

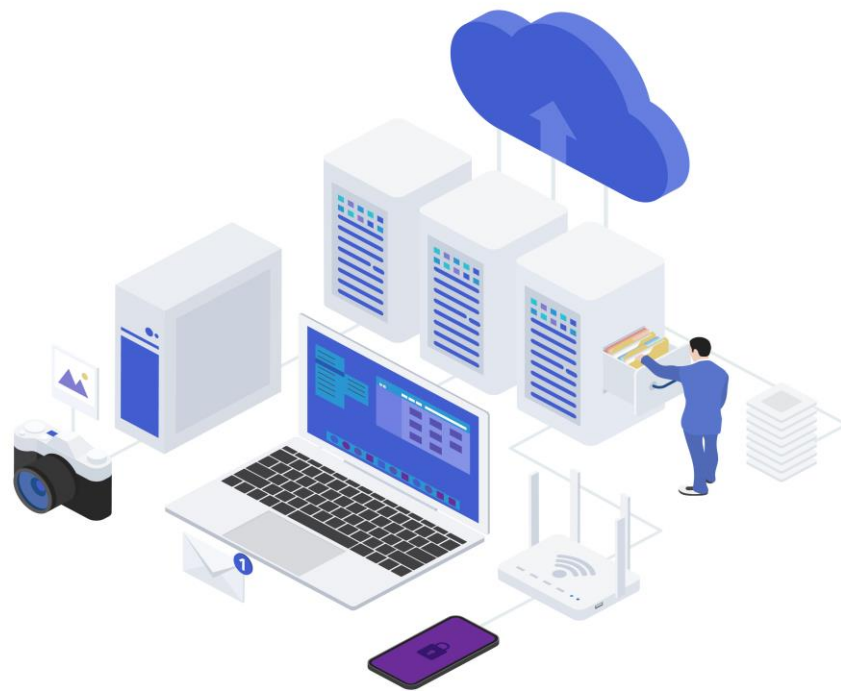
기업 데이터 분석 혁신을 위한

# Tarantula Lakehouse 소개

v3.2

UN°NET

© 2026 UNNET. 무단 복제를 금합니다. | All Rights Reserved.



# Table of Contents

## 1

### 솔루션 소개

- 1.1 타란툴라 Data Platform 제품 포트폴리오
- 1.2 타란툴라DB 소개
- 1.3 타란툴라 Lakehouse
- 1.4 Minio AIStor



## 2

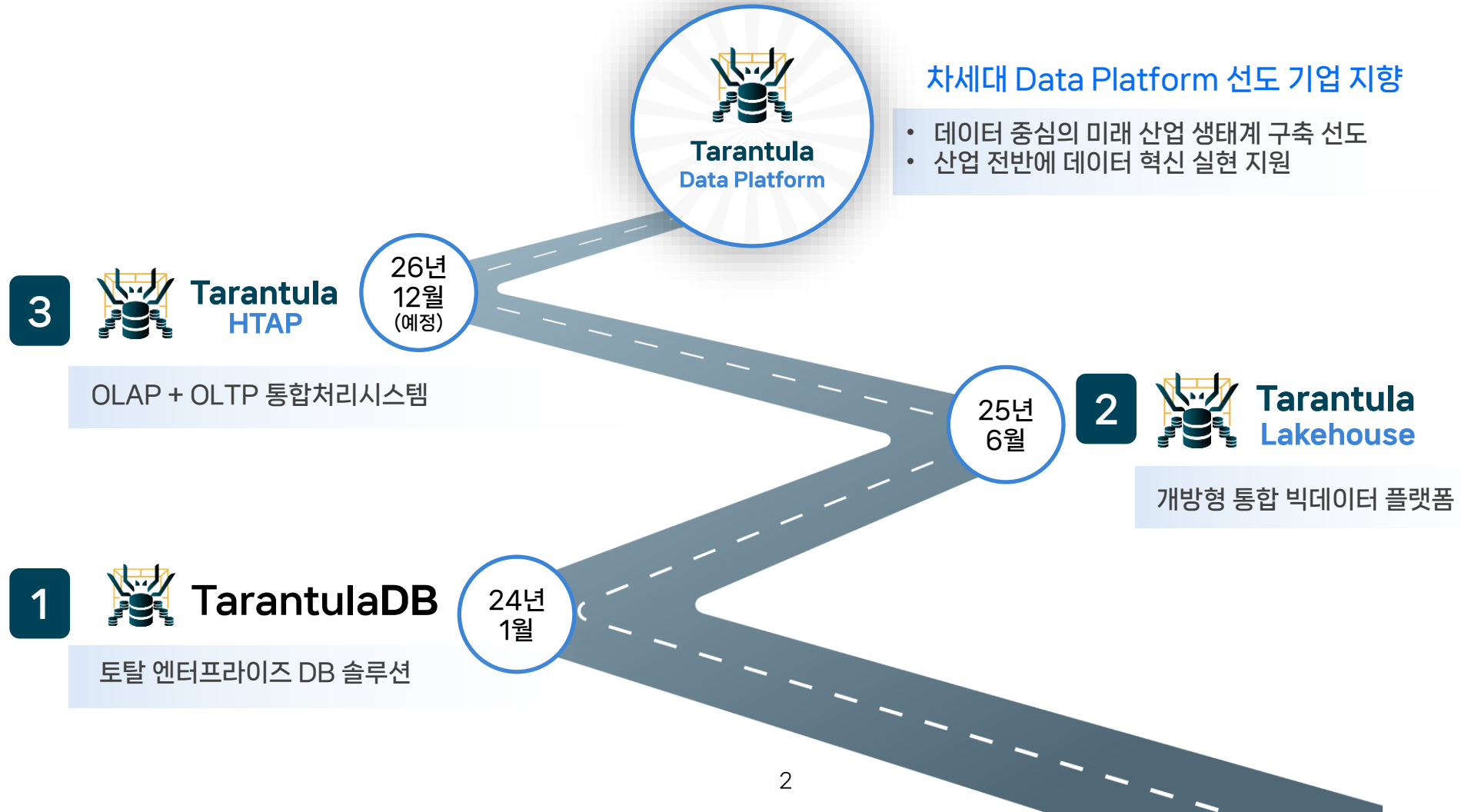
### 유엔넷 소개

- 2.1 회사소개 및 연혁
- 2.2 조직도 및 인력현황
- 2.3 사업 영역

## 1. 솔루션 소개

### 1.1 타란툴라 Data Platform 포트폴리오

타란툴라 Data Platform은 타란툴라DB 및 Lakehouse, HTAP(개발예정)에 이르는 유엔넷의 데이터 솔루션 브랜드



## 1.2.1 타란툴라DB 소개

PostgreSQL을 기반으로 기업에서 요구하는 핵심 기능과 국내 최고 기술 지원을 제공하는 '토탈 엔터프라이즈 DB 솔루션'

### 기반

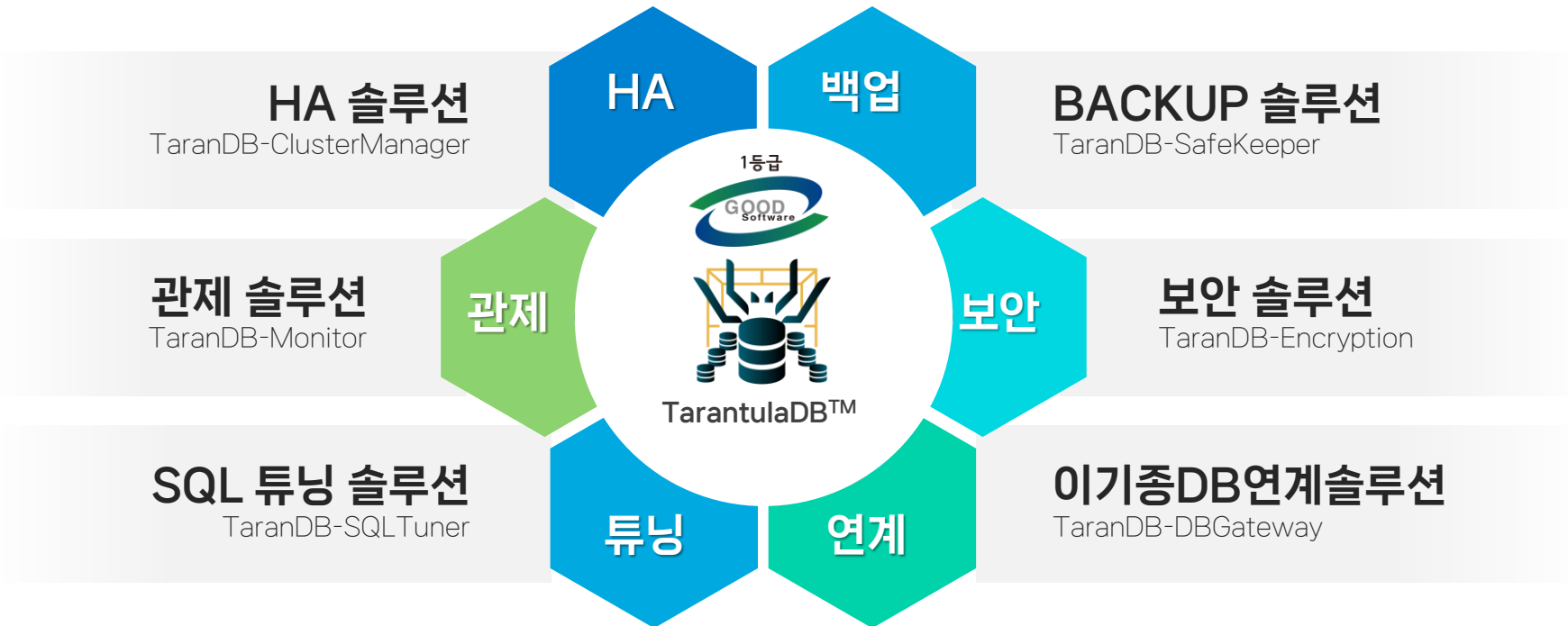


가장 빠른 성장세를 보인 DBMS

25년 글로벌 전체  
DBMS 시장 4위  
(오픈소스 중 1위)

글로벌 개발자 선호도 1위 DBMS

### 핵심 기능을 추가한 통합 패키지



## 1.2.2 타란툴라DB 레퍼런스

### 금융 부문



### 엔터프라이즈 부문

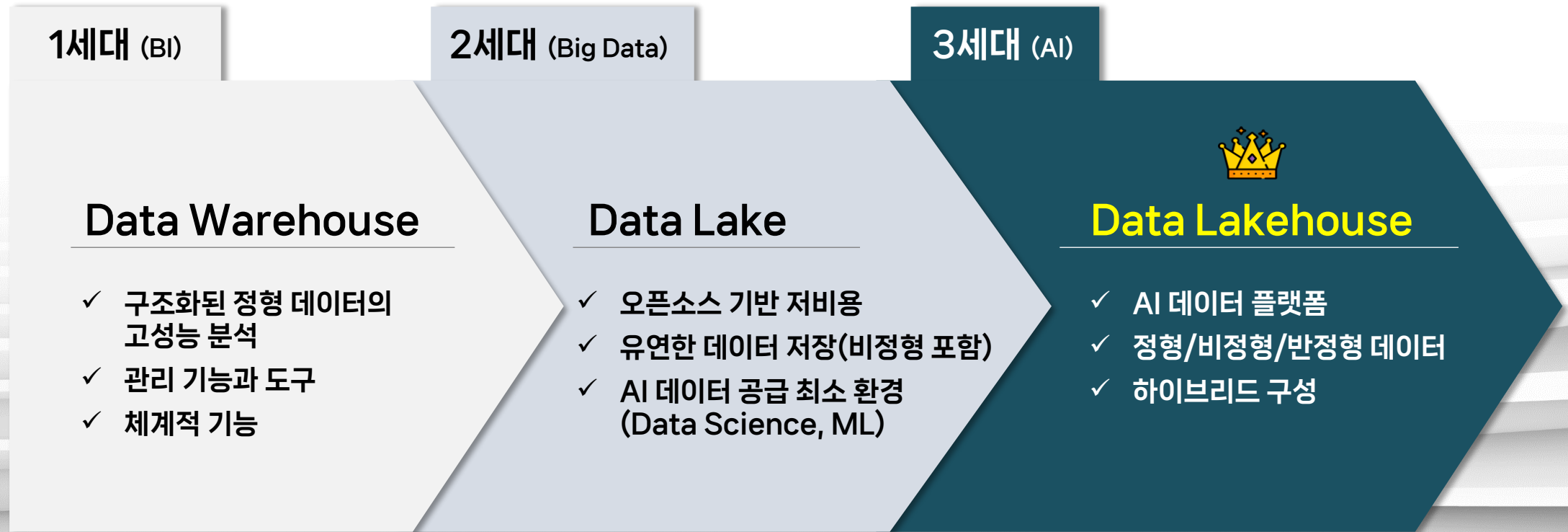


### 의료 부문



## 1.3.1 데이터 저장소의 진화

전통적인 1세대 Data Warehouse에서 빅데이터를 위한 2세대 Data Lake로 발전해 왔으며, 현재는 이 두 기술의 장점을 통합한 3세대 Data Lakehouse가 등장해, 분석과 AI를 아우르는 새로운 패러다임을 제시



## 1.3.2 타란툴라 Lakehouse 개요

데이터 레이크의 민첩성과 비용효율성, 데이터 웨어하우스의 강력한 분석 기능을 결합한 **통합 빅데이터 플랫폼**  
AI 시대 최적의 데이터 플랫폼으로, BI 연동 뿐 아니라 snowflake와 AI/ML 연동까지 자유로운 **국내 유일 솔루션**

*On-premise*



### 핵심 기능

- ✓ 통합된 SQL 쿼리 엔진
- ✓ 고급 쿼리 최적화
- ✓ ACID 트랜잭션 지원
- ✓ **스토리지-컴퓨팅 영역분리 아키텍처**



### 주요 장점

- ✓ 비용 효율성 극대화
- ✓ 개방성(클라우드 연동)
- ✓ 대규모 분석 등 성능 우수성
- ✓ 유연성(AI/ML 연동)



## 1.3.3 두 아키텍처의 장점을 하나로: 차세대 통합 저장소

### AS-IS

용도에 따라 별개의 시스템 사용

→ 막대한 도입비용과 유지비용

#### 데이터 웨어하우스(DW)

- 오라클 Exadata
- Teradata 등

#### 한계

- ① 증설/교체 고비용
- ② 대량 처리시 성능 ↓
- ③ 비정형 처리 제약

#### 데이터 레이크 (DL)

- 클라우데라 / 하둡
- Amazon S3 등

- ① 저렴하지만 속도 느림
- ② 스키마, 파티션 관리 등 운영 부담
- ③ 데이터정합성 이슈 (비정상 종료시)

### TO-BE

#### DW와 DL을 통합한 '3세대' 데이터 플랫폼

→ Lakehouse를 통해 TCO 절감, AI 기반 확대

#### 타란툴라 Lakehouse



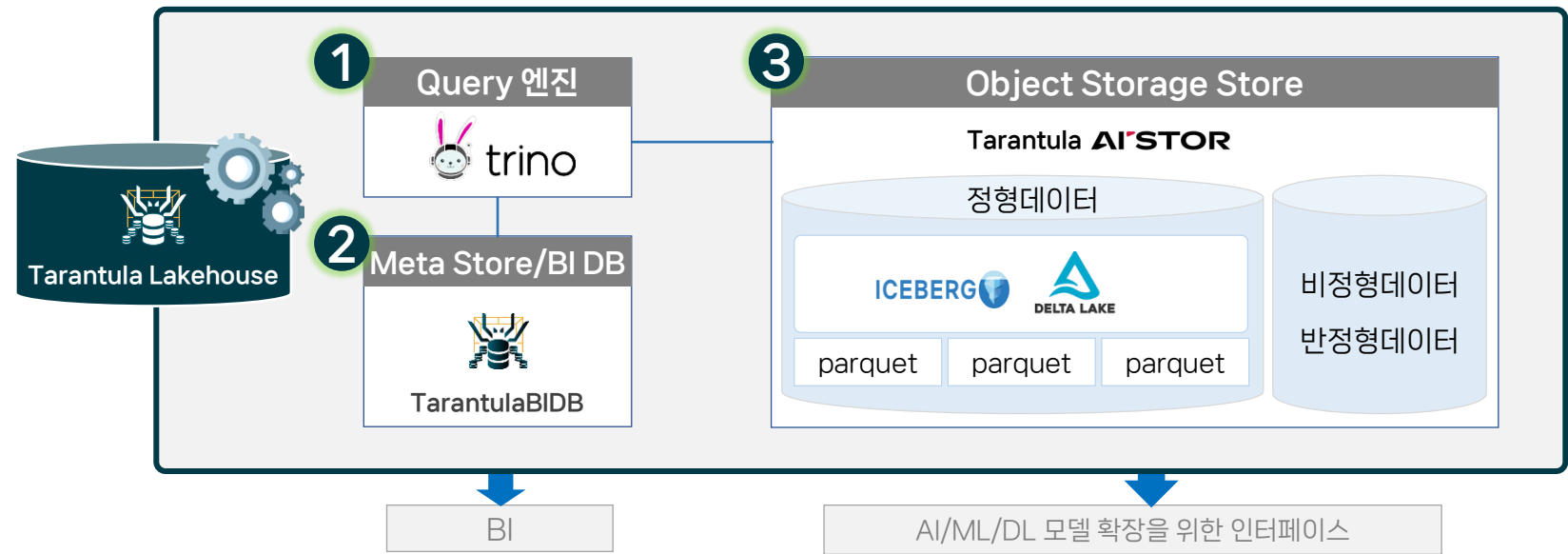
#### 기존 한계 극복

- 1 상용 대비 90% 비용 절감
- 2 대량 처리시 탁월한 성능 ↑
- 3 정형 + 비정형 모두 처리
- 4 메타데이터 자동 관리로 운영 부담 감소
- 5 데이터 정합성 보장
- 6 AI/ML 데이터 연동 용이



### 1.3.4 솔루션 구성요소

데이터 초고속 쿼리와 메타데이터 관리, 저장까지 각 영역에 특화된 구성요소를 결합한 통합 데이터 아키텍처로, 대규모 데이터 환경에서도 신뢰성과 성능을 균형 있게 제공

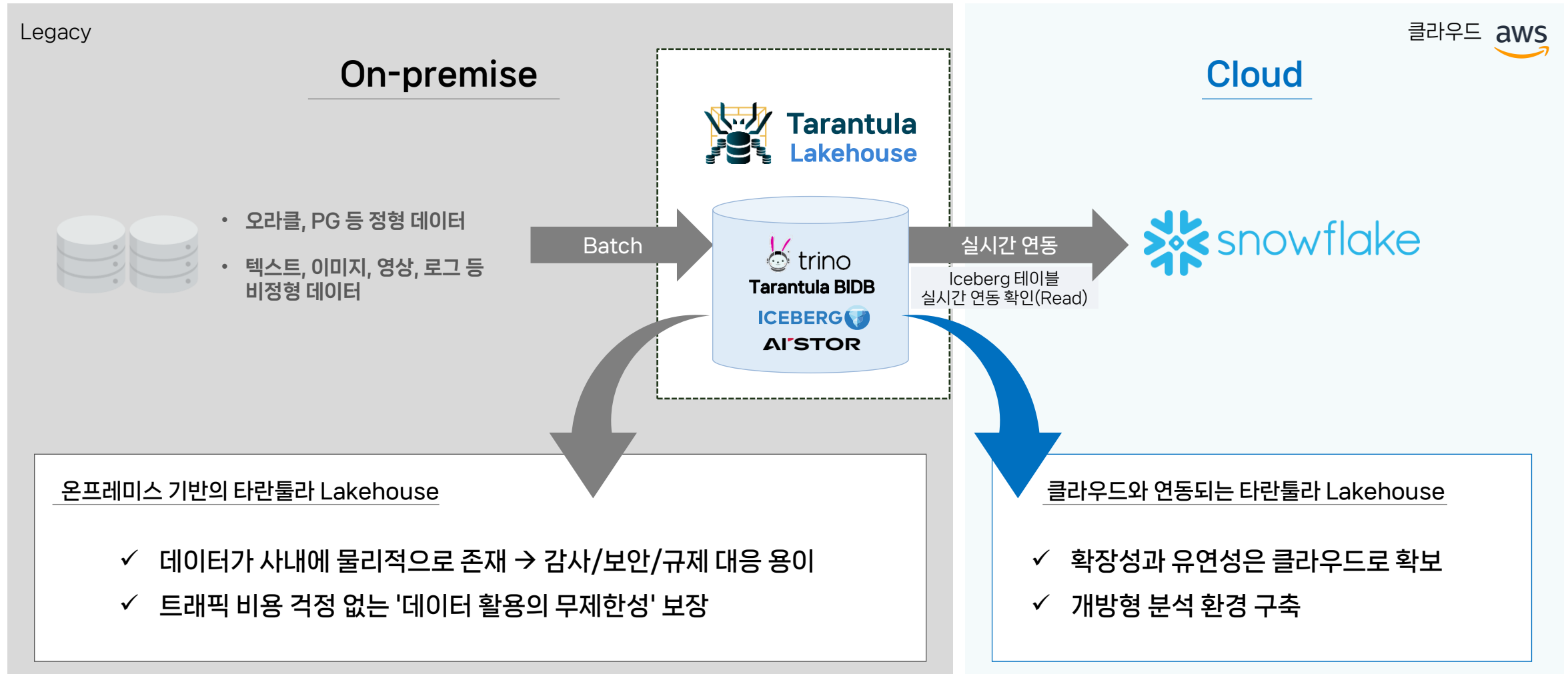


#### ■ 구성요소

핵심 컴포넌트	솔루션명	역할 및 기능
1. Query 엔진	trino	<ul style="list-style-type: none"><li>다양한 데이터 소스를 대상으로 SQL을 사용해 빠르게 분석하는 분산형 *MPP SQL 쿼리 엔진</li><li>* MPP, Massively Parallel Processing: 대규모 병렬 처리</li></ul>
2. Meta Store/BI DB	TarantulaBIDB	<ul style="list-style-type: none"><li>ICEBERG 카탈로그 저장 Meta DB</li><li>ICEBERG 테이블 조회 용도(Trino Bridge 자체 개발) 및 향후 Vector DB Extension</li></ul>
3. Object Storage Store	Tarantula AISTOR	<ul style="list-style-type: none"><li>AI 및 대규모 데이터 워크로드를 위해 설계된 온프레미스 고성능 오브젝트 스토리지 플랫폼으로, 다양한 AI 기능을 통합하여 단일 인프라에서 효율적인 데이터 관리와 분석 제공</li></ul>

### 1.3.5 특징점 – 1) 온프레임과 클라우드를 포괄하는 데이터 통합 플랫폼

온프레미스의 강력한 보안성과 클라우드의 유연한 확장성을 결합하여, 기업의 모든 데이터를 한곳에서 연결하고 비즈니스 가치로 전환하는 통합 데이터 허브



## 1.3.5 특징점 – 2) 뛰어난 쿼리성과 저장공간 효율성

N사 PoC 결과 타 솔루션 대비 뛰어난 쿼리성과 저장공간 압축 효율성을 검증 (2025년 5월 PoC 결과)

### 뛰어난 쿼리 성능

수억~수십억건의 쿼리를 안정적으로 처리

\* 2개의 테이블, 매매체결 Raw 데이터에 대해 업무 쿼리

단위: 초

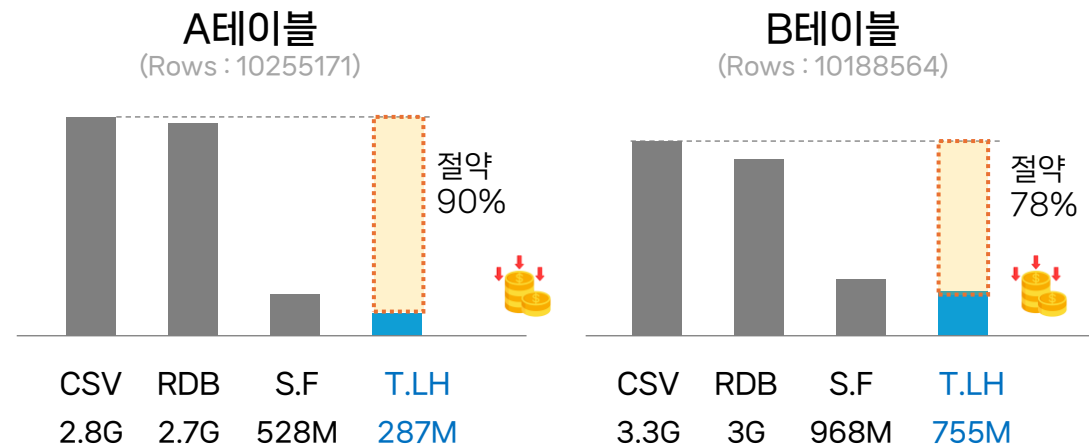
쿼리 부하 유형	A사	B사	타란툴라 Lakehouse	
A 하루 집계 (천만)	8(*23)	0.7(*3)	0.4	1등
A 한달 집계 (수억)	13	1	0.8	1등
A 일년 집계 (수십억)	26	2	4	2등
A,B 조인 하루 집계	12	1.2	1	1등
A,B 조인 한달 집계	29	3.3	2.5	1등
A,B 조인 일년 집계	52	15	19	2등

✓ 조인은 DW 방식을 사용

\* 데이터 캐싱 전 측정치

### 월등한 저장공간 효율성

같은 데이터라도 가장 적은 공간 사용



CSV: 값을 쉼표로 구분한 오픈포맷

S.F : Snowflake

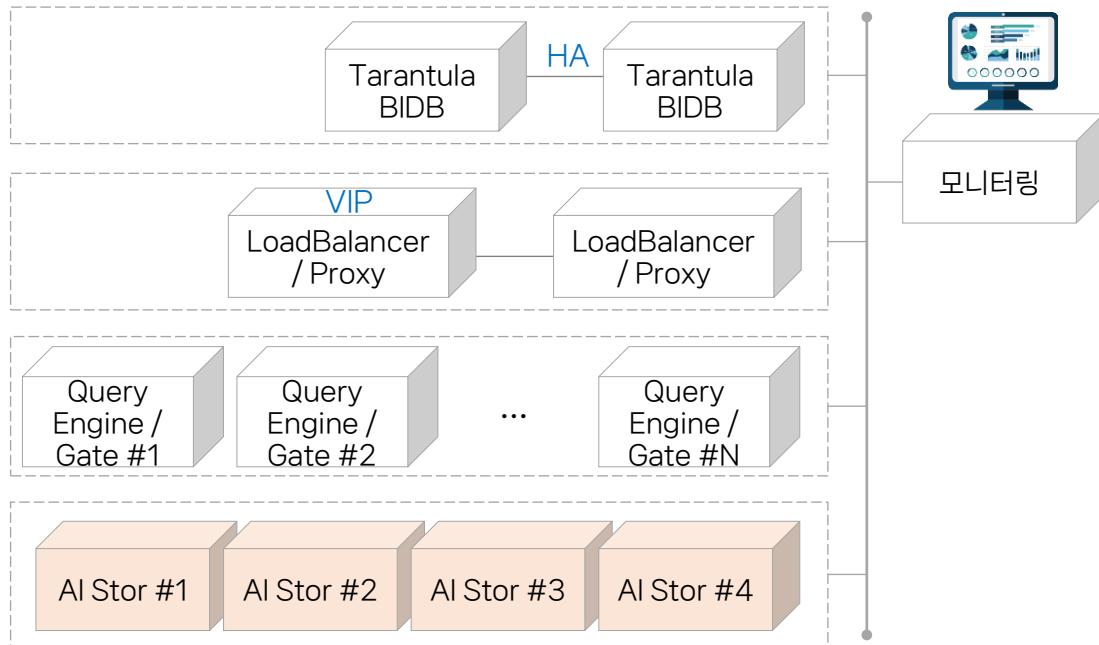
RDB : 관계형 데이터

T.LH : 타란툴라 Lakehouse

## 1.3.6 서버 구성 – 1) 권장 구성도 / 최소화 구성도

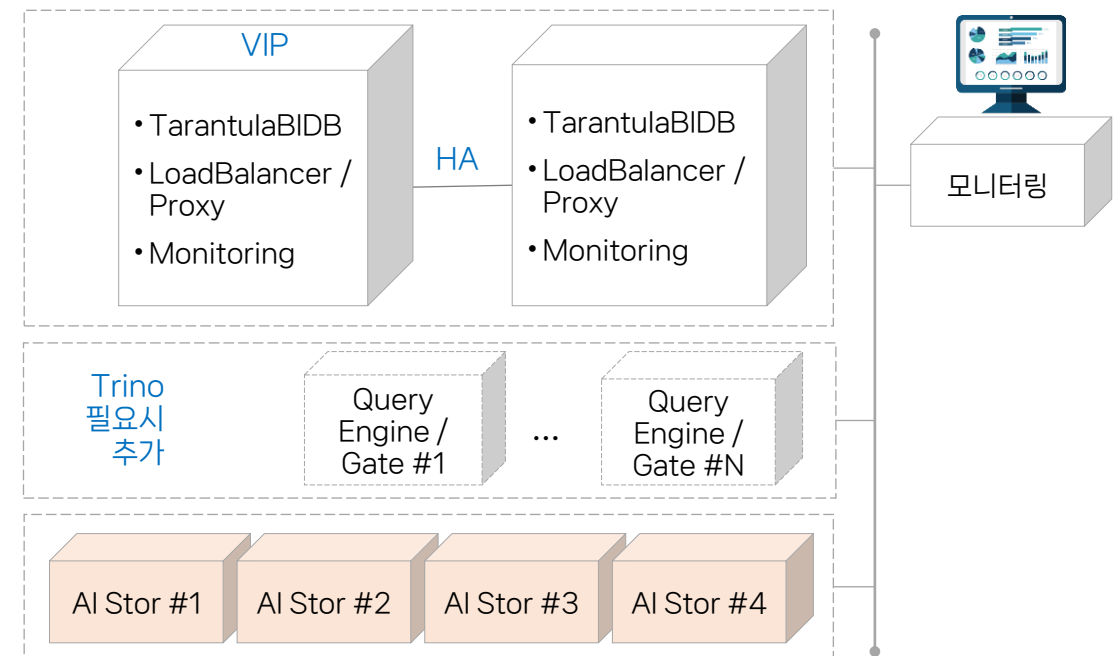
일반적인 권장 구성이며, 단계적 확장을 고려하여 초기 비용 부담을 줄인 최소화 구성안 선택 가능

1안 – 권장 구성도



- 성능과 안정성 최대화 구성

2안 – 최소화 구성도



- 하나의 서버에서 동작할 수 있는 여러 기능들을 묶어 서버의 수량을 최소화 한 구성 (3가지 기능을 하나의 서버에 구성)
- 운영상 문제는 없으며 데이터 처리범위를 단계적으로 확장 시, 초기 단계 구성에 적합

## 1.3.6 서버 구성 – 2) 서버 권장사양 (Specifications)

※ 서버 OS 는 모두 RHEL9 으로 구성. (서브스크립션 필요), AISTOR 서버를 제외한 모든서버는 가상서버로 구성 가능

구분	항목	권장사양	수량	
TarantulaBIDB 서버	프로세서	16 Core	1	<ul style="list-style-type: none"> <li>HA 구성으로 동일사양 <a href="#">2대 필요</a></li> <li>Pacemaker 필요</li> <li>공유 스토리지 영역 필요</li> </ul>
	메모리	128 GB Mem	1	
	OS 디스크	480GB SSD	2	
	NIC 1	4포트 1GbE	2	
	NIC 2	듀얼 포트 10/25GbE	2	
	HBA	듀얼 포트	2	
Load Balancer/ Proxy 서버	프로세서	8 Core	1	<ul style="list-style-type: none"> <li>동기화 구성으로 동일사양 <a href="#">2대 필요</a></li> </ul>
	메모리	64 GB Mem	1	
	OS 디스크	480GB SSD	2	
	Data 디스크	240GB SSD	2	
	NIC 1	4포트 1GbE	2	
	NIC 2	듀얼 포트 10/25GbE	2	
Query Engine/ Gate 서버	프로세서	8 Core	1	<ul style="list-style-type: none"> <li>동일사양 <a href="#">최소 2대 필요</a> (데이터 부하에 따라 가변적)</li> </ul>
	메모리	256 GB Mem	1	
	OS 디스크	480GB SSD	2	
	Data 디스크	240GB SSD	2	
	NIC 1	4포트 1GbE	2	
	NIC 2	듀얼 포트 10/25GbE	2	
AISTor 서버	프로세서	16 Core	1	<ul style="list-style-type: none"> <li>동일사양 <a href="#">4대 필요</a></li> <li>Data 영역은 별도 Raid 구성 필요 없음 (MinIO 자체 Parity 설정)</li> <li>전체 용량 중 약 63% 정도 Usable (100TB Usable 필요 시 약 160TB 물리 디스크 필요)</li> <li>제안 스토리지 <a href="#">가용영역은 최소 100TB</a> 기준임</li> <li>NVMe 디스크 사용 권장 (서버당 NVMe PCIe 4.0 디스크 11개 장착)</li> <li>100Gbps 초고속 네트워크 지원</li> </ul>
	메모리	256 GB Mem	1	
	OS 디스크	480GB SSD	2	
	Data 디스크	3.84TB NVMe	11	
	NIC 1	4포트 1GbE	2	
	NIC 2	듀얼 포트 10/25GbE	2	
Monitoring 서버	프로세서	8 Core	1	<ul style="list-style-type: none"> <li><a href="#">1대</a> 단독 서버</li> </ul>
	메모리	128 GB Mem	1	
	OS 디스크	480GB SSD	2	
	Data 디스크	1.92TB SSD	2	
	NIC 1	4포트 1GbE	2	
	NIC 2	듀얼 포트 10/25GbE	2	

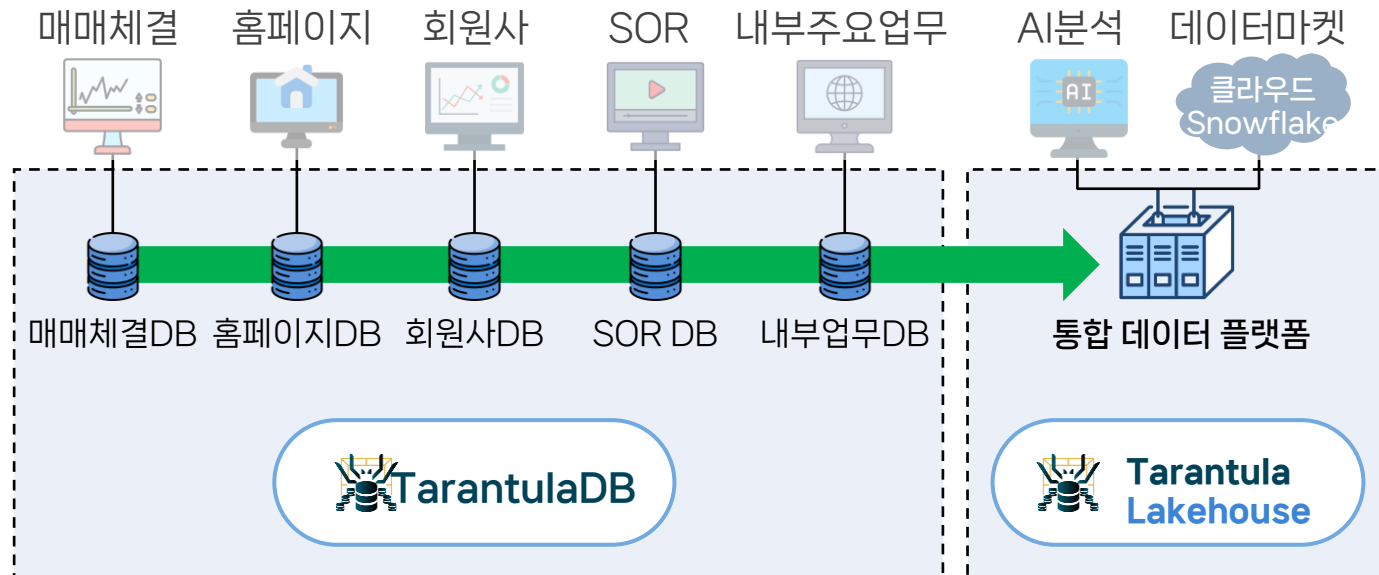
## 1.3.7 라이선스 유형

구분		Enterprise Edition	Standard Edition
용도		금융, 제조, 유통 등 엔터프라이즈용	병원 연구소, 지자체, 중소기업 등 프로젝트 연구 개발용
권장 데이터 용량		50TB 이상 권장	500GB ~ 10TB
시스템 구성	<b>Computing 영역</b> <ul style="list-style-type: none"> <li>• MetaDB (타란툴라BIDB)</li> <li>• Query Engine (Trino)</li> <li>• Proxy</li> </ul>	<b>이중화 구성</b> (서비스 연속성, 운영안정성 확보)	<b>단일 구성</b> (서비스 다운타임 발생 가능)
	<b>Storage 영역</b> <ul style="list-style-type: none"> <li>• MINIO</li> </ul>	<b>MINIO 엔터프라이즈 버전 (AISTOR)</b> (MINIO 공식 지원제품)	<b>MINIO Community Version</b> (사내 용도만 사용 가능)
기술지원		24 x 7 기술지원 / 4시간 이내 (UNNET, MINIO)	8 x 5 기술지원 / 1Day (UNNET Only)

## 1.3.8 도입사례 - 넥스트레이드 (1/2)



타란툴라DB가 표준DB로 선정되어 전체 시스템의 DB로 도입



✓ 일평균 10조원 대규모 거래처리 (25년)

✓ 금융 IT 에서 최고 수준의 신뢰성 요구

✓ 주가, 등락률, 거래량 등 실시간 데이터 처리

✓ 25년 3월 오픈 후 무중단 안정적 운영중

### ✓ 타란툴라DB를 선택한 이유

1. 상용 대비 약 90% TCO 절감
2. PoC를 통한 성능과 기술력 검증
3. 타사 대비 뛰어난 HA 이중화

### ✓ 타란툴라 Lakehouse를 선택한 이유

1. AI, BI, Snowflake 모두 연동되는 국내 최초 솔루션
2. 민감한 데이터는 On-prem 저장
3. 오라클, Snowflake 등 글로벌 기업과 PoC 통해 안정성/성능 검증



## 1.3.8 도입사례 - 넥스트레이드 (2/2)

넥스트레이드는 향후 AI 활용 및 데이터 판매를 위한 데이터 플랫폼으로 타란툴라 Lakehouse를 채택하였으며, 이를 통해 84%의 저장 비용 절감하고 데이터 처리에 따른 운영 부하 리스크를 제거하는 효과를 달성

### 도입 배경

#### 1) 대량 데이터 활용기반 필요

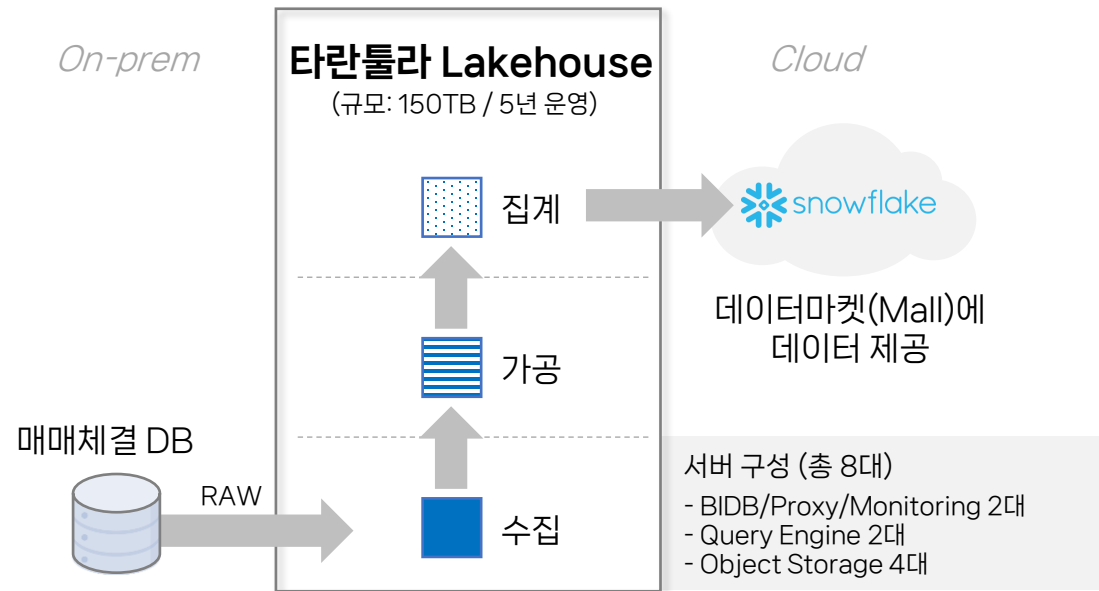
일일 적재량이 120GB 수준  
→ 연 30TB 데이터를 Tape 보관 시 과거 데이터는 활용이 어려움

#### 2) 데이터 보관비용 절감 필요

#### 3) 운영리스크 최소화 필요

통계 데이터 판매를 통해 수익을 창출하려 하나, 데이터 생성에 따른 시스템 부하와 운영 리스크 고민

### 적용 사례 (2025년 11월)



Lakehouse 내 저장 공간을 분할하여 단계적으로 통계생성  
(→ 데이터 품질, 보안성, 재사용성 확보)

### 기대효과

#### ✓ 혁신적 비용 절감

기존 시스템 증설 대비 84% 비용 절감

#### ✓ 데이터 활용방안 증대

- LLM용 데이터
- 과거데이터 통계작성 용이
- 데이터 판매
- 클라우드 연동

#### ✓ 운영리스크 최소화

운영 DB와 분리함으로써 부하에 따른 리스크 제거

#### ✓ 확장성 확보

데이터가 늘어나도 확장 용이

## 1.3.9 글로벌 기업의 Iceberg기반 Lakehouse 적용 사례

### NETFLIX

- 1. 저장 비용 90% 절감:**  
기존 테이블 포맷 대비 메타데이터 최적화 및 파일 관리 개선
- 2. 쿼리 성능 대폭 향상:**  
메타데이터 푸시다운으로 데이터 파일 스캔 완전 제거, 일부 쿼리는 1시간 → 수 초로 단축
- 3. 증분 처리 효율화:**  
전체 재처리 대신 변경분만 처리하여 컴퓨팅 비용 대폭 절감

#### 기존 저장소의 한계

- **Hive 기반 시스템의 한계** 대규모 데이터셋에서 스키마 진화, 동시성, 성능 문제 발생
- **엑사바이트급 데이터** 웨어하우스의 확장성 및 신뢰성 요구
- **지연 데이터 처리와 증분** 처리의 효율화 필요

개선

#### Lakehouse 도입 후

- ☑ Exabytes급 데이터 처리의 안정성 확보
- ☑ 비용 절감
- ☑ 성능 향상
- ☑ 업무 효율화

### Apple

- 1. 유지보수 작업 시간 98% 단축:**  
2시간 → 수 분으로 감소
- 2. 집계 쿼리 99.9% 단축:**  
1시간 이상 → 수 초로 개선, 메타데이터 푸시다운 활용
- 3. 규제 준수 비용 대폭 절감:**  
대규모 데이터셋에서 효율적인 행 단위 업데이트 가능

#### 기존 저장소의 한계

- **GDPR, DMA 등 규제 준수** 행 단위 업데이트 필요
- **전통적 Hive 한계** 파티션 단위로만 업데이트 가능해 대규모 데이터셋에서 비효율적
- **페타바이트급 테이블** 유지보수 작업 시간 단축 필요

개선

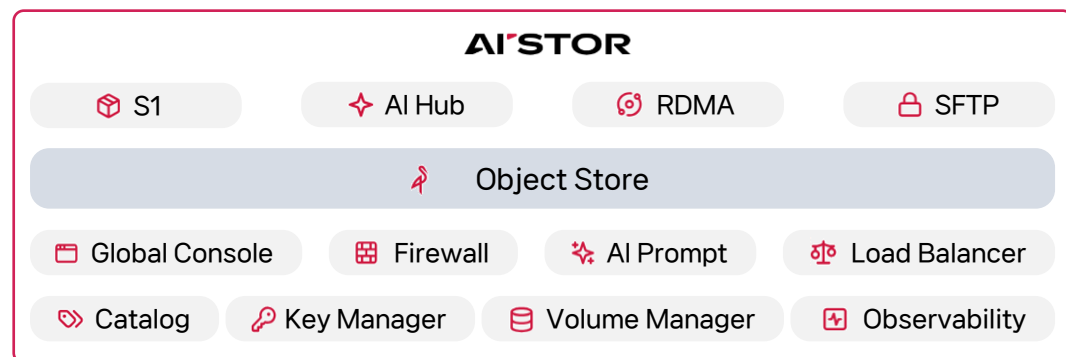
#### Lakehouse 도입 후

- ☑ 규제 준수환경 개선
- ☑ 대규모 데이터셋 처리환경 구축
- ☑ 유지보수 시간 단축

## 1. 솔루션 소개 ▶ 1.4 Minio AIStor

### 1.4.1 MinIO AIStor 소개

AI 워크로드 지원을 위한 대규모 데이터 인프라 운영에 필요한 모든 구성 요소를 포함하는 오브젝트 스토리지 솔루션



- ✓ 세계에서 가장 빠른 오브젝트 스토어 – 실운영환경 **초당 2.2 TiB\***처리량
- ✓ ARM에서 x86까지 **모든 하드웨어에서 실행** – Cloud Native 표준
- ✓ **암호화**와 IAM, ILM, 로드밸런싱, 캐싱, 가시성 등 엔터프라이즈급 제공
- ✓ **수천 개의 분산 노드**에 걸쳐 단일 네임스페이스 생성 – 놀라운 확장성
- ✓ AWS S3, HDFS 대비 **40% ~ 70%** 까지 **TCO 절감**

#### ■ S3와 비교 시 우위 포인트

솔루션 비교	AWS S3	타란툴라 Lakehouse AIStor	우위 포인트
배포 방식	AWS 클라우드 전용	온프레미스 / 클라우드 / 엣지	✓ 유연한 배포 환경
성능	네트워크와 AWS 내부 한계 존재	초고속 (하드웨어 직접 제어 가능)	✓ 로컬 디스크 최적화, 초고성능
비용 구조	저장량·트래픽에 따라 비용 증가	무제한 저장 (서버만 있으면 됨)	✓ 비용 예측 가능, 대규모 저장에 유리
S3 API 호환성	표준	완벽 호환 (AWS S3와 동일한 API)	✓ 이식성 우수, 기존 S3 기반 앱과 호환
고가용성	리전·AZ 분산으로 고가용성	4노드 이상 분산 구성 지원	✓ 자체적 고가용성 클러스터
오프라인 사용	불가능 (AWS 연결 필요)	<b>가능 (내부망에서 독립적 운영)</b>	✓ 금융기관에 적합

#### ■ 레퍼런스



## 1.4.2 커뮤니티 버전 vs 엔터프라이즈 버전 (AISTOR)

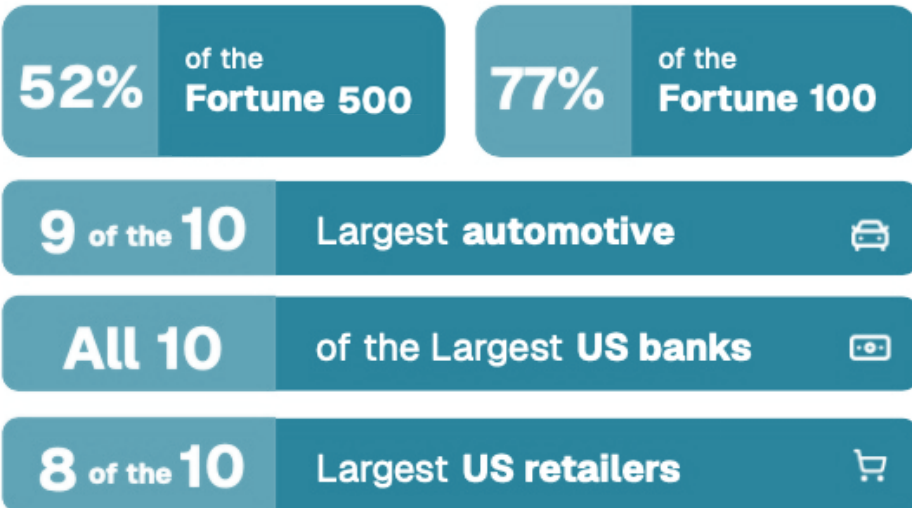
커뮤니티 버전은 테스트 및 개발용으로 기본 기능만 제공하며 실제 비즈니스 운영에서 필요한 보안, 규제준수, 성능 최적화와 전문 지원을 위해 Enterprise Edition인 AISTOR 사용을 권장

구분	MINIO OBJECT STORE Community Edition	MINIO AISTOR Enterprise Edition
라이선스	GNU AGPL v3 라이선스	Enterprise License
용도	테스트 및 개발용	실제 운영환경용 (Production)
보안	네트워크 암호화 : TLS 1.2/1.3	<ul style="list-style-type: none"> <li>최신 취약점(CVE) 평가 및 보완</li> <li>네트워크 암호화 : TLS 1.2/1.3</li> <li>저장 데이터 암호화 : SSE-KMS, SSE-S3, SSE-C</li> <li>암호화 알고리즘 : AES-256-GCM, ChaCha20-Poly1305, and AES-CBC</li> <li>KMS : AISTOR KMS 및 외부 서드파티 KMS</li> <li>데이터 방화벽</li> </ul>
규제준수	해당 없음	FIPS 140-3, 침투테스트 SOC2, ISO 27001, SEC 17a-4(f), FINRA 4511(c), CFTC 1.31(c)-(d)
지원	<ul style="list-style-type: none"> <li>커뮤니티 슬랙(Slack) 채널</li> <li>자원봉사자 지원</li> </ul>	<ul style="list-style-type: none"> <li>SUBNET 엔터프라이즈 지원 포털</li> <li>SLA: 연중무휴 24/7, 응답시간 4시간 이내, 중요 이슈(PO) 즉각 대응</li> </ul>
기능	기본 기능 (지속적 투자 차원)	기본 기능 + 추가 기능(적극적 투자): 콘솔, 암호화, 사이트 간 복제, QoS, 카탈로그, 모니터링, 감사 로그, 헬스 진단, RDMA, GPU Direct, AI 포털, AI 허브
K8s 통합	업스트림 쿠버네티스	업스트림 + Red Hat OpenShift, Rancher RKE2/K3S, VMware Tanzu (TKG)
최적화	해당 없음	소규모 객체 최적화, 대량 삭제, 리스트 작업, 낮은 TTFB, 분산 캐시, 빠른 부팅

### 1.4.3 레퍼런스

글로벌 No.1 오브젝트 스토리지로 전세계 많은 기업들에 도입된 바 있으며, 유엔넷은 국내 최초 OEM 계약을 통해 판매 및 서비스하며 향후 일본 등 아시아 시장 진출 예정

#### MinIO에서 가동되는 전세계 기업들



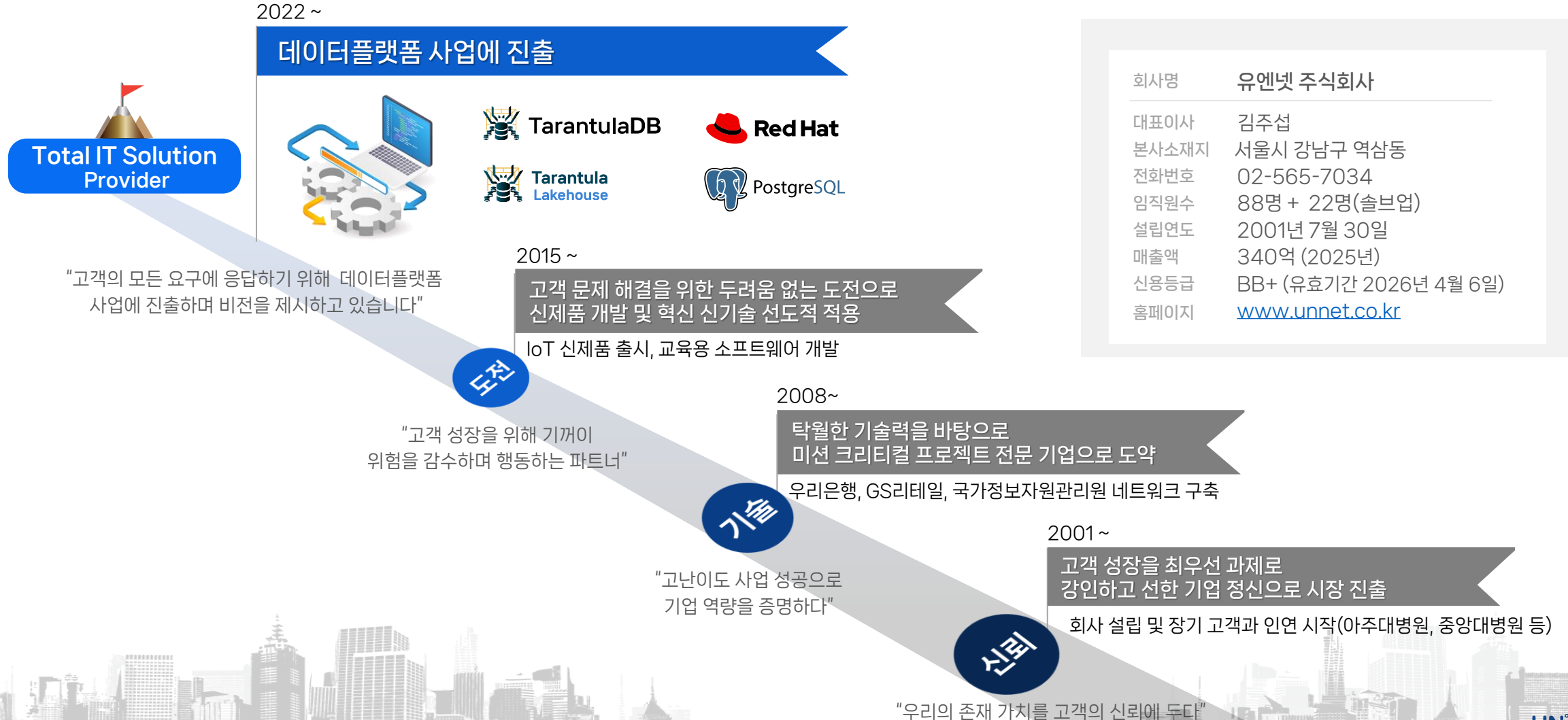
#### 도입 사례



## 2. 유엔넷 소개

### 2.1 회사소개 및 연혁

2001년 설립 후 25년간 다양한 IT 사업 경험을 축적하며 **종합 IT 솔루션 기업**으로 성장



## 2. 유엔넷 소개

### 2.2 조직도 및 인력현황

4개 본부 하에 연구개발 자회사를 포함하여 전체 110명의 임직원으로 구성

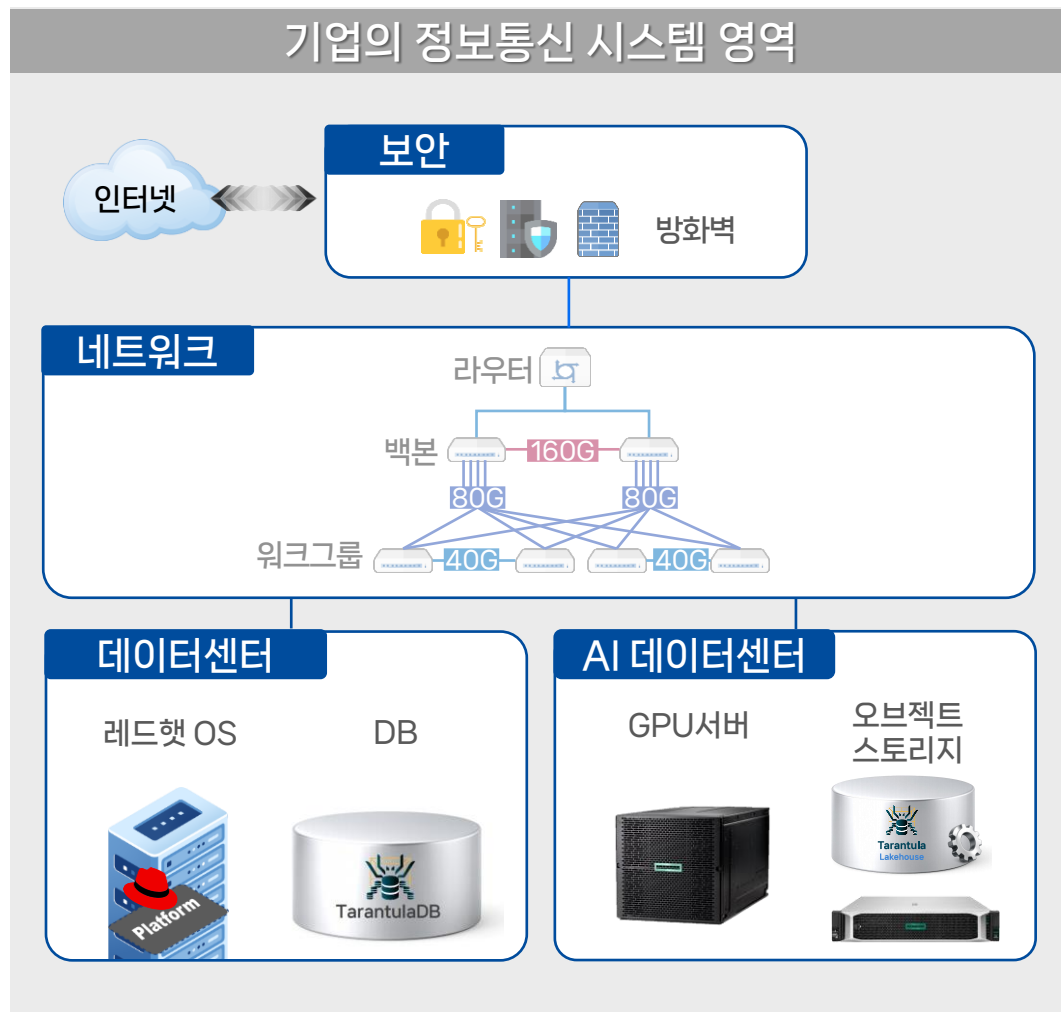




## 2. 유엔넷 소개

### 2.3 사업 영역

클라우드 / 데이터센터의 "IT Infrastructure"에 대한 토탈 솔루션 제공



#### AI 솔루션

GPU서버 및 RoCE 네트워크 등 AI 인프라  
벡터DB, 레이크 하우스 등 AI 데이터 솔루션



#### 데이터 솔루션

오픈소스 1등 PostgreSQL기반 기업용 DBMS  
처리속도 1등 데이터 레이크 하우스 솔루션 및 벡터DB



#### SW 솔루션

오픈소스 점유율 1등 RedHat OS 및 가상화 플랫폼  
오픈소스 기반의 오브젝트 스토리지 솔루션



#### 인프라 솔루션

AI기반 유/무선 네트워크 및 데이터센터/SDN 솔루션  
서버, 스토리지 및 보안 솔루션



# 감사합니다.

유엔넷 주식회사

## 솔루션 문의



이강일 전무

전화 010-2717-7408

이메일 ki.lee@unnet.co.kr



정지원 이사

전화 010-4121-1486

이메일 j1jung@unnet.co.kr



신기태 이사

전화 010-8407-7102

이메일 skt0609@unnet.co.kr

## 유엔넷 연락처

[HQ] 서울 강남구 테헤란로20길 12 안타워 2층

[여의도 Office] 서울 영등포구 여의나루로 71 동화빌딩 7층



02-565-7034



www.unnet.co.kr



www.tarantuladb.com

