

White Paper

The Hidden Costs of Developing Internal Storage Reporting Tools

Everyone knows that storage analytics are extremely valuable, providing critical data about on-going capacity, performance and allocations that are vital to sustaining business operations in today's datacenters. This data is required to understand usage, resolve issues and forecast growth. Vendor-supplied management tools typically provide all of this information on demand, however, the complexity comes when trying to provide a consolidated view across not only multiple arrays but also, in many cases, multiple vendors and their associated tools.

As the information from storage systems is freely available, the expectation from IT management is that there should be little or no cost for consolidating and analysing this information. This document shows the areas where this cost assumption not only dramatically underestimates the value of this information but also introduces significant risk to the business.

Internal Tools: The Logical Next Step?

A storage array in isolation provides a wealth of information via a GUI or associated Command Line Interface (if you prefer to extract your information that way). Gathering information from multiple arrays either requires brute force or some form of automated scripting and a system to enter and display that data. The most popular option is Microsoft Excel and the Pivot table where multiple sheets take care of daily information and can be summarised and distributed to a wider audience.

It sounds good in theory, but in the real world this process soon gets out of control and manifests itself in various issues associated with its usefulness and value to the business. There are many other customer specific problems with this method that will not be outlined here.

Issue: Internal Skills Development

Internal reporting tool projects usually start off as skunk-works projects by storage administrators that eventually require programming skills development requiring more time being spent on developing the internal tool. When new storage technologies emerge such as storage virtualization, thin-provisioning, auto-tiering and data deduplication, the internal tool requires a significant update to incorporate this new information. If there is a multi-vendor storage strategy, the difficulty of representing this data consistently and in a meaningful way is multiplied.

Issue: Silos of Information

Collecting data from storage arrays does not provide any context, so integration is required to understand the connectivity between applications and storage. If servers have been virtualized then it is very difficult and complex to connect that information to the appropriate storage arrays. There are also huge difficulties in multi-location environments that may have network access issues and different teams responsible for administration.

The vast majority of these tools are owned by the storage team and provide little or no value to a wider audience including CIO, capacity planning, procurement, audit and compliance officers. In addition, the tools are generally not open for scrutiny or auditing for accuracy.

Issue: Data Atrophy

Unless data collection is automated information very quickly loses its value through atrophy. Even if this process is automated the accuracy of collected data deteriorates over time due to the ever-changing nature of storage infrastructure. This results in bursts of activity in re-validating data, usually elicited by an audit or capacity planning exercise. This reactive working practice is disruptive to normal operations and can be costly to the business.

Issue: Corporate Risk

An internal tool adds significant risk to the business. The daily working and development of the tool will be reliant on a small number of staff or a single developer who can easily decide to exercise their skills elsewhere (skills that they have been acquiring outside the normal role of a storage administrator on the company's time). There have been many examples where tools have had to be reverse-engineered by a new team. This is time that storage reporting will not be taking place and has a high chance of failure.

Commercially Available Products

Software products are available on the markets today that have been developed over a long period of time to provide enterprise-class, highly available reporting solutions. Benefiting from many thousands of days of testing and deployment in diverse environments means that the time taken to deliver real business value to the organisation is much shorter. This extensive development and testing capability leads to a number of key advantages that will never appear in home-grown tools.

Feature: Data Interpretation

Collecting data from storage devices is only part of the solution. Each vendor represents its data in different ways with diverse naming conventions. At the same time, technologies such as thin provisioning, virtualization and data deduplication make understanding what is used, allocated or available very difficult to deduce in a holistic way. Commercial products have already developed and tested this information to make sure that it is presented accurately across different vendors.

Feature: Dissemination of Information

Commercial products have multiple ways of publishing collected data. The best products typically have a multi-tenanted web-based portal that provides granular access to information. For example, auditors can be given access to compliance reports which can reduce the amount of time they have to spend with the operations team. Other ways to publish include email, documentation, conditional alerting and integration with other reporting frameworks.

Feature: Historical data

The use of relational databases allows information to be analysed on an on-going basis to identify recurring issues, repeat errors and capacity problems. If your operations team is relying on manual analysis, there will be at a huge time cost without the guarantee of identifying trends or dilemmas. Historical data provides the basis for accurate future forecasting, allowing users to reduce the storage overhead and avoid emergency purchases.

Choosing the right vendor

When considering the right solution you will need to consider some of the following differentiators:

- **Support of existing and future storage products**—Independent vendors typically have a much broader set of supported devices than storage vendor owned products.
- **Integration with business metrics**—Make sure that the solution provides integration with your business specific information such as cost centers, departments, owner, location etc. This allows reports to be generated that are relevant to your business and creates a much more valuable solution
- **Flexibility**—Most users have reports that they would like to see in a specific format, so make sure the solution can deliver the reports in the way you want to see them. Changing your business process to fit or manually manipulating reports quickly reduces the value and longevity of the solution.

Powerful storage reporting solutions can provide a wealth of information that helps you better manage the storage environment. However, home-grown tools can be more trouble than they're worth, adding layers of complexity and management overhead to an already complex process. Organizations are better served by embracing commercial solutions that provide a holistic view of the storage environment in real time and present the information in a rich yet intuitive manner. This will allow them to get a better handle of complex storage infrastructures without requiring additional time, budget or staffing.