VIOLIN
Better Business Performance

SOLUTION BRIEF

HIGHLIGHTS

- Appropriate applications can benefit from significant capacity savings with data reduction tools.
- Applications that are poor fits for data reduction need to be treated differently
- Always "on" data reduction may actually hurt overall data center operations

GRANULAR CONTROL

- MAESTRO provides inline data reduction that is switchable by share or file
- VIOLIN All Flash Arrays provide post process data reduction on a file by file basis

Optimize Data Reduction Technology with Granular Control

Data reduction technologies can be powerful for the right applications, but it is not useful for all applications. Granular control is the best way to use these tools.

When data reduction technology is appropriate for the workload, the results can be stunning. Imagine reducing the amount of storage required by 5x, 10x or more? The effective cost of storage plummets when data reduction is a good fit. Although data reduction technology has been available for a long time, the arrival of all-flash arrays has brought it to the forefront of storage discussions.

What is Data Reduction?

Simply put, "data reduction" is the ability to avoid writing redundant blocks of data to reduce the amount of storage required. Instead of writing redundant data, systems use pointers to the original unique blocks of data. This can result in significant savings in the amount of storage used by reducing writes. There are two kinds of data reduction: inline (which is processed before writing data), or post process (where data is reduced after the initial write, to consolidate space needed).

Think of data reduction as a four step process:

- Incoming data is broken into smaller blocks
- Data blocks are then summarized as a mathematical expression (hash)
- Duplicate hash blocks are identified (easier to do with hash than actual blocks)
- Unique data is written to storage, and duplicate blocks are identified with a pointer to the original unique version

How Does Granular Control Work?

Data reduction works very well in the right environment, and can provide big savings for your storage budget.

Examples of a great fit for data reduction include virtual desktop infrastructure (VDI) and virtual server infrastructure (VSI). A great fit for data reduction is found when the files are mostly the same, so eliminating redundant blocks will have a big impact. When the files are all the same, there is no need for data reduction, a clone is created and replicated as needed. The best fit for identical files is really fast storage, like the VIOLIN All Flash Arrays. Adding data reduction where it is not needed just adds latency, and makes things less productive. For instance, when dealing with VDI environments, persistent desktops are mostly the same with some unique elements, and are great fit for data reduction. For VDI environments with non-persistent desktops, the desktops are all the same, and a poor fit for data reduction. The lesson learned is to use data reduction where it adds value, and don't use it where it just adds latency.

A survey by IDC looking at user workloads identifies VDI and VSI as a small portion of the typical data center workload. Workloads that are typically not a good fit for data reduction would include databases, transaction-oriented applications, and analytics. In particular, consider data reduction's impact on databases tables. Tables are usually eliminated with data reduction, since the data exists elsewhere in the database. Tables exist to speed performance of the database, by reducing lookup time. By eliminating the tables in a database, performance could suffer. The right solution is to use the technology where it is a benefit, and not where it doesn't add a benefit.

There is a reason that some all-flash arrays have "always-on" data reduction. All-flash arrays come in two varieties: SSD-based and backplane based. The need for data reduction is different for the two architectures. SSD-based designs have two levels of management: one in the SSD itself, and an array-level controller. Data reduction can help SSD-based arrays manage their flash resiliency by limiting write traffic, since it cannot coordinate writes between its SSD and array-level controllers very well.

One of the few examples of a backplane design is VIOLIN's all flash architecture which does not use SSDs. Instead VIOLIN puts flash chips on specially designed cards that plug into an optimized backplane and unites the management into a coherent unit. Hence there is no need for continual data reduction to manage writes, and the overhead of data reduction processing does not need to be incurred in a VIOLIN array resulting in better performance and resiliency.

VIOLIN provides data reduction that allows you to decide on a share by share, or file by file basis what needs to be data reduction, so no unnecessary latency is incurred by an unneeded, or even detrimental data reduction process. VIOLIN All Flash Array provides post process data reduction that also allows you to choose which files to leave untouched. You have the option to use the technology, or not, as the application demands.

Results

Because VIOLIN's approach allows you to turn data reduction on and off, you can assess the value of data reduction technology for your applications.

VIOLIN puts you in control of critical data reduction capability, allowing you to use it where appropriate, and to avoid the performance penalty where it is not needed.

