

오브젝트 스토리지에서의 OSD 분배 구조에 따른 성능 Trade-Off 분석

주정현, 손재기, 김동민*

*한국전자기술연구원 메디컬IT융합연구센터
jujeong0108@keti.re.kr, jgson@keti.re.kr, *dmkim@keti.re.kr

Performance Trade-off Analysis of OSD Distribution Structure in Object Storage Systems

Jeonghyeon Ju, Jaegi Son, Dongmin Kim

*Korea Electronics Technology Institute Medical IT Convergence Research Center

요 약

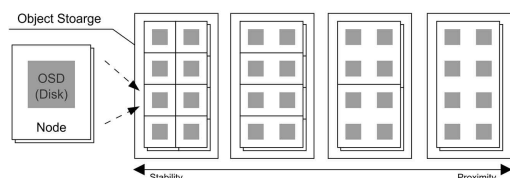
분산 오브젝트 스토리지는 대규모 데이터 처리 환경에서 확장성과 가용성을 동시에 확보하기 위한 핵심 기술로 발전해왔다. 그러나 실제 산업 및 연구 환경에서의 하드웨어 자원은 한정되어 있기에 이론적으로 제시되는 오브젝트 스토리지의 확장성 모델에는 한계가 존재한다. 이에 제한된 자원 내에서 현실적으로 최적의 성능을 확보할 수 있는 OSD(Object Storage Daemon) 분배 지표와 평가 기준의 필요성이 존재한다. 본 연구에서는 동일한 디스크 자원 내에서 노드 간 OSD 분배 구조 조정을 통해 스토리지의 성능을 최대한 유지하면서 안정성을 확보할 수 있는 최적의 분배 구조를 분석한다. 이를 통해 한정된 자원 내에서 적용 가능한 효율적인 오브젝트 스토리지 구조 설계 방안을 제시한다.

1. 서 론

클라우드 환경의 확산과 함께, 대용량 데이터를 안전하고 효율적으로 저장 및 관리가 가능한 스토리지 시스템의 중요성이 커지고 있다. 특히 분산 오브젝트 스토리지는 데이터를 여러 노드에 분산 저장하고 복제하는 구조를 통해 고가용성과 무결성을 확보한다[1]. 이러한 구조는 일반적으로 저장 노드 수와 데이터 단위 수의 확장에 따라 처리 성능이 향상되도록 설계되어 있다[2]. 그러나 실제 환경에서는 네트워크 부하, 자원 불균형, 복제 지연 등의 요인으로 인해 노드와 저장 단위 간의 분배 구조가 전체 성능에 큰 영향을 미친다. 특히, 하드웨어 자원이 한정된 환경에서 OSD가 배포되는 디스크의 분포 밀도는 전체 처리량과 지연 시간에 직접적인 영향을 미친다. 본 논문에서는 오브젝트 스토리지를 구성하는 노드 간 OSD 분배 구조 조정 실험을 통해 한정된 자원 내에서 적용 가능한 최적의 자원 분배 구조 설계 방안을 제시한다.

2. 본 론

본 논문에서는 한정된 자원 환경에서의 OSD 구조 설계를 중심으로 성능 평가 실험을 수행하였다.



<그림 1> OSD 분배에 따른 안정성·성능 Trade-Off

오브젝트 스토리지에서의 데이터 복제·복구 측면의 안정성을 유지하면서 처리량, 복제 지연 및 응답 지연 측면에서의 성능 저하를 최소화할 수 있는 최적의 OSD 분배 구성을 도출하기 위해, OSD 수를 동일하게 유지한 환경에서 다양한 배치 시나리오를 설계한다.

2-1. 실험 환경 구성

한정된 자원 환경에서의 OSD 분배 구조가 성능에 미치는 영향을 분석하기 위해 동일한 OSD 수를 기준으로 노드 간 분배 구조를 변경하며 실험을 수행하였다. 표 1은 실험에 사용된 시스템 사양 및 구성을 나타낸다.

Category	Spec.
vCPU	4 Core(per VM)
Mem.	8 GB(per VM)
Storage	50GB QEMU vDisk(SCSI) × 24
Network	82599ES 10-Gigabit SFI/SFP+
OS	Ubuntu 24.04.5
Ceph Version	Rook-Ceph:v1.18.4

<표 1> 시스템 및 오브젝트 스토리지 사양

실험은 총 24개의 OSD를 기준으로 단일 노드 집중 배치에서부터 24노드 분산 배치까지의 시나리오를 구성하였다. 각 노드는 동일한 하드웨어 사양을 유지하였으며, OSD 분배 구조가 성능에 미치는 영향을 독립적으로 분석하기 위해 MON, MGR 등의 관리 데몬은 별도의 고정된 노드에 고정하여 실험 결과에 영향을 주지 않도록 제어하였다. 표 2에서는 실험에 사용된 주요 환경 변수 및 실험 파라미터를 정리하였다.

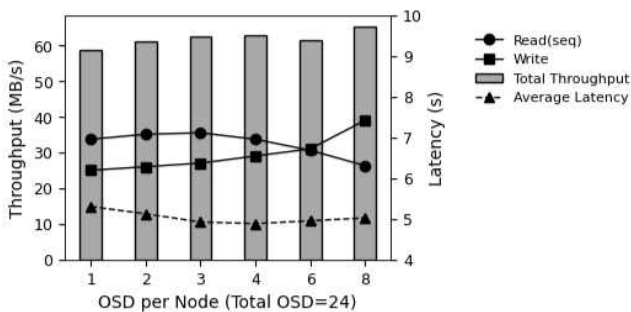
Category	Value	Description
Object Size	4 MB	Recommended base object size
Threads	16	Recommended thread amount
Operation	seq/write	Benchmark modes for read/write
Test Duration	60s	Stable throughput averaging time
Page Number	128	Fixed to minimize OSD imbalance
Replication Size	1 & 3	Performance comparison purpose

<표 2> OSD 분배 구조 성능 실험 환경 변수

2-2. 실험 절차 및 결과

해당 실험에서는 OSD 분배 구조 변화가 성능에 미치는 영향을 정량적으로 분석하기 위해 동일한 하드웨어 및 소프트웨어 환경에서 반복 측정을 수행하였다. 모든 측정은 Ceph의 내장 벤치마크 도구인 Rados Bench를 활용하여 수행하였다. 각 시나리오별 테스트는 순차적 읽기(seq) 및 쓰기 작업으로 구분하여 수행하였으며, 각 구조별 처리량(Throughput)과 객체 단위 지연시간(Latency)을 기준으로 비교·분석하였다. 또한, 오브젝트 스토리지의 일반적인 구성인 Replica=3 환경을 기준으로 하되, OSD 분포 구조의 영향을 독립적으로 검증하기 위해 복제 부하 요인을 배제한 Replica=1 환경을 추가적으로 구성하여 비교 실험을 진행하였다.

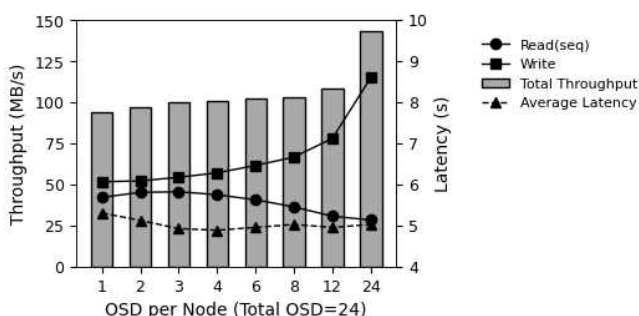
OSD Distribution Efficiency (Replica=3)



<그림 2> OSD 분포에 따른 클러스터 처리 성능 비교(Replica=3)

그림 2는 Replica=3 환경에서 노드당 OSD 수에 따른 처리 성능 변화를 나타낸다. OSD 분산도가 높을수록 읽기 성능이 향상되는 경향을 보였으나, 과도한 분산의 경우 성능이 점차 저하되는 양상을 보였다. 이는 OSD 병렬 처리가 일정 수준까지는 읽기 효율을 향상시키지만, 네트워크 대역폭 한계를 초과하는 과도한 분산으로 인한 병목 현상이 성능 저하를 유발한 것으로 판단된다. 반면, 쓰기 성능은 전반적으로 OSD 분산도가 낮을수록 증가하는 경향을 보였다. 이는 동일 노드 내 통신을 통해 복제 및 데이터 확인 과정의 오버헤드가 감소해 처리 효율이 향상된 결과로 해석된다.

OSD Distribution Efficiency (Replica=1)



<그림 3> OSD 분포에 따른 클러스터 처리 성능 비교(Replica=1)

그림 3은 Replica=1 환경에서 노드당 OSD 수에 따른 처리 성능 변화를 나타낸다. 쓰기 성능의 경우 전반적으로 Replica=3 환경보다 높게 나타났으며, 특히 단일 노드에 집중될수록 급격하게 향상되는 경향을 보였다. 이는 데이터 복제 및 노드 간 검증 과정이 생략되면서 통신 오버헤드가 사라지고, 모든 I/O 처리가 단일 노드 내에서 이루어진 데에 기인한다. 그러나 이러한 구성은 단일 노드 장애 발생 시 전체 데이터 손실로 이어질 가능성이 높아 실제 서비스 환경에서는 적용이 권장되지 않는다[3]. 읽기 성능의 경우 Replica=3 환경과 전반적으로 유사한 수준으로 유지되었다. Latency의 경우 일정 수준까지의 OSD 집중은 지연 감소에 유리하였으나, 지나친 집중은 지연 시간이 증가하는 경향을 보였다. 이는 노드 내에서의 과도한 OSD 집중으로 인해 I/O 큐 부하가 증가한 결과로 해석된다.

3. 결론

본 연구에서는 한정된 자원 환경에서의 오브젝트 스토리지 성능 최적화를 목표로, 노드 간 OSD 분배 구조 변화가 시스템 처리 효율에 미치는 영향을 분석하였다. 실험 결과, 읽기 성능은 OSD 분산도가 높을수록 향상되는 경향을 보였으나, 과도한 분산은 네트워크 병목으로 인해 성능 저하를 유발하였다. 반면, 쓰기 성능은 OSD가 소수의 노드에 집중될수록 향상되는 경향을 보였다. 이러한 결과는 스토리지의 용도와 접근 패턴에 따라 최적의 분배 구조가 달라질 수 있음을 시사한다. 즉, 특정 목적에 맞게 설계된 스토리지의 경우 읽기 또는 쓰기 성능의 최대치를 목표로 한 구조적 편향이 가능하나, 일반적인 환경에서는 균형 잡힌 분배 구조의 설계가 필요하다. 따라서 스토리지의 목적과 운용 환경에 따라 OSD 분배 구조의 유연한 설계가 필요하며, 본 연구는 그 판단의 기초 자료로 활용될 수 있다.

ACKNOWLEDGMENT

이 논문은 2025년 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임(No. RS-2025-02215256), AI 반도체 클라우드를 위한 고속 병렬 파일시스템 기술 개발

참 고 문 헌

- [1] Ceph Project, Ceph Architecture Documentation (Reef Release), Ceph Official Documentation, 2024. Available: <https://docs.ceph.com/en/reef/architecture/>
- [2] Zhu, Hang, et al. "Harmonia: Near-linear scalability for replicated storage with in-network conflict detection." arXiv preprint arXiv:1904.08964 (2019).
- [3] Ramabhadran, Sriram, and Joseph Pasquale. "Analysis of durability in replicated distributed storage systems." 2010 IEEE International Symposium on Parallel & Distributed Processing (IPDPS). IEEE, 2010.