salbnet: A Self-Adapting Load Balancing Network

Jörg Jung
University of Potsdam
Institute for Computer Science
Operating Systems and Distributed Systems
February 2014
Outline

1 Introduction 3
2 Credit based SLB 5
3 salbnet Implementation 8
4 Measurements and Evaluation 10
5 Conclusions and Future Work 16
1 Introduction

Dispatcher based Server Load Balancing (SLB): scalable, flexible and fault tolerance services
Motivation

Measurements in [Zinke and Schnor 2013] show the influence of weights.

Sophisticated algorithms are required for heterogenous workloads and heterogenous back end servers of ISPs:

Self-adapting credit based SLB algorithms for better performance.

Simulations in [Lehmann et al. 2008] show the advantages of the credit based SLB algorithms.

→ Efficient implementation for credit based SLB required.

→ Measurements to compare traditional and credit based SLB algorithms:
  Weighted Round Robin (WRR) and Dynamic Pressure Relieve (DPR).
2

Credit based SLB

Application independent *implicit* metrics are used to calculate *credits*

Back end server *push* credits to the LB

Credits represent the number of connections
Reporting Algorithms:
Dynamic Pressure Relieve (DPR) and DPR-Quantize (DPR-Q)

→ Reporting credits based on the (amount of processed) credit metric (data)
Credit Metric: TCP Backlog
salbnet Implementation

*salbd* implements metric collecting and credit reporting (runs on the LB and the back end servers)

*LVS scheduler* module implements the credits scheduling

*libnetmsg* implements network abstraction for sending messages over Ethernet and InfiniBand

*libnethook* hooks into (socket) system calls in back end servers
Dispatcher

- salbd (Server)
- libnetmsg
- libc
- librdmacm

LVS scheduler

TCP/UDP

IP

IB CM

Verbs

librdmacm

RDMA

libnetmsg

libnethook

libc

Server

- salbd (Client)
- libnetmsg
- librdmacm
- libc

httpd/named

LD_PRELOAD

Shared Memory

User space

Kernel space

TCP/UDP

IP

IPoIB

IB CM

Verbs

IB CM

Verbs

ioctl()
4 Measurements and Evaluation


Dispatcher based SLB scenario: two armed, NAT based and using route path with heterogeneous hardware and homogeneous software versions

3 heterogenous back end servers require weights for the traditional WRR algorithm
4 Measurements and Evaluation

Workload: Reduced Wikipedia Traces

Number of requests from the first ten minutes of the (filtered and reduced) Wikipedia trace from 12. November 2007 (available from [Pierre 2010])

<table>
<thead>
<tr>
<th>Factor</th>
<th>Requests</th>
<th>Mean req/s</th>
<th>Max req/s</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\frac{1}{32}$</td>
<td>49,532</td>
<td>82.55</td>
<td>91</td>
</tr>
<tr>
<td>$\frac{1}{16}$</td>
<td>99,063</td>
<td>165.12</td>
<td>183</td>
</tr>
<tr>
<td>$\frac{1}{8}$</td>
<td>198,125</td>
<td>330.21</td>
<td>366</td>
</tr>
<tr>
<td>1</td>
<td>1,584,996</td>
<td>2,641.66</td>
<td>2,925</td>
</tr>
</tbody>
</table>
Results: (First) Response Time

- **Factor $\frac{1}{32}$**
  - WRR
  - DPR
  - DPR-Q

- **Factor $\frac{1}{16}$**
  - WRR
  - DPR
  - DPR-Q

- **Factor $\frac{1}{8}$**
  - WRR
  - DPR
  - DPR-Q

*Normalized (First) Response Time*
Results: (Request) Errors

<table>
<thead>
<tr>
<th>Factor</th>
<th>WRR</th>
<th>DPR</th>
<th>DPR-Q</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\frac{1}{32}$</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>$\frac{1}{16}$</td>
<td>0.1</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>$\frac{1}{8}$</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
</tbody>
</table>

Normalized (Request) Error
Measurement Metrics

(First) Response Time, (Request) Errors and Duration are combined into single \textit{lower is better} penalty values

\textit{SLB ISP Penalty} $p_{ISP}$ used for comparison

\[
p_{ISP} = \left( \frac{\text{response}_{\text{mean}}}{\text{response}_{\text{max}}} \right) \times \left( \frac{\text{request\_error}_{\text{mean}}}{\text{requests}_{\text{total}}} \right)
\]
Results: SLB Penalty

Factor $\frac{1}{32}$

Factor $\frac{1}{16}$

Factor $\frac{1}{8}$

SLB ISP Penalty
Conclusions and Future Work

salbnet implementation for *credit* based SLB introduced

Previous simulations are confirmed:

- DPR and DPR-Q outperform traditional WRR

DPR-Q variant is slightly better than DPR, for higher workloads

Next step: salbnet and DNS, without InfiniBand and RDMA