Paul Womer

Data Deduplication

MIS 2501

By implementing Deduplication we stand to save $96,800 per year worth of physical storage. In addition to decreasing our physical storage we will increase the speed of backups and data retrieval. Additional savings will result because fewer disks will have to be deployed, powered, and cooled. The removal of redundant data from our disks will also be beneficial to our disaster recovery process. The end result of Deduplication will be an efficient utilization of our data storage resources.

Deduplication is a form of compression, which reduces the amount of data that needs to be stored. Once a full system backup is done, all future backups will only record changes to that data. This will eliminate our current redundant system that records all data weekly, thus reducing the time it takes to back up our data. Having the same data in a condensed form will decrease the time it takes us to retrieve archival data making us more efficient and save us the cost of powering and cooling the extra disks. We will also see a positive impact in our disaster recovery process; less network connectivity is required since each input/output operation carries a larger data payload. An upgrade to Deduplication will save us time, space, and money.

We currently do a full back up on a weekly basis; each of these backups is 2TB and we store the data for one year. On average only 5% of data changes per week which means we would need 2TB for the initial data and only .1TB for the weekly data that is changed, for a yearly total of 7.2TB as opposed to 104TB. Our storage cost is $1 per GB; the lower storage capacity requirement represents $96,800 savings.

Works Cited

Rinnen, Pushan, and Dave Russell. “Hype Cycle for Storage Technologies, 2012.” *Gartner* 5 Jul. 2012. Web 18 Sept. 2012.

Salamone, Salvatore. “Making the Business Case for Deduplication.” *eweek* 9 Jun. 2010. Web 19 Sept. 2012.

“NetApp Deduplication and Compression.” *Netapp* n.d. Web 19 Sept 2012.