

Ukazatele pro hodnocení výkonnosti a spolehlivosti



Výkonnost a spolehlivost – KIV/VSP

Richard Lipka

22.9.2015



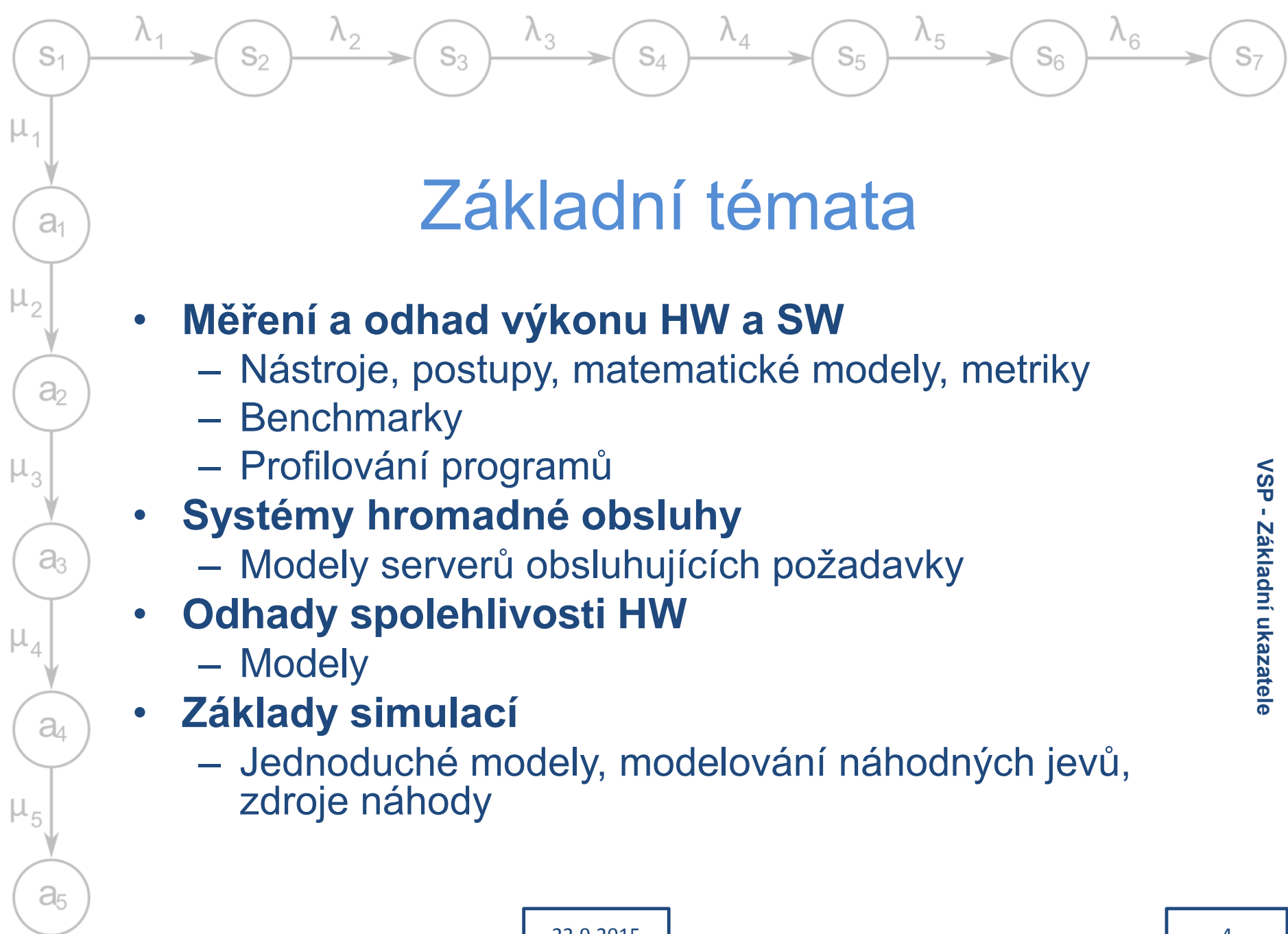
O předmětu

- Ing. Richard Lipka, PhD. (UN 306)
 - Kontakt: lipka@kiv.zcu.cz
 - KH: Po 14:00 – 15:00, Út 11:00 – 12:00
- Přednášky i cvičení v US-207
 - V út. 29.9. v UP 115
 - Cvičení konzultační, zaměřená na samostatné práce
 - Materiály na Courseware, aktualizován během týdne



Podmínky absolvování

- Průběžné příklady (5 příkladů po max. 4 bodech)
 - odevzdávání přes portál, v něm budou vidět termíny
- Semestrální práce (20 bodů, až 30 bodů za perfektní zpracování)
- Pro získání zápočtu min. 25 bodů
 - zapsán u zkoušky
- Zkouška písemná, max. 60 bodů
 - informace o zkoušce a ukázkové zadání na portálu
 - hodnocení – součet bodů ze samostatné práce a testu
 - předtermín na poslední přednášce (doporučen)



Základní témata

- **Měření a odhad výkonu HW a SW**
 - Nástroje, postupy, matematické modely, metriky
 - Benchmarky
 - Profilování programů
- **Systemy hromadné obsluhy**
 - Modely serverů obsluhujících požadavky
- **Odhady spolehlivosti HW**
 - Modely
- **Základy simulací**
 - Jednoduché modely, modelování náhodných jevů, zdroje náhody



K čemu to je?

- Měření výkonu jako druh testování SW
 - Výsledky záleží na SW i HW
 - Postupy blízké přírodním vědám
 - **Jak konstruovat experiment?**
 - **Jaké má mít vlastnosti?**
- Význam podle druhu aplikace
 - Nemusí jít jen o servery!
 - Nepodcenit – lidé neradi čekají na výsledky





Kdy je potřeba?

- **Rozsáhlé vědecké a technické výpočty**
 - Velké množství dat ke zpracování, často složitý výpočet
→ *Jaký HW musím použít?*
- **Grafika**
 - Všechno musí být spočteno včas
→ *Kolik věcí si mohu dovolit počítat najednou?*
- **Obsluha zákazníků**
 - Narůstá obliba tenkých klientů a chytrých serverů
→ *Kolik lidí najednou zvládnou obsloužit?*
→ *Jaké zdroje potřebuji pro zajištění kvality služeb?*
- **Mobilní aplikace**
 - Nefunguje dostatečně rychle? Zkusím jinou ...



Co když se něco nepovede?

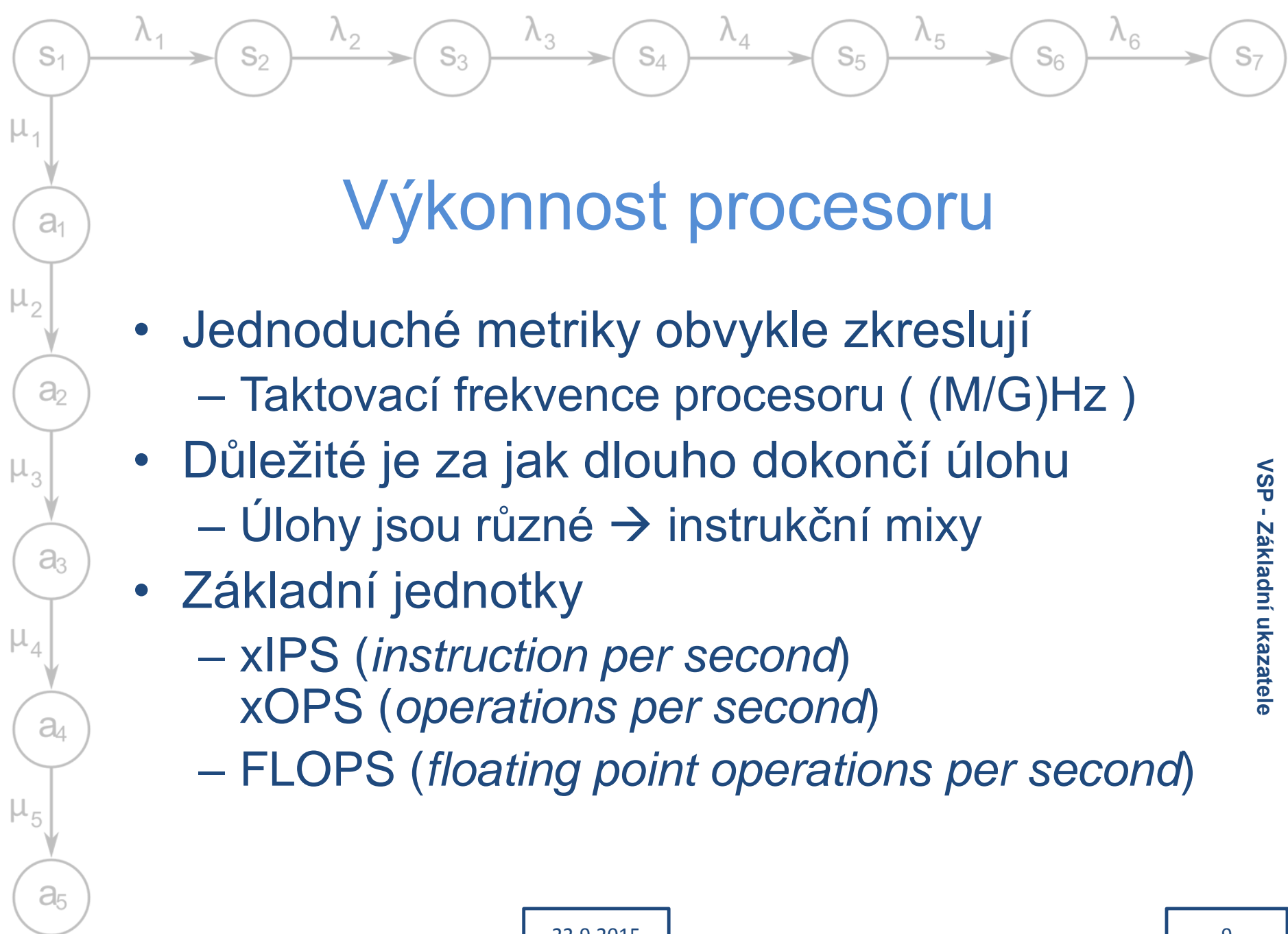
- **Předzápisy v portálu a podobné davové akce**
 - Nevyplatí se mít výpočetní kapacitu pro nárazovou zátěž, částečně řeší cloud
- **1987 – zhroucení burzy na Wall Street**
 - Systém nestihl zpracovat příkazy k ukončení prodeje
- **2011 – Bank of America online**
 - Systém po update nefunkční 6 dní po sobě
- **2012 – opakované selhání registru vozidel**
 - Nedostatečně odhadnutá potřebná kapacita pro nový systém



Výkonnostní parametry HW

- Jak měřit jak je počítač rychlý?
 - Obvykle zaměřené na jednotlivé komponenty systému
 - nejsou vypovídající pro celý systém





Výkonnost procesoru

- Jednoduché metriky obvykle zkreslují
 - Taktovací frekvence procesoru ((M/G)Hz)
- Důležité je za jak dlouho dokončí úlohu
 - Úlohy jsou různé \rightarrow instrukční mixy
- Základní jednotky
 - xIPS (*instruction per second*)
 - xOPS (*operations per second*)
 - FLOPS (*floating point operations per second*)

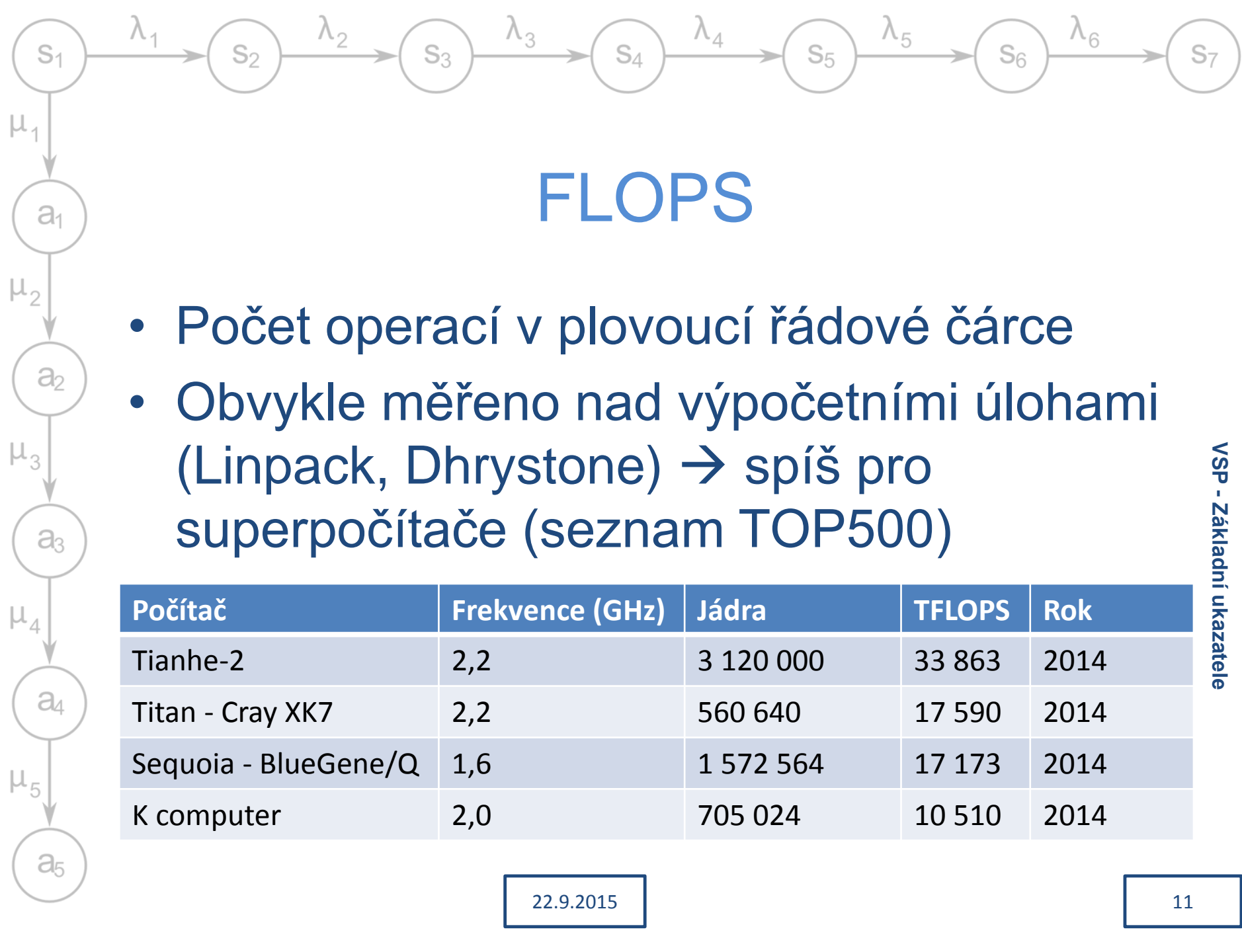


xIPS / xOPS

- Obvykle založeno na instrukčním mixu s celočíselnými instrukcemi (SPECint, Dhrystone) → není plně realistický

Procesor	Frekvence (GHz)	Jádra / vlákna	MIPS	Rok
Intel Core i7 2600K	3,4	4 / 8	128,3	2011
AMD FX-8350	4,2	8 / 8	97,2	2012
Intel Core i7 3770k	3,9	4 / 8	106,9	2013
Intel Core i7 4770k	3,9	4 / 8	127,3	2013

VSP - Základní ukazatele

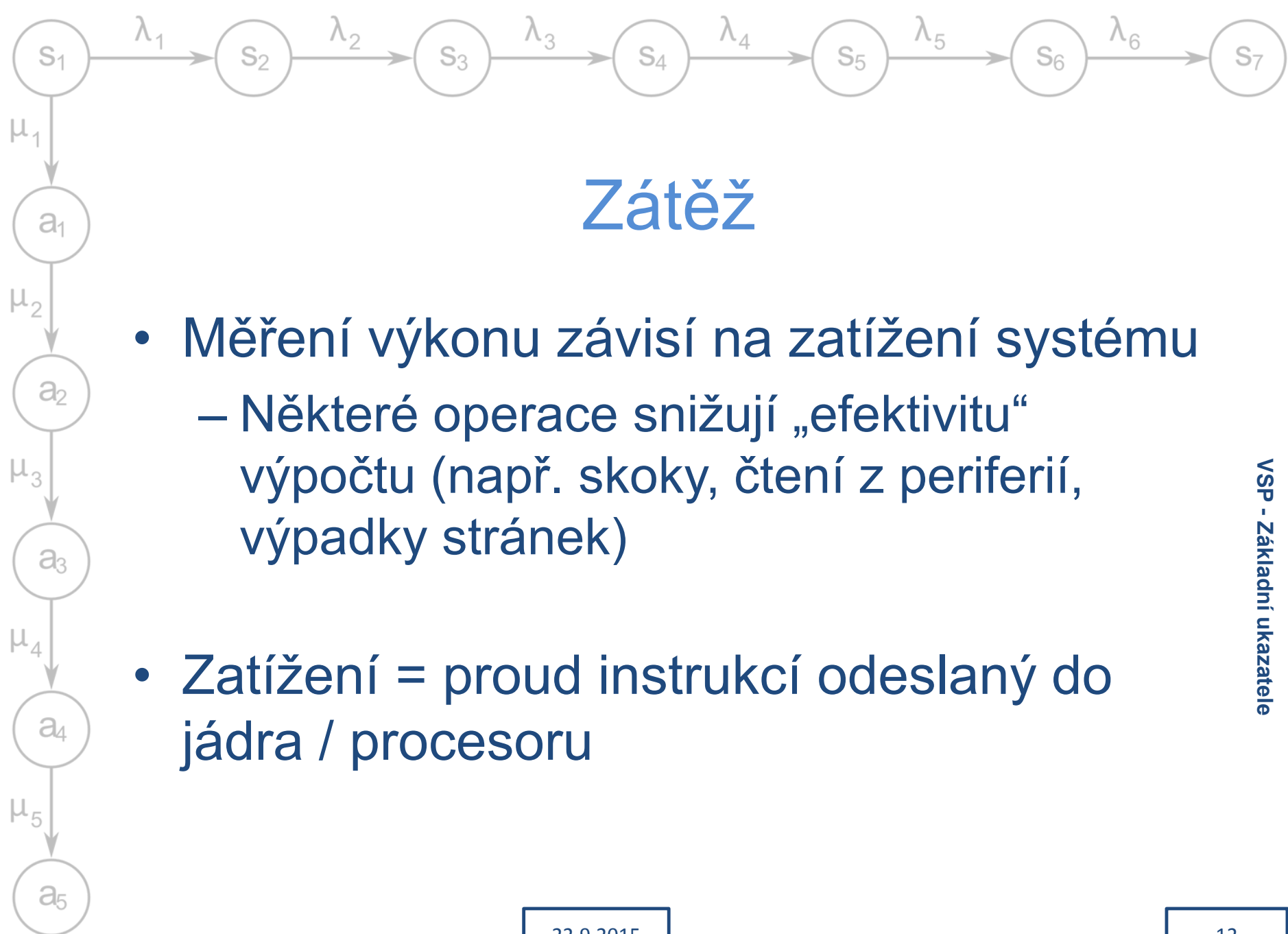


FLOPS

- Počet operací v plovoucí řádové čárce
- Obvykle měřeno nad výpočetními úlohami (Linpack, Dhrystone) → spíš pro superpočítače (seznam TOP500)

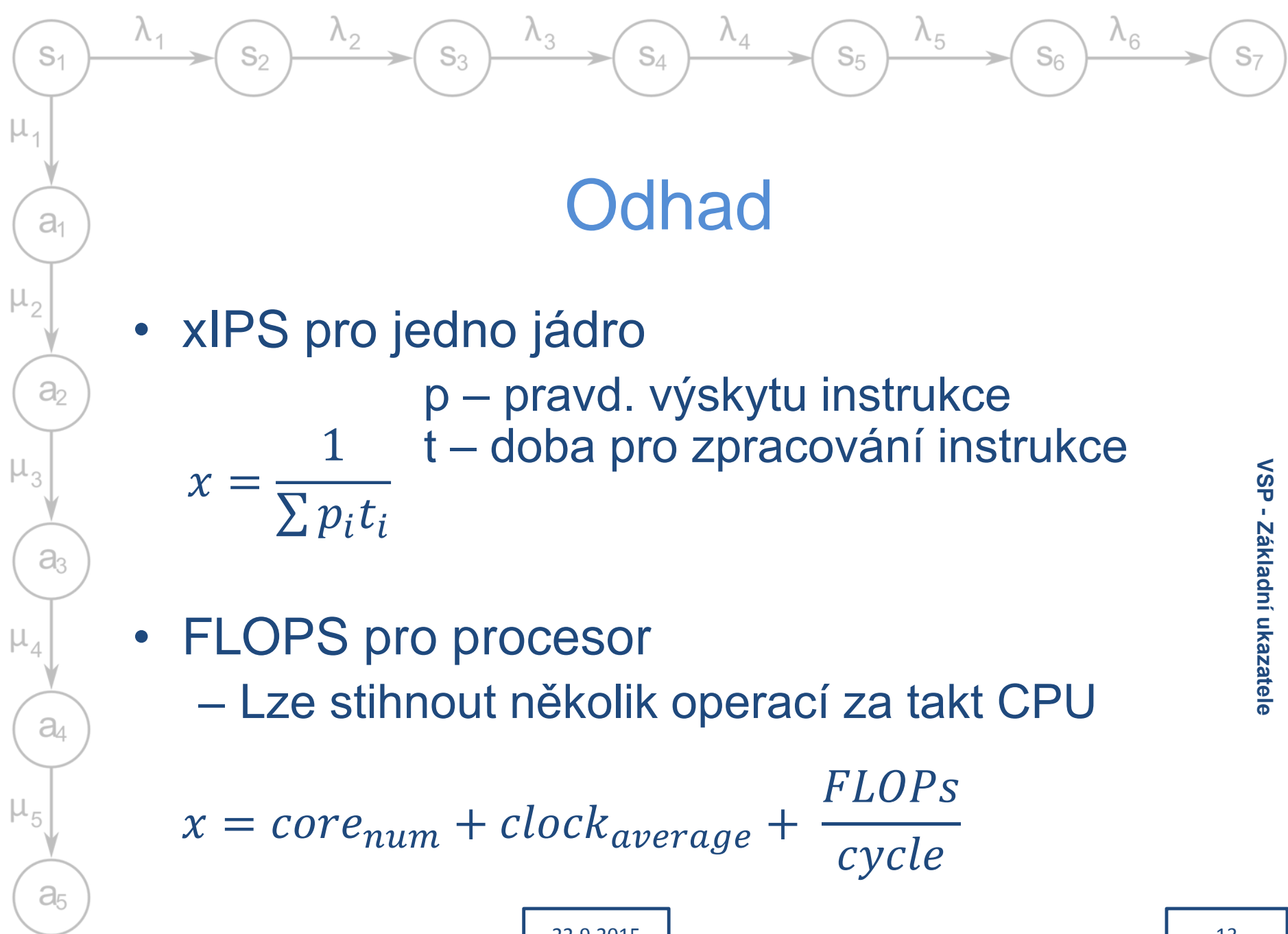
VSP - Základní ukazatele

Počítač	Frekvence (GHz)	Jádra	TFLOPS	Rok
Tianhe-2	2,2	3 120 000	33 863	2014
Titan - Cray XK7	2,2	560 640	17 590	2014
Sequoia - BlueGene/Q	1,6	1 572 564	17 173	2014
K computer	2,0	705 024	10 510	2014



Zátěž

- Měření výkonu závisí na zatížení systému
 - Některé operace snižují „efektivitu“ výpočtu (např. skoky, čtení z periferií, výpadky stránek)
- Zatížení = proud instrukcí odeslaný do jádra / procesoru



Odhad

- xIPS pro jedno jádro

p – pravd. výskytu instrukce
 t – doba pro zpracování instrukce

$$x = \frac{1}{\sum p_i t_i}$$

- FLOPS pro procesor
 - Lze stihnout několik operací za takt CPU

$$x = core_{num} + clock_{average} + \frac{FLOPs}{cycle}$$

VSP - Základní ukazatele



Výkonnost systému

- System složen z více prvků
- Instrukce / operace moc malé
 - Zajímá mě „počet užitečných výpočtů“

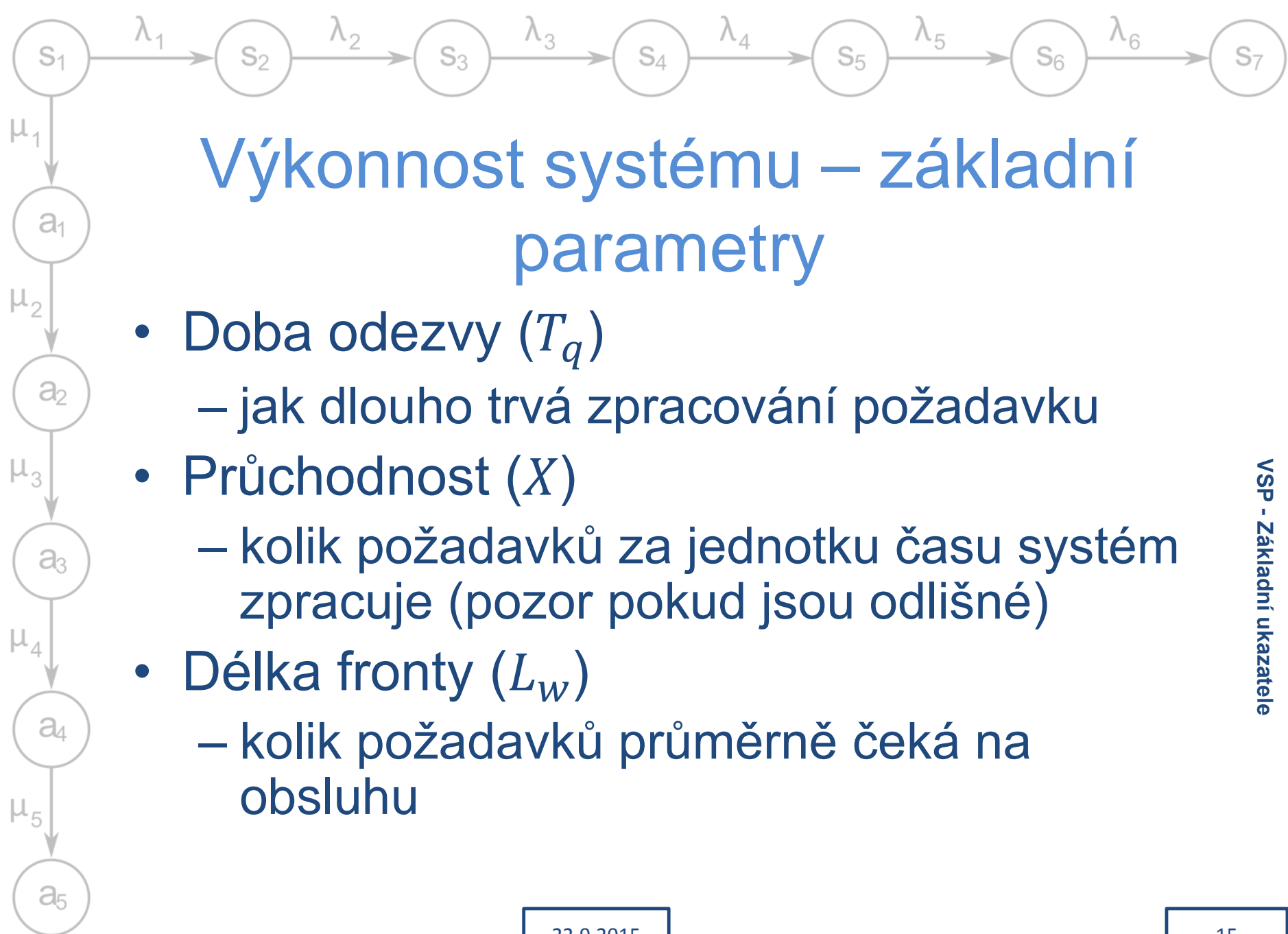
→ FPS (grafika)

→ transakce (SŘBD)

→ procesy (cloud)

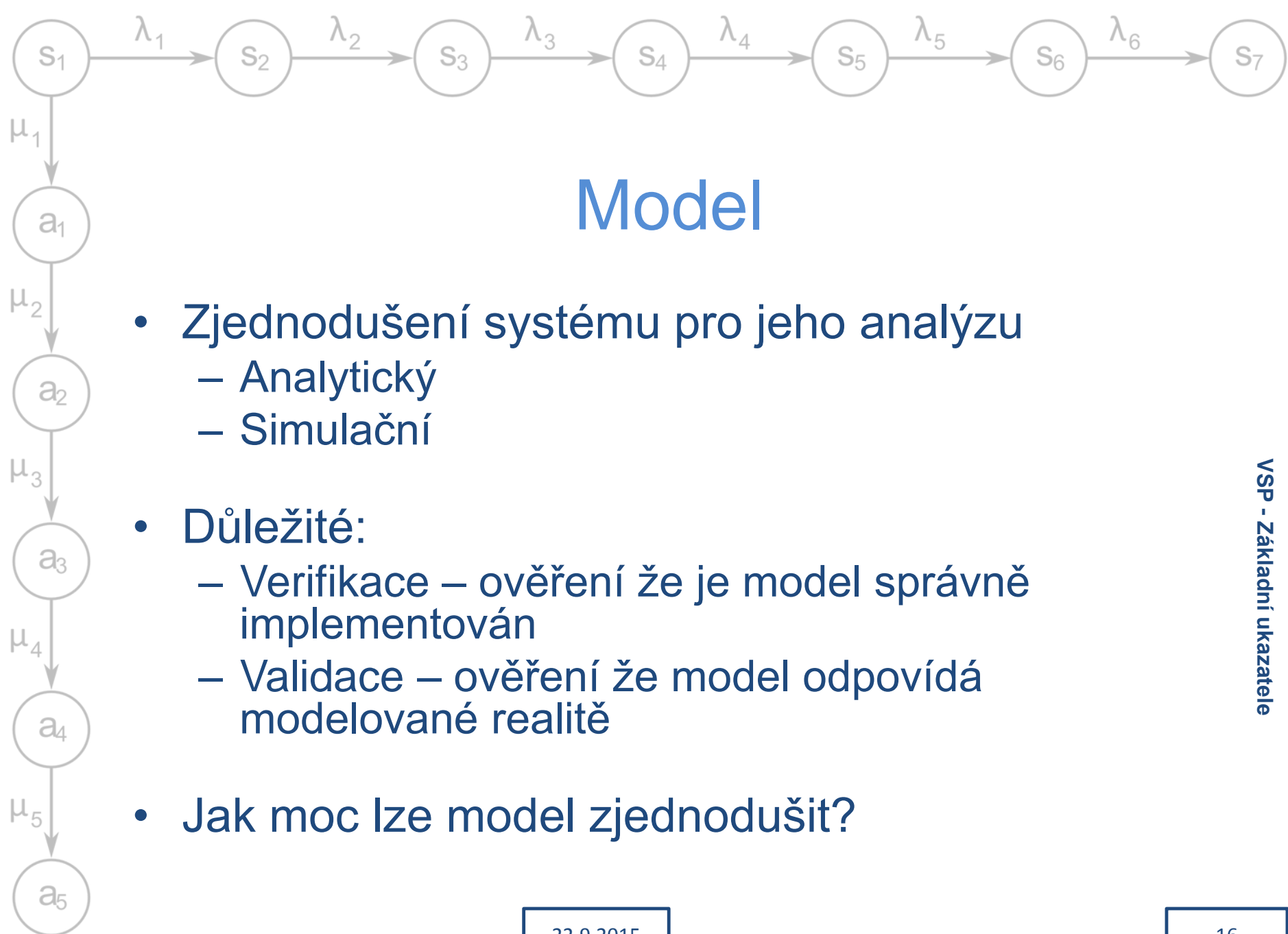
} požadavky





Výkonnost systému – základní parametry

- Doba odezvy (T_q)
 - jak dlouho trvá zpracování požadavku
- Průchodnost (X)
 - kolik požadavků za jednotku času systém zpracuje (pozor pokud jsou odlišné)
- Délka fronty (L_w)
 - kolik požadavků průměrně čeká na obsluhu



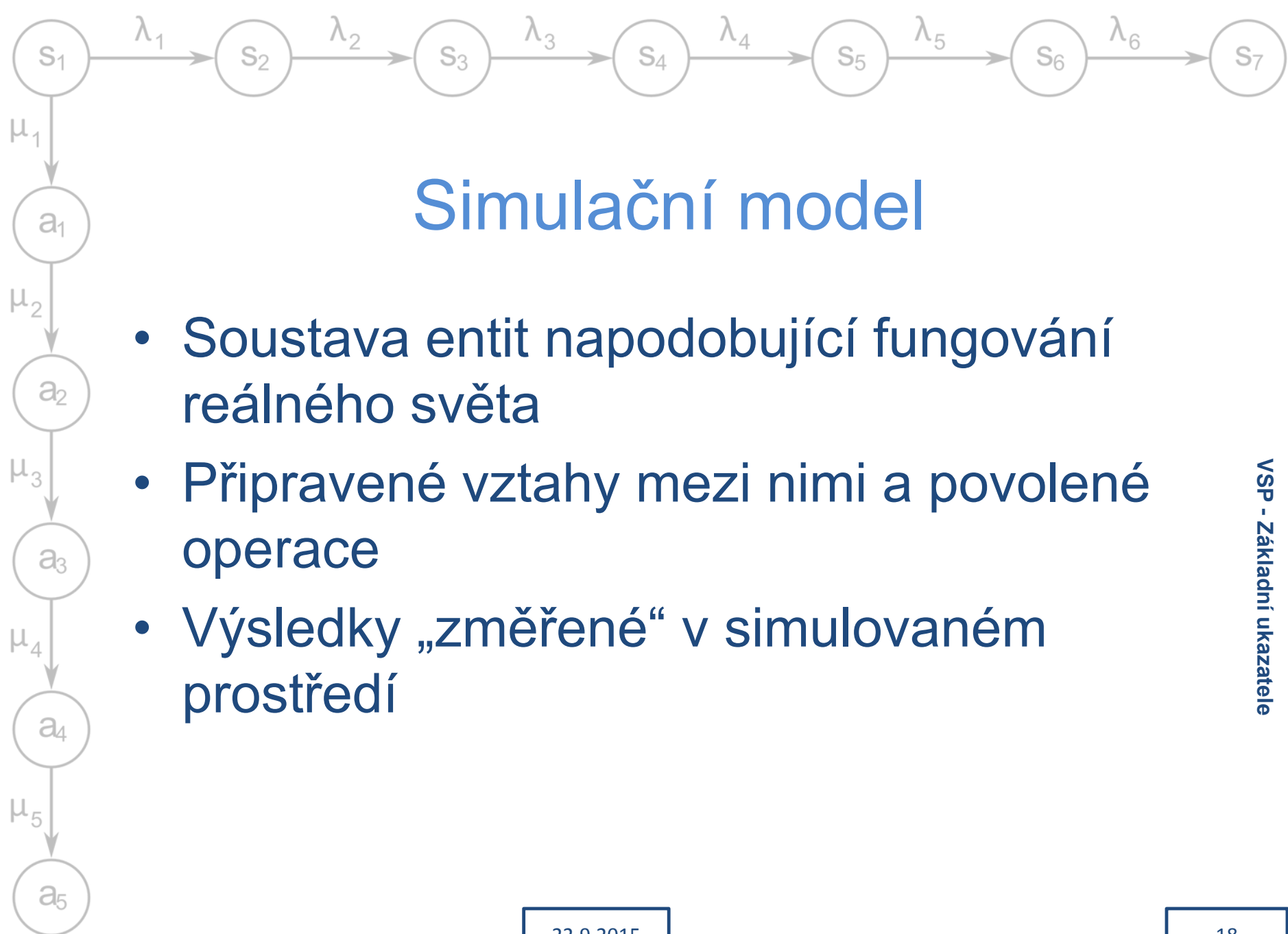
Model

- Zjednodušení systému pro jeho analýzu
 - Analytický
 - Simulační
- Důležité:
 - Verifikace – ověření že je model správně implementován
 - Validace – ověření že model odpovídá modelované realitě
- Jak moc lze model zjednodušit?



Analytický model

- „uzavřený tvar“ – popis systému vzorcem (typicky soustavou rovnic)
 - Např. newtonovská mechanika, eletrostatika ...
- Řešením rovnic získám výsledek
- Obtížné pro nestacionární systémy



Simulační model

- Soustava entit napodobující fungování reálného světa
- Připravené vztahy mezi nimi a povolené operace
- Výsledky „změřené“ v simulovaném prostředí



Srovnání



Analytický model

- + rychle získám výsledek
- + známé modely pro řadu problémů
- pro složité systémy obtížné sestavení / řešení

Simulační model

- + funguje dobře i pro složité systémy
- + Snadná nestacionarita
- složitější příprava
- delší výpočet

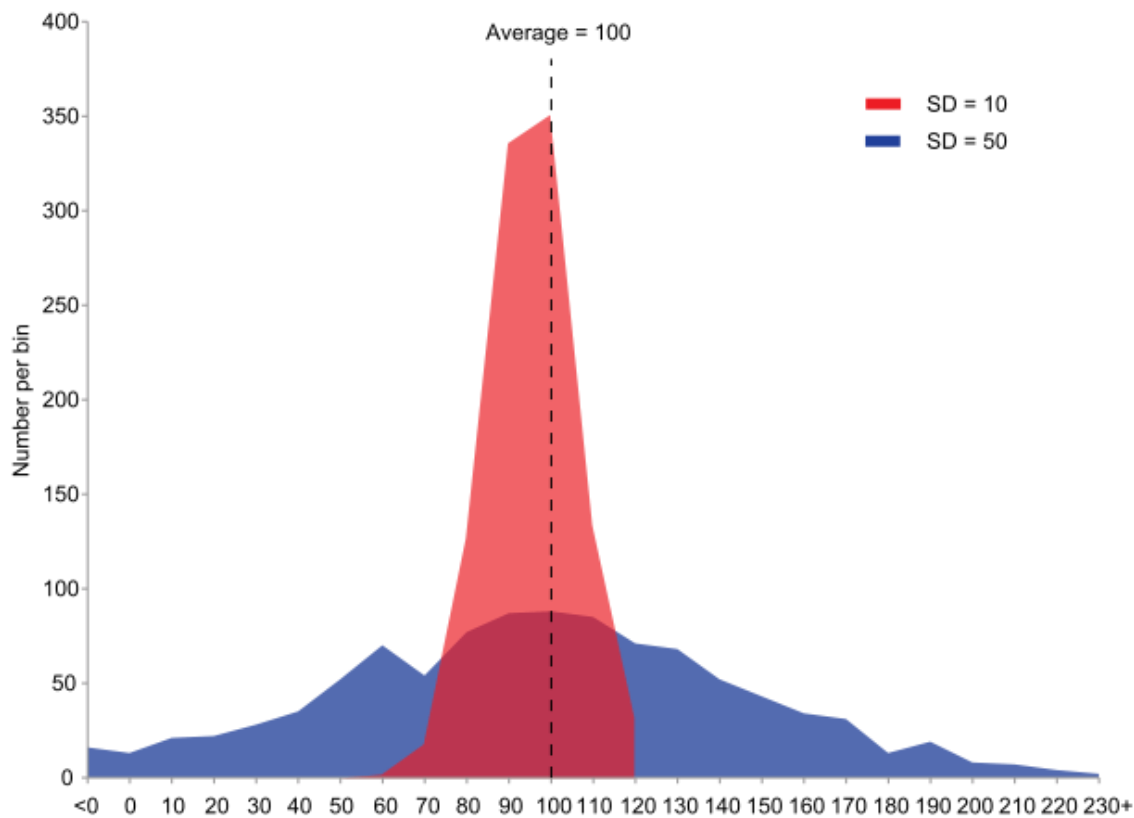


Trocha statistiky

- Náhodná veličina X
 - Měřitelná opakovaně v čase, zpracovávána statistickými metodami
- Charakteristiky
 - Střední hodnota – $D(X)$ (obvykle vážený průměr)
 - Rozptyl (střední kvadratická odchylka) - $D(X) = \sum_{i=1}^n [x_i - E(X)]^2 p_i$
 - Směrodatná odchylka - $\sigma = \sqrt{D(X)}$



Trocha statistiky – směrodatná odchyłka



VSP - Základní ukazatele



Trocha statistiky – rozdělení

- Pravidlo přiřazující každému jevu určitou pravděpodobnost
 - Diskrétní veličina
 - Spojitá veličina
- Lepší charakteristika než střední hodnota nebo odchylka, parametrizovatelné



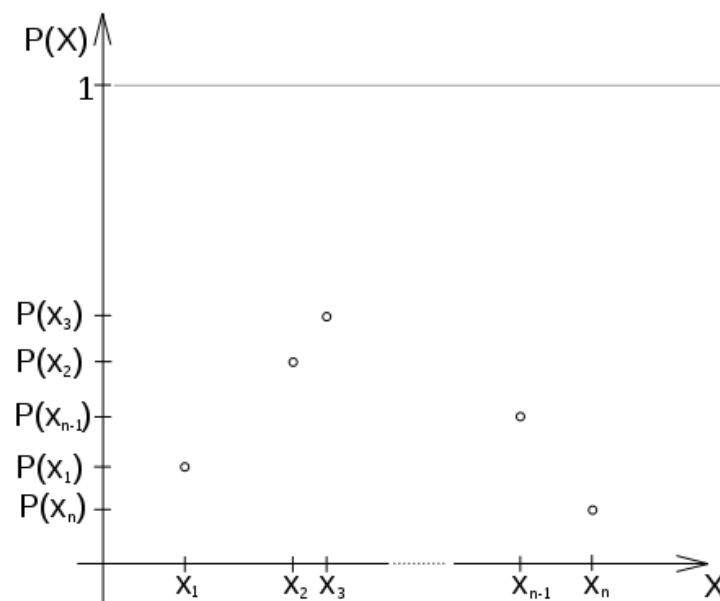
Diskrétní rozdělení

- Pravděpodobnostní funkce $P(x)$

$$\sum_{i=1}^n P(x_i) = 1$$

- Distribuční funkce $F(X) = P[X \leq x]$

$$= \sum_{t \leq x} P(t)$$



VSP - Základní ukazatele



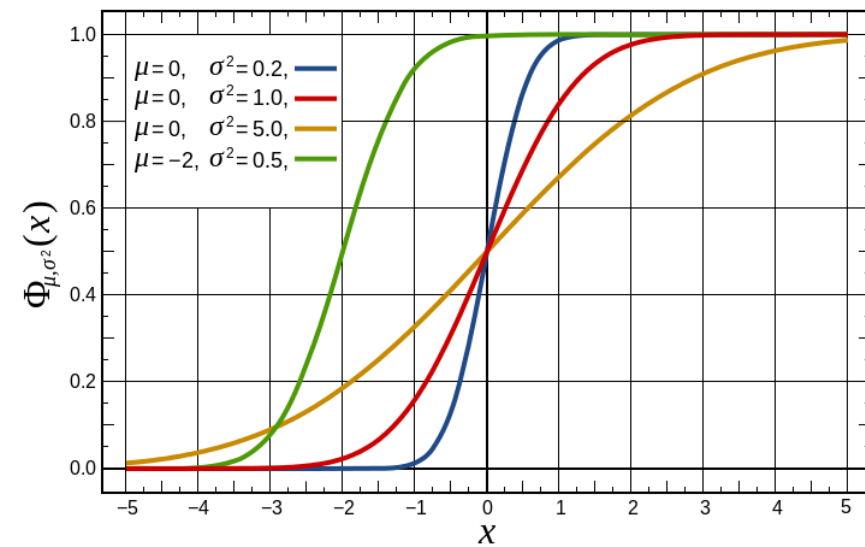
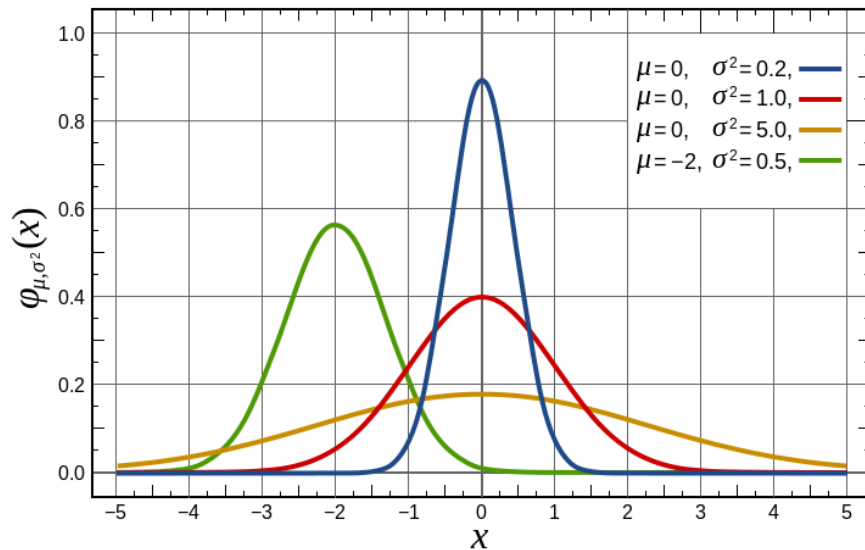
Spojité rozdělení

- Hustota pravděpodobnosti $\rho(x)$

$$\int_{\Omega} \rho(x) dx = 1, P[x_1 \leq X \leq x_2] = \int_{x_1}^{x_2} \rho(x) dx$$

- Distribuční funkce

$$F(X) = P[X \leq x] = \int_{-\infty}^x \rho(t) dt$$



- Hustota pravděpodobnosti

- Distribuční funkce

VSP - Základní ukazatele



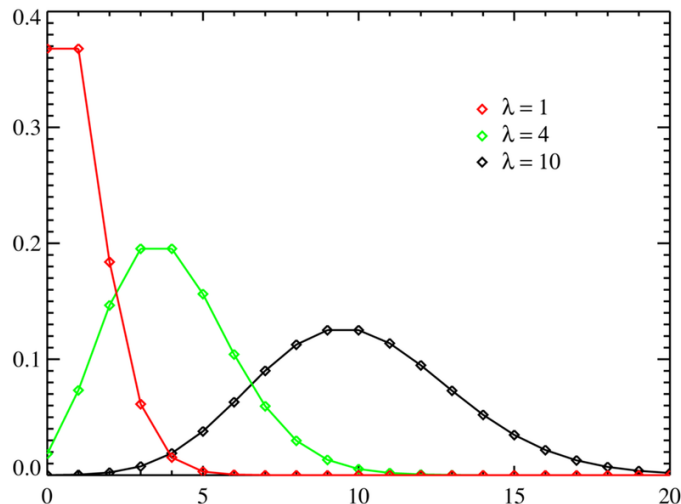
Poissonovo rozdělení

- Počet jevů v určitém časovém intervalu
 - Kolik požadavků přijde za 10 minut
 - Kolik aut projede za hodinu
 - ...
- Všechny výskyty jevu mají stejnou pravděpodobnost
- Parametr λ – střední počet událostí za jednotku času
- Ověřeno 1898 na počtu Pruských vojáků zabitých náhodným kopnutím koně

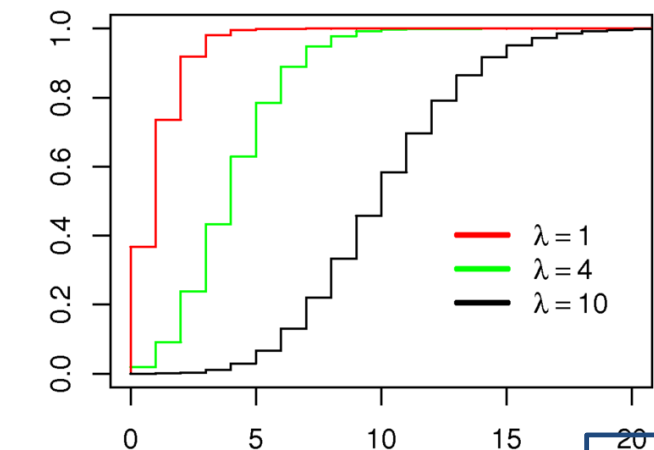




Poissonovo rozdělení



- Pravděpodobnostní funkce

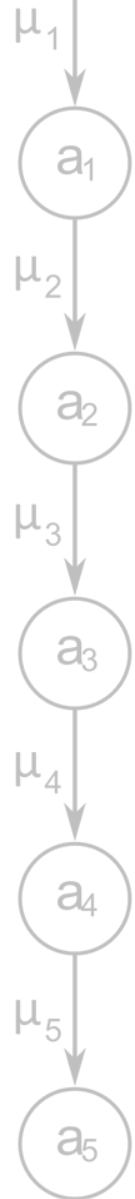


- Distribuční funkce



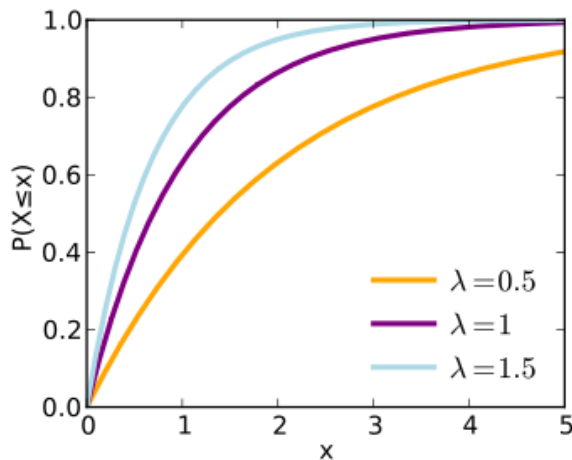
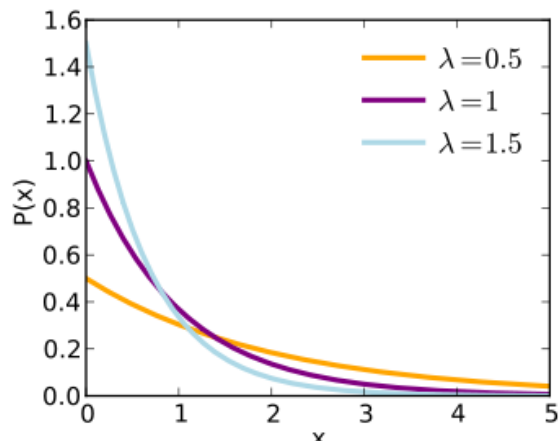
Exponenciální rozdělení

- Délka intervalu mezi dvěma událostmi (jejichž pravděpodobnost má Poissonovo rozdělení)
 - Jak dlouhá doba je mezi dvěma požadavky
 - Kolik času uplyne než projede další auto
 - ...
- Spojité rozdělení (události jsou diskrétní, ale čas mezi nimi je spojitý)
- Parametr λ – počet událostí za jednotku času (\rightarrow střední doba mezi událostmi je $\frac{1}{\lambda}$)





Exponenciální rozdělení



- Hustota pravděpodobnosti

- Distribuční funkce



Gaussovo (normální) rozdělení

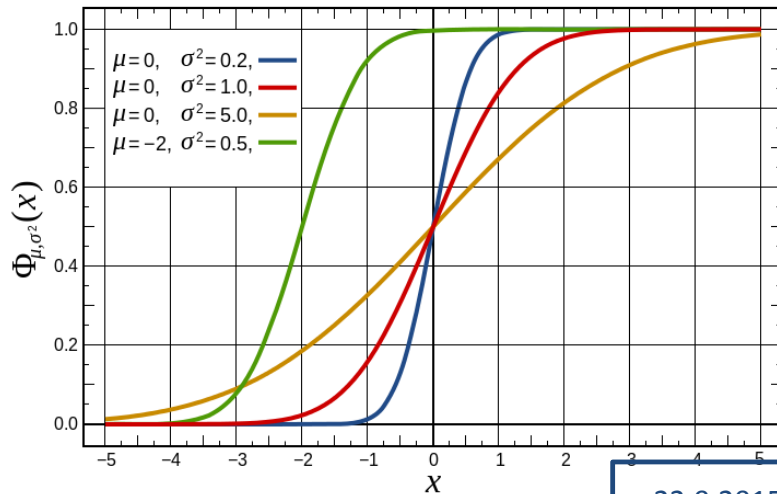
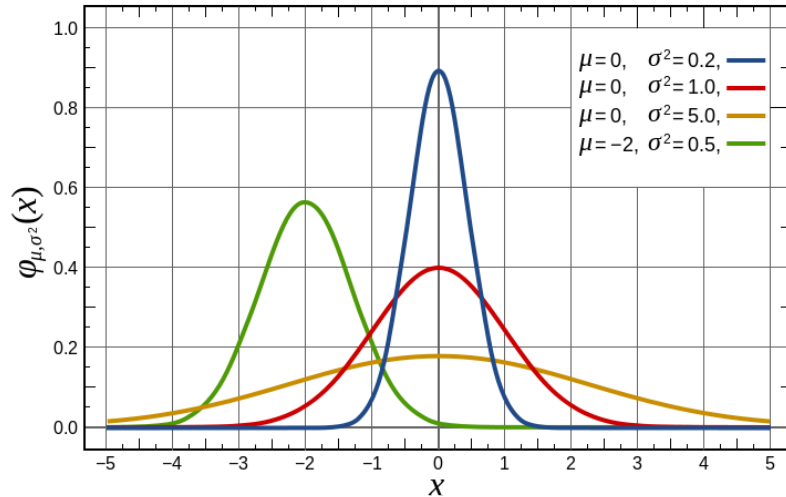
- Výsledek velkého množství různých jevů s různým rozdělením (někdy zkreslený)
 - Výška populace (a spousta dalších biologických ukazatelů – váha, rychlost růstu, délka končetin, krevní tlak, ...)
 - Odchyvky měření
 - Hodnocení (známky ve škole, skóre v IQ testech)
 - ...
- Aproximace řady jiných rozdělení





Gaussovo rozdělení

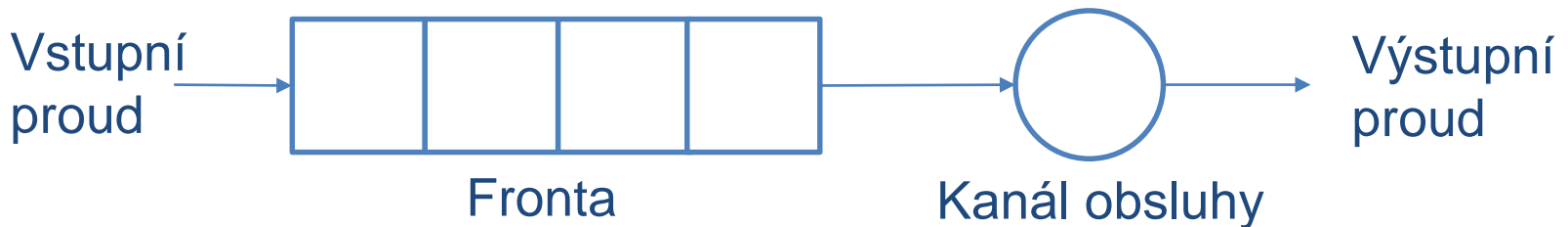
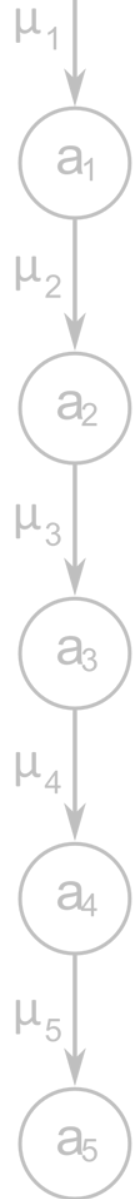
- Hustota pravděpodobnosti



- Distribuční funkce



Základní model systému – elementární SHO



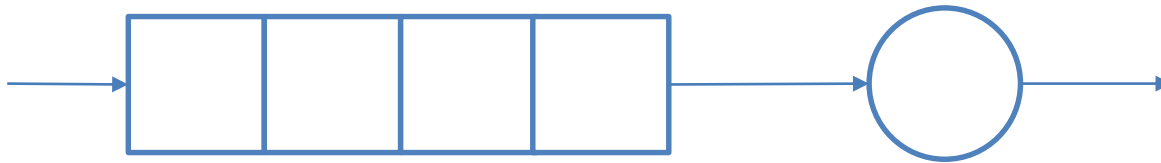
- Fronta kde čekají požadavky
– Konečná / nekonečná
- Server pro obsluhu požadavků
- Zátěž systému = proud přicházejících požadavků

VSP - Základní ukazatele



Základní model systému - zátěž

$$F_a(t) = P\{\tau \leq t\}$$



$$F_s(t) = P\{t_s \leq t\}$$

- Charakterizována
 - Četností příchodu požadavků
 - Rychlostí jejich obsluhy
- Obvykle náhodná veličina s daným rozdělením





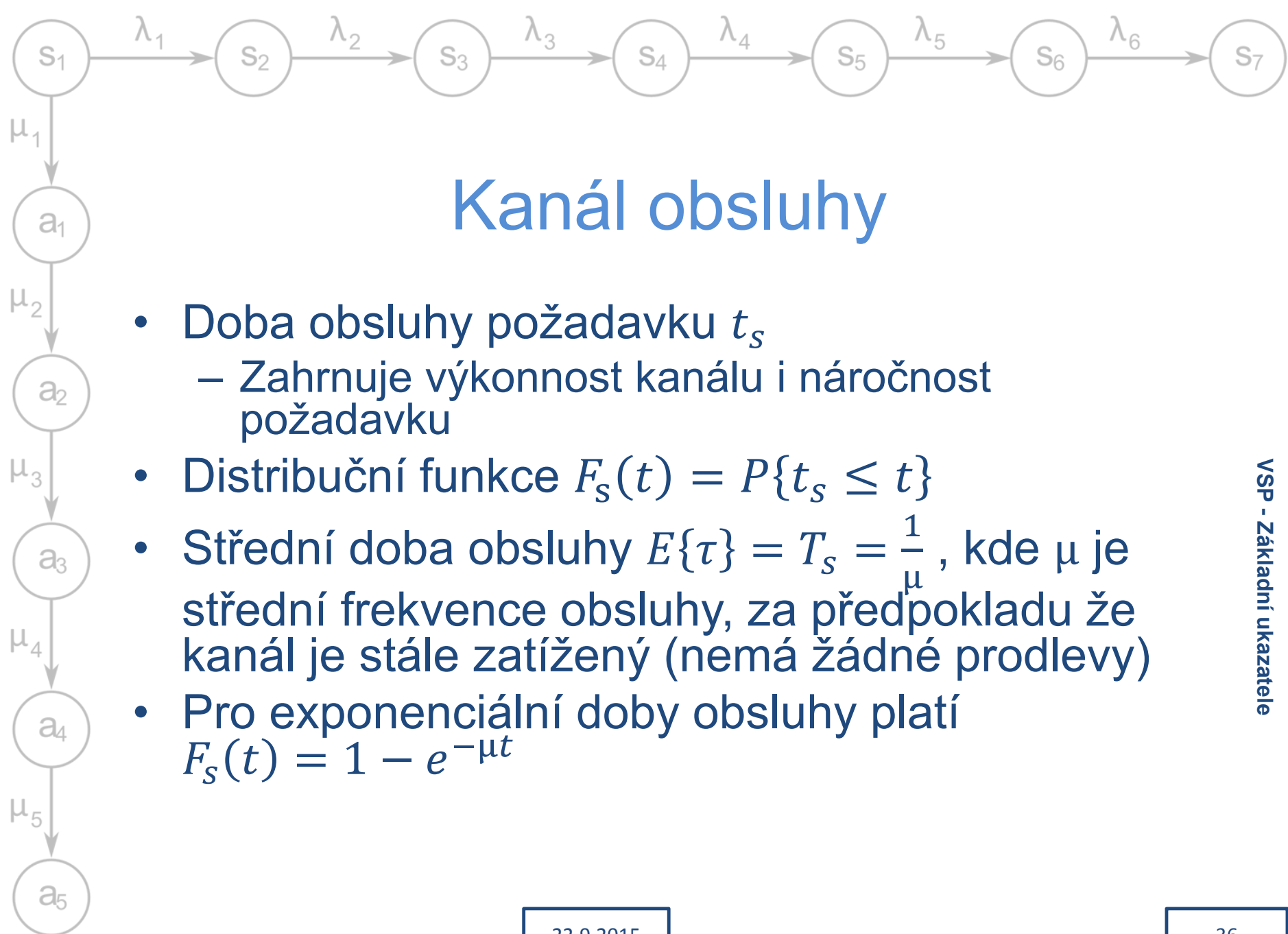
Příchody požadavků

- Distribuční funkce intervalu příchodů požadavků - $F_a(t) = P\{\tau \leq t\}$
- Interval příchodů (náhodná veličina) $\tau_k = t_k - t_{k-1}, k > 0$
- Střední hodnota intervalů mezi příchody (střední perioda příchodů) - $E\{\tau\} = T_a = \frac{1}{\lambda}$ (λ je střední frekvence příchodů požadavků)
- Pro poissonovský proud (exp. rozdělení) platí $F_a(t) = 1 - e^{-\lambda t}$, lze ho tedy popsat jedním parametrem



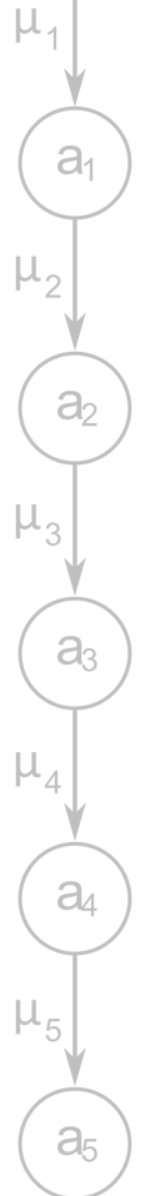
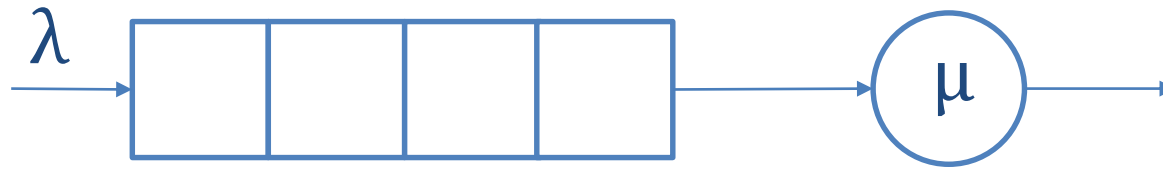
Fronta požadavků

- Popsána maximální délkou a způsobem vybírání (FIFO / LIFO ...)
- Okamžitý počet požadavků ve frontě (náhodná funkce času) - w
- Střední délka fronty - $E\{w\} = L_w$
- Doba čekání (náhodná veličina) - t_w
- Střední doba čekání požadavků ve frontě - $E\{t_w\} = T_w$

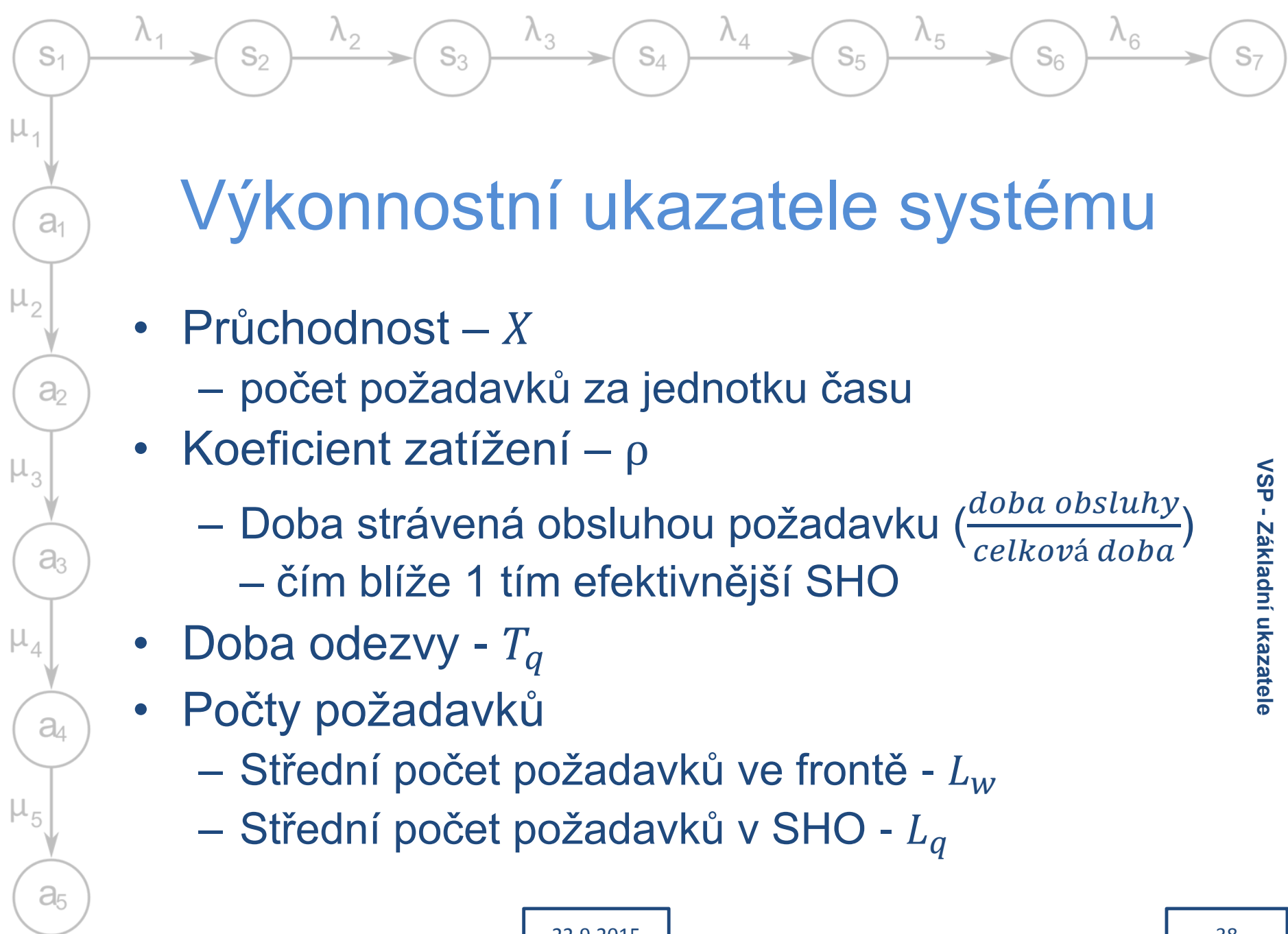




Model SHO



- Typický model SHO s poissonovským rozdělením pravděpodobnosti příchodů požadavků

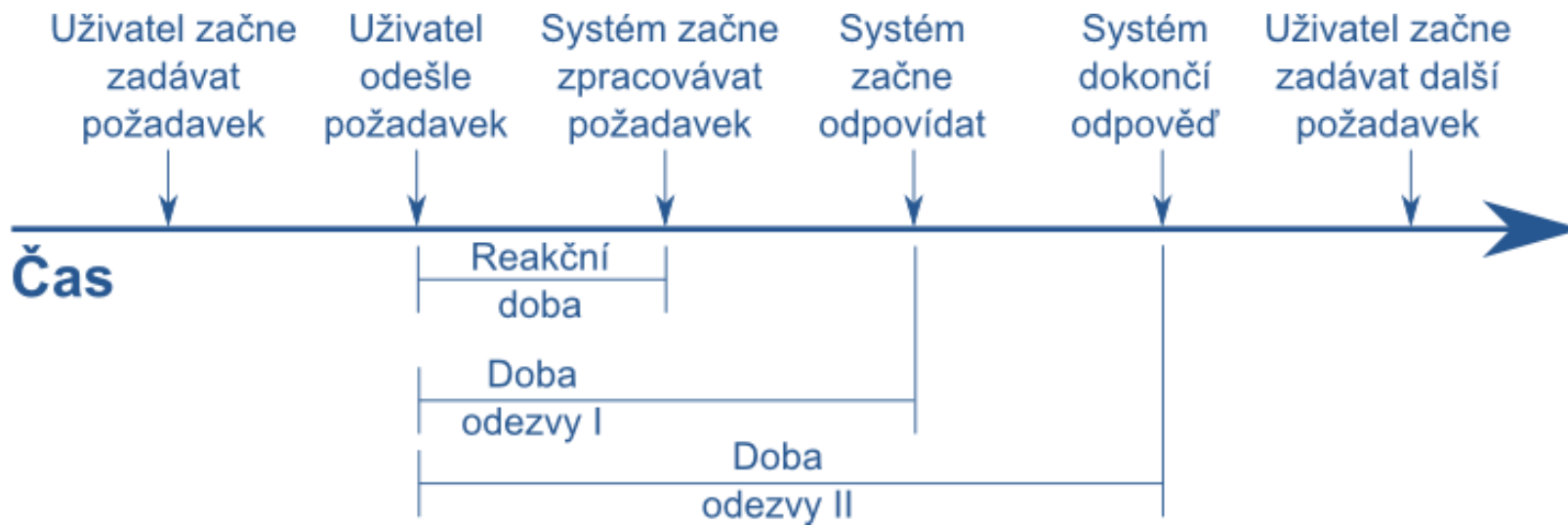


Výkonnostní ukazatele systému

- Průchodnost – X
 - počet požadavků za jednotku času
- Koeficient zatížení – ρ
 - Doba strávená obsluhou požadavku ($\frac{\text{doba obsluhy}}{\text{celková doba}}$)
 - čím blíže 1 tím efektivnější SHO
- Doba odezvy - T_q
- Počty požadavků
 - Střední počet požadavků ve frontě - L_w
 - Střední počet požadavků v SHO - L_q



Výkonnostní ukazatele systému



VSP - Základní ukazatele



Měření ukazatelů systému

- Alternativa k modelování
- Vzorkování na reálném systému
- Nezbytné k ověření platnosti modelu





Spolehlivost počítačových systémů

„Spolehlivost je obecná vlastnost objektu spočívající ve schopnosti plnit požadované funkce při zachování hodnot stanovených provozních ukazatelů v daných mezích a v čase podle stanovených technických podmínek“

ČSN 010102

- Je celkem obtížné stanovit jeden měřitelný číselný parametr odpovídající normě
- Zavádí se vyhodnocovatelné ukazovatele spolehlivosti





Základní pojmy



- Stav systému
 - Poruchový
 - Stálá porucha
 - Občasná porucha (bohužel častější)
 - Bezporuchový
- Provoz systému
 - Obnovovaný (= opravovaný)
 - Obnova = přechod z poruchového do bezporuchového stavu
 - Neobnovovaný



Spolehlivost 1 prvku

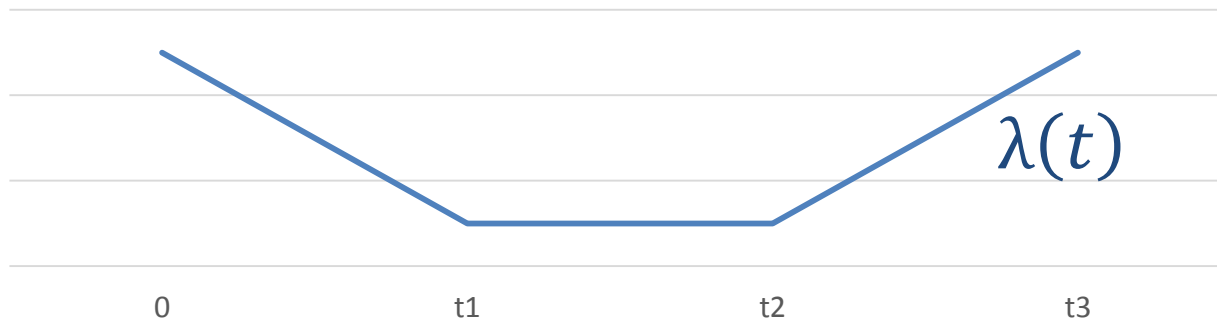
- Definovat hranice prvku
 - Klopný obvod, procesor, počítač ...
- Porucha jako náhodný jev
- Pravděpodobnost poruchy do času t – $Q(t)$ (=distribuční funkce)
- Pravděpodobnost bezporuchového stavu v čase t – $R(t) = 1 - Q(t)$
- Střední doba bez poruchy – T_s



Spolehlivost 1 prvku

- Obvykle se udává intenzita poruch - $\lambda(t)$ - podmíněná hustota poruchy v čase t pokud k poruše ještě nedošlo, nestacionární v čase

Intenzita poruch



- t_1 - týdny ($\approx 6 - 10$ týdnů)
- t_2 - roky (≈ 10 let)

VSP - Základní ukazatele



Spolehlivost více prvků

VSP - Základní ukazatele

- Záleží na jejich uspořádání
- Pokud jsou známy vlastnosti jednotlivých prvků, lze odvodit vlastnosti složeného systému ($Q(t)$, $R(t)$ a T_s)
 - Paralelní vs. sériové zapojení
- Určení vlivu záloh (více prvků stejného typu)
 - Studená záloha
 - Horká záloha (záložní systém aktivní, jen nevyužitý)





Opravované systémy

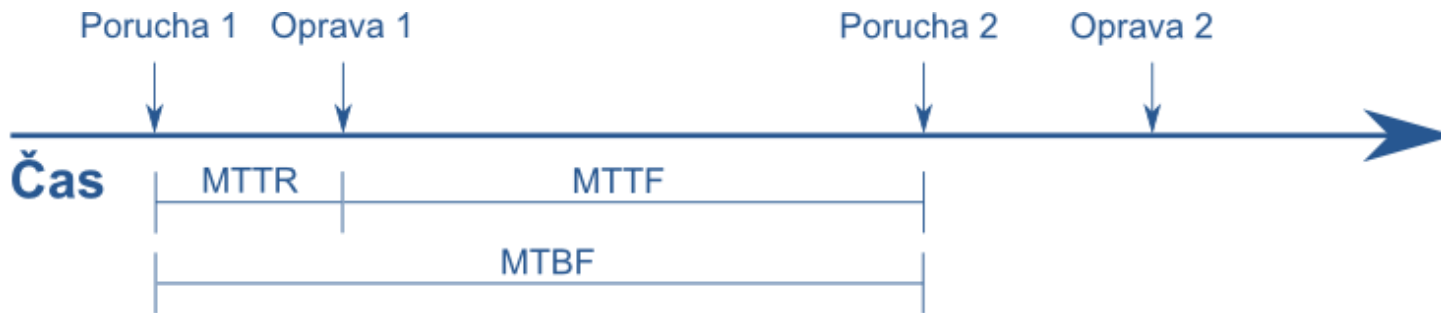
- U obnovovaného systému střídání poruchového a bezporuchového stavu
 - Intenzita poruch λ
 - Intenzita oprav μ

- Koeficient pohotovosti $K_p = \frac{t_p}{t_p+t_0} = \frac{\mu}{\mu+\lambda}$
 - Poměrná část provozuschopné doby v ustáleném stavu (stacionární systém)



Opravované systémy

- Střední doba mezi poruchami - $T_s = \frac{t_p}{n}$ (MTTF)
- MTBF - $T_s + \frac{T_0}{n}$ (MTTF + MTTR)



VSP - Základní ukazatele



Děkuji za pozornost

- Příště generátory náhodných čísel a metoda Monte Carlo