Towards a Quantitative Comparison of Location-Independent Network Architectures

Zhaoyu Gao Arun Venkataramani James Kurose Simon Heimlicher
{gaozy, arun, kurose, heimlicher}@cs.umass.edu
School of Computer Science, University of Massachusetts Amherst

ABSTRACT

This paper presents a quantitative methodology and results comparing different approaches for *location-independent* communication. Our approach is empirical and is based on real Internet topologies, routing tables from real routers, and a measured workload of the mobility of devices and content across network addresses today. We measure the extent of network mobility exhibited by mobile devices with a home-brewed Android app deployed on hundreds of smartphones, and measure the network mobility of Internet content from distributed vantage points. We combine this measured data with our quantitative methodology to analyze the different cost-benefit tradeoffs struck by location-independent network architectures with respect to routing update cost, path stretch, and forwarding table size. We find that more than 20% of users change over 10 IP addresses a day, suggesting that mobility is the norm rather than the exception, so intrinsic and efficient network support for mobility is critical. We also find that with purely name-based routing approaches, each event involving the mobility of a device or popular content may result in an update at up to 14% of Internet routers; but, the fraction of impacted routers is much smaller for the long tail of unpopular content. These results suggest that recent proposals for *pure* name-based networking may be suitable for highly aggregateable content that moves infrequently but may need to be augmented with addressing-assisted approaches to handle device mobility.

Categories and Subject Descriptors

C.4.3 [**Computer Systems Organization**]: COMPUTER-COMMUNICATION NETWORKS—*Network Architecture and Design*

Keywords

Location-independence; mobility; network architecture

SIGCOMM'14, August 17–22, 2014, Chicago, IL, USA.

ACM 978-1-4503-2836-4/14/08 ...$15.00.
http://dx.doi.org/10.1145/2619239.2626333 .

1. INTRODUCTION

Providing an abstraction of *location-independent communication*—enabling communication using fixed names without concern for changing network locations—has been a long-time goal of networking research. For example, proposals designed to provide seamless host mobility [36] target an abstraction of the form `connect(host_id)`; a number of proposals for information-centric networking [28, 13, 22, 16, 19] target an abstraction of the form `get(content_name)`.

Why is the current TCP/IP Internet seen as falling short of this goal? A common criticism is the so-called location-identity conflation problem [42]. The Internet uses an IP address to identify an interface as well as its network location. As a result, connections break when an endpoint changes network addresses, requiring application-layer workarounds to provide a semblance of seamless mobility. Advocates of information-centric networking argue that the Internet, having inherited a century-old, tethered, device-to-device communication abstraction from the wired telephony world, is poorly-suited for an Internet dominated by content traffic and a communication abstraction requiring endpoints to first obtain the network location of a host serving the requested content instead of simply procuring a copy from *any* convenient location. In response, researchers have proposed a number of designs to refactor naming, addressing, and routing in order to enable location-independent communication [22, 12, 36, 11, 28, 13, 22, 16, 19].

Our work is motivated by the observation that, although many architectural proposals share location-independence as a key design goal, there has been little prior research quantitatively comparing the different cost-benefit tradeoffs struck by these architectures in achieving that goal. One reason for the paucity of cross-architectural comparisons may be that network architecture is considered by some as part science and in good part *art* [51]. Another is that until recently, most Internet architecture efforts rarely went beyond paper designs, so a lack of a reasonably complete design specification and protocol-level implementation made it hard to justify investing research effort comparing them; however programs such as GENI[4], FIA[6], and FIRE [3] are changing this state of affairs. We believe that a quantitative comparison of different approaches for location-independent communication is timely and indeed critical for gaining a deeper understanding of cross-cutting architectural principles.

Our contribution is a quantitative methodology and empirical results comparing different location-independent architectures based on a common set of metrics, namely, routing update cost, path stretch, and forwarding table size, in a context where both devices and content may be mobile. We find that a number of existing approaches for location independence can be categorized into one of three "puristic" classes (§2): *indirection routing* (e.g., Mobile IP or GSM), *name-to-address resolution* (e.g., DNS and IP), and *name-based routing* (e.g., TRIAD [22], ROFL [12] or NDN [28]). Our methodology is empirical, comparing these three pure approaches using realistic Internet topologies, routing tables from real routers, and a measured workload of content and device mobility across addresses in today's Internet (§3).

To measure *device mobility* across network locations, we have developed an Android app, NomadLog, deployed on over 372 Android devices and has been collecting device mobility data for over 14 months (§4). To measure *content mobility,* we deployed a system across distributed Planetlab nodes to estimate the rate of change of network addresses of popular content domain names (including those delegated to CDNs) as well as unpopular content domain names. Combining this measured mobility workload with our evaluation methodology, we arrive at the following key findings.

1. Pure name-based routing entails a prohibitively high router update cost to handle device mobility today, e.g., some routers may be impacted by up to 14% of all device mobility events today.
2. Over 20% of users change network addresses over 10 times a day, suggesting that high mobility is the norm rather than the exception. The median user spends around 25% of a day at ASes at least two AS hops away from the dominant AS, implying a commensurate path stretch for indirection routing approaches.
3. Pure name-based routing imposes a much lower update cost for content mobility today, e.g., with best-port forwarding, routers are impacted by at most 6% of popular content mobility events, and are hardly impacted by mobility of the long tail of unpopular content.

A key implication of our findings is that name-based routing approaches in their puristic form may be better suited for content alone, but need to be augmented with addressing-assisted approaches such as DNS, Mobile IP, or a next-generation name resolution service [49] in order to serve as a general-purpose replacement for the TCP/IP Internet. Our findings also show the important differences between device and content mobility, as well as the emerging importance of the strategy layer [28, 55] in content-oriented architectures.

The remainder of this paper is structured as follows. We begin with a discussion of background and related work (§2) followed by a detailed description of our evaluation methodology (§3); the NomadLog app to measure network mobility of devices (§4); an expository analytical model for the path stretch vs. update cost trade-off (§5); the results of our empirical evaluation of the cost-benefit trade-offs of different approaches for location independence in the context of device (§6) and content (§7) mobility; a confession of limitations and open questions (§8); and conclusions (§9).

2. BACKGROUND AND RELATED WORK

Despite the enormous body of work on refactoring naming, addressing, and routing so as to enable location-independent communication, most known approaches fundamentally take one of just three different approaches: (1) indirection routing; (2) name resolution; (3) name-based routing.

The fact that there are not too many different approaches to enable location independence should not be surprising – in order to enable a communication abstraction of the form `connect(B,A)` where B and A are fixed names of endpoint principals (e.g., hosts, content, or services), each of which could be simultaneously multi-homed at different locations at any point in time and the set of these locations could suddenly change, endpoint B must resort to one of three options in order to send the first packet to A—(1) send the packet to one (or a small number of) network router(s) that know(s) A's current location; (2) acquire knowledge of A's current location through an extra-network service and send the packet to that location; or (3) send the packet stamped with the name A to any router, trusting a coordinated routing and forwarding strategy across routers to deliver the packet to A's current location. These three approaches are illustrated in Figure 1. Let us next consider how a number of proposed network architectures embody these approaches.

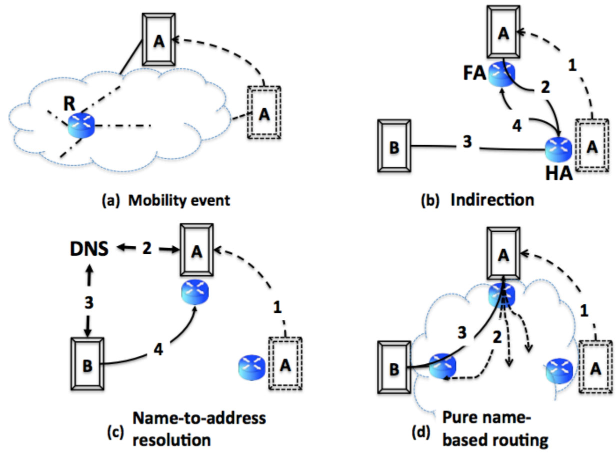


Figure 1: A mobility event (a) and three purist approaches (b, c, d) to handle them.

Indirection routing translates the problem of fixed-name routing to fixed-address routing as the first step. A name resolution infrastructure may be infrequently queried to translate an endpoint's name to a *home* address that rarely changes by design. A home agent in the home network as in GSM, Mobile IP and other architectures (e.g., [30]) or a randomly chosen rendezvous location in i3 [46] is responsible for maintaining up-to-date *visited* addresses of its subscribers and detouring packets to them. The main strength of indirection is simplicity – an endpoint remains completely oblivious to endpoint mobility. The downside is path stretch – in order to remain oblivious, a sender must detour all packets through the destination's home agent [41], even if the two endpoints happen to be in the same local network. The update cost of indirection routing is minimal because an endpoint only needs to update its home agent upon each mobility event. The forwarding table size at routers remains unchanged and is equal to the number of disaggregated prefixes (e.g., ≈400K at core routers in today's Internet).

Name resolution approaches rely on an extra-network service like DNS as the first step in all network communication in order to resolve a destination's name to its current address. A name resolution service is an integral part of the Internet as well as a number of network architectures such as HIP, AIP, LISP, Nimrod, MobilityFirst, and XIA. As in indirection routing, the update cost of handling a mobility event is minimal because it suffices to make a single update at a logically centralized service like DNS. Network routing is based on structured addresses, so the forwarding table size at routers depends on the "aggregateability" of endpoint IP addresses, e.g., if all endpoints move randomly, then routers would have to store on the order of 4B IP addresses instead of 400K prefixes today. The path stretch depends only upon the underlying network routing; a shortest-path routing network has no stretch by definition while policy-driven interdomain routing as in the Internet can incur significant path inflation [44]. Enabling location-independence through name resolution only adds a lookup latency at connection setup time, but does not add additional data path stretch compared to underlying network routing.

Name-based routing, unlike both approaches above, routes directly over names without using addresses at all. Examples include flat-label routing architectures such as ROFL [12] or SEATTLE [29] (for a single enterprise network) or information-centric routing architectures such as TRIAD [22] and NDN [28]. From an endpoint's perspective, eliminating addresses is a silver bullet as the network comes with intrinsic support for location-independent communication. However, achieving this abstraction purely at the network layer without inducing significant stretch or long outage times upon mobility events is nontrivial. Quantifying this trade-off is an important focus of our work.

All three broad approaches above have a number of other advantages and disadvantages (e.g., incremental deployability, manageability, security, handoff outage times, etc.). We have omitted a discussion of these in line with our goal of a nuanced analysis of the cost-benefit tradeoffs with respect to quantitative metrics such as path stretch, update cost, and forwarding table size. An explicit non-goal of our work is to determine the "best" among existing location-independent architectures or to propose yet another new one.

2.1 Related work

The contributions of our paper lie at the intersection of network measurement—empirically characterizing the changing connectivity of both devices and content to the Internet—and network architecture–analyzing the cost-benefit tradeoffs with respect to several metrics for different approaches for location-independence). Consequently, there is related past research in several areas.

Network mobility measurement. Many studies have empirically characterized physical human mobility among access points or base stations and discussed the impact of physical user mobility patterns on network performance and design. Human mobility traces have been collected from diverse access networks such as WLANs [31, 25, 14], Bluetooth networks [14], and cellular networks [23, 39, 26]. Some [23, 39, 15] have related human mobility patterns to AP and base station resource use, and have found [23, 39] that the extent of users' physical mobility is low and concentrated among a small number base stations within a provider's network with infrequent visits to other base stations in that network. Zarifis et al. [57] characterize metro-level path inflation (rather than mobility itself) experienced by mobile users accessing Google, identifing inter-domain routing, peering, and carrier topology as possible causes. Similarly, other studies have focused on the measured performance (throughput or delay) of WiFi or cellular connections in the wild [37, 43, 18] but focus on connection performance rather than on mobility itself. Acculoc [53] and CelloScope [9] describe applications that take measurements capturing location-specific information about cellular connections, but do not focus on mobility across locations or access networks as in our work.

Individual WiFi and cellular measurements include data from a single type of network, and more importantly characterize *physical* user/content mobility among access points or base stations, rather than changing points of attachment to the Internet (i.e., as characterized by changing IP addresses and AS numbers). Our NomadLog measurements characterize this latter aspect of mobility *across* networks, rather than physical mobility in space—a critical distinction. For example, a physically mobile user might maintain the same IP address as they move among base stations in a provider network; for our purposes, this user is considered stationary as their IP address does not change despite physical mobility. Conversely, a user may "move" from one network (cellular) to another (WiFi) while hardly moving physically.

In-network name resolution. A number of studies have considered the performance of in-network name resolution. [55] compares the performance of an instantiated NDN forwarding plane [28] with traditional IP forwarding, with an emphasis on security and congestion mitigation. [10] compares the performance of network-based name resolution in an instantiated NDN context versus a logically centralized approach as in MobilityFirst's GNS approach [49], considering forwarding table size as a function of an abstract, parameterized model of name aggregation. Neither of these works consider either content or device mobility—the key consideration in our present paper.

Several recent efforts have considered name-based content retrieval in a mobile environment. [50] considers information dissemination in a linear V2V network using NDN, focusing primarily on the impact of protocol timer values on performance; our present paper focuses on mobility among multiple networks with general topologies and is aimed at a broad comparison between different location-independent network architectures. Proxies and/or indirection points (such as the HLR in cellular networks and home agent in Mobile IP) have been a common feature of many architectures supporting mobility, including recent proposals for NDN-like architectures. [33, 24] both adopt a proxy-based approach and rely on underlying tunneling or the existence of IP addresses to deliver content. Most recently, Kim [30] proposes the use of an indirection point where mobile content publishers and subscribers can register (content publishers) mobility-related name changes or query (content subscribers) for new names associated with a mobile publisher. [30] presents a protocol for NDN-like architectures for real-time, single-sender-single-receiver scenarios, conjecturing that "*... content providers, and their locations are relatively stable. Hence, the mobility problem for the 'stored contents' is limited to the scope of user-side mobility*". Our measurement results (§7) suggest otherwise. Further, our focus here is on pure name-based routing (or in-network name resolution), rather than proxies or indirection points.

Mobile IP and indirection. Much of the analysis of Mobile IP [40] has focused on analyzing proposed enhancements that improve handoff performance (e.g., [17]) or minimize signaling overhead (e.g., [58, 52]). Our goal, instead, is to empirically characterize location update rates as nodes change their point of attachment to the Internet, and the average "distance" from their home network—two key performance considerations for any architecture with a home-agent-like component. As noted above, there are many measurement studies of mobility among individual access points and base stations, but none of these characterize the rate at which users change their IP address or AS-affiliation, and it is this latter aspect of mobility (not intra-network mobility that is "invisible" outside that network) that determines update rates at home agents.

The triangle routing problem in Mobile IP [41], which results in longer paths between a sender and receiver when routed through an indirection point (the home agent) is well known, and enhancements to allow direct routing have been proposed [41]. The tradeoff between providing shortest-path routes versus the overhead entailed (e.g., in forwarding table size) is characterized by compact routing results [32, 47]. For example, with N flat identifiers, to be within 3x stretch of shortest-path, each router needs to maintain $\Omega(N)$ forwarding entries.; for up to 5x stretch, it is $\Omega(\sqrt{(N)})$. These lines of work, that seek to either develop new protocols to minimize path stretch due to indirection or theoretically characterize path stretch versus table-size tradeoffs, address different challenges than our work that empirically characterizes a mobile device's distance from "home."

3. EVALUATION METHODOLOGY

In this section, we explain our methodology to evaluate the cost-benefit tradeoffs struck by different location-independent approaches to handle mobility. The metrics of interest are routing update cost, forwarding table size, and path stretch for mobility of devices as well as content.

3.1 Intradomain device or content mobility

Consider a simple shortest-path routing network as shown in Figure 1(a) and a singly-homed endpoint A that changes its address from 22.33.44.55 belonging to the 22.33.44.0/24 subnet to 22.33.88.55 belonging to the 22.33.0.0/16 subnet. With name-to-address resolution (indirection routing), A simply needs to update DNS (its home agent) and data traffic can subsequently flow directly (indirectly) to A's new location. With a purely network-layer approach, some routers such as R may need to update their forwarding behavior in order to continue routing packets correctly to A. Whether R needs to update its forwarding table depends on whether R's shortest-path forwarding entries for 22.33.44.0/24 and 22.33.0.0/16 point to different output ports. If they do, say to ports 5 and 3 respectively, then R must introduce another entry `[22.33.44.55/32, 3]` in its forwarding table so that longest-prefix matching continues to route correctly to A. More generally, a router needs to update its forwarding table if an endpoint moves from one longest-matching prefix to another in its forwarding table, each pointing to different output ports. We refer to such mobility events as an endpoint A being *displaced* with respect to router R.

Displaced content names induce an update at name-based routers by moving across hierarchical name spaces in a manner analogous to endpoints moving across IP address spaces. An example is shown in Figure 2(b), wherein say `/20thCenturyFox/StarWars-EpisodeIV` moves to `/Disney/StarWars-EpisodeIV` because of a distribution rights transfer. Router Q must update its forwarding table if Q maintained different output ports for the corresponding longest-matching prefixes `/20thCenturyFox/*` and `/Disney/*`, i.e., if `StarWars-EpisodeIV` is displaced with respect to Q.

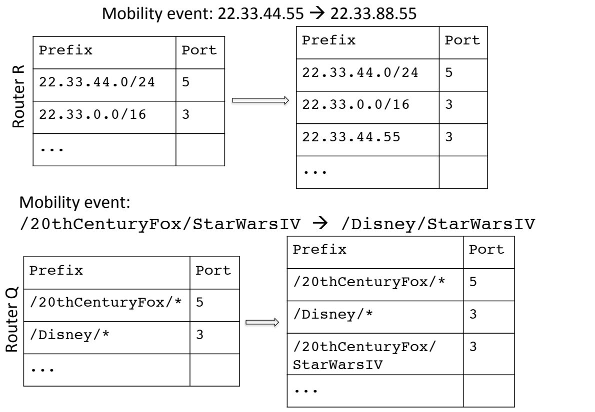


Figure 2: Example of a device or content *displaced* w.r.t. a router's forwarding table.

3.2 Interdomain mobility

From the example above, it is clear that the *update cost* of mobility depends on at least three factors: (1) the nature of network mobility patterns of endpoints; (2) the physical topology of the network; and (3) route selection policy. The policy used to select routes, e.g., shortest-path routing or BGP-style policy-driven route selection, matters because that is what determines the forwarding table at a router. Unlike shortest-path intradomain routing for which it is straightforward to answer whether or not a mobility event causes an update at any given router, it is much harder to answer this question in an interdomain network like the Internet driven primarily by policy routing.

One strawman is to use publicly available intradomain [35] and interdomain [2] topologies and combine them with a simple model of Internet routing, namely, prefer customers over peers over providers, then use AS path length to break ties, then use early-exit routing to further break ties, and so on. Unfortunately, given that even our knowledge of the Internet's physical AS-level as well as intradomain topology is incomplete [38], these models serve as a poor substitute for real Internet routing that is messier, e.g., prior studies [34] have found that interdomain routes predicted by such a model had barely 30% predictive accuracy.

Consequently, in this work, we work with a small set of real Internet routers whose RIBs (route information bases) and some route preference metrics are available to us. These RIBs already incorporate the global impact of the Internet's physical topology and import and export policy decisions made by other routers. We use these RIBs to derive the corresponding FIBs and ask, in a manner analogous to intradomain routing above, whether or not the router would update its forwarding table in response to a mobility event without having to simulate global Internet routing.

3.3 Multihomed device or content mobility

The methodology above implicitly assumed a singly-homed device or content principal. We next extend it to model the impact of mobility of multihomed principals on the update cost and forwarding table size at routers. For ease of exposition, we explain the model in the context of content mobility, but note that it applies to both device and content mobility.

3.3.1 Update cost

In order to assess the update cost of a mobility event, we begin with a *mobility workload* consisting of domain names and the set of all IP addresses to which they resolve (as measured for content in §7 from distributed vantage points). Consider a domain d, e.g., `graphics.nytimes.com`, and let $\mathit{Addrs}(d, t_1)$ denote the set of all IP addresses to which it resolves at time t_1. A mobility event refers to a change in the set to $\mathit{Addrs}(d, t_2)$ at a future time t_2. Does this mobility event cause a content router R to update its forwarding behavior? That is, is there a difference between the sets $\mathit{FIB}(R, d, t_1)$ and $\mathit{FIB}(R, d, t_2)$, where $\mathit{FIB}(R, d, t)$ denotes the set of all eligible output ports to which router R could choose to forward packets destined to d at time t?

To answer the above question, we distinguish between two *forwarding strategies— best-port forwarding* and *controlled flooding* —that respectively forward packets on at most a single best output port (like today) and forward packets to more than one eligible output port. The set of all eligible ports, $F(R, d, t)$, is determined as the set of output ports at R corresponding to the set of IP addresses $\mathit{Addrs}(d, t)$ at that time, each of which in turn is computed using R's FIB as for interdomain mobility above (§3.2). Best-port forwarding picks a single output port denoted as $\mathit{best}(\mathit{FIB}(R, d, t)$. The update cost depends on which of the two forwarding strategies is in use and is defined as follows.

The *update cost* of a mobility event at a router is 1 for best-port forwarding if there is a change in its best forwarding port, i.e., $\mathit{best}(\mathit{FIB}(R, d, t_1)) \neq \mathit{best}(\mathit{FIB}(R, d, t_2))$ for the mobility event $\mathit{Addrs}(d, t_1)$ to $\mathit{Addrs}(d, t_2)$; otherwise it is 0. The update cost is 1 for controlled flooding if there is a change in the set of all eligible output ports, i.e., $\mathit{FIB}(R, d, t_1) \neq \mathit{FIB}(R, d, t_2)$ for the mobility event $\mathit{Addrs}(d, t_1)$ to $\mathit{Addrs}(d, t_2)$; otherwise it is 0.

3.3.2 Forwarding table size and aggregateability

The forwarding table size at a content router depends on two factors: (1) forwarding strategy (as defined above) and (2) *aggregateability*, a metric that captures the extent to which forwarding tables can be compacted by relying on longest-prefix matching and inherent network locality in the content name space. To formally define this metric, we introduce some notation. Consider a set S of hierarchically organized domain names such as `yahoo.com, cnn.com, mit.edu, travel.yahoo.com`, etc. and a router R employing some forwarding strategy to route to these domains. For each domain $d \in S$, let $\mathit{FIB}(R, d)$ denote the (set of) forwarding output port(s). We refer to the set of forwarding entries $\{[d, \mathit{FIB}(R, d)]\}_{d \in S}$ as the *complete forwarding table.* Let $d_1 \prec d_2$ mean that d_1 is a strict subdomain of d_2, e.g., `travel.yahoo.com` $\prec$ `yahoo.com`. If $d_1 \prec d_2$ and $\mathit{FIB}(R, d_2) = \mathit{FIB}(R, d_1)$, then we say that the forwarding entry for d_1 is *subsumed* by d_2 with longest-prefix matching. For example, as shown in Figure 3, the entry `[travel.yahoo.com, 2]` is subsumed by the entry `[yahoo.com, 2]` as longest-prefix matching obviates storing an entry explicitly for the former, but a separate entry is needed for `[sports.yahoo.com, 5]`. We refer to the forwarding table consisting of the subset of entries in the complete forwarding table that excludes all subsumed entries as the *LPM forwarding table.*

We define *aggregateability* at a router as the ratio of the size of the complete forwarding table to the size of the corresponding LPM forwarding table.

Prefix	Port
yahoo.com	2
~~travel.yahoo.com~~	~~2~~
sports.yahoo.com	5
cnn.com	2
mit.edu	4
...	

Figure 3: Example of a content forwarding entry subsumed because of longest-prefix matching.

3.3.3 Limitations and extensions

Our simple model of controlled flooding implicitly focuses on control plane costs, not on forwarding traffic. In particular, it implies that the update cost of controlled flooding increases with the rate of mobility events and is at least as high as that of best-port forwarding. However, there exist forwarding strategies that trade off update cost against increased traffic in the forwarding plane for which neither implication is true. For example, consider a hypothetical strategy wherein a router R computes $\mathit{FIB}(R, d, t)$ based on the union of all past addresses observed for destination d, i.e., $\bigcup\{\mathit{Addrs}(d, t_i)\}_{t_i < t}$ instead of just $\mathit{Addrs}(d, t)$. In this case, the update cost at R attributable to d will soon approach 0 if d rarely visits completely new network locations no matter how frequently it flits across previously visited locations. R's controlled flooding strategy could simply forward across all of the corresponding output ports in $\mathit{FIB}(R, d, t)$, ensuring that at least one copy reaches d.

The strategy above reduces update cost (and, potentially, path stretch) but increases the costs of forwarding traffic and forwarding table size. Recognizing the fungibility of these costs—update cost, forwarding table size, and forwarding traffic—allows for other intriguing architectural combinations that have not been considered in this paper. Extending our evaluation to incorporate more general trade-offs including forwarding traffic is an interesting line of future research.

4. NETWORK MOBILITY MEASUREMENT

In this section, we describe an Android app, NomadLog [5], a development effort we undertook motivated both by the intractable or simplistic nature of purely theoretical analyses (refer §5) as well as the lack of public data on *network mobility* (unlike geographic mobility) of mobile devices.

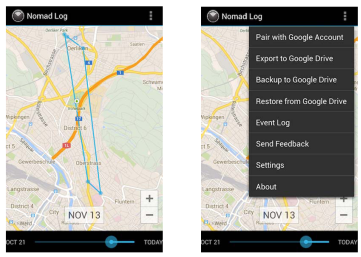


Figure 4: Screenshots of the NomadLog app.

NomadLog is a lean app explicitly designed to measure the changing IP addresses of Android devices and does little else. The value proposition for potential downloaders is that they get cool visualizations of their network mobility statistics on a map, but realistically we expect people with a bent for citizen science or personal analytics to be more likely to install the app. The app runs in the background and attempts to record the public-facing IP address upon each connectivity event. A *connectivity event* is when either network interface wakes up and successfully connects to a cellular or WiFi network or disconnects from a network. Upon a connectivity event, the app contacts a server we maintain in order to determine its public-facing IP address, so addresses are logged only if they are usable for Internet connectivity, so automatic connects to WiFi APs blocked by a paywall or authentication page will not be logged unless successful.

The app is designed to be as inconspicuous as possible and is conservative in using battery power and data traffic as follows. First, its event-driven design obviates polling the network interfaces. Second, except for a single small message to infer its public-facing IP address, the app stores all data logs locally until it is both connected to power and WiFi; at this point, it attempts to transfer previously untransferred log files to a postgresql database that we maintain. Each entry logged in the database is in the following format.

device_id | *time* | *ip_addr* | *net_type* | *(lat, long)* ···

device_id is the hashed device identifier used to track the user while providing limited privacy; *time* is the event's timestamp; *net_type* indicates the type of the connected network, WiFi or cellular; and *(lat, long)* is the geolocation.

The user's geolocation is recorded only with user permission at install time and is collected only if the GPS is already on and has obtained a recent reading on behalf of some other app, i.e., NomadLog itself does not consume GPS resources. Users can either visualize their mobility statistics through the app or use their device identifier in order to access their data from the app's website from any device. Except for the (hashed) *device_id*, we do not maintain any other information directly identifying a user in our database.

We have acquired 372 users, mostly from the United States, Europe, and South America through word-of-mouth publicity alone. The user recruitment and data collection spanned over 14 months from Mar 2013 to May 2014. Because different users downloaded the app at different times and a small fraction uninstalled the app quickly (a learning experience that helped us engineer the resource and data usage optimizations), we removed users who ran the app for less than a day. Our analysis is based on daily statistics of network mobility, so our conclusions are unlikely to be biased significantly by the differing measurement period across users.

5. PATH STRETCH VS. UPDATE COST: AN EXPOSITORY ANALYTIC MODEL

In this section, we develop a simple analytic model to quantify the tradeoff between path stretch and update cost with two goals. The first is expository and helps us better appreciate the fundamental nature of the tradeoffs evaluated empirically in this paper. The second is to suggest that anything beyond simple toy topologies is difficult to analyze theoretically, making a stronger case for empirical analysis.

Our model and results are similar in spirit to theoretical work on compact routing (see [32] for a survey) that has focused on path stretch vs. forwarding table size tradeoffs, but the difference is that we model endpoint mobility and its update costs. Informally, compact routing results say that in order to achieve small stretch over shortest-path routing in a network with arbitrary (or flat) endpoint identifiers, roughly all routers must maintain an entry for roughly all endpoint identifiers. The question we ask is: in order to achieve small

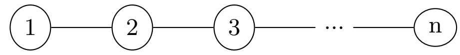


Figure 5: A simple chain network topology.

stretch over shortest-path routing, how many routers need to be updated when endpoints move across routers?

Intuitively, if every router is updated upon every mobility event, then the path stretch could be minimized. Also intuitively, if updates are restricted to at most one router (like a home agent) per mobility event, the path stretch could be as high as the diameter of the network as packets to the mobile endpoint must go through the only router that knows its whereabouts. We formally model these tradeoffs for several toy topologies and explain one of these in detail below.

5.1 Chain topology

Consider a chain network topology as shown in Figure 5 with routers numbered from 1 to n, and a user u that randomly hops from one router to another. This mobility can be modeled using a discrete-time Markov process as follows. Let $L_t(u)$ be a random variable representing u's location at time slot t. If the transition probability $P(L_{t+1}(u) = j | L_t(u) = i) = \frac{1}{n}$, then the steady-state distribution of $L_t(u)$ is uniform, i.e., $P(L_t(u) = i) = \frac{1}{n}$.

5.1.1 Indirection routing

Let $H(u)$ denote u's home agent that keeps track of its current location. We define $path\_stretch_t$ as the hop-count distance from the home agent to an endpoint's location at time t. If $H(u)$ were chosen randomly (as would be the case in a network where different nodes were equally likely to be homed at any router and moved around randomly), the stretch is derived as follows. Below $dist(u, v)$ refers to the hop count distance between u and v.

$$\begin{aligned} & E[path\_stretch_t] = E[dist(H(u), L_t(u))] \\ & = \textstyle\sum_{i=1}^{n} P(H(u) = i) \sum_{j=1}^{n} P(L_t(u) = j)|i - j| \\ & = \textstyle\sum_{i=1}^{n} \frac{1}{n} \sum_{j=1}^{n} \frac{1}{n}|i - j| \\ & = \frac{n^2-1}{3n} \simeq \frac{n}{3} \qquad \text{(refer [21] for details)} \end{aligned}$$

Thus, with indirection, the expected path stretch is $\frac{n}{3}$ and the update cost is 1 per mobility event. Note that we define path stretch as the additive distance over the shortest-path length (as opposed to a multiplicative factor) as this matches what we are able to measure empirically (§6.3).

5.1.2 Name-based routing

With name-based routing, the path stretch is 0 if we assume that routers are designed to always maintain forwarding tables corresponding to shortest-path routing. What is the *aggregate update cost*, i.e., the fraction of routers that must be updated, to achieve this minimal stretch? We derive it as follows. Suppose each router has three ports, a left (right) port connecting to the leftwise (rightwise) adjacent router, and a local port connected to the local subnet. Then a router i must update its forwarding table whenever an endpoint either moves from any leftward router to a rightward router or vice-versa, or moves from any router other than i to i or vice-versa. The expected update cost at router k is

$$\begin{aligned} & E[update\_cost_k] = P(L(u) < k) \cdot P(L(u) \geq k) \\ & \quad + P(L(u) = k) \cdot P(L(u) \neq k) + P(L(u) > k) \cdot P_k(L(u) \leq k) \\ & = \frac{k-1}{n}\frac{n-k+1}{n} + \frac{1}{n}\frac{n-1}{n} + \frac{n-k}{n}\frac{k}{n} \end{aligned}$$

The expected aggregate update cost across all n routers is

$$\begin{aligned} & E[update\_cost] = \frac{1}{n}\textstyle\sum_{k=1}^{n} E[update\_cost_k] \\ & = \frac{n^3+3n^2-n}{3n^3} \simeq \frac{1}{3} \qquad \text{(refer [21] for details)} \end{aligned}$$

Thus, for name-based routing, the aggregate update cost per mobility event is 1/3 and the path stretch is 0.

5.1.3 *Summary of results*

We have similarly quantified the aggregate update cost vs. path stretch tradeoff for other toy topologies and the simplistic random mobility model. The proofs are deferred to a

Topology	Indirection		Name-based routing	
	stretch	update cost	stretch	update cost
Chain	$n/3$	$1/n$	0	$1/3$
Clique	1	$1/n$	0	1
Binary tree	$2\log_2 n$	$1/n$	0	$\frac{2\log_2 n}{n-1}$
Star	2	$1/n$	0	$\frac{1}{n+1}$

Table 1: Path stretch vs. aggregate update cost.

technical report [21]. Note that we have omitted the analysis for a DNS-based approach above as the data path stretch is 0 (ignoring a constant lookup overhead in the connection initiation step) and the expected update cost is simply $O(1)$ (to the DNS), irrespective of topology.

6. DEVICE MOBILITY

In this section, we combine the measured NomadLog data in §4 with the methodology in §3 to evaluate the cost-benefit tradeoffs of different approaches to handle mobility of devices across networks. We begin by analyzing the extent of network mobility across devices in the NomadLog data.

6.1 Extent of device mobility across networks

Figure 6 shows the distribution across users of the average number of distinct network locations per day visited by a user[1]. The trace consists of 372 users each of which is present for at least one day in our trace. The median number of ASes, IP prefixes, and IP addresses visited per day are 2, 2, and 3 respectively. This observation is consistent with the expectation that users typically move across a cellular, home, and work address in the course of a day.

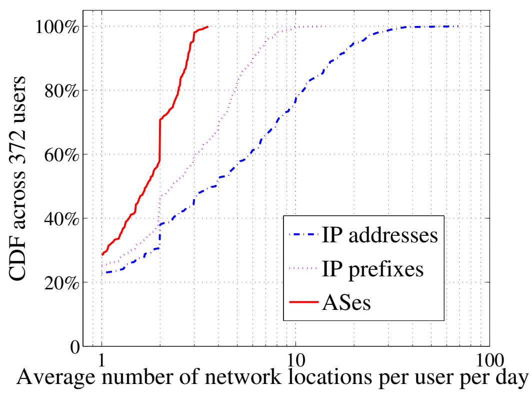


Figure 6: Average number of distinct network locations per day visited by users.

Figure 7 shows the distribution of the average number of transitions across network locations per day by a user. The number of AS transitions shows a lot more variation compared to the number of distinct ASes in Figure 6, which is

[1]We use the terms *user* and *device* interchangeably.

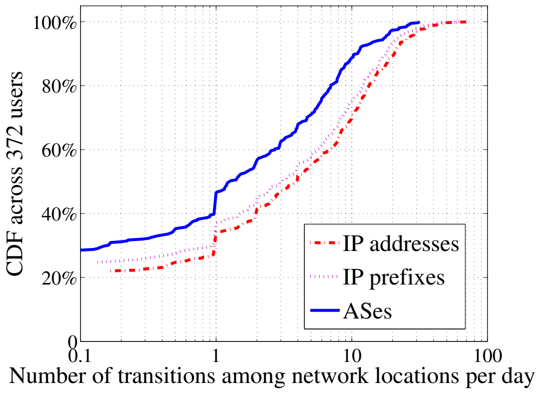


Figure 7: Average number of transitions across network locations per day made by users.

because a user can switch many times between a small number of ASes. The number of transitions depends upon the user's physical mobility, network performance or outage patterns, and behavioral patterns, e.g., some users may prefer to use WiFi for some apps but use LTE for others or make these choices depending on current network quality. The maximum and minimum numbers of average AS transitions per day are 31.6 and 0.25 respectively. The median user transitions across roughly one AS and three IP addresses.

6.2 Update cost of device mobility

To provide the abstraction of location-independence communication, a device must update its changing network addresses *somewhere*, either in DNS, or at its home agent(s), or at routers. For the first two cases, the update cost is straightforward and directly corresponds to the rate of address transitions as shown above. For the third case, we estimate the update cost using the methodology described in §3.1 and §3.2 respectively.

6.2.1 *Using Routeviews data*

To this end, we use RIBs from 12 BGP-speaking Routeviews routers [8]. The set of routers includes four in Oregon (labeled Oregon-1 to Oregon-4) and one each in Virginia, California, Georgia, Mauritius, London, Tokyo, Sydney and Sao Paulo. A single entry in a router's RIB lists several attributes of a single inter-domain route towards a given prefix. Typically, there are several routes to any given prefix and the set of all prefixes covers the entire IP address space.

ip_prefix | next_hop | local_pref. | metric | AS path | ··· |

To construct the FIB from a router's RIB, we need to compute a rank ordering of all of the routes for a single prefix. We apply the following rules in priority order based on typical BGP policies and the priority order suggested [8].

1. A route with a higher local_preference value is preferred. As local preference values are not available for most of the router dumps, we simply rely on the customer > peer > provider policy using standard techniques for inferring AS relationships [20].
2. A route with a shorter AS path is preferred.
3. A route with a smaller MED value is preferred.

Note that the above rules approximate typical BGP policies. The numerical value of `local_preference` is uniformly 0 in these RIBs, so we use AS relationships instead in the first rule. The RIBs also do not have sufficient information to implement early- or late-exit policies that are typically of higher priority than multi-exit discriminator values.

6.2.2 Update cost of name-based routing

To determine if the movement of a user from one address to another results in a change in a router's forwarding behavior, we must determine whether the the mobility event induces a change in the output port corresponding to the highest ranked route for the user's address. We use the next_hop AS path attribute as a proxy for the output port, implicitly assuming that the forwarding output port changes if and only if the next_hop attribute changes, an assumption believed to hold more often than not [45]. In practice, different next_hop addresses can correspond to the same output port, and the same next_hop can correspond to different output ports at different times because of intradomain concerns, so we may under- or over-estimate the actual update cost.

Figure 8 shows the update cost at all of the routers using the RIB data published on Mar 31, 2014. The update cost is shown as the fraction of all mobility events that induce a forwarding update at the router, also referred to as the *update rate*. The results show that the update rate can be as high as 14% at some routers. The routers at Mauritius and Tokyo experience hardly any updates, which is unsurprising as most of our users are located in the USA, Europe, and South America, so their mobility is less likely to impact distant routers. We verified that the Georgia router has a much lower next-hop degree compared to the Oregon routers, which could plausibly explain its lower update rate.

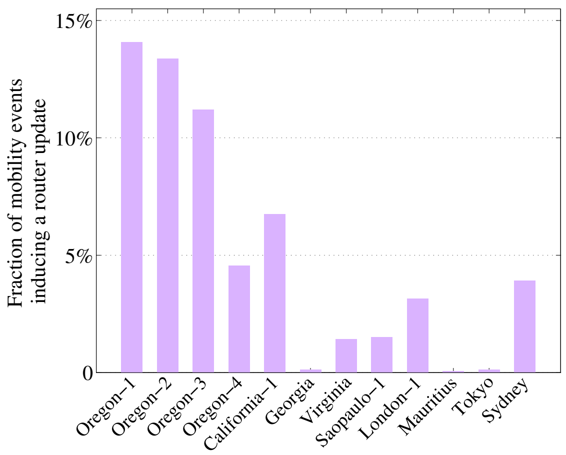


Figure 8: Fraction of device mobility events inducing a routeviews router update.

Sensitivity analysis. We further investigated the sensitivity of the above results to: (1) time, (2) the set of routers, and (3) the mobility workload. For the first, we repeated the experiment for 20 different days and found that, at every router, the standard deviation of the update rate is less than 0.005 (or 0.5%). For the second, we repeated the experiment using 13 RIPE routers [7] located in 13 different cities, 10 of which are distinct from the RouteViews set, and found qualitatively similar conclusions: the update rate at the median (most affected) router in the RouteViews and RIPE sets were 3.15% (14%) and 2.74% (11.3%) respectively.

For the third, we resorted to a significantly larger dataset [54] consisting of 7137 users from UMass IMAP servers that recently became available to us. These traces measure *user* mobility in a sense distinct from *device* mobility (an exception to the footnote in §6.1) as observed from a specific (but universally used) application's perspective, so the two traces are not directly comparable. Our preliminary analysis as above, using user mobility as a proxy for device mobility, shows that the update rates observed at all 25 RouteViews and RIPE routers are highly correlated with those for the NomadLog data, with a correlation coefficient of 0.88 (with more details deferred to the technical report [21]).

Back-of-the-envelope calculations.

Update cost: Combining the above results with the results in the previous section, we can arrive at a crude estimate of the absolute rate of updates induced at routers because of user mobility today. For example, if 2 billion smartphones change network addresses three (seven) times per day like our median (mean) user, and 3% of these mobility events induce an update at a router, the corresponding update rate is 2.1K/sec (4.8K/sec). These numbers are prohibitively high for even high-end routers today. Although it is possible to redesign router control planes to handle such high update rates using more compute resources and a software-defined control plane, it is difficult to justify this computation cost and the bandwidth cost of propagating these updates to a large number of Internet routers. In comparison, it is straightforward to handle this aggregate load by distributing it across a large number of DNS servers or home agents.

Forwarding table size: Combining the typical update rate of 3% with the fact that users typically spend 30% of a day away from the dominant IP address (see §6.3 below) suggests that a typical router would have to maintain extra forwarding entries for ≈1% of all devices that are *displaced* (as defined in §3.1) with respect to it at any given time.

6.3 Data path stretch with device mobility

The update cost analysis above induces no data path stretch (over underlying Internet routing) for name-based routing or for a DNS-based approach. However, indirection routing inflates the data path because of triangle routing via the home agent. Next, we quantify this path stretch overhead.

6.3.1 Displacement from dominant location

We introduce the notion of a *dominant location*, i.e., the network location where a user spends the largest fraction of time compared to all other locations in the course of a single day. Figure 9 shows the distribution across all days and all users of the percentage of time spent in the dominant location. For example, the plot shows that over 40% of users spend around 70% of their day at the dominant IP address and around 85% of their day at the dominant AS.

The dominant location is a natural candidate for a home agent in an indirection routing architecture. In order to compute path stretch, we need to determine C→H→M/C→M, where C→H→M is the sum of the network latency from a correspondent C to the home agent H and that from H to the mobile M, and C→M is the network latency of routing directly from C to M. We do not have a dataset of correspondents initiating communication with mobile devices (because it is largely not possible today to initiate communication to mobile devices), so instead we simply quantify the displacement of mobile users from their home agents in network distance, i.e., the latency of the path from H→M.

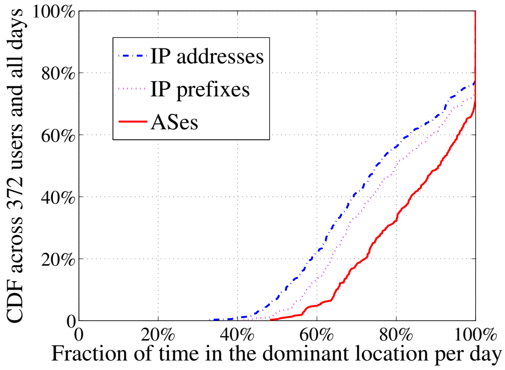


Figure 9: The CDF of time that each user spends in the dominant location.

6.3.2 *Path stretch of indirection routing*

In order to determine the network distance from a user's dominant location (or home) to their current location in the NomadLog trace, we rely on iPlane, a system that uses daily traceroute measurements from a large number of distributed vantage points to stub networks in order to predict the route (and its latency) between an arbitrary pair of IP addresses. Although using iPlane is convenient, it comes with two severe caveats for our analysis. First, iPlane returns valid responses for only 5% of the dominant and current IP address pairs in our trace; this is because it is designed to return responses only if it has sufficient traceroutes that enable it to constructed a predicted route using segments of measured routes. Second, even when iPlane returns a response, the predicted route may be inaccurate.

Figure 10 shows the distribution of network latencies across the dominant-to-current IP address pairs for which iPlane returned a predicted route. The median displacement delay from the dominant location is around 50ms and the corresponding AS hop count is 4. Recognizing the limitations of the estimates obtained using iPlane, we use a different technique to estimate a lower bound on the AS hop count of the displacement from home. We compute the length of the shortest AS path from the home to the current location using the Internet's AS-level physical topology [2] (even if this route may not exist in the AS-level routing topology). The median AS hop count of this shortest AS path is 2, suggesting that mobile users typically wander two or more ASes away from the home AS.

7. CONTENT MOBILITY

In this section, we evaluate the cost-benefit tradeoffs in terms of update cost, forwarding table size, and path stretch for content mobility. We begin by describing the procedure used to measure content mobility today.

7.1 Content mobility measurement

We begin with two sets of content domain names: a *popular* set and an *unpopular* set. The former is the set of the top 500 domains ranked according to popularity by Alexa [1] and the set of all of their subdomains. The latter is the least popular 500 domains and their subdomains in a list of the top 1 million domain names also ranked by popularity. We explicitly obtain a list of subdomains because Alexa ranks "websites" or top-level enterprise domains, e.g., `nytimes.com` or `yahoo.com`, but not their subdomains like `graphics.nytimes.com` or `travel.yahoo.com`. More importantly, the distribution of popular, bulky content that is ideally suited to name-based routing techniques is often outsourced to CDNs, and a common technique to achieve this is to CNAME-alias subdomains, e.g., `graphics.nytimes.com` is aliased to the canonical name `static.nytimes.com.edgesuite.net` that is in turn aliased to `a1158.g1.akamai.net` that finally gets dynamically resolved to one or more IP addresses close to the querying client or its local name server.

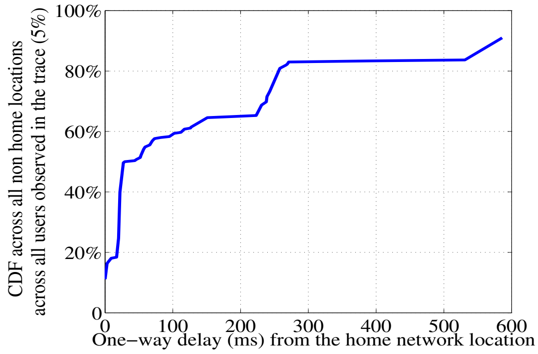


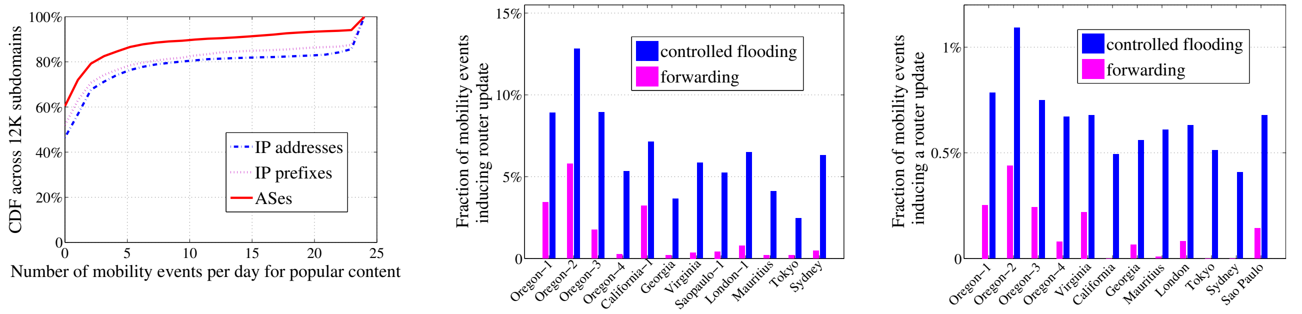
Figure 10: Distribution of delay for the IP addresses pairs that get response from iPlane.

The dynamic nature of resolution of domain names to IP addresses (either because they are resolved by a CDN delegate in a locality-aware manner or because of DNS-based load balancing employed by the origin server) means that any single vantage point will see only a subset of all IP addresses from where a domain's content may be potentially served. Our methodology to assess content mobility (in §3.3) relies on monitoring any changes to the set of all IP addresses corresponding to a domain name. This methodology implicitly assumes that a purely name-based routing network will announce a content domain name from all of the locations (including CDN locations) where it resides today.

In order to measure a reasonably complete set of IP addresses to which each domain name maps, we conduct a measurement distributed across 74 Planetlab nodes that are chosen from as many different countries as possible and all continents (except Africa where Planetlab nodes were unavailable). We conducted the measurements for a three week period from May 1 to May 22, 2014. Each node resolves each domain name once every hour, thereby observing a subset of the domain's IP addresses at that time. A central controller node collects measurements obtained from all of the vantage points and merges them in time so that for each domain name for each hour, the set of IP addresses is the union of all IP addresses obtained from all vantage points for that domain. As the measurement interval once per hour, precise time synchronization is not necessary. The measurement is done just once per hour per domain because our list of subdomains corresponding to the 500 most popular Alexa domains contains 12,342 entries, so a much higher rate would overwhelm some nodes or trigger security alarms.

7.2 Update cost of content mobility

Figure 11(a) shows the extent of daily mobility of popular content (i.e., the 12,342 subdomains obtained from the most



(a) The average number of transitions for popular content mobility events (b) Fraction of popular content mobility events inducing a router update (c) Fraction of unpopular content mobility events inducing a router update

Figure 11: Results on the extent of content mobility and its impact on the update cost at routers.

popular 500 domains). The median number of changes in the set of IP addresses per day is 2 (the similarity to device mobility being just coincidental) and the maximum is bounded at 24 because of our hourly measurement procedure.

Figure 11(b) shows the update rate at each of the twelve routers because of mobility events involving popular content. The plot shows that up to 13% of content mobility events can induce an update at some routers when controlled flooding (i.e., forwarding on all ports matching any of the domain's IP addresses) is used. However, at most 6% of the mobility events induce an update at any of the routers in our dataset when best-port forwarding is used. The reason is that although there may be some flux in the set of addresses corresponding to a domain name, the address that is the closest to any given router rarely changes because most of the addresses in the set remain unchanged, i.e., unlike devices that jump seemingly randomly from one address to an unrelated address, content locations do not change arbitrarily.

Figure 11(c) shows the corresponding result for unpopular content or the least popular 500 domains and their subdomains with a popularity rank of near about one million. The update cost for unpopular content is dramatically lower than that for popular content; at most 1% of updates induce an update at any of the routers even with controlled flooding. With best-port forwarding, almost none of the routers experience any update during the course of our measurement period (the median is 0.08%). This result is not surprising as unpopular content is unlikely to be delegated to CDNs and is probably served only from a small number of network locations that rarely change; these multiple locations if at all are chosen mainly for fault-tolerance or load balancing purposes rather than proximity to clients, so they rarely change. We further explicitly estimated [21] the fraction of domains delegated to CDNs in our trace for unpopular content to be only 1.6% compared to 24.5% for popular content.

7.3 Forwarding table size

Figure 12 shows the aggregateability (as defined in §3.3.2) for the 500 most popular domain names on 12 routeviews routers by using the best-port forwarding strategy. We see that the aggregateability at different routers varies between 2× to 16×, which suggests a commensurate reduction in the forwarding table sizes at these routers compared to the total number of popular content domain names. Unpopular content domain names in our dataset have hardly any sub-

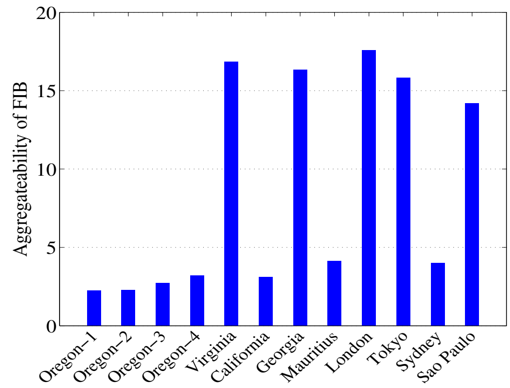


Figure 12: FIB aggregateability of popular content.

domains, which implies that content routers would have to nominally store one entry each for the long tail of unpopular content domain names, unless a different ontological structure that helps compact routing information is used to rename them. With respect to forwarding table size, unpopular content domain names present a challenge similar to device names, i.e., they both entail one forwarding entry per principal at a router unless a location-aware scheme (like IP addresses or geo-location) is used to "rename" them.

Back-of-the-envelope calculations.

Performing a calculation similar to that at the end of §6, if we assume 1B content domain names (noting that DNS has ≈150M domains), an update rate of 2/day, and a 0.5% likelihood of inducing an update at a router, the router would receive at most 100 updates/sec. Furthermore, for the vast majority of long-tail domains ranked below 1M, the update cost is likely to be even lower even if controlled flooding is used as the forwarding strategy. Finally, with best-port forwarding, the router update cost due to the mobility for vast majority of (unpopular) content is near-zero.

8. LIMITATIONS AND OPEN ISSUES

We began with an ambitious goal but have only managed to scratch the surface. Below, we list the limitations of our evaluation, caveats attached to our findings, and open issues.

The hundreds of users in our device mobility dataset may not be a representative sample of smartphone users even today, leave alone in a future Internetwork. However, given the lack of existing datasets on network mobility (unlike physical mobility) of devices, we believe our measurements are an important and necessary first step for developing meaningful mobility datasets in order to compare different architectures. An alternative might have been to develop an abstract model (e.g., the equivalent of a random-waypoint model [56] for network mobility of devices or content), but this abstract model would hardly be any more convincing unless validated through measurements (such as those presented here).

The unrepresentativeness critique applies also to other aspects of the data in our evaluation including the use of today's routing topologies, routing policies, access network diversity, content locations, etc. that may be radically different in the future and may indeed be influenced by the very architectures we seek to evaluate, were they to be widely adopted. The empirical nature of our evaluation comes with the necessary caveat that our findings are limited to network environments like today's Internet; but we do note that our findings are unlikely to change qualitatively if the extent of device or content mobility were perturbed by large factors.

Our characterization of content mobility by measuring the change in resolved IP addresses associated with domain names implicitly assumes that a purely name-based routing network will announce content domain names from all of the locations—including CDN locations—where they resides today. It is possible, even likely, that popular content will be announced from far fewer primary locations with commensurately less flux in a name-based architecture that can leverage on-path caching (as in NDN). However, that would only strengthen our favorable assessment of name-based routing for handling the update cost of content mobility. What was not clear to us when we began this work was whether routing directly over a hierarchical domain name space would scale even in today's Internet; our findings lean towards the affirmative (unlike analyses based on abstract models [10]).

Our scope of evaluation as well as goals suffer from some limitations. First, we have evaluated three pure strategies for location-independence but not the many possible combinations of these strategies in a network architecture (e.g., exploiting indirection points within a name-based routing system [27]). Second, network architecture itself is indeed part science and part art. Not everything may be easily quantifiable; what is easily quantifiable may not be the most pressing concern. Nevertheless, our position is that pushing the envelope of what is quantifiable is valuable for scientific discourse, and our work is a first step towards that goal.

Even within the confines of the three purist approaches, our empirical focus constrained the nature of the trade-offs we could evaluate. For name-based routing, we could empirically evaluate metrics such as update cost and forwarding table size, but not other control plane metrics such as routing convergence delays or data plane metrics such as forwarding traffic or user-perceived path stretch with on-path caching. We note however that on-path content caching can benefit most architectures, including ones based on name resolution like the Internet (e.g., transparent caching today) or MobilityFirst [48], but does not suffice to ensure reachability to at least one copy of the requested content.

Our methodology considers one class of controlled flooding strategies that generalize routing and forwarding strategies used in today's Internet. However, as noted in §3.3.3, the fungibility of costs between update cost, forwarding table size, and forwarding traffic allows for other exotic architectural alternatives whose cost-benefit trade-offs are harder to analyze. Investigating these more sophisticated forwarding strategies is an interesting avenue of future research, as also alluded to by calls for a stateful forwarding plane [55].

9. CONCLUSIONS

The intellectual pursuit of a location-independent communication abstraction has long intrigued networking researchers, and has in no small part influenced the design of many clean-slate Internet architectures. However, despite sharing this common goal, there has been little prior work on quantitatively comparing the different cost-benefit tradeoffs struck by these architectures in accomplishing this goal.

As a first step towards addressing this gap, we have developed a quantitative methodology to empirically evaluate three puristic approaches that drive the designs of a number of location-independent network architectures. We combine this methodology with measured traces of device mobility and content mobility on the Internet using realistic physical and routing topologies. Based on measured network mobility patterns of hundreds of devices of NomadLog, an Android app we developed explicitly for this goal, and hundreds of content domains including those delegated to content distribution networks, we find that pure name-based routing induces a prohibitively high update cost at routers because of device mobility, but induces a far lower update cost in conjunction with simple forwarding strategies for most of today's content that happens to exhibit high locality. Taken together, our results suggest that recent proposals for name-based networking in their puristic form are better suited for content distribution alone, but may need to be augmented with addressing-assisted approaches like DNS or Mobile IP in order to handle device mobility, so as to serve as a general-purpose replacement for the TCP/IP Internet.

Acknowledgments. This research was funded in part by CNS-1040781. We thank the paper's reviewers, the shepherd Scott Shenker, and Antonio Rocha, Dave Oran, Lixia Zhang, and participants at NSF FIA meetings for their feedback.

10. REFERENCES

[1] Alexa: `http://www.alexa.com/`.

[2] CAIDA: `http://www.caida.org`.

[3] FIRE: Future Internet Research and Experimentation: `http://cordis.europa.eu/fp7/ict/fire/`.

[4] GENI: `http://www.geni.net`.

[5] Nomad Log: `https://nomadlog.net/`.

[6] NSF Future Internet Project: `http://www.nets-fia.net/`.

[7] RIPE: `http://www.ripe.net/`.

[8] RouteViews: `http://www.routeviews.org/`.

[9] Celloscope: `http://celloscope.net`, 2014.

[10] A. Baid, T. Vu, and D. Raychaudhuri. Comparing Alternative Approaches for Networking of Named Objects in the Future Internet. In *IEEE INFOCOM NOMEN Workshop*, 2012.

[11] H. Balakrishnan, K. Lakshminarayanan, S. Ratnasamy, S. Shenker, I. Stoica, and M. Walfish. A Layered Naming Architecture for the Internet. In *ACM SIGCOMM*, 2004.

[12] M. Caesar, T. Condie, J. Kannan, K. Lakshminarayanan, I. Stoica, and S. Shenker. ROFL: Routing on Flat Labels. In *ACM SIGCOMM*, 2006.

[13] A. Carzaniga and A. L. Wolf. Content-Based Networking: A New Communication Infrastructure. In *NSF Workshop on Developing an Infrastructure for Mobile and Wireless Systems*. Springer-Verlag, 2002.

[14] A. Chaintreau, P. Hui, J. Crowcroft, C. Diot, R. Gass, and J. Scott. Impact of Human Mobility on Opportunistic Forwarding Algorithms. *IEEE Trans. on Mobile Computing*, 6(6):606–620, 2007.

[15] Y.-C. Chen, J. Kurose, and D. Towsley. A Mixed Queueing Network Model of Mobility in a Campus Wireless Network. In *IEEE INFOCOM*, 2012.

[16] C. Dannewitz. NetInf: An Information-centric Networking Architecture. *Computer Communications*, 36(7), 2013.

[17] S. Das, A. Misra, and P. Agrawal. TeleMIP: Telecommunications-enhanced Mobile IP Architecture for Fast Intradomain Mobility. *IEEE Personal Communications*, 7(4):50–58, Aug 2000.

[18] P. Deshpande, X. Hou, and S. Das. Performance Comparison of 3G and Metro-Scale WiFi for Vehicular Network Access. In *ACM SIGCOMM IMC*, 2010.

[19] N. Fotiou, P. Nikander, D. Trossen, and G. C. Polyzos. Developing Information Networking Further: From PSIRP to PURSUIT. In *Broadband Communications, Networks, and Systems*, pages 1–13. Springer, 2012.

[20] L. Gao. On Inferring Autonomous System Relationships in the Internet. *IEEE/ACM Trans. on Networking*, 9(6), 2001.

[21] Z. Gao, A. Venkataramani, J. Kurose, and S. Heimlicher. Towards a Quantitative Comparison of Location-Independent Network Architectures. *UMass SCS Technical Report, 2014*. `http://web.cs.umass.edu/publication/`.

[22] M. Gritter and D. R. Cheriton. An architecture for content routing support in the internet. In *USENIX Symposium on Internet Technologies and Systems (USITS)*, 2001.

[23] E. Halepovic and C. Williamson. Characterizing and Modeling User Mobility in a Cellular Data Network. In *ACM Workshop on Performance Evaluation of Wireless Ad Hoc, Sensor, and Ubiquitous Networks*, 2005.

[24] F. Hermans, E. Ngai, and P. Gunningberg. Global Source Mobility in the Content Centric Networking Architecture. In *ACM Workshop on Emerging Name-Oriented Mobile Networking Design - [...] (NoM)*, 2012.

[25] W. Hsu, D. Dutta, and A. Helmy. Structural Analysis of User Association Patterns in University Campus Wireless LANs. *IEEE Trans. on Mobile Computing*, 11(11):1734–1748, Nov. 2012.

[26] S. Isaacman, R. Becker, R. Cáceres, M. Martonosi, J. Rowland, A. Varshavsky, and W. Willinger. Human Mobility Modeling at Metropolitan Scales. In *ACM SIGMOBILE MobiSys*, 2012.

[27] V. Jacobson and R. L. Braynard et al. Custodian-based Information Sharing. *IEEE Communications Magazine*, 50(7):38–43, 2012.

[28] V. Jacobson, D. K. Smetters, J. D. Thornton, M. F. Plass, N. H. Briggs, and R. L. Braynard. Networking Named Content. In *ACM SIGCOMM CoNEXT*, 2009.

[29] C. Kim, M. Caesar, and J. Rexford. Floodless in Seattle: A Scalable Ethernet Architecture for Large Enterprises. In *ACM SIGCOMM*, 2008.

[30] D.-h. Kim, J.-h. Kim, Y.-s. Kim, H.-s. Yoon, and I. Yeom. Mobility Support in Content Centric Networks. In *ACM Workshop on Information-Centric Networking*, 2012.

[31] M. Kim, D. Kotz, and S. Kim. Extracting a Mobility Model from Real User Traces. In *IEEE INFOCOM*, 2006.

[32] D. Krioukov, K. Fall, A. Brady, et al. On Compact Routing for the Internet. In *ACM SIGCOMM*, 2007.

[33] J. Lee, D. Kim, M. wuk Jang, and B.-J. Lee. Proxy-based Mobility Management Scheme in Mobile Content Centric Networking (CCN) Environments. In *IEEE International Conference on Consumer Electronics*, 2011.

[34] H. V. Madhyastha, E. Katz-Bassett, T. E. Anderson, A. Krishnamurthy, and A. Venkataramani. iPlane Nano: Path Prediction for Peer-to-Peer Applications. In *USENIX NSDI*, 2009.

[35] R. Mahajan, N. Spring, D. Wetherall, and T. Anderson. Inferring Link Weights Using End-to-end Measurements. In *ACM Sigcomm Workshop on Internet Measurment*, 2002.

[36] R. Moskowitz, P. Nikander, P. Jokela, and T. Henderson. RFC 5201: Host Identity Protocol, Apr 2008.

[37] A. Nikravesh, D. R. Choffnes, E. Katz-Bassett, Z. M. Mao, and M. Welsh. Mobile Network Performance from User Devices: A Longitudinal, Multidimensional Analysis. In *Passive and Active Measurement Conference (PAM)*, 2014.

[38] R. Oliveira, D. Pei, W. Willinger, B. Zhang, and L. Zhang. The (in) completeness of the observed internet as-level structure. *IEEE/ACM Trans. on Networking*, 18(1), 2010.

[39] U. Paul, A. Subramanian, M. Buddhikot, and S. Das. Understanding Traffic Dynamics in Cellular Data Networks. In *IEEE INFOCOM*, 2011.

[40] C. Perkins. RFC 3220: IP Mobility Support for IPv4, 2002.

[41] C. Perkins and D. Johnson. Route Optimization for Mobile IP. *Cluster Computing*, 1(2):161–176, 1998.

[42] J. Saltzer. RFC 1498: On the Naming and Binding of Network Destinations, Aug. 1993.

[43] J. Sommers and P. Barford. Performance Comparison of 3G and Metro-Scale WiFi for Vehicular Network Access. In *ACM SIGCOMM Internet Measurement Conference*, 2012.

[44] N. Spring, R. Mahajan, and T. Anderson. The Causes of Path Inflation. In *ACM SIGCOMM*, 2003.

[45] N. Spring, R. Mahajan, and D. Wetherall. Measuring ISP Topologies with Rocketfuel. In *ACM SIGCOMM*, 2002.

[46] I. Stoica, D. Adkins, S. Zhuang, S. Shenker, and S. Surana. Internet Indirection Infrastructure. In *ACM SIGCOMM'02*.

[47] M. Thorup and U. Zwick. Compact Routing Schemes. In *ACM Symposium on Parallel Alg. and Arch. (SPAA)*, 2001.

[48] A. Venkataramani, J. Kurose, D. Raychaudhuri, K. Nagaraja, M. Mao, and S. Banerjee. MobilityFirst: A Mobility-Centric and Trustworthy Internet Architecture. *ACM SIGCOMM Computer Comm. Review (CCR)*, 2014.

[49] A. Venkataramani, A. Sharma, X. Tie, H. Uppal, D. Westbrook, J. Kurose, and D. Raychaudhuri. Design Requirements of a Global Name Service for a Mobility-Centric, Trustworthy Internetwork. In *COMSNETS*, 2013.

[50] L. Wang, A. Afanasyev, R. Kuntz, and R. Vuyyuru et al. Rapid Traffic Information Dissemination Using Named Data. In *ACM Workshop on Emerging Name-Oriented Mobile Networking Design - [...] (NoM)*, 2012.

[51] J. Wroclawski. All Hat No Answers: Some Issues Related to the Evaluation of Architecture. In *Spring 2013 NSF FIA PI meeting, Salt Lake City, http://www.nets-fia.net/Meetings/Spring13/FIA-Arch-Eval-JTW.pptx*.

[52] J. Xie and I. Akyildiz. A Novel Distributed Dynamic Location Management Scheme for Minimizing Signaling Costs in Mobile IP. *IEEE Trans. on Mobile Computing*, 1(3):163–175, 2002.

[53] Q. Xu, A. Gerber, Z. M. Mao, and J. Pang. AccuLoc: Practical Localization of Performance Measurements in 3G Networks. In *ACM SIGMOBILE MobiSys*, 2011.

[54] S. Yang, S. Heimlicher, J. Kurose, and A. Venkataramani. User Transitioning Among Networks-a Measurement and Modeling Study. *UMass SCS Technical Report*, 2014. `http://web.cs.umass.edu/publication/`.

[55] C. Yi, A. Afanasyev, I. Moiseenko, L. Wang, B. Zhang, and L. Zhang. A Case for Stateful Forwarding Plane. *Elsevier Computer Communication*, 36(7):779–791, Apr. 2013.

[56] J. Yoon, M. Liu, and B. Noble. Random Waypoint Considered Harmful. In *IEEE INFOCOM*, 2003.

[57] K. Zarifis, T. Flach, S. Nori, D. Choffnes, R. Govindan, E. Katz-Bassett, Z. M. Mao, and M. Welsh. Diagnosing Path Inflation of Mobile Client Traffic. In *Passive and Active Measurement Conference (PAM)*, 2014.

[58] X. Zhang, J. G. Castellanos, and A. T. Campbell. P-MIP: Paging Extensions for Mobile IP. *Kluwer Academic Mobile Networks and Applications*, 7(2):127–141, 2002.