

# 네트워크 기반 컴퓨팅에서 신인도에 따른 서비스 수준 보상 기법

이진성\* 최창열\* 박기진\*\*

아주대학교 정보통신전문대학원\* 아주대학교 공과대학 산업정보시스템공학부\*\*  
(cslee77,clchoi)@ajou.ac.kr\* kiejin@ajou.ac.kr\*\*

## Service Level Compensation Mechanism by Credibility in Network-based Computing

Jinsung Yi\* Changyeol Choi\* Kiejin Park\*\*

Graduate School of Information and Communication, Ajou University\*  
Division of Industrial & Information Engineering, Ajou University\*\*

### 요 약

고속 네트워크의 등장으로 관리 영역을 초월한 계산 자원의 공유가 가능하게 되었다. 유휴 상태에 있는 자원들의 자발적인 참여로 인한 자원 제공은 네트워크로 상호연결된 컴퓨터 시스템 즉, 그리드 컴퓨팅을 가능하게 하였다. 그리드 환경에 포함된 각 자원들은 이질적이고, 이질적인 자발적 참여 자원들에게는 자원의 참여 결과로서 나타나는 신인도라는 차원에서 적절한 보상이 제공되어야 한다. 본 논문에서는 그리드 컴퓨팅에서 각 노드들이 자발적인 협력체계에 참여하여 얻게되는 결과인 신인도에 따라 보상을 받을 수 있는 기술을 연구하고 그 타당성을 검토한다. 각 노드들이 작업을 하기전 시스템 정보인 정적정보와 주어진 작업을 마친 후 반환 값의 정확도와 응답시간과 같은 동적정보를 고려하여 얻게된 신인도를 고려하여 발런티어에게 피드백을 제공하는 보상체계를 구축한다. 신인도에 따른 보상체제는 유휴 노드들의 자발적 참여율을 높이고 전반적인 시스템 성능을 향상시킨다.

## 1. 서론

그리드 컴퓨팅은 지리적으로 분산된 컴퓨터 시스템, 대용량 저장 장치 및 데이터베이스, 첨단 실험 장비 등의 자원들을 고속 네트워크에 연결해 상호 공유 이용할 수 있도록 하는 디지털 신경망 구조의 차세대 인터넷 서비스이다. 예를 들어 대량의 계산 및 고성능을 필요로 하는 프로그램들을 실행하기 위해서는 지리적으로 분산되어 있는 여러 대의 슈퍼 컴퓨터를 동시에 사용할 필요성이 있다. 그리드는 한번에 한 자원에만 연결할 수 있는 웹과는 달리 여러 자원과 연결하여 작동하는 네트워크 구조로 이루어져 있으며 고성능의 유휴 컴퓨팅 자원을 서로 공유하여 효율적으로 작업을 수행하는 것을 목적으로 한다[1].

이질적인 그리드 환경에서는 동종의 구성 요소로 이루어졌던 기존의 시스템과 달리 네트워크로 연결된 환경에서 자율적 참여 노드의 자원 발견 및 자원 선택 문제가 중요시 되어 자원 관리에 관한 다양한 연구가 이루어졌다[2]. 하지만 자발적 참여로 유휴자원을 제공한 발런티어에 대한 효율적 관리 기법에 관한 연구나 보상체계에 대한 연구는 실질적으로 이루어지지 않았다. 이는 그리드 컴퓨팅이 거대 조직으로 성장할 수 있는 잠재력에 대한 장애물로 작용하고 있다. 따라서 그리드 컴퓨팅 환경에서는 유휴 자원을 사용해 계산 문제에 참여한 발런티어들에게 적절한 보상을 해주으로써

계속 발런티어가 계산 문제를 해결할 수 있도록 유지하는 것이 중요하다. 다시말해서 그리드 컴퓨팅에서의 이러한 보상 기술에 관한 연구가 미미한 실정으므로 본 연구에서는 데이터베이스 기술을 이용해 웹 서비스 형태로 계산에 참여한 수많은 발런티어를 보상하고 관리하는 메커니즘을 제안한다.

## 2. 관련 연구

그리드에서 발런티어 노드들에 대한 관리 기법이나 악의적 발런티어를 차단할 수 있는 기술에 관한 많은 연구가 있었으며 본 연구에서는 기존의 신인도(Credibility) 모델을 바탕으로 발런티어 노드들에 대해 순위를 부여하고 부여된 순위에 따라 적절한 보상체계를 제공하기 위한 연구들을 소개한다. 발런티어에 대한 보상체제는 신인도에 따라 서비스 계약 수준(SLA: Service Level Agreement) 기법과 차별화된 서비스 기법(Diffserv)를 적용한다.

### 2.1 신인도에 따른 자원 순위 기법

그리드 컴퓨터 환경에서 각 발런티어들은 시스템 하드웨어 장비의 정적 정보(Static Information)와 작업을 처리한 결과로부터 얻게되는 응답시간율(Response Time Rate), 에러율(Error Rate), 정확도(Correctness)와 같은 동적 정보로 구성된 신인도를 가지며, 이를 이용하여 관리자는 자원 분배시 정책값으로 사용하게 된다. 이러한 동적 정보는

본 연구는 한국학술진흥재단 신진교수연구과제(KRF-2003-003-D00331; 지원으로 수행되었음.

작업이 수행될 때마다 결과에 따라 바뀐 신인도는 다시 자원 분배 정보로 사용된다.

### 2.2 서비스 계약 수준(Service Level Agreement)

SLA란 서비스를 제공하는 자가 서비스 가입자에게 사전에 정의된 수준의 서비스를 제공하기로 협약을 맺는 것을 의미한다. 대개 어떤 서비스가 제공될 것인지를 측정 가능한 조건으로 명시한 것으로, 많은 수의 인터넷 서비스 공급 회사들이 SLA와 같은 형태의 계약을 참여 노드에게 제공하며 본 연구에서는 발런티어의 신인도에 따라 서비스를 제공한다. 실제 서비스는 웹 콘텐츠, 네트워크 관리 시스템을 통한 시스템 문제점 진단 및 해결 서비스와 같은 방법으로 다양하게 제공될 수 있으며, SLA에 반드시 포함되어야 할 척도로 서비스 될 수 있는 시간 비율, 동시에 서비스할 수 있는 사용자 수, 실제 성능을 주기적으로 비교할 수 있는 명확한 기준, 사용자들이 영향을 받을 수 있는 네트워크의 변경 작업 등을 사전에 고지하기 위한 일정, 다양한 종류의 문제에 대한 고객 상담실의 응답시간, 제공될 사용량 통계, 다이얼업 접속 가능성 등을 이용할 수 있다[3][4][5].

### 3. 시스템 설계

본 연구 모델은 기존의 마스터(Master)와 참여자(Worker)로 구성된 그리드 컴퓨팅 구조를 이용하며 그림 1과 같다[6]. 각 발런티어들은 정적 정보와 동적 정보의 결합으로 구성된 신인도에 따라 그룹화가 되어 있으며 함께 순위가 부여되어 있다.

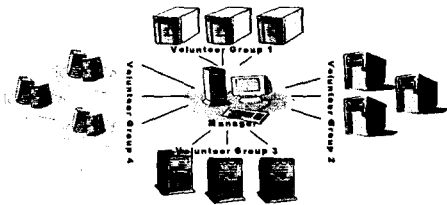


그림 1. 신인도에 따라 그룹화 되어 있는 시스템 구조

### 4. 신인도에 따른 서비스 분할 제공 기법

제안한 방법은 각 발런티어들이 신인도를 가지고 있으며, 자원 사용이 요구될 때 매니저를 통하여 신인도를 통하여 각종 콘텐츠 서비스나 혹은 유휴 자원을 제공받을 수 있다. 매니저는 각 발런티어들의 신인도를 관리하고 신인도에 따라 작업을 부여하는 역할과 함께 휴지상태의 자원들을 동적인 방법으로 분할하여 발런티어들의 신인도에 따라 서비스를 제공하게 된다.

#### 4.1 그룹화 및 차별화 서비스 제공을 위한 매커니즘

시스템은 크게 태스크 매니저(Task Manager), 서비스 매니저(Diffserv Manager), 데이터베이스(Database), 리소스

매니저(Resource Manager), 자원 풀(Resource Pool)로 구성되어 있으며 구조도는 그림 2와 같다.

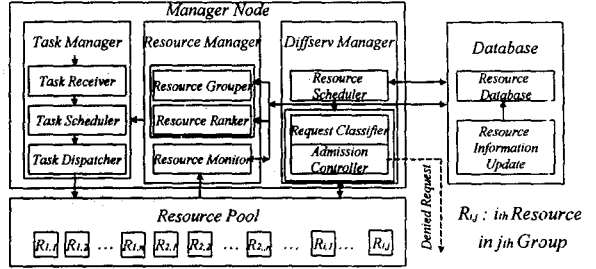


그림 2. 서비스 수준 보상을 위한 마스터 노드 구조도

태스크 매니저는 수행될 태스크를 받는 부분(Receiver), 신인도에 따라 태스크를 매칭시키는 부분(Scheduler), 실제로 태스크를 할당하는 부분(Dispatcher)으로 구성되어 있다. 자원관리기(Resource Manager)는 자원을 그룹핑하는 자원 그룹퍼(Resource Group)와 순위를 나누는 자원순위기(Resource Ranker), 자원의 상태변화를 관리하는 자원 모니터부(Resource Monitor)부로 구성되며 발런티어들의 보상체계를 위한 서비스 관리기(Diffserv Manager)는 신인도에 따라 서비스 범위를 한정하는 자원매치(Resource Matcher)와 발런티어의 서비스 요구를 받아들이는 서비스 모니터(Service Monitor)부와 서비스를 제공하는 서비스 제공부(Service Provider)로 구성되어 있다.

시스템은 스위치와 후위서버(Back-end Server)로 구성되어 있으며, 스위치는 사용자의 요구를 실제로 서비스를 처리하는 가용 자원을 가진 발런티어로 보낸다. 본 시스템에서 사용자 요구의 증가에 따른 시스템 과부하 상황에서도 신인도에 따라 QoS를 고려한 차별화된 서비스를 제공하기 위하여 웹 스위치에 QoS 조절기를 사용한다. 요구 분류기(Request Classifier)는 사용자별 요구를 분류하여 해당 큐로 보내는 역할을 한다. 승인제어기(Admission Control)는 서버노드의 과부하가 발생시 우선 순위에 따라 사용자 요구를 거부하게 되며 서비스 스케줄러(Service Scheduler)는 분류된 요구를 스케줄링 정책에 의해 해당 발런티어로 분배해주는 역할을 한다. 자원관리기(Diffserv Manager)는 분할 노드 정책에 따라 서버 노드의 자원을 관리하는 부분으로 유휴 노드가 많이 존재할 경우 발런티어의 신인도에 따라 알맞은 노드를 할당함으로써 효율적인 차별화 서비스가 가능하고, 고신인도를 가진 발런티어에 더 많은 서버 및 서비스를 할당함으로써 SLA를 만족시키도록 한다.

#### 4.2 신인도와 차별화 서비스

본 논문에서 신인도는 시스템의 하드웨어 정적 정보와 작업 수행의 결과로 형성된 동적 정보로 구성이 된다. 신인도에 따라 서비스를 제공하기 위해서는 유휴자원과 발런티어의 요구상황을 모니터링하는 것이 필요하며, 발런티어의 서비스 요구 순위는 실제로 처리해야하는 작업의 우선순위보다 낮다고 가정한다. 자원의 서비스는 서비스를 요구하는

시점에서 휴지상태에 있는 자원의 양과 서비스를 요구하는 발런티어의 요구 수 및 각 발런티어의 신인도에 따라 발런티어에게 제공되는 서비스의 질(QoS)을 결정할 수 있다.

본 논문에서는 차별화된 서비스를 위한 기법으로 두 가지를 고려한 서버 노드의 동적 분할 기법을 제안하며 그림 3 과 같은 스케줄링 기법이 있다. 첫 번째로 높은 신인도를 가진 상위 계층의 발런티어에게 하위 계층의 서비스보다 더 좋은 서비스를 제공하도록 한다. 두 번째로 서버 노드의 과부하가 발생할경우 사용자 요구의 거절을 통하여 SLA 를 보장할 수 있는 서비스를 제공하도록 한다.

- 선입선출(First In First Out) 기법 : 자원을 요구하는 발런티어의 신인도에 관계없이 대기큐에 먼저 들어온 작업에게 자원을 먼저 할당한다.



그림 3-1. 선입선출 스케줄링 기법

- 라운드로빈(Round Robin) 기법 : 선입선출과 유사하지만 주어진 시간 할당량 안에 작업을 마쳐야 하며, 마치지 못할 경우 다른 발런티어에게 자원을 양보하게 된다.



그림 3-2. 라운드로빈 스케줄링 기법

- HRN(Highest Response-Ratio Next) 기법 : 우선 순위에 의한 방법으로 우선순위는 신인도에 따라 결정된다.

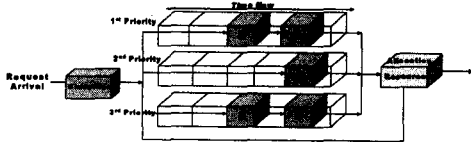


그림 3-3. 신인도에 따른 우선순위별 스케줄링 기법

### 5. 성능 평가

시스템 성능은 현재 유휴상태에 있는 노드와 신인도를 이용하여 선입선출 기법, 라운드 로빈 기법, 우선순위를 신인도로 결정하여 HRN 기법에 의하여 세가지 방법을 이용하여 스케줄링 기법과 버려진 패킷의 수(Dropped Packet), 사용자 응답시간(Response Time)과의 관계를 살펴보았다.

본 실험에서 각 서비스를 요청하는 발런티어의 신인도는 0 에서 1 사이의 임의의 값을 갖고 있으며, 가용 서버 자원은 100 에서 1000 사이의 임의의 수를 가지며, 서비스 레벨수를 변화시키면서 각 스케줄링 기법에 따른 지연시간(Latency

Time)의 관계를 살펴보면 그림 4 와 같다. HRN 기법을 이용하여 신인도에 따라 서비스를 할 때 신인도가 높은 발런티어에 대해서 지연시간이 짧아지므로 제대로 보상이 이루어짐을 확인할 수 있다.

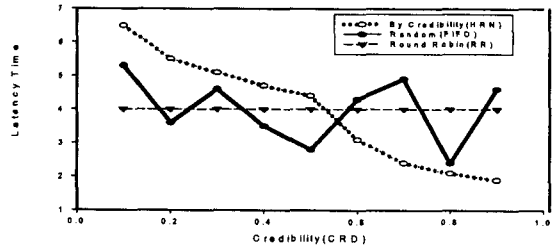


그림 4. 신인도와 스케줄링 기법에 따른 지연시간 그래프

### 6. 결론 및 향후 방향

본 논문에서는 이질적인 자원을 가진 그리드 환경에서 각 발런티어들의 정적정보와 동적 정보로 구성된 신인도에 따라 서비스를 적절히 보상하는 기법을 제공한다. 자발적 참여 노드들은 아무런 혜택없이 자신의 자원을 공개하려고 하지 않으며, 시스템의 보안에 관한 내용을 신뢰한다는 것은 어려운 일이므로 적절한 보상체계를 갖추어야 하며, 신뢰성을 제공하여야 한다. 본 연구에서는 이러한 보상체계의 일원으로서 신인도에 따라 우선순위를 부여하고 유휴노드, 즉 가용한 자원에 대해서 서비스를 제공하는 기법을 제안하였다. 이는 유휴 노드들의 자발적인 참여를 더욱 활성화시킬 것으로 기대된다.

### 7. 참고문헌

- [1] I. Foster, and K. Kesselman, "Globus: A Metacomputing Infrastructure Toolkit," The International Journal of Supercomputer Applications and High Performance Computing, Vol. 11, No. 2, pp. 115-128, 1997.
- [2] Chuang Liu, Lingyun Yang, Ian Foster, and Dave Angulo, "Design and Evaluation of a Resource Selection Framework for Grid Applications," Proceedings of the 11th IEEE Symposium on High-Performance Distributed Computing, pp. 63-72, July 2002.
- [3] D. Andresen, T. Yang, V. Holmedahl, and O. H. Ibarra, "SWEB: Towards a Scalable WWW Server on MultiComputers," Proceedings of the IEEE International Symposium on Parallel Processing, pp. 850-856, Apr. 1996.
- [4] C. Aurrecochea, A. T. Campbell, and L. Hauw, "A Survey of QoS Architectures," ACM/Springer-Verlag Multimedia Systems Journal, Special Issue on QoS Architecture, Vol. 6(3), pp. 138-151, May 1998.
- [5] N. Bhatti and R. Friedrich, "Web Server Support for Tiered Services," IEEE Network, Vol. 13(5), pp. 64-71, Sep. 1999.
- [6] Jean-Pierre Goux, Sanjeev Kulkarni, Michael Yoder, and Jeff Linderth, "An Enabling Framework for Master-Worker Applications on the Computational Grid," 9th IEEE International Symposium on High Performance Distributed Computing (HPDC'00), Aug. 2000.