

TECHNICAL HIGHLIGHTS



Cirrus Data's
Storage Consumption Planner for Azure

CirrusData

www.cirrusdata.com

Cirrus Data Cloud's Microsoft Azure integration has been expanded to automatically recommend a volume type, size and performance parameters based on the actual performance of the source volume.

What is the Cirrus Data Storage Consumption Planner for Microsoft Azure and how does it work?

Cirrus Data Cloud (CDC) is integrated with various storage providers to automatically allocate destination storage volumes that match the migration source storage volumes. Unlike integrations with other storage providers, where users must specify additional volume properties during the auto allocation process (such as performance tiers, throughput and IOPS, etc.), Cirrus Data Cloud's Microsoft Azure integration has been expanded to automatically recommend a volume type, size and performance parameters based on the actual performance of the source volume.

Knowledge of the performance statistics and IO access information used by the existing source volumes is crucial to accurately recommend destination volume sizing and performance. When a user installs Cirrus Migrate Cloud (CMC) on a Windows or Linux host, it immediately begins sampling IO statistics every second on all attached volumes. CMC collects data about the IOPS, throughput, latency, types of IO (read/write/UNMAP), IO size, randomness, etc. These samples are then consolidated and downsampled by the minute. By default, after 30 days these IO statistics are rotated to minimize the storage usage on the volume on which CMC is installed. All I/O statistics and access information are stored locally inside user's hosts.

When a user starts the Storage Consumption Planner for Microsoft Azure, they can specify various tunable parameters such as the amount of time CDC should consider when aggregating statistics from the source host. They can also specify whether they expect their performance and capacity requirements to be changed post-migration. Finally, they can select a recommendation profile among "Low Cost," "Balanced," and "High Performance." Once the user customizes these parameters (if necessary) and clicks the *recommend* button, CDC begins performing a series of tasks to recommend an accurate volume size and performance to the user.

Cirrus Data Cloud will first fetch an appropriate amount of data points for the source volume and aggregate them into minimum, maximum, average, and standard deviation for each data type

(IOPs, throughput, etc.). Then, the destination host running in Microsoft Azure will be queried for information that will be used to determine what disk types are eligible for recommendation. CMC uses the local metadata service along with Azure's public API to query information about the destination host. Information that may affect recommendation includes location, availability zones, machine type, ultra disk compatibility, shared disk compatibility, etc.. All of this information is gathered and then delivered to Cirrus Data Cloud's recommendation service that is responsible for finding an appropriate recommendation and estimating the cost.

For Microsoft Azure, there are a number of disk types with different characteristics, ranging from Standard HDD to Ultra Disk.

For purposes of calculation the disk types are split into two categories: non-scaling and scaling disk types:

- **Non-scaling** disk types such as Standard HDD and Premium SSD have performance tiers that are based on the capacity of the created disk. When creating these types of disks, performance characteristics are not customizable. Instead, users must create the disks with enough capacity to meet the minimum capacity of various performance tiers.
- **Scaling** disk types such as Premium SSD v2 and Ultra Disk are also allocated and performance-bound based on capacity of the volumes. However, for each GiB allocated, users are allowed to provision a certain amount of maximum IOPS and for each IOPS provisioned users can allocate up to a certain amount of maximum throughput.

Cirrus Data Cloud's Storage Consumption Planner has full awareness of the different disk types and disk tiers in Azure, including minimum capacity, how fast IOPS/throughput can scale, as well as performance characteristics of each disk tier.

While there are limited selections from the user-facing management interface, each of the aforementioned disk tiers and types is a unique SKU and each of them has a number of price meters associated with it. For example, a Premium SSD v2 disk has an independent meter for capacity, throughput, and IOPS. The prices of these meters also

Cirrus Data Cloud's Storage Consumption Planner has full awareness of the different disk types and disk tiers in Azure, including minimum capacity, how fast IOPS/throughput can scale, as well as performance characteristics of each disk tier.

Each time the service performs a recommendation, it always returns the lowest price configuration that satisfies the recommendation target values.

vary across regions. When providing recommendations, Cirrus Data Cloud utilizes the Microsoft Azure Retail Rates Prices API to gather current location-specific prices of all the price meters involved and takes into account different pricing factors including the “free” IOPS/Throughput that accompanies Premium SSD v2 and Ultra Disk.

Using the gathered source volume statistics, destination host metadata, and user configured parameters, Cirrus Data Cloud’s recommendation service then generates a set of target values for each of the recommendation profiles that serve as a minimum when used for matching. The target values are then checked against every disk type and tier available to the selected destination host. If the disk type and tier can accommodate the target values, the recommendation service will then calculate the real performance values and the corresponding price.

For example, (only using Premium SSDs in this example), if the target values determined were 80 GiB, 400 IOPs, 120 MB/sec., the P6 Tier would not qualify as its maximum IOPS and throughput are too low. Instead, P10 and above would be considered and CDC would determine that P10 would be the cheapest option. Therefore, the finalized values used for recommendation would be the minimum disk capacity for P10 and the IOPs/throughput of P10. Even though the original source volume was only 80 GiB, the recommendation service would recommend a 128 GiB capacity in order to achieve the needed performance.

When recommending the correct non-scaling disks, Cirrus Data Cloud’s recommendation service will simply check the performance bounds of every disk tier and find the lowest cost option.

Scaling disks, however, are inherently much more sophisticated, since capacity, IOPS and throughput can all be tweaked individually and each has a cost associated with it. In order to calculate the optimal parameters and price for the recommendation, the recommendation service works backwards. CDC understands that users get up to X IOPS per GiB allocated and up to Y throughput per IOP. Starting with the target value for throughput, CDC determines if the target IOPs is enough to satisfy the desired throughput. If not, it adjusts the final recommended IOPS and capacity and re-calculates. Once the recommendation is determined, it uses the pricing information to calculate the price of the real recommendation including free throughput/IOPS.

Each time the service performs a recommendation, it always returns the lowest price configuration that satisfies the recommendation target values. However, since different target values are generated based on the recommendation profile selected by the user, recommendations may include those more or less performance-oriented than needed, or those with costs too high for particular users.

If the result of the recommendation service is not desirable, users can tweak settings such as expected growth and the recommendation profile in order to get a more appropriate recommendation. Users may also override the recommendation and manually select their own disk parameters. In these cases, Cirrus Data Cloud's Storage Consumption Planner will still perform a cost estimation based on the manually entered parameters.

Once the recommendation is finalized and accepted, Cirrus Data Cloud will automatically allocate the volumes, map them to the destination volumes, and create migration sessions that have a minimum impact on the production workload.

ABOUT CIRRUS DATA

Cirrus Data Solutions Inc. (CDS) is a leading technology provider of next-generation solutions for data migration, data acceleration, and data protection. The company distributes its solutions through systems integrators, managed service providers, channel resellers, and partners including HPE, IBM, Dell/EMC, Insight, Pure Storage, Infinidat, Align, SHI, ePlus, Computacenter, Mainline, Sirius, WWT and many others. CDS is headquartered in Syosset, New York and has offices located in Dublin, Ireland, and Nanjing with sales and support offices in Boston and Chicago.

For more information, visit CDS online at www.cirrusdata.com.

Cirrus Data USA Headquarters
6800 Jericho Turnpike Suite 213W
Syosset, NY 11791

EMEA Headquarters
Unit 19, The Exchange - Calmount Park Calmount
Avenue, Dublin D12 F5X0
IRELAND