Retrieve Provenance

# Introduction

Retrieving provenance is an important step in the provenance lifecycle. There are several ways to retrieve provenance information from ProvEn. It uses SPARQL query language to get information from ProvEn’s semantic store. ProvEn provides REST services wrapping some of the commonly used SPARQL queries. User can write native SPARQL queries and use ProvEn’s REST service to send the query to ProvEn’s semantic store and retrieve the information as well. ProvEn team has developed Python script to retrieve the E3SM provenance for a simulation. This script will retrieve all the provenance information for a simulation and re-creates the files used for that simulation. It also creates a README file which lists the details of the simulation (e.g. simulation name, machine name)

# Prerequisites

* Python 2.7
* Access to ProvEn server

# Procedure

## Dependencies on other Quick Start Guides

In order to retrieve provenance information from ProvEn, the provenance information should be harvested and stored in ProvEn store. Refer quick start guides on Harvesting Provenance and Storing provenance.

## Steps

1. Identify the retrieve provenance script. It’s located at <https://github.com/pnnl/ProvenanceEnvironment/blob/master/examples/E3SM/download_provenance.py>
2. Identify the target directory where the provenance information will be downloaded
3. Identify the simulation name for which you need to retrieve provenance
4. Know the Proven server URL from which you want to download provenance
5. Run the python script

e.g.

python download\_provenance.py --simname="Try1.Run1.ne4\_ne4" --targetdir="/export/raju1/provenance\_download" --server=<http://localhost:28080>

## Next Quick Start Guide

Reproduce simulation

# Questions?

Contact: [Bibi.Raju@pnnl.gov](mailto:Bibi.Raju@pnnl.gov), [Todd.Elsethagen@pnnl.gov](mailto:Todd.Elsethagen@pnnl.gov), [Eric.Stephan@pnnl.gov](mailto:Eric.Stephan@pnnl.gov)