

# Replication & Sync Mail Storage Farm Tech Refresh

## INDUSTRY

Internet  
Cloud Computing  
Online Advertising

## RESULT

- 50% engineering time savings
- 60% productivity gain (6 months reduced to 2.5 months)
- 75% TCO savings

## THE CHALLENGE

- Management complexity with traditional solutions
- Inability to achieve rapid migration convergence with Lots of Small Files (LOSF)
- Lack of scale

## THE SOLUTION

- API enabled automation
- Scale-out architecture for robust operation & high-availability

## BACKGROUND

A tier-1 hyperscaler has been using NetApp FASes (aka Filers or more generally called Network Attached Storage (NAS)) in large numbers to store its users' email, both the messages and attachments. In 2017, the Storage Operations of the hyperscaler needed to retire many old, End-of-Lifed (EOLED) NetApp FASes and consolidate what they stored to fewer number of new NetApp FAS models. The hyperscaler provides one of the most widely used free email services, so the number of messages stored in those retiring NetApp FASes is really immense. The FASes were arranged in groups and each FAS' storage was divided into several volumes. Each volume stored often time hundreds of millions of mail message bodies, which often are very small text files of a few KB each. The most valuable elements of the migration setup are the following:

- The migration software tool must handle Lots of Small Files (LOSF; 20+ - 100s of millions)
- The software tool must be easy to set up and to program for automated migration. It also must provide robust operation in general and offer excellent diagnosis messages, check pointing, and job resumption in case of unexpected interruptions
- The incremental replication must provide very efficient and rapid file system scanning for fast migration convergence

## PROBLEM STATEMENT

Need a software migration tool that provides reliable operations for LOSF and still provides fast migration convergence.

## LIMITATION OF CURRENT SOLUTIONS

The Storage Operations of the hyperscaler tried both free tools such as the commonly used rsync, pftool and the vendor provided data migration tool. But none of them could meet the above 3 basic requirements. Some are difficult to setup since they were never designed for such demanding scenarios. Insufficient reliability either. Even the vendor tool crashed from time to time.

## ZETTAR SOLUTION

The migration task was carried out in the same data center. For high-availability migration the hyperscaler's Storage Operations eventually deployed Zettar zx on two 1U stock servers on the retiring NetApp FAS side, and two 1U stock servers on the new NetApp FAS side, leveraging zx's scale-out capability. zx provided everything that was required and completed the migration, including the desired final convergence and the speed of convergence. Zettar Engineering team members worked with the Storage Operation counterparts to tune zx such that even at the low transfer rate, the software worked solidly, unattended, throughout the duration of the migration.

## MAJOR FEATURES EMPLOYED

Scale-out; LOSF handling; different transfers; built-in WebUI and RESTful API; parallel processing of multiple storage systems and volumes

## ALSO APPLICABLE

- Tech refresh of NASes of other brands other than NetApp.
- Migrate data between a file system and a set of NASes
- Large scale file system migration (between the same kind or different kinds)