

Performance Evaluation of Tape Storage Systems

Valentin Honoré¹, Bertrand Simon²

¹ ENSIIE & Samovar, Evry - Palaiseau ² IN2P3 Computing Center (CNRS), Villeurbanne

Presentation of HPDA/PDS Research Projects



Tape usage today



(or 11  football fields)

≈ 20TB on 1000s × 1km read at 10m/s – 100s MB/s

<https://commons.wikimedia.org/wiki/File:LT02-cart-wo-top-shell.jpg>
https://commons.wikimedia.org/wiki/File:Usain_Bolt_Rio_100m_final_2016i-cr.jpg



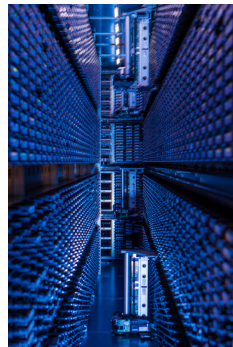
Primordial for HTC (High Throughput Computing)

e.g.,    (100s PB)

also: media companies, cloud archive. . .

😊 Impressive technology improvements
density: + 30% / year (vs HDD: + 8%)

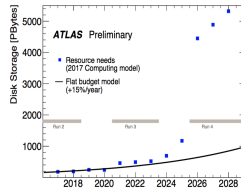
☹ high latency (mount, load, position → few mn)
Adapted for Write Once Read Many



Why using tapes?



up to 6-10 times cheaper overall (before 2020)



[Xin Zhao, HEPiX 2018]



air gap, power failure, lifetime



energy-efficient

Context of the project

- ▶ Reading order of files on tape matters
 - Writing and update operations done separately
 - Most of time spent on reading (many there & back of reading head)
- ▶ Computer Science problem :

Problem (Linear Tape Scheduling Problem)

Given a tape and a list of requests on files on this tape, find a schedule to read the files that minimize average service time (fairness)

- ▶ What we already did in [1]
 - tape models and optimization problem
 - exact algorithm + suboptimal but faster heuristics
 - input extraction from production logs
 - simulation code (in Python) to evaluate different algorithms

Description of the project

- ▶ Go through the following steps:
 - ④ Reproduce the results in [1]
 - ② Implement a performance-oriented simulator to evaluate different strategies
 - ③ Propose algorithmic enhancements for exact algorithm (pruning, branch and bound etc)
 - ④ Analyze real tape system production log (focus on temporality)
- ▶ Taste for programming, algorithmic analysis and development
- ▶ Supervision with Bertrand SIMON (CNRS researcher in IN2P3 Computing Center)
- ▶ We will be in Palaiseau this thursday & friday, do not hesitate to contact us!



V. Honoré, B. Simon, and F. Suter, “An Exact Algorithm for the Linear Tape Scheduling Problem,” in The 32nd International Conference on Automated Planning and Scheduling vol. Proceedings of the Thirty-Second International Conference on Automated Planning and Scheduling, no. 32, Singapore, Singapore, Jun. 2022, pp. 151–159. [Online]. Available: <https://cnrs.hal.science/hal-03482022>