ANALYSIS AND IMPROVEMENT OF VALIANT ROUTING IN LOW-DIAMETER NETWORKS

Mariano Benito
Pablo Fuentes
Enrique Vallejo
Ramón Beivide

With support from:

4th IEEE International Workshop of High-Performance Interconnection Networks in the Exascale and Big-Data Era (HiPINEB)
Vienna, Austria, 24-Feb, 2018.
Index

1. Background and motivation
2. Improvements to Valiant routing
3. Performance evaluation
4. Discussion and conclusions
1. Background and motivation

- Valiant routing
  - Randomized Routing mechanism originally proposed by Leslie Valiant for Hypercubes in [1] and square mesh, d-way shuffle and shuffle-exchange graphs networks in [2].
  - Diverts traffic to an intermediate router
  - Double path length on average wrt minimal routing
  - Bounded worst-case permutation time
  - Oblivious

1. Background and motivation

- Valiant has been used in low-diameter system networks
- Highly-scalable, low-cost topologies

- Low diversity of minimal paths
- Concentration (multiple nodes per switch)
- Valiant routing avoids such patterns of congestion
  - Often implemented as part of an adaptive routing mechanism.

---

1. Background and motivation


Index

1. Background and motivation
2. Improvements to Valiant routing
   1. Restricted intermediate router selection
   2. Path recomputation
3. Performance evaluation
4. Discussion and conclusions
2. Improvements to Valiant Routing

2.1 Intermediate router selection

• A variant of the turn-around problem may occur without packets visiting switches twice.
  • Packets leave and return to a given network partition.

Dragonfly (with global trunking)

Flattened Butterfly
2. Improvements to Valiant Routing

2.1 Intermediate router selection

1. Determine network partitions
2. When both source & dest. nodes belong to the same partition
   1. Select intermediate node inside the partition
3. Otherwise
   1. Select intermediate node anywhere in the network.
2. Improvements to Valiant Routing

2.2 Recomputation

- Congestion situations may appear despite the randomization mechanism
  - Particularly when Valiant is used as part of an adaptive routing mechanism
- **Valiant with recomputation** makes a new intermediate node selection when the output port is not available
2. Improvements to Valiant Routing

2.2 Recomputation

• Valiant with Recomputation (VAL-Recomp) is no longer oblivious
  • But it is not completely adaptive
• According to the taxonomy by P. Graz et al in [8], Valiant with recomputation is *adaptive*, *congestion-oblivious*

Index

1. Background and motivation
2. Improvements to Valiant routing
3. Performance evaluation
   1. Simulation setup
   2. Restricted Valiant
   3. Valiant with recomputation
   4. Impact of number of injection buffers
4. Discussion and conclusions
3. Performance evaluation

3.1 Simulation setup


<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Router size</td>
<td>23 ports (h=6 global, p=6 injection, 11 local)</td>
</tr>
<tr>
<td>Group size</td>
<td>12 routers, 72 computing nodes</td>
</tr>
<tr>
<td>System size</td>
<td>73 groups, 876 routers, 5,256 computing nodes</td>
</tr>
<tr>
<td>Latency (ns)</td>
<td>40/400 (local/global links), 200 (router pipeline)</td>
</tr>
<tr>
<td>Buffer size (KB)</td>
<td>100 KB (transit queues), 200 (injection buffers)</td>
</tr>
<tr>
<td>Router</td>
<td>2x frequency speedup, Virtual Cut-Through, iterative input-first separable allocator</td>
</tr>
<tr>
<td>Routing mechanisms</td>
<td>Minimal (MIN) Valiant (VAL) Restricted Valiant (RVAL) Valiant-Recomp (VAL-Recomp) Restricted Valiant-Recomp (RVAL-Recomp)</td>
</tr>
</tbody>
</table>

Topology

- Uniform traffic
- Adversarial traffic
- Adv-local traffic
- Hot-Region traffic
- Permutation traffic

Strict permutation: No endpoint congestion
Random: pattern differs for each simulation, but is fixed during each simulation

3.2 Restricted Valiant (RVAL)

Random Uniform

Adversarial-local

- Average packet latency (us)
- Accepted load (%)

Graphs showing performance under different load conditions for MIN, VAL, and RVAL.
3.2 Restricted Valiant (RVAL)

Benito, Fuentes, Vallejo, Beivide

Analysis and improvement of Valiant routing...
3.2 Restricted Valiant (RVAL)

- Partial conclusions:
  - **Restricted Valiant** in the Dragonfly is **highly beneficial** for intra-group traffic (Adversarial-local)
  - Very small benefit (no penalty) in other cases (~1%).
3.3 Valiant with recomputation

**Random Uniform**

- Average packet latency (us)
  - Offered load (%)
- Accepted load (%)
  - Offered load (%)

**Adversarial**

- Average packet latency (us)
  - Offered load (%)
- Accepted load (%)
  - Offered load (%)
3.3 Valiant with recomputation

<table>
<thead>
<tr>
<th>MIN</th>
<th>RVAL</th>
<th>RVAL-Recomp</th>
</tr>
</thead>
</table>

**Hot-Region**

- Average packet latency (us) vs. Offered load (%)
- Accepted load (%) vs. Offered load (%)

**Random permutation**

- Average packet latency (us) vs. Offered load (%)
- Accepted load (%) vs. Offered load (%)
3.3 Valiant with recomputation

**Partial conclusions:**

- Valiant with recomputation **improves:**
  - **Stability** of the results (much less oscillations)
  - **Latency** before saturation
  - **Peak throughput**
- The recomputation mechanism is **negative for random permutations of traffic in the saturation regimen**
  - It increases congestion issues after saturation
3.4 Number of injection buffers

- When multiple buffers are used:
  - No significant difference for latency before saturation
  - Traffic injection after saturation increases with the number of buffers, what increases congestion
  - Typical behavior for 1-4 buffers under UN or ADV traffic:
3.4 Number of injection buffers

- Hot-región traffic: more buffers increase throughput with a DEST policy.

DEST policy: injection buffer selected by destination

RANDOM policy: injection buffer selected randomly, between available
3.4 Number of injection buffers

- **Random-permutation traffic**: more buffers severely increase congestion with a RANDOM policy.

DEST policy: injection buffer selected by destination

RANDOM policy: injection buffer selected randomly, between available
3.4 Number of injection buffers

• Partial conclusions:
  • Increasing the number of injection buffers increases the amount of injected traffic.
    • Increased congestion under UN, ADV and PERM
    • Reduces endpoint-congestion effect under HOT-REGION (traffic with endpoint congestion) with a per-destination buffer selection policy.
Index

1. Background and motivation
2. Improvements to Valiant routing
3. Performance evaluation
4. Discussion and conclusions
4. Conclusions and future work

- **Restricted Valiant**
  - The performance improvement is dramatical for traffic internal to a partition (a Dragonfly group).

- **Valiant with recomputation**
  - Improves the stability of the results, latency and peak throughput.
  - More throughput also increases congestion

- **Number of virtual channels:**
  - More injection channels increase congestion
  - HoLB reduction is effective in cases of endpoint congestion (Hot-Region traffic)
4. Conclusions and future work

• Our proposal for Restricted Valiant relies on *network partitions*.
  - *How to specify useful partitions for a given (nontrivial) topology?*

• How to define (proof) when the behavior of Restricted Valiant is “correct”?
  - Example: Restricted Valiant in the Dragonfly proposed by Kim *et al* in [4]
    - Denoted *Valiant-global* in [10]
    - Pathological performance under adversarial traffic identified in [11]

• L. Valiant studies the *consumption time of a worst-case permutation*.
  - Should we use this analysis?
  - Is this equivalent to minimum throughput at saturation (per router)?

---

4. Conclusions and future work

• **Other issues** we are exploring:
  • How does Restricted Valiant and Valiant with Recomputation behave when using **adaptive routing**?
  • How should we implement them in an interconnect that implements **table-based routing**?
ANALYSIS AND IMPROVEMENT OF VALIANT ROUTING IN LOW-DIAMETER NETWORKS

Mariano Benito
Pablo Fuentes
Enrique Vallejo
Ramón Beivide

With support from:

4th IEEE International Workshop of High-Performance Interconnection Networks in the Exascale and Big-Data Era (HiPINEB)
Vienna, Austria, 24-Feb, 2018.