
클라우드 환경에서 응용에 따른 일관성 기준의 요구 사항

김치연*

*목포해양대학교

Requirements of Consistency Criteria for Cloud Computing Environments

Chi-yeon Kim*

*Mokpo National Maritime University

E-mail : gegujang2@mmu.ac.kr

요약

클라우드 컴퓨팅은 자원을 사용한만큼 비용을 지불하는 모델을 채용하고 있는 새로운 패러다임이다. 클라이언트들은 개인이 소유하지 않은 자원이라 할지라도 다양한 자원을 사용할 수 있다. 이미 아마존이나 구글, 그리고 마이크로소프트사와 같은 IT 산업 분야의 큰 업체들은 클라우드 컴퓨팅에 대한 많은 응용을 개발하였다. 이 논문에서는 클라우드 컴퓨팅에서 데이터 일관성을 위한 요구조건에 대해 기술한다. 클라우드 컴퓨팅에서 데이터는 중복, 분산, 대규모라는 특징을 가지고 있고, 가용성과 일관성을 동시에 제공하기 어려운 특징이 있다. 이 논문에서는 클라우드 컴퓨팅이 적용가능한 다양한 응용들을 분류하고, 그 응용들에 대한 일관성 기준의 요구 사항에 대해 기술한다. 이 연구를 통해서, 추후 구체적인 일관성 기준을 만드는 데 기반이 될 것으로 생각된다.

ABSTRACT

Cloud computing is a new paradigm that adopts a pay-as-you-go business model. So, clients can use the various resources, although they have not own the resources. Already, three big players of IT industry, namely Amazon, Google and Microsoft, develop the many applications for cloud computing. In this paper, we describe the data consistency requirements for cloud computing. Data characteristics of cloud computing is replicated, distributed and large-scaled. And consistency and availability of data cannot be satisfied simultaneously. In this paper, we categorized the applications of cloud computing, and describe requirements of consistency criteria for applications. With this result, we can make the base of consistency criteria that can be adapted for cloud computing, in the near future.

키워드

cloud computing, utility computing, consistency, serializability

1. 서론

클라우드 컴퓨팅은 개인이나 기업이 직접 보유하지 않은 IT 자원을 빌려 사용할 수 있는 유연한 환경이다. 스마트폰이나 태블릿 PC와 같은 이동성이 보장되는 기기의 보급으로 사용자들은 언제 어디서나 인터넷을 사용하는 것이 가능해졌다. 이러한 이동성에 추가하여, 장치가 직접 보유

하지 않은 플랫폼이나 하부구조, 소프트웨어에 대한 사용 요구도 증가하였는데, 이것을 가능하게 하는 패러다임이 클라우드 컴퓨팅이다.

클라우드 컴퓨팅에서 서비스를 제공하는 방법(배포 모델)은 크게 3가지로 구분된다[1, 2, 3]. SaaS(Software as a Service)는 인터넷을 통해 소프트웨어를 서비스로 제공하는 모델로 고객자원 관리나 웹 콘텐츠 관리에 적합하고, PaaS

(Platform as a Service)는 고객이 소프트웨어를 다운받아 설치하지 않고도 사용할 수 있도록 모든 자원을 공급하는 모델이다. 마지막으로 IaaS(Infrastructure as a Service)는 고객이 필요로 하는 하드웨어를 서비스로 제공하는 모델이다. IaaS에서는 서버, 네트워크 장비, 메모리 등 다양한 하드웨어를 제공하고, 사용자는 사용한만큼 비용을 지불한다.

클라우드 컴퓨팅 환경은 페이스북처럼 시간에 따라 작업 부하의 변동 폭이 큰 응용에서 최대한의 IT 자원을 보유하지 않고도 부족한 자원을 빌려 쓸 수 있는 방법을 제공한다[1]. 따라서 규모의 확장에 유연하게 대처할 수 있어서 다양한 응용에서 채용되고 있다.

클라우드 컴퓨팅 환경에서 관리되는 데이터들은 고도의 분산, 대규모, 중복 데이터라는 특징을 가지고 있다. 이러한 환경에서는 전통적인 데이터베이스 시스템에서 사용되던 트랜잭션의 개념과 ACID 특성을 적용하기 어렵다[4]. 따라서 새로운 기준으로 데이터 일관성을 다뤄야 한다. 이를 위해 이 논문에서는 클라우드 컴퓨팅 환경에서 수행가능한 응용들의 형태를 분류하고, 응용에 따라 적용가능한 일관성 기준들이 만족해야 할 요구 조건들에 대하여 논의하고자 한다.

이 논문의 구성은 다음과 같다. 2장에서는 데이터 일관성을 다룬 기존의 연구들에 대하여 살펴보고, 3장에서 클라우드 컴퓨팅의 응용별 일관성 요구 조건에 대하여 기술한다. 마지막으로 4장에서는 결론을 기술한다.

II. 관련 연구

이 장에서는 클라우드 컴퓨팅 환경에서 적용될 수 있는 데이터 일관성을 다룬 연구들에 대하여 기술한다.

클라우드 컴퓨팅의 구성 요소는 클라이언트 컴퓨터와 분산 서버, 그리고 데이터 센터로 구성된다[5, 6]. 클라이언트는 스마트폰이나 아이폰과 같은 모바일 클라이언트, 단순한 정보 표시 기능만 가진 썬(thin) 클라이언트, 그리고 웹 브라우저를 가진 일반 컴퓨터인 씩(thick) 클라이언트로 분류할 수 있다. 분산 서버는 보안과 가용성의 측면에서 필요하다. 데이터 센터는 클라우드 컴퓨팅에서 가장 핵심적인 부분으로, 클라우드라는 용어가 데이터 센터를 의미한다고 할 수 있다. 즉, 고객들이 사용하는 응용이 설치된 서버를 지칭한다. [그림 1]은 클라우드 컴퓨팅 환경의 구조를 보여준다.

CAP 이론[4]에서는 분할(Partition)이 발생할 수 있는 분산 환경에서는 일관성(Consistency)과 가용성(Availability)을 동시에 달성하기 어렵다는 것을 지적하였다. 네트워크에 분할이 존재할 때, 높은 일관성을 제공하기 위해서는 강한 일관성 기준을 적용해야 하므로 높은 가용성을 달성하기

어렵다. 높은 가용성을 위해서는 일관성을 완화해야 한다. 이 이론은 대부분의 분산, 대규모 환경에서 받아들여지고 있고, 클라우드 컴퓨팅 역시 적용된다. 클라우드 컴퓨팅 환경에서의 많은 응용들은 일관성 보다는 가용성이 더 중요한 응용들이 많다. 따라서 일관성 기준의 완화와 더불어 클라우드 컴퓨팅에 맞는 일관성 기준의 고안이 필요하게 되었다.

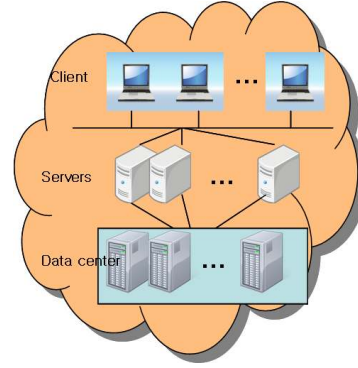


그림 1. 클라우드 컴퓨팅 구조

[7]에서는 클라이언트-서버 구조에 적용될 수 있는 다양한 일관성 기준에 대하여 기술하였다. 가장 강한 기준인 강한(strong) 일관성, 비일관적 윈도우(inconsistent window)를 허용하는 약한(weak) 일관성, 약한 일관성의 변형인 궁극적(eventual) 일관성, 또한 이것의 변형인 캐주얼(casual) 일관성과 세션(session) 일관성 등 다양한 일관성 기준에 대하여 소개하였다. 아래에 주요한 일관성 기준들을 나열하였다.

strong consistency : 갱신이 완료된 후에는 반드시 갱신된 값을 판독한다.

weak consistency : 갱신 후 연속된 접근이 반드시 갱신된 값을 반환함을 보장할 수 없다. 갱신이 완전히 반영될 때까지의 간격을 비일관적 윈도우라 한다.

eventual consistency : 약한 일관성의 변형으로, 객체에 대해 새로운 갱신이 없다면 궁극적으로 모든 접근은 가장 최근의 값을 반환할 것을 보장한다. DNS가 그 예이다.

casual consistency : 데이터를 갱신한 프로세스와 통신하는 프로세스는 갱신된 값을 접근하고, 캐주얼 관계가 없는 프로세스가 접근하는 값은 궁극적 일관성에 따른다.

read-your-write consistency : 캐주얼 일관성의 특별한 경우로, 갱신한 이후에는 항상 갱신된 값만 접근해야 하며, 절대 이전 값을 볼 수 없다.

session consistency : read-your-write의 실용적 버전으로, 하나의 세션 안에서 read-your-write 일관성을 보장한다.

[8]에서는 클라우드 컴퓨팅의 응용 분야를 세 개의 카테고리로 분류하고 가용성과 일관성을 적절하게 유지하기 위하여 각 카테고리에 적합한 일관성 기준을 제시하였다. 카테고리 A는 은행

계정 시스템과 같이 일관성의 위배가 큰 패널티를 가져오는 경우로, 가장 강한 일관성인 직렬성(serializability)이 적용되어야 한다. 카테고리 C는 고객의 구매 선호도나 로깅 데이터와 같이 임시적인 비일관성이 발생하더라도 큰 문제가 되지 않는 경우로, 세션 일관성을 적용할 수 있음을 보였다. 논문에서 주로 다루는 카테고리는 B이다. 상품의 재고 수량(또는 비행기 좌석 예약)과 같은 데이터는 그 수량이 많을 때는 어느 정도의 비일관성이 문제되지 않으나, 재고 수량이 0에 가까워지면 초과 판매를 하지 않기 위해서 강한 일관성을 적용해야 한다. 따라서 세션 일관성과 직렬성이라는 두 개의 일관성 기준을 필요에 따라 적용시킬 필요가 있는 데이터를 카테고리 B로 분류하였다.

III. 일관성 요구 사항

이 장에서는 데이터 일관성의 관점에서 SaaS 형태의 응용들에 대하여 분류하고, 응용별 데이터 일관성의 요구 사항에 대하여 기술한다. 먼저 클라우드 컴퓨팅의 응용들은 다음과 같이 분류할 수 있다.

- 1) 응용 목적에 따른 분류
 - 사무용 어플리케이션
 - 멀티미디어 편집 도구
 - 이메일 패키지
 - 가상 회의 도구
 - 고객 관계, 인사, 프로젝트 관리
 - 웹 사이트 생성 도구
 - 온라인 결제
 - 온라인 쇼핑몰
- 2) 사용 주체에 따른 분류
 - 개인 고객
 - 기업 고객
- 3) 일관성 정도에 따른 분류
 - 강한 일관성을 요구하는 응용
 - 약한 일관성을 요구하는 응용
 - 강약을 동적으로 적용할 수 있는 응용
 - 다중 버전을 필요로 하는 응용

클라우드 응용에서 데이터 일관성이 필요한 데이터들은 공유되거나 중복된 데이터이다. 개인의 블로그 정보와 같이 여러 사람에게 의해 읽혀지지만 대부분의 수정이나 기록이 한 사람에게 의해 수행되는 경우는 일관성을 유지하는 방법보다 네트워크나 신뢰도나 보안 요소가 데이터 일관성을 좌우하는 중요한 요소가 될 수 있다. 따라서 데이터 일관성은 여러 사람에게 의해 동시에 접근될 수 있는 데이터에 대하여 고려한다.

구글 독스[9]와 같은 사무용 어플리케이션이나 포토샵 익스프레스와 같은 멀티미디어 편집 도구의 경우, 하나의 문서를 여러 사람이 공동 작업

한다면 각자의 변경을 제대로 반영하기 위한 일관성 유지 방법이 필요하다. 이 경우에는 생성되는 여러 버전에 대한 관리도 필요하여 일사본 직렬성이 아닌 다중 버전 직렬성에 대한 고려도 필요하다. 이메일 도구의 경우도 개인이 생성하는 데이터이므로 높은 일관성 기준은 필요하지 않다.

가상 회의 도구나, 고객 관계 관리(CRM), 프로젝트 관리 등과 같이 기업에서 주로 사용하는 응용의 경우는 위의 경우보다는 좀 더 강한 일관성 기준을 필요로 한다. 각 응용들은 그룹에 속한 보다 많은 사람에 의해 접근되며 갱신 연산의 수행 빈도가 높으므로 그룹 단위의 일관성을 보장하는 것이 필요하다. 그룹 안에서 수행되는 작업의 순서를 보장할 수 있는 일관성 기준이 필요하다. 이러한 응용의 경우는 데이터 일관성보다도 보안이 더 중요한 요소가 될 수 있다. 그래서 이러한 기업 데이터의 경우는 공용(public) 클라우드가 아닌 하이브리드 클라우드에 저장되는 것이 더 바람직하다.

온라인 결제와 같은 고도의 데이터 정확성을 필요로 하는 응용은 엄격한 기준의 일관성인 직렬성을 적용해야 한다. 웹 응용의 많은 부분을 차지하는 온라인 쇼핑과 같은 응용들은 세션 일관성을 적용하여, 한 세션을 단위로 일관성을 적용하는 것이 가용성을 해치지 않으면서 일관성을 효율적으로 유지하는 방법이다.

사용 주체에 따라 분류한 데이터의 성격은 개인 고객보다는 기업 고객이 다루는 데이터에 더 강한 일관성이 적용되어야 한다.

기존의 전통적인 데이터베이스 시스템에서 사용되던 기준과 비교하면 클라우드 컴퓨팅에서의 일관성 기준은 완화된 일관성 기준을 필요로 하는 응용들이 더 많다는 것을 알 수 있다. 이는 엄격한 기준을 적용하기 위해 제약이 강한 일관성 기준을 적용하는 것은, 데이터의 가용성을 감소시키는 원인이 되므로 클라우드 환경의 특징에는 적합하지 않기 때문이다.

지금까지의 클라우드 응용들을 살펴보면, 가장 완화된 일관성으로는 궁극적 일관성과 제약이 가장 강한 직렬성 및 다중 버전에 이르기까지 응용의 성격에 따라 다양한 일관성 기준을 적용할 수 있음을 알 수 있다.

V. 결 론

이 논문에서는 최근 새롭게 나타나고 있는 클라우드 컴퓨팅의 특징에 대하여 기술하고, 클라우드 응용에 대한 데이터 일관성 문제에 대하여 기술하였다.

클라우드 컴퓨팅에서의 데이터는 전통적인 데이터베이스 시스템처럼 다수의 사용자에 의해 공유되거나 여러 곳에 중복 저장되는 데이터들이 많다. 따라서 이 데이터들의 정확성을 보장하기

위해서는 데이터 일관성을 위한 기준들이 필요하다. 하지만 이를 위해 전통적인 데이터베이스 시스템의 제약이 강한 일관성 기준을 적용하는 것은 바람직하지 않다.

이 논문에서는 클라우드 컴퓨팅에서 수행될 수 있는 응용들을 분류하고, 응용별 필요한 일관성 요구 사항에 대하여 기술하였다. 하나의 사본에 대한 일관성 기준 뿐 아니라 다중 버전을 유지해야 하는 응용들을 보였고, 많은 응용들에서 완화된 일관성을 적용할 필요가 있음을 지적하였다.

클라우드 컴퓨팅과 같은 대규모의 분산된 노드들이 하나의 시스템을 구성하는 경우, 데이터의 일관성과 더불어 데이터 가용성의 제공은 매우 중요하다. 하지만 언제든지 네트워크 분할이 발생할 수 있는 환경에서 두 가지 특징을 모두 만족시키는 것은 어려운 일이다. 사용자의 입장에서는 가용성 또한 매우 중요한 요소이다. 따라서 추후 연구에서는 클라우드 컴퓨팅에서 데이터 일관성과 가용성을 동시에 고려할 수 있는 방법에 대하여 연구하고자 한다.

참고문헌

- [1] Michael Armbrust, Armando Fox, Rean Griffith, Anthony D. Joseph, Randy Katz, Andy Konwinski, Gunho Lee, David Patterson, Ariel Rabkin, Ion Stoica, Matei Zaharia, "A View of Cloud Computing," Communications of the ACM, Volume 53 Issue 4, April 2010.
- [2] 이경하, 최현식, 정연돈, "클라우드 컴퓨팅에서의 대용량 데이터 처리와 관리 기법에 관한 조사," 정보과학회논문지, 데이터베이스 제38권, 제2호, pp.104-125, 2011.
- [3] 테이코산업연구소, "클라우드 컴퓨팅, 차세대 컴퓨팅 기술/시장 동향과 사업 전략," 테이코산업 연구소, pp. , 2010.
- [4] Gilbert, S. and Lynch, N., "Brewer's Conjecture and Feasibility of Consistent, Available, Partition-tolerant Web Services Towards Robust Distributed Systems," ACM SIGACT News 33, 2, pp. 40-44, 2002.
- [5] 앤서니 T 벨트, 토비 J 벨트, 로버트 엘센피터, "미래코드 클라우드 컴퓨팅," 전자신문사, pp. 28-33, 2011.
- [6] Donald Kossmann, Tim Kraska, Simon Loesing, "An Evaluation of Alternative Architecture for Transaction Processing in the Cloud," Proceedings of the 2010 international conference on Management of data, pp. 579-590, 2010.
- [7] Werner Vogels, "Eventually Consistent," Communications of ACM, 52(1):40-44, 2009.
- [8] Tim Kraska, Martin Hentschel, Gustavo Alonso, Donald Kossmann, "Consistency Rationing in the Cloud : Pay only when it matters," Proceedings of the VLDB, Volume 2, August 2009.
- [9] 크리스토퍼 버넷, "클라우드 컴퓨팅," 미래의 창, pp. 66-99, 2010.