

# Large File System Optimisation and User Education

Philipp Zigann

CERN IT-DSS

28th May 2012



- 1 What do we want to optimise?
- 2 Why do we want to optimise it?
- 3 How do we want to optimise it?
- 4 Process Chain
- 5 Requirements

# What?

Optimisation of our storage system called EOS

- Contains physics data for experiments
- Used for analysis jobs

What to optimise:

- Expanding automated detection of system anomalies
- Adaption for real requirements/typical use cases
- Improving user requests

In short: An infrastructure for early and automated detection of inefficient system access and errors



# Why?

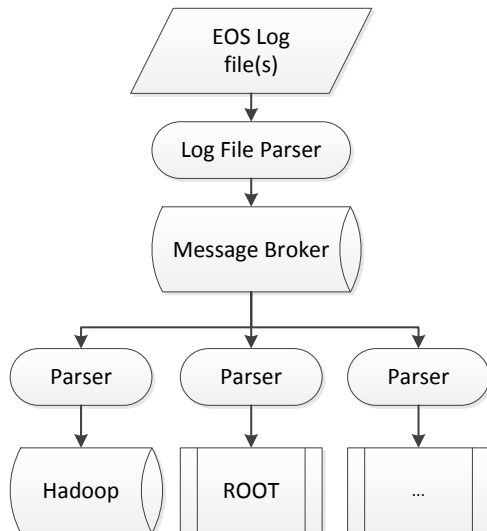
- Faster detection of errors increases speed of solving
- Reducing required resources (traffic, system calls,...)
- Reducing system response time
  - By optimisation for main use cases
  - Makes physicists happier
- Helping users to get best performance
  - Based on user education approaches for databases



# How?

- Collecting access information from storage system
- Analysing and detecting usual access behaviour of users/groups
- Setting up significant metrics (Throughput, Open Duration of Files, ...)
- Generating useful feedback for admin AND user
- Automation of process chain!





- Intelligent metrics and filters
  - Throughput per user/file
  - Read Ratio
  - Call Size
  - Opened Files Per Job
  - ...
- Correct/clear definition of use cases
  - Support/feedback from experiments
  - Feedback from us about the jobs to user

