콘텐츠 중심의 네트워크 기술

Content Centric Networking Technology

네트워크 기술의 미래 전망 특집

김창임 (J.I. Kim) 인터넷미래서비스연구팀 책임연구원
정희영 (H.Y. Jung) 인터넷미래서비스연구팀 책임연구원
박우구 (W.G. Park) 인터넷미래서비스연구팀 팀장

목  차

Ⅰ. 서론
Ⅱ. 콘텐츠 중심의 네트워크
Ⅲ. 결론

* 본 연구는 지식경제부 및 한국산업기술평가관리원의 산업협력기술개발사업(정보통신)의
  일환으로 수행하였음 (10036245, 미래인터넷에서의 아동보호 및 네트워크 탑続け 지원
  구조 연구)

1990년 초 웹의 등장과 함께 인터넷의 전세계 확산으로 'Everything over Internet'은 시대적
요구가 되었다. 그러나 현재인터넷은 시대적 요구를 수용하기에는 인터넷 구조에 근본적인 문제
가 제기되고 있어, 새로운 혁신적 접근(clean-slate approach) 방식의 미래인터넷에 대한 연구
가 세계적으로 활발히 진행되고 있다. 본 고에서는 이러한 혁신적 미래인터넷 연구 기술의 하나
인 콘텐츠 중심의 네트워킹(content centric networking) 기술에 대해 분석한다. 콘텐츠 중심의
네트워크는 현재 빠른 속도로 증가하고 있는 데이터 서비스에 대해 보다 콘텐츠 중심의 전송 방
식을 제공함으로써 더 빠르고 더 네트워크 공격에 강한 서비스를 제공하고자 하는 기술이다.
1. 서론


모던 인터넷 데이터 서비스는 웹 페이지(예, 'Google', 'Naver')와 같이 수백만의 사용자가 동일한 서비스를 요구하는 동일한 데이터라는 특징을 가진다. 이에 비해, 현재의 인터넷 전송 방법은 데이터가 무엇인가도 상관없이 송수신 호스트의 IP 주소를 이용하여 서비스를 제공함으로 동일한 데이터들이 네트워크상에 사용자 수만큼 반복 전송되는 방식으로 동작하고 있는 비효율성이 있다. 이러한 데이터 반복 전송의 비효율성을 피할 수 있도록, 콘텐츠 중심 기술은 데이터 배포(dissemination) 개념을 도입하여 사용할 것을 제안한다. 즉, 제안하는 방식은, 현재 인터넷의 IP 주소를 사용하지 않고 대신, 데이터 이름을 사용하여 네트워크에서 데이터 전달을 수행한다. 또한, 데이터 전송 채널과 데이터 저장소(container)를 보안하는 종래의 보안 방식에서 벗어나, 데이터 자체를 보안하는 새로운 방식을 제안한다.

콘텐츠 중심의 서비스 구조와 모던 인터넷에 비교하면 다음과 같은 특성을 가지고 있다.

먼저, 모던 인터넷은 데이터에 대한 송신자와 수신자 가 존재하고, 송신자가 혼잡 제어(congestion control), 호흡 제어(flow control) 등의 네트워크 제어를 수행하고, 일대일 대화 인증을 통해서 데이터를 전송하는 채널과 데이터 전달에 대해 보안을 수행하고, 수신자의 호스트 이름을 네트워크 제어의 해더에 IP로 포함하여 네트워크 라우터는 수신자 호스트로 데이터 콘텐츠가 무엇인지 모르고 라우팅을 수행하는 구조를 가진다[2]. 반면, 콘텐츠 중심 기술은 지원하는 네트워크는 송수신자 대신 데이터를 생성한 콘텐츠 생성자(publisher)와 데이터를 수신하여 소비하는 콘텐츠 소비자(consumer)의 개념을 도입하여 네트워크는 IP 주소 대신 콘텐츠 이름을 사용하여 라우팅을 수행하며, 라우터는 필요한 경우 특정 데이터 를 저장하고, 데이터를 요구하는 인근의 새로운 소비자들에게 데이터를 배포하는 방식을 사용한다.

콘텐츠 중심의 연구의 예로 CCN, DONA, PSIRP, NetInf를 들 수 있다.

CCN 기술은 PARC에서 근무하는 Van Jacobson 을 중심으로, DONA는 FIND 프로젝트에서 Teemu Koponen을 중심으로, PSIRP는 FP7 프로젝트, Net-
Inf는 4WARD의 Subproject로 연구가 진행되고 있다.

콘텐츠 중심 기술들에서 제안하는 방식은 현재의 인터넷 서비스가 송신자의 일대일 통신으로 반복적으로 전송과 보안이 수행되는 방식에 비해, 같은 데이터의 반복 전송을 피할 수 있는 기회, 끝된 서비스 제공의 기회, 송신자 관점에서 반복적인 보안을 피할 수 있는 기회를 제공하는 장점을 가지고 있어 최근 미래 인터넷을 위한 유망한 기술로 주목을 받고 있다.

본 교에서는 콘텐츠 중심 기술의 연구들 중에서 Van Jacobson에 의해 제안된 CCN 기술의 주요 내용을 분석한다.

II. 콘텐츠 중심의 네트워크

1. CCN 프로토콜

CCN은 현재인터넷의 구조가 과거 고가의 제한된 네트워크 자원을 공유하며 소량의 데이터를 교환하는 것으로 설계되었다는 것을 직시하고, 현재 낮은 비용으로 제공 가능한 네트워크 자원을 이용하여 대량의 데이터 서비스를 효율적으로 지원하는 새로운 인터넷 구조의 전이가 필요함을 주장하고 있다[3].

CCN 기술이 제안하는 배포의 방식은, 현재인터넷 IP 주소를 이용하는 “어느 곳(where)” 대신, 데이터 이름을 사용하여, 즉, “무엇을(what)”을 사용하여 네트워크에서 데이터 전달을 수행하고, 데이터 자체를 보안하는 새로운 방식이다.

(그림 2)는 현재인터넷과 CCN 기술이 제안하는 프로토콜 구조를 보여준다. (그림 3)에서와 같이 시스템 기본 요구사항이 현재인터넷에서 IP 프로토콜의 사용이 반면, CCN 기술에서는 이론으로 떨어지는 콘텐츠 청크(chunk), 정책(strategy), 보안 계층의 사용이라는 차이를 가진다.

(그림 3) CCN 데이터 배포의 예

현재인터넷의 기술의 경우 송신 호스트가 패킷의 IP 헤더(header)에 송신한 호스트의 IP 주소를 명시하여 데이터를 전송한다. 반면, CCN 기술은 데이터
를 제작한 생성자와 소비자가 IP 하더 데이터를 사용한다. 즉, 소비자는 데이터를 요구하는 패킷에 콘텐츠 이름을 적어서 방송하고, 라우터는 데이터의 해당 요구 패킷을 저장하며, 데이터를 갖고 있는 노드가 콘텐츠를 요구하는 패킷을 수신하면, 콘텐츠 요구 패킷을 전송한 노드에 데이터를 응답으로 전송한다. CCN 차수의 콘텐츠 생성자가 반드시 데이터를 전송한다는 제한을 두지 않고, 데이터를 갖고 있는 입의의 네트워크를 배포할 수 있게 함으로써, 빠른 서비스와 같은 데이터가 네트워크에서 전송되는 횟수를 줄이는 장점을 가진다.

(그림 3)은 CCN에서 client 1과 client 2가 Google 데이터를 요구할 때, client 1이 먼저 데이터를 요구한 경우, client 1이 인근에 있는 client 2가 현재 인터넷 끊으므로 더 빠르고, 더 작은 대역폭을 사용하여 데이터를 수신할 수 있는 장점을 가진다.[4]

(그림 3a) 콘텐츠 생성자와 소비자 사이의 모든 노드들이 콘텐츠를 가지고 있지 않는 상황에서, (그림 3b) client 1이 데이터 요구 패킷을 방송하면, 데이터 요구 패킷이 Google까지 전송되고, (그림 3c) 노드 A, B, D, H, client 1에게 데이터가 응답되고, (그림 3d) 노드 A, B, D, H는 응답된 데이터를 저장하고, (그림 3e) client 2가 방송한 동일한 데이터 요구 패킷은 노드 D까지 전송되어, (그림 3f) 요구된 동일 데이터가 노드 D로부터 client 2에게 응답으로 전송될 수 있는 장점을 가진다.

2. CCN 메시지 패킷

CCN은 interest packet과 data packet 두 개의 메시지 패킷을 가지고 있다.[5]. interest packet은 콘텐츠 생성자의 데이터 요청 패킷이고, data packet은 요청된 데이터를 갖고 있는 노드가 요구된 데이터를 응답으로 전송하는 패킷이다.

(그림 4)에서 interest packet과 data packet의 구조를 보여준다. interest packet은 콘텐츠 네임(content name), 선택자(selector), 랜덤 임시 변수(nonce)로 구성되고[6], data packet은 콘텐츠 네임, 서명(signature), 사인 정보(signed info), 데이터(data)로 구성한다[7].

interest packet의 선택자와 data packet은 공개키 다이제스트(public key digest)를 포함한다. 공개키 다이제스트는 요청한 패킷에 대해 응답 데이터가 제대로 수신되었는지를 확인하는 유도로 사용한다.

Interest packet의 선택자 항목은 interest packet의 콘텐츠 네임에 기술되지 않은 네임 항목을 자세히 기술하는 항목, 주어진 interest packet의 data packet의 여러 CS에 일치하는 경우, 선풍화되는 CS에 대한 정보를 제공하는 항목, 보안 등의 관리 목적으로 제공되는 정보 항목으로 구성된다.

랜덤 임시 변수는 랜덤하게 생성된 byte string으로서 interest packet이 방송되므로 서로 다른 정보를 통해 루프를 형성할 수 있으며, interest packet 을 식별하여 루프 형성을 막을 데 사용한다.

Data packet의 사인 정보는 공개키의 다이제스트, 타임 스탠드, 데이터 존재 여부와 암호화 여부 등의 데이터 타입, 데이터의 버전에 따라 의미가 있는 데이터 유호 기간, 데이터가 여러 개의 세그먼트로 구성된 경우 최종 불록 식별자의 정보, 키 위치로 구성된다.

키 위치는 데이터 암호화 키가 어디에 있는가를

김정일 외 / 콘텐츠 중심의 네트워크 기술
알려주고, 생성자가 모든 소비자가 다른 방법으로 키를 찾을 수 있다면 생각할 수 있으며, 또는 키 자체가 포함되어 있을 수 있다.

3. CCN 포워딩 엔진

CCN은 CS, PIT, FIB로 포워딩 엔진을 구성한다. (그림 5)에서 CCN 포워딩 엔진의 구성요소는 볼 수 있다. 포워딩 엔진은 방에 페이스(face)는 CCN과 IP 인터페이스와의 차이, 즉 CCN은 다중 연결을 지원하나, IP는 다중 연결을 지원하지 못하는 차이를 포함하기 위해 사용한다.

CS는 라우터의 버퍼 메모리와 같이, 데이터 콘텐츠와 콘텐츠 리스트를 저장하고 있으며, 효율적인 배포를 위해 LRU 또는 LFU 등을 교환 정책으로 운영한다. 1)

CCN 기술은 IP 주소 대신 interest packet의 경로를 관리하여 라우팅을 수행한다. 즉, PIT는 수신된 interest packet에 대한 응답으로 콘텐츠를 전송할 수 있는 노드로 interest packet을 전달 후, 어느 페이지로 interest packet이 입력되고 출력되었는지를 기록하여 관리한다. 각각 PIT 항목을 breadcrumb로 하여, interest packet이 전송된 노드에 PIT 항목이 생성되는 것을 breadcrumb가 남겨진 것으로 하고, 콘텐츠를 가지고 있는 interest packet이 전송되면, 콘텐츠를 가진 노드로부터 interest packet을 생성한 노드까지 interest packet의 경로의 역방향으로 콘텐츠가 전달되면서, interest packet이 삭제된다. 즉, interest packet의 경로에 breadcrumb를 남겨두어, 응답 data packet이 breadcrumb의 경로가 생성된 반대 방향으로 breadcrumb를 소비하면서 소비자에게 데이터를 전달한다. PIT는 일정 시간 안에 데이터가 도착하지 않으면, interest packet을 재전송해야 한다. 재전송은 책임은 소비자에게 있다. 재전송 후에도 일정 시간 후에 데이터가 도착하지 않으면, 해당 interest packet의 항목을 삭제한다.

각각의 노드는 정책 프로토콜을 두어, 재전송과 데이터 흐름, 혼잡 제어 등을 노드(호) 단위로 동작한다. FIB는 interest packet을 포워딩하는 데 사용한다. FIB는 콘텐츠 생성자가 근처 CCN 코어에 등록 등록(register operation)을 수행하여 FIB 항목이 생성된다. 광고 에이전트(announcement agents)는 정책에 따라 FIB 항목을 광고하는 범위를 정한다. 즉, 보안과 대역폭 효율 등을 고려하여 정책을 결정하고, 정책 프로토콜은 결정된 정책에 따라 FIB의 리스트를 생성하고 삭제를 수행한다.

Interest packet과 data packet이 노드에 도착하면 CS, PIT, FIB 순으로 longest match lookup이 수행된다.

CS의 네임과 interest packet의 이름이 일치하면, 데이터가 interest packet이 입력된 페이지로 전달되고, interest packet은 삭제된다.

Interest packet의 이름이 PIT에서 검색되면, PIT 항목에 interest packet의 입력 페이지가 추가되는 데, 이는 이미 interest packet이 발생하였고 데이터를 기다리는 동안 새로운 소비자가 등장하여 발생한 것으로 해석하는 것이다. Interest packet의 랜덤 임시 변수로 새로운 소비자인지 확인할 수 있다. Interest packet의 네임이 FIB에서 검색되면, interest

1) 현재인터넷은 MRU(Most Recently Used) 교환 정책(replacement policy), CCN은 다른 LRU(Least Recently Used) 또는 LFU(Least Frequent Used) 교환 정책으로 운영한다.
packet을 방송으로 포워딩하고, FIB 항목에서 삭제하고, PIT 항목을 추가한다. Interest packet의 내임이 FIB에서도 검색되지 않으면, 해당 데이터는 처리되지 못하는 서비스이므로, interest packet은 삭제된다.

Data packet이 노드에 도착하면, CS, PIT, 포워딩 정보 베이스 순으로 longest match lookup이 수행된다. 데이터 내임이 CS의 내임에 일치하면, 도착한 데이터가 존재하는 데이터와 같이므로 삭제한다. 데이터 내임이 FIB에서 검색되었다는 것은 데이터 내임이 PIT에서 검색되지 않았다는 것을 뜻하며, 도착한 데이터는 요구되지 않은 데이터이므로 삭제한다. 데이터 내임이 PIT에서 검색되면, 요청한 데이터가 도착한 것이므로, CS에 데이터를 저장하고, interest packet이 도착한 페이지로 데이터를 전송한다.

4. CCN 라우팅

CCN 포워딩 모델은 IP 포워딩 모델을 포함하는 더 큰 집합(superset)이며, IP와 마찬가지로 데이터 내임은 계층적이고, longest match lookup을 수행하는 라우팅을 수행하므로, IP에 대해 잘 동작하는 임의의 라우팅 방식도 CCN 기술에 대해 잘 동작할 것으로 예상된다. 그러나, IP 헤더에 비하여 CCN 헤더가 매우 크고, 서비스마다 내임의 길이가 다르다는 차이점이 있다.

또한, IP 계층과 CCN 기술의 FIB가 유사하여, IP FIB들을 생성하는 분산 라우팅 장비가 쉽게 CCN FIB들을 생성해 쉽게 적용할 수 있어, 개념적으로 CCN 기술은 존재하는 종래의 라우팅을 사용할 수 있으며, 존재하는 기반 시설에 CCN 라우팅을 추가적으로 배치하여 IP와 양립할 수 있다.

또한, CCN prefix는 무엇을 의미하고, IP prefix는 어디를 의미하기에 그 내용은 다르지만, 도메인 내의 라우팅(IS-IS와 OSPF)에서 IP는 TLV 방식을 사용하는데, CCN도 TLV 방식을 사용할 수 있어 총론의 도메인 내부의 변화 없이 CCN 기술이 수용될 수 있다.

CCN은 강한 정보 보호 모델을 제공하므로, CCN을 사용하는 것은 라우팅 기반시설 보호가 거의 자동으로 되게 해주는 장점을 가진다고 할 수 있다.

5. CCN 보안

현재 인터넷은 어느 호스트, 어떤 채널을 통해서 어떤 데이터가 전달되는가에 대해 보안을 수행한다. 이러한 보안 방식은 중요한 문제로 인식되었다. 어떠한 방식으로도 보안이 안정된 것 같지만, 사라지는 일시적 보안 관계이므로, 최소한 데이터를 검색하여 수신한 호스트가 다른 호스트에게 영향을 주지 않는다. 즉, 연결마다 신뢰가 보장되므로, 같은 서비스 데이터에 관심이 있는 호스트들은 항상 생성자와 신뢰 관계를 맺어야 하므로 생성자 관계에서 보안 절차가 반복된다.

최근 많은 연구들이 생성자가 반복적으로 보안 절차를 수행하지 않도록 호스트 내에서 데이터 자체를 증명(보안)하는 구조로 변경해야한다고 제안하고 있다. 즉, 자세 인증 네임(self-certifying name)을 암호화된 데이터의 암호화된 다이제스트 또는 디지털 서명(digitally sign) 용도의 생성자 키로부터 구성할 것을 제안하였다. 자체 인증 네임은 위치에 상관없이 세계적으로 고유하고, 간단하고, 완전히 자동적으로 생성되는 장점이 있다. 그러나 자체 인증 네임은 비계층적(Flat)이라 특별 네임에 해당하는 데이터의 주변 복사본을 효율적으로 데이터를 검색하여 수집하는 방법을 만드는 것이 어렵고 또한, 자체 인증 네임은 이메일 주소 또는 호스트 내역과 같은
직관에 의한 네임 처리가 어렵다. 비제충적이고, 이해하기 어려운 (opaque) 네임 시스템은 DNS와 같은 간접 구조를 필요로 하여 자체 인증 네임 구조도 본래 콘텐츠 보안 문제를 갖는다. 즉, 만약 메이핑이 보안이 이루어지지 않는 경우 호스트는 잘못된 데이터를 수신한다.

CCN은 이러한 자체 인증 네임의 문제가 없는 네임과 콘텐츠의 관계를 인증하는 방식을 제안하였다. 생성자가 콘텐츠를 생성한 후, N은 콘텐츠 C의 네임을 인증하는 과정으로, 즉 P는 N과 C 사이의 메이핑에 대해 디지털 서명을 한다. 이는 다음과 같이 표현할 수 있다.

\[ M(N;P;C) = (N:C; SignP(N:C)) \]

데이터 소비자는 임의의 네임 N에 대해 데이터를 수신하고, 콘텐츠를 확인하는 고유한 식별자에 대한 정보 또는 생성자의 식별자를 미리 알고 있지 않기도, 콘텐츠의 네임 관계의 메이핑을 인증한다. 이러한 방식은 무결성, 출처, 적절성을 가진 요구사항을 만족한다.

CCN 라우팅의 두 가지 특징이 네트워크 공격으로부터 예방하고 보호한다. 첫째, 소비자의 interest packet에 대한 응답으로 data packet으로 전송되므로 DDoS 공격을 쉽게 허용하지 않는다. 대응되는 interest packet이 없는 data packet은 삭제되고, interest packet을 이용하여 flooding attack을 시도하는 경우, 광범위한 영역에서 많은 수의 interest packet을 전송하면 다수의 interest packet이 라우터에서 하나로 결합되어, 하나의 데이터만 응답되어 DDoS 공격이 쉽지 않다. CCN은 모든 콘텐츠의 라우팅과 정책 정보를 포함하는 모든 콘텐츠를 인증하고, 데이터가 변조되지 않게 보호한다. CCN 메시지는 콘텐츠에서 대해서만 가시하므로, 특정 호스트를 묘표로 악의의 메시지를 보내는 것이 어렵다.

두번째로, 라우팅은 정책 기반 라우팅 (policy-based routing)으로 interest packet의 포워딩 도메인 영역을 조절할 수 있어 네트워크 공격이 어렵다. 즉, interest packet을 특정 도메인에만 보내추어 특정 도메인 내에서만 data packet이 응답하여 외부 도메인에서 악의의 메시지를 보내는 것을 차단할 수 있다.

CCN은 특정 서명 알고리즘들을 사용해서 서명을 생성해야 한다는 제한을 갖고 있지 않다. 서명을 생성하는 방법은 각각의 블록에 표준 공개키 알고리즘들을 사용하여 디지털 서명을 하는 방법과 여러 개의 블록에 대해 서명을 하는 두 가지 방법 모두 이용할 수 있다. 콘텐츠의 생성자가 서명의 길이, 시간 지연, 서명 계산 비용 등을 고려하여 알고리즘들을 선택한다. 앞에서 설명하였듯이, 서명은 다이제스트 알고리즘, 주인 (witness), 서명비트로 구성될 수 있다. 주인은 여러 개의 블록에 대해 서명하여 시그니처를 생성한 경우, 시그니처를 확인할 때 필요한 정보이다.

CCN 기술은 특정 콘텐츠 압호화 알고리즘들을 사용해야 한다는 제한을 갖고 있지 않다. 따라서 콘텐츠 소비자와 생성자가 합의한 임의의 알고리즘들을 사용할 수 있다.

### III. 결론

본 논문에서는 미디어넷 핵심 연구 분야의 하나로 인정되고 있는 콘텐츠 중심 기술 중에서 CCN 기술을 소개하고, 고속으로 증가하고 있는 데이터 서비스에 대해, 현재인터넷 보다 대역폭 효율을 향상시키고, 더 강화된 보안 기술을 제공할 수 있음을 소개하였다.

이러한 목적을 달성하기 위해, CCN 기술은 네트워크 대역폭 효율 향상을 위해, IP 대신 데이터의 이
콘텐츠 중심의 네트워크(CCN) 기술: 인터넷에서 데이터 전송을 IP 주소의 위치(where) 개념을 벗어나, 콘텐츠 이름(label)의 개념으로 수행하고, 콘텐츠 자체에 보안을 수행하는 기술

参考文献