Renewables Integration Model Concept of Operations

ST Elbert
H Kirkham

December 2009

Pacific Northwest
NATIONAL LABORATORY

Proudly Operated by **Battelle** *Since 1965*

Renewables Integration Model Concept of Operations

ST Elbert
H Kirkham

December, 2009

Pacific Northwest National Laboratory
Richland, Washington 99352

RIM Concept of Operations
Stephen Elbert
Harold Kirkham

Approved by:	Approved by:	Approved by:
Ross Guttromson		
Project Manager	*Principal Investigator*	
Approved by:	Approved by:	Approved by:
Subsystem Manager	*Subsystem Manager*	*Subsystem Manager*
Approved by:	Approved by:	Approved by:
Subsystem Manager	*Subsystem Manager*	*Subsystem Manager*

Change Log

Contact for revisions and proposed changes:

Harold Kirkham

DATE	SECTIONS CHANGED	REASON FOR CHANGE	BY	REVISION
12/01/2009			STE	Initial Release

Table of Contents

List of Figures

1.0 Introduction

The purpose of the Renewables Integration Model (RIM) is to allow a user to analyze quantitatively the impacts of new wind capacity on power system planning and operation. The user may wish to evaluate existing and future generation mix, and identify the flexibility/controllability attributes required including cycling, re-dispatch, ramping, and reserve. The impact of wind integration on cost may be of interest, as well as other factors such as how much more reserve is needed for load following and regulation, and what is the change on generation ramp requirements.

Since the wind is not very predictable, some users may wish to investigate the effect of forecast errors. For such studies, the user may need to specify the times of run for a Monte Carlo simulation.

It is intended that the RIM be a small and fast version of the "real world." When a simulation has been done, the user will be allowed to "pick over" the results, and extract some of the information for further processing.

This document, the RIM Concept of Operations, describes the way that the model is intended to operate. User interaction, normal and abnormal operation are described.

1.1 Purpose

The purpose of this document is to give the RIM designers and the designers of the systems that deliver other services in RIM's infrastructure an overview of how the system will work and what its capabilities are expected to be. This is not a requirements document, although some requirements can be derived from it. There is a separate RIM Requirements Document.

While the design is still evolving, this document can change. Changes are controlled using the procedures outlined by Harold Kirkham.

This document addresses:

1) The method of initiating interaction with RIM by the User.

2) The approach used for controlling the interaction once it has been initiated.

3) The autonomous means in RIM for detecting and dealing with abnormal conditions.

4) The methods to be used to terminate user interaction.

5) The expected capabilities of the RIM.

6) The means to be used when the capability of the system is being approached or exceeded.

7) The means to be used to fix bugs and improve the software.

8) User roles and responsibilities.

1.2 Scope

This document is applicable to the expert Users wishing to operate the RIM and to the designers tasked with implementing the RIM.

1.3 Objectives

The objectives for publishing this Concept of Operations document are:

1) To present the overall picture of the evolving design of the RIM, so that the RIM designers can refer to it. This is intended to aid coordination of the work involving the development and use of the software modules.

2) To allow the Users of the RIM to understand the way it is intended to work, and thereby aid coordination of the efforts to design and deliver a productive model.

2.0   Technical Background

2.1   Overview

The primary purpose of the RIM is to evaluate the impacts of variable output renewable generation on the bulk electric power system in the presence of a complex complementary systems such as hydrological systems, wind generation systems, solar photovoltaic systems, and energy storage systems. The single simulation environment will combine power system tools such as unit commitment, economic dispatch, load following, regulation and reserves with meteorological tools such as meso-scale wind models. Initially the RIM will be applied to analysis of the BPA operating area shown in Figure 1. This will include power system models of the infrastructure within the BPA service area, models of the hydrological system, as well the necessary meso-scale wind databases. The RIM can in principle be applied to any region as long as the necessary datasets and computational resources are available.

Figure 1. The BPA Service Area with Transmission Lines and Federal Dams[i]

2.2   Basic trade-offs

Power systems operations have numerous aspects that are rarely simulated in a single environment and the RIM is no exception. For example, dynamic simulations and the unit commitment problem are never combined in a single software package. This is due primarily to the limitations of computing power. Unit commitment and economic dispatch are two tools widely used by the power industry to provide the most economical schedule of committing and dispatching generating units to meet forecasted demand levels and spinning reserve requirements while all generating unit constraints are satisfied. The unit commitment problem is a large non-linear, non-convex and mixed-integer problem that belongs to a class of problem in computational complexity theory know as non-deterministic polynomial-time hard (NP-hard).

These are among the most computationally difficult problems known. For N generating units there will be $(2^N-1)^T$ possible commitment combinations for T hour planning. Computers capable of performing a dynamic simulation, while simultaneously incorporating unit commitment and economic dispatch over a 1 year period do not currently exist.

Therefore the RIM is designed to examine many operational issues, but it is not designed to analyze dynamic stability, voltage collapse, or cascading failures.

2.3 Capabilities

The high level capabilities of RIM include the ability to:
- Answer questions regarding the impact of high penetrations of variable output resources.
- Perform simulations up to 3 years in length with 1 minute resolution.
- Model the power system operations in the presence of a hydrological system and wind resources (Assessment of existing conditions as well as design for potential future conditions).

Power system operations include, but are not limited to:
- Incremental Models for power flows
- Unit Commitment (UC) and Economic Dispatch (ED)
- Load following and regulation
- Load growth projections as inputs to the model
- Reserves
- Automatic Generation Control (AGC)
- Energy storage

Hydrological operations include, but are not limited to:
- Models of fore bay and tail water constraints
- Models of tributary inflows and losses due to evaporation and irrigation
- Flow rate limitation
- Flow rate rates of change limitations

Wind operations include, but are not limited to:
- Geographic placement of wind farms within the area covered by the weather model (currently from 42°11′30″ to 49°20′30″ north latitude and from-124°59′30″ to -107°5′30″ west longitude)
- Ability to change turbine types and farm size

2.4 Constraints

To be compatible with realistic computational resources, a number of simplifying assumptions have been made regarding the size and scope of the RIM. This means the RIM cannot do:
- Dynamic simulations or evaluations of system wide dynamic stability.
- Detailed transmission level congestion analysis, *i.e.,* evaluating the congestion of a specific transmission line or set of transmission lines. The RIM does manage inter-zonal congestion, however.

- Evaluation of Volt Amp Reactive (VAR) compensation or evaluations of voltage stability.
- Analysis of cascading blackouts.

Additional operational constraints include:

- The electric power system uses an incremental model for determining power flows between zones. The RIM will be able to accommodate up to 200 zones with 1,000 discrete generators and 1000 discrete loads.
- The wind data set is currently based on a 3 year (2004, 2005, and 2006) historic meso-scale sampling of weather data provided by 3Tier. The maximum simulation time will be for these three years (26,304 hours). With a minimum time step of one minute, a maximum of 1,578,240 time steps is possible.
- The wind data set has a native spatial resolution of approximately 2,000 meter (one minute of latitude and longitude in the area covered).
- The wind data set has a 10 minute temporal resolution that is extrapolated to a 1 minute temporal resolution.
- The hydrological system model will dispatch only major hydroelectric dams. Other hydroelectric dams will have their outputs determined based on a schedule.
- Each hydroelectric dam is treated as a single generator.

2.5 Computational Infrastructure

Due to the large size of the RIM databases (the weather data occupies nearly 3 terabytes), principal computational resources will reside at PNNL. The principal computational engine will initially be the HP Superdome[ii] computer in the Electricity Infrastructure Operations Center (EIOC)[iii] at PNNL. Once a user account on Superdome has been obtained, external access will be via a PNNL approved VPN that requires a SecureID token issued by PNNL. The Superdome operates on a 24x7 schedule.

The users will have access to RIM user files through their home directory on Superdome. Users will not have access to other RIM files such as the proprietary weather data. The Superdome has 64 dual-core 1.6 GHz Itanium II processors with 256 GB of RAM, 12.6 TB of disk and runs Red Hat 4.1.2. The weather data are hosted on a separate Linux system with a 24.2 TB data vault, which will be cross-mounted to the Superdome via a dedicated gigabit Ethernet connection. Only the RIM will have access to this disk from Superdome.

The Superdome is not an architecture supported by MATLAB.

There is a two-user ILOG CPLEX license on Superdome that may be used by the Unit Commitment or other functions requiring Simplex-style optimizations.

2.6 High-level abstraction

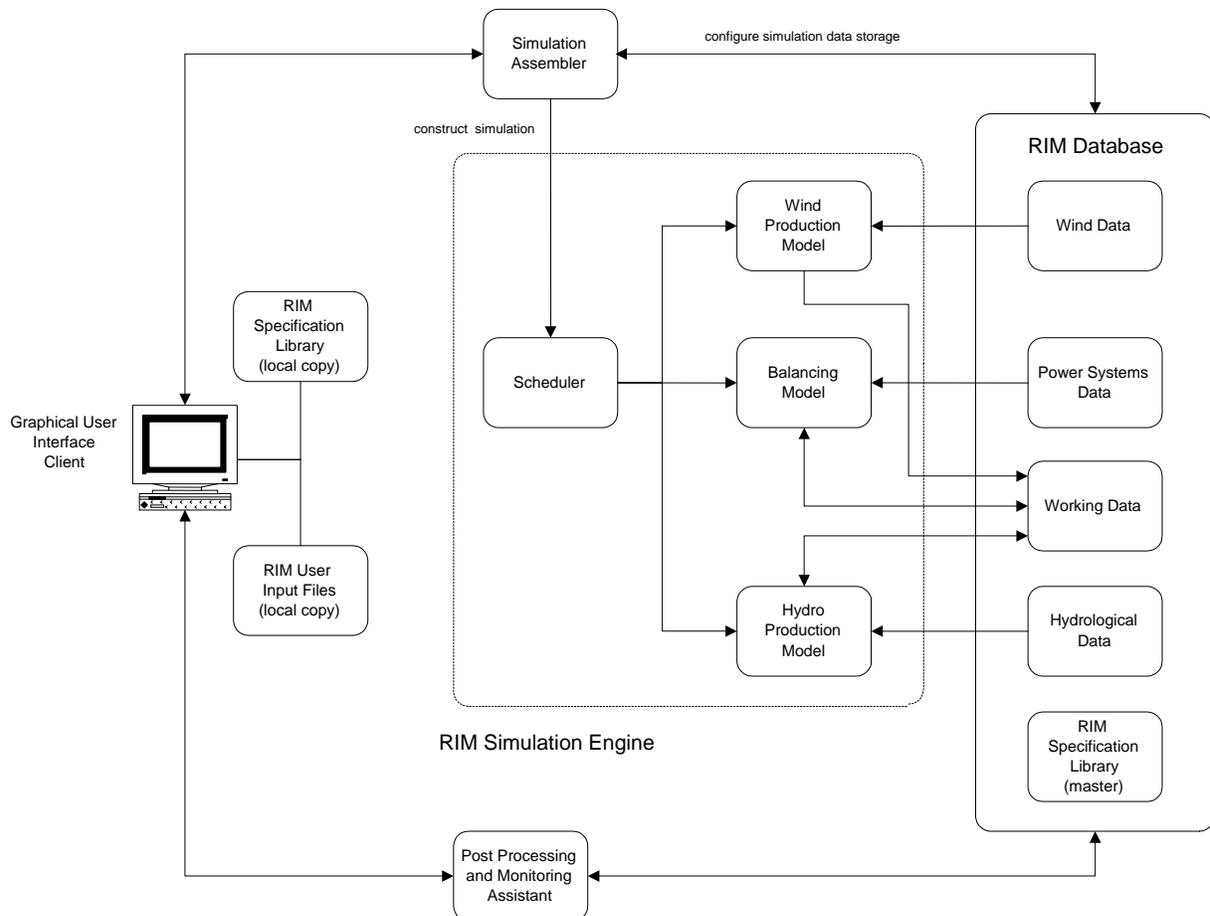The RIM simulation is done in three phases: setup, execution and post-processing.

Figure 2. Major RIM Components

At the highest level, as shown in Figure 2, the RIM contains five major components:

    1) Graphical User Interface (GUI): The GUI is the front end through which the user interacts with the RIM. The GUI is where the user can provide input, set up simulation scenarios, initiate simulations, monitor simulation progress, terminate, suspend and restart a simulation, and access the user files produced by the RIM. It will have no interaction with an active simulation and will have a read-only interface to the database to monitor an active simulation's progress. With the aid of the GUI, the user defines a simulation by providing a description unique to the scenario of interest in the form of an input file. This input file may be original or may be extracted from a local collection of previously saved input files. A local copy of the RIM Specification Library provides the user reference data, such as default values, standard constants, and RIM-specific configuration parameters, for transforming input files into complete simulation definitions compatible with the RIM simulation engine. (The local copy of the RIM Specification Library can be updated periodically from the master copy maintained by the RIM host.) The transformed input file constitutes an input dataset that the user can save locally for subsequent editing or transmission to RIM for execution. In the latter case, the GUI will pass the information to the simulation assembler which will prepare and start the simulation. During the simulation, the GUI monitors activity by reading interim results from the database, with the exception of sending a terminate, suspend, or restart signal.

2) Simulation Assembler: The simulation assembler is responsible for creating the simulation, defining simulation-specific fields in the working data store within the RIM database, establish interconnections between RIM models and the database, starting the scheduler, and defining the GUI's connection to the database for monitoring and post processing activities. The simulation assembler performs these functions as directed in the input dataset provided by the user through the GUI. The wind data, power systems data, and hydrological data depicted in Figure 2 represent the data structures created by the simulation assembler on behalf of the user to contain specifications provided in the input datasets.

3) Simulation Engine: The simulation engine is responsible for generating results from the simulation. Once the simulation assembly assembler disconnects from the simulation engine, the scheduler will initiate processing. The scheduler is responsible for controlling the execution of the simulation. The scheduler manages its progression through the defined time period by triggering the appropriate actions at each simulated time step. At the end of each simulated time step, the collected data will be stored in the database and retained for the next time step's computation.

4) Data Subsystem: The database contains parameters necessary for simulation processing and provides storage for simulation results. Necessary parameters include predefined data like wind and hydrological forecasts, initial values for simulation algorithms, generator specifications, etc. Storage space for interim and final simulation results is allocated by the simulation assembler prior to execution of the simulation. This storage is identified in Figure 2 as "Working Data." Storage allocation is based on specifications for the simulation provided by the user. During execution of the simulation, the user can monitor progress by querying interim data. After execution, the user can analyze simulation results through post processing.

5) Post processing may be performed by the users by transferring the output files back to their own system and using their own tools. In cases where this is not practical due to the number or size of the files, options exist for processing the data at PNNL but this is currently not in the RIM scope. The Post Processing and Monitoring Assistant will mediate the movement of data from the database to user's system via the GUI.

2.7 Model abstraction

The RIM distinguishes three types of models: The Weather and Site Model, the Hydrological Model, and the Balancing Model.
The Weather and Site Model is used to convert temperature, pressure, and wind speed values for a given set of turbines at a specified set of sites (expressed in latitude and longitude coordinates) into real power outputs. This may be accomplished by proprietary algorithms not available to the users or PNNL. Initially this will be done for one year at a time with ten minute resolution using weather data and algorithms supplied by 3Tier. 3Tier uses a proprietary algorithm (SCORE) written in MATLAB (supplied as protected p code), which can be run on the system that hosts the 3Tier data, but not on the Superdome. The resulting output files must then be passed to Superdome. The 3Tier code limits the number of turbines per site to 3, 6, or 10 and uses a file naming convention that allows files to overwrite each other. Eventually other weather data providers will be included. The ten minute resolution data then needs to be

interpolated by RIM, using appropriate probability distribution functions, to one minute steps and the power outputs summed to one value per zone.

The Hydrological Model to be used initially is a highly simplified one consisting of three dams and reservoirs as depicted in Figure 3, although a 10-dam system will be included in the final model. Each of the dams includes an upstream reservoir and associated storage, inflows to the reservoir, reservoir losses du to evaporation, and extractions for irrigation. Flows through each of the dams will be for hydropower production or as spill to satisfy other requirements, *e.g.,* flood control.

Operations for each of the dams are determined by pre-defined reservoir operating rules (i.e., rule curves). Rule curves are designed to guide and manage the reservoir system such that releases from each of the dams are in the best interests of the overall system objectives and consistent with current storage levels and forecast inflows. There are three seasonal operating periods: September to December, January to April, and April to August. The reservoirs will be operated on a 1 hour time step, and the time associated with the streamflow forecasts (TFC) will be pre-set, ranging from 3 days to 7 days.
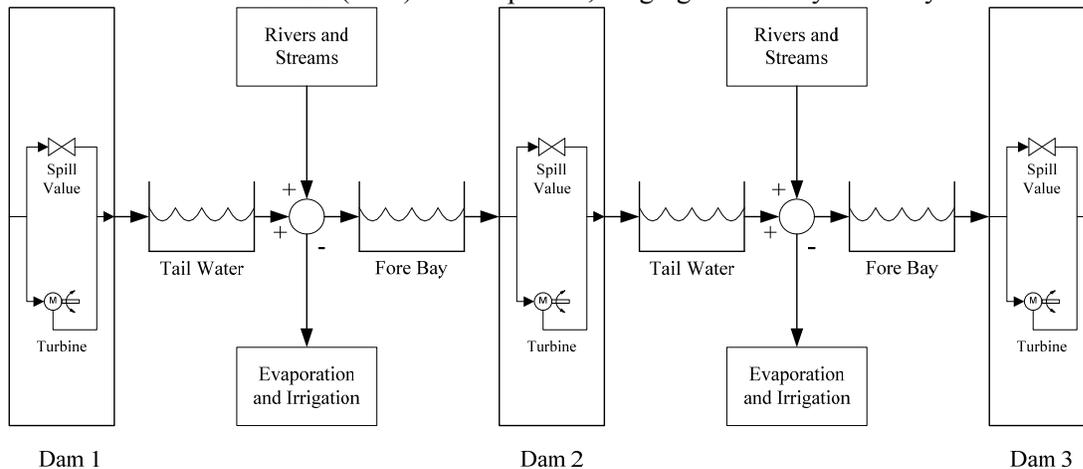

Figure 3. Simple 3 dam system

The Balancing Model is where all calculations dealing with power system operations are performed. Information is collected from the weather (wind turbine generator output) and hydrological models (available hydroelectric resources based on flow constraints—the balancing model will determine generator setting) and integrated into information that can be utilized by the balancing model. In the balancing model the megawatt output of the wind farms, from the weather model, and the available hydroelectric resources, from the hydrological model, along with existing generators are used to model the operations of the power system using Unit Commitment (UC) and security constrained Economic Dispatch (ED). The UC/ED function is used to commit the generators for one hour at a time to meet the expected load. Finer grain control of the system is further provided using UC/ED at the ten minute level and at the one minute level. At the ten-minute level this is called load following. At the one-minute level it is referred to as regulation.

In order for the balancing model to properly simulate power system operations it is necessary for it have access to models that properly represent the electric power system. The electric power system model contains a zonal model of the electrical network as well as detailed information about the load composition and electrical generation characteristics.

The power system planning models for the Western Electricity Coordinating Council (WECC) contain approximately 14,000 electrical buses and thousands of generators and loads. While these network models are well suited for power flow analysis their size and complexity present a significant challenge to other more computationally intensive operations, which are necessary to properly model power system operations. With this in mind the RIM will use a *zonal* model of the desired region that includes up to 200 zones. The model will also support up to 1,000 discrete dispatchable generators and 1000 discrete loads. The zonal model is based on the two following definitions:

1. Area: defined as a balancing authority (BA)/control area.
2. Zone: Any region (within an Area) that has uncongested transmission capacity between any two interior points.

A Zone will never contain more than 1 Area (BA), but a single Area can have multiple zones.
A model of the WECC would have as a minimum 35 zones, one for each BA. If there is any inner-BA congestion, which is likely, then there will more zones, and the RIM allows for 200.
In lieu of a system model that implements a power flow algorithm to determine the flow of power between Zones an Incremental Model will be employed. The power flows between Zones in the Incremental Model do not correspond to flows on specific transmission lines; they simply indicate aggregate inter-zone flows.

2.8 Statistical modeling

The RIM is intended to support statistical sampling of the model solutions through the use of Monte Carlo type simulations. Probability distribution functions may be applied to appropriate elements of the model and a random number generator will select a set of values that will control a specific simulation, with an appropriate number of independent simulations (initially up to 200) being made. The results from all the simulations will be available for post processing.

3.0 Operation

3.1 Startup

The User can initiate access to RIM by logging in to Superdome and executing the RIM startup. This will bring up the RIM GUI that will guide the input process, initiate the execution of the simulation(s) and monitor simulation progress.

3.2 Input

The following information must be available before the simulation can begin.
1. The Load for each Zone for every minute of the period covered by the simulation. The master load information consists of all 92 Zones in the WECC for the years 2004, 2005, and 2006 on an hourly basis. At the minute level, this is 6 MB of load data per zone or 552 MB for all 92 zones. The forecast loads will be generated from this data.
2. The Generators for each Zone and the information needed to calculate the cost of their use and the power they can produce as well as emission information. A cubic function times the fuel cost is assumed along with a variable operation and maintenance cost with additional start up and shut down costs.
3. The Time Period the simulation will cover. Initially this will be 2004, 2005, or 2006. Later, it can be two or three consecutive years.
4. The Wind data specifications. This includes the number and type of wind turbine at each site specified by a latitude and longitude. The zone of the site needs to be specified as well. Initially, additional information such as turbine height may also need to be specified.
5. The Hydrological data specifications, which include elevation-storage-area relationships, discharge limits, reservoir elevation limits, rule curves, irrigation withdrawals, abnormal operating events, and inflow forecasts.
6. The Interconnect data specifications, including a minute by minute value for the WECC Area Control Error (ACE) and related control performance standards (CPS).
7. The Case Identification that can be used to distinguish one model run (or set of related

runs) from another.

8. The Model Run definition data, which provides model options such as the number and type of Monte Carlo runs to be made.

3.3 Normal Operation

Running RIM involves each of the major RIM components described in Figure 2.

1) The GUI guides the User through the choice of input data, including the kind of model to be run.
2) The simulation assembler subsystem takes the input information collected by the GUI, constructs the data objects the simulation will use and connects them to the simulator and the GUI. An important reason for the existence of the simulation assembler is that it may take some time to stage some of the data, in particular the wind data, and the simulator will run more efficiently if all of the data it needs are present at the beginning of its operation. For Monte Carlo runs the assembler will spawn the appropriate number of simulator runs.
3) Once the simulator begins running it will keep as much data as possible in memory and periodically dump the time varying simulation results to disk. Communication on the progress of the simulator with the GUI will be through data on disk, which is more fault resilient than direct communication through memory or message passing with minimal performance impact. Should a fault external to RIM occur, the GUI can be used to rollback, with the help of the simulation assembler subsystem, the simulator to the last valid time step in the data.
4) The GUI can be used to select the appropriate output data for post processing and move it to a location specified by the user.

3.3.1 Staging 3Tier Wind Data

The wind data used to generate the power output of a given set of turbines at specified locations is provided by third-party suppliers, but, because of the volume of the data, this information must be computed before the RIM is run.

The initial wind data has been supplied by 3Tier in the form of one netCDF[iv] format file per day for each of the years 2004, 2005, and 2006, with separate folders for each year. Each file is 5.213 GB all 1096 files take 5.578 TB.

The area covered by the 3Tier data is shown in Figure 4.

North to south the grid covers 430 values of north latitude: from 49°20′30″ (49.341667), which is just north of the border between the US and Canada, to 42°11′30″ (42.191666), which is just north of the southern borders of Oregon and Idaho.

West to east the grid covers 1075 values of west longitude, from 124°59′30″ (-124.99167), which is just off the western coast of Washington and Oregon to 107°5′30″ (-107.09167), which is in eastern Montana. The proprietary 3Tier SCORE algorithm is used to produce a set of comma separated value (.csv) files, one for each height at a site, with a 19-character time stamp, the wind speed in m/s, and the power output in MW, with one set of values every ten minutes. SCORE also produces an ancillary set of similar files that contain the time stamp, wind speed, temperature and pressure used to calculate the power output.

Figure 4. Area covered by 3Tier wind data (corners indicated by ⊕)

The RIM must then read these files and interpolate to one minute while summing the power of all sites in the same zone, resulting in a netCDF file that contains array with a column for each zone and a row for each minute of the year being simulated. The interpolation program provided by 3Tier uses a different probability distribution function for large farms (>20 turbines) and small ones.

3.3.2 Staging AWS Truewind Data

PNNL is currently negotiating with AWS Truewind for the additional wind data. It is possible that the power output calculation would be performed on their premises and the resulting data sent over the network to PNNL, requiring additional staging time.

3.4 Abnormal Operation

RIM makes many attempts to validate and verify the correctness of its operation, particularly with respect to the input data. When values are out of range or missing, an error message sufficient to understand the problem is generated. Processing is expected to continue, however, as long as feasible while continuing to check the input for correctness. This allows multiple errors to be uncovered in a minimum number of attempts rather than serially discovering multiple errors. When an error does occur, output generated during the simulation may require manual cleanup, assisted by the GUI. It may be determined that the simulation may be restarted.

3.5 Fault recovery

When a system failure occurs, such as a power interruption, the time series data on disk may be used to continue the simulation from the last recorded time step.

3.6 Simulation termination

The simulation will normally continue executing until all time steps have been completed. At this point the simulation engine will mark the database files as completed and the GUI monitoring mechanism will indicate the simulation has completed normally.

In addition to monitoring the progress of the simulation, the GUI will also have a mechanism to terminate the simulation. Terminating the simulation can be done along with the removal of all working files or not. When the working files are not removed it may be possible to restart the simulation. When the working files are removed the simulation cannot be restarted.

3.7 Output data

The simulator periodically writes the data objects in memory to disk. The time series data available after the simulation finishes may be selected through the GUI and uploaded to the user's system.
Some of the output data that will be available at the end of the simulation run include:
individual generator output of each dispatch
load following reserve of individual generation unit
regulation reserve of individual generation unit
contingency reserve of individual generation unit
frequency response reserve of individual generation unit
number of start/stops of individual generation unit
emissions of individual generation unit
heat rate of individual thermal unit, gallons of water/MWh of individual hydro unit
wear and tear cost of individual unit
power flow on the zone tie-lines

3.8 Post processing

The users are responsible for converting the output files they have uploaded to the proper format for their own post processing and analysis. The RIM project will make available tools for converting the files to popular formats such as csv.
Some of the data that could be available after post processing includes:
the aggregated cost of generator output change caused by redispatch
the aggregated cost of actual load following
the aggregated system level actual load following reserve
aggregated system cost and level of actual regulation reserve
aggregated system cost and level of actual contingency reserve
aggregated system cost and level of frequency response reserve
the cost of the inefficiency of traditional generation
aggregated system level of number of start/stops
aggregated system level of emissions
the efficiency of traditional generation
the aggregated system cost and level of wear and tear cost
total and average production cost for generation
Tie-lines power flow with congestion flag (congested, close to limit, uncongested).

Analysis by generation unit type is also possible.

3.9 Maintenance

There are two aspects to software maintenance.
1) Regular scheduled releases incorporating new or enhanced features.

2) Corrections that resolve errors with a range of severity. Fixes for severe errors (ones known to cause incorrect results or threaten the integrity of the system) will be released as soon as the correction is known and validated. Fixes for other errors may be postponed until the next scheduled release.

The operation of the RIM is dependent on the computational infrastructure. The real-time performance of the RIM depends on a variety of factors including the scope of the models being run and resource contention. Resource contention can include the available processing power (system load), disk usage, network availability, and scheduling priorities.

3.10 User Roles and Responsibilities

The Users are subject to the standard acceptable use and cyber security policies of PNNL. The Users are assumed to be expert enough to determine the appropriate scientific or engineering use of the RIM. The Users are responsible deleting or removing the data from previous simulations that are no longer relevant.

References

[i] 2008 BPA Facts, http://www.bpa.gov/corporate/about_BPA/Facts/

[ii] http://superdome.pnl.gov/

[iii] http://eioc.pnl.gov/

[iv] netCDF Users' Guide, http://www.unidata.ucar.edu/software/netcdf/docs/netcdf/.