Line-Rate Map-Reduce in Data Plane

Paper ##, 12 pages, November 17, 2016 9:31 PM

ABSTRACT

p4mr is a framework to allow applications to offload computation to the data plane. It enables Map-Reduce style data flow computation in the datacenter fabric. p4mr enables information aggregation in the data plane, which reduces the traffic and lessons the load in servers.

1. INTRODUCTION

We cite dctcp [**?**].

Fixed function computation is plenty in existing data plane, e.g. counting, routing, forwards, and state machines. They are done at line rate of the networking device. These fixed data plane functions are traditionally distinct from user applications. The emergence of programmable data plane blurs this distinction, and unlocks the possibility of doing user compuation tasks at line rate in the data plane.

Prior works in programmable data plane focus on network functions. A function, like Paxos algorithm, that requires programming multiple switches, are done by hand by P4 programmers. User applications usually takes more than one switch, and to program and debug many switches is inherently difficult. Furthermore, scalabiity and extensibility of multi-target functions is a concern. After developing a program in P4, if the user want to add more switches to carry out this function, the user will have to do it by hand. In general, programming in P4 is analogous to programming in assembly for a large network of targets.

We therefore aim to develope a programming framework for multi-target functions. In particular, we build a Map-Reduce framework for P4, p4mr.

p4mr take user program as input and compile it into mulitple P4 programs for a network of P4 targets. Based on the prior knowledge of topology and hardware capability, p4mr customize P4 codes to the corresponding hardware. Specifically, p4mr adds appropriate routing between mappers and reducers, and places mapper/reducer code to appropriate targets.

Online data parallel (OLDP) applications are important catagory of data center application, such as web search, machine learning, and stream processing. This type of applications are built with a tree based algorithm and adopt a divide-and-conquer strategy. The construction of this type of applications results in many-to-one communication pattern in each layer, which may lead to congestion at the edge switch.

2. BACKGROUND

2.1 Data plane compuation: fixed functions

There are many compuation functions that are already in fixed function switches.

2.2 Programmable data plane and P4

With P4, data plane becomes more and more programmable.

2.3 Programming models for parallel computing

Many programming models for parallel computing on distributed devices. We consider circuit, actor, dataflow models are suitable for a parallel programming framework for P4.

3. SYSTEM OVERVIEW

4. IMPLEMENTATION

5. EVALUATION

6. CONCLUSION

7. REFERENCES

[1] M. Alizadeh, A. Greenberg, D. A. Maltz, J. Padhye, P. Patel, B. Prabhakar, S. Sengupta, and M. Sridharan, "Data center tcp (dctcp)," *ACM SIGCOMM Computer Communication Review*, vol. 41, no. 4, pp. 63–74, 2011.