2. Probabilistische Modelle

- **Modellannahmen**

  Zu zufälligen Zeiten A_i (Ankunftszeit) treffen Anforderungen an den Scheduler in zufälligem Umfang B_i (geforderter Bedarf nach dem betrachteten Betriebsmittel, Bedienungszeit) ein.

  Anwendungsbereich: interaktiver Betrieb, verteilte Systeme

- **Mathematisches Instrumentarium**

  Bedienungstheorie

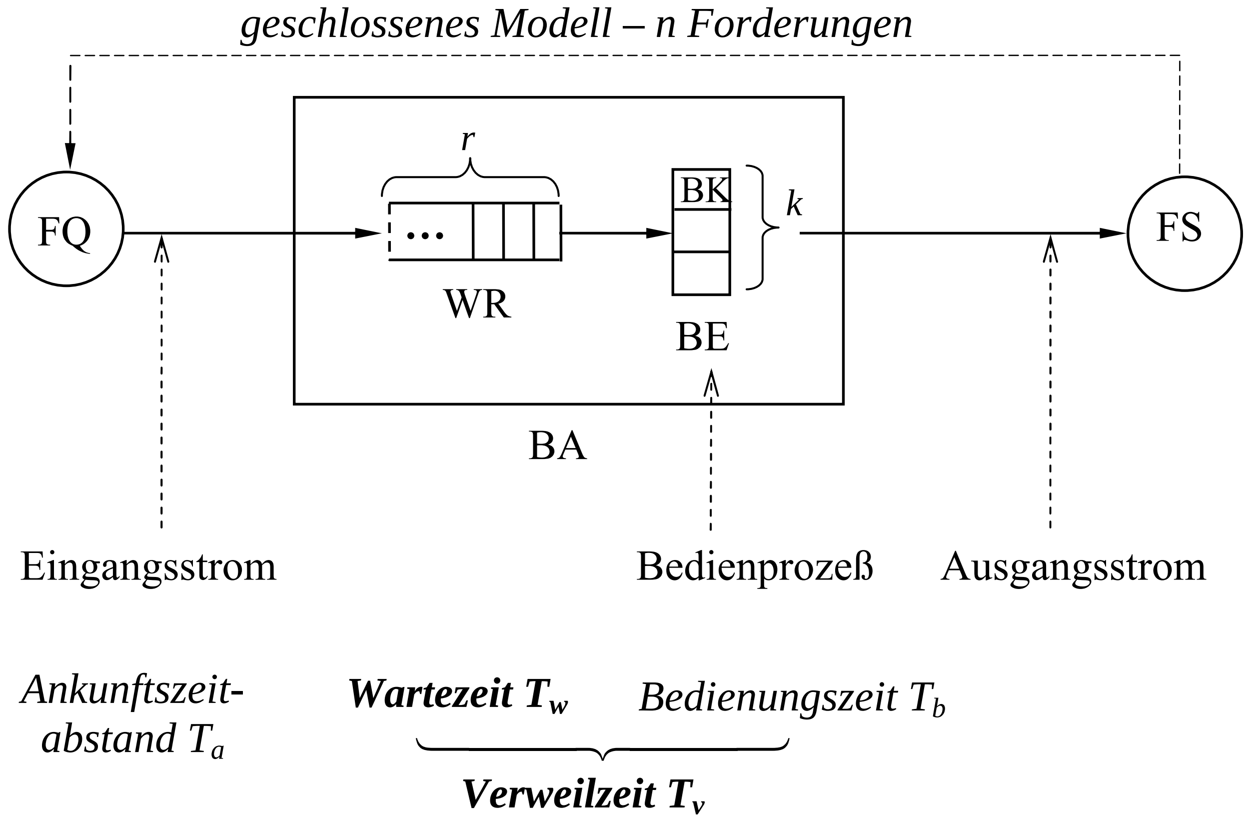
  Queueing (systems) theory

  Теория массого обслуживания

- **Historie**

1908	ERLANG	Telegraphenamt Kopenhagen
1910	PALM	Mehrmaschinen-Bedienung
1940 ... 1970	KANTOROWITSCH, GNEDENKO, KOWALENKO	
1965 ... 1980	KLEINROCK, COFFMAN, DENNING	

2.1. Bedienungssysteme – Grundlagen



BA Bedienungsanlage

WR Warteraum

BE Bedienungseinrichtung

BK Bedienungskanal

FQ Forderungsquelle

FS Forderungssenke

- **Strukturbeschreibung**

– ***Größe r des Warteraums***

$r = 0$	(reines) Verlustsystem
$0 < r < \infty$	(gemischtes) Warte-Verlust-System
$r = \infty$	(reines) Wartesystem

– ***Struktur der Bedienungseinrichtung***

Einkanalbedienung

Mehrkanalbedienung

Mehrphasenbedienung

– ***Bedienungsstrategie (Warteschlangendisziplin)***

FIFO, LIFO

RANDOM

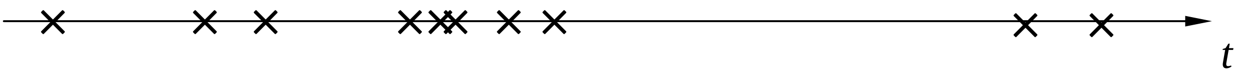
Prioritäten (absolut/relativ, statisch/dynamisch)

z. B. SPT, LJN, HRN, ...

– ***Offenes/geschlossenes Modell***

- **Lastbeschreibung**

– ***Ankunftsprozeß***



N_a zufällige Anzahl der pro Zeiteinheit eintreffenden Forderungen

T_a zufälliger Abstand zwischen dem Eintreffen zweier Forderungen

oft: POISSONscher Ankunftsprozeß mit dem Parameter $\lambda > 0$:

$$\boldsymbol{P}(N_a = i) \;=\; \frac{\lambda^i}{i!} \cdot \mathrm{e}^{-\lambda} \quad i \in \mathbb{N}$$ POISSON-Verteilung

$$\boldsymbol{P}(T_a \le t) \;=\; \begin{cases} 1 - \mathrm{e}^{-\lambda t} & t \ge 0 \\ 0 & t < 0 \end{cases}$$ Exponential-Verteilung

Es gilt:

$$\boxed{\lambda = \boldsymbol{E}(N_a) = \frac{1}{\boldsymbol{E}(T_a)}}$$ Eingangsstromintensität, Ankunftsrate

– ***Bedienprozeß***

T_b zufällige Dauer der Bedienung einer Forderung

oft: exponentiell verteilte Bedienzeit, Bedienrate μ

– *Weitere wichtige Verteilungen*

ERLANG-Verteilung k-ter Ordnung, $k \in \mathbb{N}^+$

$$f(t) = \frac{\alpha^k t^{k-1}}{(k-1)!} \cdot e^{-\alpha t} \qquad (t \geq 0)$$

Hyperexponentialverteilung

$$f(t) = \sum_{i=1}^{k} p_i \cdot \alpha_i e^{-\alpha_i t} \qquad p_i \in [0,\ 1], \quad \sum p_i = 1$$

Hypoexponentialverteilung

$$f(t) = \frac{\mu_1 \mu_2}{\mu_1 - \mu_2}\left(e^{-\mu_1 t} - e^{-\mu_2 t}\right) \qquad \mu_1, \mu_2 > 0, \quad \mu_1 \neq \mu_2$$

- **Zusammengefaßte Beschreibung**

– ***KENDALLsche Notation***

A/B/k/r

Warteraumkapazität, $r \in \mathbb{N} \cup \{\infty\}$

Anzahl der Bedienungskanäle, $k \in \mathbb{N}^+$

Bedienprozeß

Ankunftsprozeß

Wichtige Werte für A, B:

M	POISSON-Prozeß
E_k	ERLANG-Verteilung k-ter Ordnung
D	Deterministische Verteilung
G	Allgemeiner Fall

– ***Bedienungsstrategie!***

– ***Geschlossenes Modell***

n: Anzahl der Forderungen im System (statt r)

- **Bewertungsgrößen für offene Bedienungssysteme GI/G/*k*/*r***

– ***Unmittelbar abgeleitete Größen***

$\boldsymbol{E}(T_a)$ mittlerer Ankunftszeitabstand

$\boldsymbol{E}(T_b)$ mittlere Bedienzeit

$\lambda = \dfrac{1}{\boldsymbol{E}(T_a)}$ Ankunftsrate

$\mu = \dfrac{1}{\boldsymbol{E}(T_b)}$ Bedienrate

$$\boxed{\rho = \frac{\lambda}{\mu} = \frac{\boldsymbol{E}(T_b)}{\boldsymbol{E}(T_a)}}$$ Verkehrswert (Angebot, **Be**lastung)

Forderung: $\rho < k$ für nicht-kritischen Zustand

– ***Bewertungsgrößen i.e.S.***

N_w	Warteschlangenlänge
N_b	Anzahl der aktuell bedienten Forderungen
$N_v = N_w + N_b$	Anzahl der im System befindlichen Forderungen
T_w	Wartezeit
$T_v = T_w + T_b$	Verweilzeit

Für GI/G/k/∞-Systeme gilt: $\boxed{\boldsymbol{E}(T_v) = \boldsymbol{E}(N_v)\cdot\boldsymbol{E}(T_a)}$ (LITTLE)

Ferner:

p_0	Stillstands-, Leerwahrscheinlichkeit
p_V	Verlustwahrscheinlichkeit
η	Auslastung der Bedieneinrichtung

Grundlage: Zustandswahrscheinlichkeiten

$$\boxed{p_i = \boldsymbol{P}(N_v = i), \quad i \in \mathbb{N}}$$

– ***Allgemeines Vorgehen***

Ermitteln des Zustandsgraphen

Aufstellen der Zustandsgleichungen

Lösen des Gleichungssystems

Berechnen der Zustandswahrscheinlichkeiten

Berechnen der Bewertungsgrößen

2.3. Ausgewählte Standardsysteme

- **M/M/1/∞-System**

– ***Systembeschreibung***

Eingangsstrom: POISSON-Prozeß, Intensität λ

Bedienungszeit: exponentiell verteilt, Parameter μ

Bedienungsstrategie: FIFO, LIFO, gleichverteilt

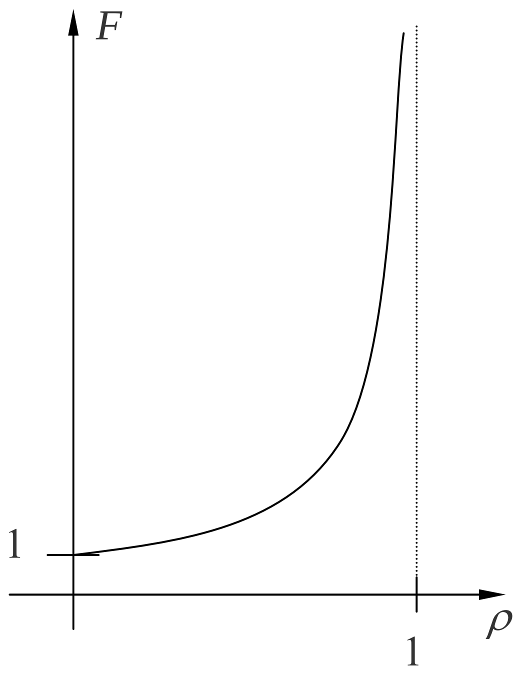
Verkehrswert $\rho = \frac{\lambda}{\mu} < 1$!

– ***Zustandswahrscheinlichkeiten***

$p_i = \rho^i \cdot (1 - \rho) \qquad i = 0,1,...$

– ***Systemverhalten***

Verweilzeitfaktor $\boxed{F = \frac{\boldsymbol{E}(T_v)}{\boldsymbol{E}(T_b)} = \frac{1}{1-\rho}}$



ρ	0,4	$\rightarrow$	0,44
F	1,67	$\rightarrow$	1,79

ρ	0,9	$\rightarrow$	0,94
F	10	$\rightarrow$	16,7

ρ	0,9	$\rightarrow$	0,99
F		$\rightarrow$	

- **M/G/1/∞-System**

	M/M/1/∞	**M/G/1/∞**
$\mathrm{E}(T_w)$ =	$\dfrac{\rho}{\mu(1-\rho)}$	$\dfrac{\lambda\cdot \mathrm{E}(T_b^2)}{2(1-\rho)}$
$\mathrm{E}(T_v)$ = $\mathrm{E}(T_w)$+ $\mathrm{E}(T_b)$ =	$\dfrac{1}{\mu(1-\rho)}$	$\dfrac{\lambda\cdot \mathrm{E}(T_b^2)}{2(1-\rho)}+\mathrm{E}(T_b)$
$\mathrm{E}(N_w)$ = $\lambda\cdot\mathrm{E}(T_w)$ =	$\dfrac{\rho^2}{1-\rho}$	$\dfrac{\lambda^2\cdot \mathrm{E}(T_b^2)}{2(1-\rho)}$
$\mathrm{E}(N_v)$ = $\mathrm{E}(N_w)$ + $\mathrm{E}(N_b)$ =	$\dfrac{\rho}{1-\rho}$	$\dfrac{\lambda^2\cdot \mathrm{E}(T_b^2)}{2(1-\rho)}+\rho$

$$\mathrm{E}(T_w^{\,2}) = \begin{cases} \dfrac{2\rho}{\mu^2(1-\rho)^2} & \dfrac{\lambda\cdot \mathrm{E}(T_b^2)}{3(1-\rho)}+\dfrac{(\lambda\cdot \mathrm{E}(T_b^2))^2}{1-\rho} & \text{FIFO} \\[2ex] \dfrac{2\rho}{\mu^2(1-\rho)^3} & \mathrm{E}(T_{w,FIFO}^2)\cdot\dfrac{1}{1-\rho} & \text{LIFO} \\[2ex] \dfrac{2\rho}{\mu^2(1-\rho)^2(1-\frac{\rho}{2})} & \mathrm{E}(T_{w,FIFO}^2)\cdot\dfrac{1}{1-\frac{\rho}{2}} & \text{Random} \end{cases}$$

Wartezeit-Verteilung für M/M/1/∞-FIFO:

$$F_{T_w}(t) = \mathbf{P}(T_w \le t) = 1 - \rho\cdot e^{-(\mu-\lambda)\mathrm{t}} \qquad (t \ge 0)$$

- **M/M/*k*/∞-System**

$$p_0 = \left[\sum_{i=0}^{k-1} \frac{\rho^i}{i!} + \frac{\rho^k}{(k-1)!(k-\rho)} \right]^{-1}$$

$$\boldsymbol{E}(N_w) = \frac{\rho^{k+1}}{(k-1)!(k-\rho)^2} \cdot p_0 \qquad \boldsymbol{E}(N_b) = \rho \qquad \eta = \frac{\rho}{k}$$

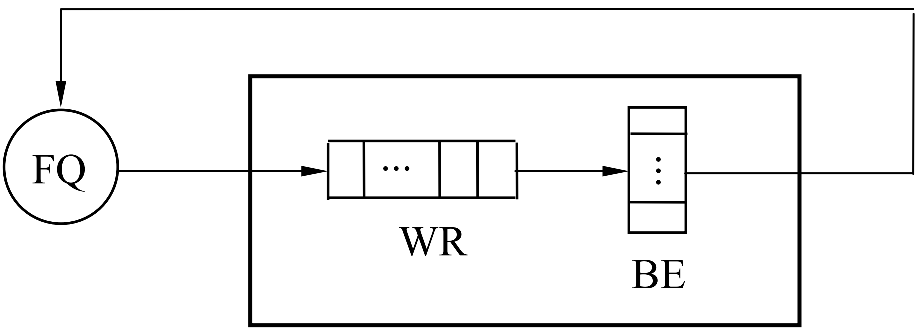
- **M/G/*k*/0-System** ERLANG 1908, СЕВАСТЬЯНОВ 1957

$$p_0 = \left[\sum_{i=0}^{k} \frac{\rho^i}{i!} \right]^{-1}$$

$$p_i = \frac{\rho^i}{i!} \cdot p_0 \quad \text{bzw.} \quad p_i = \frac{\rho}{i} p_{i-1} \qquad i = 1,\ldots,k$$

$$\boxed{p_V = p_k = \frac{\rho^k}{k!} \cdot p_0 \qquad \boldsymbol{E}(N_b) = \boldsymbol{E}(N_v) = \rho\cdot(1 - p_V)}$$

- **Geschlossenes M/M/1/*n*-System**



T_q Quellverweilzeit T_b Bedienungszeit

Für $k = 1$ gilt:

$$p_0 = \left[\sum_{i=0}^{n} \frac{n!}{(n-i)!} \cdot \rho^i \right]^{-1} \qquad \text{mit} \qquad \rho = \frac{\mathrm{E}T_b}{\mathrm{E}T_q}$$

$$p_i = \frac{n!}{(n-i)!} \cdot \rho^i p_0, \qquad i = 1,\ldots,n$$

2.4. Ausgewählte Scheduling-Strategien

- **Bewertungsmaß und Strategien**

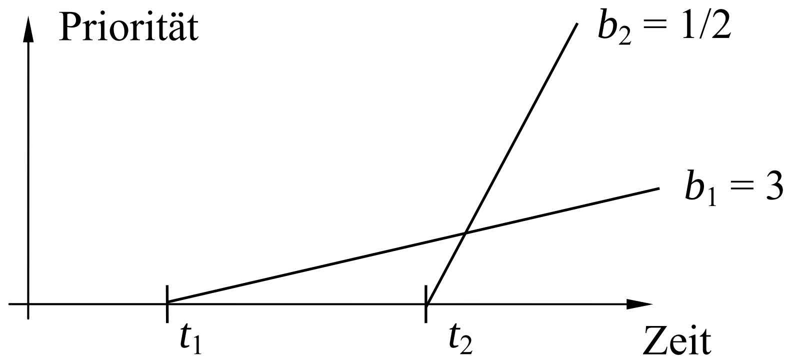
  Abhängigkeit der mittleren Wartezeit $\boldsymbol{E}(T_w)$ vom Bedienzeitwunsch b der Forderungen für die Strategien (ohne Entzug)

– ***FIFO – LIFO – Random***

– ***SJN (SPT) „Shortest Job Next“***

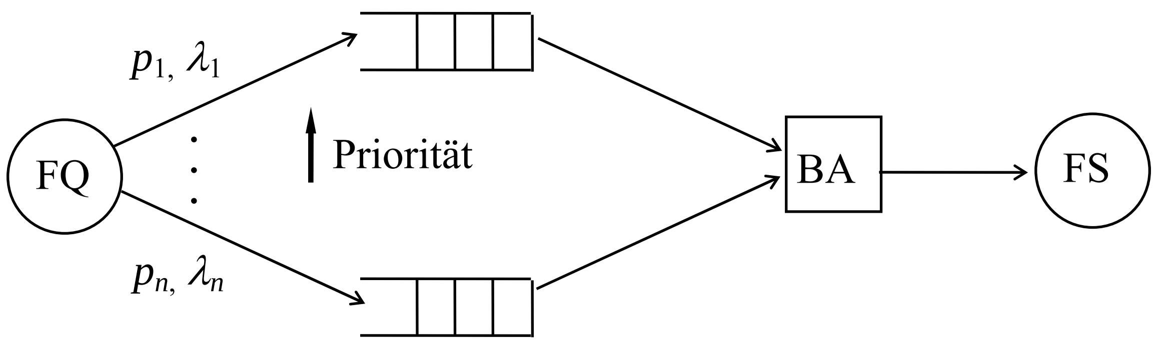
– ***HRN „Highest Response Ratio Next“***

$$\text{Priorität} = \frac{\text{aktuelle Verweilzeit}}{\text{Bedienzeitwunsch}} = \frac{t - t_i}{b_i}$$



– ***FEP „Fixed External Priorities“***

  Aufträge werden in Prioritätsklassen eingeordnet; innerhalb einer Klasse FIFO, Bedienrate klassenunabhängig



- **Basis**

  M/G/1/∞-Modell mit dynamischen Prioritäten

- **Ergebnisse**

– ***SJN***

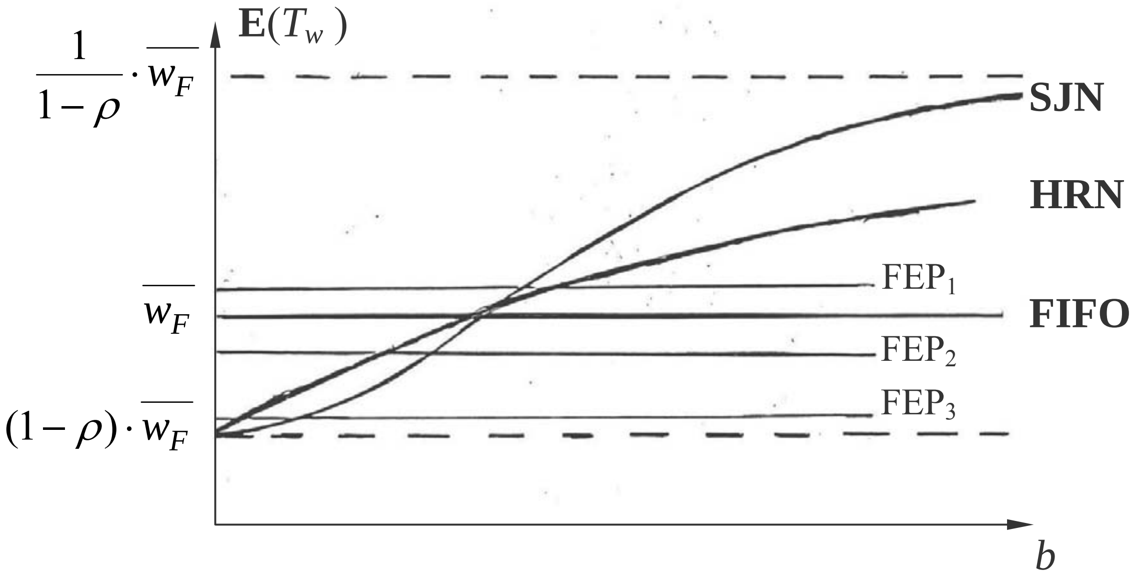
$$\boldsymbol{E}(T_{w,\mathrm{SJN}}) = \frac{\lambda \cdot E(T_b^2)}{2(1-r(b))^2}, \qquad r(b) = \lambda \cdot \int_0^b t \cdot f_{T_b}(t)dt$$

– ***HRN***

$$\boldsymbol{E}(T_{w,\mathrm{HRN}})(b) = \begin{cases} \frac{\lambda}{2} \cdot E(T_b^2) + \frac{b}{2} \cdot \frac{\rho^2}{1-\rho} & \text{für} \quad b \le \frac{E(T_b^2)}{E(T_b)} \\ \frac{\lambda}{2} \cdot \frac{E(T_b^2)}{1-\rho} \cdot \left(1-\rho+\lambda \cdot \frac{E(T_b^2)}{b}\right)^{-1} & \text{für} \quad b > \frac{E(T_b^2)}{E(T_b)} \end{cases}$$

– ***FEP***

$$\boldsymbol{E}(T_{w,\mathrm{FEP}})(i) = \frac{\rho/\mu}{(1-\sigma_i)(1-\sigma_{i-1})} \quad \text{mit} \quad \sigma_0 = 0, \ \ \sigma_i = \sigma_{i\text{-}1} + \rho_i\,, \ \ i = 1,\ldots,n$$



- **Bewertung**

– FIFO: einfach. Gleichbehandlung aller Aufträge.

– SJN: Bevorzugung kürzerer Aufträge.
$\bar{t}_v$ wird bei $R = \emptyset$ minimal, falls alle b bekannt.

– HRN: größere Gerechtigkeit.

– FEP: statisch. Prioritätszuordnung?

Bei allen Strategien wird $\boldsymbol{E}(T_w)$ durch
$(1-\rho)\cdot\boldsymbol{E}(T_{w,FIFO})$ nach unten und durch
$(1-\rho)^{-1}\cdot\boldsymbol{E}(T_{w,FIFO})$ nach oben begrenzt.

– RR: Gerechtigkeit.

- **Operationale Analyse**