Effective DDoS Mitigation in Distributed Peering Environments

Prepared by Cisco Systems and Arbor, the security division of NETSCOUT Systems

Two trends are driving the design of today's network edge: the growth in unicast streaming video, and the continued growth of Distributed-Denial-of-Service (DDoS) attacks. The Cisco Visual Networking Index™ has shown a 1270-percent rise in Internet traffic over the last 12 years and projects a threefold increase in the next five years. Cisco estimates that peering bandwidth, in particular, is growing at a steady 30-percent Compound Annual Growth Rate (CAGR), and some carriers are already planning for growth as high as 50 percent per year. This is mostly driven by a steady increase in unicast video streaming and other content sources versus traditional peering expansion with other network providers.

The move to distributed peering

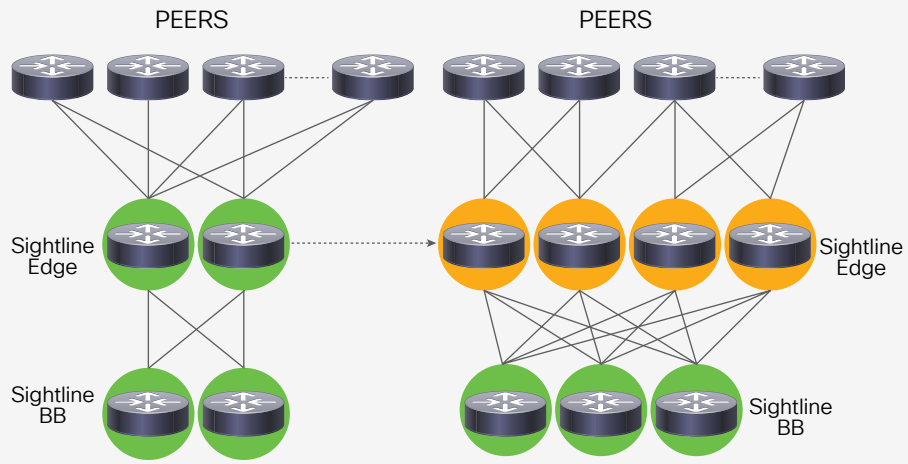
In order to keep up with the increased demand for content, many carriers are re-architecting their networks to bring users closer to content sources and minimize long-haul links to content providers. As content providers continue to invest in deploying regional content caches, private network interconnects are now often being delivered at a regional level. Deployment of distributed fabric architectures based on Clos topologies means carriers can scale peering capacity within a site and across regional sites.

Capacity planning for this growth in legitimate traffic is challenging enough for operators; scaling the network edge with many more distributed peering points also risks exposing your network to inbound DDoS attacks. This new distributed architecture increases the overall threat surface at the network edge and can potentially expose the network core to much larger DDoS attacks.

As events in early 2018 have already proven, terabit-scale volumetric DDoS attacks are becoming more frequent and more damaging. These types of attacks can easily overwhelm internal network capacity and even centralized DDoS mitigation scrubbing facilities. The potential risks to a highly distributed peering architecture are significant since they can expose the network core to significantly higher volumes of malicious traffic than previous architectures.

Contents

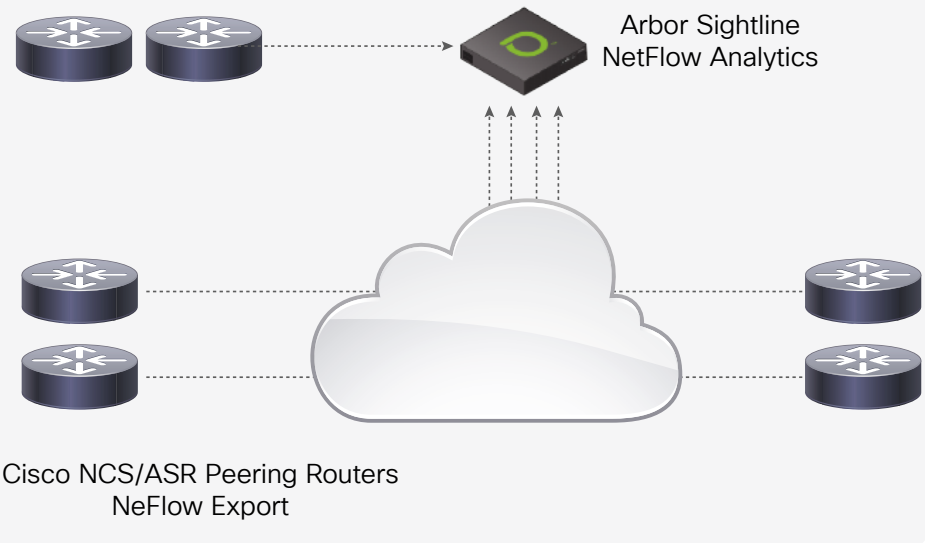
Figure 1. Examples of distributed peering points



Controlling the risk of DDoS

DDoS attacks have long presented challenges for network operators, and especially for Internet service providers (ISPs). Historically, most of these attacks targeted individual users or services on the provider network; DDoS was not considered a direct threat to the network infrastructure itself. However, in today's peering environments, new attacks such as the recent Memcached (a 1.7 Tb/second attack was recorded in early 2018) prove multi-terabit traffic floods are now possible. These attacks present a direct threat to network infrastructure and service availability, and challenge traditional assumptions about DDoS prevention, including relying solely on centralized scrubbing to absorb all DDoS traffic.

Figure 2. Distributed DDoS detection using NetFlow telemetry from edge routers



The DDoS threat landscape

DDoS attacks are a constant threat and are heavily used to attack Internet-based services and networks, with the goal to block access to these services and to disrupt Internet access. The motivations behind these attacks vary widely—from pranks to state-sponsored to cybercrime. A good overview can be found in NETSCOUT Arbor's 13th Annual Worldwide Infrastructure Security Report (Arbor WISR).

According to Arbor's Active Threat Level Analysis System (ATLAS®), which monitors more than one-third of all Internet traffic, attack frequency is on the upswing and is estimated to reach eight million in 2018, a 16-percent increase over 2017. In addition, the proportion of multi-vector DDoS attacks (those combining volumetric, application layer, and state exhaustion vectors) has gone up significantly and is estimated to continue increasing 20 percent year-over-year, due primarily to the increased weaponization and automation of DDoS attacks tools.

Refer to "Appendix 1: Understanding the DDoS threat landscape, motivations and targets" for more detail.

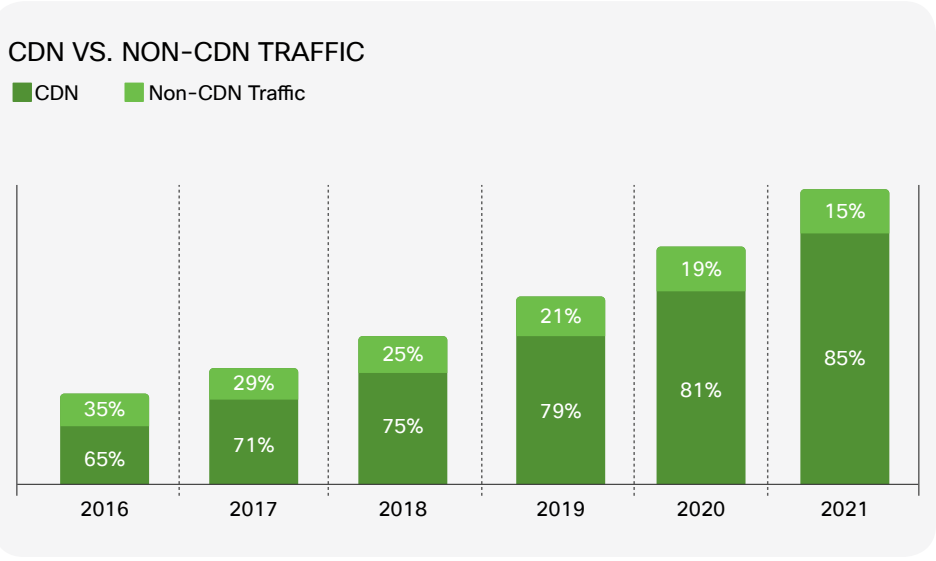
As peering architectures become more distributed and regional, edge routers must play a more active role as the first line of defense against these destructive volumetric attacks. The key is rapid detection using NetFlow-based telemetry combined with mitigation at the edge using Border Gateway Protocol Flow Specification (BGP FlowSpec) and hardware-based blocking in the edge routers to contain the attacks before they consume critical resources deeper in the network.

Keep calm, and binge on

There is a silver lining to the traffic growth trends: the increase in peering capacity is driven primarily by content, but content providers are not the source of DDoS attacks.

Inbound Internet traffic can be coarsely separated into two categories: content-peering traffic coming from content providers and Content Delivery Networks (CDN), and public-peering traffic originating from residential and business endpoints connected through traditional ISPs. Netflix, Google, Akamai, and Facebook are just a few examples of content sources that fall into the first category. These content sources dominate today, comprising 60 percent of all Internet traffic in 2018, and are projected to exceed 70 percent of all Internet traffic by 2021.

Figure 3. CDN versus non-CDN traffic



Content networks have not been a source of DDoS attack traffic in the past because they are special-purpose networks built for content distribution; their services are typically tightly controlled and do not provide transit for other networks. The primary sources of DDoS attack traffic today are compromised endpoints on residential and business networks that can be harnessed for reflection and amplification attacks. However, the threat of volumetric DDoS attacks is actually moderate when compared to the growth

Cisco + NETSCOUT Arbor: Proven solutions for infrastructure DDoS protection

- Cisco is a proven peering solution, with the Cisco® Network Convergence System (NCS) 5500 deployed at more than 40 service providers in a peering role today and thousands of Cisco IOS XR powered routers used for peering across the world. These high-density edge routers with scalable DDoS mitigation capabilities have proven to be effective against the landscape of volumetric attacks.
- Arbor's Threat Mitigation System (TMS®)—an intelligent detection and mitigation system (IDMS)—incorporates NetFlow-based analytics and detection, FlowSpec-based and DPI-based DDoS mitigation, and end-to-end workflows in a single integrated solution with Cisco routers.
- NetFlow-based detection plus BGP FlowSpec for DDoS mitigation is widely deployed, scales to multi-terabit network topologies, and utilizes industry-standard protocols.

in overall peering bandwidth. As indicated in Figure 3, public peering capacity (non-CDN) is growing at a much more moderate rate compared to content peering capacity (CDN) and the overall level of Internet transit traffic exposed to the worldwide Internet is on the decline.

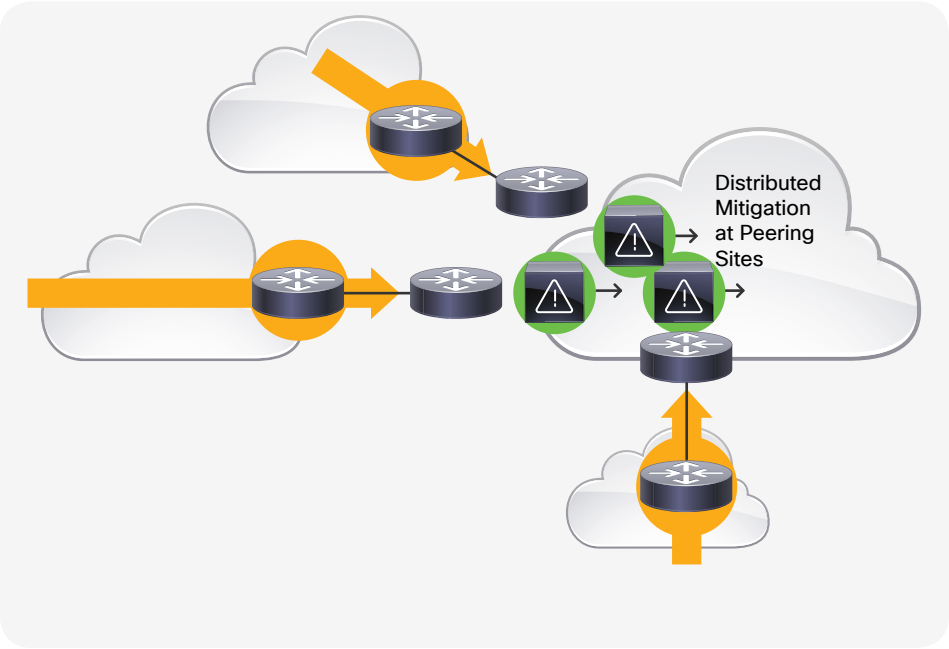
Continued traffic growth at more than 30 percent CAGR, comprised primarily of content provider traffic, requires a scalable and cost-effective Internet peering solution. The peering solution must also integrate with a DDoS mitigation solution employing best practices to mitigate all types of DDoS attacks. Unproven techniques such as applying advanced edge filtering to cover all ingress peering traffic incurs additional cost to known good traffic (i.e., reliably not a source of DDoS).

Decentralized peering needs decentralized DDoS protection

Attacks such as the Memcached reflection attack in early 2018 prove that a single attack can sustain multi-terabit traffic levels. This is more than enough bandwidth to overload the capacity of most network operators' internal infrastructure and jeopardize network availability.

This adds a new twist to classic DDoS attacks and requires a re-evaluation of DDoS mitigation architectures. While traditional techniques focus primarily on diverting large DDoS flows to active mitigation scrubber farms in order to protect users and services from attack, today's new decentralized peering architectures need an augmented approach that supports both decentralized and centralized DDoS mitigation where it makes sense.

Figure 4. Decentralized DDoS attack mitigation using the peering edge router infrastructure



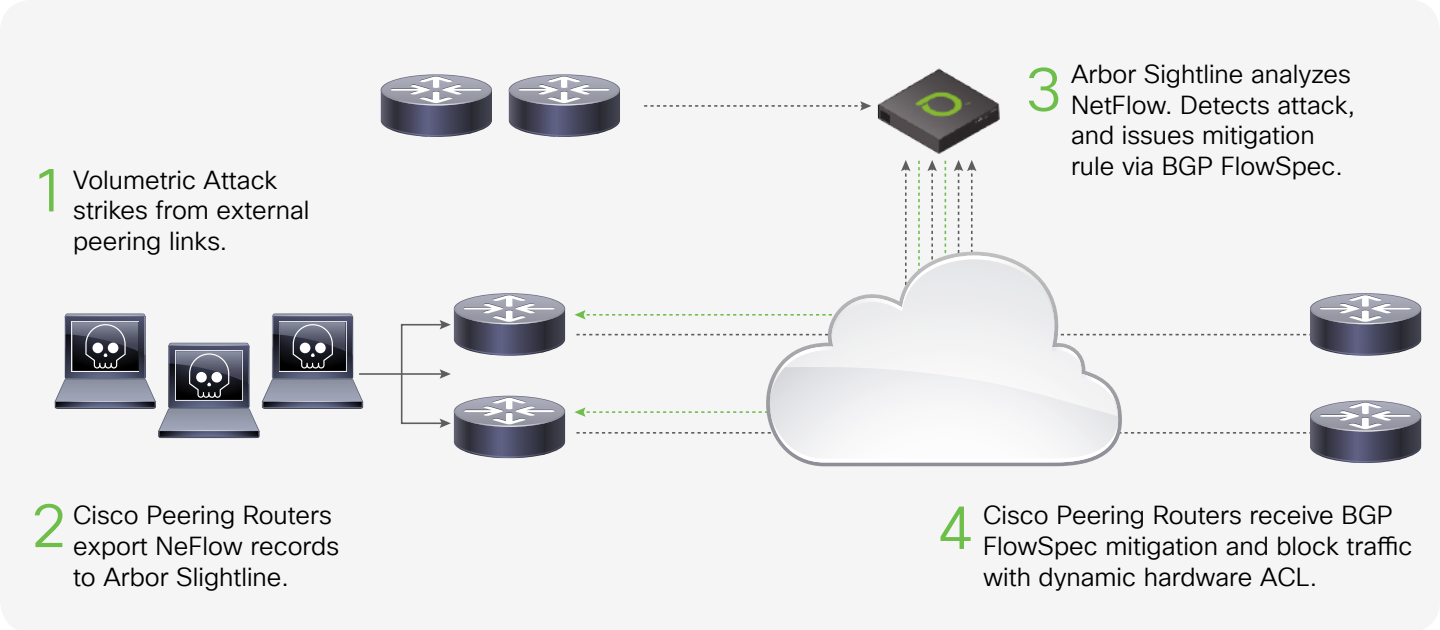
The motivations behind the attacks are complex (and discussed in more detail later in this paper), but in many cases large volumetric attacks are either targeting peering infrastructure directly (including internal core bandwidth) or serving as a smoke-screen to cover up more sophisticated and surgical attacks, such as penetration attempts on exposed services. In the latter case, the attackers are motivated to cause as much operational disruption and distraction as possible—including monitoring and rapidly mutating their attacks to evade static mitigation techniques.

A traditional DDoS mitigation strategy has been to implement strong (and computationally costly) application layer traffic controls to essentially "ride it out" and absorb significant DDoS attacks. While this is the right strategy for attacks targeting the state machine or application layer of services on the network, these attacks represent a smaller overall proportion of DDoS traffic. These application-targeted attacks are best handled with traditional active mitigation devices within the network. This comes at a cost, however: the capital infrastructure requirements for building and maintaining large-scale scrubbing facilities, as well as the core bandwidth required to backhaul large volumes of attack traffic to centralized locations in the backbone. Recent attack sizes demonstrate this approach won't scale as the cost of backhauling attack traffic for centralized scrubbing becomes prohibitive and transiting these traffic volumes adds risk to core network operations. Instead, carriers need a more effective means to quickly block this offending traffic at distributed peering locations.

Speed and automation: the keys to effective mitigation at the network edge

The rise in flash attacks has proven that attackers can rapidly change their attack patterns to make implementing an effective defense as difficult as possible. In many large service provider networks today, DDoS mitigation often involves a fully or partially manual process of programming traffic controls into peering routers (either through manual configuration, or manually-triggered BGP FlowSpec advertisements). In order to maintain a successful ongoing attack, an attacker simply has to mutate their attack faster than a defender can identify the new attack, derive appropriate Access Control Lists (ACL) or other traffic controls, and program these mitigations into their edge routers. This rapid adaptation of attack patterns also negates the efficacy of simple filters for payload signatures or patterns.

Figure 5. Distributed detection of a volumetric DDoS attack using Netflow telemetry, and dynamic announcement of BGP FlowSpec rules that block the offending traffic in the edge routers



While older attacks often included simple signatures that were consistent across all attack patterns (such as non-standard use of TCP flags), attackers have rapidly updated their tools to prevent providers from easily filtering known-bad attack patterns. Any solution to deal with motivated modern attackers must quickly detect new patterns (both in sources and protocols, as well as payload patterns), and mitigate before the attackers can modify their attack vectors. Attackers are relying on the premise that they can morph their attacks faster than providers can adapt their mitigations. The only way to block these attacks is to quickly detect and automate mitigations.

The best solution is to automate the process of both detecting new potential attacks through intelligent network analytics and to provide an automated workflow that instantiates mitigations directly into the edge routers using the BGP FlowSpec protocol's inherently dynamic capabilities.

The Cisco + NETSCOUT Arbor DDoS solution

As noted earlier, large volumetric attacks now present direct risks to service availability and network infrastructure because these attacks can easily saturate external peering links as well as core and aggregation links within the operator's network. It is therefore critical to quickly detect and control this malicious traffic at the edge in order to limit the impact on infrastructure.

In examining various DDoS mitigation architectures, it's important to distinguish protection of the network infrastructure itself from delivery of subscriber-facing DDoS mitigation services. In both scenarios, today's distributed peering architectures demand a multi-layer approach. Protection of the infrastructure itself should be distributed to peering locations using a combination of advanced flow analytics and hardware-based blocking in edge routers. For carriers that provide revenue-generating managed DDoS protection services to their downstream subscribers, a scalable, centralized mitigation solution is also required: it provides a second layer of defense against complex and application-layer attacks that manage to penetrate to the core network.

Modern DDoS attacks have proven to be highly dynamic in nature. Rapid and scalable flow analytics are therefore of paramount importance in quickly detecting potentially malicious traffic patterns, signaling appropriate mitigation strategies, and providing an automated framework for triggering mitigations, both in edge routers and intelligent DDoS mitigation systems such as Arbor's TMS.

Arbor's Sightline® Traffic Routing and Analytics platform is capable of scaling to these terabit network topologies: it consumes BGP, Simple Network Management Protocol (SNMP), and NetFlow IP Flow Information Export (IPFIX) telemetry to provide network-wide traffic visibility; it rapidly detects potential attacks; it can signal mitigations directly to edge routers using BGP FlowSpec as well as to Arbor's TMS to address the full range of volumetric and more complex state exhaustion and application layer DDoS attacks.

Edge routers must support NetFlow version 9 and IPFIX flow telemetry, as well as the ability to scale it to support the multi-hundred gigabit traffic volumes now common at large peering sites. Platforms such as Cisco's NCS 5500 and ASR 9000 Series Aggregation Services Routers are designed to deliver the scale required to support today's largest peering environments, including maintaining the high rates of NetFlow/IPFIX telemetry essential to advanced DDoS traffic analytics, detection, and mitigation. And filtering this large-scale traffic at the edge without blocking traffic from legitimate sources requires high access-control-entry scale. Cisco has implemented high-efficiency ACLs in Cisco IOS-XR Software to mitigate distributed attacks comprising thousands of source addresses.

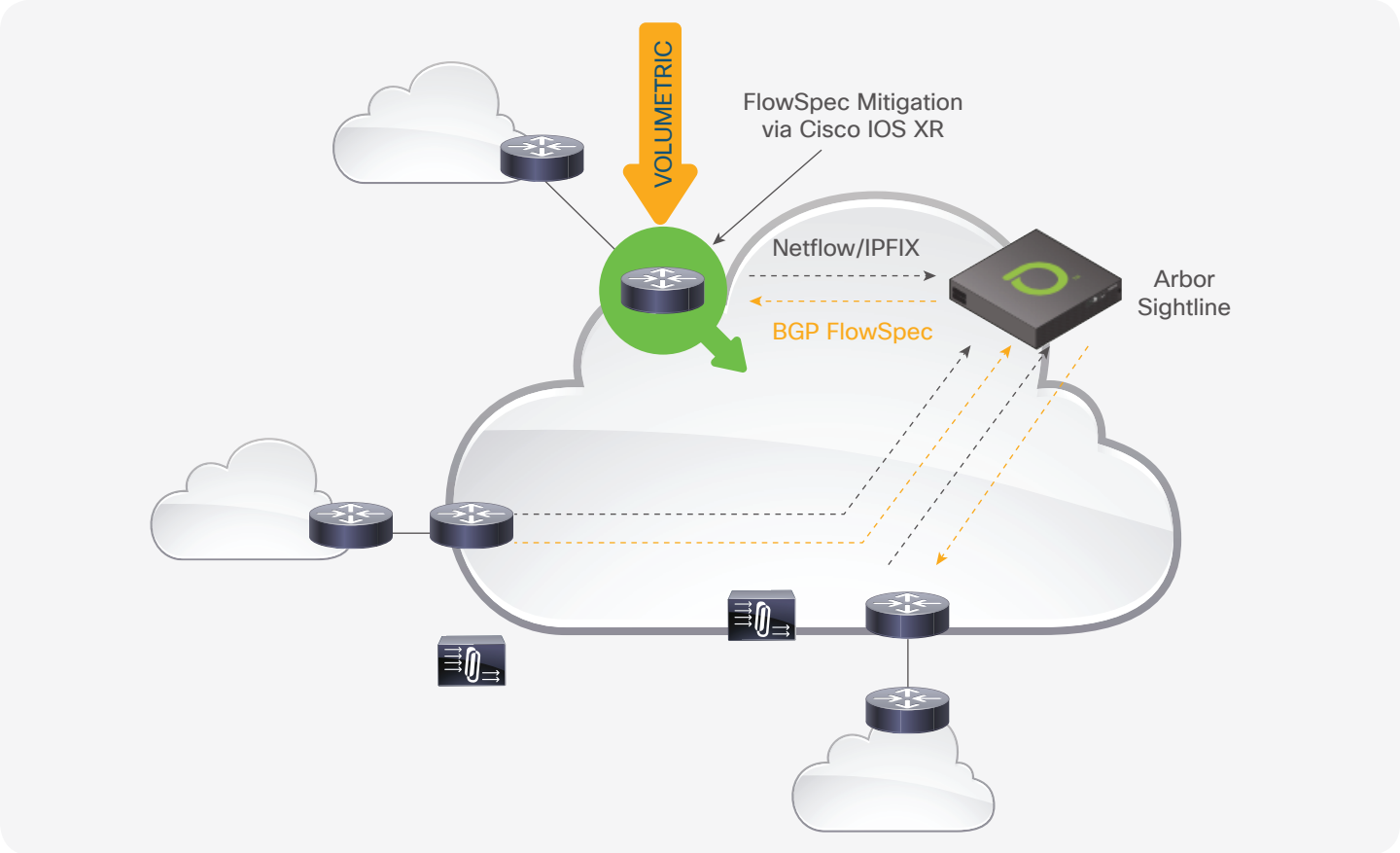
Flow-driven DDoS detection and mitigation

Reflection and amplification attacks rely on exposed and improperly secured Internet-connected devices and services, often arrayed as botnets, to deliver an asymmetric volume of attack traffic to a target. These attacks can deliver extremely high traffic volumes from a limited set of sources and protocols, and as attackers discover new unprotected services on public networks these types of volumetric attacks are likely to increase. In large-scale reflection and amplification attacks like the Memcached example, the ability to quickly detect anomalous large-scale traffic flows and signal appropriate mitigations is therefore essential. These attacks are high in volume yet often rely on unsophisticated reflection and amplification techniques to generate attack traffic, which can be matched using standard BGP FlowSpec criteria such as L3/L4 IP headers as well as packet size.

The value of BGP FlowSpec at the edge

In a volumetric DDoS attack, once malicious flows are identified by a traffic and routing analytics solution such as Arbor Sightline, FlowSpec-based mitigations can be dynamically signaled by Arbor Sightline to edge routers over BGP without the need for manual configuration of ACLs or routing policy across multiple edge routers. Flow-Specification Network Layer Reachability Information (NLRI) contained within the BGP protocol enables distributed detection platforms such as Arbor Sightline to directly announce traffic filtering and action criteria to peering edge routers. BGP FlowSpec can match on both IPv4 and IPv6 criteria and allows edge routers to dynamically build and install data-plane ACLs that selectively match and filter traffic in hardware.

Figure 6. Distributed detection of a volumetric attack and dynamic announcement of a FlowSpec mitigation to the edge routers



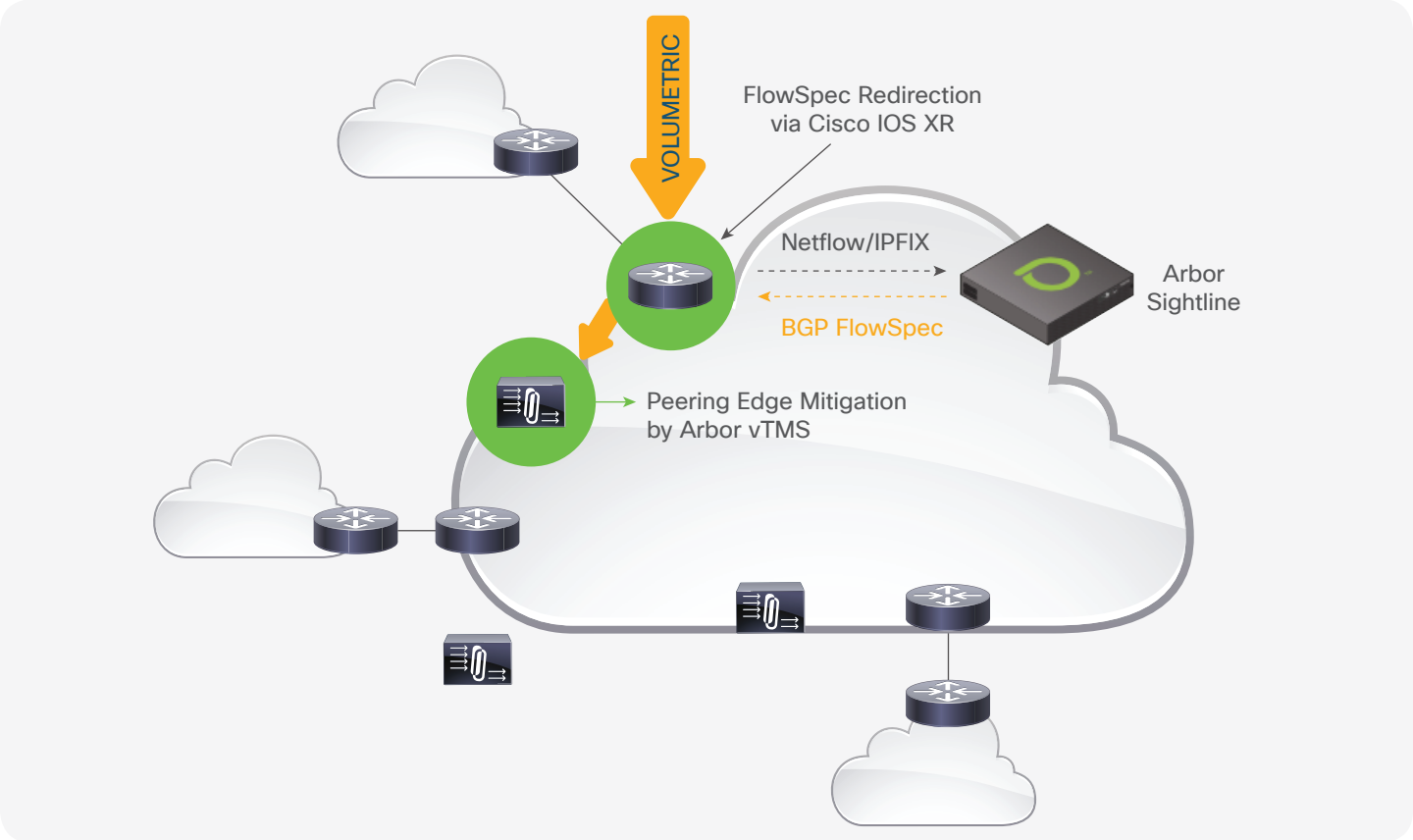
BGP FlowSpec offers multiple match conditions to identify traffic based on standard, five-tuple headers such as IP source and destination, IP protocol, Transmission Control Protocol (TCP)/User Datagram Protocol (UDP) source, and destination ports, as well as more granular criteria such as TCP headers, Differentiated Services Code Point (DSCP), and IP packet length. Together, these match criteria can identify each of the volumetric attack types currently observed in the wild.

Once matched, BGP FlowSpec can also support multiple actions, including simple packet drop, policing, DSCP marking, as well as the ability to redirect packets to a specified Virtual Route Forwarding (VRF) or IP next-hop. This provides a powerful framework for either controlling traffic directly within the edge routers or redirecting offending traffic to an external Integrated Database Management System (IDMS) for deep packet analysis and application layer traffic control. These IDMS can be located either adjacent to large peering facilities, or centrally in large-scale mitigation facilities.

Infrastructure protection using distributed mitigation

While NetFlow and FlowSpec alone can be used to mitigate many types of volumetric attacks, other more sophisticated attack types require more advanced application layer mitigation. Prime examples are botnet attacks like the Mirai IoT: tens of thousands (or more) of compromised end-user or IoT hosts orchestrated to attack through remote command and control. These discrete infected hosts can generate traffic not easily distinguishable from normal, business-critical DNS and other traffic. Detecting and mitigating these types of traffic patterns can require application layer intelligence to surgically isolate malicious traffic from normal traffic.

Figure 7. Announcing FlowSpec rules to an edge router to redirect an incoming volumetric attack to distributed intelligent DDoS mitigation resources



In a distributed mitigation architecture, IDMS devices such as Arbor's TMS are deployed adjacent to peering routers, either through dedicated appliances or virtual appliances deployed in a Virtualized Network Function (VNF) architecture. This architecture builds on the capabilities of the distributed NetFlow and FlowSpec mitigation solution, as malicious traffic can now be redirected to–and blocked by–intelligent mitigation resources local to the peering site. This solves the problem of consuming internal network capacity to transport suspected attack traffic back to centralized scrubbing centers and is a fundamentally more effective solution for infrastructure protection against advanced attacks.

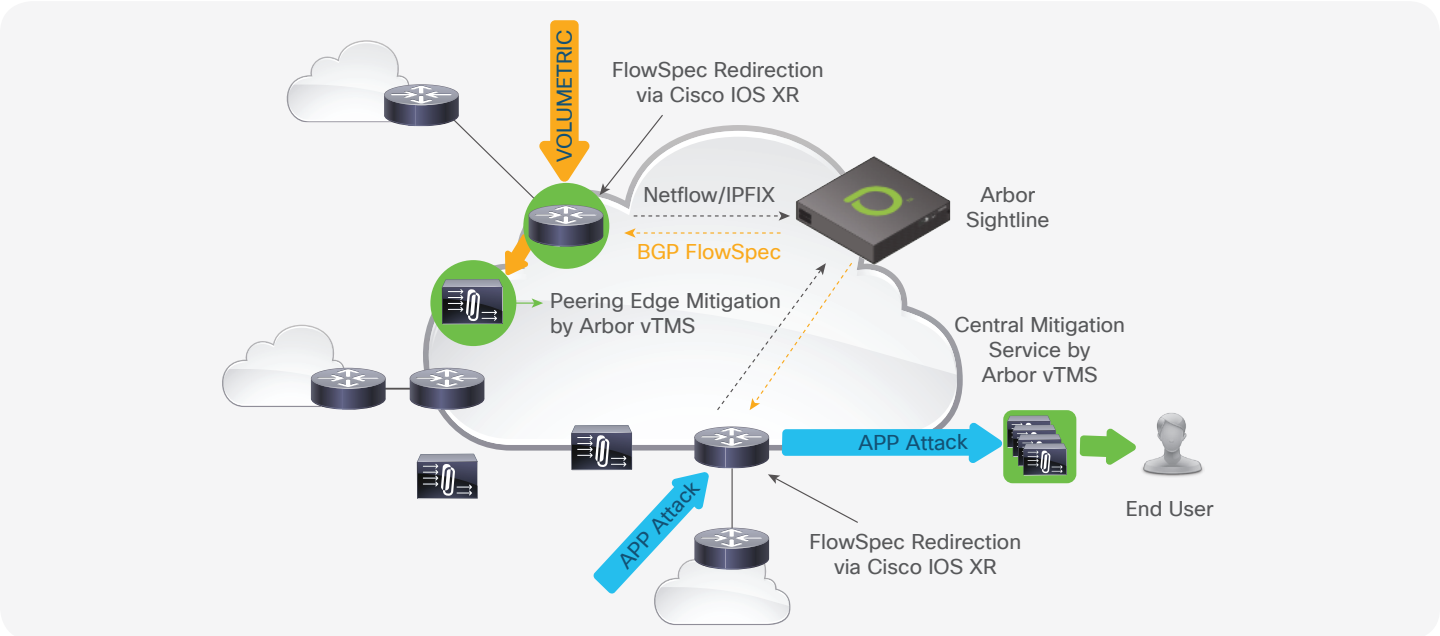
Arbor's Sightline solution can simultaneously announce real-time traffic control directly to edge routers using BGP FlowSpec and provide a second layer of protection against more complex attacks when paired with Arbor TMS.

Infrastructure protection using both distributed and centralized mitigation

While volumetric attacks pose a direct threat to network infrastructure resources, a separate class of attacks can target individual services exposed on end-user networks. These state exhaustion and application layer attacks stress the service delivery architecture of end-user services. State exhaustion attacks (such as SlowLoris) typically target the edge load balancers, firewalls, and stateful traffic inspection services of publicly-exposed services by stressing the scale of the TCP state machine of these devices. Application layer attacks target the application endpoints themselves, and typically exploit the exposed applications or APIs of public-facing services.

These attacks can easily overwhelm even large-scale enterprise services but present as much lower overall bandwidth attacks (typically less than 10 gigabits per second), so they are not normally considered a direct threat to the network provider's infrastructure. These types of attacks cannot be mitigated in stateless edge router infrastructure: they require intelligent L3-L7 mitigation front-ending individual services and applications, and centralized mitigation architected to protect the provider's infrastructure and all its downstream customers and by leveraging the economies of scale (multi-tenancy) possible with centralized scrubbing.

Figure 8. Announcing FlowSpec rules to an edge router to redirect an incoming volumetric attack to distributed intelligent DDoS mitigation resources; dynamically redirecting inbound application-layer attack traffic from edge routers to centralized scrubbing resources



Effective mitigation must be dynamic and flexible

The benefits of combining flow-based telemetry analysis with FlowSpec blocking at the network edge for volumetric attack mitigation have been thoroughly explored in this paper. Additionally, technologies like Remotely-Triggered Blackholing (RTBH) and Source-based Remotely-Triggered Blackholing (S/RTBH) can be very effective when used in a situationally appropriate manner.

Application layer and state exhaustion attacks can also be detected using security-focused flow analysis but as they are in almost all cases low-volume-type attacks, it is often necessary to use behavioral analysis or deep packet analysis to detect these kinds of attacks. This will usually require the use of an IDMS to provide visibility into the application traffic and detect the specific attack vector, so application layer traffic can be protected. The traffic is usually diverted to the IDMS using network diversion, which will then mitigate and block the attack on the device itself or offload the blocking to the edge devices.

As the attacks can be complex in nature and a determined attacker will rapidly change the attack vector when an earlier attack vector is mitigated, the IDMS uses a set of methods to analyze and block these kinds of attacks, including:

- Challenge/response mechanisms
- Behavioral analysis
- Flow and rate-based analysis
- Response analysis
- Various other approaches, depending on the application under attack and the attack vector being used

Nevertheless, recent examples of attacks by determined DDoS attackers show that even in cases where such patterns can be identified, they will simply change the attack parameters as soon as they see that the current parameters are no longer effective and launch a different type of attack.

- For example, when the Mirai IoT botnet was introduced in October 2016, Arbor published four regex patterns to help customers mitigate these attacks. Since then, most of the initial attack patterns have changed as the original flaws in the attack tool were either fixed by the attackers or the attackers moved on to new attacks. As a result, Arbor frequently updates the patterns and associated guidance to customers. Attackers usually parameterize these attack tools, so in many cases when they launch their attacks, these specific parameters can be detected and blocked using an IDMS.
- In addition, as part of Arbor's ATLAS Intelligence Feed (AIF), Arbor regularly publishes new HTTP regex patterns to match new HTTP-based botnet attacks. The number of these patterns has gradually declined as the attack tools have become more advanced, making it more and more difficult to distinguish attackers from valid users based on patterns in the attack payload.
- State exhaustion and application layer attack types are designed to be indistinguishable from legitimate traffic and operate at relatively small traffic volumes. These attacks rarely have signatures that can be matched in stateless payload filters. Active mitigations through an IDMS are typically required to identify these types of DDoS attacks.

Refer to “Appendix 2: DDoS attack types and mitigation methods” for more detail.

Arbor has already determined that actively monitored attacks can quickly change payload patterns, and attackers will continuously update their attacks to evade detection (often in timeframes as short as half an hour). This behavior has already rendered static pattern matching techniques, such as the ones proposed for ASIC-based payload matching obsolete and ineffective at mitigating these attacks.

Payload signatures are not effective mitigation

Arbor has done extensive analysis of existing DDoS attacks and has determined that using payload-based filtering against the vast majority of volumetric attacks does not bring additional value compared to simply matching header fields and packet characteristics. While competing vendors claim that in scenarios such as DNS amplification attacks (which currently represent the largest single type of volumetric DDoS attack) payload signatures can identify certain characteristics of offending DNS responses, this is likely of little additional use than traditional packet header rules for several reasons:

- The experience of large-scale ISPs shows that existing amplification attacks are mitigated using filters that identify ports and packet size
- Fields within DNS can be of varying lengths, which can render static payload signatures ineffective (which rely on matching patterns at specific offsets in the packet payload)
- Filtering based on payload signatures does not offer any additional value in blocking non-initial fragments of DNS amplification attacks comparing to traditional FlowSpec
- DNS amplification attacks can also involve IP fragmentation, where only the initial packet contains useful protocol headers. This poses a problem for any type of payload pattern matching, as IP fragments are highly unlikely to have consistent payload patterns at known offsets within the fragmented packet.

As a result, the cases where payload-based filtering might make more sense versus standard IP header rules alone do not justify investments into payload-based filtering: UDP random packet floods are not the most common attacks and all application layer attacks, including Session Initiation Protocol (SIP)/SNMP and others represent extremely low bandwidth. Moreover, both of them are not common attacks against infrastructure, residential, and SMB customers that larger service providers want to protect. Typical attacks against these markets are reflection and amplification attacks, and as analysis demonstrates, FlowSpec rules matching standard L3 and L4 headers are effective against these attacks.

Arbor has also found that when implemented in production networks, pattern matching tends to have a high rate of false positives for volumetric attacks. Not only is payload pattern matching likely be ineffective for mitigation, it also introduces operational issues such as unintended and unpredictable drops of legitimate customer traffic. As these patterns would match against packets from all sources, payload pattern packet drops would be impossible to debug in a large production network. Arbor found these issues also occurred when using regex pattern matching, which is significantly more flexible than the static payload patterns proposed for ASIC-based mitigation in competing routing platforms. Further, to effectively deploy pattern-based mitigation, it’s critical to understand and protect “known-good” traffic patterns. But in networks transporting unpredictable end-user traffic, as opposed to well-defined services, it’s nearly impossible to accurately and consistently profile “known-good” traffic.

Evolution of DDoS defenses

Modern DDoS attackers are constantly improving their attack tools and looking for new volumetric, application layer and stateful exhaustion attack techniques. They now have access to millions of vulnerable IoT devices, which allows them to launch complex attacks at scales never seen before.

Best practices to defend against these evolving types of attacks include:

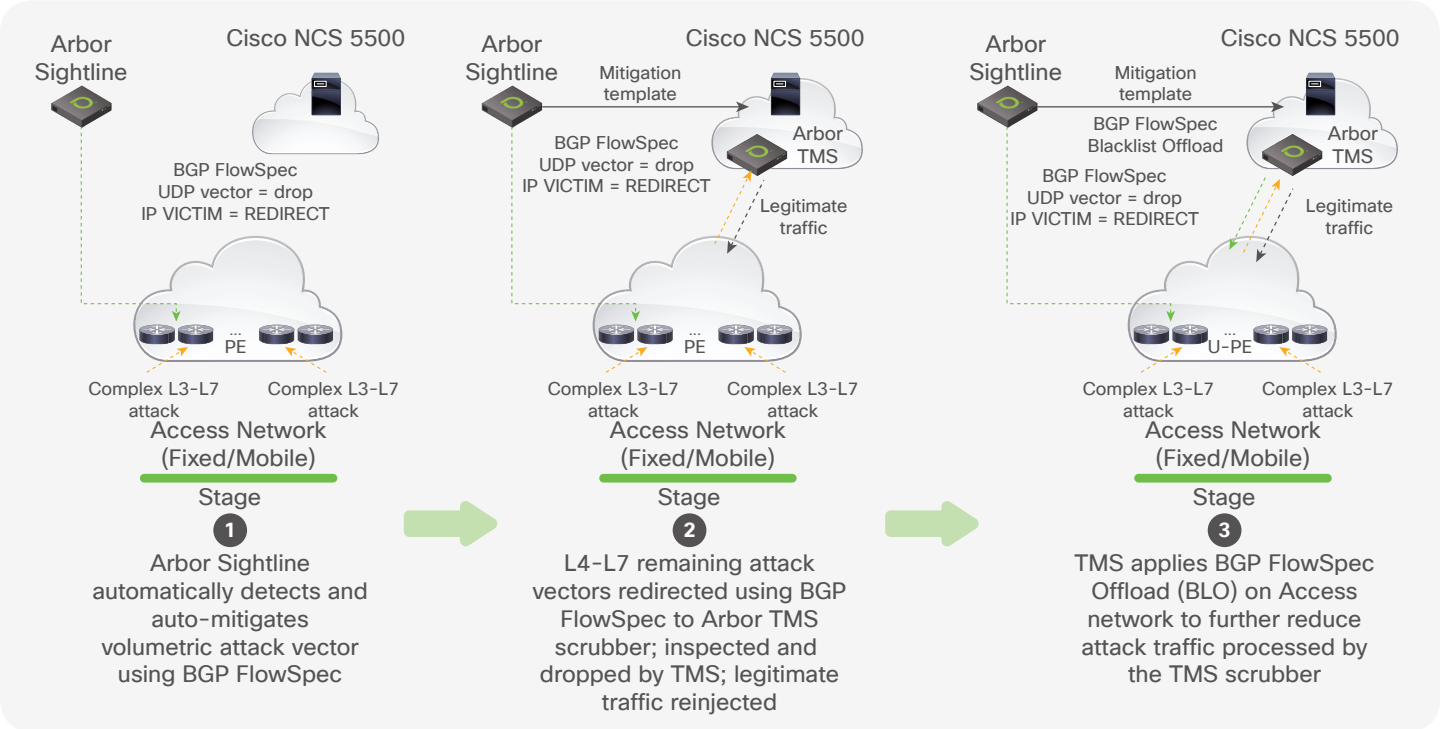
1. Using flow telemetry analysis supplemented with behavioral analysis to detect abnormalities and attacks. Focus on understanding what is normal: it will simplify identification of abnormalities.
2. Using BGP FlowSpec to activate network-based blocking at the edges of the routed network once a volumetric attack is detected.
3. Using an IDMS to detect abnormal behavior, and application layer and stateful exhaustion attacks that require advanced and active mitigation; and using this approach in conjunction with BGP FlowSpec Offload when and where appropriate.

If implemented successfully, these protections will force the attacker 'network' to behave like normal clients, rendering the DDoS attack ineffective and allowing for the use of application-level analysis to detect any abnormal traffic or usage patterns.

Implementing proven, scalable multi-layered protection

As described in the previous sections, proper defense against today's landscape of Internet-based security threats requires flexibility, scale, intelligent detection, and automated mitigation. NETSCOUT Arbor and Cisco provide a holistic solution to mitigate these threats with the deployment flexibility to handle centralized or distributed peering along with centralized or distributed traffic scrubbing when necessary.

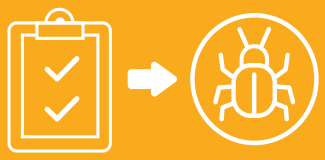
Figure 9. Three scenarios depicting Cisco + Arbor mitigation techniques for multi-vector attacks



Arbor Sightline is the foundation of intelligent attack detection, using interface statistics, NetFlow, and BGP data to identify both known and unknown attacks. Granular NetFlow data export is critical to identify attacks. Cisco, the inventor of NetFlow, provides class-leading sampled NetFlow capabilities on both its NCS 5500 and ASR 9000 Series routers. Once Arbor Sightline has detected an attack, automated attack mitigation is signaled to the Cisco peering edge to block, police, or redirect traffic to an IDMS. Using the simplified traffic engineering capabilities of Segment Routing, a router can steer traffic to a specific IDMS using either static or dynamic SR-TE Policies. Why risk using a less effective approach?

Appendix 1: Understanding the DDoS threat landscape, motivations, and targets

The DDoS attacks of today are, generally speaking, more complex and varied than the attacks seen a few years ago since DDoS defenses have become more effective. Attackers often still employ the simpler attack techniques of the past, but they are constantly adding new attack vectors, rapidly pivoting among combinations of flooding, application layer, and state exhaustion schemes to create multi-vector attacks, and increasingly focusing on application vulnerabilities instead of flooding attacks as new vulnerabilities are discovered and subsequently weaponized. Of equal concern, they are successfully automating for operation at scale. These dynamics force the defender to constantly adjust his or her defenses in order to withstand the attack.



1. Malware arms dealers are either individuals or organisations which research and develop attack tools that take advantage of security vulnerabilities. As part of their Q&A, often do live field testing.

2. The DDoS mercenaries offer DDoS services (Booters/Stresser) for hire to the attackers.

3. The attackers mostly use Booter/Stresser services to launch their attacks, though there are some exceptions.

DDoS mercenaries also offer their services for hire for launching attacks on specific customers, using a combination of “Booter” and ”Stresser” tools and by using focused attacks.

Motivations and targets behind volumetric attacks

Volumetric attacks are primarily used when the attacker is:

- Attempting to disconnect gaming users as a result of online gaming disputes. The attacks are either launched directly against the target user or against the gaming infrastructure the user is connected to. These attacks are launched using Booter and Stresser services, often with high volumes but short durations. These attackers are usually low-skilled and mostly do not monitor the results of their attacks.
- Making a point, often during political disputes where multiple low-skill attackers join forces in launching volumetric attacks, usually using Booter and Stresser services or simple attack tools. These attacks are often launched in combination with application layer attacks.
- Launching a determined attack against a specific target, usually critical service provider services or enterprise customers. A highly skilled attacker usually combines volumetric attacks with application layer attacks to hide the more focused application layer attacks, which do the real damage.

Examples of such attacks are reflection attacks using Memcached, Network Time Protocol (NTP), DNS, and Simple Service Discovery Protocol (SSDP), which flood the destination with large reply packets, filling up links and, in some cases, resulting in the collapse of the target network infrastructure.

These attacks are typically very high-bandwidth (up to 100 Gbps or more and occasionally exceeding Terabits per second) and are immediately obvious to both the target and upstream connectivity providers. This typically gets immediate attention from security and network operations teams. Because of this, determined attackers have learned to actively monitor the results of their attacks and often randomize their attack parameters as soon as defenders start to block or limit the current attack vector.

Motivations and targets behind application layer and state exhaustion attacks

Application layer and state exhaustion attacks are primarily focused on taking down specific services. For example, an attacker can launch an attack against web servers designed to constantly download large files or pictures, resulting in increased load on the back-end infrastructure. The attacker can also launch attacks against Transport Layer Security (TLS) endpoints, resulting in legitimate users being unable to connect to the services. These attacks are typically employed by determined attackers who monitor and adjust their attacks for maximum impact.

Application layer and state exhaustion attacks are usually low-volume compared to volumetric attacks since they have to conform to the protocol the application itself is using, which often involves protocol handshakes and protocol/application compliance. This means that these attacks will primarily be launched using discrete intelligent clients, usually IoT devices, and cannot be spoofed. The good news is defenders can identify those attacking hosts as a result and block them using intelligent DDoS mitigation.

But even when multi-vector attacks contain identifiable patterns, a determined attacker will monitor the results of his attack and modify it to thwart a skilled and determined defender. Because active attackers are known to continually modify payload patterns to avoid simplistic mitigations, maintaining an ongoing list of known attack patterns quickly becomes impractical due to scale issues and the rate at which this list must be updated. Further, since payload patterns bring high risk of causing collateral damage, maintaining a long-lived set of payload patterns may be unwise.

Appendix 2: DDoS attack types and mitigation methods

Volumetric attacks

Category	Frequency	Attack bandwidth (typical)	Can be mitigated using traditional FlowSpec?	Can be mitigated using IDMS?
DNS amplification	47, 9% of all volumetric attacks, according to Arbor ATLAS data	100 Gbps+	**Yes,** based on UDP ports and packet length. Exceptions are responses based on EDNS0, e.g. DNSSEC. An additional FlowSpec filter is required to block UDP fragments to the victim.	DNS amplification
NTP, SSDP, Memcached, Chargen, C-LDAP, SNMP, Portmap, MSSQL, and other amplifications	52.1% of all volumetric attacks, according to Arbor ATLAS data	100 Gbps+	**Yes,** based on ports and packet size. An additional FlowSpec filter is required to block UDP fragments to the victim.	NTP, SSDP, Memcached, Chargen, C-LDAP, SNMP, Portmap, MSSQL, and other amplifications

State exhaustion attacks

Category	Frequency	Attack bandwidth (typical)	Can be mitigated using traditional FlowSpec?	Can be mitigated using IDMS?
TCP SYN, TCP RST, TCP ACK	Vast majority of session exhaustion attacks	Less than 100Gbps	**No**	**Yes**, using a challenge/response-based approach
Idle TCP, UDP connections	Less typical attacks	Less than 10Gbps	**No**, the attack uses valid TCP and UDP sockets	**Yes**, using behavioural session analysis and dropping inactive sessions
UDP random packet flood	Less typical attacks	Less than 100Gbps	**No**, if the attack is destined to a valid active UDP socket	**Yes**, using rate-based analysis, session analysis, and challenge-response mechanisms
ICMP, GRE, and other random IP protocols	Less typical attacks	Less than 100Gbps	**Yes**, if the victim is not expecting those protocols	**Yes**

Although there are a few corner cases where these attacks can be generated using poorly crafted tools that utilize static tcp.window_size or ip.ttl, the majority of attacks employ randomized TCP fields which are indistinguishable from legitimate traffic. Therefore, patterns based on L3-L4 headers do not work. Also, Arbor's Security Engineering & Research Team (ASERT) has observed some attacks utilizing TCP SYN flood with payload, mostly launched by XOR-based tools. Those attacks are not popular because they are easily detected and mitigated without payload analysis. For example, a FlowSpec policy that drops TCP SYN traffic with a packet length of more than 40 bytes (plus necessary encapsulations) is more than enough to successfully block in this case.

Application layer attacks

Category	Frequency	Attack bandwidth (typical)	Can be mitigated using traditional FlowSpec?	Can be mitigated using IDMS?
TLS (L5)	Less typical attacks	Around 1 Gbps, usually less	**No,** behavioral analysis of TLS session setup is needed	**Yes**, by analyzing the TLS session setup, checking TLS extensions, idle sessions, and so on
HTTP	73% of Arbor's WISR enterprise respondents observed those attacks	Around 1 Gbps, usually less	**No**	**Yes**, by analyzing the rate of requests, performing challenge/response, and applying known patterns of attacks
HTTPS	68% of Arbor's WISR enterprise respondents observed those attacks	Around 1 Gbps, usually less	**No**	**Yes**, by performing decryption, provided certificates/keys are uploaded
DNS	69% of Arbor's WISR enterprise respondents observed those attacks	Around 1 Gbps, usually less	**No**	**Yes**, by analyzing the rate of requests, performing challenge/response, downloading protected DNS zone (zone transfer), and applying known patters of attacks
Others (SIP, SNMP, etc.)	68% of Arbor's WISR enterprise respondents observed those attacks	Around 1 Gbps, usually less	**No**	**Yes**, by analyzing requests, performing RFC compliance checks, and rate analysis

MedusaHTTP is an example of an HTTP botnet that uses a list of randomly chosen user-agent HTTP headers and allows using a built-in browser that is indistinguishable from real user traffic. https://asert.arbornetworks.com/medusahttp-ddos-slithers-back-spotlight/

Many DNS attacks against DNS resolver or authoritative servers are based on label prepending (so-called "water torture attacks") where attackers generate a lot of queries using randomized DNS subdomains. A good example is the famous DYN attack. It is not possible to define a pattern around illegitimate queries due to the randomness of queries. It is also not possible to define a whitelist pattern to block everything else for DNS resolvers. For DNS authoritative servers, it is possible to whitelist the protected zone, however the scale (the number of domains in a zone cut) might be an issue.