NTG

Fachausschuß 6
"Technische Informatik"

Fachausschuß 9
"Vermittlungstechnik"

DISKUSSIONSSITZUNG

«Steuerung in Datenverarbeitungs- und Vermittlungssystemen»

-Kurzfassungen der Vorträge-

23.-25 Oktober 1978

Baden-Baden

B. Walke, AEG-TELEFUNKEN, Forschungsinstitut Ulm, Postfach 1730, 7900 Ulm

Zur Leistungsbestimmung von Realzeitrechnern: Verkehrsmodelle und Abfertigungsstrategien (Zusammenfassung*)

Nach einer Einführung in die Modellbildung, der Definition charakteristischer Größen (Wartezeit-, Verweildauer-Momente, Durchsatz, usw.), von statischen und dynamischen Prioritäten (unterbrechend/nicht-unterbrechend) und Abfertigungsstrategien, werden Optimierungsziele für Realzeitrechner genannt. Solche Ziele fordern kleine bzw. günstige Werte für charakteristische Größen oder davon abgeleitete Werte. Es wird zwischen ereignis- und taktgesteuertem Realzeitrechner-Betrieb unterschieden. Auf einen kurzen Überblick über übliche Verfahren zur Leistungsbestimmung mit Hilfe von Modellen folgt die Besprechung von vier neueren Modellen für den ereignisgesteuerten Betrieb mit Poisson-Ankunftsprozessen und beliebigen Bediendauerverteilungen. Anforderungen mit gemeinsamem Ankunftsprozeß (Rate λ_i) und gemeinsamer Bediendauerverteilung (Erwartungswert β_i) werden als Typ von Anforderungen bezeichnet. In den Modellen sind verschiedene Typen von Anforderungen gleichzeitig zugelassen.

1. Terminorientierte Prioritäten

Im Modell nach Bild 3 gibt es N Typen von Anforderungen. Jeder Typ hat eine Wichtigkeit ω_i, die man als freiwillige Wartezeit bis zum Bearbeitungsbeginn ansehen kann. Alle Prioritäten sind nicht-unterbrechend. Die dynamische Priorität einer Anforderung vom Typ i ist bei der Ankunft gleich ω_i und fällt linear mit der individuellen Wartezeit. Nach Fertigstellung einer Anforderung durch den Prozessor wird als nächste die mit höchster Priorität ausgewählt. Ist die Varianz Var (ω) = 0, dann wird die Strategie FCFS (first come first serve) approximiert.

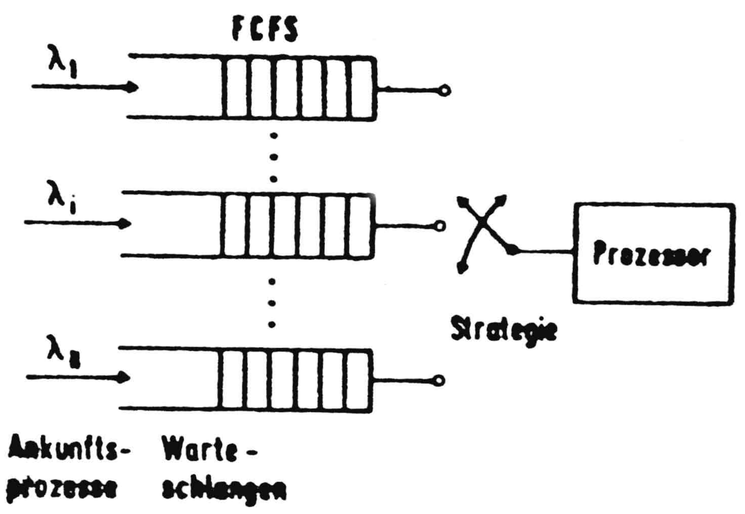

Bild 3: N Typen von Anforderungen.

*) Beitrag zur FA6/FA9 Diskussionssitzung der NTG: Steuerung in Datenverarbeitungs- und Vermittlungssystemen, 23. bis 25. Oktober 1978, Baden-Baden.

AEG TFK Vd 7923 500 000/876 Sm

Bei Var (ω)$\rightarrow\infty$ werden statische Prioritäten angenähert. Durch Wahl von Var (ω) lassen sich die charakteristischen Größen des Modells beliebig zwischen den durch diese beiden Grenzfälle gegebenen Möglichkeiten verändern. Die Bilder 5 und 6 zeigen typische Ergebnisse für einen typischen Fall Var (ω) = 125.

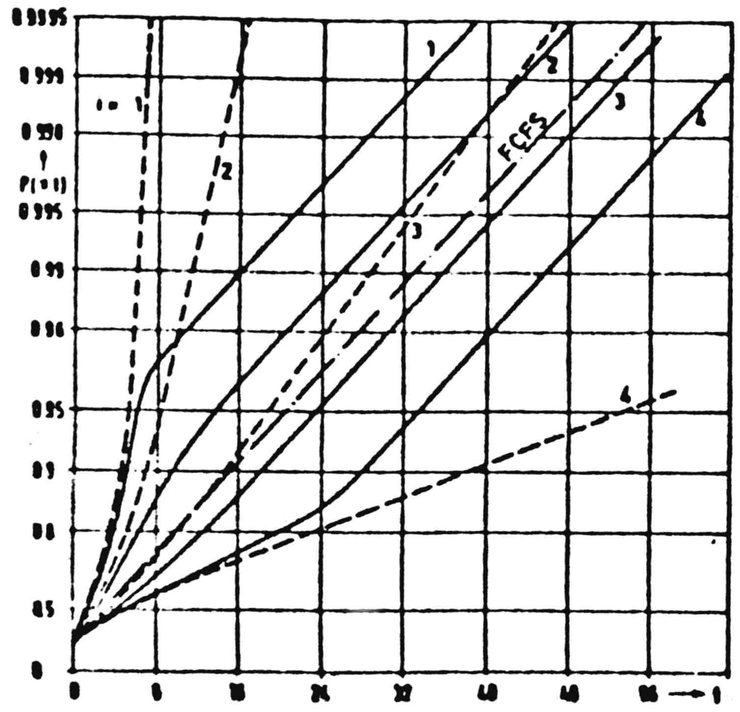

Bild 6: Mittlere Wartezeit W_i für FCFS (gestrichelt) terminorientiert (strichpunkt) und statische Prioritäten (ausgezogen).

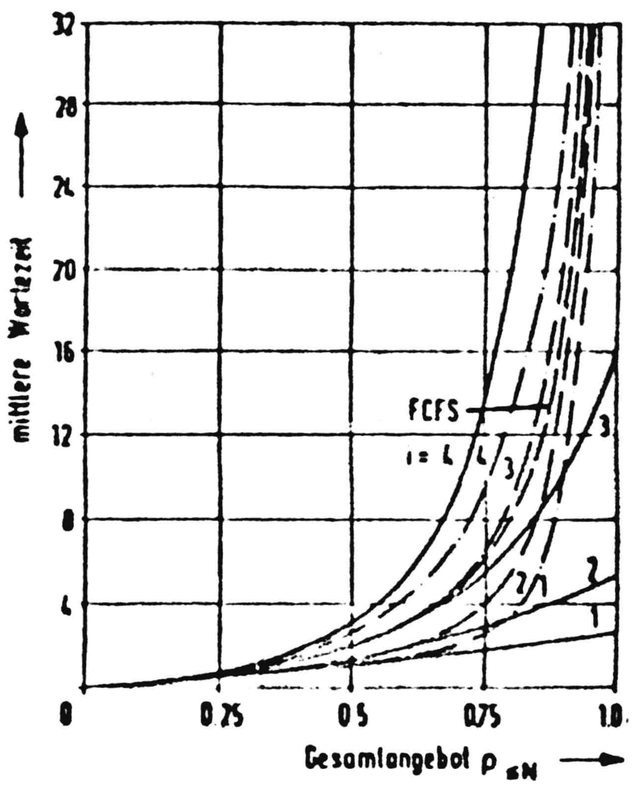

Bild 5: Wartezeitverteilungsfunktionen bei statischen (gestrichelt) und terminorientierten (ausgezogen) Prioritäten.

2. Wartezeitoptimale Bearbeitung eines Typs von Anforderungen mit individueller statischer Priorität

Bild 8 zeigt das Modell, wobei in einer beliebigen Prioritätsebene i (mit unterbrechender Prirorität gegenüber Ebenen j > i) eine wartezeitoptimale Strategie mit Zeitscheiben angewandt wird. Die mittleren Wartezeiten der anderen Ebenen bleiben dabei unverändert wie die Rechnung zeigt. Streut die Bediendauerverteilung in der Ebene i nicht mehr als eine negativ exponentielle Verteilung, dann ist dort die Strategie FCFS optimal. Andernfalls muß dort ein Zeitscheibenverfahren mit Rückspeisung (feedback) in ein Warteschlangensystem innerhalb der Ebene i angewandt werden.

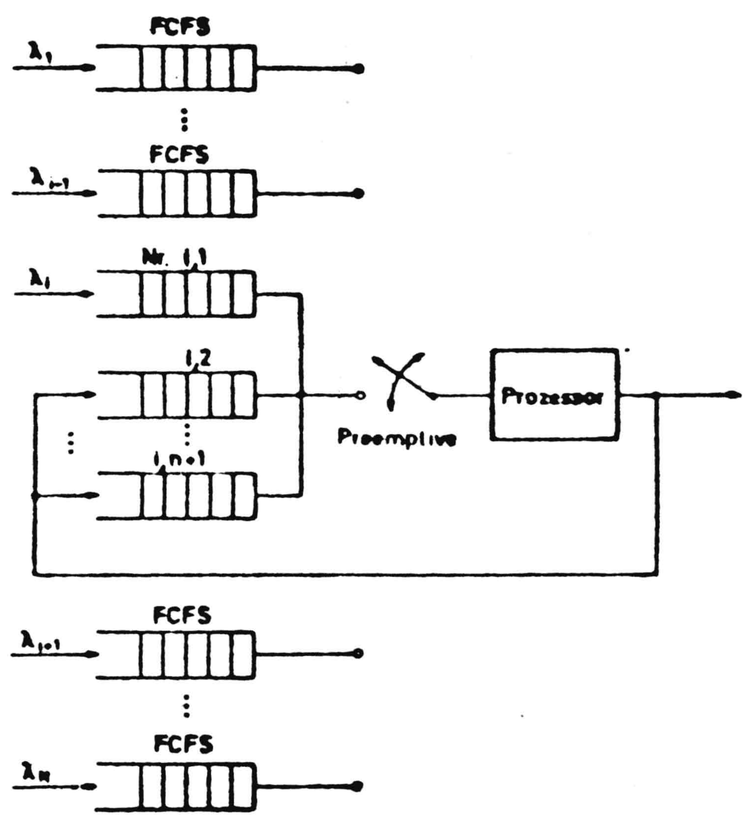

Bild 8: N Prioritätsebenen, je Ebene 1 Typ von Anforderungen und optimale Strategie.

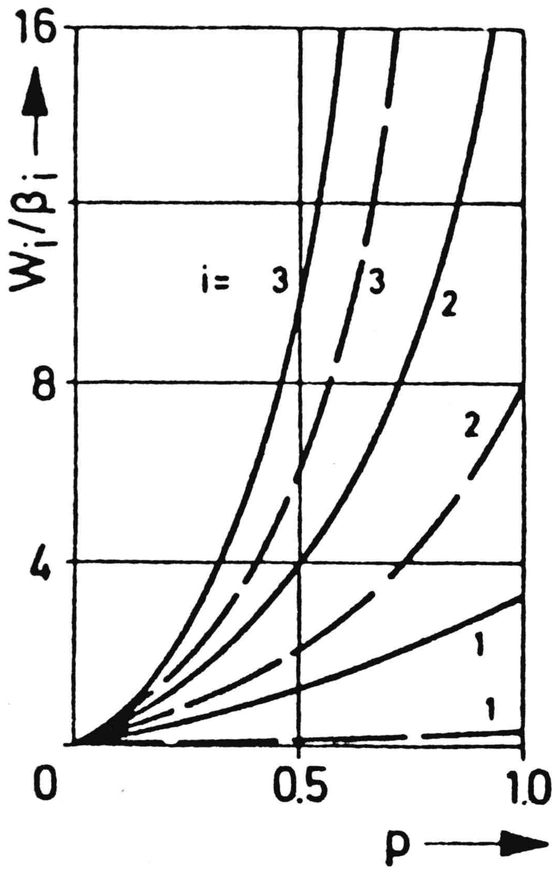

Bild 9: Bezogene mitttlere Wartezeit über dem Gesamtangebot. FCFS (ausgezogen) und optimale Strategie (gestrichelt) innerhalb jeder Ebene.

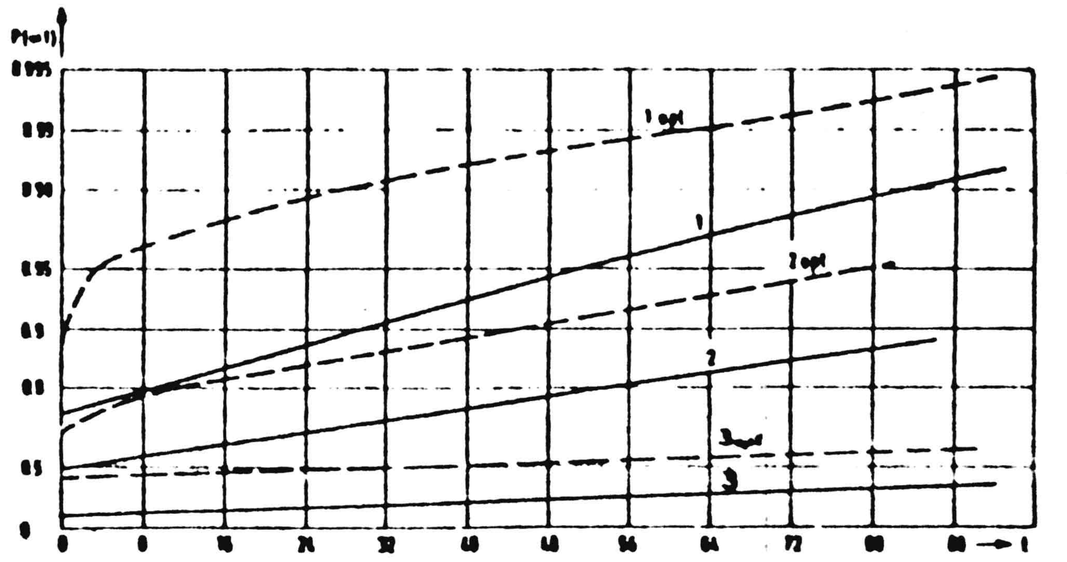

Bild 10: Wartezeitverteilungsfunktionen zum Beispiel von Bild 9.

Die Bilder 9 und 10 zeigen Ergebnisse für den Fall, daß die Bediendauerverteilungen in N=3 Ebenen gleich sind und stark streuen, und zwar den Vergleich zwischen optimaler Abfertigung (gestrichelt) und FCFS (ausgezogen) in jeder Ebene.

3. Wartezeitoptimale Bearbeitung mehrerer Typen von Anforderungen mit gleicher statischer Priorität

Bild 11 zeigt das Modell mit R Typen von Anforderungen in der Prioritätsebene i (je kleine Warteschlange ein Typ). Optimiert wird die gemeinsame mittlere Wartezeit $W_{i,\leq R}$ aller Typen von Anforderungen in der Ebene i. Nach Vorstellung von Regeln für eine diesbezüglich optimale Vergabe von statischen Prioritäten und den zugehörigen Voraussetzungen wird ein Beispiel behandelt, wo diese Voraussetzungen verletzt sind. Dafür wird das Resultat bei statischen bzw. optimalen dynamischen Prioritäten verglichen, Bilder 12 und 13. Auch in diesem Modell werden die mittleren Wartezeiten anderer Ebenen $j \neq i$ nicht durch die Strategie in der Ebene i betroffen.

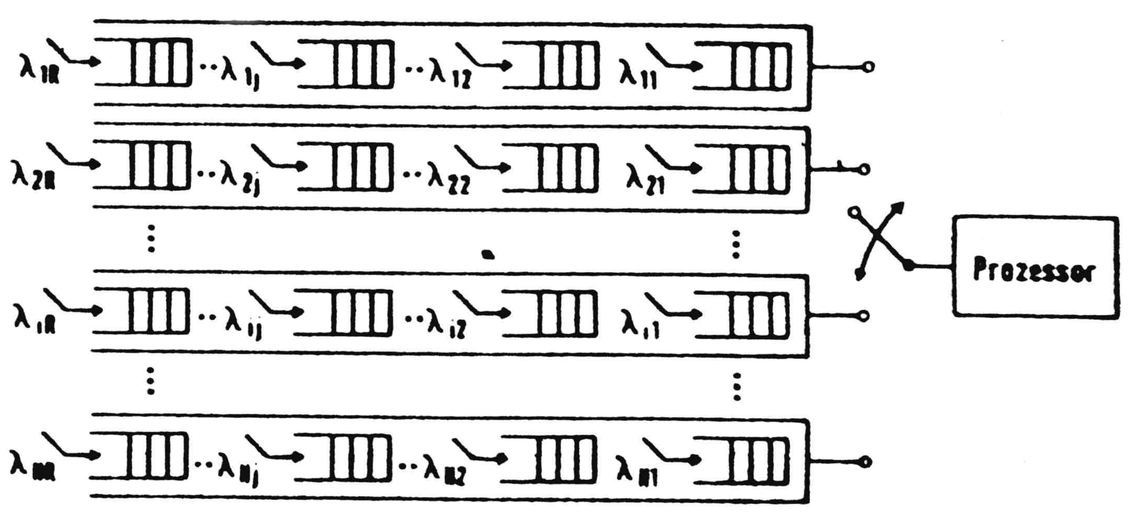

Bild 11: N Prioritätsebenen und R Typen von Anforderungen je Ebene.

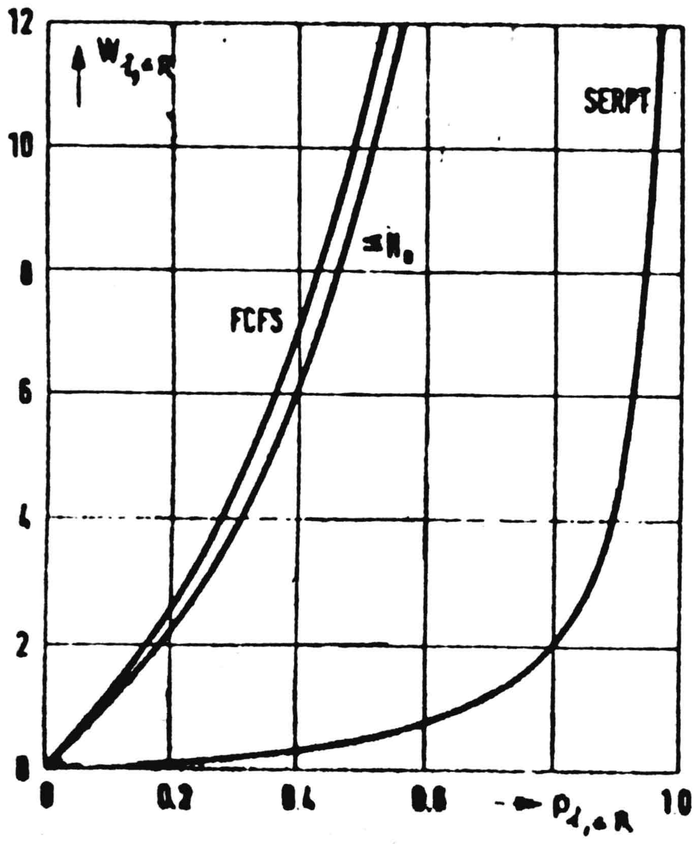

Bild 12: "optimale" statische ($\leq N_0$) und optimale dynamische Prioritäten (SERPT), sowie FCFS.

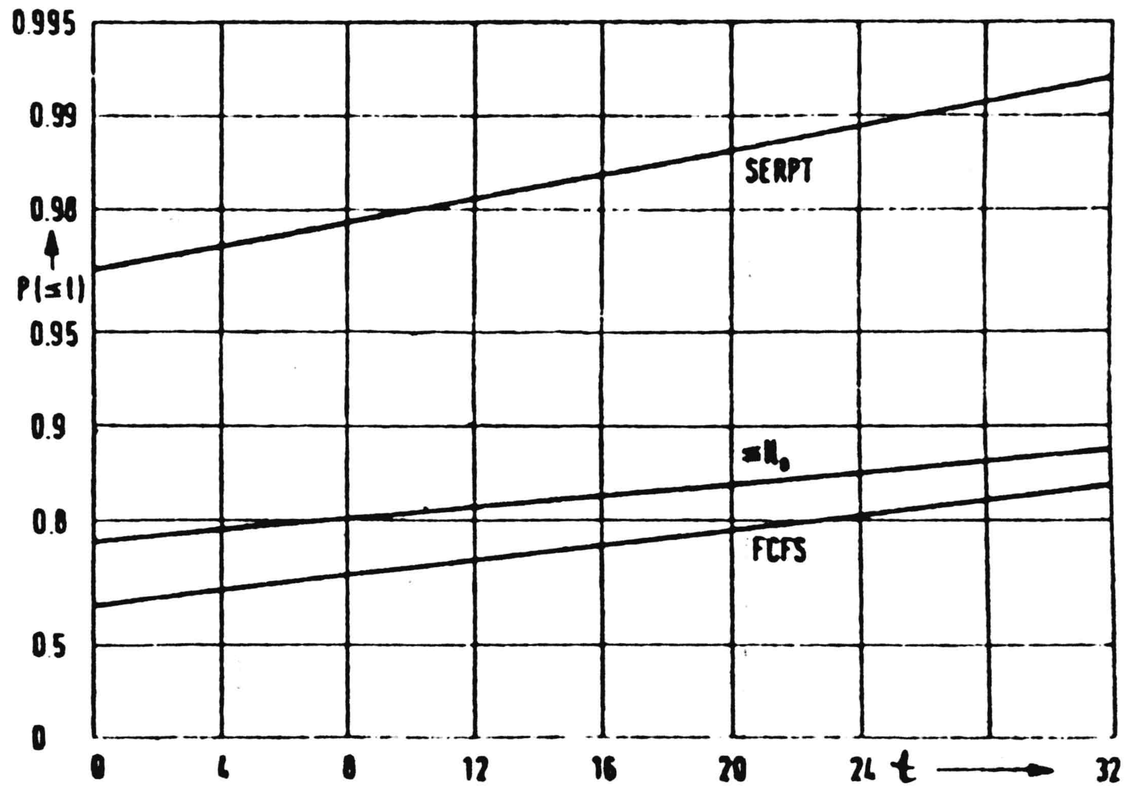

Bild 13: Wartezeitverteilungsfunktionen zu Bild 12. Gesamtangebot $\rho_{i,\leq R} = 0.55$.

4. Unterbrechungs-Distanz - Prioritäten mit Verwaltungszeit

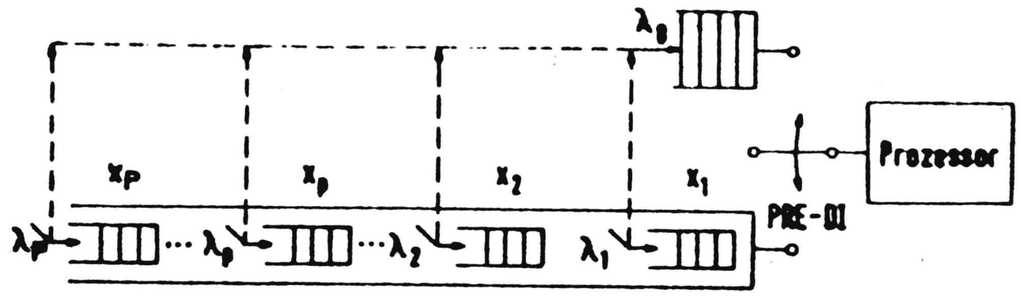

Bild 15: Unterbrechungsdistanz Prioritäten mit Verwaltungszeit für Interrupts.

Bild 15 zeigt das Modell. Für Anforderungen vom Typ p ist vorgebbar, welche Typen p-k, (k=1,2,...,r) nicht-unterbrechende statische Priorität haben sollen, bevor alle noch wichtigeren p-r-j, (j=1,2...) unterbrechende Priorität haben. Die "Unterbrechungsdistanz" r kann für jeden Typ von Anforderungen individuell verschieden sein. Die Ankunft jeder Anforderung wird durch einen Interrupt mit beliebiger Bediendauerverteilung und dem Erwartungswert β_0 ausgelöst. Interrupts unterbrechen sofort alle anderen Anforderungen. Die mittleren Warte- und Verweilzeiten sind für jeden Typ p berechenbar mit β_0 als Parameter.

5. Periodische Anforderungen

Für den taktgesteuerten Betrieb wird eine Regel angegeben, wie bei bekannten Periodendauern T_i von Anforderungen vom Typ i statische unterbrechende Prioritäten zugewiesen werden müssen, damit alle Anforderungen jeweils vor Beginn der nächsten Periode fertig werden.

Bei beliebigen Verhältnissen der Periodendauern zueinander ist die maximale zugelassene Prozessorbelastung bekannt.

Ein Beispiel, Bild 16, zeigt, wie mit terminorientierter Abfertigung auch in diesem Fall alle Anforderungen rechtzeitig bedient werden können, wobei der Prozessor voll ausgelastet sein darf.

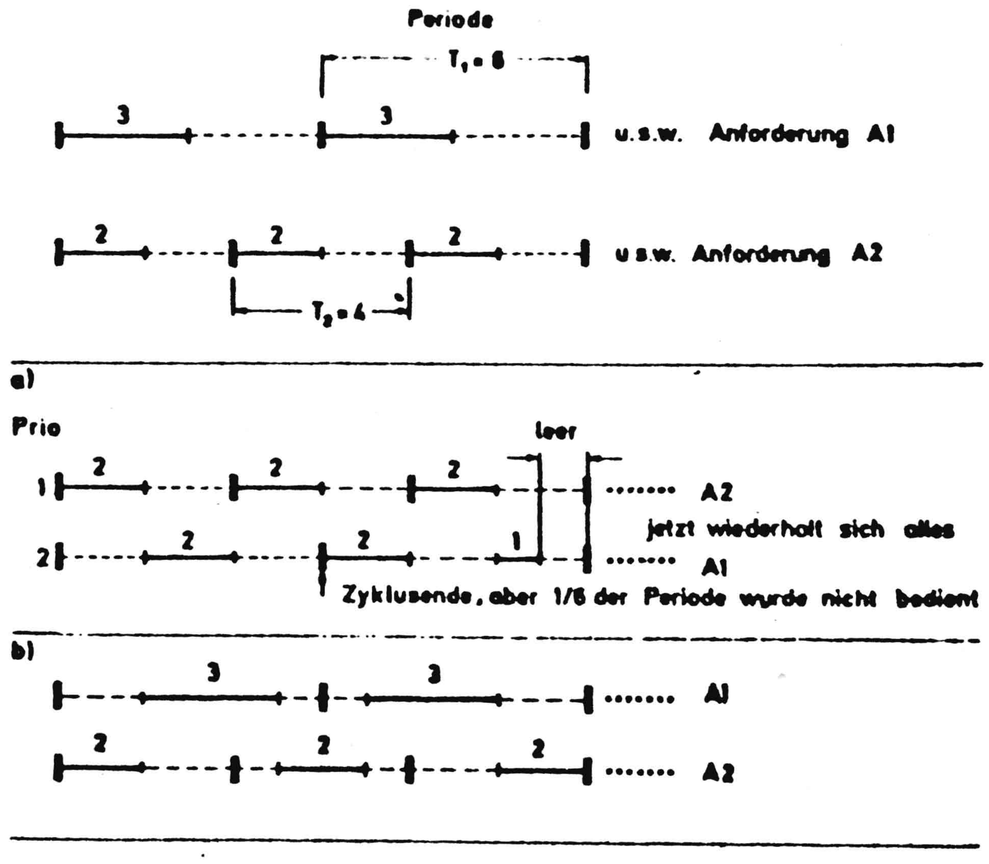

Bild 16: N=2 periodische Anforderungen, kritische Situation a) opt. statische, b) terminorientierte Prioritäten.

6. Wertung

Die Anwendbarkeit der Modelle zur Beschreibung auch viel komplizierterer Abläufe bei Realzeitrechnern wird erläutert. Möglichkeiten und Grenzen werden gezeigt.

Literatur: B. Walke, Realzeitrechner-Modelle, Theorie und Anwendung, R. Oldenbourg -Verlag, Reihe Datenverarbeitung, München 1978.