

CS492 Distributed Systems & Algorithms: Mr.CL

20030767 최종욱, 20050145 김준기, 20060080 김민국

2009년 12월 24일

1 동기

GPU를 일반 과학 연산에 활용하기 시작하면서 많은 성능 향상을 얻고 있지만 scalability의 한계점이 있었음

분산처리특강을 통해 Hadoop을 접하면서 map-reduce 방식이 그런 점을 극복하는 데 도움이 되리라 생각하고 GPU 가속을 결합시켜보고자 함

2 알고리즘 설계

2.1 Hama 프로젝트

TODO: Hama 프로젝트를 선택한 이유, 하지만 도입에 실패한 이유 설명 (HBase의 성능 제약 등)

TODO: 병목 지점이 Data I/O에 있음을 설명하고 다른 방식의 접근이 필요함을 설득

2.2 Outer Product Algorithm

TODO: outer product algorithm 설명

3 구현

TODO: 최종적으로 Hadoop 위에 바로 코딩해서 올린 구조 설명

4 성능 측정

TODO: 성능 측정 계획 설명

5 결론

TODO: 부딪힌 문제점들 설명 (Hama와 HBase의 설치 중 삽질한 것, CUDA 드라이버 관련 삽질한 것, JCublas에서 double/float 삽질한 것 등)

TODO: 앞으로 개선할 점과 계속해서 연구한다면 어떤 것들을 해보고 싶다 등등