SRE Data Durability Current State and Future Plans
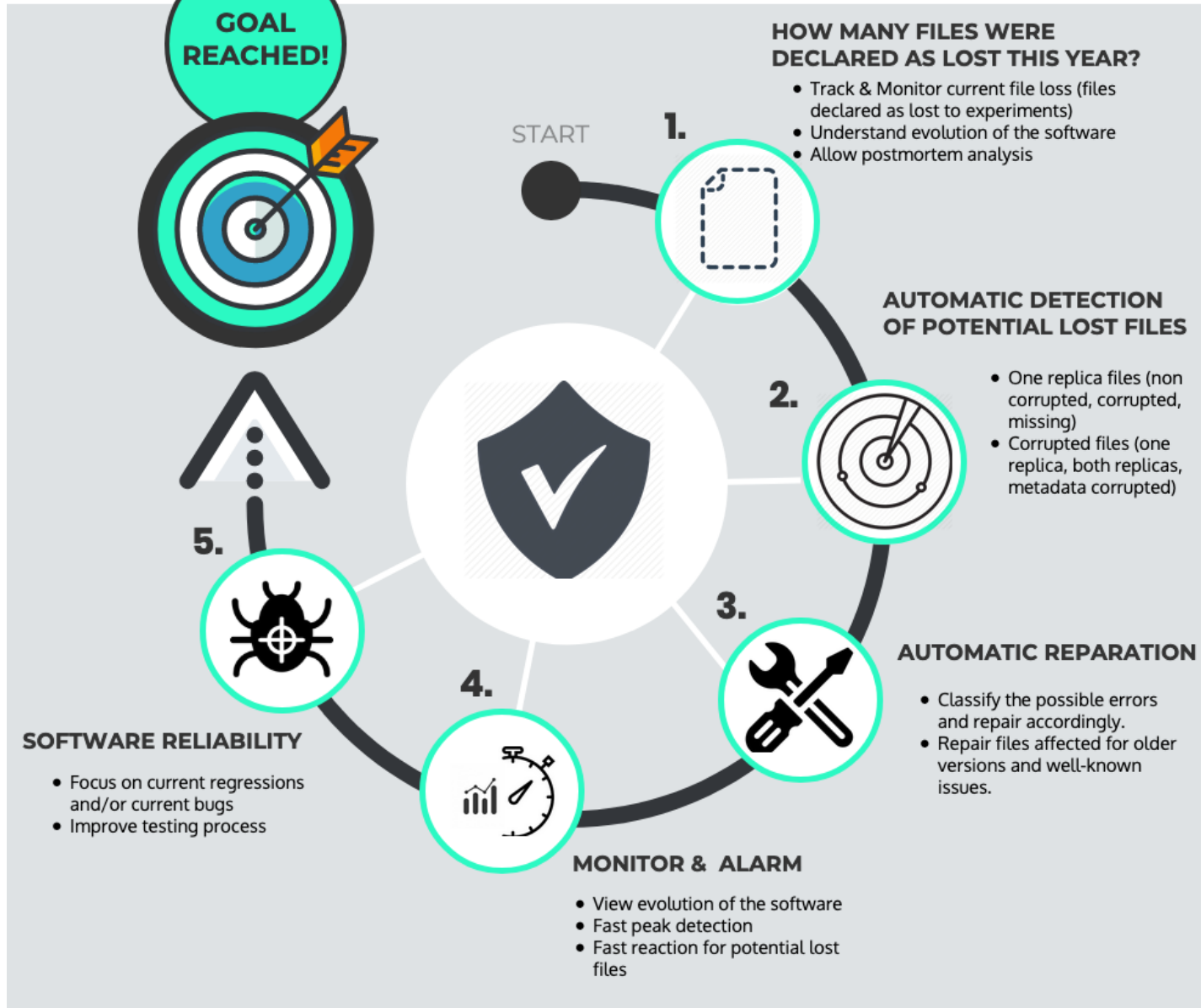
Information Technology Department

Index

- Goal and workflow

- Status of the top three objectives

- Additional improvements

- Next top three objectives

DATA DURABILITY

Data is never lost or compromised
(except double disk failure)

GOAL REACHED!

START

1.

HOW MANY FILES WERE DECLARED AS LOST THIS YEAR?

- Track & Monitor current file loss (files declared as lost to experiments)
- Understand evolution of the software
- Allow postmortem analysis

2.

AUTOMATIC DETECTION OF POTENTIAL LOST FILES

- One replica files (non corrupted, corrupted, missing)
- Corrupted files (one replica, both replicas, metadata corrupted)

3.

AUTOMATIC REPARATION

- Classify the possible errors and repair accordingly.
- Repair files affected for older versions and well-known issues.

4.

MONITOR & ALARM

- View evolution of the software
- Fast peak detection
- Fast reaction for potential lost files

5.

SOFTWARE RELIABILITY

- Focus on current regressions and/or current bugs
- Improve testing process

EOS

CERN

Top three objectives
(from previous meeting)

1. Two people team for investigation (dev + op)
   - Understand no-rep problems in home instances
2. Repair one replica files in Alice
3. Automatic drain repair

Top three objectives
(from previous meeting)

1. Two people team for investigation (dev + op):
   Elvin + Maria
   - Understand no-rep problems in home instances
   - Status: Done the majority of cases
     - Already three main classifications:
       - Checksum 0 with size!=0
         » Client sending the size to namespace, but not writing in fsts
         » Previous version not sending error to the user, now it is fixed
         » Clean up – I requested the feature to Georgios, it has to be done carefully
           • Subset of cases which actually have a copy in restic (one case found yesterday!, so first check in backup and restic!)
       - Checksum 0 with size = 0
         » Similar story, nothing written in fsts
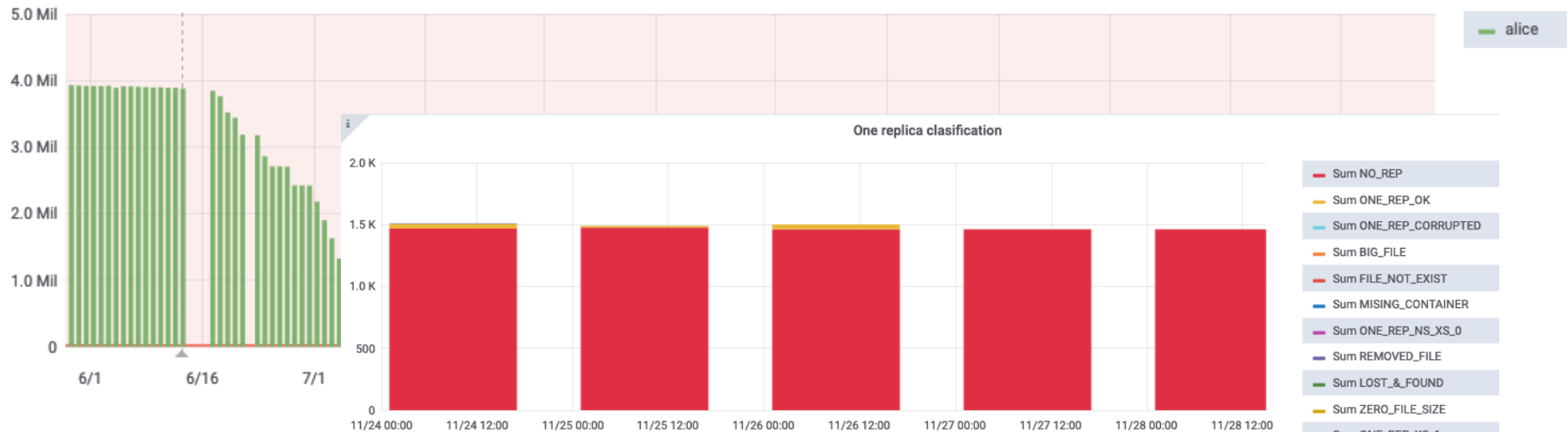   » Cleaning them up
       - One replica layout

Information Technology Department

Top three objectives
(from previous meeting)

2. Repair one replica files in Alice
   – Status: Done
     • Included in the automatic daily reparation
     • Now trying to recover no reps

One Replica Files per Instance

One replica clasification

Information Technology Department

Top three objectives (from previous meeting)

3. Automatic drain repair
   - Status: Done in all instances
   - Cronjob works every day at 20:00
   - Raw information store in cernbox per instance, day and fsid:
     - e.g. /eos/project/e/eos/Durability/Draining/Classification/ALICE/ALICE2019-11-27/9048
   - Integrated in the monitoring infrastructure
     - https://filer-carbon.cern.ch/grafana/d/JzDQWU7Zz/durability-classification?orgId=1&var-instance=All&var-fsid=All

Additional Improvements

- 7 categories repair automatically previously discussed

- Now: 3 Additional ones:
  - Purge aborted files
 - Missing containers + drop cache
  - Drop ghostfiles

Additional Improvements

- Repairing – manually no replica layouts:
  - Implies:
    - getting the parent sys attributes and propagate
    - converting the file to have 2 replicas or the layout from the parents
    - Started in Home instances detected by stripediff:
      - Home-i04: 426
      - Home-i03: 471
      - Home-i02: 1404
      - Home-i01 and Home-i00 -> ongoing

Information Technology Department

Additional Improvements

- Repairing – exceeded replicas in all instances: (space economy)
  - Alice: 2.3Mill to 1.7Mill (16.26% reduction)
  - LHC (no alice) + public instances except alice - >from 360K to 40K (88.89%)
  - Home + Project instances from -> 17K to 2K (88.24% reduction)

Current Replicas vs Expected Replicas (Mayor than 2 replicas)

Information Technology Department
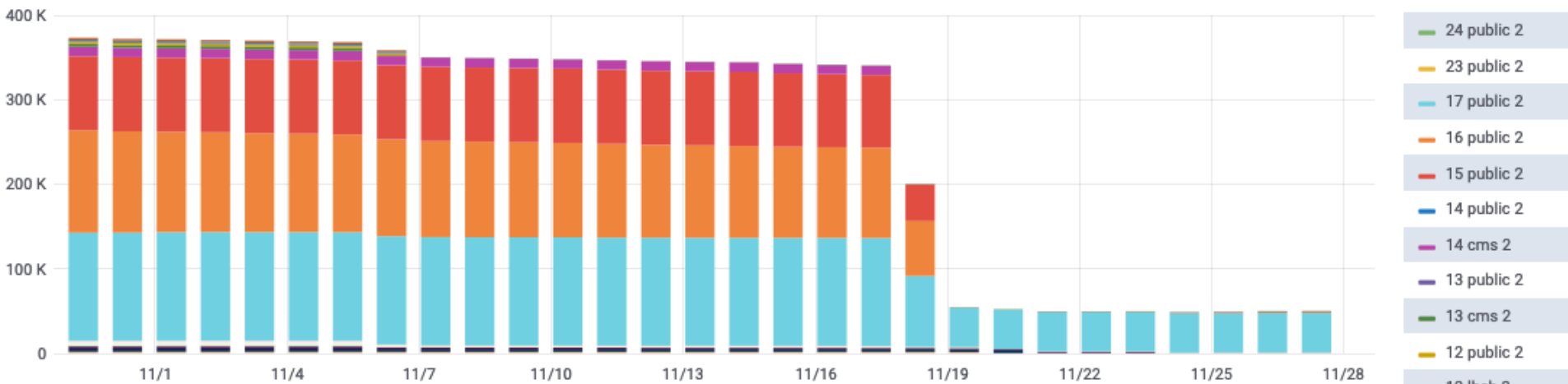
Next Top 3 Objectives

1. Automatic no_rep recovery
   - Aritz is collecting the scripts from Roberto and Remy
     - Also include backup and restic recovery for home & project instances

2. Solve draining leftovers
   - Rename replicas corrupted and provide the list to the service managers
   - Store metadata information to investigate logs

3. Test fsck in pps and collect/plotting info and later test in cms as candidate instance

Information Technology Department