LightPS: Lightweight Content-based Publish/Subscribe for Peer-to-Peer Systems

University Rovira i Virgili, Spain

Jordi Pujol Ahulló
jordi.pujol@urv.cat

Pedro García López (URV)
Antonio F. Gómez Skarmeta (UM)
Outline

- LightPS:
  - Lightweight Content-based Pub/Sub
  - Peer-to-Peer networks
  - Rendezvous-based Approach

- Introduction
- LightPS Operation
- Theoretical Evaluation
Distributed Hash Table (DHT) main properties

- Identifier of $m$ bits
  Keyspace: $[0 .. 2^m)$

- Local routing state: $O(\log N)$

- Communication cost: $O(\log N)$
Challenges of dealing with a content-based pub/sub onto DHTs

From our particular viewpoint:

- **Lightweight** proposal, avoiding to build a pub/sub overlay
- **Portable**, in order to work onto current DHTs

- **Multi-dimensional data**
  - DHT keyspace is unidimensional

- **Multiple publishers** for the same flow
  - Will we need advertisements?
Multi-dimensionality is hard to overcome

- Main approaches for DHTs:
  - Address each dimension separately
    E.g. CAN-based solutions, PastryStrings
  - Employ a dimensional reduction function
    E.g. SFC: Z-curve, Hilbert curve (LPHF)

- Our particular proposal:
  - BitMapping (BM) function:
    - An event produces one key
    - A subscription produces a set of keys
BM defines how to deal with multi-dimensional data (I)

- Assumptions:
  - Data is formed by integer values
  - PR,L monotone function (a.k.a. OPHF)

- Event mapping:

  Step 1: \{ a_1, a_2, a_3, .., a_D \} \quad // |a_i| = R

  \[ P_{R,L}(a_i) \]

  Step 2: \{ 00, 10, 01, .., 11 \} \quad // |t| = L = 2

  Bit Concatenation

  Step 3: 001001..11 \quad // |key| = (\Sigma |t|) \leq m
BM defines how to deal with multi-dimensional data (II)

Subscription mapping:

- Selectivity ratio per dimension $s_i$: $s_1 = 0.75; s_2 = 0.5 \Rightarrow s_{\text{TOTAL}} = 0.75 \times 0.5 = 0.375$
- Expected covered keyspace: $s_{\text{TOTAL}} \times 2^m$

Step 1: $\{[\text{min}_1..\text{max}_1], [\text{min}_2..\text{max}_2]\}$ // $|v| = R$

$$P_{R,L}(v)$$

Step 2: $\{[00..10], [01..10]\}$ // $|t| = L = 2$

Cartesian Product

Step 3: $\{0001, 0010, 0101, 0110, 1001, 1010\}$ // $|\text{key}| = m = 4$
BM defines a generic routing algorithm (I)

- **Assumption:**
  - Comm. cost of underlying DHT: $O(\log N)$

- **Rendezvous-based behavior:**
  - Deliver msg to key owner
  - Delivery: $O(\log N)$ hops

- **As BM(event) = key**
  - Delivery: $O(\log N)$ hops

- **Action:**
  - Start the notification

- **Consequence:**
  - Multiple publishers for the same flow
BM defines a generic routing algorithm (II)

- As BM(subscription)=keyset
  - Breath-first-like routing

- Steps:
  1) Segment detection
  2) Route message

- Delivery:
  - $O(\log N + \beta)$ hops

- Action:
  - Store subscription
  - Notify events to apps
LightPS evaluation (I)

- BM tradeoff:
  - $\sum b_i = K \leq m$ bits, but $K$ close to $m$ to balance the load through entire keyspace

- Subscription coverage cost:
  - Variables: $D, b_i, s_i$ user & app. dependant!
  - Measure: Keyset size of subscription mapping
  - Lower bound: $\Omega(\prod(2^{b_i} s_i))$
  - Higher bound: $O(\prod((2^{b_i} s_i)+1))$
LightPS evaluation (II)

\[ b_i = \frac{m}{\#\text{dims}} \]

\[ m = 64 \]

\[ m = 160 \]
LightPS evaluation (III)

- High-dimensional context effect
  - Better performance on high-dimensional contexts
  - Having low number of mapping bits ($b_i > 2$)

<table>
<thead>
<tr>
<th>Num. Nodes</th>
<th>16K</th>
<th>1M</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dimensions</td>
<td>40</td>
<td>80</td>
</tr>
<tr>
<td>25%</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>50%</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>75%</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

AVG. number of covered nodes.
$m=160$. $b_i=m/\#\text{dims.}$
Conclusions (I)

- We have proposed LighPS, a new framework for
  - Rendezvous-based pub/sub systems
  - Working onto DHTs

Properties:
- Deals with multi-dimensional data: BM
- Multiple publishers can coexist
- Lightweight: avoids building a pub/sub overlay
- Portable, leveraging the underlying DHT routing infrastructure
Conclusions (II)

- We have presented BitMapping (BM):
  - Formal performance analysis
  - Necessary tradeoffs
- Balance the load through whole keyspace
- LightPS have better performance on high-dimensional contexts
- Scales to big p2p networks
Future work

- Evaluate the work through **simulations** to check expected performance
- Application of our generic algorithms based on BM in other content distribution fields **worth exploring**
Thank you!

University Rovira i Virgili, Spain

Jordi Pujol Ahulló
jordi.pujol@urv.cat

Pedro García López (URV)
Antonio F. Gómez Skarmeta (UM)