Supporting Power Grid Optimization Competition

(Draft for discussion with ARPA-E program team only)

Pacific Northwest National Laboratory

April 27, 2015

This note outlines tasks and approaches for hosting the ARPA-E OPF competition, and the outline will be modified accordingly to future changes of the competition requirements.

It is organized as bulleted points for easy reading. It can be expanded to a more formal document if necessary.

# Scope of the competition

## Open Competition for Solving Realistic AC OPF

The scope of this competition is to search for the next generation optimization models and algorithms to solve optimal power flow problems through a series of competitions with participants from a larger community of researchers, practitioners, and individual enthusiasts.

The competition team will support ARPA-E program office to *design, host, and execute* this competition. The team will ensure an open, fair and successful competition with the goal to close the capability gaps between solving OPF models/algorithms efficiently and effectively and the power industry’s requirements.

Reference: Tim Heidel’s scoping document.

# Computing platform design and accessibility

## Computer characteristics:

* + **Operating Systems**

PNNL can support the competition on both Unix and Windows systems. We will support Windows systems in native or Virtual Machine mode.

* + **Software and Libraries**

The competition team can host a website as the main interface that includes announcement, data download, code/algorithm submission, and discussion forum. PNNL will install, if not already, necessary software and libraries. For example, optimization solvers may include CPLEX, GUROBI, SCIP, CBC, IPOPT and other relevant optimization libraries in Coin-OR. PNNL will setup environments for C++, Java and Python programming. Compilers available include CUDA, PGI, GCC, Intel, PGI, Lahey, Open64, Pathscale, and UPC. Many of these elements currently exist in PNNL’s computing infrastructure.

* + **Hardware and High Performance Computing**

PNNL can offer the stat-of-the-art hardware to support this competition. PNNL has a 308 compute node cluster (Constance) consisting of dual Intel E5-2670 (Haswell) CPUs (24 2.3 GHz cores per node) with 64 GB 2133 MHz DDR4 memory per node, 500 GB local SSD disk (>450 MB/s) with a 56 Gb/s FDR Infiniband interconnect. PNNL also has an SMP compute node (Bigmem) with 8 Xeon E7 8860 CPUS (ten 2.27 GHz cores each) with 2 TB of RAM memory. The PNNL compute nodes are backed by a multi-petabyte global Lustre file system. PNNL also has 16 compute nodes with Intel 7110P (Phi) co-processors (61 1.1 GHz cores)(Philo) and 32 compute nodes with NVIDIA Tesla M2090 (Fermi) GPUs (512 1.3 GHz cores) (Olympus/GPU). Large RAM memory (2TB) can be used to test memory intense problems.

Constance, Bigmem, Philo and Olympus/GPU are for HPC implementation.

* + **Cloud Computing**

PNNL has an OpenStack Cloud Testbed that is similar to the Amazon EC2 and Rackspace services.

* + **Backup**

PNNL can set up a competition server and a backup server.

## Accessibility

The team will set up a web site with multiple channels to exchange information with participants, ARPA-E sponsors, competition team members, and a steering committee. The multiple accesses will have different permissions to ensure dataset security and protect proprietary data and software codes while ensuring transparence of the competition.

About data access, there are two mechanisms to communicate datasets with competition participants.

Training data sets will be made public available and be hosted on PNNL’s server. The participants can download the training datasets to develop and test their models and algorithms. Participants can also use the website to upload their models onto the server and obtain the test results from the training datasets. The participants can use these datasets to become familiar with the server and compare and understand the difference of performances between the server and their native development environments.

Evaluation data will be host-only on PNNL’s server behind the lab’s firewall. Participants will submit their models/algorithms in the same way as they do for the training datasets. The real-time run logs of their algorithms will only be accessible by the competition team and ARPA-E sponsors as well as the participant who submit the algorithm. The final scoring results on designed metrics will be sent back to all participants in a standard format.

Participants can submit either executables or source codes.

## Cost

The fixed cost in the design stage will include 18 FTE months, funding 2-3 staffs (6-9 months) to set up and test the platform including the hosting server and develop the website.

Operation cost will include 8-12 FTE moths per year to maintain the server, trouble shooting, and update/retesting software for different competition stages.

* Flexible runtime configuration – The runtime environment will be configured as close as possible to that requested by the participant and executed multiple times in a dedicated environment.
* Timing measurement – Elapsed wall clock time will be recorded for each run with minimum, maximum and average times reported.
* Openness – The run logs will be available for the competition team and ARPA-E sponsors as well as the respective participant who submit the algorithm. The result, including objective value and run time will be available for the participants. The optimal solution of training datasets will be available to the participants. The optimal solution of the evaluation datasets will be accessible only to the competition team, ARPA-E sponsors, and a steering committee with proper NDA.
* Resources to leverage, including resources from partners if known

# Optimization problem and dataset design

## Optimization Problem Design

* + **Problem Selection**

The competition will involve solving a number of OPF with varying complexity. The complexity varies in three dimensions, i.e., deterministic vs stochastic, static vs temporal, and steady state vs transient. The goal is to have the final competition problem close to what the industry needs and to pave the road to future competitions like UC competition. The problems will be released in stages and winners of each stage will be announced.

|  |  |  |  |
| --- | --- | --- | --- |
| Problem | Discrete/Continuous | Deterministic/Stochastic | Static/Dynamic |
| Basic ACOPF | Continuous | Deterministic | Static |
| ACOPF with N-K | Mixed | Both | Static |
| Multi-period ACOPF | Continuous | Both | Both |

* + **Problem Design**

Competition problems are based on two formats: (1) fixed model and algorithm design and (2) open model and algorithm design. In format (1), an optimization formulation is given and participants compete on developing algorithms to solve the optimization model. In format (2), a description of the problem is given and a set of objectives is given, and participants have freedom to develop models as well as algorithms for solving their models to provide the answers to the set of objectives.

For example, solving the well-defined ACOPF is of format (1) and can be a problem for the first stage competition. Participants choose any methods to solve the ACOPF and the solutions will be evaluated for feasibility, total cost and runtime. Deciding sequential dispatching under load and generation uncertainty is of format (2). Participants can choose how to formulate this multistage problem, whether as a multi-stage stochastic program, robust optimization, or chance constraint problem and to design algorithms for their proposed models. The competition team will decide what the format of solution is for the submission and how to evaluate these solutions in a stochastic environment. For example, the metrics can be feasibility and cost for the solutions obtained within a given time limit.

## Dataset Design

The competition dataset can be divided into two groups. One set describes power systems including topology, line (admittance and limit), generator parameters (capacity, ramping rate, and costs) and other data like RAS. The other set contains scenarios for renewable generation level, load, and other scenarios like N-k contingencies. The datasets will be in human readable text format, e.g. csv.

Ideally, there will be a handful of datasets reflecting different sizes and characteristics of power systems. These power system datasets combining with scenario datasets can serve different stages of the competition.

Datasets will be tested in the competition format to ensure the existence of feasible solutions. Software tools, e.g., PSSE, PowerWorld and PNNL’s enterprise tools, will be used to test power system stability. We will also implement optimization models and algorithms to establish benchmarks.

PNNL will utilize its existing datasets and past studies to construct realistic power systems and associated parameters. For example, PNNL has collected datasets and studies for TEPPC cases, simplified WECC cases, Duke energy cases, and other synthetic datasets.

## Cost

It is estimated about 18 FTE months, including 3-5 staffs (6-9 months), to design and test the competition data. To hosting the competition, it will take 2-3 FTEs per year for developing, testing, and running benchmarks.

# Competition Execution and Scoring

## Execution of Participants’ Algorithms

Participants will use training data to test their algorithms and get familiar and calibrate with the server system. In the evaluation phase, participants can submit their codes or executables through email or web.

## Scoring Metrics

* + Metrics to evaluate competitions
* Optimality – The optimality can be measured by the gap between participants’ solutions and the benchmark solution. An alternative choices to compare total costs among the participants within a set of time limits.
* Timing measurement – Elapsed wall clock time will be recorded for each run with minimum, maximum and average times reported. More detailed in-run timing measurements can be provided if the participant’s codes are designed to do so.
* Feasibility/practicality – Solutions will be evaluated independently in other software tools to check their feasibility and stability for the given power system and scenario.

## Evaluation

The outcome from the competition can serve as the basis to start a FOA call with focused technologies and directions leading to disruptive technologies based on OPF for the power industry. The evaluation phase is the step connecting the competition to a FOA call by learning from the competition result and providing the direction to the FOA. In this step, it is important to identify required basic and applied research needs and translate them into broader requirements for the future proposals from academia, labs, and industries.



# Partnership and Outreach

PNNL will team with other organizations to support ARPA-E optimization competition throughout the *design, hosting, scoring, and evaluation* phases. PNNL will organize the phase for problem design in which it is important for all team members to discuss and generate a set of problems centering on OPF and reflecting the different aspects of OPF. The suggestions listed in Section 3.1 can serve as starting point. PNNL can host the competition to ensure the information security, required licenses for power system simulation, and sufficient power engineering, IT and hardware support. The inputs from other team members from their experience are important and will be communicated throughout the competition. Satellite hosts are feasible at the early stages of the competition to ensure availability and reachability to a large number of participants. The scoring system will be a team effort to ensure and fairness and relevance to practical problems. The past experience from the team members will be valuable to develop the scoring system that should be flexible to change accordingly to the competition needs. The team will work jointly in the evaluation phase that is a bridge to an FOA. It is important to derive from the competition results and establish broader research agendas.

To ensure success of the competition by encouraging participation, designing realistic competition problems close to industry’s need and providing fair scoring, it will be helpful to organize a steering committee that includes industrial and academia partners. The steering committee is to provide a broader view for competition problems to keep tight linkage to industry needs, to ensure the inclusion of promising optimization techniques, and to support reaching a larger community of participants.

To ensure the fairness and completeness of competition design, the competition team can organize a workshop and invite industry, government, and academia experts. The main objective of the workshop is to gain suggestions and critiques on chosen problems, datasets, scoring systems and to learn from experiences organizing competition from other communities.

# Total Cost

The total labor cost will include 4-6 full-time staffs for 6-9 months to design and setup the completion and 2-3 FTEs to operate the competition.