

Increasing throughput of server applications by using asynchronous techniques: A case study on CoAP.NET

Philip Wille

Bachelor Thesis

Supervisor:
Dr. Michael Felderer
Department of Computer Science
Universität Innsbruck

Innsbruck, 10. Januar 2022



Inhaltsverzeichnis

1	Einleitung	1
1.1	Constrained Application Protocol (CoAP)	2
1.2	Ausführungsparadigmen in der Informatik	11
1.3	Bekannte Implementierungen	16
1.4	Industrielle Motivation und Alternativen	17
1.5	Ziel der Arbeit	18
2	Implementierung	19
2.1	Struktur der Applikation	20
2.2	Besonderheiten	21
3	Messung	25
3.1	Nachrichtenverarbeitung	25
3.2	Serialisierung und Deserialisierung	26
3.3	Messaufbau	27
3.4	Messergebnisse (Nachrichtenübertragung)	27
3.5	Messergebnisse (Deserialisierung und Serialisierung)	29
4	Diskussion der Messergebnisse	34
5	Schlussfolgerung	35
Anhang		41
1	Tabellen	43
1.1	Ergebnisse (Nachrichtenübertragung)	43
1.2	Ergebnisse (Deserialisierung und Serialisierung)	44
2	Codeausschnitte	47

1 Einleitung

Heutzutage haben Webdienste einen hohen Stellenwert im Internet. Sie sind häufig ein essenzieller Bestandteil von Anwendungen und werden durch eine *Representation State Transfer* (REST)-Schnittstelle für Anwender und Entwickler angeboten. Auch das *Internet of Things* (IoT) bekommt mit fortschreitender Zeit eine immer wichtigere Rolle in der Entwicklung von Anwendungen, wie zum Beispiel in der Hausautomatisierung oder smarte Energieverwaltung, wie in den wissenschaftlichen Berichten „Implementation of Home Automation Using CoAP“ von Kodali u. a. [3], „A study and toolkit for asynchronous programming in C#“ von Okur u. a. [6], „TinyCoAP: A Novel Constrained Application Protocol (CoAP) Implementation for Embedding RESTful Web Services in Wireless Sensor Networks Based on TinyOS“ von Ludovici, Moreno und Calveras [5] oder „Implementation of coap and its application in transport logistics“ von Kuladinithi u. a. [4] thematisiert wird.

Um die REST-Architektur auch für eingeschränkte Geräte anbieten zu können, wurde mit *Constrained RESTful Environments* (CoRE) eine geeignete Form geschaffen. Somit können solche Geräte, wie zum Beispiel 8-Bit-Mikrocontroller mit begrenztem Arbeitsspeicher und Readonly-Memory (ROM), als auch Netzwerke, wie zum Beispiel *IPv6 over Low-Power Wireless Area Networks* (6LoWPANs), solche Architekturen verwenden und realisieren. Damit eine Fragmentierung von Nachrichten in solchen Netzwerken so gering wie möglich auftritt, wird im Constrained Application Protocol (CoAP) ein so niedriger Nachrichten-Overhead wie möglich angestrebt.

Mit CoAP wurde die Entwicklung eines generischen Webprotokolls für die speziellen Anforderungen dieser eingeschränkten Umgebungen, insbesondere mit Hauptaugenmerk auf Energie-, Gebäudeautomatisierungs- und andere Machine-to-Machine (M2M) Anwendungen vorangetrieben. Dabei sollte CoAP keine komprimierte Abwandlung von *Hypertext Transfer Protocol*, abgekürzt HTTP, sein, sondern vielmehr eine Submenge von HTTP mit Verwendung von REST, die für M2M-Szenarien optimiert ist. Somit könnte CoAP leicht dazu verwendet werden, einfache HTTP-Schnittstellen in ein kompaktes Protokoll überzuführen, jedoch bietet CoAP Funktionen an, die speziell in Machine-to-Machine-Anwendungen Verwendung finden. Diese Funktionen sind:

- Eingebaute Entdeckung von, im Netzwerk, angebotenen Services und Ressourcen.
- Multicast-Unterstützung.
- Asynchroner Nachrichtenaustausch.

In den, von CoAP angedachten Anwendungsgebieten, in denen eine große Menge an Daten über ein Netzwerk fließen, könnten asynchrone Techniken die passende Ergänzung sein, um eine hohe Anzahl von Nachrichten zu empfangen, zu versenden oder zu verarbeiten. Den mit steigender Digitalisierung in privaten als auch geschäftlichem Umfeld, im Form von intelligenter Haussteuerung, effizienter Nutzung von Energieerzeugern

und Energieverbrauchern¹, steigt auch der Bedarf an leistungsfähigen Anwendungen, die diese schnell anwachsende Anzahl an IoT-Geräten ressourcenschonend und effizient verwalten kann.

Jedoch eignet sich Asynchronität nicht nur zur Optimierung von Anwendungen im IoT-Bereich, sondern auch für solche mit einer grafischen Programmoberfläche (GUI). Dies hat den Grund, da der für das Zeichnen der Oberfläche verantwortliche Thread freibleibt und somit neu ankommende Eingaben des Benutzers (Mausklick, Tastatureingabe etc.) verarbeiten werden kann, ohne dass die Oberfläche „einfriert“².

Dieses Programmierparadigma ist auch nützlich, wenn mit Datenbanken oder REST-Schnittstellen interagiert oder I/O- bzw. CPU-intensive Operationen ausgeführt werden, wie Okur u. a. in deren Studie „A study and toolkit for asynchronous programming in C#“ [6] herausgefunden haben.

1.1 Constrained Application Protocol (CoAP)

Das *Constrained Application Protocol*, abgekürzt CoAP, ist ein Internetprotokoll, das speziell für M2M Anwendungen, wie zum Beispiel smarte Energieverwaltung oder Hausautomatisierung, entwickelt wurde [10]. Dabei bietet das Protokoll ein REST-ähnliches Interface für Mikrocontroller mit angeschlossenen Aktoren oder Sensoren, auch *constrained nodes* genannt, oder auch drahtlose Sensornetze, auch *constrained networks* genannt, an. Die sogenannten *nodes* besitzen meist einen angeschlossenen 8-Bit-Mikrocontroller, der nur eine kleine begrenzte Menge an Arbeitsspeicher (RAM) und Readonly-Memory (ROM) beinhaltet. Bei *constrained networks*, wie z.B. *IPv6 over Low-Power Wireless Area Networks (6LoWPANs)*, spielt die hohe Paketverlustrate und der für solche Netzwerke typische Datendurchsatz von wenigen 10 kbit/s eine große Rolle. CoAP ist durch den RFC 7252 von Shelby, Hartke und Bormann [8] spezifiziert.

Dabei bietet das *Constrained Application Protocol* ein Interaktionsmodell für Anfragen und Antworten zwischen den, in der Anwendung, definierten Endpunkten an. Auch ist eine eingebaute Entdeckung von, im Netzwerk, angebotenen Services und Ressourcen im Protokoll definiert. Zusätzlich werden Hauptbestandteile des Internets, wie *Unique Resource Identifiers* (URIs) und *Internet Media Types* (zum Beispiel *application/json*) angeboten.

Eine Unterstützung von *Multicast* ist gegeben, als auch ein sehr geringer Mehraufwand in der Datenübertragung. Ferner wurde darauf Wert gelegt, es einfach für eingebettete Systeme zu halten.

Zur Datenübertragung nutzt CoAP das *User Datagram Protocol* (UDP). Dieses unterscheidet sich zum *Transmission Control Protocol* (TCP) in den folgenden Punkten:

¹SmartSpeicher - intelligente Warmwasserspeicher zur Stabilisierung des Stromnetzes

²Beschreibung des Zustandes eines Programmes, indem die Oberfläche keine neuen Eingaben entgegennimmt

- Kein Sitzungsaufbau von Sender zu Empfänger (*Handshake*).
- Stellt nicht sicher, ob alle Pakete beim Empfänger eintreffen.
- Geht ein Paket verloren, wird dieses nicht erneut versendet.

Die folgenden Funktionen sind essenzielle Bestandteile von CoAP:

- Erfüllung von M2M Anforderungen in eingeschränkten Umgebungen.
- UDP-Verbindung mit optionaler Zuverlässigkeit, die Unicast- und Multicast Anfragen unterstützt.
- Asynchroner Nachrichtenaustausch.
- Niedriger Mehraufwand durch veränderten Header und niedriger Komplexität des Parsings.
- URI- und Content-Typen-Support.
- Einfache Proxy- und Caching-Unterstützung.
- Zustandsloses HTTP-Mapping, welches die Entwicklung von Proxys erlaubt, die den Zugriff auf CoAP Ressourcen über HTTP auf einheitliche Weise ermöglichen oder für einfache HTTP-Schnittstellen, die alternativ über CoAP realisiert werden können.
- Sicherheitsmechanismen durch das Anbinden von *Datagram Transport Layer Security* (DTLS).

Begriffe in CoAP

Um den Kontext innerhalb des *Constrained Application Protocols* zu verstehen, werden nachfolgend die wichtigsten Begriffe in CoAP kurz erklärt:

- Endpunkt (Endpoint):
 - Ein Endpunkt lebt auf einen „Knoten“. Ein Knoten ist vergleichbar mit dem Begriff „Host“, der vorwiegend in anderen Internetstandards Erwähnung findet.
 - Ein Endpunkt wird durch Multiplexing-Informationen auf der Transportschicht identifiziert, die eine UDP-Portnummer und eine Sicherheitszuordnung enthalten könnte.
- Ursprungsserver (Origin Server):
 - Jener Server, auf dem sich eine bestimmte Ressource befindet oder erstellt werden soll.
- Bestätigende Nachricht (Confirmable Message):

- Einige Nachrichten benötigen eine Bestätigung des Empfängers. Diese Nachrichten werden als *bestätigt* behandelt.
- Falls keine Pakete während der Übertragung verloren gingen, wird für jede Nachricht, die bestätigt werden muss, exakt eine Nachricht des Typs *Acknowledgement* oder *Reset* an den Sender zurückgesendet.
- Nicht bestätigende Nachricht (Non-confirmable Message):
 - Im Gegensatz zu bestätigenden Nachrichten gibt es auch Nachrichten, die nicht bestätigt werden müssen.
 - Dies betrifft Nachrichten, die für bestimmte Anwendungsanforderungen häufiger wiederholt werden müssen, wie zum Beispiel, das wiederholte Lesen eines Sensors.
- Bestätigungsnachricht (Acknowledgement Message):
 - Eine solche Nachricht bestätigt den Empfang einer bestätigenden Nachricht. Eine Bestätigungsnachricht sagt nicht aus, ob die Anfrage, die mit einer bestätigenden Nachricht versendet wurde, erfolgreich war oder nicht.
 - Jedoch enthält die Bestätigungsnachricht auch eine sogenannte *Piggybacked Response*.
- Rücksetzende Nachricht (Reset message):
 - Diese Nachricht sagt aus, dass eine spezifische Nachricht, mit dem Typ *Confirmable* oder *Non-confirmable*, empfangen wurde, jedoch einige Teile des Nachrichtenkontextes fehlen, um diese Nachricht korrekt verarbeiten zu können.
 - Dieses Verhalten tritt auf, wenn der zu empfangende Endpunkt neu gestartet wurde und somit den Zustand vergessen hat, der zur vollständigen Interpretation der Nachricht nötig war.
 - Ein absichtliches Provozieren einer rücksetzenden Nachricht *Reset Message*, zum Beispiel durch das Senden einer leeren bestätigenden Nachricht (*Empty Confirmable Message*), kann als eine kostengünstige Prüfung der Funktionsfähigkeit eines Endpunktes verwendet werden - vergleichbar mit einem *Ping*.
- „Antwort im Huckepackverfahren“ (Piggybacked Response):
 - Eine *Piggybacked Response* ist direkt in eine *Acknowledgement* (ACK) Nachricht inkludiert, die gesendet wird, um den Empfang der Anfrage für diese Antwort zu bestätigen.
- Separate Antwort (Separate Response):

- Wenn eine *Confirmable* Nachricht mit einer Anfrage mit einer *Empty* Nachricht quittiert wird (z.B. weil der Server die Antwort nicht sofort hat), wird eine separate Antwort in einem separaten Nachrichtenaustausch gesendet.
- Leere Nachricht (Empty Message):
 - Eine Nachricht mit dem Code 0.00. Die Nachricht ist weder eine Anfrage noch eine Antwort. Sie beinhaltet nur den 4-Byte-langen Kopf (*header*).

Nachrichtenübertragung

CoAP nutzt zur Nachrichtenübertragung UDP, um den Austausch von Nachrichten asynchron ausführen zu können. Dies wird dadurch erreicht, dass eine weitere Schicht auf UDP aufbauend eingefügt wird (siehe Abbildung 1). Diese Schicht kann optional auch einen Mechanismus zur Sicherstellung des Nachrichtenaustausches beinhalten. Dabei definiert CoAP vier verschiedene Arten von Nachrichten:

- Bestätigende Nachricht (*Confirmable Message*),
- Nicht bestätigende Nachricht (*Non-confirmable Message*),
- Bestätigungsnachricht (*Acknowledgement Message*),
- Rücksetzende Nachricht (*Reset Message*),

Diese vier Arten stehen orthogonal zueinander, sprich:

- Bestätigende und nicht bestätigende Nachrichten sind für Anfragen (*requests*) gedacht.
- Rücksetzende Nachrichten oder Nachrichten, die eine Anfrage bestätigen, sind für Antworten (*responses*) gedacht.

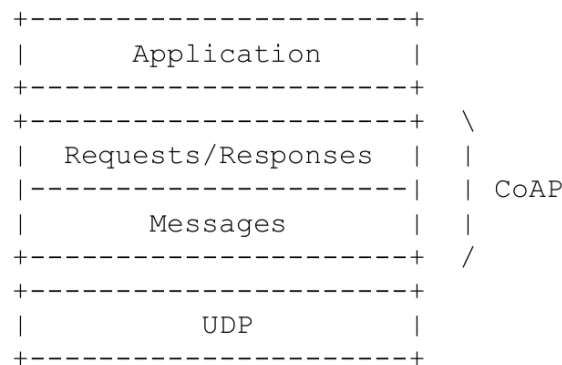


Abbildung 1: Abstrakte Darstellung der verschiedenen Schichten (Quelle: RFC 7252 [8])

Das Nachrichtenmodell des Constrained Application Protocols basiert auf den Austausch von Nachrichten über UDP. Dabei beginnt die Nachricht mit einer in der Länge fixierten

Vier-Byte-Langen Kopfzeile (*header*), gefolgt von einem optionalen Token (null bis acht Bytes lang), und null oder mehr sogenannten Optionen. Diese Optionen sind vergleichbar mit den *header fields* von HTTP. Nach den Optionen befindet sich ein sogenannter Anhang-Markierer (*Payload marker*), der einem Byte mit acht logischen Einsen, hexadezimal auch 0xFF kodiert, entspricht, jedoch nur in der Nachricht enthalten ist, wenn eine Payload mitgesendet wird. Anschließend an den *Payload marker* kommt der Anhang (*Payload*). Dieses Format ist sowohl für Anfragen als auch für Antworten dasselbe.

Damit Nachrichten als eindeutig identifiziert werden können, wird ein sogenannter Nachrichtenbezeichner (*Message ID*) verwendet. Dieser ist 16 Bit lang und erlaubt somit, bei Implementierungen mit Standardeinstellungen, bis zu 250 Nachrichten die Sekunde von einem Endpunkt zu einem anderen zu senden. Die *Message ID* wird auch für die Sicherstellung des Nachrichtenaustausches (*reliability*) benötigt. Dabei ist jedoch die *Message ID* nur zwischen zwei Endpunkten eindeutig. Kommuniziert ein Teilnehmer mit mehreren Teilnehmern gleichzeitig, dann können die *Message IDs* häufiger vorkommen. Für die eindeutige Identifizierung der Kommunikation zwischen zwei Teilnehmern wird der sogenannte Token verwendet. Dieser ist über mehrere Verbindungen hinweg eindeutig und kann als Identifikator für derzeit laufende Verbindungen gesehen werden.

Die Sicherstellung des Nachrichtenaustausches erfolgt dadurch, dass man eine Nachricht als bestätigend (*Confirmable*) markiert. Eine als *Confirmable* gekennzeichnete Nachricht wird so lange an den jeweiligen Empfänger gesendet, bis dieser eine *Acknowledgement* Nachricht mit derselben *Message ID* zurücksendet (wie in Abbildung 2 dargestellt). Wenn der Empfänger die *Confirmable* Nachricht, aufgrund fehlender Daten oder fehlendem Kontext nicht beantworten kann, sendet dieser eine *Reset* Nachricht zurück.

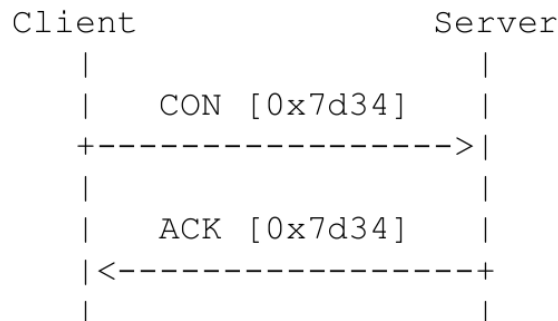


Abbildung 2: Nachrichtenaustausch mit Sicherstellung des Transfers (Quelle: RFC 7252 [8])

Jedoch wird für den Austausch von Nachrichten im CoAP Kontext kein Sicherheitsmechanismus für die Übertragung gefordert, sondern es können auch Nachrichten als *Non-confirmable* markiert werden (siehe Abbildung 3). Diese Vorgangsweise bietet sich zum Beispiel an, wenn man die Messdaten eines Sensors wiederholt ausliest. Dabei werden *Non-confirmable* Nachrichten nicht bestätigt, jedoch wird eine *Message ID* benutzt, um Duplikate zu erkennen.

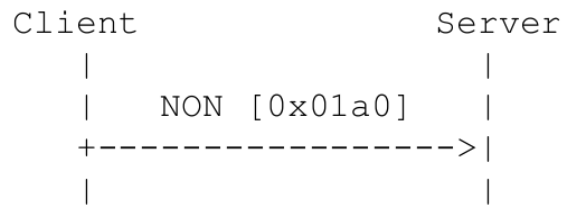


Abbildung 3: Nachrichtenaustausch ohne Sicherstellung des Transfers (Quelle: RFC 7252 [8]).

Kann eine ankommende Anfrage (*Request*), welche mithilfe einer *Confirmable* Nachricht versendet wurde, sofort beantwortet werden, wird die Antwort (*Response*) in der daraus resultierenden *Acknowledgment* Nachricht zurückgesendet. Dieses Prinzip nennt man auch *Piggybacked Response*. Das Abbildung 4 stellt diesen Mechanismus dar.

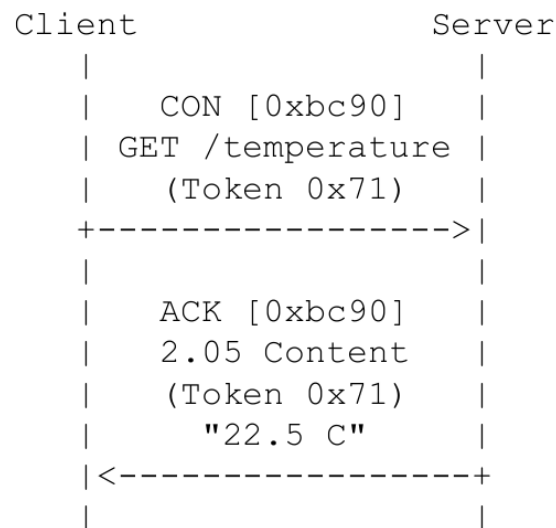


Abbildung 4: Beispiel einer erfolgreichen *Piggybacked Response* (Quelle: RFC 7252 [8]).

Ist der Server jedoch nicht sofort in der Lage die Anfrage zu beantworten, dann antwortet dieser mit einer leeren *Confirmable* Nachricht. Dies macht er, um den Client am wiederholten Senden der Anfrage zu hindern. Sind alle benötigten Daten zur Beantwortung der Anfrage vorhanden, sendet der Server die Antwort in einer neuen *Confirmable* Nachricht zurück. Dieses Prinzip wird als *separate response* bezeichnet und wird in der Abbildung 5 dargestellt.

Dabei macht CoAP von den bekannten Internetmethoden *GET*, *PUT*, *POST* und *DELETE* in einer ähnlichen Art gebrauch, wie es HTTP macht.

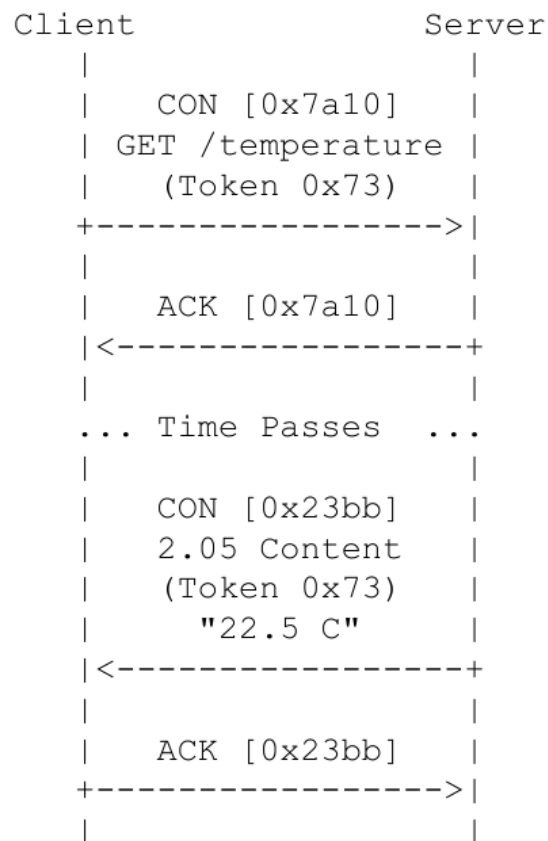


Abbildung 5: Beispiel einer *separate response* (Quelle: RFC 7252 [8])

Nachrichtenformat

Wie schon erwähnt, basiert der Nachrichtenaustausch von CoAP auf UDP. Dabei nimmt jede, über UDP versendete Nachricht ein ganzes UDP Datagramm in Anspruch. Der Aufbau einer CoAP Nachricht ist einfach gehalten und startet mit einer Vier-Byte-langen-Kopfzeile (*Header*), welche folgenden Daten beinhaltet, visualisiert in Abbildung 6:

- *Version*
- *Type*
- *Token Length*
- *Code*
- *Message ID*

Dabei repräsentieren die ersten zwei Bits des *Headers* die Versionsnummer. Die Versionsnummer gibt an, in welcher CoAP Version die Nachricht erstellt wurde bzw. verarbeitet werden kann. In dieser Arbeit beschäftigen wir uns mit CoAP Nachrichten mit der Versionsnummer 1, in Binär als 01 kodiert.

Die darauffolgenden zwei Bits entsprechen dem Typ der CoAP Nachricht. Der Typ gibt an, ob es sich um eine *Confirmable* (0 = in Binär als 00 kodiert), *Non-Confirmable* (1 = 01), *Acknowledgement* (2 = 10) oder *Reset* (3 = 11) Nachricht handelt.

Als Nächstes kommt die vier Bit lange *Token Length*, welche die Länge des *Tokens* angibt. Dabei kann der *Token* zwischen null, in Binär als 0000 kodiert, und acht, in Binär als 0111 kodiert, Bytes lang sein. *Token Lengths*, welche zwischen neun und fünfzehn Bytes lang sind, sind im RFC 7252 [8] für zukünftige Versionen reserviert.

Die nächsten acht Bit geben den *Code*, vergleichbar mit dem Statuscode bei HTTP, der CoAP Nachricht an. Dabei unterteilt sich der *Code* in eine drei Bit lange *Code Class* (*most significant bits*) und einen fünf Bit langen *Code Detail* (*least significant bits*). Dabei folgt der *Code* dem Schema „c.dd“, wobei „c“ Werte von 0 bis 7 annehmen kann und „dd“ Werte von 00 bis 31. Die *Code Class* gibt dabei an, ob es sich um

- eine Anfrage (0),
- eine erfolgreiche Antwort (2),
- eine clientseitige, fehlerhafte Antwort (4),
- oder eine serverseitige, fehlerhafte Antwort (5) handelt.

Daneben nimmt der *Code* 0.00 eine besondere Stellung ein, da dieser eine leere Nachricht (*Empty Message*) markiert. Die *Codes* gleichen sich mit einigen Statuscodes, welche man von HTTP kennt, jedoch ist nicht jeder Statuscode als CoAP *Code* abgebildet.

Der letzte Teil des *Headers* ist die sogenannte *Message Id*, die 16 Bit in Anspruch nimmt und in der *Network Byte Order*, auch unter dem Begriff *Big Endian* bekannt, angegeben wird. Dessen Aufgabe ist es, Duplikate von Nachrichten zu erkennen. Auch wird die *Message Id* dafür benutzt, um Nachrichten vom Typ *Acknowledgement* und *Reset* zu Nachrichten vom Typ *Confirmable* und *Non-Confirmable* zu verlinken. Somit besitzt der Server immer einen Überblick, zu welcher Anfrage schon eine Antwort geschickt wurde.

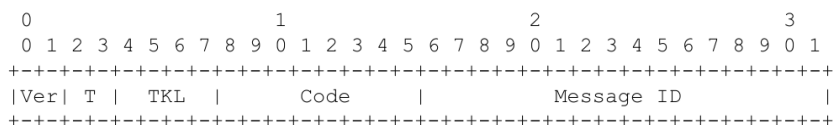


Abbildung 6: Binäre Struktur eines CoAP *Headers* (Quelle: RFC 7252 [8])

Anschließend an den *Header* kommt der *Token* der Nachricht. Dieser ist in seiner Länge variabel und hängt von der im *Header* angegebenen *Token Length* ab. Dieser ist zuständig für die Korrelation von Anfragen zu Antworten.

Nachfolgend können null oder mehr sogenannte *Options* folgen. Der Option können folgende Bestandteile einer CoAP Nachricht folgen:

- Das Ende der CoAP Nachricht (EoF),
- Eine weitere Option,
- Oder der *Payload Marker* mit anschließender *Payload*.

Ist eine *Payload* gegeben, folgt nach der Gruppe von Optionen ein sogenannter *Payload Marker*. Dieser besteht aus einem Byte voller logischer Einsen, hexadezimal als 0xFF kodiert, und markiert somit das Ende der *Options*. Alle Daten, die sich nach dem *Payload Marker* befinden, werden als *Payload* interpretiert. Dabei ist die Länge durch die *UDP Datagram* Paketgröße von 65 535 Bytes begrenzt. Werden für die Übertragung der Nachricht mehr Bytes benötigt als ein *UDP Datagram* an Größe bereitstellen kann, werden die Bytes auf mehrere *UDP Datagrams* aufgeteilt. Diesen Mechanismus nennt man auch *Blockwise transfer* und wird im RFC 7959 von Shelby und Bormann [7], der auf dem RFC 7252 von Shelby, Hartke und Bormann [8] aufbaut, beschrieben.

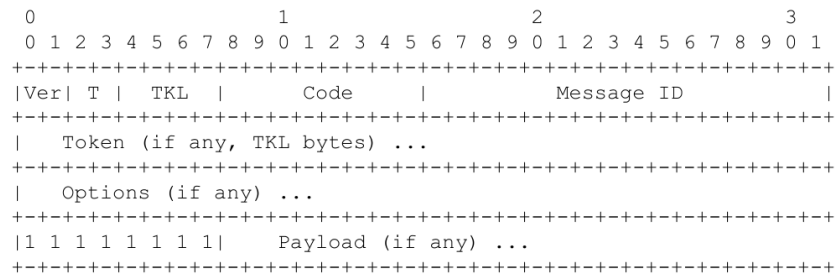


Abbildung 7: Binäre Struktur einer vollständigen CoAP Nachricht (Quelle: RFC 7252 [8])

Aufbau einer Option

Eine Option wird durch eine eindeutige Nummer identifiziert. Neben der Nummer besitzt eine Option auch einen Wert (*Value*), den diese Option hält, und einen Indikator für die Länge des Wertes. Dabei wird die Nummer nicht direkt in die Nachricht kodiert, sondern die Optionen werden zuerst aufsteigend nach ihrer Nummer sortiert und dann wird eine Deltakodierung (Differenzbildung) zwischen der aktuellen Option und deren Vorgängern gebildet. Dies geschieht dadurch, dass alle vorherigen Differenzen, auch im RFC 7252 [8] als *Delta* bezeichnet, addiert werden und dann die Differenz zur aktuellen Option gebildet wird. Für die erste Option wird der Sonderfall behandelt, dass als vorheriges Delta ein Wert von null angenommen wird. Dies resultiert daraus, dass für die erste Option die kodierte Differenz als Nummer der Option verwendet wird. Ein weiterer Sonderfall ist derjenige, wenn mehrere Instanzen der gleichen Option in der Kollektion von Optionen auftritt. Dabei ist die Differenz zwischen zwei gleichen Optionen immer null.

Eine Option fängt immer mit einem Byte an, das zwei Informationen enthält. Einmal die Differenz (*Option Delta*) und andererseits die Länge des Wertes (*Option Length*) der Option. Das *Option Delta* entspricht dabei den ersten vier Bits (*most significant bits*) und die *Option Length* den letzten vier Bits (*least significant bits*). Man kann dieses Byte auch als einen *Header* für Options bezeichnen, da dieser Informationen beinhaltet, die für das Kodierung bzw. Dekodieren von Options benötigt wird.

Um jedoch Deltas und Längen, jenseits des Wertes fünfzehn, verwenden zu können, können dem *Option Header* das *Option Delta (extended)* und die *Option Length (extended)* folgen. Diese beiden können jeweils zwischen null und zwei Bytes lang sein. Nach diesen beiden folgt der sogenannte *Option Value*, der null oder mehr Bytes betragen kann.

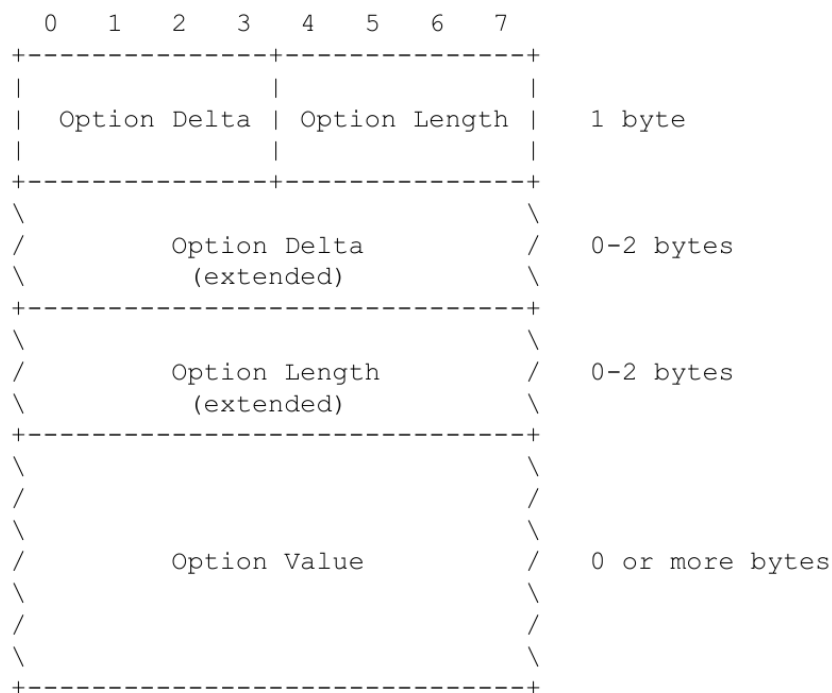


Abbildung 8: Binäre Struktur einer CoAP-Option (Quelle: RFC 7252 [8])

1.2 Ausführungsparadigmen in der Informatik

In der Informatik spricht man von zwei großen Ausführungsparadigmen, mit denen Programme bzw. Codeteile ausgeführt werden können: **synchrone** und **asynchrone Ausführung**. Diese beiden Ausführungsparadigmen unterscheiden sich in wesentlichen Punkten deutlich voneinander und bieten in verschiedenen Einsatzszenarien Vor- und Nachteile (siehe Tabelle 1). Dabei unterstützen die geläufigsten Programmiersprachen von Haus aus eine synchrone Ausführung von Programmen, jedoch bieten nicht alle eine asynchrone Ausführung.

	Synchron	Asynchron
Programmfluss	Stoppt den Programmfluss.	Kann im Programmfluss weiter gehen.
Beendigung	Überprüft periodisch, ob Funktion beendet ist.	Ein Event markiert die Beendigung der Funktion.
Main-Thread	Ist als „Blocked“ oder „Waiting“ markiert.	Ist frei für andere Aufgaben.

Tabelle 1: Vergleich zwischen synchroner und asynchroner Ausführung

Um diese Unterschiede zwischen den beiden Paradigmen zu veranschaulichen, wird dies anhand einer Client-Server-Anwendung mit einer an den Server angeschlossenen Datenbank und zwei Clients verbildlicht (siehe Abbildungen 9 und 10). Dabei ist der Server eine einfache Web-Applikation, welcher die Daten mittels SQL von der angeschlossenen Datenbank holt.

Senden nun beide Clients, in kurzem Zeitabstand zueinander, eine Anfrage an den Server, so kann der synchrone Server nur die Anfrage bearbeiten, die zuerst eintrifft - in diesem Fall die des Client 1. Dies geschieht deswegen, da der Main-Thread bzw. der für das Empfangen der Pakete zuständige Thread auf die Rückantwort des Datenbankservers wartet. Dadurch wird die Anfrage von Client 2 auf dem Server zurückgehalten und erst bearbeitet, wenn die erste Anfrage bearbeitet und zurückgesendet wurde. Dieses Verhalten skaliert schlecht mit mehreren, gleichzeitig eintreffenden Anfragen, da der sogenannte *ThreadPool*³, bei zu vielen, langandauernden Anfragen, seine Kapazitätsgrenze erreicht und somit keine neuen Anfragen / Aufgaben annehmen kann.

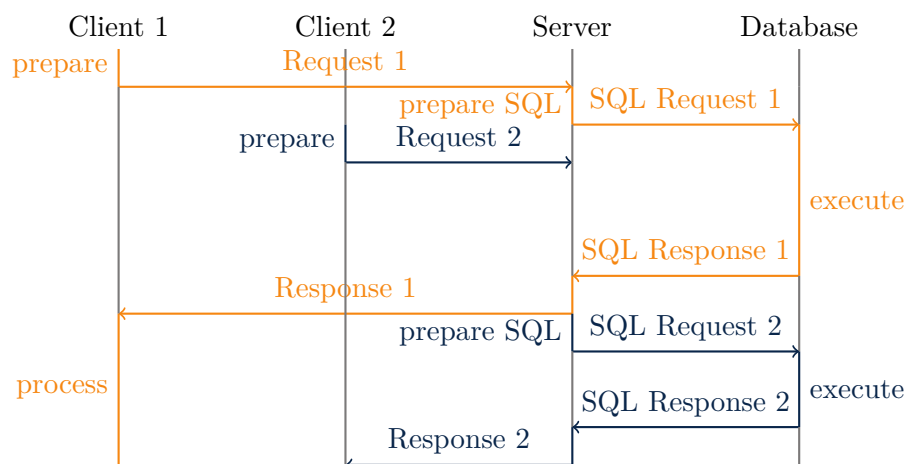


Abbildung 9: Sequenzdiagramm eines synchronen Servers

Setzt man anstelle eines synchronen Servers einen asynchronen Server ein, so ändert sich das beschriebene Szenario folgendermaßen: Trifft die Anfrage des Client 1 ein, wird diese, wie zuvor, sofort abgearbeitet. Dies geschieht jedoch auf einen anderen Thread,

³Entspricht einer Queue in der die zu bearbeitenden Aufgaben abgelegt werden.

damit der Main-Thread bzw. der für das Empfangen der Pakete zuständige Thread wieder frei wird. Wird nun vom Client 2 eine Anfrage verschickt, dann kann der Server diese entgegennehmen und auf einen weiteren, freien Thread verlagern und bearbeiten. Je nachdem welche Anfrage schneller vom Datenbankserver verarbeitet wird, in diesem Fall die des Client 1, wird der Main-Thread bzw. der Thread, der das Paket zuvor entgegengenommen hat, über die Beendigung der Datenbankanfrage benachrichtigt. Somit kann dieser Thread seinen Kontext synchronisieren und die Antwort an den Client 1 zurücksenden.

Dieses Verfahren skaliert deutlich besser mit steigender Anzahl von Anfragen, jedoch kann dies auch mehr CPU- und Speicherressourcen verursachen. Dies liegt daran, dass zum Aufrechterhalten des Zustandes eine asynchrone Zustandsmaschine konstruiert wird und der Kontextwechsel bei Beendigung des asynchronen Aufrufs mit den Main-Thread synchronisiert werden muss.

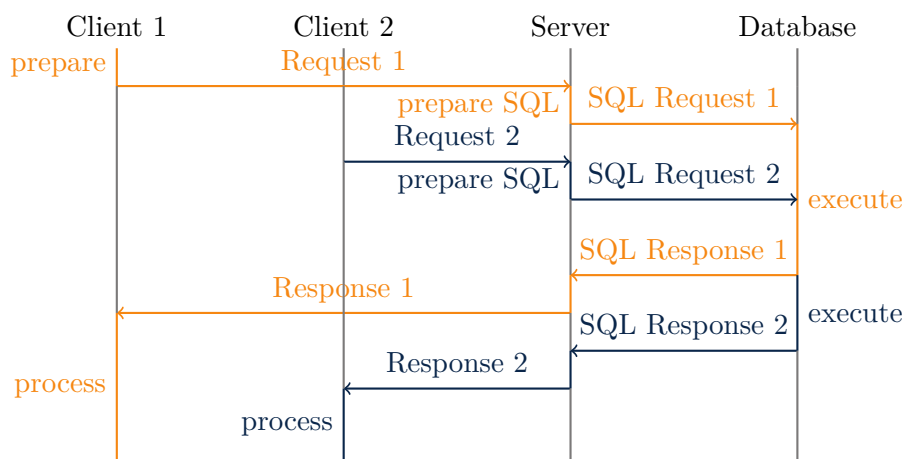


Abbildung 10: Sequenzdiagramm eines asynchronen Servers

Asynchronität in C#

Mit *Task-based Asynchronous Pattern* (TAP) ermöglicht Microsoft in C# Asynchronität. Dieses Pattern ermöglicht eine einfache Transformation von synchronen Code zu asynchronen Code ohne große Änderungen vornehmen zu müssen. Auch ist es von Haus aus im Sprachkonstrukt von C# integriert und kann somit ohne zusätzliche Konfiguration verwendet werden.

Dabei baut das asynchrone Ausführungsparadigma für C# auf folgende Komponenten auf:

- **Task:** Ermöglicht eine asynchrone Methode zu definieren.

- `Task<TResult>`: Ermöglicht einen Rückgabewert vom Typ `TResult` von einer asynchronen Methode zurückzugeben.
- `CancellationToken`: Ermöglicht es einen Aufruf einer asynchronen Methode vorzeitig zu beenden.
- `async/await`: Schlüsselwörter um asynchrone Methoden zu deklarieren und auszuführen.

Im nachfolgenden Listing 1 wird eine Instanz eines `DownloadClient` erzeugt. Dieser besitzt eine Methode `Download`, mit welcher der Inhalt einer Webseite heruntergeladen und als `string` zurückgegeben werden kann.

```

1 public string Download(Uri uri) {
2     var client = new DownloadClient();
3     var result = client.Download(uri);
4
5     return result;
6 }
```

Listing 1: Synchrone Methode in C#

Diese Vorgehensweise hat den Nachteil, dass die Funktion den Main-Thread bzw. den Thread, der diese Methode aufgerufen hat, solange blockiert bis der gesamte Webseiteninhalt der angegebenen URI heruntergeladen wurde. Bei UI-Anwendungen lässt sich dadurch die Oberfläche nicht mehr bedienen und friert ein. Bei Verwendung in einer Konsolen-Applikation ohne UI ist der Main-Thread bis auf Weiteres blockiert und somit können andere Programmteile, die nicht auf das Ergebnis dieser Methode angewiesen sind, nicht ausgeführt werden.

Damit der aufrufende Thread bzw. der Main-Thread nicht andauernd auf die Beendigung der Methode warten muss, kann man Events einsetzen, wie in Listing 2 dargestellt. Diese beenden vorzeitig die Exekution der Methode, indem der Aufrufer ein Objekt übergeben bekommt. Dieses Objekt repräsentiert den derzeitigen Zustand der Operation - also in diesem Fall das Herunterladen des Webseiteninhaltes. Dafür wird in Zeile 3 das `DownloadResult`-Objekt erzeugt. In Zeile 5 wird durch den Operator `+=` die nachfolgende anonyme Lambdafunktion `(content) => result.SetComplete(content)` als Beobachter des Events `client.DownloadComplete` registriert. Diese Lambdafunktion hat die Aufgabe den heruntergeladen Inhalt der Webseite in das `DownloadResult` zu setzen und als vollständig zu markieren. Dies passiert durch den Aufruf der Methode `SetComplete(string content)` auf dem `DownloadResult`-Objekt. Anschließend startet der `DownloadClient` die Operation mithilfe des Aufrufs `client.StartDownload(uri)`.

Ist der Client mit dem Download fertig, dann löst dieser das Event `DownloadComplete` aus und benachrichtigt somit alle darauf hörenden Empfänger - in diesem Fall die anonyme Lambdafunktion, über die Beendigung des Downloads.

Der aufzurufende Thread hat nun die Möglichkeit andere Funktionen auszuführen, solange der Download noch nicht beendet wurde.

```
1 public DownloadResult Download(Uri uri) {
2     var client = new DownloadClient();
3     var result = new DownloadResult();
4
5     client.DownloadComplete += (content) => result.SetComplete(content);
6     client.StartDownload(uri);
7
8     return result;
9 }
```

Listing 2: Eventbasierte Methode in C#

Als nächste Ausbaustufe der Funktion gibt es noch die asynchrone Variante. Hierbei wird der `DownloadClient` um eine asynchrone Downloadmethode, namentlich `DownloadAsync`, erweitert. Somit kann das Herunterladen des Webseiteninhaltes völlig asynchron und auf einem dezidiertem Thread passieren, ohne den Main-Thread bzw. aufzurufenden Thread zu beeinträchtigen. Dies kann jedoch nur solange ausgenutzt werden, solange keine andere Methode auf das Ergebnis der Downloadoperation angewiesen ist. Das Listing veranschaulicht diese asynchrone Methode und veranschaulicht auch die geringen Änderungen zur synchronen Implementierung dieser Methode (siehe Listing 3). Dieses Pattern ist auch als *Task-based asynchronous pattern*, abgekürzt TAP, bekannt.

```
1 public async Task<string> DownloadAsync(Uri uri, CancellationToken ct) {
2     var client = new DownloadClient();
3     var result = await client.DownloadAsync(uri, ct).ConfigureAwait(false);
4
5     return result;
6 }
```

Listing 3: Asynchrone Methode in C#

Ergänzend zu erwähnen ist, dass eine Methode immer durch die Schlüsselwörter `async`/`await` gekennzeichnet wird. Dabei wird nur `async` vorausgesetzt, wenn innerhalb der asyn-

chronen Methode auf ein Ergebnis einer anderen asynchronen Methode, mittels dem Schlüsselwort `await`, gewartet wird (siehe Zeile 3 im Listing 3).

1.3 Bekannte Implementierungen

Für fast jede Sprache gibt es eine Implementierung von CoAP. Die bekannteste von allen ist hierbei Californium für Java⁴. Andere Implementierungen gibt es, auszugsweise, für folgende Sprachen:

- C#: CoAP.NET
- Python: CoAPython
- C: libcoap
- Javascript: Copper
- Go: Go-Coap

Californium

Californium ist die bekannteste Implementierung von CoAP für Java. Diese Implementierung wird von der Eclipse Foundation gefördert und von mehr als 50 Entwicklern gepflegt. Aufgrund der aktiven Entwicklung und großen Abdeckung der jeweiligen RFC Standards für CoAP wurde Californium für verschiedene Sprachen portiert - z.B. mit CoAP.NET für .NET und C#.

CoAP.NET

CoAP.NET implementiert das *Constrained Application Protocol* in C# und wurde von der Organisation *smeshlink* bis Juli 2016 entwickelt [9]. Es ist ein C#-Port der CoAP-Implementation für Java *Californium*, jedoch ist es zurzeit unbetreut. Auch eine vollständige Unterstützung von asynchronen Techniken ist nicht in CoAP.NET gegeben. Dennoch zählt CoAP.NET zu den bekanntesten Implementationen im C#-Umfeld und wird auch in der Softwareentwicklungsfirma *World-Direct eBusiness solutions GmbH* in Sistrans, die auch mein derzeitiger Arbeitsgeber ist, verwendet. Da diese Bachelorarbeit in Kooperation mit dieser Firma entstanden ist, wurde der Hauptfokus auf CoAP.NET gelegt.

⁴<https://www.eclipse.org/californium/>

1.4 Industrielle Motivation und Alternativen

CoAP wird häufig im industriellen Umfeld verwendet, um Applikationen schaffen zu können, die mit eingeschränkten Geräten, wie 8-Bit-Mikrocontroller, über eine einfache, standardisierte Internetschnittstelle kommunizieren. Somit ermöglicht dies die einfache Integration von IoT-Geräten in bestehende oder neue Softwareapplikationen, um mit diesen Geräten zu interagieren, sie zu verwalten und zu steuern.

Aufgrund dessen, dass das *Constrained Application Protocol* ein einfaches zu implementierendes Protokoll darstellt und auf simple, leistungsschwache Geräte optimiert ist, im Gegensatz zum weitaus geläufigerem *Hypertext Transfer Protocol*, wird es für viele verschiedene Anwendungsszenarien verwendet. Einige von diesen werden nachfolgend beispielhaft aufgelistet:

- Hausautomatisierungen, wie von Kodali u. a. in deren Arbeit mit dem Titel „Implementation of Home Automation Using CoAP“ [3] beschrieben.
- Integration von kabellosen Sensoren mittels TinyOS von Ludovici, Moreno und Calveras [5].
- Applikation zur Nutzung von integrierten Sensoren in Transportcontainern von Kuladinithi u. a. [4].

Der Funktionsumfang von CoAP ist im Vergleich zu HTTP geringer, jedoch ist auch der Verwendungszweck vom *Constrained Application Protocol* ein anderer wie es bei HTTP der Fall ist. Das *Hypertext Transfer Protocol* kann universal eingesetzt werden und deckt viele verschiedene Anwendungsarten ab. Es wurde zwar hauptsächlich für die Übertragung von Webseiten über das Internet entwickelt, jedoch lässt es sich für Datenübertragung beliebiger Formate und Dateien einsetzen.

CoAP hat sich über die Jahre weiterentwickelt und neuere Techniken, wie Datenübertragung über TCP, Websockets und mittels TLS [1], als auch neuere Funktionen, wie blockweiser Transfer von Daten [7], implementiert. Jedoch bezieht sich diese Arbeit auf die erste Spezifikation von CoAP im Form des RFC 7252 [8], der von Shelby, Hartke und Bormann im Jahr 2014 definiert wurde.

Auch die Softwareentwicklungsfirma *World Direct eBusiness solutions GmbH*, der Kooperationspartner für diese Arbeit, verwendet CoAP in diversen Softwareprojekten, um Lösungen im IoT und Energiebereich anbieten zu können. Hier wird das Constrained Application Protocol dazu verwendet, um mit eigens entwickelten Hardwarecontrollern zu kommunizieren, die vor allem im Bereich der Regelernergie⁵ zum Einsatz kommen.

⁵Die benötigte Energie um die Unter- oder Überproduktion von Strom im Stromnetz auszugleichen.

1.5 Ziel der Arbeit

Die Forschungsfrage, welche diese Bachelorarbeit zu beantworten versucht, lautet: *Hat eine asynchrone Implementation eines Servers Einfluss auf dessen Durchsatzrate?*

Als erste Möglichkeit diese Forschungsfrage zu beantworten, versuchte man für das Open-Source-Projekt CoAP.NET, das die bekannteste Implementierung der CoAP-Spezifikation für C# darstellt, eine asynchrone Schnittstelle zu implementieren. Diese sollte dann abschließend mit der bereits bestehenden synchronen Schnittstelle verglichen werden, um zu evaluieren, ob hier die implementierte Asynchronität Vorteile bringt. Diese Idee musste jedoch verworfen werden, da zu viele Änderungen an der Codebasis von CoAP.NET notwendig gewesen wären, um das angestrebte Ziel zu erreichen. Dies hätte den zeitlichen Rahmen als auch den Arbeitsaufwand dieser Bachelorarbeit deutlich überschritten.

Aus diesem Grund entschied man sich dies zu umgehen, indem eine komplette Neuentwicklung des *Constrained Application Protocols* für C#, mit modernen Technologien und Patterns, entwickelt wurde. Hierbei wurde bei der Entwicklung auf eine synchrone als auch asynchrone Schnittstelle Wert gelegt, damit diese miteinander verglichen werden können. Dadurch sollte ermittelt werden, inwiefern eine asynchrone Implementation eines (CoAP-)Servers einen Einfluss auf dessen Durchsatzrate hat.

Das Resultat dieser Arbeit ist im Github Repository der *Firma World Direct eBusiness solutions GmbH* [world-direct/CoAP.NET](https://github.com/world-direct/CoAP.NET) einsehbar. Dies hat den Grund, da die entwickelte Bibliothek auch darüber hinaus firmenintern weiterverwendet werden soll. Jegliche Weiterentwicklungen an dieser Software passieren somit auf diesem Repository und unter der Führung von Mitarbeitern der *World Direct eBusiness solutions GmbH*.

2 Implementierung

Das Constrained Application Protocol wird für diese Bachelorarbeit in der Programmiersprache C# implementiert. Dabei wird die Implementierung nach dem .NET Standard 2.1 entwickelt. Diese Vorgehensweise ermöglicht es die Bibliothek lauffähig unter den folgenden .NET Versionen und Laufzeitumgebungen zu verwenden: .NET 5.x und .NET Core 3.x. Dadurch sind alle aktuellen und zukünftig unterstützten Laufzeitumgebungen abgedeckt. Durch .NET Core ist es möglich neben Windows auch für andere Betriebssysteme zu entwickeln und somit ist auch die Bibliothek unter Linux und Mac lauffähig.

Die Asynchronität ist, wie schon beschrieben, durch das Sprachkonstrukt von C# gegeben. Somit kann mittels *Task-based Asynchronous Pattern* (TAP) mit wenig Aufwand eine asynchrone API bereitgestellt werden. Zusätzlich wird die Drittanbieterbibliothek *Task Parallel Library*, auch TPL abgekürzt, genutzt. Diese Bibliothek erlaubt es ein sogenanntes *Dataflow-Mesh* (vergleichbar mit einer Queue, nur mit erweiterter Funktionalität und mit Asynchronität als Hauptfokus) aufzubauen und somit eine asynchrone Datenverarbeitung innerhalb einer Applikation zu ermöglichen. Dabei besteht ein solches *Dataflow-Mesh* aus sogenannten *Dataflow Blocks*. Die vordefinierten Dataflow Blocks geben entweder die Möglichkeit Daten zu puffern (*Buffering Blocks*) oder zu verarbeiten (*Execution Blocks*). Als weitere Möglichkeit bietet TPL an, eigene Dataflow Blocks zu definieren, indem man die entsprechenden Basisklassen oder Interfaces implementiert werden.

Die erforderlichen Grundlagen und *Best Practices* bezüglich asynchroner Programmierung, vor allem im Speziellen bei Verwendung von *Dataflow Blocks*, wurden aus dem Fachbuch *Concurrency in C# Cookbook: Asynchronous, Parallel, and Multithreaded Programming* von Cleary, Microsoft-MVP und Experte in asynchroner Programmierung, entnommen.

Gefundene Fehler

Nachfolgend sind hier Bugs aufgeführt, die innerhalb dieser Bachelorarbeit in CoAP.NET gefunden wurden:

- Blockweiser Transfer propagiert keine Fehler. Das heißt, dass eine Übertragung von mehreren UDP-Paketen nicht gestoppt wird, wenn ein Statuscode von 4.08 (Request Entity Incomplete) oder 4.13 (Request Entity Too Large) zurückgegeben wird.
- Der Client stoppt die Neuversendung von Anfragen nicht, obwohl eine Antwort mit der passenden MessageId zurückgesendet wurde.

2.1 Struktur der Applikation

Die Applikation gliedert sich in folgende Komponenten:

1. **Transports:** Eine Transport-Klasse übernimmt das Senden und Empfangen von CoAP-Nachrichten auf dem jeweiligen Protokoll. Zum Beispiel ist die *UdpTransport*-Klasse verantwortlich für das Senden und Empfangen von CoAP-Nachrichten, die mittels UDP übertragen wurden.
2. **Channels:** Ein *Channel* repräsentiert eine aktive Verbindung zwischen einem CoAP-Client und CoAP-Server über ein beliebiges Protokoll. Für UDP gibt es z.B. eine *UdpChannel*-Klasse, die auf einem vordefinierten Port auf UDP-Pakete horcht und über diesen Antworten an den jeweiligen Client zurücksendet.
3. **Serializers:** Ein *Serializer* bietet Methoden für das Deserialisieren als auch Serialisieren von CoAP-Nachrichten, für ein bestimmtes Nachrichtenformat und/oder eine CoAP-Version an. Dabei erhalten diese die Daten entweder von einem der *Channels* oder von einer *Ressource*, die auf dem Server registriert ist.
4. **Handlers:** Ein *Handler* kümmert sich um die Verarbeitung und Weiterleitung von CoAP-Request an die jeweiligen Ressourcen oder von CoAP-Responses an den *Serializer*.
5. **Ressourcen:** Eine *Ressource* ist vergleichbar mit einem HTTP- bzw. Controller-Endpunkt. Diese registriert sich beim Server unter einer definierten URI und gibt an, welche Methoden (GET, POST, PUT, DELETE) sie anbietet.

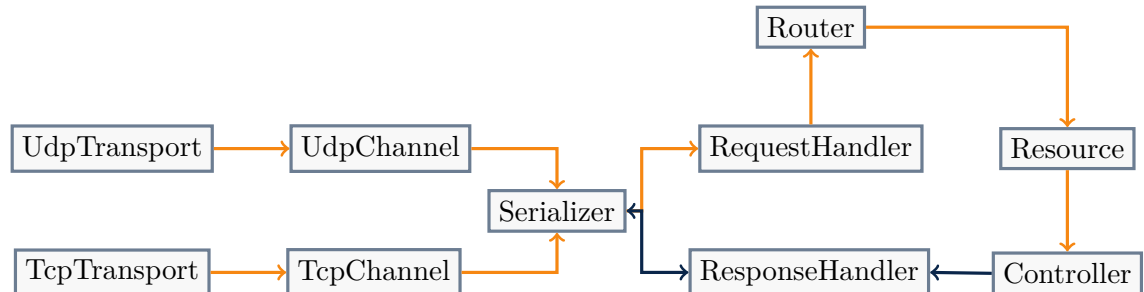


Abbildung 11: Überblick Softwarearchitektur

Dabei stellt jede Komponente einen *Dataflow Block* dar. Somit kann eine asynchrone Weiterleitung zwischen den einzelnen Komponenten sichergestellt werden. Dies geschieht dadurch, dass die einzelnen Blöcke miteinander verknüpft und auf ankommende Daten in einer asynchronen Weise gewartet werden. Dadurch können Daten, sobald diese verfügbar sind, umgehend verarbeitet werden. Auch übernimmt das *Dataflow-Mesh* die Verantwortung für die angewendete Parallelität und Synchronität der Daten. Somit kann eine flexible und anpassbare Verarbeitungskette implementiert werden, die vollkommen Asynchron arbeitet.

Die API des Serializers ist inspiriert von der API des *System.Text.Json Serializers*⁶ für JSON Dateien, der von Microsoft entwickelt wurde. Der *CoapMessageSerializer* bietet dabei folgende Schnittstellen an:

- `void CoapMessage Deserialize(ReadOnlySpan<byte> value);`
- `Task<CoapMessage> DeserializeAsync(Stream value, CancellationToken ct);`
- `void byte[] Serialize(CoapMessage message);`
- `Task SerializeAsync(Stream stream, CoapMessage message, CancellationToken ct);`

2.2 Besonderheiten

Nachfolgend werden Codeausschnitte und Teile des Projekts aufgelistet, die besondere Erwähnung finden. Wie schon beschrieben, wurde bei der Neuimplementierung von CoAP.NET darauf geachtet sowohl eine synchrone als auch eine asynchrone Schnittstelle zur Verfügung zu stellen.

Struktur der Optionen

Beim Verarbeitungsprozess der Optionen einer CoAP-Nachricht wurde darauf Wert gelegt, dass sich dieser Prozess auch für zukünftige CoAP-Versionen mit eventuell neuen Optionen eignet. Um dieses Ziel zu erreichen, wurde die Struktur der CoAP-Optionen in einer strikten Hierarchie, siehe Bild 12 im Code abgebildet.

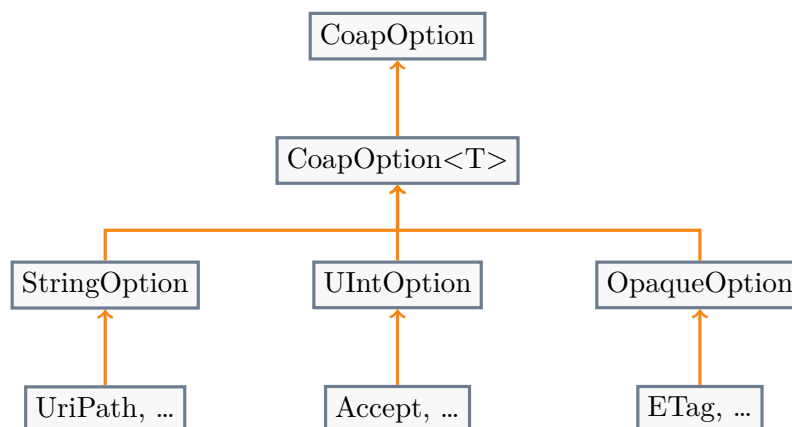


Abbildung 12: Hierarchie der CoAP-Optionen im Code

⁶Dokumentation zum System.Text.Json Serializer

Die oberste Basisklasse `CoapOption`, von der alle CoAP-Optionen erben, beinhaltet alle Basisfunktionalitäten einer Option, wie sie im RFC 7252 von Shelby, Hartke und Bormann definiert ist. Als Nächstes gibt es die Basisklasse `CoapOption<T>` die den Wert der Option mit dem Typ `T` hält. `T` kann hierbei ein beliebiger Datentyp oder Klasse sein.

Für die im RFC 7252 von Shelby, Hartke und Bormann definierten, möglichen Werte die eine CoAP-Option besitzen kann, also *string*, *unsigned integer* oder *opaque*, gibt es die jeweilige dazugehörige Basisklasse:

- `StringOption`: Für CoAP-Optionen mit den Datentyp `string` als Wert. Das bedeutet `StringOption` erbt von der Basisklasse `CoapOption<string>`.
- `UIntOption`: Für CoAP-Optionen mit den Datentyp `uint` als Wert. Das bedeutet `UIntOption` erbt von der Basisklasse `CoapOption<uint>`.
- `OpaqueOption`: Für CoAP-Optionen mit einem Byte-Array als Wert. Das bedeutet `OpaqueOption` erbt von der Basisklasse `CoapOption<ReadOnlyMemory<byte>>`.

Die konkrete Implementation der einzelnen CoAP-Optionen erben von einer der drei zuvor genannten Basisklassen. Das folgende Listing 4 zeigt dies beispielsweise anhand der Optionen-Klasse `UriHost`.

Mit dieser Struktur können einfach neue CoAP-Optionen in den Programmfluss eingebunden werden, in dem man die entsprechenden Basisklassen implementiert. Diese Struktur erleichtert auch das Lesen und Schreiben von Optionen, in dem der `OptionReader` und `OptionWriter` sich auf eine der beiden Basisklassen `CoapOption` oder `CoapOption<T>` beruft.

Für das Lesen von Optionen wird das Fabrik-Modell (Factory-Pattern) verwendet, um Optionen anhand der gelesenen Daten zu erstellen. Diese Fabriken werden durch das Interface `IOptionFactory` definiert, das die Nummer der Option und eine Methode `Create` enthält. Dadurch kann der `OptionReader` die Daten lesen, die passende Fabrik über die Nummer aussuchen und durch die `Create`-Methode die Option erstellen.

Das Listing 5 erledigt diesen Vorgang synchron mithilfe einer `while`-Schleife, die den nachfolgenden Verarbeitungsprozess ausführt, solange es Daten zu verarbeiten gibt. Wie im Kapitel 1.1 schon erklärt, gibt es zwei Bedingungen, die das Ende einer Serie von Optionen in einer CoAP-Nachricht markieren:

- Es sind keine Daten zum Lesen verfügbar (EoF).
- Der Payloadmarker, also ein Byte voller logischer Einsen, in Hexadezimal als `0xFF` geschrieben, wird gelesen.

Treten keine der beiden Bedingungen ein, dann wird im nächsten Schritt zuerst die Kopfzeile der Optionen, also das erste Byte, in dem Daten für die weitere Verarbeitung der Option beinhaltet sind, gelesen. Abhängig davon, welche Daten aus der Kopfzeile erfasst wurden, werden optional noch die erweiterte *Option Delta* und *Option Length* konsumiert, falls vorhanden. Abschließend wird dann noch der Wert der Option, anhand

der Typisierung der Option, sprich als **string**, **int** oder **bytes**, interpretiert. Die Klasse *OptionData* beinhaltet diese Informationen und stellt einen einfachen Zugriff auf diese Daten dar. Diese *OptionData*-Klasse wird durch den Methodenaufruf in Zeile 8 erzeugt. Diese Methode *ReadOptionData* liefert, wie schon erwähnt, die *OptionData* und, über das **out**-Schlüsselwort, auch die Anzahl der gelesenen Bytes zurück.

Im nächsten Schritt (Zeile 13) wird die Nummer der aktuell gelesenen Option für den nächsten Lesevorgang zwischengespeichert, da diese als Berechnungsgrundlage für die nächste *Optionen Number* verwendet wird. In Zeile 14 werden die Anzahl an gelesenen Bytes zur Variable *offset* aufsummiert. Somit entspricht die Variable *offset* der derzeitigen Position in dem Set an Daten.

In den Zeilen 16 - 19 wird die passende *IOptionFactory* ausgewählt, falls es solch eine gibt und sonst wird als Fallback die *UnrecognizedOptionFactory* gewählt, damit nicht implementierte Optionen verarbeitet werden können. Die *UnrecognizedOptionFactory* erzeugt eine *UnrecognizedOption*, die nur die gelesenen Daten beinhaltet.

In Zeile 21 wird die Option mittels der gewählten *IOptionFactory* und der *OptionData* erzeugt und in Zeile 22 zur *OptionCollection* hinzugeführt. Wenn keine Daten mehr verfügbar sind oder der Payloadmarker gelesen wurde, wird die *OptionCollection* und der bisherige Offset zurückgegeben.

Für das asynchrone Lesen von Optionen wird ein ähnlicher Prozess durchlaufen. Der einzige Unterschied liegt darin, dass alle vorher synchron verwendeten Methoden durch ihr asynchrones Gegenstück ersetzt werden. Deshalb wird auch nicht näher auf das Listing 6, das diesen Vorgang repräsentiert, eingegangen.

Für das synchrone Schreiben von Optionen wird ein ähnliches Modell verfolgt, wie zuvor beschrieben beim synchronen Lesen von Optionen. Hierbei verwendet der *OptionWriter* eine Kollektion von *IOptionSerializer*, die eine Instanz von *CoapOption* in Bytes, mittels der Methode **Serialize**, umwandelt. Da dieser Code keine große Komplexität aufweist, wird auf eine detaillierte Beschreibung des Codes verzichtet und nur auf diesen im Listing 8 verwiesen.

Dasselbe gilt auch für das asynchrone Schreiben von Optionen, wie es im Codebeispiel 8 programmiert wurde.

CoAP Code

Die Struktur von CoAP Codes ähnelt dem vorher im Kapitel 2.2 beschriebenen Konstrukt der CoAP Optionen und wird durch das Bild 13 veranschaulicht. Alle CoAP Codes erben von der Basisklasse *CoapCode*, die grundlegende Eigenschaften, wie den Class- und Detail-Bestandteil des CoAP Codes, beinhalten. Hier spaltet sich die Struktur in die Klassen *RequestCode* und *ResponseCode* auf, die jeweils CoAP-Nachrichten in Anfragen (Requests) und Antworten (Responses) unterteilt.

Request Codes im CoAP-Kontext bestehen zu einem Großteil aus den sogenannten Method Codes - sprich GET, POST, PUT und DELETE. CoAP Codes mit einem *Class*-Wert von 0 entsprechen einem Request. Hier stellt eine CoAP-Nachricht, die als *Empty Message* deklariert wurde, einen Sonderfall dar, welcher als Code 0.00 definiert wird. Eine Empty Message markiert hierbei eine CoAP-Nachricht jedoch nicht als Request oder Response, sondern als Nachricht, die folgende Eigenschaften besitzt:

- Der Code ist 0.00.
- Die Tokenlänge ist auf 0 gesetzt.
- Die CoAP-Nachricht darf keine Bytes nach dem *Message ID*-Feld besitzen.

Response Codes können in Client Response Codes (4.XX), Server Response Codes (5.XX) und Successful Response Codes (2.XX) unterteilt werden. Dabei stellen die jeweiligen Unterkategorien, wie ihre Namen schon angeben, einen Response Code für eine jeweilige Komponente dar.

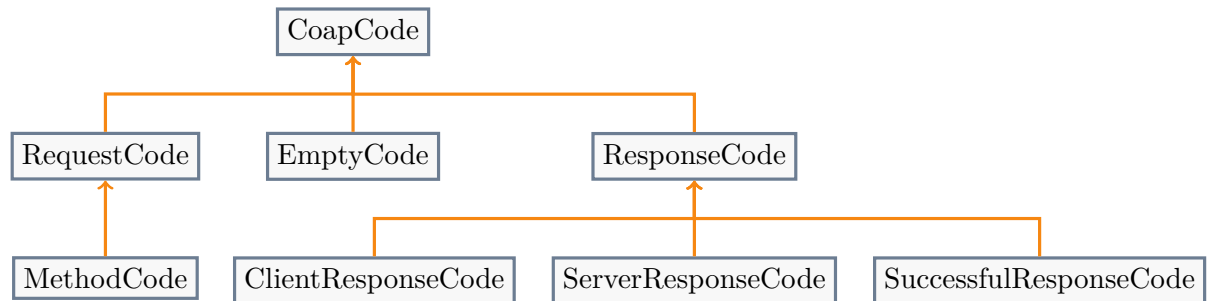


Abbildung 13: Hierarchie der CoAP-Codes

3 Messung

Zur Messung der Performance und Durchsatzrate der Implementierung betrachten wir nur den Serializer, da auf der Empfangsseite auf die bereits bestehenden Implementierungen der Sockets zur Datenübertragung von den UDP- bzw. TCP-Paketen für .NET gesetzt wird. Dabei wird auf TCP als Übertragungsweg gesetzt, da sich bei UDP die Paketgröße nicht verändern lässt bzw. auf maximal 2^{16} Bits, das 65 535 Bytes entspricht, begrenzt ist. Somit können wir die Paketgröße der CoAP-Nachricht, aufgrund der fehlenden Implementierung des blockweisen Transfers [7], nicht beliebig erhöhen. Um hier jedoch ein Szenario zu kreieren, indem wir auch sehr lange Nachrichten übermitteln können, wurde die Übertragung von CoAP-Nachrichten über TCP nach dem RFC 8323 von Bormann u. a. [1] implementiert.

3.1 Nachrichtenverarbeitung

In diesem Szenario wird die Zeit der Nachrichtenverarbeitung auf dem Server gemessen. Dabei sieht die Messung folgenden Ablauf vor:

- Die Nachrichten werden von einem Client über das jeweilige Protokoll über TCP an den CoAP-Server versendet.
- Dabei wird eine langsame Übertragung simuliert, indem nur ein Byte alle 250 Millisekunden, das einer Geschwindigkeit von 4 B/s entspricht, verschickt wird.
- Der Server wird für TCP in der asynchronen Variante auf den Port 5683 und für die synchrone Variante auf Port 5684 angesprochen.
- Es wird die Zeit gemessen, wie lange das Verarbeiten der Nachricht gedauert hat.
- Die Zeitmessung wird am Client gestartet, sobald der Client mit dem Versenden der Nachricht beginnt. Gestoppt wird die Messung, sobald der Server die Nachricht deserialisiert hat.
- Dieser Ablauf wird mehrere Male wiederholt und der Durchschnittswert dadurch ermittelt.

BenchmarkDotnet

BenchmarkDotnet ist ein Software-Tool das vom dotnet-Team⁷ entwickelt und zur Verfügung gestellt wird. Mit diesem Tool lassen sich automatisierte Laufzeittests von einem bestimmten Codeteil oder sogar von einem ganzen Programm erzeugen.

BenchmarkDotnet führt hierfür den ausführenden Codeteil in mehreren Durchläufen aus und misst bei jedem Durchlauf verschiedene Parameter, welche vom Nutzer festgelegt

⁷Open-Source-Abteilung bei Microsoft für das .NET Ökosystem

werden. Dabei wird standardmäßig die durchschnittliche Laufzeit, die Fehlertoleranz und Standardabweichung ermittelt. Auch können Parameter wie allokierten Speicher, Codeverlauf (Tracer), Anzahl von Lock's (Semaphore), Anzahl verarbeiteter Aufträge im Threadpool und vieles mehr aufgezeichnet werden.

Die Ergebnisse werden dabei in verschiedene Formate exportiert. Standardmäßig werden diese in CSV, HTML und Markdown exportiert, wobei aber auch JSON, XML oder als grafische Visualisierung in RPlot zur Verfügung stehen.

3.2 Serialisierung und Deserialisierung

In diesem Anwendungsbereich wird die Verarbeitungszeit der synchronen als auch asynchronen Variante der Serialisierungs- bzw. Deserialisierungsmethode gemessen. Dies wird mittels der Bibliothek BenchmarkDotnet⁸ durchgeführt. BenchmarkDotnet ist dabei ein Tool, dass es erlaubt nativ in C# eine vordefinierte Methode bzw. einen bestimmten Teil eines Programms einem Benchmark zu unterziehen. Somit wird ermittelt, wie sich diese beiden Varianten, unabhängig von der Netzwerkgeschwindigkeit, verhalten. Dabei wird sowohl die ermittelte Durchschnittszeit als auch der Speicherverbrauch mittels BenchmarkDotnet gemessen.

Daneben können wir in BenchmarkDotnet mehrere verschiedene Durchläufe konstruieren, welche sich in bestimmten Parametern unterscheiden. Um dabei die Länge der CoAP-Nachricht zu variieren, wird die Anzahl der Options und die Länge der Payload verändert, wobei sich die beiden Parameter in folgenden Schritten: 0, 1024, 4096, 65536, 131072, verändern. Folgende Fälle sollten somit abgedeckt sein:

- Eine CoAP-Nachricht nur mit dem Header und dem Token.
- Eine CoAP-Nachricht nur mit Optionen und keiner Payload.
- Eine CoAP-Nachricht nur mit einer Payload und keinen Optionen.
- Eine CoAP-Nachricht sowohl mit Optionen als auch einer Payload.

Auch sollte erkennbar sein, unter welchen Umständen die synchrone oder asynchrone Verarbeitung besser abschneidet.

Um eine Methode als Benchmark für BenchmarkDotnet zu kennzeichnen, muss das Attribut `[Benchmark]` der Methode hinzugefügt werden. Im Listing 9 wird ein solcher Benchmark beispielhaft implementiert.

In der Main-Methode muss diese Klasse nun nur bei BenchmarkDotnet zur Ausführung registriert werden. Dies wird, wie im Listing 10 ersichtlich ist, über die `main`-Methode des Benchmark-Programms vorgenommen.

⁸<https://benchmarkdotnet.org/index.html>

Nachrichtengenerierung für Benchmark

Die CoAP-Nachricht, die für den Benchmark benutzt werden, folgt folgendem Schema (x = Anzahl der Options; y = Länge der Payload in Bytes):

- Die CoAP-Version ist immer auf 1.
- Der Typ der Nachricht ist immer Acknowledgement.
- Die Tokenlänge ist bei acht Bytes und wird zufällig generiert.
- Der Code ist CREATED (2.01).
- Die MessageId wird zufällig generiert.
- Es werden x-mal Optionen vom Typ UriPath erstellt.
- Die Payload wird zufällig generiert und ist y Bytes lang.

Die CoAP-Nachricht wird für jeden Durchgang neu generiert und jedem einzelnen Benchmark übergeben.

3.3 Messaufbau

Die Messungen werden auf einem Rechner mit AMD Ryzen 5 2600 (6 Kerne und 12 Threads) als CPU und mit einem Arbeitsspeicher von 16 GB durchgeführt.

Die Netzwerkübertragung findet lokal statt, sprich über die Adresse 127.0.0.1 (Loopback / localhost). Mit dem Kommandozeilenbefehl `start /affinity 1 Server.exe` wird der Server nur auf einem einzelnen Kern ausgeführt, damit nur die reine Leistung des Servers betrachtet wird und nicht durch das Scheduling des Rechners bzw. der CPU verfälscht wird.

Für das Szenario der Serialisierung und Deserialisierung werden keine speziellen Einstellungen vorgenommen, da hier die Standardeinstellungen von BenchmarkDotnet verwendet werden.

3.4 Messergebnisse (Nachrichtenübertragung)

Für die Nachrichtenübertragung wurde die Anzahl der Optionen auf 100 und die Größe der Payload auf 100000 limitiert. Größere Werte haben keine bemerkenswerte Erkenntnis gebracht und wurde zwecks Übersichtlichkeit weggelassen.

Start	Ende	Laufzeit
2021-11-11T10:17:31.0524370+01:00	2021-11-11T10:19:16.5499204+01:00	00:01:45.4974834
2021-11-11T10:22:12.9707397+01:00	2021-11-11T10:23:58.8623836+01:00	00:01:45.8916439
2021-11-11T10:28:32.5175832+01:00	2021-11-11T10:30:18.3405198+01:00	00:01:45.8229366
2021-11-11T10:32:01.7660773+01:00	2021-11-11T10:33:47.4505300+01:00	00:01:45.6844527
2021-11-11T10:35:34.7139187+01:00	2021-11-11T10:37:20.3576467+01:00	00:01:45.6437280

Tabelle 2: Asynchrone Übertragung mit 100 Options und mit einer Payload von 100 Bytes.

Berechnet man nun die durchschnittliche Laufzeit aus den Ergebnissen in Tabelle 2, in der eine CoAP-Nachricht mit 100 Options und einer Payload von 100 Bytes übermittelt wurde, mittels folgender Formel. Diese Formel wird auch für die nachfolgenden Messungen verwendet, um deren durchschnittliche Laufzeit zu ermitteln.

$$t_{\text{durchschnitt}} = \frac{01:45.4974834}{5} + \frac{01:45.8916439}{5} + \frac{01:45.8229366}{5} + \frac{01:45.6844527}{5} + \frac{01:45.6437280}{5} = 105,71 \text{ Sekunden} \quad (1)$$

Daraus ergibt sich für die asynchrone Übertragung eine durchschnittliche Laufzeit von 105,71 Sekunden. Wird diese nun auch für die Messdaten in Tabelle 3 berechnet, ergibt sich hier eine durchschnittliche Laufzeit von 105,61 Sekunden.

Damit ist die synchrone Übertragung um 100 Millisekunden schneller als die asynchrone Übertragung. Durch die geringe Datenmenge von 25813 Bytes für die gesamte Nachricht ist dies nicht überraschend. Bei der synchronen Übertragung wird so lange gewartet, bis die komplette CoAP-Nachricht übertragen wurde. Im Gegensatz dazu wird bei der asynchronen Übertragung die Daten sofort verarbeitet, wenn diese verfügbar sind. Da jedoch immer nur vier Bytes pro Sekunde versendet werden, ist der Mehraufwand zum Erzeugen der asynchronen Zustandsmaschine zu groß und damit unvorteilhaft für die Performanz.

Start	Ende	Laufzeit
2021-11-11T10:17:31.0525437+01:00	2021-11-11T10:19:16.0534992+01:00	00:01:45.0009555
2021-11-11T10:22:12.9707397+01:00	2021-11-11T10:23:58.8617438+01:00	00:01:45.8910041
2021-11-11T10:28:32.5175895+01:00	2021-11-11T10:30:18.3389314+01:00	00:01:45.8213419
2021-11-11T10:32:01.7660713+01:00	2021-11-11T10:33:47.4500247+01:00	00:01:45.6839534
2021-11-11T10:35:34.7139140+01:00	2021-11-11T10:37:20.3586945+01:00	00:01:45.6447805

Tabelle 3: Synchrone Übertragung mit 100 Options und mit einer Payload von 100 Bytes.

Entnimmt man die Messdaten aus der Tabelle 6, die einen Messdurchlauf für eine asynchrone Übertragung einer CoAP-Nachricht mit 100 Options und einer Payload von 1000 Bytes darstellt, ergibt sich eine durchschnittliche Laufzeit von 341,01 Sekunden.

Für eine synchrone Übertragung derselben Nachricht, Messdaten aus Tabelle 7 entnommen, ergibt sich eine durchschnittliche Laufzeit von 341,00 Sekunden. Damit beträgt der Abstand zwischen synchron und asynchron nur 10 Millisekunden.

Erhöht man die Payload auf 10000 Bytes, ergibt sich für eine asynchrone Übertragung eine durchschnittliche Laufzeit von 2688,88 Sekunden und für eine synchrone Übertragung eine durchschnittliche Laufzeit von 2688,95 Sekunden. Die jeweiligen Daten wurden dabei von den Tabellen 8 und 9 entnommen. Damit erhöht sich der Abstand auf 70 Millisekunden, das den Trend entspricht, dass mit steigender Payloadgröße, der Vorteil der Asynchronität tragender wird.

3.5 Messergebnisse (Deserialisierung und Serialisierung)

Die detaillierten Ergebnisse zu den Messungen der Deserialisierung und Serialisierung sind im Anhang aufgeführt. Nachfolgend werden nur direkte Vergleiche zwischen der synchronen und asynchronen Implementation der jeweiligen Operation vorgenommen und auf diese verwiesen. Durch das Tool BenchmarkDotnet werden die Messergebnisse, die in diesen Tabellen aufgeführt sind, ermittelt. Dabei ergibt sich folgende Legende für die Messtabellen:

- $n_{Optionen}$: Anzahl der Optionen.
- $l_{Payload}$: Die Größe bzw. Länge der CoAP-Payload in Bytes.
- t_X : Die durchschnittliche Laufzeit der Methode X in μs .

Wie in Tabelle 4 ersichtlich, arbeitet die *DeserializeAsync*-Methode immer schneller, als ihr synchroner Gegenpart, solange keine CoAP-Optionen sich in der CoAP-Nachricht befinden. Sobald sich dies jedoch ändert, ist der Mehraufwand für den Aufbau und Verwaltung der asynchronen Zustandsmaschine nachteilig für die asynchrone Methode. Dies lässt sich dadurch erklären, dass die Größe der Optionen zu klein ist, um von der asynchronen Verarbeitung, mit ihren effizienteren Prozessen, zu profitieren.

Vergrößert man jedoch die Größe der Payload, dann profitiert der Deserialisierungsvorgang von der angebotenen Asynchronität. Obwohl sich der Unterschied in einigen 100 Mikrosekunden bewegt, kann es bei steigender Anzahl gleichzeitig eintreffender Nachrichten eine merkbare Auswirkung haben, da bei schnellerer Verarbeitung, auch schneller die nächste Nachricht abgearbeitet werden kann.

Detaillierte Daten können in den Tabellen 11 und 12 im Anhang eingesehen werden. Dort ist erkenntlich, dass der Speicherverbrauch bei der *DeserializeAsync*-Methode mit steigender Größe der Payload gleichbleibenden ist, im Gegensatz zum synchronen Gegenpart.

Eine mögliche Erklärung für den gleichbleibenden Speicherverbrauch von *DeserializeAsync* ist, dass dieser mittels eines *Stream* arbeitet, von dem ohne Allokation direkt gelesen

$n_{Optionen}$	$l_{Payload}$ [B]	$t_{DeserializeAsync}$ [μ s]	$t_{Deserialize}$ [μ s]
0	0	2.231	1.679
0	1024	3.298	3.683
0	4096	5.077	6.570
0	65536	64.490	98.533
0	131072	111.011	175.433
1024	0	1440.076	1030.051
1024	1024	1402.291	1014.093
1024	4096	1416.083	1027.275
1024	65536	1592.417	1631.388
1024	131072	1817.051	1781.418
4096	0	8416.154	6477.162
4096	1024	8147.070	6711.475
4096	4096	8737.255	6558.996
4096	65536	9057.536	6924.552
4096	131072	9543.723	7530.808
65536	0	203180.664	166612.138
65536	1024	198211.164	190898.380
65536	4096	201316.657	190316.150
65536	65536	200257.897	195944.825
65536	131072	201631.362	170582.006
131072	0	372129.381	286105.657
131072	1024	372065.709	285783.220
131072	4096	368936.160	282355.564
131072	65536	366150.575	292059.675
131072	131072	370709.611	304488.343

Tabelle 4: Vergleich DeserializeAsync und Deserialize

werden kann. Im Gegensatz dazu ist es bei *Deserialize* notwendig, die einzeln ausgelesenen Daten in eine Art Zwischenspeicher zu schreiben, um sie verarbeiten zu können. Dies übernimmt dabei die Klasse *PooledMemoryBufferWriter*, die ein gepuffertes Schreiben von Daten auf einen Speicherbereich erlaubt. Dieser Speicherbereich ist dabei durch die Klasse `Memory<byte>` modelliert. Damit der Aufrufer den *PooledMemoryBufferWriter* beschreiben kann, muss dieser sich Speicher von diesem „ausleihen“ und anschließend mit den gewünschten Daten beschreiben. Zum Schluss muss dem *PooledMemoryBufferWriter* die Anzahl der geschriebenen Bytes mitgeteilt werden, damit die Position des *PooledMemoryBufferWriters* weitergeschoben werden kann. Dies wird im Listing 11 veranschaulicht.

Der Nachteil des *PooledMemoryBufferWriters* ist, dass dieser seinen zur Verfügung stehenden Speicherbereich vergrößern muss, wenn die Kapazität erschöpft ist. Dabei vergrößert sich dieser so weit, damit der Puffer die zu schreibenden Daten aufnehmen kann.

Auch ist der *PooledMemoryBufferWriter* eine Eigenimplementierung, jedoch wurde im Laufe der Recherchen für diese Arbeit eine ähnliche Implementation von Microsoft⁹ gefunden, die dieses Problem möglicherweise besser handhabt, als die derzeitige Implementation. Da dies jedoch einen zu großen Aufwand darstellt, wurde darauf verzichtet.

Die detaillierten Messergebnisse von *Dserialize* und *DeserializeAsync* können in den Tabellen 11 und 12, als auch visualisiert durch die Bilder 14 und 15, eingesehen werden.

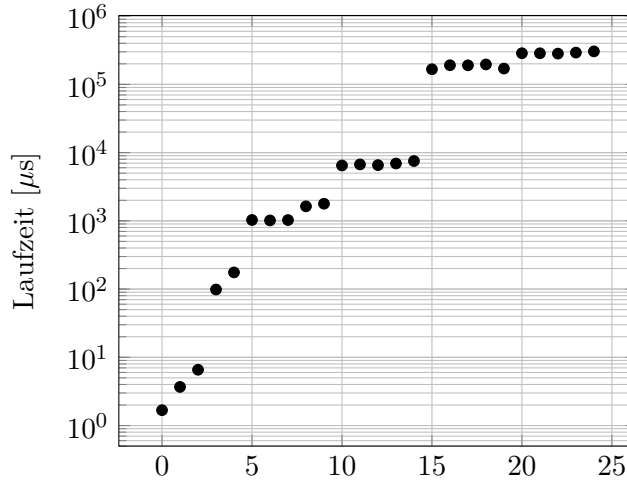


Abbildung 14: Messergebnisse von Deserialize-Methode als Diagramm

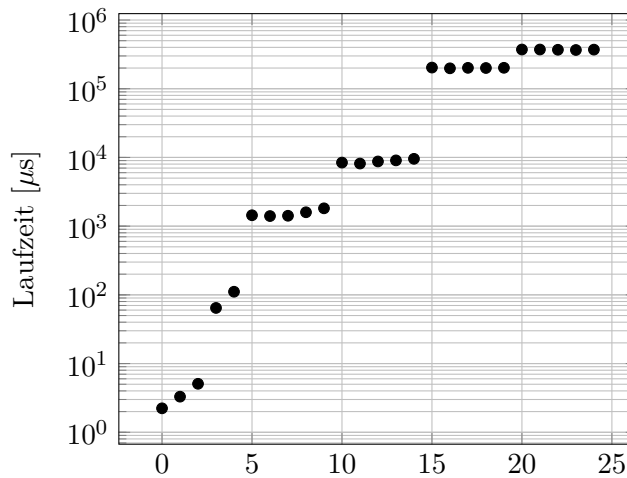


Abbildung 15: Messergebnisse von DeserializeAsync-Methode als Diagramm

Betrachtet man nun die Ergebnisse, für die *Serialize*- und *SerializeAsync*-Methode, die in der Tabelle 5 enthalten sind, ist auffällig, dass bei Nichtvorhandensein von Optionen,

⁹Dokumentation des MemoryBufferWriter

mit steigender Größe der Payload, *SerializeAsync* immer schneller ist, als sein synchroner Gegenpart. Als sich jedoch Optionen in der CoAP-Nachricht befinden, dann ändert sich das Ergebnis zugunsten der *Serialize*-Methode. Dies lässt sich damit erklären, dass durch das asynchrone Schreiben von Optionen, die nur einige Kilobytes groß sind, der Mehraufwand, zur Erzeugung der asynchronen Zustandsmaschine, überwiegt. Hier wäre eine Optimierung im Sinne eines gepufferten Schreibers möglich, da zuerst die Optionen serialisiert werden, also von der Objektstruktur in die zugrundeliegenden Bytes umgewandelt werden, in den Arbeitsspeicher und abschließend im einen Durchgang auf dem Stream asynchron geschrieben werden. Somit müsste nicht für jede Option einzeln eine asynchrone Zustandsmaschine erzeugt werden, sondern nur eine für den abschließenden Schreibvorgang vom Arbeitsspeicher auf dem Stream.

Die Daten, die in der Tabelle 5 aufgeführt sind, sind auch visuell in einem logarithmischen Diagramm dargestellt (Bilder 16 und 17).

$n_{Optionen}$	$l_{Payload}$ [B]	$t_{SerializeAsync}$ [μ s]	$t_{Serialize}$ [μ s]
0	0	5.618	5.142
0	1024	5.658	5.867
0	4096	5.710	7.186
0	65536	7.356	23.527
0	131072	9.347	49.680
1024	0	857.681	551.922
1024	1024	864.956	569.763
1024	4096	874.255	570.333
1024	65536	853.992	563.479
1024	131072	878.702	594.334
4096	0	3415.820	2758.439
4096	1024	3520.379	2659.178
4096	4096	3415.203	2596.436
4096	65536	3517.490	2507.161
4096	131072	3533.880	2559.126
65536	0	57919.596	58896.194
65536	1024	55995.639	58689.037
65536	4096	54769.243	58655.719
65536	65536	56494.925	59328.566
65536	131072	56286.964	57420.750
131072	0	112894.956	117807.577
131072	1024	110892.868	119120.364
131072	4096	109851.123	120061.250
131072	65536	111436.460	120002.202
131072	131072	112449.316	120453.645

Tabelle 5: Vergleich *SerializeAsync* und *Serialize*

Vergrößert man jedoch die Größe der Payload auf 1 GB, wie in Tabelle 10 ersichtlich, wird der Abstand zwischen *Serialize* und *SerializeAsync* sowohl im Hinblick auf die durchschnittliche Laufzeit als auch den Speicherverbrauch deutlich größer. Hierbei ist *SerializeAsync* um den Faktor 7 schneller im Durchschnitt und um den Faktor 143 effizienter im Speicherverbrauch.

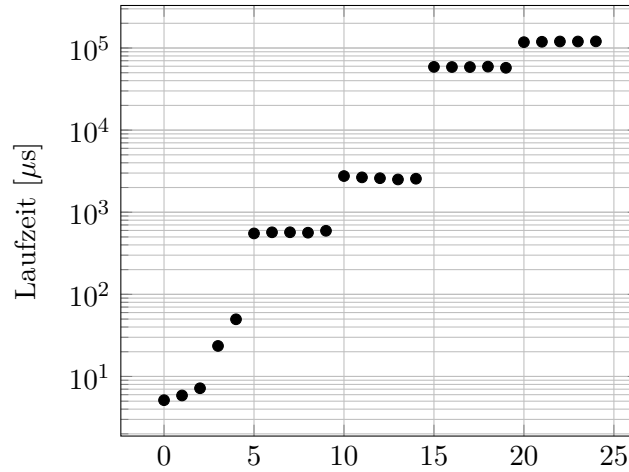


Abbildung 16: Messergebnisse von Serialize-Methode als Diagramm

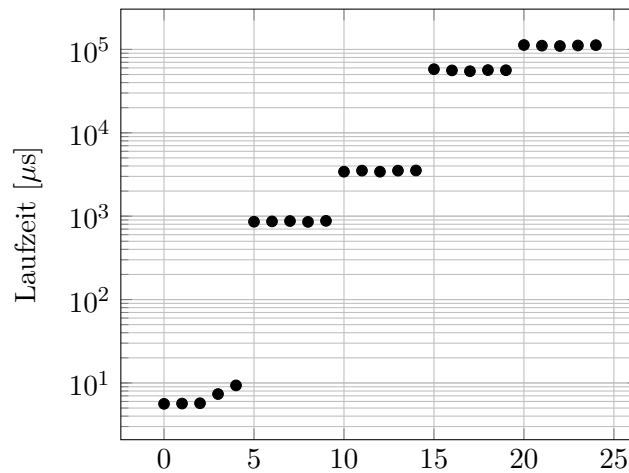


Abbildung 17: Messergebnisse von SerializeAsync-Methode als Diagramm

4 Diskussion der Messergebnisse

Betrachtet man die dargelegten Messergebnisse in Kapitel 3, sieht man, dass bei großen und aufwendigen I/O-Operationen Asynchronität besser abschneidet. Im Gegensatz dazu schlagen sich synchrone Methoden besser, wenn sich Menge der Daten im reservierten Speicher, also kein Nachladen oder Anforderung von weiteren Daten, sofort verarbeiten lässt. Dies lässt sich damit argumentieren, dass, wenn sich alle Daten im Speicher befinden, die synchrone Methode / Programm die Daten ohne weiteren Aufwand verwenden kann. Hingegen bei asynchronen Methoden ist es unvorteilhaft, wenn die Daten sehr klein sind und sich somit der Mehraufwand zum Aufbau der dafür benötigten Zustandsmaschine nicht lohnt.

Dies hat sich in beiden Messszenarien gezeigt, jedoch erst ab einer sehr großen Menge an Daten. Somit profitiert CoAP von Asynchronität, wenn eine große Menge an Daten verarbeitet werden muss. Ist jedoch die zu verarbeitende Menge an Daten klein, überwiegt der Mehraufwand der asynchronen Zustandsmaschine. Somit wirkt sich die Asynchronität bei kleineren Datenmengen negativ auf die Performanz der CoAP-Bibliothek aus.

Um diesen negativen Effekt entgegenzuwirken, gehen auch viele Entwickler von Softwareprogrammen, die gewisse asynchrone Methoden anbieten bzw. verwenden, dazu über, anhand von bestimmten Bedingungen oder Kriterien zu ermitteln, ob eine asynchrone oder eine synchrone Variante der zu implementierenden Funktion verwendet werden soll. Damit wird versucht eine asynchrone Methode dem Entwickler zur Verfügung zu stellen, die optimiert auf die jeweiligen Parameter ist und somit in jeglichen Fall die bestmögliche Leistung erbringt.

Jedoch muss der Entwickler abwägen, ob Nachrichten mit solch einer großen Payload innerhalb der Softwareapplikation zur Norm gehören. Anzumerken ist, dass sowohl die synchrone als auch asynchrone Implementierung Raum für Optimierungen offen lässt.

5 Schlussfolgerung

Hat nun die asynchrone Implementation eines Servers, in diesem Falle des *Serializers*, einen Einfluss auf dessen Durchsatzrate? Diese Frage kann nicht eindeutig beantwortet werden. Diese ist abhängig davon, wie man folgende Fragen beantwortet:

- Ist die Menge der zu verarbeitenden Daten klein? Wenn ja, dann hat Asynchronität keine große Auswirkung, sondern erzeugt eine erhebliche Mehrarbeit, aufgrund der zu erzeugenden Zustandsmaschine.
- Ist die Menge der zu verarbeitenden Daten groß? Wenn ja, dann hat Asynchronität eine große Auswirkung auf den Durchsatz, da durch die kürzere Verarbeitungszeit schneller neu ankommende Daten verarbeitet werden können.

Sieht man sich die Ergebnisse im Kapitel 3 an, bemerkt man, dass diese nicht eindeutig für Asynchronität oder Synchronität sprechen. Wenn keine CoAP-Options enthalten sind, dann verringert sich der Abstand von den asynchronen zu den synchronen Methoden mit größer werdender Payload.

Ab dem Zeitpunkt, in dem CoAP-Optionen in die Nachricht eingefügt werden, ist die synchrone Variante in den meisten Situationen schneller als die asynchrone. Dadurch, dass die Menge an Bytes, die für eine einzelne Option benötigt wird, sehr gering ist, hat *SerializeAsync* bzw. *DeserializeAsync* immer den Mehraufwand für den Aufbau seiner asynchronen Zustandsmaschine.

Erst mit einer sehr großen Payload von 1 GB zeichnet sich ein deutliches Ergebnis pro Asynchronität ab - zumindest bei der Verwendung der *SerializeAsync*-Methode. Hier ist *SerializeAsync*, wie schon festgestellt worden ist, bei der durchschnittlichen Laufzeit um den Faktor 7 schneller und um den Faktor 143 effizienter im Speicherverbrauch als die *Serialize*-Methode.

Was bedeutet dies nun für die, in dieser Arbeit gestellten Ausgangsfrage? Eine eingebaute Asynchronität in Softwareprogrammen ist nicht für jedes Problem die passende Lösung. Es kann seine Vorteile gut in Situationen ausspielen, in denen auf ein Ergebnis einer zeit- oder rechenintensiven Aktion gewartet werden muss. Diese können in Form von Netzwerkübertragungen (Anfragen bzw. Antworten mittels HTTP, Datenbankabfragen), rechenintensiven Berechnungen oder speicherintensiven I/O-Vorgängen (Lesen einer großen Datei von der Festplatte) auftreten. Für einfache und schnell auszuführende Aufgaben ist eine asynchrone Methode eher die falsche Wahl, da wie schon erwähnt, der Aufwand der asynchronen Zustandsmaschine zu groß wird.

Darum ist es abhängig von den Anforderungen und Aufgabenstellung der zu entwickelnden Software. Interagiert der betreffende Codeteil mit externen Systemen (Datenbanksystemen, REST-APIs, Internetservern) oder führt dieser rechenintensive Berechnungen durch, dann ist das Verwenden von Asynchronität zu empfehlen, wie es schon Okur u. a. in deren Studie „A study and toolkit for asynchronous programming in C#“ [6]

klargestellt haben. Damit wird der Main-Thread entlastet, was dazu führt, dass bei Applikationen ohne grafisches Interface (GUI) die Auslastung des Threadpools reduziert wird und bei Applikationen mit GUI der Main-Thread nicht blockiert und somit die Oberfläche benutzbar bleibt.

Betrachtet man nun den Fall eines Servers, dann kann man dazu tendieren auf eine asynchrone Implementierung zu setzen, da hier Daten über ein beliebiges Medium, wie zum Beispiel Ethernet oder WLAN, und beliebiges Protokoll an den Server geschickt werden. Dabei weiß der Server jedoch nicht im Vorhinein, wann diese Übertragung komplett abgeschlossen ist. Um nicht den Main-Thread des Servers zu blockieren, wie es bei einem synchronen Server passieren würde, kann durch den Einsatz von Asynchronität der Main-Thread entlastet und somit der Durchsatz gesteigert werden. Vor allem im speziellen Fall, wenn eine große Menge an Daten (siehe Tabelle 10) verarbeitet werden muss, kann die Asynchronität eines Programms deutliche Performanz- und Speichervorteile bringen. Nur durch die Verwendung von Asynchronität können sieben CoAP-Nachrichten, mit einer großen Anzahl an Options (100000) und einer großen Payload (1 GB), in der gleichen Zeit verarbeitet werden, wie als wenn man die synchrone *Serialize*-Methode verwendet.

Ist dies jedoch ein triftiger Grund, um nur asynchrone Programme zu schreiben? Nicht unbedingt. Es kommt auf die gestellten Anforderungen an die Software an, ob diese sehr performant und ressourcenschonend arbeiten sollte. Ist dies der Fall, dann sollte ein großes Augenmerk darauf gelegt werden, zu überlegen, an welchen Stellen im Programm eine asynchrone gegenüber einer synchronen Methode vorgezogen werden sollte. Hierbei bieten viele Bibliotheken, zumindest im C#-Ökosystem, sowohl synchrone als auch asynchrone Methoden an, die genau dieselbe Funktion abbilden, jedoch abhängig davon, ob es synchron oder asynchron passieren soll. Ist man jedoch selbst ein Entwickler einer solchen Bibliothek, ist es schwierig abzuwägen, welche Methoden einer asynchronen oder einer synchronen Implementierung benötigen. Betrachtet man die Referenzquelle¹⁰ von Microsoft für .NET an, sieht man, dass die Entwickler dazu tendieren, anhand von verschiedenen Parametern zu entscheiden, ob die darunterliegende Funktion synchron oder asynchron ausgeführt werden sollte. Ein fiktives Beispiel dafür wird im Listing 12 aufgeführt.

Für das CoAP-Protokoll, spezifiziert nach RFC 7252 von Shelby, Hartke und Bormann [8], erzielt Asynchronität keine bemerkenswerten Verbesserungen, da durch die Größenlimitierung von UDP-Paketen, die mit einer maximalen Länge von 65 535 Bytes festgelegt ist, die Datenmenge zu klein ist. Implementiert man die Funktion des blockweisen Datentransfers, wie im RFC 7959 von Shelby und Bormann [7] beschrieben, könnte diese Limitierung umgangen werden und größere Datengrößen erzielt werden. Man kann sich auch dazu entscheiden auf die neuere Standardisierung, wie im RFC 8323 von Bormann u. a. beschrieben [1], zu setzen, um größere Mengen an Daten, ohne großen Aufwand, zu übertragen.

¹⁰Referenzquelle von Microsoft zu .NET

Das Fazit aus dieser Arbeit ist damit, dass der Vorteil von asynchronen Server bzw. Programmen sehr davon abhängt, in welchen Szenarien diese eingesetzt werden. Müssen viele Anfragen mit großer Datenmenge gleichzeitig verarbeitet werden, dann ist ein asynchroner Server einem synchronen vorzuziehen. Dies wird dadurch noch verstärkt, dass eine asynchrone Implementierung eines Servers, zumindest in C#, durch eine einfache und durchdachte Syntax bestehend aus `async`, `await` und `Task` bzw. `Task<TResult>` ermöglicht wird. Auch ist der Trend bemerkbar, dass viele bekannte Bibliotheken eine große Anzahl an asynchronen Methoden anbieten, damit diese auch in asynchronen Konstrukten verwendet werden können.

Jedoch wird hier auch eine generelle Empfehlung zur Implementierung und Verwendung von asynchronen Methoden ausgesprochen. Die gemessenen, zeitlichen Unterschiede zwischen Asynchronität und Synchronität befinden sich im vertretbaren Rahmen. Ist man nicht auf vollkommene Optimierung der Performanz der Software fokussiert, ist dieser Unterschied vernachlässigbar.

Literatur

- [1] C. Bormann u. a. *CoAP (Constrained Application Protocol) over TCP, TLS, and WebSockets*. RFC 8323. RFC Editor, Feb. 2018. URL: <http://www.rfc-editor.org/rfc/rfc8323.txt>.
- [2] Stephen Cleary. *Concurrency in C# Cookbook: Asynchronous, Parallel, and Multithreaded Programming*. O'Reilly Media, 2019.
- [3] Ravi Kishore Kodali u. a. „Implementation of Home Automation Using CoAP“. In: *TENCON 2018 - 2018 IEEE Region 10 Conference*. 2018, S. 1214–1218. DOI: 10.1109/TENCON.2018.8650135.
- [4] Koojana Kuladinithi u. a. „Implementation of coap and its application in transport logistics“. In: *Proc. IP+ SN, Chicago, IL, USA* (2011).
- [5] Alessandro Ludovici, Pol Moreno und Anna Calveras. „TinyCoAP: A Novel Constrained Application Protocol (CoAP) Implementation for Embedding RESTful Web Services in Wireless Sensor Networks Based on TinyOS“. In: *Journal of Sensor and Actuator Networks* 2.2 (2013), S. 288–315. ISSN: 2224-2708. DOI: 10.3390/jsan2020288. URL: <https://www.mdpi.com/2224-2708/2/2/288>.
- [6] Semih Okur u. a. „A study and toolkit for asynchronous programming in C#“. In: *Proceedings of the 36th International Conference on Software Engineering*. 2014, S. 1117–1127.
- [7] Z. Shelby und C. Bormann. *Block-Wise Transfers in the Constrained Application Protocol (CoAP)*. RFC 7959. RFC Editor, Aug. 2016. URL: <http://www.rfc-editor.org/rfc/rfc7959.txt>.
- [8] Z. Shelby, K. Hartke und C. Bormann. *The Constrained Application Protocol (CoAP)*. RFC 7252. RFC Editor, Juni 2014. URL: <http://www.rfc-editor.org/rfc/rfc7252.txt>.
- [9] smeshlink. *CoAP.NET*. Version 1.1.0. 2016. URL: <https://github.com/smeshlink/CoAP.NET>.
- [10] Wikipedia contributors. *Constrained Application Protocol — Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/w/index.php?title=Constrained_Application_Protocol&oldid=1044860831. [Online; accessed 4-January-2022]. 2021.

Anhang

1 Tabellen

1.1 Ergebnisse (Nachrichtenübertragung)

Start	Ende	Laufzeit
2021-11-11T11:53:36.3360380+01:00	2021-11-11T11:59:17.2094861+01:00	00:05:40.8734481
2021-11-11T11:25:13.6291350+01:00	2021-11-11T11:30:54.8399509+01:00	00:05:41.2108159
2021-11-11T11:29:42.9196146+01:00	2021-11-11T11:35:24.2404882+01:00	00:05:41.3208736
2021-11-11T11:32:06.6750943+01:00	2021-11-11T11:37:47.7269970+01:00	00:05:41.0519027
2021-11-11T11:33:33.8796230+01:00	2021-11-11T11:39:14.4538810+01:00	00:05:40.5742580

Tabelle 6: Asynchrone Übertragung mit 100 Options und mit einer Payload von 1000 Bytes

Start	Ende	Laufzeit
2021-11-11T11:53:36.3360380+01:00	2021-11-11T11:59:17.2084637+01:00	00:05:40.8724257
2021-11-11T11:25:13.6291344+01:00	2021-11-11T11:30:54.8195384+01:00	00:05:41.1904040
2021-11-11T11:29:42.9196145+01:00	2021-11-11T11:35:24.2416850+01:00	00:05:41.3220705
2021-11-11T11:32:06.6750956+01:00	2021-11-11T11:37:47.7273639+01:00	00:05:41.0522683
2021-11-11T11:33:33.8796186+01:00	2021-11-11T11:39:14.4542465+01:00	00:05:40.5746279

Tabelle 7: Synchrone Übertragung mit 100 Options und mit einer Payload von 1000 Bytes

Start	Ende	Laufzeit
2021-11-11T11:44:51.4637384+01:00	2021-11-11T12:29:39.2904876+01:00	00:44:47.8267492
2021-11-11T11:45:48.8515388+01:00	2021-11-11T12:30:46.9714652+01:00	00:44:58.1199264
2021-11-11T11:46:46.0080509+01:00	2021-11-11T12:31:33.3480246+01:00	00:44:47.3399737
2021-11-11T11:47:52.2395272+01:00	2021-11-11T12:32:38.3763629+01:00	00:44:46.1368357
2021-11-11T11:49:06.5417421+01:00	2021-11-11T12:33:51.5103142+01:00	00:44:44.9685721

Tabelle 8: Asynchrone Übertragung mit 100 Options und einer Payload von 10000 Bytes

Start	Ende	Laufzeit
2021-11-11T11:44:51.4637476+01:00	2021-11-11T12:29:39.6267910+01:00	00:44:48.1630434
2021-11-11T11:45:48.8515384+01:00	2021-11-11T12:30:46.9709319+01:00	00:44:58.1193935
2021-11-11T11:46:46.0080501+01:00	2021-11-11T12:31:33.3486035+01:00	00:44:47.3405534
2021-11-11T11:47:52.2395282+01:00	2021-11-11T12:32:38.3767620+01:00	00:44:46.1372338
2021-11-11T11:49:06.5417372+01:00	2021-11-11T12:33:51.5109177+01:00	00:44:44.9691805

Tabelle 9: Synchrone Übertragung mit 100 Options und mit einer Payload von 10000 Bytes

1.2 Ergebnisse (Deserialisierung und Serialisierung)

Method	AmountOfOptions	LengthOfPayload	Mean	Error	StdDev	Gen 0	Gen 1	Allocated
SerializeAsync	100000	1000000000	158.1 ms	3.13 ms	5.06 ms	5000.0000	-	21 MB
Serialize	100000	1000000000	1,112.3 ms	21.24 ms	25.29 ms	3000.0000	-	3,016 MB
DeserializeAsync	100000	1000000000	1,507.7 ms	28.40 ms	29.16 ms	7000.0000	2000.0000	4,003 MB
Deserialize	100000	1000000000	1,764.1 ms	34.93 ms	52.28 ms	4000.0000	1000.0000	2,893 MB

Tabelle 10: Benchmark mit 100000 Options und mit einer Payload von 1000000000 Bytes (1 GB)

AmountOfOptions	LengthOfPayload	Mean	Error	StdDev	Gen 0	Gen 1	Gen 2	Allocated
0	0	1.679 μ s	0.0310 μ s	0.0275 μ s	0.3948	-	-	2 KB
0	1024	3.683 μ s	0.0712 μ s	0.0666 μ s	1.6327	-	-	6 KB
0	4096	6.570 μ s	0.0678 μ s	0.0530 μ s	2.9373	-	-	10 KB
0	65536	98.533 μ s	1.8525 μ s	1.9822 μ s	83.2520	83.2520	83.2520	258 KB
0	131072	175.433 μ s	1.3566 μ s	1.2026 μ s	166.5039	166.5039	166.5039	514 KB
1024	0	1,030.051 μ s	9.6583 μ s	9.0344 μ s	196.2891	97.6563	-	1,075 KB
1024	1024	1,014.093 μ s	6.1471 μ s	5.1331 μ s	191.4063	95.7031	-	1,079 KB
1024	4096	1,027.275 μ s	9.5509 μ s	8.9339 μ s	192.3828	95.7031	-	1,091 KB
1024	65536	1,631.388 μ s	38.9956 μ s	114.3674 μ s	273.4375	191.4063	76.1719	1,331 KB
1024	131072	1,781.418 μ s	35.2748 μ s	71.2567 μ s	332.0313	277.3438	144.5313	1,587 KB
4096	0	6,477.162 μ s	129.3902 μ s	163.6372 μ s	726.5625	328.1250	101.5625	4,291 KB
4096	1024	6,711.475 μ s	133.2189 μ s	259.8331 μ s	734.3750	335.9375	93.7500	4,295 KB
4096	4096	6,558.996 μ s	130.5046 μ s	339.1988 μ s	742.1875	343.7500	109.3750	4,307 KB
4096	65536	6,924.552 μ s	137.4221 μ s	261.4598 μ s	804.6875	421.8750	85.9375	4,547 KB
4096	131072	7,530.808 μ s	147.7269 μ s	175.8583 μ s	843.7500	460.9375	148.4375	4,803 KB
65536	0	166,612.138 μ s	3,300.2174 μ s	3,241.2547 μ s	11666.6667	4333.3333	666.6667	68,613 KB
65536	1024	190,898.380 μ s	3,756.8801 μ s	4,885.0053 μ s	11750.0000	4500.0000	1000.0000	68,617 KB
65536	4096	190,316.150 μ s	3,778.9048 μ s	6,716.9994 μ s	12000.0000	4750.0000	1000.0000	68,629 KB
65536	65536	195,944.825 μ s	3,911.2806 μ s	8,585.3557 μ s	11750.0000	4500.0000	1000.0000	68,869 KB
65536	131072	170,582.006 μ s	3,144.9051 μ s	6,494.7744 μ s	11666.6667	4333.3333	666.6667	69,125 KB
131072	0	286,105.657 μ s	4,145.7348 μ s	3,675.0859 μ s	22000.0000	7000.0000	-	137,225 KB
131072	1024	285,783.220 μ s	4,849.2919 μ s	4,536.0307 μ s	22000.0000	7000.0000	-	137,227 KB
131072	4096	282,355.564 μ s	4,535.0488 μ s	4,020.2027 μ s	22000.0000	7000.0000	-	137,239 KB
131072	65536	292,059.675 μ s	5,689.0265 μ s	5,587.3847 μ s	22000.0000	7000.0000	-	137,481 KB
131072	131072	304,488.343 μ s	5,960.5795 μ s	11,484.0131 μ s	22000.0000	7000.0000	-	137,735 KB

Tabelle 11: Benchmark Deserialize

AmountOfOptions	LengthOfPayload	Mean	Error	StdDev	Median	Gen 0	Gen 1	Gen 2	Allocated
0	0	2.231 μ s	0.0353 μ s	0.0483 μ s	2.214 μ s	0.3586	-	-	1 KB
0	1024	3.298 μ s	0.0602 μ s	0.0563 μ s	3.286 μ s	0.9232	-	-	4 KB
0	4096	5.077 μ s	0.1011 μ s	0.1038 μ s	5.109 μ s	2.9411	-	-	10 KB
0	65536	64.490 μ s	1.2016 μ s	1.1802 μ s	64.108 μ s	83.3130	83.3130	83.3130	258 KB
0	131072	111.011 μ s	2.0769 μ s	2.0398 μ s	110.221 μ s	166.6260	166.6260	166.6260	514 KB
1024	0	1,440.076 μ s	27.5160 μ s	30.5839 μ s	1,445.906 μ s	216.7969	107.4219	-	1,250 KB
1024	1024	1,402.291 μ s	26.1549 μ s	23.1857 μ s	1,403.606 μ s	208.9844	103.5156	-	1,254 KB
1024	4096	1,416.083 μ s	27.4978 μ s	28.2382 μ s	1,418.433 μ s	248.0469	123.0469	-	1,267 KB
1024	65536	1,592.417 μ s	31.4054 μ s	75.2453 μ s	1,574.960 μ s	285.1563	201.1719	76.1719	1,507 KB
1024	131072	1,817.051 μ s	47.0170 μ s	135.6547 μ s	1,783.732 μ s	359.3750	283.2031	146.4844	1,763 KB
4096	0	8,416.154 μ s	168.1480 μ s	396.3445 μ s	8,340.885 μ s	835.9375	367.1875	156.2500	4,995 KB
4096	1024	8,147.070 μ s	101.2302 μ s	94.6908 μ s	8,154.584 μ s	835.9375	367.1875	164.0625	4,999 KB
4096	4096	8,737.255 μ s	164.8246 μ s	236.3866 μ s	8,709.312 μ s	859.3750	359.3750	148.4375	5,011 KB
4096	65536	9,057.536 μ s	166.6596 μ s	321.0964 μ s	8,977.750 μ s	921.8750	453.1250	125.0000	5,251 KB
4096	131072	9,543.723 μ s	190.4048 μ s	463.4717 μ s	9,467.759 μ s	968.7500	500.0000	140.6250	5,507 KB
65536	0	203,180.664 μ s	3,832.5079 μ s	6,812.2788 μ s	202,695.167 μ s	13666.6667	5000.0000	666.6667	79,877 KB
65536	1024	198,211.164 μ s	3,846.5813 μ s	4,864.6937 μ s	197,809.500 μ s	13333.3333	5000.0000	666.6667	79,881 KB
65536	4096	201,316.657 μ s	3,771.5768 μ s	6,301.4575 μ s	200,897.850 μ s	13333.3333	4666.6667	666.6667	79,893 KB
65536	65536	200,257.897 μ s	3,846.1005 μ s	4,723.3569 μ s	200,426.333 μ s	13666.6667	5333.3333	666.6667	80,134 KB
65536	131072	201,631.362 μ s	3,958.6431 μ s	6,392.4708 μ s	201,162.367 μ s	13333.3333	5000.0000	666.6667	80,390 KB
131072	0	372,129.381 μ s	7,421.6676 μ s	19,289.9009 μ s	373,452.900 μ s	25000.0000	8000.0000	-	159,752 KB
131072	1024	372,065.709 μ s	7,358.5294 μ s	16,152.1504 μ s	374,064.100 μ s	25000.0000	8000.0000	-	159,756 KB
131072	4096	368,936.160 μ s	7,344.8485 μ s	16,727.9368 μ s	369,767.550 μ s	25000.0000	8000.0000	-	159,768 KB
131072	65536	366,150.575 μ s	7,271.6429 μ s	17,837.4640 μ s	360,195.900 μ s	27000.0000	9000.0000	1000.0000	160,008 KB
131072	131072	370,709.611 μ s	7,004.9653 μ s	17,574.0884 μ s	370,470.050 μ s	25000.0000	8000.0000	-	160,264 KB

Tabelle 12: Benchmark DeserializeAsync

AmountOfOptions	LengthOfPayload	Mean	Error	StdDev	Gen 0	Gen 1	Gen 2	Allocated
0	0	5.142 μ s	0.0683 μ s	0.0639 μ s	0.0687	-	-	288 B
0	1024	5.867 μ s	0.0751 μ s	0.0702 μ s	0.3281	-	-	1,384 B
0	4096	7.186 μ s	0.0671 μ s	0.0628 μ s	1.0757	-	-	4,504 B
0	65536	23.527 μ s	0.2614 μ s	0.2183 μ s	15.5945	3.1128	-	66,040 B
0	131072	49.680 μ s	0.5993 μ s	0.5606 μ s	41.6260	41.6260	41.6260	131,600 B
1024	0	551.922 μ s	9.4106 μ s	10.0692 μ s	175.7813	76.1719	76.1719	681,589 B
1024	1024	569.763 μ s	11.1684 μ s	17.7143 μ s	175.7813	76.1719	76.1719	682,612 B
1024	4096	570.333 μ s	11.2692 μ s	15.0440 μ s	175.7813	76.1719	76.1719	685,693 B
1024	65536	563.479 μ s	5.4708 μ s	4.8498 μ s	195.3125	95.7031	95.7031	747,134 B
1024	131072	594.334 μ s	6.3800 μ s	5.6557 μ s	209.9609	110.3516	110.3516	812,665 B
4096	0	2,758.439 μ s	55.0103 μ s	105.9861 μ s	808.5938	410.1563	410.1563	4,821,823 B
4096	1024	2,659.178 μ s	40.6188 μ s	37.9949 μ s	800.7813	402.3438	402.3438	4,822,917 B
4096	4096	2,596.436 μ s	46.9202 μ s	64.2249 μ s	875.0000	476.5625	476.5625	4,825,775 B
4096	65536	2,507.161 μ s	49.8926 μ s	101.9173 μ s	894.5313	496.0938	496.0938	4,887,220 B
4096	131072	2,559.126 μ s	19.1575 μ s	15.9973 μ s	890.6250	492.1875	492.1875	4,952,758 B
65536	0	58,896.194 μ s	1,159.8829 μ s	1,872.9947 μ s	7333.3333	1000.0000	1000.0000	108,594,097 B
65536	1024	58,689.037 μ s	1,169.1967 μ s	1,639.0505 μ s	7200.0000	900.0000	900.0000	108,595,122 B
65536	4096	58,655.719 μ s	1,165.0065 μ s	1,707.6514 μ s	7200.0000	900.0000	900.0000	108,598,179 B
65536	65536	59,328.566 μ s	1,168.5292 μ s	1,638.1147 μ s	7333.3333	1000.0000	1000.0000	108,659,633 B
65536	131072	57,420.750 μ s	1,129.7427 μ s	2,008.1165 μ s	7100.0000	800.0000	800.0000	108,725,176 B
131072	0	117,807.577 μ s	2,353.7162 μ s	6,482.8180 μ s	13800.0000	1200.0000	1200.0000	219,284,480 B
131072	1024	119,120.364 μ s	2,359.3323 μ s	6,132.2185 μ s	13200.0000	600.0000	600.0000	219,285,437 B
131072	4096	120,061.250 μ s	2,387.2366 μ s	5,989.1099 μ s	13800.0000	1200.0000	1200.0000	219,289,669 B
131072	65536	120,002.202 μ s	2,372.2717 μ s	5,207.1938 μ s	13800.0000	1200.0000	1200.0000	219,349,986 B
131072	131072	120,453.645 μ s	2,374.0783 μ s	5,595.9819 μ s	13800.0000	1200.0000	1200.0000	219,415,507 B

Tabelle 13: Benchmark Serialize

AmountOfOptions	LengthOfPayload	Mean	Error	StdDev	Gen 0	Allocated
0	0	5.618 μ s	0.0390 μ s	0.0345 μ s	0.0839	376 B
0	1024	5.658 μ s	0.0258 μ s	0.0215 μ s	0.0839	376 B
0	4096	5.710 μ s	0.0276 μ s	0.0231 μ s	0.0839	376 B
0	65536	7.356 μ s	0.0524 μ s	0.0464 μ s	0.0839	376 B
0	131072	9.347 μ s	0.0350 μ s	0.0327 μ s	0.0763	376 B
1024	0	857.681 μ s	5.6334 μ s	5.2695 μ s	113.2813	475,569 B
1024	1024	864.956 μ s	8.3344 μ s	7.7960 μ s	113.2813	475,570 B
1024	4096	874.255 μ s	11.4655 μ s	9.5742 μ s	113.2813	475,571 B
1024	65536	853.992 μ s	7.4185 μ s	6.9393 μ s	113.2813	475,569 B
1024	131072	878.702 μ s	7.5007 μ s	7.0161 μ s	113.2813	475,569 B
4096	0	3,415.820 μ s	20.0413 μ s	18.7467 μ s	453.1250	1,900,983 B
4096	1024	3,520.379 μ s	34.5694 μ s	32.3363 μ s	453.1250	1,900,976 B
4096	4096	3,415.203 μ s	28.1909 μ s	26.3698 μ s	453.1250	1,900,976 B
4096	65536	3,517.490 μ s	29.9689 μ s	26.5667 μ s	453.1250	1,900,981 B
4096	131072	3,533.880 μ s	41.2242 μ s	36.5442 μ s	453.1250	1,900,976 B
65536	0	57,919.596 μ s	667.7145 μ s	624.5805 μ s	7200.0000	30,409,136 B
65536	1024	55,995.639 μ s	537.3204 μ s	476.3206 μ s	7200.0000	30,409,136 B
65536	4096	54,769.243 μ s	426.8328 μ s	378.3761 μ s	7200.0000	30,409,136 B
65536	65536	56,494.925 μ s	574.7165 μ s	537.5901 μ s	7200.0000	30,409,136 B
65536	131072	56,286.964 μ s	373.1133 μ s	349.0104 μ s	7200.0000	30,409,136 B
131072	0	112,894.956 μ s	947.4691 μ s	839.9067 μ s	14400.0000	60,817,941 B
131072	1024	110,892.868 μ s	1,111.1807 μ s	1,039.3991 μ s	14400.0000	60,818,109 B
131072	4096	109,851.123 μ s	1,410.9865 μ s	1,178.2376 μ s	14400.0000	60,820,210 B
131072	65536	111,436.460 μ s	688.9756 μ s	610.7590 μ s	14400.0000	60,818,109 B
131072	131072	112,449.316 μ s	685.4489 μ s	607.6326 μ s	14400.0000	60,818,109 B

Tabelle 14: Benchmark SerializeAsync

2 Codeausschnitte

```

1  public class UriHost : StringOption
2  {
3      public const ushort NUMBER = OptionNumberRegistry.StringOptions.UriHost;
4      private const ushort MAX_LENGTH = 255;
5      private const ushort MIN_LENGTH = 1;
6
7      public UriHost(string value)
8          : base(NUMBER, value, MIN_LENGTH, MAX_LENGTH, false)
9      {
10     }
11
12     public UriHost(ReadOnlySpan<byte> value)
13         : base(NUMBER, value, MIN_LENGTH, MAX_LENGTH, false)
14     {
15     }
16
17     public class Factory : IOptionFactory
18     {
19         public int Number => NUMBER;
20
21         public CoapOption Create(OptionData src)
22         {
23             if (src.Number != this.Number)
24             {
25                 var msg = "Option number is not valid for Uri-Host factory.";
26                 throw new ArgumentException(msg);
27             }
28
29             return new UriHost(src.Value);
30         }
31     }
32 }

```

Listing 4: UriHost Implementation

```

1  public int Read(ReadOnlySpan<byte> value, out ReadOnlyOptionCollection result)
2  {
3      var options = new OptionCollection();
4      var offset = 0;
5      var previousNumber = 0;
6      while (OptionsReader.HasNext(value[offset..]))
7      {
8          var optionData = this.ReadOptionData(
9              previousNumber,
10             value[offset..],
11             out var bytesConsumed);
12
13             previousNumber = optionData.Number;
14             offset += bytesConsumed;
15
16             if (!this.factories.TryGetValue(optionData.Number, out var factory))
17             {
18                 factory = UnrecognizedOptionFactory;
19             }
20
21             var option = factory.Create(optionData);
22             options.Add(option);
23         }
24
25         result = new ReadOnlyOptionCollection(options);
26         return offset;
27     }

```

Listing 5: Synchrones Lesen von Optionen

```

1  public async IEnumerable<CoapOption> ReadAsync(
2      Stream stream,
3      [EnumeratorCancellation] CancellationToken ct)
4  {
5      var previousNumber = 0;
6      while (true)
7      {
8          if (!this.TryReadOptionHeader(stream, out var optionHeader))
9          {
10             break;
11          }
12
13          if (optionHeader.IsPayloadMarker)
14          {
15             break;
16          }
17
18          var optionData = await this.ReadOptionAsync(
19              stream,
20              optionHeader,
21              previousNumber,
22              ct)
23              .ConfigureAwait(false);
24
25          previousNumber = optionData.Number;
26          if (!this.factories.TryGetValue(optionData.Number, out var factory))
27          {
28             factory = UnrecognizedOptionFactory;
29          }
30
31          var option = factory.Create(optionData);
32          yield return option;
33      }
34  }

```

Listing 6: Asynchrones Lesen von Optionen

```

1  public void Write(PooledMemoryBufferWriter writer, ReadOnlyOptionCollection value)
2  {
3      CoapOption previousOption = null;
4      foreach (var option in value)
5      {
6          var serializer = this.serializers.SingleOrDefault(s => s.CanSerialize(option));
7          if (serializer == null)
8          {
9              continue;
10         }
11
12         serializer.Serialize(writer, previousOption, option);
13         previousOption = option;
14     }
15 }

```

Listing 7: Synchrones Schreiben von Optionen

```

1  public async Task WriteAsync(
2      Stream stream,
3      ReadOnlyOptionCollection options,
4      CancellationToken ct)
5  {
6      CoapOption previousOption = null;
7      foreach (var option in options)
8      {
9          var serializer = this.serializers
10             .SingleOrDefault(s => s.CanSerialize(option));
11             if (serializer == null)
12             {
13                 continue;
14             }
15
16             await serializer
17                 .SerializeAsync(stream, previousOption, option, ct)
18                 .ConfigureAwait(false);
19             previousOption = option;
20     }
21 }

```

Listing 8: Asynchrones Schreiben von Optionen


```

1 public class BenchmarkExample
2 {
3     // Initialisierung eines Zufallsgenerators.
4     private static readonly Random Random = new Random();
5     private readonly byte[] data;
6     private readonly SHA256 sha256 = SHA256.Create();
7     private readonly MD5 md5 = MD5.Create();
8
9     // Initialisierung der Daten für den Benchmark
10    public BenchmarkExample()
11    {
12        this.data = new byte[10000];
13        this.Random.NextBytes(this.data);
14    }
15
16    // Deklaration eines Benchmarks für SHA256 Berechnung.
17    [Benchmark]
18    public byte[] SHA256 => this.sha256.ComputeHash(this.data);
19
20    // Deklaration eines Benchmarks für MD5 Berechnung.
21    [Benchmark]
22    public byte[] MD5 => this.md5.ComputeHash(this.data);
23 }

```

Listing 9: Beispiel eines Benchmarks in BenchmarkDotnet

```

1 public class Program
2 {
3     public static void Main(string[] args)
4     {
5         var summary = BenchmarkRunner.Run<BenchmarkExample>();
6     }
7 }

```

Listing 10: Ausführen der Benchmark-Klasse

```

1  // Erzeugung des PooledMemoryBufferWriters.
2  using (var writer = new PooledMemoryBufferWriter())
3  {
4      // Anfordern eines Memory<byte> in der Größe von 2048 Bytes.
5      var buffer = writer.GetMemory(2048);
6
7      // Erzeugung eines Zufallgenerators.
8      var random = new Random();
9
10     // Befüllung des Puffers mit zufälligen Werten.
11     random.NextBytes(buffer.Span);
12
13     // Benachrichtigung des Writers, dass der gesamte Puffer beschrieben wurde.
14     writer.Advance(buffer.Length);
15 }

```

Listing 11: Verwendung des PooledMemoryBufferWriters

```

1  public async Task<byte[]> ReadAsync(Stream stream, CancellationToken ct)
2  {
3      // Ausleihen eines Speichers.
4      var owner = MemoryPool<byte>.Shared.Rent(2048);
5      var buffer = owner.Memory;
6
7      // Überprüfen, ob der gesamte Stream in den Puffer geladen werden kann.
8      if (stream.Length <= 2048)
9      {
10         // Lesen des Speichers in synchroner Variante.
11         var bytesRead = stream.Read(buffer.Span);
12         return buffer.ToArray();
13     }
14
15     // Sonst asynchrones Lesen des gesamten Streams.
16     // Das passiert folgendermaßen:
17     // 1) Lesen des Streams mittels ReadAsync.
18     // 2) Überprüfen ob Ende des Streams erreicht wurde.
19     // 2.1) Wenn ja --> den Inhalt des Puffers zurückgeben.
20     // 2.2) Wenn nein --> Puffer verdoppeln und mit Punkt 1 beginnen.
21 }

```

Listing 12: Optimierte asynchrone Methode