

DELL EMC ISILON CLOUDPOOLS

Extend the Data Lake to the Cloud with Transparent Tiering of Inactive Data

ESSENTIALS

- Seamless expansion to the cloud by enabling a cold or frozen data storage tier
- Flexible choice of public or private cloud options
- Transparent to end users and applications
- Enterprise grade encryption and compression
- Leverages Isilon SmartPools automated tiering policy engine
- Enables economic choice of CAPEX- or OPEX-based cloud storage models
- Proxy support for increased security

Data is growing astronomically and that the majority of the growth comes from unstructured data. According to the EMC® sponsored IDC study on the [Digital Universe¹](#), the amount of unstructured data in 2020 is expected to be around 44 ZB. The Dell EMC Isilon® data lake is the key enabler to address the growing challenge to store, manage, protect, and analyze all this data of an enterprise.

Businesses are looking to leverage the cloud to cut costs, simplify IT management, or gain virtually limitless storage capacity. Dell EMC Isilon® CloudPools® software provides policy-based automated tiering that lets you seamlessly integrate with the cloud as an additional storage tier for the Isilon cluster at your data center. This allows you to address rapid data growth and optimize data center storage resources by using the cloud as a highly economical storage tier with massive storage capacity for "cold" or "frozen" data that is rarely accessed. In this way, your more valuable on-premise storage resources may be used for more active data and applications, while frozen data may be retained at minimum cost and administrative burden for compliance, historical, or other business reasons. CloudPools enables you to optimize storage resources and gain cloud scale storage capacity while reducing overall storage costs.

FLEXIBLE CHOICE

With CloudPools, you have a flexible choice of tiering data to public, private or hybrid cloud options. You can select from a number of public cloud services, including Dell EMC Virtustream, Amazon Web Services (AWS) S3, or Microsoft Azure. Or, you can use a private cloud that is based on Dell EMC Elastic Cloud Storage™ (ECS) or Isilon. It is possible to integrate with more than one public or private cloud provider at the same time, thus providing maximum flexibility. For example, you may choose to tier marketing and project data to a public cloud provider but choose to tier HR data to a private on-premise option because of confidentiality or regulatory reasons.

SIMPLIFIED SETUP AND MANAGEMENT

CloudPools is simple to deploy and manage because it uses the same flexible and powerful policy engine that Isilon SmartPools® uses. CloudPools can be deployed in minutes and is easy for the administrators to set up, configure, and manage. While SmartPools software allows you to optimally place data across different storage tiers within an Isilon cluster, CloudPools software enables tiering of inactive data to the cloud based on the business rules and policies that you specify. SmartPools can seamlessly tier data from all-flash, hybrid or archive tiers. You can use SmartPools and CloudPools together to optimally place data within your Isilon cluster or the cloud. For example, SmartPools may be used to tier "warm" data to an archive tier within your Isilon cluster, while CloudPools may be used to tier "frozen" data to the cloud.

¹ <http://www.emc.com/leadership/digital-universe/2014iview/index.htm>

With CloudPools and SmartPools, you may define policies that identify the data to be tiered, the criteria for tiering, and the choice of the public or private cloud target. By default, SmartPools jobs run nightly to apply new policies to the selected data, and seamlessly tiers the files to the appropriate location based on your requirements.

SmartPools policies are dynamic, flexible and scalable which provides administrators with granular control of the data placement to meet the business objectives of your business. A policy can be based on any combination of file metadata attributes such as timestamps, file name or type, or file size. SmartPools and CloudPools allow you to efficiently manage the lifecycle of the file data on your Isilon storage without the need to deploy any external gateway or intermediary software.



SEAMLESS AND TRANSPARENT TO USERS

The use of CloudPools is transparent to end users and applications. When a file is tiered to the cloud with CloudPools, the file is replaced by a SmartLink or “stub” that contains the maps that point to the contents in the cloud. A user or application cannot tell if the data is on premise or has been tiered to the cloud. Users access the data the same way as before without changes to policies and procedures—you don’t have to do anything different to access the data. If you access a tiered file using your Windows or Linux desktop, only the relevant blocks are retrieved and delivered to you without the need to retrieve the entire file from the cloud. When you modify a tiered file, only the relevant portions of the file are rewritten to the cloud, thereby optimizing the network bandwidth usage.

CLOUD-ENABLED DATA LAKE

With a typical on-premise approach, an Isilon data lake can store and manage Petabytes of unstructured data in a data center environment. With CloudPools, you can expand your data lake with virtually unlimited cloud-scale storage capacity. By extending your data lake from the data center to the cloud with CloudPools, you can leverage the cloud as an additional, low-cost storage tier. All the metadata for a tiered

file, including its security and file attributes, is still stored as a SmartLink on the primary Isilon cluster, but the actual data is stored on the cloud. Each SmartLink that points to a tiered file is approximately 8K in size. From a namespace and data access perspective, the Isilon cluster is still the front-end view to all the data accessible to a user. Using CloudPools in this way allows you to lower data center requirements and associated costs for floor space, power, cooling, and rack space.

SECURITY IN THE CLOUD

To secure data that is tiered to the cloud, CloudPools allows you to encrypt data that is transmitted from your on-premise Isilon cluster to the cloud storage service. The data that is on the cluster is shared into chunks and sent to the cloud, allowing it to be protected.

OneFS also provides proxy support for CloudPools which increases security by removing direct external network exposure of a cluster. It avoids the need for crafting complex firewall traversal rules while allowing multiple Isilon nodes to simultaneously tier to the cloud provider of choice.

OPTIMIZED PERFORMANCE

With CloudPools, you can optimize network performance by compressing data before sending it over the wire to the cloud. This reduces the network bandwidth used and allows for faster upload of data to the cloud. Optimizing the network bandwidth also reduces the download traffic and provides faster response to the end users requests for data.

SMART ECONOMICS

The choice of cloud as a viable solution for storing frozen data is often driven by economics. Organizations may optimize storage resources and lower overall costs with a flexible choice of CAPEX and OPEX spending models that are enabled with cloud integration. For example, for your growing data storage needs, you can choose to expand storage capacity in the data center with a CAPEX approach or leverage CloudPools and public cloud storage options with an OPEX-based cost model. This allows you to gain new efficiency by embracing the cloud as an archiving storage tier for cold or frozen data while optimizing your on-premise Isilon cluster for more valuable data. With Isilon CloudPools, your data stored in the cloud remains within your data lake and is readily and transparently accessible.

SMARTPOOLS AND CLOUDPOOLS

SmartPools uses a policy-based, automated tiering approach to optimize storage resources within a single Isilon cluster and ensure that a given data set is supported with the right storage performance tier. With a similar policy-based, automated tiering approach, CloudPools seamlessly integrates the Isilon on-premise cluster with cloud-based storage. The same policy engine and framework is used in the setup and management of both CloudPools and SmartPools.

INTEGRATION WITH ONEFS SOFTWARE

CloudPools is compatible with Isilon OneFS® 8.0 and later versions. Just as CloudPools seamlessly integrates with SmartPools tiering capabilities, CloudPools also seamlessly integrates with other OneFS software modules. CloudPools integrates with Isilon SmartQuota® in that a tiered file is only charged for the space used by the SmartLink file. For Disaster Recovery purposes, CloudPools seamlessly integrates with data replication using Isilon SyncIQ®. SyncIQ is SmartLink aware and will replicate the SmartLink file to the destination cluster. During a failover scenario, the target cluster is

connected to the cloud and the users will have seamless access to on-premise and tiered files. CloudPools is also compatible with NDMP-based backup solutions that will back up the SmartLinks without recalling the files.

SOFTWARE LICENSES

CloudPools tiering software is licensed on a capacity basis depending on the choice of cloud providers. CloudPools also requires a SmartPools license to be enabled on your Isilon cluster.

By transparently tiering frozen data, CloudPools software enables you to extend your data lake by optimizing storage resources and gain cloud scale storage capacity while reducing overall storage costs.

TAKE THE NEXT STEP

Contact your Dell EMC sales representative or authorized reseller to learn more about how Isilon scale-out NAS storage solutions can benefit your organization.

[Shop Dell EMC Isilon](#) to compare features and get more information.



Dell, EMC and other trademarks are trademarks of Dell Inc. or its subsidiaries.

© Copyright 2017 Dell Inc. All rights reserved. Published in the USA. 4/17 Data Sheet H14758.1

Dell EMC believes the information in this document is accurate as of its publication date. The information is subject to change without notice.