

# 키워드 기반 전문 검색을 이용한 Access Grid에서의 효과적인 데이터 공유 방안

오현, 김진승, 이영구  
경희대학교 컴퓨터공학과  
e-mail:kevin921@empal.com

## Effective Method for Data Sharing on Access Grid Using Keyword-Based Fulltext Retrieval

Heon Oh, Jin-Seung Kim, Young-Koo Lee  
Dept of Computer Engineering, Kyung-Hee University

### 요 약

앞으로 다가올 유비쿼터스 컴퓨팅 환경에서는 일상생활 어디에서든지 컴퓨팅이 가능해질 것으로 예상되며, 원격지의 상대방과의 상호작용이 다양한 형태로 발전할 것으로 예상된다. 이를 지원하기 위하여 유비쿼터스 환경에서의 다자간 인터랙티브 협업 시스템들이 연구되고 있다. 그런데, 현재 개발된 다자간 협업 환경 시스템에서는 자료를 공유하는데 있어서, 데이터에 대한 인덱싱과 검색에 비중을 두지 않고 있다. 공유 데이터는 늘어나는 반면, 그것을 효과적으로 검색할 수 있는 방법이 마련되지 않아 데이터를 공유하고 활용하는 효율성이 떨어질 것이 예상됨에 따라서 키워드 검색 기반으로 데이터를 검색할 수 있는 방법을 마련하고자 한다.

### 1. 서론

유비쿼터스 컴퓨팅 환경에서는 현재의 컴퓨팅 인프라가 보다 확장되어, 다양한 종류의 컴퓨터가 사람, 사물, 환경 속에 내재되어 있고, 이들이 서로 연결되어, 필요한 곳 어디에서든 컴퓨팅이 가능해질 것으로 예상된다. 또한, 원격지의 상대방과의 상호작용이 현재보다 훨씬 유기적이고 실제적이며 다양한 형태로 발전할 것으로 예상된다. 이러한 흐름을 반영할 때 앞으로 진보된 컴퓨팅 시스템과 네트워크를 바탕으로 한 협업 시스템들이 주목을 받을 것으로 예상됨에 따라, 커뮤니티 컴퓨팅 환경에서의 인터랙티브 협업 시스템의 개발이 필요해졌다.

\* 본 연구는 21세기 프론티어 연구개발사업의 일환으로 추진되고 있는 정보통신부의 유비쿼터스컴퓨팅네트워크원천기반기술개발사업의 지원에 의한 것이다.

현재 미국 아르곤 국립연구소(Argonne National Laboratory: ANL)에서는 액세스 그리드(Access Grid)[1] 라는 다자간 협업환경 기술을 개발하고 있으며, 이 외에도 Stanford University의 iRoom 프로젝트[2], Washington University의 Portolano 프로젝트[3], University of Illinois at Chicago의 SAGE 프로젝트[4] 등이 진행 중에 있다.

인터랙티브 협업 시스템에서 자료의 공유는 중요한 기능 중의 하나이다. 그런데, 대부분의 협업 시스템에서 데이터 공유 기능을 제공하되 파일 이름으로만 검색이 가능하도록 되어 있다. 공유하는 데이터는 늘어나는 반면, 그것을 효과적으로 검색할 수 있는 방법이 마련되지 않아 데이터를 공유하고 활용하는 효율성이 떨어진다. 따라서 본 논문에서는 액세스 그리드에서의 키워드 검색 기반으로 데이터를 검

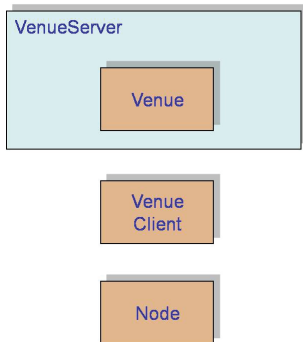
색할 수 있는 방법에 대해 제안한다.

본 논문의 구성은 다음과 같다. 2장에서는 액세스 그리드를 소개하고, 구조를 살펴본다. 3장에서는 공유데이터에 대해 키워드 기반 검색 방법을 적용하는 방법을 제안하며, 4장에서는 결론 및 향후 연구 방향을 제안한다.

## 2. 관련연구

액세스 그리드는 다자간 협업 환경 기술로서 현재 천여 개의 AG 노드에서 e-사이언스, 의료, 교육 등을 위한 원격 회의 및 강의에 활용되고 있다. 액세스 그리드는 멀티캐스트 네트워크에 기반을 둔 고품질 영상, 음성 스트리밍을 이용하여 참여자들 간의 공동작업 내용과 협업 공간 분위기를 효과적으로 전달한다. 또한 메뉴 서버(Venue Server)는 메뉴 클라이언트(Venue Client)들 간의 가상회의 공간을 안내하여 데이터/서비스/응용프로그램의 공유 기능, 안전환(secure) 그룹 통신, 이벤트 정보 전달과 같은 필수적인 협업 요소 기술들을 제공한다. 하지만 액세스 그리드는 원거리 협업환경의 네트워크 지원 측면은 강조된 반면, 스마트 미팅공간에서 요구하는 context-aware 기술, 직관적인 사용자 인터페이스를 이용한 task migration 기술 등은 고려하지 않고 있다.

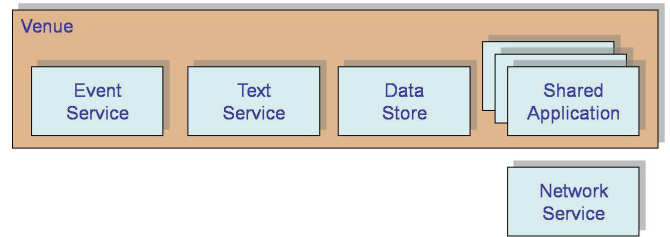
액세스 그리드의 구조는 (그림 1)에서 볼 수 있듯이 크게 메뉴 서버와 메뉴 클라이언트, 노드(Node)로 구성되어 있다.



(그림 2) 액세스 그리드의 구조

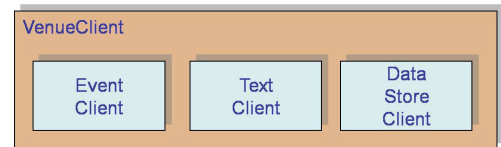
메뉴 서버에서는 메뉴라는 방을 정의하여서 메뉴 클라이언트가 메뉴에 입장하여 회의를 하고 자료를 공유하도록 한다. 메뉴는 (그림 2)에서 볼 수 있듯이 메뉴 클라이언트와 이벤트를 주고받는 이벤트 서비스(Event Service), 메뉴 클라이언트와 텍스트를 주고 받는 텍스트 서비스(Text Service), 공유 데이터를 저장하는 데이터 스토어(Data Store), 파워포인트

트, 웹 탐색기, 텍스트 등의 어플리케이션을 공유하는 어플리케이션 공유(Shared Application) 그리고 노드와의 서비스를 위한 네트워크 서비스(Network Service)로 구성되어 있다.



(그림 3) Access Grid 메뉴의 구조

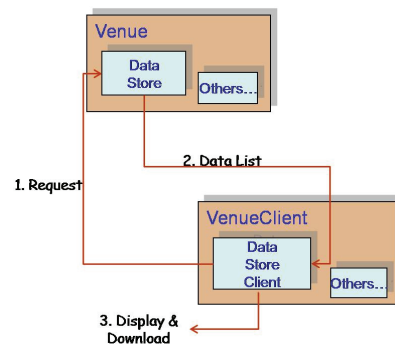
메뉴 클라이언트는 (그림 3)에서 볼 수 있듯이 이벤트 서비스(Event Service), 텍스트 서비스(Text Service), 데이터 스토어 클라이언트(Data Store Client)로 구성되어 있다.



(그림 4) Access Grid 메뉴Client의 구조

## 3. 키워드 기반 검색의 적용

기존의 액세스 그리드에서는 (그림 4)에서와 같이 메뉴 클라이언트의 데이터 스토어 클라이언트 모듈에서 메뉴의 데이터 스토어 모듈에게 데이터의 리스트를 요청하고 받아온 뒤에 데이터의 파일의 이름으로만 검색이 가능 하도록 되어 있다.



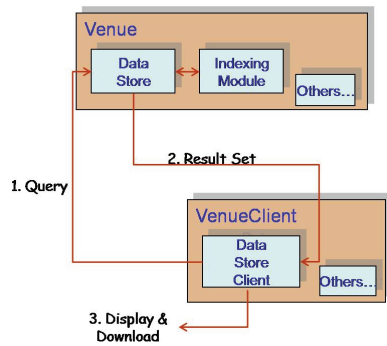
(그림 5) 기존의 검색 과정

파일의 이름으로만 검색을 하면, 원하는 데이터를 찾기가 힘들 뿐만 아니라, 데이터의 내용을 보기 위해서 해당 파일을 로컬시스템에 저장 후 확인해야 하는 불편함이 있다. 이를 해결하기 위해서 본 논문에서는 최근에 정보검색 기술로 널리 사용되고 있는 파일의 전문(fulltext)을 대상으로 한 키워드 기반 검색

색 기법을 제안한다.

(그림 5)는 키워드 기반 검색을 제공하는 데이터 스토어의 구조를 나타낸다. 액세스 그리드의 데이터 공유에 키워드 기반의 검색을 가능하게 하기 위해서 베뉴 서버의 각 베뉴에 인덱싱 모듈(indexing module)을 추가한다.

데이터 스토어에 들어오는 모든 데이터들은 인덱싱 모듈에서 전문을 대상으로 인덱싱을 한다. 인덱싱 방법으로는 Microsoft의 MSN Desktop Search[5]를 사용한다. 베뉴 클라이언트의 데이터 스토어 클라이언트에서 질의를 요청할 경우 베뉴의 데이터 스토어에서는 인덱싱 모듈의 인덱싱한 결과를 검색해서 Result Set으로 베뉴 클라이언트에 전달한다.



(그림 6) 키워드 기반 검색을 제공하는 데이터 스토어의 구조

Result Set을 받은 베뉴 클라이언트의 데이터 스토어 클라이언트에서는 받아온 Result Set을 출력 해주면 되고, 출력된 결과를 가지고 사용자가 원하는 데이터를 찾을 수 있으며, 사용자로부터 다운로드 요청이 들어왔을 때, 베뉴의 베뉴 스토어에 요청을 해서 로컬 시스템에 저장한다.

이로써 데이터의 내용을 효율적으로 검색을 할 수 있게 되며, 필요에 따라서만 데이터를 다운로드 함으로써 불필요한 네트워크 트래픽을 줄일 수 있다.

#### 4. 결론 및 향후 연구 방향

엑세스 그리드의 데이터 공유에 키워드 기반 검색을 적용해서 데이터의 인덱싱된 내용을 가져와서 사용자에게 전달을 했다. 이렇게 함으로써 데이터가 방대해짐에 따라서 생기는 데이터의 공유와 활용에 대한 검색의 효율성을 높였다.

향후 연구에서는 더욱 다양한 질의문을 처리할 수 있도록 하고, 데이터를 검색함에 있어서 데이터

의 우선순위를 정하는 알고리즘을 적용하고자 한다. 이로써 보다 정확하게 원하는 데이터를 뽑아낼 수 있도록 하며, 마지막으로 데이터에만 적용된 키워드 기반 검색을 어플리케이션 공유에 적용할 수 있도록 하고자 한다.

#### 참고문헌

- [1] AccessGrid, <http://www.accessgrid.org/>
- [2] Stanford University : iRoom  
<http://iwork.stanford.edu/>
- [3] University of Washington: Portolano - An expedition into invisible computing,  
<http://portolano.cs.washington.edu/>
- [4] University of Illinois at Chicago : SAGE  
<http://www.evl.uic.edu/cavern/sage>
- [5] MSN Desktop Search, <http://desktop.msn.co.kr/>
- [6] Google Desktop Search,  
<http://desktop.google.co.k/>
- [7] Multicast Application Sharing Tool for the Access Grid Toolkit, AHM 2005, 2005.09.20
- [8] L. Childers, T. Disz, R. Olson, M. E. Papka, R. Stevens, and T. Udeshi. Access Grid: Immersive Group-to-Group Collaborative Visualization. In Proc. Immersive Projection Technology, 2000.
- [9] Access Grid Conference Facility at Boston University, <http://scv.bu.edu/accessgrid/>
- [10] Data Store programming Example,  
[http://www-new.mcs.anl.gov/fl/research/accessgrid/documentation/tutorial/AGTk\\_2.4/DataStore/index.htm](http://www-new.mcs.anl.gov/fl/research/accessgrid/documentation/tutorial/AGTk_2.4/DataStore/index.htm)
- [11] Integrating with the Access Grid: Experiences and Issues, AHM 2005, 2005.09.20