Maglev: Loadbalancer

- NSDI 2016 paper: Daniel Eisenbud et.al
- Network load balancer
- Horizontal scaling
- Hundred of commodity servers as Load balancers
- Use consistent hashing
scalability

$ft = (S, D, P_S, P_D, prot)$

$H_1(ft)$

$H_2(ft)$

ECMP-256way L3

$H_1(ft)$ mod 5

Maglev

$H_2(ft)$ mod 11

DIPS
Scalability : Maglev changes

Key : ft does not change
Scalability: Backend changes

ft = (S, D, P_s, P_d, prot)

H_1(ft)

H_2(ft)

ECMP-256way L3

H_1(ft) mod 5

Maglev

H_2(ft) mod 11

DIPS

Key: ft does not change
Challenges

• Connections on failed backend need to be moved
• Consistent hashing mitigates movement
• Unlike data movement, connection movement means lost connections
• Maglev consistent hashing
• Each backend is assigned a permutation of $M \gg N$ (# of backends)
Example: 3 backendservers M is 7

- Server0: 3,0,4,1,5,2,6
- Server1: 0,2,4,6,1,3,5
- Server2: 3,4,5,6,0,1,2

Form a look up table to decide which server gets assigned $h_2(ft)$

- Each backend server picks one number in turn
- moves through the permutation until the first slot not already taken
Example: 3 backendservers M is 7

- Server0: 3,0,4,1,5,2,6
- Server1: 0,2,4,6,1,3,5
- Server2: 3,4,5,6,0,1,2

<table>
<thead>
<tr>
<th></th>
<th>S1</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>S0</td>
</tr>
<tr>
<td>1</td>
<td>S0</td>
</tr>
<tr>
<td>2</td>
<td>S1</td>
</tr>
<tr>
<td>3</td>
<td>S0</td>
</tr>
<tr>
<td>4</td>
<td>S2</td>
</tr>
<tr>
<td>5</td>
<td>S2</td>
</tr>
<tr>
<td>6</td>
<td>S0</td>
</tr>
</tbody>
</table>
Failure of S2

- S2 Fails;
- Server0: 3,0,4,1,5,2,6
- Server2: 3,4,5,6,0,1,2

<table>
<thead>
<tr>
<th></th>
<th>Before</th>
<th>After</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>S1</td>
<td>S0</td>
</tr>
<tr>
<td>1</td>
<td>S0</td>
<td>S0</td>
</tr>
<tr>
<td>2</td>
<td>S1</td>
<td>S0</td>
</tr>
<tr>
<td>3</td>
<td>S0</td>
<td>S0</td>
</tr>
<tr>
<td>4</td>
<td>S2</td>
<td>S2</td>
</tr>
<tr>
<td>5</td>
<td>S2</td>
<td>S2</td>
</tr>
<tr>
<td>6</td>
<td>S0</td>
<td>S2</td>
</tr>
</tbody>
</table>
Maglev: NLB

- Uses 256-way L3 ECMP
- Consistent hashing minimizes connection changes
- Can go at 10 G NIC