RuBackup

Система резервного копирования и восстановления данных

Утилита RBD

Версия 1.0 2019

Оглавление

Введение

Утилита командной строки RBD используется для дедупликации файлов. Утилита создаёт хэш-таблицу файла, на основании которой могут быть впоследствии созданы разностные копии изменённой версии файла. При помощи утилиты можно воссоздать файл из его старой и разностной копий.

Алгоритмы хэш-функций, реализованные в RDB

Наименование алгоритма	**Длина хэш, бит**	**Примечание**
streebog или GOST_R_34_11_2012	256, 512	ГОСТ Р 34.11-2012 «Информационная технология. Криптографическая защита информации. Функция хеширования» - действующий российский криптографический стандарт, определяющий алгоритм и процедуру вычисления хеш-функции. Разработан Центром защиты информации и специальной связи ФСБ России с участием ОАО «ИнфоТеКС» и введен в действие 1 января 2013 года.
sha	256, 512	Хеш-функции SHA-2 разработаны Агентством национальной безопасности США и опубликованы Национальным институтом стандартов и технологий в федеральном стандарте обработки информации FIPS PUB 180-2 в августе 2002 года
skein	256, 512	Skein — алгоритм хеширования переменной разрядности, разработанный группой авторов во главе с Брюсом Шнайером. Хеш-функция Skein выполнена как универсальный криптографический примитив, на основе блочного шифра Threefish, работающего в режиме UBI-хеширования. Основные требования, предъявлявшиеся при разработке — оптимизация под минимальное использование памяти, криптографически безопасное хеширование небольших сообщений, устойчивость ко всем существующим атакам на хеш-функции, оптимизация под 64-разрядные процессоры и активное использование обращений к таблицам
blake2b	256, 512	BLAKE2 — криптографическая хеш-функция, улучшенная версия BLAKE — одного из пяти финалистов конкурса на хеш-функцию SHA-3 (главным образом улучшено быстродействие), представлена 21 декабря 2012 года. Разработчики: Jean-Philippe Aumasson, Samuel Neves, Zooko Wilcox-O'Hearn, и Christian Winnerlein. Была создана как альтернатива широко использовавшимся в прошлом MD5 и SHA-1, в которых были найдены уязвимости.

Использование RBD

Для примера ниже нужно выбрать два каталога (test1 и test2), желательно со значительным объемом располагающихся в них файлов:

1. Создаем тестовый файл №1:

 # tar cvfz test1.tgz ~/test1/

2. Создаем тестовый файл №2:

 # tar cvfz test2.tgz ~/test1 ~/test2

3. Создаем хэш-таблицу первого файла:

 # rbd -v -a sha -i test1.tgz

В результате этой команды будет создан файл ./test1.tgz.sha_512.ht

4. Создаем хэш-таблицу второго файла и разностную копию, но основании данных, содержащихся в хэш-таблице первого файла:

 # rbd -v -a sha -i test2.tgz -c ./test1.tgz.sha_512.ht

В результате команды будут созданы два файла: разностная копия test2.tgz.pt и хэш-таблица *./test2.tgz.sha_512.ht*

5. Восстанавливаем тестовый файл №2 из тестового файла №1 и разностной копии:

 # rbd -v -a sha -i test1.tgz -x test2.tgz.pt -n restored.tgz

В результате команды будет создан файл restored.tgz, полностью идентичный тестовому файлу №2.

Описание всех опций RBD

-h справка

-a хэш-функция, можно выбрать следующие варианты: streebog или GOST_R_34_11_2012, sha, skein, blake2b

-o файл, куда будет записана хэш-таблица

-i входящий файл

-c входящая хэш-таблица (для сравнения хэш-таблиц и создания разностной копии)

-p файл, куда будет записана разностная копия (работает только с опцией -c)

-v режим расширенной информации о работе утилиты

-b размер блока данных при работе утилиты (размер по умолчанию 1048576 байта, максимальный размер блока данных 104857600)

-x входящая разностная копия (для суммирования с входящим файлом, невозможно одновременное использование с опцией -c)

-n файл, в котором будет записан результат сложения входящего файла и разностной копии (работает только с опцией -x)

-2 (256 бит длина хэш)

-5 (512 бит длина хэш)