Address and traffic dynamics in a large enterprise network

Thomas Karagiannis, Richard Mortier
Microsoft Research
thomkar@microsoft.com,mort@vipadia.com

***Abstract*—Despite the centrally-managed nature of and critical infrastructure provided by enterprise networks, analyses of their characteristics have been limited. In this paper we examine the dynamics of enterprise networks from two distinct perspectives, namely traffic and addressing. Using a large packet trace spanning approximately 3.5 weeks coupled with diverse other data sources, we pose and answer a series of questions pertinent to understanding the aforementioned aspects of today's enterprise networks. Our analysis results reveal characteristics regarding the geographical spread of traffic observed at a site in the enterprise network, the validity of the client-server model and mobility patterns within the Enterprise. Finally, we discuss the implications of our findings for tasks such as network management and dimensioning.**

I. Introduction

The surge of interest in traffic measurements and characterization over the last decade has lead to a plethora of studies in various aspects of network traffic in the wide-area Internet (e.g., [1]), tier-1 ISPs (e.g., [2]) or university campuses (e.g., [3]). However, despite their significance, enterprise networks have not been the subject of many such analyses ([4] constitutes one of the few exceptions); knowledge of their dynamics is still poor, as limitations such as data sensitivity and access restrictions have inhibited progress in studying traffic characteristics of enterprise networks when compared to, for example, the Internet. Yet, understanding enterprise network traffic patterns is a prerequisite of proper provisioning by network operators, of network dimensioning and network modeling, and can also provide the baseline for anomaly detection. This paper is a step in this direction. Our work provides a description of the address and traffic dynamics of a site that is part of Microsoft's Corporate Network.

Specifically, we first examine traffic dynamics along two perspectives: a) logical and geographical dispersion of traffic within the enterprise network, and b) the validity of the client-server distinction in terms of traffic volumes. Then, taking advantage of routing configuration files and address allocation information we examine address dynamics along two dimensions: a) the meaning of IP addresses as host identifiers and vice versa, i.e., the interpretation of name to IP mappings, and b) host mobility patterns within the larger enterprise network.

Addressing these questions constitutes the main contribution of this paper. To study traffic dynamics we first divide the observed traffic flows in four classes separating data center, local (intra-site), corporate-wide, and Internet traffic. We find that a) temporal patterns depend on the actual traffic classes, b) the majority of the traffic stays within the boundaries of the local site and traffic in the Internet class corresponds to the smallest fraction of all classes, and c) the distribution of the byte contributions per remote site appears heavy-tailed (Section III).

We highlight that defining categories of machines, such as clients, servers or proxies, to account for their traffic contributions seems meaningless within the enterprise. We find that there is a high spatial variability amongst hosts, which naturally suggests the identification of a set of "heavy" users, which contribute most to the overall traffic. This is consistent with characterizing host-behavior as drawn from a sub-exponential distribution (Section IV). However, we note that the composition of this set changes with time – a consequence of spatial variability of hosts – and is in general application specific. On the contrary, we show that defining the set of the most "connected" hosts provides a more indicative feature of the functional role of each host in the network.

With respect to address dynamics our findings include the following: a) We observe that approximately one third of IP address to host name, and host name to IP address mappings do not provide a unique identification of hosts or IPs respectively (Section V-A). b) By analyzing DNS responses and distinguishing hosts that appear in various enterprise sites over time, we provide characteristics of host mobility and travel patterns within the enterprise (Section V-B).

II. Data traces

The results presented in this paper are based principally on a single corpus of packet data collected from the network at Microsoft Research Cambridge (*MSRC*). Fig. 1 presents an overall picture of the MSRC network, and how it fits within the world-wide Microsoft Corporate network (henceforth, *CorpNet*), containing roughly 300,000 hosts connected by approximately 200 routers spread across 100 countries and 6 continents. The MSRC site contains roughly 400 hosts including a small Data Center (*DC*), wired and wireless hosts. The DC provides services for MSRC and CorpNet hosts. Hosts run Microsoft operating systems and software suites, and the site contains a mixture of researchers, admin staff, human resources, and developers. The network runs the OSPF and BGP protocols for internal routing, connecting to the Internet through proxies.

The corpus was collected over a period of 3.5 weeks beginning on Saturday August 27, 2005, and stored in 3.4 TB of disk. Packets were captured using custom tools written against the WinPCap (http://www.winpcap.org) packet capture library. Our

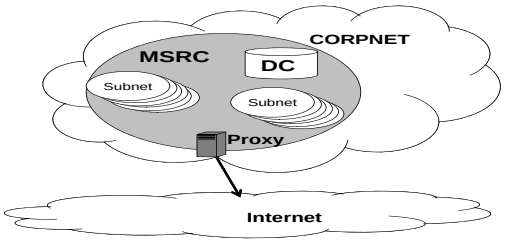

Fig. 1. A view of the MSRC network.

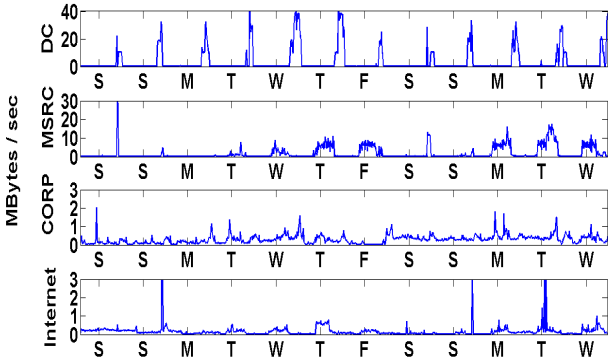

Fig. 2. Traffic over time for the first two weeks of the trace divided in 4 categories: Data center (DC), local traffic (MSRC), traffic to other enterprise sites (CORP) and traffic from and to the public Internet. The first Monday of the trace corresponds to a bank holiday.

site network is configured with each IP subnet corresponding roughly to a wing of a floor mapped to a single VLAN, and so packets were tapped from the network using VLAN-spanning on our site router. Our trace contains packets forming a data corpus of 13 billion unique packets covering 12.5 TB of data.

Finally, we further extended the analysis of the main corpus by using address allocation information and router configuration files from the same period. In particular, we inferred geographic information by extracting the OSPF configuration blocks from the router configuration files. Each such configuration block contains the IP subnets that are originated by the OSPF process at that router. Note that all routers in CorpNet are named according to a convention which encodes their location by city and country and thus traffic sources and destinations can be straightforwardly mapped to cities. The interested reader is referred to the technical report [5] for additional information regarding the collection and analysis of the corpus.

To analyze the collected trace, we constructed flow tables corresponding to 5-minute time intervals, with one record per uni-directional 5-tuple (source IP, destination IP, protocol, source port, destination port) flow observed in each 5-minute period. Each record contains the 5-tuple, the time of the first and last packets in the period, the number of bytes and packets observed, and the application inferred to have generated the traffic. Application was inferred by custom-made deep packet inspection, with care taken to track and account for MSRPC invocations appropriately. Overall, less than 3.5% of the packets in the trace could not be assigned to an application.

We observed 34,397 unique IP addresses in the trace, 591 of which were local to the capture site, MSRC. Of the observed addresses, 23,696 were sources (514 of which were local to

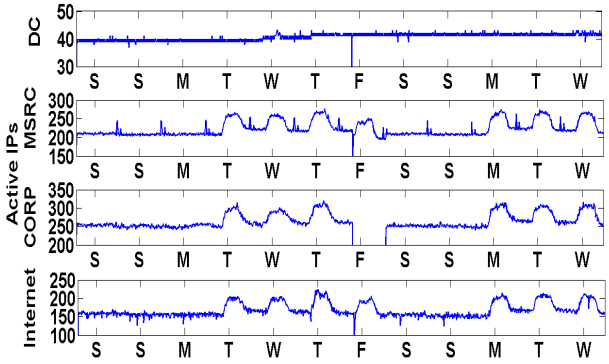

Fig. 3. Number of active local IP over time divided in the 4 categories. DC IPs are constant over time representing always-on server machines.

MSRC) and 33,885 were destinations (582 of which were local to MSRC). The 77 local addresses that received but never transmitted appear to be the result of automated security tools probing for active addresses; similarly, we observed that 9 addresses only transmitted but never received, and all appear to be single-packet aberrations.

In all, the corpus used in the remainder of this paper is a large, coherent set of data providing a useful window into the behavior of a type of network rarely studied previously.

III. Traffic Spread

A distinguishing feature of enterprise networks, when compared to campus networks for example, is that they are both large in size and typically geographically distributed. Furthermore, their configuration, security concerns and the restrictions that these impose, dictate that only a small fraction of enterprise IPs are publicly routable.

Similarly, the configuration of the MSRC network is such that all traffic to the external Internet must be routed via a hierarchy of one or more proxies (Fig. 1). On the other hand, traffic internal to CorpNet, whether it remains within the MSRC site or it goes offsite to other Microsoft installations is routed directly. Thus, in this section we address the question of traffic spread, namely, *what is the network and geographical spread of traffic observed at a site in the enterprise network?*

To this end, we formulate four different classes of observed traffic, which are separated based on subnet and proxy information.

1) *DC*: Traffic that stays within the data center, and accounts mostly for the large overnight backups.
2) *MSRC*: Traffic that stays local within MSRC, excluding the DC traffic.
3) *CorpNet*: Traffic between MSRC and CorpNet, i.e., intra-enterprise traffic.
4) *Internet*: Traffic destined for or received from the public Internet.

Excluding DC traffic, we observe that on the average 79% of the overall traffic stays within MSRC, while CorpNet and Internet amount to only 14.5% and 6.5% of the total traffic respectively (a sample of the traffic for the first two weeks is presented in Fig. 2). The fact that traffic stays mostly within the enterprise has been observed before [4]. However, we provide

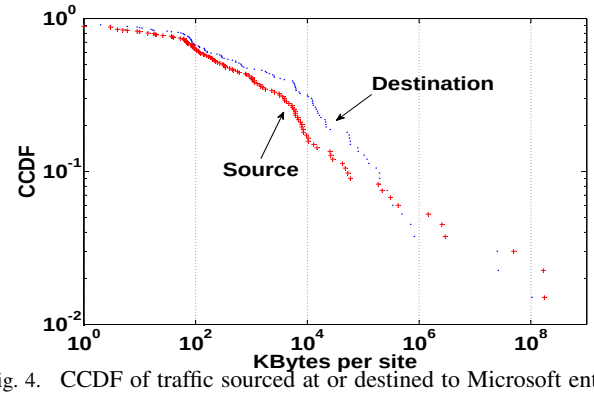

Fig. 4. CCDF of traffic sourced at or destined to Microsoft enterprise sites. The distribution shows evidence of a heavy-tailed distribution with a few sites being the largest volume contributors.

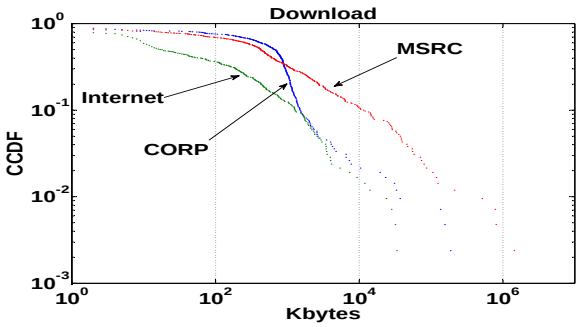

Fig. 5. CCDFs of hourly averages for downloaded bytes for MSRC, CorpNet and the Internet classes. The distributions appear heavy-tailed, with the exception of the CorpNet traffic which appears closer to the exponential distribution.

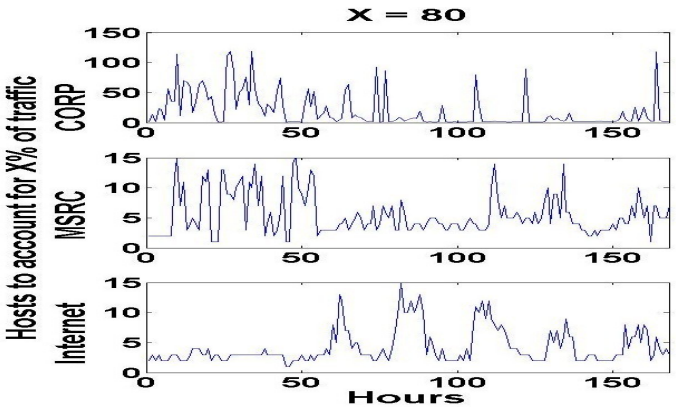

Fig. 6. Number of heaviest host to account for 80% of the total traffic in MSRC, CorpNet and Internet classes. The cardinality of the set of heaviest hosts varies significantly over time.

here a further breakdown, by showing that the majority of the traffic is local within a site of the enterprise, with "intra-enterprise" traffic representing roughly one sixth of the total.

Diurnal patterns are observed only for the MSRC and the Internet classes, while they are not as clear in the CorpNet traffic. Absence of such patterns in CorpNet traffic is due to the fact that this traffic class mostly reflects a set of applications that do not require user action (e.g., receiving email). The large occasional spikes in all classes correspond to large file transfers.

Similarly, Fig. 3 presents a breakdown with respect to the number of local IPs active for each class (i.e., for how many local IPs we observe flows from each traffic class). Here, diurnal patterns are evident in all classes except for DC where the set of active IPs is roughly constant over time. Note that the number of IPs in the CorpNet class is higher on the average when compared to the MSRC class. This occurs because of approximately 50 internal IPs that only communicate with other corporate non-local machines and represent networking equipment such as routers.

We further examined the spread of the traffic across the various sites of the enterprise using geographic information derived from router configuration files. Specifically, we examined the fraction of the traffic sourced at or destined to a particular enterprise site, thus dividing the overall traffic to origin and destination flow pairs between MSRC and remote enterprise sites. Fig. 4 presents the Complementary Cumulative Distribution Function (CCDF) of traffic volumes across all sites observed distinguishing source and destination. Roughly 95% of the traffic is destined to or originating from our local Cambridge site, while the other largest contributors are two US sites, two sites within the UK and one within Europe. The empirical distributions appear heavy-tailed for a range of values (straight line in log-log scale), suggesting a small number of "heavy sites" with respect to their traffic contributions.

IV. Distinguishing Clients from Servers

Since the bulk of the traffic in the network belongs to client-server style applications, the assignment of particular hosts to either "clients" or "servers", i.e., identifying a hosts' functional role in the network, should be straightforward based on observing the byte contributions of the various hosts in the overall traffic. In our network, machines that are physically located inside the data center tend to act predominantly as servers for one particular application. Intuitively, such main site servers (e.g., file servers, proxies, etc) should account for the majority of the traffic in the four classes.

To examine this hypothesis we examine the byte contributions per IP over time. To avoid aggregating over the whole trace which would hide shorter timescale effects, we limit the analysis in hourly intervals of the third week (which contains no network maintenance intervals).

Examining the per-host traffic contributions reveals a small number of "heavy hosts". Heavy-tailed as well as sub-exponential distributions decay more slowly than any exponential distribution. The CCDFs for the hourly average of the downloaded bytes per IP are shown in Fig. 5. Interestingly, while the MSRC and the Internet classes show signs of heavy-tailed distributions (i.e., the tails appear to follow a straight line), the CCDF of the CorpNet class shows less such evidence suggesting that traffic may be distributed more evenly among the local hosts in this class. Similar observations hold for the upload case.

Intuitively, the set of heavy hosts should consist of the various server machines. Thus, tracking the specific servers over time should allow for a comprehensive view of the overall traffic volume across the various classes. Surprisingly, this hypothesis does not hold in our data. Examining the set of heavy hosts across time reveals that not only the set comprises both server and client machines, but it is also highly dynamic with its members significantly varying over time. For example,

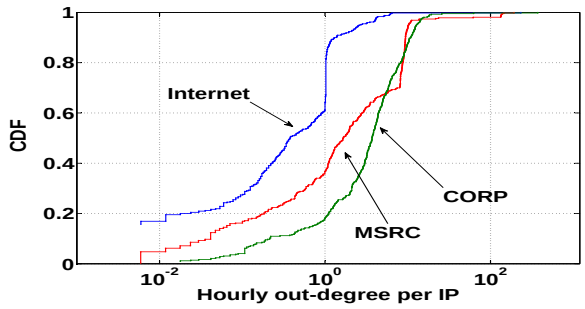

Fig. 7. Out-degree of MSRC hosts.

Fig. 6 illustrates the number of heaviest hosts required to account for 80% of the total traffic across time for the three classes; that is, tracking the cardinality of the set of heaviest hosts, where this set is defined as the machines required to capture x% of the overall traffic. In all cases, the set varies significantly over time, with diurnal patterns appearing only in the Internet traffic class. Fig. 6 suggests that attempting to predict the overall traffic volumes using a potentially static set of servers will not produce accurate estimates (similar results hold for other thresholds, e.g., 95%).

The above discussion suggests that categorizing hosts as either clients or servers in terms of traffic volumes is not straightforward. While intuitively such a distinction makes sense for a single application, it does not for individual hosts. There are two principle reasons for this: First, hosts invariably behave as both clients and servers in different applications, e.g., a web server will be a client to the directory and management services. Second, other applications may not strongly distinguish between clients and servers, e.g., in an enterprise network many machines may be clients of a central file-server while at the same time themselves acting as file-servers to other hosts.

While traffic volume is not an efficient distinctive feature to distinguish client from server hosts, activity of hosts appears more stable over time (see for example Fig. 3). In particular, connectivity information (i.e., which hosts communicate with one another) might allow for such a distinction, as intuitively servers should communicate with most of the local active clients. Thus, we define the out-degree of each host to be the number of other hosts it communicates with, and plot the corresponding CDFs in Fig. 7 by averaging the hourly out-degree per host over the whole trace. This is a similar metric as the fan-out used in [4], where the authors observe that most hosts communicate with local hosts rather than non-enterprise ones. Fig. 7 reinforces this observation by using the three classes of traffic (e.g., out-degree of local hosts to other MSRC hosts, to CorpNet hosts and Internet hosts) and introduces a further separation of local to other enterprise offsite hosts, for which the out-degree appears similar as the local MSRC one.

Close examination of Fig.7 reveals additionally a plateau in the distribution of the out-degree especially for the case of MSRC for larger values of the *x-axis*. This plateau points towards a set of hosts with comparable out-degrees that communicate with most of the internal hosts. Indeed, the IPs comprising this specific component of the distribution correspond to MSRC servers (e.g., proxies, domain controller, etc.) and is stable over time. We can further test this claim by looking at the prevalence of individual hosts in the set of most connected hosts. The prevalence is defined in a similar manner as in [1], and describes the number of intervals a host appears in the most connected set. We define this set as a percentage, x, of the most connected hosts and compare with the same percentage of the heaviest hosts in terms of bytes in Fig.8 (the two sets were calculated in hourly intervals for week three of the trace). We observe that connectivity provides a more stable set of the top hosts across time compared to the set of "heaviest" hosts. For $x = 30$, we observe that roughly 50% of nodes are never in the top-connected set (20% for bytes), while approximately 10% (less than 2% for bytes) of nodes are members of that set for all 168 hourly intervals of week three. Thus, most hosts in the connectivity case are either in or out of the "most-connected" set for all time intervals offering a clear distinction between client and server machines.

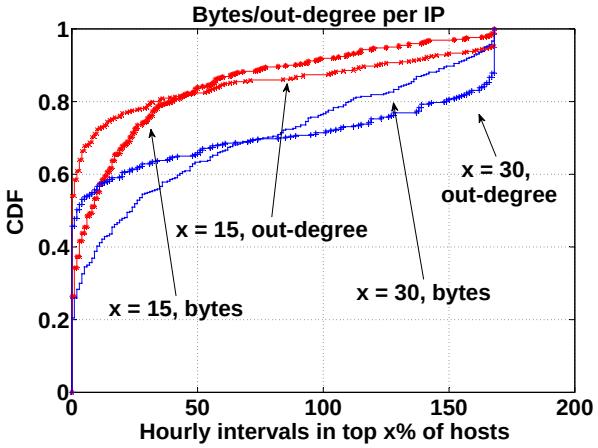

Fig. 8. Prevalence of individual hosts in the most connected set and the heaviest hosts set in terms of bytes for hourly intervals during week 3. Connectivity produces a more stable set over time.

Summarizing the discussion throughout the section, we observe considerable spatial and temporal variability, that is, both across time and across hosts with respect to individual hosts' traffic contributions. While it seems intuitive to categorize hosts as either clients or servers based on the largest traffic contributors, examination of the data suggests this is not a fruitful approach to identify the functional role of enterprise network hosts. On the contrary, connectivity information appears as a more efficient alternative since hosts appear to essentially communicate with a stable set of other internal hosts, with server machines being the most connected ones.

V. Address and Mobility Dynamics

In this section we examine address, host naming and host mobility dynamics within the enterprise. We first address the issue of how useful IPs are to uniquely identify hosts, and then we study host mobility patterns within the enterprise.

A. Name-address characteristics

It is unsurprising that a large enterprise network will provide wireless connectivity for employees' machines, and will usually

allocate addresses via DHCP for all machines, wired and wireless. Thus, a host may be assigned multiple IP addresses over time, and also an IP address may be assigned to multiple hosts. The presence of services that are provided by clusters of machines via a single name further complicates matters. The result is that an IP address does not suffice to uniquely identify a host in general, although for most desktop hosts that connect solely to a wired network it will be a stable identifier. In this section we address the following questions: *What are the characteristics of the name-address mappings in the network? How often should an IP be considered a unique identifier of an enterprise host?*

To answer these questions, we combine examination of router configuration files and DNS packet information. Specifically, we first parse DNS response packets captured in the corpus and we extract the time-varying mapping of names to addresses. Then, using the subnet allocations obtained from the router configuration files, we can map the addresses in each response to their subnet. We assume that a host's *name* tends to change very infrequently and is thus static for the duration of the trace, allowing us to use a name as a stable host identifier.

Since the name to address mapping is not a one-to-one mapping, we examine three types of mappings:

1) *Name-address* mapping, that reveals the number of unique names per IP address.
2) *Address-name* mapping, that shows how many distinct IPs we observe for a unique name.
3) *Subnet-name* mapping, that describes the number of subnets a unique name has been associated with.

Name-address: Fig. 9 displays the characteristics of name-address mappings observed in the corpus. Of the 1,757 unique addresses that were returned as the result of some name resolution, 73% mapped to a unique name, the expected common case. Of the remainder, all but one mapped to 16 names or less, the outlier appearing to be the address of a machine hosting many services in a large datacenter which is thus accessed via a variety of names.

Address-name: Of the 9,274 unique names observed, 63% map to a single address, again the expected common case. We cannot directly observe the purpose or intended use of a host, so explaining the reason why a third of hosts appear to have multiple addresses is difficult. From the subnets in question it appears that hosts with multiple addresses are either laptops with both wired and wireless addresses, or are in fact names that correspond to a service provided by a cluster of hosts (e.g., the web-proxy service provided for hosts within Europe).

Subnets-name: The 63% names that map to a single address obviously map to only a single subnet. Of the other names 30% map to just two subnets, typically one wired and one wireless, which is common behavior for employees using a laptop as their main desktop machine. Finally, the rest of the hosts map to more than two subnets: these are probably laptops moving between sites and they amount to approximately 7% of all names.

The implication of these findings is that proper identification of a host in an enterprise network might be challenging from a network trace alone. Hosts can be accessed via multiple names and single names may map to multiple addresses concurrently (where the name really refers to a service rather than a host). Even when a single name only maps to a single address at any point in time, that address may change either because the name refers to a host which leaves the network for sufficiently long time so that DHCP cannot reallocate the same address, or to a host which moves between subnets requiring a completely different address.

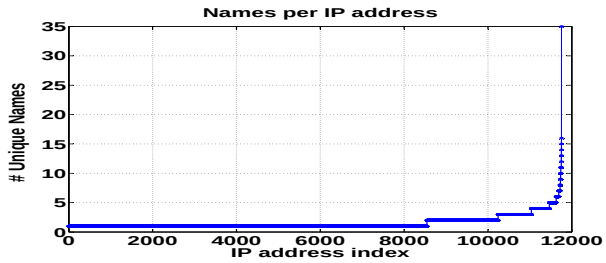

Fig. 9. IP addresses per name. 63% map to a single IP address.

B. Host mobility characteristics

Following from the previous section, it seems that at least a number of hosts move around within the enterprise network. By mapping IP addresses to subnets and then to routers, and thus to cities and countries, we can observe the travel behavior and mobility patterns of particular hosts. We see that, as suggested by the majority of hosts having addresses within a single subnet, most hosts appear to remain tethered in a single location. However, roughly 6% of hosts appear to travel to different cities, and approximately 4% travel to different countries. In this section we ask the question: *how do hosts move around the network geographically?*

Lately, there has been an increased interest in human mobility patterns in the setting of Delay Tolerant Networking (DTN) and opportunistic communications [6], [7]. While the timescales of interest here are not comparable with these studies and the setting is different, our findings in this section provide evidence of similar observations in the context of mobility within an enterprise network.

To extract the geographical location of hosts, we take the previously obtained subnet mappings for their addresses, map them to their home routers and decode the city and country codes embedded in each router's name. Note that we remove from consideration hosts with names that are known to refer to clustered service implementations such as our proxies. Overall, we are left with 712,598 name service responses to examine, involving 9,269 names in 110 cities across 63 countries.

We examine the changes in location (*trips*) visible from these data. Trips are defined as subsequent observations of a host in two different enterprise sites A and B. We can then define the *residence time*[1] at site A as $t_2 - t_1$, to represent the time a host spent in A, where t_1 is the first observation of a host in an origin site A, and t_2, with $t_2 > t_1$, is the first observation at the destination site B. Similarly, if a host follows a travel pattern of $A \rightarrow B \rightarrow A$, we regard as the *return* time to site

[1]Note that residence time is an approximate metric since it also encompasses travel times, disconnections from the enterprise network, etc.

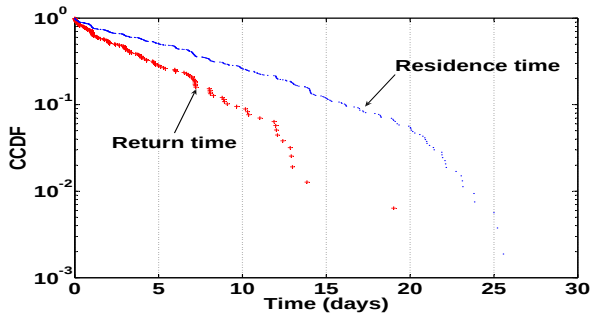

Fig. 10. The CCDF of residence time at a site and return time to a site in lin-log scale. The straight lines point towards the exponential distribution.

A (i.e., how much time a host was away from A) to be the residence time at site B.

We assume that trips with residence time less than 5 minutes are spurious and due to convergence among the many name servers in our network. We also deal with the complication of dual-ported hosts, such as laptops with a wireless and a wired interface (by preferring the wireless address), and many server hosts with two wired interfaces. The end result is 532 unambiguous observed trips, involving 344 unique names.

The distributions of residence time and return time follow the exponential distribution. Fig. 10 shows the CCDFs of residence and return time in days. The CCDFs are plotted in lin-log scale where a straight line is a sign of an exponential distribution as is the case in our data. We observe plateaus at approximately daily intervals as would be expected if trips are due to people visiting different sites. Slightly longer plateaus, indicating more trips, are observed at one, two, and three day boundaries, and at one and two week boundaries. We hypothesize that these plateaus are the result of common durations for business trips. Overall, approximately 38% of all residence times are less then three days, while the mean residence and return times are approximately 5.5 and 3.8 days respectively. Overall, our observations are consistent with [7], where the authors observe an exponential tail for the distribution of human inter-contact times and that human contacts occur in a small number of locations.

VI. Discussion - concluding remarks

Throughout this paper, we have posed and answered a series of questions with regards to characteristics that describe the underlying dynamics of modern enterprise networks. The nature of such typically geographically distributed networks that offer numerous diverse services to several clients worldwide renders them remarkably different with specific idiosyncrasies compared to traditional Internet traffic.

We believe that the implications of our observations are multifaceted. Due to space limitations, we highlight here the more direct ones.

Client vs. server distinction. Whereas intuitively a distinction of client and server machines may make sense for a single application, we observe that it does not for individual hosts within an enterprise network. Thus, inferring the functional role of hosts [8] by simply observing traffic volumes in the network does not appear feasible.

Spatio-temporal variability. The observed variability in the per-host load dictates that any system that attempts to reconstruct network-wide traffic load by sampling must track a very dynamic set of heavy users using a possibly nontrivial set of features. This implies that simple approaches where a small sample set of servers is monitored to inform traffic engineering and dimensioning will most likely fail. To this end, we believe that integration of hosts in the overall enterprise network management appears as an attractive alternative [9].

Locality. A significant fraction of the traffic stays within the enterprise network in agreement with previous work [4]. However, portion of the traffic corresponds to "cross-site" traffic within the enterprise (see also [5]), and while it will be opaque to the underlying providers of the network connectivity, it is still distributed far and wide through the network and around the globe.

Address mappings. As expected, identification of a host in an enterprise network might be challenging from a network trace alone. As mobility increases within enterprise networks and the fraction of mobile hosts grows, we believe that the issue of name-to-address mappings will be further pronounced in the future.

Mobility. We show that recent findings in opportunistic communication settings [6], [7] seem to also apply when describing mobility within an enterprise network. These observations are of general interest in understanding individual human mobility and travel patterns [10].

Overall, we believe that our observations provide valuable insights regarding the primary properties of enterprise networking to the research community, for whom such data is rarely accessible.

References

[1] V. Paxson, "End-to-end routing behavior in the Internet," in *ACM SIGCOMM Computer Communication Review*, vol. 26,4, August 1996, pp. 25–38.

[2] C. Fraleigh, C. Diot, B. Lyles, S. Moon, P. Owezarski, D. Papagiannaki, and F. Tobagi, "Design and deployment of a passive monitoring infrastructure," in *Passive and Active Measurement Workshop*, 2001.

[3] T. Henderson, D. Kotz, and I. Abyzov, "The changing usage of a mature campus-wide wireless network," in *MobiCom '04: Proceedings of the 10th annual international conference on Mobile computing and networking*, NY, USA, 2004, pp. 187–201.

[4] R. Pang, M. Allman, M. Bennett, J. Lee, V. Paxson, and B. Tierney, "A first look at modern enterprise traffic," in *Proceedings of ACM/Usenix Internet Measurement Conference (IMC) 2005*.

[5] R. Mortier, T. Karagiannis, and P. Key, "Address and traffic dynamics in a large enterprise network," MSR-TR-2008-98, Tech. Rep., July 2008.

[6] A. Chaintreau, P. Hui, J. Crowcroft, C. Diot, R. Gass, and J. Scott, "Impact of Human Mobility on the Design of Opportunistic Forwarding Algorithms," in *Infocom*, 2006.

[7] T. Karagiannis, J.-Y. L. Boudec, and M. Vojnović, "Power law and exponential decay of inter contact times between mobile devices," in *MobiCom '07: Proceedings of the 13th annual ACM international conference on Mobile computing and networking*, 2007, pp. 183–194.

[8] T. Karagiannis, K. Papagiannaki, and M. Faloutsos, "Blinc: Multilevel traffic classification in the dark." in *Proceedings of ACM SIGCOMM 2005*, Aug. 2005.

[9] T. Karagiannis, R. Mortier, and A. Rowstron, "Network exception handlers: Host-network control in enterprise networks." in *Proceedings of ACM SIGCOMM 2008*, Aug. 2008.

[10] M. C. Gonzalez, C. A. Hidalgo, and A.-L. Barabasi., "Understanding individual human mobility patterns," in *Nature*, vol. 453, Jun. 2008, pp. 779–782.