Click Ease construction of parallel packet processing tasks

Encryption, NAT, Firewall

Many packet processing tasks require more CPU cycles

Click Ease software router construction

Motivation
Packet stops at queue orToDevice.
Packet moves using function calls.
Packet enters configuration at FromDevice.
A router is a collection of interconnected Elements.

Fig 1 – Click
Queue needs to be SMP safe
- Queue hands packets from one CPU to another
- Each FromDevice or ToDevice can run on a separate CPU

Configuration Exhibits Natural Parallelism
- Expose more parallelism with simple changes
- Take advantage of natural parallelism in configurations
- User need not worry about race conditions
- Uniprocessor configurations should work correctly
Adaptive scheduler maps PROMISC and TODEVICE onto CPUs.

- Synchronous queue transfers packets between CPUs.
- FromDEVICE and TODEVICE elements run on separate CPUs.

Implementation
Adaptive scheduler balances load across CPUs.

Work initiated by FromDevice elements are more expensive than those ofToDevice.

Scheduling
Adapts to traffic pattern
Provide good load balance

Benefits

Assigns element to thread with smallest total cost so far
Sorts all schedulable elements based on cost
Periodically (e.g., 1 second), runs bin-packing algorithm:

SMP Click samples execution time of elements

Adaptive Load Balancing
- SMP Click allows programmer to specify assignment
- Sometimes assignments trigger more cache misses
- Cannot statically determine common path from configuration

Limitations of the Adaptive Approach
Implementations and performance tuning took 3 months.

Adaptive scheduler: 290 lines of C++ code and comments

14 out of 19 only needed atomic access for counters

19 out of 161 elements need to be changed
Total of 12 streams.

Each host sends 64 byte packets to 3 other hosts.
each packet is near the cost of cache misses.

- SMP Cilck does not payoff when cost of computation on

  Cost of moving a packet between two CPUs is 3.0

  On 1 CPU, SMP Cilck forwards each packet in 3.5

  Number of cache misses increases.

- Cost of enqueue increases due to lock contention

---

IP Router Does Not Scale Well
SMIP CIck pays off when packet processing cost < cost of cache misses.
Untuned configuration may not exhibit enough parallelism

E.g. IPSec

Speed up improves as cost of computation increases

E.g. IP router

Speed up is limited by cost of cache misses

SMP Cilk Usefulness
How can we expose more parallelism?

Configuration exhibits enough parallelism for 2 CPUs.

Lack of parallelism in Configuration.
HashDemux breaks traffic into 3 streams on per flow basis.
Effects of Rewriting the Configuration
http://www.cs.mit.edu/click/

- SMP Click is freely available at
- Users need not to worry about synchronization issues
- Simple changes to configuration can expose more parallelism
  Cost of cache misses limits speed up
- Speed up improves with higher cost of computation
- SMP Click extracts available parallelism from configuration
- A Click router configuration often has natural parallelism

Conclusion