File: wide-draft-ideon-overlay-discussion-02.pdf
Title: Discussions on Multi-Overlay Architecture over the Internet
Authors:
Yusuke DOI, ydoi@isl.rdc.toshiba.co.jp
Kenji SAITO, ks91@sfc.wide.ad.jp
Date: 2005-01-27

Abstract

We argue that the overlay network approach is a solution to numerous issues concerning the Internet such as mobility support, identification of entities without computing resources, security and multi/anycasting. The major advantage is that an overlay solution is essentially independent of other overlay solutions. Thus, there is no risk of conflict between solutions. Moreover, overlay can introduce tailor-made addressing and routing schemes for each application.

To realize overlays in the form of modular components for network applications, we focus on the interface between overlays. We propose an interconnection model that enables an overlay network for network exchange to act as a market for other networks.

We conclude that overloading the network layer to resolve issues is not only difficult, but also inimical to other solutions. To let the network layer as commons for future innovations, it will be necessary to consider alternative approaches such as an overlay model before we decide to overload.

1 Overlay Networking as a Solution

This paper discusses employment of overlay networks to resolve contemporary issues concerning the Internet. The issues include the mobility of entities that do not fit in a physical node, identification of targets without computing resources such as people, security in an untrustworthy network, trust management among a multitude of unknown nodes, and node-finding (rendezvous) in an ubiquitous networking environment.

Fundamental to our overlay approach is the management of identifier spaces that are independent of IP addresses. Such identifier spaces should be capable of providing applications with appropriate ways for rendezvous, location and routing.

We argue that decoupling identifiers and the Internet protocols is preferred to resolving the above issues within the network layer, which sometimes contradicts against the design of the layer and/or overload it.

1.1 Our Communication Model

The following is a description of the communication model we use at IDEON (Integrated Distributed Environment with Overlay Network), a working group in the WIDE Project that pursues autonomy in the designs of distributed systems.

We think that network designs should encourage self-generation of activities which utilize resources spread among different locations (hence, *integrated distributed environment*) by allowing spontaneous creation of layers of abstract network over the network layer (hence, *with overlay network*) .

Putting more stress on autonomy changes how the three ingredients of communication are performed:

1. Rendezvous (or how to identify the peer)

   The word "rendezvous" means a prearranged meeting place. In computer communications, such a meeting place can be a name space or a space for identifiers. Rendezvous performed autonomously allows spontaneous naming and resolution among the participating nodes.

2. Location (or how to locate the peer)

   This is to locate the node that represents the identifier. The node is typically identified by the identifier in the lower layer of the network, an IP address in most cases. Autonomous location involves identifying the closest copy of information among redundant copies spread over the network with help from participating nodes in the vicinity.

3. Routing (or how to reach the peer)

   This is to traverse the topology of the network so that a message can reach the peer. Location and routing can be done in the same procedure because it is necessary to traverse the topology of the network to locate the peer. Autonomous routing will involve creation of topology in an ad-hoc manner.

We propose alternative networking designs so that each of these can be performed in unrestrained and imaginative ways.

By "unrestrained and imaginative" we mean that no restraint should be imposed by the network as to which object can become the target for communication, without intervention of any authorities or privileged intermediate nodes, and that new ways of communication can be developed by the creativity of the participants of the network.

Since autonomy implies that there is no authority to guarantee the truthfulness of information (or that such an authority is weak), trust becomes an important issue.

In addition, the recent ubiquitous network environment makes the Internet heterogeneous and connected with real space. The only network we know of that has a variety of nodes is the current Internet, and the trend toward greater variety will continue as the mass of nodes increases. Real space is also integrated in the network. At least at 1991, we had a sensor connected to the worldwide network, the famous coffee pot[5] in Cambridge University. Many other projects, such as tangible bits[6], integrates real space and network.

2 Issues concerning the Internet

We think that an integrated distributed environment with overlay network constitutes an effective way of redesigning the current Internet. Thus, We consider issues concerning the current Internet to be opportunities to introduce new designs. In this section we focus on such problems and solutions.

2.1 Endpoint Naming, Locating, and Routing

There are many issues concerning message routing and endpoint granularity. In the current Internet architecture, almost all messages are routed into a physical network interface card (NIC), and the computer on the NIC hands the payload of the message to a process to which a port number is assigned

Binding between message routing and physical NIC is an essential constraint of the Internet. To overcome it, Roma Project[13], for example, creates its own routing scheme to route the message to a person regardless of its connectivity. This is an outstanding example of an application that uses the overlay network to decouple the endpoint of a message from the underlay network endpoint to satisfy the needs of applications.

Another kind of application introduces a more symbolic and abstract location such as "temperature of the room" or "energy consumption of this building." In many cases, endpoints of this kind do not exist. Some works of sensor fusion, eg. IrisNet[4], create virtual and abstracted nodes (organizing agent nodes) decoupled from physical sensor nodes. The virtual nodes collect data from one or many physical node dynamically. The data can be altered or integrated in the virtual node. And an overlay network structure provides naming, locating, and routing between client and such virtual nodes.

Decoupling of identifier spaces is a fundamental approach for node mobility, too. Mobile IP(MIP)[1] and LIN6[8] are efficient approaches for mobility and they decouple node addresses and network addresses. Although MIP is a standard of the Internet Protocol suite, MIP does not decouple name spaces of nodes and networks completely. On the other hand, LIN6 is desgined to decouple node identifiers and network identifiers. MIP and LIN6 are efficient because their routing is tightly coupled with the IP routing.

For mobilify, there are many alternative approaches that satisfy various needs. Due to strong relation with the IP-layer routing, MIP and LIN6 are not applicable to non-node mobility. For example, processes, contexts, or sessions can not move in the case of these solutions.

Some of contemporary approaches to node mobility uses an overlay network. Internet Indirection Infrastructure(i3)[12] is one of the best approaches for mobility using an overlay network. Their approach creates a concrete overlay network that is capable of naming, locating, and routing. In i3, each node forms a network of a distributed hash table and works as a message router over the network. Any nodes can "listen to" any points in the network. This approach decouples not only naming, but also locating and routing.

The major advantage over other mobility approaches is its flexibility. Because it has an independent routing and naming mechanism, not only physical hosts but also processes or other objects can handle its session as an overlay node. Moreover, ROAM(Robust Overlay Architecture for Mobility)[14], a mobility extension of the i3 approach, introduces a control mechanism of tradeoff between efficiency and privacy by controlling trigger insertion randomness.

2.2 Security and Overlay

Security and trust involve numerous issues that are difficult to resolve. In this paper we focus on two issues concerning security and trust, namely, node quarantine to keep infected or malicious nodes out of the network, and service selection

among unfamiliar and unknown nodes.

A node quarantine model for IPv6 node security was proposed in an internet-draft[7]. It involves control of the datalink layer such as VLAN or internet layer association (route advertisement in IPv6) to quarantine dubious nodes and protect other nodes from attacks. The approach is tightly coupled with the IP or datalink layer. Using some security audit mechanism, a node will be certified as "clean." And then it can enter the network.

Our approach can be applied in the same manner, but it utilizes a higher layer and decouple the secured state from the network layer. The regular network layer "outside" is considered to be dirty and the overlay network inside is kept clean.

Another issue concerning security and trust is service selection. To select a trustworthy service provider from all the providers around a client node, some sort of trust management among providers is necessary. At the same time, a service provider needs to distinguish malicious client nodes among accessing client nodes.

To satisfy these requirements, each node must be able to manage its own trust and that of others. A trust network, *i*-WAT[11], creates a peer-to-peer style overlay network to maintain and exchange trust between nodes. Using such a network, a node can decide which nodes to trust, or select a better one among unfamiliar nodes.

2.3 Overloading Causes Contradiction

As pointed out in the previous sections, the Internet involves numerous issues and needs. Clearly, the Internet Protocol is incapable of satisfying all those needs simultaneously. For example, mobility on or above the transport layer and addressing and locating a non-IP object such as a person are solved beyond the network layer.

Although design of network layer is limited to identification, locating, and routing of network interface, there are many demands to overload the Internet Protocol and IP addresses.

For example, an IPv6 address is a unique identifier in the Internet. Thus, re-using an IPv6 address as an identifier of a device, node, or person is an attractive idea. But such overloading is an abuse of the IPv6 address, which is not designed to support it.

Meanwhile, multicasting and anycasting overloads addressing and routing of IP. In multicasting, an address is shared among some set of interfaces and a special routing protocol is required to send messages to all of the interfaces. This essencially means that identifier of NIC is overloaded as a group of NICs, and the overloading enforces special handling of the identifier in every node participating in the Internet.

For example, multicasting requires PIM or another multicast routing protocol to be usable between sender and listener. Because not all routers support PIM and there is no evidence that PIM scales well against Internet-size, enabled area of multicasting is somewhat limited, or tunneling, a very primitive overlay, is applied to bridge between those enabled areas.

Anycasting is similar but involves more complex overloading against addressing and routing. In anycasting, interfaces share an address, as they do in

multicasting. The difference is that only one interface at a time can receive a message sent to the address.

In [10], it is pointed out that anycasting essentially involves issues of security and scalability. Although well-known anycast addresses for well-known services is an attractive idea for a service finding scheme, it also causes overloading of the IP routing.

Overloading is not always undesireble. However, it sometimes causes contradiction between two or more overloading technologies.

For example, ingress filter[3] conflicts with the earlier version of Mobile IP. Ingress filter prevents malicious attackers from spoofing its source address. However, a node with the earlier version of Mobile IP sends messages with its home address, and the messages are dropped by the filter. Due to this contradiction, Mobile IP for IP version 4 is required to use tunneling between a home agent and a mobile node to communicate. With redesign involving the network layer, Mobile IP version 6 behaves well against ingress filtering.

Great care must be exercised in overloading the network layer. Otherwise, contradiction with other technologies will undermine efficiency or, in the worst case, result in failure of deployment. At the same time, we believe this kind of overloading technology requires long and difficult standardization process (at present, mobile IPv6 draft revision is 24) to make sure that the technology is safe and harmless to other technology. Moreover, once standardized, the overloading technology becomes another obstacle to the introduction of new technologies.

3 Coordination of Multi-Thin Overlays

3.1 Multi-Thin vs. Single-Ultimate

It is necessary to investigate two architectures, namely, multi thin (MT) overlay network and single ultimate (SU) overlay network, and select whichever is better. The advantage of MT overlay network is its extendibility. For each new application, an overlay network can be created that has a naming and routing system optimized for the application.

At the same time, MT overlay introduces complexity. Lack of interconnectivity between MT overlays would be inconvenient. The resources on the overlay cannot be accessed from outside the overlay. To overcome this restriction, a syndication mechanism or a set of syndication mechanisms would be needed.

Pros and cons of SU overlay network are the verse of those of MT overlay. Since SU overlay introduces a widely applicable naming and routing framework, most applications would perform well within the framework. At the same time, however, a user wishing to use an application beyond the framework must modify the application or construct a new framework suitable for the application. One outstanding example of SU overlay network is Project JXTA (see `http://www.jxta.org/`).

3.2 Future Network Architecture with MT Style Overlays

To envision future network architecture, we believe MT style overlays are required in order to solve issues without introducing any contradction. However,

isolated multi-thin overlays do not utilize resources on the net well. With MT overlays a node on an MT overlay cannot use resources on another overlay.

Thus, interconnection of MT overlay is required to enable users to find better resources from all over the network. There are two approaches to interconnect MT overlays, namely the gateway approach and the client-side approach.

A gateway approach such as FLAPPS[9] introduces a protocol translation gateway or a proxy node in an overlay to gain access to other overlays.

Although gateway approaches are advantageous in that interconnection is transparent to clients, and thus implementation of a client software becomes easy, there is also a disadvantage.

There is a risk of contradction as same as overloading. If an application's requirement spoils another application's requirement for the gateway, implementation of the gateway becomes difficult.

In a client-side approach, everything is left up to the client process. There is nothing to interconnect multi overlays except user and pointer. URP (Uniform Rendezvous Pointer)[2] is a uniform pointer format for overlay networks. A URP is sufficient for a client to indicate which network to connect, how to search a resource in the network, etc. A URP has sufficient messages to make network be connected, to perform effective resource search, and so on.

A client program is capable of accessing resources on overlay network with plug-in software. When a client (or user) requires resource pointed by a URP, and it requires access to a network other than that which the client is currently on, it can obtain a new connection via corresponding plug-in. To trigger the interaction, resources in the overlays can contain URP to other networks.

The following is an example of a URP scenario. In the scenario, nodes A and B are on file exchange network X, and nodes A and C are on file exchange network Y. There is an *i*-WAT network to exchange values.

As node A gives its own work to node B, node A requests a certain amount of payment over the *i*-WAT network. The request contains a URP that describes how node B can access *i*-WAT node of node A. With the URP node B can complete its payment.

Then, node A wants another file from node C. Because node A has received a payment from node B, node A can pay for the file with node B's payment. With those interaction, *i*-WAT works as a market regardless of which file exchange network is used.

Another set of networks would be integrated to another kind of application. The advantage of client-side interconnection over the gateway approach is that an application developer can select a set of overlays including a newly created tailor-made network. Also, unlike in gateway approaches, since the interconnection point is isolated in client software, there is no risk of contradiction arising.

4 Conclusion and a Vision of the Future Internet

In this paper we discussed a vision of the future Internet. Our model of the renovated Internet consists of many kinds of independent overlays to fulfill requests for applications in a satisfactory manner.

We have described our communication model, which is a fundamental model of overlay, and have applied it to resolve various issues concerning the Internet.

Although overloading the network layer is an efficient way of resolving issues, we argue that overloading sacrifices innovation to some extent. Contradictions will arise between other technologies and the overloading technology, dictating the choice of an ineffective way of finishing the standardization process.

From our perspective, the Internet is sufficient if the followings are satisfied.

- Routeability: messages sent to an interface should arrive at the interface
- Efficiency: routing between each interface reflects the structure of the datalink layer configuration as precisely as possible
- Scalability: more than 10 to 1000 times the Earth's population are able to connect to one another

The final item is rough estimation of the ideal ubiquitous network environment and sensor network scenario. In the ideal scenario, every location has intelligence to support the user's activity.

Technologies to overload the Internet Protocol, paticularly in regard to routing, are inimical to such tremendous scalability. Prior to clarifying an approach capable of realizing scalability of this order, it will be necessary to consider alternative approaches such as an overlay model.

The software development process requires a modular structure to enable readability and suppress maintenance costs, rather than monolithic efficiency. Spurred by the recent introduction of more powerful computer hardware, networked applications can become more modularized to enable flexibility and simplicity. If the modularized approach proves to be effective, it will be necessary to focus on the interface between protocols, not on the protocol itself. The overlay approach is found to be highly effective for modularizing protocols. Thus, thorough investigation of the interface between overlays can facilitate the sound development of the Internet.

Copyright Notice

References

[1] C.Perkins. Mobility support in ipv6. Internet-Drafts (draft-ietf-mobileip-ipv6-24.txt), June 2003.

[2] Yusuke DOI, Takaaki ISHIDA, Kenji SAITO, and Youki KADOBAYASHI. Uniform Rendezvous Pointer: Identifier space for overlay network interreference (in Japanese). In *IPSJ SIGNotes Quality Aware Internet*. IPSJ, July 2003.

[3] Paul Ferguson and Daniel Senie. Network ingress filtering: Defeating denial of service attacks which employ ip source address spoofing. IETF RFC 2827, May 2000.

[4] Phillip B. Gibbons, Brad Karp, Yan Ke, Suman Nath, and Srinivasan Seshan. IrisNet: An architecture for a worldwide sensor web. *IEEE Pervasive Computing*, 2(4):22–33, October-December 2003.

[5] D. Gordon and M. Johnson. The trojan room coffee pot. http://www.cl.cam.ac.uk/coffee/coffee.html.

[6] Hiroshi Ishii and Brygg Ullmer. Tangible bits: Towards seamless interfaces between people, bits and atoms. In *CHI*, pages 234–241, 1997.

[7] Satoshi KONDO and Shinsuke SUZUKI. Quarantine model overview for ipv6 network security. Internet-Drafts (draft-kondo-quarantine-overview-00.txt), Feb 2004.

[8] Mitsunobu Kunishi, Masahiro Ishiyama, Keisuke Uehara, Hiroaki Esaki, and Fumio Teraoka. LIN6: A new approach to mobility support in IPv6. In *International Symposium on Wireless Personal Multimedia Communication*, 2000.

[9] B. Scott Michel, J. Dharap, R. Xu, and P. Reiher. General purpose infrastructure for networked peer-to-peer services. In *IEEE INFOCOM 2003.*, 2003.

[10] Masafumi OE and Suguru YAMAGUCHI. Implementation and evaluation of ipv6 anycast. In *Proc. of INET 2000*, June 2000.

[11] Kenji Saito. Peer-to-peer money: Free currency over the Internet. In *Proceedings of the Second International Human.Society@Internet Conference (HSI 2003), Lecture Notes in Computer Science*. Springer-Verlag, 2003.

[12] I. Stoica, D. Adkins, S. Zhuang, S. Shenker, and S. Surana. Internet indirection infrastructure. In *Proceedings of ACM SIGCOMM*, August 2002.

[13] E. Swierk, E. Kiciman, V. Laviano, and M. Baker. The roma personal metadata service. In *Third IEEE Workshop on mobile Computing Systems and Applications*, December 2000.

[14] Sehlly Q. Zhuang, Kevin Lai, Ion Stoica, Randy H. Katz, and Scott Shenker. Host mobility using an internet indirection infrastructure. In *First Internet Conference on Mobile Systems, Applications, and Services (ACM/USENIX Mobisys)*, May 2003.