Towards a new generation of information-oriented Internetworking architectures

C. Esteve Rothenberg, Fábio L. Verdi and Maurício F. Magalhães
University of Campinas, Brazil
{chesteve, verdi, mauricio}@dca.fee.unicamp.br
Outline

“out-of-the-TCP/IP-box” thinking

data / content / information networking

large flat labels

SPSwitch: fast scalable forwarding
Re-Architecturing efforts

- information-orientism
- content-centric networking
- clean-slate
- 4Ward
- RTFM
- PSIRP
- DONA
- New (flat) identifier spaces
- incentives ($)?
- feasibility work (011101010...01)
- patching
- overlays
- CDN
- IPv6
- P2P
- middleboxes
- IPv4

- ROFL
Rethinking concepts

<table>
<thead>
<tr>
<th>Original Internet</th>
<th>Information-Oriented / Content-Centric Internetworking</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sender</td>
<td>Content producer (publisher)</td>
</tr>
<tr>
<td>Receiver</td>
<td>Content consumer (subscriber)</td>
</tr>
<tr>
<td>Sender-based control</td>
<td>Receiver-based control</td>
</tr>
<tr>
<td>Client/Server communications</td>
<td>Publish/Subscribe</td>
</tr>
<tr>
<td></td>
<td>Sender and Receiver uncoupled</td>
</tr>
<tr>
<td>Host-to-host</td>
<td>Service access / Information retrieval</td>
</tr>
<tr>
<td>Topology / Domain</td>
<td>Information scope</td>
</tr>
<tr>
<td>Unicast</td>
<td>Unified uni-, multi- and anycast</td>
</tr>
<tr>
<td>Explicit destination</td>
<td>Implicit destination</td>
</tr>
<tr>
<td>End-to-End (E2E)</td>
<td>End-to-Data (E2D)</td>
</tr>
<tr>
<td>Host name (look-up oriented)</td>
<td>Data/Content name</td>
</tr>
<tr>
<td></td>
<td>(“search” activity)</td>
</tr>
<tr>
<td>Secure channels, host authentication</td>
<td>Integrity and trust derived from the data</td>
</tr>
</tbody>
</table>
Challenges & Paradigm

• Common **challenge** in data-oriented paradigms:
  – Take switching decisions
    • at *wire speed* (Gbps)
    • on a *large* universe (e.g., 256-bit hash values)
    • of *flat* (non-aggregatable) identifiers

• Let’s take **advantage** of the data-oriented paradigm:
  – Pub/sub inherently tolerates *false positives*
  – Opportunistic *caching*
# Longest IP prefix vs. flat label matching

<table>
<thead>
<tr>
<th></th>
<th><strong>IPv4</strong></th>
<th><strong>IPv6</strong></th>
<th><strong>256-bit flat label</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dec.</strong></td>
<td>81.216.171.106</td>
<td>ca12:b9fa:655a:0000:0000:ac2f:ccef:f0ab</td>
<td>08090A0B0D0E0F10121314151718191A1C1D1E1F21222324262728292B2C2D2E</td>
</tr>
<tr>
<td><strong>Bin.</strong></td>
<td>01010001 11011000 10101011 01101010</td>
<td>11001010 00100101 10111001 11110110 11001010 01011010 00000000 00000000 10101100 00000000 00000000 01011111 11001100 11101111 11110000 10101011</td>
<td>11111010 11001010 01011010 00000000 00000000 10101110 00000000 00000000 00000000 00000000 01101100 00000000 00000000 00000000 00000000 10101100 01101111 11110000 10101011</td>
</tr>
</tbody>
</table>

256-bit flat ID matching is *expensive* @ wire speed
RTFM Architecture*

- **Rendezvous**: Matches subscriptions to publications.
- **Topology**: Creates and maintains delivery trees used for forwarding traffic.
- **Forwarding**: Actual data delivery operations. (label switching and fast forwarding tables)
- **Mediation**: Node-to-node link data transfer & More (opportunistic caching, collaborative and network coding)

- Metadata and hash-based identifiers

**Identifiers space (approx.)**
- $10^{15}$ rendezvous identifiers (256-bit RiD)
- $10^{10}$ scope identifiers (256-bit SiD)
- Forwarding identifies (256-bit FiD)

* [Särelä et al. 2008]
The role of Bloom and family

• Well known Bloom filters
  – Efficient *data aggregators*
  – Performance: $f$ (memory / # elements)

• High-speed router requirements
  – Low (bounded) packet processing *time* (constant time to hash)
  – High-speed *memory* limitations

• Our first *naïve* p-bank Bloom-filter-based switching approach:
  – Bloom filter *membership-problem*

*Is label* $x$ *in outport* $P_y$? 

110010100110010 ... 011111001010
Assumption!

Given an incoming packet identified by a flat ID, which is the output port/interface?
Limitations of standard BFs

a) lack of associated \textit{values}
b) expensive \textit{deletion}
c) no notion of \textit{time}
d) \textit{unbalanced} usage of memory per outport

\textbf{We need a more \textbf{flexible} (probabilistic) data structure!}

\rightarrow \textit{d-left Fingerprint Compressed Filter (FCF)}*

*recent results by Bonomi et al (2006)
Statefull d-left FCF approach

InsertDBRdBF(s, p):

Label (s)   Port-Out (p)

011 ... 101000110001101101011001010101110110 10011011

H₀(s)       H₁(s)       H₂(s)

d = 0       d = 1       d = 2

h

b

Bucket (b)

Count: 3
Experimental results

Table 2: Analytical and experimental comparison of different data structures for the switching procedures.

<table>
<thead>
<tr>
<th>Method</th>
<th>Mem. access</th>
<th>Mem. size M</th>
<th>(Mbits)**</th>
<th>(bpc)</th>
<th>False positive</th>
<th>(predicted)**</th>
<th>(actual)**</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard Table</td>
<td>$O(n) - O(1)^*$</td>
<td>$n \times (s + p)$</td>
<td>253.68</td>
<td>266.0</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>Fingerprint Table</td>
<td>$O(n) - O(1)^*$</td>
<td>$n \times (f + p)$</td>
<td>28.61</td>
<td>30.00</td>
<td>$2^{-f}$</td>
<td>9.54 $\times 10^{-7}$</td>
<td>-</td>
</tr>
<tr>
<td>P-bank BF</td>
<td>$O(1)$</td>
<td>$2^p \times m^{***}$</td>
<td>43.63</td>
<td>45.75</td>
<td>$\approx 2^p \times 0.62^{M/n}$</td>
<td>2.91 $\times 10^{-7}$</td>
<td>4.33 $\times 10^{-3}$</td>
</tr>
<tr>
<td>d-left FCF</td>
<td>$O(1)$</td>
<td>$d \times b \times h \times (f + p)$</td>
<td>42.92</td>
<td>45.00</td>
<td>$&lt; d \times h \times 2^{-f}$</td>
<td>1.72 $\times 10^{-5}$</td>
<td>1.51 $\times 10^{-5}$</td>
</tr>
<tr>
<td>d-left FCF DBR</td>
<td>$O(1)$</td>
<td>$d \times b \times (h \times (f + p) + c)$</td>
<td>43.63</td>
<td>45.75</td>
<td>$&lt; d \times h \times 2^{-f'}$</td>
<td>3.57 $\times 10^{-6}$</td>
<td>3.46 $\times 10^{-8}$</td>
</tr>
</tbody>
</table>

* Assumes a perfect hash function. ** Parameters: $n = 1.000.008$; $d = 3$; $b = 83.334$; $f = 20$; $p = 10$; $h = 6$; $c = 3$; $s = 256$.
*** Total memory of the p-bank Bloom filters equal to the value $M$ of the d-left FCF DBR. $m = M/2^p$; $k_{opt} = 31$.

20-bit fingerprint + 10-bit port

![Graph showing false positive rate vs fingerprint bits (f)](image)

- Standard
- d-left FCF
- Hash table*

false positive rate [log]

fpr = $O(10^{-6})$
Future work
Conclusions
Q&A procedure

Question

Answer known?

Yes

Response

No

Thank audience

Claim the time is over

Yes

No