Bachelorarbeit im Studiengang B.Sc. Wirtschaftsinformatik

Realisierung einer bidirektionalen Chat-Anwendung durch HTML5 WebSockets mit Java und Angular 4 zum Einsatz im seminaristischen Kontext

Andre Weinkötz

15. Februar 2018

Fakultät für Informatik und Mathematik
Hochschule München

Betreuer: Prof. Dr. Mandl

Abstract

Hier steht am Ende was genau eigentlich mit dem Unsinn hier erreicht werden soll.

Inhaltsverzeichnis

1 Einleitung

Die Anforderungen an Webanwendungen haben sich in den vergangenen Jahren stark verändert. Mobile Geräte wie Smartphones oder Tablets ersetzen stationäre Systeme, Webanwendungen sollen zur Kollaboration eingesetzt werden und Buchungssysteme oder Finanzanwendungen verlangen die Bearbeitung tausender Anfragen mit minimaler Verzögerung. Das Hypertext Transfer Protocol (HTTP) bietet hier keine zufriedenstellende Lösung. Um den Anforderungen an moderne Webanwendungen gerecht zu werden, spezifizierten das World Wide Web Consortium (W3C) und die Internet Engineering Taskforce (IETF) das WebSocket-Protokoll mit zugehöriger JavaScript-API.

Die Entwickler des WebSocket Protokolls machten es sich zur Aufgabe, einen Mechanismus zu schaffen, der es browserbasierten Anwendungen ermöglicht bidirektional zu kommunizieren ohne dabei auf mehrere HTTP Verbindungen zu öffnen [FM11]. Dabei sollte auf zusätzliche Methoden - wie beispielsweise den Einsatz von XMLHttpRequests, iframes oder long polling - verzichtet werden.

2 Das WebSocket Protokoll

Die Arbeit an der Spezifikation des WebSocket Protokolls begann bereits 2009, als es von Google in Zusammenarbeit mit Apple, Microsoft und Mozilla im Rahmen ihrer Kooperation WHATWG[1] der Internet Engineering Task Force (IETF) vorgeschlagen wurde. Bis Mai 2010 wurden noch zahlreiche Verbesserungen hinzugefügt, bis es schließlich in Version 76 [Hic10] im August 2010 an die BiDirectional or Server-Initiated HTTP (HyBi) Task Force der IETF zur Weiterentwicklung übergeben wurde [Fet10]. Neben vielen anderen Fortschritten wurde das Protkoll um die Möglichkeit erweitert binäre Dateien auszutauschen, sowie Sicherheitslücken geschlossen, die in Verbindung mit Proxy-Servern entstehen konnten. Im Dezember 2010 wurde das WebSocket Protokoll von der IETF zu dem Request for Comment (RFC) 6455 erklärt [FM11]. Dieser definiert auf über 70 Seiten neben dem Lebenszyklus und den verschiedenen Frames der WebSockets noch zahlreiche weitere technische Details und Defintionen. In diesem Kapitel sollen die wichtigsten Grundlagen der Protokollspezifikation dargelegt werden.

2.1 WebSocket-Frames

Die Spezifikation eines WebSocket-Frames sieht eine Zweiteilung vor, wie sie bei Protokollnachrichten üblicherweise vorgenommen wird. Ein WebSocket-Frame besteht demnach aus einem Header Bereich der Kontrolldaten enthält, sowie dem Payload, welcher die Nutzdaten beinhaltet. Anhand des Aufbaus ist leicht erkennbar, dass die Kommunikation über WebSockets mit

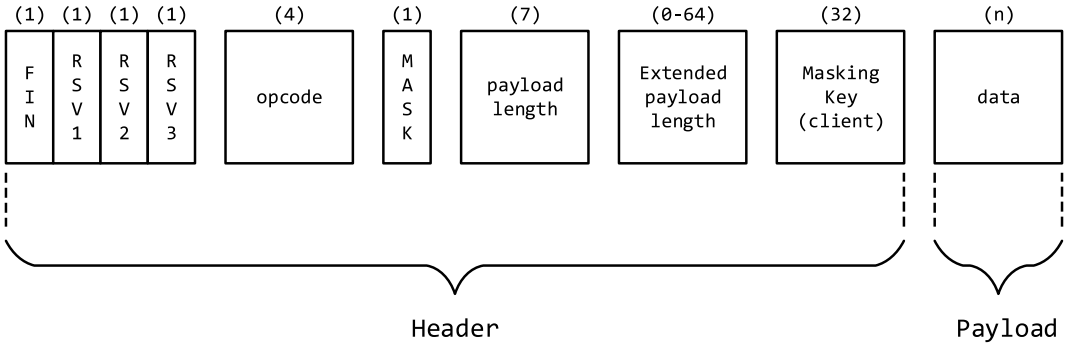

Abbildung 1: Vereinfachte Darstellung eines WebSocket Frames

deutlich geringeren Paketgrößen arbeiten kann als HTTP. Eine Nachricht mit einfacher Längenangabe, die von Server zu Client geschickt wird, erzeugt durch ihren Header lediglich einen Overhead von zwei Byte. Da WebSocket-Frames auf umgekehrtem Weg zusätzlich noch maskiert werden müssen, entspricht die minimale Länge eines clientseitigen Frames sechs Byte. Im Vergleich dazu beträgt der Overhead bei einer einfachen GET-Anfrage über HTTP/1.1 beim Aufruf des Hosts unter cs.hm.edu 35 Byte. Die Angabe des Hostnamens in HTTP/1.1 ist obligatorisch und trägt maßgeblich zur Größe des HTTP-Frames bei [Lea+99, S. 128].

[1] Web Hypertext Application Technology Working Group unter: www.whatwg.org

2.1.1 Header-Felder

Im nachfolgenden Abschnitt werden die Felder sowie deren jeweilige Aufgabe detailliert erläutert:

FIN (1 Bit) Das FIN-Flag gibt an, ob es sich bei dem empfangenen Frame um ein Fragment handelt. Kann der Payload nicht auf einmal übertragen werden, so wird die Übertragung über mehrere Frames gesteuert. Das FIN-Flag aller unvollständigen Fragmente weist den Wert 0 auf, wohingegen das finale Fragment mit einer 1 gekennzeichnet ist. Wenn alle Nutzdaten mit einem einzigen Frame übertragen werden können, so trägt dessen FIN-Flag ebenfalls den Wert 1.

RSV1, RSV2, RSV3 (jeweils 1 Bit) Diese Flags sind für zukünftige Implementierung reserviert und finden aktuell keine standardisierte Anwendung. Die Ausnahme stellt das RSV1-Bit dar, welches verwendet wird um anzuzeigen, ob eine Nachricht komprimiert wurde oder nicht [Int11]. Daher tragen sie meist den Wert 0. Ist eines dieser Flags gesetzt und die empfangende Stelle hat keine Erweiterung, welches das jeweilige Flag interpretieren kann, so muss die WebSocket-Verbindung geschlossen werden [FM11, S. 27].

opcode (4 Bit) Das opcode-Feld enthält die Angabe wie der Payload zu interpretieren ist. Neben den sogenannten Datenframes bzw. Non-Control-Frames, welche die Art der übertragenen Daten kennzeichnen, sind Control-Frames definiert, die zur Steuerung und Überprüfung der Verbindung verwendet werden. Wird ein unbekannter Opcode empfangen, so wird die WebSocket-Verbindung geschlossen.

- **0x0** definiert ein Fortsetzungs-Frame
- **0x1** definiert ein Text-Frame
- **0x2** definiert ein Binary-Frame
- **0x3-7** reserviert für weitere Non-Control-Frames
- **0x8** definiert ein Verbindung beenden-Frame
- **0x9** definiert ein Ping-Frame
- **0xA** definiert ein Pong-Frame
- **0xB-F** reserviert für weitere Control-Frames

MASK (1 Bit) Ist das MASK-Flag gesetzt, so ist der Payload maskiert. Frames, die clientseitig initiiert wurden, müssen maskiert sein, wohingegen serverseitige Frames keine Maskierung vornehmen dürfen. Der verwendete Masking-Key muss im gleichnamigen Header-Feld hinterlegt werden.

payload length (7 Bit) Hier wird die Länge der Nutzdaten angegeben. Sind diese nicht größer als 125 Bytes, so können sie mit nur einem Frame übertragen werden. Nutzdaten mit einem höheren Speicherplatzbedarf werden je nach Größe mit dem Wert 126 bzw. 127 angegeben. Diese Angaben führen zur Verwendung des Header-Felds *Extended payload length*.

Extended payload length (16 Bit oder 64 Bit) Überschreitet die Länge des Payloads 125 Byte, so wird dieses Feld verwendet. Je nach Wert im Header-Feld *payload length* umfasst die *Extended payload length* zwei Byte oder acht Byte. Dementsprechend beträgt die maximale Länge der Nutzdaten in einem WebSocket-Frame $2^{64} - 1$ Byte [GLN15, S. 41]

Masking Key (0 oder 32 Bit) Ein Frame, der von einem Client an einen Server gesendet wird, muss mit einem 32-Bit Schlüssel maskiert werden. Dieser wird vom Client generiert und im gleichnamigen Header-Feld abgelegt. Nachrichten, die von einem Server an einen Client gesendet werden, enthalten dieses Feld nicht.

data (n Byte) Hier befinden sich die eigentlichen Nutzdaten. Ihre Größe wird in den entsprechenden Längenfeldern des Headers hinterlegt. Werden lediglich Steuerungsinformationen übertragen, wird kein Payload mitgesendet.

2.1.2 Non-Control-Frames

Durch das Header-Feld *opcode* wird die Art der Nutzdaten angegeben. Es können sowohl textbasierte Daten als auch Binärdaten übertragen werden. Bei der Kommunikation über WebSockets gibt es hier einige zusätzliche Mechanismen, die für den fehlerfreien Ablauf einer Übertragung notwendig sind. Der folgende Abschnitt handelt von der bereits aus anderen Protokollen bekannten Fragmentierung, die WebSocket-Frames zur sequentiellen Übertragung unvollständiger Daten nutzt sowie der aufgrund von Sicherheitsproblemen eingeführten Maskierung clientseitig initierter Nachrichten.

Fragmentierung

In Abschnitt 5.4. des RFC6455 ist die Fragmentierung von WebSocket-Frames beschrieben. Primäre Anwendung erfährt dieses Verfahren, wenn die Daten zum Zeitpunkt der Übertragung noch nicht vollständig vorliegen oder nicht zwischengespeichert werden können. Wird die vom Server oder einer vermittelnden Stelle festgelegte Puffergröße überschritten, so wird ein Nachrichtenfragment mit dem Inhalt des Puffers gesendet [FM11, S. 32].

Folgende Grundsätze müssen bei der Fragmentierung beachtet werden. Erweiterungen werden hierbei außer Acht gelassen:

- Unfragmentierte Nachrichten werden mit einem einzelnen Frame übertragen (FIN-Bit = 1, opcode $\neq 0$)
- Die Übertragung fragmentierter Nachrichten erfolgt über mindestens einen Frame mit FIN-Bit = 0 und opcode $\neq 0$ und wird mit einem Frame FIN-Bit = 1 und opcode $\neq 0$ abgeschlossen. Die Summe der Payloads aller Fragmente entspricht dem Payload eines Einzelframes bei entsprechender Puffergröße.
- Control-Frames (Abschnitt 2.1.3) dürfen nicht fragmentiert werden. Sie können zwischen der Übertragung einer fragmentierten Nachricht gesendet werden, um beispielsweise die

Latenzzeit eines Pings möglichst gering zu halten. Die Verarbeitung zwischengesendeter Control-Frames muss von allen Endpunkten unterstützt werden.

- Der Empfänger muss die Fragemente einer Nachricht in der gleichen Reihenfolge erhalten, in welcher sie vom Sender aufgegeben wurden.
- Vermittler dürfen an der Fragmentierung keine Änderungen vornehmen, falls eines der RSV-Bits gesetzt ist und dessen Bedeutung dem Vermittler nicht bekannt ist.

Daraus folgt, dass die Verarbeitung sowohl fragmentierter als auch unfragmentierter Nachrichten von allen Endpunkten unterstützt werden muss. Da Control-Frames nicht fragmentiert werden dürfen, können Fragmente lediglich von den Typen Text, Binary oder einem der reservierten opcodes (0x3-7) sein. Die Angabe des Typs erfolgt durch den opcode des ersten Fragments. Alle weiteren Fragmente tragen den opcode 0x0, der dem Empfänger mitteilt, dass die empfangenen Nutzdaten an den Payload des vorangegangen Frames angehängt werden sollen [She17].

Maskierung

WebSocket Frames, von einem Client zu einem Server gesendet, müssen maskiert werden. Dieser Mechanismus war im ursprünglichen Draft für das WebSocket Protokoll nicht vorgesehen. In Folge einer Untersuchung durch eine Gruppe von Forschern rund um ein Team der Carnegie Mellon Universität wurden Sicherheitslücken in dem Konzept aufgedeckt. Bei der Verwendung transparenter Proxy-Server konnten sie durch deren fehlerhafte Implementierung des, bei dem Aufbau einer WebSocket Verbindung genutzten, *Upgrade* Mechanismus den Cache des Proxies infizieren [Hua+11]. Daraufhin wurde der Draft überarbeitet und um die Maskierung über eine bitweise XOR Verknüpfung ergänzt.

Da der zur Maskierung des Payloads verwendete Schlüssel im Frame enthalten ist, steht hier nicht die Vertraulichkeit der gesendeten Daten im Vordergrund. Vielmehr sollen Proxy-Server daran gehindert werden, den Inhalt des Payloads zu lesen. Somit wird verhindert, dass Angreifer den Payload manipulieren und diesen für einen Angriff gegen einen Proxy einsetzen können [GLN15]. Im Folgenden soll der Ablauf der Maskierung einer Textnachricht an einem Beispiel verdeutlicht werden:

Zunächst wählt der Client einen bisher nicht verwendeten 32 Bit Schlüssel aus, der zur Maskierung verwendet werden soll. Die Nachricht, die maskiert werden soll lautet "cs.hm.edu", konvertiert in hexadezimale Darstellung "63 73 2e 68 6d 2e 65 64 75". Als Schlüssel wird "3c 2e 3f 4a"verwendet. Dieser wird nun zyklisch auf den zu maskierenden Payload mit der XOR-Operation angewendet.

Die maskierte Nachricht wird mit dem Schlüssel in dem WebSocket-Frame abgelegt. Der Server kann diese nach dem Empfang durch erneute Anwendung der selbstinversen XOR-Operation demaskieren [GLN15].

Unmaskierte Bytes	63	73	2e	68	6d	2e	65	64	75
Operator	⊕	⊕	⊕	⊕	⊕	⊕	⊕	⊕	⊕
Schlüssel	3c	2e	3f	4a	3c	2e	3f	4a	3c
Maskierte Bytes	5f	5d	11	22	51	00	5a	2e	49

Tabelle 1: Maskierung durch bitweise XOR-Verknüpfung

Maskierte Bytes	5f	5d	11	22	51	00	5a	2e	49
Operator	⊕	⊕	⊕	⊕	⊕	⊕	⊕	⊕	⊕
Schlüssel	3c	2e	3f	4a	3c	2e	3f	4a	3c
Unmaskierte Bytes	63	73	2e	68	6d	2e	65	64	75
Nachricht	c	s	.	h	m	.	e	d	u

Tabelle 2: Demaskierung durch bitweise XOR-Verknüpfung

Die Übertragung von Server zu Client darf hingegen nicht maskiert werden. Empfängt ein Server eine unmaskierte bzw. ein Client eine maskierte Nachricht, so muss in beiden Fällen die Verbindung geschlossen werden. Dabei kann der im RFC6455 spezifierte Statuscode *1002 (protocol error)* beim Verbindungsabbau angegeben werden [FM11, S. 26].

2.1.3 Control-Frames

Control-Frames werden dazu genutzt, den Status einer WebSocket-Verbindung zu steuern und zu überwachen. Der RFC6455 unterscheidet dabei drei Arten von Control-Frames, welche anhand des *opcodes* unterschieden werden. Der Close-Frame wird zur Einleitung und Bestätigung einer schließenden Verbindung eingesetzt. Ping- bzw. Pong-Frames werden genutzt, um festzustellen ob der jeweilige Endpunkt noch erreichbar ist oder um eine Verbindung aktiv zu halten (keep-alive) [FM11, S. 35-36]. Im Folgenden werden die verschiedenen Arten der Control-Frames sowie deren Aufgaben erläutert.

Ping- und Pong-Frames

Wie eingangs erwähnt, dienen Ping- und Pong-Frames zur Verwaltung offener WebSocket-Verbindungen. Ping-Frames tragen den *opcode* 0x9, Pong-Frames hingegen den *opcode* 0xA. Das Übertragen von Nutzdaten bis zu 125 Byte durch einen Ping- oder Pong-Frame ist durch den RFC6455 zwar vorgesehen, kommt aber typischerweise nicht zum Einsatz [GLN15, S. 48]. Enthält ein Ping-Frame Nutzdaten, so müssen diese von einem Pong-Frame übernommen und zurückgesendet werden. Empfängt ein Endpunkt einen Ping-Frame, so muss schnellstmöglich mit einem Pong-Frame darauf reagiert werden. Werden mehrere Ping-Frames empfangen, bevor ein Pong-Frame zurückgesendet wurde, so kann lediglich der zuletzt eingetroffene Ping-Frame beantwortet werden [FM11, S. 36]. Ping- und Pong-Frames dienen aktuell als rein interner Mechanismus und können nicht über eine Schnittstelle direkt versendet werden [GLN15, S. 49].

	Ping-Frame	**Pong-Frame**
Byte	89 00	8a 80 (85 e1 ef 27)
Binär	1000 1001	1000 1010 (...)
Detail	FIN = 1, RSV n. gesetzt opcode = 0x9	FIN = 1, RSV n. gesetzt opcode = 0xA

Abbildung 2: Ping- und Pong-Frames zur Verbindungskontrolle

Bei den in Abbildung 2 dargestellten Frames handelt es sich um den Wireshark-Mitschnitt eines Heartbeats zwischen einem Client und dem Host unter *ws://echo.websocket.org*. Der Server sendet dabei einen Ping-Frame, um festzustellen, ob der Client noch erreichbar ist. Der Client antwortet mit einem korrespondierenden Pong-Frame. Da es sich um eine Nachricht von einem Client zu einem Server handelt, ist dieser Frame maskiert.

Close-Frames

Der Abbau einer WebSocket-Verbindung erfolgt durch den Versand eines Close-Frames. Der zugehörige *opcode* lautet 0x8. Empfängt ein Endpunkt einen Close-Frame, ohne bereits selbst einen solchen versendet zu haben, so muss er ebenfalls mit einem Close-Frame antworten. Wurde der initiale Close-Frame mit einem Status-Code versehen, so wird dieser in der Antwort übernommen [FM11, S. 35]. Wird ein Close-Frame empfangen, so können ausstehende Fragmente noch vor dessen Bestätigung gesendet werden. Allerdings kann dabei nicht garantiert werden, dass der Endpunkt, der den Verbindungsabbau eingeleitet hat, diese Nachrichten noch verarbeitet. Haben beide Endpunkte je einen Close-Frame gesendet und empfangen, so müssen sowohl die WebSocket- als auch die TCP-Verbindung getrennt werden. Der Server muss dies unverzüglich durchführen, wohingegen der Client auf den Server warten sollte [FM11, S. 36]. Sollte der Verbindungsabbau von beiden Endpunkten gleichzeitig eröffnet werden, so kann angenommen werden, dass die Verbindung geschlossen ist.

Byte	`88 82 (2a 1f fc f2) 29 f7`
Binär	`10001000 10000010 (...) 00101001 11110111`
Detail	`FIN = 1, RSV n. gesetzt`
opcode	`0x8`
Maskierung	`1`
Payload length	`2`
Payload (maskiert)	`)÷`

Abbildung 3: Close-Frame initiiert von Client

Abbildung 3 zeigt einen maskierten Close-Frame, der von einem Client abgesetzt wurde. Der maskierte Payload enthält dabei den Status-Code *1000*. Dieser wurde in der Antwort des Servers ebenfalls zurückgemeldet. Hierbei handelt es sich um die Angabe *Normal Closure*, welche besagt, dass die Verbindung ordnungsgemäß beendet wurde. Der RFC6455 definiert vier Status-Gruppen

Byte	`88 02 03 e8`
Binär	`10001000 00000010 00000011 11101000`
Detail	`FIN = 1, RSV n. gesetzt`
opcode	`0x8`
Maskierung	`0`
Payload length	`2`
Payload (unmaskiert)	`1000`

Abbildung 4: Close-Frame Antwort des Servers

mit folgender Bedeutung [FM11, S. 46]:

0-999 Dieser Bereich wird nicht genutzt.

1000-2999 Diese Status-Codes sind für die Definition innerhalb des RFC6455 sowie für zukünftige Überarbeitungen und Erweiterungen reserviert.

3000-3999 Status-Codes in diesem Bereich sind für Verwendung in Frameworks, Bibliotheken und Anwendungen gedacht und bedürfen einer Registrierung bei der Internet Assigned Numbers Authority (IANA).

4000-4999 In diesem Bereich können eigene Status-Codes definiert werden. Diese benötigen keine Registrierung und müssen den kommunizierenden Anwendungen bekannt sein.

Die Spezifikation des WebSocket-Protokolls definiert bereits einige Status-Codes innerhalb des zweiten Bereichs. Allerdings werden diese nur als Vorschlag angegeben, was inzwischen zu einem gewissen Wildwuchs und damit verbunden zu einigen Problemen führte [GLN15, S. 53]. Bisher wurden nur wenige Status-Codes neben jenen aus der Spezifikation registriert. Die bisher offiziellen Registrierungen finden sich auf den zugehörigen Seiten der IANA.[2] Tabelle 3 (S.12) zeigt einen kurzen Auszug der wichtigsten Status-Codes. Es fällt auf, dass alle Neuregistrierungen durch Alexey Melnikov erfolgten, welcher bereits an der Spezifikation des Protokolls beteiligt war. In einer der betreffenden Anfragen zur Registrierung eines Status-Codes fragt Melnikov auch, ob diese Status-Codes denn überhaupt genutzt werden und ob noch ein Interesse an deren Registrierung besteht [Mel12].

2.2 WebSocket Kommunikationsphasen

Das WebSocket-Protokoll basiert auf dem Transport Control Protocol (kurz: TCP) und durchläuft wie dieses verschiedene Phasen der Kommunikation. Zu Beginn wird ähnlich zu TCP ein Opening-Handshake zum Verbindungsaufbau durchgeführt. Auch am Ende der Datenübertragung wird ein solcher Handshake eingesetzt, um die Verbindung ordnungsgemäßg zu beenden. Die Ports, welche vom WebSocket-Protokoll verwendet werden, sind Port 80 (ungesicherte Verbindungen) und Port 443 (gesicherte Verbindungen). Aktuell wird deren Verwendung jedoch vom

[2] unter https://www.iana.org/assignments/websocket/

Status-Code	Bedeutung	Beschreibung	Spezifiert durch
1000	Normal Closure	Zeigt einen ordnungsgemäßen Verbindungsabbau an.	RFC6455
1002	Protocol Error	Der Endpunkt beendet die Verbindung aufgrund eines Protokollfehlers. Beispielsweise wird der Empfang einer nicht maskierten Nachricht von einem Server durch selbigen mit diesem Status-Code beendet.	RFC6455
1006	Abnormal Closure	Dieser Status darf nicht in einem Close-Frame verwendet werden. Anwendungen können durch diesen signalisieren, dass die Verbindung unerwartet und auf unbekannte Weise beendet wurde.	RFC6455
1010	Mandatory Ext.	Ein WebSocket-Client kann diesen Status-Code verwenden, wenn er die Verbindung aufgrund fehlender Erweiterungen beendet. Ein Server verwendet diesen Status-Code nicht, da er aufgrund mangelnder clientseitiger Erweiterungen den Opening-Handshake abbrechen kann.	RFC6455
1012	Service Restart	Gibt an, dass der Dienst neustartet. Verbundene Clients sollen im Falle einer Wiederherstellung der Verbindung eine zufällige Zeit zwischen 5-30 Sekunden warten [Mel12].	Alexey Melnikov
1013	Try Again Later	Signalisiert, dass der Dienst überlastet ist. Sollte das Ziel über mehrere IP-Adressen verfügen, kann eine andere verwendet werden. Ist nur eine IP-Adresse verfügbar, so soll der Benutzer entscheiden, ob er sich erneut verbinden möchte [Mel12].	Alexey Melnikov

Tabelle 3: Auszug der registrierten Status-Codes

HTTP(S)-Protokoll dominiert und sie müssen daher geteilt werden. Der Verbindungsaufbau wird daher als HTTP-Upgrade Anfrage initiiert. Zukünftige Implementierungen könnten dies jedoch durch einen einfacheren Mechanismus über einen dedizierten Port ablösen, ohne das WebSocket-Protokoll umgestalten zu müssen [FM11, S. 3]. Im Folgenden Abschnitt werden die notwendigen Schritte zum Verbindungsaufbau und -abbau erläutert, sowie ein beispielhafter Ablauf der Übertragung über eine WebSocket-Verbindung gezeigt.

2.2.1 Verbindungsaufbau

Da Verbindungen über die Ports 80 bzw. 443 in den meisten Firewalls zugelassen werden, bedarf es keiner neuen Freigaben bei der Verwendung von WebSockets. Allerdings sind diese Ports oftmals durch einen HTTP-Server belegt. Aus diesem Grund bedarf es beim Aufbau einer WebSocket-Verbindung der Kooperation mit HTTP. Der Client sendet dazu einen HTTP-

Upgrade-Request, der um spezielle Felder erweitert wurde.

```
GET /dako-backend/advancedchat HTTP/1.1
Host: localhost:8080
Connection: Upgrade
Upgrade: websocket
Origin: localhost:8080
Sec-WebSocket-Version: 13
Sec-WebSocket-Key: 1zUsRVYbBDvRKqqZq9cRsg==
Sec-WebSocket-Extensions: permessage-deflate; client_max_window_bits
(...)
```

Abbildung 5: Opening-Handshake des Clients

Der Client stellt in Abbildung 5 einen solchen HTTP-Upgrade-Request an den Endpunkt des WebSocket-Servers unter `localhost:8080`. Durch `Connection: Upgrade` wird dem Server mitgeteilt, dass auf ein anderes Protokoll gewechselt werden soll, in diesem Fall `websocket`. Die Versionsnummer entspricht den unterschiedlichen IETF Drafts. Über `Sec-WebSocket-Extensions` werden die Erweiterungen angefragt, die der Client unterstützt. Die übermittelte Erweiterung `permessage-deflate` wird in Kombination mit dem Parameter `client_max_window_bits` angegeben, der die Sliding Window Größe limitiert, mit der der Client die Nachricht komprimiert. So kann der zu reservierende Speicherbedarf bei der serverseitigen Dekompression reduziert werden [Yos15, S. 18]. Der Client generiert den Base64-codierten `Sec-WebSocket-Key`, der anschließend vom Server zur Empfangsbestätigung des Handshakes verwendet wird.

```
HTTP/1.1 101
Connection: Upgrade
Upgrade: websocket
Sec-WebSocket-Accept: mi8r8dWsVBJS1B3908SLbnd5tHM=
Sec-WebSocket-Extensions: permessage-deflate; client_max_window_bits=15
(...)
```

Abbildung 6: Opening-Handshake des Servers

Damit der Server bestätigen kann, dass der Handshake eingetroffen ist, wird der übermittelte `Sec-WebSocket-Key` mit dem Globally Unique Identifier (GUID) `258EAFA5-E914-47DA-95CA-C5AB0DC85B11` verbunden und ein SHA-1 Hashwert gebildet. Dieser Wert wird in einer Base64-Kodierung im Feld `Sec-WebSocket-Accept` übermittelt. Der Client kann beim Empfang der Nachricht anhand des gleichen Vorgehens einen SHA-1 Hash erzeugen und mit dem empfangenen vergleichen.

1. Kombinierter Wert `1zUsRVYbBDvRKqqZq9cRsg==258EAFA5-E914-47DA-95CA-C5AB0DC85B11`

2. SHA-1 Hashwert `9A2F2BF1D5AC541252D41DFDD3C48B6E7779B473`

3. Base64 Kodierung `mi8r8dWsVBJS1B3908SLbnd5tHM=`

Des weiteren teilt der Server dem Client mit, dass die Erweiterung `permessage-deflate` unterstützt wird. Die maximale Größe eines Sliding Windows beträgt demnach 2^{15} `Byte` $= 32$ `KB`. Als

HTTP-Status wird der Code `101 Switching Protocols` angegeben. Dieser wird eingesetzt, um zu signalisieren, dass absofort das WebSocket-Protokoll genutzt wird. Weitere Anwendung findet dieser Status-Code beim Wechsel von HTTP-Versionen [Lea+99]. Meldet der Server einen anderen Status, so wird angenommen, dass der WebSocket Handshake nicht vollständig durchgeführt werden konnte und die Semantik von HTTP gilt [FM11, S. 7]. Ebenso wird verfahren, wenn der vom Client ermittelte Hashwert nicht mit dem vom Server übertragenen `Sec-WebSocket-Accept` übereinstimmt.

Optionale Felder können ebenfalls übertragen werden. Hierzu zählt beispielsweise das Header-Feld `Sec-WebSocket-Protocol`, das angibt welches Sub-Protokoll für die Übertragung genutzt werden soll. Hat der Server den Opening Handshake erfolgreich zurückgesendet, so gilt die WebSocket-Verbindung als geöffnet und Daten können gesendet werden.

2.2.2 Datentransfer

Befindet sich die WebSocket-Verbindung nach dem erfolgreichen Opening-Handshake im Status `OPEN`, so können Daten gesendet werden. Wie in Abschnitt 2.1.1 erläutert, definiert das Header-Feld `opcode` eines WebSocket-Frames die übertragene Art von Nutzdaten. Aktuell sind lediglich Text- und Binärdaten offiziell spezifiert.

Byte	81 0e [57 65 62 53 6f 63 6b 65 74 20 32 30 31 38]
Header	FIN = 1, RSV = 0, opcode = 0x1, Payload length = 14
[Payload]	WebSocket 2018

Abbildung 7: Datentransfer von Server zu Client

In Abbildung 7 ist eine Nachricht mit dem Inhalt `WebSocket 2018` von einem Server an einen Client gesendet worden. Als `opcode` wurde `0x1` angegeben, welcher besagt, dass es sich bei der Art des Payloads um eine Textnachricht handelt. Die Größe der Nutzdaten beträgt 14 Byte und die Nachricht konnte ohne Fragmentierung vollständig übertragen werden.

2.2.3 Verbindungsabbau

Der Abbau einer WebSocket-Verbindung erfolgt anhand eines Closing-Handshakes. Er soll den `FIN/ACK` Closing-Handshake von TCP ergänzen, da dieser insbesondere bei der Verwendung von Proxy Servern oder Vermittlungsstellen nicht garantiert zwischen genau zwei Endpunkten abläuft [FM11, S. 9]. Laut RFC6455 sollen so Szenarien vermieden werden, in denen Daten unnötigerweise verloren gehen.

Die Einleitung eines Verbindungsabbaus kann sowohl vom Client, als auch vom Server erfolgen. Dazu wird ein Frame mit dem `opcode`-Wert `0x8` an den Kommunikationspartner versendet. Der versendende Endpunkt wechselt mit dem Absenden des Closing-Frames vom `OPEN` in den `CLOSING` Status. Ebenso gilt dies beim Empfang eines Closing-Frames insofern selbst noch keiner versendet wurde. Danach eintreffende oder zu sendende Nachrichten müssen verworfen werden. Haben

beide Endpunkte sowohl einen Closing-Frame gesendet, als auch empfangen, so gilt die Verbindung als beendet und die darunterliegende TCP-Verbindung kann geschlossen werden. Dies sollte stets vom Server begonnen werden, damit er im `TIME WAIT` Status verbleibt. Der Client kann so die Verbindung durch das Senden eines neuen `SYN` unverzüglich wieder eröffnen [FM11, S. 41].

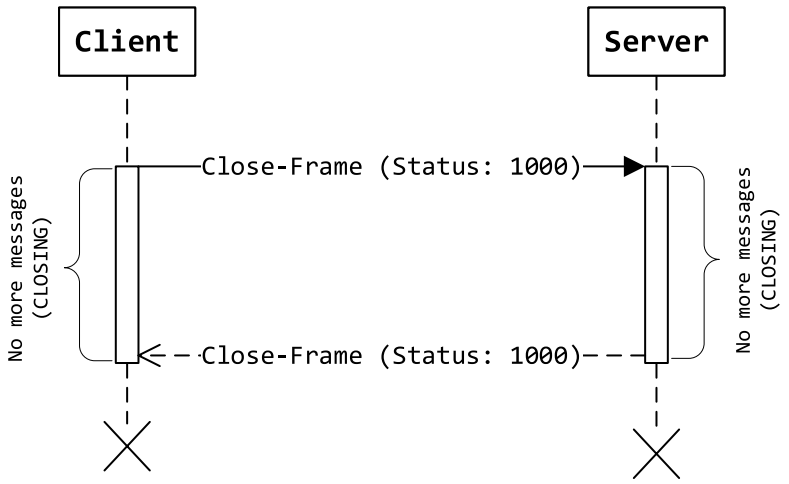

Abbildung 8: Clientseitiger Closing-Handshake

Abbildung 8 zeigt den schematischen Ablauf des Closing-Handshakes. Der Client initiiert diesen, indem er einen Closing-Frame mit dem Statuscode 1000 versendet. Der Server reagiert mit einer entsprechenden Antwort, wobei der Status aus dem ersten Frame übernommen wurde. Nachdem der Closing-Handshake abgeschlossen wurde befinden sich beide Endpunkte im Status `CLOSED` und die zugrundeliegende TCP-Verbindung kann beendet werden.

2.3 Alternative Verfahren

Die Kommunikation mittels HTTP basiert auf einem Request/Response Verfahren. Dabei sendet immer der Client eine Anfrage an einen Server. Dieser antwortet mit einer Statuszeile und entsprechendem Statuscode. Anschließend gilt die Übertragung als beendet und kann durch eine erneute Anfrage des Clients neu gestartet werden. Eine Möglichkeit Daten vom Server an den Client zu senden, ist in der Spezifikation von HTTP nicht vorgesehen. Verändert sich der Zustand einer serverseitigen Ressource, so kann der Client darüber bei einer reinen HTTP-Verbindung nur durch erneutes Anfragen informiert werden. Allerdings benötigen viele Anwendungsfälle diese Funktion. Aus diesem Grund wurde eine Vielzahl von Verfahren entwickelt, die diese Mechanismen abbilden sollen. Im Folgenden werden einige dieser Verfahren kurz vorgestellt und hinsichtlich ihrer Vor- und Nachteile gegenüber WebSocket analysiert.

2.3.1 Polling

Durch die vom W3C standardisierte API XMLHttpRequest (XHR) wurde im Jahr 2012 eine Möglichkeit geschaffen, mittels JavaScript HTTP-Requests zu konstruieren [WHA06]. Ein XHR-Objekt registriert verschiedene Methoden, welche bei einer Änderung des Objekts asynchron aufgerufen werden. Dadurch können Anfragen gesendet werden, ohne den weiteren Programmablauf zu blockieren. Dies bildet die technische Grundlage vieler moderner Web-Frontends, die unter dem Oberbegriff Ajax (*Asynchronous JavaScript and XML*) zusammengefasst werden [GLN15, S. 26].

Beim Polling-Verfahren wird eine Anfrage an den Server in festgelegten Zeitabständen wiederholt. Durch die Verwendung eines XHR-Objekts blockiert die Anwendung zwischen dem Absenden und Empfang der Serverantwort nicht. Liegt keine Änderung vor, so sendet der Server eine leere Antwort zurück.

Dieses Verfahren wird von allen Browsern unterstützt und bedarf keiner besonderen Kenntnisse. Allerdings ist der erzeugte Overhead je nach Einsatzszenario nicht zu unterschätzen. Auch die Belastung des Servers bei der Verarbeitung der eintreffenden Anfragen kann problematisch sein. Treten demnach viele Änderungen in kurzen Zeitabständen auf, so ist von der Verwendung eines einfachen Polling-Verfahrens abzuraten.

2.3.2 Long-Polling

Die regelmäßigen Abfragen des einfachen Pollings führen zu einem enormen Overhead. Um diesen zu reduzieren wird beim Long-Polling nur eine Antwort vom Server zurückgesendet, wenn eine Änderung vorliegt. Der Client reagiert umgehend mit einem neuen Request auf die Antwort des Servers. Die Verbindung bleibt dabei geöffnet, solange keine Änderung auftritt. Läuft die vordefinierte Wartezeit ab, so antwortet der Server auch hier mit einer leeren Nachricht [GLN15, S. 29].

Das Long-Polling reduziert den Overhead gegenüber dem einfachen Polling-Mechanismus deutlich. Auf Grund des sofortigen Wiederaufbaus durch den Client wird eine nahezu andauernde Verbindung simuliert. Dennoch ist auch in diesem Fall der Overhead deutlich größer als bei der Verwendung einer WebSocket-Verbindung. Mit steigender Frequenz von Nachrichten nähert sich dieser an den Overhead des einfachen Pollings.

2.3.3 Comet

Auf Grund diverser Limitierungen für Mehrbenutzeranwendungen, entwickelten sich Programmiermodelle, welche unter der Bezeichnung *Comet* [Rus06] zusammengefasst wurden. Da es sich hierbei nicht um einen spezifizierten Standard handelt, wurde im Rahmen des CometD-Projekts der Dojo Foundation das Bayeux-Protokoll ins Leben gerufen. Ziel des Projekts ist die Reduzierung der Komplexität bei der Implementierung von Comet Anwendungen [Rus+07]. Um dies zu

erreichen wurden Referenzimplementierungen des Bayeux-Protokolls in verschiedenen Programmiersprachen bereitgestellt.

Neben der Verwendung des Long-Polling Mechanismus zur ereignisgesteuerten Nachrichtenübertragung von Server zu Client, sieht das Bayeux-Protokoll auch Streaming-Techniken vor. Dabei wird die HTTP-Verbindung nach der Antwort des Servers auf den Request des Clients nicht geschlossen, sondern anhand des Connection-Headers `keep-alive` weiterhin offen gehalten. Somit können weitere Nachrichten vom Server an den Client gesendet werden, ohne dass eine erneute Anfrage durch den Client gestellt werden muss.

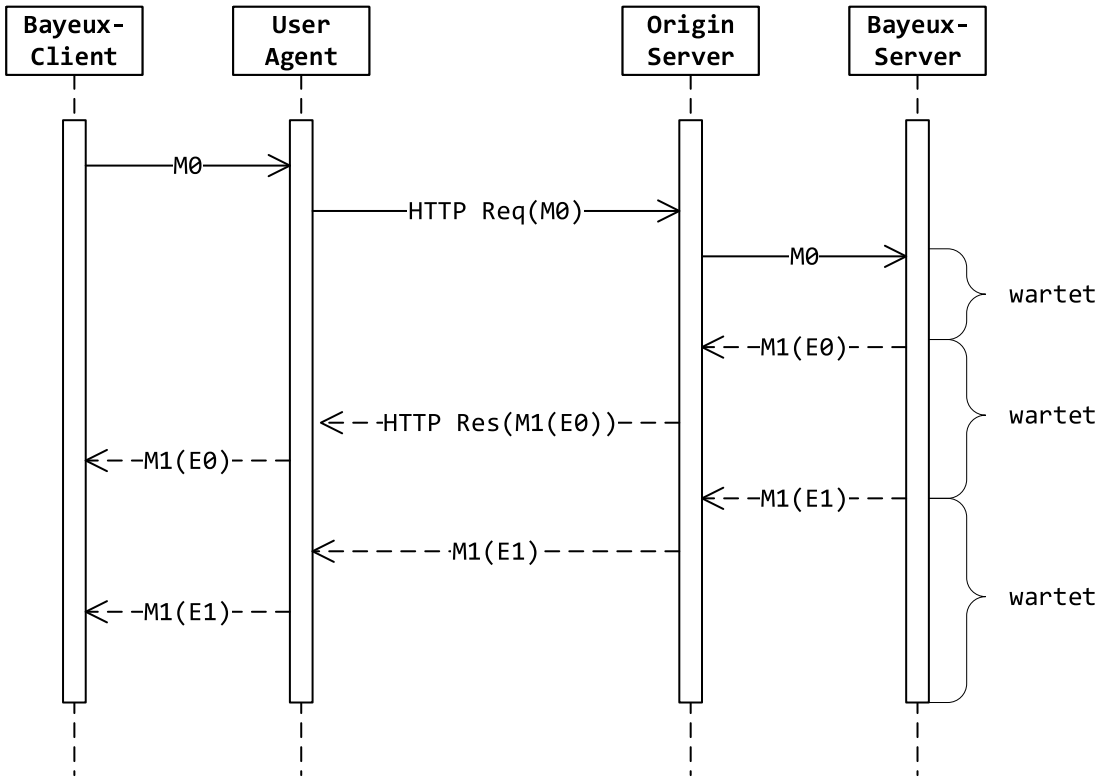

Abbildung 9: HTTP-Streaming im Bayeux-Protokoll

Allerdings muss die Verwendung von HTTP-Streaming von den User Agenten unterstützt werden, da unvollständige HTTP Responses für dieses Verfahren genutzt werden [Rus+07]. Um bidirektional kommunizieren zu können verwendet das Bayeux-Protokoll zwei HTTP-Verbindungen, wohingegen das WebSocket-Protokoll mit einer einzigen Verbindung arbeitet. Der größte Nachteil des Bayeux-Protokolls ist jedoch die fehlende Standardisierung durch eine etablierte Organisation.

2.3.4 Server-Sent Events

Die bislang vorgestellten Verfahren bieten keine standardisierte API, die es Servern ermöglicht Daten über HTTP an eine Website zu senden. In HTML5 wurde dafür vom W3C die sogenannten *Server-Sent Events (SSE)* eingeführt, welche ein `EventSource`-Objekt bereitstellen. Dieses sorgt auf der Clientseite dafür, dass Benachrichtigungen des Servers empfangen werden können. Dazu stellt der Browser drei Event-Handler bereit, die DOM-Events an die Anwendung weiterleiten [GLN15, S. 31]. Bei der Öffnung einer SSE-Verbindung wird der Event-Handler `onopen` aufgerufen, im Falle eines Fehlers der `onerror`-Handler. Kommen Nachrichten des Servers an, so wird der `onmessage`-Handler aktiviert [Hic09]. Zu jedem Event-Handler werden entsprechende Methoden registriert, welche die Bearbeitung des jeweiligen Events übernehmen.

Server-Sent Events erzeugen keinen Overhead durch wiederholte Anfragen an den Server. Die Kommunikation von Server zu Client wird ermöglicht, ohne mehrere HTTP-Verbindungen zu öffnen. Allerdings handelt es sich hier um einen unidirektionalen Kanal von Server zu Client. Dies ist für Anwendungsfälle ausreichend, in denen lediglich der Server Informationen zu übermitteln hat. Beispiele hierfür wären u.a. das Hardware-Monitoring eines Server oder ein Newsticker. Sollen jedoch Daten in beide Richtungen übertragen werden, so ist die Verwendung von WebSockets die bessere Alternative. Insbesondere Mehrbenutzeranwendungen wie Spiele oder Chats, die auf kurze Latenzen angewiesen sind, profitieren von derem bidirektionalem Kommunikationskanal.

3 Die WebSocket API

3.1 Client-seitige Implementierung

3.2 Server-seitige Implementierung

4 Fachliche Anforderungen

4.1 Analyse des DaKo-Frameworks

4.2 Anforderungen an WebSocket-Chat

4.3 Prototypische Implementierung

5 Technische Anforderungen

5.1 Entwurf des WebSocket-Chats

5.2 Vergleich der Implementierungsformen

5.3 Implementierung des WebSocket-Chats

6 Evaluation

6.1 Test-Umgebung

6.2 Leistungsanalyse WebSocket-Chat

7 Zusammenfassung und Ausblick

7.1 Einsatzgebiete

Literatur

[Fet10] Ian Fette. *The WebSocket Protocol*. 31. Aug. 2010. URL: https://tools.ietf.org/html/draft-ietf-hybi-thewebsocketprotocol-01 (besucht am 19. 12. 2017).

[FM11] Ian Fette und Alexey Melnikov. *The WebSocket Protocol*. 1. Dez. 2011. URL: https://tools.ietf.org/html/rfc6455 (besucht am 19. 12. 2017).

[GLN15] Peter Leo Gorski, Luigi Lo Iacono und Hoai Viet Nguyen. *WebSockets: moderne HTML5-Echtzeitanwendungen entwickeln*. OCLC: 901007083. München: Hanser, 2015. 269 S. ISBN: 978-3-446-44371-6 978-3-446-44438-6.

[Hic09] Ian Hickson. *Server-Sent Events*. 23. Apr. 2009. URL: https://www.w3.org/TR/2009/WD-eventsource-20090423/ (besucht am 28. 12. 2017).

[Hic10] Ian Hickson. *The WebSocket Protocol*. 6. Mai 2010. URL: https://tools.ietf.org/search/draft-hixie-thewebsocketprotocol-76 (besucht am 19. 12. 2017).

[Hua+11] Lin-Shung Huang u. a. „Talking to Yourself for Fun and Profit“. In: *Proceedings of W2SP* (1. Jan. 2011), S. 1–11.

[Int11] Internet Assigned Numbers Authority. *WebSocket Protocol Registries*. 21. Okt. 2011. URL: https://www.iana.org/assignments/websocket/websocket.xhtml#opcode (besucht am 25. 12. 2017).

[Lea+99] Paul J. Leach u. a. *Hypertext Transfer Protocol – HTTP/1.1*. 1. Juni 1999. URL: https://tools.ietf.org/html/rfc2616 (besucht am 20. 12. 2017).

[Mel12] Alexey Melnikov. *Additional WebSocket Close Error Codes*. 16.5.12. URL: http://www.ietf.org/mail-archive/web/hybi/current/msg09670.html (besucht am 25. 12. 2017).

[Rus+07] Alex Russell u. a. *The CometD Reference Book – 3.1.3*. 1. Jan. 2007. URL: https://docs.cometd.org/current/reference/#_bayeux (besucht am 28. 12. 2017).

[Rus06] Alex Russell. *Comet: Low Latency Data for the Browser – Infrequently Noted*. 3. März 2006. URL: https://infrequently.org/2006/03/comet-low-latency-data-for-the-browser/ (besucht am 28. 12. 2017).

[She17] Eric Shepherd. *Writing WebSocket Servers*. 16.8.17. URL: https://developer.mozilla.org/en-US/docs/Web/API/WebSockets_API/Writing_WebSocket_servers (besucht am 22. 12. 2017).

[WHA06] WHATWG. *XMLHttpRequest Standard*. 1. Jan. 2006. URL: https://xhr.spec.whatwg.org/ (besucht am 27. 12. 2017).

[Yos15] Takeshi Yoshino. *Compression Extensions for WebSocket*. 24. Aug. 2015. URL: https://tools.ietf.org/html/draft-ietf-hybi-permessage-compression-28#section-4 (besucht am 26. 12. 2017).

Abbildungsverzeichnis

Tabellenverzeichnis