Klastrovanie Presents

KLASTRONODE VERTEX INTRODUCTION

klastrovanie March 2026

이 매뉴얼은 국가 지식재산 거래 플랫폼에 업로드 목적으로 작성되었습니다.

본 기술은 단순한 특허 개념이 아니라, 실제 금융 거래 네트워크 분석을 위해 구현된 AML(자금세탁 방지) 분석 플랫폼입니다.

소프트웨어의 전체 매뉴얼, 설치 패키지, Docker 이미지 및 상세 기술 문서는 지식재산 보호를 위해 공개되지 않습니다.

기술 검토 또는 구매를 희망하는 기관 및 기업은 NDA 체결 후 상세 매뉴얼 제공 및 시스템 시연 예약이 가능합니다.

klastroNode Vertex

Version 0.0.1.0

1. Introduction

klastroNode Vertex는 **klastrovanie**의 **klastroNode** 제품군의 일부로, 금융 커뮤니티 내 구조를 분석하고 식별하도록 설계되었습니다.

- **금융 네트워크 데이터에 대해, klastroNode Vertex는 다음을 파악합니다.:**
    - 커뮤니티 내부의 작은 서브 금융 커뮤니티
    - 전체 금융네트워크를 대표하는 주요 금융 리더

- **klastroNode®: 귀하의 데이터 보안을 보장합니다.**
  - **설계 단계 부터 데이터 프라이버시 보호**
    - **klastroNode Vertex** 는 **온프레미스** 환경에서만 작동하여 귀하의 데이터가 로컬 시스템을 벗어나지 않도록 보장합니다. 모든 처리는 로컬에서 이루어지며, 저희는 어떤 단계에서도 귀하의 데이터에 접근하지 않습니다.

  - **암호화된 열 이름 사용**
    - 사용자가 제공한 데이터는 암호화된 열 이름을 사용하여 처리되며, 데이터 구조나 내용을 알 필요가 없습니다. 이를 통해 민감한 정보가 보호되면서도 안전한 분석이 가능합니다.

  - **라이선스 인증만 필요**
    - **klastroNode Vertex**는 소프트웨어 인증을 위한 라이선스 토큰만 필요로 합니다. 데이터 처리는 이 과정과 완전히 분리되어 있어 데이터가 외부로 전송될 일이 없습니다.

  - **개인 정보 불필요**
    - 사용자는 데이터 **인덱스와 이에 해당하는 가중치만 제공**하면 됩니다. 개인 정보나 민감한 정보를 포함할 필요가 없어 개인 데이터 유출 위험이 완전히 제거됩니다.

2. 설치 환경

이 패키지는 실행 중 토큰 기반 인증이 필요합니다. 토큰 유효성 검증 및 필요한 리소스에 접근하기 위해 활성화된 **인터넷 연결이 필수**입니다.
klastroNode Vertex를 사용하기 위해서 다음의 소프트웨어 및 하드웨어 환경이 요구됩니다.

✻ 필수 소프트웨어

1. Docker (컨테이너 작동환경에 필수)
2. NVIDIA Docker Toolkit (nvidia-docker2) 컨터이너 내의 GPU 가속 지원
3. CUDA 12.9 호환 되는 드라이버 (GPU 기반의 가속 연산)
4. Ubuntu 24.04 / Windows WSL2 Ubuntu 24.04

✻ 지원되는 GPU 리스트

만약 CUDA 12.9 호환되는 GPU 장비가 호스트 머신에서 발견이 안된 경우, klastroNode®는 에러를 발생합니다.
아래 해당 되는 NVIDIA GPU 장비들은 지원 됩니다.

❖ 전문가 시리즈 (추천)

- NVIDIA RTX A 시리즈 (RTX A6000, RTX A5000, RTX A4000, etc)
- NVIDIA Quadro 시리즈 (Quadro RTX 8000, Quadro RTX 6000, Quadro RTX 5000, etc)
- NVIDIA Tesla 시리즈 (Tesla V100, Tesla P100, etc)

❖ 일반 사용자 시리즈

- NVIDIA Geforce RTX 시리즈 (RTX 5090, RTX 4090, RTX 4080, RTX 4070, RTX 4060, RTX 4050, RTX 3090 Ti, RTX 3090, etc)
- NVIDIA Geforce GTX 시리즈 (GTX 1650, GTX 1660Ti, etc)

3. klastroNode Vertex 도커별 역할

klastroNode Vertex는 분산형 컨테이너 아키텍처를 통해 그래프 기반 금융 데이터에서 실시간 이상 징후를 탐지할 수 있도록 합니다.
각 컨테이너는 프로파일링, 정밀 분석, 모니터링, 시각화 등 특정 작업에 최적화되어 있으며, 모든 포트가 열려 있어야 원활히 작동합니다.

Docker Name	Role	Description	Port	Computing Unit
Vertex	그래프 프로파일러	금융 거래 데이터를 기반으로 초기 그래프 기반 커뮤니티 탐지 및 그래프 분석을 수행합니다.	**58000**	**GPU**
Edge	집중 분석기	특정 고위험 커뮤니티에 대한 2차 정밀 분석을 수행합니다. 초기 점수가 의심스러운 경우에만 실행됩니다.	**58001**	**GPU**
Watcher	모니터링 에이전트	들어오는 결과를 지속적으로 모니터링하며, 필요 시 edge 컨테이너를 실행하는 컨트롤러 역할을 합니다.	**58002**	**CPU**
Dashboard	시각화 인터페이스	vertex 및 edge 컨테이너가 생성한 결과를 실시간으로 시각화하여 커뮤니티 구조, 네트워크 중심성, 거래 흐름 등의 정보를 보여줍니다.	**58003**	**CPU**
Report	자동 보고서 작성	**klastroNode Vertex는** 다층 송금 네트워크 분석을 통해 주요 송금 허브의 위험도 랭킹을 산출하고, 상위 노드에 대한 상세 분석 보고서를 자동 생성합니다.	**58004**	**CPU**

4. 분석 지표 (Metrics)

klastroNode Vertex 는 네트워크 구조를 평가하고 심층 분석 여부를 판단하기 위해 다양한 지표를 조합하여 사용합니다.

핵심 분석 지표 (Key Metrics):
- Degree Centrality (정점 중심성)
- Network Cohesion Scores (네트워크 응집도 점수)
- Network Centrality Indicator (네트워크 중심성 지표)

🔍 Degree Centrality 란?

Degree Centrality(정점 중심성)는 특정 금융 계정(또는 지점, 사용자 등)이 직접적으로 연결된 거래 상대방의 수를 측정하는 지표입니다.
금융 자산 추적의 관점에서, 이 값이 높다는 것은 해당 노드가 다수의 거래 흐름의 중심에 있으며, 허브 또는 분산 계좌로 기능할 가능성이 높다는 것을 의미합니다.

- **높은 Degree Centrality**는 다음과 같은 역할을 시사합니다:
  - 다수의 계좌와 빈번하게 거래하는 중개 계좌 또는 라우팅 허브
  - 자금 흐름의 분기 또는 집결 지점
  - 의심 거래가 집중되는 비정상적 계정 구조
- 이 지표는 **vertex**의 초기 분석 단계에서 의심도 높은 노드 식별에 활용되며, 필요 시 **Edge**에서 심층 재분석 대상으로 자동 전환될 수 있습니다.

🔍 Network Cohesion Score 란?

Network Cohesion Score(네트워크 응집도 점수) 는 특정 금융 거래 네트워크 내에서 자금 흐름이 얼마나 응집된 구조로 형성되어 있는지를 정량적으로 평가하는 지표입니다.

- 🔥 **0.8 초과: 아주 의심스러운 네트워크 응집력 수준:**
  커뮤니티 내부 간 연결이 매우 강하고 외부 노드와의 상호작용이 거의 없는 경우를 나타냅니다.
  - 실제 금융 네트워크에서 이러한 수준의 응집성 (cohesion)은 다음과 같은 정황을 시사할 수 있습니다:
    - 폐쇄형 루프 운영
    - 유령 회사 또는 사기 조직
    - 자금 세탁 활동
    - 인위적으로 구성된 커뮤니티

    📌 **권장 사항**
    - 해당 결과를 **고위험**으로 표시하세요.
    - **Edge 컨테이너 혹은 Vertex로 얻어진 정보를 수집하세요.**
    - **고위험 전용 테이블에 저장하여 감사 및 교차 확인에 대비하세요.**
    - **관련 계좌 동결 조치 또는 자동화된 규제 보고서 (예: STR – 의심거래보고서) 생성도 고려하세요.**

- **1.0**: 이 값은 이론적으로 가능한 최대 점수로, 외부 연결이 전혀 없는 완전히 독립적인 커뮤니티를 나타냅니다.
  이러한 구조는 실제 금융 시스템에서는 극히 드물거나 존재하지 않으며, 관측될 경우 상당히 인위적이거나 의심스러운 구조로 간주됩니다.
- **0 또는 0에 가까운 값**: 랜덤 그래프를 나타냅니다. 구조적 패턴이 거의 없으며 무작위 연결을 가진 네트워크입니다.
- **0 미만**: 응집도가 낮은 커뮤니티를 나타냅니다. 데이터셋 내 모든 개체가 다른 개체와 유의미한 상호작용이 없음을 의미 합니다.

- **0.5 초과, 0.8 미만**: 양호한 네트워크 응집력 점수로, 강한 커뮤니티 구조를 의미하며 실제 사회에서도 발견되는 패턴입니다. 점수가 0.5에서 0.8 사이일 경우, 추후 조사를 위해 해당 정보를 별도의 데이터 테이블에 저장해두는 것을 권장합니다.

- **0.1 초과 0.5 미만**: 중간 수준의 응집도를 나타내며, 이는 겉으로 드러나지 않은 협조 관계나 의심스러운 네트워크가 초기 단계에서 구조화되고 있을 가능성을 시사합니다. 이 경우, **klastroNode Vertex**는 **Watcher** 시스템을 자동으로 트리거하며, 이후 **Edge** 컨테이너가 활성화되어 보다 심층적인 조사가 수행됩니다. 이를 통해 완전히 응집된 구조가 형성되기 이전 단계에서부터 잠재적인 의심 활동을 선제적으로 추적할 수 있습니다. 사용자는 이 분석을 트리거하는 하한 및 상한 범위를 자유롭게 조정할 수 있습니다.

🔍 Network Centrality Indicator란?

Network Centrality Indicator(네트워크 중심성 지표) 는 전체 금융 거래 네트워크 내에서 **비정상적으로 큰 영향력을 가진 계정이나 주체**를 식별하는 데 사용됩니다. 단순한 거래량이나 연결 수(degree)와는 달리, 이 지표는 **다른 고중심성 노드들과의 연결을 통해 자금 흐름을 통제하거나 중개하는 역할을 하는 노드**를 강조합니다.

금융 네트워크에서 높은 중심성 지표는 어떤 의미일까요?
- **중앙 자금 라우터**: 해당 노드는 네트워크의 서로 다른 영역 사이에서 **자금을 전달하는 중간 경로**로 자주 활용되며, **전체 자금 흐름에서 핵심적인 중개 역할**을 합니다.
- **간접적 영향력**: 본인이 직접 많은 금액을 거래하지 않더라도, **영향력 있는 다른 계정들과 연결되어 있음으로써 실질적인 영향력이 매우 큰 계정**일 수 있습니다.
- **자금세탁의 연결 고리**: 중심성 높은 노드는 **세탁 과정의 여러 단계를 연결하는 다리 역할**을 하며, 자금의 **흐름을 복잡하게 분산시켜 추적을 어렵게 만듭니다.**
- **잠재적 차단 지점**: 이러한 노드를 모니터링하거나 차단함으로써, **의심 자금의 추적 가능성 또는 유통 자체에 중대한 영향을 줄 수 있습니다.**

금융 범죄 탐지에서 왜 중요한가요?
- **자금 세탁 중개자 식별**: 최종 수취인이 아닌, **불법 자금 분산·세탁에 핵심적인 역할을 하는 계정**을 찾아냅니다.
- **비가시적 리더 식별**: 거래량이 많지 않더라도 구조적으로 중요한 위치에 있는 숨겨진 영향력자를 식별할 수 있습니다.
- **감사 및 자원 집중 우선순위 설정**: 법 집행기관 또는 내부 감사팀이 **중요도가 높은 노드를 우선적으로 모니터링 또는 동결**할 수 있도록 합니다
- **조기 경고 신호 제공**: **중심성 지표가 급격히 상승**하는 경우, 이는 **신규 세탁 경로의 구축**이나 **페이퍼 컴퍼니 활동 시작**을 암시할 수 있습니다.

✅ Vertex 컨테이너만 사용하는 고객용 고효율 운용 지침

상시 모니터링 권장 항목:

• network_centrality_indicator 기준 상위 10%에 해당하는 핵심 계좌 또는 노드는 항상 주의 깊게 감시하십시오.

• 이 상위 노드들과 직접 또는 간접적으로 연결된 클러스터(동일 커뮤니티 ID) 내 계좌들을 우선 감사 대상으로 삼는 것이 효과적입니다.

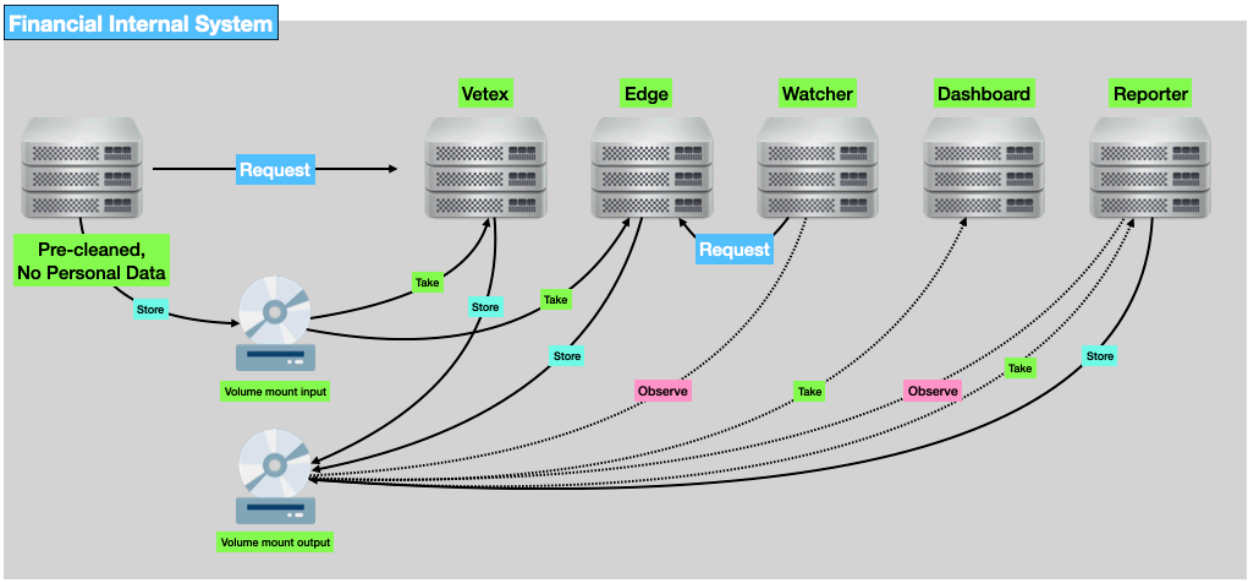
이유:

• 이들은 네트워크 내 자금 흐름에서 중추적 역할을 수행하며,

• 실제 금융 이상 탐지 사례에서도 이 상위 10%의 노드가 전체 리스크의 대부분을 유발합니다.

• 클러스터 단위로 접근할 경우 의심 트랜잭션의 맥락을 구조적으로 파악할 수 있습니다.

5. 시스템 구조 다이어그램



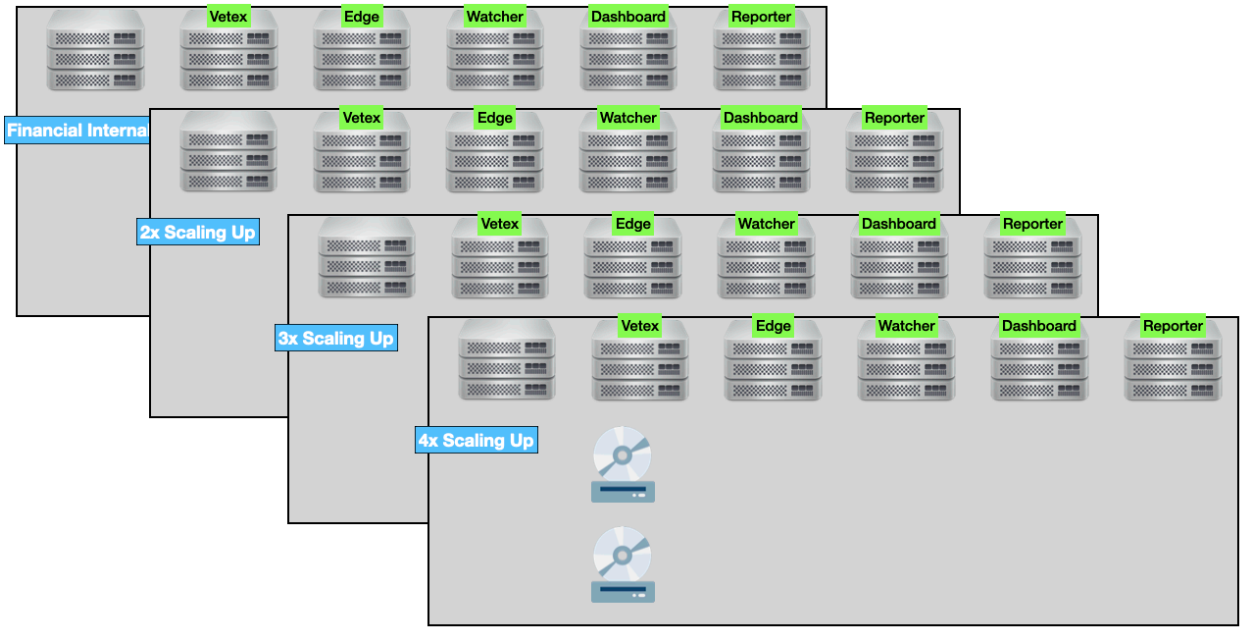
klastroNode Vertex 아키텍처는 보안성, 모듈화, 그리고 실시간 그래프 기반 분석을 통해 금융 거래 데이터를 효과적으로 처리할 수 있도록 설계되어 있습니다.

개인 정보가 제거된 사전 정제된 데이터셋은 공유 볼륨을 통해 마운트됩니다. 워크플로우는 금융 내부 시스템이 Vertex 컨테이너에 요청을 보내면서 시작되며, Vertex는 초기 커뮤니티 탐지 및 그래프 알고리즘 기반 프로파일링을 수행합니다. 이상 징후가 감지될 경우, Watcher 컨테이너가 결과를 관찰한 후 고위험 그룹에 대한 집중 분석을 위해 Edge 컨테이너를 트리거합니다.

판매자는 제공되는 컨테이너 소스에 판매자가 사용중인 인증 코드 모듈을 내장할 수 있으며, 판매자 자체의 JWT 기반 인증 및 자체 인증을 사용하여 판매자가 사용하는 자체 폐쇄 클라우드로 확장사용이 가능합니다.

Dashboard는 공유 출력 볼륨에서 결과를 지속적으로 읽어들여 커뮤니티 구조, 영향력, 거래 흐름에 대한 실시간 시각화를 제공합니다.

6. 수평 확장 아키텍처 및 처리 능력



klastroNode Vertex 시스템은 **수평 확장 (Horizontal Scalability)** 을 고려해 설계되었습니다.
각 유닛(Vertex, Edge, Watcher, Dashboard)은 **독립적으로 복제**할 수 있어, 더 큰 데이터셋이나 높은 분석 수요를 유연하게 처리할 수 있습니다.

예를 들어, 거래 데이터를 0–10만, 10만–20만처럼 **서브셋으로 나누어 각 유닛 세트에 할당함으로써 시스템을 수평적으로 확장**할 수 있습니다.
이 아키텍처는 **병렬 처리 파이프라인**을 지원하며, 추가된 각 유닛은 **서로 다른 데이터 구간**을 분석하여, 금융 기관이 대규모 거래 그래프를 **실시간**으로 처리할 수 있도록 돕습니다.

또한, **무상태 (stateless) 컨테이너 설계를 채택**하여, 각 Vertex–Edge–Watcher–Dashboard-Reporter 세트는 **독립적으로 작동**하며, 입력 및 출력은 **분리된 볼륨 마운트**를 사용합니다. 입력 파일을 읽고, 출력 파일을 남기며 종료하는 stateless 구조는 향후 **동일 입력에 대해 언제든지 동일한 결과를 재현할 수 있기 때문에, 감사 추적성 및 법적 증거 능력을 보장**하는 핵심 요소입니다.

만약 시스템을 확장하는 과정에서 **대시보드를 제거하거나 시각화 방식을 직접 커스터마이징**하고 싶다면, 자유롭게 가능합니다.
모든 분석 결과는 **JSON 및 CSV 형식으로 제공**되므로, 고객은 **자체 대시보드를 구축**하거나 **기존 내부 시스템에 통합**할 수 있습니다.

단일 처리 단위는 Vertex 컨테이너와 Edge 컨테이너로 구성된 분석 노드이며, 각 노드는 16GB VRAM GPU 환경에서 약 30초당 100,000건의 거래 데이터를 처리할 수 있습니다.

이는 단일 노드 기준으로 연간 약 1,000억 건 이상의 거래 데이터를 처리할 수 있는 수준입니다.

또한 본 시스템은 컨테이너 단위로 노드를 추가하여 확장할 수 있으므로,
고가의 엔터프라이즈 GPU 장비 없이도 **수평 확장을 통해 연간 조 단위 규모의 거래 처리 능력을 확보할 수 있습니다.**

7. 대시보드

대시보드는 분석 결과에 대한 **종합적인 시각화 인터페이스를 제공**합니다. 초기 프로파일링 단계에서 계산된 **Initial Network Cohesion Score**를 표시하며, 이 값은 네트워크의 구조적 결합력을 나타냅니다. 만약 의심스러운 금융 네트워크가 탐지되면, Watcher 컨테이너가 자동으로 Edge 컨테이너를 호출하여 2차 정밀 분석을 수행합니다. 이 과정에서 산출되는 Focus Network Cohesion Score는 해당 네트워크에 대해 **추가 조사나 규제 보고(예: STR)**가 필요한지를 판단하는 데 도움을 줍니다. 또한, 클러스터 분석 과정에서 Edge 모듈은 단순히 전역(Global) 리더 노드만 찾는 것이 아니라, 각 **고위험 커뮤니티 내의 로컬 리더(Local Leader)**도 함께 식별합니다. 이 정보는 수사기관이나 금융기관이 중점적으로 조사할 주요 계좌 목록을 확보하는 데 유용하게 활용됩니다. 마지막으로, 금융 기관은 결과 출력 볼륨 마운트에 저장된 JSON 파일을 내부 서버로 전송함으로써, 모든 분석 결과를 자체 보관 및 후속 처리할 수 있습니다.

klastroNode Vertex는 **localhost:58001**에서 접근할 수 있는 분석 대시보드를 제공합니다. 만약 klastroNode Vertex와 대시보드가 사용자 본인의 머신이 아닌 다른 머신에서 실행되고 있다면, 사용자는 SSH 포트 포워딩을 이용하여 안전하게 대시보드에 접근할 수 있습니다.

Shell

```
sudo ssh -L 58002:localhost:58002 -i <pem-key> user-account-id@remote-machine-ip-address
```

이 명령어를 실행한 후, 사용자는 웹 브라우저에서 다음 주소를 열어 대시보드에 접근할 수 있습니다.

Web

```
http://localhost:58002
```

<pem-key>를 개인 키 파일의 경로로 변경하십시오 (필요한 경우).

user@<remote-machine-ip>를 실제 사용자 이름과 원격 머신의 IP 주소로 변경하십시오.

SSH의 -L 플래그는 로컬 포트 포워딩을 설정하여 사용자가 원격 대시보드를 안전하게 자신의 로컬 머신으로 터널링할 수 있도록 합니다.

대시보드에서 대용량 데이터 처리

- **대용량 Parquet 파일에 대한 안내**

입력된 Parquet 파일이 100,000개 이상의 행을 포함할 경우, 대시보드는 자동으로 샘플링을 적용하여 성능을 최적화합니다.

- 데이터셋이 100,000개 이상의 행을 포함하는 경우, 시각화를 위해 일부 데이터만 표시됩니다.
- 데이터셋이 100,000개 이하인 경우, 전체 데이터셋이 샘플링 없이 사용됩니다.

이 샘플링 과정은 대규모 데이터셋이 브라우저를 멈추거나 대시보드 성능을 저하시킬 위험을 방지합니다.

전체 데이터셋을 활용한 상세한 시각화가 필요한 경우, 외부 BI 도구(예: Tableau, Power BI) 또는 별도의 분석 파이프라인을 사용하여 처리하는 것이 권장됩니다.

대시 보드 스크린 샷


KlastroNode Dashboard
Select Result Folder
2025-03-27-12-10-47 → transactions-2025-02-01-08-30-00.parquet
Select Result Folder (Focus)
2025-03-27-12-10-51 → transactions-2025-02-01-08-30-00.parquet
All Communities
CommunityGroupID
vertex
Distribution of Network Centrality
count
Network Centrality Indicator
Transaction Flow
Network Centrality Distribution
Network Centrality Indicator
vertex
Initial Network Cohesion Score
0.14
Focus Network Cohesion Score
0.524
🚀 **Focus Analysis Found**
Network Cohesion Score: 0.5240
- **Input Source**: `/data/transactions-2025-02-01-08-30-00.parquet`
- **Scale Factor**: `1000`
- **Timestandard**: `minute`
- **Scaler Option**: `scale_and_round`
- **Global Leader Vertex**: `741`
- **Weights**: {
"amount": 1.0,
"delta_time": 0.0001
}
Logs:
- Cluster 0.0: Node 32.0 has the highest cluster-normalized centrality of 0.0085
- Cluster 1.0: Node 404.0 has the highest cluster-normalized centrality of 0.0027
- cluster_id: 0, vertex: 32, cluster_centrality: 0.0085, global_centrality: 0.0006
- cluster_id: 1, vertex: 404, cluster_centrality: 0.0027, global_centrality: 0.0005
- First high centrality node: 32.0

8. Metabase와의 통합

8.1 Metabase는 무엇인가?

Metabase는 원시 데이터를 직관적인 대시보드와 시각화를 통해 실행 가능한 인사이트로 변환하는 선도적인 오픈소스 비즈니스 인텔리전스(BI) 도구입니다. Fortune 500 기업을 포함하여 전 세계 수천 개의 조직에서 신뢰받고 있으며, 기존 BI 솔루션의 복잡성 없이 엔터프라이즈급 분석 기능을 제공합니다.

주요기능:

- **시각적 쿼리 빌더**: 드래그 앤 드롭 인터페이스를 사용하여 복잡한 쿼리 생성
- **SQL 편집기**: 고급 사용자를 위한 사용자 정의 SQL 쿼리 작성
- **대화형 대시보드**: 필터링 및 드릴다운 기능이 있는 실시간 데이터 시각화
- **다중 데이터 소스**: PostgreSQL, MySQL, MongoDB 등 20개 이상의 데이터베이스 유형에 연결
- **공유 및 협업**: 세분화된 권한을 통한 안전한 대시보드 공유
- **모바일 반응형**: 모든 기기에서 대시보드 접근 가능

8.2 분석을 위해 Metabase를 선택하는 이유?

8.2.1 완전한 데이터 제어

클라우드 기반 BI 솔루션과 달리 Metabase는 온프레미스에 완전히 배포할 수 있어 다음을 보장합니다:

- **데이터 주권:** 민감한 데이터가 인프라를 벗어나지 않음
- **클라우드 의존성 제로:** 외부 서비스나 제3자 제공업체에 대한 의존 없음
- **완전한 규정 준수:** GDPR, HIPAA, SOX 및 기타 규제 요구사항 충족

8.2.2 엔터프라이즈급 보안

- **역할 기반 접근 제어:** 세분화된 수준에서 사용자 권한 정의
- **데이터 샌드박싱**: 사용자 그룹별 특정 데이터 하위 집합에 대한 접근 제한
- **감사 로깅**: 모든 사용자 활동 및 데이터 접근 추적
- **SSO 통합**: SAML, LDAP, JWT, Google 인증 지원

8.2.3 비용 효율적인 솔루션

- **오픈소스 코어**: 사용자당 라이선스 비용 없이 무료 사용
- **자체 호스팅**: 지속적인 클라우드 구독 비용 제거
- **확장 가능**: 소규모 팀에서 전사적 배포까지 성장
- **벤더 종속 없음**: BI 인프라에 대한 완전한 제어

8.2.4 입증된 신뢰성

- **프로덕션 준비:** 전 세계 수천 개 조직에서 사용
- **활발한 개발:** 정기적인 업데이트 및 보안 패치
- **커뮤니티 지원:** 대규모 개발자 및 사용자 커뮤니티
- **엔터프라이즈 지원:** 미션 크리티컬 배포를 위한 유료 지원 옵션 제공

사용자는 온프레미스 머신이나 원격 워크스테이션에서 완전한 데이터 제어 및 개인정보 보호와 함께 Metabase를 쉽게 구현할 수 있으며,
민감한 비즈니스 정보가 보안 인프라를 벗어나지 않도록 보장합니다.

8.2.5 보안 데이터 파이프라인을 위한 권장 아키텍처

최대 보안 및 자동화를 위해 다음 데이터 파이프라인 아키텍처 구현을 권장합니다:

보안 볼륨 마운트 설정
보안 인트라넷 폴더 내에 전용 볼륨 마운트 드라이브를 설정하는 것이 권장됩니다. 이는 원활한 데이터 흐름을 가능하게 하면서 엄격한 보안 경계를 유지하는 격리된 데이터 교환 계층을 생성합니다.

자동화된 데이터 통합
MySQL 서버는 자동화된 모니터링 시스템을 통해 지정된 볼륨 마운트 드라이브에서 분석 데이터를 획득합니다. 제공되는 스크립트는 다음을 수행합니다:

- 필요한 데이터베이스 및 테이블 구조 자동 생성
- 새로운 분석 결과에 대한 볼륨 마운트 드라이브 지속적 모니터링
- 새로운 데이터 파일이 사용 가능해질 때 처리 및 가져오기
- 파이프라인 전반에 걸쳐 데이터 무결성 및 일관성 유지

원활한 Metabase 접근
데이터 파이프라인이 구축되면, 사용자는 Metabase에서 MySQL 서버에 직접 접근할 수 있어 보안 인프라 내에서 완전한 데이터 주권을 유지하면서 분석 결과에 대한 실시간 가시성을 제공받을 수 있습니다.

8.3 설치 사전 요구사항

- **운영 체제**: Linux (Ubuntu 18.04+, CentOS 7+, RHEL 7+), macOS, 또는 Windows
- **메모리**: 최소 2GB RAM (프로덕션 환경에서는 4GB+ 권장)
- **CPU**: 2개 이상 코어 권장
- **저장 장치**: 10GB 사용 가능한 디스크 공간
- **데이터베이스**: MySQL 8.0+ (메타데이터 저장용)

8.3.2 데이터베이스 구성 옵션

Metabase가 메타데이터 저장을 위해 사용하는 MySQL 데이터베이스에 대한 두 가지 배포 옵션이 있습니다:

Option 1: 온프레미스 분석 서버 (최대 보안을 위해 권장)

- 전용 온프레미스 분석 서버에 MySQL 8.0+ 설치
- Metabase가 내부 네트워크(사설 IP)를 통해 연결
- 인프라 내에서 완전한 데이터 격리
- 완전한 데이터 제어가 필요한 고도로 규제된 환경에 이상적

Option 2: 클라우드 기반 데이터베이스

- 선호하는 클라우드 제공업체에 MySQL 8.0+ 배포 (AWS RDS, Google Cloud SQL, Azure Database)
- Metabase가 보안 암호화 연결을 통해 연결
- 자동 백업 및 업데이트가 있는 관리형 데이터베이스 서비스
- 클라우드 인프라가 있는 하이브리드 환경에 적합

8.3.3 네트워크 요구사항

- **포트 3000**: 기본 Metabase 웹 인터페이스 (구성 가능)
- **데이터베이스 접근:**
  - **내부 네트워크**: 온프레미스 분석 서버에 대한 MySQL 포트 3306 접근
  - **클라우드 데이터베이스**: 클라우드 데이터베이스 서비스에 대한 HTTPS/SSL 암호화 연결
- **데이터 소스 연결**: 분석을 위한 비즈니스 데이터베이스에 대한 네트워크 접근
- **인터넷 접근**: 초기 설정 및 업데이트에 필요 (설치 후 제한 가능)

8.3.4 종속성

Docker 배포 (권장):

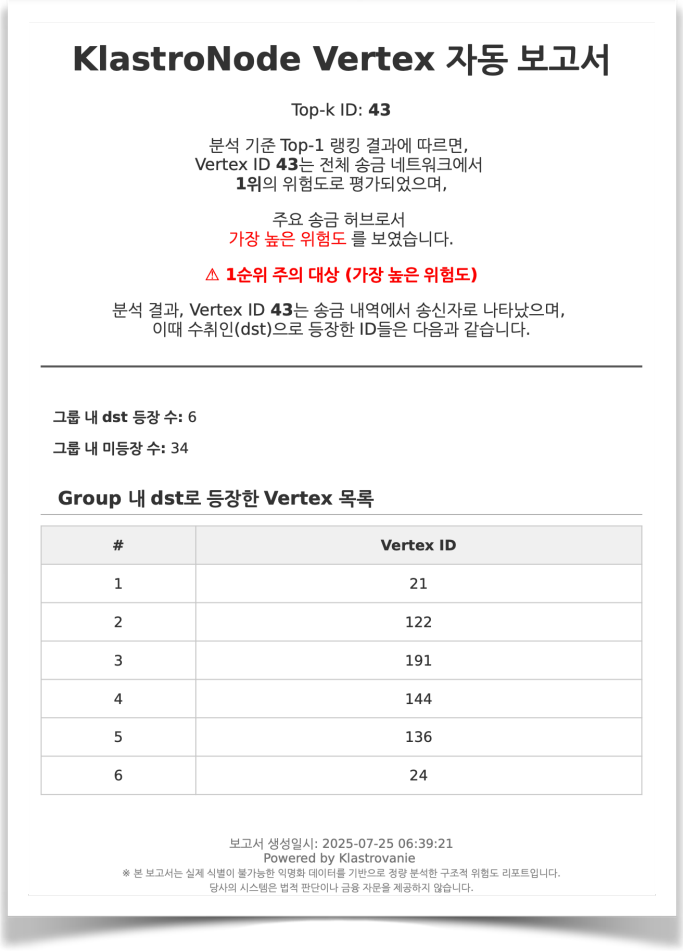
- **Docker:** 컨테이너화된 배포를 위한 버전 20.10+ 및 Docker Compose
- **Java 설치 불필요:** Java 런타임이 Docker 컨테이너 내에 포함됨

JAR 파일 배포:

- **Java:** OpenJDK 11 또는 Oracle JDK 11+ (직접 JAR 실행에만 필요)
- **Docker (선택사항):** 데이터베이스 컨테이너를 위해 JAR 배포와 함께 사용 가능

9. Reporter 컨테이너

klastroNode Vertex는 다층 송금 네트워크 분석을 통해 주요 송금 허브의 위험도 랭킹을 산출하고, 상위 노드에 대한 상세 분석 보고서를 자동 생성합니다. 포맷은 PDF 입니다. 해당 작업 폴더 내에 하위 report 폴더에 생성됩니다. 가장 높은 위험도를 보인 Top-k 송금자 리스트들에 각각 연결된 수취인을 기록하여 나타냅니다.



KlastroNode Vertex 자동 보고서

Top-k ID: **43**

분석 기준 Top-1 랭킹 결과에 따르면,
Vertex ID **43**는 전체 송금 네트워크에서
1위의 위험도로 평가되었으며,

주요 송금 허브로서
가장 높은 위험도 를 보였습니다.

⚠ **1순위 주의 대상 (가장 높은 위험도)**

분석 결과, Vertex ID **43**는 송금 내역에서 송신자로 나타났으며,
이때 수취인(dst)으로 등장한 ID들은 다음과 같습니다.

그룹 내 dst 등장 수: 6

그룹 내 미등장 수: 34

Group 내 dst로 등장한 Vertex 목록

#	Vertex ID
1	21
2	122
3	191
4	144
5	136
6	24

보고서 생성일시: 2025-07-25 06:39:21
Powered by Klastrovanie
※ 본 보고서는 실제 식별이 불가능한 익명화 데이터를 기반으로 정량 분석한 구조적 위험도 리포트입니다.
당사의 시스템은 법적 판단이나 금융 자문을 제공하지 않습니다.

위 샘플 보고서는 Vertex ID 43이 전체 네트워크에서 1위의 위험도를 나타낸 사례로, 해당 노드가 송신자로서 등장한 송금 내역을 기반으로 수취인(dst) 그룹 내 등장 여부를 시각화한 결과입니다

- 상위 위험 노드별 dst 분석 요약
- 그룹 내 등장/미등장 비교
- 위험도에 따른 시각 강조 및 설명

해당 기능은 .parquet 형식의 송금 내역 파일을 입력으로 사용하며, 보고서는 PDF 형태로 자동 저장됩니다. 또한, 보고서는 Top-k 노드 또는 Focus 대상 노드 기준으로 생성됩니다.

※본 지침은 네트워크 기반 분석 도구 운용에 대한 일반적인 권고 사항이며, 법적 판단이나 금융 자문을 제공하지 않습니다.
의심 정황이 식별되더라도 별도의 회계 감사, 법적 검토 및 기관 판단을 통해 해석되어야 합니다.

9. 기술 지원 및 문제 해결

🔐 **보안 안내**
- 업로드하는 데이터셋에는 개인 식별 정보(PII) 가 포함되지 않도록 주의해 주세요.

🔐 시스템의 설치 또는 사용 중 문제가 발생하거나 추가 기능 요청이 필요한 경우, 아래 연락처를 통해 지원을 받을 수 있습니다.
구매자가 구매자 시스템의 완전히 변경하여 적용한 경우에는 판매자인 Klastrovanie Co., Ltd. 는 지원해드릴 수 없습니다.

📧 **Contact**
- Email: contact@klastrovanie.com
- Website: www.klastrovanie.com