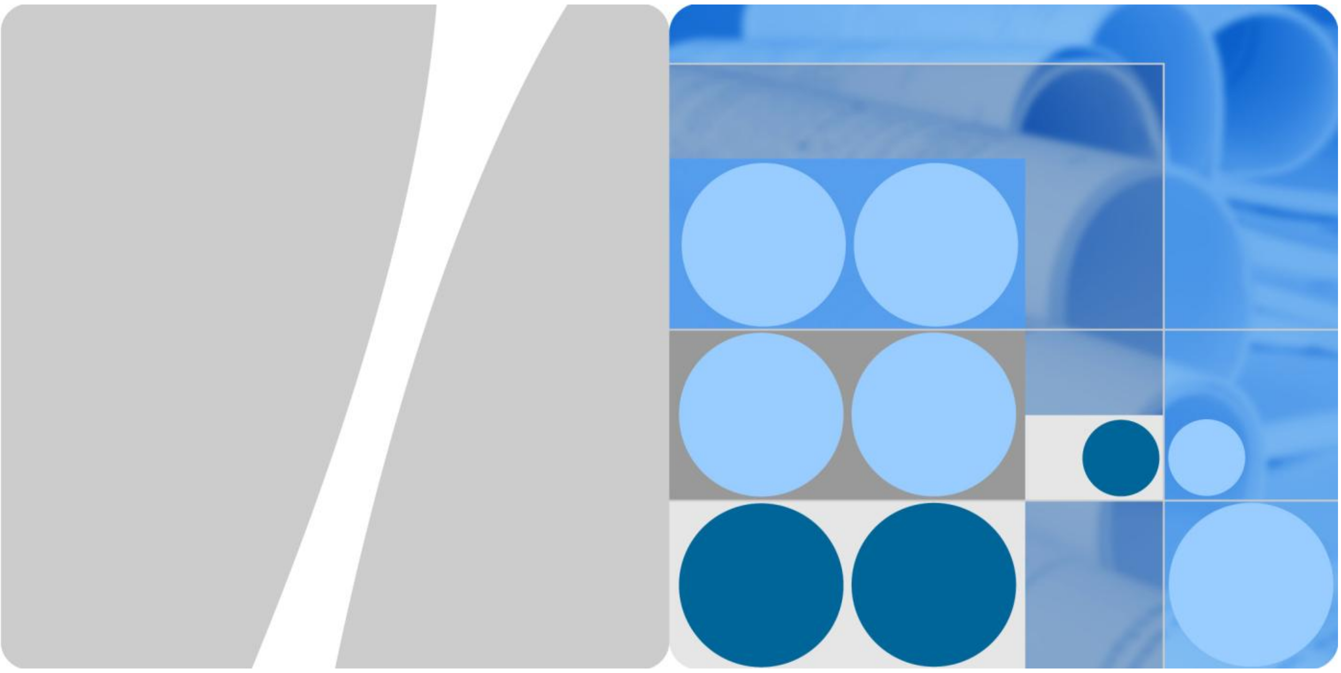


HUAWEI USG6000 Series Next-Generation Firewall

Intelligent Aware Engine (IAE) Technical White Paper

Issue	V1.1
Date	2014-03-14

HUAWEI TECHNOLOGIES CO., LTD.

Trademarks and Permissions

Notice

The purchased products, services and features are stipulated by the contract made between Huawei and the customer. All or part of the products, services and features described in this document may not be within the purchase scope or the usage scope. Unless otherwise specified in the contract, all statements, information, and recommendations in this document are provided "AS IS" without warranties, guarantees or representations of any kind, either express or implied.

The information in this document is subject to change without notice. Every effort has been made in the preparation of this document to ensure accuracy of the contents, but all statements, information, and recommendations in this document do not constitute a warranty of any kind, express or implied.

Huawei Technologies Co., Ltd.

Address: Huawei Industrial Base
Bantian, Longgang
Shenzhen 518129
People's Republic of China

Website: http://www.huawei.com

Email: support@huawei.com

Contents

HUAWEI Secospace USG6000 Series

IAE Technical White Paper

Keywords: IAE, NGFW

Abstract: This document describes the application background, functions, technical methods, and working mechanisms of Huawei IAE.

List of acronym and abbreviations:

Acronym and Abbreviation	Full Spelling
NGFW	Next-Generation Firewall
IAE	Intelligent Aware Engine
IPS	Intrusion Prevention System
AV	Antivirus
SA	Service Awareness
UTM	Unified Threat Management
RBL	Real-time Blackhole List

1 Technical Background

With the increasing of security threats on the Internet, content security requirements of security products become more demanding. Enabling multiple detection functions of UTM products greatly degrades the performance. To meet the requirements on performance and functionality, Huawei has launched an IAE to implement the core security functions of the USG6000 series and bring about new user experience.

The IAE provides a stable and integrated processing framework of content security services for gateway products and incorporates security knowledge databases to help you customize, expand, integrate, and rapidly release content security services.

The IAE meets the following requirements of gateway products:

- High performance independent from the function enabling. Enabling all functions greatly degrades the performance of traditional UTM products because of technical limitations.
- Powerful IPS or UTM with more bandwidths at the egress of the cloud center or large enterprise network
- Application identification, control, and virtualization
- Integrated IPS and automatic detection and handling
- Intelligent interworking (IPS, AV, and URL filtering)
- Content security virtualization (IPS, AV, and URL filtering)
- Platform sharing of multiple products that have the content security function

IAE competitiveness is as follows:

- High performance

The IAE merges the detection requirements of services and implements application-layer parsing and extraction at a time. With the unified policies, enabling all threat prevention features degrades performance by no greater than 50%.

- Immediate delivery and service expansion of multiple products

The IAE focuses on key content security capabilities and processes services effectively based on an appropriate architecture to meet various product requirements.

- Independent evolution and development of core security capabilities

The SA database, IPS signature database, AV database, URL category database, and reputation database can be decoupled from specific products and independently developed and evolved.

2 Definition and Mechanism

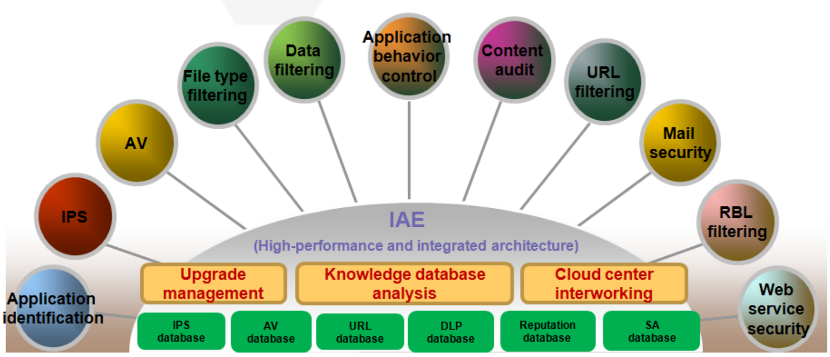
2.1 IAE Definition

IAE is the combination of an integrated processing framework of content security services and a series of security features or components, incorporates a variety of security knowledge databases, and interworks with the Security Intelligence Center to help you customize, expand, integrate, and rapidly release content security services.

The following formula can be used to describe the IAE: IAE = Security framework (flow + proxy) + Security features or components + Compatible knowledge database + Interworking with the Security Intelligence Center

Figure 2-1 shows the features that the IAE supports.

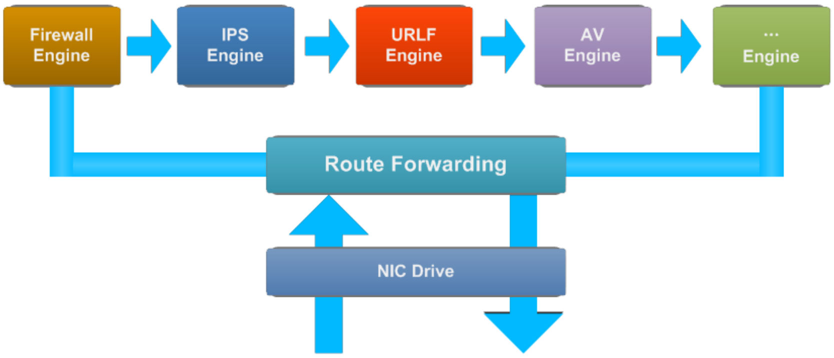
Figure 2-1 Schematic diagram of the IAE feature



2.2 IAE Architecture

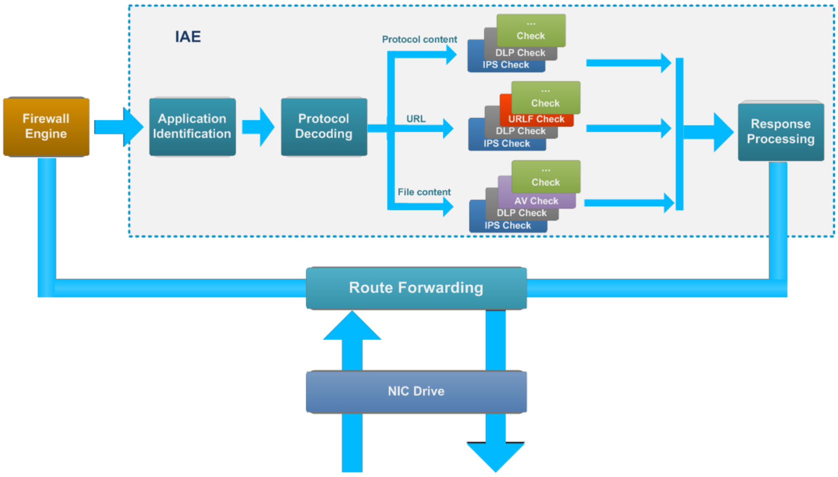
Traditional content security products such as UTM use the architecture shown in Figure 2-2, and all security engines detect network packets in serial mode. This architecture is simple. However, when multiple services are enabled, service modules repeatedly parse data packets, match policies, and take response actions. In the scenario where multiple functions are enabled on a security device, performance is drastically degraded. The security device may become unavailable.

Figure 2-2 Schematic diagram of the UTM architecture



The IAE architecture, based on the unified design, implements protocol decoding on network packets at a time and sends the obtained data including the protocol content, URL, and file content to three processing branches shown in Figure 2-3.

Figure 2-3 Schematic diagram of the IAE architecture



Multiple services of one processing branch can be merged. If the matching objects and processing phases are the same, their state machines can be normalized. If the check objects and conditions can be unified, the IAE checks the objects only once. By processing security services in a unified manner, the IAE improves detection efficiency and reduces repeated parsing, matching, and check. The processing branches send the processing requests to the response module. The response module then takes processing actions by priority.

This architecture greatly improves detection performance when multiple service functions are enabled. Processing multiple services with one parsing action is the core concept of the IAE architecture.

2.3 Key IAE Functions

Table 2-1 lists key IAE functions.

Table 2-1 Key IAE function list

Feature	Function	Description
Service Awareness (SA)	IPv4/IPv6 application identification	Supports application identification in the IPv4/IPv6 environment.
	Multi-channel application identification	Supports accurate identification of multi-channel protocols.
	User-defined application identification	Supports application identification based on user-defined rules.
Intrusion Prevention System (IPS)	Vulnerability-based signature detection	Detects threats based on signatures.
	Botnet/Trojan horse/worm detection	Detects botnets, Trojan horses, and worms.
	User-defined signature detection	Supports detection based on user-defined signatures.
	Correlation rule detection	Supports detection based on correlation rules.
Antivirus (AV)	Virus detection of seven protocols	Detects viruses transmitted over HTTP, FTP, IMAP, POP3, SMTP, SMB, and NFS.
	Response actions including alarming, blocking, advertisement, and attachment deletion	Supports response actions including alarming, blocking, advertisement, and attachment deletion
	Application exception	Exempts some applications from detection.
	Virus signature exception	Adds virus signature exceptions in rules to avoid false positives.
URL filtering	Predefined URL category filtering	Supports URL filtering based on predefined categories.
	Malicious URL filtering	Supports malicious URL filtering.
	One URL in multiple categories	Supports the adding of a URL to multiple categories.

	User-defined URL blacklist and whitelist	Supports user-defined URL blacklist and whitelist.
	User-defined URL categories	Supports user-defined URL categories.
	HTTPS URL category filtering	Supports URL category filtering for HTTPS
File type filtering	Filtering by real file type	Supports identification and filtering of multiple file types.
Data filtering	User-defined file data filtering	Supports user-defined file data filtering.
	File data filtering based on default keywords, such as bank card, credit card, Social Security number, ID card number, and confidential information	Supports file data filtering based on default keywords, such as bank card, credit card, Social Security number, ID card number, and confidential information.
	Web page keyword filtering	Supports web page keyword filtering.
	Search engine keyword-based filtering	Supports keyword filtering of the Google, Yahoo, Baidu, and Bing search engines.
	Microblog keyword-based filtering	Supports keyword filtering of the Twitter, Sina and Tencent microblogs.
	File name filtering	Supports file name filtering over HTTP or FTP.
	Mail title, body, and attachment keyword filtering	Supports data filtering by mail title, body, and attachment keyword.
Application behavior management and control	Refined HTTP behavior management and control	Supports control over file transfer, POST operation, web page browsing, and HTTP proxy.
	Refined FTP behavior management and control	Supports control over the file upload and download and the size of the uploaded or downloaded file.
Content audit	HTTP behavior audit	Audits accessed URLs, file transfer behaviors, browsed web page titles, posted content, and microblog content.
	FTP behavior audit	Audits incoming and outgoing emails.

Real-time Blackhole List (RBL)	Local blacklist and whitelist	Supports local blacklist and whitelist.
	Remote RBL query	Supports remote RBL query.
Mail content filtering	Mail address filtering	Filters the sender and receiver email address during the sending and receiving.
	Anonymous mail filtering	Supports anonymous mail filtering.
	Mail attachment quantity control	Controls the number of email attachments during the sending and receiving.
	Mail attachment size control	Controls the size of email attachments during the sending and receiving.

3 Services and Functions of SA

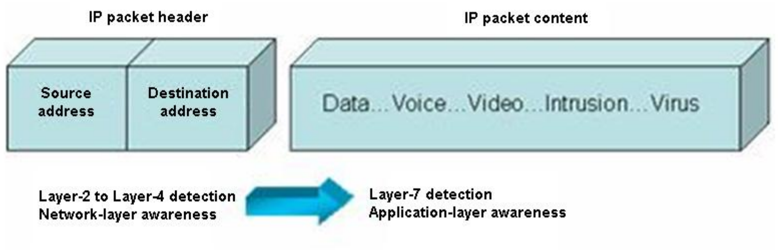
SA is the basis of content security services. The IAE SA identifies protocols and applications of Layer 4 and higher layers, including standard protocols (such as HTTP, FTP, and SMTP), private application-layer protocols, and applications Skype and eMule. The IAE starts decoding based on the identified protocols and defines security policies based on the identified applications.

3.1 Services and Functions

Policies of traditional firewalls are mainly defined on the basis of Layer 3 and Layer 4 features, such as the IP address and port. With the increasing of Internet applications, many applications such as IM, P2P, and online video use port 80. Traditional firewalls cannot distinguish the applications from web browsing. Many applications use dynamic ports and cannot be identified by traditional firewalls.

Besides detection over common packets, SA of the IAE adds application-layer analysis and identifies applications and contents to determine the real services of packets. Figure 3-1 shows the SA mechanism.

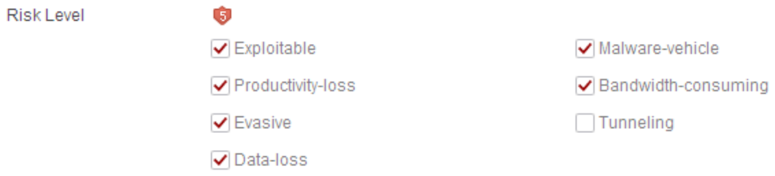
Figure 3-1 Schematic diagram of SA



Application-specific policies can be configured on the USG6000 series that uses the IAE. SA identifies application types of packets based on data characteristics at the application layer. You can define policies based on the application types to prevent port sharing and dynamic ports from causing failures in application identification.

Huawei SA identifies various applications covering those in China, Europe, Middle East, Latin America, and other regions and has high identification ratio and accuracy for encrypted P2P, IM, and VoIP traffic. As shown in Figure 3-2, Huawei SA has defined multiple characteristics of each application. You can easily identify the application that degrades productivity and consumes excessive bandwidths. Based on the characteristics, you can set the security level of each application. A higher security level indicates more security risks. You can filter applications by security level and characteristic.

Figure 3-2 SA application risk level



Risk Level 5

☑ Exploitable

☑ Productivity-loss

☑ Evasive

☑ Data-loss

☑ Malware-vehicle

☑ Bandwidth-consuming

□ Tunneling

3.2 Key Technologies

3.2.1 Signature Identification Technology

Signature identification accurately identifies the type of network traffic based on the signatures of Layer 7 data, such as GET and HTTP/1.1 in HTTP packets. Huawei SA uses a layer-by-layer signature identification technology. When applications are operating over some protocols at the application layer, the SA identifies their relationships layer by layer. For example, the SA can identify such applications as BitTorrent, eMule, and Skype over HTTP.

Signature identification of the SA uses the PCRE regular expression syntax to express complex matching logic and the hardware acceleration engine to match patterns, ensuring high performance.

Users can define applications and rules that support the PCRE regular expression syntax. When new applications emerge or application characteristics change after the upgrade, you can define rules before the new signature database is released. In addition, you can define rules to identify private applications on the enterprise network.

The IAE SA must support IPv6 networks because of the network evolution trend. The SA focuses on Layer-7 protocol load, but not lower-layer differences between IPv4 and IPv6. In some scenarios where differences between the IPv4 and IPv6 addresses must be considered during the parsing of address information, operation correlation table, and IP-domain mapping table, the SA supports IPv6 address processing.

3.2.2 Correlation Identification Technology

Correlation identification is based on the relationships among connections. For example, FTP has a control channel that can be identified by the signature identification technology. During the file transfer, FTP sets up a temporary data channel. The data channel does not have any signatures and cannot be identified by the signature identification technology. However, during the data channel setup, the control channel negotiates IP and port information of the data channel. The SA parses negotiation information and associates the source and destination addresses of the control channel to accurately identify the data channel. Huawei SA supports multi-channel protocols, such as MSN, H323, SIP, MGCP, MEGACO, FTP, MMS, RTSP, GoogleTalk, and H245.

3.2.3 Behavior Identification Technology

Behavior identification mainly identifies the encrypted traffic. Since data is encrypted and its patterns are blurred, signature identification cannot identify the data. Data of some applications is encrypted from the first packet and cannot be identified on the basis of the relationship between the control channel and the data channel. Therefore, behavior identification must collect more information to assist in identification, covering the connection quantity of an IP address, ratio of upstream and downstream traffic, packet transmission frequency, and packet length change pattern. For example, if the application is a VoIP application, the voice data packet length is stable and the transmission frequency is fixed. If the application is a P2P application, a large number of single-IP connections exist, the ports of each connection are different, the file sharing packet is large, and the packet length is stable.

Huawei SA applies heuristic identification based on characteristics and correlation relationships of data flows to accurately identify the encrypted BitTorrent, eDonkey/eMule, and Thunder data.

3.2.4 Application Tracking and Analysis

Network applications are blooming and changeable. The SA must constantly track application updates and network traffic changes. When new applications emerge or application characteristics change after the upgrade, you must analyze the new applications or changes timely to update the signature database.

Huawei Security Intelligence Center has an automated application tracking system and a large number of application analysis experts. The application tracking system automatically tracks the release of application software, downloads and installs the software, simulates the software use, and identifies traffic generated by the software. If traffic cannot be identified or the identification ratio is low, the application tracking system sends the traffic to an automated signature obtaining system for processing. If no signature is obtained from the traffic, application analysis experts analyze the traffic and obtains signatures. The obtained signatures are imported to an automated verification system for verification.

In addition, Huawei has deployed SA on thousands of network devices to monitor traffic changes. If unknown traffic is identified, the SA sends the traffic to the Security Intelligence Center after the customer approval. The Security Intelligence Center then analyzes characteristics of the traffic, identifies new applications, and obtains patterns.

Huawei SA releases a signature database monthly to add new applications and updates the changed applications. The experts can provide emergency response services within 24 hours.

4 Services and Functions of IPS

IPS is an important security function of gateways. IPS analyzes and identifies attacks, blocks malicious traffic, implements proactive and real-time defense against attacks, and provides enterprise networks with virtual patches. It discovers such web security threats as SQL injection, XSS attack, and Trojan horse websites. In addition, IPS detects Trojan horses, worms, botnets, and spyware to purify network traffic for enterprises and enhance security.

4.1 Services and Functions

4.1.1 Signature-based Threat Detection

Signature-based threat detection can be applied in the following scenarios.

Defense Against Vulnerability-based Attacks

Vulnerability refers to a security defect in the computer hardware and software, protocol implementation, or policy. If some vulnerabilities exist on devices, intruders may access unauthorized files, obtain sensitive information, or run programs.

Zero-day vulnerability is utilized or exploited before the system provider releases patches. The vulnerability information does not spread in a large scope and is not obtained by most users. In addition, the system provider does not release any patch for the vulnerability. Hackers may take advantage of the vulnerability to launch attacks.

Web Security Defense

Similar to web application programs, web services also have vulnerabilities, which can be exploited by SQL injection and XSS attacks. Different from traditional web pages, web services become more open. A large number of attacks are launched on desktop client software, and users frequently access the applications and data of an enterprise. Attack risks are increased, and web services become inviting targets. Web services are running over HTTP and allow cross-website communications without requiring firewall reconfiguration. Traditional firewalls cannot analyze web service communications over HTTP, which is risky.

Attackers use HTTP or HTTPS applications to evade firewall detection and HTML evasion technologies to attack web security servers. A large number of insecure desktop application programs are developed, and attackers launch more and more attacks on the ActiveX controls, browser plug-ins, components, and JavaScript of desktop clients. Client protection becomes more important. Hackers obtain the control permissions on web servers to tamper with web page contents or intercept sensitive data. Some hackers even inject malicious code into web pages and use web page Trojan horses to infect more clients. Hackers can control website visitors and employees' computers to intercept bank accounts and confidential information. Web-based Trojan horses are easy to develop, and network vulnerabilities are inevitable. Web page Trojan horses that exploit website vulnerabilities become prevailing and are widely used by hackers to launch attacks and spread Trojan horses.

Defense Against Malicious Code

Botnet refers to a network where a controller infects many hosts with the Bot program by means of one or more spreading methods. The botmaster and the infected hosts (zombies) form a one-to-many control network.

- Bot: Bot, short for Robot, can automatically execute pre-defined functions or be controlled by predefined commands. A Bot may not be malicious. However, in Botnet, bots are designed for malicious purpose.
- Zombie: Zombie is the host on which malicious Bots or other remote control programs are running.
- Command & Control Server: The IRC server that connects to the IRC Bot is called the Command & Control Server (C&CS). The attacker uses this server to deliver commands and control the IRC Bot.
- Botnet: Botnet is a network that consists of many hosts with malicious Bot programs and is controlled by an attacker.

Driven by economic interests in recent years, botnets develop rapidly, become more complex and concealed, and have diversified control means. Detecting and controlling botnets become demanding challenges.

A Trojan horse is an application installed on a computer to intercept data and launch attacks from within. The Trojan horse is a program that provides some useful or appealing functions as well as copies files or steals passwords. Once being injected into a computer, Trojan horses intercept important files and data and closely monitor all user operations. In addition, the computer can be remotely controlled by an attacker to launch attacks.

A worm is a program that independently runs without human intervention. The worm continuously obtains partial or whole control rights of the computers with vulnerabilities to spread. Different from a virus, a worm can run by themselves and spread rapidly, without requiring user participation.

Spyware is a software that installs backdoors and collects user information without customer's knowledge. Spyware performs the following operations:

- Deteriorates user experience, privacy, and system security capabilities.
- Uses system resources, including programs installed on computers.
- Collects, uses, and spreads personal or sensitive information.

IPS mainly uses the misuse detection model, integrates intrusion signatures into the knowledge database, and matches data flows with signatures in the knowledge database to discover threats. The major advantages of this method are high detection efficiency, low false positive ratio, and low detection costs. This method relies on the accumulation of signatures in the knowledge database. Therefore, long-term maintenance over signature databases is required, and unknown threats cannot be defended against.

The attack defense mechanism analyzes vulnerability patterns, extracts common signatures, and matches patterns to detect exploits. Web security protection defends against common system vulnerabilities as well as protects HTTP applications. It implements URL anti-evasion, prevents HTML confusion, and matches patterns to detect attacks on HTTP applications. The IAE analyzes Trojan horses, worms, and botnets, extracts communication features, and identifies roles based on the features to discover threats.

4.1.2 Threat Detection by User-Defined Signature

In practice, attack signatures are released later than the emergence of new attacks. Some users who deeply understand the new attacks can define some signatures by themselves to rapidly defend against the attacks.

4.2 Key Technologies

4.2.1 Protocol Identification Technology

To implement intrusion detection, virus detection, and data filtering on application-layer data, you must identify application-layer protocol types and provide protocol-specific handling methods. Common protocol identification is to identify protocol type by port defined in RFC. However, this method is low in accuracy. The IAE also analyzes data of subsequent packets to comprehensively identify the protocol type.

To implement IPS, the IAE identifies actual protocols (see the SA service) and carries out protocol identification and threat scanning at the same time. The IAE detects attack threats even if they use ephemeral ports. For example, the IAE detects HTTP attacks that are launched over port 3128. To meet requirements on threat detection, IPS analyzes and identifies multiple protocols and file types and incorporates SA capabilities that identify hundreds of transport protocols to detect real traffic.

4.2.2 Accurate Protocol Decoding

IPS carries out in-depth application identification, protocol decoding, and deep threat detection to effectively identify attacks. Protocol decoding is an essential step of deep threat detection to reduce the computing workload of signature matching, identify and deal with evasion behaviors, detect protocol anomaly attacks, and improve threat detection accuracy.

The IAE decodes hundreds of protocol variable fields. Decoding protocol variable fields is based on research in network attacks and analysis on protocol information of signature databases.

In the protocol decoding phase, IPS normalizes anti-evasion technologies, including application protocol packet fragmentation, flow segmentation, RPC fragmentation, HTML confusion, and URL confusion.

Based on protocol decoding, IPS detects protocol anomaly attacks. Hackers attack network servers because the servers are not perfect in design and protocol anomalies are not fully considered. Hackers send non-standard or overflowed protocol data to the servers to control or break down them. The IAE detects protocol anomalies and applies in-depth protocol decoding to identify behaviors that intrude application servers and clients by severity, including behaviors that violate RFC regulations, overlength fields, inappropriate interaction sequences of protocols, incorrect protocol parameters.

Protocol anomaly detection covers more than 40 protocols, including HTTP, SMTP, FTP, POP3, IMAP, MSRPC, NETBIOS, SMB, TDS, TNS, TELNET, IRC, and DNS.

4.2.3 File-based Detection Technology

IPS uses file-based anomaly detection to detect a file as a "protocol" and identify malicious files. The IAE also identifies actual file types (see the file type filtering service) and carries out file identification and threat scanning at the same time. The IAE detects attack threats even if they change file name extensions to evade detection. For example, the IAE detects a PDF file even if the file name is changed from **attack.pdf** to **attack.txt**.

IPS detects the files transferred over most Internet protocols, including HTTP, SMB, FTP, SMTP, POP3, IMAP, and NFS. The built-in file identification engine of the IAE identifies hundreds of file types including PE, ZIP, OFFICE, PDF, JPG, AVI, and SWF and detects malicious files.

4.2.4 Pattern Matching Technology Based on Network Features

Network traffic contains the features of intrusion attacks, Trojan horse and virus spread, vulnerability exploits, and communication behaviors of Trojan horses and botnets. The IAE analyzes these features and generates network behavior feature codes to detect abnormal traffic and block malicious network behaviors. Most malicious network behaviors can be detected by packet feature. IPS provides the multi-pattern matching technology and supports regular expressions to improve rule flexibility and accuracy.

4.2.5 Detection Technology Based on Correlation Analysis

Only the single-packet pattern matching technology for network features cannot accurately detect intrusion behaviors and malicious traffic of zombies, Trojan horses, and worms. The IAE must analyze a flow or implement correlation analysis on packets of multiple flows. The IAE collects patterns of attack behaviors, forms features rules, and inspects network traffic for matching attack behaviors. The feature can be a protocol field, such as length, value, or content, sequence of key parameters, feature appearance quantity, or relationship between features. The IAE detects more accurate results based on multiple feature rules.

4.2.6 Protocol Anomaly Detection Technology

Protocol anomaly detection is a basic intrusion detection method. Hackers attack network servers because the servers are not perfect in design and protocol anomalies are not fully considered. Hackers send non-standard or overflowed protocol data to the servers to control or break down them.

IPS detects protocol anomalies and applies in-depth protocol analysis to identify behaviors that intrude application servers and clients by severity, including behaviors that violate RFC regulations, overlength fields, inappropriate interaction sequences of protocols, incorrect protocol parameters.

IPS also considers an abnormal file structure as a protocol anomaly. In this case, IPS can detect buffer anomaly attacks or scripting attacks hidden in file content.

4.2.7 Hardware Acceleration Detection Technology

IPS must have high detection accuracy and performance. In addition to high-performance processors, IPS must use a high-speed multi-pattern matching engine and packet decompression engine. IAE uses industry-leading MIPS64 processors of Cavium (a multi-core MIPS and ARM processor provider) to provide a high-performance pattern matching engine for IPS and hardware decompression capabilities to decompress ZIP packages. The IAE enables high-performance IPS detection over files in the compressed package.

4.2.8 Web Attack Behavior Detection Technology

In addition to anomaly detection over HTTP traffic, the USG6000 series restores user data, detects behavior features, and checks whether the data belongs to a legitimate user or the SQL injection and XXS attacks.

4.2.9 Comprehensive Anti-Evasion Technology

Because of network protocol complexity and TCP/IP openness, attackers deform protocol traffic. Unintentional traffic deformation cannot be distinguished from malicious traffic deformation by attackers. Permitting or denying all deformed traffic may cause attacks or affect normal services. In this case, you must use the IAE to normalize traffic (by shaping) as follows:

1. IP fragment reassembly: caches and reassembles out-of-order fragments of packets to ensure that the first fragment reaches first and subsequent fragments reach in order.
2. TCP flow reassembly: includes TCP status maintenance, overlapped TCP segment processing, overlapped data discarding, and TCP option check.
3. RPC fragment reassembly and multiple request binding
4. Normalization for the inserted character, encoding, and path of URLs
5. Processing of inserted FTP characters
6. NetBIOS and SMB anti-evasion
7. HTTP anti-evasion

4.2.10 Protection Using User-Defined Signature

In addition to predefined rules, IPS enables users to define signatures by themselves. IPS provides refined customization based on common fields of protocols. The protocols include HTTP, FTP, DNS, SMTP, POP3, IMAP, NETBIOS, SMB, DCERPC, SUNRPC, MYSQL, TNS, TDS, and FILE.

4.2.11 Signature Database Update

Huawei IPS team closely traces security bulletins of the renowned security organizations and software vendors, analyzes and verifies network threats, and generates the signature database that protects the software system including the operating system, application programs, and database. In addition, the information collection system captures packets of the latest attacks, worms, and Trojan horses, extracts their signatures, and discovers threat trends. The IAE can rapidly obtain the latest signatures to defend against zero-day vulnerability exploits.

The IPS signature database of the IAE supports the following update modes:

- Scheduled automatic update: Updates can be immediately implemented without user intervention to defend against new attacks. This mode applies to the devices that connect to the update server. To verify that the downloaded signature database is secure and available, you can regularly download a new version and apply the version only after confirmation.
- Real-time update: When a version is released but the automatic update time does not approach, you can update the signature database immediately. The advantage is high timeliness. You can know the update results immediately.
- Local update: When a device cannot connect to the update server or must be rolled back to an earlier version, you can perform local update to switch the current version to the target version.
- Version rollback: The current version can be rolled back to a target version. If the false positive ratio of the current version is too high, the detection ratio is too low, or other improper elements exist, the version can be rolled back to a normal version.

5 Services and Functions of AV

AV is a major function of security protection gateways. AV of a gateway detects files transferred over protocols and applications and deals with the detected viruses according to the specified action to block virus files from entering the protected network.

5.1 Services and Functions

5.1.1 PE Virus Detection

According to statistics on virus file types, Portable Executable (PE) file viruses are prevailing. During the evolution of network viruses, PE file viruses are destructive. PE file viruses implement virus functions and behaviors, and other viruses assist in the spread of PE file viruses.

To greatly improve network security status, you must control the spread of PE file viruses.

Huawei AV detects the following PE file viruses:

- Trojan horse programs
- Worms
- Backdoor programs
- Downloaders
- Droppers
- Dialers
- Bots
- Clickers
- RootKit programs
- Adware
- Spyware

5.1.2 Immediate Coverage over Viruses

Huawei AV timely covers pandemic viruses to provide high performance and refresh detection:

- Huawei cooperates with security vendors and organizations inside and outside China to obtain the latest threat information.
- Huawei virus analysis team researches new vulnerabilities and viruses on networks, analyzes the latest threat data, summarizes the data in real time using threat analysis technologies, and generates virus signatures.

- The update center timely releases new virus signatures for the managed devices.

The threat analysis system enables the virus signature database to cover the latest, prevailing, and destructive viruses and secure the protected network.

5.2 Key Technologies

5.2.1 File Identification

AV of a network gateway detects transferred files and identifies file data in network traffic. Huawei AV identifies common file transfer and sharing protocols, such as HTTP, HTTPS, FTP, SMTP, POP3, IMAP, SMB, and NFS, decodes traffic of these protocols, and obtains file data. After obtaining the file data, Huawei AV also identifies file types and detects PE file viruses.

5.2.2 Flow-based Virus Signature Matching

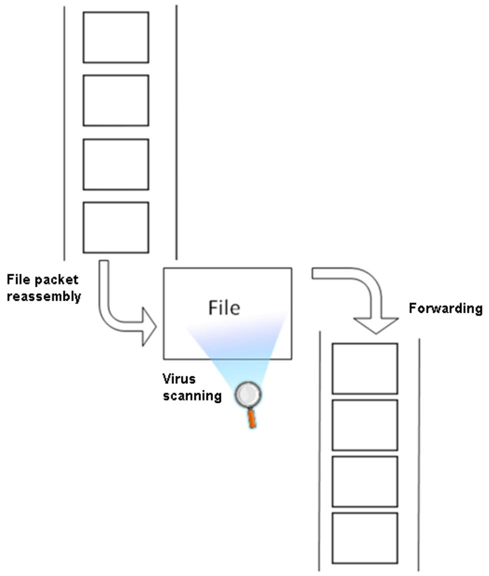
As shown in Figure 5-1, signature-based virus detection is implemented on an entire file. The detection engine opens the file, extracts data at the position specified by a signature, and matches the data with the signature.

If a gateway is deployed, the gateway must reassemble data packets into a complete file and start virus detection.

A large amount of memory and CPU resources must be used to reassemble files.

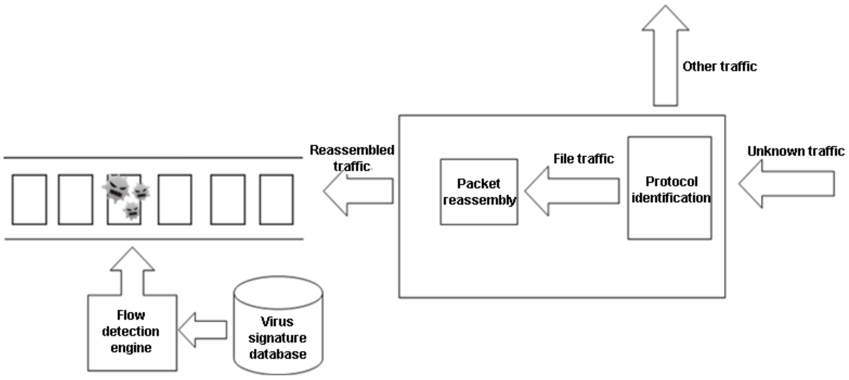
File reassembly, scanning, and retransfer of the gateway prolong the file transfer duration, reduce data throughput, and degrade user experience.

Figure 5-1 File-based AV scanning



The virus flow signature obtaining technology of Huawei detects viruses based on data packets to resolve the previous problems. This technology helps the USG6000 series obtain signatures from the data packets of virus files. When detecting network traffic for viruses, the USG6000 series identifies virus files if the packets match a virus signature. This technology avoids performance deterioration caused by file reassembly and retransfer and improves detection performance. Figure 5-2 shows flow-based AV scanning.

Figure 5-2 Flow-based AV scanning



The USG6000 series uses protocol identification to identify traffic of transport protocols. File traffic is reassembled according to the offset sequence of the file. The reassembled packets reach the detection engine. The detection engine matches the packets with signatures in the virus signature database. If a match is found, the detection engine considers the file as a virus file. The USG6000 series processes the data flows of the file based on user configuration, such as blocking or alarming. If no match is found, the USG6000 series permits the packets.

5.2.3 Flow-based Heuristic Virus Detection

Heuristic detection has enhanced the feature matching technology and can defend against unknown viruses. Heuristic detection is the primary method of dealing with unknown viruses. Based on flow-based detection, Huawei AV supports statically heuristic detection to analyze tens of millions of malicious program samples, extracts the code logic of malicious files, and generates patterns. If file data matches the code logic of a pattern, the USG6000 series considers the file as a virus file.

5.2.4 Virus Exception Handling

To avoid false positives (that may fail the file transfer) caused by AV detection, Huawei AV provides virus exceptions. If a rule generates false positives, you can enter the rule ID to disable the rule, so that the normal file transfer is not interrupted.

5.2.5 Virus Signature Database Update

The virus signature database updates at least once every 24 hours to push new virus signatures to gateways and prevent the virus spread.

The update modes of the virus signature database are as follows:

- Scheduled automatic update: Updates can be immediately implemented without user intervention to defend against new attacks. This mode applies to the devices that connect to the update server. To verify that the downloaded signature database is secure and available, you can regularly download a new version and apply the version only after confirmation.
- Real-time update: When a version is released but the automatic update time does not approach, you can update the signature database immediately. The advantage is high timeliness. You can know the update results immediately.
- Local update: If the USG6000 series fails to connect to the update server, you can download the signature database update package from the update website and load the package on the USG6000 series.
- Incremental update: Virus signatures are incrementally updated without downloading oversized files to improve efficiency and reduce bandwidth consumption.

6 Services and Functions of URL Filtering

The rapid development of the Internet enriches human life and brings about more social issues. There are two categories of social issues, vulgar information flooding and privacy security threat.

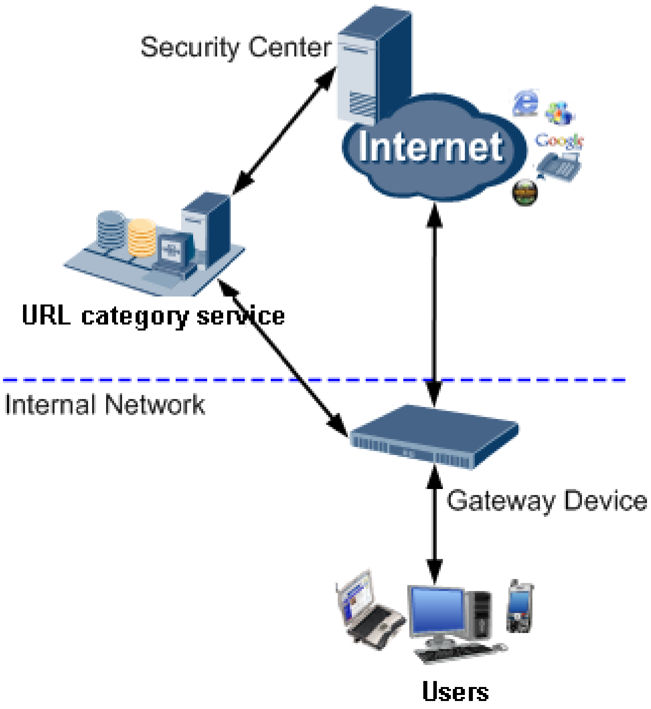
Nowadays, pornography, violence, and other bad taste websites are increasing and accessible to any user. According to the survey, the number of pornographic web pages exceeds 370 million and about 20 thousand pornographic photos emerge on the Internet per day. Young netizens are suffering from vulgar information and unhealthy content. Preventing the spread of vulgar information and unhealthy content becomes a demanding issue. URL filtering of the USG6000 series reduces the vulgar information and unhealthy content.

Protecting user privacy also becomes a major concern. With the development of e-commerce, people make diversified transactions on the Internet. If key information such as account and password leaks, people may encounter severe property loss, which hinders the healthy development of the Internet. The common methods of user privacy theft include disseminating and using malware such as Trojan horses and tricking users with phishing websites. URL filtering identifies and blocks malicious websites to minimize the possibility of Trojan horse infection. URL filtering is also the best method to defend against phishing websites.

The Internet is a huge resource warehouse and business platform of enterprises. More and more enterprises rely on the Internet to carry out businesses. If you prevent employees from browsing non-work-related websites at working hours by disconnecting the Internet, the loss outweighs the gain. URL filtering allows employees to access some websites and traces their online behaviors to ensure legal compliance, understand organization status, and improve user management, meeting requirements on online behavior management of employees.

Figure 6-1 shows the deployment of the URL categorization service.

Figure 6-1 URL categorization deployment



In the remote query deployment, you can integrate Huawei IAE into devices such as the NGFW, but not deploy a URL query server additionally.

The IAE communicates with Huawei URL category remote query server cluster to implement user authentication and URL category query services. The update center periodically updates the URL category database. The IAE also provides the cache management function for local URL category data to improve overall performance.

The remote query deployment is simple and low in costs. Users can obtain URL category services only if the Internet is connected. URL category query uses the Internet. Service quality is greatly affected by network status even if query servers are deployed in multiple areas. Therefore, remote query deployment applies to the scenario that does not have demanding requirements on data processing delays.

6.1 Services and Functions

6.1.1 Management and Control Based on Page Content

Huawei URL filtering is based on the URL category database that contains more than 85 million URLs in 16 languages with 90% URLs over the host level. The URLs fall into more than 40 categories and more than 100 subcategories by page or host content, web page layout, and link. Common URL categories are P2P, download, culture, gaming, entertainment, religion, sex, job hunting and recruitment, search/portal, government/politics, education/science, press/media, tourism, law, IT, forum, shopping, business/economics/finance, vulgar/thrilling content, gambling, drug, malicious website, pornography, crime, weapon, fraud, and cult. You can define access policies on the USG6000 series based on URL categories to block access to undesirable information and protect user privacy. In addition, the USG6000 series provides many configuration templates, hierarchical security protection and data filtering, URL filtering policy templates from basic security control to advanced data filtering. You can choose a template and modify it when necessary.

6.1.2 Malicious URL Management and Control

Huawei URL category database contains malicious URLs, covering prevailing malicious websites and phishing websites. Filtering out malicious URLs plays an important role in protecting user privacy.

6.1.3 User-Defined URL, Blacklist, and Whitelist

URL filtering allows user-defined URL categories, URLs, blacklist entries, and whitelist entries. User-defined URL categories can include predefined URL categories in the system. User-defined URLs can contain wildcards. The user-defined URL categories, URLs, blacklist entries, and whitelist entries enable flexible configurations.

6.2 Key Technologies

6.2.1 Hierarchical Cache Mechanism

The URL category database is huge and cannot be saved on your device. Huawei URL filtering uses a multi-level cache query system that saves some hotspot URLs on the device. Hotspot URLs can be dynamically updated on the basis of user access status. Hotspot URLs accounts for 80% access requests. Users can query other URLs that are not covered on the remote URL query server. The device sends URLs to the query server, and the server returns URL category information. Huawei has deployed a large number of query servers around the world. In normal network conditions, the remote query delay is less than 100 ms.

6.2.2 URL Collection and Categorization

The speed of the URL increase and content change is much faster than the increase of protocols and applications. Huawei Security Intelligence Center and categorization system continuously collect and discover URLs and update the attributes of existing URLs.

The collection and discovery process is as follows:

1. Based on the collected website information, the Security Intelligence Center continuously refreshes new websites on the Internet and senses URL changes.

2. After obtaining website contents, the URL categorization platform uses an automatic categorization algorithm to categorize websites based on page contents, images, and links.

3. After obtaining website categories, the URL category database releases the categories to the update center, so that the URL query server on the customer network can obtain the latest URL categories.

6.2.3 Categorization Algorithm

Based on the latest research findings in fields such as machine learning, natural language processing, and image processing, the categorization algorithm of the IAE URL category database considers various features such as the text content, images, layout, and links on web pages.

The enormous amount of frequently-changing websites poses a great challenge to the storage, error-tolerance, and computing capabilities of the data processing platform. The Security Intelligence Center, based on the data processing analysis system over the distributed cluster, stores and analyzes the content and images on millions of web pages and effectively supports the production flow of the IAE URL category database.

The IAE URL category database uses automatically categorized website contents provided by the Security Intelligence Center. Algorithm engineers in the IAE URL category database team participate in training and set up categorization models to verify massive data and constantly apply more accurate categorization models to the actual production environment.

7 Services and Functions of File Type Filtering

With the rapid development of network technologies, the loss of enterprises' confidential information and users' personal information has become a major concern in information security. Viruses infect files or attach to files to evade detection and penetrate firewalls. File security is a growing concern, but traditional firewalls and UTM devices cannot meet requirements on file security.

In this context, file type filtering technologies come into being. These technologies filter files by type.

Confidential information and viruses are found in specific types of files. For example, confidential information is stored in document files and viruses are attached to executable files. File type filtering blocks specified types of files to reduce the risks of information leaks and virus infections on an intranet.

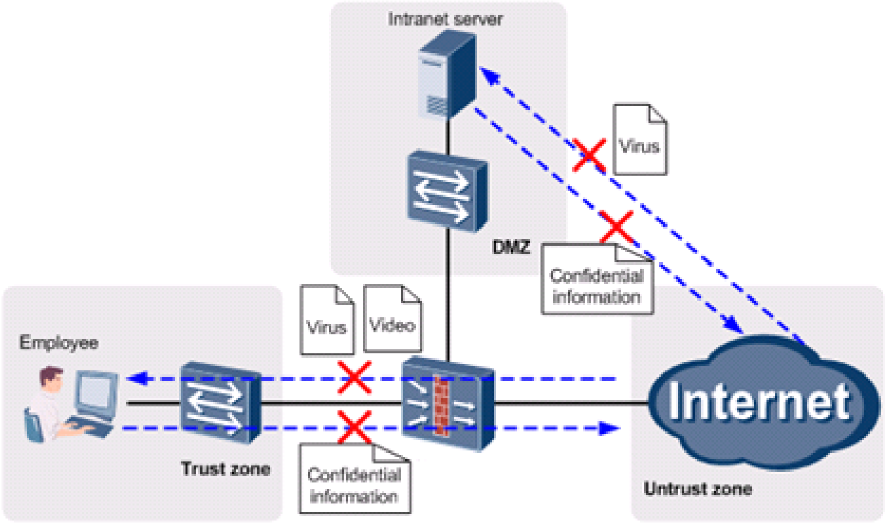
To further reduce the risks of information leaks and virus infections on the intranet, you can apply data filtering or AV with file type filtering.

7.1 Services and Functions

7.1.1 File Type Filtering

Figure 7-1 shows an application scenario of file type filtering.

Figure 7-1 File type filtering



The functions of file type filtering are as follows:

- Reduces the risks of confidential information leaks.

Confidential information is saved in documents and compressed files. If confidential information is not adequately protected, it may be uploaded to the Internet by employees or stolen by hackers. File type filtering blocks the upload of documents and compressed files to the Internet and prevents Internet users from downloading documents and compressed files on the intranet server. Therefore, the risks of information leaks are greatly reduced.

To prevent information leaks, the USG6000 series prohibits employees from uploading common documents, R&D files (such as C, CPP, and JAVA files), and compressed files to the intranet server or the Internet.

- Reduces the risks of virus infections on the intranet.

Viruses are generally attached to executable files and are increasingly difficult to detect and prevent. File type filtering prevents intranet users from downloading executable files from the Internet and blocks Internet users from uploading executable files to the intranet server. Therefore, the risks of virus infection are greatly reduced.

To reduce the risks of virus infections, the USG6000 series prohibits employees from downloading executable files from the Internet and users on the Internet from uploading executable files to the intranet server.

Some employees download non-work-related videos and images that may exhaust bandwidth and compromise employee productivity. File type filtering prevents intranet users from downloading non-work-related videos, images, and compressed files to ensure sufficient bandwidth for normal services and high productivity.

7.2 Key Technologies

7.2.1 File Type Identification Technology

Huawei file type filtering identifies the application that carries files, file transfer direction, and real file type.

File Type Identification

Application: Files are transferred over application protocols such as HTTP, FTP, SMTP, POP3, NFS, SMB, and IMAP.

File transfer direction: The value can be uploaded or downloaded.

Real file type: The real file type can be identified by the USG6000 series. For example, the file type of the Word document whose name is changed from **file.doc** to **file.exe** is still identified as **doc**.

File Anomaly Identification

The global configuration of file blocking defines the action for handling anomalies in file type identification. If the results of file type identification are normal, global configuration is not required. The possible anomalies in file type identification are as follows:

- Mismatched file name extension: The file type and file name extension are not matched.
- Unidentifiable file type: The file type cannot be identified and no file name extension exists.
- Damaged file: The file is damaged and its type cannot be identified.

Handling Policy After File Identification

The IAE determines the matching of file blocking rules and matching conditions according to the results of file type identification and anomaly handling.

To implement the matching of file blocking rules, the IAE matches file attributes (application, direction, file type, and file name extension) with the rules defined in the file blocking profile.

If the file attributes match all rules, the file matches the file blocking profile successfully. If one condition is not met, the IAE matches the file attributes with the next rule. If no rule is matched, the IAE allows the transfer of the file.

If the file matches a rule, the IAE implements the action defined in the rule. If the action is **Block**, the IAE blocks the file transfer. If the action is **Alert**, the IAE allows the file transfer and generates a log.

8 Services and Functions of Data Filtering

Enterprises have spent huge amounts of money to deploy traditional security products, such as firewalls, IDS/IPS devices, and antivirus products because of frequent attacks from the Internet. The products scan only the data that enters an intranet for abnormal content, but not the sensitive data transmitted between intranets or from the intranet to the Internet. This issue becomes a bottleneck of enterprise security protection. Data filtering monitors network behaviors and uses the IAE to prevent information leaks that utilize emails, web pages, and transferred files, which ensures data security and helps enterprises protect intellectual properties.

Data filtering consists of file data filtering and protocol data filtering. File data filtering uses a flow-based technology that filters file data. Protocol data filtering detects and filters the sensitive information and keywords of protocols.

8.1 Services and Functions

8.1.1 Protocol Data Filtering

You can configure policies to check whether sensitive data is transmitted by a specific application protocol. Protocol data filtering covers the following information:

- User-defined keywords that support regular expressions
- Web page data
- Search engine keywords
- Microblog keywords
- HTTP/FTP upload or download file names
- Email title, body, and attachment keywords

8.1.2 File Data Filtering

You can configure policies to check whether sensitive data exists in the file to be transferred. Protocol data filtering provides the following functions:

- Data filtering on Microsoft Office files
- File data filtering based on default keywords, such as bank card, credit card, Social Security number, and ID card number
- Data filtering on the compressed files

- Data filtering on the files transferred over protocols HTTP, FTP, SMTP, POP3, NFS, SMB, and IMAP

8.2 Key Technologies

8.2.1 Flow-based File Data Filtering Technology

Traditional file data filtering receives all file data, reassembles the data into a complete file, parses the file, and matches the file with keywords. After the detection is complete, the traditional file data filtering resolves the file into packets and forwards them.

Flow-based file data filtering starts to parse packets and match them with patterns when some data packets of a file are received. In this case, packet receiving, detection, and sending are concurrently carried out, which greatly improves the packet detection speed and forwarding speed and reduces resource consumption. Figure 8-1 shows the traditional and flow-based file data filtering technologies.

Figure 8-1 Traditional and flow-based file data filtering technologies

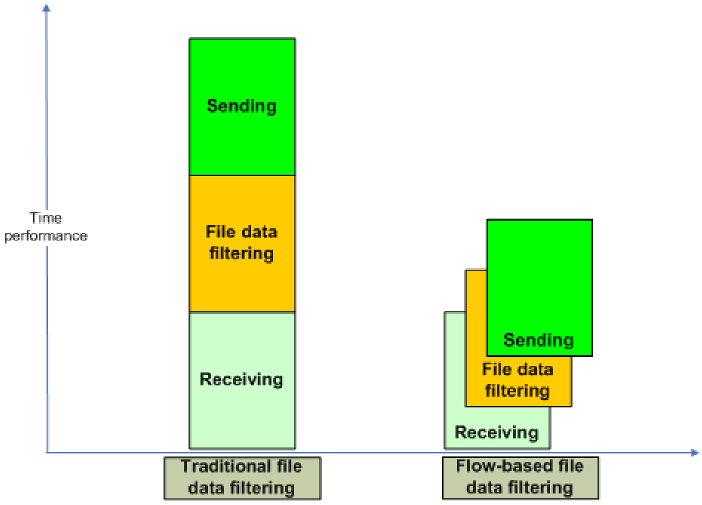
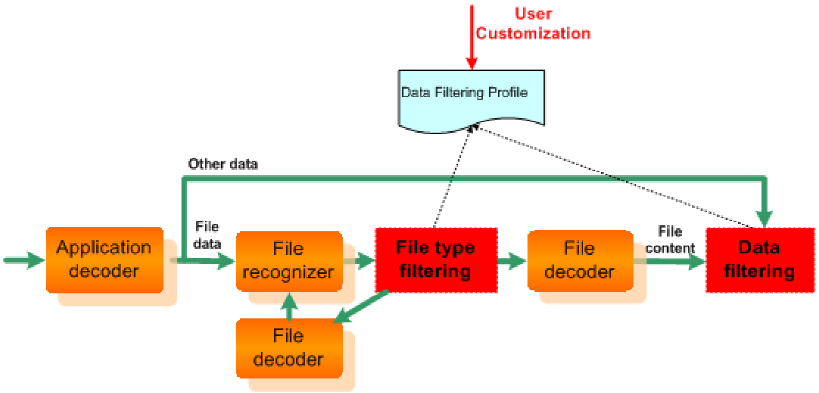


Figure 8-2 shows the detailed detection flow.

Figure 8-2 Detailed detection flow



Data filtering uses in-depth content identification and contextual analysis technologies to identify and monitor the data transmitted over networks in real time and applies sensitive data protection policies to manage and control user behaviors. Data filtering is based on flows and can be applied without any reassembly to greatly reduce costs in system resources.

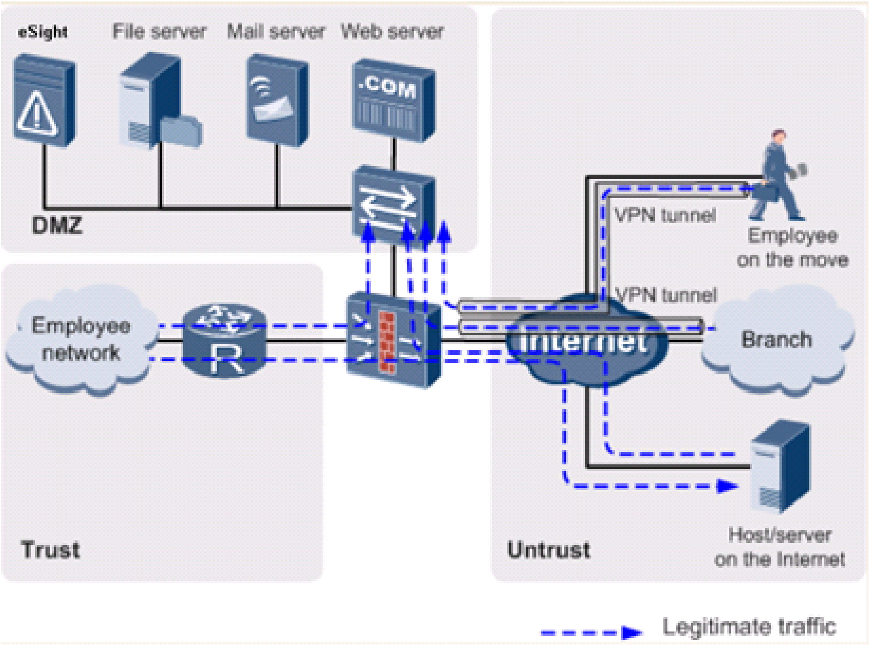
9 Operation and Deployment

The IAE can be deployed on security gateways such as Huawei NGFWs and routers and applies to various deployment scenarios:

9.1 Application-Layer Security Protection at the Enterprise Border

Figure 9-1 shows the USG6000 series deployed at the egress of an intranet.

Figure 9-1 Security protection at the enterprise border



In this scenario, the USG6000 series implements traditional network isolation, VPN communication, anti-DDoS, and traffic control and provides the following application-layer security protection functions:

- Identify, manage, control, and audit applications.
- Enable content security defense for corresponding services provided for external users. For example, enable file blocking and data filtering on the file server, mail filtering on the mail server, and antivirus and intrusion prevention on all servers.

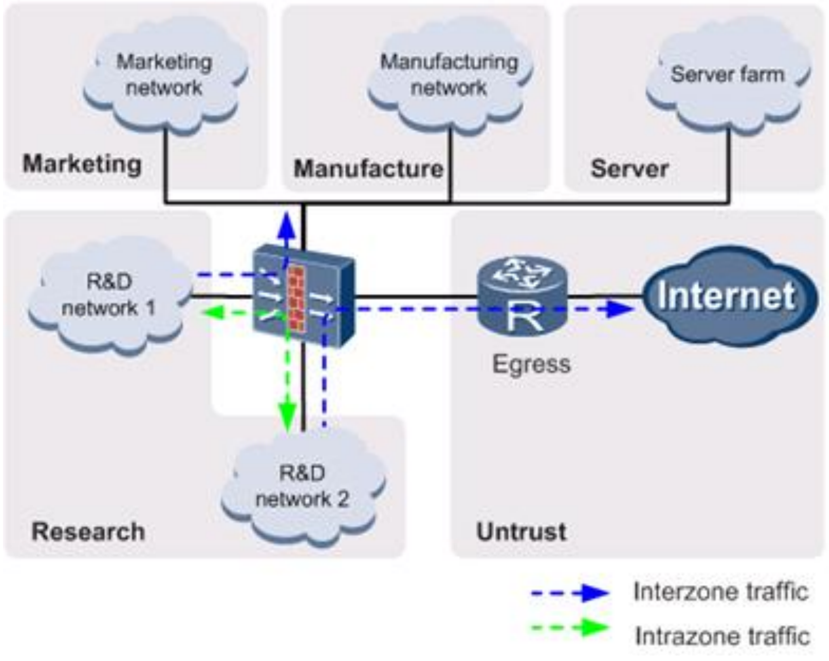
- Enable such functions as URL filtering, file type filtering, data filtering, antivirus, and application behavior control to defend against Internet threats and prevent information leaks to ensure network security.
- Record network operating logs to help administrators adjust configurations, identify risks, and audit traffic.

9.2 Intranet Control and Security Isolation

A large or medium-sized enterprise isolates networks by security level to ensure security. Figure 9-2 shows the typical deployment. For example, the USG6000 series isolates the R&D network, production network, and marketing network and monitor traffic among the networks to implement the following:

- Take different security policies for the service types and security risks of the networks.
- Control traffic among the networks to avoid information leaks.
- Isolate networks to prevent the spread of viruses.
- Divide networks to reduce detection load and improve detection efficiency for network connectivity because most traffic is generated within one network and the traffic within one network does not require much intervention.

Figure 9-2 Typical networking of intranet control and security isolation



In addition to user permission management and bandwidth control, the USG6000 series provides application-layer security protection as follows:

- The networks of the same security level are divided into the same security zone and a few security functions are configured. For example, R&D networks 1 and 2 belong to security zone **Research**, and the packet filtering, blacklist and whitelist, and antivirus functions can be applied between the two networks.

- The networks of different security levels are divided into different security zones and security functions are configured according to actual service requirements. For example, only some R&D hosts can access the marketing network, and the antivirus, file type filtering, and data filtering functions are applied between the R&D network and the marketing, production, and server networks.
- The intrusion prevention, antivirus, file type filtering, data filtering, URL filtering, and application behavior control functions are applied between the intranet security zones and the Internet.

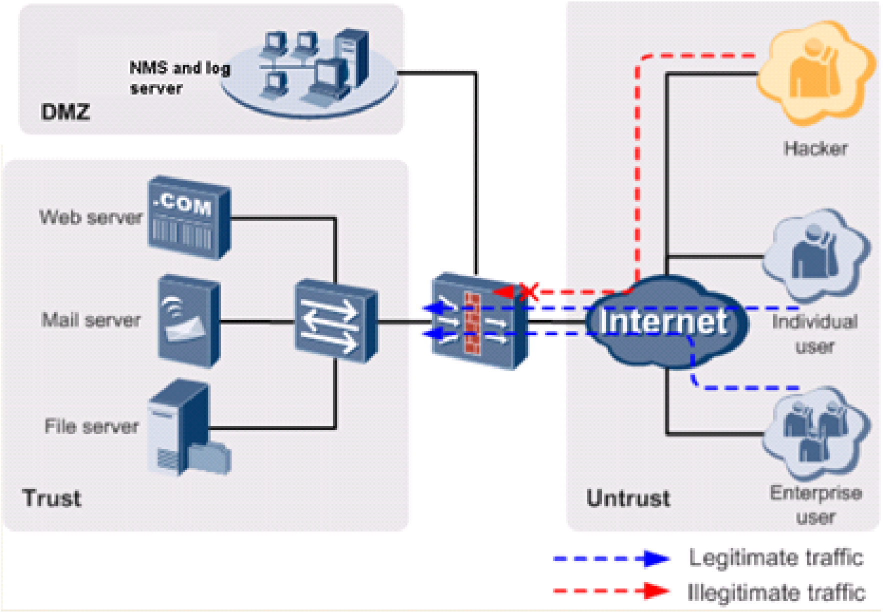
9.3 Border Protection for an IDC

The network structure of an IDC has the following features:

- Servers of the IDC must be protected.
- Servers from multiple vendors may be deployed in an IDC and are prone to attacks.
- The IDC provides network services for external users. To ensure extranet users' access to the servers in the IDC even when attacks are launched on the IDC, the border protection device must provide high performance and comprehensive reliability.
- The IDC traffic is complex. The administrator cannot adjust configurations effectively if the traffic is not unclear.

Figure 9-3 shows the scenario where the USG6000 series provides border protection for an IDC.

Figure 9-3 Typical networking of border protection for an IDC



In addition to traffic monitoring and network-layer anti-DDoS, the IAE provides the following security protection functions:

- Identify and control application traffic to avoid network congestion and service interruption.
- Enable intrusion prevention and antivirus to protect servers from viruses, Trojan horses, and worms.
- Enable file filtering and data filtering to prevent information leaks.
- Record network operating logs to help administrators adjust configurations, identify risks, and audit traffic.

9.4 VPN Remote Access and Mobile Working

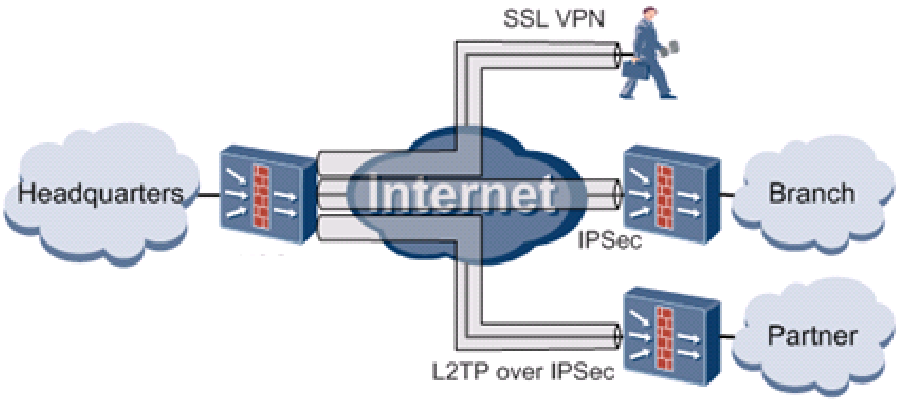
Nowadays, enterprises generally establish branches or cooperate with remote organizations around the world. Branches, partners, and employees on the move need to remotely access the headquarters. The secure and low-cost remote access and mobile working can be implemented through VPN technologies.

Remote access and mobile working have the following features:

- Branches need to access the headquarters network seamlessly and implement operations uninterruptedly.
- Partners must be flexibly authorized to limit the accessible network resources and transmittable data types according to the services.
- The locations, IP addresses, and access time of employees on the move are unfixed. In addition, employees on the move are not protected by information security measures. Strict access authentication must be implemented on employees on the move, and their accessible resources and permissions must be accurately controlled.
- Encryption protection must be implemented on data of remote access communications to prevent network eavesdropping, tampering, forgery, and replay as well as information leaks on the application and content planes.

Figure 9-4 shows the typical application scenario of VPN access.

Figure 9-4 Typical networking of VPN remote access and mobile working



In addition to network security services such as IPSec and SSL, the IAE provides the following application protection functions:

- The intrusion prevention, antivirus, file filtering, data filtering, and anti-DDoS functions are configured to prevent remote access users from introducing network threats as well as information leaks.
- User behavior audit is configured to discover risks in time for future tracking.