

분산 환경에서 동적 미전송 Push & Pull 모델 기반 효율적인 데이터 전송 기법

김남현*, 유헌창**

*고려대학교 컴퓨터정보통신대학원

**고려대학교 대학원 컴퓨터학과

e-mail : {lamyun10, yuhc}@korea.ac.kr

An Efficient Data Transmission Technique Based on Adaptive Non-Transmission Push & Pull Model in Distributed Computing

NamHyun Kim*, Heonchang Yu**

* Graduate School of Computer & Information Technology, Korea University

** Department of Computer Science and Engineering, Korea University

요 약

시스템 자원의 효율성에 대한 관심이 높아짐에 따라 성능데이터의 효과적인 전송에 대한 연구가 지속되고 있다. 본 논문에서는 기존 성능데이터의 Push & Pull 전송 방식에 변동률 기반 예측 알고리즘을 적용함으로써 성능데이터의 중복 수치 처리하여 효율적인 성능데이터를 전송하도록 하였다.

1. 서론

분산 환경으로 인해 이기종 자원의 복잡한 처리로 기존의 모니터링보다 복잡한 형태의 모니터링이 요구되고 있다. 이에 따라 성능데이터 전송은 자원의 특성과 중요도에 따라 Push 방식과 Pull 방식으로 전송되었다. 하지만 전송대상의 증가로 인해 기존 전송방식에 따른 성능부하에 대해 논의되었다. 따라서 본 논문은 동적 P&P(Adaptive Push & Pull)의 효율적인 전송을 위하여 중복 데이터의 전송을 제거한 동적 미전송 P&P(Adaptive Non-Transmission Push & Pull Model) 전송 모델을 제안한다.

2. 관련 연구

2.1 시계열 분석 알고리즘

성능데이터는 시간의 흐름에 따라 수치적으로 기록되는 성질이 있으므로 시계열자료이다. 따라서 시계열 분석을 통해 예측이 가능하다. 이를 반영하여 이전 연구에서 데이터의 변동에 대해 보정하였으나[1], 변동에 대한 평균차이를 고려하지 않았다.

2.2 Adaptive Push & Pull 모델

동적 P&P(Adaptive Push & Pull) 방식은 Push 방식과 Pull 방식을 동적 전환하여 전송하는 방식으로 웹 분야에 적용되어 증명되었다[2]. 하지만 웹 분야의 사례로써 성능데이터를 대상으로 동적 P&P 모델을 적용하는 것은 부적절하다. 또한 성능데이터 대상으로 예측 알고리즘을 반영한 동적 P&P 모델이 연구되었지만[3], 변동률에 대해 고려되지 않았으며 중복적 수치에 대한 처리에 대해 언급하지 않았다.

3. 동적 미전송 P&P 모델

3.1 변동률

시간에 따라 성능데이터가 변동하므로 평균에 대한 편차로 비교하는 것은 기준이 부적합하다. 더불어 편차는 평균값이 비교적 큰 값으로 커지는 경향이 있어 객관적인 변동률 측정이 어렵다. 따라서 평균값 차를 보완하기 위해 적용한 변동률은 <식 1>과 같다.

<식 1> 변동률 계산

$$V = \text{성능데이터 값}$$

$$\mu = \text{성능데이터 평균}$$

$$\sqrt{E((V - \mu)^2)} = \text{표준편차}$$

$$\text{변동률(CV)} = \frac{\sqrt{E((V - \mu)^2)}}{\mu} \times 100$$

3.2 가중 이동 평균(Weighted Moving Average)

성능데이터는 최근 변동률에 대해 영향을 받으므로, 최근변동률을 근거로 예측된 결과로 미전송을 수행하기 위해 가중 이동 평균법을 사용한다.

3.3 동적 미전송 P&P 전송방법

동적 미전송 P&P 에서 Push 과 Pull 방식의 전환은 임계값과 변동률에 따라 결정되며 (그림 1)의 세가지로 구분되어 수행된다. (a)는 예측변동률이 임계값과 정의한 근사치 구간에 포함된다면 일정함을 예측하고 데이터 전송을 수행하지 않는다. (b)는 변동률이 임계값보다 큰 경우 성능데이터의 변동이 크기 때문에

Push 방식으로 전송한다. 마지막으로 (c)는 변동률이 임계값보다 작거나 같은 경우로 변동의 폭이 작기 때문에 전송 횟수를 줄일 수 있는 Pull 방식으로 전환한다.

```

1  WHILE TRUE
2    Calculate 변동률
3    IF isPull //Pull 전환여부확인
4      Calculate 가중이동평균 계산
5      IF |임계 값-예측 값|<= 근사치
6        (a) 미전송
7      ELSE
8        IF 변동률> 임계 값
9          (b) Push 방식
10       ELSE // 변동률 <= 임계 값
11         (c) Pull 방식
12       END IF
13     END IF
14   ELSE //Push 전환여부확인
15     Push 방식
16   END IF
17 END WHILE
    
```

(그림 1) 동적 미전송 P&P 전송 알고리즘

4. 실험

4.1 실험방법

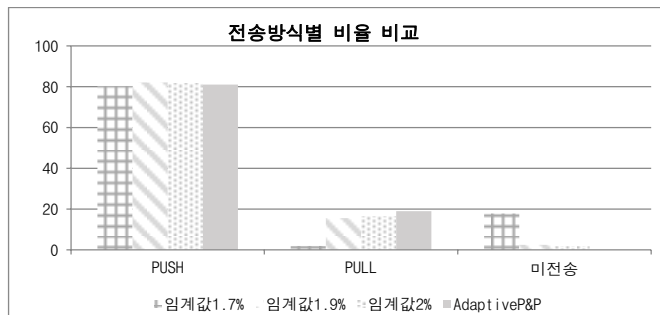
본 논문의 실험대상으로 일정한 수치를 유지하는 특성을 갖고 있는 메모리 성능 데이터를 실험대상으로 선정한다. 실험과정은 다음과 같다.

- 1) 동일 대상으로 동적 P&P 와 동적 미전송 P&P 방식으로 전송한다.
- 2) 동적 P&P 와 동적 미전송 P&P 방식의 전체 전송횟수를 비교한다.
- 3) 동적 P&P 와 동적 미전송 P&P 방식의 미전송 횟수 및 정확도를 비교한다.

4.2 실험 결과 및 평가

<표 2> 전송횟수 결과값

전송방식	동적 P&P	임계값 1.7%	임계값 1.9%	임계값 2%
Push 방식	1167	1157	1182	1177
Pull 방식	273	27	224	237
계	1440	1184	1406	1414



(그림 2) 동적 미전송 P&P 전송방식 별 횟수 비교

<표 1>의 실험 결과와 같이 임계값이 작을수록 전체 전송 횟수가 작아지며, 반대로 임계값이 커질수록

전송횟수가 커진다는 것을 확인할 수 있다. 또한 (그림 2)에서 동적 P&P 와 Push 방식의 비율은 큰 차이가 없지만, Pull 방식의 경우 임계값이 감소할수록 비율이 감소하는 반면, 미전송의 비율이 높아짐을 확인할 수 있다. 이는 임계값이 작을수록 변동률을 크게 반영하여 미전송 횟수가 증가하고, 임계값이 클수록 변동률을 작게 반영하여 미전송 횟수가 줄어드는 결과를 보여준다.

<표 2> 미전송 횟수 비교

항목	동적 P&P	임계값 1.9%	임계값 2%
Pull 전송횟수	273	224	237
(a) 미전송 횟수	0	35	26
(b) 미전송 대상 횟수	0	33	24
(c) 비율	0.00	12.8	9.5

<표 2>는 전체 Pull 방식 대비 미전송 결과값이다. (a)는 전체 미전송 횟수로 임계값 1.9%에서는 35 개, 임계값 2%은 26 개이다. (b)는 변동률이 낮은 미전송 가능 횟수로 임계값 1.9%일 경우 33 개가 미전송 되었으며, 임계값 2%의 경우 24 개의 데이터가 미전송 되었다. 또한 (c)에서 각 Pull 전송비율은 12%, 8.79%이며, 이는 중복적인 데이터의 전송을 감소되었음을 보여 준다.

5. 결론 및 향후 과제

본 논문에서 효율적인 성능데이터 전송을 위해 동적 미전송 P&P 전송 방식을 제안하였다. 데이터의 변동률을 이용하여 Push 전송과 Pull 전송 방식 전환을 조정하였으며, 예측 알고리즘을 통해서 중복 데이터 전송에 대한 미전송 처리를 수행하였다. 동적 P&P 방식에 비해 데이터의 변동률에 따른 전체적인 전송횟수와 중복 데이터의 전송 작업을 줄일 수 있었다. 향후, 연속적인 Push 전송의 중복 데이터 처리와 패턴 분석을 통해 Push 와 Pull 간의 효과적인 전환 방법을 연구할 계획이다.

참고문헌

- [1] Z. Xiao, W. Song, Q. Chen, "Dynamic resource allocation using virtual machines for cloud computing environment", IEEE Transactions on Parallel and Distributed Systems, Vol.24, No.6, pp.1107-1117, June. 2013.
- [2] Manish Bhide, Pavan Deolasee, Amol Katkar, Ankur Panchbudhe, Krithi Ramamritham, and Prashant Shenoy, " Adaptive Push-Pull Disseminating Dynamic Web Data", IEEE Transactions on Computers, Vol. 51, No. 6, June 2002.
- [3] Ying Jiang, Heng Sun, Jiaman Ding and Yingli Liu , "A Data Transmission Method for Resource Monitoring under Cloud Computing Environment", International Journal of Grid Distribution Computing, Vol.8, No.2, pp.15-24, 2015.