SANDVINE

# Shine a light on the darkening internet: How to thrive despite Encryption

## An exploration of the impact of encryption on ANI use cases

### OVERVIEW

**Today, anywhere from 50-90% of traffic on networks are encrypted (depending on the type of network and the geographic location) based on industry reports and Sandvine's Global Internet Phenomena data. The public nature of security breaches has pushed application vendors and content providers to increasingly implement encryption to protect consumer privacy - and this is a very good thing for consumers. Why? Because encryption protects your data - credit card numbers, passwords, personal information - from anyone who gains access to your traffic.**

A majority of traffic on most networks is now encrypted. An example from a customer network below shows 57% of traffic is encrypted, with the majority still using TLS 1.2, but with TLS 1.3 beginning to grow, mainly powered by Facebook.

Figure 1

ALL ENCRYPTION STATUS(S)

| Encryption Status | Distribution |
|---|---|
| Encrypted | 57.05% |
| Unencrypted | 42.95% |

ALL ENCRYPTION TYPE(S)

| Encryption Status | Distribution |
|---|---|
| Unencrypted | 42.95% |
| TLS 1.2 | 40.08% |
| App_Encryption | 9.03% |
| SSL | 3.16% |
| QUIC | 2.85% |
| TLS 1.0 | 0.91% |
| TLS 1.3 | 0.26% |
| TLS 1.1 | 0.04% |

ENCRYPTION TYPE: TLS 1.3

| Protocol | Distribution |
|---|---|
| TLS 1.3 | 40.55% |
| Facebook | 31.92% |
| Facebook Video | 25.08% |
| Facebook Messenger | 1.27% |
| Instagram | 0.88% |
| Instagram Video | 0.17% |
| Tumblr | 0.10% |
| Slacker Radio | 0.01% |
| Origin Games | 0.01% |
| WhatsApp Media | 0.00% |
| WhatsApp | 0.00% |
| Xbox Live | 0.00% |
| 9GAG | 0.00% |
| Symantec Live Update | 0.00% |
| Google | 0.00% |
| Skype PC to PC | 0.00% |
| Skype Generic | 0.00% |

All units measured in bytes

A previous Sandvine whitepaper on the trends that are driving encryption discussed sites being rewarded in Google search rankings if they support HTTPS, while application vendors are also rewarded through the continued use of their applications by consumers. This trend is a challenge to any Network Intelligence (NI) solution that is deployed to classify applications and network traffic. This whitepaper explores the impact of encryption on common NI use cases.

# Application Identification in the Era of Encryption and Virtualization

## WHAT IS NETWORK INTELLIGENCE IN THE ERA OF THE ENCRYPTED INTERNET?

Wikipedia defines Network Intelligence as:

**"Network Intelligence (NI) is a technology that builds on the concepts and capabilities of Deep Packet Inspection (DPI), Packet Capture and Business Intelligence (BI) It examines, in real time, IP data packets that cross communications networks by identifying the protocols used and extracting packet content and metadata for rapid analysis of data relationships and communications patterns. Also, sometimes referred to as Network Acceleration or piracy. NI is used as a middleware to capture and feed information to network operator applications for bandwidth management, traffic shaping, policy management, charging and billing (including usage-based and content billing), service assurance, revenue assurance, market research mega panel analytics, lawful interception and cyber security. It is currently being incorporated into a wide range of applications by vendors who provide technology solutions to Communications Service Providers (CSPs), governments and large enterprises. NI extends network controls, business capabilities, security functions and data mining for new products and services needed since the emergence of Web 2.0 and wireless 3G and 4G technologies."**

(Wikipedia citation of https://en.wikipedia.org/wiki/Network_intelligence).

On the encrypted Internet, it is no longer enough to claim "deep packet inspection" capabilities since looking at an encrypted packet in isolation is likely to reveal very little about the traffic other than standard Layer 3 information - required by the network to get traffic from point A to point B - which can be done by almost any network element. Not only is the "metadata" - the main source of intelligence used to identify applications by many DPI or NI offerings - which is now encrypted, but basic header information is also encrypted with increased frequency with TLS.

| Characteristic | Traditional DPI | Encrypted Internet |
|---|---|---|
| Header Information Visible | Layer 7 information and metadata available to classify applications | Header information encrypted, not revealing application directly |
| Full Hostname/uRL available | "Site" signatures and full URL path available for classification | Hostname or SNI may or may not be available for site-level classification |
| Netflow/IPF ix/logging visibility | Traditional packet logging can reveal application or content accessed | Packet logging identifies traffic as TLS, HTTPS, VPN etc |
| Behaviour analysis required | Only for certain applications (encrypted messaging, P2P) | For most applications (except CDN-identifiable content) |

While encryption is becoming prevalent, Network Functions Virtualization (NFV) is rapidly becoming the architecture of choice for network operators of all types. NFV introduces further application identification challenges, as NFV/SDN networks can be highly distributed, dynamically change their behavior based on network conditions and are deployed on COTS hardware that rarely has hardware assist/acceleration capabilities for packet processing offload. Crucially, some DPI and NI systems rely on network processor hardwares for flow segmentation or even for their packet inspection; true NFV implementations require virtual network functions (VNFs) to run on any hardware throughout the network.

Both of these trends fundamentally change how NI engines must operate to successfully deliver their core function - Application Identification. Let's look at some of the challenges in this new network landscape.

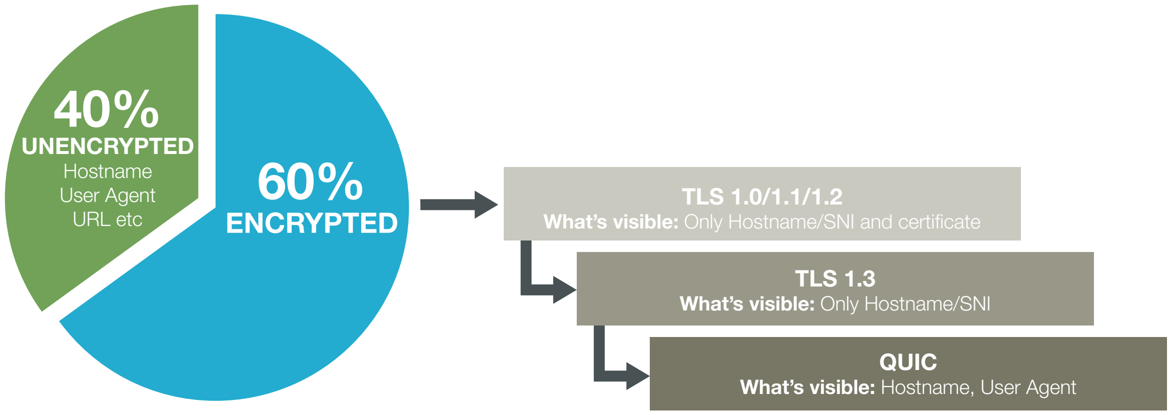
# Application Identification challenges

## ENCRYPTION

As already discussed, the secret sauce for any NI engine is in identifying the core application behind a traffic flow regardless of the protocols used to transport traffic. An application identification engine needs to be able to differentiate between application traffic with a high degree of granularity using multiple techniques, with heavy emphasis on the use of heuristics to identify flow behavior as well as other "clues" that may be present in the other active flows in the subscriber's traffic stream. For example, Facebook video has different requirements than Facebook browsing, and this level of granularity may be critical for certain use cases.

The graphic below illustrates how traffic is changing from unencrypted http to TLS 1.3 with encrypted DNS.

Figure 2


40%
UNENCRYPTED
Hostname
User Agent
URL etc
60%
ENCRYPTED
TLS 1.0/1.1/1.2
What's visible: Only Hostname/SNI and certificate
TLS 1.3
What's visible: Only Hostname/SNI
QUIC
What's visible: Hostname, User Agent

As shown, as traffic increasingly becomes encrypted the "metadata" that in the past made identifying HTTP simple is disappearing. To cope with this, NI engines need to leverage sophisticated flow analysis and hueristics that do not rely on header or payload information, so encryption does not affect classification. Machine Learning techniques to refine and maintain application signatures have become mandatory in the age of encryption.

## ASYMMETRY

The distributed nature of virtualized networks introduces a high likelihood that subscriber traffic flows may take different paths through the network. This is known known as asymmetry. Asymmetry introduces significant challenges for application identification since flow classification is based on "seeing" all parts of a connection to perform the heuristics on the traffic. Many current systems solve this challenge by using traffic re-routing with their hardware solutions to ensure that all traffic for a subscriber goes to a single CPU for flow correlation. However in a SDN/NFV environment that requirement introduces unnecessary complexity into the network that complicates operations and troubleshooting. Virtual solutions must adapt and use other methods to communicate necessary flow state information between systems without forcing re-architecting of the network.

Accurately identifying applications requires the correlation of control plane and data plane connections

## FLOW CORRELATION

Many applications utilize the concept of parent/child connections. Accurately identifying these applications requires the correlation of control plane and data plane connections. FTP and SIP are two commonly used examples of these type of applications. While today the control plane is often unencrypted, the data plane is encrypted. It is therefore important to correlate both channels as part of the same conversation for accuracy, especially in a security and/or charging use case.

## LOCALIZATION

We live in a global world that has people traveling on a daily basis, visiting and moving to new countries. Every country's subscribers have favorite applications that may differ from the favourite applications of subscribers in other countries. Any application identification solution must support localized application signatures, not only to satisfy the local market but also to support popular applications that are used around the world and are being used by travelers.

An application that is popular in one specific region of the world may get almost no native population use in another region. However, a local expatriate population that uses it may be an important niche market for an operator. The best application identification solution for a North American operator may not be the best solution for a Southeast Asian operator if it cannot identify Chinese applications well. It is common for a solution that has not been exposed to a particular region to have a high unknown ratio when first tested by an operator and a good test for vendors is how quickly that they can get the unknown percentage down to acceptable levels.

There are several methods to quickly reduce the percentage of unknown traffic

There are several methods to quickly reduce the percentage of unknown traffic, which range from the vendor implementing new signatures based on network captures, to enabling the operator to create their own signatures based on regional applications. Vendors should also be aware that many applications especially voice, video and messaging applications, often operate differently from region to region; sometimes based on the Content Delivery Networks (CDNs), sometimes based on techniques for evading detection and blocking, based on regional regulations.

## SPEED OF CHANGE

Applications evolve rapidly in the Internet economy. Some applications become more efficient, some improve to deliver a better quality of experience and some change as new technologies become available. As a result NI vendors need to be able to rapidly update their signatures based on the latest releases of the application, since a change by a major application (Netflix, YouTube, Skype etc) could cause huge inaccuracies in analytics or if zero-rating is being applied to specific applications or classes of applications.

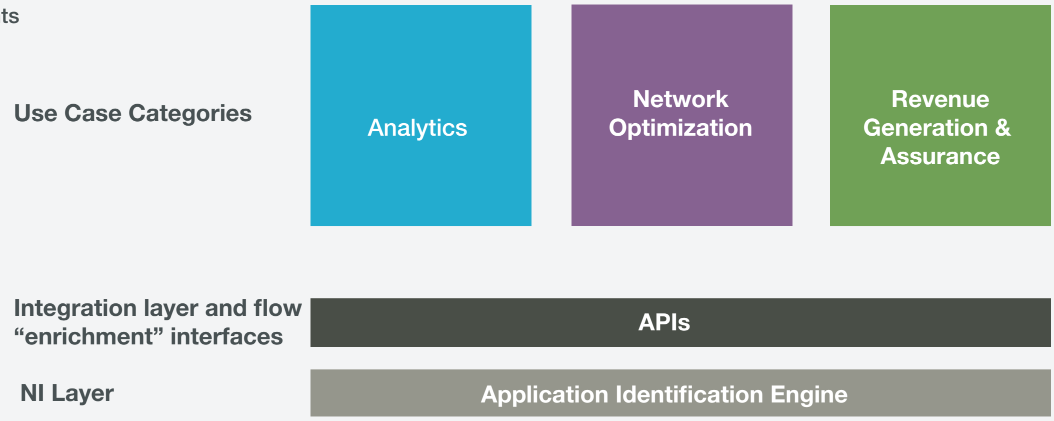
For some applications the turn around on signatures should be hours or days. For most others it should be less than a week. If the normal signature release cycle for a vendor is much longer than a single week, your network is at risk of missing important updates in key applications or that hot new application (like Pokemon Go!) with which your subscribers are suddenly obsessed.

## BUILDING NI-ENABLED USE CASES AND COPING WITH ENCRYPTION

Once the fundamental application identification challenges have been solved, use cases become the proof points in how NI-enabled solutions can adapt to the encrypted Internet. NI was initially deployed in many networks specifically to deal with traffic management for peer-to-peer file sharing applications, but has since evolved to support many more use cases. As shown in the diagram below there are several layers that need to be added to the NI engine to implement the most common use cases, since simply identifying the application on the network does not provide as much value as correlated or contextual data that includes other attributes.

Figure 3

NI integration points and use cases


Use Case Categories
Analytics
Network Optimization
Revenue Generation & Assurance
Integration layer and flow “enrichment” interfaces
APIs
NI Layer
Application Identification Engine

Most NI solutions add an integration layer onto their engine to add this “enrichment” to the application-centric data delivered by the NI engine. The most common example of this is subscriber integration, where the traffic from a subscriber is not only identified by IP Address, but also by a unique identifier (email, phone number, MAC Address, account name, etc.) for many use cases. Other attributes that increase the actionability of NI data may include location, topology, device type (for mobile networks), service plan, content classification, route information and QoE scoring or metrics for that subscriber. By combining all of this information, the common use cases for Analytics, Network Optimization and Revenue Generation & Assurance are actionable in real-time, rather than just being appropriate for historical use cases or requiring extensive Big Data integrations to be useful.

| Use Case | Impact |
|---|---|
| Analytics | Low to Medium |
| Network Optimization | Low |
| Revenue Generation & Assurance | Low to High |

Once this integration layer is in place, it is possible to implement both “Save Money” and “Make Money” use cases that utilize NI solutions. The key use case categories to explore are Analytics, Network Optimization and Revenue Generation & Assurance. Each one of the main use case categories will be affected differently by encryption, so let’s take explore each class of use case.

# Impact Analysis: Analytics Use Cases **Low to Medium**

**A foundational use case for NI solutions is to provide detailed analytics and reporting on network traffic. Almost all NI solutions deployed in operator networks today contribute analytics data to Big Data projects. Sometimes analytics are delivered natively using the NI vendor's solutions, but are also increasingly integrated with larger Big Data projects. NI is providing a wealth of network and subscriber intelligence, including network quality information, which has become critical to operators. Each use case will be affected differently, so we will explore each one individually.**

| Analytics Use Case | Impact | Notes |
|---|---|---|
| Capacity Planning | Low | • Network capacity planning driven by high profile applications and subscriber totals<br>• Very little loss of visibility due to encryption |
| User Behavior and Demographic Analysis | Medium | • Service planning relies on visibility into applications and content accessed by subscribers<br>• Some applications will be harder to track, requiring better NI solutions |
| Performance and Operational Monitoring | Low/ Medium | • Service QoE levels can be measured without MOS metrics gathered from signaling layer<br>• More emphasis on QoE KPIs regardless of application<br>• Distinguishing between Facebook Video and Facebook browsing is more challenging |
| Regulatory Analytics | Medium | • Regulatory analyitics will lose metadata visibility for web traffic and applications<br>• Loss of granularity for applications that are now encrypted (no URLs etc) |

## CAPACITY PLANNING

Capacity Planning analytics is used to forecast how network capacity needs to be expanded. This use case relies on visibility into network consumption and system load in specific locations and is increasingly combined with visibility into not only how many subscribers are using the network at that location, but the applications that they are using and what they need from the network to deliver a high QoE. This use case will not be affected much by encryption, as high bandwidth applications (like streaming video) will still be identifiable using multiple methods of identification with a good NI engine.

## USER BEHAVIOR AND DEMOGRAPHIC ANALYSIS

User Behavior and Demographic Analytics is used by marketing teams to create new services based on how their subscribers are consuming their network bandwidth. Marketing teams need a better understanding of the applications and content that subscriber's view as integral to their network experience, in order to create attractive packages. There will be some impact to this use case by encryption as some applications, especially Peer-to-Peer style applications (some VOIP and messaging apps use this technique for communication and are the biggest challenge) require heuristics to reliably identify.

## PERFORMANCE AND OPERATIONAL MONITORING

Telecom regulators are increasingly focusing on the QoE delivered by networks as a "truth-in-advertising" metric. Both the US FCC and the EU BEREC have stipulations in their Network Neutrality rulings that require regulators to monitor the performance that their networks deliver, primarily for throughput, latency and packet loss. The FCC has even proposed a Broadband Labeling Initiative that encourages operators to specify the expected performance for their networks. This use case will not be impacted by encryption, as the KPIs have nothing to do with encryption, but simply measuring network performance (but measuring how the network is delivering QoE to specific applications certainly does).

### REGULATORY ANALYTICS

Many governments have enacted legislation that require network operators to log certain subscriber activities, traditionally for law enforcement or security purposes. This use case will be heavily affected by encryption, as the use case-relevant "metadata" will become encrypted and unavailable for analysis. This will require a shift in tactics by law enforcement agencies, since the encryption of the metadata is considered a good thing for both consumer privacy and in making identity theft and hacking harder for cyber criminals.

## Impact Analysis: Network Optimization Use Cases **Low to Medium**

**The initial use case for NI-enabled solutions in many operator networks was traffic management, primarily to manage Peer-to-Peer file sharing. P2P sharing wreaked havoc on early generation networks as it would use all of the bandwidth that was available, allowing a few users to create a bad network experience for everyone else. P2P file sharing has decreased in many parts of the world as streaming audio and video have become readily available, but the core problem of bandwidth contention is still an important use case. Network Neutrality rulings around the world have supported the use of NI for network optimization purposes. However, these rulings include specific guidelines on how traffic management should be implemented on the network and encryption has a low-level impact on the common network optimization use cases.**

| TM Use Case | Impact | Notes |
|---|---|---|
| Congestion Management | Low | • No loss of location visibility<br>• High bandwidth applications still identifiable |
| Video Streaming Management | Medium | • Video Streaming applications leverage multiple encrypted protocols<br>• Different content providers may require different bandwidth rates for different resolutions due to content, introducing more optimization opportunities |
| Fair Usage | Low | • Service Tiers integration unaffected<br>• Often application-agnostic traffic management, so no impact |
| Peering and Transit Link Management | Low | • No loss of BGP visibility<br>• High bandwidth applications still identifiable |

### CONGESTION MANAGEMENT

Expecting congestion on their network, every network operator implements some form of traffic management in each network node. Due to the combination of application, subscriber and location awareness, NI solutions are often used to enable congestion management that can limit specific applications during times of congestion for specific classes of users. Although Network Neutrality has limited this traditional use case, there are still many operators whose terms of service clearly state that some applications may be managed during times of congestion. Furthermore, some operators offer unlimited plans that specifically authorize the operator to rate-limit applications that are non-interactive like software updates or downloads during times of congestion for heavy users. The applications traditionally targeted for this use case are already encrypted, so there is very little additional impact in implementing location-aware congestion management.

## VIDEO STREAMING MANAGEMENT

The traditional congestion management use case has evolved into a smarter implementation focused on ensuring that Video QoE is maintained while also reducing the overall traffic volume on the network. To deliver on this use case, a solution needs to be both subscriber and application aware and utilize advanced traffic management techniques to ensure a high QoE for streaming video while recognizing the specific streaming application. It must also be able to measure the impact of congestion management that matches the KPIs required by Network Neutrality, namely throughput, latency and loss. Encryption will have a medium impact on this use case, as it requires not only that individual applications still be identified, but also that each streaming application is managed to the correct bandwidth to deliver the promised resolution and to ensure a high QoE.

## FAIR USAGE

Many operators are choosing to implement basic fair usage in order to comply with Network Neutrality regulations. Fair usage is designed to ensure that every subscriber has equal access to bandwidth during times of congestion. In the majority of cases, encryption will have no impact on this use case, however there are service offering options that allow individual subscribers to prioritize applications when Fair Usage is being enforced. Those deployments would have traditional application identification requirements for the applications or application classes that subscribers are allowed to select for prioritization.

## PEERING AND TRANSIT LINK MANAGEMENT

Peering points are increasingly becoming a critical link in delivering a high QoE to subscribers. Peering can also be very expensive for operators whose subscribers access a majority of their content from outside of their network (or country, as is the case in many developing countries). Many operators use analytics to determine what applications or content drives usage on peering links and then manage each link to ensure that the cost of the link is minimized while still delivering a high QoE for real-time application. NI solutions are increasingly being used for this traffic management use case and encryption will have a small impact on this use case, as it still requires good application identification for high bandwidth applications. However most of these applications are already encrypted, so there will be minimal change for the use case.

## Impact Analysis: Revenue Generation & Assurance Use Cases

### Low to High

**Revenue Generation and Assurance is the use case category that will be most affected by encryption; the risk level for false positive or false negative application detection needs to be minimized when policies are applied to traffic that result in filtering or charging actions. Although the term Policy and Charging Enforcement Function (PCEF) is most commonly associated with mobile deployments, the same solution is applied to fixed, satellite and even WiFi networks to support differentiated services offerings.**

| Revenue Generation & Assurance Use Case | Impact | Notes |
|---|---|---|
| Advanced Data Services | Low | • No application visibility required |
| Application-Based Services | Low/High | • Some applications easily identifiable despite encryption<br>• Achieving 100% certainty on some applications not possible |
| Zero Rating | Medium/High | • Some applications easily identifiable despite encryption<br>• Achieving 100% certainty on some applications not possible |
| Parental Controls | Medium | • Domain-level visibility can still be maintained easily<br>• Blocking specific URLs is not possible |
| Data Fraud Management | Medium/High | • Some applications easily identifiable despite encryption<br>• Encryption opens new fraud channels for zero rating |

Working directly with application vendors can help ensure accurate identification

### ADVANCED DATA SERVICES (VOLUME-BASED CHARGING)

Volume-based charging or usage caps have become a standard service offering for fixed and mobile broadband operators. Many operators have implemented volume-based charging using DPI solutions, in some cases because they are also offering some of the application-based plans discussed below while in other cases they intend to evolve their offerings in the future to include these service plans as an option. This use case is not affected by encryption at all, since no application identification is needed for basic byte counting service plans.

### APPLICATION-BASED SERVICES

Application-based service plans enable users to consume more (or occasionally unlimited amounts) of a specific application or class of application for an additional monthly fee. An example of a service like this would be to give the consumer unlimited audio or video streaming for an additional $10/month - whether that uplift fee was an à la carte offering or bundled into a higher rate for the service package. NI solutions are used for this type of offering because of the ease in selecting applications or classes of applications where the signatures are updated on a regular basis.

Encryption will have a low to high impact on this use case. While some applications will still be easily identifiable despite the addition of encryption, others will become much harder to identify. For those applications where identification with encryption becomes less certain, operators must exercise caution in including them in these type of bundles; any time charging is involved, consumers expect 100% accuracy.

In these scenarios, the operators need to work closely with their vendor to ensure that they understand the accuracy SLAs that can be established for the applications selected, as well as assessing their ability to rectify any billing disputes over identification accuracy should an application change behavior and subsequently be identified incorrectly during a billing cycle. Working directly with the application vendors can help ensure accurate identification through proactive notifications of any change to the application’s behavior, thus reducing any concerns about false positives or false negatives and mimizing revenue leakage. Sophisticated hueristics that leverage extensive machine learning models on the back end are mandatory in today’s quick changing environment to maintain accuracy.

Encryption is already proving to be a challenge for regulators

### ZERO-RATING

Zero-Rating is a variant on an application-based service plan where rather than charging an uplift fee for an application or class of application, users are allowed to unlimited use with no additional fees. T-Mobile's BingeOn and MusicFreedom plans, as well as iiNet's FreeZone are examples of these types of service offering. Access to the zero rating is usually restricted to a specific set of service plans, but the principle remains the same for the consumer - purchase of that plan means that they can consume unlimited volume (not bandwidth) of data for that application. Encryption will have the same restriction as described above in the application-based service plans section, but the need for accuracy is increased when a consumer believes that they can use unlimited data because the associated billing risk is higher due to increased usage. Again working closely with application vendors and your NI supplier before these service offerings are affected by encryption is highly recommended.

Content identified by traffic source/destination will be less affected by encryption, but with the extensive use of content delivery networks that can change over time, there is still risk in a false negative identification of traffic if changes occur due to capacity expansion, business model changes, or even network issues.

### PARENTAL CONTROLS

Many operators around the world offer parental control services as part of their service packages, sometimes as a value-added offering and sometimes as part of a government-mandated requirement. Parental control services allow the consumer to select specific categories of content (pornography being the most common target) to be blocked on their service offering.

Network-based parental control services are limited to being able to block these sites only on their infrastructure; if a mobile device can access another provider's WiFi network, blocking can not be extended to another access network.

This use case will lose a great deal of granularity at the URL level, but site-level visibility will be maintained. This translates to not being able to block specific sections or URLs on websites that host both regular content as well as questionable content (like Tumblr for example), so parental control offerings in the future should be specific about the risks that can be addressed by the service packages.

### DATA FRAUD MANAGEMENT

From a services perspective, zero-rating is a popular addition to the lineup of mechanisms network operators employ to attract and retain users. However, the introduction of zero-rating unfortunately creates a potential avenue for unscrupulous users to exploit, which directly impacts operators' revenue. To benefit with more certainty in the promise of zero-rating, network operators need to detect and mitigate various methods of zero-rated fraud, primarily HTTP header injection, domain fronting, and DNS spoofing.

Traffic classification that goes far beyond what traditional and embedded NI systems can deliver, including advanced application fingerprinting is required to ensure accuracy and prevent fraud. Advanced reporting and analytics that provide insight into the prevalence of data fraud and the impact of management policies enables an operator to understand how much revenue is at risk from fraud.

## Sandvine's Application Identification Strategy for Use Case Mitigation

**Sandvine's industry-leading Active Network Intelligence engine is ideally positioned to help operators ensure that they do not lose visibility into their networks. With over 2.1B subscribers in over 100 countries worldwide covered by Sandvine solutions, we have a broad view of how applications are affecting regional traffic trends that we report on regularly with our Global Internet Phenomena Report. In fact, our customers tell us that one of the top reasons that they choose our solution is our rapid response to signature detection issues; not only our frequent signature release cycle, but also the accuracy that they see when our solutions are deployed on their networks.**

| NI Challenges | Sandvine Strengths |
|---|---|
| Encryption | Almost 3/4 of over 3000 signatures are for encrypted applications, separate analysis engine for encrypted traffic based on heuristics rather than application metadata. Aggressively engineering more heuristics capabilities to handle the shifts to encrypted DNS and TLS 1.3, which will have a huge impact on existing "DPI solutions from vendors. |
| Asymmetry | Clustering naturally extends to distributed virtual deployments and maintains visibility for asymmetric encrypted applications without any need for traffic re-routing or flow balancing. Clustering ensures that classification is not affected by encryption and do not affect the signature development process. |
| Flow Correlation | Strong support for parent/child flow association, child flow can inherit properties and policies from the parent flow. |
| Localization | Deployments in over 100 countries worldwide, broad support for applications in multiple languages. Localized signature development when Revenue Generation applications are deployed, especially charging services. |
| Speed of Change | Frequent signature updates, support for customer created Virtual Services signatures for rapid deployment of signatures for local OTT applications. |

### ASYMMETRY FOR VIRTUALIZED DEPLOYMENTS

There is one area that needs a bit more detail and that is the issue of asymmetry and the distributed deployments that will become the norm in NFV and SDN. As mentioned in the introduction section, NFV/SDN networks can be highly distributed, dynamically change their behavior based on network conditions and are deployed on COTS hardware that rarely has hardware assist/acceleration capabilities for packet processing offload. This is important because some systems rely on network processors for flow segmentation to send all traffic from a single subscriber to the same CPU.

Sandvine's solution does not impose this requirement on the network - freeing the operator to design the NFV/SDN network that meets their needs - not those of the DPI system. No extra packet routing or load balancing is required and network troubleshooting does not have to take "policy-based routing" when trying to determine the cause of a network or customer issue. Clustering enables state synchronization across multiple systems to simplify application identification and enables queue synchronization to simplify traffic management in networks where multiple systems are used to manage the exit points of the network. This is a huge simplicity benefit for operators and ensuring that dynamic changes in the network can be automatically dealt with for NI deployments without the extra policy-based routing configurations that some vendors require. If each system in a service chain needs its own flow-based load balancer, it will double the number of systems required to deploy services, which is an unacceptable use of CPU cycles in an optimized NFV/SDN deployment.

## INNOVATING NI TECHNOLOGY IN THE ERA OF ENCRYPTION

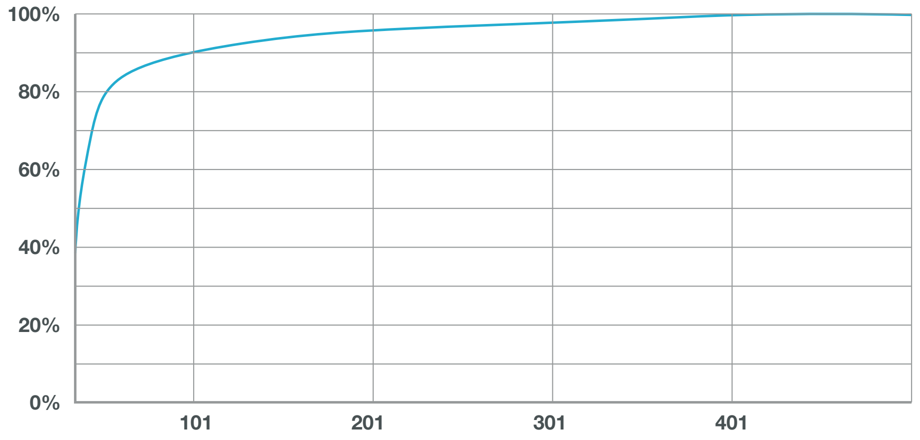
In addition to the above solutions that are already in the product, Sandvine continues to innovate and develop our technology to adapt to the ever changing landscape of the encrypted Internet. As application and underlying Internet infrastructure changes, application identification techniques must adapt in order to maintain the visibility needed to ensure that network operators can deliver a high QoE to their subscribers. The methods described are part of Sandvine's plan of record to enhance our ability to detect applications and give operators visibility into their networks.

## REAL-TIME ENDPOINT CLASSIFICATION

Today the server hostname is visible via the SNI field, allowing the "long tail" of ~ 1 billion URLs to be accurately identified even with HTTPS ("direct") activated. If SNI becomes encrypted in the future, then a method of IP endpoint mapping will need to be used to "infer" hostname on the top 1000s to 10,000s of sites and the remainder will be "dark" and unknown. Sandvine is implementing a new database of "short tail" IP endpoints that cover the most commonly used CDNs and content servers on the Internet, covering virtually all HTTP/HTTPS bandwidth but not the "long tail" of 1 billion URLs. This database will map these sites to IP Endpoint ranges and be updated in near real-time. When this is combined with the existing output from our ANI engine, it will increase our accuracy for traffic that can usually be characterized by source/destination - like video and audio streaming, websites, etc.

Figure 4

Example traffic distribution by URL

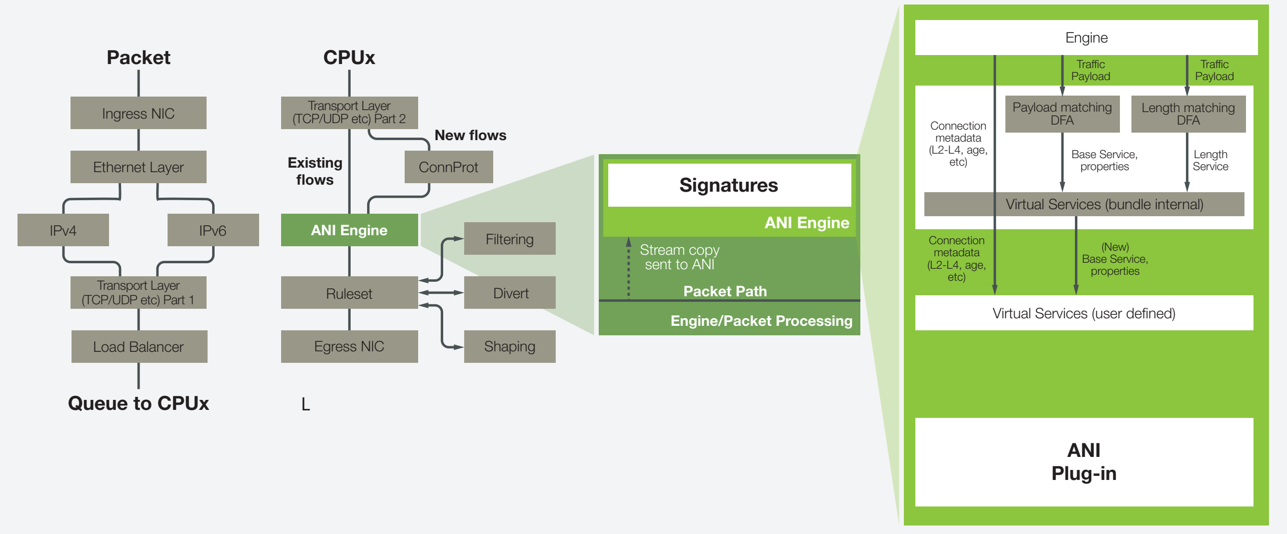


Research that Sandvine has undertaken with our installed base of customers shows that nearly 100% of web traffic (by volume) can be classified in the first few hundred URLs.

## RAPIDLY ADAPTABLE HEURISTICS ENGINES (RAHE)

Sandvine’s ANI engine also has the built-in ability to make calls to other modular classification engines. This capability is how the ContentLogic solution works where once traffic is identified as HTTP, the ContentLogic database is consulted to provide content classification that can be used for analytics and revenue generation use cases (like Parental Control). We are already developing rapidly adaptable heuristics engines (RAHE) that leverage existing behavior flags in (i.e. random-looking/encrypted and VoIP-like) to enhance our ability to gain better classification of specific encrypted traffic types that are required for some of the use cases described above.

Figure 5



## GRANULAR VISIBILITY: THE FOUNDATION OF ACTIVE NETWORK INTELLIGENCE

Building and maintaining a world class network is hard and getting harder due to the prevalence of encryption on the internet. Preserving network visibility has been identified by network operators as critical in evolving their infrastructure to next generation technologies, whether their roadmap includes 5G, NFV, or even automation solutions. Without the right data to feed the decision making process, network operators will struggle to meet the conflicting demands of subscriber's expectations of network quality and the demands of financial markets to reduce CAPEX and OPEX.

Automation of network infrastructure requires signficantly enhanced network intelligence that exceeds the capabilities of most network intelligence systems deployed on networks today. With the maturation of virtualization technology, the CAPEX and OPEX cost to implement network intelligence has dropped and the flexibility to deploy network intelligence solutions. The challenges introduced by encryption for the use cases described in this whitepaper highlight areas that network operators should be aware of when designing their next generation infrastructure that needs to adapt to encrypted applications and maintain the quality levels that users have come to expect.

Sandvine's Active Network Intelligence is taking the challenge of simplifying the solutons needed to engineer and operate world class networks by leveraging our unique network intelligence and closing the loop with automation. Unlike other Network Intelligence solutions on the market, Sandvine's Active Network Intelligence has the ability to sit in-line on the network and make real-time policy changes to improve network performance. Sandvine can deploy this level of automation today across multiple solution verticals, and we are continually innovating with new use cases that leverage our analytics foundation to make better decisions in real-time to enable machine-driven decisions and actions rather than requiring human intervention.

Sandvine is helping our customers maintain visibility in the era of the encrypted Internet and to deliver competitive service offerings to their subscribers. Our industry-leading application identification technology is used in 100 countries by more than 150 Tier 1 & 2 operators worldwide servicing over 2.1B subscribers. Let us help you maintain visibility into the darkening Internet!

v20180711

## ABOUT SANDVINE

Sandvine helps organizations run world-class networks with Active Network Intelligence, leveraging machine learning analytics and closed-loop automation to identify and adapt to network behavior in real-time. With Sandvine, organizations have the power of a highly automated platform from a single vendor that delivers a deep understanding of their network data to drive faster, better decisions. For more information, visit **sandvine.com** or follow Sandvine on Twitter at **@Sandvine.**

SANDVINE

**USA**
47448 Fremont Blvd,
Fremont,
CA 94538,
USA
T. +1 510.230.2777

**EUROPE**
Birger Svenssons
Väg 28D
432 40 Varberg,
Sweden
T. +46 340.48 38 00

**CANADA**
408 Albert Street,
Waterloo,
Ontario N2L 3V3,
Canada
T. +1 519.880.2600

**ASIA**
RMZ Ecoworld,
Building-1, Ground Floor,
East Wing Devarabeesanahalli,
Bellandur, Outer Ring Road,
Bangalore 560103, India
T. +91 80677.43333