

Parameter Estimation using a Windows Cluster



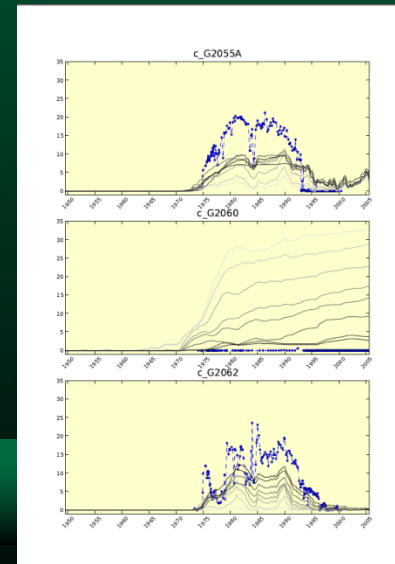
U.S. Department of the Interior
U.S. Geological Survey

Christian Langevin - Hydrologist
Edgar Hernandez – IT Specialist

Fort Lauderdale, FL

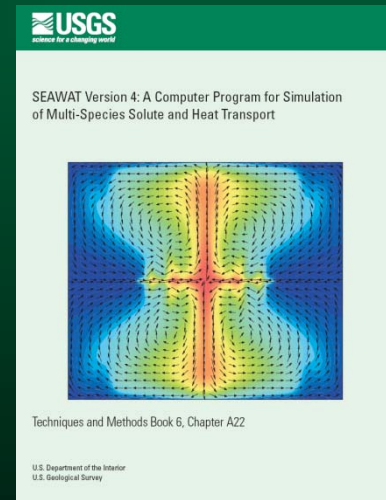
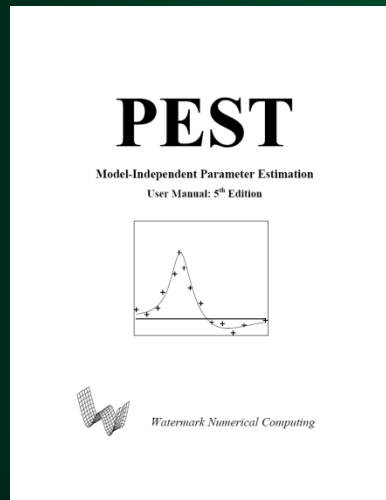
Parameter Estimation

- Automated procedure for adjusting model input parameters until simulated values match with observed values
- Requires numerous runs of the forward model



Software (32 or 64 bit)

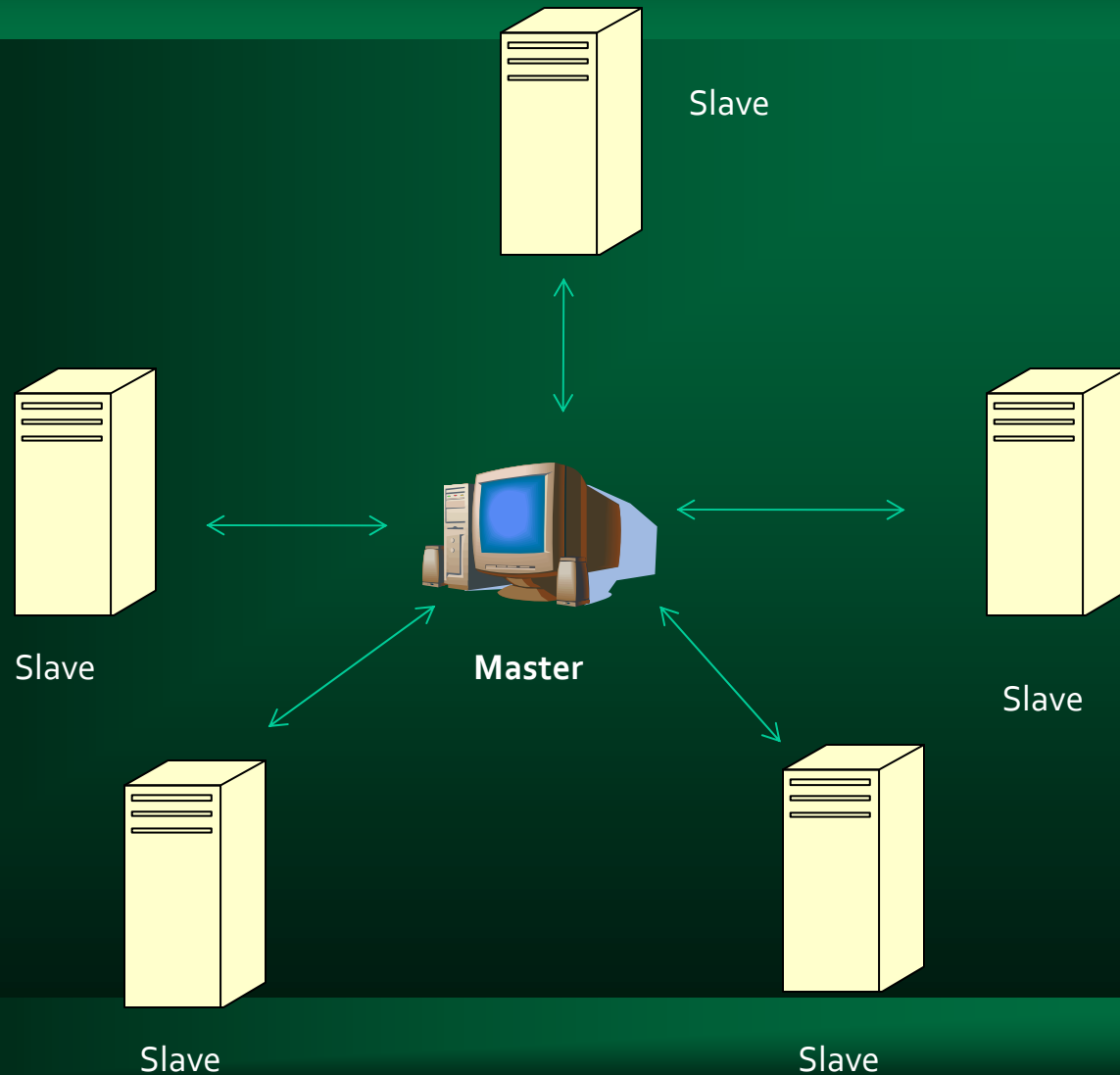
- PEST
- MODFLOW/MT3DMS/SEAWAT



Parallel PEST

- Uses “master and slave” approach with signal files to manage forward model runs
- Configurations
 - Single multi-core PC
 - Networked PCs
 - Cluster

Networked PC Approach



Windows Cluster Approach

- Advantages
 - Price per node
 - No need for administrative privileges
 - Multiple users
 - Familiar environment
- Disadvantages
 - Upfront capital investment
 - More complicated



Job Scheduler

Job Management

All Jobs

Configuring

Active

Finished

Failed

Canceled

My Jobs

Configuring

Active

Finished

Failed

Canceled

Job Templates

Default

Test Template

Clusrun Commands

Temporary View

All Jobs (24)

Filter:

Job name

Owner

Submit time

Project name

Job ID	Job Name	State	Owner	Priority	Submit Time	Requested Reso
180	ND\WTP jking...	Finished	GS\jking	Normal	11/17/2009 9:08:34 AM	224-224 Cores
179	ND\WTP jking...	Canceled	GS\jking	Normal	11/16/2009 10:24:07 AM	224-224 Cores
178	ND\WTP jking...	Canceled	GS\jking	Normal	11/14/2009 3:07:08 PM	224-224 Cores
177	ND\WTP jking...	Canceled	GS\jking	Normal	11/13/2009 9:28:11 AM	224-224 Cores
176	ND\WTP jking...	Canceled	GS\jking	Normal	11/12/2009 5:16:46 PM	224-224 Cores
99	test	Failed	GS\jking	Normal	9/3/2009 11:38:40 AM	Auto-Auto Cores
98	test	Failed	GS\jking	Normal	9/3/2009 11:38:24 AM	Auto-Auto Cores
97	test	Failed	GS\jking	Normal	9/3/2009 11:36:24 AM	Auto-Auto Cores
96	test	Failed	GS\jking	Normal	9/3/2009 11:34:48 AM	Auto-Auto Cores
95	test	Failed	GS\jking	Normal	9/3/2009 11:34:02 AM	Auto-Auto Cores
92	test	Failed	GS\jking	Normal	9/3/2009 11:32:40 AM	Auto-Auto Cores
91	test6	Failed	GS\jking	Normal	9/3/2009 11:31:57 AM	Auto-Auto Cores
90	startCN.8	Failed	GS\jking	Normal	9/3/2009 11:29:11 AM	Auto-Auto Cores
89	startCN.7	Failed	GS\jking	Normal	9/3/2009 11:27:17 AM	Auto-Auto Cores
88	startCN.6	Failed	GS\jking	Normal	9/3/2009 11:26:04 AM	Auto-Auto Cores
87	startCN.5	Failed	GS\jking	Normal	9/3/2009 11:25:04 AM	Auto-Auto Cores

Actions

Pivot To

Nodes for the Jobs

Job Submission

New Job ...

New Single-Task Job ...

New Parametric Sweep Job ...

Create New Job from Description ...

Job Actions

View Job ...

Modify Job ...

Add Task to Job ...

Copy Job ...

Submit Job

Submitting a Job

The image displays three sequential screenshots of a 'New Job' dialog box, illustrating the steps to submit a job.

Job Details Tab: This tab contains fields for 'Job name', 'Job template' (set to 'Default'), 'Project', and 'Priority' (set to 'Normal'). It also includes 'Job run options' with checkboxes for 'Do not run this job for more than:', 'Run job until cancelled or run time expires', and 'Fail the job if any task in the job fails'. The 'Run time' is set to 0 days, 0 hours, and 1 minute. Under 'Job resources', the user selects 'Core' as the resource type. The 'Minimum' and 'Maximum' resource types are set to 'Auto-calculate'. There is a checkbox for 'Use assigned resources exclusively for this job'.

Task List Tab: This tab is for entering tasks. It features a table with columns 'Task Name', 'Command Line', and 'Requested Resources'. Buttons for 'Add', 'Edit...', 'Copy', 'Delete', 'Export to File...', and 'Dependency...' are available. Below the table is a 'Task Properties' section with icons for adding, deleting, and saving tasks.

Resource Selection Tab: This tab allows selecting resources. It includes a 'Node preferences' section with checkboxes for 'Run this job only on nodes that are members of all the following groups:' and 'Run this job only on nodes in the following list:'. The 'Available node groups' list includes 'ComputeNodes', 'BatchNodes', and 'Top Nodes'. The 'Selected node groups' list is empty. Below this is a table of node details:

Node Name	Cores	Memory	State
IGSBAMESCN001	8	16379	Online
IGSBAMESCN002	8	16379	Online
IGSBAMESCN003	8	16379	Online
IGSBAMESCN004	8	16379	Online
IGSBAMESCN005	8	16379	Online
IGSBAMESCN006	8	16379	Online
IGSBAMESCN007	8	16379	Online
IGSBAMESCN008	8	16379	Online
IGSBAMESCN009	8	16379	Online

Hardware preferences are also specified, including 'Minimum memory (MB)' and 'Minimum cores', both set to 0. There is a checkbox for 'Prefer nodes with:' and a dropdown menu for 'More Memory'.

In our case, a job starts a batch file on every compute node

Job XML File to Start Slaves

[illegible]

Parallel Computing Options

First



Second



\$750 / core

Present

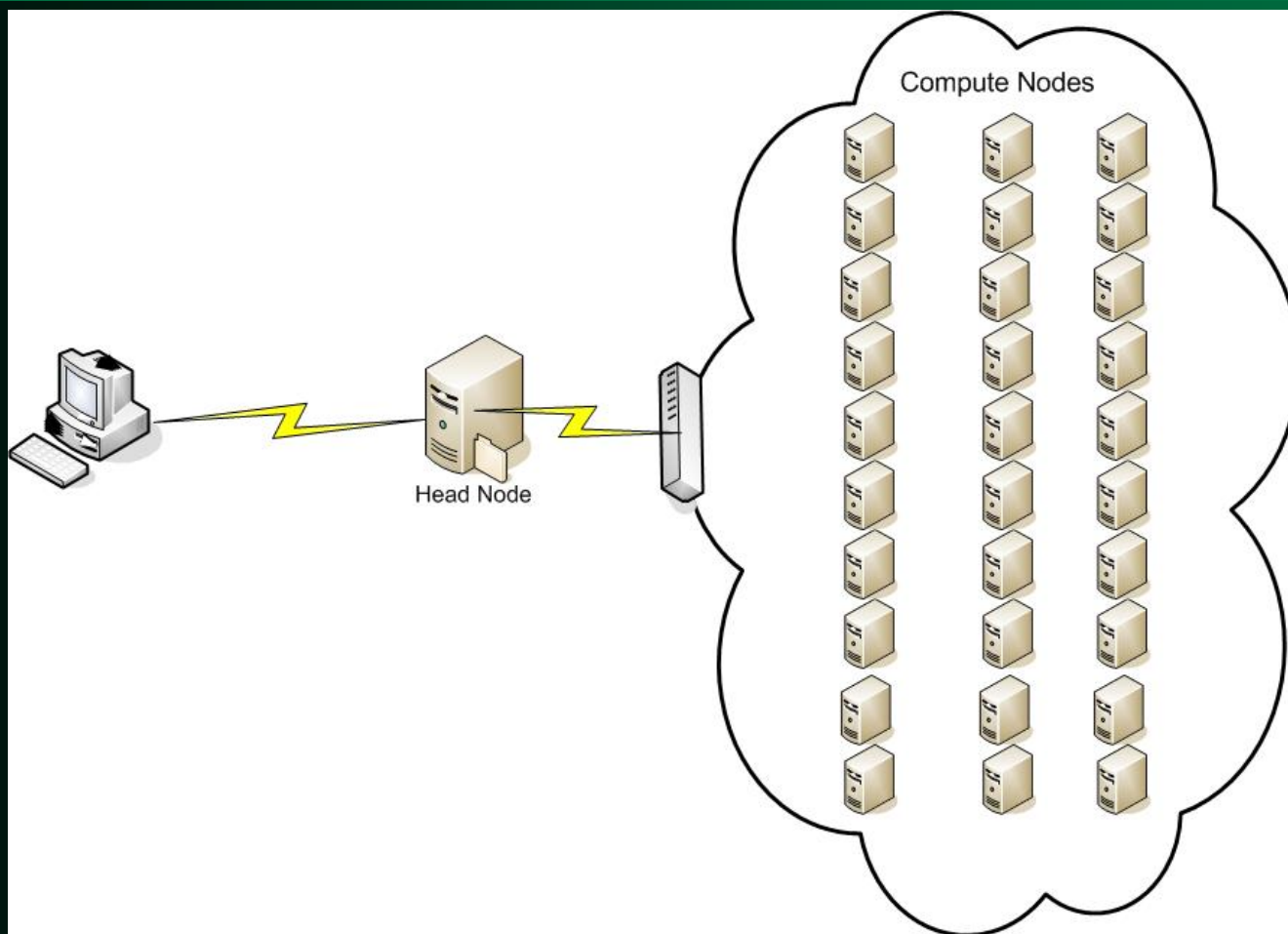


\$500 / core₁₀

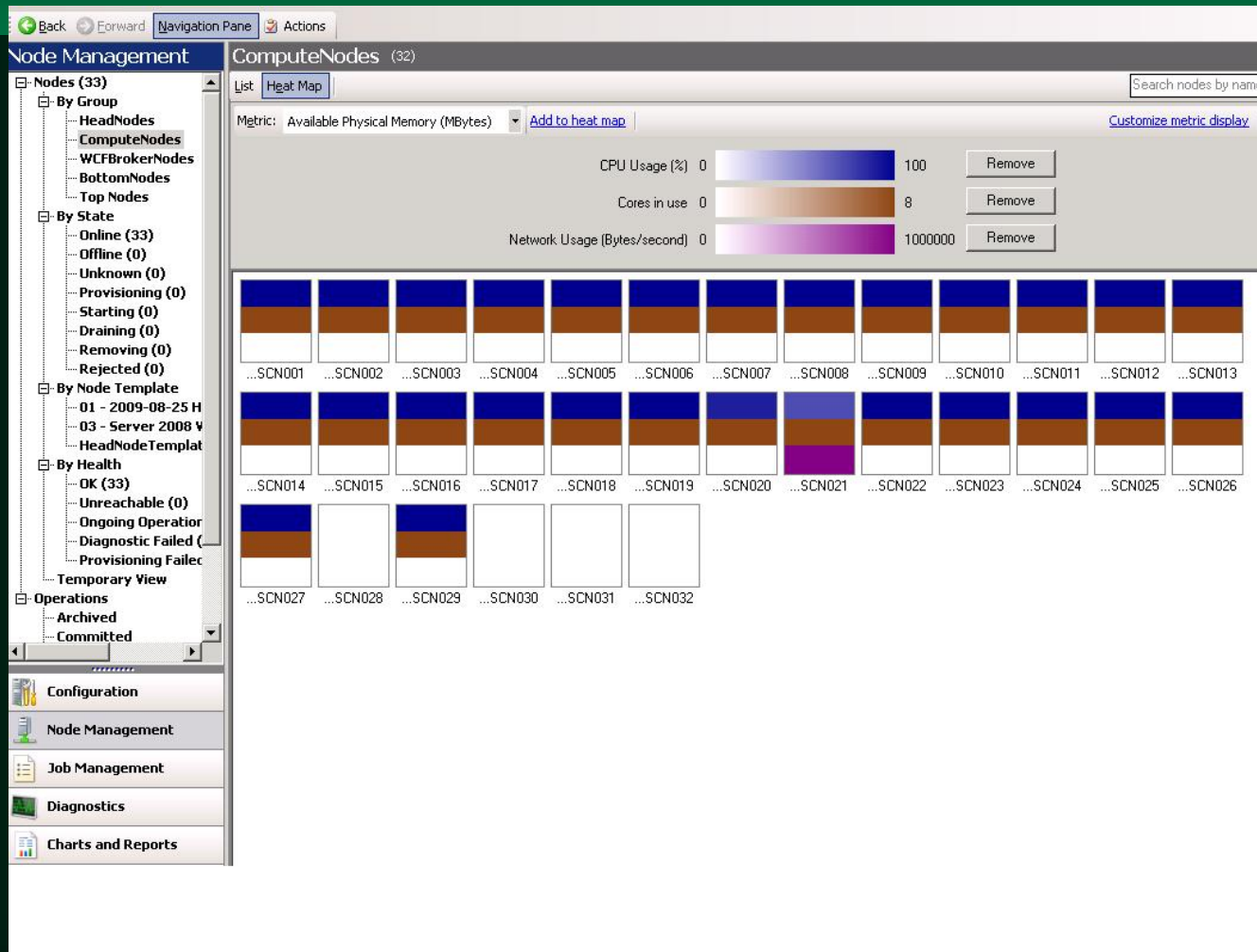
Cluster Specifications

- GS Domain
- Head node
 - 64 bit
 - 6 TerraByte RAID5
 - 8 cores
 - Windows Server HPC 2008
 - 16 GB RAM
- 34 compute nodes
 - 64 bit
 - 272 total cores (8 each)
 - 500 GB hard drives each
 - 16 GB RAM each
- Automatic node deployment
- Automatic software updates
- Required significant power and A/C upgrades in the computer room

Windows Server HPC 2008



Heat Map



Summary

- Transition from single PCs to Windows Cluster environment
- What used to take 83 days now takes 7.5 hours on the cluster
- Time and capital investment was required for success

Questions?

