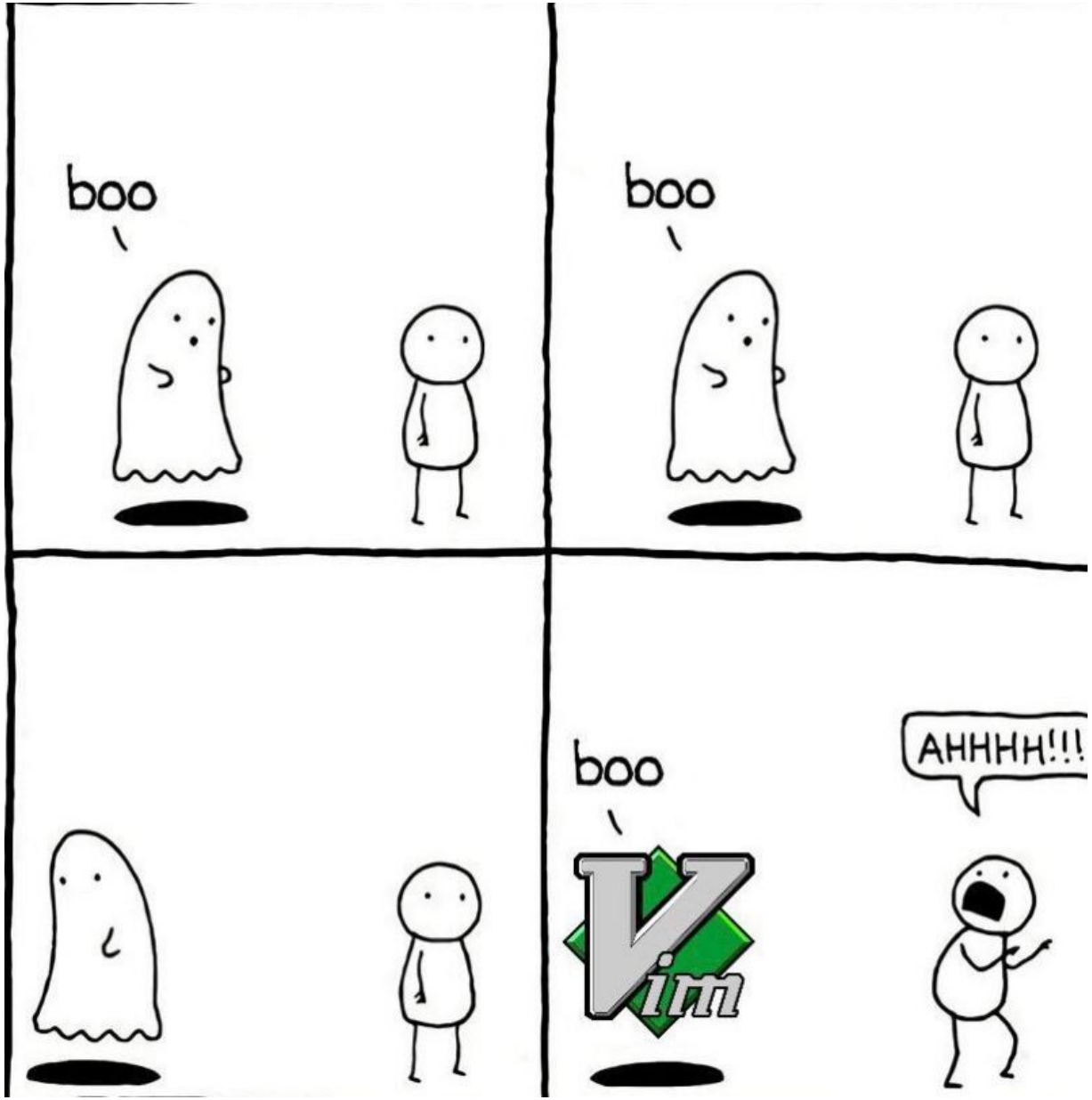
Open Science Data Federation - OSDF

Fabio Andrijauskas - fandrijauskas@ucsd.edu

UCSD


boo
boo
boo
Vim
AHHHH!!!

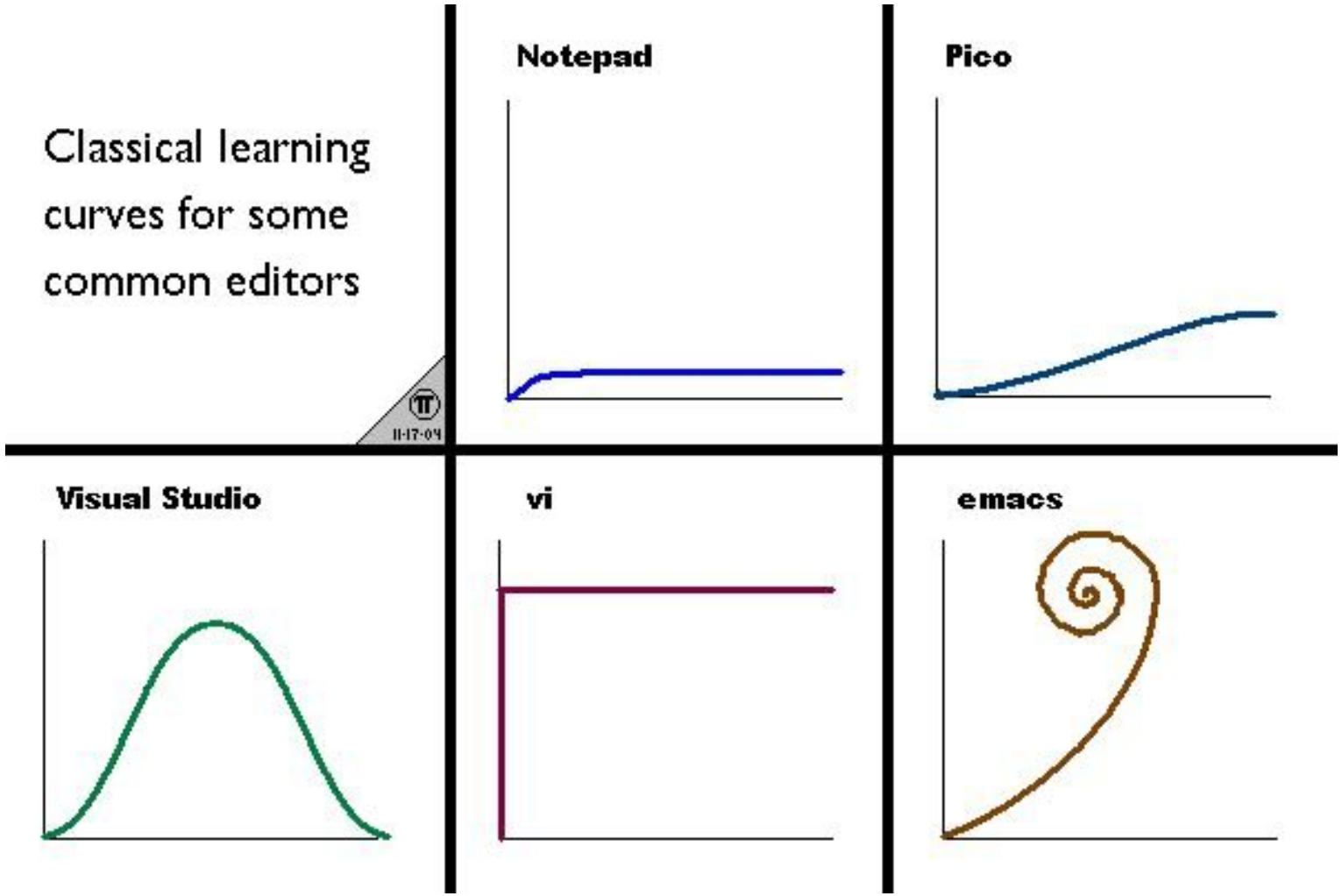
CHOOSE YOUR WEAPON

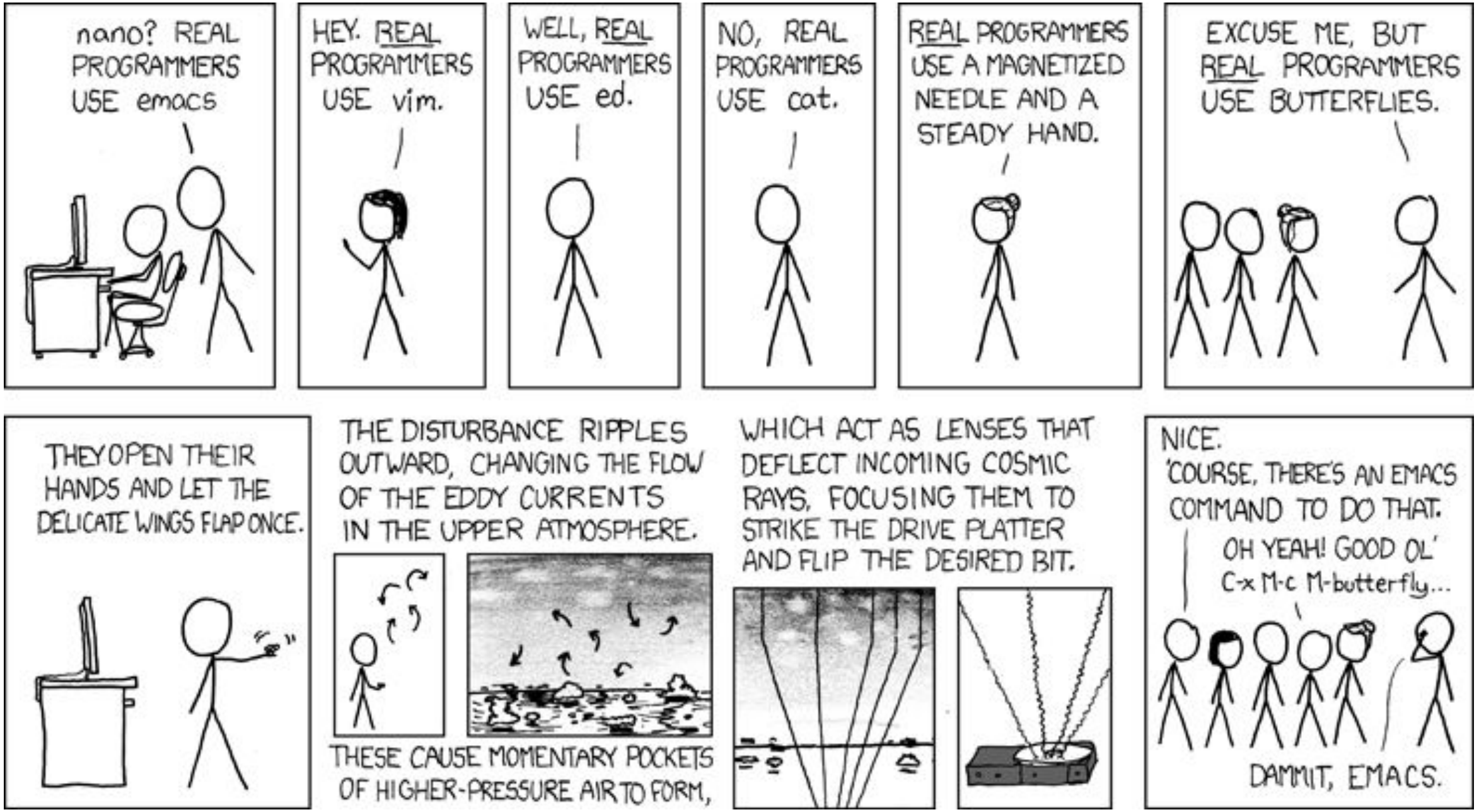
Emacs

VIM

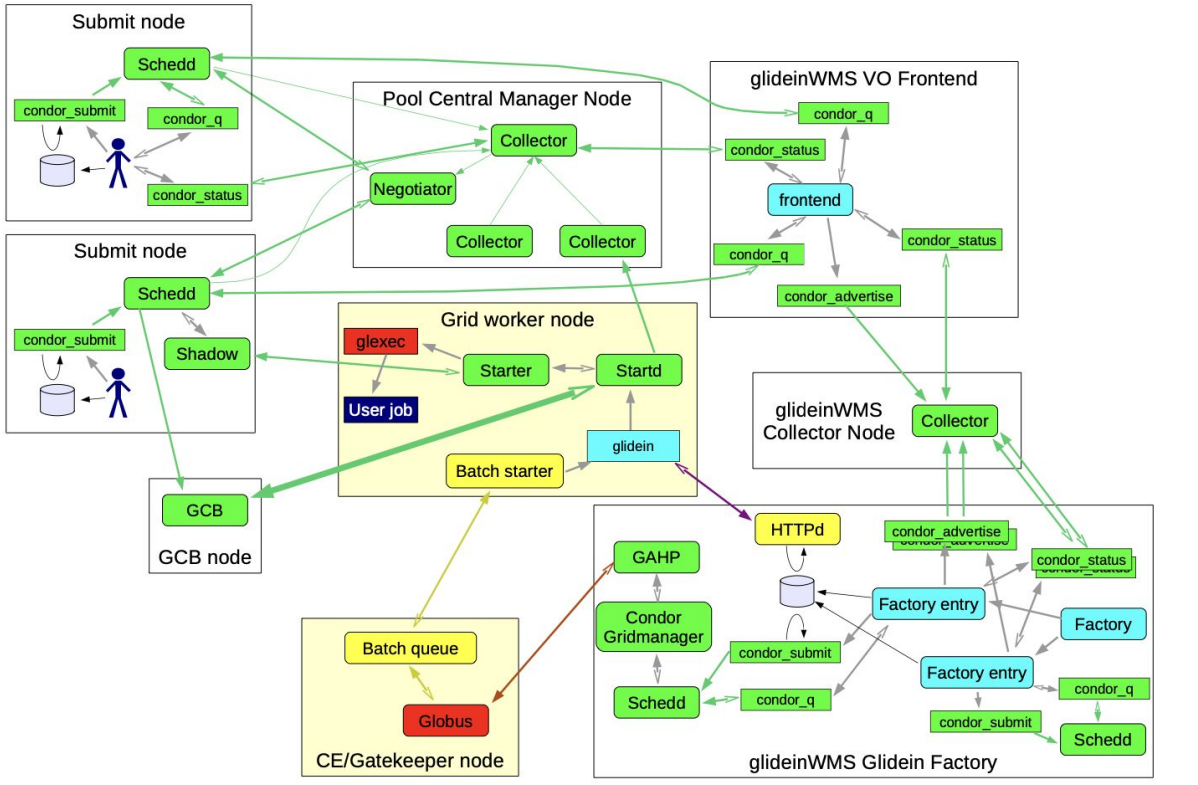
MS Notepad

Classical learning curves for some common editors

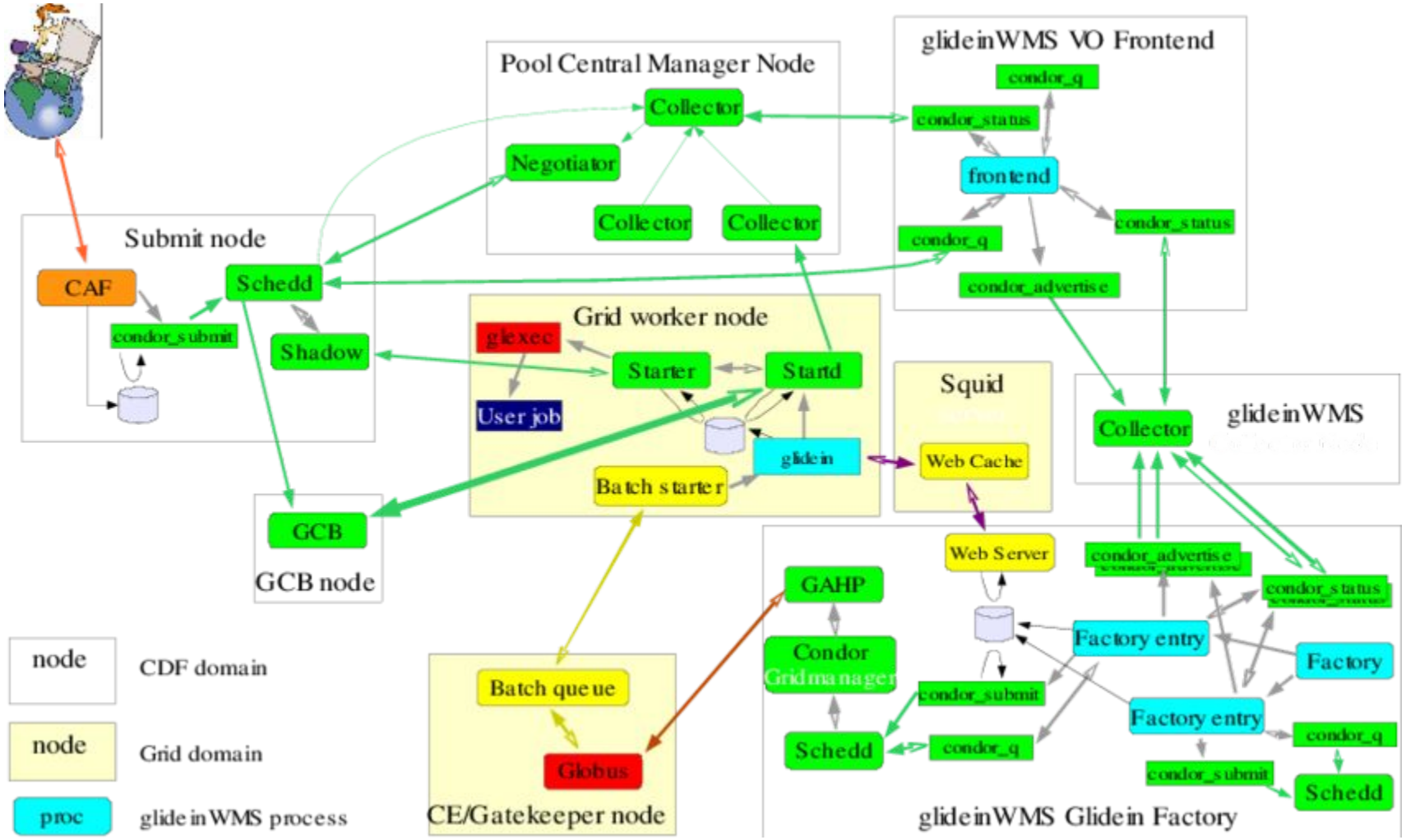



nano? REAL PROGRAMMERS USE emacs
HEY. REAL PROGRAMMERS USE vim.
WELL, REAL PROGRAMMERS USE ed.
NO, REAL PROGRAMMERS USE cat.
REAL PROGRAMMERS USE A MAGNETIZED NEEDLE AND A STEADY HAND.
EXCUSE ME, BUT REAL PROGRAMMERS USE BUTTERFLIES.
THEY OPEN THEIR HANDS AND LET THE DELICATE WINGS FLAP ONCE.
THE DISTURBANCE RIPPLES OUTWARD, CHANGING THE FLOW OF THE EDDY CURRENTS IN THE UPPER ATMOSPHERE.
THESE CAUSE MOMENTARY POCKETS OF HIGHER-PRESSURE AIR TO FORM,
WHICH ACT AS LENSES THAT DEFLECT INCOMING COSMIC RAYS, FOCUSING THEM TO STRIKE THE DRIVE PLATTER AND FLIP THE DESIRED BIT.
NICE.
'COURSE, THERE'S AN EMACS COMMAND TO DO THAT.
OH YEAH! GOOD OL' C-x M-c M-butterfly...
DAMMIT, EMACS.

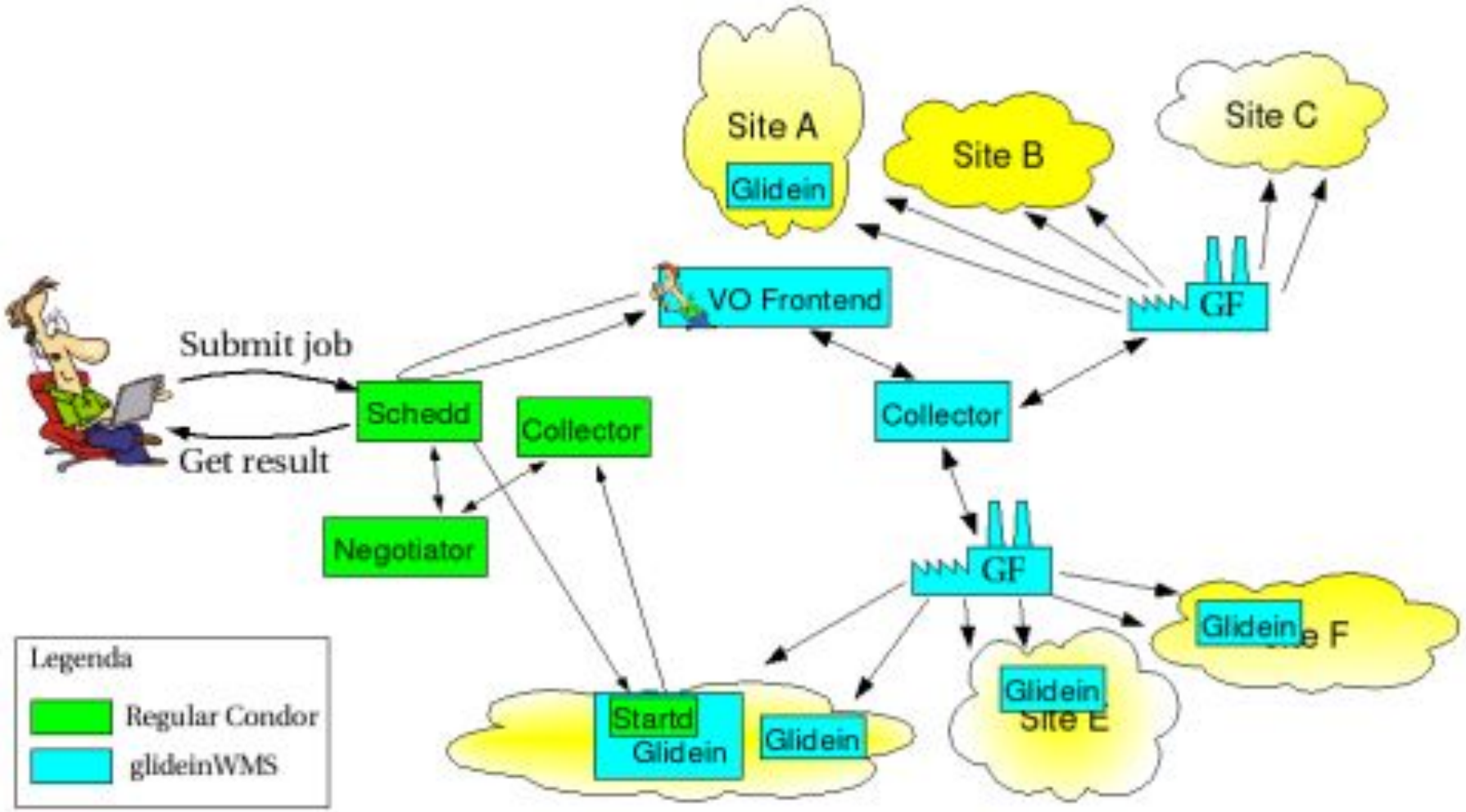
OSG


Submit node
Schedd
condor_submit
condor_q
condor_status
Submit node
Schedd
condor_submit
Shadow
GCB
GCB node
Pool Central Manager Node
Collector
Negotiator
Collector
Collector
Grid worker node
glexec
Starter
Startd
User job
glidein
Batch starter
glideinWMS VO Frontend
condor_q
condor_status
frontend
condor_q
condor_status
condor_advertise
glideinWMS
Collector Node
Collector
HTTPd
condor_advertise
GAHP
condor_status
Condor
Gridmanager
Factory entry
Factory
condor_submit
Factory entry
Schedd
condor_q
condor_q
condor_submit
Schedd
glideinWMS Glidein Factory
Batch queue
Globus
CE/Gatekeeper node
Legenda:
Condor daemon
Condor cmdline
glideinWMS daemon
glideinWMS cmdline
Grid software
System software
Execute command
Execute command and get output
Condor command
Condor 2-way communication
Long lived TCP connection
GRAM protocol
HTTP protocol
Site-dependent protocol
Disk storage
Interaction with FS
User submitting jobs
User job
User job
node
glideinWMS domain
node
Grid domain

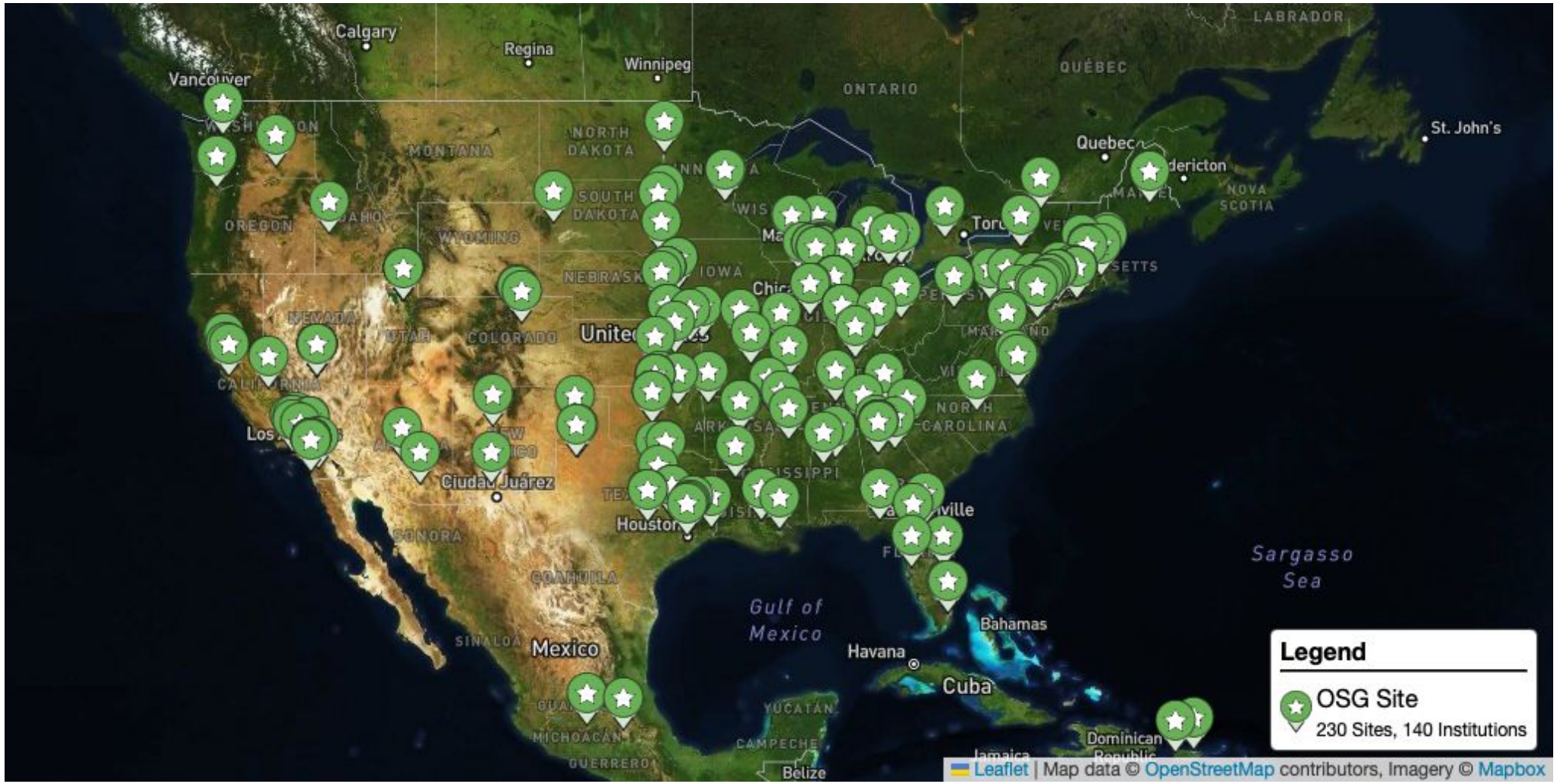
OSG


Pool Central Manager Node
Collector
Negotiator
Collector
Collector
glideinWMS VO Frontend
condor_q
condor_status
frontend
condor_q
condor_status
condor_advertise
Submit node
CAF
Schedd
condor_submit
Shadow
Grid worker node
glexec
Starter
Startd
User job
glidein
Batch starter
Squid
Web Cache
Collector
glideinWMS
GCB
GCB node
Web Server
condor_advertise
condor_status
GAHP
Condor
Gridmanager
Factory entry
Factory
condor_submit
Factory entry
condor_q
Schedd
condor_q
condor_submit
Schedd
Batch queue
Globus
CE/Gatekeeper node
glideinWMS Glidein Factory
node
CDF domain
node
Grid domain
proc
glideinWMS process

OSG


Site A
Glidein
Site B
Site C
VO Frontend
GF
Submit job
Get result
Schedd
Collector
Collector
Negotiator
GF
Glidein
Glidein
Startd
Glidein
Glidein
Glidein
Legenda
Regular Condor
glideinWMS

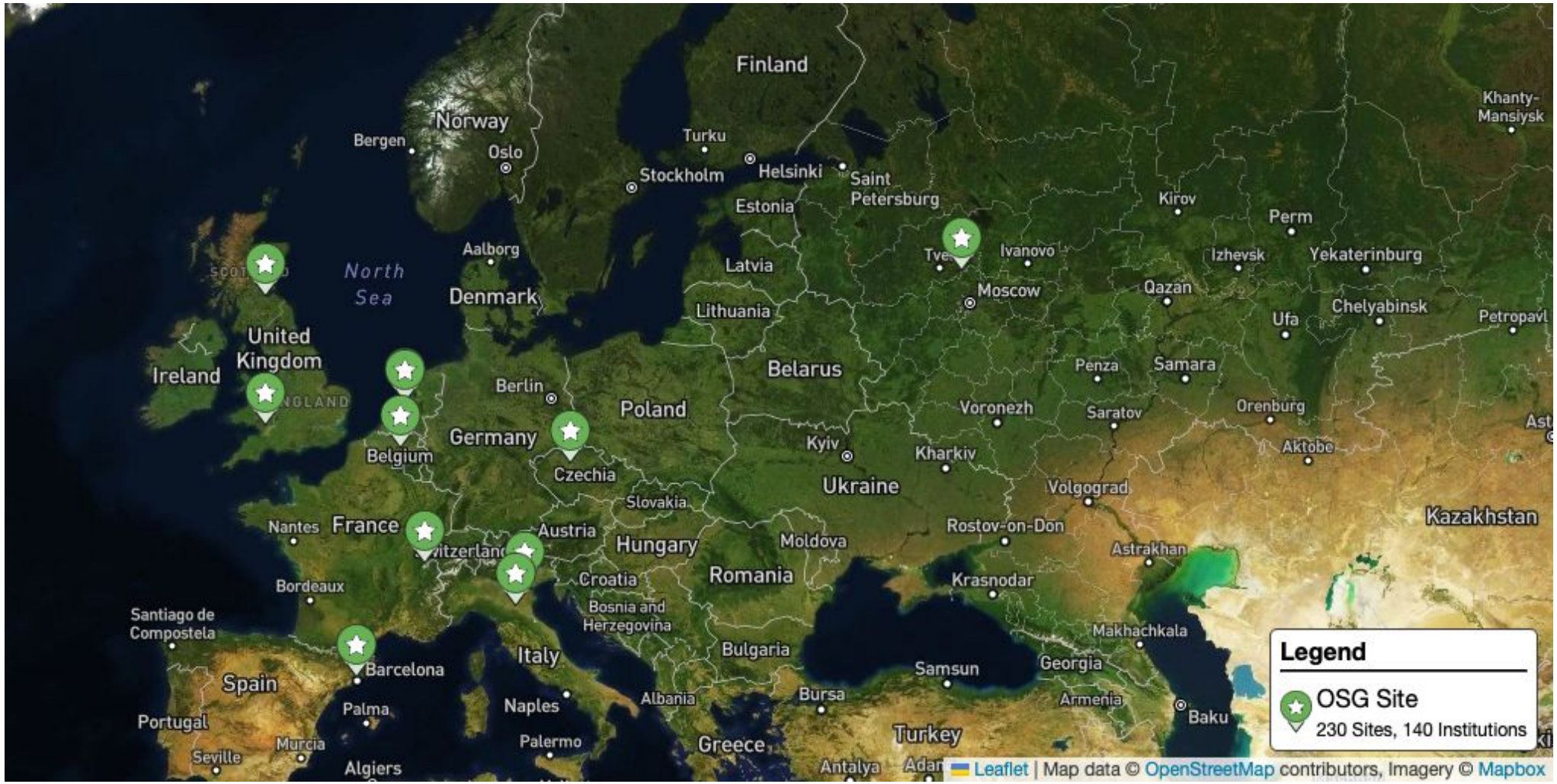
OSG


Calgary
Regina
Winnipeg
Vancouver
ONTARIO
QUÉBEC
LABRADOR
St. John's
Quebec
MONTANA
NORTH DAKOTA
SOUTH DAKOTA
OREGON
WYOMING
NEBRASKA
IOWA
NEVADA
UTAH
COLORADO
Ciudad Juárez
Houston
Gulf of Mexico
Sargasso Sea
Bahamas
Havana
Cuba
Mexico
SONORA
COAHUILA
SINALOA
YUCATÁN
CAMPECHE
MICHOACÁN
GUERRERO
Belize
Dominican Republic
Jamaica
NOVA SCOTIA
Legend
OSG Site
230 Sites, 140 Institutions
Leaflet | Map data © OpenStreetMap contributors, Imagery © Mapbox

OSG

Quito
Guayaquil
Manaus
Belém
São Luís
Georgetown
AMAZONAS
PARÁ
MARANHÃO
CEARÁ
PIAUÍ
PARAÍBA
ALAGOAS
ACRE
RONDÔNIA
Peru
Brazil
Lima
Cusco
BAHIA
Salvador
MATO GROSSO
La Paz
Bolivia
GOIÁS
Arica
MINAS GERAIS
MATO GROSSO DO SUL
Vitória
Tarija
Antofagasta
Paraguay
Salta
PARANÁ
Chile
Resistencia
SANTA CATARINA
South Atlantic Ocean
La Serena
Porto Alegre
Cordoba
Santiago
Uruguay
Buenos Aires
Talca
Montevideo
Argentina
Temuco
Neuquén
Legend
OSG Site
230 Sites, 140 Institutions
Leaflet | Map data © OpenStreetMap contributors, Imagery © Mapbox

OSG


Finland
Norway
Bergen
Oslo
Turku
Stockholm
Helsinki
Saint Petersburg
Estonia
Kirov
Perm
Khanty-Mansiysk
Aalborg
Latvia
Ivanovo
Izhevsk
Yekaterinburg
North Sea
Denmark
Moscow
Qazan
Lithuania
Chelyabinsk
Ufa
Petropavl
United Kingdom
Ireland
Belarus
Penza
Samara
Berlin
Poland
Voronezh
Saratov
Orenburg
Belgium
Germany
Kyiv
Kharkiv
Aktobe
Czechia
Ukraine
Volgograd
Slovakia
Nantes
France
Austria
Rostov-on-Don
Kazakhstan
Hungary
Moldova
Astrakhan
Bordeaux
Croatia
Romania
Krasnodar
Bosnia and Herzegovina
Santiago de Compostela
Makhachkala
Italy
Bulgaria
Barcelona
Samsun
Georgia
Spain
Albania
Bursa
Armenia
Naples
Palma
Baku
Portugal
Turkey
Murcia
Palermo
Greece
Seville
Algiers
Antalya
Legend
OSG Site
230 Sites, 140 Institutions
Leaflet | Map data © OpenStreetMap contributors, Imagery © Mapbox

Open Science Data Federation

- The Open Science Data Federation (OSDF) is an OSG service designed to support the sharing of files staged in autonomous “origins”, for efficient access to those files from anywhere in the world via a global namespace and network of caches.

- The OSDF may be used either from within OSG or independent of OSG

Example OSDF Use Cases

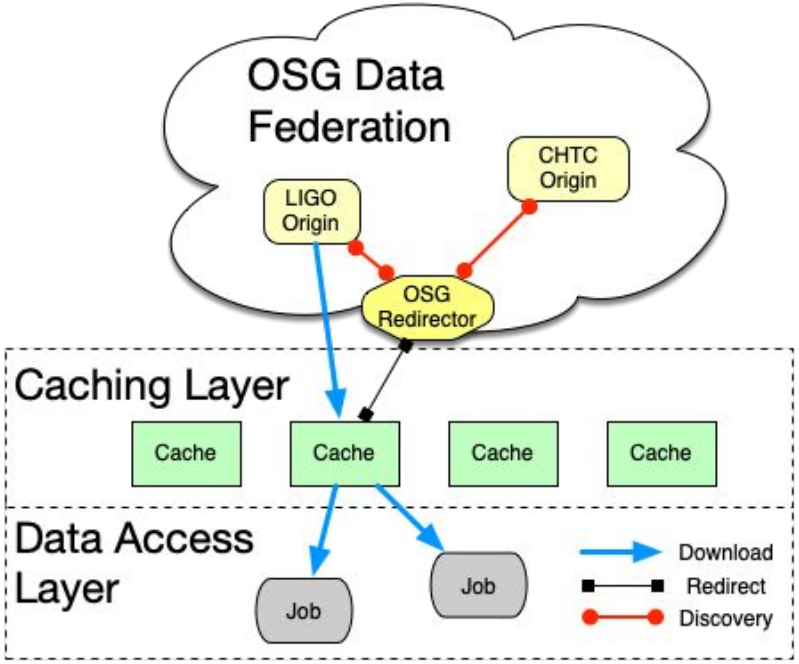
- A researcher wants to share a dataset with their community such that others may process it.

- A researcher produces data on the OSG that they need to store for future processing or sharing with the community.

- A researcher has a GB to TB-scale dataset that they want to analyze. Their workflow processes the same data many times, thus benefiting greatly from the caching within OSDF.

Open Science Data Federation

- Origin: Storage of the data from multiple Virtual Organizations (VO);
- Redirector: Process the data request to direct the request to the appropriate origin;
- Cache: Storage to provide data geographically close to the execution points and access points;
- Data access: It is possible to use different tools. Using OSDF commands is possible to fetch the files from the closest cache (GeoIP);



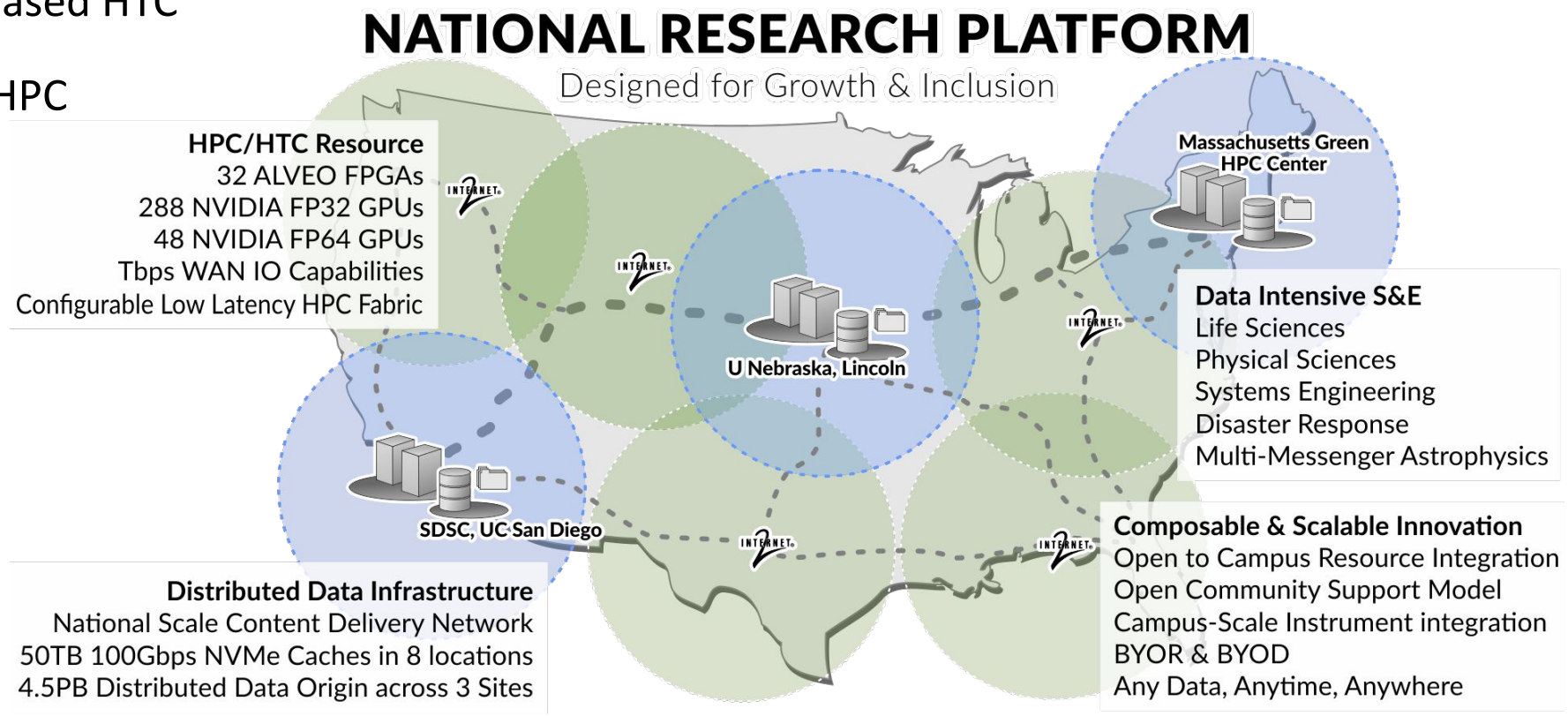
Cache, origins, and redirector based on the XRootD technology: https://xrootd.slac.stanford.edu

Open Science Data Federation

- 1 Cache (50TB for cache) and one origin (1.6PB for origin storage): San Diego Super Computer Center - San Diego - California.
- 1 Cache (42TB for cache) and one origin (1.2 PB for origin storage): University of Nebraska-Lincoln - Lincoln - Nebraska.
- 1 Cache (29TB for cache) and one origin (1.2PB for origin storage): Massachusetts Green High-Performance Computing Center - Holyoke - Massachusetts.
- 1 Cache: Internet2 - Boise - Idaho (42TB for cache).
- 1 Cache: Internet2 - Houston - Texas (in installation).
- 1 Cache: Internet2 - Jacksonville - Florida (42TB for cache).
- 1 Cache: Internet2 - Denver - Colorado (42TB for cache).
- 1 Cache: Northeastern University - Boston - Massachusetts (in installation).

System architecture

- A distributed system spanning multiple locations
- Three semi-independent components:
  - FP32-optimized GPU-based HTC
  - GPU and FPGA-based HPC
  - Storage
- All managed though a unified Kubernetes setup



Open Science Data Federation

Providing data access and transfer services for Open Science

BYTES READ
29,570 TB
↑938 MB/s Last 1 Year

FILES READ
1,188,355,803
↑38/s Last 1 Year

https://osdf.osg-htc.org

OSDF Client

StashCP is the client tool to download or upload files on the Open Science Data Federation.

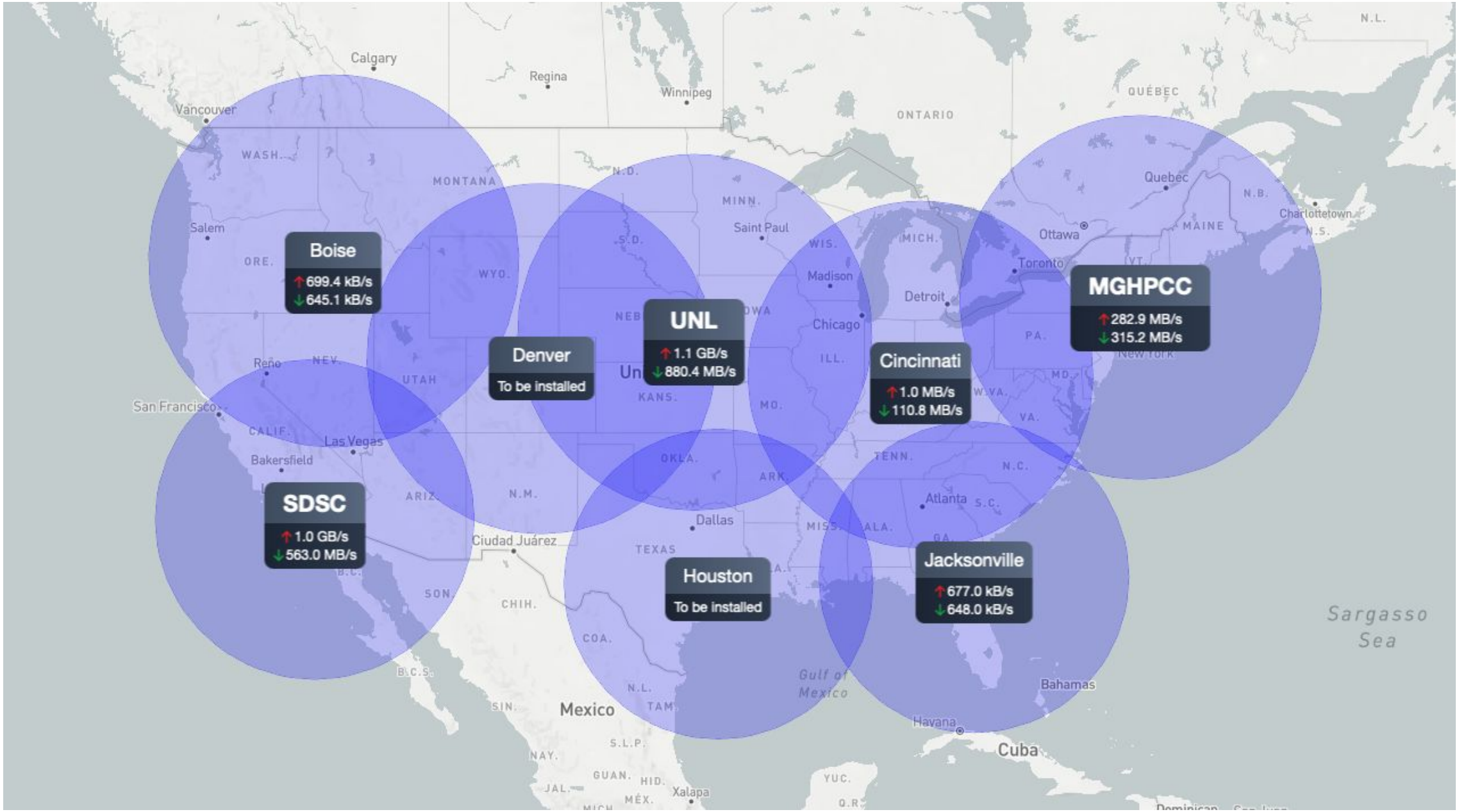
Download StashCP for your platform
rpm, deb, exe, static binary

View the Source </>

Terminal

```
$ stashcp /osgconnect/public/dweitzel/blast/queries/query1 ./
query1 done!
```

https://osdf.osg-htc.org/client

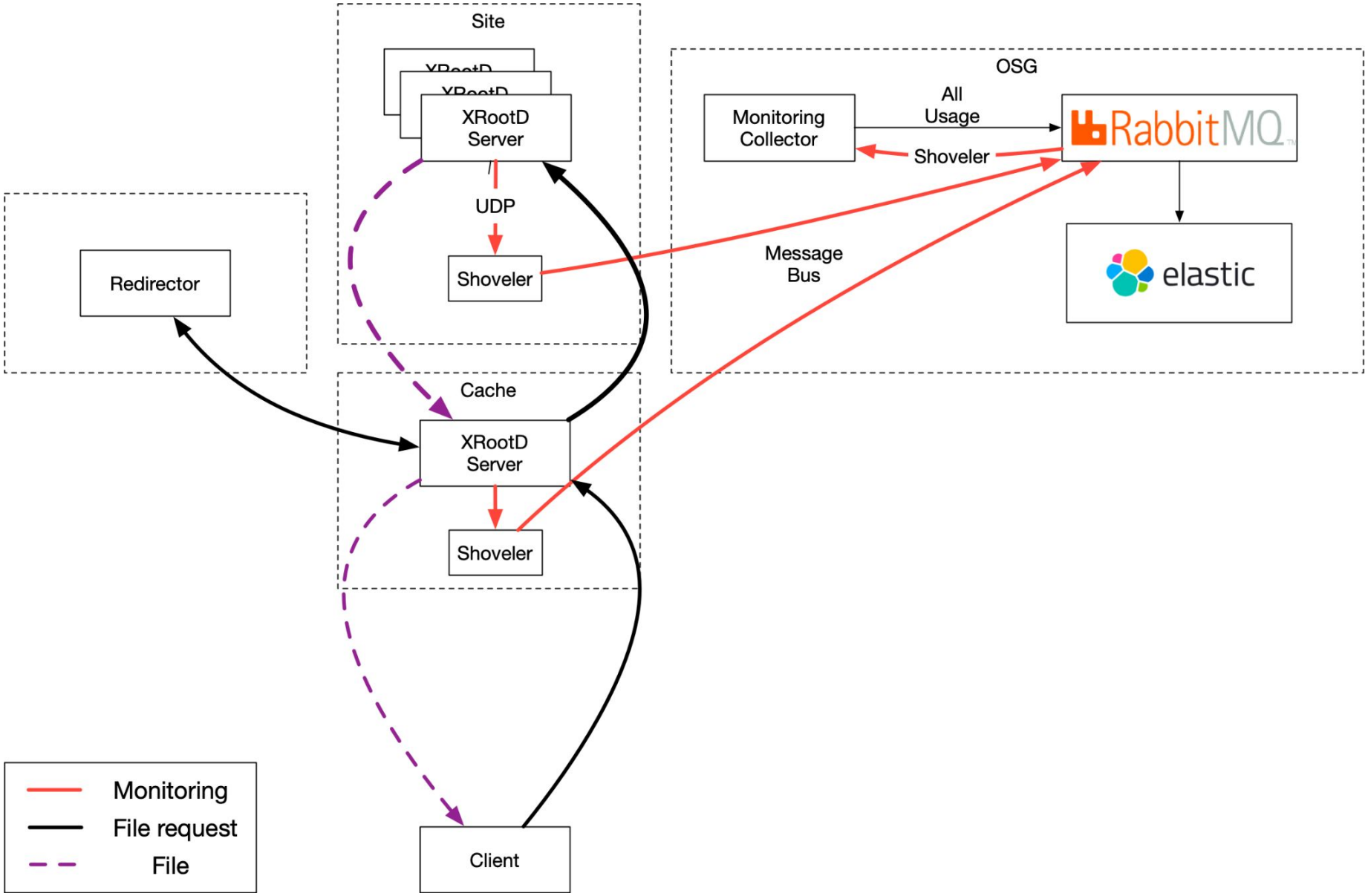


https://nrp-website.vercel.app

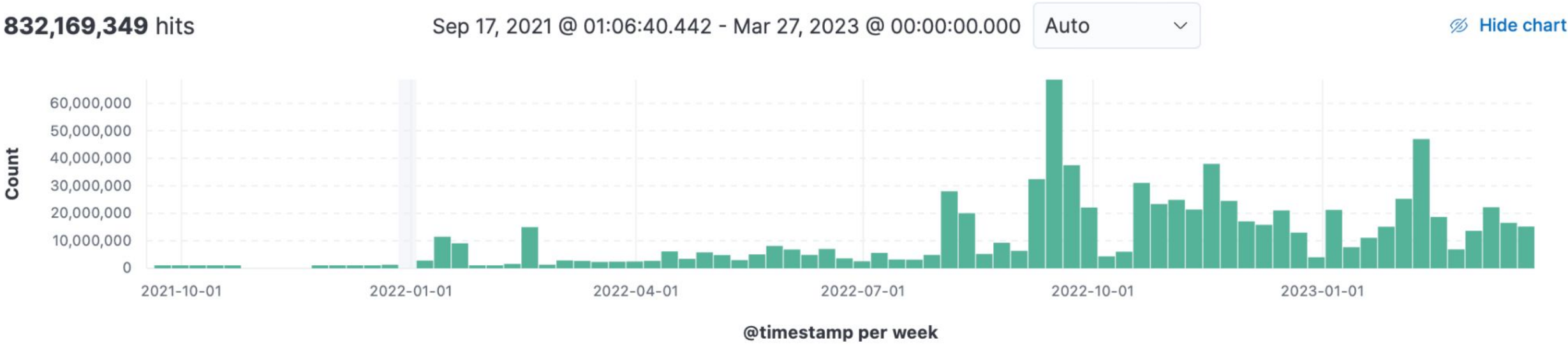
https://opensciencegrid.org/about/osdf/

https://opensciencegrid.org/about/osdf/

OSDF


Site
XRootD
Server
UDP
Shoveler
Redirector
OSG
Monitoring
Collector
All
Usage
Shoveler
RabbitMQ
Message
Bus
elastic
Cache
XRootD
Server
Shoveler
Client
Monitoring
File request
File

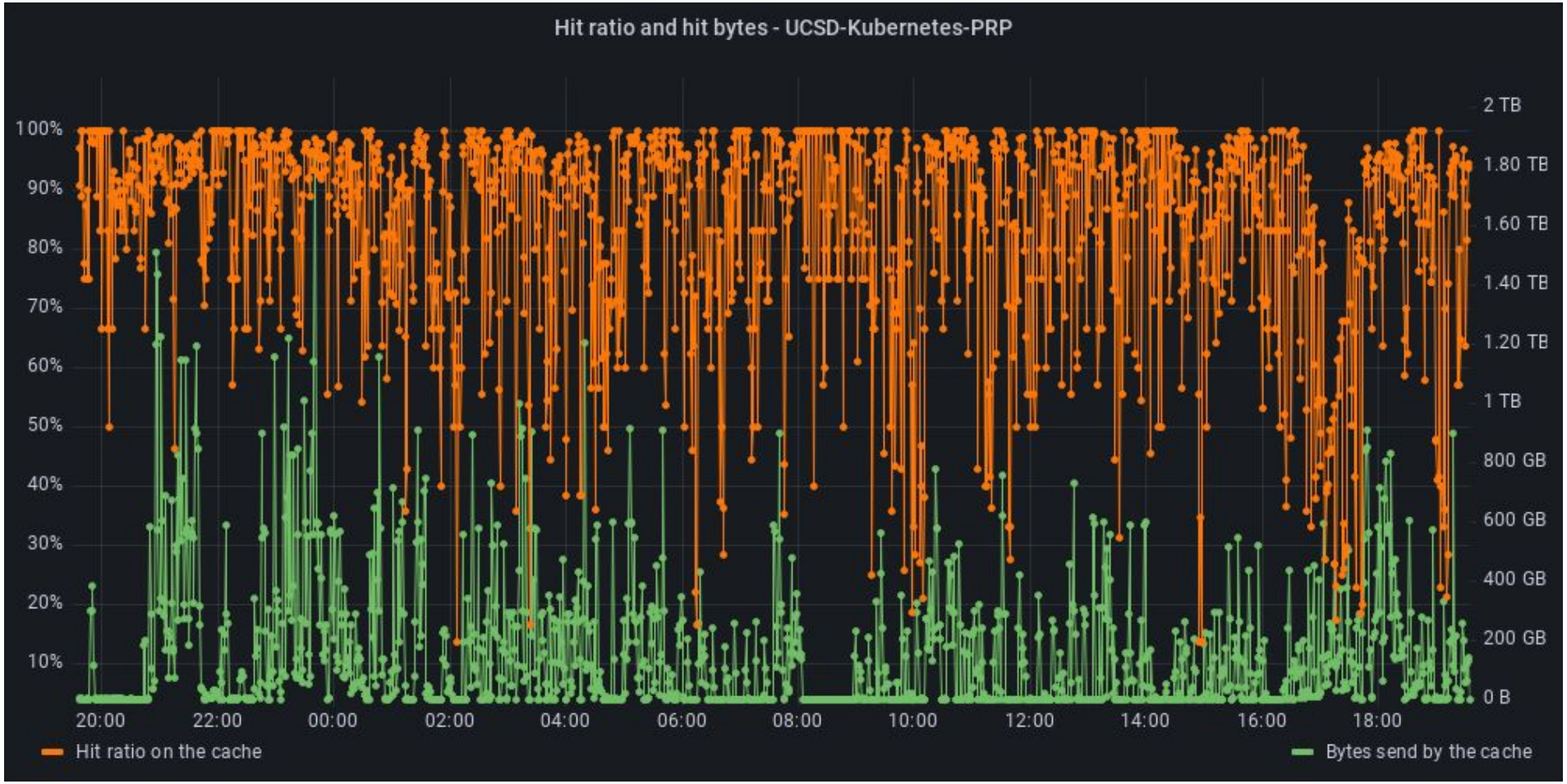
OSDF - Monitoring



OSDF - Monitoring

Hit ratio on the caches

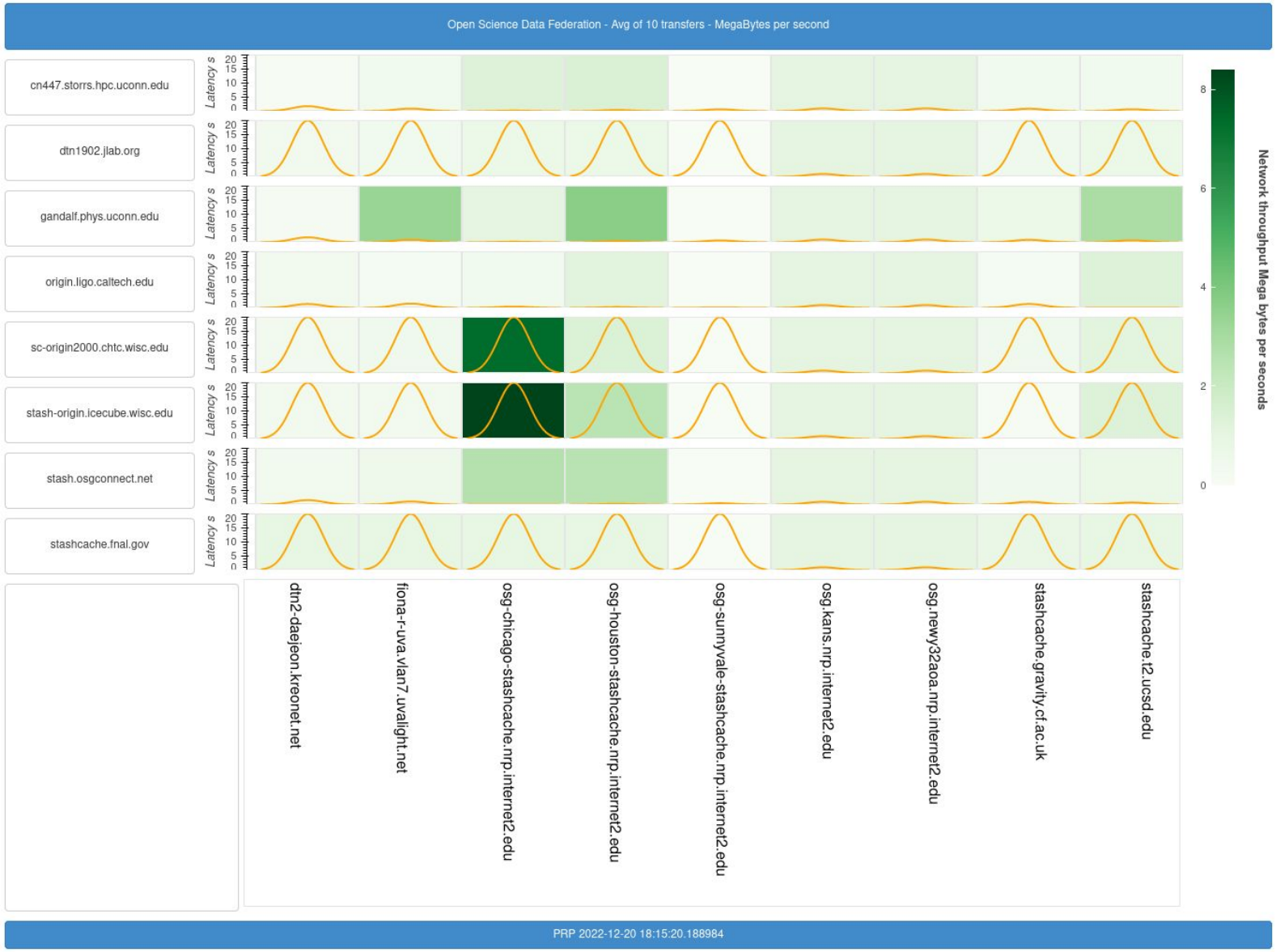
UCSD cache



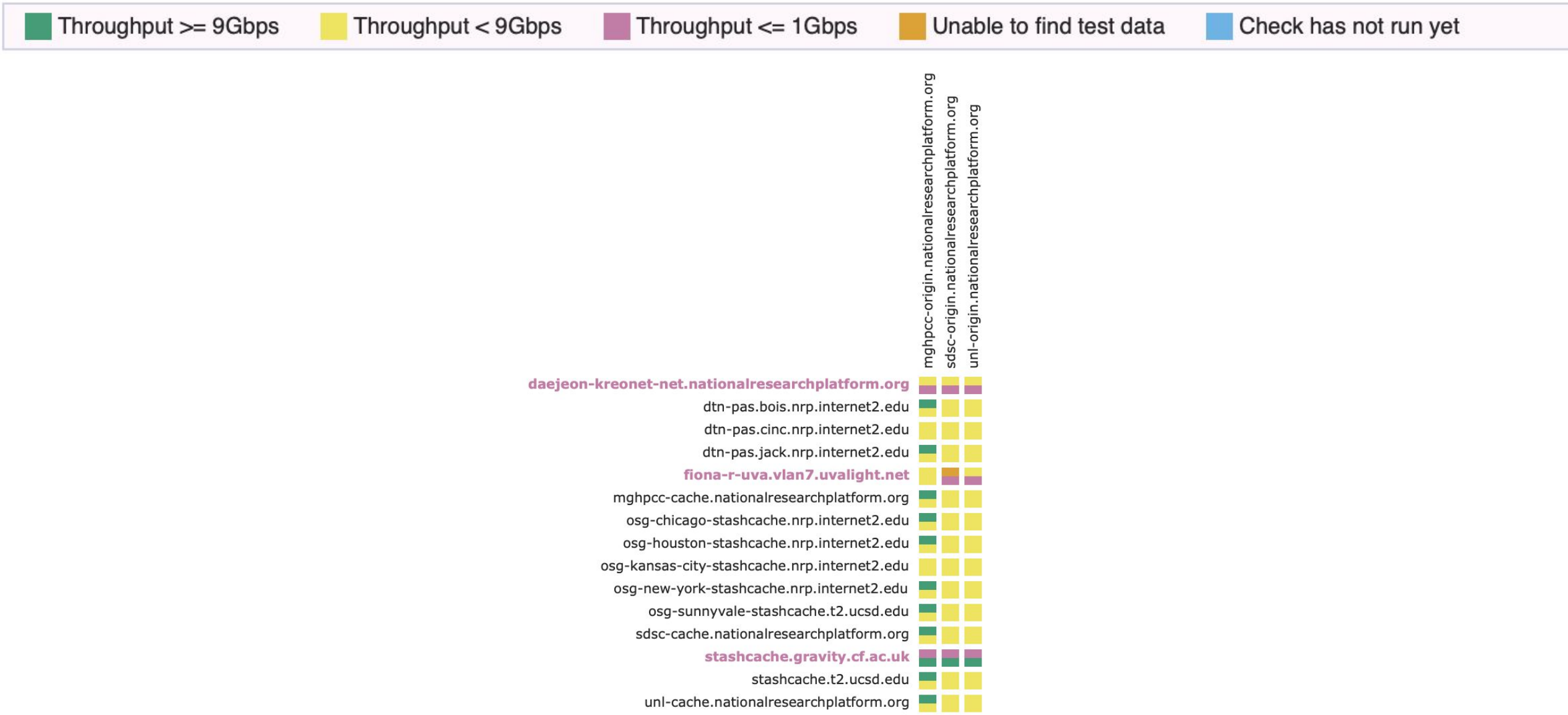
OSDF - Monitoring

OK	Check SSL certificate		OK - Certificate 'stashcache.t2.ucsd.edu' will expire on Sat Nov 4 23:59:59 2023 +0000.	17 h	12.8 s	
OK	HTTP Check CVMFS Access		HTTP OK: HTTP/1.1 200 OK - 94 bytes in 0.085 second response time	17 h	12.8 s	84.6 ms
OK	Read and Write check with Scitoken Authorization		file on origin test0.2924439967576046: file on origin test0.2924439967576046: HTTP/1.1 200 OK	2023-03-24 22:36:39	40.8 s	
OK	Shoveler processed msg		shoveler_packets_received = 7108974	25 h	52.8 s	
OK	Stash cache auth			17 h	12.8 s	
OK	Stash cache copy check with authentication		[' % Total % Received % Xferd Average Speed Time Time Time Current	17 h	7.78 s	
OK	Stash cache reports to GRACC in the past 4 days		GRACC received 727843 records	2023-02-06 21:25:02	12.8 s	
OK	Stash cache shoveler status		shovler_queue_size = 0	2023-03-23 16:31:51	12.8 s	
OK	XRootD copy		[13B/13B][100%][==========] [13B/s] [13B/13B][100%] [==========][13B/s]	17 h	3.79 s	
OK	XRootD copy per host		[8MB/10MB][80%][==========>][8MB/s] [10MB/10MB][100%][==========] [10MB/s] [10MB/10MB][100%] [==========][10MB/s]	17 h	3.79 s	
OK	XRootD StashCache cache Ping		OK	17 h	52.8 s	

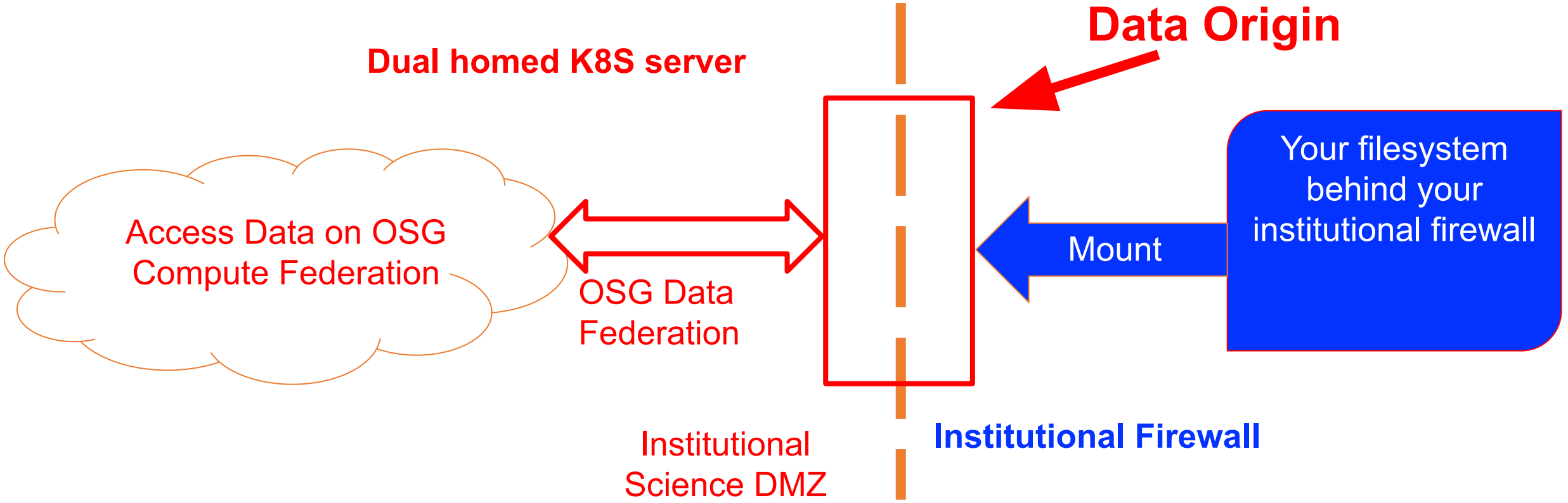
OSDF - Monitoring



OSDF - Monitoring



Campus Cyberinfrastructure (CC*)



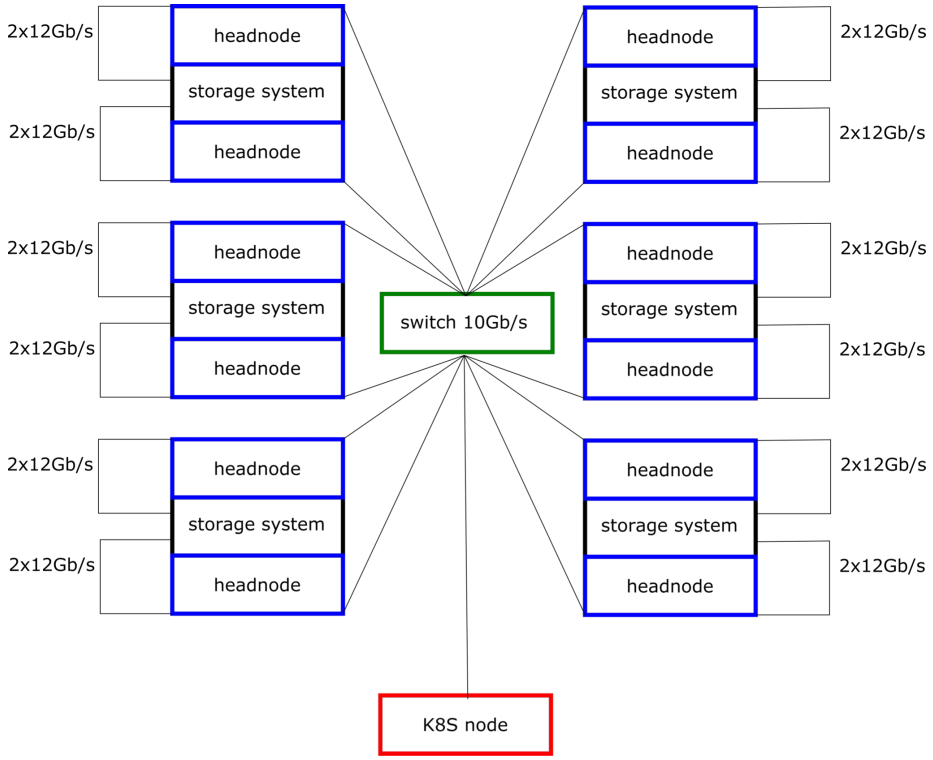
The devOps model and the actual containers used to make NSF 22-582 data origins possible were developed by IRIS-HEP for use by the LHC.

Campus Cyberinfrastructure (CC*)

CC* awards made via this solicitation will be supported in two program areas:

Data Storage awards will be supported at up to $500,000 total for up to 2 years

We created Storage Architecture suggestion to help the researcher to reach this award.



Storage Architecture suggestion

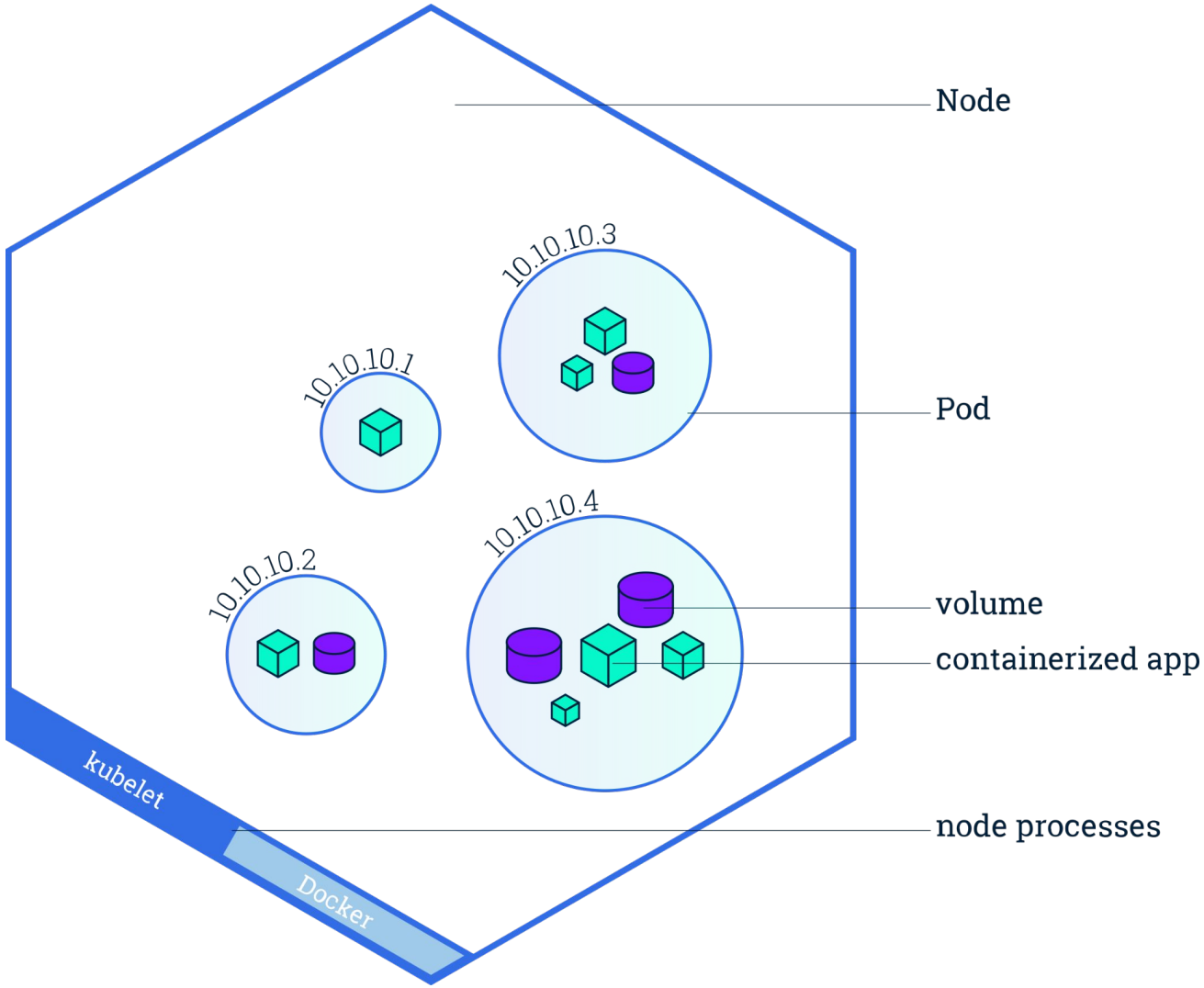
Open Science Data Federation - OSDF
Kubernetes - docker - containers

Fabio Andrijauskas

UCSD

K8S

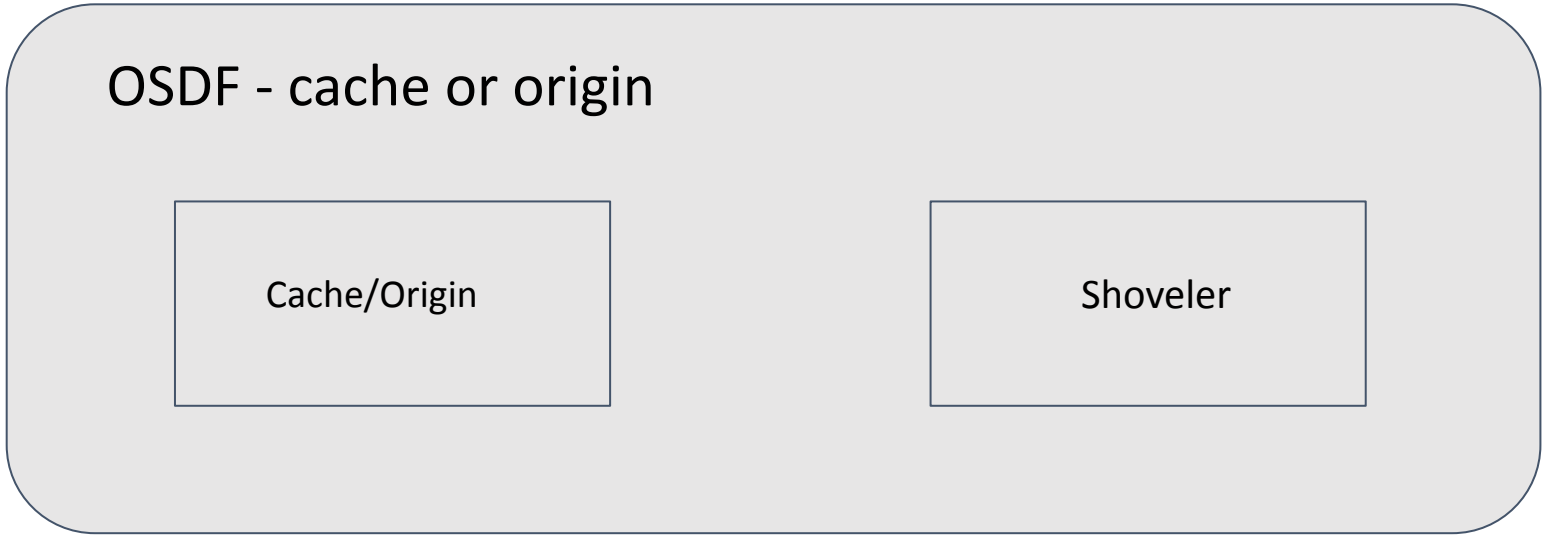
K8S


Node
10.10.10.3
10.10.10.1
Pod
10.10.10.4
10.10.10.2
volume
containerized app
kubelet
Docker
node processes

K8S - OSDF

```
- env:
    - name: XC_SHOVELER
      value: dtn-pas.jack.nrp.internet2.edu:9993
    - name: XRDHOST
      value: dtn-pas.jack.nrp.internet2.edu
    - name: XC_ROOTDIR
      value: /xcache
    - name: SITENAME
      value: JACKSONVILLE_INTERNET2_OSDF_CACHE
    - name: XC_RESOURCENAME
      value: JACKSONVILLE_INTERNET2_OSDF_CACHE
    - name: CACHE_FQDN
      value: dtn-pas.jack.nrp.internet2.edu
    - name: XC_RAMSIZE
      value: 10g
    - name: OSG_RESOURCE_NAME
      value: SDSC_NRP_OSDF_CACHE
  image: opensciencegrid/stash-cache:3.6-release
  imagePullPolicy: Always
  name: cache
  ports:
  - containerPort: 1094
```

K8S



K8S - OSDF

Good things:

- Easy way to change versions;
- Standard machine monitoring;
- Template to create new hosts;
- Easy form to rebuild the system;
- “Only” 2% overhead

K8S - OSDF

“Bad” things:

- Hostname and certificates (quick pool: who hates certs??)
- “Need” to use HostNetwork
- Log persistence

Acknowledgement

This material is based upon work supported by the National Science Foundation under Grant No. 2030508. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.