

박필성
수원대학교

AM-16

A synchronous/asynchronous hybrid parallel method for some eigenvalue problems on distributed systems

오늘날 단일 슈퍼컴퓨터로는 처리가 불가능한 거대한 문제들의 해법이 시도되고 있는데, 이들은 지리적으로 분산된 슈퍼컴퓨터, 데이터베이스, 과학 장비 및 디스플레이 장치 등을 초고속 통신망으로 연결한 GRID 환경에서 효과적으로 실행시킬 수 있다. GRID는 1990년대 중반 과학 및 공학용 분산 컴퓨팅의 연구 과정에서 등장한 것으로, 점차 응용분야가 넓어지고 있다. 그러나 GRID 같은 분산 환경은 기존의 단일 병렬 시스템과는 많은 점에서 다르며 이전의 기술들을 그대로 적용하기에는 무리가 있다. 기존 병렬 시스템에서는 주로 동기 알고리즘(synchronous algorithm)이 사용되는데, 직렬 연산과 같은 결과를 얻기 위해 동기화(synchronization)가 필요하며, 부하 균형이 필수적이다. 그러나 부하 균형은 이질 클러스터(heterogeneous cluster) 처럼 프로세서들의 성능이 서로 다르거나, 지리적으로 분산된 계산자원을 사용하는 GRID 환경에서는 이기종의 문제뿐 아니라 네트워크를 통한 메시지의 전송 지연 등으로 유희시간이 길어질 수밖에 없다. 이처럼 동기화의 필요성에 의한 연산의 지연을 해결하는 하나의 방안으로 비동기 반복법(asynchronous iteration)이 나왔으며, 지금도 활발히 연구되고 있다. 이는 알고리즘의 동기점을 가능한 한 제거함으로써 빠른 프로세서의 유희 시간을 줄이는 것이 목적이다. 즉 비동기 알고리즘에서는, 각 프로세서는 다른 프로세서로부터 갱신된 데이터가 올 때까지 기다리지 않고 계속 다음 작업을 수행해 나간다. 따라서 동시에 갱신된 데이터를 교환한 후 다음 단계로 진행하는 동기 알고리즘에 비해, 미처 갱신되지 않은 데이터를 사용하는 경우가 많으므로 전체적으로는 연산량 대비의 수렴 속도는 느릴 수 있다. 그러나 각 프로세서는 거의 유희 시간이 없이 연산을 수행하므로 wall clock time은 동기 알고리즘보다 적게 걸리며, 때로는 50%까지 빠른 결과도 보고되고 있다. 그러나 현재까지의 연구는 모두 어떤 수렴조건을 만족하는 선형 시스템의 해법에 국한되어 있으며 비교적 구현하기 쉬운 공유 메모리 시스템에서의 연구만 보고되어 있다. 본 연구에서는 행렬의 주요 고유쌍을 구하는 데 있어 비동기 반복법의 적용 가능성을 타진하기 위해 우선 이론적으로 단순한 멱승법을 사용하여 실험하였고 그 결과 순수한 비동기 반복법은 수렴하기 어렵다는 결론을 얻었다. 그리하여 동기 알고리즘에 비동기적 요소를 추가한 혼합 병렬 알고리즘을 제안하고, MPI(Message Passing Interface)를 사용하여 수원대학교의 Hydra cluster에서 구현하였다. 그 결과 특정 노드의 성능이 다른 것에 비해 현저하게 떨어질 때 전체적인 알고리즘의 수렴 속도가 떨어지는 것을 상당히 완화할 수 있음이 밝혀졌다.