

NASA Workflow Tool

User Guide

September 29, 2010

[NASA Workflow Tool](#)

[User Guide](#)

[1. Overview](#)

[2. Getting Started](#)

[Preparing the Environment](#)

[3. Using the NED Client](#)

[Common Terminology](#)

[Workflow Configuration](#)

[Groups](#)

[Variables](#)

[Basic Operations](#)

[Logging In](#)

[Loading A Workflow](#)

[Choosing Who To Run As](#)

[Configuring a Workflow](#)

[4. Running a Workflow](#)

[Monitoring A Workflow](#)

[Saving A Workflow](#)

[5. The NED Interface](#)

[Main Window GUI Components](#)

[Configuration Tree](#)

[Configuration](#)

[Status](#)

[Making Tree Selections](#)

[Other Tree Functions](#)

[Menu Bar](#)

[File Menu](#)

[Search Menu](#)

[Tools Menu](#)

[Run Menu](#)

[Views Menu](#)

[Help](#)

[Tool Bar](#)

[Open Workflow *](#)

[Save Workflow *](#)

[Find In Workflow *](#)

[Submit With Validation *](#)

[Submit Without Validation *](#)

[Validate Workflow *](#)

[6. Advanced NED Features](#)

[SSH Key Management](#)

[External Tool Management](#)

[Differencing Workflows](#)

1. Overview

The Workflow tool was created by NASA Goddard Space Flight Center's [Software Integration and Visualization Office](#) through its MAP Modeling Environment work. The tool is designed to configure, run, monitor and manage complex model experiments for NASA scientists. Users manipulate experiments via an easy-to-use GUI rather than configuring scripts and executing system commands. The system takes care of scheduling the individual tasks that make up the workflow to run in the proper order. Such tasks include source check-out, building the executable, execution, post-processing, visualization, etc. Experiments can be customized for many systems and models, and run from one convenient location.

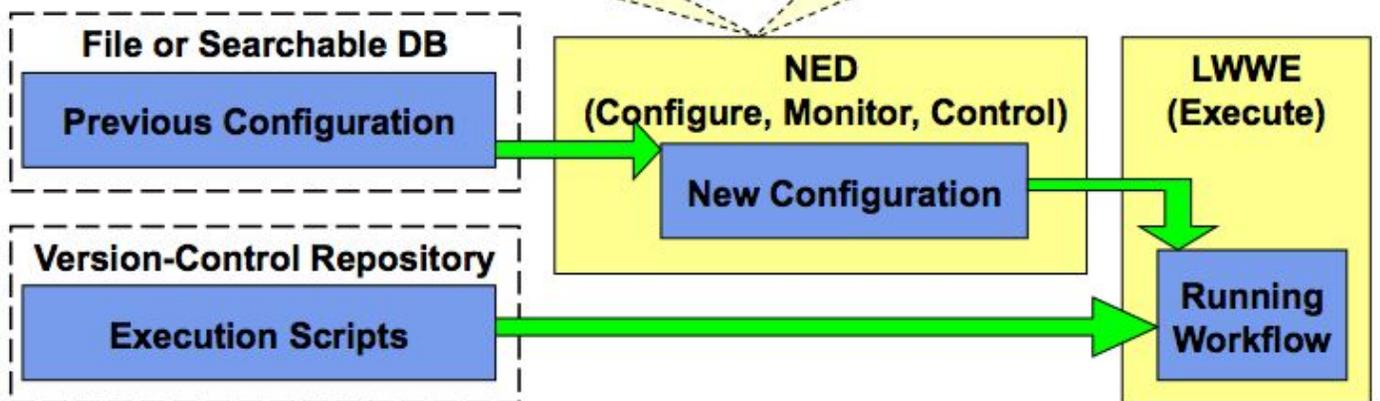
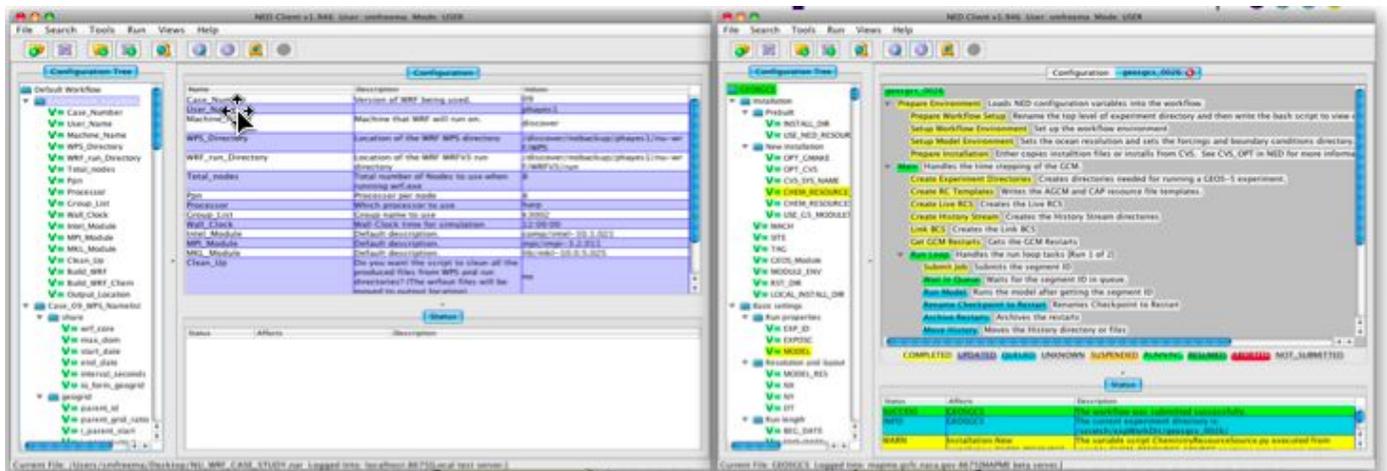


Figure 1: NED Client and LWWE Workflow Server

The tool is designed to be flexible, with beta versions running within NCCS on the mapme server as well as running outside NCCS on the cirrus/nimbus “desktop supercomputers”.

2. Getting Started

Before getting started using the tool, make sure your personal account is set up on whatever system the NED server software is executing.

Note: If you also plan to run workflows at the NCCS, you’ll need access to their “mapme” server, and quite probably accounts on HPC systems that may be used by workflows (e.g. discover, dirac, progress, and possibly cvsact). Please contact NCCS for an account on these systems.

NCCS Contact Information

- **Website:** <https://nccs.nasa.gov>
- **Phone:** 301-286-9120

- Email: support@nccs.nasa.gov

Preparing the Environment

There are two ways to run a workflow, controlled by the generic user option underneath the File menu.

1. **Generic User** - One way is to use the “generic” user, which will submit and run workflows on your behalf. This is generally the most convenient method for running a workflow as the workflow account is already set up with everything necessary to successfully execute a workflow.
2. **As Yourself** - The second method is to run as “yourself”. This method can require more effort as it is up to the user to ensure that their account is set up correctly and has shared keys. Such settings may include establishing correct ssh keys for issuing commands to remote HPC systems. Different environments may require different settings.

Note: The primary means for job submission and communicating with remote machines is ssh. As this is the case, it is important to avoid the use of “echo” commands in your login scripts (`.login`, `.profile`, and `.bashrc`) as these are known to cause ssh issues.

IMPORTANT: The `JAVA_HOME` environment variable must be set correctly and that the java binaries are added to your path for proper compilation and operation.

3. Using the NED Client

The NASA Experiment Designer (NED) workflow tool enables a user to quickly and easily set up and run an existing workflow. The NED GUI application provides a simple user interface for configuring the workflow, along with the ability to monitor the submitted workflow as it runs.

So what is a workflow? At a high level a workflow is nothing more than a series of dependant and/or independent steps that are executed in order to accomplish a goal. In the case of the NED workflow tool, a workflow is an abstracted set of scripts and programs that are run in a defined order to perform a set of tasks. The NED GUI allows the user to adjust parameters and other settings, which are then used by these scripts and programs to execute the tasks the way the user wants.

The NED application itself is a thick-client application, which communicates with a server either located locally or remotely. The server supports both several methods of authentication (including RADIUS) and uses SSL for secure communications.

Common Terminology

The NED application has some basic terminology one should be familiar with in order to work with the software.

Workflow Configuration

The workflow configuration is the fundamental construct of the NED application. It contains all the configuration information for a particular workflow.

The workflow configuration file can take on several forms. It can be either an XML file with a .ned extension, a “sectioned” properties text file with a .nif extension, or a compressed file that contains a .ned or .nif file. Compressed files are stored using the .nar extension.

Compressed files are just zipped files. Using an unzip program will extract the .ned or .nif file from the archive.

Groups

Groups are basically a container with a set of properties to help organize the workflow configuration. Each group can contain other groups and/or variables.

Variables

Variables are the basic unit of the workflow. A variable contains a name that is associated with some sort of value. Values can be anything from simple text to complex objects. Depending on how a workflow designer has set up the variable, user's may see different representations of the variable in the GUI.

Basic Operations

The operation of NED for the standard user is pretty straightforward.

Logging In

The first thing the NED application does when starting up is try to connect to a server. If the user has connected to a server before, then NED attempts to use the last used server. If not, then NED will try to use the default server. If that fails, NED warns the user that it is not connected. In this circumstance the user will need to go the login menu item under the tools menu and select or create a server configuration to login to.

Once NED connects to a server, the user will be prompted for a username and credentials. If the credentials are valid, the user will be able to use all the remote capabilities of NED.

Loading A Workflow

There are several ways a workflow configuration can be loaded. By default, NED will try to load the last opened workflow automatically. The user may also open a workflow from disk by using the file open option, or load a workflow from the remote server if they have appropriate permissions to do so.

The user may also load a workflow configuration from a database. There are two ways which this can be done. By selecting the open database option, a list will be presented of all the

experiments in the database. A user may select one from the list and load the workflow configuration.

The other option is to use the find in database option under the Search Menu. This presents the user with a list of properties to search for in the database. If no workflow configurations match

the criteria, the user is informed of this. If there are matches in the database, a list displaying those matches is shown and the user may choose to load a workflow configuration.

Choosing Who To Run As

A workflow can either be submitted as yourself or as a generic user. The generic user is the easiest to use as no additional setup is required. If the user chooses to run as themselves, then the user must have all the necessary SSH keys established. If the SSH keys are not set up correctly for a particular machine and the workflow utilizes that machine, then the workflow will fail.

Configuring a Workflow

After loading a workflow configuration into the NED client, the user can edit the configuration.

Variables in a workflow configuration come with a default setting. You can change them by selecting the variable and then clicking on the value field in the table. An edit box, combo box or some other widget will allow the user to change the value to something else.

NED also allows you to reference other NED variables when specifying a value. This is done by using `#...#` as a beginning and end demarcation for the variable reference. So if a user had a Variable BAR in Group FOO, then the user would reference that variable by specifying `#FOO.BAR#` in the variable value.

A user can search for specific variables by entering in keywords in the Find In Workflow window under the Search menu. The user can then select a variable from the list and the GUI will update to show that variable to the user.

The user selects and alters variables to match what they want. If real-time validation is selected from the Preferences, then NED will validate the workflow after every change and notify the user of potential issues. If issues are present, the GUI will highlight what workflow items have the problems.

As noted previously, how a variable is edited depends on how the designer set up the variable. Some variables contain multiple values. Some variables represent tables. Some variables may even have custom editors for handling complex objects, such as the GEOS History Editor. Different variable editors may be displayed to the user for simpler editing depending on how the configuration is set up.

4. Running a Workflow

When the user is done configuring the workflow, the user may then submit the workflow to the NED server for execution. If the user submitted with validation then if the server detects a validation problem it will notify the user and ask them if they want to continue. If the user submits without validation, then the server will attempt to submit the workflow regardless of validation issues.

However, if there are errors present in the workflow then that will always stop a submission from taking place, regardless of whether or not the user has submitted with validation. These errors can range from connection problems to integrity failures (workflow configuration structural issues). These status messages are reported back to NED and displayed in the Status pane.

Upon a successful submission, the Status pane will contain a message indicating where the workflow is running from.

Monitoring A Workflow

After a workflow has been successfully submitted, NED will display a task tree view of the submitted workflow. This view updates at a regular interval, displaying a color coded status of the executing workflow. This view also allows the user to control the workflow execution to some extent. The control options are available by right clicking on the monitoring window.

The task tree views can be closed at any time. If the user wishes to reopen a workflow view they may do so by going to the Views menu and selecting the "View Other Workflows" and selecting the appropriate engine and workflow ID.

Users can also monitor logs that are generated by a running workflow by right-clicking tasks or the general task tree. After selecting a log to view, the Status pane will display the log. The log display is updated at regular intervals just like the task tree view. This additional information can be useful in the event that a problem is encountered.

For legacy SMS based workflows, users may also log into the server and examine their workflows using the XCDP application. This is an X-Windows application that interacts directly with SMS which is a scheduler engine provided by ECMWF. More information about SMS and XCDP is available here: <http://www.prism.enes.org/Software/WSS/sms/manual/course/index.html>.

NOTE: Current workflows should NOT use SMS. SMS is not very portable and is currently only set up on the MAPME server. SMS support will be deprecated in the future.

Saving A Workflow

At any time the user may save the current workflow to file or to the database using the menu options provided. The database has the additional benefit of being searchable which may make retrieving specific workflow configurations easier.

For file based workflow configurations, NED keeps a history of the last five workflows that it has opened for quick retrieval.

5. The NED Interface

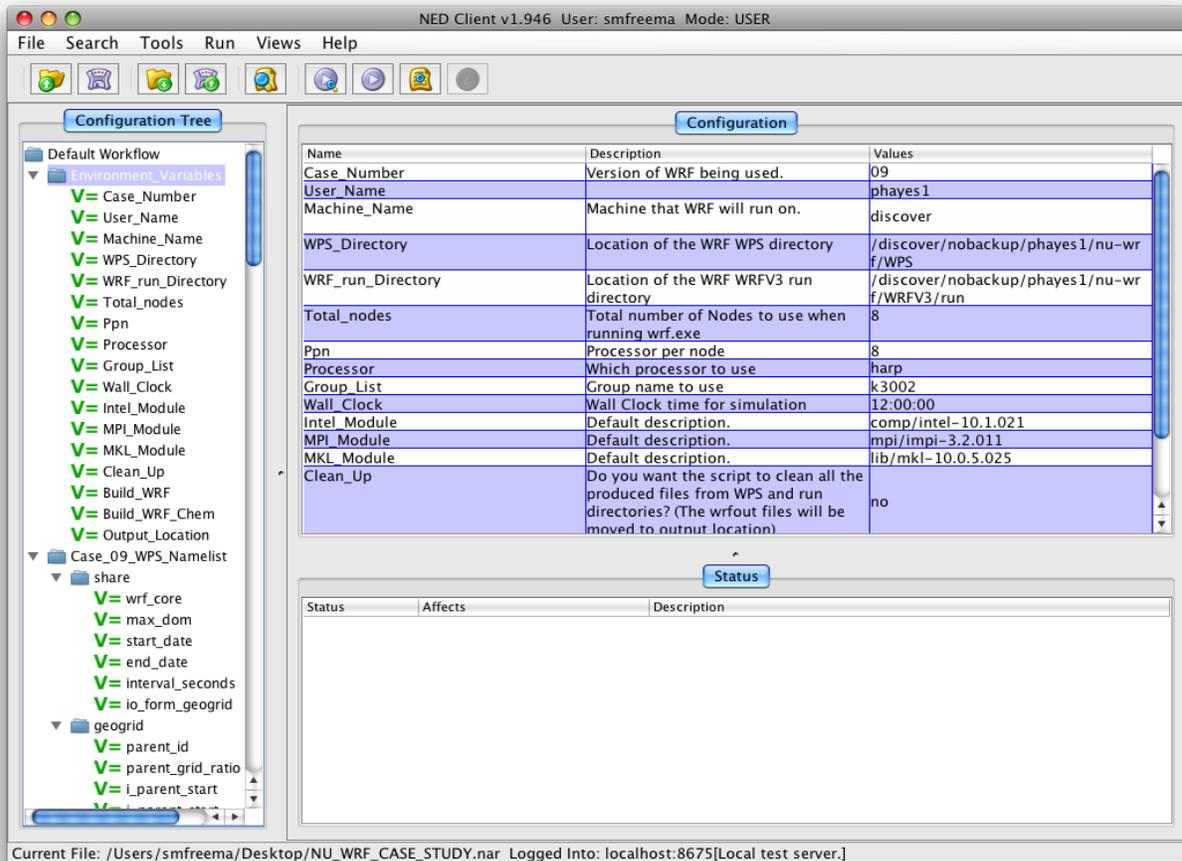


Figure 2: A typical example of what the NED GUI looks like when a user starts the application. The main GUI window is divided into three components.

Main Window GUI Components

Configuration Tree

The configuration tree is the main navigational element of the GUI. The tree represents the hierarchical structure of the current workflow configuration. By selecting an element in the workflow tree, the GUI changes to display the properties of the selected element. The tree also allows users to enable or disable groups if the workflow configuration allows.

Configuration

The configuration pane displays the selected object's properties. Selecting the workflow element at the top of the tree will display the overall workflow properties. Selecting a group will display a list of the variables and their properties that are contained underneath the group. Selecting a variable will highlight the selected variable.

Status

The last pane is the status pane. This will display information in regards to the workflow, such as submission issues, validation warnings, logs, and/or other related information.

Making Tree Selections

By selecting different elements in the tree, the configuration pane changes.

Selecting the root of the tree brings up the main properties for the workflow. Some of these properties are editable, while others are automatically filled in for the user.

- **Name:** The user friendly name of the workflow.
- **Unique ID Prefix:** This is the prefix used for naming the submission. The application uses this to create submission directories and to keep track of the number of submissions for a particular experiment to prevent inadvertent overwriting.
- **Description:** A general description of the workflow.
- **Modified By:** The last user who edited the workflow. This field is automatically updated by the application and is not editable.
- **Modification Date:** When the workflow was last modified.
- **Group Code:** Based on your credentials, this field will allow you to select a group code that will be used to submit workflows. Depending on the workflow, this field may be ignored.
- **Notes:** A place to put any other information that is relevant to the workflow or submission.

Variable groups are represented in the tree by folder icons. By selecting a group within the tree, contents are displayed. If a group only contains other groups, then those groups are listed. If a group contains groups and variables, then only the variables are displayed.

Groups may also have the ability to be enabled or disabled. This is indicated by a check box next to the group name in the tree. A disabled group will not generate configuration information for its contents, effectively making it invisible to a workflow. An enabled group will be included in the workflow configuration.

Variables are represented in the tree by a green "V=" icon. By selecting a variable in the tree, the variable is highlighted in the Configuration pane. The variable's value can then be edited by the user. A general user can only alter the variable value.

Other Tree Functions

The tree has several right-click options. Most of these are self explanatory. Expand and collapse options will fully expand and collapse the tree. Enable and disable all groups will recursively enable or disable all groups

Menu Bar

The NED application has a menu bar that contains a fair amount of functionality.

File Menu

The file menu contains the options for loading and saving workflows, along with other functionality.

- **Open:** Brings up the open workflow window that enables a user to select a file to open as a workflow. This can also be brought up by pressing CTRL O.
- **Save:** Saves the workflow to disk if the workflow was opened from a file on disk. Otherwise, this will bring up the Save As window for the user to select or enter the name of a file to save to. This can also be done by pressing CTRL-S.
- **Save As...:** Brings up the Save As window the allows the user to specify the file to save the workflow to.
- **Open From Database:** Presents the user with a dialog for selecting a workflow to load from the database.
- **Save To Database:** Saves the current workflow to the database. The user may select from a list of databases.
- **Delete From Database:** Allows the user to delete a workflow from the database. The user is only allowed to delete workflows that they either created or modified.
- **Open Remote:** Allows a user to open a remote file stored on the server. The user must have read permissions and be connected to a server in order to open the file.
- **Save Remote:** Allows a user to save their current configuration to a file on the remote server. The user must have a write permissions and be connected to a server to write files on the server.
- **Design Mode:** Allows the user to switch between USER and DESIGN mode. DESIGN mode enables advanced functionality that should only be used by people creating new workflows. Most users will not need to use this option.
- **Runtime Account:** This sets how the NED application will run a submitted workflow. The generic account is not affiliated with the user's systems accounts. When running as "Yourself", you must have all the appropriate SSH keys set up for communication as required by the workflow being run. NED has an SSH tool that can handle key setup.
- **Previous Workflows:** Displays a list of recently opened workflows. The user can select one, and if that workflow is still present then the NED application will load it.
- **Exit:** Exits the NED application. If you have made changes to a workflow, then the user is prompted to save the changes before exiting. This can also be accessed by using CTRL-W.

Search Menu

The Search Menu contains options for searching workflows or databases.

- **Find In Workflow:** Brings up a search window that allows the user to search within the workflow of a specified string. The user may then click on results and NED will take them to where in the workflow the result was found. This may also be accessed by pressing CTRL-F.
- **Find In Database:** Brings up a window that allows the user to search the database for a particular workflow. The user can fill in the fields to search for. If matches are found, they are returned for the user to select from.

Tools Menu

This menu can vary depending on if the user has added their own tools. Below are the standard menu items for this menu.

- **Login To Server:** Brings up a window that allows the user to switch to a different server.
- **NED Preferences:** Displays a window that allows the user to alter some of the look and feel parameters of the GUI.
- **Manage External Tools:** Allows the user to add, remove, or configure their own tools that they wish to run from within the NED application.
- **Manage SSH Keys:** This option allows the user to add, remove, or refresh their SSH keys that are necessary for their systems.
- **Difference Workflows:** Allows the user to compare the current workflow to another one from file or from the database. The differences between the workflows are highlighted, and some differences can be copied to the current workflow.

Run Menu

This menu contains options for submitting and validating a workflow.

- **Submit Workflow:** Submits a workflow with validation. If validation issues are detected, those issues are displayed in the status window and the user will be asked whether or not they wish to fix the problems before submission. It is usually a good idea to fix validation errors.
- **Submit Workflow Without Validation:** Submits a workflow regardless of validation issues.
- **Validate Workflow:** Runs through the validation logic in within the workflow. Any issues are reported in the status window.

Views Menu

This menu contains options for controlling the different views NED can display. The standard configuration view is the default and cannot be closed. However NED can have one or more workflow task tree views open which displays the state of the submitted workflows.

The menu options available depend on the number of workflow engines available. Here are the current options

- **Clear Status Window:** Clears the status window of messages.
- **SMS and LWWE:** Both of these have the same options which operate on their respective workflows
 - **View Most Recent:** This option will pull up the most recent workflow submitted based on the workflow loaded.
 - **View Other Workflows:** Displays a window allowing the user to select from a list of workflows. This list will be dependent on the server connection shown at the very bottom of the NED window. Upon selection, the task tree view of the workflow is displayed.
 - **Close All Workflow Views:** Closes all open workflow task tree views.

Help

This menu option only has one item: Help. This will bring up a browser directed to the MAPME Workflow Tool page on the NASA Modeling Guru website.

Tool Bar

The NED application has a tool bar that allows single-click operation for select features. The tool bar can be dragged around or off the application. The tool bar itself is comprised of tool bars which allow the user to move and rearrange the tool layout. This section will only cover tool bar items available in user mode.

Open Workflow

Opens a selected workflow. This behaves the same way as the file open menu option.

Save Workflow

Saves the current workflow. This behaves the same way as the save file menu option.

Open Remote

Opens a remote workflow stored on the server if the user has the correct permissions.

Save Remote

Saves the current workflow to the server if the user has the correct permissions.

Find In Workflow

Finds workflow elements with the user specified text. This behaves the same way as the find in workflow menu option.

Submit With Validation

Submits a workflow with validation. This behaves the same ways as the submit with validation menu option .

Submit Without Validation

Submits a workflow without validation. This behaves the same way as the submit without validation menu option.

Validate Workflow

Validates the workflow. This behaves the same way as the validate workflow menu option.

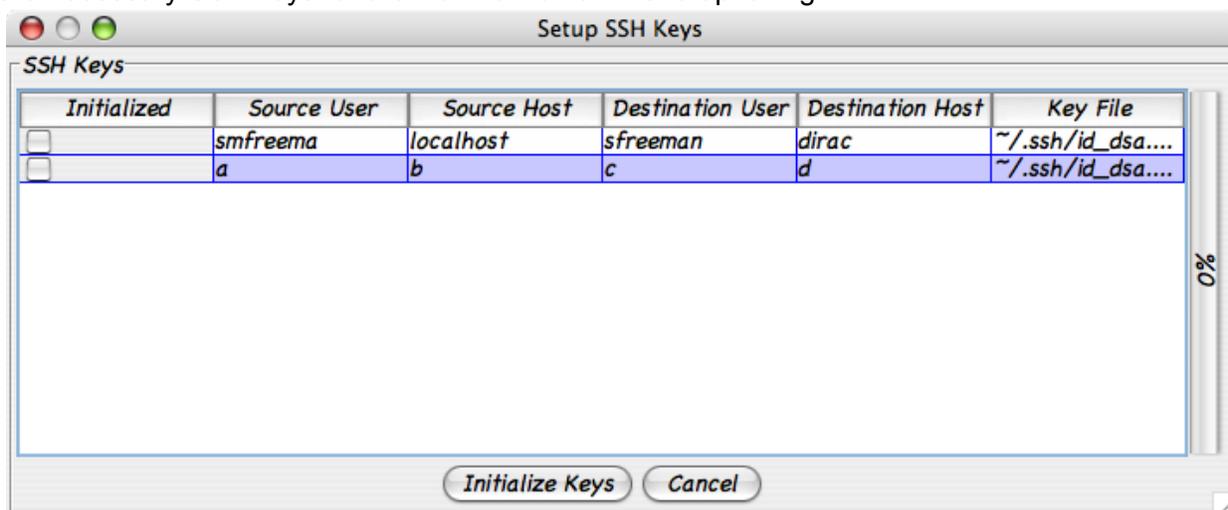
Info

If the workflow configuration has defined additional information, this will display it to the user.

6. Advanced NED Features

SSH Key Management

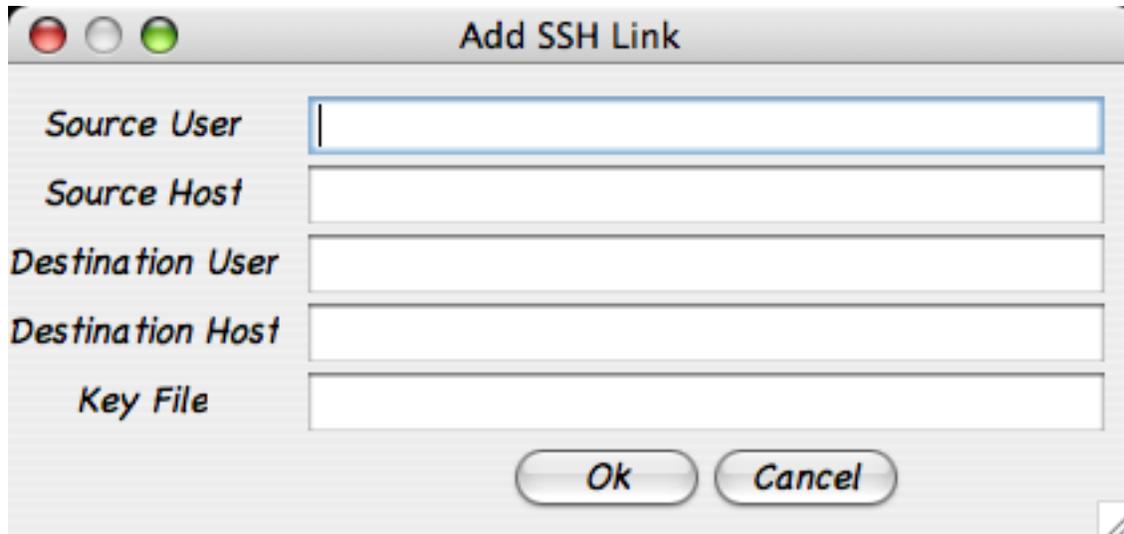
When a user submits a workflow, the user can choose to either submit it as a generic user or as him/herself. If the user chooses to submit as themselves, then they will need to have set up all the necessary SSH keys for the workflow or it will end up failing.



The above image shows the SSH window displayed when setting up or refreshing keys. As it goes through the keys, it updates with the status of where it is in the process. It will also notify the user of problems if encountered.

The SSH Key Manager allows the user to add, remove, or refresh their SSH keys. When a user

starts NED, they will be asked to set up their SSH keys. The SSH Key window will prompt them for any credentials needed for setting up the SSH keys. The SSH tool will either use existing files if it finds them or create new keyfiles if they don't exist. Upon a successful completion, the user should have all required keys on all the necessary systems.



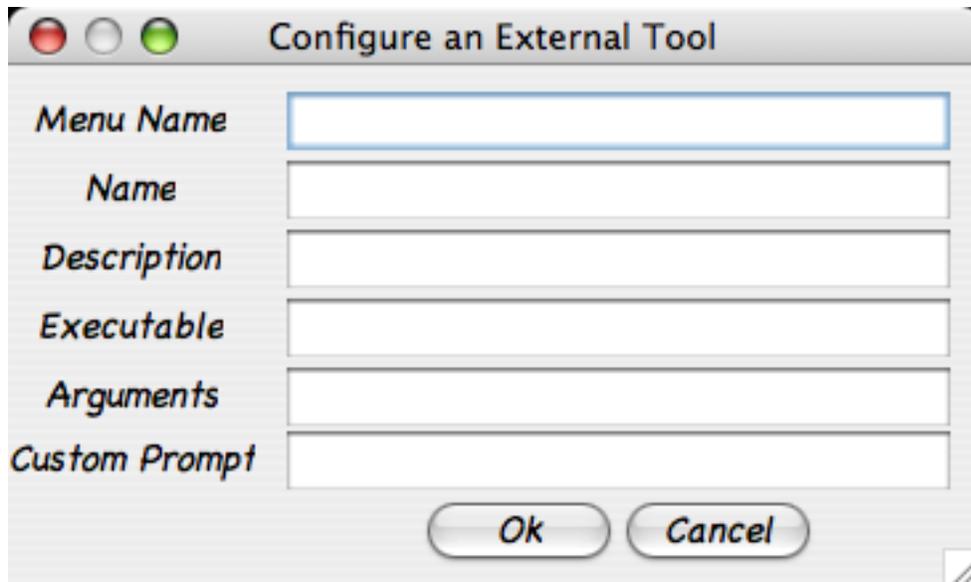
The above image shows the window used to add a key. A user specifies the username and host for the source and destination machines. The user may also specify a particular keyfile to use. Leaving the field blank will use the default keyfile.

Sometimes changes occur on the systems a workflow runs on and the keys will need to be refreshed. This can always be performed by going to the tools menu and selecting the refresh keys option under the manage SSH keys menu.

At present, this operation can only be performed by running the NED client on the hosting machine (currently mapme).

External Tool Management

NED enables the user to add any number of external tools so that they may be invoked from the GUI. The add external tool option under the manage external tool menu allows the user to specify how a tool should be invoked and how it appears in the NED GUI. The user may also choose to edit and existing external tool or remove external tools.



This image shows the external tool configuration window. A user enters the menu they want the tool to appear under. The user can specify a comma separated list of names to create a menu chain. The name field is what the tool will be called. The description is optional, but if filled out a tooltip will display when hovered over. The executable is the tool to run. Arguments are arguments to the tool, which can include special keywords for accessing NED values (using the #...# format). The last field is a custom prompt. This allows a user to specify arguments to prompt the user for.

There are two other special indicators. If %prompt% is used in the arguments field, the user will be prompted for arguments when the tool is invoked. If %browser% is used in the executable field, then the tool will launch a the systems default browser when executed. In this case, only the first argument is used to specify the URL the browser should launch.

Differencing Workflows

The Difference tool allows users to compare two similar configurations to see exactly what parameters are different between the two. Select Tools/Difference Workflows and a new window will display allowing you to compare another workflow. This comparison can either be done against a workflow from file or a workflow from the database.

The differencing window highlights the differences between two workflows, and also allows copying differences to the current workflow.