Providing Secure Access to OAuth-based Services in HTCondor Jobs

Jason Patton
Center for High Throughput Computing

Motivation

- HTCondor's file transfer mechanism allows for direct HTTP(S) downloads from public URLs
  - `transfer_input_files = https://wisc.edu/~jpatton/input.data`
  - Good: Users don't have to store input on the submit host
  - Bad: Users need someone to run a webserver
  - Ugly: Users' data are available to anyone with the URL
- How can HTCondor provide secure downloads from and uploads to file storage services that require credentials?

Motivating Example

- Suppose the pool admin has configured HTCondor to get tokens from some service named “cloudstore,” and a “cloudstore” file transfer plugin exists to download and upload files from that service:

```
executable = do_science.sh
arguments = input.data output.data

use_oauth_services = cloudstore
transfer_input_files = cloudstore://input.data
transfer_output_remaps = “output.data = cloudstore://output.data”

queue
```

Solution: Credd and Credmon Architecture

1. OAuth Credmon Webserver
   - WSGI application that places OAuth tokens in secure storage after sending users through an OAuth 2.0 authorization flow
2. OAuth Credmon
   - HTCondor daemon that monitors and refreshes (as needed) the OAuth tokens in secure storage
3. Credd
   - HTCondor daemon that fetches credentials (e.g. OAuth tokens) from secure storage and pushes them to job sandboxes

Point of Clarification

- In this presentation, “OAuth tokens” refer to the credentials that are gathered via an OAuth 2.0 authorization flow and that are used as bearer tokens when communicating with protected resources.

CONFIGURING SUBMIT HOST

Configuring HTCondor Submit Host

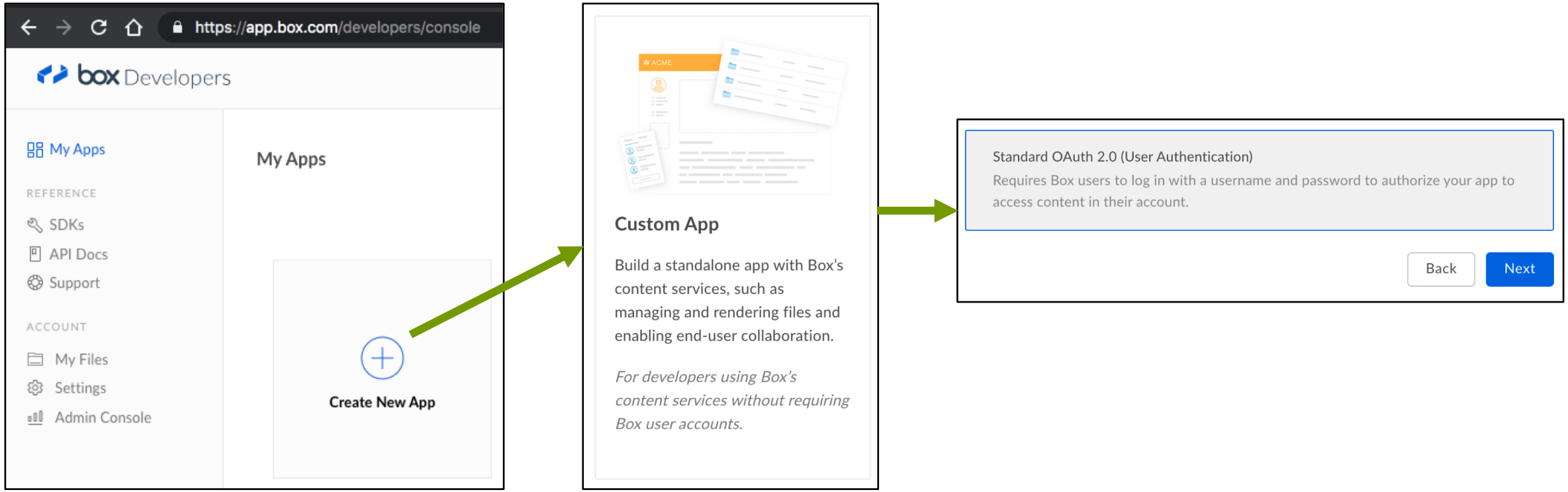
- Assumptions:
  - RHEL 7-based OS
  - HTCondor 8.9.9 or later
  - HTCondor pool configured to use authentication and encryption
  - Apache configured with SSL/TLS certificate (LetsEncrypt works)

Configuring HTCondor Submit Host

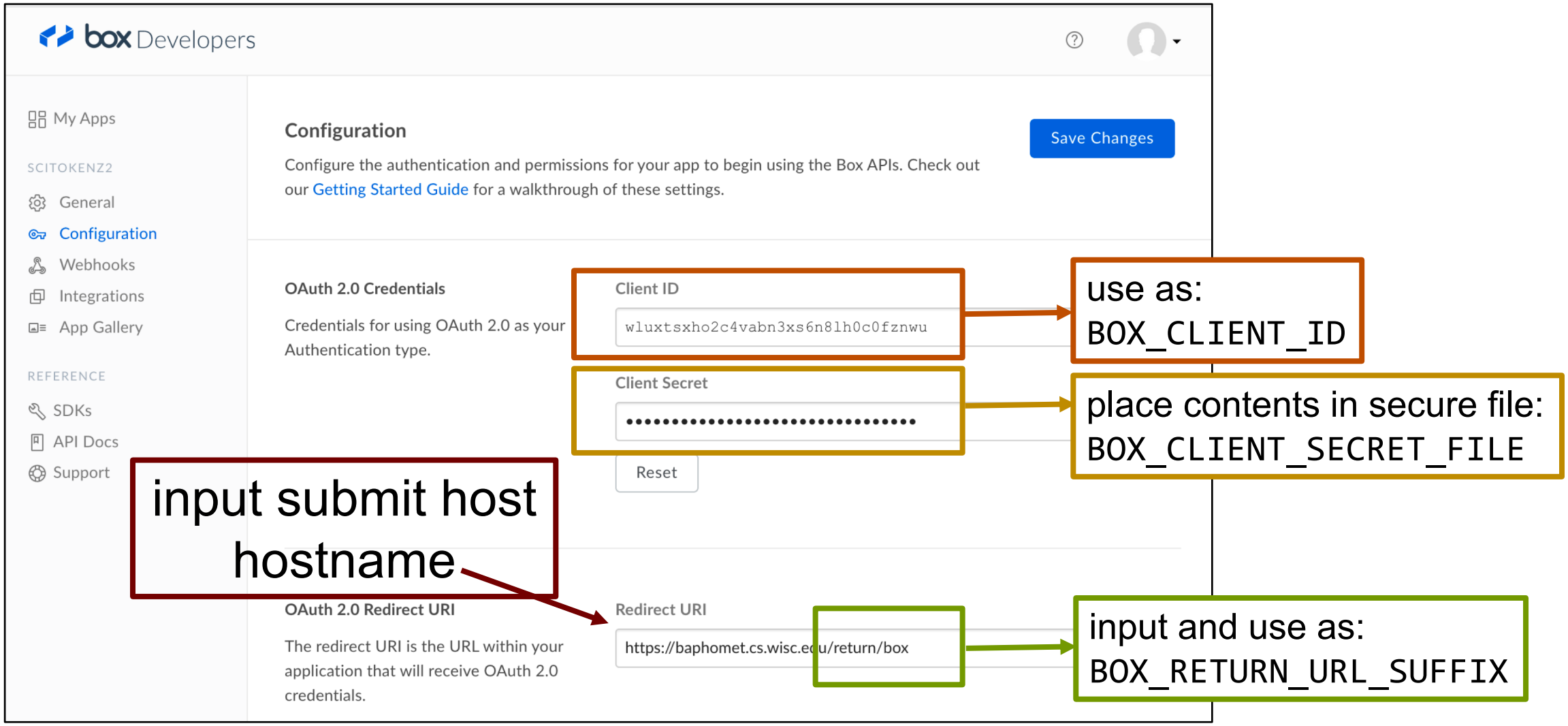
- Step 1: Install and enable the OAuth Credmon and OAuth Credmon Webserver:
  - `yum install condor-credmon-oauth`
    - Drops `/etc/condor/config.d/40-oauth-credmon.conf`, which contains the line: `use feature : OAUTH`
  - `cp /usr/share/doc/condor-credmon-oauth-8.9.9/condor_credmon_oauth/config/apache/condor_credmon_oauth.conf /etc/httpd/conf.d/condor_credmon_oauth.conf`
    - Adjust `condor_credmon_oauth.conf` as needed, installs the WSGI application under the webserver root by default!
  - Restart `condor` and `httpd`

Configuring HTCondor Submit Host

- Step 2: Set up an OAuth API client with desired service
- Example: Box.com (https://app.box.com/developers/console)



Configuring HTCondor Submit Host



Configuring HTCondor Submit Host

- Step 3: Configure HTCondor with keys (and API URLs*)
- Edit `/etc/condor/config.d/40-oauth-tokens.conf` with API client details, e.g. for Box.com:

```
# Box.com client
BOX_CLIENT_ID = wluxtsxho2c4vabn3xs6n8lh0c0fznwu
BOX_CLIENT_SECRET_FILE = /etc/condor/.secrets/box
BOX_RETURN_URL_SUFFIX = /return/box
```

- `condor_reconfig`

*If not already set by default, see `condor_config_val -dump TOKEN_URL AUTHORIZATION_URL`

SUBMITTING JOBS WITH TOKENS

Submitting Jobs with OAuth Tokens

- Suppose the admin has configured the submit host with Box.com per the previous slides
- Minimal example (`simple_box.submit`):

```
executable = do_science.sh

use_oauth_services = box

queue
```

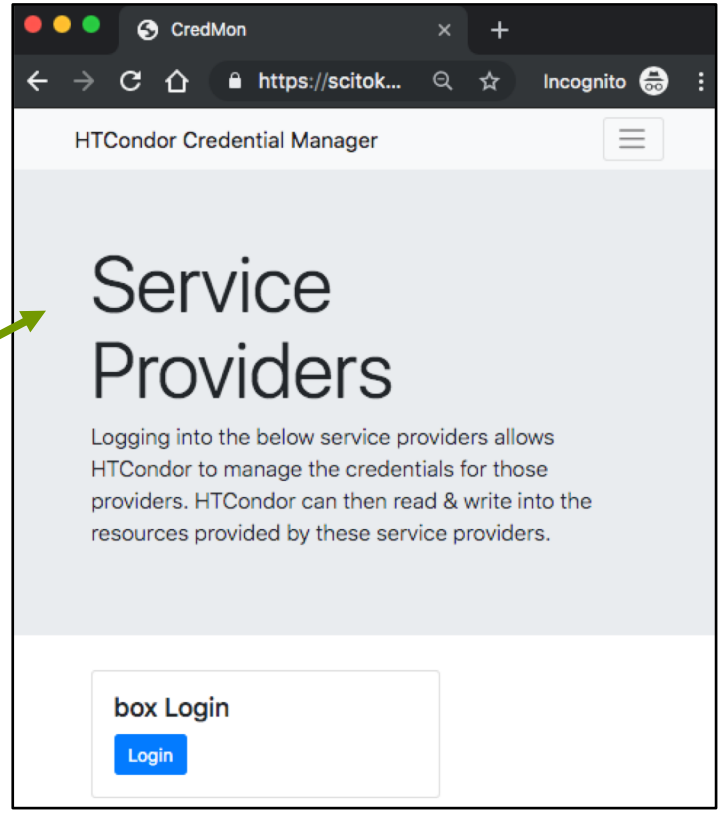
Submitting Jobs with OAuth Tokens

› Case 1: User doesn't have a Box.com token stored yet

```
[cndruser@baphomet ~]$ condor_submit simple_box.submit
Submitting job(s)
Hello, cndruser.
Please visit:
https://baphomet.cs.wisc.edu/key/151f2837e906c5107c25f
a201bc7f385e33df4ac5674700158746d4caede9355

[cndruser@baphomet ~]$ condor_q -totals

-- Schedd: baphomet : ... @ 01/01/20 00:00:00
Total for cndruser: 0 jobs; 0 completed, 0 removed, 0
idle, 0 running, 0 held, 0 suspended
```



Submitting Jobs with OAuth Tokens

› Case 1: User doesn't have a Box.com token stored yet

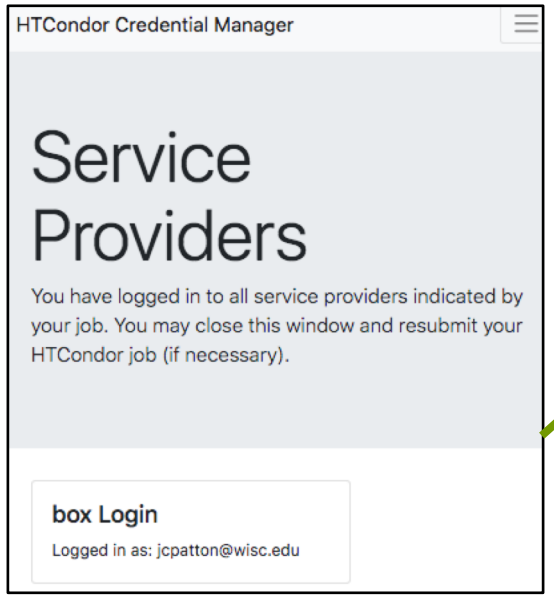

```
[cndruser@baphomet ~]$ condor_q -totals

-- Schedd: baphomet : ... @ 01/01/20 00:01:00
Total for cndruser: 0 jobs; 0 completed, 0 removed, 0 idle, 0 running, 0 held, 0 suspended
```

Still no job submitted yet!

Submitting Jobs with OAuth Tokens

› Case 1: User doesn't have a Box.com token stored yet




```
[cndruser@baphomet ~]$ condor_submit simple_box.submit
Submitting job(s).
1 job(s) submitted to cluster 1.
```

Only now has the job been submitted!

Submitting Jobs with OAuth Tokens

- What happened to the token?
  - Access tokens are copied to a restricted directory in the job sandbox, `$_CONDOR_CREDS`, under `<service_name>.use`
  - In the previous example, the token file could be accessed from within the job environment via `$_CONDOR_CREDS/box.use`
  - Only access tokens are sent to the job sandbox, not refresh tokens!

Submitting Jobs with OAuth Tokens

- Case 2: Using a file transfer plugin
- HTCondor ships with plugins for Box.com (`BOX_...`), Google Drive (`GDRIVE_...`), and Microsoft OneDrive (`ONEDRIVE_...`)

```
executable = do_science.sh
arguments = input_$(ProcId).txt output_$(ProcId).txt

use_oauth_services = box

transfer_input_files = box://my_input_files/input_$(ProcId).txt
transfer_output_remaps = “output_$(ProcId).txt = box://my_output/output_$(ProcId).txt”

queue 100
```

(Soon to be) FAQs

- How can tokens be scoped?
  - Use `<service>_oauth_permissions`, for example:

```
use_oauth_services = uwtokens
uwtokens_oauth_permissions = read:/shared
```

- How can tokens be tied to a specific resource?
  - Use `<service>_oauth_resource`, for example:

```
use_oauth_services = uwtokens
uwtokens_oauth_permissions = read:/shared
uwtokens_oauth_resource = https://mironlab.wisc.edu
```

(Soon to be) FAQs

- How can multiple tokens be obtained and used from the same token provider?
  - Add _<handle> suffix to each command, for example:

```
use_oauth_services = uwtokens

uwtokens_oauth_permissions_read = read:/shared
uwtokens_oauth_resource_read = https://mironlab.wisc.edu/

uwtokens_oauth_permissions_write = write:/home/jpatton
uwtokens_oauth_resource_write = https://jpatton.wisc.edu/
```

Accessed in job sandbox via:

```
$_CONDOR_CREDS/uwtokens_read.use
$_CONDOR_CREDS/uwtokens_write.use
```

(Soon to be) FAQs

- How can multiple tokens be obtained and used from the same token provider if scopes aren't needed? How do I use multiple tokens with file transfer plugins?
  - Add `_<handle>` suffix to `permissions` but leave value blank, use `<handle>+<service>` as the URL protocol, for example:

```
use_oauth_services = box

box_oauth_permissions_public =
box_oauth_permissions_private =

transfer_input_files = public+box://input.txt
transfer_output_remaps = "output.txt = private+box://output.txt"
```

(Soon to be) FAQs

- What if I put the OAuth Credmon Webserver under a different directory?
  - Set `CREDMON_WEB_PREFIX = https://<hostname>/<path>`
- Box.com Example:
  - `CREDMON_WEB_PREFIX = https://schedd.hostname/path/to/credmon-webserver`
  - `BOX_RETURN_URL_SUFFIX = /return/box`
  - Set Box.com API OAuth 2.0 Redirect URI: `https://schedd.hostname/path/to/credmon-webserver/return/box`

Questions?

- Documentation coming to the HTCondor manual in 8.9.9
  https://htcondor.readthedocs.org
- Jason Patton - jpatton@cs.wisc.edu
- Zach Miller - zmiller@cs.wisc.edu