# Übung zu Betriebssystembau

Zeitscheibenscheduling

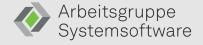
10. Dezember 2024

#### **Alexander Krause**

Arbeitsgruppe Systemsoftware Technische Universität Dortmund

(Mit Material vom Lehrstuhl 4 der FAU)





# **Kooperative Ablaufplanung**



## **Unterbrechende Ablaufplanung**



# **Unterbrechende Ablaufplanung**



## **Unterbrechende Ablaufplanung**

Aufgabe: Präemptives Scheduling mittels Timer.



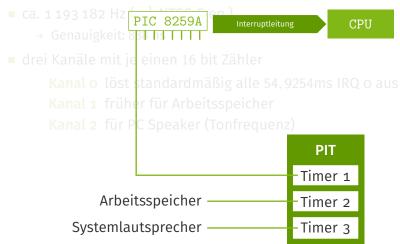
- Standardtimer seit 1981 (IBM-PC, Intel 8253/8254)
- ca. 1 193 182 Hz (=  $\frac{1}{3}$  NTSC-Freq.)
  - → Genauigkeit: 838 ns
- drei Kanäle mit je einen 16 bit Zähler

Kanal o löst standardmäßig alle 54,9254ms IRQ o aus

Kanal 1 früher für Arbeitsspeicher

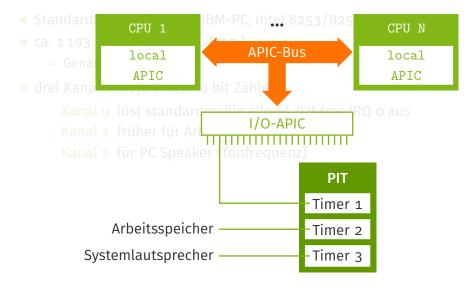
Kanal 2 für PC Speaker (Tonfrequenz)

```
Standardtimer seit 1981 (IBM-PC, Intel 8253/8254)
```





ak





- Standardtimer seit 1981 (IBM-PC, Intel 8253/8254)
- ca. 1 193 182 Hz (=  $\frac{1}{3}$  NTSC-Freq.)
  - → Genauigkeit: 838 ns
- drei Kanäle mit je einen 16 bit Zähler

Kanal o löst standardmäßig alle 54,9254ms IRQ o aus Kanal o früher für Arbeitsspeicher Kanal o für PC Speaker (Tonfrequenz)

via PIC bzw. I/O APIC → (relativ) langsam

- Standardtimer seit 1981 (IBM-PC, Intel 8253/8254)
- ca. 1 193 182 Hz (=  $\frac{1}{3}$  NTSC-Freq.)
  - → Genauigkeit: 838 ns
- drei Kanäle mit je einen 16 bit Zähler

Kanal o löst standardmäßig alle 54,9254ms IRQ o aus Kanal o früher für Arbeitsspeicher Kanal o für PC Speaker (Tonfrequenz)

- via PIC bzw. I/O APIC → (relativ) langsam
- ausreichend für frühere BSB-Varianten, geht aber besser.
- → machine/pit.h



### Real Time Clock (RTC)

- seit 1984 (IBM-PC/AT)
- 32 768 Hz (= 2<sup>15</sup> Hz, Verwendung in Uhren)
  - Standardmäßig Interrupts bei 1024 Hz (fast 1 ms)
  - 12 weitere Möglichkeiten von 2 bis 8 192 Hz durch Vorteiler
  - IRQ 8
- für Zeit & Datum
- Betrieb im ausgeschalteten Zustand mittels Batterie

## Time Stamp Counter (TSC)

- seit 1993 (Pentium)
- 64 bit, auslesbar über Assemblerinstruktion rdtsc
- Taktfrequenz wie CPU
  - ursprünglich Erhöhung mit jedem Clock-Signal
  - unterschiedliche Takte abhängig vom Stromsparmodus
  - bei neueren Versionen: konstante Rate entsprechend nominaler Geschwindigkeit
- kann keinen Interrupt auslösen
- → machine/tsc.h



## **ACPI Power Management Timer**

- seit es ACPI-Mainboards gibt (1996)
- 3 579 545 Hz (= NTSC-Freq.)
- ein 24 oder 32 bit Zähler
  - besser als alte (nicht konstante) TSC
  - Zugriff über I/O Port
- kann auch keinen Interrupt auslösen

## **High Precision Event Timer (HPET)**

- von Intel und Microsoft 2005 als PIT- & RTC-Ersatz veröffentlicht
- > 10 MHz
  - → Genauigkeit: 100 ns oder besser
- ein 64 bit Zähler
  - min. drei 32 oder 64 bit breite Vergleichseinrichtungen
  - konfigurierbarer Interrupt bei Gleichheit

#### **LAPIC Timer**

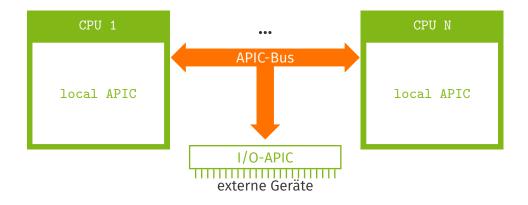
- > 100 MHz
  - → Genauigkeit: 10 ns oder besser
- 32 bit Zähler
- verwendet Busfrequenz
  - abhängig vom System
  - aber unabhängig von Stromsparmodus
  - Interrupt geht nur an den entsprechenden Kern
  - 8 Möglichkeiten (bis  $\frac{1}{128}$  Busfrequenz) durch Vorteiler

#### **LAPIC Timer**

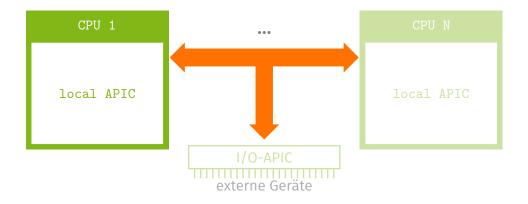
- ≥ 100 MHz
  - → Genauigkeit: 10 ns oder besser
- 32 bit Zähler
- verwendet Busfrequenz
  - abhängig vom System
  - aber unabhängig von Stromsparmodus
  - Interrupt geht nur an den entsprechenden Kern
  - 8 Möglichkeiten (bis  $\frac{1}{128}$  Busfrequenz) durch Vorteiler

#### Perfekt für unsere Bedürfnisse

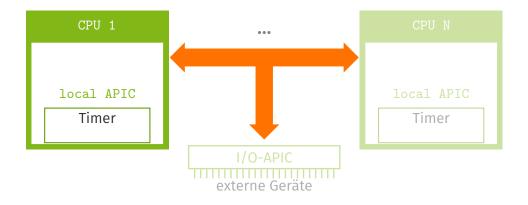




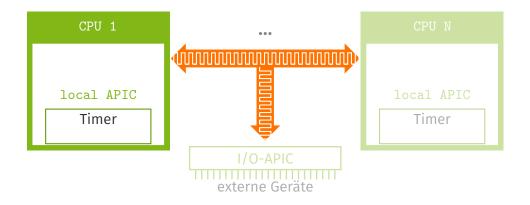




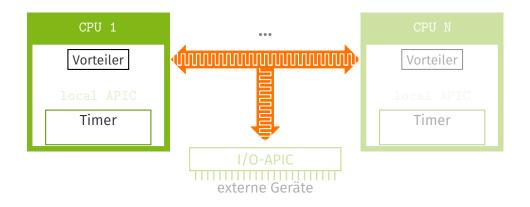




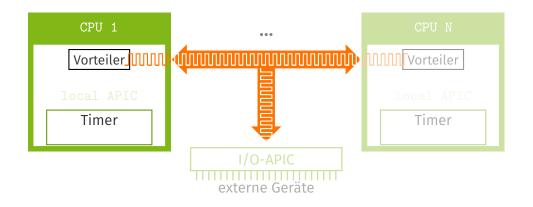




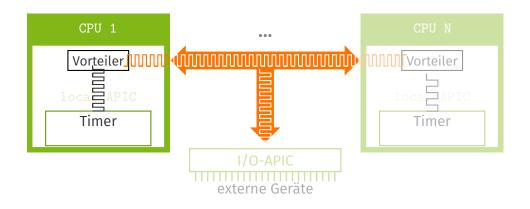






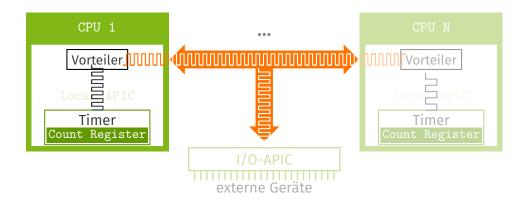




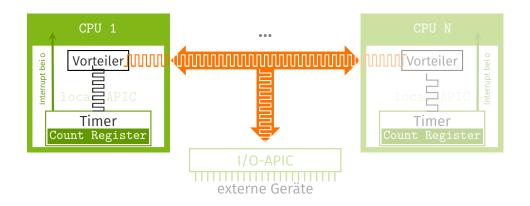




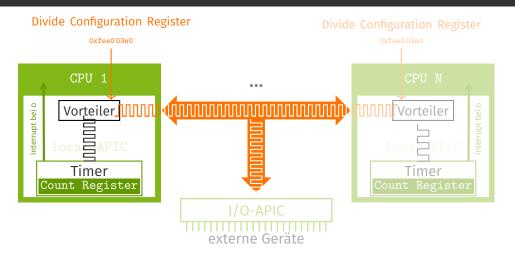
9/21



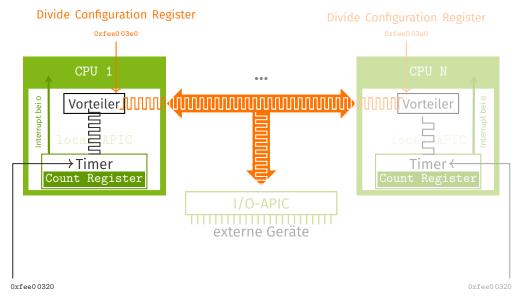




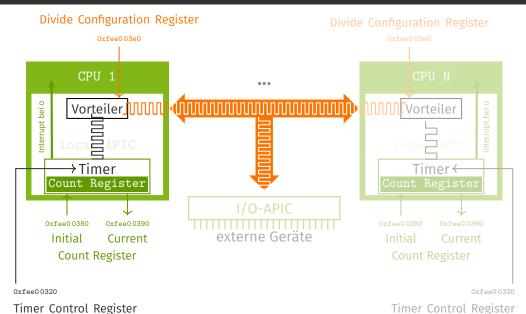








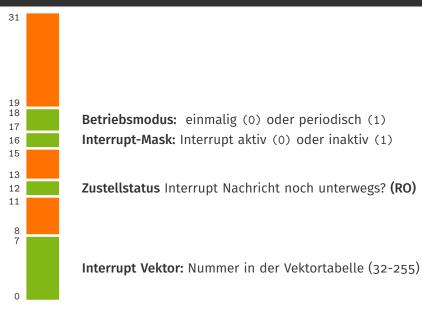






ak

# Aufbau des Timer Control Register Eintrags





## **Zusammenfassung LAPIC Timer**

- jede CPU hat einen eigenen 32bit Timer
- Änderung am INITIAL COUNT REGISTER startet den Timer
- zu Beginn wird der initiale Startzählwert aus dem INITIAL COUNT REGISTER in das CURRENT COUNT REGISTER kopiert
- welches im Bustakt dekrementiert wird
- bei o wird sofern aktiviert ein Interrupt ausgelöst
- je nach Betriebsmodus wird gestoppt oder wieder neu begonnen

```
windup stellt das Unterbrechungsintervall ein
(z.B. alle 1000 Mikrosekunden)

activate setzt den Timer und aktiviert Interrupts

prologue fordert Epilog an
(und kann zu Testzwecken eine Ausgabe tätigen)

epilogue wechselt die Anwendung mittels scheduler.resume()
```



1. LAPIC-Timer kalibrieren



Umsetzung

#### 1. LAPIC-Timer kalibrieren

in der Funktion LAPIC::Timer::ticks(Funktion gibt die Anzahl der Ticks in einer Millisekunde zurück)



#### 1. LAPIC-Timer kalibrieren

- in der Funktion LAPIC::Timer::ticks
   (Funktion gibt die Anzahl der Ticks in einer Millisekunde zurück)
- benötigt zur Konfiguration die Funktion LAPIC::Timer::set, Aufruf mit
  - maximalem Wert für den LAPIC-Zähler,
  - als ONE SHOT und
  - ohne Unterbrechungen

#### 1. LAPIC-Timer kalibrieren

- in der Funktion LAPIC::Timer::ticks
   (Funktion gibt die Anzahl der Ticks in einer Millisekunde zurück)
- benötigt zur Konfiguration die Funktion LAPIC::Timer::set, Aufruf mit
  - maximalem Wert für den LAPIC-Zähler,
  - als ONE SHOT und
  - ohne Unterbrechungen
- unter Verwendung des PIT
  - Wartezeit von mehreren MS einstellen: PIT::set
  - mittels PIT::waitForTimeout warten
  - Start- und Endwert des LAPIC-Zählerregisters merken

#### 1. LAPIC-Timer kalibrieren

- in der Funktion LAPIC::Timer::ticks
   (Funktion gibt die Anzahl der Ticks in einer Millisekunde zurück)
- benötigt zur Konfiguration die Funktion LAPIC::Timer::set, Aufruf mit
  - maximalem Wert für den LAPIC-Zähler,
  - als ONE SHOT und
  - ohne Unterbrechungen
- unter Verwendung des PIT
  - Wartezeit von mehreren MS einstellen: PIT::set
  - mittels PIT::waitForTimeout warten
  - Start- und Endwert des LAPIC-Zählerregisters merken
- Hilfsstrukturen in lapic\_timer.cc & lapic\_registers.h





#### 2. Initialen Wert und Vorteiler korrekt setzen

■ Interrupt soll alle *n* Mikrosekunden ausgelöst werden

- Interrupt soll alle *n* Mikrosekunden ausgelöst werden
- Startwertzähler  $initial = \frac{n \cdot \text{LAPIC} : : \text{Timer} : : \text{ticks}()}{Vorteiler}$  berechnen, dabei gilt  $Vorteiler = 2^{\times}$  mit  $x \in \{0, \dots, 7\}$



- Interrupt soll alle *n* Mikrosekunden ausgelöst werden
- Startwertzähler  $initial = \frac{n \cdot \text{LAPIC} : : \text{Timer} : : \text{ticks}()}{Vorteiler}$  berechnen, dabei gilt  $Vorteiler = 2^{\times}$  mit  $x \in \{0, \dots, 7\}$
- Möglichst kleiner Vorteiler, aber kein Überlauf (32 bit!)

- Interrupt soll alle *n* Mikrosekunden ausgelöst werden
- Startwertzähler  $initial = \frac{n \cdot \text{LAPIC} : : \text{Timer} : : \text{ticks}()}{Vorteiler}$  berechnen, dabei gilt  $Vorteiler = 2^{\times}$  mit  $x \in \{0, \dots, 7\}$
- Möglichst kleiner Vorteiler, aber kein Überlauf (32 bit!)
- Beispiel: watch.windup(5000000);

```
n = 5000000 \, \mu s = 5 \, s
```

LAPIC::Timer::ticks 1000000 ms<sup>-1</sup>

- Interrupt soll alle *n* Mikrosekunden ausgelöst werden
- Startwertzähler  $initial = \frac{n \cdot \text{LAPIC} : : \text{Timer} : : \text{ticks}()}{Vorteiler}$  berechnen, dabei gilt  $Vorteiler = 2^{\times}$  mit  $x \in \{0, \dots, 7\}$
- Möglichst kleiner Vorteiler, aber kein Überlauf (32 bit!)
- Beispiel: watch.windup(5000000);

```
n = 50000000 \,\mu s = 5 \,s
LAPIC::Timer::ticks 10000000 \,m s^{-1}

Vorteiler 1 = 2^0
```



- Interrupt soll alle *n* Mikrosekunden ausgelöst werden
- Startwertzähler  $initial = \frac{n \cdot \text{LAPIC} : : \text{Timer} : : \text{ticks}()}{Vorteiler}$  berechnen, dabei gilt  $Vorteiler = 2^{\times}$  mit  $x \in \{0, \dots, 7\}$
- Möglichst kleiner Vorteiler, aber kein Überlauf (32 bit!)



- Interrupt soll alle n Mikrosekunden ausgelöst werden
- Startwertzähler  $initial = \frac{n \cdot \text{LAPIC} : : \text{Timer} : : \text{ticks}()}{Vorteiler \cdot 1000}$  berechnen, dabei gilt  $Vorteiler = 2^{\times} \text{ mit } x \in \{0, \dots, 7\}$
- Möglichst kleiner Vorteiler, aber kein Überlauf (32 bit!)



- Interrupt soll alle *n* Mikrosekunden ausgelöst werden
- Startwertzähler  $initial = \frac{n \cdot \text{LAPIC} : : \text{Timer} : : \text{ticks}()}{Vorteiler}$  berechnen, dabei gilt  $Vorteiler = 2^{\times}$  mit  $x \in \{0, \dots, 7\}$
- Möglichst kleiner Vorteiler, aber kein Überlauf (32 bit!)



- Interrupt soll alle *n* Mikrosekunden ausgelöst werden
- Startwertzähler  $initial = \frac{n \cdot \text{LAPIC} : : \text{Timer} : : \text{ticks}()}{Vorteiler \cdot 1000}$  berechnen, dabei gilt  $Vorteiler = 2^{\times} \text{ mit } x \in \{0, \dots, 7\}$
- Möglichst kleiner Vorteiler, aber kein Überlauf (32 bit!)



- Interrupt soll alle *n* Mikrosekunden ausgelöst werden
- Startwertzähler  $initial = \frac{n \cdot \text{LAPIC} : : \text{Timer} : : \text{ticks}()}{Vorteiler}$  berechnen, dabei gilt  $Vorteiler = 2^{\times}$  mit  $x \in \{0, \dots, 7\}$
- Möglichst kleiner Vorteiler, aber kein Überlauf (32 bit!)

```
■ Beispiel: watch.windup(5000000);

n 5000000 µs = 5 s

LAPIC::Timer::ticks 1000000 ms<sup>-1</sup>

Vorteiler 2 = 2<sup>1</sup>

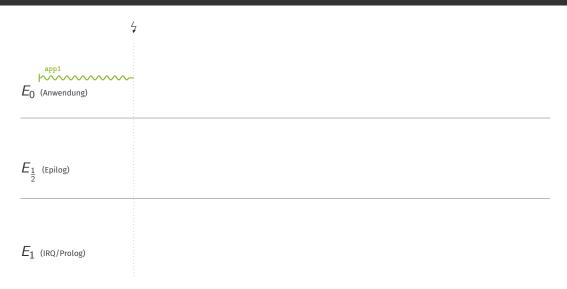
initial 2500000000
```



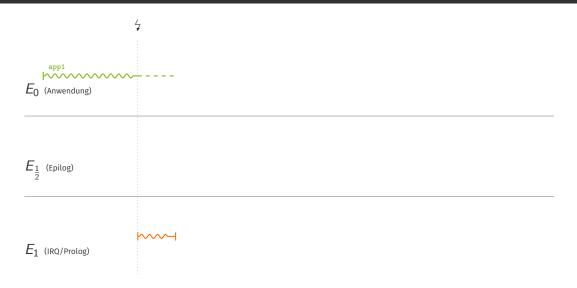
 $E_0$  (Anwendung)

 $E_{rac{1}{2}}$  (Epilog)

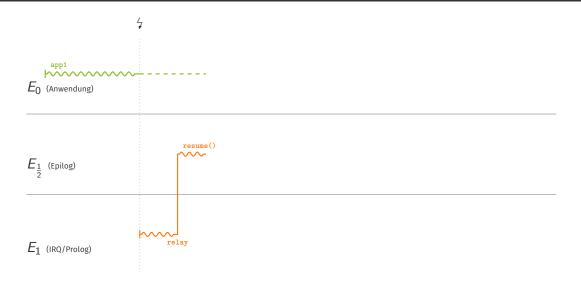












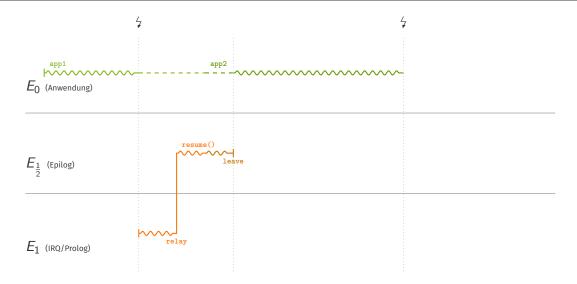




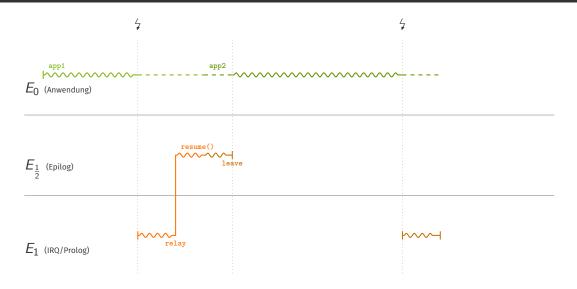




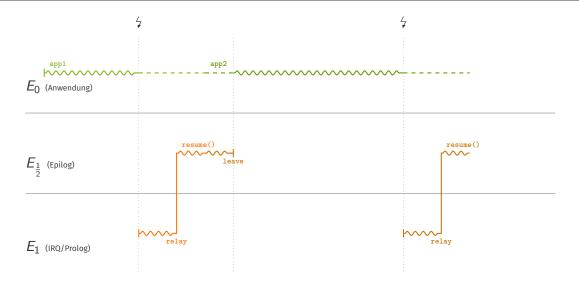




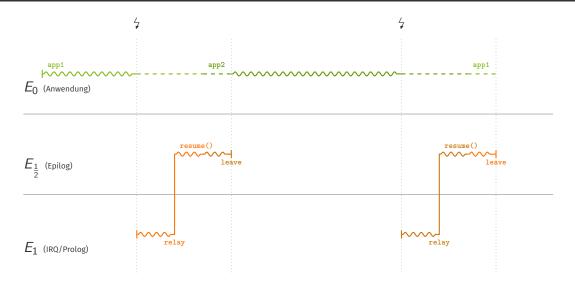




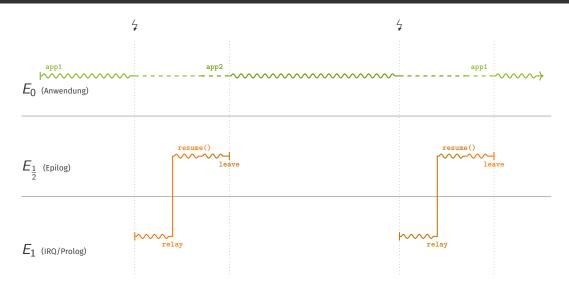
















$$E_{rac{1}{2}}$$
 (Epilog)



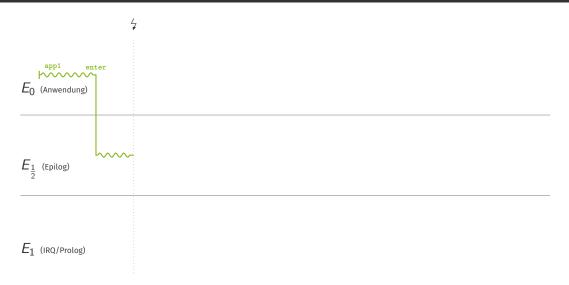


Umsetzung





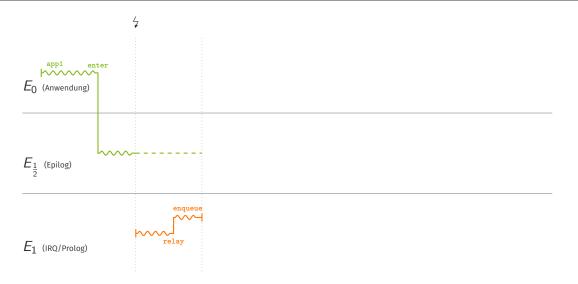




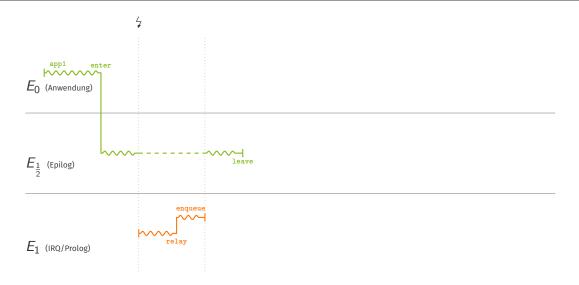




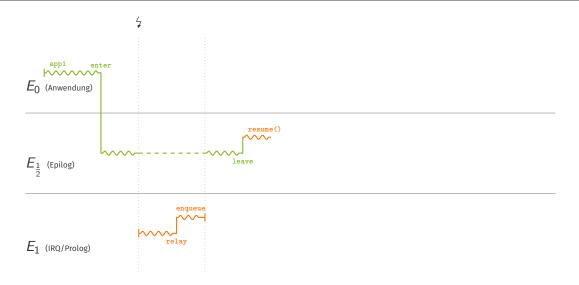




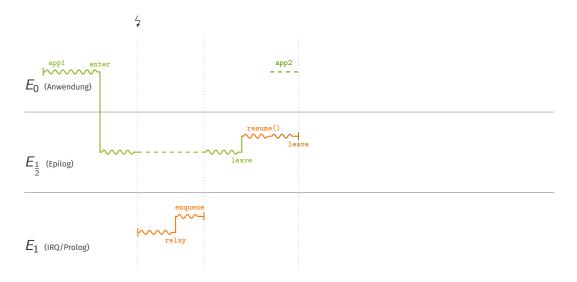




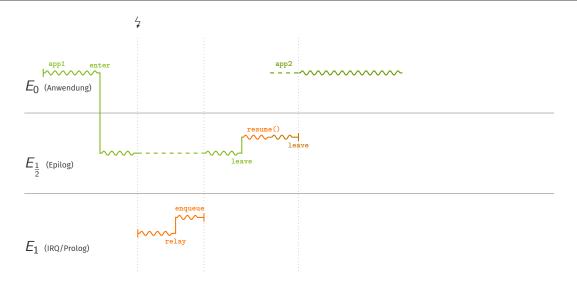




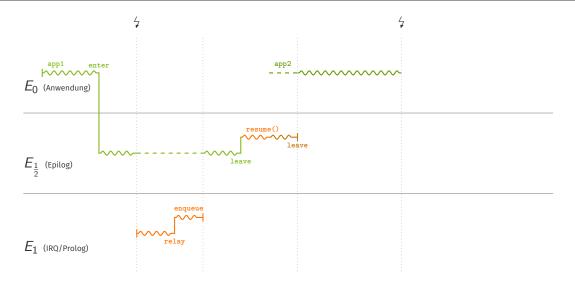




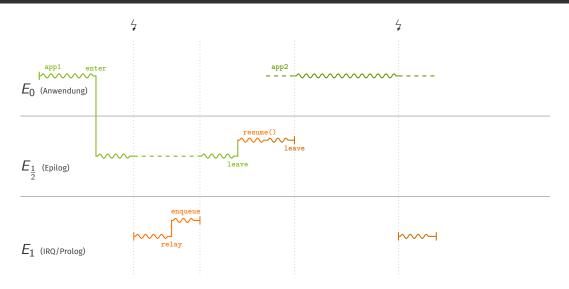




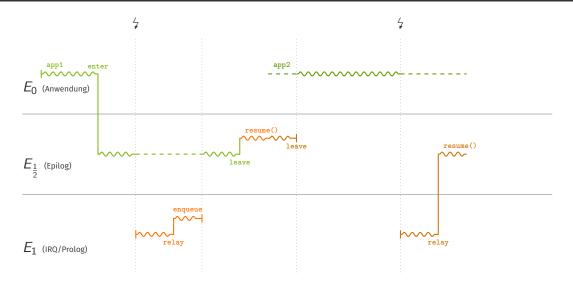




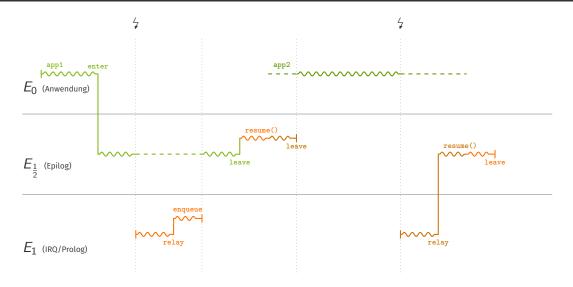






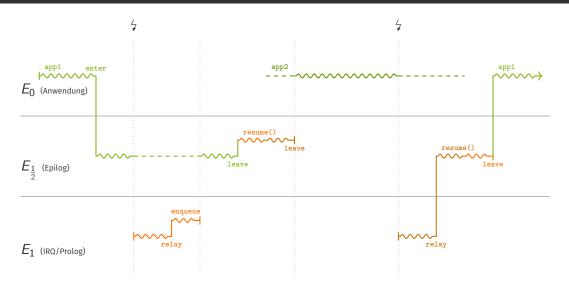








# Ablaufbeispiel bei Faden in Systemebene



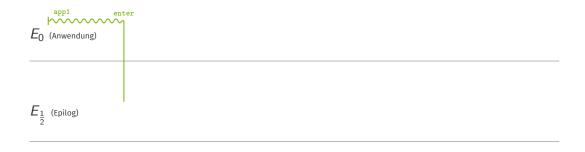


 $E_0$  (Anwendung)

 $E_{rac{1}{2}}$  (Epilog)

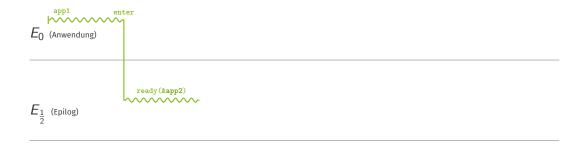
 $E_1$  (IRQ/Prolog)





 $E_1$  (IRQ/Prolog)





 $E_1$  (IRQ/Prolog)





 $E_1$  (IRQ/Prolog)

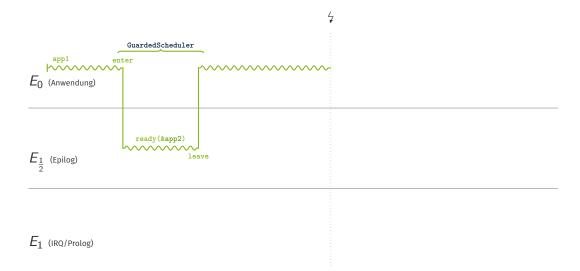




Umsetzung

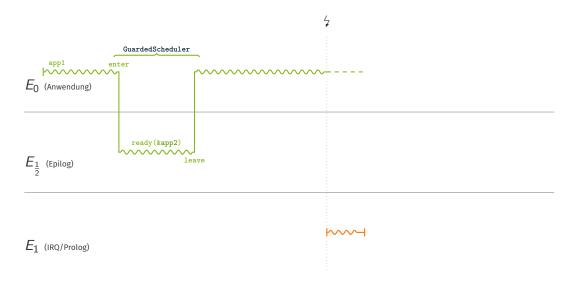
 $E_1$  (IRQ/Prolog)





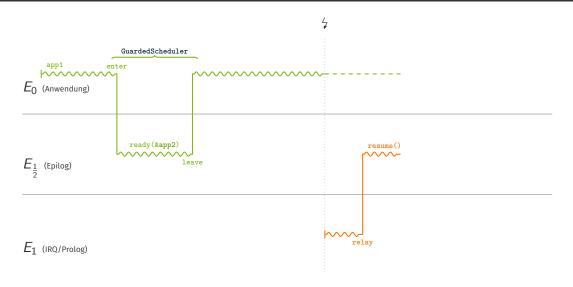


ak





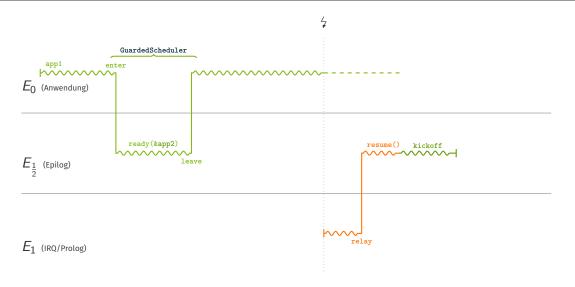
Umsetzung





ak

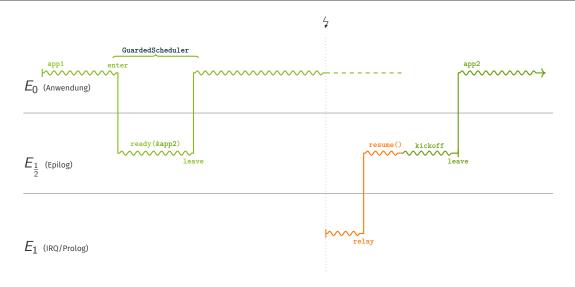
Umsetzung





Umsetzung

16/21





ak Umsetzung

16/21

Präemptives Beenden mittels Scheduler::kill(Thread&)



Präemptives Beenden mittels Scheduler::kill(Thread&)

OOStuBS keine Änderung

Präemptives Beenden mittels Scheduler::kill(Thread&)

OOStuBS keine Änderung: aus der Ready-Liste entfernen bzw. Kill-Flag setzen und bei resume prüfen



Präemptives Beenden mittels Scheduler::kill(Thread&)

OOStuBS keine Änderung: aus der Ready-Liste entfernen bzw. Kill-Flag setzen und bei resume prüfen

Präemptives Beenden mittels Scheduler::kill(Thread&)

OOStuBS keine Änderung: aus der Ready-Liste entfernen bzw. Kill-Flag setzen und bei resume prüfen

MPStuBS der Anwendungsfaden kann gerade auf einer anderen CPU laufen

Kill-Flag setzen (wie gehabt)

Präemptives Beenden mittels Scheduler::kill(Thread&)

OOStuBS keine Änderung: aus der Ready-Liste entfernen bzw. Kill-Flag setzen und bei resume prüfen

- Kill-Flag setzen (wie gehabt)
- falls er nicht in der Ready-Liste ist, läuft er wohl gerade auf einer anderen CPU

Präemptives Beenden mittels Scheduler::kill(Thread&)

OOStuBS keine Änderung: aus der Ready-Liste entfernen bzw. Kill-Flag setzen und bei resume prüfen

- Kill-Flag setzen (wie gehabt)
- falls er nicht in der Ready-Liste ist, läuft er wohl gerade auf einer anderen CPU
- diese andere CPU muss benachrichtigt werden

Präemptives Beenden mittels Scheduler::kill(Thread&)

OOStuBS keine Änderung: aus der Ready-Liste entfernen bzw. Kill-Flag setzen und bei resume prüfen

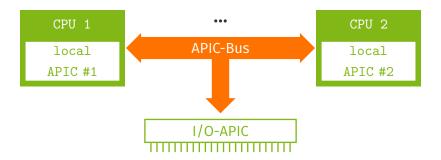
- Kill-Flag setzen (wie gehabt)
- falls er nicht in der Ready-Liste ist, läuft er wohl gerade auf einer anderen CPU
- diese andere CPU muss benachrichtigt werden
  - → INTER PROCESSOR INTERRUPT (IPI)

Präemptives Beenden mittels Scheduler::kill(Thread&)

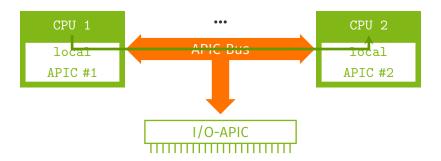
OOStuBS keine Änderung: aus der Ready-Liste entfernen bzw. Kill-Flag setzen und bei resume prüfen

- Kill-Flag setzen (wie gehabt)
- falls er nicht in der Ready-Liste ist, läuft er wohl gerade auf einer anderen CPU
- diese andere CPU muss benachrichtigt werden
  - → INTER PROCESSOR INTERRUPT (IPI)
- die angesprochene CPU muss dann das Kill-Flag des aktuellen Prozesses prüfen

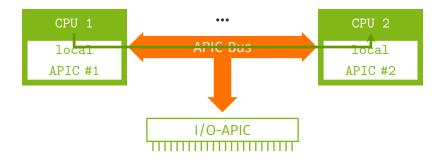






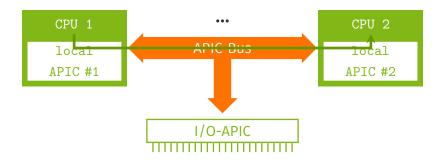






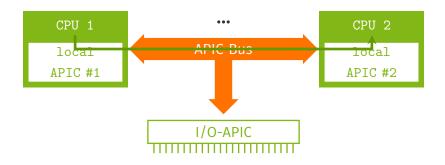
LAPIC::IPI::send(destination, vector);



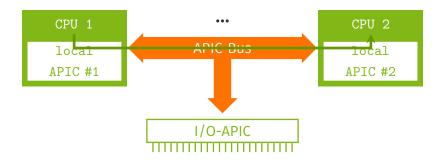


LAPIC::IPI::send(destination, vector);
Interrupt









```
destination = APIC::getLAPICID(cpu);
LAPIC::IPI::send(destination, vector);
                    Empfänger
```

Interrupt

#### Der Teufel steckt im Detail

#### Was kann hier schon schief gehen?

```
1 lock[Core::getID()] = true;
```



#### Was kann hier schon schief gehen?

```
1 lock[Core::getID()] = true;
```

#### Viel - diese Zeile wird nicht atomar ausgeführt:

```
; Array lock an Adresse 0x2000
```

```
2 call <Core::getID()>
3 mov [rax+0x2000], 0x1
```

#### Was kann hier schon schief gehen?

```
1 lock[Core::getID()] = true;
```

#### Viel - diese Zeile wird nicht atomar ausgeführt:

```
1 ; Array lock an Adresse 0x2000

2 call <Core::getID()> 5 Scheduler Interrupt
3 mov [rax+0x2000], 0x1
```



#### Was kann hier schon schief gehen?

```
1 lock[Core::getID()] = true;
```

#### Viel - diese Zeile wird nicht atomar ausgeführt:

```
1 ; Array lock an Adresse 0x2000

2 call <Core::getID()> 5 Scheduler Interrupt
3 mov [rax+0x2000], 0x1
```

Was passiert nun, wenn der Anwendungsfaden anschließend auf einer anderen CPU eingeplant wird?



## Fragen über Fragen

• Kann nun eine fehlerhafte Anwendung unser Betriebssystem blockieren?



## Fragen über Fragen

- Kann nun eine fehlerhafte Anwendung unser Betriebssystem blockieren?
- Unter welchen Umständen wäre es effizienter, den Timer abzuschalten (Stichwort tickless)?

## Fragen über Fragen

- Kann nun eine fehlerhafte Anwendung unser Betriebssystem blockieren?
- Unter welchen Umständen wäre es effizienter, den Timer abzuschalten (Stichwort tickless)?

Abgabe der Aufgabe bis Mittwoch, den 08. Januar

Gibt es noch Fragen?