

---

# 노드의 가용성을 고려한 하둡 태스크 할당 정책

류우석

부산가톨릭대학교

## Task Assignment Policy for Hadoop Considering Availability of Nodes

Wooseok Ryu

Catholic University of Pusan

E-mail : wsryu@cup.ac.kr

### 요 약

하둡 맵리듀스(MapReduce)는 사용자가 요청한 잡을 하둡 클러스터에서 효과적으로 병렬 분산 처리하기 위한 프레임워크이다. 맵리듀스의 태스크 스케줄러는 사용자의 잡 태스크들을 여러 노드에 할당하기 위한 기법이다. 하지만, 기존의 스케줄러는 노드의 가용 상태에 따라 규모가 동적으로 변화하는 하둡 클러스터를 고려하지 않음으로써 클러스터의 자원을 충분히 활용하지 못하는 문제가 있다. 본 논문에서는 노드의 가용성을 고려하여 잡 태스크를 효과적으로 할당함으로써 하둡 클러스터의 활용성을 높이는 태스크 할당 정책을 제시한다.

### 요 약

Hadoop MapReduce is a processing framework in which users' job can be efficiently processed in parallel and distributed ways on the Hadoop cluster. MapReduce task schedulers are used to select target nodes and assigns user's tasks to them. Previous schedulers cannot fully utilize resources of Hadoop cluster because they does not consider dynamic characteristics of cluster based on nodes' availability. To increase utilization of Hadoop cluster, this paper proposes a novel task assignment policy for MapReduce that assigns a job tasks to dynamic cluster efficiently by considering availability of each node.

### 키워드

하둡, 맵리듀스, 가용성, 태스크 할당

## 1. 서 론

하둡(Apache Hadoop)은 대용량으로 생성되는 빅 데이터를 분산 노드들에 효과적으로 저장하고 이를 병렬 분산 처리하기 위한 병렬 분산 플랫폼이다. 4차 산업혁명의 도래에 따라 빅 데이터의 가치가 보다 높아짐에 따라 다양한 응용 환경에서 하둡 플랫폼을 활용하여 빅 데이터를 분석하려는 요구가 급속도로 증가하고 있다[1].

하둡 플랫폼을 구축하기 위해서는 다수의 노드로 구성된 클러스터를 구축하여야 한다. 하둡 플랫폼은 클러스터의 규모를 수 개의 노드부터 수 천 개의 분산 노드까지 지원하는 규모 확장성을

지원하고 있다[2]. 하지만, 중소병원과 같은 중소 규모 기관의 경우 다양한 빅 데이터 분석 요구[3]에도 불구하고 비용 문제 등과 같은 여러 문제에 따라 해당 요구에 대응하는 대규모 분석 전용 클러스터를 구축하기가 어렵다.

이러한 문제를 해결하기 위한 기존 연구로서 해당 기관이 기존의 업무용 전산 자원을 활용하여 비교적 저렴한 비용으로 하둡 분산 클러스터를 구성하기 위한 노드의 시간-계약 기반 동적 하둡 클러스터 관리 기법[4] 및 노드의 가용성을 고려한 하둡 클러스터의 탄력적 관리 기법[5]이 연구되었다. 이는 하둡 분산 파일시스템(Hadoop Distributed File System, HDFS)에 대해 노드의

가용성을 고려하여 저장 가능한 노드 클러스터의 규모를 동적으로 확장하는 연구이다. 하지만, 기존 연구들은 사용자의 응용 프로그램을 병렬 분산 처리하기 위한 맵리듀스 프레임워크에 대한 고려는 전혀 이루어지고 있지 않은 문제가 있다. 본 논문에서는 노드의 가용성을 고려하여 맵리듀스의 잡 태스크들 효율적으로 할당하기 위한 방법을 제시하고자 한다.

## II. 시간-제약 기반 태스크 할당 정책

하둠 클러스터 내의 노드들에 대한 시간-제약 조건은 가상 랙(virtual rack)에 명시된 *operationalTime* 속성에 따라 가용 시간이 지정 가능하다[3]. 이때, 시간-제약 조건이 지정된 가상 랙에 위치한 노드들은 24시간 가용한 상태가 아니라 특정 시간에만 사용이 가능한 노드들이다. 해당 노드들이 시간-제약에 따라 사용 가능한 상태가 되면 기존 태스크 스케줄러들은 사용자가 제출한 잡 태스크를 해당 노드들에게 할당할 수 있다. 하지만 해당 노드들은 *operationalTime* 속성에 명시된 가용시간의 한계에 다다르면 더 이상 실행되지 않고 중지된다. 이때, 기존의 하둠은 이들을 장애 노드로 처리하며 해당 태스크들은 하둠의 장애 복구 정책에 따라 다른 노드에서 처음부터 다시 수행된다.

노드들에 시간-제약 조건이 주어진 상황에서 한계 가용시간까지 태스크가 수행되는 것을 막기 위한 태스크 할당의 기본 정책은 맵리듀스의 스케줄러에 *time\_limit* 속성을 새로 추가하고 이를 이용하여 태스크가 할당될 수 있는 시간을 설정하는 것이다. 이 속성은 하나의 태스크가 안전하게 실행을 완료할 수 있는 시간으로 정의할 수 있다. *operationalTime* 에서 제시된 허용가능시간의 타임스탬프가  $t_c$ 부터  $t_e$ 까지라면 해당 노드가 잡 태스크를 할당받을 수 있는 시간은  $t_c$ 부터  $t_e-t_1$ 이 된다. 이때  $t_1$ 은 *time\_limit* 속성으로 명시된 시간 간격을 의미한다. 이 값은 하둠 클러스터의 운용에 따라 구동되는 태스크들의 평균 및 최대 수행 시간을 고려하여 하둠 관리자가 설정할 수 있는 속성 값이다.

## III. 가상 랙을 고려한 태스크 할당 정책

가상 랙에 위치한 노드들은 실제 랙에 위치하고 있는 잘 구조화되어 운용되는 노드들과는 몇 가지 다른 특징이 있다. 단일 맵리듀스 잡이 여러 태스크를 생성하여 실행되는 과정에서 태스크들이 여러 가상 랙(virtual rack)을 거쳐서 실행되는 경우 노드 가용시간이 서로 다를 뿐만 아니라 태스크 실행 결과의 네트워크 전송 비용이 매우 높은 문제가 발생한다. 이 문제를 해결하기 위해서

는 가능한 잡 태스크들의 랙 지역성(rack locality)를 최대한 보장하여야 한다. 본 논문에서는 랙 지역성을 고려한 태스크 할당 정책을 다음과 같이 제시한다.

- 첫째, 잡의 실행이 요청되면 전체 클러스터 내의 랙들을 대상으로 잡의 실행에 필요한 데이터 블록 복제본을 모두 가지고 있는 랙을 탐색하고, 그중 하나의 랙의 노드들을 선택하여 해당 랙의 노드들에게 태스크를 할당한다. 이때 랙의 선택 우선순위는 1) 가상 랙, 2) 작은 규모의 랙이다.
- 둘째, 잡이 요구하는 태스크들을 단일 랙 안의 노드들에게 할당되지 못하는 경우 두 개 이상의 랙에 걸쳐서 할당해야 한다. 이 경우 잡이 요구하는 태스크를 가장 많이 수용할 수 있는 랙을 우선적으로 선택한 후 나머지 태스크들을 수용할 수 있는 랙을 추가로 선택하여 할당한다.
- 셋째, 위 정책에서 공통적으로 적용할 사항으로 노드의 선택 시 선택 가능한 노드는 이미 사용 중인 노드와 앞에서 기술한 *time\_limit* 속성에 따라 더 이상 할당할 수 없는 노드를 제외한 유휴 노드로 한정한다.

제안한 기법은 랙 지역성 및 데이터의 지역성을 고려하되, 노드의 가용성을 최대한 활용하는 정책이다. 가상 랙의 노드들은 항상 가용 상태에 있는 것이 아니라 시간-제약에 따라 선택적으로 사용 가능한 노드들이므로 가상 랙의 노드들이 가용 상태가 되면 이를 가능한 우선적으로 활용하는 태스크 할당 정책이라 할 수 있다.

이 정책은 하둠에서 제공하는 기본 태스크 스케줄러인 FIFO 스케줄러를 수정하여 구현이 가능하다. FIFO 스케줄러는 큐에 있는 잡들이 순차적으로 실행되고 각각의 잡은 클러스터내 모든 자원을 점유하는 정책이다. 하지만, 제안한 기법은 순서상으로 우선순위가 떨어지는 잡에 대해서도 위 정책을 통해 태스크를 수행할 노드들이 정해지면 바로 수행할 수 있으므로 FIFO 스케줄러보다 성능이 우수하다고 할 수 있다.

## IV. 결 론

본 논문에서는 하둠 클러스터의 규모가 시간 가용성에 따라 동적으로 변화하는 환경에서 사용자의 맵리듀스 잡을 효율적으로 처리하기 위한 잡 태스크 할당 정책을 제시하였다. 제안한 기법은 가용성 조건이 주어진 노드들을 데이터 분석에 적극적으로 활용하기 위해 노드의 시간-제약 조건 및 가상 랙의 지역성을 고려하여 태스크를 할당하는 정책이다. 이를 통해 기존의 유휴 전산 자원을 분산 클러스터에 편입시키고 빅 데이터 분석에 적극적으로 활용할 수 있으므로 비용 효

울적으로 하둡 클러스터를 활용할 수 있다. 본 논문의 향후 연구는 제안한 태스크 할당 정책을 하둡 태스크 스케줄러에 구현하고 기존의 태스크 할당 정책과 그 성능을 비교 평가하는 것이다.

## V. ACKNOWLEDGEMENT

이 연구는 2016년도 정부(미래창조과학부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. NRF-2016R1C1B1012364).

### 참고문헌

- [1] Schwab, K., "The Fourth Industrial Revolution", Penguin UK, 2017.
- [2] Shvachko, K. et al, "The hadoop distributed file system," In 2010 IEEE 26th symposium on mass storage systems and technologies (MSST), pp. 1-10, 2010.
- [3] Raghupathi, W. and Raghupathi, V., "Big data analytics in healthcare: promise and potential," Health Information Science and Systems, vol. 2, article 3, pp. 1 - 10, 2014.
- [4] Ryu, W., "Increasing Scalability of Hadoop Cluster Exploiting Time-Constraints of Distributed Nodes," The Third Asia Workshop on IT Convergence of KIICE, pp. 31-33, 2017.
- [5] Ryu, W., "Flexible management of data nodes for hadoop distributed file system," In Proc. Int. Conf. on Big Data, Small Data, Linked Data and Open Data (ALLDATA 2017), pp. 1-2, 2017.