UK Interconnect White Paper

Newport networks

UK Interconnect White Paper

Introduction

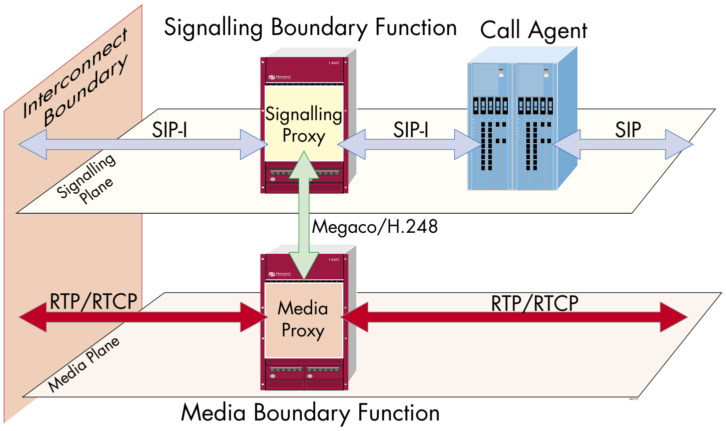
The UK will probably have the first regulated Inter-Service Provider VoIP interconnect anywhere in the world. The standards for this interconnect are being created under the auspices of the UK Network Interoperability Consultative Committee (NICC) and is the result of the combined efforts of interested parties in the main Service Providers, both mobile and fixed, and vendors that operate within the UK.

The boundary conditions that have been set for the service offered at the interconnect are intended to mirror those available for PSTN service within the UK. This is quite natural since the service being provided by the IP system is a PSTN replacement service.

Consequently, one of the key objectives is to have service restored within half a second of a failure in the interconnect. This means that not only does a resilient multiple node architecture have to be used, but also the IP protocols and networking technologies have to be specially configured to meet this stringent requirement.

Router components in the interconnect network are avoided, with interconnection being performed at layer 2, so that questions of ownership of IP layer products and systems are easier to resolve. In the longer run, the evolution of the interconnect specification will take in VLAN technology which is beneficial in economic and management terms.

In the tradition of standards bodies, the definition of the interconnect takes on the form of distributed logical functions, so that no particular vendor is favoured. However, most of the border functions correspond closely to the functionality provided by session border controllers. The interconnect is divided into signalling interconnect and media interconnect and thus requires SBC to be able to do the same.

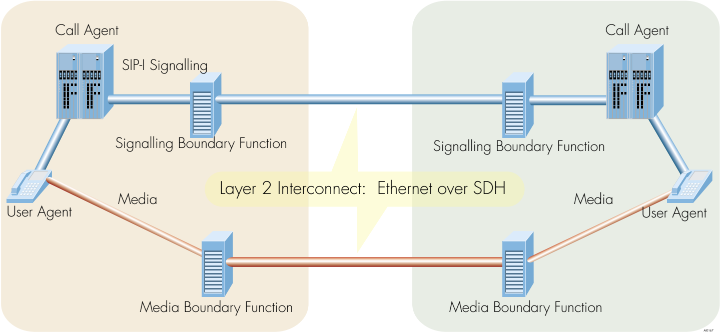


The protocol used is ISUP over SIP (SIP-I), as defined in the ITU document Q.1912.5 (Mode C), with some detailed profiling to meet the needs of the existing UK SS7 based service model.

Since the service is a PSTN simulation service, many of the issues are resolved through reference to the UK implementation of ISUP, and indeed many of the contributors to the interconnect definition documents are from this background.

Networking Models

The basic network model that can be used to realise the interconnect is shown below. At the boundary of the network are signalling and media firewall functions, designed to protect the Service Provider's infrastructure, while offering the required quality of service levels.

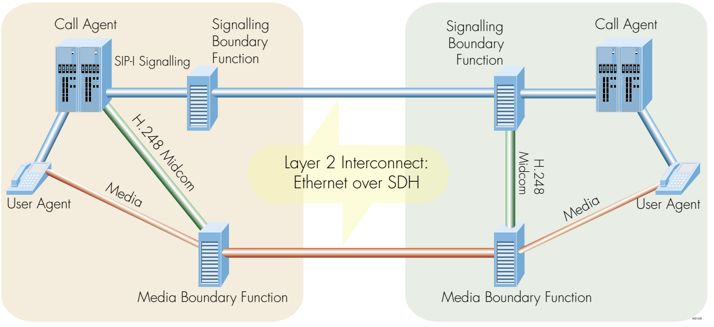


In the notional schematic above, the User Agent may be an IP client, a media gateway or indeed a normal PSTN phone connected through the media gateway. Their signalling (MGCP/Megaco) will be translated by a Call Agent to SIP-I, but the media will flow directly from the IP client / media gateway device to the media border function.

However, this diagram, while showing the border functions, has a missing functional link, that is the link between either the Call Agent or Signalling Boundary Function to the Media Boundary Function. This link is necessary to open media pinholes for the calls to flow through the Media Boundary Function.

Traditional firewall functions for media are impractical because of the huge and variable number of sources of media and the fact that there are incoming calls. Any traditional firewall would have to be opened to virtually any incoming traffic. Thus, the Service Providers network would be open to their peers. This would be unacceptable to Service Providers.

There are two possible network architectures that can be used, as shown below in the two sides of the schematic:



Option 1 Call Agent Control

Option 2 Signalling Border Function Control

In the first option, shown on the left hand side above, the Call Agent controls the Media Boundary function using a Midcom protocol, such as the ETSI standardised H.248 package. In this option, the Signalling Boundary function could be a normal firewall. However, the Call Agent has to present a publicly routable IP address to the peer Service Providers and there is no protection at the signalling layer within the Service Providers network. Access to other parts of the Service Providers infrastructure is prevented by the firewall.

More importantly in the context of this discussion document, it requires a H.248 control connection from the Call Agent to the Media Boundary Function at the network edge. Today, this protocol technology has not yet been developed, so this option is not readily available for early adopters.

In Option 2, the same protocol is used from a back-to-back User Agent (B2BUA) implementation in the Signalling Boundary Function to control the Media Boundary Function. Fortunately, this has been developed and is readily available in the 1460 session border controller, and so can be implemented by early adopters.

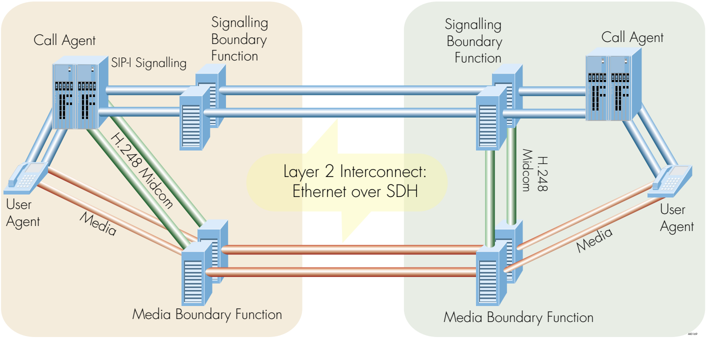
The Newport 1460 SIP SignallingProxy is aligned with the Signalling Boundary Function and the 1460 MediaProxy is the Media Boundary Function.

The 1460 product offers and easy evolution path between the two architectures allowing for cost effective deployment and growth. Additionally, the 1460 SignallingProxy can be deployed as a SIP signalling firewall, further protecting and securing the Service Provider network.

Creating Resilient Interconnects

In this diagram, both the Signalling and Media Boundary Functions are duplicated to provide a resilient interconnect. The Call Agents decide which interconnect function to use depending on load and whether or not the particular boundary function is operational. It is also possible, in Option 2 on the right hand side of the diagram below, that a Signalling Boundary Function could control many Media Boundary Functions within a Service Provider's network (not shown here).

Resilient Interconnect Scenario



Option 1 Call Agent Control

Option 2 Signalling Border Function Control

Signalling Protocols

The signalling protocol chosen by the NICC is SIP–I in the interconnect space. The Call Agent or the Signalling Boundary Function inspects the protocol and creates a pinhole to open in the interconnect for media flows as appropriate using the Megaco Gate Control Protocol package. The media pinhole is open for the duration of the call.

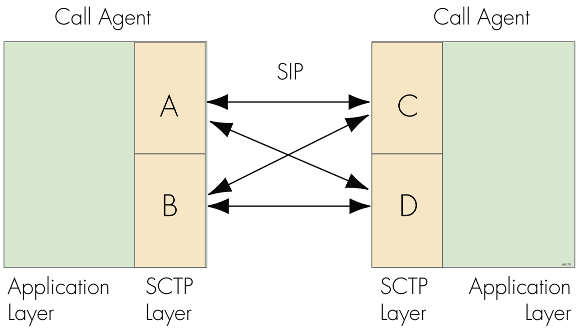
SIP-I can be carried over UDP or TCP, with SCTP having also been proposed within the NICC, as a feature rich alternative.

At present it looks like UDP or TCP will be chosen by the NICC initially, albeit that SCTP would be a superior choice. This is because of the pressure to deploy combined with the current lack of availability of SCTP capable SIP devices.

SCTP has the benefit of being capable of supporting the required 500 millisecond failure detection times and also supports a layer 4 based multi-homing capability that is intended to offer easy to use standby switchover in the case of failure. TCP and UDP, on the other hand will be much slower to respond without careful configuration and do not support multi-homing at the protocol level. They are, however, much more widely used and available.

While the multi-homing capability within networks is seen as quite useful in Service Provider interconnect applications, it does not deliver the commercial benefits that can be delivered by a least cost routing scheme that supports alternate routes at the application layer (e.g. the Softswitch). There is a school of thought that the choosing of an alternate path should be done by the Softswitch/ Call Agent and that the choice of secondary path should be made on grounds of cost.

In the SIP-I environment, TCP has few advantages over UDP. If the message size exceeds 1300 bytes, by convention, either TCP is used or the message is fragmented. In UDP environments fragmentation is regarded as being unreliable due to the possibility of out of sequence messages. However, SIP will discover bad messages and a repeat transmission of the message will be made. So unless there are a lot of bad messages above 1300 bytes in length that are discarded, there is little downside to using UDP. It is not foreseen that the SIP-I messages in the interconnect space will ever exceed 1300 bytes, so the choice of TCP or UDP may in come cases come down to the capability of the source / destination Softswitch.

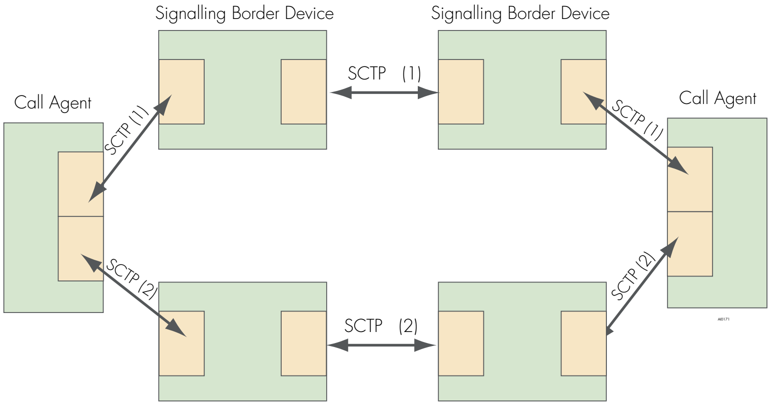


In SCTP, there is a multi-homing option which allows an SCTP source to define two locations from where it can transmit and receive data. These can be used to provide very rapid circuit protection at the IP transport layer. In the diagram above, both ends of the connection, (the Call Agents) have chosen to present primary and secondary addresses.

The SCTP layer in each Call Agent decides to send packets from and to either interface as it pleases. While normally there is the concept of primary and secondary, some SCTP implementations send alternate packets from primary and secondary as shown above.

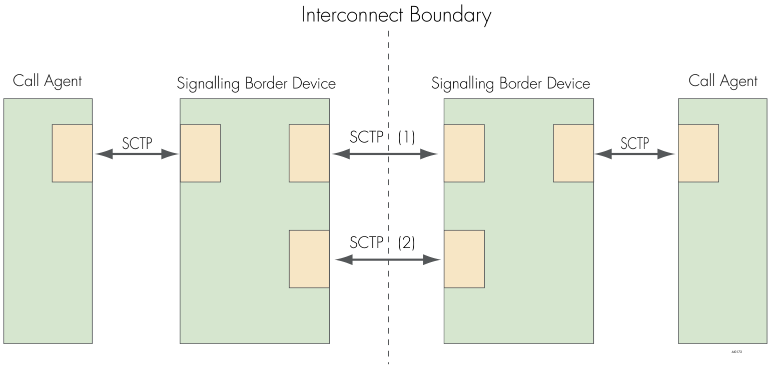
This introduces some difficulties for Session Border Controllers. The SBC terminates the transport layer, re-assembles the SIP messages, creates media pinholes and then re-forms the SIP messages with changed SDP and source details and sends them on.

To do this the SBC has to be able to see all multi-homed streams, and of course in the distributed SBC scenario this is not possible as shown below. The two multi homed SCTP streams don't come together in the border devices.



So a workable architecture, if SCTP is used, is to have the SCTP terminated and then resourced as a multi-homed stream across the interconnect as illustrated below.

In this case the media path can be successfully controlled because the Signalling Border devices can reassemble at the Transport Layer and can thus understand the SIP signalling.



Separate Signalling and Media

In the general model, media and signalling boundary functions can be provided in either separate or combined devices. However, while a loss of signalling can be easily detected (if a little slowly when using TCP as the transport protocol), the loss of media due to some fault in the interconnect is less easy to determine. So when the link fails between two peer Service Providers, the callers will know before the Switching System that the media connection has gone and will clear down. However, when the caller tries to re-establish the call, because the switching system has no knowledge of the failure it will still try to use the same resources to connect the media, and the call will probably fail.

What is required is for the Call Agent to know, as rapidly as possible, that the connectivity has failed using techniques such as O&AM capabilities in layer 2 interconnect. Prior to the widespread availability of suitable standards, such as Bidirectional Forwarding Detection (BFD-which is still at the IETF draft stage - draft-ietf-bfd-base-03.txt July 2005), the best approach is, in the short term, to pass media and signalling down the same link so that when the link fails the signalling fails also and the Call Agent can take appropriate actions to avoid the failure.

As the standards evolve to support comprehensive link failure detection, the 1460 can be re-deployed to with split signalling and media architectures, such as those being defined by the ETSI TISPAN initiative.

Overload Control and Interconnect Congestion Management

The interconnect functions will be supported by a limited capacity defined by the bandwidth of the interconnecting links. In order to prevent the overloading of these links and the consequential decline in call quality, a bidirectional session admission control system, such as provided by the Newport Networks 1460 should be used.

This will regulate the use of the interconnect, but it cannot effectively deal with overloads related to focus events. These events, often initiated by television programmes that involve subscriber voting, consist of very high volume, short calls. These events are very financially rewarding and calls should be handled efficiently whenever possible. At the same time other calls such as emergency calls should be preferentially allowed to pass.

This 'Session Admission Control Plus' capability, which can regulate call volumes and bandwidth usage to target ranges of telephone numbers rather than simple destination IP addresses, not only caps interconnect usage but also allows the early back-off of damaging call volumes, through the selective and increasing use of SIP 503 messages (this message allows the SBC to provide a server unavailable message but also to ask the sender to retry after time x) to achieve traditional call gapping functionality.

Conclusion

The 1460 product family includes both Signalling and Media Border Functions that can be deployed in a number of ways to meet the requirements of NICC and UK service provider PSTN emulation VoIP interconnect. The H.248 capabilities of the product allow a range of evolutionary deployment scenarios as standards become available and deployed. The 1460 products enables carrier resilience to be deployed with a number of multi-node mechanisms to provide a PSTN equivalent service using VoIP, which is as good as or better than current PSTN practice, and at a much lower cost.

Newport
networks

www.newport-networks.com

info@newport-networks.com

Whilst every effort has been made to ensure that the information included in this publication is accurate at the time of going to press, Newport Networks Ltd assumes no responsibility for the accuracy of the information. Newport Networks Ltd reserves the right to change their specifications at any time without prior notice. Some features described in this document may be planned for future releases and may not be available in the current product. Newport Networks Ltd reserves the right to modify its product development schedule without notice and does not guarantee that any feature or product mentioned in this document will be produced or produced in the form described.