SI3000 Border Gateway

Application of SI3000 Border Gateway

Solving security and interconnection issues when using SIP

Iskratel, July 2012

Contents

1 Introduction 2

2 Interconnecting SIP UA 3

2.1 SIP/SDP Framework 3

2.1.1 Session Initiation Protocol (SIP) 3

2.1.2 Session Description Protocol (SDP) 3

2.2 Service Security and Realiability Problems 3

2.3 FW/NAT Problem 5

2.4 How to Find the Solution? 5

3 SI3000 Border Gateway (BGW) as a Security Network Element 7

3.1 Topology hiding 7

3.2 Call Admission Control (CAC) 7

3.3 DoS and DDoS Protection 7

3.3.1 Call duration monitoring 8

3.3.2 Handling Emergency Calls 8

3.4 Access lists (ACL) 8

3.4.1 Statically assigned ACLs 9

3.4.2 Dynamically assigned ACLs 9

3.5 Far-End NAT Traversal 9

3.6 SIP Upper-registration 10

4 Applications of SI3000 Border Gateway 11

4.1 General overview 11

4.2 Protecting Service Provider´s Network 12

4.3 Protecting Corporate´s Network 13

4.4 Connecting SI3000 BGW to IP Network 14

4.5 Implementation of Security policies 14

5 Conclusion 16

6 Abbreviations 17

1 Introduction

In the world there is growing interest in multimedia communication services which are offered on top of the Internet, using connection-less protocols. Signalling plays the key role when enabling these services. To implement desired multimedia services many protocols have to be used altogether:

- SIP prevailed as the core signalling protocol at the application-layer
- SDP for description and exchanging data about multimedia sessions
- RTP for delivering the content of the session
- UDP for exchanging datagrams
- IP to succesfully route messages with the content of communication through the Internet

Because all these protocols uses IP network as a delivery path between different users they are sensitive to all security problems in the IP network. Hence, how to provide integrity, authentication, authorization, confidentiality and service availability is the major task when deploying multimedia services in such environment. In addition, SIP and SDP are cleartext based protocols which facilitates attacks caused by malicious users (attackers).

The other problem when implementing multimedia services is, that service is divided into two stages:

- First - to find the user we want to communicate with
- Second - to establish a session where media (voice, video, fax, IM, presence status, …) is exchanged

Because these two stages are independent and application-layer signaling does not ensure reliable exchange of the media, the use of firewalls (FW) in the network can causes that users which previously successfully exchanged data about needed multimedia sessions cannot exchange media.

In this application paper we will show how Iskratel´s SI3000 Border Gateway can be used to solve presented problems. Our general interest will be SIP VoIP solutions.

2 Interconnecting SIP UA

2.1 *SIP/SDP Framework*

2.1.1 Session Initiation Protocol (SIP)

Typical VoIP solutions which rely on SIP includes:

- SIP protocol used as an application-level signaling protocol
- SIP User Agents (UAs) as a client which can generate or receive calls
- SIP registrar which registers UAs on the network, to enable location services
- SIP proxy server which forwards signaling traffic between UAs

We can understand SIP as a tool for users to find each other and to distribute information about sessions and for providing a way for setting up different kinds of multimedia sessions, including VoIP calls.

Because SIP was designed as a protocol operating at the application-level it is independent of the service of underlying transport protocols which it can use (UDP, TCP, SCTP).

2.1.2 Session Description Protocol (SDP)

In multimedia sessions, the object of communication (e.g. voice or video) is coded on the originating side of communication channel and then sent as a “media stream”, usually with the use of underlaying protocols (RTP over UDP) through the IP network interface. On the receiving side a media stream is received through the IP network interface and then decoded to restore and reproduce the original data. For coding and decoding original data different codecs (codec=COder/DECoder) are used regard different type of communication.

For successful exchange of the content of the multimedia session many information should be negotiated and agreed before a media stream can be sent:

- What kind of media will be transmitted and received
- How the media will be coded
- Where the media stream should be sent

SIP itself does not provide services to negotiate needed parameters about media stream, mentioned above. This is the purpose of Session Description Protocol (SDP):

- To find a common format for describing sessions
- To provides a protocol for delivering session descriptions among participants

SDP in is carried within the payload of SIP.

2.2 *Service Security and Realiability Problems*

In this chapter some well-known security issues with providing multimedia sessions will be exposed.

IP Network used as a Transport Network

Multimedia services uses IP networks as a transport channel and are sensitive to all typical security attacks in IP networks. Hence, usual security measures should be provided to prevent this kind of attacks (e.g. preventing UDP flood, ICMP flood, TCP SYN flood,…).

The use of Textual Protocols

SIP and SDP are entirely textual protocols and as such exposed to malicious attacks. If someone could get access to the content of the signaling messages, integrity and confidentiality of communication between participants in communication could be compromised. Measures for user authentication and authorization to use multimedia services of the participants in communication should be provided.

Message Tampering

SIP messages used in multimedia services have no built-in means to insure integrity. An attacker can intercepts and modifies packets exchanged between participants of a communication. This kind of attack can occur through registration hijacking, proxy impersonation or an attack on any component trusted to process SIP messages.

Registration Hijacking

When using SIP, a registration is normally performed using UDP which makes it easier to spoof requests. An attacker impersonates a valid UA to a registrar and replaces the legitimate registration with its own address. This attack causes all incoming calls to be sent to the UA registered by the attacker.

Proxy Impersonation

Proxy impersonation occurs when an attacker tricks one of SIP UAs or proxies into communicating with a rogue proxy. A rogue proxy inserts itself into the signaling stream usually done with DNS spoofing or ARP cache spoofing. Then an attacker has access to all SIP messages and in in complete control of the call.

Denial-of-Service (DoS)

An attacker can perform Denial-of-Service with simple flooding REGISTER or INVITE storm, sending malformed packets, manipulating SIP states. For example, flooding the firewall with BYE messages, possibly tearing down UDP ports opened for legitimate calls.

Session Tear Down

An attacker observes the signaling for a call, and then sends spoofed SIP BYE messages to the participating UAs. A side effect of this kind of attack is that the SIP proxy may not be aware of the calls being tear down and will not have proper call records.

2.3 FW/NAT Problem

VoIP signalling protocols and media stream are not able to traverse NATs and firewalls without a help.

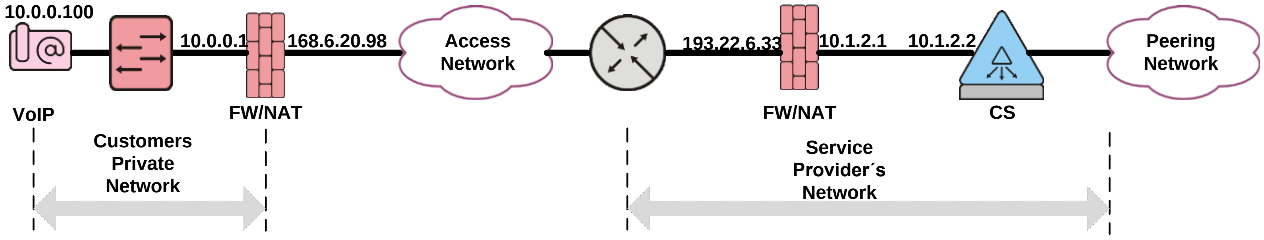

Figure 1: FW/NAT Traversal Problem

FW/NAT at the border of the customer´s private network blocks inbound call signalling. The CS in the SP network is not able to send a SIP INVITE request to the VoIP phone, because the IP phone is not addressable from the SP.

The FW/NAT also blocks incoming corresponding media stream. When the customer VoIP phone makes a call, it sends a SIP INVITE with an SDP body containing its local IP address, which is again not addressable from public network. Session parameters are negotiated in advance and when session parameters are agreed media can start. When traversing FW/NAT, parameters negotiated through SDP differs from parameters to reliable deliver datagrams used within transport protocols. Because FW/NAT does not know about negotiated media streams, media streams will be broken.

The issue how the media stream will be delivered should be solved when a session is setting up.

2.4 How to Find the Solution?

We examined major issues which should be solved when VoIP services are deployed. Here we will set requirements on which typical solutions are based:

- To enable traversing NATs:
  - “some-device” in the signaling path should understand and correctly interpret SIP handshake
  - “some-device” should take care about pinholes in FW/NATs devices – to tell FW/NAT which pinholes and when should be controlled
- To enable reliable service deployment of VoIP services DoS and DDoS attacks should be prevented:
  - Monitoring inbound and outbound SIP messages for application-level attacks should be performed – analaysing of SIP messages
  - Discarding of malicious packets should be performed
  - Monitoring for unusual calling patterns and rejecting calls with such patterns if needed
  - Performing granular Call Admission Controll (CAC)
- To enable integrity and confidentiality of communication:
  - Authentication and authorization of users should be performed
  - Encryption alghoritms and other standard-based security measures should be taken within signaling and media streams

To meet all these requirements full control over whole multimedia session has to be taken.

When we take into consideration all requests we see that we need new functions on existing elements which task is to assure security or new type of device in the network to be placed on the border between users and SP´s network is needed.

3 SI3000 Border Gateway (BGW) as a Security Network Element

The SI3000 Border Gateway (BGW) acts as a session border controller for SIP sessions in the “Access” scenarios. Typically, BGW is deployed at the edge of service provider´s network, or it can be used at the edge of an enterprise network, as a border element which provides secure interconnection for SIP services and a way to solving Far-End NAT traversal issues, initially implementing solutions together with SI3000 Call Server and SI3000 cCS. Security features which are provided by BGW will be explained in this chapter.

3.1 Topology hiding

Signalling messages convey information that can allow the recipient to determine both the internal topology of a network, and the route taken by a call across that network (and possibly out the other side). For example, the Via headers in SIP signalling messages carry this kind of information. It is undesirable to expose this information to users outside a network.

To solve this problem, BG removes sensitive information by rewriting the VoIP headers in the signalling messages that are sent across the network boundary. BG achieve this by acting as B2BUA (RFC3261).

3.2 Call Admission Control (CAC)

CAC is basic functionality offered by all BGW. With using this function BGW controls which calls may be signalled through the network, gracefully rejects calls when necessary. CAC allows BGW to guarantee and police SLAs. This serves to protect the CS in the SP’s network.

The BGW achieves this by:

- rate-limiting the calls that are set up through it, per subscriber and per group of subscribers, and also by rate-limiting or blacklisting calls
- tracking the bandwidth being consumed in the access network It rejects new calls from that subscriber if such calls would exceed the bandwidth limit set in their SLA.
- monitoring the total number of calls per group of subscribers, to prevent exceeding the limits set in their SLA.

Limits can be enforced at global scope (sum of total on BG) or per interface.

3.3 DoS and DDoS Protection

BGW protects SP network from malevolent endpoints. It monitors signalling and media traffic and dynamically detect potential attacks without disrupting the other services it provides. These attacks may then be blocked.

DoS and DDoS protection is complementary functionality than that provided by the CAC functionality, which gracefully limit/control the resources used by UACs. CAC policy configuration prevents UACs from extending their own service provision beyond the limits of their service agreements, whereas dynamic blacklisting performed by DoS/DDoS function of BGW prevents malicious endpoints from attempting to prevent provision of service to others.

For DoS/DDoS prevention, BGW uses dynamic blacklisting. To defend CS in operator´s network BGW uses rate-limitig:

- A continuous stream of signalling messages
- A continuous stream of badly-formed signalling messages
- A massive, continuous stream of media packets

If rates are exceeded, then the BGW can be configured to shut down all traffic received from outside the network for a period. Incidents regarding DoS/DDoS are written into log files and appropriate alarms are logged.

The shut-down runs for a configurable time period and then ceases, provided that rogue messages are not still being received. BGW can be configured, that shut-down of traffic from the user which have sent malicious traffic is performed permanently.

3.3.1 Call duration monitoring

BGW allows to set time limitation of the duration of calls. The administrator of BGW can configure a time limit and BGW monitors call duration and gracefully tears down calls that exceed the configured time limit. In this case it sends SIP BYEs to the endpoints of SIP call branches. By default, calls are not time limited.

3.3.2 Handling Emergency Calls

The BGW places emergency services calls on a permanent “whitelist” of calls that will always be permitted, except possibly from addresses or interfaces that have been shut down because of suspected DoS attacks. Emergency calls are processed ahead of non-emergency calls. BGW provides rate-limiting function for emergency calls, to prevent DoS attack caused by emergency calls. Operator can also admin, that in case of already blacklisted subscriber, emergency call will still be processed.

3.4 Access lists (ACL)

Firewall is the major part of BGW. It protects SP´s network and BGW applications from unwanted, malicious and heavy traffic.

To prevent unwanted traffic processing the uses statically and dynamically set of ACLs BGW (typically closes all traffic or certain types of traffic from a specific IP address or IP network).

By default FW part of BGW:

- Permits:
    - Protocols needed for management, maintenance and HA on dedicated interfaces
    - SIP signalling on dedicated interfaces
- All media flows are prevented

3.4.1 Statically assigned ACLs

For permanent security policies, the SP can apply statically assigned ACLs (“white lists” and “black lists”) on signalling and media part of BGW. For example:

- To prevent traffic flows form particular peer(s) or network(s)
- To prevent interfaces intended for management to be used for signalling or media traffic
- To put subscribers on blacklist or on whitelist

3.4.2 Dynamically assigned ACLs

BGW uses dynamic whitelisting/blacklisting regards to configured security and CAC/DoS policies for SIP signalling and Media proxy on BGW.

BGW dynamically sets ACLs to set up pinholes on local FW on BGW:

- For SIP signalling
- For RTP and RTCP flows
- To response when DoS/DDoS attack occurs

Pinholes in local BGW´s FW for media RTP and RTCP flows are assigned dynamically (per UA or per call basis). The BGW sets up and tears down media pinholes on FW part of BGW on a call-by-call basis according to per call SIP signalling information.

Pinholes for SIP signalling are assigned dynamically when CAC and DoS policies are appended on particular interfaces.

3.5 *Far-End NAT Traversal*

To solve the VoIP firewall/NAT traversal problem, the BGW replaces the provider’s FW/NAT, as shown:

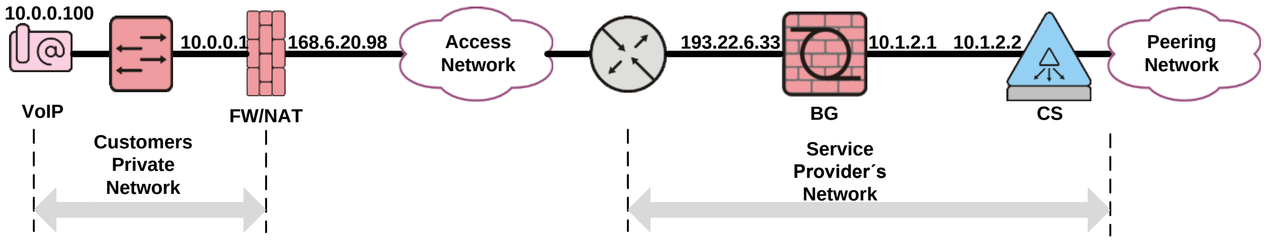

Figure 2: Far-End NAT Traversal

The signalling pinhole is created when the IP phone first comes online (REGISTER), and then kept open until the phone goes offline again. Media pinholes are created when the IP phone first sends a media packet on each established media session.

The pinholes for signalling and media have different lifetimes:

- The signalling pinhole, once created, is reused for all call signalling.
- The media pinhole is created anew for each media stream, because the source and destination ports of the media stream are dynamically allocated per call.

3.6 SIP Upper-registration

The BGW is, for security reasons, placed in the path between CS and UACs which wants to get services from CS. The BGW itself does not act as a SIP registrar. It tracks SIP REGISTER messages from SIP UAC when it tries to register itself to CS. For UAC, the BGW acts as a CS.

The BGW changes the REGISTER messages – it inserts its own data instead of UAC´s which was originally sent by the UAC. On the other hand then BGW acts as the UAC and tries to register on CS. The BGW then tracks respond form CS and forwards respond to UAC which tries to register. If the CS responds positively to registration messages and identified UAC as a valid user, then BGW allows UAC to access to SIP services on CS, otherwise BGW blocks further UAS requests´s.

The BGW tracks all currently registered subscribers, so that traffic to or from subscribers can be whitelisted during a DoS attack.

4 Applications of SI3000 Border Gateway

4.1 General overview

BGW is a demarcation point between SIP clients and CS. All SIP protocol traffic between clients and CS is routed through the BGW. Indeed, for SIP clients BGW acts as a CS. For security purposes, SIP client should never direct access to CS.

When terminating and re-originating SIP traffic, BGW acts as a “back-to-back” (B2BUA) user agent. BGW treats SIP requests that it receives as separate transactions from the requests that it propagates. Additionally, the “untrusted” side of a call is treated as a separate SIP dialog from the “trusted” side of the call.

The operator selects interfaces on which subscribers may registers and request services (ingress).

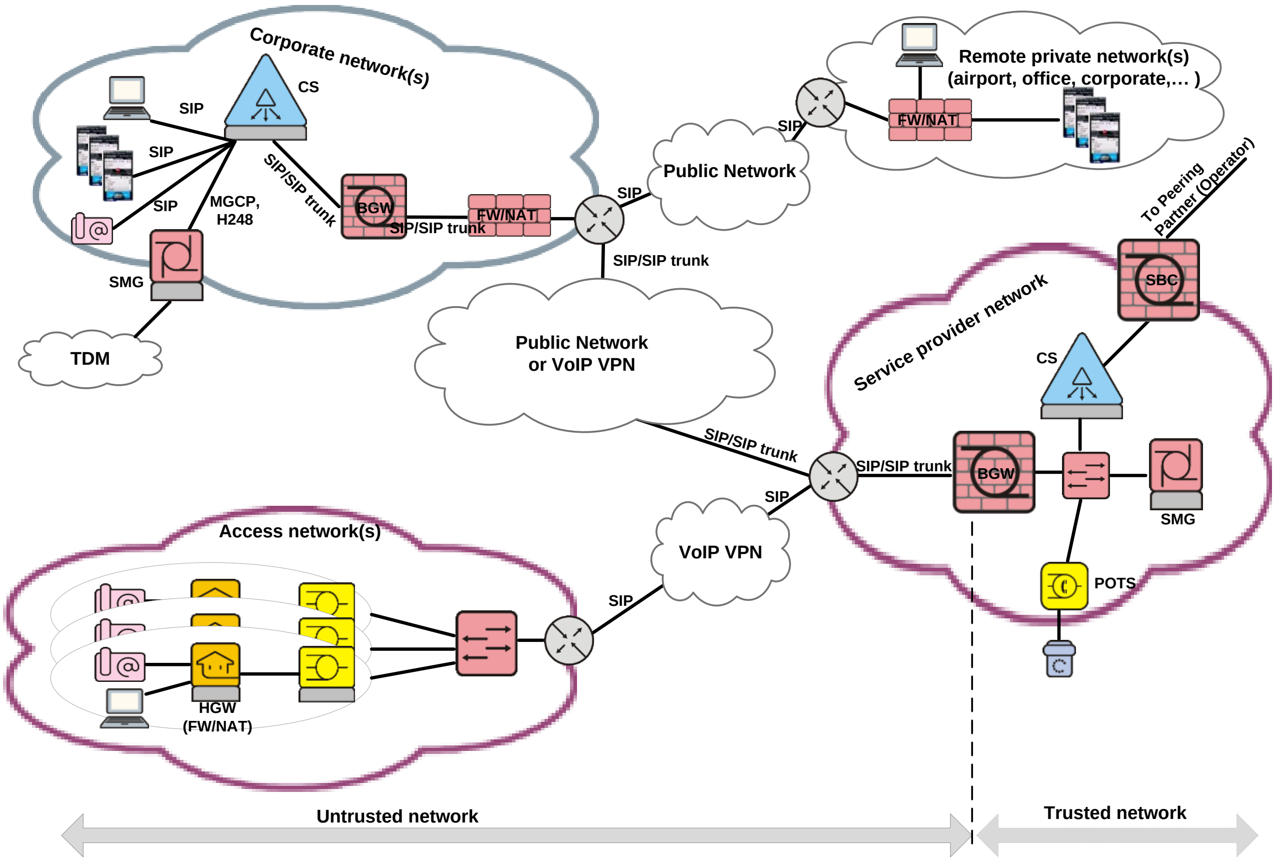

Figure 3: SI3000 Border Gateway as an Access SBC

For media flows, the BGW acts as a media proxy (RTP proxy), and so it terminates the media of a call on both the internal and external network sides. The media part (RTP proxy) of BGW supports proxying of Voice, Video and FAX RTP media packets.

The ports that it uses to send and receive media on each side are allocated dynamically when the call is established. The operator can set the range of RTP /RTCP ports, that should be used.

By default, the BGW always routes the media for calls that it handles through a Media part (also known as »media latching«). Signalling part of BGW rewrites the SDP in the SIP messages that it forwards, to ensure that the media of a call is routed through the appropriate BGW Media part.

The BG media part supports the following modes of operation (configurable):

- All RTP/RTCP streams must pass the RTP proxy (default configuration)
- RTP proxying is disabled (no one RTP/RTCP stream will be routed through proxy)

4.2 Protecting Service Provider´s Network

When used to protect SP´s network, SI3000 Border Gateway can be installed on different places in the network. Mostly we distinguish two models: central (Figure 4) and de-centralized model (Figure 5).

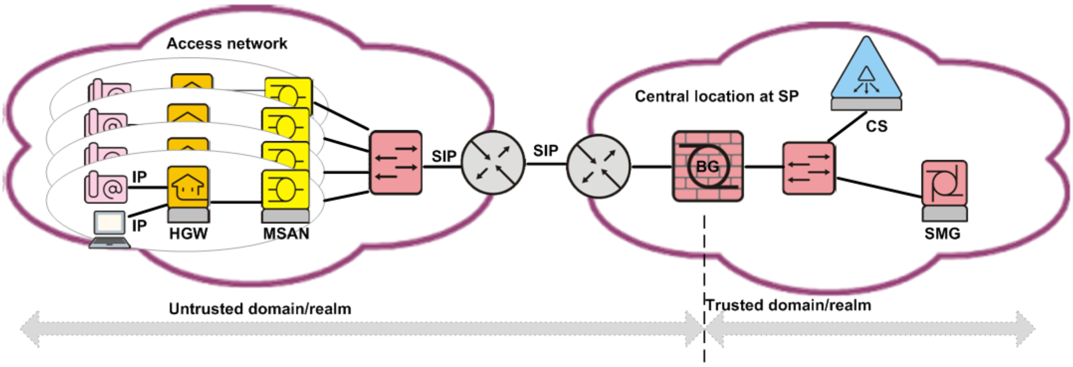

Figure 4: In SP´s Network – Central Location

When BGW is installed in central location it is placed together with CS on SP´s central premises. . In this case aggregation of subscribers traffic is done on central location and is much greater than in the other case when BGW is in installed at the border of aggregation network.

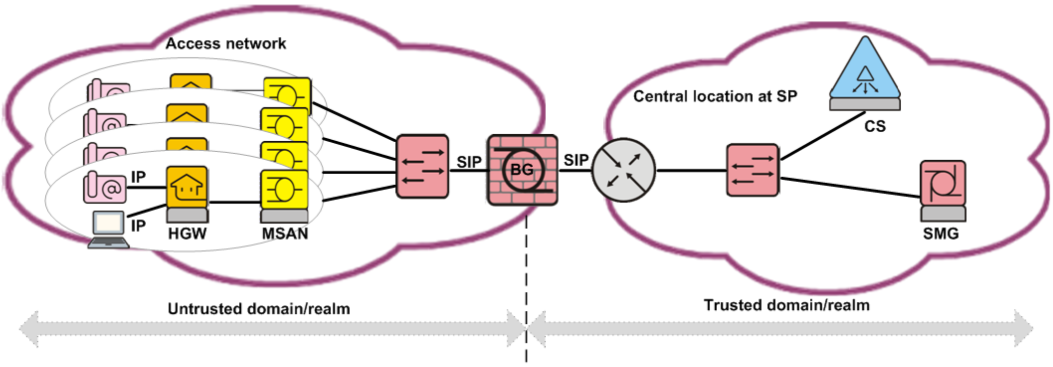

Figure 5: In SP´s Network – De-centralizied Model

In the first case, RTP traffic from all subscribers is routed to SP´s central location what can be the scenario which is not desired, particularly in cases where locations with subscribers are geographically wide apart from the central location. SI3000 Border Gateway should then be placed at the edge of the access network or even it is possible to install it in MSAN shelf together with broadband access network elements.

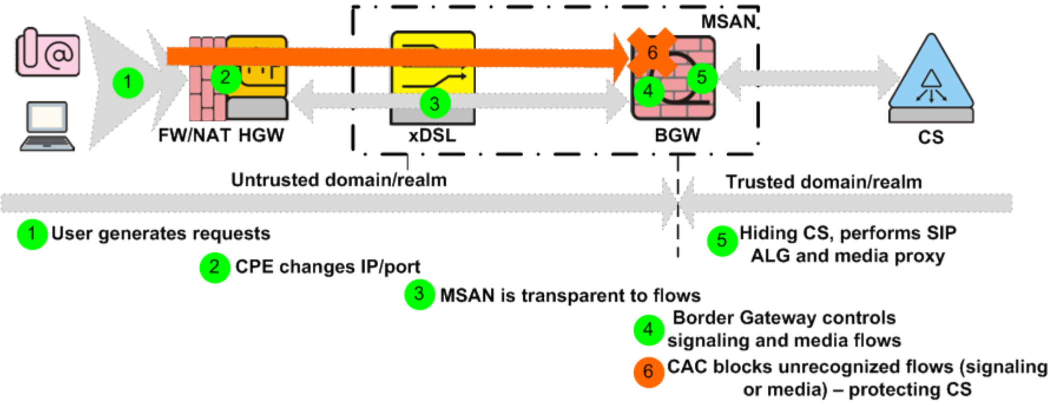
MSAN
FW/NAT HGW
xDSL
BGW
CS
1
2
3
4
5
6
Untrusted domain/realm
Trusted domain/realm
1 User generates requests
2 CPE changes IP/port
3 MSAN is transparent to flows
4 Border Gateway controls signaling and media flows
5 Hiding CS, performs SIP ALG and media proxy
6 CAC blocks unrecognized flows (signaling or media) – protecting CS

Figure 6: BGW installed in MSAN

4.3 *Protecting Corporate´s Network*

In corporate networks the SI3000 Border Gateway can be used to protect local corporate´s IP PBX (Figure 7).

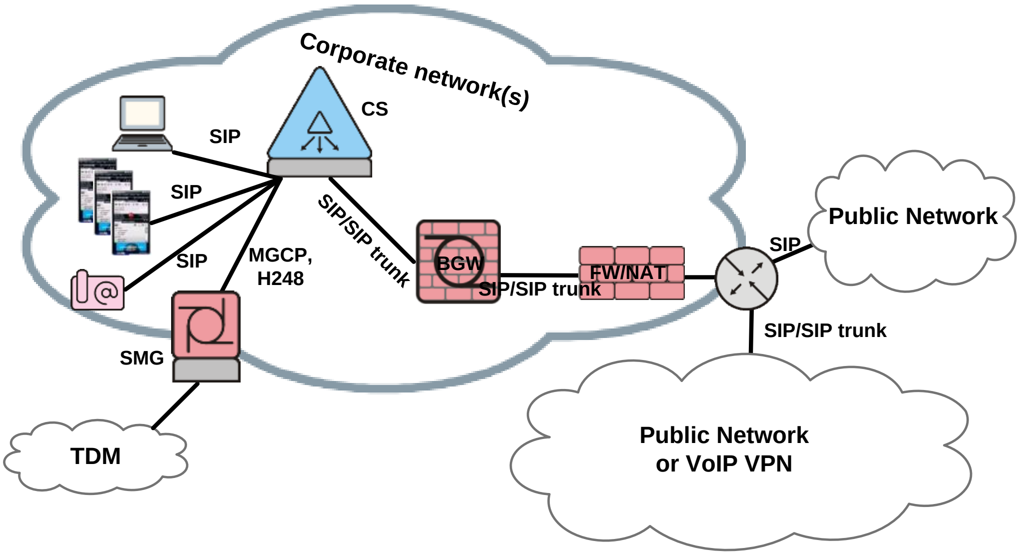

Figure 7: At the Edge of Corporate´s Network

In this scenario it has two functions:

- To protect corporate´s IP PBX from malicious traffic form the public network
- To protect corporate´s IP PBX from attacks inside corporate´s network

When IP PBX supports nomadic users to register on local IP PBX then it is exposed to the public network (Internet), because nomadic users can connect form “anywhere”. To follow best practices, dedicated interface on BGW should be used to allow this users to get desired services. It is also recommended to use FW at the border of corporate network.

Usually corporate uses SIP business trunk to connect to SP´s network and to aggregate all outgoing calls from its network. BGW can supervise the amount of traffic exchanged. The advantage of SI3000 Border Gateway is that it can be installed together with IP PBX in the same rack (Figure 8).

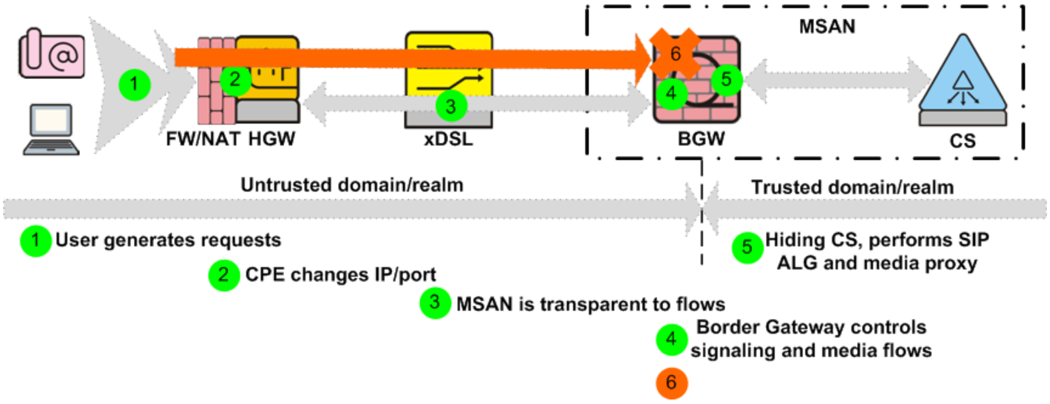
MSAN
1
2
3
4
5
6
FW/NAT HGW
xDSL
BGW
CS
Untrusted domain/realm
Trusted domain/realm
1 User generates requests
2 CPE changes IP/port
3 MSAN is transparent to flows
4 Border Gateway controls signaling and media flows
5 Hiding CS, performs SIP ALG and media proxy
6

Figure 8: BGW installed together with CS

It should be highlighted, that the installation shown above is probably the most often used case.

4.4 Connecting SI3000 BGW to IP Network

The BGW interconnects to IP network over multiple IPv4 interfaces. It supports multiple WAN and multiple LAN interfaces - in the minimum configuration with at least four interfaces (“VoIP WAN”, “VoIP LAN”, Management interface and HB - Heartbeat).

To ensure redundant network connectivity, at minimum two Ethernet ports with 1Gbps bandwidth are used.

To logical separate network interfaces, BGW uses VLANs (802.1Q). The operator can provide appropriate policies on each particular interface. On network elements (routers, switches) which provide BGW´s connectivity the network administrator has to apply policies which assure that unwanted crossing between VLANs is protected.

To ensure better security on L3, BGW doesn´t support IP routing protocols. It acts as a L3 network device with static routing used.

4.5 Implementation of Security policies

By default the SI3000 Border Gateway blocks all traffic coming from untrusted network. The administrator has to set rules how to process incoming traffic on BGW. If those rules are not set, all services for subscribers will not be available from the SP´s network.

Requests for service usually comes from different kind of subscribers which may be located in different networks. Therefore it is necessary to apply different security measures for different kind of subscribers and different services they request. To provide BGW´s administrator options to apply security measures per different groups of subscribers on BGW different groups of subscribers are associated into realms (Figure 9)

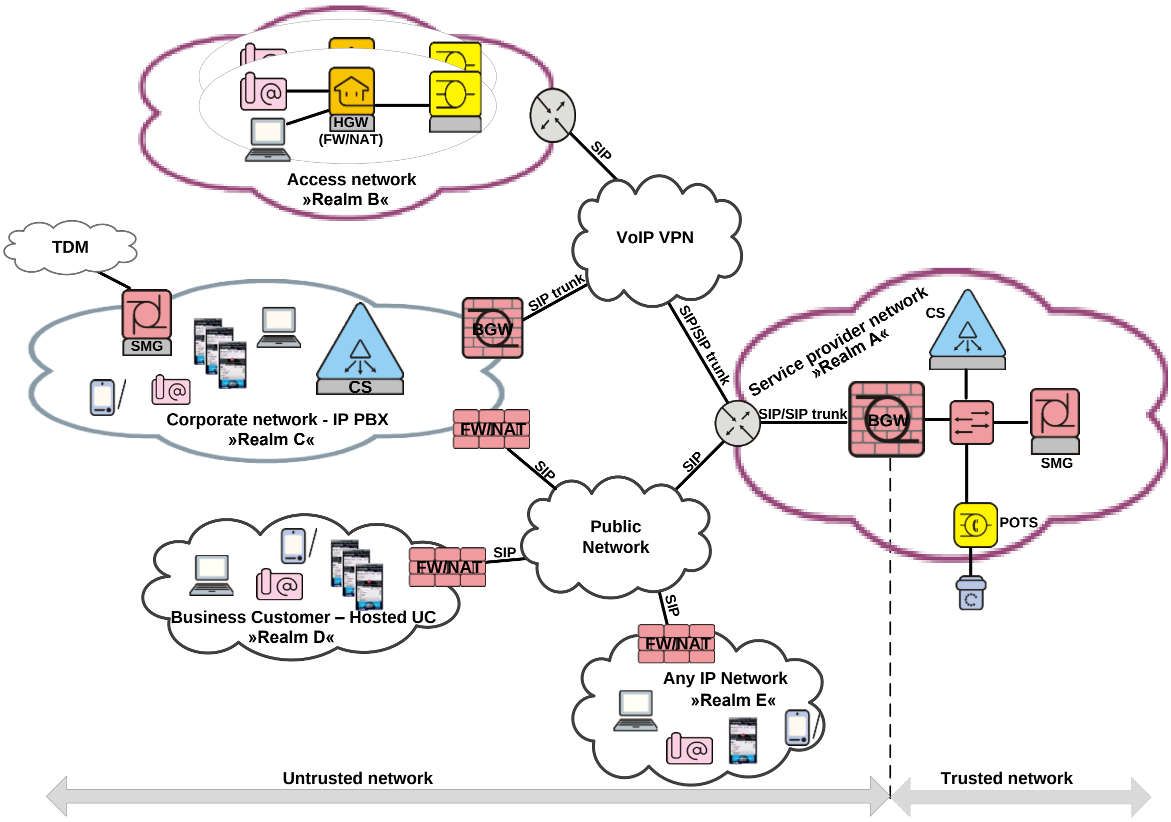

Figure 9: Subscriber Aggregation

A realm is an IP network address or an IP host address from which user´s traffic is allowed in accordance to policies applied for each particular realm. Typically, different realms are created for Access network(s), for Corporate network(s) and for Public network (Internet). Therefore a realm represents a remote peer or a remote peer-network.

The BGW´s administrator can create a realm for each corporate network or for each access network particularly. On this way, traffic from different networks can be separated and be processed in different manner. For each realm at least one SIP interface on BGW has to be created and for each SIP interface and for each Realm CAC and/or DoS policy has to be applied.

An BGW´s administrator creates CAC and DoS policies where it defines how ingress user traffic should be treated:

- Limiting the amount of bandwidth which can be used
- Limiting the amount of subscribers that are permitted to request services…
- Signaling rate limitation
- Far-End NAT traversal rules
- Etc…..

Then admin applies policies per each Realm or BGW´s SIP interface. With this action rules how to treat subscribers traffic from different networks is done and services form SP´s network should be enabled.

5 Conclusion

For successful deploy of multimedia services every service provider has to solve problems with reliable and secure service implementation. Using cleartext protocols, the lack of secure authentication and authorization, blocking data streams with the usage of firewalls are major shortages which have to be sloved.

The SI3000 Border Gateway with described feature set in this paper is a network device which enables a service provider to solve problems efficiently. It is designed to be installed at the border of the service provider´s network or at the border of the corporate´s network and is destined to be a part of Iskratel´s solutions together with other SI3000 products, especially with SI3000 CS and SI3000 cCS.

6 Abbreviations

ALG	Application Layer Gateway
ARP	Address Resolution Protocol
BGW	Border Gateway
CAC	Call Admission Controll
CS	Call Server
DNS	Domain Name System
DoS	Denial of Service
DDoS	Distributed DoS
FW	Firewall
IM	Instant Messaging
IP	Internet Protocol
MSCN	Multi service Control Node
NAT	Network Address Translation
RTP	Real Time Protocol
SBC	Session Border Controller
SCTP	Stream Control Transmission Protocol
SDP	Session Description Protocol
SIP	Session Initiation Protocol
SLA	Service Level Agreement
SMG	Signalling and Media Gateway
SP	Service Provider
TCP	Transmission Control Protocol
UA	User Agent
UAC	User Agent Client
UAS	User Agent Server
UDP	User Datagram Protocol
VoIP	Voice over IP

Iskratel d.o.o., Kranj

Ljubljanska c. 24a, SI 4000 Kranj, Slovenia
phone: +386 (0)4 207 2000, fax: +386 (0)4 207 2712

e-mail: info@iskratel.si
www.iskratel.com

Iskratel Electronics, Ljubljanska cesta 24a, SI 4000 Kranj, Slovenia, phone: +386 (0)4 207 34 96, fax: +386 (0)4 207 29 91, e-mail: info-ite@iskratel.si, www.iskratel-electronics.si
Iskrateling, Ljubljanska cesta 24a, SI 4000 Kranj, Slovenia, phone: +386 (0)4 207 62 76, fax: +386 (0)4 207 62 77, e-mail: info@iskrateling.si, www.iskrateling.com
Monis, Oktyabrskoy revolucii str. 99, UA – 61157 Harkov, Ukraine, phone: +380 577 15 80 00, fax: +380 577 15 80 16, e-mail: monis@monis.com.ua, www.monis.com.ua
Iskrauraltel, Komvuzovskaya str. 9a, 620137 Yekaterinburg, Russian Federation, phone: +7 343 210 69 51, fax: +7 343 341 52 40, e-mail: iut@iskrauraltel.ru, www.iskrauraltel.ru
Iskrabel, Harkovskaya str. 1/601, BY - 220073 Minsk, Belarus, phone: +375 17 213 03 36, fax: +375 17 251 74 59, e-mail: pihtin@iskrabel.by
Iskracom, Naurizbay batyra 17, office 213, 050004 Almaty, Kazakhstan, phone: +7 327 2917 166, fax: +7 327 2917 166, e-mail: a.nikonov@mail.ru
ITS Iskratel Skopje, Kej 13 Noemvri, Kula 4, 1000 Skopje, Macedonia, phone: +389 2 323 53 00, fax: +389 2 323 53 99, e-mail: info@its-sk.com.mk, www.its-sk.com.mk