The ACTION Project: Application coordinating with Transport, IP and Optical Networks

Naoaki Yamanaka, Satoru Okamoto, Yukihiro Imakiire, Mai Arase
Keio University, Department of Information and Computer Science, Japan

Eiji Oki
The University of Electro-Communications, Japan

Malathi Veeraraghavan
University of Virginia, USA

ABSTRACT
Internet traffic volume continues to increase at a high rate every year. To handle unexpected traffic surges, link failures and growth in traffic volume, network link capacities are typically over-provisioned. To reduce the extent of over-provisioning, we proposed an application and network coordinated multi-layer architecture called ACTION. Application QoE (Quality of Experience) are translated into network QoS requirements, and the multi-layer architecture is leveraged to meet these QoS requirements. To determine the quality experienced by application flows, Hadoop logs are communicated to the network controller, and edge devices are programmed to detect elephant flows automatically. Virtual networks that offer applications flows their required QoS are realized by configuring physical network switches to achieve the required QoT (Quality of Transport). ACTION is a multi-QoE, application-centric, energy-efficient network architecture that leverages flexible elastic optical network technologies. This project is supported by both NSF, USA, and NICT, Japan, under the JUNO (joint collaboration between Japan and US) program. This paper provides a brief overview of the project.

Keywords: Quality of Experience, Quality of Service, Quality of Transport, flexible elastic network

1. INTRODUCTION
As increasing amounts of content are added to servers on the Internet, user access to this content results in increasing levels of network traffic [1]. Networks need to handle a variety of flows: high-throughput flows that generate traffic spurts close to link capacity, long-duration audio/video streaming flows, and short messages. QoS requirements and traffic characteristics vary for each application type. Mirroring datasets between physically separate data centers requires high bandwidth but such transfers occur just a few times per day. Streamed video flows require relatively low bit rates but could be long lasting. Business-related data flows require high reliability, while gaming applications are major consumers of the Internet best-effort service. Today’s Internet supports this varied set of applications by simply over-provisioning links, e.g., 30% - 40% link utilization levels are quite common [2]. However, low network utilization implies low energy efficiency. To increase energy efficiency, future network designs should offer solutions for supporting multiple classes of service without over-provisioning.

The state-of-the-art elastic optical network technologies enable dynamic and flexible bandwidth assignment for each wavelength [3-5]. These technologies could be used to provide multi-QoS pipes between edge nodes to support various classes of service. As optical network switches consume lower levels of energy when compared to electronic packet switches, our multi-layer design should achieve better energy efficiency than IP networks.

We proposed a new application-centric network architecture based on elastic optical network technology [3]. In the proposed ACTION architecture, edge nodes called ACTION edges monitor application QoE (Quality of Experience) and assign flows to various multi-QoS (Quality of Service) virtual optical networks. Two coordination methods between application and network are described. The first method uses Hadoop resource manager to coordinate both network and server resources. The second method uses a dual LRU (Least Recently Used) technique to identify elephant flows and mice flows.

In the proposed ACTION network, QoS monitoring will be used to identify bottleneck links, and the bandwidth of these links will be changed automatically leveraging elastic optical network technologies. In addition, techniques for automatically aggregating traffic onto small number of links will be designed so that one or more links can be powered off using the MiDORi methods [10]. Therefore, high QoE measures can be achieved while simultaneously ensuring energy efficiency. This is the main objective of ACTION.

2. ACTION Architecture
The Internet has proven to be a highly scalable network infrastructure, and IP has successfully supported different services. TCP has played a critical role in enabling a fair sharing of link bandwidth among multiple connections. However, certain links can become bottlenecks on end-to-end paths, even while other links are under-utilized. There is potential for energy savings on these under-utilized links.
As a solution, we proposed ACTION, a new application-centric network as represented in Fig. 1 [3]. ACTION enables the creation and use of multiple slices in an optical network. Each slice can be configured to provide a different QoS. Elastic optical networks can support pipes of different capacities between electrical switching nodes.

Fig. 1 ACTION architecture overview with multi-QoS virtual optical network slices using elastic optical networks

ACTION edge nodes monitor the QoE of certain application flows. If QoE degradation is detected, an application flow can be mapped to a slice with better QoS automatically. On the other hand, if QoE is well above the required level, the flow could be directed to a lower-QoS slice.

The multi-QoS optical slices are provisioned using a self-sizing elastic network technique illustrated in Fig. 2. The elastic network can adjust link capacities easily. All the virtual link utilization levels are monitored, and bandwidth levels are adjusted to hit the desired utilization targets, e.g., 70%. With this technique, link bandwidth can be adjusted automatically to match traffic levels, optionally using transmission across parallel routes if needed [6], and hence this technique is referred to as “self-sizing.”

3. ACTION Edge

The ACTION Edge node is typically an IP router, Ethernet switch or a TCP/IP endpoint (host). The ACTION Edge can be programmed to monitor the QoE of certain applications, and to initiate requests to modify the transport network bandwidth automatically if QoE is below the desired level. In the ACTION network, a complete mesh of virtual-optical pipes is provisioned between ACTION Edges. Users and applications can thus enjoy high QoE as the network automatically provisions the required bandwidth. To achieve this automatic QoE, and required QoS levels, we propose two methods. In the first method, an application explicitly sends a signal to a network controller. In the second method, flows are classified as elephant flows and mice flows. Each flow is directed to a different QoS network slice. Details are described in section 4.

4. Application and network coordination method

Experiments for application triggered network slice provisioning are described. Method-1 is a data center network based experiment. Fig. 2(a) shows a data center network, which consists of multiple of servers controlled by a Hadoop manager (modified YARN scheduler). The Hadoop manager computes traffic characteristics of applications. The manager distinguishes computing/communications-heavy vs. computing/communications-light tasks, and assign server resources as well as network resources. The manager allocates multi-layer network resources according to traffic values. Heavy tasks use the optical network, while light tasks use the electrical Layer-2 (L2) packet network. Therefore, the Hadoop manager considers both computing resources and network resources.
For determining the effectiveness of the proposed idea, we used two kinds of tasks, heavy task and light task. For the heavy task, the TeraSort benchmark program was used, and for the light task, the wordcount benchmark program was used. All heavy tasks were directed to use the optical network and light tasks were directed to use the electrical network automatically as also shown in Fig. 2(b). Peak traffic value of optical network is 108 bps and that of electrical network is 106 ~ 107 bps. We can see the effectiveness of application coordinated networking. We need further studies for better assignment of server resources and network resources to tasks.

In method-2, traffic classification by flow-monitoring was developed. Figure 3 shows automatic traffic classification with a dual LRU (Least Recently Used) scheme modified from the advanced cache method [14] [15]. Flow information is stored in LRU-1 and moved to the bottom as new flows are added. As packets from a cached flow arrive, the corresponding flow number is moved to the top of the LRU-1 cache. Flows at the bottom of the LRU-1 cache will be shifted out. Stored flow number has a counter. If a cached flow moves to the top of the LRU-1 cache, its counter value is incremented. If the counter value of flow K is larger than a threshold (Th), flow K is moved to LRU-2 cache. Flows in LRU-2 cache may be elephant flows. Operation of LRU-2 cache is same as that of LRU-1 cache. However, only elephant flows are stored in the LRU-2 cache. Mice flows are not moved from LRU-1 to LRU-2. Design parameters include LRU shift register lengths L1, L2 and threshold value Th.

Basic evaluation results of proposed traffic classification method are shown in Fig. 4 using Chicago-San Jose CAIDA data. CAIDA offers unsampled anonymized packet traces collected at two OC192 monitors [16]. These monitors are deployed on WAN links, and hence this traffic may not be representative of data center traffic. For this first step in our research, we used CAIDA traces just to illustrate the concept. Specifically the trace obtained from the Chicago OC192 monitor The source and destination IP addresses define a flow. We defined an elephant flow as one that sends more than 2.5 Mbyte of data. The threshold value was set to 8 times...
flow during 50ms. Two metrics, FNR (false negative rate) and FPR (false positive rate) for elephant-flow identification are shown in Fig. 4. These elephant flows are directed to the optical network. However, we need real traffic traces from data center networks to properly select L1, L2 and Th.

A false negative occurs if an elephant flow is never moved to LRU-2. This can happen if packets from the elephant flow do not arrive at a high enough rate, and hence the flow keeps getting removed from LRU-1. A false positive occurs if after observing 8 flows /flow in 50ms in LRU-1, the flow is moved to LRU-2, but the flow size never reaches 2.5 MB.

The size of LRU-1 was about 12K packet and the size of LRU-2 was infinite for this evaluation. The specific CAIDA data trace analyzed in this work was collected on March 20th, 2014 from 22:01:00 ~ 22:01:59 of Chicago to Sanjose CAIDA data.

5. Conclusion

This paper described ACTION, a new application-centric network architecture. ACTION uses a dynamic flexible optical sliced network with multi-QoS capability. Two application and network coordination methods are described. In one method, the application triggers flow direction to an appropriate QoS slice, while in a second method, an ACTION edge node monitors the QoE of certain application flows, and maps these flows to an appropriate QoS optical slice. The first method is tested using a Hadoop manager that determines the computation and communication requirements of tasks for scheduling heavy tasks to optical networks, and light tasks to electrical packet networks. The second method is tested using a dual LRU technique for elephant-flow identification. Both methods are useful in the ACTION application and network coordination architecture.

Acknowledgments

This work was supported in part by the National Institute of Information and Communications Technology (NICT), Japan and the NSF grant (CNS-1405171), USA.

REFERENCES