

Collaborative Multi-Granularity Distributed Registry Planning for Fast Container Image Pulling

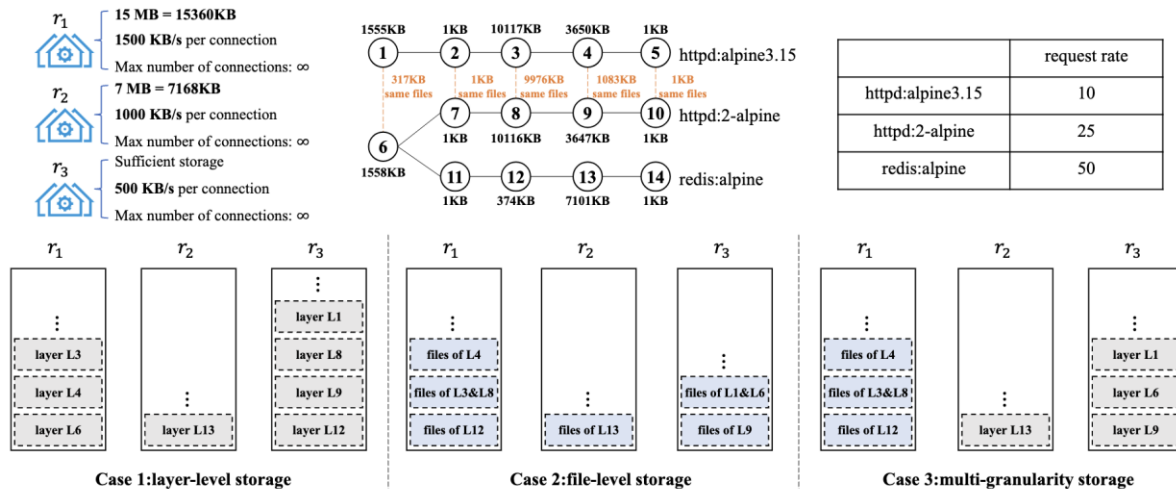
Ziyou SI, Lin GU, Yunzhuo JU, Deze ZENG, Hai JIN

Frontiers of Computer Science, DOI: [10.1007/s11704-025-50350-y](https://doi.org/10.1007/s11704-025-50350-y)

Problems & Ideas

- Problems of Container Image Distributed Storage and Pulling:
 - Registries are geographically distributed, with independent and heterogeneous storage resources.
 - Container images are pulled in layers, but can be stored at different levels of granularity, i.e., layer-level or file-level, each with varying storage requirement and pulling latency.
- Ideas: MIS determines the storage granularity and schedule image storage collaboratively, aiming to reduce the image pulling latency while improving the storage utilization.

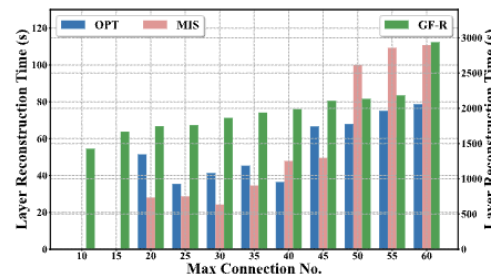
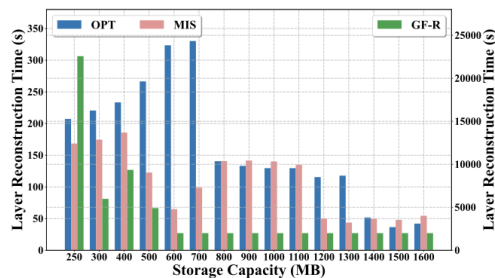
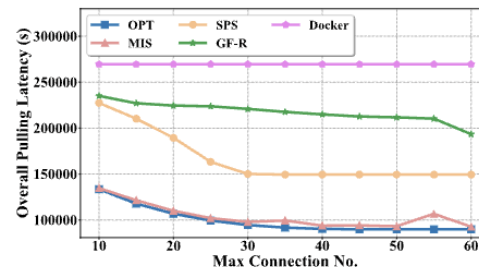
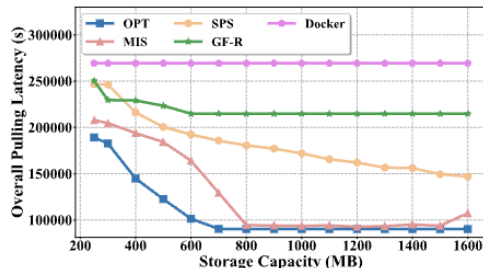
Case	Total Pulling Latency (s)
Case 1	928.15
Case 2	622.50
Case 3	606.75



An example of storing images in three distributed registries at different levels of granularity. The layer reconstruction latency for files is 0.21s. Case 3 yields optimal results by factoring in both image popularity and layer reconstruction latency.

Main Contributions

- Contributions:
 - We are the first to investigate the multi-granularity image storage (MIS) planning problem at both layer-level and file-level for distributed registries. We formulate MIS problem into a non-linear mixed-integer programming (NL-MIP) form and prove its NP-Hardness.
 - We propose a low computational complexity algorithm via randomized rounding with a guaranteed approximation ratio.
 - Experimental results demonstrate that MIS effectively reduces image pulling latency while balancing storage utilization.



Overall pulling latency and layer reconstruction time under different storage capacity and connection number yielded by MIS versus different planning strategy.