



Methoden zur Beschreibung und Einordnung heterogener Hardware Infrastrukturen für Scientific Workflows

Robert Balink

robert.balink@campus.tu-berlin.de

16.Dezember 2021

Bachelorarbeit

Distributed and Operating Systems

Technische Universität Berlin

Erstprüfer: Prof. Dr. habil. Odej Kao

Betreuer: Jonathan Bader

Zweitprüfer: Prof. Dr. Dr. h.c. Sahin Albayrak

Eidestattliche Erklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig und eigenhändig sowie ohne unerlaubte fremde Hilfe und ausschließlich unter Verwendung der aufgeführten Quellen und Hilfsmittel angefertigt habe.

Die selbständige und eigenständige Anfertigung versichert an Eides statt:

Berlin, den 2.9.2021

TODO: Unterschrift

Anerkennung

I would like to thank the following persons for their great support when writing my thesis at the Mobile Cloud Computing Research Group (MCC) at Technische Universität Berlin.

INITIALS

Zusammenfassung

Abstract Style

Inhaltsverzeichnis

Eidestattliche Erklärung	I
Anerkennung	III
Zusammenfassung	V
List of Figures	IX
List of Tables	X
List of Listings	XI
1 Einführung	1
1.1 Problematik	1
1.2 Ziel	2
1.3 Interessante Latex Möglichkeiten	2
2 Hintergrund	5
2.1 Scientific Workflows	5
2.2 Computer Architektur	5

2.2.1	Aufbau	5
2.2.2	Prozessor	6
2.2.3	Arbeitsspeicher	7
2.2.4	Sekundärspeicher	9
2.2.5	Weitere Ein und Ausgabegeräte	9
2.2.6	Zusammenspiel der Komponenten für die Performance	10
2.3	Metriken	11
2.3.1	Anforderungen an eine Metrik	11
2.3.2	Historische Metriken	12
2.3.3	Industrie Benchmarks	12
2.3.4	openSource Benchmarks Metriken	12
2.3.5	Fehlerquellen	12
3	Methodik	13
3.1	Anforderungen an das Programm	13
3.2	Online-Datenbanken	14
3.2.1	cpu-world	14
3.2.2	userBenchmark	14
3.2.3	openBenchmarking	14
3.3	Berechnung des Scores	14

4	Auswertung	15
4.1	Validierung der Ergebnisse	15
4.2	Abweichungen	15
4.3	Klassifizierung der Ergebnisse	15
5	Stand der Forschung	17
6	Diskussion	19
7	Fazit	21
A	Test	23
B	Abkürzungsverzeichnis	25
C	Lexikon	27
D	Listings	29
D.1	Configuration for Node A	29
	Literaturverzeichnis	31

List of Figures

1.1	Example Setup of Three Machines Agreeing on Values Based on the Paxos	
	Algorithm	3

List of Tables

1.1	Influence of Operations on a Data Item's Version Vector	2
1.2	Selection of Possible Devices in a Smart Home	3

List of Listings

D.1 Configuration for Node A	29
--	----

Chapter 1

Einführung

!Einführungstext eventuell!

1.1 Problematik

Für das Ausführen wissenschaftlicher Workflow Management Systeme, wie beispielsweise nextflow.io laufen verschiedene Arbeitspakete auf unterschiedlichen Computern. Die Arbeitspakete können allerdings zeitintensiv sein, da sie viele Ressourcen benötigen. Ohne das Wissen, welche Computereinheit leistungsstärker ist und diese dementsprechend nicht priorisieren zu können welcher Computer genutzt werden soll, k Für den Anwender wäre es daher von Vorteil zu wissen, welcher Computer von Anfang an Leistungsfähiger ist. Durch dieses Wissen spart der Anwender sehr viel Zeit und kann effektiver die Arbeitspakete aufteilen.

1.2 Ziel

Das Ziel dieser Abschlussarbeit ist es, die Performance relevanten und kritischen Bereiche des Computers ausfindig zu machen, die gängigen wissenschaftlichen Performance Metriken zu ermitteln und daraufhin ein eigenes Kommandozeilen Tool zu schreiben, welches die Werte aus dem Computer ausliest, analysiert und die Performance des Computers einschätzt in Relation zu einem Referenz Computer. Am Ende sollen die berechneten Werte validiert und gruppiert werden.

1.3 Interessante Latex Möglichkeiten

This chapter gives a general introduction into the content of and the motivation for this thesis. This is done by initially outlining the motivation and problem in section 1.1, defining the goal in section 1.2, and explaining the structure of the thesis in section 1.3.

Operation related to data item Consequence for vector of data item	
Update on site S_i	Increment v_i by one
Delete or rename on site S_i	Keep vector and increment v_i by one, remove data item value
Reconcile version conflict	Set each v_i to maximum v_i from all vectors used for reconciliation. In addition, increment v_i of site that initiated reconciliation by one
Copy to new site	Augment vector to include new site

Table 1.1: Influence of Operations on a Data Item's Version Vector

Thus, many executives feel fear to become unconscious Also, I want to buy **smart home devices** such as the ones listed in table 1.2.

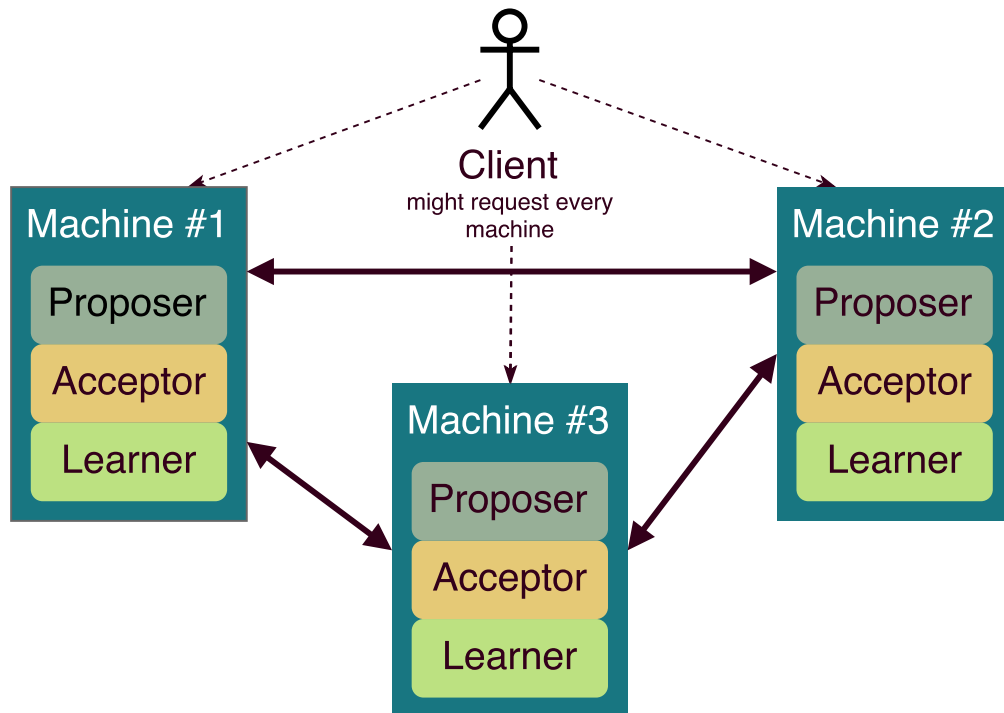


Figure 1.1: Example Setup of Three Machines Agreeing on Values Based on the Paxos Algorithm

Table 1.2: Selection of Possible Devices in a Smart Home

Device	Example Activity
Sensors	
Brightness sensor	Measure the current brightness to identify whether lights should be turned on.
Air quality sensor	Detect if a windows should be opened based on the current air quality.
Movement sensor	Detect movement in a room to determine if residents are present.
Temperature sensor	Measure the current temperature to detect if a heater should be turned on.
Contact sensor	Determine whether a window or door is open.

Continued on next page

Table 1.2 – continued from previous page

Device	Example Activity
Visual sensor	Count the number of people in a room to perform specific actions.
Actuators	
Outlet	Turn on/off connected devices.
Thermostat	Increase/decrease the room temperature.
Motors	Open/close windows or curtains.
Mail-Server	Send notifications to residents based on recorded data.
Infrared emitter	Control devices (e.g., a TV) that are not fully integrated into the smart home system.

1. Analysis ...
2. Design ...
3. Prototypical implementation of the design.

The **client**

Chapter 2

Hintergrund

2.1 Scientific Workflows

2.2 Computer Architektur

2.2.1 Aufbau

Ein moderner Personal Computer (PC) besteht aus dem Zusammenspiel von Prozessor, Arbeitsspeicher, sowie einiger Ein und Ausgabegeräten, welche untereinander mit einem Bus-System verbunden ist. Dieser Aufbau entspricht der Von-Neumann-Architektur, welche aus den 1940er Jahren stammt und vom ungarisch-amerikanischen Mathematiker John von Neumann entwickelt wurde. In den folgenden Unterkapiteln wird näher auf die einzelnen Bestandteile des Computers eingegangen. Allen voran dem Prozessor sowie dem Arbeitsspeicher. Unter dem Sammelbegriff Ein und Ausgabegeräte versteht man hierbei alle Komponenten, die nicht unmittelbar zur Von-Neumann-Architektur gehören. In den folgenden Unterkapiteln wird dabei vor allem auf die Sekundärspeicher eingegangen, die eine wichtige Rolle spielen für die Performance sowie ein kurzer Überblick gegeben über die weitere typ-

ische Ein und Ausgabegeräte, die allerdings nicht von wesentlicher Bedeutung sind für die Performance. ISA Bus Industry Standard Architecture PCI Bus Peripheral Component Interconnect

2.2.2 Prozessor

Das Herzstück des Computers bildet der Prozessor, auch CPU (Central processing unit) genannt. Der Prozessor kümmert sich um die Ausführung von Programmen, die er aus dem Arbeitsspeicher erhält. Die CPU besteht aus vier Funktionseinheiten. Dem Steuerwerk, dem Register, dem Rechenwerk und den Bussen. Das Steuerwerk ist für die Koordination des gesamten Prozessors verantwortlich. Es kontrolliert hierbei die Befehle und Operationen, die das Rechenwerk als nächstes ausführen soll. Bei dem Register handelt es sich um Datenregister, das abhängig von ihrer Funktion unter anderem Adressen, aber auch Daten speichern kann. Sie zeichnen sich vor allem damit aus, dass auf diese Speicherbereiche schnell zugegriffen werden kann [8]. Das Rechenwerk, (Arithmetic Logic Unit, ALU) nutzt die vom Steuerwerk geprüften Operationen und Befehle, um diese dann auf den Daten des Registers auszuführen. Bei den Operationen und Befehlen die Verarbeitet werden handelt es sich, wie der Name ALU schon verrät um arithmetische und logische Funktionen [8]. Alle Komponenten sind untereinander über ein Leitungssystem, auch Bus genannt verbunden. Die Arbeitsweise des Prozessors geht dem Abrufen-Dekodieren-Ausführen-Zyklus nach [8]. – Pipelining etwas verschnellern aber nicht wesentlich. Hierfür benötigt man mehr Prozessoren

Mehrprozessorsysteme sind Computersysteme, die mehr als einen Prozessor haben. Sie werden meist für (MIMD /SIMD) Statt mehr Horizontal und vertikal Skalieren erklären == Mehrprozessorsysteme vs Multicomputersysteme bspw Blockchain Minen Multiprozessorsysteme – einfach programmieren allerdings schwer zu verbinden [8] Multicomputersysteme – einfacher zu bauen [8]

Durch die stetige Entwicklung von Computerchips und der damit verbundenen Erhöhung

von Transistoren auf diesen steigt die Leistungsfähigkeit von Prozessoren weiter an [8]. Fast 5 Jahrzehnte galt der Grundsatz von Gordon Moore, dass sich ungefähr alle 18 Monate die Anzahl der Transistoren pro Flächeneinheit (Integrationsdichte genannt) in einem Computerchip verdoppeln würden. Dieser Grundsatz ist als das Mooresche Gesetz bekannt. Doch auch dieses soll demnächst an seine physikalischen Grenzen kommen, da man sich im Bereich nm befindet [10].

Von CPU-Herstellern wird oft die hohe Taktfrequenz beworben. Der numerische Wert lässt Endkonsumenten suggerieren, dass es sich um eine Leistungsstärkere CPU handelt. Dies ist allerdings ein Trugschluss. Eine 350 Megahertz (MHz) CPU ist nicht zwangsläufig leistungsstärker als eine CPU mit 300 MHz. Die Zahl ignoriert nämlich die tatsächliche Berechnung die bei jedem Zyklus durchgeht [3] Bei alltäglicher Nutzung eines Computers beträgt die Auslastung in den wenigsten Fällen 100 Prozent, vorallem beim Hochfahren des Computers wird die CPU unter Stress gesetzt [6].

Information über (Cache)

2.2.3 Arbeitsspeicher

Während sich der Prozessor um die Bearbeitung des Programms kümmert, befindet sich das geöffnete Programm auf dem Speicher des Arbeitsspeichers. Es bietet die Fläche für das zu bearbeitende Programm. Der Arbeitsspeicher fungiert als eine Art Zwischenspeicher, der für die Ausführung von Programme benötigt wird damit diese ausgeführt werden kann. Im Gegensatz zu einer Festplatte existieren die Daten nur zur Laufzeit und würden, wenn diese nicht im Hintergrund auf einer Festplatte hinterlegt werden alle wieder gelöscht, sobald der Computer ausgeschaltet wird. Im allgemeinen Wortgebrauch wird der Arbeitsspeicher auch als Hauptspeicher oder Memory bezeichnet. Der englische Begriff !Random Access Memory! (kurz RAM) beschreibt den Arbeitsspeicher eigentlich ziemlich treffend als !Direktzugriffsspeicher!, da dieser nämlich Zugriff auf jeden möglichen Speicherplatz des Computers hat,

diesen direkt adressieren kann und im Austausch mit dem Prozessor steht [1]. Der RAM ist wie auch der Prozessor elementarer Bestandteil. Der Arbeitsspeicher ist aufgebaut als eine Anordnung von vielen Zellen, welche Bits in sich halten, (0 oder 1) [1].

Im Gegensatz zu Prozessoren, die sich in ihrer Leistung und damit ihrer Geschwindigkeit steigern, bekommen Arbeitsspeicher immer mehr Kapazität [8]. Technisch wäre es zwar möglich, den Arbeitsspeicher auch deutlich schneller zu machen, allerdings wäre das unnötig, da die Verbindung über den Bus zwischen dem Arbeitsspeicher und dem Prozessor zu langsam ist. Für die Optimierung der Zugriffsgeschwindigkeit nutzt man daher den sogenannten Cache (Bedeutung), einen Speicher, der direkt in oder neben dem Prozessor eingebaut ist [8]. Der Cache ist ein kleiner und schneller Speicher, der die am häufigsten verwendeten (Wörter?) aufbewahrt, um eine schnellere Zugriffszeit für die CPU [8]. Während bei dem Prozessor die Geschwindigkeit untereinander verglichen werden kann, ist bei dem Arbeitsspeicher vor allem wichtig, dass die Kapazität der Programmanforderung entspricht. Sollte dies nämlich nicht der Fall sein und das Programm benötigt mehr Speicher, als der Arbeitsspeicher vorweisen kann, wird die Festplatte benötigt, was dazu führt, dass der gesamte Prozess schlagartig verlangsamt wird [6]. Giakamozis (1999) [1] vergleicht den Arbeitsspeicher mit dem Kurzzeitgedächtnis eines Menschen, während der Sekundärspeicher das Langzeitgedächtnis repräsentiert. Sollte etwas nicht im Kurzzeitgedächtnis sein, müsste es von daher bei dem Arbeitsspeicher wichtig, genug Kapazität bereitzustellen. Ansonsten wird der RAM zum Flaschenhals des Systems. Der leistungsstärkste Prozessor bringt einem keinen Mehrwert, wenn der Arbeitsspeicher nicht genug Platz anbietet.

DDR Wenn der Arbeitsspeicher voll ausgelastet ist, wird die Festplatte genutzt, was den gesamten Prozess extrem verlangsamt [6]. Je langsamer der Speicher ist, desto mehr Zyklen muss die CPU warten [8].

2.2.4 Sekundärspeicher

Als Sekundärspeicher werden alle weiteren Speichermedien zusammengefasst, die im Gegensatz zum Arbeitsspeicher nicht in unmittelbarer Verbindung zum Prozessor stehen. Der Sekundärspeicher kann im Gegensatz zum RAM die Daten persistent speichern. Der Sekundärspeicher steht in Verbindung mit dem Arbeitsspeicher und kann zum Interagieren mit diesem Daten Lesen und Schreiben. In der von-Neumann-Architektur gehören Sekundärspeicher zu den Ein- und Ausgabegeräten. Typische eingebaute Sekundärspeicher im Computer ist die Festplatte, auch Hard Drive Disk (kurz HDD) genannt, sowie die Solid State Drive (kurz SSD). Weitere Sekundärspeicher wären beispielsweise CDs, DVDs, Blu-rays, USB-Sticks oder auch Disketten. Im Gegensatz zum Arbeitsspeicher ist die Kapazität der eingebauten Sekundärspeichern enorm, dafür allerdings die Zugriffszeit deutlich langsamer. Hauptaugenmerk liegt hier vor allem bei den Festplatten und SSDs. Eine Festplatte ist ein magnetisches Speichermedium bestehen aus mehreren runden Aluminiumscheiben mit magnetisierbarer Beschichtung [8] SSDs sind im Gegensatz zu Festplatten Solid State Drive's sind deshalb auch deutlich schneller als HDDs Zugriffszeiten Daher sind SSDs im Jahr 2021 auch vierfach so teuer wie HDDs für die gleiche Kapazität zugriffszeit vergrößert sich Auch in den nächsten Jahren wird sie die Kapazität von Festplatten und SSDs vergrößern. Wesentliche Erkenntnis für die Sekundärspeicher ist, dass für den gesamten Prozess die Zugriffsgeschwindigkeit von wichtiger Bedeutung ist, da diese ansonsten zum Flaschenhals werden kann. Neben der Zugriffsgeschwindigkeit muss selbstverständlich auch dafür gesorgt werden, dass auf dem Medium genügend Speicher vorhanden ist, um die Daten persistieren zu können.

2.2.5 Weitere Ein und Ausgabegeräte

Da unter dem Sammelbegriff Ein und Ausgabegeräte im Grunde alle Geräte inbegriffen sind, die in der von-Neumann-Architektur nicht in unmittelbarer Verbindung zum Prozessor stehen gibt es hier viele Beispiele für Ein und Ausgabegeräte. Beispielhafte Eingabegeräte,

die wohl die meisten Computer besitzen sind hierbei: Maus, Tastatur oder auch Mikrofon. Ausgabegeräte wären beispielsweise Monitor, Drucker oder Lautsprecher. Ein weiteres Ausgabegerät, welches sich typischerweise in Computern befindet und für alle Anwender wichtig ist, ist die Grafikkarte. Ohne die Grafikkarte, auch Videokarte oder Graphics processing unit (kurz GPU) genannt, würde man kein Bild auf dem Monitor angezeigt bekommen. Es ist die Schnittstelle über die Menschen mit dem Computer interagieren können. Wie bei einer CPU befindet sich in einer Grafikkarte ein Mikroprozessor. Während GPUs vor über einem Jahrzehnt hauptsächlich nur in der Lage waren 3D-Rendering-Prozesse zu verarbeiten, verfügen Grafikkarten Prozessoren mittlerweile über enorme arithmetische Kapazitäten, die sogar die CPU übertreffen [4]. Sie sind optimiert für das parallele Berechnen von repetitiven arithmetischen Aufgaben, wie das Darstellen von Pixeln in einem Computerspiel. Die Grafikkarte muss nämlich ununterbrochen jede Visualisierung in einem Computerspiel berechnen. In den letzten Jahren kamen vor allem zwei weitere Trends hinzu, für die sich die Funktionsweise der GPU als vorteilhaft erwiesen hat und das, obwohl diese Aufgaben keine grafischen sind. Das ist auf der einen Seite das sogenannte Minen für Blockchains, wie die Kryptowährung Bitcoin und auf der anderen Seite die Berechnung für künstliche neuronaler Netze von Machine Learning Prozessen [5] [9]. Die Kombination aus Parallelisierung und der starken arithmetischen Kapazität des Mikroprozessors der Grafikkarte macht sie daher sehr beliebt für Prozesse die enorm viele Aufgaben berechnen muss. Gewöhnliche Office-Computer oder Laptops benötigen dahingegen nicht die zusätzliche Grafikleistung und haben daher einen integrierten Grafikprozessor auf der Hauptplatine oder in der CPU direkt integriert. Dieser reicht aus, wenn keine grafisch intensiven Arbeiten verrichtet werden.

2.2.6 Zusammenspiel der Komponenten für die Performance

technologischer Fortschritt Auch wenn der Prozessor hauptsächlich für die Performance verantwortlich ist, so sind die anderen Komponenten im Zusammenspiel dennoch wichtig für die Gesamt-Performance des Computers. Während sich Prozessoren durch ihre Geschwindigkeit

überbieten können. Kommt es bei den Speichermedien vor allem auf die Mindestkapazität (Arbeitsspeicher) und die Zugriffszeit (Sekundärspeicher) an. Sie sind also weniger ein Beschleuniger der CPU Leistung, sondern vielmehr potentielle Flaschenhälse, welche sich negativ auf die gesamte Performance auswirken können.

2.3 Metriken

2.3.1 Anforderungen an eine Metrik

Die Bewertungskriterien für einen Performance Test werden in einer Metrik festgehalten. Allerdings erweist es sich nicht als trivial nach welchen Kriterien die Leistung eines Computers bewertet werden soll, da kein wissenschaftlicher Konsens bis dato herrscht. Historisch betrachtet gibt es dementsprechend sowohl Unklarheiten als auch Uneinigkeiten über die richtige Methodik, da Metriken oftmals missverständlich interpretiert werden oder auch falsch angewendet wurden [3]. Daher stellt sich die Frage, was eine aussagekräftige Metrik für Voraussetzungen haben muss. Nach David J Lilja sollte eine Metrik folgendes mitbringen: - Linear - Zuverlässig - Wiederholbar - Einfachheit der Messbarkeit - Konsistenz - Unabhängigkeit nach Huppler - Relevanz - Wiederholbar - Fair and Portable - Verifizierbar - Economical

Eine faire und repräsentative Benchmark ist wichtig [2] (2020SSB)

Schon im Jahr 1988 [7] (1988JES) galt es als kontrovers die Performance eines Computers auf eine einzige Zahl zu reduzieren, da man sich im wissenschaftlichen Diskurs einig war, dass die Performance multidimensional ist und sich daher durch mehrere Zahlen repräsentiert [7]

Wie in dem CPU-Kapitel ausgeführt, bietet sich die Taktfrequenz der CPU als ungeeignetes Kriterium einer Metrik an, da ein höherer Megahertz-Wert keine sinnvolle Aussage trifft über die Performance [3].

2.3.2 Historische Metriken

Eine historisch oft genutzte und bis heute immer noch beliebte Maßeinheit für die Berechnung der Computer Performance ist Millionen Instruktionen pro Sekunde (MIPS). MIPS MFLOPS QUIPS GigaFlops CPU Performance Equation

2.3.3 Industrie Benchmarks

SPEC Benchmark System Performance Evaluation Cooperative Aufgrund der verschiedenen Einheiten die

2.3.4 openSource Benchmarks Metriken

2.3.5 Fehlerquellen

Errors Genauigkeit des Timers.

Chapter 3

Methodik

3.1 Anforderungen an das Programm

Link zu dem github Projekt: Programmiersprache Virtual Machine : Google Cloud Platform

Das Kommandozeilen Tool wurde in der Programmiersprache Java geschrieben Das Tool muss auf Linux Systemen funktionieren und soll dabei alle relevanten Hardware-Informationen des Computers auslesen, diese dann auf einer externen Datenbank speichern. Gleichzeitig soll anhand der Hardware-Informationen weitere relevante Benchmark Werte aus Hardware Datenbanken ausgelesen und ebenfalls gespeichert werden. Danach soll ein Ergebnis berechnet werden, wie schnell der Computer im Vergleich zu einem Referenz Computer performed. Es soll außerdem möglich sein, aus der Datenbank eine Gruppierung der vorhandenen Hardware Daten zu erstellen. Für das Auslesen der Hardware Information wurden die Standard Kommandozeilen Befehle `lscpu` und `xyz` genutzt. Als externe Datenbank zum speichern der relevanten Hardware Informationen wurde Postgresql gewählt. Die Datenbank läuft extern auf einer virtuellen Maschine, welche von der Google Cloud Platform bereit gestellt wurde. Somit ist es möglich das Programm überall ausführen können um die Daten zu speichern. Als Hilfstool zum bearbeiten der Datenbank wurde DBeaver genutzt. Im Internet existieren

eine Vielzahl von Online-Datenbanken für Hardware Informationen. Allerdings sind die meisten spezialisiert auf die Performance von Computerspielen und daher auch sehr Grafikkarten lastig. Hinzu kommt, dass keine einzige gefundene Online-Datenbank eine Programmierschnittstelle anbietet. Daher wurde die HtmlUnit Bibliothek für Java benutzt zum sogenannten scraping von Webseiten. Dies bedeutet, dass auslesen des gesamten Html Textes von Webseiten. Mithilfe des Document Object Model (kurz DOM) ist es möglich, direkt Elemente aus dem Html Text, wie beispielsweise Tabellen direkt anzusprechen und zu verarbeiten.

Es wurden zur Testen über 10 Computer genutzt. Zwar bietet Google Cloud Platform auch virtuelle Maschinen an, das Problem bei diesen ist allerdings, dass durch die virtualisierung keine standardmäßigen Hardware Komponenten genutzt werden, sondern für virtualisierung prozesse angepasste Komponenten die von den Hardware Herstellern extra entwickelt worden sind und dementsprechend auch in kaum einer Online-Datenbank zu finden sind. Hierfür wurden dennoch interne Mikrobenchmarking Tests gemacht.

<http://cpudb.stanford.edu/processors> : wissenschaftliche Datenbank für CPUs, allerdings SPEC Daten von 2006, sprich viele CPUs Benchmarks gar nicht vorhanden.

3.2 Online-Datenbanken

3.2.1 cpu-world

3.2.2 userBenchmark

3.2.3 openBenchmarking

3.3 Berechnung des Scores

Chapter 4

Auswertung

4.1 Validierung der Ergebnisse

4.2 Abweichungen

4.3 Klassifizierung der Ergebnisse

Chapter 5

Stand der Forschung

Während der Bearbeitung gab es einige Hürden. Wie schon in dem Kapitel (METRIKEN) beschrieben, gibt es bis dato keine gängige Metrik die im wissenschaftlichen Diskurs angewendet wird. Vor allem (insert alte Metrik MFLOPS) wird noch genutzt wobei die Nachteile hier deutlich überwiegen. Oftmals werden in wissenschaftlichen Arbeiten wiederum eigene Metriken oder Benchmarks kiiert, so dass es oft gar nicht dazu kommen kann das sich Metriken etablieren. (evntl auch in den Fazit/Diskussionsteil) Dadurch das Computerspiele, aber auch das Mining von Computern für Kryptowährungen sehr populär geworden ist, gibt es unzählige Benchmark Werte für die Leistung der Komponenten von Computerspielen, vor allem spielt dort die Grafikkarte nochmal einen höheren Nutzen, also für (produktion / entwicklung). Es gibt kaum wissenschaftliche Hardware Online Datenbanken. Leider gibt es dazu keine API Schnittstellen bei den restlichen Online Datenbanken, sodass die extraktion relevanter Daten deutlich aufwendiger war.

Eine spannende Thematik in dem Bereich Benchmarking liefert aber das Thema Performance Prediction durch Machine Learning Ansätze : beschrieben –

Chapter 6

Diskussion

Chapter 7

Fazit

Appendix A

Test

Appendix B

Abkürzungsverzeichnis

AWS Amazon Web Services

IoT Internet of Things

Appendix C

Lexikon

Byzantine failure A malfunction of a component that leads to the distribution of wrong/-conflicting information to other parts of the system is called Byzantine failure . The name is based on the Byzantines Generals Problem, in which three Byzantine generals need to agree on a battle plan while one or more of them might be a traitor trying to confuse the others.

Appendix D

Listings

This is the appendix for code, that does not need to be provided directly inside the thesis.

D.1 Configuration for Node A

Listing D.1: Configuration for Node A

```
{
  "nodeID" : "nodeA",
  "publicKey" : "<public key>",
  "encryptionAlgorithm" : "RSA",
  "machines" : [ "192.168.0.132", "192.168.0.165" ],
  "publisherPort" : 8000,
  "messagePort" : 8010,
  "restPort" : 8080,
  "location" : "52.515249, 13.326476",
  "description" : "Raspberry Pi Cluster in EN 004"
}
```


Literaturverzeichnis

- [2] Samuel Kounev, Klaus-Dieter Lange, and Jóakim von Kistowski. *Systems Benchmarking*. Springer, 2020.
- [3] David J Lilja. *Measuring computer performance: a practitioner's guide*. Cambridge university press, 2005.
- [4] John D Owens et al. “GPU computing”. In: *Proceedings of the IEEE* 96.5 (2008), pp. 879–899.
- [5] Daniel Schlegel. “Deep machine learning on Gpu”. In: *University of Heidelber-Ziti* 12 (2015).
- [6] Rajesh Singh. “An approach to enhance performance of Computer”. In: *International Journal of Science & Engineering Development Research* 1.2 (2016), pp. 298–302.
- [7] James E. Smith. “Characterizing computer performance with a single number”. In: *Communications of the ACM* 31.10 (1988), pp. 1202–1206.
- [8] Andrew S. Tanenbaum. *Computerarchitektur: Strukturen, Konzepte, Grundlagen*. 5. Auflage. Pearson Studium, 2006.
- [9] Michael Bedford Taylor. “The evolution of bitcoin hardware”. In: *Computer* 50.9 (2017), pp. 58–66.
- [10] Thomas N Theis and H-S Philip Wong. “The end of moore’s law: A new beginning for information technology”. In: *Computing in Science & Engineering* 19.2 (2017), pp. 41–50.

Literaturverzeichnis (Online)

- [1] Peter Haugen et al., eds. *A Basic Overview of Commonly Encountered types of Random Access Memory (RAM)*. 2000. URL: <https://www.rose-hulman.edu/Class/ee/yoder/ece332/Papers/RAM%20Technologies.pdf> (visited on 10/10/2021).