

Efficient teletraffic models for optimizing the performance of communication networks

Michael D. Logothetis
mlogo@upatras.gr

University of Patras
Department of Electrical & Computer Engineering
Division of Telecommunications & Information Technology
Wire Communications & Information Technology Laboratory
265 04 Patras, Greece

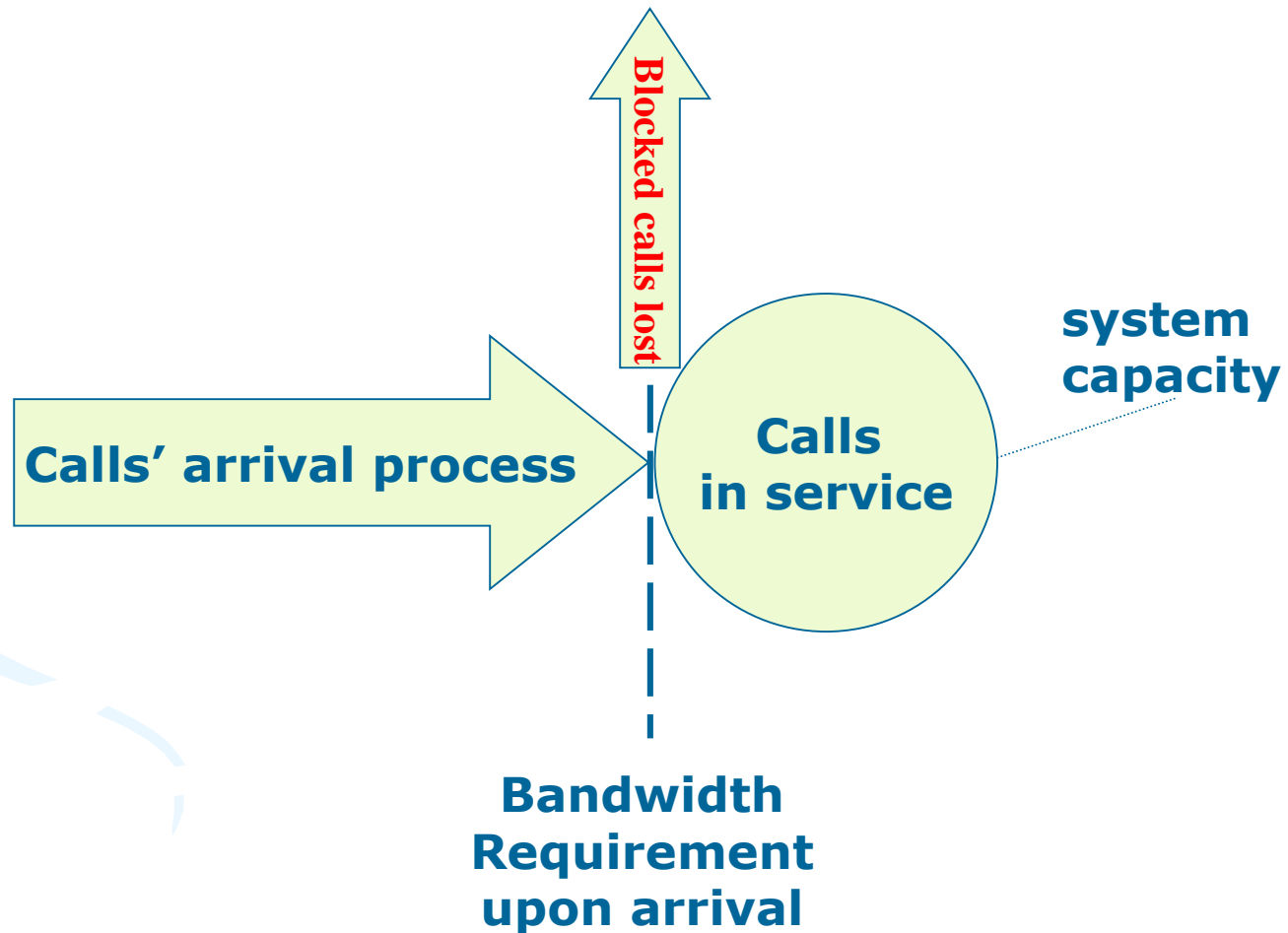
<http://www.wcl.ece.upatras.gr/teletraffic/mlogo/>

Structure

- Preamble
- Performance-oriented network management
- The importance of a teletraffic model
- Applicability to SDN-based 5G networks
- Classification of teletraffic models
- Efficient teletraffic loss models
 - ❑ Teletraffic models of random input
 - *The Erlang Multirate Loss Model (EMLM)*
 - *The Connection Dependent Threshold Model (CDTM)*
 - *The Extended Connection Dependent Threshold Model (E-CDTM)*
 - ❑ Teletraffic models of quasi-random input
 - *The Extended Engset Multi-rate Loss Model (E-EnMLM)*
 - *The Extended Finite Connection Dependent Threshold Model (CDTM)*
 - ❑ Teletraffic models of batched Poisson input
 - *The Batched Poisson ON-OFF Model (BP-ON-OFF)*
- Conclusion – Summary

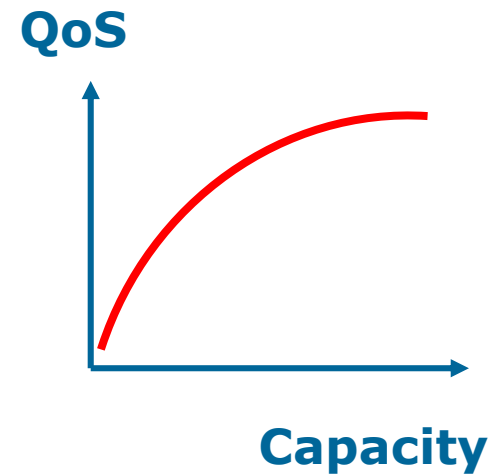
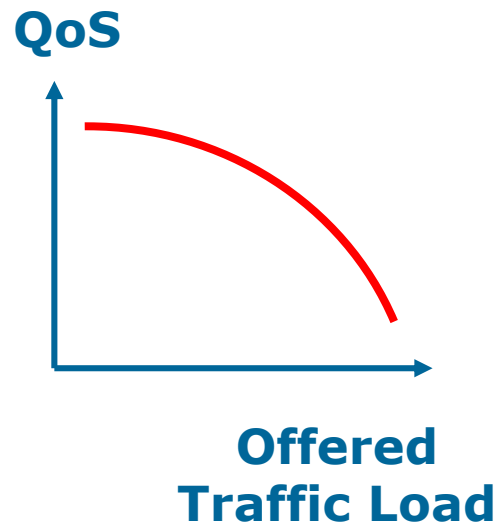
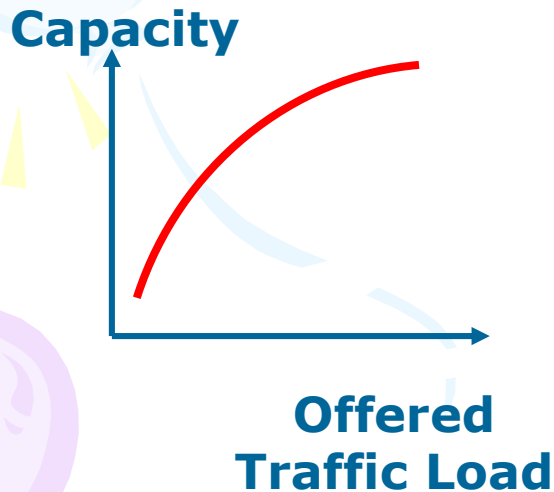
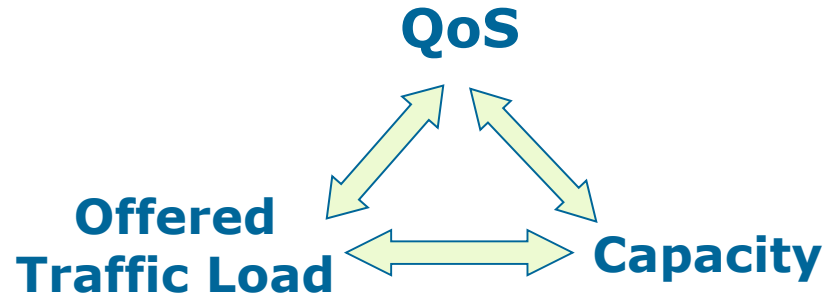
Preamble

A Loss Service System



Teletraffic Loss Models

*Models =
mathematical
formulas*



Performance-oriented network management

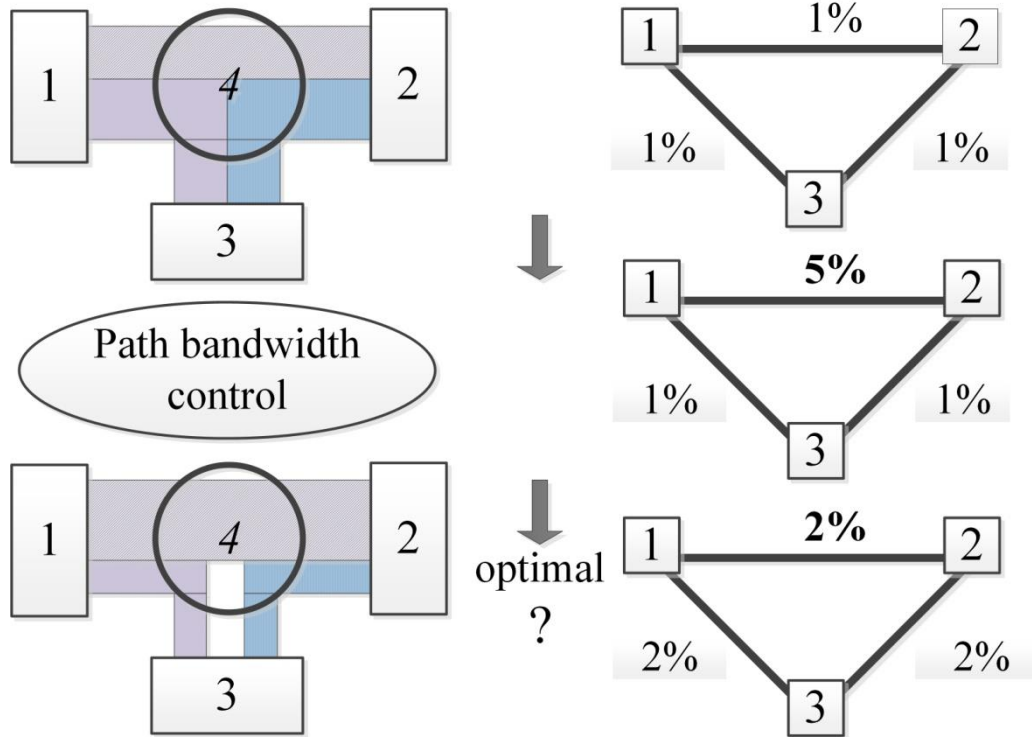
- Periodical performance evaluation (based on traffic measurements) and
- Adaptive resource assignment
 - **most suitable strategy for network planning under demand uncertainty**

Nature of traffic

- No single/constant rate
 - very different bandwidth per call requirements
 - several alternative contingency bandwidth requirements (e.g., multimedia traffic)
- In-service calls may
 - have adaptive features of bandwidth and holding time
 - experience bandwidth compression-expansion
- Random - Bursty traffic
 - Poisson arrivals
 - Quasi-random arrivals
 - ON-OFF traffic
 - Batched Poisson arrivals

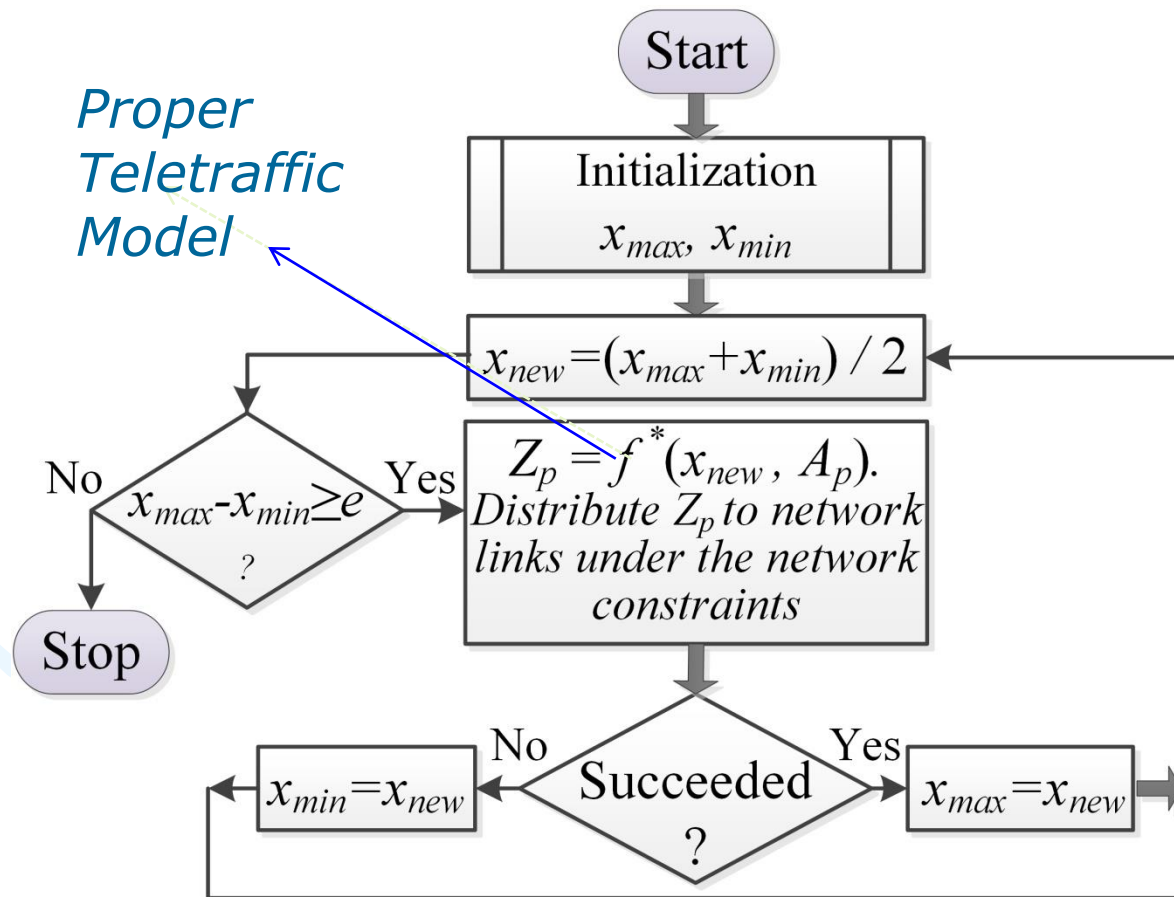
Performance-oriented network management

The Problem



Performance-oriented network management

Global network optimization – Algorithm



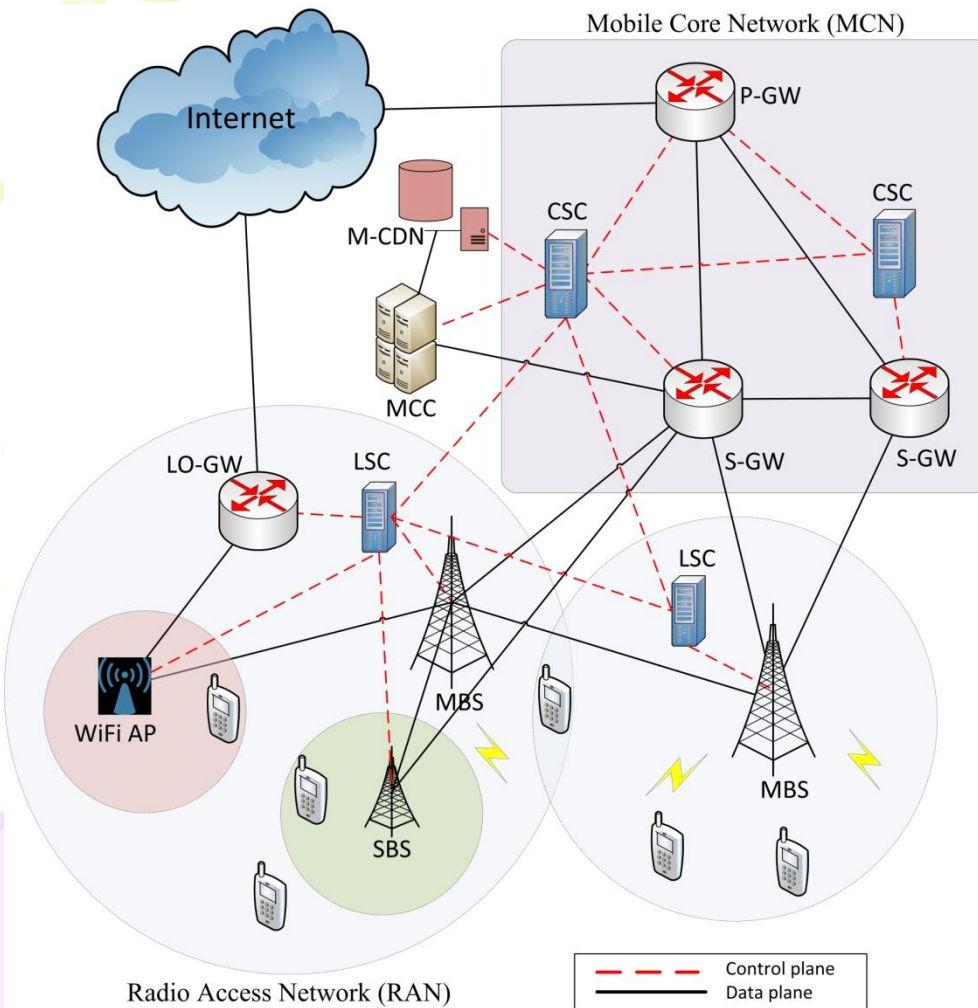
Teletraffic Models – Why?

- **Importance of QoS assessment through teletraffic models:**
 - Bandwidth allocation among service-classes → QoS per service guarantee.
 - Avoidance of too costly over-dimensioning of the network.
 - Prevention of excessive network throughput degradation, through traffic engineering mechanisms.
- **A sine qua non of teletraffic loss models:**

The efficient calculation of Call Blocking Probability → Recursive formula.
- **Applicability:**
 - Connection Oriented Communication Networks, in general.
 - IP based networks with resource reservation capabilities (IntServ - DiffServ).
 - Cellular networks (e.g., UMTS).
 - All-optical core networks (MP λ S/GMPLS).
 - 5G networks.

Applicability to SDN-based 5G networks

SDN/NFV based 5G architecture



Software Defined Network (SDN):
*completely programmable network
by decoupling the control and data planes.*

Network Function Virtualization (NFV):
*allows executing the SDN functions on general-
purpose hardware, reducing the network cost.*

P-GW – Packet Data Network Gateway

CSC – Core SDN Controller

LSC – Local SDN Controller

S-GW – Serving Gateway

M-CDN – Mobile Content Delivery Network

MCC – Mobile Cloud Computing

MBS – Macro cell Base Station

SBS – Small cell Base Station

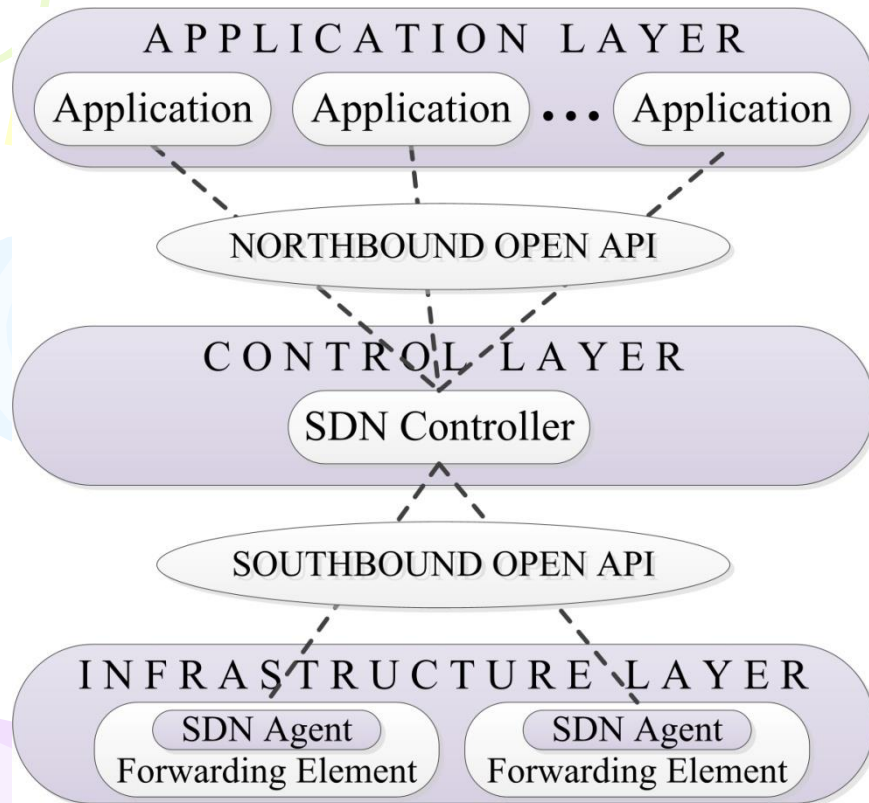
WiFi AP – Access Point with WiFi protocol

LO-GW – Local Offload Gateway

MU – Mobile User

Applicability to SDN-based 5G networks

Layering concept in SDN

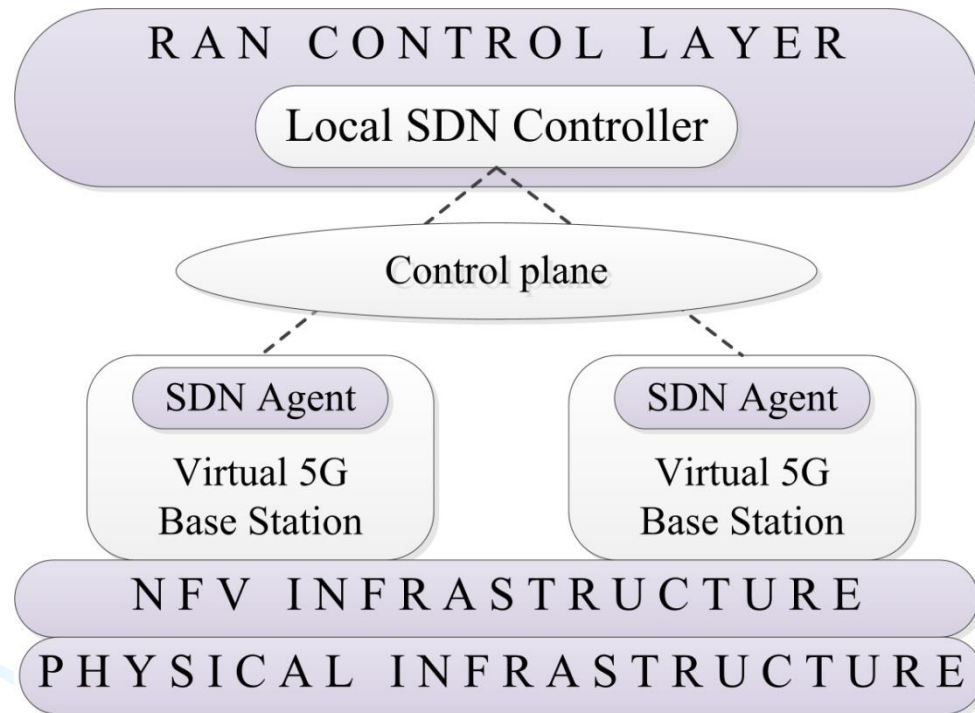


The SDN controller provides a global view of the available underlying resources to network applications (Application Layer) by the Northbound Open API.

The SDN controller configures the Forwarding Elements (located at the Infrastructure Layer) by sending control messages to the SDN Agents (located within the FEs) through the Southbound Open API.

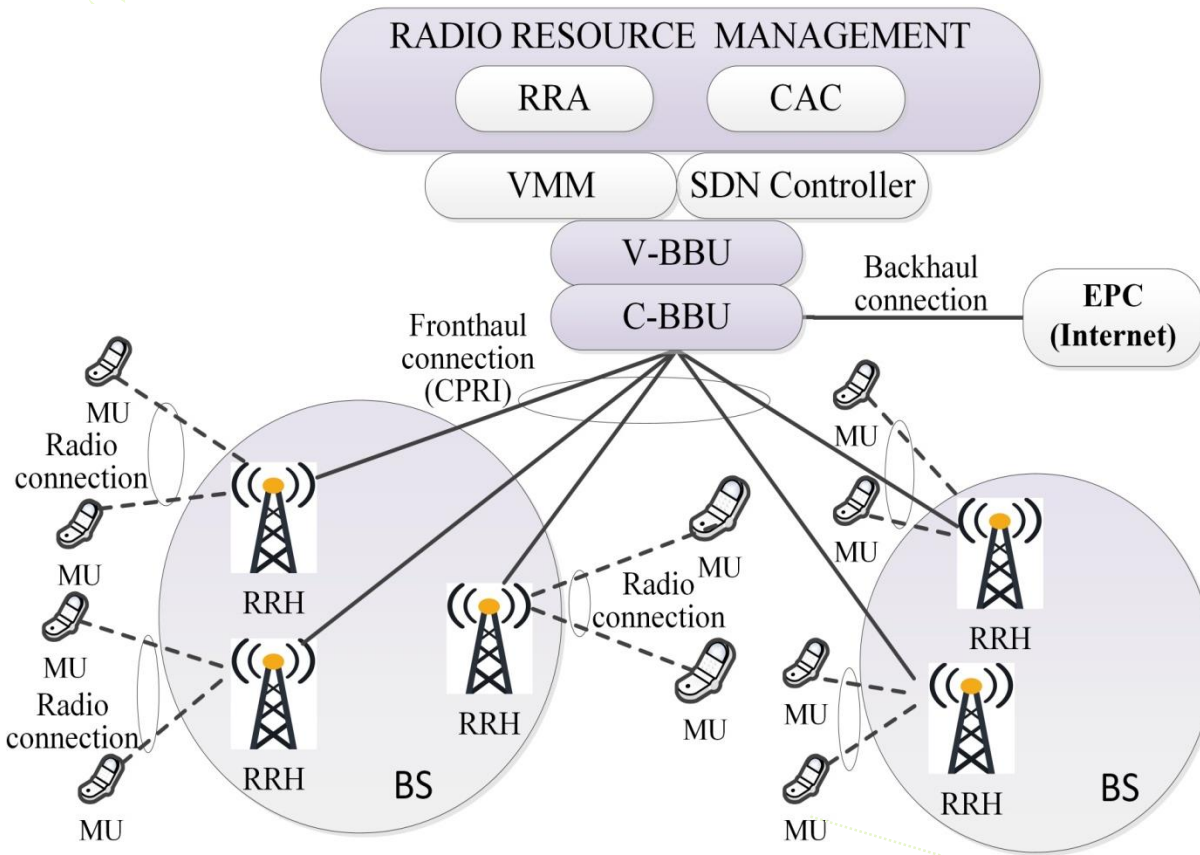
Applicability to SDN-based 5G networks

SDN/NFV based RAN



Applicability to SDN-based 5G networks

Cloud-RAN architecture



RRA – Radio Resource Allocation
CAC – Connection Admission Control
VMM – Virtual Machine Monitor
V-BBU – Virtual BaseBand Units
(signal processing servers)
C-BBU – Centralized BaseBand Units
(central pool of data center resources)
EPC – Evolved Packet Core
CPRI – Common Public Radio Interface
RRH – Remote Radio Head
(radio frequency components–antennas)
BS – Base Station

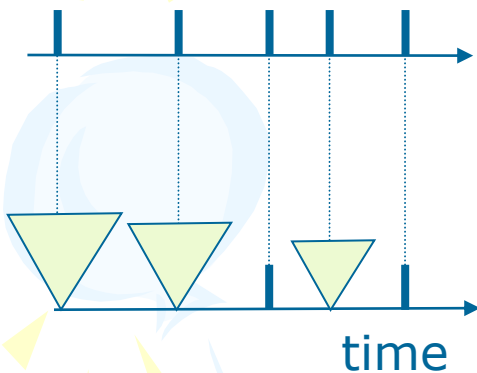
A call requires a radio resource unit from RRH and a computational resource unit from V-BBU.
CAC checks the availability of resources to accept the call.

Classification of teletraffic models

Key considerations:

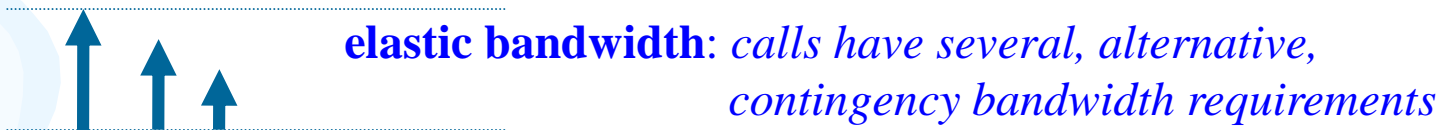
- The call arrival process.
- The service-classes
 - Bandwidth requirement upon call arrival.
 - The behavior of in-service calls regarding the amount of occupied b.u. per call over time.
- Bandwidth sharing policy
 - Complete sharing policy
 - Bandwidth/Trunk reservation policy
 - Threshold Policy

Call Arrival Process

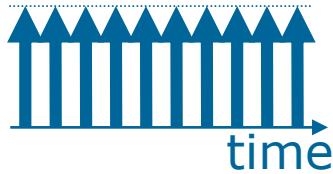


- ❖ Random arrivals – traffic (*infinite number of traffic sources*).
- ❖ Quasi-random arrivals – traffic (*finite number of traffic sources*).
- ❖ Batched Poisson arrivals (*infinite number of traffic sources*).
Calls from different service-classes arriving in batches,
while batches arriving randomly.

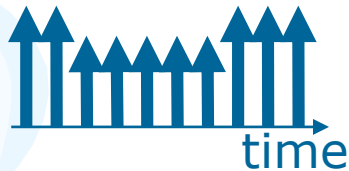
Bandwidth requirement upon call arrival



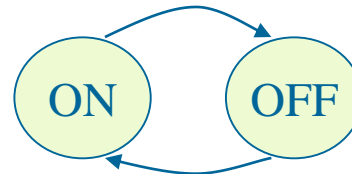
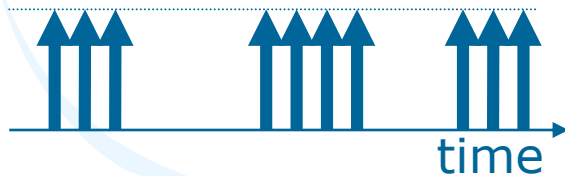
Call's behavior while in service



constant-bit-rate
(stream traffic)



bandwidth compression/expansion
(elastic traffic)



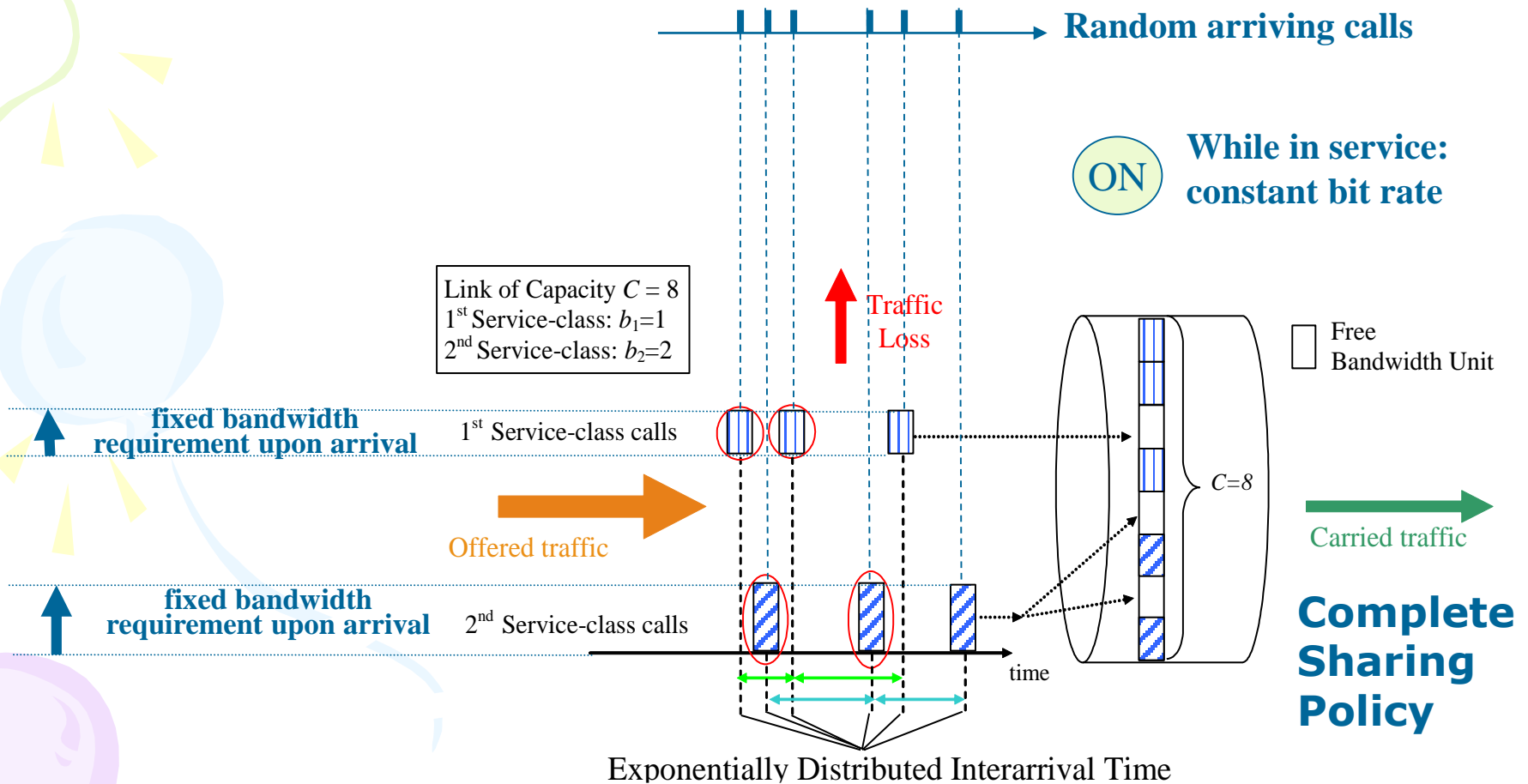
ON-OFF traffic

Efficient teletraffic loss models

- **Teletraffic models of random input**

- Random arriving calls with fixed or elastic bandwidth requirements, and *fixed bandwidth allocation during service*.
- Random arriving calls with fixed or elastic bandwidth requirements, and *elastic bandwidth during service*.
- Random arriving calls with fixed or elastic bandwidth requirements, and *ON-OFF traffic behavior during service*.

The Erlang Multi-rate Loss Model (EMLM)

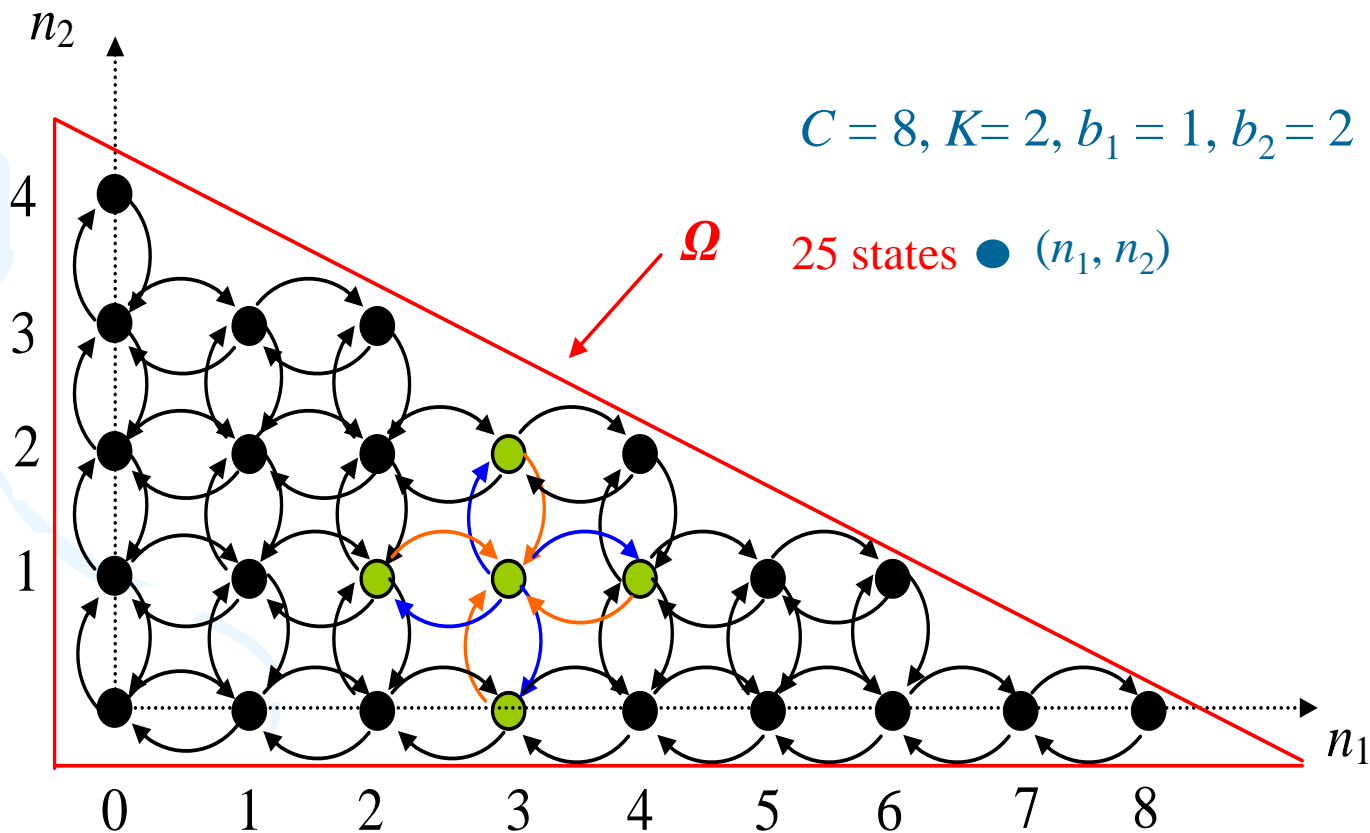


EMLM Analysis – Classical Method

State Space Ω

Complete Sharing Policy – A coordinate convex policy

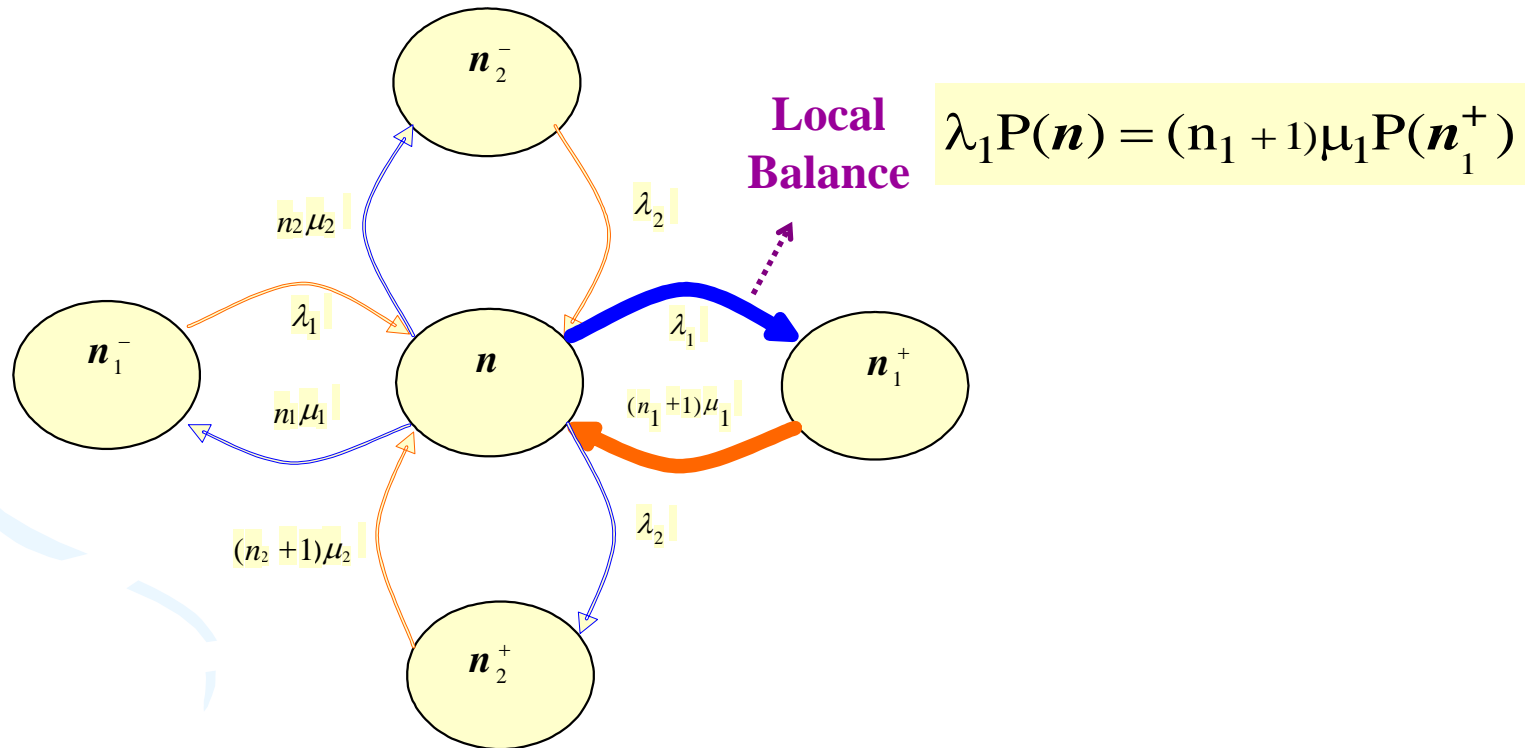
Global Balance (rate_in=rate_out) - Statistical equilibrium



EMLM Analysis – Classical Method (cont.1)

Global Balance (Rate_in = rate_out)

Local Balance (Rate_up = rate_down)



EMLM Analysis – Classical Method (cont.2)

Call Blocking Probability Determination – Classical Method

$$K=2, b_1 = 1, b_2 = m$$

$$P_{b_1} = P_{00} \sum_{j=0}^s \frac{\alpha_1^{C-mj}}{(C-mj)!} \frac{\alpha_2^j}{j!}$$

$$P_{b_2} = P_{00} \left(\frac{\alpha_2^2}{s!} \sum_{i=0}^k \frac{\alpha_1^i}{i!} + \sum_{j=0}^{s-1} \sum_{i=C-mj-m+1}^{C-mj} \frac{\alpha_1^i}{i!} \frac{\alpha_2^j}{j!} \right)$$

Example of formulas
for Call Blocking
Probability Calculation

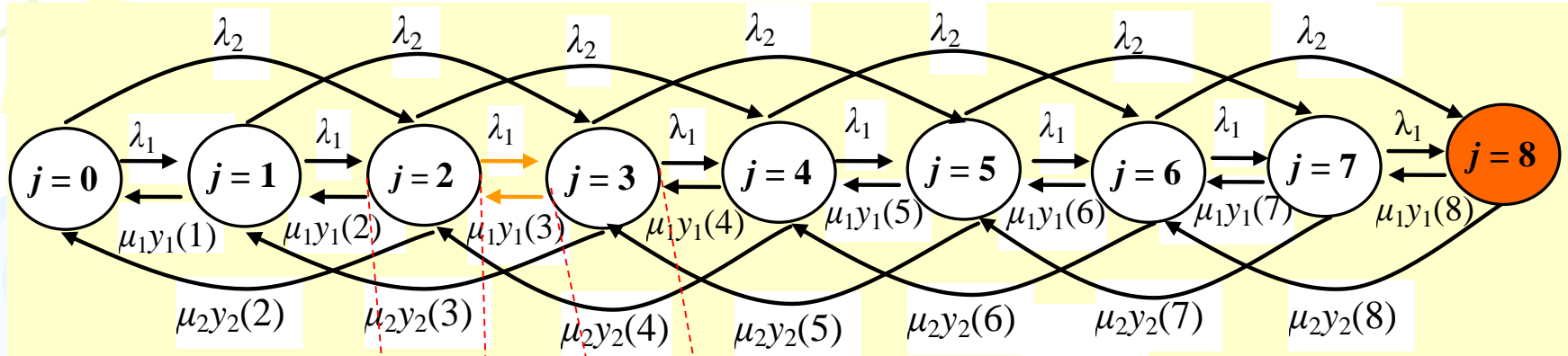
where $k = C \text{ (mod } m)$

(Necessity for recursive formulas)

EMLM Analysis – New Method

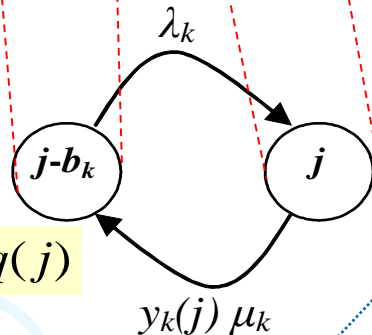
Macro-states – One-dimensional Markov chain

$C = 8, K=2, b_1 = 1, b_2 = 2$ Macro-state $j=n_1b_1+n_2b_2$ denotes the total number of in-service calls



local balance

$$\lambda_k q(j - b_k) = y_k(j) \mu_k q(j)$$



$$q(j) = \begin{cases} 1 & \text{for } j = 0 \\ \frac{1}{j} \sum_{k=1}^K a_k b_k q(j - b_k) & \text{for } j = 1, \dots, C \\ 0 & \text{otherwise} \end{cases}$$

“Kaufman / Roberts Recursion”

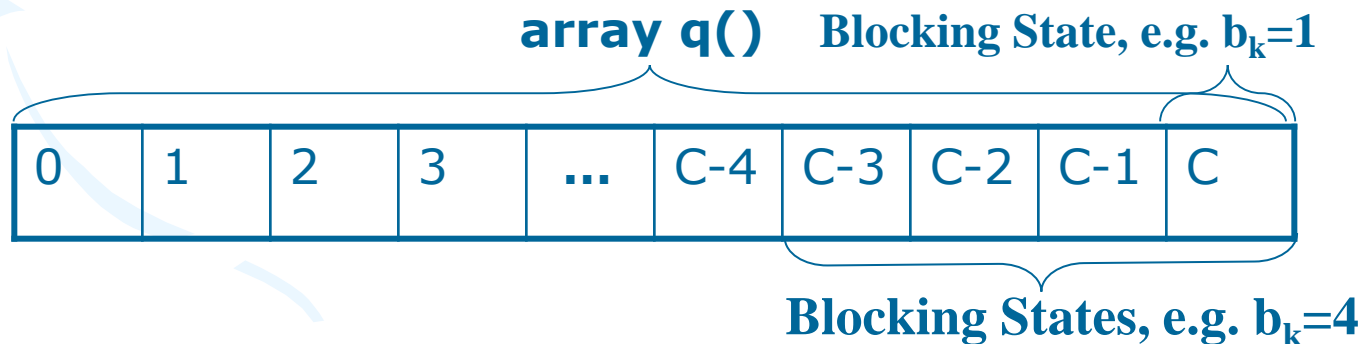
Un-normalized

Call Blocking Probability – Recursive Calculation

Call Blocking Probability: $P_{b_k} = \sum_{j=C-b_k+1}^C G^{-1} q(j)$ where $G = \sum_{j=0}^C q(j)$

**Accurate calculation especially
when service-classes have
equal mean service times!**

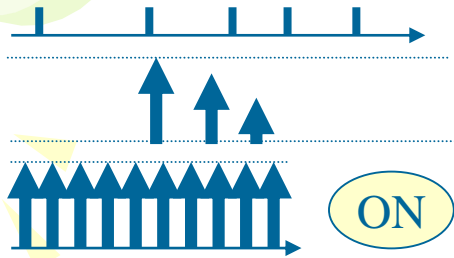
$q(j)/G$ – Macro-state Probabilities



Link Utilization:

$$U = \sum_{j=1}^C j q(j)$$

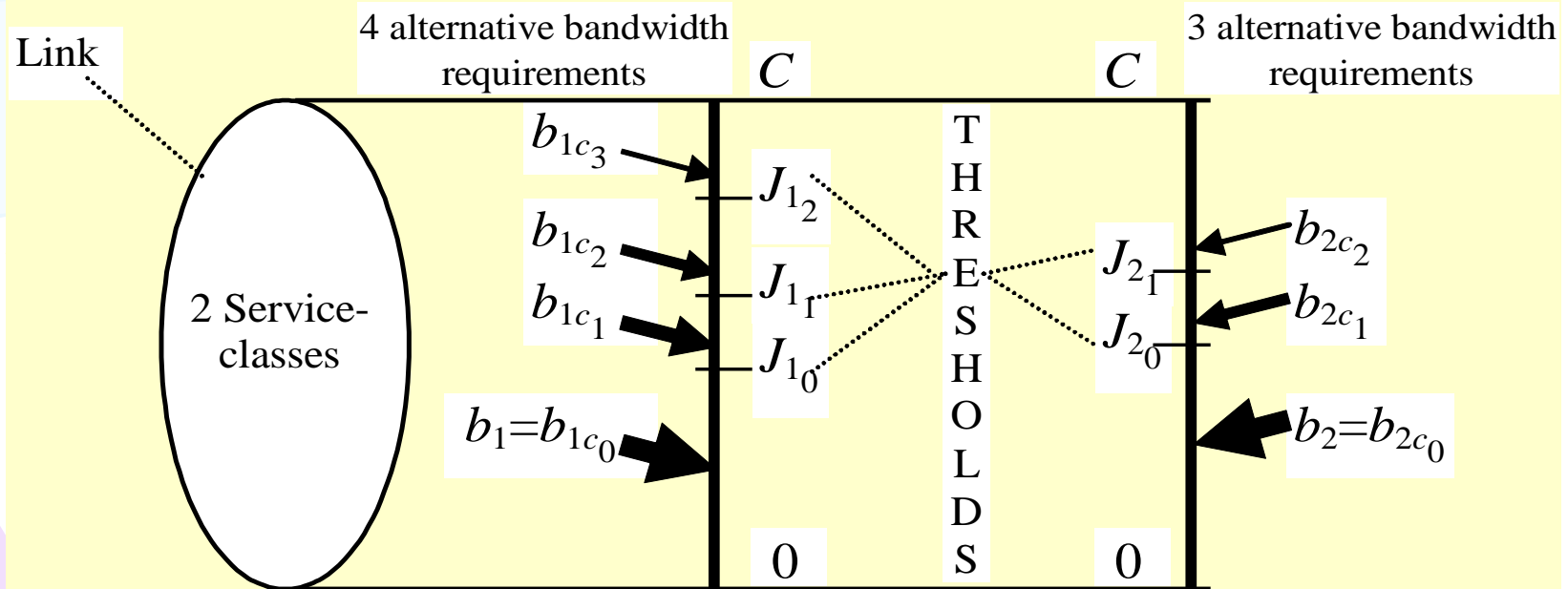
The Connection Dependent Threshold Model (CDTM)



Random arrivals

Elastic bandwidth requirements

Constant bit rate (stream traffic)



~~Local Balance~~



~~Product Form Solution~~



$\approx P_{bk}$

CDTM - The analytical model

Assumptions – Approximations

- 1) Local Balance
- 2) Migration Approximation, M.A ($\delta_{kc_s}(j)$)
- 3) Upward migration Approximation, U.A ($\delta_k(j)$)

$$q(j) = \begin{cases} 1 & \text{for } j = 0 \\ \frac{1}{j} \left(\sum_{k=1}^K a_k b_k \delta_k(j) q(j - b_k) + \sum_{k=1}^K \sum_{s=1}^{S(k)} a_{kc_s} b_{kc_s} \delta_{kc_s}(j) q(j - b_{kc_s}) \right) & \text{for } j = 1, \dots, C \\ 0 & \text{otherwise} \end{cases}$$

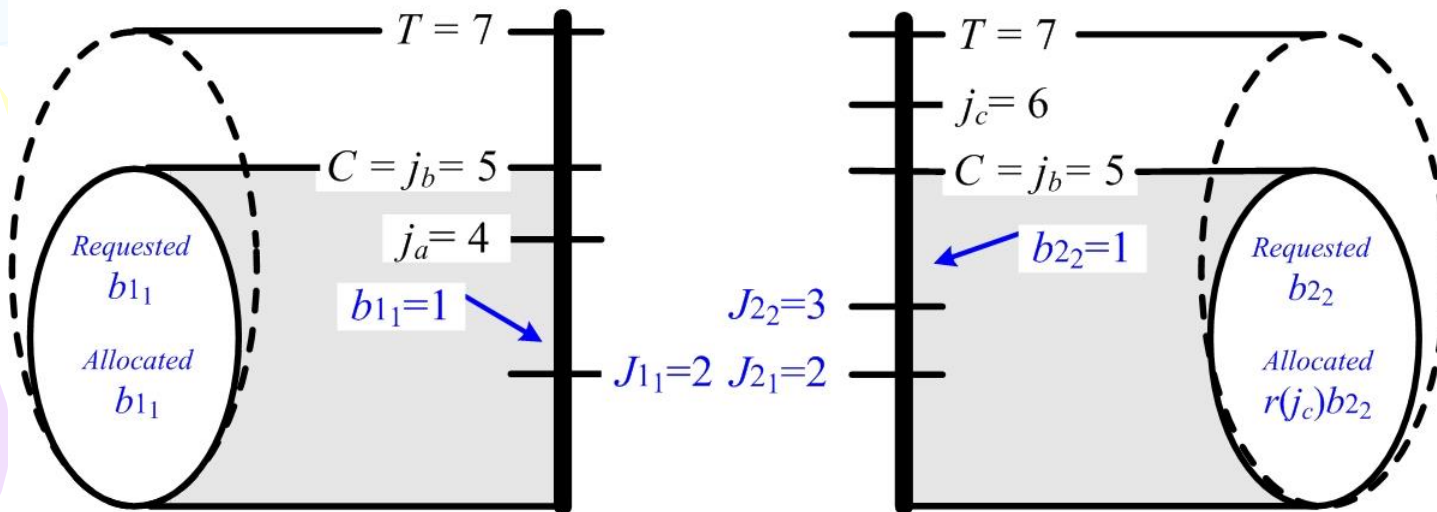
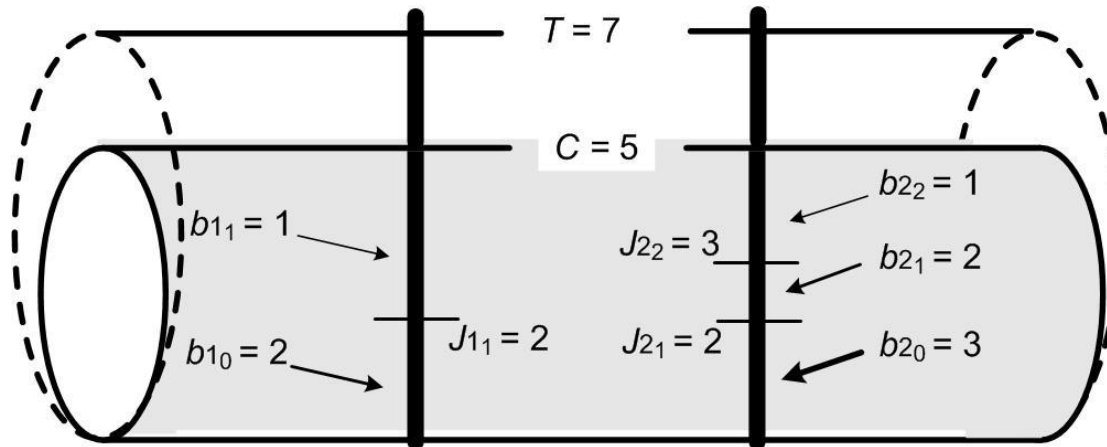
$$a_{kc_s} = \lambda_k \mu_{kc_s}^{-1} \quad \delta_k(j) = \begin{cases} 1 & \text{(if } 1 \leq j \leq J_{k0} + b_k \text{ and } b_{kc_s} > 0) \text{ or (if } 1 \leq j \leq C \text{ and } b_{kc_s} = 0) \\ 0 & \text{otherwise} \end{cases} \quad \text{U.A}$$

$$\delta_{kc_s}(j) = \begin{cases} 1 & \text{if } J_{ks} + b_{kc_s} \geq j > J_{ks-1} + b_{kc_s} \text{ and } b_{kc_s} > 0 \\ 0 & \text{otherwise} \end{cases} \quad \text{M.A}$$

$$\text{Call Blocking Probability: } P_{b_k} = \sum_{j=C-b_{kc_{S(k)}}+1}^C G^{-1} q(j) \quad \text{where } G = \sum_{j=0}^C q(j)$$

The Extended Connection Dependent Threshold Model (E-CDTM)

example



Compression rate = $C/j = 5/6$

E-CDTM – The analytical model for elastic and adaptive service-classes

Link occupancy distribution

$$q(j) = \begin{cases} 1 & \text{for } j = 0 \\ \frac{1}{\min(C, j)} \sum_{k \in K_e} \sum_{l=0}^{S_k} \alpha_{k_l} b_{k_l} \delta_{k_l}(j) q(j - b_{k_l}) + \\ \quad + \frac{1}{j} \sum_{k \in K_a} \sum_{l=0}^{S_k} \alpha_{k_l} b_{k_l} \delta_{k_l}(j) q(j - b_{k_l}) & \text{for } j = 1, \dots, T \\ 0 & \text{otherwise} \end{cases}$$

un-normalized

$$G = \sum_{j=0}^T q(j)$$

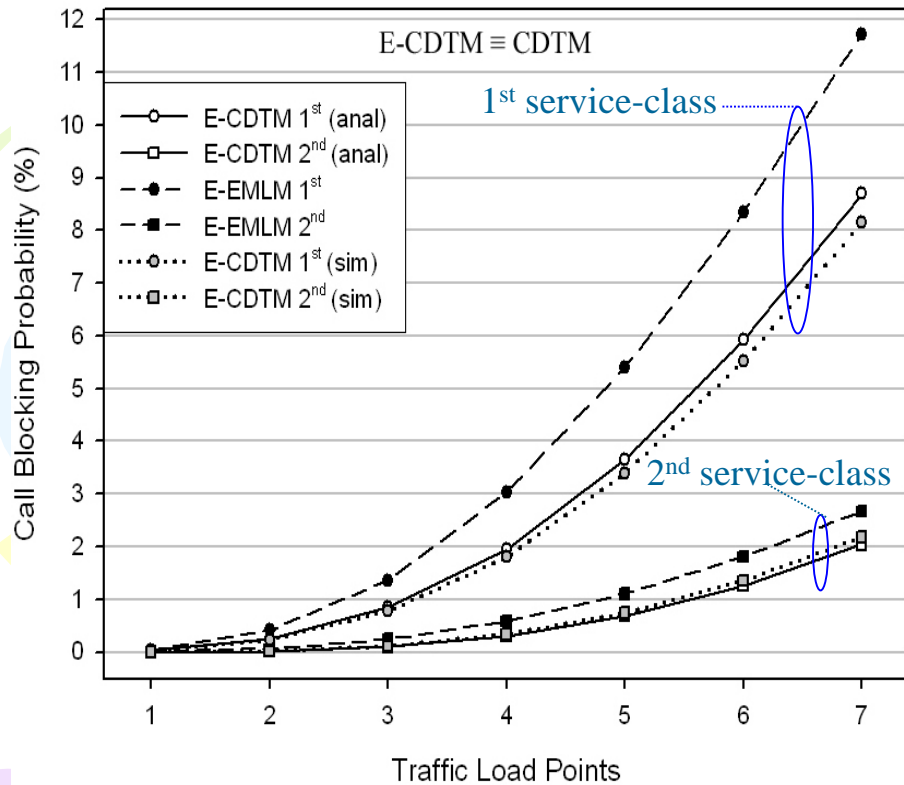
Call Blocking Probability

$$P_{b_k} = \sum_{j=T-b_k S_k+1}^T G^{-1} q(j)$$

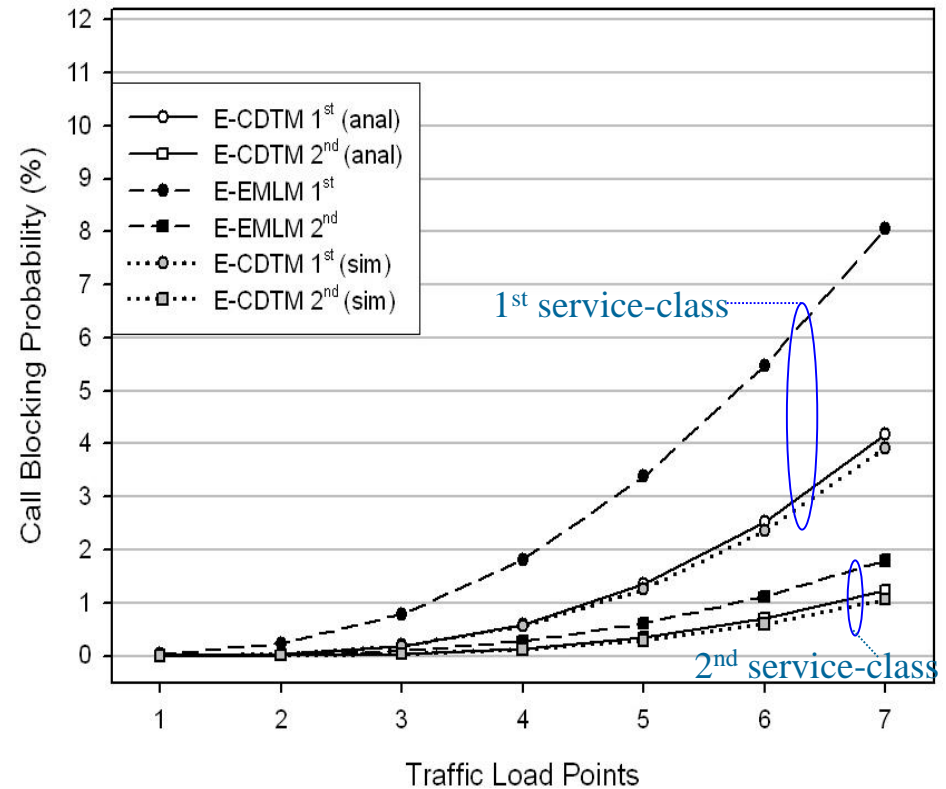
Link Utilization

$$U = \sum_{j=1}^C j G^{-1} q(j) + \sum_{j=C+1}^T G^{-1} C q(j)$$

E-CDTM versus E-EMLM



$$C=T=80$$



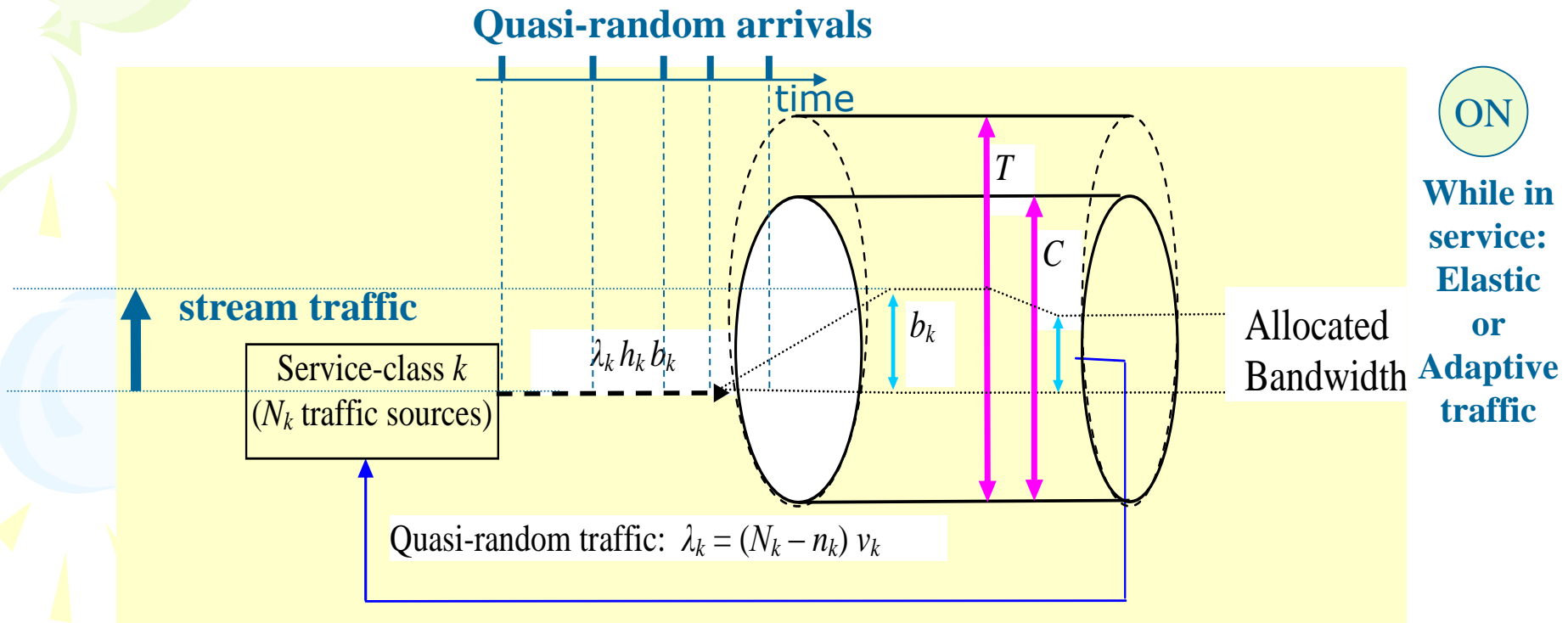
$$T=C+10$$

Efficient teletraffic loss models (cont.1)

- **Teletraffic models of quasi-random input**

- Quasi-random arriving calls with fixed or elastic bandwidth requirements and *fixed bandwidth allocation during service*.
- Quasi-random arriving calls with fixed bandwidth requirements and *elastic bandwidth during service*.
- Quasi-random arriving calls with fixed bandwidth requirements and *ON–OFF traffic behavior during service*.

The Extended Engset Multi-rate Loss Model (E-EnMLM)



h_k : holding (service) time of service-class k calls

If compression: “Bandwidth * Service-time” \Rightarrow constant \Rightarrow elastic traffic

j : total bandwidth demand ($0 \leq j \leq T$)

T : maximum total bandwidth demand ($T \geq C$)

s : real bandwidth allocation ($0 \leq s \leq C$)

E-EnMLM – The analytical model for elastic and adaptive service-classes

Link occupancy distribution

$$q(j) = \begin{cases} 1 & \text{for } j = 0 \\ \frac{1}{\min(C, j)} \sum_{k \in K_e} (N_k - n_k + 1) \alpha_k b_k q(j - b_k) + \\ + \frac{1}{j} \sum_{k \in K_a} (N_k - n_k + 1) \alpha_k b_k q(j - b_k) & \text{for } j = 1, \dots, T \\ 0 & \text{otherwise} \end{cases}$$

un-normalized

$$G = \sum_{j=0}^T q(j)$$

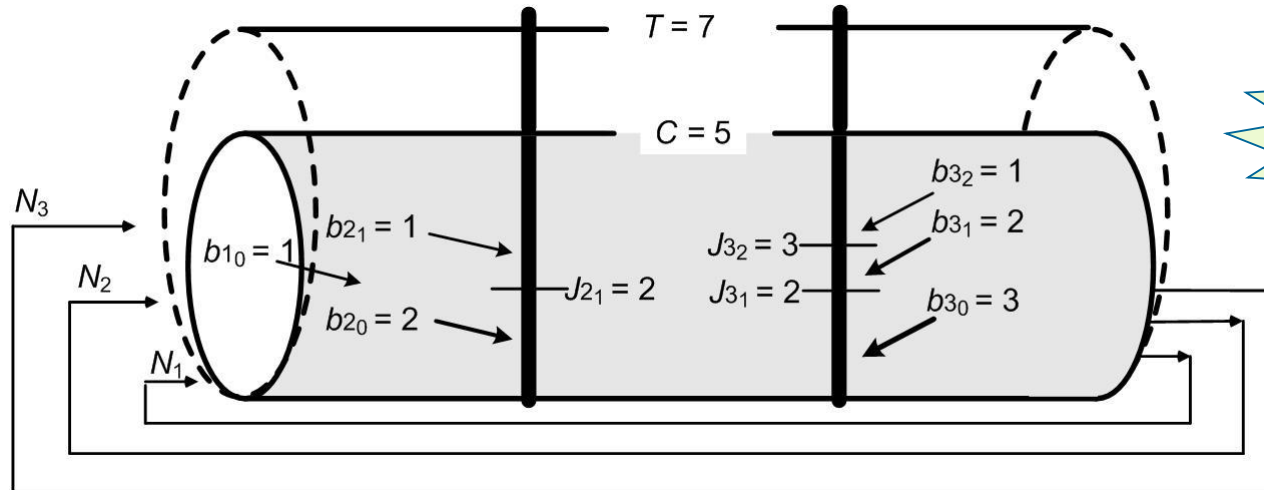
Time Congestion Probability

$$P_{b_k} = \sum_{j=T-b_k+1}^T G^{-1} q(j)$$

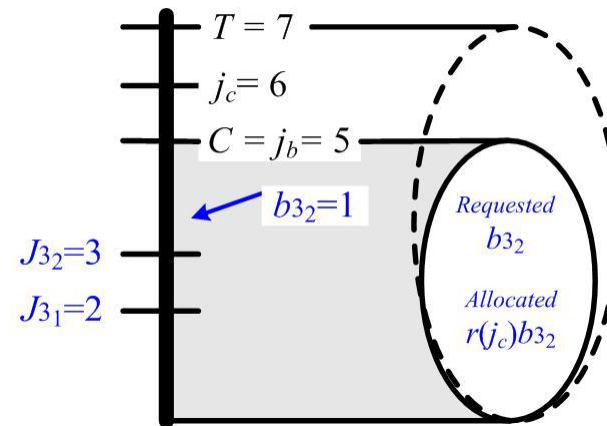
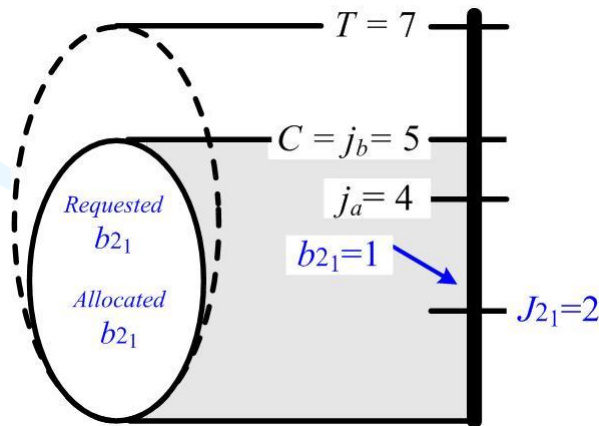
Link Utilization

$$U = \sum_{j=1}^C j G^{-1} q(j) + \sum_{j=C+1}^T G^{-1} C q(j)$$

The Extended Connection Dependent Threshold Model for finite population (Ef-CDTM)



example



Compression rate = $C/j = 5/6$

Ef-CDTM – The analytical model

Link occupancy distribution

$$q(j) = \begin{cases} 1 & \text{for } j = 0 \\ \frac{1}{\min(C, j)} \sum_{k \in K_e} \sum_{l=0}^{S_k} (N_k - \sum_{l=0}^{S_k} n_{k_l} + 1) \alpha_{k_l} b_{k_l} \delta_{k_l}(j) q(j - b_{k_l}) + \\ + \frac{1}{j} \sum_{k \in K_a} \sum_{l=0}^{S_k} (N_k - \sum_{l=0}^{S_k} n_{k_l} + 1) \alpha_{k_l} b_{k_l} \delta_{k_l}(j) q(j - b_{k_l}) & \text{for } j = 1, \dots, T \\ 0 & \text{otherwise} \end{cases}$$

un-normalized

$$G = \sum_{j=0}^T q(j)$$

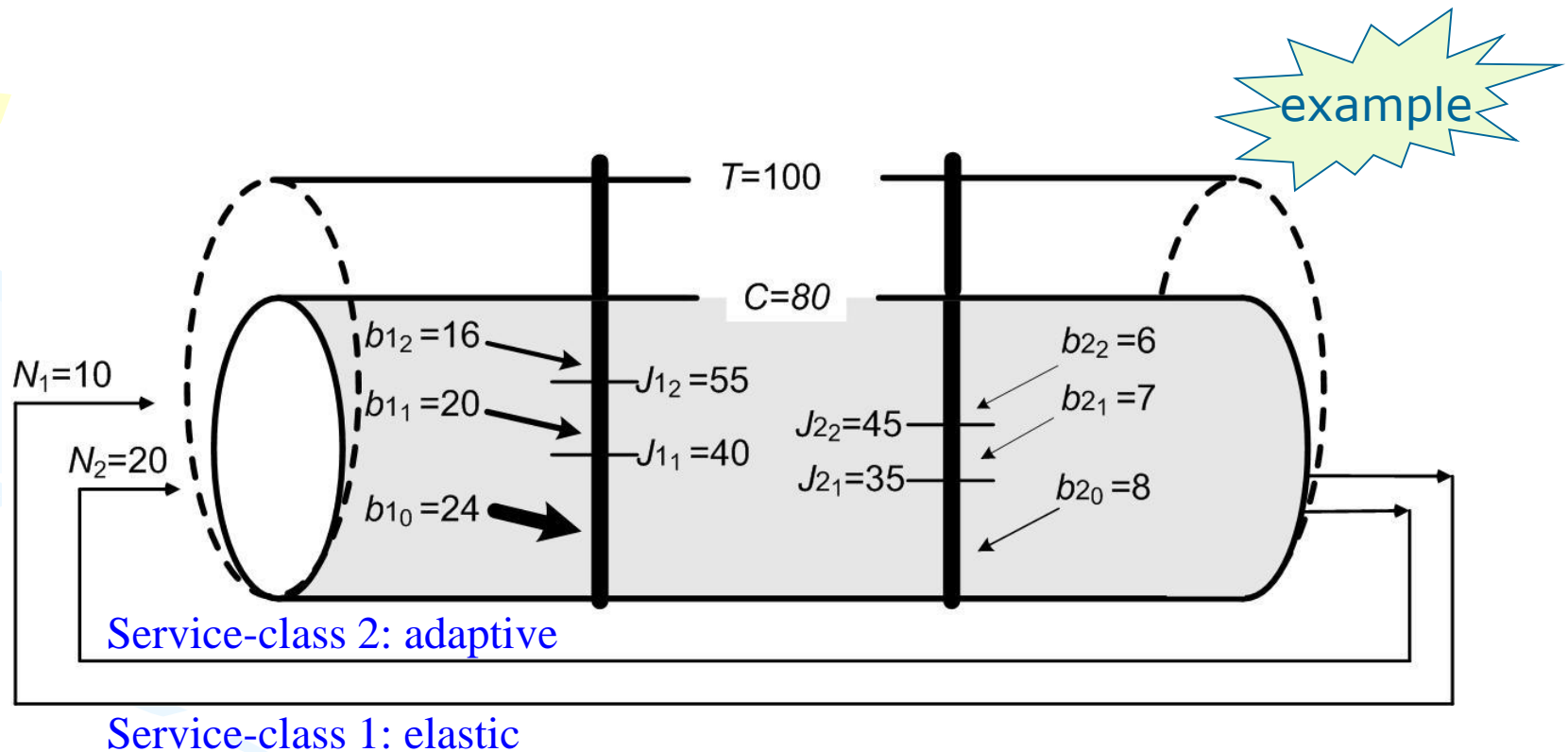
Time Congestion Probability

$$P_{b_k} = \sum_{j=T-b_k S_k+1}^T G^{-1} q(j)$$

Link Utilization

$$U = \sum_{j=1}^C j G^{-1} q(j) + \sum_{j=C+1}^T G^{-1} C q(j)$$

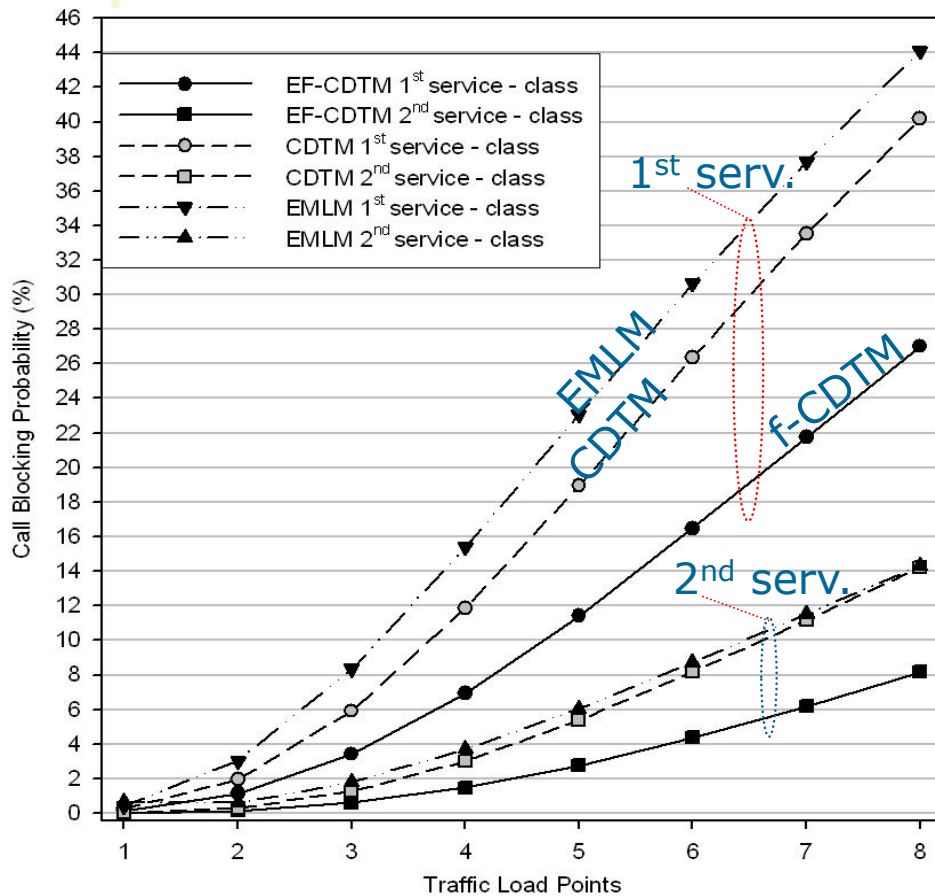
Ef-CDTM comparison with other models: EMLM, CDTM, E-CDTM



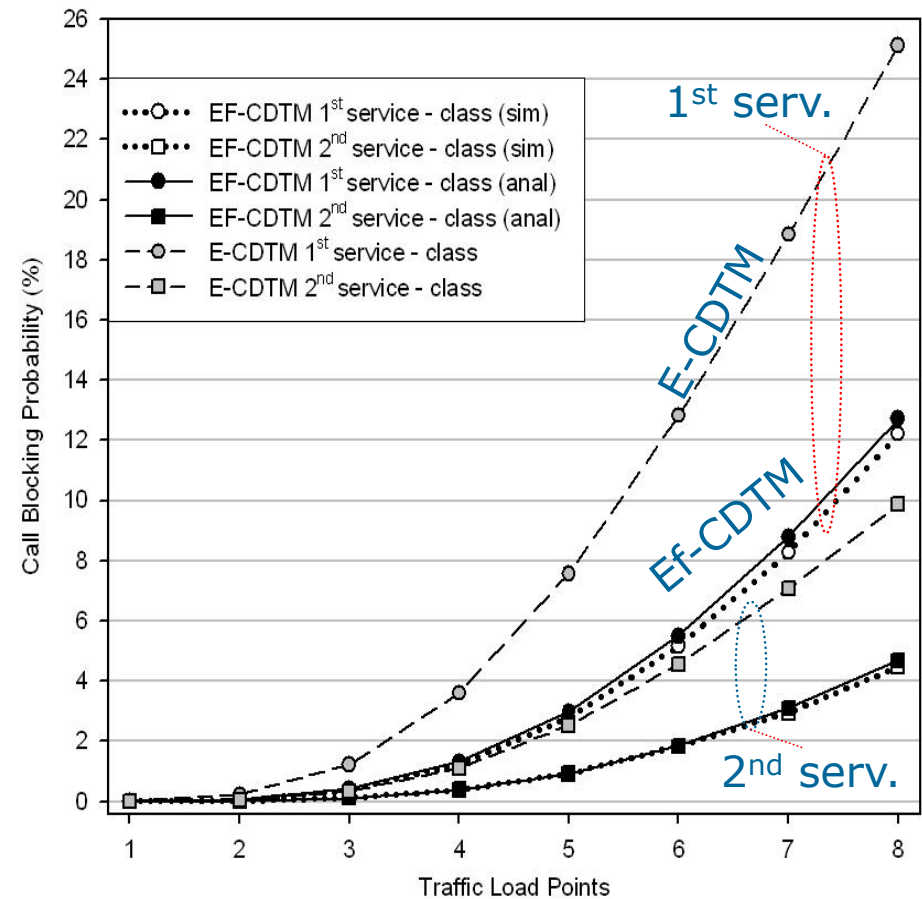
Offered Traffic-Load per idle source = 0.025 erl
Consequently, it increases by 0.025 erl

Ef-CDTM comparison with other models: EMLM, CDTM, E-CDTM (cont.)

T=C



T=C+20

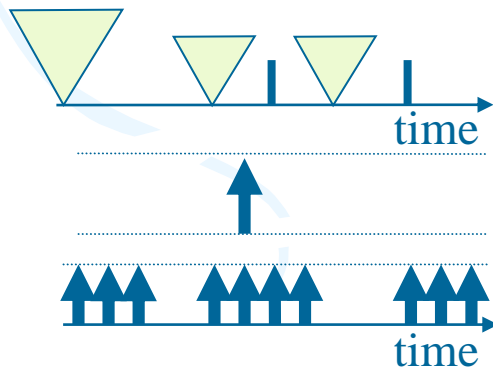


Ef-CDTM ↔ f-CDTM

Efficient teletraffic loss models (cont.2)

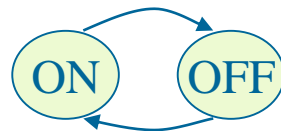
- **Teletraffic models of batched Poisson input**

- Batched Poisson arriving calls with fixed bandwidth requirements and *fixed bandwidth allocation during service*.
- Batched Poisson arriving calls with fixed bandwidth requirements and *elastic bandwidth during service*.
- Batched Poisson arriving calls with fixed bandwidth requirements that, *when in service, alternate between transmission periods (ON) and idle periods (OFF)*.



Batched Poisson arriving calls

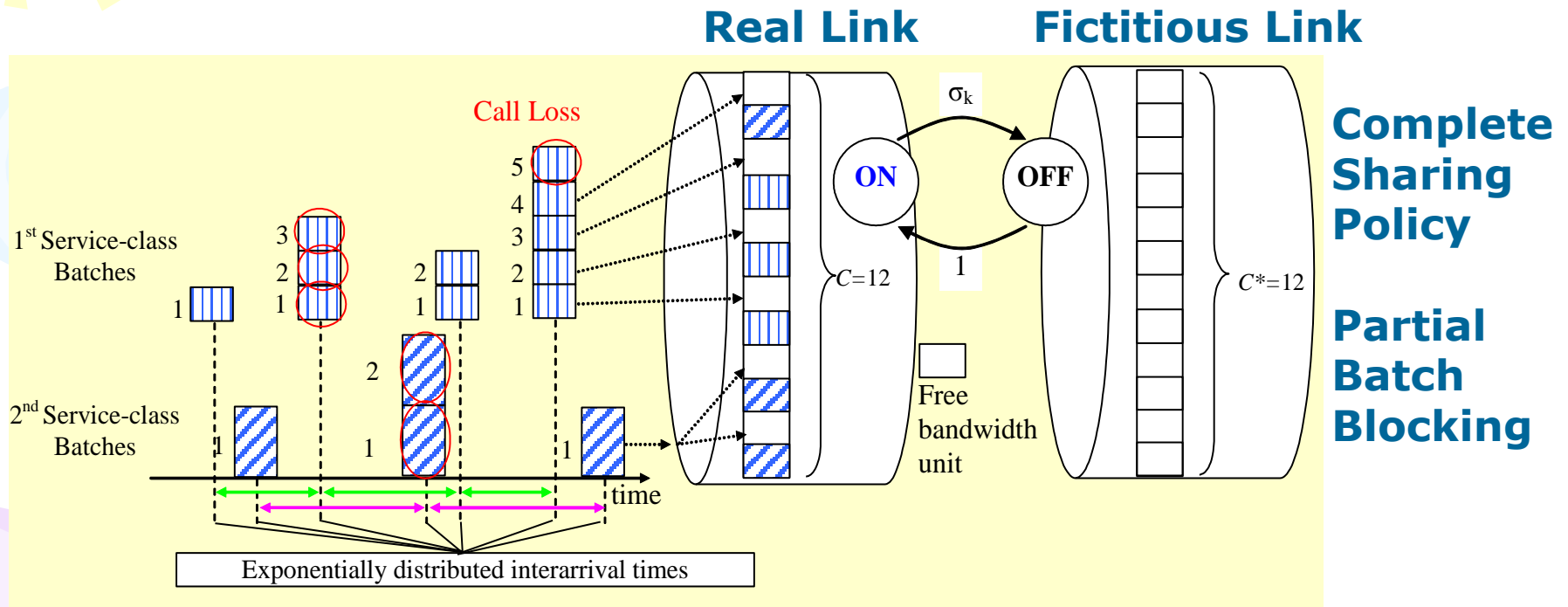
Fixed bandwidth requirements upon arrival



ON-OFF traffic, while in service

The Batched Poisson ON-OFF Model (BP-ON-OFF)

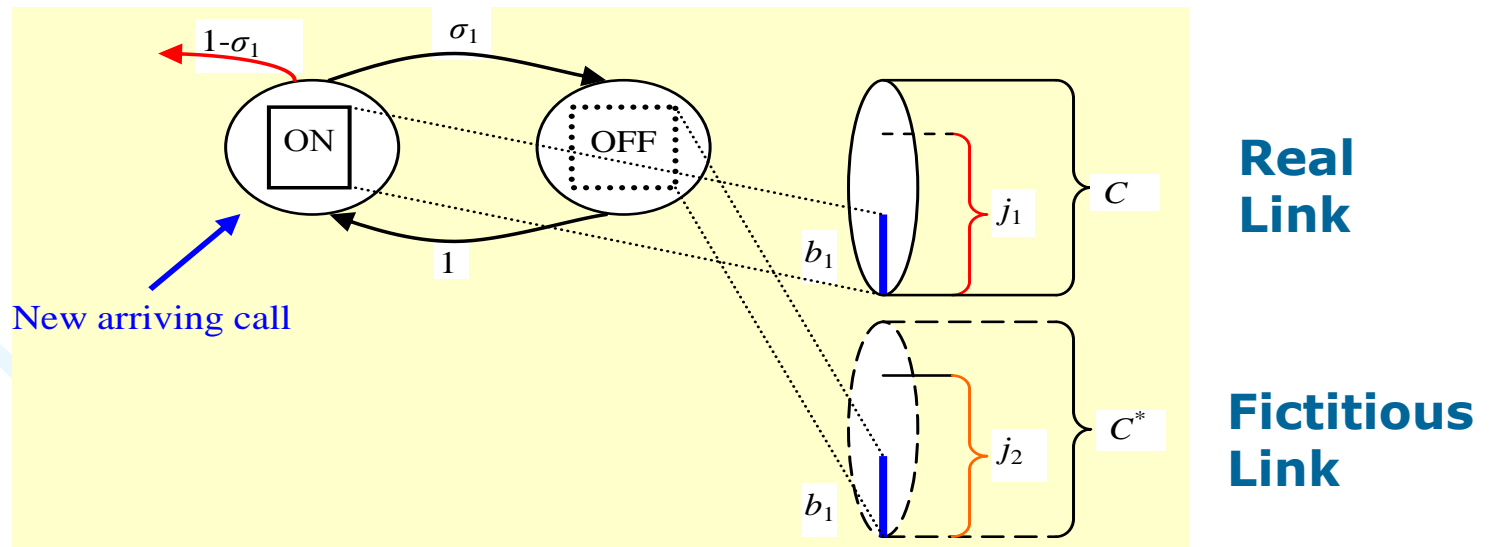
B_{kr} probability that there are r calls in an arriving batch of service-class k .



The incorporated ON-OFF model

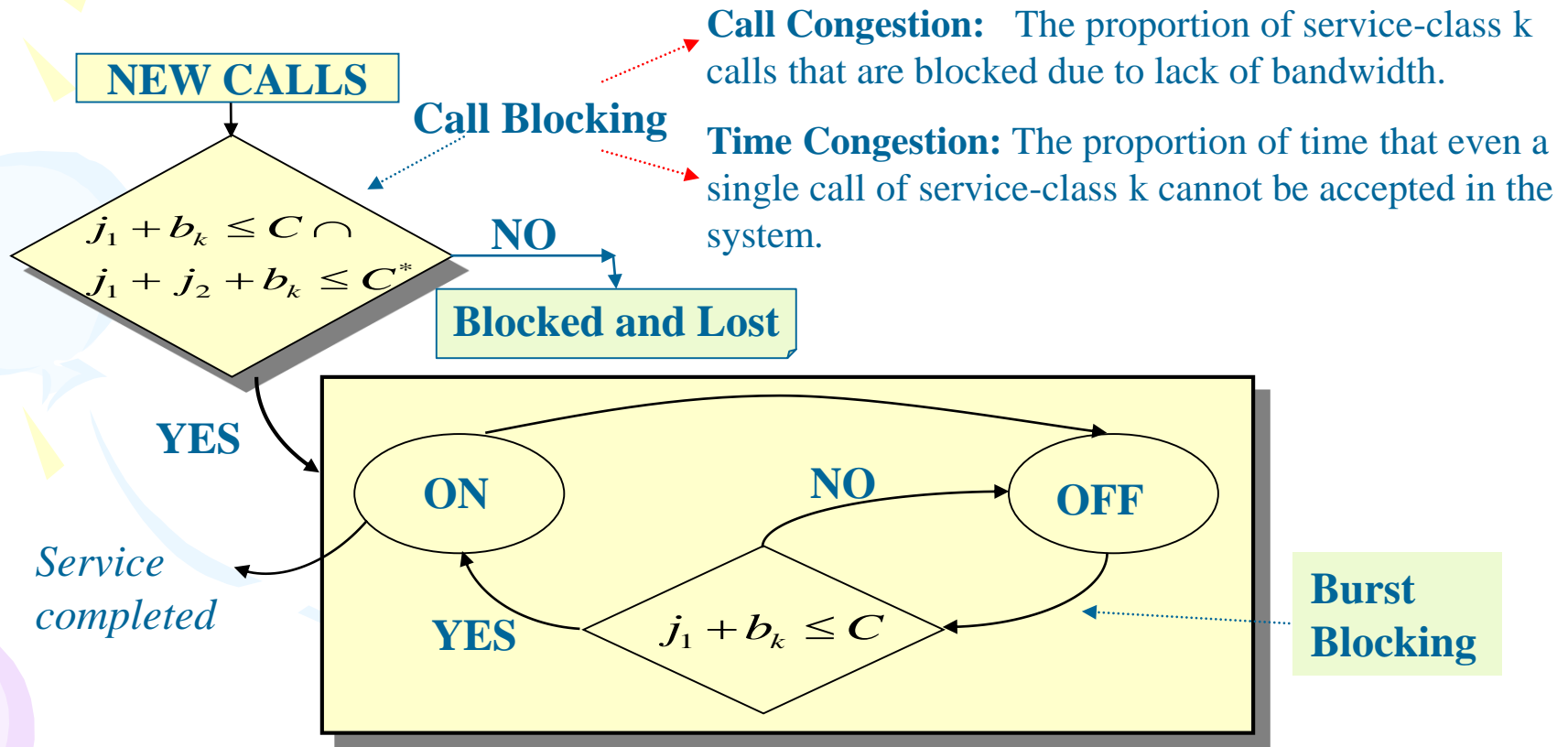
The system

- ✓ Real and Fictitious Link
- Fixed Bandwidth Requirement upon Arrival



The BP ON-OFF Model – CAC

Each call is serviced according to the following scheme:



The analytical BP-ON-OFF Model

Local Balance does not exist but “Local Flow Balance” does exist

$i = 1 \Rightarrow$ state ON

$i = 2 \Rightarrow$ state OFF

$$n^i = (n_1^i, \dots, n_k^i, \dots, n_K^i)$$

$$n_{k+l}^i = (n_1^i, \dots, n_k^i + l, \dots, n_K^i)$$

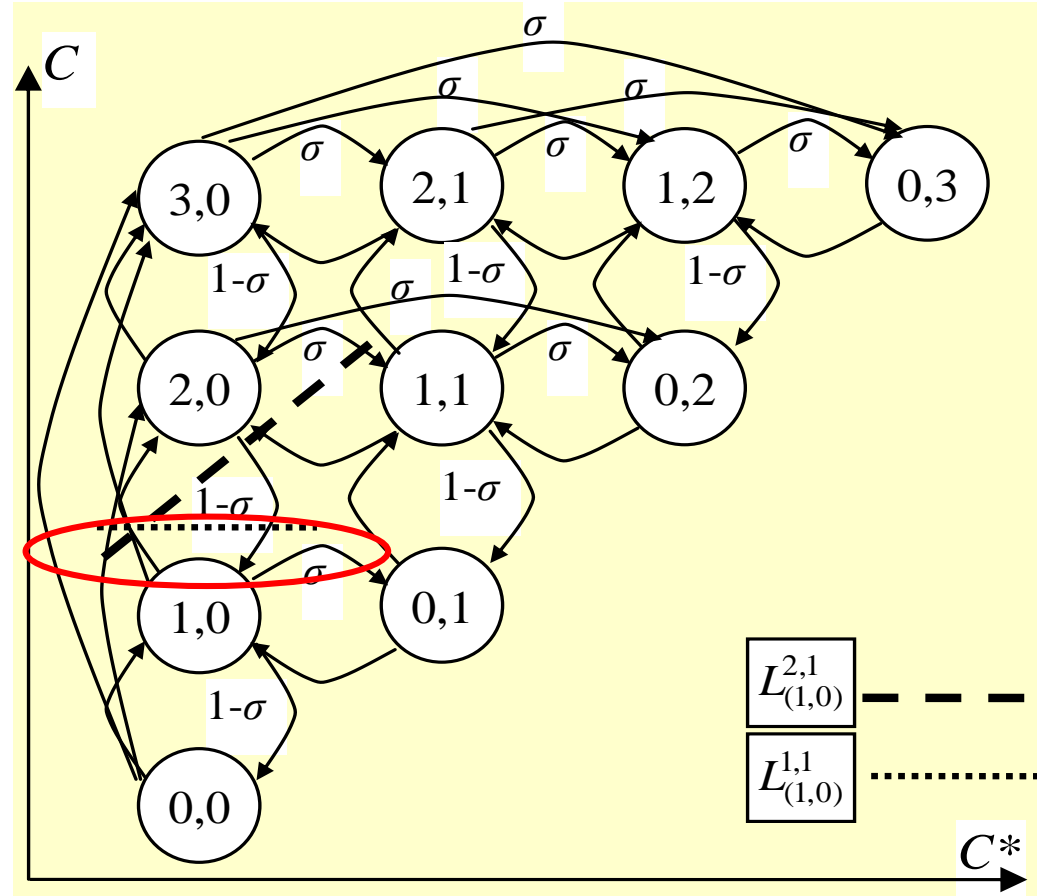
$$n_{k-l}^i = (n_1^i, \dots, n_k^i - l, \dots, n_K^i)$$

$$\vec{n}_{k+l}^1 = (n_{k+l}^1, n^2)$$

$$\vec{n}_{k-l}^1 = (n_{k-l}^1, n^2)$$

$$\vec{n}_{k+l}^2 = (n^1, n_{k+l}^2)$$

$$\vec{n}_{k-l}^2 = (n^1, n_{k-l}^2)$$



$L_n^{1,k}$ the level that separates the state-vector \vec{n} from the state \vec{n}_{k+1}^1

$L_n^{2,k}$ the level that separates the state-vector \vec{n} from the state \vec{n}_{k+1}^2

The analytical BP-ON-OFF Model (cont.1)

Flow_up = Flow_down (for each level)

$$f^{(up)}(\vec{L}_n^{1,k}) = \sum_{l=0}^{n_k^1} P(\vec{n}_{k-l}^1) \lambda_k \hat{B}_l^{(k)} \quad \text{upward probability flow across the level } \vec{L}_n^{1,k}$$

$$f^{(down)}(\vec{L}_n^{1,k}) = \mu_{1k} (n_k^1 + 1) (1 - \sigma_k) P(\vec{n}_{k+1}^1) \quad \text{downward probability flow across the level } \vec{L}_n^{1,k}$$

$$f^{(up)}(\vec{L}_n^{2,k}) = \sum_{l=0}^{n_k^2} P(\vec{n}_{k-l}^2) \lambda_k \sigma_k \hat{B}_l^{(k)} \quad \text{upward probability flow across the level } \vec{L}_n^{2,k}$$

$$f^{(down)}(\vec{L}_n^{2,k}) = \mu_{2k} (n_k^2 + 1) (1 - \sigma_k) P(\vec{n}_{k+1}^2) \quad \text{downward probability flow across the level } \vec{L}_n^{2,k}$$

$$f^{(up)}(\vec{L}_n^{1,k}) = f^{(down)}(\vec{L}_n^{1,k})$$

$$f^{(up)}(\vec{L}_n^{2,k}) = f^{(down)}(\vec{L}_n^{2,k})$$

Local Flow Balance equations

$$\begin{aligned} & f^{(up)}(\vec{L}_{n_{k-1}}^{1,k}) + f^{(down)}(\vec{L}_n^{1,k}) + f^{(up)}(\vec{L}_{n_{k-1}}^{2,k}) + f^{(down)}(\vec{L}_n^{2,k}) = \\ & = f^{(up)}(\vec{L}_n^{1,k}) + f^{(down)}(\vec{L}_{n_{k-1}}^{1,k}) + f^{(up)}(\vec{L}_n^{2,k}) + f^{(down)}(\vec{L}_{n_{k-1}}^{2,k}) \end{aligned}$$

Global Balance equation

The analytical BP-ON-OFF Model (cont.2)

Local Flow Balance \Rightarrow Product Form Solution

$$P(\vec{n}) = \frac{\prod_{i=1}^2 \prod_{k=1}^K P(n_k^i)}{G}$$

$$P(n_k^i) = \begin{cases} \sum_{l=1}^{n_k^i} p_{i,k} \frac{P(n_k^i - l) \hat{B}_{l-1}^{(k)}}{n_k^i} & \text{for } n_k^i \geq 1 \\ 1 & \text{for } n_k^i = 0 \end{cases}$$

where

$$\hat{B}_l^k = \sum_{r=l+1}^{\infty} B_r^k$$

complementary
batch size
distribution

and

$$p_{ik} = \frac{e_{ik}}{\mu_{ik}} = \begin{cases} \frac{\lambda_k}{(1-\sigma_k)\mu_{1k}} & \text{for } i=1 \\ \frac{\lambda_k \sigma_k}{(1-\sigma_k)\mu_{2k}} & \text{for } i=2 \end{cases}$$

utilization
of state i

The analytical BP-ON-OFF Model (cont.3)

Link occupancy distribution

Un-normalized

$$q(\vec{j}) = \begin{cases} 1 & \text{for } \vec{j} = 0 \\ \frac{1}{j_s} \sum_{i=1}^2 \sum_{k=1}^K b_{i,k,s} p_{ik} \sum_{l=1}^{\lfloor j_s/b_k \rfloor} \hat{B}_{l-1}^k q(\vec{j} - lB_{i,k}) & \text{for } j_1 = 1, \dots, C \text{ (if } s=1) \text{ or for } j_2 = 1, \dots, C^* - j_1 \text{ (if } s=2) \\ 0 & \text{otherwise} \end{cases}$$

Performance measures

Time Congestion probability

accurate

$$P_{b_k} = \sum_{\{\vec{j} | j_1 + j_2 + b_k > C\}} G^{-1} q(\vec{j})$$

$$G = \sum_{\vec{j} \in \Omega} q(\vec{j})$$

$$\Omega = \left\{ \vec{n} : \sum_{i=1}^2 \sum_{k=1}^K n_k^i b_k \leq C \right\}$$

Call Congestion probability

accurate

$$C_{b_k} = \frac{(p_{1k} + p_{2k}) \hat{B}_k - (\bar{n}_k^1 + \bar{n}_k^2)}{(p_{1k} + p_{2k}) \hat{B}_k}$$

Burst Blocking probability

approximate

$$P_{b_k}^* = \frac{\sum_{(\vec{j} \in \Omega^*)} y_{2k}(\vec{j}) q(\vec{j}) \mu_{2k}}{\sum_{(\vec{j} \in \Omega)} y_{2k}(\vec{j}) q(\vec{j}) \mu_{2k}}$$

$$y_{ik}(\vec{j}) = \frac{p_{i,k} \sum_{l=1}^{\lfloor j_s/b_k \rfloor} \hat{B}_{l-1}^k q(\vec{j} - lB_{i,k})}{q(\vec{j})}$$

BP-ON-OFF Model: Numerical example

$K = 2$ service-classes

$b_1 = 1, b_2 = 12$ b.u. per call

$C = C^* = 60$ b.u.

The batch size, s_k , is given by the geometric distribution, i.e. $P_r(s_k=r) = (1 - \beta_k)\beta_k^{r-1}$. In this example $\beta_1=0.2, \beta_2=0.5$.

Arrival rate: $\lambda_1=10, \lambda_2=2$

Call holding time, exponentially distributed:

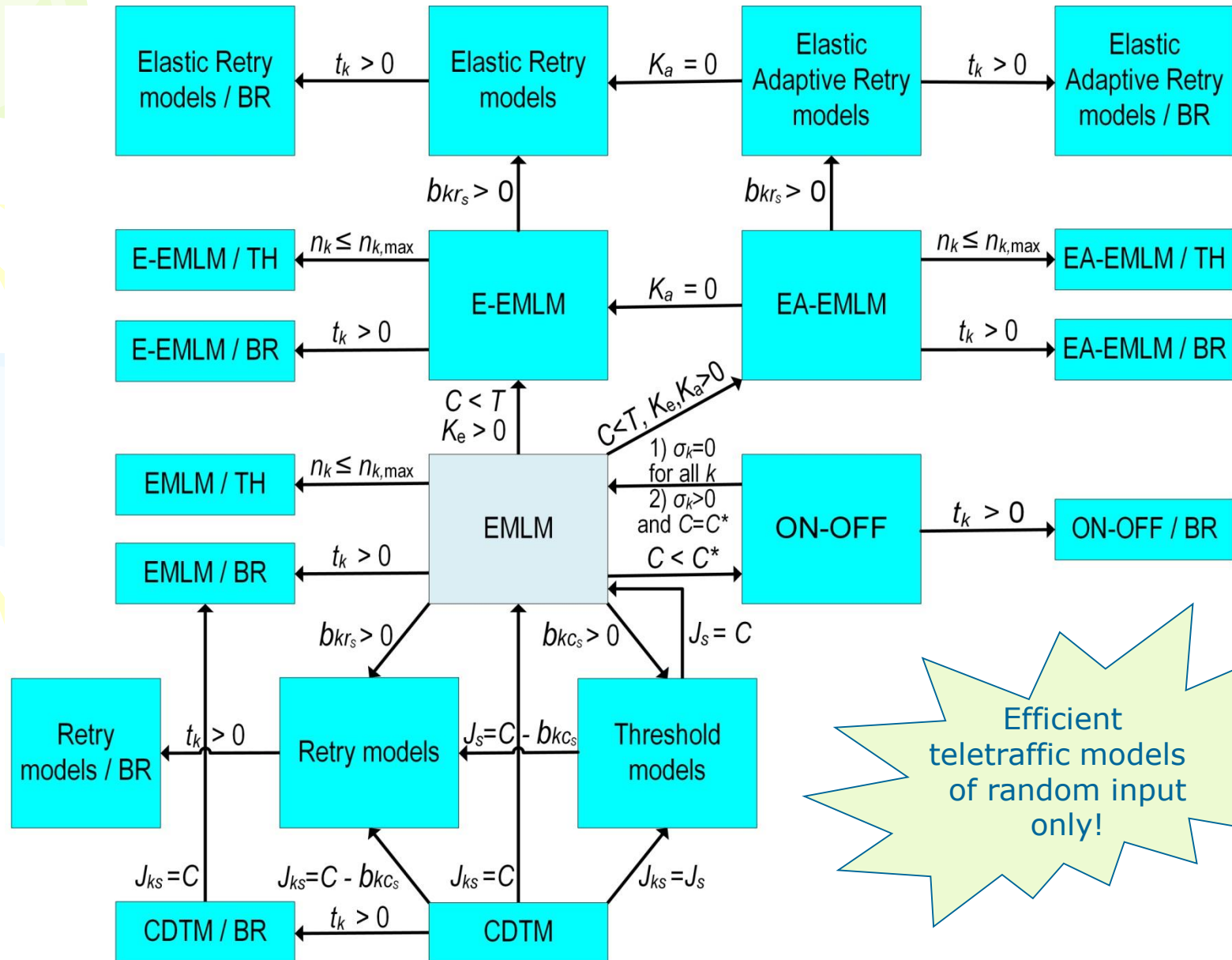
1st service-class: $\mu^{-1}_{11}=0.0405$ (state ON), $\mu^{-1}_{21}=0.01$ (state OFF).

2nd service-class: $\mu^{-1}_{12}=0.0405$ (state ON), $\mu^{-1}_{22}=0.01$ (state OFF).

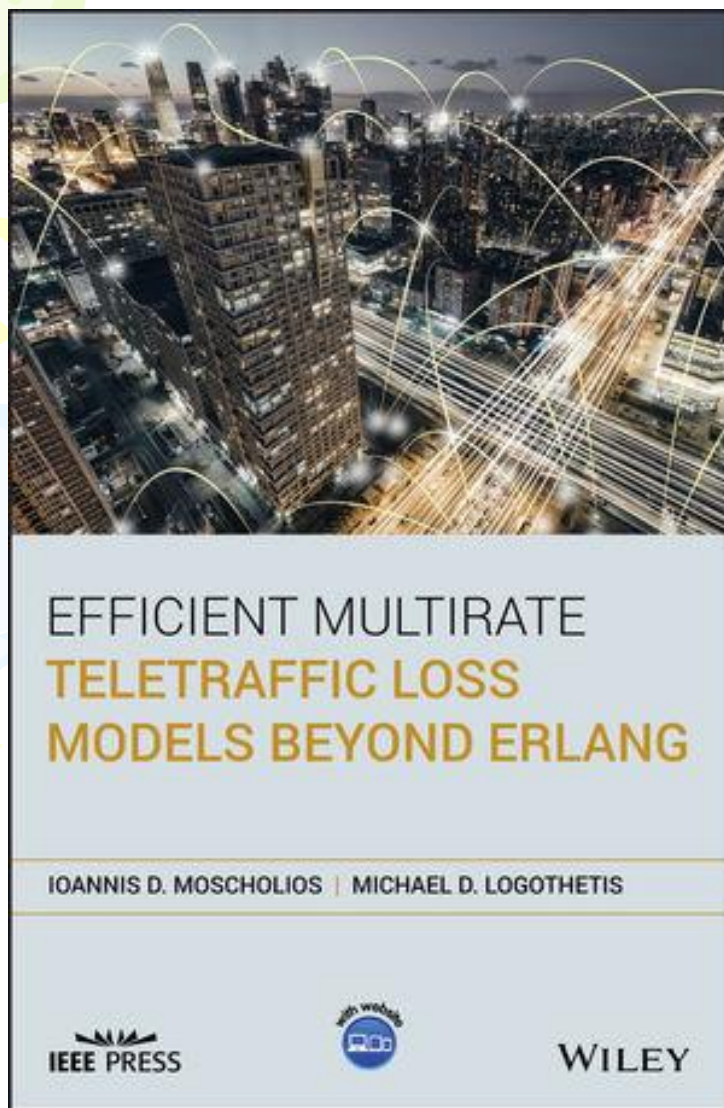
BP-ON-OFF Model: Numerical example

	Time Congestion (%)		Simulation results (%)	
λ_2	1 st class	2 nd class	1 st class	2 nd class
2	2.52	25.56	2.51 ± 0.06	25.55 ± 0.23
1.8	2.20	22.99	2.19 ± 0.07	23.02 ± 0.16
1.6	1.88	20.32	1.86 ± 0.07	20.29 ± 0.12
1.4	1.57	17.54	1.57 ± 0.06	17.55 ± 0.14
1.2	1.26	14.69	1.24 ± 0.03	14.65 ± 0.10
1.0	0.97	11.79	0.99 ± 0.01	11.77 ± 0.15
0.8	0.70	8.92	0.70 ± 0.01	8.92 ± 0.18

SUMMARY



References



Collaborators

- ❑ **Prof. Ioannis D. MOSCHOLIOS** (Associate Professor, UOP, Greece.)
- ❑ **Dr. Vassilis G. VASSILAKIS** (Lecturer, York University, U.K.)
- ❑ **Dr. Ioannis S. VARDAKAS** (Iquadrat S.A., Spain)
- ❑ **Mr. Georgios A. KALLOS** (British Telecom, U.K.)
- ❑ **Mr. Michael K. SIDIROPOULOS** (INTRACOM S.A., Greece.)

Thank You !