Self-Adapting Load Balancing for DNS

Jörg Jung, **Simon Kiertscher**, Sebastian Menski, and Bettina Schnor

University of Potsdam
Institute of Computer Science
Operating Systems and Distributed Systems

Before we start …

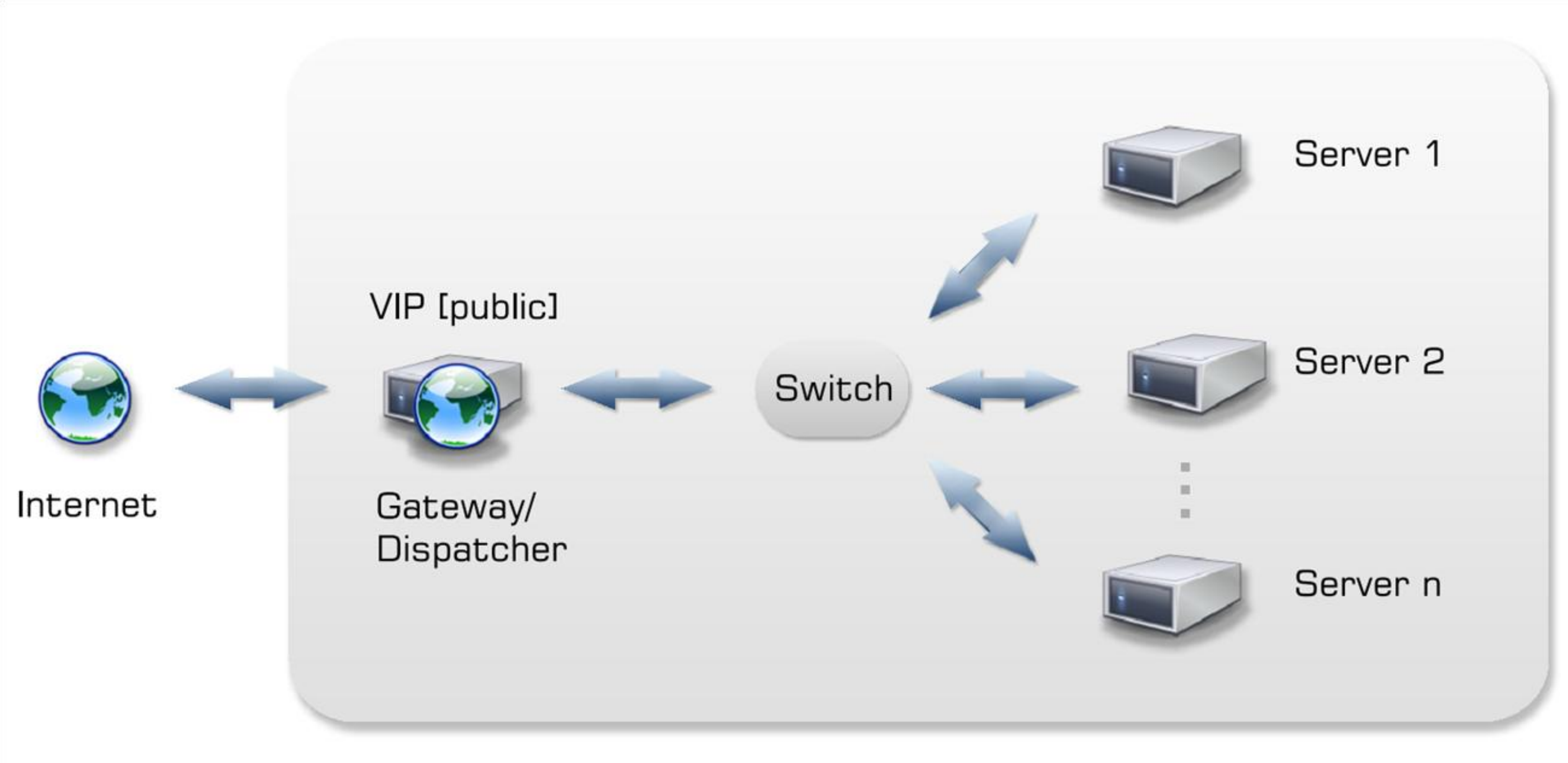
Outline

1. Introduction
2. Credit based SLB
3. Implementation
4. Measurements and Evaluation
5. Conclusion and Future Work

Introduction

Dispatcher based Server Load Balancing (SLB):

scalable, flexible, and fault tolerance services



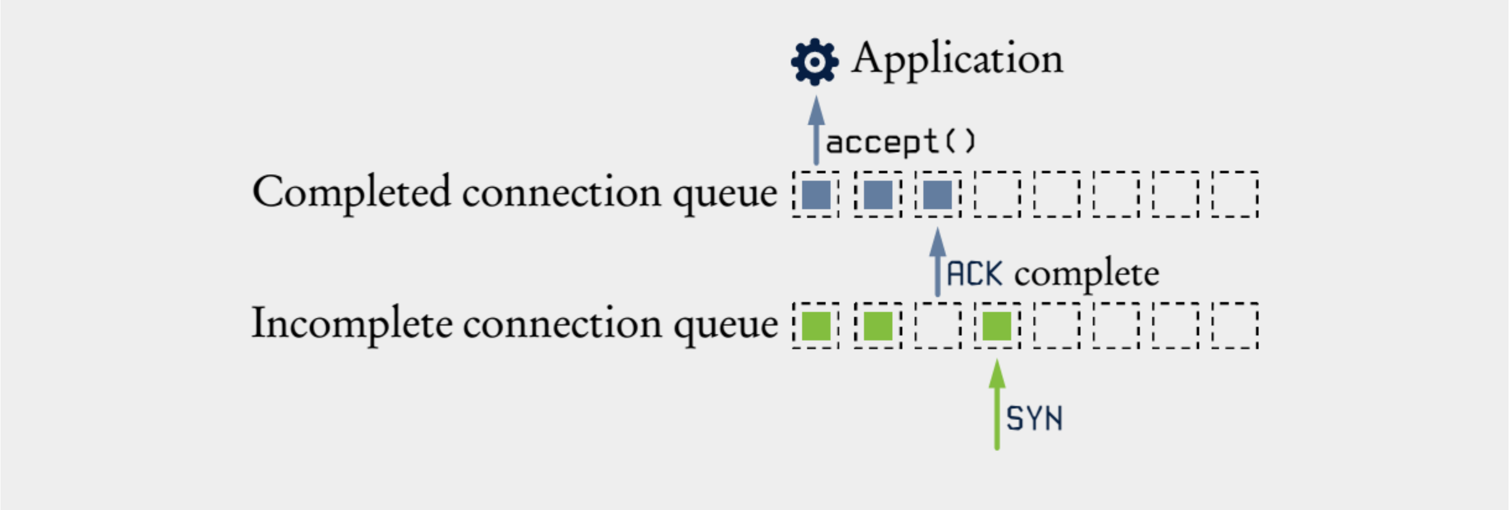
Motivation

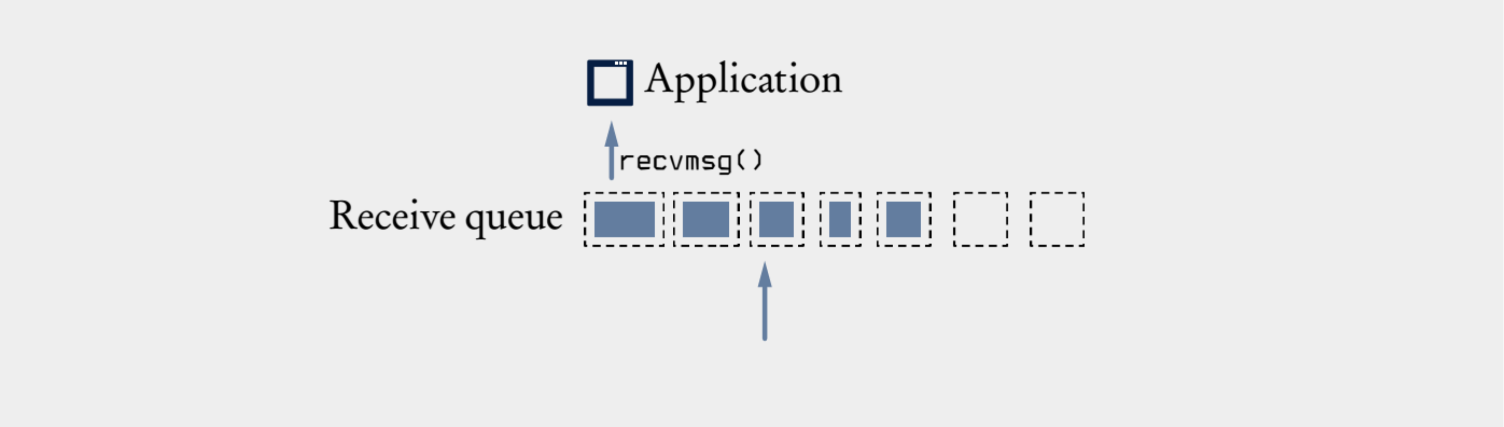
- Sophisticated algorithms are required for **heterogeneous** workloads and **heterogeneous** back end servers
- **No weights** to determine/tune
- Self-adapting ***credit*** based SLB for better performance
- Measurements show the advantages of credit based SLB for **TCP services** like HTTP [Jung et al. 2014]
- contribution:
  - → Efficient implementation for **UDP services** like DNS required, using a **suited credit metric**
  - → Measurements to compare traditional and credit based SLB algorithms: WRR vs. DPR

Credit based SLB

- introduced in [Jung et al. 2014]:
  - Application independent metrics are used to calculate *credits*
  - Back end server *push* credits to the LB
- Linux environments
- Credits represent the free capacity of the UDP receive queue / TCP backlog queue
  (in number of UDP requests or TCP connections)

receive queue vs backlog queue

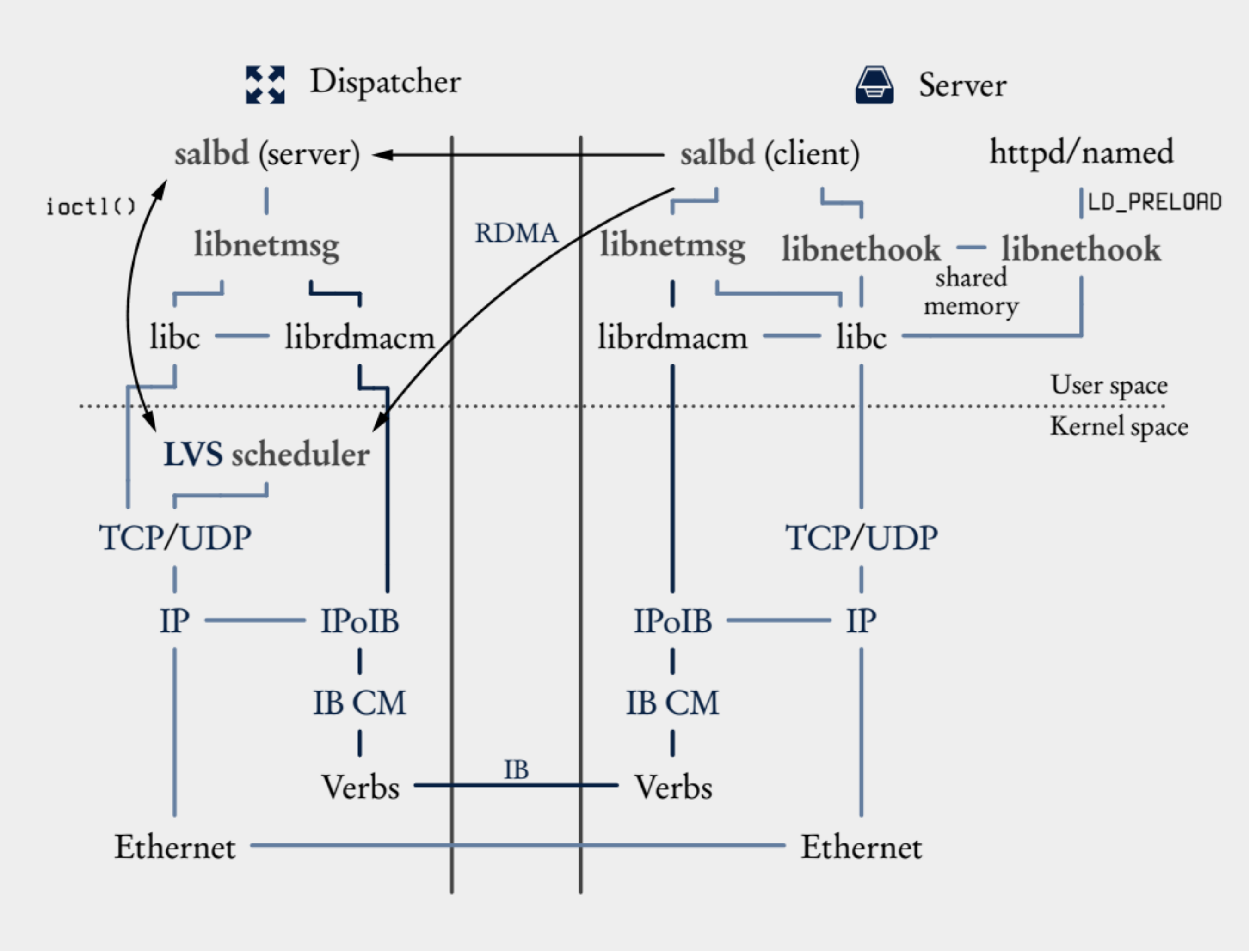




Credit Metric: UDP Receive Queue

- UDP receive queue as credit metric implicitly reflects the current load on the back end servers:
  - A filled receive queue indicates a busy application
  - With a full queue the application if fully engaged and might not respond to requests

Implementation - Overview of salbnet



Implementation

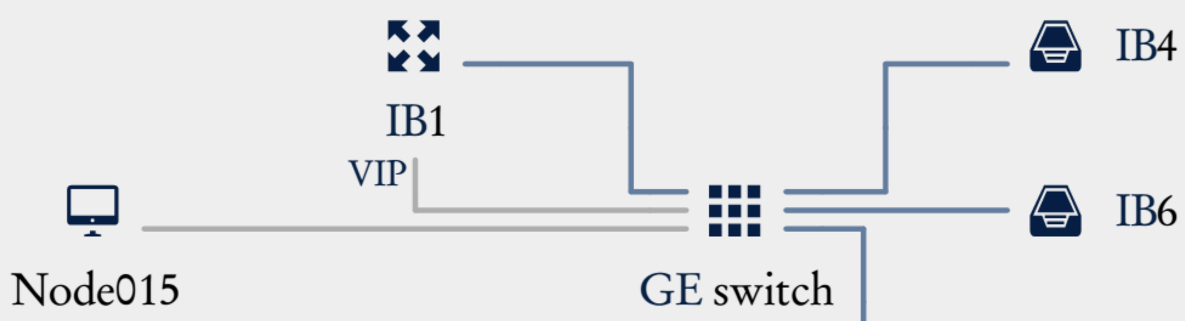
- Using salbnet libnethook for recvmsg() system call interception
  - → get size of the received message (payload)
  - → transparent for UDP applications, like DNS server (BIND)
- UDP packet size and required space in memory needs to be predicted
  - → Find out free capacity
  - → Experimentally determined for Linux kernel
  - → Proof of concept prototype: dependent on specific kernel version

Measurements and Evaluation

- Measurements in Internet Service Provider (ISP) like SLB environment using BIND instances
- Heterogeneous hardware and homogenous software versions
- WRR vs. DPR
- 3 heterogeneous back end servers require weights for traditional WRR algorithm

  → Determined in beforehand measurements [Zinke and Schnor 2013]

Environment: Setup



Universität Potsdam — Mathematisch-Naturwissenschaftliche Fakultät

Environment: Hardware

Hostname	CPU
LVS and IB4	Dual 1.8 GHz AMD Opteron 244
IB6	2.8 GHZ Intel Pentium 4
IB8	1.86 GHz Dual Core Intel Xeon 3040

- All machines have 4 GByte memory and GBit links
- Client (load generator) runs with 12 GByte memory

Benchmark: servload [Zinke et al. 2012]

- Existing DNS benchmarks like *queryperf* or *DNSPerf* and *ResPerf* are not able to replay queries
- In contrast benchmark *servload* tries to replay real user sessions

  → Factor feature replicates real user sessions

  → Support for DNS protocol added to *servload*

Environment: Software

- 3 BIND DNS servers v9.3.6
- LVS LB with ipvsadmin v1.24
- Client with servload v0.5.1
- OS LB and Servers: CentOS Linux 5.7 with kernel 2.6.18
- OS Client: Debian Linux with kernel 2.6.26
- Monitoring: SNMPv1 requests once a minute from LB (in addition to the credits)

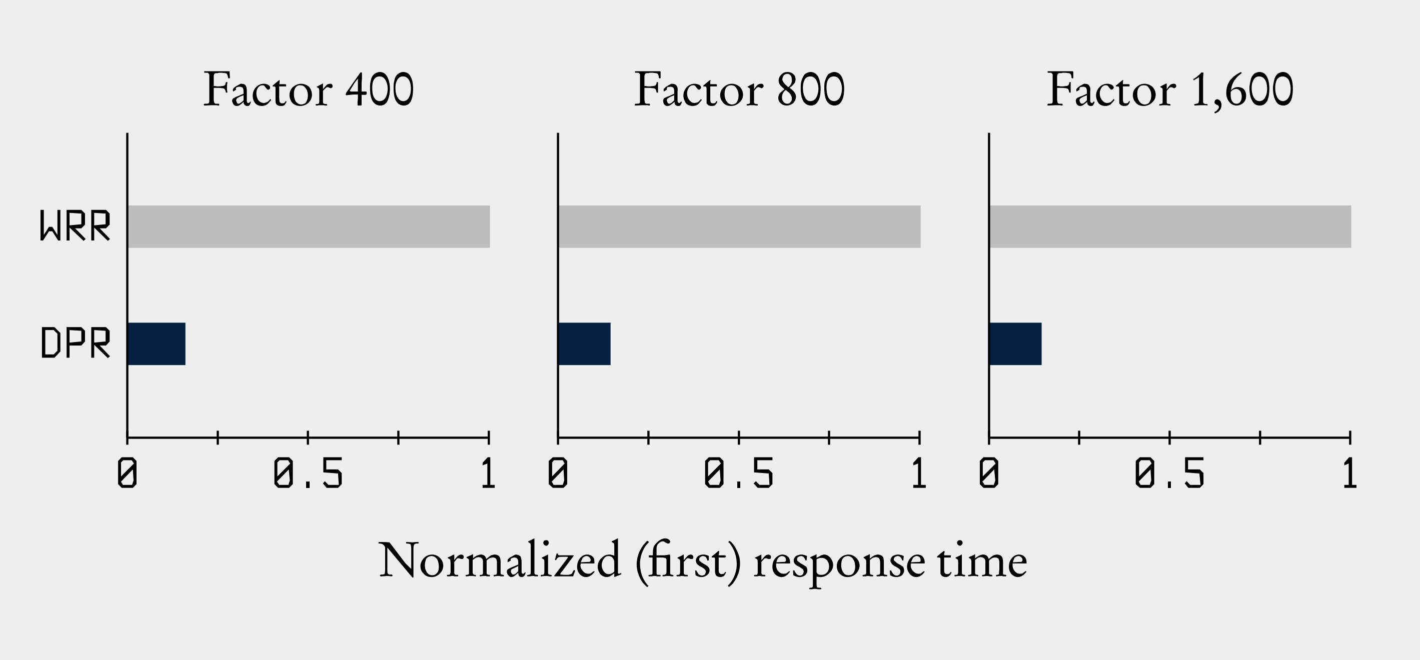
Workload: Increased DNS Traces

Number of requests from five minutes of a trace from the 29th September 2011 from DNS servers running

haiti.cs.uni-potsdam.de

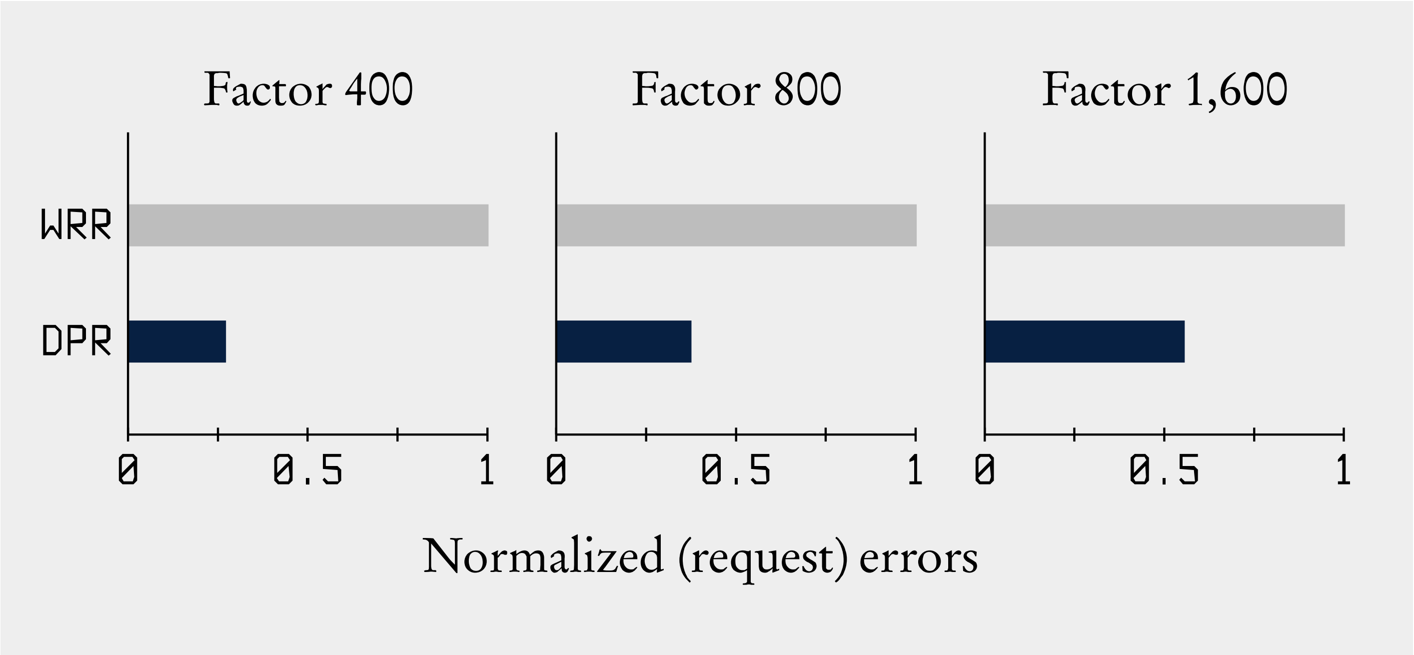
Factor	Requests	Sessions	Mean req/s	Max req/s
1	22,594	33	75.31	204
400	9,037,600	13,200	30,125	81,600
800	18,075,200	26,400	60,250	163,200
1,600	36,150,400	52,800	120,501	326,400

Results: (First) Response Time



Method	F 400	F 800	F 1600
WRR	4.05 ms	4.33 ms	4.24 ms
DPR	0.64 ms	0.62 ms	0.60 ms

Results: (Request) Errors



Method	F 400	F 800	F 1600
WRR	1273 (0.01%)	2038 (0.01%)	2910 (0.01%)
DPR	343 (0.00%)	759 (0.00%)	1612 (0.00%)

Conclusions and Future Work

- Credit based SLB extended to support UDP services
- servload benchmark extended to work with DNS
- Previous TCP (HTTP) measurements are confirmed
- DPR outperforms traditional WRR with already tuned weights
- by 15% (FRT) and 25%-50% (error rate)
- salbnet scales
- Next: transition from proof concept to mature solution

Universität Potsdam Mathematisch-Naturwissenschaftliche Fakultät

Thank you for your attention!
Any questions?

Contact:

jozinke@cs.uni-potsdam.de

www.cs.uni-potsdam.de

References

- queryperf is part of the BIND distribution available from: [ISC 2014] ISC. BIND | Internet Systems Consortium, May 2014. URL http://www.isc.org/downloads/bind/. Accessed May 2014.
- DNSPerf and ResPerf are available from: [Nominum 2014] Nominum. Measurement Tools | Nominum, February 2014. URL http://www.nominum.com/support/measurement-tools/. Accessed May 2014.
- [Zinke et al. 2012] Jörg Zinke, Jan Habenschuß, and Bettina Schnor. servload: Generating Representative Workloads for Web Server Benchmarking. In Proceedings of the 15th International Symposium on Performance Evaluation of Computer and Telecommunication Systems, pp. 82–89. SPECTS 2012. IEEE Communications Society. Society for Modeling & Simulation International (SCS), Genoa, Italy, July 2012.
- [Zinke and Schnor 2013] Jörg Zinke and Bettina Schnor. The Impact of Weights on the Performance of Server Load Balancing Systems. In Proceedings of the 16th Internation- al Symposium on Performance Evaluation of Computer and Telecommunication Systems, pp. 541–548. SPECTS 2013. IEEE Communications Society. Society for Modeling & Simulation International (SCS), Toronto, Canada, July 2013.
- [Jung et al. 2014] Jörg Jung, Bettina Schnor, and Sebastian Menski. salbnet: A Self-Adapting Load Balancing Network. In Proceedings of the 12th IASTED International Conference on Parallel and Distributed Computing and Networks, pp. 249–257. PDCN 2014. ACTA Press. IASTED, Innsbruck, Austria, February 2014.