Echtzeitbetriebssysteme

Oliver Jack

Ernst-Abbe-Hochschule Jena Fachbereich Elektrotechnik und Informationstechnik

Sommersemester 2025



- Lernziele dieser Vorlesung
- Q Grundbegriffe
 - Annahmen und Bezeichnungen
 - Anforderungen
 - Präzedenzen
- Planen
 - Pläne
 - Strategien
- Eigenschaften von Plänen
 - Bewertung
 - Anomalien
- Zusammenfassung

- Lernziele dieser Vorlesung
- ② Grundbegriffe
 - Annahmen und Bezeichnungen
 - Anforderungen
 - Präzedenzen
- 3 Planer
 - Pläne
 - Strategien
- 4 Eigenschaften von Plänen
 - Bewertung
 - Anomalien
- 5 Zusammenfassung



Lernziele

- Kenntnis von Grundbegriffen der Echtzeitplanung
- Kenntnis von Präzedenzsystemen
- Kenntnis von Plänen (Schedules) und Planungsstrategien

- Lernziele dieser Vorlesung
- **Grundbegriffe**
 - Annahmen und Bezeichnungen
 - Anforderungen
 - Präzedenzen
- 3 Planer
 - Pläne
 - Strategien
- 4 Eigenschaften von Plänen
 - Bewertung
 - Anomalien
- 5 Zusammenfassung



- Lernziele dieser Vorlesung
- Grundbegriffe
 - Annahmen und Bezeichnungen
 - Anforderungen
 - Präzedenzen
- - Pläne
 - Strategien
- Eigenschaften von Plänen
 - Bewertung
 - Anomalien



Annahmen

Die Ablaufgeschwindigkeit der physikalischen Vorgänge im technischen Prozess bestimmt notwendige Reaktionszeiten des Rechners und legt Antwortzeiten fest.

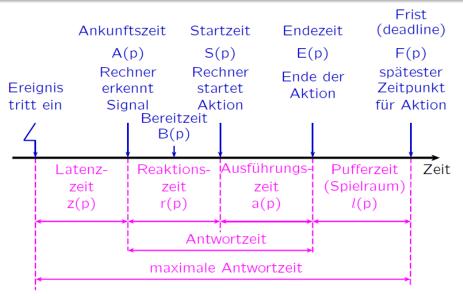
Es treten periodische und spontane Anforderungen (Zeitpunkte, Signale, Alarme), harte und weiche Zeitbedingungen auf.

Planung für Prozessoren und Betriebsmittel

Klassifikation

- statische / dynamische Planung
- explizit / implizite Planung
- offline / online Planung

Zeitablauf eines Ereignisses im Prozess p



- Großbuchstaben: Zeitpunkte
- Kleinbuchstaben: Zeitdauern
- Reaktionszeit und Ausführungszeit abhängig von der Umgebung, z. B. von anderen wichtigeren Prozessen
- Latenzzeit: u. a. Verzögerungen durch lokale Hardware, Übertragung über Netze
- Bereitzeit: Der Prozess darf vor seiner Bereitzeit nicht gestartet werden, da z. B. dann erst benötigte Daten verfügbar; in der Regel B(p) = A(p)

- Lernziele dieser Vorlesung
- Grundbegriffe
 - Annahmen und Bezeichnungen
 - Anforderungen
 - Präzedenzen
- - Pläne
 - Strategien
- Eigenschaften von Plänen
 - Bewertung
 - Anomalien



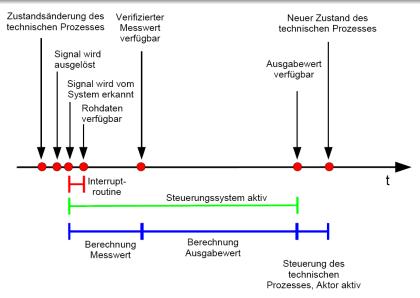
Forderungen an Echtzeitplanung

- A priori zeigen, dass stets Zeitrandbedingungen erfüllt, d. h. $\forall p : E(p) < F(p)$
- Einplanung nicht nur der CPU, sondern auch anderer Betriebsmittel
- Zeitbedingungen einhalten bzgl. paralleler Vorgänge (Rechenprozesse)
- Berücksichtigung von Kausalzusammenhängen (Synchronisation. Vorranggraphen, Präzedenzsysteme)
- Einhalten von Zeitbedingungen bei der Verteilung von Information (Kommunikation, gemeinsames Führungsziel)
- Präemptive Prozessorzuteilung

Nicht trivial, da technische Systeme vielfach komplex und schwer überschaubar



Beispiel technischer Prozess



Dimensionierung der Rechnerleistung

Aufstellen der Worst Case-Analyse

- Bedarf aus bekannten periodischen Anforderungen
- Bedarf aus erwarteten spontanen Anforderungen
- Zuschlag von 100% oder mehr zum Auffangen von Lastspitzen

Unterschied zu konventionellen Betriebszielen

- keine maximale Auslastung des Prozessors,
- keine Durchsatzoptimierung,
- Abläufe sollten determiniert abschätzbar sein.



Weitere Bezeichnungen

Algorithmen zur korrekten Zuteilung von Rechenleistung bei mehreren zu bearbeitenden Aktionen (Rechenprozessen) P_1, P_2, \ldots, P_n müssen für jeden Prozess P_i beinhalten:

- $F(p_i)$ spätester Zeitpunkt (Frist) für Antwort an technischen Prozess
- $B(p_i)$ Bereitzeitpunkt (frühester Startzeitpunkt)
- $a(p_i)$ Gesamtausführungszeit

Weitere wichtige Größen, die von der aktuellen Zeit t abhängig sind:

- a(p_i, t) Restausführungszeit des Prozesses P_i zur Zeit t; wobei $a(p_i, B(p_i)) = a(p_i) \ a(p_i, F(p_i)) = 0$, falls Prozess zeitgerecht beendet
- $\bullet \Delta t(p_i, t)$ Zeitspanne bis zur Frist des Prozesses P_i ; wobei $\Delta t(p_i, t) = F(p_i) - t$
- I(p_i, t) Zur Zeit t noch verfügbarer Spielraum für Prozess P_i : $I(p_i, t) = \Delta t(p_i, t) - a(p_i, t)$ falls: $I(p_i, t) < 0 \rightarrow P_i$ kann nicht rechtzeitig fertiggestellt werden Echtzeitforderung verletzt!

Annahmen

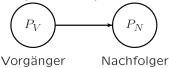
- Laufzeit $a(p_i)$ von Programmen bei Echtzeitumgebung bekannt, da Aufgaben und Aufgabenprofile festgelegt (Analyse des Codes, Messungen)
- Ggf. Präzedenzen der Aktionen P_i beachten (z. B. P_i muss vor P_i ausgeführt werden)

- Lernziele dieser Vorlesung
- Grundbegriffe
 - Annahmen und Bezeichnungen
 - Anforderungen
 - Präzedenzen
- - Pläne
 - Strategien
- Eigenschaften von Plänen
 - Bewertung
 - Anomalien



Präzedenzsysteme

- Folgen von untereinander abhängigen Aktionen (z. B. pipes)
- Bereitzeiten, Antwortzeiten oder Fristen ergeben sich evtl. erst aus vorherigem Ablauf
- Beispiel f
 ür Graph-Darstellung



• Formel-Notation $p_V \prec p_N$

Beispiel

Bei der Planung mit Präzedenzsystemen muss also berücksichtigt werden, dass auch Folgeprozesse noch rechtzeitig fertig werden



Verfahren zur Planung von Präzedenzsystemen

Es wird ein normalisiertes Präzedenzsystem PS' anstelle des ursprünglichen Präzedenzsystems PS eingeführt

Normalisierung (Durchführung zur Zeit t)

- $a'(p_i) = a(p_i)$ und $a'(p_i, t) = a(p_i, t)$
- $: N(p_i) = \emptyset$ • $F'(p_i, t) = \begin{cases} F(p_i) & \text{if } (p_i) \\ \min \left(F(p_i, \min_{q \in \mathcal{N}(p_i)} (F'(q) - a(q, t)) \right) \text{ : sonst } \end{cases}$
- wobei $\mathcal{N}(p_i)$ die Menge der unmittelbaren Nachfolger von p_i ist.

Rekursive Berechnung, beginnend mit Prozessen ohne Nachfolger im Präzedenzgraphen



Normalisierung zu $B'(p_i)$

- Hängen die Bereitzeiten von externen Vorgängen ab, die unabhängig vom Scheduling der Prozesse sind, dann gilt $B'(p_i) = B(p_i)$
- Hängen die Bereitzeiten von der Beendigung vorheriger Prozesse ab, dann ist $B'(p_i)$ abhängig vom konkreten Scheduling.
- Ahnlich wie in der Netzplantechnik lassen sich früheste und späteste Start-Termine für die Prozesse angeben. Diese sind aber für das Scheduling ohne Bedeutung, da die konkreten Präzedenzen berücksichtigt werden müssen.

Satz

Ein Präzedenzsystem ist nur dann planbar, falls das zugehörige normalisierte PS' planbar ist, d. h. es dafür einen zulässigen Plan gibt.

- Lernziele dieser Vorlesung
- ② Grundbegriffe
 - Annahmen und Bezeichnungen
 - Anforderungen
 - Präzedenzen
- Planen
 - Pläne
 - Strategien
- 4 Eigenschaften von Plänen
 - Bewertung
 - Anomalien
- 5 Zusammenfassung



- Lernziele dieser Vorlesung
- 2 Grundbegriffe
 - Annahmen und Bezeichnungen
 - Anforderungen
 - Präzedenzen
- Planen
 - Pläne
 - Strategien
- 4 Eigenschaften von Plänen
 - Bewertung
 - Anomalien
- Zusammenfassung



Pläne

Phasen der Planung

- Test auf Einplanbarkeit (feasibility check)
- Planberechnung (schedule construction)
- Umsetzung auf Zuteilung im Betriebssystem (dispatching)

Gesucht

Plan mit aktueller Start- und Endezeit für jede Aktion P_i

Pläne

Darstellung z. B. als nach der Zeit geordnete Liste von Tupeln $(P_i, S(P_i), E(P_i), BM)$; wobei BM die Menge der benötigten Betriebsmittel (entfällt, wenn nur eines geplant wird)

Falls Prozesse unterbrochen werden: mehrere Tupel je P_i mit Anfangs- und Endzeitpunkten von aktiven Intervallen

Begriffe für Planung

Zulässiger Plan

Ein Plan, bei dem alle Prozesse einer Prozessmenge eingeplant werden und dabei keine Präzedenzrestriktionen und keine Zeitanforderungen verletzt werden

Optimales Planungsverfahren

Verfahren ist optimal, wenn es für jede Prozessmenge unter den gegebenen Randbedingungen einen zulässigen Plan findet, falls ein solcher existiert

Begriffe für Planung

Statische Planung

- Erfordert Kenntnis aller Prozesse P_i und von $B(p_i)$, $a(p_i)$ und $F(p_i)$
- Abgeschlossenes System
- Exakte Algorithmen gehören in diese Gruppe

Dynamische Planung

- Unvorhergesehen eintreffende Aktionen (Prozesse)
- Unbekannte Prozessparameter
- Keine Gesamtoptimierung



- Lernziele dieser Vorlesung
- ② Grundbegriffe
 - Annahmen und Bezeichnungen
 - Anforderungen
 - Präzedenzen
- Planen
 - Pläne
 - Strategien
- 4 Eigenschaften von Plänen
 - Bewertung
 - Anomalien
- Zusammenfassung



Grundstrategien zur Prozessbearbeitung

Präemptives (unterbrechbares) Abarbeiten

- Aktionen (Prozesse) haben Prioritäten
- Prioritäten statisch oder dynamisch berechnet
- Ausführung einer Aktion wird sofort unterbrochen, sobald Aktion mit höherer Priorität eintrifft
- unterbrochene Aktion wird an der Unterbrechungsstelle fortgesetzt, sobald keine Aktion höherer Priorität ansteht
- typisch für Echtzeitaufgaben (mit Ausnahme von Programmteilen, die zur Sicherung der Datenkonsistenz nicht unterbrochen werden dürfen)
- Nachteil: häufiges Umschalten reduziert Leistung

Nichtpräemptives (ununterbrechbares) Abarbeiten

- Eine begonnene Aktion wird zu Ende ausgeführt, selbst wenn während der Ausführung Aktionen höherer Dringlichkeit eintreffen
- z. B. Schreiben in Druckerpuffer, Schließen eines Ventils
- Nachteil: evtl. Versagen (zu hohe Reaktionszeit) des Systems beim Eintreffen "unvorhergesehener" Anforderungen

- Lernziele dieser Vorlesung
- ② Grundbegriffe
 - Annahmen und Bezeichnungen
 - Anforderungen
 - Präzedenzen
- 3 Planer
 - Pläne
 - Strategien
- 4 Eigenschaften von Plänen
 - Bewertung
 - Anomalien
- 5 Zusammenfassung



- Lernziele dieser Vorlesung
- ② Grundbegriffe
 - Annahmen und Bezeichnungen
 - Anforderungen
 - Präzedenzen
- 3 Planer
 - Pläne
 - Strategien
- Eigenschaften von Plänen
 - Bewertung
 - Anomalien
- Zusammenfassung



Bewertung von Plänen

- Performanz von Planungsalgorithmen (Scheduling-Algorithmen) wird durch Kostenfunktionen auf der Prozess-/Task-Menge bestimmt.
- In der Regel gilt es, diese Kosten zu minimieren.
- Minimiert werden z. B. die mittlere Antwortzeit, die Gesamtausführungszeit, die maximale Verspätung.

Mittlere Antwortzeit (Average response time)

$$ar{t_r} := rac{1}{n} \sum_{i=1}^n (E(p_i) - A(p_i))$$

Gesamtabschlusszeit (Total completion time)

$$t_c := \max_i E(p_i) - \min_i A(p_i)$$

Gewichtete Summe der Abschlusszeiten (Weighted sum of completion times)

$$t_w := \sum_{i=1}^n w_i E(p_i)$$

Die Gewichte w_i geben für jeden Prozess die Wichtigkeit einer fristgerechten Ausführung an.



Bewertung von Plänen

Maximale Verspätung (Maximum lateness)

$$L_{\mathsf{max}} = \max_{i} (E(p_i) - F(p_i))$$

Maximale Zahl verspäteter Tasks (Maximum number of late tasks)

$$N_{\mathsf{late}} := \sum_{i=1}^{n} \mathsf{miss}(E(p_i))$$

wobei

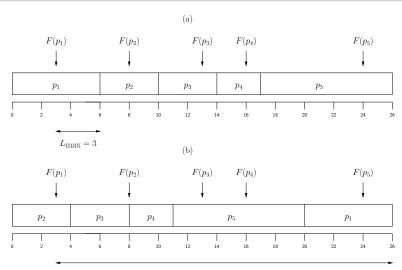
$$\mathsf{miss}(E(p_i)) := \left\{ egin{array}{l} 0 : \mathsf{falls} \ E(p_i) \leq F(p_i) \ 1 : \mathsf{sonst} \end{array} \right.$$



- Das Performanzmaß für einen Scheduling-Algorithmus bestimmt wesentlich die Performanz des Echtzeitsystems.
- Eine geeignete Wahl des Performanzmaßes hängt von der spezifischen Anwendung ab.
- Die mittlere Antwortzeit ist bei Echtzeitanwendungen meist nicht relevant.
- Die Minimierung der Gesamtabschlusszeit ist ebenfalls bei Echtzeitanwendungen selten zielführend.
- Die gewichtete Summe der Abschlusszeiten ist bei Echtzeitanwendungen relevant, falls einzelne Tasks wichtiger für das System sind als andere.
- Maximale Verspätung kann bei Echtzeitanwendungen in der Entwurfsphase relevant sein.



Beispiel Schedule



$$L_{\text{max}} = 23$$

Beispiel Schedule

- Minimierung der maximalen Verspätung minimiert nicht generell die maximale Zahl verspäteter Tasks!
- Plan (a) minimiert die maximale Verspätung, aber alle Tasks sind verspätet, d. h. überschreiten ihre Frist.
- Plan (b) hat eine größere maximale Verspätung, aber vier von fünf Tasks sind vor ihrer Frist beendet.

Lerneinheit 4. Echtzeitplanung

- Lernziele dieser Vorlesung
- ② Grundbegriffe
 - Annahmen und Bezeichnungen
 - Anforderungen
 - Präzedenzen
- 3 Planer
 - Pläne
 - Strategien
- Eigenschaften von Plänen
 - Bewertung
 - Anomalien
- 5 Zusammenfassung



Satz:

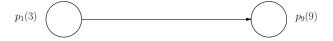
If a task set optimally scheduled on a multiprocessor with some priority assignment, a fixed number of processors, fixed execution times, and precedence constraints, then increasing the number of processors, reducing execution times, or weakening the precedence constraints can increase the schedule length.

R.L. Graham. Bounds in the performance of scheduling algorithms. In *Computer and Job Scheduling Theory*, pp. 165–227. John Wiley and Sons, 1976

 $p_{7}(4)$

 $p_5(4)$

Scheduling Anomalien





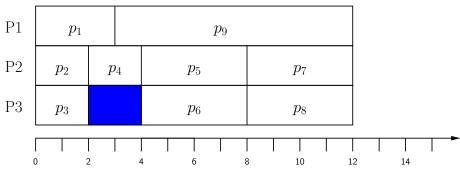
$$\operatorname{prio}(p_i) < \operatorname{prio}(p_j) \forall i < j$$

Zahlen in Klammern: $a(p_i)$

$$p_4(2)$$
 $p_6(4)$

◆ロト ◆団ト ◆豆ト ◆豆 → のへで

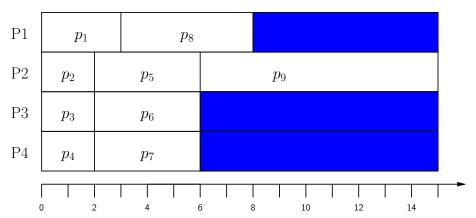
 $p_3(2)$



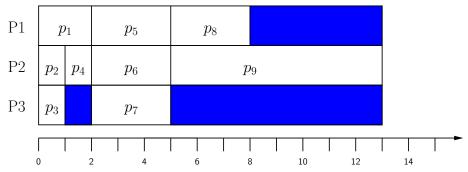
Optimaler Plan für drei Prozessoren

- Bedeutung prio (p_i) : kleiner Wert entspricht hoher Priorität.
- Einplanung: Prozesse hoher Priorität vor Prozessen niedriger Priorität.
- Einplanung: Präzedenzen einhalten.

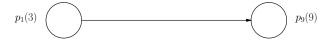




Plan für vier Prozessoren: Erhöhung der Gesamtabschlusszeit um drei Einheiten!



Plan für drei Prozessoren mit Ausführungszeiten jedes Tasks um eine Einheit reduziert: Erhöhung der Gesamtabschlusszeit um eine Einheit!





$$\operatorname{prio}(p_i) < \operatorname{prio}(p_j) \forall i < j$$

Zahlen in Klammern: $a(p_i)$

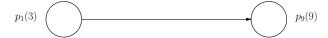
$$p_4(2)$$
 $p_6(4)$



 $p_{7}(4)$



 $p_3(2)$



 $p_8(4)$

$$\operatorname{prio}(p_i) < \operatorname{prio}(p_j) \forall i < j$$

Zahlen in Klammern: $a(p_i)$

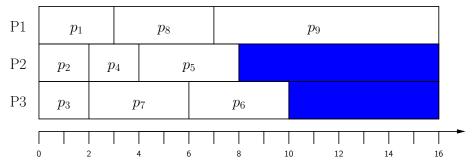
$$p_3(2)$$
 p_4 p_4

$$p_4 \prec p_7 p_4 \prec p_8$$

$$p_{7}(4)$$

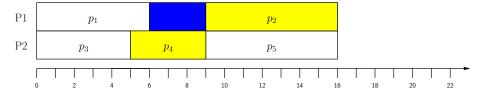
 $p_5(4)$

$$p_4(2)$$
 $p_6(4)$



Plan für drei Prozessoren mit entfernten Präzedenzen: Erhöhung der Gesamtabschlusszeit um vier Einheiten!

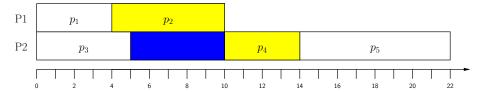
Scheduling Anomalien bei exklusiv genutzten Betriebsmitteln



- Prozesse p_2 und p_4 benutzen ein Betriebsmittel, das nur exklusiv genutzt werden kann..
- Präzedenzen: $p_1 \prec p_2, p_4 \prec p_5$.



Scheduling Anomalien bei exklusiv genutzten Betriebsmitteln



- Prozesse p_2 und p_4 benutzen ein Betriebsmittel, das nur exklusiv genutzt werden kann.
- Reduktion der Ausführungszeit von p₁ um zwei Einheiten: Erhöhung der Gesamtabschlusszeit um fünf Einheiten!

Lerneinheit 4. Echtzeitplanung

- Lernziele dieser Vorlesung
- Grundbegriffe
 - Annahmen und Bezeichnungen
 - Anforderungen
 - Präzedenzen
- 3 Planer
 - Pläne
 - Strategien
- 4 Eigenschaften von Plänen
 - Bewertung
 - Anomalien
- Zusammenfassung



Zusammenfassung

- Ablaufplanung ist ein zentraler Dienst von Echtzeitbetriebssystemen.
- Wichtige Größen für die Echtzeitplanung sind Bereitzeitpunkt, Frist und Gesamtausführungszeit.
- Planung erfolgt häufig in Präzedenzsystemen.
- Planung kann statisch und dynamisch erfolgen.
- Prozessbearbeitung kann präemptiv und nicht-präemptiv erfolgen.
- Performanz von Planungsalgorithmen wird durch Kostenfunktionen auf der Prozess-/Task-Menge bestimmt.