

# HDFS 이기종 스토리지 전략을 위한 스몰파일 아카이빙 기법

## Small File Archiving Strategy for Heterogeneous Storage in HDFS

박주원  
Juwon Pak, Ta



### 요 약

Hadoop의 HDFS는 대용량 데이터를 효율적으로 처리하기 위해 설계된 분산파일시스템이다. Hadoop에서는 스몰파일을 효율적으로 처리하기 위해 HDFS에 대한 아카이빙 기법들을 제공하여, 스몰파일들을 대용량 파일처럼 처리한다. 하지만 현재의 스몰파일 아카이브는 HDFS의 이기종 스토리지 전략을 수행하기 위해 각 정책에 맞게 스몰파일을 디렉토리 별로 분류하고 각 디렉토리에서 아카이빙 과정을 거쳐야 하는 등 SSD 등이 포함된 이기종 스토리지 전략을 유연하게 사용할 수 없다. 본 논문에서는 이기종 스토리지 전략을 효율적으로 수행하기 위해 스몰파일의 액세스 빈도에 따라 온도를 분류하고, 한 아카이브 내에서 분류한 온도에 맞게 HDFS 스토리지 정책을 설정할 수 있는 아카이빙 기법을 제안한다. 프로토타이핑을 통해 제안 기법을 구현하고, 실험을 통해 기존 방식 대비 성능을 분석하였다.

키워드: HDFS, Small File, Hot Data, Heterogeneous Storage

### 1. 서론

빅데이터 플랫폼인 Hadoop[1]은 대용량 데이터를 여러 서버에 분산 저장하는 파일시스템인 HDFS[2]에서 스몰파일을 효율적으로 처리하기 위해 스몰파일 여러 개를 병합하여 단일파일로 아카이빙(archiving)하는 기법들[3-5]을 제공하고 있다.

그러나 RAM, SSD(solid-state disk), HDD, ARCHIVE 등을 함께 사용하는 이기종 스토리지에서 기존 아카이빙 기법들은 HDFS 스토리지 정책을 효율적으로 적용하기 어렵다. Hadoop은 2.6.0 버전 이후로 이기종 스토리지를 지원하며, HDFS 스토리지 정책(default)은 각 데이터 블록을 세 개를 복사하여 각각 다른 노드의 디스크에 저장한다. 이기종 스토리지의 역할은 여러 유형의 스토리지와 함께 더 많은 스토리지 전략을 제공하는 것이다. 데이터 액세스 빈도가 높은 데이터의 경우 3개의

복제본을 모두 SSD 저장하고 데이터 액세스 빈도가 낮은 데이터의 경우 3개의 복제본을 모두 HDD에 저장하는 등 특정 디렉토리에 알맞은 정책을 설정하여 HDFS 스토리지를 보다 유연하고 효율적으로 구성할 수 있다. 따라서 데이터들을 액세스 패턴을 기반으로 데이터의 선호 정도를 측정한 데이터 온도[6]로 분류하게 된다면, 효과적인 전략 수립이 가능하다.

하지만 아카이브된 스몰파일의 경우 한 디렉토리 내에 존재하므로 하나의 정책만 설정할 수 있어 액세스 패턴에 따라 이기종 스토리지 전략을 충분히 수행할 수 없다. 즉, 기존 아카이빙 기법의 경우 HDFS의 이기종 스토리지 전략을 수행하기 위해 각 정책에 맞게 스몰파일을 분류하고 알맞은 디렉토리로 이동시켜 각각 아카이빙 과정을 거쳐야 하므로 번거롭고 시간 비용이 효율적이지 않다.

\* 교신저자

&lt;표 1&gt; HDFS 스토리지 정책 [7]

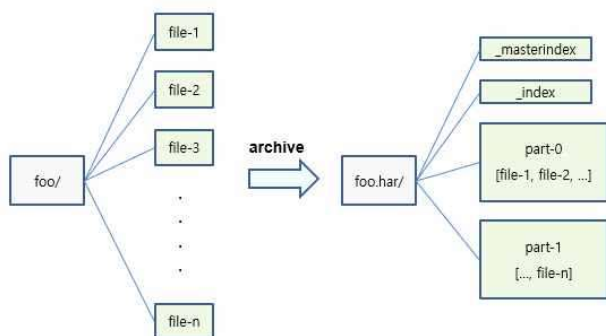
정책	블록 배치 (n replicas)
Lazy_Persist	RAM_DISK: 1, DISK: n-1
All_SSD	SSD: n
One_SSD	SSD: 1, DISK: n-1
Hot (default)	DISK: n
Warm	DISK: 1, ARCHIVE: n-1
Cold	ARCHIVE: n

본 논문에서는 스몰파일에 효율적으로 이기종 스토리지 전략을 적용할 수 있는 기법을 제안한다. 제안하는 기법에서는 Hadoop에서 제공하는 기존 아카이빙 기법인 HAR의 아카이브 구조를 변형하여 구현했다.

## 2. 관련 연구

본 장에서는 HDFS에서 스몰파일에 대한 처리 성능을 향상하기 위해 Hadoop에서 제공하는 아카이빙 기법에 대해 알아보고 이기종 스토리지 환경에서 데이터의 액세스 패턴에 따라 데이터 온도를 정의하고 각 온도에 맞는 HDFS 스토리지 정책을 적용한 연구를 알아본다.

HAR (Hadoop archive)[3]은 스몰파일을 효율적으로 처리하기 위해 여러 스몰파일을 병합하여 하나의 큰 파일로 아카이브 하는 기법으로 Hadoop에서 제공하는 기법들 중 하나다. 그림 1은 HAR에서 스몰파일을 아카이빙 하는 과정이다. HAR을 통해 생성된 \*.har에는 \_masterindex와 \_index 파일로 총 2개의 인덱스 파일과 실제 스몰파일 데이터가 저장되어있는 part-\*로 구성되어 있다. \_masterindex에는 해시와 오프셋이 저장되어있고 \_index에는 스몰파일들의 메타데이터가 저장되어있다.



(그림 1) HAR에서 스몰파일 아카이브 과정 [8]

Subramanyam[6]는 이기종 스토리지 환경에서 HDFS 스토리지 정책(표 1)에 맞게 데이터를 저장하기 위해, 데이터의 액세스 패턴을 기준으로 데이터 온도를 설정하는 연구를 진행했다. 데이터마다 액세스 패턴을 기준으로 데이터 온도를 분류하면 액세스 빈도에 따라 알맞은 스토리지 정책을 설정할 수 있다. 표 2는 이 연구에서 데이터 온도를 분류하기 위한 데이터 저장 시간과 액세스 패턴의 기준이다. d, w, m은 1일, 1주, 1달 동안 데이터 액세스 횟수이다. 이 기준으로 데이터 온도를 분류하고 SUMMER의 데이터는 All\_SSD, FALL의 데이터는 One\_SSD, WINTER의 데이터는 Warm, FROZEN의 데이터는 Cold의 정책을 따른다. 그 외의 데이터의 경우에는 Hot 정책을 따른다.

&lt;표 2&gt; 액세스 패턴에 따른 데이터 온도 [6]

데이터 저장 시간	액세스 패턴	온도
<1 week	d>30	SUMMER
>1 week, <1 month	d>15, w>30	FALL
>1 month, <3 months	w=0, m>0	WINTER
>3 months, <1 year	m=0	FROZEN

스몰파일에 대한 이기종 스토리지 전략 수행을 위해 기존 아카이빙 기법인 HAR을 사용하면, HDFS 스토리지 정책에 따라 분류된 디렉토리를 생성한 후 스몰파일들을 각 디렉토리에 이동하고 각 디렉토리에서 아카이빙을 해야 한다. 이는 HDFS 스토리지 정책이 디렉토리를 기준으로 설정될 수 있기 때문이다. 이 경우, 스몰파일 이동 시간과 각 정책마다 아카이빙 과정이 필요하기 때문에 시간 비용이 효율적이지 않다. 따라서 스몰파일에 대해 효율적으로 이기종 스토리지 전략을 적용하는 연구가 필요하다.

## 3. 이기종 스토리지 전략을 위한 아카이빙 기법

본 장에서는 제안 기법인 이기종 스토리지에서 스몰파일의 액세스 패턴에 따라 온도를 분류하고, 이에 맞는 HDFS 스토리지 정책을 효율적으로 적용하기 위한 스몰파일 아카이브 기법을 설명한다.

### 3.1 스몰파일의 온도 분류

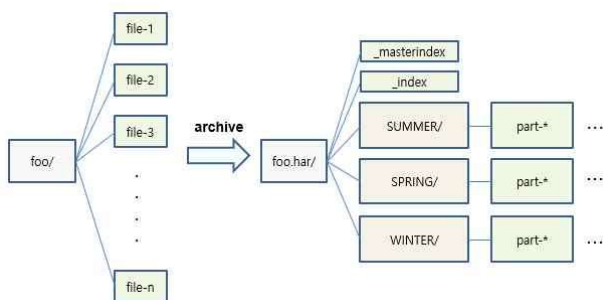
본 제안에서는 스몰파일의 온도를 분류하기 위해 스몰파일의 메타데이터 구조를 변형했다. 스몰파일의 기존 메타데이터에 온도를 분류할 수 있는 int

타입(4 bytes)의 변수인 스몰파일 액세스 횟수를 추가했다. 액세스 횟수는 스몰파일이 액세스 될 때 마다 1씩 증가한다.

스몰파일의 온도를 분류할 때 Subramanyam의 데이터 온도 분류 연구를 참고하여, 본 제안 조건에 적합하도록 수정하였다. HAR을 통해 생성된 아카이브는 한 번 재생성 되면 수정 불가능 하기 때문에 실시간으로 액세스 패턴을 반영하여 파일을 정책에 맞게 이동시키기 어렵다. 즉, 스몰파일의 액세스 패턴에 따라 온도를 분류하기 위해 아카이브를 재생성해야 한다. 따라서 본 연구에서는 HDD와 SSD 스토리지만을 고려하여 아카이브 재생성 주기를 1달로 정했다. Subramanyam의 연구에 따르면 ARCHIVE 스토리지를 사용하는 Warm과 Cold 정책의 경우 저장 시간이 3달 및 1년 이상인 데이터들에 맞는 정책이기 때문이다. 표 3은 과거 1달 동안 스몰파일의 액세스 횟수에 따른 온도 분류 기준과 각 온도의 정책이다. 아카이브 파일의 재생성 주기를 1달로 설정함으로써 아카이브 파일이 재생성될 때 스몰파일의 과거 1달간의 액세스 횟수를 기준으로 온도를 분류했다. 따라서 초기 1달은 모든 스몰파일이 WINTER의 온도로 분류된다.

<표 3> 액세스 횟수(1달)에 따른 스몰파일 온도 분류

과거 액세스 횟수	온도	정책
>=900	SUMMER	All_SSD
>=450	SPRING	One_SSD
>=0	WINTER	Hot(default)



(그림 2) 이기종 스토리지 전략을 위한 아카이브 구조

### 3.2 아카이브 구조

본 연구에서 제안하는 아카이브 구조는 Hadoop에서 제공하는 아카이빙 기법인 HAR의 아카이브

구조 중 part 부분을 변형하여 구현했다. 그림 2는 제안하는 아카이브 구조이며 분류한 온도에 맞는 HDFS 스토리지 정책을 설정하기 위해 아카이브 내에 각 온도의 디렉토리를 구성하였다. 온도 디렉토리 내에는 해당 온도로 분류된 실제 스몰파일들이 저장되어있는 part-\* 파일로 구성하였다. 따라서 제안된 기법은 한 번의 아카이빙 과정을 수행할 때 각 온도에 따라 디렉토리를 생성하고 해당 디렉토리에 스토리지 정책을 설정하는 과정을 포함한다.

## 4. 실험

본 장에서는 제안 기법과 HAR을 기반으로 수행한 실험 결과를 비교 분석한다.

### 4.1 실험 환경

실험 환경은 Hadoop 3.2.1 버전에서 네임노드 1개와 데이터노드 6개로 구성하였으며 각 노드의 성능은 표 4와 같다. 데이터노드의 경우 HDD 서버 3개와 SSD 서버 3개로 구성했다. 실험에 사용한 데이터는 1MB 크기의 스몰파일을 사용하였고 개수는 3000개, 4000개, 5000개로 늘려가며 실험하였다. 아카이브를 생성할 때는 Replication Factor를 default인 3으로 설정하였다. 스몰파일 온도의 경우 데이터노드의 HDD와 SSD의 용량 비율을 토대로 SUMMER 10%, SPRING 30%, WINTER 60% 비율로 구성하였다.

<표 4> 노드 성능

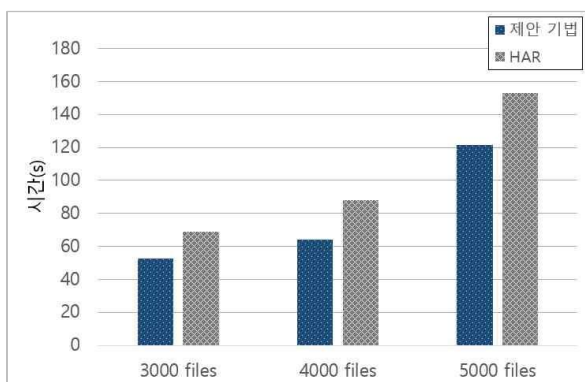
CPU	Intel i7-8700
Memory	32 GB
HDD	1 TB
SSD	256 GB
Network Bandwidth	1 Gbps
OS	Ubuntu 18.04

### 4.2 실험 결과

본 실험에서는 스몰파일들이 액세스 패턴에 따라 SUMMER, SPRING, WINTER의 온도로 이미 분류되었다고 가정하였다. 기존 기법인 HAR의 경우 온도마다 총 개수의 10%, 30%, 60%의 스몰파일들을 각각 아카이빙하여 3번의 아카이브를 생성한 시간을 측정하였고, 제안 기법의 경우 전체 스몰파일

을 아카이빙하여 1개의 아카이브를 생성한 시간과 아카이브 내부에 온도 디렉토리 생성 및 정책 설정 시간을 포함하여 측정하였다.

그림 3은 스물파일 개수 별 제안 기법과 기존 기법인 HAR을 통해 이기종 스토리지 전략에 맞는 아카이브 생성되는 시간을 비교한 실험 결과이다. 3000개의 파일에서 HAR은 약 69초, 제안 기법의 경우 약 52초의 시간으로 이기종 스토리지 전략에 맞는 아카이브를 생성했다. 4000개와 5000개의 스물파일에서도 제안 기법의 경우 약 64초, 122초의 시간이 소요되었지만 HAR의 경우에는 약 88초, 153초의 시간이 소요되었다. 이 실험을 통해 제안 기법이 기존 기법 대비 이기종 스토리지 전략을 적용한 아카이브 생성에 대해 시간 효율이 좋음을 확인하였다.



(그림 3) 이기종 스토리지 전략 수행을 위한 아카이브 생성 시간 비교 실험

## 5. 결론 및 향후 연구

본 논문에서는 스물파일에 이기종 스토리지 전략을 효과적으로 수행시키기 위한 기법을 제안했다. 실험을 통해 제안 기법은 이기종 스토리지 전략을 효율적으로 수행할 수 있음을 확인했다. 향후에는 아카이브 파일을 재생성할 필요 없는 구조를 연구하고 데이터 온도를 분류하는 기준을 개선하는 연구를 진행할 것이다.

## Acknowledgement

본 연구는 과학기술정보통신부 및 정보통신기획평가원의 대학ICT연구센터지원사업(IITP-2020-2018-0-01431)과 한국연구재단 기초연구사업(2018R1D1A1B07043858)의 지원을 받아 수행된 연구임

## 참고문헌

- [1] Tom White, "Hadoop: The Definitive Guide: Storage and Analysis at Internet Scale 4th Edition," O'Reilly Media, 2015.
- [2] K. Shvachko, H. Kuang, S. Radia, and R. Chansler, "The hadoop distributed file system," in Mass storage systems and technologies (MSST), 2010 IEEE 26th symposium, pp. 1 - 10, 2010.
- [3] Joe Sremack, "Big Data Forensics - Learning Hadoop Investigations," pp.41-43, 2015.
- [4] Hrishikesh Vijay Karambelkar, "Apache Hadoop 3 Quick Start Guide: Learn about big data processing and analytics," pp.78-80, 2018.
- [5] Mehul Nalin Vora, "Hadoop-HBase for large-scale data," International Conference on Computer Science Network Technology, IEEE, vol. 1, pp. 601-605, 2011.
- [6] R. Subramanyam, "HDFS heterogeneous storage resource management based on data temperature," International Conference on Cloud and Autonomic computing (ICCAC), Springer, pp. 232 - 235, 2015.
- [7] Archival Storage, SSD & Memory, "https://hadoop.apache.org/docs/current/hadoop-project-dist/hadoop-hdfs/ArchivalStorage.html," Accessed: 2020.07.
- [8] C. Vorapongkitipun, N. Nupairoj, "Improving performance of small-file accessing in hadoop," Computer Science and Software Engineering (JCSSE), 11th International Joint Conference, IEEE, pp. 200 - 205, 2014.