# The Cost of Data Loss and How to Avoid It

**Dan Olds**

**Evolution**

## Executive summary

Almost every business is now data driven and many are realizing that their data is one of their most valuable assets. But how well are they protecting that data from IT problems that can cause data loss? In this report, we discuss the value of data and take a look at a readily available solution designed to protect against data loss.

## Introduction

What if your organization lost a day's worth of operational data? What would that mean to your company? Could you bill your customers? Could you run your payroll? Could you sell to new customers? Would you be able to deliver your products to your existing customers? Some of the answers to these questions are probably a resounding "no." And that's a big problem.

Everyone knows that business data is highly valuable, right? From various sensor readings, to customer information, to gathered market analytics, data is rapidly becoming a critical asset. Organizations are continuously finding new, innovative ways to utilize this information, allowing them to be more effective in the market.

## What's your data worth?

There are ongoing efforts to quantify the actual value of the data that an organization possesses. OrionX is finding that the value of data is much more than most financial analysts realize.

For example, in 2016, Microsoft purchased LinkedIn for $26 billion, which represented about 91 times LinkedIn's earnings at the time. For comparison sake, Alphabet (parent company of Google and a highly profitable company) currently has a price/earnings ratio of about 33. If you applied Alphabet's P/E ratio to the purchase of LinkedIn, Microsoft should have paid about $9.5 billion for the company. Some would argue that a large portion of the $26 billion minus the 'reasonable' price of $9.5 billion represents the value of LinkedIn's vast trove of data.

Others believe that the value of data in an organization is the difference between the tangible value of the company (the value of assets such as the buildings, cash, machinery, etc.) and its intangible value (such as branding, or goodwill). The worth would be measured as the total value of the company minus the value of tangible assets – not exactly a precise measure.

Another way to measure the value of data is to look at how companies are affected by the temporary loss of data access and its effect on the business; in other words, the impact of IT downtime.  According to IT research house, IHS Markit, the average cost of unplanned downtime for a mid-sized company (100-1,000 employees and $100 million in revenue) is $1 million per year. The annual costs of downtime for large companies (13,000 employee or greater, and an average of around $2 billion in revenue) are more than $60 million per year.

One recent example is the downtime that Delta Airlines experienced. Back in September 2016, the airline experienced a five-hour IT failure that forced them to cancel a large number of flights and also affected their reservation system. Delta estimates that this outage cost them upwards of $150 million. Southwest Airlines experienced IT failures that forced them to cancel flights over a three-day period in July 2016, at an estimated cost of $177 million. The largest airline IT-related shutdown grounded the entire British Airline fleet for 36 hours and cost upwards of 500 million Euros.

Experts point to a lack of adequate disaster recovery mechanisms as one reason why these types of outages were so pervasive and costly.

Calculating the cost of downtime or not having access to data is tricky business; it varies according to industry and is highly dependent on which systems experience the outage. However, it's safe to say that any IT outage has the potential to cause chaos in the organization and become expensive.

## The True Value of Data

The true value of data is a function of a dynamic set of attributes attached to that data. Within your organization, data usage determines its criticality. Data is an asset, and just like any other asset, it has different values at different times and under different conditions. While the value of your data assets is market driven, it has attributes that indicate its value to the organization. So, what are these attributes?

We've listed the most common characteristics driving the value of data in Figure 1.

The first characteristic is **Size**, which is somewhat self-explanatory. The larger the amount of data lost, the larger the impact to current operations. One must also consider the cumulative value of many pieces of low value data being lost.

The **Category** of the data is key to ascertaining the value of the data lost. For example, losing a large piece of the current quarter financial data is potentially devastating, while losing a like-sized portion of employee email probably isn't nearly as damaging.

**Convertibility** refers to the actual monetary value of the data that is lost. If an entire day's worth of order transactions is lost, the value of those orders is an absolute monetary loss which can be measured by the average daily revenue of the firm. A similar approach could also be applied to losing a database of prospective customers, although the value of this list is more difficult to quantify.

**Reproducibility** is an issue that often comes up when dealing with data loss. Is it possible to reproduce the lost data by piecing together data from other sources? How long would it take and how much effort would be involved in reproducing the data? The less replicable the data is, the higher the value of the data.

The **Timeliness** of the data dynamically changes as organizational needs shift. Archived data may not have much immediate value until it is needed by, for example, a legal team who is trying to prepare for a trial. Historical sales records might not be considered vital until the sales department is working on a proposal for an important new procurement from an existing customer.

## Data Loss Drivers

- Size
- Category
- Convertibility
- Reproducibility
- Timeliness
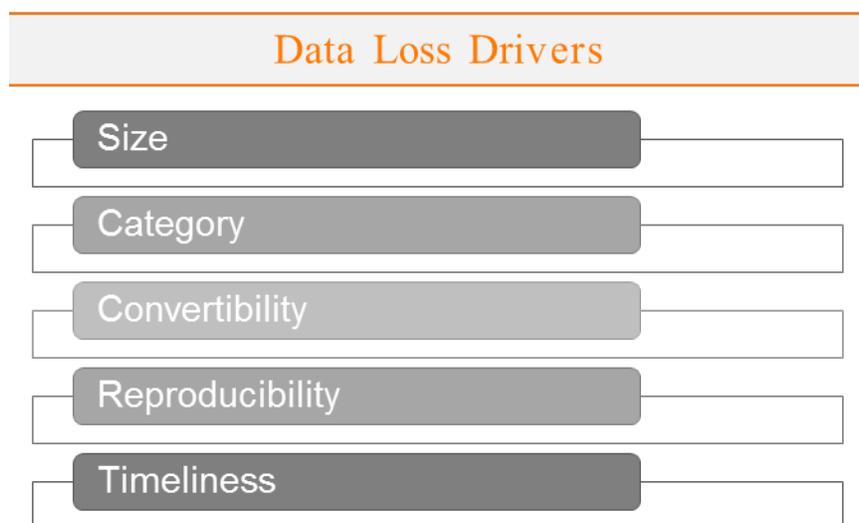
*Figure 1: Data Loss Drivers. Source: ©2017 OrionX.net Research*

## Data Loss Costs

Figuring the exact costs arising from data loss is a complex undertaking. We've found that these costs typically fall into some or most of the categories listed below.

**Lost revenue** is typically one of the easier costs to understand and track. It is simply the amount of revenue that the company doesn't receive due to the data loss. This gap can be orders that aren't recovered at a later date or the cost of shipping products to customers after due dates.

**Lost time**, the cost of **regeneration**, and cost of **absorption** are a bit more difficult to quantify. These areas include employee time taken up by the data loss. Some examples of this lost productivity include IT employee time used to quantify and qualify the data loss, and their attempts to recover the data. It would also include employee time consumed by contacting customers and other parties to reconstruct the data. The costs associated with an outside loss consultant would also fall into this category.

### Data Loss Costs

- Lost revenue
- Lost time
- Cost of regeneration
- Cost of absorption
- Impact on Brand
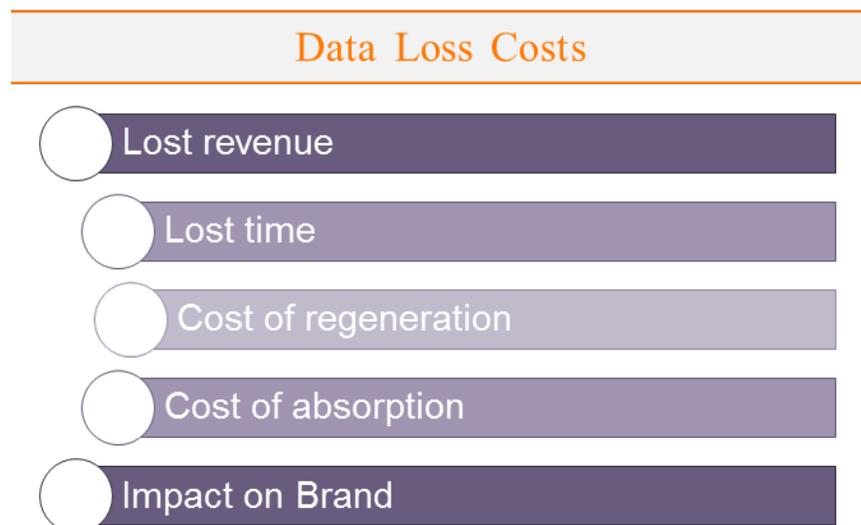
*Figure 2: The Cost Categories of Data Loss. Source: ©2017 OrionX.net Research*

Perhaps the largest impact could be on the **brand** of the company. A sufficiently large loss of critical data could have wide-ranging impacts on the organization's customers and other stakeholders, resulting in a loss of stock market value, the loss of customers, and even regulatory investigations.

## Data Loss vs. Downtime

Data loss and IT downtime are often confused and jumbled together. While data loss often coincides with IT downtime, this isn't necessarily the case. Not every instance of downtime causes significant data loss, and not every instance of data loss is accompanied by IT downtime.

Even companies with solid back-up plans can experience significant problems from data loss. Let's take the example of a company that backs up their systems every day at midnight. If they experience an IT disaster that takes down all of their systems at 11:59pm, for instance, they could lose an entire day's worth of processing.

## How to Avoid Data Loss and Data Inaccessibility

Every system is vulnerable to data loss and experiencing conditions where data isn't accessible. A magic operating system or application software that doesn't ever fail simply doesn't exist.

Back-up can help reduce the data loss that occurs with unplanned downtime; however, it all depends on when the last back-up was taken. If you're lucky, you might only lose a few hours' worth of data; if you're unlucky, you might miss an entire business day.

However, there are software packages and features that can be used to radically reduce the impact of IT outages, data loss, and data inaccessibility. Most of these are proprietary hardware or software solutions that come with eye watering price tags, long implementation cycles, and reduced flexibility. While these solutions might be the right answer in some cases, we believe that there is a more effective approach that should be explored first.

## Open Source Solutions

A major trend in IT is the emergence of open source solutions which reduce costs and leverage the creativity of a community. Once they reach maturity, they become a lower-risk choice for organizations than proprietary options since the functionality is acceptable -- if not leading -- the source code is available, the licensing process is less complicated, and the solution is immune from the whims of a specific company. Today, there are open source solutions in nearly every technology category.

When it comes to data protection, one open source solution is already built into the Linux operating system. It's called LINBIT DRBD and it ensures that data is replicated from one system to another location (or locations) quickly and safely.

The LINBIT DRBD software follows a simple concept:  when someone saves a file to one location, whether it's their own local drive or a NAS, that data is also copied onto another

location or locations – perhaps locally, perhaps at a hot back-up site or a set of disaster recovery systems.

Using LINBIT DRBD software eliminates single points of failure that can cause data loss by replicating data onto many systems. This way, no matter what, your data still exists, regardless of whatever IT disaster might befall a particular datacenter.
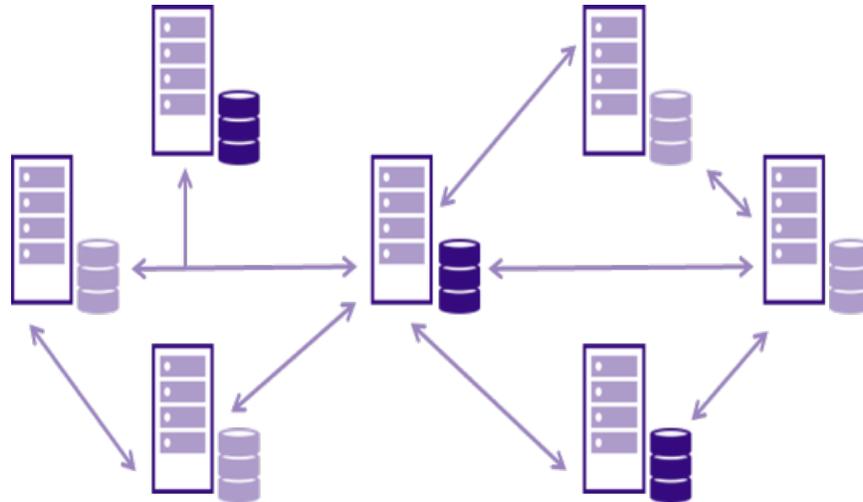


*Figure 3: One open source solution is already built into the Linux operating system.*

The LINBIT DRBD software specializes in providing high availability (HA) for Linux, and soon, Microsoft. LINBIT DRBD is the lead developer of the DRBD software, which has the goal of providing the highest degree of data availability.

Their position as project lead has given them the unique ability to drive development, and, incidentally, to become the world's leader in Linux HA and Disaster Recovery software.

With LINBIT's DRBD software, customers can easily mirror data over long distances to one or more datacenters. As can be seen in the picture above, the data from the central system is being replicated onto six other systems – simultaneously, and without operator intervention.

This software can be run in an on-premises data center, in public or private cloud and on commodity hardware. It also works with most virtualization platforms. It works on all major Linux operating systems and is fully compatible with nearly all applications, devices, and services.

Customers can set up LINBIT's DRBD to sync data volumes in multiple locations in real time. With their DRDB Proxy, files can be compressed before replication for faster performance when bandwidth is limited.

The business value here is pretty simple to understand: your valuable data is saved to multiple locations in real time, thus making sure that you always have other copies just in case you need them.

## Summary & Recommendations

Companies should take a hard look at their data and work to understand the value of it. The next step is to make sure that, at the very least, their most valuable data is protected against loss by replication. We would also recommend that customers investigate DRBD and how companies like LINBIT have improved this mechanism in order to make it more powerful and user friendly.

Please visit OrionX.net/research for additional information and related reports.