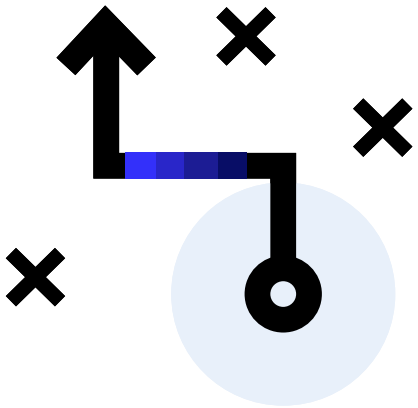
WHITE PAPER

From Compliance to Resilience: The Case for Purple Teaming

GUIDEPOINT®
SECURITY

TABLE OF CONTENTS



An Introduction to Purple Teaming

Many organizations treat cybersecurity testing as a once-per-year check-box exercise that fulfills a regulatory requirement. However, penetration tests are one of the most powerful tools in your arsenal – when done right. Purple teaming exercises can elevate your security testing from a simple compliance task to a truly proactive, collaborative effort that strengthens your organization's resilience against real-world threats. By combining the offensive tactics of a red team with the defensive expertise of a blue team, purple teaming transforms traditional testing into an ongoing learning opportunity. You'll reveal gaps, sharpen detection and response capabilities, and ensure that your security investments deliver measurable improvements in your ability to withstand attacks.

In this whitepaper, we explore the strategic value of purple teaming, discuss common approaches to the practice, and dig into what makes a purple team successful. We will round out the discussion with some ideas to streamline purple team engagements, and we'll discuss how to incorporate this strategic exercise into your ongoing efforts to reduce risk and improve your security posture.

WHAT IS PURPLE TEAMING?

A Purple Team Assessment (PTA) is a collaborative engagement between a red team (offensive security experts) and blue team (defensive security teams) to test, evaluate, and improve an organization's security posture. Unlike traditional penetration testing, which often isolates offensive and defensive efforts, purple teaming emphasizes real-time collaboration and shared learning. By simulating realistic cyberattacks and observing how defenses respond, both teams work together to identify gaps, refine detection capabilities, and enhance incident response processes. The goal of a purple team is to create a continuous feedback loop that strengthens the organization's ability to detect, respond to, and recover from threats.

A PTA is different from an internal penetration test or red team assessment. The primary goal is not to identify and assess vulnerabilities commonly managed through vulnerability management programs. Nor will the resulting deliverables of a PTA include findings with severity levels. Instead, purple team deliverables will focus on metrics such as detections, alerts, preventions, and response times. It may also provide guidance on the creation of detection logic and threat hunting playbooks.

Below, we have a more detailed outline showing the differences between internal penetration tests vs. red team assessments vs. purple team assessments.

Internal Penetration Test

- Emulation of realistic actors is not a focus
- Assessment scopes and rules of engagement are defined and adhered to
- Interacting with the blue team is typically limited to clearing security alerts
- Identifying misconfigurations and security vulnerabilities is a key focus

Red Team Assessment

- Realistic threat actors and TTPs are emulated
- Engagement guardrails should be limited
- Blue team evasion is a key principle
- Goals of a red team assessment are specifically tailored prior to beginning the engagement

Purple Team Assessment

- Realistic threat actors and TTPs are emulated
- CTI assists with creating realistic campaigns
- Red and blue teams collaborate to test the efficacy of logging and detection capabilities
- The goal is to exercise and advance the blue team with logging, detection, and preventing capabilities

"Prevention is a goal, but detection and response are a REQUIREMENT"

Jorge Orchilles, Author: *Running Your First Purple Team Exercise - Understand the Kill Chain, Emulation & Response*

Core Components of a Purple Team Assessment

A successful purple team assessment relies on several key elements working in unison. These components ensure that offensive and defensive security efforts are strategically aligned, measurable, and beneficial to the organization's security posture. With the blue team, red team, and team coordinators all working within the same scope, toward the same goals, and measuring based on pre-determined objectives, organizations can quickly assess their security tools and practices and make meaningful improvements toward security maturity.

THE BLUE TEAM

The blue team represents the organization's frontline threat defenders. They play an active role in purple team engagements by collaborating with the red team to develop and refine attack campaigns based on public cyber threat intelligence, internal intelligence, and an understanding of the existing threat landscape. Most importantly, the blue team will grant access to the environment, oversee security controls, and are involved with detection engineering tasks. The goal of a purple team assessment is to strengthen the defensive capabilities of the organization; thus, the blue team is considered the most critical component of the exercise.

THE RED TEAM

The red team emulates threat actors and adversary tactics in close coordination with the blue team. Their role includes developing the necessary tools and infrastructure, conducting research, and executing planned procedures that simulate real-world attacks. Every step, along with all low-level details and observable outcomes, are documented in detail by the red team in a central system for post-operation review. The red team's insights are essential to testing and enhancing the organization's detection and response readiness.

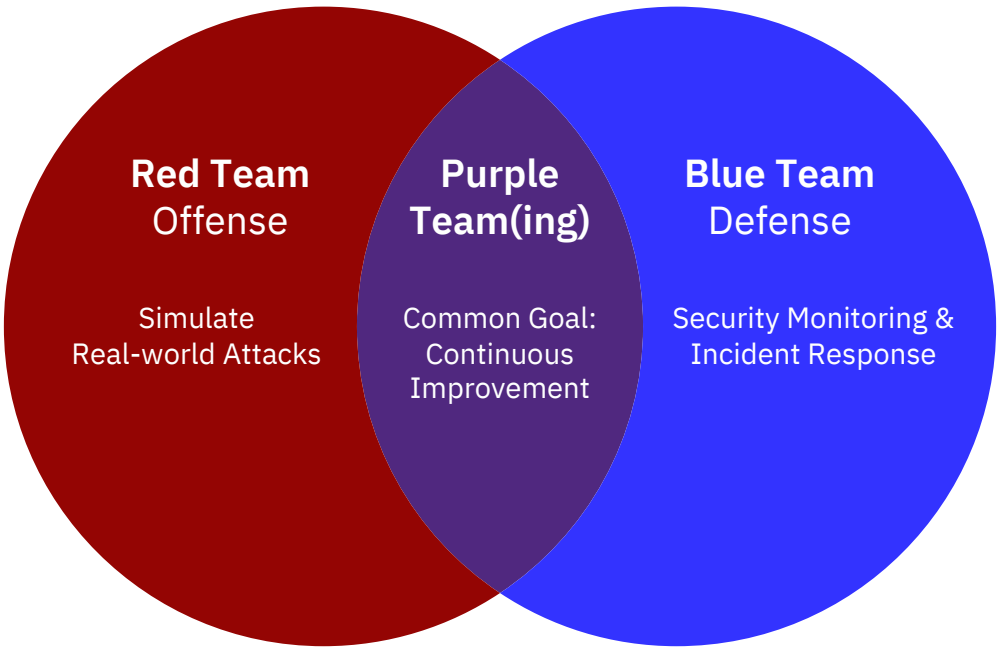


Figure: Purple Teaming brings together offensive and defensive security strategies to facilitate continuous improvement

SCOPING, GOALS, & OBJECTIVES

Clear scoping and defined objectives are essential to maximizing the value of a purple team assessment. This process begins with understanding the existing threat landscape, current security controls, and responsible teams. Security gaps, such as limited visibility into east-west network traffic, misconfigured policies or tools, or broken SIEM logs are identified and captured as known issues during the scoping exercise. This level of detailed information, if provided early in the planning process, ensures all aspects of the existing security program are accounted for when creating campaigns, emulating threat actors, and prioritizing test procedures. Once the scope of the engagement is determined, the team develops goals and objectives. It is important to differentiate goals from objectives. Goals describe what the blue team hopes to improve or achieve through the course of the assessment. Objectives, on the other hand, are specific, measurable expectations tied to individual campaigns or procedures. These measurable objectives are vital for assessing success and guiding remediation efforts post-engagement.

TEAM COORDINATORS

Team coordinators act as neutral facilitators to keep the assessment on track and aligned to its predefined goals. The coordinators prevent distractions and deviations from critical operations. Ideally, having three coordinators allows for a tiebreaker in the event of differing opinions. These individuals can also hold other roles within the red or blue teams, as full-time dedication to coordination is not always necessary.

Developing Meaningful Metrics

Measuring the success of a purple team assessment goes far beyond counting alerts or simply timing responses. To truly understand and improve an organization's security posture, metrics must be aligned with the goals of the engagement and the maturity of the security program. The following introduces how event categorization and performance metrics inform key categories used to assess detection and response, identify gaps, and guide continuous improvement.

EVENT TYPES

Understanding and categorizing different types of events allows the purple team to frame discussions around the effectiveness of detection and response playbooks and procedures. Events can be broken down into two rudimentary categories:

- **Detectable Events:** These are actions that should generate an alert or be immediately flagged by defensive technologies. For example, the detonation of a known malicious executable should trigger a detection event.
- **Forensic Events:** These are actions that may not prompt immediate alerts, but that do leave behind evidence that can be uncovered through forensic investigation. A common example is an operating system logging process creation. If an untrusted executable runs without triggering endpoint detection and response (EDR), its presence may only be revealed in this log post-facto.

PERFORMANCE METRICS

In addition to event categorization, two commonly used performance metrics help track how well security teams are detecting and responding to threats:

- **Mean Time to Detect (MTTD):** The duration from the start of an attack to when defenders actively triage the first alert.
- **Mean Time to Respond (MTTR):** The duration from the start of an attack to when the threat is fully contained or the corresponding response playbook is completed.

IDENTIFYING SUCCESS CRITERIA

These metrics should not be viewed in isolation. Purple team assessments are dynamic by design. Successful purple team engagements rely on flexibility so that they can evolve based on real-time discoveries, shifting priorities, or unexpected challenges. As such, relying solely on traditional metrics can present a skewed picture, especially when procedures are paused, adjusted, or deprioritized mid-assessment.

To ensure metrics remain meaningful, it's important to define success up front. Success criteria should reflect the nature of the engagement and the specific goals of the team. Consider the following two examples:

1. A suspicious PowerShell command creates an event that is ingested into a SIEM, a detection is automatically created in a ticketing system, and it is triaged quickly by the defenders.

2. An exception was created for an executed payload and the resulting C2 traffic was correctly categorized by NDR.

Each of these scenarios reflects success, just through a different lens. One emphasizes endpoint detection, and the other highlights network-layer defense. Ultimately, aligning definitions of success with team expectations, organizational maturity, and existing security controls is critical.

CONSIDER THIS

Logging is one functional area that often reveals room for improvement throughout the course of a purple teaming engagement. Incomplete visibility, delayed log ingestion, and insufficient log retention all reveal gaps in security policies and practices that lead to inefficient or ineffective detection and response. By gaining early access to logging infrastructures and corresponding controls, purple teams can define metrics around the logging information that is available so that they can accurately design a campaign.

The Role of Cybersecurity Frameworks in Purple Teaming

There are many different frameworks that can be used by a purple team to make the engagement more efficient. Rather than discuss them all, we will summarize two of the most widely used industry-recognized frameworks.

THE CYBER KILL CHAIN

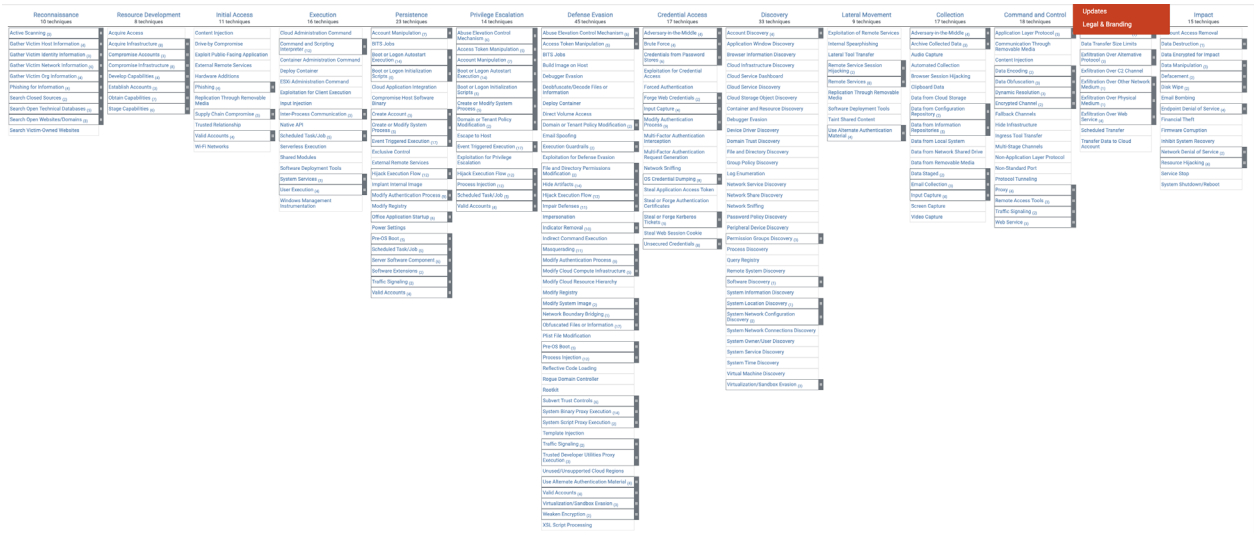
Created in 2011, Lockheed Martin developed a methodology that is widely applied in the cybersecurity industry today. The following describes the distinct stages outlined by the "Cyber Kill Chain" and categorizes an incident into rudimentary stages when observed from a high-level, enabling an organization to focus resources on specific stages and tactics of a cyberattack.

Figure: The Cyber Kill Chain

An attacker conducts reconnaissance to gather target information and identify security gaps. Once an opportunity is identified, an attacker will construct weaponized tooling that will be delivered and executed by exploiting a vulnerability. Persistence on a system is achieved by loading and installing additional code, applications, and services. Upon successful execution, the attacker will focus on establishing a stable command-and-control (C2) session with the target asset. In doing so, it enables the attacker to leverage advanced capabilities and conduct subsequent actions on objectives.

MITRE ATT&CK

The MITRE ATT&CK Framework serves as a cornerstone for any security program and enables cybersecurity professionals to review tactics, techniques and procedures (TTPs) in a single knowledgebase. This equips organizations with a roadmap of high-quality intelligence for an evolving threat landscape. Organizations benefit from this tool by leveraging the information to prioritize defensive strategies and efficiently allocate resources.



Source: *https://attack.mitre.org/*

Security testing teams use this framework to drive comprehensive security assessments and evaluate the logging and detection posture of an organization. Specific procedures from the framework are selected based on predefined objectives and executed inside the target environment. Results are logged in detail and matched to the corresponding MITRE Tactics and Techniques in the ATT&CK matrix. Additionally, this structured approach allows for a thorough yet flexible evaluation of logging and detection security controls across both broad and isolated attack vectors.

COMMON PROCEDURES AND CAPABILITIES

The following capabilities map to specific MITRE tactics and techniques.

- Lateral movement opportunities and how they can be leveraged for privilege escalation.
- The use of untrusted portable executables (PEs) and scripting interpreters.
- Execute malicious operations in memory to avoid artifacts being saved to disk.
- The stealing of credentials from a system or domain controller.
- Leveraging command-and-control (C2) operations for subsequent attacks, Active Directory exploitation, and data exfiltration opportunities.
- Bypassing endpoint security controls to avoid detection from defenders.
- Violating access policies to gain unauthorized access to data for destruction or exfiltration.
- Password spraying attacks to gain visibility into successful initial access by an actor.
- Use built-in tools (LOLBins) and functionality to enumerate the environment and establish persistence.

Tools of the Purple Team

Purple team assessments rely on a diverse and flexible set of tools to simulate realistic attacks and evaluate an organization's detection and response capabilities. This toolkit spans both offensive and defensive domains to ensure a full-spectrum assessment.

Common tools and capabilities include:

- **MITRE ATT&CK-mapped testing** using platforms like Atomic Red Team to validate specific techniques.
- **Custom command-and-control (C2) infrastructure**, including tools such as loaders and Beacon Object Files (BOFs) for stealthy and evasive operations.
- **Adversary emulation platforms** like CALDERA, which allow for complex, scenario-based simulations aligned with real-world threat actors.
- **Vulnerability scanners and exploitation frameworks** to uncover and safely exploit security weaknesses.
- **Network traffic generators** to test monitoring and response across east-west and north-south traffic paths.
- **Custom scripts and utilities** tailored to the target environment, allowing teams to simulate edge cases or bypasses specific to organizational configurations.
- **Threat intelligence platforms** to inform the prioritization of techniques, actors, and TTPs based on relevance to the organization.
- **Blue team visibility tools,** such as SIEMs and EDR consoles, to capture and analyze defensive responses.

Tool selection is just as critical as the tools themselves. The red team must evaluate each engagement's objectives, constraints, and scope to determine which tools will provide the most meaningful results. In many cases, the same technique is executed using multiple methods to test detection and depth of resilience.

CONSIDER THIS

A Kerberoasting attack that, when executed using a custom BOF, may generate entirely different indicators of compromise (IoCs) than an attack performed with a well-known, open-source script. Running both methods helps the blue team identify coverage gaps, refine detection rules, and reduce the risk of false negatives.

By tailoring tools to the context of each engagement, and by varying techniques to mimic real adversary behaviors, purple teams ensure that their assessments are both realistic and actionable.

THE "PYRAMID OF PAIN"

While most purple team tools are technical in nature, some of the most valuable resources are conceptual. **The Pyramid of Pain**, originally developed by David Bianco, is one such tool. It provides a framework for understanding the relative difficulty and impact of detecting and disrupting different types of attacker indicators. For purple teams, the Pyramid of Pain offers a powerful lens through which to design, execute, and evaluate threat actors. By aligning red team actions across all levels of the pyramid, it is possible to determine how responders can inhibit attackers in a meaningful way.

As represented in the Pyramid of Pain graphic, indicators such as domain names, IP addresses, and payload signatures are easily reworked by attackers when identified by defenders. This is because an attacker can quickly and easily regenerate payloads and infrastructure when needed to advance a campaign.

However, TTPs, tooling, and the IOCs involved with attacks are harder to rework in the middle of a campaign. These capabilities are often set up and equipped for specific tasks and operations, and therefore overhauling these components increases the level of difficulty, risk of detection, and overall time taken by an attacker.

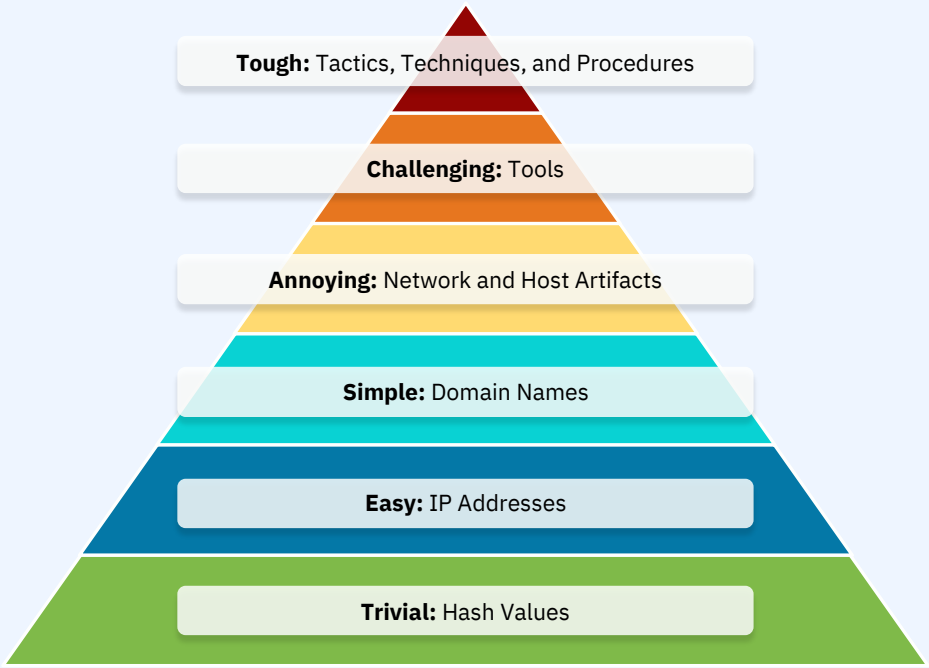


Figure: Pyramid of Pain

The Purple Team Lifecycle: Prepare, Execute, Identify, and Remediate

The goal of a purple team engagement is the creation of a continuous feedback loop that drives improvement. The "Plan-Do-Check-Act" (PDCA) methodology (the "Deming wheel") is a common visual that defines continuous improvement over time. As such, it serves as a foundation for defining the purple team lifecycle: The "Prepare, Execute, Identify, and Remediate" (PEIR) model.

Each stage of the PEIR model represents an important step in the purple team engagement process:

Prepare: During the preparation stage team roles are defined, and expectations are aligned among all participants. Discussion includes relevant TTPs and threat actors pertinent to the organization. One of the initial steps in cyber threat intelligence is aligning procedures with real-world threats. This involves using a GRIT-provided report to identify relevant threat actors, which helps the team build effective and timely campaigns. This information is crucial for teams to plan their procedures accurately and respond to potential threats more effectively.

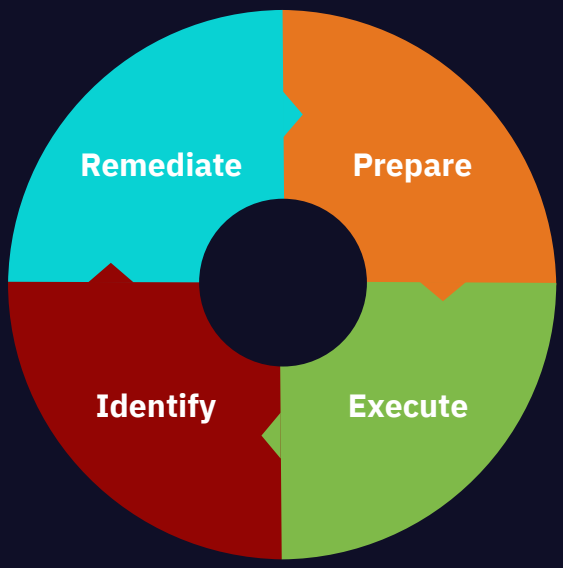
The selected procedures are presented to the red team for their preparation, while the blue team will ensure their readiness and alignment with security tooling. Additionally, using this stage of the assessment to map procedures to MITRE ATT&CK helps develop a holistic view for understanding the scope of the engagement.

Execute: During this phase, the red team executes techniques, and the blue team detects and responds to threats as they arise. Both teams document all actions and responses to ensure a comprehensive record of the procedures, results of the operations, and the subsequent steps taken. This stage emphasizes real-time engagement opportunities, enabling teams to test and refine their strategies to improve resilience.

Identify: After execution, both teams collaboratively review and validate the outcomes. Detection and prevention capabilities are assessed against expectations. Gaps and missed detections are identified and prioritized based on their potential impact and exploitability. During this phase, the blue team may share their screen to walk through indicators of compromise, detections, and log data. Security control performance is rated, and action items are created to guide future improvements. All findings are carefully documented to inform remediation plans and strengthen organizational resilience.

Remediate: The final stage focuses on improving the environment based on previously identified issues. Teams validate the current state of security controls, implement necessary fixes, and re-run selected procedures to confirm the improvements made.

Unlike traditional penetration tests, purple teams don't stop after initial remediation. Instead, the process repeats, testing the improvements made and applying stress to other systems through follow-on engagements. This cyclical process of testing and tuning ensures gaps are not only addressed but also validated. Over time, this strengthens the organization's ability to detect, respond to, and ultimately prevent advanced threats.



- **Remediate:** Implement quick-wins and validate observations
- **Prepare:** Identify tactics, techniques and procedures
- **Identify: I**dentify and priorotize security gaps
- **Execute:** Execute procedures and document observations

Purple Teaming Checklist

The following checklist can be used as a guide for standing up your purple team.

Plan and Align

- [ ] Define clear goals and success criteria for the engagement.
- [ ] Establish priorities and scope (systems, teams, timeframes).
- [ ] Align stakeholders (blue team, red team, leadership).

Delegate Tasks and Define Responsibilities

- [ ] Establish rules of engagement and communication channels.
- [ ] Assign blue and red team members.
- [ ] Designate team coordinators.

Design Threat Models & Campaigns

- [ ] Map procedures to the chosen frameworks.
- [ ] Develop red team tooling and infrastructure.
- [ ] Select relevant threat actors and TTPs for the red team to emulate.
- [ ] Agree on a centralized platform that can be used to track the progress of the engagement.
- [ ] Validate blue team readiness and visibility.

Prep Infrastructure and Test Access

- [ ] Stand up your C2 infrastructure prior to the assessment start date.
- [ ] Determine what categories of malware you will introduce during the engagement.
- [ ] Check connectivity and ensure all systems are available prior to beginning the assessment.
- [ ] Disable EDR and NDR on systems being tested so that they don't introduce delays.
- [ ] Validate all network access and user credentials.

Execute & Iterate

- [ ] Conduct planned campaigns with real-time collaboration.
- [ ] Document all actions, responses, and gaps.
- [ ] Debrief after each scenario or phase.

Conduct Post-Op Sync After Procedures/Operations

- [ ] Prioritize findings and remediation actions.
- [ ] Re-test key improvements (validate "quick wins").
- [ ] Deliver final report and improvement recommendations.

Conclusion

Purple team engagements are most effective when they're structured, collaborative, and tailored to the unique risks and realities of your organization. At GuidePoint Security, our seasoned experts bring deep offensive and defensive experience to help you uncover gaps, improve detection, and elevate your security posture. Whether you're just getting started or looking to mature an existing program, our purple team assessments are designed to deliver actionable insights and measurable outcomes. Let GuidePoint be your trusted partner in building a more resilient, threat-informed defense.

Contact us today to learn more about a customized purple team engagement that fits your organization's unique needs.

GUIDEPOINT®
SECURITY

1900 Reston Metro Plaza, Suite 701, Reston, VA 20190
guidepointsecurity.com • info@guidepointsecurity.com • (877) 889-0132