Federico Gargiulo

Tutor: Pasquale Arpaia – co-Tutor: Nicola Mazzocca

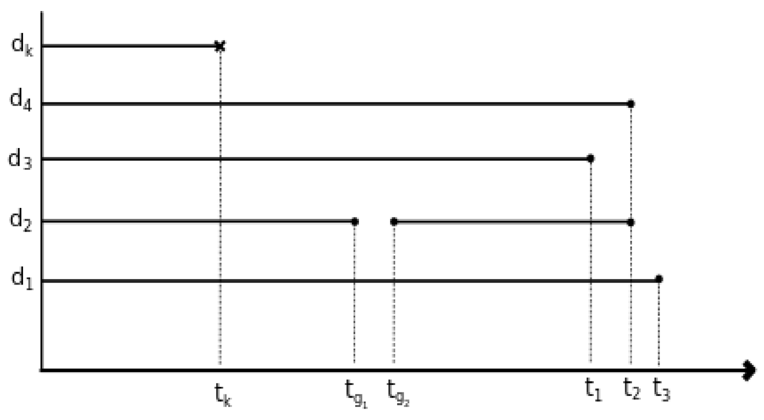
XXIX Cycle - II year presentation

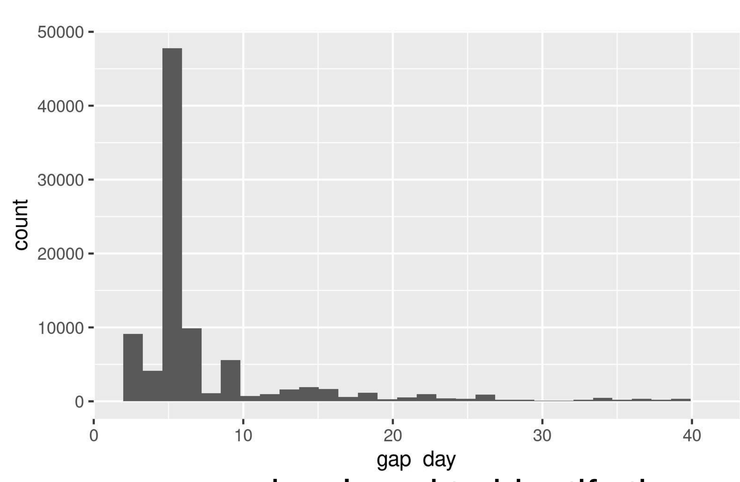
Fault Diagnostics by artificial intelligence with heterogeneous measures

Magnetic hard disks are among the most frequently failing components of cloud storage systems. Disk failures cause about 78% of cloud infrastructure failures. These events lead to several problems. As unscheduled interventions, they often require rapid human intervention, costly consistency checks, potentially validation of recent processing steps by both service providers and users. In the worst case, disk failures can lead to permanent data loss and hence economical damage to users and providers. To overcome the unexpected failure risks, drive manufacturers implement the Self-Monitoring, Analysis and Reporting (SMART) technology which is a software for drive health monitoring. SMART collects and provides a set of health parameters for the manual inspection of the health status and provides a self-test to warn the user when a disk is about to fail.
The SMART test is usually positive when a disk is close to getting failed and can present false positives and false negatives. To overcome the limitations of the SMART test, a data preprocessing approach and supervised machine learning technique have been used to build a classifier with high accuracy and precision for predicting a failure up to 7 days in advance.

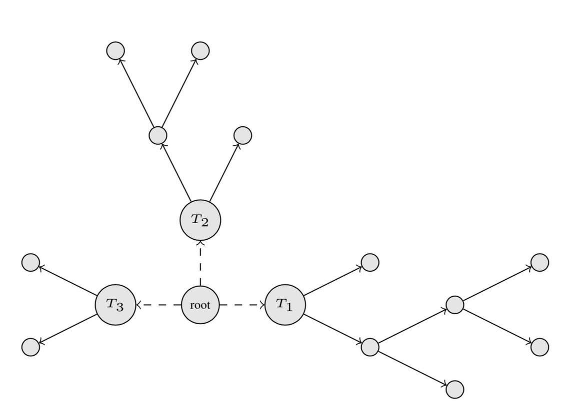
The methodology proposed hereafter has been validated on a case study at CERN in Geneva on a population of 65,000 hard disks. The results obtained are also reported below.

A probe has been develped and put in production for acquiring SMART measures several times per day in all servers or computer nodes.
Given that disk hardware and firmware differs significantly between vendors and even between different models produced by the same vendor, the SMART metrics may differ in calibration, range or even semantics.
A disk can be turned off because of several reasons (failures, retirements/replacing at the end of planned lifetime, space relocation etc.).
All incorrect measures associated with errors, missing values, exceeding values,etc., must be preliminarily removed.





A preliminary study on the gaps between measures was developed to identify the maximum temporary stop period of hard disks.

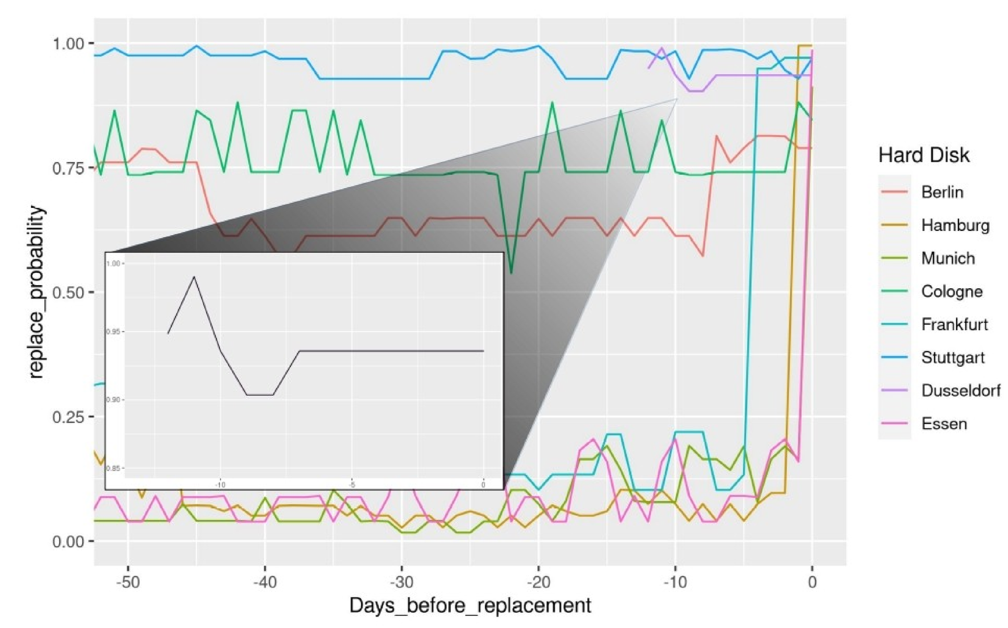


Numerous ML classifier models have been tested for this purpose, the best results have been obtained using the Regularized Greedy Forest. The hyperparameter tuning was done with a Designo Of Experiment. The DOE allowed to search for the best configuration in order to maximize the Positive Likelihood Ratio index.

The actual failure rate is typically around a dozen hard drives per week. Contrary to common machine learning problems, in this context it is crucial to reduce the number of false positives due to the large number of devices. In the following table the results of the five most numerous model are reported.

Model	Recall	False Positive Rate	Positive Likelihood Ratio
Ischia	96.0%	0.7%	130.9
Capri	91.6%	2.2%	41.6
Ventotene	96.4%	0.6%	160.6
Procida	89.3%	1.9%	46.2
Ponza	94.4%	4.1%	22.8

The performance of the model was evaluated on a small fraction of disks not used in the previous phases. It can be noted that some hard disks showed a significant change in the range of a week before the replacement day andother hard disks were to be considered at risk even tens of days before replacement. Moreover an actual case of infantmortality is reported in the insets of the same figures; this drive showed a high probability of required replacement sincetheir first operating day.



Note that models and disk serials have been masked for privacy reasons

- Future Work:
- Generalization of methodologies (several models, typology, models rare used etc)
- Looking for systematic factors and faults in models by vendor, in order to make predictive algorithms highly reusable.
- Analysis of impacts of company fusions, programmed obsolescences, MTTF unrealistic, Enterprises Drives VS. Customers Drives
- Involvement of the Worldwide LHC Computing Grid (WLCG) for validation on a larger sample.

For more information:
https://home.cern/science/computing
https://wlcg.web.cern.ch/

Email: federico.gargiulo@unina.it
Skype: gargiulo.federico