SAMSUNG SDS

REAL 2019

Realize your vision
through Digital Transformation

2019 . 5 . 8 . Wed . The Shilla Seoul

Key Technologies

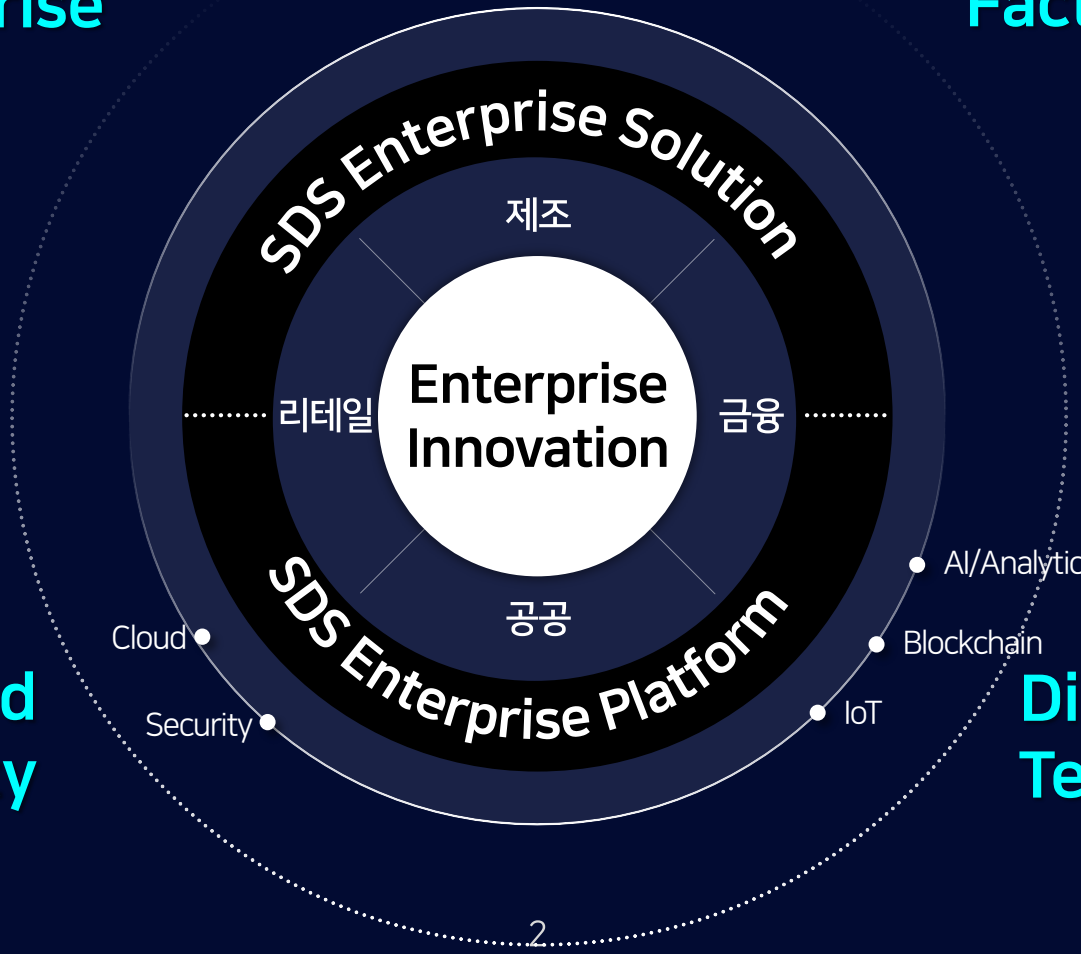
in Digital Transformation

윤심 부사장

삼성SDS의 Digital Transformation Framework

Intelligent Enterprise

Intelligent Factory



Cloud & Security

Disruptive Technologies

AI in Real World – 이상감지 / Intelligent Factory

설비가 한번 고장나면
손해가 많은데

2,200TB /월

2,000개 설비/공장

+

100~200개 센서/설비

+

수초 간격 데이터 수집

실시간 이상감지

머신러닝 정비시점 예측

고장 24시간~5일전 조기감지로 설비가동률 최대화

AI in Real World – 결함분류 / Intelligent Factory

불량유형 분류
쉽게 할 수 없을까?

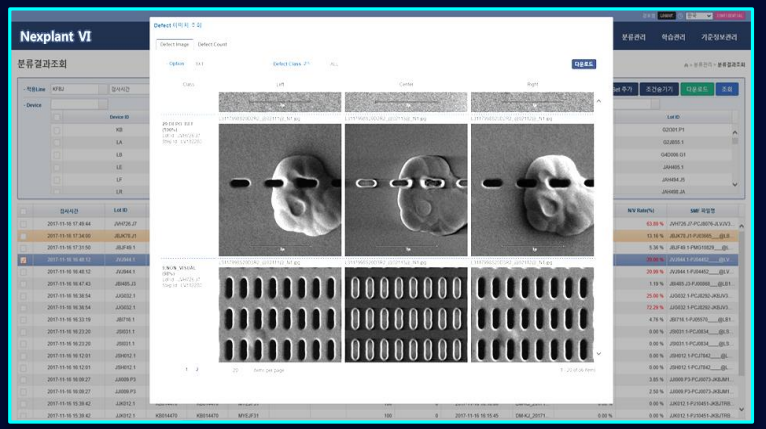
숙련된 검사원 확보 어려움
검사자 별 편차 발생

딥러닝
분류 자동화

이미지 정제 + 불량패턴 학습

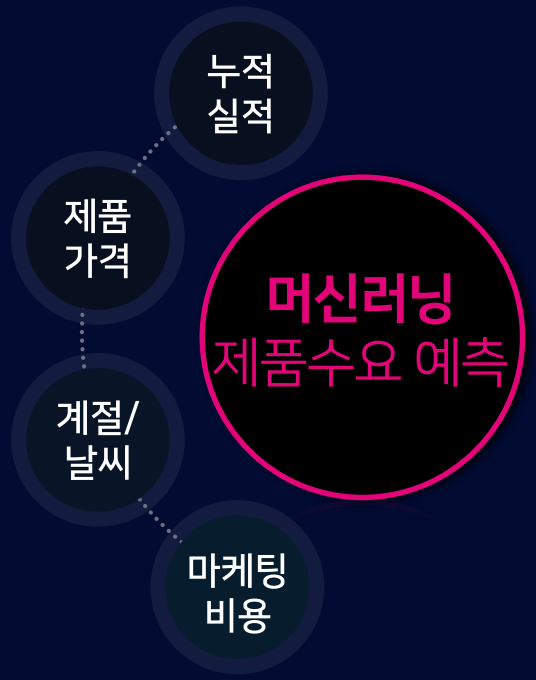
검사 ▸ 불량 검출 ▸ 유형 수동분류 ▸ 원인 파악

분류 정확도 **93%**



숙련된 품질검사원 대비 검출률 24% 향상

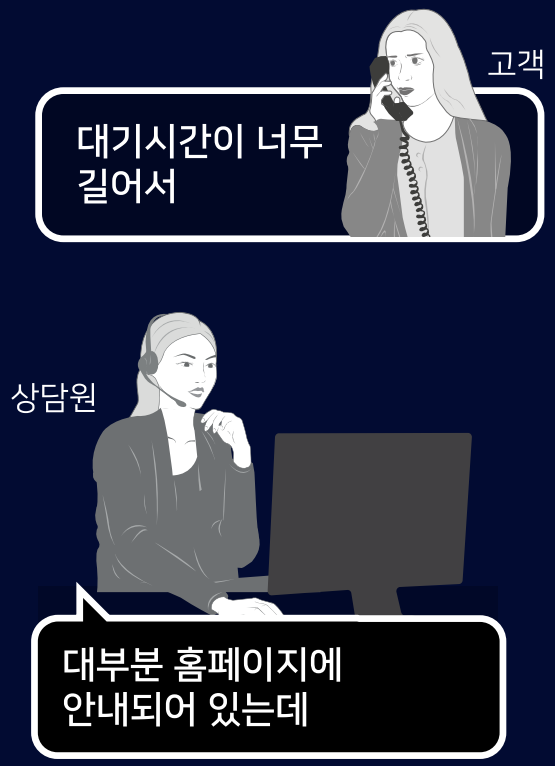
AI in Real World – 판매예측 / Intelligent Enterprise



수작업 대비 예측 정확도 25% 상승

적기 물량공급 및 재고 최소화

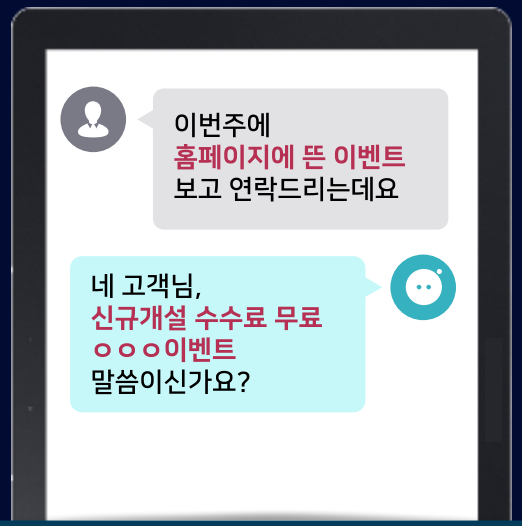
AI in Real World – 챗봇상담 / Intelligent Enterprise



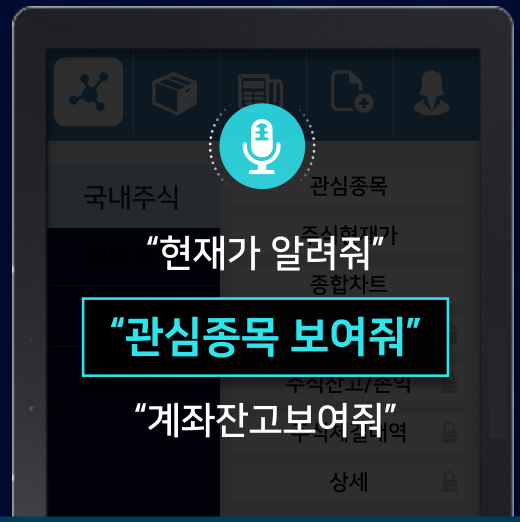
머신러닝 언어모델

단어추론 답변추천

대화의도 파악 95% 이상



의도를 이해한 자연어 대화

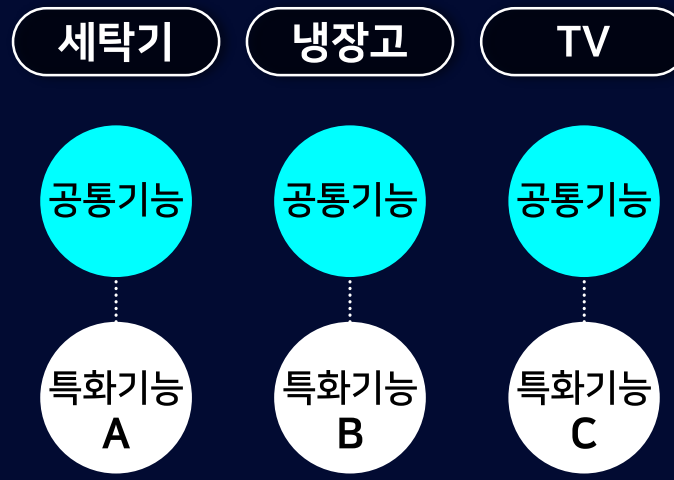


음성인식으로 대화형 서비스

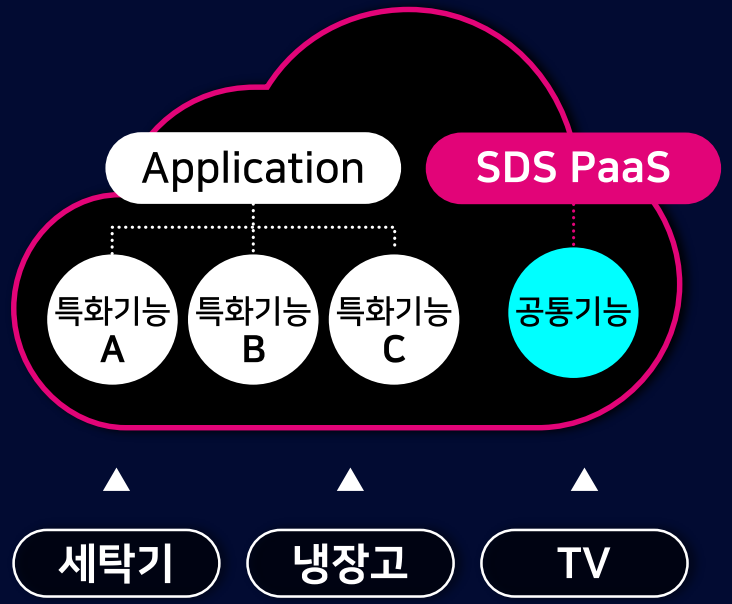
이용시간 2분에서 50초로 단축, 24시간 고객대응

Cloud & Security in Real World – Intelligent Factory

공통/특화기능
제품별 개발 및 운영



공통기능은 **PaaS** 로
특화기능은 **MSA**로 효율화



개발 및 운영 효율화로 35% 비용절감

※ MSA : 마이크로서비스 아키텍쳐

Cloud & Security in Real World – Intelligent Enterprise

사용패턴
분석

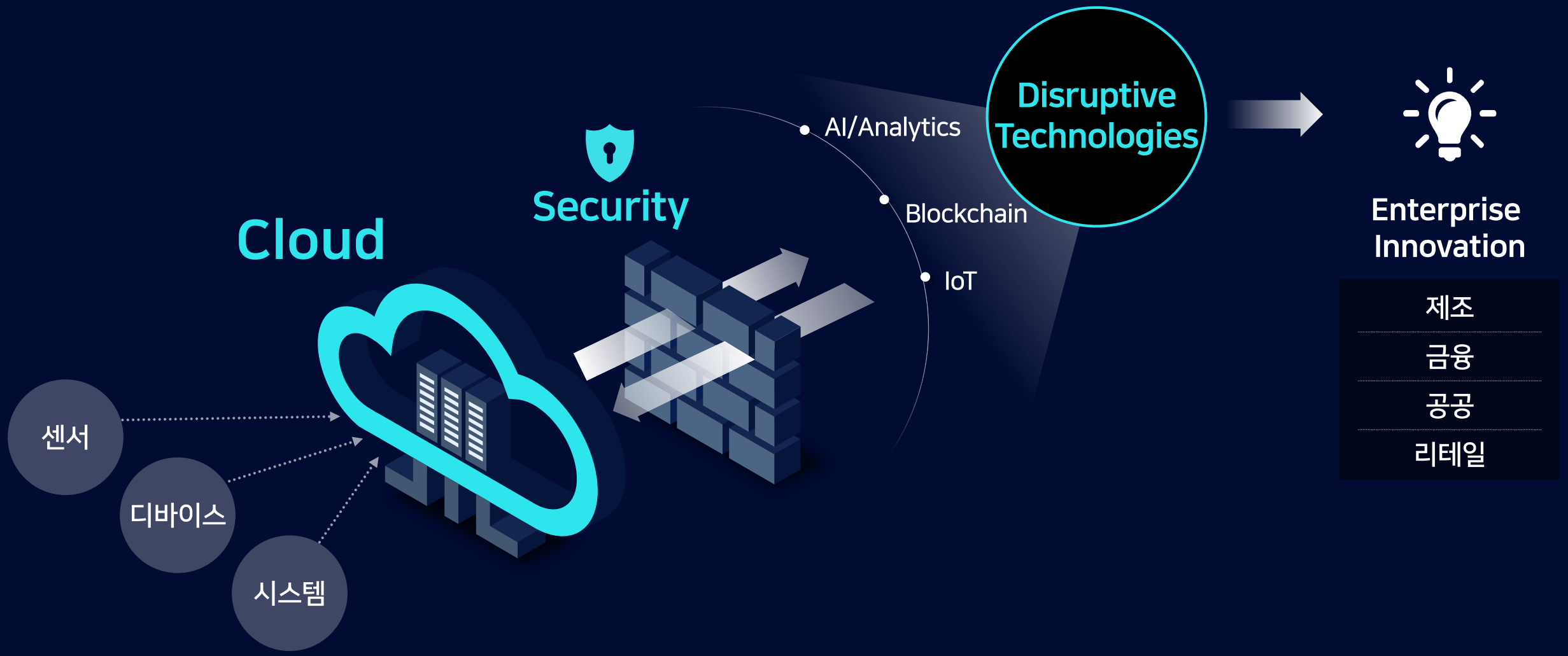
최적 클라우드 추천

Cloud A	**Cloud B**	Cloud C

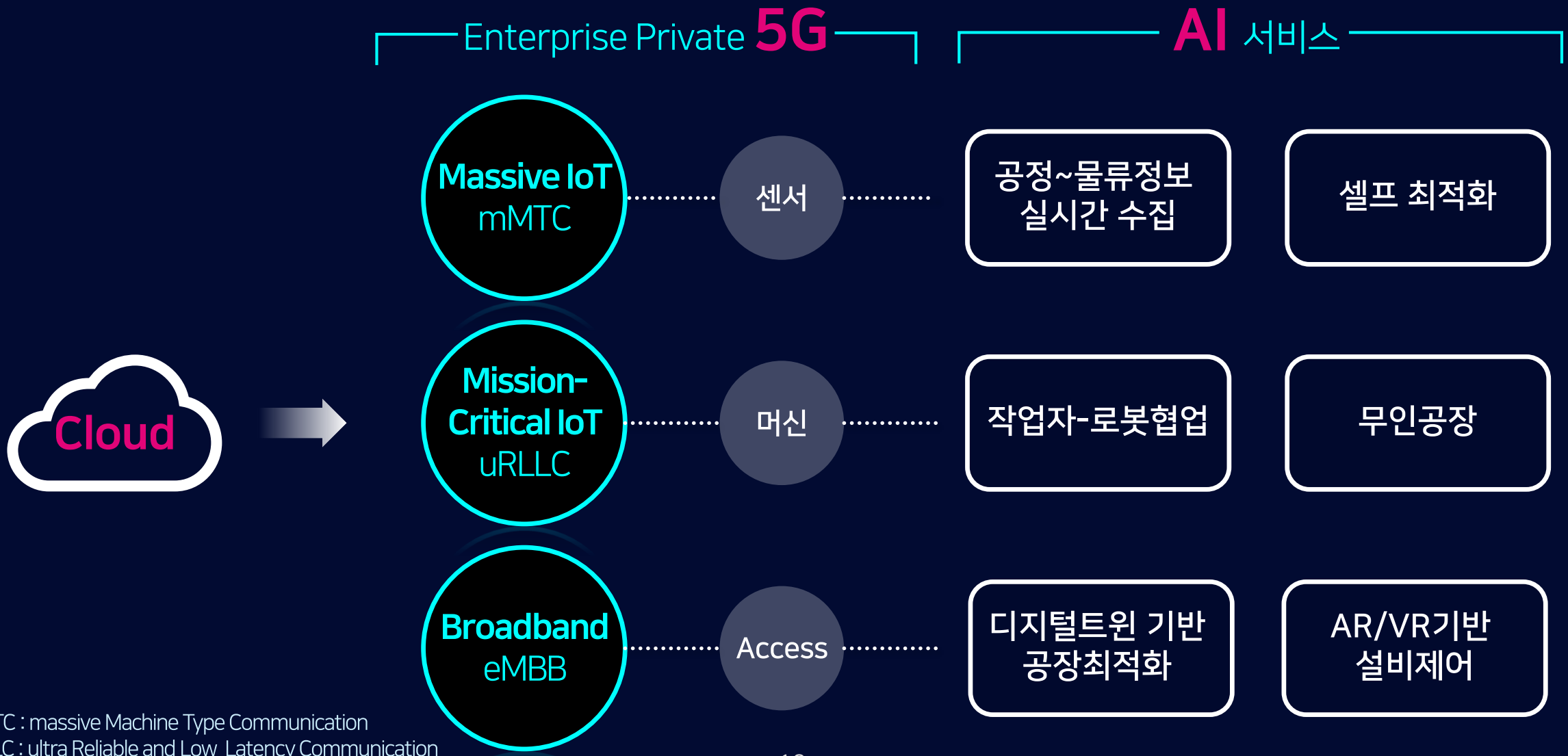
- 자원축소 추천
- 미사용 시간
  서버가동중단
- 트래픽 분석
  해킹탐지

클라우드 비용 20% 절감 및 서비스 연속성 확보

Digital Transformation을 실현하는 핵심기술



5G로 앞당기는 미래



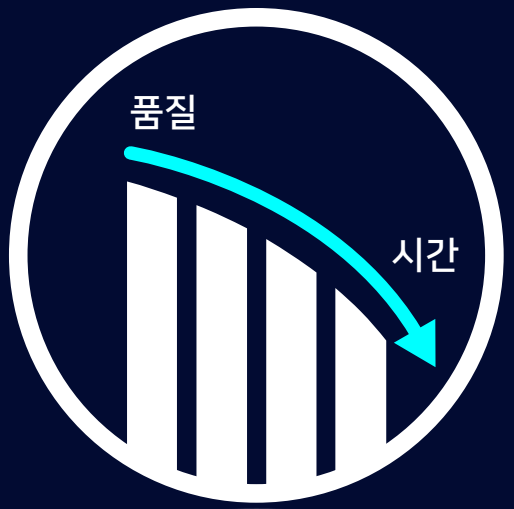
※ mMTC : massive Machine Type Communication
※ uRLLC : ultra Reliable and Low Latency Communication
※ eMBB : enhanced Mobile BroadBand

Technology Adoption - AI

Challenges in AI

데이터 사이언티스트 업무의
80%는 데이터 준비

시간이 지날수록
분석모델 품질 하락



비 전문가의
AI 수요 증가

AI 생산성 향상을 위한 Brightics AI

- **라벨생성 자동화**로 **15일 → 3일** 학습시간 단축 (15만장 기준)

- **모델 지속학습**을 통한 초기성능 유지
- 다양한 목적을 만족하는 **300종** 고성능 함수

- 검증된 재사용 모델 **100종**
- 최적 알고리즘 추천으로 **2주 → 2시간** 모델링 단축

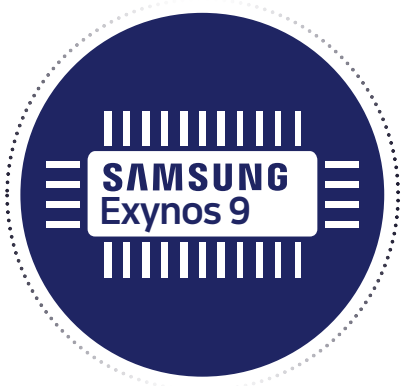
AI 트렌드와

삼성전자 종합기술원의 적용사례

삼성전자 종합기술원 심은수 전무

기술원 AI 기술 적용 사례

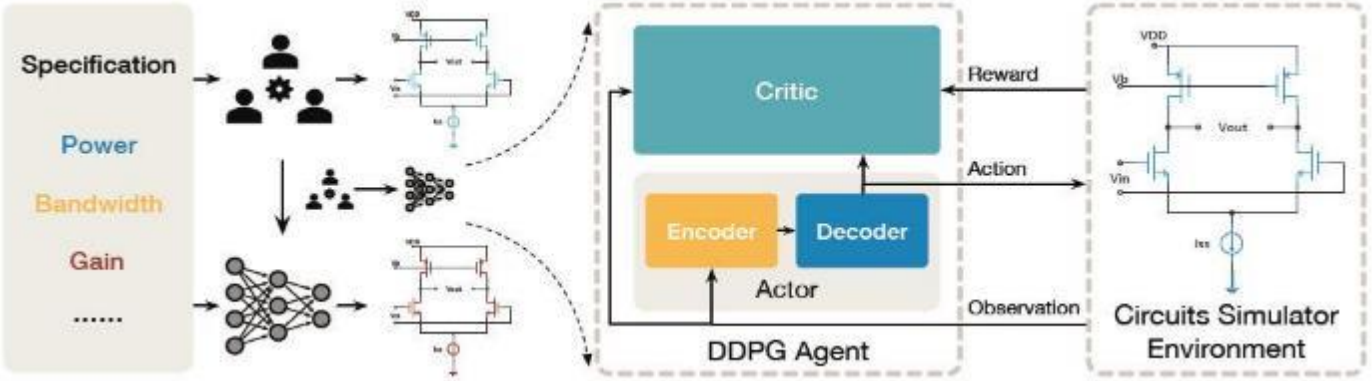
Neural Processor, Speech/Language, Biometrics



AI 기반 개발 혁신

Autonomous R&D?

Learning to Design Circuits



Autonomous Material Development

AI 기술 적용에서의 어려운 점

사람, 데이터, 컴퓨팅 인프라, 그리고 조건 변화에 따른 성능 저하

- AI 전문가 구하기 쉽지 않음
- Raw 데이터 저장 미비, 데이터 가공에 소요되는 시간
- 서버 클러스터 필요

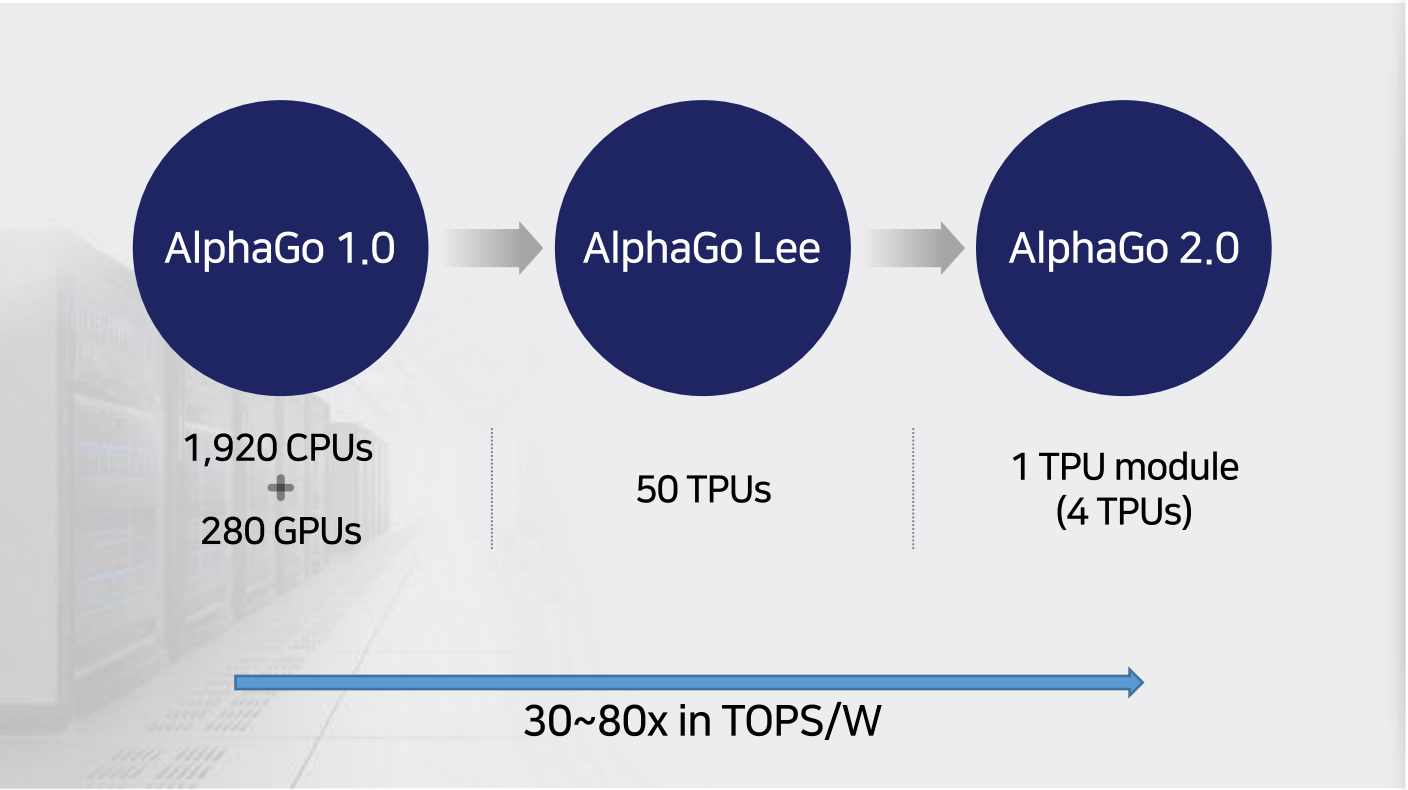
- 조건 변화
  → AI 알고리즘 성능 저하

Bad Case: 조건 변화 시간 간격 < 신규 데이터 확보 + 알고리즘 개발 소요 시간

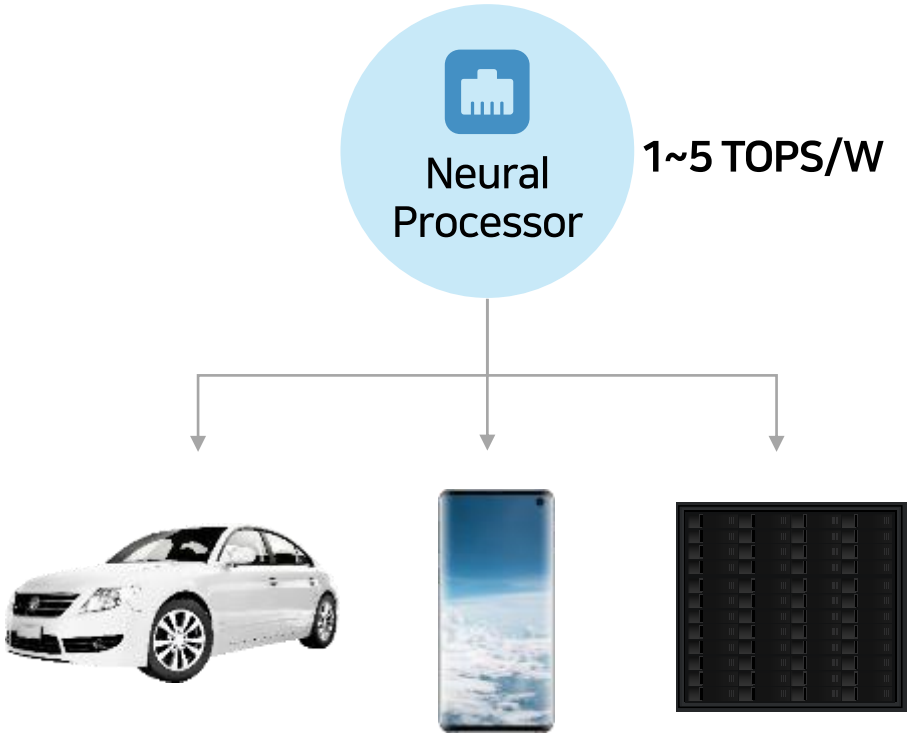
Neural Processor의 등장과 확산

AI 의 핵심 기술인 Deep Learning 연산에 최적화된 프로세서 ➔ 전력 효율 증대

Google server racks loaded with TPUs



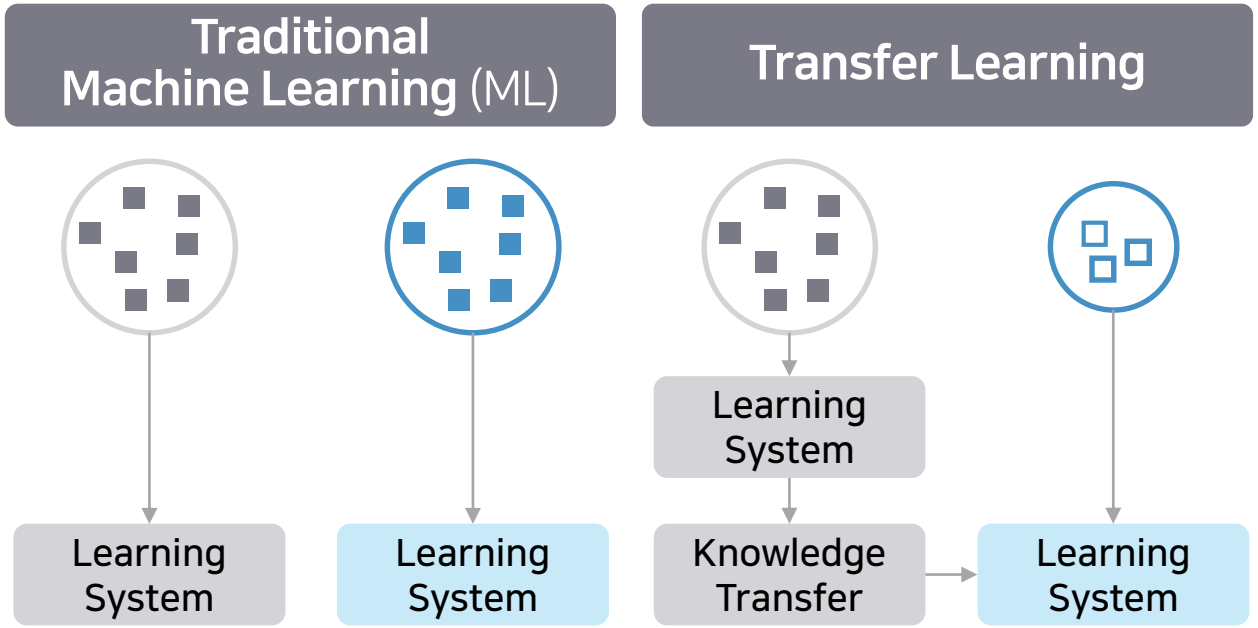
※ 출처: Google

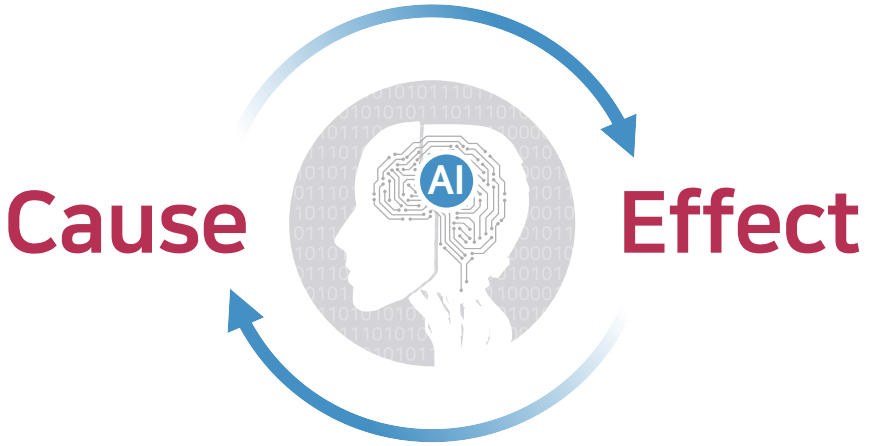


AI의 발전 방향

Narrow AI → Broad AI → General AI

Frontier in AI Research : Knowledge Transfer
Causality
High Level Abstraction & Generalization
Natural Language & Large Memory

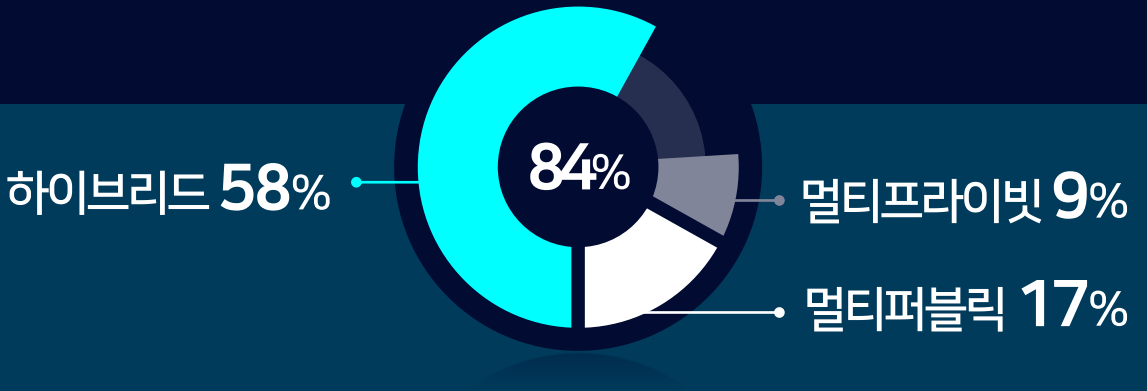




Technology Adoption - Cloud

Challenges in Cloud

글로벌 기업의
멀티 클라우드 활용률 84%



다른 클라우드로 이동 하려면
번거롭고 어려울 것 같은데?

여러 클라우드를 쓰면
비용이 증가하지 않을까?

여러 클라우드를 어떻게
효율적으로 관리할 수 있을까?

※ 출처: IDG Market Pulse, 2019 클라우드 현황 및 과제, 2019

하이브리드 클라우드 통합관리

컨테이너 기술을 활용한
자유롭고 유연한 마이그레이션

마이그레이션 기간
8주 ➡ 3주

사용패턴 분석으로
비용절감 방안 추천

GPU 가상화
25%⬇

최적 RI추천
10%⬇

Hybrid Cloud Platform
통합모니터링 및 보안관제

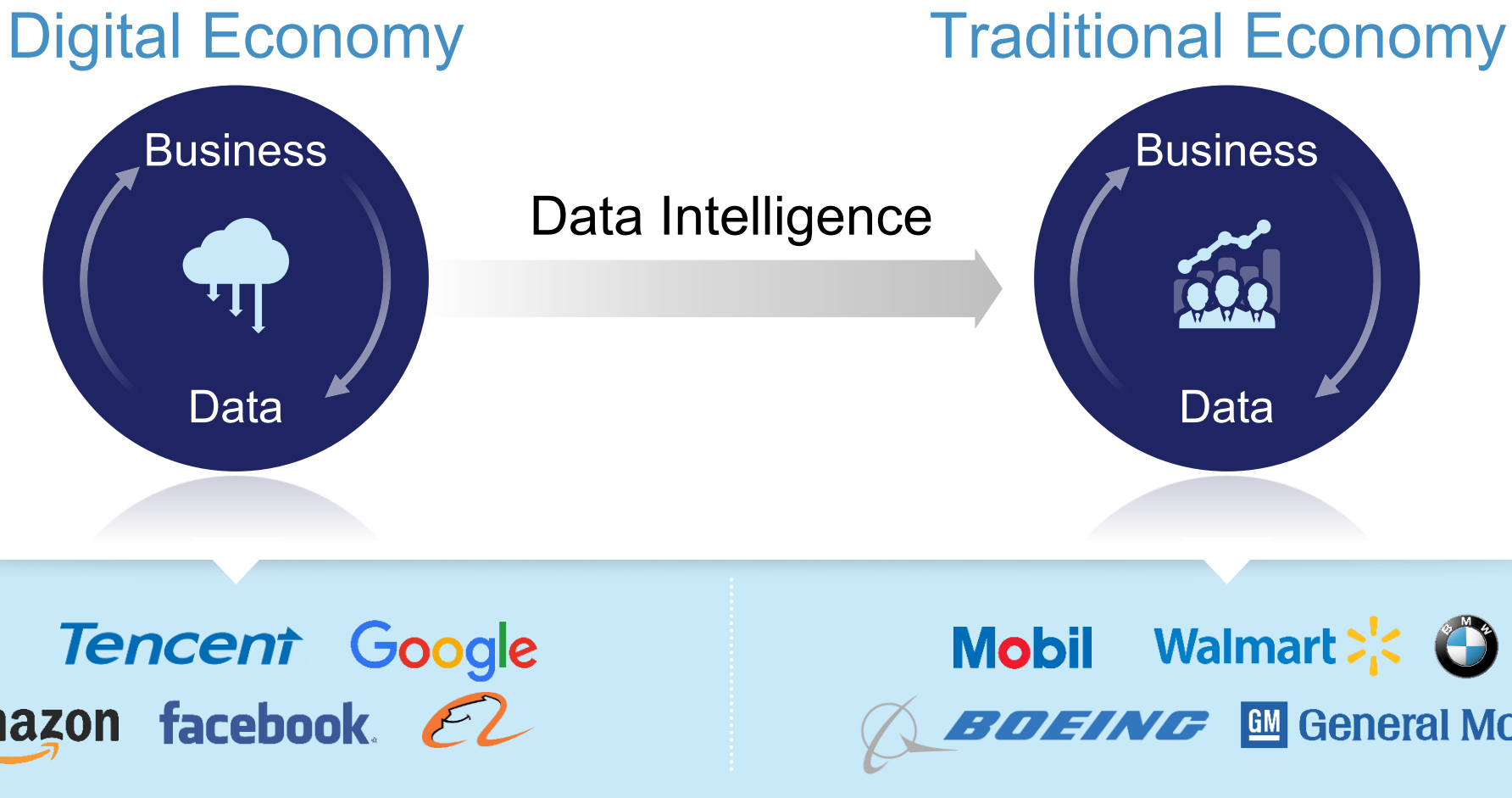
12개 글로벌 데이터 센터
(+17개 퍼블릭 파트너 지원)

The Role of Cloud Computing

in Digital Transformation

Jonathan Wu, CTO of Lenovo APAC

Data Intelligence is Powering the 4th Industrial Revolution



2019 Lenovo Internal. All rights reserved.

Smart IoT 2020

* Source: IDC 2018, Mario Morales

2019 Lenovo Internal. All rights reserved.

SMART CITY

Video Security and Surveillance Solutions

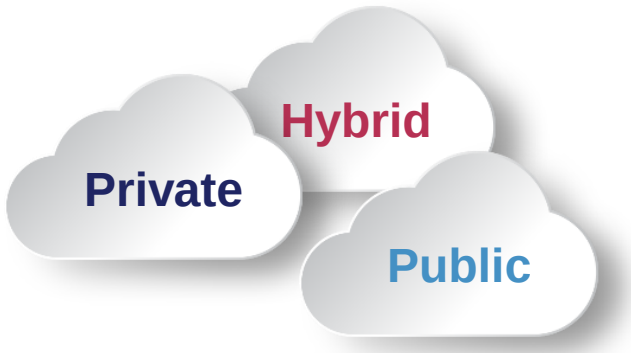
Central Management of **5000** Surveillance Cameras across **18** Boroughs in Bogotá

Improvement in Situational Awareness and Incidents Monitoring and Response

2019 Lenovo Internal. All rights reserved.

TRANSITION OF CLOUD COMPUTING

Technology → Financial → Business



Always available	Elastic Scalability	Near Zero Cost of Management	Layer of API

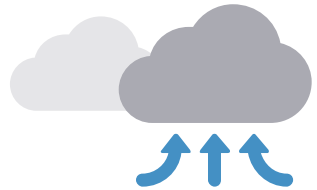
2019 Lenovo Internal. All rights reserved.

Key trends of Cloud computing

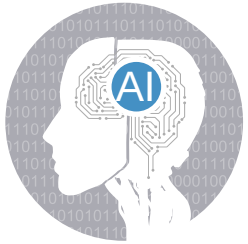
Containers, Microservices, and Cloud-native apps

DevOps & Continuous Software Delivery

On-prem & Hybrid Clouds

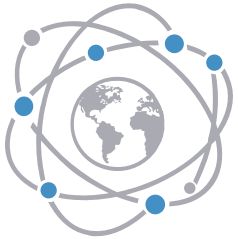


Data, AI, and Machine learning



Serverless computing

IoT/Edge Computing

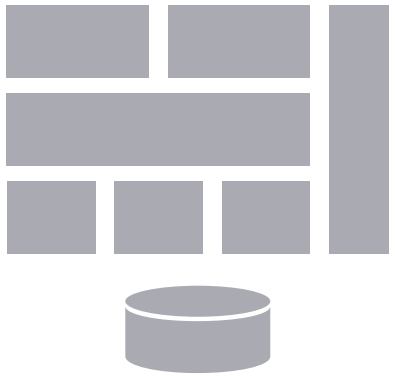


2019 Lenovo Internal. All rights reserved.

Monolithic Software vs. Micro-services

Challenges

- Complexity
- Validation costs
- Bulky updates



Monolithic/Layered

Micro-service

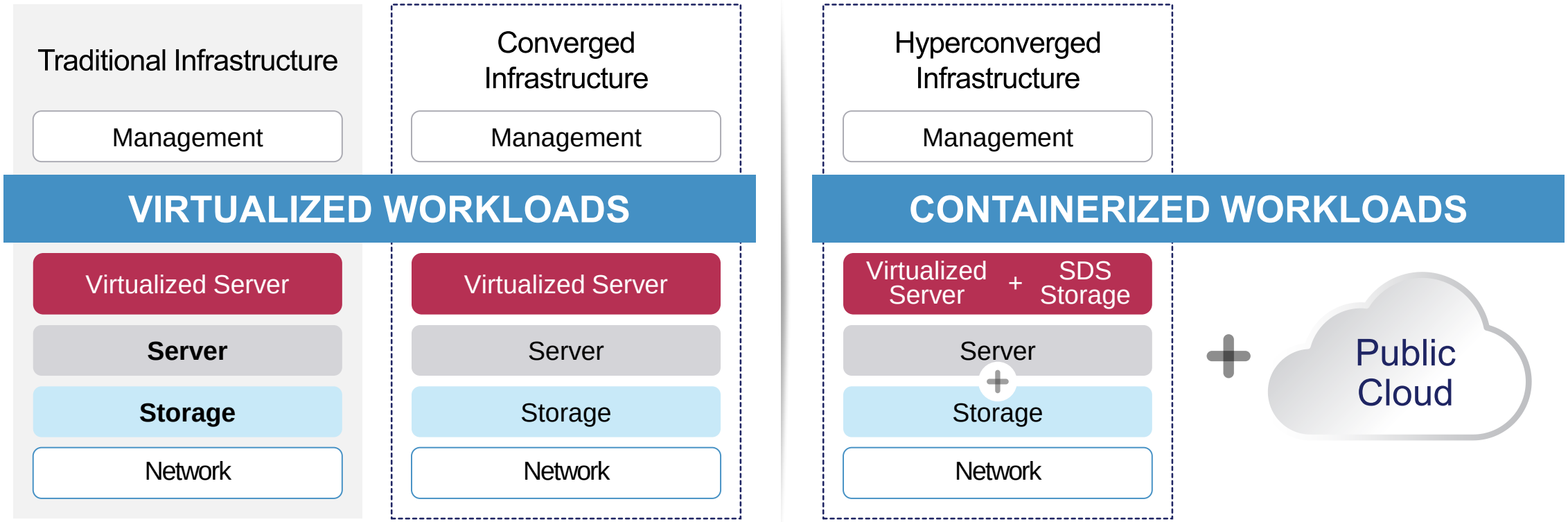
Benefits

- Code independence
- Scalability
- Portability

Traditional monolithic software re-architected to micro-services architecture

2019 Lenovo Internal. All rights reserved.

Containers and Hyperconverged Infrastructure for Cloud



Traditional/On-prem to Hyperconverged and Hybrid Cloud

2019 Lenovo Internal. All rights reserved.

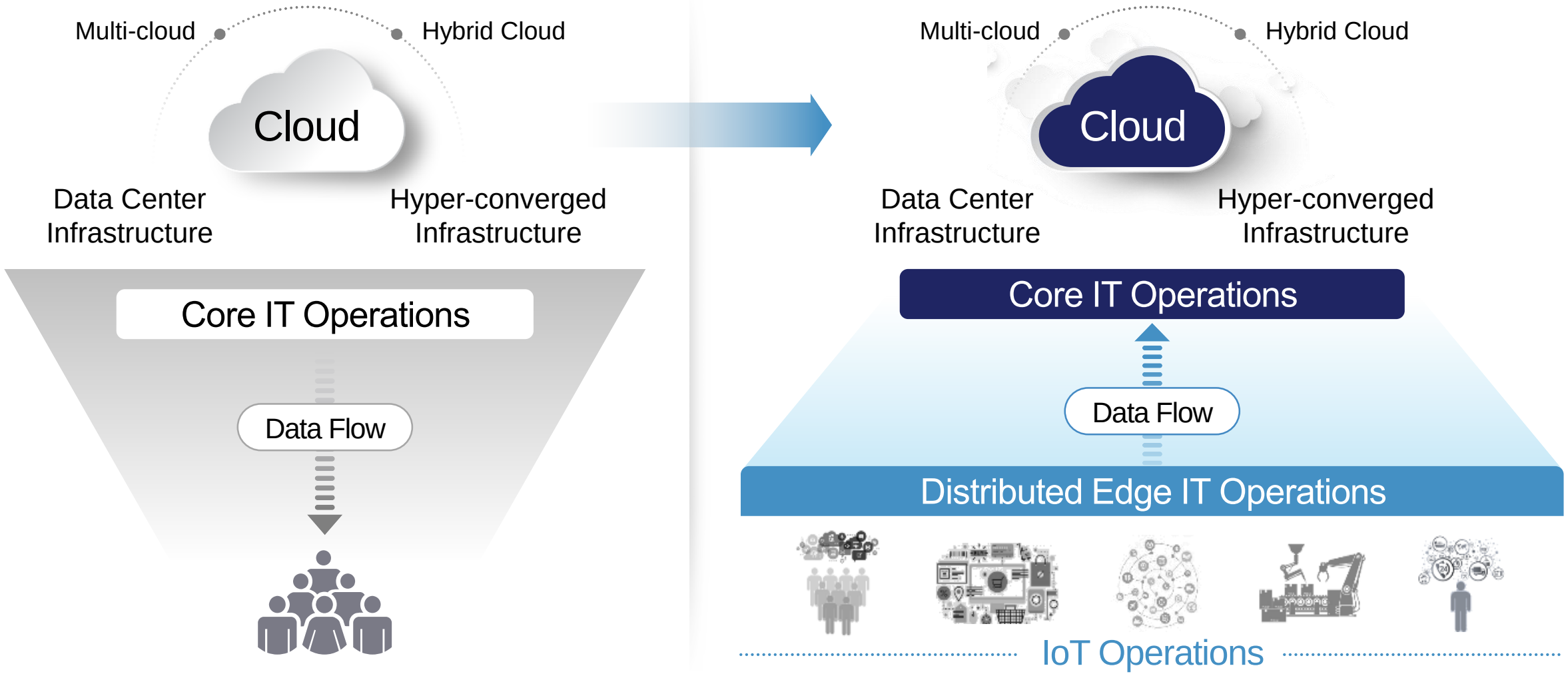
Serverless Computing

- New trend being driven primarily **by major Cloud providers**

- Does not require **dedicated compute servers to run code** (i.e. you will not be charged for compute usage)
- Code will be executed on-demand in **response to specific events**
- Cloud service provides **will charge based on the number of code invocations**

2019 Lenovo Internal. All rights reserved.

Edge to Cloud

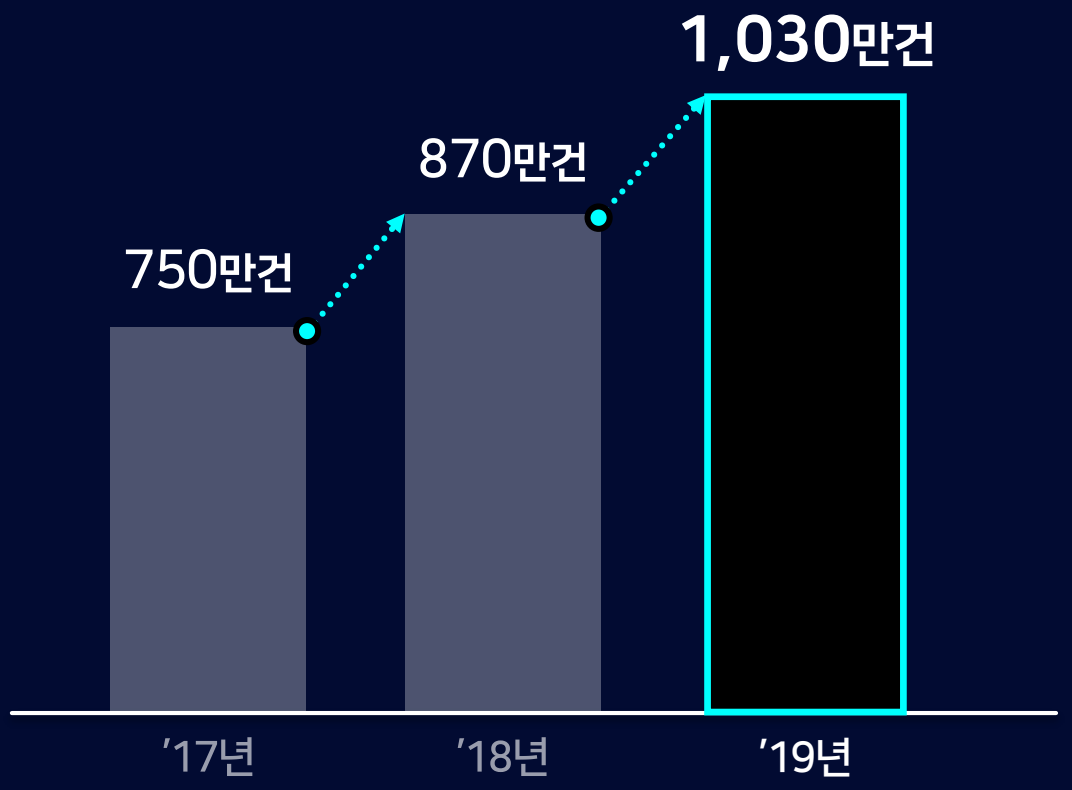


2019 Lenovo Internal. All rights reserved.

Technology Adoption - Security

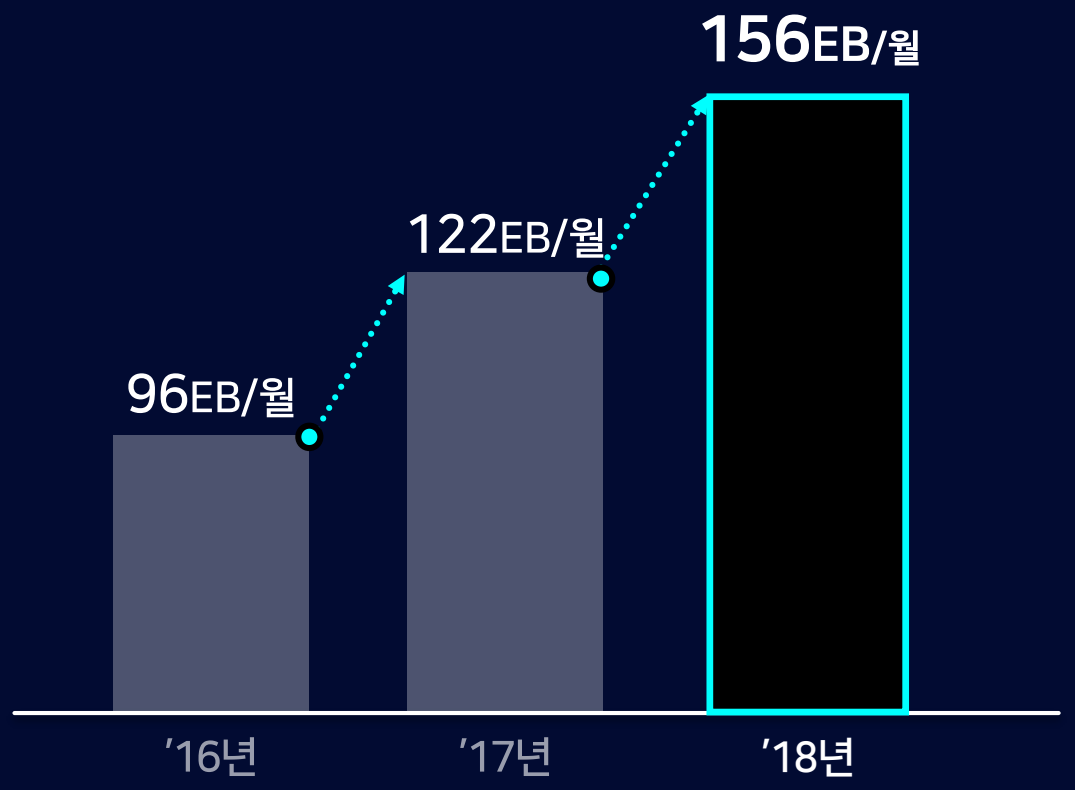
Challenges in Security

올해 전세계 DDoS 공격 **1천만 건** 이상 예상



*Source : Cisco, 비주얼 네트워킹 인덱스, 2018

네트워크 트래픽 **매년 30%** 증가



※ 1 EB (엑사바이트) = 100만 테라바이트 = 10^{18} 바이트

인프라에서 서비스까지 End-to-End 보안

이상감지로
실시간 **해킹탐지**

자원변동에 대응
보안공백 Zero

패턴분석으로
모니터링

보안설정
자동진단

국내최초, 최대규모
글로벌 **보안관제센터**

국내최초 인증
AWS Competency

해킹을 원천적으로 차단하는 암호기술

인증키 해킹차단
부채널내성암호

세계최초
전자서명 공격 원천차단

세계최고
연산속도

데이터 암호화키 해킹차단
화이트박스암호

세계최고 효율성
연산속도, 구현사이즈

모바일, IoT, 클라우드 등
50개 이상 적용

데이터를 암호화된 상태로 분석하는
동형암호

동형암호 기술의 특징

서울대학교 수리과학부 천정희 교수

암호기술의 분류

1세대 암호	2세대 암호	3세대 암호	4세대 암호
			
PASSWORD(인증기술)	**대칭키 암호**(데이터암호화)	**공개키 암호**(키 암호화)	**동형암호**(키보호 암호)
	블록암호(AES)	RSA-암호화	암호화 상태 계산이 가능한 암호

미래컴퓨팅 환경

우리가 원하는 것은 **완벽한 하인**

내가 하려는 일을 **빠르게** 대신 수행	**비밀**을 알지 못한다

사례1 개인 클라우드

스토리지 클라우드	계산 클라우드
Dropbox	DNA 계산
Google email/ calendar	헬스케어
NAS	개인성향분석 통한 추천

사례2 빅데이터 분석

개인정보기반 마케팅 (구글, 페이스북, 네이버 등)	정부 데이터베이스
	교육
	의료
	납세

동형암호 (HE : Homomorphic Encryption)

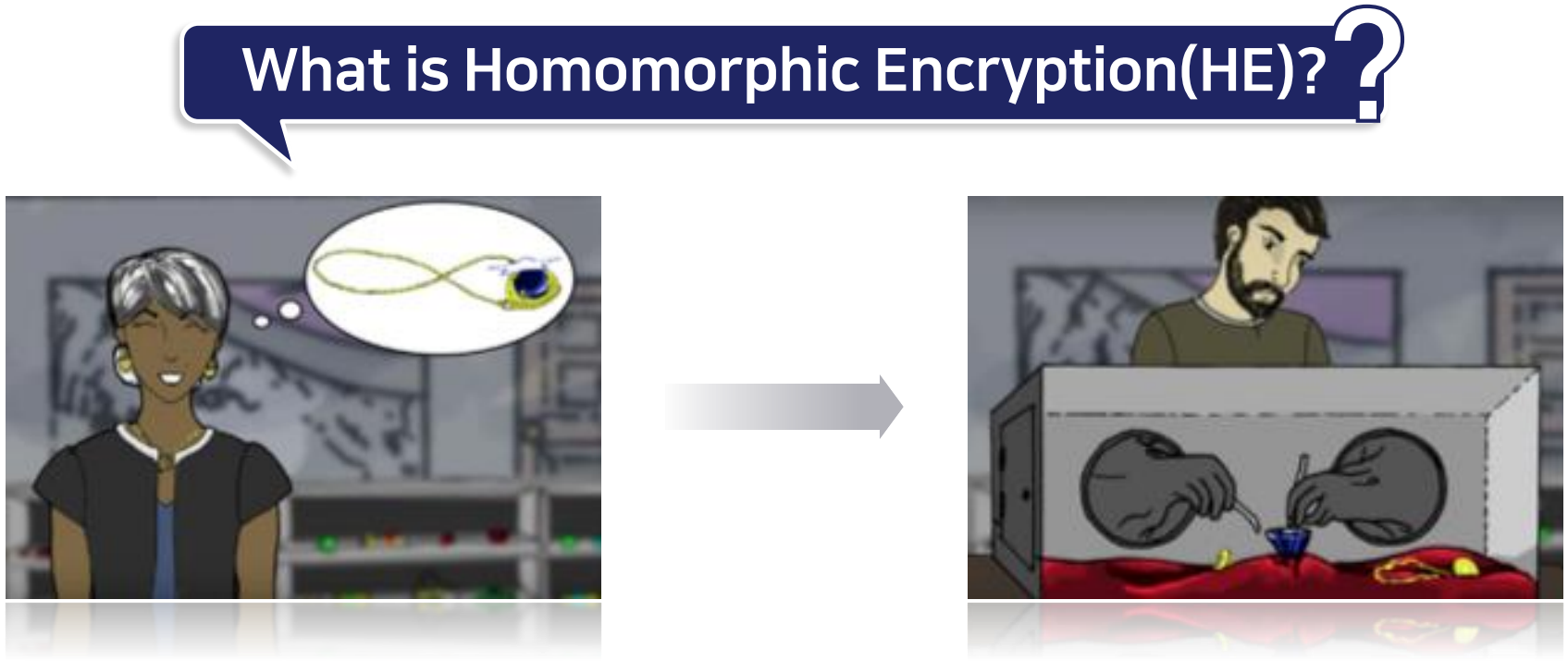
4세대 암호 : 암호화된 데이터를 복호화 없이 연산하는 암호

-10 Emerging Technologies (MIT Technical Review 2011)

Ciphering

Gentry's system allows encrypted data to be analyzed in the cloud. In this example, we wish to add 1 and 2. The data is encrypted so that 1 becomes 33 and 2 becomes 54. The encrypted data is sent to the cloud and processed: the result (87) can be downloaded from the cloud and decrypted to provide the final answer (3). Credit: Steve Moors

동형암호 (HE : Homomorphic Encryption)



Homomorphic encryption is a method of performing calculations on Encrypted data **without decryption.**

The Future of Encryption by National Science Foundation (https://youtu.be/BylWT5gsgfM)
(한글자막: http://www.math.snu.ac.kr/~jhcheon/[HD%20MP4]The%20future%20of%20encryption)

※ 출처: Youtube

동형암호의 특성

장점

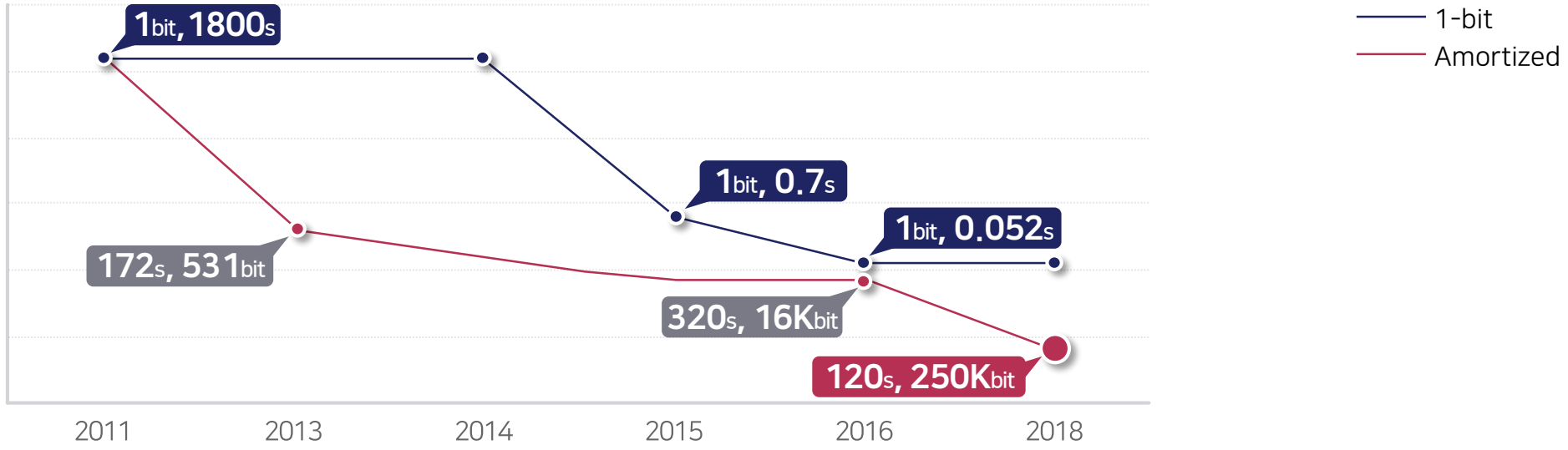
- 튜링 완전성 : 컴퓨터로 하는 모든 연산이 가능
- 암호화 후 통계처리/검색/기계학습
- 해커의 데이터 유출 원천봉쇄
- 암/복호화 속도 : 수십 ms (AES 1us, RSA 1ms)

단점

- 암호문 확장 : 수십 배
- 암호문 연산: 수백 배 (평문상태 연산대비)
- 응용연산 종류에 따른 속도의 차이가 큼

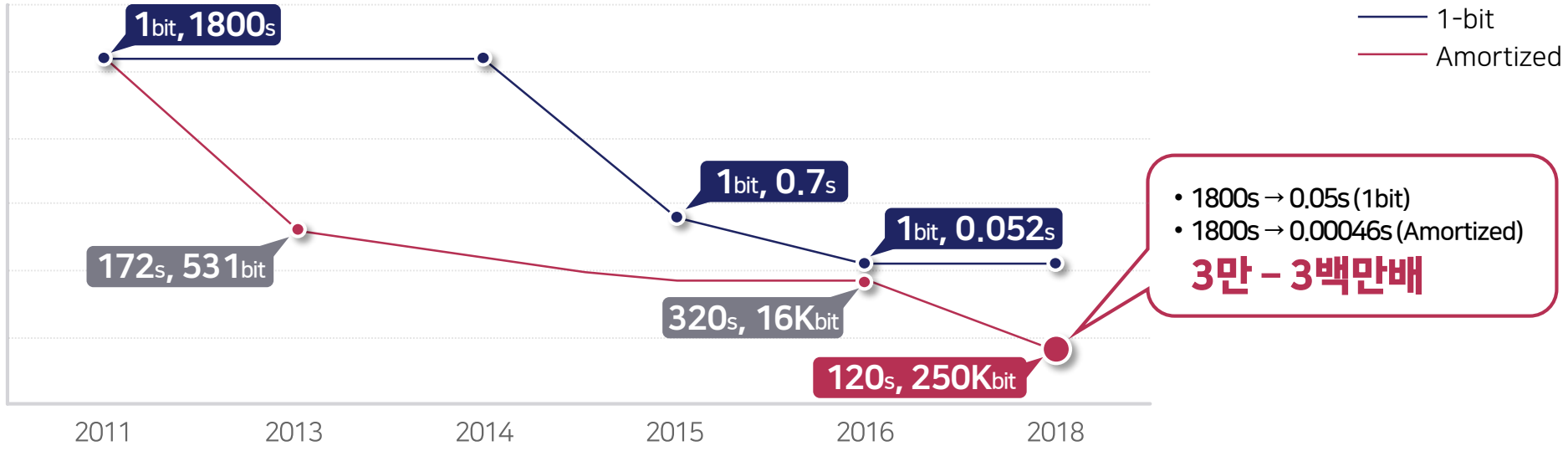
  ➡ 최적 알고리즘과 구현기술 필요

재부팅 속도로 보는 동형암호의 발전



- GH11
  - Implementing Gentry's Fully-Homomorphic Encryption Scheme, Eurocrypt 2011.
- CCK+13
  - Batch Fully Homomorphic Encryption over the Integers, Eurocrypt 2013.
- CLT14
  - Scale-Invariant Fully Homomorphic Encryption over the Integers, PKC 2014.
- HS15
  - Bootstrapping for Helib, Eurocrypt 2015.
- DM15
  - FHEW: Boostrapping Homomorpic Encryption in Less Than a Second, Eurocrypt 2015.
- CGGI16
  - Faster Fully Homomorphic Encryption: Bootstrapping in less than 0.1 Seconds, Asiacrypt 2016.
- **CHH18**
  - Faster Homomorphic Discrete Fourier Transforms and Improved FHE Bootstrapping, eprint, 1073, 2018
  - Intel Xeon CPU E5-2620 2.10GHz, 64RAM

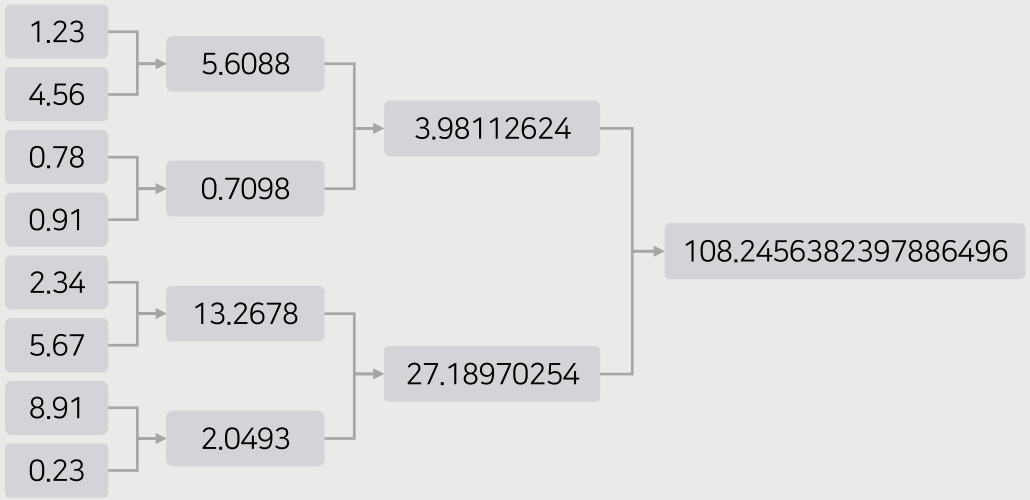
재부팅 속도로 보는 동형암호의 발전



- GH11
  - Implementing Gentry's Fully-Homomorphic Encryption Scheme, Eurocrypt 2011.
- CCK+13
  - Batch Fully Homomorphic Encryption over the Integers, Eurocrypt 2013.
- CLT14
  - Scale-Invariant Fully Homomorphic Encryption over the Integers, PKC 2014.
- HS15
  - Bootstrapping for Helib, Eurocrypt 2015.
- DM15
  - FHEW: Boostrapping Homomorpic Encryption in Less Than a Second, Eurocrypt 2015.
- CGGI16
  - Faster Fully Homomorphic Encryption: Bootstrapping in less than 0.1 Seconds, Asiacrypt 2016.
- **CHH18**
  - Faster Homomorphic Discrete Fourier Transforms and Improved FHE Bootstrapping, eprint, 1073, 2018
  - Intel Xeon CPU E5-2620 2.10GHz, 64RAM

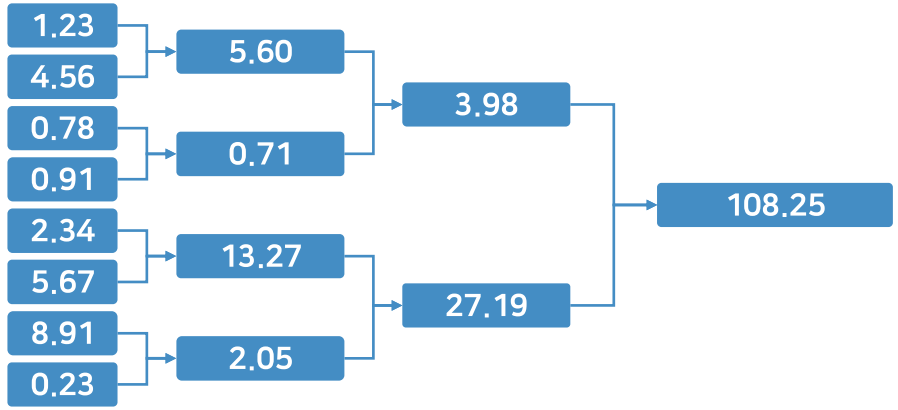
근사동형암호

AS-IS



- **평문 크기가 연산마다 2배로 증가하면 20번 후에는 백 만 비트**
  (쌀 한 톨의 비유: 한 톨로 시작하여 한달 후 $0.02g*2^{30}$ = 20톤)

TO-BE



HEAAN [CKKS17]

- **근사계산을 지원** (= 암호화된 정수값의 Re-scaling을 지원)
- **항상 같은 크기의 자릿수를 유지할 수 있음**
- **대부분의 응용분야에서는 실수계산을 필요로 함**
  (e.g 데이터 분석, 기계학습)

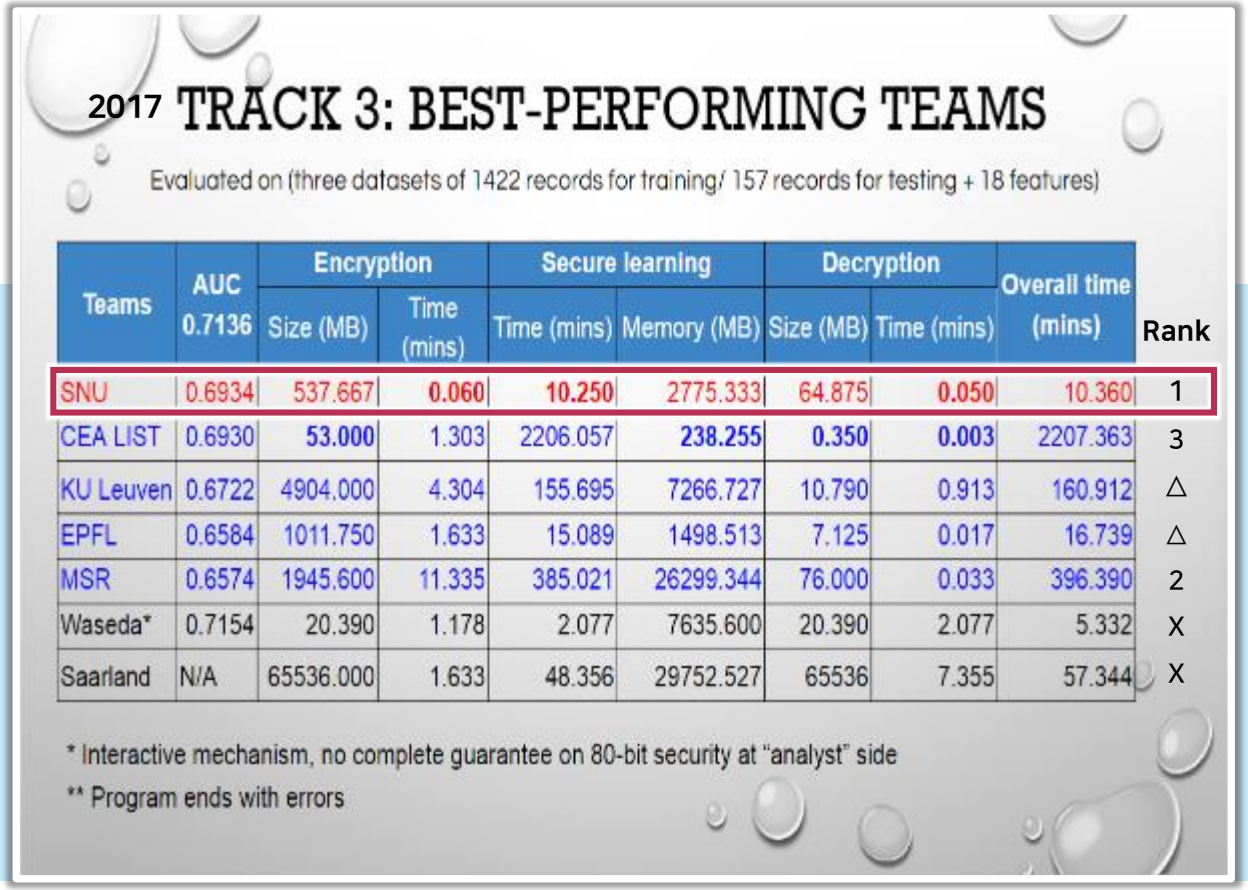
국제 게놈 보안 경진대회 (Secure Genome Analysis Competition)

iDASH

2017년 **서울대팀 우승**

근사동형암호 **혜안사용**

월등한 **계산속도**



2017 TRACK 3: BEST-PERFORMING TEAMS

Evaluated on (three datasets of 1422 records for training/ 157 records for testing + 18 features)

Teams	AUC 0.7136	Encryption		Secure learning		Decryption		Overall time (mins)	Rank
		Size (MB)	Time (mins)	Time (mins)	Memory (MB)	Size (MB)	Time (mins)		
SNU	0.6934	537.667	**0.060**	**10.250**	2775.333	64.875	**0.050**	10.360	1
CEA LIST	0.6930	**53.000**	1.303	2206.057	**238.255**	**0.350**	**0.003**	2207.363	3
KU Leuven	0.6722	4904.000	4.304	155.695	7266.727	10.790	0.913	160.912	△
EPFL	0.6584	1011.750	1.633	15.089	1498.513	7.125	0.017	16.739	△
MSR	0.6574	1945.600	11.335	385.021	26299.344	76.000	0.033	396.390	2
Waseda*	0.7154	20.390	1.178	2.077	7635.600	20.390	2.077	5.332	X
Saarland	N/A	65536.000	1.633	48.356	29752.527	65536	7.355	57.344	X

* Interactive mechanism, no complete guarantee on 80-bit security at "analyst" side

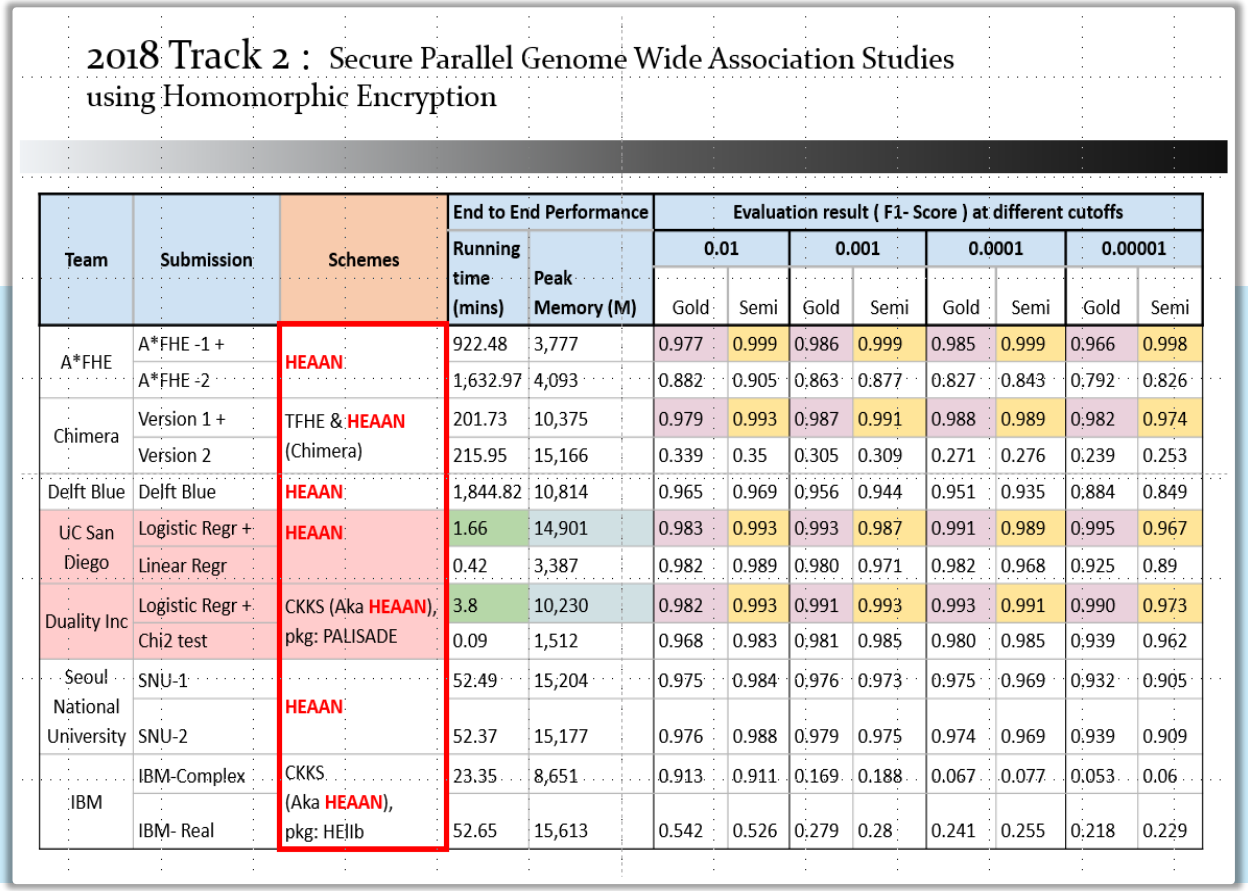
** Program ends with errors

국제 게놈 보안 경진대회 (Secure Genome Analysis Competition)

iDASH

2018 **iDASH**

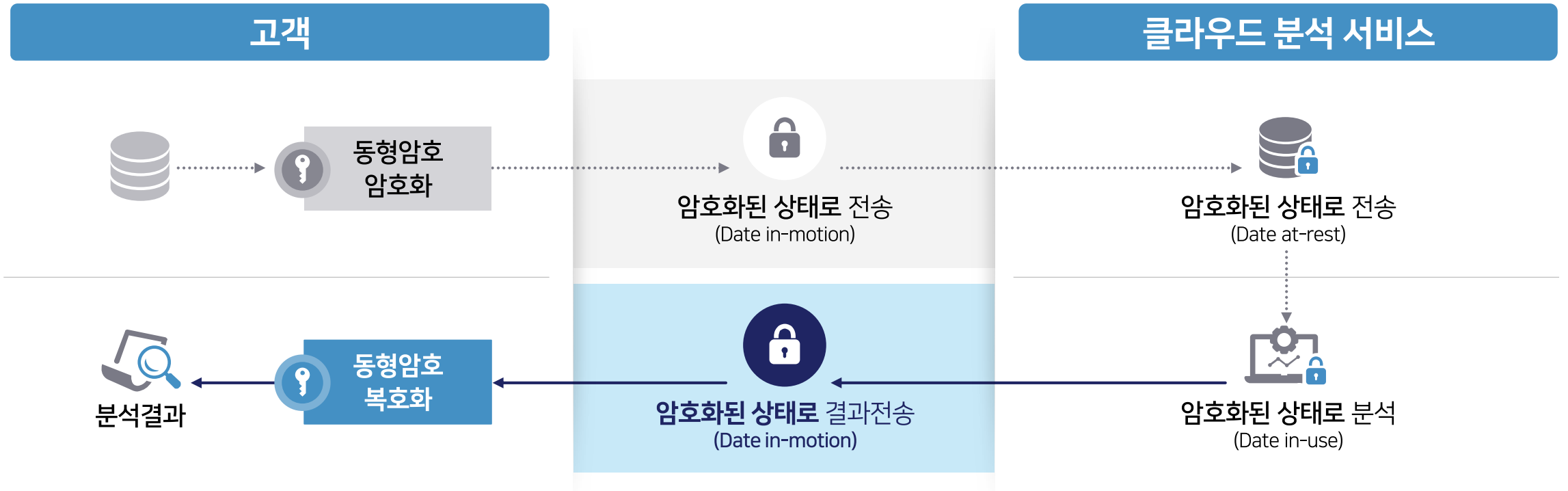
IBM등 모든 팀 **혜안사용**



2018 Track 2 : Secure Parallel Genome Wide Association Studies using Homomorphic Encryption

Team	Submission	Schemes	End to End Performance		Evaluation result (F1- Score) at different cutoffs							
			Running time (mins)	Peak Memory (M)	0.01		0.001		0.0001		0.00001	
					Gold	Semi	Gold	Semi	Gold	Semi	Gold	Semi
A*FHE	A*FHE -1 +	HEAAN	922.48	3,777	0.977	0.999	0.986	0.999	0.985	0.999	0.966	0.998
	A*FHE -2		1,632.97	4,093	0.882	0.905	0.863	0.877	0.827	0.843	0.792	0.826
Chimera	Version 1 +	TFHE & HEAAN (Chimera)	201.73	10,375	0.979	0.993	0.987	0.991	0.988	0.989	0.982	0.974
	Version 2		215.95	15,166	0.339	0.35	0.305	0.309	0.271	0.276	0.239	0.253
Delft Blue	Delft Blue	HEAAN	1,844.82	10,814	0.965	0.969	0.956	0.944	0.951	0.935	0.884	0.849
UC San Diego	Logistic Regr +	HEAAN	1.66	14,901	0.983	0.993	0.993	0.987	0.991	0.989	0.995	0.967
	Linear Regr		0.42	3,387	0.982	0.989	0.980	0.971	0.982	0.968	0.925	0.89
Duality Inc	Logistic Regr +	CKKS (Aka HEAAN), pkg: PALISADE	3.8	10,230	0.982	0.993	0.991	0.993	0.993	0.991	0.990	0.973
	Chi2 test		0.09	1,512	0.968	0.983	0.981	0.985	0.980	0.985	0.939	0.962
Seoul National University	SNU-1	HEAAN	52.49	15,204	0.975	0.984	0.976	0.973	0.975	0.969	0.932	0.905
	SNU-2		52.37	15,177	0.976	0.988	0.979	0.975	0.974	0.969	0.939	0.909
IBM	IBM-Complex	CKKS (Aka HEAAN), pkg: HElIb	23.35	8,651	0.913	0.911	0.169	0.188	0.067	0.077	0.053	0.06
	IBM- Real		52.65	15,613	0.542	0.526	0.279	0.28	0.241	0.255	0.218	0.229

클라우드 기반 데이터 분석: 삼성SDS-서울대학교



Key Value | 고객의 데이터는 고객의 손을 떠난 시점부터 분석 결과가 되어 다시 고객 손에 돌아올 때 까지 어떤 순간에도 복호화되지 않는다

개인정보를 보호하는 데이터분석

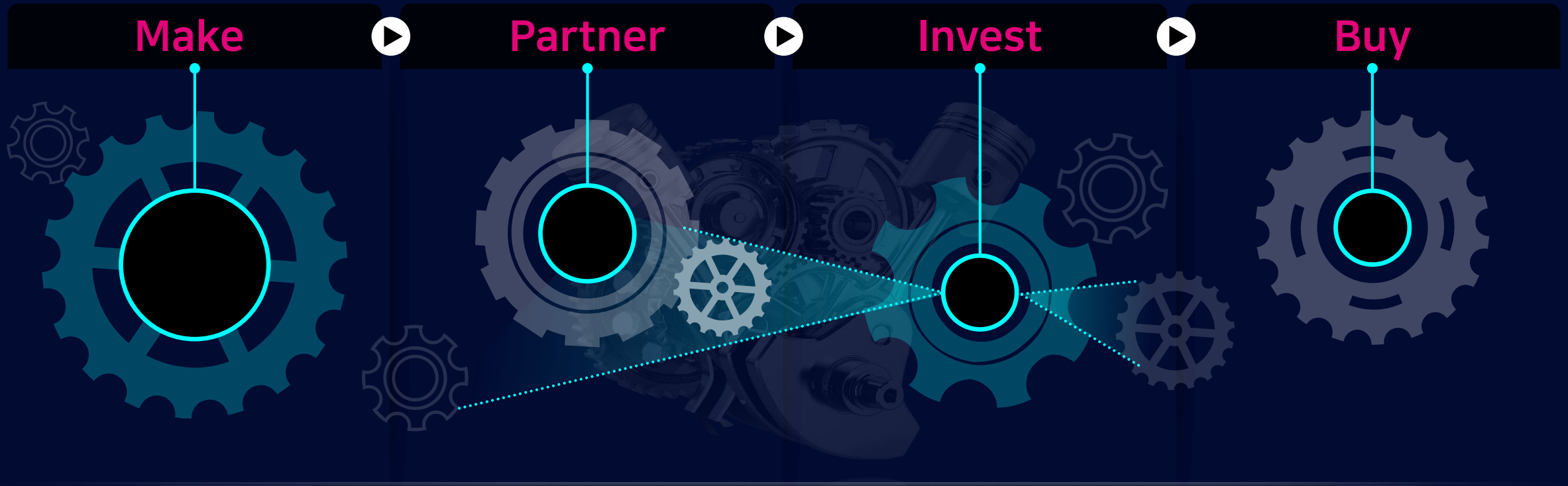
데이터 분석
알고리즘

적용 분야

의료 • • • 신용 • • • 마케팅

Technology Innovation

삼성SDS 이노베이션 프레임워크



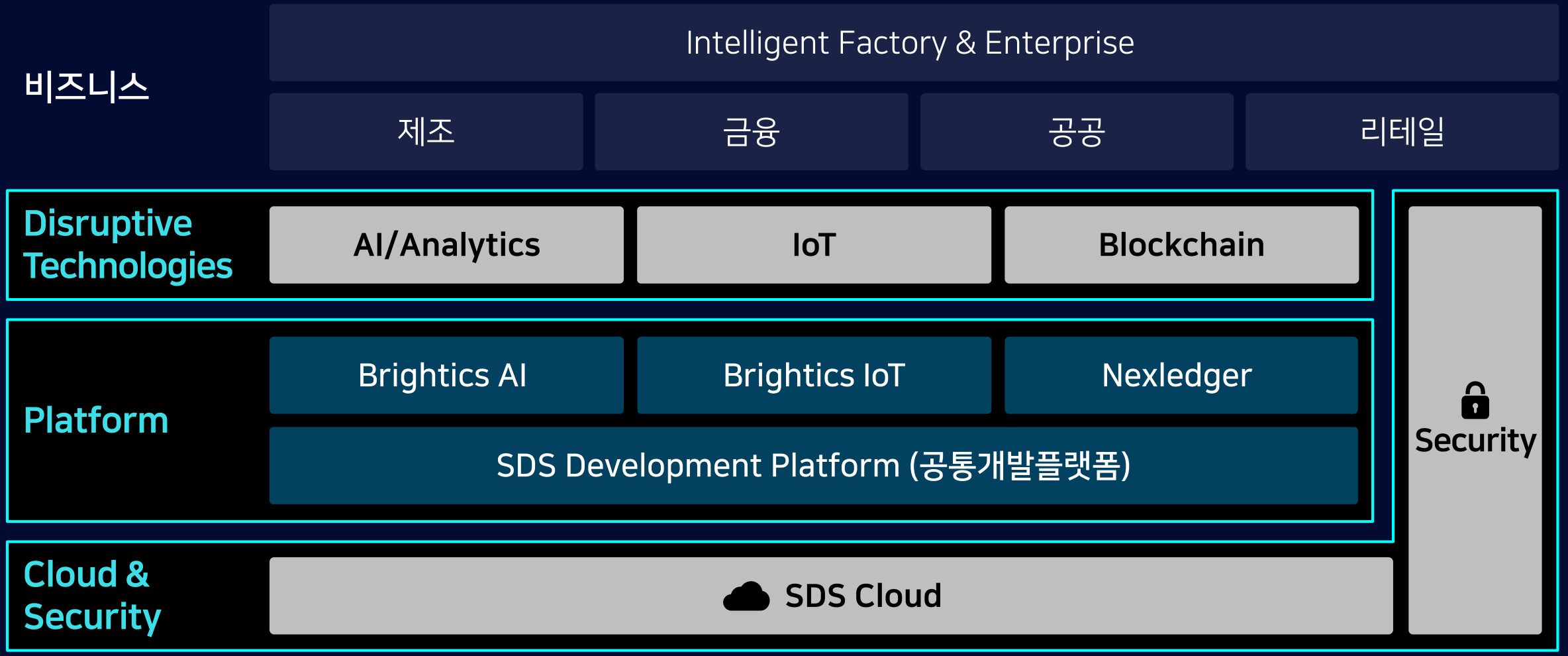
- SW엔지니어
- 신사업 아이디어 공모
- 오픈소스 커뮤니티

- 파트너십
  (솔루션도입, 산학협력 등)

- 네트워크, 보안 등 투자

- 제조, 보안관련

삼성SDS Enterprise Platform (SEP)



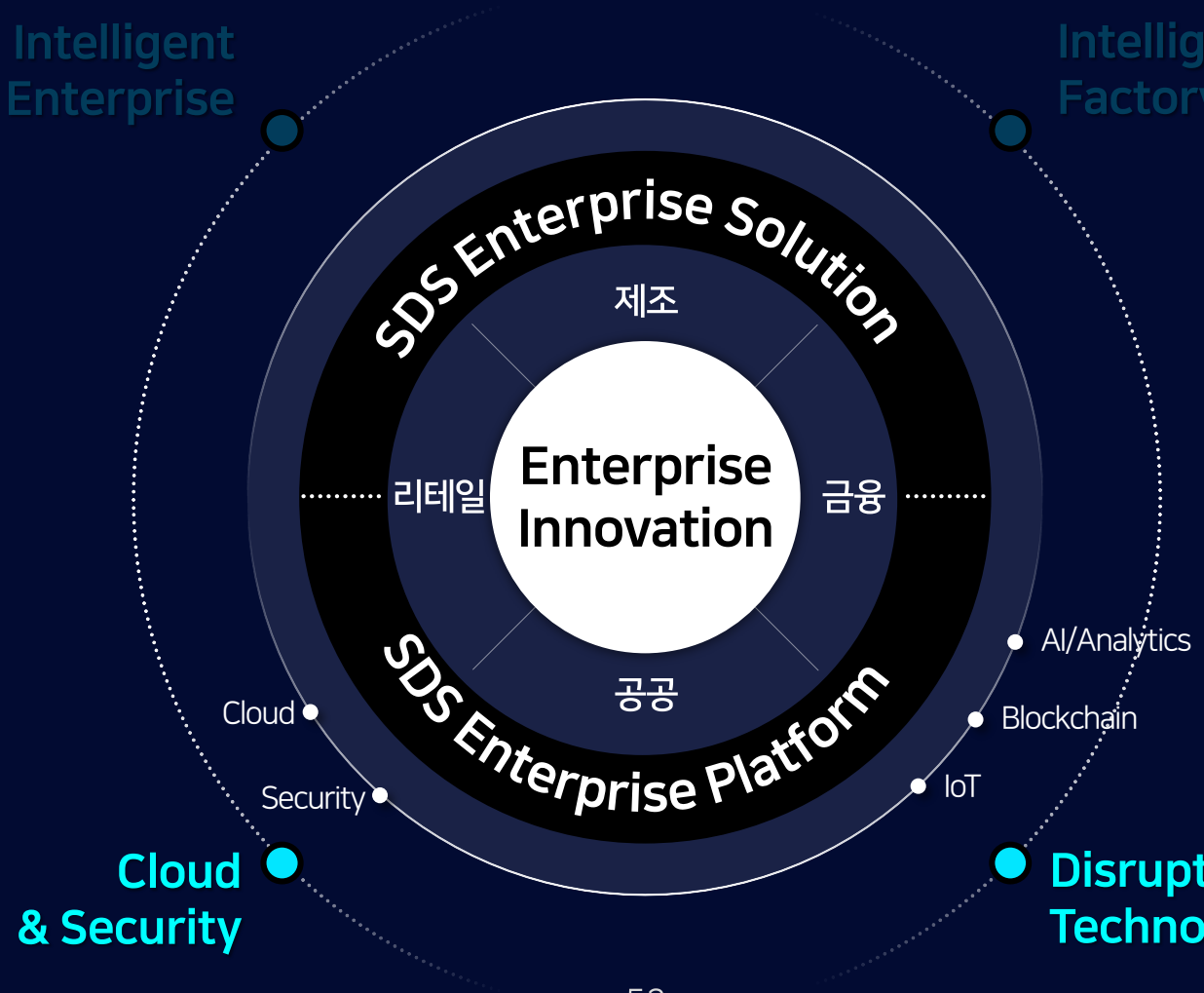
삼성SDS의 Digital Transformation Framework

Cloud

Migration
Hybrid Cloud
MSA / 컨테이너

Security

악성코드 분석
AI기반 이상탐지
해킹위협 대응



AI/Analytics

제조/마케팅/헬스사례
AI기반 고객상담

Blockchain

금융/제조사례
가속화 기술

IoT

제조/건설사례
Data Preparation

Thank you