Supercharging PlanetLab – a High Performance, Multi-application, Overlay Network Platform

Jon Turner, Patrick Crowley, John DeHart, Amy Freestone, Brandon Heller, Fred Kuhns, Sailesh Kumar, John Lockwood, Jing Lu, Mike Wilson, Charlie Wiseman, Dave Zar
Department of Computer Science and Engineering

Supported by the National Science Foundation

Overlay Hosting Service

- Shared overlay infrastructure supporting many services
- Vehicle for research and deployment
- Testbed: PlanetLab
PlanetLab

- Shared overlay network testbed
- Applications run as user-space processes in virtual machines
  - Limited throughput
  - High, unpredictable latency

Supercharging PlanetLab

- Leverages network processor technology
- Standard fast-path/slow-path application structure
- Removes performance limitations
  - Supports Internet-scale throughput
  - Supports latency-sensitive applications
- Allows existing PlanetLab applications to run unmodified

Slow Path
- runs in standard PlanetLab environment
- exception packets forwarded to slow-path

Fast Path
- runs on network processor
- handles most traffic
- supports Internet-scale throughput
- supports latency-sensitive applications
- allows existing PlanetLab applications to run unmodified
SPP Components

- Control Processor (CP)
- Switch
- Line Card (LC)
- General Purpose Processing Engine (GPE)
- Network Processing Engine (NPE)

**Conventional server which coordinates system components and synchronizes with PlanetLab**

**Conventional server blades supporting standard PlanetLab environment**

**Blade containing 10GE data switch and 1GE control switch**

**Dual Intel IXP 2850 blade which forwards packets to correct PEs**

**Dual Intel IXP 2850 blades supporting application fast-paths**

---

IXP 2850 Overview

- 16 multi-threaded MicroEngines (MEs)
  - 8 thread contexts with rapid switching capability
  - Fast nearest-neighbor connections for pipelined apps
- 3 SDRAM and 4 SRAM channels (optional TCAM)
- Management Processor (MP) for control
System Control

- Instantiate new application
- Open socket
- Instantiate fast-path

Sharing the NPE

- each application has private lookup entries
- forms key for lookup
- formats outgoing packet headers
- each application has private queues
- Forwarding Infrastructure Application Specific Code
Evaluation

- IPv4
  - Packets arrive/depart in UDP tunnels
- Internet Indirection Infrastructure (i3)
  - Packets contain triggers matched to IP addresses
  - No match at local node results in Chord forwarding

![Diagram of Data Interfaces and Switch with filters and paths]

IPv4 Throughput Comparison

- Click Modular Router
- conventional i3
- standard IP forwarding
- single trigger matching and Chord forwarding

- NPE can’t keep up with full line rate for 0 byte payloads
- 80x improvement for 0 byte payloads
- 10x improvement for 1400 byte payloads

<table>
<thead>
<tr>
<th>Payload Size (B)</th>
<th>Click Modular Router</th>
<th>NPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>120B</td>
<td></td>
<td></td>
</tr>
<tr>
<td>400B</td>
<td></td>
<td></td>
</tr>
<tr>
<td>800B</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1400B</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0B</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
IPv4/i3 Fast-Path Throughput Comparison

![Graph showing IPv4/i3 fast-path throughput comparison](image)

- Constant input rate of 5 Gb/s

IPv4 Latency Comparison

![Graph showing IPv4 latency comparison](image)

- 8 IPv4 instances
Summary

- Base platform intended for overlay hosting service
- An SPP node removes performance limitations found in conventional PlanetLab nodes
- Standard fast-path/slow-path application structure eases deployment
- Future work includes
  - More flexible IXP-based NPE implementation
  - NPEs built on other hardware
  - Automatic NPE code verification
- Targeting 2 SPP nodes available by end of 2007

Questions?