SELECTION OF TECHNIQUES AND METRICS

Gaia Maselli
maselli@di.uniroma1.it

SAPIENZA
UNIVERSITÀ DI ROMA

Overview

- Selection of an evaluation technique
- Selection of performance metrics

Selecting an evaluation technique

- Three techniques for performance evaluation
  1. Analytical modeling
  2. Simulation
  3. Measurement

- 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 value

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

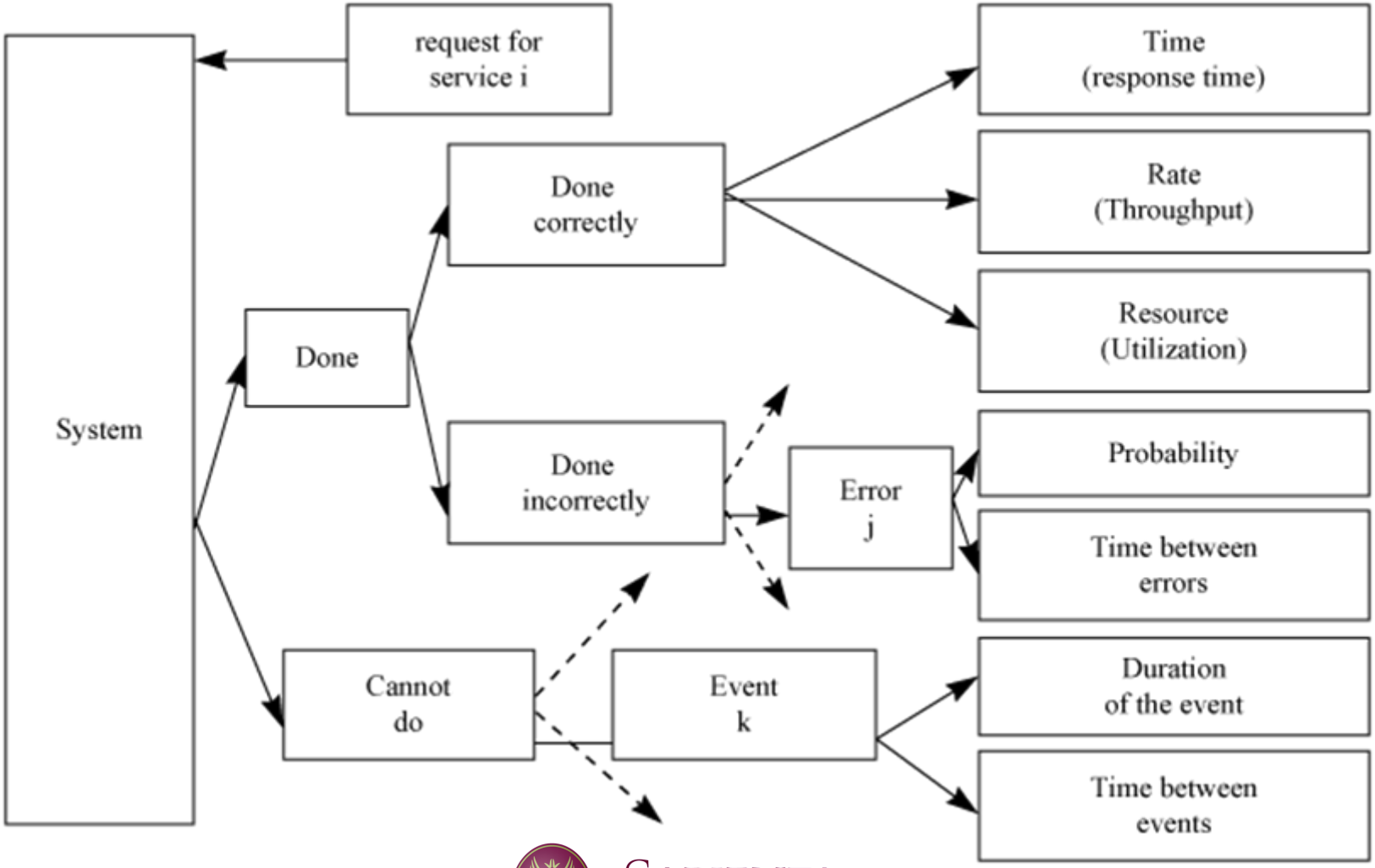
Which technique do we choose?

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.

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

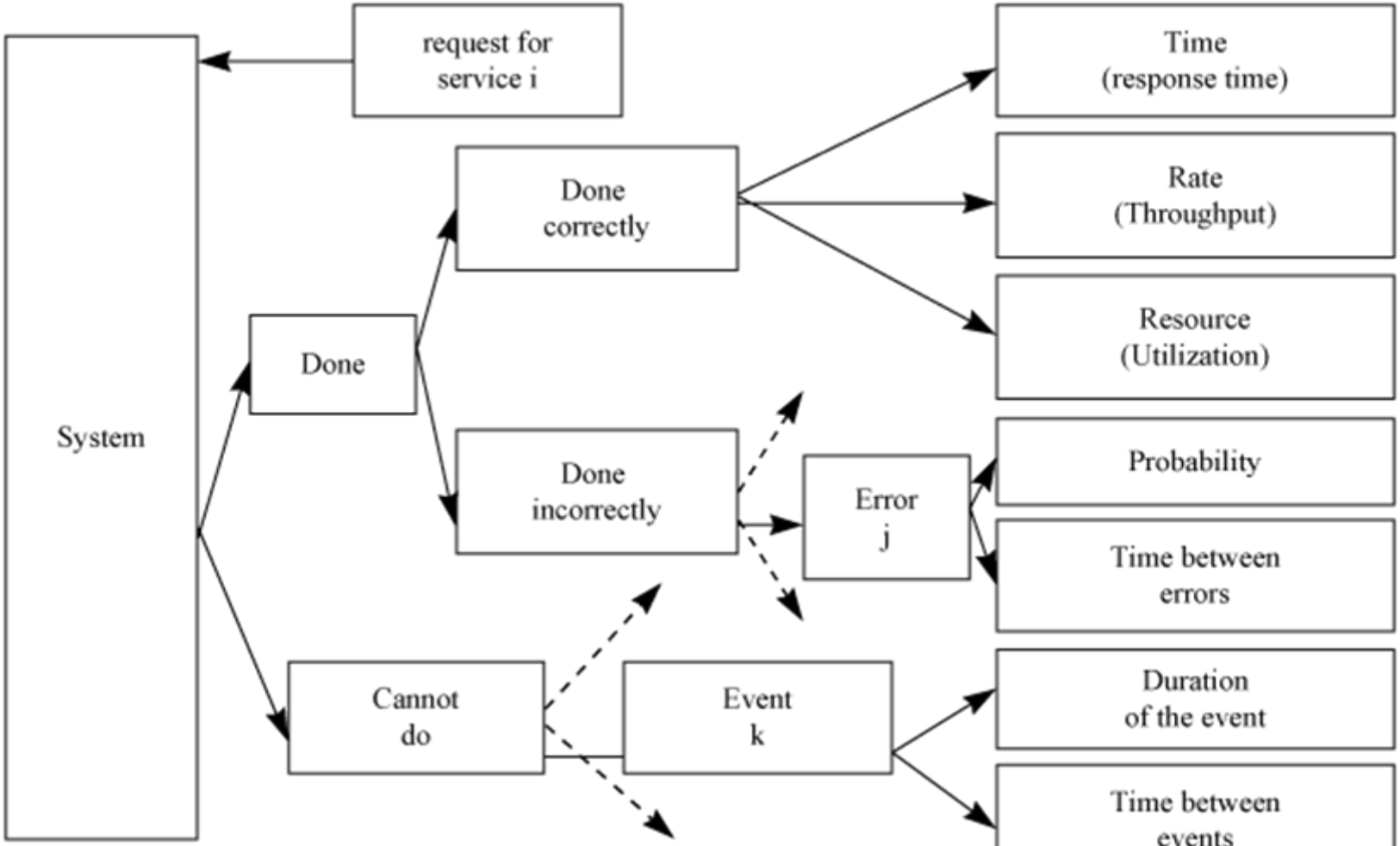
Example: a gateway in a network offers the service of forwarding packets to the specified destination. When presented a packet, it may forward the packet correctly, it may forward it to the wrong destination, or it may be down (not forward it at all)

- 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

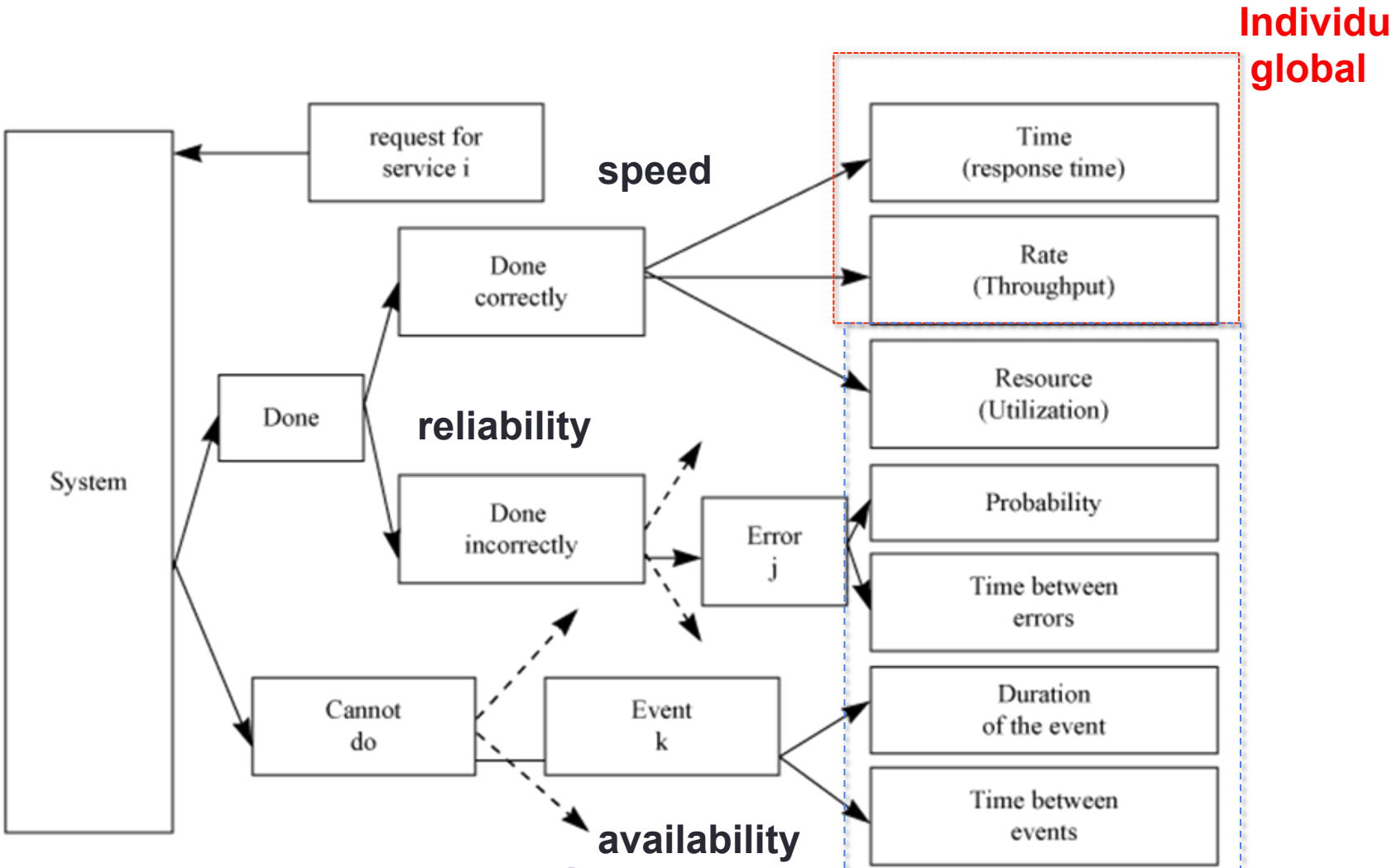
- If the system performs the service incorrectly, an error is occurred
- It is helpful to classify errors and to determine the probabilities of each class of errors.

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 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

Homework

- Revise your homework by applying the procedure for metric selection we have seen today
- Keep note of the changes:
  - Have you added any new metric?
  - Have you found any redundant metric?