Reliable and Scalable Packet Striping

Hari Adiseshu

Guru Parulkar

George Varghese

Washington University in St. Louis
Packet Striping

Introduction

- Packet Striping
- Used for performance or cost reasons
- Using multiple links as a single logical link
Packet Striping

Current Striping Techniques

- Usually at the physical level (Inverse Multiplexing)
- Physical level striping has drawbacks
- Wanted: packet striping at a logical level
- MultiLink PPP provides framing, not load sharing

Current Striping Techniques

Reverse
Multiplexing

Inverse
Packet Striping

Channel Striping Algorithm

Desirable properties:
- Load Sharing
- FIFO delivery

Assume: All Channels are FIFO

Sender

Receiver

Packet Striping Reference Model
Overview of Talk

- How to achieve load sharing
- How to achieve FIFO delivery
- How to recover from packet loss
- Experimental results
Contributions

- Provide a transformation of a class of fair queuing algorithms into striping algorithms
- Show how to recover from packet loss at the receiver
- Show how to achieve Quasi-FIFO delivery
- Provide a specific instance of such a fair queuing algorithm
- Providing load sharing
- Providing a transformation of a class of fair queuing algorithms
Packet Striping

Equal Sized Packets

Round Robin Striping

RR Striping Algorithm

100 4
100 2
100 3
100 1
No Load Sharing with unequal sized packets

Round Robin Striping

Packet Striping

Round Robin Algorithm

Striping

100 4
100 2
200 3
200 1
200 2
200 1
100 4
100 3
200 2
100 1
200 1
100 4
100 2
200 3
200 1
Use function $f(s)$ to select queue
Transmit packet at head of queue
Update state using function $g(s', d')$

Causal Fair Queuing
Load Sharing

CFQ and Load Sharing
With the same fairness properties transformed into Load Sharing algorithms, CFQ algorithms can be modeled as repeated applications of two functions. Theorem: Causal CFQ algorithms can be transformed into Load Sharing algorithms in which sent so far queue to be served depends only on packets. CFQ algorithms - FQ algorithms in which Load Sharing.
Surplus Round Robin (SRR)

Example of a CQ algorithm

Quantum and a Deficit Counter (DC)

Each input queue associated with a
Quantum and a Deficit Counter (DC)

Queues served in round robin order

At each visit to a queue, DC is incremented

By Quantum

Each queue served as long as DC is positive

Each queue served as long as DC is positive

Packet Striping

Hari Adiseshu@Washington University
Packet Scheduling

Output Channel

Quantum = 500

DC1 = DC2 = 0

Queue 1
Queue 2

Round 1
Round 2

Initialization
SRR Causal Fair Queuing

DC1 = DC2 = 0

Queue 2
Queue 1

DC1 = DC2 = 0

SRR Causal Fair Queuing
Packet Striping

Input Queue

SRR Transform for Load Sharing
SRR Results

The SRR transform provides load sharing if the Quantum associated with each channel is proportion to its capacity, and is greater than the maximum packet size. In rounds after initialization, the difference between data actually sent over a channel and the amount that actually should have been sent is bounded by Quantum.
Contributions

- Provide a transformation of a class of fair queuing algorithms into striping algorithms providing load sharing at the receiver.
- Provide a specific instance of such a fair queuing algorithm achieving Quasi-FIFO delivery.
- Provide a transformation of a class of fair queuing algorithms into striping algorithms providing load sharing.
FIFO delivery

Quasi-FIFO - FIFO in the absence of packet loss

Guaranteed FIFO - add sequence number to packets

Transparent Ethernet striping, e.g. ATM cells, not always possible.
Quasi-FIFO reception using SRR
The combination of a transformed version of a CFQ algorithm at the sender, and the CFQ algorithm at the receiver provides Quasi-FIFO delivery.
Effect of Packet Loss
Effect of Packet Loss
Packet Striping

Recovery from Packet Loss

Synchronization depends on shared state between sender and receiver.

On packet loss, the shared state is lost.

For the specific case of SRR algorithm, state consists of Round Number, and Deficit Counter values for each link.

Solution: Propagate state periodically from the sender to the receiver.

Periodically propagate these values to the receiver using marker packets.
Packet Stripping

Hari Adiseshu

@Washington University

Layer 2

IP

Convergence

Layer

Data Link

Interface

- StrIPE

- IP

StrIPE Layer
Experimental Setup
Packet Striping

NetBSD Implementation

Contributions
- RR, no logical reception
- RR, logical reception
- GRR, no logical reception
- GRR, logical reception
- SRR, no logical reception
- SRR, logical reception

ATM Channel Capacity (Mbps)

Sum of Ethernet and ATM Throughput

Application Level Throughput (Mbps)
Experimental Results

Throughput with SRR matched aggregate throughput.

Throughput more with resequencing.

Throughput of SRR greater than other variants.
Conclusions

A simple transformation of a class of FQ algorithms can be used for scalable striping, providing FIFO packet delivery with variable-sized packets and variable-speed channels.

Quasi-FIFO provides transparent striping with the marker algorithm restoring synchronization in the event of packet loss.

Scaled algorithms, providing FIFO packet delivery algorithms can be used for scalable striping.