

# Projektseminar Rechnersysteme

**Seminararbeit vorgelegt von (Robert Wiesner)**

Betreuer: Dipl.-Ing. (Changgong Li)

Beginn: (Beginndatum) | Abgabe: (Abgabedatum)

Institut für Datentechnik | Fachgebiet Rechnersysteme | Prof. Dr.-Ing. Christian Hochberger



TECHNISCHE  
UNIVERSITÄT  
DARMSTADT



---

# 1 Aufgabenstellung

Entwicklung eines "Spill and Fill"-Mechanismus um den Framestack des AMIDAR Prozessors auf einen externen DDR3 Speicher auszulagern und so die maximal mögliche Stackgröße deutlich zu erhöhen. Dazu gehört das einrichten verschiebbarer "windows" die jeweils einen Ausschnitt des Stacks beinhalten. Bei einem Methodenaufruf, dessen Stackframe den aktuellen Ausschnitt überschreiten würde soll ein Teil der im Window enthaltenen Daten auf den externen Speicher ausgelagert werden um Platz für den neuen Stackframe zu schaffen. Bei einem Methoden Rücksprung sollen die Daten wieder zurück kopiert werden um den alten Stackframe wiederherzustellen. Es sollen 4 Windows angelegt werden um die Stacks von bis zu 4 threads vorzuhalten. Es soll die Möglichkeit einer Threadverdrängung realisiert werden, damit mehr als 4 Threads ausgeführt werden können.

---

## 2 Grundlagen

---

### 2.1 Spill and Fill

---

---

### 2.2 Verwendung eines Spill and Fill Mechanismus in anderen Prozessoren

---

---

#### 2.2.1 Sun Sparc Prozessoren

---

Die Sun Sparc Prozessoren V8/V9 verfügen über ein sliding register Window mit jeweils 16 64 Bit großen Registern in 7 Registersätzen. Sliding Window bezeichnet ein Verfahren, bei dem die Registersätze im Falle eines Funktionsaufruf nicht auf dem Stack gesichert werden müssen, stattdessen wird auf den nächsten Registersatz gewechselt. Dabei wird meistens auch die Parameterübergabe realisiert. Dabei sind die Input Register in dem Registerwindow des Callers identisch mit dem Output Registern, in dem Registerwindow des Callee. Mit speziellen Bytecode Instruktionen werden im Spill an Fill Verfahren Registersätze auf einen externen Speicher gesichert, was allerdings wegen den großen Registersätzen recht viel Zeit in Anspruch nimmt. Da das Spill and Fill durch spezielle Bytecodes ausgelöst wird muss der Compiler vorher festlegen an welchen Stellen das Spill and Fill stattfinden soll.

---

### 2.3 AMIDAR

---

Bei der Klasse der AMIDAR Prozessoren handelt es sich um ein konfigurierbares System bestehend aus Function Units (FU) für die jeweiligen Aufgaben. Die FUs sind dabei untereinander mit einen Token- und einem Datennetzwerk verbunden.

---

#### 2.3.1 Funktion

---

---

#### 2.3.2 Framestack

---

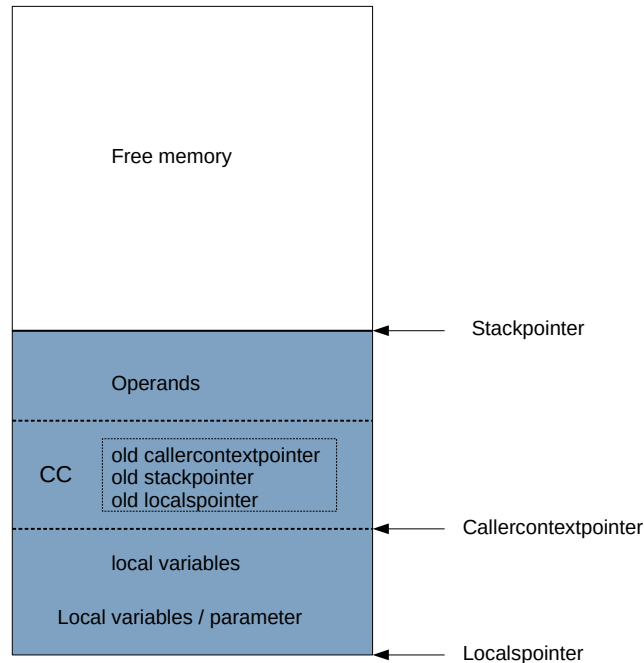
Die Framestack FU dient dazu die Funktionen des Operand Stacks und des Speichers für lokale Variablen zur Verfügung zu stellen. Der AMIDAR Framestack arbeitet bei der Stackframe Verarbeitung mit drei grundsätzlichen Zeigern. Der Stackpointer gibt die nächste freie Speicheradresse über den aktuellen Stackframe an. Der Locals Pointer gibt die unterste Lokale Variable und damit auch das untere Ende des Stackframes an. Die dritte wichtige Zeiger ist der Callercontext Pointer, der auf die unterste Adresse des Callcontext deutet.

---

#### Stackframe

---

Ein Stackframe beginnt mit den Parametern und den Lokalen Variablen. Darüber kommt der Callercontext mit den alten Localspointer, Stackpointer und Callercontextpointer. Das obere Ende des Stackframes besteht aus dem Callercontext.



before Invoke.pdf

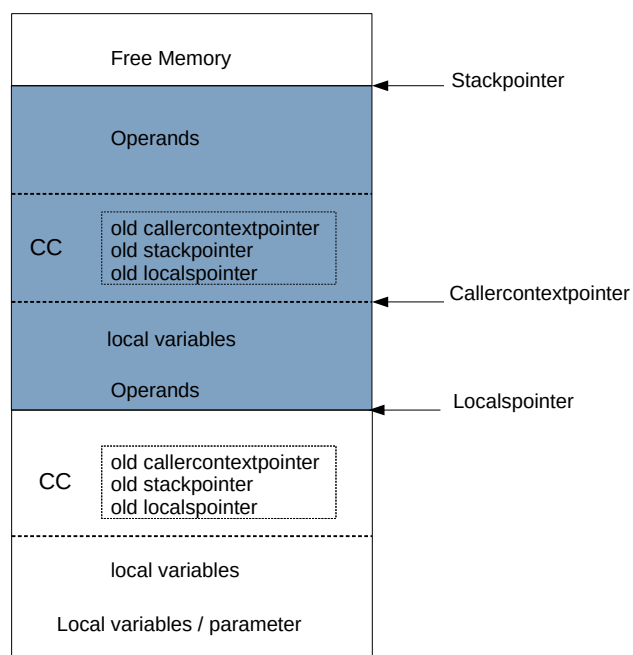
**Abbildung 2.1: Stackframe vor einen Funktionsaufruf**

## Funktionsaufrufe

Eine für dieses Projekt wichtige Funktion des Framestacks sind Funktionsaufrufe. Bei einen Funktionsaufruf werden im Framestack die die 3 wichtigen Pointer neu gesetzt. Und der alte Callercontext gesichert. Der neue Localspointer wird berechnet in dem man vom aktuellen Stackpointer die Anzahl der Parameter der aufgerufenen Methode abzieht. Der neue Callercontextpointer wird berechnet indem auf dem neuen Localspointer die mit den Token übergebene Anzahl lokaler Parameter drauf addiert wird. Der neue Stackpointer wird berechnet, in dem auf dem Callercontextpointer die größe des Callercontext drauf addiert wird. In den nächsten Takten werden die alten Localspointer, Stackpointer und in den Callercontextpointer geschrieben.

## Funktionsrückprung

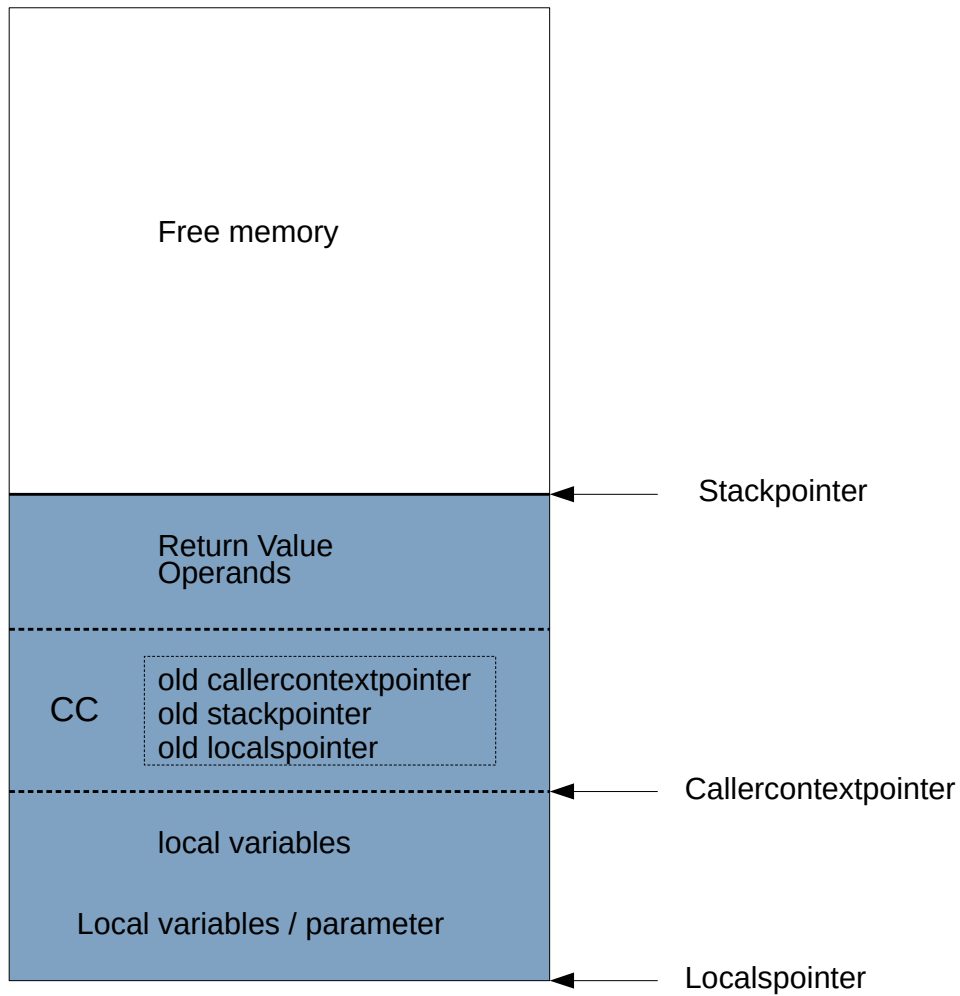
Bei einen Funktionsrückprung werden im AMIDAR Framestack die Locals- Stack- und Callercontextpointer aktualisiert. Der Vorgang beginnt wenn eines der entsprechenden Token über den AMDIAR Bus gesendet wird. Es gibt 3 Varianten des Tokens. Einer liefert keinen Rückgabewert, die anderen geben jeweils 32 oder 64 Bit zurück. Wenn der Token abgearbeitet wird, werden nacheinander die drei Pointer ausgelesen und wiederhergestellt. Anschliessend wird gegebenenfalls der Rückgabewerte gesichert dieser steht in den "top of Stack" bzw. "next of stack" Registern. Der Rückgabewert wird an stelle des obersten, im Falle eines Return64 der oberen beiden, Operanden des Ursprünglichen Funktionsaufruf gespeichert.



after Invoke.pdf

**Abbildung 2.2:** Stackframe nach einen Funktionsaufruf

after return.pdf



**Abbildung 2.3:** Stackframe nach einen Funktionsrückprung

---

## 3 Implementierung

---

### 3.1 DRAM Anbindung über AXI

---

AMIDAR ist über AXI mit 512 MB DDR Ram verbunden, der auf 800 MHz getaktet.

---

#### 3.1.1 Addressbereich des Framestacks im Hauptspeicher

---

Eine Schwierigkeit bei der Auslagerung des Framestacks an den Hauptspeicher ist die Unterschiedliche Wortbreite. Für den Garbagecollector werden zu jedem 32bit Datenwort noch 2 Statusbit gespeichert in den festgehalten wird, ob es sich bei den Daten um Metadaten, Referenzen oder Werte handelt. Damit beim Auslagern der Daten auf dem externen Speicher diese Informationen nicht verloren gehen werden die Schreibvorgänge immer in Blöcken von 16 Wörtern durchgeführt, wobei die Statusbits der 16 Wörter als weiteres 32Bit Wort in den Speicher geschrieben werden. Dies ist kein Problem, da die Größe der beim Spill and Fill Vorgang übertragenen Bereiche einen Vielfachen von 16 Wörtern entspricht.

Der Ausgelagerter Framestack wird in das obere Ende des 512MB groÄßen DDR3 Speicher gelegt und belegt:

$(numbermaxthreads * wordsperthread) * (1 + 1/16)$  Wörter.

Als Ausgangspunkt wird die höchste Speicheradresse des DDR Speichers verwendet und 2 mal nach links geschiftet, da der externe Speicher Byte adressiert, der Framestack jedoch wortadressiert ist. Davon wird die Maximale Größ des Framestacks abgezogen. Von da an aufwärts werden die Daten der einzelnen Threads gespeichert.

---

#### 3.1.2 Schreiben von Framestackinhalten in den Hauptspeicher

---

Für das Zwischenspeichern und packen der Blöcke wird ein eigenes Modul verwendet. Vom Framestack aus kann auf dieses Modul geschrieben werden. Dabei wird die Startadresse des Blocks, die länge des AXI Bursts und das erste Datenwort angelegt. Zum Zwischenspeichern dieser Daten werden jeweils Fifos verwendet. Wobei die Fifos für Burstlänge und Startadresse einen Sechzehntel der Größe des Datenfifos haben. Für jedes übertragene Datenwort werden 2 Statusbits in einen 32 Bit Register geschrieben. Nach dem 16 Datenworte in den Datenfifo geschrieben wurden wird der Inhalt dieses Registers in den Fifo geschrieben. Das Modul, indem die eigentliche AXI Übertragung verarbeitet wird wartet darauf das der Adress- und Burstlengthfifo nicht leer sind. Wenn die der Fall ist wird die Adresse und Burstlänge ausgelesen und wie folgt umgerechnet:

Anschließend werden die Daten aus dem Fifo übertragen bis die neu berechnete Burstlänge erreicht wurde. Solange der noch Einträge in den Fifos für Burslänge und Adresse liegen wird eine neue Übertragung gestartet.

---

#### 3.1.3 Lesen von Framestackinhalten aus den Hauptspeicher

---

Wenn die Daten ausgelesen werden sollen, wird das selbe Modul verwendet. Es wird die Startadresse des auszulesenden Blocks und die Burstlänge übergeben. Diese wird für den DDR Speicher umgerechnet

## Blocks in Ram.pdf

Tabelle1

Framestack data format

Address	Framestack Entry	
	2 Bit	32Bit
0	Type 0	Data 0
1	Type 1	Data 1
2	Type 2	Data 2
3	Type 3	Data 3
4	Type 4	Data 4
5	Type 5	Data 5
6	Type 6	Data 6
7	Type 7	Data 7
8	Type 8	Data 8
9	Type 9	Data 9
10	Type 10	Data 10
11	Type 11	Data 11
12	Type 12	Data 12
13	Type 13	Data 13
14	Type 14	Data 14
15	Type 15	Data 15
16	Type 16	Data 16
17	Type 17	Data 17
18	Type 18	Data 18

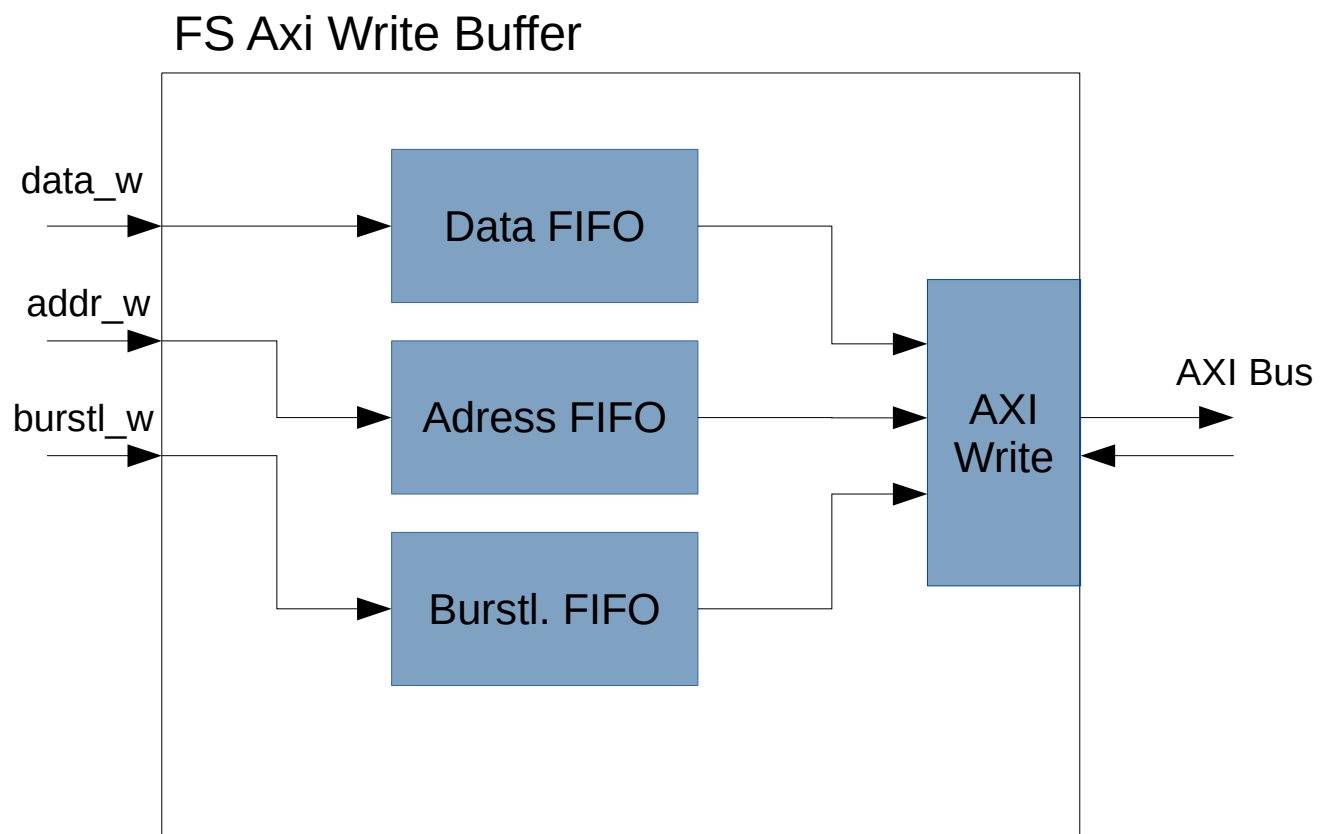
Framestack Data in external RAM

Address *	Data 32 Bit
0	Data 0
1	Data 1
2	Data 2
3	Data 3
4	Data 4
5	Data 5
6	Data 6
7	Data 7
8	Data 8
9	Data 9
10	Data 10
11	Data 11
12	Data 12
13	Data 13
14	Data 14
15	Data 15
16	Type 0-15
17	Data 16
18	Data 17
19	Data 18

\*Nicht die eigentliche Adresse im RAM, sondern der Offset im Speicherbereich des Threads

**Abbildung 3.1:** Zuordnung der Framestack Adressen im Hauptspeicher





**Abbildung 3.2:** Schema des AXI Schreib Puffers

## FS AXI Read

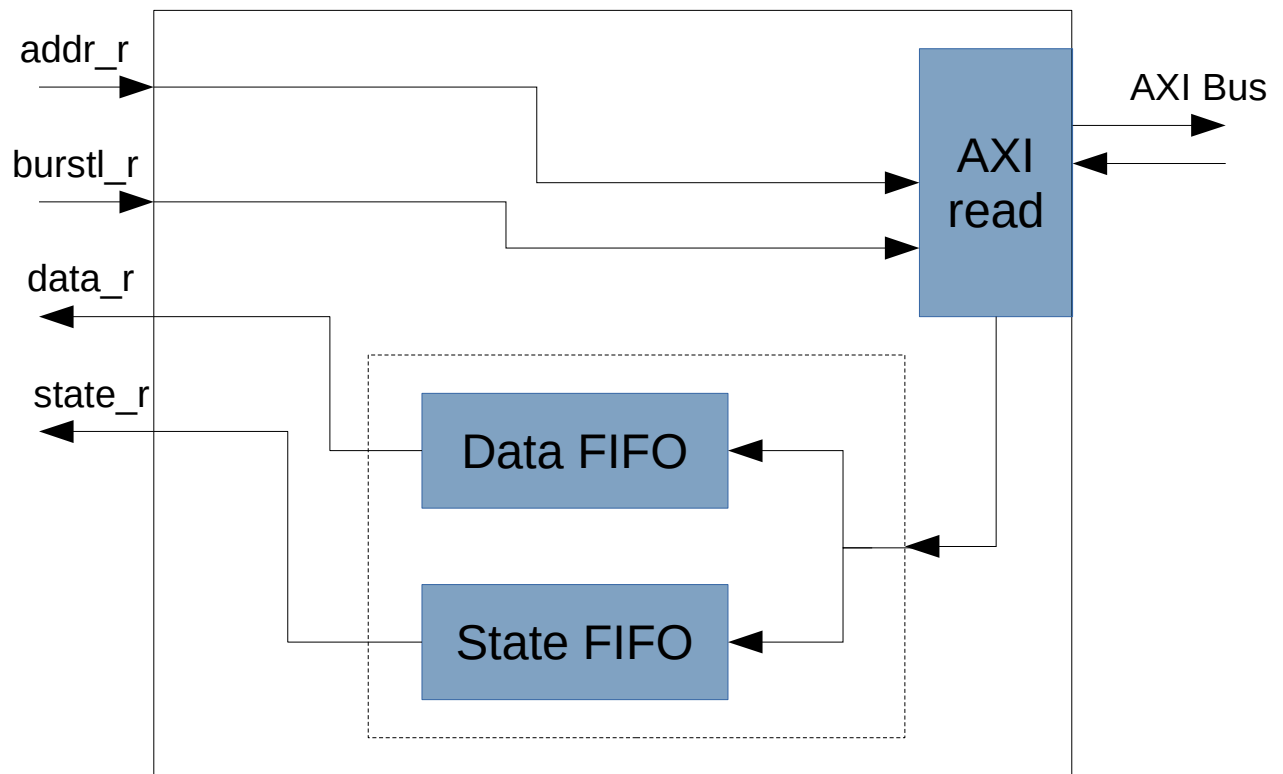


Abbildung 3.3: Schema des AXI Lese Puffers

und der Lesevorgang über AXI gestartet. Für das Zwischenspeichern der ausgelesenen Daten werden 2 Fifos verwendet. Ein Fifo wird für die ausgelesenen Daten gespeichert 32 Wörter mit 32 Bit Wortbreite. In einen weiteren Fifo ebenfalls 32Bit werden die kumulierten Statusbits gespeichert. Nachdem die ersten 16bit Datenworte in den Fifo gespeichert wurden, enthält das nächste Wort die Statusbits, dieses wird in den entsprechenden Fifo gespeichert. Sobald das Fifo für die Statusbits nicht leer ist, kann auf Seiten des Framestacks mit den Auslesen begonnen werden. Bei jeden takt in dem eine 1 an readable angelegt ist wird ein Datenwort aus dem Fifo und 2 Statusbits übertragen. Nach dem 16 Wörter aus den Datenfifo gelesen wurden wird wenn vorhanden das nächste Statuswort aus den Statusfifo gelesen.

---

## 3.2 Spill and Fill Windows

---

Bei den Spill and Fill Prozess werden Teile des aktuellen Threads und evtl die anderer Threads in sogenanntes Windows vorgehalten. Die Windows werden als Blockrams realisiert. Die Anzahl und Grösse dieser Windows kann über Parameter eingestellt werden. Für jedes Window werden die Framestack Adressen der oberen und unteren Grenze, des Adressbereichs gespeichert, der in dem Window vorgehalten wird. Dazu wird jeweils die Basis Adresse im Window gespeichert, die angibt an welcher Window Adresse die unterste Framestackadresse liegt. Das Window selber wird als Ringspeicher organisiert. Für jedes Window wird in einen Register gespeichert welcher Thread in diesen momentan vorgehalten wird. Die Adressen werden für das Window wie folgt umgerechnet:

---

### 3.3 Framestackteile des aktiven Threads auslagern (Spill)

---

Stackteile des aktiven Threads auf den externen Speicher auslagern ist eines der zwei Hauptbestandteile des Spill and Fill Mechanismus. Nämlich Spill. Bei dieser Implementierung wird der Spill Mechanismus ausgelöst, wenn beim Funktionsaufruf festgestellt wird, dass der Stack das aktuelle Window überschreiten würde.

---

#### 3.3.1 Überprüfen des freien Speichers im Window

---

Bei jeden Funktionsaufruf wird der Platzbedarf im Stack abgeschätzt und geprüft ob noch Platz dafür im aktuellen Window ist.

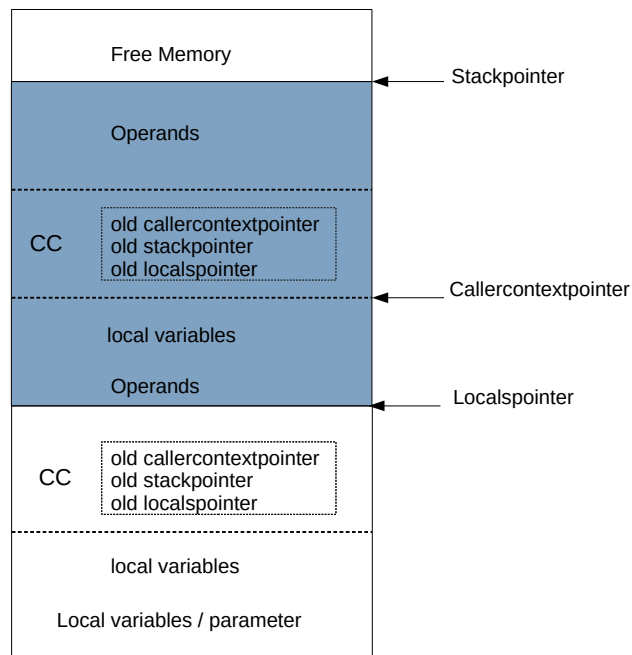
Ein Stackframe in AMIDAR beginnt mit den Argumenten und lokalen Variablen, darüber kommt der Callercontext bestehend aus: Localspointer, Callercontextpointer und Stackpointer. Zum bestimmen ob der im Window vorhandene Platz noch ausreicht wird vom aktuellen Stackpointer die Anzahl der Parameter der aufgerufenen Funktion abgezogen und Anzahl lokaler Variablen die größe des Callercontexts darauf addiert und eine Konstante für die Anzahl möglicher Parameter drauf addiert. Wenn das Ergebnis noch im aktuellen Adressbereichs des Windows liegt, wird ein normaler Invoke durchgeführt, andernfalls wird ein Spill durchgeführt.

---

#### 3.3.2 Spill

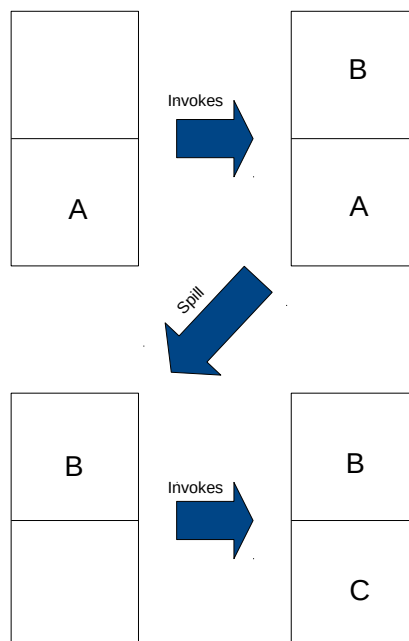
---

Zu Beginn des Spill Vorgangs wird die Burstlänge für die AXI Übertragung auf 128 gesetzt und die Bottom Adresse des unteren Windows wird in ein Register zwischengespeichert, das als Laufvariable für den folgenden Prozess verwendet wird und in ein Register kopiert die Basis des aktuellen 16er Block angibt. Anschließend wird darauf gewartet, dass das AXI Modul geschrieben werden kann, ist dies der Fall wird der Speicher im Window mit den Wert im Address Register adressiert. Nachdem die Daten aus dem Blockram gelesen wurden wird ein in das oben beschriebene AXI Modul geschrieben. Dafür wird die Threadadresse zur Framestackadresse umgerechnet. Gleichzeitig wird der Wert im Register mit der Laufvariable um eins erhöht. In den folgenden Takten wird dieses Vorgang wiederholt, bis die Laufvariable um 15 über der Basisadresse des aktuellen Blocks liegt. Ist dies der Fall wird geprüft, ob das Ende des zu übertragenen Window Parts erreicht ist. Ansonsten wird der Vorgang mit geänderter Adresse neu gestartet. Am Ende des Spill Vorgangs werden die Register für die bottom and top Adressen aktualisiert. Im darauf folgenden Takt wird die Ausführung der Invoke Instruktion neu gestartet.



after Invoke.pdf

**Abbildung 3.4:** Stackframe nach einen Funktionsaufruf



**Abbildung 3.5:** Spill: Verschieben des Windows

---

### 3.4 Ausgelagerte Framestackteile des aktiven Threads wiederherstellen (Fill)

---

Wenn der Stack durch Rücksprünge schrumpft und nahe dem unteren Ende des vorgehaltenen Adressbereichs kommt muss der ausgelagerte Framestack wiederhergestellt werden.

---

#### 3.4.1 Ablauf eines return Vorgangs mit fill

---

Im ersten Takt wird der aktuelle callercontext pointer im Register `öld_callercontext_pointer` zwischengespeichert und es wird der Callercontext des wiederherzustellenden Stackframes ausgelesen beginnend mit den `localspointer`. Anschließend wird der ausgelesene `localspointer` gespeichert und der Stackpointer zum Auslesen adressiert. Der `localspointer` stellt das untere Ende des neuen Stackframes da. Nach dem der Stackpointer wiederhergestellt wurde wird überprüft ob, der `localspointer` kleiner, als die unterste Adresse im Window ist, ist dies der Fall muss der Fill Vorgang gestartet werden. Ansonsten geht der return Vorgang normal weiter mit der Wiederherstellung des Callercontextpointers, der Rückgabe des neuen Programmcounters über den AMIDAR Bus und dem aktualisieren der `"Top_of_Stack"` und `"Next_of_Stack"` Register.

---

#### 3.4.2 Ablauf des eigentlichen Fill Vorgangs

---

Bevor das kopieren der Stackdaten passieren kann werden müssen die `-Bottom -Top` und `Baseadressregister` des aktuellen Windows angepasst werden. Im nächsten Takt muss der vorher ausgelesene Programmcounter in einen Register zwischengespeichert werden, damit dieser später zurück gegeben werden kann. Anschließend wird die neue untere Adresse des Windows umgerechnet und der Lesevorgang im AXI Modul gestartet. In dem nächsten Takt wird gewartet bis der Lesebuffer gefüllt ist. Wenn dies der Fall ist wird mit jeden weiteren Takt ein Datenwort und die dazu gehörigen Status Bits ausgelesen und von unten beginnend in das Window gespeichert. Nach jeden gespeicherten 16 Bit block wird überprüft ob ein ganzes Segment übertragen wurde oder der ganze Burst ausgelesen wurde. Wenn ersteres der Fall ist wird der Vorgang beendet. Falls der Burst komplett übertragen wurde das aktuelle Segment jedoch nicht vollständig übertragen wurde wird ein neuer Leseburst mit der aktuellen Adresse gestartet.

---

### 3.5 Multithreading mit Verdrängung

---

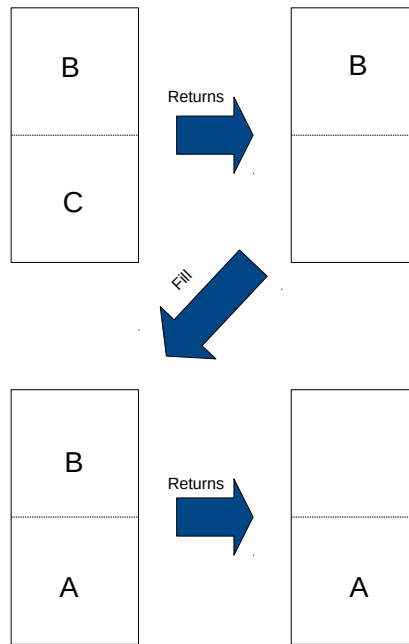
Es kann eine konstante Anzahl an Threads in den Windows vorgehalten werden. Um die Multithreading Funktionalität mit Threads weiterhin zu gewährleisten wurde eine Threadverdrängung implementiert durch die Stacks von in den Hauptspeicher ausgelagert und wiederhergestellt werden können.

---

#### 3.5.1 Zuordnung von Threads zu den Windows

---

Es gibt für jedes Window ein Register in dem die zugeordnete ThreadID gespeichert wird. Dazu kommt 1 Status Bit, mit dem angegeben wird, ob dem Window bereits eine ThreadID zugeordnet ist und 4 Bits, die für einen Least Recently Used counter genutzt werden. Nach einen Reset des Framestack Moduls werden alle Bits der Windows 1-n bis auf das erste Statusbit auf 0 gesetzt. Für Window 0, in dem der Stack des Thread 0 liegt, wird das erste Bit auf 0 gesetzt, damit es ohne explizite Threaderzeugung vom Bootloader genutzt werden kann.



**Abbildung 3.6:** fill: Verschieben des Windows

---

### 3.5.2 Änderungen am der Threaderzeugung

---

Bisher wurden Threads im Framestack angelegt in dem über das Wishbone Interface erst der Threadtable initialisiert wird und anschließend der Stackframe durch Schreiben der Rücksprungadresse und eines leeren Callercontext in den Blockram. Das Problem dabei ist durch die Auslagerung der Threads in den Framestack die Daten in den Hauptspeicher geschrieben werden müsste, was aufwendig zu implementieren wäre und sehr viel Zeit bei der Ausführung benötigen würde. Es ist außerdem nicht möglich neue Threads direkt im Window zu initialisieren, da es möglich ist das mehr Threads initialisiert werden, als Windows verfügbar sind, bevor die Ausführung der Threads zum ersten mal gestartet wird. Stattdessen wurde für die Threaderzeugung der bisher ungenutzte Framestack Opcode `ÖP_FRAMESTACK_NEWTHREAD` genutzt um den Stackframe der Threads zu initialisieren. Der Threadtable wird weiterhin über Wishbone initialisiert. Sobald der `NEWTHREAD` Opcode abgearbeitet wird geprüft ob schon ein Thread mit der zu über Wishbone übergebenen ID in einen der Windows vorhanden ist. Ist dies der Fall wird der als Parameter mit den Opcode übergebene Threadhandle an Adresse 0 des entsprechenden Windows geschrieben. Darüber wird der Callercontext initialisiert. Wenn die ThreadID noch keinen der Windows zugewiesen ist, muss der Stackframe des Threads im externen Speicher initialisiert werden. Dafür wird auch das vorher beschriebene AXI Modul verwendet. Es wird der Threadhandle und der leere Callercontext übertragen. Außerdem müssen noch 12 weitere leere Wörter übertragen werden um einen 16 Wort Block übertragen zu können.

---

### 3.5.3 Threadwechsel

---

Für einen Threadwechsel bei einen AMIDAR Prozessor werden im Framestack die Werte der Register Stackpointer, Localspointer und Callercontext aus dem Threadtable übernommen. In der neuen Implementierung wird zusätzlich geprüft ob, der Thread auf den gewechselt wird bereits in einen der Windows

---

vorgehalten wird. Ist dies der Fall kann einfach auf das entsprechende Window gewechselt werden um den Thread weiter auszuführen. Wenn der Stack des Threads noch nicht in einen der Windows sein sollte wird dieser aus dem Speicher geladen. Davor wird gegebenenfalls noch der Stack eines zu verdrängender Threads in den externen Speicher vorschoben.

---

### Ablauf des Threadwechsels

---

Der Threadwechsel beginnt, wenn der Opcode `ÖP_FRAMESTACK_THREADSWITCH` abgearbeitet wird. Die ThreadID des neuen Threads wird über den AMIDAR-Bus übergeben und in einem Register gespeichert. Zu Beginn des Vorgangs wird geprüft ob die ThreadID schon im `WINDOW_TO_THREAD` Register eingetragen ist. Ist dies nicht der Fall wird geprüft ob es noch ein Window leer ist. Wenn kein Window verfügbar ist werden die LRU Bits überprüft und der am wenigsten verwendete Thread bestimmt. Wenn auf einen schon vorgehaltenen Thread gewechselt wird, wird der LFU Counter des Threads um eins erhöht. Dabei wird überprüft ob der counter des aktuellen nächsten Threads gleich 15 ist. Ist dies der Fall wird der counter aller Threads halbiert. In dem Fall, in dem ein Thread verdrängt werden muss wird der am wenigsten verwendete Thread in den externen Speicher transferiert. Dafür wird die passende Burstlänge anhand des Stackpointers und der unteren Windowadresse berechnet. Sie muss ein Vielfaches von 16 sein. Danach findet die eigentliche Übertragung des Stacks statt. Nachdem der Stackframe des zu verdrängenden Threads gesichert wurde kann das Window mit den Stack des neuen Threads überschrieben werden. Dafür wird die untere Startadresse der Übertragung anhand des Localspointer und der größe der Windowabschnitte sodass der Locals und der Stackpointer im Window liegen. Daraufhin wird der zu übertragende Stackframe ausgelesen und in das Window kopiert.

---

### 3.6 Garbage Collector Interface

---

Der Framestack muss den Garbage Collector ein Interface zur Verfügung stellen über das, das Rootset ausgelesen werden kann. Dafür werden für jedes Datenwort im Framestack zwei Statusbits angegeben, die den Entrytype beschreiben. Der Entrytype 2'b10 gibt dabei Referenzen an. Um das Rootset zu erzeugen muss der Stack jedes einzelnen Threads ausgelesen werden, dabei wird bei jedem Wort geprüft ob der EntryType 2'b10 hinterlegt wurde, in dem Fall wird das Wort an den Garbage Collector übertragen. Da Teile des Framestacks im externen Speicher liegen müssen diese ausgelesen werden. Wenn der Eingang `gc_request_rootset` gesetzt ist beginnt der Auslese Vorgang. Der Stack wird Thread für Thread ausgelesen. Nachdem ein Thread zum Auslesen ausgewählt wurde wird geprüft, ob dieser im Window liegt. Ist das der Fall wird geprüft, ob der komplette Stack im Window liegt. Wenn Teile des Stacks ausgelagert wurden, werden diese in 16er Blöcken ausgelesen und geprüft. Sobald die untere Grenze des Windows dabei erreicht ist, kann der Rest aus dem Window gelesen werden. Sobald alle Threads nach diesen Verfahren abgearbeitet wurden ist das Rootset komplett übertragen.

---

### 3.7 Lesezugriff Wishbone

---

Lesezugriff über Wishbone auf dem Framestack musste auch umgeschrieben werden. Bisher liefen die Zugriffe direkt über den 2. Port des Blockram in dem der Framestack gespeichert wurde. Wegen des ausgelagerten Threads funktioniert das nur wenn die auszulesenden Daten im Window liegen. Wenn das nicht der Fall ist, müssen diese jeweils aus dem RAM geladen werden, was von der FramestackFSM abgearbeitet wird. Währenddessen muss allerdings das Wishbone Acknowledge Signal zurück gehalten werden, wodurch der Wishbone-Bus eine Weile blockiert wird. Da dieser Lesezugriff nur für den Debugger genutzt wird, ist die Performance bei diesen Vorgang jedoch vernachlässigbar.

---

## 4 Evaluation

Um die korrekte Funktionsweise und die Performance des geänderten Framestacks zu überprüfen wurde eine modifizierte AMIDAR Testbench verwendet. Vorallem im Fokus stand die korrekte Ausführung der SSpill and Fill Abläufe und die Threadwechsel.

---

### 4.1 Testumgebung

---

Für die Tests wurde die AMIDAR Testbench modifiziert um Funktionen mit einen größeren Framestack oder Threadwechseln zu testen. Ausserdem wurde die Hardware des Framestacks erweitert um die benötigte Anzahl an Takten für einen bestimmten Vorgang aufzuzeichnen.

---

#### 4.1.1 Hardwareänderungen zur Performance Evaluation

---

Um die Performance evaluieren zu können ein zusätzlicher Blockram angelegt in dem jeweils die Dauer eines bestimmten Vorgangs protokolliert wird. Es kann entweder die Dauer der Spill, der Fill oder der Threadswitch Vorgänge protokolliert werden. Um auf diese Daten zuzugreifen wurden dem Framestack die Wishbone Register 15, 16, 17 und 18 hinzugefügt. Mit den Register 18 wird dabei angegeben ob und welcher Vorgang protokolliert werden soll. Wird eine 0 in das Register geschrieben wird die Protokollfunktion deaktiviert. Mit Register 15 Wird angegeben aus welcher Adresse des Speichers gelesen werden soll. Die ausgelesenen Daten selber stehen in Register 18. Das Register 17 gibt die Anzahl der Einträge in dem Speicher an. Zur Protokollierung eines Vorgangs werden die Takte gezählt sobald der Startzustand eines Vorgangs erreicht ist und endet wenn der Vorgang beendet wurde und neue Tokens abgearbeitet werden können. Zu beachten dabei ist, das bei jeden Vorgang nur die Anzahl der Takte gemessen wird bei denen die Framestack FSM beschäftigt ist. Beim Spill zum Beispiel wird das zählen beendet sobald die letzten Stackdaten in den Schreibfifo geschrieben wurde. Die Zeit, in der der AXI Bus belegt ist, ist etwas länger, da der FIFO noch abgearbeitet werden muss. Der Framestack selber kann währenddessen schon weitere Token abarbeiten. Alle eben beschriebenen Änderungen lassen sich wenn nicht benötigt durch eine Präprozessordefinition deaktivieren.

---

#### 4.1.2 angepasste Testbench

---

Um die weiterhin korrekte Funktionsweise der Grundlegenden Framestack Funktionalitäten zu testen wurde die AMIDAR Testbench genutzt. Erweitert wurde diese dabei zum einen um Tests, die den Spill and Fill Vorgang testen. Dafür wurde eine Funktion benötigt, die einen möglichst grossen Stack erzeugt und dennoch eine einigermaßen kurze Laufzeit hat. Die Funktion zur Rekursiven Berechnung der Fibonacci Folge erzeugt zwar schnell einen sehr großen Stack, allerdings ist die Laufzeit dieser deutlich zu lang um als Test in Frage zu kommen. Zum Testen der Spill and Fill Funktionen wurde eine vereinfachte Ackermann Funktion genutzt, die mit 2 Parametern arbeitet. In dem neuen Testfall AckermannTest wird diese Funktion mit unterschiedlichen Parametern aufgerufen:



---

**Tabelle 4.1: Ackermanntests**

Ackermann (3,3):	Rekursionstiefe = 63 Maximale Stackgrösse = 450 Wörter Rückgabewert = 9
Ackermann (3,4)	Rekursionstiefe = 127 Maximale Stackgrösse = 890 Wörter Rückgabewert = 61
Ackermann (3,5):	Rekursionstiefe = 255 Maximale Stackgrösse = 1785 Wörter Rückgabewert = 253

Weitere wichtige Testfälle waren die schon vorhandenen, jedoch nicht einkommentierten Multithreading Tests. Genutzt wurde der "BasicThreadTest" und der "AdvancedThreadTest". Ausserdem wurde eine modifizierte Variante des BasicThreadTests genutzt, bei der der Ackermanntest in mehreren Threads ausgeführt wurde.

Eine weitere Anpassung der Testbench ist eine Methode, die über Wishbone die im Framestackmodul gemessenen Performancedaten ausliest. Zuerst wird die Anzahl der erfassten Vorgänge ausgelesen. Anschließend wird über den Speicher mit den Messdaten iteriert und der Inhalt ausgegeben, dabei wird das arithmetische Mittel der Messwerte berechnet und ebenfalls ausgegeben.

---

## 4.2 Performance Messungen

---

Bei den Performance Messungen wurde die Dauer bestimmter Vorgänge gemessen, in denen die oben beschriebenen Tests ausgeführt und die Messdaten ausgelesen worden sind. Die Tests wurden mit unterschiedlichen Konfigurationen der Spill and Fill Windows des Framestacks durchgeführt. Es wurde sowohl mit einer Window Grösse von 512 Wörtern als auch mit 1024 Wörtern getestet. 1024 Wörter entspricht dabei der Anzahl der Wörter, die im ursprünglichen Framestack einem Thread jeweils zur Verfügung standen. Das Window wird dabei in 2 Segmente geteilt, die jeweils 256 oder 512 Wörter gross sind.

---

### 4.2.1 Spill and Fill

---

Bei der Variante mit 512 Wörtern werden für einen Spill Vorgang 288 Takte benötigt. Der Wert ist konstant, da Verzögerungen der AXI Verbindung durch den Schreibpuffer ausgeglichen werden. Ein Fill Vorgang dauerte in dem Fall mindestens 363 und maximal 370, im Durchschnitt 363 Takte. Beim Durchlauf der Ackermanntests findet der Spill und der Fill Vorgang jeweils 462 mal statt. Ein Fill benötigt mehr Takte, da zum einen das Adressieren des AXI Busses für den Auslese Vorgang Zeit benötigt. Außerdem müssen nach dem Ende der AXI Übertragung noch mindestens 16 Wörter aus dem Lesebuffer gelesen werden. Wenn ein Window 1024 Wörter umfasst, werden für einen Spill 576 Takte gemessen. Für den Fill werden dabei zwischen 725 und 744 Takte gemessen. Im Durchschnitt werden 726 Takte benötigt. Es finden beim Durchlauf der Testbench jeweils 131 Spill and Fill Vorgänge statt.

---

## Interpretation

---

Bei der Variante mit 512 Wörtern werden insgesamt 300762 Takte für Spill and Fill Operationen genutzt. Mit einem 1024 Wörter großen Window werden lediglich 162702 Takte durch Spill and Fill blockiert. Bei der Variante mit einem 1024 Wörter großen Window ist in diesen Testfällen die Anzahl der Takte, die für Spill and Fill benötigt werden, beinahe halbiert, was ein größeres Performance-Vorteil andeutet. Berücksichtigt man jedoch, dass die Gesamtlaufzeit der Testfälle 17943867 bzw. 17805807 Takte beträgt, ist Spill and Fill nur für 0,91376% bzw. 0,1676% der Laufzeit der Tests verantwortlich. Gemessen wurde dabei wirklich nur die Laufzeit der Testfälle ohne Bootloader, static Initializer und UART-Ausgaben. Bei realen Anwendungen dürften die Prozentsätze deutlich niedriger liegen als bei der Ackermannfunktion.

---

### 4.2.2 Threadswitch

---

Die Dauer eines Threadwechsels ist stark schwankend, je nachdem, wie groß der Stack des nächsten Stacks ist. Wird auf einen Thread gewechselt, der bereits in einem der Windows liegt, werden lediglich 5 Takte für den Wechsel benötigt. Nach Ausführung der Multithreading-Tests werden 43 Threadwechsel gemessen. Diese benötigten zwischen 5 und 245 Takten. Im Durchschnitt werden 99 Takte. Mehrere Windows bringen bei den Threadtests der AMIDAR-Testbench jedoch keinen größeren Vorteil, da alle 10 Threads nacheinander aufgerufen werden. Bei Anwendungen, in denen wenige Threads oft, viele andere dagegen selten geweckt werden, treten Threadverdrängungen seltener auf, was zu deutlich schnelleren Threadwechseln führen dürfte.

---

## 5 Fazit

Im Zusammenhang dieses Projektseminars wurde die AMIDAR Framestack FU um Spill and Fill erweitert. Was deutlich größere Stackframes ermöglicht bei gleichzeitiger Reduzierung des benötigten Blockrams. Der Gesamteinfluss auf die Performance des Prozessors dabei ist minimal.

Um dies zu ermöglichen wurde eine Anbindung des Framestacks über AXI eingebaut. Dafür wurde ein Modul gebaut was neben der AXI Ansteuerung auch das konvertieren der Unterschiedlichen Adressbereiche von Framestack zu DDR RAM übernimmt.

---

## 6 Literaturverzeichnis

- [1] Tao Guo. Ansteuerung eines SDRAM auf einem Xilinx FPGA-Board. Master's thesis, TU Darmstadt, Fachgebiet Rechnersysteme, 2008.