CLOUD ON
Azure Managed Service

CloudON for Data Service : 6-Wk PoC

데이터 활용 방안 제안서 (요약본)

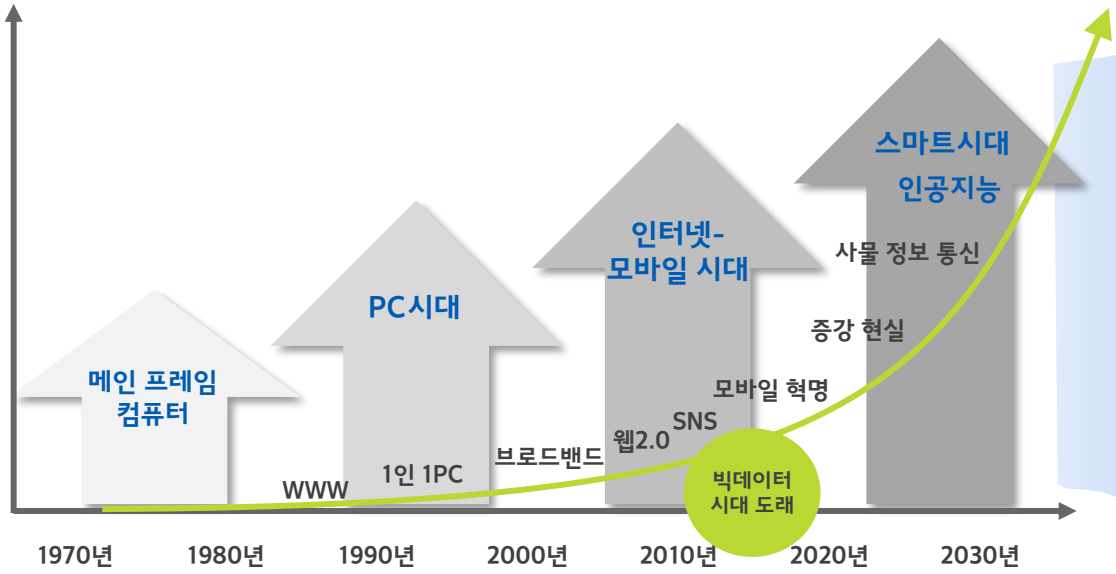
Data digital transformation using Azure Data Service

2023.08

데이터 분석 환경 필요성

CLOUD ON
Azure Managed Service

시대적 흐름에 따라 **다양한 데이터의 수집 및 분석의 필요성** 증대 되고 있습니다.



새로운 데이터 소스

- 센서, 장비, 기존응용
  웹서버, 공공 데이터 등

실시간 처리

- 데이터 발생속도 가속화
- 스트림 데이터 폭증
- 신속한 데이터 처리

대용량 데이터 규모

- 수백 Tera Byte 에서
  수십 Peta Byte

신경제학

- 초저가 대규모 프로세싱
- 분석 클라우드 컴퓨팅

다른 데이터 형식

- 비정형 데이터
- 약한 관계형 스키마
- 텍스트, 이미지, 비디오, 로그

새로운 형태의 질문과 통찰

- 이 제품은 얼마나 인기있나?
- 가장 뛰어난 광고는?
- 이 거래는 사기가 아닐까?

데이터 분석 환경 구축

효율적 의사결정	업무 효율성 향상	비즈니스 리스크 관리	타깃 이해도, 만족도 증가	마켓팅 최적화
인사이트 도출 예상 결과 구체화 신속한 의사 결정	운영 최적화 주요 지표(KPI) 추적 및 개선 비즈니스 실시간 모니터링	잠재적인 리스크 관리 개선 영역의 빠른 발견 미래예측 및 대응	구매패턴 선호도 분석 고객 충성도 만족도 향상	마케팅 성과분석 이탈율, 조회수 분석 잠재고객 타겟팅

분야별 빅데이터 활용 예시

다양한 분야에서 빅데이터 분석을 통하여 비즈니스의 개선과 새로운 인사이트를 발굴하고 있습니다.

공공

- 위험 예측 및 예방 (범죄 등)
- 복직 프로그램
  (운영비용절감, 적격자선정)
- 지능 교통 통제 관리
  (자동 텔레메틱스)
- 산불 및 물 자원관리
- 트레킹 서비스 및 자연재해 예측
  (재난안전, 예방관리, 안전지원)

다양성, 공평성 갖춘 국가현안의 효율적 개선

통신

- SNS 분석, 고객 이탈 방지,
  위치서비스 연동 등
- 마케팅 활동
  (고객 유치, Next Action 예측)
- 실시간 다중 감지 및 상황 인지
- 사이버 보안 감지
- 타 분야 시스템 실시간 연계
- 개인화 맞춤형 광고 제공 연계

다양한 분야와의 사업 연계 효과적인 마케팅

의료

- 환자 치료 품질 및 효율성 향상
- 고객 참여 네트워크 의료 서비스
  구축
- 개인 맞춤식 진단
- 유행병 조기진단 및 치료비 절감
- 원격 헬스케어 모니터링
- 유전자 연구를 통한 신약 개발

국민 건강 보호 및 증진 맞춤형 의료 서비스

제조

- 스마트 팩토리
  (고장 예측, 수율 향상, 에너지절감)
- 예측적 유지보수 및 품질향상
- 수요예측 및 재고관리
- AI기반 MES 센서 모니터링 분석
  및 운전 자동화

산업적 AI 채택으로 제조 현장 역량 강화

금융

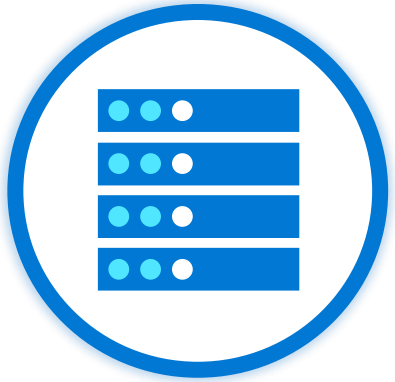
- 대고객 서비스 최적화
  (신규 상품 개발, 콜센터 개선 등)
- 부정거래 탐지 및 신용 평가
- 실시간 사기 방지 모니터링
- 주가에 대한 날씨 영향도 분석
- 시장 및 상권 데이터 분석
- 전사적 리스크 관리 활용

ESG 노력 강화 차세대 금융산업 도약

빅데이터 분석 플랫폼

CLOUD ON
Azure Managed Service

투자 비용 절감과 손쉬운 확장이 가능한 Cloud 분석 플랫폼 도입으로 비즈니스의 유연성 확보



On-Premise

Cloud

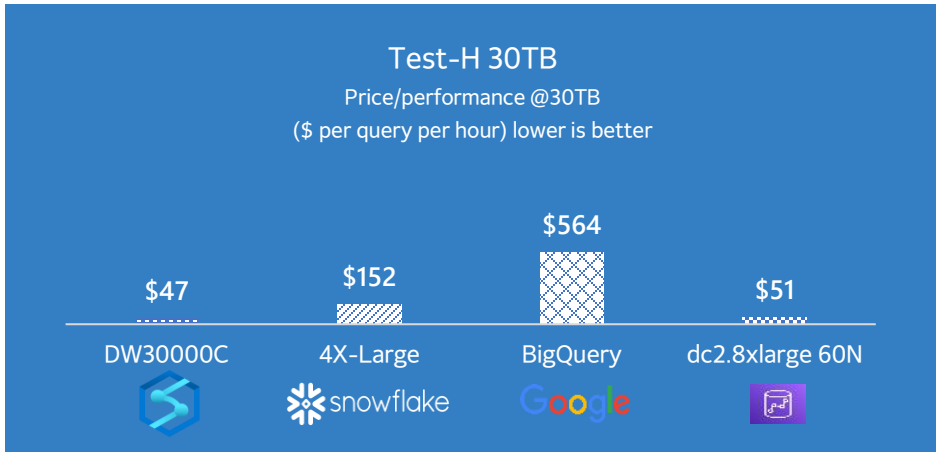
On-Premise		Cloud
하드웨어와 소프트웨어 구매비용 + 유지보수 비용	비용	사용량에 따른 종량제 가격
최대용량 기준으로 비용 산정, 비효율성	도입 효율성	최소의 비용으로 도입, 단계적인 증설
시스템 확장시에 S/W 신규 설치, 확장 및 축소 어려움	확장 편의성	손쉽게 확장 및 축소 가능
S/W 신규 릴리스, 기능 추가시 사용자 Upgrade 수행	업그레이드	주기적인 자동 업그레이드 지원
하드웨어, 소프트웨어를 각 IT 담당자 유지관리	유지보수	공급업체가 관리, 관리자는 업무에 집중
IT 팀에서 가용성, 보안, 성능 모니터링과 튜닝을 수행	모니터링/보안	IT팀과 CSP는 보안 및 관리의 책임분담
3rd Part 솔루션 추가 구매 및 유지보수 필요	백업 복구	완전자동화 백업복구, 3rd Part 솔루션 불필요

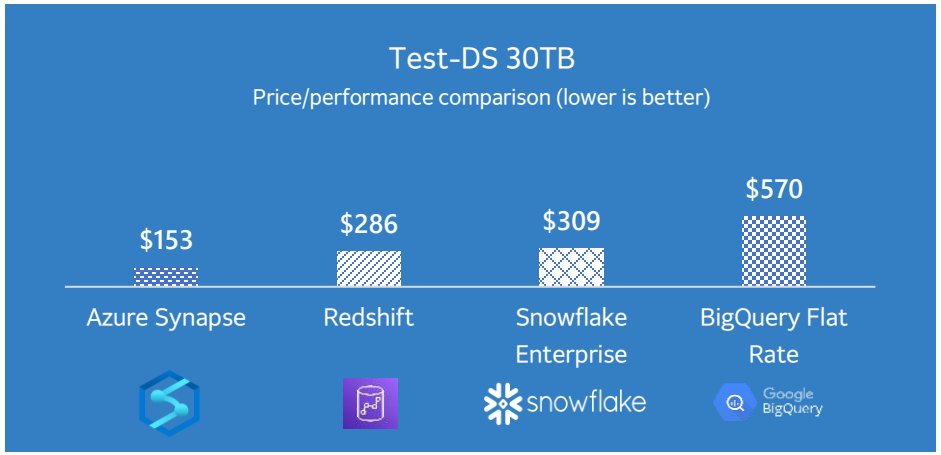
검증된 성능과 확장성

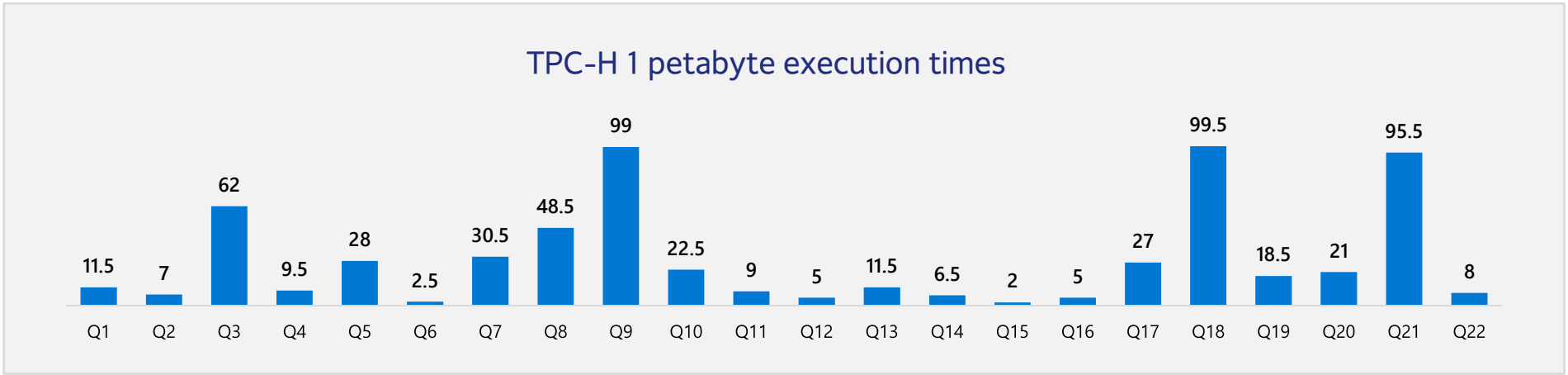
CLOUD ON
Azure Managed Service

타 클라우드 데이터 분석서비스에 비해 가격 대비 성능 우위 검증 완료

Azure Synapse Analytics는 1PB 규모에서 TPC-H를 성공적으로 완료한 유일한 쿼리 엔진입니다.







데이터 분석 플랫폼 제안

미래 지향적 Azure Cloud 데이터 분석 플랫폼 고도화를 통한 데이터 활용 기반 확보

Azure Cloud 데이터 분석 플랫폼을 이용하여 데이터 인사이트 / 비즈니스 / 융합에 이르는 서비스 구축

1 유연한 데이터 분석

- 기존의 정형분석과 더불어 최신 오픈소스 분석 언어인 R, Python을 사용한 자유로운 분석 토대 마련
- 어느 레벨의 사용자라도 분석을 수행할 수 있는 유연한 환경의 분석 플랫폼으로 다양한 비즈니스 데이터 분석 능력 함양

2 데이터 비즈니스 대응

- 각 사용자의 아이디어를 활용할 수 있도록 구축하여 인사이트 극대화
- 결과를 곧바로 사용할 수 있는 플랫폼을 구축하여 비즈니스 연결성 확보
- 가공된 결과를 저장하고 추후 데이터 비즈니스를 시작할 수 있는 기반 마련

3 비용 절감 및 운영/관리 효율 증가

- 분석 인프라 확장 시 TCO 절감 필요
- 플랫폼 및 소프트웨어 서비스를 적극 활용하여 운영 및 관리 비용 절감
- 향후 증설 혹은 확장에 대비한 단순한 사용 기반 비용 정책

"분석 유연성, 비즈니스 민첩성, 관리 효율성을 제공할 수 있는 분석 플랫폼 구축"

Data Lake 구축 로드맵

CLOUD ON
Azure Managed Service

데이터 기반의 분석플랫폼 구축 및 단계별 운영 계획 로드맵

1. POC 수행 (파일럿)

- 비즈니스의 데이터 타당성, 운영/관리 비용 개선 효과 점검
- 새로운 비즈니스의 방향성 및 **도입 리스크 사전 검토**
- 데이터 보호를 위한 **보안기능 및 정책의 유효성 점검**
- 데이터 처리 및 활용 방식의 영향도 점검

2. 플랫폼 및 센터 구축/연계

- 분석 플랫폼 및 센터 구축/연계
- **데이터 품질 관리체계 구축**
- 수요조사 및 기초 데이터 구축
- 비개방 샌드박스 서비스 구축
- **혁신서비스 아이템 발굴 및 선정**
- **인력양성 프로그램 개발**
- 데이터 거래시스템 구축

3. 안정화 및 데이터 확대

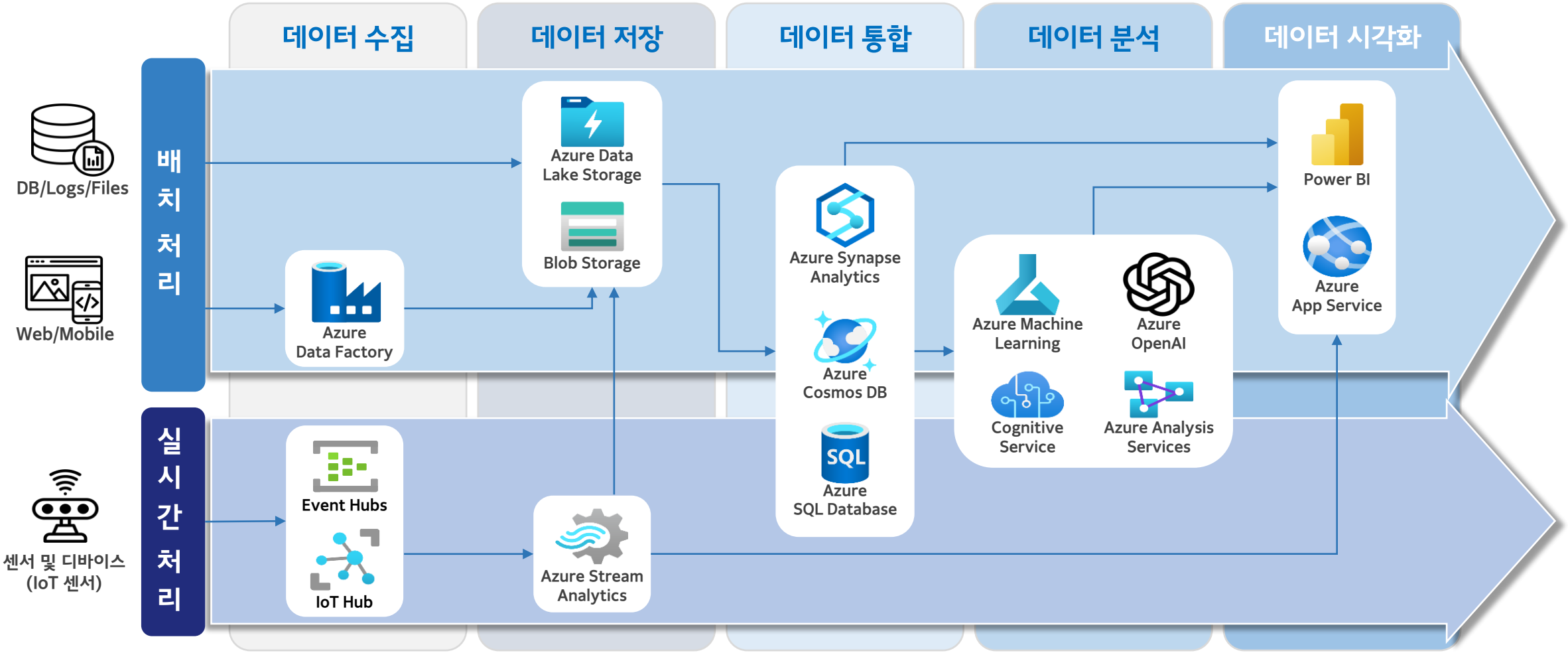
- 플랫폼 및 분석센터 고도화
- **융합데이터 구축 확대**
- 혁신서비스 개발 (신규)
- **수요업무 및 참여인력 확대**
- 샌드박스 서비스 고도화
- 인력양성 프로그램 운영
- **데이터 거래시스템 확대**

4. 서비스 확산 및 업무 자립 전환 준비

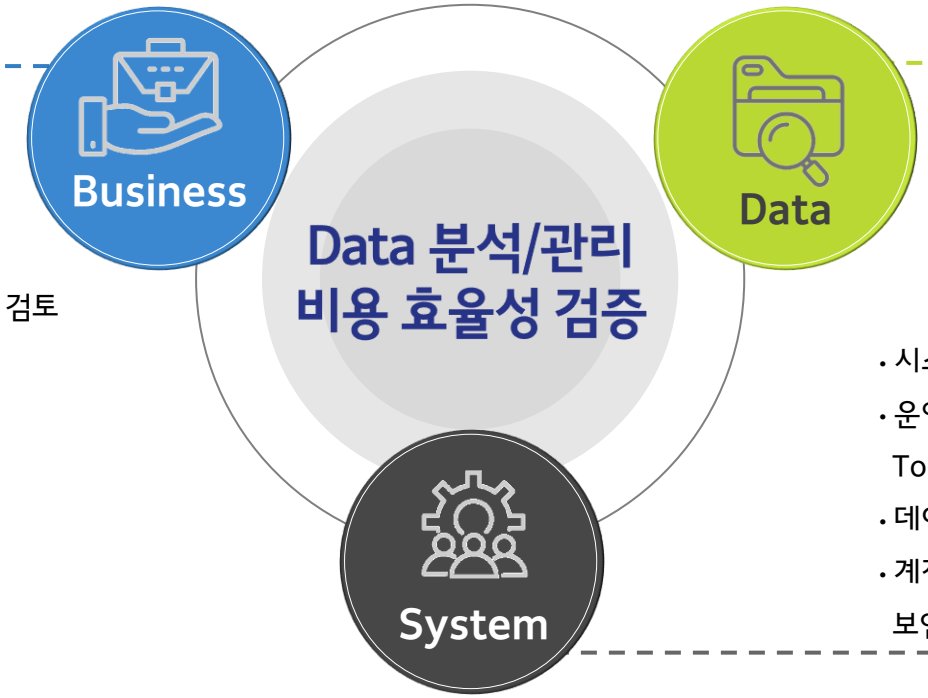
- **혁신서비스 개발 (+@)**
- 수요업무 및 참여인력 확대
- **비즈니스 연동 서비스 확대**
- 샌드박스 서비스 오픈
- 인력양성 프로그램 운영 (계속)
- **업무적 자율 분석 프로세스 수립**

빅데이터 분석 플랫폼

Azure Cloud 데이터 분석 플랫폼은 데이터 수집, 저장, 분석, 활용 등 다양한 서비스들을 제공합니다.



PoC 프로그램 소개



- PoC를 통한 잠재적인 문제 조기 식별
- 비즈니스 개선 및 요구사항 충족도 검증
- 비즈니스 요건 별 데이터 타당성, 운영 및 관리 비용 개선 효과 점검
- 새로운 비즈니스의 방향성 및 리스크 사전 검토
- 비지니스 혁신 및 시장 경쟁력 강화

- Data Architecture 점검
- 데이터 요건 충족도 점검
- 데이터 처리 및 활용 방식의 영향도 점검
- 데이터 관리 체계 개선

- 시스템 Architecture 점검
- 운영관리, 개발, 모니터링, 시각화 등 Tool의 기능/편의성 점검
- 데이터 적재 및 조회 성능 점검
- 계정관리, 접근제어 등 데이터 보호를 위한 보안기능 및 정책의 유효성 점검

다양한 분석 시나리오 검증

- 실시간, 배치분석 시나리오 등 다양한 비즈니스 시나리오에 적합한 새로운 데이터 분석 서비스를 검증
- 기존에 정형분석과 더불어 최신 오픈소스 분석 언어인 R, Python, Spark를 사용한 자유로운 분석환경 토대 마련
- Scikit Learn, TensorFlow, Keras 등 머신러닝과 딥러닝에 필요한 다양한 분석 기법과 모델 활용 및 검증

분석 서비스 성능 / 기능 검증

- 동일 노트북 내에서 Python, R, Scala, SQL을 지원하여 사용 편리성 증대, Data 분석 테스트 및 성능 검증
- 데이터 엔지니어, 사이언티스트의 협업 환경 편리성 검증
- 유연한 환경의 분석플랫폼으로 다양한 서비스 타입의 업무 연계성 최적화 테스트
- 데이터 보호 및 관리 정책의 유효성 검증

비용절감 및 운영 / 관리 효율성 검증

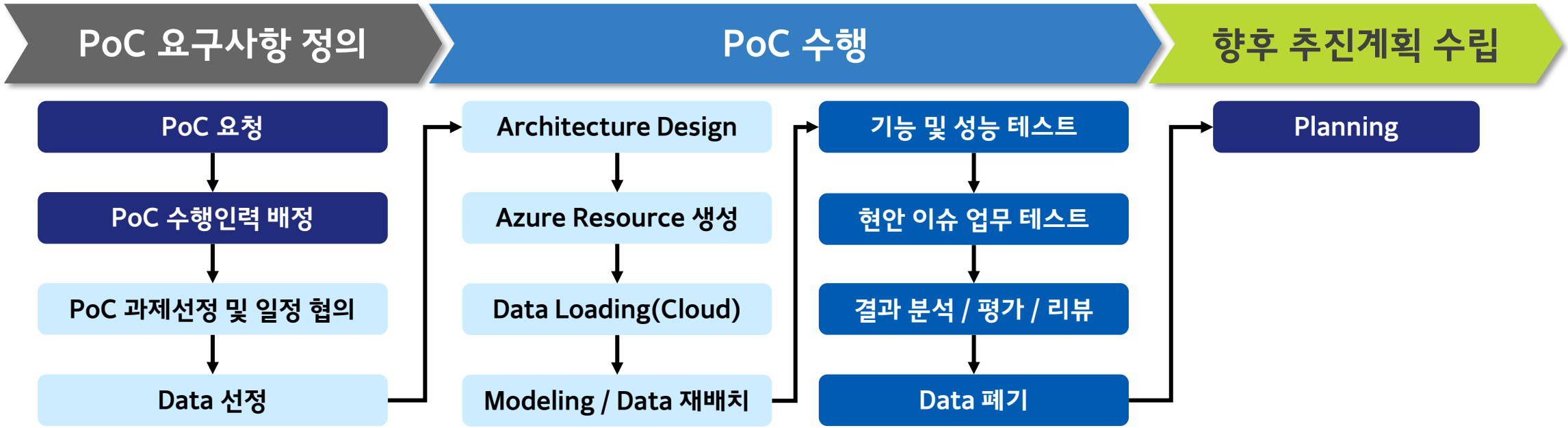
- Cloud기반 분석 서비스 구축의 예상비용 산출 및 TCO 절감을 위한 최적화 아키텍처 검증
- 향후, 시스템 증설 or 확장에 대비하여 사용량 기반 비용 관리 계획 사전검증

PoC 수행 절차 및 단계별 고려사항

CLOUD ON
Azure Managed Service

분야별 전문인력을 투입하여, 체계적이고 계획적인 PoC를 수행합니다.

각 단계별 전문인력을 배정하여 기술지원, 데이터 분석, 교육, 구축 전략 수립의 노하우를 전달합니다.



- 업무, 고객, 기술적 관점을 모두 고려한 수행과제 선정
- 명확한 수행과제와 평가방법 및 Source Data 선정 필수

- 분석 담당자를 위한 Tool 제공
- 대용량 데이터 최적화 및 분석 노하우 전달
- Data 분석 가이드 및 방법론 지원
  - ✓ 비즈니스/산업 별 특화 분석 모델 및 분석 방법
  - ✓ 레퍼런스 기반 검증된 분석 방법 활용

- Road Map 수립
  - ✓ 분석, 구축 전략 수립
- To-Be 시스템 관리, 운영비용 예상 산출
- 도입 의사결정

PoC 수행 일정 및 시나리오

CLOUD ON
Azure Managed Service

PoC를 수행하기 위한 예상 일정 및 작업을 아래와 같이 단계별로 진행합니다.

단계	작업내역	수행 일정					
		1W	2W	3W	4W	5W	6W
요구사항 정의	• PoC 수행과제 선정 • PoC 수행인력 배정 • PoC 일정협의						
시나리오 구성	• PoC 수행 데이터 선정 • 아키텍처 구성 방안 • Cloud 리소스 생성 • 보안 및 네트워크 설정						
테스트 환경준비	• 초기/전체 데이터 이관 및 적재 • 데이터 속성(스키마) 분석 • 데이터 모델링 수행						
테스트 및 검증	• PoC 기능 / 성능 테스트 • 비즈니스 시나리오 테스트 • 모니터링 및 기능 최적화						
결과 리포팅	• PoC 결과 리포트 작성 • PoC 결과 분석 / 평가 • PoC 리뷰 및 결과보고서 제출						

※PoC 검증대상데이터의양과기간에따른비용(Azure 및서비스)은다르게발생할수있습니다.

PoC 결과 리포트

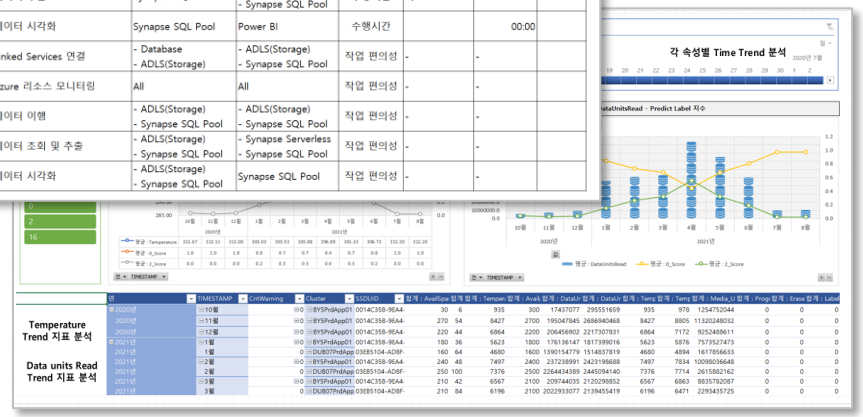
CLOUD ON
Azure Managed Service

PoC 요구사항으로 선정된 과제에 맞는 결과 샘플 리포트를 아래와 같이 제공합니다.

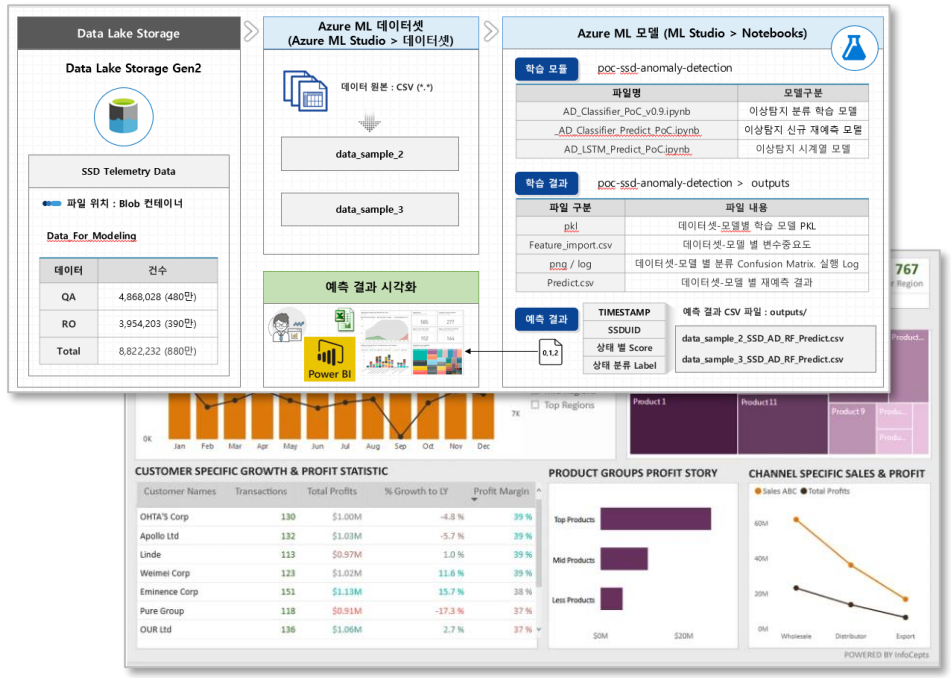
성능/기능 평가표

[성능 / 기능 종합 평가표]

구분	요구사항	Input	Output	평가항목	데이터크기	수행시간	비고
성능	초기데이터 이행	File	ADLS(Storage)	처리속도	약 500M	00:00	
	전체데이터 이행	Database	ADLS(Storage)	처리속도	약 10G	00:00	
	데이터 파티션, 분산처리	Synapse SQL Pool	Synapse SQL Pool	수행시간		00:00	
	데이터 조인	Synapse SQL Pool	Synapse SQL Pool	수행시간		00:00	
	데이터 그룹연산	Synapse SQL Pool	Synapse SQL Pool	수행시간		00:00	
	데이터 조회	Synapse SQL Pool	Synapse SQL Pool	수행시간		00:00	
	데이터 추출	Synapse SQL Pool	- ADLS(Storage) - Synapse SQL Pool	수행시간	약 500M	00:00	
	데이터 시각화	Synapse SQL Pool	Power BI	수행시간		00:00	
기능	Linked Services 연결	- Database - ADLS(Storage)	- ADLS(Storage) - Synapse SQL Pool	작업 편의성	-	-	
	Azure 리소스 모니터링	All	All	작업 편의성	-	-	
	데이터 이행	- ADLS(Storage) - Synapse SQL Pool	- ADLS(Storage) - Synapse SQL Pool	작업 편의성	-	-	
	데이터 조회 및 추출	- ADLS(Storage) - Synapse SQL Pool	- Synapse Serverless - Synapse SQL Pool	작업 편의성	-	-	
	데이터 시각화	- ADLS(Storage) - Synapse SQL Pool	Synapse SQL Pool	작업 편의성	-	-	



분석결과 리포트



CLOUD ON
Azure Managed Service

Thank you

Company : 씨앤토트(주)
Homepage : https://www.cnthoth.com
E-mail : cloudon@cnthoth.com

Copyright 2023. C&Thoth All Rights Reserved