P2PSIP: On The Road To A World Without Servers

David A. Bryan

Where will peer-to-peer SIP be most useful? The answers may surprise you.

Peer-to-Peer SIP (P2PSIP) may still be a very new technology, but it has been a topic of discussion, debate and development for some time. The first research started in late 2003, and the first published academic papers and IETF (Internet Engineering Task Force) informal meetings started in mid-2005. The work has resulted in a great deal of "buzz" in the press and at conferences, several companies have brought P2PSIP based products to market, and a flurry of new ones are expected to hit the market in the next 18 months (also see *BCR* February 2006, pp. 36–40).

With a few years of critical thought and evolution behind the P2PSIP community, now is a good time to step back and look at the state of P2PSIP in 2007. In the past few years, P2P has been explored, discussed and debated, both inside and outside the IETF, where a working group was formed in March, and where work on P2PSIP specifications are about to begin.

Notwithstanding the lack of standard specifications, researchers and vendors alike are forging ahead to explore P2PSIP ideas, products and P2P enhancements to their existing products. They are primarily motivated by the promise of P2PSIP to reduce both capex and opex in communications networks and applications. Our work on the IETF and with our clients (we develop P2PSIP software) puts us in a position to comment upon, if not definitively answer, these questions:

- Where is P2PSIP emerging as a "ready for the market" technology?
- Where is there more work to be done?
- Where is P2PSIP a bad idea, or a bad fit compared to other solutions?

In this article, we'll take a look at these questions and give you a sense of the state of P2PSIP.

David A. Bryan is the CTO and co-founder of SIPeerior Technologies (www.SIPeerior.com), a leading producer of software development kits for vendors targeting the P2PSIP market. Mr. Bryan also chairs the IETF P2PSIP working group, and is actively involved with a number of open source telephony projects. He can be reached at 757-565-0101 or dbryan@SIPeerior.com.

Why P2PSIP?

A quick look at the world of computing and communications shows that server-based technology works. So the natural question to ask is: Why are people contemplating P2PSIP solutions? The reason is to reduce costs. Eliminating servers eliminates the original cost of the servers themselves, although, as their price keeps falling, this reason alone would likely not be enough to cause the interest we are seeing in P2PSIP. The bigger motivator is to reduce the ongoing operational cost associated with powering, connecting, managing and monitoring server farms, blades and arrays.

For example, if carriers distribute call control among endpoints, they can eliminate the electricity, bandwidth and, most importantly, the staff costs that would otherwise be necessary for those servers. In enterprises that adopt P2P VOIP, staff spend less time configuring or troubleshooting the systems, because P2P-enabled devices recognize each other automatically, and their distributed functionality is inherently robust.

Hosted VOIP providers also can benefit from P2PSIP, by moving everything but the PSTN call termination function out to the subscriber or enterprise endpoints, eliminating the need for call signaling servers at the customer or provider premises. Finally, P2PSIP can enable more sophisticated call handling choices for even the smallest enterprises, for which the maintenance and operation expense of a server might have been enough to preclude a VOIP solution in the past. In the end, it is all about reducing costs, most importantly, operational costs.

Where Will P2PSIP Be Used?

While it is still too early to be certain where the technology will find a home, recent debate seems to have segregated applications into three broad categories. First, there are places where there is definite interest in using P2PSIP and most questions about how to use it seem to be answered. At the opposite extreme, there are a few areas where it seems architectural or business concerns make P2PSIP unsuitable or difficult to integrate. Finally, there are some areas where the question seems an open one. In these areas, there is a great deal of interest in and promise for a P2PSIP solution, but there are also many unanswered questions.

Where is there definite interest in using

P2PSIP? The first thing that comes to mind for most people is a Skype-like system for global communication, but using an open standard. P2PSIP systems could reduce expenditures for central equipment and operational resources, improving the bottom line for service providers. Moreover, P2PSIP could leverage existing equipment, and the resulting networks could quickly connect to existing enterprise systems that speak SIP today.

P2PSIP also offers the prospect of a global service with no carrier at all, simply a collection of clients (possibly from different vendors, all based on a well defined standard protocol) connected to one another using the Internet cloud.

Although the possibilities for a global, decentralized solution get much of the attention, it is actually in the enterprise environment where the first real inroads are being made, due to the relative simplicity of building smaller systems and the perceived rapid revenue path for vendors. Systems like the Avaya one-X Quick Edition, the Siemens HiPath BizIP product and SIPeerior's P2PSIP Endpoint Development Kit quickly recognize new phones as they are connected to the network, and they require minimal provisioning beyond entering phone extensions into the devices.

A few services, most notably voice mail, must be done in a new way using a P2PSIP system, but many traditional SIP features work exactly the same in a P2PSIP solution. The ability to interoperate with traditional SIP environments, use SIP termination (trunking or in-house gateways), and the potential to expand an entry-level P2PSIP system into a more traditional SIP solution as needed make the enterprise play for P2PSIP very interesting to some developers, and attractive to some customers.

As a result, these deployments are already beginning to make inroads in the market. Frost and Sullivan published a report in January of 2007 indicating that P2P telephony had captured 0.3 percent of the European business telephone market, a number they expect to grow to nearly 4 percent in 2012. While that number seems modest, this technology is just emerging, and is only available from a few manufacturers to date.

P2PSIP also seems to be generating much interest in a space that at first seemed unlikely—among the server vendors and their carrier customers. In the same way that P2PSIP can be used among endpoints to cluster and communicate without a central server, it could also be used among servers. For example, to make calls, endpoints would still connect to the servers, with a particular call still being handled by one server, but that server fetches some of the call information it needs from the P2P overlay before it starts handling the call. Billing records would still be written by one billing server per call, but into the overlay instead of to a large database. In this architecture, only one server handles a particular call, but as the call arrives, any of the servers could be selected to process that call. This approach could help carriers improve their scalability and redundancy.

Finally, P2PSIP is emerging in a number of other areas. Because P2PSIP systems require very little configuration, and no central servers, they are ideal for *ad hoc* systems. Essentially, the idea is that wireless access points and endpoints are turned on, then any P2PSIP devices in the area can communicate among themselves.

This zero-configuration approach has some security concerns (most notably the lack of user authentication), but in some scenarios, the tradeoff looks viable. While there are no applications being brought to the market to date, discussions in the IETF in this area have focused on *ad hoc* networks that could, for example:

- be set up quickly in the wake of a disaster,
- be used in a small conference or meeting (in this case using SIP—Session Initiation Protocol—to establish multimedia collaborative sessions such as slides or applications, rather than voice sessions),
- and in the developing world, where a communications system could be dropped into a remote area for internal communication.

P2P is a good fit with *ad hoc* networks

Where Not To Use P2PSIP

As work has continued on P2PSIP, it also has become apparent where this technology may not be as well suited. One obvious environment that won't work with P2P is in large, centralized deployments, in which tight control of the call is the primary motivation Three examples that quickly come to mind are:

1.) Carriers that want to carefully regulate (and charge for) the end user experience

2.) Large corporations that want or are legally required to monitor or control their employee's communications

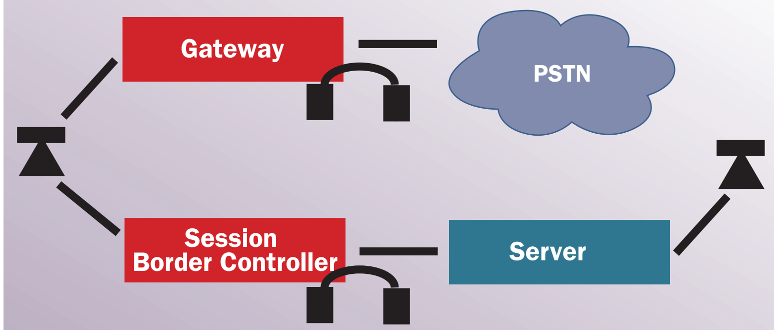
3.) Entities that need to intercept calls for law enforcement

A P2P approach could reduce operational costs in these deployments, but its distributed nature would also eliminate a convenient central control point for billing, monitoring and other such activities (Figure 1, p. 42).

Another area in which it is less clear how P2P might be put to use, is in systems that require call signaling to be closely coupled to big databases. Two examples are: the large-scale customer relationship management (CRM) VOIP systems; and interactive voice response (IVR) systems used by airlines and other organizations with heavy customer interaction. It is hard to distribute and frequently access such big databases using the distributed hash tables (DHTs) that commonly undergird P2PSIP.

DHTs are extremely efficient at distributing information among many machines, then looking up information when you know what you are

FIGURE 1 Signaling And Media Intercept In Traditional Network



Note: In traditional networks, session border controllers can be placed in the call path for IP-to-IP calls. PSTN calls are intercepted at the gateway.

looking for. For example, if you are trying to call a particular username, or get a list of devices that can provide a particular service, then a DHT is fine. But DHTs are not good when you don't know what you are looking for, e.g., "Which customers who purchased items in the last 2 weeks?" or "Give me all the users that live in Finland."

In other words, it's easy to distribute a registration, or a provisioning file, or even a billing record for a particular user, but if you want to search the records, you need a database. Thus, if a central server must exist for the application database, having it closely coupled to a centralized call control server looks to be the simpler approach for the near future. That said, however, there does appear to be a bright future using P2PSIP among remote phones for distributed call center applications, if only as a means to connect those phones back to a centralized SIP system supplying the actual call center services.

Finally, there are a number of use cases in which P2PSIP has potential, but the problem isn't well enough understood to say for certain if P2PSIP will make inroads. For example, voice carriers have eyed P2PSIP as a tool for offering low-cost second line services, often called "teen lines." Because P2PSIP is self-organizing, the systems would require less maintenance, allowing these additional lines to be sold at a lower price. Such lines could run over broadband connections or could be incorporated in next-generation mobile systems.

As wideband wireless protocols such as wireless USB gain in popularity, exploring alternatives for connecting clusters of consumer electronics devices together becomes increasingly important. Because many of these devices are designed to stream media, SIP is a logical choice, and the self-organizing behavior of P2PSIP could allow various consumer electronics devices to establish a cluster and share media streams in a well-defined, standards-based way. Similarly, set top box developers for IPTV see P2P as a simple way to reduce bandwidth and server load by having "hot" on-demand programming obtained from other nearby boxes, rather than from a central server. P2PSIP may prove valuable for this scenario, particularly if SIP is already included on the box for other end user services such as video, VOIP, or gaming. What is still up in the air is if P2PSIP—or some other kind of more general P2P—will prove best suited for these applications (also see *BCR* October 2006, pp. 48–53).

Recent P2PSIP Developments

While we are starting to see commercial P2PSIP deployments, much of the excitement and momentum behind P2PSIP has been taking place in the IETF. The IETF first discussed P2PSIP at an informal meeting in March of 2005, and has had five more meetings to date. The meetings have been very well attended and, when the group met last month in Prague, they did do so for the first time as a formal IETF working group.

The newly formed working group will have a number of tasks ahead of it, most of which it already has been working on informally. In fact, at press time, more than 20 Internet Drafts dealing with P2PSIP already had been submitted and were generating lively debate on the P2PSIP mailing list. All these drafts (and more) can be found on the P2PSIP community site, www.p2psip.org.

The IETF P2PSIP working group will first create a concepts and terminology draft, probably based on draft-willis-p2psip-concepts (currently in the fourth revision), which has been the most popular topic of discussion on the list in the last few months. Then the group has agreed to produce one or more formal drafts defining the P2PSIP protocol, and this is where more detailed choices will have to be made. The group is facing several major decisions, including:

- **Selecting the right P2P algorithm**—This is likely to prove contentious. While all the proposed DHTs basically perform the same task—locating a resource among a distributed collection of peers—each one does so in a slightly different way and offers advantages and disadvantages. The IETF is likely to take an approach where one or a small number are mandatory to implement, but additional DHTs can be easily added, leaving room for innovation in this area in the future.
- **Which protocol will put the "bits on the wire"**—Will SIP messages be used to organize the DHT? Advocates of SIP point out it already has mechanisms for routing, security and NAT traversal, will require less new protocol development, and that all P2PSIP endpoints will need to speak SIP in any case. Others suggest something besides SIP be used, to reduce message size and to potentially allow the work to be more easily leveraged for other non-SIP applications.

In some ways, the real work on P2PSIP has just begun. While the group was informal, authors proposed their own solutions and such choices were debated, but not formally resolved. Now some tough decisions need to be made, consensus reached, and a formal P2PSIP protocol defined.

Outside the IETF, acceptance within the broader community seems to be growing. A quick search for open source projects at Source Forge shows several P2PSIP projects under way, but the lack of a standard has made it difficult for one to emerge as the leading project. Instead, each is taking a slightly different approach to P2PSIP.

The lessons learned by these projects are likely to be very useful to the working group, but it may prove difficult for one to really attract the core base of developers needed until the working group's direction becomes clearer. After all, P2PSIP is not about building "just" a new VOIP protocol, but also one that will work with today's SIP infrastructure. The developers who are drawn to P2PSIP are drawn by the interoperability potential, which will emerge more fully along with a firm direction for the standard.

Progress And Prospects

So when will we see the products? The short answer is that some P2PSIP products have already reached the market. Last year, Avaya's one-X Quick Edition phone system became the first SIP-compatible P2P telephone product to reach the market, followed by Siemens. Vendors continue to develop their solutions, and many, including Avaya, are actively involved in the new standards work at the IETF.

SIPeerior Technologies is working with a number of vendors to provide P2PSIP technology for their products as well. In some cases, these products will be marketed as P2PSIP products. In other cases, they are being used as enhancement to an existing product, adding new capabilities that just happen to be implemented over P2PSIP.

Some issues have arisen as P2PSIP applications have been developed and brought to market and as standardization efforts have progressed, all of which will need to be addressed as P2PSIP matures. While P2P systems have a number of attractive features such as scalability, ease of configuration and fault tolerance, a few problems have proved more difficult to address in P2P systems than in traditional client-server solutions.

Security and identity assertion need to be addressed in a very different way with a P2P system than in a traditional client-server approach. In a centralized system, the server can decide who is allowed to place calls and which user is given which username, but one of P2P's primary design paradigms—distributed admission and call control—sometimes works against today's security and identity management solutions.

With P2P systems, peers decide whether a new peer is admitted, how calls to and from the peer are handled, and how and where to store resources. To date, no *fully* distributed system for security has been devised that appears to be as robust as a centralized solution which controls call admission. In fact, so far most of the work in the IETF proposes a limited centralized entity (contacted only at enrollment time) that provides security certificates and controls identities, while distributing as much of the rest of the system as possible, including call control, NAT traversal and offline storage.

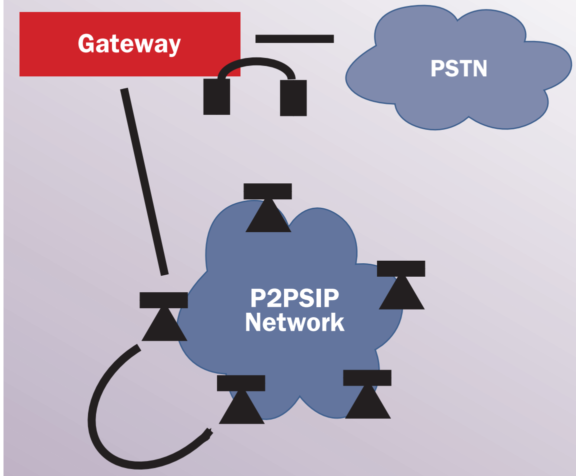
The approach most frequently discussed on the mailing lists relies on a certificate server issuing security certificates to each peer when that peer first enrolls in the system. Each peer also stores credential information from the server, which allows it to verify the authenticity of other certificates they see later. Each certificate asserts that a peer is allowed to join the network and can also provide a unique username for the user. Without a valid certificate, other peers will reject a rogue peer. Because each peer can verify the certificates, the server needs only be consulted one time, to obtain the certificate at initial enrollment. Normal operations such as a peer joining, leaving or placing a call are handled strictly among the peers.

Lawful intercept issues could also limit where and how P2PSIP can be deployed. Lawful intercept can involve obtaining only the signaling information (such as who is calling whom and for how long) or in some cases the signaling and media (the actual voice or video conversations).

In traditional SIP, call signaling information flows through a central server, and media flows directly between the endpoints. Calls flowing to the PSTN typically flow through a gateway (Figure 2). An ISP can easily comply with lawful intercept requests for signaling information, since this information resides on their server, and PSTN gateways provide a point to obtain both signaling

Some P2P products have already reached the market

FIGURE 2 With P2PSIP Signaling And Media Intercepted Only At The Gateway



Note: Within the P2PSIP network, IP-to-IP calls have no predetermined path, so inserting an SBC is difficult (or defeats the purpose of P2P), making IP-to-IP call interception difficult. PSTN calls are still intercepted at the gateway.

Lawful intercept poses a significant challenge

and media for calls going through them. Vendors such as Ditech and Acme Packet have sold session border controllers (SBCs) to service providers to provide a "choke point" at certain network boundaries through which all media and signaling will flow, enabling both to be intercepted.

With P2P, interception may be more difficult. There is no central server which is aware of all calls that were placed, and inserting SBCs into the architecture would reintroduce the central servers that a P2P approach is designed to avoid. Carriers deploying P2P systems will probably mitigate this with a hybrid solution, where a centralized authentication or billing server may be able to provide signaling information, but the question of how to do media intercept in a P2P system is a thorny one. Recently, there has been discussion about whether the FCC will force Skype to provide for CALEA compliance. It will be interesting to watch Skype to see how they handle compliance, if and when they are ordered to do so.

Troubleshooting and network management will also require a different approach in P2P systems. Since the information that controls the network is distributed among many peers, management systems will need to be developed that query all peers directly for system information. While there are no technical reasons this couldn't be done, it will require a new approach to developing these tools, and more advanced debugging tools to isolate faults that may be distributed among many peers, rather than in one central server.

Conclusion

The coming months will be critical to the future of P2PSIP. As a standard emerges and vendors begin to implement it and interoperate with one another, the true capabilities of P2PSIP will become clear. The stakes are high for the standard to be broadly applicable, resilient, scalable and easy to implement and debug. While it is likely that P2P will make headway in a number of areas such as small enterprise systems, the extent to which the technology will be adopted may only be clear once vendors have a chance to see the standard that emerges and how it can apply to the challenges they face in the marketplace□

Companies Mentioned In This Article

Acme Packet (www.acmepacket.com)

Avaya (www.avaya.com)

Ditech Networks (www.ditechcom.com)

Skype (www.skype.com)

Siemens (www.siemens.com)

SIPeerior Technologies (www.SIPeerior.com)