

# 확장된 WARD기법을 사용한 웹 GIS 클러스터링 시스템

°장용일\* 이충호\* 이재동\*\* 배해영\*  
\*인하대학교 전자계산공학과  
\*\*단국대학교 전산과  
himalia@dblal.inha.ac.kr

## Web GIS Cluster Design with Extended Workload-Aware Request Distribution(WARD) Strategy

°Yong-Il Jang\* Chung-Ho Lee\* Jae Dong Lee\*\* Hae-Young Bae\*  
\*Dept. of Computer Science and Engineering, Inha University  
\*\*Dept. of Computer Science, Dankook University

### 요 약

웹 클러스터링 시스템은 사용자의 집중에 의한 서버의 과부하에 효율적으로 대처할 수 있는 방법 중 하나이다. 공간 질의의 대부분은 인접 영역에 대한 경우가 매우 많으며, 특정 영역에 집중되는 특성을 갖는다. 타일 기반 웹 GIS 클러스터링 시스템은 이러한 공간 질의의 특성을 만족시키면서 클러스터에 포함되는 각 노드의 버퍼 재 사용률을 높이고 디스크 접근 빈도를 낮출 수 있는 기법을 사용한다. 그러나, 모든 질의가 디스패처를 거치기 때문에 병목현상이 나타날 수 있으며, 질의의 빈도가 낮은 지역 경우 할당된 지역의 범위에 의해 버퍼 관리에 문제가 생긴다.

본 논문에서는 확장된 WARD기법을 사용한 웹 GIS 클러스터링 시스템을 제안한다. 제안되는 시스템은 타일 기반 웹 GIS 클러스터링 시스템을 기반으로 디스패처를 분산시켜 병목현상을 줄이고, 각 지역을 그룹으로 묶어 클러스터를 분배하여 대용량의 공간 데이터를 위해 적절한 버퍼 관리를 한다. 또한, 질의가 집중되는 영역에 대한 레이어를 모든 노드의 코어 영역에 중복 저장하여 로컬에서 지역적으로 처리하고 다른 노드로의 질의 처리 요구에 대한 포워딩(Forwarding) 오버헤드를 줄여 기존의 서버에 비해 안정성과 확장성, 그리고 처리 능력을 향상시킬 수 있다.

## 1. 서 론

인터넷을 사용하는 사용자가 늘어나고 서버에 대한 접속 수가 많아지면서 웹 서버의 부하는 점점 더 커지게 되었다. 클러스터링 시스템은 효율적으로 서버의 확장성과 가용성을 증가시키기에 이러한 환경에서 가장 어울리는 방법이라 할 수 있다. 클러스터링 시스템은 더불어서 클러스터의 크기가 커지더라도 안정된 성능 향상을 보여야 하기 때문에 균형잡힌 질의의 분산과 몇몇 오버헤드가 중요하게 작용한다.

일반적인 웹 데이터는 요청들 사이에 밀접한 연관을 가지지 않으며, 하나의 질의 요청에 대해 반환되는 리소스 역시 많지 않다. 이에 반해, 공간 데이터는 일반적인 웹 데이터와는 다르게 데이터가 큰데다가 인접한 영역에 대한 질의가 많고, 작업 처리량이 많으며, 특정 영역에 질의가 집중이 되는 특성을 가진다. 이러한 공간데이터의 특성을 클러스터링 시스템에서 고려하지 않는다면 클러스터의 각 노드의 버퍼 사용률이 낮아지고, 디스크의 접근 빈도수가 많아져서 결국 전체적인 시스템 성능이 떨어지게 된다.

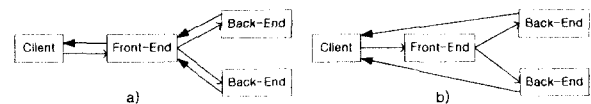
이러한 환경에서 타일 기반 웹 GIS 클러스터링 시스템은 질의의 인접성을 이용한 로드 밸런싱 기법을 써서 각 노드의 버퍼 사용률을 높이고 디스크 접근 빈도를 낮추어 시스템의 처리량을 증가시킨다[1]. 그러나, 이 방법을 사용하는 경우 단 하나의 디스패처를 사용하는 기존의 구성은 다음의 문제점을 가진다. 첫 째, 모든 질의요청이 디스패처를 거치고 결과 또한 디스패처를 거쳐야 하므로 사용자가 집중되는 경우 병목현상이 일어날 가능성이 있다. 둘째, 디스패처가 예기치 못한 사고로 정지된다면 클러스터 전체가 마비된다. 셋 째, 대부분의 공간질의는 특정 지역에 집중되는 특성을 가지는데 이러한 특성을 기존의 시스템에서 타일을 조정함으로써 그 문제를 해결하긴 하지만 이 경우 반대로 접근 횟수가 낮아 다른 노드보다 넓은 타일을 관리하게 되는 노드는 버퍼에 저장될 영역이 같이 넓어져 효율적인 버퍼의 사용에 영향을 준다.

본 논문에서는 확장된 WARD(Workload-Aware Request Distribution) 기법을 사용한 새로운 웹 GIS 클러스터링 시스템을 제안한다[2]. 제안된 기법에서는 질의를 요청한 클라이언트에게 클러스터의 각 노드가 직접 처리 결과를 전달하는 TCP handoff 전달 기법을 사용함과 동시에 디스패처를 클러스터의 각 노드에 분산시켜 기존 시스템에서의 병목현상을 줄이며 시스템의 확장성을 가져온다[3]. 그리고, 접근 빈도가 높은 영역의 레이어와 영역 데이터를 모든 노드가 공동으로 버퍼상에 놓음으로써 포워딩(Forwarding) 오버헤드를 줄이는 WARD기법을 사용하여 시스템의 전체적인 처리량을 개선한다.

## 2. 관련 연구

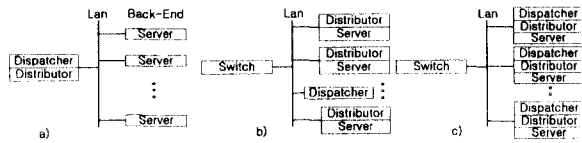
### 2.1 웹 클러스터의 구조

질의 요청을 내용에 의해 분산시키는 방법으로는 TCP TCP handoff 방식 또는 TCP TCP splicing이 있다[3]. TCP splicing은 front-end에 의해 트래픽의 최적화가 이루어지며, TCP handoff는 질의의 결과를 직접 클라이언트에게 전달하는 구조를 갖는다. TCP handoff는 TCP splicing보다 높은 확장성과 향상된 속도의 장점을 가진다.



[그림 1] 트래픽 흐름 구조 a) TCP splicing 기법 b) TCP handoff 기법  
TCP handoff방식은 아직 front-end의 병목현상을 야기시키는 문제점을 갖는다. 이러한 문제를 해결하기 위해 클러스터 서버의 구조에서 질의 분산을 담당하는 디스패처를 각 클러스터에 분산시키고 front-end에는 물리적 단계에서 빠르게 스위칭 기능을 수행하는 LAN 스위치 또는 L4

레벨의 로드 밸런싱 기기가 사용될 수 있다[4].



[그림 2] 클러스터 구조

[그림 2]의 a)는 단일 front-end를 가지는 가장 보편적인 클러스터 구조이다. 병목현상을 발생시킬 수 있으며, 단일 디스패처에 의해 중앙에 집중된 관리 형태를 가진다. b)는 a) 형태에서의 front-end를 스위치로 대체하여 병목현상을 해결하고 RR-DNS기법 등을 이용해 클러스터에 순차적으로 연결을 한다[3]. 이 구조의 동작 형태는 다음과 같다.

1) 클라이언트가 임의의 선택된 클러스터 노드의 Distributor에 접속한다. 2) Distributor에 의해 연결이 승인되고 질의를 분석한다. 3) Dispatcher에 접속하여 질의에 대한 서버 분배를 받는다. 4) Distributor에 의해 각 서버로의 질의 분배 및 처리가 이루어진다. 5) 서버들은 TCP handoff방식에 의해 직접 클라이언트로 결과를 보낸다[2,3].

마지막으로 c)의 구조는 Dispatcher를 클러스터의 각 노드에 분산시키는 형태를 띤다. b)의 구조와 비교해 현 구조는 Dispatcher를 분리함으로써 단일 Dispatcher로부터 질의 분배를 받는데 네트워크 사용에 걸리는 시간과 단일 Dispatcher의 위험성 또한 해결이 가능하다[2].

c)의 구조는 고확장성과 가용성을 얻게 하는 장점을 가진다. 그러나, 대부분의 질의 요청에 대해서 질의를 받은 노드는 Dispatcher에 의해 다른 노드로 해당하는 질의를 보내게 되는데, 이 때 네트워크를 통해 질의를 보내므로 포워딩 오버헤드가 발생하며 클러스터의 수가 많으면 많을수록 포워딩 오버헤드는 증가한다.

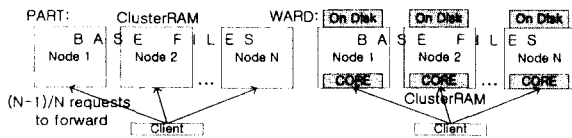
2.2 Workload-Aware Request Distribution(WARD) 기법

WARD기법은 기본적으로 [그림 2]의 c) 클러스터의 구조를 기본 바탕으로 삼는다. WARD는 호출 빈도가 가장 큰 데이터 집합을 core라는 곳에 저장하고 이 외의 데이터는 각 노드에 분산하여 part영역에 저장한다. core와 part는 모두 메모리의 영역이다. 전체 노드가 메모리 안에 core영역을 가지고 있으며 같은 데이터를 중복하여 저장한다. WARD기법은 이렇게 core영역의 데이터를 클러스터에 중복 저장하여 포워딩 오버헤드를 줄이고 디스크 접근 횟수를 줄인다.

WARD의 목적은 다음과 같다.

- 1) 분할된 파일이 각각의 서버의 모든 클러스터 메모리로부터 최대의 질의를 처리하게끔 한다
- 2) 포워딩 오버헤드를 최소화 한다.

WARD는 이러한 목적을 위한 알고리즘으로 ward-analysis 기법을 사용하여 이 알고리즘은 독립적인 파일의 접근 빈도와 데이터의 크기를 연산 매개로 사용한다.



[그림 3] WARD 기법 구조

3. 확장된 WARD기법을 사용한 웹 GIS 클러스터링 시스템 설계

3.1 질의 영역의 인접성을 이용한 로드 밸런싱 기법과 디스패처

기존에 연구되었던 질의 영역의 인접성을 이용한 로드 밸런싱 기법은 타일 기반 웹 GIS 클러스터링 시스템에 기반을 둔다[1]. 이는 이미 서버가 공간 데이터를 타일 단위로 분할하여 관리하고 있음을 의미하고, 이 분할 기준인 타일을 2차원 공간 매핑을 위해 그리드로 고려하여 힐버트

곡선(Hilbert Curve)을 사용해 공간 질의의 특성 중 하나의 질의 영역의 인접성을 적용한다. 힐버트 곡선을 사용하면 인접된 공간의 데이터 정렬이 가능하다[6].

공간 질의 디스패처는 통신 관리기의 질의 영역 추출기로부터 전달되는 공간 질의를 질의 분류 데이터를 이용하여 분류하고 질의 영역인 타일들로부터 질의 요청을 처리하기 위한 클러스터링 서버를 결정하여 통신 관리기의 질의 요청 리디렉터(Redirector)로 넘겨준다. 디스패처는 질의 요청에 대한 결과를 통계 정보 관리기로 넘기어 질의 처리 영역 및 각 타일에서 처리된 질의의 통계 정보를 기록하도록 한다[1].

3.2 확장된 WARD기법

WARD 기법은 일반적인 웹 데이터에 위한 기법이다. 이 기법을 공간 데이터에 적용하기 위해서 공간 데이터의 인접성, 데이터 크기, 질의의 집중성이 고려되어야 한다. 그러나 기존의 WARD 기법은 이러한 특성을 고려하지 않아 공간 데이터에 대해 효율적이지 못하다.

첫 째, 공간 데이터의 데이터 크기로 인해 빈도가 잦은 지역의 데이터를 모두 core안에 넣기엔 메모리의 공간이 부족할 수 있다. 그리고, 질의의 빈도가 낮아 part영역에 저장되는 객체는 공간 데이터의 특성이 고려되지 않아 성능을 낮추는 요인이 된다. 공간 데이터는 일반적으로 여러 레이어로 구성이 된다. 도로, 구경계, 강과 같은 특정 레이어는 대부분의 질의에서 표현이 되며, 이러한 특성을 고려하여 확장된 WARD기법에서는 core영역을 확장하여 질의 빈도가 높은 지역을 저장하는 영역과 특정 레이어 영역의 두 가지를 함께 저장하는 구조를 가진다.

확장된 WARD기법의 특성을 정리하면 다음과 같다.

- 1) core영역에 포함되는 공간 데이터: 일반적인 웹 데이터는 웹 페이지에 대한 접근 빈도만을 고려하지만 본 논문의 공간 데이터의 특성을 고려해 표현 빈도가 높은 기본 레이어와 접근 빈도가 높은 지역의 두 가지 정보를 함께 저장한다.
- 2) 지역별 그룹: 웹 데이터들은 공간 데이터에 비해서는 아주 작기 때문에 core영역의 크기에 대한 고려가 중요하지 않다. 그러나 공간 데이터의 데이터 크기 때문에 확장된 기법에서는 그 영역들을 core영역의 크기에 맞게 그룹으로 묶어 활용한다. 그 기준은 데이터와 분포도에 의해 달라진다.
- 3) part영역의 데이터 관리: part영역의 데이터는 기존 기법에서 Dispatcher에 의해 분산된 영역에 의한 처리만을 하지는 않는다. 버퍼의 효율적인 관리를 위하여 질의의 인접성을 고려한 로드밸런싱기법이 적용된다.
- 4) 동적인 Dispatcher의 질의 분산: 처리되는 데이터 크기에 비해 메시지는 상대적으로 작다. 기존 WARD기법에서 이전 로 그 데이터를 이용하여 질의 분산을 수행하여 5~10%의 성능 저하를 가져오던 방식보다 동적인 메시지 송수신에 의해 포워딩 오버헤드를 감소하면서 동적으로 질의 분산을 조절하는 것이 처리 데이터가 큰 현재의 환경에 보다 효율적이다.
- 5) 확장된 Dispatcher: 동적으로 질의 분산을 하는 현재의 방법은 공간 데이터의 특성을 고려하더라도 포워딩 오버헤드를 여전히 수반한다. Global Dispatcher는 클러스터 그룹에 대한 디스패처이다. 클러스터 그룹에 대한 질의 분산을 수행하며 병목현상을 없애기 위해 복수의 형태로 존재하며 병목 현상을 방지하며, URL의 내용만 판단하여 질의 분산을 수행하므로 부하를 줄인다. Local Dispatcher는 그룹 내의 Dispatcher들과 메시지 송수신을 하며 질의 영역 분산을 수행한다. 두 단계로 나누어진 Dispatcher에 의해 포워딩 오버헤드는 그룹 내로 한정된다.

3.3 Dispatcher와 Distributor 질의 분산 기법

Global Dispatcher와 Distributor는 클러스터의 그룹을 관리한다. 예를 들어, [그림 4]를 살펴보면 하나의 그룹은 하나의 시를 담당하고, 그룹 내의 각 노드들은 core영역에 시 내의 기본 레이어와 접근 빈도가 높은

영역을 저장하고 part영역은 각자 영역을 분산시켜 버퍼 관리를 수행한다.



[그림 4] 질의 영역의 집중성

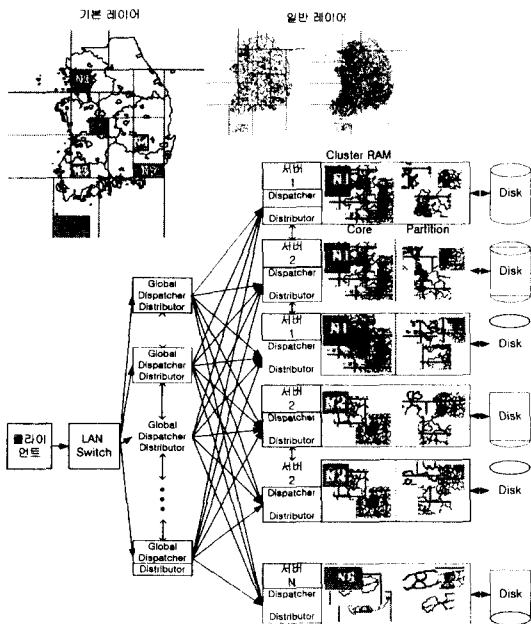
Global Dispatcher는 URL의 분석 및 통계를 통해 그룹 관리 및 질의 분산을 하며, Distributor는 분산된 질의를 실제 그룹의 노드의 주소로 알고 있으며, 이전 질의 통계를 기억한 후 RR방식으로 질의를 보낸다.

Local Dispatcher와 Distributor는 그룹이 담당하는 영역 내의 지역에 대해 담당한다. Local Dispatcher는 Global Dispatcher와는 달리 core와 part정보를 관리하며, Distributor는 로컬 노드와 원격 노드로의 질의 분산을 담당한다.

### 3.4 확장된 WARD 기법을 사용한 웹 GIS 클러스터링 시스템

확장된 웹 GIS 클러스터링 시스템의 특징은 다음과 같다.

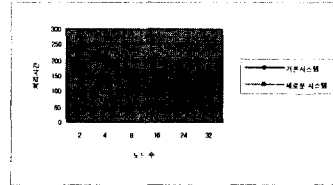
- TCP handoff 기법 사용 : 처리량 증대
- 분산된 Global Dispatcher와 Local Dispatcher : 병목 현상과 포워딩 오버헤드 감소
- 동적인 버퍼 관리 : 최적화된 버퍼의 관리 가능
- core 영역 : 접근 빈도가 높은 데이터에 대한 중복 저장을 통해 수행 속도를 높임
- part 영역 : 질의 영역의 인접성을 고려한 버퍼 관리로 버퍼 Hit률을 높이고 디스크 접근 빈도를 낮춤



[그림 4] 확장된 WARD기법을 적용한 시스템 구성도

### 4. 성능평가

본 장에서는 기존의 타일기반 웹 GIS 클러스터링 시스템과 확장된 WARD기법을 적용한 두 타입에 대한 성능평가를 하였다. [그림 6]은 노드 수가 2, 4일 때 약 20%의 성능향상을 보이고 있다. 노드 수가 많아지면 기존 시스템은 처리시간이 다시 증가하지만 제안된 시스템은 평균화 된 성능을 나타낸다.



[그림 6] 노드와 질의 수에 따른 질의 처리시간

### 5. 결론

본 연구는 기존 타일 기반의 웹 GIS 클러스터링 시스템의 문제점 보완과 향상된 성능에 초점이 맞추어졌다. 기존 시스템의 몇몇 문제들을 본 논문에서 제시하는 확장된 WARD기법을 사용한 웹 GIS 클러스터링 시스템에서 해결을 하고 있다.

단일 Dispatcher는 분산된 질의의 분산이 가능한 형태로 바뀌어 병목현상을 해결하고 사고에 대한 안정성을 높였고, TCP handoff방식의 질의 결과의 처리는 동일 시간 내에서 클러스터링 시스템이 처리할 수 있는 총 데이터의 양을 증가시켰다. 기존의 인접 영역에 대한 로드밸런싱 기법에서 공간 데이터의 중첩 레이어를 고려한 core와 part 영역의 버퍼 관리는 보다 좋은 성능을 보여준다.

제안된 웹 GIS 클러스터링 시스템은 분산된 Dispatcher와 Global Dispatcher에 의해 이전 시스템에 비해 고 확장성을 갖는다. Global Dispatcher가 클러스터의 그룹을 지정하는데 있어서 중간에 스위치를 두어 그룹을 마치 하나의 노드로 인식하게 한다면 Global Dispatcher는 더 많은 그룹들을 지정할 수 있다.

향후 연구과제로는 영역과 중첩되는 레이어 간의 관계를 고려한 로드 밸런싱 기법과 디클러스터링과 클러스터링을 중첩한 고 가용성 웹 GIS 서버에 대한 연구가 필요하다.

### 6. 참고문헌

[1] 이찬규, "Web GIS 클러스터링 시스템에서질의 영역의 인접성을 이용한 로드 밸런싱 기법", 인하대학교 전자계산공학과 석사학위 논문, 2001

[2] L. Cherkasova, M Karlsson. "Scalable Web Server Cluster Design with Workload-Aware Request Distribution Strategy WARD", In Advanced Issues of E-Commerce and Web-Based Information Systems, WECWIS 2001, Third International Workshop on.

[3] M. Aron, D. Sanders, P. Druschel, and W. Zwaenepoel. "Scalable content-aware request distribution in cluster-based network servers", In Proceedings of the USENIX 2000 Annual Technical Conference.

[4] L. Cherkasova. "FLEX: Load Balancing and Management Strategy for Scalable Web Hosting Service", In Proceedings of the Fifth IEEE Symposium on Computers & Communications (ISCC'00).

[5] H. V. Jagadish. "Linear Clustering of Objects with Multiple Attributes", In Proceedings of the 1990 ACM SIGMOD International Conference on Management of Data, 1990

[6] 조영성, "타일 저장구조와 하이브리드 공간 질의 처리기법을 이용하는 클라이언트 중심 웹 지리정보시스템", 인하대학교 전자계산공학과 박사학위 논문, 1999