

# Peer-to-Peer 기반의 공간 데이터 미러링 시스템의 개발†

김한주<sup>0</sup>, 김동오, 윤재관, 한기준  
건국대학교 컴퓨터공학과  
{hjkim, dokim, jkyun, kjhan}@cb.konkuk.ac.kr

## Development of a Spatial Data Mirroring System based on Peer-to-Peer

Han-Ju Kim, Dong-O Kim, Jae-Kwan Yun, Ki-Joon Han  
Dept. of Computer Science & Engineering, Konkuk University

### 요 약

최근 컴퓨터 기술이 발전하면서 GIS(Geographic Information System)에 근거한 다양한 형태의 GIS 응용 시스템이 연구 및 개발되고 있으며, 또한 인터넷의 발전으로 사용자 측면의 기술도 발전되어 가고 있다. 이에 따라 다양한 사용자의 요구에 대해 빠른 결과를 제공할 수 있는 지리 정보 시스템이 필요하게 되었다. 그러나, 일반적인 지리 정보 시스템은 데이터의 처리와 전송에 상당한 시스템 자원을 소모하는 단점이 있어 이를 개선할 필요성이 있다. 각각의 컴퓨터를 직접 연결하는 P2P 기술은 자원의 공유를 통한 저장 공간의 절약, 데이터 전송의 효율성, 클라이언트의 기능 활용 등의 장점을 가지고 있다. 이러한 P2P를 이용하면 이전의 방법에서 문제시되었던 과부하, 접속 지연, 전송 비용 과다 등의 문제들을 해결할 수 있다. 따라서, 본 논문에서는 기존의 클라이언트/서버 형태 지리 정보 시스템의 문제점을 파악하고, 이를 극복하기 위한 방안으로 P2P를 살펴보았다. 또한, 이를 토대로 다양한 미러링 클라이언트를 두어 서버의 과부하나 전송 비용 과다를 줄이고, 클라이언트의 활용도를 높여 기존의 클라이언트/서버 형태의 지리 정보 시스템의 문제점을 보완한 P2P 기반의 공간 데이터 미러링 시스템을 설계 및 구현하였다.

### 1. 서론

최근 컴퓨터의 대용량화 및 고성능화에 따라 GIS(Geographic Information System)에 근거한 다양한 형태의 GIS 응용 시스템이 연구 및 개발되고 있다. 이전까지의 이러한 연구 및 개발은 주로 지리 정보 시스템에 대한 데이터베이스의 측면에서 많이 진행되었지만, 인터넷의 발전으로 WWW을 중심으로 한 사용자 측면의 기술로 발전되어 가고 있다[7]. 이에 따라 다양한 사용자의 요구에 대해 빠른 결과를 제공할 수 있는 지리 정보 시스템이 필요하게 되었으나, 일반적인 지리 정보 시스템은 데이터의 처리와 전송에 상당한 시스템 자원을 소모하게 된다. 따라서 지리 정보 시스템의 대용량 공간 데이터의 처리에 대한 효율적인 처리 방안과 대안이 필요하다[2, 8].

대용량 데이터 처리에 대한 효율적인 방법으로 P2P(Peer-to-Peer) 기술[6]이 있다. 이것은 각각의 컴퓨터를 직접 연결하여 자원을 공유하는 방법으로서, 컴퓨터의 파일이나 데이터에 대한 공유 기법으로부터 네트워크로 연결된 각각의 컴퓨터를 서로 연결하여 프로세스 사이클을 공유하는 분산 처리에 이르기까지 다양한 범위를 포함한다. 이러한 P2P 기술은 자원의 공유를 통한 저장 공간의 절약, 데이터 전송의 효율성, 빠르고 정확한 결과 도출 등의 장점을 갖고 있다.

P2P 기술의 또 다른 장점은 클라이언트의 기능을 활용한다는 점에 있다. 기존의 클라이언트/서버 시스템에서 클라이언트는 단지 하나의 사용자에게 불과하지만 P2P 방식에서 클라이언트는 서버의 일부 기능을 지원하고, 또한 정보를 제공하는 호스트로서의 역할도 수행한다. 즉, 서버 중심적인 이전의 방법에서 문제시되었던 과부하, 접속 지연, 네트워크 트래픽, 그리고 전송 비용 과다 등의 문제들을 효율적으로 해결할 수 있다. 본 논문에서는 기존의 클라이언트/서버 형태 지리 정보 시스템의 문제점을 보완한 P2P 방식의 공간 데이터 미러링 시스템을 설계 및 구현하였다.

본 논문의 구성은 다음과 같다. 제 1 장 서론에 이어 제 2 장에서는 관련 연구로 P2P, GML 2.0에 관해서 알아본다. 그리고, 제 3 장에서는 공간 데이터 미러링 시스템의 전체적인 구조 및 시스템을 구성하는 클라이언트와 서버의 설계에 대해서 언급하고, 제 4 장에서는 공간 데이터 미러링 시스템의 동작 과정과 구현에 대하여 설명한다. 마지막으로, 제 5 장에서는 결론에 대하여 언급한다.

### 2. 관련 연구

본 장에서는 데이터 공유 기술인 P2P에 대해 분석하고, 본 시스템에서 지리 정보의 저장 및 전송을 위해 사용하는 GML 2.0에 대해서 소개한다.

#### 2.1 P2P

P2P는 시스템간의 상호 연결을 통해 자원과 서비스를 공유하기 위한 것으로서, 공유하는 자원과 서비스의 범주에는 정보, 프로세스 사이클, 캐쉬 스토리지, 그리고 디스크 공간의 화일이 있다. 따라서, P2P 기술의 장점은 기존의 데스크 탑 컴퓨터를 서로 연결하여 더 나은 성능을 발휘할 수 있다는 점이다. P2P 구조에서는 기존의 단순한 클라이언트로 사용되는 컴퓨터들이 각자 직접 통신하고 클라이언트와 서버로서의 기능을 동시에 할 수 있다. 이것은 중앙 서버의 부하를 줄이며 클라이언트가 메일링 리스트나 지분 시스템과 같은 전문화된 서비스를 더욱 효과적으로 수행할 수 있도록 한다[6].

P2P 기술은 서로 연결된 각각의 peer들이 서버와 클라이언트의 두 가지 역할을 모두 수행한다는 점에서 클라이언트/서버 환경과의 차이가 있다. 이러한 P2P의 방식은 세 가지가 있다. 브로커 중재형 화일 공유 방식, 브로커가 존재하지 않는 직접적인 화일 공유 방식, CPU 사이클 공유 방식이 있다.

##### 2.1.1 브로커 중재형 화일 공유 방식

브로커 중재형 화일 공유 방식은 각각의 호스트 역할을 하는 클라이언트 사이에 브로커(Broker)를 두어 데이터를 관리하는 방식이다.

† 본 연구는 한국과학재단 목적기초연구(과제번호:R01-2001-00540) 지원으로 수행되었음.

이 모델의 장점은 peer의 정보를 정확하게 관리하는 브로커가 요청한 peer에게 적합한 정보를 전달하는 것으로써 처리 속도가 매우 빠르다는 것이다. 그러나, 브로커의 기능을 담당하는 중재자가 필요하고, 이 중재자의 비중이 높아서 중재자가 기능을 제대로 하지 못하거나 peer로부터의 과도한 요청이 들어올 때 트래픽 문제가 발생한다는 단점이 있다. 본 논문에서는 브로커 중재형 화일 공유 모델을 적용하여 공간 데이터 미러링 시스템을 개발하였다.

2.1.2 직접적인 화일 공유 방식

직접적인 화일 공유 방식은 각각의 호스트 사이에 중재자나 서버가 없이 네트워크로 연결된 servant(서버와 클라이언트를 합친 것)들이 상호 데이터 정보를 등록하여 화일을 송수신하는 방식이다. 이 모델은 브로커나 중앙의 중재자가 없어도 peer의 직접 연결을 지원하기 때문에 서버의 네트워크 트래픽 문제가 발생하지 않는다. 그러나, 적합한 servant를 검색하는데 많은 시간이 걸릴 가능성이 있으며, 정보의 관리에 일관성을 유지하기 어렵고, 중복된 데이터에 대한 검증이 어렵다는 단점이 있다. 이 모델의 예로는 Gnuteller, freenet 프로젝트 등이 있다.

2.1.3 CPU 사이클 공유 방식

CPU 사이클 공유 방식은 각 호스트 클라이언트의 CPU를 활용하여 프로세스의 결과를 이용하는 방법이다. 이 모델에서는 중앙의 마스터가 복잡한 연산 과정을 여러 단위로 나누어서 각각의 peer들에게 분배해 주면, peer는 각기 분배된 연산을 수행한다. 그리고, 각 peer는 수행의 과정을 끝낸 결과 데이터를 마스터에게 보내면 마스터는 그 결과를 취합하는 방식이다. 이 모델의 예로는 SETI@HOME 프로젝트 등이 있다.

2.2 GML 2.0

OGC(OpenGIS Consortium)에서는 지리정보의 상호운영성과 웹 환경에서의 사용성을 위해 Simple Feature 명세를 제시하였고[3], 또한 이에 대한 XML 인코딩 표준기술인 GML 모델로 제시하였다[5]. 즉, GML은 Feature의 공간, 비공간 property들을 XML 형식으로 표현한다. 기존의 GML 1.0은 지리 데이터를 XML로 인코딩하기 위해 사용하는 Feature와 Syntax를 DTD(Document Type Definition)로 문서 구조를 정의하였다. 하지만, GML 2.0은 다음과 같은 4가지의 스키마를 제시하고 있다. 즉, 점, 선, 면의 기하학적 요소인 Geometry 모델을 정의한 geometry.xsd, Feature 모델을 정의한 feature.xsd, Xlink를 사용하기 위해 정의하고 있는 xlink.xsd, 그리고 앞에서 언급한 3가지 기본 스키마들을 기반으로 사용자가 GML 문서를 참조하기 위한 응용 스키마가 있다.

3. 시스템의 설계

본 장에서는 공간 데이터 미러링 시스템의 전체 구조를 설명한 후, 공간 데이터 미러링 시스템의 서버와 클라이언트를 구성하고 있는 각 모듈에 대해 소개한다.

3.1 시스템 전체 구조

본 시스템은 클라이언트에게 지도 데이터를 제공하는 서버와 각 peer간의 효과적인 정보 전송을 고려하여 설계된 클라이언트로 구성된다. 서버는 현재 접속되어 있는 클라이언트의 상태를 수시로 감시하며 클라이언트의 정보를 관리하고, 클라이언트는 서버로부터 공간 데이터를 얻은 후 미러링 사이트로 등록하고 다른 클라이언트의 요청에 응답한다. 그림 1은 시스템의 전체 구조를 보여 준다.

그림 1에서 보는 것처럼 클라이언트는 사용자의 질의가 입력되면 질의를 서버로 전송한다. 이때, 서버의 질의 관리자는 질의를 분석하여 클라이언트 정보 관리자에서 질의에 해당하는 클라이언트가 존재할 경우 클라이언트의 정보를 보내고, 그렇지 않은 경우 서버의 공간 데이터를 클라이언트에 전송하게 된다.

3.2 서버

서버는 클라이언트 정보 관리자, 질의 관리자, 서버 GML 관리자, 인덱스 관리자, 격자 관리자로 구성된다.

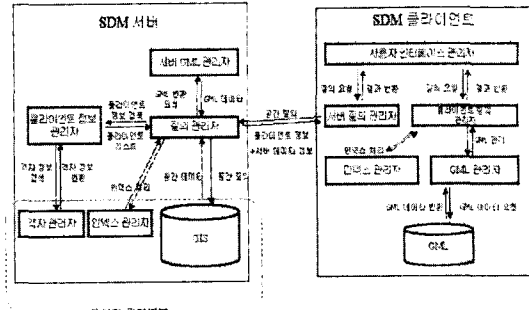


그림 1 시스템 전체 구조도

3.2.1 클라이언트 정보 관리자

클라이언트 정보 관리자는 클라이언트의 정보와 클라이언트가 보유한 공간 데이터에 대한 정보를 관리하고, 질의를 요청한 클라이언트에게 해당하는 클라이언트의 정보를 반환하는 모듈이다.

3.2.2 질의 관리자

질의 관리자는 클라이언트가 요청한 질의를 분석하여 그 결과를 반환하는 역할을 하는 모듈이다. 이것은 클라이언트로부터 받은 질의를 처리하기 위해 먼저 공간 데이터를 소유한 미러링 클라이언트를 검색하고 그 결과로 클라이언트의 IP와 Port번호를 반환한다. 만약 요구한 질의에 적합한 클라이언트를 검색하지 못한 경우 서버의 공간 데이터를 클라이언트에게 직접 반환한다.

3.2.3 서버 GML 관리자

서버 GML 관리자는 서버의 공간 데이터를 GML로 변환하는 기능을 수행하는 모듈이다. 또한, 서버 GML 관리자는 격자 관리자에서 영역별로 분할된 벡터 형태의 공간 데이터를 얻어서 GML 명세에 맞도록 변환한 후 저장한다.

3.2.4 인덱스 관리자

인덱스 관리자는 서버의 효율적인 공간 데이터 검색, 공간 데이터 분할, 격자 화일의 일관성있는 관리를 위해 공간 데이터 및 비공간 데이터에 대한 인덱스를 생성하고 관리하는 모듈이다.

3.2.5 격자 관리자

격자 관리자는 서버의 전체 공간 데이터를 영역별 격자식으로 분할하고, 격자식으로 분할된 공간 데이터를 GML 관리자에게로 전달하는 모듈이다. 격자의 분할 배수를 높일수록 더 정확한 자료에 대한 검색이 가능하게 되지만, 검색이나 질의의 복잡성은 높아지게 된다.

3.3 클라이언트

클라이언트는 사용자 인터페이스 관리자, 서버 질의 관리자, 클라이언트 질의 관리자, 공간 데이터 조합-분할 관리자, 인덱스 관리자, GML 관리자로 구성된다.

3.3.1 사용자 인터페이스 관리자

사용자 인터페이스 관리자는 서버로부터 처리된 결과를 화면에 보여 주며, 또한 사용자로부터 질의를 입력받는 모듈이다. 그리고, 영역별 질의, 이동, 확대 및 축소 기능 등을 지원한다.

3.3.2 서버 질의 관리자

서버 질의 관리자는 클라이언트에서 요청된 질의를 서버로 전달하고, 그 결과를 처리하는 모듈이다. 이것은 서버로부터 다른 클라이언트의 정보를 받거나, 또는 서버의 공간 데이터를 직접 받는다.

3.3.3 클라이언트 질의 관리자

클라이언트 질의 관리자는 서버 질의 관리자에서 얻은 클라이언트의 정보를 이용하여 다른 미러링 클라이언트에게 공간 데이터를 요청하는 모듈이다.

