SSN Project proposal - Secure off-site backup solution

Rory Breuk Pieter Lexis Razvan Oprea

November 17, 2011

Introduction

The OS3 education at the University of Amsterdam has its own infrastructure. With this comes the responsibility of maintaining this infrastructure. This also means that the OS3 Systems Administrators have to backup their servers. These backups will preferably be offsite and encrypted in transit, as well as on-disk. The OS3 team has not implemented this kind of offsite-backup. This project will research the possibilities, properties and performance of said backup possibilities.

We will focus on a solution that encompasses the needs of small business or academic faculties. Generally, the data to-be-backed up consists of many small files (e.g. system files, emails, website pages etc.) which have high-security requirements, meaning the data has to be encrypted end-to-end. One other chracteristic is the high budgetary constraints for implementing such a solution.

Intended audience

This proposal and consecutive paper will be written for system administrators with an academic (BSc or MSc) degree and an interest in the security of systems and networks. Trivial matters about security and backups are considered prior knowledge.

Description

The scope of this project is to analyze the possible solutions for a highly secure, reliable off-site backup solution within the constraints set before.

Partly due to budgetary constraints and fully in line with the OS3 philosophy (Open Standards, Open Software, Open Security), the project scope will comprise solely open-source tools and projects, excluding commercial applications.

To ensure that the recommended solution is future-proof, we will not include any deprecated technologies.

Research Question

What is the best backup solution for off-site storage with regards to performance, security and disk usage?

To better answer this question, it is necessary to answer a few sub questions. These allow us to define the parameters of the research and the final backup solution.

Hypothesis

We believe that the best solution is to use on-site backups, encrypt and send them to an off-site location using an agnostic transport and remote storage mechanism.

SubQuestion 1

How can we combine source-based encryption with incremental backups?

SubQuestion 2

What is more appropriate for our situation: encrypting the data at the source, or at the destination?

SubQuestion 3

What is the best encryption method (for instance, symmetrical or asymmetrical, choice of algorithm, etc.)?

SubQuestion 4

What is the best storage solution for off-site backups (Linux server, FTP account, cloud storage on-demand such as Amazon S3, etc.)?

Previous work and preliminary research

There is work done in the field of backups, the following list of works will be used (among others) for this research:

- “Cumulus - Filesystem Backup to the Cloud” - Michael Vrable, Stefan Savage, Geoffrey M. Voelker (University of California) - `http://dl.acm.org/citation.cfm?doid=1629080.1629084`
- “Fast and Secure Laptop Backups with Encrypted De-duplication” - Paul Anderson, Le Zhang (University of Edinburgh) - `http://www.usenix.org/events/lisa10/tech/full_papers/Anderson.pdf`
- “Content addressable storage for encrypted shared backup” - Gilroy N. Gonsalves - `http://homepages.inf.ed.ac.uk/dcspaul/publications/CASFESB.pdf`
- “Backup & Recovery: Inexpensive Backup Solutions for Open Systems” - W. Curtis Preston - Publisher: O’Reilly Media - ISBN-13: 978-0596102463

Estimated Schedule

Week 1

Task	Time	Done by
Investigate different encryption and compression implementations	10	
Investigate incremental backup methods	10	
Investigate storage solutions	10	

Week 2

Task	Time	Done by
Finish investigations and start writing the report	20	
Start benchmarking	10	

Week 3

Task	Time	Done by
Run benchmarks	20	
Analyse benchmark data	10	

Week 4

Task	Time	Done by
Finish analysing benchmark data	10	Razvan, Pieter, Rory
Finish the report	20	Razvan, Pieter, Rory
Write the presentation	5	Razvan, Pieter, Rory