Session Border Control in IMS

An analysis of the requirements for Session
Border Control in IMS networks

Jonathan Cumming
Director of VoIP Product Management
jonathan.cumming@dataconnection.com

Executive Summary

There is a lot of controversy and press coverage over both the role of Session Border Controllers (SBCs) and the design of the IP-Multimedia Subsystem (IMS). In this environment it is difficult to determine what are the real issues for each technology, let alone how they need to work together.

This white paper is aimed at equipment manufacturers looking at building SBC functionality into their product range to target the IMS market, and carriers and consultants looking to understand how an SBC fits into an IMS network.

It explains why IMS networks need session border control and what alternatives are available. It also looks at how these requirements are likely to evolve as services and access methods change, and discusses the function that products targeting this market require.

SBCs are described as both a cure-all for next generation telecommunications networks and an unnecessary attempt by carriers to stop their business becoming a simple bit-carrying commodity. This paper seeks to explain how these different views arise and the varied roles that SBCs play in IMS.

About the Author

Jonathan Cumming is Director of VoIP Product Management at Data Connection. During his 5 years at Data Connection, he has held a range of development, marketing and product management roles.

Jonathan has over 15 years' experience in the communications software industry. He holds an MBA from INSEAD and an Engineering degree from Cambridge University.

September 2005

Table of contents

1　Introduction

The IP-Multimedia Subsystem (IMS) defines the functional architecture for a managed IP-based network. It aims to provide a means for carriers to create an open, standards-based network that delivers integrated multimedia services to increase revenue, while also reducing network CapEx and OpEx.

IMS was originally designed for third-generation mobile phones, but it has already been extended to handle access from WiFi networks, and is continuing to be extended into an access-independent platform for service delivery, including broadband fixed-line access. It promises to provide seamless roaming between mobile, public WiFi and private networks for a wide range of services and devices.

This move, from a centrally managed network with control over the core and access networks to an open network with soft clients, represents a sea change in the applicability and deployment of IMS. Previously, it was aimed at centrally-managed networks with significant control over the core and access networks and the clients. Now it is moving to a much more open network model, where previous assumptions about the sorts of connecting networks and clients break down. This introduces the need for session border control at the network boundary to provide security, interoperability and monitoring.

This white paper examines these evolving requirements for IMS and where session border control fits in the IMS functional architecture. It also assesses the market for equipment targeting this space; both the evolution of existing equipment to handle these new requirements and the likely future evolution as the market and technology mature.

This chapter provides an overview of session border control and IMS.

Chapters 2 and 3 cover the requirements for session border control in IMS and how this function fits into the IMS architecture

Chapters 4, 5 and 6 discuss what products need to address these requirements, how this market is likely to change in the future, and what conclusions can be drawn from this.

Chapter 7 provides a list of references to additional information and a glossary of the acronyms used throughout this document.

Chapter 8 contains information on Data Connection and its products.

1.1 Overview of Session Border Control

Session border control is not a standardized set of functions. Instead, Session Border Controllers (SBCs) have evolved to address the wide range of issues that arise when voice and multimedia services are overlaid on IP infrastructure. These include

- security and prevention of service abuse to ensure Quality of Service (QoS)

- monitoring for regulatory and billing purposes

- maintaining privacy of carrier and user information

- resolution of VoIP protocol problems arising from the widespread use of firewalls and network address translation (NAT), and the vast array of differing protocols and dialects used in VoIP networks.

These issues are relevant for access to both carrier and enterprise networks, and on both User-Network Interfaces (UNI) to end users and access networks, and Network-Network Interfaces (NNI) to peer networks. The following diagram shows where SBC function is typically required.
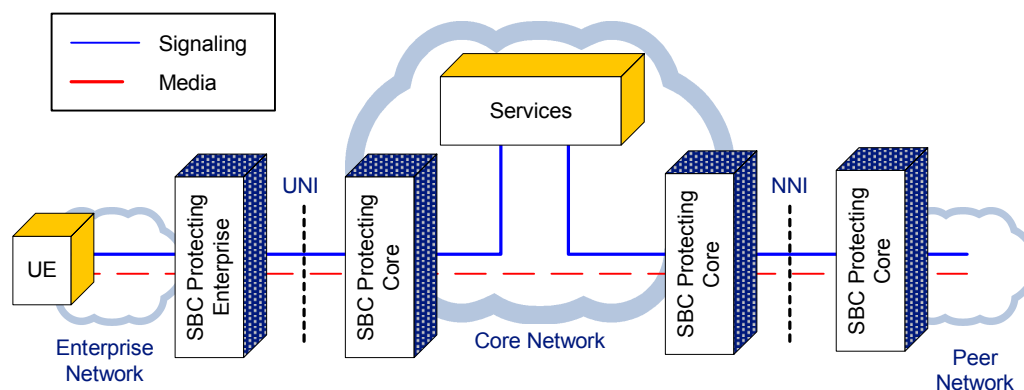
The diagram depicts a single device at the edge of each network (a traditional SBC), but there is actually great flexibility in how this function is distributed. For example,

- a device in the access network might perform initial user authentication

- an edge device might enforce access policy to limit Denial of Service (DoS) attacks and prevent bandwidth theft

- core devices might limit the total usage for a particular group of users and detect distributed DoS attacks.

The location of each function will depend on the overall system design, including the availability of processing resources and the level of trust between the different devices.

The following sections describe each of the SBC functions. For a more detailed description of Session Border Controllers, see our white paper: "Session Border Controllers: Enabling the VoIP revolution."

1.1.1 Security

An insecure network cannot charge for its use or provide a guaranteed QoS service, because unauthorized users cannot be prevented from overusing limited network resources.

SBCs can provide security and protection against

- unauthorized access into the trusted network

- invalid or malicious calls, including Denial of Service (DoS) attacks

- bandwidth theft by authorized users

- unusual network conditions, for example a major emergency.

Typical resources that require protection are bandwidth on access links and processing capacity on network servers. In general, core network links can be cheaply over-provisioned to help prevent them becoming bottlenecks.

To provide this security, the SBC

- identifies and authenticates each user and determines the priority of each call

- limits call rates and resource usage to prevent overloads

- authorizes each media flow and classifies and routes the data to ensure suitable QoS

- prevents unauthorized access for both signaling and media traffic.

QoS across the core of the network is normally handled by an aggregated classification mechanism, for example DiffServ, as this removes the overhead of reserving bandwidth for each individual flow.

The SBC may also be used to enforce QoS in the access network by signaling to the access routers or instructing the endpoint to reserve necessary resources across the access network. Alternatively, an intelligent access network may independently determine appropriate QoS for the media streams by analyzing the call signaling messages.

1.1.2 Monitoring

Network usage may need to be monitored for regulatory reasons (such as wiretapping and QoS monitoring), as well as commercial reasons (such as billing and theft-detection).

The monitoring devices need sufficient intelligence to understand the signaling and media protocols. They must also be located at a point through which all media and signaling flows.

SBCs fulfill both these requirements as all traffic passes through an SBC to enter the network. They provide a scalable, distributed solution to this processing-intensive function.

1.1.3 Maintaining Privacy

The following two types of information need to be protected.

- Information about the core network, which might provide commercially sensitive information to a competitor or details that could aid an attack.

- Information about a user that the user does not wish to be made public.

An SBC can be used to remove confidential information from messages before they leave the core network, including details of internal network topology and routing of signaling through the core network. It can also hide the real address of the user by acting as a relay for both the media and the signaling.

1.1.4 Resolution of VoIP Protocol Problems

SBCs can also act as gateways to heterogeneous networks; hiding any differences between the protocols used in the core and access networks. This can include the following.

- Hiding access network topology, including the complexity of routing through NATs and firewall and to overlapping address spaces of VPNs or private IP address spaces.

- Interworking between devices and networks of different capabilities (such as conversion between SIP and H.323 signaling, or between IPv4 and IPv6, or even different versions of H.323).

- Transcoding media flows between incompatible codecs.

Putting this function in the SBC, which is close to the access device, simplifies the core network devices by limiting the range of protocol variations that they must support.

1.2   Overview of IMS

IMS is the control plane of the 3rd Generation Partnership Project (3GPP) architecture for its next-generation telecommunications network. This architecture has been designed to enable operators to provide a wide range of real-time, packet-based services and to track their use in a way that allows both traditional time-based charging as well as packet and service-based charging.

IMS provides a framework for the deployment of both basic calling services and enhanced services, including

- multimedia messaging
- web integration
- presence-based services
- push-to-talk.

At the same time, it draws on the traditional telecommunications experience of

- guaranteed QoS
- flexible charging mechanisms (time-based, call-collect, premium rates)
- lawful intercept legislation compliance.

Network operators also hope that IMS will cut their CapEx and OpEx through the use of a converged IP backbone and the open IMS architecture.

- The IMS architecture defines many common components (for example, call control and configuration storage) so less development work is required to create a new service as this existing infrastructure can be reused.
- The use of standardized interfaces should increase competition between suppliers; preventing operators from being locked into a single supplier's proprietary interfaces.

As a result, IMS should enable new services to be rolled out more quickly and cheaply, compared with the traditional monolithic design of telephony services.

1.2.1   History and evolution

IMS was initially developed as a call control framework for packet-based services over 3G mobile networks as part of 3GPP Release 5 (2003). It was then extended to include WiFi roaming and additional services such as presence and instant messaging in Release 6 (2004/5).

Although originally designed for mobile networks, both ETSI TISPAN and the Multi-Service Switching Forum (MSF) have now also adopted the IMS architecture for their visions of fixed telecommunications networks. Discussions within these groups are driving the IMS extensions to cover fixed networks in 3GPP Release 7 (work-in-progress) and many of the session border control requirements that fixed network access introduces.

At this point, it should also be noted that the design of IMS Release 7 is not yet complete and there is ongoing disagreement over the scope and location of specific functions. However, although the names and details of the specification are likely to change, the principles and issues described in this document are unlikely to be significantly affected.

1.2.2   Drivers

Although originally developed for mobile operators, the main interest in IMS is from fixed line operators, as the existing fixed-line network is older and is due for replacement, whereas much of the mobile infrastructure has only recently been deployed.

In particular, the current generation of fixed telephone networks is limited to narrowband voice services and is at great risk of being displaced by mobile and Internet telephony services. An IMS-based network would enable fixed line operators to offer a much wider range of services, to help protect their market.

Despite the widespread industry support for IMS, many uncertainties remain over its value. The cost of providing such a QoS-enabled managed network is high compared with the Internet's stateless model. Also, as the success of Vonage and Skype and many other VoIP providers testifies, telephony services are easily provided over the public Internet and the quality is sufficient for many situations.

In order to justify the investment in IMS, the resulting service must be significantly better than that available over the Internet and people must be prepared to pay for it. Whether IMS is a commercial success will be determined over the coming years, but competition from Internet-based providers will make this a competitive market.

1.2.3   Architecture

IMS decomposes the networking infrastructure into separate functions with standardized interfaces between them. Each interface is specified as a "reference point", which defines both the protocol over the interface and the functions between which it operates. The standards do not mandate which functions should be co-located, as this depends on the scale of the application, and a single device may contain several functions.

The 3GPP architecture is split into three main planes or layers, each of which is described by a number of equivalent names: Service or Application Plane, Control or Signaling Plane, and User or Transport Plane.
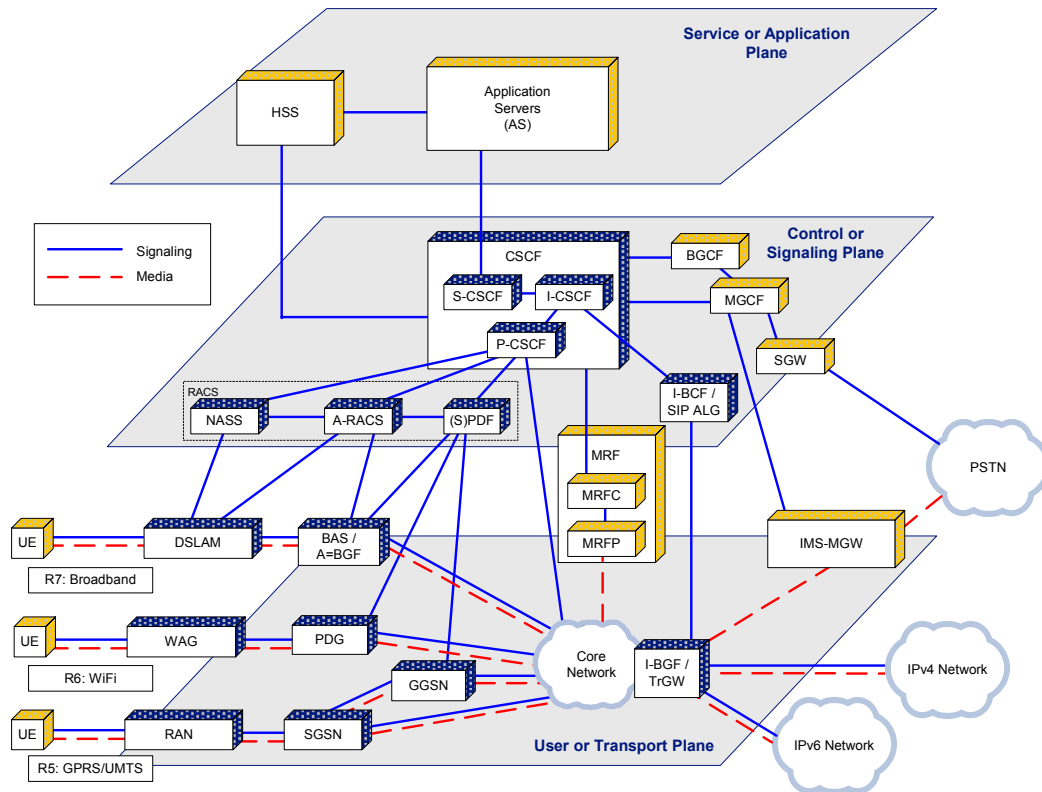
Application plane

The application plane provides an infrastructure for the provision and management of services, and defines standard interfaces to common functionality including

- configuration storage, identity management, user status (such as presence and location), which is held by the Home Subscriber Server (HSS)

- billing services, provided by a Charging Gateway Function (CGF) (not shown)

- control of voice and video calls and messaging, provided by the control plane.

Control plane

The control plane sits between the application and transport planes. It routes the call signaling, tells the transport plane what traffic to allow, and generates billing information for the use of the network.

At the core of this plane is the Call Session Control Function (CSCF), which comprises the following functions.

- The Proxy-CSCF (P-CSCF) is the first point of contact for users with the IMS. The P-CSCF is responsible for security of the messages between the network and the user and allocating resources for the media flows.

- The Interrogating-CSCF (I-CSCF) is the first point of contact from peered networks. The I-CSCF is responsible for querying the HSS to determine the S-CSCF for a user and may also hide the operator's topology from peer networks (Topology Hiding Inter-network Gateway, or THIG).

- The Serving-CSCF (S-CSCF) is the central brain. The S-CSCF is responsible for processing registrations to record the location of each user, user authentication, and call processing (including routing of calls to applications). The operation of the S-CSCF is controlled by policy stored in the HSS.

This distributed architecture provides an extremely flexible and scalable solution. For example, any of the CSCF functions can generate billing information for each operation.

The following diagram shows the routing of a typical call in an IMS environment and the two distinct uses of the NNI.

- Roaming (the left-hand NNI): This is required to access services provided by your own service provider (home network) when connected to another carrier's network (visited network).

- Interworking (the right-hand NNI): This is required when placing a call to a customer of a different carrier network.

The call signaling flows from the caller pass through the P-CSCF in the visited network to his home S-CSCF. The signaling then passes onto the called party via his S-CSCF.
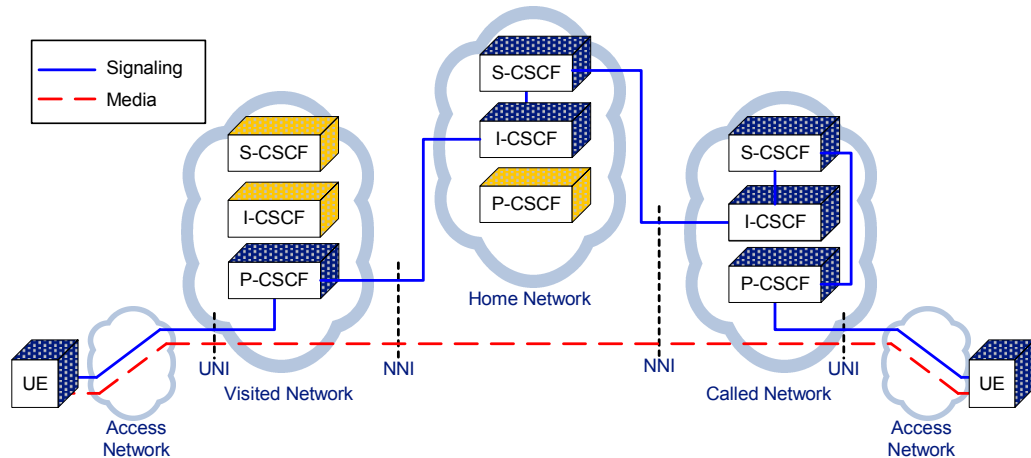
The Control Plane controls User Plane traffic through the Resource and Admission Control Subsystem (RACS). This consists of the Policy Decision Function (PDF), which implements local policy on resource usage, for example to prevent overload of particular access links, and Access-RAC Function (A-RACF), which controls QoS within the access network.

User plane

The User plane provides a core QoS-enabled IPv6 network with access from User Equipment (UE) over mobile, WiFi and broadband networks. This infrastructure is designed to provide a wide range of IP multimedia server-based and P2P services.
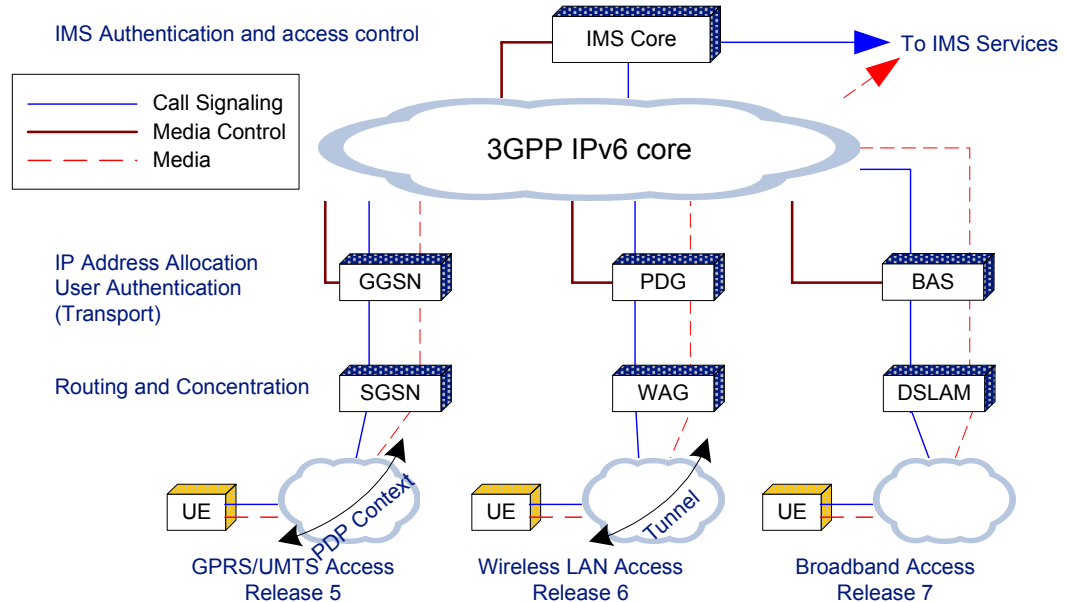
Although IPv6 is defined for this transport plane, many initial deployments are built upon existing IPv4 infrastructure and use private IPv4 addresses. This introduces NATs at the boundary of each address domain and the associated difficulties routing VoIP calls across the boundary.

Access into the core network is through Border Gateways (GGSN/PDG/BAS). These enforce policy provided by the IMS core: controlling traffic flows between the access and core networks, as follows.

- With GPRS/UMTS access, the GGSN authenticates the user equipment (UE) and controls the establishment of media channels using authenticated PDP contexts. This enforces QoS and access control through the access network to the UE.

- With Wireless LAN (WLAN) access, the Packet Data Gateway (PDG) controls the establishment of tunnels through the access network to the UE. These tunnels provide security of the message flows to the UE, but not QoS. Separately, the access network may apply QoS policy to data flowing to/from the carrier core and have a billing arrangement with the carrier to charge for use of its network.

- Release 7 adds support for IP connectivity over a range of access technologies. There is ongoing discussion over how much of this access will be covered by the core IMS specifications. For example, the ETSI TISPAN architecture envisages the IMS core connected to external networks through Border Gateways that are not part of the IMS specifications.

It is this change from a very controlled network with limited access methods in Releases 5 and 6, to a much wider range of access devices in Release 7, which introduces the need for Session Border Controllers.

2   IMS Requirements for SBC

Some of the functions provided by a Session Border Controller have always been important and inherent in IMS, given its role in providing a QoS-enabled service with detailed usage monitoring to enable charging for its use. Others are only now becoming important as Release 7 expands the range of supported access methods. This chapter looks at the requirements for this functionality and how they are changing.

The IMS architecture consists of interconnected core networks belonging to different carriers, with endpoints connected through attached access networks, and gateways to non-IMS networks. Border gateways control access into and out of each core network, monitoring and regulating the data flows on each interface.

This architecture is shown on the diagram below.
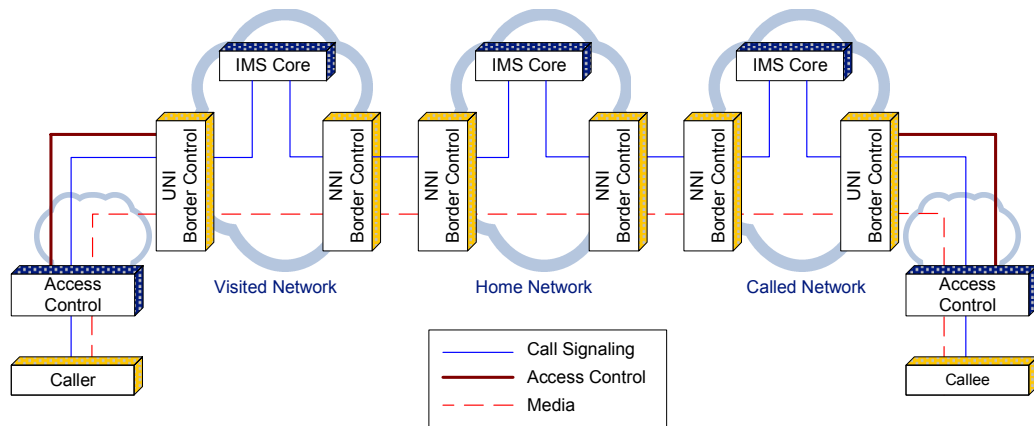
The core network needs to be protected against all of the threats described in section 1.1, but each interface imposes a different set of border control requirements due to differences in the attached devices and access networks.

The following sections describe the common requirements and the specific issues that each interface introduces.

2.1 Security

Security in IMS Releases 5 and 6 is designed around an open IPv6 core with well protected access.

- Access to the network core is protected using transport layer security on the UNI in the form of authenticated PDP contexts and tunnels.

- The NNI is an internal trusted interface within this secure core, so requires very little security.

However, IMS Release 7 and the reality of early IMS deployments have changed this model, and have expanded the range of security needed on each interface. Nevertheless over some interfaces, a subset of this functionality may be provided by the access network.

2.1.1 User-Network Interface (UNI)

The expanding range of access devices and reduced control over the access network has increased the responsibility on the border controller at the edge of the core network.

- In 3GPP IMS Release 5, only GPRS/UMTS access networks are supported. In this environment, the P-CSCF uses the GGSN to control access and bandwidth use through the entire access network all the way to the handset. Additional DoS controls on the signaling can be applied by the P-CSCF, but due to the controlled design and certification of handsets, there is limited scope for such attacks.

- The addition of WLAN access in 3GPP Release 6 does not greatly expand the range of protection required at the network border, as traffic into the core is controlled through tunnels managed by the PDG. In addition, Release 6 is primarily aimed at data roaming between 3G and WLAN, not the handling of voice calls over WLAN access, so there is limited function to protect.

- Release 7 expands both the range of supported access methods and the function, and as a result greatly expands the scope of attack. In addition, the core network now exerts little control over the access network, so the border gateway becomes its first line of defense.

2.1.2 Network-Network Interface (NNI)

Early IMS deployments have identified that security on the NNI is also required to protect the core network from malicious or unexpected behavior by a peer, and to prevent a problem in one network core affecting another.

2.2 Monitoring

Government regulations and commercial reasons both require monitoring of network use.

IMS Release 5 did not include monitoring on the NNI. However, reconciliation of inter-carrier charges, monitoring of service level agreements (SLAs), and lawful intercept of calls traversing the network, have all increased the need for monitoring on this interface.

2.3 Privacy

Privacy of both network topology and user information is required on all interfaces.

Again, network topology hiding was not considered in the design of IMS Release 5, but is considered a requirement for real deployments to protect this commercially-sensitive information from peers.

The requirements of Telco networks impose two aspects to privacy of user information.

- A caller may request for his identity to be hidden from the callee.
- The callers's identity must be available for emergency calls and lawful intercept regulations.

These requirements mean that user policy may modify the visible identification of the user to the callee, but that

- the signaling within the trusted core must continue to contain the true identity of the caller
- at the border of the trusted network, the true identity of the caller must be removed.

The border of the trusted network may be the edge of one carrier's core, or contain the core networks of several different carriers, depending on the regulations under which each operates.

2.4 VoIP Protocol Problems

VoIP protocol problems were not seen as an issue in IMS Releases 5 and 6. However the importance of this area has been raised by two factors:

- the use of IPv4 and other interoperability issues in early IMS trials
- the inclusion of NATs and a wider range of devices and network topologies in Release 7.

The scope of VoIP protocol problems seen depends on the interface and access method, as each has very different characteristics.

2.4.1 User-Network Interface (UNI)

The UNI border typically has to handle a large number of separate connections from individual users and a wide range of equipment, so it has to deal with a wide variety of protocol variants and network topologies.

It is not decided whether responsibility for NAT and firewall traversal issues is part of the IMS architecture, but functionality is required to enable interworking across such devices.

2.4.2 Network-Network Interface (NNI)

The NNI handles the signaling and media traffic between IMS carriers and through gateways to non-IMS carriers. A single interface typically handles a small number of high-volume connections with peer carriers.

The original IMS architecture envisaged a pure IPv6 IMS network core with minimal protection at the NNI boundaries. However, this model has changed due to the need to interoperate with non-IMS and pre-standard networks, such as IPv4 networks, and the requirement from carriers to protect their network core. The effect of this on the IMS architecture is to add an SBC on the NNI.

The NNI border controller may therefore provide the following.

- Interworking between signaling protocols, protocol variants and media codecs

- NAPT function.

Firewall and NAT traversal mechanisms are not required, as the peer carrier is expected to manage its own NAT/Firewall.

IPX Proxy

The GSM Association (GSMA) has identified the need for centralized interconnection of multiple carriers through an inter-carrier carrier that provides both IP connectivity and a clearinghouse for inter-carrier charges. This mimics the existing inter-GSM carrier (GRX) networks and removes the need for bi-lateral agreements between all interconnected carriers. This inter-carrier IP network is known as an IPX network.

In addition to simple connectivity, the IPX network provider can also provide a wide range of session border control functionality to its customers by providing an SBC within the IPX network. This SBC is known as an IPX proxy.

2.5   Summary

The following table summarizes the original SBC function defined in IMS Releases 5 and 6, and the new function introduced with Release 7 and early IMS deployments.

	Releases 5 and 6	**Additional requirements in Release 7 and Early IMS**
Security	Access network enforces access controls Peer networks are trusted	Border gateway enforces policy at all network boundaries
Privacy	Topology hiding not considered User privacy is handled by the P-CSCF	Topology hiding required. Media relayed to hide end-user location User privacy may be required at gateways to non-IMS networks
Monitoring	Monitoring of UNI for billing and lawful intercept	Monitoring of NNI to enforce inter-carrier agreements
VoIP Protocol problems		NAT to support IPv4 core with private addresses. NAT / firewall traversal on UNI Interoperability with devices with limited function

3 IMS architecture for SBC

As discussed earlier, session border control is inherent in the design of IMS. However, unlike architectures defined by other standards bodies, for example the Multi-service Switching Forum (MSF), the IMS architecture does not include a device labeled "SBC".

This chapter describes how SBC function fits onto the IMS-defined functional architecture, and how this architecture is evolving to handle the increasing requirements.

The IMS architecture defines separate sets of functions for each access type. However many of these functions perform similar roles in the network and each role provides the same subset of session border control function.
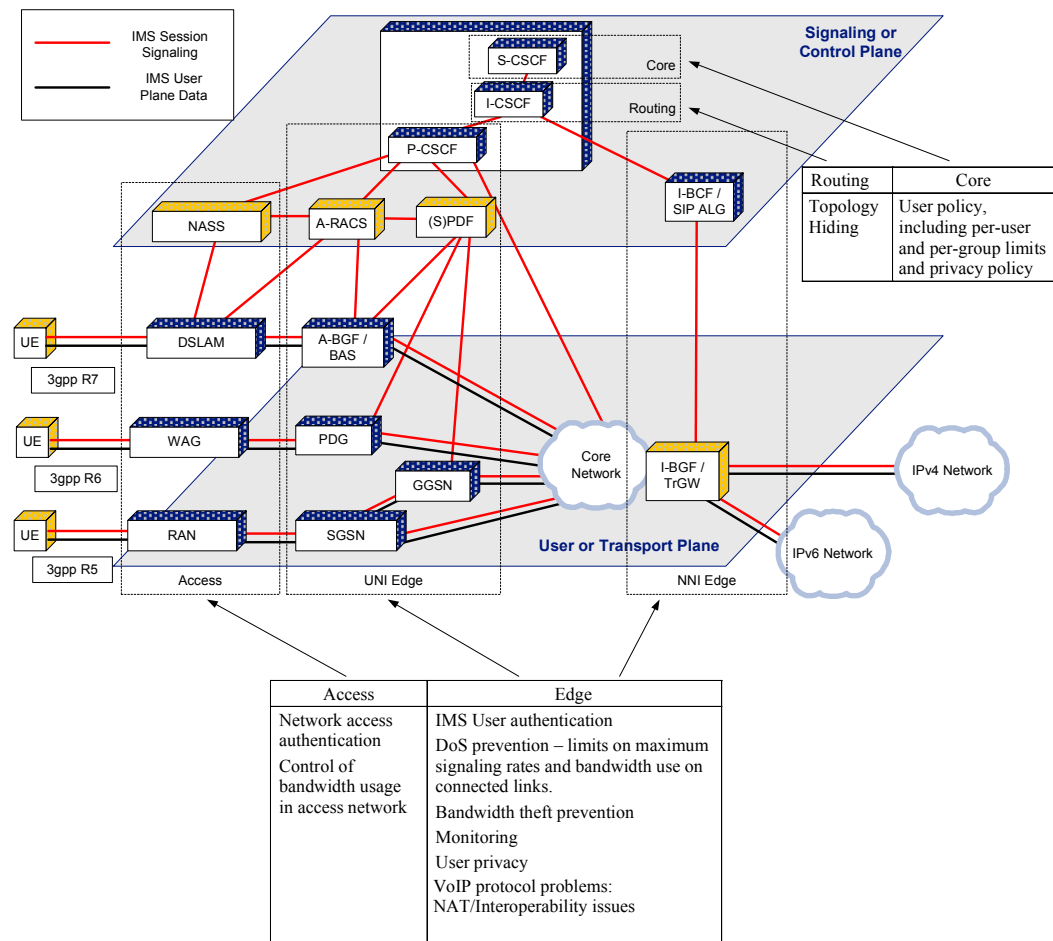

The following sections describe the differences between the function required on the UNI and NNI and the set of IMS functions that may be combined into an IMS-targeted SBC.

3.1   UNI

On the UNI, the set of IMS functions providing session border control depends on the access method. The following diagram shows how the IMS functions could be combined to build a single-box SBC for Release 6 and Release 7 network access.

The I-CSCF could also be part of the SBC in some situations.
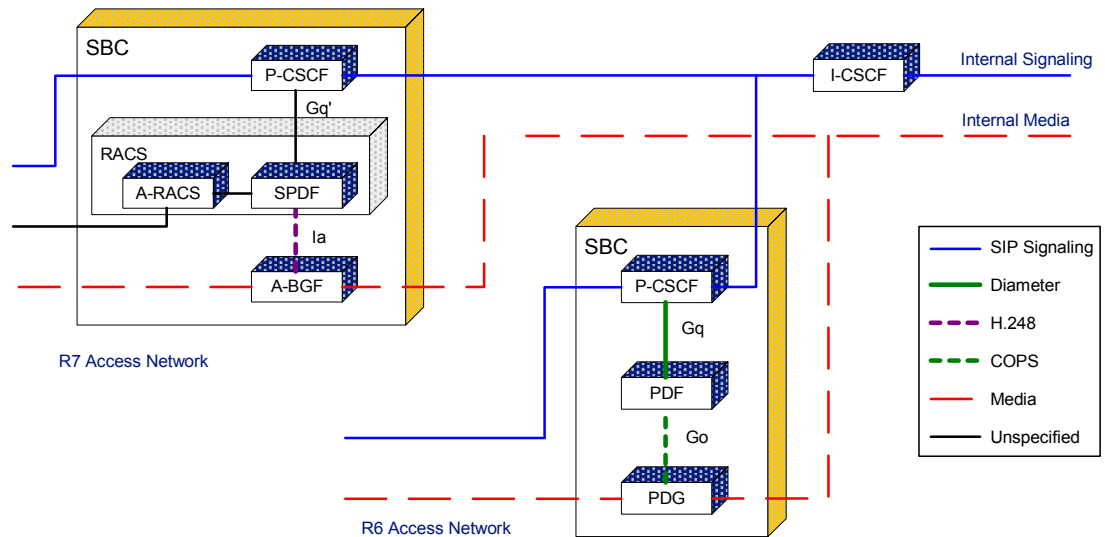
The SBC function is split between the following functions.

	R6 - GPRS/UMTS/WiFi	**Additional SBC features in R7**
P-CSCF	Controls security over the access network Tells the PDF what resources are required for the call	SIP ALG for IPv4 address translation and NAT firewall traversal
PDF / SPDF	Implements media resource allocation policy Authorizes media resource requests from BGF	Programs the A-BGF to accept media flows
A-RACS	Function incorporated into GGSN/PDG	Controls resources within the access network In IMS Release 7, the management of the access network is split out of the expanded PDF into the Access-Resource and Access Control Subsystem (A-RACS). The combination of the A-RACS, SPDF is known as the Resource and Access Control Subsystem (RACS).
A-BGF	Provides media relay for hiding endpoint address with managed pinholes to prevent bandwidth theft	Implements NAPT and NAT/Firewall traversal for media flows

3.2   NNI

The SBC-related functions within the NNI have a similar architecture to the UNI. This is shown in the following diagram. Again, the I-CSCF could also be part of the SBC in some situations.
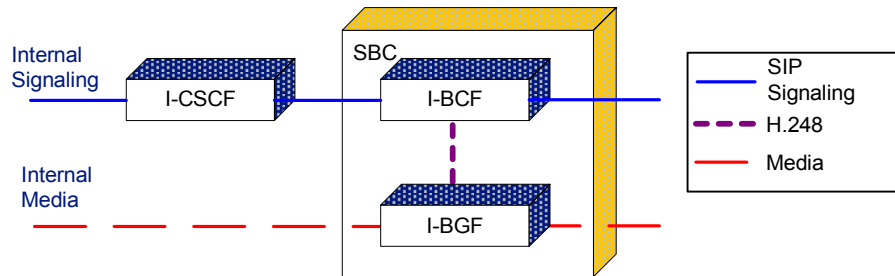
The I-BCF and BGF are new functions in IMS Release 7.

	SBC features provided
I-BCF	Transport-level security
	Tells the RACS what resources are required for the call
	NAPT function and control of NAPT in BGF
I-BGF	Media relay for hiding endpoint address
	Pinholes to prevent bandwidth theft
	NAPT and NAT/Firewall traversal for media flows

There is ongoing discussion whether the I-BCF and I-BGF function should be standardized within the IMS architecture. The ETSI TISPAN recommendation is that they remain outside the core specifications, providing a flexible gateway to other networks.

There is also discussion over the inclusion of a standardized RACS function between the I-BCF and I-BGF, as on the UNI, to mediate requests for media resources and manage local policy.

3.3   Reference Points

A "Reference Point" is a standardized interface between two IMS functions. It defines both the functions that it links and the protocol across the interface. Each reference point is denoted by the combination of one upper and one lower case letter, e.g. Gq.

The main reference points involved in session border control are described below.
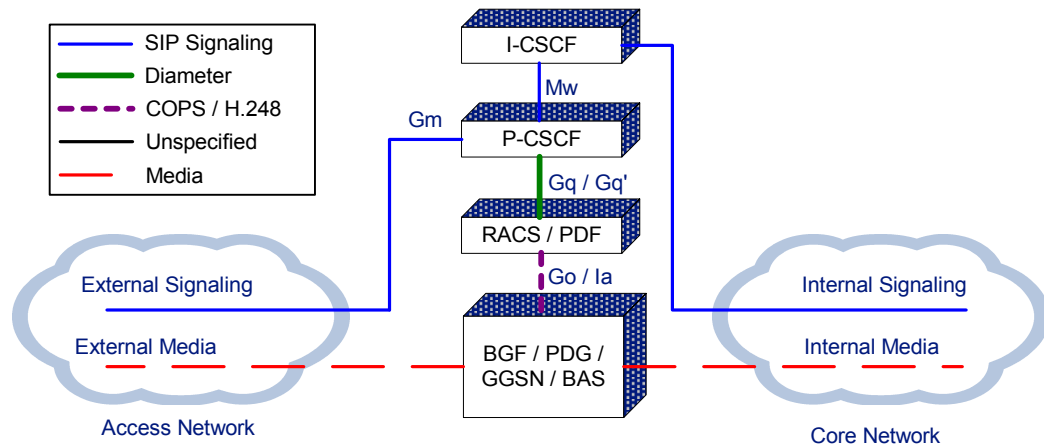
3.3.1 Call Signaling (Gm and Mw)

All call signaling in IMS uses the Session Initiation Protocol (SIP). For a comprehensive introduction to SIP, see our white paper: SIP Market Overview.

3.3.2 CSCF to PDF/RACS (Gq/Gq')

This reference point controls requests for network bandwidth from the IMS core. The original Gq interface in Release 6 enabled access control policy to be centralized in a separate Policy Decision Function (PDF). The Gq reference point is based on the Diameter protocol and enables the P-CSCF to request an authorization token from the PDF for access for a specified bandwidth.

This reference point is being extensively expanded in Release 7, to include the direct control of access. The Gq' reference point enables the P-CSCF to program the BGF to perform specific NAPT and NAT traversal function, as well as control the access network. This revised interface may be based on Diameter or H.248.

3.3.3 PDF/GGSN (Go)

In Releases 5 and 6, the Go reference point enables to GGSN/PDG to authenticate tokens received on the establishment of new media channels. This interface uses the COPS protocol.

3.3.4 PDF/BGF (Ia)

Release 7 expands the role of the Go reference point to enable direct control of the BGF to program NAPT and NAT traversal function and open pinholes in the gateway. The new Ia reference point is being defined for this purpose, probably based on H.248.

4 IMS SBC Products

The earlier chapters describe how session border control fits onto the functions defined by IMS. This chapter considers how this function is likely to be packaged into products and how these products will evolve from existing devices.

There are three different types of interface requiring control: UNI, NNI to peer IMS carriers and NNI to non-IMS carriers. Although a single device could be used for all applications, it is likely that separate products will be developed to target the functionality and scale of each application.

4.1 Scope of function

Most current SBC products are implemented as a single stand-alone device that is placed in front of existing equipment in the path of all the signaling and media traffic on an interface. This one box includes the media (BGF) and signaling processing (BCF/P-CSCF) as well as the media resource control (PDF/RACS).
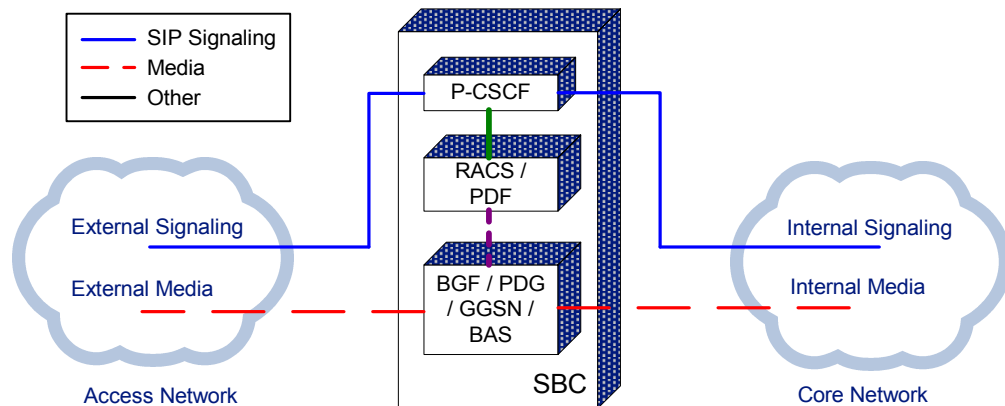
Although in some small-scale applications, it makes sense to include all these functions in a single device, this solution does not scale well to the requirements of service providers. For these applications, the media and signaling processing will often be split into separate devices, with the signaling processing centralized into regional server farms, and the media processing distributed closer to the user. This provides economies of scale on the signaling processing whilst maintaining direct media routing to minimize network transit delays.

The PDF/RACS is likely to remain colocated with the BCF/P-CSCF in most situations. However, the function might be split along this interface (Gq) to enable specialist RACS devices to be deployed to handle each access network with a common P-CSCF, or for the RACS to be used by non-IMS applications.

In addition, the PDF requires less processing than the P-CSCF, so a small number of PDF may be able to handle multiple P-CSCFs in highly-scaled installations.

4.2   Management and Control

One of the most challenging aspects of the number of different functions defined by the IMS architecture is managing the resulting proliferation of devices. User configuration is centralized in the HSS, but operator policy is implemented across many separate devices (P-CSCF, I-CSCF, S-CSCF, PDF, and others). There is currently no standardized way to control these devices, but an SNMP-based solution is likely to be the most effective way to centrally configure and monitor the system.

4.3   Product Evolution

The requirements of IMS Release 7 are bringing together three sets of products, each currently lacking the complete set of function:

- Products targeted at earlier IMS releases. These generally lack NAPT and NAT/firewall traversal function, and do not have sufficient call access control to prevent the wide range of DoS attacks that broadband devices can generate.

- SBCs targeted at the broadband access market. These may not be designed to fit the IMS architecture and generally do not support the IMS reference points and 3GPP-specific protocol extensions.

- IP and multi-service routers that have traditionally targeted the IP carrier interconnect and carrier edge markets. These products excel at high-performance routing, but need to add both the SBC function and IMS interfaces.

The short-term effect is that very different products are being promoted as fulfilling this same requirement, and the still-evolving standards are being pushed by each manufacturer to conform as closely as possible to its existing product range.

At the same time, there is huge pressure on the manufacturers to enhance their products to address the rapidly growing IMS SBC market. Depending on their situation, most are taking one of the following routes.

- Developing the function themselves, often incorporating much of the technology from independent suppliers such as Data Connection to reduce cost and improve time to market.

- Partnering with complimentary suppliers.

- Purchasing companies with the relevant expertise and then trying to merge the product lines, (for example, Juniper/Kagoor, Tekelec/IPtel, NetCentrex/NeoTIP).

Whichever method they choose, a rapidly expanding range of IMS-targeted devices incorporating SBC functionality will evolve to create an extremely competitive market. However, given the amount of investment by operators in their next generation networks, this will be extremely lucrative for those vendors that get it right.

5   The Future

Looking forward, there are two areas that we need to consider to predict the evolution and interaction of IMS and SBC products.

- Firstly, the future of IMS and its requirements for session border control.

- Secondly, the evolution of session border control function itself as VoIP technology matures.

The following sections look at each of these areas in more detail.

5.1   The Future of IMS

There is huge pressure on fixed-line operators to deploy a new architecture, and IMS offers the most attractive model available. However IMS has not yet been deployed outside trials, and the IMS standards continue to evolve to overcome flaws in the original design, as well as extensions to provide new functionality. As a result, it is likely that IMS in some form will be deployed, but that it will not be in the form currently envisaged in any of the defined releases. Instead, future networks will combine elements from each of the releases with new functionality to address new market opportunities.

The following factors will have a significant influence on the direction of IMS, and will change the nature and location of session border control within the IMS network and its derivatives.

5.1.1   Legislation

Government action to apply lawful intercept (wire tapping) and mandatory quality levels to telephony services may force all telephony service providers (including pure VoIP services) to deploy managed networks with border controls.

However, unless governments make it illegal to communicate over P2P VoIP services (as they have in China), the effect of this sort of legislation will be to increase the cost of providing a traditional telephony service and increase the use of less regulated P2P solutions.

5.1.2   Consumer pressure

Consumers will judge the value of IMS on whether the service that it provides surpasses that of alternatives. If it does, then the operator will be able to charge a premium for IMS services and reap the benefit from its investment in the IMS infrastructure. However, IMS-based operators will be at a cost disadvantage compared with operators that offer a pure IP connection without the expense and complexity of an IMS infrastructure, so they will not be able to compete on price on basic services.

Many factors influence consumers' choices, but important areas include

- reliability

- trust, including solutions to SPAM Telephony (SPIT) and SPAM Messaging (SPIM)

- convenience and simplicity

- cost.

The IMS architecture as it is currently designed is focused on an "operator knows best" model. The success of IMS and its evolution will depend on whether consumers agree with operators' choices or require a different set of features and restrictions.

5.1.3 Competitive pressure

The competitive landscape will be the primary driver for the introduction and success of IMS. If all the major carriers pursue the IMS model to prevent their service becoming a commodity, then there will be limited competition and pressure to encourage them to try more a radical business model.

However, this scenario is unlikely for number of reasons.

- One or more carriers will choose to offer an open IP network link at low cost. There are already examples of unlimited IP price plans from both fixed and mobile operators.

- If an uncompetitive market develops, anti-trust legislation will force operators to open up their networks to competitive carriers, who can provide a pure-IP service at marginal cost.

- The spread of alternative network providers (such as WiFi hotspots and urban WiFi networks) is increasing competition in the access network.

- Internet telephony is being promoted by well-known companies with deep pockets, including Ebay/Skype, Google and Microsoft.

5.2 SBC Function Evolution

The power of the Internet is the transparency of the network and the ability for services to evolve without the need to upgrade the network core. Many session border control functions break this transparency by requiring the SBC in the core to understand the media signaling. This both increases the cost of running the network and reduces the speed at which services can evolve, but provides additional security for the users.

Hopefully, some of the more intrusive SBC roles will diminish over time with the spread of IPv6 and VoIP friendly NATs, but others will remain to control and monitor access to the operator's network. The following sections discuss the likely evolution of each area.

5.2.1 Security

SBCs cannot support end-to-end security, as they need to be able to understand and modify the signaling messages. If users require higher security then they will use an encrypted P2P service across an open IP network instead. The primary driver for such end-to-end security is likely to be illegal; avoiding surveillance by the intelligence services, but it may also be used to prevent unauthorized surveillance, for example by an intermediate carrier or competitor.

The security model provided by IMS will not change – it will remain point-to-point. Users who require end-to-end security will use an alternative service, or encrypt the media to provide sufficient security for their requirements.

5.2.2 NAT and firewall traversal

Within the SOHO environment, the use of symmetric NATs[1] is likely to decrease, so the support of STUN and other NAT traversal techniques by endpoints will enable the NAT traversal technology in SBCs to be retired for many users. The use of STUN enables SIP to be used through all types of NAT except symmetric NATs.

However, in enterprise environments, SBCs will increasingly be installed at the edge of the corporate network to protect it from attack.

Corporate rules will enable NAT and firewall traversal according to corporate policy – this may limit the roll-out and availability of IMS services from within a corporate LAN. This is identical to the situation today with access to other Internet services, e.g. email and web browsing, from within a corporate LAN, and should not be subverted by the carriers.

5.2.3 IPv4 and NATs

The IMS network, particularly when IPv4 and VPN issues are included, is not an open transparent network, so SBC function is required to enable multimedia services to work. End-point NAT traversal technologies such as STUN remove the need for traversal devices at the UNI, but do not provide an end-to-end solution for media traffic that needs to traverse multiple private address spaces.

The use of IPv6 or a single global IPv4 address space throughout the network core would enable media to be routed directly between endpoints without the need for SBC function on the NNI. In addition to reducing the processing required, this would also ensure that the media takes the most direct route to its destination.

However, the current generation of core networks is based on IPv4 and many early IMS deployment will run over these networks. It will be a number of years before NATs on the NNI can be removed.

---

[1] Symmetric NATs set up separate mappings between the private IP address and port, and the public IP address and port, for each remote address. As a result, an endpoint cannot use STUN to determine a public address that a third-party can use to route media to it through a symmetric NAT. Instead, a media relay must be used.

5.2.4 QoS

QoS across core networks is already extremely good, and the Internet has been shown to be capable of handling serious disruptions to its infrastructure without significant effect on its performance.

However, QoS in access networks remains a challenge. Bandwidth availability within access networks will continue to increase, but new services will evolve to use any extra capacity, so QoS mechanisms within access networks will continue to be required. The design of these mechanisms will depend on the specific access medium: in some cases an SBC will be required to enforce the rules at the network boundary, but in others it will be possible to negotiate access end-to-end.

It is certain that differential handling of different classes of traffic over the access network will increase, however it is not clear that this will require session border control at the core network end of the link to enforce policy. A more flexible solution would be to allow the endpoint to determine the class of service to be applied to each stream using an out-of-band mechanism.

6 Conclusion

Session border control is fundamental to the IMS proposition to both operators and consumers. The exact function required will evolve as the underlying infrastructure and customers' demands change, but SBCs will always be part of any IMS solution.

There is short-term pressure for manufacturers to enhance their existing SBC and IMS products to enable them to be used as part of a Release 7 solution. This will be a competitive area with products being developed by many manufacturers, but is also an area that requires an unusually wide breadth of expertise, so will challenge the expertise of many contenders.

Longer term, the evolution of session border control in IMS networks will depend on the ability of IMS-based networks to compete with lower-cost solutions available over the Internet, and of operators to charge enough for the QoS and security that they offer.

With or without IMS, SBCs will continue to provide protection at the boundaries between managed networks. The evolution of the next generation of Telco networks will just determine where and how transparent those boundaries are.

7 Further Information

7.1 Sources

GSM Association	http://www.gsmworld.com
IETF	http://www.ietf.org
SIP Forum	http://www.sipforum.org
3rd Generation Partnership Project	http://www.3gpp.org
SIP Working Group	http://www.softarmor.com/sipwg
SIPPING Working Group	http://www.softarmor.com/sipping
ETSI TISPAN	http://portal.etsi.org/tispan
3GPP2	http://www.3gpp2.org

7.2 Reference material

IP Multimedia Subsystem (IMS)

3rd Generation Partnership Project	TS 23.228 v7.0.0	Technical Specification Group Services and System Aspects: IP Multimedia Subsystem (IMS); Stage 2 (Release 7)
GSM Association	IR.65	IMS Roaming and Interworking Guidelines

SBC

Data Connection white paper	Session Border Controllers: Enabling the VoIP Revolution

SIP

Data Connection white paper	SIP Market Overview	
IETF	RFC 3261	Session Initiation Protocol (SIP)
IETF	RFC 3489	STUN - Simple Traversal of User Datagram Protocol (UDP) Through Network Address Translators (NATs)

7.3   Glossary of Acronyms

3GPP	Third Generation Partnership Program
ALG	Application Level Gateway
BAS	Broadband Access Server
BCF	Border Control Function
BGCF	Breakout Gateway Control Function
BGF	Border Gateway Function
CSCF	Call Session Control Function
COPS	Common Open Policy Service
DoS	Denial of Service
GGSN	Gateway GPRS Support Node
I-CSCF	Interrogating CSCF
IMS	IP Multimedia Subsystem
NAPT	Network Address and Port Translation
NASS	Network Attach Subsystem
NAT	Network Address Translation
NNI	Network-Network Interface
P-CSCF	Proxy CSCF
P2P	Peer-to-Peer
PDF	Policy Decision Function
PDG	Packet Data Gateway
PLMN	Public Land Mobile Network
PoC	Push-to-Talk over Cellular
QoS	Quality of Service
RACS	Resource and Access Control Subsystem
RAN	Radio Access Network
S-CSCF	Serving CSCF
SBC	Session Border Controller
SGSN	Serving GPRS Support Node
SIP	Session Initiation Protocol
SLA	Service Level Agreement
SOHO	Small Office Home Office
STUN	Simple Traversal of UDP through NATs
THIG	Topology Hiding Inter-network Gateway
TrGW	Transition Gateway
UE	User Equipment
UNI	User-Network Interface
VLAN	Virtual Local Area Network
VPN	Virtual Private Network
WAG	Wireless Access Gateway
WLAN	Wireless Local Area Network

8   About Data Connection (DCL)

Data Connection (DCL) is the leading independent developer and supplier of software-based solutions for OEMs and Service Providers, and Softswitch solutions for the telecommunications industry. Founded in 1981 and headquartered in London UK, with US offices in Alameda, CA, Boxborough, MA, Dallas, TX, and Reston, VA, Data Connection has increased revenue and profitability each of the last 23 years. Privately held, revenue and profit for fiscal year 2003/2004 were $52M and $15M respectively. Customers include BT, Cisco, IBM, Lucent, Microsoft and Verizon

Data Connection has consistently grown its customer base by carefully managing the development and deployment of open, standard-based solutions for telecom's most challenging problems. These solutions greatly reduce OEM/Service Provider time to market, cost, and project risk while increasing revenues from end users.

The unique and well-established business model that is the cultural foundation of Data Connection has allowed us to develop and retain the top engineering talent which provides our customers with solutions of the highest quality that are backed by world-class support and corporate financial stability.

For more information on Data Connection please see www.dataconnection.com.

8.1   Data Connection Network Protocols

The Network Protocols division of Data Connection provides portable source code products for OEMs.

- VoIP software solutions include SIP, MGCP, Megaco/H.248, and Session Border Controller.

- Control plane solutions for packet and optical routing solutions include BGP, OSPF, ISIS, RIP, PIM, IGMP, MPLS, LMP, DDRP, VPN, and ATM.

8.2   Data Connection Internet Applications

The Internet Applications division of Data Connection is a leading developer and supplier of email, voicemail and unified messaging solutions, as well as audio, video and data conferencing. Its MailNGen and MeetingServer platforms are widely deployed among tier one service providers across North America and Europe. These are in turn supported by its reliable and scalable directory, DC-Directory, which is used in some of the largest deployments in the world today.

8.3   Data Connection Enterprise Connectivity

The Enterprise Connectivity division of Data Connection first started developing SNA products in 1983, and is now the world-leading vendor of SNA and SNA/IP technology. Our products provide full feature pure SNA and all the major techniques for integrating SNA systems into evolving IP networks. Customers include Cisco, IBM and HP.

8.4   MetaSwitch

The MetaSwitch division of Data Connection is the industry's leading vendor of Class 4/5 softswitch and enhanced applications solutions for packet and TDM networks. Its widely deployed call agent, media/signaling gateway and application server platform supports over 100 Class 5 features including CLASS services, IP Centrex, Unified Messaging, E911, LNP, 1-800 and CALEA, and scales from a few hundred to half a million subscribers in both integrated and distributed configurations. Customers include incumbent and competitive local exchange carriers, as well as operators of broadband wireless, cable and fiber networks.

For more information on MetaSwitch please see www.metaswitch.com.

8.5   About DC-SBC

Drawing from technology and experience from multiple divisions of Data Connection the Network Protocols division has developed a fully portable Session Border Controller (DC-SBC) software solution designed specifically for system vendors. Data Connection's extensive VoIP and IP routing heritage provides OEMs with a field-hardened SBC solution that is deployable immediately, delivering dramatic cost and time to market savings. Like all Data Connection products, DC-SBC is developed with the highest quality standards and is architected for unparalleled modularity, flexibility, scale, and reliability.

For more information on DC-SBC please see www.dataconnection.com/sbc.

Data Connection is a trademark of Data Connection Limited and Data Connection Corporation. All other trademarks and registered trademarks are the property of their respective owners.