

Even More Metrics ...

Vojtěch Horký

Peter Libič

Petr Tůma

2010 – 2021

This work is licensed under a “CC BY-NC-SA 3.0” license. Created to support the Charles University Performance Evaluation lecture. See <http://d3s.mff.cuni.cz/teaching/performance-evaluation> for details.

Contents

1	Network Traffic	1
2	Reliability and Availability	1
3	Computing Infrastructure Capacity	2
4	Information Retrieval Functionality	2
5	More ...	3

1 Network Traffic

Network Metrics

Network Communication

- Throughput – with networks called bandwidth.
- Latency.
- Jitter.
- Packet loss.

Network Services

- Connections per second.
- Transactions per second.
- Maximum concurrent connections.

2 Reliability and Availability

Reliability and Availability Metrics

Reliability

- Probability of errors
- Mean time between errors – *error-free seconds*

Availability

- $availability = \frac{uptime}{totaltime}$
- Mean uptime – *Mean Time To Failure (MTTF)*
- Mean downtime – *Mean Time To Repair (MTTR)*
- MTTF is often better indicator than availability (think short uptimes and extremely short downtimes)

3 Computing Infrastructure Capacity

Cloud Metrics

Scalability

The ability to adjust available resources.

Expressing scalability

- Maximum resource allocation limits
 - Easy in private cloud
 - But how about public cloud ?
- Resource allocation granularity
- Acquisition and release time
 - Technical
 - Accounting

1

Cloud Metrics

Elasticity

The ability to use scalability to address changing resource demands.

Expressing elasticity

- Average or total time to move from under or over provisioned
- Average or total amount of under or over provisioned resources
- In reaction to various workload fluctuation patterns
 - Gradual up or down
 - Regular fluctuations
 - You have been slashdotted !
- Converting to money ?
 - Cost of under vs over provisioning
 - Bunch of estimates
 - Apple 2015 App Store 10-12 hours down estimated 25M USD
 - Amazon 2017 S3 Service 3-4 hours down estimated 150M USD
 - Amazon 2018 Store 12-15 minutes down estimated 3M USD

2

Look at the paper by Herbst et al.: Elasticity in Cloud Computing ... <https://www.usenix.org/conference/icac13/technical-sessions/presentation/herbst>. Examine Figure 2 for intuitive definition of overprovisioning and underprovisioning.

4 Information Retrieval Functionality

Information Retrieval Metrics

Precision

Share of correct results in total results delivered.

$$precision = \frac{truepositive}{truepositive + falsepositive}$$

¹Based on Herbst et al.: Elasticity in Cloud Computing ...

²Based on Herbst et al.: Elasticity in Cloud Computing ...

Recall

Share of correct results delivered in total correct results.

$$recall = \frac{truepositive}{truepositive + falsenegative}$$

Accuracy

Share of correctly classified results.

$$accuracy = \frac{truepositive + truenegative}{all}$$

5 More ...

... Think About More

- Test coverage metrics ?
- Energy efficiency metrics ?
- Interface usability metrics ?
- Artificial intelligence metrics ?
 - Speed of learning
 - Quality of decision
- ...

Look at the blog post by Tatman: Evaluating Text Output in NLP ... <https://towardsdatascience.com/evaluating-text-output-in-nlp-bleu-at-your-own-risk-e8609665a213>.