Volume 3, Issue 3, March 2013 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

Congestion Avoidance in MPS with Flow Cut

Tejaswini N P*
Department of Computer Science & Engineering, East West Institute of Technology, VTU, India

Prasanna Kumar M
Department of Computer Science & Engineering, East West Institute of Technology, VTU, India

Abstract*: Traffic balancing is a useful technique that keep network free from congestion when link is overloaded or link failures occur in MPS. The performance of balancing the system is determined by granularity of packet splitting, whether it is based on flow-based or packet based solution. Packet based solutions requires packet reordering in the same TCP flow while flow-based solutions will not requires reordering techniques. In this paper we develop a novel flow-based scheme, namely Flow-Cut (FC) which splits each flow in to cuts at every intra-flow interval larger than the flow cut threshold of 1-4ms. Thus balances the load on finer granularity. We show that FC achieves impressive traffic balancing performance with little network cost while reducing the packet out-of-order chances to negligible level (below 10^{-6}).

Keywords- Traffic balancing, Flow-cut, switching system

I. INTRODUCTION

Multipath Switching System (MPS) plays a vital role in fabricating state-of- the-art high performance core routers. One of the major issues in designing MPS is the load-balancing problem defined as how to distribute incoming traffic across its multiple switching paths to meet the three objectives simultaneously:

1. *Uniform load-sharing*: The traffic destined for each output should be dispatched to all the switching paths uniformly.
2. *Intra-flow packet ordering*: The packets in the same flow should depart MPS as their arrival orders.
3. *Low complexity*: The load-balancing and the additional resequencing mechanisms should work fast enough to match line rate and should introduce limited hardware complexity.

In packet-based solutions, the traffic is dispatched packet by packet to optimally balance the load. However, in this packets in the same flow may be forwarded in the separate paths and experience various delays, thus violating the intra-flow packet ordering requirement. Although timestamp or sequence based re-sequencers can be added to restore packet orders, they are often shown to be costly and not scalable. By timestamp based re-sequencer [1], each packet is slowed down statically (or adaptively) by the system delay upper bound, which will impose a huge delay penalty. On the other hand, the sequence based re-sequencer [2] will need to maintain at least N re-sequencers at each output, leading to $O(N^2)$ complexity. To avoid the packet out-of-order, another choice is to use flow-based traffic-balancing algorithms [3]. Here it dispatches packets in the same flow to a fixed switching path by hashing its 5-tuple to path ID.

However, hashing solution will lead to severe load-imbalance. In this paper, we present a new scheme, namely *Flow-Cut* (FC) that perfectly achieves the three objectives defined above. Here the intra-flow packet intervals are often; say in 40-50 percentages, larger than the delay upper bound at MPS which is calculated statistically. As such, if we cut off each flow at every packet interval larger than a flow-cut threshold equaling to this bound and balance the load on the generated flow-cuts, the three objectives are met triply:

1. The traffic-balancing uniformity of FC is only moderately degraded from the optimal traffic-balancing.
2. The intra-flow packet order is kept intact as their arrivals. Exceptions only happen in a negligible level (10^{-6}).
3. The flow-cut table size to implement FC only requires 1.8MB under 40Gbps line rate, which can be placed on-chip to provide an ultra-fast access speed.

Flow-Cut

Definition: A flow-cut (FC) is a sequence of packets in a flow, where every intra-flow interval between two consecutive packets is smaller than or equal to a flow-cut threshold. Flow-cut exhibit the properties of small average packet size, light-tailed size distribution, and fewer active flow-cuts. Hence, the optimal load balance is achieved by FC.

A. *Objective of the study*

- This novel based FC scheme address the problem of congestion and also packet loss, while other solutions like the re-sequencer require additional loss detection Mechanisms.
- FC scheme can achieve optimal performance at the cost of little hardware complexity.

B. *Scope*

In our paper, FC scheme achieves comparative load-balancing performance to the optimal one by setting flow-cut threshold to 1-4ms.

II SYSTEM ANALYSIS

A. *Existing System*

In existing system packet-based solutions are used. Timestamp or sequence based re-sequencers can be added to restore packet orders. They are often shown to be costly and not scalable and it suffers from huge delay and hardware complexity. Without re-sequencer packet out-of-ordering will occur, which leads to improper traffic balance. Thus network congestion and packet loss occurred.

B. *Proposed System*

We proposed a new scheme namely Flow-Cut, based on the fact that the intra-flow packet interval is often larger than the flow-cut threshold. Due to the three positive properties of flow cut namely small average packet size, light-tailed size distribution, and fewer active flow-cuts our scheme achieves good load-balancing uniformity from congestion with little hardware overhead and timing complexity.

C. *System Specification*

Hardware Requirements:

System	: Pentium IV 2.4 GHz.
Hard Disk	: 40 GB.
Floppy Drive	: 1.44 Mb.
Ram	: 512 Mb.

Software Requirements:

Operating system	: Windows XP.
Technology	: Java(jdk 1.6),Swing, Networking
Data Base	: Access/My Sql

D. *System Architecture*

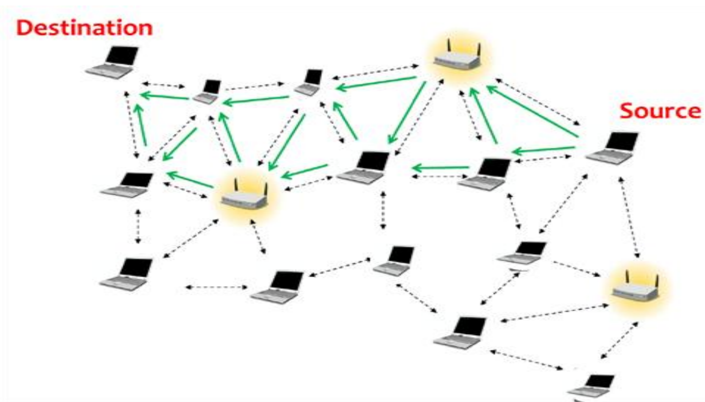

Fig.1 System Architecture

III. MODULE DESCRIPTION

A. *Multistage Multiplane Clos switches (M^2Clos)*

We consider the Multistage Multiplane Clos-network based switch by Chao et al[4] [5]. It is constructed of five stages of switch modules with top-level architecture similar to external input/output ports. The first and last stages Clos are composed of input de-multiplexers and output multiplexers, respectively, having similar internal structures as those in PPS. Stages 2-4 of M2Clos are constructed by parallel switching planes; however, each plane is no longer formed by a basic switch, but by a three-stage Clos Network to support large port count. Inside each Clos Network, the first stage is composed by k identical Input Modules (IM). Each IM is a packet switch, with each output link connected to a Central Module (CM). Thus, there are a total of m identical in second stage of the Close networks.

B. *Multipath Switching System*

In Buffer Management module, all packets are virtually queued at the output according to the flow group and the priority class in a hierarchical manner. The output scheduler fetches packets to the output line using information provided by BM. Packets in the same flow will be virtually buffered in the same queue and scheduled in discipline. Hence, intra-flow packet departure orders hold as their arriving orders at the multiplexer.

C. *Load balancing scheme*

The main idea of load balancing is to map from the part of the heavily loaded paths to some slightly loaded path to avoid congestion in the network[6]. Flow-based load balancing scheme describes FC scheme in two steps:1) an introduction of the mechanism to detect and maintain flow-cut context, and 2)an algorithm to map new flow cut into individual switching path.

First, implement hash table to maintain flow-cut information. Every table entry is expected to record the context of one active flow cut and is composed of two fields apart from the table index: Latest Arrival Timestamp (T_L), which holds the exact time of the latest packet arrival of this flow cut; and Flow-cut Destination Path (P_D), indicating the switching path taken by all the packets in this flow cut. Second, to load balance new flow cuts, a simple round robin algorithm called N-FC is applied, where N is the number of output ports in MPS. Interflow packet order is natively preserved by setting slicing threshold to the delay upper bound at MPS. Since 1)Any two packets in the same flow cut cannot be disordered as they are dispatched to the same switching path where processing is guaranteed and 2) any two packets in the same flow but different flow cuts will be in order at departure, as the earlier packet will have depart from before the latter packet arrives. Due to the fewer number of active flow cuts, the only additional overhead in the hash table, can be kept rather small, and placed on-chip to provide ultrafast access speed. This table size depends only on system line rate and will stay unchanged even if scales to more than thousand external ports, thus guarantees system scalability. The time complexity of N-FC in dispatching packet is shown to be O (1).

Advantages of N-FC are that

1. It is immune to packet loss, while other solutions like the VIQ re-sequencer require additional loss detection mechanisms.
2. It maintains a hash table to record active flow-cut context; a redirection mechanism can be added to provide robustness to system failure.

D. *Backlog generation*

Backlog is created based on the traffic load in the specific node. Backlog provides information congestion occur. If we apply flow-cut technique means, then it will take alternative path to send data to the destination.

E. *Delay bound measurement*

Here we calculate the packet delay bound at first stage. We then study delay at second-stage switches. Define native packet delay at stage m of an be delay experienced at stage m on the condition that all the preceding stages immediately send all arrival packets out without delay.

IV. IMPLEMENTATION ISSUE

The major implementation cost introduced by FC is that of hash table maintaining an active flow-cut context. The size of this hash table is a trade-off 1) a large hash table reduces hash collision probability. Here, hash collision is defined as the situation when more than one flow cuts are hashed into the same entry, which degrades load-balancing uniformity; while 2) a small hash table is desirable as the table is accessed in each load-balancing operation and on-chip SRAM with relatively small size allows higher access speed. As each FC load-balancing operation only needs one query at the hash table, O(1) timing complexity is also achieved.

V. CONCLUSION

We propose a novel load-balancing scheme, namely, Flow-Cut, based on the fact that the intra-flow packet interval is often, say in 40-50 percent, larger than the flow-cut threshold. Due to three positive properties of flow cut, our scheme achieves good load-balancing uniformity with little hardware overhead and O(1) timing complexity. By calculating delay bounds at M^2Clos, we show that when the slicing threshold is set to the smallest admissible value at 1-4ms, the FC scheme can achieve optimal performance while keeping the intra-flow packet out-of-order probability negligible (below 10^{-6}), given an internal speedup up to two.

VI. FUTURE WORK

Here we can enhance FC scheme with QoS condition. But QoS provisioning is also critical in switch designs, one of our future works will be studying FC performance under QoS conditions.

References

[1] J. S. Turner, *"Resequencing Cells in an ATM Switch," Tech. Rep.*, WUCS-91-21, Feb. 1991.

[2] D. A. Khotimsky and S. Krishnan, *"Evaluation of Open-loop Sequence Control Schemes for Multi-path Switches,"* in *Proc. IEEE ICC*, pp. 2116-2120, 2002.

[3] L. Shi, W. Li, B. Liu, and X. Wang, *"Flow Mapping in the Load Balancing Parallel Packet Switches,"* in *Proc. IEEE HPSR*, pp. 254-258, 2005.

[4] Cisco CRS-1, *http://www.cisco.com/go/crs/*, 2011

[5] vitesse Intelligent Switch Fabrics, *http://www.vitesse.com*, 2011

[6] J.Bennett, C. Patridge, and N. Shectman, *"packet reordering is not pathological network behaviour,"* IEEE/ACM Trans. Networking, vol.7,no 6, pp.789-798, dec.1999.

