



Dimenzování moderních telekomunikačních sítí

Ing. Petr Hampl, Ph.D.

e-mail: petr.hampl@fel.cvut.cz

České vysoké učení technické v Praze
Fakulta elektrotechnická
Katedra telekomunikační techniky



Obsah

- Proč se zabývat dimenzováním?
- Dimenzování sítí s přepojováním okruhů:
 - model OS,
 - charakter vstupního toku,
 - Erlangův model.
- Dimenzování služeb v IP síti:
 - nástin problémů souvisejících s dimenzováním,
 - zobecněný Erlangův model,
 - analýza procesu přidělování přenosové kapacity,
 - příklad dimenzování:
 - bez rezervace,
 - s rezervací.



Proč se zabývat dimenzováním?

- Nutnost dodržet potřebné parametry jakosti obsluhy (GoS – Grade of Service), v požadovaných mezích tak, aby byly splněny požadavky na kvalitu provozované služby (QoS – Quality of Service).
- Obecný přehled sledovaných parametru lze nalézt v doporučení ITU-T E.802, např.:
 - doba sestavení spojení,
 - zpoždění,
 - počet poruch na přípojku,
 - počet reklamací,
 - doba vyřízení stížnosti,
 - ...



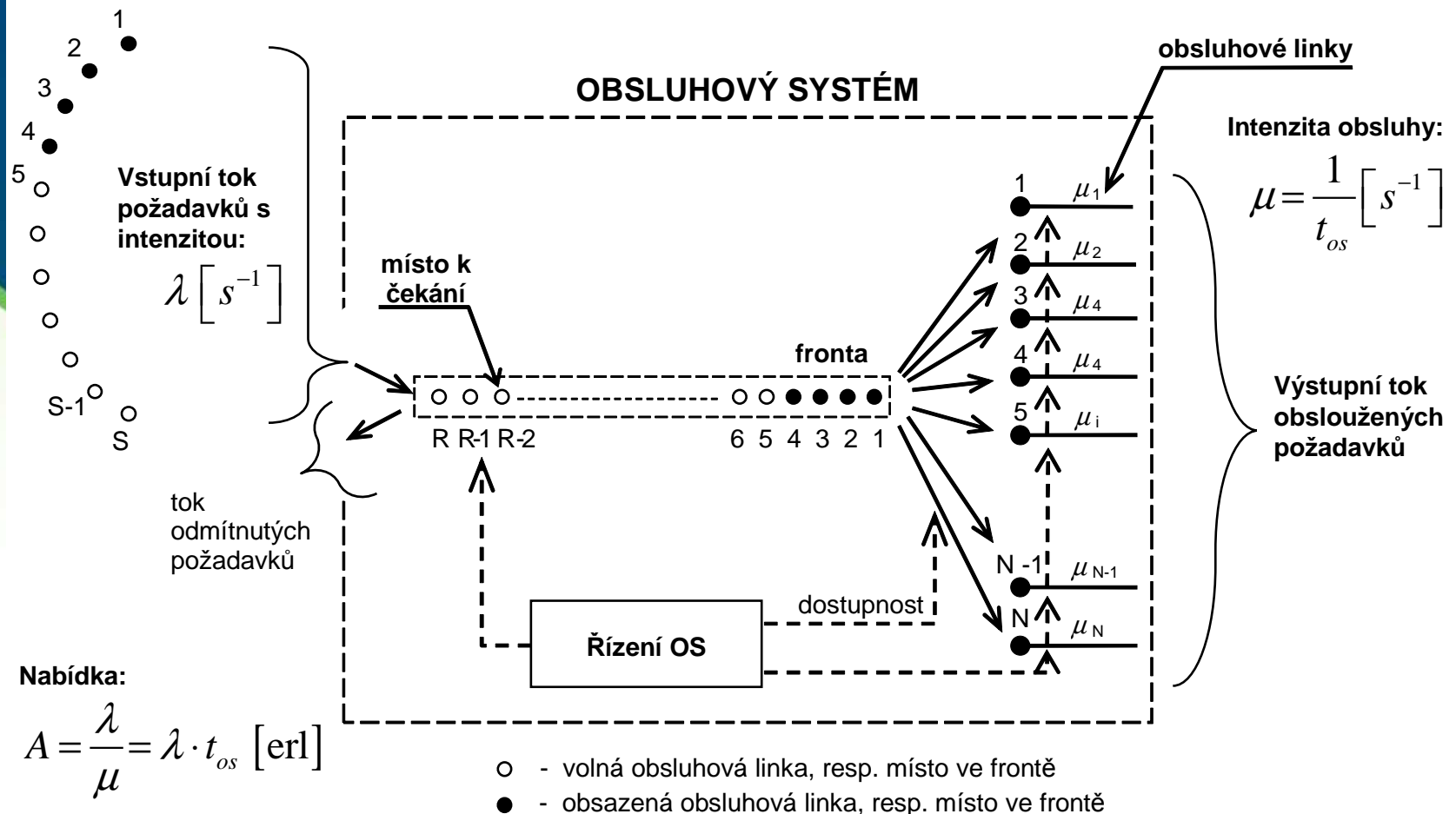
Dimenzování sítí s přepojováním okruhů



Model obsluhového systému

- Při dimenzování se vychází z klasické teorie provozního zatížení s následujícím uspořádáním modelu obsluhového systému:

zdroje





Klíčové předpoklady použití

- Na vstupu systému je homogenní poissonovský tok:
 - tj. požadavky přicházejí individuálně, náhodně a vzájemně nezávisle;
 - pravděpodobnost příchodu k požadavků v intervalu délky T je dána Poissonovým rozložením:

$$P_k(T) = \frac{(\lambda T)^k}{k!} e^{-\lambda T}; k = 0, 1, \dots$$

- rozložení intervalů mezi příchody je dáno distribuční funkcí:

$$F(t) = 1 - e^{-\lambda t}; t \geq 0$$
$$= 0; \quad t < 0$$

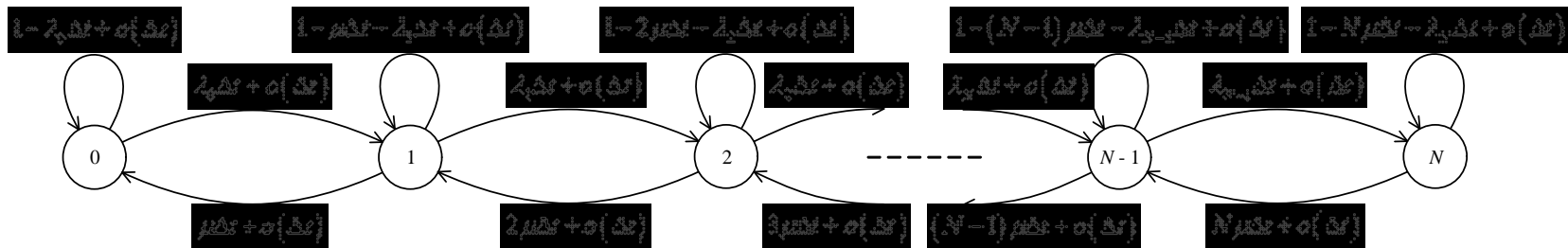
- s touto myšlenkou přišel již v roce 1909 Agner Krarup Erlang:
"The Theory of Probabilities and Telephone Conversations,,.





Erlangův model systému se ztrátou $M/M/N/0$:

- Stavový diagram modelu:



- Pravděpodobnost ztráty volání svazku je dána prvním Erlangovým vztahem B -formula - publikován roku 1917:

$$B = P_N = \frac{A^N}{\sum_{i=0}^N \frac{A^i}{i!}} = E_{1,N}(A)$$

-Vzhledem k předpokladu pevné kapacity přenosového kanálu má model v IP sítích omezené použití.



Dimenzování sítí s přepojováním okruhů

- Proces dimenzování spočívá ve stanovení :
 - potřebného počtu okruhů N ,
 - nebo kapacity fronty R , počtu agentů S v případě call centra, tak, aby byly splněny požadavky na jakost obsluhy, např.:
 - pravděpodobnost ztráty B příchozího volání (dimenzování svazků):
zpravidla $B \leq 1 \%$;
 - pravděpodobnost čekání na operátora $P\{ W \leq t \}$:
$$P\{ W \leq 30 \text{ s} \} = 0,8.$$
 - doba zdržení v systému,
 -
- Předpokládá se, že existuje **pouze jeden** typ služby s požadavky jenž obsadí kapacitu jednoho kanálu v systému po celou dobu obsluhy, např.:
- 64 kbit/s u systému PCM,
- 33,854 kbit/s (jeden Time Slot) u systému GSM,
-

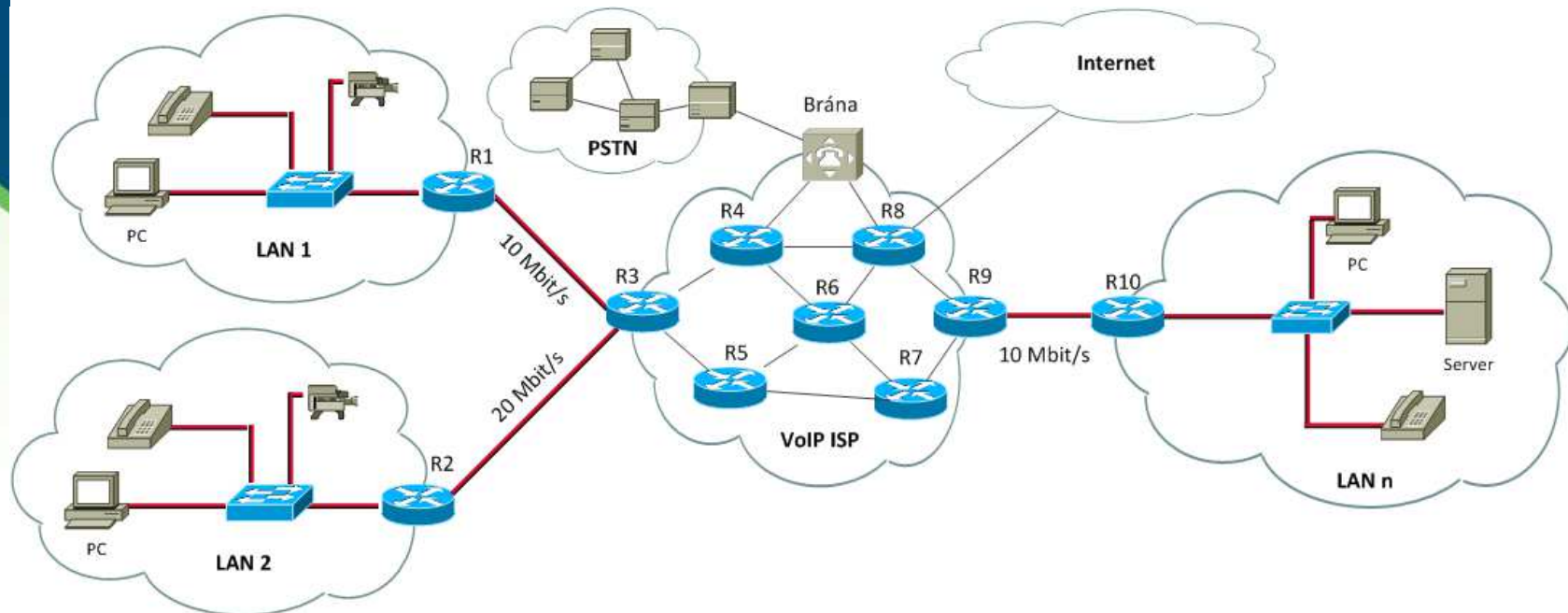


Dimenzování služeb v IP síti



Dimenzování služeb v IP síti

- Charakter chování zákazníků užívajících konverzační služby typu VoIP, videotelefonie, ... zůstává „téměř“ stejný, v závislosti na typu služby se mění pouze parametry.
- Dochází však ke změně přenosové technologie (protokol IP).



- Koexistence služeb konverzačních a interaktivních. Pro jejich vzájemné oddělení se užívá klasifikace do různých tříd.



Toky v IP sítích

- Typy služeb a jim odpovídající toky v síti:
 - konverzační (realtime) hlas, video, SSH, telnet ...
 - video na požádání (Video on Demand - VoD)
 - mission-critical aplikace (databázové systémy, ...)
 - data směrovacích protokolů (OSPF, BGP, RIP, ...)
 - data managementu sítě (SNMP, ...)
 - ostatní (bulk transfer, best-effort)
 - nežádoucí provoz (scavenger: less-than-best-effort)
- Rozdílné požadavky na:
 - přenosovou rychlost,
 - zpoždění, rozptyl zpoždění,
 - ztrátu paketů,
 - zabezpečení,
 -



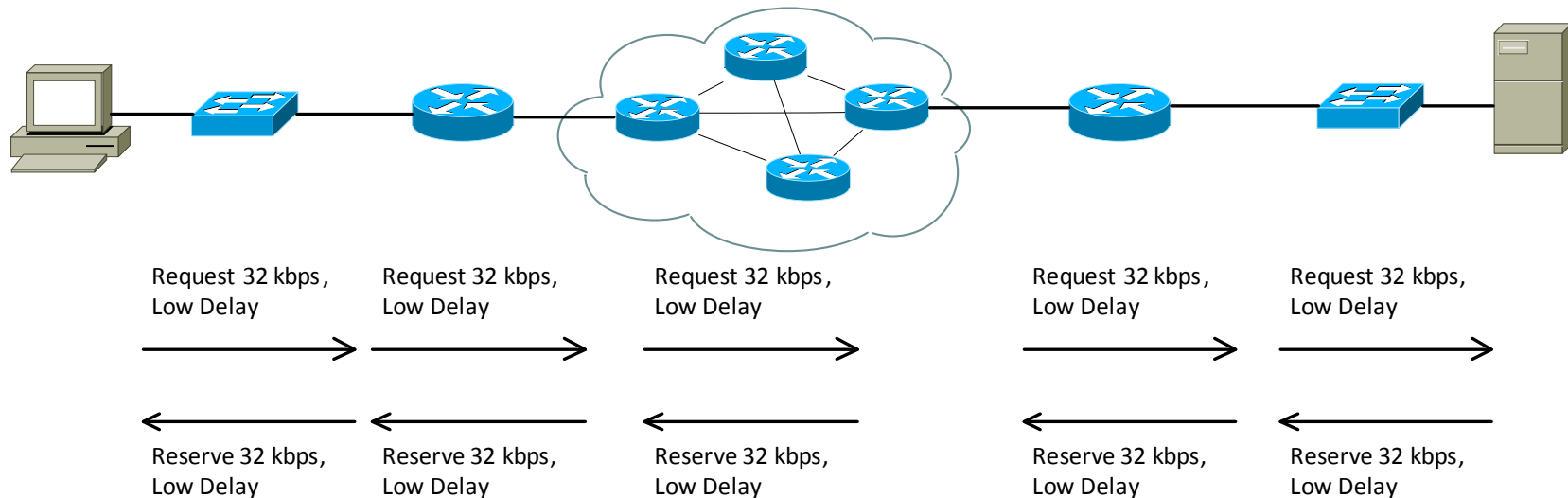
Dostupné modely QoS v IP sítích

- Best-Effort Service
 - síť negarantuje doručení do cílového uzlu sítě, není zaručena žádná úroveň kvality služeb
- Integrated Services (IntServ) RFC 1633
 - využívá Resource ReSerVation Protocol (RSVP) - RFC 2205
 - obtížně aplikovatelné na rozlehlé sítě (musí podporovat **všechny** uzly sítě)
 - detailní nastavení QoS pro každou službu (End-to-End)
- Differentiated Services (DiffServ) RFC 2474
 - pakety jednotlivých služeb jsou roztríděny (klasifikovány) do různých tříd, v každém uzlu sítě jsou stanovena pravidla jak s danou třídou naložit (Per-Hop Behaviours - PHBs)
 - využívá pole Differentiated Services Code Point (DSCP) v záhlaví IP paketů
 - aplikovatelné na rozlehlé sítě (Internet)



Integrated Services (IntServ)

- IntServ model vyžaduje:
 - možnost rezervace přenosových prostředků a možnost omezení intenzity jednotlivých služeb (Admission Control)

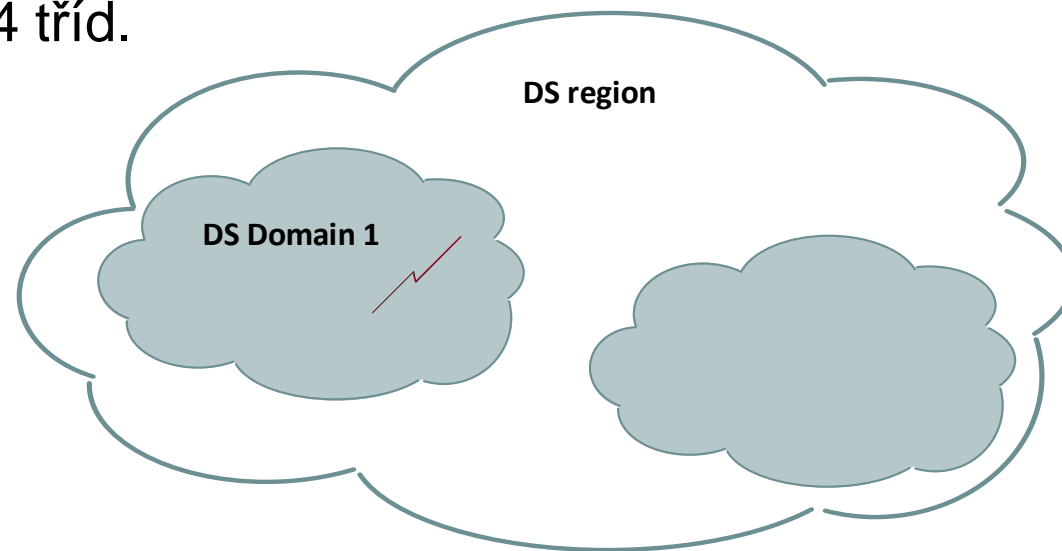


- Každý směrovač musí udržovat řídicí informace o všech spojeních \Rightarrow neaplikovatelné na rozlehlé sítě



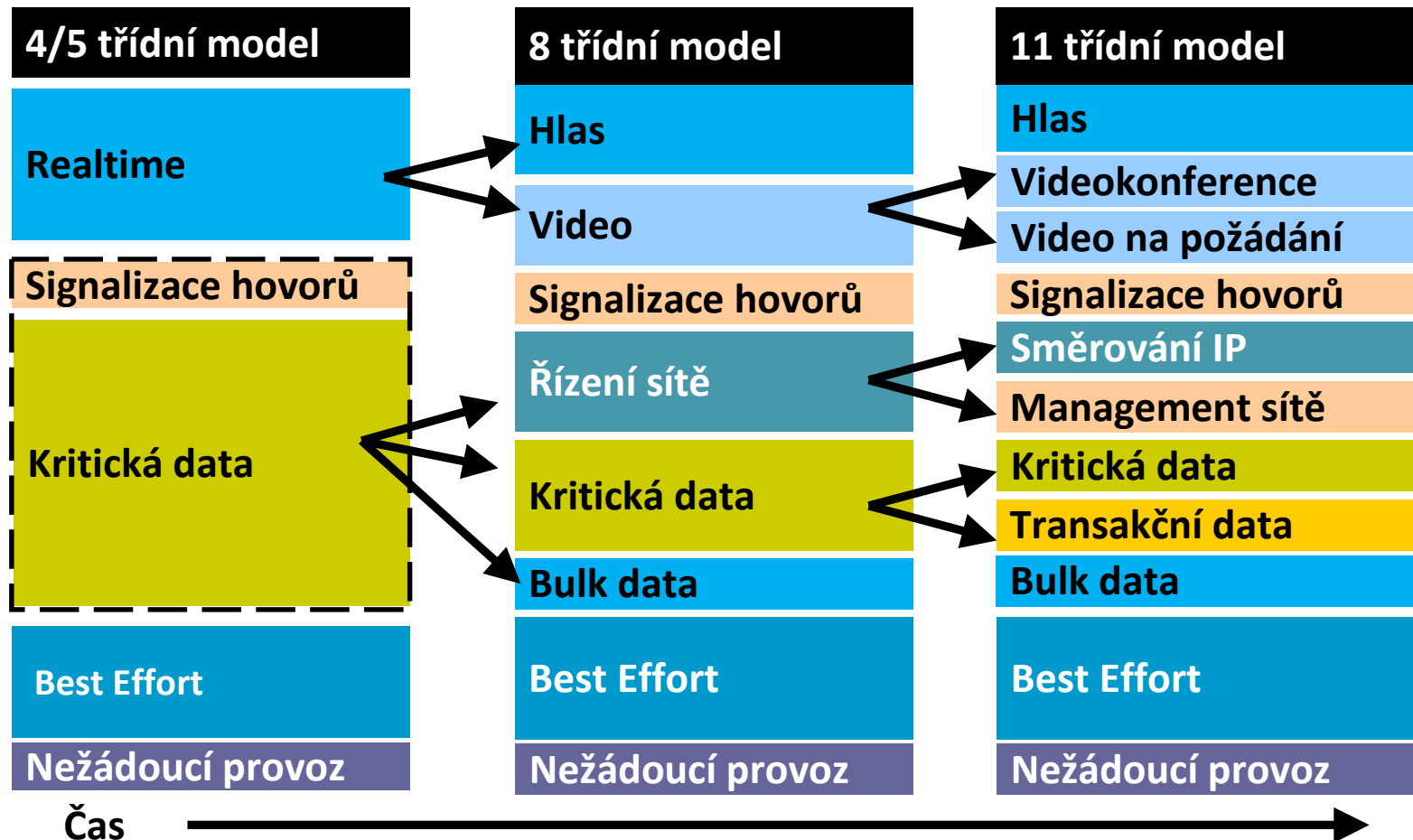
Differentiated Services

- Značkování se provádí na 3. vrstvě OSI modelu - záhlaví IP paketu, (záhlaví 2. vrstvy je nevhodné, mění se na každé trase)
- Paket může být mezi jednotlivými sítěmi (DS doménami) přeznačkován.
- Pro značkování bylo modifikováno záhlaví ToS protokolu IP: šestibitové pole DSCP-Differentiated Services Code Point, $2^6 = 64$ tříd.





Který model klasifikace zvolit?





Problémy které je nutné při dimenzování řešit?

- **Nedostatečná přen. rychlost:** vede k zhoršení parametrů (narůstající zpoždění, ztráta, ...)
- **Celková doba přenosu sítí:** pakety prochází mnoha síťovými prvky (směrovače, prepínače, ...), kde dochází k jejich zdržení.
- **Rozptyl doby zdržení (jitter):** doba zdržení kolísá v závislosti na aktuálním stavu (zatížení) jednotlivých rozhraní a síťových prvků.
- **Ztráta paketů:** V případě, že intenzita příchozích paketů přesáhne dostupnou kapacitu odchozího rozhraní dochází ke ztrátě paketů.



Přenosová rychlost a odhad jejího výpočtu

$$v_p [kbit/s] = \frac{\text{velikost paketu [bytů]} \times 8}{1000} \times \text{intenzita paketů} \left[\frac{\text{paket}}{s} \right]$$

- Velikost paketu:

- užitečná data v jednom paketu [bytů] (dáno použitým kodekem),
- **L2 záhlaví** (Ethernet: 18 bytů, FrameRelay: 6 bytů),
- **L3 záhlaví** (IPv4 20 bytes),
- **další záhlaví** (UDP, RTP, IPsec, MPLS značka),

- Příklad: kodek G.729, **L2(ethernet)**, **≥L3 (IPv4,UDP,RTP)**

- Velikost paketu=**20+18+40**=78 bytů.
- Intenzita odesílaných paketů = 1 / (20 ms) = 50 paketů/s.

$$v_p [kbit/s] = \frac{78 [bytů] \times 8}{1000} \times 50 [paketů/s] = 31,2 kbit/s$$



Parametry některých kodeků

Kodek	Algoritmus	MIPS	CBR [kbit/s]	NEB [kbit/s]	MOS
G.711	PCM		64	87,2	4,1
G.726	ADPCM	1	32	55,2	3,85
G.728	LD-CELP	30	16	31,5	3,61
GSM 06.10	RPE-LP	10	13		3,5
G.729A	CS-ACELP	11	8		3,7
G.729	CS-ACELP	20	8	31,2	3,92
G.723.1	MP-MLQ	16	6.3	21,9	3,9
G.723.1	ACELP	20	5.3	20,8	3,65

MIPS - Million Instructions Per Second (závislé na architektuře procesoru),

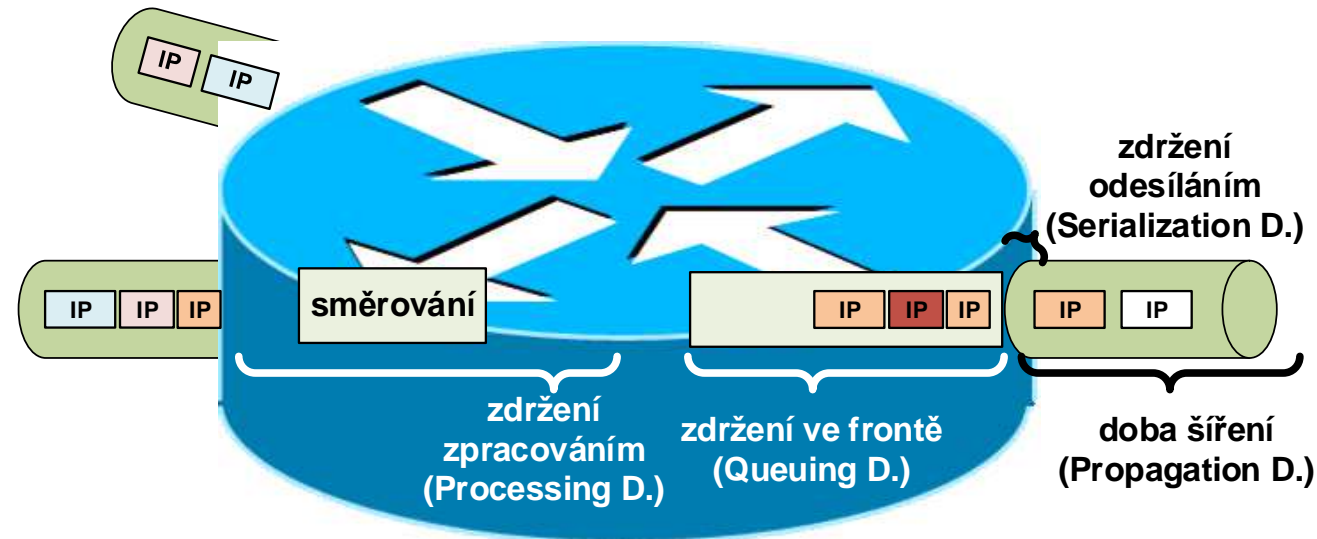
CBR - Codec Bit Rate,

NEB - Nominal Ethernet Bandwidth

MOS- Mean Opinion Score



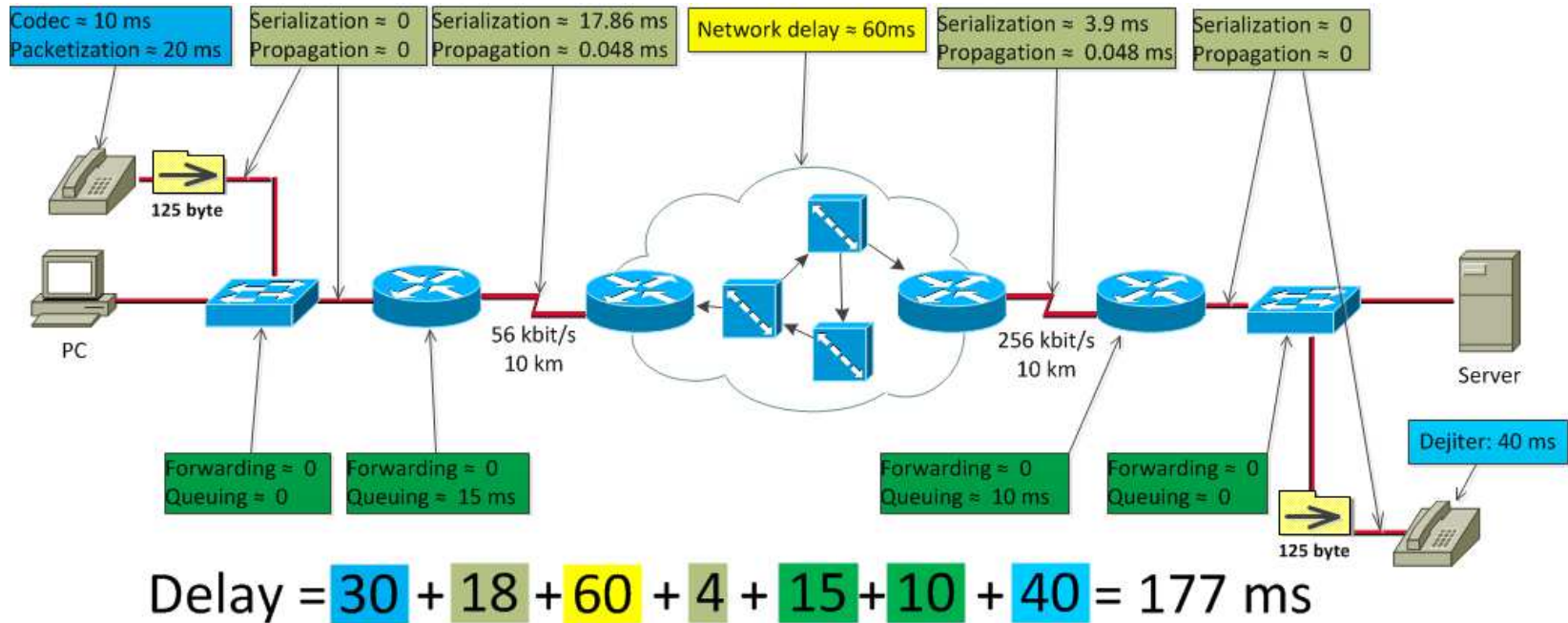
Zpoždění v IP síti



- **Proměnné – dáno síťovým prvkem (směrovač, přepínač)**
 - **Zdržení zpracování (Processing delay):** Čas potřebný pro analýzu jednotlivých záhlaví a přesun ze zásobníku (paměti) vstupního rozhraní do zásobníku výstupního rozhraní.
 - **Zdržení ve frontě (Queuing delay):** Zpoždění způsobené čekáním paketu určité priority.
- **Fixní – dáno přenosovým médii (ethernet, sériové rozhraní, DSL,)**
 - **Serializační zpoždění (Serialization delay):** Čas potřebný pro odeslání bitové posloupnosti (například ethernetového rámce) na fyzické rozhraní o dané přenosové rychlosti.
 - **Doba šíření signálu přenosovým médii (Propagation delay)**



Celková doba přenosu IP sítí

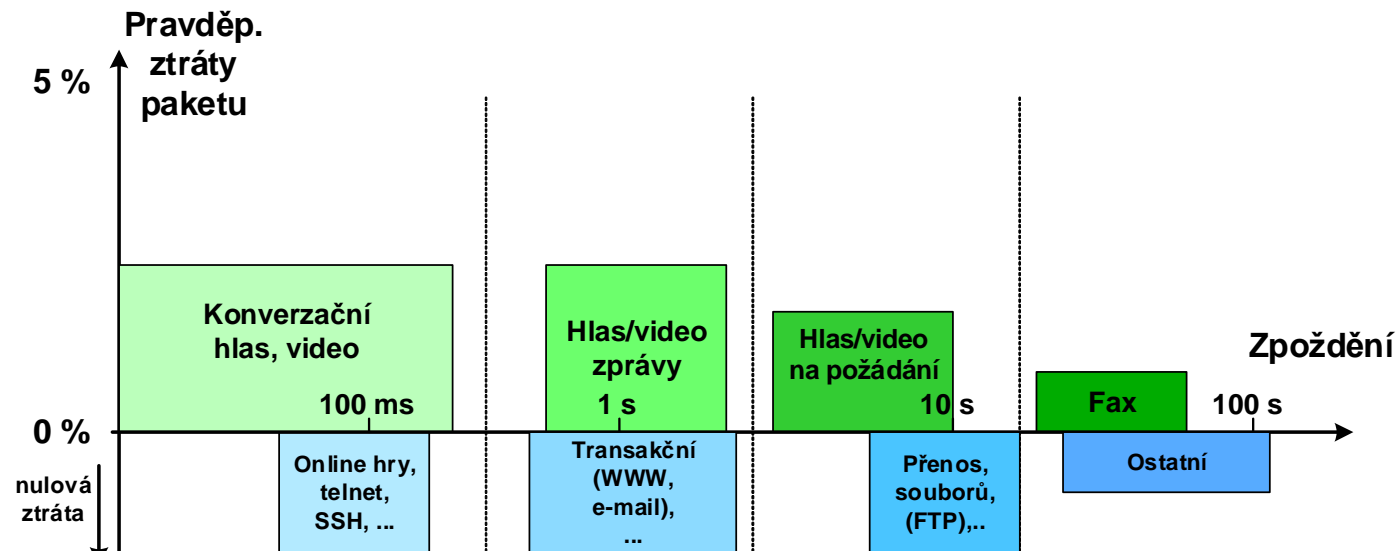


- **Zpoždění konec-konec (End-to-End)** - součet všech zpoždění.
(zdržení zpracování + zdržení ve frontě) + (serializační zpoždění + doba šíření)
- **Rozptyl celkové doby zpoždění (End-to-end).**



Orientační hodnoty zpoždění a ztráty paketů

- Z pohledu standardizačních institucí (ITU-T, IETF, 3GPP, ...) jednotlivé služby většinou dělí do následujících kategorií:
 - Konverzační (conversational): přenos multimediálních toků v reálném čase (videotelefon, iptelefon, SSH, telnet...);
 - Proudová (streaming): přenos multimediálních toků v reálném čase (video na požádání, iVysílání, YouTube,...);
 - Interaktivní (interactive): WWW, e-mail, přenos souborů, ...
 - Ostatní (background): e-mail, přenos souborů na pozadí, ...



ITU-T G.1010 – požadavky QoS z pohledu uživatele



Jak minimalizovat zpoždění a jeho rozptyl?

- **Zvýšením přenosových rychlostí jednotlivých rozhraní:**
 - nejjednodušší
 - nákladné
 - minimalizuje se serializační zpoždění
- **Prioritizací vybraných toků:**
 - U upřednostňovaných služeb (hlas, video, SSH, telnet) se situace zlepšuje na úkor ostatních služeb (HTTP, FTP, Bittorrent,)
- **Kompresí záhlaví druhé nebo třetí vrstvy:**
 - vhodné použít pro „malé“ pakety nebo rámce \Rightarrow vyšší efektivita
 - u rozhraní s nízkou přenosovou rychlostí výrazně klesne serializační zpoždění
 - vznikne nový typ zpoždění způsobený vlastní kompresí a dekompresí

např.: cRTP (RFC2508) 40 bytů \rightarrow 3-5 bytů
- **Fragmentací s prokládáním (zavedením priorit)**
 - sníží se efektivita přenosu

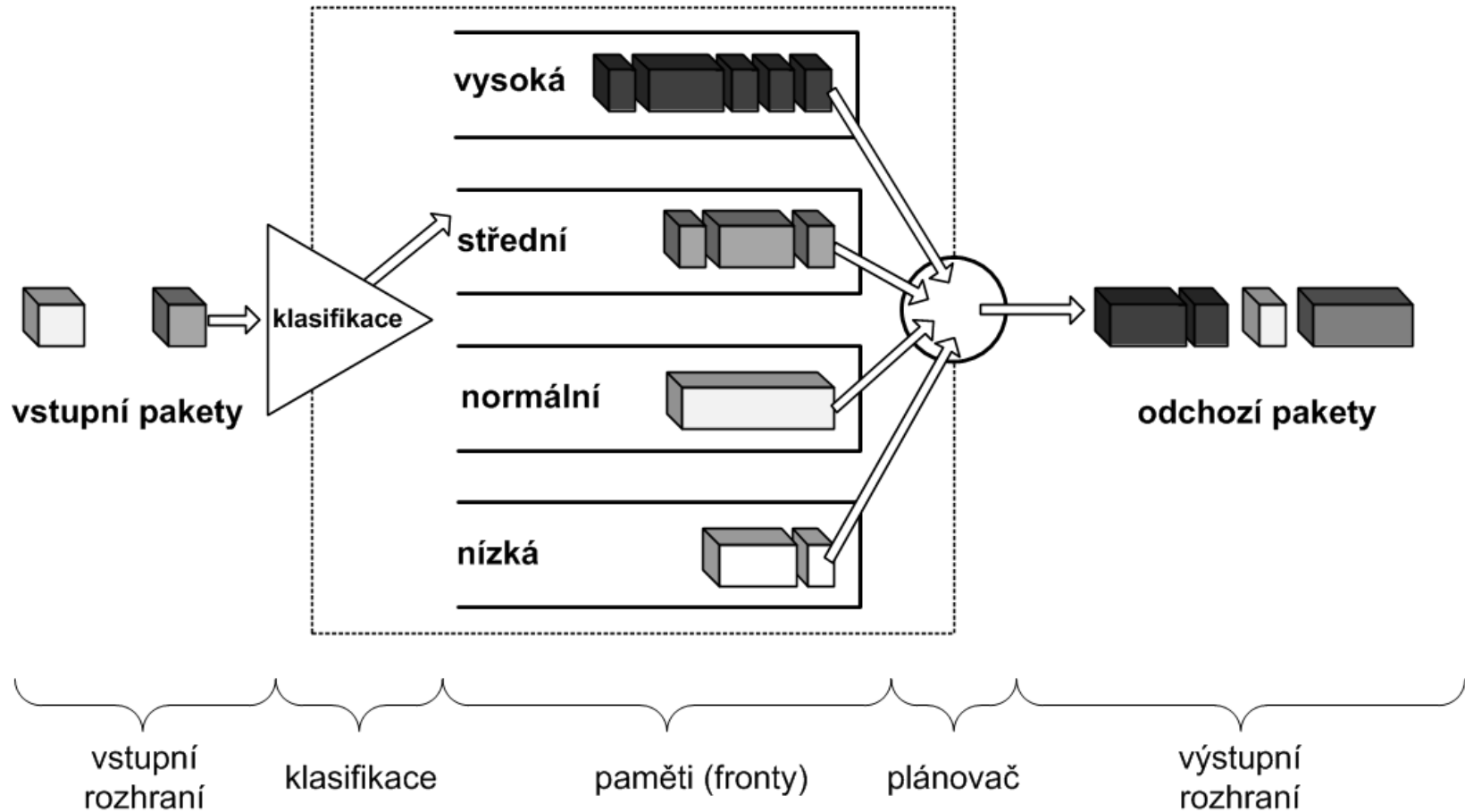


Zavedení systému s prioritami

- Požadavky na systém obsluhující služby konverzačního charakteru:
 - systém s prioritami (garance zpoždění prioritní třídy),
 - rezervace přenosové rychlosti toku nejvyšší priority z celkové přenosové rychlosti rozhraní (tj. ochrana ostatních služeb).
- Technické prostředky - směrovače umožňují zavést následující režimy fronty:
 - First In – First Out (FIFO),
 - **Priority Queuing (PQ)**,
 - Custom Queuing (CQ),
 - Weighted Fair Queuing (WFQ),
 - Class-Based Weighted Fair Queuing (CBWFQ),
 - **Low Latency Queuing (LLQ)**.



Priority Queuing



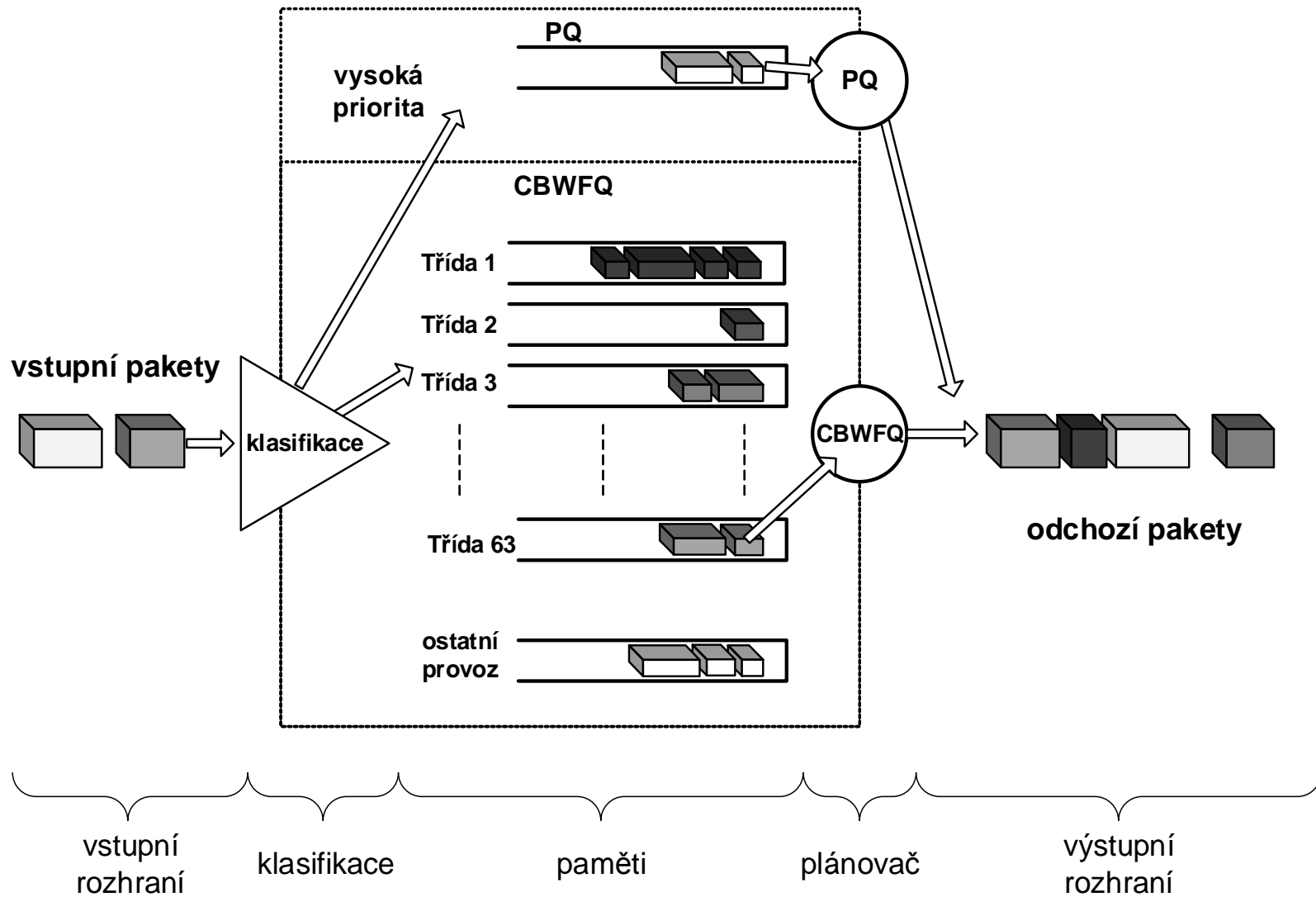


Priority Queuing

- čtyři prioritní třídy: vysoká (high), střední (medium), normální (normal), a nízká (low)
- pakety zařazené do třídy s vyšší prioritou mají vždy přednost před pakety z třídy s nižší prioritou
- pakety s nejvyšší (high) prioritou mají vždy přednost => mohou získat 100% přenosové kapacity, minimalizuje se zpoždění i rozptyl na úkor ostatních.



Low Latency Queuing



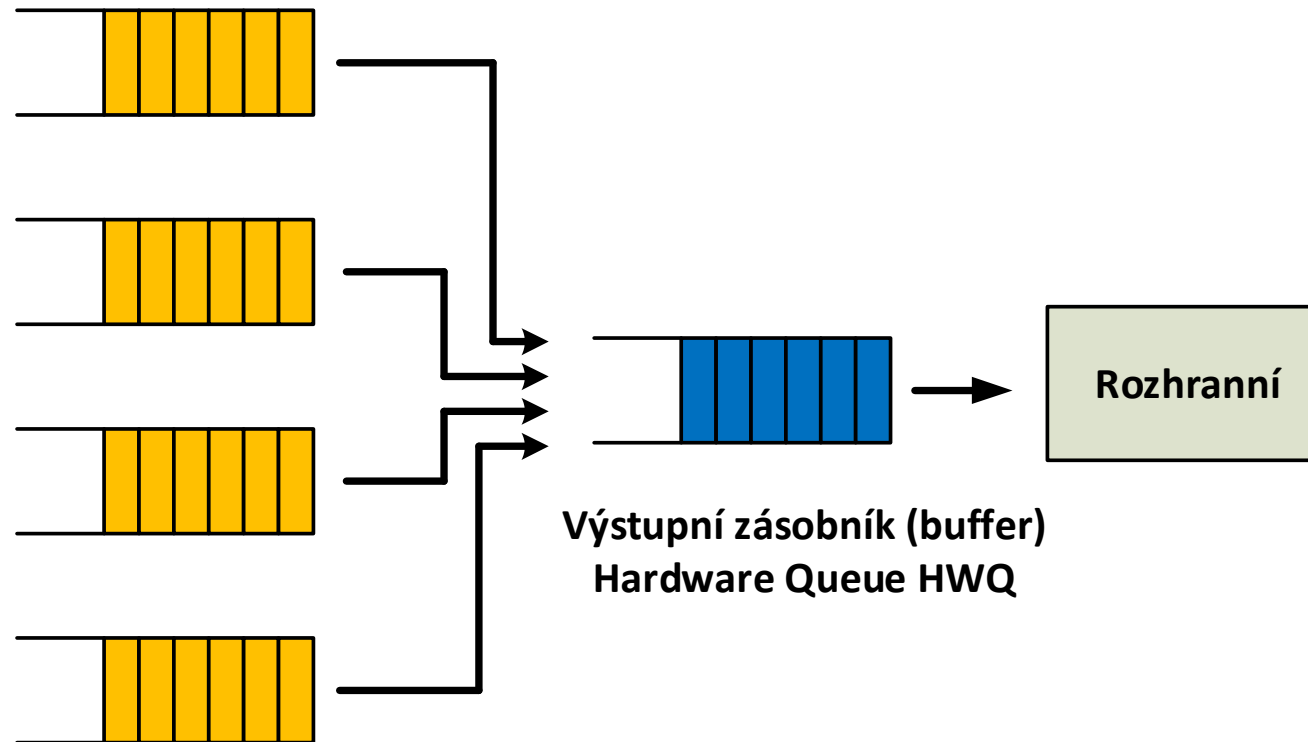


Low Latency Queuing (LLQ)

- Model CBWFQ - Class-Based Weighted Fair Queueing rozšířen o priority (PQ).
- Prioritním třídám je garantováno:
 - minimální zdržení paketů;
 - maximální dovolená přenosová rychlost;
- V případě překročení garantované přenosové rychlosti dochází ke ztrátě nadbytečných paketů i u prioritní třídy.



Hardwarová a softwarová paměť



Softwarové řazení
Software Queue SWQ

Výstupní zásobník (buffer)
Hardware Queue HWQ

- Po aktivaci SWQ se automaticky sníží kapacita HWQ na 2-3 pakety



Roztyl zpoždění (doby přenosu sítí)

- Pakety multimediálních toků vstupují do sítě většinou v pravidelných intervalech.
- V cílovém uzlu sítě však mohou být přijaty:
 - s různým zpožděním (v důsledku rozdílné cesty sítí),
 - v různém pořadí,
 - vícekrát,
 - nebo vůbec (protokol IP negarantuje doručení – best-effort).
- Rozptyl zpoždění (jitter) popisuje nepravidelnost těchto příchodů.
- Síťová infrastruktura by měla zajistit:
 - spolehlivé doručování paketů s garantovanými hodnotami jejich zpoždění a rozptylu.

Výše uvedený požadavek je **obtížně** splnitelný, pokud dochází vlivem přetížení k přeplnění vyrovnávacích pamětí směrovačů!

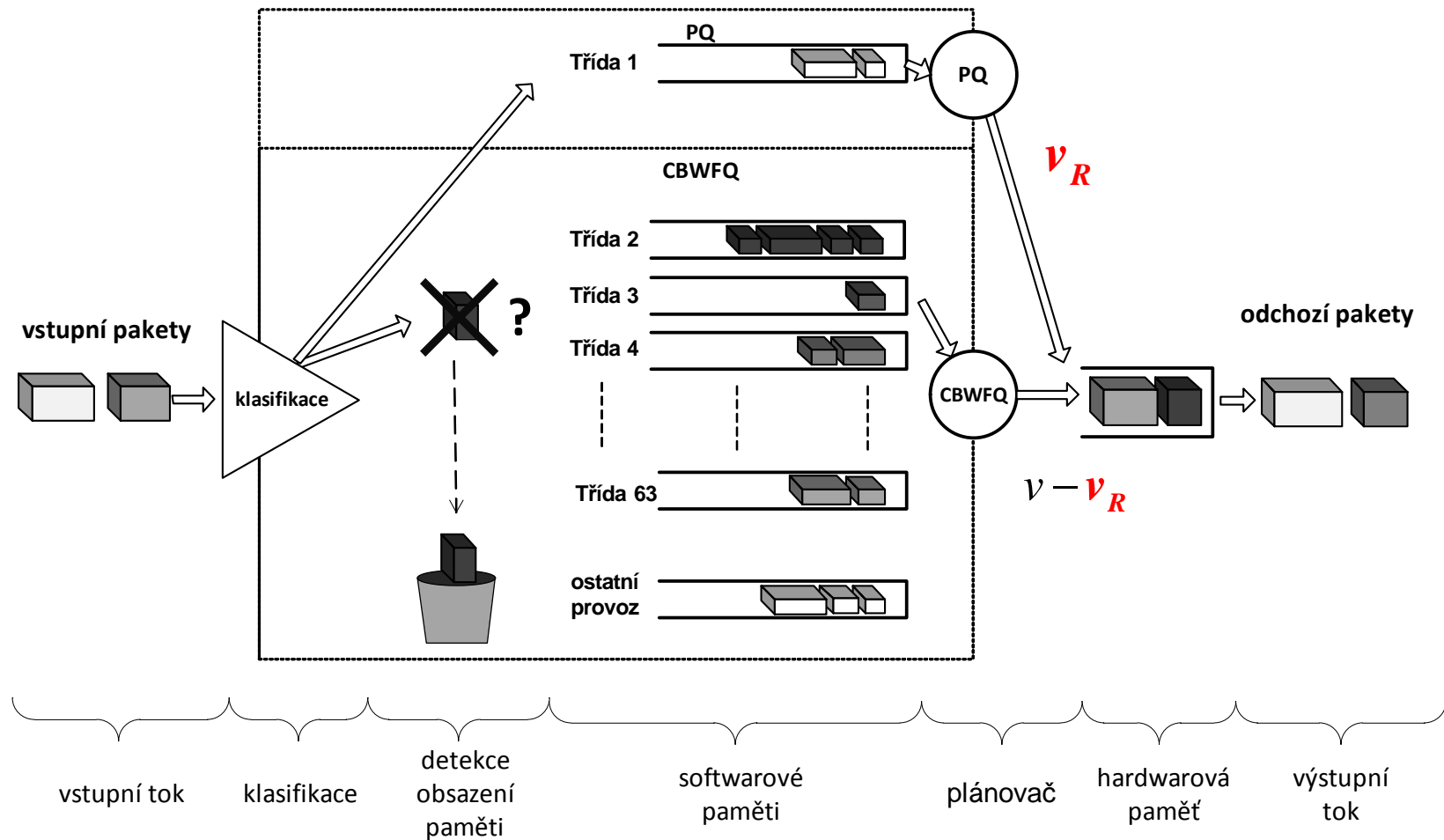


Stanovení přenosové rychlosti

- Snaha odhadnout potřebnou přenosovou rychlost v_R , kterou je nutné rezervovat pro skupinu Q konverzačních služeb, vychází z těchto předpokladů:
 - vstupní toky požadavků jednotlivých služeb lze aproximovat poissonovským tokem s intenzitou λ_q příchozích požadavků;
 - každé jednotlivé spojení q -tého toku využije z celkové dostupné kapacity v_R na rozhraní přenosovou rychlost v_q .
- V případě splnění výše uvedených předpokladů lze pro dimenzování užít **zobecněný Erlangův model**:
 - matematický model systému s čekáním (na třetí vrstvě modelu OSI) pak aproximujeme pro potřeby dimenzování modelem systému se ztrátou.

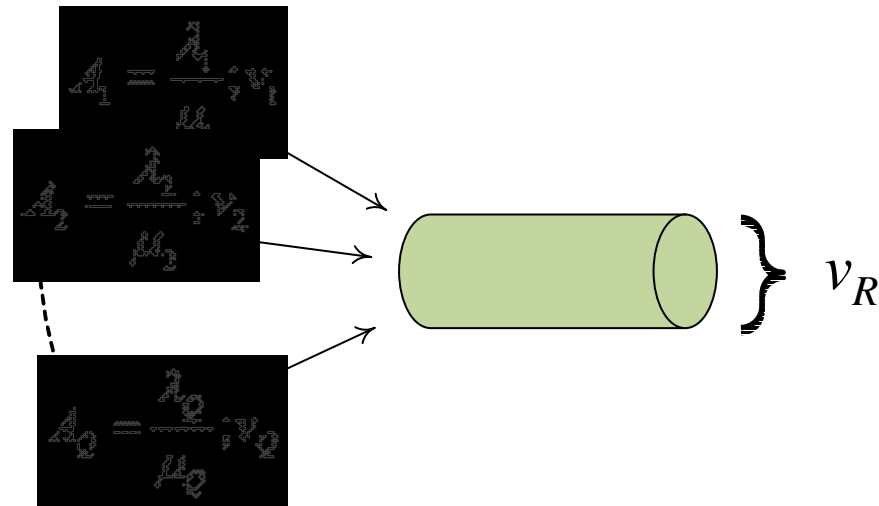


Příklad: systém Low Latency Queuing (LLQ)





Zobecněný Erlangův model pro Q toků:



- Vstupní tok tvoří Q různých služeb, každá s:
 - nabízeným zatížením:

$$A_q = \frac{\lambda_q}{\mu_q} = \lambda_q t_{os} \text{ [erl]}; \quad q = 1, 2, \dots, Q$$

kde λ_q je intenzita příchozích požadavků q -té služby [s^{-1}]

a $\mu_q = 1 / t_{os}$ je intenzita obsluhy q -té služby [s^{-1}]

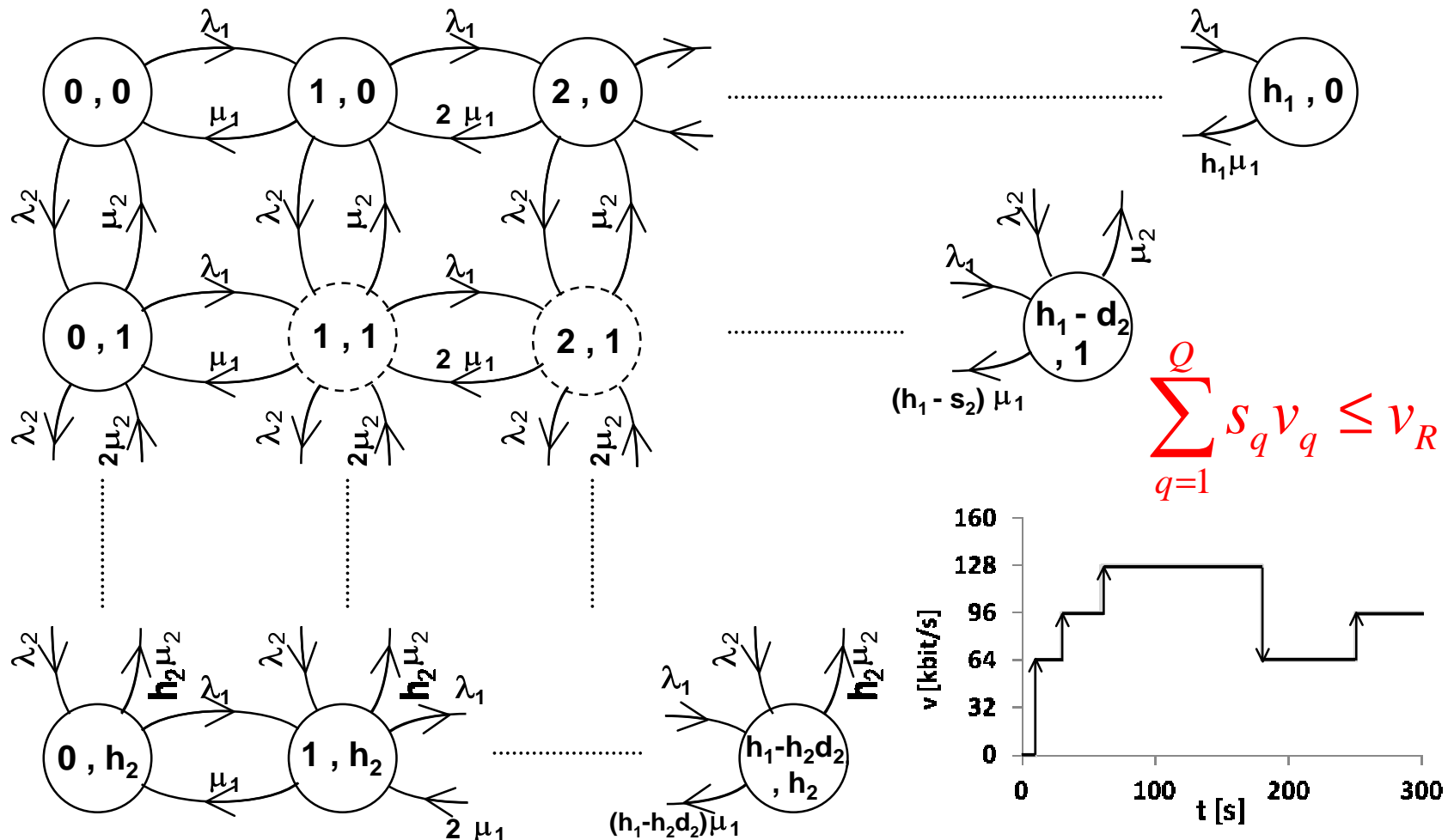
- přenosovou rychlostí pro jedno spojení:

$$v_q \text{ [kbit/s]}; \quad q = 1, 2, \dots, Q$$



Přidělování přenosové kapacity

- stav systému: (s_1, s_2) - počet požadavků z jednotlivých toků v syst.





Zobecněný Erlangův model pro Q toků:

- stacionární pravděpodobnosti stavů systému:

$$P(s_1, K, s_Q) = \frac{A_1^{s_1}}{s_1!} \cdot K \cdot \frac{A_Q^{s_Q}}{s_Q!} \cdot P(0, K, 0)$$

- kde pravděpodobnost, že je systém prázdný je dána:

$$P(0, K, 0) = \left[\sum_{0 \leq s_1 v_1 + K + s_Q v_Q \leq v_R} \frac{A_1^{s_1}}{s_1!} \cdot K \cdot \frac{A_Q^{s_Q}}{s_Q!} \right]^{-1}$$

- pravděpodobnost ztráty q -tého toku:

$$B_q = \sum_{v_1 s_1 + \dots + v_Q s_Q > v_R - v_q} P(s_1, \dots, s_Q)$$

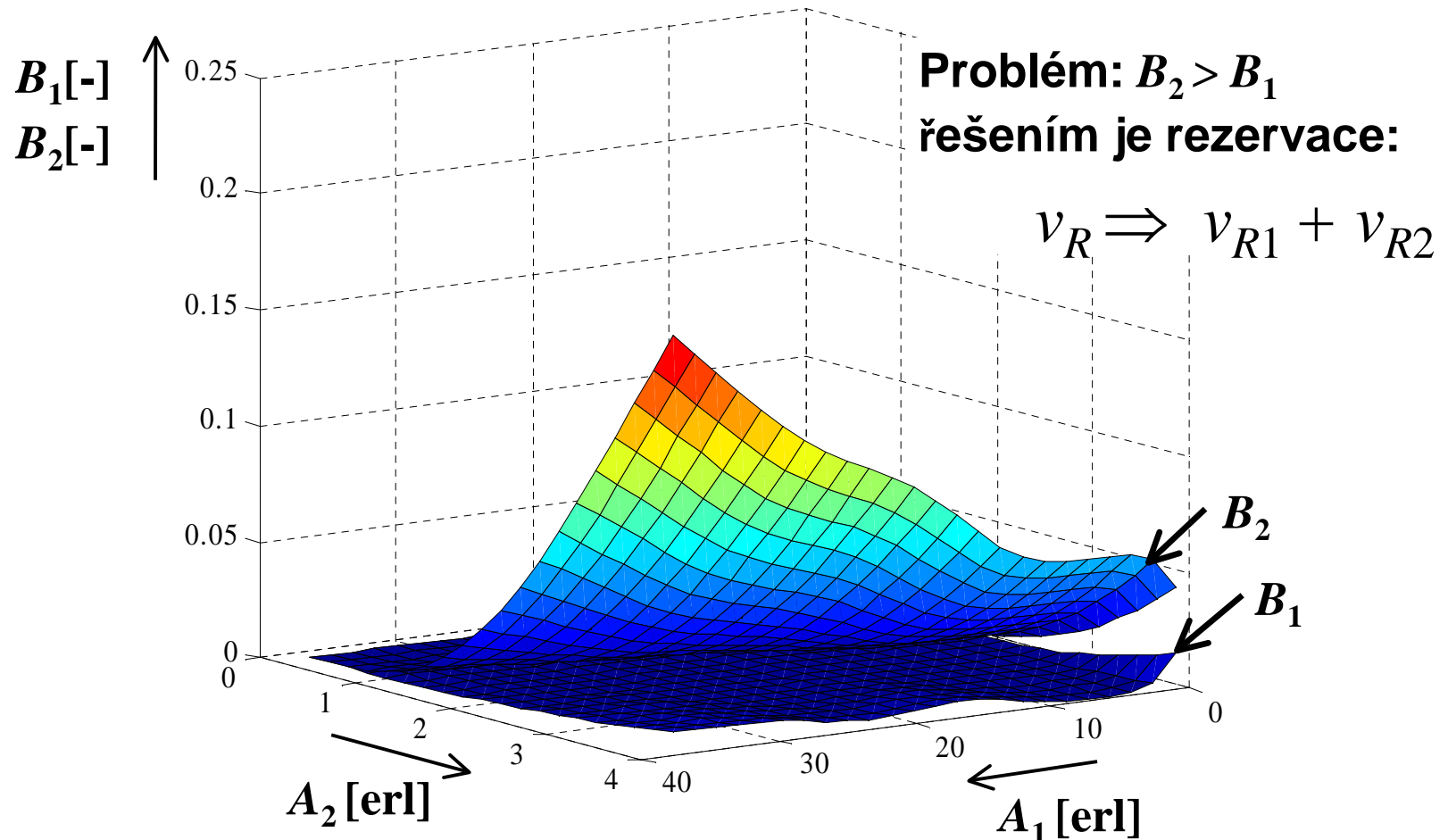
- střední hodnota pravděpodobnosti ztráty všech toků:

$$B_S = \frac{\sum_{q=1}^Q A_q B_q v_q}{\sum_{q=1}^Q A_q v_q}$$



Příklad:

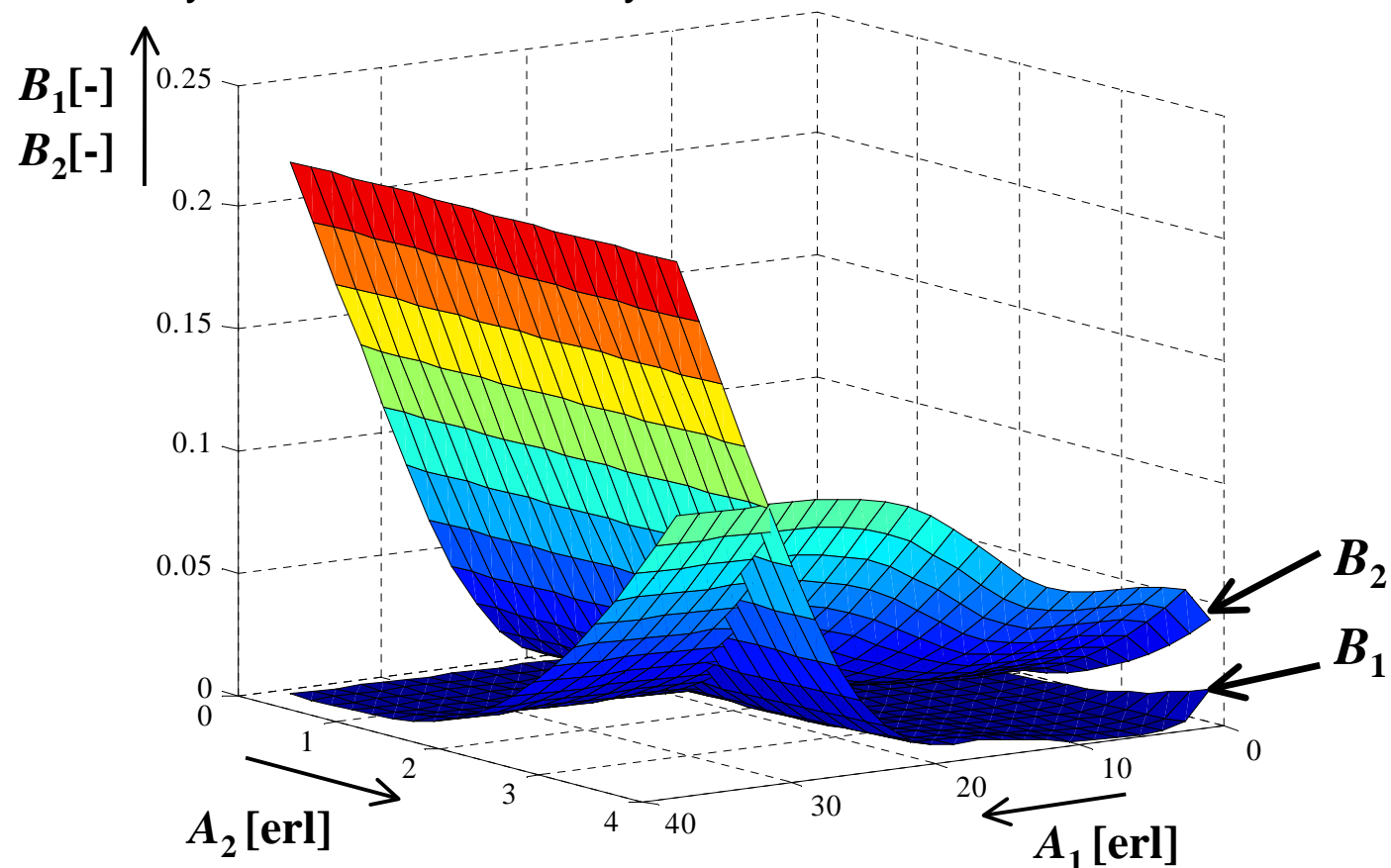
- model $M_1+M_2/M_1+M_2/v_R/0$; dvě služby:
 - VoIP; jedno spojení $v_1 = 32$ kbit/s,
 - videotelefon; jedno spojení $v_2 = 512$ kbit/s,





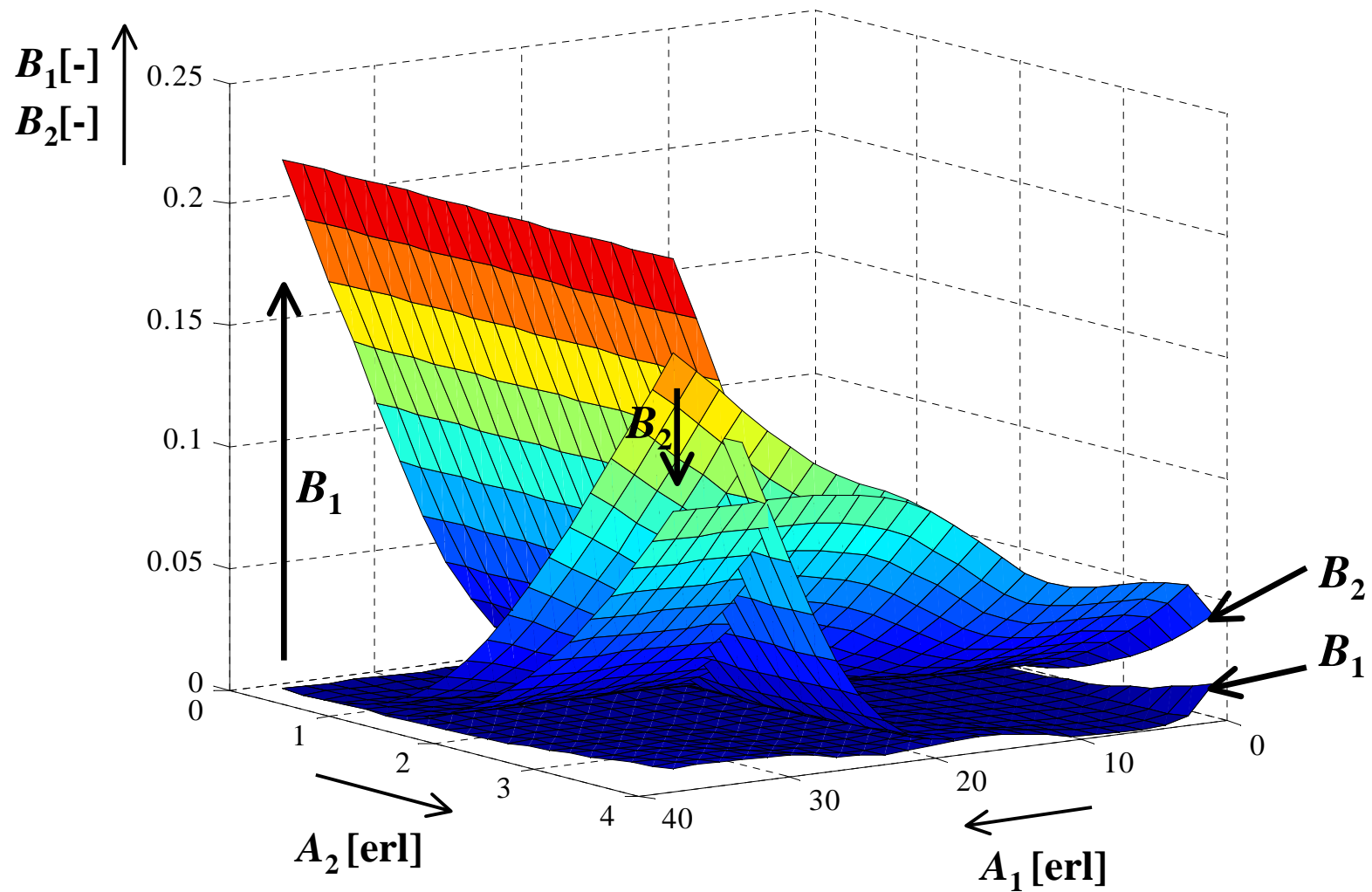
Příklad: systém s rezervací

- model $M_1+M_2/M_1+M_2/v_{R1} +v_{R2}/0$; dvě služby:
 - VoIP; jedno spojení $v_1 = 32$ kbit/s, $v_{R1} = 1024$ kbit/s max. 34 hovorů
 - videotelefon; jedno spojení $v_2 = 512$ kbit/s , $v_{R2} = 4096$ kbit/s teoreticky max. 4 video hovory současně





Příklad: oba modely současně





Příklady dimenzování (výsledky simulací)

A_1 [erl]	A_2 [erl]	ν_{R1} [Mbit/s]	ν_{R2} [Mbit/s]	distribuce	B_1 [-]	B_2 [-]
30	5	1,312	6,72	$H_2, c_v = 3$	0,0441 ± 0,9 %	0,0288 ± 1,8 %
				$M, c_v = 1$	0,0108	0,0100
				$E_2, c_v = 0,71$	0,0047 ± 0,8 %	0,0032 ± 1,7 %
				$D, c_v = 0$	0,0010 ± 1,7 %	0,0002 ± 5,2 %
60	10	2,368	11,088	$H_2, c_v = 3$	0,0472 ± 0,8 %	0,0326 ± 1,5 %
				$M, c_v = 1$	0,0108	0,0103
				$E_2, c_v = 0,71$	0,0054 ± 1,0 %	0,0040 ± 1,7 %
				$D, c_v = 0$	0,0014 ± 2,0 %	0,0005 ± 4,6 %
90	15	3,392	15,232	$H_2, c_v = 3$	0,0484 ± 0,8 %	0,3509 ± 1,0 %
				$M, c_v = 1$	0,0109	0,0100
				$E_2, c_v = 0,71$	0,0057 ± 1,0 %	0,0040 ± 2,0 %
				$D, c_v = 0$	0,0017 ± 1,6 %	0,0007 ± 3,8 %
120	20	4,416	19,136	$H_2, c_v = 3$	0,0465 ± 0,7 %	0,0393 ± 1,1 %
				$M, c_v = 1$	0,0103	0,0108
				$E_2, c_v = 0,71$	0,0056 ± 1,5 %	0,0048 ± 2,2 %
				$D, c_v = 0$	0,0018 ± 2,2 %	0,0008 ± 4,0 %

Poznámka: distribuce vyjadřuje různé rozložení intervalů mezi příchody (H_2 - hyperexponenciální, M - exponenciální, E_2 - Erlangovo 2 řádu; D -deterministické); c_v - je variační koeficient.



Závěr

- Reálné systémy jsou natolik složité, že je nelze komplexně analyticky popsat.
- Model byl zvolen z důvodů odhadu přenosové rychlosti pro jednotlivé služby.
- Charakter toků služeb nemusí vždy odpovídat předpokladům klasické teorie (poissonovský vstupní tok, exponenciální rozložení doby obsluhy,...).
- Doby zpoždění je nutné pro odhadnuté přenosové rychlosti získat měřením nebo simulací reálného modelu.
- Lze předpokládat, že uvedené modely (obzvláště pak model s rezervací) budou nabývat na význam v souvislosti s rozvojem nových sítí, například LTE.



Literatura

- M. Stasiak, M. Glabowski, A. Wisniewski, P. Zwierzykowski. *Modeling and dimensioning of mobile networks: from GSM to LTE*. Chichester, West Sussex, UK: Wiley, 2011, xxiii, 315 p. ISBN 978-047-0665-862.
- D. Gross, C. M. Harris, *Fundamentals of Queueing Theory*. 3rd edition. New York : Wiley-Interscience, 1998. 464 p. ISBN 0-471-17083-6.
- W. Odom, M. J. Cavanaugh. *Cisco QOS exam certification guide: CCVP self-study*. 2nd ed. Indianapolis: Cisco Press, 2005, 730 p. ISBN 15-872-0124-0.
- A. Ranjbar. *CCNP ONT: official exam certification guide*. Indianapolis: Cisco Press, 2007, 373 p. ISBN 978-1-58720-176-9.
- Cisco IOS Quality of Service Solutions Configuration Guide
http://www.cisco.com/en/US/docs/ios/12_2/qos/configuration/guide/fqos_c.html
- Internetworking Technology Handbook (Cisco) - Quality of Service
http://docwiki.cisco.com/wiki/Quality_of_Service_Networking



Děkuji za pozornost