

rosie-summary.txt

Erasure codes have been used in storage systems since RAID was developed in the 1980s. However the reliability provided by a single code for an array drops significantly when a disk fails. Some recent work has focused on automatically reconfiguring disk arrays after failures to increase fault tolerance in storage systems.

We propose to study erasure codes and the relationship between reliability and performance in storage systems. When a disk fails, clearly the reliability for the data that was stored on that disk is reduced until it is replaced. If the failed disk cannot be replaced quickly, restructuring the endangered data can improve its reliability by creating an additional copy on another disk. Such restructuring obviously requires available free space on the remaining disks. In addition, there will be a performance impact on the storage system as the data may need to be rebuilt from fragments on multiple disks. By combining several erasure codes, we intend to design a system that reorganizes itself to protect data using any available resources while not impacting performance requirements.