ENTENTE: Cross-silo Intrusion Detection on Network Log Graphs with Federated Learning

Author: Jiacen Xu, Chenang Li, Yu Zheng, Zhou Li

2026 Network and Distributed System Security Symposium (NDSS)

Speaker: Jiacen Xu

Background:

Entente Meaning

Dictionary

Definitions from Oxford Languages · Learn more

en·tente

/änˈtänt/

Learn to pronounce

noun

a friendly understanding or informal alliance between states or factions.

"the growing entente between former opponents"

Similar: understanding | agreement | arrangement | entente cordiale | covenant

- a group of states in an informal alliance.
  "the unsuccessful scheme to lure Greece into the war on the side of the entente"
- the understanding between Britain and France reached in 1904, forming the basis of Anglo-French cooperation in World War I.
  noun: **Entente Cordiale**; noun: **the Entente Cordiale**

Background:

Regulations and Real-world Cases

Background:

Regulations and Real-world Cases

ARTICLE

From GDPR to Global CBPR: The New Era of Data Transfer Compliance

The global game of data governance has changed

In 2025, cross-border data transfers have become one of the most complex and high-stakes challenges for legal and compliance teams. Regulatory fragmentation, evolving national security concerns, and the rise of AI-driven processing have transformed data transfers from a compliance afterthought into a strategic risk category.

Background:

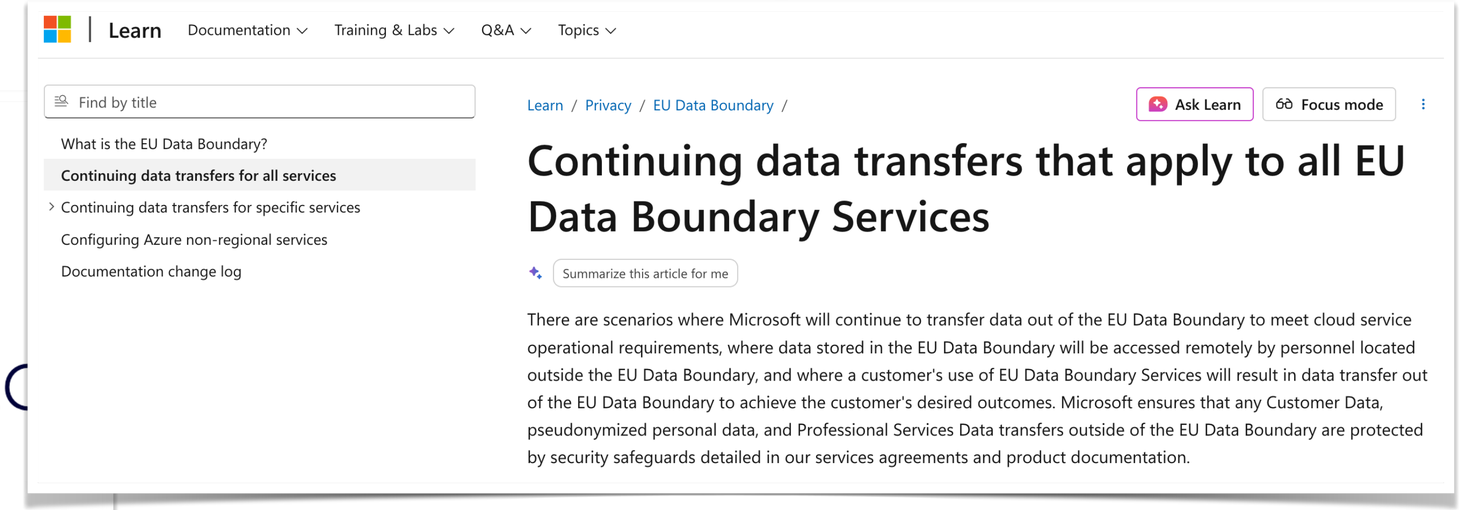
Regulations and Real-world Cases

ARTICLE

From GDPR to Global CBPR: The New Era of Data Transfer Compliance

The global game of data governance has changed

In 2025, cross-border data transfers have become one of the most complex and high-stakes challenges for legal and compliance teams. Regulatory fragmentation, evolving national security concerns, and the rise of AI-driven processing have transformed data transfers from a compliance afterthought into a strategic risk category.



Learn | Documentation | Training & Labs | Q&A | Topics

Find by title

- What is the EU Data Boundary?
- Continuing data transfers for all services
- Continuing data transfers for specific services
- Configuring Azure non-regional services
- Documentation change log

Learn / Privacy / EU Data Boundary /

Ask Learn | Focus mode

Continuing data transfers that apply to all EU Data Boundary Services

Summarize this article for me

There are scenarios where Microsoft will continue to transfer data out of the EU Data Boundary to meet cloud service operational requirements, where data stored in the EU Data Boundary will be accessed remotely by personnel located outside the EU Data Boundary, and where a customer's use of EU Data Boundary Services will result in data transfer out of the EU Data Boundary to achieve the customer's desired outcomes. Microsoft ensures that any Customer Data, pseudonymized personal data, and Professional Services Data transfers outside of the EU Data Boundary are protected by security safeguards detailed in our services agreements and product documentation.

Background:

Regulations and Real-world Cases

ARTICLE

From GDPR to Global CBPR: The New Era of Data Transfer Compliance

The global game of data governance has changed

In 2025, cross-border data transfers have become one of the most complex and high-stakes challenges for legal and compliance teams. Regulatory fragmentation, evolving national security concerns, and the rise of AI-driven processing have transformed **data transfers** from a compliance afterthought into a strategic risk category.

Learn | Documentation | Training & Labs | Q&A | Topics

Find by title

- What is the EU Data Boundary?
- **Continuing data transfers for all services**
- Continuing data transfers for specific services
- Configuring Azure non-regional services
- Documentation change log

Learn / Privacy / EU Data Boundary /

Ask Learn | Focus mode

Continuing data transfers that apply to all EU Data Boundary Services

Summarize this article for me

There are scenarios where Microsoft will continue to transfer data out of the EU Data Boundary to meet cloud service operational requirements, where data stored in the EU Data Boundary will be accessed remotely by personnel located outside the EU Data Boundary, and where a customer's use of EU Data Boundary Services will result in data transfer out of the EU Data Boundary to achieve the customer's desired outcomes. Microsoft ensures that any Customer Data, pseudonymized personal data, and Professional Services Data transfers outside of the EU Data Boundary are protected by security safeguards detailed in our services agreements and product documentation.

paloalto NETWORKS | Products | Solutions | Services | Partners | Company | More | Demos and Trials

Cyberpedia / Compliance / Data Compliance / How The Next-Generation Security Platform Contributes to GDPR Compliance

- Simplified Healthcare Compliance and Risk Management with Prisma Cloud
- What Is Data Privacy?
- How to Maintain AWS Compliance
- What Is NIST?
- What Is Data Privacy Compliance?
- **How The Next-Generation Security Platform Contributes to GDPR Compliance**

How The Next-Generation Security Platform Contributes to GDPR Compliance

14 min. read

Cybersecurity is an essential investment to protect personal data and comply with the GDPR.

The vast majority of GDPR requirements center around data management, namely data collecting and processing. There are obligations to provide notice when collecting personal data, prohibitions on unauthorized data processing, requirements to keep records of data processing, a duty to appoint a data protection officer in certain instances, and rules regarding transfer of personal data to third parties and third countries, amongst others.

Background:

Regulations and Real-world Cases

ARTICLE

From GDPR to Global CBPR: The New Era of Data Transfer Compliance

The global game of data governance has changed

In 2025, cross-border data transfers have become one of the most complex and high-stakes challenges for legal and compliance teams. Regulatory fragmentation, evolving national security concerns, and the rise of AI-driven processing have transformed data transfers from a compliance afterthought into a strategic risk category.

Learn | Documentation | Training & Labs | Q&A | Topics

Find by title

- What is the EU Data Boundary?
- Continuing data transfers for all services
- Continuing data transfers for specific services
- Configuring Azure non-regional services
- Documentation change log

Learn / Privacy / EU Data Boundary /

Ask Learn | Focus mode

Continuing data transfers that apply to all EU Data Boundary Services

Summarize this article for me

There are scenarios where Microsoft will continue to transfer data out of the EU Data Boundary to meet cloud service operational requirements, where data stored in the EU Data Boundary will be accessed remotely by personnel located outside the EU Data Boundary, and where a customer's use of EU Data Boundary Services will result in data transfer out

Data sharing across regions meets regulation and compliancy issues.

paloalto NETWORKS | Products | Solutions | Services | Partners | Company | More | Demos and Trials

Cyberpedia / Compliance / Data Compliance / How The Next-Generation Security Platform Contributes to GDPR Compliance

- Simplified Healthcare Compliance and Risk Management with Prisma Cloud
- What Is Data Privacy?
- How to Maintain AWS Compliance
- What Is NIST?
- What Is Data Privacy Compliance?
- How The Next-Generation Security Platform Contributes to GDPR Compliance

How The Next-Generation Security Platform Contributes to GDPR Compliance

14 min. read

Cybersecurity is an essential investment to protect personal data and comply with the GDPR.

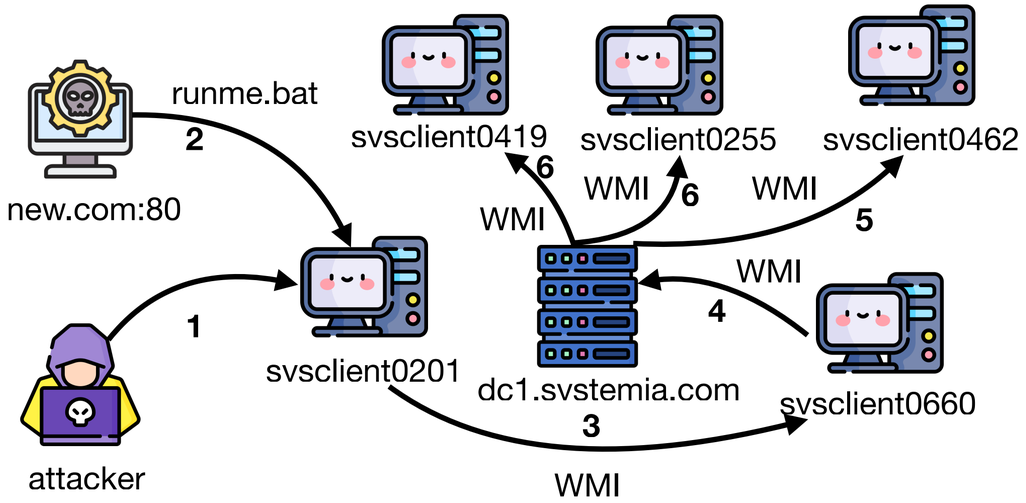
The vast majority of GDPR requirements center around data management, namely data collecting and processing. There are obligations to provide notice when collecting personal data, prohibitions on unauthorized data processing, requirements to keep records of data processing, a duty to appoint a data protection officer in certain instances, and rules regarding transfer of personal data to third parties and third countries, amongst others.

Background

Graph-based Network Intrusion Detection Systems (GNIDS)

Background

Graph-based Network Intrusion Detection Systems (GNIDS)



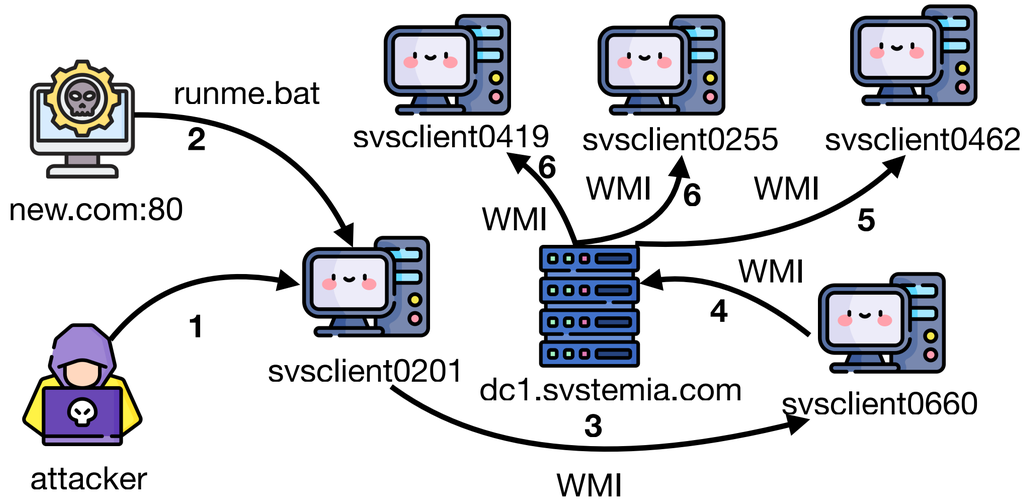
A real-world case in OpTC dataset.

Example:

Lateral Movement across clients and domain controllers.

Background

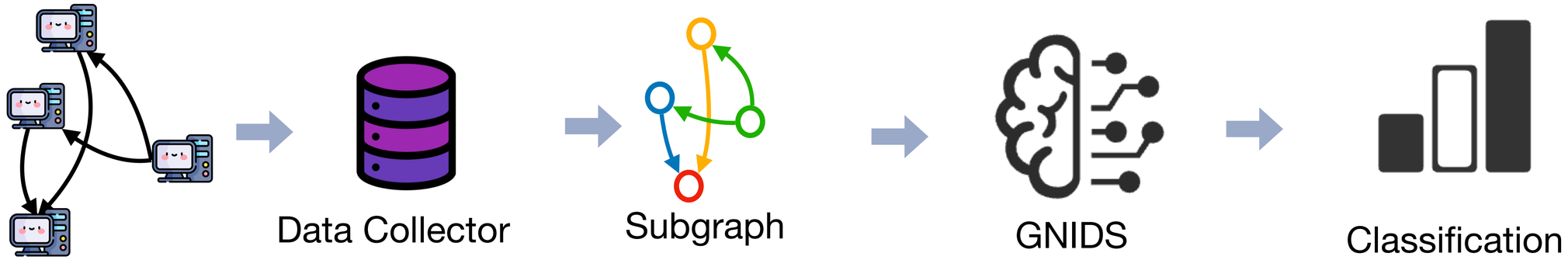
Graph-based Network Intrusion Detection Systems (GNIDS)



A real-world case in OpTC dataset.

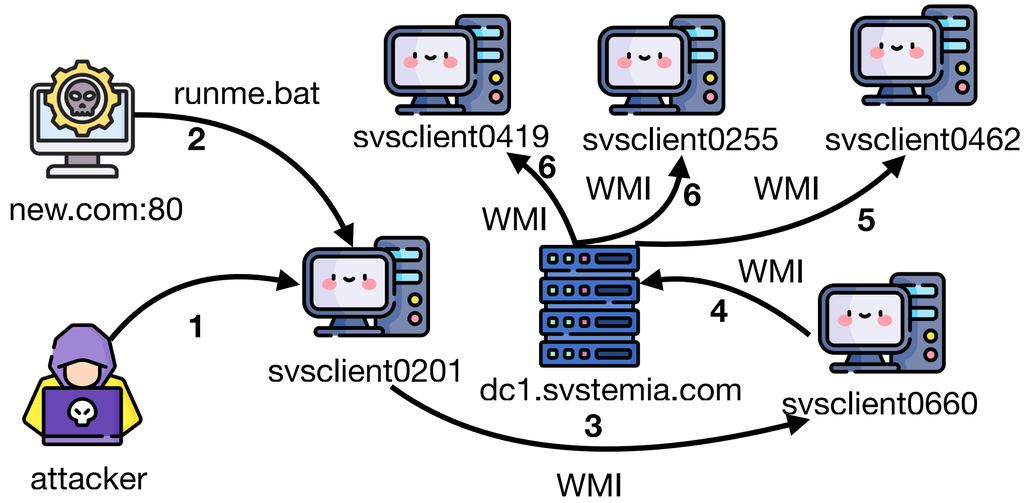
Example:

Lateral Movement across clients and domain controllers.



Background

Graph-based Network Intrusion Detection Systems (GNIDS)



A real-world case in OpTC dataset.

Example:

Lateral Movement across clients and domain controllers.

Problem Statement

- How to train a GNIDS model when:
  - Data is distributed across silos?
  - Data are **non-IID** and highly **heterogeneous**?
  - Attackers may poison the training process?

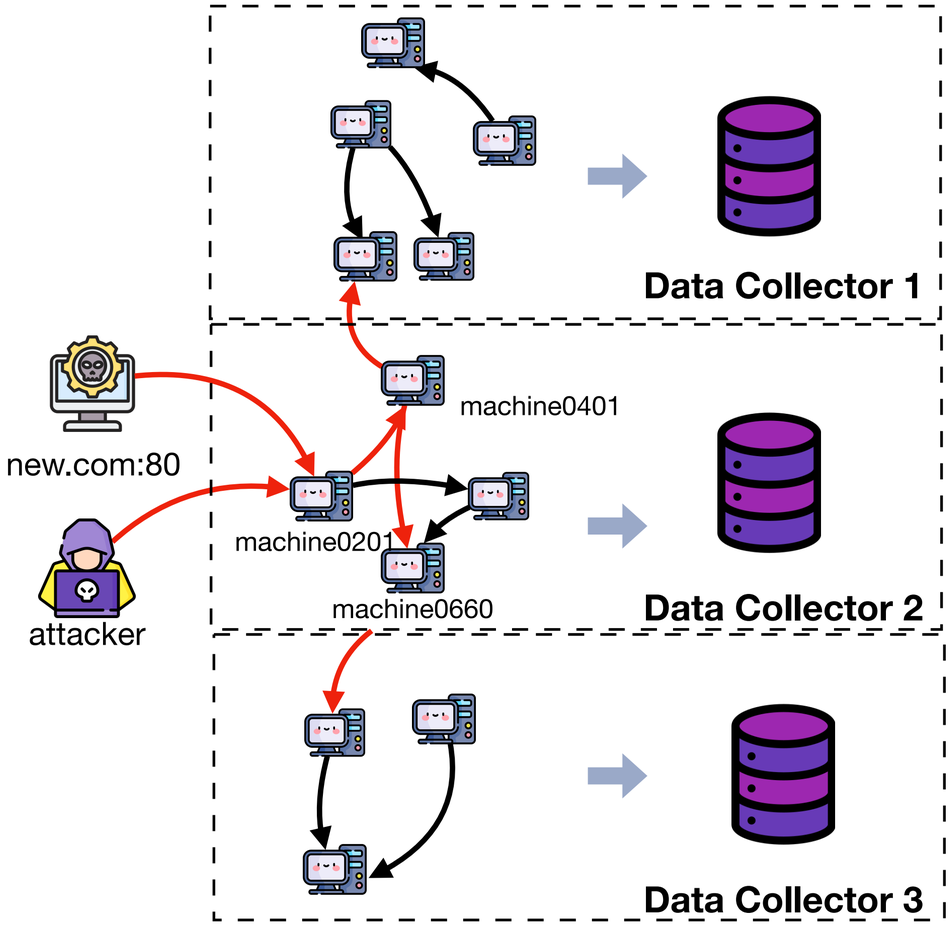
Problem Statement

- How to train a GNIDS model when:
  - Data is distributed across silos?
  - Data are **non-IID** and highly **heterogeneous**?
  - Attackers may poison the training process?

K / Setting	Node (OpTC)	Node (LANL)	Node (Pivoting)	Events (OpTC)	Events (LANL)	Events (Pivoting)
Non-FL	814	17649	1015	92M	1051M	74M
2 (SD)	475	10858	112	864K	33K	36M
3 (SD)	408	5626	139	759K	5M	34M
4 (SD)	359	5139	107	673K	6M	31M

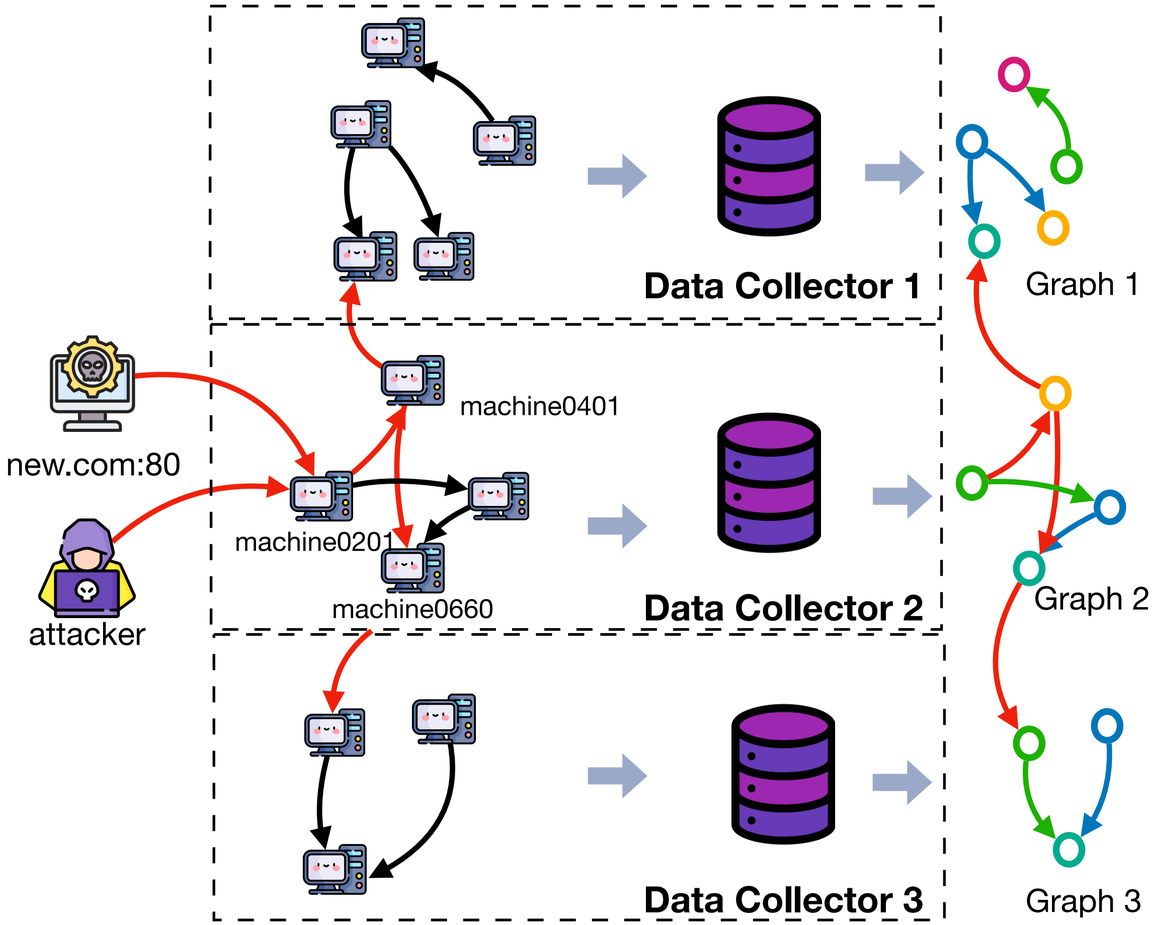
Problem Statement

How about Federated Learning?



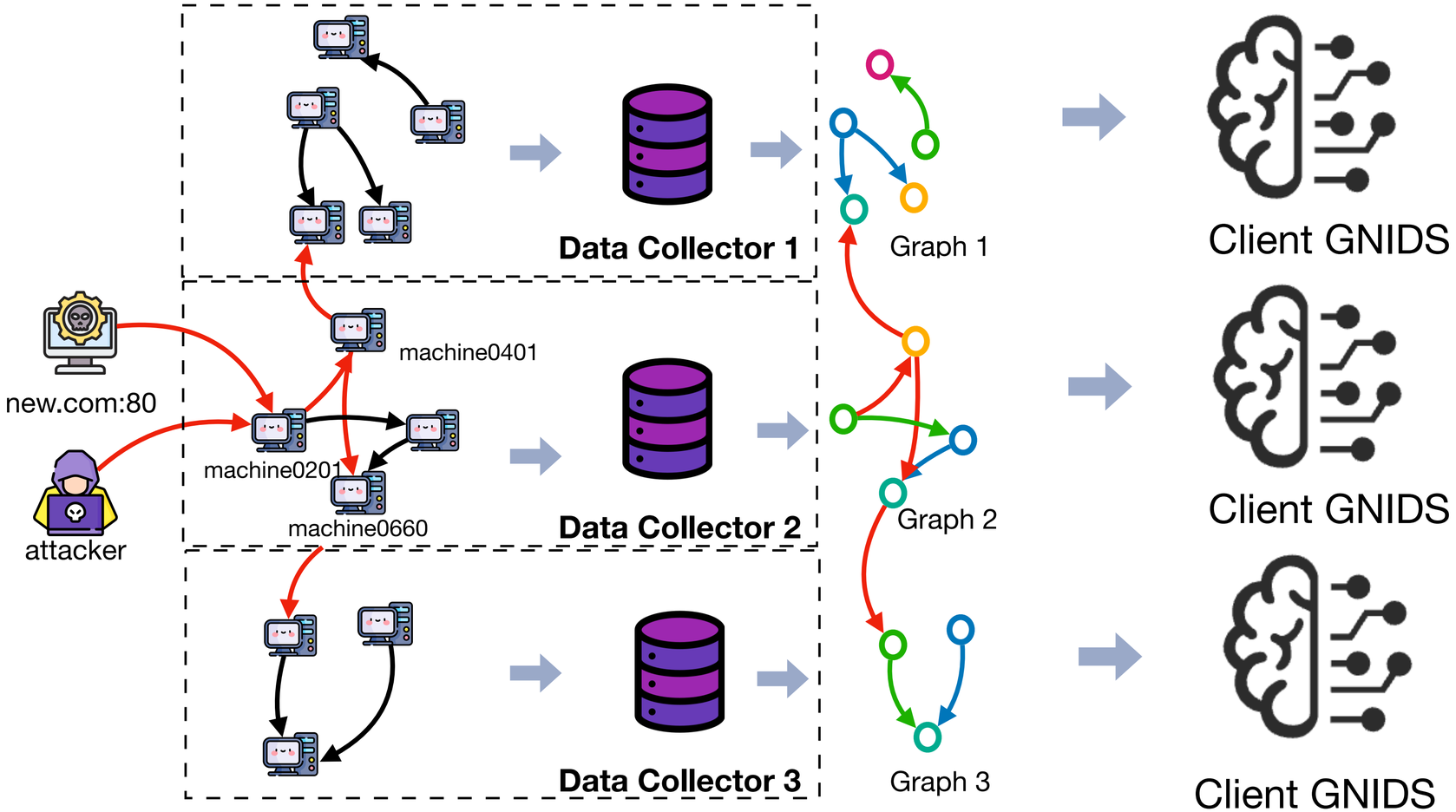
Problem Statement

How about Federated Learning?



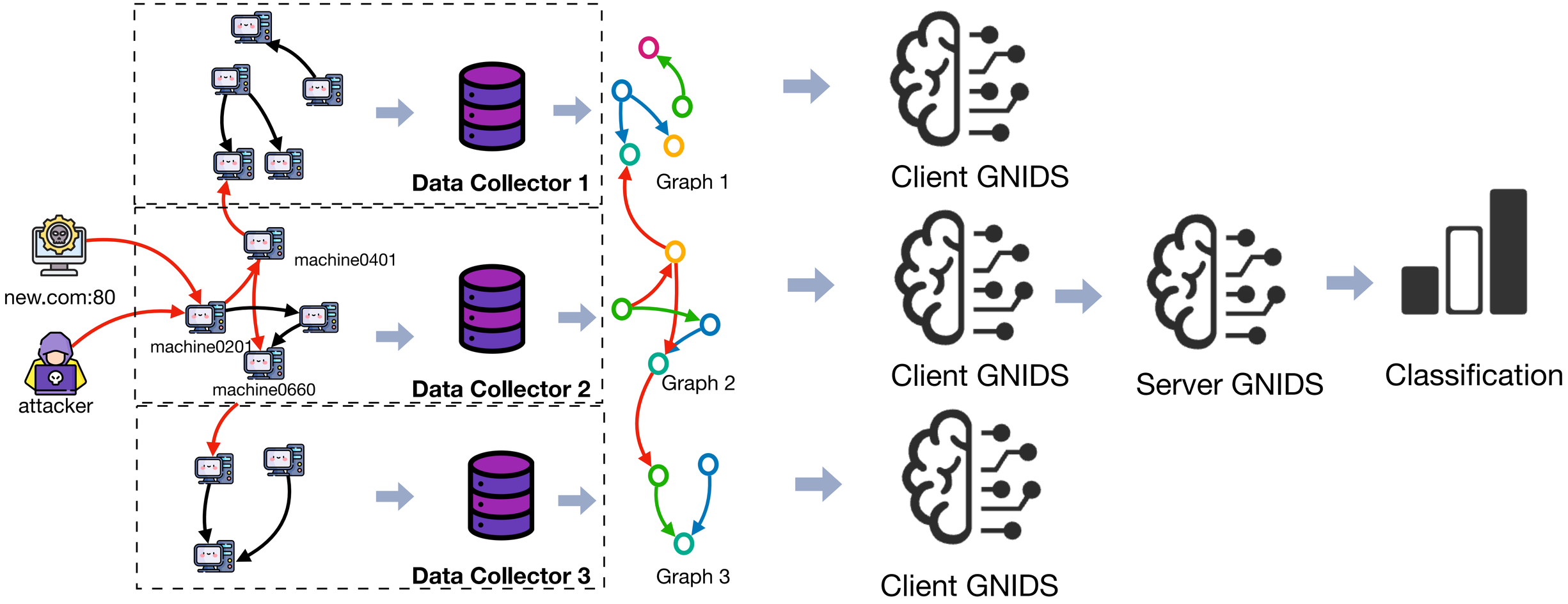
Problem Statement

How about Federated Learning?



Problem Statement

How about Federated Learning?



Problem Statement

- Existing FL methods fail on:
  - Graph heterogeneity
  - Imbalanced classes
  - Robustness vs. accuracy trade-offs

Key Idea: ENTENTE

- A federated learning framework tailored for GNIDS that achieves:
  - ✅ Effectiveness (close to centralized GNIDS)
  - ✅ Scalability (low overhead)
  - ✅ Robustness (resistant to poisoning)

System Overview

System Overview

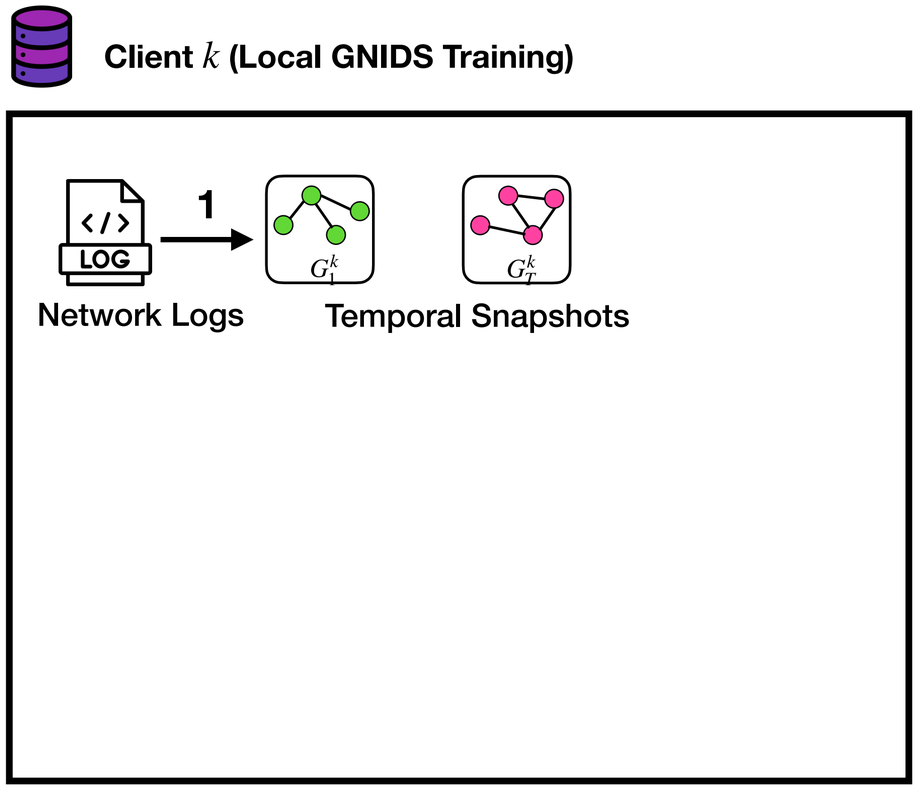
Client k (Local GNIDS Training)

System Overview

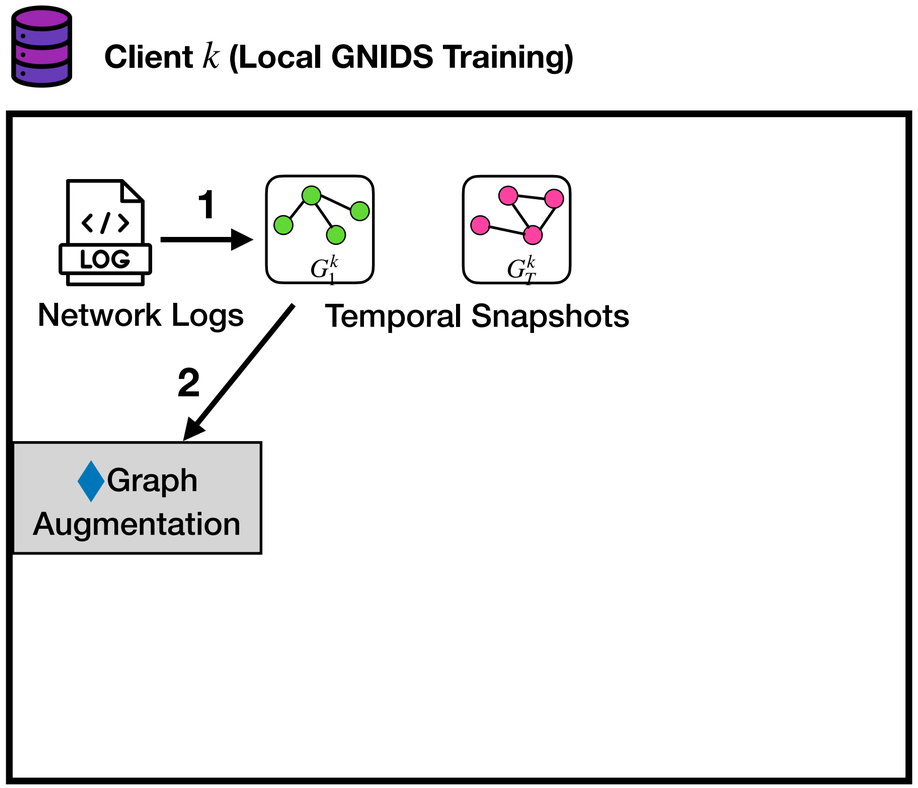
Client k (Local GNIDS Training)

Network Logs

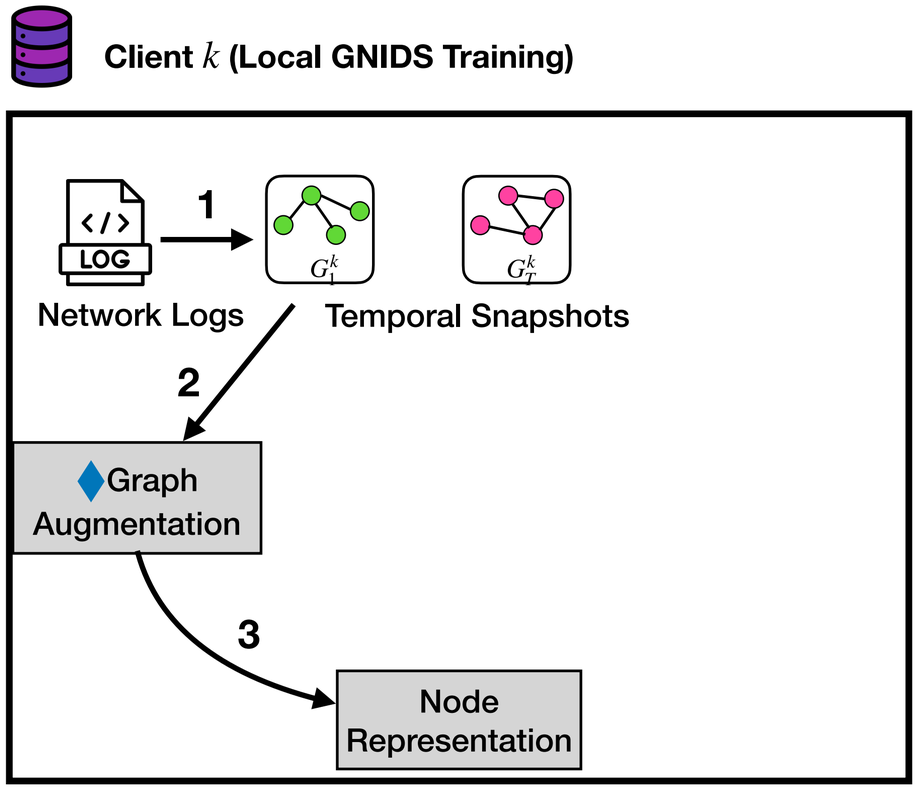
System Overview



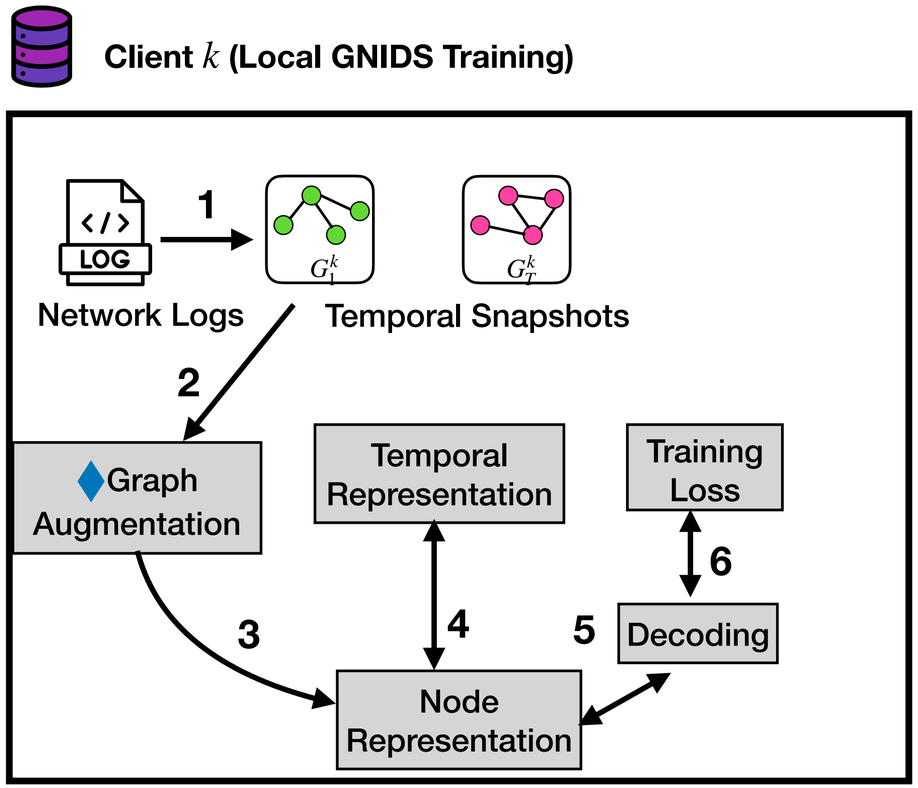
System Overview



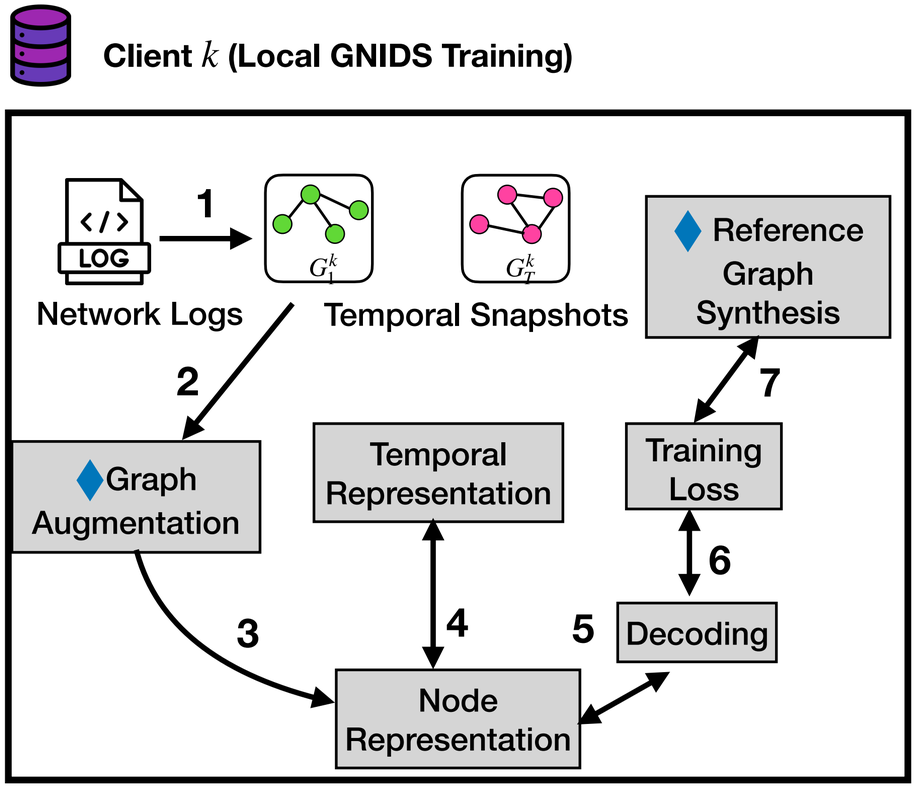
System Overview



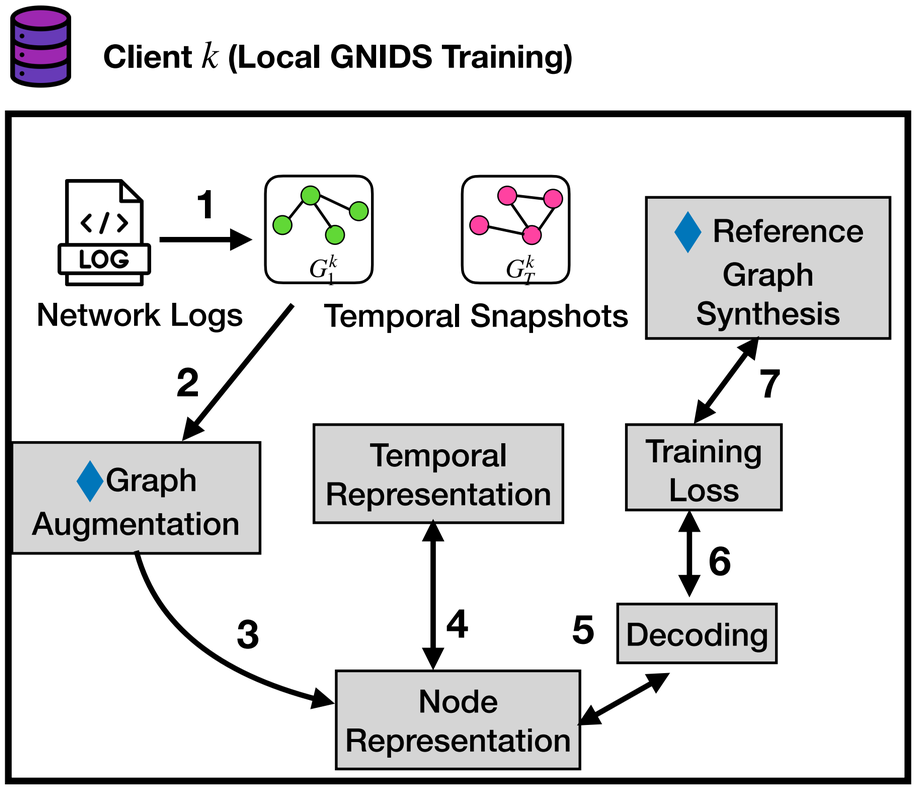
System Overview

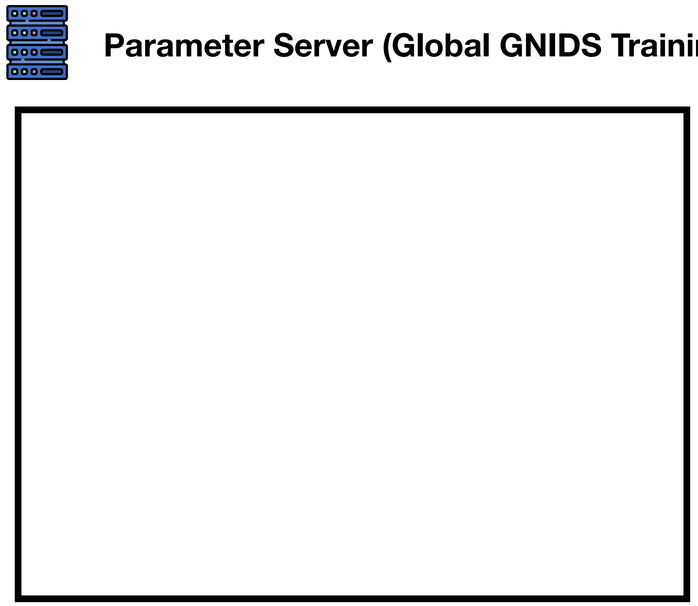


System Overview

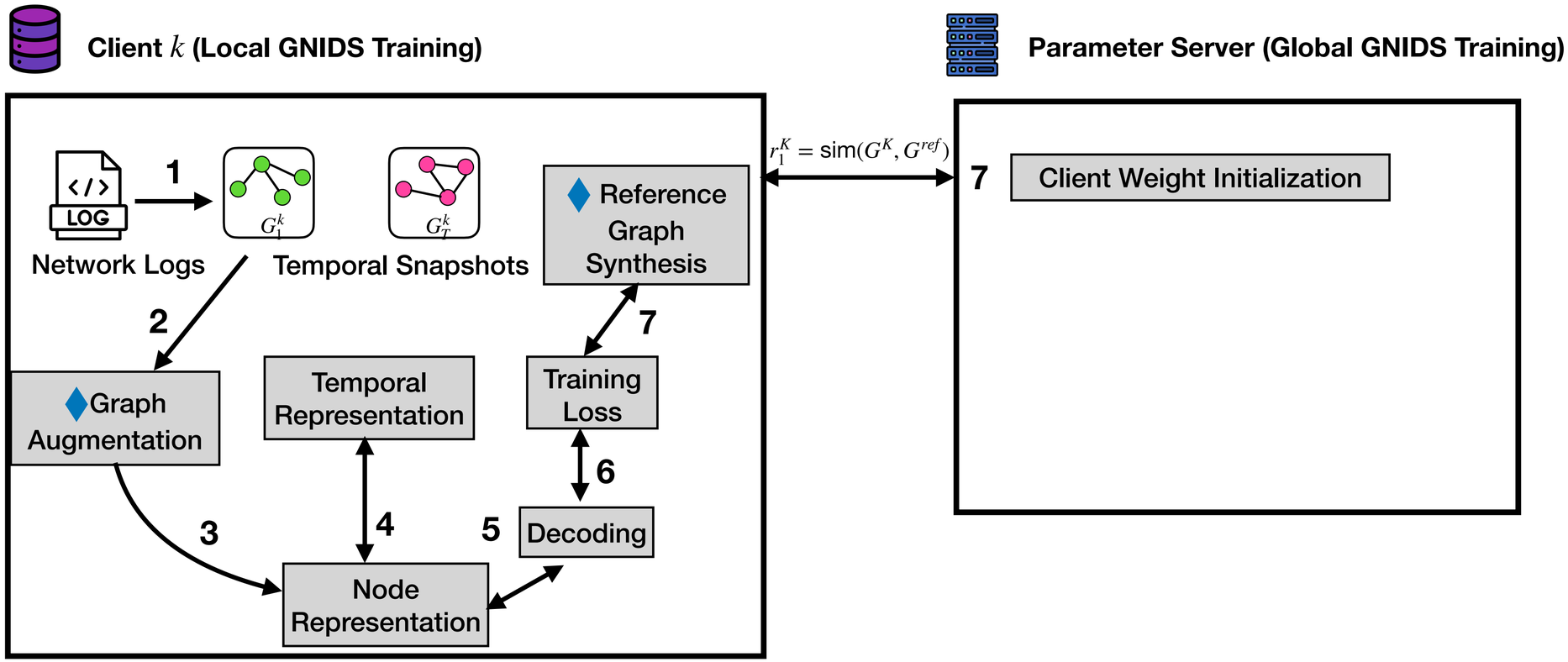


System Overview

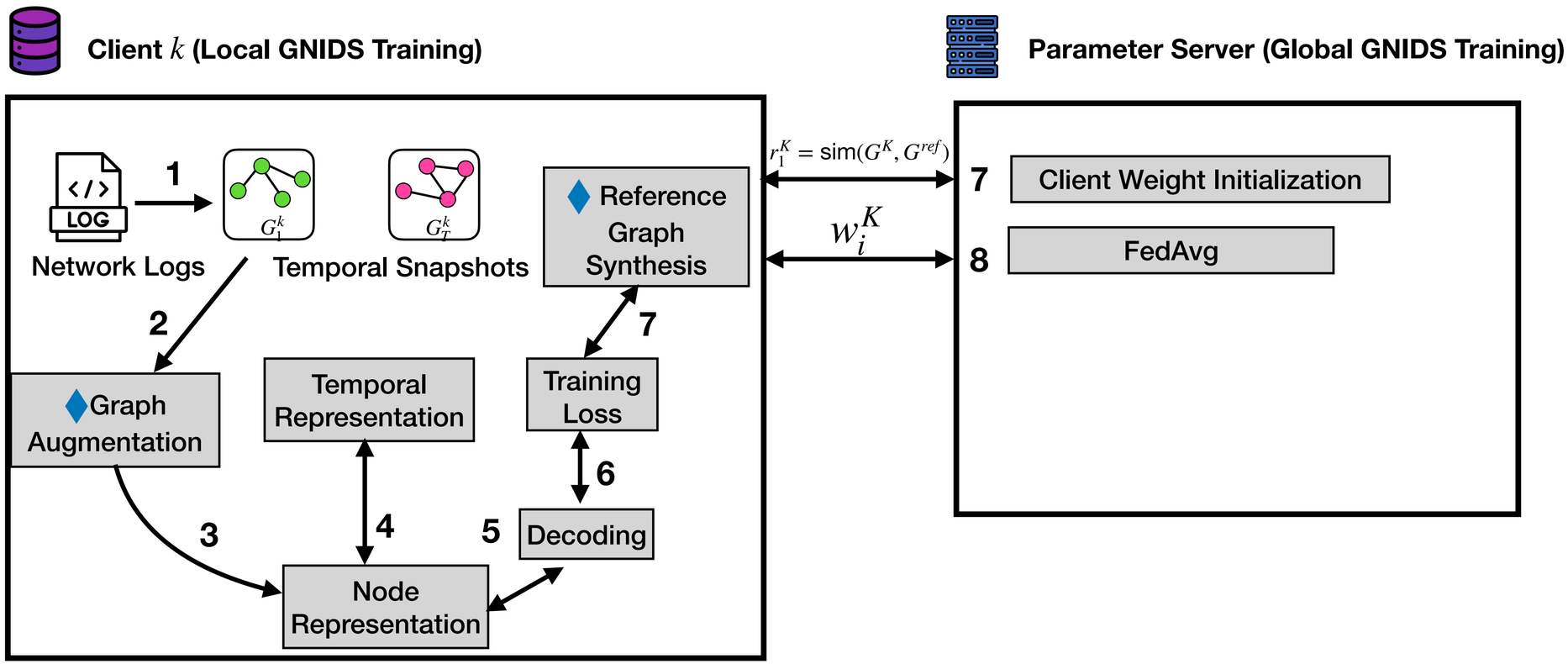




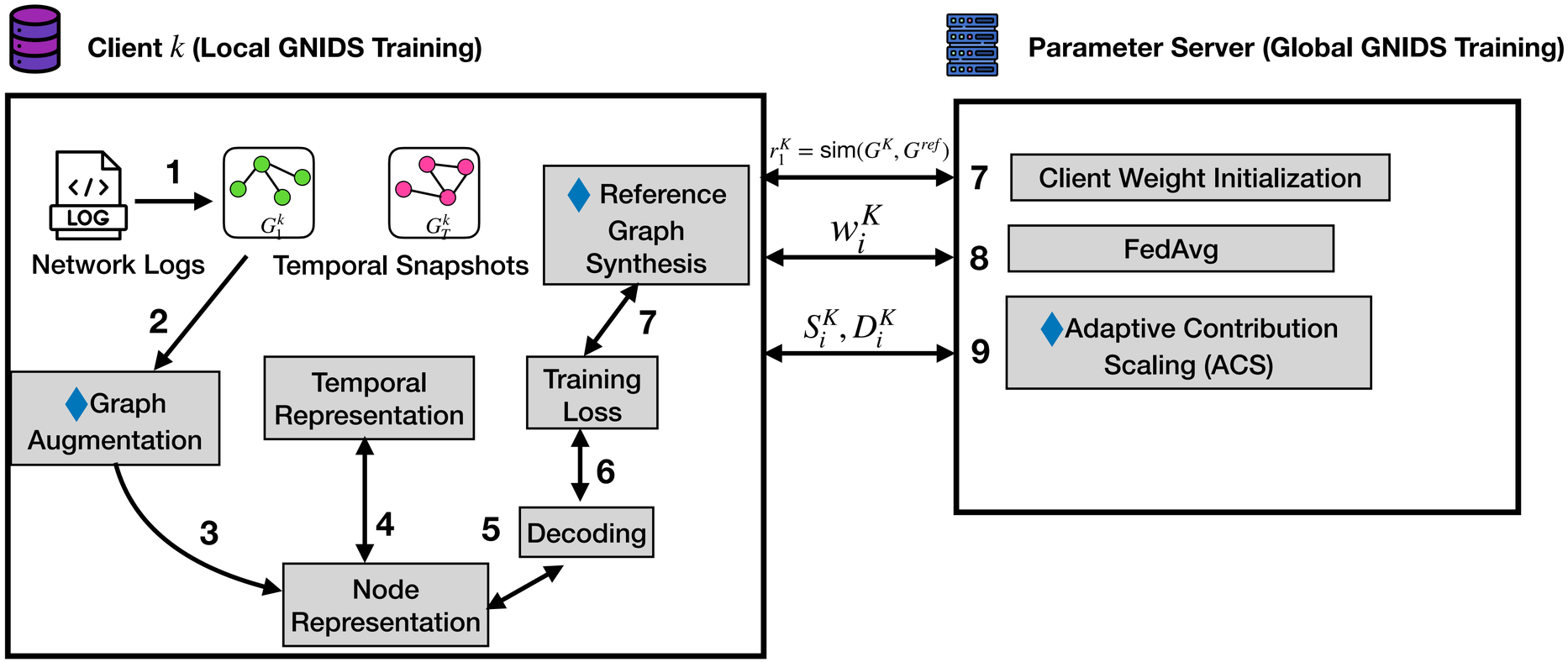
System Overview



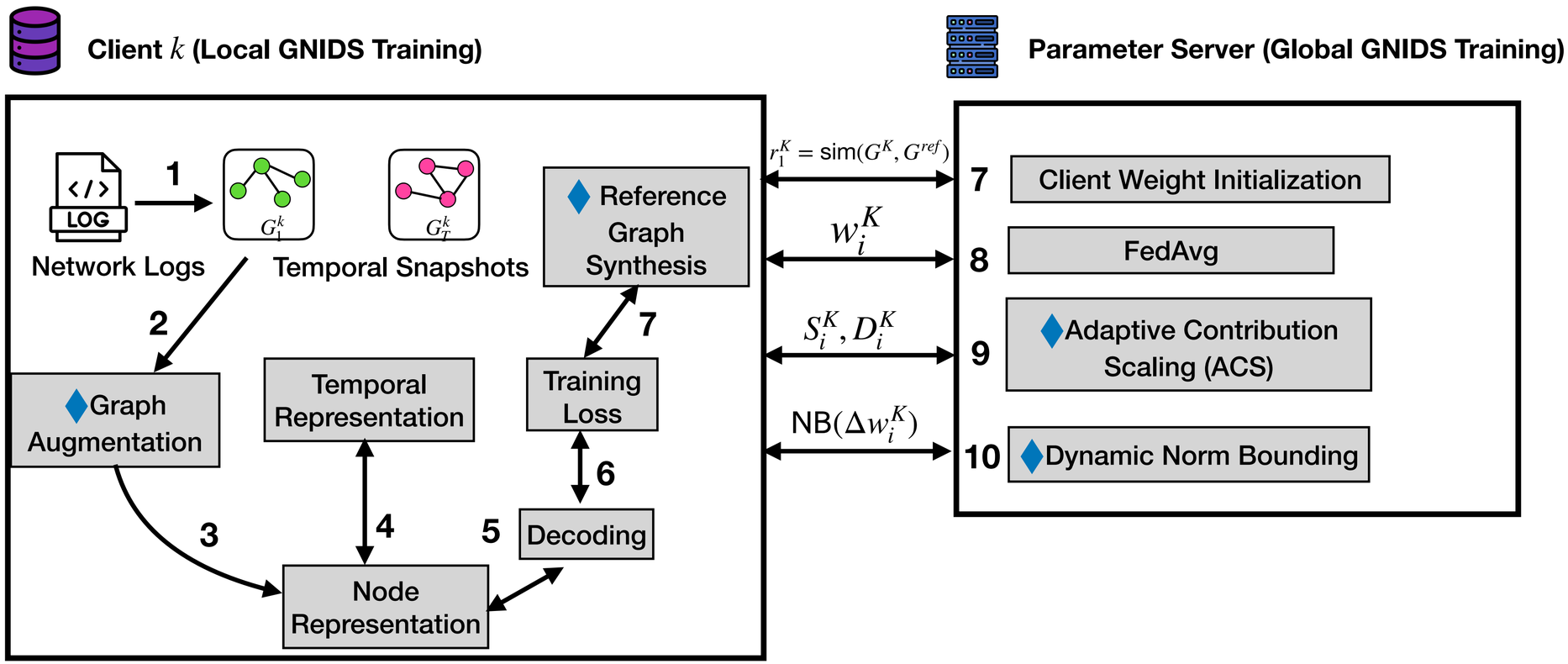
System Overview



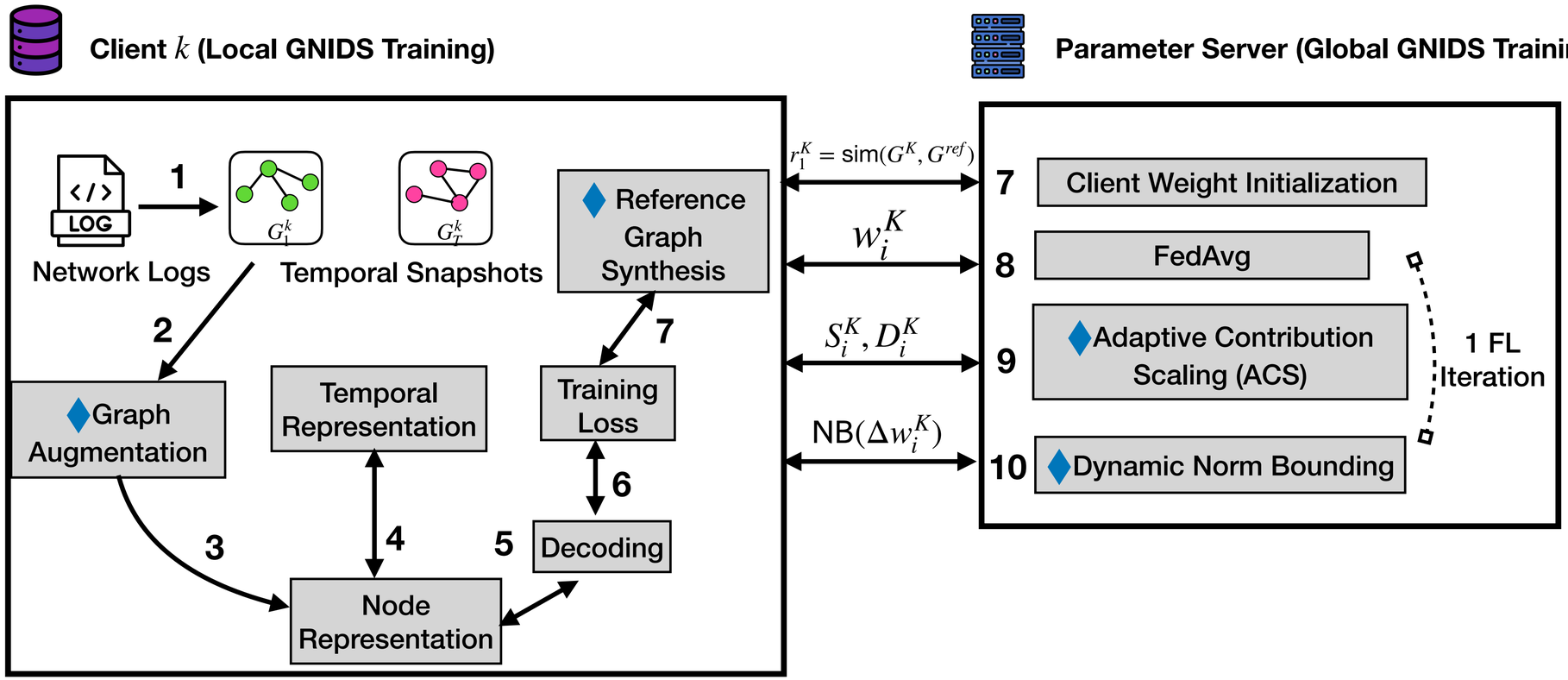
System Overview



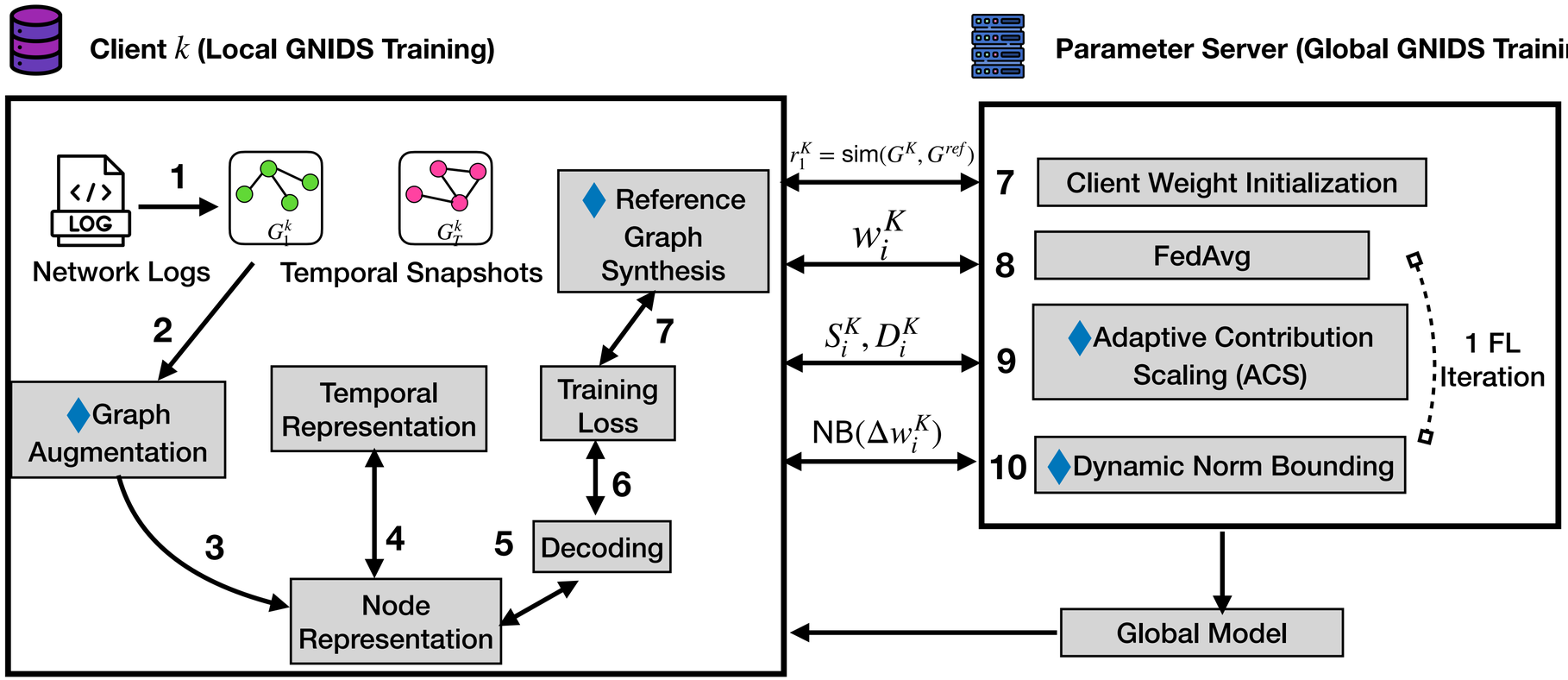
System Overview



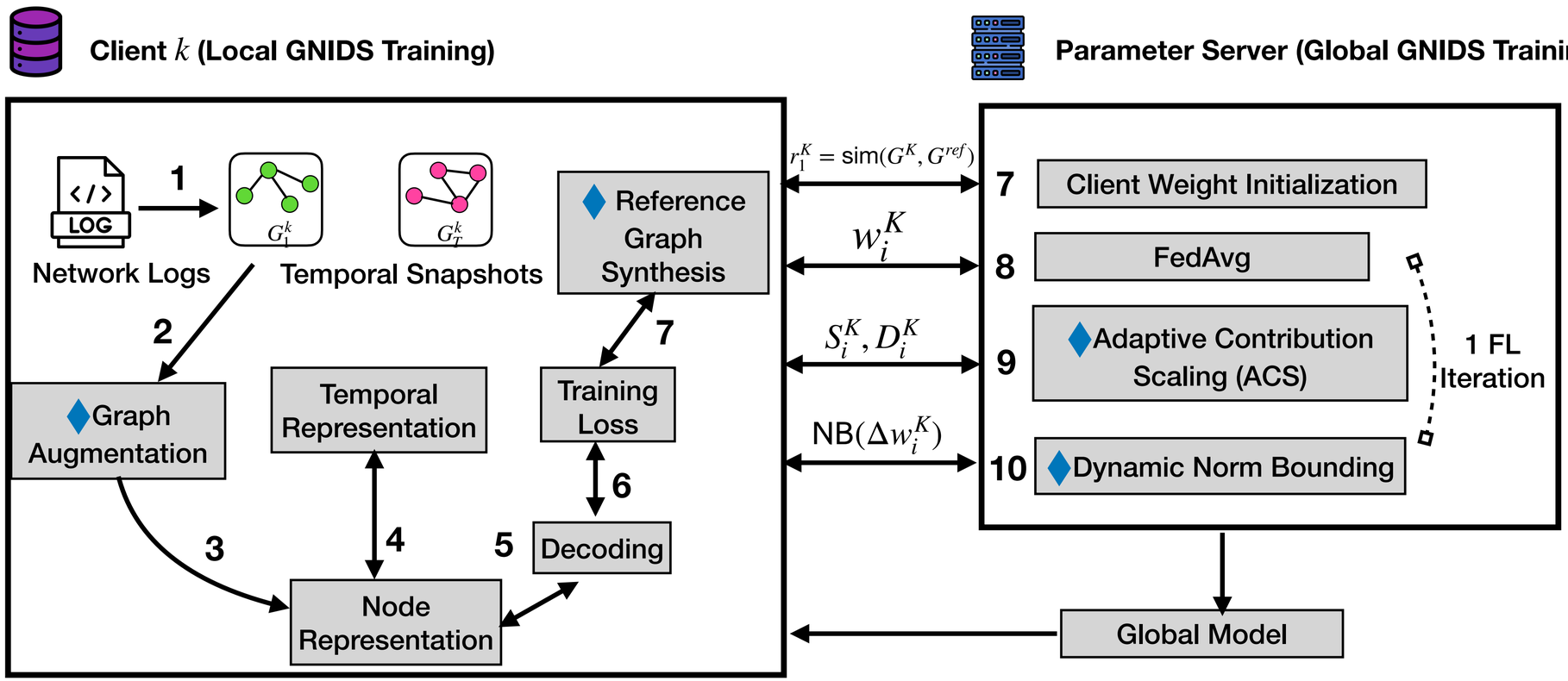
System Overview



System Overview

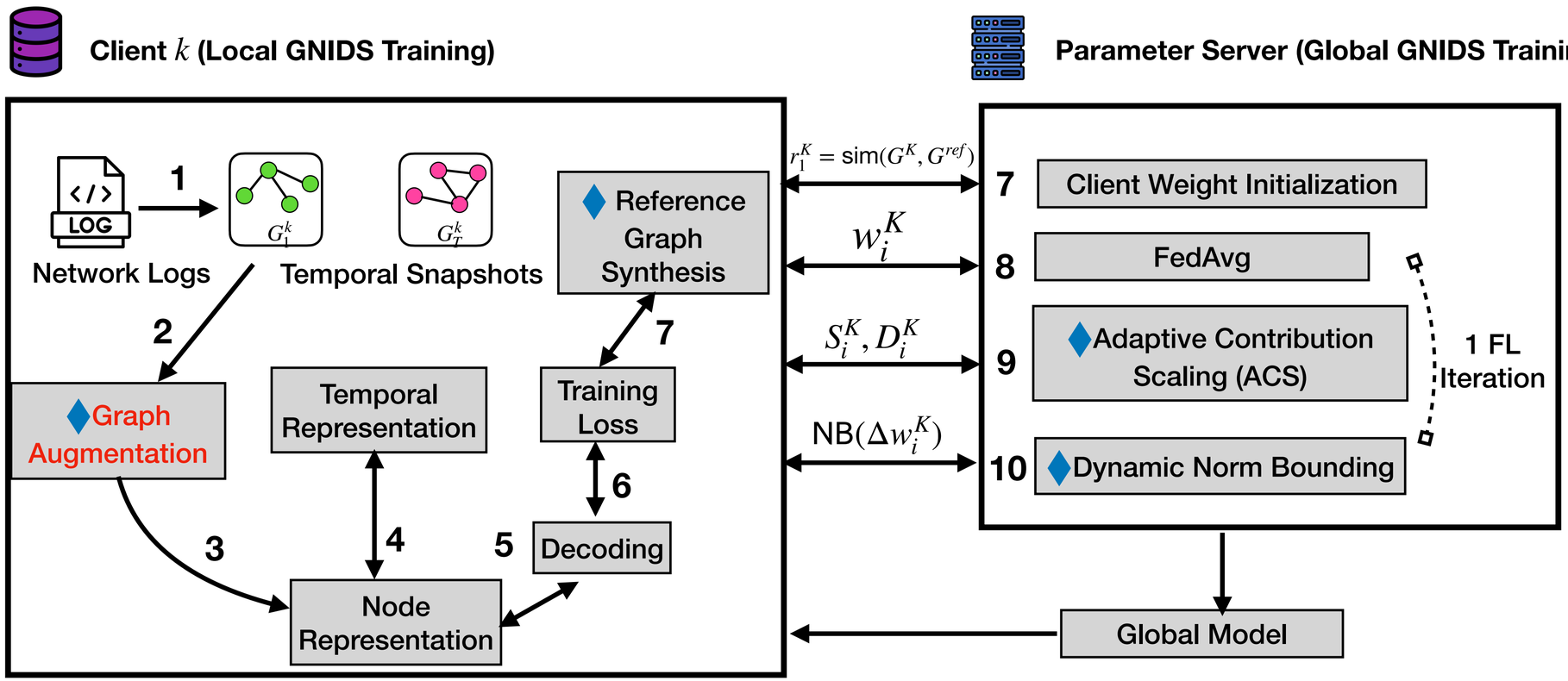


System Overview



- Key Components:
  - Graph Augmentation
  - FL with:
    - Reference Graph Synthesis
    - Adaptive Contribution scaling (ACS)
    - Dynamic Norm Bounding

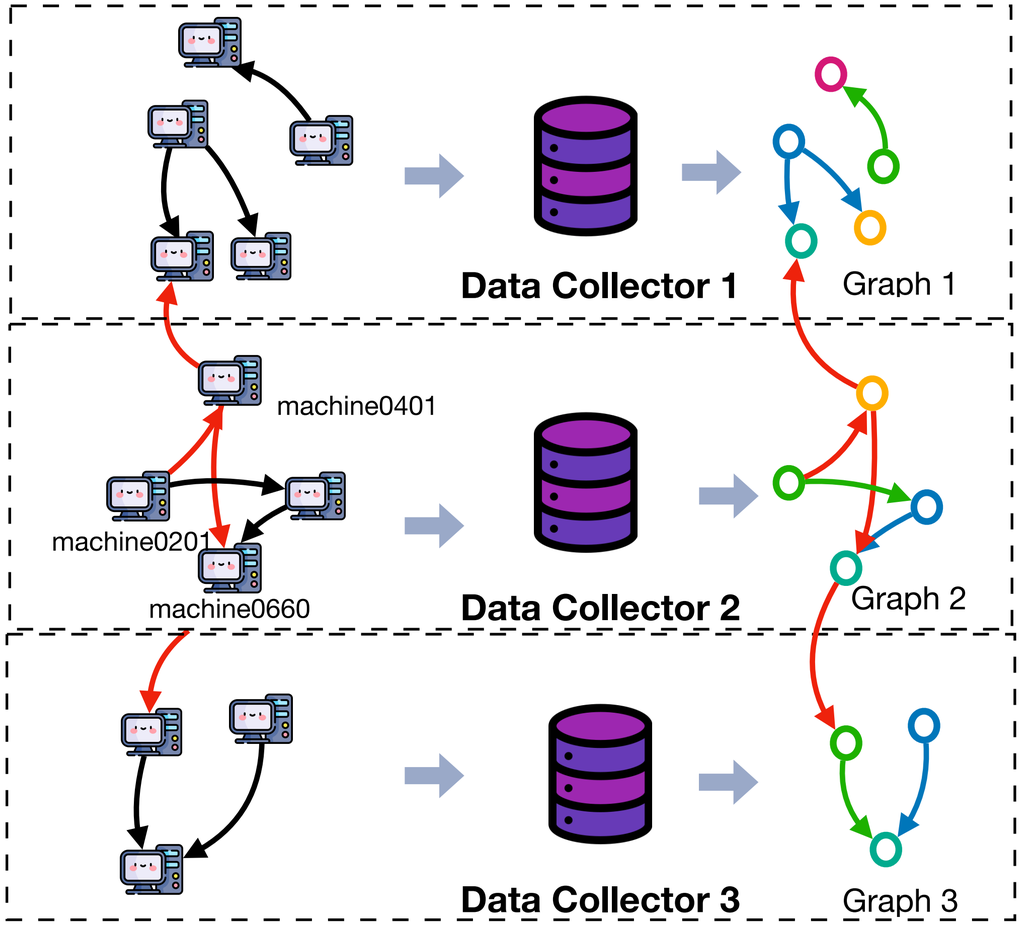
System Overview



- Key Components:
  - Graph Augmentation
  - FL with:
    - Reference Graph Synthesis
    - Adaptive Contribution scaling (ACS)
    - Dynamic Norm Bounding

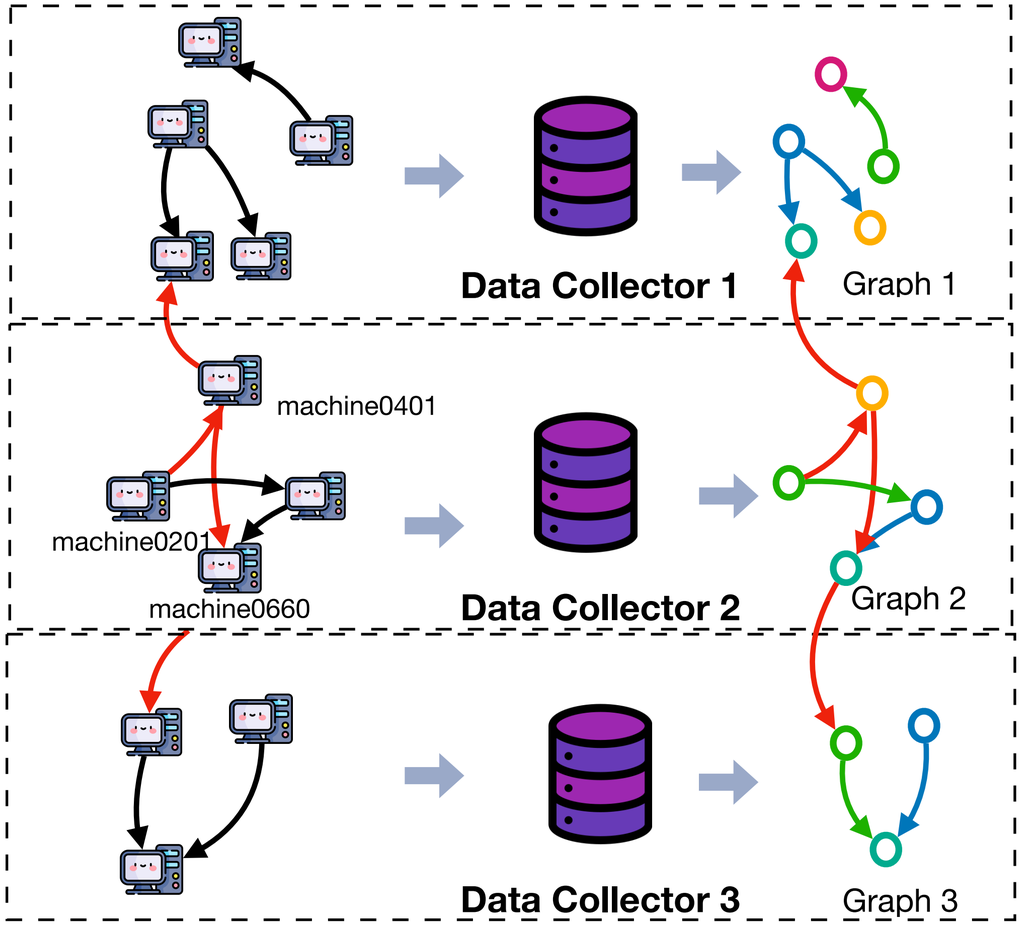
Technique 1: Graph Augmentation

- Following Centralized GNIDS, each client builds dynamic graph snapshots
  - Nodes: hosts
  - Edges: network flows / authentication events



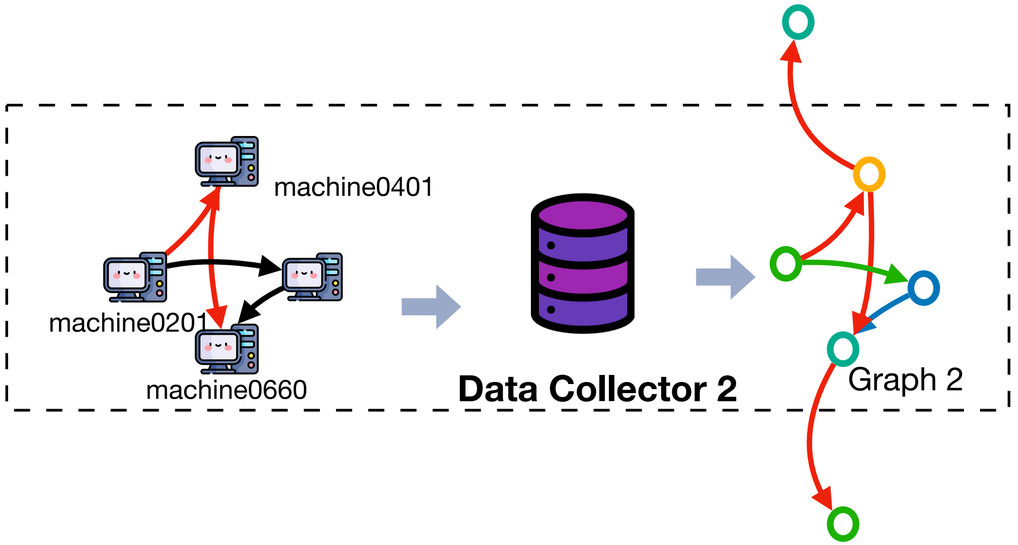
Technique 1: Graph Augmentation

- Following Centralized GNIDS, each client builds dynamic graph snapshots
  - Nodes: hosts
  - Edges: network flows / authentication events
- Insight: 1-hop cross-client augmentation
  - Firewalls already log inbound/outbound flows
  - No extra privacy leakage

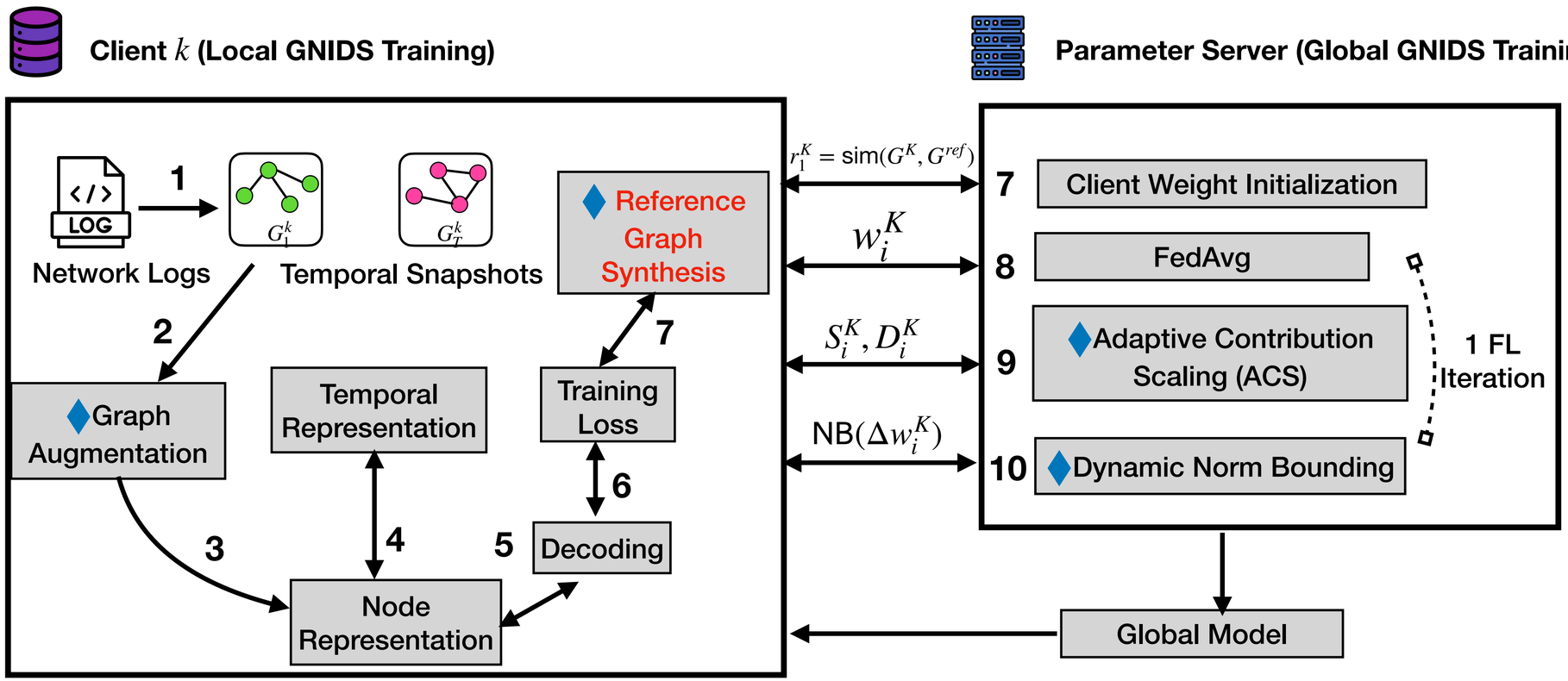


Technique 1: Graph Augmentation

- Following Centralized GNIDS, each client builds dynamic graph snapshots
  - Nodes: hosts
  - Edges: network flows / authentication events
- Insight: 1-hop cross-client augmentation
  - Firewalls already log inbound/outbound flows
  - No extra privacy leakage



System Overview



- Key Components:
  - Graph Augmentation
  - FL with:
    - Reference Graph Synthesis
    - Adaptive Contribution scaling (ACS)
    - Dynamic Norm Bounding

Technique 2: Reference Graph Synthesis

- Issue: FL uses average weights for each client which performs bad non-iid and imbalanced data.
- What is the better way to calculate client weights lightly and accurately?

Technique 2: Reference Graph Synthesis

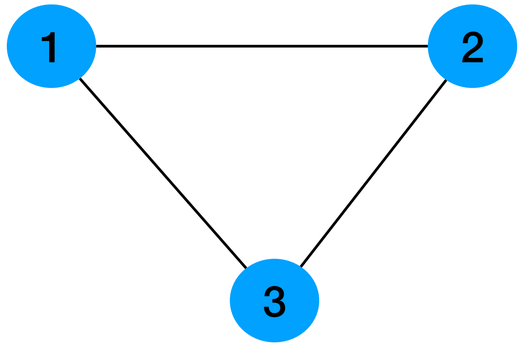
- Server knows only total node count of each client.
- Generates a **Barabási–Albert** (BA) reference graph
  - Scale-free:
    - most nodes have few connections;
    - a small number of hubs have high degree.

Technique 2: Reference Graph Synthesis

- Server knows only total node count of each client.
- Generates a **Barabási–Albert** (BA) reference graph
  - Scale-free:
    - most nodes have few connections;
    - a small number of hubs have high degree.

For example, Initial nodes: m_0=3, Edges per new node: m=2

$$P(i) = \frac{k_i}{\sum_i k_j} \quad \text{P(1)=P(2)=P(3)=}\frac{1}{3}$$

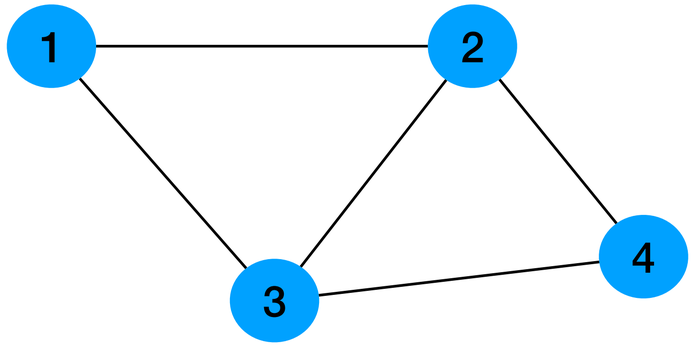


Technique 2: Reference Graph Synthesis

- Server knows only total node count of each client.
- Generates a **Barabási–Albert** (BA) reference graph
  - Scale-free:
    - most nodes have few connections;
    - a small number of hubs have high degree.

For example, Initial nodes: m_0=3, Edges per new node: m=2

$$P(i) = \frac{k_i}{\sum_i k_j}$$ P(1)=P(4)=$\frac{2}{10}$,P(2)=P(3)=$\frac{3}{10}$

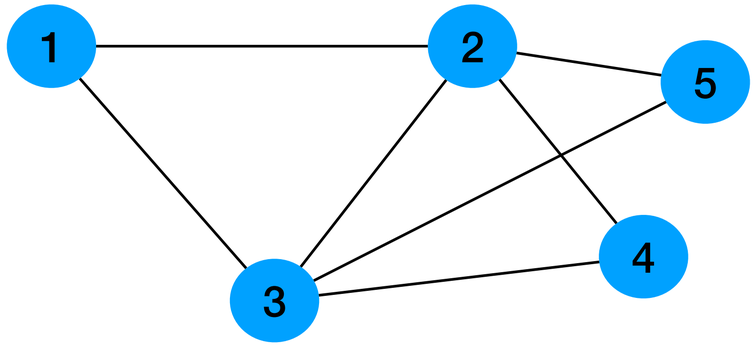


Technique 2: Reference Graph Synthesis

- Server knows only total node count of each client.
- Generates a **Barabási–Albert** (BA) reference graph
  - Scale-free:
    - most nodes have few connections;
    - a small number of hubs have high degree.

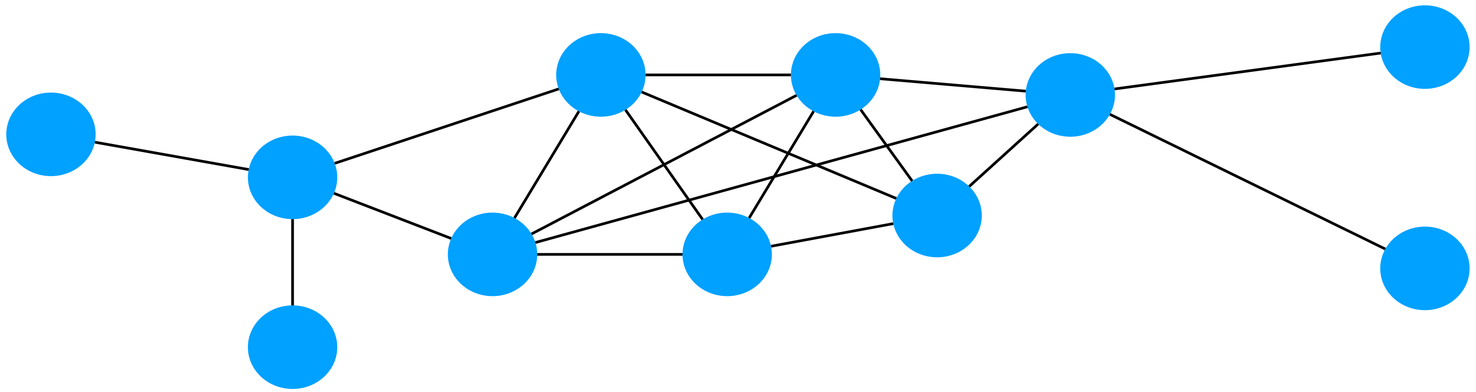
For example, Initial nodes: m_0=3, Edges per new node: m=2

$$P(i) = \frac{k_i}{\sum_i k_j} \qquad \text{P(1)=P(4)=}\frac{2}{10}\text{,P(2)=P(3)=}\frac{3}{10}$$



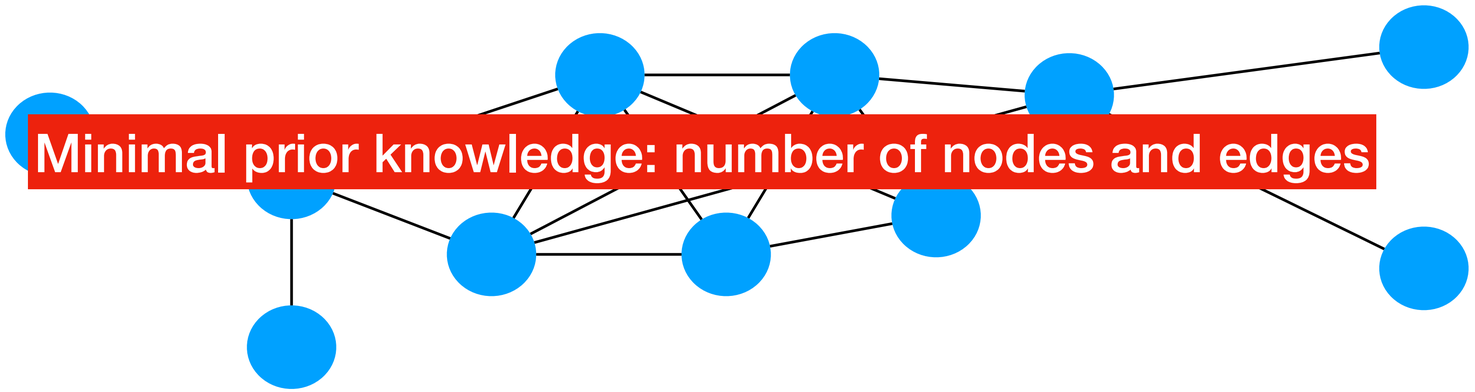
Technique 2: Reference Graph Synthesis

- Server knows only total node count of each client
- Generates a **Barabási–Albert** (BA) reference graph
  - Scale-free:
    - most nodes have few connections;
    - a small number of hubs have high degree.



Technique 2: Reference Graph Synthesis

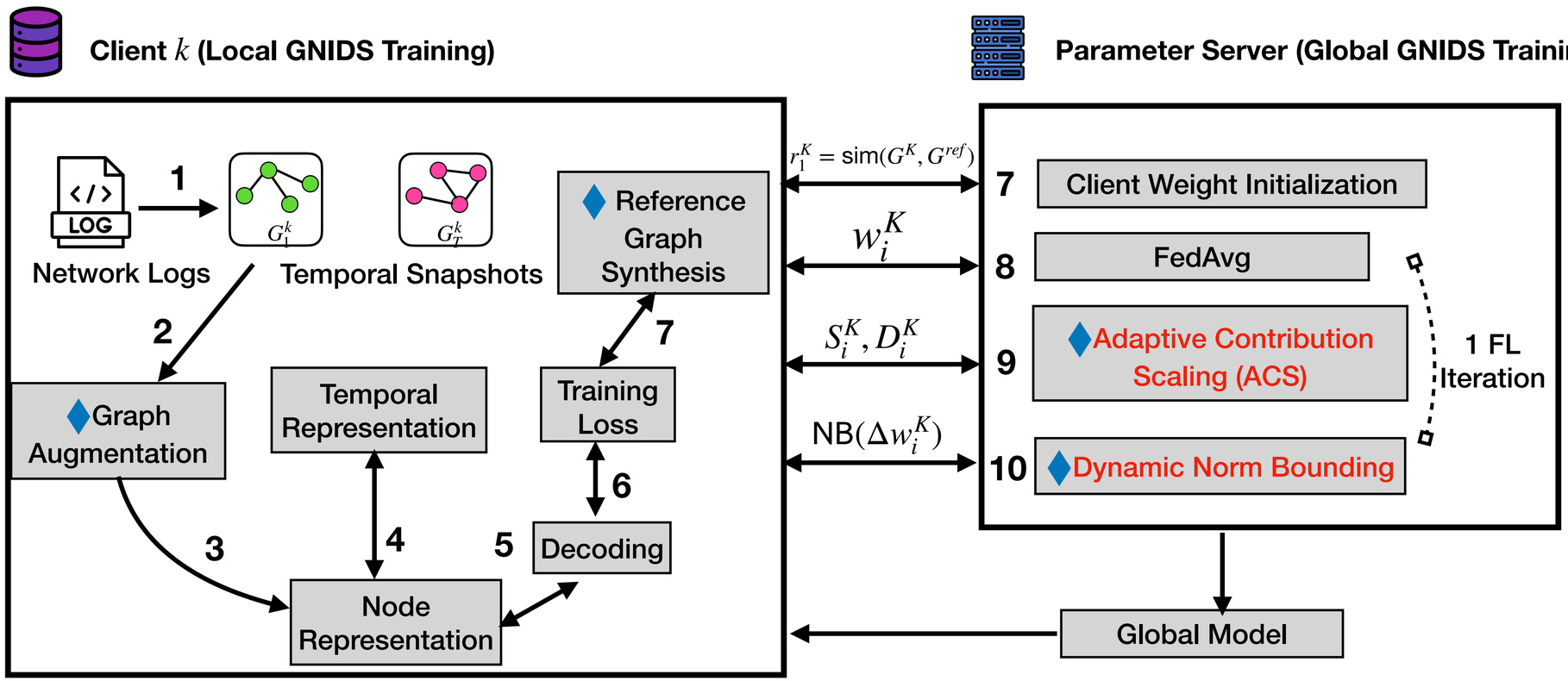
- Server knows only total node count of each client
- Generates a **Barabási–Albert** (BA) reference graph
  - Scale-free:
    - most nodes have few connections;
    - a small number of hubs have high degree.



Technique 2: Reference Graph Synthesis

- Clients compute similarity to reference graph
- Computing graph similarity is expensive
  - Use **Weisfeiler–Lehman** Histogram (WLH) in FL setting
  - a histogram/multiset local subtree patterns up to i-hop neighborhoods.
  - Compute Jaccard similarity between WLH(G_k) and WLH(G_ref)
  - Lightweight and privacy-preserving

System Overview



- Key Components:
  - Graph Augmentation
  - FL with:
    - Reference Graph Synthesis
    - Adaptive Contribution scaling (ACS)
    - Dynamic Norm Bounding

Technique 3: ACS + Norm bounding

- Each iteration adjusts client weight using:
  - S_{Jac}^{k}: structural similarity of weights
  - S_{i}^{k}: cosine similarity of weights to global model
  - D_{i}^{k}: norm bounded L2 distance of weights

Technique 3: ACS + Norm bounding

- Each iteration adjusts client weight using:
  - S_{Jac}^k: structural similarity of weights
  - S_i^k: cosine similarity of weights to global model
  - D_i^k: norm bounded L2 distance of weights
- Client weight: $r_i^k = c_1 * S_{Jac}^k + x_2 * S_i^k * D_i^k$
- Server model update: $w_{i+1} = \frac{1}{K} \sum_{k=1}^{K} r_i^k * w_i^k$

Technique 3: ACS + Norm bounding

- Each iteration adjusts client weight using:
  - S_{Jac}^{k}: structural similarity of weights
  - S_{i}^{k}: cosine similarity of weights to global model
  - D_{i}^{k}: norm bounded L2 distance of weights
- Client weight: $r_i^k = c_1 * S_{Jac}^{k} + x_2 * S_i^k * D_i^k$
- Server model update: $w_{i+1} = \frac{1}{K} \sum_{k=1}^{K} r_i^k * w_i^k$

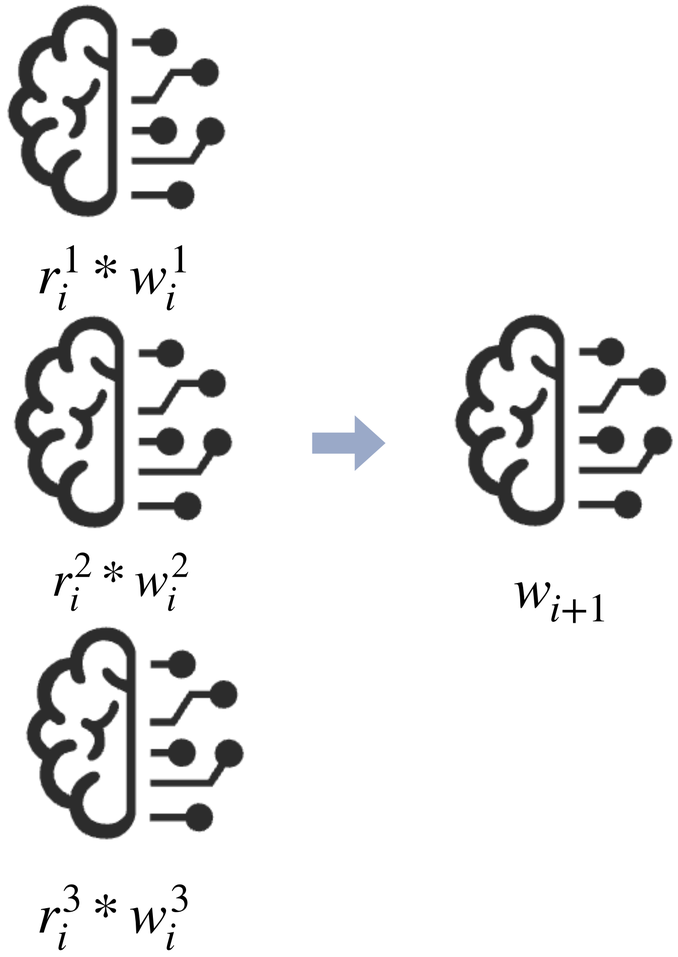
$r_i^1 * w_i^1$

$r_i^2 * w_i^2$

$r_i^3 * w_i^3$

Technique 3: ACS + Norm bounding

- Each iteration adjusts client weight using:
  - S_{Jac}^k: structural similarity of weights
  - S_i^k: cosine similarity of weights to global model
  - D_i^k: norm bounded L2 distance of weights
- Client weight: $r_i^k = c_1 * S_{Jac}^k + x_2 * S_i^k * D_i^k$
- Server model update: $w_{i+1} = \frac{1}{K} \sum_{k=1}^{K} r_i^k * w_i^k$



Technique 3: ACS + Norm bounding

- Each iteration adjusts client weight using:
  - S_{Jac}^k: structural similarity of weights
  - S_i^k: cosine similarity of weights to global model
  - D_i^k: norm bounded L2 distance of weights
- Client weight: $r_i^k = c_1 * S_{Jac}^k + x_2 * S_i^k * D_i^k$
- Server model update: $w_{i+1} = \frac{1}{K} \sum_{k=1}^{K} r_i^k * w_i^k$

Advantages:

- Bounds update norms without hurting convergence
- Preventing attackers scaling model updates to poison FL
- Theoretical guarantee: bounded iteration-wise shift

Evaluation Setup

[1] King, Isaiah J., and H. Howie Huang. "Euler: Detecting network lateral movement via scalable temporal link prediction." *ACM Transactions on Privacy and Security* 26.3 (2023): 1-36.
[2] Khoury, Joseph, et al. "Jbeil: Temporal graph-based inductive learning to infer lateral movement in evolving enterprise networks." *2024 IEEE Symposium on Security and Privacy (SP)*. IEEE, 2024.

Evaluation Setup

- Datasets:
  - OpTC (8 days, 92M events)
  - LANL (58 days, 1B events)
  - Pivoting (1 day, 74M events)
- GNIDS:
  - Euler (transductive)[1]
  - Jbeil (transductive + inductive)[2]

[1] King, Isaiah J., and H. Howie Huang. "Euler: Detecting network lateral movement via scalable temporal link prediction." *ACM Transactions on Privacy and Security* 26.3 (2023): 1-36.
[2] Khoury, Joseph, et al. "Jbeil: Temporal graph-based inductive learning to infer lateral movement in evolving enterprise networks." *2024 IEEE Symposium on Security and Privacy (SP)*. IEEE, 2024.

Evaluation Setup

- Datasets:
  - OpTC (8 days, 92M events)
  - LANL (58 days, 1B events)
  - Pivoting (1 day, 74M events)
- GNIDS:
  - Euler (transductive)[1]
  - Jbeil (transductive + inductive)[2]

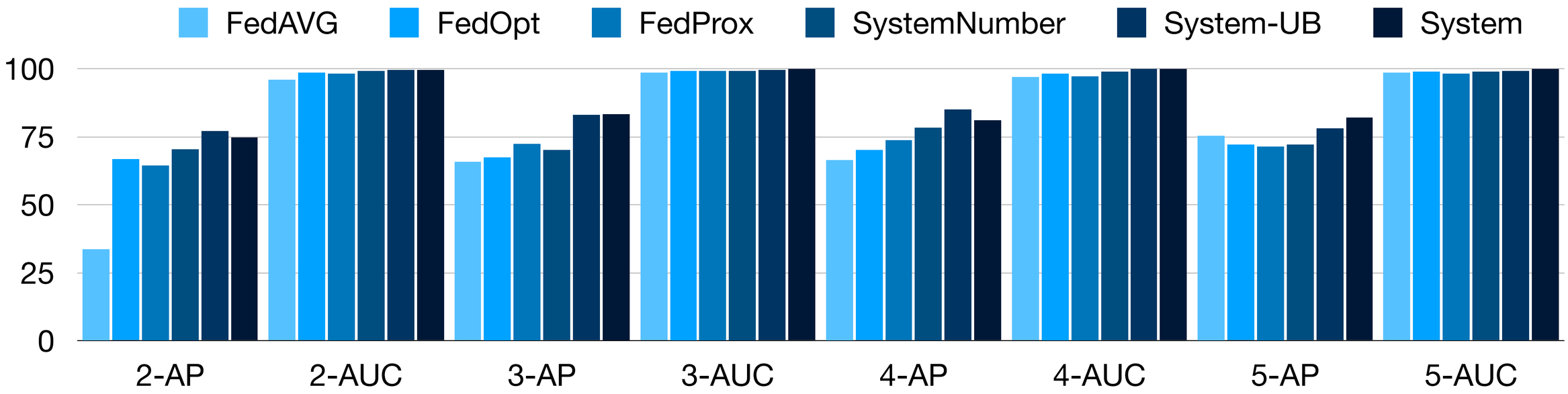
- Baselines:
  - Non-FL
  - FedAvg
  - FedAvg-N: Simple weighted averaging of client models.
  - FedOpt: Use other adaptive optimizers (e.g., Adam) to improve convergence.
  - FedProx: Don’t let local models drift too far from the global model.

[1] King, Isaiah J., and H. Howie Huang. "Euler: Detecting network lateral movement via scalable temporal link prediction." *ACM Transactions on Privacy and Security* 26.3 (2023): 1-36.
[2] Khoury, Joseph, et al. "Jbeil: Temporal graph-based inductive learning to infer lateral movement in evolving enterprise networks." *2024 IEEE Symposium on Security and Privacy (SP)*. IEEE, 2024.

Evaluation

Effectiveness

- ENTENTE achieves 74–84% AP, >99% AUC outperforms all FL baselines
- Even beats Non-FL (centralized training)
- Reason: ENTENTE handles long-sequence + heterogeneity better



Evaluation

Scalability

- Low communication overhead
- Only model weights transmitted (Euler: ~1.94MB)
- Training time scales reasonably with number of clients
- GPU memory stable (~3.2GB per client)

K	Training (s)	AP (%)	AUC (%)	CPU (MB)	GPU (MB)
5	138.51	0.67	97.16	1521.21	3343.39
10	224.38	0.77	98.72	866.13	3256.13
20	497.46	1.05	99.12	522.87	3308.8
50	5111.43	0.2	97.99	1341.3	N/A

Evaluation

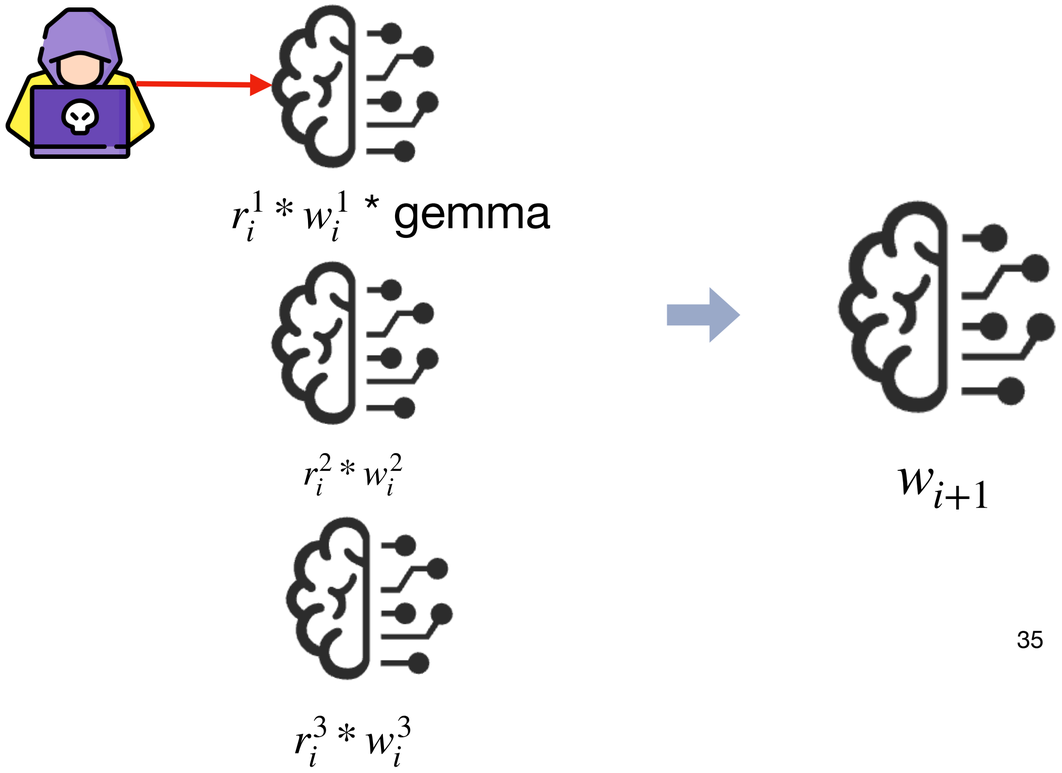
Robustness Against Poisoning

- Attack model: Model poisoning + malicious edge injection
- Attackers scale updates by gemma

Evaluation

Robustness Against Poisoning

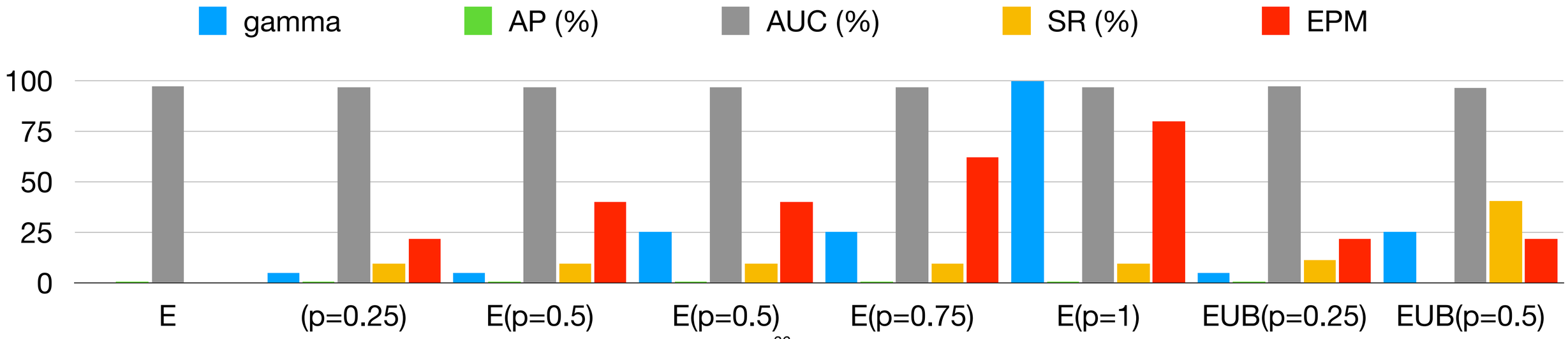
- Attack model: Model poisoning + malicious edge injection
- Attackers scale updates by gemma



Evaluation

Robustness Against Poisoning

- Attack model: Model poisoning + malicious edge injection
- Attackers scale updates by gemma
- ENTENTE bounds attack success rate to <10%
- ENTENTE-UB (no norm bounding) → training diverges



Conclusion

- ENTENTE
  - Enables cross-silo GNIDS without data sharing
  - Achieves effectiveness, scalability, robustness
  - Outperforms all FL baselines
  - Shows FL can beat centralized GNIDS in some cases

Future Work

- Better Ground Truth & Data Quality
- Real-World Distributed Deployment
- Stronger & Broader Attack Evaluation
- Improved Privacy–Robustness Tradeoff

Thanks for listening!

Back up

Limitation

1. LANL redteam labels are coarse
2. Clustering (MBM) may not reflect real org boundaries
3. No fully distributed deployment tested
4. DP integration reduces accuracy significantly

Evaluation

Scalability

- Low communication overhead
- Only model weights transmitted (Euler: ~1.94MB)
- Training time scales reasonably with number of clients
- GPU memory stable (~3.2GB per client)

K	Training (s)	AP (%)	AUC (%)	CPU (MB)	GPU (MB)
5	138.51	0.67	97.16	1521.21	3343.39
10	224.38	0.77	98.72	866.13	3256.13
20 (15 valid clients)	497.46	1.05	99.12	522.87	3308.8
50 (35 valid clients, CPU only)	5111.43	0.2	97.99	1341.3	N/A