

リコンフィギャラブルハードウェアを利用した 高速トラフィックモニタ

安田 豊[†]高島 研也[‡][†] 京都産業大学 理学部[‡] ユーテン・ネットワークス株式会社email: [†] yasuda@cc.kyoto-su.ac.jp [‡] kenya@u10networks.com

現在のインターネットのトラフィックがどのようなものかを知ることはネットワーク機器やアプリケーションを設計するにあたって極めて重要である。トラフィックモニタリングはそのための最も基礎的な技術であるが、Gigabit Ethernetの普及に代表される近年の著しいネットワーク速度の向上にモニター機器の能力が追いつかないのが現状である。本研究ではリコンフィギャラブルハードウェアを用いたモニタリングシステムを開発した。またインターネット日食中継イベントにおいて実験したのであわせて報告する。

Reconfigurable hardware based high speed traffic monitor

Yutaka Yasuda[†]Kenya Takashima[‡][†] Kyoto Sangyo University, Faculty of Science[‡] u10 Networks, Inc.email: [†] yasuda@cc.kyoto-su.ac.jp [‡] kenya@u10networks.com

It is important to keep tracking the real Internet traffic to make a reasonable design and development of network devices and applications. Traffic monitoring is one of the most primitive technologies but it is difficult to keep the performance enough according as recent network acceleration.

We developed a monitoring system experimentally which is based on the reconfigurable hardware. This paper also shows how it worked at the live streaming event of solar eclipse.

1 はじめに

トラフィックモニタリングはネットワークの運用や機器の設計に関連する基礎的で重要な技術の一つであるが、Gigabit Ethernetの普及に代表される近年の著しいネットワーク速度の向上に対応することが難しくなりつつある。

例えば Gigabit Ethernet インタフェイスはもう十分に普及しているが、これを安価な PC と UNIX ベースの OS を組み合わせてモニタリングするのは処理能力が不足しており、インテリジェントな専用ハードウェアで実現するとコストが上がりすぎる。

対して FPGA に代表されるリコンフィギャラブルなハードウェアの能力があがっており、こうした「柔軟なハードウェア技術」を用いた Gigabit レベルのトラフィック処理が可能になっている。([2], [3])

本研究ではリコンフィギャラブルハードウェアを用いたモニタリングシステムを開発し、インターネット

日食中継イベントで実験したので報告する。

以下、2章で開発したシステムの設計について説明し、3章で実験結果を示しながら、4章で考察する。

2 設計

設計目標はロストのないキャプチャ能力と安定した大量データの記録を比較的安価に実現することである。

開発したシステムはパケットを拾うプローブ部と、それを記録するストレージ部からなる。両者の接続はネットワーク (Gigabit Ethernet) による。

プローブ装置は 1000BaseT の NIC を 2 つとリコンフィギャラブルプロセッサ (IPFlex 社製) を 2 チップ搭載した既存のボードを流用したものである。このチップは 32 の要素プロセッサ (PE) からなり、外部から PE 間の配線とソフトウェアを与えることができる。

今回はこのチップをプログラムして、一方の NIC から入ってきたパケットのヘッダ部分を取り出し、タイ

ムスタンプをつけて蓄積し、一定量がたまってきたところで他方の NIC からストレージ装置に送り出す。このとき複数パケットのヘッダを一つのパケットにまとめて送出し、送受信にかかるシステム負荷を引き下げる。

計測は対象となる経路途中にタップまたはミラーポートを設けたスイッチを設置して行う。典型的なシステムの構成は図 1 のようになる。

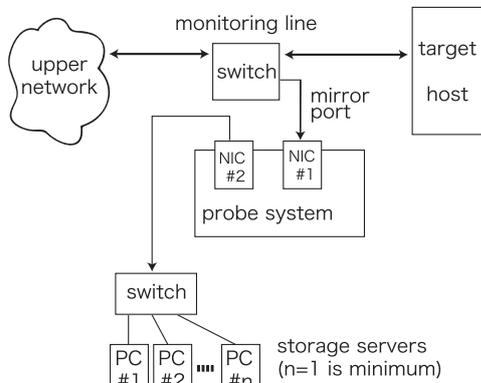


図 1: システム構成図

ヘッダに記録用のタイムスタンプや通し番号をつけているため、1Gbps の 80% ほどが短い TCP パケットだった場合には 75MB/sec 程の記録速度が要求される。そのため記録サーバは複数用意し、プローブからの宛先をラウンドロビンさせて負荷分散することを可能とした。

3 実験

非営利団体ライブ!ユニバース [4] が中心となって 2005 年 10 月 3 日に実施された日食中継プロジェクト「LIVE! ECLIPSE 2005 Annular[5]」で、主たる Web コンテンツを提供していた Web サーバを対象に計測実験を行った。ストレージ部分は今回最小構成である n=1 すなわち一台の Linux サーバだけで実施した。



図 2: 実験風景

現場での設置状況を図 2(写真) に示す。写真左側の 1U の機器は上からスイッチ、プローブ装置、記録サー

バである。プローブと記録サーバはクロスケーブルで直結した。状態監視用のノート PC や記録サーバのディスプレイ等が見えるが、なくても問題ない。

計測時間帯は 16 時 23 分 ~ 19 時 58 分。通過パケット数は 1 億 5300 万 (IPv4 のみ)、アドレスは 7 万 1000 に及ぶ。ピークでのトラフィックはおよそ 90Mbps 程であった。

4 考察とまとめ

本研究は全パケットの記録を重視しているが、事前テストの不充分さもあり実験ではそれに失敗した。キャプチャ部分ではなく、その後のプローブ=ストレージ間の通信でのロスが原因である。

本システムでは信頼できるタイムスタンプをパケット受信時にハードウェアが確実につけられることも特徴の一つである。これは負荷状態等の影響を受けやすいソフトウェアベースのシステムに対するアドバンテージであるが、しかしハードウェア開発はコスト増という問題を生じる。

本研究ではこれをリコンフィギャラブルプロセッサを利用することで実質的に既存の汎用ハードウェアとソフトウェア開発の組み合わせで解決した。

同種の研究である SCAMPI[2] も ENDACE[3] 社のリコンフィギャラブルハードウェア (ただし FPGA) による汎用ハードウェアを利用しているが、FPGA はゲートレベルのプログラミングを要し、本システムの方がソフトウェア開発がより容易で柔軟性が高い。

本開発はまだ始まったばかりである。今回の実験では一部課題も残したが現段階の成果として評価し、今後はロスのないプローブ=ストレージ間の通信方法を検討する予定である。

謝辞 実験の機会を与えてくださった LIVE! ECLIPSE 2005 Annular の関係各位に感謝します。

参考文献

- [1] 安田豊, “インターネットのトラフィック計測とその分析”, 情報処理学会研究報告, 99-DPS-92, pp.73-78, 1999
- [2] J. Coppens, et al, *SCAMPI - A Scalable Monitoring Platform for the Internet*, Proceedings of the 2nd International Workshop on Inter-Domain Performance and Simulation (IPS 2004), September 2004
- [3] ENDACE, <http://www.endace.com/>
- [4] ライブ!ユニバース, <http://www.live-universe.org/>
- [5] LIVE! ECLIPSE 2005 Annular, <http://www.live-eclipse.org/>