

오디세우스/Parallel-OOSQL 에 기반한 대규모 병렬 정보검색 서비스 시스템 아키텍처

성경복⁰ 이재길 황규영
한국과학기술원 전산학과/첨단정보기술연구센터
{kbseong, iglee, kywhang}@mozart.kaist.ac.kr

A Service System Architecture of a Large Parallel Information Retrieval System
Based on ODYSSEUS/Parallel-OOSQL

Kyung-Bok Sung⁰ Jae-Gil Lee Kyu-Young Whang

Department of Computer Science and
Advanced Information Technology Research Center
Korea Advanced Institute of Science and Technology

요 약

인터넷에 존재하는 데이터의 양이 기하급수적으로 증가함에 따라 많은 양의 데이터에 대해 정보 검색을 효율적으로 지원하기 위해 병렬 정보검색 시스템이 개발되었다. 그러나 대규모 병렬 정보검색 서비스 시스템의 운영에 관해 발표된 자료가 미비하여 실제로 대규모 병렬 정보검색 시스템을 구축하고 운영하기에는 어려움이 있다. 본 논문에서는 대규모 병렬 정보검색 서비스 시스템의 아키텍처를 제안한다. 이를 위해, 1) 병렬 정보검색 서비스 시스템을 구축하기 위한 물리적인 기기 구성을 보이고, 2) 검색 서비스 중에도 빠른 데이터 추가가 가능한 데이터 추가 방법을 고안하며, 3) 데이터베이스 재구축 중에도 지속적인 서비스가 가능한 데이터베이스 재구축 방법을 고안한다.

1. 서론

인터넷에 존재하는 데이터의 양이 증가함에 따라 많은 양의 데이터에 대해 정보검색을 효율적으로 지원해 주는 것이 더욱 중요해지고 있다. 인터넷 등장 초기에는 데이터의 양이 적어 단일 정보검색 시스템으로도 효율적인 정보검색을 지원할 수 있었으나, 인터넷이 대중화 됨에 따라 데이터의 양이 기하급수적으로 증가하여 단일 검색 시스템은 그 한계를 드러내었다. 따라서 대규모 데이터에 대한 효율적인 검색을 지원하기 위해 병렬 정보검색 시스템이 도입되었다. 실제로 구글(Google), 야후(Yahoo), 알타비스타(AltaVista)와 같은 웹 검색 시스템은 대규모 데이터에 대한 효율적인 검색을 지원하기 위해 병렬 정보검색 시스템을 사용하는 것으로 알려져 있다[1,2].

병렬 정보검색은 주어진 키워드와 매칭되는 문서를 찾기 위해 사용되는 색인 및 문서 자체를 여러 대의 컴퓨터에 나누어 저장한 후, 나누어 저장된 색인 및 문서를 병렬적으로 검색하여 질의 결과를 병합하는 방식이다[3]. 병렬 정보검색을 사용할 경우 여러 대의 컴퓨터를 사용하여 색인 및 문서를 나누어 저장하기 때문에 더 큰 저장공간을 사용할 수 있을 뿐만 아니라, 검색을 병렬로 수행함으로써 질의 처리 시간을 줄일 수 있다.

본 논문에서는 한국과학기술원 첨단정보기술연구센터에서 개발한 병렬 정보검색 시스템인 오디세우스/Parallel-OOSQL[4]을 이용하여 실제로 대규모 병렬 정보검색 서비스 시스템을 운영하기 위한 아키텍처를 제안한다.

본 논문에서 제안한 대규모 병렬 정보검색 서비스 시스템 아키텍처는 다음과 같은 특징을 가진다. 첫째, 여러 대의 서버와 디스크 어레이를 사용하여 대규모 병렬 정보검색을 효율적으로 지원하는 물리적인 기기 구성 방법을 제안한다. 둘째, 서비스 도중에도 빠르게 수행 가능한 데이터 추가 및 업데이트 방법을 제안한다. 대규모 데이터가 이미 존재하는 상황에서 데이터 추가 및 업데이트는 많은 비용이 소요되므로 검색 서비스에 미치는 영향을 최소화 하면서도 빠른 데이터 추가

및 업데이트가 가능하여야 한다. 셋째, 서비스의 중단 없이 전체 데이터베이스를 재구축 할 수 있는 기법을 제안한다. 데이터베이스 재구축은 전체 데이터베이스를 새로 구축하는 것이며, 이때 검색 서비스의 중단 없이도 데이터베이스 재구축이 가능하여야 한다.

본 논문의 구성은 다음과 같다. 제 2 절에서는 본 논문에서 대규모 병렬 정보검색 서비스 시스템을 구축하는데 사용한 기반 시스템인 오디세우스/Parallel-OOSQL 을 소개한다. 제 3 절에서는 대규모 병렬 정보검색 서비스 시스템 아키텍처를 제안한다. 마지막으로 제 4 절에서는 결론을 내린다.

2. 오디세우스/Parallel-OOSQL

오디세우스/Parallel-OOSQL[4]은 본 논문에서 병렬 정보검색 서비스 시스템을 구축하는데 사용한 기반 시스템이다. 오디세우스/Parallel-OOSQL은 그림 1과 같이 질의를 분석하는 부분인 마스터(master) 오디세우스와 실제 질의를 수행하는 부분인 슬레이브(slave) 오디세우스로 구성된다. 마스터 오디세우스는 사용자로부터 받은 질의를 분석하여 검색을 해야 하는 문서 블록(block)을 결정한다. 문서 블록이란 문서 분할의 단위로, 전체 문서는 여러 개의 문서 블록으로 분할된다. 검색을 해야 하는 문서 블록을 결정 한 후, 마스터 오디세우스는 각 슬레이브 오디세우스가 처리해야 하는 질의를 전달하여 해당 슬레이브 오디

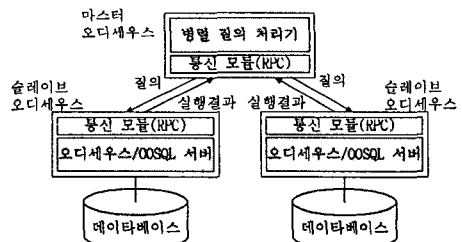


그림 1. 오디세우스/Parallel-OOSQL 의 아키텍처

* 본 연구는 첨단정보기술연구센터를 통하여 과학재단의 지원을 받았음.

세우스가 실제로 검색을 수행하도록 한다. 또한 마스터 오디세우스는 슬레이브 오디세우스로부터 검색 결과를 받아 병합하여 최종 검색 결과를 생성한 후 사용자에게 그 결과를 전달한다. 슬레이브 오디세우스는 마스터 오디세우스로부터 받은 질의를 가지고 실제로 검색을 수행한 후 검색 결과를 마스터 오디세우스에게 돌려준다. 마스터 오디세우스와 슬레이브 오디세우스는 슬레이브 오디세우스가 처리해야 하는 질의와 질의 수행 결과를 전달하기 위해 RPC(Remote Procedure Call)를 사용한다.

오디세우스/Parallel-OOSQL에서는 역색인을 분할하여 데이터베이스를 구축한다. 역색인 분할 방법에는 문서 식별자 분할 방법과 키워드 식별자 분할 방법이 있다[5]. 문서 식별자 분할 방법은 같은 문서 식별자에 대한 포스팅 리스트는 같은 역색인 블록에 포함되도록 분할하는 방법이다. 포스팅이란 키워드가 발생한 문서의 식별자와 그 문서 내에서의 발생 위치 정보를 가지는 데이터이고, 포스팅 리스트란 포스팅들의 리스트이다. 역색인 블록이란 역색인 분할의 단위를 말한다. 키워드 식별자 분할 방법은 같은 키워드 식별자에 대한 포스팅 리스트는 같은 역색인 블록에 포함되도록 분할하는 방법이다. 오디세우스/Parallel-OOSQL에서는 병렬 정보검색 서비스 시스템 구축을 위한 역색인 분할 방법 중 문서 식별자 분할 방법과 키워드 식별자 분할 방법을 모두 지원할 뿐만 아니라 이 두가지 방법을 혼합하여 사용할 수 있는 혼합 분할 방법도 지원한다. 본 논문에서는 웹 검색 시스템의 환경에 보다 더 적합한 것으로 알려진 문서 식별자 분할 방법을 사용하여 병렬 정보검색 시스템을 구축한다.

3. 대규모 병렬 정보검색 서비스 시스템 아키텍처

본 절에서는 대규모 병렬 정보검색 서비스 시스템의 아키텍처를 제안하고 제안한 아키텍처 상에서 서비스를 운영하는 방법을 제안한다. 제 3.1 절에서는 대규모 병렬 정보검색 서비스를 위한 시스템 기기 구성을 설명한다. 제 3.2 절에서는 대규모 병렬 정보검색 서비스를 위한 데이터베이스 운영 방법을 설명한다.

3.1. 병렬 정보검색 서비스 시스템의 기기 구성

본 논문에서 제안하는 대규모 병렬 정보검색 서비스 시스템의 아키텍처는 그림 2 와 같다. 마스터 오디세우스를 가지는 마스터 검색 서버가 존재하고, 슬레이브 오디세우스를 가지는 슬레이브 검색 서버가 존재한다. 그리고 데이터베이스 저장을 위한 디스크 어레이가 있다. 또한 데이터 로딩을 위한 데이터 로딩 서버와 데이터 수집을 위한 데이터 수집 서버가 별도로 존재한다. 그리고 서버들과 디스크 어레이를 연결해 주는 SAN 스위치가 존재한다. SAN 스위치란 서버와 스토리지 간에 Any-to-Any 액세스를 제공하기 위해, 스토리지 프로토콜을 사용하여 하나 이상의 서버와 스토리지를 연결하는 Storage Area Network 을 구성할 수 있도록 하는 장비이다[6].

마스터 검색 서버와 슬레이브 검색 서버는 LAN 을 통하여 연결되고, 각 슬레이브 검색 서버는 자신이 사용하는 데이터베이스 분할을 가진 디스크 어레이와 SAN 스위치를 통하여 연결된다. 데이터 수집 서버와 데이터 로딩 서버는 LAN 을 통해 연결되고, 데이터 로딩 서버는 데이터 로딩을 위해 디스크 어레이에 접근할 수 있어야 하므로 SAN 스위치와도 연결된다.

마스터 검색 서버에는 여러 개의 마스터 오디세우스가 실행될 수 있고, 하나의 슬레이브 검색 서버에는 여러 개의 슬레이브 오디세우스 셋이 실행될 수 있다. 하나의 슬레이브 오디세우스 셋에는 현재 실행 중인 마스터 오디세우스의 개수와 동일한 수의 슬레이브 오디세우스가 실행된다. 동일한 슬레이브 오디세우스 셋 내의 슬레이브 오디세우스들은 동일한 데이터베이스 분할에 대한 검색을 담당한다.

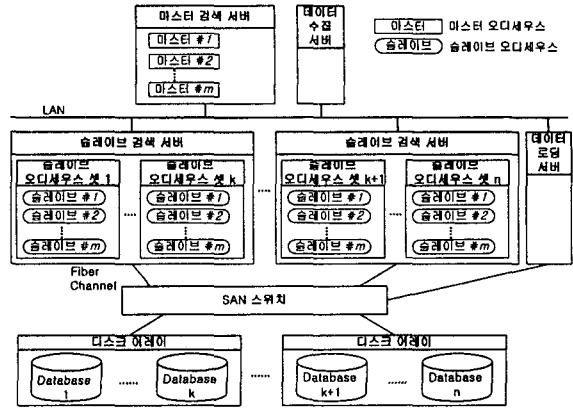


그림 2. 병렬 정보검색 서비스 시스템 아키텍처.

그림 2 의 아키텍처에서 데이터베이스 구축은 다음과 같은 순서로 이루어진다. 먼저 데이터 수집 서버를 이용하여 데이터를 수집한다. 그 후, 데이터 수집 서버에서는 수집한 데이터를 분할하고 키워드 추출을 수행하여 분할된 데이터베이스에 로딩할 데이터를 만들고 포스팅 파일을 생성한다. 데이터 분할 및 포스팅 파일의 생성이 종료되면 분할된 데이터 및 생성된 포스팅 파일을 데이터 로딩 서버에 전송하여 데이터베이스를 구축한다. 데이터베이스 구축 및 운영방법에 대해서는 제 3.2절에서 자세히 설명한다.

그림 2 의 아키텍처에서 검색은 다음과 같은 순서로 이루어진다. 먼저 마스터 오디세우스를 통해서 사용자로부터 질의를 받는다. 사용자로부터 질의가 들어오면 마스터 오디세우스는 질의를 분석하여 각 슬레이브 오디세우스에서 수행해야 하는 질의로 분할하여 LAN 을 통해 슬레이브 오디세우스로 전달한다. 각 서버의 슬레이브 오디세우스는 SAN 스위치를 통해 데이터베이스를 액세스 함으로써 마스터 오디세우스로부터 전달받은 질의를 처리하여 결과를 얻는다. 그 후, 슬레이브 오디세우스는 질의 결과를 LAN 을 통해 다시 마스터 오디세우스에 전달하고 마스터 오디세우스는 각 슬레이브 오디세우스로부터 전달받은 질의 결과를 병합하여 최종적으로 사용자에게 전달한다.

오디세우스/Parallel-OOSQL에서는 1개의 마스터 오디세우스가 동작하기 위해 필요한 슬레이브 오디세우스의 수는 데이터베이스 분할수와 동일하다. 따라서 데이터베이스 분할수를 n 이라고 할 때, m 개의 마스터 오디세우스가 존재하면, 슬레이브 오디세우스의 개수는 $(n \cdot m)$ 이 되므로, 검색에 사용되는 총 프로세스의 수는 $m + (n \cdot m) = m \cdot (n + 1)$ 이 된다. 이를 검색에 사용하는 서버의 성능을 고려하여 서버마다 프로세스를 균등하게 운영하도록 프로세스를 분배한다.

3.2. 데이터베이스 운영 방법

검색 시스템의 운영은 데이터베이스 초기구축을 통하여 처음으로 검색 시스템을 구성하고, 데이터 추가 및 업데이트와 데이터베이스 재구축을 반복하여 검색 시스템을 지속적으로 운영하게 된다. 먼저 제 3.2.1 절에서는 데이터베이스 초기구축 과정에 대해 보이고, 제 3.2.2 절에서는 데이터 추가 및 업데이트 과정에 대해 보인다. 마지막으로 제 3.2.3 절에서는 데이터베이스 재구축 과정에 대해 보인다.

3.2.1. 데이터베이스 초기구축

데이터베이스 초기구축은 검색 서비스를 시작하기 위해 데이터베이스에 데이터를 처음으로 로딩하는 과정이다. 이 과정에서는 많은 양의

문서를 데이터베이스에 로딩해야 하므로, 배치(batch)작업을 통해 빠르게 데이터베이스를 구축할 수 있도록 오디세우스/OOSQL에서 제공하는 벌크 로딩(bulk loading) 기능을 사용한다.

데이터베이스 초기구축 과정은 다음과 같다. 먼저 데이터 수집 서버를 이용하여 데이터를 수집한다. 그 후 수집된 데이터에 대해 문서 및 색인 분할 방법을 적용하여 데이터를 분할하고 포스팅 파일을 생성한다. 데이터 분할 및 포스팅 파일 생성이 완료되면 분할된 데이터와 포스팅 파일을 데이터 로딩 서버로 전송한 후, 데이터 로딩 서버에서 초기 벌크 로딩(initial bulk loading)을 이용하여 각각의 분할된 데이터베이스를 구축한다.

3.2.2. 데이터 추가 및 업데이트

데이터 추가 및 업데이트란 검색 서비스 중 새로운 데이터와 기존 데이터의 변화를 검색 시스템에 반영하기 위해 데이터를 추가하고 업데이트 하는 것을 말한다. 인터넷에 새로운 사이트가 생성되거나 기존 사이트가 업데이트 되는 일은 수없이 많다. 이와 같이 새로운 데이터가 생성되거나 기존 데이터의 업데이트가 있는 경우에는 검색 시스템은 이를 되도록 빨리 반영해주어야 한다. 그러나 이미 대량의 데이터가 존재하는 데이터베이스에 새로운 데이터를 추가하는 것은 매우 비용이 높은 연산이다.

본 논문에서는 이러한 데이터 업데이트 문제를 해결하기 위해 데이터 추가 전용 볼륨을 사용하는 방법을 제안한다. 데이터 추가 전용 볼륨을 사용하여 새로운 데이터를 추가할 때는 이미 대량의 데이터가 저장되어있는 기본 볼륨에 새로운 데이터를 추가하지 않고, 데이터 추가를 위한 별도의 작은 사이즈의 데이터 추가 전용 볼륨을 구성하여 데이터를 추가한다. 그리고 데이터를 업데이트 할 때는 이미 존재하는 데이터를 업데이트 하는 대신 삭제한 후, 데이터 추가 전용 볼륨에 업데이트된 데이터를 새로 추가한다. 검색 시에는 기본 볼륨과 함께 데이터 추가 전용 볼륨에도 검색을 수행한다. 데이터 추가 전용 볼륨을 사용할 경우의 장점은 데이터 추가를 위해 별도의 작은 볼륨을 사용하기 때문에 데이터 추가에 소요되는 시간이 줄어드는 것이다.

병렬 정보검색 시스템의 경우 데이터베이스가 분할되어 있으므로 각 분할된 데이터베이스마다 데이터 추가 볼륨을 둔다. 데이터 추가 시에는 추가할 데이터를 데이터베이스 분할수만큼 분할하여 데이터 추가를 수행함으로써 각 데이터 추가 볼륨이 동일한 크기를 유지하도록 한다. 추가할 데이터를 분할할 때 데이터 특성상 클러스터링(clustering)이 필요한 경우에는 이를 고려하여 데이터를 분할한다. 예를 들어, 본 논문에서 사용한 오디세우스/Parallel-OOSQL 병렬 정보검색 시스템의 경우, 같은 사이트에서 수집한 웹 페이지들은 클러스터링하여 저장하는 것이 검색 성능 향상에 도움이 되므로 데이터 분할 시 사이트 단위로 분할하여 데이터 추가를 수행한다.

데이터 추가 전용 볼륨을 이용하여 데이터를 추가하는 과정은 다음과 같다. 먼저 데이터 수집 서버를 이용하여 추가하려는 데이터를 수집한다. 그 후 수집된 데이터에 대해 문서 및 색인 분할 방법을 적용하여 데이터를 분할하고 포스팅 파일을 생성한다. 그 후, 분할된 데이터와 포스팅 파일을 각각의 슬레이브 검색 서버로 옮긴 후, 데이터 추가 전용 볼륨에 추가 벌크 로딩(append bulk loading)을 수행한다.

3.2.3. 데이터베이스 재구축

데이터베이스 재구축이란 새로운 데이터를 이용하여 전체 데이터베이스를 새로 구축하는 것을 말한다. 검색 시스템 운영 중 데이터의 업데이트를 위해 데이터 추가 전용 볼륨에 지속적으로 데이터를 추가할 경우, 데이터 추가 볼륨에 저장된 데이터의 크기가 커져 추가 벌크 로딩으로 데이터를 추가하는데 많은 시간이 소요된다. 또한, 데이터 추가

전용 볼륨의 여유 공간이 부족해지는 현상이 발생한다. 그러므로 주기적으로 기본 볼륨과 데이터 추가 전용 볼륨에 있는 모든 데이터를 하나로 병합하여 기본 볼륨에 다시 초기상태로 데이터베이스를 재구축하는 작업이 필요하다.

본 논문에서는 검색 서비스를 지속적으로 수행하면서 데이터베이스 재구축이 가능하도록 이중 데이터베이스(dual database)를 구성하는 방법을 제안한다. 이중으로 데이터베이스를 구성하는 것은 검색 서비스를 수행 중인 데이터베이스와 데이터베이스 재구축을 위한 데이터베이스를 분리하여 데이터베이스 재구축이 검색 서비스에 영향을 미치지 않도록 하기 위한 것이다. 이중 데이터베이스를 이용하여 데이터베이스 재구축을 수행할 때의 장점은 데이터베이스 재구축을 수행하면서도 검색 서비스가 중단되지 않는 것이다.

이중 데이터베이스를 이용하여 데이터베이스를 재구축하는 과정은 다음과 같다. 먼저 현재 데이터베이스 내의 데이터를 확인한 후, 모든 데이터에 대하여 데이터 수집 서버를 이용하여 새로 데이터를 수집한다. 즉, 기본 볼륨 내의 데이터 뿐만 아니라 데이터 추가 전용 볼륨 내의 데이터도 모두 새로 수집한다. 새로 수집된 전체 데이터에 대해 문서 및 색인 분할 방법에 따라 데이터를 분할하고 포스팅 파일을 생성한다. 그 후, 분할된 데이터 및 포스팅 파일을 데이터 로딩 서버로 전송하고 데이터 로딩 서버를 이용하여 서비스를 하고 있지 않은 데이터베이스에 초기 벌크 로딩(initial bulk loading)을 통해 새로 데이터베이스를 구축한다. 데이터베이스 구축이 끝나면 새로 구축된 데이터베이스로 전환하여 서비스를 계속한다.

4. 결론

본 논문에서는 대용량의 문서에 대하여 효율적인 정보검색을 지원하기 위한 병렬 정보검색 서비스 시스템의 아키텍처를 제안하였다. 이를 위해, 우선 여러 대의 서버, 디스크 어레이, SAN 스위치를 사용하는 물리적인 기기 구성 방법을 제안하였다. 그리고 빠른 데이터 추가를 위해 데이터 추가 전용 볼륨의 사용을 제안하였고, 데이터베이스 재구축 중에도 지속적인 서비스가 가능하도록 이중으로 데이터베이스를 구성하는 방안을 제안하였다.

참고 문헌

- [1] Zhihong L., Scalable Distributed Architectures for Information Retrieval, Ph. D. Dissertation, Computer Science Department, University of Massachusetts, Amherst, 1999.
- [2] Barroso, L. A., Dean, J., and Holzle, U., "Web Search for a Planet: The Google Cluster Architecture," *IEEE Micro*, Vol. 23, No. 2, pp. 22-28, Mar./Apr. 2003.
- [3] MacFarlane, A., McCann, A., and Robertson, S. E., "PLIERS: A Parallel Information Retrieval System Using MPI," In *Proc. Euro PVM/MPI'99*, Barcelona, Sept. 1999.
- [4] 류재준, 이재길, 이민재, 황규영, "오디세우스/Parallel-OOSQL: 오디세우스 객체 관계형 데이터베이스 관리 시스템을 사용한 병렬 정보 검색 시스템," *한국정보과학회 춘계학술발표회*, pp. 187-189, 2002년 4월.
- [5] Jeong, B. and Omicinski, E., "Inverted File Partitioning Schemes in Multiple Disk Systems," *IEEE Trans. on Parallel and Distributed Systems*, Vol. 6, No. 2, pp. 142-153, Feb. 1995.
- [6] Wilson, S., Managing a Fibre Channel Storage Area Network, Storage Network Management Working Group for Fibre Channel, Storage Networking Industry Association, Inc., Nov. 1998.