COMPUTER NETWORK PERFORMANCE

Gaia Maselli
maselli@di.uniroma1.it

SAPIENZA
UNIVERSITÀ DI ROMA

Performance evaluation: why?

Commercially

- What is the reach of the Internet?
- How many individuals are connected in a given area?
- What fraction of users have high speed Internet connection? And how many depends on dial up?
- Where should network access points be placed?
- Will users with wireless connectivity be able to access the Internet?

Socially

- Social implications of Internet use
- Understanding the amount of network activity involving various sites and protocols give considerable insight into social issues
- To characterize web site popularity and content
- Emerging protocols

Technically

- What is the delay (or latency) for a packet to traverse the networks?
- The design of network components (e.g. routers) and protocols is strongly driven by the nature of Internet workloads
  - Router designs strongly depends on the **statistical properties** of **network traffic** and **packet size distribution**
  - The statistical properties **of Web pages** influence the performance and design of Web servers and browsers
- The **popularity** of new applications can drive improvements to associated protocols (as in the case of the explosion of Web traffic, which motivated the improvement of the basic HTTP/1.0 protocol to yield HTTP/1.1)

But also..

1. I have an idea for a new network paradigm and want to test it
2. I have an idea for a new MAC (or any other) protocol, would it be better than the existing ones?

- ***If you are doing any research, whatever you are doing, you need to***
  - ***analyze that***
  - ***compare different alternatives to find the best one***
- Performance is a key criterion in the design, procurement, and use of computer networks
- Get the highest performance at a given (or the lowest) cost

Performance evaluation: goals

Two main goals

1. Assess the performance of an existing network
2. Study a new system before implementing it

Performance evaluation is required at every stage in the life cycle of a system

- Find the best design
- Even if there are no alternatives, evaluating the current system helps in determining how well it is performing and whether any improvements need to be made

How do we evaluate performance?

- Measurements
  - Real network
  - testbed
- Analytical evaluation
  - Queueing networks
  - Probabilistic models
- Simulation
  - Network Simulator (NS3)

How do we choose one of them?

Criteria for Selecting an Evaluation Technique

Criterion	Analytical Modeling	Simulation	Measurement
1. Stage	Any	Any	Postprototype
2. Time required	Small	Medium	Varies
3. Tools	Analysts	Computer languages	Instrumentation
4. Accuracy[a]	Low	Moderate	Varies
5. Trade-off evaluation	Easy	Moderate	Difficult
6. Cost	Small	Medium	High
7. Saleability	Low	Medium	High

[a] In all cases, result may be misleading or wrong.

Criterion 1: Stage

- The key consideration in deciding the evaluation technique is the *life-cycle stage* in which the system is

- New system ➔ *analytical modeling* and *simulation* are the only techniques from which to choose

- Prototype or Improved system

  ➔ *measurement* (but also modeling and simulation)

Modeling and simulation can be done anytime, measurement requires a prototype

Criteria for Selecting an Evaluation Technique

Criterion	Analytical Modeling	Simulation	Measurement
1. Stage	Any	Any	Postprototype
2. Time required	Small	Medium	Varies
3. Tools	Analysts	Computer languages	Instrumentation
4. Accuracy[a]	Low	Moderate	Varies
5. Trade-off evaluation	Easy	Moderate	Difficult
6. Cost	Small	Medium	High
7. Saleability	Low	Medium	High

[a] In all cases, result may be misleading or wrong.

Criterion 2: Time required

- Time available for evaluation has to be taken into account
- Often results are required *yesterday*

Short ➔ *analytical modeling*

Medium ➔ *simulation*

Long ➔ *measurement*

Murphy's law strikes measurements more often than other techniques

↓

Variable time

Criteria for Selecting an Evaluation Technique

Criterion	Analytical Modeling	Simulation	Measurement
1. Stage	Any	Any	Postprototype
2. Time required	Small	Medium	Varies
3. Tools	Analysts	Computer languages	Instrumentation
4. Accuracy[a]	Low	Moderate	Varies
5. Trade-off evaluation	Easy	Moderate	Difficult
6. Cost	Small	Medium	High
7. Saleability	Low	Medium	High

[a] In all cases, result may be misleading or wrong.

Criterion 3: Tools

- Availability of tools plays an important role

Modeling skills

Simulation languages

Measurement instruments

Criteria for Selecting an Evaluation Technique

Criterion	Analytical Modeling	Simulation	Measurement
1. Stage	Any	Any	Postprototype
2. Time required	Small	Medium	Varies
3. Tools	Analysts	Computer languages	Instrumentation
4. Accuracy[a]	Low	Moderate	Varies
5. Trade-off evaluation	Easy	Moderate	Difficult
6. Cost	Small	Medium	High
7. Saleability	Low	Medium	High

[a] In all cases, result may be misleading or wrong.

Criterion 4: Accuracy

- Level of accuracy desired is another important consideration

- Low ➔ *analytical modeling* may require so many simplifications and assumptions that results may be too approximate
- Moderate ➔ *simulations can* incorporate more details and and require less assumptions and thus are more closer to reality
- Variable ➔ *measurement* may not give accurate results simply because many of the environmental parameters, such as system configuration, type of workload, and time of measurement may be unique to the experiment

Criteria for Selecting an Evaluation Technique

Criterion	Analytical Modeling	Simulation	Measurement
1. Stage	Any	Any	Postprototype
2. Time required	Small	Medium	Varies
3. Tools	Analysts	Computer languages	Instrumentation
4. Accuracy[a]	Low	Moderate	Varies
5. Trade-off evaluation	Easy	Moderate	Difficult
6. Cost	Small	Medium	High
7. Saleability	Low	Medium	High

[a] In all cases, result may be misleading or wrong.

Criterion 5: Trade-off evaluation

- The goal of every performance study is to compare different alternatives or to find the optimal parameter values

Easy	➔	*analytical modeling*
Moderate	➔	*simulation*
Difficult	➔	*measurement*

Criteria for Selecting an Evaluation Technique

Criterion	Analytical Modeling	Simulation	Measurement
1. Stage	Any	Any	Postprototype
2. Time required	Small	Medium	Varies
3. Tools	Analysts	Computer languages	Instrumentation
4. Accuracy[a]	Low	Moderate	Varies
5. Trade-off evaluation	Easy	Moderate	Difficult
6. Cost	Small	Medium	High
7. Saleability	Low	Medium	High

[a] In all cases, result may be misleading or wrong.

Criterion 6: Cost

- Cost allocated for the project is also important
- Small ➔ *analytical modeling* requires only paper and pencil (in addition to the analyst's time)
- Medium ➔ *simulation* requires a simulator (often free) and some time
- High ➔ *measurement* requires real equipment, instruments and time

Criteria for Selecting an Evaluation Technique

Criterion	Analytical Modeling	Simulation	Measurement
1. Stage	Any	Any	Postprototype
2. Time required	Small	Medium	Varies
3. Tools	Analysts	Computer languages	Instrumentation
4. Accuracy[a]	Low	Moderate	Varies
5. Trade-off evaluation	Easy	Moderate	Difficult
6. Cost	Small	Medium	High
7. Saleability	Low	Medium	High

[a] In all cases, result may be misleading or wrong.

Criterion 7: Saleability

- Saleability is the key justification when considering the expense and labor of measurements.

- Low ➔ *analytical modeling* - some people are skeptical of analytical results simply because they do not understand the technique or the final results
- Medium ➔ *simulation*
- High ➔ *measurement* - It is easy to convince others if it is a real measurement

Which technique do we choose?

Three rules of validation

- Do not trust the results of a **simulation model** until they have been validated by analytical modeling or measurements.
- Do not trust the results of an **analytical model** until they have been validated by a simulation model or measurements.
- Do not trust the results of a **measurement** until they have been validated by simulation or analytical modeling.

The art of performance evaluation

- Like a work of art, successful evaluation **cannot be produced mechanically**
- Every evaluation requires an intimate knowledge of the system being modeled and a **careful selection of the methodology, workload, and tools**
- Like an artist, each analyst has a unique style.
  - Given the same problem, two analysts may choose different performance metrics and evaluation methodologies
  - Given the same data, two analysts may interpret them differently.

Basic terms

- **System**: anything that you want to study (any collection of hardware, software, and firmware components)
- **Metrics**: criteria used to evaluate the performance of a system (e.g., response time)
- **Workloads**: the requests made by the users of the system (anything that goes in the system).

A systematic approach to performance evaluation

1. **State Goals and Define the System**
   - State del goals of the study and define what constitutes the system by delineating system boundaries (given the same set of hardware and software, the definition of the system may vary depending upon the goals of the study)
2. **List Services and Outcomes**
   - Each system provides a set of services.
   - A computer network allows its users to send packets to specified destination
   - When a user requests a service, there are a number of possible outcomes.
     - Desirable (the packet arrives to the destination)
     - Undesirable (the packet gets lost)

A systematic approach to performance evaluation

3. **Select Metrics**
   - Select criteria to compare performance
   - Metrics are related to speed, accuracy, and availability of services
4. **List Parameters**
   - Make a list of all the parameters that affect the system (anything that you can change)
     - System parameters (do not vary among various installation of the system)
     - Workload parameters (depend on users' requests and vary from one installation to the next)
5. **Select Factors to Study**
   - Select the parameters that you want to study (they will change during the evaluation)
   - If you ignore significant factors you get meaningless results

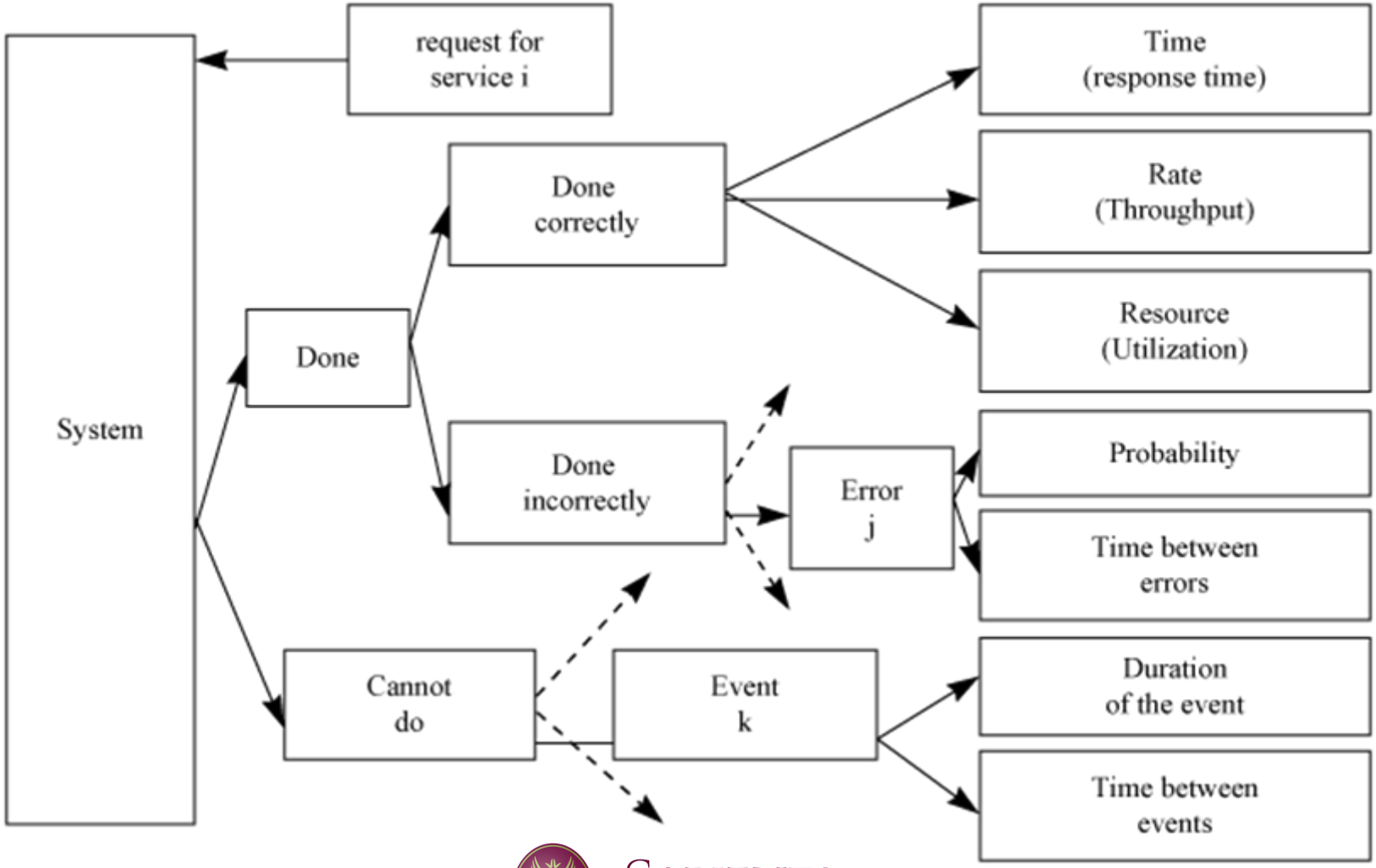
A systematic approach to performance evaluation

6. **Select Evaluation Technique**
   - Analytical modeling, simulation, or measuring?
   - The choice depends upon the time and resources available
7. **Select Workload**
   - list of service requests to the system
     - Probability of various requests
     - Trace of requests measured on a real system
8. **Design Experiments**
   - Decide a sequence of experiments that maximize information with minimal effort
9. **Analyze and Interpret Data**
   - Analysis only produces results (big amount of results) and not conclusions
   - Draw conclusions from results (interpretation of data)
10. **Present Results**
    - Make graphs (results must be easy to understood)

Repeat (go back and reconsider some decisions)

Selecting performance metrics

Selecting performance metrics



Selecting performance metrics

- For each service request the system may **perform the service correctly, incorrectly, or refuse** to perform the service

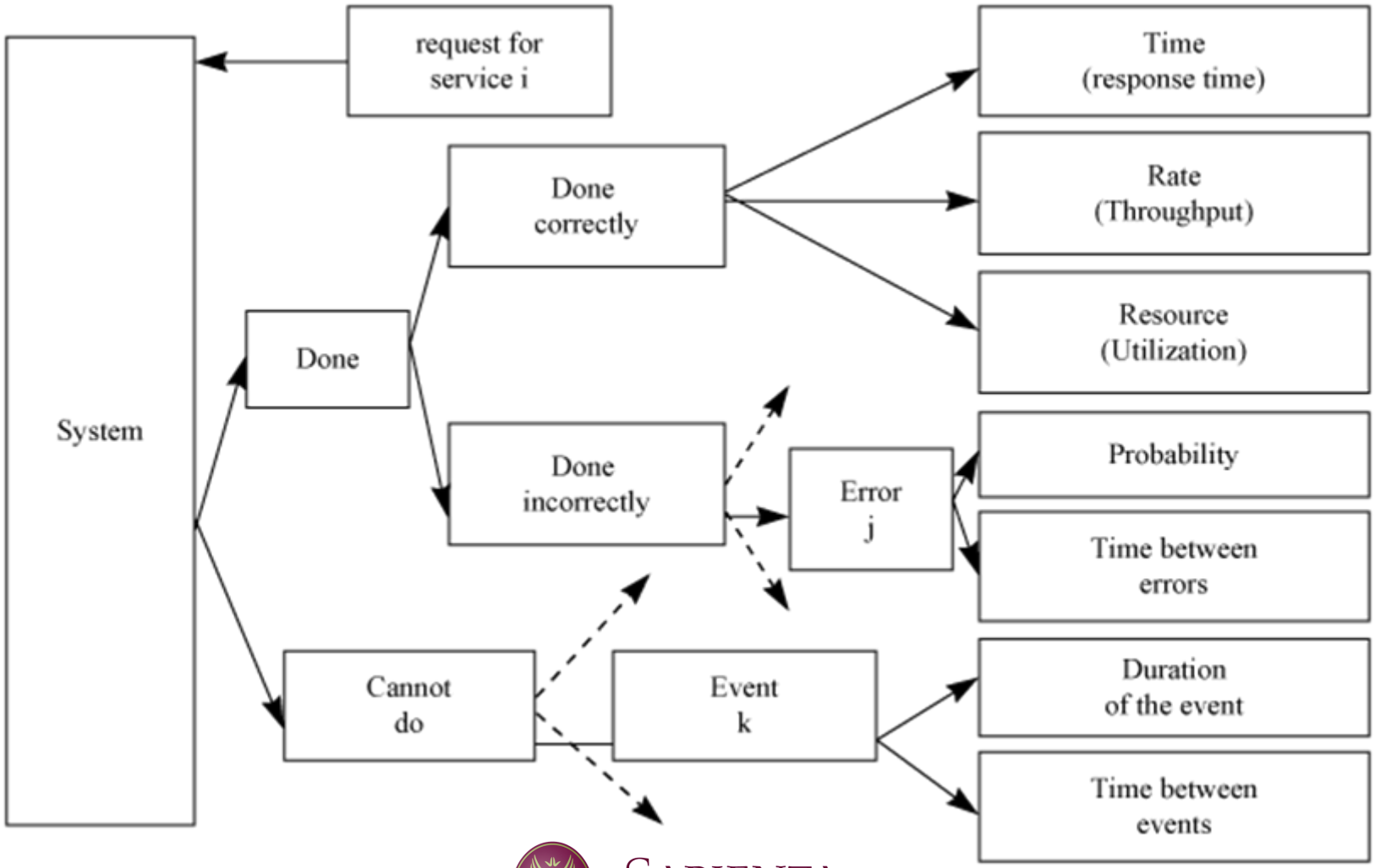
Example: a gateway in a network offers the service of forwarding packets to the specified destination. When presented a packet, it may forward the packet correctly, it may forward it to the wrong destination, or it may be down (not forward it at all)

- If the systems performs the service correctly, its performance is measured by the **time taken** to perform the service**, the rate at which** the service is performed, and **the resource consumed** while performing the service
  - **Time ➔ responsiveness**
  - **Rate ➔ productivity**
  - **Resource ➔ utilization**

Example (gateway):

- Responsiveness is the time interval between arrival of a packet and its successful delivery
- Productivity is the number of packets forwarded per unit of time
- Utilization is percentage of time gateway resources are busy for the given load level

Selecting performance metrics



Selecting performance metrics

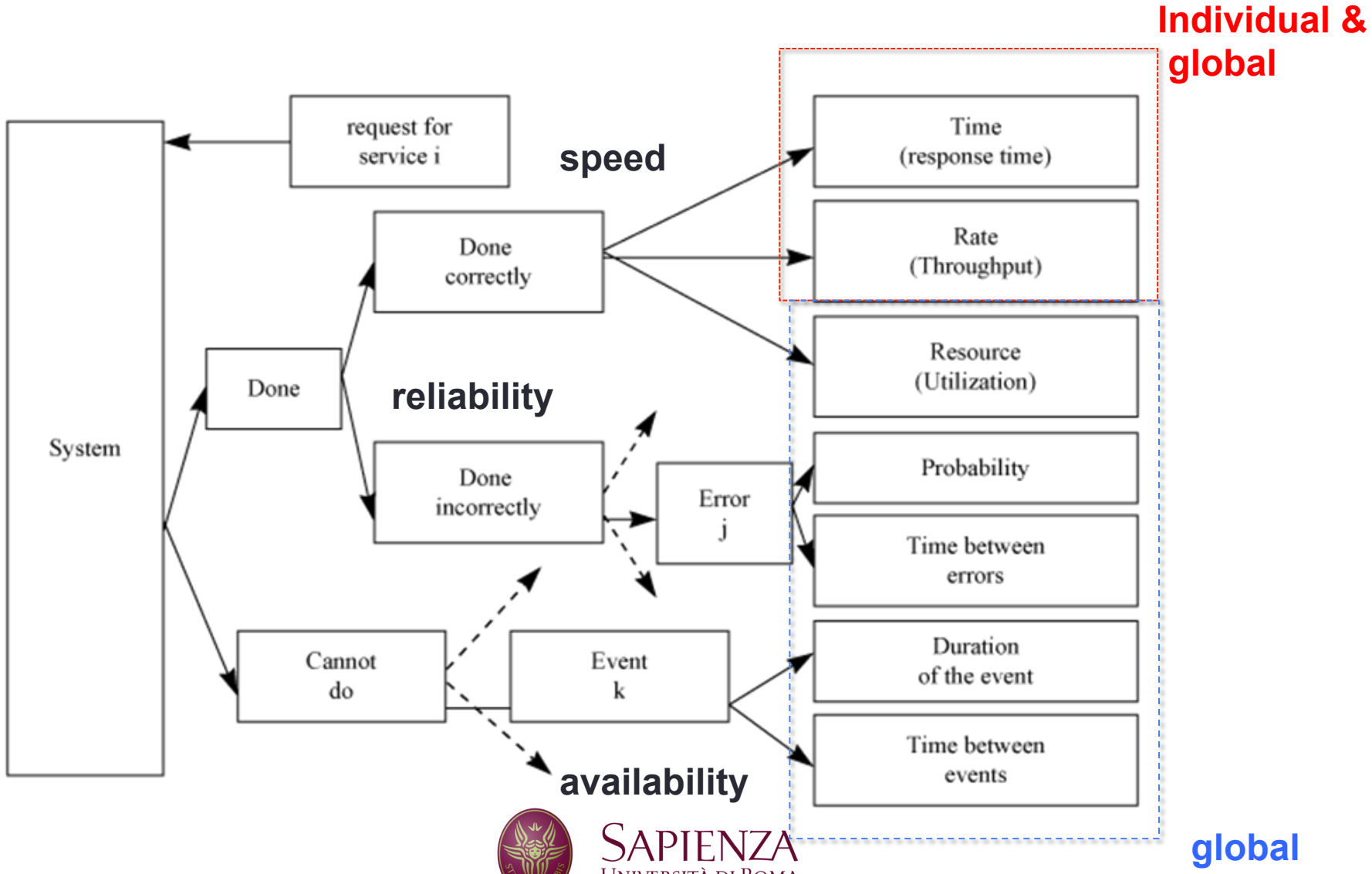
- If the system performs the service incorrectly, an error is occurred
- It is helpful to classify errors and to determine the probabilities of each class of errors.

Example (gateway): we may want to find the probability of single-bit errors and packet error

- If the system does not perform the service, it is said to be down, failed or unavailable
- It is helpful to classify the failure modes and to determine the probability of each class

Example (gateway): the gateway may be unavailable 0.01% of the time due to processor failure and 0.03% due to software failure

Selecting performance metrics



Selecting performance metrics

- Names of the metrics associated with the three outcomes
  - **successful service ➔ speed**
  - **Error ➔ reliability**
  - **Unavailability ➔ availability**
- For each service offered by the system, one would have a number of speed metrics, a number of reliability metrics, and a number of availability metrics
- Most systems offer more than one service, and thus the number of metrics grows proportionally
- As a network is shared by multiple users, two types of performance metrics need to be considered**: individual** and **global**
  - **Individua**l metrics reflect the **utility of each user**
  - **Globa**l metrics reflect the **system wide utility**
  - Some metrics are **individual and global**

N.B. there are cases when the decision that optimize individual metrics is different from the one that optimizes the system metric (e.g., throughput !!!)

Selecting performance metrics

- Given a number of metrics, use the following considerations to select a subset:
  - **Low variability**
    - helps reducing the number of repetitions
    - Metrics that are ratio of two variables generally have larger variability than either of the two variable and should be avoided
  - **Non redundancy**
    - If two metrics give essentially the same information, it is less confusing to study only one
  - **Completeness**
    - All possible outcomes should be reflected in the set of performance metrics

Selecting performance metrics: summary

- Include (by applying the procedure for metric selection):
  - Time, Rate, Resource
  - Error rate, probability
  - Time to failure and duration
- Consider including:
  - Mean and variance
  - Individual and global
- Selection criteria:
  - Low-variability
  - Non-redundancy
  - Completeness

Commonly used performance networks

Let us identify *some* common metrics in computer networks

Commonly used performance metrics

Response time: interval between a user's request and the system response

In computer networks:

Packet delay: the delay experienced by a packet as it passes through the network, given by the sum of

1. *Routing delay (processing + queuing)*: time a packet spends inside a router (time between the arrival of the trailing bit at the router and the moment the first bit of the packet is placed on the output link)
2. *Transmission delay*: time required to place a packet onto a link (packetSize/rate)
3. *Propagation delay*: time required for a packet to pass from one end of a link to the other end (distance/propSpeed)

Commonly used performance metrics

Throughput: rate at which the requests are serviced by the system

In computer networks:

Throughput: the rate at which traffic flow through the network.

- In a given interval T, the throughput is calculated as the number of packets the pass at a point *without loss* over time T, and is measured in packets per seconds (**pps**) or bits per second (**bps**)
- Throughput over a sequence of hops is determined by the element with minimum available capacity
- The maximum achievable throughput under ideal workload conditions is called **nominal capacity** of the system and corresponds to the **bandwidth**
- The ratio of maximum achievable throughput (usable capacity) to nominal capacity is called **efficiency**. Example: if the maximum throughput from a 100Mbps LAN is 85 Mbps, its efficiency is 85%

Commonly used performance metrics

Reliability: probability of errors or the mean time between errors

In computer networks:

Packet Loss: fraction of packets lost in a time period

Relative Packet loss rate: if C_n is the number of packets entering a network element in time period n, and L_n is the number of packets lost during that time period, the relative packet loss rate can be estimated as L_n/C_n

Commonly used performance metrics

A metric related to throughput and packet loss

Goodput: the rate at which the application endpoint successfully receives data

- The rate at which TCP sends packets is the load it places on the network per unit of time
- Packets being sent may be retransmissions
- Bytes are not necessarily delivered to the application at the rate that the connection sends them

Utility classification of performance metrics

- Three classes depending upon the utility function of a performance metric

1. **Higher is better** (**HB**): systems users and systems managers prefer higher values of such metrics
   - throughput is an HB metric
2. **Lower is better (LB**): systems users and systems managers prefer lower values of such metrics
   - packet delay is an HB metric
3. **Nominal is best** (**NB**): Both high and low values are undesirable. A particular value in the middle is considered best
   - Utilization