

Inter Domain Settlements

In Multi-Service IP Networks

By Dr. Matthew Lucas and Philip Mutooni

Settlements between Internet service providers (ISPs) have never moved beyond the “bill and keep” peering model that evolved from the early days of Internet privatization. Informally, this means ISPs agree to terminate each other’s traffic without charge. At first, the Internet peering model sounds absurd to those with a circuit-switched background, where carriers charge for termination/access on a per-call basis. But it makes sense for a number of reasons, the most compelling of which is global connectivity. Clearly, it is in every IP provider’s best interest to make sure that all networks are reachable - as in the PSTN where you can call anyone regardless of which long distance, CLEC or LEC they terminate with. ISPs must have the same level of connectivity; otherwise there may not be a network path to a particular web server, and who would buy that service?

The Internet peering model also makes sense due to the traditional “best effort” service provided by IP carriers. In essence, because ISPs do not give SLAs or performance guarantees for the transport of messages exchanged between them, the value of the service is not worth the trouble of the accounting. Finally, the Internet accounting models were founded on government subsidies; thus, settlement technology and the accounting infrastructure was never developed. Of course there are some glitches (e.g., many large carriers are reluctant to peer with small carriers, see *Billing World’s* “Peering Pressure: Providers Won’t Exchange Traffic with Just Anyone”, May 1999 issue), but to date it has worked well.

IP telephony, however, is rocking the settlements-free boat. In effect, carrier settlements are required for QoS-based services (where providers *must* make performance guarantees for the packets exchanged between them), or when costs are incurred beyond that of basic packet transport. IP telephony qualifies on both counts. First, telephony packets require performance guarantees. In particular, toll quality IP voice requires, across all providers, less than a three percent packet loss rate and less than 200 ms transmission delay. Compared with “best effort” service provisioned today, deliver-

ing these QoS guarantees requires ISPs to deploy additional bandwidth, traffic management systems and QoS protocols – all of which substantially increase the provider’s cost. Second, there is the additional cost of gateways (the device that takes analog voice, converts it to IP packets and back) as well as per-use PSTN access charges.

Because of the termination and transit costs, Internet telephony service providers (ITSPs) have no choice but to capture each transaction and reconcile with each other periodically (as is done with circuit carriers today), in real-time or on a pre-paid basis. Achieving a bilateral settlement infrastructure for ITSPs similar to that found between circuit carriers is a problem for a number of reasons. A standard CDR, or standardized reconciliation capabilities on the gateways (or gatekeepers), do not exist. Second, there are hundreds of ITSPs around the world in the business of terminating IP telephony calls - each using a different currency and tax structure and each with distinct regulatory constraints. Establishing connectivity and bilateral settlement capabilities with every other carrier would require order N^2 number of billing relationships (where N is the number of carriers.) This doesn’t scale and is simply unmanageable.

Wholesalers are therefore a very

important part of the ITSP landscape, because they provide a solution to the settlement, call routing and connectivity problems. An ITSP can establish connectivity and a billing relationship with the wholesaler. ITSPs that cannot terminate a call within their own network can route the call across the wholesaler's backbone to an ITSP that has termination in the destination area, a compatible gateway and meets the IP quality constraints (delay and packet loss) described above. Thus, the originating and terminating carriers do not require a billing relationship, because the wholesaler has agreed to pay for the call.

Of course, there are some problems. Gateway compatibility is a big issue, each of the 160 gateway vendors has implemented a different subset of the H.323 protocol. Quality transit is also a problem. Some wholesalers manage their network constantly, and through proprietary technology offer their customers and partners an assured level of quality for service calls, while others don't have this capability. Both of these issues are a matter of engineering (i.e., agreeing on a standard implementation) and investing the money to develop the software to tackle QoS issues. Therefore, the remaining challenge is inter-domain billing and settlements.

Scaleable Inter-domain billing and provisioning

As engineers resolve gateway interoperability and interconnection problems, ITSPs are focusing their attention on settlements. Today, the settlement problem lies primarily with international voice service, where a large number of ITSPs circumvent the international accounting rates to offer call rates at substantially lower cost than the PSTN. The trick, however, is to identify good terminating carriers within those countries, and establish peering arrangements with them. However, there are thousands of ITSPs. Quality, credit worthiness, regulatory hassles, currency rates, and taxes are all concerns. The industry must adopt a mechanism to sustain the peering arrangements

already in place. This becomes complex, because settlements are really an issue for *any* IP service (e.g., video, multicast, e-commerce) that requires QoS or incurs other per-use costs.

The settlements problem manifests itself in provisioning and billing across multiple service provider networks as follows: when two (or more) carriers establish a peering relationship, they need to determine the remote IP address of the gateways/gatekeepers in order to terminate (i.e., route) traffic to a node. Distributing these endpoints poses a significant scale problem in a multi-service provider environment, because N networks with M nodes need to be provisioned for each new ITSP. Therefore, to do inter-domain settlements would require knowledge of all endpoints on each ITSP's network. Circuit-switched networks rely on CABS and SS7 standards to solve routing and settlements issues. Unfortunately, CABS or SS7 standards equivalents do not exist in the IP world.

All the large VoIP carrier's carriers - IDT, VIP Calling, Delta Three - are looking at clearinghouse models to solve the billing and provisioning constraints. Clearinghouse technology providers include Transnexus, Gric and iPass. A clearinghouse reduces the number of billing relationships from N^2 to N, because ITSPs form bilateral relationships with the carrier. In this way, the carrier assumes the burden of negotiating multiple bilateral agreements with its ITSP partners, as illustrated in Figure 1.

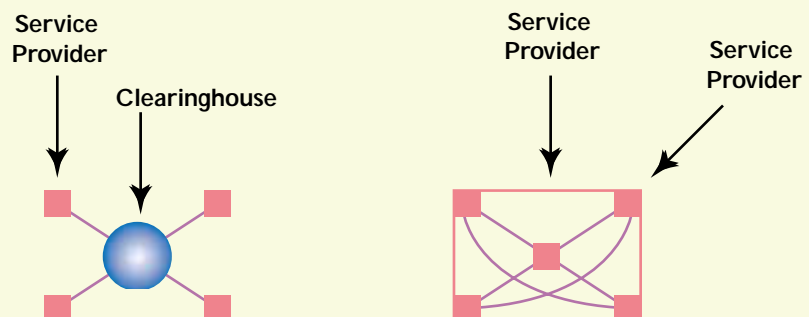
The iNow! Profile

Despite the adoption of a similar basic clearinghouse model, each carrier has a different implementation. One approach is to enable settlements in concert with inter-gateway/gatekeeper interoperability. The iNow! (interoperability NOW!) profile, pioneered by Lucent, VocalTec, and ITXC, bundles gatekeeper-clearinghouse interoperability with the session level (i.e., call setup) interoperability functionality (see iNow! Sidebar, Pg. 46). While the border elements are shown as separate functions, they could be co-resident with the gateway, gatekeeper, or clearinghouse.

The iNOW profile was developed to specifically support IP telephony. Although it did not come from a standards-setting body, Lucent and Vocaltec have incorporated it into their gateways, and used it successfully as the settlement mechanism within ITXC's network. However, as an embedded IP telephony settlements protocol, iNOW has some problems:

1. Inter-clearinghouse traffic flows: iNOW does not scale beyond one clearinghouse, which makes it unsuitable for settlements, because inter-domain billing (for traffic traversing two or more clearinghouse networks) is impossible.
2. H.323: iNOW is tied to H.323 session level interoperability; thus, service providers are limited to settlements on H.323-based services. Given the problems with H.323, many vendors and

Figure 1: The value-proposition of a clearinghouse settlement model vs. a bilateral settlement model. The clearinghouse model has four relationships, compared with 10 in the bilateral model.



Call Set-Up in an iNOW! Clearinghouse

The call flow for a call between different ITSPs that are part of a clearinghouse is as follows:

1. The originating gateway (OGTW) performs authentication and authorization on an in-bound inter-domain call.
2. The originating gatekeeper (OGK) validates the caller.
3. The OGTW requests the destination digits.
4. The OGK contacts its associated Border Element (BE), which determines the clearinghouse to use. Each clearinghouse also has associated Border Elements.
5. The BE requests a list of possible termination Border Elements (TBE) and associated termination gatekeepers (TGK).
6. The clearinghouse requests each TBE and TGK on the list (L1) to generate a termination token for an associated termination gateway (TGTW). The TGK returns the L1 with the associated tokens.
7. The clearinghouse BE forwards L1 to the OGK along with a clearinghouse token.
8. The OGK initiates call setup with TGK and establishes the call.
9. The OGTW and TGTW signaling occurs and the call is initiated.
10. When the call ends, the TGTW signals the OGTW that the call is completed.
11. The OGK and TGK signal the clearinghouse, and each gatekeeper sends along its CDR for settlement to the clearinghouse.

providers are moving towards SIP and MGCP. Further, many other services such as IP multicasting, video-streaming, unified messaging will not be based on H-series protocols.

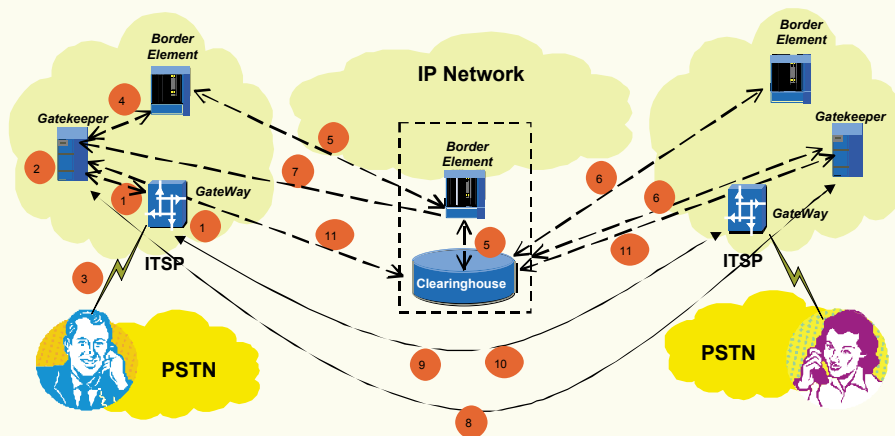
3. Gatekeeper and H.323 v2 dependence: iNOW requires gatekeeper-routed call signaling and H.323 v2 fast setup. In a network without gatekeepers (or gatekeepers that use H.323 v2), inter-domain billing cannot be done. Likewise, if the billing system acts as the radius server as well as the gatekeeper (e.g., in the Mind CTI application), a service provider would have to buy and use gatekeepers for settlements.
4. Standardization: Because iNOW takes advantage of gateway/gatekeeper interoperability and enables settlements on gatekeeper routed calls, it requires a certain degree of gateway compatibility (e.g., the same call detail record (CDR) format). Although Lucent and Vocaltec are behind the protocol, more than 160 gateway vendors have not adopted it.

The OSP solution - A Cisco Backed ETSI initiative

The iNOW approach is a bilateral-like response to the inter-domain VoIP billing problem. Such a solution can only support a cascading settlement model for voice service. This has driven the industry to investigate a standards-based settlement architecture, de-coupled from session level interoperability, that scales to multiple networks and multiple services.

The Open Settlement Protocol (OSP) involves a dedicated settlement server (an OSP server) to perform inter-domain authentication, authorization and reconciliation. VoIP companies pioneering this effort include Cisco and VIP Calling, with TransNexus and Gric Communications leading the development of OSP servers. OSP is intended to enable settlements between ITSPs that are members of a clearinghouse, and allow inter-clearinghouse settlements for any IP-based service.

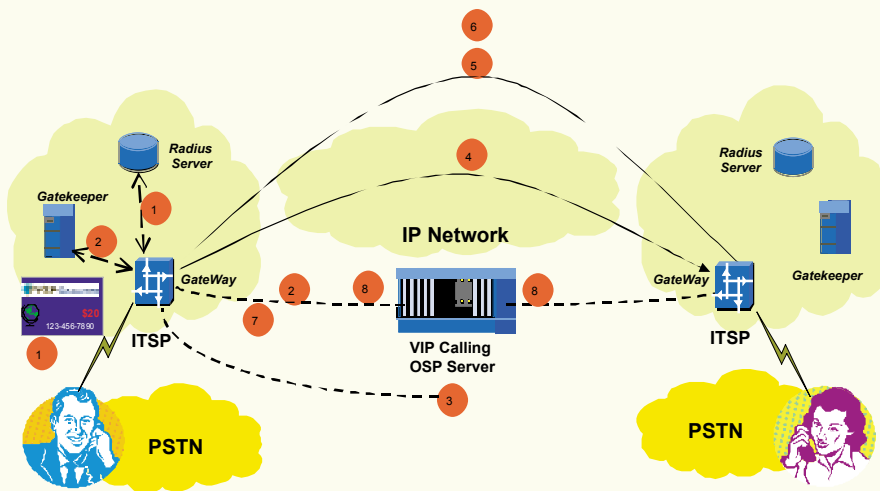
The OSP approach is being standardized by ETSI/TIPHON's working group 3.



Call Set-Up in an OSP Clearinghouse

The call flow for an OSP enabled clearinghouse call is as follows:

1. The originating gateway (OGTW) performs authentication and authorization on an in-bound inter-domain call.
2. The OGTW requests the destination digits. If the call's destination is within the ITSP's network, the local radius server handles the AAA. If the call is destined for a receiver outside the originating ITSP's network, the OGTW forwards the call to the OSP server, which performs the AAA.
3. The settlement server provides a settlement token, along with three IP addresses on which the call can be terminated, to the OGTW. These three addresses can be gateway or gatekeeper addresses, enabling an infinite number of gateways to be defined as endpoints.
4. The terminating gateway (TGTW) receives the settlement token with the destination digits and completes the call setup.
5. The OGTW and TGTW signaling occurs, the token is validated, and the call is initiated.
6. When the call ends, the TGTW signals the OGTW and settlement server that the call is completed. The TGTW then sends its CDR to the OSP server.
7. The OGTW adds detail to the settlement token confirming the transaction and sends the token to the settlement server.
8. The settlement server reconciles the call details it receives from both the OGTW and TGTW.



ETSI's (European Telecommunications Standards Institute) charter is to work on standardization of voice over IP communications, and it has established Project TIPHON — "Telecommunications and Internet Protocol Harmonization Over Networks" — to spearhead these efforts. OSP is based on the ETSI/TIPHON TS 101 321 specification for inter-domain pricing, authorization and usage exchange.

OSP provides a centralized response to the inter-domain billing problem (see OSP sidebar, left). The OSP model, unlike iNow!, natively supports a multi-service, multi-hop environment indicative of next-generation provider networks. At a high level, a call traversing multiple networks will consist of several network transactions and include various messaging elements. The typical XoIP call (where "X" is any QoS-based IP application), once originated, will require authentication, authorization, rating and accounting. This is fairly straightforward when the call has few legs, and the pricing involves AB rating in a single carrier's network.

In OSP, the following key transactions occur when a call is initiated and prior to the call's completion:

- **Authentication:** Once it is determined that the call will be terminated outside the originating network (from the dialed destination digits), the caller must first be authenticated to use the terminating network's facilities.
- **Authorization:** The settlement server needs to ascertain whether the call can be terminated on a peer network where a settlement relationship has been established. A set of possible terminating endpoints should then be provided by the server, and a token sent to a specified terminating endpoint.
- **Pricing:** Basic call rating information needs to be communicated. This may be established beforehand, but in cases where a guaranteed level of quality of service is required, it may be negotiated.
- **Usage:** A record detailing the service, including relevant session information like time stamps, endpoint IP addresses, caller ID information, and if necessary, level of QoS and Protocol is sent

to the settlement server for reconciliation. The case for the real-time communication of this service detail information is made when the level of sophistication of a service requires PSTN-like functionality.

For example:

1. *During call re-origination.* In this case, if a prepaid caller has completed a call and wants to make a follow-on call, they need not hang up. Some gateway platforms, such as the Cisco AS5300, support pound sign re-origination, allowing multiple calls on a single authentication event. This flexibility needs to be supported at the settlement server.
2. *Pre-paid card recharging:* In this case, a caller places a call but their card runs out of funds in the middle of the call. They should be able to recharge the card in the same session (i.e., by hitting *) and proceed when the card is refreshed.
3. *Fraud detection cases* — for example, simultaneous calls. In this case, use of the same PIN codes should be detectable and prevented. The same applies for velocity-fraud and other types of fraud.

OSP Messaging Protocols and Content

From an architectural point of view, the complexities that arise from supporting the functionality outlined in the transactions mentioned above and in the call flow detailed in the sidebar require a

comprehensive protocol-stack architecture. OSP uses HTTP, SSL (or TLS), TCP, and IP for the messaging required in the transactions (see below for the layered architecture defined by ETSI). In contrast with OSP, iNow! is implemented over UDP.

The actual messaging uses Extensible Markup Language (XML), which is rapidly emerging as the standard for handling client-server communication from a web-driven GUI. XML is extensible, can be parsed through firewalls, and uses the ubiquitous HTML as a subset. XML messages are broken down into an http header, the message content, and an optional digital signature. For the actual message content, the MIME (Multipurpose Internet Mail Extension) standard is used to communicate the components of transactions listed in the previous section (authorization, pricing, usage). In addition, each of these transactions follows a simple client-server protocol for bi-directional messaging.

Settlements beyond VoIP: XoIP and E-Commerce

Aside from IP voice, settlements are required at interconnection points for a wide range of value-added IP services. Of particular importance is electronic commerce, where the problem revolves around the settlement of content distribution. Because most service providers are also content aggregators, enhanced services such as messaging and telephony

can also be viewed as “content.” Making this observation, the problem can be described in the following way: when a subscriber who belongs to ISP-a’s network purchases a product or service that is available on ISP-b’s network, the subscriber currently pays the merchant who is the vendor of the particular product or service. The vendor (who is on ISP-b’s network) gets the revenue, and pays ISP-b for providing the storefront. This payment may or may not be transaction based, but is usually a flat fee. The service provider “owner” of the subscriber/purchaser (ISP-a) does not get a share of the revenue. This pricing model is unscaleable and unsustainable in a marketplace of value-added services, and bandwidth intensive content and applications.

If the Internet is ever going to move beyond an e-mail and web network, the settlement problem must be solved. By the nature of the Internet, one service provider can not hold all of the customers, nor can they provide the entire transit. There must be cooperation between ISPs, and an effective settlement, metering, and accounting infrastructure in place. IP telephony was the first application to test the Internet’s peering model. For lack of a better solution and standards, vendors such as Lucent and Vocaltec, and wholesalers such as ITXC developed iNow! to get things moving. Fortunately, the industry is making a serious effort to develop a scaleable and sustainable settlement infrastructure that addresses not only telephony, but also any IP-based service. 🌐

Philip Mutooni is the Back-Office product manager at VIP Calling, Inc. based in Burlington, Massachusetts. Philip holds two MS degrees from MIT in Electrical Engineering & Computer Science, and in Technology Policy. He can be reached at pmutooni@vipcalling.com

Dr. Matthew Lucas is an IP billing and multicast system consultant. He holds a Ph.D. in computer science from the University of Virginia and a B.S. in math/computer science from Carnegie-Mellon University. He can be reached at matt@telestrategies.com.

