iDigBio appliances

Renato Figueiredo (renato@acis.ufl.edu)

Cyberinfrastructure Working Group (CYWG) meeting
October 10, 2014

Appliances

- Complement the cyber-infrastructure core
  - Functionality desired on the client
  - Hide iDigBio APIs; user-friendly interface
- Image ingestion appliance
  - Batch upload of images to iDigBio cloud using APIs
- Virtual machine appliances
  - Package existing bio-collection tools
  - Use cases: training, technology evaluation
  - Vagrant packaging; desktops, servers (VMware, Virtualbox)
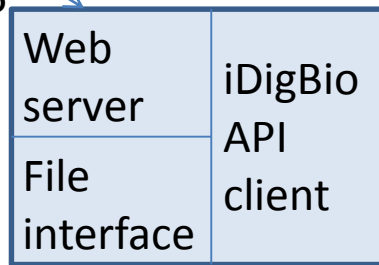
Image ingestion appliance

- Reliable approach to upload batches of thousands of images with metadata to the iDigBio repository.

- Successfully helped users to upload 290,000+ images.

Web UI browser

HTTP

Web server	iDigBio API client
File interface	

Appliance

HTTP

iDigBio Cloud API server
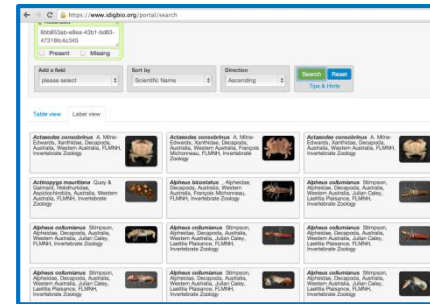
Images stored on local media

Upload Images with CSV File

Generate CSV | Upload Via CSV | Upload History

Image License: CC0 (Public Domain)

CSV File Full Path: C:\Users\winday0215\Desktop\idigbio\images\media_records.

⊙ Upload   ▤ CSV file format

Progress: (Successful:350, Skipped: 810, Failed: 0, Total to upload: 2821.)

- Upload starts with CSV file with image paths, identifier, and additional metadata
- 10 threads are used to speed up transmission.

CSV File Generation

Generate CSV	Upload Via CSV	Upload History

Upload Path *   C:\Users\winday0215\Desktop\idigbio\images

☑ Also Search Files in the Sub-directories.

GUID Syntax *   GUID = hash of image contents ▾

CSV Save Path   If left blank, the file will be saved to the image directory.

Note: Fields with * are mandatory.

📄 Generate CSV

- Appliance helps users generate a CSV file (with GUID and path) for all images within a directory hierarchy
- Optionally, users can manually edit or define new metadata fields in the CSV file

Viewing the History

Generate CSV   Upload Via CSV   Upload History

Batch Information Table

(Click on each row to see the details)

Show 10 ▼ entries   ← Previous 1 Next →

ID	CSV File Path	Start Time	Total Records	Failed Records	Skipped Records
1	C:\Users\winday0215\Desktop\idigbio\images2\media_records.csv	2014-08-17 22:15:34	336	0	0
2	C:\Users\winday0215\Desktop\idigbio\images2\media_records.csv	2014-08-17 22:15:53	336	0	0
3	C:\Users\winday0215\Desktop\idigbio\images2\media_records.csv	2014-08-17 22:16:21	336	0	220
4	C:\Users\winday0215\Desktop\idigbio\idigbio-ingestion-tool-	2014-09-07	3	0	0

- The local upload history can be viewed / saved to CSV files
- Current upload results are also shown after each upload

Virtual Appliances

- Package bio-collections software, dependences and configuration within virtual machine
  - Example: Specify appliance:
    - Ubuntu 12.04 LTS
    - MySQL
    - Java 7
    - Specify 6.5
    - Demo database
- VM runs on various O/Ss (Windows, OS X, Linux)
- Use cases:
  - Workshops/training
  - Dissemination of tools and technology preview

Specify Appliance in workshops

- Users bring their own laptop to a workshop
- During the workshop, or in advance:
  - User downloads appliance "image" from iDigBio server
    - One (large) file: ~2GB
  - User installs free virtual machine software
    - VMware (for Windows)
    - VirtualBox (for Mac)
- Saved approximately half a day of setup time in the Specify workshops where it has been used
  - No dependences with respect to software on user's computer (Java, MySQL)
  - Consistent environment across all users (interface and databases)

Specify Appliance User Interface
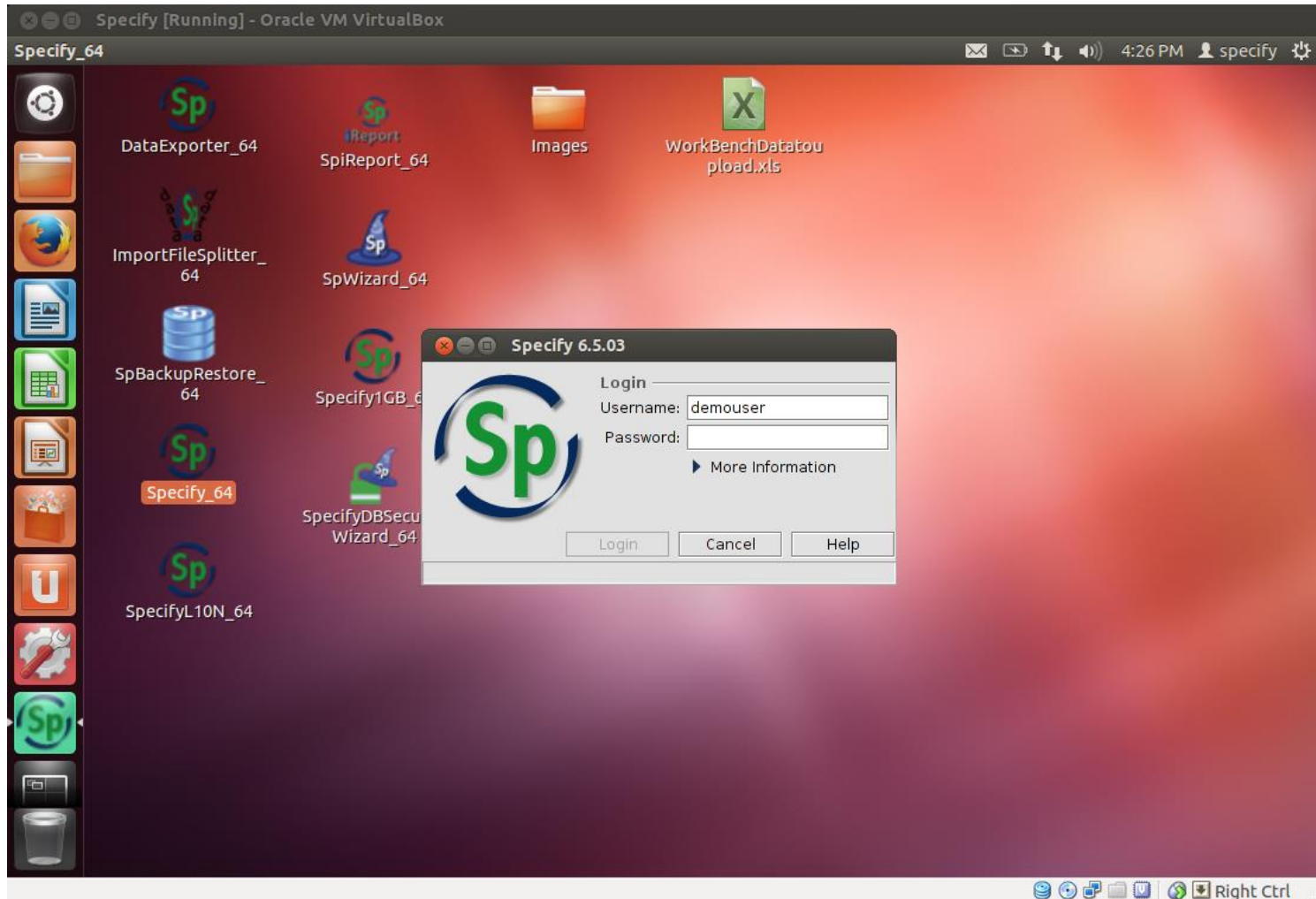
Planned features

- Image ingestion appliance
  - Enhanced collection of performance and failure metrics to guide improvements and tune parallel transfers
  - Improvements to support larger batches
  - UI improvements based on user feedback
- Virtual appliances
  - Continued improvement of Specify appliance based on experience in workshop settings
  - Work with the community to identify additional bio-collections tools that are candidates for integration in virtual appliances

Links

- Image ingestion appliance:
  - Wiki and download links:
    https://www.idigbio.org/wiki/index.php/CYWG_iDigBio_Image_Ingestion_Appliance
  - Windows 7 and up, Mac OS and Linux.
  - IE, Firefox, Chrome and Safari.
- Virtual appliances

  https://www.idigbio.org/wiki/index.php/IDigBio_Virtual_Appliances

For feedback: idigbio@acis.ufl.edu

www.idigbio.org

facebook.com/iDigBio

twitter.com/iDigBio

vimeo.com/idigbio

idigbio.org/rss-feed.xml

webcal://www.idigbio.org/events-calendar/export.ics