ISSN 2586-5021

월간 SW 중심사회

MONTHLY SOFTWARE ORIENTED SOCIETY

통권 제70호
2020 04
No.70 Apr, 2020
국가 소프트웨어 정책 Think Tank

FOCUS

코로나19와 미래사회로의 대전환

재택/원격근무와 미래의 일, 공간
포스트 코로나, 언택트 디지털 전환 기회
언택트를 콘택트처럼 만드는 5G와 실감기술의 기회

TREND

의료와 자동차 산업에서의 디지털 홀로그래피 기업 기술 동향
SW산업의 집단적 의사표시 동향
에듀테크 산업 동향 및 시사점

COLUMN

실감산업, 큰 그림을 그리자
디지털 전환 시대에 필요한 新교육·인재 전략

AI BRIEF

인공지능 최신 동향과 시사점

ISSUE

AI 성능개선을 위한 클라우드 GPU 가상화 기술의 현황과 향후방향

SPRi 소프트웨어정책연구소

코로나19와 미래사회로의 대전환 시대

코로나19로 인해 10년이 걸렸을 변화가 단 한 달 만에 일어나고 있다. 전국의 540만 명의 학생들이 온라인 수업으로 등교를 대신하고 있고, 논의만 활발하던 원격진료도 한정적으로나마 시행되고 있다. 일부에 그쳤던 유연근무나 재택근무는 코로나19 이후 일반적인 모습이 되었다.

온라인 교육, 원격진료, 재택근무 등은 코로나19로 인해 본격적으로 등장했지만, 어쩌면 미래학자 엘빈 토플러가 1980년에 '제3의 물결'에서 미래의 우리 삶의 모습으로 언급했던 바로 그 모습이다. 코로나19라는 세계적인 대유행으로 인해 40여 년 전에 예견한 미래사회가 갑작스럽게 우리 앞에 펼쳐졌다. 결국 우리는 어디로 가야하는지는 오래전부터 알았지만 언제 어떻게 가야할지 몰랐을 뿐이며, 지금 미래사회로 가는 대전환의 시대의 길목에 서있다고 해도 과언이 아니다.

사회적 거리두기로 시작한 여러 시도들은 예상치 못한 어려움을 가늠해보고, 우려해왔던 부작용들이 실제로 어느 정도인지 확인하여 더 나은 방법을 찾도록 도와 줄 것이다. 하지만 우리 사회는 전환 실패의 경험이 많다. 스마트팜은 농민단체와 갈등을 풀지 못하고 여러 시도들이 무산되었으며, 풀러스와 타다와 같은 공유 모빌리티는 기존 사업자와의 갈등으로 서비스를 중단하기에 이르렀다.

우리는 지금 미래사회로 가는 대전환의 시대에서 다시 과거로 회귀할 것인지 아니면 더 나은 미래를 만들어 나갈지 선택의 기로에 서있다. 갑작스럽게 찾아온 변화에 어떻게 적응하고 더 나은 변화를 만들어 낼 수 있는지 고민이 필요한 시점이다. 이번호에서는 이러한 고민의 실마리를 가져다줄 글을 실었다.

월간 SW 중심사회
MONTHLY SOFTWARE ORIENTED SOCIETY

2020.04 No.70 Apr

04 TREND
동향
의료와 자동차 산업에서의 디지털 홀로그래피 기업 기술 동향
Technology Trends of Digital Holography Companies in the Medical Care and Automotive Industries

SW산업의 집단적 의사표시 동향
SW Industry Labor Union Trends

에듀테크 산업 동향 및 시사점
Edutech Industry Trends and Implications

24 COLUMN
칼럼
실감산업, 큰 그림을 그리자
Time to Broaden the Horizons of the Immersive Industry

디지털 전환 시대에 필요한 新교육·인재 전략
New Education and Talent Strategies Needed in the Era of Digital Transformation

31 AI BRIEF
AI 브리프
인공지능 최신 동향과 시사점
Recent Trends of AI and Their Implications

40 PHOTO ESSAY
포토 에세이
사진산책 중간(中間)
Your Moment with Photos : The Stopover

42 FOCUS
포커스 : 코로나19와 미래사회로의 대전환
재택/원격근무와 미래의 일, 공간
Working from Home/Telecommuting and Future Work, Space

포스트 코로나, 언택트 디지털 전환 기회
Post-corona, Opportunities for Non-contact Digital Transformation

언택트를 콘택트처럼 만드는 5G와 실감기술의 기회
Opportunities for 5G and Immersive Technology to Make Non-Contact Seem Like Contact

62 ISSUE
이슈
AI 성능개선을 위한 클라우드 GPU 가상화 기술의 현황과 향후방향
Current State and Future Directions of GPU Virtualization Technology in the Cloud Computing for Improving AI Performance

90 STATISTICS
통계
SW 산업 통계
Software Industry Statistics

의료와 자동차 산업에서의 디지털 홀로그래피 기업 기술 동향

Technology Trends of Digital Holography Companies in the Medical Care and Automotive Industries

남현숙 Nam, Hyunsuk ▌선임연구원 Senior Researcher, SPRi ▌hnam@spri.kr

최근 홀로그래피 기술은 3무(안경, 어지럼증, 공간왜곡)의 한계를 극복하기 위해 활발하게 연구가 진행되고 있다. 디지털 홀로그래피 기술은 홀로그램 광학 요소(HOE), 공간 광변조기(SLM), 백라이트 등과 같은 하드웨어 광학 소자와 3차원 영상을 위한 렌더링(Rendering), 측정(Measurement), 시각화(Visualization) 등을 위한 소프트웨어 기술의 융합 시스템이다. 이 기술을 실용화하기 위하여 글로벌 디지털 홀로그래피 기업들은 홀로그램 발명자인 데니스 가보(Dennis Gabor)의 3차원 영상기록 홀로그래피 이론을 기반으로 의료, 차량, 상업용 부문 등에서 관련 제품개발을 위해 노력하고 있다.

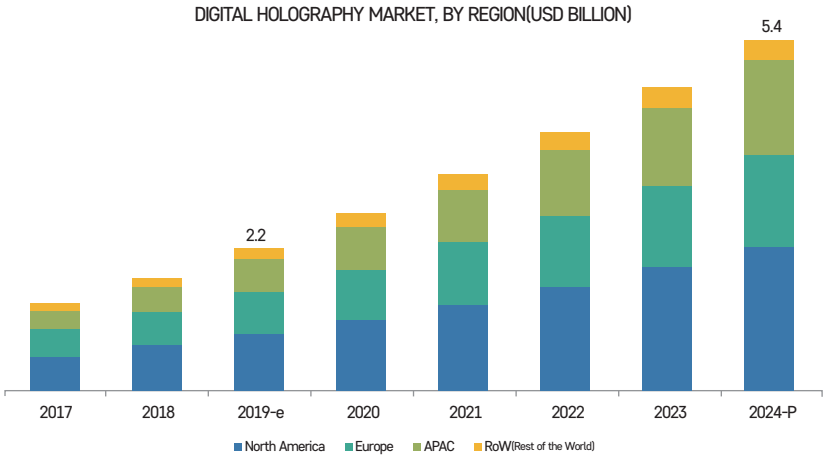
Recently, holography technology has been actively researched to overcome the limitations of 3 drawbacks(glasses, dizziness, and spatial distortion). Digital holographic technology is a fusion system of hardware optical elements, such as holographic optical elements(HOE), spatial photovoltaic modulato(SLM), and backlight, and software technology for rendering, measuring and visualizing for three-dimensional images. To make this technology practical, global digital holography companies are trying to develop related products in the medical, automotive, and commercial sectors based on the holography inventor Dennis Gabor's three-dimensional image recording holography theory.

홀로그램 개요와 시장 전망

홀로그램 기술의 역사는 1948년 노벨상 수상자인 데니스 가보(Dennis Gabor)에 의해 시작되었다. 홀로그램은 그리스에서 유래된 '전체'와 '정보'가 합쳐진 용어이고, 이는 홀로그래피로 촬영된 결과물을 말한다. 홀로그래피는 두 개의 레이저 광이 서로 만나서 일으키는 빛의 간섭 현상을 이용하여 입체정보를 기록하고 재생하는 기술을 말한다. 홀로그램 종류는 아날로그 홀로그램, 유사 홀로그램, 디지털 홀로그램 등이 있다. 이 중 디지털 홀로그램은 사물로부터 반사된 빛을 디지털화하여 기록 및 재현을 통해 현실감을 제공하는 것으로 디지털화된 홀로그램 정보를 획득·생성·전송·압축·재현하는 기술이 핵심이다.

세계 디지털 홀로그래피 시장은 [그림 1]과 같이 2019년 22억 달러에서 2024년 54억 달러로 연평균성장률(CAGR) 19.8%로 전망된다. 전 세계적으로 디지털 홀로그래피 시장의 성장요인은 의료 및 상업용 부문에서 디지털 홀로그래픽 디스플레이, 홀로그래픽 원격 실재(Telepresence) 등의 수요증가, 정밀한 측정을 위한 디지털 홀로그래피 현미경의 사용증가, 그리고 금융 부문에서 보안목적으로의 디지털 홀로그래피 채택 증가 등이 주요 원인으로 꼽힌다.

[그림 1] 디지털 홀로그래피 시장규모



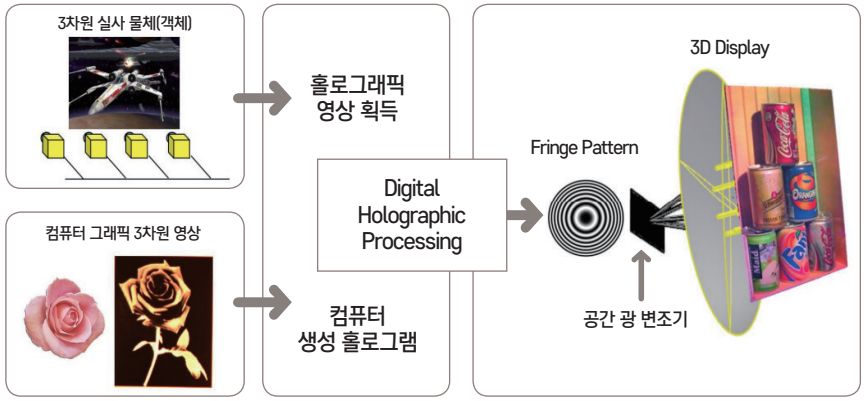
※ 자료 : MarketsandMarkets(2020), "Digital Holography Market"

디지털 홀로그램의 주요 기술

디지털 홀로그래피 기술은 [그림 2]와 같이 크게 두 가지로 구분할 수 있다. 홀로그램 영상을 획득하는 기술과 영상을 디스플레이 하는 기술이다. 이 중에서 홀로그램 영상획득 기술에는 실제 사물의

3차원 정보를 획득하는 기술과 컴퓨터 그래픽으로부터 영상을 획득하여 이를 홀로그램 영상으로 만드는 기술로 나눌 수 있다. 컴퓨터 그래픽 기반으로 홀로그램 영상을 획득하기 위한 대표적 기술로 수학적 모델링 기법을 이용한 컴퓨터 생성 홀로그램(Computer Generated Hologram : CGH) 기술이 있다. CGH 기술을 이용할 경우 광학시스템을 활용한 실물 기반의 홀로그램으로는 제작하기 어려운 다양한 가상장면들을 구현할 수 있다. CGH 기술은 복잡한 광학계 구축 문제로부터 자유롭고, 방대한 양의 파동 진행 계산을 효율적으로 수행할 수 있는 알고리즘과 병렬처리 하드웨어 기술이 급성장하면서 더욱 주목받고 있다. 특히 GPU(Graphics Processing Unit) 같은 대규모 병렬처리 하드웨어 기술의 비약적인 발전은 방대한 계산량을 요구하는 CGH의 문제점을 해결하여 디지털 홀로그램의 실용화를 촉진하고 있다.

[그림 2] 디지털 홀로그래피 주요 기술[1]



※ 자료 : 임용준, 김진웅(2018), "가상, 증강 및 혼합현실을 위한 디지털 홀로그래피 기술 연구 동향", IITP

디지털 홀로그램 영상 디스플레이 핵심기술은 공간 광변조기(Spatial Light Modulator : SLM)이다. SLM은 아날로그 홀로그램에서의 필름과 같은 역할을 한다. SLM은 빛의 진폭 또는 위상을 표현할 수 있는 소자이며 이를 기본적인 영상표시 장치로 구성하여 3차원 홀로그램 영상을 형성한다. SLM과 관련하여 홀로그래픽 디스플레이에서 더 넓은 시야각을 확보하기 위해 많은 연구가 현재 진행되고 있다.

1 임용준, 김진웅(2018), "가상, 증강 및 혼합현실을 위한 디지털 홀로그래피 기술 연구 동향", IITP

주요 산업에서 활용되는 디지털 홀로그램 기술

디지털 홀로그램 기술이 가장 활발히 상용화되는 제품에는 레이저, 카메라, 공간 필터, 프로젝션 장치 등이 포함된다. 특히 CCD 또는 CMOS[2] 카메라는 이미지 획득에 중요한 요소이다. 현재까지는 디지털 홀로그래피 기술은 의료 분야에서 현미경, 측정, 3차원 디지털 홀로그래피 디스플레이 등의 제품개발에 집중되고 있으며, 자동차 분야에서는 인포테인먼트를 위한 HUD, 디지털 홀로그래피 디스플레이 등에서 활발하게 진행되고 있다.

[표 1] 디지털 홀로그래픽 디스플레이를 활용한 산업 및 대표기업

산업	기술분야	내용	기업
자동차	AR 렌더링 HUD 시스템	• 차량용 HUD 개발을 통해 실시간 주행 정보 제공	• Continental Automotive(Leia) • WayRay • SeeReal Technology(Volkswagen)
전자기기	디지털 홀로그래픽 디스플레이	• 서브 홀로그램 기술을 통해 시청자의 위치에 따라 홀로그램 시청 가능 • Light Field 기술로 홀로그래픽 디스플레이 개발	• SeeReal Technology(VISIO 20) • Looking Glass Factory(8K Immersive Display)
의료	수술, 진단	• 의료용 현미경을 위한 디지털 홀로그래픽 디스플레이 • 홀로그램 3차원 영상을 활용한 의료진단 등에 활용	• Lyncee Tec(DHM) • RealView Imaging(HOLOSCOPE)

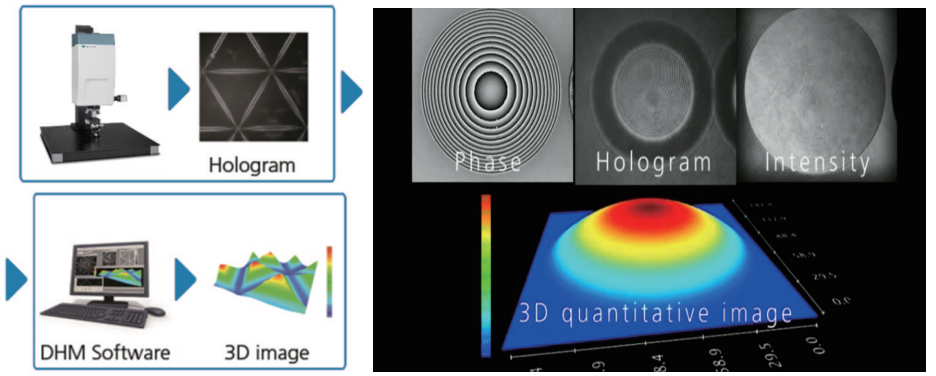
※ 자료 : 기업별 웹페이지 기반 저자 재구성

① Lyncee Tec

스위스 Lyncee Tec사는 홀로그램으로 세포의 3차원 정보를 추출하고 컴퓨터 그래픽 모델을 만들 수 있는 홀로그램 현미경을 개발하였다. [그림 3]과 같이 이 기업의 DHM(Digital Holographic Microscopy) 기술은 CCD 카메라를 사용하여 기준 파동(Reference Wave)과 표본(Specimen)에서 나오는 파동 간의 간섭에 의해 생성된 홀로그램을 기록한다. 이 기록된 이미지는 컴퓨터로 전송되어서 알고리즘이 적용되어 표본의 3차원 이미지를 재구성한다. 이 과정을 실시간 이미지 재구성(Image Reconstruction)이라고 한다. 이전의 현미경에서는 볼 수 없었던 홀로그램 이미지 획득 및 실시간 3차원 홀로그램 이미지 재구성에서는 신호처리를 위해 디지털 신호 처리방식을 이용하였다.

2 CCD, CMOS : 빛을 전기적 신호로 변환하여 저장하는 센서

[그림 3] Lyncee Tec (좌)디지털 홀로그래픽 현미경, (우)출력 이미지[3]

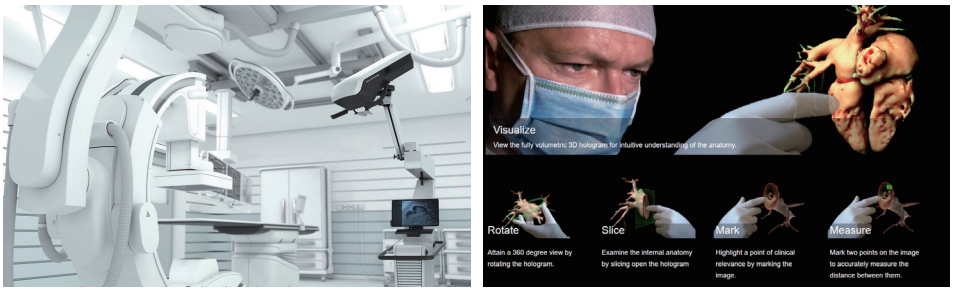


※ 자료 : Lyncee Tec 웹 페이지
Tristan Colomb et al(2007), "Enhancing the performance of digital holographic microscopy", SPIE

② RealView Imaging

이스라엘 RealView Imaging사는 의료 서비스를 위해 홀로그래픽 디스플레이 제품을 선보였다. 이 기업은 독점적인 Digital Light Shaping기술을 사용하여 복잡한 3차원 구조의 의료 이미지를 의료인이 가상경험을 통하여 쉽게 이해 할 수 있도록 하였다. HOLOSCOPE은 이를 상용화한 세계 최초의 의료 홀로그램 시스템이다. 이 제품은 3차원 스테레오스코피(Stereoscopy) 방식이 아닌 간섭 기반(Interference-Based) 홀로그래피이며, 3차원 이미지의 수렴-초점 불일치(Vergence-Accommodation Conflict)문제를 제거하여 어지럼증을 해결하였다. 이 홀로그램 이미지로 복잡한 장기 해부 및 인체의 생리적인 과정에 대해 의료인이 쉽게 이해 할 수 있도록 하였다. [그림 4]와 같이 홀로그램 이미지는 시각화, 회전, 슬라이스, 마크 및 측정 등이 가능하다.

[그림 4] RealView Imaging (좌)디지털 홀로그래피 시스템, (우)HOLOSCOPE 다양한 사례



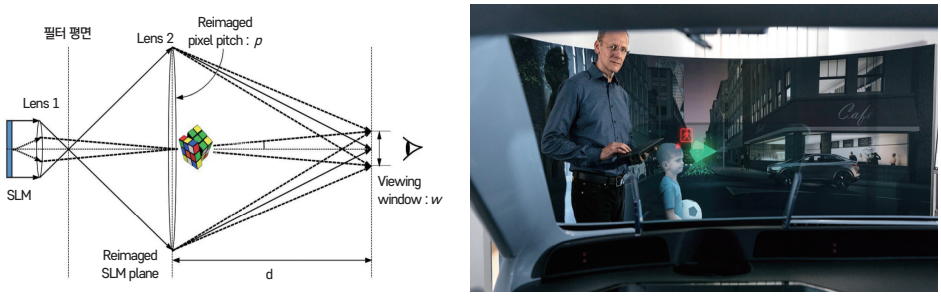
※ 자료 : RealView Imaging 웹페이지

3 Tristan Colomb et al(2007), "Enhancing the performance of digital holographic microscopy", SPIE

③ SeeReal Technologies

독일의 SeeReal Technologies사는 홀로그램 원리를 기반으로 3차원 장면을 재구성(Reconstruction)하기 위해 위상변조 디스플레이 하드웨어 및 간섭 기반 광변조를 사용한 홀로그래픽 3D 기술을 개발하였다. 일반적으로 홀로그래픽 디스플레이를 구현하기 위해서는 홀로그래픽 영상을 실시간으로 생성하고 재현하여야 하는데, 이 경우 막대한 양의 데이터를 처리해야 하므로 기술적인 어려움이 있다. 하지만 이 기업의 시야창 방식(Viewing Window) 홀로그램 기술은 [그림 5]와 같이 시청자의 동공에 해당되는 작은 시야창을 형성하고 이 영역에 대해서만 홀로그램을 생성한다. 이렇게 되면 전체 화면의 홀로그램 영역보다 처리량(Throughput)이 적어져서 처리해야 할 데이터양이 크게 줄어들기 때문에 홀로그램 영상 생성이 가능해진다.

[그림 5] SeeReal Technologies (좌)디지털 홀로그래피 디스플레이 기술 원리[4], (우)차량용 HUD 활용



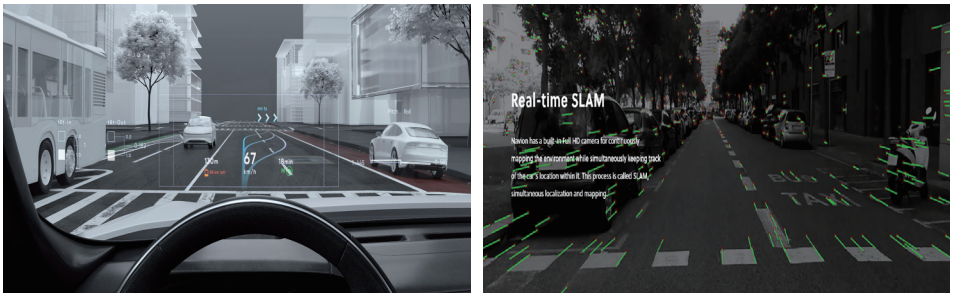
※ 자료 : SeeReal Techologies 웹페이지
임용준, 김진웅(2018), "가상, 증강 및 혼합현실을 위한 디지털 홀로그래피 기술 연구 동향", IITP

④ WayRay

WayRay사는 CES 2020에서 자율주행 자동차의 인포테인먼트 서비스를 위한 새로운 홀로그래픽 AR 디스플레이 기술을 선보였다. 이 회사는 처음으로 홀로그램 AR 디스플레이에 청색 레이저를 추가하여 완전한 RGB(Red, Green, Blue) 색의 홀로그램 기술을 보여주었다. [그림 6]과 같이 자동차 앞에 가상물체를 보여주기 위해 홀로그램 광학 요소(HOE)가 설치되어 있으며, 이 HOE는 가상이미지를 생성하여 운전자의 눈과 적절한 거리에서 경로 관련 정보를 표시해 준다. HOE를 통해 전달된 데이터는 풀(Full) HD 카메라를 통해 주변 환경에 매핑(Simultaneous Localization and Mapping : SLAM)되고 이 결과물인 가상물체를 실시간으로 디스플레이 해준다.

4 임용준, 김진웅(2018), "가상, 증강 및 혼합현실을 위한 디지털 홀로그래피 기술 연구 동향", IITP

[그림 6] WayRay (좌)가상물체를 보여주는 차량용 HUD, (우)실시간 SLAM 기술



※ 자료 : WayRay 웹페이지

시사점

디지털 홀로그래피 기술은 현재 시야각의 제약이 적은 의료 및 자동차 분야에서 상용화를 위한 개발이 활발하게 진행되고 있다. 그러나 궁극적으로 디지털 홀로그램 기술이 대중화되기 위해서는 무안경, 무어지럼증, 무왜곡을 실행할 수 있는 '3무 홀로그램 기술'이 개발되어야 할 것이다.

선도기업의 디지털 홀로그램 기술개발 및 제품개발 동향 분석을 통해 다음 세 분야의 발전이 필요한 것을 확인하였다. 첫째, 디지털 홀로그래피 디스플레이를 위한 센서, HOE, SLM, 백라이트 등과 같은 홀로그램 관련 소자 기술의 국산화가 필요하다. 둘째, 3차원 영상을 위한 렌더링, SLAM, 홀로그램 생성 알고리즘 등과 같은 소프트웨어 기술개발이 시급하다. 마지막으로 이러한 기술개발을 기반으로 교육, 엔터테인먼트, 제조 등과 같은 분야에서 활용할 수 있는 다양한 킬러콘텐츠 개발이 함께 이루어져야 할 것이다.

SW산업의 집단적 의사표시 동향

SW Industry Labor Union Trends

이종주 Lee, Jongju ▌연구원 Researcher, SPRi ▌ljj@spri.kr

SW기업 노동조합 위원장이 정치에 입문하면서 노동조합에 대한 관심이 늘고 있다. 그동안 SW산업은 고용구조와 일자리 특성으로 인해 사실상 노동조합의 불모지였다. 그런데 최근 대기업의 노동조합과 외국계 기업 지사의 노동조합이 설립되어 노총과 연대하여 활동하고 있다. 기업 단위 노동조합이지만, 이들이 연대하여 최근 정부에서 추진하려는 정책에서 SW산업 종사자들을 위한 목소리를 내는 것을 볼 수 있었다. 앞으로 정부와 업계는 SW산업에서 이들 기업 단위 노동조합과 노동조합들의 연대 움직임에 대해 관심을 가지고 대비해야 할 것으로 보인다.

Interest in labor unions is increasing as the chairman of the SW company's labor union enters politics. Until now, the SW industry had virtually no labor union activity due to its employment structure and job characteristics. However, labor unions of large corporations and foreign corporate branches have recently been established, operating in solidarity with the Federation of Korean Trade Unions. Even if they are labor unions in a corporate unit, they have united to represent the voices of workers in the field of SW regarding policies recently pursued by the government. In the future, the administration and businesses will have to pay attention to and prepare for the joint movement of these corporate-level unions and other labor unions in the SW industry.

SW산업 노동조합 위원장의 행보

2020년 4월 15일 총선을 앞두고 연동형 비례대표제에 대한 관심이 높은 가운데 정의당이 'IT산업노동특별위원장'인 류호정 위원장을 비례대표 1번으로 올린 것에 관심이 쏠리고 있다.[1] 1992년생인 류 위원장이 국회의원으로 당선되면 역대 최연소 국회의원일 수 있다는 것도 관심사지만, SW업계는 류 위원장의 과거 행보 때문에 관심을 가질 수밖에 없다.

류 위원장은 과거 e스포츠에서 선수로 활동한 경험이 있으며, 아프리카tv 등 소셜미디어 플랫폼에서 활동해온 1인 크리에이터였다. 또한, 게임회사에 근로자로 일하며 노동조합을 설립한 경험까지 있다. 역할 측면에서 SW업계에서 소위 핫한 것들을 두루두루 경험하고 정치 입문을 앞두고 있는 상황인 것이다.

류 위원장에 대한 관심은 자연스럽게 노동조합 불모지였던 IT/SW업계의 집단적 의사소통 창구에 대한 관심으로 이어지고 있다. 이번 리포트에서는 국내 SW산업의 노동조합 현황과 외국 사례 등을 통해 시사점을 찾아보고자 한다.

SW산업에서 집단적 의사표시 창구

IT/SW업계는 사용자와 노동자 사이에 대등한 구도가 형성되어 있지 않은 것은 IT/SW업계의 '일하는 방식'이라고 할 수 있을 정도로 노동조합의 무풍지대였다.[2] 그만큼 SW산업 종사자들은 자신들의 처우와 관련해서 집단적인 의사표현을 할 수 있는 창구가 없었다. 그럼에도 불구하고 SW산업에서는 다른 산업과 다르게 노동조합 같은 집단적 의사표현을 위한 창구 마련이 어렵다고 여겨왔다. 국회사무처(2013)는 IT업계에 노동조합이 활성화가 쉽지 않은 이유를 ① 영세한 중소기업 중심의 사업장 구도로 단체 결사가 어려운 점, ② IT근로자 스스로가 권리의식이 낮은 점, 그리고 ③ 노조 활동에 대한 유무형의 압력이 존재하는 점 등으로 분석하였다.[3] 카카오 노조(별칭 크루) 설립 선언문에도 SW업계에 노조가 없었던 것과 관련해 "개인주의적 분위기 때문이 아니라 탄력적인 사업구조로 인한 불안한 고용환경이 근본적인 원인"이라고 표현하고 있다. 결국, IT/SW업계는 업계 전반의 분위기와 내부 조직 문제로 인해 노동조합의 필요성은 인정되지만, 구성 및 운영 활성화는 어려웠던 것이다.

1 한국경제(2020.3.7.), "정의당 비례대표 1번은 92년생 류호정" 내용 발췌
2 한국노동연구원(2015), "업종별 고용구조 개선과 일자리창출 정책과제 발굴"
3 국회사무처(2013), "IT노동자 근로실태조사 및 법·제도 개선방안"

IT/SW 업계의 기존 노동조합으로는 2003년에 설립된 민주노총의 '한국정보통신노동조합(이하 IT노조)'이 있다. IT노조에는 IT/SW 종사자라면 가입할 수 있고, 정규직, 비정규직 및 프리랜서 등 신분상 제약이 없다. IT노조는 국회와 협업하여 다수의 IT업계 근로환경 실태조사를 진행하기도 하였으며, 주로 IT업계의 노동환경 개선을 위해 목소리를 내고 있다. 특히, SW업계의 임금체불, 부당해고 등에 대한 상담 창구를 운영하며, IT 업계의 근로환경 개선에 대해 정부와 국회에 의견을 개진하고 있다.

최근에는 안랩(Ahnlab)에 노동조합이 설립되어 한국노총에 가입된 상태이다. 안랩 노동조합의 조직화와 더불어 한국노총은 판교 일대에서 꾸준히 홍보활동을 벌이고 있다. 거대 노총에서 SW업계의 노조 조직화에 관심을 가지고 있는 것으로 해석할 수 있다.

[표 1] 한국노총의 IT/SW업계 노동조합 조직화 홍보활동

	일자	홍보 활동 명	내용
한국노총	2019.4.23.	IT노동자 조직화 캠페인 전개	판교 안랩 사거리에서 노동조합 캠페인과 직무스트레스 예방활동 진행
한국노총	2019.6.18.~ 2019.6.21.	판교 무료 법률 상담	판교 유스페이스 앞에서 무료 법률 상담 진행
한국노총	2019.9.18.	IT 노동자 대상으로 직장 내 괴롭힘 예방 및 조직화 캠페인 진행	판교역 일대에서 IT 노동자를 대상으로 '직장 내 괴롭힘 산업재해 예방과 조직화 캠페인'을 진행
한국노총	2019.11.18.	IT노동자 노동환경 실태 및 직무스트레스 개선방안 국회 토론회	한국노총 산업안전보건 연구소에서 실시한 'IT노동자 노동환경 실태와 직무스트레스 실태조사' 결과 공유

최근 SW산업 노동조합의 조직 동기 및 현황

2018년도에 대형 IT/SW기업에서 노동조합 설립 열풍이 있었다. 네이버, 카카오, 안랩 등에서 잇달아 노동조합이 설립되었다. 노조설립 동기에 대한 인터뷰 자료를 살펴보면, ① 기존의 수직적 의사결정에 대한 문제 제기, ② 불안한 고용구조 해결, ③ 근로 환경 개선 등을 위해 노동조합을 조직한 것으로 나타났다.[4] 4개의 IT/SW 대기업 노조 조합원을 대상으로 한 실태조사 결과를 보면 선호하는 의제 1순위로 네이버는 복지제도 개선, 카카오는 공정한 보상 시스템 설계, 게임회사인 넥슨과 스마일게이트는 고용안정성 강화를 꼽았다. 이는 기업의 위치와 규모, 분야 세부 산업의 특징이 반영된 것으로 해석될 수 있다.[5] 또한 동일한 보고서에는 네이버, 넥슨, 스마일게이트 조합원의 선호 의제 2순위는 공정보상 시스템으로 불공정한 시스템에 대해 개선이 필요함을 보여주고 있다.

4 조규준(2019), "2018년 화학섬유식품노조 IT기업 노조조직화의 동기와 특징". 월간노동리뷰 2019년 7월호
5 김종진(2020), "IT·게임산업 특징과 노사관계 과제: 네이버, 카카오, 넥슨, 스마일게이트". KLSI 이슈페이퍼

IT대기업의 노동조합은 안랩을 제외하고는 민주노총의 화학섬유식품연맹(이하 화섬연맹) 산하 지부로 가입되어 있어 산별 노조 형태를 취하고 있으나, 교섭 등 단체행동은 기업별로 이루어지고 있다. 또 하나의 특이점은 화섬연맹 산하의 지부들은 별칭을 가지고 있는데 기업과 노동조합의 특성을 반영한 별칭으로써 젊고 개성이 강한 SW업계 특성을 반영하고 있는 듯하다.

[표 2] IT/SW 기업의 노동조합 조직화 동기

	노조 조직화 동기		인터뷰 내용
1	수직적 의사 결정 문제		"저희 회사 노사협의회가 정상적으로 돌아가고 있지는 않아요. 회사가 일방적으로 정해서 진행하는 방법들이 대부분이었구요."(모 기업 노조 간부)
2	불안한 고용구조		"본질적으로 회사에 오래 있어야 될 이유가 없다라고 생각하는 경우가 좀 많이 있어요… 사람들의 머릿속에 정규직이라는 개념 자체가 없던 거예요. 종신고용이라는 기본적인 개념은 아예 없고, '회사에서 짤렸다. 다른 회사로 빠르게 이직하는 게 나의 능력이다'라고 생각하시는 분들도 계세요. 그런 식으로 왜곡된 생각을 갖고 계신 분들도…"(D기업 노조간부)
3	근로환경 개선	포괄임금제	"업계에서 안 그런 곳이 거의 없어요. 재작년까지만 해도 모든 회사 100%가 다 포괄임금제였고, 굉장히 극소수의 회사들이 재작년부터 폐지하기 시작했어요. 저희도 올해부터 단협으로 폐지했어요."(C기업 노조간부)
		전환배치	"저희 게임업계가 전반적으로 갖고 있는 문제이긴 하지만 프로젝트가 날아가면 사람을 전부 내보내요."(D기업 노조간부)
		정보 비대칭	"저희가 요구하고 있는게 인센티브 산정 근거랑 통계를 공개하라는 것이거든요. 왜 이 사람이 인센티브를 그만큼 받았는지에 대한 그게 없어요. 저희는 평가등급 자체가 없거든요. 아무 것도 없어요. 완전 깜깜이이고 조직별로 그냥 알아서 하는 거에요."(A기업 노조간부)

※ 자료 : 조규준(2019) 자료 발췌하여 정리

외국 SW기업의 국내 지사들도 노동조합의 조직화 바람이 불었으며, 큰 틀에서는 국내 SW대기업 노동조합의 조직화와 차이가 없다. 다만, 국내 기업보다는 조직화 사유가 '고용안정성 강화'라는데 공통적인 방향성이 있는 것으로 보인다. 한국오라클의 노동조합을 보면 '클라우드 중심으로 사업을 개편하면서 구조조정이 있었고, 이에 100여 명의 직원이 비자발적으로 회사를 퇴직한 것'이 노조 조직화의 불씨가 되었다.[6] 2019년 조직화된 SAP노동조합도 같은 이유이다.[7] 외국계 기업의 경우 본국에서와 같이 노동유연성을 이유로 조직개편에 따라 정리해고를 진행하였으며, 해고 위험에 처한 국내 지사의 종사자들은 고용안전성을 이유로 노조를 조직화 하고 있는 것으로 보인다. 이들 노동조합은 민주노총의 사무금융노동조합연맹에 소속되어 활동하고 있다.

6 아이뉴스24(2018.5.5.), "[팩트체크]외국계 IT기업 파업은 어디서 시작됐나 2000년 한국후지쯔 전면파업, 18년 뒤엔 한국오라클 재현" 내용 중 발췌

7 news1(2019.3.7.), "SAP·오라클 사업개편 진통…직원들 고용불안에 노조설립 '맞불'" 내용 중 발췌

[표 3] 2017년 이후 결성된 IT/SW산업 노동조합 현황

구분	회사명	노동조합 명칭	설립 신고	별칭
국내 대기업	카카오	민주노총 전국화학섬유식품연맹 카카오지회	2018년 1월	크루
	네이버	민주노총 전국화학섬유식품연맹 네이버지회	2018년 4월	공동성명
	넥슨	민주노총 전국화학섬유식품연맹 넥슨지회	2018년 9월	스타팅포인트
	스마일게이트	민주노총 전국화학섬유식품연맹 스마일게이트지회	2018년 9월	SG길드
	안랩	한국노총 전국IT사무서비스노련 안랩노동조합	2018년 10월	
외국계 기업	한국마이크로소프트	민주노총 전국사무금융노동조합연맹 한국마이크로소프트노동조합	2017년 7월	
	한국오라클	민주노총 전국사무금융노동조합연맹 한국오라클노동조합	2017년 10월	
	SAP코리아	민주노총 전국사무금융노동조합연맹 SAP코리아노동조합	2019년 3월	

※ 자료 : 조혁진(2020)의 표를 순서 재배열[8]

해외 SW노동조합 현황

해외에서도 IT/SW산업을 대표하는 노동조합보다는 ICT분야 노동조합이 하나의 분과로 조직되어 활동하고 있는 것으로 조사되었다. 이하에서 영국과 일본의 사례를 살펴볼 것이다.

영국 통계청 조사에 의하면 ICT 분야의 노동조합 조직률이 2017년에는 9.5%로 나타났다. 노동조합의 명부등재(Certification Officer[9])에서 IT/SW관련 노동조합을 확인할 수 있다. 우선, CWU(Communication Workers Union)는 약 20만 명의 조합원(조합 자체추산)을 보유한 영국 내 ICT 산업에서 가장 큰 노동조합이다. CWU의 주요 목표는 고용보장달성, ICT산업에서 조합 확장, 산업인프라 발전을 위한 정치적 단결, 모든 형태의 차별에 반대하기 위한 캠페인 등의 의제에 중점을 두고 활동하고 있다.[10] 두 번째로 큰 Prospects 노동조합은 기술전문가들로 구성된 약 14만 2천 명의 조합원이 있고, ICT분야 경우 1만 6천 명의 조합원이 있다. 다른 노동조합과 마찬가지로 임금, 연금, 평등, 안전한 직장 내 근로환경, 불공정한 대우금지 등의 근로환경과 관련된 문제해결을 주요 목표로 한다.

일본 경우 2018년 기준, ICT분야 추정 노조 조직률이 17.2%이고 건설업의 노조조직률이 19.7%(추정)이므로 이를 비교하면, IT 산업의 비교적 높은 노동조합 조직률을 파악할 수 있다.[11] 일본의

8 조혁진(2020), "IT산업의 노사관계 평가와 전망". 월간노동리뷰 2020년 1월호

9 Certification Officer : 노동조합의 명부를 관리하는 기관으로 노동조합 규약서, 간부명단, 본부소재지, 명칭을 제출하여야 한다. 본문의 ICT 부문 노동조합 4개 모두 이곳에 노동조합으로 등재되어 있다.

10 https://www.cwu.org/(접속일 2019.8.29.)

11 일본 노동조합법은 우리나라 노동조합법 제2조 제4호의 단서 라목의 '근로자가 아닌 자의 가입을 허용하는 경우'라는 규정이 존재하지 않기 때문에 노동조합 설립의 주체 요건으로써 노동자 이외에도 학생, 일반 시민이 조합원이 되어도 노동조합성이 부정되지 않아 노동조합 설립과 가입이 자유롭기 때문에 근로자성 다툼이 있는 자들은 연합단체나 직능조합으로 편재되어 활동함으로써 이들의 이익을 대변할 수 있게 된다(국회 환경노동위원회 전문위원, '특수형태근로종사자의 지위 및 보호에 관한 법률안 검토보고서 [조성래의원 대표발의]' 2007.6. 21-22쪽).

노동조합은 일반적으로 정규직 근로자로 구성된 기업별 노동조합 형태가 대다수이다. IT산업 종사자들의 비정규직 비율이 높은 점을 감안하면 기업별 노동조합 형태보다는 산업별 노동조합의 형태로 조직되고 있는 것으로 추정할 수 있다.

일본에서 ICT분야 산별노조로서 정보산업노동조합연합회(이하 정보노련)[12]가 조직되어 있다. 정보노련 산하에는 약 257개 조합이 있으며, 조합원 수는 약 20만 명(2018년 6월 기준)에 달한다. 정보노련은 정보통신, 정보서비스, 통신건설, 인쇄, 운수, 제조, 건축, 빌딩 유지관리 등의 업종을 포괄하고 있다. 노조의 주요 의제로는 조합원의 경제적 및 사회적 지위 향상을 도모하는 것을 목적으로 하고 있다. 그 외에도, 전일본 전기·전자·정보관련산업노동조합연합회(이하 전기렌고)[13]가 있다. 전기렌고는 231개 조합이 가입되어 있으며 조합원 수는 약 57만 명이다.(2019년 6월 기준) 전기렌고 조합은 전기, 전자, 정보관련 산업의 노동조합을 포괄하는 조직이다.

IT산업과 관련된 커뮤니티 유니온 형태(산별노조의 형태로 보임)로는 "IT UNION"이 있다.[14] 일본도 IT산업의 중층적 다단계 하도급 구조가 형성되어 있으며, 여기에 종사하는 비정규직 근로자들이 많은데 이들도 IT 유니온에 가입할 수 있다. IT 유니온은 첫째, 위장도급, 다중파견, 둘째, 장시간 근로, 위법한 재량근로제, 셋째, 프리랜서 노동문제 등에 대응하고 있다.

[표 4] 해외 국가 노동조합 사례

국가	노동 조합명	조합 가입 분야 또는 대상	주요 의제
영국	CWU	영국 내 ICT 산업 부문에서 가장 큰 노동조합으로 ICT산업에 해당하면 가입 가능	고용안정 강화, ICT분야 노조연대, 달성, 산업 인프라 발전을 위한 정치적 단결, 모든 형태의 차별에 반대
	Prospects	기술 전문가가 주축이 된 노동조합으로 ICT분야의 조합도 존재	근로환경, 불공정한 대우금지 등의 근로환경과 관련된 문제해결을 주요 목표
일본	정보산업 노동조합연합회	정보통신, 정보서비스, 통신건설, 인쇄, 운수, 제조, 건축, 빌딩 유지관리 등의 업종	조합원의 경제적 및 사회적 지위 향상을 도모
	전기·전자·정보관련 산업노동조합연합회	전기, 전자, 정보관련산업 및 그 관련 산업	근로환경 개선, 최저임금 향상, 일가정 양립 지원 등
	IT UNION	IT분야 산업 종사자	첫째, 위장도급, 다중파견, 둘째, 장시간 근로, 위법한 재량근로제, 셋째, 프리랜서 노동문제 등에 대응

12 정보노련(정보산업노동조합연합회) https://www.joho.or.jp/
13 전일본전기·전자·정보관련산업노동조합연합회(全日本電機·電子·情報関連産業労働組合連合会) https://www.jeiu.or.jp/what/organization/cat4/
14 IT UNION https://it-union.jp/

근로환경 개선과 관련한 SW업계 종사자들의 집단적 의사표시

2018년 근로기준법 개정으로 시행된 '근로시간 단축제도(주52시간 근무 상한제도)'와 관련해서 SW산업 전반이 뜨거웠다. 주요 쟁점은 유연근로제 확대와 관련된 문제였다. 국회는 IT/SW산업의 특수성을 반영하여 선택근로제 정산 기간 및 탄력근로제 산정 단위 기간을 확대하려 한 것이다. 이와 관련해서 협회 등 사업자 단체는 더 적극적으로 위 기간을 늘리고자 하였다. 민주노총 등 노동자 단체는 노동자의 건강권 등을 이유로 반대 입장을 표명하였고, 양쪽의 절충안을 마련하지 못한 채 국회의 회기가 종료하였다. 즉, 국회에서 해당 안건은 통과되지 못한 채로 마무리되었다. 결국, 정부는 SW산업을 특별연장근로 허용 업종으로 공식화하여 사업자 단체의 요구를 일단락 지었다.

이 과정에서 한국SW·ICT총연합회(KSWICT)를 필두로 한국소프트웨어산업협회(KOSA), 한국IT서비스산업협회(ITSA) 등이 사업자의 입장을 대변하였다. 노동자 입장에서는 산별 노조 형태인 민주노총 소속의 IT노조가 반대 목소리를 내었다. 그러나, 이번에는 다른 점이 있었다. 화섬연맹의 네이버·넥슨·스마일게이트·카카오 지회 등이 연대하여 공식적으로 의견을 표명한 것이다. 기업별 노조임에도 불구하고 산업 전반을 위해 목소리를 낸 첫 행보였다. 기자회견 내용 중에는 "IT업계는 고질적인 하청 구조로 인한 저임금 노동과 납기일을 맞추기 위해 야근을 밥 먹듯 하는 장시간 노동이 만연해 있다."라는 내용이 있다. 위에서 살펴본 바와 같이 노동조합의 조합원들은 고용 안전 강화, 공정한 평가 시스템 도입, 복지제도 개선에 대한 의제에 관심이 높지만, 정부를 향해 근로자 처우에 대해 목소리를 낸 것은 업계에 새로운 조짐이 있음을 보여준다.

시사점

SW산업 종사자들은 노동조합의 조직과 활동에 기대감이 없었고, 열악한 근로환경을 근로자가 각개격파 해왔다. 최근 대형 및 외국계 IT/SW기업을 중심으로 기업단위 노동조합이 설립되면서 기업 내에 머물던 목소리가 기업 밖 산업, 정부 등으로 퍼지고 있다. 이와 관련해서 국내 양대노총인 민주노총과 한국노총에서 다양한 방식으로 활동하고 있다. 앞으로 이러한 활동이 계속된다면, 그동안 산업에서 고질적인 문제로 지적되어왔던 불법파견, 장시간 근로 등의 문제에 대해서 IT/SW업계의 일방적인 목소리가 아닌, 산업에 종사하는 개발자들의 목소리가 생생하게 반영될 수 있을 것이다. 2019년도에는 대형 IT/SW기업의 노동조합들이 서로 연대하여 IT/SW업계의 근로시간과 관련된 법제도 추진에 대해서 업계와 정부에 IT/SW산업 근로자 전체를 대변하여 목소리를 낸 바 있다. 근로환경의 개선과 관련해서 정부와 업계의 노력도 반드시 필요하지만, IT/SW업계의 종사자를 대표하는 노동조합과의 대등한 협상대상자로서 대화가 필요하다. 그러므로 정부와 업계는 노동조합을 통해 불고 있는 잔잔한 파동에 대해 지속적인 관심과 연구가 필요하다.

에듀테크 산업 동향 및 시사점

Edutech Industry Trends and Implications

공영일 Kong, Youngil ▌책임연구원 Principal Researcher, SPRi ▌kong01@spri.kr

AI, AR/VR, 블록체인 등의 기술을 기존 교육에 접목한 에듀테크가 주목을 받고 있다. 에듀테크는 디지털전환이 더뎌 잠재력이 큰 시장으로 평가되고 있다. 최근 코로나19 사태로 우리사회는 온라인 개학 등 비대면 교육이라는 피할 수 없는 선택과 마주하게 되었다. 교육부문의 디지털전환에 대한 강력한 사회적 공감대 형성, 교육현장의 문제해결을 위한 에듀테크의 다양한 시도, 그리고 이를 수용할 수 있는 제도의 개선이 뒷받침이 된다면 우리는 코로나19 사태로 인한 교육의 위기를 교육분야의 디지털전환을 선도할 수 있는 기회로 만들 수 있을 것이다.

Edutech, combining education with such technologies as artificial intelligence(AI), augmented reality(AR), virtual reality(VR) and blockchains, has been gaining attention. Edutech is considered a market with huge potential as its digitalization is slow in progress. The recent Covid-19 crisis has led our society to make an inevitable choice of contactless education by reopening schools online. It can be an opportunity for the country to take the lead in transforming education into a digitalized one, if we find common ground in society toward digitalization of education, make various edutech to tackle problems in education, and improve institutions to make those changes happen.

에듀테크(Edutech), 교육에 ICT기술을 접목해 기존 서비스를 개선하거나 새로운 서비스를 제공하는 것

에듀테크(Edutech)는 교육(Education)과 기술(Technology)의 합성어로, 교육에 ICT기술을 접목해 기존 서비스를 개선하거나 새로운 서비스를 제공하는 것, 또는 교육 서비스를 개선하거나 새로운 가치를 제공하는데 활용되는 기술을 의미한다. 최근에는 인공지능, AR/VR, 블록체인 등이 기존 교육을 혁신시키는 기술로 주목을 받고 있다.

에듀테크와 유사한 개념으로 이러닝(E-Learning), 스마트 러닝(Smart Learning) 등이 있다. 이러닝은 1990년대 말 인터넷의 급속한 확산으로 등장한 개념으로 주로 디지털교과서와 온라인 학습에 방점이 주어졌다. 2010년경에는 스마트폰과 태블릿이 본격적으로 확산됨에 따라 스마트기기를 통해 교육 성과를 높이기 위한 스마트러닝이 업계 및 학계의 주목을 받았다. 이러닝과 스마트러닝의 경우 교육이 행해지는 학습수단(인터넷+데스크탑, 스마트폰, 태블릿 등)에 주요한 개념의 방점이 주어졌다. 에듀테크의 경우는 이러한 학습 수단에 대한 것(AR/VR)도 있지만, 인공지능, 빅데이터, 블록체인 등 데이터와 소프트웨어를 기반으로 학습자에 대한 분석과 의사소통, 정보관리를 용이하게 함으로써 학습 성과를 제고시키는 방향에 좀 더 무게중심이 실려 있다.

[표 1] 에듀테크의 개념과 유사 개념

구분	개념	특징
이러닝	전자적 수단, 정보통신 및 전파·방송기술을 활용하여 이루어지는 학습[1]	인터넷과 컴퓨터에 교육을 접목한 온라인 교육중심
스마트러닝	스마트폰, 태블릿PC, E-Book 단말기 등 스마트 디바이스와 이러닝 신기술이 융합된 개념[2]	스마트 기기를 활용한 교육
에듀테크	교육에 ICT기술을 접목해 기존 서비스를 개선하거나 새로운 서비스를 제공하는 것	데이터와 소프트웨어에 무게중심

에듀테크 시장, 디지털전환이 상대적으로 더딘 편

에듀테크의 기본 환경인 세계 교육 시장은 2020년 6조 5천억 달러 수준에서 2025년 8조 1천억 달러, 2030년에는 10조 달러에 달할 것으로 전망되고 있다. 이 중 에듀테크 시장은 2018년 1,530억 달러에서 2025년 3,420억 달러의 시장규모를 형성할 것으로 예측되고 있다. 특기할 만한 점은 전체

1 이러닝산업발전법 제2조
2 한경 경제용어사전

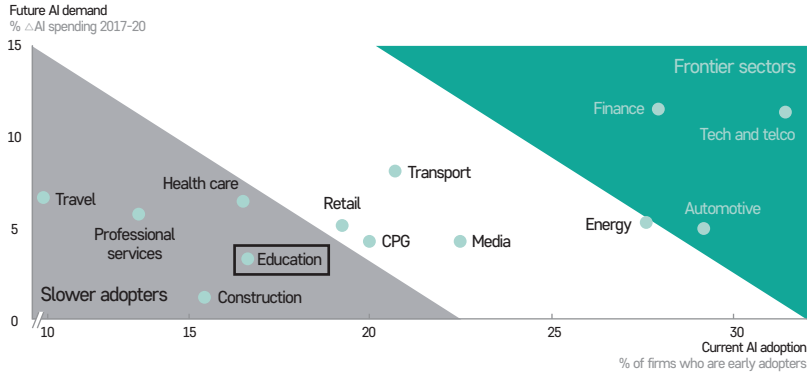
교육시장에서 에듀테크가 차지하는 비중이 낮은 수준이라는 점이다. 2018년 기준으로 에듀테크가 교육시장에서 차지하는 비중은 2.5%에 불과하다. 2025년에도 이 비중은 4.3%에 그칠 것으로 전망된다.

[그림 1] 세계 교육시장 규모 추세(좌)와 에듀테크 시장 규모(우)

※ 자료 : Holon IQ(2020.1.28.)

이는 교육부문의 디지털전환이 다른 부문에 비해 상대적으로 더디게 진행되고 있음을 시사한다. 다음의 그림에서 볼 수 있듯이 교육분야의 AI 도입률도 다른 분야에 비해 상대적으로 낮은 것을 알 수 있다.

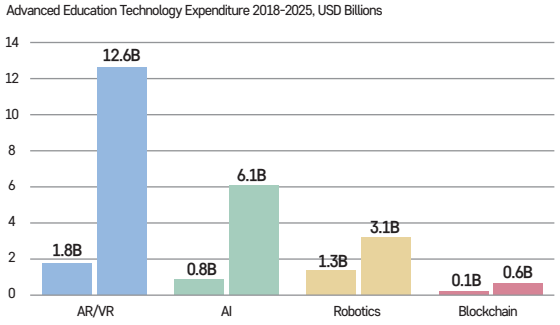
[그림 2] 주요 분야별 AI 도입 현황과 전망



※ 자료 : McKinsey Global Institute(2018)

교육분야에 적용되는 주요 기술은 AR/VR, AI, 로보틱스, 블록체인 등으로 파악된다. 체험 기반의 실감형 교육을 가능하게 하는 AR/VR 기술의 경우, 2018년 18억 달러에서 2025년 126억 달러로 7배의 높은 성장을 보일 것으로 전망된다. 빅데이터를 통해 학습자 맞춤형 교육, 교수법 및 행정관리 자동화와 지능화를 가능케하는 인공지능도 2018년 8억 달러에서 2025년 61억 달러 규모의 시장을 형성할 것으로 전망되고 있다. 블록체인도 높은 성장을 보일 것으로 전망(6배 성장)되나 다른 기술에 비해 상대적으로 작은 시장규모를 유지할 것으로 전망된다.

[그림 3] 교육분야의 주요 기술 시장 전망



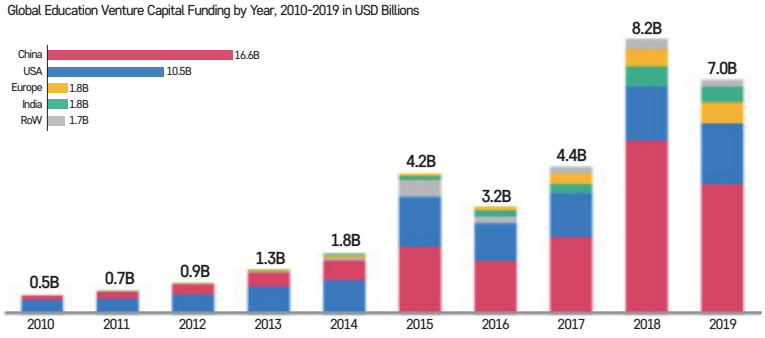
※ 자료 : Holon IQ(2020.1.28.)

벤처캐피탈(돈)은 에듀테크를 어떻게 보는가?

[그림 4]는 교육분야 글로벌 벤처캐피탈 투자 추세를 보여주고 있다. 교육분야의 벤처캐피탈 투자는 2015년과 2018년 전년대비 2배에 달하는 급격한 성장을 보였다. 2019년 투자금액은 70억 달러에 달하는 것으로 파악된다.

국가별로 2014년까지는 미국이 주도를 했고 2015년부터는 중국이 글로벌 벤처 투자를 주도하고 있음을 알 수 있다. 지난 10년 동안 교육부문의 글로벌 벤처투자의 약 50%를 중국이 차지하고 있다. 이어 미국이 33%, 유럽, 인도가 각각 5%를 차지하고 있다.

[그림 4] 교육분야 글로벌 벤처캐피탈 펀딩 추세



※ 자료 : Holon IQ(2020.1.28.)

기업가치가 10억 달러 이상으로 평가되는 비상장 기업을 유니콘 기업이라고 한다. 에듀테크 유니콘 기업은 2020년 1월 기준 14개가 있는 것으로 파악된다. 2011년에 설립된 ByJu's는 인도 방갈로에 기반을 둔 온라인교육 서비스 기업이다.[3] ByJu's는 2020년 1월, 타이거 글로벌(Tiger Global)이라는 벤처캐피탈로부터 2억 달러의 투자를 유치함으로써 기업가치가 80억 달러로 평가되고 있다.[4]

[표 2] 글로벌 에듀테크 유니콘 기업(2020년 1월 기준)

기업명	국가	분야	최근 라운드	최근 라운드 종류	가치 평가액
ByJu's	India	Tutoring	Jan 2020	PE Round	$8B
VIPKid	China	Language	Sep 2019	Corporate Round	$4.5B
Yuanfudao	China	Tutoring	Dec 2018	Series F	$3.0B
Duolingo	US	Language	Dec 2019	Series F	$1.5B
Guild Education	US	UpSkilling	Nov 2019	Series D	$1.0B+
Knowbox	China	Tutoring	May 2019	Series D	$1.0B+
Coursera	US	MOOC	Apr 2019	Series E	$1.0B+
iTutorGroup	China	Language	Feb 2019	Corporate Round	$1.0B
Zhangmen	China	Tutoring	Jan 2019	Series E	$1.0B+
Huike	China	OPM	May 2018	Series D	$1.0B
17zuoye (一起作业)	China	Tutoring	Mar 2018	Series E	$1.0B
Age of Learning	US	Online Curriculum	May 2016	Series A	$1.0B
Udacity	US	Proprietary Online	Nov 2015	Series D	$1.1B
HuJiang	China	Online Curriculum	Oct 2015	Series D	$1.0B

※ 자료 : Holon IQ(2020.1.), Wikipedia의 자료를 기반으로 작성

다음의 [표 3]에서와 같이 에듀테크 유니콘의 수와 기업가치를 다른 분야와 비교해보면 개수도 상대적으로 적고 가치평가액도 작은 것으로 파악된다. 이는 교육부문의 디지털화가 상대적으로 더딘 것과 관련이 있는 것으로 보인다. 한편, 이들 14개 기업에 대한 벤처캐피탈의 총 투자액은 66억 달러로 추정된다.[5]

3 https://en.wikipedia.org/wiki/BYJU%27S
4 https://tech.economictimes.indiatimes.com/news/internet/tiger-global-invests-200m-in-byjus-at-8-billion-valuation/73169591
5 https://www.holoniq.com/edtech-unicorns/

[표 3] 주요 분야별 유니콘 기업 개수 및 가치평가 금액 비교

분야	핀테크	인공지능	헬스케어	에듀테크
유니콘 기업 개수	61개	46개	33개	14개
가치평가 금액	1,992억 달러	1,723억 달러	862억 달러	250억 달러

※ 자료 : CB Insights(2020.4.), Holon IQ(2020.1.) 자료를 기반으로 작성

코로나19 사태로 인한 교육의 위기를 에듀테크 활성화의 기회로

이상으로 에듀테크 산업에 대한 글로벌 차원의 큰 흐름에 대해 논의했다. 요약하면 다음과 같다.

- 이러닝, 스마트러닝은 학습수단(인터넷+컴퓨터, 스마트기기)에 중점을 두고 있다. 반면 에듀테크는 학습 수단에 대한 것(AR/VR)도 있지만, 인공지능, 빅데이터, 블록체인 등 데이터와 소프트웨어를 기반으로 학습자에 대한 분석과 의사소통, 정보관리를 용이하게 함으로써 학습의 성과를 제고시키는 것에 무게중심이 실려있다.

- 에듀테크의 기본 환경인 세계 교육 시장은 규모가 크지만 다른 분야에 비해 디지털전환이 더디게 진행되고 있다. 교육에 접목되는 기술로는 AR/VR과 인공지능이 상대적으로 유망할 것으로 전망된다.

- 교육분야의 글로벌 벤처캐피탈 투자는 2015년, 2018년에 전년대비 약 2배에 달하는 급격한 성장을 보였다. 그러나 핀테크, 인공지능, 헬스케어 등 다른 분야와 비교해 볼 때 유니콘 기업 개수나 기업가치 측면에서 상대적으로 뒤쳐져 있다.

최근 전 세계는 코로나19로 인해 전대미문의 상황을 마주하고 있다. 여러 측면에서 코로나19 사태는 우리 삶에 큰 영향을 미칠 것으로 예상된다. 일하고, 여가를 보내고, 소비하고, 배우고 가르치는 등 일상의 모든 방식과 행태에서 비대면(Untact)의 중요성과 비중이 더 커질 것으로 예측되고 있다. 비대면 방식의 확산추세가 뚜렷하게 나타날 분야 중 하나가 교육부문이다. 유례없는 초중등 및 대학의 온라인 개학으로 우리사회는 비대면 교육이라는 피할 수 없는 선택을 마주하게 되었다. 우리는 비대면 환경에 학생, 교사, 학부모, 교육당국 등 모두가 준비가 되어 있지 않다는 현실을 인지하고 있다.

위기(危機)는 본질적으로 기회(機會)를 품고 있다. 코로나19 사태로 인한 방역의 위기를 우리는 'K방역'이라는 방역시스템을 세계에 확산시킬 수 있는 기회로 만들어 가고 있다. 이번 코로나19 사태는 디지털전환 시대의 우리나라 교육현실과 문제점을 적나라하게 드러내고 있다. 그러나, 교육부문의 디지털 전환에 대한 강력한 사회적 공감대 형성, 교육현장의 문제해결을 위한 에듀테크의 다양한 시도, 그리고 이를 수용할 수 있는 제도의 개선이 뒷받침이 된다면 우리는 코로나19 사태로 인한 교육의 위기를 교육 디지털화를 선도할 수 있는 기회로 만들어 갈 수 있을 것이다.

실감산업, 큰 그림을 그리자

Time to Broaden the Horizons of the Immersive Industry

한상열 Han, Sangyeal ❙ 선임연구원 Senior Researcher, SPRi ❙ syhan17@spri.kr

2016년 출시된 가상현실(Virtual Reality, VR) 헤드셋, 오큘러스 리프트(Oculus Rift)와 HTC 바이브(Vive)는 그 전과는 차원이 다른, 한마디로 “심장이 쫄깃”해지는 충격적인 경험을 선사하였고 가상현실 시장은 뜨겁게 달아올랐다. 이어서 출시된 증강현실(Augmented Reality, AR) 기반의 포켓몬 고(Pokémon GO) 게임도 출시 반년 만에 누적 다운로드 수 6억 회를 달성하는 폭발적인 인기를 누렸다. 바야흐로 실감기술(Immersive Technology) 대중화 시대의 막이 열렸다는 기대가 한껏 높아졌었다. 그러나 장시간 사용 시의 어지러움, 무게감으로 인한 착용의 불편함, 3D 콘텐츠 부족을 포함한 여러 한계로 인해 실감기술은 아쉽게도 일부 게이머(Gamer)들만의 전유물로 남겨지는 듯했다. 하지만 최근 본격화된 실감기기의 기술적 진전과 게임 외의 교육, 제조, 재난 등의 다양한 분야에서 활용사례가 늘어나면서 실감기술 재도약의 기대감이 높아지고 있다.

최근 출시된 실감기술은 이전 세대 기술과 분명히 구분된 진전을 보이고 있다. 과거 PC 기반의 VR 헤드셋(Headset)은 고사양의 PC 연결이 필요하므로 주로 고정된 장소에서만 사용이 가능했다.

하지만 최근에는 PC에 연결하지 않고 독립적인 구동이 가능한 무선 올인원(All-In-One) 형태의 VR 헤드셋들이 출시되면서 어디서든 가상현실 접속이 가능해지고 있다. 무게 측면의 진전도 뚜렷하다. 2020년 CES(Consumer Electronics Show)에서는 무거운 헤드셋 형태가 아닌, 무게도 가벼워지면서 착용감이 한결 개선된 안경형태의 VR 글래스(VR Glasses) 시제품이 등장하였다. 또한 가상현실에서 느낄 수 없는 물리적 감각을 더하기 위한 다양한 햅틱(Haptic)기기들도 전시되었다. 예로서 테슬라수트(Teslasuit)社의 햅틱기기는 장갑처럼 손에 착용하면 가상물체를 손에 쥐는 듯한 물리적 촉감까지 전달해준다. 이외에도 신발 형태, 밴드 형태의 다양한 햅틱기기를 통해 더욱 실감나는 사용자 경험이 가능해지고 있다. AR 분야에서도 안경형태의 AR 글래스(AR Glasses)와 햅틱기기 신제품들이 개발되고 있다.

실감기술이 발전하면서 활용 분야도 다양해지고 있으며 구체적인 성과도 나타나고 있다. 교육 분야에서는 실감기술을 활용할 때 학습효과가 상승하고 교육시간이 절감되는 성과가 나오고 있다. 글로벌 컨설팅社인 액센츄어(Accenture) 보고서에 따르면, VR을 이용한 수술교육을 실시한 사례에서 수술 중 실수가 기존보다 40% 감소한 바 있으며, 소매 산업에서는 실감기술을 활용하여 교육시간을 80%까지 단축시킨 사례도 있다. 제조업 등의 업무현장에서는 재직자의 업무지식을 AR로 기록하여 교육하는 방법을 도입함으로써 문서작성 시간과 교육시간을 줄일 수 있다. 의료 산업에서는 VR을 활용한 트라우마 치료 효과가 가시화되면서 의학교육 및 훈련, 재활 치료 분야에서 실감기술 활용사례가 증가하고 있다. 국방, 재난·안전 등의 공공 분야에서도 가상환경의 다양한 시나리오를 훈련하면서 실제상황에 대한 대응력을 높일 수 있다.

코로나19 확산 같은 특정 상황으로 인해 현실공간에서 충족시키기 어려운 수요가 생길 때도 실감기술은 효과적 대안으로 활용될 수 있다. 실감기술을 활용한 문화생활뿐만 아니라 가상공간을 활용한 원격회의, 쇼핑 같은 다양한 응용사례들을 만들 수 있다. 실감기술을 이용하면 원격회의의 몰입감이 더욱 증가하고 3D 모델을 통한 원거리 설계가 가능하다. 다수가 모여야 하는 신제품 품평회도 가상공간에서 진행할 수 있다. 향후 비대면(Untact) 방식이 보편화된다면 다양한 분야에서 실감기술 활용수요가 증가할 것으로 예상된다.

실감기술의 지속적인 개선과 발전, 교육, 의료, 국방, 재난·안전 등 다양한 분야에서 실감기술과의 융합이 본격화되면서 실감기술은 이제 국가의 새로운 성장을 이끄는 신성장동력으로 주목받고 있다. 영국 Innovate UK는 2018년 발표한 정책보고서 “The Immersive Economy in the UK”를 통해 전 산업으로 확대되는 실감기술의 영향력을 고려할 때 이제는 실감기술이 산업혁신과 새로운 혁신의 가치창출에 기여하는 공통된 핵심기술, 즉 범용기술(General-Purpose Tech.)에 속한다고 보았다. 이에 실감기술로 경제, 사회, 문화적 가치를 창출하는 “실감경제” 개념을 제시하고 정책역량을 집중하고 있다. 미국, 중국, 일본 등 다른 주요국들도 이미 실감기술 개발에 지속적인 투자를 하며 국방, 교육, 제조 등의 다양한 공공·산업 분야와의 융합을 추진하면서 경제성장을 위한 정책을 추진 중이다.

우리 정부도 2019년 9월「콘텐츠산업 3대 혁신전략」, 2019년 10월「실감콘텐츠산업 활성화 전략」 발표를 통해 국내 실감콘텐츠 산업발전을 위한 거시적 관점의 정책지원 방안을 마련하였다. 이제는 산업·공공분야에서 실감기술이 촉발할 혁신의 기회들을 국민들이 체감할 수 있는 구체적인 성과로 실현하고 효과성을 높이기 위한 후속 계획을 마련해야 할 시기이다. 국내 실감산업은 해외에 비하여 규모가 작고 핵심기술력이나 인재역량 측면에서 부족한 부분이 많다. 반면에 우리나라는 실감산업 발전에 필요한 이동통신 인프라와 제조·디스플레이 기술 분야에서 강점을 보유하고 있다. 세계최초로 상용화한 5G 인프라를 기반으로 산업·공공분야에서 새로운 시장을 창출하려는 정부의 정책적 의지도 강하다. 실감기술의 기회와 우리의 강점을 잘 활용한다면 경제적 불확실성이 증대되는 현 위기상황을 극복하고 혁신성장의 비전을 이룰 수 있는 신성장동력 창출이 가능하지 않을까?

실감기술의 지속적인 발전과 활용 효과 가시화에 따라 실감기술과 산업간 융합의 범위는 계속 확대되고 산업과 사회를 변화시키는 속도는 더욱 빨라질 것이다. 실감기술의 영향력을 단지 게임이나 일부 파생영역으로 한정짓는다면 우리는 다가오는 큰 변화와 발전의 기회를 놓칠 수 있다. 실감기술의 발전과 이로 인한 경제적, 사회적 변화에 대한 면밀한 관찰, 한 세대 이상의 미래를 내다보는 선제적 대응 관점에서 큰 그림을 담은 국가적 대응전략을 준비할 시기이다.

칼럼2 COLUMN

디지털 전환 시대에 필요한 新교육·인재 전략

New Education and Talent Strategies Needed in the Era of Digital Transformation

서영희 Seo, Younghee ▮ 선임연구원 Senior Researcher, SPRi ▮ yhseo@spri.kr

최근 제4차 산업혁명의 핵심인 AI, 빅데이터, 클라우드 등 SW 신기술의 융합으로 빅블러(Big Blur) 현상이 나타나고 있다. 빅블러 현상은 SW 신기술 발전에 따라 기존 산업 간, 온·오프라인 간 경계가 모호해지는 현상을 의미한다. 아마존은 이러한 빅블러 현상을 대표한다. 1994년 온라인서점으로 출발한 아마존은 전자상거래, 음악 스트리밍, 미디어 유통으로 다각화하고, 최근 배송, 무인점포, 금융 및 클라우드 기업으로 발전하였다. 검색 서비스로 시작한 구글 역시 모바일 플랫폼, 광고, AI 스피커 산업에서도 큰 성과를 내고 있으며, 자회사인 웨이모(Waymo)는 2019년 기준으로 자율주행기술기업 1위를 차지하는 등 빅블러 현상의 대표적 사례이다.

한편, 코로나19로 인해 글로벌 경제 위기가 지속되는 가운데 오프라인 관련 산업의 피해는 커지고 있으나, 온라인 상거래나 스마트워크 협업 서비스 등 온라인 서비스 관련 시장의 성장세가 가속화되는 추세이다. 미국 내 온라인 주문의 39%를 차지하고 있는 아마존은 코로나19로 인한 주문량 폭주로 인해 직원 10만 명을 추가 고용하기로 결정하였다. 온라인 협업 서비스를 제공하는 국내 기업인 알서포트의 경우, 3월 첫째 주의 영상회의 수가 지난 1월 초에 비해 1,924% 정도 증가[1]하였다고

1 IT데일리(2020.3.), 코로나19 감염 확산에 재택근무 관련 솔루션 수요 급증

밝혔으며 웍스모바일이 개발한 업무용 협업 도구인 라인웍스는 1월 20일 대비 다자간 영상통화는 28배, 음성통화는 25배, PC 화면 공유는 15배 이상 상승한 것으로 나타났다.[2]

이와 같이 급변하는 환경에 대비하여 국가 경쟁력을 제고하기 위해서는 과거와는 다른 새로운 핵심역량을 가진 인재들이 요구된다. 세계경제포럼(WEF)은 2016년, 제4차 산업혁명 시대 인재의 핵심역량으로 복합적인 문제해결 능력, 창의력, 협업, 디지털 리터러시를 제시하였다.[3] 디지털 리터러시는 디지털화된 정보를 평가 및 판단하고 자신에게 필요한 정보를 선택, 편집 및 가공하여 새로운 지식을 창출하는 능력을 말한다. 새롭게 강조되는 역량은 단순지식을 습득하는 능력보다 디지털 기반의 컴퓨팅 사고력과 문제발견 및 해결능력, 융합 및 창의력이 포함된다. 미래 일자리의 경우, 인공지능 및 빅데이터 기술의 발달로 인해 단순·반복적 업무는 자동화되거나 대체될 것이다. 또한, 데이터 분석가나 VR/AR 콘텐츠 개발자 등 전문성과 창의력이 요구되는 일자리의 수요는 증가할 것으로 예측되기도 하였다.[4]

이러한 미래 SW 역량을 갖추기 위해 해외 선도국은 SW를 디지털 전환의 핵심동력으로 인식하고 대규모 SW 인재양성을 통한 국가 경쟁력 강화 정책을 추진하고 있다. 이제는 SW 코딩뿐만 아니라 컴퓨팅 사고력 기반의 SW융합 교육 및 인공지능 교육을 적극적으로 도입하는 추세이다. 아래 표와 같이 미국, 영국, 중국은 앞 다투어 국가 차원에서 SW 및 AI 인재양성을 위해 관련 계획을 실행하고 있다.

[표 1] 국가별 SW·AI 주요 정책 및 상세 내용 정리

국가	주요 정책	상세 내용
미국	모두를 위한 컴퓨터과학, Computer Science for All (2016)	유치원생-고등학생의 컴퓨터과학 교육 실시
	연방 5개년 STEM 교육 전략 계획, Federal 5-Year STEM Education Strategic Plan(2018)	STEM(Science, Technology, Engineering, and Mathematics) 교육을 위해 향후 5년간 매년 최소 2억 불(약 2,500억 원)을 투자
	미국 AI 이니셔티브, The American AI Initiative (2019)	전문인력 양성 및 전 국민 STEM 교육 확대
영국	컴퓨팅 교육 의무화 의결(2014)	모든 초·중·고등학교(만 5세~16세)에서 컴퓨팅 교육을 필수과목으로 지정하여 수업 시작
	영국 AI 산업 발전(2017)	전문역량 공급을 개선하기 위한 6개의 권고사항 제시
	산업 전략-AI 분야 합의안(Industry Strategy-Artificial Intelligence Sector Deal(2018)	1억 파운드(약 145억 원)의 AI 장학금과 STEM 관련 교육에 4.06억 파운드(약 6천억 원) 투자[5]

2 라인웍스(2020.4.), 코로나19 이후 화상회의 28배 증가, ZDNet Korea
3 New Vision for Education : Fostering Social and Emotional Learning through Technology., World Economic Forum(2016)
4 소프트웨어정책연구소(2019), 미래 직업 예측 모델 개발 연구
5 한국보건산업진흥원(2019), 주요 국가별 인공지능(AI) 인력양성 정책 및 시사점

국가	주요 정책	상세 내용
중국	대학 AI 혁신 행동계획(2018)	2030년 세계 최고의 AI 혁신센터를 목표로 5개년 단위 인재육성 계획 수립
	대학 AI 인재 국제양성계획(2018)	5년 내 인공지능 교수 500명, 학생 5,000명 육성 계획 수립

우리 정부도 부처별로 SW 신기술 관련 인재양성을 위한 노력을 기울이고 있다. 대표적으로 과학기술정보통신부는 SW 신기술 분야 인재양성을 위한 사업을 추진하고 있고, 2020년부터 SW와 AI 역량을 높이기 위해 SW·AI 인재양성을 본격적으로 강화한다고 밝혔다. 교육부는 미래 산업을 위한 핵심 인재 및 AI 융합 교육을 위한 지도자 양성을 위한 사업을 진행하고 있다. 그리고 고용노동부는 취·창업자 및 재직자 대상 SW 신기술 교육 사업을 수행하고 있다. 아래 표와 같이 부처별로 역할을 분배하여 SW 인재 양성을 위한 교육 사업을 진행하고 있으나 부처별 사업에 대한 연계 및 중복을 고려한 협력 체계는 아직 부족한 것이 현실이다.

[표 2] 부처별 사업 및 특징

부처	사업	주요 특징
과학기술정보통신부	이노베이션 아카데미	SW 신기술 분야 인재 양성
	AI대학원	
	SW스타랩	
교육부	LINC+ 4차 산업혁명선도대학	미래 산업과 지역 사회가 필요로 하는 핵심 인재 양성 및 AI 융합 교육의 지도력 강화
	AI융합교육대학원	
고용부	4차 산업혁명 선도인력 양성	창업자 및 재직자 대상 교육 지원 사업 수행
	국가 기간·전략 산업 직종 훈련	
	국민 내일배움카드	

가속화되고 있는 디지털 전환에 우리 나라가 적절히 대응하기 위해서는 무엇을 준비해야 할까? 먼저 범부처 차원의 새로운 「SW·AI 인재 양성 혁신전략」을 마련하여 한 방향으로 자원을 모으고 나아가기 위한 국가차원의 청사진이 필요하다. 이러한 전략에는 현재 각 부처에서 개별적으로 이루어지고 있는 다양한 인력양성 사업에 대한 협력 연계방안(과기정통부-교육부-고용노동부-산업부 등)뿐만 아니라, 직업의 변화속도와 규모, 다양성에 대응할 수 있는 직업훈련, 재교육, 생애 전 주기에 걸친 AI·SW 평생교육 체계 등이 포함되어야 할 것이다. 그리고 전 국민의 AI·SW 기초 소양에 대한

교육 및 개인의 수준별 맞춤형 교육을 위한 전 생이주기 차원의 교육 플랫폼 구축이 필요하다. 미국은 다양한 수준 및 대상별 수요를 반영한 SW 교육 플랫폼인 유다시티(Udacity), 코세라(Coursera), 에덱스(edX) 등의 교육 플랫폼을 구축하고, 협업 및 실습 기능을 제공하는 등 현장에서 활용할 수 있는 실무 능력을 키우기 위해 학습범위를 확대하고 있다.

우리도 전 국민대상의 수준별 맞춤형 AI·SW 교육을 준비하기 위해서는 관련 부처가 동주공제(同舟共濟)의 마음으로 협력해야 할 것이다. 각 부처는 협력 및 연계방안에 관한 구체적인 액션플랜을 통해 개인의 생애주기별 맞춤형 교육체계를 구축하여 전 국민의 AI·SW 역량이 강화되고, 더 나아가 우리의 미래 경쟁력이 향상될 수 있기를 기대한다.

SPRI AI BRIEF

인공지능 최신 동향과 시사점

Recent Trends of AI and Their Implications

Table of Contents

- AI로 코로나바이러스 구조 분석, 바이두 등 대응책 개발 박차
- 어그테크 기업, AI+클라우드 날개 달고 고공 성장 전망
- 딥페이크를 선용하는 사례 속속 등장
- 美 에너지부, AI for Science 보고서 발표
- IDC, 코로나19로 인한 경기침체 속 AI 시장 및 고용 전망 상향 조정
- 코로나19 치료를 위한 AI 성과 속 한계 지적
- 백악관, 코로나19 학술 정보 분석 AI 경진 대회 개최
- 구글, AI 칩 설계에 AI 활용 시도

AI정책연구팀 jayoo@spri.kr 031-739-7352

AI로 코로나바이러스 구조 분석, 바이두 등 대응책 개발 박차[1]

1 코로나19의 급격한 확산에 따라 전 세계 과학·의료계가 감염 환자 치료제 및 예방 백신 개발을 위하여 AI를 활용한 바이러스 분석 진행

- 의료계는 코로나바이러스의 강한 전염성에 대응하는 백신을 만들기 위해 바이러스의 구조적 분석에 집중하고 있으며, 빠른 변이로 인해 분석이 까다로워 AI 활용이 필수적
  - 기계학습(ML)과 심층학습(DL)을 통해 바이러스의 유전체 및 분자구조를 분석하여 바이러스의 돌연변이 과정을 찾아내거나 예측*하고, 치료를 위한 새로운 화합물 발견

  * 바이러스의 게놈(Genome) 서열을 빠르고 정확하게 재생성하고 복제본을 구축, 수천 개의 다양한 조합 및 임상 수행(美 의료업체 인실리코 메디슨, AI 솔루션업체 그래픈 등)

2 中 IT기업 바이두(BAIDU)는 AI를 활용하여 코로나바이러스 구조를 분석하거나 감염 여부를 판별, 바이러스 확산을 추적하는 등 다양한 시도를 진행 중[2]

- Linearfold 알고리즘*을 활용하여 코로나바이러스의 분석시간을 기존보다 120배 더 빠른 55분에서 27초로 단축하는 연구결과 및 알고리즘을 관련 커뮤니티에 공개

  * 2019년 오레곤 주립대, 로체스터 대학교와 공동으로 발표한 바이러스 구조 예측 AI 알고리즘으로 기존의 RNA 폴딩 알고리즘 보다 빠름

- 자사의 오픈소스 딥러닝 플랫폼(PaddlePaddle)에서 구동하며 코로나19 폐렴 CT 이미지를 분석하는 AI 모델을 출시하고 92%의 정확도로 1분 이내 감염 여부를 판별*

  * 의료 데이터 분석회사 LinkingMed(베이징 소재)와 협력하여 개발하였으며, 후난성 제휴 병원의 데이터를 기반으로 테스트한 결과

- 지능형 로보콜 플랫폼으로 사람들의 최근 활동 기록을 수집* 및 AI 기반 매핑 시스템**으로 온라인 활동(검색쿼리, 소셜미디어 등)을 분석하여 바이러스의 확산 경로를 추적

  * 3백만 건 이상의 자동 전화 통화를 통한 데이터 수집

  ** 바이두 맵 및 마이그레이션 빅 데이터 플랫폼을 사용하여 바이러스 이동 흐름을 식별하고 위험이 높은 지역을 도출

3 AI기술을 활용한 지능형 보건 체계 구축을 위한 글로벌 협력 강화 필요

- AI로 돌연변이 발생률이 높은 까다로운 바이러스의 구조를 빠르고 정확하게 분석하는 것은 백신 설계 시간을 크게 단축시키고, 더 많은 사람의 생명을 구할 수 있음
  - 코로나19 사태의 조속한 해결과 피해 최소화를 위해 각국의 정부 및 기업들이 기술 및 인적 네트워크를 구축하고 정보를 공유하는 등 지속적 협업 체계 강화 필요

1 MIT Technology Review(2020.3.11.), "How Baidu is bringing AI to the fight against coronavirus"
2 Forbes(2020.3.19.), "How Artificial Intelligence Can Help Fight Coronavirus"

어그테크 기업, AI+클라우드 날개 달고 고공 성장 전망[3]

1 클라우드·AI 기술 확산에 따라 식량 산업에서 어그테크(AgTech)* 기업이 증가

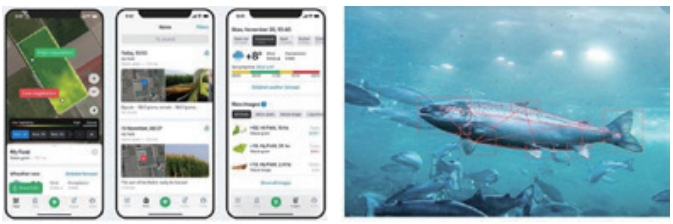
* 농업(Agriculture)과 기술(Technology)의 합성어로 농업에 최신 IT, BT 기술을 접목하는 것을 의미

- 2050년 세계 식량 수요는 현 수준의 2배까지 증가하여 농수산업의 생산성 개선 요구 증대
- 아마존 웹서비스(AWS), 마이크로소프트 애저(MS Azure) 등 클라우드를 활용한 50개 이상의 AI 스타트업 어그테크 분야에 진출(2020.2월 기준)

2 클라우드에서 학습시킨 AI를 활용해 농수산업 종사자의 생산성 개선 서비스 제공

- OneSoil AWS를 통해 농경지 위성사진, 기후, 토양수분 및 영양도, 식물 성장 상태 등의 데이터를 수집·처리한 후 AI를 학습시켜 170개 국가 13만 농가에 서비스 제공
  - 농민들은 무료 앱으로 병충해 발생, 비료·농약 투입량, 파종 및 수확시기에 대한 정보 취득
- Aquabyte AWS에서 제공하는 AI 개발도구(SageMaker)로 구축한 AI 모델을 수중 센서가 수집한 양식 고기의 상태, 수질·수온 등의 데이터로 학습시켜 양식장에 지능형 서비스를 제공
  - 노르웨이 연어 양식 기업은 수중 연어의 상태를 실시간 확인하고 사료 급여량, 해충 발생 여부, 출하시기 등을 추천받아 양식장 운영에 활용
- WeFarm 2백만 명 이상의 아프리카 지역 영세농이 가입한 AWS 기반 플랫폼으로 농민 간 매일 2만 건 이상의 지식 공유 서비스 제공
  - 플랫폼은 단순 지식 답변(Q&A) 시스템이 아닌 다국어를 해석해 가장 적합한 지식을 사용자의 언어로 추천하는 기계 학습 기반 AI 플랫폼

[그림 1] OneSoil 앱(좌)과 Aquabyte가 제공하는 양식장 수중 영상(우)[4]



3 어그테크 시장은 2019년 9억 3천만 달러에서 2024년 24억 9천만 달러로 급성장할 전망

- 식량 산업 분야의 AI 솔루션의 확대는 산업의 생산성 확대에 기여할 것
  * 마이크로소프트 어그테크 스타트업을 위한 데이터 저장 및 AI 솔루션 시장에 가세
- 종사자의 경험지식과 데이터 과학을 결합한 기술 업체 및 전문 인력 수요 예상

3 Wall Street Journal(2020.3.12.), "Agtech Startups Use Machine Learning to Analyze Farm Data"
4 Wired(2020.2.), "The Secret Ingredient for a Sustainable Food Supply : Machine Learning"

딥페이크를 선용하는 사례 속속 등장[5]

1 AI발달과 함께 딥페이크(Deepfake, AI를 이용한 가짜 디지털 콘텐츠)에 대한 우려 증가

- 딥페이크를 이용한 포르노, 가짜 정치 뉴스 등이 화제가 되면서 사회 불안감을 생성
  - * Deeptrace에 따르면, 딥페이크 영상증가율은 전년 대비 약 100% 수준, 영상의 96%가 성인물이 차지[6]
  - * 딥페이크는 합성 성인물 피해자의 인권 침해를 야기할 뿐 아니라, 가짜 정치 뉴스로 민주주의를 위협[7], 새로운 지능범죄 출현[8] 등 다양한 부작용 초래

2 그러나, 딥페이크의 산업적 가치와 선용(善用) 가능성을 보여주는 다양한 사례 존재

[표 1] 딥페이크 기술의 긍정적 활용 사례

분야	사례	세부내용	제공가치
역사/예술	Dal Lives (2019)	美 달리 박물관은 살바도르 달리의 삶에 대해 직접 들을 수 있도록, 생전 인터뷰와 기록 영상 수백 개를 학습하여 45분간의 새로운 영상을 제작	- 대중의 역사/예술에 관한 관심 고취
	살아있는 모나리자 (2019)	모스크바 삼성 AI 연구소는 다빈치의 모나리자를 딥러닝을 통해 눈, 머리, 입을 움직이는 비디오로 변환	
	JFK 암살당일 예정연설 재현(2018)	스코틀랜드 CereProc는 JFK의 암살당일(1963.11.22.) 달라스에서 연설예정이었던 '읽어버린' 연설을 딥페이크 기술로 재현	
언론/교육	로이터의 자동뉴스 제작(2019)	Reuters는 Synthesia의 딥페이크 기술을 이용, 세계 최초로 새로운 촬영 없이 기존영상을 학습하여 개인화된 뉴스영상 제작	- 개인화 뉴스 서비스 - 획기적 제작비용 절감
	베컴의 9개국어 공익광고(2019)	영국 건강 자선 단체는 딥페이크을 이용, 데이비드 베컴이 출연하는 말라리아 퇴치 공익광고를 9개 언어로 제작된 말라리아 캠페인 제작	
의료	AI학습용 모조의료 데이터 생성 (2018)	Nvidia와 의료 연구진은 10%의 실제영상만으로 학습된 '가짜' 뇌 MRI 영상이 실제영상에서만 훈련된 알고리즘과 동등한 성능을 보일 수 있음을 입증	- 프라이버시 보호 - 데이터수집 비용/기간 단축

3 딥페이크 선용을 위한 기술적·사회적 여건 마련 시작

- 딥페이크 자동탐지기술, 디지털 콘텐츠 유통 인프라 신뢰강화기술 등이 개발 중
  - – 페이스북, 구글, MS 등은 딥페이크 자동탐지기술 개발을 위한 투자를 확대 중이며, 블록체인 기술을 이용한 디지털 콘텐츠 진위성 판별 기술 등이 연구 중
- 딥페이크 허용범위의 기준마련과 디지털 콘텐츠 오용에 관한 관계 법령도 정비 중
  - – 페이스북, 유튜브, 트위터는 자체 딥페이크 규정을 마련하고 자사 서비스 통제 중
  - – 국내는 사람 얼굴-음란물 합성한 딥페이크 음란물 처벌을 강화한 '성폭력범죄의 처벌 등에 관한 특례법'이 통과(2020.3.17.)

5 Forbes(2020.3.9.), Why Deepfakes Are A Net Positive For Humanity
6 Deeptrace(2019), "The state of Deepfakes, Landscape, Threats and Impact"
7 Guardian(2019.6.22.), The rise of the deepfake and the threat to democracy
8 Verge(2020.1.29.), FTC says the tech behind audio deepfakes is getting better

美 에너지부, AI for Science 보고서 발표[9]

1 美 에너지부는 산하 국가연구소[10]를 통해 "과학을 위한 AI(AI for Science)" 보고서를 발표하여 AI의 산업적 활용을 넘어 과학의 영역까지 AI의 활용을 추진)

- 美 에너지부는 2019년 2월 트럼프 대통령의 행정명령인 『American AI Initiative』의 일환으로 에너지부 산하 인공지능기술위원회를 신설하고, 55억 달러의 R&D 예산을 AI 분야에 적극 투자하기 위한 조치로 각계의 의견 수렴을 위한 타운홀미팅 개최
- 과학을 위한 AI(AI for Science) 보고서는 향후 10년 동안 AI, 빅데이터, 고성능컴퓨팅과 관련된 과학 분야의 새로운 기회와 도전을 발굴
- 과학적 문제에 대한 AI 방법론의 개발 및 응용에 더불어 고성능 컴퓨팅 관점에서의 AI 특화 HW 연구개발과 확장가능성(Scalability) 등을 포괄

2 과학을 위한 AI 보고서는 16가지 과학 분야에 대해 최신 기술, 주요 도전과제, 로드맵 , 로드맵 달성을 위한 연구개발 방향, 기대되는 결과물로 구성

- 특히 전통적인 6대 과학 분야에서는 AI를 활용한 신소재 설계·발견·검증, 대규모 과학 데이터에서의 새로운 이론 발견, 스스로 작동하는 실험실 등 다양하고 참신한 과학적 기회가 논의됨

[표 2] 분야별 사례

분야	사례
화학 및 재료과학	배터리의 저장 용량을 늘리거나 태양광 발전 효율 향상을 위한 새로운 소재나 화합물의 발견
지구환경과학	산불, 홍수, 가뭄 등 다양한 환경 변화에 대한 위험 요소 감지 및 대응, 회복 수단 개발
생물학 및 생명과학	생물학적, 화학적 시스템의 새로운 발견을 위해 스스로 작동하는 실험실(Self-Driving Laboratories) 개발
고에너지 물리학	AI 기술을 활용한 우주 역사의 재구성, 제타바이트 수준의 데이터를 다루는 AI가 새로운 물리학의 기초를 발견
핵물리학	양성자와 원자핵의 단층 촬영, 우주의 물질/반물질에 대한 이해 향상, 핵합성(Nucleo-Synthesis)의 이해 고도화
플라즈마 물리학	플라즈마 융합 및 연소 상태에 대한 예측 이해도 향상

- 그 밖에 제조와 에너지 분야, 컴퓨터 과학과 AI가 연관되어 있는 8대 분야(AI 이론, 데이터, 영상, SW, HW, 인프라, 엣지 컴퓨팅, 생태계) 역시 각각의 도전과제와 로드맵을 제시

3 과학을 위한 AI 보고서는 미래 과학적 발견의 중추가 AI에 있음을 시사

- 에너지부 산하의 주요 R&D는 이제 AI의 활용이 선택이 아닌 필수로 전환
- 이 보고서는 미국의 AI 기술을 비롯한 과학 영역에서의 글로벌 리더십을 유지하고 선도하기 위한 일환으로 발간되어, 앞으로의 과학기술 R&D 트렌드가 AI 중심으로 발전될 가능성이 높을 것으로 전망됨

9 Argonne National Laboratory, AI for Science Report, (2020.2.28.) https://www.anl.gov/ai-for-science-report
10 Argonne National Laboratory, Oak Ridge National Laboratories, Lawrence Berkeley National Laboratories

IDC, 코로나19로 인한 경기침체 속 AI 시장 및 고용 전망 상향 조정[11, 12]

1 코로나19 감염병이 전 세계로 급격히 확산됨에 따라 경기침체 및 IT 부문의 투자 위축 전망

- 세계 주요 증시가 올 들어 30% 가까이 하락하고 실업이 급증함에 따라 국제통화기금(IMF)은 현재 상황을 2008년 금융위기에 버금가는 경제 위기로 인식
- IT 분야 전문조사기관 IDC는 경기침체의 여파로 올해 전 세계 IT 지출액 전망을 전년 대비 5.0% 성장에서 4.3% 성장으로 하향 조정

2 반면, 감염병 대응에 AI가 적극 활용되면서 AI 시장규모 및 고용은 오히려 증가할 전망

- 시장 감염병 예측·진단·치료제 개발 등 의료분야, 원격근무, 온라인교육, 가상비서 등 비대면 서비스분야를 중심으로 AI 시장 규모는 500억 달러로 전년 대비 30% 성장
  - 헬스케어 기업 자이언트(Gyant.com)가 개발한 코로나19 감염 여부 판단 및 환자 관리를 위한 AI 기반 가상비서 앱의 월 매출은 전년 대비 3배 이상 증가
  - 헬스케어 기업 케어에이아이(Care.ai)는 이미지 인식 기반 감염병 진단 장비의 올해 매출 목표를 6개월 만에 달성할 전망
  - 에듀테크 기업 에드맨텀(Edmentum)의 AI 기반 맞춤형 온라인 학습 도구는 코로나19로 폐쇄하는 초중등학교를 중심으로 도입 문의 쇄도
- 고용 기술 분야 구인 공고가 감소한 반면, 데이터 엔지니어, 데이터 과학자, 기계 학습 개발자 등 AI 일자리 수는 전년 대비 16% 증가한 97만 명에 도달할 전망
  - 미국 컴퓨터기술산업협회(CompTIA)는 지난 2월 미국의 정보기술직에 대한 구인 공고가 1월 대비 약 7만 건 감소했다고 발표
  - 하지만 CompTIA의 부사장 Tim Herbert는 "코로나19에 대응하기 위해 경제 시스템 전반에 정보의 수집, 분석, 처리의 효율화 요구가 증대되어 AI 인력 수요는 증가할 것"으로 전망
  - 실제로 IT 직업훈련기관인 리버쳐(Revature LLC)는 AI 및 기계 학습 훈련생들에 대한 고객 기업의 수요가 급증하고 있다고 발표

3 코로나19 위기를 신속하게 극복하고 경제침체 등 부작용을 최소화하는 동시에 글로벌 기술 주도권 경쟁에 앞서기 위한 AI 분야의 과감한 투자가 진행될 전망

- 미국 백악관은 29,000여 개의 코로나19 관련 연구를 기계가 읽을 수 있는 데이터로 전환하고 슈퍼컴퓨터를 제공하여 AI를 활용한 코로나19 백신 개발을 지원한다고 발표[13]
- 유럽 평의회(Council of Europe)는 미국, 중국 등이 코로나19 대응에 AI를 적극 활용하는 상황에 주목하고 유럽 국가들도 AI를 활용한 위기 돌파에 협력할 것을 촉구[14]

11 Wall Street Journal(2020.3.30.), "AI Spending Is Expected to Weather Coronavirus Slowdown"
12 Wall Street Journal(2020.3.30.), "AI Hiring Expected to Show Resilience Amid Coronavirus Slowdown"
13 White House(2020.3.16.), "Call to Action to the Tech Community on New Machine Readable COVID-19 Dataset"
14 Council of Europe, "AI and Control of COVID-19 Coronavirus"

코로나19 치료를 위한 AI 성과 속 한계 지적[15]

1 AI를 활용한 코로나19의 치료법 개발에 일부 성과 도출

- 코로나19에 대한 AI 분석결과 근육통, 간효소(ALT), 헤모글로빈 수치가 임상에 영향이 있음을 발견*하고 환자를 사망케 하는 합병증(ARDS**) 발현을 70~80% 정확도로 예측[16]
  * 미(뉴욕주립대)·중(온주병원) 공동연구로 53명 환자를 대상으로 분석한 결과
  ** ARDS : 폐에 물이 차는 급성호흡기질환증후군, 치사율 50%
  - 이는, 관련된 치료제 개발뿐 아니라 기존 독감증세와 유사하거나 증세가 뚜렷하지 않은 환자의 상태를 구별하는 것에도 활용
- 제약회사의 신약 개발을 위해 약 1만 가지 이상의 화학 제제 조사 및 실험에 소요되는 시간(5~10년)과 비용(1조 원 이상)을 획기적으로 단축(2개월)하고 치료제 후보 발견[17]*
  * 제약회사 Theragen Etex와 Synteka Bio는 AI를 활용하여 美 식품의약국(FDA)에서 승인한 약물(1,880개) 중 각각 5, 30가지 코로나19 치료를 위한 제형을 발견
- 韓 AI 기업 Lunit은 명지대병원과 협력하여 코로나 바이러스 테스트 속도를 높이는 새로운 진단 AI 프로그램(Lunit Insight CXR*)을 개발, 97~99% 정확도로 감염을 검출
  * 전 세계 80개국 3백만 장 이상의 폐CT 이미지를 학습했으며 수초 이내에 결과 도출(기존 진단키트 검출은 약 6시간 소요), 현재 국내 3천 명 이상 감염 여부 분석
- 美 의료 AI 벤처 Jvion은 공공장소에서 사람들의 온도를 자동 측정하거나, 흉부 X-ray 결과를 머신러닝을 통해 구분하여 입원이 필요한 환자들을 가려내는 시스템 제안[18]*
  * 2백만 명의 기존 환자 데이터 학습을 기반으로 고위험 환자를 구분하고, 개별 병원 및 응급실 활용을 예측하는 등의 내용을 'American Journal of Managed Care'에서 발표(2020.1.)

2 한편, AI를 활용한 코로나19 대응 결과의 불완전성에 대한 문제들이 제기

- UN의 AI 이니셔티브인 Global Pulse, 세계보건기구(WHO), 밀라퀘벡 AI연구소 등은 AI를 통한 코로나19 관련 결과 값에 대하여 충분한 의료·과학적 검증과 검토가 부족함을 경고
  - 코로나19에 대한 AI 활용은 진단, 치료법 발견, 사회적 확산의 세 가지 범주로 구분되며 각 분야별로 과학적 엄격함과 품질 관리 메커니즘을 통한 평가 필요성 제기
- 美 방사선학회(ACR) 및 질병통제센터(CDC) 등은 코로나19 진단을 위한 AI 학습에 고(高)품질의 이미지가 활용되어 초기 징후 탐지 및 유사 증상 구별에 한계를 지적
  - 고품질의 감염 이미지만을 학습한 AI는 바이러스의 감염 여부가 불분명한 초기 징후 탐지에 효과적이지 못할 수 있으며 유사증상을 구별하는 데에 오류를 범할 위험성을 내포
- AI 학습에는 데이터의 절대적인 양과 정확함이 필요하고 현재의 코로나19를 대응하기 위한 AI 시스템의 수준은 당장 현장에 투입하기 어려워 과도한 기대는 경계할 필요[19]

15 OnZero(2020.3.27.), UN Warns Most A.I. Isn't Ready to Fight the Coronavirus
16 NYpost(2020.4.1.), AI tool finds 3 coronavirus signs that often lead to a severe case
17 TheKoreaBizwire(2020.4.2.), Drugmakers Introduce AI to Speed Up Development of Coronavirus Cure
18 TheNextWeb(2020.4.3.), How AI can determine which coronavirus patients require hospitalization
19 TechCrunch(2020.3.27.), AI and big data won't work miracles in the fight against coronavirus

백악관, 코로나19 학술 정보 분석 AI 경진 대회 개최[20]

1 백악관 과학기술정책실(OSTP)은 코로나19 대응의 일환으로 민간 테크기업과 협업해 의료 학술정보를 분석해 해법을 찾는 AI 경진 대회 추진

- 백악관 과기정책실은 민간 테크 기업들과 협업*을 통해 코로나19 관련 학술 데이터베이스(DB)를 구축하고, 오픈데이터 플랫폼을 활용한 AI 경진대회 개최(3.16.)
  * 기계학습용 학술정보DB 구축(Allen Institute for AI, AI2), 데이터 검색 및 분석도구 제공(Microsoft), 기계학습 AI 경진대회 플랫폼(구글 Kaggle), 데이터 저장소 제공(아마존)
- 코로나19 바이러스 관련된 연구 결과를 빠르게 분석하는 AI도구 개발을 촉진하고 의료 커뮤니티에 적시에 필요한 관련 의학 지식 제공이 목적
  - 코로나19 바이러스 발현 원인, 전이, 확산 과정 파악 및 처치 방법, 치료제 개발 촉진

2 구글 캐글에서 "COVID-19 Open Research DataSet(CORD-19)" 경진대회 진행[21]

- **데이터셋** MERS-Cov, SARS-COV-2 등 코로나19 관련 33,000편의 논문 전문(Full-Text)을 포함한 45,000편 이상의 학술 논문이 등록
  * 일부 상업적 출판물들에 대해서는 정부 기관과 테크기업들이 논문 출판업체들의 활용에 대한 법적 승인을 받은 상태
- **요구과업** 의료 커뮤니티에서 주로 논의되는 코로나바이러스 관련 중요한 질문들에 신속하게 해답을 찾을 수 있는 텍스트·데이터 마이닝 툴의 개발과 해법 제시
  - 전미과학·공학·의학한림원(NASEM)과 세계보건기구(WHO)의 코로나19 대응관련 주요 10대 질문*을 과업으로 제시
  * 바이러스의 전염, 배양, 환경안정성, 코로나19의 위협인자, 백신과 치료법, 진단 및 감시, 지리적 영향, 정보 공유와 협력 체계, 윤리 및 사회과학적 고려사항 등
- **상금** 상위 3명의 입상자에게 각각 과업당 $1,000을 캐글에서 제공(4월 3일 응모 마감)

3 최신 과학 지식을 현장에 신속히 적용할 수 있도록 돕는 AI의 역할 조명

- 현장 대응에 여력이 없는 의료진들이 빠르게 업데이트 되는 관련 지식을 적시에 인지하고 활용할 수 있는 정보 체계의 필요성 증가
  - 코로나19 유전자 분석, 변이 예측에 AI가 활용되고 있었으나 이번 경진대회를 통해 방대한 학술 정보를 분석해 현장 전문가에 적시 제공할 수 있는 AI의 역할도 주목
- AI를 활용해 지식 커뮤니티 사이의 최신 정보 공유를 촉진시키고, 집단지성(Crowd)을 통해 다양한 사회적 문제에 대한 창의적 해결 방안 발굴 시도
  - 해결하고자 하는 사회적 문제를 제안하고, 해법 발굴을 위해 민간의 분산된 AI자원(데이터, SW, 플랫폼, 전문가 집단 등)을 통합·매개하는 정부의 촉진자 역할 필요

20 Whitehouse.gov(2020.3.16.), Call to Action to the Tech Community on New Machine Readable COVID-19 Dataset
21 https://www.kaggle.com/allen-institute-for-ai/CORD-19-research-challenge(2020.3.31. 접속)

구글, AI 칩 설계에 AI 활용 시도[22]

1 구글 AI의 수장을 역임한 제프리 딘(Jeffrey Dean)은 새로운 AI 칩 설계를 위해 AI를 활용한 연구 결과를 발표

- 현재 AI의 트렌드는 더 큰 규모의 데이터와 인공신경망을 활용하는 접근이 가속화되고 있으나, 이를 학습시킬 수 있는 컴퓨팅 파워의 발전 속도는 이에 크게 미치지 못해* AI에 최적화된 칩 설계의 필요성이 증가

  * 연산처리장치의 발전 속도는 다양한 개발 상의 난제로 인해 2배의 성능을 달성하기 위해 20년이 필요한 반면, 새로운 AI 모델이 요구하는 계산 양은 3.43개월 마다 2배의 성능을 요구

- 제프 딘은 2020년 2월 개최된 국제 반도체 학회(International Solid-State Circuits Conference, ISSCC)의 기조연설에서 심층학습으로 대변되는 최신 AI기술을 활용해 차세대 컴퓨터 구조와 AI 칩 설계 연구 결과를 소개

2 구글은 AI 칩 설계에 강화 학습을 적용하여 새로운 AI 칩 설계도를 도출한 결과 전문 인력이 설계한 칩보다 20% 이상 효율이 높고, 설계 시간을 단축

- 칩 설계는 다양한 부품과 회로를 배치(Placement)하고 배치된 요소들을 배선(Routing)하는 것으로 볼 수 있는데, 배치와 배선의 형태에 따라 칩의 효율이 변화
  - 구글 연구진은 배치와 배선을 일종의 게임으로 바라보고 강화 학습을 적용해 최적의 배치와 배선을 통한 칩 설계를 시도
- 칩 설계의 배치와 배선에 강화 학습을 적용한 결과 전문 인력 대비 효율이 증대[23]
  - 실제 저전력 기계학습 칩 설계에 강화 학습을 적용한 결과 전문 인력의 설계도 대비 복잡도는 20.2% 낮아지고, 총 배선 길이는 26.7%가 절감되어 전반적인 효율이 증가
  - 구글의 AI 전용 HW인 TPU(Tensor Processing Unit)의 연산 코어와 메모리에 대한 배치와 배선 실험 결과 전문 인력은 6~8명이 일주일간 개발했으나, AI를 적용한 결과 24시간 만에 전문 인력이 개발한 것과 유사한 성능을 갖는 설계도를 도출

3 AI 칩 설계에 AI 활용하는 전략은 빠르게 변화하는 AI 시장의 수요에 대응할 수 있는 수단이며, 구글의 연구결과는 비단 AI 칩 설계뿐만 아니라 컴퓨터 시스템 전반에 적용될 가능성도 있음을 시사

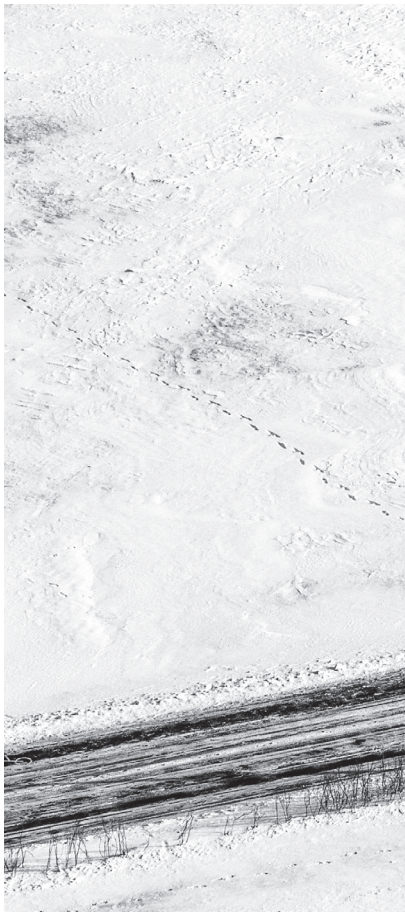
- AI 분야에서 발간된 논문의 수는 지난 2년간 2배가 증가하여 급격한 발전을 겪고 있으나, 이 수요에 대응할 수 있는 AI 전용 칩의 개발은 통상 1.5년에서 2년이 소요되어 이를 단축할 수 있는 수단으로 구글은 AI를 활용한 칩 설계를 시도
- 구글의 연구 결과는 배치와 배선과 같이 휴리스틱 기반의 의사결정이 존재하는 컴파일러나 운영체제에도 적용할 수 있는 가능성이 높아, AI 시스템의 효율 향상을 위해 다시 AI를 활용할 수 있음을 시사

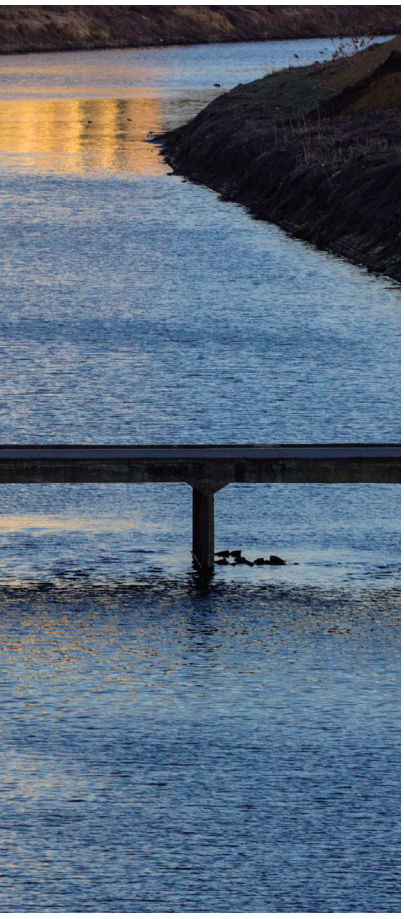
22 Jeffrey Dean(2020.2.20.), The Deep Learning Revolution and Its Implications for Computer Architecture and Chip Design, ISSCC(International Solid-State Circuits Conference)
23 The Next Platform(2020.2.20.), Google teaches AI to play the game of chip design

안양천

이호준 Lee, Ho-Joon I 언론학박사 I ighwns@hanmail.net

우정사업본부에서 근무하고 있으며, 대한항공 여행사진 공모전에서 2회 수상하고, 세 차례의 개인전과 단체전 3회를 개최했다.





사진산책

중간 中間

디지털과 아날로그의 중간.
일상과 쉼의 중간 같은 사진.
특정 주제나 내용에 치우치지 않는 누구나 공감할 수 있는 사진

128년 전통의 '파리-브레스트-파리(PBP)'는 1,230km를 90시간 안에 완주해야 하는, 세상에서 가장 오래된 아마추어 자전거 대회다. 4년마다 열리는 이 대회에 참가하기 위해, 전 세계의 수많은 랜도너들이 파리로 모여든다. 상당한 비용과 시간을 들여 이 고단한 레이싱에 참가하려는 이유는 무엇일까? 이 대회에 참가했던 한 선수는 "내 삶이 넘어지지 않도록 균형을 잡아주는 것이 '도전과 성취'인데, 그것을 두 바퀴를 통해 확인하기 위해서"라고 말한다. 자전거는 '온전히 내 몸이 만들어내는 에너지만으로 달려야' 하는 지극히 단순한 작동 원리를 따른다. '균형을 잡고 넘어지지 않기 위해서는 끊임없이 움직여야' 하는 것이다. 페달을 밟을 때마다 몸과 마음이 앞으로 나아가듯, 끊임없이 움직이고 노력하는 사람만이 자기 인생을 제대로 살아갈 수 있다는 것을 자전거는 우리에게 알려주는 듯하다.

여의도

SPRi FOCUS

코로나19와 미래사회로의 대전환

· 재택/원격근무와 미래의 일, 공간 이명호 디자이너((재)여시재 솔루션)

· 포스트 코로나, 언택트 디지털 전환 기회 방은주 선임기자, 부장(지디넷코라아)

· 언택트를 콘택트처럼 만드는 5G와 실감기술의 기회 양병석 대표(코믹스브이)

* 이 원고는 필진의 개인 견해이며,
소프트웨어정책연구소의 공식의견과 다를 수 있습니다.

재택/원격근무와 미래의 일, 공간

Working from Home/Telecommuting and Future Work, Space

코로나19에 대한 효과적인 대책은 결국 '사회적 거리두기'였다고 할 수 있다. 백신이 없고 치료제가 없는 상태에서 전염병에 대한 대책은 결국 가장 초보적인 방식인 우리 행동을 바꾸는 것이었다. 사회적 거리두기는 여러 형태로 일어났다. 개학 연기, 예배 등 종교행사 중단, 대중 이용시설 폐쇄, 온라인 쇼핑, 재택근무 등. 이 중에서 우리 사회가, 직장인들이 새롭게 경험한 것은 재택근무일 것이다. 재택근무는 정부의 강제적인 조치가 아니었지만, 코로나19 확진자가 나오면 건물이 봉쇄되는 상황에서 기업체들이 단 하루라도 회사가 봉쇄되는 상황을 막고 어떻게든지 업무의 연속성을 유지하는 방안으로 채택한 방법이 재택근무였다.

아직 코로나19가 가져 온 재택근무 현황에 대한 신뢰할 만한 통계가 나오지 않았지만, 한 인맥관리 업체가 이용자들을 대상으로 한 조사에 의하면 응답자의 61.3%는 '코로나19로 인해 근무형태가 변했다'고 답했고, 이 중 22.2%는 유연 단축 근무제를, 19%는 전원 재택근무를, 17.2%는 임산부 및 유증상자 등 일부 인원만 재택근무를 시행하고 있

이명호
Lee, Myungho
(재)여시재 솔루션 디자이너
Yeosijae, Solution Designer
lee.myungho@gmail.com

다고 답했다. 대략적으로 기업체의 36% 정도가 한시적이지만 재택근무를 시행했다고 볼 수 있다. 이는 엄청난 경험이라고 할 수 있다. 사실 다른 선진국들과 기업들이 오래전부터 재택근무를 채택하는 경향 속에서도 우리나라는 재택근무가 거의 도입되지 못하고 있는 상황이었다. 미국은 대략 전체 노동자 중에 3%가 재택근무를 하고 있는 것으로 알려져 있으며, EU 전체가 5%이고 가장 높은 비율의 국가는 네덜란드로 13.7%가 재택근무제를 활용하는 것으로 조사되고 있다. 일반적으로 원격근무제(재택근무)란 조직의 근무자들이 적어도 주 1회 이상 집, 위성사무실, 원격근무센터 등 기존의 사무실 중심 근무현장 이외의 장소에서 정보통신장비를 사용하여 일하는 대안 근무를 의미한다.

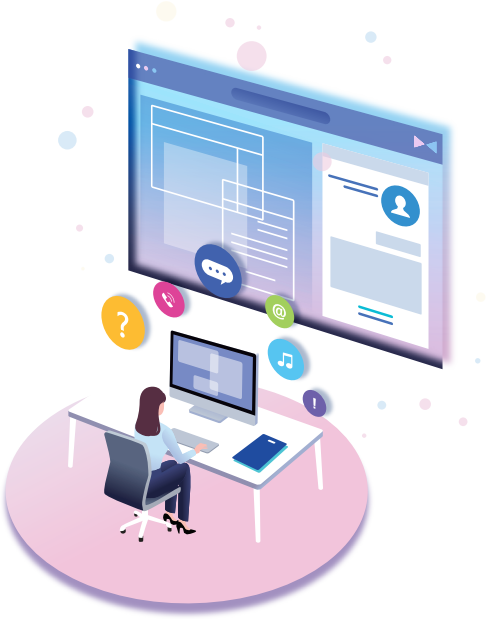
공식적으로 우리나라 정부 통계에서 재택근무라는 근무형태가 집계되지 않고 있으며, 통계청에 따르면 유연근무제(활용여부)는 11% 정도 도입되어 있는 것으로 조사되고 있다. 유연근무제는 재택 및 원격근무제, 근로시간단축근무제, 시차출퇴근제, 선택적 근무시간제, 탄력적 근무제, 기타유형(재량근무 등)을 모두 포함하고 있으며, 전체 유연근무제 활용 중에서 재택 및 원격근무제 비중은 5.8%(2017년 8월 기준)에 불과하다(통계청 KOSIS). 공무원을 대상으로 한 조사에서도 유연근무제 활용 인원이 16.1%에 불과하고, 이 중에서 재택근무유형 비중은 1%도 안되는 0.6%로 조사되었다. 이러한 통계를 바탕으로 어림잡아 우리나라에서 공공과 민간을 포함하여 재택근무 실시율은 0.01%~0.64% 정도의 극히 미미한 수준이라고 할 수 있다. 그런데 이번 코로나19를 계기로 한시적이나마 재택근무를 도입하거나 체험한 경우가 36%에 달했다는 것은 문화적 충격이었을 것이다. 재택근무를 체험한 직원들은 재택근무가 정착되기를 바라는 의견이 많았다.

앞으로 재택근무가 돌이킬 수 없는 근무형태로 자리잡을 것인가에 대해서는 전망이 갈리고 있다. 재택근무는 재난적 상황에서의 한시적인 조치에 불과하고 여러 문제점이 드러났기 때문에 재택근무는 해프닝으로 끝날 것이라는 전망도 있지만, 재택근무는 미래의 근무형태이고 디지털 트랜스포메이션(Digital Transformation)이라는 흐름 속에서 정착될 것이라는 전망도 나오고 있다. 특히 후자의 입장은 재택근무를 일찍 시행해온 선진국에서 많이 주장되고 있다. 필자는 미래의 근무형태, 디지털 트랜스포메이션이라는 측면에서 재택근무의 가능성과 한계, 재택근무가 정착되기 위해서는 무엇이 필요한지를 살펴보고자 한다.

재택근무, 익숙한 개념 그러나 실행은 더디게 진행

재택근무(Telework) 또는 원격근무(Remote Work)라는 개념이 최초로 등장한 것은 1975년이다. 1970년대에 미국 LA로 출근하던 로켓 과학자 잭 닐스(Jack Nilles)는 러시아워로 길에서 시간과 비용을 낭비하는 출퇴근을 하지 않아도 되는 방법이 없을까를 생각했다. 장거리 차량 통근은 교통 정체와 환경 오염을 일으키고, 낭비적인 주거지의 교외 확산을 야기하며, 비효율적이다(2016년 기준으로 한국 직장인의 통근시간은 평균 40분(편도)이다. 이는 OECD 평균 통근시간 23분 보다 약 두 배에 달한다.). 국립과학재단(NSF)의 지원을 받아 닐스는 LA의 한 보험회사가 재택근무("Telecommuting"이라는 용어를 사용)를 시행하는 것이 가능할지 알아보는 연구를 진행했다. 닐스는 재택근무 시행이 가능하다는 결론을 내렸고, 그 회사는 재택근무를 시도하였다. 그러나 시행되자마자 폐기됐다. 경영자들은 재택근무하는 직원들을 전과 같은 방식으로 통제할 수 없었기 때문이었다. 닐스가 우려했던 데로 상사는 직원을 통제할 수 없을 것이고 직원들은 사무실 생활에서 오는 사회적 분위기를 잃을 것이라는 문제를 극복하지 못했다(니킬 서발, 2015).

재택근무의 개념을 더욱 발전시킨 선구자는 앨빈 토플러였다. 토플러는 〈제3의 물결〉(1982년)에서 "지식 근로자들이 전자 오두막(Electronic Cottage, 자기 집에서 통신



장비를 마련해 일하는 생활 양식)에서 일하게 된다. 퍼스널 컴퓨터와 영상장치, 통신장비 등을 이용해 새 유형의 네트워크를 만들 수 있다."고 전망하였다. 닐스가 출퇴근의 문제에 주로 주목했다면 토플러는 '일하는 방식'의 변화, 네트워크로 연결된 컴퓨터로 일하는 시대가 되면 사무실이 아닌 재택근무가 가능해질 것으로 보았다. 1990년대에 접어들며 ICT 기술이 발달하면서 재택근무와 원격근무를 검토하고 도입하는 기업들이 나타나게 된다.

기업이 재택근무를 도입하는 동기는 조직 구성원에게 일과 삶의 균형 유지(Work-Life Balance), 일-가정 충실화(Work-Family Enrichment)에 도움을 주어 우수 인력 유치와 유지(이직 의지 감소), 직무 성과 및 생산성(Productivity) 향상, 사기 진작, 직무 만족도 증가, 자율성 증가, 조직 몰입(Organizational Commitment), 무단결근(Absenteeism) 감소, 스트레스 감소 등의 긍정적 효과를 얻고자 하는 측면이 강하였다. 즉, 재택근무(원격근무)는 ICT를 기반으로 한 인적자원관리(ICT-Based Human Resource Management)를 위한 유연근무제(Flexible Work Arrangement)의 한 형태로 도입되기 시작하였다고 할 수 있다. 이러한 인적자원 관리 방식으로서의 재택근무의 효과는 수 많은 연구들을 통하여 입증되고 있으나, 여전히 논란이 있는 부분은 개인이 느끼는 재택근무의 효능이 조직과 기업의 성과로 이어지는 관계가 불명확하고, 혁신과 창의성을 증진시키지는 못한다는 점이다. 이는 기업이 재택근무에 맞는 직원들의 직무관리 방식을 확신하지 못한 상태에서 재택근무를 꺼리는 요인으로 작용하고 있다. 필자는 인적자원 관리 관점의 재택근무에서 미래의 일과 공간이라는 관점에서 재택근무, 원격근무를 검토해야 제대로 원격근무의 장점을 살릴 수 있다고 본다.

디지털 시대의 일, 일하는 방식은 온라인으로 진화 중

일, 일하는 방식의 변화를 보기 위해서는 일의 결과물(생산물), 일하는 도구, 일하는 조직, 일하는 사람, 일하는 공간 등에서 어떤 변화가 일어나고 있는지를 봐야 한다. 산업시대와 디지털 시대로 구분해서 검토해 보기로 하자. 산업시대의 범용기술(생산 및 경제 활동의 기반이 되는 기술)은 엔진이었고, 디지털 시대는 컴퓨터(인터넷)으로 변하였다. 산업시대에는 동일한 기계가 여러 대 모여있는 공장에서 엔진으로 동력을 얻는 기계를 작동하여 유형의 제품을 생산하는 육체 노동자들이 대규모의 수직적인 기업 조직에 속해서 동시에 일하는 경제활동 방식이 주를 이루었다. 이 시대의 재택근무란 공장 생산과 관련된 서류 작업을 집에 가서 하는 것이었다고 할 수 있다. 재택근무는 공장제 근무의 종속된 형태에서 벗어날 수 없었다.

디지털 시대에는 일하는 방식에 근본적인 변화가 일어났다. 범용기술은 컴퓨터(인터넷, 모바일)로 변하였다. 기계 또는 도구에 엔진과 동시에 컴퓨터(IoT : 사물인터넷)가 들어가는 것이 일반화되고 있다. 일의 형태는 컴퓨터라는 도구를 이용하여 무형의 콘텐츠(알고리즘)를 만드는 정신 노동으로 바뀌었고, 직원들은 소규모의 수평적 조직(팀이라고 할 수 있다)으로 구성되어 사무실이라는 공간에서 서로 다

른 연관된 업무를 비동시적으로 일하는 방식으로 바뀌었다. 노트북 하나면 어떤 정보든지 입수 가능하고, 어떤 업무 프로세스도 접근할 수 있고, 통합적으로 업무를 처리할 수 있게 되었다. 이는 디지털 시대의 일을 공간적 귀속성, 시간적 귀속성을 완화시키고 유연화시키고 있다. 한마디로 일의 디지털 전환이라고 할 수 있다. 일의 디지털 전환은 일하는 도구의 디지털화 단계에서 일 자체의 디지털화로 넘어가고 있다. 일의 결과물, 일 자체가 디지털이 되고 있다. 초기 단계의 업무 전산화는 일(사무)의 일부분을 컴퓨터로 처리하는 수준이었다. 그러나 지금은 업무 자체가 디지털화되는 단계로 나아가고 있다. 회사 조직이 클라우드 플랫폼 위에서 움직이고 모든 업무가 디지털 도구에 의하여 처리되고, 협업과 업무의 연계도 디지털에 의해 이뤄지고, 결과물도 디지털로 나오는 업무 형태가 등장한 것이다.

이러한 디지털 기업의 대표적인 재택근무 사례는 깃랩(Gitlab)이다. 깃랩은 개발자들이 소스코드를 쉽게 공유할 수 있는 소프트웨어를 운영하는 회사로 1,200명 직원 모두가 원격근무를 하고 있다. 전 세계 CMS(콘텐츠 관리시스템)의 60%를 차지하는 '워드프레스'를 서비스하고 있는 오토매틱도 직원의 5%만 사무실에서 일한다. 오토매틱 CEO는 "사람마다 집중 잘 되는 시간, 휴식 취하는 시간이 다르다. 언제, 어디서 일하느냐 보다 똑똑하게 일하는 것이 중요한 시대다."라고 재택근무 도입의 이유를 밝히고 있다. 미국의 디자인 소프트웨어 회사 인비전(InVision)은 2011년 설립할 때부터 근무형태를 '100% 원격근무'로 정했고 20개국에 흩어져 있는 1,000명의 직원들이 원격근무 중이다(매일경제, 2020.3.16.). 우리나라도 100% 원격근무를 지향하는 기업들이 등장하고 있다. 온라인 스터디 사업을 하는 스터디파이도 2018년 창업을 하면서 전 직원 원격근무를 실시하고 있으며, 많은 스타트업들이 주도적으로 원격근무 등의 유연근무를 채택하고 있는 것도 주로 디지털 기업이라는 특성을 갖고 있기 때문이다.

결국 디지털 전환이 근무의 유연성을 높이고 공간(사무실)에 대한 종속성을 완화(해방)시키고, 집이나 어느 공간에서도 네트워크로 연결된 노트북만 있으면 업무를 볼 수 있는 환경이 되면서 재택근무, 원격근무가 기본적인 업무 형태로 자리잡아 가고 있다고 봐야 할 것이다. 기업의 입장에서도 이번 코로나19와 같은 천재지변의 상황이 발생하더라도 업무의 단절없이 업무가 지속적으로 처리될 수 있는 인프라의 조성이라는 측면에서도 원격근무는 기업의 중요한 위기관리 역량으로 요구되고 있다. 또한 기후변화와 지구온난화는 기업들에게 재생에너지 사용 등 친환경 정책을 요구하고 있는 것도 고려할 필요가 있다.

[표 1] 산업시대와 디지털시대의 일의 변화

구분	산업시대	디지털시대
범용기술	증기기관(엔진)	컴퓨터(인터넷, 모바일)
생산물	유형의 제품	무형의 콘텐츠(알고리즘)
도구	기계	컴퓨터
직원	육체 노동자	정신 노동자
조직	대규모 수직적 조직	소규모 수평적 조직
공간	공장	사무실
시간	동시적 작업	비동시적 작업

조직문화, 혁신성과 창조성의 문제를 어떻게 극복할 것인가

재택근무, 원격근무에 대한 가장 큰 반대 주장은 기업, 조직 측면에서의 혁신성과 창의성을 저하시킨다는 것이다. 원격근무의 원조라고 할 수 있는 IBM은 1990년부터 시행해오던 원격근무를 2017년 2월에 전격적으로 폐지하였다. 전자상거래 업체 자포스도 원격근무를 허용하다가 2013년부터 금지하였다. 원격근무 반대 주장들은 인간의 대면(Face-To-Face)이 가져오는 효과를 강조하고 있다. "팀이 될 때 더 강력해지고 창의적으로 된다, 더 마주치고 한마디라도 더 어울려야 혁신이 가능하다, 네트워크 시대에는 이메일과 채팅으로 아이디어를 개발할 수 있다는 것은 착각이다." "복도에서, 카페에서, 우연히 마주친 사람과의 즉흥적인 대화에서 아이디어가 나온다. 이는 오로지 사무실에서만 가능하다." 많은 연구도 "모여서 일한 직원들의 창의적인 업무수행능력이 원격근무자들보다 뛰어났다."고 보고하고 있다.

그럼 재택근무의 효과로 입증된 개인의 생산성이라는 것이 조직의 생산성으로 이어지지 않는 것인지, 조직의 생산성으로 이어지지만 혁신성과 창조성은 다르다는 것인지를 구분할 필요가 있다. 일반적인 연구결과는 재택근무가 조직 생산성을 저하시키지 않고 높인다고 보고 있다. 쟁점은 조직 혁신성과 창조성이다. 개인적인 대면이 가져오는 접촉의 밀도와 우연성이 조직(단체)의 혁신성과 창조성에 긍정적인 영향을 미친다고 본다. 그리고 그 사이에는 협력성과 연대감이 작동한다고 본다. 인사관리(HR)을 연구하는 사람들은 공간을 1) 물리적 공간(Physical Distance), 2) 업무적 공간(Operational Distance), 3) 감정적 공간(Affinity Distance)으로 나눠질 수 있고, 업무 효율성에 영향을 미치는 공간은 감정적 공간 〉 업무적 공간 〉 물리적 공간 순으로 영향력이 다르다고 본다. 즉 감정적 공간이 업무 효율성에 더 중요하다. 직원들이 한 공간에 모여 있지 않아도 신뢰, 협력과 연대감을 키울 수 있느냐가 재택근무와 원격근무가 성공할 수 있느냐를 가르는 분기점이라고 할 수 있다. 원격근무를 지원하는 협업 툴도 중요하지만, 결국 조직문화가 관건이다. 우리나라에서도 디지털화가 많이 진행되었으나 원격근무 도입이 늦어진 이유도 수직적 통제방식의 조직문화 때문이라고 할 수 있다.

재택근무, 원격근무에 성공할 수 있는 조직문화는 다음과 같이 정리할 수 있다. 과정에 대한 감독이 아니라 성과와 결과 중심의 업무(평가)문화가 필요하다. 눈 앞의 직원이 열심히 일하는 듯(?)이 보여야 안심하는 업무관리 방식으로는 원격근무에 성공할 수 없다. 자율적인 업무 문화가 필요하다. 업무 위임이 명확하고 책임과 자율권이 주어져야 떨어져서 독립적, 주도적으로 일을 수행할 수 있다. 투명하고 명확한 커뮤니케이션 역량이다. 업무 목표와 지시가 명확하게 글(문서)로 전달되고 이에 대한 논의와 대화가 기록되어 다른 팀원들도 볼 수 있게 해야 한다. 물론 전화 통화가 더 효율적일 수 있으면 그렇게 하더라도 통화 기록은 정리되어 글로 공유돼야 한다. 업무 진행 과정을 투명하게 공개해 모두 열심히 일하고 있음을 인식시키는 신뢰 형성이 중요하다. 마지막으로는 팀의 범위를 넘어선 조직 전체 차원의 커뮤니케이션 확대이다. 앞에서도 언급하였듯이 혁신성과 창의성은 우연성에 기인하는 것이 많다. 따라서 팀을 벗어난 전체 게시판, 메신저, 이메일 등을 활성화할 필요가 있다. 대면보다 접촉의 밀도는 떨어지겠지만, 온라인 소통은 접촉의 범위를 넓혀주는 장점이 있다.

O2O : Online Work와 Offline Work의 융합이 최적의 솔루션

사무실 출근 방식의 근무를 하다가 재택근무로 전환하는 것은 생활방식을 바꾸는 것이기 때문에 쉽지 않다. 많은 재택근무 체험자들이 출퇴근 시간이 절약되어 개인적인 시간이 늘어나고, 방해 받지 않고 집중할 수 있어서 좋은 반면

에, 혼자 일하는 외로움, 의사소통의 어려움, 불명확한 업무 지시, 일과 삶의 균형의 붕괴, 근무시간 이외의 근무지시, 논다고 생각하지 않을까 하는 자기검열에서 오는 스트레스, 가족과 자녀의 업무 방해 등의 부정적 효과를 경험하고 있다. 물론 재택근무가 정착된 기업들은 재택근무의 장점을 잘 살리고 있다. 재택근무라는 것이 업무를 넘어 생활습관을 바꾸는 것이기 때문에 점진적으로 몸에 배는 시간, 적응 기간이 필요하다. 재택근무 도입을 결정한 기업들은 이러한 전환 기간 동안에 발생할 수 있는 문제를 해결해 나가야 할 것이다.

가장 좋은 재택근무, 원격근무 방식은 무엇일까? 한 연구에 의하면 원격근무를 잘하는 회사의 특징은 회의실 같은 오프라인 공간에서 긴밀하게 다양한 의견과 아이디어를 모으고, 해결책을 정한 후 역할 분담을 해서 각자 흩어져 자신의 공간(집, 카페, 공유 오피스 등)에서 맡은 역할을 하는 방식이었다. 온라인과 오프라인, 협력, 합의와 개인 작업, 사무실과 원격 공간, 집중과 분산 등을 자유자재로 전환하는 기업이었다고 할 수 있다. 실리콘밸리의 HR 스타트업 Humu도 자신의 고객사들을 분석한 결과 사무실 근무와 재택근무를 번갈아 하게 했을 때 직원들이 가장 만족하고 생산성도 높았다고 보고하고 있다. 일주일에 3-4일은 사무실, 1-2일은 재택 등 자신이 원하는 곳에서 일하게 하는 것이 가장 이상적인 것으로 나타났다. 업무 공간과 업무 시간, 업무 방식의 자율권을 줄 때 직원들은 가장 동기부여를 받게 된다. 혁신성과 창의성의 조건이 다양한 우연성이라고 했을 때, 온라인 워크와 오프라인 워크의 결합은 다양성을 증진시켜 조직이 더욱 혁신성과 창의성을 발휘하게 해 줄 것이다.

코로나19가 누구도 못했던 디지털전환을 가속시키고 있다고 한다. 재택근무, 원격근무란 디지털 시대의 일하는 방식, 일에 대한 전환을 의미한다. 전환의 과정은 쉽지 않겠지만, 분명히 우리나라 기업문화의 혁신, 생산성과 창의성이라는 성과를 가져다 줄 것이다.

참고 자료

여시재(2017), 디지털 기술의 발달에 따른 일과 직주공간의 미래, 여시재 Issue Report 2017.9.

이명호(2018), 미래의 일자리와 도시 공간, 여시재 Issue Report 2018.1.

이명호(2020), 디지털이 미래다(08) : 기업과 노동의 미래, 여시재 인사이트, 2020.2.25.

박시진, 유상엽(2017), 정부조직 원격근무제 논리모형 : 미국 연방정부 사례를 중심으로

국회입법조사처(2016), 공무원 유연근무제의 현황과 개선방안

통계청, 유연근무제 활용여부, (http://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT_1DE7100)

통계청, 유연근무제 활용형태(복수응답), (http://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT_1DE7100S&vw_cd=MT_ZTITLE&list_id=101_B1A&seqNo=&lang_mode=ko&language=kor&obj_var_id=&itm_id=&conn_path=MT_ZTITLE)

TTimes(2020), 재택근무 없던 시절로 돌아가긴 어려울 것

TTimes(2020), 스티브 잡스는 왜 재택근무를 미친 짓이라 했을까?

한국경제, '워라밸' 정착 네덜란드 7명중 한명꼴 재택근무… 한국도 바뀔까(https://www.hankyung.com/international/article/201806236430Y)

매일경제, 1,000명 직원 모두가 원격근무를 하는 유니콘 기업 '인비전'(https://www.mk.co.kr/news/business/view/2020/03/269098/)

IT조선, [슬기로운 재택근무백서] "원격근무 경험이 이력서 빛내는 시대"(http://it.chosun.com/site/data/html_dir/2020/03/06/2020030602792.html)

매일경제 미라클랩, Silicon Valley Original : 코로나19 사태 이후 우리는, 어떻게 일을 잘 할 수 있을까(https://stibee.com/api/v1.0/emails/share/mXuU0IqT-rXlJd5uEo4eQ3Jco03APQ==)

리멤버, 코로나19가 미친 영향 직장인 설문조사, 2020.3.31. (https://www.zdnet.co.kr/view/?no=20200331083426)

Deloitte(2019), The future of work in technology, Deloitte Insights

포스트 코로나, 언택트 디지털 전환 기회

Post-corona, Opportunities for Non-contact Digital Transformation

중국 우한에서 발생한 신종 코로나 바이러스(코로나19)로 세계가 경악에 빠졌다. 4월 13일 오후 8시 현재 코로나19 사망자가 세계적으로 11만 3,000여 명에 달하고 확진자가 18만 명에 육박했다. 처음 중국에서 발생할때만 해도 '블랙 스완(Black Swan, 어쩌다 한번 일어나는 큰 사건)'인줄 알았는데 '뉴 노멀(새로운 사회 현상)'로 확대됐다. 이제 코로나19는 세계 산업을 재편하고 정부에 새로운 역할을 요구한다. 사람 간 소통 방식도 바꾸고 있다.

특히 코로나19가 팬더믹으로 진화하면서 세계적으로 시행되고 있는 '물리적 거리두기'는 '비대면(Untact, 언택트)'이라는 새로운 문화를 형성하면서 글로벌 산업 지형도도 바꿔놓고 있다. 난세에 영웅이 태어나는 법이다. 미증유의 코로나19는 새로운 경제 및 산업 영웅을 낳을 것이다. 그 싹은 이미 심어졌다. 코로나19 위기 속에 원격의료, 원격 근무, 영상회의, 클라우드, 로봇, 전자상거래(e커머스), 스트리밍 분야는 각광을 받았다. 반면 자동차, 항공 같은 전통 제조업과 공유오피스 같은 대면 서비스업은 고전을 면치 못했다.

방은주
Bang, Eunjoo
지디넷코라아 선임기자, 부장
ZdNet Korea senior reporter
ejbang@zdnet.co.kr

코로나19가 세계 산업계에 던진 메시지는 뚜렷하다. 언택트 산업과 시장을 잡으라는 거다. 향후 세계는 점점 더 언택트 문화가 확산될 것이다. 기업, 학교, 정부에서 일하는 방식이 바뀌고 노는 문화가 달라질 것이다. 각 주체의 언택트화는 곧 디지털화를 뜻한다. 그동안 4차 산업혁명이라는 이름으로 진행되온 각 분야의 디지털화는 코로나19라는 '태풍'을 만나 더 급물살을 탈 것이다.

이미 변화는 시작됐다. 그동안의 코로나19 기승은 벌써 세계부호 순위를 바꾸게 했다. 온라인 유통과 온라인 게임, 온라인 교육 같은 온라인 분야 창업주들이 억만장자 대열에 새로 합류했다. 반면 전통 제조업체를 창업한 거부들은 자산이 눈에 띄게 줄었다.

변화의 방향은 명확하다. 언택트고 자동화고 디지털이다. 코로나19 이후 산업별 희비도 이 시장 장악 여부에 따라 달라질 것이다. 사실 미래를 예측한다는 건 '위험'한 일이다. 잘못된 예측은 혹세무민이 되기 때문이다. 하지만 코로나19 이후에 대해 확실히 말할 수 있는 건 있다. 코로나19는 언젠가 진정될 것이고, 이후 세계 산업 질서는 이전과 다를 것이라는 점이다. 몇 가지 키워드로 코로나19가 초래한 산업 현황을 살펴봤다.

코로나19 최대 수혜자는 e커머스

집에 머무는 시간이 늘면서 온라인 유통과 전자상거래(e커머스) 분야는 호황을 맞았다. 근착 영국 파이낸셜타임즈(FT)도 "코로나19 시대 최대 승자는 온라인 유통업체"라고 진단했다. 코로나19 이전에도 오프라인 유통은 온라인 유통에 야금야금 시장을 빼앗겼다. 코로나19는 여기에 쐐기를 박았다. 코로나19 사태가 끝나면 소비자들이 이전처럼 오프라인 매장을 찾을까? 이전과 다를 것이라는 게 중론이다. 미증유의 코로나19 사태에도 e커머스 거인 아마존은 단연 돋보였다. 생필품 공급처를 넘어선 아마존은 의료기기 보급, 사재기 방지 같은 정부와 병원 역할도 일부 했다. 미 정부 요청으로 미국 전역에 손세정제와 마스크를 보급하기도 했다. 늘어나는 온라인 배송 수요를 충족하기 위해 아마존은 미국에서만 10만 명의 직원을 추가로 고용하겠다고 밝혔다.

미국뿐 아니다. 중국 전자상거래 시장의 55%를 차지하고 있는 알리바바 역시 코로나19로 수혜를 입었다. 코로나19 기간 중 매출이 상승세를 보였다. 우리나라도 마찬가지다. 산업부 발표에 따르면 지난 2월 온라인과 오프라인 유통업체 매출 비중은 각각 49 대 51이였다. 작년 2월에는 이렇지 않았다. 온라인 비중이 39%였다. 코로나19로 온라인 매출이 껑충 뛴 것이다. 이 추세라면 정부가 발표하는 3월 매출에는 온라인 매출이 사상 처음으로 오프라인 매출을 넘어설 가능성이 높다. 더구나 온라인에서 판매하는 제품은 계속 늘고 있다. 최근에는 자동차와 오토바이도 온라인 판매 품목에 가세했다.

세계 15억7600만명이 휴교령.... 세계는 원격 교육 바람

코로나19 확산 방지를 위해 우리나라를 비롯해 미국, 유럽 등이 학교 수업 대신 원격 수업을 시행했다. 유네스코에 따르면 세계 188개국에서 15억 7,600만 명 학생이 휴교령 영향을 받았다. 세계 학생의 91.3%에 달하는 수치다. 유네스코는 "이런 규모 교육 차질은 역사상 유례없는 일"이라고 놀라워했다.

프랑스는 200여 년 전통의 유명한 대학입학시험 '바칼로레아'를 올 6월 치르지 않기로 결정한 바 있다. 코로나19로 세계 교육이 요동치고 있는 것이다. 물론, 코로나19 이전에도 교육 분야에서 여러 혁신 모델이 있었다. 온오프라인으로 교육하는 블렌디드 교육과 '플립 교육'이 그 예다. 여기에 물리적 캠퍼스가 없는 온라인 대학 '미네르마스쿨'과 '무크(MOOC, Massive Open Online Course)'라 불리는 대학의 온라인 수업도 새로운 교육 풍속도를 형성했다. 하지만 코로나19 여파로 현재 세계 교육 시장에 불어닥친

원격 교육은 규모 면에서 이전 혁신 모델을 압도한다. 그만큼 파장도 크다. 코로나19가 교육 분야에 던진 메시지는 여러 장소에서 수시로 지식을 전수하는 시스템을 만들라는 것이다. 이 니즈를 잡는 곳이 코로나19 이후 교육 시장의 강자가 될 것이다.

원격진료 부상…"10년 걸릴 변화를 코로나19가 1주일 만에 해결"

코로나19로 극적 변화를 맞은 대표적 분야가 원격진료(원격의료)다. 그동안 원격진료에 꿈쩍도 않던 우리나라도 정부가 지난 2월 말 한시적으로 이를 허용했다. 환자가 의료기관을 직접 방문하지 않아도 전화 상담과 처방을 받을 수 있게 했다.

우리나라뿐만이 아니다. 일본도 4월초 도쿄, 오사카 등 7개 지역에 한 달간 비상 사태를 발효하며 초진 원격진료를 허용했다. 이 조치로 온라인 상담과 신체검사 없는 처방전 발급이 일본에서 가능해졌다. 일본 언론은 "진료 차트의 65%가 종이 문서로 발급되고 병원비 결제도 현금이 대세인 보수적인 일본에서 획기적 변화"라고 평했다.

미국 정부도 지난 3월 말 약 600만 명에 해당하는 메디케어(65세 이상 고령자와 장애인을 위한 공공보험) 가입자에게 원격의료 서비스를 허용했다. 시장조사회사 포레스터리서치는 3월 미국 원격의료 수요가 지난해 같은 달보다 30% 증가했고, 올해 3,600만 건의 원격진료가 이뤄질 것으로 내다봤다.

미국내 800개 병원과 제휴하고 있는 원격진료기업 반얀메디컬 최고경영자(CEO) 토니 부다는 "코로나19 유행으로 원격진료 인식에 전환이 나타나고 있다"면서 "대면 진료가 아니면 정확한 진단을 내리기 어렵다며 내심 꺼리던 의사들도 많이 달라졌다"고 소개했다.

영국은 코로나19 사태 이전 연간 3억 5,000만 회에 이르는 방문 진료 중 원격의료 비중은 불과 1%였다. 그런데 지난 3월에는 수천 곳의 병원이 원격진료를 시행했다. 런던의 한 의사는 "10년 걸릴 변화가 1주일 만에 일어났다."고 말했다.

이밖에 3월 말부터 3주간 전국 봉쇄를 실시한 인도도 전

화, 왓츠앱, 스카이프 등으로 처방전을 받는 일을 허용했다.

특히 이번 코로나19 사태에서 의료 인공지능(AI)이 감염병 확산에 큰 활약을 했고, 이런 움직임은 앞으로 더 거세질 것이다.

재택근무 확산... 영상 회의 솔루션 각광

코로나19는 국내외에서 재택근무 열풍을 낳았다. 땅이 넓은 미국은 기업에 따라 재택근무를 활발히 도입했다. 하지만 지난해 IBM이 재택근무를 철회, 미국에서 재택근무 효과에 대한 논란이 있었고, 이번 코로나19는 이를 한 방에 날려버렸다. 마이크로소프트(MS)를 비롯해 구글, 줌 등이 내놓은 재택근무 솔루션 사용이 급증했다.

MS가 만든 영상회의 솔루션 '팀스(Teams)'의 경우 지난 3월 20일 기준 하루 사용자가 4,400만 명으로 일주일 새 사용자가 1,000만 명 이상이 늘었다. 구글 지스위트의 영상회의 서비스 구글 미트도 사용자가 급증, 3월 사용자가 지난 1월보다 20배 이상 늘었다.

중국계 에릭 위안이 창업한 원격영상회의 전용 솔루션 줌도 3월 한 달간 일일 트래픽이 전월보다 53%나 증가했다. 줌은 보리스 존슨 영국 총리와 앨런 그린스펀 전 미 연방준비제도(연준. Fed) 의장도 애용자였는데 최근 보안 문제가 불거져 사용자 증가세가 주춤거리고 있다.

영상회의 솔루션과 함께 미국 넷플릭스 시청자도 크게 늘었다. 시장조사기업 센서타워에 따르면 이탈리아와 스페인에서 3월 한 달간 넷플릭스 앱 다운로드 건수가 각각 66%, 35% 증가했다. 대부분이 넷플릭스에 가입한 미국에서도 신규 가입자가 9%나 늘었다.

기업 채용 시장에도 언택트 바람

코로나19는 기업 채용 문화도 바꿔 놓았다. 일부 대기업이 면접은 물론 필기시험도 언택트로 진행했다. 대표적인 곳이 SK이노베이션이다. 이 회사는 지난 3월 말 신입사원을 채용하면서 필기시험을 온라인 심층역량검사로 대체했다. SK텔레콤도 통신업계 최초로 '언택트 채용' 방식을 도입해 신입 사원을 뽑았다.

현대해상의 경우 4월 초 이뤄진 6급 신입사원 채용에 온라인 AI면접을 도입했다.

현금 없는 사회 가속화

코로나19로 결제 문화에도 지각변동이 일고 있다. 특히 전통적으로 현금 선호 현상이 강한 유럽에서 캐시리스 바람이 불고 있다. 유럽중앙은행(ECB)에 따르면 2018년 기존 유로존(유로화 사용 19개국)의 평균 현금 결제 비중은 53.8%였다. 미국(26%), 한국(18.8%)보다 두 배 이상 높다. 코로나19로 타격이 큰 스페인과 이탈리아는 현금 결제 비중이 각각 87%와 86%에 달했다. 이런 문화가 바뀌고 있다. 바이러스 감염 우려로 지폐 사용을 꺼리기 때문이다.

현금 결제 감소는 코로나19 이전에도 있었지만 코로나19는 이를 더 촉발하고 있다. EU 집행위원회와 ECB는 "지폐와 동전에 코로나 바이러스가 묻어 있을 수 있으니 카드와 모바일결제를 활용하라."고 권고하기도 했는데, 영국 파이낸셜타임즈(FT)는 "코로나19가 현금 없는 시대를 빠르게 앞당기고 있다."고 전했다. 우리나라도 예외가 아니다. 코로나19가 급속히 퍼지기 직전인 지난 2월 첫째 주 2조 920억이었던 신용카드 온라인 결제 승인액은 3주 후에 2조 7,611억 원으로 32%나 급증했다.

사람을 필요로 하지 않는 시대가 온다

코로나19 사태로 3월 중순 이후 2주 동안 미국에서 발생한 실업자는 1,000만 명이 넘는다. 코로나19가 진정되고 경제활동이 재개되면 이들이 다시 일터로 돌아갈 수 있을

까? 그렇지 않을 것이라는 전망이 우세하다. 실제, 미 브루킹스 연구소는 노동자가 빠져나간 자리를 로봇이 채울 것으로 예상했다. 1970~1980년대만 해도 경기 침체 때 사라진 일자리가 경기 회복으로 다시 살아났다. 하지만 1990년대 들어 이런 움직임이 둔화됐고, 이번 코로나19 사태는 이를 다시한번 확인해 줄 것이다.

글로벌 공급망 재편 불가피... 세계화 종말?

세계가 지난 20년간 호황을 보인데는 1980년대 이후 정보기술을 기반으로 발전한 고도화된 글로벌 공급망(SCM, Supply Chain Management)이 한 몫 단단히 했다. 하지만 코로나19로 글로벌 공급망은 재편이 불가피해졌다. 바이러스 창궐로 공급망에 차질이 빚어졌기 때문이다. 특히 세계 공장 역할을 해 온 중국은 코로나19 진원지인데다 코로나19 대응 과정에서 신뢰하기 어려운 파트너라는 인식을 세계에 심어줬다. 중국을 중심으로 한 공급망에 균열이 불가피하다. 애플은 중국 폭스콘 공장 폐쇄로 1분기 아이폰 출하량이 최대 10%감소할 것으로 예측하기도 했다.

중국에 집중된 글로벌 공급망은 재편이 불가피하고, 주요 기업은 공급 불확실성을 차단하기 위해 역내에서 주요 제품을 생산할 것이다. 한편으로는 해외로 옮긴 제조시설을 본국으로 옮기는 리쇼어링도 활발해질 것이다. 미국 한 언론은 "코로나19로 공급 사슬이 망가지자 비용을 더 지불하고서라도 국내 공급자들을 찾아 나선 기업들이 생겨나고 있다."고 전했다.

글로벌 공급망 붕괴는 지난 20년간 줄기차게 달려온 세계화의 종언과 맞닿아 있다. 팬데믹(세계적 대유행)으로 번진 코로나19는 초연결적인 세계화 시대의 취약점과 부정적 얼굴을 고스란히 드러냈고, 이는 세계화의 종말로 이어질 것이라고 일각에서 주장하고 있다.

페놉티콘 사회 오나 우려도

코로나19 감염 확산을 차단하기 위해 각국이 개인 정보를 활용하면서 '빅브라더'와 '페놉티콘'(감시자가 보이지

않은 원형감옥) 우려도 나왔다. 특히 전통적으로 기술과 편리성보다 인권을 앞세웠던 유럽에서 프라이버시 정책에 변화의 조짐이 보이고 있다.

사생활 가치를 금과옥조처럼 여겨온 유럽은 코로나19 확산 방지를 위해 휴대전화 위치 정보를 방역에 활용하기 시작했다. 스페인 정부는 이동통신사에서 휴대전화 위치정보를 받아 전국 이동제한령 감시와 방역의 기초자료로 활용하고, 이탈리아는 한국 방식의 감염자 동선 추적 앱을 제시했다. 또 독일의 경우 IT 기술을 방역에 활용하되 정부가 만든 앱을 시민들이 내려 받으면 무작위로 아이디가 생성되는 방안을 제시했다. IT 전문가들은 "전염병 방지 확산을 위해 개인 정보를 어떻게 분석하고 활용하는지가 국가 및 산업 경쟁력을 좌우하는 시대가 왔다"고 내다봤다.

코로나19와 로봇

로봇도 코로나19 이후 예상되는 대표적 수혜주다. 코로나19 창궐 현장에서 로봇은 수많은 의료진 사이를 누비며 사람을 대신해 방역을 하고, 환자들을 돌봤다. 덴마크 로봇회사 UVD가 만든 로봇이 대표적이다. 이 로봇은 코로나19 사태에서 가장 큰 활약을 했다. UVD는 코로나19 사태가 나기 전인 2018년에 개발됐다.

자외선(UV)을 이용해 병실과 수술실을 소독할때 사용한다. 미증유의 코로나19 사태에서 빛을 발했다.

코로나19 발원지인 중국에서도 로봇은 큰 역할을 했다. 팔과 내시경이 달린 로봇이 면봉으로 환자 목 안의 검체를 체취했고, 내시경을 환자 목에 집어 넣어 기관지 상태도 살폈다. 중국 선양시에서는 카메라로 환자 얼굴을 인식해 의사가 처방한 약품을 전달하는 간호 로봇이 운영됐다.

국내서도 코로나19를 계기로 일부 병원에서 로봇을 도입했다. 로봇학계와 산업계는 코로나19처럼 감염력이 높은 바이러스와의 전쟁이 장기화되면 다양한 로봇이 더 등장할 것으로 예상했다.

에필로그

가문 가슴에, 어둡고 막막한 가슴에
푸른 하늘 열릴 날이 있을 거야
길이 없다고, 길이 보이지 않는다고
그대, 그 자리에 머물지 말렴
길이 끝나는 곳에서 길은 다시 시작되고
그 길 위로 희망의 별 오를 테니
(중략)
겨울나무처럼 그대는 고단하게 서 있지만
길은 끝나지 않았어, 끝이라고 생각될 때
그때가 바로, 다시 시작해야 할 때인 걸

–백창우, 「길이 끝나는 곳에서 길은 다시 시작되고」

백창우 시인의 말처럼 길이 끝나는 곳에서 길은 다시 시작된다. 코로나19는 언젠가 종식될 것이고, 인류는 새로운 길을 갈 것이다. 지난날의 역사는 우리에게 말한다. 위기를 두려워하지 말라고. 14세기 흑사병(페스트)이나 1차 세계대전 직후 스페인 독감, 그리고 2009년 이후 신종플루와의 싸움에서 승자는 항상 인류였다.

위기는 기회다. 실제 위기라는 말은 위험과 기회가 합친 말이다. 위험 속에 기회가 있다는 뜻이다. 2001년 9.11테러를 비롯해 지난 20년간 몇 차례 경제 위기가 있었다. 하지만 그때마다 인류는 잘 이겨냈고 디지털 바람은 거셌다. 코로나19 이후도 마찬가지 일것이다. 투자면에서 보면, 인공지능(AI) 등 디지털 투자는 더 늘어날 것이다. 자동차 경주에서 선두가 뒤바뀌는 곳은 직선 구간이 아니다. 곡선 구간, 즉 변곡점이다. 코로나19 이후 세계는 새로운 산업질서가 펼쳐질 것이고, 그 승자는 변곡점을 준비한 자가 될 것이다.

언택트를 콘택트처럼 만드는 5G와 실감기술의 기회

Opportunities for 5G and Immersive Technology to Make Non-Contact Seem Like Contact

2020년 1월 8일 한국에서 코로나19 감염 의심환자가 처음 확인되었다. 그후로 3개월이 지난 요즘은, 벚꽃놀이를 하며 많은 사람들이 거리로 나올 시기가 되었지만, 확진자 수가 1만 명을 넘고, 정부는 여의도의 벚꽃놀이를 통제 하고 있다. 거리의 가게에는 손님들을 찾아볼 수 없으며 엘리베이터를 탈 때는 왠지 모르게 서로에게 불편하다. 한국에 처음 코로나19가 알려졌던 1월에 요즘의 풍경을 상상할 수 있었던 사람들은 별로 없었을 것이다. 사람과 사람의 만남이 사라진 시대, 이러한 흐름을 일컬어 비대면, 영어로는 언콘택티드(Uncontacted) 이를 줄여서 언택트(Untact)시대라는 말이 실감나는 요즘이다.

언택트가 가져온 위기

경제는 사람과 사람의 만남을 기본으로 한다. 사람과 사람이 만나야 거래를 하고, 돈과 재화가 돈다. 코로나19에 의한 반 강제적인 언택트의 흐름은 소비심리를 위축시키고 만남에 기반한

양병석
Yang, Byongseok
코믹스브이 대표
ComixV, Representative
fstory97@comixv.com

많은 산업들을 위기로 몰고 가고 있다. 식당은 문을 닫았으며, 그나마 문을 연 곳은 마스크를 쓴 점주들이 한산한 식당에서 걱정스러운 표정을 짓고 있다. 여행업이나 전시에 기반한 MICE산업, 그리고 많은 이들을 모여서 하는 교육행사 심지어는 학교 같은 교육기관까지 사실상 거의 멈춰버렸다고 해도 과언이 아니다.

학교의 개학은 4월 6일로 한차례 미루어졌다가 4월 중순과 말 이후로 온라인으로 순차적으로 개학을 하게 되었다. 올해 처음 초등학교 입학을 앞두고 있던 8살 아들은 새 책가방을 싸두어 놓고 "나쁜 코로나19"를 연신 외친다. 평일에도 아빠는 집에 있지만 놀아주지 못해 야속하게 쳐다본다.

우리집뿐만 아니라 아마도 이러한 풍경은 2020년 코로나19가 가져온 새로운 풍경일 것이다. 2차 세계대전에 비견할 정도의 위기라는 미국의 발표만큼 전쟁 상황에 가깝게 느껴진다. 그리고 이러한 위기는 멈춰 버린 산업의 기업과 근로자들에게는 더욱 큰 위기로 다가오고 있고, 소득수준이 낮은 이들 일수록 직장을 잃어버렸을 가능성이 높고, 좁은 집에서 아이들을 돌보기는 더 어려운 상황일 것 같아 안타깝다.

그나마 다행인 것은 소비활성화 및 긴급 생계지원으로 정부에서 하위 70%에게 보조금을 지급하기로 했지만 코로나19가 언제 끝날지는 알 수 없는 상황이다.

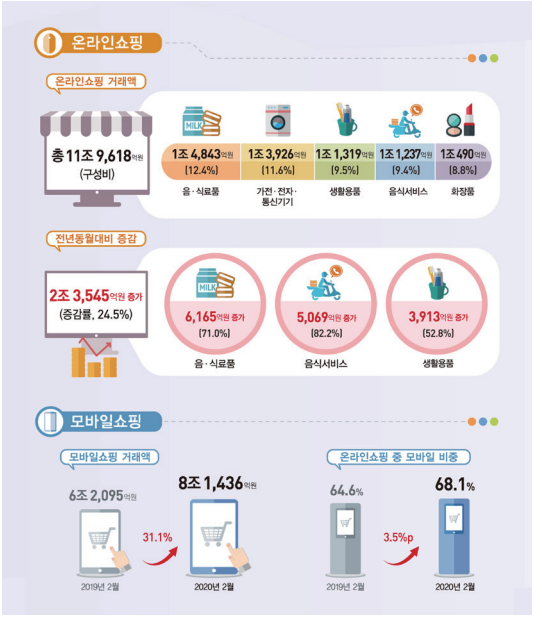
온라인 커머스로 보는 언택트의 흐름

하지만, 모든 업종들이 위기인 것은 아니다. 모바일 커머스와 새벽배송으로 성장을 하고 있던 쿠팡은 주문 폭주로 서버가 멈추는 일도 있었다.

통계청에 의하면 2020년 2월 온라인쇼핑 거래액은 11조 9,618억 원으로 전년동월대비 24.5% 증가, 모바일쇼핑 거래액은 8조 1,436억 원으로 31.1% 증가했다. 그중 세부 상품을 보자면 배달의 민족으로 대표되는 음식서비스는 82.2%가 증가했으며 쿠팡의 새벽배송으로 대표되는 농축수산물은 103.7%, 음·식료품 71.0%, 생활용품 52.8%나 증가했다.

하지만, 모든 온라인 쇼핑이 증가한 것은 아니다. 오프라인 상품을 온라인에서 거래하는 문화 및 레저서비스는 -60.0%, 여행 및 교통서비스는 -46.2% 로 감소하여 언택트 사회 변화를 극단적으로 보여주고 있다.

[그림 1] 2020년 2월 온라인 쇼핑 동향



※ 자료 : 통계청

코로나19의 위기는 한마디로 대면서비스의 위기라고 할 수 있다. 하지만, 대면서비스의 위기가 코로나19 때문이라고만 할 수 있을까? 쿠팡의 경우 기존 대면서비스로만 이루어졌던 신선 제품 구매 서비스를 온라인으로 끌어와 비대면 서비스로 변화시켰으며, 배달의 민족은 오프라인의 전단지 서비스를 모바일 온라인 플랫폼으로 변화시킨 것이 핵심이다.

또한 스타트업계에서 이슈가 되었던 '타다'의 경우, 평가가 엇갈리기는 하지만, 택시를 대체하는 좋은 서비스로 평가받는 이유 중 하나가 택시 기사의 불친절함과 안전의 문제를 해결했다는 점이었다. 특히 젊은 여성들에게는 택시 이용에 대한 안 좋은 경험이 있으며, 이들이 원하는 것은 택시 기사의 친절한 인사나 대화보다는 대화자체를 하지 않았으면 하는 언택트에 가까운 요구사항이었다. 현대사회에서 모르는 이를 만난다는 것은 스트레스로 여겨지고 있으며, 코로나19뿐만 아니라 다양한 이유로 개인의 안전 문제로 여겨지며 언택트의 요구가 커지고 있음을 반증하는 증거다.

이러한 증거는 쿠팡의 택배기사의 직고용에 대한 사회적 인식에서도 나타났었다. 쿠팡은 쿠팡맨을 만들어 낯선 택배 기사의 개인의 신뢰에서 회사의 신뢰 문제로 바꾸어 극복했다. 비대면 서비스의 증가는 사실 매슬로우의 2단계의 욕구인 "안전"에 대한 근본적인 욕구에 기반한다. 복잡해진 현대사회는 필연적으로 너무 많은 모르는 이들과 대면해야 되는 상황을 만들었으나 최근 스타트업들은 이를 해결하기 위한 다양한 아이디어들을 내놓았고, 그 근간에는 네트워크와 모든 이들이 하나씩 가지게 된 스마트폰이 큰 역할을 했다.

코로나19가 만든 뜻밖의 혁신

앞서 말했듯이 코로나19 이전에도 있었던 언택트의 흐름이지만, 콘택트가 너무 익숙해서 혹은 오프라인에 기반한 이권 문제로 언택트로 이관되지 못했던 것들이 있다. 언택트의 장점은 알지만, 사회적 관습으로 절대 넘어갈 수 없었던 장벽처럼 느껴지던 것들 이다.

대표적인 것이 "원격 근무와 미팅"이다. 화상회의 솔루션들은 대부분의 큰 기업들과 공공기관에는 이미 보급이 되어 있었지만 그 이용율은 그리 높지 않았다. "원격 업무와 미팅"은 근거리 업무와 미팅이 불가능한 특수상황에서만

하는 것이었지 일상적인 업무의 한가지로 받아들이기에는 어려웠다. 하지만, 코로나19가 변화시키고 있다. 피치 못하게 모두들 원격 미팅과 업무를 진행해야 하는 상황에 온 것이고, 거꾸로 피치 못할 때만 콘택트를 하는 상황으로 바뀌었다. 모든 업무는 언택트를 중심으로 다시 설계되고 있으며 이를 통해 원격 근무와 미팅의 장점과 단점들을 확실하게 발견하게 된다.

특히 원격 근무와 미팅이 어려웠던 이유 중 하나는 일이 잘되지 않을지 모른다는 걱정과 업무 프로세스의 익숙하지 않음에 있었는데 코로나19는 원격 근무가 가능한 모든 기업들이 원격 근무를 강제 학습하고 있다. 그리고 한번 학습된 이 경험은 원격 근무의 장점을 코로나19 이후에도 버리지 않을 가능성이 높다. 시간과 공간의 분리, 그리고 비동기적인 업무 프로세스는 그동안 그 가능성은 높게 평가하고 있었지만, 일부 최첨단 IT기업들만 하는 것으로 믿어졌던 것들이다. 하지만 이번 기회를 통해 기업이나 근로자나 반강제적으로 학습을 하고 있다.

그리고 이런 상황이 지속되면 업무를 잘하는 이들이 더

높은 평가를 받게 될 것이고, 원격 근무의 장점도 매우 탁월하기 때문에 기업의 인재들은 원격 근무를 이해하고 잘 하는 이들이 보상을 받게 되어 정착이 될 것으로 보인다.

디지털 교과서나 온라인 학습은 어떨까? 가장 변화가 느리다는 교육계가 반강제적으로 온라인 수업을 하지 않을 수 없게 되었다. 교사들은 강제적으로 동영상 강의를 찍어야 하는 상황에 내몰렸다. 한국이 디지털 교과서 정책을 처음 만든 것은 2007년으로 13년이 지난 지금까지도 진전은 크지 않았다. 하지만 코로나19는 단 한번에 교육계 전체를 디지털 기반의 온라인으로 전환시키고 있다.

단단한 철옹성 같던 의료계 역시 2020년 2월 22일 보건복지부 정례브리핑을 통해 원격진료 시행방안 등을 구체화한 '전화상담·처방 및 대리처방 한시적 허용방안'을 발표했다. 국민이 의료기관을 이용하는 과정에서 혹시라도 코로나19에 감염되는 것을 방지하기 위해 한시적으로 '특례'를 인정하겠다는 취지로 밝히고 있지만, 과연 코로나19 종료 이후에 한번 경험한 원격진료에 대한 수요를 외면할 수 있을지 모르겠다.

이러한 교육계나 의료계의 언택트 서비스의 증가는 공급자들은 오프라인이 아닌 온라인 전체에서 경쟁을 하게 되므로 오프라인 지역에 한정되던 기존과 비교할 수 없을 정도로 경쟁 시장으로 내몰리게 되며 서비스 질의 증가를 가져온다. 태풍 후에 풍어라고, 각종 습관과 이권이 엮인 제도로 단단하던 둑이 코로나19의 큰 쓰나미로 체질 개선을 바랄 수 있지 않을까 싶다.

유효한 콘택트의 경험 가치

하지만 콘택트가 가치가 없는 것은 아니다. 사람은 사회적 동물이다. 안전도 혼자 존재함으로서 확인 하는 것이 아니라 신뢰할 만한 누군가와 함께 있을 때 확인된다. 때문에 사회적 거리두기의 요즘의 또 하나의 문제는 혼자 생활하는 이들은 오히려 사회적으로 고립되어 안전에 대한 불안을 느낄 수 밖에 없다.

또한 경제적으로 사회가 성숙할수록 소유의 가치보다는 경험의 가치로 이동한다. 그리고 그 경험은 오감을 이용한 활동으로 여가활동과 오프라인의 각종 문화 활동이 더 큰 가치로 평가된다. 영화관의 가치가 고품질 독점 콘텐츠만이 있는 것이 아니다. 그 공간에서 경험하는 연인, 혹은 가족과의 유대경험, 그리고 그곳에서만 먹을 수 있는 먹거리와 분위기가 바로 영화관의 가치다. 영화관이 콘텐츠 가치로만 평가 받았다면, 온라인 고품질 독점 콘텐츠를 만드는 넷플릭스의 성장은 고스란히 영화관 산업에 타격으로 돌아왔을 것이다. 물론 실제 타격도 있고 그 규모가 증가하고는 있지만, 아직 영화산업 규모에 비해 넷플릭스의 오리지널 콘텐츠가 차지하는 비중은 미미하며 영화산업은 넷플릭스 성장 중에도 함께 성장했었다. 또 하나의 사례를 들자면 스타벅스가 커피가 아닌 문화를 판다고 했던 이유가 그것이다. 때문에 오프라인 공간에 대한 경험설계가 중요하고, 소비자가 접근할 수 있는 상권이 매우 중요하다. 하지만 코로나19는 이 가치를 무색하게 할 정도로 고강도의 위험이 되어버려 현재는 최대 피해 업종이 되었다. 이는 가치가 손실됐다기 보다는 위협이 가치를 압도해버렸다고 해석하는 것이 타당할 것이다.

또한 온라인과 비대면이 익숙한 젊은 층을 타겟한 대면 서비스의 스타트업이 성공을 거두는 흐름도 있다. 트래바리, 문토, 버핏서울 등 관심사를 중심으로 한 온/오프라인의 콘택트를 기반으로 한 스타트업들이 있다. 18세기 중반 프랑스, 지성인과 예술가가 한데 모여 토론을 펼치고 지식을 나누던 사교 집회를 뜻하는 "살롱 문화"를 테마로 한 스타트업들이다.

무분별한 콘택트는 온라인에도 존재한다. 가볍고 무분별하고 소비적인 SNS에서 피로해진 젊은 층들이 신뢰할 만한 정보와 사람들을 찾아가고 있는 흐름이다. 분명한 것은

온라인이 그동안 가벼운 것을 추구해가던 흐름에서 좀 더 심도 있는 관계와 좋은 정보를 추구하기 시작했다는 점이다. 하지만 역시 코로나19의 위협은 이러한 오프라인의 모임에 타격을 주고 있는 것은 유효하다.

언택트를 콘택트처럼 만드는 기술들의 기회

그럼 콘택트의 가치를 언택트로 가져올 수는 없을까? 이에 해당하는 대표적인 현재의 솔루션은 화상회의다. 코로나19로 인해 미국의 주가가 바닥을 치고 있는 가운데에도 화상회의 솔루션인 Zoom만은 상승세에 있다. 원격과 재택근무가 증가하면서 디지털 기기 판매도 증가했다. 이런 흐름에 따라 최근 인터넷이 느리다는 소비자의 불만도 늘어나고 있는 것으로 보아, 언택트를 콘택트처럼 만드는 기술들은 큰 수요증가를 이루고 있는 것이 분명하다.

그러나 현재의 PC/모바일의 기술들은 작은 디스플레이를 사이에 두고 업무를 할 수 밖에 없다. 동시에 여러 사람과 일을 하거나 협업을 한다면 한사람당 보여지는 해상도는 더욱더 줄어든다. 다수의 원격의 음성을 동기화하여 합성하고, 제때에 다수의 영상을 모든 클라이언트에 보여주는 것도 현재의 네트워크 상황에서는 쉬운 일이 아니다. 그나마 초고속 통신망이 발달한 한국에서의 상황이 이 정도이니 인터넷이 발달하지 못한 해외의 상황은 더 말할 것도 없다.

언택트를 콘택트처럼 만드는 5g 실감미디어

다행히도 한국은 확실히 IT부문에서는 해당 분야의 선두국가이며, 큰 기회로 만들 수 있다. 쿠팡의 로켓프레시는 아마존 프레임 프레시를 벤치마킹한 패스트팔로우 서비스지만 잘 발달된 운송망과 스마트폰 이용율, 대다수가 도시에 주거하는 만큼 더 수준 높은 서비스를 더 많은 지역에서 받을 수 있으며 우아한 형제의 광고 카피처럼 언택트 음식 서비스인 배달 문화가 잘 발달한 나라다.

특히 5g를 세계최초로 상용화한 국가로서 세계 최고수준의 초고속 인터넷망을 가지고 있으며 5g에서 밀고 있는 VR/AR 곧 실감 서비스는 PC/모바일 보다 언택트를 콘택트처럼 만들 수 있는 기술로서 그 기회가 있다.

더 빠른 응답속도와 대용량 데이터 처리가 가능한 네트워크는 Zoom을 뛰어넘는 더 많은 이들과의 동기화 서비스를 가능하게 해주며, 실감 서비스에 필요한 초고해상도영상과 대량의 3D정보들도 전달이 가능하게 해준다.

VR HMD를 직원 수 만큼 가지고 있는 코믹스브이에서는 어디까지 이러한 경험이 가능할까 싶어, 페이스북의 오큘러스고를 이용하여 VR 원격 회의를 진행해보았다.

서로의 아바타를 이용해 동시에 문서를 볼 수 있는 솔루션을 원했는데 딱 맞는 솔루션을 찾지 못해, 오큘러스고의 파티기능을 이용해 갤러리 앱을 사용해보았다. 여기서 말하는 파티는 온라인게임의 파티와 유사한 개념으로 오큘러스 플랫폼 내에서 파티를 맺은 이용자끼리 음성대화를 하거나 파티기능을 지원하는 특정 앱에서는 함께 서비스를 즐길 수 있다. 온라인 회의를 위해 발표용 문서를 이미지파일로 저장하고, 갤러리에 팀(파티)원들을 초대하여 원격회의를 진행했다. 갤러리는 회의에 최적화된 솔루션이 아니다 보니 이미지 순서를 수동으로 정렬해야 했고, 동시에 4명밖에 접속하지 못한다는 것을 확인했다.

하지만 효과는 놀랄 만큼 좋았다. 적어도 옆에 팀원들이 함께 있다는 기분이 들었고, 전면에 100인치 정도 펼쳐진 발표 이미지는 한눈을 팔 수가 없었다. HMD를 벗기 전에는 애시당초 다른 곳에 눈 돌릴 곳이 없었다. 약 20분에 걸친 회의 끝에 배터리가 방전되어 회의가 종료되었지만, 다들 재밌는 경험이었다고들 고백했다.

아직은 다수의 사람들이 동시에 접속해서 가볍게 쓸 수 있는 HMD가 없고, 해상도도 한정적이었지만 2020년부터 출시 예정인 글래스타입의 고해상도 VR HMD가 나오면 훨

[그림 2] 코로나19로 화상회의하는 코믹스브이(컨셉이미지)

씬 더 나은 환경을 기대할 수 있을 것 으로 기대된다.

VR/AR, 실감현실은 원격 공간에서 같은 공간에 있는 것처럼 느끼게 하며, 작은 디스플레이의 한계를 벗어나는 일은 언택트를 콘택트로 만들 수 있는 핵심 기술 중에 하나임이 분명하다. 더군다나 최근 페이스북에서 손을 트랙킹하는 NUI까지 등장하였으니 조금 더 자연스러운 언택트가 될 것으로 기대된다.

더불어 한국이 4차 산업혁명의 키워드 중 하나로 5g와 VR/AR을 선정한 것은 탁월한 선택이었다고 생각한다.

한국의 힘을 믿는다.

코로나19는 분명 매우 큰 위기다. 하지만 위기는 늘 변화를 이끌고 그 변화에 적응하는 개체들은 큰 번영을 누렸다는 것이 지구가 가르친 진화의 역사다. 다행스러운 것은 한국이 그 변화 속에서 살아남을 수 있는 다양한 무기를 가졌으며 패스트팔로우가 아닌 퍼스트무버로서 인정을 받고 있다는 것이다. 위기극복을 취미로 여기는 민족이라는 인터넷의 농담처럼, 국난에 강한 힘을 가졌음은 분명하다.

전 세계의 위기인 코로나19는 한국에게는 단순히 위기가 아니라 얼마나 슬기롭게 극복하느냐에 따라 세계 선도국가가 될 수 있는 큰 기회라고 믿는다.

AI 성능개선을 위한 클라우드 GPU 가상화 기술의 현황과 향후방향

Current State and Future Directions of GPU Virtualization Technology in the Cloud Computing for Improving AI Performance

Executive Summary

GPU(Graphic Processing Unit)를 클라우드 컴퓨팅에 활용하는 연구와 상용화가 점차 확산되고 있다. 인공지능이 주목받게 되면서, 이를 동작시키기 위한 인프라이자 데이터를 공급하는 저장소인 클라우드도 수요와 서비스가 늘고 있다. 인공지능의 성능은 데이터의 학습량과 그 학습에 필요한 컴퓨팅 파워에 따라 달라져서 고성능 클라우드 컴퓨팅이 요구된다. GPU는 다수의 코어를 가져 병렬 처리에 유리하며, CPU 대비 단위 코어당 가격이 저렴하다. 이에 따라 GPU를 클라우드 컴퓨팅 자원으로 활용하는 방법이 등장했다. 그런데 클라우드 컴퓨팅은 가상화가 전제되는 기술이기 때문에 GPU를 온전한 클라우드 자원으로 활용하기 위해서는 역시 가상화가 필요하다. 그간 GPU 가상화는 어렵고 복잡한 기술이었으나, 여러 연구를 거쳐 최근에는 하드웨어에서 가상화를 지원하는 제품이 출시되고 있다. 이 보고서에서는 GPU 가상화 기술의 등장배경과 그간의 현황에 대하여 살펴보고, 클라우드 GPU 가상화 연구에서 기술 수준을 발전시킬 수 있는 과제와 방향을 논의하고자 한다.

안성원
Ahn, Sungwon
선임연구원
Senior Researcher, SPRi
swahn@spri.kr

Research into using GPU(Graphic Processing Unit) for cloud computing and its commercialization is gradually spreading. AS AI gains attention, cloud computing that is the infrastructure to operate AI and the storage for providing data is also increasing in demand and services. The performance of AI depends on the amount of data learned and the computing power required to learn, which is why AI requires high performing cloud computing. GPU has multiple cores, which makes it more advantageous for parallel processing, and its price per core is lower than that of CPU. This has led to GPU being used as cloud computing resources. Since cloud computing is a technology for which virtualization is a prerequisite, virtualization is required to make the most of GPU for cloud computing. GPU virtualization has been a difficult and complex technology, but after many researches, products that support virtualization in hardware have been released. This paper examines the background and the current state of GPU virtualization technology, and discusses the challenges and directions to advance the level of technology in cloud GPU virtualization research.

1 논의배경

AI의 핵심 인프라로써 점점 더 높은 성능이 요구되는 클라우드 컴퓨팅

클라우드 컴퓨팅은 인공지능(Artificial Intelligence, AI)의 성능향상을 위한 학습데이터를 제공하는 저장소이자 컴퓨팅 파워를 제공하는 필수 인프라다. 클라우드 컴퓨팅은 이기종의 컴퓨팅 노드를 활용하여 컴퓨팅 인프라의 소유 및 유지비용을 줄이고, 더 높은 성능과 범용성, 에너지 효율성을 높이는 장점이 있다. 이러한 장점들은 AI의 성능을 올리는 데에 매우 적합하여, 많은 AI 연구와 서비스에 클라우드 컴퓨팅이 활용되고 있다. 최근 AI를 활용한 서비스가 우리생활 곳곳에 실현되고 있고 니즈(Needs)가 점차 다양해짐에 따라 클라우드 컴퓨팅의 보다 안정적이고 우수한 성능이 요구되고 있다.

클라우드 컴퓨팅 성능향상을 위해 가성비 좋은 GPU 활용의 확대

높은 클라우드 컴퓨팅 성능을 위해서는 빠른 작업 처리가 가능한 고성능 프로세스 유닛(Process Unit, 이하 코어(Core))이 다수 필요하다. 특히 딥러닝(Deep Learning)과 같은 AI는 많은 수의 데이터를 학습할수록 성능이 좋아지는데, 이를 위해서는 다중 코어를 갖춘 병렬컴퓨팅 환경이 사실상 필수적이다. 다중 코어를 갖춘 병렬컴퓨팅 환경을 구축하기 위한 최근의 패러다임은 강력한 성능의 범용코어가 있는 멀티코어(Multi-Core) CPU[1]와 단순하지만 더 많은 코어를 가지는 GPU[2]와 같은 가속기(Accelerator)의 두 가지 아키텍처를 결합하는 방식을 추구하고 있다. 이러한 이기종 결합의 컴퓨팅 시스템은 비용과 성능의 효율성을 모두 높일 수 있는 방식으로 지난 2016년 구글의

1 두 개 이상의 독립 연산 코어를 갖는 중앙처리장치(Central Processing Unit)로 보통 다수의 코어를 하나의 칩에 집적해 놓은 형태, 코어 수는 2개에서 72개까지 다양함
2 Graphic Processing Unit, 그래픽처리장치

알파고(AlphaGo)에도 활용되었다. 그래픽처리장치인 GPU를 클라우드 컴퓨팅에 활용하는 이유는 비용 효율성이 좋기 때문이며, 어려운 이기종 시스템의 결합이 있어야 한다.

이 보고서에서는, AI를 위한 필수 인프라인 클라우드 컴퓨팅에서 GPU를 활용하는 목적과 GPU를 활용하기 위해 연구해왔던 기술들을 살펴보고, 향후 기술발전을 위한 연구 주제와 방향을 논의하고자 한다.

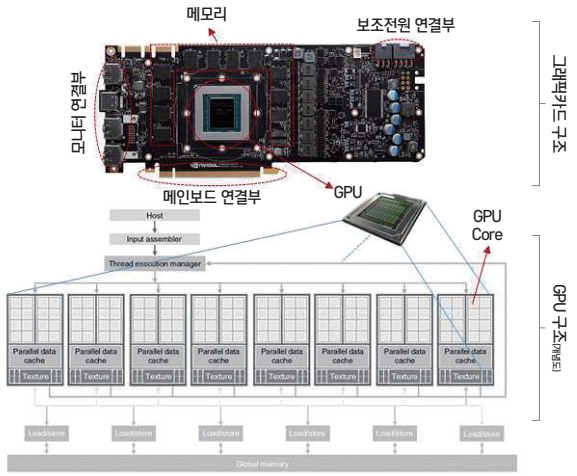
2 GPU 기술

2.1. GPU의 구조

○ GPU는 컴퓨터의 그래픽을 처리하기 위한 장치로 다수의 코어를 보유

GPU는 컴퓨터 모니터의 수천~수백만 개의 픽셀(Pixel[3])로 투영되는 그래픽의 처리를 위한 장치로 병렬연산에 특화되어 있다. 때문에 GPU는 구조적으로 다수의 코어를 보유하고 있으며, [그림 2-1]과 같다. [그림 2-1] 상단부는 PC에 일반적으로 사용하는 그래픽카드의 구조를 나타낸다. 중앙에 GPU가 탑재되어 있으며, 그래픽연산에 필요한 메모리, 컴퓨터 메인보드의 슬롯(PCIe[4])연결부, 그래픽 처리 결과를 모니터에 송신하는 연결부로 구성되어 있다.

[그림 2-1] 그래픽카드와 GPU의 구조



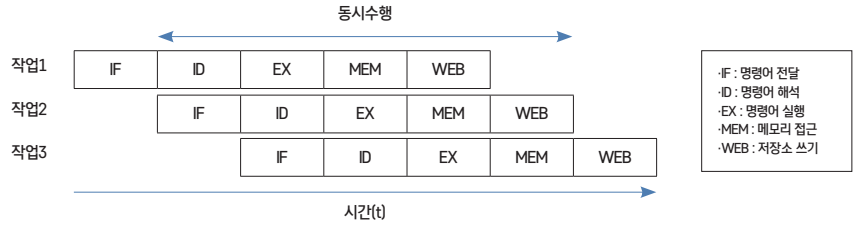
※ 자료 : NVIDIA(nVidia GTX 1080 ti 모델의 사례)

3 모니터에서 화면을 구성하고 있는 가장 작은 단위인 네모 모양의 작은 점, 화소
4 Peripheral Component Interconnect Express, 메인보드(Main Board, Mother Board)에 그래픽, 사운드, TV카드 등 각종 확장카드를 사용하기 위해 만들어진 규격

그래픽 카드의 중앙에 위치한 GPU는 그림 하단부와 같이 다중 코어로 구성되어 있다. GPU는 병렬 응용 프로그램을 효율적으로 실행하기 위해 기존의 멀티 코어 프로세서와는 근본적으로 설계가 다르다. 기존의 멀티 코어 프로세서는 지연시간(Latency)을 줄여서 처리 속도를 높이는 것을 목표로 설계되었다. 주로 단일 작업에 대해 빠른 속도의 연산 수행을 한다. 반면, GPU는 처리량(Throughput) 중심으로 설계가 되어 있다. 처리량을 높이기 위해 비교적 구조가 간단한 수천 개의 코어들과 높은 대역폭의 메모리를 탑재한 형태를 띠고 있다. 이러한 설계는 병렬로 동작하는 프로그램이 GPU에서 적절히 분배되어 실행되게 함으로써 응용 프로그램의 실행 처리량을 극대화할 수 있다.

예를 들어, 실행되고 있는 일부 쓰레드(Thread)[5]가 메모리 접근을 위해서 긴 대기 시간을 갖는다고 가정할 때, 이 작업이 끝날 때까지 기다렸다가 다른 작업을 이어서 수행하는 것은 전체 작업량의 기준으로 볼 때 비효율적이다. 이러한 작업 중 대기 시간을 숨기기 위해 GPU의 하드웨어 스케줄러(Scheduler)는 실행 가능한 또 다른 쓰레드를 스케줄링(실행, 접근 등) 함으로써 병렬 처리를 가능하게 할 수 있다. 이렇게 하면 각 쓰레드의 개별 실행 시간은 길어질 수도 있지만, 전체 프로그램 입장에서 총 실행 처리량은 향상시킬 수 있는 장점이 있다. 이를 파이프라인(Pipe-Line) 기법이라고 하며, 파이프라인은 멀티코어 CPU에서도 활용되는 방식이지만, GPU는 더 많은 수의 병렬 코어를 통해 훨씬 더 방대한 양의 병렬처리를 가능하게 한다.

병렬처리(Parallel processing)

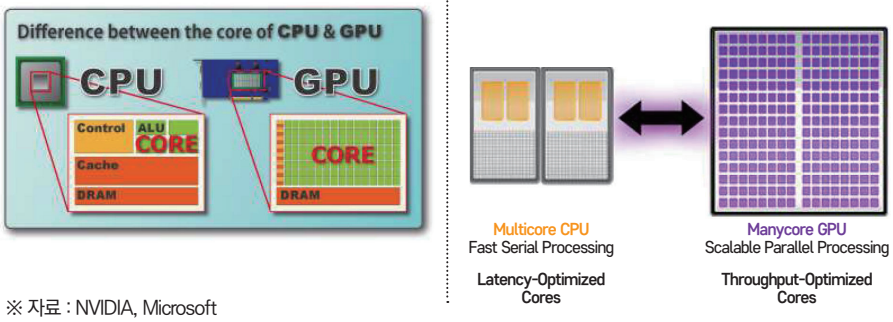


[그림 2-2] 병렬처리 기법의 예시(파이프라이닝 기법)

- 병렬처리는 작업을 최대한 동시에 실행하여 시스템의 성능을 올리는 방법으로 파이프라인, 멀티쓰레드, 분산컴퓨팅 외에도 여러 종류가 있음

5 프로그램이 실행될 때의 일련의 과정을 프로세스(Process)라고 하며, 이 프로세스 내에서 실행되는 세부 흐름의 단위를 쓰레드(Thread)라고 함. 또한, 하나의 프로세스 내에서는 다수의 쓰레드가 실행될 수 있으며, 이러한 실행 방식을 멀티쓰레드(Multi-thread)라고 함.

CPU와 GPU의 비교



※ 자료 : NVIDIA, Microsoft

[그림 2-3] CPU와 GPU의 구조 비교

- CPU는 빠른 일련의 작업(Fast Serial Processing)을 목표로 하여 지연시간에 최적화된 코어를 보유한 단일 또는 멀티코어(Multi-Core) 구조
- GPU는 매니코어(Many-Core)의 형태로 확장 병렬 프로세싱(Scalable Parallel Processing)을 목표로 하여 출력값에 최적화된 코어로 구성

2.2. 왜 GPU를 컴퓨팅 자원으로 쓰는가?

○ 그래픽 처리가 아닌 다른 일반적인 목적으로 GPU를 활용하는 GPGPU

앞서 살펴본 바와 같이 GPU는 CPU보다 많은 수의 코어(Core)를 가지고 있다. GPU가 보유한 코어들은 일반적으로 CPU의 코어보다 성능은 낮지만, 이러한 코어를 수백~수천 개 가지고 있어서 CPU보다 단위 코어 당 가성비가 높다. 이는 같은 비용으로 더 많은 수의 계산 가용한 코어를 보유할 수 있다는 의미이다.

[표 2-1] 상용 CPU와 GPU의 단위 코어 당 가격 비교

구분	모델명		코어 수(개)	가격($)	코어 당 가격($)
CPU	Intel Core-i9 9900k(3.6Ghz) (2018.10. 출시)		8	549	68.6
GPU	nVidia Geforce RTX 2080 Ti (2018.9. 출시)		4,352	1,199	0.275

※ 자료 : Intel, NVIDIA

특히, 높은 성능의 단일 코어가 필요한 계산 업무보다, 단일 코어대비 속도는 약간 떨어지더라도 훨씬 많은 양의 좀 더 단순한 계산 업무를 수행하는 데에 유리하다. 좀 더 직관적인 예를 들어, 4명의 대학원생과 100명의 유치원생 그룹이 있다고 가정할 때, 1개의 미적분 문제를 푸는 것은 대학원생 그룹이 유리할 수 있으나, 100개의 편지봉투에 우표를 붙이는 작업은 유치원생 그룹이 훨씬 유리하다. 앞선 예에서 대학원생 그룹은 CPU, 유치원생 그룹은 GPU에 비유할 수 있다. 이처럼, 본디 그래픽처리가 목적인 GPU를 다른 일반적인 목적으로 활용하는 것을 GPGPU[6]라고 한다. GPU를 범용적인 컴퓨팅 자원(Resource)[7]으로 활용하면 상대적으로 낮은 비용을 들여 컴퓨팅 파워를 확보할 수 있다.

○ 한계에 다다른 CPU의 성능

현재의 CPU는 코어 수가 점차 증가하고 있음에도 불구하고, 정교한 제어 로직과 대형 캐시 메모리를 사용하여 순차 프로그램의 지연시간을 줄이는 것을 목표로 해왔다. 이러한 설계는 각 코어에서 순차적으로 수행되는 코드의 실행 시간을 줄이기 위한 것에 최적화되어 있으므로 하나의 프로세서 패키지에 더 적은 코어가 탑재되는 대신 각 코어에 복잡성은 증가하게 된다. 더군다나, CPU의 집적도[8]를 높이기 위한 나노 공정은 현재 기술로 7nm[9]가 한계이며, 그 이하의 간격으로 집적을 할 경우 간섭이 발생하여 성능이 떨어진다.

또한, CPU의 처리속도를 나타내는 클럭(Clock) 수는 4Ghz 내외 수준을 효율성의 한계치로 보고 있다. 그 이유는 집적도 향상 및 클럭 수 향상에 따른 전력소비의 증가와 발열문제를 해결하는 것에 한계가 있기 때문이다. 이로 인해 현재의 CPU의 단일 코어 당 성능은 한계수준에 다다랐으며, 근래에 출시되는 CPU는 코어를 여러 개 두는 것을 통해 성능을 높이는 것으로 설계 방향이 잡혀있다. 물론, 여전히 복잡한 구조를 갖는 고성능 칩셋이기에 [표 2-1]에서 살펴본 바와 같이 단위 코어 당 가격도 비싸다.

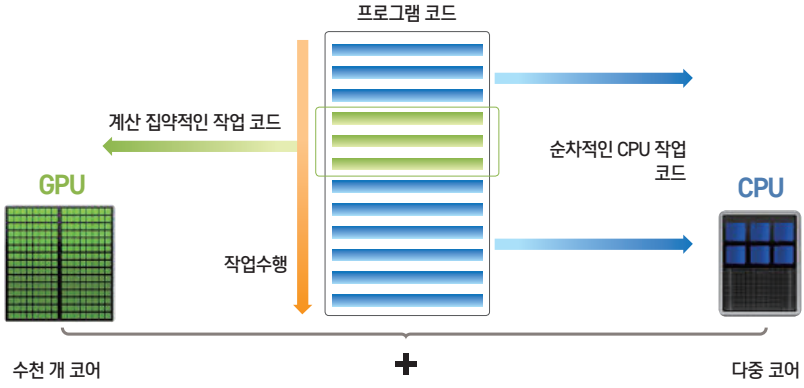
○ 고성능 컴퓨팅 패러다임의 변화 : CPU + GPU

최근 고성능 컴퓨팅(HPC[10]) 연구자들은 프로그램의 수행 시간을 단축시키기 위한 방법으로 강력한 성능의 범용 CPU 코어가 있는 멀티 코어 프로세서와 단순하지만 많은 코어를 가지는 그래픽처리 가속기인 GPU의 두 가지 아키텍처를 결합한 컴퓨팅을 추구하고 있다. 멀티 코어 프로세서와 GPU를

6 General Purpose computing on GPU, 컴퓨터 그래픽 처리를 목적으로 하는 GPU를 전통적으로 CPU가 맡아왔던 응용프로그램의 계산에 활용하는 기술
7 CPU, 메모리, 네트워크, 서버, 스토리지, 애플리케이션 등 컴퓨터에서 가용한 요소
8 Level of Integration, 실리콘 칩(반도체)에 집적할 수 있는 게이트(소자)의 수로 하나의 실리콘 칩에 논리 소자가 몇 개나 포함되어 있는지를 나타내는 정도
9 Nanometer(나노미터), 10억 분의 1미터(m)
10 High Performance Computing

결합한 이기종 시스템은 제어 집약적 구성 요소와 고도의 데이터 병렬 구성 요소를 모두 갖추고 있어서 고성능 컴퓨팅 응용 프로그램의 다양한 요구 사항을 충족시킬 수 있다. [그림 2-4]는 GPU와 CPU의 결합한 연산 방식을 나타낸다.

[그림 2-4] GPU + CPU 아키텍처의 작업수행 사례



응용프로그램을 수행할 때 대량의 계산 작업이 필요한 계산 집약적인 부분을 GPU로 넘기고, 나머지의 코드는 CPU에서 처리하는 것으로 프로그램의 수행속도를 비약적으로 증가시킬 수 있다. 쉬운 이해를 위해 앞서 예로든 대학원/유치원생의 사례를 다시 응용해 보자. 프로그램 코드의 작업이 미적분 문제 3개를 풀다가, 우표붙인 편지봉투 500개를 만들고, 다시 2차방정식 6문제를 푸는 일련의 과정으로 되어 있다고 가정한다. 4명의 대학원생이 이 일련의 과정을 모두 마치는데 드는 시간은 제법 걸릴 것이다. 가장 많은 시간이 소요되는 병목구간은 편지봉투에 우표붙이기 작업으로, 1인당 125개의 편지봉투에 각각 우표를 붙여야 한다. 이 작업만 따로 떼어서 100명의 유치원생들에게 나눠준다면 1인당 5개씩만 만들면 되므로, 작업속도는 월등히 빨라진다. 병목 구간의 작업 시간이 단축될 뿐만 아니라, 유치원생들이 우표를 붙이는 동안 대학원생들은 나머지 방정식 문제를 풀고 있으면 된다. 전체 작업의 완료 시간이 크게 단축될 수 있다.

GPU를 사용하는 컴퓨팅 시스템이 점차 확산되고 있다는 점은 최신 Top500 목록에서 나타난다. 현재 슈퍼컴퓨터의 27% 이상[11]이 CPU와 GPU를 결합한 시스템이다. 또한, 이러한 패러다임은 가속화(19%(2017) → 27%(2019)) 되고 있다.

11 nVIDIA(2019.11.19.), "Record 136 NVIDIA GPU-Accelerated Supercomputers Feature in TOP500 Ranking"

Top500 10위 내 슈퍼컴퓨터

- Top500 목록 중 10위권 내의 슈퍼컴퓨터들의 절반이 GPU를 활용하는 것으로 나타남(Summit, Sierra, Piz Daint, ABCI, Lassen 등)

[표 2-2] 전 세계 Top10 슈퍼컴퓨터

순위	이름	국가	시스템	코어 수(개)	성능(TFlop/s)
1	Summit	미국	IBM Power System AC922, IBM POWER9 22C 3.07GHz, ***NVIDIA Volta GV100***, Dual-Rail Mellanox EDR Infiniband	2,414,592	148,600.0
2	Sierra	미국	IBM Power System AC922, IBM POWER9 22C 3.1GHz, ***NVIDIA Volta GV100***, Dual-Rail Mellanox EDR Infiniband	1,572,480	94,640.0
3	Sunway TaihuLight	중국	Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway	10,649,600	93,014.6
4	Tianhe-2A	중국	TH-IVB-FEP Cluster, Intel Xeon E5-2692v2 12C 2.2GHz, TH Express-2, Matrix-2000	4,981,760	61,444.5
5	Frontera	미국	Dell C6420, Xeon Platinum 8280 28C 2.7GHz, Mellanox Infiniband HDR	448,448	23,516.4
6	Piz Daint	스위스	Cray XC50, Xeon E5-2690v3 12C 2.6GHz, Aries interconnect, ***NVIDIA Tesla P100***	387,872	21,230.0
7	Trinity	미국	Cray XC40, Xeon E5-2698v3 16C 2.3GHz, Intel Xeon Phi 7250 68C 1.4GHz, Aries interconnect	979,072	20,158.7
8	AI Bridging Cloud Infrastructure (ABCI)	일본	PRIMERGY CX2570 M4, Xeon Gold 6148 20C 2.4GHz, ***NVIDIA Tesla V100 SXM2***, Infiniband EDR	391,680	19,880.0
9	SuperMUC-NG	독일	ThinkSystem SD650, Xeon Platinum 8174 24C 3.1GHz, Intel Omni-Path	305,856	19,476.6
10	Lassen	미국	IBM Power System AC922, IBM POWER9 22C 3.1GHz, Dual-rail Mellanox EDR Infiniband, ***NVIDIA Tesla V100***	288,288	18,200.0

※ 자료 : Top500.org
※ 주 : 시스템 내 GPU 활용은 볼드 이탤릭체 표기
※ TFlop/s : 테라플롭스, 슈퍼 컴퓨터의 연산성능 척도로써, 1TFlop/s는 초당 1조(1012) 회의 부동소수점(Flop)연산을 수행

2.3. 클라우드에 GPU컴퓨팅이 필요한 이유

○ 클라우드 인프라에 대한 수요 증가 및 고성능 요구

AI 시대를 맞이하여 전통의 제조업, 서비스업 기업들도 디지털 전환을 추진하는 사례는 이미 많이 있다. 후발 주자들이 글로벌 경쟁력을 갖추기 위해 클라우드 서비스를 새롭게 제공하거나 인프라 확보를 추진하는 사례도 많다.

특히, AI 서비스의 오류를 줄이고 학습 성능을 끌어올려 완성도 있는 시스템을 확보하기 위해서는 클라우드가 필수적이다. 머신러닝(Machine Learning, 기계학습)의 방법론 중 하나인 딥러닝(Deep Learning)은 수많은 데이터의 지속적인 반복학습을 통해 특징을 추출하고 판단의 근거로 삼는다. 여기서 '반복학습' 작업을 수행하는 데에는 많은 계산능력을 필요로 한다. 클라우드는 여러 개의 계산 자원들을 묶어서 활용할 수 있기 때문에 딥러닝에 요구되는 큰 계산능력을 단일 컴퓨터 보다 쉽게 제공할 수 있다. 이 때문에 클라우드 인프라에 대한 수요는 꾸준히 증가하고 있다. 또한, 쏟아지는 빅데이터를 정제하고 관리하며, 이를 기반으로 새로운 AI 서비스의 학습양분으로 활용하는 사례가 늘면서 더 빠르고 효율적인 클라우드 컴퓨팅에 대한 요구도 증가하고 있다.

보다 큰 계산능력과 효율성을 제공하기 위해서는 기존의 CPU로만 구성된 머신들을 통합하는 것만으로는 한계가 있다. 앞서 언급한 반복학습 작업도 당연히 최대 10개 내외의 코어를 갖는 CPU 보다 수천 개의 코어를 갖는 GPU가 동시 다발적으로 작업을 수행하는 것이 유리하다.

○ 가격대비 성능의 효율성 확보

클라우드의 높은 컴퓨팅 능력 요구사항을 만족시키기 위해 GPU 수준의 수천 개 코어로 이루어진 시스템을 CPU 칩으로만 구축한다면 천문학적인 비용이 수반된다. 지난 2012년 구글은 구글브레인(Google Brain)이라는 신경망 시스템을 개발했다. 이 시스템은 고양이 이미지 1천만 개를 학습해서 고양이를 분별해 내는 인지 능력을 보여주었다. 그런데 이 시스템은 CPU 서버 1천 개를 병렬로 연결한 딥러닝 클라우드 인프라를 활용하였다. 당시 인프라 구축비용은 50억 원 이상이었으며, 전력소모는 60만 와트(W)에 달하여 구축과 운영에 많은 비용이 소요되었다. 이후 GPU 제조 업체인 엔비디아(NVIDIA)에서 GPU를 활용한 서버 3대로 동일한 성능을 제공할 수 있음을 입증[12]했고, 구축비용은 3,300만 원 수준에 전력 소모량도 4천 와트에 불과했다.

상대적으로 복잡한 매커니즘을 갖는 CPU들 간의 업무 분배 및 통합관리를 위한 시스템의 구현 또한 어려운 일이다. 일례로, CPU 제조업체인 인텔(Intel)에서 Xeon Phi[13]라는 72개(4세대 기준)의 CPU 코어로 이루어진 병렬처리 전문 제품을 상용화 하였으나, 병렬컴퓨팅 시장에서 GPU 제품 대비 효율성 측면의 경쟁에서 밀렸으며, 제품 생산을 중단을 발표한 것이 현실이다.[14]

앞서 언급하였지만, 빅데이터를 전처리하고 가공하는 일과 같은 작업은 코어의 클럭 수가 높은 것 보다, 코어의 숫자가 많아야 유리한 작업에 속한다. GPU 코어의 클럭 수가 현재의 CPU 코어 대비 낮더라도 코어의 숫자가 압도적으로 많고, 최근 출시되는 GPU 코어의 클럭 수(1.6Ghz)[15]는 수년 전의 CPU 클럭 수 못지않은 성능을 가지고 있다. 이것은 하드웨어의 발전은 속도 대비 가격은 상대적으로

12 NVIDIA/스탠포드(2013), "Deep Learning with COTS HPC Systems", ICML
13 Intel에서 제작한 병렬 연산용 마이크로 프로세서
14 TechPowerUp(2018.7.), "Intel is Giving up on Xeon Phi – Eight More Models Declared End-Of-Life"
15 nVidia Geforce RTX 2080 Ti 모델 기준(최대 클럭수 : 1635 Mhz)

떨어지는 무어의 법칙을 따른다고 볼 수 있다. 물론 최근 들어 하드웨어 발전 속도의 증가가 무어의 법칙을 넘어섰으나, 여전히 큰 흐름에서 성능과 가격의 반비례는 적용된다.

○ 구글, MS, IBM, Apple 등 현재 AI 시장을 리딩하는 클라우드 기업들은 모두 GPU를 활용하여 컴퓨팅 파워를 확보

글로벌 클라우드 기업들이 자사의 클라우드 인프라에 GPU를 활용 사례는 쉽게 찾아볼 수 있다. 글로벌 자산가치 순위 Top5에 드는 기업들은 공교롭게도 모두 클라우드 인프라를 보유하고 있거나 서비스 하고 있는 기업들이며, 이들은 AI 시장도 리딩하고 있다. [표 2-3]은 글로벌 기업가치 순위를 나타낸다. 2위부터 4위까지의 기업들은 자체 클라우드 인프라를 보유하고 있고, 이를 활용한 다양한 서비스를 제공한다. 1위와 5위인 애플, 페이스북의 경우는 조금 다른데, 자체 인프라를 활용한다기보다는 타 회사의 클라우드 인프라를 활용하여 서비스를 제공하고 있다. 그러나 최근에는 클라우드 서비스를 제공하는 업체와 직접 협력하거나, 관련 업체를 인수하며 클라우드 인프라를 확보하는 방향으로 사업영역을 확장하고 있다. 애플은 자체 클라우드 인프라 구축을 위해 인재를 채용하는 등 사업을 추진하고 있으며(2020.2.), 페이스북의 경우 최근 AI 클라우드를 구축을 위해 중국의 알리바바와 협력(2019.9.)하고 유럽의 클라우드 게임업체인 플레이기가를 900억에 인수(2019.12.) 하기도 했다.

[표 2-3] 전 세계 기업가치 순위 Top5

순위	기업	국가	시가총액(십억 달러)	클라우드 인프라/서비스
1	Apple	미국	1,190	iCloud
2	Microsoft	미국	1,155	Azure
3	Google(Alphabet)	미국	905	Google Cloud Platform(GCP)
4	Amazon	미국	901	AWS
5	Facebook	미국	576	Intel, AWS, MS 클라우드 활용

※ 주 : 2019.11.30. 기준
※ 자료 : PWC(2019.7.), "Global Top100 companies by market capitalisation"
iWeblists(2019.11.30.), "US commerce Stock Market Capitalization of the 50 Largest American Companies"
CorporateInformation(2019), Top100 list

Apple의 클라우드 스토리지이자 서비스인 iCloud는 Azure, AWS, GCP를 기반으로 서비스 하며, iCloud는 이들의 최대 고객 중 하나이다. MS의 클라우드 Azure는 N시리즈를 통해서 GPU 컴퓨팅을 제공하고 있다. Google은 알파고 당시 GPU 클러스터를 활용했으며, GPU와 유사한 병렬 구조의 텐서플로우(TensorFlow)[16] 전용 TPU(Tensor Processing Unit)를 개발하여 클라우드에 적용하고

16 2015년 11월 구글에서 공개, 머신러닝 프로그래밍용 공개SW기반 라이브러리

있다. Amazon은 GPU 제조업체 엔비디아와 협력하여 GPU 가속 클라우드 서비스인 AWS EC2 P3를 보유 및 서비스 하고 있다.

이처럼 현재의 클라우드 인프라에는 GPU가 활용되고 있으며, 이에 대한 수요는 AI 및 클라우드 서비스와 함께 커지고 있다. 클라우드는 공급자 측면의 '컴퓨팅 자원의 효율적 활용'과 수요자 측면의 '필요 시 컴퓨팅 자원 대여'라는 특성과 장점을 가짐과 동시에, '대표적인 병렬컴퓨팅 인프라'로써 활용되고 있다.

3 GPU 가상화 기술과 분류

3.1. GPU 가상화의 필요성과 클라우드 적용

○ GPU도 자원의 효율적인 활용을 위해 가상화가 필요

클라우드는 기본적으로 가상화(Virtualization)[17] 기술을 기반으로 구현되는데, GPU 컴퓨팅을 효율적으로 활용하기 위해서도 가상화 기술이 필요하다. GPU 자원을 가상화하는 이유는 '자원의 효율적 활용'이라는 측면에서 가상화 기술의 목적과 맥락에 일치한다. 일반적인 경우, 시스템에서 GPU의 활용률은 평균 15% 수준에 불과[18]한 경우가 종종 있고, 자원 요구가 급증하는 특정시간(Peak Time)에는 컴퓨팅 자원이 부족한 경우가 발생한다.

GPU 자원을 가상화 하게 되면 매번 필요한 만큼만 자원을 최적화하여 활용할 수 있도록 하여 자원의 비효율적인 낭비 또는 부족현상을 최소화할 수 있다. 또한, 유동적으로 조절하는 컴퓨팅 능력은 머신러닝 및 딥러닝의 작업 생산성을 향상시킨다. 가상화를 통해 기존의 가상머신(Virtual Machine, VM)[19]과 호환성도 보장할 수 있다.

○ GPU 가상화는 그동안 어려웠던 일

그동안 GPU의 가상화는 제품 제조업체의 하드웨어(HW) 및 소프트웨어(SW) 지원이 없어서 매우 어려운 문제였다. 가상화는 소프트웨어적으로 구현되고 동작하는 경우가 대부분이지만, 하드웨어차원의 호환이 없다면 구현이 어려운 이슈이다. 하드웨어를 생산하는 업체가 해당 하드웨어의 구조를 가장

17 가상화 기술은 기존 하드웨어 장비를 마치 여러 개의 장치가 동작하는 것처럼 논리적으로 나누거나, 여러 개의 하드웨어 장비를 묶어 하나의 장치인 것처럼 동작하게 하며, 하드웨어 장치가 가진 고정된 능력을 보다 유연하게 활용하여 자원효율성을 높이는 기술(상세내용 – 소프트웨어정책연구소(2018.12.), "클라우드 가상화 기술의 변화" 참조)

18 TowardsDataScience(2019.3.), "Monitor and Improve GPU Usage for Training Deep Learning Models"

19 HW의 가상화를 통해 컴퓨팅 환경을 에뮬레이션 하는 것으로 실제 HW와 유사하게 동작하며, OS나 응용프로그램을 설치 및 실행시킬 수 있음. 경우에 따라 다수의 VM이 하나의 장치 상에 공존할 수 있음(상세내용 – 소프트웨어정책연구소(2018.12.), "클라우드 가상화 기술의 변화" 참조)

잘 알고 있고, 하드웨어가 잘 동작하게끔 펌웨어(firmware)[20]를 탑재하거나, 시스템 내에서 유연하게 동작할 수 있도록 드라이버(Driver)[21] 소프트웨어를 배포하기 때문이다. 즉, 제품에서 지원하지 않는 가상화 기능을 강제로 구현하기 위해서는 제품의 하드웨어적인 동작 요소와 구조를 모두 알아야 하며, 드라이버나 펌웨어를 수정(대부분의 경우에는 수정이 불가능한)해야만 한다.

현재는 클라우드, AI의 수요가 증가하고 있고, 이들과 연계된 빅데이터(Bigdata), 사물인터넷(Internet of Things, IoT) 등의 확산 또한 가속화되고 있어서 GPU 제조 업체들은 클라우드 컴퓨팅에 잘 호환될 수 있도록 가상화를 지원하기 시작했다.

3.2. GPU 가상화 기술의 종류

GPU 가상화 기술은 가상화를 어디에 구현하느냐에 따라 크게 세 가지 기술로 분류할 수 있다. 이 기술들은 각각 API 리모팅, 반가상화 및 전가상화, 하드웨어 지원 가상화에 해당된다.

1 API 리모팅 : GPU 벤더가 가상화를 지원하지 않을 때 가상화를 구현하는 방법

클라우드 환경에서 GPU를 가상화하는 작업은 네트워크 카드 또는 디스크와 같은 I/O[22] 장치를 가상화하는 것보다 복잡하고 어려운 작업이다. 그 이유는 다음과 같다. 첫째, NVIDIA와 같은 GPU 공급 업체들은 상업적인 이유로 GPU 드라이버의 세부 구현 사항과 소스 코드를 공개하지 않고 있다. GPU를 가상화하는데 GPU 드라이버의 소스코드는 필수적인 정보이므로 이는 큰 제약사항이 된다. 둘째, 리버스 엔지니어링(Reverse Engineering)[23]과 같은 방법으로 GPU 디바이스 드라이버를 일일이 분석하여 수작업으로 설계를 만든다고 해도, GPU 공급업체들은 새로운 GPU 드라이버를 주기적으로 배포하고 있으므로 이를 매번 따라가기에는 시간과 노력이 많이 든다. 따라서 리버스 엔지니어링으로 만들어진 디바이스 드라이버는 시간이 흐름에 따라 구형의 GPU만 지원하게 된다. 마지막으로, 일부 운영체제나 가상화 솔루션 배포 업체들은 가상화를 지원하는 독점적인 GPU 드라이버를 배포하기도 하지만, 이는 모든 운영체제에서 사용할 수 없는 드라이버가 된다. 결국, GPU에 접근하기 위한 표준적인 인터페이스가 없기 때문에, GPU 디바이스 드라이버의 제작이 어려고, GPU를 가상화하는 것이 어려워진다. API[24] 리모팅 방식은 현재까지 GPU 가상화에 가장 많이 쓰이는 보편적인 방법이었으며, 앞에서 언급한 제한 사항을 극복할 수 있는 방법으로 등장하였다. [그림 3-1]은 가상화 환경에서 API 리모팅 방식이 동작하는 원리를 도식화한 것이다.

---

20 HW에 포함된 SW로 HW제조사에서 제공하며, 해당 HW기기가 주어진 기능을 수행할 수 있도록 함(예 : 리모콘, 전기밥솥, 냉장고 등). 기본적으로 응용프로그램으로써 수정은 불가능하고, 업데이트가 가능한 경우는 있음.

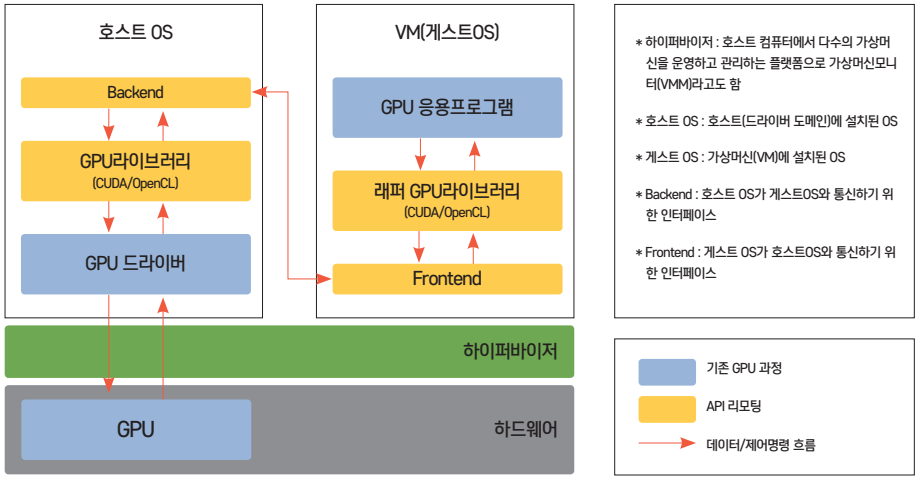
21 특정 HW를 제어하기 위해 시스템의 운영체제(OS) 내에 설치/동작하는 프로그램

22 Input/Output, 입출력 장치

23 역공학, 구조분석을 통해 시스템의 원리를 발견하는 과정

24 API(Application Programming Interface)는 운영체제(OS)나 프로그래밍 언어에서 제공하는 기능을 응용프로그램이 제어할 수 있도록 만든 인터페이스

[그림 3-1] 가상화 환경에서 API 리모팅 기법



API 리모팅의 기본 원칙은 게스트 운영체제(Guest OS)에 네이티브(Native) 환경(호스트 OS)에서 쓰이는 GPU라이브러리(Library)[25]와 동일한 API를 가진 래퍼(Wrapper) 라이브러리를 제공하는 것이다. 래퍼 라이브러리는 GPU 응용프로그램이 GPU에 대한 연산 함수를 호출할 때 그 호출이 직접 GPU 드라이버에 도달하기 전에 가로채는 역할을 맡게 된다. 이렇게 가로채어진 호출은 공유 메모리를 통해서 같은 시스템 내의 호스트 운영체제로 전달되거나 또는 네트워크를 통해 GPU를 갖고 있는 원격의 시스템에 전달된다. 호스트 운영체제나 다른 원격 시스템으로 전달된 GPU 호출은 원격으로 처리되므로 API 리모팅이라는 이름이 붙여졌다. 원격에서 GPU 호출이 처리된 후 호출의 결과 값만 다시 래퍼 라이브러리로 전달되게 된다. 즉 결론적으로 API 리모팅은 게스트 운영체제에 GPU를 직접 노출시키지 않고서도 게스트 운영체제에 GPU가 있는 것처럼 에뮬레이션 시킬 수 있는 효율적인 방법을 제공한다. 이러한 접근 방식은 라이브러리 수준에서 GPU를 가상화함으로써 소스 코드가 공개되지 않는 상용 GPU 드라이버를 우회할 수 있다.

25 컴퓨터 프로그램의 조직화된 집합으로 데이터, 문서, 도움말 자료, 메시지 틀, 미리 작성된 코드, 서브루틴(함수), 클래스, 값, 자료 형 사양 등 이들을 묶어서 실행에 활용할 수 있도록 하는 기능

API 리모팅 방식의 GPU 가상화

• API 리모팅 방식을 사용하는 기법들은 GViM, vCUDA, rCUDA, GVirtuS, GVM, Pegasus, Shadowfax, VOCL, DS-CUDA 등이 있음([표 3-1])

[표 3-1] API 리모팅 방식의 GPU 가상화 기법

기술명	특징
GViM	• Xen[26] 하이퍼바이저를 기반으로, NVIDIA社의 GPU 라이브러리인 CUDA를 통해 가상화를 구현하여 가상머신상에 CUDA 응용프로그램(App) 및 API를 탑재 • GPU App이 데이터를 대량으로 전송할 수 있는 공유메모리(Xenstore) 활용
vCUDA	• Xen 하이퍼바이저 기반으로 게스트OS에 가상 GPU(vGPU)와 CUDA 래퍼(Wrapper) 라이브러리를 제공하는 형태로 구현
rCUDA	• 클라이언트-서버 구조를 기반으로 하여 원격(Remote)으로 호스트 서버 컴퓨터의 GPU를 게스트OS에 가상화 시켜주는 기법으로 자체 통신프로토콜 및 100G 대역폭 활용 • 기존 같은 물리머신 상의 호스트의 GPU를 에뮬레이션 함으로 발생하는 오버헤드 및 성능저하를 개선하기 위해 개발
GVirtuS	• Xen외에도 KVM, VMware 등과 같은 다양한 하이퍼바이저를 지원하는 GPU 가상화 • 게스트-호스트 간의 통신 병목현상을 해결하기 위해 각 하이퍼바이저가 제공하는 고속 통신 채널을 활용할 수 있는 통신모듈(Communicator)을 탑재
GVM	• GPU App의 성능을 예측하기 위한 가상화 지원 모델로, 사용자 API, GPU 가상화 관리자 및 가상 공유메모리로 구성된 자체 가상화 인프라를 제공
Pegasus	• 기존 GViM의 업그레이드 버전으로, 가상화된 GPU(vGPU)를 스케줄링 하며 공유가 가능한 단위로 관리
Shadowfax	• Pegasus의 업그레이드 버전으로, 연산요구량이 많아질 때 호스트 상의 GPU 뿐만 아니라 원격의 GPU를 포함하여 다양한 가상 플랫폼을 구성할 수 있도록 함
VOCL	• 공개SW기반의 병렬 컴퓨팅 프레임워크인 OpenCL을 지원하는 GPU 가상화 기술로, rCUDA와 유사한 형태로 원격 GPU 가상화를 지원
DS-CUDA	• rCUDA와 유사하게 원격 GPU 가상화를 지원하며, 두 개의 서로 다른 호스트 GPU가 동일한 연산을 수행하여 연산의 결과를 비교 검증하는 기능을 제공

※ 자료 : 참고문헌 목록 참조(국외 참고자료 9~17)

다만 이 방식은 기존의 가상화 환경에서 게스트 OS의 프로그램들을 수행하는 데에 있어서 성능 저하를 유발할 수 있다. 게스트 OS의 응용프로그램은 하드웨어를 활용할 때, 반드시 호스트 OS를 거쳐서 접근 및 실행할 수 있도록 되어 있다. 이는 자원을 공유하는 과정에서 호스트 OS가 관리자(Manager) 역할을 수행하기 때문이다. 따라서 [그림 3-1]과 같은 Backend-Frontend 통신이 수반된다. 이는 가상화가 없는 상태의 GPU를 활용하는 것에 비해 단계가 추가됨을 의미하여, 곧 성능 저하로 이어질 수 있다. 이를 해결하기 위한 방법으로 하드웨어 제조사들은 가상환경에서 게스트 OS가 호스트 OS를 거치지 않고 직접 하드웨어와 통신할 수 있는 다양한 방법(Intel-VT 기술 등 - 후술)들을 제시하고 있다.

26 여러 게스트 운영체제를 하나의 호스트 컴퓨터에서 동작하게끔 하는 대표적인 하이퍼바이저, 케임브리지 대학교에서 개발이 시작되어 2003년 첫 공개버전 발표. 초기에는 게스트 운영체제를 수정해야 하는 반가상화만 지원하였으나 현재는 전가상화도 지원

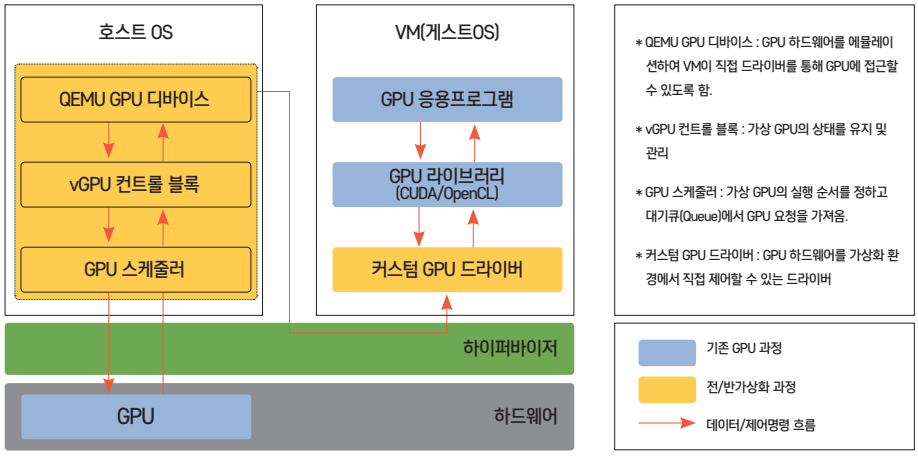
API 리모팅 방식은 새로운 기능이 GPU 공급 업체에 의해 원래 라이브러리에 추가될 때 래퍼 라이브러리 또한 업데이트 되어야 하며, 이는 상당히 번거로운 작업이다. 이로 인해서, API 리모팅 솔루션은 최신의 GPU 라이브러리 버전을 사용하지 못하는 경우가 많다.

2 전가상화 및 반가상화 기법을 통한 커스텀 GPU 드라이버 가상화

API 리모팅 방식의 최신성 유지 측면의 한계를 개선하기 위해서 GPU 드라이버를 가상화하는 방식이 등장했다. 가상화의 방식에는 전가상화(Full-Virtualization)와 반가상화(Para-Virtualization)가 있는데, 전가상화는 GPU 드라이버에 대한 수정 작업 없이 통째로 에뮬레이션 하는 방법이고, 반가상화는 GPU 드라이버를 수정하는 방식으로 VM의 직접 제어가 어느 정도 가능하게 하여, 성능을 좀 더 올릴 수 있는 방식이다.

최근에 CPU 및 GPU 칩셋 제조사중 하나인 AMD[27]는 특정 모델에 한해 GPU 아키텍처에 관한 공개 문서를 발표했다. 또한 일부 개발자는 리버스 엔지니어링 기술을 사용하여 NVIDIA GPU와 호스트 간의 인터페이스를 추측했다. 이러한 노력 덕분에 AMD 또는 NVIDIA GPU를 제어 할 수 있는 커스텀(Custom) GPU 드라이버가 개발되었다. 이러한 커스텀 드라이버의 출현은 반가상화 기술과 전가상화 기술에 대한 많은 연구를 촉진시켰다. [그림 3-2]는 커스텀 GPU를 활용한 가상화 방법을 나타낸다.

[그림 3-2] 커스텀 GPU 드라이버 가상화



27 글로벌 시장에서 CPU 부문 Intel(Intel, 1위) 다음, GPU 부문 NVIDIA(nVidia, 1위) 다음으로 시장을 점유

이 아키텍처에서 게스트 운영체제는 수정되지 않은 또는 수정 된 커스텀 GPU 드라이버를 활용하고 있다. 우선, 커스텀 GPU 드라이버는 기존의 상용화된 드라이버가 아닌 가상화를 위한 튜닝(Tuning)을 거친 드라이버이다. 이후, 별다른 수정 없이 GPU 드라이버를 완전하게 에뮬레이션 하거나(전가상화), 하이퍼바이저를 통한 VM에 의해 제어될 수 있도록 수정(반가상화)을 하여 가상화를 구현한다. 이 방식은 VM이 직접 GPU 드라이버를 제어하는 방식이다.

게스트 운영체제의 커스텀 드라이버에서 나오는 GPU 접근 및 실행 요청은 하이퍼바이저에서 관리하는 공유 메모리를 통해 호스트 운영체제의 QEMU[28] GPU 장치로 전달된다. QEMU 장치는 장착된 GPU 하드웨어를 에뮬레이션하고 GPU가 장착된 PCIe 주소값(BAR[29])을 게스트 드라이버에 노출한다. 이렇게 하면 게스트 드라이버는 QEMU 장치를 실제 GPU로 간주하게 된다. 가상 GPU(vGPU) 제어 블록은 각 가상 GPU의 상태를 유지하고 QEMU 장치에서 발행 된 GPU 요청을 대기 큐에 보관한다. GPU 스케줄러는 다음에 수행할 가상 GPU를 선택하여 실행하고 관련 대기 큐에서 VM으로부터 온 GPU 요청을 가져온다. 게스트 운영체제의 GPU 요청은 GPU에서 직접 처리되고 결과는 역 경로를 통해 응용 프로그램에 반환된다.

이 구조의 장점은 기존 GPU 라이브러리를 재사용할 수 있으며 GPU 가상화의 구현 지점이 게스트 운영체제의 드라이버 계층으로 내려가므로 향후 라이브러리의 변경에 대비할 수 있다. 또한 게스트 운영체제의 GPU 요청을 하이퍼바이저가 모니터링하고 조정할 수 있으므로 API 리모팅과 비교하여 실시간 작업이동(Live Migration)[30]과 같은 필수 가상화 기능을 어려움 없이 지원할 수 있다. 반면, 이 구조의 단점은 커스텀 GPU 드라이버에 크게 의존한다는 것이다. 새로운 GPU 마이크로 아키텍처가 출시되었을 때 리버스 엔지니어링 또는 공개 문서가 아직 없다면 해당 커스텀 드라이버의 개발에 상당한 부담을 줄 수 있다.

28 Quick EMUlator의 약자로, 하드웨어 가상화 기능을 갖춘 오픈소스 기반의 에뮬레이터
29 Base Address Register
30 가상화 환경에서 가상머신의 실행환경을 유지시키는 기능. 호스트 시스템에 하드웨어 및 소프트웨어 상의 업데이트와 같은 이벤트가 발생하더라도 가상머신의 복제, 이전 등을 통해 실행환경을 유지

전/반가상화 방식의 GPU 가상화

- 전가상화 기법으로는 GPUvm, gVirt, gHyvi, gScale 등이 있고, 반가상화 기법으로는 VMware SVGA II, LoGV, VGRIS, AMD HSA 솔루션 등이 있음([표 3-2])

[표 3-2] 전/반가상화 방식의 GUP 가상화 기법

기술명		특징
전가상화	GPUvm	• 리눅스 기반의 공개OS인 우분투(Ubuntu)의 그래픽드라이버 nouveau를 활용하고, Xen 하이퍼바이저 기반의 GPU 가상화를 제공하며 전가상화와 반가상화를 모두 지원
	gVirt	• Xen 하이퍼바이저 환경에서 Intel CPU 내장형 GPU에 대한 전가상화를 지원하며, GPGPU 보다는 그래픽 가속 자체에 중점을 둠
	gHyvi	• gVirt의 성능을 개선하기 위한 후속 기술로, 가상머신에서 구동되는 GPU App의 빈번한 메모리 접근 및 업데이트 과정을 개선
	gScale	• gVirt의 확장성을 개선하기 위한 후속 기술로, 기존 최대 4개까지의 vGPU를 제공하던 것을 12~15개 수준으로 확장한 버전
반가상화	VMware SVGA II	• 시스템 가상화 업체 VMware社의 GPU 가상화 제공 기술로, AMD社의 GPU 및 공개 문서를 기반으로 구축한 게스트OS용 GPU 드라이버 • 하이퍼바이저가 가상머신과 실제 GPU 하드웨어간의 접근에 보다 적극적으로 개입하고 기본적인 가상화 기술 중 하나인 실시간 마이그레이션(Live Migration) 기능을 지원하며 API 리모팅 방식이 갖는 한계점을 개선
	LoGV	• NVIDIA社의 GPU 드라이버를 리버스 엔지니어링을 통해 오픈소스로 구현한 드라이버인 PathScale를 통해, KVM[31] 하이퍼바이저 기반의 GPU 가상화를 지원
	VGRIS	• VMware SVGA II에서 구현한 반가상화 기술을 기반으로 하며, 각 가상머신의 성능을 모니터링하고 GPU 스케줄러에 전송하는 관리 에이전트를 제공
	AMD HSA 솔루션	• AMD社에서 2016년에 개발한 가상화 기술로 KVM 하이퍼바이저를 기반으로 하며, GPU와 CPU를 동시에 활용하되 이들간의 통신 오버헤드를 완화시키는 기능을 제공

※ 자료 : 참고문헌 목록 참조(국외 참고자료 18~21, 22~25)

3 하드웨어 지원 GPU 가상화(하드웨어 벤더가 가상화를 지원)

하드웨어 지원 가상화는 GPU 또는 메인보드를 제작하는 벤더들이 제품상에서 가상화 기능을 제공하는 것이다. 기존 벤더들이 가상화 기능을 제공하지 않음으로 인해 API 리모팅 기법으로 우회하거나, 리버스 엔지니어링을 통해 수정된 GPU 드라이버를 제어하는 방식 보다 가상화가 쉽고 빠르다. 이 방식에서 가상머신은 GPU 하드웨어에 직접 액세스 할 수 있다.

이 접근법은 Intel, AMD, NVIDIA 등 CPU 및 메인보드 칩 제조업체 또는 GPU 공급 업체가 제공하는 Intel VT-d, AMD-Vi, NVIDIA-GRID 등과 같은 I/O 가상화를 위한 하드웨어 확장 기능을 활용한다. 하드웨어 지원 GPU 가상화 기법은 하드웨어 장치로 전송되는 데이터 또는 인터럽트

31 Kernel-Based Virtual Machine(커널기반가상머신), 리눅스 커널(운영체제 내에서 시스템을 통제하는 핵심 프로그램 – 각주40 참조)을 하이퍼바이저화 한 가상화 지원 플랫폼으로, QEMU라는 가상머신 실행 프로그램을 기반으로 구동되며 전가상화와 반가상화 모두 지원

(Interrupt)[32] 등과 같은 제어명령들을 처리하거나 담아두기 위한 장치 내 메모리 공간을 게스트 운영체제에 직접 매핑(Mapping)[33]해주는 방식으로 GPU를 가상화한다. 즉, 가상머신이 직접 GPU 하드웨어에 접근하여 읽기/쓰기/실행 등의 명령을 수행하도록 하는 방식이다.

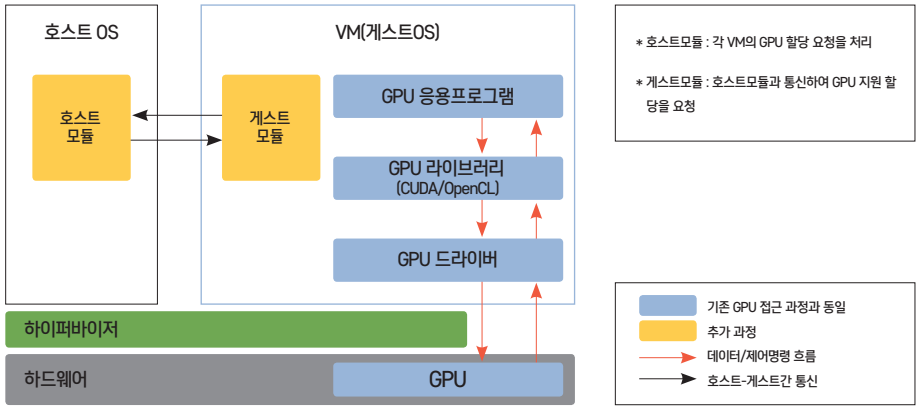
스케줄링(Scheduling)과 가상화

- 기본적으로 컴퓨터에서 응용프로그램이 하드웨어 장치에 접근하는 것은 엄격한 시스템 감시와 관리 하에 이루어짐
  - 이는 다양한 응용 프로그램들이 동시에 한 장치를 활용하고자 할 때 발생할 수 있는 충돌을 방지하고자 하는 것임
  - 자원의 효율적인 활용을 위해 꼭 필요한 기능
- 때문에 가상화를 통해 공유되는 자원(GPU 등)은 반드시 한 순간에 하나의 인가된 VM에서 접근하도록 하는 방식을 취함
  - 결국, 드라이버 도메인(호스트 OS)에서 이러한 관리역할을 담당
  - 이러한 구조적인 한계는 결국 성능저하, 호환성 결여 등과 같은 문제를 야기

① 단일 가상머신 지원형

하드웨어 지원방식의 가상화는 하이퍼바이저의 개입 없이 GPU 장치에서 가상 머신의 메모리 공간으로 데이터를 직접(Direct) 전송할 수 있으며, 이 과정에서 발생하는 인터럽트 또한 게스트 운영체제에 직접 전달된다. [그림 3-3]은 하드웨어 지원 GPU 가상화 시스템의 예를 나타내며, Intel VT-d 또는 AMD-Vi로 구현된 하드웨어 지원 GPU 가상화를 채택한 시스템과 같다.

[그림 3-3] 하드웨어 지원 GPU 가상화



32 실행 중인 명령을 잠시 중단하고 다른 명령을 먼저 실행하는 것
33 입력 데이터를 원하는 목적지에 배치하는 것으로, 여기서는 할당(Allocation)의 의미

이 아키텍처에서 가상머신의 GPU 접근/처리과정은 라이브러리나 드라이버를 수정하지 않고도 GPU를 활용할 수 있다. 이렇게 하면 게스트 운영체제는 DMA[34]와 인터럽트가 가상머신에 직접 매핑되기 때문에 GPU와 통신하는 동안 하이퍼바이저를 우회하게 되며, 하이퍼바이저와의 통신 오버헤드가 사라지면서 성능을 올릴 수 있다.

하드웨어 지원 GPU 가상화를 활용하는 대표적인 최초 사례는 2010년에 발표된 아마존(Amazon)의 EC2(Elastic Compute Cloud)가 있다. EC2는 Intel VT-d 기술을 사용하여 클라우드 사용자에게 GPU 가상화를 통한 인프라 환경을 제공한다. EC2는 가상머신 당 2개의 NVIDIA Tesla GPU를 제공하는 CGI(Cluster GPU Instance)라는 기술을 도입했다. CGI는 각 게스트 운영체제에 GPU를 직접 연결하여 대규모 병렬 처리 능력이 필요한 하이퍼포먼스 컴퓨팅(HPC) 응용 프로그램을 지원할 수 있다. 이후, 다양한 하드웨어 지원형 GPU 가상화 기법들이 연구 및 발표되었는데, [표 3-3]과 같다.

하드웨어 지원 방식의 GPU 가상화(단일머신 지원형)

[표 3-3] 하드웨어 지원 방식의 GUP 가상화 연구들

연도	특징
2013	• Amazon EC2에서 32개의 클러스터로 구성된 CGI 성능 평가 연구 – 가상화 환경과 자체 클러스터를 사용하는 기본환경에서 성능 비교 – 계산 집약적인 프로그램이 EC2와 같은 설정에서 더 유리함을 증명
2013	• 하드웨어 지원형 GPU 가상화 시스템에서 클라우드 게임 성능 평가 – 일부 프로그램에서 GPU 가상화를 활용하게 되면 성능이 저하되는 원인을 분석한 결과 빈번한 하이퍼바이저와 가상머신간의 작업 전환 요청에 의한 것임을 입증
2014	• Xen과 KVM환경에서 Intel VT-d 기술 기반의 GPU 패스스루 시스템 구현 – 위 두 하이퍼바이저에서 VT-d 기술을 사용했을 때 GPU 성능이 네이티브 수준에 도달하는 것을 증명
2014	• GPU 하드웨어 가상화 성능 측정 연구 – Xen 하이퍼바이저 환경에서 Intel VT-d 기술을 활용하면, 가상화 환경에서 성능저하가 API 리모팅 기법(40%) 대비 성능 저하가 1.2%수준에 불과함을 실험
2014	• 다양한 하이퍼바이저를 기반으로 하는 Intel VT-d 기술의 실험 – VMWare ESXi, KVM, Xen 및 Linux 컨테이너(LXC)의 성능을 실험한 결과 네이티브에 근접하며 가장 일관된 성능은 KVM, 평균 수준은 Xen, 최고 성능은 LXC 임을 밝힘
2015	• SR-IOV[35]를 활용한 고성능 컴퓨팅 부하 실험 – GPU 가상화 환경에 대한 부하실험 결과 네이티브 환경 대비 1.5% 수준의 오버헤드만 보임

※ 자료 : 참고문헌 목록 참조(국외 참고자료 26~31)

34 Direct Memory Access(직접 메모리 접근), 입출력 장치가 CPU에 의한 제어 없이 직접 자료의 이동 및 교신을 할 수 있도록 하는 방식

35 Single Root – I/O Virtualization, Intel에서 제공하는 네트워크 인터페이스 카드(NIC)의 하드웨어 지원형 가상화 기술로, 가상화 환경에서도 거의 네이티브 수준(10Gb)의 초고속 네트워크 대역폭을 제공

하드웨어 지원 가상화 방식의 장점은 하이퍼바이저와 같은 추가 소프트웨어 계층 없이 GPU 가상화를 실현하면서 거의 네이티브 환경에 근접하는 성능을 달성할 수 있다는 것이다. 반면, 단점으로는 GPU 작업이 호스트 운영체제 또는 하이퍼바이저를 우회하기 때문에 시스템상의 전체적인 GPU 스케줄링 정책을 적용하는 것이 불가능할 수 있다는 것이다. 아울러, API 리모팅 방식과 마찬가지로 실행 체크 포인트, 실시간 마이그레이션 및 내결함성 실행 기능을 구현하기가 어렵다. 이 부분의 구현은 전적으로 하드웨어 제공 업체의 지원 여부에 달려있다.

기존의 Intel VT-d 및 AMD-Vi는 이러한 방식으로 가상화 시스템 환경의 성능을 높일 수 있었다. 그러나, Intel VT-d 및 AMD-Vi에서는 게스트 운영체제가 부팅 될 때 지정된 가상머신에 GPU가 정적(Static)으로 할당된다. 즉, 이것은 특정 시간에 하나의 가상머신에게만 GPU 가상화를 지원한다는 의미이다. 이렇게 되면 여러 게스트 운영체제 간에 단일 GPU 공유를 지원할 수 없는 한계점이 생긴다.

② 다중 가상머신 지원형

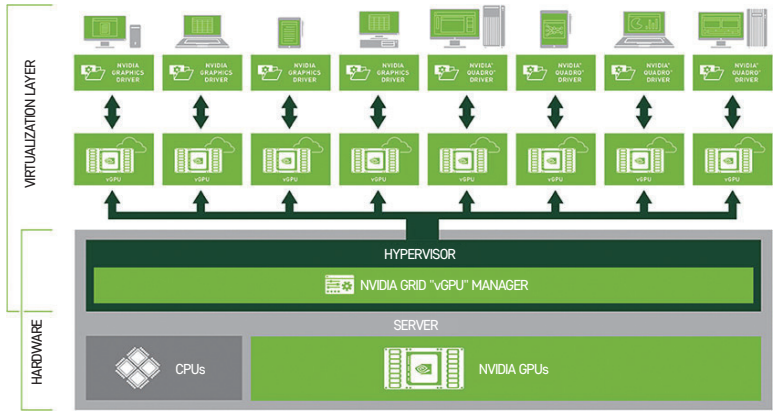
최근에는 가상머신간의 GPU 자원을 동적으로 활용할 수 있도록 핫 플러그(Hot Plug)[36] 기능을 탑재한 솔루션이 출시되었다. 2013년에 여러 가상 머신들이 단일 GPU를 공유할 수 있는 기법 연구가 시도되었다.[37] GPU를 가상머신에 직접 연결·할당하는 기능인 패스스루(Pass Through)를 확장하는 방법이다. 기존의 패스스루 방식은 오직 하나의 가상머신에만 하나의 GPU 하드웨어를 직접 할당하였기 때문에, 해당 가상머신이 종료되기 전까지 GPU를 다른 가상머신에 할당 할 수 없었다. 이 연구에서는 GPU 장치를 동적으로 설치하거나 제거할 수 있는 PCIe 채널의 핫 플러그 기능을 활용하여 GPU를 여러 가상머신이 공유할 수 있게 구현하였다. 이 구현이 가능하기 위해서는 GPU를 구동하기 위한 래퍼 라이브러리를 각 가상머신에 제공해야하며, 이들을 모니터링 해야만 한다.

현재 시점에서 다중 가상머신에 GPU 자원을 공유하고 할당하는 최신 기술은 NVIDIA의 GRID 솔루션이 있다. [그림 3-4]는 NVIDIA GRID 아키텍처를 나타낸다. GRID 솔루션은 여러 가상머신 간의 GPU 공유를 지원하면서 좀 더 나은 확장성과 성능을 제공한다. 이 솔루션은 클라우드 환경을 타겟으로 개발된 최신의 NVIDIA GPU에서 지원 된다. [그림 3-4]의 NVIDIA GRID vGPU MANAGER는 가상 GPU의 할당과 해제, 제어를 위한 하이퍼바이저 역할을 수행한다. 이 하이퍼바이저는 각 게스트 운영체제가 GPU를 활용하기 위해 갖는 가상 메모리 영역을 GPU의 물리적 메모리에 개별적으로 매핑하고 변환할 수 있도록, 새로운 I/O 메모리 관리 장치(IOMMU)를 가지고 있다.

36 전원이 유지된 채 시스템의 구동 상태에서 하드웨어 장치를 활용/교체 할 수 있도록 하는 기능으로 윈도우의 Plug and Play 기능과 유사

37 Jo, Heeseung, et al(2013), “Exploiting GPUs in virtual machine for BioCloud.” BioMed research international 2013

[그림 3-4] NVIDIA GRID 아키텍처 개념도



※ 자료 : nvidia.com

또한 각 게스트 운영체제가 제어 요청에 대해 별도의 입력 버퍼를 가질 수 있도록 했다. 이 두 가지 아키텍처 변경을 통해 GRID를 지원하는 NVIDIA GPU는 하드웨어 차원에서 GPU를 공유하는 것을 실현하였으며, 동시에 각 게스트 운영체제의 성능을 네이티브 환경처럼 유지할 수 있다. 현재 NVIDIA GRID는 클라우드 환경을 목표로 하는 NVIDIA GRID K1, K2, NVIDIA Tesla M6 및 M60 GPU 등에서 지원이 된다.

4 향후 기술방향과 시사점

4.1. 클라우드 GPU 가상화 기술의 향후방향

클라우드 컴퓨팅을 위한 GPU 가상화 기술들은 비교적 최근에 등장한 기술들이다. 지금껏 어려웠던 GPU를 가상화하는 방법에 대해 많은 부분이 해결되기도 하였다. 그러나 아직 GPU 가상화 환경의 성능과 기능 측면에서 다양한 해결 과제가 남아 있다. 향후 더 다양해질 AI 서비스와 고객 요구를 충족시키기 위해 GPU 가상화 기술은 다음과 같은 발전방향이 고려되어야 한다.

○ 컨테이너 기반 경량 가상화의 실현이 필요

대표적인 경량화 사례인 Linux 기반 컨테이너[38]는 프로세스 수준의 가벼운 가상화를 제공하는

38 Container, '물체를 격리하는 공간'이라는 사전적 의미처럼, 컴퓨팅 환경에서 모듈화되고 격리된 컴퓨팅 공간 또는 컴퓨팅 환경을 의미, 프로세스를 격리하여 모듈화된 프로그램 패키지로써 수행하는 것을 의미함(상세내용 – 소프트웨어정책연구소(2018.12.), "클라우드 가상화 기술의 변화" 참조)

새로운 클라우드 기술이다. 컨테이너 기술에서는 하나의 리눅스 커널에 의해 여러 개의 컨테이너가 생성 및 관리되기 때문에 컨테이너는 GPU를 가상화하기 위해 앞서 살펴본 추가적인 래퍼 라이브러리나 프론트/백엔드 드라이버 모듈을 필요로 하지 않는다. 따라서 컨테이너를 사용하면 네이티브 환경에 가까운 성능을 얻을 수 있다.

하지만 컨테이너를 사용한 GPU 가상화에 대한 현재의 연구는 아직 초기 단계에 있다. 지금까지의 연구는 컨테이너와 다른 가상화 솔루션 간의 성능을 비교하는 것이 주를 이루고 있다. 클라우드 컴퓨팅에서 GPU가 장착된 컨테이너를 활용하려면 공정하고 효과적인 GPU 스케줄링 알고리즘이 필요하다. 대부분의 GPU 스케줄러는 GPU 호출을 가로 채기 위해 API 래퍼 라이브러리 또는 드라이버 변경이 필요하지만, 이는 컨테이너 환경에서 무시할 수 없는 오버 헤드를 일으킬 수 있다.

○ 가상화를 통한 시스템의 확장성을 보장

시스템 가상화의 주목적은 컴퓨팅 자원의 사용률을 높이고 소유 또는 운영 비용을 줄이는 것이다. 이 목표는 데이터 센터의 각 물리적 시스템에 다수의 가상머신을 통합하여 클라우드화 함으로써 달성할 수 있다. Xen 하이퍼바이저는 호스트 당 500개가 넘는 가상 CPU를 생성할 수 있도록 설계되었다. 그러나 GPU 가상화의 경우에는 가상머신에게 상대적으로 적은 수의 가상 GPU를 제공한다.

앞서 살펴본 gScale 프레임워크가 확장성을 높이는 데 상당한 노력을 기울인 기술이지만, 이 프레임워크는 Linux 가상머신용으로 최대 15개의 vGPU만 제공할 수 있다. CPU와 GPU의 리소스(=자원) 통합 성능 간에는 아직도 상당한 불균형이 존재하며 그 격차를 메우는 것은 여전히 어려운 과제이다. 따라서 방대한 vGPU를 생성할 경우 GPU 장치 메모리 크기, GPU 컨텍스트 전환(Context Switching)[39] 빈도 및 캐시(Cache)[40] 등 관련하여 성능에 영향을 미치는 모든 사항을 고려해야한다.

○ 가상화 환경에 필요한 보안책 마련

하이퍼 바이저의 중요한 기능 중 하나는 가상머신 간의 안전한 격리(Isolation)를 제공하는 것이다. 이 작업을 수행하기 위해 앞서 살펴보았던 LoGV, GPUvm, gVirt 및 gScale 등은 한 가상머신이 다른 가상머신의 GPU 주소 공간을 매핑하지 못하게 한다. 이러한 보호 메커니즘에도 불구하고 GPU 가상화 프레임 워크는 DoS(Denial of Service)[41]와 같은 공격에 취약한 상태이며 악의적인 가상머신이 높은 빈도로 GPU 명령을 백엔드(호스트 운영체제의 Backend)에 연속적으로 보내게 되면 전체 시스템을 위험에 빠뜨릴 수 있다.

39 문맥전환이라고도 하며, CPU(또는 GPU)가 다른 작업을 수행하기 위해 현재 수행중인 작업을 보관하고 새로운 작업으로 전환하는 과정

40 컴퓨터에서 작업을 위한 데이터나 값을 미리 복사해 두는 임시 저장소로, 보통 CPU 칩 내부나 바로 옆에 탑재하여 하드디스크나 메인메모리의 데이터에 접근하는 시간이 오래 걸리는 경우 활용

41 서비스 거부 공격, 악의적으로 시스템의 자원을 부족하게끔 하여 제 기능을 못하도록 만드는 공격

이 문제를 해결하기 gVirt는 GPU가 중단될 경우 GPU를 재설정하고 각 가상머신의 실행 상태를 검사한 후 의심스러운 가상머신을 제거한다. 하지만 이로 인해 서비스가 일시 중지되어 정상적인 가상머신에 피해가 갈 수 있다. GPU 재설정을 피하려면 악의적인 가상머신이 시스템을 위협하기 전에 악성 가상머신의 실행속도를 지연시킬 수 있는 세분화된 접근 제어 기법이 필요하다.

○ 전력 소비 효율성 향상

가상화하지 않은 GPU 플랫폼에서 에너지 효율에 대한 연구는 많이 이루어졌지만, 가상화 환경에서의 연구는 아직 별로 없다. GPU 가상화 환경에서 에너지 효율성을 높이려는 연구는 pVOCL 정도가 있는데, 앞서 살펴본 API 리모팅 기법 중 VOCL의 전력효율화 버전이다. GPU 간 최대 전력 소비를 원격에서 작업량에 따라 유동적으로 제어함으로써 GPU 클러스터의 전체적인 에너지 효율을 향상시킨다. 이 방식 외에도 호스트 측에서 전력 효율에 대한 연구도 필요하다. 가상머신 간의 다양한 GPU 사용 패턴을 모니터링하고 작업 부하에 따라 GPU의 전원 상태를 동적으로 조정하는 시스템이 향후 연구개발 기술로 적합하다.

○ 자원의 공유 및 활용성 증대

최신의 GPU들은 프로세스 또는 가상머신이 단일 GPU 하드웨어에서 여러 GPU 커널(Kernel)[42]을 동시에 실행시킬 수 있도록 지원한다. 이러한 GPU 하드웨어에 대한 공간 다중화(Spatial Multiplexing, SM)[43] 기법은 각각의 여러 커널이 GPU 내의 분할된 공간을 완전히 활용함으로써, GPU 활용도를 향상시킬 수 있다. 그러나 대부분의 GPU 스케줄링 방법은 여러 가상머신의 GPU 커널이 GPU에서 순차적으로 실행되는 시간 다중화(Time Multiplexing) 기법을 활용한다.

이 방법은 철저한 스케줄링 기술이 필요하며, 많은 작업을 동시에 진행해야 하는 워크로드의 경우 낮은 사용률을 보일 수 있다. 또한 많은 수의 병렬 코어들을 활용하는 데에 있어서 공간 다중화가 시간 다중화보다 유리한 경우가 발생할 수도 있다. 때문에 GPU 스케줄링에서 높은 GPU 사용률과 공정성을 모두 달성하려면 두 가지 방법을 함께 사용하는 방법을 고려해야 한다.

○ 작업의 유동적인 실시간 이동성(마이그레이션)에 대한 지원

NVIDIA GRID의 출현으로 GPU에서 진정한 하드웨어 기반 가상화가 가능해졌다. NVIDIA GRID를 사용하면 GPU 장치가 하드웨어 수준에서 여러 별도의 가상 장치로 나타난다. 그러나 이 기술은 가상화 계층(하이퍼바이저)을 우회하기 때문에 클라우드 컴퓨팅에서 가장 중요한 가상화 기능 중 하나인 가상머신의 실시간 작업이동(Live Migration)[44]에 문제를 일으킬 수 있다.

---

42 운영체제 내에서 시스템을 통제하는 핵심 영역이자 프로그램으로, 운영체제의 많은 서비스 기능들과 응용 프로그램 수행에 필요한 서비스들을 제공

43 하드웨어 내의 메모리 및 레지스터 등을 공간적으로 분리하여 활용하는 멀티프로세스 방법, 다른 방법으로는 점유시간을 분할하는 시분할(Time Multiplexing) 기법도 있으며, 컴퓨팅 및 통신분야에서 자원을 공유하는 방식으로 널리 쓰임

44 라이브 마이그레이션, 각주 28 참조

하드웨어 기반 가상화에서 실시간 마이그레이션을 사용하려면 하드웨어 상태를 캡처하여 원격 시스템에서 복원할 수 있는 새로운 매커니즘이 필요하다. 최근 NVIDIA는 GPU 가상화 환경에서 실시간 마이그레이션을 지원할 수 있는 기술을 제안했다.[45] 가상머신에서 수행중인 작업을 다른 물리서버의 가상머신으로 옮기는 데에는 시스템의 가동 중단 시간의 최소화, 데이터 손실 방지, 접근 권한의 이동 등 고려해야 할 복잡한 요소들이 모두 충족되어야 한다.

○ 통신 오버헤드의 최적화

API 리모팅 방식은 게스트 운영체제에서 GPU 호출을 가로 채서 호스트 운영체제 또는 원격 시스템으로 전달하여 GPU 가상화를 제공한다. API 리모팅 방식은 응용 프로그램이 GPU를 많이 사용하지 않는 경우 오버 헤드가 적다. 그러나 최근의 GPU 애플리케이션은 크기는 작거나(예 : 100μs 미만의 수행 시간을 갖는 경우) 반복적인 GPU 커널 수행 작업을 하는 경우가 많다.

이럴 경우 매 작업마다 GPU 호출을 전달하는 것은 심각한 통신 병목 현상(Bottle Neck)이 발생할 수 있다. 더구나 GPU의 공정기술과 성능이 점차 발전함에 따라 커널 단위의 실행 시간은 더욱더 낮아질 것으로 예상되고, 이는 곧 잦은 통신 오버 헤드의 증가를 야기할 것으로 예상된다. API 리모팅 방식에서 GPU 집약적인 응용프로그램을 수용하려면 훨씬 최적화된 통신 방법이 필요하다.

○ 통합된 형태의 CPU-GPU 칩 개발

기존의 시스템은 CPU와 GPU가 분리되어 있으며, 이들 간의 통신은 시스템 버스(System Bus)[46]를 통해 이루어진다. 인공지능의 발달로 작업의 복잡성과 크기는 점차 증가하고 있기 때문에, 현재의 PCIe 시스템 버스의 대역폭 용량(16Gbps)으로는 데이터 전송 오버헤드를 해결하지 못한다. 또한, 기본적으로 분리된 데이터 공간을 활용하는 CPU와 GPU 때문에 프로그래밍의 어려움을 겪게 된다.

이러한 문제를 해결하기 위해 통합된 CPU-GPU 칩이 출현했으며 이 칩은 두 프로세서 간에 통합된 메모리 공간을 제공한다. 예를 들면 Intel의 통합 CPU-GPU 칩, AMD의 HSA 아키텍처 등이 있다. 이러한 새로운 아키텍처는 두 프로세서 간에 상당한 통신량이 필요한 큰 크기의 데이터 응용프로그램의 성능을 향상시킬 수 있다. 기존의 연구들 중 Intel 및 AMD의 통합칩에 대한 솔루션 개발이 있었지만(gVirt), 두 칩의 작업 부하를 균형있게 조정(Load Balancing)하는 것은 부족했다. 두 프로세서를 최적화하여 사용할 수 있는 정교한 스케줄링 알고리즘이 필요하다. 또한, 최근 NVIDIA는 GPU와 CPU 사이의 통신에 높은 대역폭(80~200Gbps)을 가능하게 하는 NVLink[47]라는 기술을 출시했는데, 데이터 집약적인 애플리케이션의 고성능을 달성하기 위해서는 통합 메모리 기법과 NVLink의 결합과 같은 연구개발도 필요하다.

45 https://www.nvidia.com/ko-kr/data-center/virtualization/virtual-gpu-migration/ 참조
46 컴퓨터의 주요 구성요소인 주변기기들을 연결하는 데이터 및 제어명령의 통로이며, CPU의 빠른 처리 속도 대비 상대적으로 느린 주변기기들 간의 속도차를 보정해주기 위한 기능들을 탑재하고 있음
47 https://www.nvidia.com/en-us/data-center/nvlink/ 참조

4.2. 요약과 시사

○ 고성능 컴퓨팅 패러다임의 변화와 GPU 가상화 기술 수요 증대

4차 산업혁명 시대의 인공지능의 확산과 여기에 맞물린 클라우드 컴퓨팅에 대한 수요가 늘어나면서, 서비스 벤더 및 연구자들은 비용 대비 높은 성능을 제공할 수 있는 컴퓨팅으로 GPU 활용을 도입했다. 이러한 CPU + GPU 조합의 이기종 컴퓨팅은 지난 몇 년 동안 고성능 컴퓨팅(HPC) 및 클라우드 플랫폼에서 높은 성능과 자원 활용도를 상대적으로 낮은 운영비용으로 제공할 수 있는 패러다임으로 주목받아왔다. 클라우드 컴퓨팅에서 가상화는 전제가 되는(Basic Technology) 기술이다. 곧, 이기종의 머신들이 있는 클라우드 데이터 센터에서 GPU의 가상화는 사용자들 또는 서비스별 가상머신들 간에 GPU 장치를 효과적으로 공유하는 방법으로써 충분한 컴퓨팅 파워를 제공하기 위한 핵심 기술이 된다.

○ GPU 가상화를 위한 다양한 시도와 벤더들의 태세전환

GPU 가상화를 구현하는 것은 최근까지 어려운 일이었다. GPU를 제조·공급하는 벤더들은 GPGPU라는 개념이 등장하기 전까지, 가상화에 큰 관심이 없었다. 인공지능에 대한 관심이 늘어나고, 그 기반기술로 클라우드 컴퓨팅이 재조명되면서, 본격적으로 GPU 벤더들은 자체적으로 클라우드 컴퓨팅을 위한 가상화 지원 기술을 개발·출시를 가속화했다. GPU를 가상화하기 위한 방법은 크게 API 리모팅 방식, 전/반가상화를 활용한 방식, 하드웨어 지원 방식으로 나눌 수 있다. [표 4-1]은 앞서 살펴본 세 가지 방식에 대한 특징을 정리한 것이다.

[표 4-1] GPU 가상화 방식의 특징(요약)

<table>
<tr><th colspan="2">구분</th><th>특징</th><th>장점</th><th colspan="2">단점</th></tr>
<tr><td colspan="2">API 리모팅</td><td>• GPU 하드웨어가 지원하지 않는 GPU 가상화의 구현
– GPU 라이브러리의 복제인 래퍼를 가상머신에 제공, 원격처리</td><td>• 드라이버가 공개되지 않은 상용 GPU에 대해 드라이버 우회를 통한 가상화 실현</td><td colspan="2">• 성능저하
• 지속 업데이트 불가
• 어렵고 번거로움</td></tr>
<tr><td rowspan="2">하이퍼바이저 활용</td><td>전가상화</td><td rowspan="2">• 가상화 플랫폼인 하이퍼바이저를 활용한 GPU 가상화
– GPU 드라이버를 가상화</td><td>• GPU 드라이버의 수정 없이 에뮬레이션</td><td rowspan="2">• 드라이버 의존도 높음</td><td>• 다소 낮은 성능</td></tr>
<tr><td>반가상화</td><td>• 가상머신이 직접 드라이버 제어 및 성능향상</td><td>• 드라이버 수정 필요</td></tr>
<tr><td colspan="2">하드웨어 지원</td><td>• GPU 및 CPU 벤더들의 제품에서 가상화 기술 지원</td><td>• 높은 성능 및 호환성</td><td colspan="2">• 높은 하드웨어 종속성</td></tr>
</table>

○ GPU 가상화 기술의 상용화는 이제부터 본격화

GPU를 그래픽 처리 이외의 목적으로 활용하는 GPGPU 개념이 등장하게 되면서, 컴퓨팅 가성비가 상승하는 효과를 불러왔다. GPU는 AI학습을 위한 데이터 병렬처리에 최적화된 구조를 가지고 있으며, AI 및 클라우드 확산에 따른 GPGPU 수요가 증가하면서 칩제조사들은 GPU 가상화를 지원하고 나섰다. GPU 가상화 영역은 향후 기술의 성능향상 및 확장성, 최적화에 더 다양한 연구개발 필요한 분야이다.

특히, GPU의 활용은 클라우드 컴퓨팅의 인프라를 구축 및 확보하고 다양한 서비스를 가능하게 하는 기반 기술로 연계된다. AI의 안정적인 성능확보 및 활용 확산을 위해서 인프라에 대한 투자는 반드시 필요하다.

정부도 클라우드 컴퓨팅 활성화를 위한 기본계획[48]을 발표한 바 있고, 최근에는 AI 클러스터 구축[49]을 위한 사업을 추진하며 국가전략[50]을 발표하는 등 다양한 노력을 기울이고 있다. 이에 따라 GPU 컴퓨팅과 같은 AI 및 클라우드와 연계되는 인프라 원천기술의 확보를 위해, 관계부처와 산·학·연의 적극적인 협력이 있어야 한다.

참고자료

1. 국내

[1] 소프트웨어정책연구소(2018.12.), "클라우드 가상화 기술의 변화"

[2] 엔비디아, https://www.nvidia.co.kr/

[3] AMD, https://www.amd.com/ko

[4] 인텔, https://www.intel.co.kr/

2. 국외

[1] http://Top500.org

[2] nVIDIA(2019.11.19.), "Record 136 NVIDIA GPU-Accelerated Supercomputers Feature in TOP500 Ranking"

[3] Coates, Adam, et al(2013), "Deep learning with COTS HPC systems." ICML

[4] PWC(2019.7.), "Global Top100 companies by market capitalisation"

[5] iWeblists(2019.11.30.), "US commerce Stock Market Capitalization of the 50 Largest American Companies"

[6] CorporateInformation(2019), Top100 list

[7] TechPowerUp(2018.7.), "Intel is Giving up on Xeon Phi - Eight More Models Declared End-Of-Life"

[8] TowardsDataScience(2019.3.), "Monitor and Improve GPU Usage for Training Deep Learning Models"

48 과학기술정보통신부(2018.12.28.), "4차 산업혁명 체감을 위한 클라우드 컴퓨팅 실행(ACT) 전략 - 제2차 클라우드 컴퓨팅 발전 기본계획"

49 인공지능 중심 산업융합 집적단지 조성사업(광주), 2019.8.

50 관계부처합동(2019.12.), "인공지능 국가전략"

[9] Gupta, Vishakha, et al(2009), "GViM : GPU-accelerated virtual machines." Proceedings of the 3rd ACM Workshop on System-level Virtualization for High Performance Computing. ACM

[10] Shi, Lin, et al(2012), "vCUDA : GPU-accelerated high-performance computing in virtual machines." IEEE Transactions on Computers 61.6 : 804-816

[11] Duato, José, et al(2010), "rCUDA : Reducing the number of GPU-based accelerators in high performance clusters." High Performance Computing and Simulation(HPCS), 2010 International Conference on. IEEE

[12] Montella, Raffaele, et al(2017), "On the virtualization of CUDA based GPU remoting on ARM and X86 machines in the GVirtuS framework." International Journal of Parallel Programming 45.5 : 1142-1163

[13] Li, Teng, et al(2011), "GPU resource sharing and virtualization on high performance computing systems." Parallel Processing(ICPP), 2011 International Conference on. IEEE

[14] Gupta, Vishakha, et al(2011), "Pegasus : Coordinated scheduling for virtualized accelerator-based systems." 2011 USENIX Annual Technical Conference(USENIX ATC'11)

[15] Merritt, Alexander M., et al(2011), "Shadowfax : scaling in heterogeneous cluster systems via GPGPU assemblies." Proceedings of the 5th international workshop on Virtualization technologies in distributed computing. ACM

[16] Xiao, Shucai, et al(2012), "VOCL : An optimized environment for transparent virtualization of graphics processing units." Innovative Parallel Computing(InPar), IEEE

[17] Oikawa, Minoru, et al(2012), "DS-CUDA : a middleware to use many GPUs in the cloud environment." High Performance Computing, Networking, Storage and Analysis(SCC), 2012 SC Companion. : IEEE

[18] Suzuki, Yusuke, et al(2014), "GPUvm : Why not virtualizing GPUs at the hypervisor?." USENIX Annual Technical Conference

[19] Tian, Kun, Yaozu Dong, and David Cowperthwaite(2014), "A Full GPU Virtualization Solution with Mediated Pass-Through." USENIX Annual Technical Conference

[20] Dong, Yaozu, et al(2015), "Boosting GPU Virtualization Performance with Hybrid Shadow Page Tables." USENIX Annual Technical Conference

[21] Xue, Mochi, et al(2016), "gScale : Scaling up GPU Virtualization with Dynamic Sharing of Graphics Memory Space." USENIX Annual Technical Conference

[22] Dowty, Micah, and Jeremy Sugerman(2009), "GPU virtualization on VMware's hosted I/O architecture." ACM SIGOPS Operating Systems Review 43.3 : 73-82.

[23] Gottschlag, Mathias, et al(2013), "LoGV : Low-Overhead GPGPU Virtualization." HPCC/EUC.

[24] Qi, Zhengwei, et al(2014), "VGRIS : Virtualized GPU resource isolation and scheduling in

cloud gaming." ACM Transactions on Architecture and Code Optimization(TACO) 11.2 : 17

[25] Huang, Yu-Ju, et al(2016), "Building a kvm-based hypervisor for a heterogeneous system architecture compliant system." ACM SIGPLAN Notices. Vol. 51. No. 7. ACM

[26] Cloud, Compute(2009), "Amazon Elastic Compute Cloud."

[27] Expósito, Roberto R., et al(2013), "General purpose computation - on GPUs for high performance cloud computing." Concurrency and Computation : Practice and Experience 25.12 : 1628-1642

[28] Yang, Chao-Tung, et al(2014), "Implementation of GPU virtualization using PCI pass-through mechanism." The Journal of Supercomputing 68.1 : 183-213

[29] Shea, Ryan, and Jiangchuan Liu(2013), "On GPU pass-through performance for cloud gaming : Experiments and analysis." Proceedings of Annual Workshop on Network and Systems Support for Games. IEEE Press

[30] Younge, Andrew J., and Geoffrey C. Fox(2014), "Advanced virtualization techniques for high performance cloud cyberinfrastructure." 2014 14th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing(CCGrid). IEEE

[31] Younge, Andrew J., et al(2015), "Supporting high performance molecular dynamics in virtualized clusters using iommu, sr-iov, and gpudirect." ACM SIGPLAN Notices. Vol. 50. No. 7. ACM

[32] Walters, John Paul, et al(2014), "GPU passthrough performance : A comparison of KVM, Xen, VMWare ESXi, and LXC for CUDA and OpenCL applications." Cloud Computing (CLOUD), 2014 IEEE 7th International Conference on. IEEE

[33] Jo, Heeseung, et al(2013), "Exploiting GPUs in virtual machine for BioCloud." BioMed research international 2013

[34] Vu, Lan, Hari Sivaraman, and Rishi Bidarkar(2014), "GPU virtualization for high performance general purpose computing on the ESX hypervisor." Proceedings of the High Performance Computing Symposium. Society for Computer Simulation International

[35] Herrera, Alex(2014), "NVIDIA GRID : Graphics accelerated VDI with the visual performance of a workstation." Nvidia Corp

[36] Hong, Hua-Jun, et al(2014), "GPU consolidation for cloud games : Are we there yet?." Proceedings of the 13th Annual Workshop on Network and Systems Support for Games. IEEE Press

자문

· 중앙대학교 전자전기공학부 홍철호 교수

SW 산업 통계

국내 SW 생산 현황

2019년 11월 전체 SW생산액은 4.61조 원이고 패키지SW 생산액은 0.84조 원, IT서비스 생산액은 2.88조 원, 게임SW 생산액은 0.89조 원이다. 11월 SW생산액은 2018년에 비하여 전체적으로 감소(9.2% 감소)하였는데 이는 IT서비스와 게임SW의 생산액이 감소(각각 9.5%, 14.5% 감소)한 것이 원인으로 분석된다. SW생산액은 매년 4분기에 집중되는 경향성이 있었으나, 2019년 10월, 11월 SW생산액은 이 같은 경향이 보이지 않고 1분기, 2분기, 3분기와 유사한 수준이다.

11월 패키지SW 생산액은 전월 대비 3.8% 증가하였으나 전년 동기 SW생산액보다 1.6% 감소한 0.84조 원이다. 세부 품목별로는 시스템SW 생산이 0.4조 원으로 전년 같은 기간 대비 1.3% 증가하였으며 응용SW 생산은 전년 동기 대비 4.1% 감소한 0.4조 원이다. 11월 패키지SW 생산 주요 증감 품목으로는 산업범용SW(2.2% 증가), 보안SW(2.7% 감소), 산업특화SW(9.9% 감소), 미들웨어SW(24% 감소) 등이 있다. IT서비스 생산액은 전월 대비 2.9% 증가하였으나 전년 동기 대비 9.5% 감소한 2.88조 원이다. 11월 IT서비스 생산 주요 증감 품목으로는 IT시스템 관리 및 지원서비스(4.2% 감소), IT컨설팅 및 시스템 개발(14% 감소) 등이 있다. 게임SW 생산액은 전월 대비 3.7% 증가하였으나 전년 동기 대비 14.5% 감소한 0.89조 원이다.

[표 1] 국내 SW 생산 현황 (단위 : 조 원)

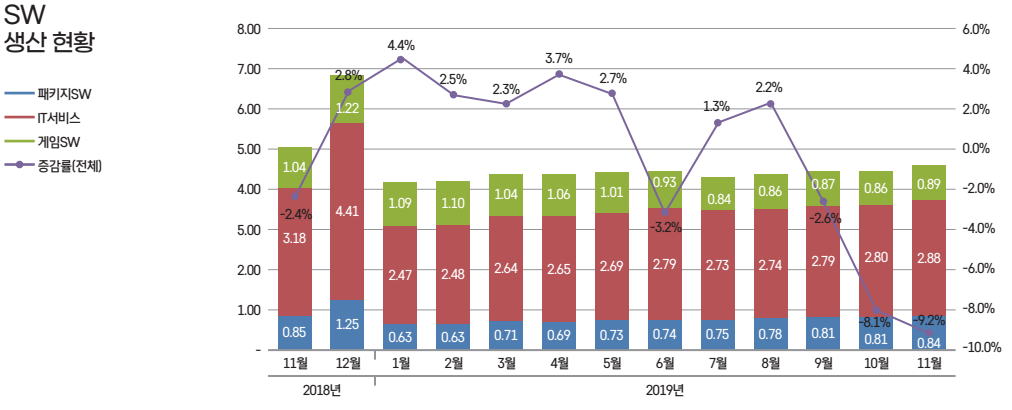
구 분			2018년			2019년					
			11월	12월	연간	1분기	2분기	3분기	10월	11월	11월 누적
패키지 SW	**생산액**		**0.85**	**1.25**	**9.45**	**1.97**	**2.17**	**2.34**	**0.81**	**0.84**	**8.13**
	증감률	동기대비	24.0%	11.3%	6.8%	0.4%	-1.8%	-0.7%	-0.8%	-1.6%	-0.9%
		전월대비	4.7%	46.8%	-	-	-	-	-0.4%	3.8%	-
IT 서비스	**생산액**		**3.18**	**4.41**	**33.98**	**7.59**	**8.13**	**8.25**	**2.80**	**2.88**	**29.65**
	증감률	동기대비	-5.1%	0.5%	0.7%	4.0%	2.9%	0.3%	-5.6%	-9.5%	0.3%
		전월대비	7.4%	38.5%	-	-	-	-	0.5%	2.9%	-
게임 SW	**생산액**		**1.04**	**1.22**	**12.10**	**3.22**	**3.01**	**2.57**	**0.86**	**0.89**	**10.55**
	증감률	동기대비	-10.1%	3.0%	5.8%	2.6%	-2.0%	0.7%	-20.4%	-14.5%	-3.0%
		전월대비	-3.5%	16.8%	-	-	-	-	-1.8%	3.7%	-
전체	**생산액**		**5.08**	**6.88**	**55.53**	**12.78**	**13.31**	**13.16**	**4.47**	**4.61**	**48.33**
	증감률	동기대비	-2.4%	2.8%	2.8%	3.0%	1.0%	0.2%	-8.1%	-9.2%	-0.7%
		전월대비	4.5%	35.5%	-	-	-	-	-0.1%	3.2%	-

출처) 과학기술정보통신부 ICT주요품목동향조사(2020.4.)
※ 월별 조사는 패널조사 결과를 합산한 잠정치로 연간조사를 통해 확정치로 대체 예정(2018년 : 2020.5., 2019년 : 2021.5.)

전이슬
JEON, Yiseul
연구원
Researcher, SPRi
jys@spri.kr

그림 1-1

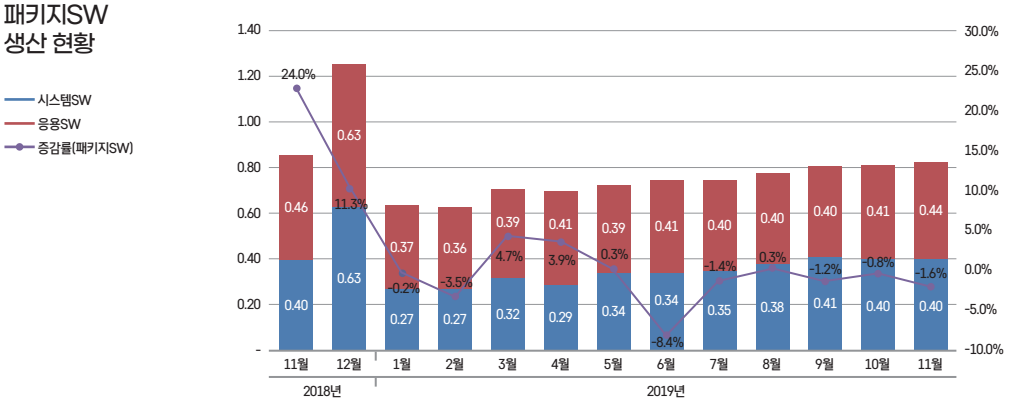
SW 생산 현황



단위 : 조 원, 전년 동기 대비 증감률

그림 1-2

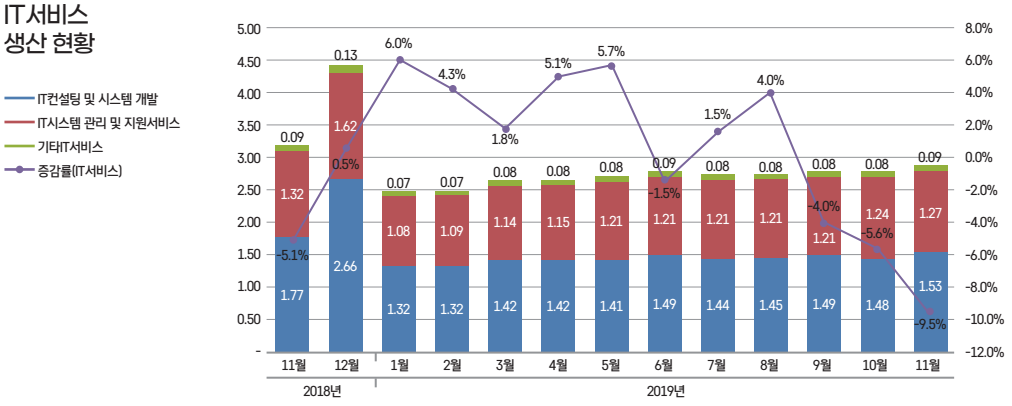
패키지SW 생산 현황



단위 : 조 원, 전년 동기 대비 증감률

그림 1-3

IT서비스 생산 현황



단위 : 조 원, 전년 동기 대비 증감률

국내 SW 수출 현황

2019년 3분기 전체 SW수출액은 31.95억 달러이고 패키지SW 수출액은 3.91억 달러, IT서비스 수출액은 17.38억 달러, 게임SW 수출액은 10.66억 달러이다. 3분기 SW수출액은 2018년 3분기 대비 13.8% 증가하였는데 이는 패키지SW 수출액이 전년 동기 대비 증가(53.8% 증가)한 것이 원인으로 분석된다. 2019년 1분기, 2분기, 3분기 SW수출액은 2018년 같은 기간 대비 전반적으로 높은 수준(1분기 7.3% 증가, 2분기 13.3% 증가, 3분기 13.8% 증가)을 기록하고 있으며 2019년 4분기 SW수출액 또한 2018년 4분기 대비 높을 것으로 전망된다.

3분기 패키지SW 수출액은 2분기 대비 13.4% 감소하였으나 전년 동기 대비는 53.8% 증가한 3.91억 달러이다. 세부 품목별로는 시스템SW 수출이 0.4억 달러로 전년 동기 대비 79.4% 증가하였으며 응용SW 수출은 전년 같은 기간 대비 51.3% 증가한 3.51억 달러이다. 3분기 패키지SW 수출 주요 증감 품목으로는 미들웨어SW(70.6%), 산업범용SW(27.1%), 보안SW(0.9% 증가), 산업특화SW(53.6% 감소) 등이 있다. IT서비스 수출액은 2분기 대비 3.1% 증가하였고 전년 동기 대비 14.3% 증가한 17.38억 달러이다. 3분기 IT서비스 수출 주요 증감 품목으로는 IT컨설팅 및 시스템 개발(39.1% 증가), IT시스템 관리 및 지원서비스(8.6% 증가) 등이 있다. 게임SW 수출액은 2분기 대비 0.9% 증가하였고 전년 동기 대비 3.2% 증가한 10.66억 달러이다.

[표 2] 국내 SW 수출 현황

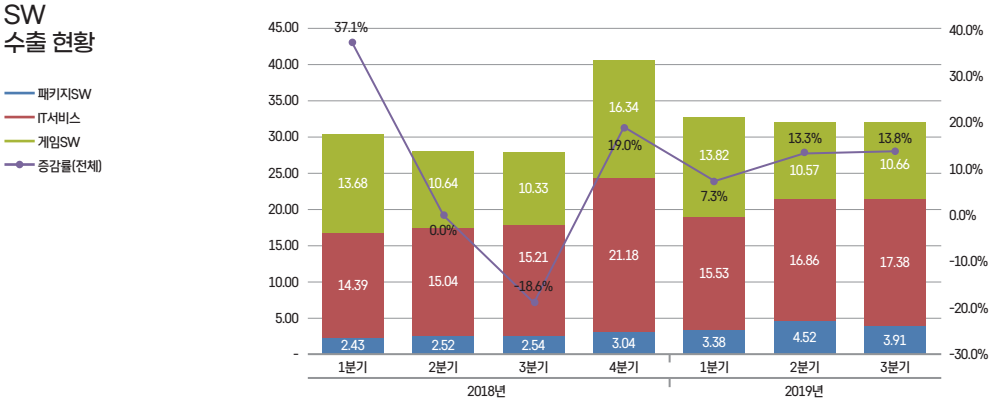
(단위 : 억 달러)

구 분			2018년			2019년			
			3분기	3분기 누적	연간	1분기	2분기	3분기	3분기 누적
패키지 SW	수출액		**2.54**	**7.50**	**10.53**	**3.38**	**4.52**	**3.91**	**11.81**
	증감률	동기대비	-3.9%	-0.7%	7.2%	39.2%	79.2%	53.8%	57.6%
		전분기대비	1.0%	-	-	11.4%	33.5%	-13.4%	-
IT 서비스	수출액		**15.21**	**44.65**	**65.83**	**15.53**	**16.86**	**17.38**	**49.77**
	증감률	동기대비	-7.8%	0.3%	7.0%	7.9%	12.1%	14.3%	11.5%
		전분기대비	1.1%	-	-	-26.7%	8.6%	3.1%	-
게임 SW	수출액		**10.33**	**34.65**	**50.98**	**13.82**	**10.57**	**10.66**	**35.04**
	증감률	동기대비	-32.8%	5.2%	6.8%	1.0%	-0.7%	3.2%	1.1%
		전분기대비	-3.0%	-	-	-15.4%	-23.5%	0.9%	-
합계	수출액		**28.08**	**86.79**	**127.34**	**32.73**	**31.95**	**31.95**	**96.63**
	증감률	동기대비	-18.6%	2.1%	7.0%	7.3%	13.3%	13.8%	11.3%
		전분기대비	-0.4%	-	-	-19.3%	-2.4%	0.0%	-

출처) 소프트웨어정책연구소 SW수출통계조사(2020.2.)
※ 분기별 조사는 패널조사 결과를 합산한 잠정치로 연간조사를 통해 확정치로 대체 예정(2018년 : 2020.5., 2019년 : 2021.5.)
※ SW수출액은 매월 입금된 금액을 기준으로 집계

그림 2-1

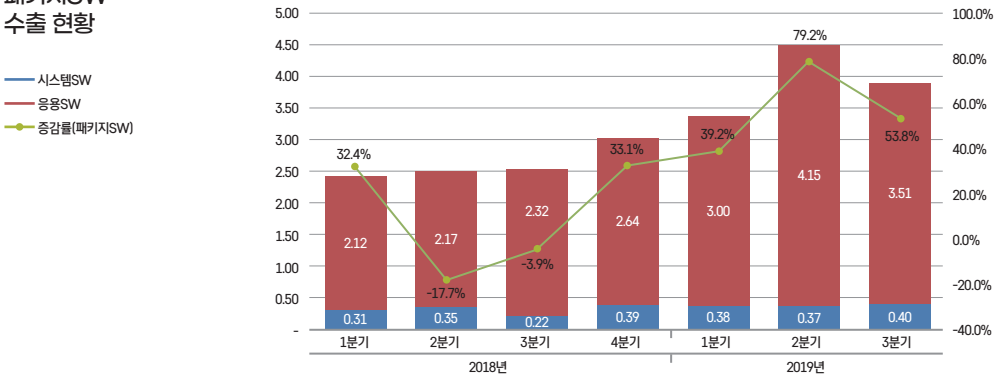
SW 수출 현황



단위 : 억 달러, 전년 동기 대비 증감률

그림 2-2

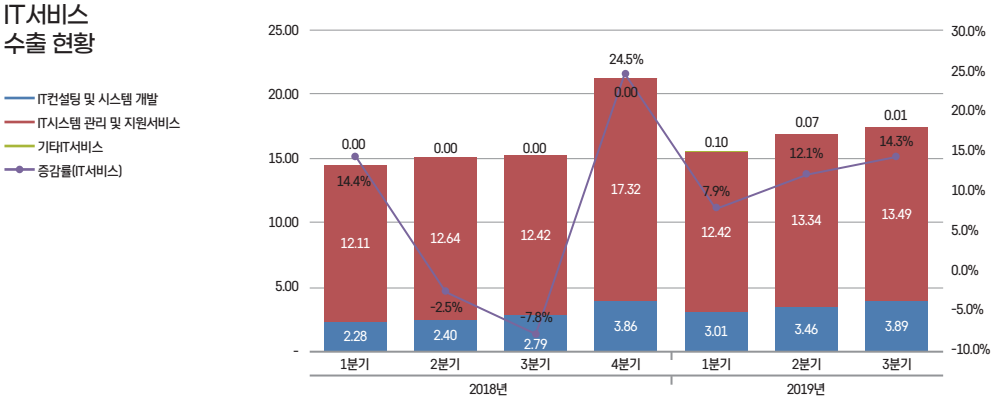
패키지SW 수출 현황



단위 : 억 달러, 전년 동기 대비 증감률

그림 2-3

IT서비스 수출 현황



단위 : 억 달러, 전년 동기 대비 증감률

월간 SW 중심사회
MONTHLY SOFTWARE ORIENTED SOCIETY

발행인 박현제
편집위원회 박태형 책임연구원, 임영모 책임연구원(소프트웨어정책연구소)
김덕현 교수(세종사이버대학교), 방은주 기자(ZDNet), 장윤희 교수(단국대학교)
발행처 소프트웨어정책연구소(Software Policy & Research Institute)
경기도 성남시 분당구 대왕판교로 712번길 22 글로벌 R&D센터 연구동(A)
홈페이지 www.spri.kr
전화 031.739.7300(+82-31-739-7300)
디자인·제작 (주)늘품플러스 www.npplus.co.kr

사명 Mission

소프트웨어 정책 연구를 통한 국가의 미래전략을 선도함
Leading Nation's Future Strategy through Research on Software Policy

미래상 Vision

국민 행복과 미래사회 준비에 기여하는 소프트웨어 정책 플랫폼
Software Policy Platform contributing to the public happiness and future society

핵심 가치 Core Values

역할 Roles

건강한 소프트웨어 산업 생태계 육성
To build a fair Ecosystem for Software Industry

소프트웨어 융합을 통한 사회 혁신
To innovate a Society through Software Convergence

국가 소프트웨어 통계 체계의 고도화
To advance the National Software Statistics System

개방형 소프트웨어 정책 연구 플랫폼 구축
To establish an Open Research Platform for Software Policy

소프트웨어정책연구소
Software Policy & Research Institute

이호준
Lee, Ho-Joon
언론학박사
ighwns@hanmail.net

우정사업본부에서 근무하고 있으며,
대한항공 여행사진 공모전에서
2회 수상하고, 세 차례의 개인전과
단체전 3회를 개최했다.

SPRi
소프트웨어정책연구소

경기도 성남시 분당구 대왕판교로 712번길 22 글로벌 R&D센터 연구동(A)
Global R&D Ceneter 4F, 22, Daewangpangyo-ro 712beon-gil, Bundang-gu, Seongnam-si, Gyeonggi-do
www.spri.kr

04
9 772586 502001
ISSN 2586-5021