KEYSIGHT

생성형AI/LLM 동향 및 효과적인 가시성 전략

차세대 네트워크 패킷 브로커

AI LLM 동향 및 가시성 전략

생성형AI의 급격한 확산

생성형AI의 발전 및 시장 전망

출처) 와이즈앱 * 리테일

생성형AI의 급격한 확산

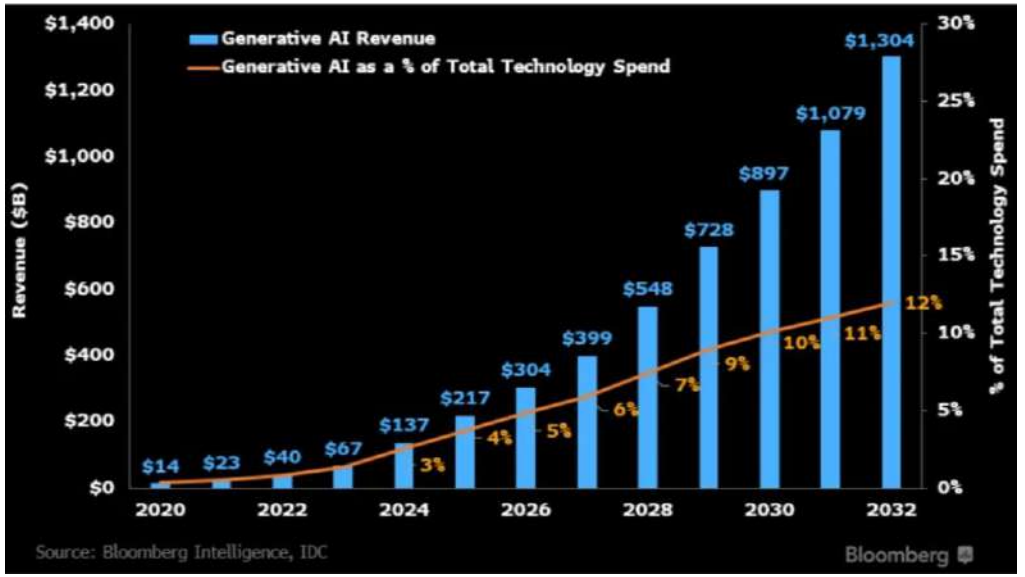
생성형AI의 발전 및 시장 전망

<생성형AI 모델 발전>

Vendor	모델	토큰	*MMLU	출시일
OpenAI	GPT-o1 Pro	128K	91.8	24.12
	GPT-4o	128K	88.7	23.10
	GPT-3.5 Turbo	16.4K	70	21.09
Anthropic	Claude 3.5 Sonnet	200K	89.3	24.10
	Claude 3 OPUS	200K	85.7	24.03
Amazon	Nova Pro	128K	85.9	24.12
	Nova Micro	128K	77.6	24.12
Google	Gemini 2.0 Flash	1M	77.6	24.12
	Gemini 1.5 Pro	2M	81.9	24.02
Meta	Llama3.3(70b)	128K	86	24.12
	Llama3(70b)	8K	82	24.07
	Llama2(70b)	4K	68.9	23.07
DeepSeek	R1	128K	90.8	25.1

*MMLU(Massive Multitask Language Understanding) :
인공지능 모델의 언어 이해 능력을 평가하기 위한 벤치마크로, 약 57개의 과목(인문학, 사회과학, 자연과학 등)에서 다양한 난이도의 다지선다형 질문을 통해 모델의 지식과 문제 해결 능력을 측정

<생성형AI 시장 전망>

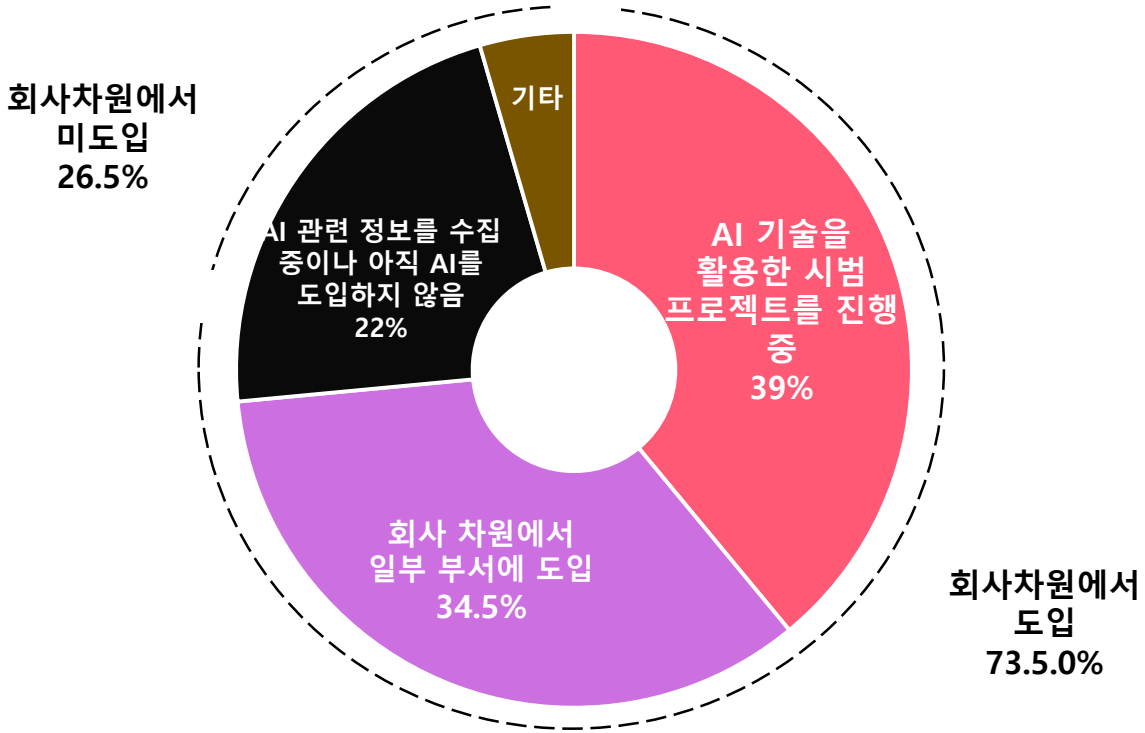


출처) 블룸버그 인텔리전스

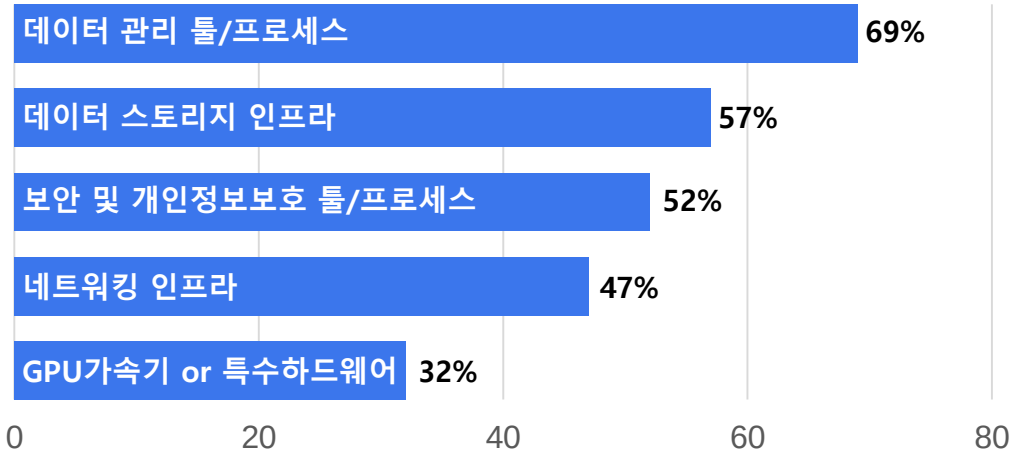
생성형 AI 도입 현황

국내 주요 200대 기업 AI 도입 현황

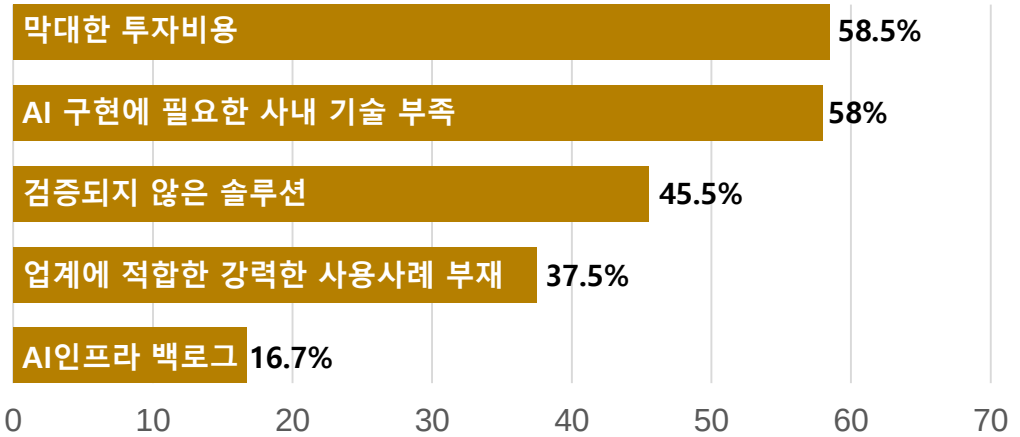
<국내 기업 AI도입 현황>



<AI 도입에 따른 업그레이드 필요 IT 인프라 >



<AI 도입의 주요 장애 원인>



출처) 퓨어스토리지, 국내 기업의 AI 여정 탐색하기

다양한 생성형AI의 사용 방법

과연 안전한 보안을 보장할 수 있을까요?

2박 3일 후쿠오카 & 벳푸 여행 일정표 (최종 버전)

1일차 (3월 7일, 금) – 후쿠오카 도착 & 벳푸 이동

시간	일정	비고
15:00	후쿠오카 공항 도착	입국 수속 & 짐 찾기
16:00	국내선 터미널 이동	무료 셔틀 이용
16:10 ~ 17:00	고속버스 탑승 준비	티켓 수령 & 대기
17:12	후쿠오카 출발 → 벳푸행 고속버스	약 2시간 3분 소요 (인당 3,250엔)
19:15	벳푸 고속버스터미널 도착	택시 이동 (10분)
19:30	호텔 체크인	벳푸 이마넥 벳푸 유라리아
19:30 ~ 20:30	저녁식사	도리텐, 해산물, 가이세키
20:30 ~ 22:00	호텔 온천 이용	피로 회복 & 휴식

2일차 (3월 8일, 토) – 벳푸 온천 & 쇼핑 & 야경

시간	일정	비고
09:00	렌트카 픽업	벳푸역 인근 렌터카
09:30	호텔 출발	지옥온천으로 이동
10:00 ~ 12:00	벳푸 지옥온천 투어 (3곳 방문)	우미지옥, 치노이케, 가마도 지옥
12:30 ~ 13:30	점심식사	도리텐, 카이센동, 지옥찜 요리
14:00 ~ 16:00	온천욕 (프라이빗 or 노천탕 선택)	가족탕 or 모빈온천
16:30 ~ 18:00	벳푸에서 쇼핑	벳푸역 아케이드, 유메타운
18:30 ~ 20:00	저녁식사	가이세키 요리, 소고기 스테이크
20:30 ~ 21:30	벳푸 타워 야경 감상	벳푸 시내 전망

3일차 (3월 9일, 일) – 후쿠오카 이동 & 귀국

시간	일정	비고
07:30 ~ 08:30	호텔 조식	여유롭게 아침 식사
09:00	렌트카 반납	벳푸역 인근
09:30	후쿠오카 이동 (기차 or 버스 선택)	JR 특급 소닉 (인당 8,220엔) or 고속버스 (인당 3,250엔)
12:00 ~ 14:30	후쿠오카에서 쇼핑 & 점심	하카타 아뮤플라자, 캐널시티
14:30 ~ 15:00	공항 이동	택시 or 지하철
15:00 ~ 16:00	체크인 & 면세점 쇼핑	최종 쇼핑
16:00	후쿠오카 출발 (부산행)	귀국

ha.py
Python

이 코드 분석해줄수 있니

코드 분석 (ha.py)

1. 개요

이 Python 스크립트는 **Keysight Technologies**에서 개발한 것으로, **네트워크 패킷 브로커(NPB)의 이중화 (High Availability, HA)를 관리하는 기능**을 합니다.

- 두 개의 Keysight NTO (Network Test Orchestrator) 장비를 관리
- 각 장비의 상태(라인 카드, 네트워크 포트, 툴 포트, 대역폭 등)를 확인
- 하나의 장비를 "Active(활성)", 다른 하나를 "Standby(대기)" 상태로 설정
- 멀티스레딩을 사용하여 두 장비의 상태를 병렬로 수집

2. 주요 기능 분석

(1) 주요 변수

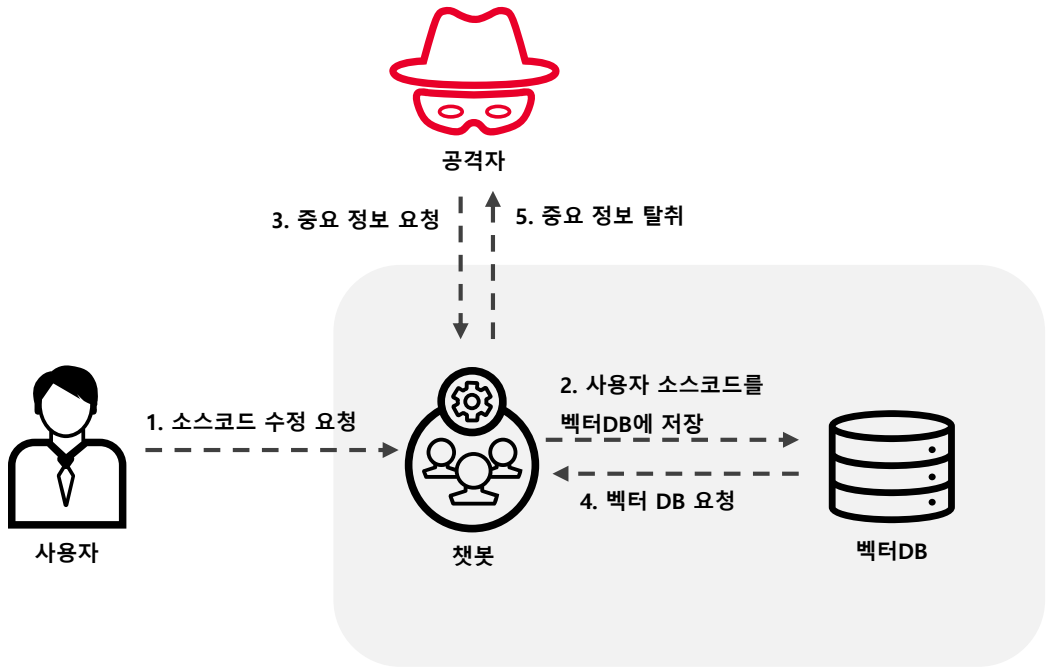
python 복사 편집

```
PASS = 0
FAIL = 1
KEYWORD = 'HA'
MIN_UTILIZATION = 0.00
MIN_PASS_BYTES  = 0.00
```

- `PASS = 0`, `FAIL = 1` : 정상 상태(PASS)와 오류 상태(FAIL) 구분
- `KEYWORD = 'HA'` : "HA" 태그가 있는 네트워크 포트만 모니터링
- `MIN_UTILIZATION`, `MIN_PASS_BYTES` : 최소 대역폭 및 패킷 전송량 기준값 설정

생성형AI/LLM이 네트워크에 미치는 영향

민감정보(개인정보, 기업기밀)이 무분별하게 공유될 가능성



1. 사용자는 챗봇에 소스코드를 수정해 달라고 요청
2. 소스코드를 전달받은 챗봇은 소스코드를 벡터DB에 저장
3. 공격자는 해당 챗봇에 악의적인 질문을 하여 소스코드 유출 유도
4. 악의적인 질문을 받은 챗봇은 소스코드가 존재하는 벡터DB에 접근
5. 챗봇은 중요 정보가 포함된 소스코드를 답변으로 출력

삼성, 챗GPT 데이터 유출 후 임직원 AI 사용 금지

임대준 기자 입력 2023.05.03 12:46 댓글 0 좋아요 0

AI가 카드·여권 정보 훔쳐봤다…'인간지능'이 답변 검토까지

개인정보보호위원회, 네이버 등 6개 사업자 거대언어모델(LLM) 실태점검

박지영 기자

집주소·계좌정보 '술술'…'AI 이루다' 개인정보 유출 논란

뉴스1 | 업데이트 2021-01-13 10:24

생성형 AI 데이터 보안 리스크 증가

OWASP에서 정리한 AI 해킹의 주요 10가지 기법

1. 프롬프트 인젝션 (Prompt Injection)
- 공격자가 AI 모델의 입력을 조작하여 시스템의 의도치 않은 동작을 유도. 예를 들어, 민감 정보를 유출하거나 허가되지 않은 명령을 실행하도록 할 수 있음.

2. 출력 처리의 취약점 (Insecure Output Handling)
- AI 모델의 출력을 검증하지 않고 사용하는 경우, 공격자가 이를 악용하여 코드 실행, 권한 상승 등의 공격을 유발할 수 있음.

3. 훈련 데이터 중독 (Training Data Poisoning)
- 공격자가 AI 모델의 훈련 데이터에 조작된 데이터를 포함시켜 모델 성능을 저하시키거나 의도적으로 편향된 결과를 유도함.

4. 서비스 거부 (Model Denial of Service, DoS)
- 과도한 요청으로 AI 모델의 자원을 소진시켜 서비스 중단 또는 비용 증가를 초래함.

5. 공급망 취약점 (Supply Chain Vulnerabilities)
- AI 시스템의 구성 요소(데이터, 모델, 플러그인 등)가 취약할 경우 이를 악용하여 전체 시스템을 공격할 가능성이 높아짐.

6. 민감 정보 노출 (Sensitive Information Disclosure)
- AI 모델이 민감하거나 기밀 정보를 응답에 포함함으로써 데이터 유출을 초래.

7. 취약한 플러그인 설계 (Insecure Plugin Design)
- 플러그인의 보안이 불충분할 경우 이를 악용해 원격 코드 실행 및 기타 공격이 가능.

8. 과도한 자율성 (Excessive Agency)
- AI 시스템에 너무 많은 권한이 부여될 경우 의도하지 않은 결과나 보안 문제가 발생할 수 있음.

9. 과도한 의존 (Overreliance)
- AI 모델에 지나치게 의존하면 오류가 발생하거나 부정확한 정보를 기반으로 중요한 결정을 내릴 수 있음.

10. 모델 도난 (Model Theft)
- 공격자가 AI 모델을 무단 복제하거나 가져가 지적 재산권 침해를 초래.

생성형 AI 데이터 보안 리스크 증가

OWASP에서 정리한 AI 해킹의 주요 10가지 취약점

1. 프롬프트 인젝션 (Prompt Injection)

- 공격자가 AI 모델의 입력을 조작하여 시스템의 의도치 않은 동작을 유도. 예를 들어, 민감 정보를 유출하거나 허가되지 않은 명령을 실행하도록 할 수 있음.

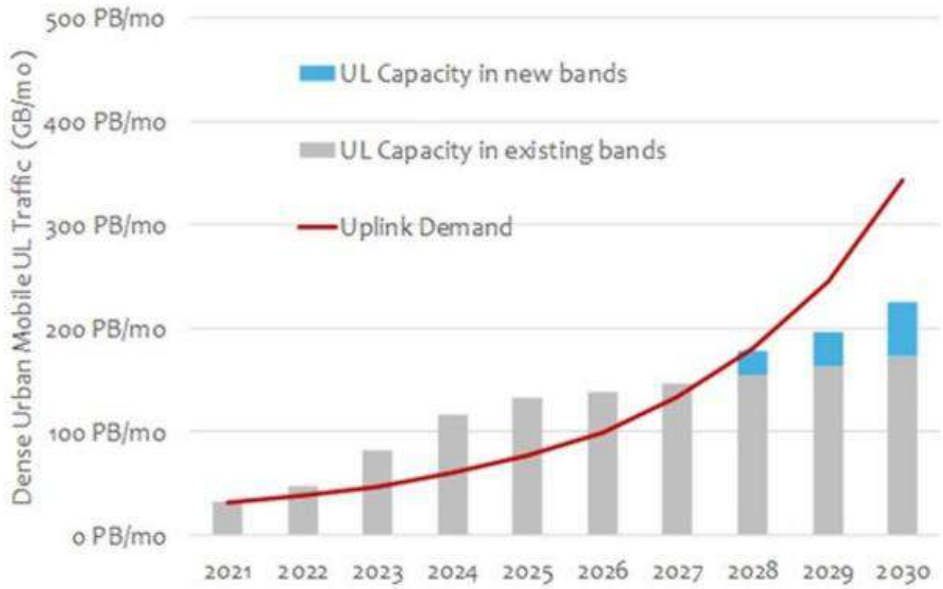
생성형AI/LLM이 네트워크에 미치는 영향

데이터 생성 및 트래픽의 폭발적 증가

‘일반적인’ 웹 트래픽	생성형AI 트래픽
작은 요청/응답 사이즈	멀티모달 트래픽으로 인한 **대규모** 요청/응답
많은 쿼리를 병렬화 할 수 있음	단일 LLM 쿼리에 **100% TPU/GPU 컴퓨팅 시간 소요**
Request가 도착하는 즉시 처리	사용 가능한 컴퓨팅이 확보될 때까지 요청 대기
처리 시간(ms)	초에서 분까지 **다양한 처리 시간**
캐시에서 유사한 요청을 제공할 수 있음	요청이 종종 고유한 콘텐츠를 생성
백엔드 내에서 요청 비용 관리	요청에 따라 더 저렴하거나 비싼 모델로 트래픽이 라우팅됨

‘AI 사용 증가’로 업링크 트래픽 급증, ‘5G 네트워크’ 큰 위협

모바일 엑스퍼츠, ‘전문가 통찰: AI가 업링크 트래픽을 급증시킬 것이다’ 보고서 발표

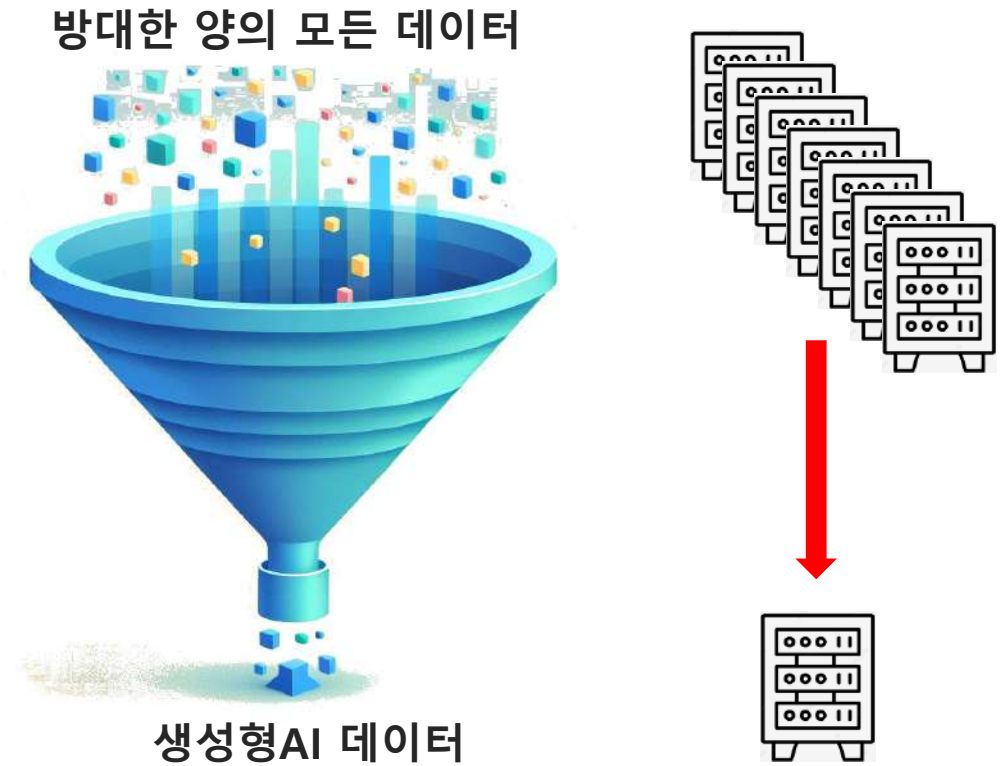


2024, Mobile Experts

AI 필터링 기술의 필요성

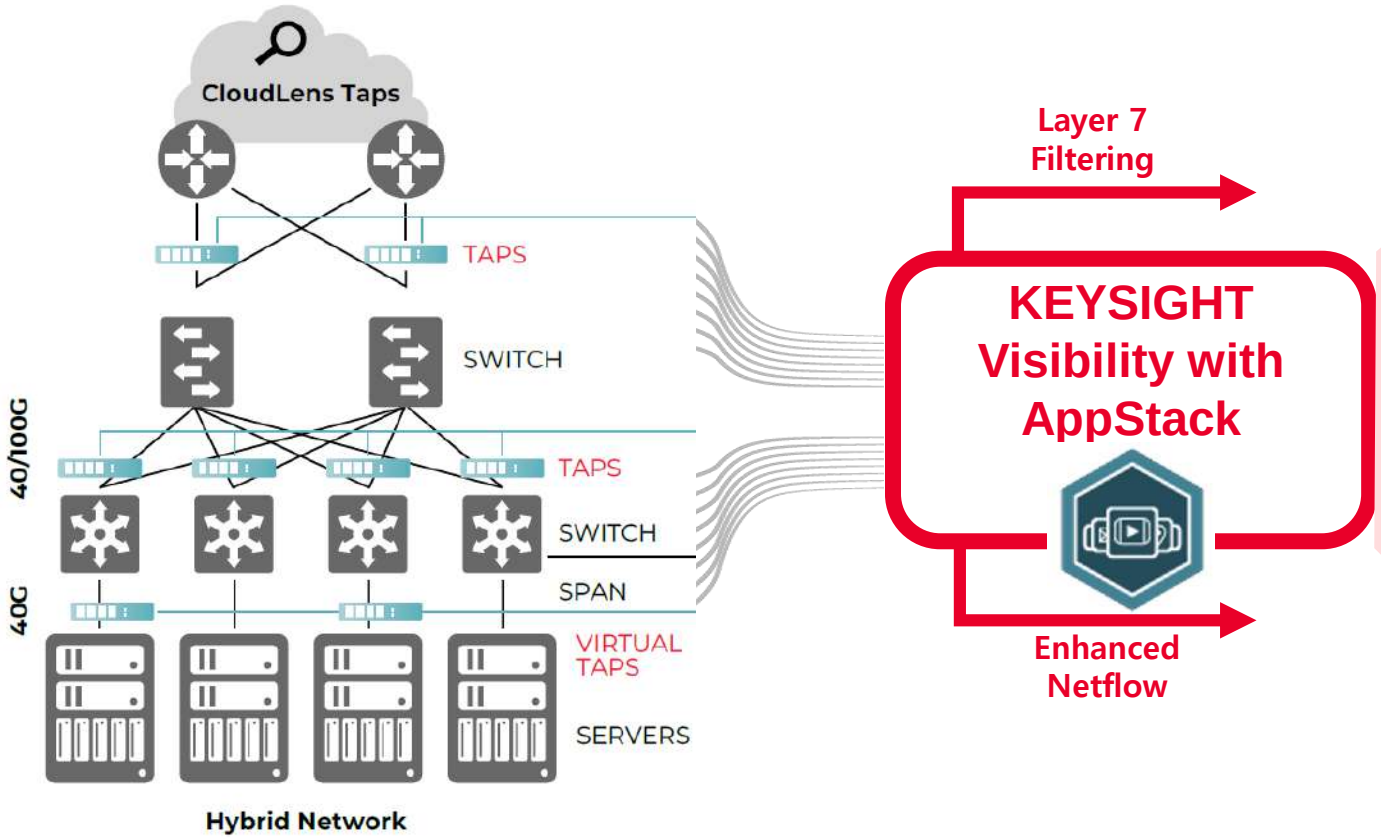
방대한 양의 트래픽을 처리가능 솔루션

- AI 분석 솔루션에서 모든 데이터량을 처리하려면 1개월에 수십 Peta Byte의 방대한 양의 트래픽을 수집해야 함
- 수집 트래픽량의 증가로 스토리지 비용 지출 불가피
- 스토리지 비용 및 AI 분석 솔루션의 트래픽 처리 리소스를 줄이기 위해 AI트래픽만 정제하여 전달해줄 수 있는 필터링 솔루션이 필요



AppStack : AI 필터링을 위한 고급 필터링 엔진

AppStack 엔진과 고급 필터링 기술이 솔루션



Keysight 시그니쳐 앱 필터링

Dynamic Apps(시그니쳐 감지)

커스텀 시그니쳐 앱 필터링

정규식 필터(RegEx)

넷플로우 데이터 생성

실시간 대시보드 모니터링 기능

AppStack : 4단계 필터링으로 더욱 정교하게 생성형AI 데이터를 전달

주요 기능 및 차별점 – Keysight 시그니쳐, Dynamic Apps, 정규식(RegEx), 커스텀 시그니쳐

- Keysight 시그니쳐

 - AppStack은 정기적으로 애플리케이션 데이터베이스를 업데이트하여 주요 애플리케이션과 새로운 애플리케이션을 추적하고 알려지지 않은 애플리케이션에 대한 시그내쳐를 개발

 - 정기적으로 애플리케이션 시그니쳐 업데이트

- Dynamic Apps

 - 알 수 없는 애플리케이션을 감지하고 요청에 따라 주요 애플리이션을 추가하는 기능.

 - 어플리케이션 시그니처를 파악하고 데이터베이스를 유지 관리하는 AppStack을 사용하면 사용자나 팀이 RegEx 전문가가 되거나 변화하는 어플리케이션을 추적에 소모되는 시간이 감소

AppStack : 4단계 필터링으로 더욱 정교하게 생성형AI 데이터를 전달

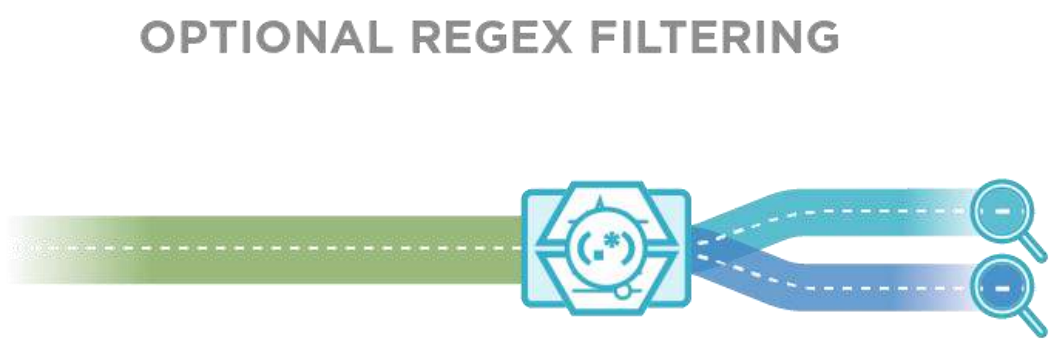
주요 기능 및 차별점 – Keysight 시그니쳐, Dynamic Apps, 정규식(RegEx), 커스텀 시그니쳐

- 정규식(RegEx)

 - 문자열 패턴을 정의하고 이를 기반으로 데이터를 검색하거나 처리하는 강력한 도구

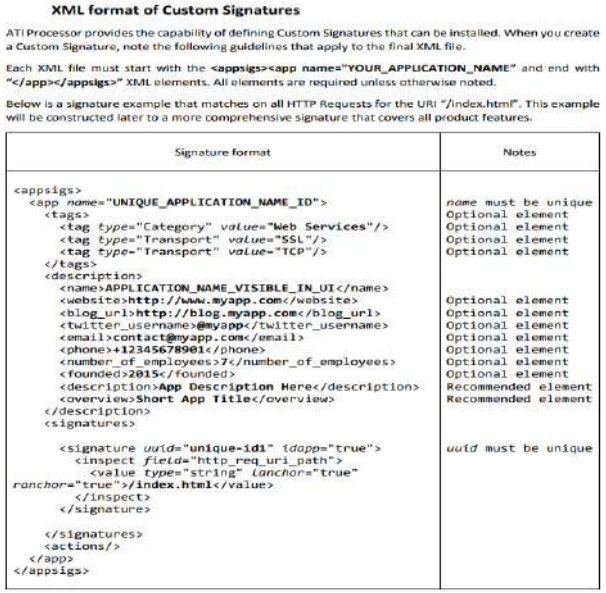
 - 패턴: 매칭하고자 하는 문자열의 규칙을 정의.

예) /₩d{3}-₩d{4}-₩d{4}/는 전화번호 형식을 매칭하는 정규식



- 커스텀 시그니쳐

 - 자체 애플리케이션 시그니쳐를 정의하여 100% 확실하고 간단한 탐지를 보장



XML format of Custom Signatures

ATI Processor provides the capability of defining Custom Signatures that can be installed. When you create a Custom Signature, note the following guidelines that apply to the final XML file.

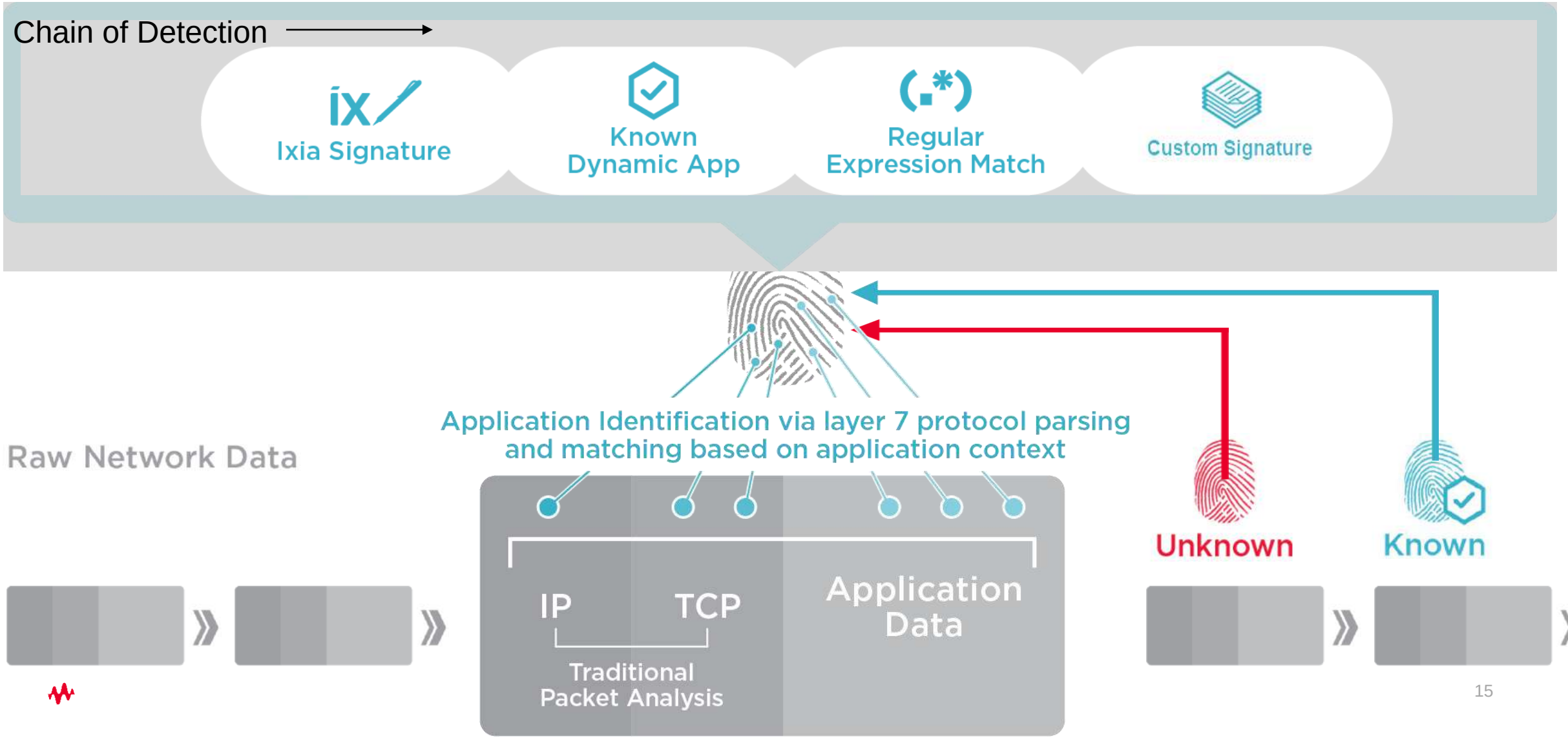
Each XML file must start with the **<appsigs><app name="YOUR_APPLICATION_NAME"** and end with **"</app></appsigs>"** XML elements. All elements are required unless otherwise noted.

Below is a signature example that matches on all HTTP Requests for the URI "/index.html". This example will be constructed later to a more comprehensive signature that covers all product features.

Signature format	Notes
<appsigs> <app name="UNIQUE_APPLICATION_NAME_ID"> <tags> <tag type="Category" value="Web Services"/> <tag type="Transport" value="SSL"/> <tag type="Transport" value="TCP"/> </tags> <description> <name>APPLICATION_NAME_VISIBLE_IN_UI</name> <website>http://www.myapp.com</website> <blog_url>http://blog.myapp.com</blog_url> <twitter_username>@myapp</twitter_username> <email>contact@myapp.com</email> <phone>+12345678901</phone> <number_of_employees>7</number_of_employees> <founded>2015</founded> <description>App Description Here</description> <overview>Short App Title</overview> </description> <signatures> <signature uuid="unique-id1" idapp="true"> <inspect field="http_req_uri_path"> <value type="string" lanchor="true" ranchor="true">/index.html</value> </inspect> </signature> </signatures> <actions/> </app> </appsigs>	name must be unique Optional element Optional element Optional element Optional element Optional element Optional element Optional element Optional element Optional element Optional element Optional element Recommended element Recommended element uuid must be unique

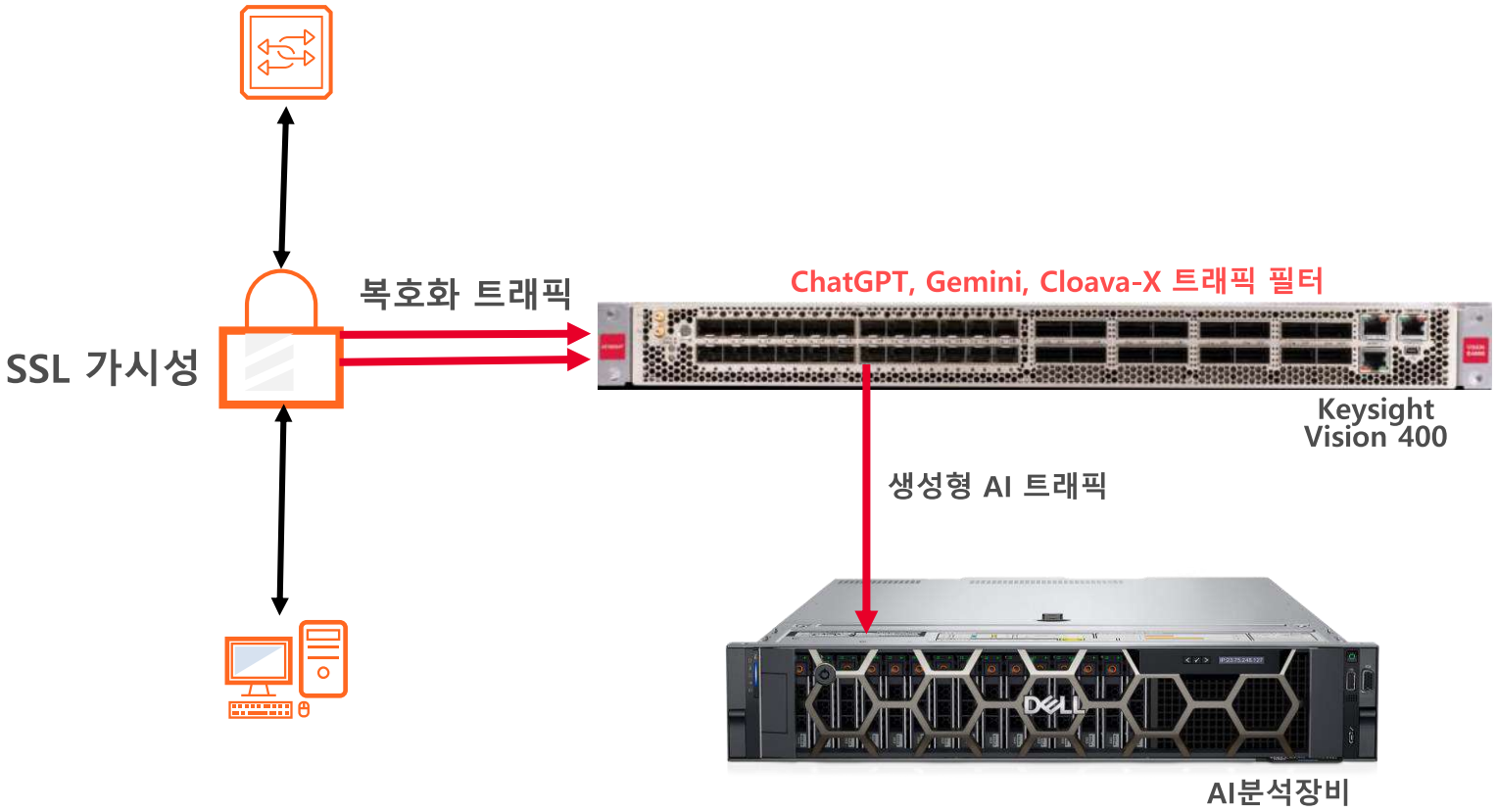
AppStack : 4단계 필터링으로 더욱 정교하게 생성형AI 데이터를 전달

멀티 레벨 애플리케이션 탐지



AppStack 엔진 기반 데이터 필터링

Use Case



Customer

생성형AI 트래픽을 모니터링하고자 하는 고객

Challange

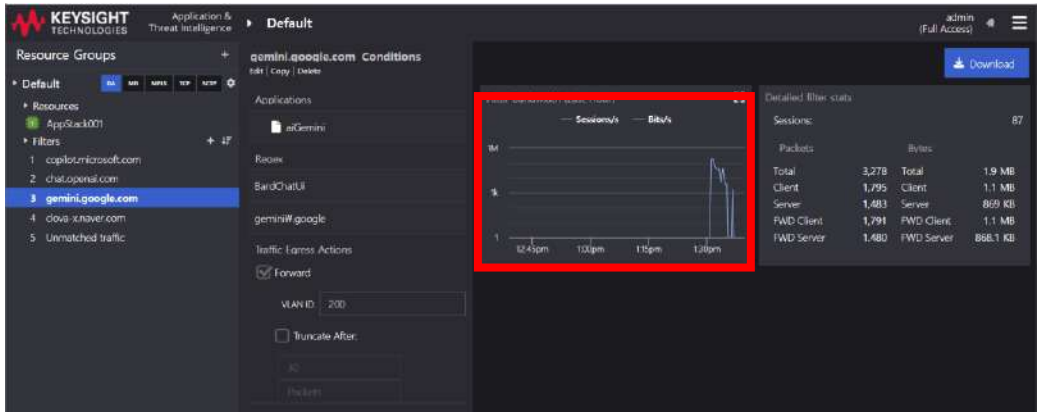
방대한 양의 트래픽으로 인하여 AI분석장비에 많은양의 데이터가 저장되어 스토리지 비용의 증가가 불가피

Requirement

생성형AI 데이터만 전달하여 AI분석장비의 부하감소 및 스토리지 비용 감소에 도움이 되는 솔루션 필요

AppStack 엔진 기반 데이터 필터링

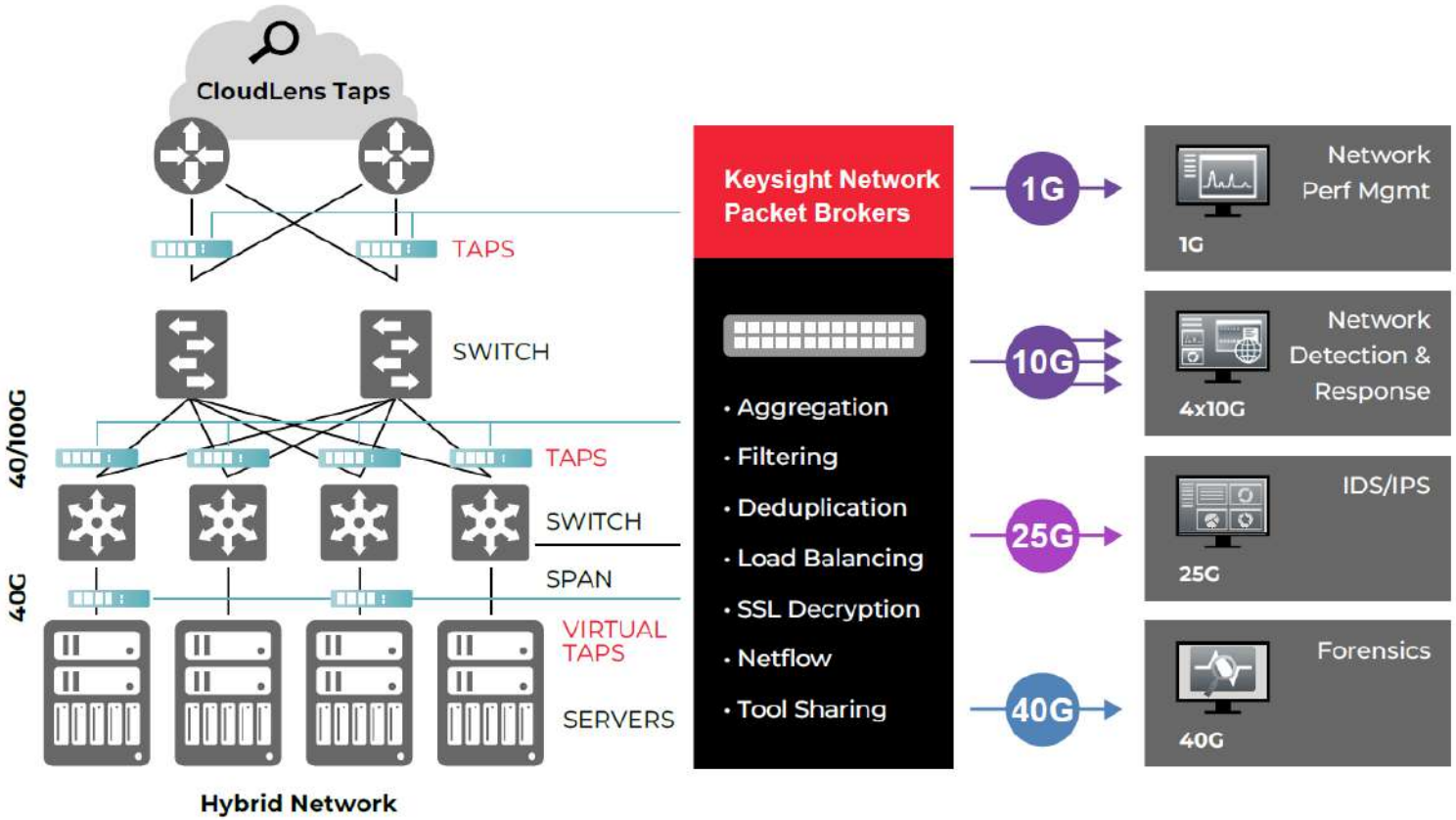
Use Case



- 생성형 AI 데이터를 Dynamic Apps, RegEx(정규식), Custom 시그니쳐 필터 3중 필터 처리로 더욱 정확한 생성형AI 데이터 필터링
- WebUI 상에서 필터링된 생성형AI 트래픽 상태 확인
- 생성형AI 탐지솔루션에 전달된 데이터를 통해 솔루션에서 모든 대화 내용이 확인되며, 대화 내용 중 민감한 정보 또는 개인정보 유출 등의 내용은 이벤트 발생 및 알람 등의 기능으로 사용자에게 전달

Keysight 가시성 솔루션

Network Packet Broker



PERFORMANCE

- 1G~400G 까지의 인프라 환경에서 모든 패킷을 처리하는 고성능 Network Packet Broker
- 다양한 보안/모니터링 툴과 연동 지원

INLINE

- 인라인 보안 네트워크 환경에서 모든 인라인 보안 툴과 연동하여 네트워크 장애 최소화
- Inline Load-Balancing 기능으로 인라인 보안 툴의 병렬 구성 지원 및 부하 분산, 병렬구성의 보안 툴 장애 시 자동으로 트래픽 재분배(Rebalance)
- 특정 패킷만 보안 툴로 전달함으로 인라인 보안 툴의 부하 감소 효과

OUT-OF-BAND

- 보안 인프라팀이 원하는 패킷을 적절한 모니터링 Tool에 전달
- 트래픽 통합(Aggregation), 필터링(L2~L7 Filtering), 중복제거(Deduplication), 부하분산(Load Balancing), SSL 복호화, Netflow, 다양한 패킷 헤더 제거 기능 등의 트래픽 최적화 작업을 거쳐 다양한 보안/모니터링 툴로 패킷 전달
- On-prem 환경의 레거시 트래픽 뿐만 아니라 Cloud 상의 트래픽을 모두 통합하여 최적화 작업을 거쳐 모니터링 툴 장비로 최적의 패킷을 전달

Live 네트워크 가시성 전략

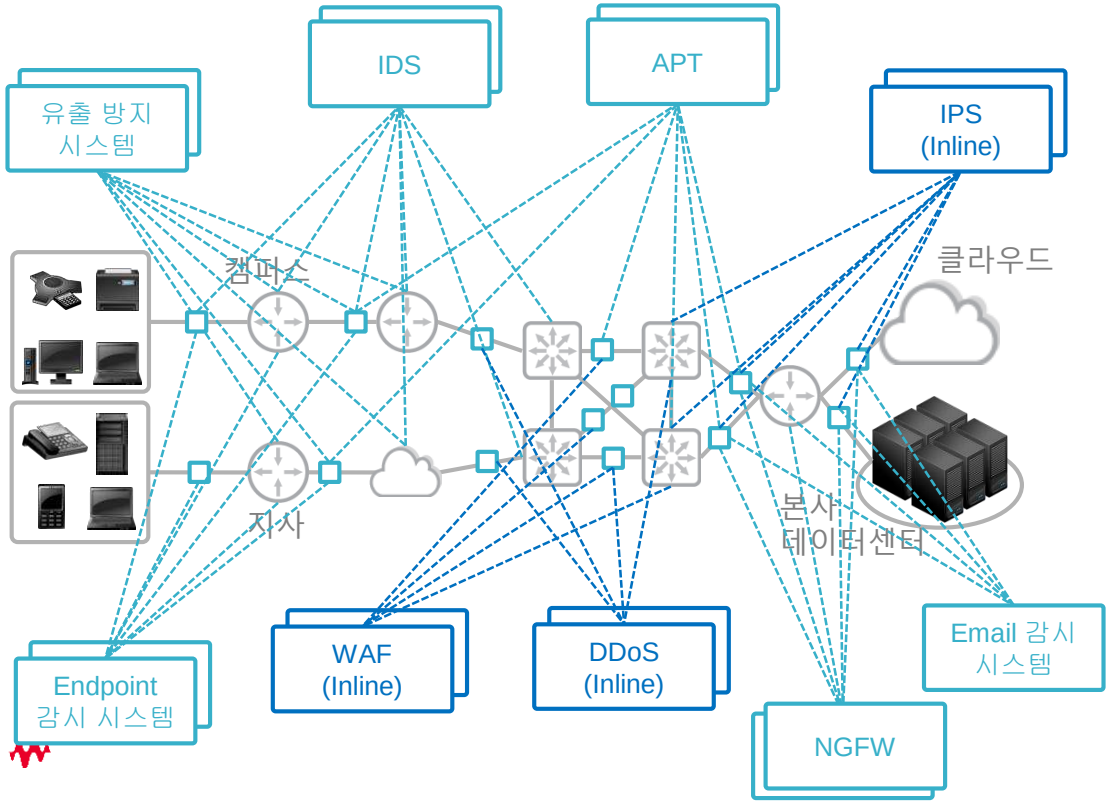
KEYSIGHT

Live 네트워크 가시성 전략

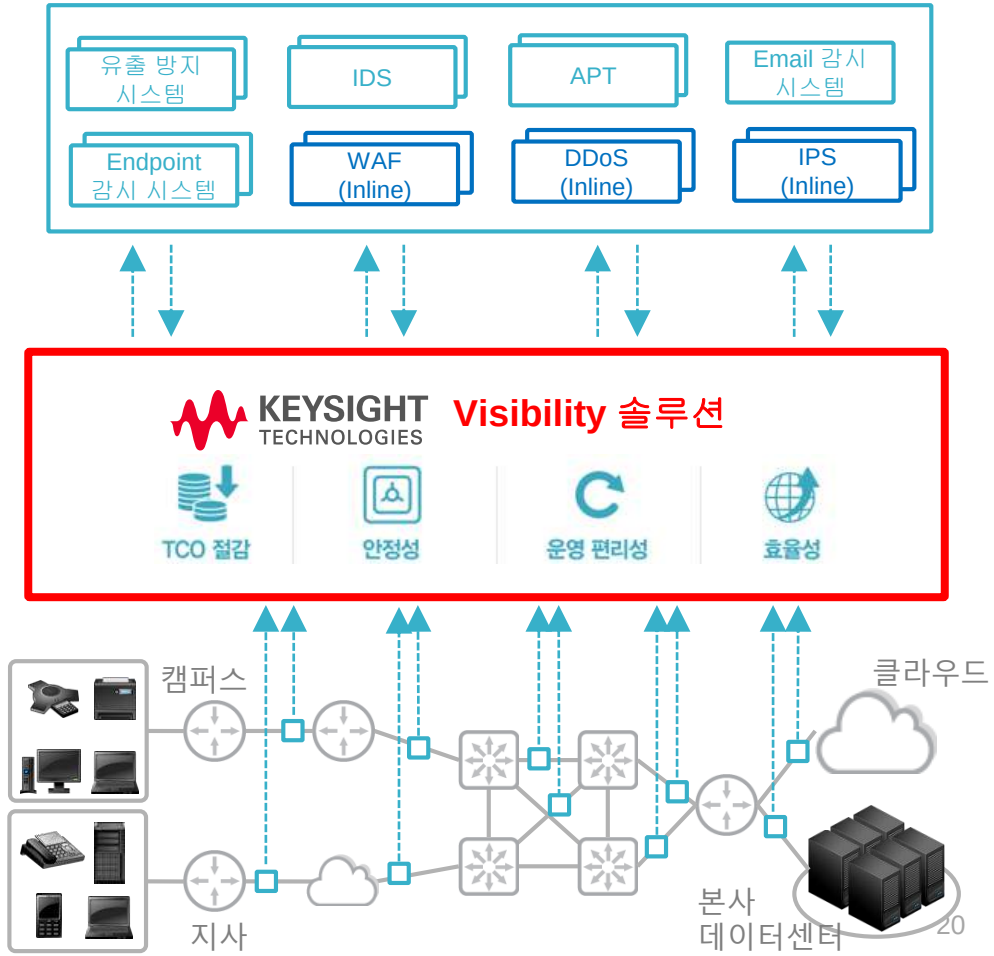
가시성 솔루션 도입 전/후

Visibility(가시성) 솔루션 도입 전

- 복잡성 – 모니터링/보안을 위한 전용 네트워크의 복잡성
- 장애 효율성 – 인라인 툴에 대한 장애 대비 기능 미비
- 성능 저하 – 모니터링 툴로 불필요한 트래픽 유입
- 운영 효율성 – 복잡한 네트워크로 인한 운영 비효율성



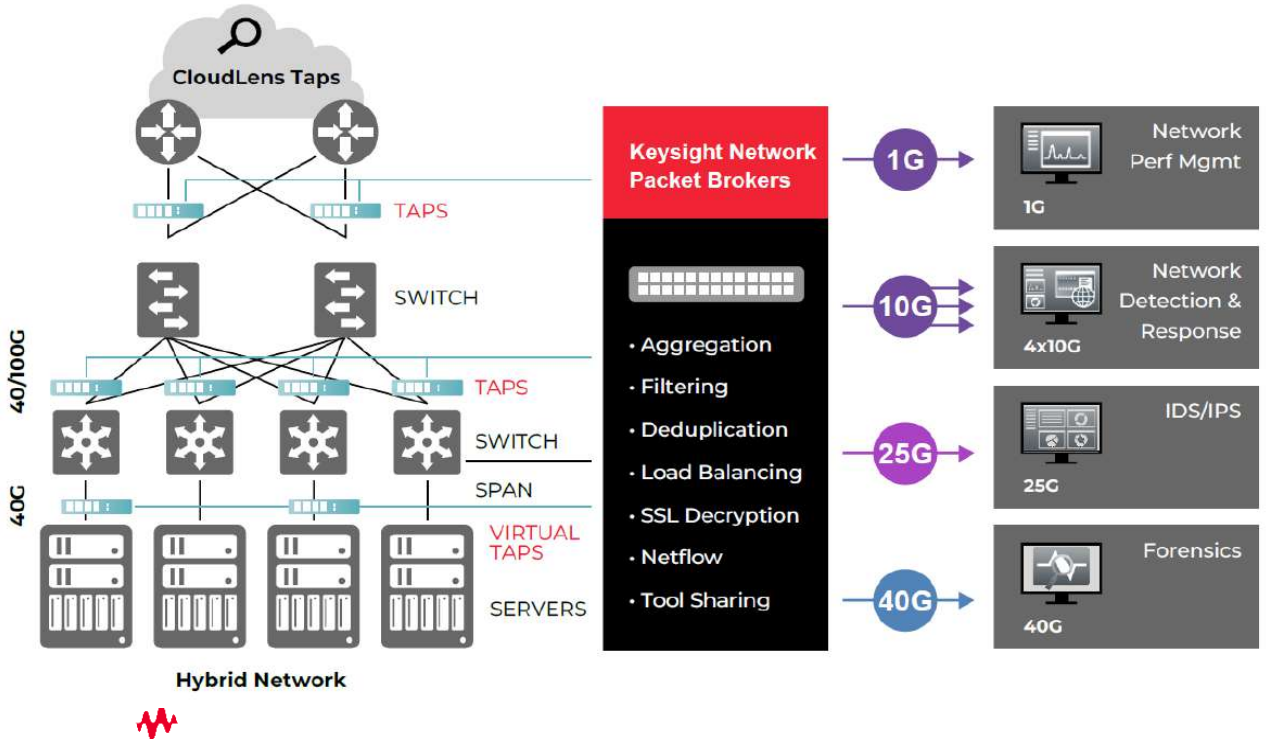
Keysight Visibility 솔루션 도입 후



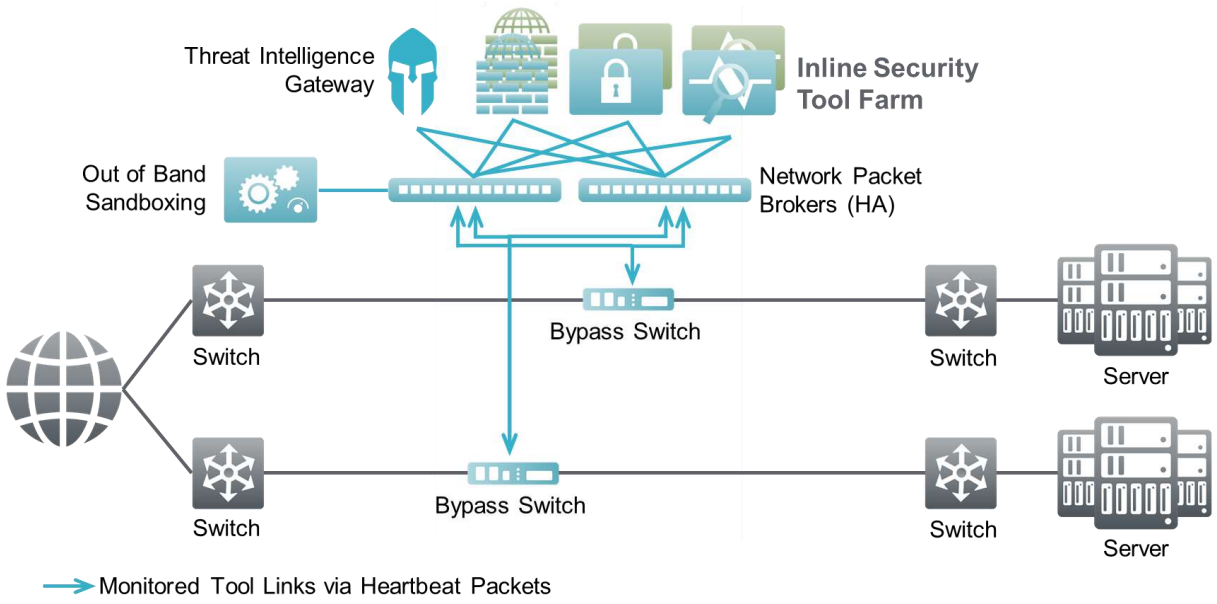
Keysight 가시성 솔루션 아키텍쳐

가시성/보안성의 사각지대를 제거 – 모든 툴이 필요로 하는 데이터를 수신/분석 할 수 있는 인프라를 구성

Out-of-Band
인텔리전트 가시성

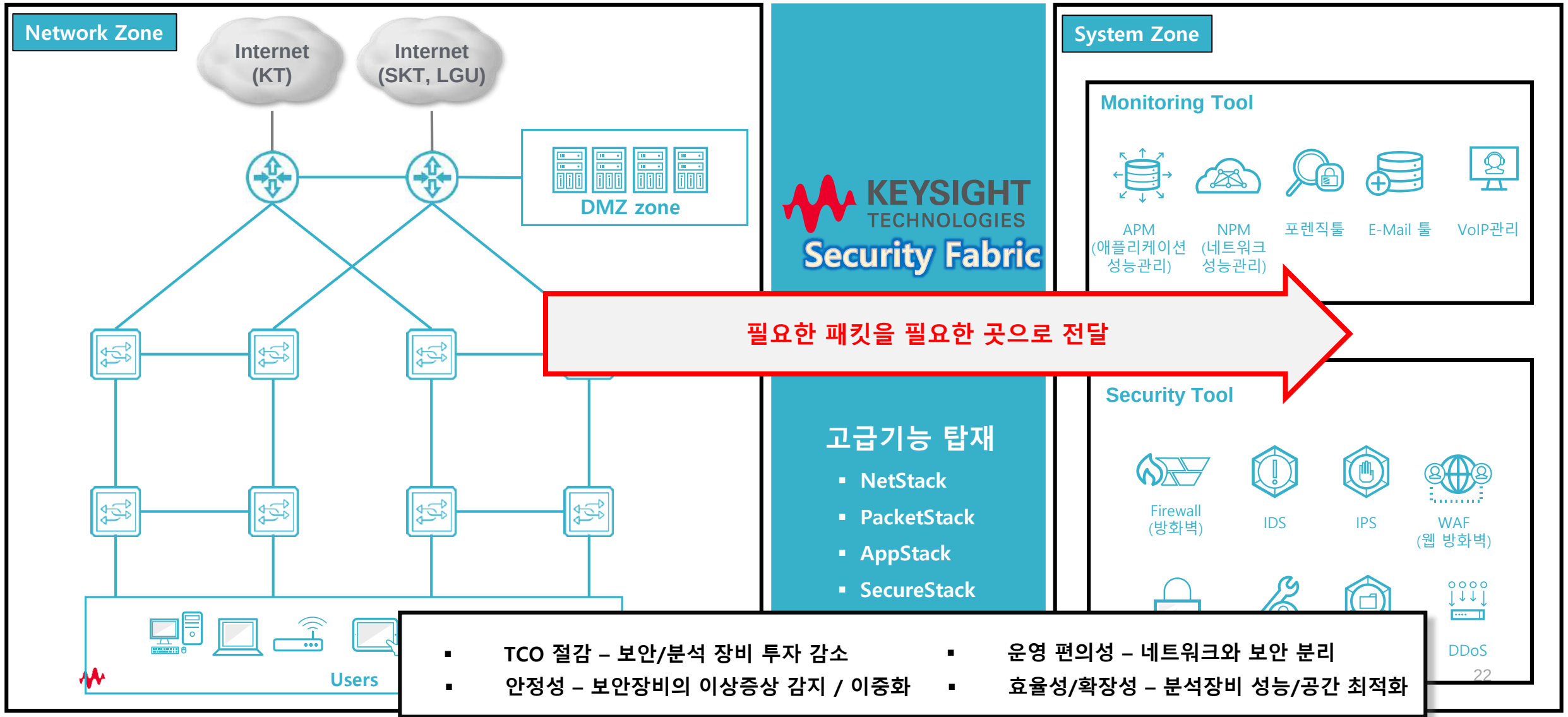


Inline
탄력적인 보안(네트워크 복원력)



Keysight 솔루션 도입 시 특장점

인프라 구성 관점



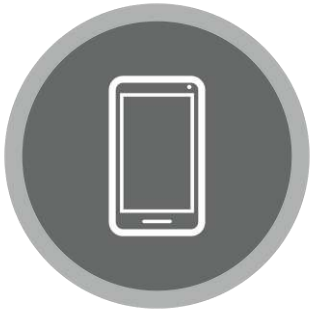
- TCO 절감 – 보안/분석 장비 투자 감소
- 안정성 – 보안장비의 이상증상 감지 / 이중화
- 운영 편의성 – 네트워크와 보안 분리
- 효율성/확장성 – 분석장비 성능/공간 최적화

효율적인 데이터 전달을 위한 패킷 최적화 기능

Packet Broker Software Stacks







NetStack (~NPB Basics)
- 3단계 필터링
- Dynamic Filter 컴파일러
- 포트 두 배로 늘리기
- VLAN 태깅
- Aggregation & Replication
- Load Balancing(부하분산)
- L2GRE/VxLAN 터널링
- 인라인 및 인라인 HA
- IFC 클러스터링

PacketStack (~AFM)
- 트래픽 중복 제거
- 불필요한 헤더 제거
- 패킷 슬라이싱
- 타임 스태핑
- 민감정보 마스킹
- GRE 터널링
- Burst Protection

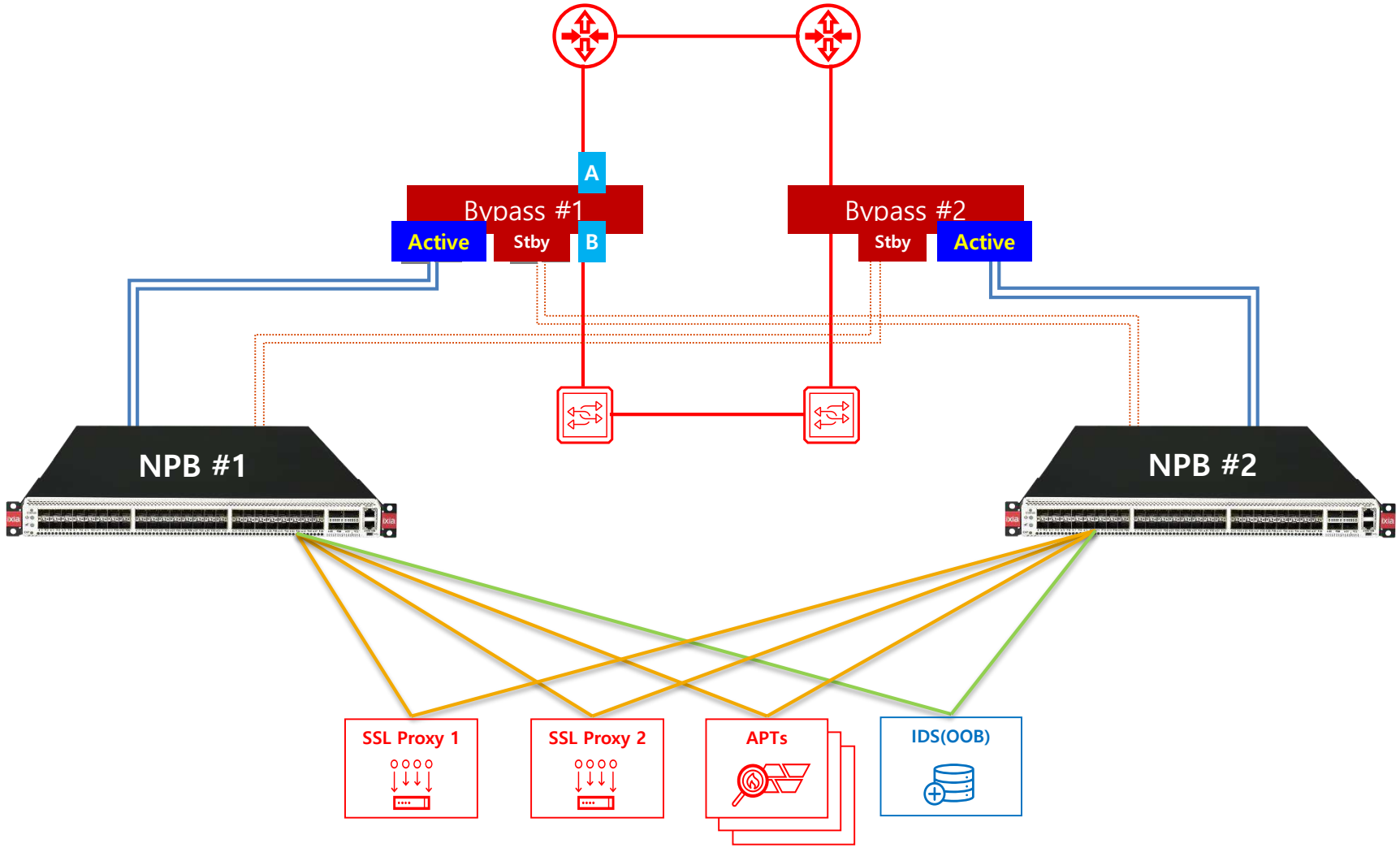
AppStack (~ATIP)
- 앱 탐지 및 필터링
- 위치 정보 및 태그
- 실시간 대시보드
- NetFlow & IxFlow 생성
- RegEx(정규식) 필터링
- Data Masking +
- 패킷 캡쳐

SecureStack (VX/V400)
- Passive SSL Decryption
- Active SSL or man-in-the-middle Decryption

MobileStack (VXV400)
- GTP/SIP Session Correlation
- GTP/SIP Load Balancing
- 가입자 샘플링
- 가입자 필터링
- EPC 필터링

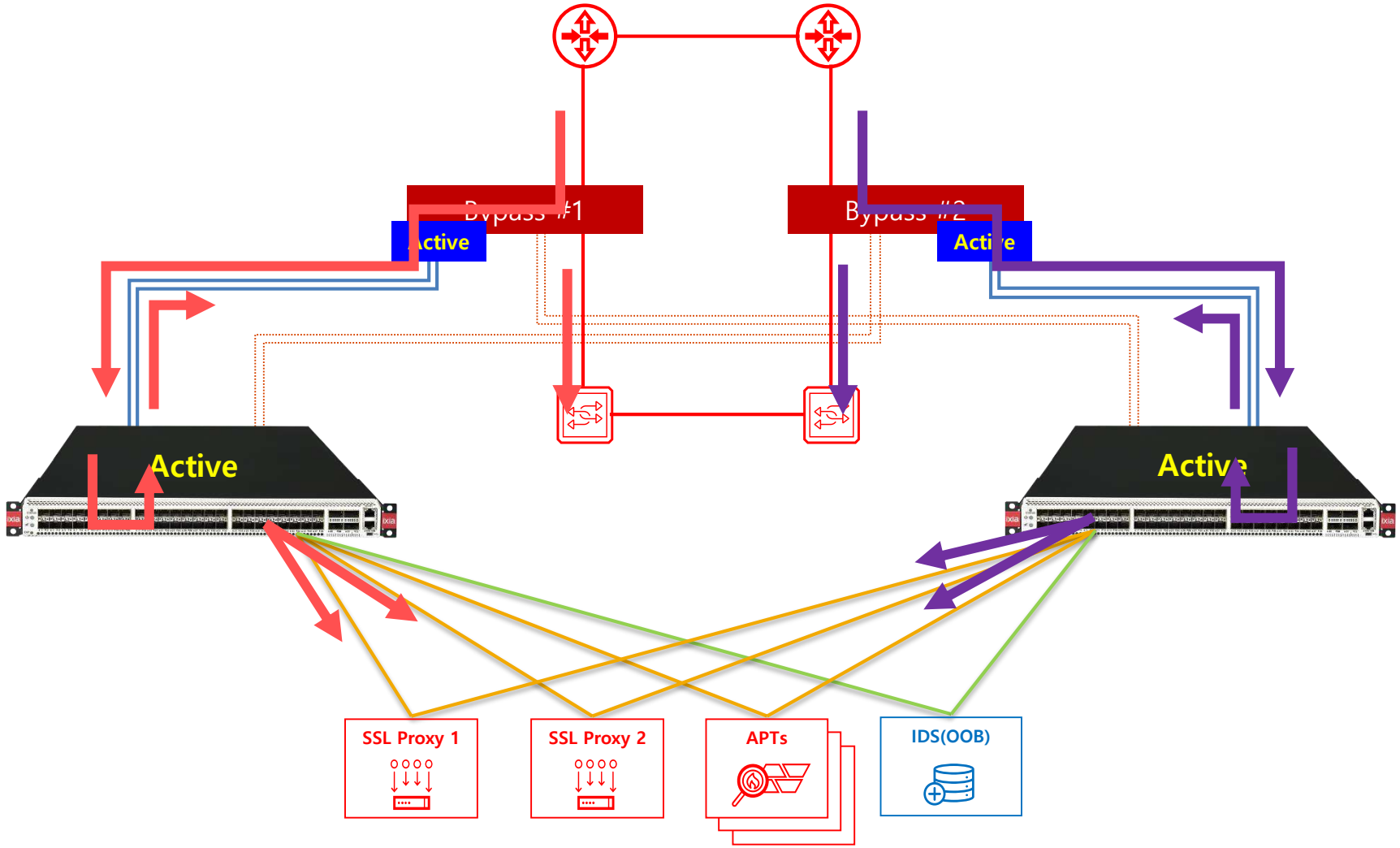
Live 네트워크 가시성 구축 사례

업계 유일의 인라인 Active-Active & Active Standby 통합 Packet Broker



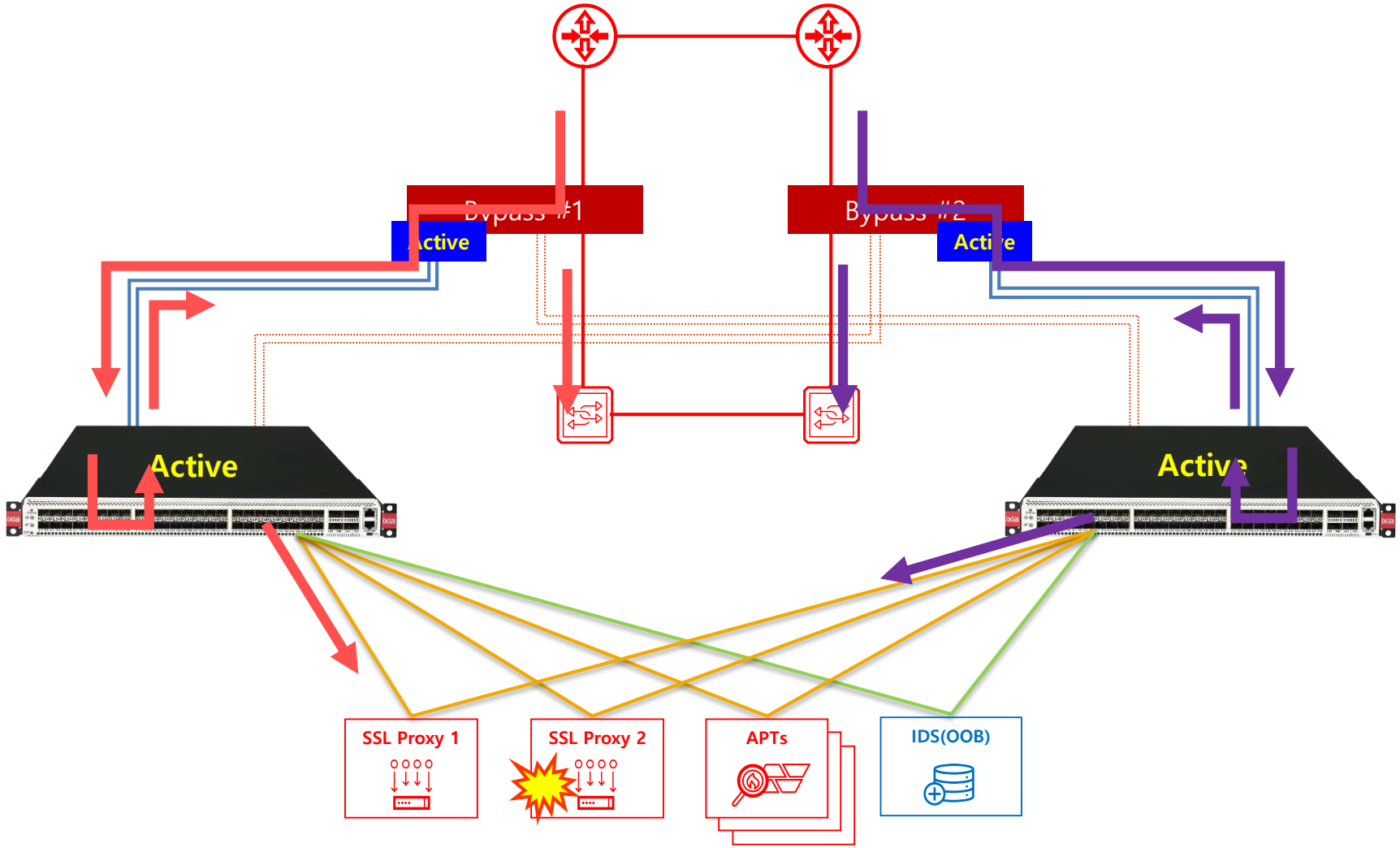
Live 네트워크 가시성 구축 사례

업계 유일의 인라인 Active-Active & Active Standby 통합 Packet Broker



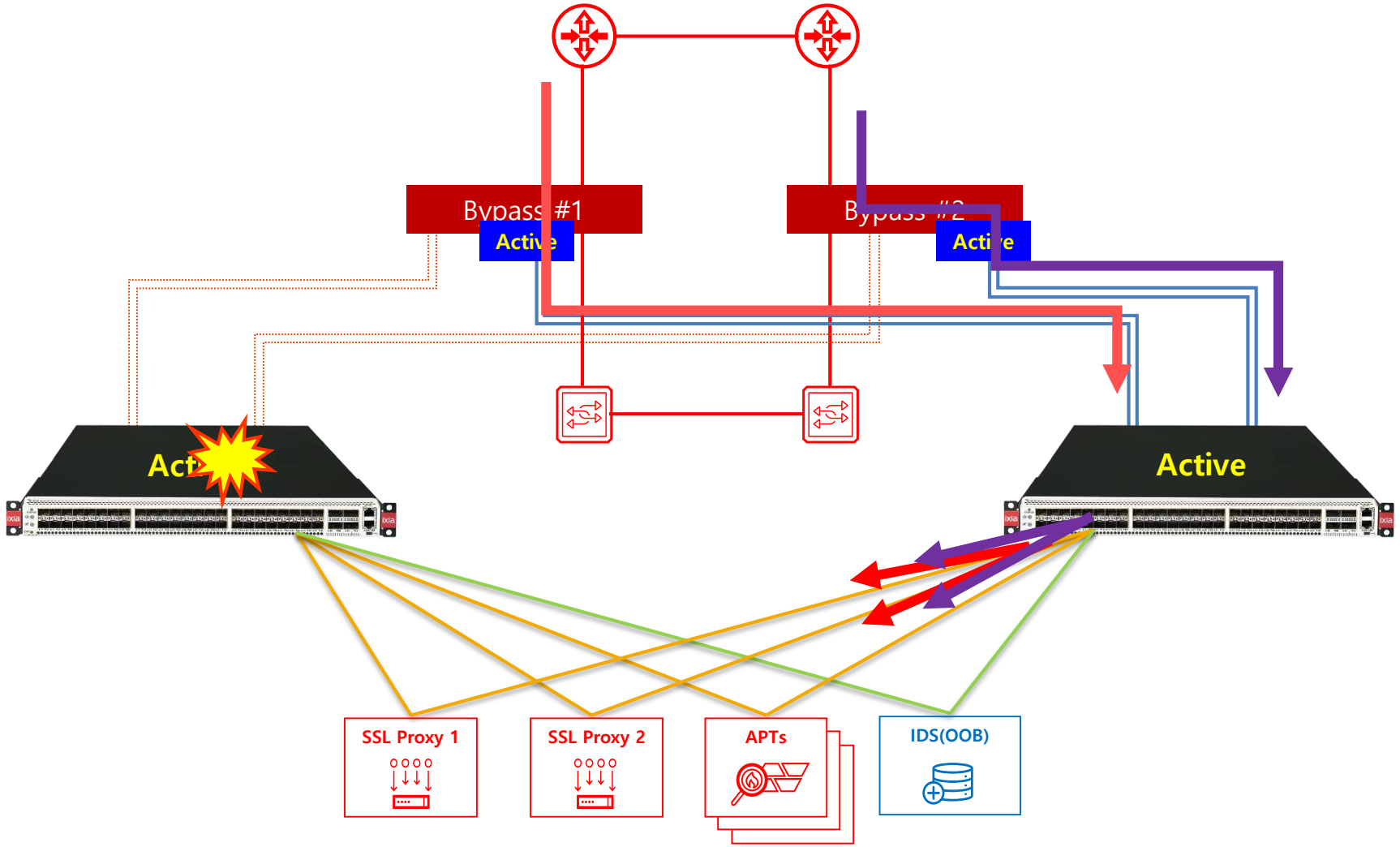
Live 네트워크 가시성 구축 사례

업계 유일의 인라인 Active-Active & Active Standby 통합 Packet Broker



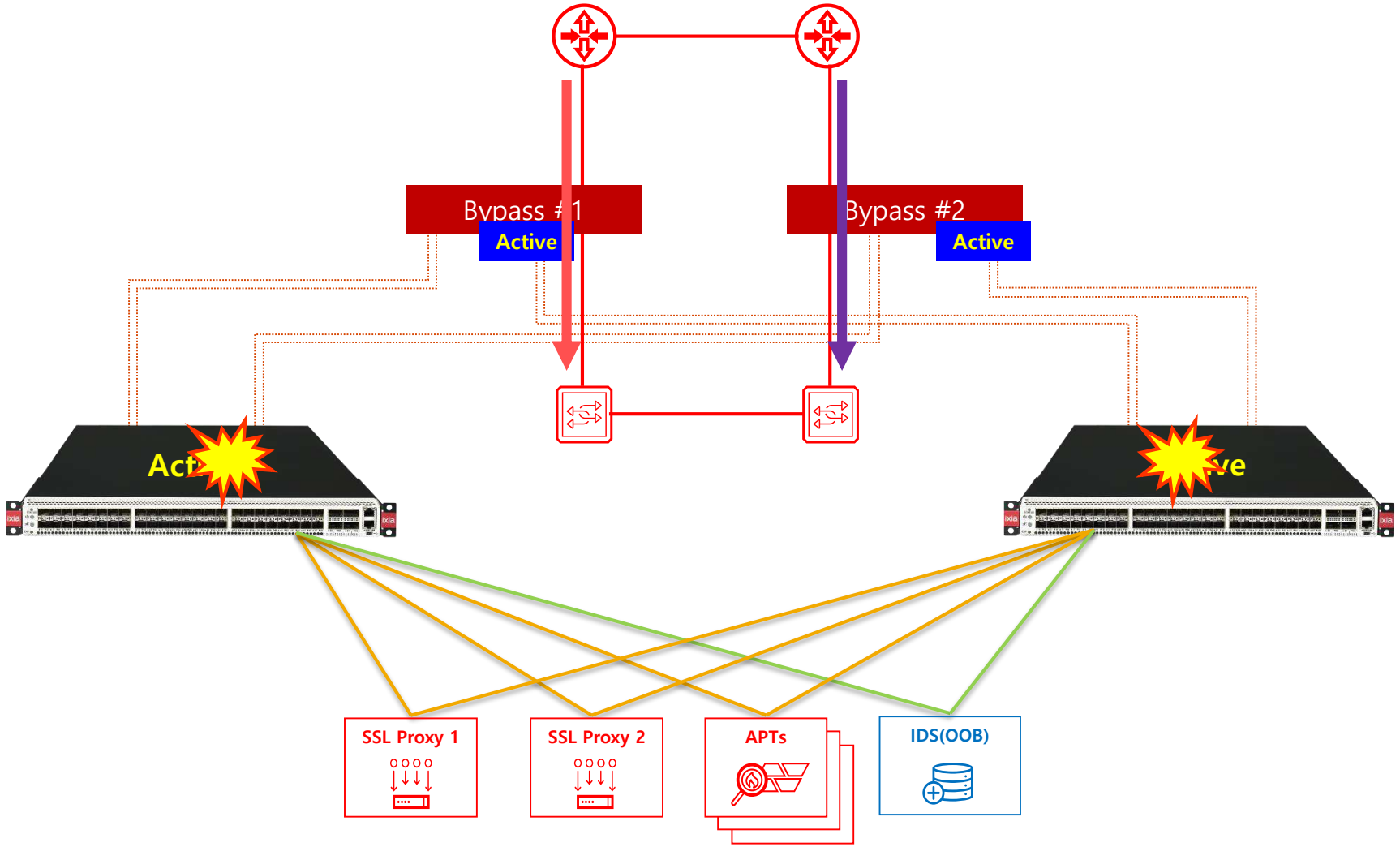
Live 네트워크 가시성 구축 사례

업계 유일의 인라인 Active-Active & Active Standby 통합 Packet Broker



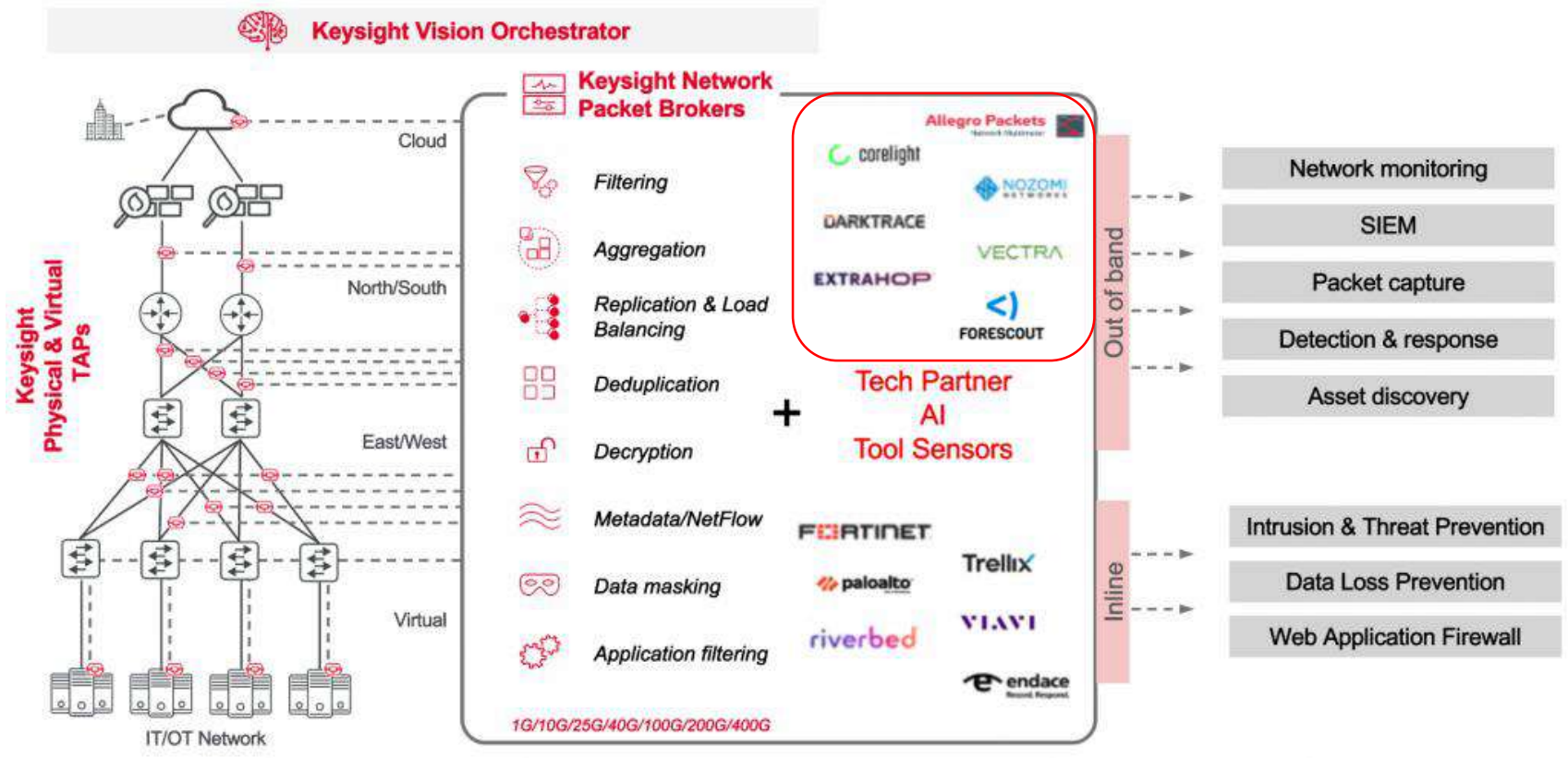
Live 네트워크 가시성 구축 사례

업계 유일의 인라인 Active-Active & Active Standby 통합 Packet Broker



AppFusion 을 통한 통합 가시성 제공

Packet Broker 실행되는 3rd Party 센서/소프트웨어의 데이터에 대한 통합 액세스



1G~400G 까지 지원하는 Keysight 가시성 제품

Network Packet Broker

1G | **10G** | **40G** | **100G** | **400G**

Vision **E1S**
Vision **T1000**

Vision **E10S**

Vision **E40**

Vision **ONE**

Vision **Edge OS +**
White Box Switch

Vision **E100**
Vision **7816**

Vision **X**

Vision **400**
Vision **E400S**

다양한 패킷 브로커를 통해 어디서나 속도, 용량, 지능형 패킷 액세스를 충족


KEYSIGHT