Pushing Data to Galaxy

Ido M. Tamir CSF, Vienna

May 26, 2011

Background

- CSF: Campus Science Support Facilities
- since 2011
- Sequencing, Fly library (VDRC), MS, Bioinformatics, Imaging, Phenotyping,...
- Sequencing should provide at least primary and secondary data analysis
- Tools for users to analyze data

Galaxy was a Closed System

User Upload

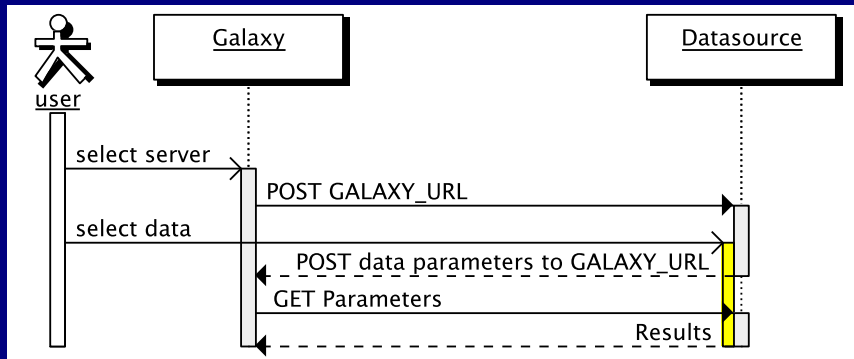## Upload File

- file
- http
- ftp

User Upload
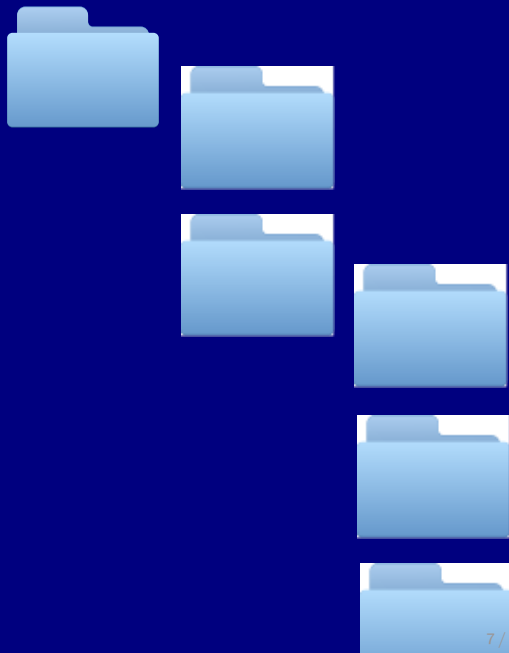
Upload File

- file
- http
- ftp

User Upload

Upload File

- file
- http
- ftp

A/Synchronous Data Depositing
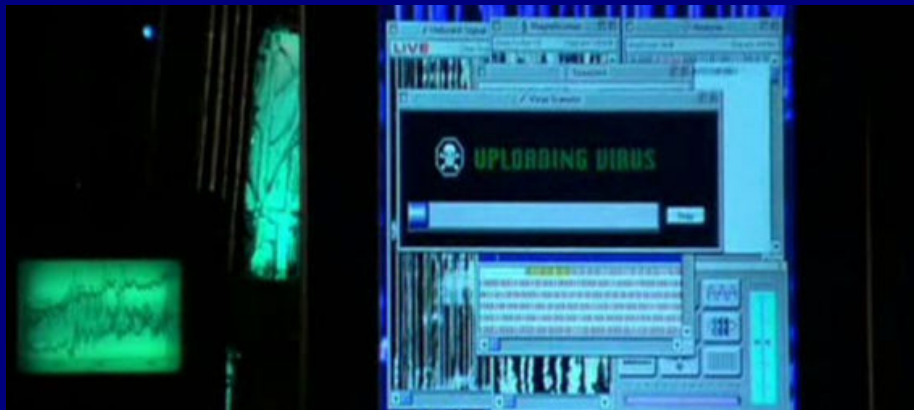
RESTful API β

/api/controller/...

- GET retrieval
- POST creation/modification
- PUT creation
- DELETE deletion

RESTful API β

usage

- enable in universe_wsgi.ini
- create api-keys for user
- use scripts in GALAXY_HOME/scripts/api
- scripts for libraries, samples

Use Cases: Data Pushers

- machine + LIMS
- server (data repository)
- workflows
- genome browser
- embedded bioinformatician

Use Cases: Data Pushers

- machine + LIMS
- server (data repository)
- workflows
- genome browser
- embedded bioinformatician

Use Cases: Data Pushers

- machine + LIMS
- server (data repository)
- workflows
- genome browser
- embedded bioinformatician

Use Cases: Data Pushers

- machine + LIMS
- server (data repository)
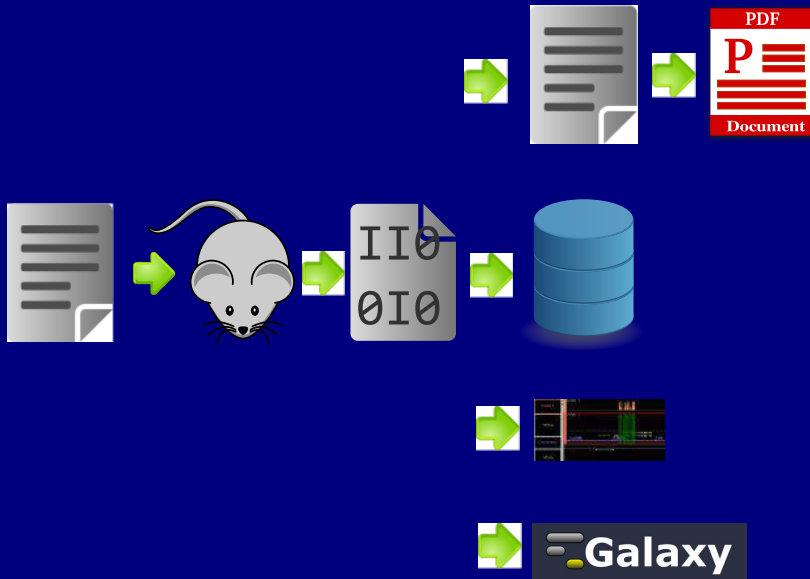- workflows
- genome browser
- embedded bioinformatician

Use Cases: Data Pushers

- machine + LIMS
- server (data repository)
- workflows
- genome browser
- embedded bioinformatician

Use Cases: Data Pushers

results store

results store

Modifications

- hg branch pusher
- registered new resource for api_mapper in lib/galaxy/web/buildapp.py
- put some more undocumented python code in lib/galaxy/web/api and scripts
- added hidden tool that executes the requested action in tools

Generated HTTP Request

Overview

- POST
- domain http://yourdomain/galaxy
- /api/fileposts?key=df96...
- {"poster" : "who", "message" : "encrypted & encoded" }

Generated HTTP Request

message, encrypted & encoded

parameter	description
action*	ln cp zcat scpuser@server
filepath*	/path/of/file
as*	name in galaxy
dbkey	genome identifier
filetype*	bed, bam, gff
username*	the user you want to push to
historyname	which history to use
description	from string to xhtml (e.g. table)

* = mandatory

transport capabilities (action)

- cp - copy to local from local file
- ln - symbolic link from local file
- zcat - gunzip to local (+cp)
- scp - scp remote file to local

transport capabilities (action)

- cp - copy to local from local file
- ln - symbolic link from local file
- zcat - gunzip to local (+cp)
- scp - scp remote file to local

transport capabilities (action)

- cp - copy to local from local file
- ln - symbolic link from local file
- zcat - gunzip to local (+cp)
- scp - scp remote file to local

transport capabilities (action)

- cp - copy to local from local file
- ln - symbolic link from local file
- zcat - gunzip to local (+cp)
- scp - scp remote file to local

Generated HTTP Request

```
POST /api/fileposts?key=df96f277488364f141b8f19ca58cc8f0 HTTP/1.1
Accept-Encoding: identity
Content-Length: 1316
Host: localhost:8080
Content-Type: application/json
Connection: close
User-Agent: Python-urllib/2.6

{"poster": "pusher1", "message": "MmQzZjYyNWI4YzkzODY0MWQ4YjdjMTY1OTUzMGY1MmE1YzM1ZTM2Y
2EzZDdmNzYxMzNkOGJmYjgxYzYxN2ViYTMxNTk2NDVhMDA5OTc2MzZhMWZiMzg0ZDBlNmFjYjlmYjBiMWY5MWQ5
MjRjYjg0YzI5YmRiYmNhNDMwMWI4ZDc1MmRjOTkyMDQxNDkyNzUyZDc2NmUxZTRmNGUwMzU0NDJiMWRlMTU4MTl
kNGU2ODE2YjBhNjFjZWFjOTA4OGM5NjA5NDg1OWQ0MDIyZmJkZDY2M2M3YzA1Yjk0MjViMWFjMWQzNmNiNjM3NT
gx0WQzNTc4ZGYwM2FlMDQzZDY3ZmY5YThlNzQwNTE3NWExN2M3MWY2MzQ5N2E2N2RlZmJlNGM4NWQ0MWI2NDJmO
DE2YmI2NDY2ZjJlE5YzY2YWJiNWQ3MGIyNmI10DZhMzk4MjlkZWY4ZjQ3NTcxMzE5NjVm0Dkz0DAw0TY1YWFjYzZm
NThkMmUz0WZm0TBiZmFmZWI0ZDNiMWNhMzhlYzUyMDI2Yzc0ZGFkNWMzZGIzM2VhZDk5NTg3ZjBlNzliNTZhNDJ
mZGYxZDRlN2E4YjZkYTRhZTg2MGYyYmRiMjk5NGQ0ZmRl0WZmNjk1NTNjZDJjZGFjZjBiMzZhN2U3NTQ0NWU50T
U4NGQzYWQ1NGQ5MDI2NjNmYjQ5ZWZkMTQ2YmM3YmNlZTgyMWRiZDJiNmU3YjUxNTQxNTU2NTk2NjQ4MzRjMGYzY
zhiZGQzNTJmNmI2YTZkZDNiMWNhMzhlYzUyMDI2YzljMjBmMDI1NWJhMWZjYjI0MTU1Nj U5NjY0ODM0YzBmMDJk
0WEwZTdkZjFmYWI0NmU5YzNjMTNiMGVlMDEwNTNlZDU0OWZkZjZjNGVkODRiMjczMTQ1M2RhZDdjZjdjZWI1YWQ
4NGI5ZTA40GJhNmI0MmIzZGUyZTFmYTRlNTc1MDZjYTVjNTFhZGYzNDUyM2EyZDRjNzNhNTVlMjIyYTNhNTQw0T
MzZmI5MDY40WJlNzAzMjQ4YjA4MWFhNDcyNmRk0GMz0GZhYTAw0TQ0NDM4ZjI0NzY10TY2NWNkMGEyZmQzZGZkM
TlmZjNhMDMxNDkyNjY5NTU4MTQ1N2Y5MzY20GU1YjRjMzdiYjNhMzZi0WZjYTU2NDNhMWRh0TBkZTllYzBj0DIy
Mjc5YTA10TAw0TMw0DlhYzQ5ZTE1MWMzNDBmZjgxNmNkNDY0ZjBkNWY1MDdh0DM5NjNhMWY4YzUyNGE1NWM3MWM
4YWIyYTkw"}
```

Example Pushed Dataset

Editing of Annotation not Nice

Annotation / Notes:

```
<table>

 <tr>

  <td>

   sampleNr

  </td>

  <td>

   4592

  </td>
```

Add an annotation or notes to a dataset: annotations are...

Why not Libraries?

- not pushing to himself!
- "pushed to" user sees his results
- linking, scp,... possible
- structure
- who is allowed (not just admins)
- complete message encrypted/tamperproof
- have to revisit libraries and sample tracking

Why not Libraries?

- not pushing to himself!
- "pushed to" user sees his results
- linking, scp,... possible
- structure
- who is allowed (not just admins)
- complete message encrypted/tamperproof
- have to revisit libraries and sample tracking

Why not Libraries?

- not pushing to himself!
- "pushed to" user sees his results
- linking, scp,... possible
- structure
- who is allowed (not just admins)
- complete message encrypted/tamperproof
- have to revisit libraries and sample tracking

Why not Libraries?

- not pushing to himself!
- "pushed to" user sees his results
- linking, scp,... possible
- structure
- who is allowed (not just admins)
- complete message encrypted/tamperproof
- have to revisit libraries and sample tracking

Why not Libraries?

- not pushing to himself!
- "pushed to" user sees his results
- linking, scp,. . . possible
- structure
- who is allowed (not just admins)
- complete message encrypted/tamperproof
- have to revisit libraries and sample tracking

Why not Libraries?

- not pushing to himself!
- "pushed to" user sees his results
- linking, scp,... possible
- structure
- who is allowed (not just admins)
- complete message encrypted/tamperproof
- have to revisit libraries and sample tracking

Why not Libraries?

- not pushing to himself!
- "pushed to" user sees his results
- linking, scp,... possible
- structure
- who is allowed (not just admins)
- complete message encrypted/tamperproof
- have to revisit libraries and sample tracking

Acknowledgements

- Galaxy Team
- CSF Vienna
- Andreas Sommer
- Andras Aszodi, Benjamin Almeida, Markus Jaritz, Wolfgang Lugmayer