SIP vs MINET for managed VoIP telephony applications

The major differences between the SIP protocol and the Minet protocol can be traced back to the radically different philosophies adopted when the two protocols were first developed.

SIP (Session Initiation Protocol)

SIP (Session Initiation Protocol) was developed at the IETF (Internet Engineering Task Force) for computing platforms to make media (voice) connections directly with other computing platforms. With SIP, the endpoints act as peers. SIP is a “peer-to-peer” protocol because developers believed endpoints could be made increasingly smart and so there was never intended to be a third party mediating or supporting their connections. Consequently, SIP is in no way optimized for hosted service. It can be argued that SIP was actually intended to make service providers unnecessary. Of course service providers do a lot more for customers than just connect calls. Businesses need service providers to keep records, manage resources, simplify services, manage change, add value, and just plain keep things working. Because of the philosophy behind its design, SIP does not make those things easy for service providers. As evidence of the inefficiency of the SIP protocol in hosted Voice over IP (VoIP) environments, note that all major VoIP PBXs work better when their customers choose that vendor’s propriety VoIP protocol over SIP.

Minet

In contrast to SIP, Minet™ was developed by Mitel™ so that an intelligent server is always in charge of managing telephone calls. The Minet philosophy is that there is a controlling server, and all of the intelligence naturally resides there. That’s where functionality is most conveniently maintained and updated. The user simply presses a key on their phone and the server is told which key was pressed, and then the server uses its latest software and configuration to decide the appropriate actions. Minet is thus a “stimulus-response” protocol. Stimulus protocols minimize phone software and don’t attempt to set the phone up as a peer to the server.

Who’s in Control?

When you buy a hosted VoIP solution for your business, you expect the service provider to manage all of the technical details. You only want to add and modify users. You only need control over the end-user features: forwarding, speed-dials, parking, intercom, group features, etc. But even for those end-user functions you need the service provider to be in a position to see everything and assist immediately when there are difficulties. For all of those reasons business VoIP providers, SIP-based or not, try to control their services from hosted servers. These VoIP servers “in the cloud” give the service providers both control and visibility. But SIP-based hosted service providers are always at a disadvantage, because phones using the SIP peer-to-peer protocol inherently hide things from the service provider.

For example:

- They collect dialed digits and decide on their own when to invoke the server.
- More generally, buttons pressed on a SIP phone are not seen directly by the associated server or provider. End users can make procedural mistakes in handling a call that the service provider cannot see.
- They hold critical configuration information, commonly provided when the phone is shipped or updated, and the configuration on those phones can change or be out-of-date without the service provider knowing.

Providers of SIP-based services have gone to great lengths to mitigate all of these issues, but there is still no way of their knowing for sure exactly what is happening on the phone or what might have been done that altered the phone. This creates support problems for both the end customer and the service provider.
Stimulus-Response Protocol

With a stimulus-response protocol like Minet, the server is always in control. No matter what you or your users do to their phone, all the server is ever responding to is “what button was pushed”? It’s not possible for the phone to initiate service requests without the server knowing, because service requests start and end with those individual key presses. It’s not possible to configure the phone to do something different because all the “somethings” are done at the server.

The UniVoIP OfficeConnect user seemingly configures the phone, but is actually configuring the server’s response to the phone. Anyone with sufficient access to the server configuration (e.g. a UniVoIP OfficeConnect tenant administrator) can see exactly what their user has done “to their phone”. Furthermore, in extremely unusual cases where things don’t seem to be working properly, the service provider simply monitors the protocol at the server and “sees” every key the user is pressing. An analyst can determine exactly how the user expected the phone to work and exactly what happened in response. Determining an appropriate fix is so much easier.

The Disadvantages of “Intelligent” Phones

With stimulus protocols like Minet, phones do not have to be smart. They are making no decisions on their own. A button is pressed, a message is sent to the server, and the server makes the decisions. It tells the phone everything about how to change its display and lamps, when to generate tones, and when and where to connect audio. The phone is dumb. Instead of investing money in phone processors and phone software development, developers of stimulus-response phones can invest in better physical design, better phone construction, better displays, and better sound quality.

Whenever intelligence resides in the phone, as is the case with SIP, not only does each device have to have more processor power and more software, but the phone must also be locally configured and specialized for the role it is serving. Hosted VoIP service providers dealing with SIP phones have invested significant effort figuring out ways to deliver the right custom configuration files to SIP phones. These custom configuration files usually start by assuming a few standard employee roles: manager, administrative assistant, receptionist, etc. Those stock configuration files are further customized by assuming the “best” phone model for each particular role. Then you as a customer must match each of your employees to one of these pre-defined roles and phones. The provider helpfully personalizes the configuration file for the designated user’s extensions. Then the fully customized ‘config’ file is transmitted to the user’s specific phone. Then the user can try it out. The delivery of a replacement config file is possible, but the phone sometimes needs to be rebooted for the new configuration to take effect. This process of matching a user to a pre-defined template, customizing, delivering, and booting with the new configuration is repeated for each of your users.

All this effort to achieve SIP phone customization also creates an obstacle to portability and mobility. Because the customization must be deployed to the specific phone, when the user moves to a new phone, the customization has to be moved, too. There are many reasons users want to move to a new phone - replacing defective hardware, visiting at a different office, working from home, or upgrading to a newer model. In each of these cases customization done for the current phone must somehow be copied or ported to the next. If someone else has been using the new phone and will need to use it again, the config file on that phone will have to change twice. If the new phone is a different model, the config file will likely have to be modified. SIP-service providers have gone a long way toward making this all easier than it used to be, but all of the steps described are still required and the processes are still intrusive.

UniVoIP OfficeConnect takes advantage of its Minet stimulus protocol so that each company and every user can customize their “phone” endlessly. The configuration changes are actually to the way the server responds to key presses and take effect immediately, no phone rebooting required. Since changing the “phone” does not actually change the phone, there is no need to do a download or a reboot just to try out a new function. Every set of phone keys can be customized endlessly, with or without assistance from the service provider. Pre-conceived user templates (key programming templates) are possible and indeed helpful, but they don’t lock anyone permanently into any part of their pre-conceived role.
Furthermore, the portability problems created by “intelligent” (i.e., SIP) phones are practically non-existent with stimulus phones. Logging your extension into a new stimulus phone just tells the server to start responding “my way” to key presses from this phone. Logging back in as the previous owner tells the server to go back to responding the “old way”. This is what UniVoIP OfficeConnect has called “hot desking”.

Change programming this morning and log into a new phone this afternoon and all this morning’s programming immediately works here, too. If the phone is an upgraded model with more keys, all the old phone’s keys show up in intuitive places on the new phone. If the borrowed phone has fewer keys than the old phone, the most commonly used keys from the larger model are preserved.

Changes to functionality, replacement phones, upgrades, teleworking, and roaming between offices are all made easy with a stimulus-response protocol like Minet. Distributing intelligence and configuration information to phones, as required by SIP, complicates all of those processes.

TCP vs. UDP

TCP is a part of the IP protocol stack that protects the delivery of information. Minet uses TCP. If Minet packets drop, TCP takes care of retransmitting them so that they always get to their destination. Furthermore, every TCP packet is sequenced from the time the TCP connection is built (when the phone boots up), until the connection ends (at the next power-down). TCP retransmission makes packet loss problems rare and TCP sequencing makes trouble-shooting more coherent. SIP services usually use UDP, not TCP. While SIP has its own retransmission strategy to make up for the fact UDP does not retransmit, UDP is also “connectionless” which means it is harder for the server to detect that it has lost contact with the phone. “Connectionless” also means it is easier in principle for an intruder to “spoof” the SIP phone’s packets and steal service.

Security

While it is not possible to quantify this benefit, the fact that SIP is a widely-used standard with its content transported as readable text makes it an attractive target for hackers. Furthermore, broadband service providers have been known to search for SIP messages and restrict their use on their networks. Minet is an unpublished proprietary standard, far less widely used. It also uses digitally encoded name/value pairs to transmit its information. These facts combine to make Minet far less interesting to hackers and to those looking to regulate traffic on their networks.

Data Models

Because SIP endpoints are customized for the individual user, inside many SIP-based VoIP servers the data model tightly connects user to specific phones (MAC address), and specific phone number. This further complicates the processes of changing phones and phone numbers. UniVoIP OfficeConnect has taken advantage of the fact that the phone itself does not have any customization for a user. The user is not modeled as being equivalent to their phone or their phone number. Having separate data models for users, their (current) phones, and their (current) phone numbers makes changing those associations much easier.