

가상화와 컴퓨터 네트워크의 활용 1 : 가상화와 데이터센터

최영락 휴레이포티즈브 선임연구원 & 오픈플로우코리아 기술매니저

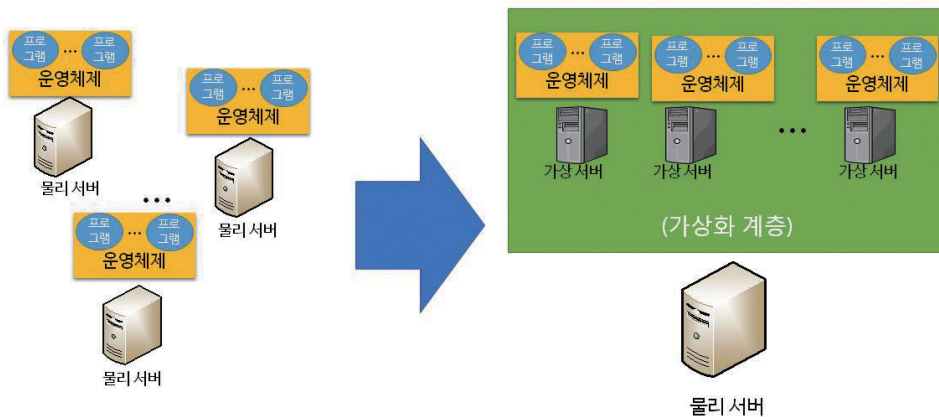
- 연재 목록 -

- | | |
|------------------------|-----------------------|
| 1. 가상화와 데이터센터 | 6. 네트워크 스토리지 기술 - (2) |
| 2. 데이터센터 네트워크 구성 - (1) | 7. 활용 사례 - (1) |
| 3. 데이터센터 네트워크 구성 - (2) | 8. 활용 사례 - (2) |
| 4. 가상화와 네트워크 스토리지 | 9. 활용 사례 - (3) |
| 5. 네트워크 스토리지 기술 - (1) | |

지난 '컴퓨터 네트워크의 이해 및 활용' 연재를 통해서 IP 주소를 기반으로 통신이 이루어지는 컴퓨터 네트워크에 대해 전체적으로 살펴보았습니다. IP 기반의 컴퓨터 네트워크 기술이 확산되면서 관련하여 빠질 수 없는 관련 분야로 '가상화'가 있습니다. 가상화를 기반으로 기존 서버-클라이언트 중심의 인프라는 데이터센터 또는 클라우드 환경으로 변화하고 있으며, 네트워크 및 스토리지 가상화 또한 점차 적용되는 추세에 있습니다. 본 연재에서는 가상화의 발전에 따른 관련 기반 인프라와 컴퓨터 네트워크망이 어떻게 변화하고 있는지, 그리고 이와 관련된 기술 및 활용 방식을 설명하고자 합니다. 첫 연재에서는 '가상화와 데이터센터'를 제목으로 하여 구체적인 내용을 살펴보겠습니다.

가상화란, 그리고 서버 가상화

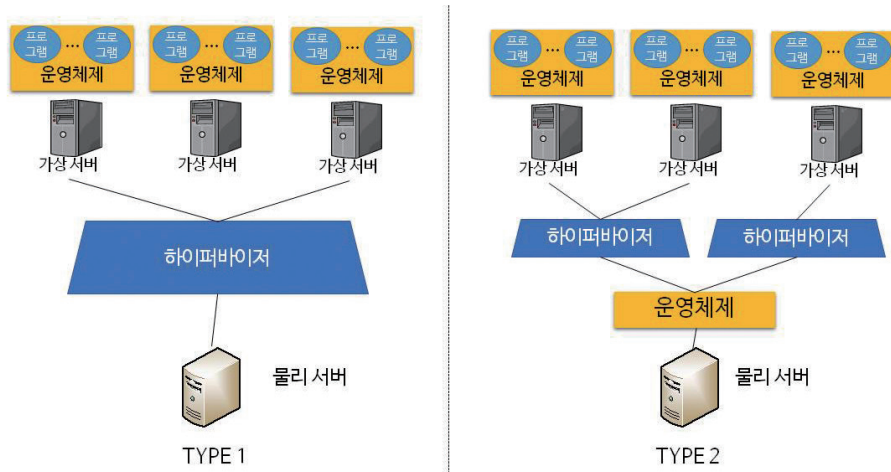
가상화에 대해 이야기할 때, 일반적으로는 [그림 1]과 같은 '서버 가상화'를 예로 들어 설명합니다. 가상화라는 기술이 오늘날처럼 사용되기 이전에는 물리적인 서버를 각각 구입하고, 그 위에 운영체제를 설치하고 프로그램을 구동하는 식으로 사용했었습니다. 실제로, 이와 같은 인프라는 여전히 방송 서버실에서 볼 수 있는 환경입니다. 그런데 가상화를 적용하면 기존에 사용하던 물리 서버를 '가상 서버'라는 단위로 만들고, 이를 성능이 매우 좋은 하나의 물리 서버 위에 구동이 가능합니다. 이때, 물리적으로 보면 하나의 물리 서버가 존재하지만, 실제 해당 1대의 물리 서버에는 기존에 운영체제 및 프로그램이 설치되던 3대의 각 서버 단위로 동작하는 3대의 가상 서버가 있으며, 해당 가상 서버 위에 기존에 수용하던 운영체제 및 프로그램들이 실행됩니다.



[그림 1] 서버 가상화

가상화를 한마디로 정의하자면, '물리적, 논리적 자원을 추상화시켜 실제와 같은 자원으로 사용 및 관리 가능하도록 하는 기반 환경 및 기술'이라고 이야기할 수 있습니다. 그리고 [그림 1]에서 예를 들었던 물리 서버를 가상 서버로 동작시키는 가상화를 '서버 가상화'라고 합니다. 이때, 물리 서버를 가상화 환경으로 제공하기 위해, 1대의 서버에서 여러 가상 머신 및 운영체제를 동시에 실행하기 위한 논리적인 플랫폼, 즉 서버 가상화를 실현하는 플랫폼을 '하이퍼바이저(Hypervisor)'라고 이야기합니다.

하이퍼바이저는 크게 TYPE1과 TYPE2 2가지 방식으로 구분됩니다. TYPE1 방식은 물리 서버 위에 하이퍼바이저가 직접 설치되고, 그 위에서 가상 서버 및 운영체제가 실행되는 방식이고, TYPE2 방식은 운영체제 위에 하이퍼바이저가 설치되고, 그 위에서 가상 서버 및 운영체제가 실행되는 방식입니다. TYPE1 하이퍼바이저는 Native 또는 Bare-metal 방식의 하이퍼바이저라고도 하며, 하드웨어 계층 바로 위에서 하이퍼바이저가 동작하여 가상 서버를 제공합니다. 물리 서버의 성능을 최대한 발휘하면서 가상 서버의 성능을 제공하기에, 성능이 많이 감소되지 않으면서 가상 서버를 제공하지만, 물리 서버 바로 위에 설치된 하이퍼바이저를 직접적으로 관리하기에는 어려움이 있습니다. 반면, TYPE2 하이퍼바이저는 Hosted 하이퍼바이저라고도 하며, 기존 물리 서버에 설치된 운영체제 위에 하이퍼바이저가 설치되어 동작하므로, 하나의 물리 서버 위에 여러 종류의 하이퍼바이저를 실행할 수 있습니다. 또한, 하이퍼바이저를 직접 실행하기에 편리한 인터페이스 화면을 제공하지만, 반면 이에 따른 성능 감소가 생긴다는 단점이 있습니다.

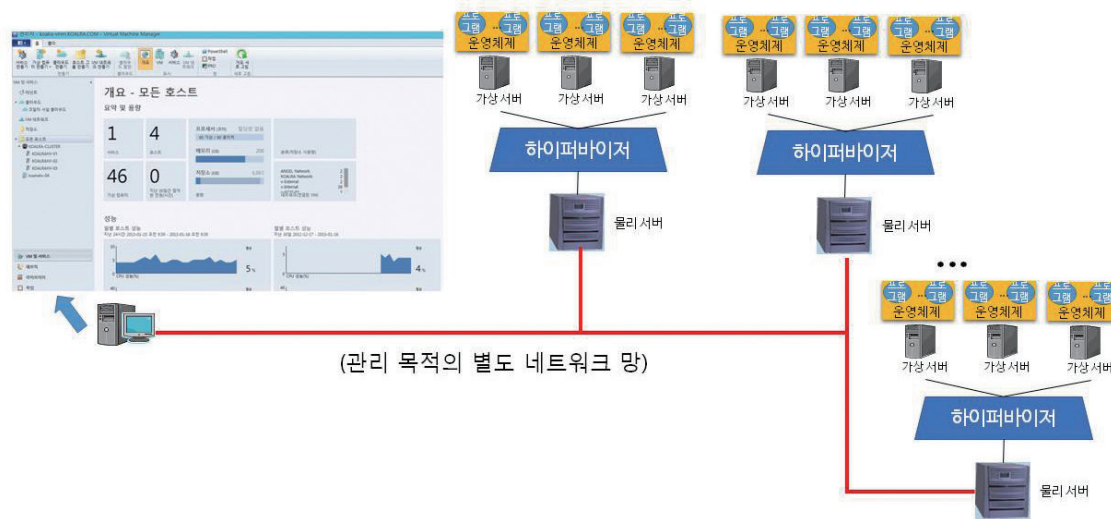


[그림 2] 하이퍼바이저 구분 : TYPE1과 TYPE2

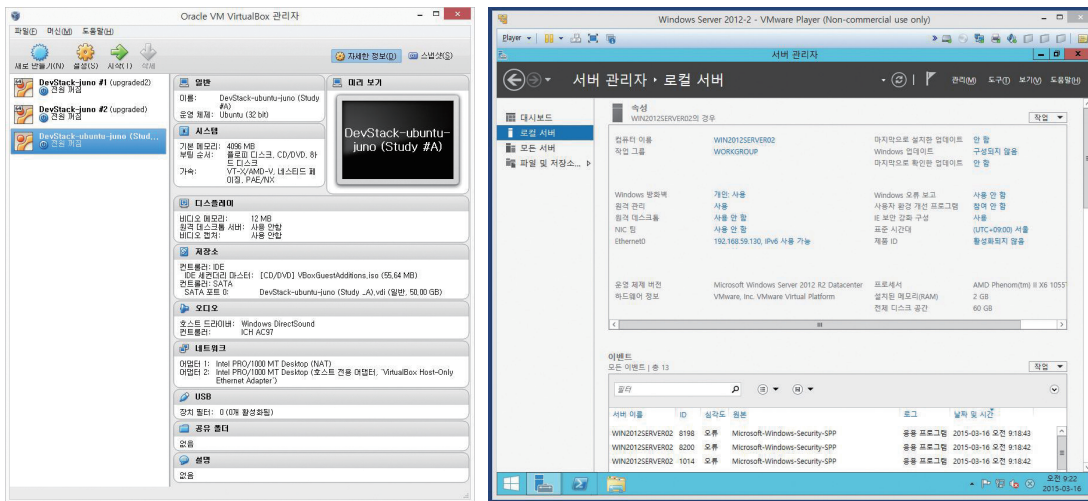
하이퍼바이저 유형	주요 제품
TYPE1	VMware ESX/ESXi, Microsoft Hyper-V, Citrix XenServer, KVM
TYPE2	VMware Workstation, VMware Player, VirtualBox, QEMU, Microsoft Virtual PC, Parallels Desktop

[표 1] 하이퍼바이저 유형에 따른 제품

실제로 TYPE1 하이퍼바이저는 성능에 대한 이점이 있어 서버 인프라 환경을 구성할 때 많이 사용됩니다. 단, 하이퍼바이저를 관리하기 위해 [그림 3]과 같이 별도의 관리용 네트워크 망을 구성하여 특정 컴퓨터에서 하이퍼바이저에 접속하여 관리하는 형태로 사용합니다. 반면, TYPE2 하이퍼바이저는 주로 가정용 컴퓨터 또는 운영체제가 이미 설치된 컴퓨터에서 여러 운영체제를 설치하여 테스트하는 목적으로 사용 가능합니다. 대표적인 TYPE1 및 TYPE2 하이퍼바이저 제품은 [표 1]과 같으며, [그림 4]는 무료로 사용할 수 있는 TYPE2 하이퍼바이저에 해당하는 VirtualBox와 VMware Player 화면입니다. [그림 4]의 왼쪽 VirtualBox를 살펴보면 왼쪽 탭에 선택된 가상 머신에 대한 정보가 오른쪽에 나타나 있는데, 메모리(RAM), 그래픽카드, 오디오, 네트워크, USB, 하드디스크 등 해당 가상 머신을 이루는 실제 많은 구성 요소들이 있음을 볼 수 있습니다. 또한 오른쪽에 실행 중인 VMware Player에 대한 스크린샷과 같이, 창 모드로 해당 가상 머신의 화면을 볼 수 있어 사용자에게 편리성을 제공하고 있습니다.



[그림 3] TYPE1 하이퍼바이저 구성 예시 (사진: System Center 2012, koalra.com)



[그림 4] TYPE2 하이퍼바이저 예시 (왼쪽: VirtualBox, 오른쪽: VMware Player)

서버 가상화는 기존 물리 서버 위에서 또 다른 여러 개의 가상 서버를 생성하는 과정이 생기기에 일정 수준의 성능 저하가 발생함에도 인프라를 유연하게 사용할 수 있다는 큰 장점이 생깁니다. 초기 서버 가상화 환경이 등장하였을 때는 성능이 확보되지 않아 가상화가 오늘날과 같이 보급되지 못하였습니다. 그러나 하드웨어, 특히 중앙처리장치로 불리는 CPU에서 '하드웨어 가상화' 기술을 지원하면서 성능이 어느 정도 확보가 되어 오늘날 가상화 환경은 많은 인프라로 확산되어 특히, 클라우드 컴퓨팅 인프라를 구축한 회사들은 모두 서버 가상화 환경으로 구축되어 있다고 봐도 과언이 아닐 정도입니다. 하드웨어 가상화는 [그림 5]와 같은 프로그램으로 Intel CPU의 경우 VT-x, AMD CPU의 경우 AMD-V를 지원하는지 확인할 수 있습니다.



[그림 5] CPU 하드웨어 가상화 지원 여부 확인 프로그램

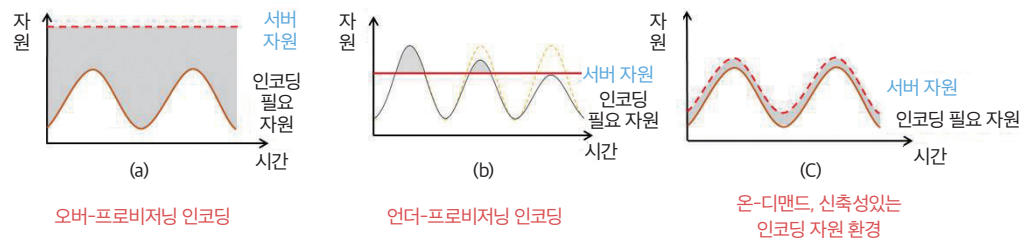
가상화와 클라우드 컴퓨팅

‘클라우드 컴퓨팅’이란 용어에서 클라우드는 구름을 의미하는데, 이 구름은 마치 컴퓨팅이 일어나는 곳이 어디인지 정확하게 알지 못하는 상황을 ‘구름’에 비유한 것이 아닐까 생각합니다. 언제 어디서나 사용한 만큼 지불이 가능한 컴퓨팅 환경을 클라우드 컴퓨팅 이라고 하는데, 이때 클라우드 컴퓨팅을 사용하는 사람은 실제 계산이 어디에서 일어나는지 알지 않아도 되는, 즉 사용자 입장에서 는 일종의 구름 위의 어딘가에서 계산이 이루어지는 상황을 빗대어 클라우드 컴퓨팅이란 용어가 등장하지 않았나 생각합니다.

클라우드 컴퓨팅은 가상화와 매우 밀접한 관계가 있습니다. 언제 어디서나 클라우드 컴퓨팅을 사용 가능하도록 하는 인프라는 컴 퓨터 네트워크 및 인터넷 인프라와 밀접한 관련이 있으며, 사용한 만큼 지불이 가능한 컴퓨팅 환경과 관련해서는 클라우드 컴퓨팅 의 확장성과 신축성이라는 두 요소가 필요합니다. 동영상 인코딩을 하는 경우를 예로 들어, 확장성과 신축성을 살펴봅시다.

먼저, 확장성은 동영상 인코딩 서버를 증설하는 부분과 관련되어 있습니다. 1대의 서버에서 10시간 정도 소요되는 용량이 매우 큰 동영상이라고 가정하면, 만약 해당 서버의 처리 능력을 2배 향상시킨다면 동영상 인코딩 작업은 절반 정도로 줄어들 것입니다. 이렇게 서버 처리 능력을 확장할 수 있는 부분은 확장성과 관련이 있습니다.

반면, 신축성은 해당 서버 처리 능력을 일정 기간만 잠깐 몇 배로 향상시키기도 하고 다시 축소시키기도 하는 것을 이야기합니다. [그림 6]에서 보면, (a)의 경우 인코딩을 하는데 2시간 정도 소요된다고 했을 때, 서버 자원이 너무 많아 회색 부분이 다른 인코딩 작 업 등에 할당 가능한 자원의 양이 될 것입니다. 이를 클라우드 컴퓨팅에서는 ‘오버-프로비저닝’ 되었다고 이야기합니다. 반면, (b)의 경우에는 ‘언더-프로비저닝’ 되어 인코딩이 완료되기까지 시간이 더욱 걸릴 것입니다. 그런데 (c)와 같이 서버 자원을 인코딩을 필 요로 하는 자원 양 만큼만 확장 및 축소가 가능하다면 여유 서버 자원을 다른 곳에 배치하여 다른 작업에 활용할 수 있을 것입니다. 이와 같이 확장성과 신축성을 지원하여, 클라우드 컴퓨팅 환경에서는 필요로 하는 서버 자원만큼만 사용 가능하도록 환경을 구성 하고, 사용자가 사용한 만큼만 지불하도록 인프라를 구성합니다.



[그림 6] 클라우드 컴퓨팅 프로비저닝 예: 인코딩 작업

이때, 클라우드 컴퓨팅의 확장성과 신축성은 ‘가상화’에 의해 이루어집니다. 물리 서버 위에 가상 서버를 구성하였기에, 해당 가상 서버에 할당된 CPU, 메모리 등의 자원을 늘리거나 줄이는 것이 가능해집니다. 또한, 물리 서버 대수를 늘리기 위해서는 관련된 물 리 시설을 준비해야 하고, 해당 물리 서버에 구동되는 운영체제를 새로 설치해야 하지만, 가상 서버 환경에서는 ‘스냅샷’이라는 기 능을 통해 현재 구동 중인 가상 서버 상태를 여러 개 복사하여 가상 서버 대수를 늘릴 수 있습니다. 이와 같이 서버를 늘리는 방식을 스케일 업과 스케일 아웃이라고 하는데, 클라우드 컴퓨팅에서는 가상화를 통해 확장성과 신축성이 [표 2]와 같이 이루어집니다.

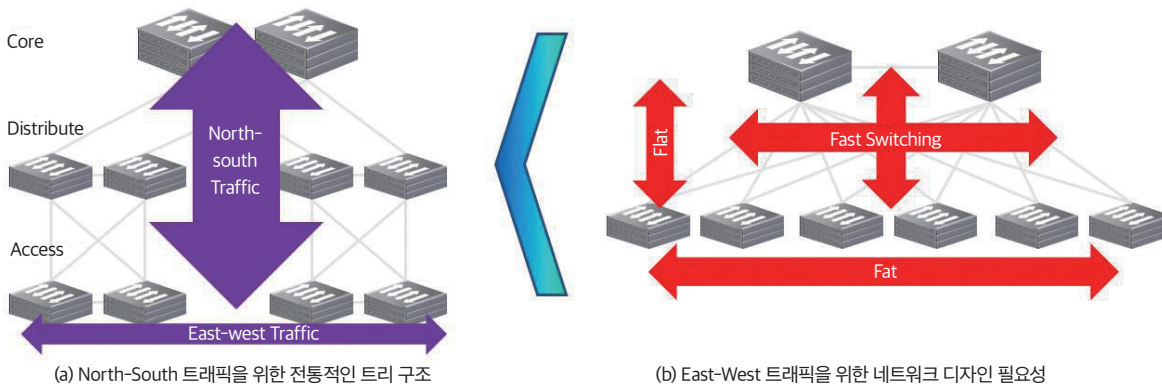
구분	정의	가상화에서의 적용
스케일 업	서버 자체를 증강하는 것으로 프로세서 (CPU), 메모리 등을 늘림	1) 가상 서버에 할당된 CPU, 메모리 등을 증설하여 확장 2) 물리 서버 용량에 한계가 있는 경우, 물리 용량이 더 큰 서버로 가상 서버를 이동(마이그레이션) 후 스케일 업이 이루어짐
스케일 아웃	서버 대수를 늘려 처리 능력을 향상	한 가상 머신 파일 또는 스냅샷을 복사 등을 통해 공유한 후, 여러 물리 서버에 동시에 실행

[표 2] 스케일 업과 스케일 아웃

데이터센터와 네트워크

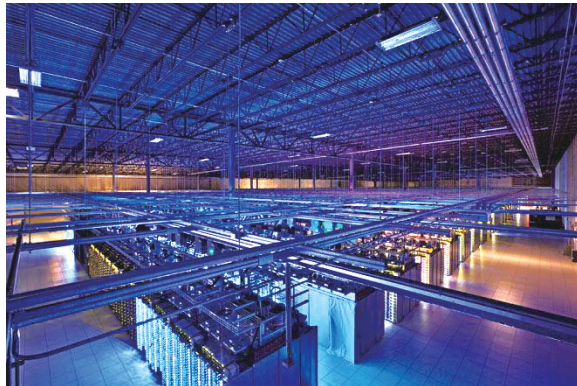
클라우드에서 서버 가상화 환경이 필수로 도입되면서 생긴 큰 변화로 인프라 시설이 거대해졌다는 데 있습니다. 서버 등 IT 인프라를 중앙 집중식 환경으로 전용 건물에 구축하고 24시간 365일 해당 환경을 운영 및 관리하는 장소를 데이터센터라고 합니다. 클라우드 컴퓨팅이 등장하기 이전까지, 데이터센터에 대해 인터넷 데이터센터(IDC, Internet Data Center)란 용어를 일반적으로 사용하였습니다. 데이터센터 내에 여러 물리 서버들이 놓여 있으며, 해당 물리 서버들은 보통 데이터센터 외부의 클라이언트에서 요청이 있는 경우 요청을 물리 서버에서 처리한 후 결과를 클라이언트에 주는 트래픽 유형이 주를 이루었습니다. 그런데, 클라우드 컴퓨팅과 가상화 환경이 도입되면서, 데이터센터에서 오가는 네트워크 트래픽의 패턴이 변화하였습니다.

[그림 7]은 이러한 데이터센터 트래픽 패턴의 변화를 보여주고 있습니다. (a)는 전통적인 데이터센터 트래픽의 패턴으로, 외부 인터넷망은 코어 스위치와 연결되어 있으며, 물리 서버들은 액세스 스위치에 연결되어 있습니다. 물리 서버들은 외부 클라이언트의 요청을 주로 거리하기에, 코어 스위치에는 외부 인터넷망과 통신을 하는 트래픽이 많아지며, 이에 따라 코어 스위치들은 많은 트래픽을 수용 가능한 값비싼 스위치들을 필요로 하였습니다. 반면, (b)는 클라우드 컴퓨팅 및 가상화 환경으로 인해 변화하는 데이터센터 트래픽 패턴에 해당합니다. 오늘날 데이터센터에서는 내부에서 먼저 트래픽을 처리하는 경우가 많습니다. 예를 들어, 빅데이터의 경우 축적된 자료를 데이터센터 내에서 여러 방향으로 먼저 분석이 이루어져야 할 것입니다. 그리고 인코딩의 경우 이전에는 요청이 있을 때 해당 인코딩 포맷으로만 인코딩하기도 하였으나, 오늘날에는 미리 여러 인코딩 포맷 및 비트레이트에 따라 인코딩을 수행하여 N-스크린과 같은 여러 단말기들을 지원하기도 합니다. 이렇게 데이터센터 내부에서 트래픽이 더 많이 발생하는 경우가 많아지고 있으며 이를 수용하기 위해 새로운 데이터센터 네트워크 디자인을 필요로 하였습니다.



[그림 7] 데이터센터 트래픽 패턴의 변화 / 참조 : 서영석, 이미지, “SDN 입문”, 영진출판사, 2014년 1월

또한 데이터센터의 규모가 점차 커지면서 네트워크 인프라를 디자인하는데 기존과는 다른 새로운 여러 방식들을 필요로 하였습니다. [그림 8]과 같이 해외에 있는 주요 데이터센터들은 수십만 대 가량의 물리 서버들을 두고 있습니다. 해당 데이터센터를 활용해 구글은 YouTube 동영상 서비스를 포함한 많은 인터넷 서비스들을 제공하고 있습니다. 반면, 아마존의 경우 2014년 12월, 자체 클라우드 컴퓨팅 플랫폼을 기반으로 UHD 화질의 인터넷 스트리밍 방송 서비스를 하겠다고 발표하기도 하는 등 인프라를 활용한 기존과는 다른 유형의 여러 서비스가 탄생하고 있습니다. 이러한 대규모의 데이터센터 플랫폼 내에 새로운 유형의 복잡한 서비스를 제공하기 위해 데이터센터 내 네트워크 인프라에 대한 많은 새로운 기술들이 활용되고 있습니다. 이때, 서버 가상화를 통해 유연함을 제공하여 물리 서버 관리에 대한 이점을 주었던 것처럼, 네트워크 또한 네트워크 가상화라는 이름으로 여러 관련 기술이 등장하고 있습니다. 네트워크 가상화를 통해 데이터센터에서는 네트워크 관리자에게 네트워크의 유연성뿐만 아니라 관리적인 여러 이점을 제공하는데, 해당 사항에 대해서는 추후 연재에서 다루고자 합니다.



Google Data Center (Iowa, US)

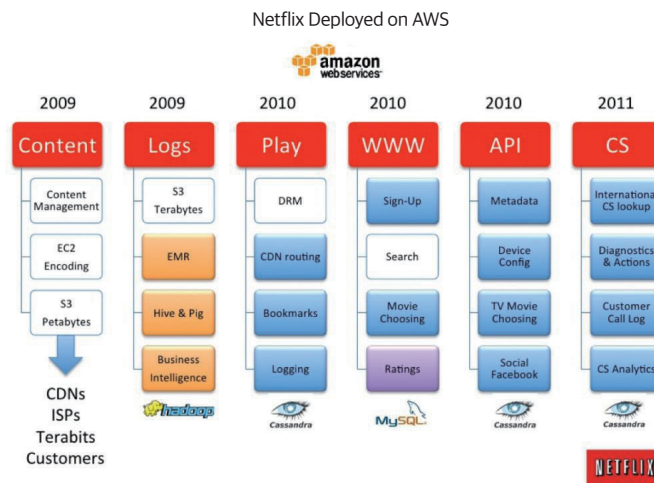


Microsoft Data Center (Dublin, Ireland)

[그림 8] 해외 주요 데이터센터 사진

일반적으로 사내 보안이 중요한 데이터를 처리하거나 또는 외부에 공개하기 어려운 주요 업무를 클라우드 컴퓨팅을 통해 처리하고자 하는 경우에는 자체적으로 데이터센터를 구축하는 방식을 통해 별도의 물리 서버와 스토리지 자원을 구축합니다. 반면, 직접 데이터센터를 구축하기 어려운 상황인 경우에는 아마존, 구글, 마이크로소프트 등과 같은 클라우드 제공 업체에 가입 후 가상 서버들을 만들어 사용할 수도 있습니다. (클라우드 컴퓨팅 관점에서는 이를 클라우드 공개 범위로 구분 지어, 사설 클라우드와 공용 클라우드로 구분합니다.)

한 예로, 미국에서 온라인 스트리밍을 위주로 서비스하는 넷플릭스(Netflix)라는 업체가 있습니다. 해당 업체는 캐나다, 멕시코, 유럽 일부 국가 등 해외 여러 국가에 온라인 스트리밍 서비스를 제공하고 있습니다. 아직 국내에서 서비스를 하고 있지 않으나, 4000만 장 이상의 영상 콘텐츠를 보유하고 있어, 미국에서는 저녁 프라임타임 때 넷플릭스 트래픽이 1/3까지 차지한다고 할 정도로 인기가 많습니다. 이 업체의 경우, 자체 데이터센터 관리에 대한 부담 등을 이유로 아마존 공용 클라우드에 가입하여 사용하고 있습니다. 직접 데이터센터를 구축하여 서비스를 제공하는 대신 공용 클라우드 위에 서비스를 구축하고, 이를 자동화된 방식으로 관리 가능하도록 구성하여 가상 서버 확장 및 축소, 그리고 여러 해외 국가로의 서비스 확장을 유연하게 처리하도록 인프라를 구성하였습니다.



[그림 9] 아마존 클라우드에 배포된 넷플릭스 / 출처: <http://goo.gl/9fJPSH>

이번 호에서는 가상화와 데이터센터를 주제로 하여 서버 가상화와 하이퍼바이저, 클라우드 컴퓨팅과 함께 데이터센터에서 변화하는 트래픽 패턴에 대해 살펴보았습니다. 다음 호에서는 데이터센터 내 네트워크에 대해 살펴볼 예정입니다. ☺