

Scalability Analysis of the TurfNet Naming and Routing Architecture

Jordi Pujol, Stefan Schmid, Lars Eggert, Marcus Brunner and Jürgen Quittek

NEC Europe Ltd., Network Laboratories, Kurfürstenanlage 36, 69115 Heidelberg, Germany
{pujol | schmid | eggert | brunner | quittek}@netlab.nec.de

Abstract — TurfNet is a novel internetworking architecture that enables communication among highly autonomous and heterogeneous network domains. The architecture uses a global identity namespace and does not require global addressing or a shared internetworking protocol. It integrates the new concept of dynamic network composition with other recent architectural concepts, such as decoupling locators from identifiers. This paper examines whether TurfNet’s naming and inter-domain routing architecture can scale to networks of the size of the global Internet. The paper uses existing research into the topology of the Internet’s autonomous system graph and related results that quantify typical traffic patterns to analyze the scalability and performance of the TurfNet architecture on similar internetwork topologies.

I. INTRODUCTION

The basic principles of the original Internet architecture include end-to-end addressing, global routeability and a single namespace of IP addresses that are locators and host identifiers at the same time. These principles are suitable for static and well-managed flat network hierarchies. However, as the Internet evolved from a small research network to a worldwide information exchange, a growing diversity of commercial, social, ethnic, and governmental interests led to increasingly conflicting requirements among the competing stakeholders. These conflicts create tensions that the original Internet architecture struggles to withstand. Clark *et al.* refer to this development as “tussles in cyberspace” [1]. This development has prompted research into different internetworking architectures, such as FARA [2], Plutarch [3], Triad [4], IPNL [5] or 4+4 [10].

The TurfNet architecture addresses this trend by enabling interoperation between otherwise autonomous networks [11][13]. These autonomous networks are modularized according to the inherent boundaries drawn by the different interests of the involved stakeholders. As proposed by Braden [6], TurfNet “minimizes the degree of required global architectural consistency” by relaxing the requirement that all connected networks must use the same internetworking protocol together with a globally coordinated addressing scheme. Instead, TurfNet enables autonomous networks that use different network protocols and addressing schemes to interoperate.

Earlier papers [11][13] have described the details of the TurfNet architecture and Section II briefly summarizes the key mechanisms. The focus of this paper is an analytical

scalability study of TurfNet as a large-scale internetworking architecture. It studies the performance of the TurfNet naming and inter-domain routing mechanisms when they operate on realistic, large-scale internetwork topologies with billions of nodes. The internetwork topology and communication patterns are based on existing analyses of the Internet’s autonomous system (AS) graph and inter-AS communication. Section III discusses these Internet characteristics, Section IV presents the model used as the basis for this scalability analysis and Section V analyzes the behavior of the TurfNet architecture under the given model. Finally, Section VI summarizes and concludes this paper with an outlook on future work.

II. TURFNET OVERVIEW

One key architectural feature of the TurfNet architecture is the explicit separation of host identities and host locators, similar to HIP [7], multi6 [8], SNF [9] or DOA [12]. TurfNet introduces a new host identity namespace that enables the use of different addressing and routing mechanisms in each individual autonomous network.

Network composition is a second, new concept central to the TurfNet architecture. Two different variants of composition exist, *vertical* composition and *horizontal* composition. When Turfs compose vertically, one of the composing networks takes on a connectivity provider role for the other “customer” Turfs in the composition. Vertical network composition encapsulates administrative, control and routing functionalities, and isolates network-internal structures. Horizontal composition is an alternative way for networks to compose. It is the preferred composition variant when networks do not have an intrinsic customer-provider relationship. Horizontal composition is therefore also referred to as “peering” and is preferred, for example, between two personal-area networks or between service provider networks that establish a direct peering agreement.

A common node registration and lookup service that operates in and across all TurfNets enables inter-TurfNet communication – regardless of potential technological and administrative differences. The registration mechanism announces the reachability of TurfNodes outside their local Turfs. The lookup mechanism locates an announced TurfNode before communication takes place. Both mechanisms contribute to the dynamic creation of end-to-end communication paths; different registration and lookup mechanisms result in different routing paths, increasing the flexibility of the architecture. Inter-Turf gateways perform locator and protocol translation on packets that traverse between the different autonomous Turfs. End-to-end communication across Turf boundaries is thus a product of the following processes: node registration and node lookup. The

This paper is a byproduct of *Ambient Networks*, a research project partly supported by the European Commission under its *Sixth Framework Program*. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the *Ambient Networks* project or the European Commission.

performance of these two mechanisms significantly affects the scalability behavior of the overall architecture.

A TurfNode becomes reachable to other TurfNodes by registering its local address with its Turf-local lookup service. This registration propagates vertically through the “hierarchy of composed Turfs” to achieve Turf-external reachability (see Fig. 1). Turfs always forward non-local registration messages to their vertically composed parents, resulting in a system where lookups are guaranteed to succeed at the root level, but may of course succeed at a lower level, if it caches the looked-up registration. In addition, Turfs can also forward registrations horizontally to peer Turfs as an optimization [13]. The *scope* of such a horizontal lookup limits the propagation of the request to peers that are within a given number of hops. Note that a Turf may change the scope of a request that it forwards as it sees fit.

Likewise, for end-to-end communication between TurfNodes located in different Turfs, the local lookup services propagate any lookup request for non-local peer nodes to the vertically composed parent Turfs, which then try to resolve the requested host identity within their respective domains. As with registrations, Turfs may also forward a lookup request to horizontally composed peers with a certain scope as an optimization. For successful node resolutions, the Turfs along the lookup path configure their gateways to allocate proxy addresses and install the necessary translation state between the different address spaces and/or network protocols. Fig. 1 illustrates how these mechanisms work together to establish an end-to-end communication path through a hierarchy of composed TurfNets.

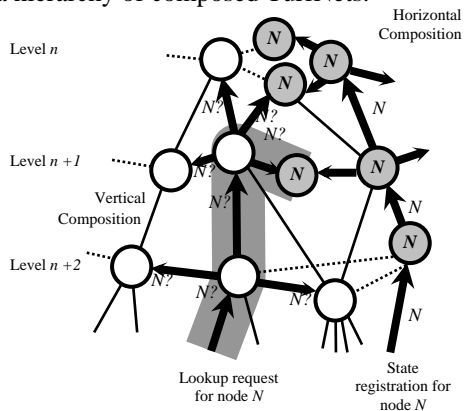


Fig. 1. Successful registration and lookup operations pin a routing path through a TurfNet hierarchy.

III. INTERNET CHARACTERISTICS

To some extent, Turfs are comparable to the Internet’s autonomous systems. Both are administratively independent sub-domains in a global structure. TurfNet gateways are similar to the Internet’s border gateways that connect together different autonomous systems. Furthermore, the relationship between two autonomous systems – *peering* or *customer/provider* – is similar to TurfNet’s horizontal and vertical compositions. However, the main difference between the two architectures is that TurfNet gateways may also perform locator and protocol translation, which increases the autonomy of individual Turfs, because they may use different internal network protocols and address allocation mechanisms.

This paper analyses how TurfNet performs when the hierarchy of composed TurfNets is similar in size and structure to the Internet’s AS-level topology. It consequently models the AS-level topology of the Internet as a TurfNet hierarchy that represents every AS with a separate Turf. This modeling process uses existing studies of the Internet topology [14-16]. In addition, this paper assumes that the inter-Turf communication patterns in such a large-scale topology are similar to those observable in the Internet [17-20].

Existing research results extract the AS-level topology of the Internet from the routing tables of its Border Gateway Protocol (BGP) [21][22]. Other papers apply heuristics to the AS-level graph to determine the type of inter-AS relationships (peering or customer/provider) between any two connected autonomous systems [22]. This information can help classify autonomous systems into *tiers* or *levels*, from 1 (the highest) to 5 (the lowest). Thus, end systems (hosts) are located at the bottom level 5. Table 1 illustrates some properties of the hierarchical graph obtained through this process.

Level	Number of Autonomous Systems	Mean Path Length	Diameter	Mean Number of Peers per AS
1 (highest)	22	1.25	2	24.2
2	215	3.90	10	5.7
3	1391	1.98	11	1.0
4	1421	0	0	0
5 (lowest)	13872	0	0	0

Table 1. Properties of each AS level.

Table 1 illustrates that the autonomous systems at the first level are almost fully connected. Thus, a level-1 AS can usually reach any other level-1 AS in one hop; in rare cases, two hops may be required, because the graph diameter is 2. At level 2, 193 out of the 215 ASs belong to the same connected cloud, even though the mean path length is significantly longer (3.9 vs. 1.25). Level 3 consists of a single cloud encompassing all 243 ASs, but most ASs have zero peers (456 ASs) or just a single peer (269 ASs). At levels 4 and 5, no intra-level peering exists; hence, the mean path length and diameter properties are 0.

In order to make realistic assumptions about traffic patterns in a large-scale TurfNet topology, the analysis uses the results of previous studies that have investigated corresponding traffic in the Internet [17-20]. They show that the mean end-to-end communication only traverses between 3 and 4 ASs. This paper assumes that equivalent probability profile is likely to be present in a comparable TurfNet topology. Table 2 shows the communication distribution used in the remainder of this analysis; it was taken from [17].

AS Path Length	1	2	3	4	5	6+
Percentage of Communication Instances [%]	3.9	18.1	36.4	31.8	8.3	1.5

Table 2. Communication distance distribution in the Internet, adopted from [17].

IV. MODEL

In order to study whether the TurfNet naming and inter-domain routing architecture is scalable to topologies of similar size to the Internet, this paper models TurfNet behavior on a Turf hierarchy of similar size and structure. The

previous section has outlined existing work that investigates the structure of the Internet. This section uses this information and presents a model of TurfNet, focusing on metrics, assumptions and the overall model. Section V then analyzes TurfNets scaling properties in this model.

A. Metrics

The TurfNet architecture has two main possible scalability bottlenecks: registration table size and aggregate lookup frequency. At higher levels in the hierarchy of TurfNets, the registration tables maintained by Turfs at that level are increasing, because they contain information about all the TurfNodes in the sub-hierarchy below them. The size of these registration tables could thus render the TurfNet architecture infeasible for increasing sizes of the networks.

The frequency of address lookups for remote TurfNodes is a second potential scalability bottleneck. Turfs propagate lookup requests that they cannot resolve locally up the TurfNet hierarchy. Forwarding of lookup requests adds latency and increases the lookup load on higher levels of the hierarchy. The analysis in Section V thus investigates both registration table sizes and lookup request frequencies for a global TurfNet.

B. Assumptions

TurfNet uses a registration/lookup scheme to initiate communication. The analysis in Section V evaluates lookup scalability and ignores registration requests. This simplification is legitimate, because the impact of registrations on the system is similar to lookups. If lookups are much more frequent than registrations, this overhead can be ignored in this preliminary analysis.

One important aspect with respect to the communication pattern is the rate at which hosts establish connections to hosts in other TurfNets that they are not already communicating with, or have been communicating in the recent past. Because such new connections require a lookup operation, this rate is particularly important for a scalability analysis. The analysis does not include requests for hosts inside the same TurfNet, because they do not involve inter-Turf operations that can affect global system performance. Moreover, once a lookup has succeeded, all TurfNets along the path have installed the necessary protocol and locator translation state. Successive communication between these hosts does not require additional lookup operations. The analysis in Section V uses a value of 0.01 unique lookup requests per second and host for Turf-external peers. This means that every node of a Turf establishes a new connection to a remote peer every 100 seconds.

A third aspect of the analysis is the total number of nodes in the network. The analysis in Section V assumes one billion hosts to exist at level 5 of the hierarchy, which significantly exceeds current estimates of the size of the Internet [23]. These hosts are distributed evenly across the level-5 Turfs. A future revision of the analysis will consider more realistic distributions of hosts.

C. Modeling

Based on these assumptions, the analysis of the model in Section V attempts to quantify the scalability properties of the TurfNet architecture. The analysis computes the required registration table sizes and aggregate lookup frequencies for

every Turf in the hierarchy. Because all hosts are located at level 5 of the hierarchy, only nodes at that level generate and consume traffic.

The aggregate lookup frequency at a Turf depends on the frequency of communication initiations by any of its local nodes or the nodes in any of its lower-level “customer” Turfs. The analysis assumes that every node at level 5 tries to randomly communicate with other nodes according to the connection establishment rate and the communication pattern described above.

A Turf first attempts to resolve any lookup request locally. If it cannot, it forwards the request vertically to its provider Turfs. These Turfs then attempt to resolve the lookup request in the same manner. This recursive lookup procedure is guaranteed to terminate at level 5 due to the hierarchical registrations used in TurfNet.

The fraction of lookup requests that a Turf can locally resolve depends on its level in the overall hierarchy. A Turf can resolve lookup requests for all local nodes and any nodes beneath it, *i.e.*, nodes that are known to any of its customer Turfs. Consequently, higher-level Turfs are more likely in general to be able to resolve lookup requests, because their registration tables are larger.

The scope of horizontal registration and lookup requests enables the TurfNet architecture to increase the fraction of lookup requests that a given level of its hierarchy can resolve. Similarly, the registration table size of a Turf depends on the number of its local nodes and on the number of nodes in all of its customer Turfs. For example, with a registration scope of two, a Turf learns about all nodes known to any of its immediate neighbors as well as their immediate neighbors. Obviously, larger scopes make known larger fractions of the overall population, but also increase registration tables. Section V attempts to analyze the impact of different scopes on the TurfNet performance by investigating scenarios that use different scopes.

V. ANALYSIS

This section mathematically analyzes the scalability properties of the TurfNet architecture. The basis of this analysis is the hierarchical TurfNet topology modeled after the Internet’s AS-level graph, as described in Section IV. The main metrics of this preliminary analysis are the amount of registration state required and the lookup frequency at different levels of the topology.

The bar graphs below show registration table sizes and lookup frequencies for three scenarios that use different combinations of scopes. The first scenario is the baseline case, where TurfNet operates with scopes of zero at all levels, *i.e.*, without using horizontal peerings for lookup requests. The second scenario uses a scope of 1 at levels 1, 2 and 3. Consequently, Turfs at these levels forward lookup requests to their immediate horizontal neighbors. The third scenario uses a scope of 2 at levels 1, 2 and 3. Here, Turfs forward node registration requests to horizontally peering Turfs that are at most 2 hops away. Due to space restrictions, this paper does not present other combinations of scopes, although the analysis included them as well. Also note that although these scenarios use the same fixed scope at all

levels, Turfs are free to modify the scope of a forwarded request as they see fit.

Note that the figures in this section include a level-0 Turf. It is required by the TurfNet architecture to terminate all lookup forwarding and does not exist in the corresponding Internet hierarchy. Even for TurfNet, the level-0 Turf will not often be required, because peering at level-1 can remove the need for a single root Turf, as illustrated by the results below.

This analysis first investigates the lookup request frequency at different levels. Fig. 2 shows the mean arriving lookups per second and Turf at different levels of the TurfNet hierarchy. Fig. 3 shows the aggregate lookup frequency across all Turfs at a particular level. Fig. 4 and Fig. 5 illustrate how many of the arriving lookup requests the Turfs at a given level can resolve, again per second and Turf and as an aggregate.

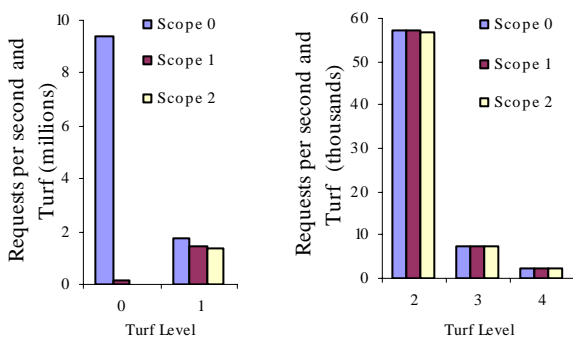


Fig. 2. Mean lookup requests arriving different levels of the Turf hierarchy for different scopes.

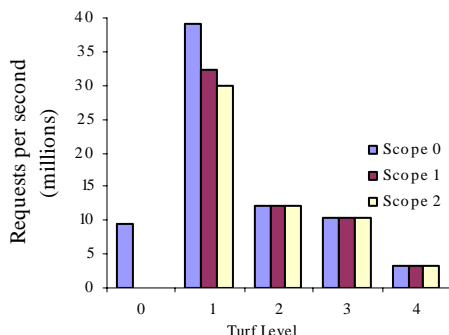


Fig. 3. Mean aggregate lookup requests arriving at different levels of the Turf hierarchy for different scopes.

A first conclusion is that with a scope of 0, the lookup volume at the virtual level-0 Turf is extremely high with approximately 9 million lookups/second. Introducing any peering scope reduces this load significantly. The main reason for this drastic reduction is that with a scope of even 1, any level-1 Turf can obtain the registration information of the vast majority of nodes in the system due to the small diameter of the level-1 graph (see Table 1).

At lower levels, increasing the scope has not such a dramatic effect. At level 2, increasing the scope already has only a minor impact. At levels below 2, the use of peering through increasing scopes has virtually no effect. The characteristics of the topology explain this difference. The average node at level 1 has over 20 peering links, whereas nodes at lower levels have much fewer (5-6 at level 2 and around 1 at level

3; none at the levels below 3). This means that the opportunities for benefiting from an increased scope are much smaller.

A second observation concerns the aggregate lookup frequencies. Fig. 3 illustrates that the arriving lookup load increases at higher levels. This result is surprising, because one expects that due to the locality inherent in the communication patterns, lower levels will resolve the majority of requests. However, TurfNet forwards any unresolved requests to *all* higher-layer Turfs. This amplifies the lookup load at higher layers, even when individual Turfs filter out duplicate requests – the same request can arrive at different Turfs at the same layer.

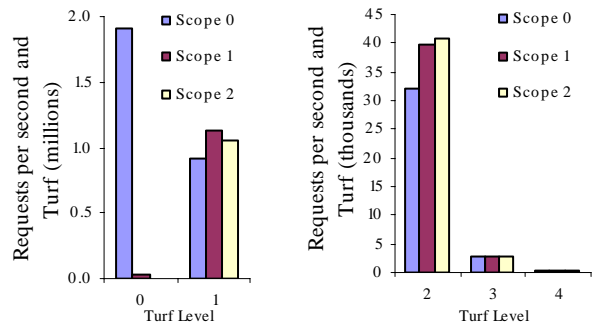


Fig. 4. Mean lookup requests resolved at different levels of the Turf hierarchy for different scopes.

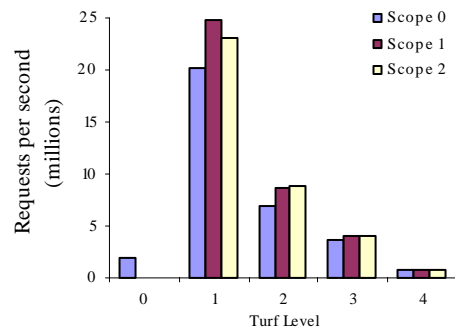


Fig. 5. Mean aggregate lookup requests resolved at different levels of the Turf hierarchy for different scopes.

An extension of the current TurfNet architecture attempts to address this inefficiency through several strategies that may reduce the number of lookup requests at layer 1; early result indicate that they can lead to reductions of almost 60%. The basic idea behind these extensions is to forward lookup request only to a small number of selected higher-level Turfs and to improve the resulting the path before communication.

In the third scenario, which uses a scope of 2 at levels 1, 2 and 3, no lookup request reach the virtual level-0 Turf. The level-1 Turfs resolve all incoming requests, because the diameter at level 1 is equivalent to the scope, *i.e.*, every Turf maintains the registration information of the whole population. This illustrates one advantage of peering: it can reduce lookup load by trading it against an increase in registration state. Another benefit of peering is that it enables improvements to the end-to-end path, *i.e.*, by using shorter paths that link Turfs more directly across horizontal peerings instead of passing vertically through the hierarchy.

A second metric that affects the scalability behavior of TurfNet is the amount of registration state that Turfs need to

maintain. Fig. 6 illustrates the mean registration table size for a Turf at a given level. In the baseline scenario with a scope of 0, the level-0 Turf must hold the registrations of all nodes in the system. Even with a scope of 1, this is still the case. Furthermore, the level-1 Turfs now replicate much of the overall registration state, significantly increasing their table sizes. With a scope of 2, the level-0 Turf becomes superfluous, because the level 1 Turfs resolve all lookup requests among themselves, without the need to forward up the hierarchy.

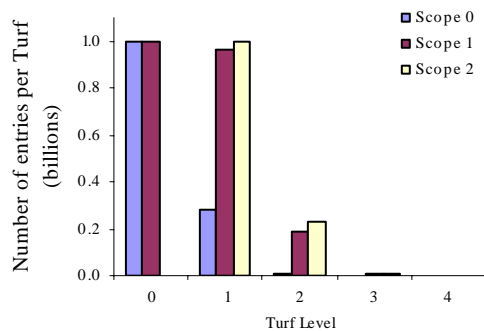


Fig. 6. Mean registration table size per Turf at different levels of the TurfNet hierarchy for different scopes.

The results in Fig. 6 indicate that the top-level Turfs must be able to handle large numbers of node registrations. Solutions that can maintain data sets of this order exist and have shown to perform well [24][25]. Additionally, in the investigated scenarios, all nodes always register for global reachability. A future refinement of the model will incorporate idle times, during which node registrations expire. This is expected to significantly reduce table sizes. Furthermore, not all nodes need to register for global reachability as TurfNet enables selective registration in Turfs of interest.

VI. CONCLUSION

TurfNet is a novel internetworking architecture that enables communication among highly autonomous and heterogeneous network domains. The architecture uses on a global identity namespace and does not require global addressing or a shared internetworking protocol. It integrates the new concept of dynamic network composition with other recent architectural concepts, such as decoupling locators from identifiers.

This paper examined whether the TurfNet architecture can scale to networks of the size of the global Internet. Based on existing work that investigates the Internet's autonomous system topology and develops heuristics to deduce the relationships between autonomous systems, this paper created a billion-host TurfNet topology that models the structure of the Internet. This topology is the basis for a preliminary scalability analysis that evaluates the performance of TurfNet's node registration and lookup mechanisms, which form the basis for inter-domain routing.

An ongoing effort investigates additional scalability aspects and extends the model presented in this paper. Furthermore, an experimental evaluation of a prototype implementation of TurfNet will attempt to experimentally validate the analytical scalability results presented in this paper.

REFERENCES

- [1] D. Clark, J. Wroclawski, K. R. Sollins and R. Braden, "Tussle in Cyberspace: defining tomorrow's Internet," Proc. *ACM SIGCOMM*, Pittsburgh, PA, USA, August 19-23, 2002, pp. 347-356
- [2] D. Clark, R. Braden, A. Falk and V. Pingali, "FARA: reorganizing the addressing architecture," Proc. *ACM SIGCOMM Workshop on Future Directions in Network Architecture (FDNA)*, Karlsruhe, Germany, August 2003, pp. 313-321
- [3] J. Crowcroft, S. Hand, R. Mortier, T. Roscoe and A. Warfield, "Plutarch: an argument for network pluralism," Proc. *ACM SIGCOMM Workshop on Future Directions in Network Architecture (FDNA)*, Karlsruhe, Germany, August 2003, pp. 258-266
- [4] D. R. Cheriton and M. Gritter, "TRIAD: a scalable deployable NAT-based Internet architecture," Stanford Computer Science Technical Report, January 2000
- [5] P. Francis and R. Gummadi, "IPNL: a NAT-extended Internet architecture," Proc. *ACM SIGCOMM*, San Diego, CA, USA, August 2001, pp. 69-80
- [6] R. Braden, D. Clark, S. Shenker and J. Wroclawski, "Developing a next-generation Internet architecture," Whitepaper, available at <http://www.isi.edu/newarch/DOCUMENTS/WhitePaper.ps>, July 2000
- [7] R. Moskowitz and P. Nikander, "Host Identity Protocol architecture," Work in Progress (draft-ietf-hip-arch-02.txt), January 2005
- [8] J. Abley, B. Black and V. Gill, "Goals for IPv6 site-multihoming architectures," RFC 3582, August 2003
- [9] Andreas Jonsson, Mats Folke and Bengt Ahlgren, "The split naming/forwarding network architecture," Proc. *First Swedish National Computer Networking Workshop (SNCNW)*, Arlandastad, Sweden, September 8-10, 2003
- [10] Z. Turanyi, A. Valko and A. Campbell, "4+4: an architecture for evolving the Internet address space back towards transparency," *ACM SIGCOMM Computer Communication Review*, Vol. 33, No. 5, October 2003, pp. 43-54
- [11] S. Schmid, L. Eggert, M. Brunner and J. Quittek, "TurfNet: an architecture for dynamically composable networks," Proc. *First IFIP TC6 WG6.6 International Workshop on Autonomic Communication (WAC 2004)*, Berlin, Germany, October 18-19, 2004
- [12] M. Walfish, J. Stribling, M. Krohn, H. Balakrishnan, R. Morris and S. Shenker, "Middleboxes no longer considered harmful," Proc. *USENIX Symposium on Operating Systems Design & Implementation (OSDI)*, San Francisco, CA, USA, December 2004, pp. 215-230
- [13] S. Schmid, L. Eggert, M. Brunner and J. Quittek, "Towards Autonomous Network Domains," Proc. *8th IEEE Global Internet Symposium*, Miami, FL, USA, March 17-18, 2005
- [14] H. Chang, R. Govindan, S. Jamin, S. Shenker, and W. Willinger, "On inferring AS-level connectivity from BGP routing tables," Proc. *IEEE INFOCOM 2002*, pp. 618-627
- [15] Z. Ge, D.R. Figueiredo, S. Jaiswal, and L. Gao, "On the hierarchical structure of the logical Internet graph," Proc. *SPIE ITCOM*, Denver, Colorado, USA, August 2001
- [16] Z. Mao, D. Johnson, J. Rexford, J. Wang, and R. Katz, "Scalable and accurate identification of AS-level forwarding paths," *IEEE INFOCOM*, Vol. 23, No. 1, Hong Kong, China, March 2004, pp. 1606-1616
- [17] "CAIDA web site," <http://www.caida.org>
- [18] A. Broido, E. Nemeth, and KC Claffy, "Internet expansion, refinement and churn," *European Transactions on Telecommunications*, Vol. 13, No. 1, January-February 2002, pp. 33-51
- [19] D. Krioukov, K. Fall, and X. Yang, "Compact routing on Internet-like graphs," Proc. *IEEE INFOCOM*, March 2004
- [20] X. Yang, "NIRA: a new Internet routing architecture," *ACM SIGCOM Workshop on Future Directions in Network Architecture (FDNA)*, Karlsruhe, Germany, August 2003, pp. 301-312
- [21] L. Gao, "On inferring Autonomous System Relationships in the Internet," *IEEE/ACM Transactions on Networking*, Vol. 9, No. 6, December 2001, pp. 733-745
- [22] L. Subramanian, S. Agarwal, J. Rexford, and R.H. Katz, "Characterizing the Internet hierarchy from multiple vantage points," Proc. *IEEE INFOCOM*, NY, USA, June 2002, pp. 618-627
- [23] "CIA World Factbook," <http://www.cia.gov/cia/publications/factbook>
- [24] M. F. Kaashoek and D. R. Karger, "Koorde: a simple degree-optimal distributed hash table". Proc. *2nd International Workshop on Peer-to-Peer Systems*, Berkeley, CA, USA, February 2003, pp. 98-107
- [25] I. Stoica, R. Morris, D. Karger, M. Frans Kaashoek, and H. Balakrishnan, "Chord: a scalable peer-to-peer lookup service for Internet applications," *IEEE/ACM Transactions on Networking*, Vol. 11, No. 1, February 2003, pp. 17-32