

Multi-Class Job 모델을 위한 Size-Interval 기반 할당 시스템 분석

문용혁*, 권혁찬*, 윤찬현**

*한국전자통신연구원

**한국과학기술원

e-mail : {yhmoon, hckwon}@etri.re.kr; chyoun@kaist.ac.kr

Analysis on Size-Interval Based Dispatching System for Multi-Class Job Model

Yong-Hyuk Moon*, Hyeokchan Kwon*, Chan-Hyun Youn**

*ETRI

**KAIST

요 약

본고에서는 Multi-class Jobs 을 Dispatching system 에서 처리하는 경우, Cost performance 을 점근적으로 해석하는 과정에 대해 논의한다. 구체적으로, Job 할당 시스템은 Size-Interval 기반의 스케줄링 기법을 이용하고, Resource failure 에 대비하여 Job duplication 전략을 활용하는 것으로 가정한다.

1. Multi-class Job 모델

기본적으로 Dispatching System (DS)은 Jobs (또는 Tasks)을 최적의 Resources 에 할당하는 역할을 수행하기 위해 고안되었다. 따라서, DS 는 서비스의 'Load balancing, Utilization, Fairness' 와 같은 성능을 보장시켜주는 Quality-of-Service (Qos) Broker 로서의 중요한 기능을 수행한다.

DS 가 운용되는 환경 및 해당 응용의 요구사항은 Job 의 형태로 구현되므로, Job 의 주된 요구사항, 예를 들어 최소 시간 또는 비용 등을 고려하여 Resource 할당을 전략을 실행해야 한다. 그러나, 동일 응용 환경 내에서도 Job 은 상이한 형태의 요구사항을 갖는 경우가 대부분이다 (Multi-class jobs). 더불어, 실제 분산 네트워크 환경에서 Resource 의 가용성은 Heterogeneous 한 특성을 갖고 있으며, 예상치 못한 Failures 로 인한 성능 손실이 발생함에 따라 요구된 QoS 을 만족 시키기 어려운 문제점을 안고 있다. 이를 해결하기 위해 대표적으로 Replication 전략을 활용한 Job Duplication 기법이 적용되고 있다. 그러나, Job Duplication 기법은 결합비용에 대한 강건함은 제공할 수 있으나, 비용대비 성능 (Cost Performance)에 취약한 단점을 갖고 있다. 따라서, 본 기법을 적용하기 전에 DS 시스템을 Long-term 관점에서 관측할 수 있는 모델의 필요성이 크게 요구되고 있는 실정이다.

그러므로, 본고에서는 'Multi-class jobs, Heterogeneous resource capacity, Resource failures'와 같이 요약될 수 있는 DS 운용환경에서 Job duplication 모델을 적용한 경우 QoS 만족도를 점근적인 관점에서 모델링하고 비용대비 성능을 평가하기 위한 방법에 대해 논의하

고자 한다. 이를 위해 먼저 1) Job 의 상이한 요구사항을 Job size 로 단편화하여 구분하고, 2) 이기종의 자원을 상이한 Capacity 를 갖는 복수개의 M/M/1 Queueing 모델로 표현하며, 3) 기존 Size-Interval (SI) Job 할당 알고리즘 [1, 2]을 Queueing 모델에 적용하여 DS 시스템의 Dispatching rule 을 형성하고 4) 이로부터 Cost performance factor 을 산출하는 과정을 설명한다.

2. Size-Interval 기반 할당 시스템

SI 전략은 Job size 를 N 개로 구분하고 (N-class jobs), 각 Size 에 해당하는 Job 을 처리하는 별도의 Queue ($R_1 \sim R_N$)에 Job 을 할당하는 구조를 갖는다. 그러나, SI 와 같은 정적 Job 할당 기법은 자원의 가용성 변동 사항 및 Job 의 요구사항 변화를 수용하기 보다는 최초의 요구사항 및 가용성을 고려하여 Job 을 할당하는데 초점을 맞추고 있는 Job 스케줄링 방법론이다. 따라서, 본 기법에 기초한 점근적 성능 평가를 위해서는 별도의 평가 모델의 고안이 요구된다.

3. SI 기반 Job Duplication 의 점근적 해석

본 절에서는 예상치 못한 결합으로 인해 발생할 수 있는 Job failures 에 대한, 대표적인 대처 기법으로 Job duplication 적용 시 상기 소개한 SI 할당 기법의 Dispatching 시스템의 예상 성능에 대해서 점근적으로 해석하는 방법에 대하여 논의한다.

먼저, 본 Job duplication 기법은 임의 Job 의 Size 를 처리할 수 있는 유사 Queue d 개를 각 자원 사이트로부터 식별하고, DS 은 각 Queue 로 동일 Job 을 할당한다. 이후 DS 는 각 Queue 자원을 모니터링하며, 결합

이 발견된 Job 은 중단시키고, 하나라도 성공한 Job 이 발생하면, 나머지 Jobs 을 중단시키는 방법으로 구현 될 수 있다. 따라서, 각 평균 Service rate (μ)은 Resource 의 결합확률에 따라 발생하는 시간지연을 각 Queue 에서 처리함에 따른 오버헤드 (ξ)를 고려하여 Effective service rate (μ') 형태로 다음과 같이 변형되어야 한다.

$$\mu' = \mu - \lambda \cdot d \cdot \xi, d > 1 \quad (1)$$

여기에서 λ 은 Arriving rate 을 의미한다. 그러므로, 각 Queue 에 할당될 평균 Load (ρ)는 다음과 같이 구해질 수 있으며, 본 연구에서는, 시스템이 항상 안정 상태 ($\rho < 1$)에 있다고 가정한다. 다음으로, 각 자원의 확률적 Queue size 을 계산하기 위해 먼저, i 개의 Queued job 을 가지고 있는 자원의 개수를 $s_i(t)$ 라 정의하고, 이로부터 최소 i 개 이상의 Queue size 인 자원이 존재할 확률 $p_i(t)$ 를 다음과 같이 산출한다.

$$p_i(t) = \sum_{k=i}^{\infty} \frac{s_k(t)}{s} \quad (2)$$

따라서, $p_i(t)$ 를 이용하면, 시간 t 일 때 예상되는 평균 Queue size 는 아래와 같이 표현될 수 있다.

$$\sum_{i=1}^{\infty} p_i(t) = \frac{\sum_{i=1}^{\infty} i \cdot s_i(t)}{s} \quad (3)$$

또한, p_i 는 [3]의 증명 과정에 따라 수식 (4)와 같이 Queue load (ρ)의 형태로 유도될 수 있다.

$$\frac{dp_i}{dt} = \rho(p_{i-1}^d - p_i^d) - (p_i - p_{i+1}) \quad (4)$$

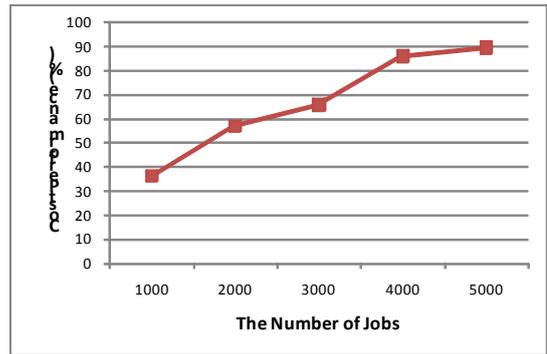
결과적으로, SI 기반 Job duplication 기법에서 평균 응답 시간은 Little 의 법칙 [4]에 따라 하단과 같이 구해진다.

$$E[T_d(\lambda)] = \frac{1}{\mu - \lambda \cdot d \cdot \xi} \sum_{i=1}^{\infty} \rho^{\frac{d-d}{d-1}} \quad (5)$$

그러므로, Makespan 은 $\text{Max} \{E[T_d(\lambda_j)], \text{for } j = 1, 2, 3, \dots, m\}$ 와 같이 획득할 수 있으며, 여기에서 m 은 주어진 Resource 의 개수를 의미한다. 다음 그림 1 은 Job duplication 의 정도를 (r^*) 2~3 으로 가정하고, Job 의 개수가 증가할 경우 이에 따른 비용대비 성능 변화추이를 50 개의 Heterogeneous Resources 를 대상으로 상기 논의한 점근적 분석 기법을 통해 시뮬레이션한 결과이다. 특히, Job duplication 기법과의 성능 비교를 위해 Resource 결합 시 이를 감내하고 동일 Resource 에서 Job 을 재할당시키는 Delayed Job 수행기법을 고려한다. 따라서, Cost-performance (CP) ratio 는 다음과 같이 표현될 수 있다.

$$CP\ Ratio = \frac{MS / MS_d}{R_d / R} \times 100(\%) \quad (6)$$

여기에서, R 및 MS 은 Delayed job 기법 수행 시 각 Job 에 할당된 Resource 개수 및 Makespan 을 R_d 및 MS_d 은 Duplicated job 기법 수행 시 각 Replicated job 에 할당된 Resource 개수 및 Makespan 을 의미한다. 다음 그림 1 과 같이 Replicated jobs 을 처리하기 위한 추가 Resources 사용에 따른 비용대비 성능은 Job 개수가 1,000 ~ 5,000 개로 증가하면서 36% ~ 90%으로 크게 개선되는 추이를 보이고 있다.



(그림 1) Cost Performance 에 대한 점근적 분석

4. 결론 및 추후연구

본 고에서는 Multi-class Job 을 SI 기반 할당 시스템으로 Heterogeneous Resources 에 적용하였고, Resource 결합 시 Job duplication 기법이 적용되었을 경우 이의 점근적 성능을 분석하기 위한 모델을 제안하였다.

Acknowledgement

“본 연구는 방송통신위원회의 Beyond Smart TV 기술 개발사업의 연구결과로 수행되었음 (2011년도 사업)”.

참고문헌

- [1] M.E. Crovella, M. Harchol-Balter, C.D. Murta, “Task Assignment in a Distributed System: Improving Performance by Unbalancing Load, in: ACM SIGMETRICS, 1998, pp. 268-269.
- [2] M. Harchol-Balter, M.E. Crovella, C.D. Murta, “On Choosing a Task Assignment Policy for a Distributed Server System”, J. Parallel Distribut. Comput. 59(1999) 204-228.
- [3] M. Mitzenmacher, “The Power of Two Choices in Randomized Load Balancing”, IEEE Trans. Parallel Distrib. Syst., Vol. 12, No. 10, pp.1094-1104, Oct. 2001.
- [4] Donald Gross, John F. Shortle, James M. Thompson, and Carl M. Harris, “Fundamentals of Queueing Theory”, 4th ed., published by John Wiley & Sons, Inc. (ISBN: 978-0-471-79127-0)