# End-Middle-End Architecture for the Internet

Olli-Pekka Lamminen
TKK Networking Laboratory
*olli-pekka@netlab.tkk.fi*

## Abstract

*End-to-end traffic model in the Internet has been broken by firewalls, NAT boxes and such. This has created a need for new Internet architectures accounting the presence of these middleboxes and involving them in connection negotiations. This paper reviews the work of IRTF EME research group involving end-middle-end architectures, and a promising EME architecture called NUTSS. We discuss these works in light of our own ideas, and provide reasoning for trust being an essential component in the success of any new Internet architectures.*

**Keywords** Internet architecture, future Internet, end-middle-end, middleboxes

## 1 Introduction

The Internet has evolved beyond the boundaries of its original design. The networks have grown and the number of stakeholders in traffic has increased drastically. The original end-to-end traffic model has been broken by techniques like NAT, and devices such as firewalls and proxies wanting to have their say in the traffic outside endpoints' control.

This kind of development in networks was at first viewed as negative [8], but lately the attitudes have changed, and many believe that it is not only inevitable, but actually required functionality [1, 11]. Various research groups have been researching ways to have devices and applications along the network data path to be more involved in connections and connection establishment [7, 9, 11].

The purpose of this paper is to give an overview of the work done by different research groups on end-middle-end (EME) architectures. Our original contribution is a feasibility analysis of EME architectures as described in workings of IRTF EME research group (IRTF EME) [6, 9].

In the following section, section 2, we give a brief description of EME architecture and current work status of IRTF EME. In section 3 we review a prominent EME architecture called NUTTS [7] and discuss its feasibility against IRTF EME guidelines. In section 4 we give our considerations for EME architectures and argue that trust relations will become a key component in future networks. We conclude this paper in section 5 with discussion of related work and our current research plans.

## 2 End-Middle-End Architecture

End-middle-end architecture offers endpoints a level of control over which middleboxes their traffic is passed through. This control can be extended to cover routing but is mainly intended for enabling network provided services, like firewalls, proxies and NAT.

In this section we first give a brief overview of EME architecture and then present the goals and guidelines of the IRTF EME Research Group.

### 2.1 Architecture Overview

The idea behind EME architecture is to enhance the level of control offered by networks to better meet the needs of the current Internet. Running out of IPv4 addresses and subnets has caused network and service providers to adapt technologies to combat address exhaustion. One suggested solution, IPv6 [3], has not yet had its breakthrough, and it is still uncertain if it ever will. Industry has adopted NAT as a solution of choice instead of IPv6. NAT, however, breaks end-to-end addressing semantics of the Internet [4].

Number of stakeholders in network connections has grown. Not only are the endpoints involved, but the network itself has become increasingly aware of the traffic passing through it. Middleboxes with NAT and firewalls are used to police and control network traffic, while operators implement cache proxies to reduce the volume of traffic between networks. A way for endpoints to discover, request and utilise these and other services provided by the network is required.

Various different approaches to accommodate middleboxes into Internet architecture have been proposed. The suggestions vary from new signalling and flow establishment methodologies [7, 9] to redesigns of Internet naming and name resolution [5, 10]. Common themes among these suggestions are enhanced access control and authentication, and extensions to naming conventions. Most of the problems arise from feasibility of the deployment, especially if the proposed changes are not incrementally deployable.

### 2.2 IRTF EME Research Group

IRTF End-Middle-End Research Group was formed in late 2006. The intent of this group is to function as a forum for EME related discussion and research. The group's charter identifies three common root problems that EME research should try to solve. First on the list of problems is IP addresses no longer being globally-unique or stable. Second problem is with transport port numbers having no clear semantics beyond the endpoint opening connection. Third problem is endpoints not being aware of middleboxes along traffic route and thus not being able to control or even know what is happening to their traffic [6, 9].

### 2.3 EME Architecture Requirements

IRTF EME has set evaluation of feasibility and desirability of any architectural changes proposed to solve these problems as its goal. To achieve these goals IRTF EME has come up with a list of requirements a suggested Internet architecture should fulfil. This list includes requirements for authentication, security, privacy, control, steering, mobility, interoperability, any and multicast, and performance [6].

First up on the list is a requirement for globally-unique, long-term stable, and user friendly endpoint identities to allow easier policy control for endpoint users and network administrators alike. Next on the list are requirements for access control and authentication. Requirement for access control is that the network allows endpoint administrators control over which other endpoints and middleboxes the endpoint can communicate with. Authentication is required between endpoints and middleboxes and must be provided in both ways. These requirements are in place with network policing in mind, and to provide a level of security against spoofing attacks and impersonation.

The list continues with a privacy requirement stating that any confidential information should be only revealed to trusted parties. The endpoints need to be allowed to communicate anonymously, or the middleboxes need to provide a way to anonymise the traffic between endpoints. Endpoints must not be required to reveal their network addresses to untrusted parties, and all parties must be allowed to require flow level encryption if full path authorisation cannot be guaranteed.

Requirements related to network control include middlebox discovery, flow redirection, protocol negotiation and multi-homing. Middlebox discovery includes ability for endpoints to discover middleboxes and request their services when desired. Middleboxes which require their services to be used, e.g. NAT devices or firewalls, need to inform endpoints of their services.

Flow redirection requirements are built with mobility in mind. Endpoints must be allowed to redirect incoming flows without the need for the initiating application to intervene. Middleboxes must be able to redirect inbound traffic to an alternate endpoint or an alternate address for an endpoint without intervention from the endpoint applications. Finally, the network must be able to maintain and reroute flows between mobile endpoints.

Protocol negotiation means that endpoints must be able to negotiate the protocol stack, e.g. UDP or TCP delivery, for a flow based on application requirements and network policy. Multi-homing requirement states that both endpoints and middleboxes must be able to specify the routes for flows, and the network has to support multiple simultaneous routes.

The network must be able to support multicast flows, with a fallback option when IP multicast functionality is not present. The network also needs to support fast flow establishment, i.e. be optimised for short flows on high-latency networks. If possible, the first transmitted packet should be allowed to contain application payload.

The final requirement on the IRTF EME list is possibility for incremental deployment of the new architecture. IRTF EME recognises that an overnight overhaul of the Internet is not a valid option. Instead all changes to endpoints and middleboxes to accommodate the new architecture need to be implemented gradually. Also an incentive for migration has to be present, if not initially then at least for any architecture close to deployment.

## 3 NUTSS

NUTSS [7] is an architecture and a protocol designed to satisfy EME naming and addressing requirements. Its core idea is to give endpoints user- friendly names, and use signalling protocols to later bind these names to 5-tuple transport flows. The transport flows in NUTSS are short-lived and renegotiated using both on-path and off-path signalling protocols.

Connection establishment in NUTSS is initiated by one endpoint via off-path signalling protocol followed by on-path signalling to open data path through possible middleboxes on route. The initiating signalling uses both endpoints' names combined with requesting application's name. These names are mapped into 5-tuple flow addresses by policy-aware boxes near both endpoints. These addresses are then conveyed to both endpoints, who use the addresses to form a data connection.

NUTSS achieves separation between identification and network location by using stable names separated from flow addresses. This allows the architecture to support both mobility and multi-homing. The signalling protocols used both on-path and off-path allow

endpoints negotiate the used protocol stack with each other and any middleboxes involved in the communications, allowing flexible managing of multiple network layers.

This section describes the basic architecture of NUTSS, possible extensions as outlined by the original authors [7], and discusses its feasibility in regard to IRTF EME guidelines.

## 3.1 NUTSS Architecture

NUTSS uses user-friendly, long-term stable, and location independent naming. These names are associated with applications and services running on endpoint devices, and can be further associated to individual user level. Connection establishment is initiated via a NUTSS socket using only these names. Opening the socket triggers name-based signalling, that authenticates the endpoints and establishes a data path with 5-tuple IP addressing via a series of middleboxes.

The NUTSS architecture has two major components, called P-Boxes (policy-boxes) and M-Boxes (middleboxes). These components are deployed in the networks as well as in end-hosts. Any network on data path wishing to enforce a policy must deploy both P-Boxes and M-Boxes. P-Boxes form an overlay network to carry name-routed signalling from end to end but do not relay data flows. P-Boxes are used to make policy decisions, and direct data flows through M-Boxes if required. Flow data is transmitted via M-Boxes, either because the M-Box lies on the direct data path, or because the initiating signalling has negotiated an alternative route through an M-Box.

Signalling messages may traverse both P-Boxes and M-Boxes. When no IP address is known, name-routed signalling traverses via P-Boxes. Otherwise the signalling is routed via M-Boxes similar to regular data flows. A name-routed path via P-Boxes between endpoints must always exist, even when techniques like NAT are used. This allows end-to-end signalling to reach both endpoints before data connection is established.

### 3.1.1 Naming and Name-Routed Signalling

NUTSS uses 3-tuple endpoint names, consisting of *user*, *domain* and *service* name. The *user* is a user-friendly name that is not globally-unique, e.g. *alice*. The *domain* is a user-friendly, hierarchical, globally-unique DNS name, identifying the endpoint machine, e.g. *bob.org*. Together the *user* and *domain* identify the *principal*, that is considered to own the endpoint. *User* may be empty, in which case the *principal* identifies to endpoint machine. The *service* is a globally-unique, user-friendly name identifying the service provided by the endpoint.

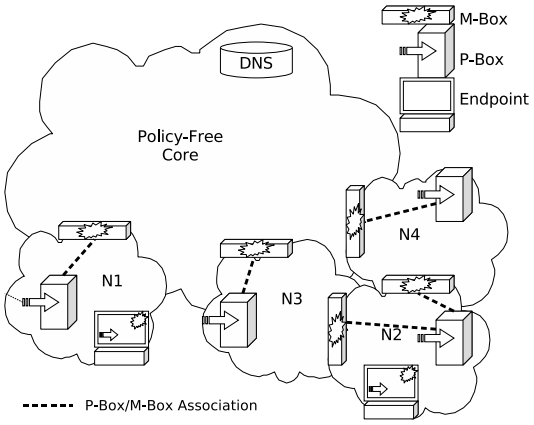


**Figure 1:** NUTSS network topology.

All names are independent of network location. Access control policies are defined in relation to these names.

NUTSS models the Internet as a group of policy-aware edge networks connected by a policy-free core. Edge networks deploy both P-Boxes and M-Boxes, while the core is P-Box free. Every network M-Boxes lie usually between networks connecting them to each other, but can be deployed inside the networks as well. Each M-Box is configured with the name and address of its associated P- Box. NUTSS assumes the presence of DNS or similar name-resolution service in the network. The name service contains addresses for contact P-Boxes in every top level edge network called domain. These P-Boxes must be globally addressable. Every endpoint contains an inbuilt P-Box and M-Box as well. NUTSS network topology is pictured in figure 1.

P-Boxes are organised in a tree-like hierarchy, forming an overlay network. Every P-Box in a network not directly connected to the core has a designated parent P-Box. The parent P-Boxes are discovered through M-Box referral mechanism, where the child P-Box asks sends an address-routed message to a public address, and the parent network's M-Box intercepts this message and responds with the information of the parent P-Box.

Upon joining a network, an endpoint registers with its local P-Box. When receiving registration request, the P-Box adds the endpoint into its registration table, and associates an address with the endpoint. If the P-Box has any parents, it propagates this address mapping recursively to all of its parents. If the topmost P-Box is a contact P-Box of the domain the endpoint wants to associate with, the registration is now complete. If not, then the topmost P-Box forwards the registration request to the contact P-Box of the endpoint's chosen domain.

Name-routed signalling happens via the overlay network formed by P-Boxes. When an endpoint wishes to

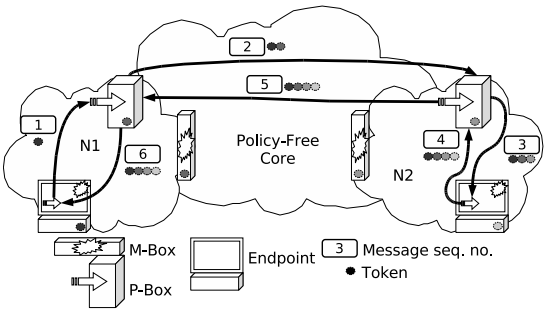


**Figure 2:** Connection initiation with name-routed signalling.

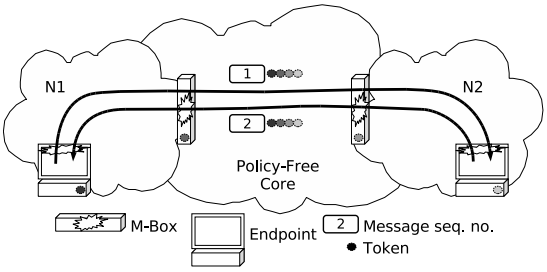


**Figure 3:** Data flow establishment over address-routed signalling.

initiate a connection with another endpoint it sends a initiation message to its parent P-Box. If the parent P-Box does not know the location of the destination endpoint, it forwards the initiation message to either its parent P-Box, or if not applicable, the contact P-Box of the destination's name-routed domain, in which case the tree traversal proceeds downwards towards child P-Boxes. This is repeated until the message reaches a P-Box, which knows the destination's location. Before the forwarding, the P- Box adds a next-hop token to the initiation message. Once the destination endpoint's location is known the initiation message contains tokens for the reversed name-route between the endpoints. The process of connection initiation is shown in figure 2.

#### 3.1.2 Address-Routed Messages

Address-routed path is established by endpoints using peer address and the next-hop tokens added by P-Boxes during name-routed signalling. During the path establishment, any necessary per-flow state changes to M-Boxes on route are installed. Each M-Box on route checks the presence of a next-hop token corresponding to it, and forwards the data based on these tokens. This process is presented in figure 3.

### 3.2 Feasibility

NUTSS incorporates many of the aspects required by IRTF EME [6]. The names used in name-based routing are both user-friendly and stable, and have globally-unique parts to ease policy management. Access control is implemented by P-Boxes, and authentication, while not addressed in the design paper [7], is implementable with standard techniques, such as challenge-response protocols like DIAMETER [2], in the registration and signalling protocols. Privacy aspects are not discussed in the design paper, but can be addressed with P-Box and M-Box policies as well.

Middlebox control is built in the signalling protocols NUTSS uses. Route discovery reveals any NUTSS-enabled middleboxes on the data route. Service requests can be achieved with next-hop tokens. Steering control and mobility are achieved by short-lived flow addresses and rapid flow renegotiation. This does not cater for long flows, and as such might require redesign of current protocol stacks. The name-route and address-route signalling protocols allow negotiation of protocol stacks as well. This makes it possible for NUTSS to support any current protocols running over IP.

Multi-homing follows naturally from the separation of name and location. Multicast is not discussed in the original design, but a multicast extension for the naming is mentioned. Basically this involves extending the 3-tuple name to a 4-tuple, where the fourth field defines multicast group.

We feel that the performance is the main stumbling block of the NUTSS architecture. Double signalling during connection establishment speaks against fast flow establishment and optimisation for short flows, and the route discovery mechanism does not allow transmitting payload in the initial packets. Meanwhile the short-lived flow addresses prohibit long data flows. We feel that this kind of design might cause unnecessarily high signalling overhead and complexity.

The incremental deployment scheme suggested in [7] has three phases. In the first phase only the endpoints are NUTSS-aware and a public P-Box services are provided by third parties. In the second phase the network middle is gradually made NUTSS-aware by deploying P-Boxes in networks. During the final phase, networks replace legacy middleboxes with M-Boxes. It is unclear to us in what amounts these phases may overlap each other, and how much work is involved in the final phase of implementation, if the M-Boxes need to be made aware of any legacy middleboxes in the network.

# 4 Considerations for EME Architecture

With the ongoing development in Internet it has become apparent that middleboxes have come to stay. This means that the need for including middleboxes in the connection negotiation process has become necessary. This section discusses a few key aspects of middlebox-aware networks and gives complementing thoughts for EME architectures, especially when trust issues should be considered.

## 4.1 On Naming and Name Resolution

A network needs two types of names for endpoints; user-friendly names, that are easy to remember and provide stability even when the endpoint changes location, and machine-friendly names (e.g. addresses), which are fast to process by networked applications. Names should also be globally-unique, but this requirement brings with it a constraint for the network: for a name to be unique there needs to be an authority which allows and registers the use of a name within the network. Without such an authority no method to guarantee an endpoint authenticity exists. This requirement makes the presence of a name resolution service mandatory.

Mobility is a growing trend among network devices. This creates an additional requirement for a name resolution service: the devices need to be able to update their location in the network. This means that either the device itself needs to update its information, or that the network needs to be aware of the connecting devices and update their information accordingly. As the basic end-to-end network functionality requires the network to be aware of a connecting device, and as it is easier for a name resolution service to trust a network instead of an individual device, it should be the network who updates the name resolution. For added security, the network should have means to authenticate the connecting device either by itself or via a third party service provider.

## 4.2 On Middlebox and Connection Control

Middleboxes are a great way for a network to police and provide additional services for passing traffic. A network willing to provide additional services should be offered a way to advertise them to endpoints. It is worth remembering that while a network might want to advertise the middleboxes it deploys, a network cannot be prevented from deploying hidden middleboxes. The possibility of a network advertising middleboxes is enough to require any new Internet architecture to provide means for controlling and requesting middleboxes along the data path.

Requesting middleboxes comes with the price of increased signalling. If endpoints want services from middleboxes, the middleboxes need to be provided with means to advertise themselves, and the endpoints need to be provided with means to request these services. Fast flow negotiation can be achieved by including a mechanism to re-route data flows to pass through middleboxes whose service has been requested after the initial data flow has been formed. If a network on the data path deploys mandatory middleboxes it can be argued that the endpoints should be notified of such middleboxes. The possibility of hidden middleboxes makes this argument moot; as such notifications cannot be enforced, it is enough for polite middleboxes to inform the endpoints of their existence during the flow formation, and thus allow the endpoints to tear down the connection if they feel the network cannot be trusted anymore. If the endpoints cannot trust the network it is their responsibility, and not the network's, to provide sufficient protection for their connection.

## 4.3 On Trust, Security and Privacy

Globally-unique names, updatable name resolution service, middlebox services, and other advanced network services bring forth questions of security and trust. Negotiating trust and trust relations are a key components in providing a secure networking environment. Enabling mobile endpoints to seamlessly move between networks enhances this problem even more. The trust needs to be placed somewhere but it is unclear where. The endpoints cannot always fully trust the networks they are connected in, and the networks cannot trust the endpoints even that far. One solution for establishing trust is to include third party trust authorities in the network. Service providers are a natural choice for this role.

Another thing complicating trust is privacy. An endpoint might want to shield its identity from the network, much like NAT [4] does today. Such shielding of identity might even be extended to the network the endpoint is directly connected in. Once again a third party trust authority can be used to facilitate the connection in a manner both participants can be comfortable with.

# 5 Summary and Future Work

Current work of the IRTF EME research group focuses on the design of signalling protocols, specifically NUTSS. This solution complements IP, but leaves most of the underlying problems of IP networks untouched. The groups intent is to create an efficient way to request and identify middleboxes on the data path, allowing the middle to be fully incorporated in the network architecture.

Our current and future research focuses on creating an operator friendly alternative for IP networks, with considerations of security, network services and trust in mind. While replacing IP might not be feasible in consumer networks it can be feasible in core networks, where IP can be transported over this new technology where necessary. Key objectives of our research are separating services from the transport layer and establishing trust relations between networks and endpoints to achieve carrier-grade transport.

## Acknowledgements

We would like to thank Raimo Kantola and Marko Luoma for the inspiring discussions on post-IP networks that gave us a new perspective for future Internet architectures. We would also like to thank Mika Ilvesmäki for his comments and review.

## References

[1] BLUMENTHAL, M. S., AND CLARK, D. D. Rethinking the Design of the Internet: The End-to-End Arguments vs. the Brave New World. ACM Transactions on Internet Technology, Vol. 1, No. 1, pp. 70-109. August 2001.

[2] CALHOUN, P., LOUGHNEY, J., GUTTMAN, E., ZORN, G., AND ARKKO, J. RFC 3588 - Diameter Base Protocol. September 2003

[3] DEERING, S., AND HINDEN, R. RFC 2460 - Internet Protocol, Version 6 (IPv6) Specification. December 1998.

[4] EGEVANG, K., AND FRANCIS, P. RFC 1631 - The IP Network Address Translator (NAT). May 1994.

[5] GRITTER, M., AND CHERITON, D. R. TRIAD: A New Next-Generation Internet Architecture. [http://www-dsg.stanford.edu/triad/] *(ref. 2007-10-11)*. July 2000.

[6] GUHA, S., AND FRANCIS, P. Requirements for the End-Middle-End Research Group. [http://www3.tools.ietf.org/group/irtf/trac/wiki/EME_Charter] *(ref. 2007-11-21)*. November 2006.

[7] GUHA, S., AND FRANCIS, P. An End-Middle-End Approach to Connection Establishment. In Proceedings of SIGCOMM'07. August 2007.

[8] HAIN, T. RFC 2993 - Architectural Implications of NAT. November 2000.

[9] HANDLEY, M., AND FRANCIS, P. Charter for the End-Middle-End Research Group (EME). [http://www3.tools.ietf.org/group/irtf/trac/wiki/EME_Charter] *(ref. 2007-11-21)*. November 2006.

[10] KOPONEN, K., CHAWLA, M., CHUN, B-G., ERMOLINSKIY, A., KIM, K. H., SHENKER, S., AND STOICA, I. A Data-Oriented (and Beyond) Network Architecture. In Proceedings of SIGCOMM'07. August 2007.

[11] WALFISH, M., STRIBLING, J., KROHN, M., BALAKRHISNAN, H., MORRIS, R., AND SHENKER, S. Middleboxes No Longer Considered Harmful. In Proceedings of OSDI'04. December 2004.