

A New Approach To Data Backup

Jay Hutchinson

Dr. Gordon Springer, Thesis Advisor

Abstract

Backing up data in order to protect against its possible loss is important. As more and more data is stored the burden of storing large amounts of backup data becomes significant. This thesis discusses the design and implementation of a new easy to use command line backup utility called Jac. Jac makes use of recently developed deduplication techniques in order to significantly reduce the amount of storage space required for backup data in many cases. Testing was done using Jac to determine its ability to reduce the storage requirements using several different data sets and to determine Jac's performance compared to more traditional backup tools. Results indicate that for many common backup data sets Jac can significantly reduce the storage requirements. Performance was more varied, depending heavily upon the data's content and size.