Rapid Application Development & Grid Computing Using GAMS

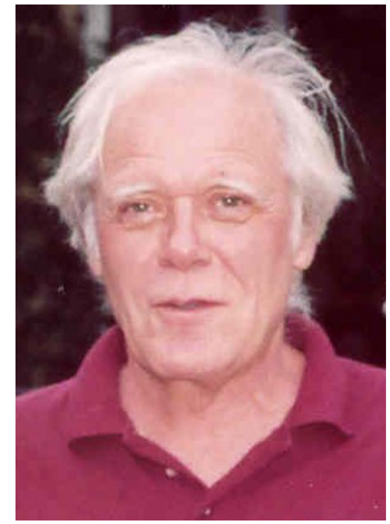
Software Demonstration

INFORMS San Francisco 2005

Demo Objectives

- Basics / Algebraic Modeling
- Data Exchange / Charting
- Tying things together (procedural elements)
- Grid Computing
- Model Deployment

Company Profile



- **G**eneral **A**lgebraic **M**odeling **S**ystem
- Started as a **research project** at the World Bank 1976
- Went **commercial** in 1987
- **Offices** in Washington, D.C and Cologne
- Professional **software tool provider**
- Operating in a **segmented niche market**
- Broad **academic** and **commercial** user base

Modeling Systems

- Describe problems to a computer system in the same way that people describe problems to each other.
- Simplify the model building and solution process
- Create maintainable models
- Adapt models quickly to new situations

Basic Technical Principles

- Separation of model and data
- Separation of model and solution methods
- Computing platform independence
- Multiple model types, solvers, platforms
- Balanced mix of declarative and procedural approaches
- Model is a data base operator and/or object

Basic Technical Principles II

- Open architecture and interfaces to other systems:
  - GUI
  - Excel, Databases
  - Programming Languages etc.
- Maintainable models and protection of investments

Multiple Model Types

- LP - Linear Programming
- MIP - Mixed Integer Programming
- QCP - Quadratically Constrained Programming
- NLP - Nonlinear Programming
- CNS – Constrained Nonlinear Systems
- MINLP - Mixed Integer Nonlinear Programming
- MPEC - NLP with Complementarity Constraints
- MPSGE - General Equilibrium Models
- Stochastic Optimization

Multiple Solvers & Platforms

Solver/Platform availability - 22.0 August 1, 2005										
	Intel MS Windows	x86_64 MS Windows	Intel Linux	x86_64 Linux	Sun Sparc SOLARIS	HP 9000 HP-UX 11	DEC Alpha Digital Unix 4.0	IBM RS-6000 AIX 4.3	SGI IRIX	Mac PowerPC Darwin
BARON 7.4	✔	32bit	✔	32bit				✔		
BDMLP	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
COIN	✔	✔	✔	✔						✔
CONOPT 3	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
CPLEX 9.1	✔	✔	✔	✔	✔	✔	8.1	✔	✔	
DECIS	✔	✔	✔	✔	✔	✔	✔	✔	✔	
DICOPT	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
KNITRO 4.0	✔	32bit	✔	✔						
LGO	✔	✔	✔	✔	✔	✔	✔		✔	✔
MILES	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
MINOS	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
MOSEK 3.2	✔	✔	✔	✔	✔	✔				✔
MPSGE	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
MSNLP	✔	✔	✔	✔	✔	✔			✔	✔
OQNLP	✔	32bit	✔	32bit						
OSL V3	✔	32bit	✔	32bit	✔	V2		✔	V2	
OSLSE	✔	32bit	✔	32bit	✔			✔		
PATH	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
SBB	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
SNOPT	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
XA	✔	32bit	✔	✔	✔	✔	✔	✔		
XPRESS 15.30	✔	32bit	✔	32bit	✔	15.20		15.20		

Basic Portfolio Selection

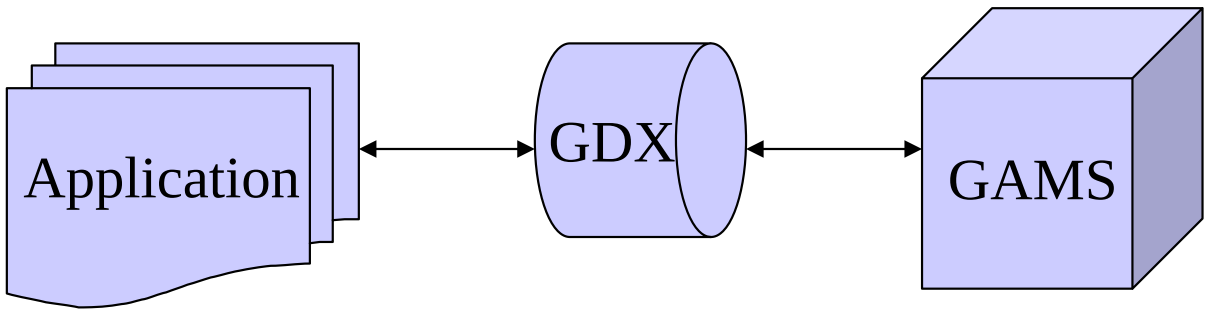
- Markowitz
- Closing Values of Stocks over Time
- Calculate Return, Deviation, Co-variance using Closing Values
- Select Portfolio to
  - Minimize Variance
  - s.t. Some Return Goal
  - Minimum Portfolio Size per Stock

- Basic Portfolio Selection Model QP1

- Data from Excel QP2

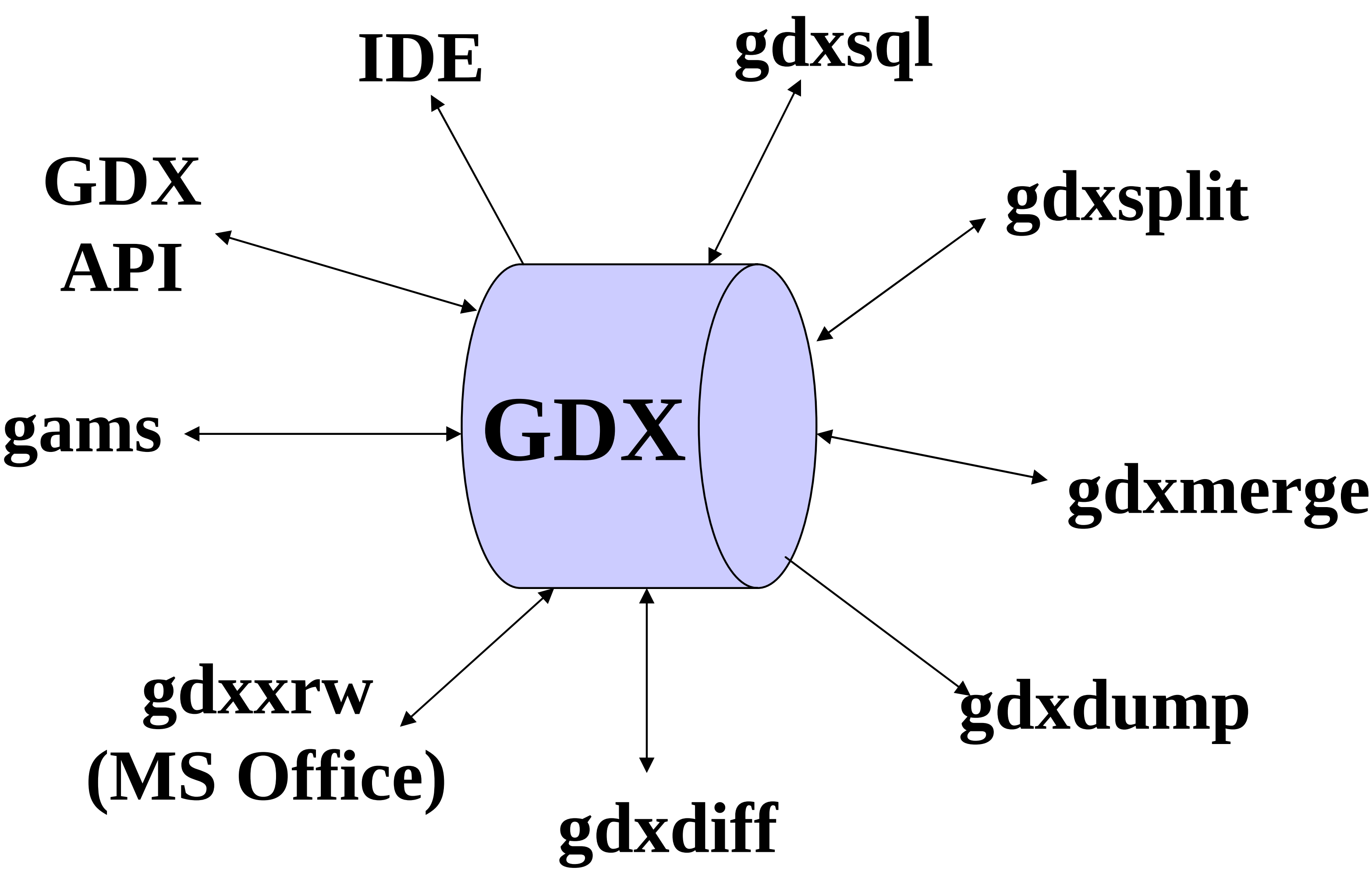
Gams Data eXchange

- Separation of Responsibility for Data and Model (Data contract)
- Gams Data eXchange (GDX):



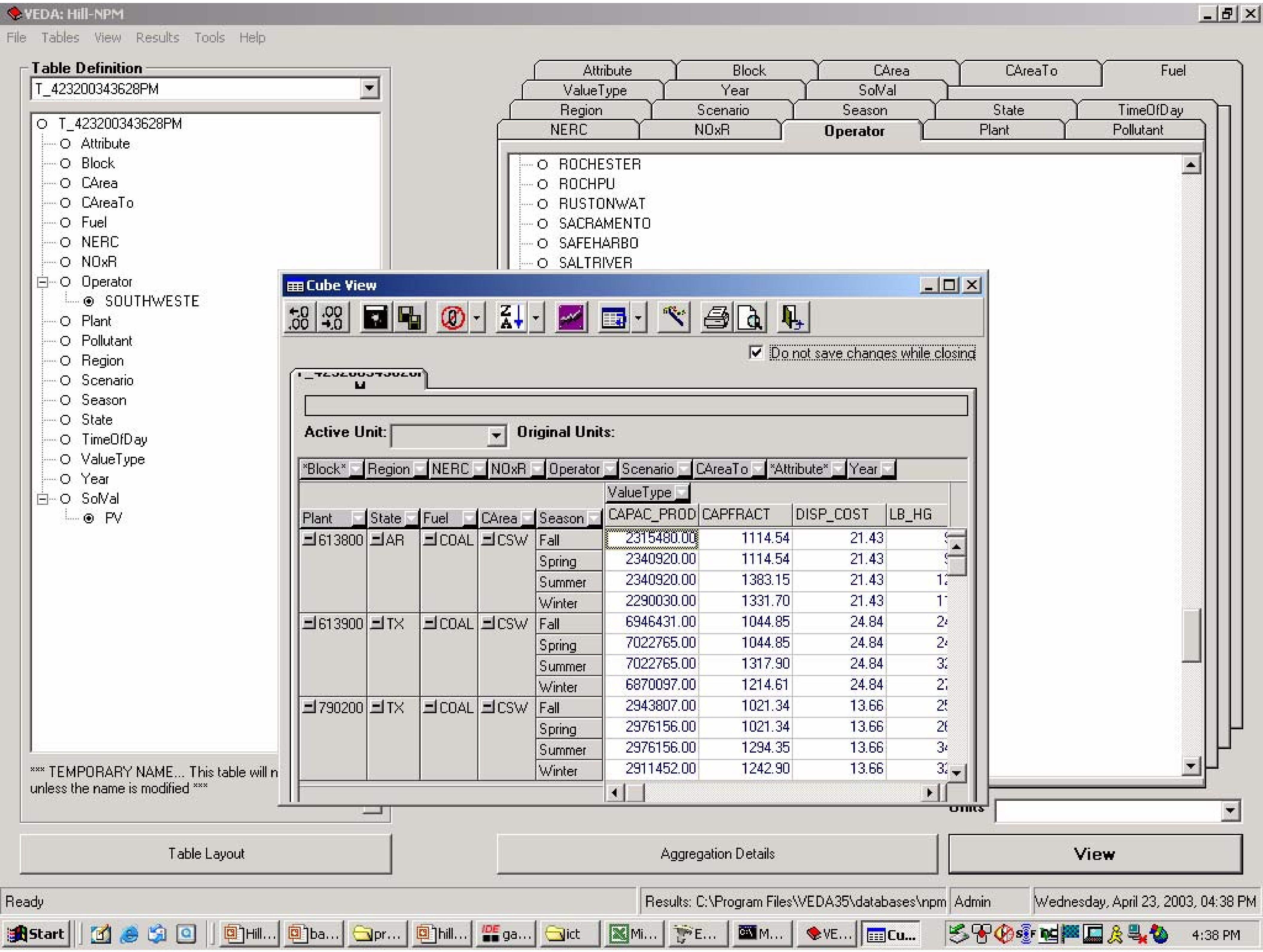
- A priori validation of data contract
- Complements the ASCII text data input

GDX Tools


IDE
gdxsql
GDX
API
gdxsplit
gams
GDX
gdxmerge
gdxxrw
(MS Office)
gdxdump
gdxdiff

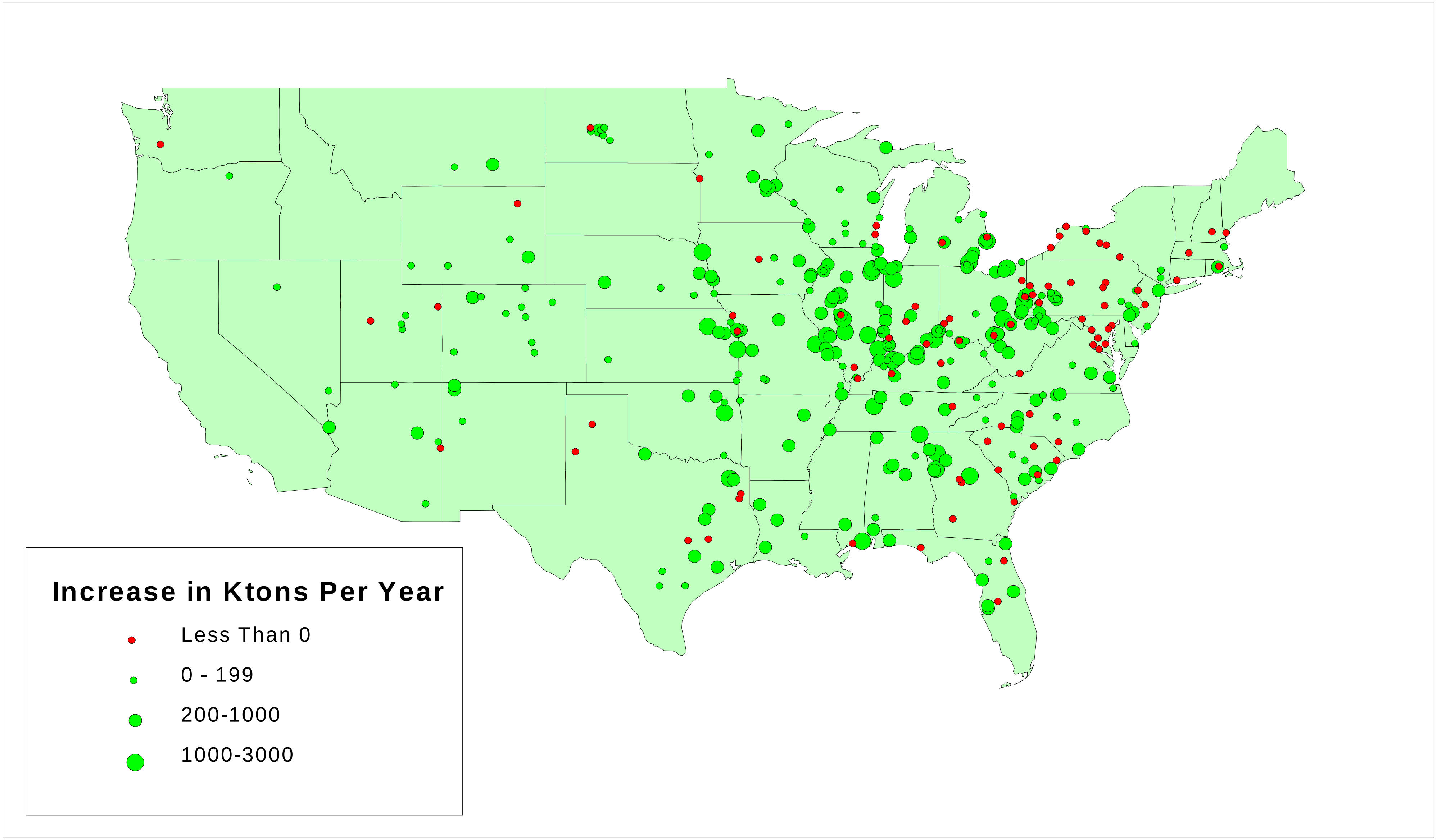
Reporting

- Integrated Standard Reporting Facilities
  - text files
  - GDX: MS Office, DB, Charting, …
- Specialized Reporting
  - Scenario Management/Data Cube: VEDA
  - Geographical Information System: MapInfo
  - Visualization with MATLAB

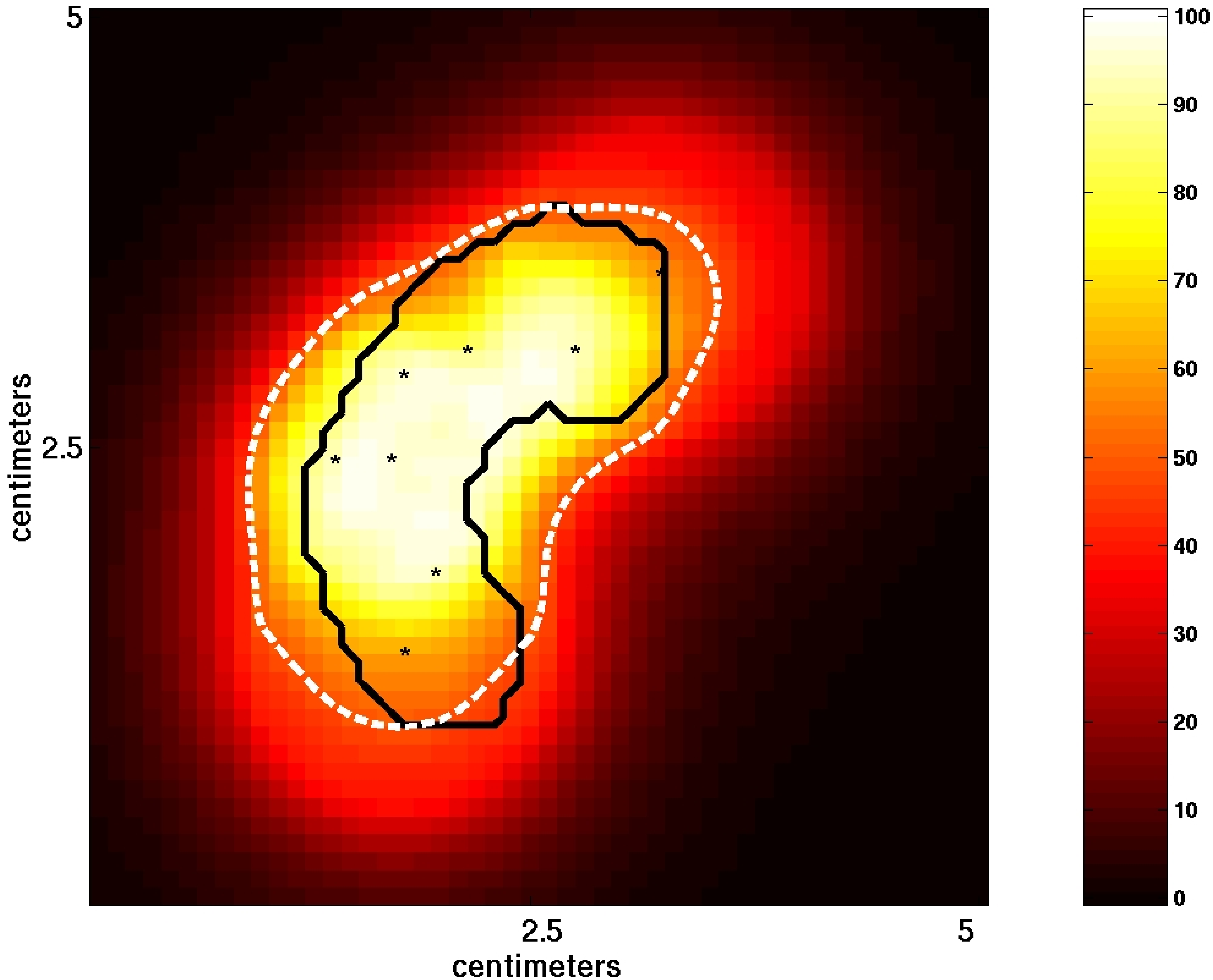


VEDA: Hill-NPM
File Tables View Results Tools Help
Table Definition
T_423200343628PM
T_423200343628PM
Attribute
Block
CArea
CAreaTo
Fuel
NERC
NOxR
Operator
SOUTHWESTE
Plant
Pollutant
Region
Scenario
Season
State
TimeOfDay
ValueType
Year
SolVal
PV
*** TEMPORARY NAME... This table will n
unless the name is modified ***
Attribute
Block
CArea
CAreaTo
Fuel
ValueType
Year
SolVal
Region
Scenario
Season
State
TimeOfDay
NERC
NOxR
Operator
Plant
Pollutant
ROCHESTER
ROCHPU
RUSTONWAT
SACRAMENTO
SAFEHARBO
SALTRIVER
Cube View
Do not save changes while closing
Active Unit:
Original Units:
Block Region NERC NOxR Operator Scenario CAreaTo *Attribute* Year
ValueType
Plant State Fuel CArea Season CAPAC_PROD CAPFRACT DISP_COST LB_HG
613800 AR COAL CSW Fall 2315480.00 1114.54 21.43
Spring 2340920.00 1114.54 21.43
Summer 2340920.00 1383.15 21.43
Winter 2290030.00 1331.70 21.43
613900 TX COAL CSW Fall 6946431.00 1044.85 24.84
Spring 7022765.00 1044.85 24.84
Summer 7022765.00 1317.90 24.84
Winter 6870097.00 1214.61 24.84
790200 TX COAL CSW Fall 2943807.00 1021.34 13.66
Spring 2976156.00 1021.34 13.66
Summer 2976156.00 1294.35 13.66
Winter 2911452.00 1242.90 13.66
Units
Table Layout
Aggregation Details
View
Ready
Results: C:\Program Files\VEDA35\databases\npm
Admin
Wednesday, April 23, 2003, 04:38 PM
Start
4:38 PM

GAMS/MapInfo



GAMS/MATLAB



5
2.5
centimeters
2.5
5
centimeters
100
90
80
70
60
50
40
30
20
10
0

Procedural Elements

- Contrast pure declarative approach:
  - e.g. AMPL, MPL, OPL
- Scripting added outside language
  - AMPL Script, OPL script
  - Programming Languages API (OptiMax)
- Experience shows importance of procedural paradigm
- Part of the GAMS Language

- Procedural Elements QP3

What is Grid Computing?

- A pool of connected computers managed and available as a common computing resource
  - Allows parallel task execution
  - Allows effective sharing of CPU power
  - For us, not necessarily distributed
  - Scheduler handles management tasks
  - Can be rented or owned in common
  - E.g. Condor, Sun Grid Engine, Globus

GAMS

Economics of Grid Computing

- Yearly cost, 2-CPU workstation: $5200
  - Hardware - $1200
  - Software - $4000
- Hourly cost on the grid: $2
  - $1/hour for CPU time (to grid operator)
  - $1/hour for software (GAMS, model owner)
- 1 workstation == 50 hrs/week grid time
- Up-front vs. deferred, as-needed costs

Energy planning problems

- Given expected energy demands, costs, and supplies, one can quickly form an optimal plan.
- Uncertain data ➔ many possible scenarios
- Planning time increases dramatically
  - 30 sec/scen * 4K scen = 33 hours
- Limits the number of scenarios allowed

Solution A: Optimize for Speed

- Modify model to batch scenarios
- Eliminate generation time
  - GAMS generates each model from scratch
  - Fraction of total time used is non-trivial
  - Requires significant programming
  - Increased chance of error
  - Not maintainable/flexible/layered
- Speedup: 1.5 (at most!), or 22 hours

Solution B: Use a Grid

- Solve the scenarios in parallel, e.g.
  - 100 CPUs: 20 minutes
- Marginal cost is $66
- No programming required
- Model stays maintainable
- Separation of model and solution maintained

Grid Specifics

```
echo "#!/bin/bash"               > ${3}runit.sh
echo $1 $2                      >> ${3}runit.sh
echo gmscr_ux.out $2            >> ${3}runit.sh
echo mkdir ${3}finished >> ${3}runit.sh
chmod 750 ${3}runit.sh
${3}runit.sh > /dev/null &
```

- Parallel Submission QP4

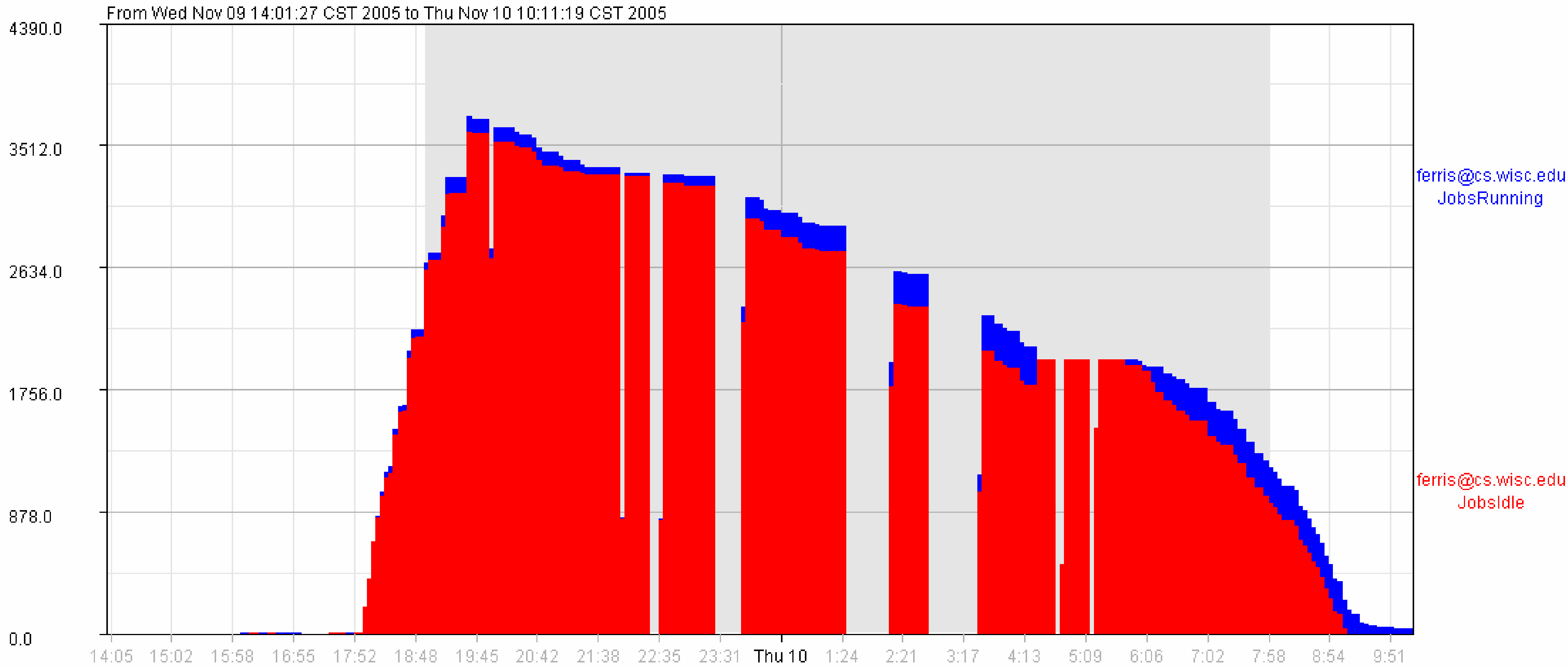
Massively Parallel MIP

- MIP/B&C Algorithm ideal to parallelize
  - Seymour problem solved that way
  - Software: FATCOP/Condor, BCP/PVM
  - A priori subdivision in *n* independent problems
- Open Pit Mining Problem (openpit in GAMS Model library)
  - Used integer variables to subdivide Model into into 4096 sub-problems
  - Experiments (Ferris) at UW using Condor Pool

Results

- Submission start Nov. 9 at 6 pm
- All job submitted by Nov. 9 at 7:30 pm
- All jobs returned by Nov 10, 10:15 am
  - 16 hours wall time, 1812 CPU hours
  - Peak number of CPUs: 270
- Different Instance:
  - 24 hours wall time, 3000 CPU hours

Condor Statistics



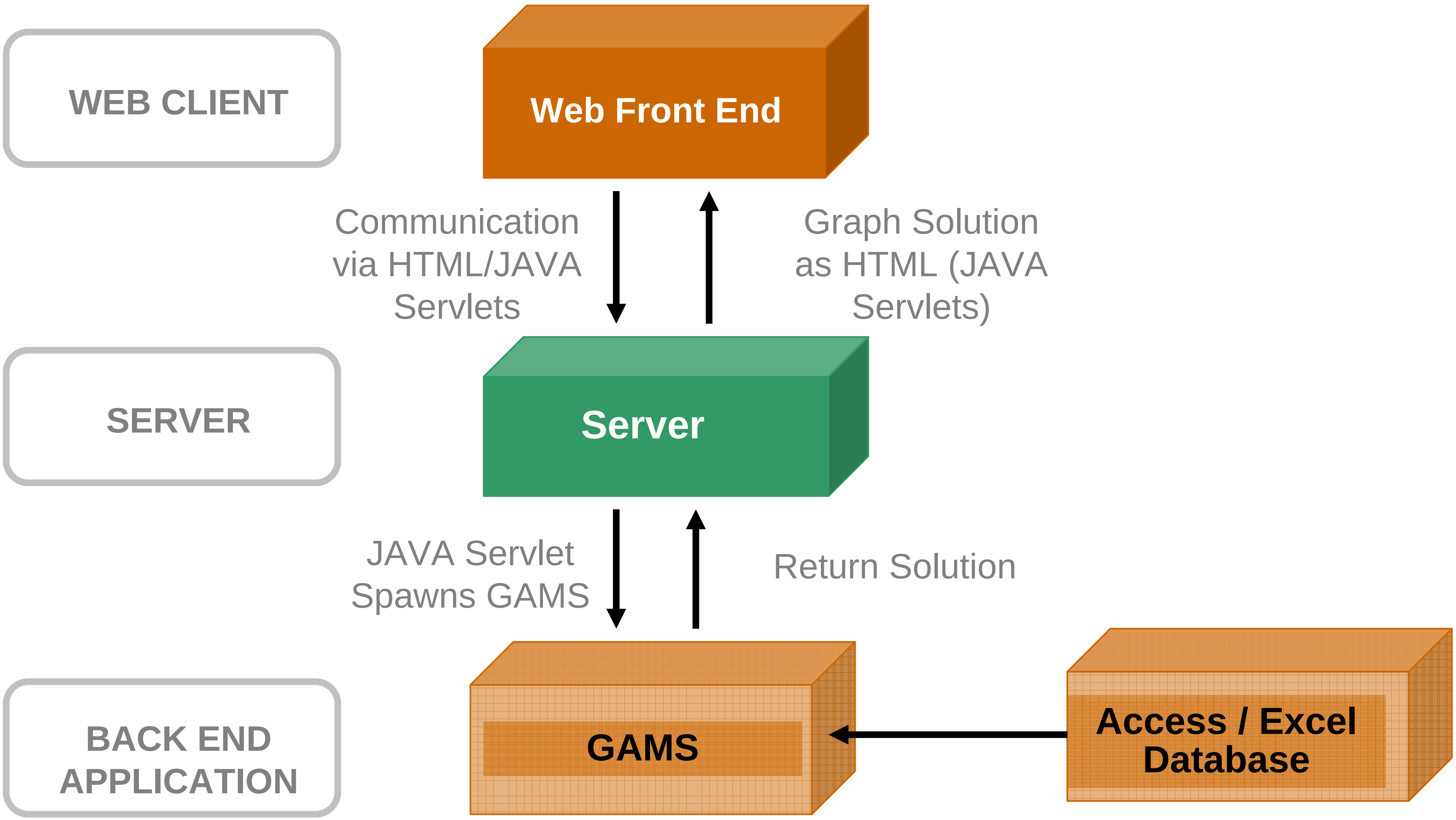
Other Examples

- Student-centric scheduling at USMA (West Point)
  - Models used intensively 4 times/year
  - 4000 independent MIP models
- Hill & Associates (energy/coal experts)
  - National Power Model (coal & electricity)
  - Multi-period (20 yr) model: runs 10 hours
  - Tens of runs & quick turnaround required

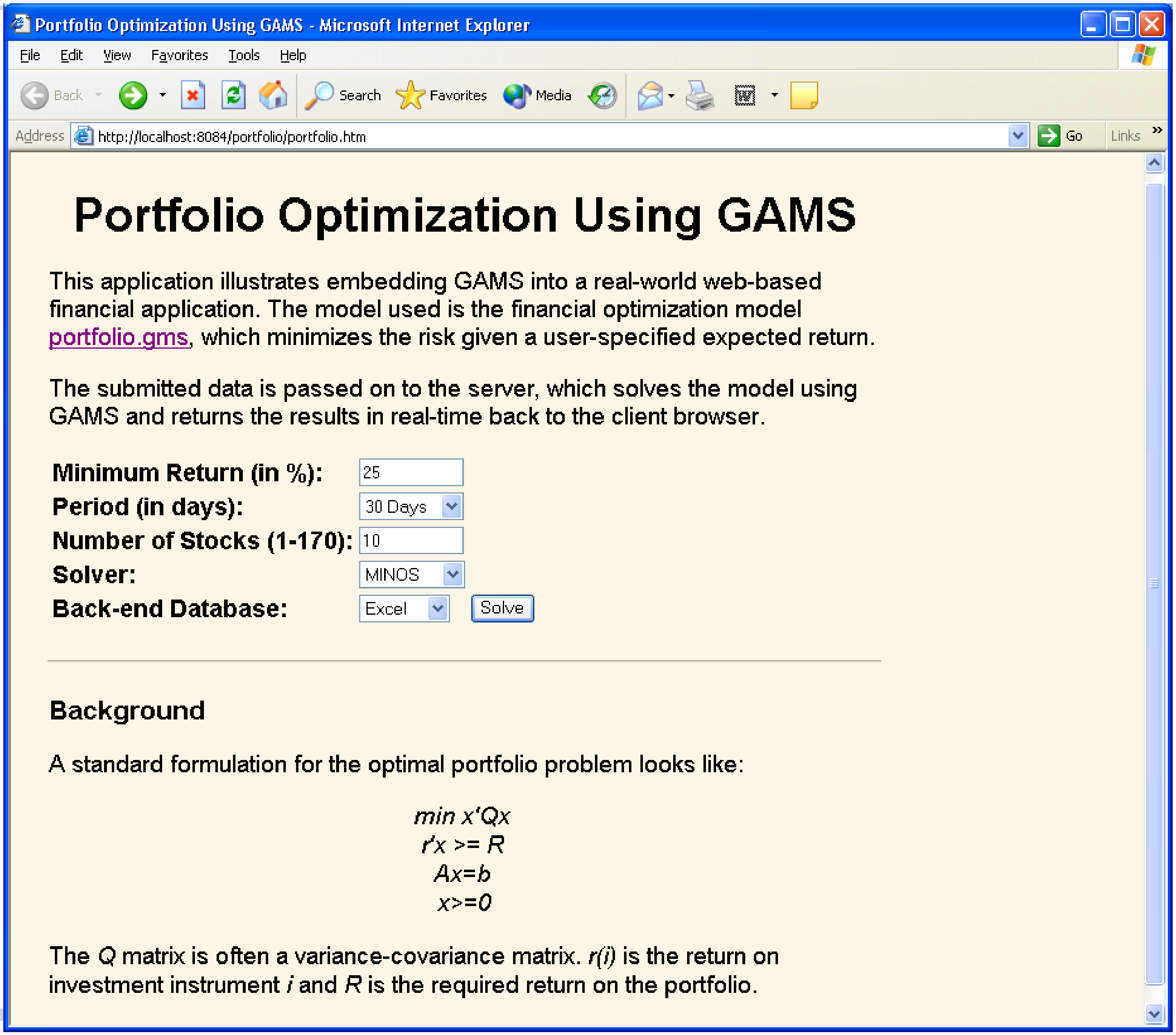
Model Deployment

- All options are available
  - Spreadsheets, databases
  - Web services
  - GUI (IDE, custom built)
  - On the Grid
- Independent of model or data sources
- Basic philosophy: layers of separation

Fully Embedded Application



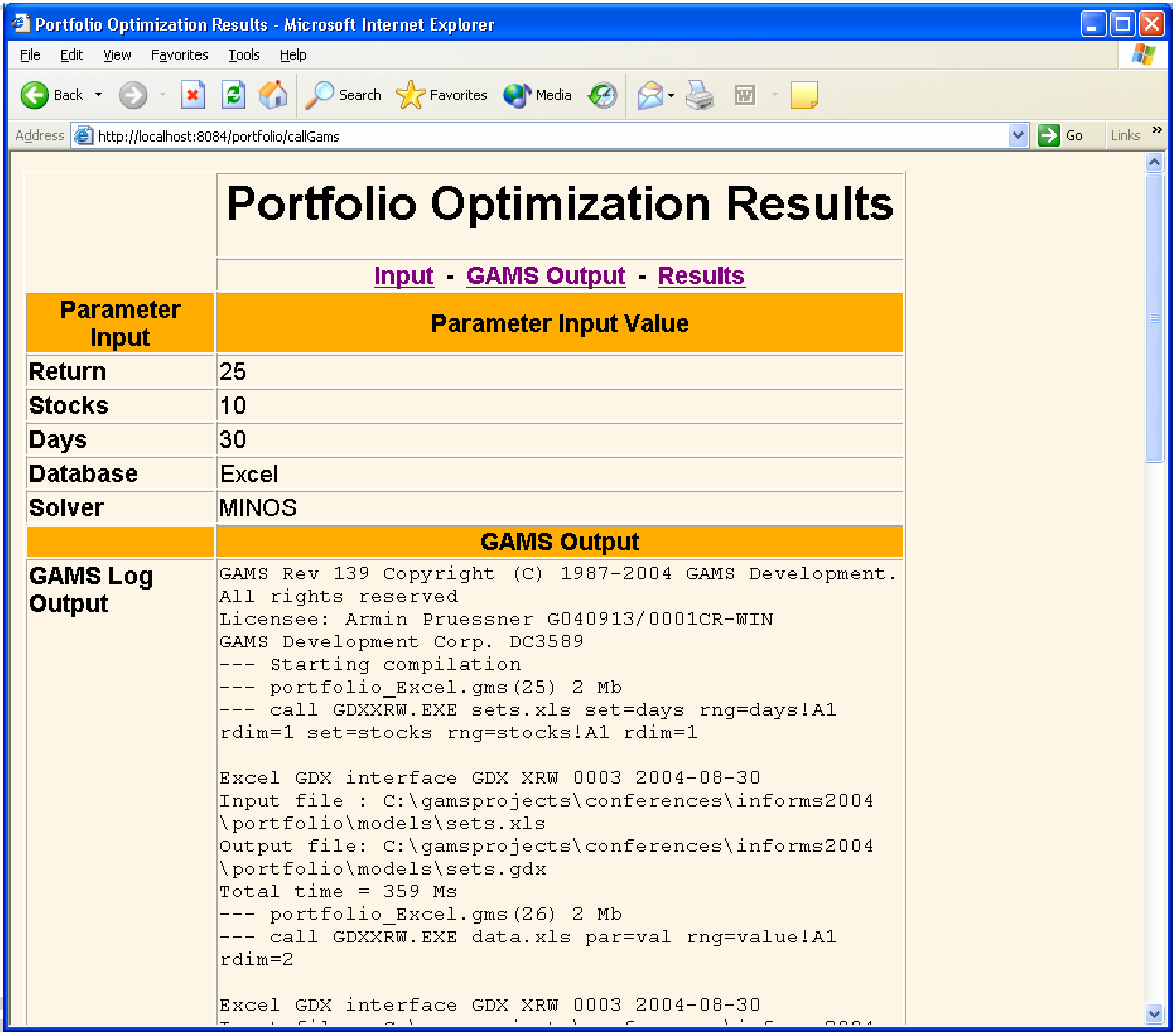
Web Front End



User may specify:

- Expected Return
- Investment Period (Days)
- Number of Stocks
- Solver
- Database (Excel or Access)

Output



Portfolio Optimization Results - Microsoft Internet Explorer

http://localhost:8084/portfolio/callGams

Portfolio Optimization Results

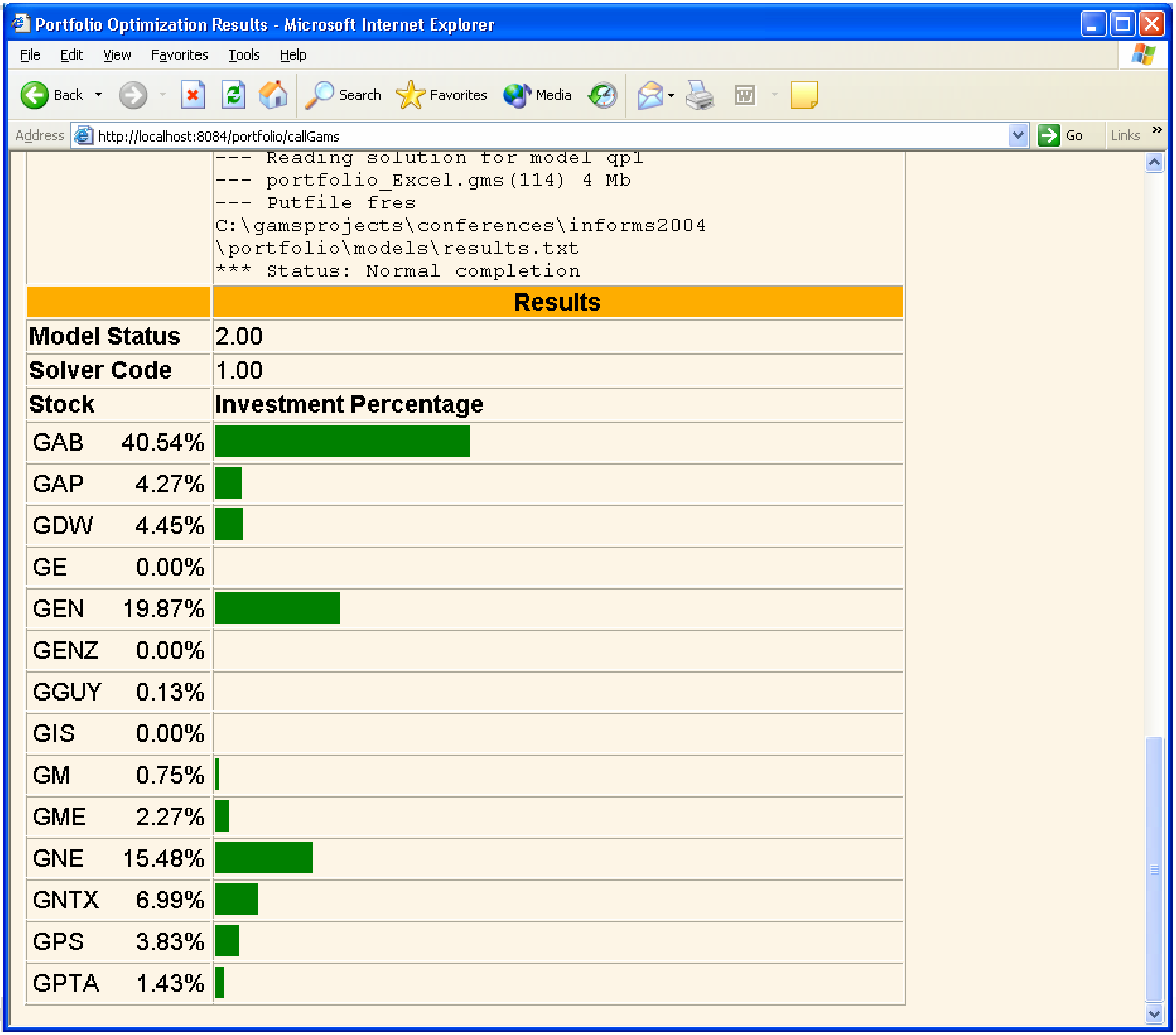
Input - GAMS Output - Results

Parameter Input	Parameter Input Value
Return	25
Stocks	10
Days	30
Database	Excel
Solver	MINOS
	GAMS Output
GAMS Log Output	GAMS Rev 139 Copyright (C) 1987-2004 GAMS Development. All rights reserved Licensee: Armin Pruessner G040913/0001CR-WIN GAMS Development Corp. DC3589 --- Starting compilation --- portfolio_Excel.gms(25) 2 Mb --- call GDXXRW.EXE sets.xls set=days rng=days!A1 rdim=1 set=stocks rng=stocks!A1 rdim=1 Excel GDX interface GDX XRW 0003 2004-08-30 Input file : C:\gamsprojects\conferences\informs2004 \portfolio\models\sets.xls Output file: C:\gamsprojects\conferences\informs2004 \portfolio\models\sets.gdx Total time = 359 Ms --- portfolio_Excel.gms(26) 2 Mb --- call GDXXRW.EXE data.xls par=val rng=value!A1 rdim=2 Excel GDX interface GDX XRW 0003 2004-08-30

Output:

- User input parameters
- Real-time GAMS output of model solve

Output (continued)



Output:

- GAMS Return Codes
- Investment distribution
- Graphical output