# KCCD ENTERPRISE DATA BACKUP SOLUTION

Version 2.0

Update: 6/2/2023

Table of contents:

## OVERVIEW

The KCCD Enterprise backup system currently backs up the following resources across the district.

1. Physical servers
2. Virtual Machines and Virtual Appliances
3. AWS Cloud EC2 Instances(Servers and Virtual Appliances)
4. CIFS Shares
5. SAN Volumes

KCCD enterprise data backup system consists of the following components:

1. Veritas NetBackup Enterprise Data Management Solution
2. NetApp SAN units on campuses
3. AWS Cloud Backup as a service (BaaS)
4. AWS Cloud native backup tools

### *Basic Architecture*

(See diagram below)

The heart of the KCCD enterprise data backup system is the coordinated workings of Netbackup Enterprise Data Management Solution, NetApp SAN, and AWS cloud storage.  Netbackup manages most of the backup operations while data is stored on the SAN and in the AWS Cloud.

The NetApp SAN units, one on each campus and the DO(see note below regarding exception), also have a configured backup feature and process of their own(storage snapshots), which backs up Virtual Machines and other resources which are built on their data volumes. This servs as a second data backup method and resource for data recoveries.

(Note: BC does not have a SAN yet but a unit is scheduled to be installed in June, 2023).

The general concept of the KCCD's backup system design is that data from each of KCCD's three colleges and the DO(VMs, File Server contents, Network Shares, etc.) is initially backed up to the local disk of the media server residing in the Data Center of the District Office. This disk repository serves as a temporary cache before the data is then sent to the AWS cloud S3 buckets for storage. The amount of time the data is stored in the cloud and type of S3 bucket storage used depends on type of data and from which campus the data came from. For example, VM backups from CC and BC are currently stored in faster class of S3 storage for faster retrievals since these two campues do not have local SAN snapshots yet for fast VM restores. On the other hand, data from DO and PC are stored in slower class of S3 buckets because these locations have readily available SAN snapshots for faster restores.

[redacted]

## SYSTEMS, SOFTWARE, AND SERVICES

### *Netbackup Servers*

The KCCD Enterprise Data Backup system employs two servers, which are VMware Virtual Machines(VM).

[redacted]

*Netbackup Software*

Netbackup version 9.1

*Storage Area Network(SAN)*

There is at least one SAN unit each at the Cerro Coso College(CC), Porterville College(PC), and at the DO. A SAN unit is being installed at Bakersfield College and should be in place by the end of June, 2023.

[redacted]

*BaaS Service (Backup as a Service)*

AWS 3 Storage Buckets

**AWS Standard IA Storage**

The standard IA(Infrequent Access) S3 storage offer s slower speed than standard S3 storage  for data retrieval but 6 times faster than glacier class storage.   This is used for backing up CC and BC Virtual Machines which do not have restore from SAN snapshot recovery capability yet.

**AWS Glacier Storage**

The Glacier storage is used for data that doesn't require fast recovery such as Virtual Machine backups for DO and PC, and CIFS data, which have SAN snapshots to recovery quickly from.

*AWS EC2 EBS Volume Snapshots*

EBS Volume Snapshots are used to back up EC2 instances(VMs) in the AWS cloud.  Each EC2 instance has one or more storage volumes(EBS volumes)  associated with it.  These volumes are backed up to back up the entire EC2 instance.

## *Data Security*

Data in transit sent to AWS S3 through Netbackup is fully encrypted and is protected once there by S3's Object Lock, Write Once Read Many(WORM) feature, which is activated on KCCD's backup buckets.  This feature offers object level immutable data protection.  More details of Object Lock protection can be found at the following internet link:

https://aws.amazon.com/s3/features/object-lock/

## BACKUP AND RESTORE OPERATIONS

Depending on the type of data and the campus where it is located, the backup process for that data takes a unique path and process.  The following is backups by location and data source:

*Bakersfield College(BC)*

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

*Porterville College(PC)*

[redacted]

[redacted]

*District Office(DO)*

[redacted]

[redacted]

*AMAZON AWS EC2*

The EC2 instances in each of the KCCD accounts in the AWS cloud are backed up by AWS's EBS Volume Snapshot function. The function is executed by custom Lambda function scripts which choose the backup frequency and retention time on each EC2 instance based on Tags assigned to the instance.

The EC2 snapshots are stored in AWS Compute's storage, to which customers do not have direct access.

### *Restorations*

VMs are restored either from the SAN or Netbackup, depending on which resource is available. [redacted]

To restore a VM using Netbackup, a restore job from the Netbackup Management Console is initiated after selecting date and time to restore from. The restore job will pull data back from either the "do-dedupe-bucket" or "bc-cc-dedupe-bucket" S3 bucket, depending on where the VM backup was saved. Once the restore job is complete, a new restored server will appear in VMware.

Restoring individual files from backup works much the same way using Netbackup. The date from which data is desired to be restored from is selected and then the restore job initiated. A choice is given whether to replace existing data in the folder with the restored data or to place the restored data in a new folder.

If restoring VMs or files from SAN snapshots, the restore process is to create a temporary SAN volume from a snapshot of the volume which contains the VM or file, and based on backup date of choice. Once the temporary volume is created, the volume is presented to the VMware host or file server. The VM files or data files are then retried to re-establish the VM or to restore the file.

SIFS restores can be accomplished from Netbackup or directly from the SAN using SAN snapshots, just like with restoring VMs and individual non-CIFS files. Each CIFS directory has a snapshot on the SAN to which it can be restored from.

EC2 instance restores are done by restoring the EBS volume(s) associated with each instance. The EBS volumes are restored using the native storage tools available in AWS. For example, a server's system volume can be restored directly from any saved snapshot of it by choosing to "Restore Root Volume" in AWS Management Console. This will effectively, restore then entire server to the chosen date and time.

## BACKUP SCHEDULES AND RETENTION TIMES

Below is the schedule for backups and their retention times for SAN volumes and VMs, which includes local databases installed on those VMs.

| RESOURCE | TYPE | BACKUP FREQUENCY | LOCATION | RETENTION |
|---|---|---|---|---|
| | | | | |
| BC, CC Virtual Machines | [redacted] | Inc: Daily Full: Every other Friday | [redacted] | 105 days |
| PC, DO Virtual Machines | [redacted] | Inc: Daily Full: Every other Friday | [redacted] | 105 days |
| PC, DO Virtual Machines | [redacted] | Daily | [redacted] | 1 month |

| | | | | |
|---|---|---|---|---|
| BC, CC File Servers | [redacted] | Inc: Daily Full: Every other Friday | [redacted] | 105 days |
| CIFS | [redacted] | Inc: None Full: Every Tu 4th Wk | [redacted] | 105 days |
| CIFS | [redacted] | Daily | [redacted] | 1 month |
| EC2 Snapshots | [redacted] | Daily/Weekly/Monthly(Individual Schedule) | [redacted] | 3 months |
| CC SAN Volumes | [redacted] | Daily | [redacted] | 1 year |
| DO SAN Volumes | [redacted] | Daily | [redacted] | 1 month |
| PC SAN Volumes | [redacted] | Daily | [redacted] | 1 month |

Below is the list of backups for specific services:

| **SERVICE** | **TYPE** | **BACKUP FREQENCY** | **LOCATION** | **RETENTION** |
|---|---|---|---|---|
| | | | | |
| do-swd (Voic mail) | [redacted] | Daily | [redacted] | 105 days |
| do-symetry (door lock) | [redacted] | Daily | [redacted] | 105 days |
| Aruba | [redacted] | Daily | [redacted] | 105 days |
| Clearpass | [redacted] | Daily | [redacted] | 105 days |
| Airwave | [redacted] | Daily | [redacted] | 105 days |
| Pulse Secure VPN | [redacted] | Daily | [redacted] | 105 days |
| APC Struxureware | [redacted] | Daily | [redacted] | 105 days |
| | | | | |

## FINAL NOTES

The KCCD Enterprise Data Backup system is a work in progress, as all IT systems are, as new Software and hardware features and technologies continually become available. In addition, the system is constantly being expanded to meet its maximum potential, such as the installation of a SAN storage at BC to expand SAN backups and enabling of SAN snapshots at CC for faster data restores.

In the AWS Cloud, the use of Netbackup's CloudPoint on EC2 instances is being investigated as a way to bring the EC2 instances into the Netbackup system. This would allow KCCD's cloud VMs to be backed up both by the AWS native EBS Volume snapshots and by Netbackup.

As the design and implementation of the KCCD Enterprise Data Backup System changes, this document will be updated accordingly.