



Development of a New MARKAL Post-Processor

Joe DeCarolis, Kevin Hunter
North Carolina State University

MARKAL/TIMES Symposium
Research Triangle Park, NC
September 28-29, 2009

jdecarolis@ncsu.edu

Motivation

Energy future fraught with uncertainty; results should reflect that reality

Methods to address uncertainty:

- sensitivity analysis

- uncertainty propagation (Monte Carlo simulation)

- stochastic optimization (TIMES)

- Modeling-to-Generate Alternatives (MGA)

Options can also be used serially:

Sensitivity analysis → stochastic optimization → MGA

Running in a cluster environment provides capability for rigorous analysis

Cluster details

NC STATE UNIVERSITY

11 nodes, each with 2 AMD quad-core Opteron processors
(2.0 GHz with 512 KB Cache/core)

1.8 TB of storage

176 GB memory

OpenSuse 10.3 (Linux)

FLOPS = 704 Gigaflops

(1.1 Petaflops (LANL);

70 GFLOPS (PC))

1 Gbit interconnect



MARKAL Default Workflow

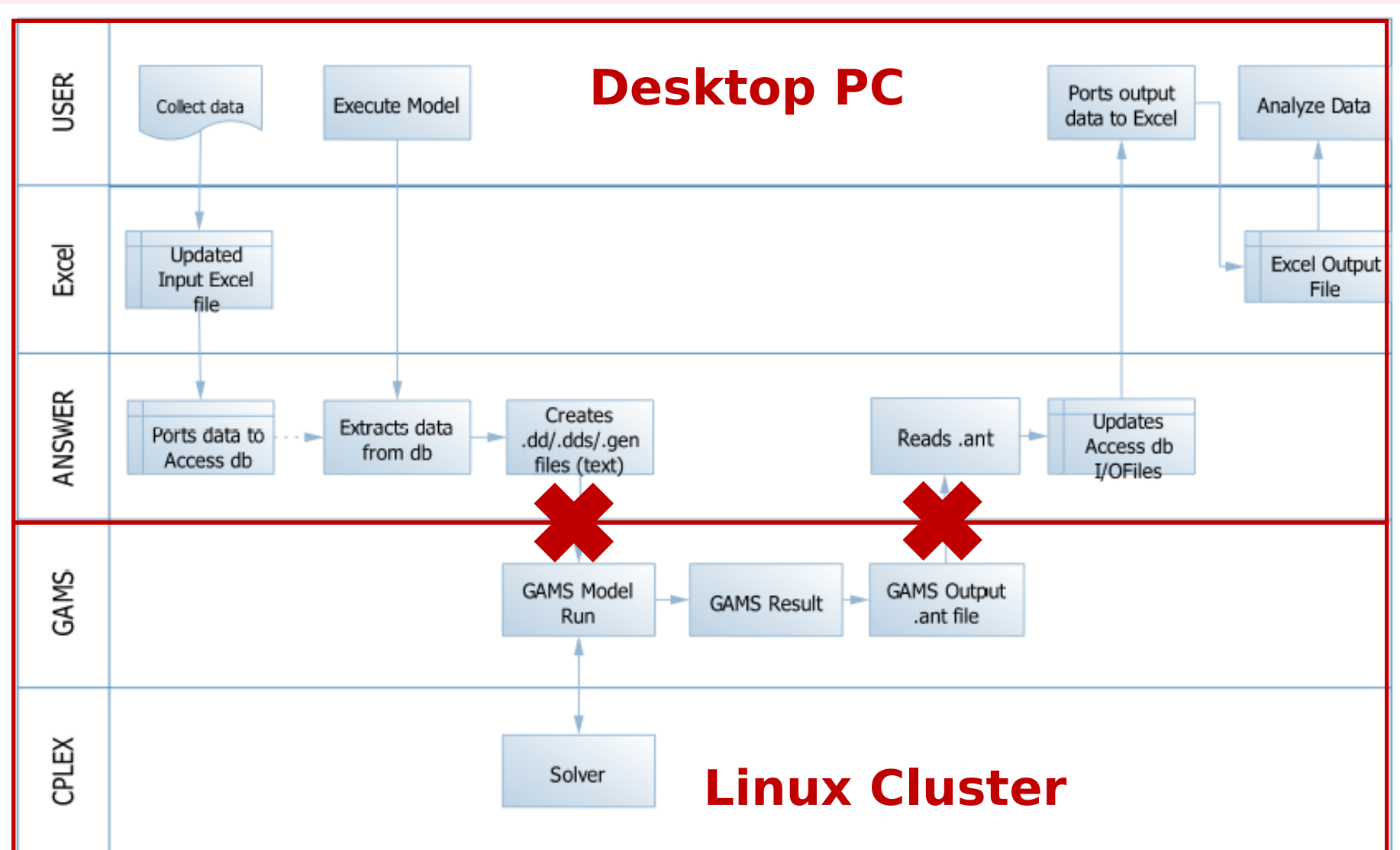


Diagram provided courtesy of Rajat Kapoor

Model Iteration

- There are 3 basic approaches to iterating the model:
1. Call high level script that invokes GAMS and builds the model.
 2. Fork model; use GAMS “grid computing” facility to manage model instances
 3. Fork model; manage model instances with a user-defined script

Our problem is “embarrassingly parallel”

All options involve using ANSWER or VEDA to create an initial set of GEN/DD/DDS files.

GAMS Performance

NC STATE UNIVERSITY

Use PROFILE feature of GAMS

e.g., “-PROFILE 2” to end of GAMS call to GEN file
(Ans_gams.cmd)

Test of U.S. EPA NMD08:

```
--- Generating LP model MRK
--- Starting Equ EQ_PRICE: elapsed 0:00:21.719
--- Starting Equ EQ_ADRAT1: elapsed 0:00:21.887
--- Starting Equ EQ_ADRAT2: elapsed 0:00:42.137
...
--- Starting Equ EQ_SRM: elapsed 0:00:52.782
...
--- Job TST.GEN Stop 09/24/09 21:31:42 elapsed 0:01:18.719
```

Managing Output Data

Suppose we perform sensitivity analysis and generate 1000 ANT files...

We don't want to move back to local machine; need a post-processor on the cluster to efficiently handle output data

Also, an opportunity to streamline the workflow associated with MARKAL data processing.

Database Motivation

What do we want to do on the back end?

Extract data from 100s-1000s of ANT files, organize, and store.

Want to be able to look across arbitrary sets of runs and generate statistics, tables, and graphics on-the-fly

Archive past runs that can be easily accessed and compared to new results

Sidestep: What is a DB?

- A database (DB) is
 - more than just a datastore
 - structured relationships among data
 - SQL interaction – Structured Query Language
 - SELECT carbon_intensity FROM data WHERE ...
 - ad-hoc, on-the-fly snapshots of data
 - Example: SELECT query

What we (currently) do

(current) MARKAL interaction paradigm:

Excel → Answer/VEDA (+DB) → MARKAL → Answer/
VEDA (+DB) → Excel → Analysis →
Presentation/Paper

Major drawbacks:

- lots of human interaction
- error-prone, very inefficient

What we (should) do

(our) MARKAL interaction paradigm:

Excel → Answer/VEDA (+DB) → Copy to Linux → MARKAL → Postprocessor → Database → Report script → Presentation

Drawbacks:

- New workflow (have to learn it)
- New, maybe incorrect, debugging needed
- Need a larger calculator?

Benefits: less human interactivity (errors, time)

Workflow Interaction

Current Workflow

1. Excel →
2. Answer/VEDA (+DB) →
3. MARKAL →
4. Answer/VEDA (+DB) →
5. Excel →
6. Analysis →
7. Presentation/Paper

Our Workflow

1. Excel →
2. Answer/VEDA (+DB) →
3. Copy to Linux →
4. MARKAL →
5. Postprocessor (+DB) →
6. Report script →
7. Presentation/Paper

Workflow Interaction

Current Workflow

1. Excel →
2. Answer/VEDA (+DB) →
3. MARKAL →
4. Answer/VEDA (+DB) →
5. Excel →
6. Analysis →
7. Presentation/Paper

Our Workflow

1. Excel →
2. Answer/VEDA (+DB) →
- 3. Copy to Linux →**
- 4. MARKAL →**
- 5. Postprocessor (+DB) →**
- 6. Report script →**
7. Presentation/Paper

How did we do it?

- An amalgam of tools:
 - PostgreSQL – Petascale capable, statistics base
 - Python – Script “glue” to make it all work
 - Gnuplot – Graphic utility
 - SSH – Communication from afar
 - Linux – The base, makes it possible

Questions and Comments?

NC STATE UNIVERSITY

”[P]eople seem to think that performance is about doing the same thing, just doing it faster, and that is not true. ... If you can do something really fast, really well, people will start using it differently.” – *Linus Torvalds*
on GIT