

DATA DEDUPLICATION ASSESSMENT



CONTACT

Jim Smid
Practice Director
Data Center
303.526.6760
jim.smid@ironbow.com

IRON BOW TECHNOLOGIES

Iron Bow Technologies forwards a history more than 20 years as a leading full-lifecycle IT solutions provider. Leveraging strong resources and a collaborative approach, the company delivers targeted solutions for data center, collaboration, network infrastructure, information security and client computing.

ASSESSMENT OVERVIEW

Understanding your backup, recovery and archiving requirements and current limitations is the first step in planning and designing an infrastructure to address your business challenges. Iron Bow's Data Deduplication Assessment (File Commonality Scan) provides intelligent data analysis that is designed to be a non-disruptive, time-efficient and a cost-effective way to assess the current state and performance of your existing backup environment.

The Data Deduplication Assessment is a collection of scripts designed to emulate the backup process (without actually writing the data to disk). Data is gathered on deduplication change rates and backup times, to generate a real world deduplication ratio based upon actual customer data. The assessment results provide invaluable data that is used to properly architect and size a deduplication solution.

ASSESSMENT SCOPE

Through this service, you will gain specific operational and procedural insights into your current environment and practices, along with recommendations for improvement. As a result, you will be better prepared to design and implement a solution that allows you to reduce costs and meet the data-protection, retrieval and recovery expectations of your business. The Data Deduplication Assessment will provide a summary of your backup environment, and report on the expected deduplication ratios of the backup infrastructure. This will enable the calculation of savings, which would result from decreased tape usage and faster backup times upon implementation of the recommended deduplication solution.

The deliverable is a presentation or report that will highlight any redundant data stored across the servers and provide an understanding of the potential effectiveness of deduplication for your data. Various types of data can deduplicate in a different manner, so it is important to understand the data in your environment to ensure the proper sizing and expectation of performance of the solution you select.

ASSESSMENT SCOPE (CONT.)

There are three key phases that are conducted during assessment activities:

Consultancy – Iron Bow will work with your organization to outline the scope and objectives of the Data Deduplication Assessment and to identify the servers that are to be assessed. Additionally, our engineers will conduct interviews with your IT team and identified stakeholders to gather information that will provide a robust and accurate assessment of your operating environment.

Assessment – As part of the service, Iron Bow employs a collection of utilities and scripted tools, which are used to scan, collect data and report on usage in the customer's environment.

Data deduplication scans are required to run (scheduled via the Data Profiling tool) over multiple days in order to capture a good snapshot of data activity in the environment. The first commonality scan takes about the same amount of time as a normal full backup. Depending upon the amount and size of the files in the scan set and the total data collection the period will run from three to seven days. Once the data collection has been completed, the output files are placed in a temporary folder, whereupon they will be compressed and sent back to Iron Bow for analysis.

Analysis and Recommendations – Upon receipt of the output files, Iron Bow will analyze the collected data and will document our findings.

Our engineers will deliver a comprehensive overview of the assessment results and provide recommendations for any deduplication scenarios and infrastructure improvements that should be considered.

(Note: No actual customer data will reside on the data collection server.)

ASSESSMENT TIMELINE

The initial onsite visit - which includes stakeholder interviews and script installation - will run no longer than one day. Data collection will last about one week and once the data has been collected, data analysis will be completed. The presentation of the assessment's findings and final data analysis will be conducted by Iron Bow at a place and time of the customer's choosing and will run approximately two hours.

Please contact an Iron Bow Technologies representative to schedule your Data Deduplication Assessment today.