On the Inter-domain Scalability of Route-by-Name Information-Centric Network Architectures

Konstantinos V. Katsaros*, Xenofon Vasilakos†, Timothy Okwii*, George Xylomenos†,
George Pavlou* and George C. Polyzos†
*University College London, UK
E-mail: {k.katsaros, timothy.okwii.11,g.pavlou}@ucl.ac.uk
†Athens University of Economics and Business, GR
E-mail: {xvas, xgeorge, polyzos}@aueb.gr

***Abstract*—Name resolution is at the heart of Information-Centric Networking (ICN), where names are used to both identify information and/or services, and to guide routing and forwarding inside the network. The ICN focus on information, rather than hosts, raises significant concerns regarding the scalability of the required Name Resolution System (NRS), especially when considering global scale, inter-domain deployments. In the route-by-name approach to NRS construction, name resolution and the corresponding state follow the routing infrastructure of the underlying inter-domain network. The scalability of the resulting NRS is therefore strongly related to the topological and routing characteristics of the network. However, past work has largely neglected this aspect. In this paper, we present a detailed investigation and comparison of the scalability properties of two route-by-name inter-domain NRS designs, namely, DONA and CURLING. Based on both real, full-scale inter-domain topology traces and synthetic, scaled-down topologies, our work *quantifies* a series of important scalability-related performance aspects, including the distribution of name-resolution state across the Internet topology and the associated processing and signaling overheads. We show that by avoiding DONA's exchange of state across peering links, CURLING results in deployment costs proportional to the total number of downstream customers of each Autonomous System. This translates to a 62-fold global state size reduction, at the expense of a 2.78-fold increase in lookup processing load, making CURLING a feasible approach to ICN name resolution.**

I. Introduction

Information-Centric Networking (ICN) is a new networking paradigm that reflects the currently prevailing usage of the Internet i.e., the retrieval and dissemination of information/content[1]. Focusing on information exchange rather than on pair-wise communication between hosts, ICN promises a series of advantages compared to the current Internet architecture, including in-network caching, multicast, anycast, mobility support, privacy and security [1]. In this context, a considerable amount of research effort has been devoted to the design of alternative ICN architectures [2], [3], [4], [5], [6].

Name resolution lies at the heart of all proposed ICN architectures. Information is organized in the form of named *Information Objects* (IOs)[2], where names both identify information and guide routing and forwarding decisions within the network. Unfortunately, the size of the IO name space is expected to be enormous. Considering that the number of unique web pages indexed by Google is greater than 1 trillion [7] and that billions of devices (phones, sensors, home appliances) will be offering additional content in the near future [8], any name resolution design will have to handle unique IOs in the order of 10^{13}; some studies raise this estimate even to 10^{16} [9]. As a result, significant scalability concerns are raised for any ICN *Name Resolution System* (NRS), related to the amount of network and computational resources required to register, update and lookup IO names.

Several research efforts have attempted to address this scalability issue by employing *Distributed Hash Tables* (DHTs), due to their logarithmic scalability properties (e.g., [9], [10], [11]). However, with DHTs name-resolution paths are far from optimal [12], [13], even in the presence of extensive caching of name resolution information[3] [11], and they often violate the underlying inter-domain routing policies. In addition, the stochastic nature of DHTs arbitrarily spreads state across the inter-network, forcing requests for local content to rely on NRS servers at faraway network locations.

In contrast to DHT designs, the *Data-Oriented Network Architecture* (DONA) [6] and the *Content-Ubiquitous Resolution and deLivery Infrastructure for Next Generation services* (CURLING) [4] were designed to conform to the underlying routing policies. Both the registration of name resolution state and the name resolution process follow customer-provider, and (possibly) peer-to-peer links. As a result, name resolution paths are compliant to Border Gateway Protocol (BGP) routing and state is only placed/replicated as mandated by the inter-domain routing relationships of the interconnected *Autonomous Systems* (ASes). However, this replication process obviously raises concerns regarding the scalability of the resulting NRS, especially when considering the expected size of the name space and the size of the inter-domain topology, which currently exceeds 40,000 ASes [14].

The identified scalability concerns have received so far limited attention. Koponen et *al.* [6] provided some insights

[1]We will use both terms interchangeably in the remainder of this paper.
[2]For brevity, we use the term IO to refer to both information and services.
[3]For brevity, we use the term *state* to refer to name resolution information.

on expected state size and processing overheads, but only for a limited part of the Internet (i.e., Tier-1 domains). The actual overheads for the remainder of the ASes are inherently coupled with the structure of the inter-domain topology and the routing relationships. In previous work [11], [15] we investigated this interplay for DONA, finding a heavily skewed distribution of state across the inter-domain topology, which is significantly affected by peering links.

In this paper, we take a step further by presenting a detailed scalability investigation and comparative study of DONA and CURLING. We assess the impact of the inter-domain topology on state and processing overheads, revealing and, most importantly, *quantifying* the associated design trade-offs in terms of the resulting resource requirements (and corresponding deployment costs). To this end, we employ an analysis of the AS-level Internet graph inferred by the CAIDA BGP traces [14], backed by detailed packet-level simulations on scaled-down Internet-like graphs. Our performance evaluation reveals that the majority of ASes, 81% (DONA) and 90% (CURLING) to be exact, present moderate to low resource requirements. By not exchanging state across peering links, CURLING further reduces the state size 62-fold compared to DONA, at the cost of an average 2.78-fold increase of lookup processing load. Moreover, while DONA shortens name resolution paths by 16% compared to CURLING, it requires 684% more traffic for name registrations. To the best of our knowledge, this is the first comprehensive investigation of the actual scalability of non-DHT, BGP-compliant, ICN NRSs.

The remainder of this paper is structured as follows. In Section II we provide background information on name resolution for ICN, as well as a review of past work in the evaluation of route-by-name inter-domain NRSs to motivate this study. We describe our evaluation framework in Section III and then present and discuss our findings in Section IV. Finally, we summarize our results and conclusions in Section V.

II. Background and Motivation

ICN focuses on the information itself, rather than on the endpoints of communication. The network acts as a mediator that decouples content providers (*a.k.a.* publishers) from content consumers (*a.k.a.* subscribers) by (a) interacting with publishers, which supply information about the availability and the location of IOs, and (b) accepting end-user requests referring to IO names with the purpose of locating the corresponding content or service and enabling its delivery. The NRS matches content availability with user requests, therefore it must maintain state that binds IO names to the information required to locate the corresponding IOs.

Maintaining such state at a global Internet scale becomes a significant challenge when considering the size of the IO name space. This has triggered a series of research efforts trying to design highly scalable overlay systems [6], [4], [9], [10], [11]. Overlays target the more easily available computing resources of the application layer, rather than the more restricted network layer environment, where approaches such as *Content Centric Networking* (CCN) [3] are based (see, for instance, [16]).

Resource requirements are also largely affected by the structure of the name space. As in the case of domain names and the IP address space, hierarchical name spaces can easily support name aggregation[4] which results in lower state size overheads. On the other hand, flat, self-certifying names, better support persistence and authentication. Furthermore, the perceived benefits of aggregation may not materialize in an ICN name space, where there may exist a huge number of prefixes with unknown popularities, unlike in DNS and IP. Though the debate on hierarchical *vs.* flat name spaces is still on-going, the latter seem to have attracted the attention of researchers with most proposed ICN architectures being based on them [2], [4], [6], [11].

Within this body of work, two main resolution models have emerged. In the *lookup-by-name* model, the NRS takes the form of a large scale distributed database maintaining mappings between IO names and the most suitable source location (e.g., the closest) of the corresponding content. In the *route-by-name* model, name resolution requests are routed hop-by-hop until they reach the location of the corresponding content i.e., the resolution request reaches the content itself, rather than the node aware of the content location.

A. Route-by-name approaches

In the classic route-by-name approach, DONA [6], each IO is associated with a *principal* that can be considered as its owner. An IO identifier consists of the cryptographic hash of the principal's public key and a label unique to the principal, forming a self-certifying, globally unique name. The DONA design involves an overlay of *Resolution Handlers* (RHs) with at least one logical RH placed at each AS. The role of RHs is to register and maintain name resolution state, as well as to propagate name resolution requests until they are resolved. To this end, RHs interconnect following the hierarchical interconnection of their ASes, forming a corresponding RH hierarchy. The RH hierarchy is further enhanced with peering links i.e., RHs of peering domains are also linked in the RH overlay[5].

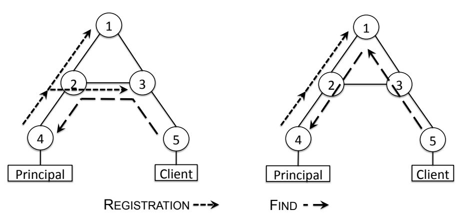
Principals issue REGISTER messages towards their local RH(s) to advertise their IO's to the network. A local RH propagates a REGISTER message upwards in the inter-AS hierarchy and to RHs at peering ASes, thus setting up the name resolution state throughout the network, as shown in Fig. 1(a). The propagation of REGISTER messages stops at the top-most AS level i.e., at Tier-1 RHs (see Section III-B). ASes not willing to transit name resolution requests and/or data do not propagate REGISTER messages received over peering links. Since Tier-1 ASes all peer with each other, each tier-1 AS will be aware of all IOs in the network. The resolution state at each RH is in the form of `<IO name, next hop RH>` pairs, i.e., mappings between advertised IO names and a pointer to the previous RH in the corresponding REGISTER propagation path.

[4]Using common name prefixes to represent different IO names e.g., using the name /a/b/ to represent both /a/b/c and /a/b/d.

[5]Since peering links introduce cycles in the inter-domain network graph, the topology is not strictly hierarchical.

Clients (i.e., end hosts) issue name resolution requests in the form of FIND messages submitted to their local RH(s). FIND messages are also propagated upwards in the domain hierarchy according to the inter-domain routing policies, but not over peering links, as shown in Fig. 1(a). In the worst case, a FIND message may have to reach a Tier-1 AS in order to locate an IO, or determine that no such IO exists. Upon a name match with an RH entry, FIND messages follow the reverse registration path to reach the local RH of the appropriate principal, which triggers the data transfer.

CURLING follows a similar approach, also adapting to the inter-domain topology structure for name registration and resolution. However, in CURLING, *Content Resolution Servers* (CRS)[6] propagate registration and resolution requests only to their provider ASes, as shown in Fig. 1(b). In effect, both REGISTER and FIND messages follow a subset of the underlying routing relationships, as they do not cross peering links. As a result, when a FIND request reaches a Tier-1 AS, it will have to be broadcasted to all other Tier-1 ASes to guarantee resolution. CURLING however allows optimizing the data paths to allow the utilization of peering links in the delivery of the content itself. Finally, it is worth noting that CURLING also offers *scoping* and *filtering* features, where a name resolution request may explicitly define the ASes allowed (or not) to act as sources of the desired content. These features are beyond the scope of this work.



(a) DONA (b) CURLING

Fig. 1. Example of registration and resolution processes.

B. Lookup-by-name approaches

The *Domain Name System* (DNS) is the closest existing equivalent to a lookup-by-name NRS. Nevertheless, it is considered inadequate for the purposes of a global ICN NRS, even for ICN architectures that use hierarchical, rather than flat, names. One reason is its susceptibility to security attacks [17], due to the limited redundancy in name-servers and the fact that many servers have a single point of attachment to the Internet [12]. In addition, the load is not equally balanced between root servers, since names are not equally distributed among top level domains. These concerns gain additional importance in the context of ICN, where the load is expected to be orders of magnitude higher, since in ICN individual IOs, rather than the servers hosting them, need to be resolved. There may also exist a huge number of name prefixes, that will need to be resolved anywhere on the network, rather than on a static set of root name servers. An extension of the DNS for ICN was recently proposed [18], but that work did not delve into the scalability properties of the resulting design.

DHTs have drawn the attention of researchers in the past [12], [13], in an effort to address the limitations of DNS. With the advent of the ICN paradigm, DHTs appeared again as a promising way to address the increased needs for scalability. Dannewitz et *al.* [9] proposed MDHT, a multilevel DHT based NRS that aggregates IO registration entries at higher levels of the inter-domain hierarchy. MDHT provides an indirection mechanism which allows the NRS to resolve content provider names, with the resolution of the content itself taking place at a lower level. A similar approach is followed by Rajahalme et *al.* [10], where an indirection level is used to map *scopes* of information to lower level resolution nodes. The presented performance evaluation study is based on significant abstractions e.g., intra-domain routing overhead and caching have been coarsely modeled based on observations made in different contexts, which do not reflect the proposed architecture's intrinsic characteristics. In both approaches, the routing inefficiency of DHTs is circumvented by the aggregation of information at higher levels of the inter-domain structure, raising however scalability concerns.

The use of a hierarchical version of the SkipNet [19] overlay has also been proposed by Dannewitz et *al.* in [9] to achieve low resolution delays. However, the presented performance evaluation is based on an oversimplified abstraction of the inter-domain topology i.e., a full k-ary tree, overlooking the effect of multihoming and peering links. DHT-NRS [11] is based on H-Pastry [20], a multi-level version of the Pastry DHT [21] that tries to adapt to the underlying network hierarchy by taking physical network proximity, administrative domain boundaries and inter-domain routing policies into account. Although DHT-NRS exhibits better performance than other DHT-based NRSs, it reduces but does not eliminate the intrinsic problems of DHTs [11].

C. Motivation

A closer look at the two alternative NRS design approaches for ICN, reveals a fundamental tradeoff between routing efficiency and state distribution. In DHT-based lookup-by-name approaches, state is equally spread among nodes, independently of the structure of the inter-domain topology. However, this comes at the cost of inefficient name resolution: DHTs yield stretched overlay paths, leading to increased delays compared to shortest-path routing. The common workarounds for this problem are caching and increasing the average node degree in the DHT (e.g., [13]). As shown by Katsaros et *al.* [11], overlay routing in DHTs also violates inter-domain routing policies i.e., the routes followed when resolving a name may not respect the routing policies of the underlying physical network. Finally, the stochastic nature of DHTs offers very limited control over the placement of the resolution state: the node hosting the resolution state of an IO may reside in a network far from the IO itself, introducing security issues

[6] For simplicity, we will use the term RH for both DONA and CURLING, as RHs and CRSs offer similar functionality with respect to the aspects investigated.

and further increasing resolution delays. These issues become more important in the context of ICN, where resolution traffic is expected to be dramatically higher than today.

Route-by-name NRS designs like DONA and CURLING achieve shortest-path routing, providing full compliance with the underlying inter-domain routing policies. Unfortunately, they require extensive replication of state across the inter-domain topology, with multiple RHs along the registration paths maintaining name-resolution entries for the same IOs. While it is easy to deduce how state accumulates at Tier-1 ASes, especially with DONA where all Tier-1 NRS nodes are burdened with the entire name resolution state [6], the remainder of the topology is not so simple to analyze. This is because the state accumulation process heavily depends on the structure of the underlying inter-network: since REGISTER messages propagate upwards along BGP paths, the state accumulated at each AS depends on its position in the inter-domain hierarchy. In the case of DONA, this state overhead further depends on the state accumulated at peering domains. Consequently, assessing the scalability properties of DONA and CURLING requires a detailed topological investigation.

In our previous work we examined this tradeoff by comparing DONA with a DHT-based NRS using scaled-down Internet-like topologies [11]. Our investigation showed that in terms of routing efficiency DONA outperforms the DHT, even when the DHT is supported by extensive caching. At the same time, we found that the state distribution across the ASes of the inter-domain topology was extremely skewed. We later investigated this phenomenon in more detail using the full CAIDA Internet topology [15]. In this paper we expand our research on the scalability properties of route-by-name inter-domain NRSs in ICN, not only by using an extended and updated set of topologies, but also by covering both DONA and CURLING, so as to investigate the impact of state exchange across peering links. As we later show in this paper, the different treatment of peering links in CURLING has a dramatic effect on the resource requirements of ASes, especially when considering the gradual increase of peering links in the inter-domain topology.

III. Evaluation Framework

A. Performance metrics

Evaluating an NRS, as any other system, calls for the careful selection of the performance metrics. In the following, we provide a detailed description of the metrics used in our evaluation and a discussion of their importance for this study. In essence, these metrics directly reflect different aspects of the scalability properties of the considered NRSs.

1) State size: The most important challenge in an ICN NRS is the IO population to be resolved. For this reason, the size of the state maintained by the NRS is of paramount importance, as it determines the resources required in terms of memory and lookup processing load. Unless otherwise stated, we assume that each IO entry consists of a 40 byte IO identifier and a 2 byte pointer to the next RH i.e., a bitmap for the RH's interfaces to neighbour RHs [6].

We evaluate state size on a per RH and on a per AS tier (see Section III-B) basis. In the former case, we attempt to quantify the overheads imposed to individual ASes, so as to assess the deployment costs of each considered NRS. In the latter case, we aim to provide insights on the expected burden at different levels of the AS hierarchy. To achieve these goals, we first measure the number of IO entries per AS and normalize it as a percentage of the overall state size (i.e., unique IO entries) throughout the inter-network. Then, we quantify the level of replication in each considered name resolution scheme by defining the *multiplier* m as the ratio of the total number of registration entries maintained throughout the inter-domain topology (RE_{total}) to the total number of unique information objects (RE_{unique}), or $m = \frac{RE_{total}}{RE_{unique}}$.

2) Processing overhead: To estimate the processing load imposed by the name resolution process, we define the *lookup overhead* (LO) metric as the sum of the total number of resolution requests handled (forwarded or locally resolved) by each RH. Our target is to capture the impact of the design choices in both DONA and CURLING, *wrt.* the forwarding of registration and resolution requests. In addition, it is important to note that the amount of state at each node further provides an indication of the processing requirements per lookup.

3) Resolution delay: As in this work we focus on the impact of routing/forwarding on the name resolution delay, we express delay as the hop count of name resolution requests i.e., the number of inter-domain hops taken by a name resolution request until the requested IO is found at an NRS node. We do not take into account lookup processing, queuing and transmission delays, so as to isolate the impact of state placement across the hierarchy in each considered scheme.

4) Registration overhead: Considering the vast size of the information space, the registration of IOs may result in a significant overhead on the control plane. We investigate this aspect by measuring the aggregate number of single hop transmissions needed for all REGISTER messages to reach their target(s) and register all IOs in the inter-network.

B. Network topology

As previously discussed, the name resolution performance of DONA and CURLING heavily depends on the structure of the underlying inter-domain topology. Studies indicate that the entire inter-AS graph consists of approximately 45,000 ASes and approximately 200,000 annotated links [14], making such a performance evaluation a technical challenge on its own. We have therefore considered two complementary approaches. We first used the AS-level Internet graph inferred by the CAIDA BGP traces[7] [14]. This graph was used by a Java-based custom-written DONA/CURLING simulator which simulates the registration of IOs across the domains by traversing the appropriate parts of the graph. As mentioned before, we neglect processing, queueing and transmission delays, to enhance the scalability of the simulator. For the same reason,

[7]The traces are created by first collecting traceroute-like IP-level topology data from several vantage points in the Internet and subsequently identifying the involved ASes from the traced IP addresses.

our measurements consider only one RH per AS. Due to these limitations, we employ this graph only to study the state distribution across the inter-domain topology, as metrics requiring the simulation of name resolution requests require enormous amounts of computational resources at this scale.

The Internet inter-domain AS graph is actually far from hierarchical, due to the prevalence of multihomed ASes and peering links. In this paper we classify the ASes appearing in the CAIDA trace set into four tiers based on the size of their customer *cone*, i.e. the total number of their downstream customers [22]. The four tiers are:

1) *Stub* networks, i.e. networks with no more than 4 customer networks, which includes *all* access networks,
2) *Small ISPs*, i.e. small Internet Service Provider ASes that have a cone size between 5 and 50,
3) *Tier-1* ASes, i.e. ASes at the highest level of the hierarchy that do not act as customers for another AS,
4) *Large ISPs*, a category which includes the remaining ASes that have a larger cone than small ISPs but are *not* Tier-1 members.

We then employed scaled-down inter-domain topologies generated by the algorithm of Dimitropoulos et *al.* [23], which present a manageable size for our evaluation purposes while maintaining the same properties of the original CAIDA graph. Specifically, we employ a domain topology of 400 ASes interconnected with multi-homing and peering links[8]. On top of this topology, we deploy a population of 4000 RHs uniformly across the domains i.e., 10 nodes per domain. We neglect any intra-domain communication overheads and focus our study on the effects of the inter-domain topology structure on the performance of the NRS. In these scaled-down topologies we classify the ASes according to their minimum hop-count distance to the top level of the hierarchy.

In both setups we consider REGISTER messages generated only by stub networks. This reflects (i) the currently dominant service/business model where content providers may connect at any level of the inter-domain hierarchy to push their content, without providing transit services, and (ii) the cases of user generated content, with users connected under stub networks of the hierarchy. In multihomed ASes, RHs forward registrations to the RH of their randomly selected default provider AS.

C. Workload

Considering the simulator scalability limitations discussed above, we again consider two different workload models, one per topology model. For the large scale CAIDA trace, we reduce the computational requirements by having each stub AS generate a single distinct REGISTER message. This convention obviously corresponds to the simplified, uniform distribution of content across the Internet, and as such it can only provide a rough insight on the distribution of state across the hierarchy. A more realistic evaluation would necessitate an uneven distribution of content across the topology, however, to the best of our knowledge, no such model currently exists. Moreover, to further reduce the processing and memory requirements, we do not consider resolution requests, as the added overhead is at least of the same order of magnitude as for registrations.

We address these limitations with the workload for the smaller topology, which considers a detailed mixture of various traffic types (e.g., Web, Video, P2P). For this purpose we employed the GlobeTraff [24] traffic generator tool[9]. To estimate the number of resolution requests, we derived the actual number of IOs for each traffic type by dividing the corresponding data volume with the median[10] IO size for that traffic type as measured in relevant studies (see [24] and references therein). We thus ended up with the traffic mix of Table I, where for each traffic type we show which percentage of the data and control plane traffic it represents, what the median object size is, how object sizes are distributed and how object popularities are distributed. We notice that for the Video and P2P traffic types, the control plane traffic is considerably low, as the large object sizes result in fewer resolution requests for a certain amount of data traffic, compared to the (even more than one order of magnitude) smaller object sizes in the Web and Other traffic types. We then generated a workload corresponding to 25 GB of traffic, resulting in an average[11] of 2,430,379 subscription messages for 1,032,030 IOs. This size limit was imposed by the resource limitations of the simulation environment. We considered again REGISTER message generation to be uniformly distributed across stub domains. Since end hosts typically reside in access networks i.e., domains that have no customer domains, each resolution request is injected from a randomly chosen access network.

IV. PERFORMANCE RESULTS

A. Full-scale topology

Figures 2(a) and 2(b) show how state is distributed to ASes across all topology tiers and per tier, for DONA and CURLING, respectively, based on the 2013 CAIDA topology model. The y-axis shows the cumulative fraction of ASes that together maintain the percentage of the total state size indicated on the x-axis. We also provide the average and median percentage of total state held by each AS across all tiers and per tier in Table II. The replication factors are $m_{DONA} = 1702.64$ and $m_{CURLING} = 27.34$, indicating that CURLING provides very large improvements against DONA in the state required per AS, although, as the last column of Table II shows, the gains are very dependent on the AS category: while the average x-fold improvement of CURLING over DONA is 62.97, it ranges from 1.67 for Tier-1 ASes to 679.67 for stub ASes. More than 81% of ASes in DONA

[8]Similar results where obtained for other topology sizes, which are not presented due to length limitations.

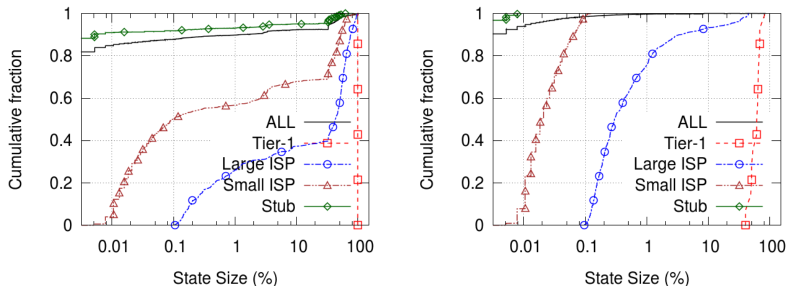
[9]We did not use DNS traces as they reflect the current Internet architecture. For example, they omit requests sent directly to servers, like HTTP requests.

[10]The choice of the median rather than the mean object size was made to avoid skewing the results due to the long-tail characteristics of some distributions (e.g., the Pareto tail of Web object size distribution).

[11]We used different workload instances to increase randomness.

Traffic type	Data plane fraction	Control plane fraction	Object size median	Object size distribution	Object popularity distribution
Web	35.10%	36.40%	10.386 KB	Lognormal-Pareto	Zipf
P2P	15.85%	$2.56\text{x}10^{-4}$%	650.11 MB	Sampling	Mandelbrot-Zipf
Video	19.54%	$37.04\text{x}10^{-3}$%	7.6 MB	Concatenated Normal	Weibull
Other	29.51%	63.57%	5 KB	Normal	Zipf

TABLE I
TRAFFIC MIX CHARACTERISTICS, BASED ON GLOBETRAFF [24].



(a) DONA (b) CURLING

Fig. 2. Distribution of state size across the AS-level topology (CAIDA 2013 traces).

	DONA		CURLING		
Type	**Average**	**Median**	**Average**	**Median**	**Avg. gain**
All Tiers	3.778%	0.003%	0.060%	0.003%	62.97%
Tier-1	100.00%	100.00%	59.895%	61.769%	1.67%
Large ISP	36.701%	42.687%	2.758%	0.298%	13.31%
Small ISP	15.599%	0.097%	0.029%	0.018%	537.90%
Stub	2.039%	0.003%	0.003%	0.003%	679.67%

TABLE II
STATE SIZE PER AS EXPRESSED AS A PERCENTAGE OF THE TOTAL STATE SIZE THROUGHOUT THE INTER-NETWORK (%).

and 90% in CURLING are only burdened with their own registration entries, while the remainder of the ASes (8150 and 4335 ASes in total, for DONA and CURLING, respectively) are disproportionally loaded with state, reaching even the entire available state in the network in the case of Tier-1 ASes in DONA. This is a direct consequence of the structure of the inter-domain topology: the vast majority of ASes (92.87%) belong to the Stub tier, with only a small fraction (9.17%) having more than one (and up to 4) customers that contribute their registration entries.

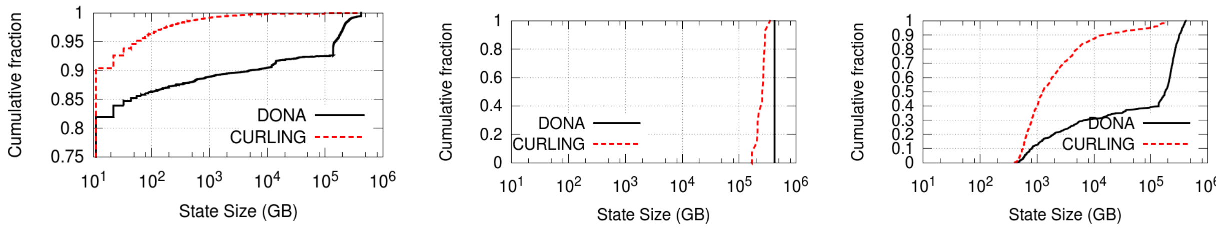
	DONA		CURLING	
Type	**Average**	**Median**	**Average**	**Median**
Tier-1	26,250	26,250	15,723	16,215
Large ISP	9,635	11,206	724	79
Small ISP	4,095	26	8	5
Stub	536	1	1	1

TABLE III
NUMBER OF 16 GB RAM SERVERS REQUIRED TO HOLD STATE IN RAM.

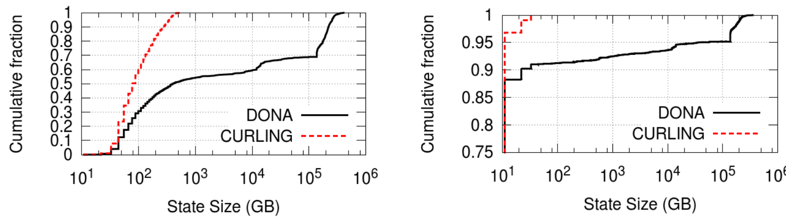
Figures 3(a) to 3(e) present a more detailed comparison between DONA and CURLING. These plots show the cumulative distribution of state to ASes, but expressed in GB and corresponding to a total of 10^{13} IOs in the inter-network. We can see that the overall state in the network is in the order of several hundred TB (420 TB for 10^{13} IOs), and this is also the total state size handled by each Tier-1 AS in DONA. As expected, CURLING accumulates considerably less state per AS compared to DONA: in the case of Tier-1 ASes, the requirements remain in the order of TBs, but they are 40% lower. Approximately 90% of the Large ISPs would be required to maintain no more than ten TB of state in CURLING, while this only holds for approximately 30% of the Large ISPs in the case of DONA. The difference between the two schemes becomes more evident at lower tiers, where DONA increases state size by up to 3 or 4 orders of magnitude. Table III translates the observed state sizes into the corresponding hardware resource requirements, expressed as the number of 16 GB RAM servers required to hold the resolution state in RAM at each AS, per tier[12]. Clearly, both schemes require the deployment of large data centers at Tier-1 to cope with the state size; fortunately, this is required at only a few ASes.

The reduced state size with CURLING, also reflected in the substantial difference between the m_{DONA} and $m_{CURLING}$ values reported above, is the direct result of not exchanging IO entries between peering domains. On closer inspection, each tier is defined by the *cone* size of the participating ASes (see Section III-B), which in turn determines the size of the

[12]For simplicity, we assume that all 16 GB are used to hold name resolution state and ignore indexing overheads.

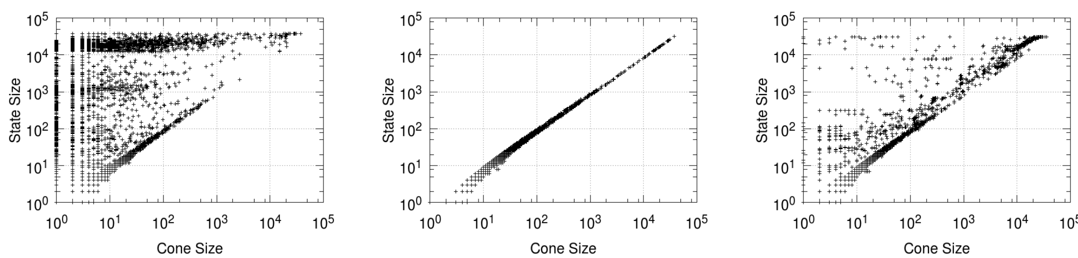


(a) All tiers (b) Tier-1 (c) Large



(d) Small (e) Stub

Fig. 3. Comparison of state distribution per tier (CAIDA 2013 traces).



(a) DONA, CAIDA 2013 traces (b) CURLING, CAIDA 2013 traces (c) DONA, CAIDA 2011 traces

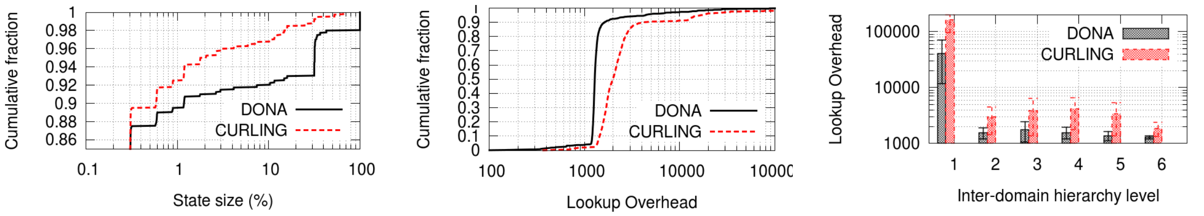
Fig. 4. Relation between AS cone size and accumulated registration state in log scale.

accumulated (distinct) registrations created by cone members. This state is augmented in DONA by the entries received from peering ASes, as clearly demonstrated in Figure 4, which shows the relation between AS cone size and state size. The large concentration of points in the diagonal shows that the amount of state accumulated by ASes is mainly affected by the size of their cone. The scattered points above the diagonal in Figure 4(a) show the impact of peering relationships in DONA. This is most evident at the upper right part of the graph which depicts the full mesh connectivity at Tier-1. It is also important however to note the existence of several cases where ASes with a relatively small cone size maintain state of disproportionate size in DONA (upper left area of Figure 4(a)), reflecting the establishment of peering relationships with ASes at higher tiers (and therefore large cone sizes), which results in the reception of large amounts of state. In contrast, Figure 4(b) shows no points above the diagonal for CURLING. For the vast majority of ASes, CURLING results in a state overhead largely determined by their cone size. This is particularly important, since smaller ASes, residing lower in the hierarchy, are not forced into large infrastructure deployments only to resolve large content volumes residing at other domains. The points below the diagonal correspond to ASes whose cone also contains non Stub ASes, which do not generate registrations.

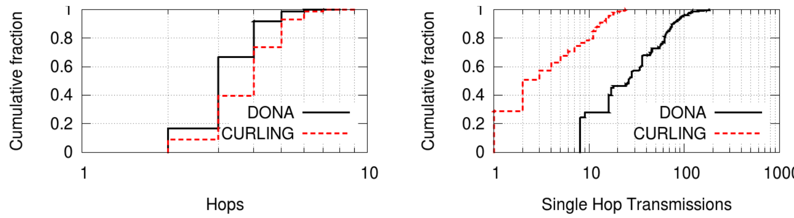
The impact of peering relationships on state overhead is of particular importance, due to the evolution of the inter-domain topology graph as a result of the increasing establishment of peering links, especially by large content providers such as Google [25]. Unfortunately, the identification of peering links remains a difficult task. As argued by Chen et *al.* [26] the followed CAIDA methodology fails to reveal up to 90% of existing peering links. For the purposes of our study, we repeated our measurements over the 2011 CAIDA traces [27], in which the number of peering links reported is only 3,523, rather than the 66,617 peering links in the more recent traces [14]. The impact of the additional peering links becomes obvious by comparing Figures 4(a) and 4(c), where we can see how the additional peering links have multiplied the points above the diagonal.

B. Scaled-down topology

As discussed in Section III-B, we have resorted to scaled-down topologies for the evaluation of more dynamic aspects of the considered NRSs, such as processing and signaling overheads, and resolution delays. Figure 5(a) shows the same trend for the cumulative distribution of state size per node in DONA and CURLING, as in the case of the CAIDA trace based measurements (see Figure 3(a)), confirming the validity



(a) Cumulative state overhead (b) Cumulative lookup overhead (c) Distribution of lookup overhead across hierarchy levels



(d) Cumulative distribution of hops per resolution request (e) Cumulative distribution of single hop transmissions per registration

Fig. 5. Simulation based performance results (scaled down topology).

of our scaled-down approach. There is a sharp increase in the set of ASes that present a relative state size larger than approximately 25-30% of the total state size in the case of DONA, which is also observed with the CAIDA trace, due to the exchange of state over peering links with higher tier ASes.

Figure 5(b) shows the cumulative distribution of the lookup overhead per node as defined in Section III-A. On average, CURLING incurs a 2.78-fold increase on lookup processing load compared to DONA, as resolution requests must be forwarded to the closest (according to BGP) common ancestor in the topology, even when the requested item resides in a peering AS (or one of the ASes in its cone). Figure 5(c) further shows how lookup overhead is distributed across the levels of the hierarchy. In both cases, the top-most domains receive a substantially higher load than their lower level counterparts. This is because the top-most domains end up always serving requests that could not be resolved lower in the hierarchy, which is the typical case when requests and IOs reside in different areas of the inter-network. Obviously, the exchange of state over peering links (in DONA) does not suffice to balance the load among the hierarchy levels. Nevertheless, we notice that lookup overheads are consistently higher for CURLING at all levels. Since CURLING does not allow the exchange of state across peering links, FIND messages reach higher levels in the hierarchy more frequently. In the case of the top-most domains, the workload increase is 3.9-fold on average.

Figure 5(d) shows the cumulative distribution of hop counts for name resolution, showing that DONA requires less hops to resolve an IO identifier as RHs maintain resolution entries from peering domains, i.e FIND messages often reach their destination without reaching a Tier-1 domain. On average, DONA requires 3.26 hops per resolution against 3.86 hops in CURLING i.e., DONA leads to a 16% average reduction in hops. Finally, Figure 5(e) shows the cumulative distribution of single hop transmissions for registrations in both DONA and CURLING. Registering an IO with DONA requires 35.57 single hop transmissions on average, against only 5.20 in the case of CURLING i.e., an overhead of more than 684%, due to the propagation of registrations over peering links. Following Koponen et *al.* [6], we can estimate a total average/median load of 66.13/66.13 Gbps (DONA) and 39.61/40.85 Gbps (CURLING) of registration traffic at Tier-1 domains, for a total number of 10^{13} IOs with an average lifetime of two weeks and a REGISTRATION size of 1 KB. These are obviously non-negligible data rates for control plane operations, even though they only refer to a few Tier-1 ASes, and are within reach of current technology capabilities. The expected rates drop for Large ISPs, Small ISPs and Stub networks to 24.28/28.23 Gbps, 10.32 Gbps/64.15 Mbps, 1.35 Gbps/1.98 Mbps (DONA) and 1.82 Gbps/197.08 Mbps, 19.18/11.90 Mbps, 1.98/1.98 Mbps (CURLING).

V. SUMMARY AND CONCLUSIONS

In this paper we have attempted to shed some light on an important and still open research challenge in the area of ICN i.e., inter-domain name resolution. We focused on DONA and CURLING, two route-by-name inter-domain NRSs, that combine the benefits of both BGP-compliant, low latency routing/forwarding and flat, semantic-free name spaces. Our goal was to reveal the effects of the inter-domain topology structure on the distribution of state across the inter-network, further quantifying these effects based on the real topology of the Internet. Our findings can be summarized as follows:

- Both DONA and CURLING lead to extensive replication of state across the inter-domain topology, reaching a replication factor of 1702.64 (DONA) and 27.34 (CURLING).
- The distribution of state across the topology is heavily skewed. Tier-1 ASes must maintain 100% (DONA) and 59.89% (CURLING) of the entire state in the inter-network, leading to the requirement for the deployment

of data centers at the scale of 20K servers. However, more than 81% (DONA) and 90% (CURLING) of all ASes only hold state for local content leading to moderate deployment costs for the vast majority of ASes.

- By *not* forwarding registrations across peering ASes, CURLING reduces the state size 62-fold across the inter-domain topology, and even up to 679-fold for stub domains, compared to DONA. As a result, small ASes are not flooded by state from higher tier peers and CURLING deployment costs become proportional to AS (cone) sizes.
- The state reduction in CURLING comes at the cost of an average 2.78-fold increase of lookup processing load compared to DONA. The topmost domains pay the highest penalty i.e., 297% additional lookups on average.
- By allowing name resolution paths to follow peering links, DONA shortens name resolution paths by 16% compared to CURLING. Nonetheless, this comes at the cost of a 684% higher total traffic overhead for name registrations. This translates to bandwidth requirements in the order of several Gbps for Tier-1 domains and Large ISPs in DONA, while CURLING reduces this requirement to a few hundred Mbps for most Large ISPs.

Based on these results, we can conclude that while DONA's ability to scale to Internet-wide deployments is questionable, especially in terms of the state size required, CURLING's requirements are far more modest and, possibly, feasible to support. Though this comes at a slight performance degradation for the end users due to the increase of name resolution path lengths, it is still considered encouraging, as route-by-name NRSs, by respecting BGP policies and AS peering agreements, are easier to deploy than their lookup-by-name competitors. To this end, we have previously discussed the feasibility of leveraging resource virtualization and scalable storage abilities of private, public or hybrid cloud facilities for the cases of larger (Tier-1/Large) and smaller ASes, respectively [15]. However, reducing the amount of state aggregated at the higher levels of the hierarchy, as well as registration traffic, especially at Tier-1 providers, is still important. As part of ongoing work on this problem, we are currently studying the feasibility of Bloom filter-based aggregation schemes for name resolution state.

ACKNOWLEDGMENT

This research has been co-financed by the European Union (European Social Fund-ESF) and Greek national funds through the Operational Program Education and Lifelong Learning of the National Strategic Reference Framework (NSRF)-Research Funding Program: Aristeia II/I-CAN, as well as the EPSRC UK project COMIT (grant no. EP/K019589/1) and the CHIST-ERA / EPSRC UK project CONCERT (grant no. EP/L018535/1).

REFERENCES

[1] G. Xylomenos *et al.*, "A Survey of Information-Centric Networking Research," *IEEE Communications Surveys and Tutorials*, vol. 16, no. 2, pp. 1024–1049, 2014.

[2] PURSUIT Project, *PURSUIT Project Home Page*, http://www.fp7-pursuit.eu, 2011.

[3] V. Jacobson *et al.*, "Networking named content," in *Proc. of the ACM CoNEXT*, 2009, pp. 1–12.

[4] W. K. Chai *et al.*, "Curling: Content-ubiquitous resolution and delivery infrastructure for next-generation services," *IEEE Communications Magazine*, vol. 49, no. 3, pp. 112 –120, 2011.

[5] C. Dannewitz, "NetInf: An Information-Centric Design for the Future Internet," in *Proc. of the GI/ITG KuVS Workshop on The Future Internet*, 2009.

[6] T. Koponen *et al.*, "A data-oriented (and beyond) network architecture," in *Proc. of the ACM SIGCOMM*, 2007, pp. 181–192.

[7] Google, *We knew the web was big, July 2008*. [Online]. Available: http://googleblog.blogspot.com/2008/07/we-knew-web-was-big.html

[8] D. Evans, "The Internet of Everything: How More Relevant and Valuable Connections Will Change the World," White Paper, February 2012. [Online]. Available: http://www.cisco.com/web/about/ac79/docs/innov/IoE.pdf

[9] C. Dannewitz *et al.*, "Hierarchical DHT-based name resolution for information-centric networks," *Elsevier Computer Communications*, vol. 36, no. 7, pp. 736–749, 2013.

[10] J. Rajahalme *et al.*, "On name-based inter-domain routing," *Elsevier Computer Networks*, vol. 55, pp. 975–986, 2011.

[11] K. V. Katsaros *et al.*, "On inter-domain name resolution for information-centric networks," in *Proc. of the IFIP TC 6 Conference on Networking*, 2012, pp. 13–26.

[12] V. Ramasubramanian *et al.*, "The Design and Implementation of a Next Generation Name Service for the Internet," in *Proc. of the ACM SIGCOMM*, 2004, pp. 331–342.

[13] M. Walfish *et al.*, "Untangling the web from DNS," in *Proc. of the USENIX NSDI*, 2004, pp. 17–17.

[14] CAIDA, "The IPv4 Routed /24 AS Links Dataset," August 2013. [Online]. Available: http://www.caida.org/data/active/ipv4_routed_topology_aslinks_dataset.xml

[15] X. Vasilakos *et al.*, "Cloud Computing for Global Name-Resolution in Information-centric Networks," in *Proc. of the Symposium on Network Cloud Computing and Applications (NCCA)*, 2012, pp. 88–94.

[16] D. Perino *et al.*, "A Reality Check for Content Centric Networking," in *Proc. of the ACM SIGCOMM Workshop on Information-centric Networking (ICN)*, 2011, pp. 44–49.

[17] M. A. Kyle Schomp *et al.*, "DNS Resolvers Considered Harmful ," in *Proc. of the ACM Workshop on Hot Topics in Networks (Hotnets)*, 2014.

[18] S. Sevilla *et al.*, "iDNS: Enabling information centric networking through The DNS," in *Proc. of the IEEE INFOCOM Workshop on Name Oriented Mobility*, 2014, pp. 476–481.

[19] N. J. A. Harvey *et al.*, "SkipNet: A Scalable Overlay Network with Practical Locality Properties," in *Proc. of the USENIX Symposium on Internet Technologies and Systems (USITS)*, 2003, pp. 9–9.

[20] N. Fotiou *et al.*, "H-pastry: An inter-domain topology aware overlay for the support of name-resolution services in the future internet," *Computer Communications*, no. 0, pp. –, 2015.

[21] A. Rowstron *et al.*, "Pastry: Scalable, decentralized object location and routing for large-scale Peer-to-Peer systems," in *Proc. of the Middleware Conference*, 2001, pp. 329–350.

[22] R. Oliveira *et al.*, "The (in)completeness of the observed internet AS-level structure," *IEEE/ACM Transactions on Networking*, vol. 18, pp. 109–122, 2010.

[23] X. Dimitropoulos *et al.*, "Graph annotations in modeling complex network topologies," *ACM Transactions on Modeling and Computer Simulation*, vol. 19, pp. 17:1–17:29, 2009.

[24] K. Katsaros *et al.*, "GlobeTraff: A Traffic Workload Generator for the Performance Evaluation of Future Internet Architectures," in *Proc. of the International Conference on New Technologies, Mobility and Security (NTMS)*, 2012, pp. 1–5.

[25] S. M. Besen *et al.*, "The evolution of Internet interconnection from hierarchy to "Mesh": Implications for government regulation," *Information Economics and Policy*, vol. 25, no. 4, pp. 235 – 245, 2013.

[26] K. Chen *et al.*, "Where the sidewalk ends: extending the internet as graph using traceroutes from P2P users," in *Proc. of the ACM CoNEXT*, 2009, pp. 217–228.

[27] CAIDA, "The IPv4 Routed /24 AS Links Dataset," January 2011. [Online]. Available: http://www.caida.org/data/active/ipv4_routed_topology_aslinks_dataset.xml