Chapter 4.1 Loss System M/M/n

Performance Evaluation of the Internet of Things (IoT)

Module Course: Performance Evaluation of Distributed Systems

Prof. Tobias Hoßfeld, Summer Semester 2022



Disclaimer and Copyright Notice

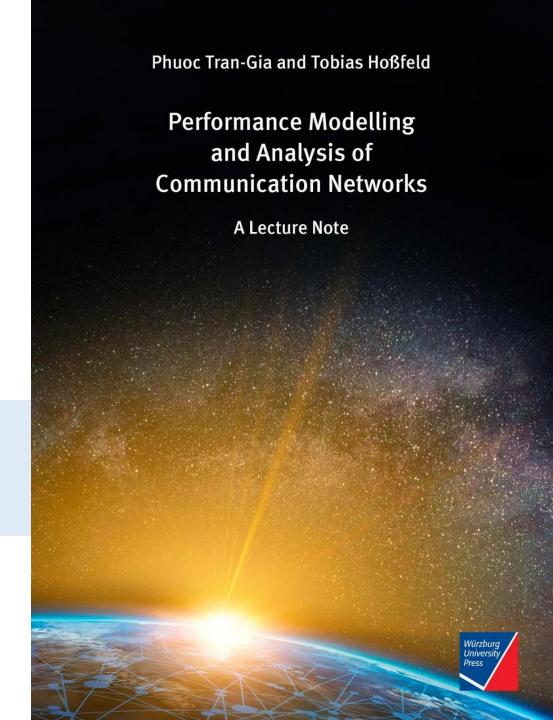
Lecture slides, figures, and scripts are based on the open access text book "Performance Modeling and Analysis of Communication Networks". The book and scripts are licensed under the Creative Commons License Attribution-ShareAlike 4.0 International (CC BY-SA 4.0). If you remix, transform, or build upon the material, you must distribute your contributions under the same license as the original.

The book must be cited and the disclaimer attached when using lectures slides or scripts.

Tran-Gia, P. & Hossfeld, T. (2021).
Performance Modeling and Analysis of Communication
Networks - A Lecture Note. Würzburg University Press.
https://doi.org/10.25972/WUP-978-3-95826-153-2

Website to download book, exercises, slides and scripts: https://modeling.systems/





Chapter 4

4 Analysis of Markovian Systems

- 4.1 Loss System M/M/n
 - 4.1.1 Model Structure and Parameters
 - 4.1.2 State Process and State Probabilities
 - 4.1.3 Other System Characteristics
 - 4.1.4 Generalization to Loss System M/GI/n
 - 4.1.5 Modeling Examples and Applications
- 4.2 Delay System M/M/n
 - 4.2.1 Model Structure and Parameters
 - 4.2.2 State Process and State Probabilities
 - 4.2.3 Other System Characteristics
 - 4.2.4 Delay Distribution
 - 4.2.5 Example: Single Server Delay System

- 4.3 Loss System with Finite Number of Sources
 - 4.3.1 Model Structure and Parameters
 - 4.3.2 State Process and State Probabilities
 - 4.3.3 Example: Mobile Cell with Finite Number of Sources
- 4.4 Customer Retrial Model with Finite Number of Sources
 - 4.4.1 Model Structure and Parameters
 - 4.4.2 Recursive Analysis Algorithm
 - 4.4.3 Calculation of Traffic Flows
 - 4.4.4 Example: Mobile Cell with Customer Retrials
- 4.5 Processor Sharing Model M/M/1-PS

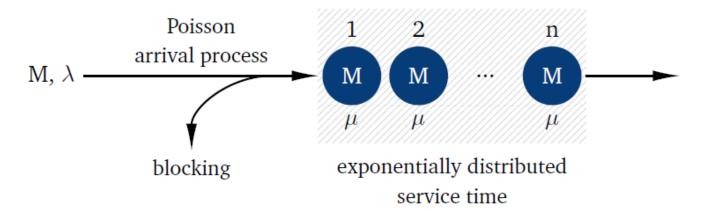


Model Structure and Parameters





Loss System M/M/n



- Interarrival time A with arrival rate λ
- Service time B with service rate μ
- Offered traffic $a = \frac{\lambda}{u}$ in pseudo-unit Erlang [Erl]

$$A(t) = P(A \le t) = 1 - e^{-\lambda t}, \quad E[A] = \frac{1}{\lambda}$$

$$A(t) = P(A \le t) = 1 - e^{-\lambda t}, \quad E[A] = \frac{1}{\lambda},$$

$$B(t) = P(B \le t) = 1 - e^{-\mu t}, \quad E[B] = \frac{1}{\mu}.$$

- Pure loss or blocking operation mode
 - arriving customer finding all servers occupied upon arrival will be blocked
 - blocked customers leave system: no further impact on the system state process





STATE PROCESS AND STATE PROBABILITIES

Erlang formula for loss systems

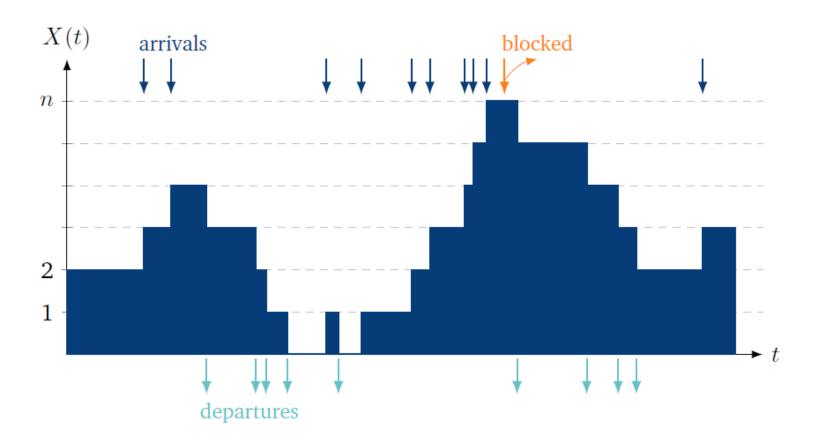




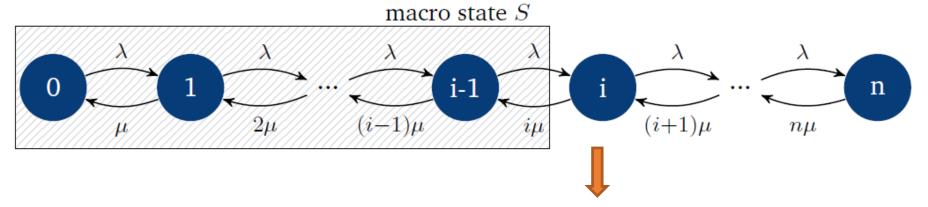
State Process of M/M/n Loss System

- \blacktriangleright X(t) number of customers or busy servers in system at time t
- Stationary system $\lim_{t \to \infty} X(t) = X$
- Steady-state probabilities (in statistical equilibrium)

$$x(i) = P(X=i)$$
, $i=0,1,...$



State Transition Diagram

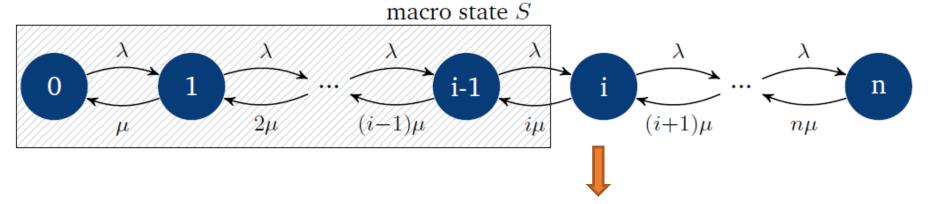


(micro) state [X = i] customers in the system

- ► Customer arrival: $[X = i] \rightarrow [X = i + 1]$ with rate λ
 - if arriving customer is accepted (i = 0, 1, ..., n 1)
- ► Customer departure or service termination: $[X = i] \rightarrow [X = i 1]$ with rate $i\mu$
 - service time of one of the i customers ends (i = 1, 2, ..., n)



Micro State Equations



- Rate for **leaving** state [X = i]
- $x(i) \cdot \lambda + x(i) \cdot i \cdot \mu$

Rate for **reaching** state [X = i]
$$x(i-1) \cdot \lambda + x(i+1) \cdot (i+1) \mu$$

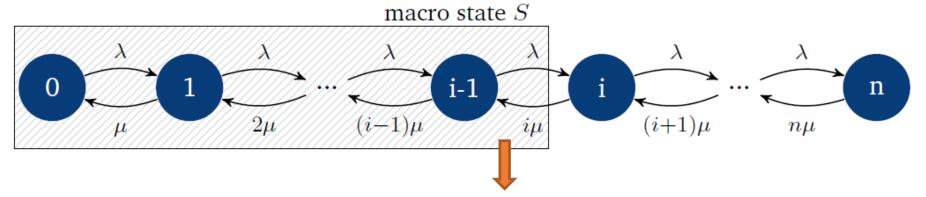
Micro state equation for [X = i]
$$x(i) \lambda + x(i) i\mu = x(i-1) \lambda + x(i+1)(i+1)\mu$$

▶ Special cases for [X = 0] and [X = n]
$$x(0)\lambda = x(1)\mu$$
 and $x(n-1)\lambda = x(n)\mu$

Normalization condition

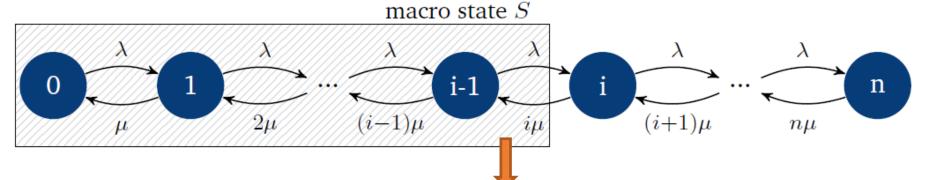
$$\sum_{i=1}^{n} x(i) = 1$$

Macro State Equations



- ► Macro state $S = \{0,1,...,i-1\}$ for i = 1,2,...,n
- Macro state equations $\begin{cases} \lambda x(i-1) = i \mu x(i), & i = 1, 2, ..., n \\ \sum_{i=0}^{n} x(i) = 1 \end{cases}$

Macro State Equations: Solution



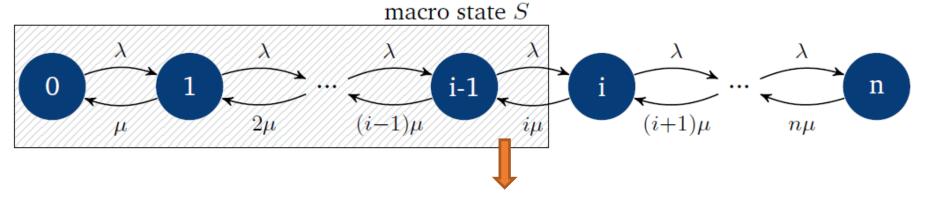
Macro state equations

$$\lambda x(i-1) = i \mu x(i)$$
, $i = 1, 2, ..., n$

$$\sum_{i=0}^{n} x(i) = 1$$



Macro State Equations and Erlang Formula for Loss Systems



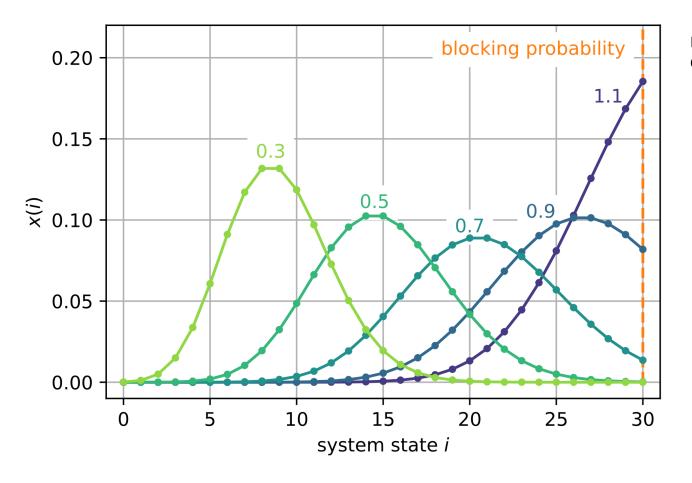
- ► Macro state $S = \{0,1,...,i-1\}$ for i = 1,2,...,n
- Macro state equations $\begin{cases} \lambda x(i-1) = i \mu x(i), & i = 1, 2, ..., n \\ \sum_{i=0}^{n} x(i) = 1 \end{cases}$
- ► Erlang formula for loss systems with offered traffic $a = \frac{\lambda}{\mu}$

$$x(i) = \frac{\frac{a^{i}}{i!}}{\sum_{k=0}^{n} \frac{a^{k}}{k!}}$$



Steady State Distribution

► M/M/30-0 with normalized offered load $\rho = \frac{a}{n} = \frac{\lambda}{n\mu}$



normalized offered load ρ

1.1 --- 0.9 --- 0.7 --- 0.5 --- 0.3

PASTA property

$$x_A(i) = x(i)$$

Blocking probability

$$p_B = x_A(n) = x(n)$$



State Probability at Arbitrary and Arrival Times

- ▶ State probability at **arrival time** $x_A(i)$
 - arriving customers sees i other customers in the system
- ► State probability at **arbitrary time** $x^*(i)$
 - stationary system is in state i at arbitrary time
 - time-averaged steady state probability
 - analysis of Markov state processes leads to $x(i) = x^*(i)$
- ▶ Due to memoryless property of the Poisson arrival process
 - $x_A(i) = x(i), i = 0, 1, ..., n$
 - PASTA property: Poisson arrivals see time averages





OTHER SYSTEM CHARACTERISTICS

Blocking probability (Erlang-B formula), carried traffic, utilization





System Characteristics

- ► Erlang-B formula
 - blocking probability
 - PASTA property utilized $x_A(i) = x(i)$
- Carried traffic Y
 - mean number of occupied servers Y = E[X]
 - derived using Little's law and traffic flows

- offered traffic $a = \frac{\lambda}{\mu}$
- normalized offered traffic $\rho = \frac{a}{n}$

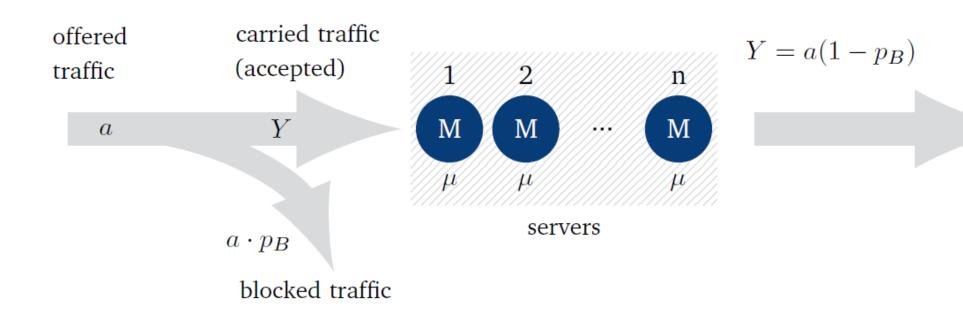
$$p_{B} = x_{A}(n) = x(n) = \frac{\frac{a^{n}}{n!}}{\sum_{k=0}^{n} \frac{a^{k}}{k!}}$$

$$Y = \sum_{i=0}^{n} i x(i) = \lambda (1 - p_B) \frac{1}{\mu} = a (1 - p_B)$$

$$\frac{Y}{n} = \rho \cdot (1 - p_B)$$



Traffic Flows in M/M/n Loss System



- ▶ Little's law $E[X] = \lambda' \cdot E[T]$
 - arrival rate $\lambda(1 p_B) = \lambda'$
 - mean sojourn time E[B] = E[T]
 - mean number of customers $Y = E[X] = a \cdot (1 p_B)$





GENERALIZATION TO M/GI/N-0

Insensitivity or robustness property





Insensitivity or Robustness Property

- Erlang formula provides steady state probabilities for M/M/n loss system
- Generalization to M/GI/n-0
 - Erlang formula is also valid for loss systems M/Gl/n
 - proof in literature, e.g. by Syski
- **Insensitivity or robustness property**
 - extends applicability of Erlang formula and Erlang-B formula considerably in practice

$$x(i) = \frac{\frac{a^{i}}{i!}}{\sum_{k=0}^{n} \frac{a^{k}}{k!}}$$

$$x(i) = \frac{\frac{a^{i}}{i!}}{\sum_{k=0}^{n} \frac{a^{k}}{k!}} \qquad p_{B} = x_{A}(n) = x(n) = \frac{\frac{a^{n}}{n!}}{\sum_{k=0}^{n} \frac{a^{k}}{k!}}$$



Example: M/GI/1-0



Modeling Examples and Applications

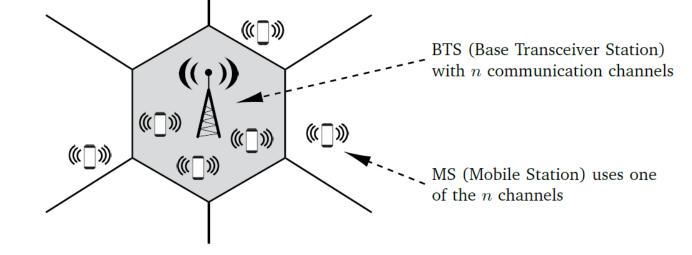
Dimensioning, economy of scale





Dimensioning of Systems

- Trunk size in telephone networks
 - best known and earliest application of M/GI/n loss system
 - dimension number of trunk groups
- ► GSM cell with n communication channels
 - dimension number of channels
- ▶ DHCP server in ISP environment
 - dimension number of IP addresses

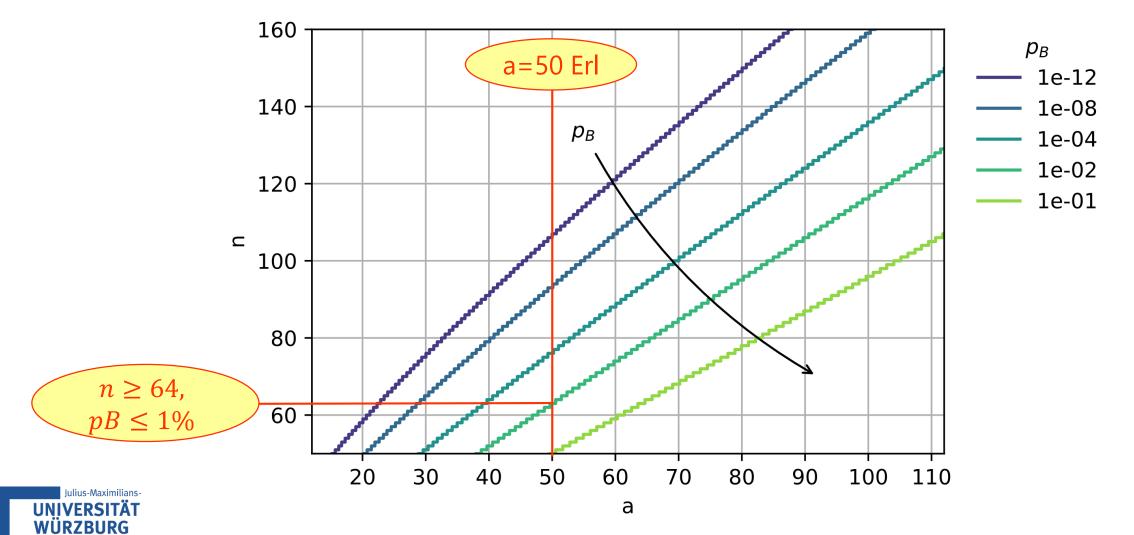


- Dimensioning means
 - finding number n of servers such that the
 - blocking probability is below a predefined value according to Service Level Agreements (SLAs) $p_B < \epsilon$



Numerical Solution: Dimensioning

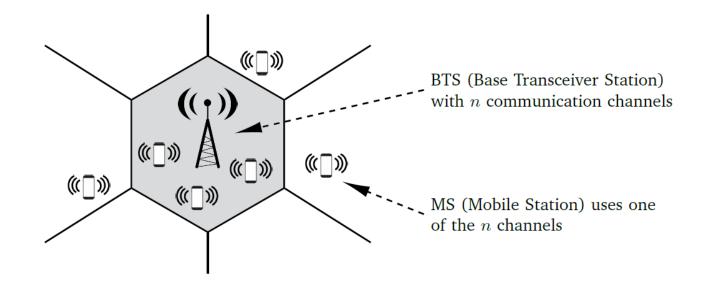
► Efficient implementation of Erlang-B formula: see script "4.1 M/M/n-0 loss system: Erlang-B formula" https://modeling.systems/





Example: GSM Cell

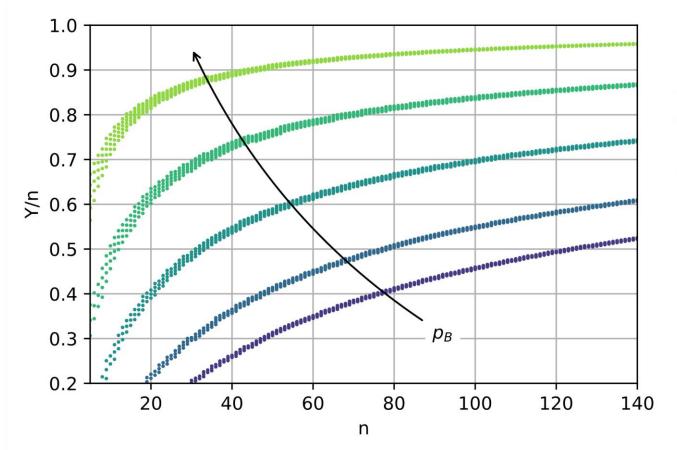
- \triangleright *n* number of channels
- ► B call duration with E[B] = 90 s
- λ arrival rate of Poisson process with $\lambda = 1$ call/s
- $\begin{array}{cc} \bullet & a & \text{offered load} \\ a = 90 \text{ Erl} \end{array}$



- ▶ Goal: Quality of Service (QoS) threshold $p_B \le 10^{-3}$
- Result: n = 117 channels required



Economy of Scale

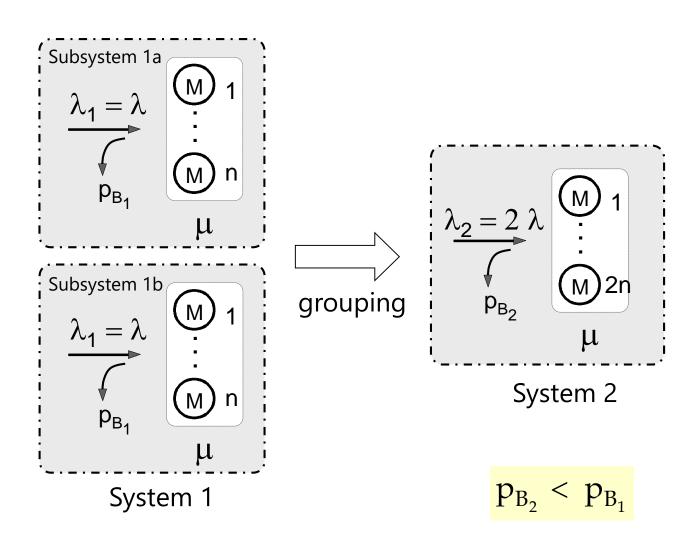


- *p_B* 1e-01
- 1e-02
- 1e-04
- 1e-08
- 1e-12
 - The server utilization $\frac{Y}{n}$ increases with number of servers, when we keep required QoS (p_B) constant
 - Larger trunk sizes are more economical
 - ► Increase of factor $\frac{Y}{n}$ corresponds to the **economy of scale**.
 - ► Increase is limited (flatter shape)





Economy of Scale in Loss Systems





Example: Several Cells

