MOVES Architecture: From the Ground Up

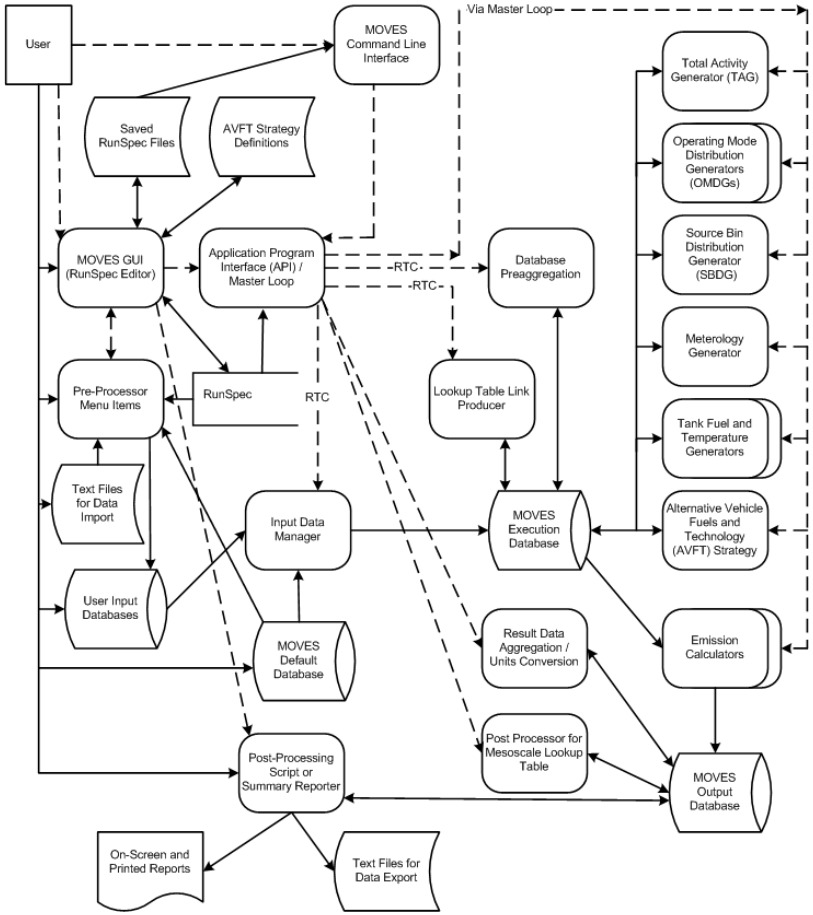
Air Quality and Modeling Center
Office of Transportation and Air Quality
U.S. Environmental Protection Agency

MOVES

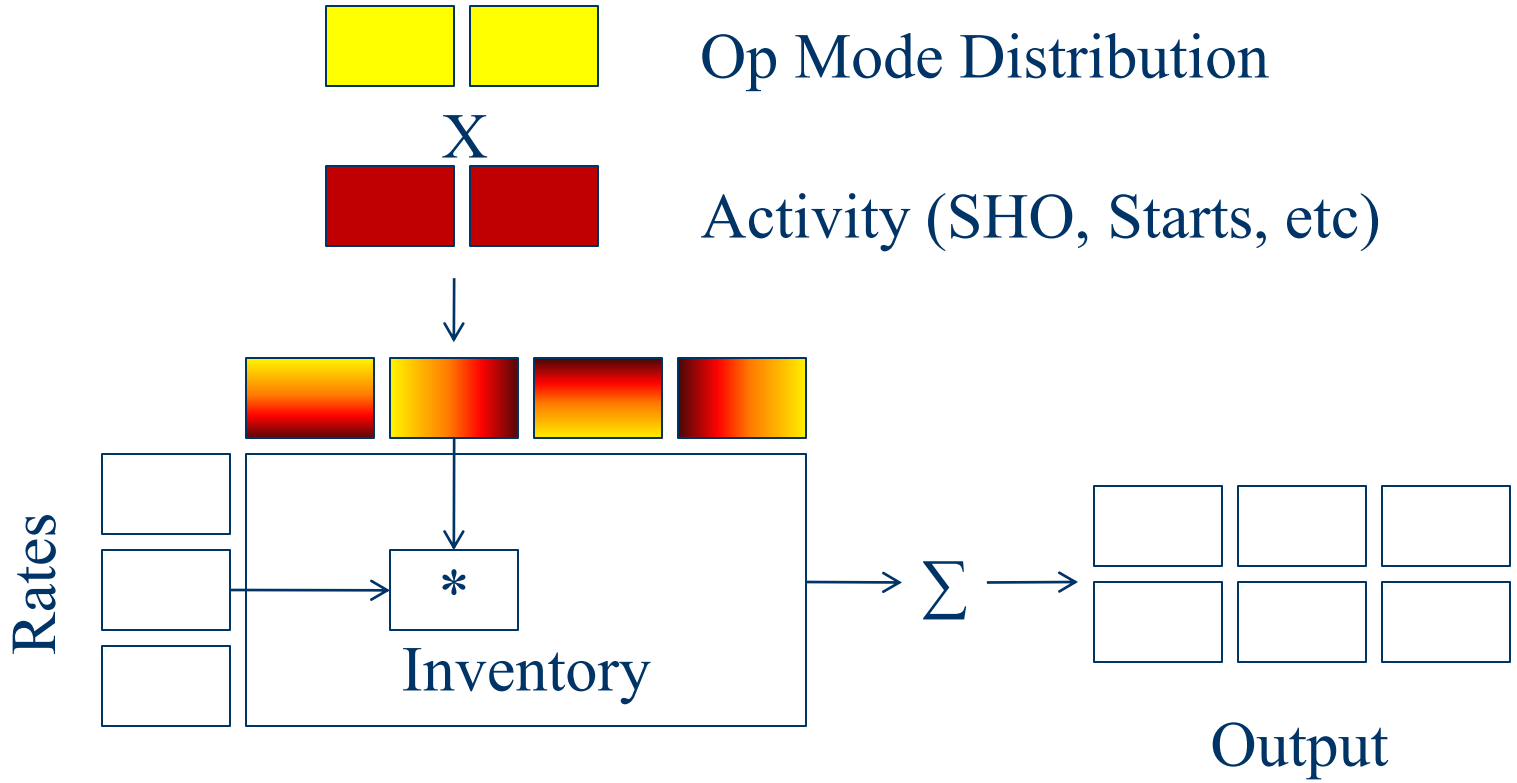
Outline

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

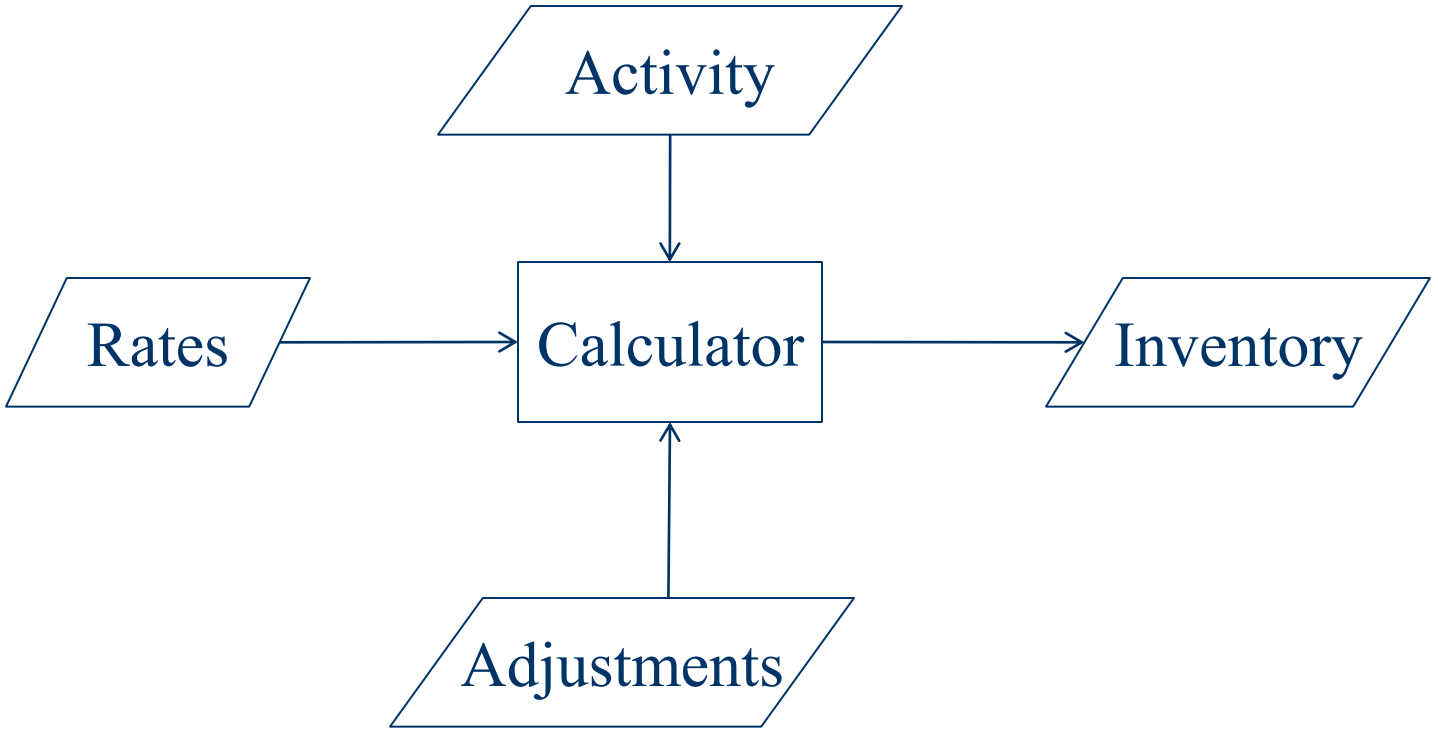
MOVES Architecture Details



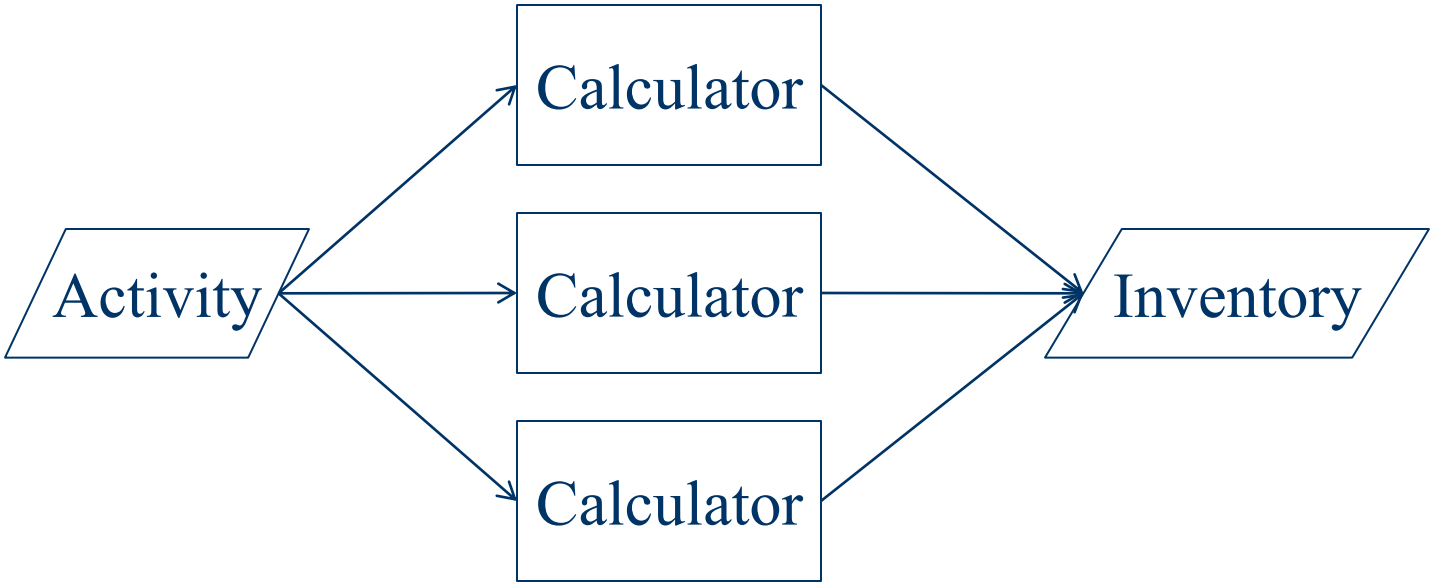
One Calculator, Internals



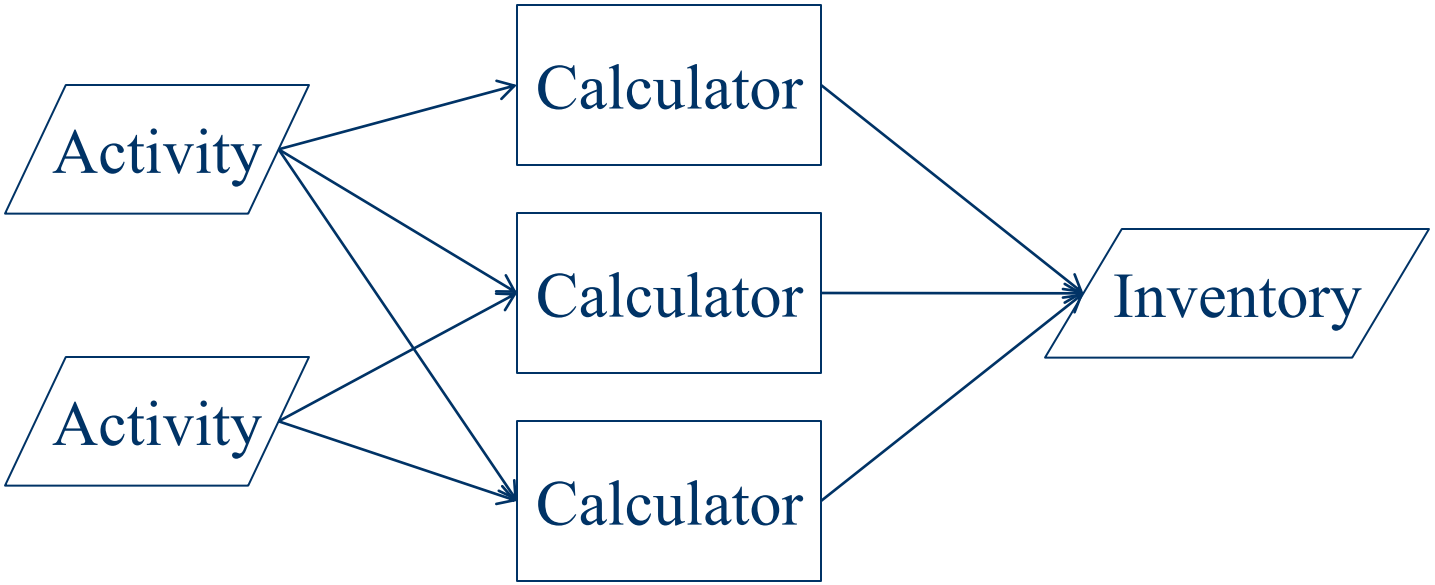
One Calculator



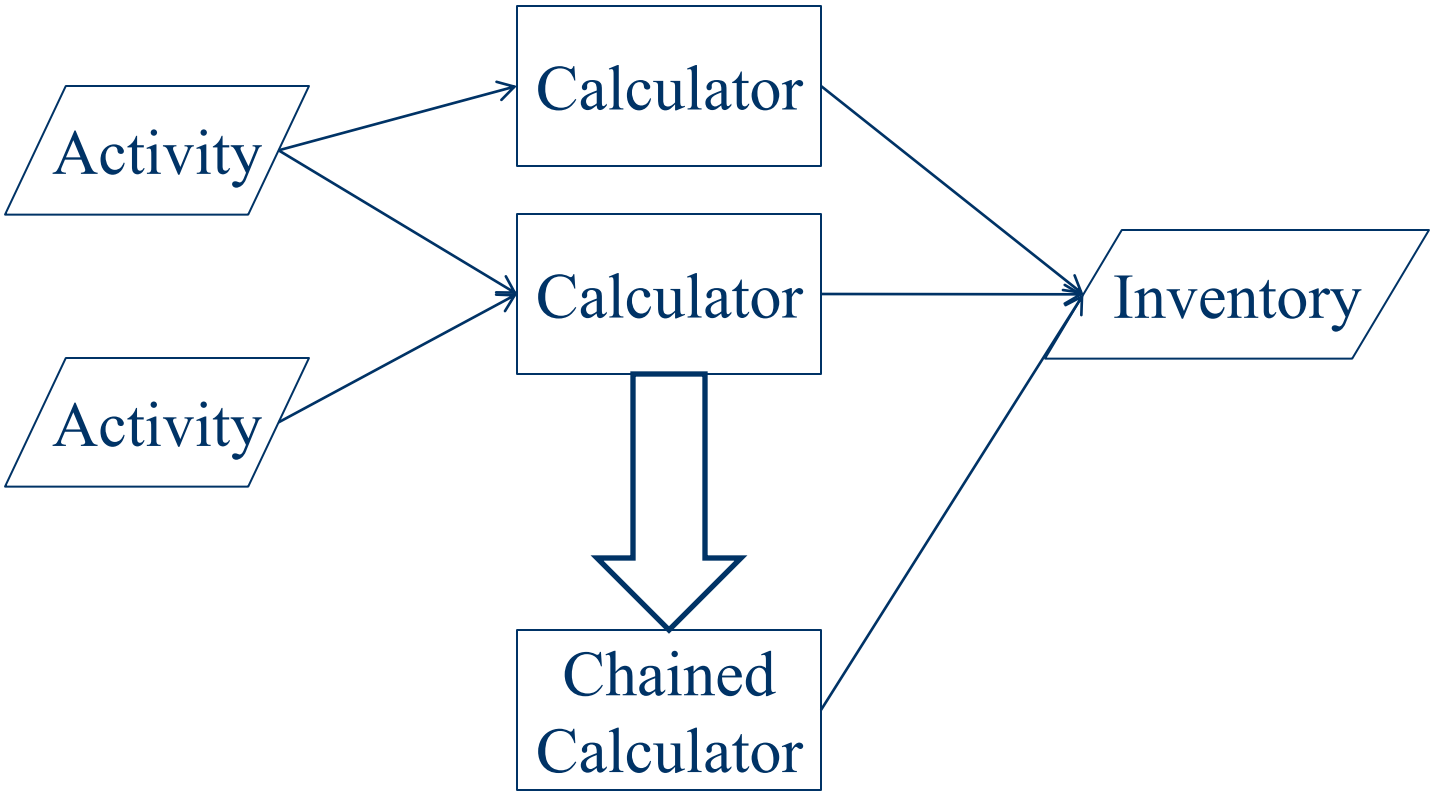
Many Calculators



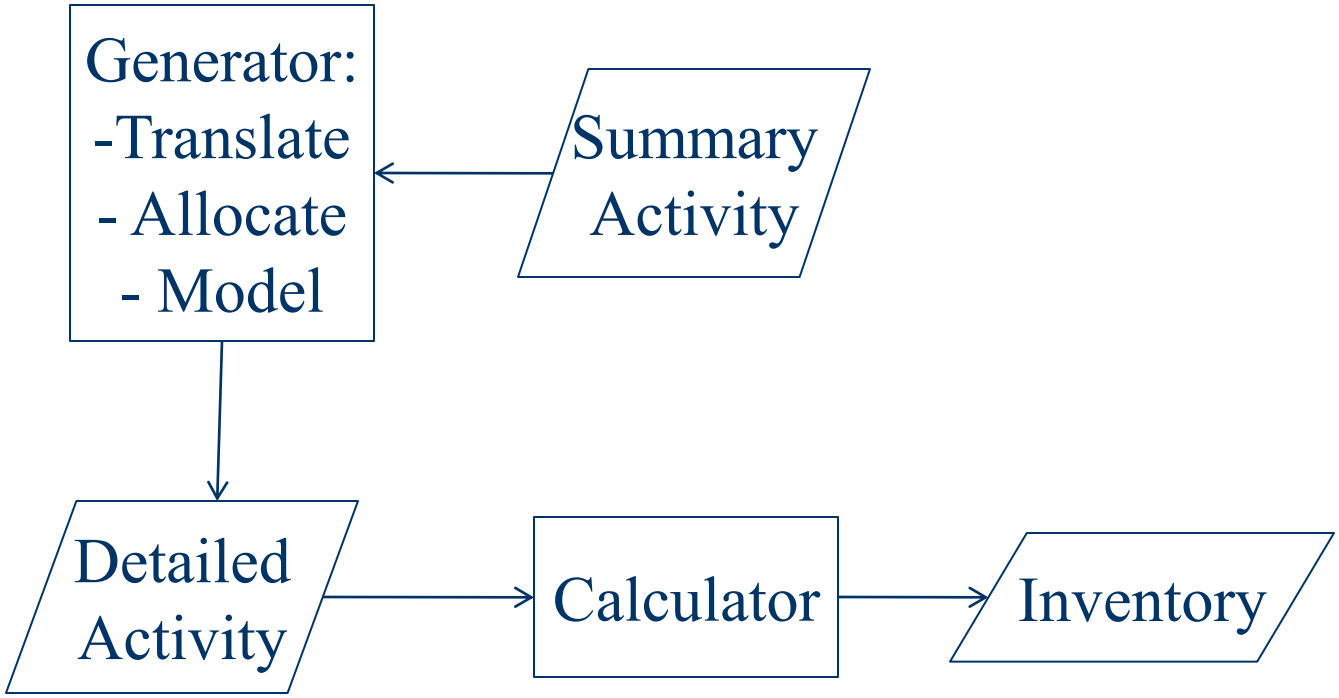
Many Activities



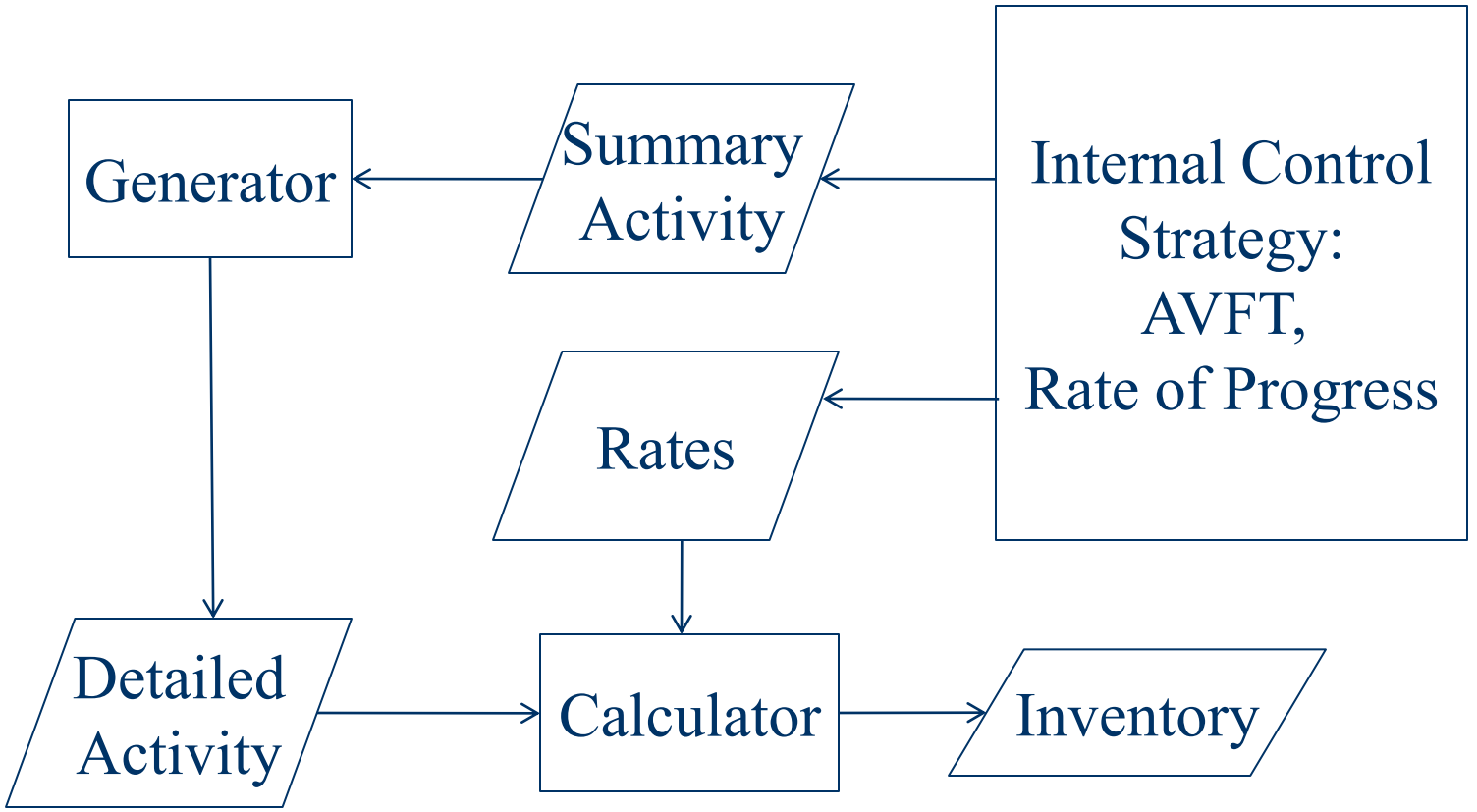
Chaining


Activity
Activity
Calculator
Calculator
Chained Calculator
Inventory

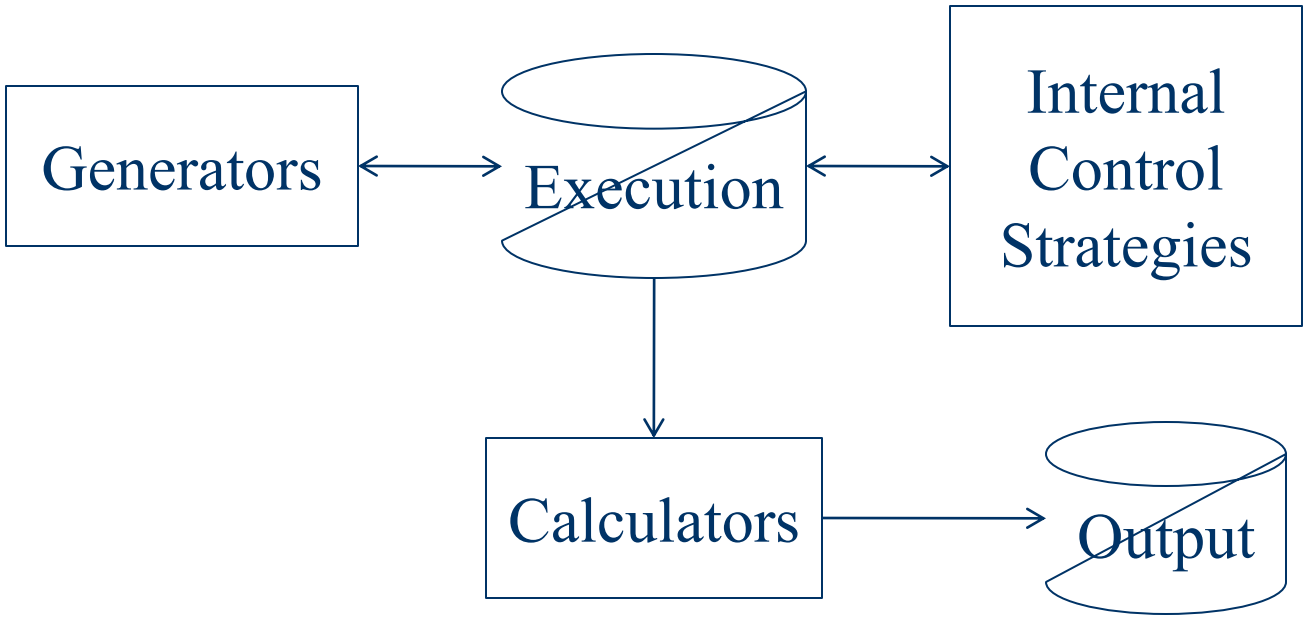
Generators


Generator:
-Translate
- Allocate
- Model
Summary
Activity
Detailed
Activity
Calculator
Inventory

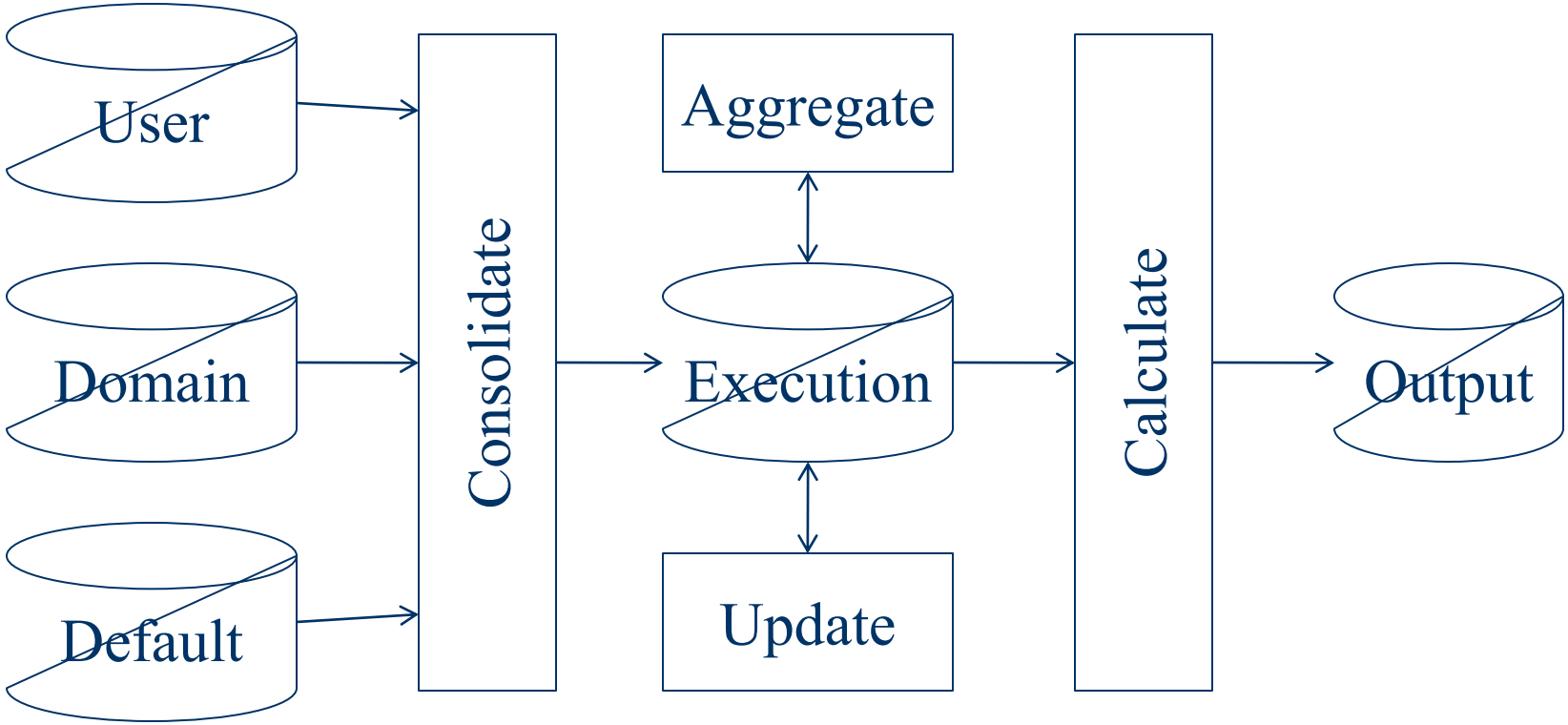
Internal Control Strategies



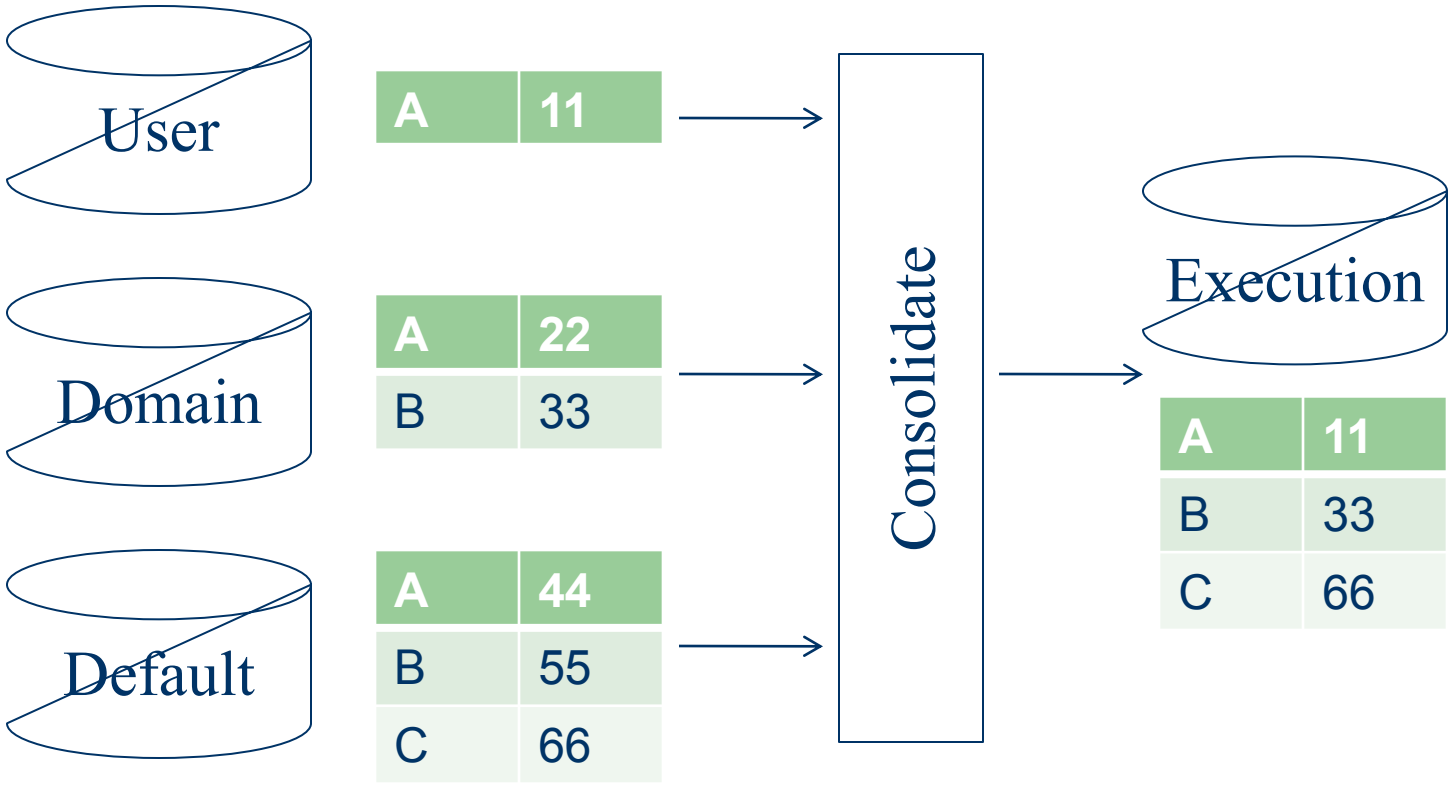
Databases


Generators
Execution
Internal Control Strategies
Calculators
Output

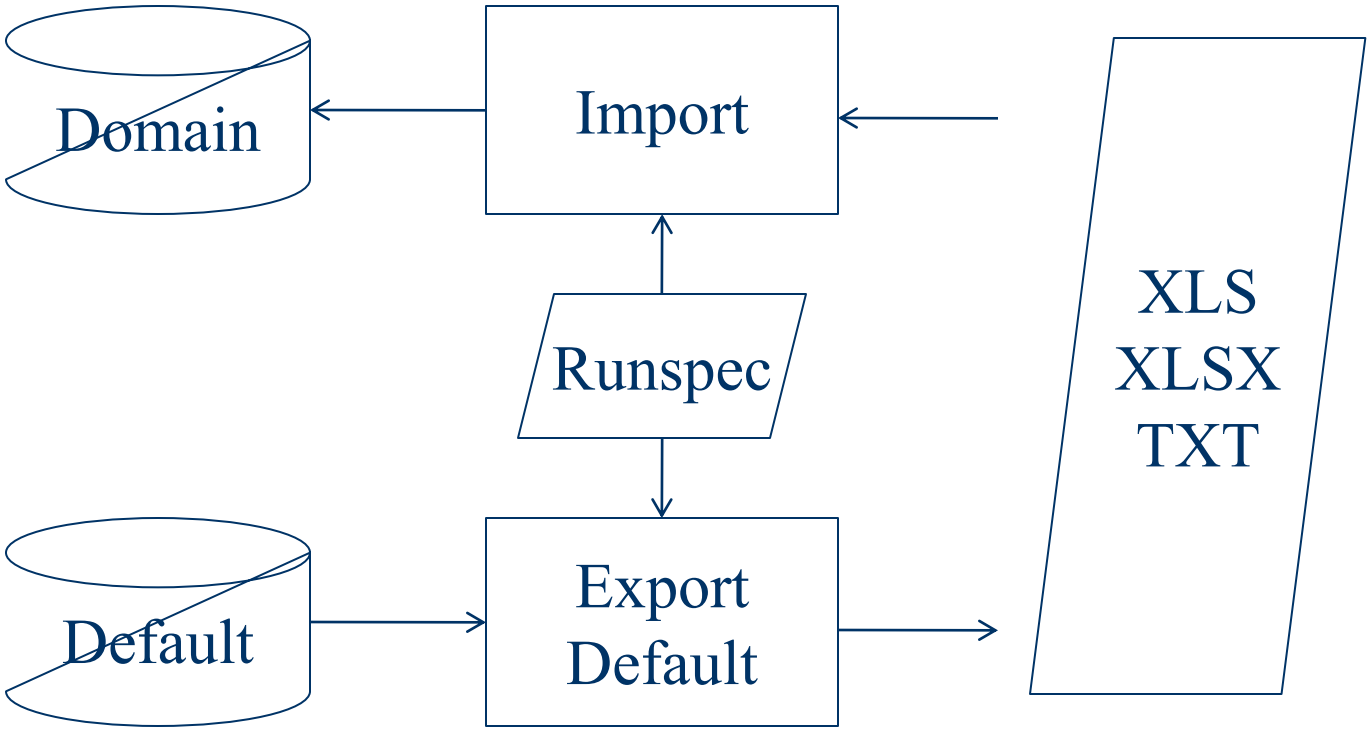
Databases, cont.



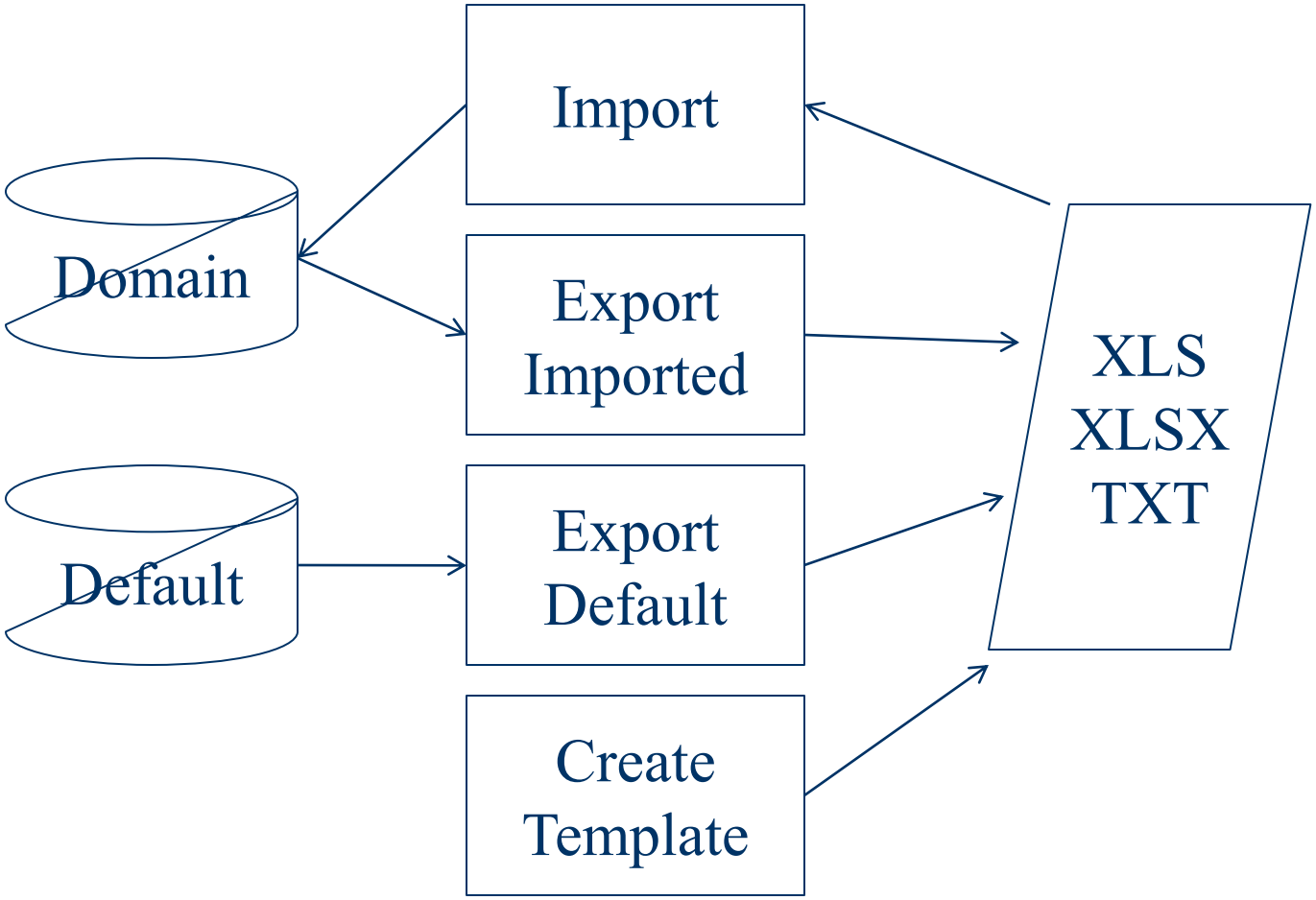
Data Priority



Importers



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

MOVES

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

MOVES shares the load

Master

- Make work assignment "bundles"
- Do Generator-level calculations
- Do Internal Control Strategies
- Little or no aggregation
- Cache data needed by multiple workers

Worker

- Do Calculator-level calculations
- Cross-joins to expand data
- Aggregate data
- Do chaining

MOVES Architecture revisit