Adaptive Contextual Threat Benchmark (ACT-Bench): AI-Enhanced Benchmarking Solution for IDS/IPS Performance

---

1. Problem Statement Description

Intrusion Detection Systems (IDS) and Intrusion Prevention Systems (IPS) are essential for monitoring network traffic, identifying suspicious activities, and protecting against unauthorized access or malicious threats. However, as cyber threats grow more sophisticated, the need for adaptive and dynamic benchmarking methods becomes critical. Traditional IDS/IPS benchmarking typically relies on static attack patterns and known traffic profiles, which do not accurately reflect the diverse, evolving nature of modern cyber threats. This limitation means that traditional tests may not adequately assess an IDS/IPS's effectiveness against real-world, adaptive threats.

In today's environment, cyber attackers are increasingly using advanced techniques that adapt to detection mechanisms, making it challenging for static benchmarking methods to fully capture the IDS/IPS's ability to handle these dynamic attacks. Attackers often employ polymorphic tactics, where the characteristics of malicious traffic change over time to evade detection. This necessitates a benchmarking approach that can simulate attacks which are not only realistic but also capable of evolving and adapting in response to the IDS/IPS's defenses.

Another challenge lies in balancing detection speed with minimal impact on network performance. An IDS/IPS must respond to threats rapidly, but high detection responsiveness can introduce latency, potentially slowing down legitimate network traffic. Excessive latency can impact user experience and network efficiency, creating a trade-off between security and performance. To meet real-world needs, a benchmark must evaluate how well an IDS/IPS can maintain security standards without compromising network performance, especially during high-traffic situations.

---

2. Solution Proposed in Reference RFC

The solution for benchmarking IDS/IPS devices is based on **RFC 9411**, which provides a structured methodology for evaluating network security devices, including IDS/IPS systems. RFC 9411 focuses on reproducibility and transparency in benchmarking by defining performance metrics and standard testing environments. It outlines several key aspects that are essential for a comprehensive evaluation of IDS/IPS performance, including throughput, latency, and packet handling under varied traffic profiles.

Key Aspects of RFC 9411

1. **Throughput and Latency Benchmarking**:
    - RFC 9411 emphasizes the importance of measuring data throughput and latency as primary indicators of IDS/IPS performance. Throughput reflects the amount of data that an IDS/IPS can process in a given timeframe, while latency indicates the delay introduced by the IDS/IPS as it inspects packets.

- These metrics are crucial because they determine how well an IDS/IPS can manage high-traffic loads and impact network performance. By focusing on throughput and latency, RFC 9411 aims to establish baseline measurements that can be reproduced and compared across different devices.

2. **Traffic Profiles**:
    - RFC 9411 recommends testing IDS/IPS systems using a variety of traffic profiles like real-world conditions. These profiles may include regular traffic patterns, attack traffic, and a mixer of both. The purpose of using diverse traffic profiles is to evaluate how well an IDS/IPS can handle different types of network loads and identify potential performance limitations.
    - The recommendation aligns with ACT-Bench's objective to create a dynamic testing environment, where adaptive and evolving attack patterns (simulated by GANs and RL models) stress-test the IDS/IPS in a realistic manner. Although RFC 9411 does not explicitly cover adaptive or AI-driven profiles, it establishes a foundation for evaluating the impact of traffic complexity on IDS/IPS performance.
3. **Reproducibility and Transparency**:
    - A core principle of RFC 9411 is that benchmarking tests should be reproducible and transparent. This means that all test configurations, metrics, and results must be documented in detail, allowing other testers to replicate the benchmarking process.
    - ACT-Bench follows this principle by maintaining clear documentation of all implemented features, including traffic generation, latency measurement, and model performance metrics. By ensuring that tests are reproducible, ACT-Bench can provide consistent insights into IDS/IPS performance across different test environments and scenarios.
4. **Performance Degradation and Reporting**:
    - RFC 9411 suggests tracking performance indicators over time to observe trends in IDS/IPS performance. This is especially relevant when testing under increasing traffic loads, where performance degradation (e.g., increased latency or packet drops) becomes evident.
    - In ACT-Bench, this is extended through the **Latency vs. Action Responsiveness (LAR) Score** and **Cumulative Detection Degradation (CDD) Metric**. These metrics help measure and visualize how well an IDS/IPS balances detection responsiveness with latency and how performance may degrade under heavy loads or complex attacks.

---

3. Proposed Solution/Optimization by the Team

The ACT-Bench solution introduces AI-driven optimizations that extend traditional IDS/IPS benchmarking to address modern, dynamic cyber threats. This solution integrates advanced

machine learning techniques, including **Generative Adversarial Networks (GANs)** and **Reinforcement Learning (RL)**, to create adaptive threat simulations. Additionally, the **Latency vs. Action Responsiveness (LAR) Score** helps balance detection speed with minimal impact on legitimate network performance. Together, these features create a more robust, realistic, and insightful benchmarking framework.

a. AI-Driven Threat Simulation

Purpose: The AI-Driven Threat Simulation is designed to test the IDS/IPS against realistic, evolving attack patterns that mirror real-world adversarial tactics. Traditional testing with static attacks does not accurately represent the adaptive nature of modern cyber threats. By introducing GANs and RL, the team aims to create a simulation environment that rigorously challenges IDS/IPS systems, pushing them beyond conventional scenarios.

1. **Generative Adversarial Networks (GANs)**:
    - **Role in Threat Simulation**: GANs are used to generate synthetic attack data that simulates real-world tactics, techniques, and procedures (TTPs) employed by attackers. The GAN framework consists of a generator and a discriminator:
        - **Generator**: Creates synthetic attack patterns, aiming to make them indistinguishable from actual cyber threats.
        - **Discriminator**: Attempts to differentiate between real attack patterns and the generated ones, pushing the generator to create increasingly sophisticated and realistic attacks.
    - **Focus on Polymorphic Attacks**: GANs are particularly well-suited for producing polymorphic attacks, which continuously change in appearance and behavior. This forces the IDS/IPS to adapt to dynamic and evolving attack patterns, testing its ability to detect threats that evade traditional signature-based detection.
2. **Reinforcement Learning (RL)**:
    - **Role in Threat Simulation**: RL introduces an adaptive element where the attack simulation learns from the IDS/IPS's responses, creating an environment where threats continuously adjust to evade detection.
    - **Mechanism**: The RL model, acting as an “agent,” interacts with the IDS/IPS system, receiving feedback on whether its attack patterns were detected or not. Based on this feedback, the RL agent modifies its tactics to improve its chances of bypassing detection.
    - **Real-Time Adaptation**: This adaptive approach enables the IDS/IPS to face threats that respond in real-time to its defenses, closely mimicking the behavior of real-world attackers who adjust their strategies based on observed defensive mechanisms.

These AI-driven simulations allow ACT-Bench to provide a diverse, challenging threat landscape that goes beyond static testing, creating a comprehensive assessment of IDS/IPS resilience.

b. Latency vs. Action Responsiveness (LAR) Score Optimization

Purpose: The LAR Score measures how well an IDS/IPS balances detection effectiveness with minimal impact on legitimate network traffic. This feature addresses the need for real-time security solutions that do not disrupt regular network operations, a challenge that most IDS/IPS systems face in dynamic network environments.

1. **LAR Score Calculation using Deep Q-Networks (DQN)**:
    - **Role in Optimization**: The LAR Score calculation uses reinforcement learning to dynamically balance IDS/IPS responsiveness with latency. This metric allows the IDS/IPS to optimize its settings for maximum detection accuracy without sacrificing network performance.
    - **Mechanism**: Using DQN, the IDS/IPS learns to maximize its LAR Score by balancing the trade-off between quick detection and minimal latency. Positive feedback is given when the IDS/IPS detects a threat without introducing significant latency, helping the system learn the optimal configuration for different traffic conditions.
2. **Predictive Latency Modeling**:
    - **Role in Anticipating Latency**: The latency model forecasts potential latency spikes based on factors such as current network load, traffic type, and IDS/IPS settings. This proactive approach allows the IDS/IPS to adjust before latency impacts network performance.
    - **Mechanism**: A regression model (like Decision Trees or LSTM) is trained on historical traffic data to predict latency under various conditions. By providing real-time insights into expected latency, the IDS/IPS can preemptively modify configurations to maintain responsiveness without compromising legitimate traffic flow.
    - **Proactive Tuning**: The combination of predictive modeling and LAR Score optimization allows the IDS/IPS to fine-tune its settings in real time, balancing detection speed and system responsiveness dynamically.

Together, the LAR Score and latency modeling ensure the IDS/IPS can handle high-traffic situations efficiently, prioritizing threat detection while keeping network latency within acceptable limits.

c. Tools and Frameworks for Implementation

To achieve these optimizations, the team uses a variety of tools and frameworks, selected for their compatibility with machine learning and network security applications.

1. **Machine Learning Libraries**:
    - **PyTorch and TensorFlow**: Both libraries provide flexibility and support for implementing GANs and RL models, with PyTorch being especially suitable for RL due to its dynamic computation graph.
    - **Keras**: As a high-level API built on TensorFlow, Keras facilitates rapid prototyping of GANs and manages the training loop for generator and discriminator networks.

2. **GAN-Specific Libraries**:
    - **Pytorch-GAN or TensorFlow GAN**: These libraries include pre-built architectures and utilities for working with GANs, accelerating the development and testing process.
3. **Reinforcement Learning Libraries**:
    - **OpenAI Gym**: Provides a standardized environment for training RL models, with the ability to create custom environments that mimic IDS/IPS responses.
    - **Stable-Baselines3**: A PyTorch-based library offering reinforcement learning algorithms like Deep Q-Networks (DQN) and Proximal Policy Optimization (PPO), which are adaptable for adaptive threat simulations and LAR score optimization.
4. **Data Simulation and Attack Modeling Tools**:
    - **MalGAN**: A GAN-based tool designed for malware generation, useful for creating specific malware-like attack simulations.
    - **Cybersecurity Datasets (CICIDS 2017, UNSW-NB15)**: These datasets offer a range of attack data that can train GAN and RL models, ensuring that the generated threats are realistic and representative of real-world attacks.
5. **Development and Testing Environment**:
    - **Docker**: Used to containerize the simulation, ensuring easy dependency management and consistent deployment across different environments.
    - **Kubernetes**: Supports distributed testing and scalability, enabling the simulation of multiple IDS/IPS configurations or handling high-traffic scenarios.
6. **Evaluation and Benchmarking Tools**:
    - **Scikit-learn**: Provides tools for evaluating model performance, including precision, recall, and F1-score, which help measure the effectiveness of IDS/IPS against simulated attacks.

These tools and frameworks streamline the implementation process, allowing the team to efficiently develop, test, and refine the GAN, RL, and latency optimization models.

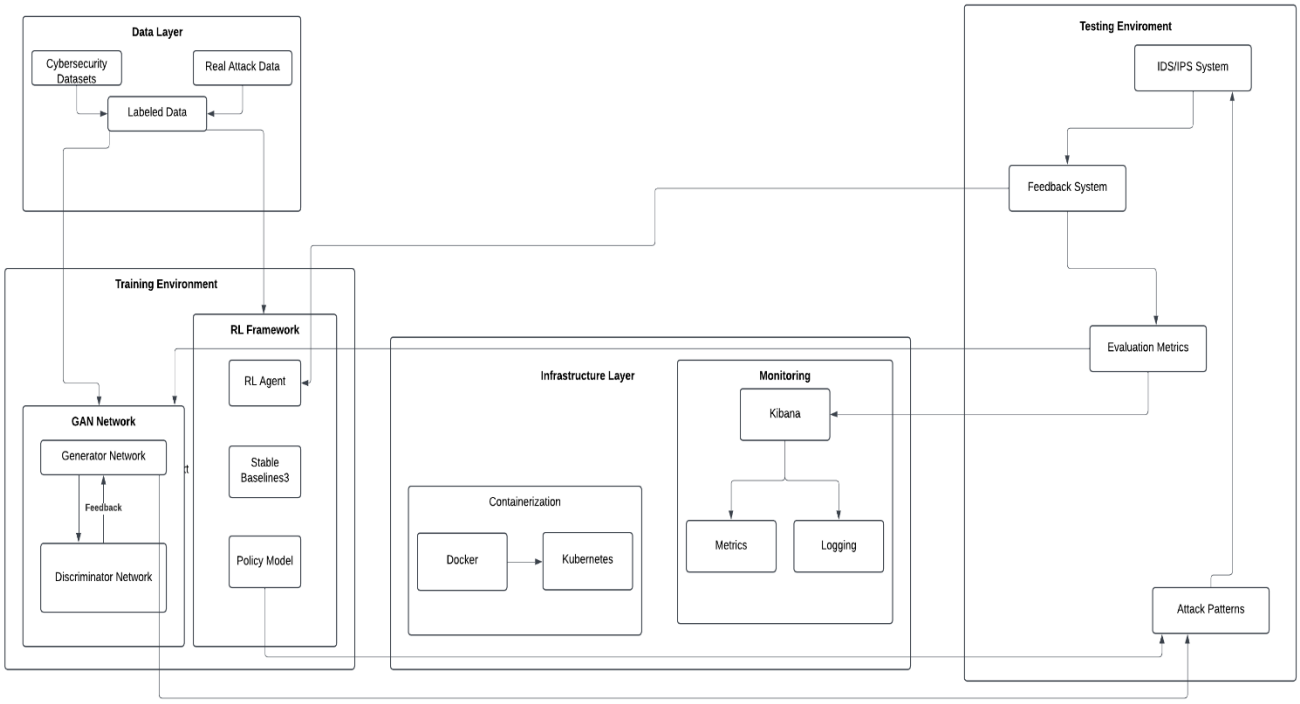
Extensions to RFC 9411 in the ACT-Bench Solution

While RFC 9411 establishes a comprehensive benchmarking foundation, the ACT-Bench solution extends its scope in the following ways:

1. **AI-Driven Threat Simulation**:
    - Unlike the static traffic profiles recommended in RFC 9411, ACT-Bench introduces AI-driven models like **Generative Adversarial Networks (GANs)** and **Reinforcement Learning (RL)**. GANs are used to create realistic, evolving attack patterns that

simulate adaptive cyber threats, while RL introduces real-time adaptability, allowing attack patterns to adjust in response to IDS/IPS defenses.

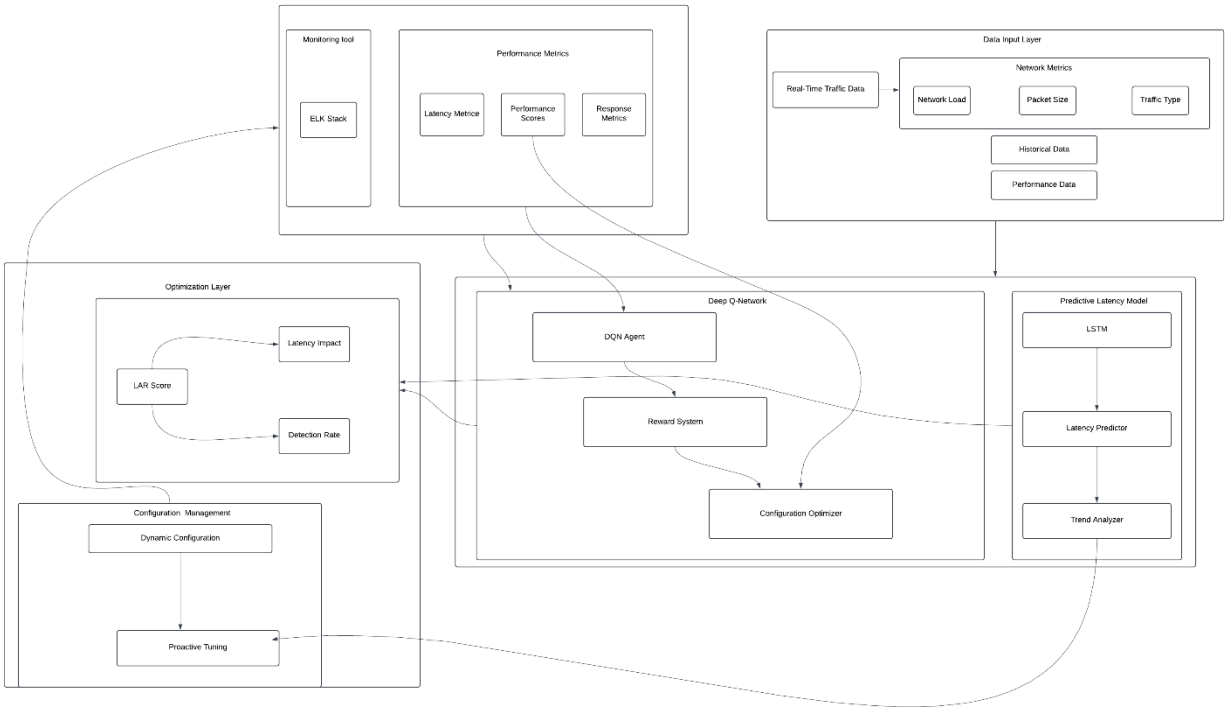
- This extension addresses the gap in RFC 9411 concerning the evaluation of IDS/IPS under evolving, adaptive threats, offering a more realistic testing scenario for modern network security challenge



2. **Latency vs. Action Responsiveness (LAR) Score**:

- Although RFC 9411 focuses on throughput and latency as separate metrics, it does not provide a metric that dynamically balances detection speed with latency impact. The **LAR Score** is introduced in ACT-Bench to measure this balance, using reinforcement learning techniques to optimize IDS/IPS performance while minimizing latency.

- This metric reflects real-world requirements, where an IDS/IPS must respond quickly to threats without significantly delaying legitimate traffic. The LAR Score provides a more nuanced view of IDS/IPS performance by taking both security and network efficiency into account.



Summary of Proposed Solution

The proposed ACT-Bench solution combines AI-driven threat simulations and a latency-aware scoring system to evaluate IDS/IPS performance under conditions that closely mimic real-world cyber threats. By utilizing GANs and RL for adaptive threat modeling, and the LAR Score to balance latency and detection responsiveness, the solution provides a comprehensive and realistic assessment of IDS/IPS systems. This approach demonstrates the team's commitment to pushing the limits of traditional benchmarking and creating a tool that reflects the complexities of modern cybersecurity environments.

---

5. Timeline of Delivery

Timeline	Task	Description
11/11/24-14/11/24	Project Setup & Environment Configuration	Set up the development environment using Docker for containerization and configure Kubernetes for distributed testing. Install necessary libraries (PyTorch, TensorFlow, etc.).
15/11/24-18/11/24	Implement Initial GAN Model	Develop and train the initial GAN model using available cybersecurity datasets (e.g., CICIDS 2017, UNSW-NB15) to generate synthetic attack patterns.

Timeline	Task	Description
19/11/24-21/11/24	Train GAN for Polymorphic Attack Generation	Refine the GAN to generate polymorphic, evolving attack patterns that adapt over time. Adjust generator and discriminator parameters for realistic threat simulation.
22/11/24-23/11/24	Implement Reinforcement Learning (RL) Model	Set up the RL environment to allow interaction with the simulated IDS/IPS. Train the RL agent to create adaptive attack strategies based on IDS/IPS feedback.
24/11/24-25/11/24	Integrate GAN and RL for Combined Threat Simulation	Integrate GAN and RL models to create an adaptive threat simulation environment. Test combined model by evaluating the IDS/IPS's response to evolving attacks.
26/11/24-27/11/24	Develop Latency vs. Action Responsiveness (LAR) Score	Implement the LAR Score metric using a Deep Q-Network (DQN) model to balance detection responsiveness with minimal latency.
28/11/24	Implement Predictive Latency Modeling	Train a latency prediction model (using Decision Trees or LSTM) to anticipate network latency based on traffic and IDS/IPS configuration.
29/11/24	Initial Testing & Performance Evaluation	Test GAN and RL-generated attacks against the IDS/IPS. Measure performance metrics (e.g., precision, recall) and evaluate LAR Score under different traffic loads.
30/11/24	Dashboard Setup & Data Visualization	Create a visualization dashboard using Kibana/Grafana. Display key metrics like LAR Score, latency, and IDS/IPS effectiveness against adaptive threats.
30/11/24	Final Testing, Documentation, and Report Preparation	Conduct final end-to-end testing, prepare the project documentation, and compile reports and insights for presentation at the hackathon.

---

6. References

1. **RFC 9411** - *Benchmarking Methodology for Network Security Device Performance*. March 2023. Internet Engineering Task Force (IETF).

2. **PyTorch & TensorFlow** - Core deep learning libraries used to implement Generative Adversarial Networks (GANs) and reinforcement learning models for threat simulation.
3. **Stable-Baselines3** - A reinforcement learning library compatible with PyTorch, providing various algorithms such as Deep Q-Network (DQN) for optimizing the LAR score.
4. **OpenAI Gym** - A platform providing an environment for training reinforcement learning models, with the flexibility to create custom environments for IDS/IPS benchmarking.
5. **Cybersecurity Datasets (CICIDS 2017, UNSW-NB15)** - These datasets provide realistic attack data for training GAN and RL models to simulate adaptive, real-world threats.
6. **Scikit-Learn** - A machine learning library for implementing regression models, performance metrics, and evaluation tools for latency and detection measurement.
7. **ELK Stack (Elasticsearch, Logstash, Kibana)** - A widely-used suite for monitoring and visualizing network data, integrated to display LAR scores, latency trends, and IDS/IPS performance metrics.
8. **Docker & Kubernetes** - Tools for containerizing and scaling the project environment, allowing distributed testing with multiple IDS/IPS configurations.

---

7. Conclusion

The ACT-Bench project introduces a modern, AI-enhanced solution for benchmarking IDS/IPS systems by simulating real-world, adaptive threats and measuring their impact on network performance. Traditional IDS/IPS benchmarking often relies on static attack patterns, which do not fully capture the sophistication of contemporary cyber threats. By incorporating Generative Adversarial Networks (GANs) and Reinforcement Learning (RL), ACT-Bench creates a dynamic, evolving threat environment that tests the resilience of IDS/IPS systems beyond conventional scenarios.

Additionally, the **Latency vs. Action Responsiveness (LAR) Score** provides a new metric to balance security with network efficiency, ensuring that IDS/IPS measures do not disrupt normal operations. With predictive latency modeling and a comprehensive visualization dashboard, ACT-Bench gives users actionable insights into IDS/IPS performance under diverse traffic conditions, highlighting areas for potential optimization.

In sum, ACT-Bench's innovative approach to adaptive threat simulation and latency-aware scoring sets a new standard for IDS/IPS benchmarking, offering valuable insights for cybersecurity teams seeking to enhance network defenses. This project is not only technically robust but also aligns with the evolving needs of modern network security, making it a valuable contribution to the field of IDS/IPS evaluation.