N-TAP: A Platform of Large-Scale Distributed Measurement for Overlay Network Applications

Kenji Masui        Youki Kadobayashi
Nara Institute of Science and Technology
8916-5 Takayama, Ikoma, Nara 630-0192, Japan
{kenji-ma, youki-k}@is.naist.jp

Abstract

To sustain a large-scale overlay network, knowledge about network characteristics is indispensable. However, the collection of such information is often a burden on the developers: why should they struggle to obtain the information even though it is not their intended purpose? As one solution to this problem, this paper presents N-TAP, a distributed measurement infrastructure. N-TAP is an independent service for overlay network applications that provides network characteristics. N-TAP itself forms the measurement overlay network on which measurement nodes can cooperate in measurement activity. Consequently, overlay network applications can take full advantage of novel measurement methodologies. Through a discussion on the problems of distributed measurement, in this paper we explore the capabilities and the future direction of N-TAP.

1. Introduction

Although one merit of the Internet is that end nodes can simply communicate without knowing the status of intermediate nodes and networks, recent studies on overlay networks have revealed that such knowledge is still capable of empowering applications on end nodes. In this paper, network characteristics denotes the information that characterizes a network and its components such as nodes and links, with round-trip time (RTT) and IP topology being instances of network characteristics. For example, in the case of application layer multicast (ALM) [1, 2, 3], topology information enables an application to generate a distribution tree with the following purposes: preserving alternative routes in case of node failure, redistributing contents among proximal nodes, and so on. As peer-to-peer overlay applications become more popular, interest in discovering and exploiting large-scale network characteristics continues to grow.

Meanwhile, the measurement methodology for collecting network characteristics is shifting toward a distributed manner because the measurement capability of a solo node is limited and the centralized aggregation of the results from isolated monitoring nodes is not often to scale, especially in the case of large-scale measurement. In distributed measurement, monitoring nodes utilize the data that other nodes collected, or communicate with others for cooperative measurement, then perform analysis and estimation to calculate network characteristics. There have been some past works in this area, including Vivaldi [4] and Doubletree [5]. In overlay networks, measurement is often performed across multiple administrative domains distributed over the Internet without preliminary knowledge about their networks; therefore, their assumptions of the measurement environment will suit this situation.

As a provider of network characteristics for large-scale overlay network applications, in this paper we propose a novel distributed measurement platform, named N-TAP. The fundamental feature of N-TAP is the collection of network characteristics on the overall Internet; however, N-TAP contains some unique features and appearances. N-TAP is equipped with shared storage for collected network characteristics. This storage consists of monitoring nodes for distributing the storage cost. Additionally, since the monitoring nodes in N-TAP can communicate with each other, they can cooperate in collecting network characteristics. For applications, N-TAP appears to be an independent service that provides requested network characteristics to them.

Our motivation for proposing N-TAP is to enable application developers to handle network characteristics easily and to reduce the measurement cost on the Internet by abstracting measurement procedures into one independent service. Current network applications perform their measurement by themselves, and these operations are indispensable especially for large-scale overlay network applications in order to effectively maintain their networks. However, application developers need to write lines of code for such measurement, and this is often a significant burden on them. Moreover, this situation is a waste of resources (CPU, network bandwidth, etc.) because some types of required network characteristics are common among applications and portions of collected data may be reusable with other applications. It is also reported [6] that the measurement traffic, which is generated from experiments on overlay networks, consumes considerable amounts of...
bandwidth in PlanetLab [7]. If applications can obtain network characteristics from an integrated service, and if collected data can be shared and reused within the service, it will be possible to reduce the impact of these problems. Additionally, there have already been several works on distributed measurement; therefore, the provision of the infrastructure where these methodologies can be easily deployed allows applications to take full advantage of them.

Our contributions in this paper continue as follows: Section 2 describes the requirements of N-TAP. In Section 3, we propose the design and the architecture of N-TAP to meet these requirements. We also refer to our current implementation in the same section. Section 4 discusses the future direction of N-TAP. We mention some past works on distributed measurement in Section 5, and conclude this paper in Section 6.

2. Requirements

First, we describe the necessary characteristics for N-TAP in order to become a provider of network characteristics for overlay network applications. The main requirements are cooperation, independence, and decision making.

Cooperation Cooperation among monitoring nodes provides the opportunity to obtain the data that cannot be measured with a solo node, and it is considered an effective measurement methodology for large-scale overlay networks. N-TAP must prepare both a mechanism for communication among monitoring nodes and accessibility to the collected network characteristics on the overall system. This is because distributed measurement often involves communication among other nodes and the utilization of collected data, as stated in Section 1.

Independence Independence means that any kind of application can utilize this service through a network in order to obtain network characteristics. N-TAP must include an interface for exchanging messages with other applications to accept their requests and provide network characteristics to them. Moreover, the formats of messages and network characteristics must preliminarily be defined to ensure the portability of this system.

Decision making N-TAP has to interpret requests from applications carefully and make a decision about what action to take to collect network characteristics. For instance, topology data collected ten years ago does not represent the current topology and thus is worthless for an application that intends to reconstruct an overlay multicast tree. However, the RTT data that has been collected continuously during a week and shows its periodicity would be useful for RTT-based proximity-aware node selection. As a consideration to the purpose of each application, we also need to take into account other factors such as measurement cost and response time.

3. Architecture and implementation

The system of N-TAP is composed of N-TAP agents, programs that run on monitoring nodes. Except for N-TAP agents, N-TAP consists of no extra facility or program. In other words, it is just the N-TAP agents on monitoring nodes that perform all the system’s procedures. We assume, in principle, that all the overlay application nodes, which use N-TAP, run N-TAP agents locally. However, this assumption is not necessary in cases that the required network characteristics do not require measurement from the local node.

As Fig. 1 shows, N-TAP agent can broadly be divided into four components: a network characteristics database, a network characteristics provider, a network characteristics collector, and an N-TAP network manager. The network characteristics database is a repository of collected network characteristics. The network characteristics provider is an interface between an N-TAP agent and other applications. This component accepts the requests from the applications, and then based on the requests, it decides how to collect the requested network characteristics. We describe this decision-making process later. Applications can determine what kind of network characteristics they want and some conditions for them through the interface. The network characteristics collector performs measurement in order to collect the network characteristics requested by the network characteristics provider. The N-TAP network manager is responsible for forming N-TAP network, the measurement overlay network among N-TAP agents. It also searches other N-TAP agents that other components (collector or provider) required for cooperation, and sets up a place for their rendezvous.

Here we present the decision-making process of network characteristics provider. To reduce measurement costs and improve responsiveness, the process adopts the basis of local-first. Initially, the network characteristics provider searches for network characteristics that satisfy the request from an application in the local database. If the requested data are not found in the database, it checks whether the agent can collect the data or not. If the agent can, it requests the network characteristics collector to collect the data; otherwise, it searches the data in the shared database. If the data does not exist in the shared database, by asking to the N-TAP network manager, the network characteristics provider tries to find other agents that can collect the data. If one or more agents is found, the network characteristics provider makes a request to the network characteristics collector to collect the data. If not, it replies to the application that the requested data are unavailable. The provider that accepted the request from another agent performs the same decision-making process, but it does not forward the request to other agents and immediately replies in the case of unavailability. Through the simple steps described above, the network characteristics provider decides what to do to collect the requested network characteristics.
In our actual implementation of N-TAP, we utilize the Chord [8] technique, which is one implementation of a distributed hash table (DHT), to form an N-TAP network. N-TAP currently uses the 160-bit keys obtained from the SHA-1 cryptographic hash function. Each agent has its own agent ID that locates the agent in the Chord ring, and the length of the ID space is same as the keys.

N-TAP constructs its shared database for collected network characteristics by using the nature of Chord, for the reason that the storage cost, which will be enormous in the case of large-scale measurement, can be distributed among monitoring nodes. After an agent has collected the network characteristics, the item is stored in the node that collected the data as well as the (other) nodes that are responsible for the hash values obtained from the respective indexes in the collected data, as the keys. The indexes of each type of network characteristics are preliminarily defined in order to make the collected data accessible. For example, if node A collects the RTT between node A and node B, the data are stored in node A as well as the nodes that are responsible for \( \text{hash}(\text{IP}(A)) \) and \( \text{hash}(\text{IP}(B)) \) in the Chord ring, where \( \text{IP}(N) \) denotes the IP address of node \( N \) and \( \text{hash}(x) \) is the hash value calculated from a key \( x \). As above, N-TAP forms the shared database where N-TAP agents can deposit and retrieve the collected data with some indexes. Additionally, each agent deposits the information about itself into the shared database so that other agents can find the agent for cooperation. The agents use their IP addresses, netmasks, and fully qualified domain names (FQDNs) as the indexes of the deposited information, thus the N-TAP network manager can search other agents by using these indexes.

Regarding the protocol that applications use for making requests to N-TAP agents, we adopt XML-RPC [9] because of its widespread deployment and description capability. To request network characteristics, applications call the agent’s methods for collecting the target data with specifying the type of network characteristics and certain conditions for the data. The collected data are stored with additional information such as time stamps, the ID of the agent that collected the data, its collection method, and so on. They are the criteria for judging whether the data can show the actual state of network entities, and the judgment depends on the conditions offered by an application.

Our current implementation of N-TAP contains the fundamental features presented above for meeting the requirements stated in Section 2. N-TAP agents now work on FreeBSD, Mac OS X, and Linux on PlanetLab. As the collection methods, current N-TAP can \text{ping} to measure the RTT between a monitoring node and another node, and can also \text{traceroute} to obtain the IP topology. Furthermore, one N-TAP agent can request other agents to \text{ping} or \text{traceroute}, and it is a primitive method of cooperative measurement from the standpoint that only one node cannot measure the RTT or the topology whose start points are not the measurement node. The implementation of existing methodologies for distributed measurement on N-TAP is still ongoing. Due to space limitations, we have skipped the details of the protocols, the format of collected data, etc. here. However, in future we also plan to release N-TAP software and documentation on N-TAP.

4. Discussion

From the aspect of our goal, the most important proposition is whether N-TAP really reduces a developers’ burden and promotes the utilization of network characteristics. As stated in Section 1, measurement is often a burden on overlay application developers. Our answer to reduce such burden is to prepare an independent measurement service that provides network characteristics to them. Compared to writing own codes for measurement, N-TAP will surely reduce the burden; they just need to specify what kind of network characteristics to collect and request N-TAP to collect them even if the measurement procedure that N-TAP performs is very complicated. However, such
measurement framework would be worthless if N-TAP does not provide what they want, e.g., collectable network characteristics, measurement methodologies and their parameters. We need to continue the survey of such requirements and verify the proposition by actually implementing some applications that utilize N-TAP.

On the other hand, one of the challenges of distributed measurement we recognize is to explore the a reasonable point of trade-off as an infrastructure for overlay network applications. There are some indexes to characterize measurement methodology, but each index affects another as a trade-off. If agents attempt to extend the target where they collect network characteristics, due to the limitation of their measurement capacity they risk the possibility of losing timeliness. Timeliness is one of the most important factors because its increase will improve the reusability of collected data, which will in turn reduce the measurement cost on both each agent and the overall Internet. Meanwhile, excess measurement activity often changes the primary characteristics of network entities. Therefore, we should study on a strategy of the collection.

Another challenge of N-TAP is extracting significant data from a massive amount of the collected data in a distributed manner, which is also a unique feature of N-TAP. The entire N-TAP system can be compared to one huge shared repository for network characteristics. Each agent analyzes the data in the repository and makes the decision about what action to take in order to provide network characteristics to an application. Comparing with a centralized type of system, distributed analysis has the advantage of reduced analysis cost because each agent does not have to process all requests in N-TAP, only ones from a part of the applications in question. Meanwhile, if retrieval of the data that the agent uses for the analysis takes a long time, this makes turnaround time to the application longer. We need to verify whether N-TAP can process the request within an acceptable timeframe under such a situation.

5. Related work

Some has already been conducted on the methodology of distributed measurement. Vivaldi [4] is a decentralized coordinate system based on a physical mass-spring system, and can estimate the RTT between two nodes with few measurements. GNP [10], NPS [11], PIC [12], and Light-house [13] are also included as coordinate systems. Meanwhile, Doubletree [5] is an algorithm that reduces the cost of traceroute by exploiting the common portion of IP topology. By deploying these methodologies on N-TAP, applications can take advantage of them and researchers will have opportunities to perform their experiments on an actual network environment.

Projects for collecting network characteristics such as CAIDA [14] and DIMES [15] also exist, and they are basically working on scientific and statistical analysis. Though their objectives are slightly different from the one of N-TAP, the fundamentals of their studies, such as the style of infrastructure and the analysis methodologies, will be informative for N-TAP, too.

6. Conclusion

In this paper, we presented the architecture and the implementation of N-TAP, a platform for large-scale distributed measurement with the consideration for overlay network applications. We believe that such an infrastructure is indispensable for sustaining application networks and promoting their continued growth even though we need more proof from both theoretical and practical aspects. To guard against the excessive traffic derived from the activities of overlay network applications, our system will contribute toward solving such problems. Furthermore, we hope that our system aids application developers in handling network characteristics, and accelerates the studies on the methodology of distributed measurement. Future work will involve continuing research on the validity of distributed measurement for overlay network applications on actual network environments.

References