

가상화 기술과 클라우드 컴퓨팅

1. 개요

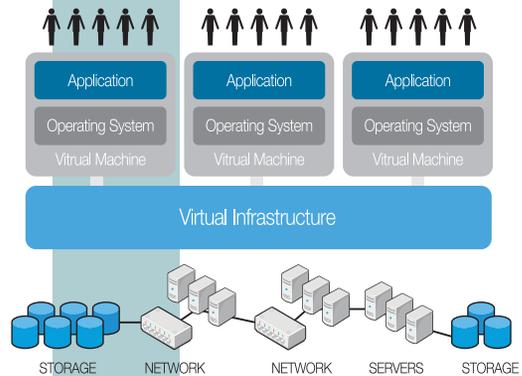
가상화(Virtualization)는 실제 물리적인 정보 기술자원의 논리적 재구성을 통해, 시스템 일관성 · 확장성 · 효율성을 증대시키기 위한 제반 기술을 의미한다. 이러한 정보 기술자원에는 CPU, Memory, Storage, Network, Software 및 Service 등 다양한 유형이 존재한다.

가상화 기술은 1960년대 IBM 메인프레임에서 사용되기 시작하여, 점차 다양한 분야로 확산되기 시작했다. 특히, 최근에는 x86기반 서버에 가상화 기술이 본격적으로 도입됐고, 스토리지 및 네트워크 분야 등에서도 중요한 이슈가 되고 있다. 이와 함께 소프트웨어 분야(Application 및 Service 등)에도 적용되면서 그 영역을 점차 넓혀나가고 있으며, 이에 따른 다양한 비즈니스 모델도 수립됐다.

가상화 기술의 주요 도입목적 중 가장 큰 것이 정보자원 활용률 극대화이다. 과거 x86기반의 서버들은 하나의 비즈니스 관점에서 애플리케이션을 장착한 형태로 구성했다. 예를 들어, 메일 서버는 하나의 x86 아키텍처 하드웨어 위에 메일이라는 애플리케이션을 장착한 형태이다. 자원 활용 측면에서 보면, 전체 시스템 자원의 약 10~20%만 활용하며, 나머지에 해당하는 자원은 거의 무사용 상태인 셈이다. 이것이 현재 대부분 기업용 x86기반의 시스템 자원 활용 상황이다. 게다가 새로운 기술이 탑재된 제품이 속속 출시됨에 따라 시스템 자원의 비효율적 사용 사례는 지속적으로 증가하는 추세이다.

이러한 시스템의 유휴자원을 좀 더 효율적으로 활용할 수 있는 방안을 제시하는 것이 바로 가상화 기술이다. 이는 다수의 물리적 정보 기술자원을 사용자 관점에서 유용한 다수의 세부자원으로 논리적 재구성한다. 이를 통해 기존 시스템의 유휴자원을 가상 시스템 전용 자원으로 재활용하며 물리적인 하나의 하드웨어 장비의 시스템 자원 활용률을 향상시킨다. 아울러 재구성된 개별 논리적 세부자원에 대해 자원 상호간 독립성을 보장한다.

여기에서는 주요 가상화 기술의 구체적인 기술유형 등에 대해서 살펴보고자 한다. 특히, IT 인프라 부문에서는 Server와 Network, Storage 가상화를 중점적으로 살펴볼 것이며, Application 부분에서는 SBC(Server-Based Computing)에 대해서 알아보겠다. 아울러 최근의 가상화 기술인 클라우드 컴퓨팅(Cloud computing)에 대해서도 살펴보겠다.



2. Server 가상화

가장 일반적인 가상화 영역이다. 흔히, 가상화라고 하면 이를 떠올릴 만큼 가장 오랜 기간 발전해 왔다. Server 가상화는 CPU, 메모리, 입출력 등의 다양한 물리적 서버 자원을 하나의 집합으로 묶어서 사용자의 필요에 따라 자유롭게 나누어 사용하는 개념이다. 이는 여러 대의 작은 서버들을 보다 쉽게 관리하고, 활용률을 높이는데 주로 사용된다. 이러한 서버 가상화를 통해, 각각의 서버에 대한 이용률을 증대시킬 수 있다는 것이 큰 이점이라 할 수 있다. 서버 가상화의 대표적인 기술이 파티셔닝(Partitioning) 이다.

2-1. 파티셔닝(Partitioning)

파티셔닝은 다수 서버에 대해 하나 이상의 독립적인 운영환경을 구현할 수 있도록 시스템 자원을 물리적 또는 논리적으로 분할하는 기술이라고 할 수 있다. 이러한 파티셔닝 기술을 통해 기존 서버 자원의 활용도를 높일 수 있으며, 운영되는 서버의 수를 감소시켜 유지보수 비용을 절감할 수 있다. 또한, 관리해야 할 서버 수 감소에 따른 전력비용과 상면비용 절감도 가능하다.

[파티셔닝 유형]

Physical Partitioning	- 서버 H/W 기반으로 서버 분할하는 방식 - 서버의 System board(또는 Cell board) 단위로 수행
Logical Partitioning	- 시스템 Microcode(firmware) 단위로 파티셔닝 수행 - H/W 부분은 공유되고, microcode 단위로 별도 파티션 구현
Software Partitioning	- 서버 Hardware/Firmware와 운영체제 사이에 구현된 S/W 코드에 따른 파티셔닝 수행 - S/W 코드별로 OS Kernel을 load하고 Hardware를 연계하는 방식

2-2. 하이퍼바이저(Hypervisor)

파티셔닝의 대표적인 상용화 기술에는 hypervisor가 있다. 이는 하드웨어를 논리적으로 분할하여 복수 파티션으로 나누고, 각 파티션 간의 보안성과 안정성을 관장하여 강력한 독립성을 보장하는 기술이다. 가상 머신 모니터(virtual machine monitor)라고도 하며, 하나의 호스트 컴퓨터 상에서 동시에 다양한 운영체제(OS)를 구동시킬 수 있는 가상화 플랫폼을 제공한다.

[hypervisor의 구현방식]

Type 1	- native 또는 bare-metal. 물리 하드웨어 상에서 해당 하이퍼바이저가 직접 실행 - 운영 체제는 하이퍼바이저 위에서 실행
Type 2	- hosted 기반. 물리 하드웨어 상의 호스트 운영체제 상에서 일반 프로그램과 같이 하이퍼바이저 실행 - 게스트 운영체제는 하이퍼바이저 상에서 실행

하이퍼바이저 기술 선정시, 크게 다음의 요소를 고려해야 한다. 시스템 측면으로는 물리적 자원의 분할 가능 개수, 파티션 관리 용이성 및 장애 발생시 상호 영향 최소화 등의 요소들을 점검한다. 이와 함께, 기술이 적용될 비즈니스 분야의 요건을 충분히 고려해야 한다. 이를 통해 비즈니스에 적합한 최적의 시스템 자원을 제공할 수 있다.

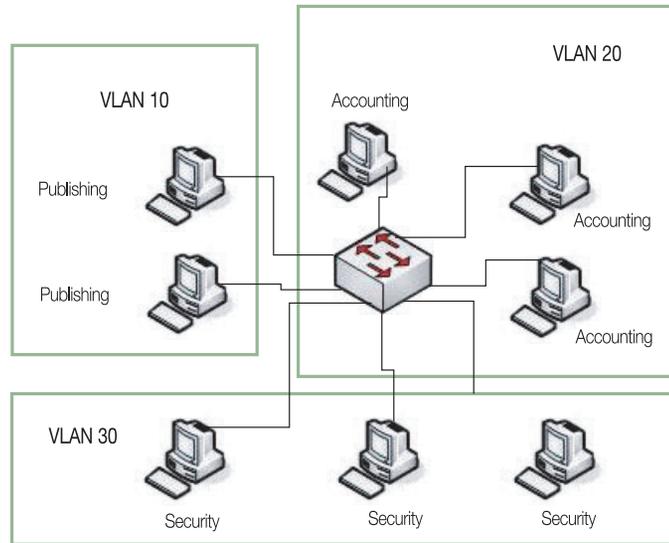
3. Network 가상화

네트워크 가상화는 네트워크를 통해 여러 사용자들을 논리적인 가상 환경으로 묶어주는 기술이다. 이는 여러 명의 사용자가 같은 가상공간을 공유하면서 협력 작업을 가능하게 한다. 네트워크에서 공유되어 사용할 수 있는 자원은 대표적인 네트워크 장비들인 라우터, 스위치, 방화벽 등이 있다. 이러한 자원들을 효율적으로 사용하기 위해, 사용자에 대한 물리적 자원을 분리할 수 있어야 하고 오류 발생시에 가용성을 제공하여 신뢰성 있는 복구가 행해져야 한다.

3-1. V LAN(Virtual LAN)

VLAN은 임의의 LAN을 논리적인 관점에서 분할한 스위치 네트워크를 의미한다. 최근 네트워크 의존도가 높아지고, 트래픽이 증가했으며, 보안 요구도 증대되면서 대부분의 스위치 네트워크에서 VLAN을 사용하게 됐다.

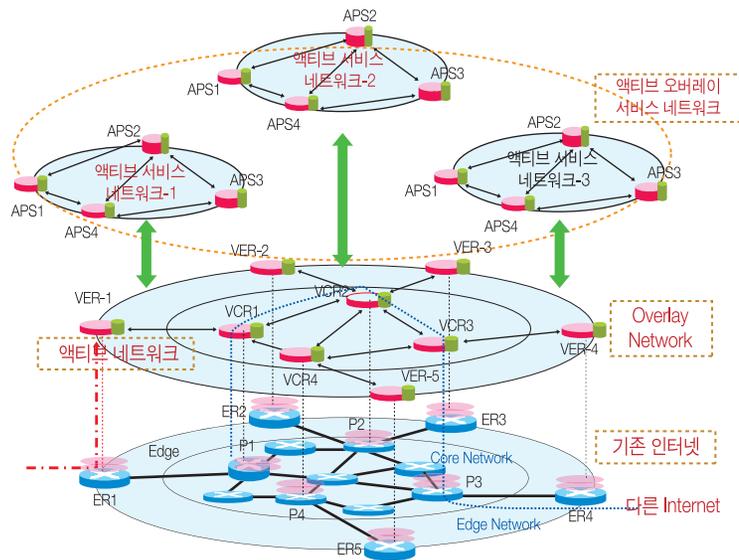
VLAN은 스위치의 전체 인터페이스들을 다수의 동일 그룹 브로드캐스트 도메인으로 분류하여 설정함으로써, 해당 그룹 내의 도메인 내에서 통신할 수 있도록 한다. 따라서, 원하는 포트들로만 프레임 전송을 가능하게 할 수 있다. 이를 통해, 전체적인 네트워크 트래픽 감소 및 성능향상과 보안성 증대, 부하분산(load balancing) 등의 다양한 목적을 동시에 달성할 수 있다.



3-2. Overlay Network

기존의 설치된 네트워크를 이용하여 부가적인 네트워크 서비스를 제공하기 위한 가상화 네트워크이다. 기존 네트워크의 물리적 토폴로지(physical topology)를 기반으로 하여 새로운 논리 토폴로지(logical topology)를 구성함으로써, 네트워크 사용자에게 가치 있는 네트워크 서비스를 제공하기 위한 가상화 네트워크이다. 이는 기존의 물리적 노드들 사이에 논리적 링크(logical link)를 새로이 재구성함으로써 이루어지며, 기존 네트워크를 최대한 활용하여 보다 효율적인 네트워크 서비스를 제공할 수 있다.

대표적인 것이 P2P 서비스가 될 것이다. 물리적 네트워크에 연결된 Peer들 간에 P2P서비스를 통해 상호 연결되어 통신을 수행하게 된다. P2P 오버레이 네트워크상에서 Peer들 간 콘텐츠 전송 및 다양한 통신을 가능하게 함으로써, 기존에는 제공하지 못했던 고품질 네트워크 서비스를 제공하게 된다.



4. Storage 가상화

스토리지 가상화는 개별 저장장치의 낮은 활용도와 관리복잡도 증대, 중복투자에 따른 문제점을 극복하고 효율적인 스토리지 통합을 위해 등장한 기술이다. 이는 물리적인 스토리지 박스 단위를 뛰어넘어 여러 대의 스토리지를 논리적인 스토리지 저장 풀로 구성함으로써 사용자 별 손쉬운 용량 배분 할당, 여러 기기에 걸친 볼륨 구성, 이기종 서버 간 볼륨 단위 데이터 공유 등을 구현한다. 보통 다음과 같은 기술들을 많이 언급한다.

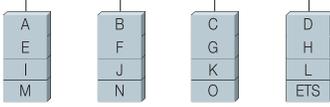
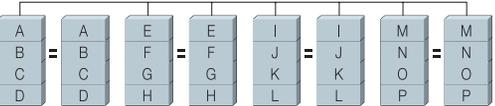
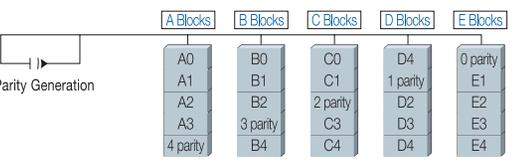
4-1. RAID(Redundant Array of Independent Disk)

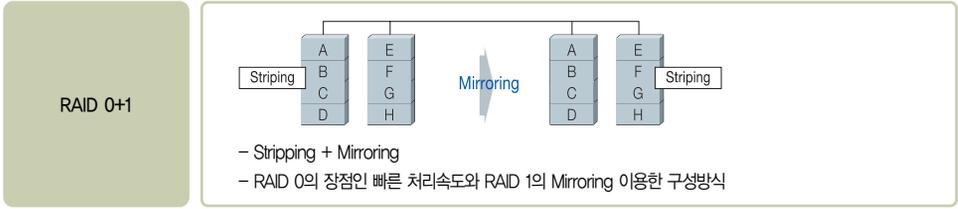
여러 디스크를 병렬로 구축하여 물적으로는 여러 개의 디스크들로 존재하나 논리적으로는 하나의 디스크처럼 보이도록 하는 스토리지 가상화 기술이다. 복수의 드라이브 집합을 하나의 저장장치처럼 다룰 수 있고, 장애발생시 데이터의 안정적인 회복을 보장하며, 개별 디스크 크가 독립적으로 동작할 수 있도록 구현된 저장장치 기술이다.

[RAID 기술의 원리]

Data Striping	- 데이터를 여러 조각으로 나누어 여러 Disk들로 분산하여 저장하고, 동시 Access를 가능하게 함(Disk Interleaving)
Redundancy	- 일부 Data가 손실된 경우 복구할 수 있도록 Data 중복 저장

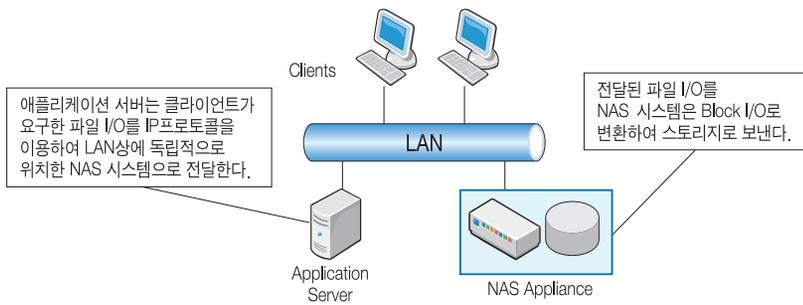
[자주 이용되는 RAID 유형]

RAID 0	 <ul style="list-style-type: none"> - Striping. 데이터를 Segment 단위로 각각의 디스크에 분산 저장 - 높은 I/O 성능, 데이터는 병렬로 Write, 디스크는 비동기 동작 - 장애 발생에 대비한 여분의 공간을 갖지 않으므로 데이터 복구 기능 부재
RAID 1	 <ul style="list-style-type: none"> - Mirroring. 모든 디스크마다 하나의 중복 디스크를 설치하는 방법 - Fault Tolerant에 유용하며 금융기관 등 중요 데이터 처리시 사용. 비경제적인 방법
RAID 5	 <ul style="list-style-type: none"> - Rotating Independent Disk Array - 데이터 복구를 위한 Parity를 함께 Striping하여 각 디스크에 저장 - 높은 처리속도를 보장하며 I/O가 빈번한 곳에 적합

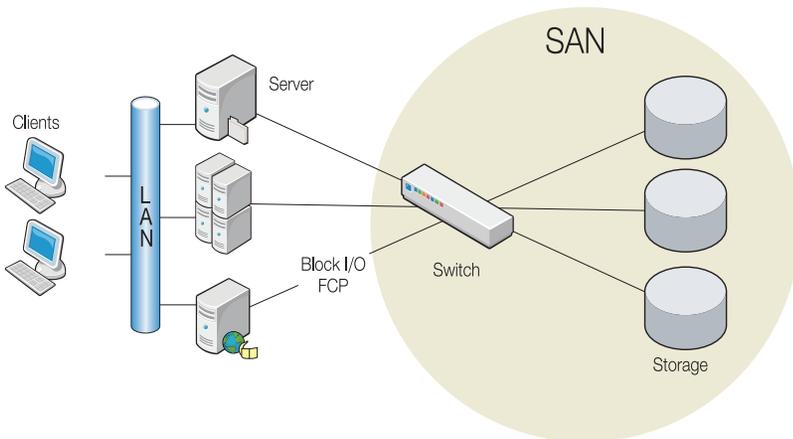


4-2. NAS(Network Attached Storage)

전용 파일 서버를 두고, LAN이나 WAN과 같은 Ethernet Interface를 통해 접속하여 Storage의 내용을 활용하는 네트워크 저장장치이다. 파일 I/O 프로토콜을 이용하여, 원격 저장장치에 I/O Operation을 수행하여 다양한 파일을 저장할 수 있다.



4-3. SAN(Storage Area Network)



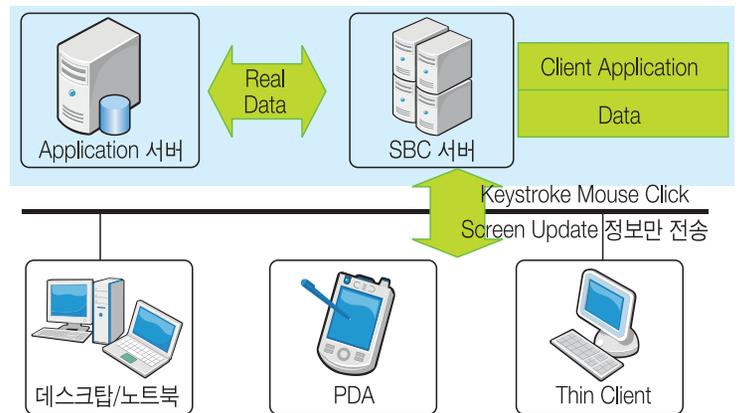
SAN은 개별 스토리지를 연결해 별도의 스토리지 전용 네트워크를 구성하는 방식이다. 전용 네트워크 스위치가 있으며, 이를 통해 복수의 이기종 서버들이 분산된 스토리지에 접속할 수 있다. 저장장치의 데이터 접근시 기존 네트워크와는 별도의 채널을 이용하므로 기존 네트워크 부하가 최소화된다. 이 기술 환경 하에서 스토리지는 수많은 서버들을 지원하는 일종의 대규모 데이터 '풀(Pool)'로 인식할 수 있다.

별도의 데이터 전달 통로를 통해 스토리지 시스템에 액세스하므로, 일반 네트워크 소통량에 영향을 받지 않고 신속한 데이터 액세스가 가능하다. 또한, 스토리지 관리 유연성과 확장성, 가용성이 우수하다. 아울러, 여러 대의 서버와 스토리지 Array를 연동하기 쉽다는 점에서 확장성이 높고 투자 회수 효과(ROI)가 높은 편이다.

반면, 초기 투자비용이 높다는 점이 단점으로 지적된다. 아울러, NAS에 비해 파일 공유 기능이 다소 부족한 편이다. SAN 제조업체 간 장비 비호환성과 네트워크 구성 거리의 한계가 있지만, 최근 IP-SAN 기술의 도입으로 이러한 한계는 어느 정도 극복됐다.

5. 어플리케이션 가상화 – SBC(Server Based Computing)

서버 기반 컴퓨팅은 서버 중심 전산환경으로 모든 Application과 Data를 서버에 두고 필요할 때마다 접속하여 이를 사용할 수 있도록 함으로써, Client에게 IT 서비스를 제공하는 기술이다. 이 방식은 모든 처리가 100% 서버에서 이루어지고, 팻 클라이언트(Fat Client)와 반대된다는 뜻에서 썬 클라이언트(Thin Client) 컴퓨팅이라고 표현하기도 한다. 대표적인 것으로 Windows 운영 체제에서 제공하는 원격 데스크톱(Remote Desktop)을 들 수 있다.



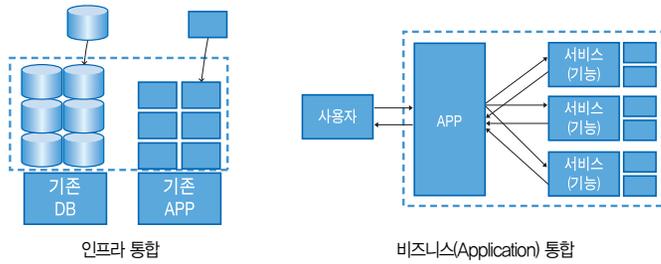
클라이언트 사양과는 무관하게 서버에서 모든 응용 프로그램을 실행하므로 충분한 성능을 보장할 수 있으며, 어플리케이션 실행환경 변화에 따른 클라이언트 업그레이드가 불필요하다. 또한, 기업 문서유출 등의 보안 사고에도 대비가 가능하다.

6. 최근의 가상화 경향 – Cloud computing

앞에서 언급한 가상화 기술들은 그 각각으로써 뿐만 아니라, 서로 통합·연계되어 다양한 형태로 활용된다. 이러한 가상화 기술들의 총체적인 구현기술로써, 최근 클라우드 컴퓨팅이 많이 언급되고 있다.

클라우드 컴퓨팅은 과거의 메인 프레임, Client/Server 및 웹 환경으로의 발전 과정에서 다양한 패러다임과 가상화 기술들이 결합되어 등장한 기술이다. 이는 서로 다른 물리적 위치에 있는 다양한 IT 자원들을 가상화 기술로 통합하여 사용이 가능하도록 한 컴퓨팅 환경이라고 할 수 있다. 좀 더 구체적으로 표현하자면, 다양한 프로그램이나 문서를 인터넷으로 접속할 수 있는 대형 컴퓨터에 저장하고, PC/휴대폰과 같은 각종 단말기로 원격에서 원하는 작업을 수행할 수 있는 컴퓨팅 사용 환경을 의미한다.

구름(Cloud)이란 표현은 소프트웨어나 데이터가 저장된 대형 서버가 물리적으로 고정되어 있는 것이 아니라 네트워크로 연결되어 마치 구름처럼 형태가 없음을 의미한다. 복잡한 시스템 인터페이스나 환경 등을 ‘구름(Cloud)’ 속에 숨겨 놓고, 정보 시스템 자원을 활용하기 위해 사용자는 시스템의 내부 구성에 대한 정보와는 전혀 상관없이 비즈니스에 필요한 서비스를 선택하고 활용할 수 있게 하는 기술이 바로 Cloud Computing이라 할 수 있다.



[클라우드 컴퓨팅의 가상화 기술]

인프라 측면에서는 내부 서버들 간 리소스를 분배해 동적 할당과 활용이 가능하도록 한다. 외부에서는 필요한 리소스의 양만 요청하게 되고, 클라우드 컴퓨팅 환경 내부적으로 요청된 리소스를 제공하기 위한 컴퓨팅 자원의 할당을 통해 작업을 수행하게 된다.

비즈니스(Application) 통합 측면에서 전체 애플리케이션과 데이터들은 모두 하나의 시스템에 존재하는 것처럼 활용할 수 있다. 기반이 되는 개별 애플리케이션들은 사용자가 접근해 사용할 수 있는 서비스의 형태로 외부에 인터페이스를 제공하게 된다. 이러한 서비스들을 효과적으로 통합하여 연계(mash-up) 형태로 사용자에게 적합한 Application 기능을 제공하게 된다. 이러한 클라우드 컴퓨팅의 핵심은 사용자에게 IT 서비스를 활용할 수 있는 플랫폼을 제공하는 것이다.

[플랫폼을 제공하는 각종 기술]

분산 데이터 관리	<ul style="list-style-type: none"> - 분산 데이터 관리 : 정형화된 대용량 데이터를 네트워크로 연결된 다수의 서버에 분산 저장하고 관리 - 분산 파일 시스템 : 네트워크 상의 여러 서버에 파일 데이터를 블록 단위로 나누어 관리
클러스터 관리 기술	<ul style="list-style-type: none"> - 고가용성(High-Availability) 및 부하분산(Load-Balancing)을 위한 클러스터를 통해 시스템 안정성 확보 - 서비스 프로비저닝(Service Provisioning) : 운영체제(OS), 소프트웨어, 환경정보, 스크립트 등을 관리하며, 필요한 순간에 동적으로 자원 할당을 수행해 서비스를 생성하고 제공할 수 있는 기술. 이를 통한 서비스들의 관리 및 배치 프로세스 자동화
분산 컴퓨팅 기술	<ul style="list-style-type: none"> - 다수의 작업 태스크(Task)가 몇 대의 서버에 나누어져 실행될 때, 서버 간 메시지를 전달하면서 병렬로 컴퓨팅 수행 - 메시지는 데이터, 송·수신지 주소로 구성

대부분 분산화된 컴퓨팅 자원을 일관성 있고 안정적으로 관리하는 기술들이며, 이러한 요소기술들을 통해 사용자에게 최적의 컴퓨팅 자원을 제공하게 된다. 클라우드 컴퓨팅에서도 해결해야 할 과제는 있다. 우선, '보안' 측면이다. 기업 데이터가 기업의 관리범위가 아닌 클라우드 컴퓨팅 업체에 존재하므로 이에 대한 비밀성, 무결성, 접근의 가용성 등을 보장할 수 있는 방안을 제공해야 한다. 또한, '안정성' 측면에서 클라우드 컴퓨팅 서비스를 제공하는 업체가 얼마나 끊어짐 없이 서비스를 제공할 수 있는지도 고려해야 한다. 이러한 점들을 충분히 확신시켜 줄 수 있어야 클라우드 컴퓨팅이 좀 더 대중화되어 사용될 수 있을 것이다.

[클라우드 컴퓨팅의 사례]

Google Apps	<ul style="list-style-type: none"> - 웹 브라우저로 이용할 수 있는 일반적인 비즈니스 응용 프로그램들을 온라인으로 제공하는 서비스 소프트웨어와 데이터는 서버에 저장됨 - App Engine : Google의 웹 응용 프로그램 개발 및 호스팅 플랫폼으로 사용자들이 자신의 하드웨어 인프라 없이도 트래픽이 매우 많은 웹 응용 프로그램을 개발하여 서비스할 수 있게 함 - Google Browser(Chrome), 구글폰(안드로이드) 등은 클라우드 컴퓨팅 인프라와 서비스를 보다 다양한 경로에서 접근하며, 이에 따른 비즈니스 영역을 확대하기 위한 요소 기술에 해당
Amazon	<ul style="list-style-type: none"> - 사용자 또는 고객 요구에 따라 기능 및 용량 변경이 수월한 컴퓨팅 자원을 제공하는 웹 서비스로서, 클라우드 컴퓨팅 기술로 구축된 인프라 상에서 운영 - 뉴욕타임즈는 아마존의 클라우드 컴퓨팅 서비스로 1,100만 건의 기사를 PDF로 전환하는 과정에서 아마존의 웹 서비스를 활용했고, 240달러의 비용으로 하루 만에 작업을 끝내 클라우드 컴퓨팅의 가능성 보여줌

7. 결론

비즈니스를 원활하게 지원하기 위한 기업의 IT 자원을 최대한 적은 비용으로 제공하기 위한 노력은 이후에도 계속될 것이다. 이를 위해 현재의 Silo 방식으로 구성되어 낭비되는 유휴 시스템 자원을 하나의 Pool로 묶어 제공하는 가상화 기술은 큰 기여를 할 수 있다.

[가상화 기술 구현시 기대효과]

총 소유 비용(TCO) 절감	- 시스템 복잡도 최소화에 따른 유지보수 및 관리비용 절감
구매 관리 최소화	- 시스템 구매 발주 건수 감소, 시스템 및 S/W 도입비용 감소
효율적 용량 관리	- 비즈니스 요구에 적합한 최적의 IT 자원 제공 가능
보안 강화	- 백업 및 접근제어를 통해 기업의 보안 수준 강화 가능
Green IT 기여	- 장비 증설을 최대한 억제하여, 운영에 필요한 전력소모량 최소화

하지만, 가상화 기술 도입시 몇 가지 고려할 요소가 있다. 우선, 안정성 측면이 가장 큰 고려요소로 지적된다. 제품을 제공하는 Vendor별로 충분한 기술검증(Proof of Concept)을 수행하여 해당 비즈니스의 제한 요건들을 잘 충족하는지 여부를 파악해야 한다. 특히, 가상화 기술에서 장애 발생시 이에 대한 즉각적인 대응과 복구가 가능한지 여부를 고려한다.

기존에 도입된 S/W의 라이선스 정책과 가상화 기술이 적절히 부합하는지 검토해야 한다. 여전히 소프트웨어 라이선스 사용권한이 있음을 확실히 해야 한다. 만일, 부합되지 않을 경우, 무료 또는 최소한의 비용으로 계약 조건을 개정함으로써 차후 있을 분쟁소지를 최소화 한다. 아울러, 신규 S/W 사용계약 체결시 도입된 가상화 기술에 대해서 추가 비용 없이 사용할 수 있도록 함으로써, 가상화 기술의 가장 근본적인 목표인 비용절감을 극대화하도록 노력해야 한다.

궁극적으로 시스템 구축은 기술적인 복잡함을 숨기고 단순하고 빠른 구현이 가능한 형태로 발전될 것이다. 사용자 측면에서는 업무 변경에 즉각적으로 대응하면서 쉽게 사용이 가능한 환경이 될 것이다. 그리고, 이러한 기술의 실현을 위해 가상화 기술은 핵심적인 역할을 할 것으로 기대된다.