Using Amazon S3 for storing binary data

Christopher Nadeau - 2017-11-09 - Comments (0) - Deskpro Legacy

Deskpro allows storing binary data (such as file attachemnts, emails, and images) in Amazon S3. This has two primary advantages:

1) Amazon S3 is highly available and highly scalable. This takes the burden off of you to manage servers.

2) For helpdesks that use multiple web servers (for example, with use with a node balancer), S3 is the perfect way to store blobs because each server can be "dumb". Without S3 you would need to use the database to store blobs (which is inefficient), or use a network file system (which is just another thing to worry about).

Configuring Amazon S3

To use Amazon S3, you must register with Amazon Web Services: http://aws.amazon.com/s3/

Getting your Access Key and Secret Key

1) Once registered, go to the "AWS Management Console" (available from the menu in the top-right corner of the screen).

2) In the top right corner you will see a menu with your name. In this menu, click "Security Credentials".

3) Click the [+] icon next to the "Access Keys" section. Click "Create New Access Key" and copy the Access Key and Secret Key. It's very important to copy the Secret Key now, you will not be able to retrieve it once you close the modal window.

Note: You may wish to create a separate user with specific permissions instead of using your main AWS user (which has full AWS permission to do everything). Refer to Amazon's Best Practices guide for more information.

Creating a bucket

1) Still in the AWS Management Console, open the "Services" menu in the top left corner of the screen and click on "S3".

2) Click on "Create Bucket" and choose a bucket name (for compatibility with DNS, you should use only lower-case letters, numbers and dashes - see the Naming Buckets and Keys section here for details). You should choose a location that is closest to your actual web

server.

3) Select your new bucket in the list and make sure the [Properties] tab in the right pane is selected. Expand the [Permissions] section and click "Edit Bucket Policy" and copy+paste the following policy into the box:

```
{
	"Version": "2008-10-17",
	"Statement": [
		{
			"Sid": "AllowPublicRead",
			"Effect": "Allow",
			"Principal": {
				"AWS": "*"
			},
			"Action": [
				"s3:GetObject"
			],
			"Resource": [
				"arn:aws:s3:::YOUR_BUCKET_NAME/*"
			]
		}
	]
}
```

(Replace "YOUR_BUCKET_NAME" with the bucket you just created).

This policy makes the bucket public so Deskpro can serve blobs directly from Amazon S3 servers.

Enable the Amazon S3 Storage Mechanism

To enable Amazon S3 storage, go to Admin > Server > File Uploads. Click the "Storage Mechanism Options" button and switch the mechanism to Amazon S3. Enter your S3 details you configured above and click the "Save" button.

Move existing blobs to S3

Deskpro will begin to automatically move existing blobs to S3. This is done slowly in small batches every minute as part of the default scheduled task (the same one that is responsible for processing email).

This can be rather slow, especially if you have many blobs. If you would like to speed up the process you can use the dedicated command-line tool:

```
$ cd /path/to/deskpro
$ php cmd.php dp:move-blobs    # Shows an overview
```

```
$ php cmd.php dp:move-blobs --run    # Actually starts moving blobs
```

FAQ

Q: What happens if S3 goes down?

A: If S3 goes down, links to existing blobs will not work. For example, if there was an image thumbnail saved to S3, that image would not load in the browser. However, the helpdesk as a whole will remain fully operational. Deskpro is designed to "fallback" to database storage of attachments if S3 uploads fail.

Q: Can I move my files off of S3 later?

A: Yes. You can re-enable filesystem or database storage at any time form the same page in the Admin Interface. When you change your helpdesk's storage mechanism, your existing files are gradually copied over. This happens automatically and your helpdesk can still run while it's happening.

Q: Isn't it insecure to have every file public on S3?

A: No. The URL for each file on S3 includes a very long string of random and unguessable characters that essentially acts as a password. Anyone with the full URL will be able to download the file, but to actually get the URL you need to be an authorised user. E.g., an authorised user would need to copy the URL and give it to someone else. (Which is no different than an authorised user just downloading the file manually and sending it to someone else).