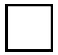
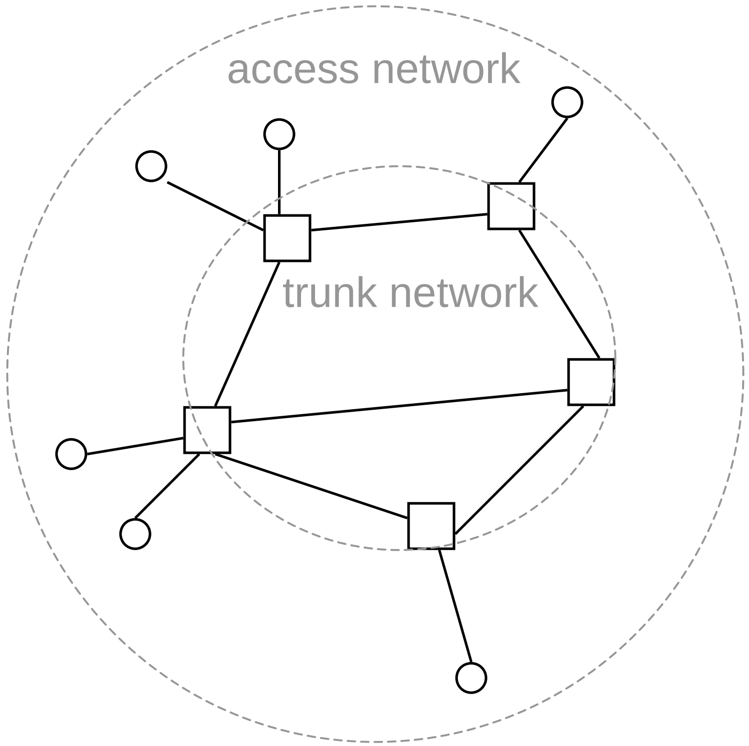
10. Network planning and dimensioning

Contents

- Introduction
- Network planning
- Traffic forecasts
- Dimensioning

Telecommunication network

- A simple model of a telecommunication network consists of
  - **nodes**
    - terminals 
    - network nodes 
  - **links** between nodes
- **Access network**
  - connects the terminals to the network nodes
- **Trunk network**
  - connects the network nodes to each other



Why network planning and dimensioning?

- "The purpose of dimensioning of a telecommunications network is to ensure that

the expected needs will be met in an economical way

both for subscribers and operators."

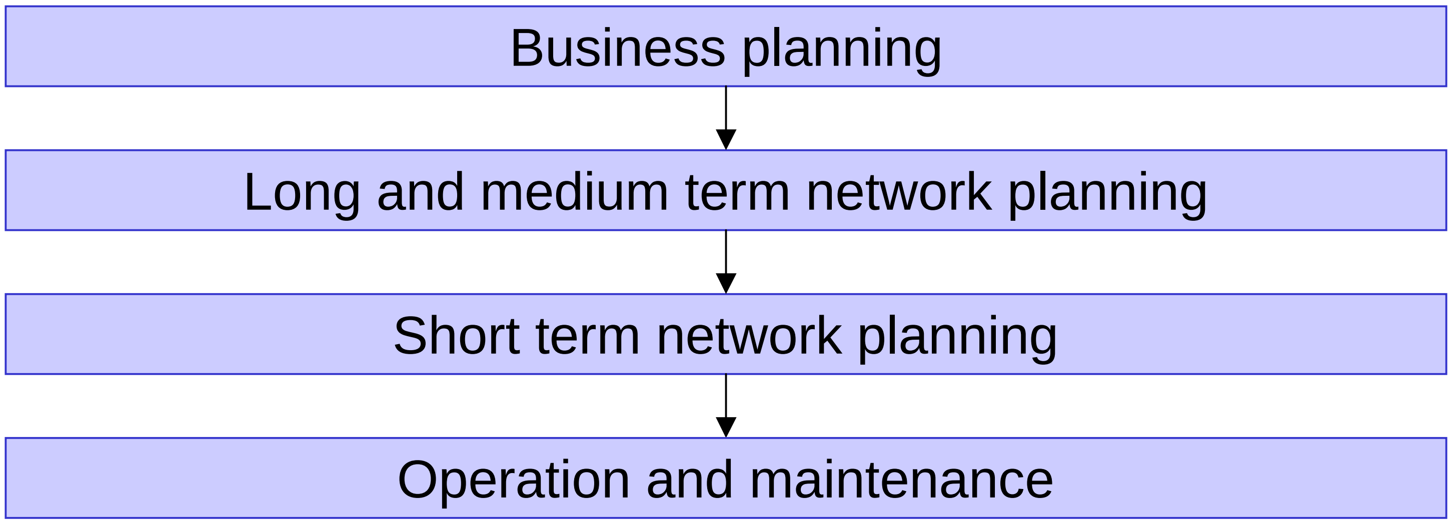
Source: [1]

Contents

- Introduction
- Network planning
- Traffic forecasts
- Traffic dimensioning

Network planning in a stable environment (1)

- Traditional planning model:

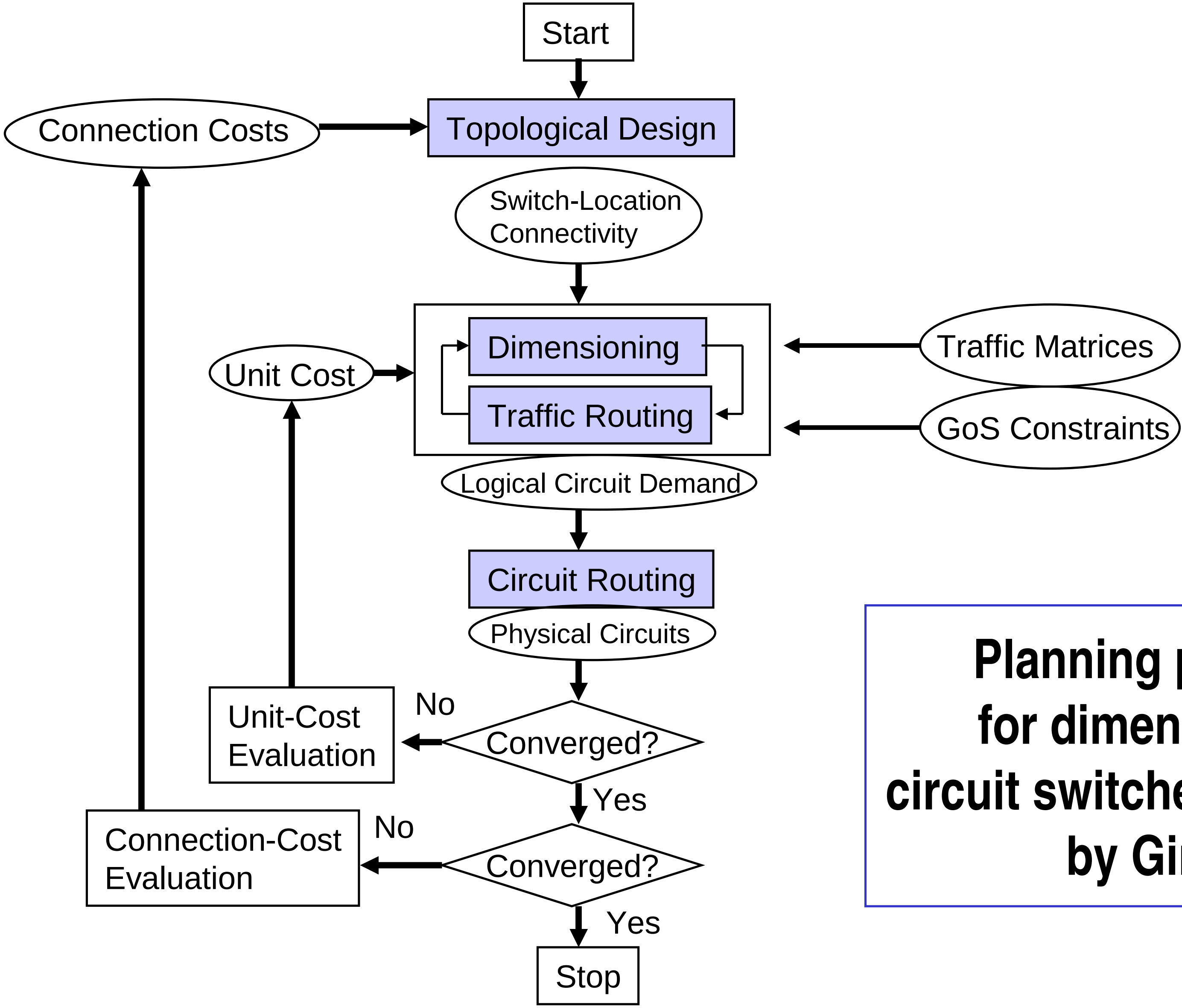


Source: [1]

Network planning in a stable environment (2)

- Traffic aspects
  - Data collection (current status)
    - traffic measurements
    - subscriber amounts and distribution
  - Forecasting
    - service scenarios
    - traffic volumes and profiles
- Economical aspects
- Technical aspects
- Network optimisation and dimensioning
  - hierarchical structure and topology
  - traffic routing and dimensioning
  - circuit routing

Source: [1]



Planning process for dimensioning circuit switched networks by Girard

Source: [2]

Network planning in a turbulent environment (1)

- Additional decision data are needed from the following areas:
  - The market, with regard to a specific business concept
    - due to competition!
    - operator's future role (niche): dominance/co-operation
  - Customer demands:
    - new services: Internet & mobility (first of all)
    - new business opportunities
  - Technology:
    - new technology: ATM, xDSL, GSM, CDMA, WDM
  - Standards:
    - new standards issued continuously
  - Operations and network planning support:
    - new computer-aided means
  - Costs:
    - trends: equipment costs going down, staff costs going up

Source: [1]

Network planning in a turbulent environment (2)

- Safeguards for the operator:
  - Change the network architecture so that it will be more **open**, with generic **platforms**, if possible
  - Build the network with a certain prognosticated overcapacity (**redundancy**) in generic parts where the marginal costs are low

- New planning situation (shift of focus to a strategic-tactical approach):

Business planning; Strategic-tactical planning of network resources for **flexible use**

↓

Business-driven, dynamic network management for **optimal use** of network resources

Source: [1]

Contents

- Introduction
- Network planning
- Traffic forecasts
- Traffic dimensioning

Need for traffic measurements and forecasts

- To properly dimension the network we need to

estimate the traffic offered

- If the network is already operating,
  - the current traffic is most precisely estimated by making traffic measurements
- Otherwise, the estimation should be based on other information, e.g.
  - estimations on characteristic traffic generated by a subscriber
  - estimations on the number of subscribers
- Long time-span of network investments $\Rightarrow$
  - it is not enough to estimate only the current traffic
  - forecasts of future traffic are also needed

Traffic forecasting

- Information about future demands for telecommunications
  - an estimation of future tendency or direction
- Purpose
  - provide a basis for decisions on investments in network
- Forecast periods
  - time aspect important (reliability)
  - need for forecast periods of different lengths

Source: [1]

Traffic forecast

- Traffic forecast defines
  - the estimated traffic growth in the network over the planning period
- Starting point:
  - current traffic volume during busy hour (measured/estimated)
- Other affecting factors:
  - changes in the number of subscribers
  - change in traffic per subscriber (characteristic traffic)
- Final result (that is, the forecast):
  - **traffic matrix** describing the **traffic interest** between exchanges (traffic areas)

Traffic matrix

- Traffic matrix $T = (T(i,j))$
  - describes traffic interest between exchanges
  - N^2 elements (N = nr of exchanges)
  - element $T(i,i)$ tells the estimated traffic within exchange i (in Erlangs)
  - element $T(i,j)$ tells the estimated traffic from exchange i to exchange j (in Erlangs)
- Problem:
  - easily grows too big: 600 exchanges $\Rightarrow$ 360,000 elements!
- Solution: hierarchical representation
  - higher level: traffic between traffic areas
  - lower level: traffic between exchanges within one traffic area

Example (1): one local exchange

- **Data**:
  - There are 1000 private subscribers and 10 companies with their own PBX's in the area of a local exchange.
  - The characteristic traffic generated by a private subscriber and a company are estimated to be 0.025 erlang and 0.200 erlang, respectively.
- **Questions**:
  - What is the total traffic intensity a generated by all these subscribers?
  - What is the call arrival rate λ assumed that the mean holding time is 3 minutes?
- **Answers**:
  - a = 1000 * 0.025 + 10 * 0.200 = 25 + 2 = **27 erlang**
  - h = 3 min
  - $\lambda = a/h$ = 27/3 calls/min = **9 calls/min**

Example (2): one local exchange

- **Data**:
  - In a 5-year forecasting period the number of new subscribers is estimated to grow linearly with rate 100 subscribers/year.
  - The characteristic traffic generated by a private subscriber is assumed to grow to value 0.040 erlang.
  - The total nr of companies with their own PBX is estimated to be 20 at the end of the forecasting period.
- **Question**:
  - What is the estimated total traffic intensity a at the end of the forecasting period?
- **Answer**:
  - a = (1000 + 5*100)* 0.040 + 20 * 0.200 = 60 + 4 = **64 erlang**

Example (3): many local exchanges

- **Data**:
  - Assume that there are three similar local exhanges.
  - Assume further that one half of the traffic generated by a local exchange is local traffic and the other half is directed uniformly to the two other exchanges.
- **Question**:
  - Construct the traffic matrix T describing the traffic interest between the exchanges at the end of the forecasting period.

- **Answer:**
  - $T(i,i)$ = 64/2 = 32 erlang
  - $T(i,j)$ = 64/4 = 16 erlang

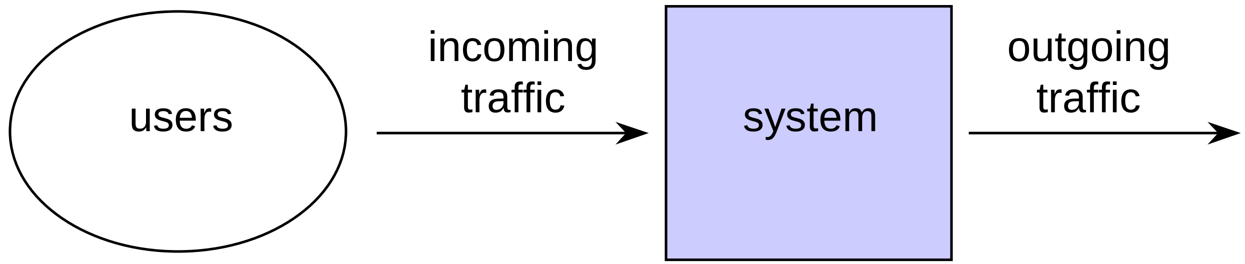
area	1	2	3	sum
1	32	16	16	64
2	16	32	16	64
3	16	16	32	64
sum	64	64	64	192

Contents

- Introduction
- Network planning
- Traffic forecasts
- Traffic dimensioning

Traffic dimensioning (1)

- Telecommunications system from the traffic point of view:



- Basic task in **traffic dimensioning**:

Determine the minimum **system capacity** needed
in order that the incoming **traffic** meet
the specified **grade of service**

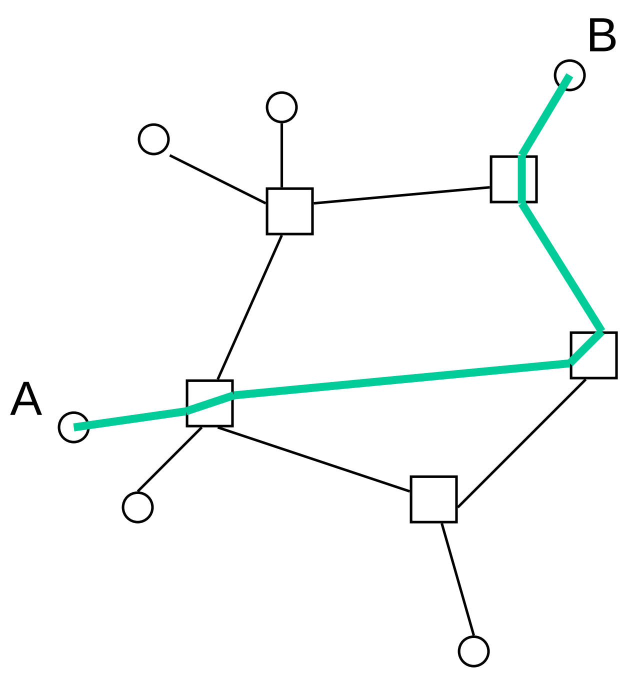
Traffic dimensioning (2)

- Observation:
  - Traffic is varying in time
- General rule:
  - Dimensioning should be based on peak traffic not on average traffic
- However,
  - Revenues are based on average traffic
- For dimensioning (of telephone networks),
  peak traffic is defined via the concept of busy hour:

Busy hour ≈ the continuous 1-hour period
for which the traffic volume is greatest

Telephone network model

- Simple model of a telephone network consists of
  - network nodes (exchanges)
  - links between nodes
- Traffic consists of **calls**
- Each call has two phases
  - first, the connection has to set up through the network (**call establishment** phase)
  - only after that, the information transfer is possible (**information transfer** phase)

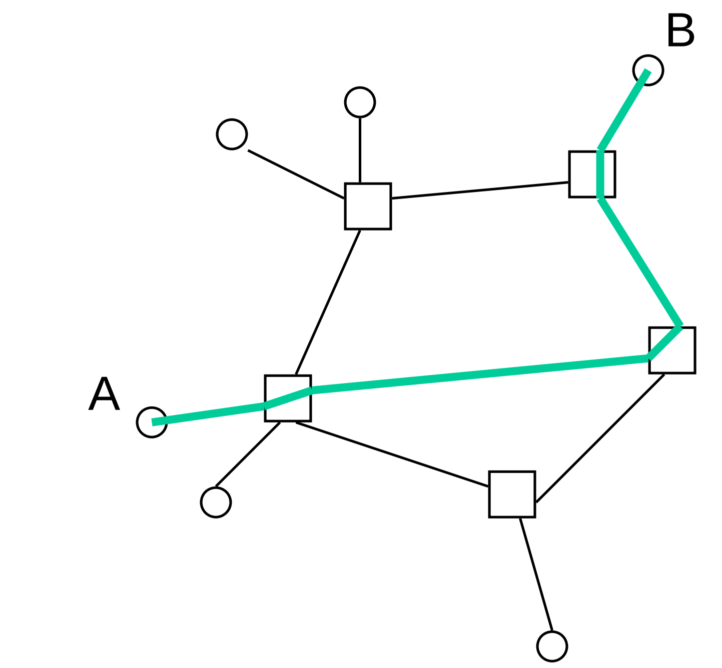


Two kinds of traffic processes

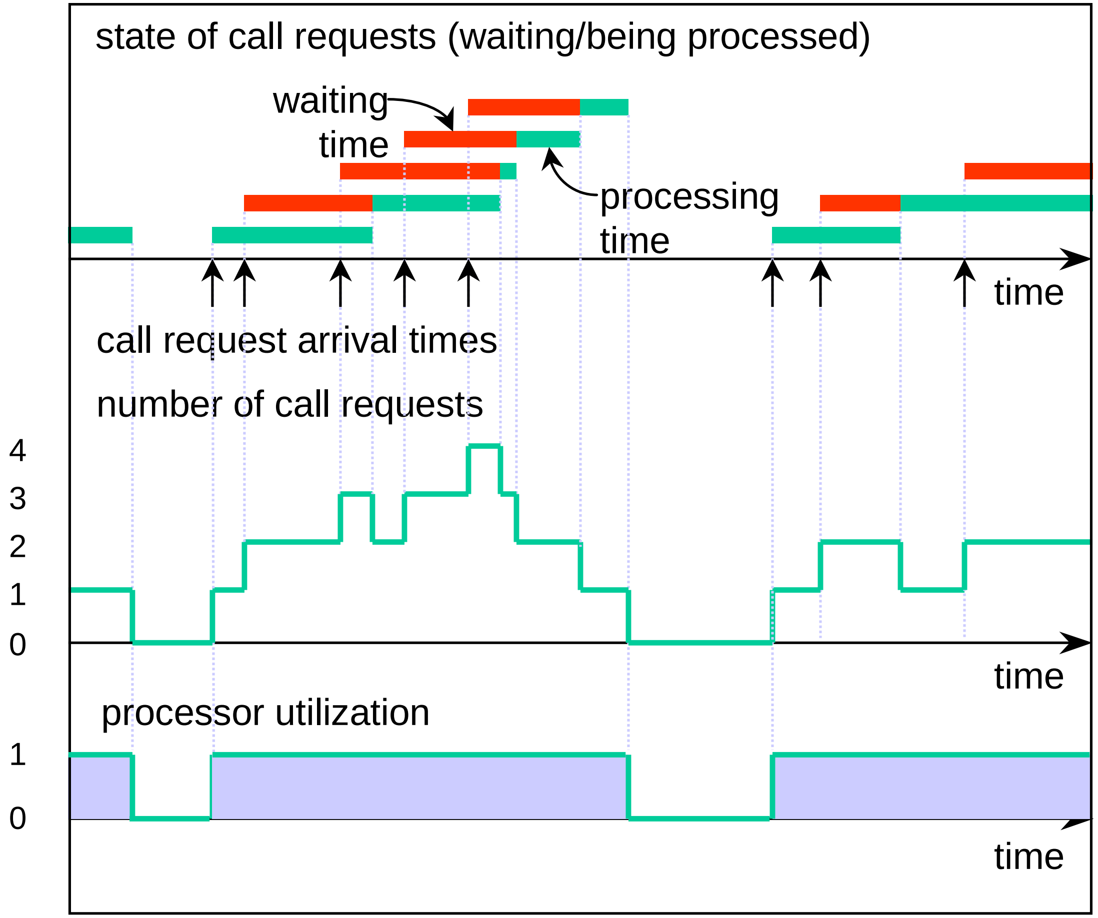
- Traffic process in each **network node**
  - due to **call establishments**
  - during the call establishment phase
    - each call needs (and competes for) processing resources in each network node (switch) along its route
  - it typically takes **some seconds** (during which the call is processed in the switches, say, **some milliseconds**)
- Traffic process in each **link**
  - due to **information transfer**
  - during the information transfer phase
    - each call occupies one channel on each link along its route
  - information transfer lasts as long as one of the participants disconnects
    - ordinary telephone calls typically hold **some minutes**
- **Note:** totally **different time scales** of the two processes

Simplified traffic dimensioning in a telephone network

- Assume
  - fixed topology and routing
  - given traffic matrix
  - given GoS requirements
- Dimensioning of network nodes: Determine the required **call handling capacity**
  - max number of call establishments the node can handle in a time unit
- Dimensioning of links: Determine the required **number of channels**
  - max number of ongoing calls on the link



Traffic process during call establishment (1)



Traffic process during call establishment (2)

- Call (request) arrival process is modelled as
  - a Poisson process with intensity λ
- Further we assume that call processing times are
  - IID and exponentially distributed with mean s
    - typically s is in the range of **milliseconds** (not minutes as h)
    - s is more a **system parameter** than a traffic parameter
- Finally we assume that the call requests are processed by
  - a single processor with an infinite buffer
    - $1/s$ tells the processing rate of call requests
- The resulting traffic process model is
  - the **M/M/1 queueing model** with traffic load $\rho = \lambda s$

Traffic process during call establishment (3)

- Pure delay system $\Rightarrow$

Grade of Service measure = Mean waiting time $E[W]$

- Formula for the mean waiting time $E[W]$ (assuming that ρ <1):

$$E[W] = s \cdot \frac{\rho}{1-\rho}$$

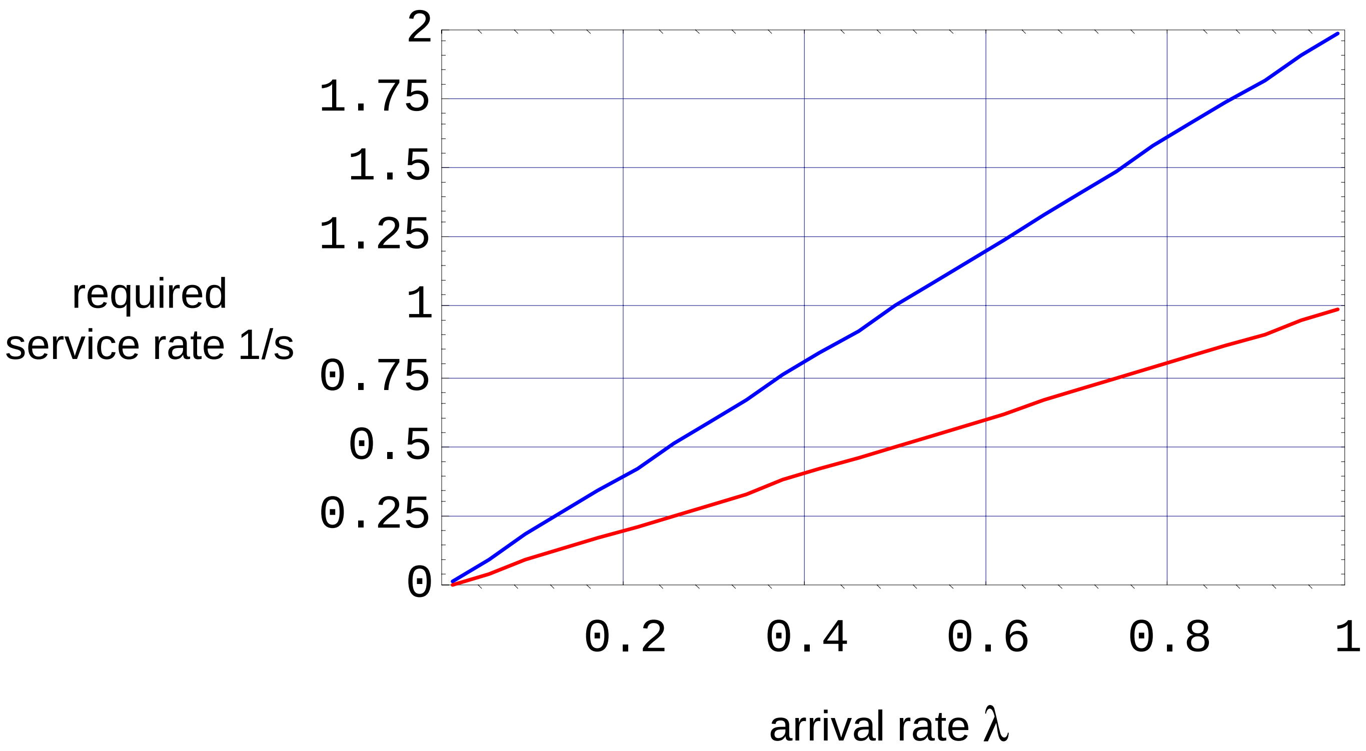
  - ρ = λs
  - **Note**: $E[W]$ grows to infinity as ρ tends to 1

Dimensioning curve

- Grade of Service requirement: $E[W] \leq s$

  $\Rightarrow$ Allowed load $\rho \leq 0.5 = 50\% \Rightarrow \lambda s \leq 0.5$

  $\Rightarrow$ Required service rate $1/s \geq 2\lambda$ (blue line)



Dimensioning rule

- To get the required Grade of Service (the average time a customer waits before service should be less than the average service time) …

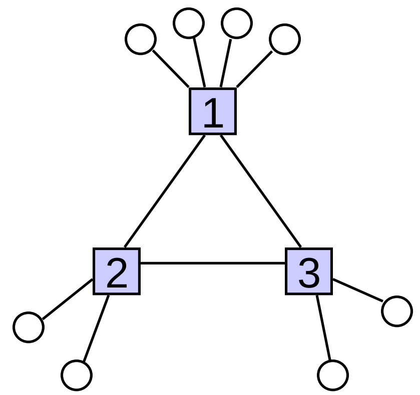
… Keep the traffic load less than 50%

- If you want a less stringent requirement, still remember the **safety margin** …

Don’t let the total traffic load approach to 100%

- Otherwise you’ll see an explosion!

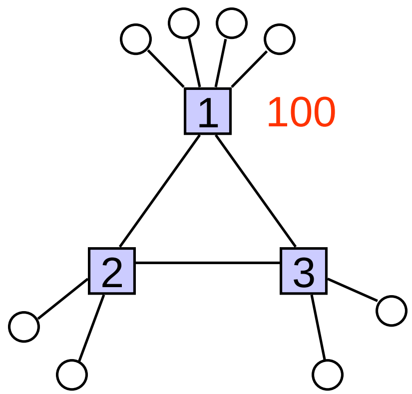
Example: dimensioning the nodes (1)



area	1	2	3	sum
1	60	15	15	90
2	30	30	15	75
3	30	15	30	75
sum	120	60	60	240

- **Assumptions**:
  - 3 local exchanges completely connected to each other
  - Traffic matrix T describing the busy hour traffic interest (in erlangs) given below
  - Fixed (direct) routing: calls are routed along shortest paths.
  - Mean holding time $h = 3$ min.
- **Task**:
  - Determine the call handling capacity needed in different network nodes according to the GoS requirement $\rho < 50\%$

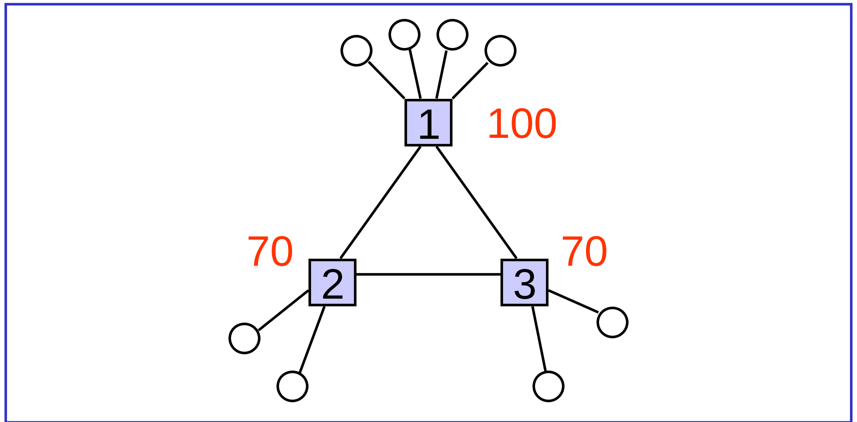
Example: dimensioning the nodes (2)



area	1	2	3	sum
1	60	15	15	90
2	30	30	15	75
3	30	15	30	75
sum	120	60	60	240

- **Node 1**:
  - call requests from own area: $[T(1,1) + T(1,2) + T(1,3)]/h = 90/3 = 30$ calls/min
  - call requests from area 2: $T(2,1)/h = 30/3 = 10$ calls/min
  - call requests from area 3: $T(3,1)/h = 30/3 = 10$ calls/min
  - total call request arrival rate: $\lambda(1) = 30+10+10 = 50$ calls/min
  - required call handling capacity: $\rho(1) = \lambda(1)/\mu(1) = 0.5 \Rightarrow$ $\boldsymbol{\mu(1)} \geq 2*\lambda(1) =$ **100 calls/min**

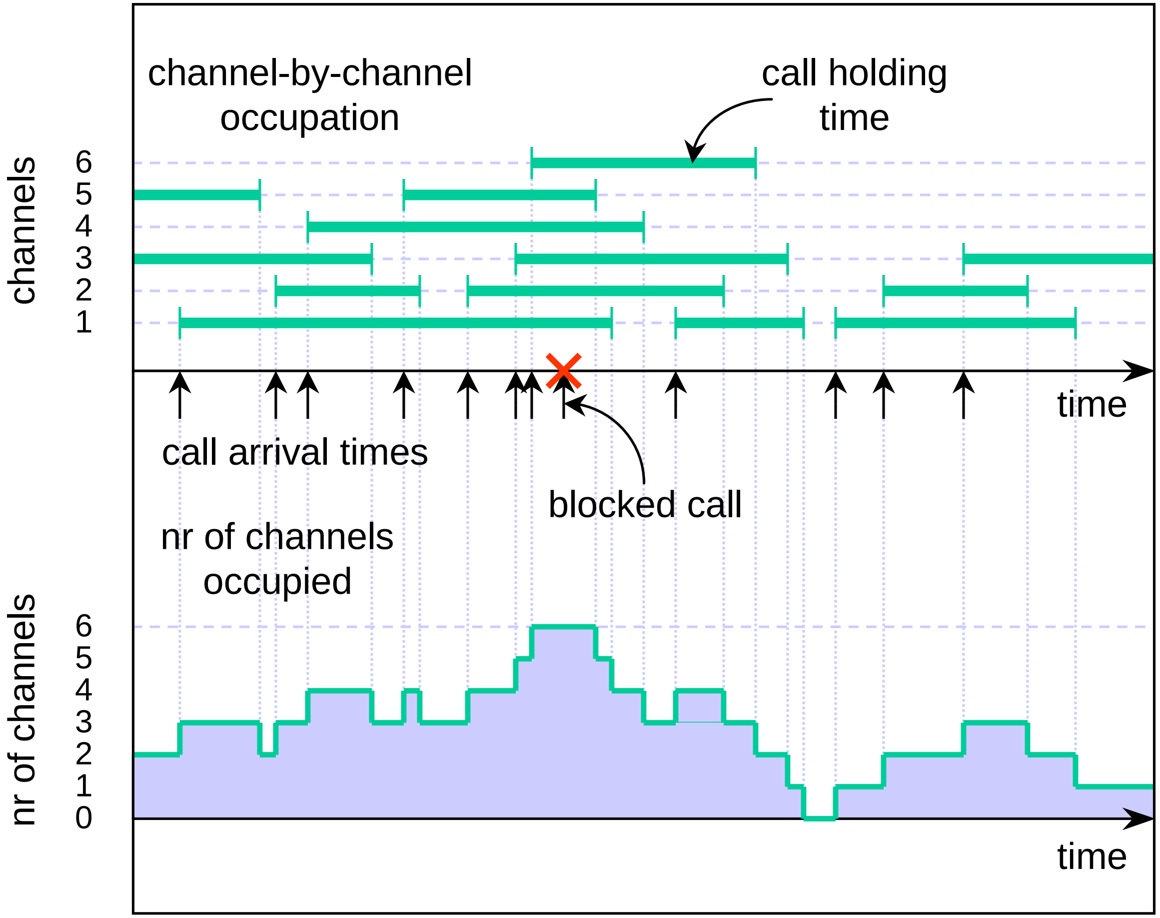
Example: dimensioning the nodes (3)



area	1	2	3	sum
1	60	15	15	90
2	30	30	15	75
3	30	15	30	75
sum	120	60	60	240

- **Node 2**:
  - total call request arrival rate:
    $\lambda(2) = [T(2,1) + T(2,2) + T(2,3) + T(1,2) + T(3,2)]/h$
    $= (75+15+15)/3 = 35$ calls/min
  - required call handling capacity:
    $\mu(2) \geq 2*\lambda(2) =$ **70 calls/min**
- **Node 3**:
  - total call request arrival rate :
    $\lambda(3) = [T(3,1) + T(3,2) + T(3,3) + T(1,3) + T(2,3)]/h$
    $= (75+15+15)/3 = 35$ calls/min
  - required call handling capacity:
    $\mu(3) \geq 2*\lambda(3) =$ **70 calls/min**

Traffic process during information transfer (1)



Traffic process during information transfer (2)

- Call arrival process has already been modelled as
  - a Poisson process with intensity λ
- Further we assume that call holding times are
  - IID and generally distributed with mean h
    - typically h is in the range of **minutes** (not milliseconds as s)
    - h is more a **traffic parameter** than a system parameter
- The resulting traffic process model is
  - the **M/G/n/n loss model** with (offered) traffic intensity $a = \lambda h$

Traffic process during information transfer (3)

- Pure loss system ⇒

Grade of Service measure = Call blocking probability B

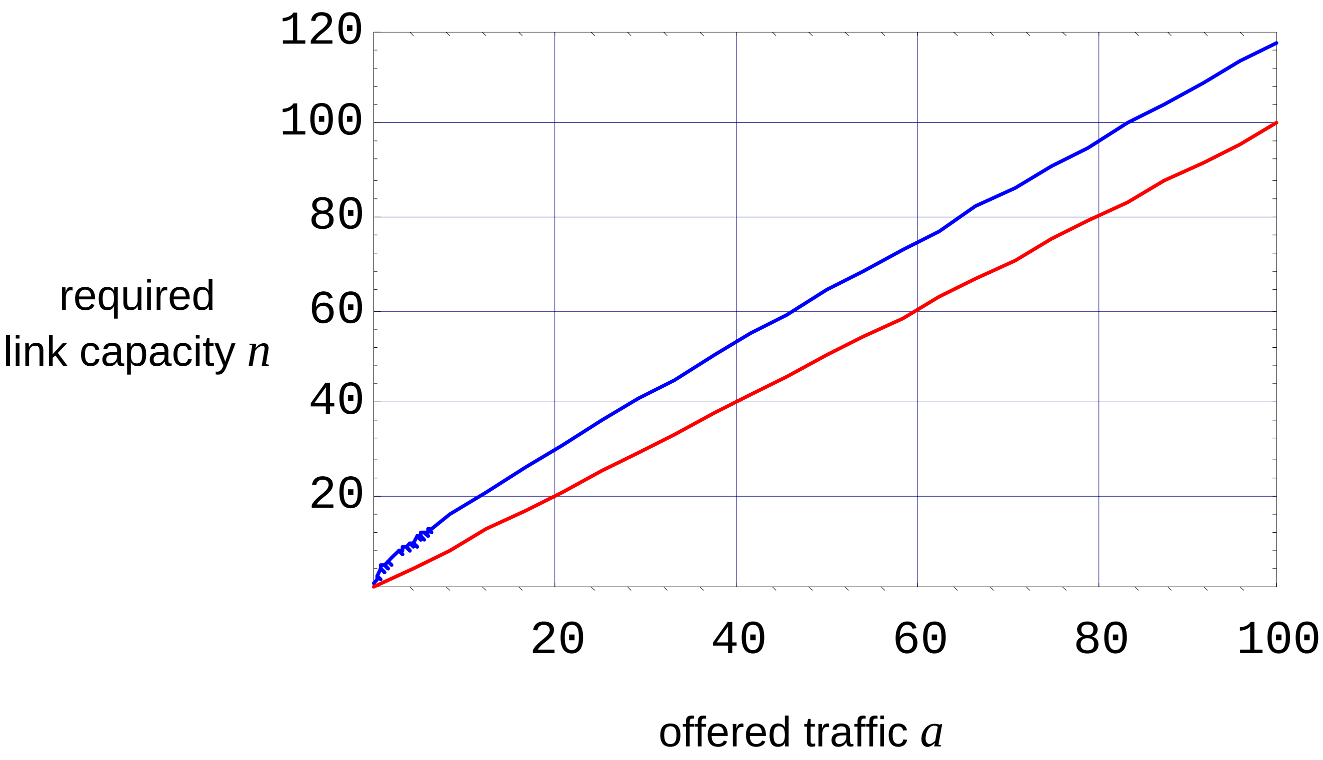
- **Erlang's blocking formula**:

$$B = \mathrm{Erl}(n, a) = \frac{\frac{a^n}{n!}}{\sum_{i=0}^{n} \frac{a^i}{i!}}$$

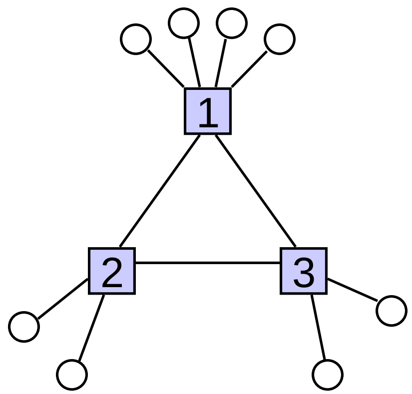
  - $a = \lambda h$
  - $n! = n(n-1)(n-2)\ldots 1$

Dimensioning curve

- Grade of Service requirement: $B \leq 1\%$
  $\Rightarrow$ Required link capacity: $n = \min\{i = 1,2,\ldots \mid \mathrm{Erl}(i,a) \leq B\}$



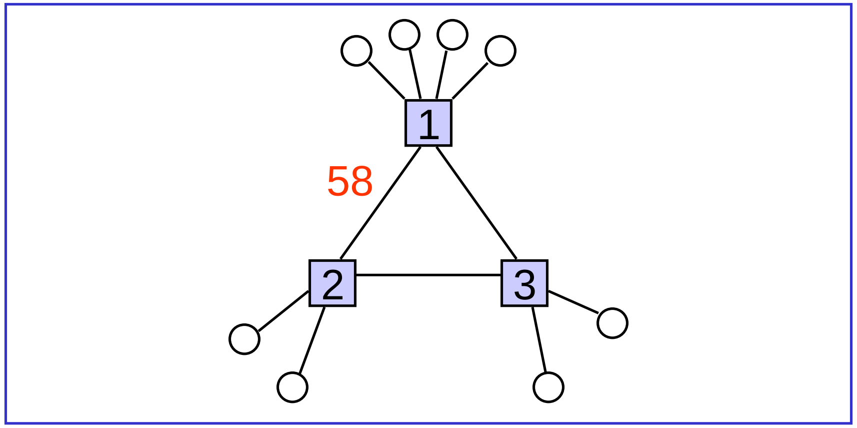
Example: dimensioning the links (1)



area	1	2	3	sum
1	60	15	15	90
2	30	30	15	75
3	30	15	30	75
sum	120	60	60	240

- **Assumptions**:
  - 3 local exchanges completely connected to each other with two-way links
  - Traffic matrix T describing the busy hour traffic interest (in erlangs) given below
  - Fixed (direct) routing: calls are routed along shortest paths.
  - Mean holding time h = 3 min.
- **Task**:
  - Dimension trunk network links according to the GoS requirement $B \leq 1\%$

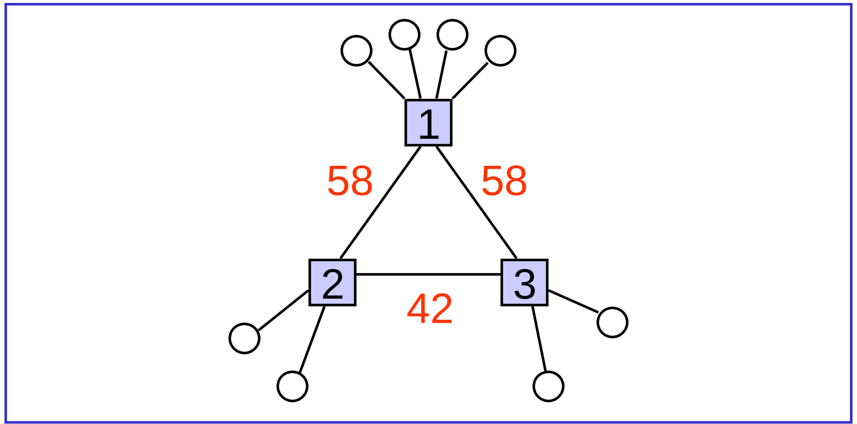
Example: dimensioning the links (2)



area	1	2	3	sum
1	60	15	15	90
2	30	30	15	75
3	30	15	30	75
sum	120	60	60	240

- **Link 1-2** (betw. nodes 1 and 2):
  - offered traffic 1 $\rightarrow$ 2:
    $a(1,2) = T(1,2) = 15$ erlang
  - offered traffic 2 $\rightarrow$ 1:
    $a(2,1) = T(2,1) = 30$ erlang
  - total offered traffic:
    $a(1–2) = T(1,2) + T(2,1)$
    $= 15+30 = 45$ erlang
  - required capacity:
    $n(1\text{-}2) \geq \min\{i \mid \text{Erl}(i,45) \leq 1\%\}$
    $\Rightarrow$ **$n(1\text{-}2) \geq 58$ channels**

Example: dimensioning the links (3)



area	1	2	3	sum
1	60	15	15	90
2	30	30	15	75
3	30	15	30	75
sum	120	60	60	240

- **Link 1-3**:
  - total offered traffic:
    $a(1\text{-}3) = T(1,3) + T(3,1)$
    $= 15+30 = 45$ erlang
  - required capacity:
    $n(1\text{-}3) \geq \min\{i \mid \mathrm{Erl}(i,45) \leq 1\%\}$
    $\Rightarrow$ ***n*(1-3) ≥ 58 channels**
- **Link 2-3**:
  - total offered traffic:
    $a(2\text{-}3) = T(2,3) + T(3,2)$
    $= 15+15 = 30$ erlang
  - required capacity:
    $n(2\text{-}3) \geq \min\{i \mid \mathrm{Erl}(i,30) \leq 1\%\}$
    $\Rightarrow$ ***n*(2-3) ≥ 42 channels**

Table: $B = \mathrm{Erl}(n,a)$

- $B = 1\%$

n:	a:
35 channels	24.64 erlang
36 channels	25.51 erlang
37 channels	26.38 erlang
38 channels	27.26 erlang
39 channels	28.13 erlang
40 channels	29.01 erlang
41 channels	29.89 erlang
42 channels	30.78 erlang
43 channels	31.66 erlang
44 channels	32.55 erlang
45 channels	33.44 erlang

$B = 1\%$

n:	a:
50 channels	37.91 erlang
51 channels	38.81 erlang
52 channels	39.71 erlang
53 channels	40.61 erlang
54 channels	41.51 erlang
55 channels	42.41 erlang
56 channels	43.32 erlang
57 channels	44.23 erlang
58 channels	45.13 erlang
59 channels	46.04 erlang
60 channels	46.95 erlang

End-to-end blocking probability

- Thus far we have concentrated on the single link case, when calculating the call blocking probability B_c
- However, there can be many (trunk network) links along the route of a (long distance) call. In this case it is more interesting to calculate the total **end-to-end blocking probability** B_e experienced by the call. A method (called **Product Bound**) to calculate B_e is given below.
- Consider a call traversing through links $j = 1, 2, \ldots, J$. Denote by $B_c(j)$ the blocking probability experienced by the call in each single link j. Then

$$B_e = 1 - (1 - B_c(1)) * (1 - B_c(2)) * \ldots * (1 - B_c(J))$$

$$B_c(j)\text{'s small} \Rightarrow B_e \approx B_c(1) + B_c(2)) + \ldots + B_c(J)$$

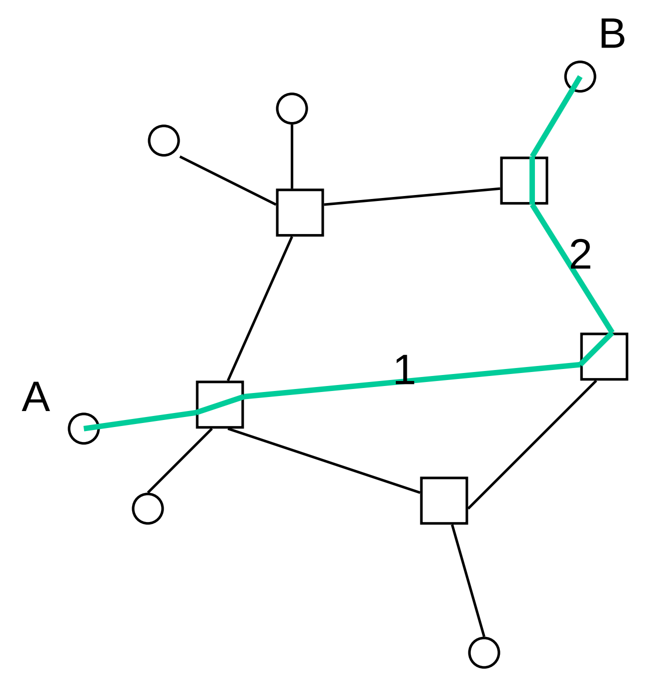
Example

- The call from A to B is traversing through trunk network links 1 and 2
- Let $B_c(1)$ and $B_c(2)$ denote the call blocking probability in these links
- Product Bound (PB):

$$B_e = 1 - (1 - B_c(1))(1 - B_c(2))$$
$$= B_c(1) + B_c(2) - B_c(1)\, B_c(2)$$

- Approximately:

$$B_e \approx B_c(1) + B_c(2)$$



Literature

1 A. Olsson, ed. (1997)

- “Understanding Telecommunications 1”
- Studentlitteratur, Lund, Sweden

2 A. Girard (1990)

- “Routing and Dimensioning in Circuit-Switched Networks”
- Addison-Wesley, Reading, MA