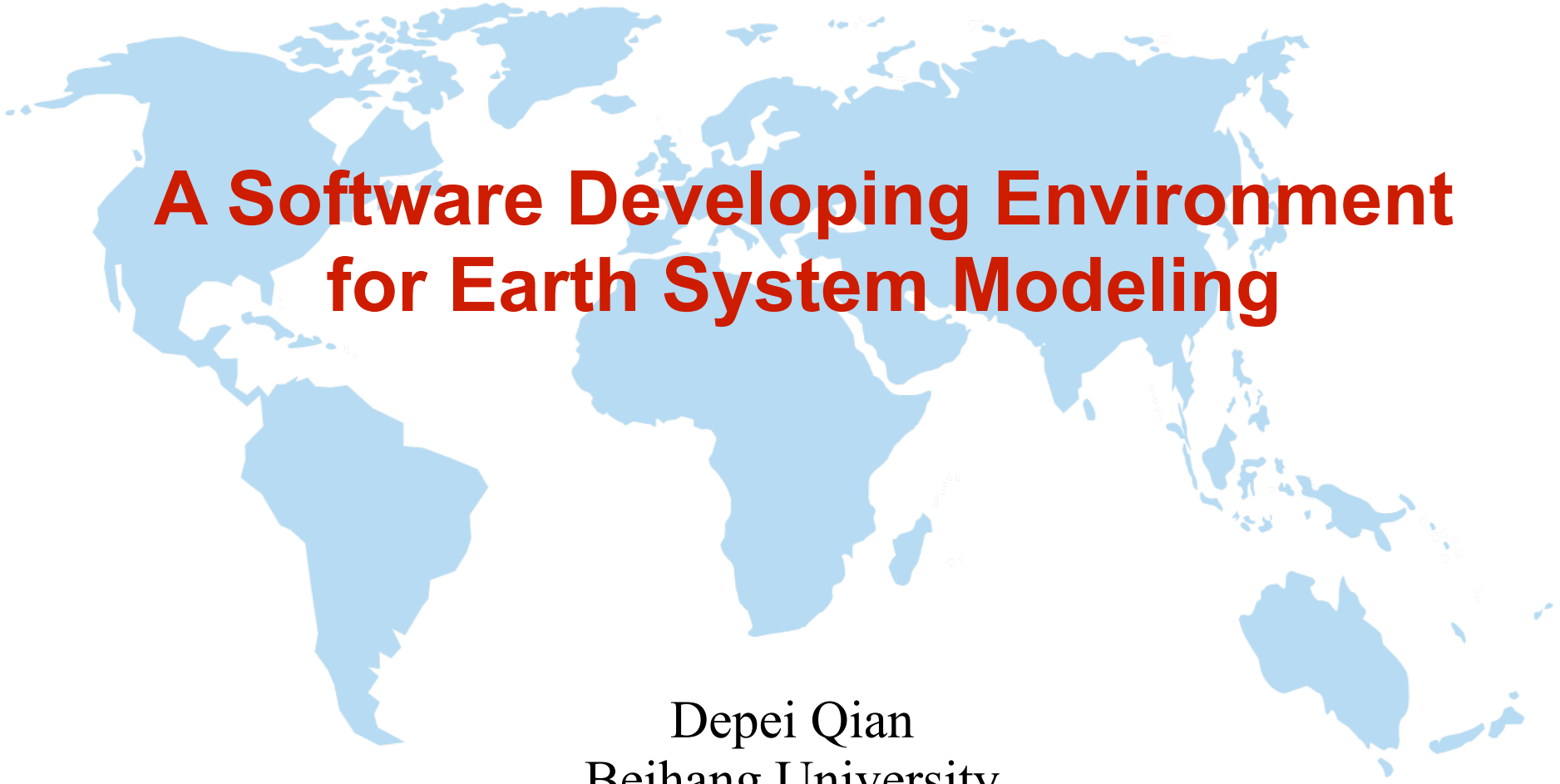


A Software Developing Environment for Earth System Modeling

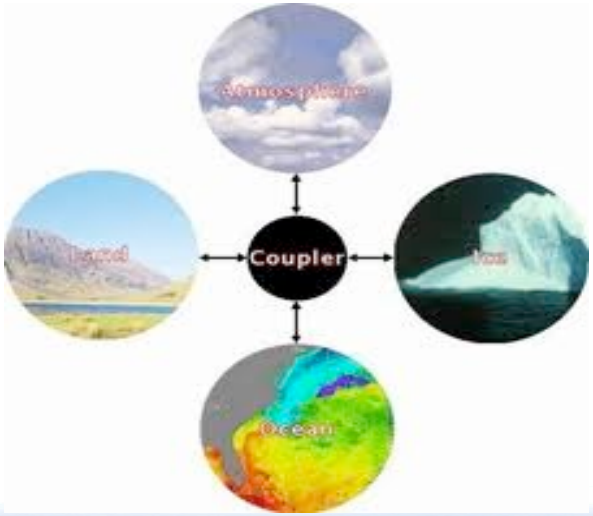
Depei Qian
Beihang University
CScADS Workshop, Snowbird, Utah
June 27, 2012

Outline

- Motivation
- Purpose and Significance
- Research Contents
- Technology Roadmap
- IDE design & development functions
- ESM Module Encapsulation
- Runtime Support
- Deployment

Motivation

- CCSM/CESM
  - Coupled climate models used to simulate the Earth's climate system
  - Includes four separate models: the atmosphere, oceans, land surface and sea ice
  - Support the research of global climate change in the past, present and future



Motivation (cont'd)

Lack of integrated HPC software development environment for earth system modeling

- Primitive program development approaches
- Various model development tools
- Program development coupled with machine hardware / software

Graphical IDE is in urgent need

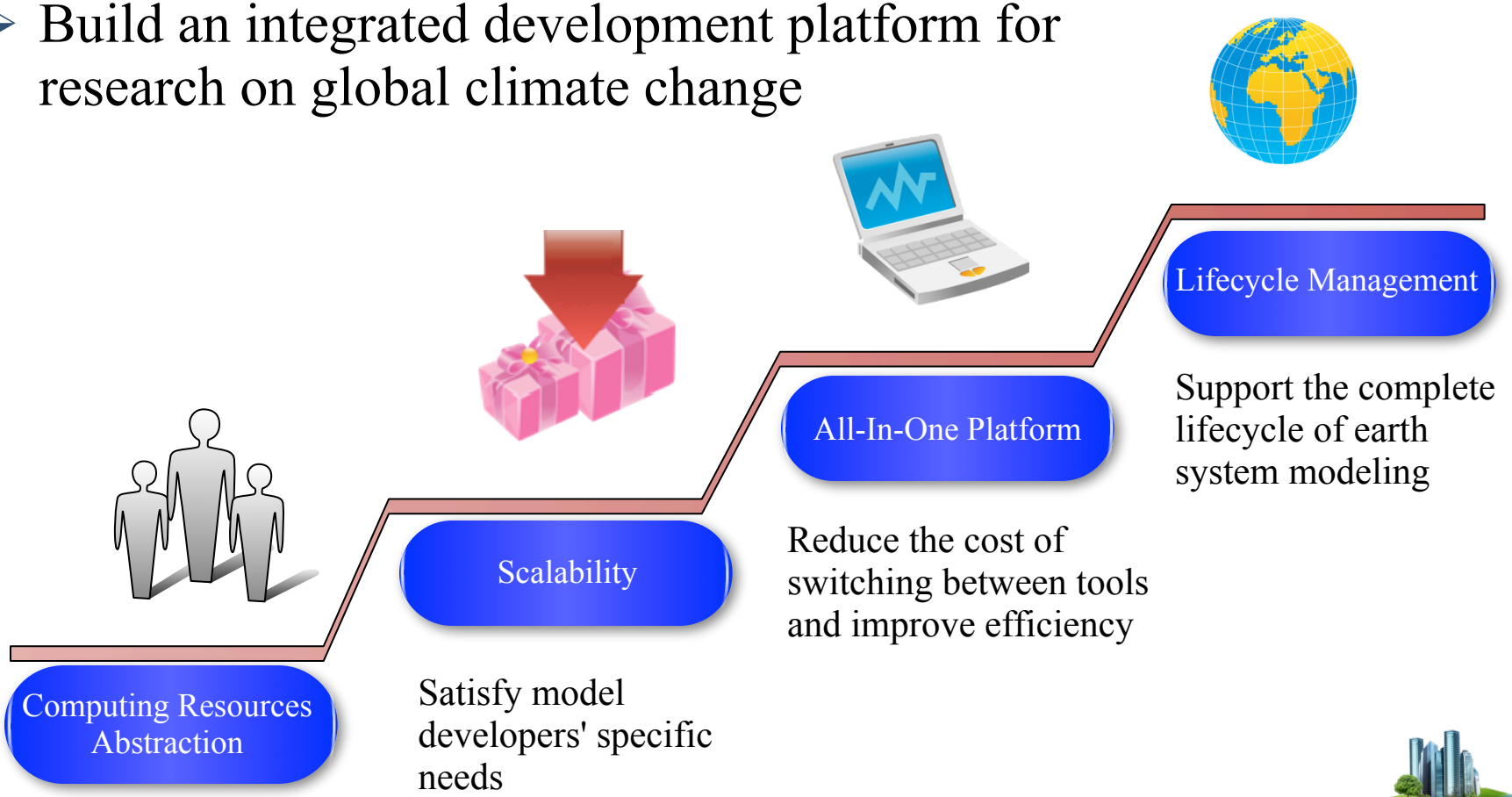
1. All-in-one platform, comprehensive tool library and the module / template library, to improve development efficiency
2. Decoupled with software/hardware details to allow researchers to focus on model development
3. Customized development environment to provide lifecycle support in developing the earth system models

Outline

- Motivation
- Purpose and Significance
- Research Contents
- Technology Roadmap
- IDE design & development functions
- ESM Module Encapsulation
- Runtime Support
- Deployment

Purpose and significance

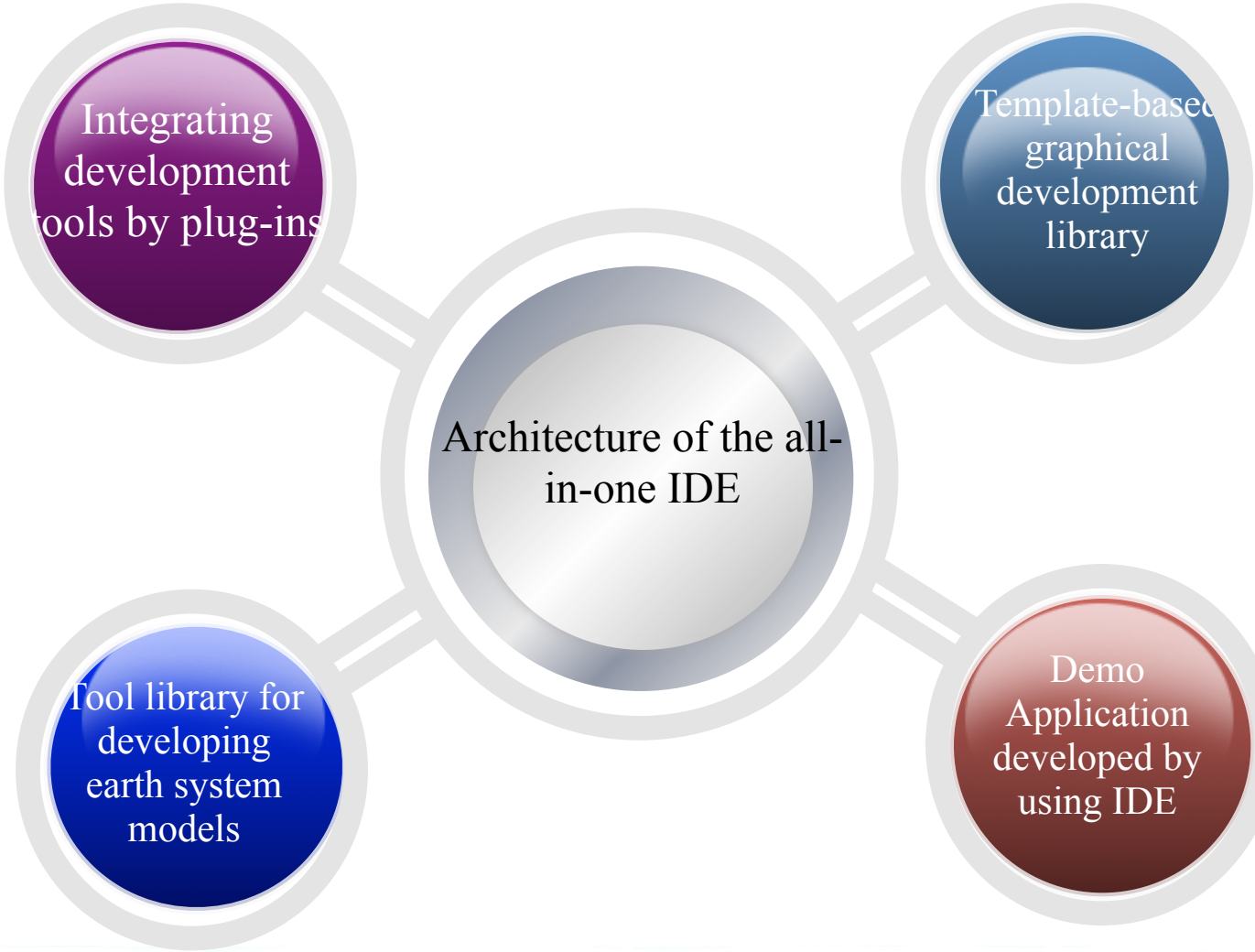
- Focus on support to the development of earth system models
- Build an integrated development platform for research on global climate change



Outline

- Motivation
- Purpose and Significance
- Research Contents
- Technology Roadmap
- IDE design & development functions
- ESM Module Encapsulation
- Runtime Support
- Deployment

Research content



Outline

- Motivation
- Purpose and Significance
- Research Contents
- Technology Roadmap
- IDE design & development functions
- ESM Module Encapsulation
- Runtime Support
- Deployment

Technology roadmap

Key Technologies

- Remote code organization and management
- Reusable code blocks for model development
- Description of heterogeneous HPC resources
- Parallel support, e.g, MPI & JASMIN
- Portability, scalability and feasibility of platform
- Remote compilation, debugging and execution
- Graphical drag&drop -based program coding
- Integrating existing tools

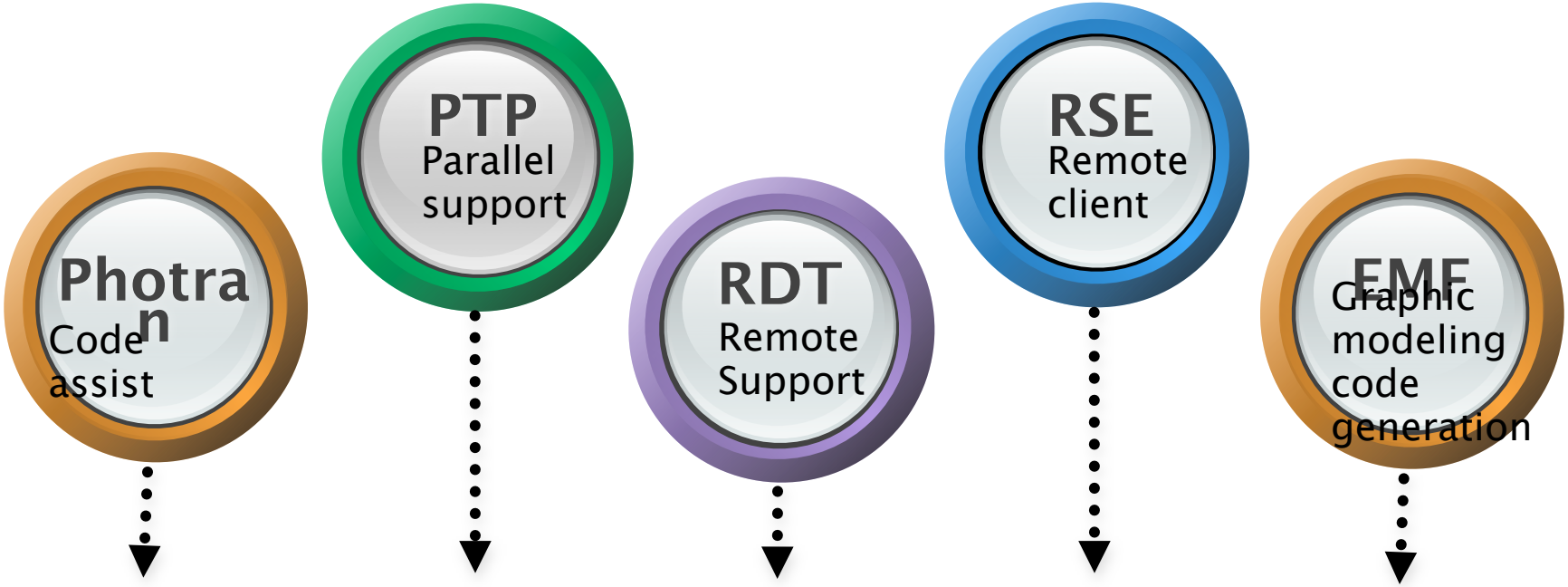
research outcomes

- Module / Template Library
- Tool Library
- IDE interface

Plug-in Integration &
Graphical interface support

Integrated Development Platform

Technology roadmap



Overall strategy

Based on the open source Eclipse: mature community, abundant references, plenty of free plug-ins, wide range of languages

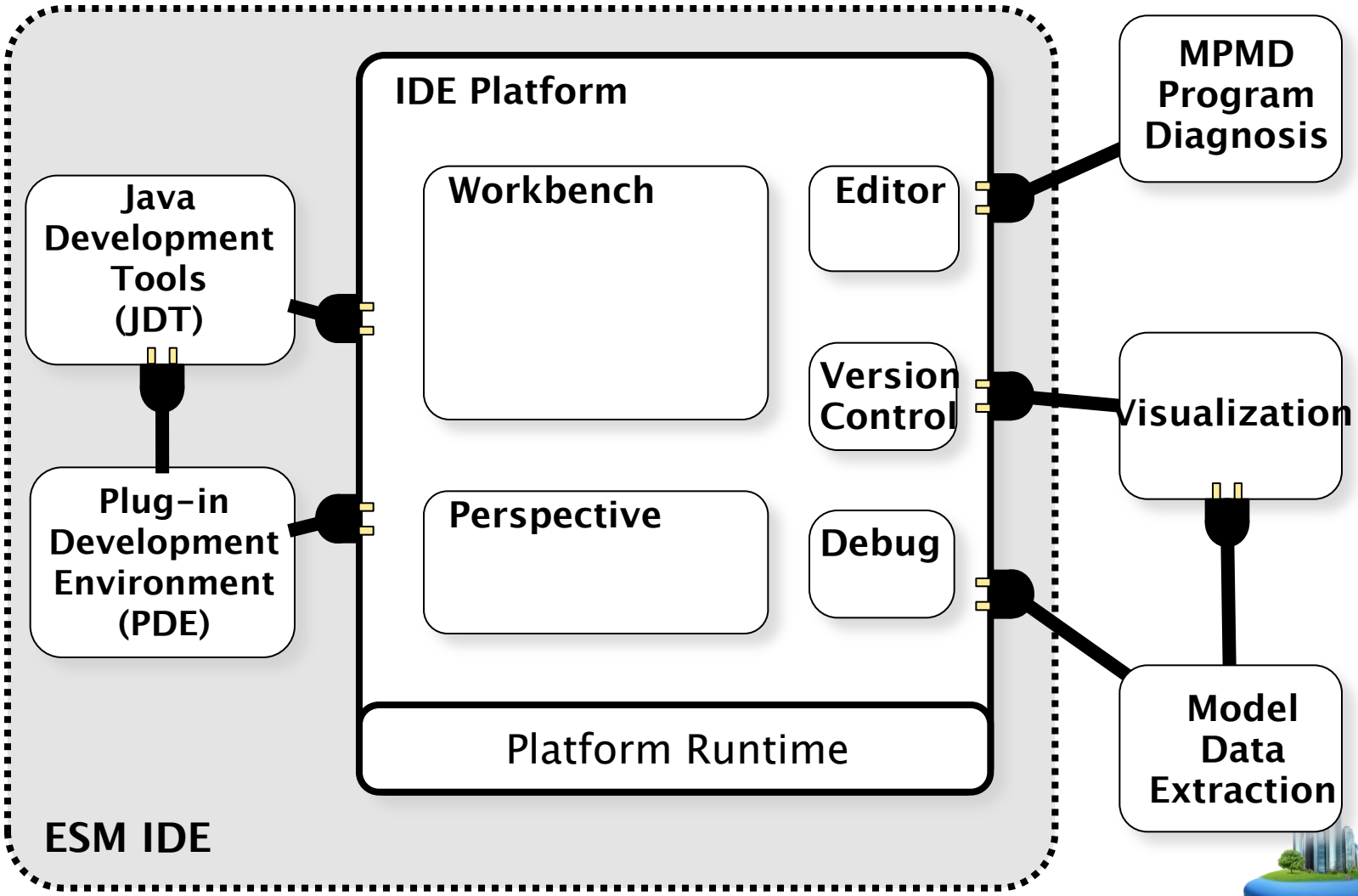
Outline

- Motivation
- Purpose and Significance
- Research Contents
- Technology Roadmap
- IDE design & development functions
- ESM Module Encapsulation
- Runtime Support
- Deployment

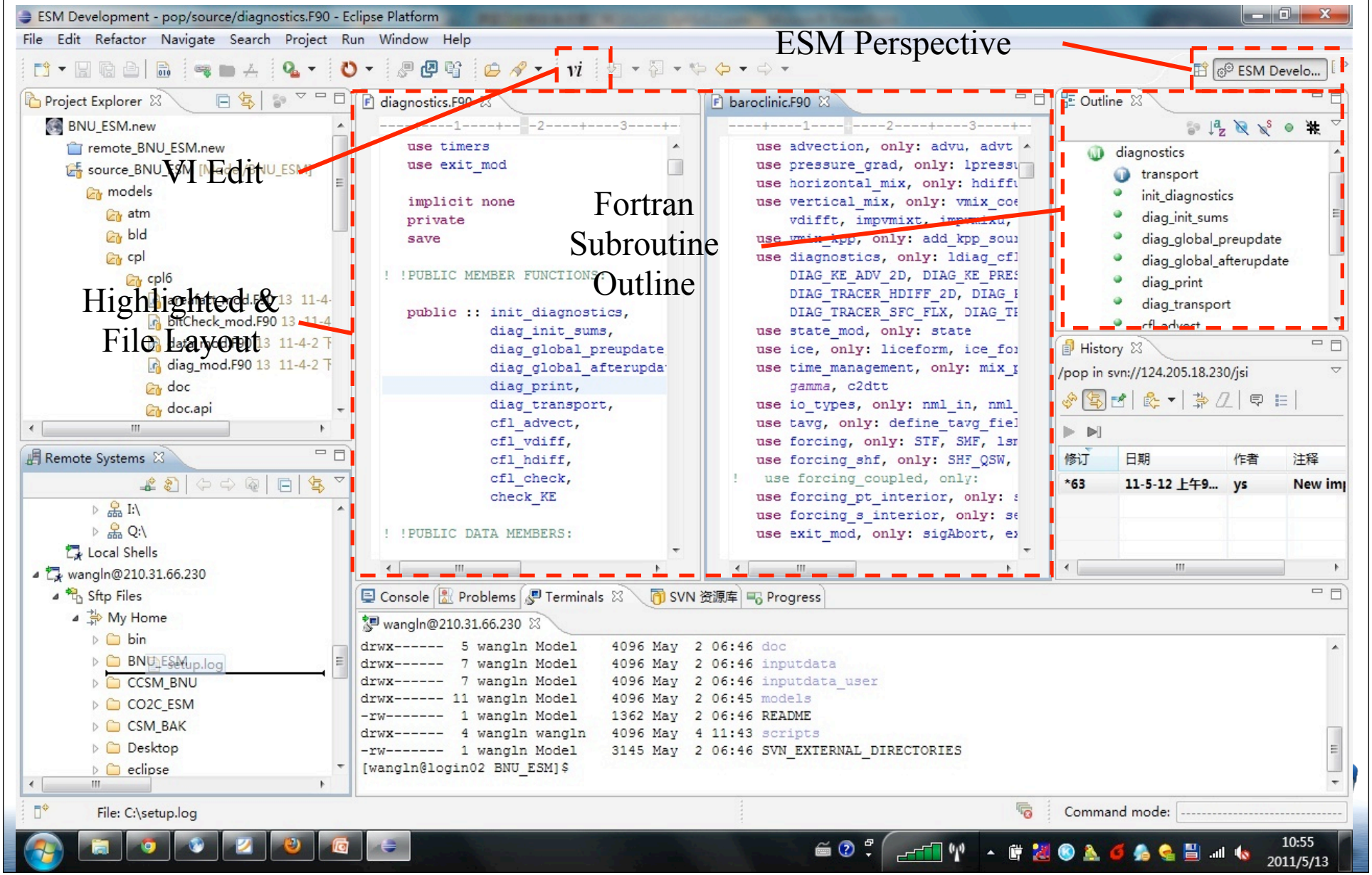
IDE design & development functions

- Works have been done：
  - IDE overall design
  - Fortran code assistant
  - Source code version control
  - Remote online multi-collaboration
  - Integration of other useful tools

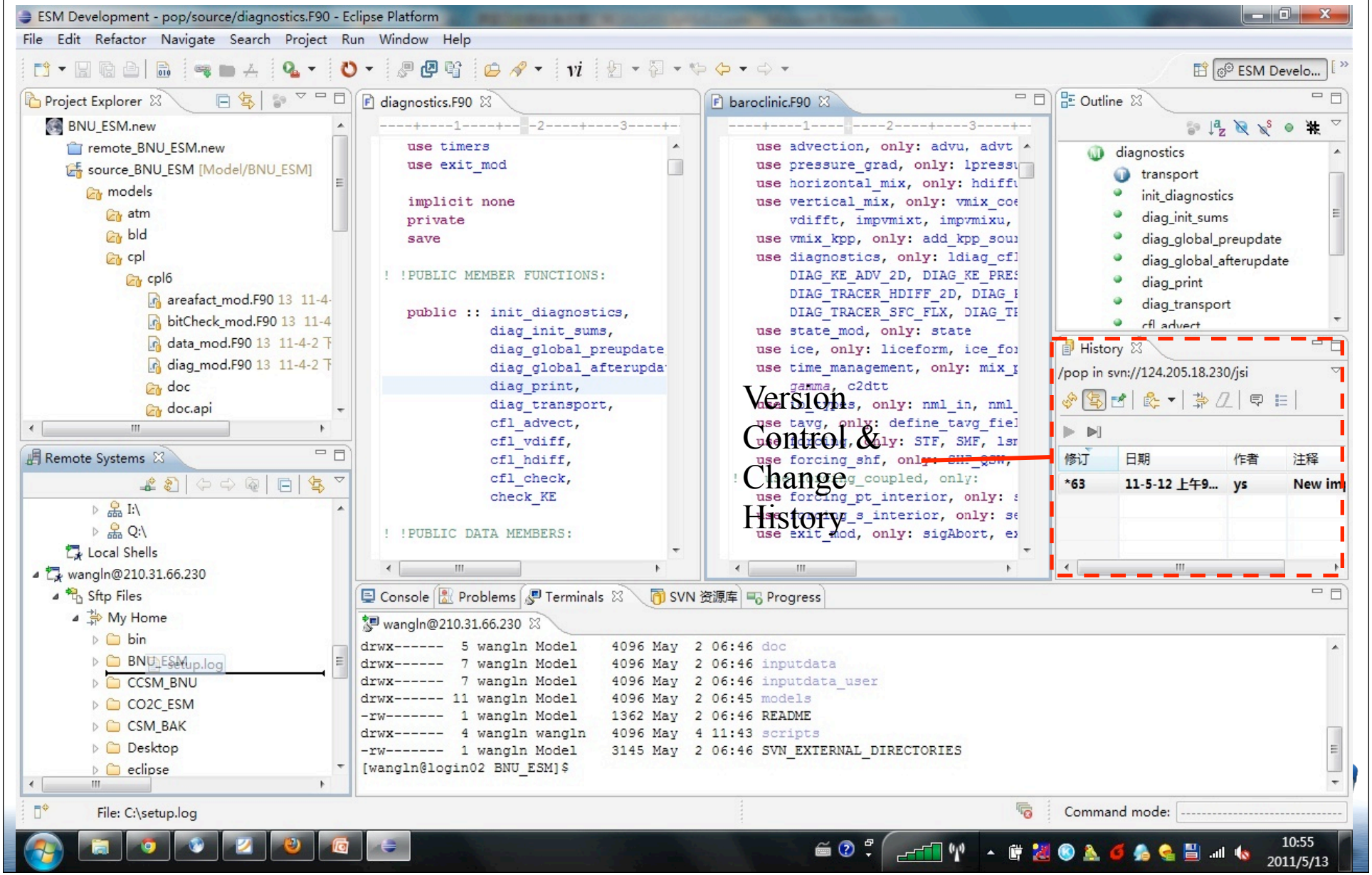
IDE overall design



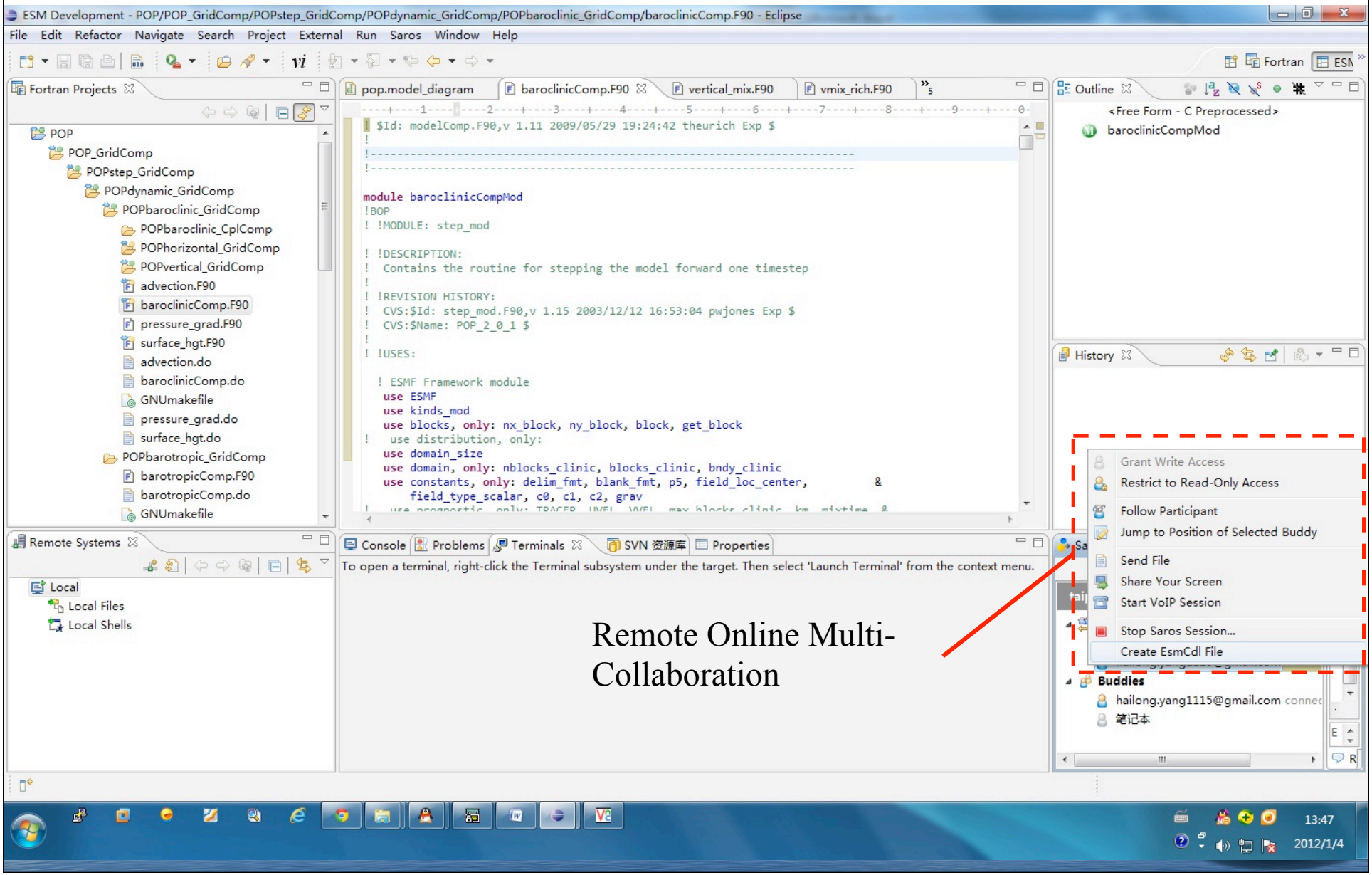
Fortran code assistant



Source code version control

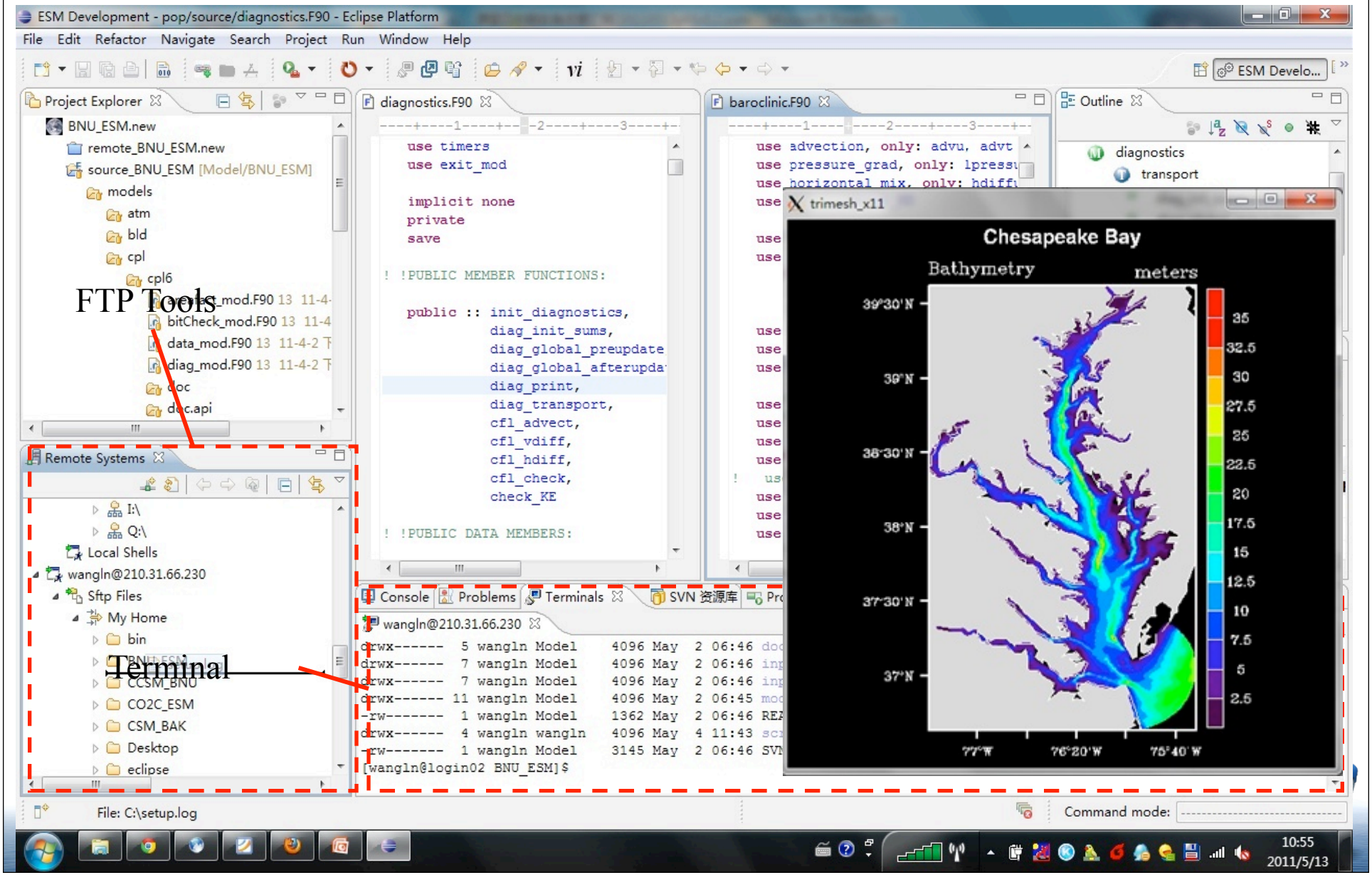


Remote online multi-collaboration



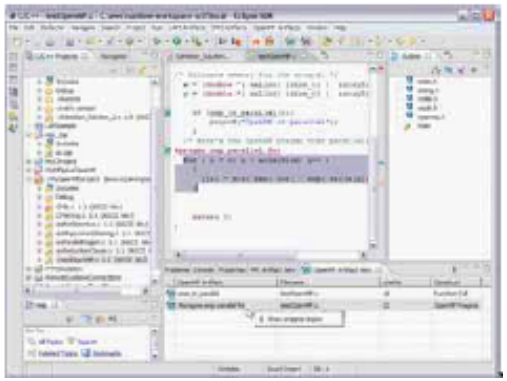
Remote Online Multi-Collaboration

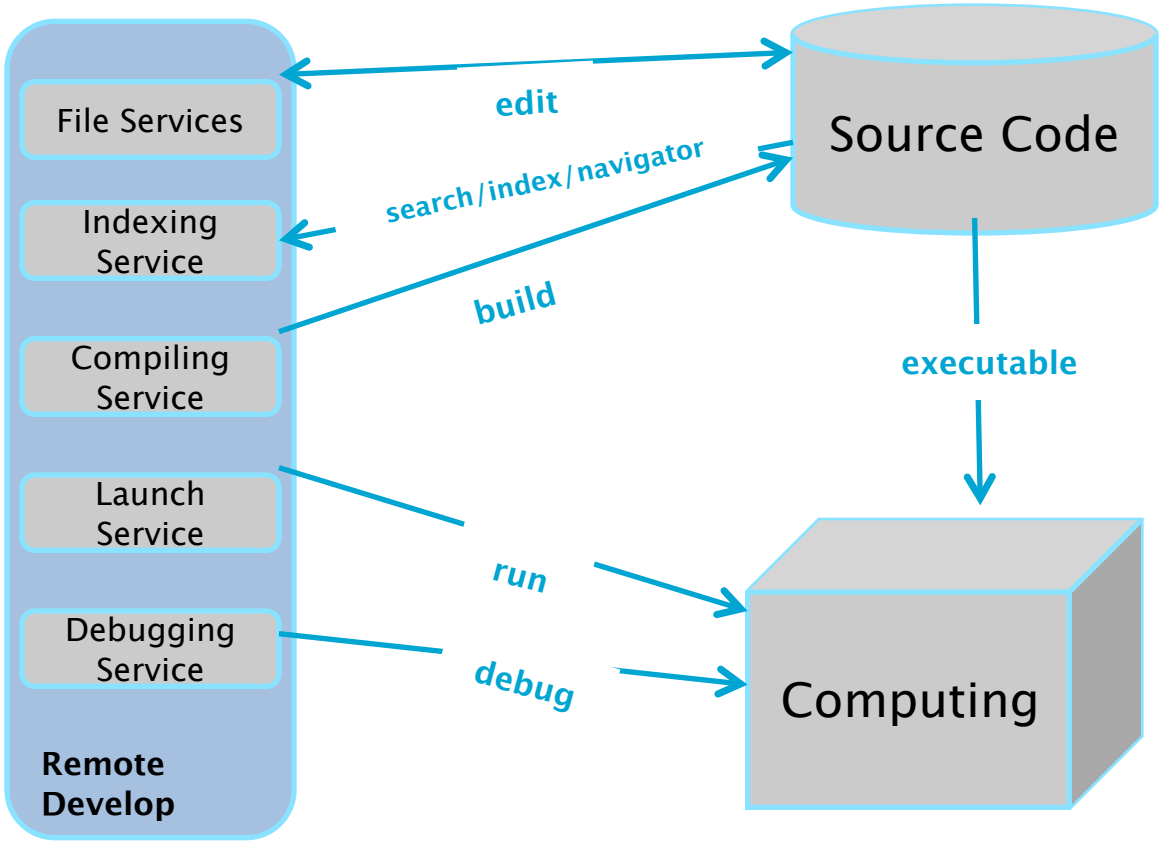
Integrated tools



Remote function invocation

- Remote editing, compilation, debugging and execution



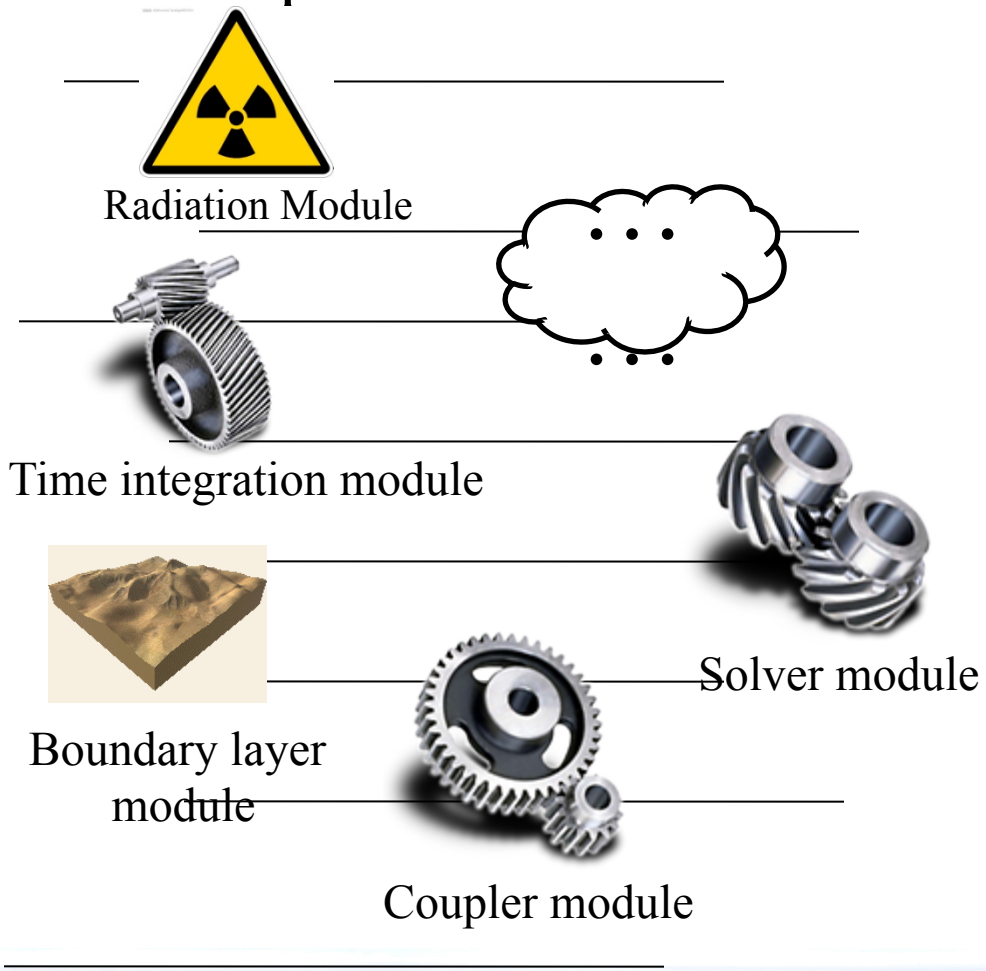


Outline

- Motivation
- Purpose and Significance
- Research Contents
- Technology Roadmap
- IDE design & development functions
- ESM Module Encapsulation
- Runtime Support
- Deployment

ESM module encapsulation

- Encapsulation of reusable code blocks for model development



Modular Encapsulation Specification

module unit
- High-performance
- Reusable

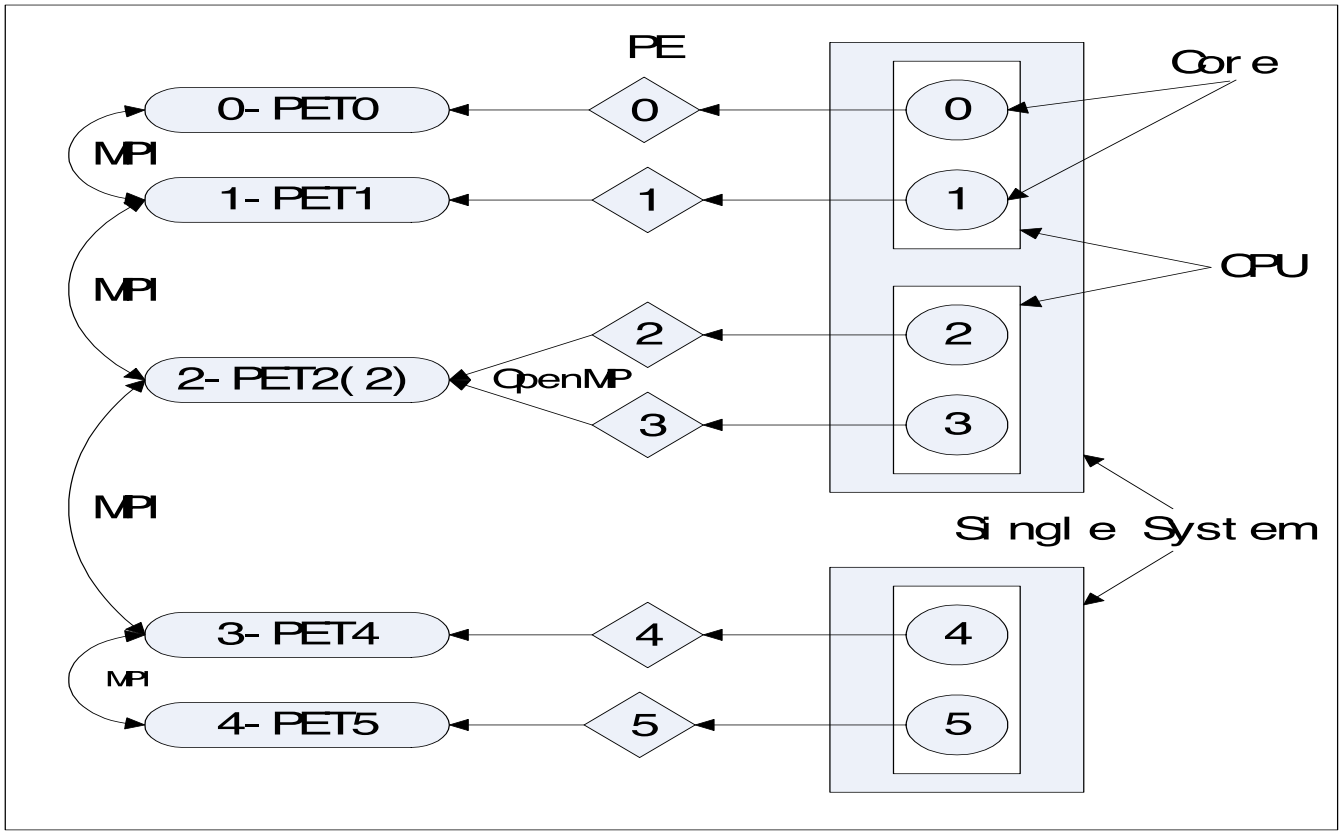
model module library
- Enrichment
- Optimization

Program decomposition

- Decomposition of POP (Parallel Ocean Program) into modules

Resource allocation

- Parallel computing resources allocation for POP components



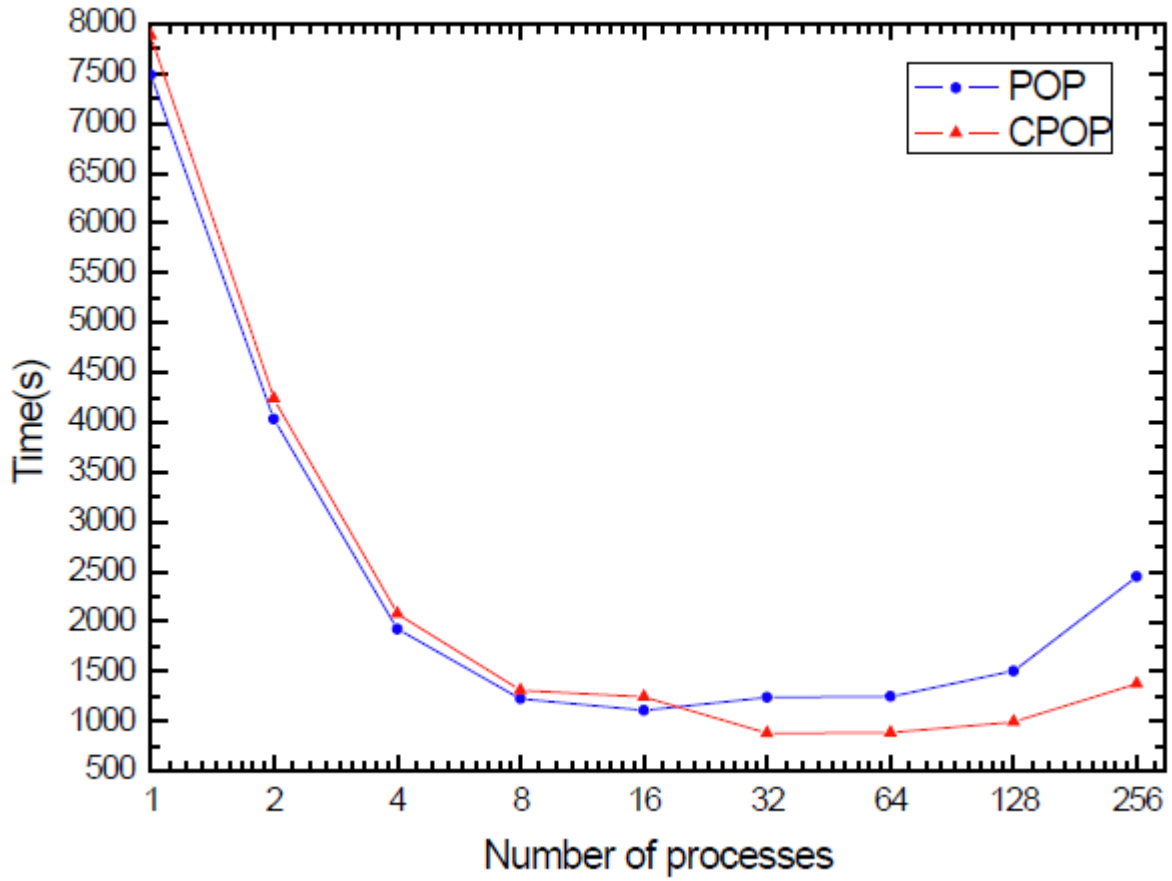
Dynamic process organization

- Dynamic process organization based on grid partition

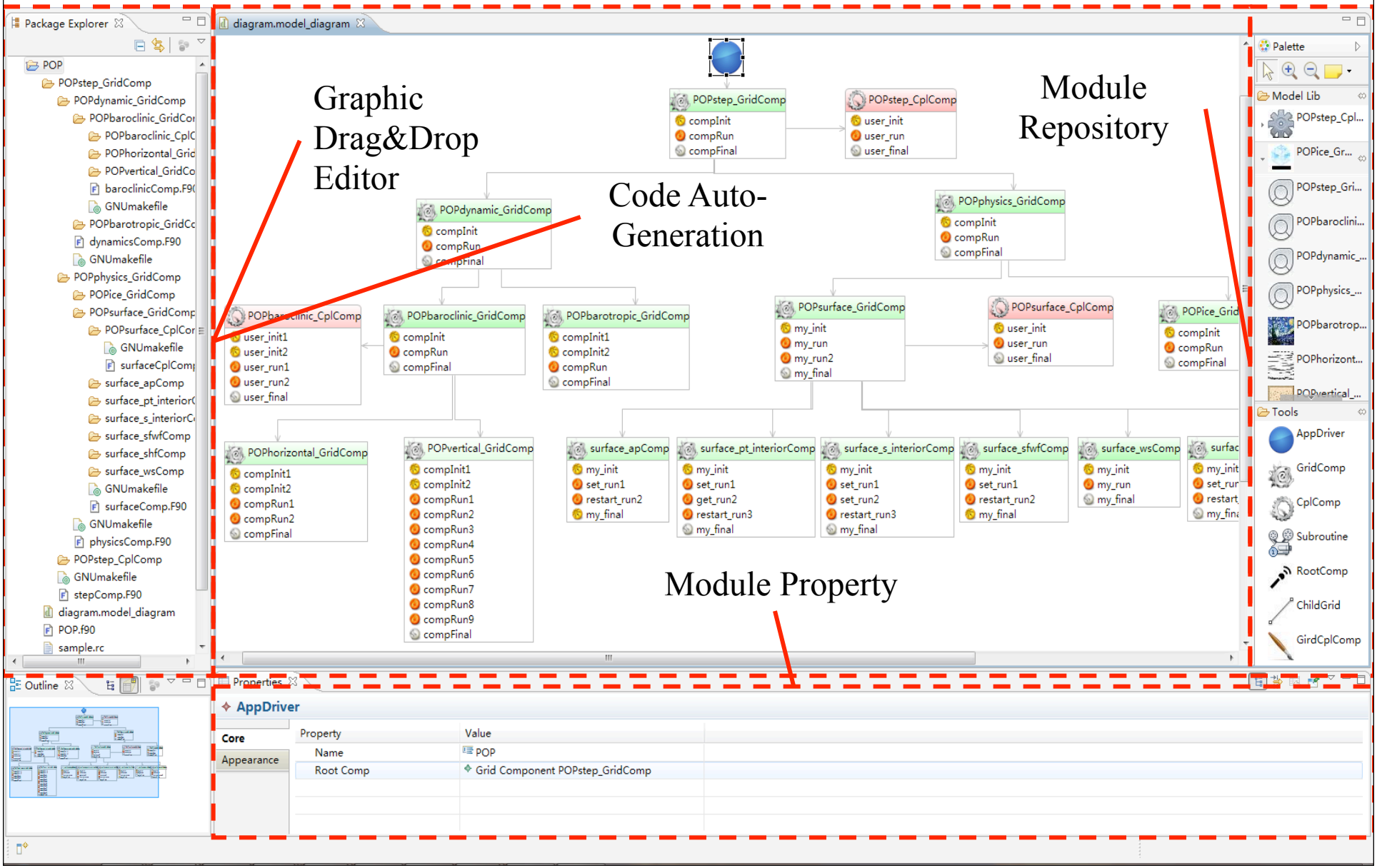
Validation and performance evaluation

- Functional validation and performance evaluation

Number of Porcesses
1
2
4
8
16
32
64
128
256



Code generation



Component specification

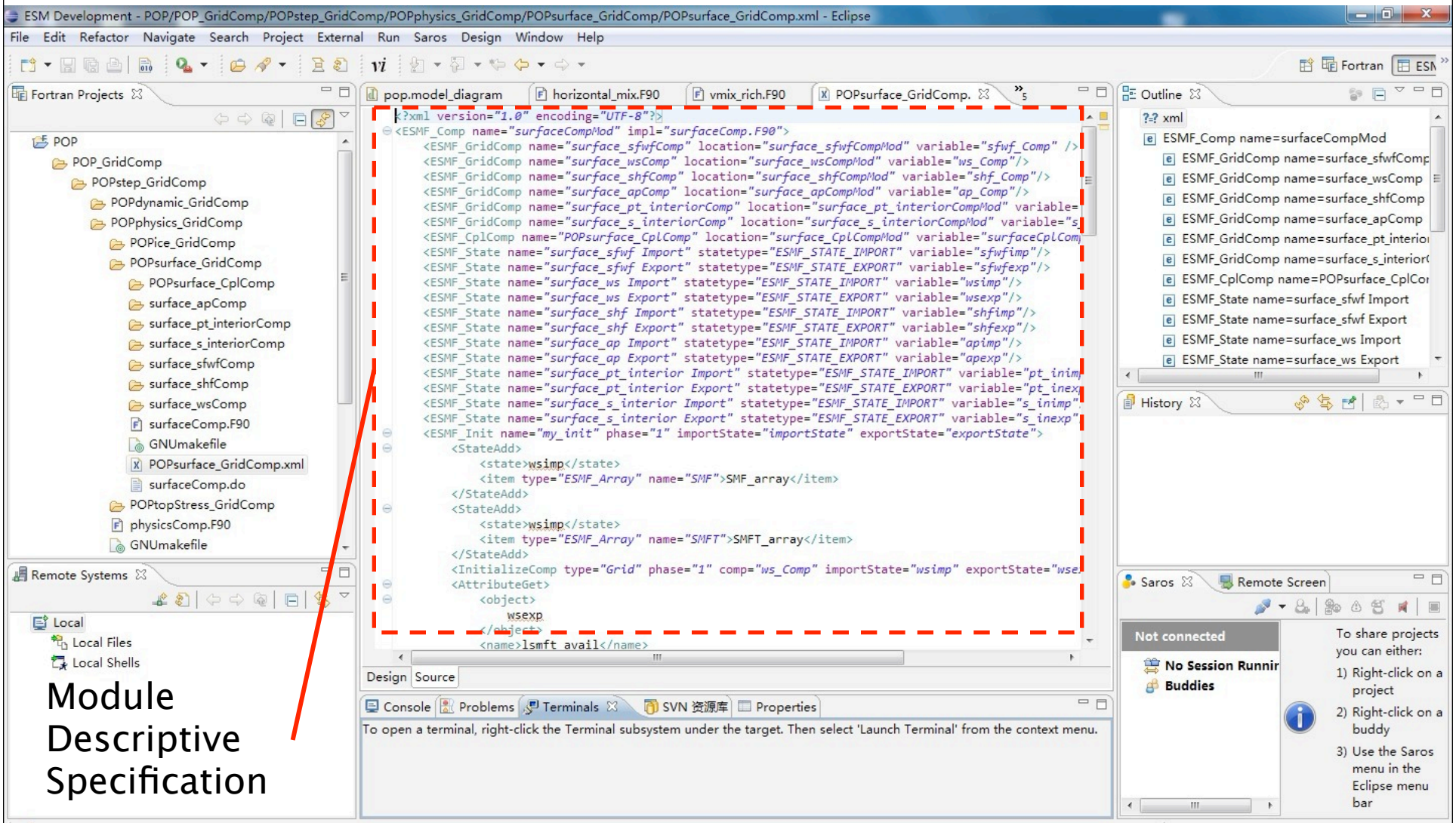
- Specification for component encapsulation
  - Templates to guide component development
  - Each component is associated with a xml-based description file before importing to the component library

GridComponent

Core | Appearance

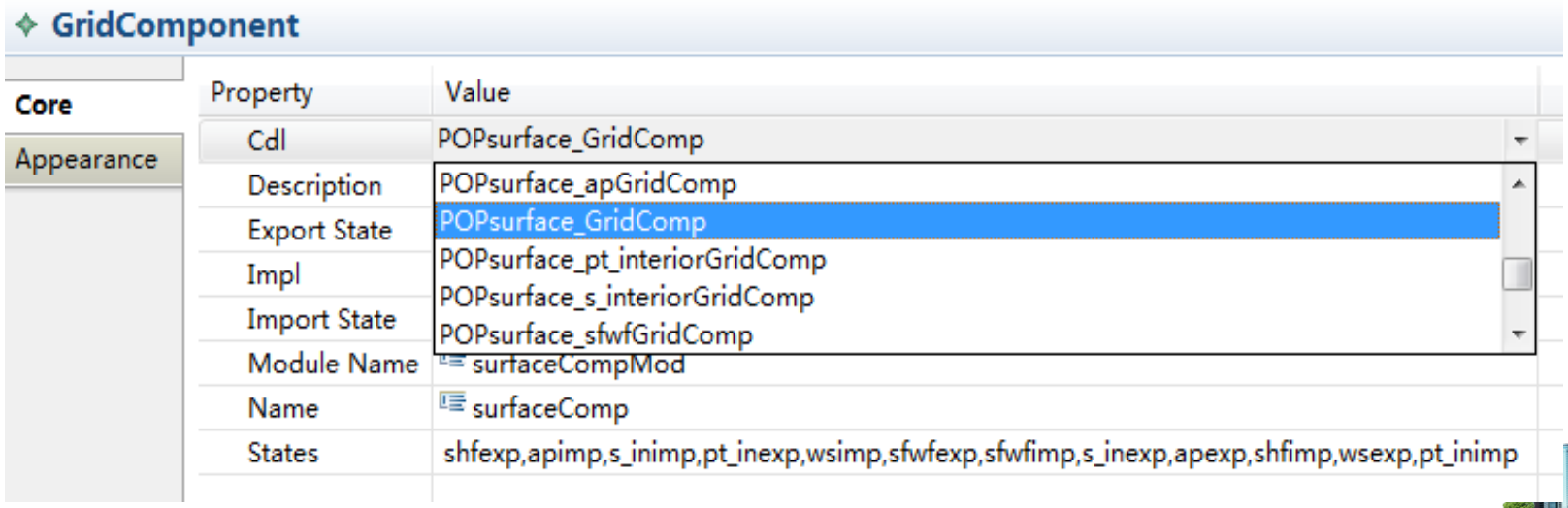
Property	Value
Cdl	POPsurface_GridComp
Description	POPsurface_GridComp
Export State	
Impl	POPsurface_GridComp
Import State	
Module Name	surfaceCompMod
Name	surfaceComp
States	shfexp,apimp,s_inimp,pt_inexp,wsimp,sfwfexp,sfwfimp,s_inexp,apexp,shfimp,wsexp,pt_inimp

Component specification (cont'd)

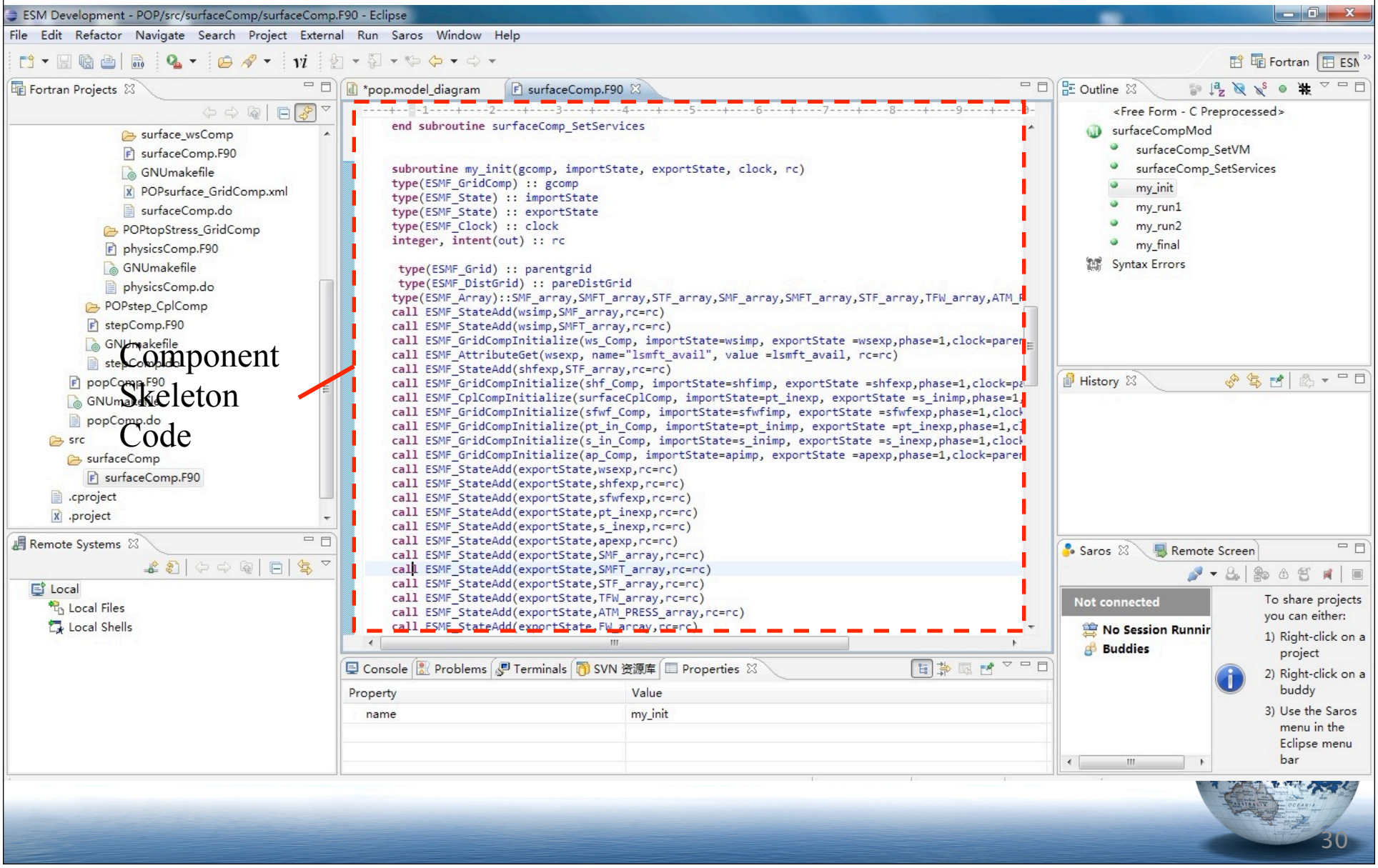


Code generation (cont'd)

- Drag-Drop based component usage
  - Automatically locate and intercept the information about the component based on CDL, and then fill in the panel of attributes about this component
  - Component substitution can be done easily through selection

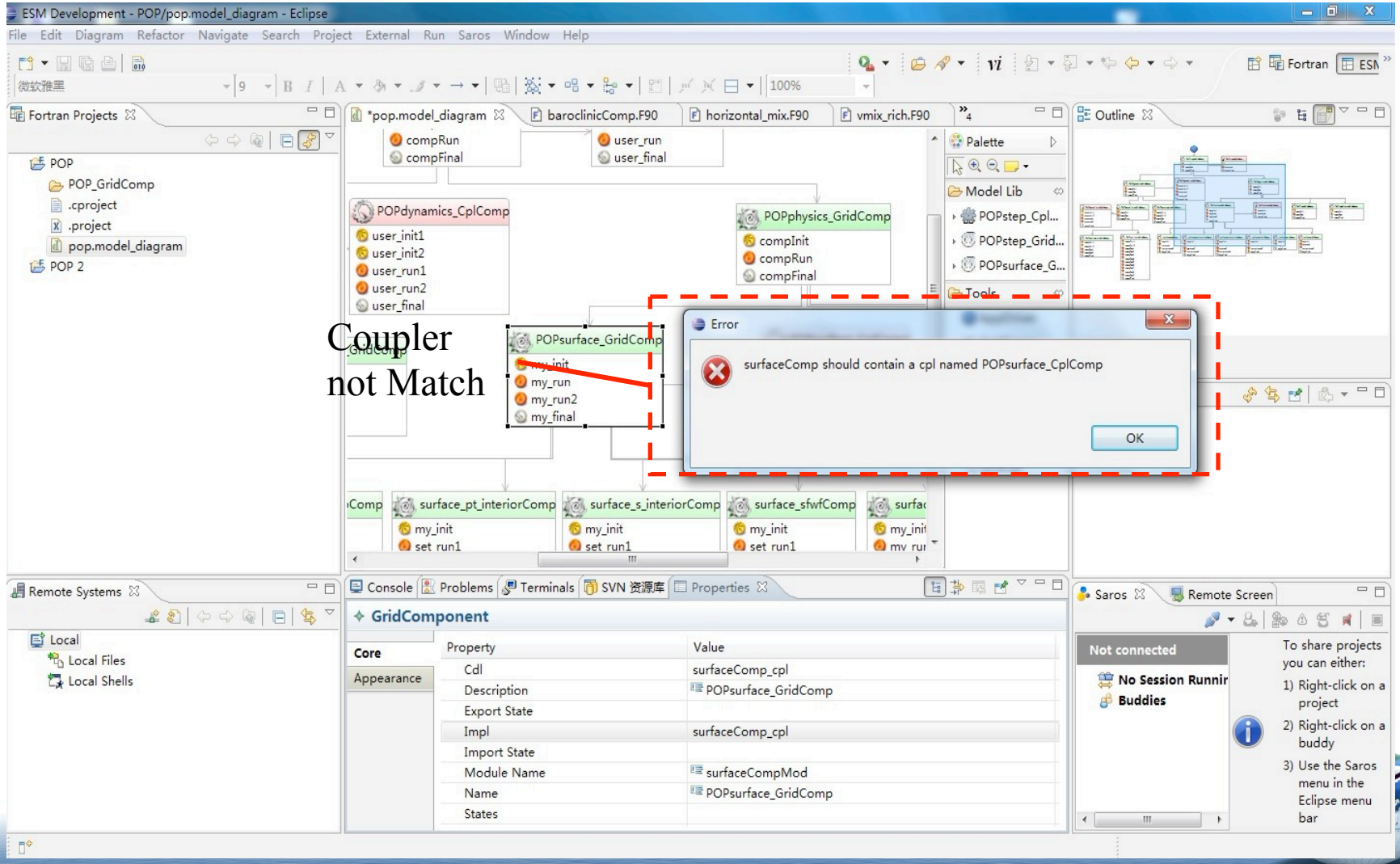


Code generation (cont'd)



Module auto-verification

- Primary auto-verification for module validity



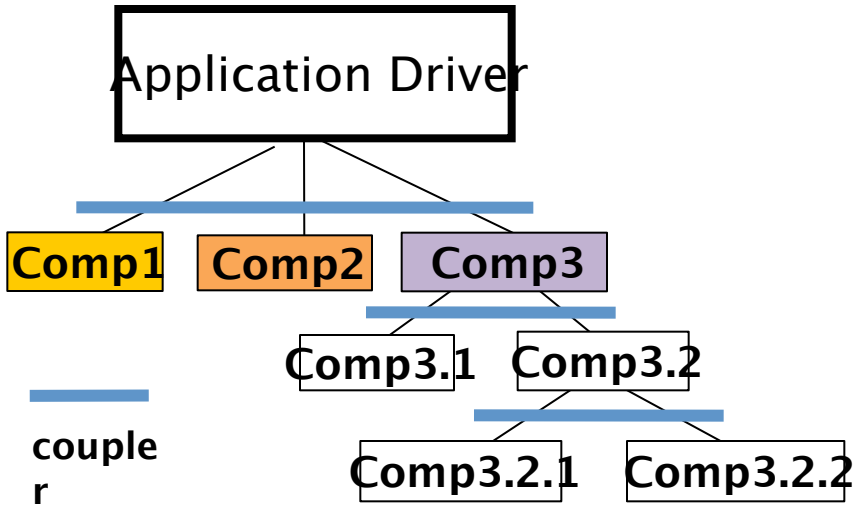
Model components substitution

- Model component encapsulation allows the components developed by different groups or communities to be inter-changeable, reducing duplicated programming efforts
  - Component substitution is demonstrated between POP and MOM
  - MOM
    - a three-dimensional ocean circulation model designed primarily by the Geophysical Fluid Dynamics Laboratory (GFDL) of the National Oceanic and Atmospheric Administration
  - POP
    - developed and supported by Los Alamos National Laboratory (LANL)

POP vs MOM

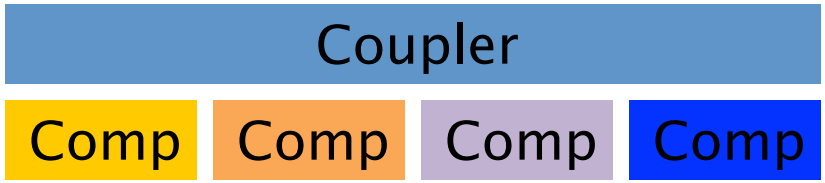
- Framework Comparison
  - POP is based on ESMF, while MOM is based on FMS

ESMF



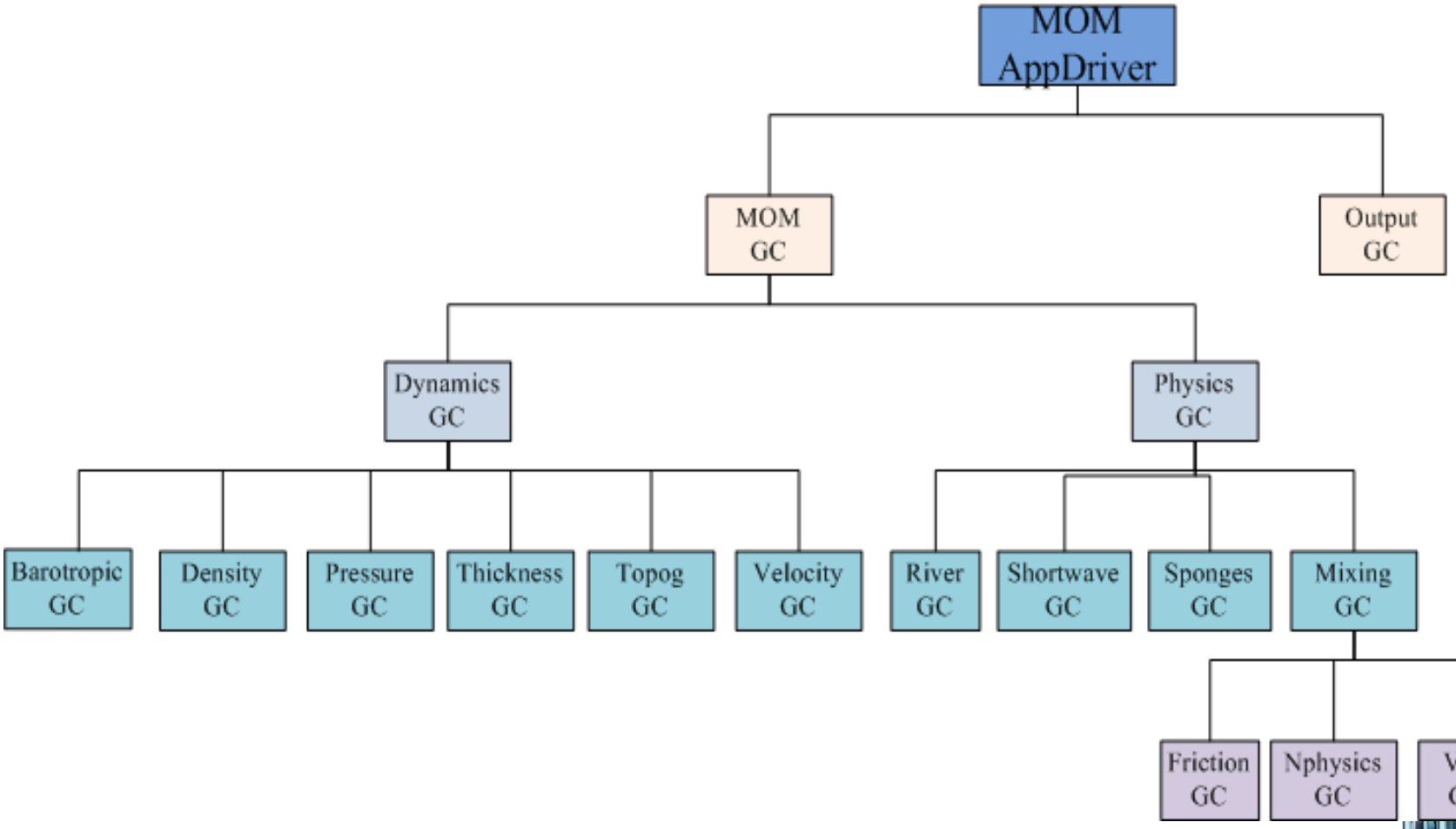
- Child components are executed simultaneously
- Invocation among components is nested
- Multiple couplers
- Data is transferred through data structure called state

FMS



- Components are executed sequentially
- All component resides in flat
- Single coupler
- Data is transferred through shared public variables

MOM code structure analysis



MOM code structure analysis

- Analyzing the code structure of MOM
  - MOM component interface
    - init() #initialize and configure
    - run()
      - update() #update data, time step, status and so on
      - restart() #reload configuration and restart
    - final(): #release the resouces

The above interface would make sure that the components derived from MOM, which is based on FMS specification, are compatible with components derived from POP which is based on ESMF, so that the components are substitutable

Outline

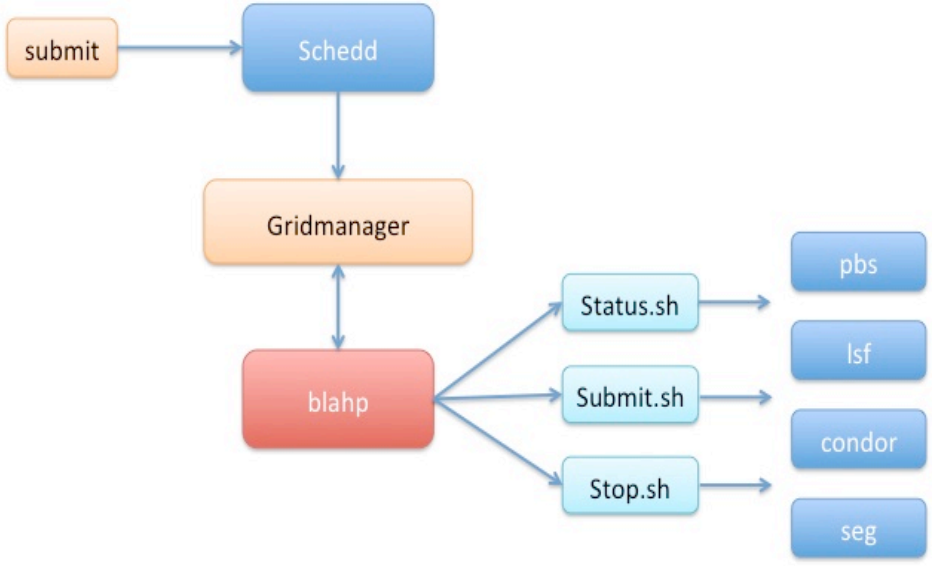
- Motivation
- Purpose and Significance
- Research Contents
- Technology Roadmap
- IDE design & development functions
- ESM Module Encapsulation
- Runtime Support
- Deployment

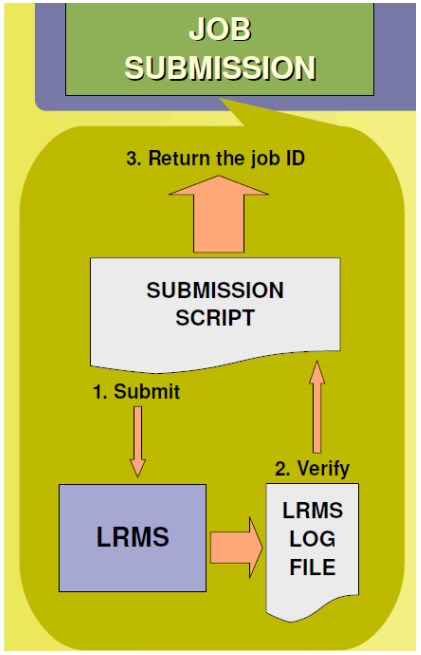
Runtime support

- Work have been done：
  - Seamless integration with Slurm batch system with an easy-to-use management portal
  - Support ESM program execution monitoring and trace analysis

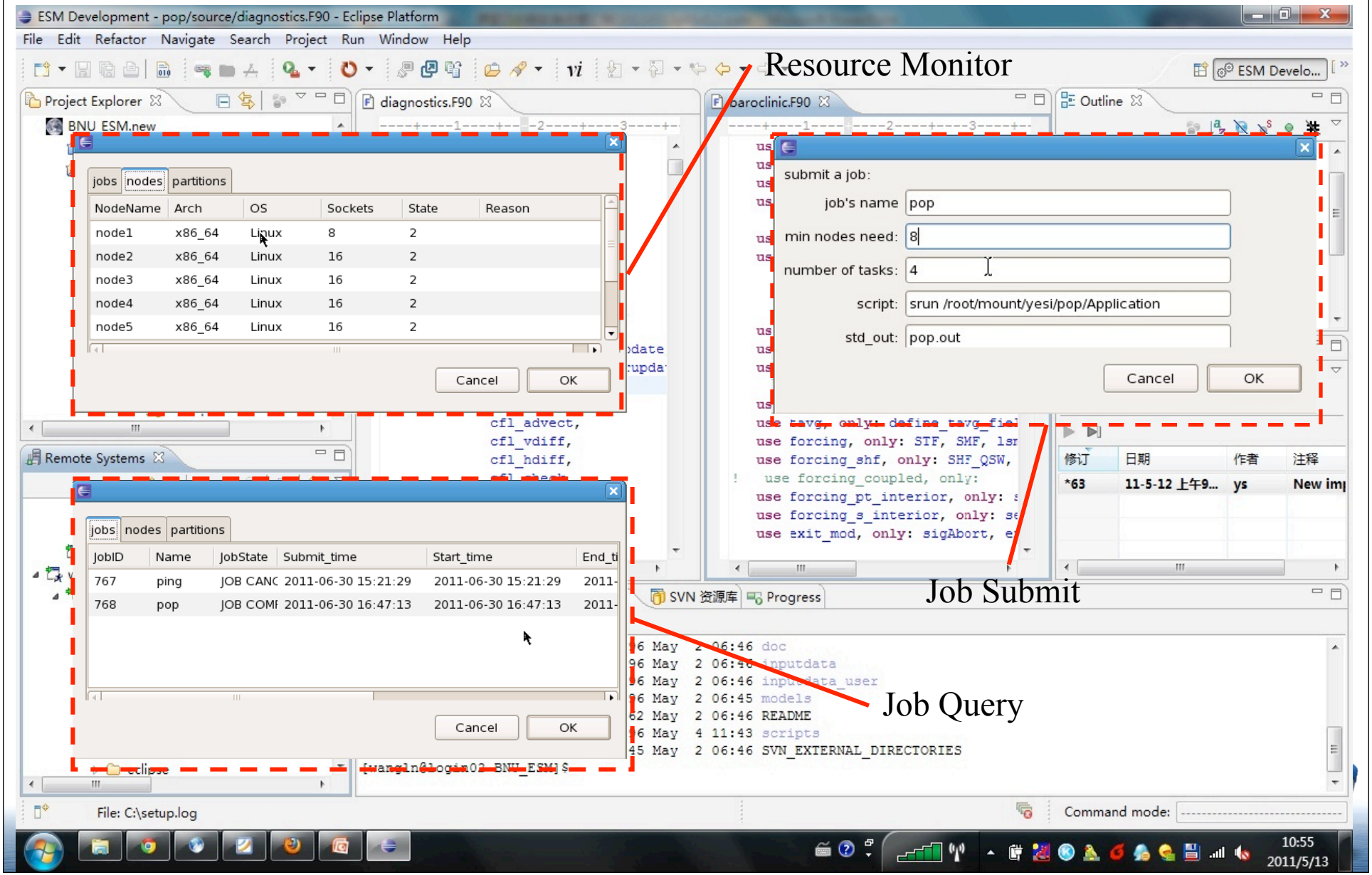
Integration with the batch system

- Integration with Condor to provide persistent execution environment
- The jobs submitted to the IDE are delegated to underlying slurm batch system
  - Extend condor to talk to slurm
  - Submit, query and management

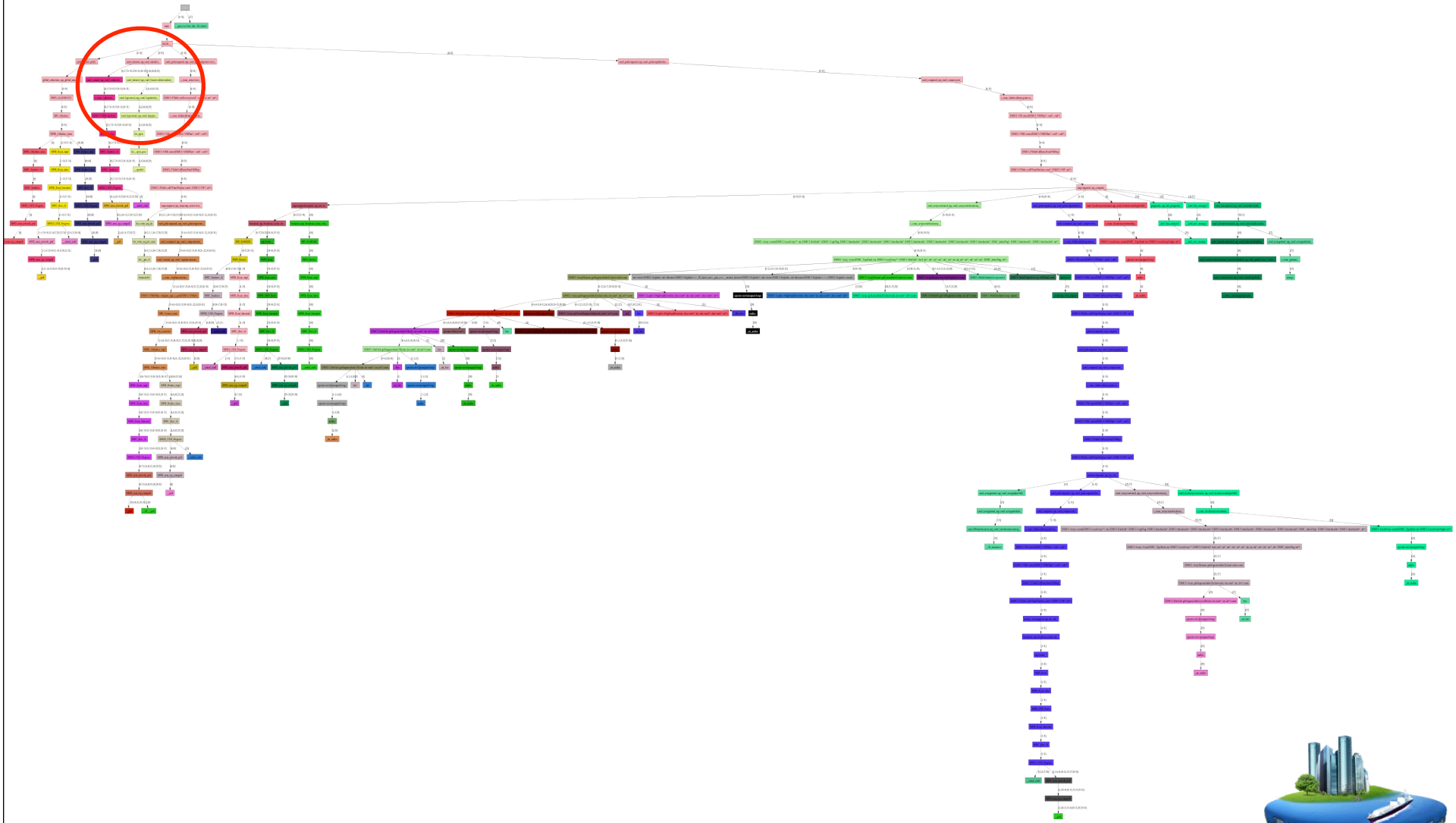




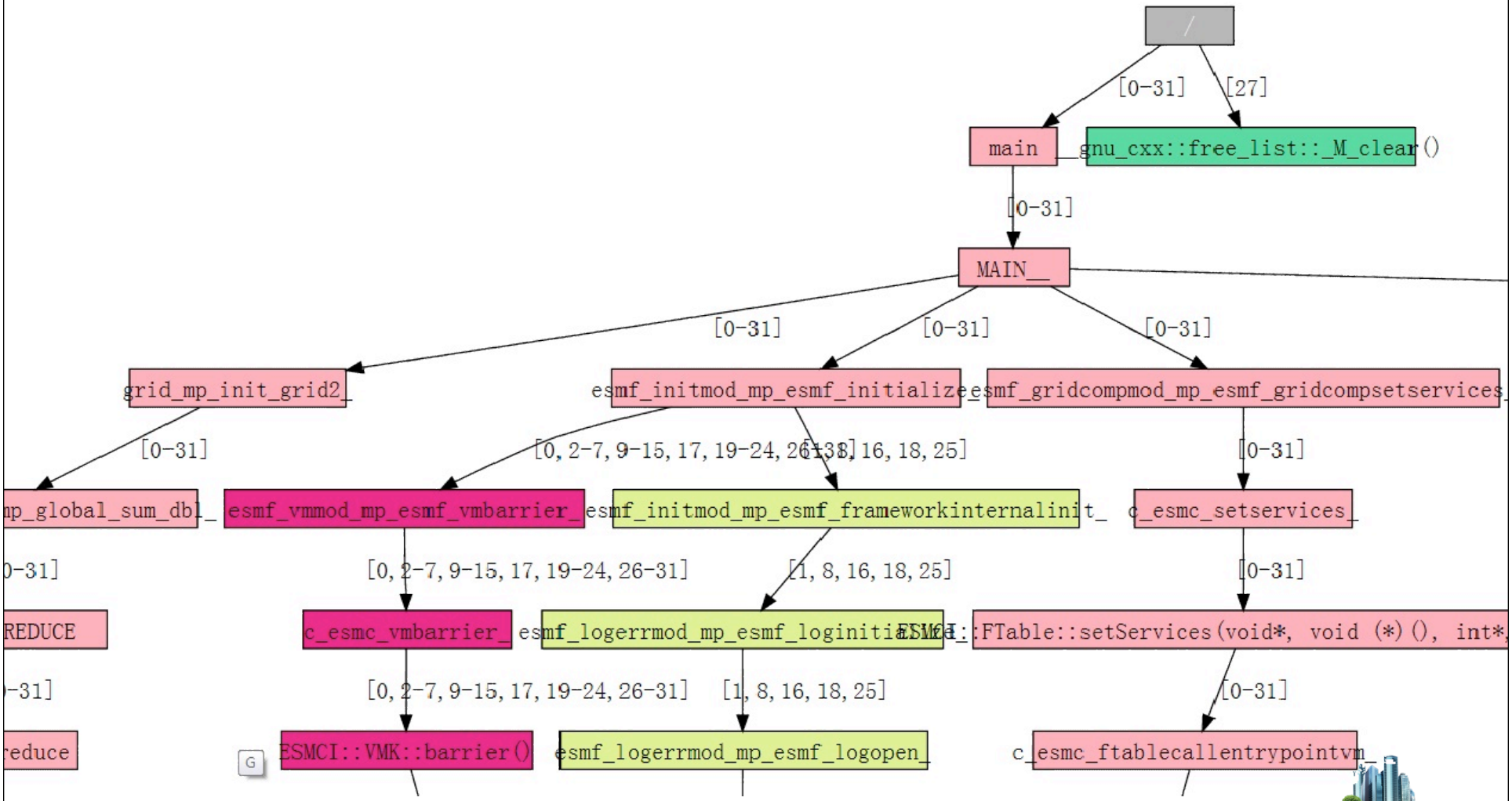
Program execution monitoring



Trace analysis based on STAT



Trace analysis based on STAT (cont'd)



Outline

- Motivation
- Purpose and Significance
- Research Contents
- Technology Roadmap
- IDE design & development functions
- ESM Module Encapsulation
- Runtime Support
- Deployment

Deployment

- Alpha version released (March 2011)
- Five versions (v0.2, v0.4, v0.6, v1.0, v2.0) released so far
- Deployed in Beijing Normal University (since March 2011)
  - Pilot usage of four versions (v0.4, v0.6, v1.0, v2.0)
  - Used in BNU_ESM development
  - Latest version is v2.0
- Deployed in Institute of Atmospheric Physics, Chinese Academy of Sciences (since March 2011)
  - Latest version is v2.0
- Deployed in National Climate Center (since June 2011)
  - Latest version is v2.0
- IDE and model encapsulation approach adopted by the coupler development project (since July 2011)

Thank You!


AUSTRALIA
OCEANIA
NEW ZEALAND
VANUATU
TONGA
Northern Territory
Queensland
Chesterfield
Norfolk I.
Lord Howe
Tasman Sea