Running HEP Workloads on Distributed Clouds

R.Seuster, F. Berghaus, K. Casteels, C. Driemel
M. Ebert, C. R. Leavett-Brown, M. Paterson,
R.Sobie, T. Weiss-Gibson

2017 Fall HEPiX meeting, Tsukuba
16. - 20. Oct. 2017

Distributed clouds

What do we mean by distributed cloud ?

- required services run locally at UVic
  - in our case HTCondor and CloudScheduler
- real workloads / pilots run “anywhere” in the world
  - on OpenStack clouds operated by Compute Canada at UVic (west), Sherbrooke (east)
  - on commercial (MS, EC2) and opportunistic clouds in North America and Europe (none in Asia … yet ?)
- (usually) UVic’s storage element used for data, CERN for European clouds
- for experiments we look like any other grid site

Workflow

Workflow CloudScheduler

0) queues empty, no worker nodes running

- HTCondor and CloudScheduler (CS) are running at UVic

Workflow CloudScheduler

0) queues empty, no worker nodes running

- HTCondor and CloudScheduler (CS) are running at UVic

1) user submits jobs into queuing system

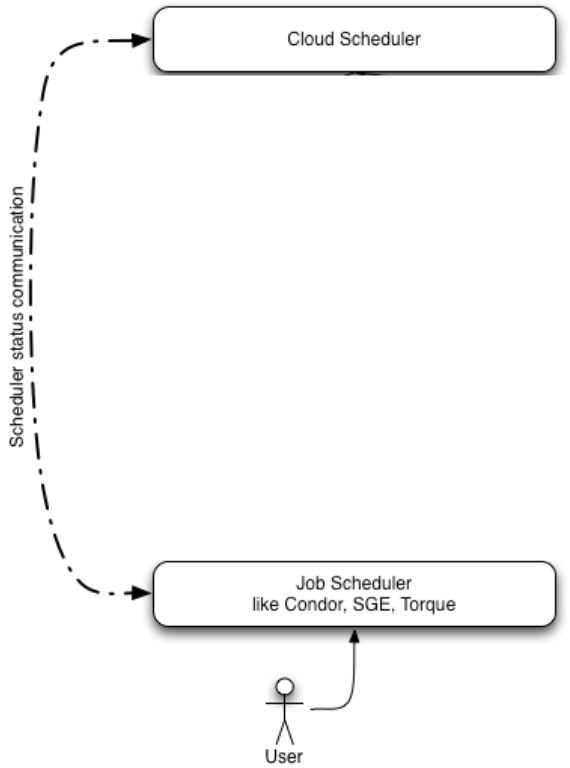
Workflow CloudScheduler

0) queues empty, no worker nodes running

- HTCondor and CloudScheduler (CS) are running at UVic

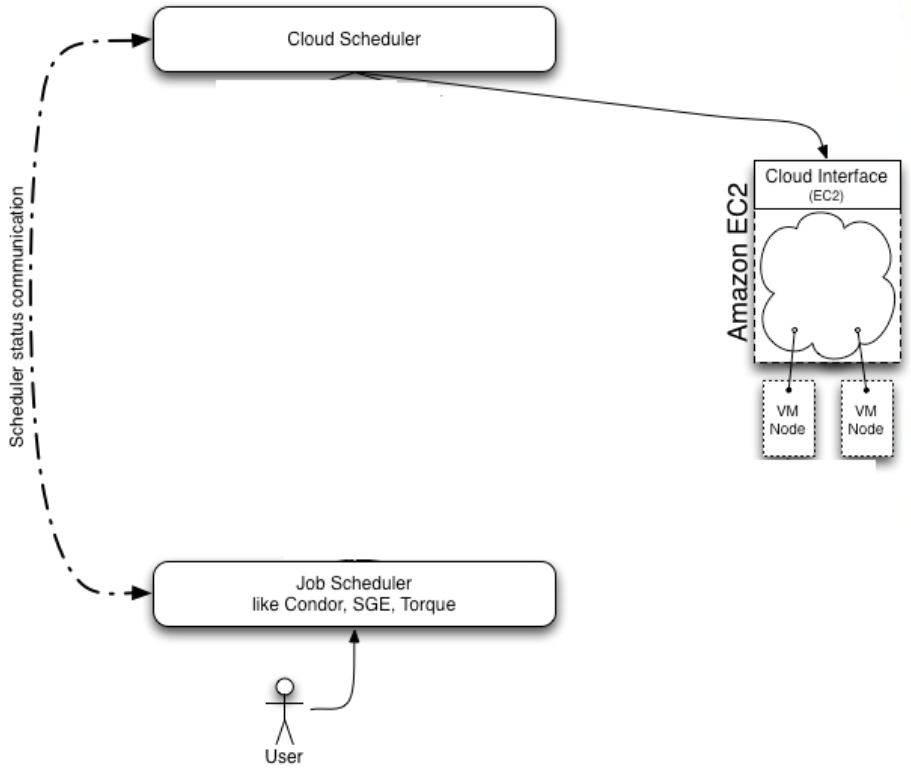
1) user submits jobs into queuing system

2) CS sees jobs in queue, and knows no (free) resources are available



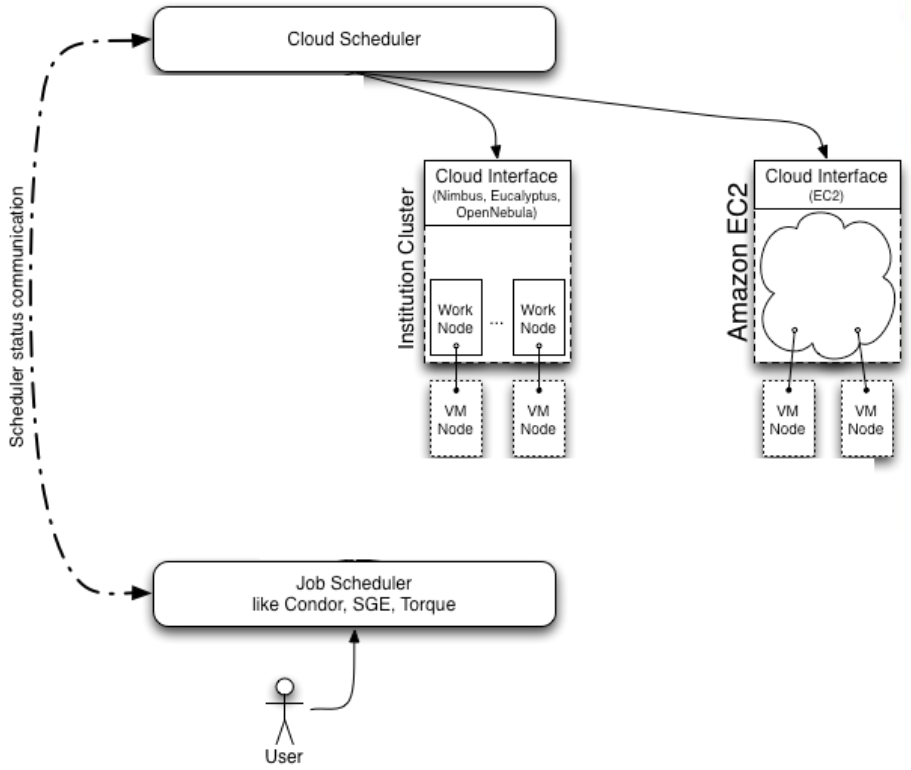
Workflow CloudScheduler

0) queues empty, no worker nodes running
   - HTCondor and CloudScheduler (CS) are running at UVic
1) user submits jobs into queuing system
2) CS sees jobs in queue, and knows no (free) resources are available
3) CS boots VM on any cloud which has matching resources
   - VMs could be anywhere in world



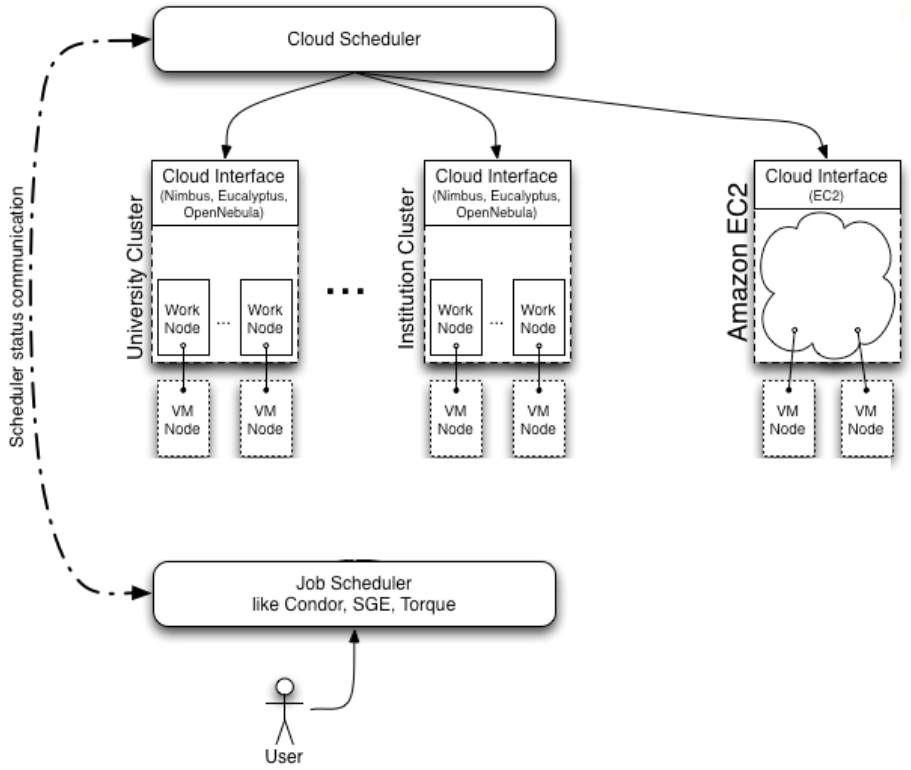
Workflow CloudScheduler

0) queues empty, no worker nodes running
   - HTCondor and CloudScheduler (CS) are running at UVic
1) user submits jobs into queuing system
2) CS sees jobs in queue, and knows no (free) resources are available
3) CS boots VM on any cloud which has matching resources
   - VMs could be anywhere in world



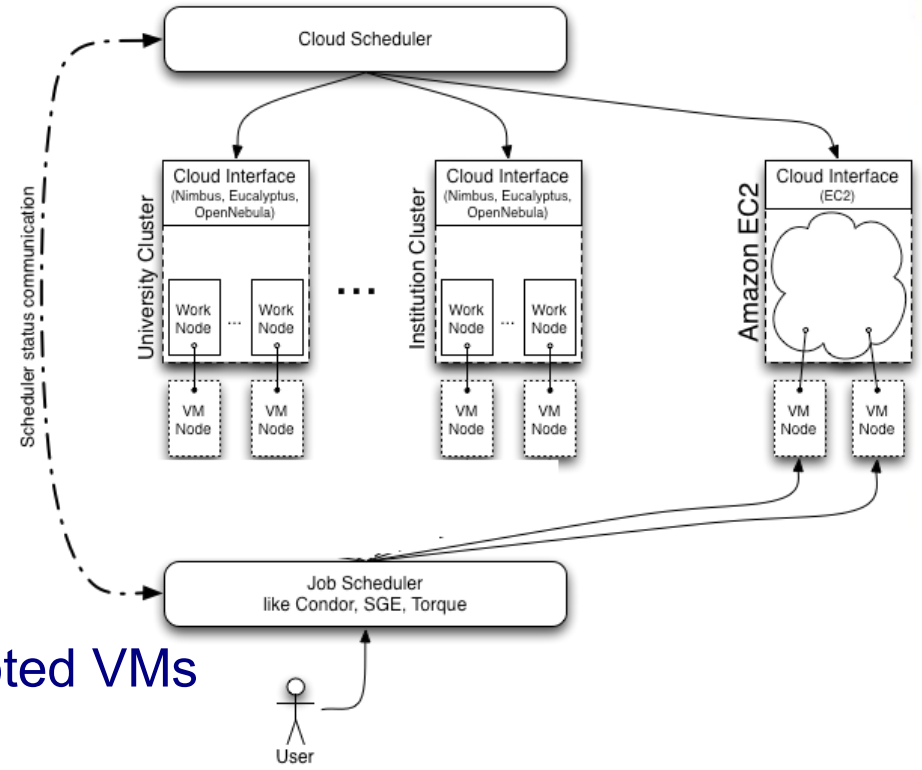
Workflow CloudScheduler

0) queues empty, no worker nodes running
   - HTCondor and CloudScheduler (CS) are running at UVic
1) user submits jobs into queuing system
2) CS sees jobs in queue, and knows no (free) resources are available
3) CS boots VM on any cloud which has matching resources
   - VMs could be anywhere in world



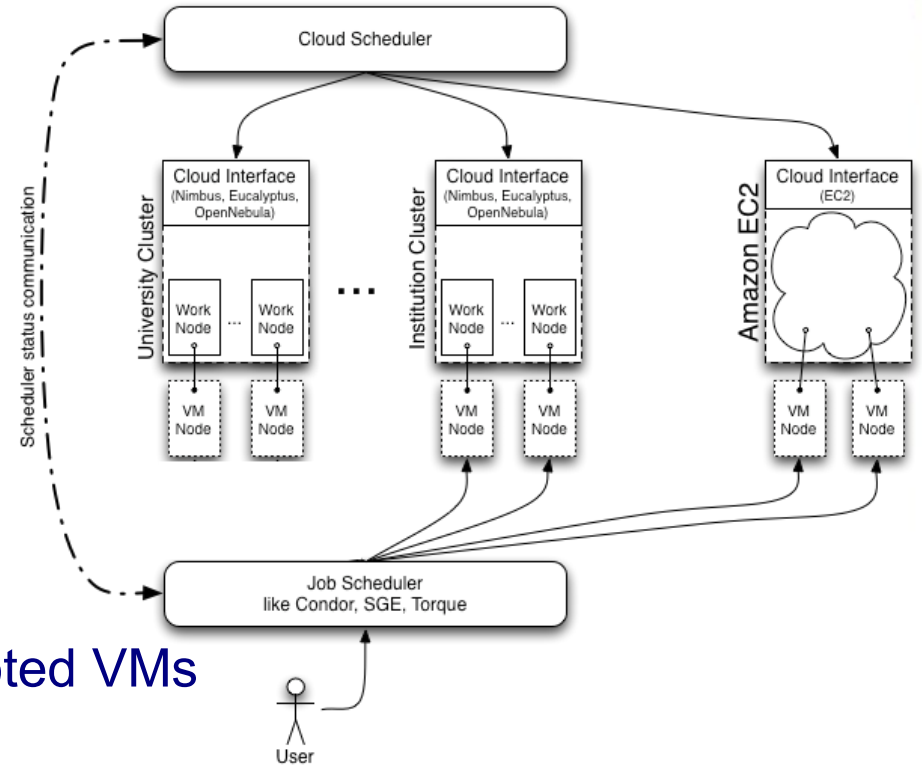
Workflow CloudScheduler

0) queues empty, no worker nodes running
   - HTCondor and CloudScheduler (CS) are running at UVic
1) user submits jobs into queuing system
2) CS sees jobs in queue, and knows no (free) resources are available
3) CS boots VM on any cloud which has matching resources
   - VMs could be anywhere in world
4) VMs boot and start HTCondor during their startup
   - jobs will start automatically on booted VMs



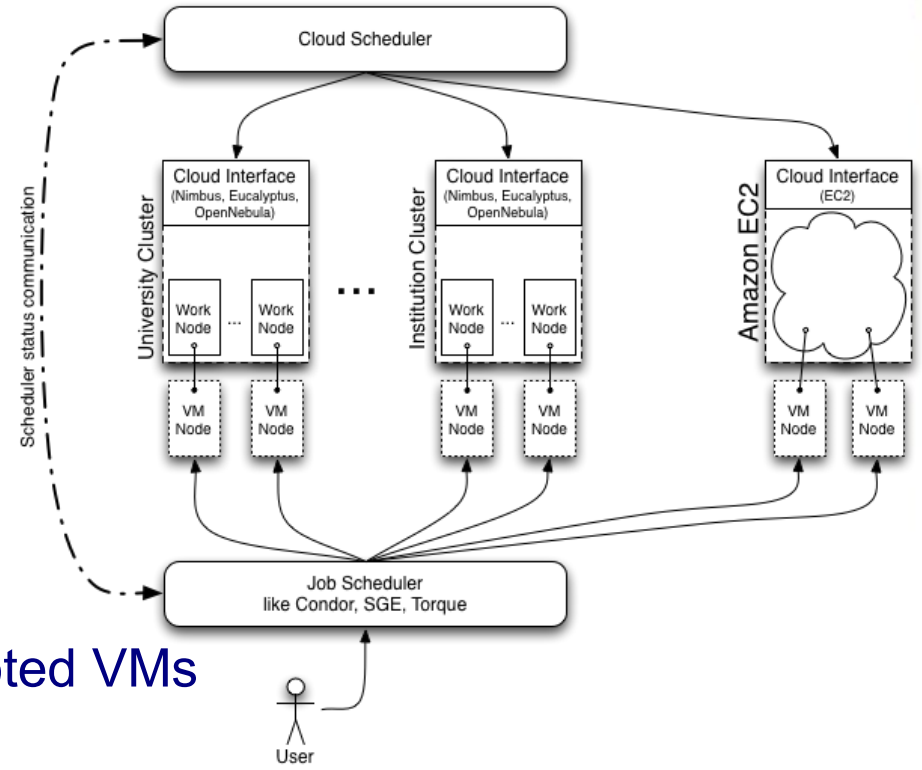
Workflow CloudScheduler

0) queues empty, no worker nodes running
   - HTCondor and CloudScheduler (CS) are running at UVic
1) user submits jobs into queuing system
2) CS sees jobs in queue, and knows no (free) resources are available
3) CS boots VM on any cloud which has matching resources
   - VMs could be anywhere in world
4) VMs boot and start HTCondor during their startup
   - jobs will start automatically on booted VMs



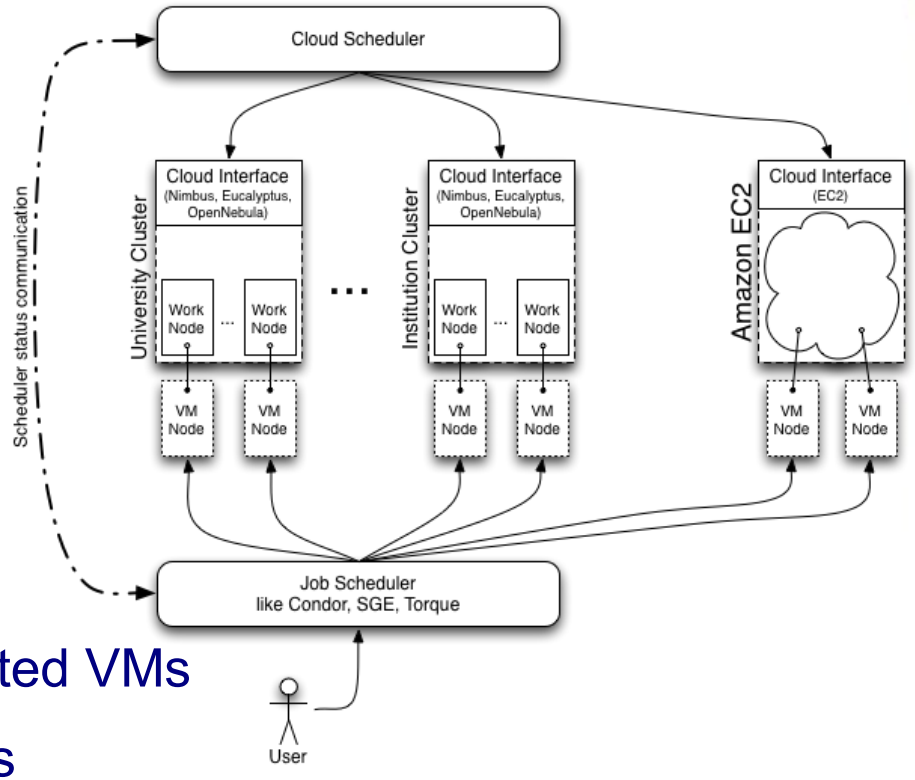
Workflow CloudScheduler

0) queues empty, no worker nodes running
   - HTCondor and CloudScheduler (CS) are running at UVic
1) user submits jobs into queuing system
2) CS sees jobs in queue, and knows no (free) resources are available
3) CS boots VM on any cloud which has matching resources
   - VMs could be anywhere in world
4) VMs boot and start HTCondor during their startup
   - jobs will start automatically on booted VMs



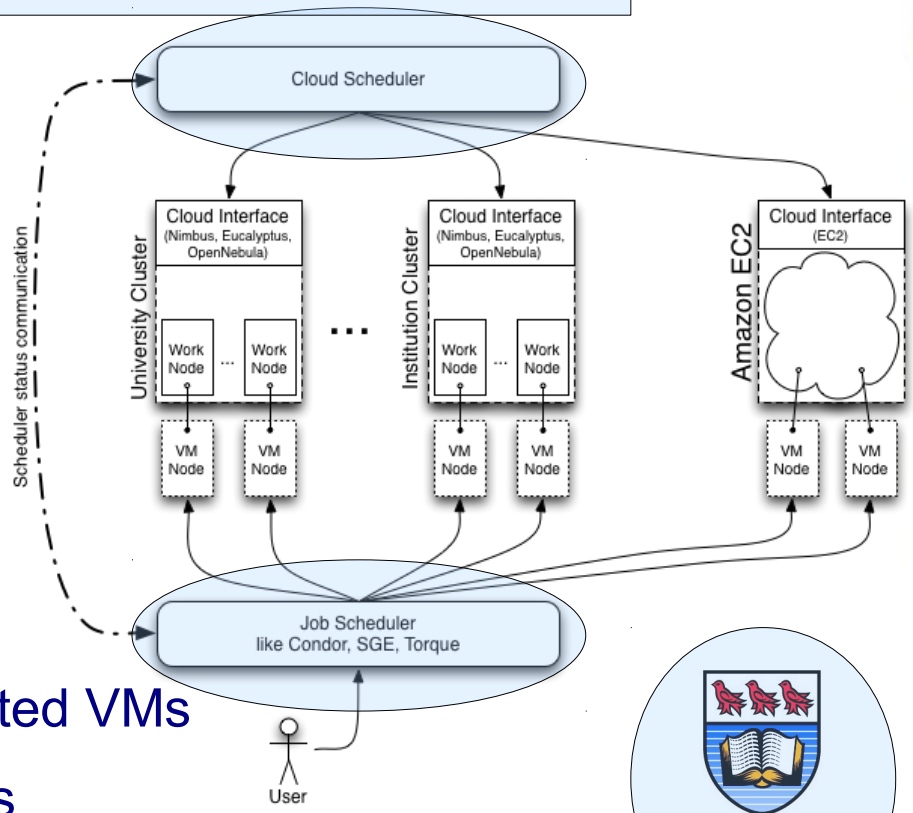
Workflow CloudScheduler

0) queues empty, no worker nodes running
   - HTCondor and CloudScheduler (CS) are running at UVic
1) user submits jobs into queuing system
2) CS sees jobs in queue, and knows no (free) resources are available
3) CS boots VM on any cloud which has matching resources
   - VMs could be anywhere in world
4) VMs boot and start HTCondor during their startup
   - jobs will start automatically on booted VMs
5) more jobs would populate empty job slots on running VMs or new VMs are booted by CS



Workflow CloudScheduler

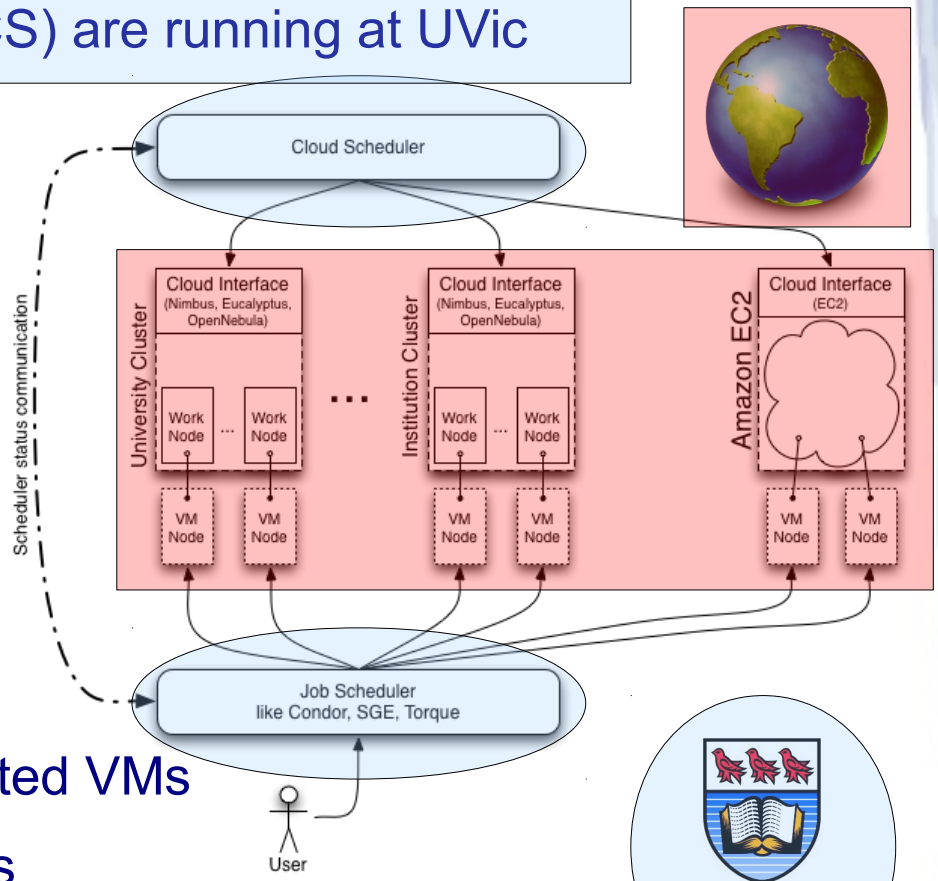
0) queues empty, no worker nodes running
   - HTCondor and CloudScheduler (CS) are running at UVic
1) user submits jobs into queuing system
2) CS sees jobs in queue, and knows no (free) resources are available
3) CS boots VM on any cloud which has matching resources
   - VMs could be anywhere in world
4) VMs boot and start HTCondor during their startup
   - jobs will start automatically on booted VMs
5) more jobs would populate empty job slots on running VMs or new VMs are booted by CS



University of Victoria

Workflow CloudScheduler

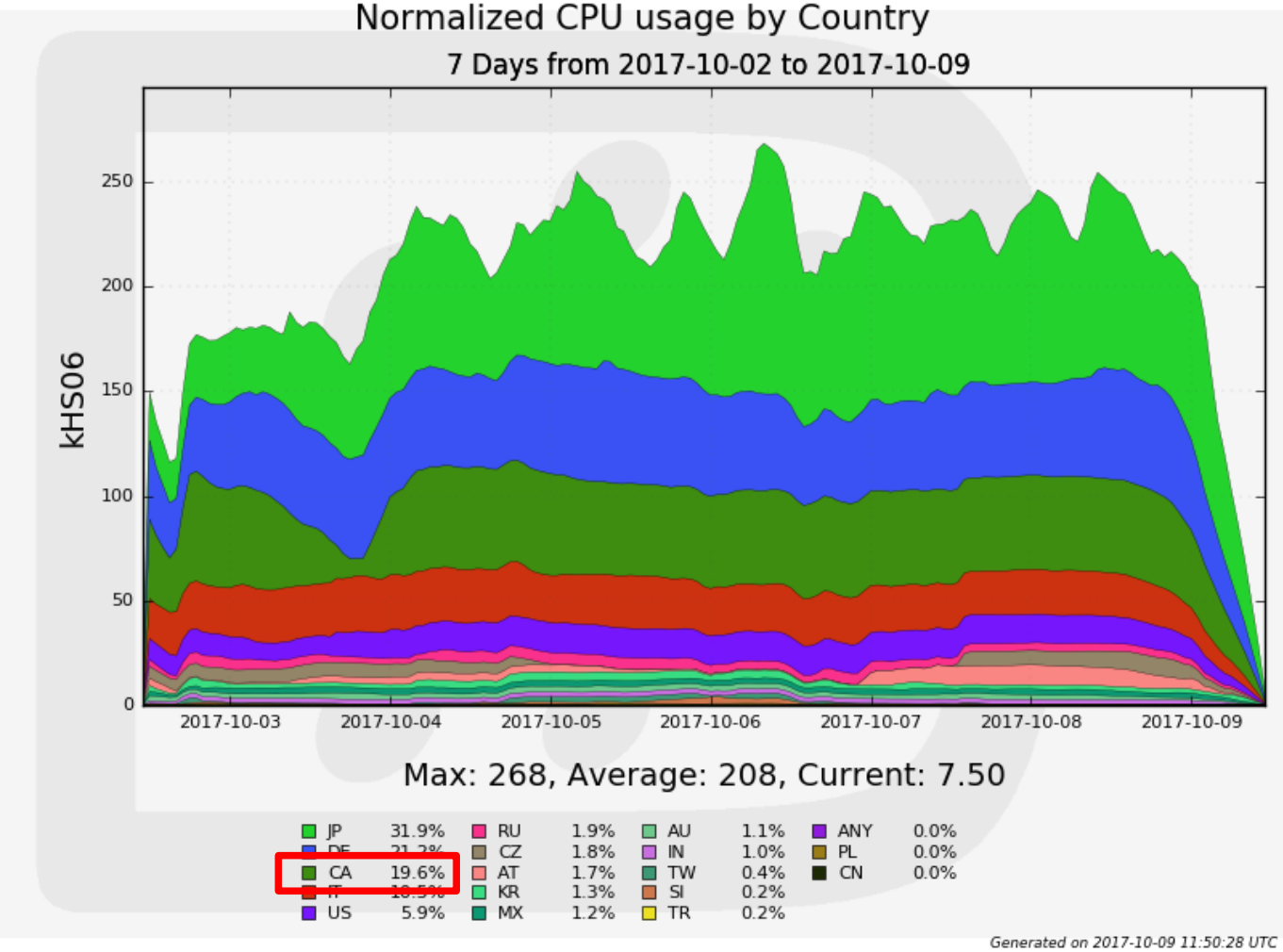
0) queues empty, no worker nodes running
   - HTCondor and CloudScheduler (CS) are running at UVic
1) user submits jobs into queuing system
2) CS sees jobs in queue, and knows no (free) resources are available
3) CS boots VM on any cloud which has matching resources
   - VMs could be anywhere in world
4) VMs boot and start HTCondor during their startup
   - jobs will start automatically on booted VMs
5) more jobs would populate empty job slots on running VMs or new VMs are booted by CS



University of Victoria

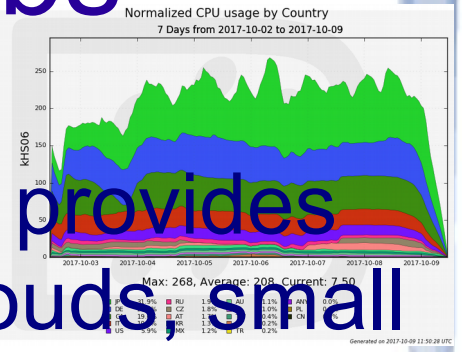
How ‘good’ are we ?

Number of Concurrent Jobs



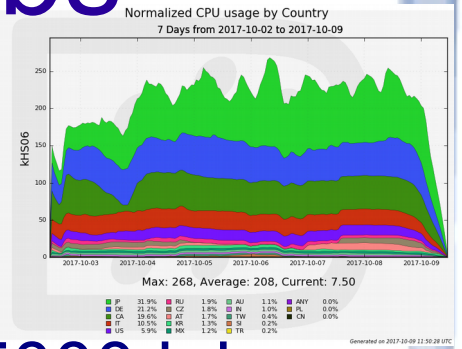
- for e.g. Belle II, from DIRAC. Canada provides most computing via our distributed clouds, small contribution from McGill from conventional CPU

Number of Concurrent Jobs

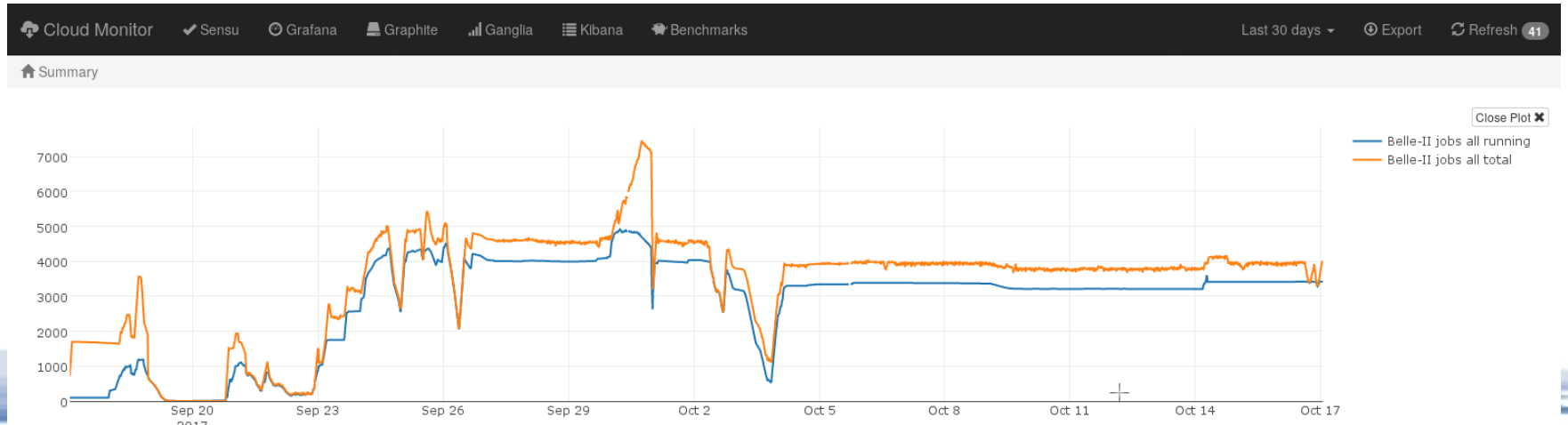


- for e.g. Belle II, from DIRAC. Canada provides most computing via our distributed clouds, small contribution from McGill from conventional CPU
- for last MC campaigns, we typically provided 20-25% of all Belle II computing resources in terms of CPU resources
- Belle II: all single core jobs
- ATLAS: mostly 8 (some 4) CPU multi-core jobs, small number of single core jobs

Number of Concurrent Jobs



- from our monitoring for Belle II
  - scale according to demand, maximum ~5000 jobs
  - one limit seems to be HTCondor (we are on WAN !)
  - for ATLAS we run two separate CloudScheduler with similar number of utilized number of CPU cores:
    one instance at CERN to operate European clouds
    one at UVic to operate North American clouds



How to make our life more easy

Automation

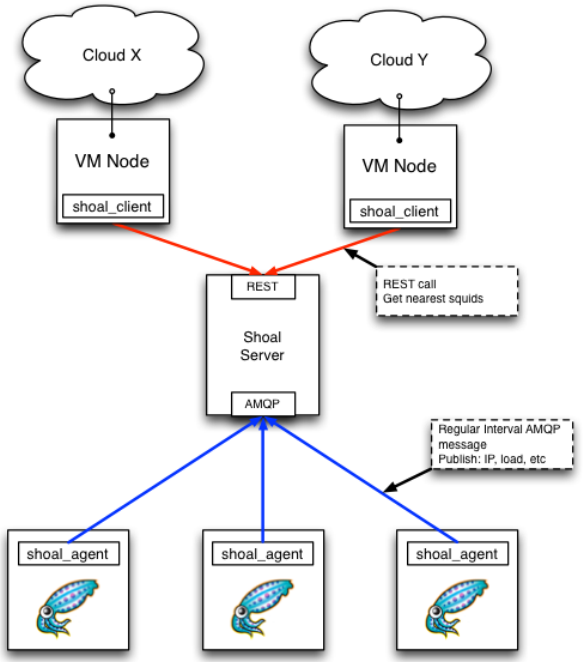
- to reduce manual intervention, we automate as much as possible at various stages:
  - cernvm images, configured for ATLAS/Belle II jobs
  - configure closest squid for Software and Condition Data via shoal
  - configure closest SE for input and (later also output) data via DynaFed
  - Accounting: Benchmarks and uptime
  - for Belle II also get application success / failure rate
  - far future: automatically detect and cure application problems on clouds

CERNVM images

- CERNVM get almost all (OS and experiments) software via CVMFS, a http read only filesystem that requires to configure close-by squids for efficient operation→ shoal
  - ATLAS publishes conditions data (MC) via CVMFS
- CloudScheduler injects contextualization via cloud-init into vanilla CERNVM images
- main configuration we require:
  - connect to HTCondor at UVic
  - ssh keys to allow for debugging problems
  - smaller cloud specific ‘adjustments’

Shoal – “School of Squids”

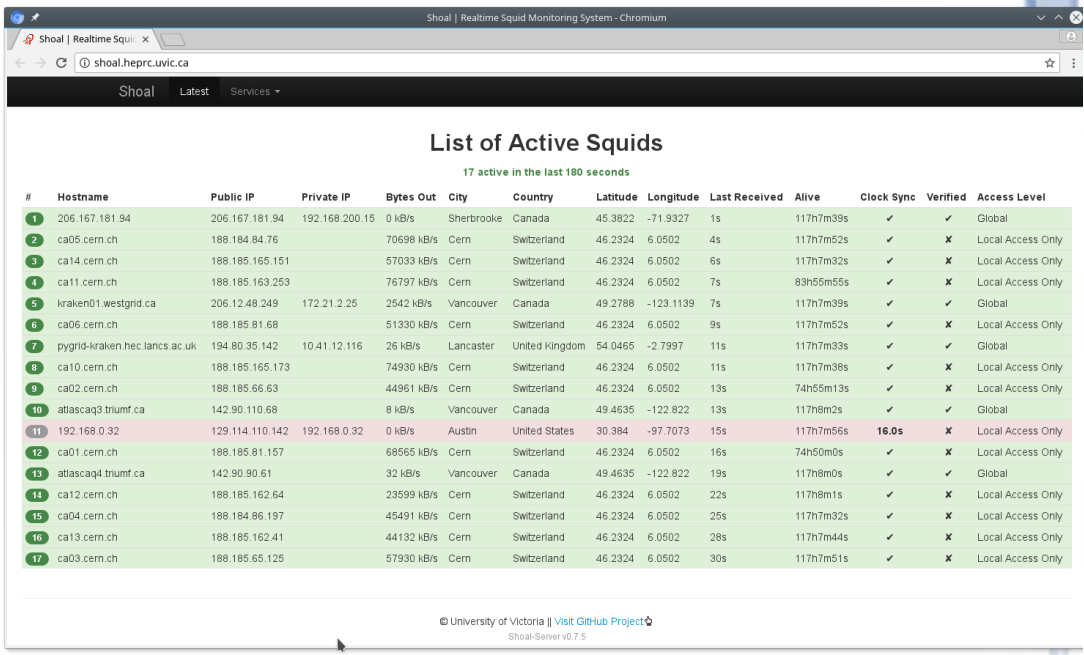
- all squid servers we utilize register with shoal-server at UVic http://shoal.heprc.uvic.ca/
  - register IP and load of squid server
- all VMs contact shoal-server to find their nearest squid - ‘nearest’ based on GeoIP DB
  - some squids are ‘private’ i.e serving only local cloud (e.g. due to private IP of VMs)
- basically it runs “curl http://shoal.heprc.uvic.ca/nearest”
  - e.g. at UVic: finds squid in Vancouver
  - amazon-east (Virginia): finds squid in Sherbrooke/CA



https://github.com/hep-gc/shoal

Shoal – “School of Squids”

- all distributed squids we are utilize register with shoal-server at UVic http://shoal.heprc.uvic.ca/
  - register IP and load
- all VMs regularly contact shoal-server to configure their nearest squid: distance based on GeoIP DB
  - some squids are ‘private’ i.e serving only local cloud (e.g. due to private IP of VMs), typically squid and VMs have same or similar public IP



https://github.com/hep-gc/shoal

DynaFed

- dynamic federation of storage elements
  - can federate many different SE protocols: http(s), webdav, S3, MS Azure
  - presents http or webdav interface to users, hides authentication to e.g. S3 via temporary tokens
- returns list of ‘closest’ locations based on GeoIP DB
- also supports writing, but ‘lower’ on our todo list

DynaFed

- dynamic federation of storage elements
  - can federate many different SE protocols: http(s), webdav, S3, MS Azure
  - presents http or webdav interface to user, hides authentication to e.g. S3 via temporary tokens
- ATLAS: test instance with own storage behind, trying to integrate into ATLAS DDM
- Belle II: also trying to integrate, but looking also at using DynaFed to store background files locally on clouds we utilize

→ more in Marcus' Talk on Thursday

Accounting

Benchmark and Accounting

- running fast benchmark from working group
- also reporting uptime and user time for VMs
  - every 15 minutes, because VMs might get killed due to spot market pricing
  - allows us to identify clouds which provide ‘unreliable’ resources, usually requires manual investigation, doesn’t happen often
- would allow us to provide resource utilization if required (cloud xyz belongs to region abc...)

https://bmkwg.web.cern.ch/bmkwg/

Application Success/Failure for Belle II

- recent addition: Store JobID of all jobs and inquire DIRAC about application status
- cloud dependent: allows us to identify clouds with frequent failures
  - most recent: beam background files caused lots of problems on small number of clouds, but causing vast amounts of failures – jobs stalled for ~30mins vs successful jobs ran for >12h

Application Success/Failure for Belle II

- recent addition: Store JobID of all jobs and inquire DIRAC about application status

21-09-2017
cloud-1: Completed:0 Deleted:0 Done:78 Failed:32 Killed:0 Running:0 Stalled:0
cloud-2: Completed:0 Deleted:0 Done:0 Failed:16 Killed:0 Running:0 Stalled:0
cloud-3: Completed:0 Deleted:0 Done:192 Failed:80 Killed:0 Running:0 Stalled:0
cloud-4: Completed:0 Deleted:0 Done:1390 Failed:1300 Killed:0 Running:1 Stalled:0
cloud-5: Completed:0 Deleted:0 Done:0 Failed:72 Killed:0 Running:0 Stalled:0
Overall: Completed:0 Deleted:0 Done:1660 Failed:1500 Killed:0 Running:1 Stalled:0

- similar feature for ATLAS implemented but not in production yet

The Future …

CloudScheduler V2

- re-write to bring software development into 21st century:
  - unit tests, Jenkins, CI, ...
- overcome scaling issue at O(10^4) jobs
  - could flocking help for far clouds ?
- update ‘dependencies’ of software like rabbitmq
- common interfaces with other cloud provisioning solutions to ease integration ?
- containers for software disitrbution + workload

More Docs as blogs

shoal:
http://heprc.blogspot.com/2017/10/shoal-squid-proxy-discovery-and.html

glint:
http://heprc.blogspot.com/2017/08/glint-version-2-enters-production.html

adding squid to shoal:
http://heprc.blogspot.com/2017/07/adding-your-squid-to-shoal.html

azure usage:
http://heprc.blogspot.com/2017/06/microsofts-azure-cloud-for-high-energy.html

dynafed:
http://heprc.blogspot.com/2017/06/grid-mapfile-based-authentication-for.html

Summary and Outlook

- long history of successful operation of distributed clouds – experiments don't realize, they run on clouds
- always looking for new resources, latest addition OpenNebula Cluster in Munich
  - operate computing for small sides ? 'just' installation of OpenStack / OpenNebula
  - boinc like service for Belle II ?

Backup Slides

shoal nearest at UVic

$ curl http://shoal.heprc.uvic.ca/nearest

{"0": {"load": 17069, "domain_access": true, "squid_port": 3128, "global_access": true, "verified": true, "last_active": 1508218317.190254, "created": 1508037771.9229009, "external_ip": null, "geo_data": {"city": "Vancouver", "region_code": "BC", "area_code": 0, "time_zone": "America/Vancouver", "dma_code": 0, "metro_code": null, "country_code3": "CAN", "latitude": 49.27879999999999, "postal_code": "V6B", "longitude": -123.1139, "country_code": "CA", "country_name": "Canada", "continent": "NA"}, "hostname": "kraken01.westgrid.ca", "public_ip": "206.12.48.249", "private_ip": "172.21.2.25", "max_load": 122000, "distance": 0.0024112797351785562}, "1": {"load": 2, "domain_access": true, "squid_port": 3128, "global_access": true, "verified": true, "last_active": 1508218338.649739, "created": 1508037767.3524401, "external_ip": null, "geo_data": {"city": "Vancouver", "region_code": "BC", "area_code": 0, "time_zone": "America/Vancouver", "dma_code": 0, "metro_code": null, "country_code3": "CAN", "latitude": 49.46350000000001, "postal_code": "V6T", "longitude": -122.822, "country_code": "CA", "country_name": "Canada", "continent": "NA"}, "hostname": "atlascaq3.triumf.ca", "public_ip": "142.90.110.68", "private_ip": null, "max_load": "122000", "distance": 0.0030458533010027382}, "2": {"load": 0, "domain_access": true, "squid_port": 3128, "global_access": true, "verified": true, "last_active": 1508218335.9935689, "created": 1508037770.3463099, "external_ip": null, "geo_data": {"city": "Vancouver", "region_code": "BC", "area_code": 0, "time_zone": "America/Vancouver", "dma_code": 0, "metro_code": null, "country_code3": "CAN", "latitude": 49.46350000000001, "postal_code":"V6T", "longitude": -122.822, "country_code": "CA", "country_name": "Canada", "continent": "NA"}, "hostname": "atlascaq4.triumf.ca", "public_ip": "142.90.90.61", "private_ip": null, "max_load": "122000", "distance": 0.0030458533010027382}, "3": {"load": 0, "domain_access": true, "squid_port": 3128, "global_access": true, "verified": true, "last_active": 1508218312.902683, "created": 1508037768.167448, "external_ip": "206.167.181.94", "geo_data": {"city": "Sherbrooke", "region_code": "QC", "area_code": 0, "time_zone": "America/Montreal", "dma_code": 0, "metro_code": null, "country_code3": "CAN", "latitude": 45.382200000000012, "postal_code": "J1K", "longitude": -71.932699999999997, "country_code": "CA", "country_name": "Canada", "continent": "NA"}, "hostname": "206.167.181.94", "public_ip": "206.167.181.94", "private_ip": "192.168.200.15", "max_load": "122000", "distance": 0.096127539420955088}, "4": {"load": 0, "domain_access": true, "squid_port": 3128, "global_access": true, "verified": true, "last_active": 1508218336.935617, "created": 1508037782.22193, "external_ip": null, "geo_data": {"city": "Lancaster", "region_code": "H2", "area_code": 0, "time_zone": "Europe/London", "dma_code": 0, "metro_code": null, "country_code3": "GBR", "latitude": 54.046500000000009, "postal_code": "LA1", "longitude": -2.7997000000000014, "country_code": "GB", "country_name": "United Kingdom", "continent": "EU"}, "hostname": "pygrid-kraken.hec.lancs.ac.uk", "public_ip": "194.80.35.142", "private_ip": "10.41.12.116", "max_load": "122000", "distance": 0.18320660378475304}}

shoal nearest on Amazon-east

ssh 54.224.171.5 curl http://shoal.heprc.uvic.ca/nearest

{"0": {"load": 1, "domain_access": true, "squid_port": 3128, "global_access": true, "verified": true, "last_active": 1508218596.104032, "created": 1508037768.167448, "external_ip": "206.167.181.94", "geo_data": {"city": "Sherbrooke", "region_code": "QC", "area_code": 0, "time_zone": "America/Montreal", "dma_code": 0, "metro_code": null, "country_code3": "CAN", "latitude": 45.382200000000012, "postal_code": "J1K", "longitude": -71.932699999999997, "country_code": "CA", "country_name": "Canada", "continent": "NA"}, "hostname": "206.167.181.94", "public_ip": "206.167.181.94", "private_ip": "192.168.200.15", "max_load": "122000", "distance": 0.02095165777968146}, "1": {"load": 0, "domain_access": true, "squid_port": 3128, "global_access": false, "verified": false, "last_active": 1508218619.0026579, "created": 1508207893.932693, "external_ip": "129.114.110.142", "geo_data": {"city": "Austin", "region_code": "TX", "area_code": 512, "time_zone": "America/Chicago", "dma_code": 635, "metro_code": "Austin, TX", "country_code3": "USA", "latitude": 30.383999999999986, "postal_code": "78758", "longitude": -97.707300000000004, "country_code": "US", "country_name": "United States", "continent": "NA"},"hostname": "192.168.0.32", "public_ip": "129.114.110.142", "private_ip": "192.168.0.32", "max_load": "122000", "distance": 0.051928479934898311}, "2": {"load": 0, "domain_access": true, "squid_port": 3128, "global_access": true, "verified": true, "last_active": 1508218618.119164, "created": 1508037767.3524401, "external_ip": null, "geo_data": {"city": "Vancouver", "region_code": "BC", "area_code": 0, "time_zone": "America/Vancouver", "dma_code": 0, "metro_code": null, "country_code3": "CAN", "latitude": 49.46350000000001, "postal_code": "V6T", "longitude": -122.822, "country_code": "CA", "country_name": "Canada", "continent": "NA"}, "hostname": "atlascaq3.triumf.ca", "public_ip": "142.90.110.68", "private_ip": null, "max_load": "122000", "distance": 0.093165780218136679}, "3": {"load": 0, "domain_access": true, "squid_port": 3128, "global_access": true, "verified": true, "last_active": 1508218615.4703889, "created": 1508037770.3463099, "external_ip": null, "geo_data": {"city": "Vancouver", "region_code": "BC", "area_code": 0, "time_zone": "America/Vancouver", "dma_code": 0, "metro_code": null, "country_code3": "CAN", "latitude": 49.46350000000001, "postal_code": "V6T", "longitude": -122.822, "country_code": "CA", "country_name": "Canada", "continent": "NA"}, "hostname": "atlascaq4.triumf.ca", "public_ip": "142.90.90.61", "private_ip": null, "max_load": "122000", "distance": 0.093165780218136679}, "4": {"load": 505, "domain_access": true, "squid_port": 3128, "global_access": true, "verified": true, "last_active": 1508218596.708499, "created": 1508037771.9229009, "external_ip": null, "geo_data": {"city": "Vancouver", "region_code": "BC", "area_code": 0, "time_zone": "America/Vancouver", "dma_code": 0, "metro_code": null, "country_code3": "CAN", "latitude": 49.27879999999999, "postal_code": "V6B", "longitude": -123.1139, "country_code": "CA", "country_name": "Canada", "continent": "NA"}, "hostname": "kraken01.westgrid.ca", "public_ip": "206.12.48.249", "private_ip": "172.21.2.25", "max_load": 122000, "distance": 0.093691056410059564}}