

# 스파크 기반 분산 환경에서 슬레이브 노드의 개수에 따른 성능 분석과 예측

박봉우\*, 명노영\*,정광식\*\* 유현창\*, 최숙경\*

\*고려대학교 컴퓨터학과

\*\*한국방송통신대학교 컴퓨터학과

e-mail:{bongwoo, mry1811, yuhc, csukyong}@korea.ac.kr

kchung0825@knou.ac.kr

## Performance evaluation and prediction for number of slave nodes in Spark

Bongwoo Bak\*, Rohyoung Myung\*, KwangSik Chung\*\*, Heonchang Yu\*, Sukyong Choi\*

\*Dept of Computer Science and Engineering, Korea University

\*\*Dept of Computer Science, Korea National Open University

### 요 약

최근 빅 데이터를 이용한 시스템들이 여러 분야에서 활발히 이용되기 시작하면서 대표적인 빅 데이터 저장 및 처리 플랫폼인 하둡(Hadoop)의 기술적 단점을 보완할 수 있는 분산 시스템 플랫폼 스파크(Apache Spark)가 등장하였다. 본 플랫폼을 바탕으로 슬레이브 노드들에게 작업을 분산하여 대용량 연산을 수행한다. 하지만 요구하는 성능을 내기 위해 어느 정도 규모의 슬레이브 노드가 필요한지, 각각의 컴퓨팅 능력은 얼마나 필요한지를 예측하는데 어려움이 있다. 본 논문에서는 스파크에서 원하는 성능을 내기 위해 어떤 조건을 충족해야 하는지, 현재 환경에서는 어느 정도 성능을 낼 수 있는지 실험을 통해 모델을 만들어 예측한다.

다루고 3장에서는 실험 환경과 실험방법 그리고 예측 모델에 대해 소개한 후에 4장에서는 실험 결과에 대하여 분석하고 마지막으로 5장에서는 연구에 대한 평가와 향후 연구 계획에 대하여 다룰 것이다.

### 1. 서론

아파치 재단의 스파크(Spark) [1]는 최근 각광받는 클러스터 컴퓨팅 프레임워크이다. 기존에 있던 하둡(Hadoop) [4] 플랫폼의 속도 저하 단점을 보완하기 위해 인 메모리 처리를 지원하여 대용량 데이터를 효율적으로 처리할 수 있다. 이 때문에 여러 기계학습 알고리즘에서 유용하게 사용되고 있다. 하지만 빅 데이터를 기반으로 하는 스파크는 그 환경, 클러스터(cluster) 구성에 따라 그 성능이 쉽게 바뀐다. 즉 맵리듀스 작업을 슬레이브 노드들에게 분산 시켜 연산을 수행하기 때문에 슬레이브 노드들의 수, 컴퓨팅 능력에 따라 분산 환경 전체의 성능이 좌우된다. 그래서 원하는 성능의 빅 데이터 처리 플랫폼을 구성하기 위해서는 환경을 적절하게 구성해야 하는데 그 기준이 모호하며 예측하기 어렵다. 본 논문에서는, 스파크 플랫폼에서 슬레이브 노드의 수에 따른 빅 데이터 처리 성능을 실험을 통해 예측한다. 본론에 앞서 2장에서는 본 연구에서 수행하는 맵리듀스와 실험에서 수행한 K-Means 알고리즘에 대해

### 2. 배경 지식

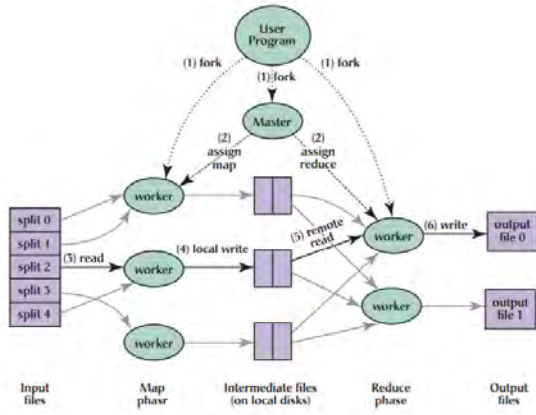
가. 맵리듀스 (MapReduce)

맵리듀스(MapReduce) [2]는 Google에서 정보검색을 위한 데이터 가공을 목적으로 개발된 분산 환경에서의 병렬 데이터 처리 기법이자 프로그램 모델이다. 페타바이트 이상의 대용량 데이터를 신뢰도가 낮은 컴퓨터로 구성된 클러스터 환경에서 병렬처리를 지원하기 위해서 개발되었다. 크게 Map과정과 Reduce과정으로 이루어져 있는데, Map 과정은 임의 키-값 쌍을 읽어서 이를 필터링 하거나 다른 값으로 변환하는 작업을 수행하고, Reduce과정은 Map을 통해 출력된 값들을 새 키 기준으로 그룹화 한 후에 여기에 집계 연산을 수행한 결과를 출력한다. 이는 빅 데이터를 분산 환경에서 처리하는데 널리 사용되고 있다. 본 논문에서는 각 슬레이브 노드에 작업을 분배할 때 맵리듀스를 통해 작업을 처리한다.

나. K-Means

K-Means[3] 알고리즘은 주어진 데이터를 k개의 클러스터로 묶는 알고리즘으로, 각 클러스터와 거리 차이의 분산을 최소화시키는 방식으로 동작한다.

이 논문은 정부(미래부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임.  
(NRF-2016R1610351)



(그림 1) 맵리듀스 과정

빅 데이터를 처리하는 기계 학습의 일종으로 EM 알고리즘을 이용한 클러스터링과 비슷한 구조를 가지고 있다. K는 클러스터의 개수를 의미하며 이 값은 사용자에게 의해 결정된다. 각 클러스터들의 평균값과 개별 데이터들의 거리를 최소화하도록 평균값을 계속해서 갱신한다. 그렇게 각 클러스터에 포함된 데이터들과 해당 클러스터의 평균값의 비용을 최소화시키는 것이 목표이다. 본 논문에서는 K-Means 알고리즘의 연산들을 슬레이브 노드들에게 분산시켜 작업을 수행한다.

### 3. 슬레이브 노드의 수에 따른 성능 예측하기

#### A. 실험환경

본 연구에서는 빅 데이터를 K-Means 알고리즘으로 수행하기 위해 한 대의 마스터 노드와 여러 대의 슬레이브 노드를 두었다. 스파크를 기반으로 맵리듀스를 수행하였으며 슬레이브 노드의 개수는 1대에서 7대까지 사용하였다. 실험 환경은 다음과 같다.

- Master node
- DELL optiplex 3020 Micro PC
- OS : Ubuntu 14.04
- 8GM RAM
- intel Core i3 CPU

- Slave node
- raspberry pi 1 B
- OS : Rassbian jessi
- 512MB RAM
- CPU : ARMv7

스파크 2.0.2 버전을 설치하여 K-Means 알고리즘의 연산을 슬레이브 노드들에게 분산시켰다.

#### B. 맵 리듀스를 통한 K-Means 알고리즘 연산

K-Means 알고리즘을 수행하는데 UCI 3D Road network

data [5]를 사용하였다. 클러스터는 2000개로 설정하였고 1000번 반복하여 클러스터링을 진행하였다. 슬레이브 노드를 1개에서 7개까지 늘려가면서 최적화된 클러스터의 평균값을 구하는데 걸리는 시간을 측정하였다.

#### C. 슬레이브 노드 수에 따른 성능 예측 모델

가장 이상적인 병렬 컴퓨팅 성능을 생각했을 때 슬레이브 노드가 1개일 때 특정 작업에 대한 소요시간이  $T$ 라고 하면 슬레이브 노드가  $N$ 개 일 때 그 소요시간은  $\frac{T}{N}$ 이다.

하지만 실제로 마스터 노드와 슬레이브 노드를 구성했을 때는 다를 수밖에 없다. 하지만 비슷한 동향일 것으로 판단되어 다음과 같은 모델을 써서 회귀 분석하였다.

$x$ 축이 슬레이브의 개수이고  $y$ 축이 작업을 처리하는데 걸린 시간이라고 해보자. 그리고 다음과 같은 꼴의 식에 실험 데이터를 점목시켜서 가장 오차가 작은 식을 도출해낸다.

$$t = \frac{(ax + c)}{(x - b)} \tag{1}$$

수식 (1)의 형태를 바꿔 보면

$$\begin{aligned} (x - b)y &= ax + c \\ xy - by &= ax + c \\ xy &= ax + by + c \end{aligned}$$

이다. 이때 실험 데이터와 이상적인 값의 오차는 다음과 같다.

$$F(a, b, c) = \sum_i (x_i y_i - ax_i - by_i - c)^2 \tag{2}$$

$F(a, b, c)$ 는 오차 값이다. 실험 데이터를 기반으로 오차 값이 가장 작은  $a, b, c$ 를 구하면 된다. 이렇게 만들어진 값을 바탕으로 해서 슬레이브 노드의 개수에 따른 성능을 예측할 수 있다.

### 4. 실험 결과

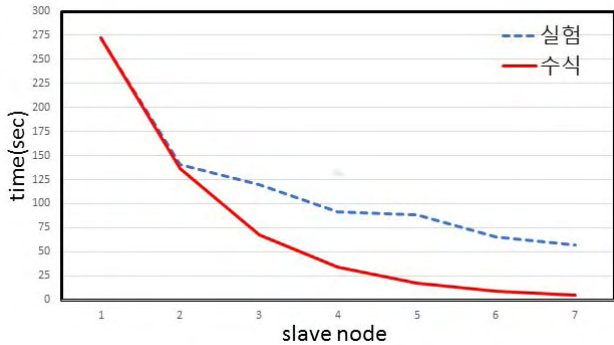
#### A. 슬레이브 노드에 따른 실제 작업 시간과 이상적인 수치 비교

슬레이브 노드인 라즈베리파이 보드를 마스터 노드와 같은 스위치에 연결하여 슬레이브 노드가 2개일 때부터 수행 시간을 측정하였다. (그림 3)에서 점선은 슬레이브 노드의 개수에 따른 작업 수행 시간을 나타낸 것이다. 그리고 실

선은  $t = \frac{T}{N}$ 을 따르는 이상적인 수행 시간이다. (그림 3)

에서 볼 수 있듯이 슬레이브 노드가 2개일 때는 이상적인 수치와 실제 실험 데이터와 비슷하지만 슬레이브 노드가 3개일 때는 오차가 있다. 이렇게 슬레이브 노드의 개수에 따른 수행 시간은 이상적인 수치와 거리가 멀다. 실제 환경에서는 노이즈도 많고 결함이 발생할 수도 있기 때문이다. 이를 통해 슬레이브 노드의 개수에 따른 실제 수행 시간은 단순히 예측하기 힘들다는 점을 볼 수 있다. 본 논문에서는 Nonlinear Regression을 통해 주어진 데이터를

모델링 하여 성능을 예측한다.



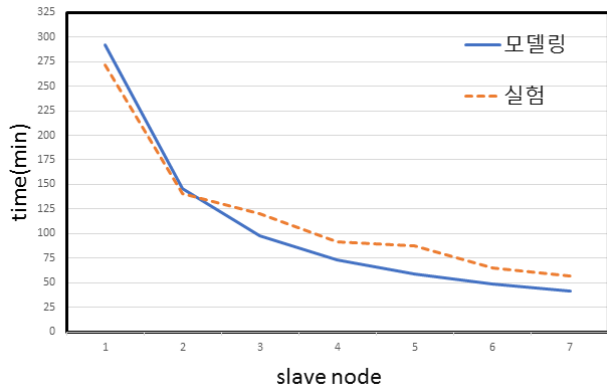
(그림 3) 슬레이브 노드에 따른 실제수행시간과 이상적인 수행시간 비교

B. 실험 데이터를 바탕으로 한 성능 예측 모델

수식 (2)에서  $F(a,b,c)$ 의 값이 가장 최소가 되는  $a, b, c$ 를 구하면 다음과 같다.

$$a = 0, b = 0, c = 292.417$$

이다. 이 모델과 실제 데이터 값을 비교하면 다음과 같다.



(그림 4) 슬레이브 노드의 개수에 따른 실제 수행시간과 모델링 결과 비교

(그림 4)에서 볼 수 있듯이 실제 데이터와 모델링 결과의 오차가 크지 않다. 슬레이브 노드가 7개 일 때뿐만 아니라 8개 이상 있을 때 도 작업의 수행 시간을 모델링 결과를 통해 예측할 수 있다.

## 5. 결론

본 논문은 스파크 분산 환경에서 슬레이브 노드에 따른 K-means 알고리즘의 빅 데이터 처리 시간을 분석하여 슬레이브 노드에 따른 성능을 모델링 하였다. 이를 통하여 요구하는 성능을 내기 위해 분산 환경을 어떻게 구성해야 하는지 예측할 수 있는데 의의가 있다. 향후 연구 과제는 서로 다른 연산 능력을 가진 슬레이브 노드들로 시스템을 구축 할 때의 성능을 예측하는 것이다. 슬레이브의 개수도 더 늘리고 어플리케이션의 특성까지 고려하여 더 정밀한 모델을 접목시킬 예정이다.

## 참고문헌

[1] Apache Spark Research 2014  
<https://spark.apache.org/research.html>.  
 [2] Dean, Jeffrey, and Sanjay Ghemawat. "MapReduce: simplified data processing on large clusters." Communications of the ACM 51.1 (2008): 107-113.  
 [3] Hartigan, John A., and Manchek A. Wong. "Algorithm AS 136: A k-means clustering algorithm." Journal of the Royal Statistical Society. Series C (Applied Statistics) 28.1 (1979): 100-108.  
 [4] Shvachko, Konstantin, et al. "The hadoop distributed file system." Mass storage systems and technologies (MSST), 2010 IEEE 26th symposium on. IEEE, (2010).  
 [5] 3D Road Network (North Jutland, Denmark) Data Set :  
[https://archive.ics.uci.edu/ml/datasets/3D+Road+Network+\(North+Jutland,+Denmark\)](https://archive.ics.uci.edu/ml/datasets/3D+Road+Network+(North+Jutland,+Denmark))