

# Úvod do SHO – sítě front



Výkonnost a spolehlivost – KIV/VSP

Richard Lipka

20.10.2015



# Malá odbočka k pravděpodobnosti

- Diskrétní rozdělení (každé pro které máme konkrétní hodnoty se chová jako diskrétní!!!)
  - Jednotlivé hodnoty, každá má nenulovou pravděpodobnost
  - $E\{X\} = \frac{1}{n} \sum_{k=1}^n x_k$ , ale existuje jen několik ( $i$ ) různých  $x_k$  (tříd), každé s četností  $m_k$  a pravděpodobnost výskytu  $x_k$  je tedy  $\frac{m_k}{n} = p_k \rightarrow E\{X\} = \sum_{k=0}^i x_k p_k = \sum_{k=1}^i x_k \frac{m_k}{n}$
- Příklad „vážený vs. normální průměr“:
  - $X = \{1, 1, 1, 2, 2, 3, 3, 3\}$  - tři třídy,  $\{1, 2, 3\}$
  - $p_1 = \frac{3}{8}, p_2 = \frac{2}{8}, p_3 = \frac{3}{8}$
  - $E\{X\} = \frac{1}{8} (1 + 1 + 1 + 2 + 2 + 3 + 3 + 3) = \frac{1}{8} (1 + 1 + 1) + \frac{1}{8} (2 + 2) + \frac{1}{8} (3 + 3 + 3) = 1 \cdot \frac{3}{8} + 2 \cdot \frac{2}{8} + 3 \cdot \frac{3}{8}$



# Malá odbočka k pravděpodobnosti

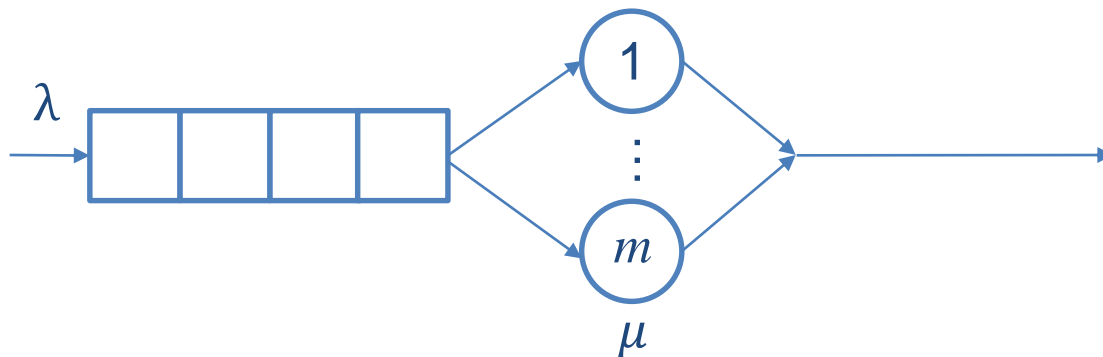
- Spojité rozdělení (teoretické, ne konkrétní hodnoty)
  - Nekonečně mnoho různých hodnot  $x_k$ ,  $m$  výskytů konkrétní hodnoty  $x_k$   
 $\rightarrow p_k = \frac{m}{\infty} \rightarrow 0$ , při konkrétní realizaci typicky  $m \rightarrow 1$  (nebo velmi malé)
  - Pravděpodobnost lze určit jen pro jednotlivé intervaly:  $p_{k,k+1} = \int_k^{k+1} f(x) dx$
- Střední hodnota definována:

$$E\{X\} = \int_{-\infty}^{\infty} x f(x) dx$$

- $x$  – konkrétní hodnota,  $f(x)$  - „pravděpodobnost této hodnoty“  
 $\rightarrow$  podobné jako  $E\{X\} = \sum_{k=0}^i x_k p_k$  ve spojitém světě



# Elementární SHO - opakování



- Koeficient zatížení  $\rho = \frac{1}{m} \cdot \frac{\lambda}{\mu} = \frac{1}{m} \cdot \frac{T_s}{T_a}$ 
  - Stacionární režim:  $\rho < 1$
- Počet požadavků v systému  $L_q \cong \frac{m\rho}{1-\rho^m}$ 
  - Pro  $m \in \{1, 2\}$  přesné, jinak odhad
- Doba odezvy (obsluhy)  $T_q \cong \frac{T_s}{1-\rho^m}$
- Littleovy vztahy:  $L_q = \lambda \cdot T_q, L_w = \lambda \cdot T_w, T_w = L_w \cdot T_a$



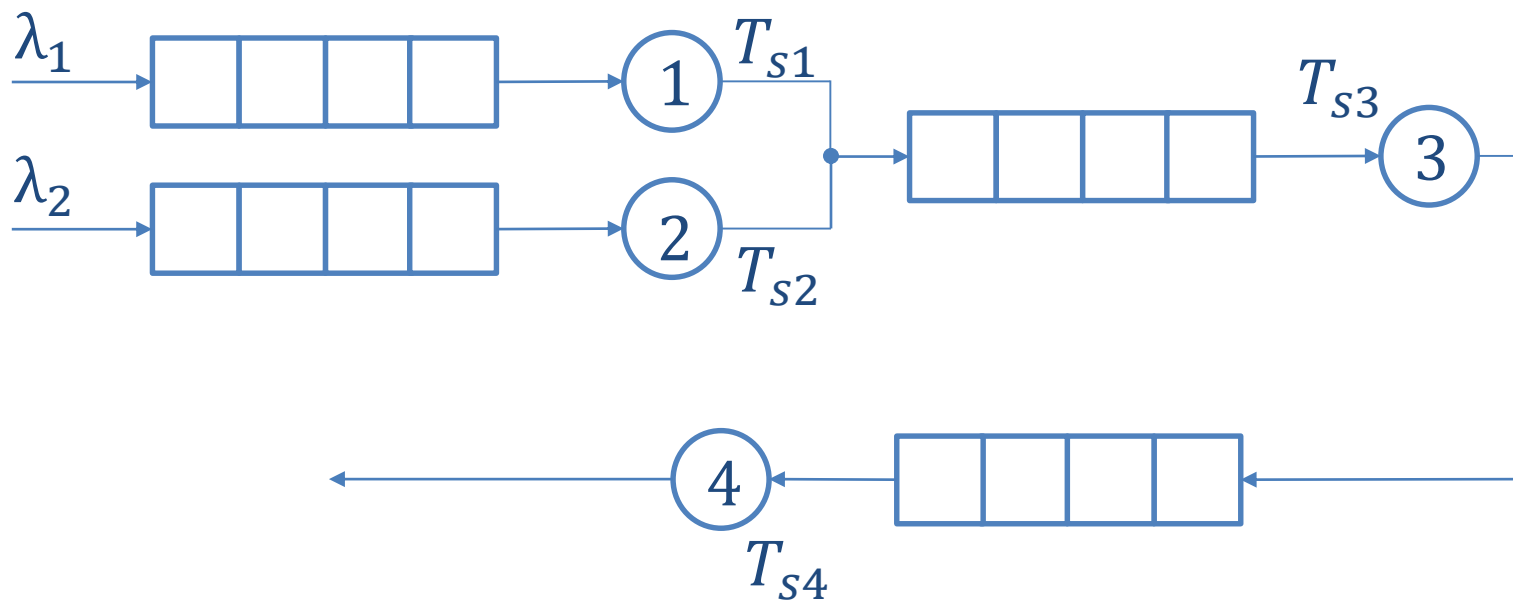
# Sítě SHO



- Obvykle bližší realitě než elementární SHO
  - Požadavky postupně prochází skrz několik systémů  
*uživatel → aplikační server → databáze → aplikační server → uživatel*
  - Dokážu-li charakterizovat dílčí subsystémy, jak se bude chovat celý systém?
- Model lze stále sestavit z jednotlivých SHO
  - Někdy je možné ho rozložit a řešit SHO samostatně
  - Někdy je nutné řešit model jako celek (zpětné vazby – požadavky se vrací do systému)
- Orientovaný graf
  - Uzly: elementární SHO
  - Hrany: toky mezi elementárními SHO

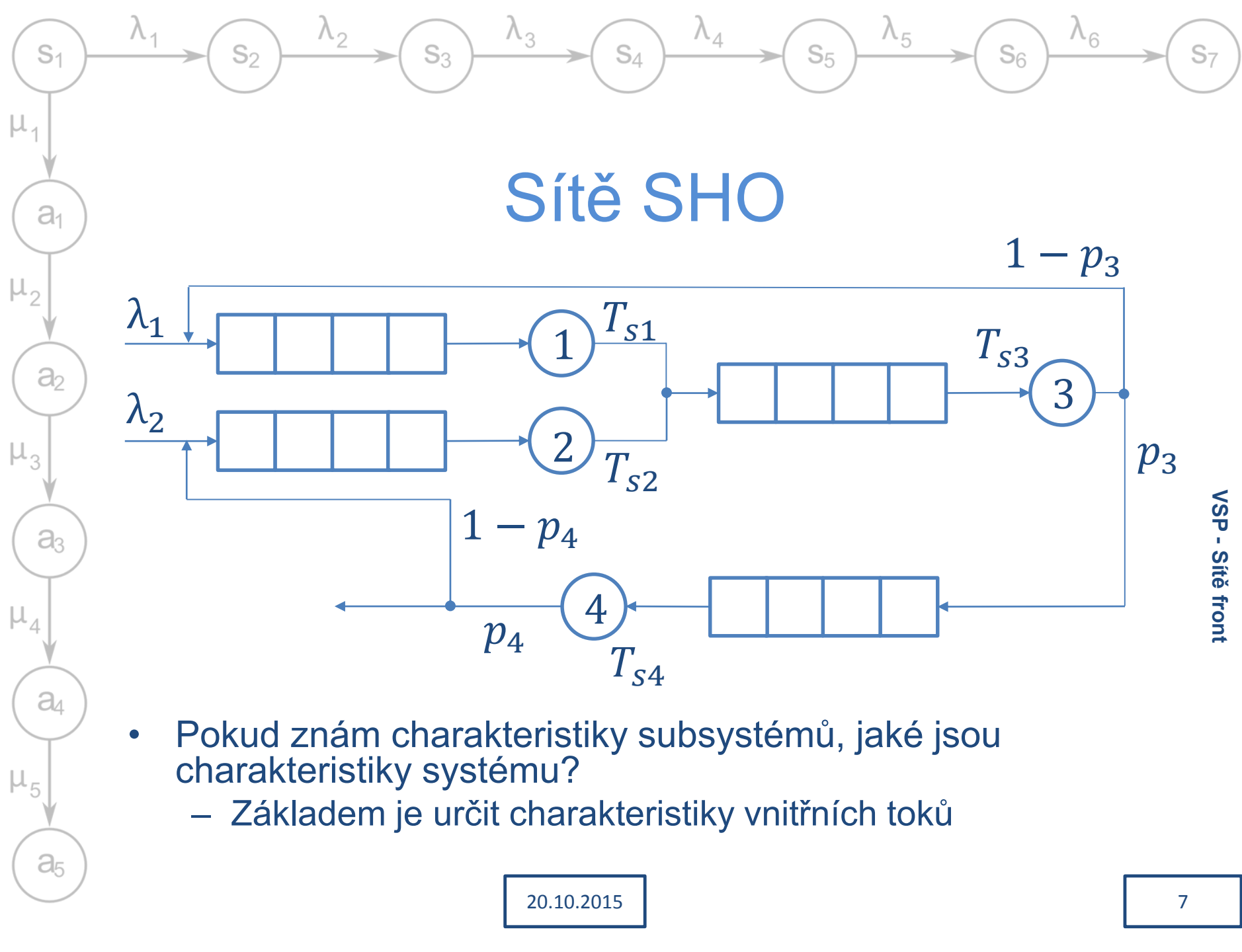


# Sítě SHO



VSP - Síť front

- Jednotlivé systémy lze řešit odděleně
  - Do uzlu 3 vstupuje součet toků 1 a 2 – lze zjistit jejich analýzou
  - Do uzlu 4 vstupuje výstupní tok uzlu 3

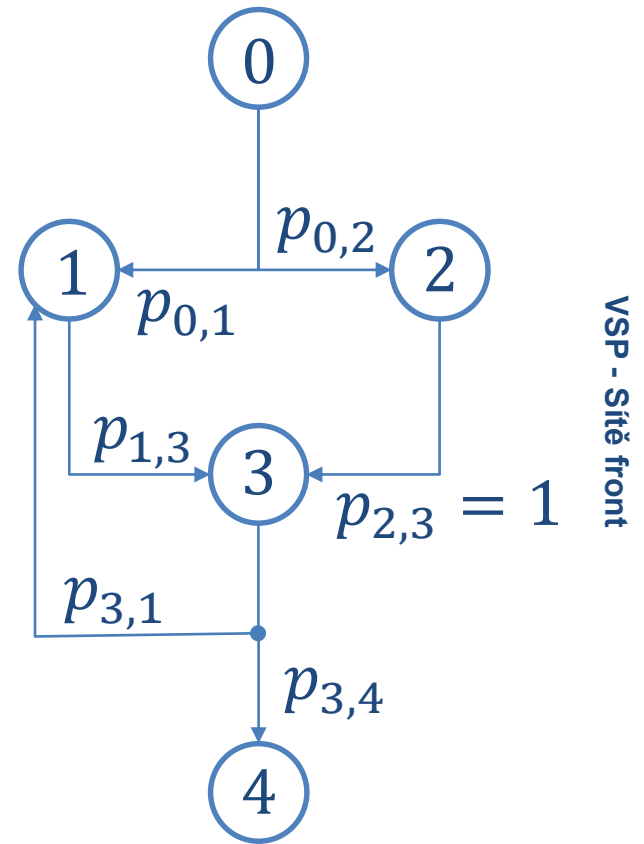




# Popis SHO

## Orientovaný graf

- Uzly – elementární SHO,  $n$  uzlů
- Hrany – toky mezi nimi  
( $\rightarrow$  hrana vede z výstupu obslužného kanálu do fronty jiného kanálu)
  - Intenzita toku do uzlu  $\Lambda_i$  - součet dílčích toků do  $i$
  - Váha hrany: pravděpodobnost přechodu z uzlu  $i$  do uzlu  $j$  -  $p_{i,j}$
- Okolí lze modelovat dalšími uzly (zdroj a stok požadavků)
- Zjevně platí  $p_{3,1} + p_{3,4} = 1$ 
  - $\sum_{i=0}^n p_{k,n} = 1$  pro každý uzel  $k$





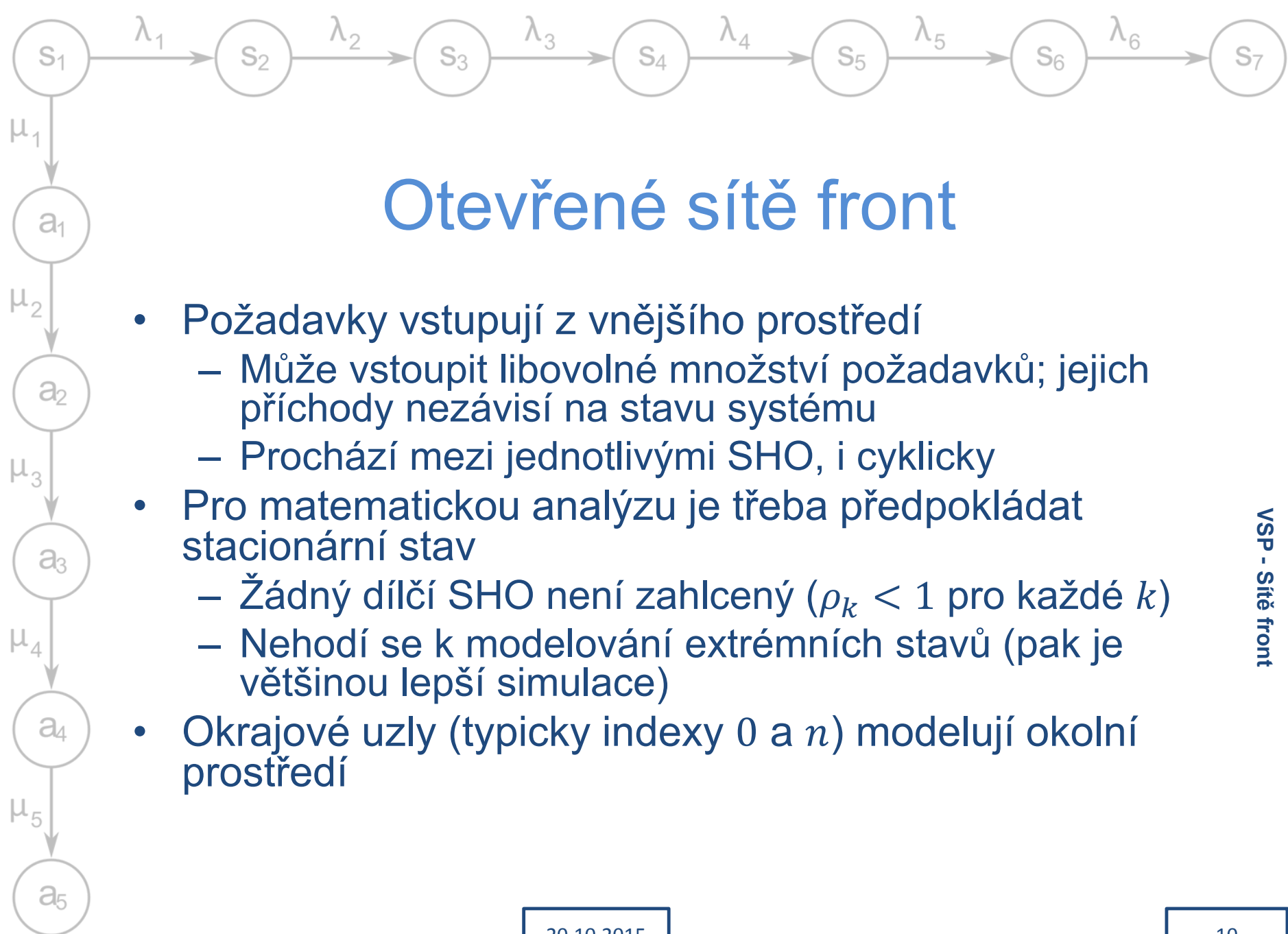


# Obecné vlastnosti sítí SHO

VSP - Síť front

- Zákon zachování toků
  - „co do uzlu nebo křížení vteče, musí vytéct“
  - $\sum_{i=0}^n p_{k,n} = 1$  pro každý uzel  $k$
  - Neplatí pro okrajové uzly v otevřených sítích – okolní prostředí
  - Spojování toků zachovává exponenciální rozdělení
- Výstupní tok z M/M/m
  - Stejná frekvence a intenzita jako u vstupního toku, charakter záleží na zatížení (očekáváme  $\rho < 1$ )
- „Jacksonův zákon“
  - Jsou-li všechny vstupní toky poissonovské a obsluhy exponenciální, jsou i výstupní toky exponenciální  
 → jednotlivé dílčí SHO lze řešit jako M/M/m (pokud dokážeme určit vnitřní toky)



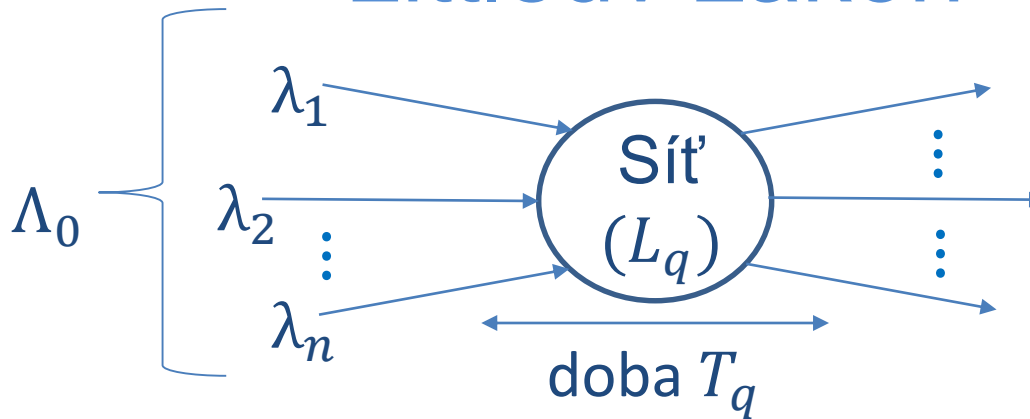


# Otevřené sítě front

- Požadavky vstupují z vnějšího prostředí
  - Může vstoupit libovolné množství požadavků; jejich příchody nezávisí na stavu systému
  - Prochází mezi jednotlivými SHO, i cyklicky
- Pro matematickou analýzu je třeba předpokládat stacionární stav
  - Žádný dílčí SHO není zahlcený ( $\rho_k < 1$  pro každé  $k$ )
  - Nehodí se k modelování extrémních stavů (pak je většinou lepší simulace)
- Okrajové uzly (typicky indexy 0 a  $n$ ) modelují okolní prostředí



# Littleův zákon



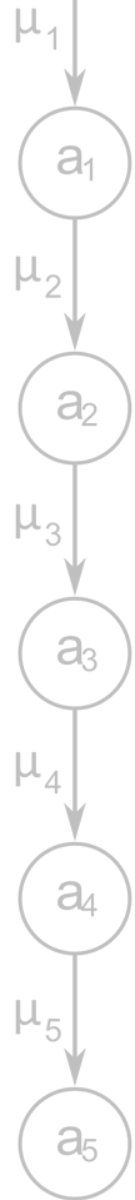
- $L_q = \Lambda_0 \cdot T_q$ 
  - počet požadavků v systému je úměrný intenzitě vstupů a době sledování
  - Stejné důvody jako u elementárních SHO (pokud je ve stacionárním stavu – nutno kontrolovat pro každý uzel zvlášť)

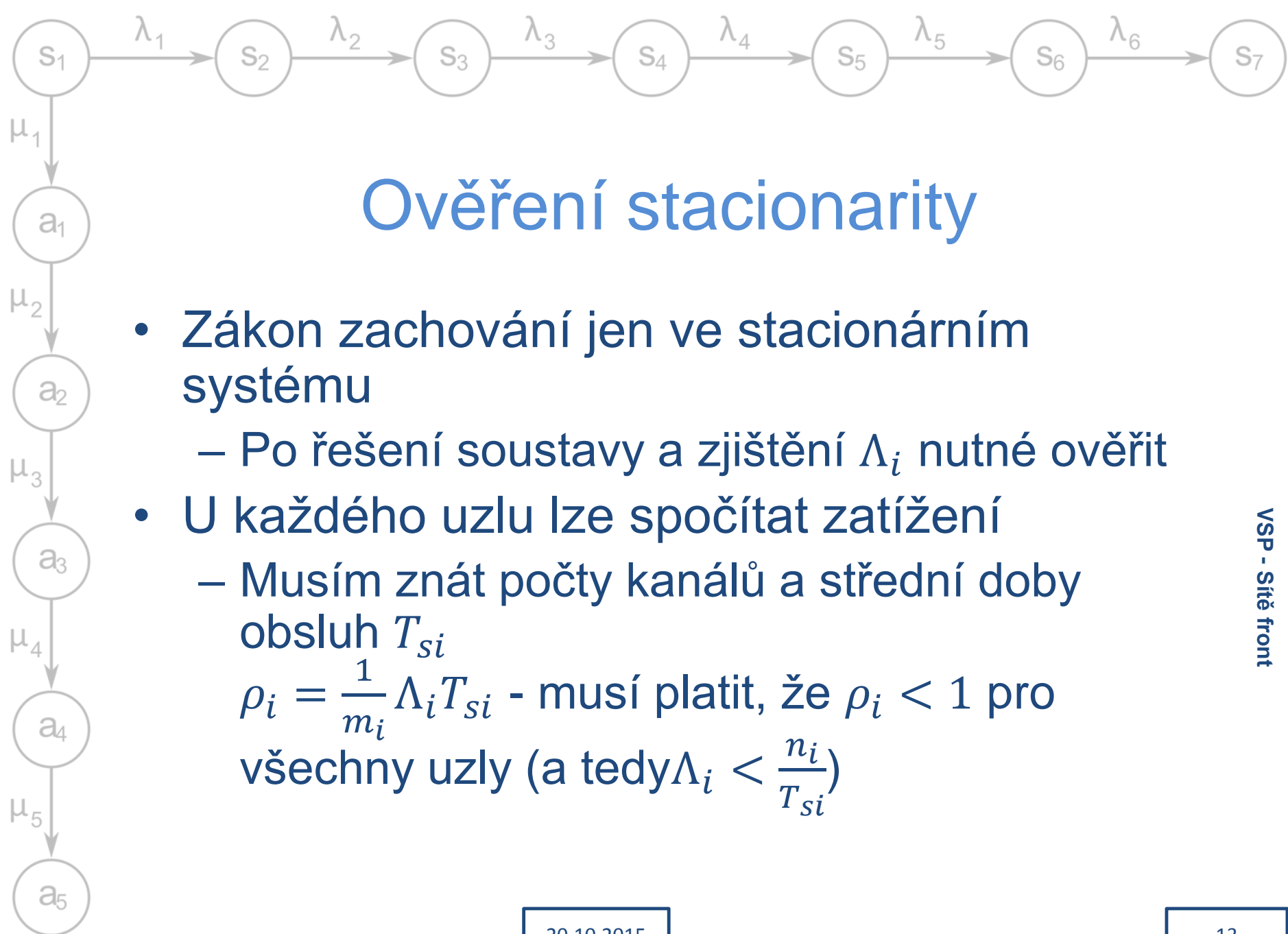


# Střední frekvence a zatížení

- Nezávisí na rozdělení, jen na středních frekvencích
- $\Lambda_i$  - Vnitřní frekvence toku v uzlu
  - $\Lambda_0$  - souhrnný vstupní tok do systému
- Tok požadavků z  $i$  do  $j$ :  $\lambda_{i,j} = \Lambda_i p_{i,j}$ 
  - Na vstupu uzlu  $i$  se sčítají frekvence na přicházejících hranách
- Stacionární režim a zákon zachování  $\rightarrow$ 

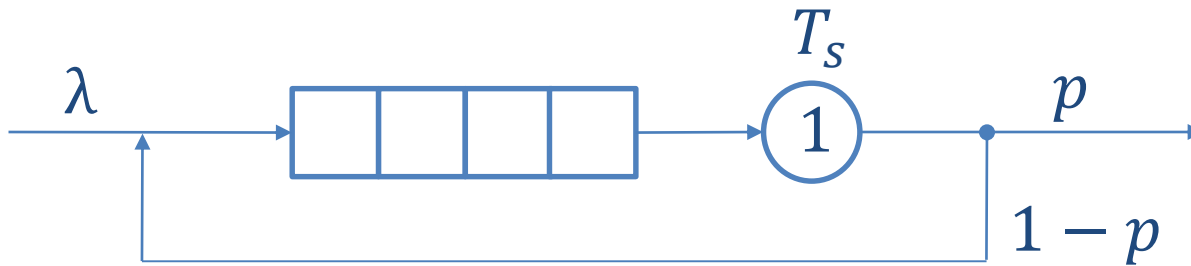
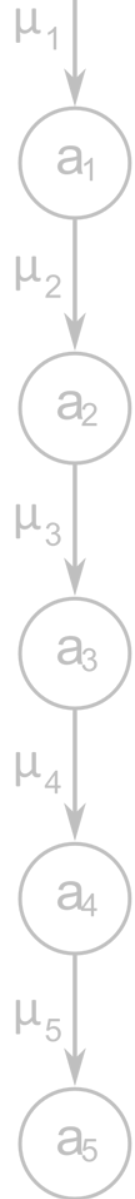
$$\sum_k \Lambda_k p_{k,i} = \Lambda_i = \sum_j \Lambda_j p_{i,j}$$
  - $k$  – uzly ze kterých vede hrana do  $i$
  - „všechno co do uzlu vtéká z něj vytéká a nic navíc“
  - Řešením soustavy  $n$  rovnic získám vnitřní toky  $\Lambda_i$







## Příklad – finanční úřad



- Přichází 15 lidí za hodinu, obsluha trvá průměrně 3 minuty  
 $\rightarrow \lambda = 15 \text{ hod}^{-1}, T_s = 0,05 \text{ hod}$
- Průměrně 1 formulář z 10 vyplněn špatně  
 $\rightarrow p = 0,9$

$$\Lambda_1 = \lambda + \Lambda_1(1 - p)$$

$$\Lambda_1 = \frac{\lambda}{p} \cong 16,66$$

$$\rho_1 = \Lambda_1 T_s \cong 0,833$$

stacionární - OK

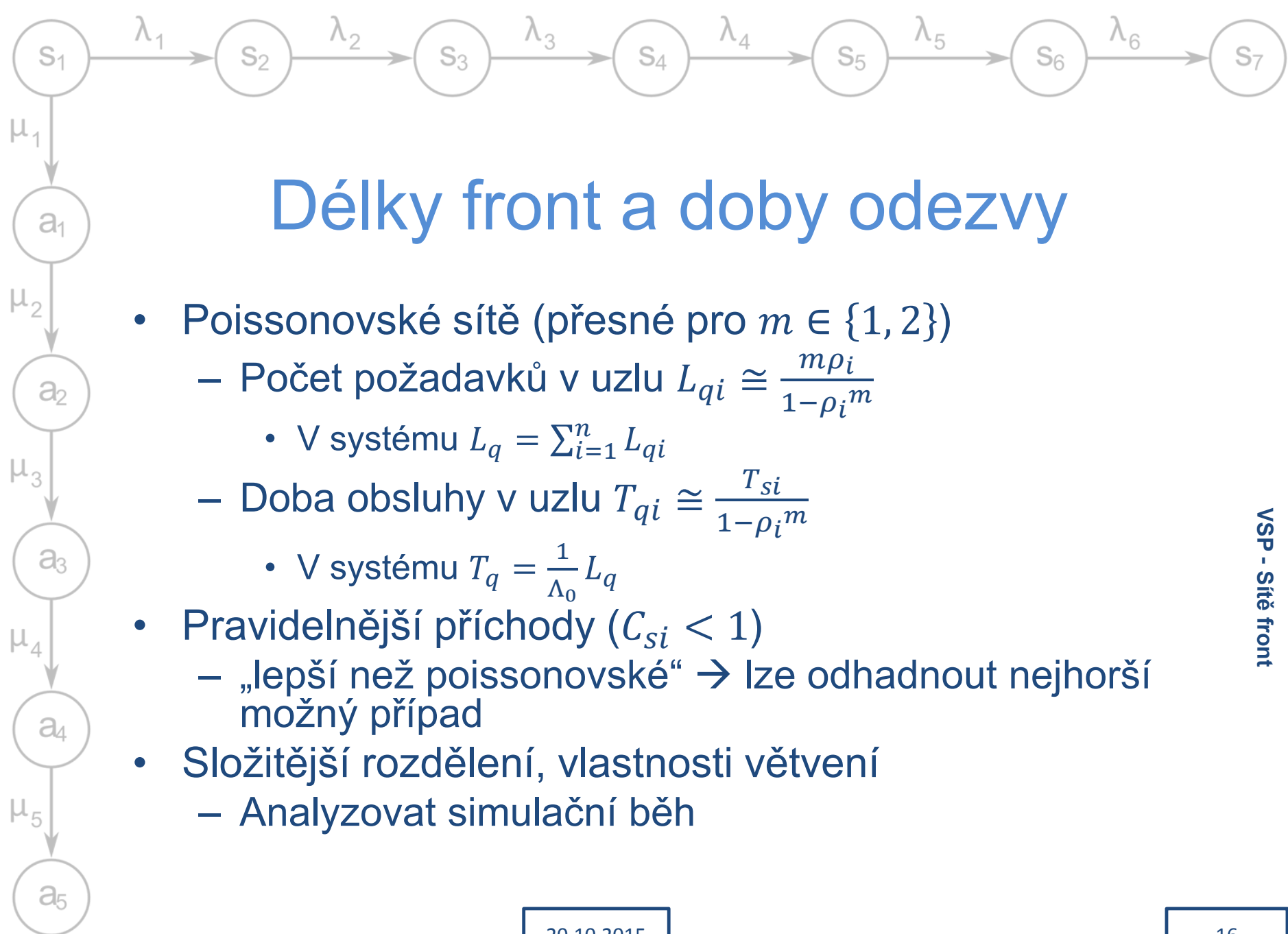


# Délky front a doby odezvy

- Souvisí s charakterem rozdělení
  - Délky front narůstají s nepravidelností intervalů (synchronizované systémy fronty nevytváří)
- Pro analytické řešení je třeba splnit Jacksonovy podmínky
  - Toky požadavků z okolí mají poissonovský charakter
  - Všechny doby obsluhy mají exponenciální rozdělení
  - Přejít mezi elementárními SHO je bez zpoždění, při větvení je náhodný, s pravděpodobností  $p_{i,j}$

→ výsledné toky budou také poissonovské a jednotlivé uzly (elementární SHO) lze analyzovat odděleně

  - Frekvence vstupního toku  $\Lambda_i$ , střední doba obsluhy  $T_{si}$



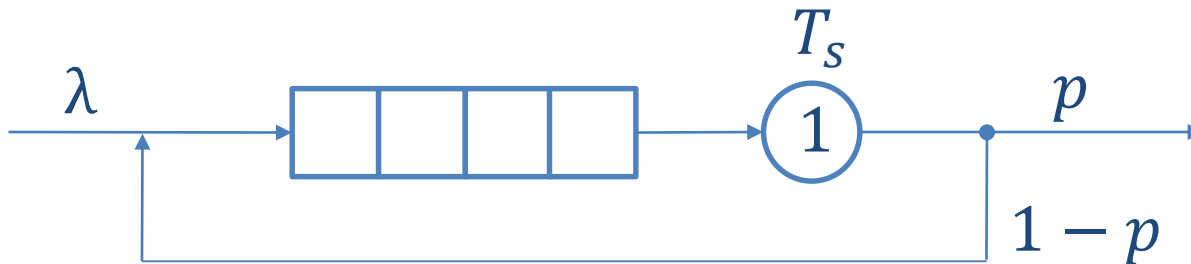
# Délky front a doby odezvy

- Poissonovské sítě (přesné pro  $m \in \{1, 2\}$ )
  - Počet požadavků v uzlu  $L_{qi} \cong \frac{m\rho_i}{1-\rho_i^m}$ 
    - V systému  $L_q = \sum_{i=1}^n L_{qi}$
  - Doba obsluhy v uzlu  $T_{qi} \cong \frac{T_{si}}{1-\rho_i^m}$ 
    - V systému  $T_q = \frac{1}{\Lambda_0} L_q$
- Pravidelnější příchody ( $C_{si} < 1$ )
  - „lepší než poissonovské“ → lze odhadnout nejhorší možný případ
- Složitější rozdělení, vlastnosti větvení
  - Analyzovat simulační běh





## Příklad – finanční úřad



- Přichází 15 lidí za hodinu, obsluha trvá průměrně 3 minuty  
 $\rightarrow \lambda = 15 \text{ hod}^{-1}, T_s = 0,05 \text{ hod}, \mu = \frac{1}{T_s} = 20 \text{ hod}^{-1}$
- Průměrně 1 formulář z 10 vyplněn špatně  
 $\rightarrow p = 0,9$
- Všechny doby exponenciální (pravidelnější - lepší výsledek)

$$L_q = \frac{\rho}{(1-\rho)} \cong \frac{0,833}{0,166} \cong 5,018$$

$$L_w = L_q - \rho \cong 4,185$$

$$T_q = L_q / \lambda \cong 0,334 \text{ hod} \cong 20 \text{ min}$$



## Uzavřené sítě front

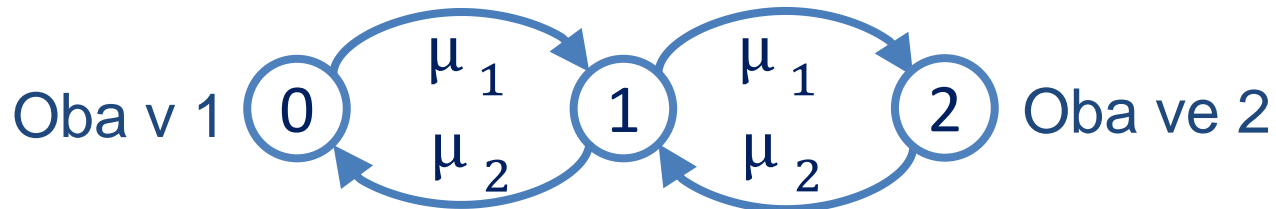
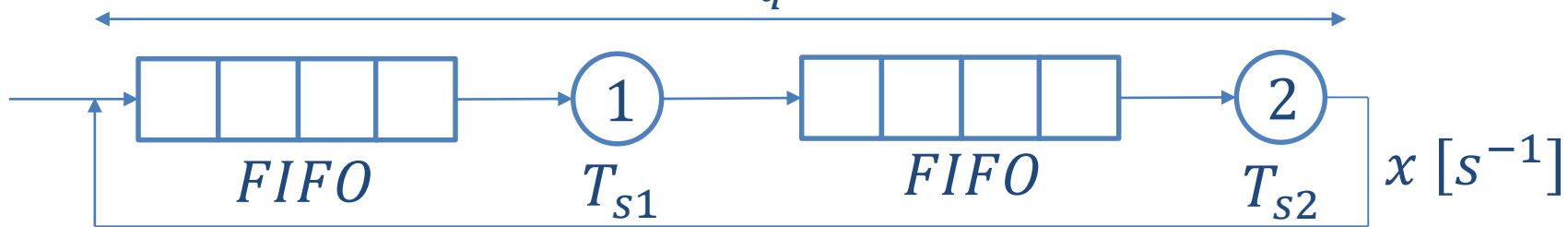


- Žádné vstupy požadavků z okolí
  - Pevná množina požadavků „krouží“ sítí  
→ fronty nemohou růst do nekonečna
- Hledáme typicky
  - Střední frekvenci průchodů nějakým místem (*propustnost*)
  - Střední doba průchodu mezi dvěma body v síti
- Analytické řešení jen pro poissonovské sítě
  - Lze řešit s využitím markovských modelů
  - Obecné řešení je složité



# Jednoduchý příklad

$T_q$

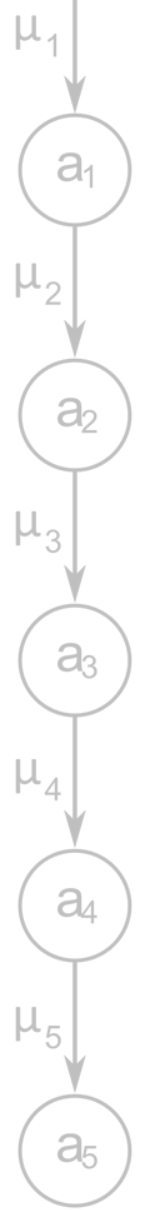
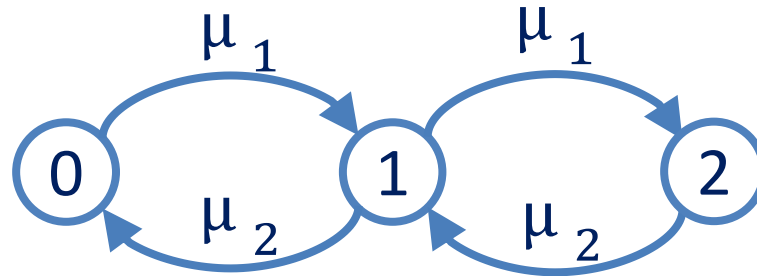


VSP - Síť front

- 2 požadavky  $\rightarrow$  jen 3 stavy, lze určit pravděpodobnosti
- $x$  – průchodnost
- Intenzity obsluh:  $\mu_1 = 1/T_{s1}, \mu_2 = 1/T_{s2}$



# Jednoduchý příklad



- Průchodnost

$$x = p_1 \cdot \mu_2 + p_2 \cdot \mu_2$$

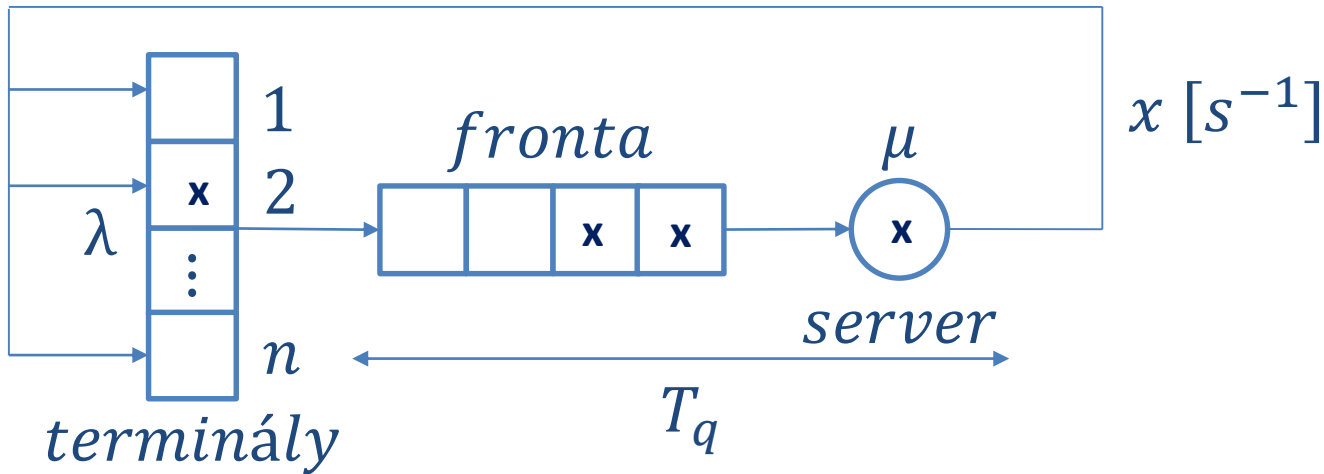
- Délky front v tabulce

$$L_{w1} = 1 \cdot p_0 + 0 \cdot p_1 + 0 \cdot p_2$$

Stav	Fronta 1	Fronta 2
0	1	0
1	0	0
2	0	1



# Příklad – interaktivní systém

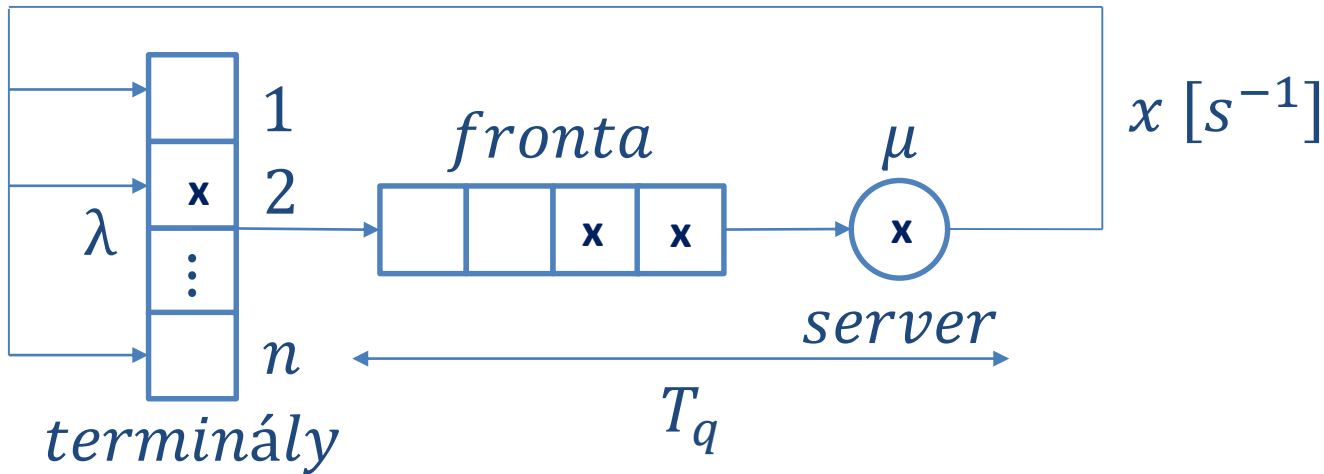


VSP - Síť front

- Server +  $n$  terminálů
  - Jednovláknové zpracování, obsluha s intenzitou  $\mu$
  - Terminály stále v provozu, uživatelé zadávají požadavky s frekvencí  $\lambda$
  - Exponenciální rozdělení
- Požadavek – práce uživatele
  - Přemýšlí ( $\lambda$ ), čeká ve frontě, čeká na výsledek ( $\mu$ )



# Příklad – interaktivní systém

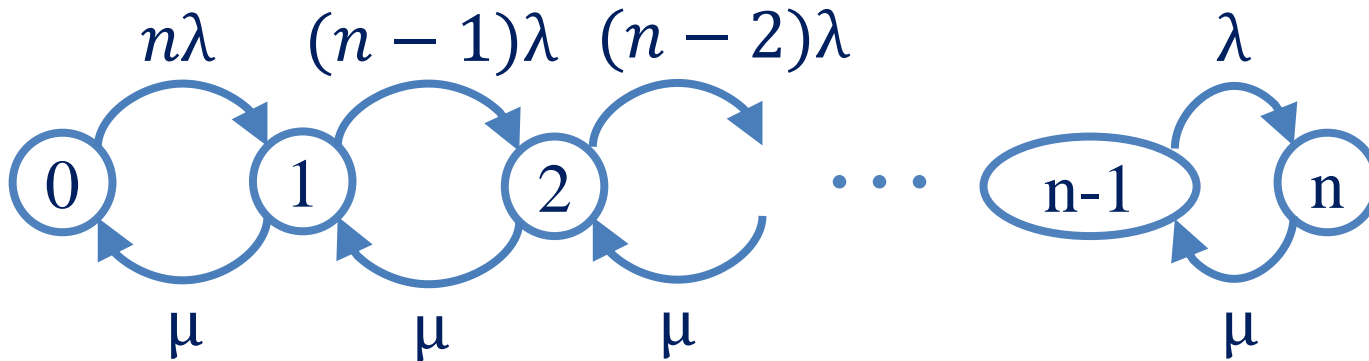
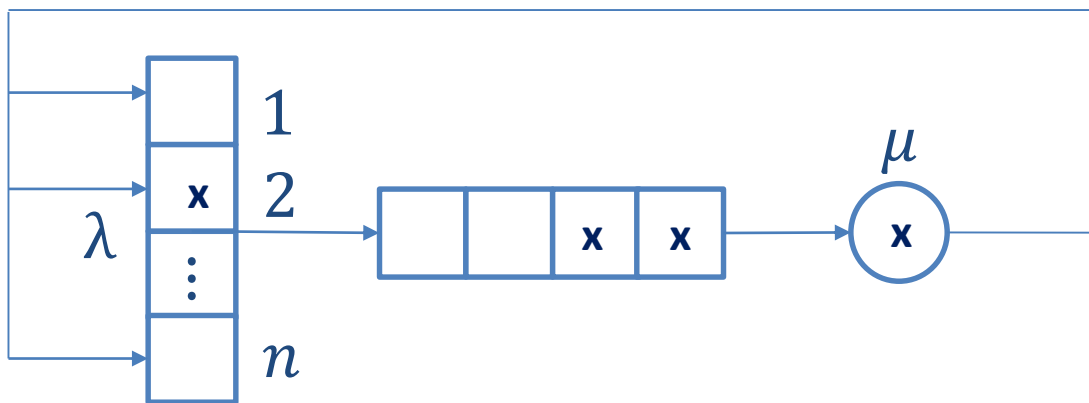


VSP - Síť front

- Střední doba přemýšlení  $T_t$ ,  $\lambda = 1/T_t$
- Střední doba obsluhy  $T_s$ ,  $\mu = 1/T_s$
- Poměr  $\frac{T_s}{T_t} = \frac{\lambda}{\mu} = \rho$  – nemá význam zatížení jako v otevřených SHO
- Stav systému – počet požadavků v SHO

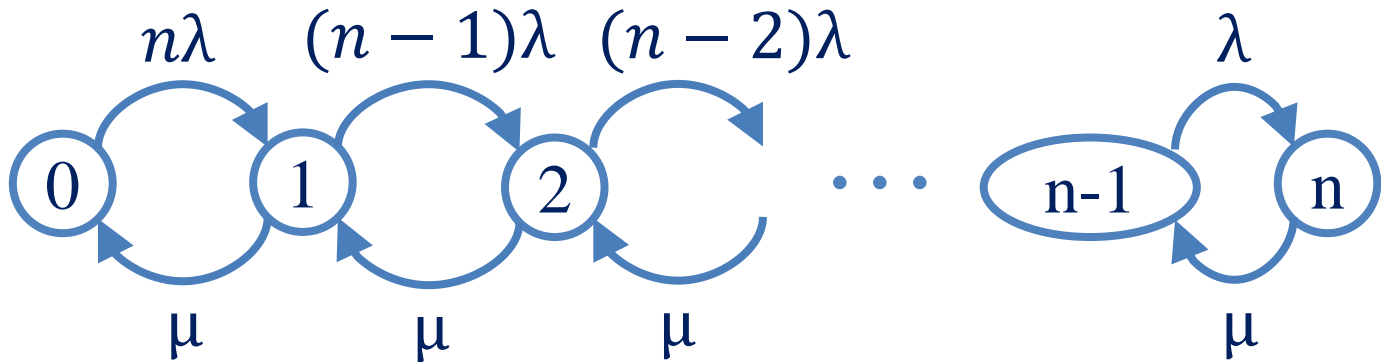


# Markovský model interaktivního systému





# Markovský model interaktivního systému



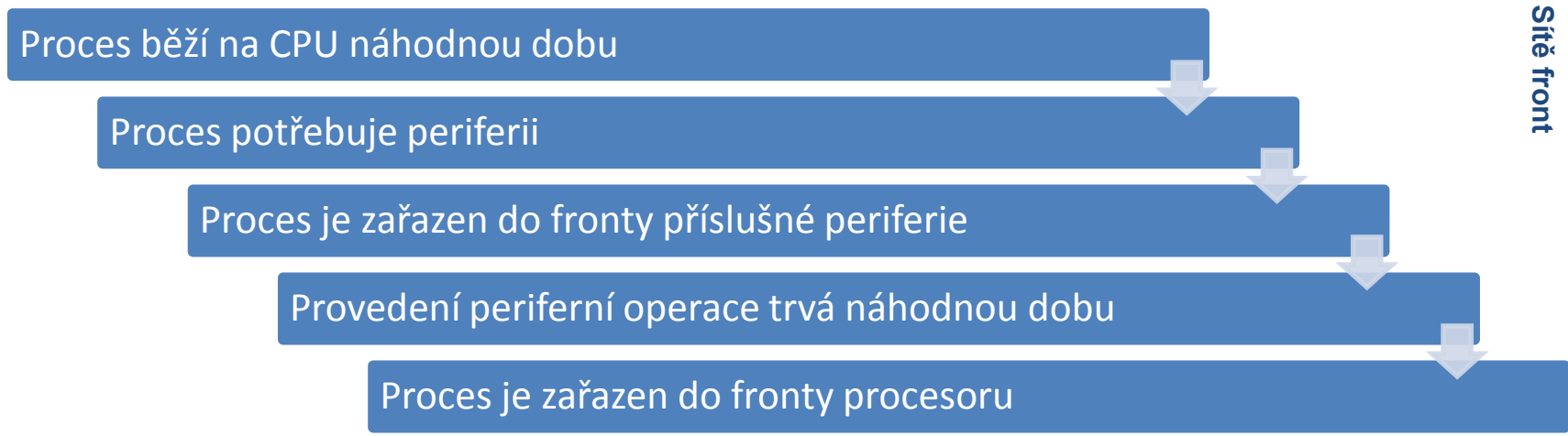
- Lze určit ustálené pravděpodobnosti (předpokládáme že existují – soustava lin. rovnic)  $p_0 \dots p_n$  a z nich určit vlastnosti
- $x = \frac{1}{T_s} (1 - p_0)$  ( $1 - p_0$  - pravděpodobnost že je kanál obsazený)
  - Nepodmíněná frekvence obsluhy (podmíněná -  $\mu$ )
- $T_q = \left( n \frac{T_s}{1-p_0} \right) - T_t = \frac{n}{x} - T_t$

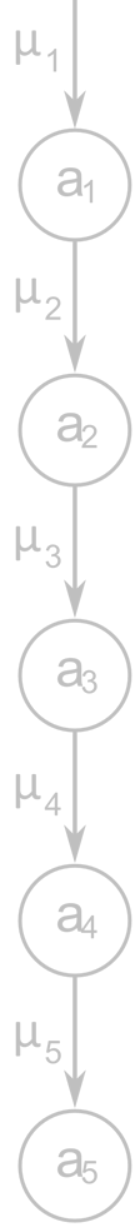




# Příklad – OS s více procesy

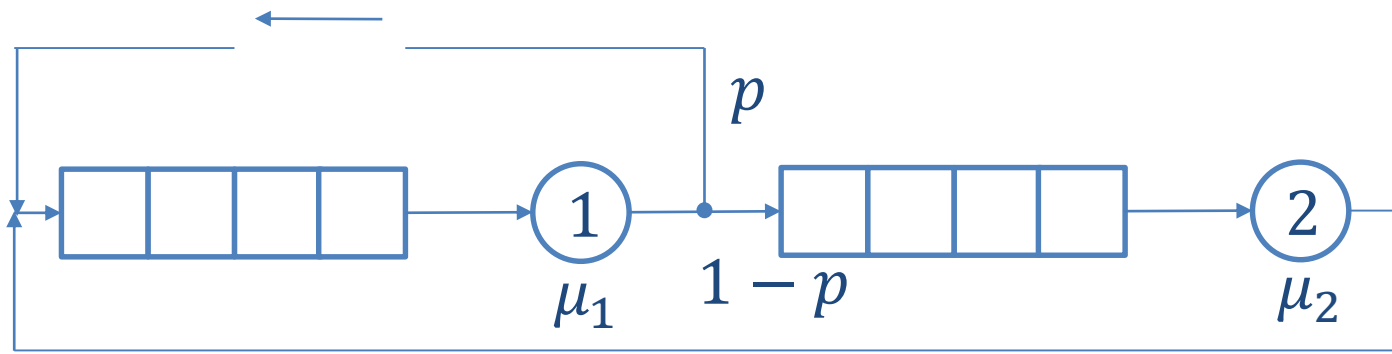
- 2 elementární SHO:
  - Procesor – elementární SHO
  - Řadič periferií
- Pevně daný počet najednou zpracovávaných úloh
- Ukončený proces je ihned nahrazen dalším (dávka požadavků ve vnější paměti)
  - počet průchodů přes procesor = propustnost





# Konkrétní příklad – správa paměti

- Součásti:
  - CPU (1)
  - Řadič paměti (stránkování, 2)
- 3 procesy najednou, stejné vlastnosti
- $p$  – pravděpodobnost že opuštění procesoru znamená ukončení programu
- Exponenciální rozdělení
  - Dobré pro konflikty stránek, pro disk lépe Gaussovské nebo rovnoměrné

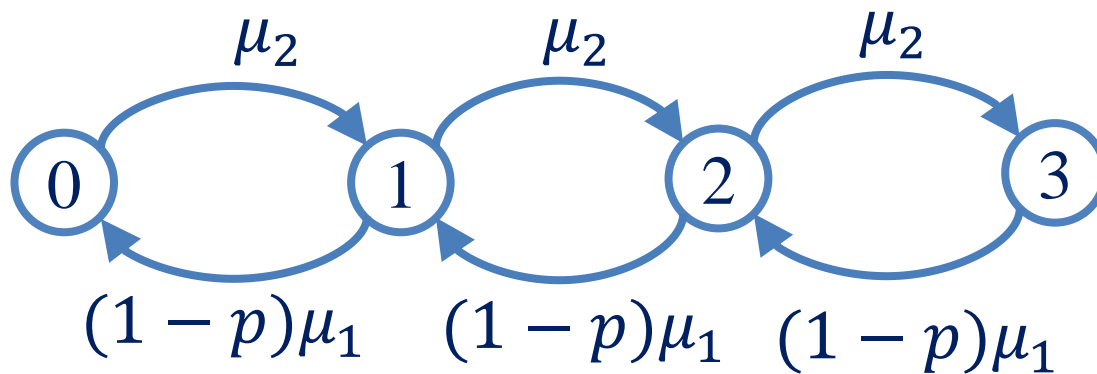
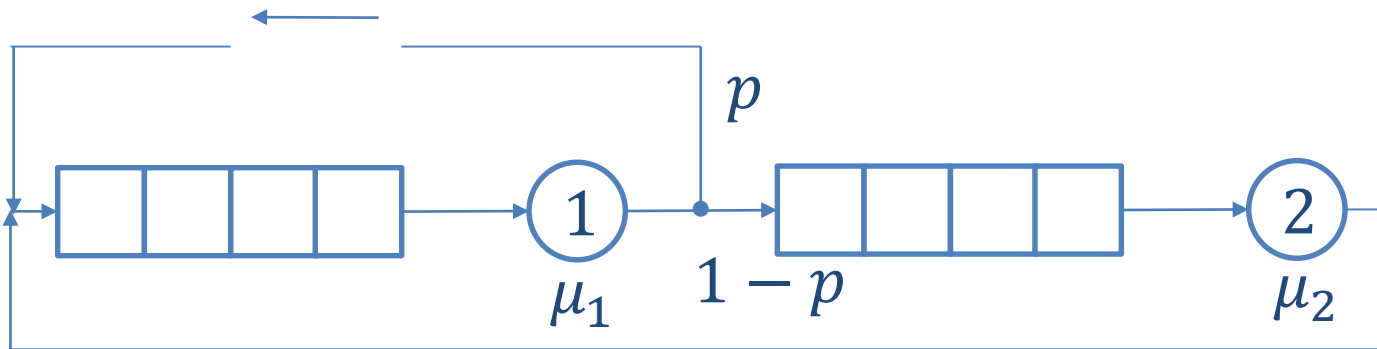


VSP - Síť front



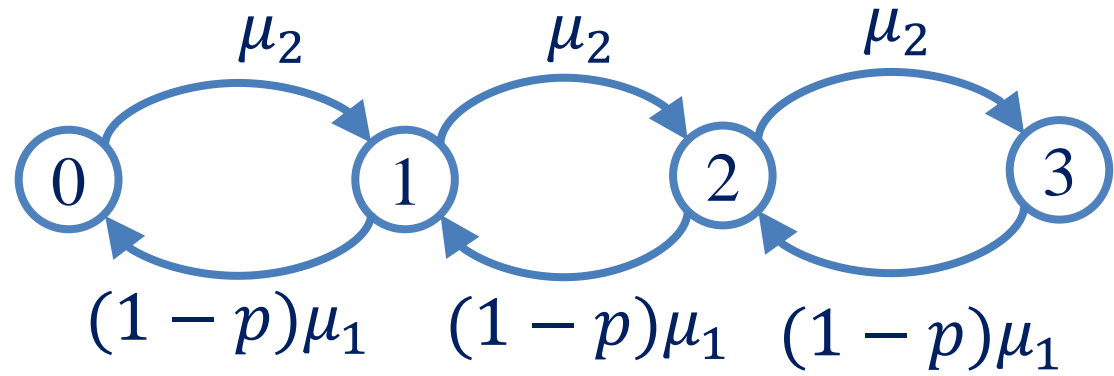
## Konkrétní příklad – správa paměti

- Lze sestavit markovský model bez absorbčních stavů





# Konkrétní příklad – správa paměti

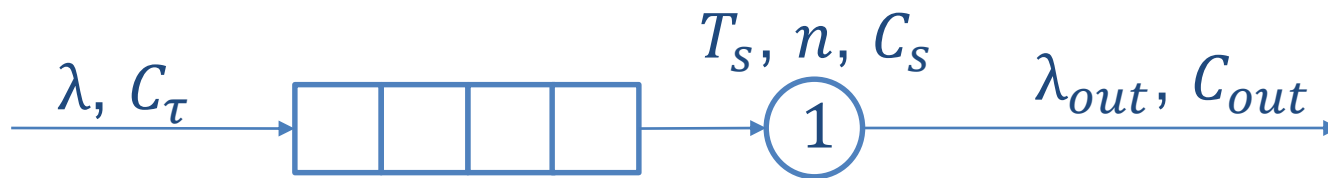


- Průchodnost  $x = \mu_1 p (p_1 + p_2 + p_3)$ 
  - $p$  – pravděpodobnost ukončení procesu
  - Může nastat jen když je aktuálně nějaký proces v CPU
  - Neodpovídá žádné explicitní hraně v markovském modelu
- Střední doba obsluhy  $T_q = \frac{n}{x} = \frac{3}{x}$ 
  - Proces není nikde zdržen, neodečítáme žádné  $T_t$



# Nepoissonovské sítě front – přibližné řešení

- Obecné toky  $\rightarrow$  charakterizovány frekvencí  $\lambda$  a koeficientem variace  $C_\tau$
- Otevřená síť front



- $\lambda_{out} = \lambda$  (jinak nebude stacionární)
- $C_{out}^2 = 1 + \rho^2(C_s^2 - 1) + (1 - \rho^2)(C_\tau^2 - 1)$   
–  $\rho$  určuje vliv rozdělení dob vstupů a dob obsluh
- Nelze snadno sčítat a dělit toky



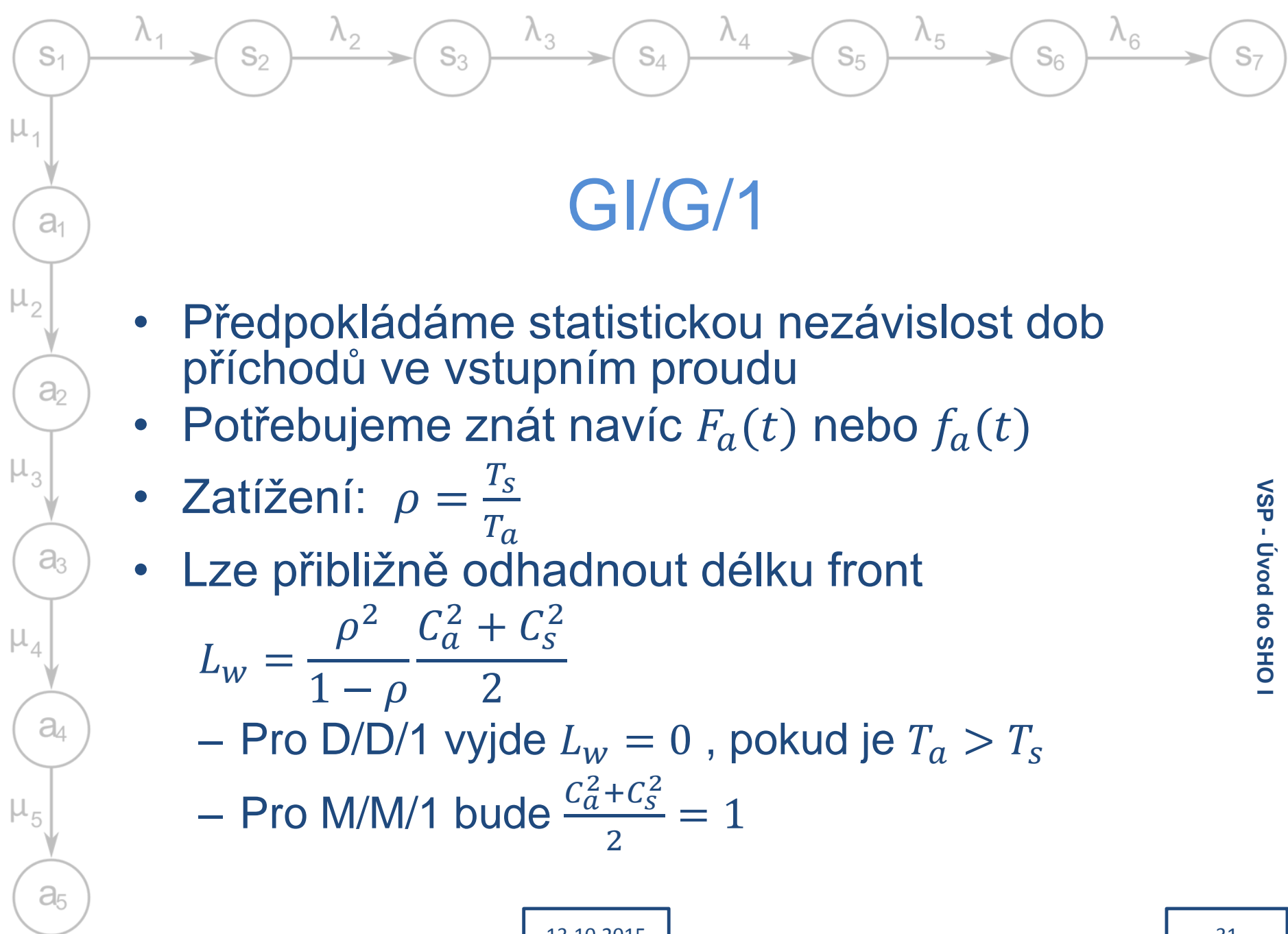
VSP - Síť front



# Nepoissonovské sítě front – toky

- Dělení toku s  $\lambda, C_\tau$  na  $n$  toků
  - $\lambda_i = \lambda p_i$
  - $C_i^2 = 1 + p_i(C_\tau^2 - 1)$
- Slučování toků s  $\lambda_i, C_i$ 
  - $\lambda = \sum_{i=1}^n \lambda_i$
  - $C^2 = 1 + \sum_{i=1}^n \left(\frac{\lambda_i}{\lambda}\right)^2 (C_i^2 - 1)$
- Pro jednotlivé elementární SHO lze použít vztahy pro GI/G/n

VSP - Sítě front



# GI/G/1

- Předpokládáme statistickou nezávislost dob příchodů ve vstupním proudu
- Potřebujeme znát navíc  $F_a(t)$  nebo  $f_a(t)$
- Zatížení:  $\rho = \frac{T_s}{T_a}$
- Lze přibližně odhadnout délku front

$$L_w = \frac{\rho^2}{1 - \rho} \frac{C_a^2 + C_s^2}{2}$$

– Pro D/D/1 vyjde  $L_w = 0$  , pokud je  $T_a > T_s$

– Pro M/M/1 bude  $\frac{C_a^2 + C_s^2}{2} = 1$



# Samostatná práce

- Síť s alespoň 4 uzly a 2 zpětnými vazbami
  - Nerozložitelná – zpětné vazby přes celý systém
  - Návrh na přednášku 10. listopadu
- Parametry volit tak, aby byla ve stacionárním režimu, ale se zátěží ( $\rho > 0,5$ )
- Určit toky v uzlech, L a T parametry v uzlech i pro celou síť
- Porovnat vypočtené a odsimulované hodnoty
- Sledovat další charakteristiky
  - Počet požadavků které projdou zadaným uzlem
  - Charakteristiky toku mezi vybranými uzly (histogram, střední hodnota, směrodatná odchylka)
  - Průměrný počet obrátek v systému
  - ...





Děkuji za pozornost



- Příště simulační analýzy SHO