Improving Science Gateways usage reporting for XSEDE

Amit Chourasia, Scott Sakai, Michael Shapiro, Steven Gallo & XCI team

https://xsede-xdcdb-api.xsede.org/api/gateways
Background

**Gateway attributes** provide complimentary information about jobs that were run under a community account on XSEDE

- Report for whom the job was actually run (aka gateway user)

XSEDE mandates Gateways to report this information

Trivia: Gateway attributes reporting was originally developed during TeraGrid project.
Background: Attribute compilation

Service Providers report job info
- Job ID – Slurm/PBS
- Username - Community account
- Usage
- Allocation
- Node info, etc

Gateways report job attributes
- Job ID – Slurm/PBS
- Gateway name
- Gateway user name

XDCDB
Combines info received from SP & Gateway attributes based on Job ID
Importance of Gateway attributes

For Gateways
- Better visibility of SU burn rate
- View, report and share usage data via XDMoD interface with Sponsor

For each Service Provider (SP) aka site hosting the cluster
- Number of Gateways users using a system
- Usage by individual gateway users
- Better understanding of how Gateways are using the cluster
- Report this information to sponsor

For XSEDE
- Same as for SPs
- Better handle on how various clusters are being used to meet community needs
- Visibility into the breadth of users using XSEDE

XSEDE user Trend

Active users
Red: HPC users
Orange: Gateway users
Shortcomings of existing reporting

Two ways to report gateway attributes
- a) direct writing to the database (legacy, but still possible to use)
- b) post using web services with two different scripts

XSEDE Service Providers

Must act as brokers and enable submission via custom scripts that they
- Install
- Manage
- Document
- No ability to audit who submitted the attributes

Gateway Operators

- Submission methods vary between SPs, and Gateway operators must know this
- Can only submit attributes from login nodes only
- Issues with data validation and consistency checks
- No feedback to Gateway operator on submission
- Cannot check prior submission (submit a ticket)
Refinements: Server side

Key effort to

- Better match submission semantics
- Add ability to identify and enforce data consistency and strengthen security by storing additional information
- Provide better and actionable feedback on incorrect usage
Refinements: Gateway friendliness

Cater well to the needs of the gateway operators

- Self service
- Single documentation point for submission on any XSEDE resource
- Semantic arguments for submission
  (hide cryptic stuff XA-header, XA-agent, XA-Resource ID, etc.)
- Origin/tool agnostic submission
Refinements

- Eliminate hassle for SP’s to broker and manage submission setup
- Better integration with XDMoD
Implementation

A new RESTful interface developed with following features

- Register Gateway via XSEDE single sign-on
- Assignment of unique API key to each Gateway
- Submission is origin/tool agnostic
- The API provides feedback for success or failure.
- Stores the API version to keep track of attribute submission lineage.
Implementation

Following set of attributes are supported

- Gateway User – The Gateway user who initiated computation
- XSEDE Resource Name – Identity of the resource where the computation was performed
- Job ID – Identity of the compute job as set by the resource provider
- Submit Time – The time job was submitted for computation
- Software – Optional list of software that were used in computation
- Usage – Time consumed by the compute job. This is needed for Cloud resource like JetStream that does not report granular usage information.
- VMID – Identity of the virtual machine where the computation was performed, needed to supported Cloud resource like JetStream
Software design and delivery documentation

XCI-419- Gateway attribute submission improvements

Reference
https://software.xsede.org/display/XCI-419
Real submission example

```bash
$ curl -XPOST --data @$HOME/.xsede-gateway-attributes-apikey \
--data-urlencode "gatewayuser=John Doe" \
--data-urlencode "xsederesourcename=comet-gpu.sdsc.xsede" \
--data-urlencode "jobid=12345" \
--data-urlencode "submittime=2019-01-31 23:45 PST" \
--data-urlencode "software=MATLAB:MATLAB R2019a, NumPy:NumPy v1.17" \
https://xsede-xdcdb-api.xsede.org/gateway/v2/job_attributes
```

Sample response

Job: 12345
Job submitted: 2019-01-31 23:45:00-0800
Resource Name: comet-gpu.sdsc.xsede
Gateway Name: Tesseract
Gateway User: John Doe
Software: MATLAB:MATLAB R2019a, NumPy:NumPy v1.17
Outcomes

- New implementation rolled out to production few weeks back
- Single Documentation Source https://xsede-xdcdb-api.xsede.org/api/gateways
- Poster at PEARC19
- Paper only in PEARC19 proceedings

SP/XSEDE benefits

- How Gateways are using a system?
- Number of Gateways users using a system.
- Usage by individual gateway users.
- Ability to view highly used software(s).
- View and share reporting of above data via XDMoD interface with Sponsor.
Gateway operator benefits

- Consistent & portable way to report across all XSEDE resources.
- Consistent documentation.
- New ability to report software(s) used.
- View and share reporting of above data via XDMoD interface with Sponsor.

Project flashback

1. Discovery
2. Gateway operator Point of View (PoV)
3. XDCDB PoV
4. SP PoV
5. XDMoD PoV
6. ECSGW PoV
7. Globus PoV
8. New Design and consensus building
9. Development and deployment requirements
10. Private implementation & testing
11. XCI – Design, review, implementation, testing, deployment
12. Production deployment
Tasks ahead

• Shepherd gateway operators to switch to new method by Nov 30, 2019 deadline

Two webinars scheduled Sep 18, Oct 1

• Hand off information to XDMoD team

• JetStream support needs to be tackled
  • Discussions underway
Acknowledgements

• Nancy Wilkins-Diehr (project initiator), Marlon Pierce, Robert Sinkovits

• (Alphabetic) Christopher Irving, David Hart, Eric Coulter, Eroma Abheysinghe, Ester Soriano, Lee Leming, Mahidhar Tatineni, Maytal Dahan, Nicole Wolter, Sudhakar Pamidighantam, Victor Hazelwood

• XCI Team: JP Navarro, Shava Smallen, Choonhan Yoon, Jim Basney and other reviewers