

Betriebssysteme

7. Tutorium - Paging

Peter Bohner

12. Dezember 2023

ITEC - Operating Systems Group

- Rechnungen haben noch Übungsbedarf

- Rechnungen haben noch Übungsbedarf
- PML5, PML4, PDP, PDT, PT → Das ist x86-spezifischer Schrott. Sagt dazu 1., 2., ... level page table

- Rechnungen haben noch Übungsbedarf
- PML5, PML4, PDP, PDT, PT → Das ist x86-spezifischer Schrott. Sagt dazu 1., 2., ... level page table
- P1 (strings): string-split mit einer allocation. Frage: Wie?

- Rechnungen haben noch Übungsbedarf
- PML5, PML4, PDP, PDT, PT → Das ist x86-spezifischer Schrott. Sagt dazu 1., 2., ... level page table
- P1 (strings): string-split mit einer allocation. Frage: Wie?
- P2 (mmu): Validität von Einträgen prüfen. Ein globaler access counter

- Rechnungen haben noch Übungsbedarf
- PML5, PML4, PDP, PDT, PT → Das ist x86-spezifischer Schrott. Sagt dazu 1., 2., ... level page table
- P1 (strings): string-split mit einer allocation. Frage: Wie?
- P2 (mmu): Validität von Einträgen prüfen. Ein globaler access counter

Allgemein

- Ich habe die slides hochgeladen, erinnert mich, wenn etwas fehlt.

- Rechnungen haben noch Übungsbedarf
- PML5, PML4, PDP, PDT, PT → Das ist x86-spezifischer Schrott. Sagt dazu 1., 2., ... level page table
- P1 (strings): string-split mit einer allocation. Frage: Wie?
- P2 (mmu): Validität von Einträgen prüfen. Ein globaler access counter

Allgemein

- Ich habe die slides hochgeladen, erinnert mich, wenn etwas fehlt.
- Heute: auch etwas Wiederholung vom letzten mal

- Rechnungen haben noch Übungsbedarf
- PML5, PML4, PDP, PDT, PT → Das ist x86-spezifischer Schrott. Sagt dazu 1., 2., ... level page table
- P1 (strings): string-split mit einer allocation. Frage: Wie?
- P2 (mmu): Validität von Einträgen prüfen. Ein globaler access counter

Allgemein

- Ich habe die slides hochgeladen, erinnert mich, wenn etwas fehlt.
- Heute: auch etwas Wiederholung vom letzten mal
- Wer ist nächste Woche da?

d

Betrachten Sie ein System, das mittels einer hierarchischen Seitentabelle virtuelle in physische Speicheradressen übersetzt. Der virtuelle Adressraum umfasst 512 GiB, die Seitengröße ist 4 KiB, eine Seitentabelle beinhaltet 512 Einträge. Berechnen Sie die Anzahl der Stufen der Seitentabellenhierarchie.

e

Ein Programm kopiert einen 4 MiB Puffer im virtuellen Adressraum. Quell- und Zielpuffer überlappen sich nicht und sind an Seitengrenzen ausgerichtet. Die CPU verfügt über keinen Cache, jedoch einen leeren TLB. Die Wortbreite beträgt 8 Bytes. Es werden hierarchische Page Tables mit 3 Stufen verwendet. Wie viele Speicherzugriffe sind für den Kopiervorgang mindestens nötig?

Page Fault Handling

What is Demand-Paging and Pre-Paging?

Demand-Paging:

What is Demand-Paging and Pre-Paging?

Demand-Paging:

- Load pages on-demand, right when they are needed

What is Demand-Paging and Pre-Paging?

Demand-Paging:

- Load pages on-demand, right when they are needed

Pre-Paging:

- Loaded Pages speculatively in batches, even *before* you need them

What is Demand-Paging and Pre-Paging?

Demand-Paging:

- Load pages on-demand, right when they are needed

Pre-Paging:

- Loaded Pages speculatively in batches, even *before* you need them

Why would you (not?) use Demand-Paging?

What is Demand-Paging and Pre-Paging?

Demand-Paging:

- Load pages on-demand, right when they are needed

Pre-Paging:

- Loaded Pages speculatively in batches, even *before* you need them

Why would you (not?) use Demand-Paging?

- + Only loads needed data \Rightarrow Less memory wasted

What is Demand-Paging and Pre-Paging?

Demand-Paging:

- Load pages on-demand, right when they are needed

Pre-Paging:

- Loaded Pages speculatively in batches, even *before* you need them

Why would you (not?) use Demand-Paging?

- + Only loads needed data \Rightarrow Less memory wasted
- Generates lots of page faults before working set is in memory

Why would you (not?) use Pre-Paging?

Why would you (not?) use Pre-Paging?

- + Might reduce number of page faults

Why would you (not?) use Pre-Paging?

- + Might reduce number of page faults
- Loads more than needed \Rightarrow Wasteful

Why would you (not?) use Pre-Paging?

- + Might reduce number of page faults
- Loads more than needed \Rightarrow Wasteful
- More I/O \Rightarrow Slower?

Why would you (not?) use Pre-Paging?

- + Might reduce number of page faults
- Loads more than needed \Rightarrow Wasteful
- More I/O \Rightarrow Slower?
- + HDDs a lot faster when reading chunks

Different kind of page faults

Not all pages are created equal. Do you have any idea what types of page faults typically exist?

On-Demand Paging

Different kind of page faults

Not all pages are created equal. Do you have any idea what types of page faults typically exist?



Kernel Space

The diagram consists of a solid purple rectangle labeled 'Kernel Space'. Below it is a dashed-line rectangle labeled 'No access'.

No access

On-Demand Paging

Different kind of page faults

Not all pages are created equal. Do you have any idea what types of page faults typically exist?

Kernel Space
No access
Buffer
No access
Stack

On-Demand Paging

Different kind of page faults

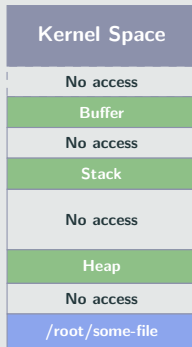
Not all pages are created equal. Do you have any idea what types of page faults typically exist?

Kernel Space
No access
Buffer
No access
Stack
No access
Heap

On-Demand Paging

Different kind of page faults

Not all pages are created equal. Do you have any idea what types of page faults typically exist?



On-Demand Paging

Different kind of page faults

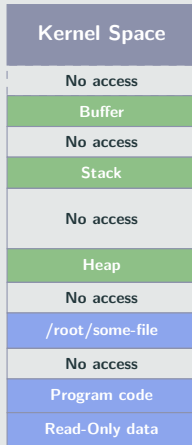
Not all pages are created equal. Do you have any idea what types of page faults typically exist?

Kernel Space
No access
Buffer
No access
Stack
No access
Heap
No access
/root/some-file
No access

On-Demand Paging

Different kind of page faults

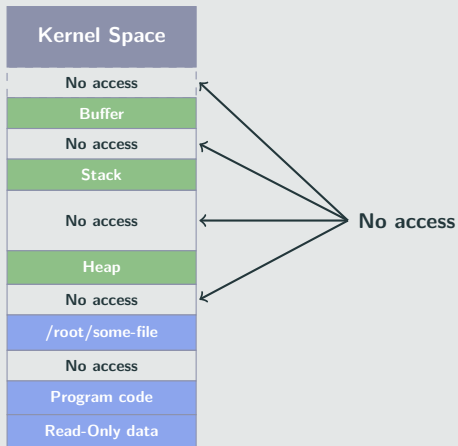
Not all pages are created equal. Do you have any idea what types of page faults typically exist?



On-Demand Paging

Different kind of page faults

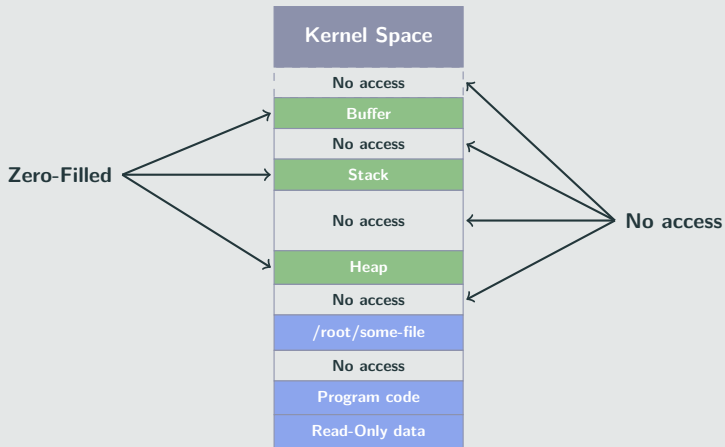
Not all pages are created equal. Do you have any idea what types of page faults typically exist?



On-Demand Paging

Different kind of page faults

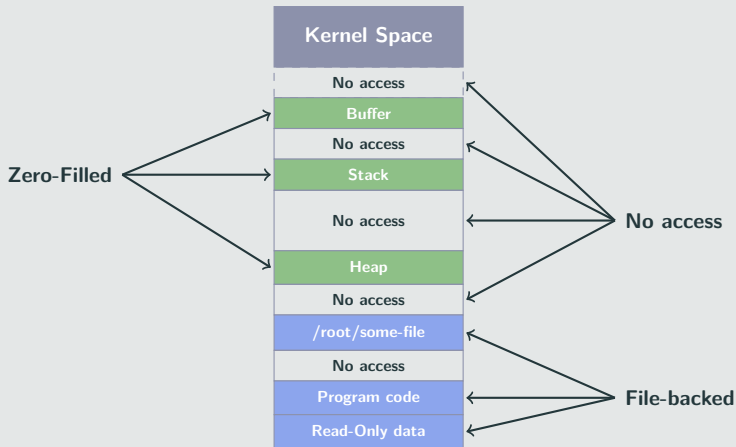
Not all pages are created equal. Do you have any idea what types of page faults typically exist?



On-Demand Paging

Different kind of page faults

Not all pages are created equal. Do you have any idea what types of page faults typically exist?



Different kind of page faults

Why are pages to generic memory zero filled?

Different kind of page faults

Why are pages to generic memory zero filled? It could leak other processes' memory otherwise (called a [Covert Channel](#)).

Different kind of page faults

Why are pages to generic memory zero filled? It could leak other processes' memory otherwise (called a [Covert Channel](#)).

And there is one other kind of page fault...

Different kind of page faults

Why are pages to generic memory zero filled? It could leak other processes' memory otherwise (called a [Covert Channel](#)).

And there is one other kind of page fault...

...The page was stolen by the OS and swapped out!

Different kind of page faults

Why are pages to generic memory zero filled? It could leak other processes' memory otherwise (called a [Covert Channel](#)).

And there is one other kind of page fault...

...The page was stolen by the OS and swapped out!

Also supported on some systems: *Purgable memory*. Stolen from [Apple](#) and also implemented in [SerenityOS](#) in [this video](#).

What kind of information does the page fault handler need?

What kind of information does the page fault handler need?

- Access flags: Can the user perform the operation on this page?

What kind of information does the page fault handler need?

- Access flags: Can the user perform the operation on this page?
- Where to find the most recent version (different for zero filled, file backed, etc.)

How could you implement Copy-on-Write memory?

How could you implement Copy-on-Write memory?

- Mark memory as read-only on fork
- Add an additional CoW flag: When a page fault is raised check it, copy the page and clear the CoW and ro flag

Page replacement

The pager in some systems tries to keep „spare pages“. Why?

The pager in some systems tries to keep „spare pages“. Why?

- OS needs free (pre-zeroed) frames to assign to processes

The pager in some systems tries to keep „spare pages“. Why?

- OS needs free (pre-zeroed) frames to assign to processes
- What happens when there are none and a page fault occurs?

The pager in some systems tries to keep „spare pages“. Why?

- OS needs free (pre-zeroed) frames to assign to processes
- What happens when there are none and a page fault occurs?

⇒ Needs to write dirty pages back to disk

⇒ Slow

If you need to swap out a page, what pages do you search for a victim?

If you need to swap out a page, what pages do you search for a victim?

Local page replacement algorithms:

- Only look through the frames *of that process*

Global page replacement algorithms:

- Look through *all* frames, even that of other processes

What are the pro and cons of local algorithms?

What are the pro and cons of local algorithms?

- + Guaranteed number of pages per application

What are the pro and cons of local algorithms?

- + Guaranteed number of pages per application
- + Potentially faster

What are the pro and cons of local algorithms?

- + Guaranteed number of pages per application
- + Potentially faster
- Guaranteed number of pages per application \Rightarrow Can not adapt as easily to changing demands

What are the pro and cons of local algorithms?

- + Guaranteed number of pages per application
- + Potentially faster
- Guaranteed number of pages per application \Rightarrow Can not adapt as easily to changing demands
- Difficult selection of optimal number of frames per application (see point above)

What are the pro and cons of local algorithms?

- + Guaranteed number of pages per application
- + Potentially faster
- Guaranteed number of pages per application \Rightarrow Can not adapt as easily to changing demands
- Difficult selection of optimal number of frames per application (see point above)

Every process should get an equal number of pages. Do you use a Global or Local Algorithm?

What are the pro and cons of local algorithms?

- + Guaranteed number of pages per application
- + Potentially faster
- Guaranteed number of pages per application \Rightarrow Can not adapt as easily to changing demands
- Difficult selection of optimal number of frames per application (see point above)

Every process should get an equal number of pages. Do you use a Global or Local Algorithm?

Local

What is the working set?

- The set of pages that a process accessed in the last Δ page references

What is the working set?

- The set of pages that a process accessed in the last Δ page references

What is Thrashing?



What is the working set?

- The set of pages that a process accessed in the last Δ page references

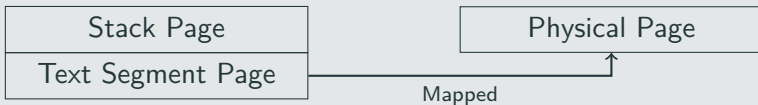
What is Thrashing?



What is the working set?

- The set of pages that a process accessed in the last Δ page references

What is Thrashing?



What is the working set?

- The set of pages that a process accessed in the last Δ page references

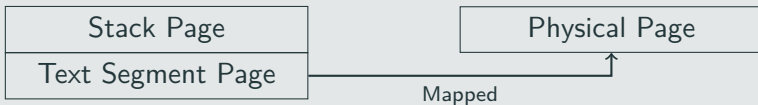
What is Thrashing?



What is the working set?

- The set of pages that a process accessed in the last Δ page references

What is Thrashing?



What is the working set?

- The set of pages that a process accessed in the last Δ page references

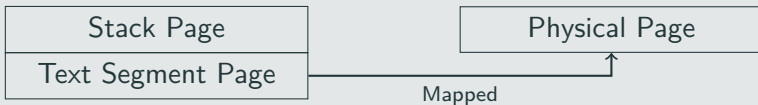
What is Thrashing?



What is the working set?

- The set of pages that a process accessed in the last Δ page references

What is Thrashing?



What is the working set?

- The set of pages that a process accessed in the last Δ page references

What is Thrashing?



What is the working set?

- The set of pages that a process accessed in the last Δ page references

What is Thrashing?



- Not enough frames to fit the working set

⇒ Pages will be stored to disk and reloaded very often

What are those?

We need to find a victim page to replace with a new one.

What are those?

We need to find a victim page to replace with a new one.

Common strategies

- FIFO:

What are those?

We need to find a victim page to replace with a new one.

Common strategies

- **FIFO:** First page to be loaded is the first to be unloaded
- **LRU:**

What are those?

We need to find a victim page to replace with a new one.

Common strategies

- **FIFO:** First page to be loaded is the first to be unloaded
- **LRU:** Least recently used page is evicted
- **Optimal:**

What are those?

We need to find a victim page to replace with a new one.

Common strategies

- **FIFO:** First page to be loaded is the first to be unloaded
- **LRU:** Least recently used page is evicted
- **Optimal:** Evict the page that is used the *furthest into the future*. Feasible?

What are those?

We need to find a victim page to replace with a new one.

Common strategies

- **FIFO:** First page to be loaded is the first to be unloaded
- **LRU:** Least recently used page is evicted
- **Optimal:** Evict the page that is used the *furthest into the future*. Feasible? Used as a *benchmark*
- **Clock:**

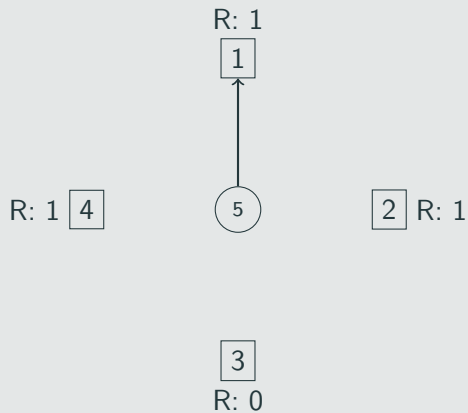
What are those?

We need to find a victim page to replace with a new one.

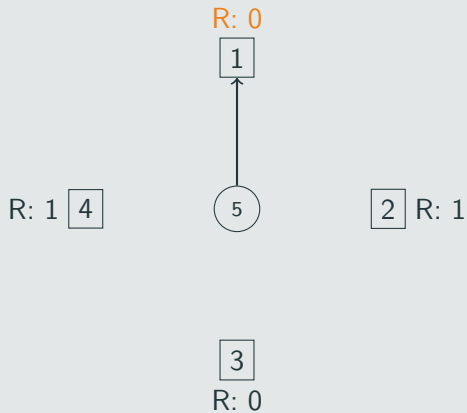
Common strategies

- **FIFO:** First page to be loaded is the first to be unloaded
- **LRU:** Least recently used page is evicted
- **Optimal:** Evict the page that is used the *furthest into the future*. Feasible? Used as a *benchmark*
- **Clock:** Uff, let's talk about that on its own page

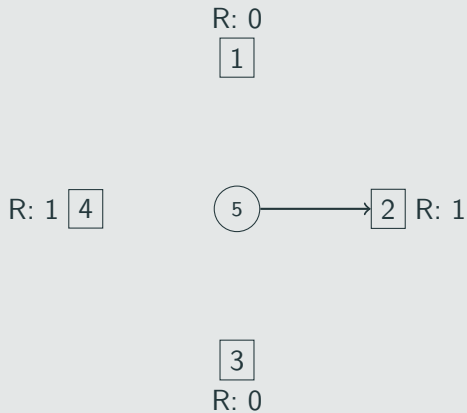
Clock



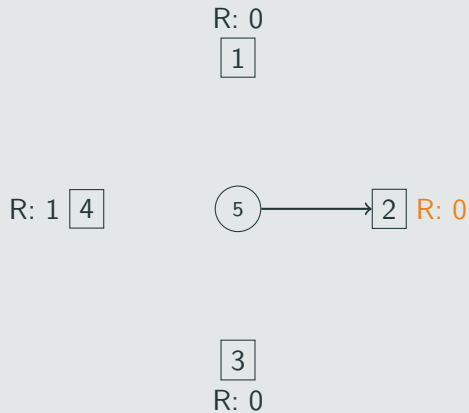
Clock



Clock

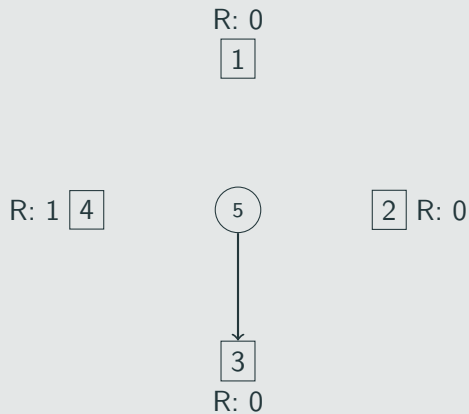


Clock



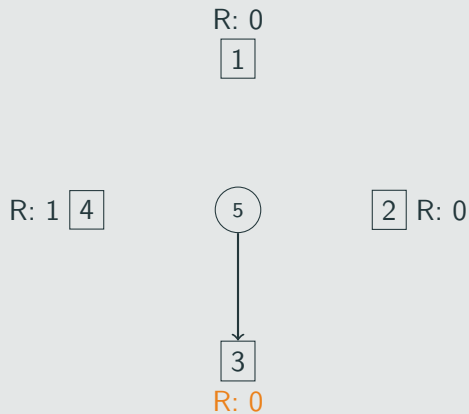
Page Replacement Policies – Clock

Clock



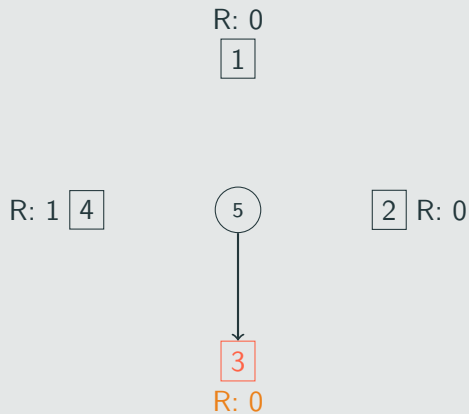
Page Replacement Policies – Clock

Clock

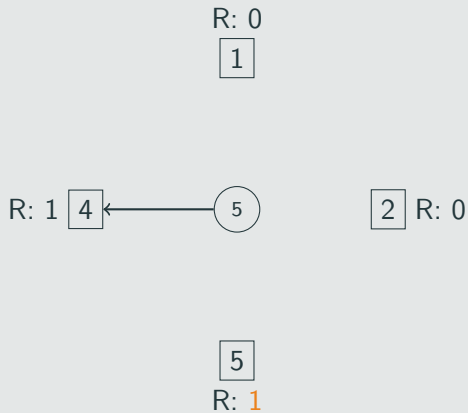


Page Replacement Policies – Clock

Clock



Clock



Where did that R come from?

Where did that R come from?

- Referenced (and modified) bits are set by the CPU when accessing or writing to the page

Where did that R come from?

- Referenced (and modified) bits are set by the CPU when accessing or writing to the page
- Referenced bits are periodically cleared by the kernel using timer interrupts

Do it

- Clock: Ordered by Load time ASC, Head is on Frame 3
- Reference order: 4, 0, 0, 0, 2, 4, 2, 1, 0, 3, 2

Frame	Virtual page	Load time	Access time	Referenced	Modified
0	2	60	161	0	1
1	1	130	160	0	0
2	0	26	162	1	0
3	3	20	163	1	1

How well does LRU work for the Stack, the Heap and the Code segment?

How well does LRU work for the Stack, the Heap and the Code segment?

- Stack: Beautifully. The older it is the longer it will take to be reached again

How well does LRU work for the Stack, the Heap and the Code segment?

- Stack: Beautifully. The older it is the longer it will take to be reached again
- Code: Mostly, loops are mostly linear and it follows certain patterns

How well does LRU work for the Stack, the Heap and the Code segment?

- Stack: Beautifully. The older it is the longer it will take to be reached again
- Code: Mostly, loops are mostly linear and it follows certain patterns
- Heap: Well, the heap has more random access patterns. So not that well, probably

Pointer Arithmetic

Are numbers!

```
1  int  hello    = 20;    // no worries
```

Are numbers!

```
1  int  hello    = 20;      // no worries
2  int* hPointer = &hello; // no worries
```

Are numbers!

```
1  int  hello      = 20;      // no worries
2  int* hPointer   = &hello;  // no worries
3  int* pointer    = 0x200;   // no worries
```

Are numbers!

```
1  int  hello      = 20;      // no worries
2  int* hPointer   = &hello;  // no worries
3  int* pointer    = 0x200;   // no worries
```

And you can convert between them

```
1      int  hello          = 20;
2      int* aPointer       = &hello;
```

Are numbers!

```
1  int  hello      = 20;      // no worries
2  int* hPointer   = &hello;  // no worries
3  int* pointer    = 0x200;   // no worries
```

And you can convert between them

```
1      int  hello      = 20;
2      int* aPointer    = &hello;
3      intptr_t asInt   = (intptr_t) aPointer;
```

Are numbers!

```
1  int  hello      = 20;      // no worries
2  int* hPointer   = &hello;  // no worries
3  int* pointer    = 0x200;   // no worries
```

And you can convert between them

```
1      int  hello          = 20;
2      int* aPointer       = &hello;
3      intptr_t asInt      = (intptr_t) aPointer;
4      int* backToPointer  = (int*) asInt;
```

And what do operations do?

```
1  int array[5] = {0, 1, 2, 3, 4};  
2  printf("Current value: %d", *array);
```


And what do operations do?

```
1  int array[5] = {0, 1, 2, 3, 4};  
2  printf("Current value: %d", *array); // 0  
3  array++;  
4  printf("Current value: %d", *array);
```

And what do operations do?

```
1  int array[5] = {0, 1, 2, 3, 4};  
2  printf("Current value: %d", *array); // 0  
3  array++;  
4  printf("Current value: %d", *array); // 1
```

Wait, weren't integers more than one byte?

- + and - on a pointer decrease it by *the size of its type*

And what do operations do?

```
1  int array[5] = {0, 1, 2, 3, 4};
2  printf("Current value: %d", *array); // 0
3  array++;
4  printf("Current value: %d", *array); // 1
```

Wait, weren't integers more than one byte?

- + and - on a pointer decrease it by *the size of its type*

```
1  int* pointer;
2  pointer++; // is the same as
```

And what do operations do?

```
1  int array[5] = {0, 1, 2, 3, 4};
2  printf("Current value: %d", *array); // 0
3  array++;
4  printf("Current value: %d", *array); // 1
```

Wait, weren't integers more than one byte?

- + and - on a pointer decrease it by *the size of its type*

```
1  int* pointer;
2  pointer++; // is the same as
3  intptr_t asInt = (intptr_t) pointer;
```

And what do operations do?

```
1  int array[5] = {0, 1, 2, 3, 4};
2  printf("Current value: %d", *array); // 0
3  array++;
4  printf("Current value: %d", *array); // 1
```

Wait, weren't integers more than one byte?

- + and - on a pointer decrease it by *the size of its type*

```
1  int* pointer;
2  pointer++; // is the same as
3  intptr_t asInt = (intptr_t) pointer;
4  asInt += sizeof(int);
5  pointer = (int*) asInt;
```

Multiplication? Division?

- Probably a compiler error

Multiplication? Division?

- Probably a compiler error
- Sounds a bit useless

We can also *cast* them!

```
1  int hello[5] = {0};
2  char* start = (char*) hello;
3  *start = 1;
4  printf("First value: %d\n", hello[0]);
```


We can also *cast* them!

```
1  int hello[5] = {0};
2  char* start = (char*) hello;
3  *start = 1;
4  printf("First value: %d\n", hello[0]); // 1!
```

We can also *cast* them!

```
1  int hello[5] = {0};
2  char* start = (char*) hello;
3  *start = 1;
4  printf("First value: %d\n", hello[0]); // 1!
5  // Why is it 1 and not 2147483648?
```

Some interesting things are writable

DoubleForker

DoubleForker!

Modifying Stackframes in GDB

Capture the flag!

```
1  #include <stdio.h>
2  #include <string.h>
3
4  int readIt() {
5      char buffer[10];
6      int result = 1;
7
8      printf("Enter the password: ");
9      gets(buffer);
10
11     if(strcmp(buffer, "Secret") == 0) {
12         result = 0;
13     }
14     return result;
15 }
16
17 void flag() {
18     printf("FLAG!\n");
19 }
20
21 int main() {
22     if (readIt() == 0) {
23         flag();
24     } else {
25         printf("Nope!\n");
26     }
27     return 0;
28 }
```

Modifying Stackframes in GDB

Useful Resources

- [Stackframe layout \(for overflowing\)](#)
- [Input non printable characters in GDB](#)
- Run it in GDB, Compile with `-fno-stack-protector`
- `lscpu` to find Endianness of your computer (probably Little Endian)

Modifying Stackframes in GDB

Useful Resources

- [Stackframe layout \(for overflowing\)](#)
- [Input non printable characters in GDB](#)
- Run it in GDB, Compile with `-fno-stack-protector`
- `lscpu` to find Endianness of your computer (probably Little Endian)

Spoiler Example Solution

Modifying Stackframes in GDB

Useful Resources

- [Stackframe layout \(for overflowing\)](#)
- [Input non printable characters in GDB](#)
- Run it in GDB, Compile with `-fno-stack-protector`
- `lscpu` to find Endianness of your computer (probably Little Endian)

Spoiler Example Solution

Created on my computer while running in GDB. Addresses may vary.

```
r <<< $(python2 -c 'print
    "A"*14 # Padding to overflow to base pointer
    + "\x00\x00\x7f\xff\xff\xff\xdd\x02"[:-1] # overwrite BP
    + "\x00\x00\x55\x55\x55\x55\x51\xbd"[:-1] # overwrite ret addr
    ')
```

Modifying Stackframes in Real Life

Why that won't work 1: gets

- We needed to inject some null bytes to pad our addresses to 64 bit.
- `gets` just reads until EOF/newline and doesn't care about size or zeroes

⇒ *Really* dangerous

⇒ Removed from the C standard (but it still probably compiles)

Why that won't work 2: Stack-Canary

- The compiler inserts a random number between the local variables and the return address
- If it is changed the program will die with a nice message:
`*** stack smashing detected ***: terminated`
- Disabled with `-fno-stack-protector`

Why that won't work 3: ASLR

- We injected *absolute addresses* to jump to (or set the base pointer to)
 - Modern Operating Systems employ **A**ddress **S**pace **L**ayout **R**andomization
- ⇒ Your code is loaded at a random address and absolute addresses don't work
- GDB disables ASLR for the loaded program so we ran it in there

Code! And where is code stored?

Functions are...

Code! And where is code stored?

In memory. So let's point to it.

```
1 #include <stdio.h>
2
3 void sayHi(char* name) {
4     printf("Hello, %s!\n", name);
5 }
6
7 int main() {
8     (&sayHi)("John");
9     return 0;
10 }
```

How can we declare a variable?

```
1   ??? myFunction = &sayHi;
```

How can we declare a variable?

```
1   ??? myFunction = &sayHi;  
2   // We basically copy the declaration  
3   void sayHi(char* name) = &sayHi;
```

How can we declare a variable?

```
1   ??? myFunction = &sayHi;
2   // We basically copy the declaration
3   void sayHi(char* name) = &sayHi; // Nearly!
4   void (sayHi)(char* name) = &sayHi;
```

How can we declare a variable?

```
1   ??? myFunction = &sayHi;
2   // We basically copy the declaration
3   void sayHi(char* name) = &sayHi; // Nearly!
4   void (sayHi)(char* name) = &sayHi; // Closer!
5   void (* myName)(char* name) = &sayHi;
```

How can we declare a variable?

```
1   ??? myFunction = &sayHi;
2   // We basically copy the declaration
3   void sayHi(char* name) = &sayHi; // Nearly!
4   void (sayHi)(char* name) = &sayHi; // Closer!
5   void (* myName)(char* name) = &sayHi;
6   //~~~ ^ ~~~~~ ~~~~~~
7   // | Pointer   | Parameters delimited with ,
8   // Return Type |
9   //           Var name
```


How can we declare a variable?

```
1 #include <stdio.h>
2
3 void greetMultiple(char* name, char* other) {
4     printf("Hello, %s and %s!\n", name, other);
5 }
6
7 int main() {
8     void (*myFunction)(char* name, char*) = &greetMultiple;
9     myFunction("Peter", "Jane");
10    return 0;
11 }
```

Aufgabe

1. Schreibe eine Insertion-Sort die ein int array, einen compare-Funktionspointer und was man sonst so braucht entgegennimmt
2. Schreibe eine compare-Funktion, die Integer absteigend sortiert und nutze sie
3. Schreibe eine compare-Funktion, die Integer basierend auf der Anzahl an 1-Bits sortiert

```
1  int main() {
2      int array[] = {5, 3, 2, 4, 1};
3      int length = sizeof(array) / sizeof(int);
4
5      insertionSort(array, length, compareDesc);
6
7      for(int i = 0; i < length; i++) {
8          printf("%d - %d\n", i, array[i]);
9      }
10     return 0;
```



XKCD 313 - Insomnia

F R A G E N?



<https://forms.gle/9CwJSKidKibubran9>

Bis nächste Woche