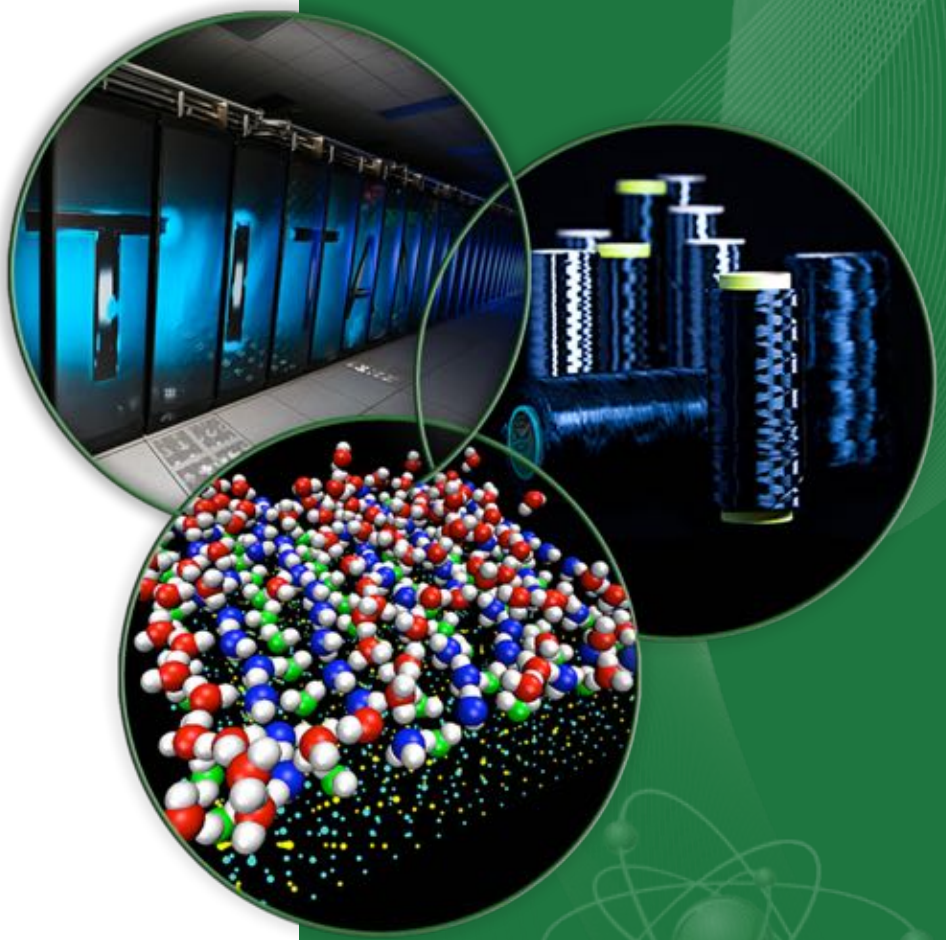
Scientific Software Development with Eclipse



A Best Practices for HPC Developers Webinar

Gregory R. Watson

ORNL is managed by UT-Battelle
for the US Department of Energy

OAK RIDGE National Laboratory

Contents

- Downloading and Installing Eclipse
- C/C++ Development Features
- Fortran Development Features
- Real-life Development Scenarios
  - Local development
  - Using Git for remote development
  - Using synchronized projects for remote development
- Features for Utilizing HPC Facilities

What is Eclipse?

- An integrated development environment (IDE)
- A platform for developing tools and applications
- An ecosystem for collaborative software development

Getting Started

Downloading and Installing Eclipse

- Eclipse comes in a variety of packages
  - Any package can be used as a starting point
  - May require additional components installed
- Packages that are best for scientific computing:
  - Eclipse for Parallel Application Developers
  - Eclipse IDE for C/C++ Developers
- Main download site
  - https://www.eclipse.org/downloads

Eclipse IDE for C/C++ Developers

- C/C++ development tools
- Git Integration
- Linux tools
  - Libhover
  - Gcov
  - RPM
  - Valgrind
- Tracecompass

Eclipse for Parallel Application Developers

- Eclipse IDE for C/C++ Developers, *plus:*
  - Synchronized projects
  - Fortran development tools
  - Job scheduler support
  - Remote monitoring
  - Remote console

Installation

- First, install Java 1.8
  - Check if it is installed using `java -version` from command line
  - Follow procedure for your operating system
- Download Eclipse package
  - Zip for windows
  - Tar.gz for Linux
  - Dmg for Mac OS X
- Uncompress and move to installed location
- Launch Eclipse application

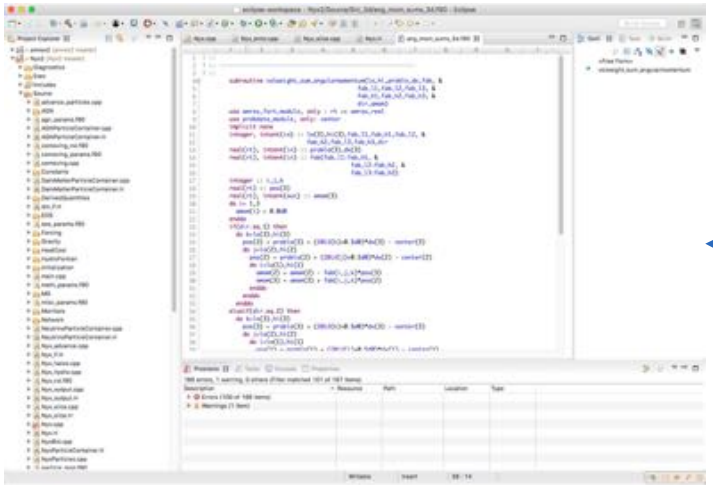
Adding Features

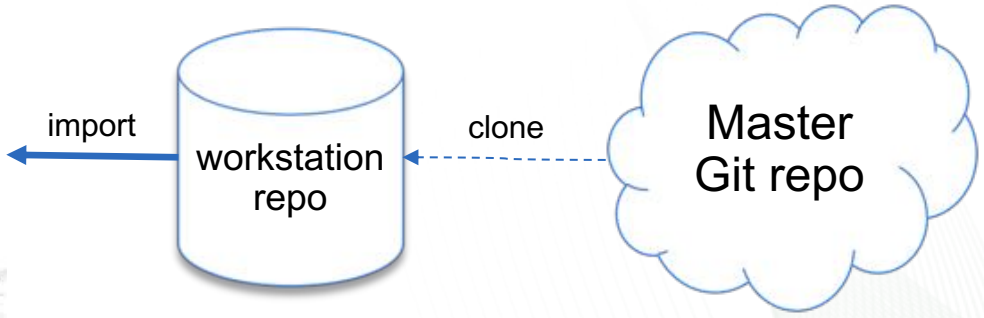
- Eclipse Marketplace
  - Over 1600 packages available
  - Ability to search and browse
  - **Help > Eclipse Marketplace…**
- Eclipse update sites
  - Good for updating installed software to latest version
  - Or if you know the URL
  - **Help > Install New Software…**

Developing with Eclipse

C/C++ Development

- Works best on local projects with hierarchical directory structure
- Supports Makefile/CMake based projects
- Can import directly from a Git repository
- Can manage multiple Git repositories



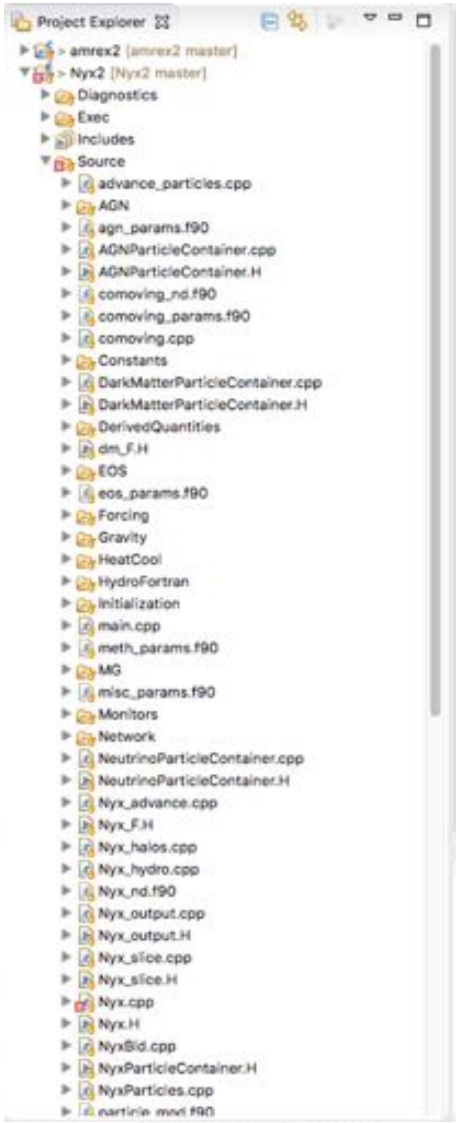


Importing from Git

- Select **File > Import…**
- Select the **Git > Projects from Git** import wizard
- Clone URI
  - https://github.com/AMReX-Astro/Nyx.git
- Once cloned, choose **Import as general project** Wizard then **Finish**
- Then select the project, right click, and choose **New > Convert to a C/C++ Project (Adds C/C++ Nature)**
- Pick **Makefile project** from *Project type*

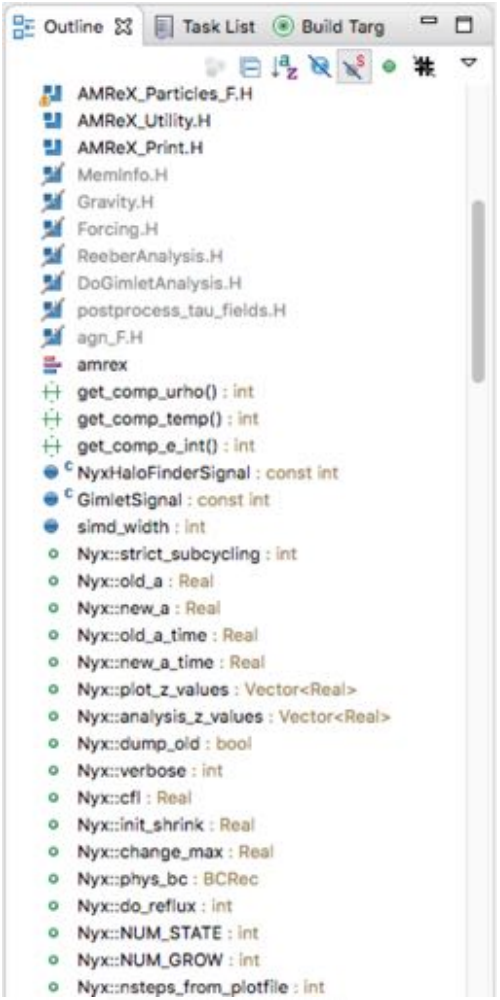
Project Explorer

- Shows project tree structure
- Virtual nodes showing
  - Include paths
  - Libraries
  - Binaries and executables
- File nodes can be expanded to show
  - Preprocessor symbols and includes
  - Type and variable declarations
- Compound types can be expanded to show
  - Fields
  - Methods



Outline View

- Shows structure of current file in editor
  - Preprocessor symbols and includes
  - Type and variable declarations
- Compound types can be expanded to show
  - Fields
  - Methods
- Can filter what is being shown using buttons or dropdown menu



Editor Features

- Syntax coloring
- Line numbers
- Folding
- Content assist
- Hover help
- Block selection

- Code activation based on preprocessor directives
- Formatting
  - Can be run from the command line
- Display revision information

Formatting and Refactoring

- Formatting
  - Generate Getters and Setters
  - Add/Organize Includes
  - Implement Method
  - Toggle Comment
- Refactoring
  - Rename
  - Extract Constant
  - Extract Local Variable
  - Extract Function
  - Toggle Function Definition
  - Hide Method

And many other features…

Fortran Development[1]

- Fortran editor
  - Similar to C/C++ editor
- Fortran perspective
  - Gathers together various Fortran specific views
  - Adds Fortran declaration view
- Fortran feature search
  - Search for language features

[1] Requires Parallel Application Developers Package

Fortran Editor

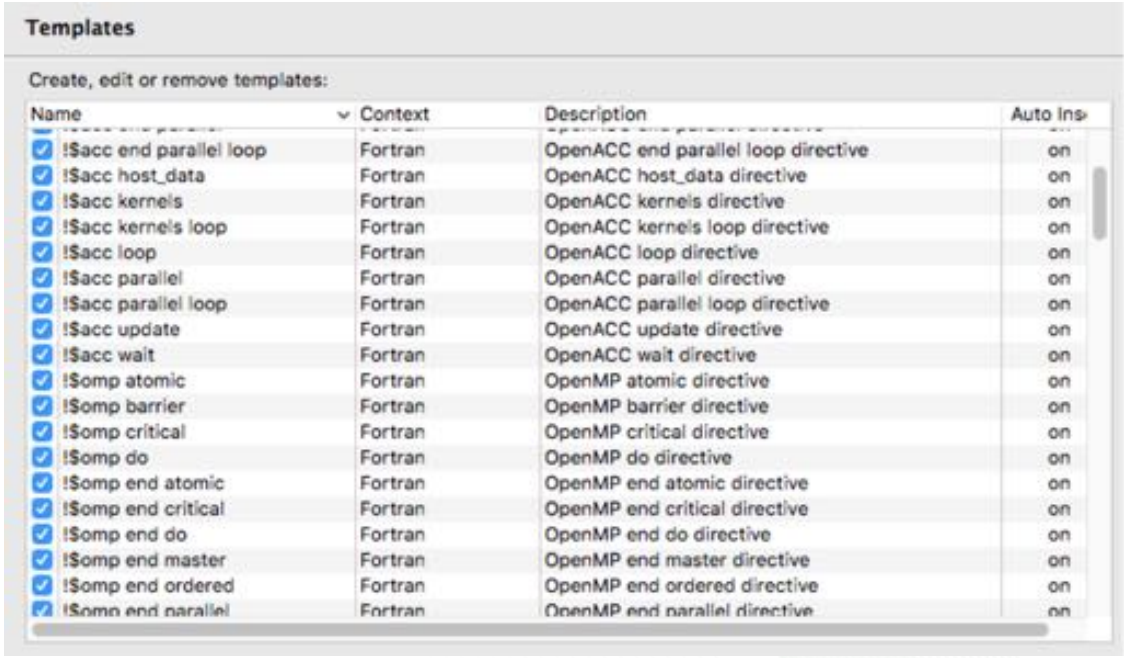
- Supports free and fixed formats
- Opens for any file ending in Fortran suffix
  - .f, .F, etc.: fixed source form
  - .f08, .f90, etc.: free source form with INCLUDE
  - .F08, .F90, etc.: free source form with C preprocessor
- Syntax coloring
- By default, only basic editing features are enabled

Advanced Fortran Development

- Fortran analysis/refactoring is disabled by default
- If not already a Fortran project
  - Right click on project > **Convert to Fortran Project**
- Open project properties
- Select **Fortran General > Analysis/Refactoring**
- Check **Enable Fortran analysis/refactoring**
- Choose analysis properties

Advanced Editor Features

- Folding
- Content assist
- Hover help
- Code templates



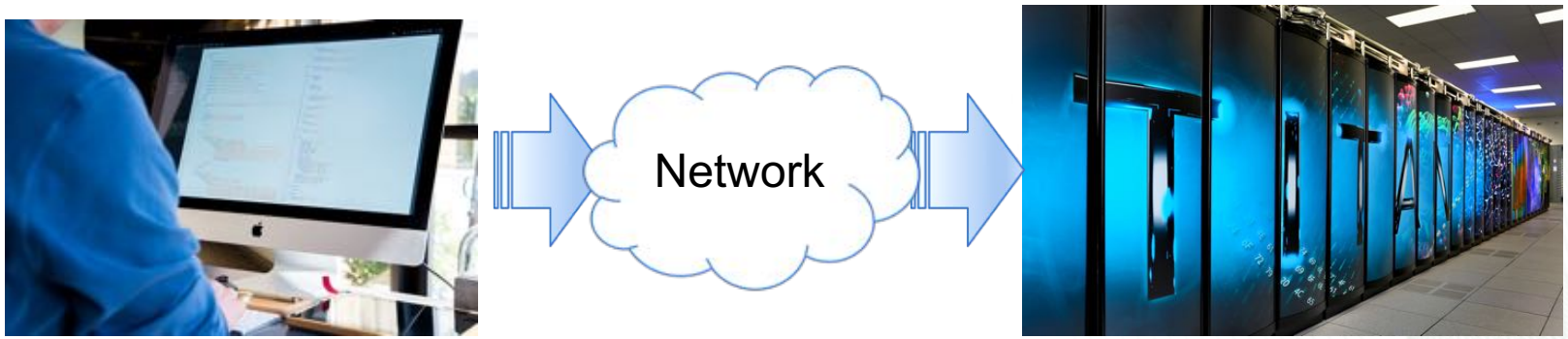
Real-life Development Scenarios

Example Scenarios

- Local development – already covered
- Using Git for remote development
- Using synchronized projects for remote development

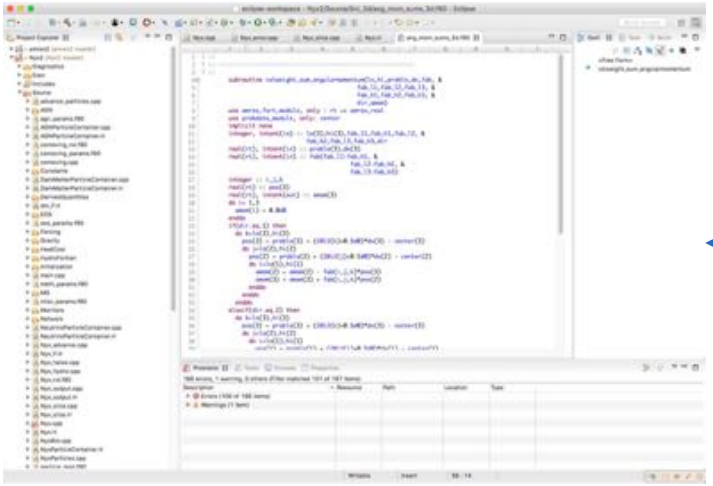
Remote Development

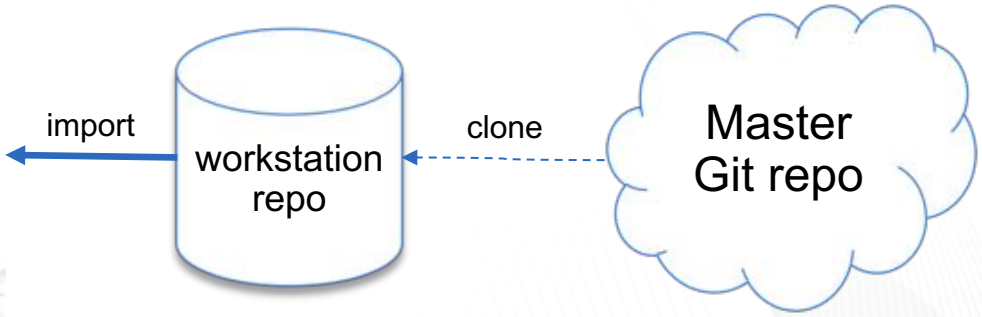
- In scientific computing, application code is normally compiled and run on remote system
- Local machine rarely has same environment, libraries, etc. as target system
- May have different architecture, utilize GPUs, etc.
- Also usually need to submit job via batch scheduler



Remote Development Using Git

- Clone repository to workstation either through UI or command line
- Import into Eclipse as before
- Clone repository on target system if it is not already there

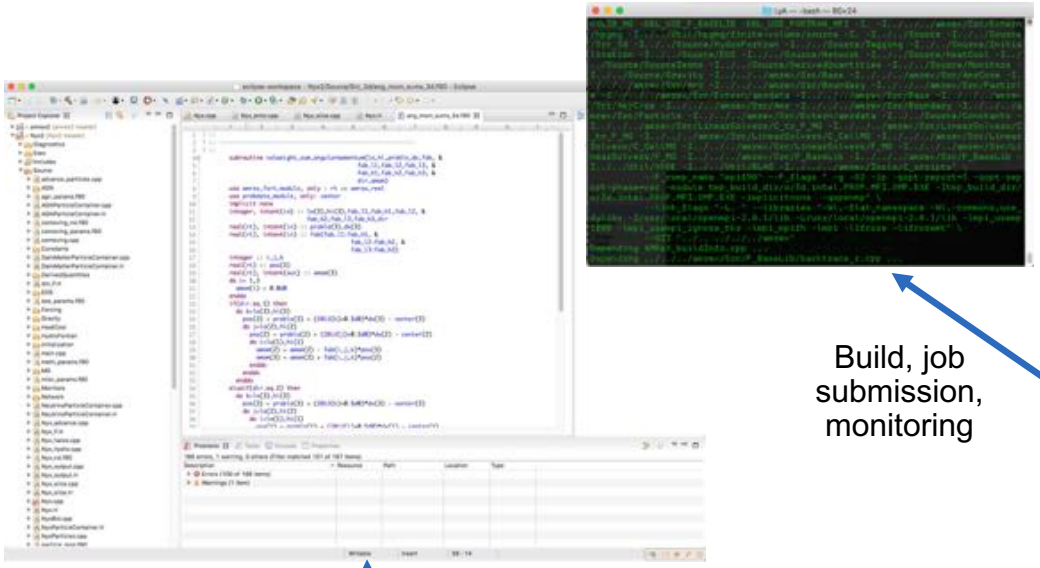




Remote Development Using Git Cont…

- Changes committed to workstation repository
  - Push to central repo (e.g. GitHub) or directly to target system (if allowed)
  - Can utilize code reviews (e.g. Gerrit) and continuous integration if required
- Pull changes into repository on target machine
- Manually run build
- Manually submit to job scheduler

Remote Development Using Git Cont...



Build, job submission, monitoring

Editing, static analysis, search, navigation

workstation repo

Target system repo

Master Git repo

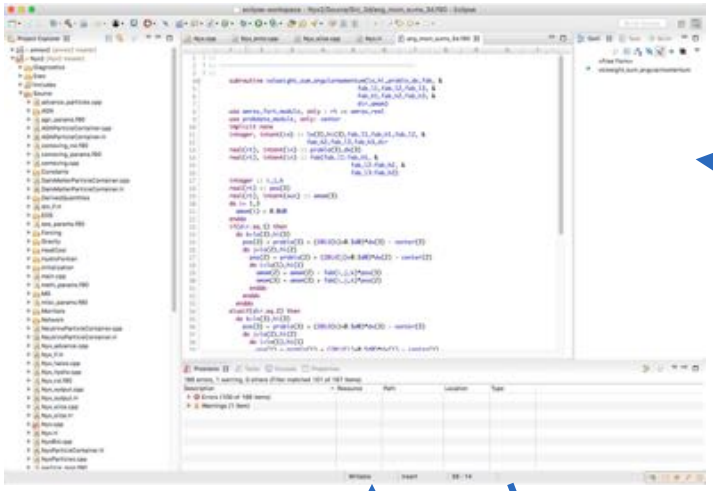
push

pull

Remote Development using Synchronized Projects

- Rather than using Git, Eclipse can manage the synchronization for you
  - Any changes made locally will be automatically synchronized
  - Changes made remotely can be manually synchronized or will be picked up at next sync point
  - Can configure filters to avoid copying large files
- Orthogonal to Git, so both can be used
- Can start with either local or remote source

Synchronized Projects



Build, job submission, monitoring

Editing, static analysis, search, navigation

Workstation copy

Target system copy

Git repo

sync

Starting with Local Source

- Create project as before (e.g. from Git)
- **New > Other**
- **Other > Convert to Synchronized Project**
- Choose project
- Choose connection and remote directory
- After synchronize
  - Go to project properties
  - **C/C++ Build > Tool Chain Editor**
  - Set the current toolchain for the target system (change current build back to “Sync Builder” if necessary)

Remote Building

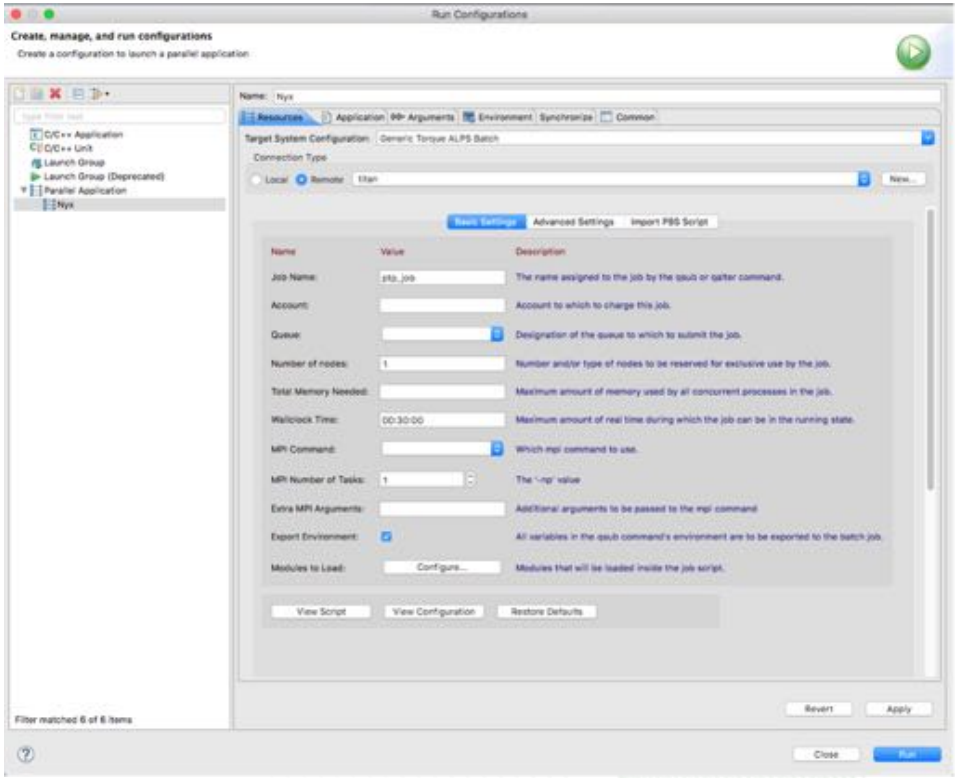
- Synchronized projects automatically set up remote build
- Clicking on the build button will run the build command remotely (normally “make”)
- Add build targets to run “make whatever”
- Can run more complex build commands also

Other Parallel Application Developer Features

- Job submission
- Monitor system/queues
- Remote console

Job Submission

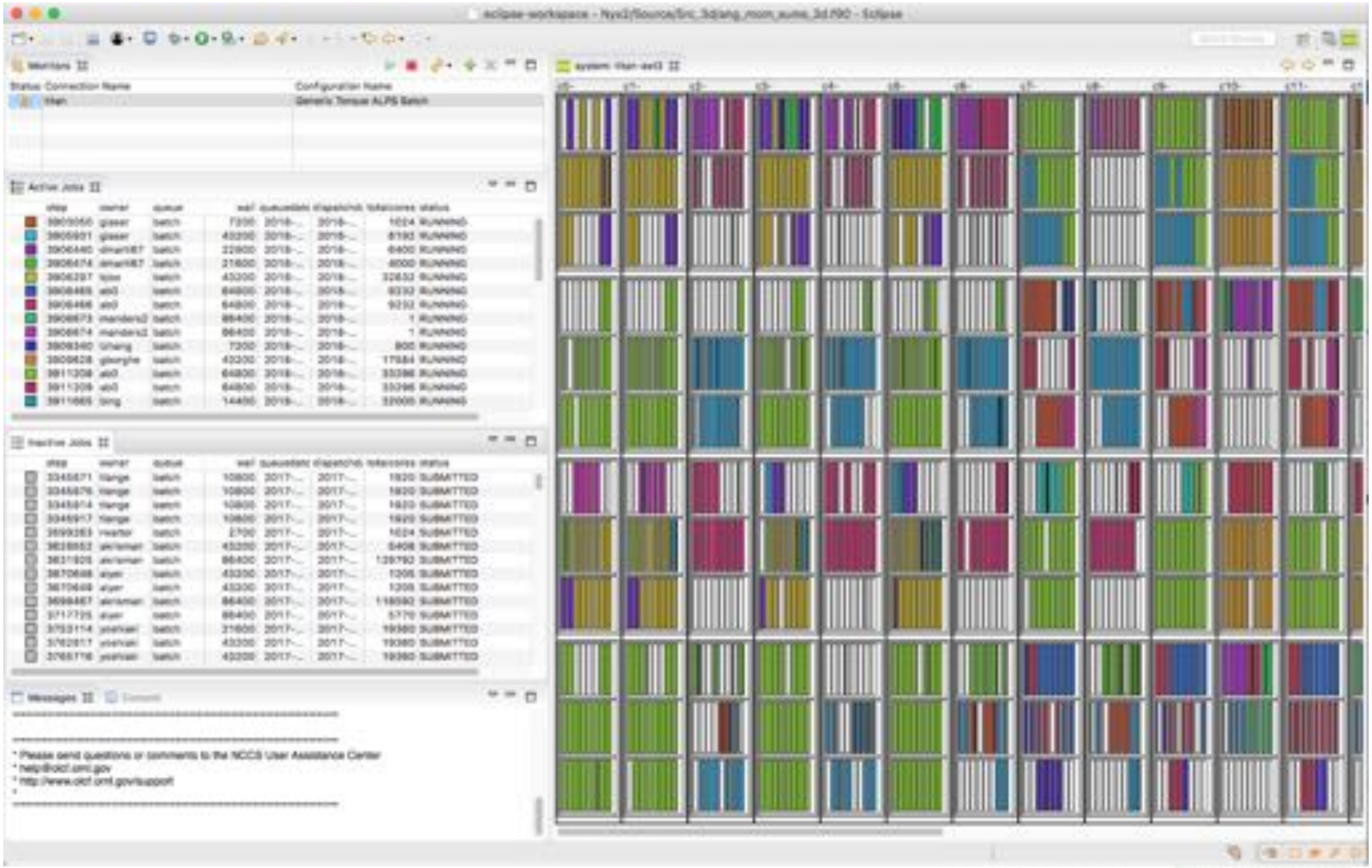
- Use the “Parallel Application” run configuration type
- Comes pre-configured with many generic- and system-specific configurations
- Supports most common job schedulers and runtimes
- Can use to launch remote commands also



System and Job Monitoring

- Comes pre-configured with many generic- and system-specific configurations
- Switch to “System Monitoring” perspective
- Can monitor multiple systems simultaneously
- Jobs launched through PTP can be controlled
- Once job is completed, stdout/stderr is accessible from the UI

System Monitoring Perspective



Remote Console

- Select Console view using tab
- Click on open console button and choose “Command Shell Console”
- Select the Connection Type and Connection name you want to use
- Click OK
- You will now have a shell on the target machine
- Open as many consoles as you like

Environment Modules

- Many HPC systems use environment modules
  - Allow different compilers/libraries to be selected
- Environment modules are integrated with the Parallel Application Developer package
  - Modules can be selected before the project is built
  - Modules can be selected before the code is submitted to the job scheduler

Summary

- Eclipse provides a variety of features to support scientific software development
  - C/C++/Fortran development
  - Local/remote project management
  - Integration with Git
  - Support for job submission and monitoring
  - Environment module support
- Allows developers who prefer IDEs to pick and choose how they wish to develop
- Supports complex workflows and provides both automatic and manual configuration options

Additional Material

C/C++

Importing Dependencies (Optional)

- Nyx depends on AMReX
- Repeat the same process for the AMReX repo
  - https://github.com/AMReX-Codes/amrex.git
- Only needed if you
  - Want to build locally
  - Want to resolve include files and types

Project Configuration

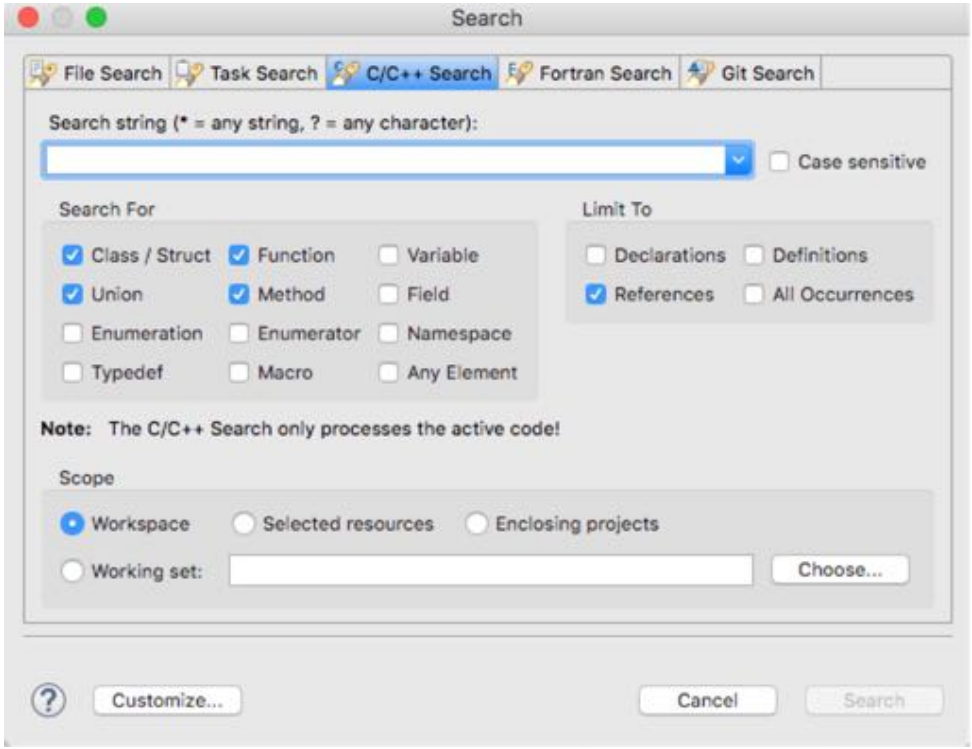
- Some settings are worked out automatically
  - Include paths
  - Compiler defined macros
- Usually need to add includes from dependent libraries manually
  - Open project properties
  - Go to **C/C++ General > Preprocessor Include Paths**
  - Add appropriate entries

Managing Code Analysis

- Code analysis (codan) requires headers to be configured correctly
- If the automatic configuration misses some header files you can add these manually
- You can also disable codan
  - Open project properties
  - Go to **C/C++ General > Code Analysis**
  - Select “Use project settings”
  - Uncheck problems you don’t wish to see

Search

- Search for
  - Class/struct/union
  - Function/method
  - Variable/field
  - Namespace
  - Typedef
  - Macro
- Limit to
  - Declarations
  - References
  - Definitions

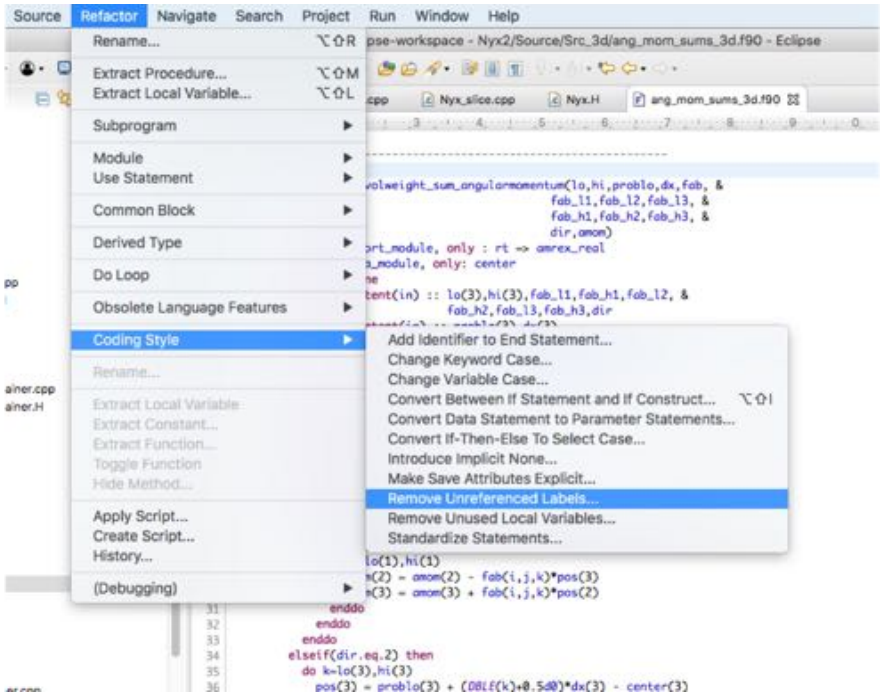


Other Features

- C/C++ Unit Testing
- Visual debugging
- Multicore debugging
- LLVM support
- And more…

Fortran

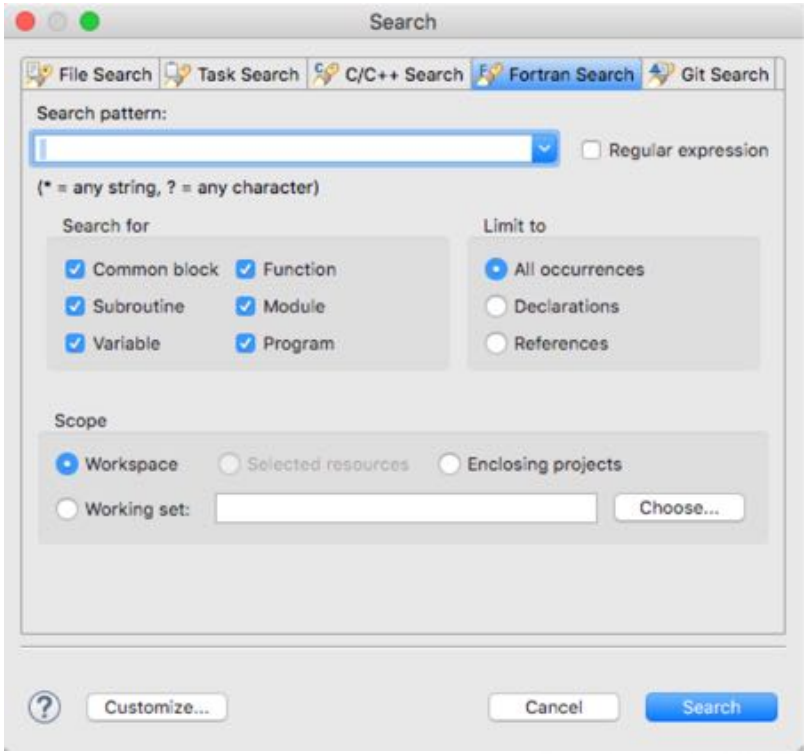
Refactoring



- Rename
- Extract procedure
- Extract local variable
- Make private entity public
- Add subprogram parameter
- Make common block names consistent
- Unroll loop
- Introduce implicit none
- And more…

Search

- Search for
  - Common block
  - Subroutine
  - Variable
  - Function
  - Module
  - Program
- Limit to
  - Declarations
  - References



Synchronized projects

Starting with Remote Source

- **New > Synchronized C/C++ Project**
- Pick project name (can be different from remote)
- Pick remote connection or create a new one
- Browse for remote directory
- Pick project type (normally Makefile > Empty Project)
- Select toolchains for local and remote copies
- Remote source will be automatically copied to a local project

Configuring Synchronized Projects

- Advanced editing features can be used because there is a local copy of the source
- It would be useful if the editor reflected the remote environment
  - System/library include files
  - Architecture specific macro definitions
- This information can be gathered from
  - Automatically from compilers on the remote system
  - Manually from compilers on the remote system (macros file)
  - Entered manually

Automatic Configuration (GCC only)

- From project properties
  - **C/C++ General > Preprocessor Include Paths, Macros, etc.**
  - Click on “Providers”
  - Select
    - Sync GCC Build Output Parser
    - Sync GCC Builtin Compiler Settings
  - Check “Allocated console in Console View” if you want to see the commands that are run
- Should trigger a re-index of the project

Manual Configuration (compiler generated)

- Generate macro definitions by running the appropriate compiler command
  - E.g. gcc -E -P -v –dD file.c > macros
  - Synchronize the project so that “macros” is copied to local
- From project properties
  - **C/C++ General > Preprocessor Include Paths, Macros, etc.**
  - Click on “Entries” and select “CDT User Setting Entries”
  - Click “Add”
  - Choose “Preprocessor Macros File”
  - Navigate to and select the file from the project

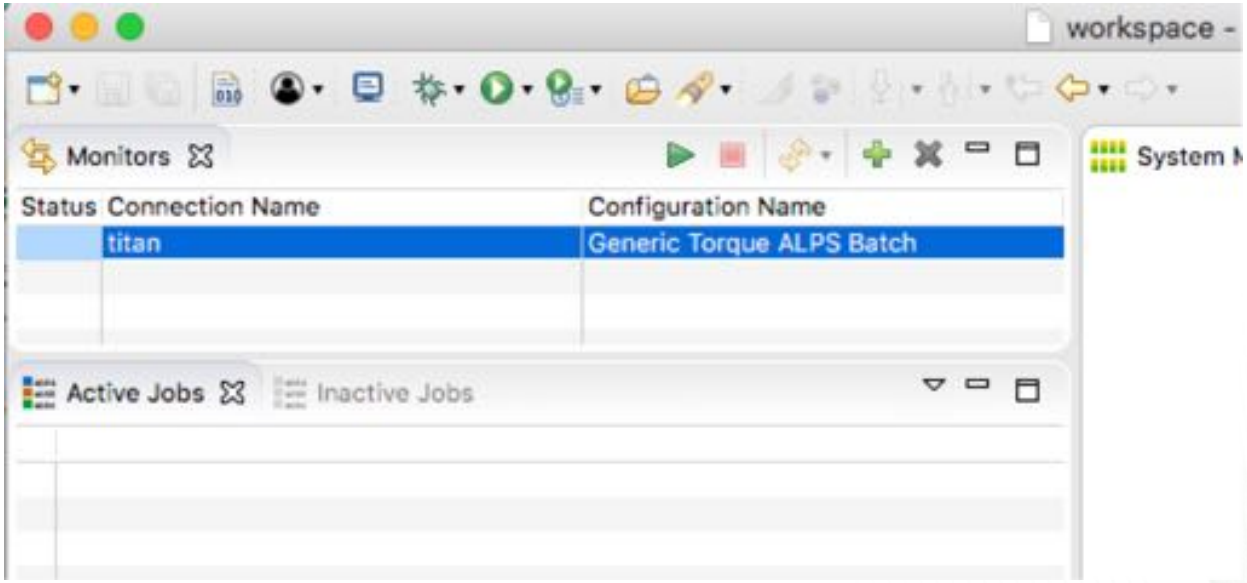
Manual Configuration

- From project properties
  - **C/C++ General > Preprocessor Include Paths, Macros, etc.**
  - Click on “Entries” and select “CDT User Setting Entries”
  - Click the “Add” button
  - Add an include directory or preprocessor macro using the dialog
- Unfortunately only one include or macro can be entered at a time

System Monitoring

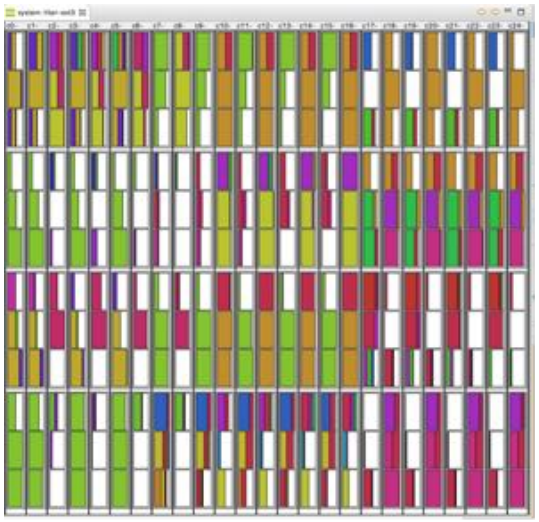
Managing Systems

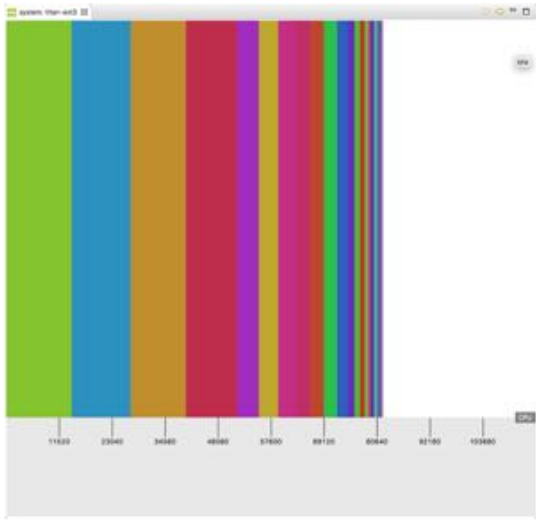
- Switch to the System Monitoring perspective
- Add/delete systems in the “Monitors” view



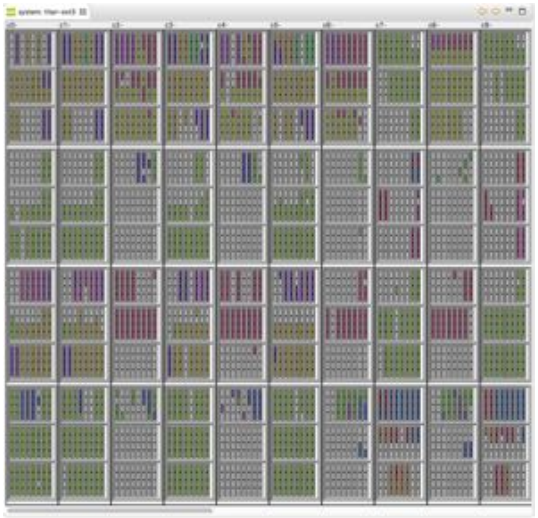
System View

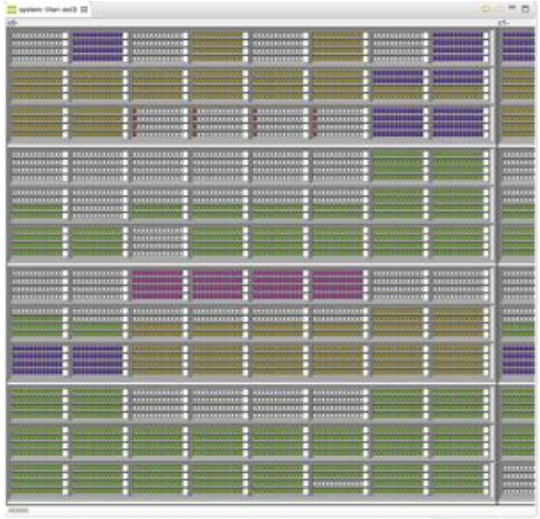
Zoom Out





Zoom In





Environment Modules

Using Environment Modules

- For the build:
  - Open project properties
  - Click on “Synchronize”
  - Select the remote configuration
  - Check the “Use an environment management system to customize the remote build environment”
- When submitting job:
  - Open run configuration for target machine
  - If supported, find the ”Modules to Load” entry and click “Configure”
  - Check the “Use an environment management system to customize the remote build environment”