Identity Assurance with Voice over IP

Andrew Rutherford

Iagu Networks

<andrewr@iagu.net>

PRESENTATION NOTES

With both companies and individuals using VoIP for cost reduction and/or increased flexibility, we find the line between computers and the Public Switched Telephone Network significantly blurred. Identity issues over the Internet will be discussed. Also discussed is that this new way of implementing telephony challenges many of the assumptions inherent in telephony about identity.

There are two main VoIP signalling systems operating over the Internet today, the International Telecommunication Union's H.323 and the Internet Engineering Task Force's Session Initiation Protocol (SIP). H.323 is a binary protocol and thus must be reformulated to add extensions, while SIP is a text based protocol.

H.323 was based on the H.320 point-to-point videoconferencing over ISDN standard, and version one is very much a "lets just get this working" standard. Two is much more usable, but was based on carrier requirements and so does not have PABX features such as Call Transfer - so all the vendors have their proprietary extensions on top of H.323v2, making them incompatible. Version 3 was never really deployed because of some ambigutities in the specification, and version 4 is what version 3 should have been, including features such as Call Transfer, to try and merge the various H.323 flavours back together.

SIP was designed after H.323, and the aim in design was to avoid the mistakes of H.323. Rather than reinventing the wheel, intentionally a lot of the work borrows on SMTP and HTTP, to make implementation familiar to developers. Headers can be added - if you don't understand a header, you can just ignore it and process what you do know. There is the capacity for option support announcement and option negotiation if one needs to know if the other end supports a header or is just ignoring it.

The example shows a minimal INVITE packet. The URI shows the address we are trying

```
--- Begin Sample SIP Packet ---
INVITE sip:+61884252203@iagu.net SIP/2.0
Via: SIP/2.0/UDP 203.32.218.40:5060;rport;branch=c86a23a64f95a00c43f6b8c4a3d9959c
Via: SIP/2.0/UDP 203.32.218.40:1028;branch=z9hG4bK16df486e
From: "Tony Clark" <sip:tony@sip.adl.rsp.com.au>;tag=001120dfe59f00241c877e37-41c26c94
To: <sip:0884252203@sip.adl.rsp.com.au>
Call-ID: 001120df-e59f0021-146846f8-231d6c1b@10.5.4.126
CSeq: 101 INVITE
Contact: sip:tony@203.32.218.40:1028
Date: Thu, 12 Aug 2004 02:28:21 GMT
Expires: 180
Remote-Party-ID: "Tony Clark" <sip:tony@203.32.218.40:1028>;party=calling;id-type=subs
criber;privacy=off;screen=no
Max-Forwards: 69
User-Agent: CSCO/7,Iagu-Slipper/3.2b5
Content-Type: application/sdp
Content-Length: 246

v=0
o=Cisco-SIPUA 14618 19823 IN IP4 10.5.4.126
s=SIP Call
c=IN IP4 203.32.218.40
t=0 0
m=audio 25312 RTP/AVP 18 8 0 101
a=rtpmap:18 G729/8000
a=rtpmap:8 PCMA/8000
a=rtpmap:0 PCMU/8000
a=rtpmap:101 telephone-event/8000
a=fmtp:101 0-15

--- End Sample SIP Packet ---
```

Figure 1: Sample SIP Packet

to contact. The Via headers show the path this packet has come so far. The From and To headers show the original source and destination of the call, and the ";tag=" is a unique identifier for the call leg. The "Call-ID" is similar to a "Message-ID", but may persist for a single logical call with multiple call legs (eg, Transfers, Conferences, etc). The CSeq (Call Sequence) is a mechanism for separating out multiple messages about the same call leg. The Contact header informs the recipient as to the URI to use for initation of messages in the reverse direction, eg BYE messages to terminate the call, or REFER to initiate a call transfer. The Remote-Party-ID will be discussed further, but maps directly onto ISDN and SS7 signalling.The SDP (Session Description Protocol) message body. It describes the IP address of the end-party, the port to be used for audio media, and a preference list of codecs.

Via headers are analogous to Received headers in email. Responses (180 Ringing, 200 OK, 302 Redirect, 404 Not Found, 486 Busy) are sent back using the "Via" path, and these must correctly identify the host, or the IP and port to send the audio media to are not known, thus they cannot be faked without active assistance from those in the path, or packet sniffing.

This does not protect you from compromised servers, such as the "zombies" beloved of crackers.

The Remote-Party-ID header is a direct map of a number of things normally set within the carrier network. This normally carries the A-party identification used for billing purposes, but can also indicate if a number is suppressed or not. Note that like "From" headers in email, additional information beyond the addressing information can be embedded in the line.

The commonly used SIP methods are REGISTER, which is a network login equivalent to a GSM phone letting the network know which cell it is on, INVITE to start a call, CANCEL to abort setup (eg, user has hung up before you answered), BYE to terminate an established call, REFER to do call transfer, MESSAGE to send a message of any MIME Content-Type, and SUBSCRIBE to request notification of events and NOTIFY to deliver that notification for thing such as presence or Message Waiting Indicators for voicemail.

The SIP call process consists of an INVITE being sent, with multiple 1xx replies before the 200 OK when the call is set up, followed by the intiating end sending an acknowledgement of call setup. The RTP media stream is then set up end to end, before being terminate by one end sending a BYE, followed by the other end acknowledging with a 200 OK.

User authentication, required for most chargeable services, is usually done by MD5 Digest authentication, using the exact same method as Digest authentication in HTTP. In fact, the SIP standard tells one to refer to the HTTP standard rather than replicate it in the SIP standard. An accounting system can also trust an IP address, or a certificate used in TLS (sips), although the latter functionality is not widely supported, especially given it interferes with NAT fixups when crossing a NAT boundary. In a carrier system, one also has to watch out for Diverted or Transferred Calls. For example, a person an organisation A calls a person at organisation B, who has their phone diverted to a PSTN gateway operated by carrier C, who has a relationship with both A and B. They must be careful to bill B and not A. Transferred calls are more of a problem, as if B transfers a call to C, B sends a message to A asking them to establish a new call leg directly with C.

As RTP is sent end-to-end, proxy servers are not involved in an ongoing call and so won't know if a call terminates by both devices being simultaneously removed from the network. If charging, you must control one end and have it generate BYE's that are proxy-visible if the end-to-end RTP stream shuts down unexpectedly.

Large organisations such as Universities often have internal emergency service numbers through security or similar, because supplying caller-ID to normal emergency services won't give them any useful location information. With VoIP, it is possible to give out a few specific Caller-ID numbers, say one per building or per floor, only to emergency services, over-riding normal Caller-ID, and pre-declare to emergency services through the carrier the location of these numbers. Embedding textual information is currently a one-way street - Telstra may do it to declare a number "OVERSEAS", but we can't send them anything. Putting in detailed location information would be useful in an emergecy services context. Carriers limit the incoming Caller-ID to only those from the blocks allocated to that service, which prevents one from placing calls via alternative gateways with the correct Caller-ID if a gateway is down.

The ACA is currently investigating the possibility of using non-geographic numbers for VoIP services, where each number would have to be declared as to location. This would prove useful to allow a call placed in one location to be routed to emergency services via some other location if, say, the local gateway was damaged in the emergency. Power must also be addressed, for example the 802.3af inline power for phones, UPS's, etc.

Normally a phone call is routed via carriers, but with SIP although the IP packets may follow that path, call signalling logically goes direct between the two end points. In the PSTN world, the fundamental routing unit is the phone number, but with VoIP it's the IP address, so the phone number is no longer bounded by any topological conditions. ENUM and TRIP allow for automatic discovery of where your call to a phone number is best routed over the internet, ENUM using DNS and TRIP using a BGP-like routing protocol. People can also advertise SIP addresses like email addresses or phone number on business cards, .signatures, etc. Thus like email, a phone call may come from some device you do not have a pre-existing trust relationship with, unlike the traditional phone network where there is some element of trust between each connecting party, forming chain of trust that law enforcement officers have a window to. Issues like the contextual problems of Call Transfer become more intricate when the three parites involved don't have a pre-existing trust relationship.

Like CGI's for HTTP, setting up VoIP applicaations is easy. Many such systems include call transfer features, for exaample a voicemail system that allows you to press 1 for reception or 2 for that person's mobile, and often these are written to quickly fix a problem without a view towards security. This can create problems as unsavoury types may get poorly written applications to annoy others or rack up charges.

The upcoming problem is going to be Phone SPAM, which could end up being much worse than email SPAM due to the phone's intrusive nature. Via headers, unlike Received headers, limit spoofing, but that doesn't counter the growing threat of zombies, which don't have to maintain anonymity of the VoIP device to manage to hide the originator of the call.