MOVES Architecture:
From the Ground Up

Air Quality and Modeling Center
Office of Transportation and Air Quality
U.S. Environmental Protection Agency
Outline

- Architecture details
- Calculators
- From Calculator to Generator and more
- Databases
- Importing data
- Designed for sharing the work
- Architecture revisit
One Calculator, Internals

Op Mode Distribution

Activity (SHO, Starts, etc)

Rates

Inventory

Output

\[ \sum \]
One Calculator

Activity

Rates → Calculator → Inventory

Adjustments
Many Calculators
Many Activities
Chaining

Activity

Calculator

Calculator

Inventory

Activity

Chained Calculator
Generators

Generator:
- Translate
- Allocate
- Model

Summary Activity

Detailed Activity → Calculator → Inventory
Internal Control Strategies

Generator → Detailed Activity → Calculator → Inventory

Summary Activity

Rates

Internal Control Strategy: AVFT, Rate of Progress
Databases

Generators → Execution → Calculators → Output

Internal Control Strategies
Databases, cont.

User → Consolidate → Aggregate → Execution → Update → Calculate → Output

Domain
Default
Data Priority

User

Domain

Default

Consolidate

Execution

1. Default
   A: 11
   B: 33
   C: 66

2. Domain
   A: 22
   B: 33

3. User
   A: 11
   B: 33
   C: 66

Consolidate
Importers

Domain -> Import -> Runspec -> Export Default

Default

XLS
XLSX
TXT
Importers, cont.
Design: How much work?

- 3,223 Counties
- 1 Zone per county
- 5 Road types
- 1 Calendar Year
- 12 Months
- 2 Day types
- 24 Hours
- 13 Source use types
- 4 Fuel types (pending)

- 31 Model Years
- 15 Operating Modes
- 25 Pollutants (pending)
- 6 Processes
- 2 I/M (with and without)
Design: How much math?

That’s:

67,333,368,960,000 combinations!

Each matched with location and time-specific:

- Activity
- Fuel formulations
- Meteorology
- Population

500 - 1,000 Trillion math operations required
Design: How much data?

- **24 bytes / output record MINIMUM**
  - No SCC
  - No indexes
  - Standard precision

- **1,469 Terabytes with no chained pollutants**

- **6,000 Terabytes with all pollutants (and growing!)**
  - At 100 MBytes/second write speed to disk…
    - 728 days of continuous, uninterrupted time to write to disk
    - Only 178 days with no chained pollutants
How MOVES Reduces the work

- **Reduces dimensions in output**
  - No Operating Mode
  - No speed or acceleration
  - No I/M distinction
  - No fuel formulation

- **Minimizes calculation time**
  - Aggregate internally (ex: Speed bins)
  - Cache calculations
  - Chain when possible
  - Chain as late as possible
  - Add SCC after major calculations
MOVES shares the load

<table>
<thead>
<tr>
<th>Master</th>
<th>Worker</th>
</tr>
</thead>
<tbody>
<tr>
<td>● Make work assignment “bundles”</td>
<td>● Do Calculator-level calculations</td>
</tr>
<tr>
<td>● Do Generator-level calculations</td>
<td>● Cross-joins to expand data</td>
</tr>
<tr>
<td>● Do Internal Control Strategies</td>
<td>● Aggregate data</td>
</tr>
<tr>
<td>● Little or no aggregation</td>
<td>● Do chaining</td>
</tr>
<tr>
<td>● Cache data needed by multiple workers</td>
<td></td>
</tr>
</tbody>
</table>
MOVES Architecture revisit

- XLS XLSX TXT
  - Import
  - User
  - Domain
  - Default

- Execution
  - Internal Control Strategies
  - Generators
  - Calculators Bundle Work

- Shared Work Folder
  - Calculators Pickup Results
  - Worker
  - Worker
  - Worker

- Output