Universität Potsdam
Mathematisch-Naturwissenschaftliche Fakultät

Optimizing Energy Efficiency and Quality of Service in Large Scale Web Server Environments

2016-10-07

Fachgruppe Betriebssysteme - Herbsttreffen

Simon Kiertscher, Bettina Schnor

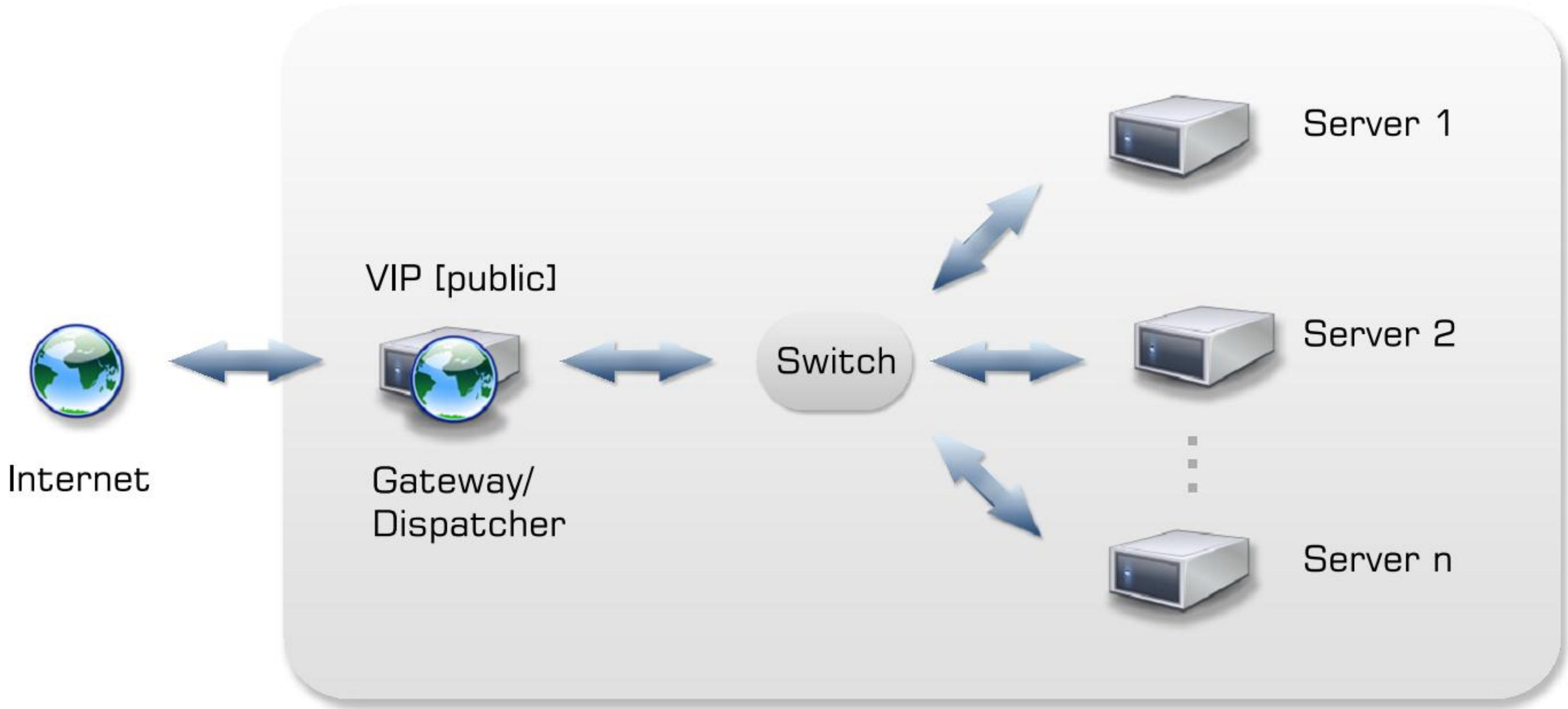
University of Potsdam

Universität Potsdam Mathematisch-Naturwissenschaftliche Fakultät

Outline

- Motivation
- Energy Saving Daemon (CHERUB)
- Cluster Simulator (ClusterSim)
- Parameter Study
- Conclusion

Components of a Server-Load-Balancing (SLB) Cluster



Universität Potsdam Mathematisch-Naturwissenschaftliche Fakultät

Outline

- Motivation
- Energy Saving Daemon (CHERUB)
- Cluster Simulator (ClusterSim)
- Parameter Study
- Conclusion

Motivation

- 30% of servers world-wide are *comatose* according to [1] (2015, Stanford) and [2] (2008, Uptime Institute)
- Corresponds to 4GW
  The most power full nuclear power plant block on earth generates 1.5GW

Source:
https://de.wikipedia.org/wiki/Datei:Chooz_Nuclear_Power_Plant-9361.jpg

Motivation

- Energy has become a critical resource in cluster designs
- Demand of energy is still permanently rising
- Strategies for saving energy:
  1. **Switch off unused resources**
  2. Virtualization
  3. Effective cooling
     (e.g. build your cluster in north Sweden like Facebook did)

Outline

- Motivation
- Energy Saving Daemon (CHERUB)
- Cluster Simulator (ClusterSim)
- Parameter Study
- Conclusion

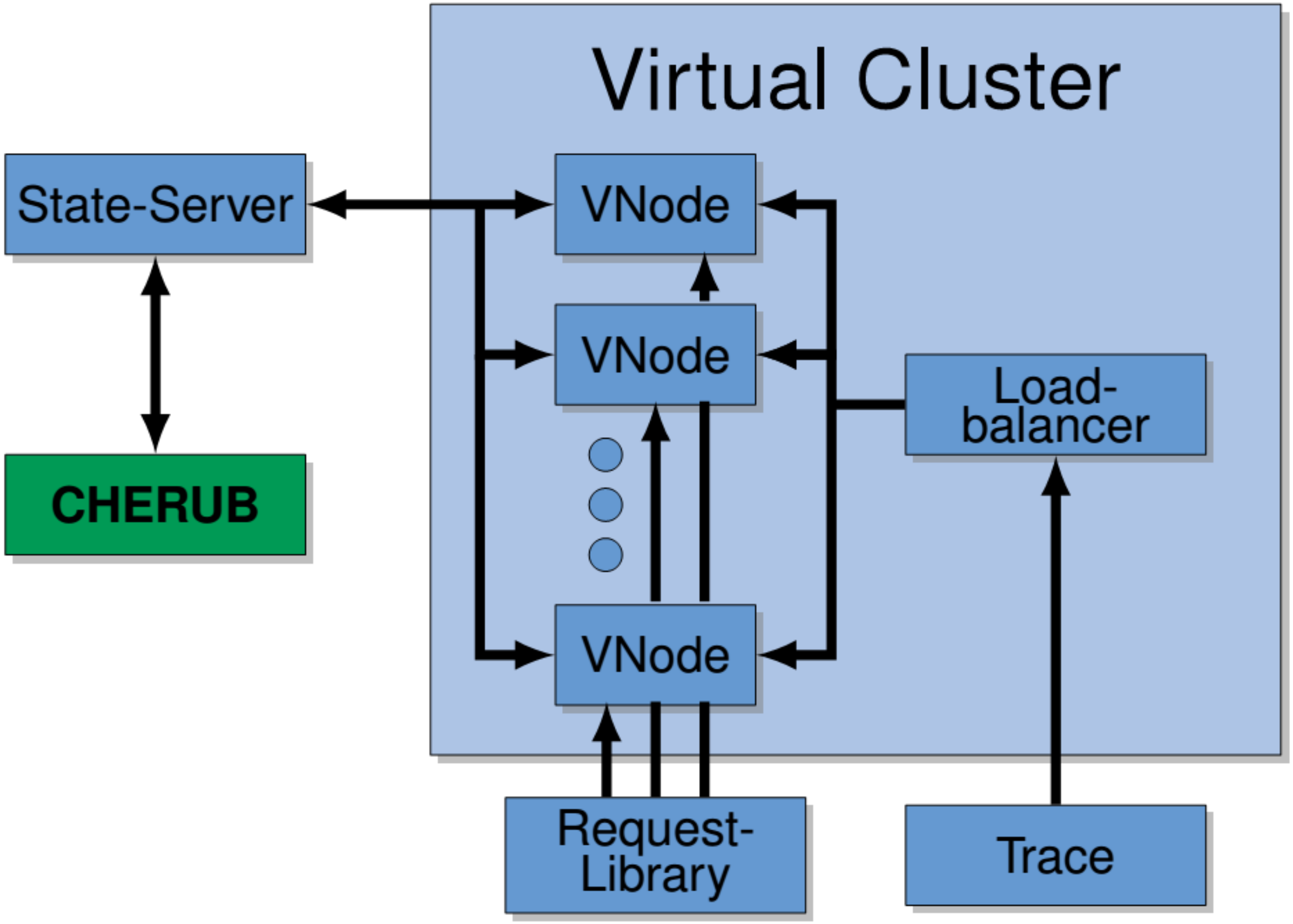
Cherubs functionality

- Centralized approach - no clients on back-ends
- Daemon located at master node polls the system in fixed time intervals to analyze its state
  - Status of every node
  - Load situation
- Depending on the state of the nodes, saved attributes and the load prediction, actions are performed for every node
- Online system - we don’t need any information about future load

Outline

- Motivation
- Energy Saving Daemon (CHERUB)
- Cluster Simulator (ClusterSim)
- Parameter Study
- Conclusion

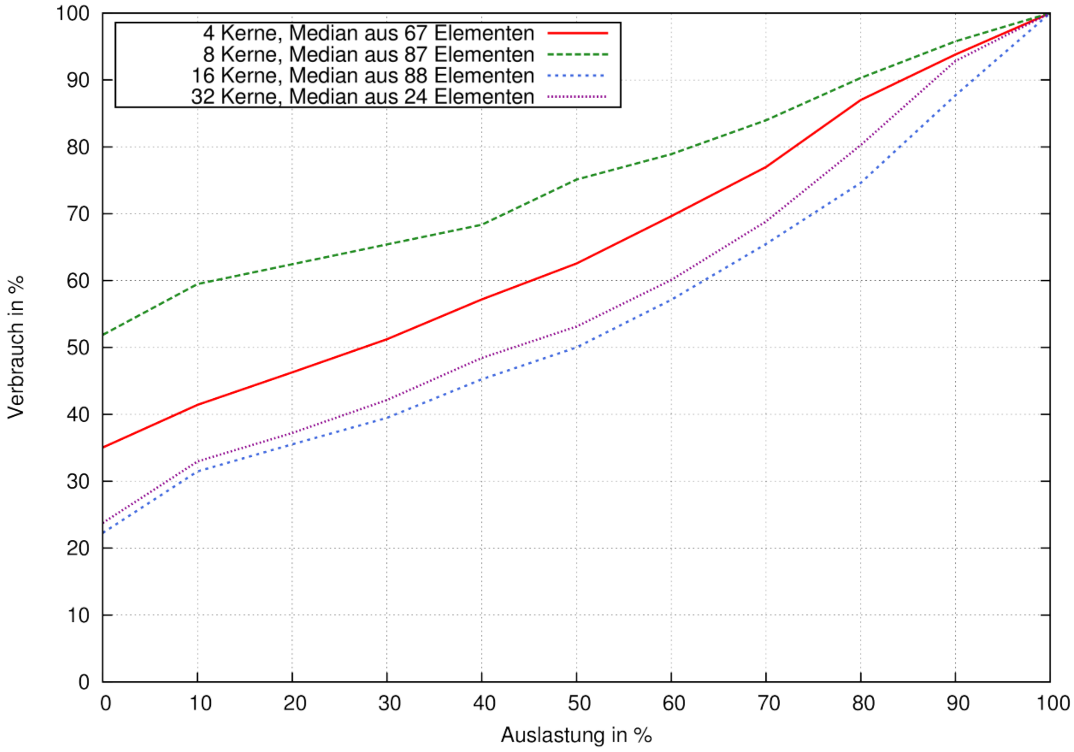
Simulation - ClusterSim Architecture



Scalability Evaluation of an Energy-Aware Resource Management System for Clusters of Web Servers
Kiertscher, Schnor
International Symposium on Performance Evaluation of Computer and Telecommunication Systems (SPECTS), Chicago, USA, July 2015

Simulation - Energy accounting

- Using real data from SPECpower_ssj 2008 Benchmark (Systems from 2007-2015)
- No data about STR, Boot or Shutdown consumption



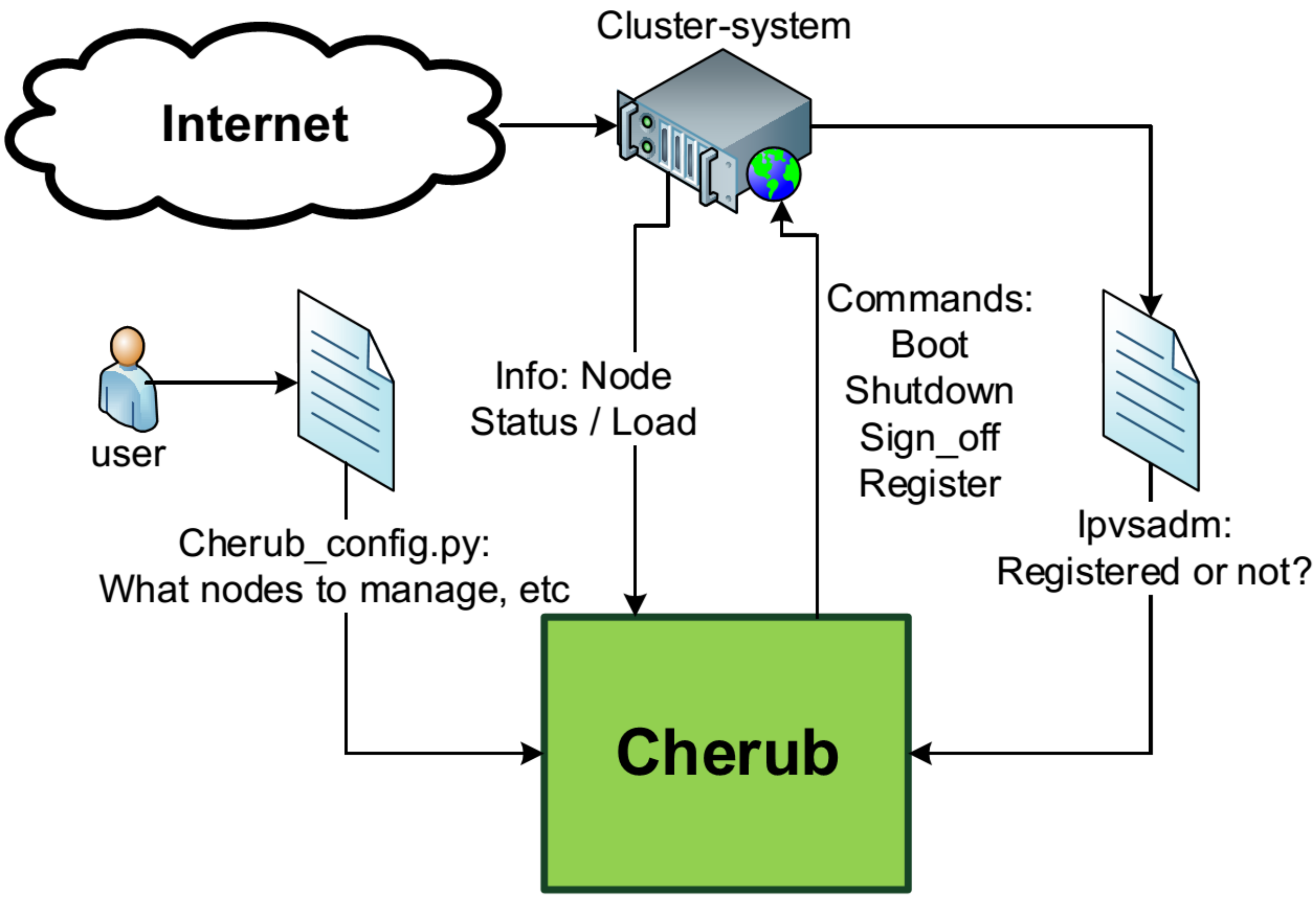
ClusterSim - Features

- Round-robin scheduler with 100ms time quantum
- Simulation of the Apache MPM Modules
- Bulk arrivals and TCP-Backlog Queue (BLQ) checks every second
(no typical discrete event driven simulation)
- Energy modeled based on utilization and real data

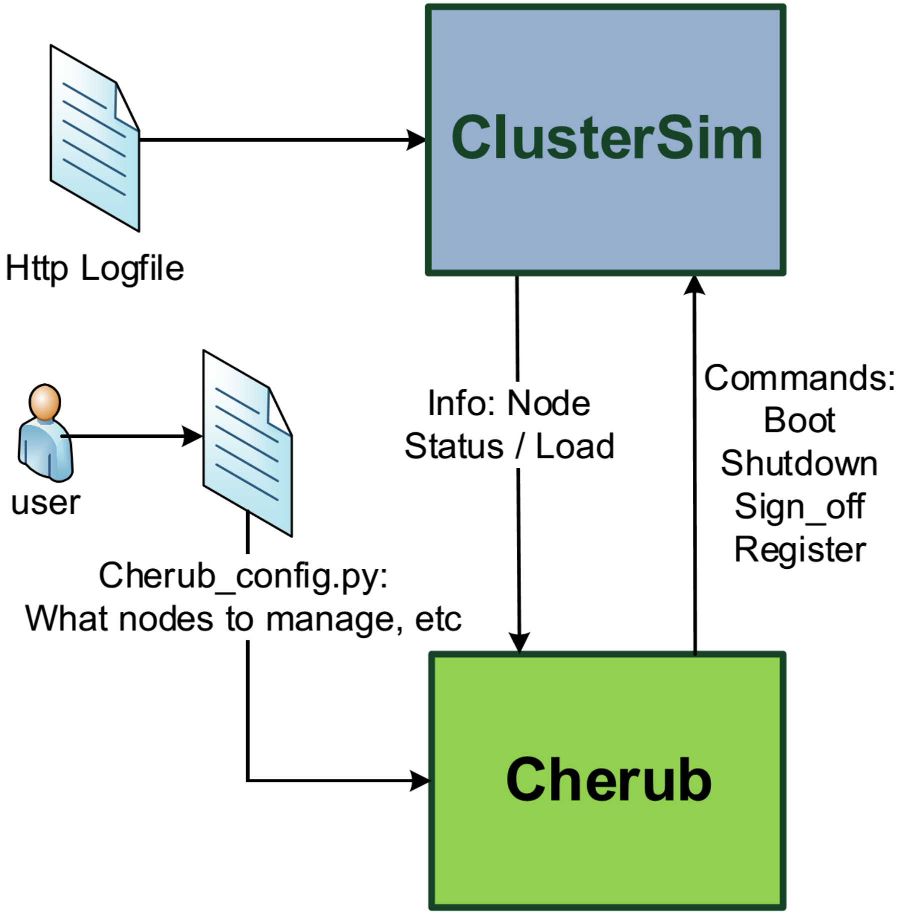
ClusterSim - Missing Features

- No modeling of system noise (easy to integrate)
- No concurrent resource access

Normal Setup



Simulation Setup



Outline

- Motivation
- Energy Saving Daemon (CHERUB)
- Cluster Simulator (ClusterSim)
- Parameter Study
- Conclusion

Used Metrics

- Quality of Service (QoS) in %
  using a 5 second timeout
- Request duration (RD) in milliseconds
  including waiting and processing time
- Energy consumption (EC) in Wh / Energy saved (ES) in %
- Number of physical state changes (PSCs) defined as the process to either turn on or turn off a node
- Score, a weighted ranking of the other 4 metrics

Score Strategies

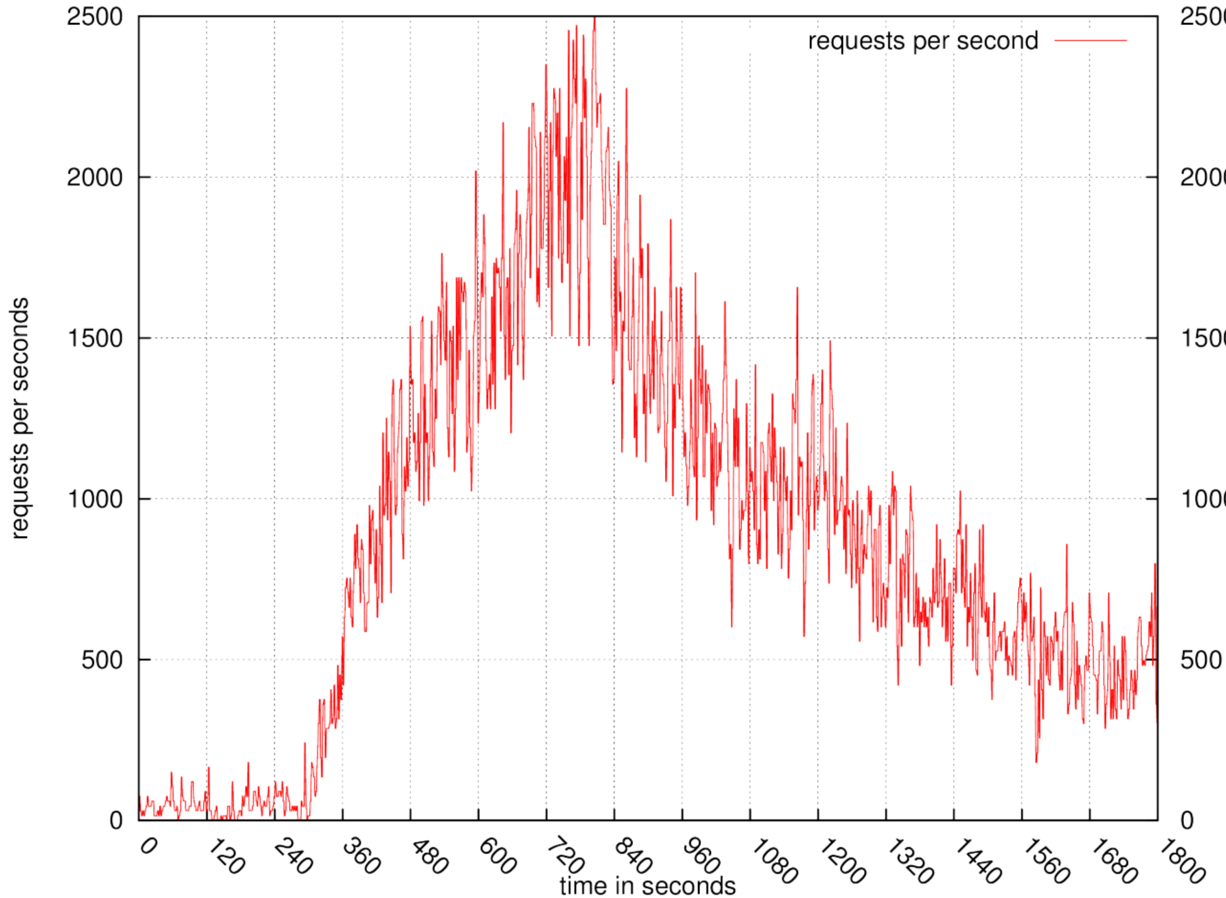
- Scores weighted ranking is done according to 3 different strategies
  - High Performance Provider (HP) prioritizes QoS and RD
  - Low Cost Provider (LC) prioritizes EC/ES and PSC
  - Balanced System (B) prioritizes QoS and EC/ES

Varied Factors

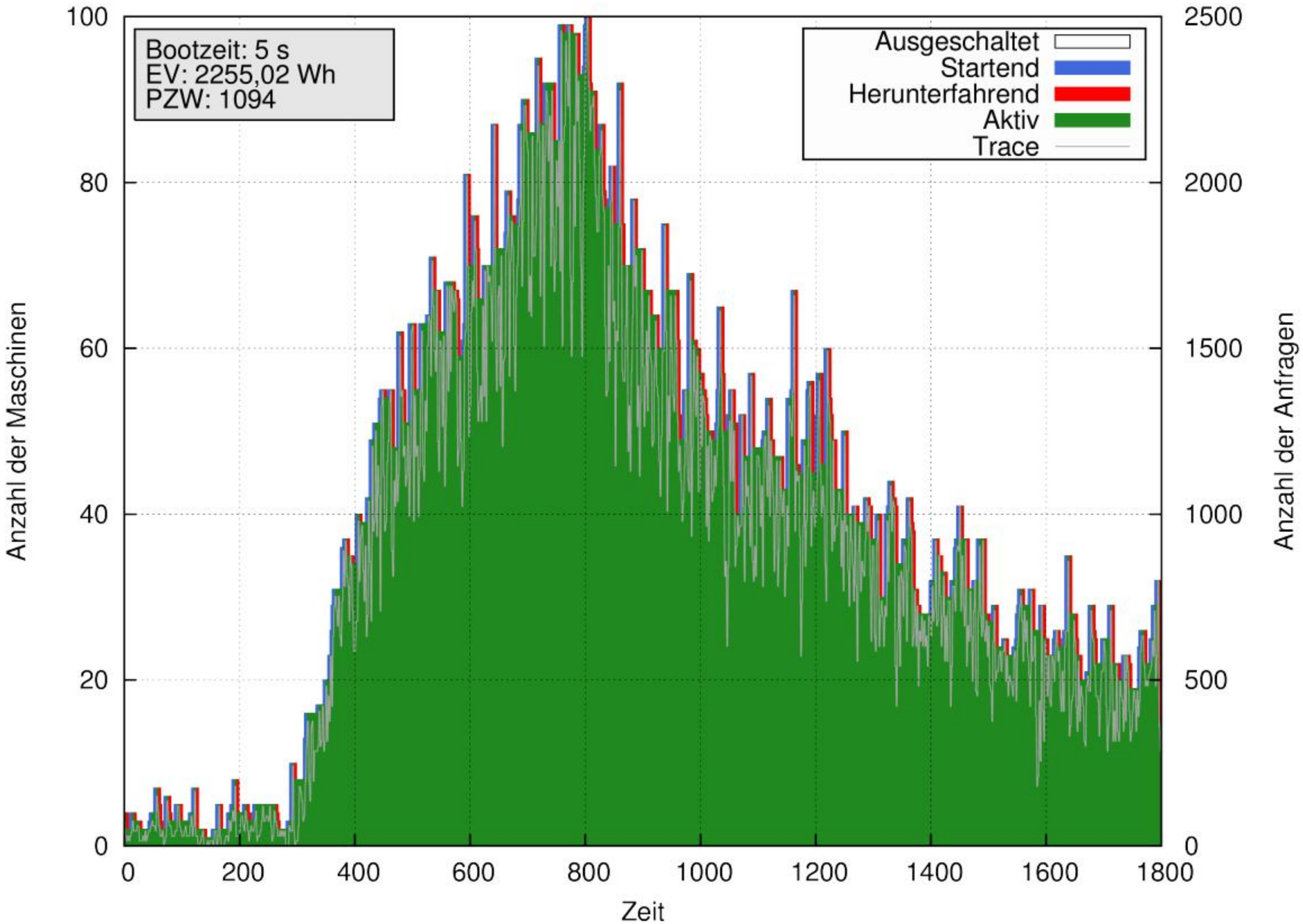
- Boot duration of the nodes:
  5, 30, 60, 120, 180 seconds
- Used backup:
  0, 5, 10, 25 %
- Shutdown strategy:
  aggressive or one-by-one
- Explicit wait before boot:
  0 or 1 minute

Used Workload and Optimum regarding EC/ES

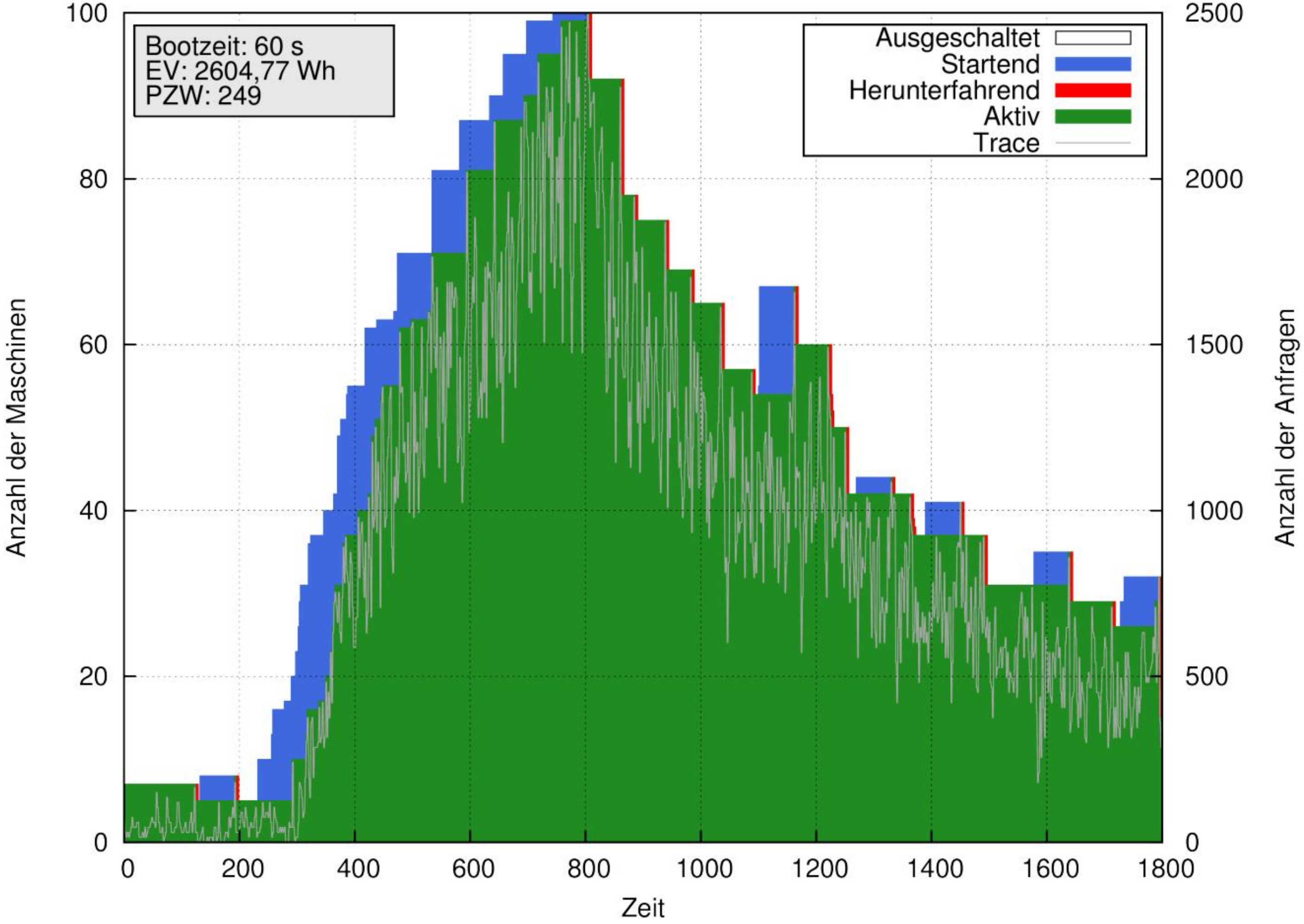
- Peak load situation → Worst Case
- Derived from real trace



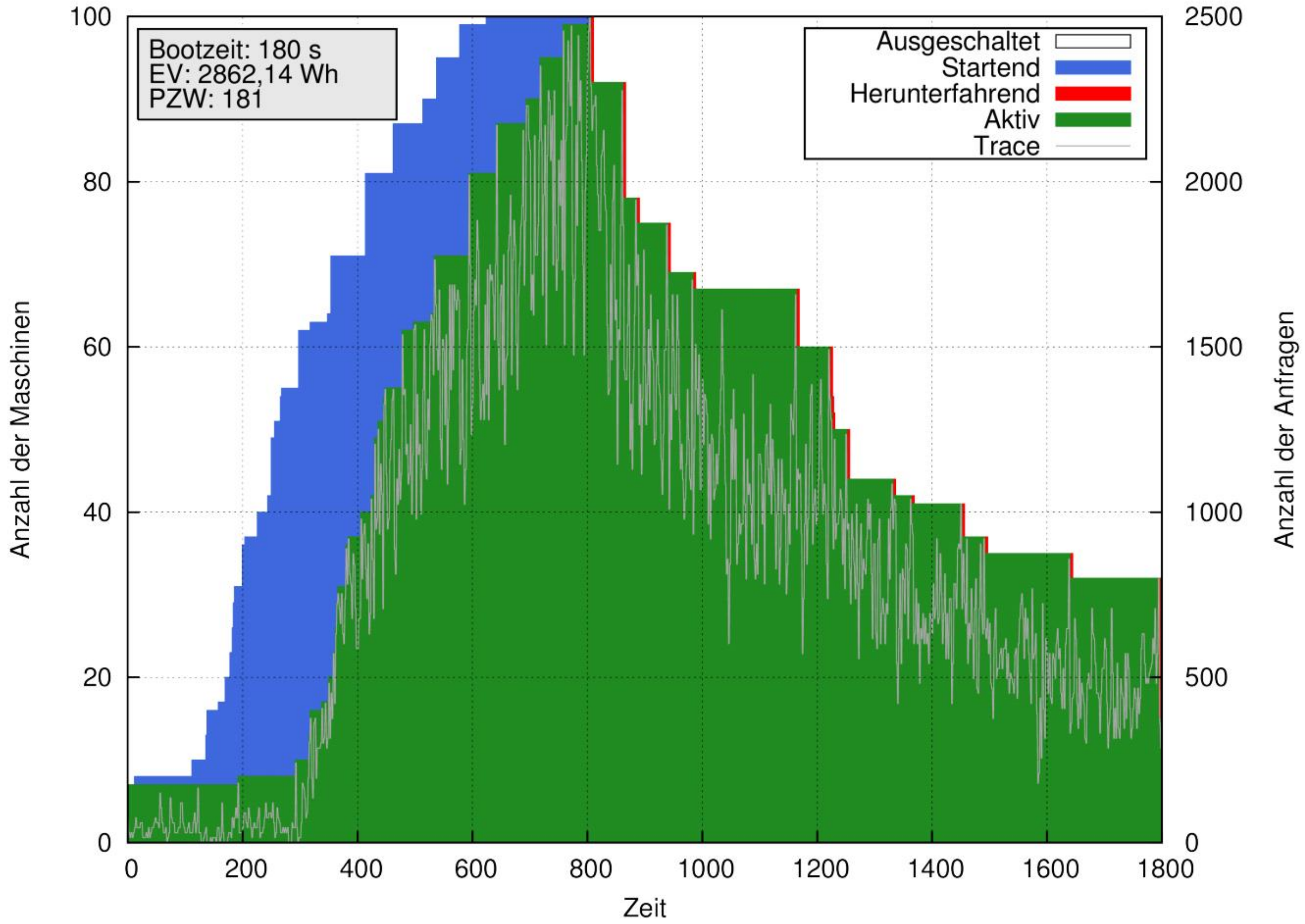
Used Workload and Optimum (5 seconds boottime)



Used Workload and Optimum (60 seconds boottime)



Used Workload and Optimum (3 minutes boottime)

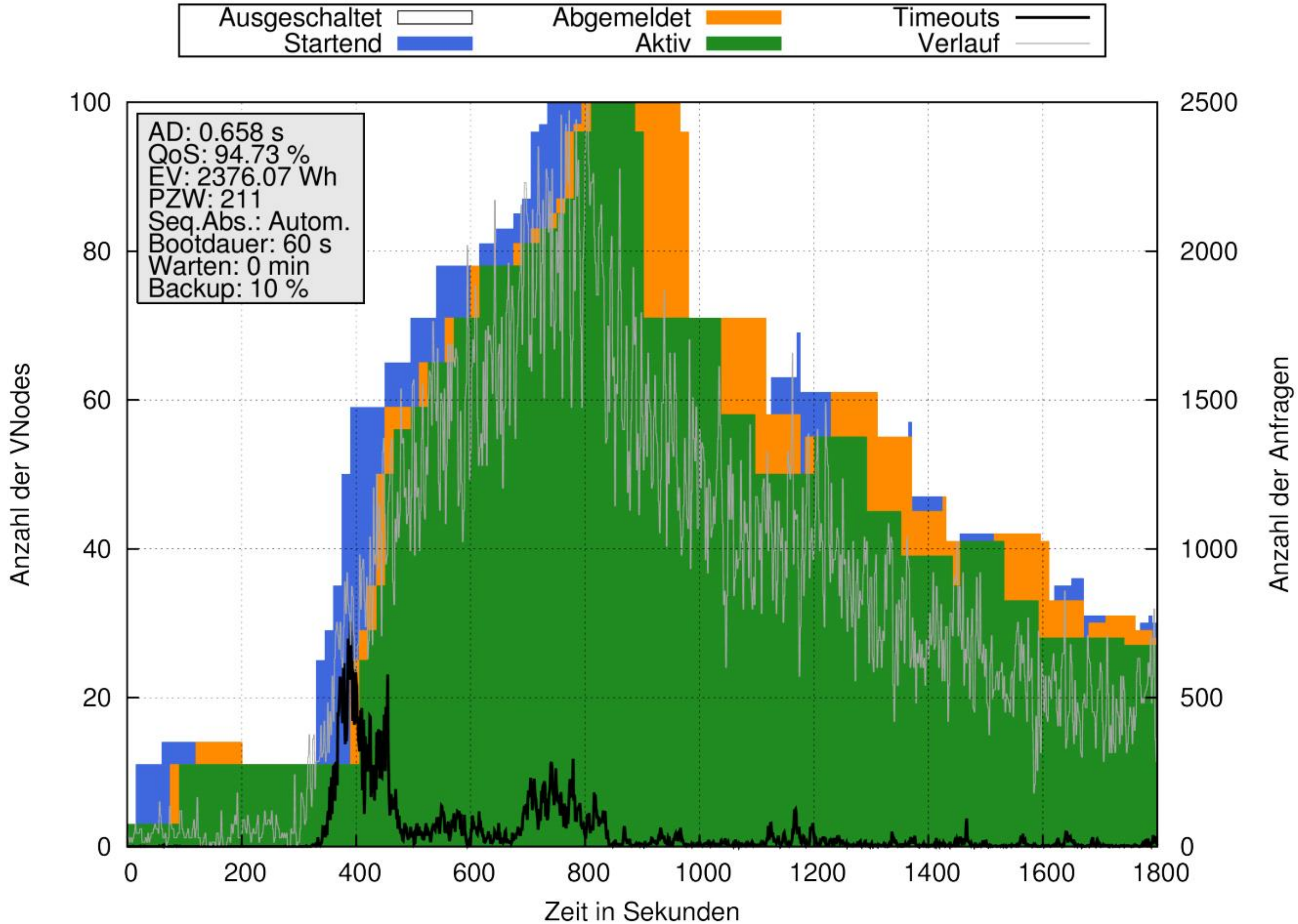


Reference without CHERUB

- All nodes active and running achieves QoS 98,67 % and EC of 3214,5 Wh

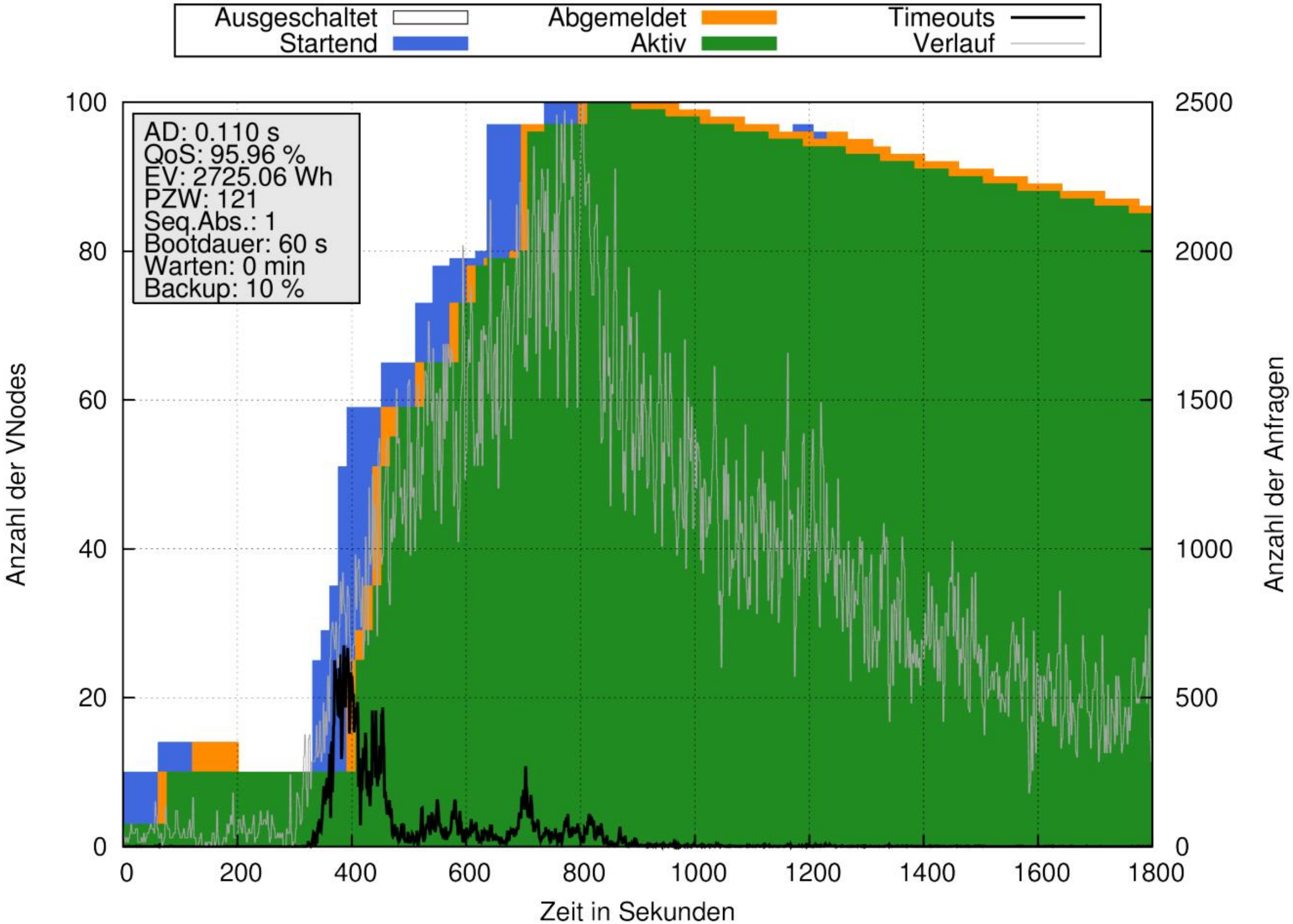
	QoS	EC in Wh	ES in %	PZW
Reference	98,67 %	3214,5	-	-
Optimum 5	-	2255,0	29,9	1094
Optimum 30	-	2477,3	22,9	404
Optimum 60	-	2604,8	19,0	249
Optimum 120	-	2734,2	14,9	215
Optimum 180	-	2862,1	11,0	181

One out of 80



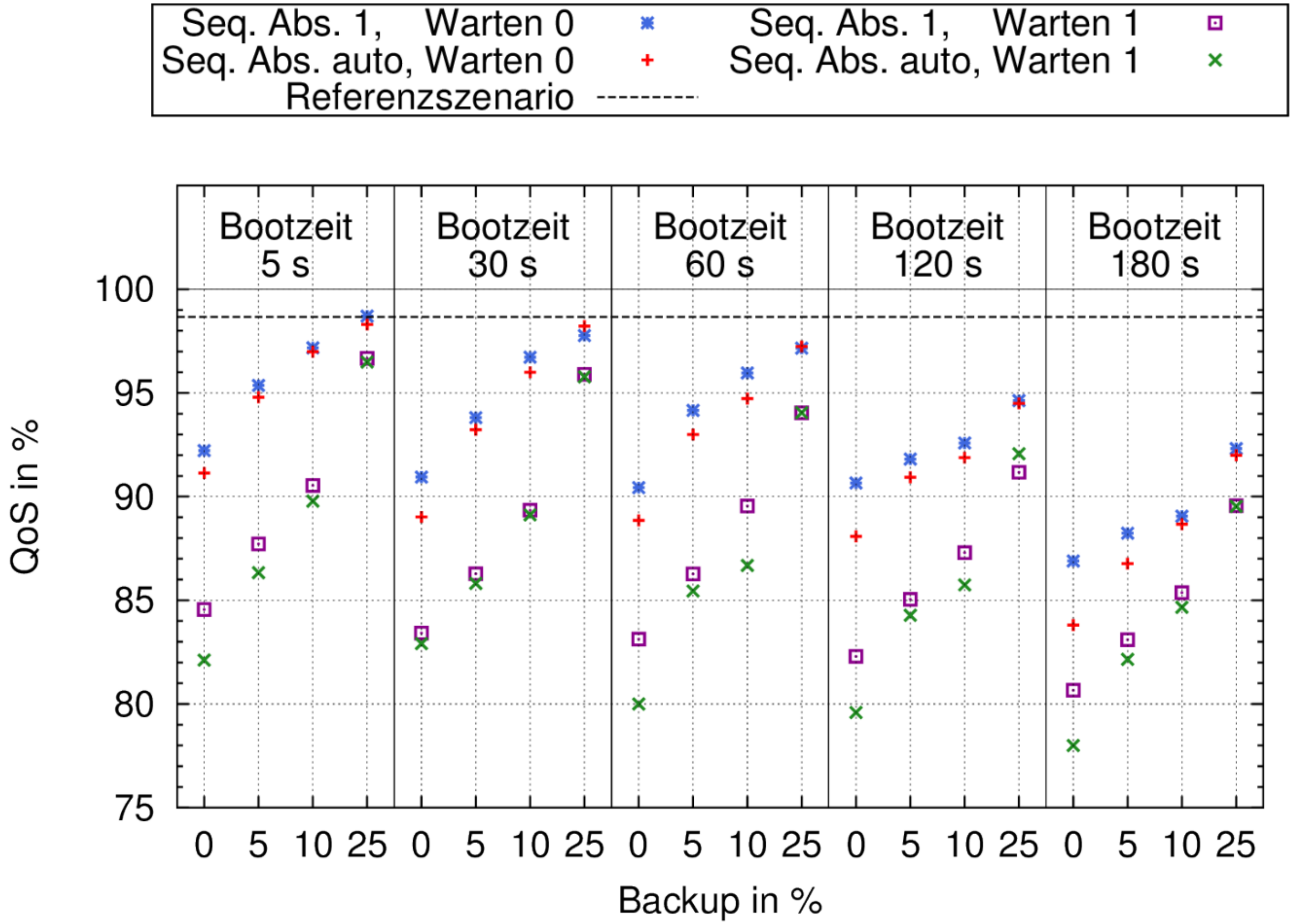
Ausgeschaltet
Startend
Abgemeldet
Aktiv
Timeouts
Verlauf
AD: 0.658 s
QoS: 94.73 %
EV: 2376.07 Wh
PZW: 211
Seq.Abs.: Autom.
Bootdauer: 60 s
Warten: 0 min
Backup: 10 %
Anzahl der VNodes
Anzahl der Anfragen
Zeit in Sekunden
0
20
40
60
80
100
500
1000
1500
2000
2500
200
400
600
800
1000
1200
1400
1600
1800

One out of 80



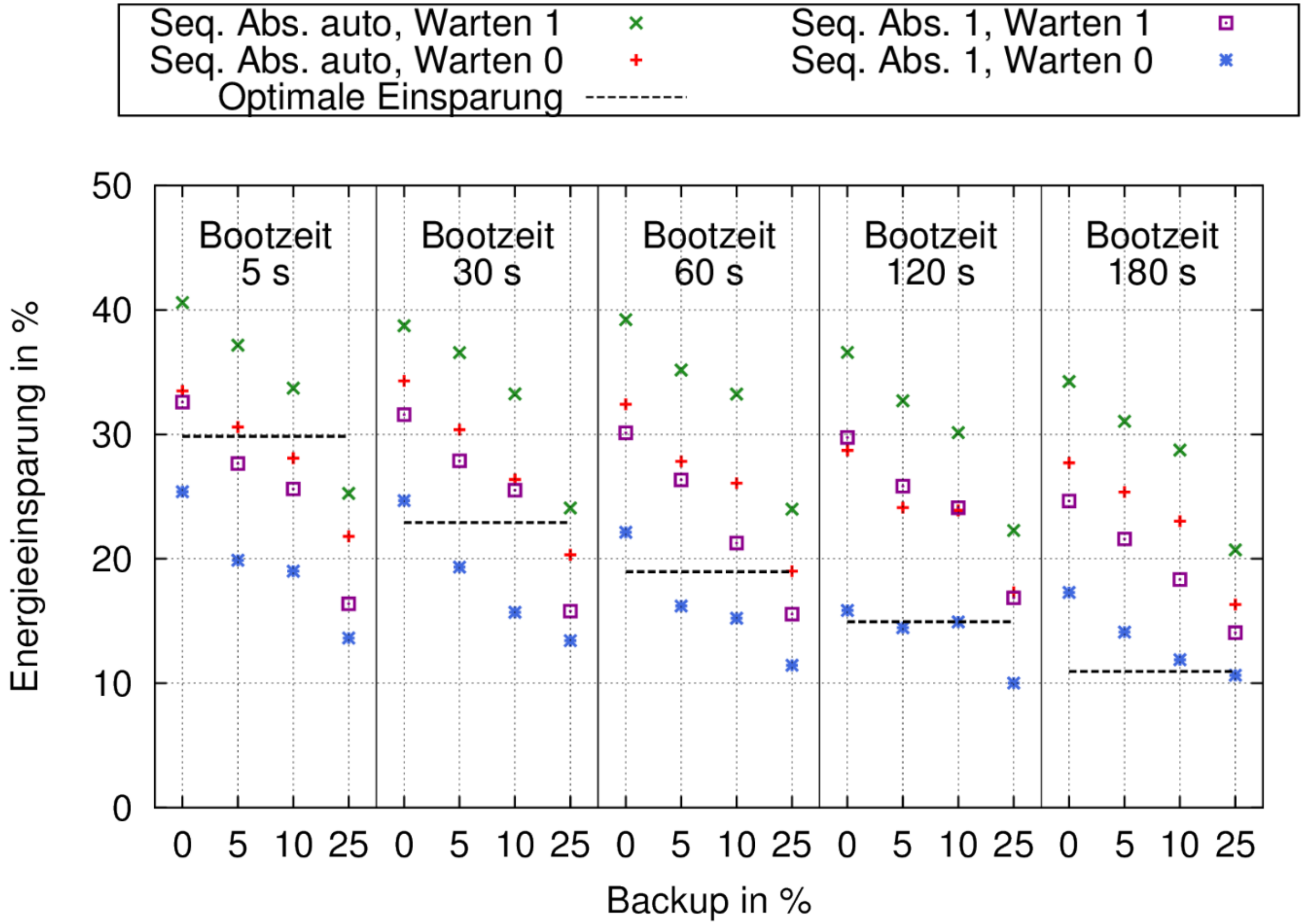
Ausgeschaltet
Startend
Abgemeldet
Aktiv
Timeouts
Verlauf
AD: 0.110 s
QoS: 95.96 %
EV: 2725.06 Wh
PZW: 121
Seq.Abs.: 1
Bootdauer: 60 s
Warten: 0 min
Backup: 10 %
Anzahl der VNodes
Anzahl der Anfragen
Zeit in Sekunden
0
20
40
60
80
100
500
1000
1500
2000
2500
200
400
600
800
1000
1200
1400
1600
1800

Results QoS

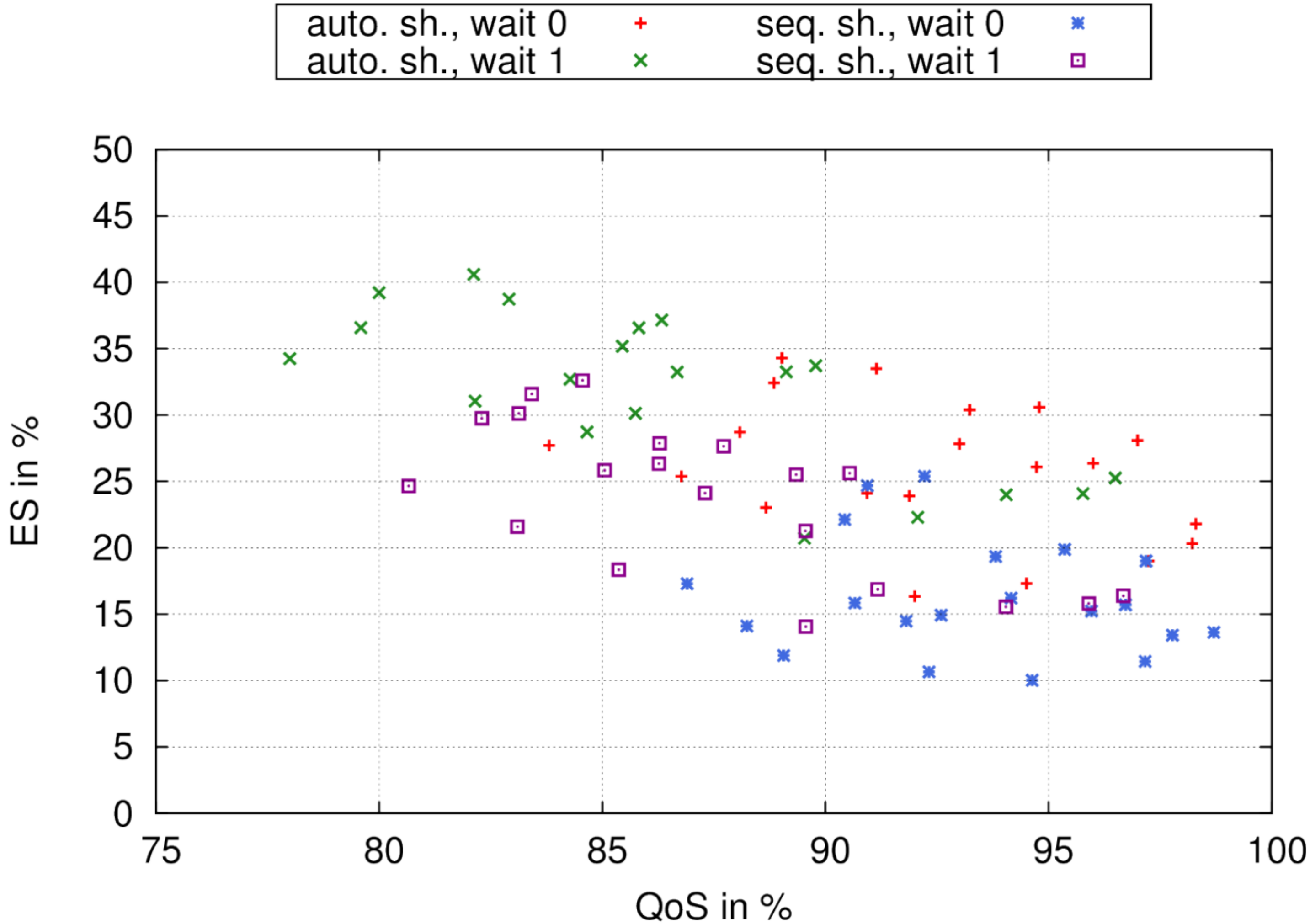


Seq. Abs. 1, Warten 0
Seq. Abs. auto, Warten 0
Referenzszenario
Seq. Abs. 1, Warten 1
Seq. Abs. auto, Warten 1
Bootzeit 5 s
Bootzeit 30 s
Bootzeit 60 s
Bootzeit 120 s
Bootzeit 180 s
QoS in %
100
95
90
85
80
75
0 5 10 25 0 5 10 25 0 5 10 25 0 5 10 25 0 5 10 25
Backup in %

Results Energy Saving



ES vs QoS



auto. sh., wait 0
auto. sh., wait 1
seq. sh., wait 0
seq. sh., wait 1
ES in %
QoS in %
0
5
10
15
20
25
30
35
40
45
50
75
80
85
90
95
100

Score - Parameter Settings

strategy	boot duration in s	score	shutdown	wait in min	backup in %
HP	5	70.3	seq.	0	25
	30	66.5	seq.	0	25
	60	63.6	seq.	0	25
	120	56.9	seq.	0	25
	180	53.1	seq.	0	25
LC	5	61.0	seq.	1	0
	30	60.5	seq.	1	0
	60	56.3	seq.	1	0
	120	49.2	seq.	1	0
	180	43.1	auto.	1	0
B	5	55.4	auto.	0	10
	30	50.8	auto.	0	10
	60	47.8	auto.	0	10
	120	41.1	auto.	1	25
	180	33.9	auto.	1	10

Score - Results

strategy	boot duration in s	score	QoS in %	ES in %	Optimum Saving	RD in ms	PSCs	shutdown	wait in min	backup in %
HP	5	70.3	98.70	13.63	29.9	67	123	seq.	0	25
	30	66.5	97.77	13.42	22.9	79	135	seq.	0	25
	60	63.6	97.16	11.44	19.0	78	159	seq.	0	25
	120	56.9	94.63	10.02	14.9	81	156	seq.	0	25
	180	53.1	92.32	10.64	11.0	80	148	seq.	0	25
LC	5	61.0	84.56	32.60	29.9	548	94	seq.	1	0
	30	60.5	83.42	31.58	22.9	430	89	seq.	1	0
	60	56.3	83.13	30.12	19.0	401	105	seq.	1	0
	120	49.2	82.30	29.76	14.9	527	124	seq.	1	0
	180	43.1	78.00	34.25	11.0	1452	180	auto.	1	0
B	5	55.4	96.99	28.08	29.9	574	199	auto.	0	10
	30	50.8	96.00	26.37	22.9	619	219	auto.	0	10
	60	47.8	94.73	26.08	19.0	658	211	auto.	0	10
	120	41.1	92.07	22.29	14.9	211	174	auto.	1	25
	180	33.9	84.66	28.74	11.0	626	173	auto.	1	10

Score - Results

strategy	boot duration in s	score	QoS in %	ES in %	Optimum Saving	RD in ms	PSCs	shutdown	wait in min	backup in %
HP	5	70.3	98.70	13.63	29.9	67	123	seq.	0	25
	30	66.5	97.77	13.42	22.9	79	135	seq.	0	25
	60	63.6	97.16	11.44	19.0	78	159	seq.	0	25
	120	56.9	94.63	10.02	14.9	81	156	seq.	0	25
	180	53.1	92.32	10.64	11.0	80	148	seq.	0	25
LC	5	61.0	84.56	32.60	29.9	548	94	seq.	1	0
	30	60.5	83.42	31.58	22.9	430	89	seq.	1	0
	60	56.3	83.13	30.12	19.0	401	105	seq.	1	0
	120	49.2	82.30	29.76	14.9	527	124	seq.	1	0
	180	43.1	78.00	34.25	11.0	1452	180	auto.	1	0
B	5	55.4	96.99	28.08	29.9	574	199	auto.	0	10
	30	50.8	96.00	26.37	22.9	619	219	auto.	0	10
	60	47.8	94.73	26.08	19.0	658	211	auto.	0	10
	120	41.1	92.07	22.29	14.9	211	174	auto.	1	25
	180	33.9	84.66	28.74	11.0	626	173	auto.	1	10

Score - Results

strategy	boot duration in s	score	QoS in %	ES in %	Optimum Saving	RD in ms	PSCs	PSCs in Optimum Case	shutdown	wait in min	backup in %
HP	5	70.3	98.70	13.63	29.9	67	123	1094	seq.	0	25
	30	66.5	97.77	13.42	22.9	79	135	404	seq.	0	25
	60	63.6	97.16	11.44	19.0	78	159	249	seq.	0	25
	120	56.9	94.63	10.02	14.9	81	156	215	seq.	0	25
	180	53.1	92.32	10.64	11.0	80	148	181	seq.	0	25
LC	5	61.0	84.56	32.60	29.9	548	94	1094	seq.	1	0
	30	60.5	83.42	31.58	22.9	430	89	404	seq.	1	0
	60	56.3	83.13	30.12	19.0	401	105	249	seq.	1	0
	120	49.2	82.30	29.76	14.9	527	124	215	seq.	1	0
	180	43.1	78.00	34.25	11.0	1452	180	181	auto.	1	0
B	5	55.4	96.99	28.08	29.9	574	199	1094	auto.	0	10
	30	50.8	96.00	26.37	22.9	619	219	404	auto.	0	10
	60	47.8	94.73	26.08	19.0	658	211	249	auto.	0	10
	120	41.1	92.07	22.29	14.9	211	174	215	auto.	1	25
	180	33.9	84.66	28.74	11.0	626	173	181	auto.	1	10

Outline

- Motivation
- Energy Saving Daemon (CHERUB)
- Cluster Simulator (ClusterSim)
- Parameter Study
- Conclusion

Conclusion

- Strategy works in 100 node SLB-Setup
- Results are very close to the optimum (with fast hardware)
- Boot duration is a critical factor
- Backup has a linear influence on QoS and EC
- Aggressive shutdown can save up to 12,9 % extra energy in the peak load scenario
- Extra waiting time is not necessary if load forecasting is used

Thank you for your attention!
Any Questions?

Contact:

kiertscher@cs.uni-potsdam.de

www.cs.uni-potsdam.de

Sources

[1] “New data supports finding that 30 percent of servers are ‘Comatose’, indicating that nearly a third of capital in enterprise data centers is wasted” by Jonathan Koomey and Jon Taylor, 2015

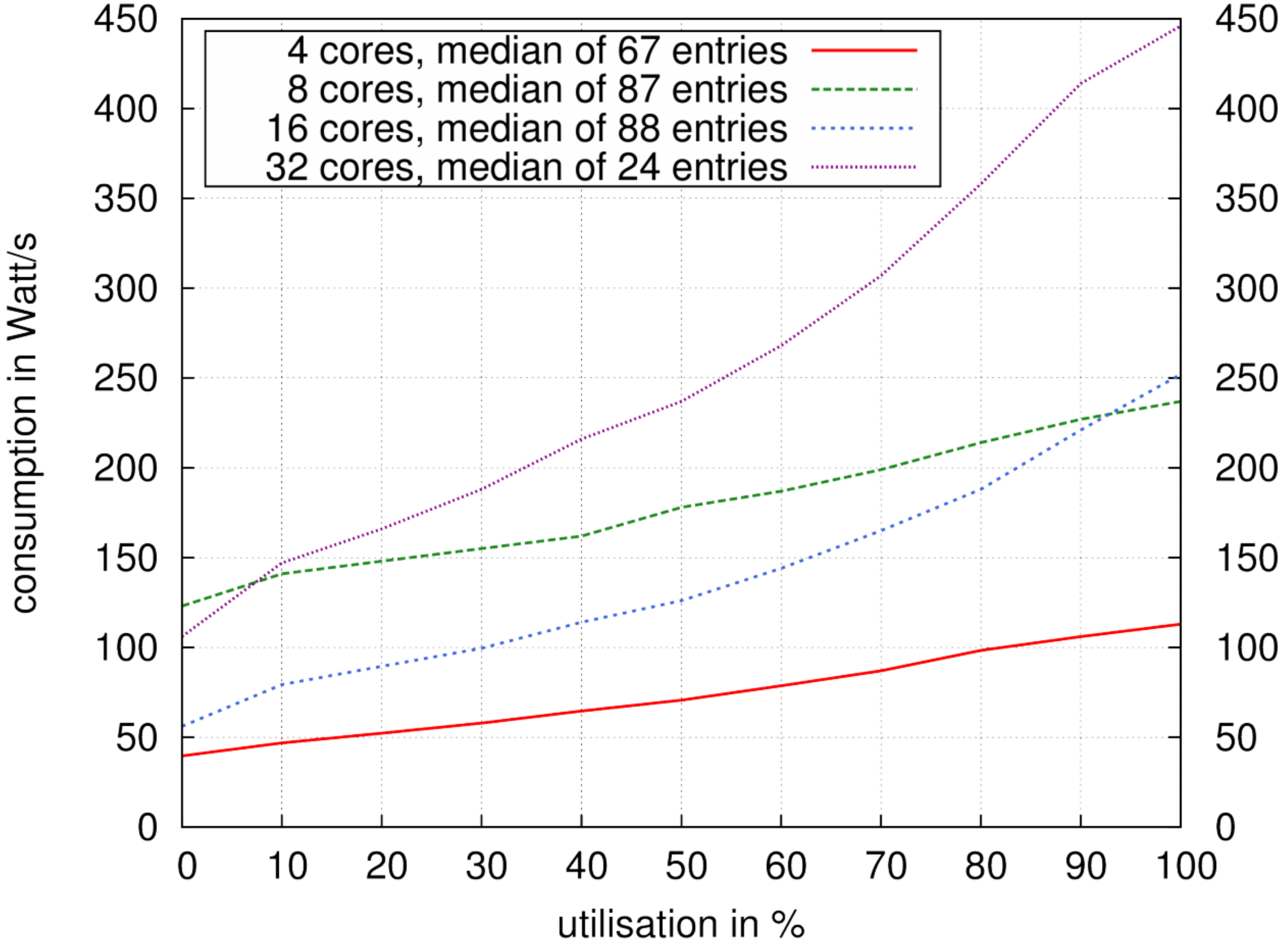
[2] “Revolutionizing Data Center Energy Efficiency” by James Kaplan, William Forrest, Noah Kindler, 2008

Backup

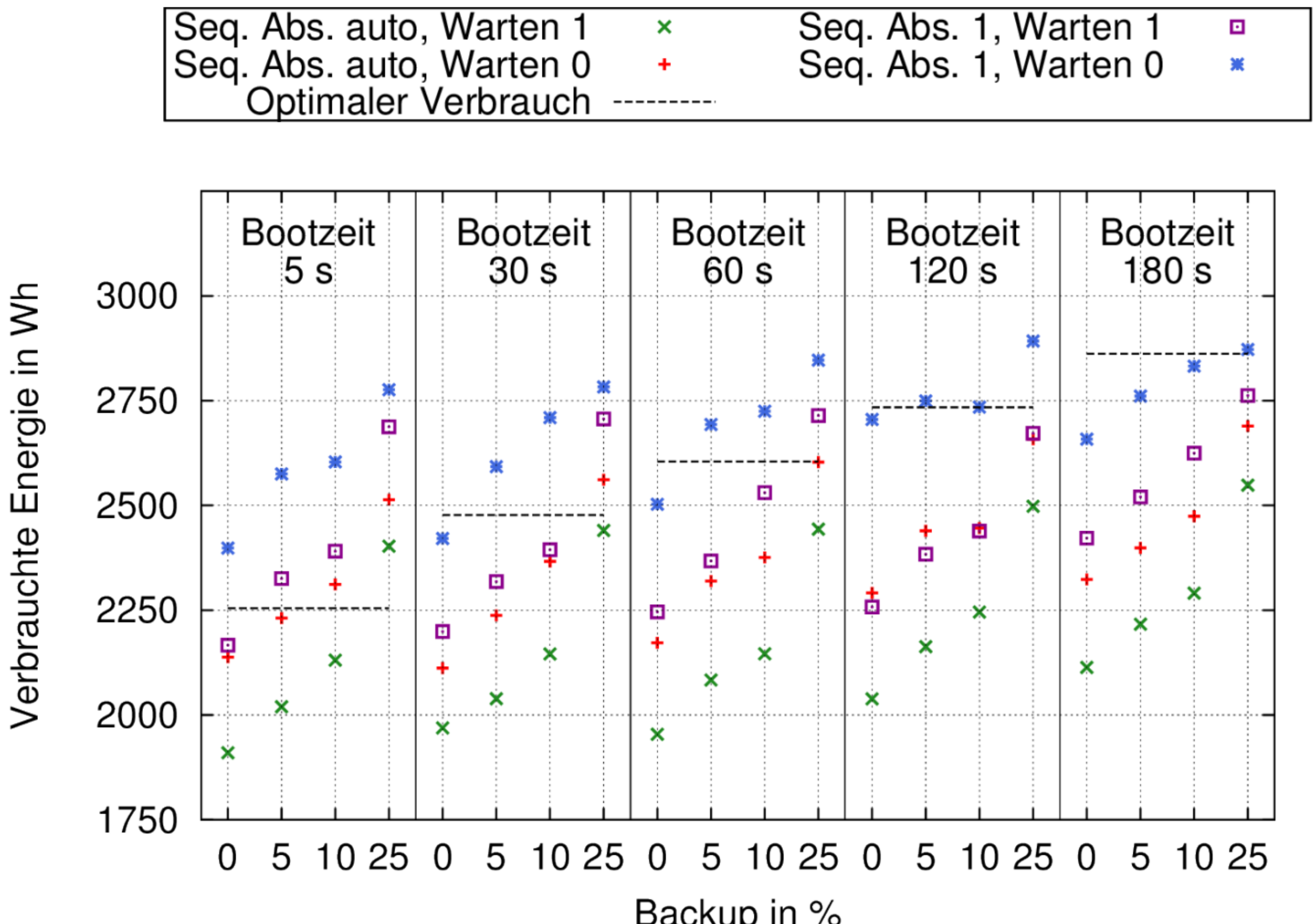
Cloud ?

- Smallest dedicated hardware instance at Amazon (Linux on m4.large Dedicated) costs 5747 USD for on month (50 % utilization)

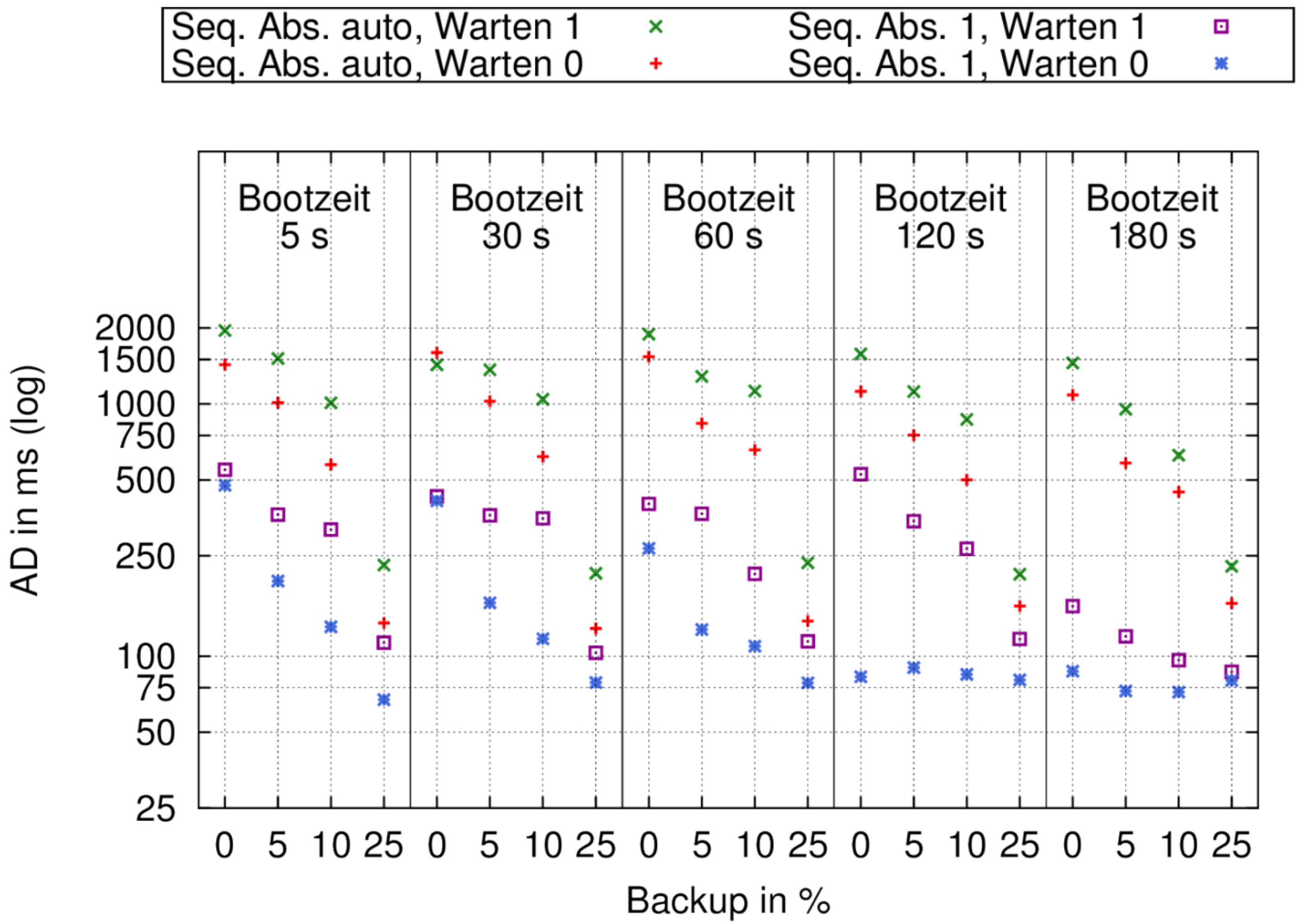
Consumtion in Watt/s



EC


Seq. Abs. auto, Warten 1
Seq. Abs. auto, Warten 0
Optimaler Verbrauch
Seq. Abs. 1, Warten 1
Seq. Abs. 1, Warten 0
Bootzeit 5 s
Bootzeit 30 s
Bootzeit 60 s
Bootzeit 120 s
Bootzeit 180 s
Verbrauchte Energie in Wh
3000
2750
2500
2250
2000
1750
0 5 10 25 0 5 10 25 0 5 10 25 0 5 10 25 0 5 10 25
Backup in %

FRT



Seq. Abs. auto, Warten 1
Seq. Abs. auto, Warten 0
Seq. Abs. 1, Warten 1
Seq. Abs. 1, Warten 0
Bootzeit 5 s
Bootzeit 30 s
Bootzeit 60 s
Bootzeit 120 s
Bootzeit 180 s
AD in ms (log)
2000
1500
1000
750
500
250
100
75
50
25
0 5 10 25 0 5 10 25 0 5 10 25 0 5 10 25 0 5 10 25
Backup in %

Mathematisch-
Naturwissenschaftliche Fakultät

PSC