

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



Galaxy Initiates Import



User Upload

Upload File

- file
- http
- ftp

User Upload

Upload File

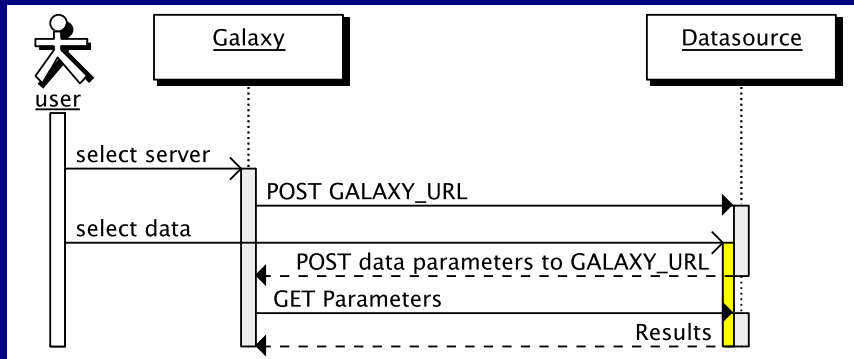
- file
- http
- ftp

User Upload

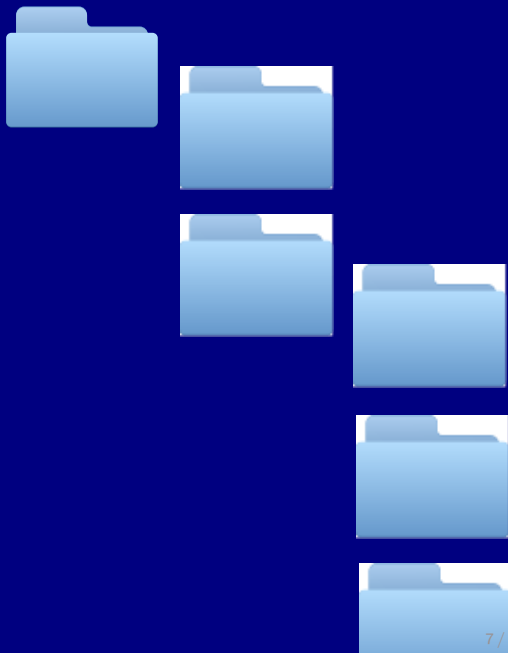
Upload File

- file
- http
- ftp

A/Synchronous Data Depositing



Galaxy Importer



RESTful API β

/api/controller/...

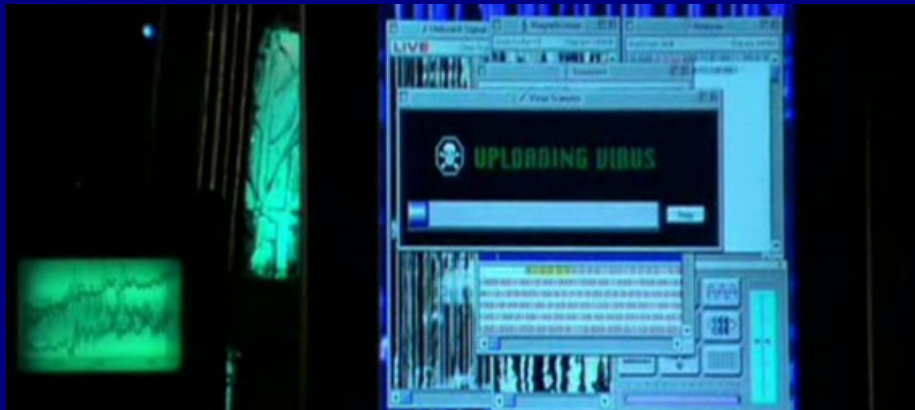
- GET retrieval
- POST creation/modification
- PUT creation
- DELETE deletion

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

RESTful API β



Big Picture: Interoperability



LIMS

 Galaxy

The Galaxy logo consists of three horizontal bars of varying lengths (grey, grey, yellow) to the left of the word 'Galaxy' in a white, bold, sans-serif font, all contained within a dark blue rectangular background.

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



Use Cases: Data Pushers

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*	In 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": "MmQzZjYyNWlY4YzkzODY0MWQ4Yj djMjY1OTUzMGY1MmE1Yz M1ZTM2Y
2EzZDdmNzYxMzNkOGJmYjgxYzYxN2ViYTMxNTk2NDVhMDA5OTc2MzZhMWZiMzg0ZDBlNmFjYjlmYjBiMwY5MwQ5
MjrjYjg0YzI5YmRiYmNhNDMwMwI4ZDc1MmRjOTkyMDQxNDkyNzUyZDc2NmUxZTRmNGUwMzU0NDJiMWRlMTU4MTl
kNGU20DE2YjBhNjFjZWJjOTA4OGM5NjA5NDg10WQ0MDIyZmJkZDY2M2M3YzA1Yjk0MjVIMWFiMwQzNmNiNjM3NT
gx0WQzNTc4ZGYwM2FLMDQzZDY3ZmY5YThlNzQwNTE3NWExN2M3MwY2MzQ5N2E2N2RlZmJlNGM4NWQ0MwI2NDJmO
DE2YmI2NDY2ZjE5YzY2YwJiNwQ3MGIyNmI10DZhMzk4MjlkZWY4ZjQ3NTcxMzE5NjVmODkzODAwOTY1YWFjYzZm
NThkMmUzOWZmOTBiZmFmZWl0ZDNiMwNhMzhlYzUyMDI2Yzc0ZGFkNWZmZGIzZmVhZDk5NTg3ZjBlNzljNTZhNDJ
mZGYxZDRlN2E4YjZkYTRhZTg2MGYyYmRiMjA5NDg10WQ0ZmRlOWZmNjA1NTNjZDZjZGFjZjBiMzZhN2U3NTQ0NWU5OT
U4NGQzYWQ1NGQ5MDI2NjNmYjQ5ZmZkMTQ2YmM3YmNlZTgyMWRiZDZjInM3YjUxNTQxNTU2NTk2NjQ4MzRjMGYzY
zhIZGQzNTJmNmI2YTZkZDNiMwNhMzhlYzUyMDI2YzljMjBmMDI1NWJhMWZjYjI0MTU1NjU5NjY0ODM0YzBmMDJk
0WEwZTdkZjFmYWI0NmU5YzNjMTNiMGVlMDEwNTNlZDU0OWZkZjZjNGVkdDRiMjczMTQ1M2RhZDdjZj djZWI1YWQ
4NGI5ZTA4OGJhNmI0MmIzZGUyZTFmYTRlNTc1MDZjYTVjNTFhZGYzNDUyM2EyZDRjNzNhNTVlMjIyYTNhNTQwOT
MzZmI5MDY4OWJlNzAzMjQ4YjA4MWFhNDcyNmRkOGMzOGZhYTAwOTQ0NDM4ZjI0NzY1OTY2NWNkMGEEyZmQzZGZkM
TlMzJnhMDMxNDkyNjY5NTU4MTQ1N2Y5MzY2OGU1YjRjMzdiYjNhMzZiOWZjYTU2NDNhMWRhOTBkZTllYzBjODIy
Mjc5YTA1OTAwOTMwODlhYzQ5ZTE1MwMzNDNmZjgxNmNkNDY0ZjBkNWY1MDdhODM5NjNhMwY4YzUyNGE1NW3MwM
4YWlyYTkw"}}
```

Example Pushed Dataset






   

Methylation

3:   

627CUAAXX 3 20110304a 4592 2
0110307 bowtie

~9,100,000 regions
format: bed, database: ?

Annotation:

sampleNr 4592
creator Ido Tamir
lims 4592
report report

view in [GeneTrack](#)

1.Chrom	2.Start	3.End	4.Name	5	6.Str
chr1	7890909	7890945	.	1	-

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 av

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