Progresswiz

Sharing Static Data with Amazon S3

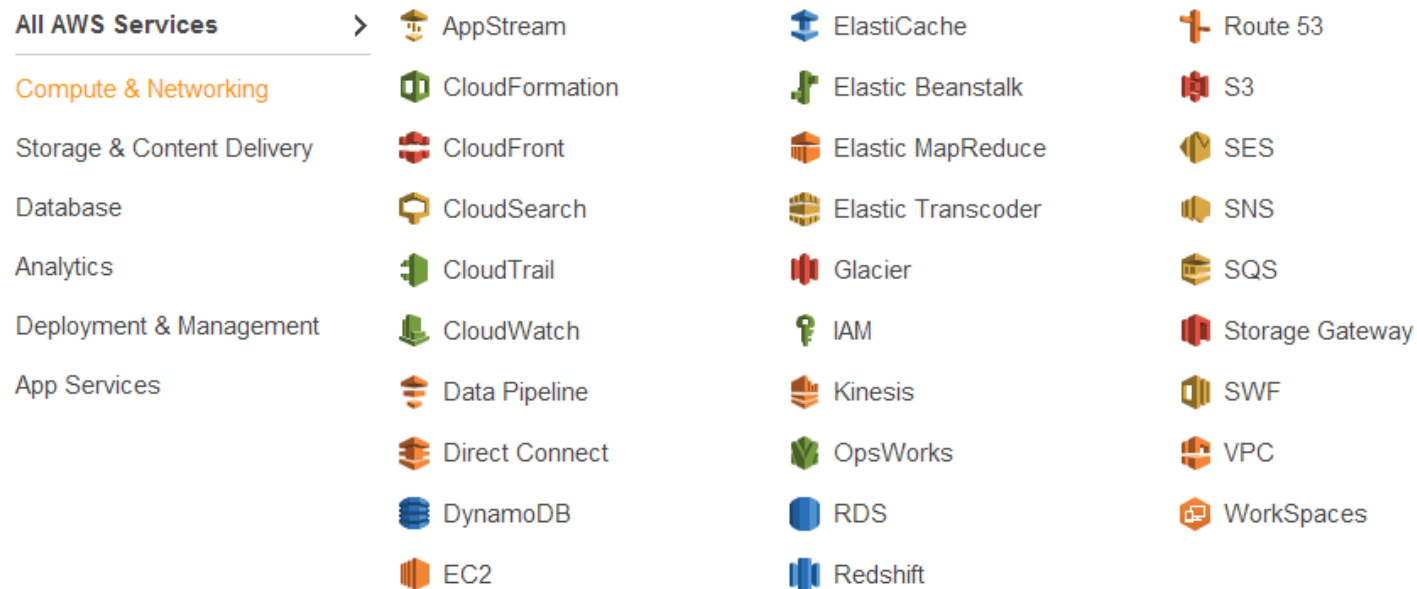**PAUL KOUFALIS**
PRESIDENT
PROGRESSWIZ CONSULTING

Agenda

- What is this Amazon Web Services?

- How Do I Get Started?

- Can It Do Magic Tricks?

- My Boss is Cheap – What Does it Cost?

- I'm Bored Now – How Do I Use It?

- Are We Done Yet? I Have a Question

What Is Amazon AWS?

- Amazon currently provides 29 infrastructure services over the web
- The most popular are EC2 and S3

All AWS Services >	AppStream	ElastiCache	Route 53
Compute & Networking	CloudFormation	Elastic Beanstalk	S3
Storage & Content Delivery	CloudFront	Elastic MapReduce	SES
Database	CloudSearch	Elastic Transcoder	SNS
Analytics	CloudTrail	Glacier	SQS
Deployment & Management	CloudWatch	IAM	Storage Gateway
App Services	Data Pipeline	Kinesis	SWF
	Direct Connect	OpsWorks	VPC
	DynamoDB	RDS	WorkSpaces
	EC2	Redshift	

Amazon EC2

- EC2 is that whole "deployment in the cloud" thing

- You run Windows and Linux Servers on Amazon's infrastructure
  - All over the world

- Download *Getting Started in the Cloud* from http://progresswiz.com/downloads

What is Amazon S3?

- Simple Storage Service (Somewhere)

What is Amazon S3?

- It's just file system space with a few bells and whistles

- Files can be accessed using a standard web URL:

http://pugchallenge.s3.amazonaws.com/talisker.jpeg

What is Amazon S3?

- You can share files with whomever you want
  - Public
  - Private
  - Somewhere in between

Progresswiz

How Do I Get Started

- First you need to open an account

- http://aws.amazon.com
  - Have your credit card ready

- For the first year, some of the stuff is free
  See http://aws.amazon.com/free/

Progresswiz

Now What?

- **How about a little security?**
  - Users and groups?

Users and Groups

Group Security

Group Security

- The security policy is for all AWS services
- Ex.: The RW group can do everything EXCEPT manage users:

```
{"Statement":
  [{
      "Effect": "Allow",
      "NotAction": "iam:*",
      "Resource": "*"
  }]
}
```

- IAM = Identity and Access Management

Buckets

- S3 data is stored in buckets

- http://s3.amazonaws.com/***bucket***/...
   OR

- http://***bucket***.s3.amazonaws.com

- For the 2nd format to work must follow DNS-compliant naming rules
  - All lowercase

Buckets

Sharing Static Data with Amazon S3

Progresswiz
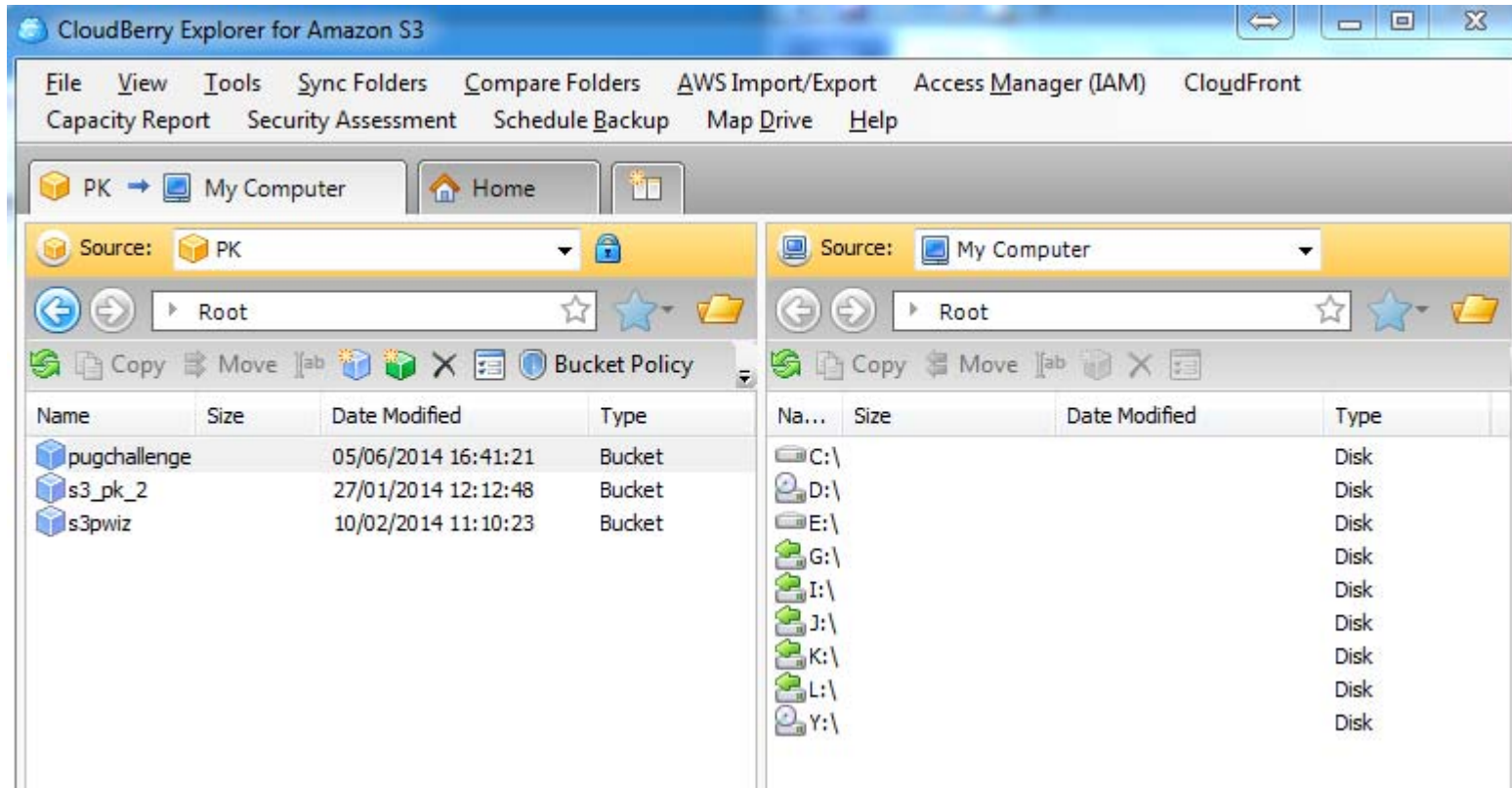
Quick Recap

- Open an Amazon account

- Create a couple of users and groups

- Create your buckets

- Install CloudBerry Explorer for S3

- Upload your files

- Share them

- Total Effort: 30 minutes

Very Nice. What Else?

- **S3 provides all kinds of cool features:**
    - You can pay…or have the downloader pay
    - Peer-to-peer access (BitTorrent)
    - File versioning
    - Static Web Sites
    - Automatic life cycle management

That Was Cool! Keep Talking…

- **Pre-signed URLs**
- **Try** http://pugchallenge.s3.amazonaws.com/talisker.jpeg
  - FAIL: Access Denied
- **I can generate a URL with an embedded access key**

- **I can expire that key**
  - The URL is only valid for a limited time

Progresswiz

Nothing in Life is Free

- All pricing is per month

- Storage: $0.03/Gb for first TB
  - $0.01 for Glacier (infrequent access: backups)
  - Subsequent TB are cheaper
- Requests (PUT, GET…): $0.005 per 1000
- Data Transfer In: Free
- Data Transfer Out: 1 Gb Free
  - $0.12/Gb for first TB

Example Pricing

- 100 Gb of stuff $3.00/month
- 50 Gb of downloads $6.00/month
- 250K requests $1.25/month

- TOTAL: $10.25/month

Integrating S3

- **Unfortunately there is no AWS SDK for OpenEdge**
  - Maybe one day…
- **Existing SDKs:**
  - PHP, .NET, Java, JavaScript, Python, Ruby

- **Luckily someone did the heavy lifting for us**

Progresswiz

s3cmd

- http://s3tools.org/s3cmd

*S3cmd is a **free command line tool** and client for uploading, retrieving and managing data in Amazon S3 and other cloud storage service providers that use the S3 protocol, such as Google Cloud Storage.*

s3cmd

- Written in Python
- Python available for most platforms
  - Windows, Mac, Linux
  - Also HPUX, Solaris, AIX and others

- Supports most (all?) of the SDK functions
  - s3cmd –help returns 256 lines!

Installing s3cmd

- **Install Python**
  - https://www.python.org/downloads/
- **Install s3cmd**
  - http://s3tools.org/download

- **On Windows, create s3cmd.bat**
```
@echo off
PATH=C:\Apps\Python27;%PATH%
python c:\apps\s3cmd\s3cmd %* 2>NUL
```

Progresswiz

Configure s3cmd

- If you don't want to always pass usernames and passwords on cmd line

- s3cmd --configure

- Creates s3cmd.ini in $HOME
  - /home/paul/.s3cfg
  - C:\Users\Paul\Application Data\s3cmd.ini

Progresswiz

Using s3cmd

- I'm warning you: this is easy
  - Really

- Wait…let me show you how to integrate with the OpenEdge ABL

```
OS-COMMAND SILENT s3cmd …
```

Using s3cmd

- ## Make a bucket

```
$ s3cmd  mb s3://pugch_1
Bucket 's3://pugch_1/' created
```

- ## CAREFUL: Bucket names are shared by everyone
  - I.e. *bucket*.s3.amazonaws.com
  - Don't be surprised if bucket already exists

Progresswiz

Using s3cmd

- ## List your buckets

```
$ s3cmd ls
2014-06-05 20:41  s3://pugchallenge
2014-01-27 17:12  s3://s3_pk_2
2014-02-10 16:10  s3://s3pwiz
2014-06-06 19:52  s3://pugch_1
```

Progresswiz

Using s3cmd

- ## Upload a file to a bucket

```
$ s3cmd put toto.txt s3://pugch_1
toto.txt -> s3://pugch_1/toto.txt  [1 of 1]
 29 of 29   100% in    0s   269.81 B/s  done
```

- ## List it:

```
$ s3cmd ls s3://pugch_1
2014-06-06 19:54        29   s3://pugch_1/toto.txt
```

- ## Download:

```
$ s3cmd get s3://pugch_1/toto.txt toto2
s3://pugch_1/toto.txt -> toto2  [1 of 1]
 29 of 29   100% in    0s   570.09 B/s  done
```

What about from a Web Browser?

- http://pugch_1.s3.amazonaws.com/toto.txt

```
<Error>
<Code>AccessDenied</Code>
<Message>Access Denied</Message>
<RequestId>4656EB4745639875</RequestId>
<HostId>
SHIOL/sCMuQ5k0UYBBxd4aG3EwZrpP1tktKlGI2UXzE
</HostId>
</Error>
```

Progresswiz

Pre-Signed URL

- s3cmd signurl <url> <expiry_epoch>

- Epoch time = seconds since 1970/01/01
- 2014/06/11 18:00:00 EDT = 1402524000

```
$ s3cmd signurl s3://pugch_1/toto.txt 1402524000
```

```
http://pugch_1.s3.amazonaws.com/toto.txt?AWSAcces
sKeyId=AKIAIJEH6H5UBWEBZWKA&Expires=1402524000&Si
gnature=cV%2BPEO4e9W%2FF0QZi5Izzh1sf6sU%3D
```

Progresswiz

Pre-Signed URL

- ## What happens after expiry?

```
<Error>
<Code>AccessDenied</Code>
<Message>Request has expired</Message>
<RequestId>AE15D0420933B946</RequestId>
<Expires>2014-06-06T20:21:48Z</Expires>
<HostId>
IHGdBBU1r3v/HhYe6ozirCbbeTCevpzgjKDpjDGn4mqyW
</HostId>
<ServerTime>2014-06-06T20:21:53Z</ServerTime>
</Error>
```

Progresswiz

File Info

```
$ s3cmd info s3://pugch_1/toto.txt
```

```
s3://pugch_1/toto.txt (object):
   File size: 109
   Last mod:  Fri, 06 Jun 2014 20:20:15 GMT
   MIME type: text/plain
   MD5 sum:   cc11bc75234e59df7f33cfbadd1316ad
   SSE:       NONE
   policy: none
   ACL:       pkoufalis: FULL_CONTROL
```

Disk Usage

```
$ s3cmd du

138      s3://pugch_1/
31399    s3://pugchallenge/
4635280  s3://s3_pk_2/
525373   s3://s3pwiz/
--------
5192190  Total
```

Progresswiz

What Else?

- Synchronize a local dir and bucket dir

- Enable access logging

- Upload large files in multiple parts

- And of course the usual stuff
  - Move
  - Delete
  - Copy

Are the Gears Turning?

- I hope this presentation has you thinking about possibilities

  - The advantages are HUGE
  - The ROI is inarguable

Questions?

Sharing Static Data with Amazon S3

Progresswiz

Progresswiz Consulting

- Based in Montréal, Québec, Canada

- Providing technical consulting in Progress®, UNIX, Windows, MFG/PRO and more

- Specialized in performance tuning, system availability and business continuity planning

- …and security of Progress-based systems

Progresswiz