할 예정입니다. 자세한 내용은 CTO Chris Wright의 메시지에서 참조하십시오.

차례

1장. CEPH 구성의 기본 사항

스토리지 관리자는 Ceph 구성을 보는 방법과 Red Hat Ceph Storage 클러스터에 대한 Ceph 구성 옵션을 설정하는 방법에 대해 기본 지식이 있어야 합니다. 런타임에 Ceph 구성 옵션을 보고 설정할 수 있습니다.

1.1. 사전 요구 사항

- Red Hat Ceph Storage 소프트웨어 설치.

1.2. CEPH 구성

모든 Red Hat Ceph Storage 클러스터에는 다음을 정의하는 구성이 있습니다.

- 클러스터 ID
- 인증 설정
- Ceph 데몬
- 네트워크 설정
- 노드 이름 및 주소
- 인증 키 경로
- OSD 로그 파일의 경로
- 기타 런타임 옵션

cephadm 과 같은 배포 도구는 일반적으로 사용자를 위한 초기 Ceph 구성 파일을 만듭니다. 그러나 배포 툴을 사용하지 않고 Red Hat Ceph Storage 클러스터를 부트스트랩하려는 경우 직접 만들 수 있습니다.

추가 리소스

- **cephadm** 및 Ceph 오케스트레이터에 대한 자세한 내용은 *Red Hat Ceph Storage 운영 가이드를 참조하십시오*.

1.3. CEPH 구성 데이터베이스

Ceph Monitor는 전체 스토리지 클러스터에 대한 구성 옵션을 저장하여 구성 관리를 중앙 집중화하는 Ceph 옵션의 구성 데이터베이스를 관리합니다. 데이터베이스에서 Ceph 구성을 중앙 집중화함으로써 스토리지 클러스터 관리를 단순화합니다.

Ceph에서 옵션을 설정하는 데 사용하는 우선순위 순서는 다음과 같습니다.

- 컴파일된 기본값
- Ceph 클러스터 구성 데이터베이스
- 로컬 **ceph.conf** 파일
- **ceph 데몬** ***DAEMON-NAME*** **구성 세트 또는 ceph tell** ***DAEMON-NAME*** **injectargs** 명령을 사용하여 런타임 덮어쓰기

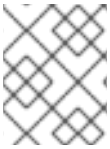
여전히 로컬 Ceph 구성 파일에 정의할 수 있는 몇 가지 Ceph 옵션은 기본적으로 **/etc/ceph/ceph.conf** 입니다. 그러나 **ceph.conf** 는 Red Hat Ceph Storage 5에서 더 이상 사용되지 않습니다.

cephadm 은 Ceph 모니터 연결, 인증 및 구성 정보 가져오기를 위한 최소 옵션 세트만 포함하는 기본 **ceph.conf** 파일을 사용합니다. 대부분의 경우 **cephadm** 은 **mon_host** 옵션만 사용합니다. **mon_host** 옵션에 대해서만 **ceph.conf** 를 사용하지 않으려면 DNS SRV 레코드를 사용하여 모니터를 사용하여 작업을 수행합니다.

중요

Red Hat은 assimilate **-conf** 관리 명령을 사용하여 **ceph.conf** 파일에서 유효한 옵션을 구성 데이터베이스로 이동하는 것이 좋습니다. assimilate **-conf에 대한** 자세한 내용은 관리 명령을 참조하십시오.

Ceph를 사용하면 런타임에 데몬 구성을 변경할 수 있습니다. 이 기능은 디버그 설정을 활성화 또는 비활성화하여 로깅 출력을 늘리거나 줄이는 데 유용할 수 있으며 런타임 최적화에도 사용할 수 있습니다.



참고

동일한 옵션이 구성 데이터베이스와 Ceph 구성 파일에 있는 경우 구성 데이터베이스 옵션은 Ceph 구성 파일에 설정된 옵션보다 우선 순위가 낮습니다.

섹션 및 마스크

데몬 유형별로 Ceph 옵션을 전역적으로 구성하거나 Ceph 구성 파일의 특정 데몬을 구성할 수 있으므로 다음 섹션에 따라 구성 데이터베이스에서 Ceph 옵션을 구성할 수도 있습니다.

섹션	설명
global	모든 데몬 및 클라이언트에 영향을 미칩니다.
mon	모든 Ceph 모니터에 영향을 미칩니다.
mgr	모든 Ceph Manager에 영향을 미칩니다.
osd	모든 Ceph OSD에 영향을 미칩니다.
mds	모든 Ceph 메타데이터 서버에 영향을 미칩니다.
클라이언트	마운트된 파일 시스템, 블록 장치, RADOS 게이트웨이를 포함하여 모든 Ceph 클라이언트에 영향을 미칩니다.

Ceph 구성 옵션에는 마스크가 연결되어 있을 수 있습니다. 이러한 마스크는 옵션이 적용되는 데몬 또는 클라이언트에 추가로 제한할 수 있습니다.

마스크에는 두 가지 형식이 있습니다.

type:location

type 은 CRUSH 속성입니다(예: **랙** 또는 **호스트**). **위치는** 속성 유형의 값입니다. 예를 들어 **host:foo** 는 옵션을 **foo** 호스트에서 실행되는 데몬 또는 클라이언트로만 제한합니다.

예제

예제

```
ceph config set osd/host:magna045 debug_osd 20
```

class:device-class

device-class 는 clusterd **또는 ssd** 와 같은 CRUSH 장치 클래스의 이름입니다. 예를 들어, **class:ssd** 는 옵션을 SSD(반도체 드라이브)에서 지원하는 Ceph OSD로만 제한합니다. 이 마스크는 클라이언트의 비 OSD 데몬에 영향을 미치지 않습니다.

예제

```
ceph config set osd/class:hdd osd_max_backfills 8
```

관리 명령

Ceph 구성 데이터베이스는 하위 명령 **ceph config작업으로 관리할 수 *있습니다*.** 수행할 수 있는 작업은 다음과 같습니다.

ls

사용 가능한 구성 옵션을 나열합니다.

dump

스토리지 클러스터에 대한 옵션의 전체 구성 데이터베이스를 덤프합니다.

Get 설문 *조사*

특정 데몬 또는 클라이언트의 구성을 덤프합니다. 예를 들어 *md* **s.a** 와 같은 데몬일 수 있습니다.

OPENSHIFT *VALUE 설정*

Ceph 구성 데이터베이스의 구성 옵션을 설정합니다. 여기서 WHO 는 대상 데몬이고 OPTION 은 설정할 옵션입니다. VALUE 는 원하는 값입니다.

표시

실행 중인 데몬에 대해 보고된 실행 구성을 표시합니다. 사용 중인 로컬 구성 파일이 명령줄 또는 런타임에 재정의된 경우 이러한 옵션은 Ceph 모니터에 의해 저장된 옵션과 다를 수 있습니다. 또한 옵션 값의 소스는 출력의 일부로 보고됩니다.

assimilate-conf -i INPUT_FILE -o OUTPUT_FILE

INPUT_FILE 에서 구성 파일을 연결하고 유효한 옵션을 Ceph 모니터의 구성 데이터베이스로 이동합니다. OUTPUT_FILE 에 저장된 약어 구성 파일에서 Ceph Monitor 반환을 통해 인식되지 않거나 유효하지 않거나 제어할 수 없는 모든 옵션은. 이 명령은 레거시 구성 파일에서 중앙 집중식 구성 데이터베이스로 전환하는 데 유용할 수 있습니다. 구성을 결합하고 모니터 또는 기타 데몬에 동일한 옵션 세트에 대해 설정된 다른 구성 값이 있는 경우 최종 결과는 파일이 포함되는 순서에 따라 달라집니다.

도움말 옵션 -f json-pretty

JSON 형식의 출력을 사용하여 특정 옵션의 도움말을 표시합니다.

추가 리소스

- *명령에 대한 자세한 내용은 런타임에 특정 구성 설정을 참조하십시오.*

1.4. CEPH 메타 변수 사용

Metavariables는 Ceph 스토리지 클러스터 구성을 획기적으로 단순화합니다. metavariable이 구성 값에 설정된 경우 Ceph는 메타 변수를 구체적인 값으로 확장합니다.

메타 변수는 Ceph 구성 파일의 **[global], [osd], [mon]** 또는 **[client]** 섹션에서 사용할 때 매우 강력합니다. 하지만 관리 소켓과 함께 사용할 수도 있습니다. Ceph 메타 변수는 Bash 쉘 확장과 유사합니다.

Ceph에서는 다음과 같은 메타 변수를 지원합니다.

$cluster

설명

Ceph 스토리지 클러스터 이름으로 확장합니다. 동일한 하드웨어에서 여러 Ceph 스토리지 클러스터를 실행할 때 유용합니다.

예제

/etc/ceph/$cluster.keyring

Default

Ceph

$type

설명

인스턴트 데몬의 유형에 따라 **osd** 또는 **mon** 중 하나로 확장됩니다.

예제

/var/lib/ceph/$type

$id

설명

를 데몬 식별자로 확장합니다. **osd.0** 의 경우 **0** 이 됩니다.

예제

/var/lib/ceph/$type/$cluster-$id

$host

설명

는 인스턴트 데몬의 호스트 이름으로 확장됩니다.

$name

설명

$type.$id 로 확장됩니다.

예제

/var/run/ceph/$cluster-$name.asok

1.5. 런타임 시 CEPH 구성 보기

Ceph 구성 파일은 부팅 시 및 런타임에 볼 수 있습니다.

사전 요구 사항

- Ceph 노드에 대한 루트 수준 액세스.
- 관리자 인증 키에 대한 액세스.

절차

절차

1. 런타임 구성을 보려면 데몬을 실행하는 Ceph 노드에 로그인하여 다음을 실행합니다.

   구문

   ```
   ceph daemon DAEMON_TYPE.ID config show
   ```

   osd.0의 구성을 보려면 osd.0 이 포함된 노드에 로그인하고 다음 명령을 **실행합니다.**

   예제

   ```
   [root@osd ~]# ceph daemon osd.0 config show
   ```

2. 추가 옵션은 데몬과 **도움말을** 지정합니다.

   예제

   ```
   [root@osd ~]# ceph daemon osd.0 help
   ```

1.6. 런타임 시 특정 구성 보기

Red Hat Ceph Storage의 구성 설정은 Ceph Monitor 노드에서 런타임에 볼 수 있습니다.

사전 요구 사항

- 실행 중인 Red Hat Ceph Storage 클러스터.
- Ceph 모니터 노드에 대한 루트 수준 액세스.

절차

1. Ceph 노드에 로그인하여 다음을 실행합니다.

   구문

   ```
   ceph daemon DAEMON_TYPE.ID config get PARAMETER
   ```

   예제

   ```
   [root@mon ~]# ceph daemon osd.0 config get public_addr
   ```

1.7. 런타임에 특정 구성 설정

런타임에 특정 Ceph 구성을 설정하려면 **ceph config set** 명령을 사용합니다.

사전 요구 사항

- 실행 중인 Red Hat Ceph Storage 클러스터.
- Ceph 모니터 또는 OSD 노드에 대한 루트 수준 액세스.

절차

1. 모든 모니터 또는 OSD 데몬에서 구성을 설정합니다:

   구문

   ```
   ceph config set DAEMON CONFIG-OPTION VALUE
   ```

   예제

   ```
   [root@mon ~]# ceph config set osd debug_osd 10
   ```

2. 옵션과 값이 설정되었는지 확인합니다.

   예제

   ```
   [root@mon ~]# ceph config dump
   osd      advanced debug_osd  10/10
   ```

   - 모든 데몬에서 구성 옵션을 제거하려면 다음을 수행합니다.

     구문

     ```
     ceph config rm DAEMON CONFIG-OPTION VALUE
     ```

     예제

     ```
     [root@mon ~]# ceph config rm osd debug_osd
     ```

   - 특정 데몬에 대한 구성을 설정하려면 다음을 수행합니다.

     구문

     ```
     ceph config set DAEMON.DAEMON-NUMBER CONFIG-OPTION VALUE
     ```

     예제

     ```
     [root@mon ~]# ceph config set osd.0 debug_osd 10
     ```

   - 지정된 데몬에 구성이 설정되었는지 확인하려면 다음을 수행합니다.

     예제

     ```
     [root@mon ~]# ceph config dump
     osd.0      advanced debug_osd     10/10
     ```

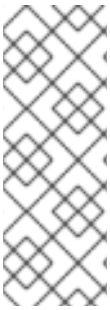
   - 특정 데몬의 구성을 제거하려면 다음을 수행합니다.

     구문

     ```
     ceph config rm DAEMON.DAEMON-NUMBER CONFIG-OPTION
     ```

예제

```
[root@mon ~]# ceph config rm osd.0 debug_osd
```



참고

구성 데이터베이스의 옵션 읽기를 지원하지 않거나 ***ceph.conf*** *를 사용하여 다른 이유로 클러스터 구성을 변경해야 하는 클라이언트를 사용하는 경우 다음 명령을 실행합니다.*

```
ceph config set mgr mgr/cephadm/manage_etc_ceph_ceph_conf false
```

스토리지 클러스터에 ***ceph.conf*** *파일을 유지 관리하고 배포해야 합니다.*

1.8. OSD 메모리 대상

bluestore는 ***osd_memory_target*** *구성 옵션을 사용하여 OSD 힙 메모리 사용량을 지정된 대상 크기 미만으로 유지합니다.*

osd_memory_target *옵션은 시스템에서 사용 가능한 RAM에 따라 OSD 메모리를 설정합니다. TCœoc가 메모리 할당자로 구성된 경우 이 옵션을 사용하고 BlueStore의* ***bluestore_cache_autotune*** *옵션이* ***true*** *로 설정된 경우 이 옵션을 사용합니다.*

블록 장치가 느리면 Ceph OSD 메모리 캐싱이 더 중요합니다(예: 캐시 적중의 이점은 솔리드 스테이트 드라이브의 경우보다 훨씬 높기 때문입니다. 그러나 HCI(하이퍼 컨버지드 인프라) 또는 기타 애플리케이션과 같은 다른 서비스와 OSD를 배치하기로 결정해야 합니다.

1.8.1. OSD 메모리 대상 설정

osd_memory_target *옵션을 사용하여 스토리지 클러스터의 모든 OSD에 대한 최대 메모리 임계값을 설정하거나 특정 OSD에 대한 메모리 임계값을 설정합니다.* ***osd_memory_target*** *옵션이 16GB로 설정된 OSD는 최대 16GB의 메모리를 사용할 수 있습니다.*

참고

개별 OSD의 구성 옵션이 모든 OSD의 설정보다 우선합니다.

사전 요구 사항

- *실행 중인 Red Hat Ceph Storage 클러스터.*
- *스토리지 클러스터의 모든 호스트에 대한 루트 수준 액세스.*

절차

- *스토리지 클러스터의 모든 OSD에* ***osd_memory_target*** *을 설정하려면 다음을 수행합니다.*

  구문

  ```
  ceph config set osd osd_memory_target VALUE
  ```

  VALUE 는 스토리지 클러스터의 각 OSD에 할당할 GBytes 메모리 수입니다.

- *스토리지 클러스터에서 특정 OSD에* ***osd_memory_target*** *을 설정하려면 다음을 수행합니다.*

구문

```
ceph config set osd.id osd_memory_target VALUE
```

. ID는 OSD의 ID이고 VALUE 는 지정된 OSD에 할당할 메모리의 수입니다. 예를 들어, 최대 16GB의 메모리를 사용하도록 ID 8이 있는 OSD를 구성하려면 다음을 수행합니다.

예제

```
[ceph: root@host01 /]# ceph config set osd.8 osd_memory_target 16G
```

- *개별 OSD가 최대 메모리 양을 1개 사용하도록 설정하고 나머지 OSD가 다른 양을 사용하도록 구성하려면 먼저 개별 OSD를 지정합니다.*

예제

```
[ceph: root@host01 /]# ceph config set osd osd_memory_target 16G
[ceph: root@host01 /]# ceph config set osd.8 osd_memory_target 8G
```

추가 리소스

- *OSD 메모리 사용량을 자동 조정하도록 Red Hat Ceph Storage를 구성하려면 Operations Guide의 OSD 메모리 자동 튜닝을 참조하십시오.*

1.9. OSD 메모리 자동 튜닝

*OSD 데몬은 **osd_memory_target** 구성 옵션에 따라 메모리 사용을 조정합니다. 옵션 **osd_memory_target** 은 시스템에서 사용 가능한 RAM을 기반으로 OSD 메모리를 설정합니다.*

*Red Hat Ceph Storage가 다른 서비스와 메모리를 공유하지 않는 전용 노드에 배포된 경우 **cephadm** 은 총 RAM 용량 및 배포된 OSD 수에 따라 자동으로 OSD 소비를 조정합니다.*

중요

*기본적으로 Red Hat Ceph Storage 5.1에서 **osd_memory_target_autotune** 매개 변수가 **true** 로 설정됩니다.*

구문

```
ceph config set osd osd_memory_target_autotune true
```

*스토리지 클러스터를 Red Hat Ceph Storage 5.0으로 업그레이드하면 OSD 추가 또는 OSD 교체와 같은 클러스터 유지 관리를 위해 시스템 메모리에 따라 **osd_memory_target_autotune** 매개변수를 **true** 로 설정하는 것이 좋습니다.*

*Cephadm은 분수 **mgr/cephadm/autotune_memory_target_ratio** 로 시작합니다. 기본값은 시스템에서 전체 RAM의 **0.7** 로 설정되고, 비OSDS와 같은 비-자동 조정 데몬에서 사용하는 메모리를 제거하고 **osd_memory_target_autotune** 이 false인 OSD의 경우 나머지 OSD로 나눕니다.*

***osd_memory_target** 매개변수는 다음과 같이 계산됩니다.*

구문

```
osd_memory_target = TOTAL_RAM_OF_THE_OSD * (1048576) * (autotune_memory_target_ratio) / NUMBER_OF_OSDS_IN_THE_OSD_NODE - (SPACE_ALLOCATED_FOR_OTHER_DAEMONS)
```

SPACE_ALLOCATED_FOR_OTHER_DAEMONS 는 선택적으로 다음과 같은 데몬 공간 할당을 포함할 수 있습니다.

- Alertmanager: 1GB
- Grafana: 1GB
- Ceph Manager: 4GB
- Ceph 모니터: 2GB
- node-exporter: 1GB
- Prometheus: 1GB

예를 들어 노드에 OSD가 24개 있고 251GB RAM 공간이 있는 경우 **osd_memory_target** 은 **7860684936** 입니다.

최종 대상은 옵션을 사용하여 구성 데이터베이스에 반영됩니다. **MEM LIMIT** 열의 **ceph orch ps** 출력에서 각 데몬에서 사용하는 제한 및 현재 메모리를 볼 수 있습니다.

참고

Red Hat Ceph Storage 5.1에서 **osd_memory_target_autotune true** 의 기본 설정은 컴퓨팅 및 Ceph 스토리지 서비스가 함께 배치되는 하이퍼컨버지드 인프라에 적합하지 않습니다. 하이퍼컨버지드 인프라에서 **autotune_memory_target_ratio** 를 **0.2** 로 설정하여 Ceph의 메모리 사용을 줄일 수 있습니다.

예제

```
[ceph: root@host01 /]# ceph config set mgr mgr/cephadm/autotune_memory_target_ratio 0.2
```

스토리지 클러스터에서 OSD에 대한 특정 메모리 대상을 수동으로 설정할 수 있습니다.

예제

```
[ceph: root@host01 /]# ceph config set osd.123 osd_memory_target 7860684936
```

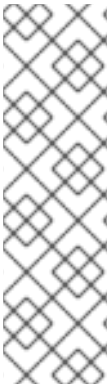
스토리지 클러스터에서 OSD 호스트에 대한 특정 메모리 대상을 수동으로 설정할 수 있습니다.

구문

```
ceph config set osd/host:HOSTNAME osd_memory_target TARGET_BYTES
```

예제

```
[ceph: root@host01 /]# ceph config set osd/host:host01 osd_memory_target 1000000000
```



참고

osd_memory_target_autotune *을 활성화하면 기존 수동 OSD 메모리 대상 설정을 덮어씁니다.* ***osd_memory_target_autotune*** *옵션 또는 기타 유사한 옵션이 활성화된 경우에도 데몬 메모리가 튜닝되지 않도록 하려면 호스트에서* ***_no_autotune_memory*** *레이블을 설정합니다.*

구문

```
ceph orch host label add HOSTNAME _no_autotune_memory
```

autotune 옵션을 비활성화하고 특정 메모리 대상을 설정하여 메모리 자동 튜닝에서 OSD를 제외할 수 있습니다.

예제

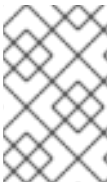
```
[ceph: root@host01 /]# ceph config set osd.123 osd_memory_target_autotune false
[ceph: root@host01 /]# ceph config set osd.123 osd_memory_target 16G
```

1.10. MDS 메모리 캐시 제한

MDS 서버는 ***cephfs_metadata*** *라는 별도의 스토리지 풀에 메타데이터를 유지하고 Ceph OSD의 사용자입니다. Ceph File Systems의 경우 MDS 서버는 스토리지 클러스터 내의 단일 스토리지 장치뿐만 아니라 전체 Red Hat Ceph Storage 클러스터를 지원해야 합니다. 따라서 특히 워크로드가 중간 규모의 작은 파일로 구성된 경우 특히 메타데이터와 데이터 간의 비율이 훨씬 높은 경우 메모리 요구 사항이 상당할 수 있습니다.*

예제: ***mds_cache_memory_limit*** *를 2000000000 바이트로 설정합니다.*

```
ceph_conf_overrides:
  osd:
    mds_cache_memory_limit=2000000000
```



참고

메타데이터를 많이 사용하는 워크로드가 있는 대규모 Red Hat Ceph Storage 클러스터의 경우 MDS 서버를 다른 메모리 집약 서비스와 동일한 노드에 배치하지 마십시오. 이를 통해 MDS에 메모리를 더 많은 메모리를 할당할 수 있습니다(예: 100GB보다 큰 크기).

추가 리소스

- *Red Hat Ceph Storage 파일 시스템 가이드의 메타데이터 서버 캐시 크기 제한을 참조하십시오.*

1.11. 추가 리소스

- *특정 옵션 설명 및 사용법에 대해서는 부록 A 의 일반 Ceph 구성 옵션을 참조하십시오.*

2장. CEPH 네트워크 구성

스토리지 관리자는 Red Hat Ceph Storage 클러스터가 작동하는 네트워크 환경을 이해하고 그에 따라 Red Hat Ceph Storage를 구성해야 합니다. Ceph 네트워크 옵션을 이해하고 구성하면 전체 스토리지 클러스터에서 최적의 성능과 안정성을 보장할 수 있습니다.

2.1. 사전 요구 사항

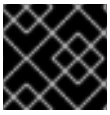
- 네트워크 연결.
- Red Hat Ceph Storage 소프트웨어 설치.

2.2. CEPH의 네트워크 구성

네트워크 구성은 고성능 Red Hat Ceph Storage 클러스터를 구축하는 데 중요합니다. Ceph 스토리지 클러스터는 Ceph 클라이언트 대신 요청 라우팅 또는 디스패치를 수행하지 않습니다. 대신 Ceph 클라이언트는 Ceph OSD 데몬에 직접 요청합니다. Ceph OSD는 Ceph 클라이언트를 대신하여 데이터 복제를 수행하므로 복제 및 기타 요인이 Ceph 스토리지 클러스터의 네트워크에 추가 로드를 적용합니다.

Ceph에는 모든 데몬에 적용되는 하나의 네트워크 구성 요구 사항이 있습니다. Ceph 구성 파일은 각 데몬에 대해 **호스트** 를 지정해야 합니다.

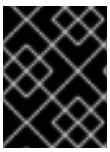
cephadm 과 같은 일부 배포 유틸리티는 구성 파일을 생성합니다. 배포 유틸리티가 이를 수행하는 경우 이러한 값을 설정하지 마십시오.



중요

host 옵션은 FQDN이 아닌 노드의 짧은 이름입니다. IP 주소가 아닙니다.

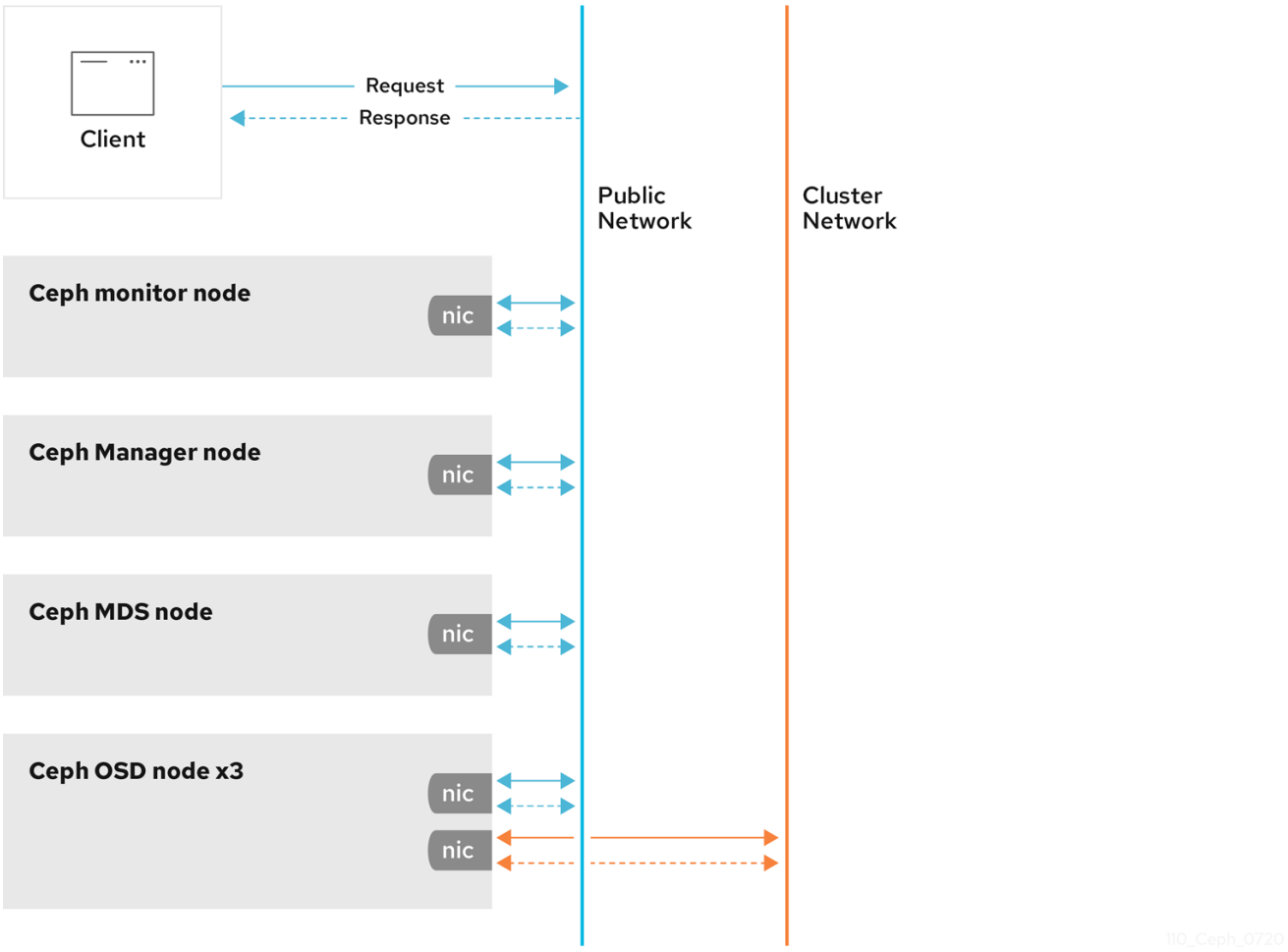
모든 Ceph 클러스터에서 공용 네트워크를 사용해야 합니다. 그러나 내부 클러스터 네트워크를 지정하지 않는 한 Ceph는 단일 공용 네트워크를 가정합니다. Ceph는 공용 네트워크에서만 작동할 수 있지만, 대규모 스토리지 클러스터에서는 클러스터 관련 트래픽만 수행할 수 있도록 두 번째 사설 네트워크로 인해 성능이 크게 향상됩니다.



중요

Red Hat은 두 개의 네트워크가 있는 Ceph 스토리지 클러스터를 실행하는 것이 좋습니다. 공용 네트워크 1개, 사설 네트워크 1개.

두 개의 네트워크를 지원하려면 각 Ceph 노드에 NIC(네트워크 인터페이스 카드)가 두 개 이상 있어야 합니다.



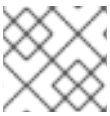
두 개의 개별 네트워크를 운영하는 데에는 몇 가지 이유가 있습니다.

- ***성능:*** *Ceph OSD는 Ceph 클라이언트의 데이터 복제를 처리합니다. Ceph OSD가 두 번 이상 데이터를 복제할 때 Ceph OSD 간에 네트워크 부하가 쉽게 드러납니다. Ceph 클라이언트와 Ceph 스토리지 클러스터 간의 네트워크 로드가 용이합니다. 그러면 대기 시간이 발생하고 성능 문제가 발생할 수 있습니다. 복구 및 리밸런싱으로 인해 공용 네트워크에 상당한 대기 시간이 발생할 수 있습니다.*

- ***보안****: 대부분의 사람들은 일반적으로 민간이지만 일부 행위자들은 서비스 거부(DoS) 공격으로 알려진 상황에 참여하게 됩니다. Ceph OSD 간 트래픽이 중단되면 피어링에 실패할 수 있으며 배치 그룹이 더 이상* ***활성 + 클린*** *상태를 반영하지 않을 수 있으므로 사용자가 데이터를 읽고 쓰는 것을 방지할 수 있습니다. 이러한 유형의 공격을 없애는 좋은 방법은 인터넷에 직접 연결되지 않는 완전히 분리된 클러스터 네트워크를 유지하는 것입니다.*

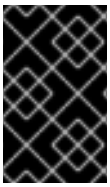
네트워크 구성 설정이 필요하지 않습니다. 공용 네트워크가 Ceph 데몬을 실행하는 모든 호스트에 구성되어 있다고 가정하면 Ceph가 공용 네트워크에서만 작동할 수 있습니다. 그러나 Ceph를 사용하면 공용 네트워크에 대한 여러 IP 네트워크 및 서브넷 마스크를 포함하여 훨씬 더 구체적인 기준을 설정할 수 있습니다. OSD 하트비트, 오브젝트 복제 및 복구 트래픽을 처리하는 별도의 클러스터 네트워크를 구축할 수도 있습니다.

구성에 설정한 IP 주소와 공용 방향 IP 주소 네트워크 클라이언트가 서비스에 액세스하는 데 사용할 수 있는 IP 주소를 혼동하지 마십시오. 일반적인 내부 IP 네트워크는 대부분 ***192.168.0.0*** *또는* ***10.0.0.0*** *입니다.*



참고

Ceph는 서브넷(예 ***: 10.0.0.0/24)에 CIDR 표기법을 사용합니다****.*



중요

공용 또는 사설 네트워크에 대해 둘 이상의 IP 주소와 서브넷 마스크를 지정하는 경우 네트워크 내의 서브넷이 서로 라우팅할 수 있어야 합니다. 또한 IP 테이블에 각 IP 주소와 서브넷을 포함하고 필요에 따라 포트를 열어야 합니다.

네트워크를 구성할 때 클러스터를 재시작하거나 각 데몬을 다시 시작할 수 있습니다. Ceph 데몬은 동적으로 바인드되므로 네트워크 구성을 변경하는 경우 한 번에 전체 클러스터를 다시 시작할 필요가 없습니다.

추가 리소스

- *특정 옵션 설명 및 사용법은 Red Hat Ceph Storage Configuration Guide , 부록 B 를 참조하십시오.*

2.3. CEPH 네트워크 바우처

collectd은 Ceph 네트워크 계층 구현입니다. Red Hat은 다음과 같은 두 가지 유형을 지원합니다.

- ***simple***
- ***async***

*Red Hat Ceph Storage 4 이상에서 **async** 는 기본 유형입니다. 아카이브 유형을 변경하려면 Ceph 구성 파일의 **[global]** 섹션에 **ms_type** 구성 설정을 지정합니다.*

참고

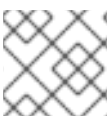
*비동기식 **의** 경우 Red Hat은 **posix** 전송 유형을 지원하지만 현재 **rdma** or **dpdk** 는 지원하지 않습니다. 기본적으로 Red Hat Ceph Storage 4 이상의 **ms_type** 설정은 **async+posix** 를 반영합니다. 여기서 **async** 는 jboss type이고 **posix** 는 전송 유형입니다.*

SimpleMessenger

***SimpleMessenger** 구현은 소켓당 두 개의 스레드가 있는 TCP 소켓을 사용합니다. Ceph는 각 논리 세션을 연결과 연결합니다. 파이프는 각 메시지의 입력 및 출력을 포함하여 연결을 처리합니다.*
***SimpleMessenger** 는 **posix** 전송 유형에 효과적이지만, **rdma** or **dpdk** 와 같은 다른 전송 유형에는 효과적이지 않습니다.*

AsyncMessenger

*결과적으로 **AsyncMessenger** 는 Red Hat Ceph Storage 4 이상을 위한 기본 유형입니다. Red Hat Ceph Storage 4 이상에서는 **AsyncMessenger** 구현에서 연결에 고정된 스레드 풀이 있는 TCP 소켓을 사용하며, 이 소켓은 가장 많은 수의 복제본 또는 삭제 코드 청크와 같아야 합니다. CPU 수가 낮거나 서버당 OSD 수가 많기 때문에 성능이 저하되는 경우 스레드 수를 더 낮은 값으로 설정할 수 있습니다.*



참고

*Red Hat은 현재 **rdma** or **dpdk** 와 같은 다른 전송 유형을 지원하지 않습니다.*

추가 리소스

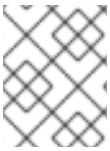
- *Red Hat Ceph Storage 구성 가이드 의 AsyncMessenger 옵션, 특정 옵션 설명 및 사용에 대한 부록 B 를 참조하십시오.*
- *Ceph 버전 2 프로토콜로 유선 암호화를 사용하는 방법에 대한 자세한 내용은 Red Hat Ceph Storage Architecture 가이드를 참조하십시오.*

2.4. 공용 네트워크 구성

Ceph 네트워크를 구성하려면 **cephadm** 쉘 내에서 **config set** 명령을 사용합니다. 네트워크 구성에서 설정한 IP 주소는 네트워크 클라이언트가 서비스에 액세스하는 데 사용할 수 있는 공용 방향 IP 주소와 다릅니다.

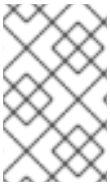
Ceph는 공용 네트워크에서만 완벽하게 작동합니다. 그러나 Ceph를 사용하면 공용 네트워크에 대한 여러 IP 네트워크를 포함하여 훨씬 더 구체적인 기준을 설정할 수 있습니다.

OSD 하트비트, 오브젝트 복제 및 복구 트래픽을 처리하기 위해 별도의 프라이빗 클러스터 네트워크를 설정할 수도 있습니다. 사설 네트워크에 대한 자세한 내용은 사설 네트워크 구성을 참조하십시오.



참고

Ceph는 서브넷에 CIDR 표기법을 사용합니다(예: 10.0.0.0/24). 일반적인 내부 IP 네트워크는 192.168.0.0/24 또는 10.0.0.0/24인 경우가 많습니다.



참고

공용 또는 클러스터 네트워크에 대해 둘 이상의 IP 주소를 지정하는 경우 네트워크 내의 서브넷이 서로 라우팅할 수 있어야 합니다. 또한 IP 테이블에 각 IP 주소를 포함하고 필요한 경우 포트를 열어야 합니다.

공용 네트워크 구성을 사용하면 공용 네트워크의 IP 주소와 서브넷을 구체적으로 정의할 수 있습니다.

사전 요구 사항

- Red Hat Ceph Storage 소프트웨어 설치.

절차

1. **cephadm** 쉘에 로그인합니다.

   예제

   ```
   [root@host01 ~]# cephadm shell
   ```

2. 서브넷을 사용하여 공용 네트워크를 구성합니다.

   구문

   ```
   ceph config set mon public_network IP_ADDRESS_WITH_SUBNET
   ```

   예제

   ```
   [ceph: root@host01 /]# ceph config set mon public_network 192.168.0.0/24
   ```

3. 스토리지 클러스터에서 서비스 목록을 가져옵니다.

   예제

   ```
   [ceph: root@host01 /]# ceph orch ls
   ```

4. *데몬을 다시 시작합니다. Ceph 데몬은 동적으로 바인드되므로 특정 데몬의 네트워크 구성을 변경하는 경우 한 번에 전체 클러스터를 재시작할 필요가 없습니다.*

   예제

   ```
   [ceph: root@host01 /]# ceph orch restart mon
   ```

5. *선택 사항: 클러스터를 다시 시작하려면 admin 노드에서 root 사용자로 **systemctl** 명령을 실행합니다.*

   구문

   ```
   systemctl restart ceph-FSID_OF_CLUSTER.target
   ```

   예제

   ```
   [root@host01 ~]# systemctl restart ceph-1ca9f6a8-d036-11ec-8263-fa163ee967ad.target
   ```

추가 리소스

- *특정 옵션 설명 및 사용법은 Red Hat Ceph Storage Configuration Guide , 부록 B 를 참조하십시오.*

2.5. 사설 네트워크 구성

네트워크 구성 설정이 필요하지 않습니다. 사설 네트워크라고도 하는 클러스터 네트워크를 구체적으로 구성하지 않는 한 Ceph 에서는 모든 호스트가 작동하는 공용 네트워크를 가정합니다.

클러스터 네트워크를 생성하는 경우 OSD 는 클러스터 네트워크를 통해 하트비트, 오브젝트 복제 및 복구 트래픽을 라우팅합니다. 이렇게 하면 단일 네트워크 사용과 비교하여 성능이 향상될 수 있습니다.

중요

보안을 강화하기 위해 공용 네트워크 또는 인터넷에서 클러스터 네트워크에 연결할 수 없어야 합니다.

*클러스터 네트워크를 할당하려면 **cephadm bootstrap 명령에 --cluster-network** 옵션을 사용합니다. 지정한 클러스터 네트워크는 CIDR 표기법으로 서브넷을 정의해야 합니다(예: 10.90.90.0/24 또는 fe80::/64).*

*boostrap 후 **cluster_network** 를 구성할 수도 있습니다.*

사전 요구 사항

- *Ceph 소프트웨어 리포지토리에 액세스할 수 있습니다.*
- *스토리지 클러스터의 모든 노드에 대한 루트 수준 액세스.*

절차

- *스토리지 클러스터에서 모니터 노드로 사용할 초기 노드에서 **cephadm 부트스트랩 명령**을 실행합니다. 명령에 **--cluster-network** 옵션을 포함합니다.*

구문

```
cephadm bootstrap --mon-ip IP-ADDRESS --registry-url registry.redhat.io --registry-username USER_NAME --registry-password PASSWORD --cluster-network NETWORK-IP-ADDRESS
```

예제

```
[root@host01 ~]# cephadm bootstrap --mon-ip 10.10.128.68 --registry-url registry.redhat.io --registry-username myuser1 --registry-password mypassword1 --cluster-network 10.10.0.0/24
```

- *부트스트랩 후* ***cluster_network*** *를 구성하려면* ***config set*** *명령을 실행하고 데몬을 다시 배포합니다.*

  1. ***cephadm*** *쉘에 로그인합니다.*

     예제

     ```
     [root@host01 ~]# cephadm shell
     ```

  2. *서브넷을 사용하여 클러스터 네트워크를 구성합니다.*

     구문

     ```
     ceph config set global cluster_network IP_ADDRESS_WITH_SUBNET
     ```

     예제

     ```
     [ceph: root@host01 /]# ceph config set global cluster_network 10.10.0.0/24
     ```

  3. *스토리지 클러스터에서 서비스 목록을 가져옵니다.*

     예제

     ```
     [ceph: root@host01 /]# ceph orch ls
     ```

  4. *데몬을 다시 시작합니다. Ceph 데몬은 동적으로 바인드되므로 특정 데몬의 네트워크 구성을 변경하는 경우 한 번에 전체 클러스터를 재시작할 필요가 없습니다.*

     예제

     ```
     [ceph: root@host01 /]# ceph orch restart mon
     ```

  5. *선택 사항: 클러스터를 다시 시작하려면 admin 노드에서 root 사용자로* ***systemctl*** *명령을 실행합니다.*

     구문

     ```
     systemctl restart ceph-FSID_OF_CLUSTER.target
     ```

     예제

     -

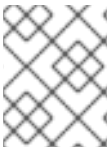
```
[root@host01 ~]# systemctl restart ceph-1ca9f6a8-d036-11ec-8263-fa163ee967ad.target
```

추가 리소스

- ***cephadm* 부트스트랩** 호출에 대한 자세한 내용은 Red Hat Ceph Storage 설치 가이드의 새 스토리지 클러스터 부트스트랩 섹션을 참조하십시오.

2.6. 기본 CEPH 포트에 대해 방화벽 규칙 구성

기본적으로 Red Hat Ceph Storage 데몬은 TCP 포트 6800-s를 사용합니다.7100: 클러스터의 다른 호스트와 통신합니다. 호스트의 방화벽이 이러한 포트에서 연결을 허용하는지 확인할 수 있습니다.



참고

네트워크에 전용 방화벽이 있는 경우 다음 절차에 따라 구성을 확인해야 할 수 있습니다. 자세한 내용은 방화벽 설명서를 참조하십시오.

자세한 내용은 방화벽 설명서를 참조하십시오.

사전 요구 사항

- 호스트에 대한 루트 수준 액세스.

절차

1. 호스트의 **iptables** 구성을 확인합니다.
   a. 활성 규칙을 나열합니다.

   ```
   [root@host1 ~]# iptables -L
   ```

   b. TCP 포트 6800에서 연결을 제한하는 규칙이 없는지 확인합니다.7100.

   예제

   ```
   REJECT all -- anywhere anywhere reject-with icmp-host-prohibited
   ```

2. 호스트의 **firewalld** 구성을 확인합니다.
   a. 호스트에서 열려 있는 포트를 나열합니다.

   구문

   ```
   firewall-cmd --zone ZONE --list-ports
   ```

   예제

   ```
   [root@host1 ~]# firewall-cmd --zone default --list-ports
   ```

   b. 범위가 TCP 포트 6800이 포함되어 있는지 확인합니다.7100.

2.7. CEPH MONITOR 노드의 방화벽 설정

네트워크에서 모든 Ceph 트래픽에 대한 암호화를 활성화할 수 있으며, 버전 2 프로토콜이 도입되었습니다. v2의 **보안** 모드 설정은 Ceph 데몬과 Ceph 클라이언트 간의 통신을 암호화하여 포괄적인 암호화를 제공합니다.

collectd v2 프로토콜

Ceph의 온프레미스 프로토콜인 **msgr2** 의 두 번째 버전에는 다음과 같은 몇 가지 새로운 기능이 포함되어 있습니다.

- 보안 모드는 네트워크를 통해 이동하는 모든 데이터를 암호화합니다.
- 인증 페이로드 캡슐화 개선.
- 광고 및 협상 기능 개선.

Ceph 데몬은 레거시, v1- 호환 및 새로운 v2 호환 Ceph 클라이언트가 동일한 스토리지 클러스터에 연결할 수 있도록 여러 포트에 바인딩됩니다. Ceph Monitor 데몬에 연결되는 Ceph 클라이언트 또는 기타 Ceph 데몬은 가능한 경우 **v2** 프로토콜을 먼저 사용하려고 하지만 그렇지 않은 경우 레거시 **v1** 프로토콜을 사용합니다. 기본적으로 **v1 및 v 2** 프로토콜 모두 활성화됩니다. 새 v2 포트는 3300 이고 레거시 v1 포트는 기본적으로 6789 입니다.

사전 요구 사항

- 실행 중인 Red Hat Ceph Storage 클러스터.
- Ceph 소프트웨어 리포지토리에 액세스할 수 있습니다.
- Ceph 모니터 노드에 대한 루트 수준 액세스.

절차

1. 다음 예제를 사용하여 규칙을 추가합니다.

```
[root@mon ~]# sudo iptables -A INPUT -i IFACE -p tcp -s IP-ADDRESS/NETMASK --dport
6789 -j ACCEPT
[root@mon ~]# sudo iptables -A INPUT -i IFACE -p tcp -s IP-ADDRESS/NETMASK --dport
3300 -j ACCEPT
```

   a. **IFACE** 를 공용 네트워크 인터페이스(예: **eth0**,**eth1** 등)로 바꿉니다.

   b. **IP-ADDRESS** 를 공용 네트워크의 IP 주소로 바꾸고 **NETMASK** 를 공용 네트워크의 넷마스크로 바꿉니다.

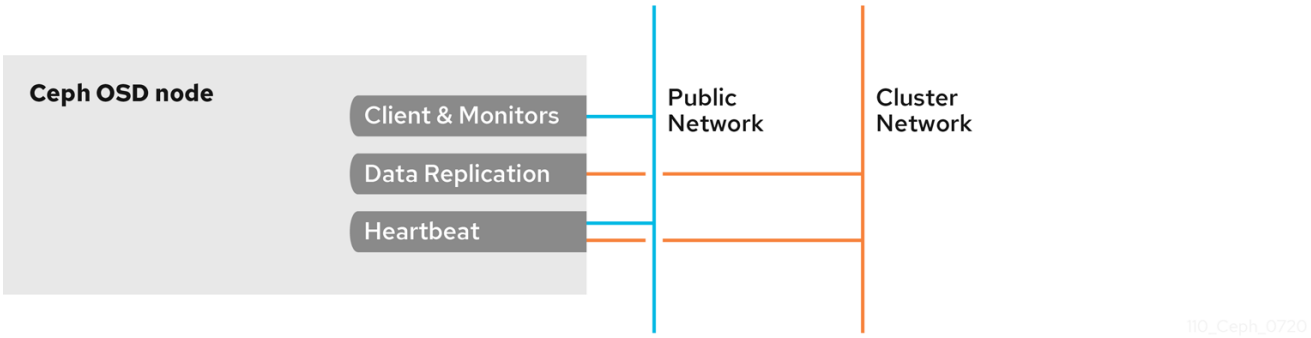
2. **firewalld** 데몬의 경우 다음 명령을 실행합니다.

```
[root@mon ~]# firewall-cmd --zone=public --add-port=6789/tcp
[root@mon ~]# firewall-cmd --zone=public --add-port=6789/tcp --permanent
[root@mon ~]# firewall-cmd --zone=public --add-port=3300/tcp
[root@mon ~]# firewall-cmd --zone=public --add-port=3300/tcp --permanent
```

2.8. CEPH OSD의 방화벽 설정

기본적으로 Ceph OSD는 포트 6800부터 Ceph 노드의 첫 번째 사용 가능한 포트에 바인딩됩니다. 노드에서 실행되는 각 OSD의 포트 6800에서 최소 4개의 포트를 열어야 합니다.

- *하나는 공용 네트워크의 클라이언트 및 모니터와 통신하기 위한 것입니다.*
- *클러스터 네트워크의 다른 OSD에 데이터를 보내는 것입니다.*
- *클러스터 네트워크에서 하트비트 패킷을 보내는 2개.*



포트는 노드별로 다릅니다. 그러나 프로세스가 다시 시작되고 바인딩된 포트가 릴리스되지 않는 경우 해당 Ceph 노드에서 실행 중인 Ceph 데몬에 필요한 포트 수보다 많은 포트를 열어야 할 수 있습니다. 재시작된 데몬이 새 포트에 바인딩되어 포트를 해제하지 않고 데몬이 실패하는 경우 몇 가지 추가 포트를 여는 것이 좋습니다. 또한 6800의 포트 범위를 여는 것이 좋습니다. 각 OSD 노드의 7300.

별도의 공용 네트워크와 클러스터 네트워크를 설정하는 경우 클라이언트가 공용 네트워크를 사용하여 연결되고 다른 Ceph OSD 데몬이 클러스터 네트워크를 사용하여 연결되므로 공용 네트워크와 클러스터 네트워크 모두에 대한 규칙을 추가해야 합니다.

사전 요구 사항

- *실행 중인 Red Hat Ceph Storage 클러스터.*
- *Ceph 소프트웨어 리포지토리에 액세스할 수 있습니다.*
- *Ceph OSD 노드에 대한 루트 수준 액세스.*

절차

1. *다음 예제를 사용하여 규칙을 추가합니다.*

```
[root@mon ~]# sudo iptables -A INPUT -i IFACE  -m multiport -p tcp -s IP-ADDRESS/NETMASK --dports 6800:7300 -j ACCEPT
```

   a. ***IFACE*** *를 공용 네트워크 인터페이스(예:* ***eth0,eth1*** *등)로 바꿉니다.*

   b. ***IP-ADDRESS*** *를 공용 네트워크의 IP 주소로 바꾸고* ***NETMASK*** *를 공용 네트워크의 넷마스크로 바꿉니다.*

2. ***firewalld*** *데몬의 경우 다음을 실행합니다.*

```
[root@mon ~] # firewall-cmd --zone=public --add-port=6800-7300/tcp
[root@mon ~] # firewall-cmd --zone=public --add-port=6800-7300/tcp --permanent
```

클러스터 네트워크를 다른 영역에 배치하는 경우 해당 영역의 포트를 적절하게 엽니다.

2.9. 추가 리소스

- 특정 옵션 설명 및 사용에 대해서는 부록 B 의 Red Hat Ceph Storage 네트워크 구성 옵션을 참조하십시오.
- Ceph 버전 2 프로토콜로 유선 암호화를 사용하는 방법에 대한 자세한 내용은 Red Hat Ceph Storage Architecture 가이드를 참조하십시오.

3 장. CEPH 모니터 구성

스토리지 관리자는 Ceph 모니터의 기본 구성 값을 사용하거나 원하는 워크로드에 따라 사용자 지정할 수 있습니다.

3.1. 사전 요구 사항

- Red Hat Ceph Storage 소프트웨어 설치.

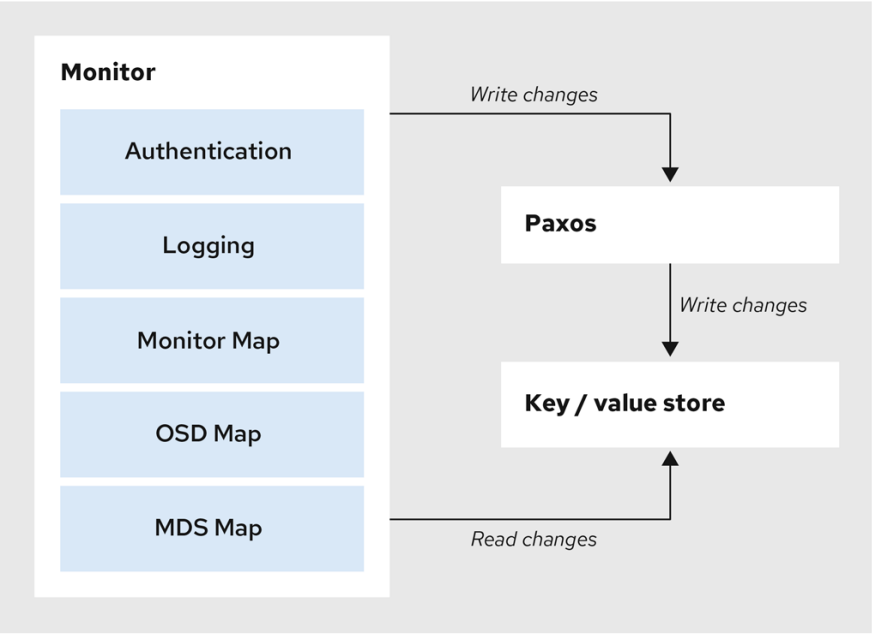
3.2. CEPH 모니터 구성

신뢰할 수 있는 Red Hat Ceph Storage 클러스터를 구축하는 데 있어 Ceph Monitor 를 구성하는 방법을 이해하는 것이 중요합니다. 모든 스토리지 클러스터에는 모니터가 하나 이상 있습니다. Ceph 모니터 구성은 일반적으로 상당히 일관되게 유지되지만 스토리지 클러스터에서 Ceph 모니터를 추가, 제거 또는 교체할 수 있습니다.

Ceph 모니터는 클러스터 맵의 "마스터 복사"를 유지합니다. 즉, Ceph 클라이언트는 하나의 Ceph 모니터에 연결하고 현재 클러스터 맵을 검색하여 모든 Ceph 모니터 및 Ceph OSD 의 위치를 확인할 수 있습니다.

Ceph 클라이언트가 Ceph OSD 에서 읽거나 쓰기 전에 먼저 Ceph 모니터에 연결해야 합니다. 현재 클러스터 맵과 CRUSH 알고리즘의 복사본을 사용하여 Ceph 클라이언트에서 모든 오브젝트의 위치를 계산할 수 있습니다. 개체 위치를 계산하는 기능을 통해 Ceph 클라이언트가 Ceph OSD 와 직접 통신할 수 있습니다. 이 기능은 Ceph 의 높은 확장성과 성능에 매우 중요한 부분입니다.

Ceph Monitor 의 주요 역할은 클러스터 맵의 마스터 복사본을 유지 관리하는 것입니다. Ceph 모니터는 인증 및 로깅 서비스도 제공합니다. Ceph 모니터는 모니터 서비스의 모든 변경 사항을 단일 Paxos 인스턴스에 작성하고 Paxos 는 키-값 저장소에 변경 사항을 작성하여 강력한 일관성을 유지합니다. Ceph 모니터는 동기화 작업 중에 최신 버전의 클러스터 맵을 쿼리할 수 있습니다. Ceph 모니터는 **rocksdb** 데이터베이스를 사용하여 키-값 저장소의 스냅샷과 반복자를 활용하여 저장소 전체 동기화를 수행합니다.



3.2.1. Ceph Monitor 구성 데이터베이스 보기

구성 데이터베이스에서 Ceph Monitor 구성을 볼 수 있습니다.

참고

*이전 Red Hat Ceph Storage 릴리스는 **/etc/ceph/ceph.conf** 에서 Ceph Monitor 구성을 중앙 집중화합니다. 이 구성 파일은 Red Hat Ceph Storage 5에서 더 이상 사용되지 않습니다.*

사전 요구 사항

- *실행 중인 Red Hat Ceph Storage 클러스터.*
- *Ceph 모니터 호스트에 대한 루트 수준 액세스.*

절차

1. ***cephadm** 쉘에 로그인합니다.*

```
[root@host01 ~]# cephadm shell
```

2. ***ceph config** 명령을 사용하여 구성 데이터베이스를 확인합니다.*

   예제

```
[ceph: root@host01 /]# ceph config get mon
```

추가 리소스

- ***ceph config 명령에 사용할 수 있는 옵션에 대한 자세한 내용은 ceph config -h** 를 사용합니다.*

3.3. CEPH 클러스터 맵

클러스터 맵은 모니터 맵, OSD 맵 및 배치 그룹 맵을 포함한 복합 맵입니다. 클러스터 맵은 다음과 같은 여러 중요한 이벤트를 추적합니다.

- *Red Hat Ceph Storage 클러스터에 **있는** 프로세스는 무엇입니까.*
- *Red Hat Ceph Storage 클러스터에 **있는** 프로세스가 **시작되어** 실행 중이거나 **중단된** 프로세스는 무엇입니까.*
- *배치 그룹은 **활성** 또는 **비활성** 그룹이며 **정리** 또는 다른 상태입니다.*
- *클러스터의 현재 상태를 반영하는 기타 세부 정보는 다음과 같습니다.*
  - *총 스토리지 공간 또는*
  - *사용된 스토리지의 양입니다.*

*클러스터 상태가 크게 변경되는 경우(예: Ceph OSD가 다운되면 배치 그룹이 성능이 저하된 상태가 됩니다. 클러스터 맵이 업데이트되어 클러스터의 현재 상태를 반영합니다. 또한 Ceph 모니터는 클러스터의 이전 상태 기록도 유지합니다. 모니터 맵, OSD 맵 및 배치 그룹 맵은 각각 해당 맵 버전의 기록을 유지합니다. 각 버전을 **epoch** 라고 합니다.*

Red Hat Ceph Storage 클러스터를 작동하는 경우 이러한 상태를 유지하는 것은 클러스터 관리에서 중요한 부분입니다.

3.4. CEPH MONITOR 쿼럼

클러스터는 단일 모니터로 충분히 실행됩니다. 그러나 단일 모니터는 단일 장애 지점입니다. 프로덕션 Ceph 스토리지 클러스터에서 고가용성을 보장하려면 여러 모니터가 있는 Ceph를 실행하여 단일 모니터의 오류가 발생하면 전체 스토리지 클러스터가 실패하지 않습니다.

Ceph 스토리지 클러스터에서 고가용성을 위해 여러 Ceph 모니터를 실행하는 경우 Ceph 모니터는 Paxos 알고리즘을 사용하여 마스터 클러스터 맵에 대한 합의점을 구축합니다. 합의를 적용하려면 클러스터 맵에 대한 합의를 위해 쿼럼을 설정하기 위해 실행되는 대부분의 모니터가 필요합니다. 예를 들어 1, 3개 중 2개, 5개 중 3개, 6개 중 4개 등입니다.

고가용성을 보장하기 위해 Ceph 모니터가 3개 이상 있는 프로덕션 Red Hat Ceph Storage 클러스터를 실행하는 것이 좋습니다. 여러 모니터를 실행할 때 스토리지 클러스터의 멤버여야 하는 초기 모니터를 지정하여 쿼럼을 설정할 수 있습니다. 이렇게 하면 스토리지 클러스터가 온라인 상태가 되는 데 걸리는 시간이 단축될 수 있습니다.

```
[mon]
mon_initial_members = a,b,c
```

참고

*스토리지 클러스터의 대부분의 모니터는 에서 서로 도달하여 쿼럼을 설정할 수 있어야 합니다. 모니터의 초기 수를 줄여 **mon_initial_members** 옵션을 사용하여 쿼럼을 설정할 수 있습니다.*

3.5. CEPH 모니터 일관성

*Ceph 구성 파일에 모니터 설정을 추가하는 경우 Ceph 모니터의 아키텍처 측면을 알고 있어야 합니다. 클러스터 내에서 다른 Ceph 모니터를 검색할 때 Ceph 모니터에 대한 엄격한 일관성 요구 사항을 적용합니다. Ceph 클라이언트 및 기타 Ceph 데몬은 Ceph 구성 파일을 사용하여 모니터를 검색하는 반면, 모니터는 Ceph 구성 파일이 아닌 모니터 맵(**monmap**)을 사용하여 서로 검색합니다.*

Red Hat Ceph Storage 클러스터에서 다른 Ceph 모니터를 검색할 때 Ceph 모니터는 항상 모니터 맵의 로컬 사본을 참조합니다. Ceph 구성 파일 대신 모니터 맵을 사용하면 클러스터가 중단될 수 있는 오류가 방지됩니다. 예를 들어 모니터 주소 또는 포트를 지정할 때 Ceph 구성 파일의 오타입니다. 모니터 맵은 검색을 위해 모니터 맵을 사용하고 클라이언트 및 기타 Ceph 데몬과 모니터 맵을 공유하므로 모니터 맵은 모니터에 합의가 유효하다는 엄격한 보장을 제공합니다.

모니터 맵에 업데이트를 적용할 때 엄격한 일관성

Ceph 모니터의 다른 업데이트와 마찬가지로 모니터 맵에 대한 변경은 항상 Paxos라는 분산된 합의 알고리즘을 통해 실행됩니다. Ceph 모니터는 Ceph 모니터 추가 또는 제거와 같은 모니터 맵의 각 업데이트에 동의해야 쿼럼의 각 모니터 버전에 동일한 버전의 모니터 맵이 있는지 확인해야 합니다. 모니터 맵에 대한 업데이트는 점진적으로 되어 Ceph Monitor에 동의한 최신 버전 및 이전 버전 집합을 사용할 수 있습니다.

내역 유지

기록을 유지하면 이전 버전의 모니터 맵이 있는 Ceph Monitor가 Red Hat Ceph Storage 클러스터의 현재 상태를 따라잡을 수 있습니다.

모니터 맵을 통해 Ceph 구성 파일을 통해 Ceph 모니터를 서로 발견하면 Ceph 구성 파일이 자동으로 업데이트되고 배포되지 않기 때문에 추가 위험이 발생할 수 있습니다. Ceph 모니터는 이전 Ceph 구성 파일을 실수로 사용하거나, Ceph 모니터를 인식하지 못하거나, 쿼럼이 중단되거나, Paxos가 시스템의 현재 상태를 정확하게 확인할 수 없는 상황을 개발할 수 있습니다.

3.6. CEPH 모니터 부트스트랩

*대부분의 구성 및 배포 사례에서 Ceph를 배포하는 툴(예: **cephadm**)은 모니터 맵을 생성하여 Ceph 모니터를 부트스트랩하는 데 도움이 될 수 있습니다.*

Ceph 모니터에는 몇 가지 명시적 설정이 필요합니다.

- ***파일 시스템 ID**: **fsid** 는 오브젝트 저장소의 고유 식별자입니다. 동일한 하드웨어에서 여러 스토리지 클러스터를 실행할 수 있으므로 모니터를 부트스트랩할 때 오브젝트 저장소의 고유 ID를 지정해야 합니다. **cephadm** 과 같은 배포 도구를 사용하면 파일 시스템 식별자가 생성되지만 **fsid** 를 수동으로 지정할 수도 있습니다.*
- ***모니터 ID**: 모니터 ID는 클러스터 내의 각 모니터에 할당된 고유 ID입니다. 관례적으로 ID는 모니터의 호스트 이름으로 설정됩니다. 이 옵션은 **ceph** 명령 또는 Ceph 구성 파일에서 배포 툴을 사용하여 설정할 수 있습니다. Ceph 구성 파일에서 섹션은 다음과 같이 구성됩니다.*

  예제

  ```
  [mon.host1]

  [mon.host2]
  ```

- ***키**: 모니터에는 비밀 키가 있어야 합니다.*

추가 리소스

- ***cephadm** 및 Ceph 오케스트레이터에 대한 자세한 내용은 Red Hat Ceph Storage 운영 가이드를 참조하십시오.*

3.7. CEPH 모니터의 최소 구성

*Ceph 구성 파일의 Ceph 모니터 베어 최소 모니터 설정에는 DNS 및 모니터 주소에 대해 구성되지 않은 경우 각 모니터의 호스트 이름이 포함됩니다. Ceph 모니터는 기본적으로 포트 **6789** 및 **3300** 에서 실행됩니다.*

중요

Ceph 구성 파일을 편집하지 마십시오.

참고

*모니터의 최소 구성은 배포 툴에서 **fsid** 및 **mon.** 키를 생성한다고 가정합니다.*

다음 명령을 사용하여 스토리지 클러스터 구성 옵션을 설정하거나 읽을 수 있습니다.

- ***Ceph 구성 덤프** - 전체 스토리지 클러스터에 대한 전체 구성 데이터베이스를 덤프합니다.*
- ***Ceph config generate-minimal-conf** - 최소 **ceph.conf** 파일을 생성합니다.*
- ***Ceph config get** œ - Ceph 모니터의 구성 데이터베이스에 저장된 대로 특정 데몬 또는 클라이언트에 대한 구성을 덤프합니다.*
- ***Ceph 구성 세트 OPTION VALUE** - Ceph 모니터의 구성 데이터베이스에 구성 옵션을 설정합니다.*
- ***Ceph config show the** directory - 실행 중인 데몬에 대해 보고된 실행 구성을 표시합니다.*

- ***Ceph config assimilate-conf -i INPUT_FILE -o OUTPUT_FILE*** - *입력 파일에서 구성 파일을 가장하고 유효한 옵션을 Ceph Monitors의 구성 데이터베이스로 이동합니다.*

여기에서 는 섹션 또는 Ceph 데몬의 이름일 수 있으며 OPTION 은 구성 파일이며 VALUE 는 ***true*** *또는* ***false*** *일 수 있습니다.*

중요

구성 저장소에서 옵션을 가져오기 전에 Ceph 데몬에 config 옵션이 필요한 경우 다음 명령을 실행하여 구성을 설정할 수 있습니다.

```
ceph cephadm set-extra-ceph-conf
```

이 명령은 모든 데몬의 ***ceph.conf*** *파일에 텍스트를 추가합니다. 해결방법이며 권장되는 작업이 아닙니다.*

3.8. CEPH의 고유 식별자

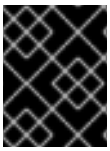
*각 Red Hat Ceph Storage 클러스터에는 고유한 식별자(fs****id)가****있습니다. 지정된 경우 일반적으로 구성 파일의* ***[global]*** *섹션 아래에 표시됩니다. 배포 도구는 일반적으로* ***fsid*** *를 생성하고 모니터 맵에 저장하므로 값이 구성 파일에 표시되지 않을 수 있습니다.* ***fsid*** *를 사용하면 동일한 하드웨어에서 여러 클러스터에 대해 데몬을 실행할 수 있습니다.*

참고

이를 수행하는 배포 툴을 사용하는 경우 이 값을 설정하지 마십시오.

3.9. CEPH 모니터 데이터 저장소

Ceph는 Ceph에서 데이터를 모니터링하는 기본 경로를 제공합니다.



중요

프로덕션 Red Hat Ceph Storage 클러스터에서 최적의 성능을 위해 Ceph OSD의 별도의 호스트 및 드라이브에서 Ceph 모니터를 실행하는 것이 좋습니다.

Ceph 모니터는 ***fsync()*** *함수를 자주 호출하여 Ceph OSD 워크로드를 방해할 수 있습니다.*

Ceph 모니터는 해당 데이터를 키-값 쌍으로 저장합니다. 데이터 저장소를 사용하면 Ceph 모니터에서 Paxos를 통해 손상된 버전을 복구할 수 없으며, 다른 이점 중에서 하나의 원자성 배치에서 여러 가지 수정 작업을 수행할 수 있습니다.

중요

Red Hat은 기본 데이터 위치 변경을 권장하지 않습니다. 기본 위치를 수정하면 구성 파일의 ***[mon]*** *섹션에 설정하여 Ceph 모니터를 균일하게 설정합니다.*

3.10. CEPH 스토리지 용량

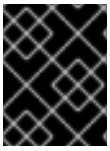
Red Hat Ceph Storage 클러스터가 최대 용량에 근접하게 되면(***mon_osd_full_ratio*** *매개 변수로 지정됨) Ceph는 데이터 손실을 방지하기 위해 Ceph OSD를 참조하거나 읽을 수 없습니다. 따라서 Red Hat Ceph Storage 클러스터가 전체 비율을 달성할 수 있도록 하는 것은 고가용성을 희생하기 때문에 좋은 방법이*

*아닙니다. 기본 전체 비율은 **.95** 또는 95% 용량입니다. 이 설정은 OSD 수가 적은 테스트 클러스터에 매우 공격적인 설정입니다.*

작은 정보

*클러스터를 모니터링할 때 **거의 전체** 비율과 관련된 경고에 경고해야 합니다. 즉, 하나 이상의 OSD가 실패하면 일부 OSD가 실패하면 서비스가 일시 중단될 수 있습니다. 스토리지 용량을 늘리려면 OSD를 추가하는 것이 좋습니다.*

테스트 클러스터의 일반적인 시나리오에는 시스템 관리자가 Red Hat Ceph Storage 클러스터에서 Ceph OSD를 제거하여 클러스터 재조정을 감시하는 작업이 포함됩니다. 그런 다음 Red Hat Ceph Storage 클러스터가 전체 비율 및 잠금에 도달할 때까지 다른 Ceph OSD를 제거합니다.



중요

Red Hat은 테스트 클러스터에서도 약간의 용량 계획을 권장합니다. 계획을 사용하면 고가용성을 유지하기 위해 필요한 예비 용량을 측정할 수 있습니다.

*클러스터가 Ceph OSD를 즉시 교체하지 않고 **활성 + 정리** 상태로 복구할 수 있는 일련의 Ceph OSD 실패를 계획하는 것이 좋습니다. **활성 + 성능이 저하된** 상태에서 클러스터를 실행할 수 있지만 일반적인 운영 조건에는 이상적이지 않습니다.*

*다음 다이어그램에서는 호스트당 하나의 Ceph OSD가 있는 33개의 Ceph 노드가 포함되어 있으며 각 Ceph OSD 데몬이 3TB 드라이브에서 읽고 쓰는 각 Ceph OSD 데몬이 포함된 간단한 Red Hat Ceph Storage 클러스터를 보여줍니다. 따라서 이 예제의 Red Hat Ceph Storage 클러스터는 최대 99TB의 실제 용량을 갖습니다. mond **의 전체 비율이** 4.85인 **경우** Red Hat Ceph Storage 클러스터가 5TB의 나머지 용량으로 떨어지면 클러스터에서 Ceph 클라이언트가 데이터를 읽고 쓸 수 없습니다. 따라서 Red Hat Ceph Storage 클러스터의 운영 용량은 99TB가 아닌 95TB입니다.*

*이러한 클러스터의 경우 하나 또는 두 개의 OSD가 실패하는 것이 일반적입니다. 덜 빈번하지만 합리적인 시나리오에서는 랙의 라우터 또는 전원 공급이 실패하여 여러 OSD(예: OSD 7-12)가 동시에 중단됩니다. 이러한 시나리오에서는 작동 상태를 유지하고 **활성 + 클린** 상태가 될 수 있는 클러스터를 위해 계속 노력해야 합니다. 즉, 추가 OSD가 있는 몇 개의 호스트를 짧은 순서로 추가하는 경우에도 마찬가지입니다. 용*

량 사용률이 너무 높으면 데이터가 손실되지 않을 수 있지만 클러스터의 용량 사용률이 전체 비율을 초과하면 장애 도메인 내에서 중단을 해결하는 동안 데이터 가용성을 희생할 수 있습니다. 이러한 이유로 최소한 대략적인 용량 계획을 수립하는 것이 좋습니다.

클러스터의 두 숫자를 확인합니다.

- *OSD 수*
- *클러스터의 총 용량*

클러스터 내에서 OSD의 평균 용량을 결정하려면 클러스터의 총 용량을 클러스터의 OSD 수로 나눕니다. 일반 작업(이 비교적 적은 수) 동안 동시에 실패할 것으로 예상되는 OSD 수를 곱하는 것이 좋습니다. 마지막으로 최대 운영 용량에 도달하기 위해 클러스터 용량에 전체 비율을 곱합니다. 그런 다음 적절한 전체 비율에 도달하지 못하는 OSD에서 데이터 양을 뺀다. OSD 실패 수가 많은 수(예: OSD 랙)를 통해 포지셔닝 프로세스를 반복하여 거의 전체 비율을 위해 적절한 개수로 도착합니다.

3.11. CEPH HEARTBEAT

Ceph는 각 OSD에서 보고가 필요하고 인접 OSD의 상태에 대해 OSD에서 보고서를 수신하여 클러스터에 대해 알 수 있습니다. Ceph는 모니터와 OSD 간의 상호 작용을 위한 적절한 기본 설정을 제공하지만 필요에 따라 수정할 수 있습니다.

3.12. CEPH MONITOR 동기화 역할

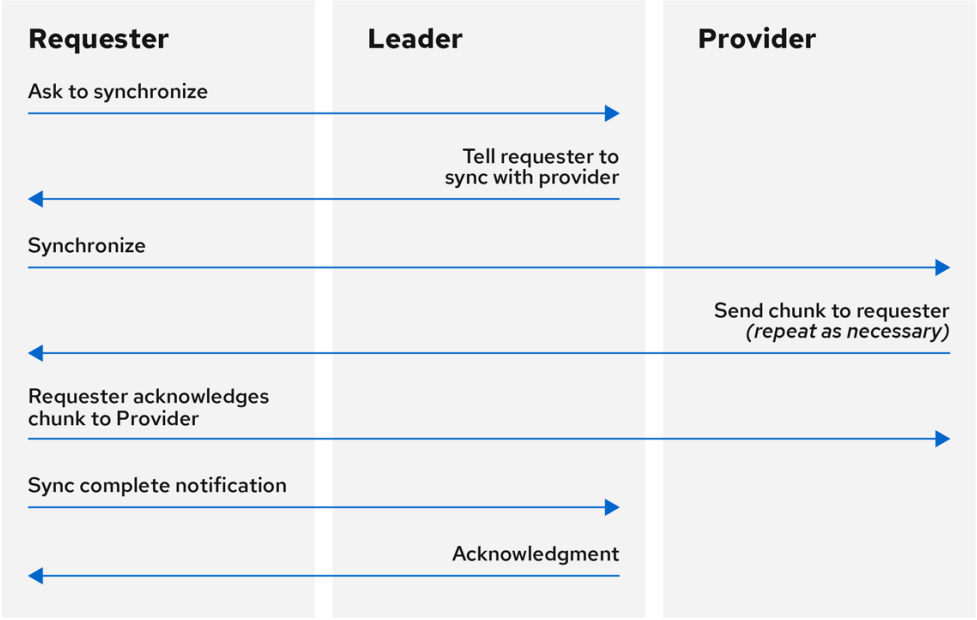
여러 모니터가 있는 프로덕션 클러스터를 실행하는 경우 각 모니터는 인접 모니터에 더 최신 버전의 클러스터 맵이 있는지 확인합니다. 예를 들어, 인스턴트 모니터 맵의 최신 epoch보다 하나 이상의 epoch 번호가 있는 인접 모니터에 있는 맵입니다. 주기적으로 클러스터의 한 모니터가 쿼럼을 벗어나야 하는 지점으로 떨어지고, 동기화하여 클러스터에 대한 최신 정보를 검색한 다음 쿼럼에 다시 참여할 수 있습니다.

동기화 역할

동기화 목적을 위해 모니터는 다음 세 가지 역할 중 하나를 가정할 수 있습니다.

- ***리더**: Leader는 클러스터 맵의 최신 Paxos 버전을 달성하는 첫 번째 모니터입니다.*
- ***공급자**: 공급자는 최신 버전의 클러스터 맵이 있지만 가장 최신 버전을 제공하는 데 첫 번째 버전이 아닌 모니터입니다.*
- ***요청자:** 요청자는 리더 뒤에 떨어지고 쿼럼에 다시 참여하기 전에 클러스터에 대한 최신 정보를 검색하기 위해 동기화해야 하는 모니터입니다.*

이러한 역할을 사용하면 리더가 동기화 작업을 프로바이더에 위임하여 동기화 요청이 리더에 과부하되고 성능이 향상되지 않도록 할 수 있습니다. 다음 다이어그램에서 요청자는 다른 모니터 뒤에서 떨어지는 것을 배웠습니다. 요청자는 리더에게 동기화하도록 요청하고 리더는 요청자에게 공급자와 동기화하도록 지시합니다.



동기화 모니터링

새 모니터가 클러스터에 참여하면 동기화가 항상 발생합니다. 런타임 작업 중에 모니터는 다른 시간에 클러스터 맵에 대한 업데이트를 수신할 수 있습니다. 즉, leader 및 provider 역할은 한 모니터에서 다른 모니터로 마이그레이션할 수 있습니다. 예를 들어 동기화 중에 공급자가 리더 뒤에 속하는 경우 공급자는 요청자와 동기화를 종료할 수 있습니다.

동기화가 완료되면 Ceph에서 클러스터 전체에서 트리밍해야 합니다. 트리밍을 수행하려면 배치 그룹이 ***활성 + 정리되어야 합니다****.*

3.13. CEPH 시간 동기화

Ceph 데몬은 중요한 메시지를 서로 전달합니다. 이 메시지는 데몬이 시간 제한 임계값에 도달하기 전에 처리해야 합니다. Ceph 모니터의 시계가 동기화되지 않으면 여러 예외가 발생할 수 있습니다.

예를 들면 다음과 같습니다.

- *수신된 메시지를 무시하는 데몬(예: 오래된 타임스탬프).*
- *메시지가 시간 내에 수신되지 않았을 때 너무 빨리 지연이 트리거되었습니다.*

작은 정보

Ceph 모니터 호스트에 NTP를 설치하여 모니터 클러스터가 동기화된 클록으로 작동하는지 확인합니다.

불일치가 아직 취약하지 않더라도 NTP에서는 클럭 드리프트가 여전히 눈에 띄게 표시될 수 있습니다. NTP가 적절한 수준의 동기화를 유지하지만 Ceph 클럭 드리프트와 클럭 스큐 경고가 트리거될 수 있습니다. 이러한 상황에서 클럭 드리프트를 늘리는 것은 허용되지 않을 수 있습니다. 그러나 워크로드, 네트워크 대기 시간, 기본 시간 초과 구성 및 기타 동기화 옵션과 같은 여러 요소가 Paxos 보장을 손상시키지 않고 허용되는 클럭 드리프트 수준에 영향을 미칠 수 있습니다.

추가 리소스

- *자세한 내용은 Ceph 시간 동기화 섹션을 참조하십시오.*

3.14. 추가 리소스

- *특정 옵션 설명 및 사용에 대해서는 부록 C 의 모든 Red Hat Ceph Storage Monitor 구성 옵션을 참조하십시오.*

4장. CEPH 인증 구성

스토리지 관리자는 Red Hat Ceph Storage 클러스터의 보안에 사용자와 서비스를 인증하는 것이 중요합니다. Red Hat Ceph Storage에는 Cephx 프로토콜, 기본 암호화 인증, 스토리지 클러스터에서 인증을 관리하는 툴이 포함되어 있습니다.

4.1. 사전 요구 사항

- Red Hat Ceph Storage 소프트웨어 설치.

4.2. CEPHX 인증

cephx 프로토콜은 기본적으로 활성화되어 있습니다. 암호화 인증에는 일반적으로 상당히 낮은 컴퓨팅 비용이 있지만 몇 가지 컴퓨팅 비용이 있습니다. 클라이언트와 호스트를 연결하는 네트워크 환경이 안전하다고 간주되고 인증 컴퓨팅 비용을 감당할 수 없는 경우 이를 비활성화할 수 있습니다. Ceph 스토리지 클러스터를 배포할 때 배포 툴은 **client.admin** 사용자와 인증 키를 생성합니다.

중요

Red Hat은 인증 사용을 권장합니다.

참고

인증을 비활성화하면 중간자 공격이 클라이언트 및 서버 메시지를 변경하여 심각한 보안 문제로 이어질 수 있습니다.

Cephx 활성화 및 비활성화

Cephx를 활성화하려면 Ceph 모니터 및 OSD용 키를 배포해야 합니다. Cephx 인증을 켜거나 끄는 경우 배포 절차를 반복할 필요가 없습니다.

4.3. CEPHX 활성화

cephx 가 활성화되면 Ceph는 **/etc/ceph/$cluster.$name.keyring** 이 포함된 기본 검색 경로에서 인증 키를 찾습니다. Ceph 구성 파일의 **[global]** 섹션에 **인증 키** 옵션을 추가하여 이 위치를 재정의할 수 있지만 권장되지는 않습니다.

인증이 비활성화된 클러스터에서 **cephx** 를 활성화하려면 다음 절차를 실행합니다. 귀하 또는 배포 유틸리티에서 키를 이미 생성한 경우, 키 생성과 관련된 단계를 건너뛸 수 있습니다.

사전 요구 사항

- 실행 중인 Red Hat Ceph Storage 클러스터.
- Ceph 모니터 노드에 대한 루트 수준 액세스.

절차

1. **client.admin 키를 생성하고 클라이언트** 호스트에 대한 키 사본을 저장합니다.

```
[root@mon ~]# ceph auth get-or-create client.admin mon 'allow *' osd 'allow *' -o
/etc/ceph/ceph.client.admin.keyring
```

주의

*이렇게 하면 기존 **/etc/ceph/client.admin.keyring** 파일의 내용이 지워집니다. 배포 툴이 이미 수행된 경우 이 단계를 수행하지 마십시오.*

2. *모니터 클러스터에 대한 인증 키를 생성하고 모니터 시크릿 키를 생성합니다.*

```
[root@mon ~]# ceph-authtool --create-keyring /tmp/ceph.mon.keyring --gen-key -n mon. --
cap mon 'allow *'
```

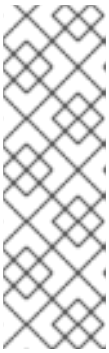
3. *모니터 인증 키를 모니터 **mon 데이터** 디렉터리의 **ceph.mon.keyring 파일에 복사합니다**. 예를 들어 클러스터 **ceph** 의 **mon.a** 에 복사하려면 다음을 사용합니다.*

```
[root@mon ~]# cp /tmp/ceph.mon.keyring /var/lib/ceph/mon/ceph-a/keyring
```

4. *모든 OSD에 대한 비밀 키를 생성합니다. 여기서 **ID** 는 OSD 번호입니다.*

```
ceph auth get-or-create osd.ID mon 'allow rwx' osd 'allow *' -o
/var/lib/ceph/osd/ceph-ID/keyring
```

5. *기본적으로 **cephx** 인증 프로토콜은 활성화됩니다.*



참고

*이전에 인증 옵션을 **none** 으로 설정하여 **cephx** 인증 프로토콜이 비활성화된 경우 Ceph 구성 파일(**/etc/ceph/ceph.conf**)의 **[global] 섹션에서 다음 행을 제거하면 cephx 인증 프로토콜을 다시 활성화합니다.***

```
auth_cluster_required = none
auth_service_required = none
auth_client_required = none
```

6. *Ceph 스토리지 클러스터를 시작하거나 다시 시작합니다.*

중요

***cephx** 를 활성화하려면 클러스터를 완전히 다시 시작해야 하거나 클라이언트 I/O 가 비활성화된 동안 해당 클러스터를 종료한 다음 시작해야 하므로 다운타임이 필요합니다.*

스토리지 클러스터를 재시작하거나 종료하기 전에 이러한 플래그를 설정해야 합니다.

```
[root@mon ~]# ceph osd set noout
[root@mon ~]# ceph osd set norecover
[root@mon ~]# ceph osd set norebalance
[root@mon ~]# ceph osd set nobackfill
[root@mon ~]# ceph osd set nodown
[root@mon ~]# ceph osd set pause
```

***cephx** 가 활성화되고 모든 PG가 활성 상태이고 정리되면 플래그를 설정 해제합니다.*

```
[root@mon ~]# ceph osd unset noout
[root@mon ~]# ceph osd unset norecover
[root@mon ~]# ceph osd unset norebalance
[root@mon ~]# ceph osd unset nobackfill
[root@mon ~]# ceph osd unset nodown
[root@mon ~]# ceph osd unset pause
```

4.4. CEPHX 비활성화

다음 절차에서는 Cephx를 비활성화하는 방법을 설명합니다. 클러스터 환경이 상대적으로 안전할 경우 인증을 실행하는 데 드는 계산 비용을 상쇄할 수 있습니다.

중요

Red Hat은 인증 활성화를 권장합니다.

그러나 설정 또는 문제 해결 중에 인증을 일시적으로 비활성화하는 것이 더 쉬울 수 있습니다.

사전 요구 사항

- *실행 중인 Red Hat Ceph Storage 클러스터.*
- *Ceph 모니터 노드에 대한 루트 수준 액세스.*

절차

1. *Ceph 구성 파일의 **[global]** 섹션에 다음 옵션을 설정하여 **cephx** 인증을 비활성화합니다.*

   예제

   ```
   auth_cluster_required = none
   auth_service_required = none
   auth_client_required = none
   ```

2. Ceph 스토리지 클러스터를 시작하거나 다시 시작합니다.

4.5. CEPHX 사용자 인증 키

인증을 활성화한 상태에서 Ceph를 실행하면 Ceph **관리** 명령 및 Ceph 클라이언트가 Ceph 스토리지 클러스터에 액세스하려면 인증 키가 필요합니다.

이러한 키를 **ceph** 관리 명령 및 클라이언트에 제공하는 가장 일반적인 방법은 **/etc/ceph/** 디렉터리에 Ceph 인증 키를 포함하는 것입니다. 파일 이름은 일반적으로 **ceph.client.admin.keyring** 또는 **$cluster.client.admin.keyring** 입니다. **/etc/ceph/** 디렉터리에 인증 키를 포함하는 경우 Ceph 구성 파일에 **인증 키** 항목을 지정할 필요가 없습니다.

중요

client.admin 키가 포함되어 있으므로 관리 명령을 실행할 노드에 Red Hat Ceph Storage 클러스터 인증 키 파일을 복사하는 것이 좋습니다.

이를 수행하려면 다음 명령을 실행합니다.

```
# scp USER@HOSTNAME:/etc/ceph/ceph.client.admin.keyring /etc/ceph/ceph.client.admin.keyring
```

USER 를 호스트에 사용된 사용자 이름으로 **client.admin** 키로 변경하고 **HOSTNAME** 을 해당 호스트 이름으로 바꿉니다.

참고

ceph.keyring 파일에 클라이언트 시스템에 설정된 적절한 권한이 있는지 확인합니다.

권장되지 않는 키 설정을 사용하여 Ceph 구성 파일에서 키 자체를 지정하거나 **key file** 설정을 사용하여 키 파일에 대한 경로를 지정할 수 있습니다.

4.6. CEPHX 데몬 인증 키

관리 사용자 또는 배포 툴은 사용자 인증 키를 생성하는 것과 동일한 방식으로 데몬 인증 키를 생성할 수 있습니다. 기본적으로 Ceph는 데이터 디렉터리 내부에 데몬 인증 키를 저장합니다. 기본 인증 키 위치 및 데몬이 작동하는 데 필요한 기능.

참고

모니터 인증 키에는 키가 포함되지만 기능은 없으며 Ceph 스토리지 클러스터 **인증** 데이터베이스의 일부가 아닙니다.

데몬 데이터 디렉터리 위치는 기본적으로 양식의 디렉터리입니다.

```
/var/lib/ceph/$type/CLUSTER-ID
```

예제

```
/var/lib/ceph/osd/ceph-12
```

이러한 위치를 재정의할 수 있지만 권장되지는 않습니다.

4.7. CEPHX 메시지 서명

Ceph는 세부적인 제어를 제공하므로 클라이언트와 Ceph 간에 서비스 메시지의 서명을 활성화하거나 비활성화할 수 있습니다. Ceph 데몬 간 메시지의 서명을 활성화하거나 비활성화할 수 있습니다.

중요

Red Hat은 Ceph에서 초기 인증에 설정된 세션 키를 사용하여 엔터티 간의 모든 진행 중인 메시지를 인증하는 것이 좋습니다.

참고

Ceph 커널 모듈은 서명을 아직 지원하지 않습니다.

4.8. 추가 리소스

- 특정 옵션 설명 및 사용법에 대해서는 부록 D 의 모든 Red Hat Ceph Storage Cephx 구성 옵션을 참조하십시오.

5장. 풀, 배치 그룹 및 CRUSH 구성

스토리지 관리자는 풀, 배치 그룹 및 CRUSH 알고리즘에 Red Hat Ceph Storage 기본 옵션을 사용하거나 원하는 워크로드에 맞게 사용자 지정할 수 있습니다.

5.1. 사전 요구 사항

- Red Hat Ceph Storage 소프트웨어 설치.

5.2. 풀 배치 그룹 및 CRUSH

풀을 생성하고 풀의 배치 그룹 수를 설정하면 Ceph는 기본값을 구체적으로 재정의하지 않을 때 기본값을 사용합니다.

중요

일부 기본값을 재정의하는 것이 좋습니다. 특히 풀의 복제본 크기를 설정하고 기본 배치 그룹 수를 재정의합니다.

pool 명령을 실행할 때 이러한 값을 설정할 수 있습니다.

기본적으로 Ceph는 3개의 오브젝트 복제본을 만듭니다. 오브젝트의 복사본 4개를 기본값, 기본 사본 및 3개의 복제본 사본으로 설정하려면 **osd_pool_default_size** 에 표시된 대로 기본값을 재설정합니다. Ceph가 성능 저하 상태에 적은 수의 복사본을 작성하도록 허용하려면 **osd_pool_default_min_size** 를 **osd_pool_default_size** 값보다 작은 숫자로 설정합니다.

예제

```
[ceph: root@host01 /]# ceph config set global osd_pool_default_size 4  # Write an object 4 times.
[ceph: root@host01 /]# ceph config set global osd_pool_default_min_size 1  # Allow writing one copy in a degraded state.
```

실제 개수의 배치 그룹이 있는지 확인합니다. Red Hat은 OSD당 약 100개를 권장합니다. 예를 들어 OSD의 총 수가 100을 복제본 수, 즉 **osd_pool_default_size** 로 나눈 값입니다. 10개의 OSD 및 **osd_pool_default_size** = 4의 경우 약 100 * 10) / 4 = 250을 권장합니다.

예제

```
[ceph: root@host01 /]# ceph config set global osd_pool_default_pg_num 250
[ceph: root@host01 /]# ceph config set global osd_pool_default_pgp_num 250
```

5.3. 추가 리소스

- 부록 E 에서 특정 옵션 설명 및 사용법에 대한 모든 Red Hat Ceph Storage 풀, 배치 그룹, CRUSH 구성 옵션을 참조하십시오.

6장. CEPH OSD(오브젝트 스토리지 데몬) 구성

스토리지 관리자는 원하는 워크로드에 따라 Ceph OSD(오브젝트 스토리지 데몬)를 중복 및 최적화하도록 구성할 수 있습니다.

6.1. 사전 요구 사항

- Red Hat Ceph Storage 소프트웨어 설치.

6.2. CEPH OSD 구성

모든 Ceph 클러스터에는 다음을 정의하는 구성이 있습니다.

- 클러스터 ID
- 인증 설정
- 클러스터의 Ceph 데몬 멤버십
- 네트워크 설정
- 호스트 이름 및 주소
- 인증 키 경로
- OSD 로그 파일의 경로
- 기타 런타임 옵션

cephadm 과 같은 배포 도구는 일반적으로 사용자를 위한 초기 Ceph 구성 파일을 만듭니다. 그러나 배포 툴을 사용하지 않고 클러스터를 부트스트랩하려는 경우 직접 클러스터를 생성할 수 있습니다.

편의를 위해 각 데몬에는 일련의 기본값이 있습니다. 대부분 **ceph/src/common/config_opts.h** 스크립트에 의해 설정됩니다. monitor **tell** 명령을 사용하거나 Ceph 노드의 데몬 소켓에 직접 연결하여 이러한 설정을 Ceph 구성 파일로 재정의하거나 런타임 시 이를 재정의할 수 있습니다.

중요

Red Hat은 Ceph 문제를 나중에 해결하기가 더 어렵기 때문에 기본 경로를 변경하는 것을 권장하지 않습니다.

추가 리소스

- **cephadm** 및 Ceph 오케스트레이터에 대한 자세한 내용은 Red Hat Ceph Storage 운영 가이드를 참조하십시오.

6.3. OSD 스크럽

Ceph는 개체 복사본을 여러 개 만드는 것 외에도 배치 그룹을 수정하여 데이터 무결성을 보장합니다. Ceph 스크럽은 오브젝트 스토리지 계층의 **fsck** 명령과 유사합니다.

Ceph는 각 배치 그룹에 대해 모든 오브젝트의 카탈로그를 생성하고 각 기본 오브젝트와 해당 복제본을 비교하여 오브젝트가 누락되거나 일치하지 않는지 확인합니다.

가벼운 스크럽(daily)은 개체 크기 및 특성을 확인합니다. 심층 스크럽(주별)은 데이터를 읽고 체크섬을 사용하여 데이터 무결성을 보장합니다.

스크럽은 데이터 무결성을 유지하는 데 중요하지만 성능을 줄일 수 있습니다. 다음 설정을 조정하여 스크럽 작업을 늘리거나 줄입니다.

추가 리소스

- 자세한 내용은 Red Hat Ceph Storage Configuration Guide 의 부록에서 Ceph 스크럽 옵션을 참조하십시오.

6.4. OSD 백필

클러스터에 Ceph OSD를 추가하거나 클러스터에서 제거하면 CRUSH 알고리즘에서 배치 그룹을 Ceph OSD로 이동하거나 균형을 복원하여 클러스터를 리밸런싱합니다. 배치 그룹 및 오브젝트를 마이그레이션하는 프로세스는 클러스터 운영 성능을 크게 줄일 수 있습니다. 운영 성능을 유지하기 위해 Ceph는 'backfill' 프로세스를 사용하여 이 마이그레이션을 수행합니다. 이 프로세스를 사용하면 Ceph에서 백필 작업을 요청에서 읽기 또는 쓰기 데이터에 대한 우선 순위보다 낮은 우선 순위로 설정할 수 있습니다.

6.5. OSD 복구

클러스터가 시작되거나 Ceph OSD가 예기치 않게 종료되면 OSD는 쓰기 작업이 발생하기 전에 다른 Ceph OSD와 피어링하기 시작합니다.

Ceph OSD가 충돌하여 다시 온라인 상태가 되면 일반적으로 배치 그룹의 최신 오브젝트 버전이 포함된 다른 Ceph OSD와 동기화되지 않습니다. 이 경우 Ceph OSD는 복구 모드로 전환되고 최신 데이터 복사본을 가져와서 해당 맵을 최신 상태로 가져오려고 합니다. Ceph OSD가 다운된 기간에 따라 OSD의 개체 및 배치 그룹이 상당히 오래되었을 수 있습니다. 또한 장애 도메인(예: 랙)이 하나 이상의 Ceph OSD가 동시에 온라인 상태가 될 수 있습니다. 이를 통해 복구 프로세스 시간이 많이 소요되고 리소스를 많이 사용할 수 있습니다.

운영 성능을 유지하기 위해 Ceph는 복구 요청, 스레드 및 개체 청크 크기를 제한하여 Ceph가 성능이 저하된 상태에서 제대로 수행할 수 있도록 합니다.

6.6. 추가 리소스

- 특정 옵션 설명 및 사용에 대해서는 부록 F 의 모든 Red Hat Ceph Storage Ceph OSD 구성 옵션을 참조하십시오.

7장. CEPH 모니터 및 OSD 상호 작용 구성

스토리지 관리자는 안정적인 작업 환경을 보장하기 위해 Ceph 모니터와 OSD 간의 상호 작용을 올바르게 구성해야 합니다.

7.1. 사전 요구 사항

- Red Hat Ceph Storage 소프트웨어 설치.

7.2. CEPH 모니터 및 OSD 상호 작용

초기 Ceph 구성을 완료한 후에는 Ceph를 배포하고 실행할 수 있습니다. **ceph 상태 또는 ceph -s** 와 같은 명령을 실행하면 Ceph Monitor에서 Ceph 스토리지 클러스터의 현재 상태를 보고합니다. Ceph Monitor는 각 Ceph OSD 데몬에서 보고서를 필요로 하고, Ceph OSD 데몬에서 옆에 있는 Ceph OSD 데몬의 상태에 대한 보고서를 수신하여 Ceph 스토리지 클러스터에 대해 알고 있습니다. Ceph Monitor에서 보고서를 받지 않거나 Ceph 스토리지 클러스터의 변경 사항 보고서를 수신하는 경우 Ceph Monitor에서 Ceph 클러스터 맵의 상태를 업데이트합니다.

Ceph는 Ceph 모니터 및 OSD 상호 작용을 위한 적절한 기본 설정을 제공합니다. 그러나 기본값을 재정의할 수 있습니다. 다음 섹션에서는 Ceph 스토리지 클러스터를 모니터링하기 위해 Ceph Monitors 및 Ceph OSD 데몬이 상호 작용하는 방법에 대해 설명합니다.

7.3. OSD 하트비트

각 Ceph OSD 데몬은 6초마다 다른 Ceph OSD 데몬의 하트비트를 확인합니다. 하트비트 간격을 변경하려면 런타임 시 값을 변경합니다.

구문

```
ceph config set osd osd_heartbeat_interval TIME_IN_SECONDS
```

예제

```
[ceph: root@host01 /]# ceph config set osd osd_heartbeat_interval 60
```

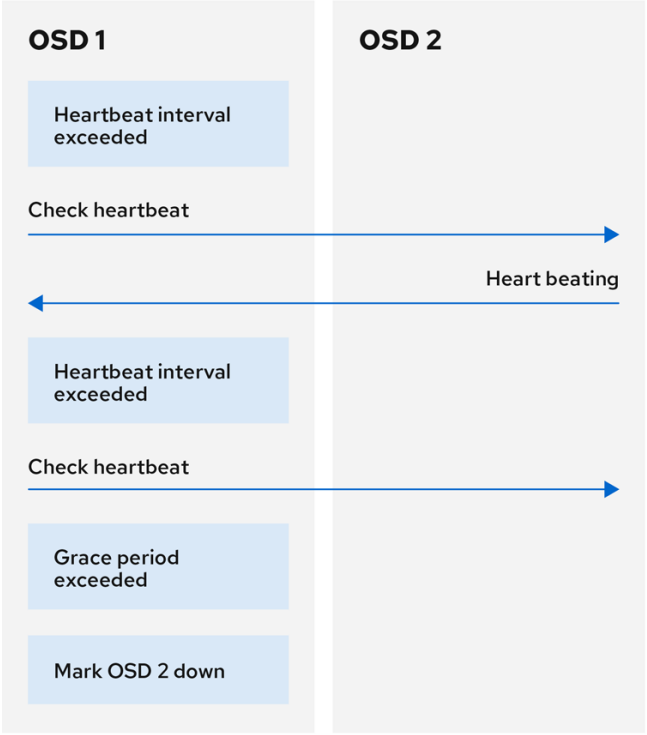
인접 Ceph OSD 데몬이 20초의 유예 기간 내에 하트비트 패킷을 전송하지 않으면 Ceph OSD 데몬에서 인접 Ceph OSD 데몬을 **중단할** 수 있습니다. Ceph 클러스터 맵을 업데이트하는 Ceph Monitor에 다시 보고할 수 있습니다. 유예 기간을 변경하려면 런타임 시 값을 설정합니다.

구문

```
ceph config set osd osd_heartbeat_grace TIME_IN_SECONDS
```

예제

```
[ceph: root@host01 /]# ceph config set osd osd_heartbeat_grace 30
```



7.4. OSD를 아래로 보고

기본적으로 **다른 호스트의** Ceph OSD 데몬 2개가 Ceph OSD 데몬이 다운되었음을 확인하기 전에 Ceph OSD 데몬이 **다운** 되었음을 보고해야 합니다 **.**

그러나 모든 OSD에서 오류를 보고하는 것은 잘못된 스위치로 인해 OSD 간 연결 문제가 발생하는 랙의 다른 호스트에 있을 가능성이 있습니다.

Ceph는 "false 알람"을 방지하기 위해 피어가 실패를 "하위 클러스터"의 프록시로 보고하는 것으로 간주합니다. 이러한 경우가 항상인 것은 아니지만 관리자가 제대로 수행하지 못하는 시스템의 하위 집합에 유예 수정 사항을 로컬라이제이션하는 데 도움이 될 수 있습니다.

Ceph는 **mon_osd_reporter_subtree_level** 설정을 사용하여 CRUSH 맵의 공통 부모 유형별로 피어를 "subcluster"로 그룹화합니다.

기본적으로 다른 Ceph OSD Daemon을 보고 **하려면 다른 하위 트리** 의 보고서 2개만 필요합니다 **.** 관리자는 런타임 시 mon_osd_ **min_ down _reporter 및 mon_osd_reporter_subtree_level 값을 설정하여 Ceph OSD 데몬을 Ceph Monitor로 보고하는** 데 필요한 고유한 하위 트리와 공통 상위 유형에서 reporters의 수를 변경할 수 있습니다.

구문

```
ceph config set mon mon_osd_min_down_reporters NUMBER
```

예제

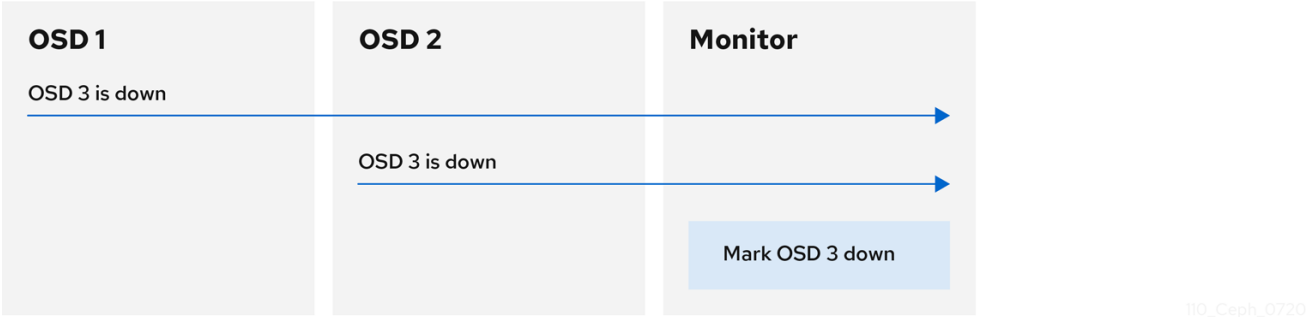
```
[ceph: root@host01 /]# ceph config set mon mon_osd_min_down_reporters 4
```

구문

```
ceph config set mon mon_osd_reporter_subtree_level CRUSH_ITEM
```

예제

```
[ceph: root@host01 /]# ceph config set mon mon_osd_reporter_subtree_level host
[ceph: root@host01 /]# ceph config set mon mon_osd_reporter_subtree_level rack
[ceph: root@host01 /]# ceph config set mon mon_osd_reporter_subtree_level osd
```



7.5. 피어링 실패 보고

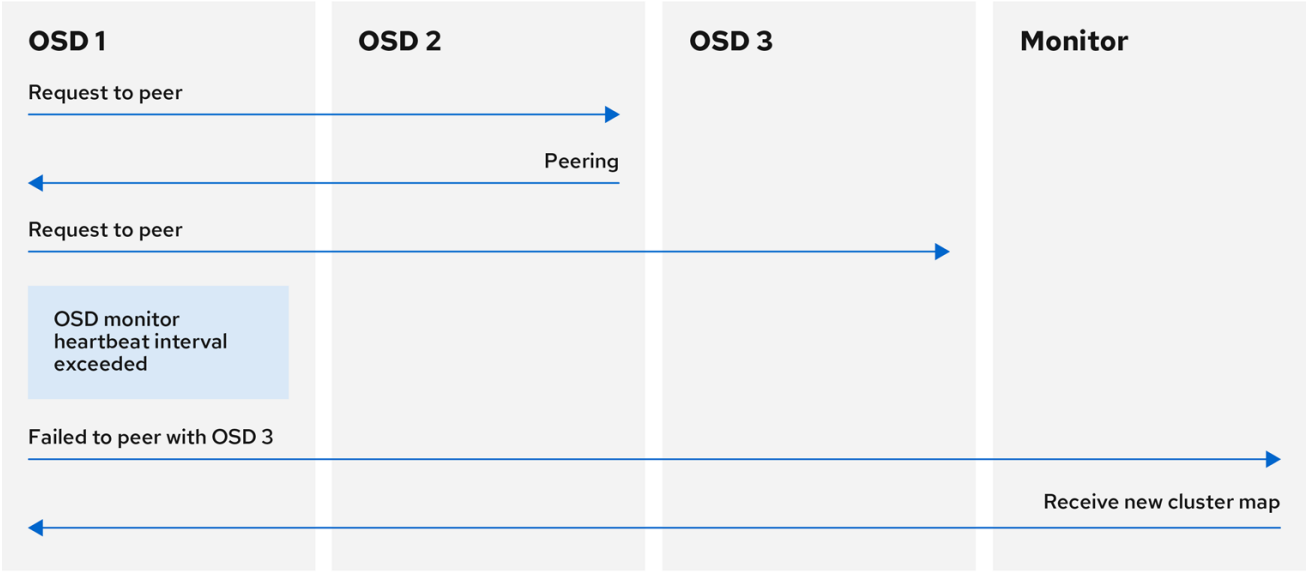
Ceph OSD 데몬이 Ceph 구성 파일 또는 클러스터 맵에 정의된 Ceph OSD 데몬과 피어링할 수 없는 경우 30초마다 클러스터 맵의 최신 복사본에 대해 Ceph Monitor를 ping합니다. 런타임 시 값을 설정하여 Ceph Monitor 하트비트 간격을 변경할 수 있습니다.

구문

```
ceph config set osd osd_mon_heartbeat_interval TIME_IN_SECONDS
```

예제

```
[ceph: root@host01 /]# ceph config set osd osd_mon_heartbeat_interval 60
```



7.6. OSD 보고 상태

Ceph OSD 데몬이 Ceph 모니터에 보고되지 않는 경우, Ceph 모니터는 **mon_osd_report_timeout** 후 Ceph OSD 데몬을 표시합니다. 이 시간은 900초입니다. **Ceph OSD 데몬은 실패와 같은 보고 가능한 이벤트, 배치 그룹 통계 변경, up_thru 가 변경되거나 5초 이내에 부팅되는 경우 Ceph 모니터로 보고서를 보냅니다.**

런타임에 osd_mon_report_interval 값을 설정하여 Ceph OSD 데몬 최소 보고서 간격을 변경할 수 있습니다.

구문

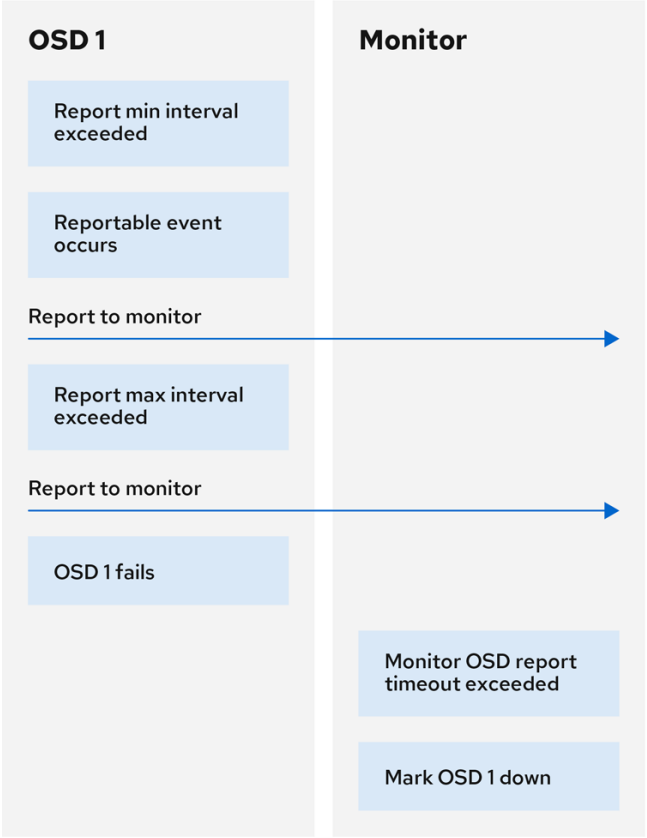
```
ceph config set osd osd_mon_report_interval TIME_IN_SECONDS
```

다음 예제를 사용하여 구성을 가져오고, 설정하고, 확인할 수 있습니다.

예제

```
[ceph: root@host01 /]# ceph config get osd osd_mon_report_interval
5
[ceph: root@host01 /]# ceph config set osd osd_mon_report_interval 20
[ceph: root@host01 /]# ceph config dump | grep osd

global                advanced  osd_pool_default_crush_rule              -1
  osd                 basic     osd_memory_target                        4294967296
  osd                 advanced  osd_mon_report_interval                  20
```



7.7. 추가 리소스

- 특정 옵션 설명 및 사용에 대해서는 부록 G 의 모든 **Red Hat Ceph Storage Ceph Monitor** 및 **OSD** 구성 옵션을 참조하십시오.

8장. CEPH 디버깅 및 로깅 구성

스토리지 관리자는 **cephadm** 에서 디버깅 및 로깅 정보의 양을 늘려 **Red Hat Ceph Storage** 문제를 진단할 수 있습니다.

8.1. 사전 요구 사항

- **Red Hat Ceph Storage** 소프트웨어가 설치되어 있습니다.

8.2. 추가 리소스

- **cephadm** 문제 해결에 대한 자세한 내용은 **Red Hat Ceph Storage** 관리 가이드에서 **Cephadm** 문제 해결을 참조하십시오.

- **cephadm** 로깅에 대한 자세한 내용은 **Red Hat Ceph Storage** 관리 가이드의 **Cephadm** 작업을 참조하십시오.

부록 A. 일반 구성 옵션

다음은 Ceph의 일반적인 구성 옵션입니다.

참고

일반적으로 cephadm 과 같은 배포 도구를 통해 자동으로 설정됩니다.

fsid

설명

파일 시스템 ID입니다. 클러스터당 하나씩.

유형

UUID

필수 항목

아니요.

Default

해당 없음. 일반적으로 배포 도구로 생성됩니다.

admin_socket

설명

Ceph 모니터가 쿼럼을 설정했는지 여부와 관계없이 데몬에서 관리 명령을 실행하는 소켓입니다.

유형

문자열

필수 항목

없음

Default

/var/run/ceph/$cluster-$name.asok

pid_file

설명

*모니터 또는 **OSD**가 **PID**를 쓰는 파일입니다. 예를 들어, **/var/run/$cluster/$type.$id.pid** 는 **ceph** 클러스터에서 **id**가 실행 중인 **mon** 에 대해 **/var/run/ceph/mon.a.pid**를 만듭니다. **pid** 파일은 데몬이 정상적으로 중지되면 제거됩니다. 프로세스가 데몬화되지 않은 경우(**- f** 또는 **- d** 옵션으로 실행됨) **pid** 파일이 생성되지 않습니다.*

유형

문자열

필수 항목

없음

Default

없음

chdir

설명

*디렉터리 **Ceph** 데몬은 실행 중이 되면 해당 데몬이 로 변경됩니다. 기본 / 디렉토리 권장.*

유형

문자열

필수 항목

없음

Default

/

max_open_files

설명

*설정된 경우 **Red Hat Ceph Storage** 클러스터가 시작될 때 **Ceph**는 **OS** 수준에서 **max_open_fds** 를 설정합니다(즉, 최대 **#**의 파일 설명자). **Ceph OSD**가 파일 설명자를 실행하지 못하도록 합니다.*

유형

64비트 정수

필수 항목

없음

Default

0

fatal_signal_handlers

설명

*설정된 경우 **SEGV, ABRT, BUS, ILL, FPE, XCPU, XFSZ, SYS** 신호에 대한 신호 처리기를 설치하여 유용한 로그 메시지를 생성합니다.*

유형

부울

Default

true

부록 B. CEPH 네트워크 구성 옵션

다음은 Ceph의 일반적인 네트워크 구성 옵션입니다.

public_network

설명

공용(전면) 네트워크(예: 192.168.0.0/24)의 IP 주소 및 넷마스크. [global] 에 설정합니다. 쉼표로 구분된 서브넷을 지정할 수 있습니다.

유형

<ip-address>/<netmask> [, <ip-address>/<netmask>]

필수 항목

없음

Default

해당 없음

public_addr

설명

공용(전면) 네트워크의 IP 주소입니다. 각 데몬에 대해 설정됩니다.

유형

IP 주소

필수 항목

없음

Default

해당 없음

cluster_network

설명

클러스터 네트워크의 IP 주소 및 넷마스크(예 : 10.0.0.0/24). [global] 에 설정합니다. 쉼표로 구분된 서브넷을 지정할 수 있습니다.

유형

<ip-address>/<netmask> [, <ip-address>/<netmask>]

필수 항목

없음

Default

해당 없음

cluster_addr

설명

*클러스터 네트워크의 **IP** 주소입니다. 각 데몬에 대해 설정됩니다.*

유형

주소

필수 항목

없음

Default

해당 없음

ms_type

설명

*네트워크 전송 계층의 **jboss** 유형입니다. **Red Hat**은 **posix** 의미 체계를 사용하여 단순하고 비동기식 유형을 지원합니다.*

유형

문자열.

필수 항목

아니요.

Default

async+posix

ms_public_type

설명

공용 네트워크의 네트워크 전송 계층에 대한 **jboss** 유형입니다. **ms_type** 과 동일하게 작동하지만 공용 또는 전면 네트워크에만 적용할 수 있습니다. 이 설정을 사용하면 **Ceph**에서 공용 또는 전면, 클러스터 또는 후면 네트워크에 다른 유형을 사용할 수 있습니다.

유형

문자열.

필수 항목

아니요.

Default

없음.

ms_cluster_type

설명

클러스터 네트워크의 네트워크 전송 계층에 대한 **jboss** 유형입니다. **ms_type** 과 동일하게 작동하지만 클러스터 또는 후면 네트워크에만 적용할 수 있습니다. 이 설정을 사용하면 **Ceph**에서 공용 또는 전면, 클러스터 또는 후면 네트워크에 다른 유형을 사용할 수 있습니다.

유형

문자열.

필수 항목

아니요.

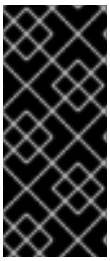
Default

없음.

호스트 옵션

선언된 각 모니터 아래에 **mon addr** 설정을 사용하여 **Ceph** 구성 파일에서 하나 이상의 **Ceph Monitor** 를 선언해야 합니다. **Ceph**는 **Ceph** 구성 파일에서 선언된 각 모니터, 메타데이터 서버, **OSD** 아래에 호스

트 설정을 예상합니다.



중요

localhost 를 사용하지 마십시오. **FQDN**(정규화된 도메인 이름)이 아니라 노드의 짧은 이름을 사용합니다. 노드 이름을 검색하는 타사 배포 시스템을 사용할 때 호스트의 값을 지정하지 마십시오.

mon_addr

설명

클라이언트가 **Ceph** 모니터에 연결하는 데 사용할 수 있는 **<hostname>:<port>** 항목 목록입니다. 설정되지 않은 경우 **Ceph**는 **[mon.*]** 섹션을 검색합니다.

유형

문자열

필수 항목

없음

Default

해당 없음

호스트

설명

호스트 이름입니다. 특정 데몬 인스턴스에 이 설정을 사용합니다(예: **[osd.0]**).

유형

문자열

필수 항목

예, 데몬 인스턴스의 경우.

Default

localhost

TCP 옵션

Ceph는 기본적으로 TCP 버퍼링을 비활성화합니다.

ms_tcp_nodelay

설명

Ceph는 ms_tcp_nodelay 를 활성화하여 각 요청이 즉시 전송되도록 합니다(버거 없음). Nagle의 알고리즘을 비활성화하면 네트워크 트래픽이 증가하여 정체가 발생할 수 있습니다. 많은 수의 작은 패킷이 있는 경우 ms_tcp_nodelay 를 비활성화할 수 있지만 비활성화하면 일반적으로 대기 시간이 증가합니다.

유형

부울

필수 항목

없음

Default

true

ms_tcp_rcvbuf

설명

네트워크 연결의 수신 끝에 있는 소켓 버퍼의 크기입니다. 기본적으로 비활성되어 있습니다.

유형

32비트 정수

필수 항목

없음

Default

0

ms_tcp_read_timeout

설명

클라이언트 또는 데몬이 다른 **Ceph** 데몬에 요청하고 사용되지 않는 연결을 삭제하지 않으면

tcp 읽기 제한 시간이 지정된 초 후 연결을 유휴 상태로 정의합니다.

유형

서명되지 않은 64 비트 정수

필수 항목

없음

Default

900 15분.

바인딩 옵션

bind 옵션은 Ceph OSD 데몬의 기본 포트 범위를 구성합니다. 기본 범위는 6800:7100 입니다. Ceph 데몬이 IPv6 주소에 바인딩되도록 활성화할 수도 있습니다.

중요

방화벽 구성을 통해 구성된 포트 범위를 사용할 수 있는지 확인합니다.

ms_bind_port_min

설명

OSD 데몬이 바인딩할 최소 포트 번호입니다.

유형

32비트 정수

Default

6800

필수 항목

없음

ms_bind_port_max

설명

OSD 데몬이 바인딩할 최대 포트 번호입니다.

유형

32비트 정수

Default

7300

필수 항목

아니요.

ms_bind_ipv6

설명

Ceph 데몬이 IPv6 주소에 바인딩할 수 있도록 합니다.

유형

부울

Default

false

필수 항목

없음

비동기식 옵션

*이러한 **Ceph** 옵션은 **AsyncMessenger** 의 동작을 구성합니다.*

ms_async_transport_type

설명

AsyncMessenger 에서 사용하는 전송 유형입니다. **Red Hat**은 **posix** 설정을 지원하지만 현재는 **the dpdk** 또는 **rdma** 설정을 지원하지 않습니다. **POSIX**는 표준 **TCP/IP** 네트워킹을 사용하며 기본값입니다. 다른 전송 유형은 실험적이며 지원되지 않습니다.

유형

문자열

필수 항목

없음

Default

posix

ms_async_op_threads

설명

*각 **AsyncMessenger** 인스턴스에서 사용하는 초기 작업자 스레드 수입니다. 이 구성 설정은 **SHOULD** 가 복제본 수 또는 삭제 코드 청크 수와 동일하지만 **CPU** 코어 수가 낮거나 단일 서버의 **OSD** 수가 높은 경우 더 낮을 수 있습니다.*

유형

***64**비트 서명되지 않은 정수*

필수 항목

없음

Default

3

ms_async_max_op_threads

설명

*각 **AsyncMessenger** 인스턴스에서 사용하는 최대 작업자 스레드 수입니다. **OSD** 호스트에 **CPU** 수가 제한된 경우 더 낮은 값으로 설정하고, **CPU** 활용도가 낮은 경우 증가합니다.*

유형

***64**비트 서명되지 않은 정수*

필수 항목

없음

Default

5

ms_async_set_affinity

설명

***AsyncMessenger** 작업자를 특정 **CPU** 코어에 바인딩하려면 **true** 로 설정합니다.*

유형

부울

필수 항목

없음

Default

true

ms_async_affinity_cores

설명

***ms_async_set_affinity** 가 **true**인 경우 이 문자열은 **AsyncMessenger** 작업자가 **CPU** 코어에 바인딩되는 방법을 지정합니다. 예를 들어 **0,2** 는 작업자 **#1** 및 **#2**를 각각 **CPU** 코어 **#0** 및 **#2**에 바인딩합니다. 알림: 선호도를 수동으로 설정하는 경우 하이퍼 스레딩 또는 유사한 기술의 영향으로 생성된 가상 **CPU**에 작업자를 할당하지 않도록 합니다. 이는 실제 **CPU** 코어보다 느려지기 때문입니다.*

유형

문자열

필수 항목

없음

Default

(empty)

ms_async_send_inline

설명

***AsyncMessenger** 스레드에서 보내는 대신 메시지를 생성한 스레드에서 직접 보냅니다. 이 옵션은 **CPU** 코어가 많은 시스템에서 성능을 낮추는 것으로 알려져 있으므로 기본적으로 비활성화*

됩니다.

유형

부울

필수 항목

없음

Default

false

부록 C. CEPH 모니터 구성 옵션

다음은 배포 중에 설정할 수 있는 Ceph 모니터 구성 옵션입니다.

mon_initial_members

설명

시작 중에 클러스터에 있는 초기 모니터의 ID입니다. Ceph를 지정하면 초기 쿼럼(예: 3)을 형성하기 위해 홀수의 모니터가 필요합니다.

유형

문자열

Default

없음

mon_force_quorum_join

설명

맵에서 이전에 제거된 경우에도 모니터를 쿼럼에 참여하도록 강제 적용

유형

부울

Default

False

mon_dns_srv_name

설명

모니터 호스트/주소의 DNS를 쿼리하는 데 사용되는 서비스 이름입니다.

유형

문자열

Default

ceph-mon

fsid

설명

클러스터 ID입니다. 클러스터당 하나씩.

유형

UUID

필수 항목

네, 필요합니다.

Default

해당 없음. 지정하지 않은 경우 배포 도구에서 생성할 수 있습니다.

mon_data

설명

모니터의 데이터 위치.

유형

문자열

Default

/var/lib/ceph/mon/$cluster-$id

mon_data_size_warn

설명

모니터의 데이터 저장소가 이 임계값에 도달하면 Ceph는 클러스터 로그에 HEALTH_WARN 상태를 발행합니다. 기본값은 15GB입니다.

유형

정수

Default

*15*1024*1024*1024**

mon_data_avail_warn

설명

*모니터 데이터 저장소의 사용 가능한 디스크 공간이 이 백분율 이상이면 **Ceph**가 클러스터 로그에서 **HEALTH_WARN** 상태를 발행합니다.*

유형

정수

Default

30

mon_data_avail_crit

설명

*모니터 데이터 저장소의 사용 가능한 디스크 공간이 이 백분율 이하이면 **Ceph**가 클러스터 로그에서 **HEALTH_ERR** 상태를 발행합니다.*

유형

정수

Default

5

mon_warn_on_cache_pools_without_hit_sets

설명

*캐시 풀에 **hit_set_type** 매개 변수가 설정되지 않은 경우 클러스터 로그에서 **HEALTH_WARN** 상태를 **Ceph**에서 발행합니다.*

유형

부울

Default

True

mon_warn_on_crush_straw_calc_version_zero

설명

***CRUSH**의 ¢_ calc_version이 **0**이면 **Ceph**에서 클러스터 로그에 **HEALTH_ WARN** 상태를*

발행합니다. 자세한 내용은 CRUSH 튜닝 가능 항목을 참조하십시오.

유형

부울

Default

True

mon_warn_on_legacy_crush_tunables

설명

CRUSH 튜닝 가능 항목이 너무 오래된 경우 Ceph는 클러스터 로그에 HEALTH_WARN 상태를 발행합니다(mon _min_c_required_version이전).

유형

부울

Default

True

mon_crush_min_required_version

설명

이 설정은 클러스터에 필요한 최소 튜닝 가능 프로필 버전을 정의합니다.

유형

문자열

Default

Hammer

mon_warn_on_osd_down_out_interval_zero

설명

Leader가 noout 플래그 가 설정된 경우 유사한 방식으로 동작하기 때문에 mon_osd_down_out_interval 설정이 0이면 Ceph는 클러스터 로그에서 HEALTH_WARN 상태를 발행합니다. 관리자는 noout 플래그를 설정하여 클러스터 문제를 더 쉽게 해결할 수 있습니다. Ceph는 관리자가 설정이 0임을 알리기 위해 경고를 발행합니다.

유형

부울

Default

True

mon_cache_target_full_warn_ratio

설명

***cache_target_full* 및 *target_max_ object* 의 비율 간에 *Ceph*에서 경고를 발행합니다.**

유형

교체

Default

0.66

mon_health_data_update_interval

설명

쿼럼의 모니터가 해당 상태를 피어와 공유하는 빈도(초)입니다. 음수는 상태 업데이트를 비활성화합니다.

유형

교체

Default

60

mon_health_to_clog

설명

*이 설정을 사용하면 *Ceph*가 클러스터 로그에 상태 요약을 주기적으로 보낼 수 있습니다.*

유형

부울

Default

True

mon_health_detail_to_clog

설명

이 설정을 사용하면 **Ceph**에서 상태 세부 정보를 클러스터 로그에 주기적으로 보낼 수 있습니다.

유형

부울

Default

True

mon_op_complaint_time

설명

업데이트 없이 **Ceph Monitor** 작업이 차단되는 시간(초)입니다.

유형

정수

Default

30

mon_health_to_clog_tick_interval

설명

모니터가 클러스터 로그에 상태 요약을 보내는 빈도(초)입니다. 양수가 아닌 숫자를 사용하면 비활성화됩니다. 현재 상태 요약이 비어 있거나 마지막과 동일한 경우 모니터는 클러스터 로그에 상태를 보내지 않습니다.

유형

정수

Default

60.000000

mon_health_to_clog_interval

설명

모니터가 클러스터 로그에 상태 요약을 보내는 빈도(초)입니다. 양수가 아닌 숫자를 사용하면 비활성화됩니다. 모니터는 항상 클러스터 로그에 요약을 보냅니다.

유형

정수

Default

600

mon_osd_full_ratio

설명

***OSD**가 전체로 간주되기 전에 사용된 디스크 공간의 백분율입니다.*

유형

플로트:

Default

.95

mon_osd_nearfull_ratio

설명

***OSD** 전에 사용된 디스크 공간의 백분율은 거의 전체로 간주됩니다.*

유형

교체

Default

.85

mon_sync_trim_timeout

설명, 유형

두 배

Default

30.0

mon_sync_heartbeat_timeout

설명, 유형

두 배

Default

30.0

mon_sync_heartbeat_interval

설명, 유형

두 배

Default

5.0

mon_sync_backoff_timeout

설명, 유형

두 배

Default

30.0

mon_sync_timeout

설명

모니터가 동기화 공급자에서 다음 업데이트 메시지를 기다린 후 다시 공급 및 부트 스트랩을 대기하는 시간(초)입니다.

유형

두 배

Default

60.000000

mon_sync_max_retries

설명, 유형

정수

Default

5

mon_sync_max_payload_size

설명

동기화 페이로드의 최대 크기(바이트)입니다.

유형

***32**비트 정수*

Default

1045676

paxos_max_join_drift

설명

*모니터 데이터 저장소를 먼저 동기화해야 하기 전에 최대 **Paxos** 반복됩니다. 모니터에서 피어가 너무 앞서 있다는 것을 알게 되면 먼저 데이터 저장소와 동기화됩니다.*

유형

정수

Default

10

paxos_stash_full_interval

설명

***PaxosService** 상태의 전체 복사본을 저장하는 빈도(커밋)입니다. 현재 이 설정은 **mds,mon,auth** 및 **mgr PaxosServices**에만 영향을 미칩니다.*

유형

정수

Default

25

paxos_propose_interval

설명

맵 업데이트를 제안하기 전에 이 시간 간격의 업데이트를 수집합니다.

유형

두 배

Default

1.0

paxos_min

설명

*유지해야 할 최소 **paxos** 상태 수*

유형

정수

Default

500

paxos_min_wait

설명

일정 기간 동안 비활성화된 후 업데이트를 수집하는 최소 시간입니다.

유형

두 배

Default

0.05

paxos_trim_min

설명

트리밍하기 전에 허용되는 추가 제안 수

유형

정수

Default

250

paxos_trim_max

설명

한 번에 트리밍할 최대 추가 제안 수

유형

정수

Default

500

paxos_service_trim_min

설명

트리밍을 트리거 할 최소 버전 (0 비활성화)

유형

정수

Default

250

paxos_service_trim_max

설명

단일 제안 중 트리밍할 최대 버전 수 (0 비활성화)

유형

정수

Default

500

mon_max_log_epochs

설명

단일 제안 중에 트리밍할 최대 로그 양

유형

정수

Default

500

mon_max_pgmap_epochs

설명

*단일 제안 중에 트리밍할 최대 **pgmap epoch** 수*

유형

정수

Default

500

mon_mds_force_trim_to

설명

*모니터가 이 지점으로 **mdsmaps**를 트리밍하도록 강제합니다(**0** 비활성화). 위험, 주의해서 사용)*

유형

정수

Default

0

mon_osd_force_trim_to

설명

*지정된 **epoch**에 **PG**가 정리되지 않더라도이 시점에 **osdmaps**를 강제 실행 (비활성화합니다. 위험, 주의해서 사용)*

유형

정수

Default

0

mon_osd_cache_size

설명

***osdmaps** 캐시의 크기, 기본 저장소의 캐시에 의존하지 않음*

유형

정수

Default

500

mon_election_timeout

설명

*선택 제안자의 경우 모든 **ACK**의 최대 대기 시간(초)입니다.*

유형

교체

Default

5

mon_lease

설명

모니터 버전의 리스를 길이(초)합니다.

유형

교체

Default

5

mon_lease_renew_interval_factor

설명

***Mon lease * mon lease**의 갱신 간격 요소는 리더가 다른 모니터의 임대를 갱신하는 간격이 됩니다. 인수는 **1.0** 보다 작아야 합니다.*

유형

교체

Default

0.6

mon_lease_ack_timeout_factor

설명

*리더는 **mon lease * mon lease * mon lease** 시간 초과 요소를 통해 공급자가 리스 연장을 승인할 것입니다.*

유형

교체

Default

2.0

mon_accept_timeout_factor

설명

*리더는 **mon lease * mon**이 요청자가 **Paxos** 업데이트를 수락할 때까지 시간 초과 인수를 허용합니다. 유사한 용도로 **Paxos** 복구 단계 중에도 사용됩니다.*

유형

교체

Default

2.0

mon_min_osdmap_epochs

설명

*항상 유지할 최소 **OSD** 맵 **epoch** 수입니다.*

유형

***32**비트 정수*

Default

500

mon_max_pgmap_epochs

설명

*모니터가 유지해야 하는 최대 **PG** 맵 수입니다.*

유형

***32**비트 정수*

Default

500

mon_max_log_epochs

설명

모니터가 유지해야 하는 최대 로그 수입니다.

유형

32비트 정수

Default

500

clock_offset

설명

*시스템 클록을 오프셋하는 양. 자세한 내용은 **Clock.cc** 를 참조하십시오.*

유형

두 배

Default

0

mon_tick_interval

설명

모니터의 틱 간격(초)입니다.

유형

32비트 정수

Default

5

mon_clock_drift_allowed

설명

모니터 간에 허용되는 시간(초)입니다.

유형

교체

Default

.050

mon_clock_drift_warn_backoff

설명

시계 드리프트 경고에 대한 기하급수적인 백오프.

유형

교체

Default

5

mon_timecheck_interval

설명

리더에 대한 시간 점검 간격(시계 드리프트 확인)(초)입니다.

유형

교체

Default

300.0

mon_timecheck_skew_interval

설명

리더에 대한 스큐(초)가 있는 시간 점검 간격(시계 드리프트 확인).

유형

교체

Default

30.0

mon_max_osd

설명

*클러스터에 허용되는 최대 **OSD** 수입니다.*

유형

***32**비트 정수*

Default

10000

mon_globalid_prealloc

설명

*클러스터의 클라이언트 및 데몬에 대해 사전 할당할 글로벌 **ID** 수입니다.*

유형

***32**비트 정수*

Default

10000

mon_sync_fs_threshold

설명

*지정된 개수의 오브젝트를 작성할 때 파일 시스템과 동기화합니다. 비활성화하려면 **0** 으로 설정합니다.*

유형

***32**비트 정수*

Default

5

mon_subscribe_interval

설명

서브스크립션의 새로 고침 간격(초)입니다. 서브스크립션 메커니즘을 사용하면 클러스터 맵과 로그 정보를 얻을 수 있습니다.

유형

두 배

Default

86400.000000

mon_stat_smooth_intervals

설명

Ceph는 최근 NPG 맵에 대한 원활한 통계가 될 것입니다.

유형

정수

Default

6

mon_probe_timeout

설명

모니터가 부트스트랩하기 전에 피어를 찾기 위해 대기하는 시간 (초)입니다.

유형

두 배

Default

2.0

mon_daemon_bytes

설명

*메타데이터 서버 및 **OSD** 메시지(바이트)에 대한 메시지 메모리 제한입니다.*

유형

***64**비트 정수가 서명되지 않았습니다*

Default

400UL 20

mon_max_log_entries_per_event

설명

이벤트당 최대 로그 항목 수입니다.

유형

정수

Default

4096

mon_osd_prime_pg_temp

설명

*외부 **OSD**가 다시 시작될 때 이전 **OSD**를 사용하여 **PGMap** 기본 설정을 활성화하거나 비활성화합니다. **true** 설정을 사용하면 클라이언트는 **OSD**에서 **PG** 피어**ed**로 새로 마운트될 때까지 계속 이전 **OSD**를 사용합니다.*

유형

부울

Default

true

mon_osd_prime_pg_temp_max_time

설명

*외부 **OSD**가 다시 시작될 때 모니터가 **PGMap**을 기본으로 만드는 데 소비해야 하는 시간(초)입니다.*

유형

교체

Default

0.5

mon_osd_prime_pg_temp_max_time_estimate

설명

*모든 **PG**를 병렬로 사용하기 전에 각 **PG**에서 사용된 최대 시간 추정.*

유형

교체

Default

0.25

mon_osd_allow_primary_affinity

설명

***osdmap**에 **primary_affinity** 를 설정할 수 있습니다.*

유형

부울

Default

False

mon_osd_pool_ec_fast_read

설명

*풀에서 빠른 읽기를 켭니다. **fast_read** 가 생성 시 지정되지 않은 경우 새로 생성된 삭제 풀의 기본 설정으로 사용됩니다.*

유형

부울

Default

False

mon_mds_skip_sanity

설명

*계속 진행하려는 버그의 경우 **FSMap**에 대한 보안 어설션을 건너뜁니다. 모니터는 **FSMap** 온전성 검사가 실패할 경우 종료되지만 이 옵션을 활성화하여 비활성화할 수 있습니다.*

유형

부울

Default

False

mon_max_mdsmap_epochs

설명

*단일 제안 중에 최대 **mdsmap epochs** 양.*

유형

정수

Default

500

mon_config_key_max_entry_size

설명

***config-key** 항목의 최대 크기(바이트)입니다.*

유형

정수

Default

65536

mon_scrub_interval

설명

저장된 체크섬과 모든 저장된 키 중 계산된 체크섬을 비교하여 모니터가 저장소를 스크럽하는 빈도(초)입니다.

유형

정수

Default

3600*24

mon_scrub_max_keys

설명

매번 스크럽할 최대 키 수입니다.

유형

정수

Default

100

mon_compact_on_start

설명

ceph-mon start에서 **Ceph Monitor** 저장소로 사용한 데이터베이스를 작게 합니다. 수동 압축을 사용하면 모니터 데이터베이스를 축소하고 정기적인 압축이 작동하지 않는 경우 성능이 향상됩니다.

유형

부울

Default

False

mon_compact_on_bootstrap

설명

부트스트랩에서 **Ceph Monitor** 저장소로 사용되는 데이터베이스를 압축합니다. 모니터는 부트스트랩 후 쿼럼을 생성하기 위해 서로 검색하기 시작합니다. 쿼럼에 참여하기 전에 시간 초과하면 다시 시작하여 부트스트랩합니다.

유형

부울

Default

False

mon_compact_on_trim

설명

*이전 상태를 정리할 때 특정 접두사(**paxos** 포함)를 압축합니다.*

유형

부울

Default

True

mon_cpu_threads

설명

*모니터에서 **CPU** 집약적인 작업을 수행하는 스레드 수입니다.*

유형

부울

Default

True

mon_osd_mapping_pgs_per_chunk

설명

*배치 그룹에서 청크의 **OSD**로의 매핑을 계산합니다. 이 옵션은 청크당 배치 그룹의 수를 지정합니다.*

유형

정수

Default

4096

mon_osd_max_split_count

설명

" involved" OSD당 최대 개수로 분할할 수 있습니다. 풀의 pg_num 을 늘리면 배치 그룹이 해당 풀을 제공하는 모든 OSD에서 분할됩니다. PG 분할 시 극심한 승수를 피하고자 합니다.

유형

정수

Default

300

rados_mon_op_timeout

설명

***rados** 작업에서 오류를 반환하기 전에 모니터에서 응답을 기다리는 시간(초)입니다. 0은 제한 시 또는 대기 시간을 의미합니다.*

유형

두 배

Default

0

추가 리소스

- *풀 값*
- ***CRUSH** 튜닝 가능 항목*

부록 D. CEPHX 구성 옵션

다음은 배포 중에 설정할 수 있는 Cephx 구성 옵션입니다.

auth_cluster_required

설명

활성화된 경우 Red Hat Ceph Storage 클러스터 데몬인 ceph-mon 및 ceph-osd 가 서로 인증되어야 합니다. 유효한 설정은 cephx 또는 none 입니다.

유형

문자열

필수 항목

없음

Default

cephx.

auth_service_required

설명

활성화된 경우 Red Hat Ceph Storage 클러스터 데몬은 Ceph 서비스에 액세스하려면 Ceph 클라이언트가 Red Hat Ceph Storage 클러스터를 인증해야 합니다. 유효한 설정은 cephx 또는 none 입니다.

유형

문자열

필수 항목

없음

Default

cephx.

auth_client_required

설명

Ceph 클라이언트가 활성화된 경우 Ceph 클라이언트에서 Red Hat Ceph Storage 클러스터

*를 인증해야 합니다. 유효한 설정은 **cephx** 또는 **none** 입니다.*

유형

문자열

필수 항목

없음

Default

cephx.

인증 키

설명

인증 키 파일의 경로입니다.

유형

문자열

필수 항목

없음

Default

/etc/ceph/$cluster.$name.keyring,/etc/ceph/$cluster.keyring,/etc/ceph/keyring,/etc/ceph/keyring.bin

키 파일

설명

키 파일의 경로(즉, 키만 포함하는 파일).

유형

문자열

필수 항목

없음

Default

없음

key

설명

키(즉, 키 자체의 텍스트 문자열). 권장되지 않음.

유형

문자열

필수 항목

없음

Default

없음

ceph-mon

위치

$mon_data/keyring

capabilities

Mon '허용 *'

ceph-osd

위치

$osd_data/keyring

capabilities

Mon 'allow profile osd' osd 'allow *'

radosgw

위치

$rgw_data/keyring

capabilities

mon 'allow rwx' osd 'allow rwx'

cephx_require_signatures

설명

true 로 설정하면 *Ceph* 클라이언트와 *Red Hat Ceph Storage* 클러스터 사이의 모든 메시지 트래픽과 *Red Hat Ceph Storage* 클러스터가 구성된 데몬 간 서명이 필요합니다.

유형

부울

필수 항목

없음

Default

false

cephx_cluster_require_signatures

설명

true 로 설정하면 *Ceph*는 *Red Hat Ceph Storage* 클러스터를 구성하는 *Ceph* 데몬 간의 모든 메시지 트래픽에 서명이 필요합니다.

유형

부울

필수 항목

없음

Default

false

cephx_service_require_signatures

설명

true 로 설정하면 *Ceph* 클라이언트와 *Red Hat Ceph Storage* 클러스터 간의 모든 메시지 트래픽에 서명이 필요합니다.

유형

부울

필수 항목

없음

Default

false

cephx_sign_messages

설명

***Ceph** 버전이 메시지 서명을 지원하는 경우 **Ceph**는 모든 메시지에 서명하여 스푸핑할 수 없도록 합니다.*

유형

부울

Default

true

auth_service_ticket_ttl

설명

***Red Hat Ceph Storage** 클러스터가 **Ceph** 클라이언트에 인증 티켓을 보내면 클러스터에서 실시간으로 티켓을 할당합니다.*

유형

두 배

Default

60*60

부록 E. 풀, 배치 그룹, CRUSH 구성 옵션

풀, 배치 그룹 및 CRUSH 알고리즘을 제어하는 Ceph 옵션.

mon_allow_pool_delete

설명

모니터가 풀을 삭제할 수 있도록 합니다. RHCS 3 이상 릴리스에서는 모니터가 데이터를 보호하기 위해 기본적으로 풀을 삭제할 수 없습니다.

유형

부울

Default

false

mon_max_pool_pg_num

설명

풀당 최대 배치 그룹 수입니다.

유형

정수

Default

65536

mon_pg_create_interval

설명

동일한 Ceph OSD 데몬에서 PG 생성 간격(초)입니다.

유형

교체

Default

30.0

mon_pg_stuck_threshold

설명

PG를 정지된 것으로 간주할 수 있는 시간(초)입니다.

유형

32비트 정수

Default

300

mon_pg_min_inactive

설명

mon_pg _stuck_threshold보다 비활성 상태인 PG 수가 이 설정을 초과하면 Ceph는 클러스터 로그에서 HEALTH_ ERR 상태를 발행합니다. 기본 설정은 하나의 PG입니다. 양수가 아닌 경우 이 설정을 비활성화합니다.

유형

정수

Default

1

mon_pg_warn_min_per_osd

설명

클러스터의 OSD당 평균 PG 수가 이 설정보다 작으면 Ceph는 클러스터 로그에서 HEALTH_WARN 상태를 발행합니다. 양수가 아닌 경우 이 설정을 비활성화합니다.

유형

정수

Default

30

mon_pg_warn_max_per_osd

설명

클러스터의 OSD당 평균 PG 수가 이 설정보다 큰 경우 Ceph는 클러스터 로그에 HEALTH_WARN 상태를 발행합니다. 양수가 아닌 경우 이 설정을 비활성화합니다.

유형

정수

Default

300

mon_pg_warn_min_objects

설명

클러스터의 총 오브젝트 수가 이 수보다 낮은 경우 경고하지 마십시오.

유형

정수

Default

1000

mon_pg_warn_min_pool_objects

설명

개체 번호가 이 번호 아래에 있는 풀에는 경고하지 마십시오.

유형

정수

Default

1000

mon_pg_check_down_all_threshold

설명

Ceph가 모든 PG를 확인하여 중단 되거나 오래되지 않았는지 확인하는 백분율별 OSD 임계값.

유형

교체

Default

0.5

mon_pg_warn_max_object_skew

설명

*풀의 평균 오브젝트 수가 **mon pg warn max** 오브젝트 스큐보다 큰 경우 모든 풀의 평균 오브젝트 수를 초과하는 경우 **Ceph**는 클러스터 로그에서 **HEALTH_WARN** 상태를 발행합니다. 양수가 아닌 경우 이 설정을 비활성화합니다.*

유형

교체

Default

10

mon_delta_reset_interval

설명

***Ceph**가 **PG delta**를 **0**으로 재설정하기 전의 비활성 시간(초)입니다. **Ceph**는 각 풀에 사용된 공간을 추적하여 관리자가 복구 및 성능 진행 상황을 평가하는 데 도움을 줍니다.*

유형

정수

Default

10

mon_osd_max_op_age

설명

***HEALTH_WARN** 상태를 발행하기 전에 작업을 완료하는 데 사용할 최대 기간(초)입니다.*

유형

교체

Default

32.0

osd_pg_bits

설명

Ceph OSD 데몬당 배치 그룹 비트.

유형

32비트 정수

Default

6

osd_pgp_bits

설명

배치용 PGP(배치 그룹의 Ceph OSD 데몬)당 비트 수.

유형

32비트 정수

Default

6

osd_crush_chooseleaf_type

설명

CRUSH 규칙에서 select leaf 에 사용할 버킷 유형입니다. 이름 대신 서수 등급을 사용합니다.

유형

32비트 정수

Default

1. 일반적으로 하나 이상의 Ceph OSD 데몬을 포함하는 호스트입니다.

osd_pool_default_crush_replicated_ruleset

설명

복제된 풀을 생성할 때 사용할 기본 CRUSH 규칙 세트입니다.

유형

8비트 정수

Default

0

osd_pool_erasure_code_stripe_unit

설명

*코딩된 풀에 대한 오브젝트 스트라이프 청크의 기본 크기(바이트)를 설정합니다. 크기 S의 모든 개체는 N 스트라이프로 저장되며 각 데이터 청크는 스트라이프 단위 바이트를 수신합니다. N * 스트라이프 단위 바이트의 각 스트라이프 는 개별적으로 인코딩/디코딩됩니다. 이 옵션은 삭제 코드 프로필의 스트라이프_unit 설정으로 재정의할 수 있습니다.*

유형

서명되지 않은 32 비트 정수

Default

4096

osd_pool_default_size

설명

풀에서 오브젝트의 복제본 수를 설정합니다. 기본값은 ceph osd 풀 세트 {pool-name} 크기 {size} 와 동일합니다.

유형

32비트 정수

Default

3

osd_pool_default_min_size

설명

클라이언트에 대한 쓰기 작업을 승인하기 위해 풀에서 오브젝트의 최소 쓰기 복제본 수를 설정합니다. 최솟값이 충족되지 않으면 Ceph에서 클라이언트에 대한 쓰기를 인식하지 못합니다. 이

설정은 성능이 저하된 모드에서 작동하는 경우 최소 복제본 수를 보장합니다.

유형

32비트 정수

Default

0, 이는 특정 최소값이 없음을 의미합니다. 0 인 경우 최소 크기는 size - (size / 2) 입니다.

osd_pool_default_pg_num

설명

풀의 기본 배치 그룹 수입니다. 기본값은 mkpool 을 사용하여 pg_num 과 동일합니다.

유형

32비트 정수

Default

32

osd_pool_default_pgp_num

설명

풀에 대한 배치를 위한 기본 배치 그룹 수입니다. 기본값은 mkpool 을 사용하여 pgp_num 과 동일합니다. PG와 PGP는 동일해야 합니다.

유형

32비트 정수

Default

0

osd_pool_default_flags

설명

새 풀의 기본 플래그입니다.

유형

32비트 정수

Default

0

osd_max_pgls

설명

나열할 최대 배치 그룹 수입니다. 많은 수의 클라이언트가 Ceph OSD 데몬을 연결할 수 있습니다.

유형

서명되지 않은 64 비트 정수

Default

1024

참고

기본값은 정상이어야 합니다.

osd_min_pg_log_entries

설명

로그 파일을 정리할 때 유지 관리할 최소 배치 그룹 로그 수입니다.

유형

32 비트 Int 서명되지 않음

Default

250

osd_default_data_pool_replay_window

설명

OSD가 클라이언트가 요청을 재생할 때까지 기다리는 시간(초)입니다.

유형

32비트 정수

Default

45

부록 F. OSD(오브젝트 스토리지 데몬) 구성 옵션

다음은 배포 중에 설정할 수 있는 Ceph OSD(오브젝트 스토리지 데몬) 구성 옵션입니다.

osd_uuid

설명

Ceph OSD의 UUID(Universally Unique Identifier).

유형

UUID

Default

UUID입니다.

참고

osd uuid 는 단일 Ceph OSD에 적용됩니다. fsid 는 전체 클러스터에 적용됩니다.

osd_data

설명

OSD 데이터 경로입니다. Ceph를 배포할 때 디렉터리를 만들어야 합니다. 이 마운트 지점에 OSD 데이터의 드라이브를 마운트합니다.

> **IMPORTANT:** Red Hat does not recommend changing the default.

유형

문자열

Default

/var/lib/ceph/osd/$cluster-$id

osd_max_write_size

설명

쓰기의 최대 크기(MB)입니다.

유형

32비트 정수

Default

90

osd_client_message_size_cap

설명

메모리에 허용된 가장 큰 클라이언트 데이터 메시지.

유형

64비트 정수가 서명되지 않았습니다

Default

*500MB 기본값. 500*1024L*1024L*

osd_class_dir

설명

RADOS 클래스 플러그인의 클래스 경로입니다.

유형

문자열

Default

$libdir/rados-classes

osd_max_scrubs

설명

Ceph OSD의 최대 동시 스크럽 작업 수입니다.

유형

32비트 Int

Default

1

osd_scrub_thread_timeout

설명

스크럽 스레드를 시간 초과하기 전 최대 시간(초)입니다.

유형

32비트 정수

Default

60

osd_scrub_finalize_thread_timeout

설명

스크럽이 스레드를 종료하기 전에 최대 시간(초)입니다.

유형

32비트 정수

Default

*60*10*

osd_scrub_begin_hour

설명

경량 또는 심층 스크럽을 시작할 수 있는 가장 빠른 시간입니다. osd 스크럽 종료 시간 매개 변수와 함께 사용하여 스크럽 타임 창을 정의하고 제한 사항을 오프 피크 시간에 스크럽할 수 있습니다. 설정은 24시간 주기의 시간을 지정하는 데 정수를 사용합니다. 여기서 0 은 오전 12:01부터 오전 1:00까지 시간을 나타내고, 13은 오후 1:01부터 오후 2:00까지 시간을 나타냅니다.

유형

32비트 정수

Default

0:12:01 - 1:00.

osd_scrub_end_hour

설명

경량 또는 심층 스크럽을 시작할 수 있는 가장 최근 시간입니다. 이 매개 변수는 **osd** *스크럽 시작 시간 매개 변수와 함께 사용하여 스크럽 시간 창을 정의하고 구성 요소가 오프 피크 시간에 스크럽되도록 허용합니다. 설정은* **24***시간 주기의 시간을 지정하는 데 정수를 사용합니다. 여기서* **0** *은 오전* **12:01***부터 오전* **1:00***까지 시간을 나타내고,* **13***은 오후* **1:01***부터 오후* **2:00***까지 시간을 나타냅니다. 종료 시간은 시작 시간보다 커야 합니다.*

유형

32*비트 정수*

Default

24:11:01 p.m. - a.m.

osd_scrub_load_threshold

설명

최대 부하. 시스템 부하(**getloadavg()** *함수에서 정의한 대로)가 이 수 보다 높으면* **Ceph***가 스크럽되지 않습니다. 기본값은* **0.5** *입니다.*

유형

교체

Default

0.5

osd_scrub_min_interval

설명

Red Hat Ceph Storage *클러스터 로드가 낮을 때* **Ceph OSD***를 스크럽하는 최소 간격(초)입니다.*

유형

교체

Default

하루에 한 번. **60*60*24**

osd_scrub_max_interval

설명

*클러스터 부하에 관계없이 **Ceph OSD**를 스크럽하는 최대 간격(초)입니다.*

유형

교체

Default

*주당 한 번. **7*60*60*24***

osd_scrub_interval_randomize_ratio

설명

*비율을 가져와서 **osd** 스크럽 최소 간격과 **osd** 스크럽 최대 간격 간에 예약된 스크럽을 임의화합니다.*

유형

교체

Default

0.5.

mon_warn_not_scrubbed

설명

***osd_scrub_interval** 후 스크럽되지 않은 모든 **PG**에 대해 경고하는 시간(초)입니다.*

유형

정수

Default

***0** (경고 없음).*

osd_scrub_chunk_min

설명

*오브젝트 저장소는 해시 경계에서 끝나는 청크로 분할됩니다. 청크 스크럽의 경우 **Ceph**는 해당 청크에 대해 쓰기가 차단된 상태에서 한 번에 하나의 청크를 제거합니다. **osd scrub** 청크 **min***

설정은 스크럽에 대한 최소 청크 수를 나타냅니다.

유형

32비트 정수

Default

5

osd_scrub_chunk_max

설명

스크럽할 최대 청크 수입니다.

유형

32비트 정수

Default

25

osd_scrub_sleep

설명

심층 스크럽(deep 스크럽) 사이에 대기하는 시간입니다.

유형

교체

Default

0 (또는 꺼짐).

osd_scrub_during_recovery

설명

복구 중에 스크럽 가능.

유형

부울

Default

false

osd_scrub_invalid_stats

설명

추가 스크럽을 통해 통계를 유효하지 않은 것으로 수정합니다.

유형

부울

Default

true

osd_scrub_priority

설명

스크럽 작업의 대기열 우선 순위와 클라이언트 I/O를 제어합니다.

유형

서명되지 않은 32 비트 정수

Default

5

osd_scrub_cost

설명

대기열 스케줄링을 위한 메가바이트 단위의 스크럽 작업 비용.

유형

서명되지 않은 32 비트 정수

Default

52428800

osd_deep_scrub_interval

설명

*모든 데이터를 완전히 읽는 심층 스크럽 간격입니다. **osd** 스크럽 로드 임계값 매개 변수는 이 설정에 영향을 미치지 않습니다.*

유형

교체

Default

*주당 한 번. **60*60*24*7***

osd_deep_scrub_stride

설명

심층 스크럽을 수행할 때 크기 읽기.

유형

***32**비트 정수*

Default

512KB. 524288

mon_warn_not_deep_scrubbed

설명

***osd_deep_scrub_interval** 후 스크럽되지 않은 모든 **PG**에 대해 경고하는 시간(초)입니다.*

유형

정수

Default

***0** (경고 없음).*

osd_deep_scrub_randomize_ratio

설명

***scrubs**가 무작위로 깊은 스크럽이 됩니다(**osd_deep_scrub_interval** 이 전달되기 전에도 해당).*

유형

교체

Default

0.15 또는 15%.

osd_deep_scrub_update_digest_min_age

설명

스크럽하기 전에 전체 오브젝트 다이제스트를 업데이트하기 전에 이전 오브젝트가 몇 초입니까.

유형

정수

Default

7200 (120시간)

osd_op_num_shards

설명

*클라이언트 작업의 **shard** 수입니다.*

유형

32비트 정수

Default

0

osd_op_num_threads_per_shard

설명

*클라이언트 작업에 대한 **shard**당 스레드 수입니다.*

유형

32비트 정수

Default

0

osd_op_num_shards_hdd

설명

HDD 작업의 shard 수입니다.

유형

32비트 정수

Default

5

osd_op_num_threads_per_shard_hdd

설명

HDD 작업을 위한 shard당 스레드 수입니다.

유형

32비트 정수

Default

1

osd_op_num_shards_ssd

설명

SSD 작업의 shard 수입니다.

유형

32비트 정수

Default

8

osd_op_num_threads_per_shard_ssd

설명

SSD 작업의 shard당 스레드 수입니다.

유형

32비트 정수

Default

2

osd_client_op_priority

설명

클라이언트 작업에 설정된 우선 순위입니다. osd 복구 op 우선 순위 와 상대적입니다.

유형

32비트 정수

Default

63

유효한 범위

1-63

osd_recovery_op_priority

설명

복구 작업에 설정된 우선 순위입니다. osd 클라이언트 op 우선 순위 와 상대적입니다.

유형

32비트 정수

Default

3

유효한 범위

1-63

osd_op_thread_timeout

설명

Ceph OSD 작업 스레드 시간 초과(초).

유형

32비트 정수

Default

15

osd_op_complaint_time

설명

지정된 시간(초)이 경과하면 작업에 대한 불만이 발생합니다.

유형

교체

Default

30

osd_disk_threads

설명

스크럽 및 스냅샷 트리밍과 같은 백그라운드 디스크 집약 OSD 작업을 수행하는 데 사용되는 디스크 스레드 수입니다.

유형

32비트 정수

Default

1

osd_op_history_size

설명

추적할 완료된 작업의 최대 수입니다.

유형

32 비트 서명되지 않은 정수

Default

20

osd_op_history_duration

설명

추적할 가장 오래된 작업.

유형

32 비트 서명되지 않은 정수

Default

600

osd_op_log_threshold

설명

한 번에 표시할 작업 수입니다.

유형

32비트 정수

Default

5

osd_op_timeout

설명

OSD 작업을 실행하는 시간(초)입니다.

유형

정수

Default

0

중요

*클라이언트가 결과를 처리할 수 없는 경우 **osd op timeout** 옵션을 설정하지 마십시오. 예를 들어 가상 시스템이 이 시간 초과를 하드웨어 오류로 해석하기 때문에 가상 시스템에서 실행 중인 클라이언트에 이 매개 변수를 설정하면 데이터가 손상될 수 있습니다.*

osd_max_backfills

설명

*단일 **OSD**에 또는 단일 **OSD**에 허용된 최대 백필 작업 수입니다.*

유형

***64**비트 서명되지 않은 정수*

Default

1

osd_backfill_scan_min

설명

백필 검사당 최소 오브젝트 수입니다.

유형

***32**비트 정수*

Default

64

osd_backfill_scan_max

설명

백필 검사당 최대 오브젝트 수입니다.

유형

32비트 정수

Default

512

osd_backfill_full_ratio

설명

Ceph OSD의 전체 비율이 이 값보다 높은 경우 백필 요청을 수락하지 않습니다.

유형

교체

Default

0.85

osd_backfill_retry_interval

설명

백필 요청을 다시 시도하기 전에 대기하는 시간(초)입니다.

유형

두 배

Default

30.000000

osd_map_dedup

설명

OSD 맵에서 중복 제거를 활성화합니다.

유형

부울

Default

true

osd_map_cache_size

설명

OSD 맵 캐시의 크기(MB)입니다.

유형

32비트 정수

Default

50

osd_map_cache_bl_size

설명

OSD 데몬의 인메모리 OSD 맵 캐시 크기입니다.

유형

32비트 정수

Default

50

osd_map_cache_bl_inc_size

설명

인메모리 OSD의 크기는 OSD 데몬의 캐시 증분입니다.

유형

32비트 정수

Default

100

osd_map_message_max

설명

MOSDMap 메시지당 허용된 최대 맵 항목입니다.

유형

32비트 정수

Default

40

osd_snap_trim_thread_timeout

설명

스냅 트리 스레드를 시간 초과하기 전 최대 시간(초)입니다.

유형

32비트 정수

Default

*60*60*1*

osd_pg_max_concurrent_snap_trims

설명

병렬 스냅샷 트리밍/PG의 최대 수입니다. 이는 한 번에 정리할 PG당 오브젝트 수를 제어합니다.

유형

32비트 정수

Default

2

osd_snap_trim_sleep

설명

PG 문제마다 트리밍 작업을 삽입합니다.

유형

32비트 정수

Default

0

osd_max_trimming_pgs

설명

최대 PG 트리밍 수

유형

32비트 정수

Default

2

osd_backlog_thread_timeout

설명

백로그 스레드를 시간 초과하기 전 최대 시간(초)입니다.

유형

32비트 정수

Default

*60*60*1*

osd_default_notify_timeout

설명

OSD 기본 알림 시간 제한(초).

유형

32 비트 정수가 서명되지 않았습니다.

Default

30

osd_check_for_log_corruption

설명

로그 파일에서 손상이 있는지 확인합니다. 비용이 많이 들 수 있습니다.

유형

부울

Default

false

osd_remove_thread_timeout

설명

***OSD** 스레드 제거를 시간 초과하기 전 최대 시간(초)입니다.*

유형

***32**비트 정수*

Default

60*60

osd_command_thread_timeout

설명

명령 스레드의 시간 초과 전 최대 시간(초)입니다.

유형

***32**비트 정수*

Default

10*60

osd_command_max_records

설명

반환할 손실 오브젝트 수를 제한합니다.

유형

32비트 정수

Default

256

osd_auto_upgrade_tmap

설명

*이전 객체 의 **omap**에 **t map** 을 사용합니다.*

유형

부울

Default

true

osd_tmapput_sets_users_tmap

설명

*디버깅에만 **tmap** 을 사용합니다.*

유형

부울

Default

false

osd_preserve_trimmed_log

설명

정리된 로그 파일을 보존하지만 디스크 공간을 더 많이 사용합니다.

유형

부울

Default

false

osd_recovery_delay_start

설명

피어링이 완료되면 오브젝트 복구를 시작하기 전에 지정된 시간(초) 동안 **Ceph**가 지연됩니다.

유형

교체

Default

0

osd_recovery_max_active

설명

한 번에 **OSD**당 활성 복구 요청 수입니다. 요청이 많으면 복구 속도가 빨라지지만 요청에 따라 클러스터에서 부하가 증가합니다.

유형

32비트 정수

Default

0

osd_recovery_max_chunk

설명

내보낼 데이터 청크의 최대 크기입니다.

유형

64비트 정수가 서명되지 않았습니다

Default

8388608

osd_recovery_threads

설명

데이터 복구를 위한 스레드 수입니다.

유형

32비트 정수

Default

1

osd_recovery_thread_timeout

설명

복구 스레드의 타이밍을 줄이기 전 최대 시간(초)입니다.

유형

32비트 정수

Default

30

osd_recover_clone_overlap

설명

*복구 중에 복제 중복을 유지합니다. 항상 **true**로 설정해야 합니다.*

유형

부울

Default

true

rados_osd_op_timeout

설명

RADOS가 RADOS 작업에서 오류를 반환하기 전에 OSD에서 응답을 기다리는 시간(초)입니다. 값 0은 제한이 없음을 의미합니다.

유형

두 배

Default

0

부록 G. CEPH 모니터 및 OSD 구성 옵션

하트비트 설정을 수정할 때 Ceph 구성 파일의 [global] 섹션에 이를 포함합니다.

mon_osd_min_up_ratio

설명

Ceph OSD Daemon의 최소 비율은 Ceph OSD Daemon을 낮춥니다.

유형

두 배

Default

.3

mon_osd_min_in_ratio

설명

Ceph OSD 데몬에 비해 Ceph OSD 데몬의 최소 비율은 Ceph OSD Daemon 을 표시합니다.

유형

두 배

Default

0.750000

mon_osd_laggy_halflife

설명

지연 추정 시간 (초)은 감소합니다.

유형

정수

Default

60*60

mon_osd_laggy_weight

설명

laggy 추정 감소의 새 샘플의 가중치.

유형

두 배

Default

0.3

mon_osd_laggy_max_interval

설명

laggy 추정치의 최대 laggy_interval 값(초). 모니터는 적응형 접근법을 사용하여 특정 OSD 의 laggy_interval 을 평가합니다. 이 값은 해당 OSD의 유예 시간을 계산하는 데 사용됩니다.

유형

정수

Default

300

mon_osd_adjust_heartbeat_grace

설명

true 로 설정하면 Ceph가 지연 추정 에 따라 확장됩니다.

유형

부울

Default

true

mon_osd_adjust_down_out_interval

설명

true 로 설정하면 Ceph가 laggy 추정에 따라 확장됩니다.

유형

부울

Default

true

mon_osd_auto_mark_in

설명

Ceph는 모든 부팅 Ceph OSD 데몬을 Ceph Storage Cluster와 같이 표시합니다.

유형

부울

Default

false

mon_osd_auto_mark_auto_out_in

설명

Ceph는 클러스터와 같이 Ceph Storage 클러스터에서 자동으로 표시된 Ceph OSD 데몬 을 부팅 으로 표시합니다.

유형

부울

Default

true

mon_osd_auto_mark_new_in

설명

Ceph는 새 Ceph OSD 데몬을 Ceph Storage Cluster와 같이 부팅하는 것을 표시합니다.

유형

부울

Default

true

mon_osd_down_out_interval

설명

Ceph OSD 데몬이 응답하지 않는 경우 Ceph OSD 데몬 을 종료하기 전에 대기 하는 시간(초) 입니다.

유형

32비트 정수

Default

600

mon_osd_downout_subtree_limit

설명

Ceph가 자동으로 표시하는 가장 큰 CRUSH 장치 유형입니다 .

유형

문자열

Default

rack

mon_osd_reporter_subtree_level

설명

이 설정은 보고 OSD의 상위 CRUSH 장치 유형을 정의합니다. OSD에서 응답하지 않는 피어를 찾으면 모니터에 실패 보고서를 보냅니다. 모니터는 유예 기간 후에 보고된 OSD를 아래로 표시한 다음 종료할 수 있습니다.

유형

문자열

Default

호스트

mon_osd_report_timeout

설명

*응답하지 않는 **Ceph OSD** 데몬을 선언하기 전 유예 기간(초)입니다 .*

유형

32비트 정수

Default

900

mon_osd_min_down_reporters

설명

***Ceph OSD** 데몬을 보고하는 데 필요한 최소 **Ceph OSD** 데몬 수입니다.*

유형

32비트 정수

Default

2

osd_heartbeat_address

설명

*하트비트를 위한 **Ceph OSD** 데몬의 네트워크 주소입니다.*

유형

주소

Default

호스트 주소입니다.

osd_heartbeat_interval

설명

***Ceph OSD** 데몬이 피어를 **ping**하는 빈도(초)입니다.*

유형

32비트 정수

Default

6

osd_heartbeat_grace

설명

Ceph OSD Daemon이 Ceph 스토리지 클러스터가 다운 하는 하트비트를 표시하지 않은 경과 시간입니다.

유형

32비트 정수

Default

20

osd_mon_heartbeat_interval

설명

Ceph OSD Daemon 피어가 없는 경우 Ceph OSD 데몬이 Ceph 모니터를 ping하는 빈도입니다.

유형

32비트 정수

Default

30

osd_mon_report_interval_max

설명

Ceph OSD 데몬이 Ceph 모니터에 보고해야 하는 최대 시간(초)입니다.

유형

32비트 정수

Default

120

osd_mon_report_interval_min

설명

Ceph OSD 데몬이 Ceph 모니터에 보고하기 전에 시작 또는 보고 가능한 이벤트에서 대기할 수 있는 최소 시간(초)입니다.

유형

32비트 정수

Default

5

유효한 범위

osd mon 보고서 간격 최대보다 작어야합니다

osd_mon_ack_timeout

설명

Ceph 모니터가 통계 요청을 승인할 때까지 대기하는 시간(초)입니다.

유형

32비트 정수

Default

30

부록 H. CEPH 스크럽 옵션

Ceph는 배치 그룹을 제거하여 데이터 무결성을 보장합니다. 다음은 스크럽 작업을 늘리거나 줄이기 위해 Ceph에서 조정할 수 있는 옵션입니다.

osd_max_scrubs

설명

Ceph OSD 데몬에 대한 최대 동시 스크럽 작업 수입니다.

유형

integer

Default

1

osd_scrub_begin_hour

설명

스크럽이 시작되는 특정 시간입니다. osd_scrub_end_hour 과 함께 스크럽이 발생할 수 있는 시간 창을 정의할 수 있습니다. osd_scrub_begin_hour = 0 및 osd_scrub_end_hour = 0 을 사용하여 하루 전체를 스크럽할 수 있습니다.

유형

integer

Default

0

허용되는 범위

[0, 23]

osd_scrub_end_hour

설명

스크럽이 종료되는 특정 시간입니다. osd_scrub_begin_hour 과 함께, 스크럽이 발생할 수 있는 시간 창을 정의할 수 있습니다. osd_scrub_begin_hour = 0 및 osd_scrub_end_hour = 0 을 사용하여 하루 동안 스크럽을 수행할 수 있습니다.

유형

integer

Default

0

허용되는 범위

[0, 23]

osd_scrub_begin_week_day

설명

*스크럽이 시작되는 특정 날입니다. **0 = Sunday, 1 = Monday, etc.***
"osd_scrub_end_week_day" 와 함께 스크럽이 발생할 수 있는 시간 창을 정의할 수 있습니다.
osd_scrub_begin_week_day = 0 과 osd_scrub_end_week_day = 0 을 사용하여 일주일 동안 스크럽을 수행할 수 있습니다.

유형

integer

Default

0

허용되는 범위

[0, 6]

osd_scrub_end_week_day

설명

*이는 스크럽이 끝나는 날을 정의합니다. **0 = Sunday, 1 = Monday, etc.***
osd_scrub_begin_week_day 와 함께, 스크럽이 발생할 수 있는 시간 창을 정의합니다.
osd_scrub_begin_week_day = 0 과 osd_scrub_end_week_day = 0 을 사용하여 일주일 동안 스크럽을 수행할 수 있습니다.

유형

integer

Default

0

허용되는 범위

[0, 6]

osd_scrub_during_recovery

설명

*복구 중에 스크럽을 허용합니다. **false** 로 설정하면 활성 복구가 있는 동안 새 **scrub** 및 **deep-scrub** 예약이 비활성화됩니다. 이미 실행 중인 **scrubs**는 계속해서 사용 중인 스토리지 클러스터의 부하를 줄이는 데 유용합니다.*

유형

boolean

Default

false

osd_scrub_load_threshold

설명

*정규화된 최대 로드입니다. **getloadavg() / 온라인 CPU** 수에 정의된 대로 시스템 로드가 정의된 번호보다 높으면 스크럽이 발생하지 않습니다.*

유형

부동 값

Default

0.5

osd_scrub_min_interval

설명

***Ceph** 스토리지 클러스터 로드가 낮을 때 **Ceph OSD** 데몬을 스크럽하는 최소 간격(초)입니다.*

유형

부동 값

Default

1일

osd_scrub_max_interval

설명

*클러스터 로드와 관계없이 **Ceph OSD** 데몬을 스크럽하는 최대 간격(초)입니다.*

유형

부동 값

Default

7일

osd_scrub_chunk_min

설명

*단일 작업 중에 스크럽에 대한 최소 오브젝트 저장소 청크 수입니다. **Ceph** 블록은 스크럽 중에 단일 청크에 쓰기를 차단합니다.*

type

integer

Default

5

osd_scrub_chunk_max

설명

단일 작업 중에 스크럽을 스크럽할 최대 오브젝트 저장소 청크 수입니다.

type

integer

Default

25

osd_scrub_sleep

설명

다음 청크 그룹을 스크럽하기 전에 잠 자는 시간입니다. 이 값을 늘리면 클라이언트 작업에 영향을 미치지 않도록 스크럽의 전체 속도가 느려집니다.

type

부동 값

Default

0.0

osd_deep_scrub_interval

설명

*딥 스크럽을 위한 간격은 모든 데이터를 완전히 읽습니다. **osd_scrub_load_threshold** 는 이 설정에 영향을 미치지 않습니다.*

type

부동 값

Default

7일

osd_scrub_interval_randomize_ratio

설명

*배치 그룹에 대한 다음 **scrub** 작업을 예약할 때 **osd_scrub_min_interval** 에 임의의 지연을 추가합니다. 지연은 **osd_scrub_min_interval * osd_scrub_interval_randomized_ratio** 보다 작은 임의의 값입니다. 기본 설정은 **[1, 1.5] * osd_scrub_min_interval** 의 허용된 시간 창에 걸쳐 스크럽을 분배합니다.*

type

부동 값

Default

0.5

osd_deep_scrub_stride

설명

심층 스크럽을 수행할 때 크기 읽기.

type

크기

Default

512KB

osd_scrub_auto_repair_num_errors

설명

이러한 많은 오류가 발견된 경우에는 자동 복구가 수행되지 않습니다.

type

integer

Default

5

osd_scrub_auto_repair

설명

*이 값을 **true** 로 설정하면 **scrubs** 또는 **deep-scrubs**에서 오류를 찾을 때 자동 배치 그룹(PG)을 복구할 수 있습니다. 그러나 **osd_scrub_auto_repair_num_errors** 오류를 초과하는 경우 복구가 수행되지 않습니다.*

type

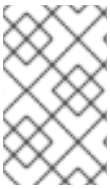
boolean

Default

false

부록 I. BLUESTORE 구성 옵션

다음은 배포 중에 구성할 수 있는 Ceph BlueStore 구성 옵션입니다.



참고

이 목록은 완료되지 않았습니다.

rocksdb_cache_size

설명

kmsDB 캐시의 크기(MB).

유형

32비트 정수

Default

512