EAR-GUI User Guide

v1.0

1 Introduction

Energy Aware Runtime (EAR) package provides monitoring and energy saving solutions for super computers based on MPI and SLURM.

Lenovo Intelligent Computing Orchestration (LiCO) is an infrastructure management software for high- performance computing (HPC) and artificial intelligence (AI). It provides features like cluster management and monitoring, job scheduling and management, cluster user management, account management, and file system management.

As the first step for LiCO to integrate EAR, EAR-GUI provide a web-based user interface to create or modify EAR configuration file. This not only provides easy-to-use configuration tools for EAR system administrator, but also helps the LiCO team learn more about EAR.

2 Installation

EAR-GUI release as a Docker image, it can work independently. There are two ways to get the image:

1. Access the Antilles project from the Github, and build the Docker image by yourself. (Antilles is the LiCO open source version)
2. Access the EAR-GUI Docker Hub Tag directly.

Execute the follow command to launch EAR-GUI service:

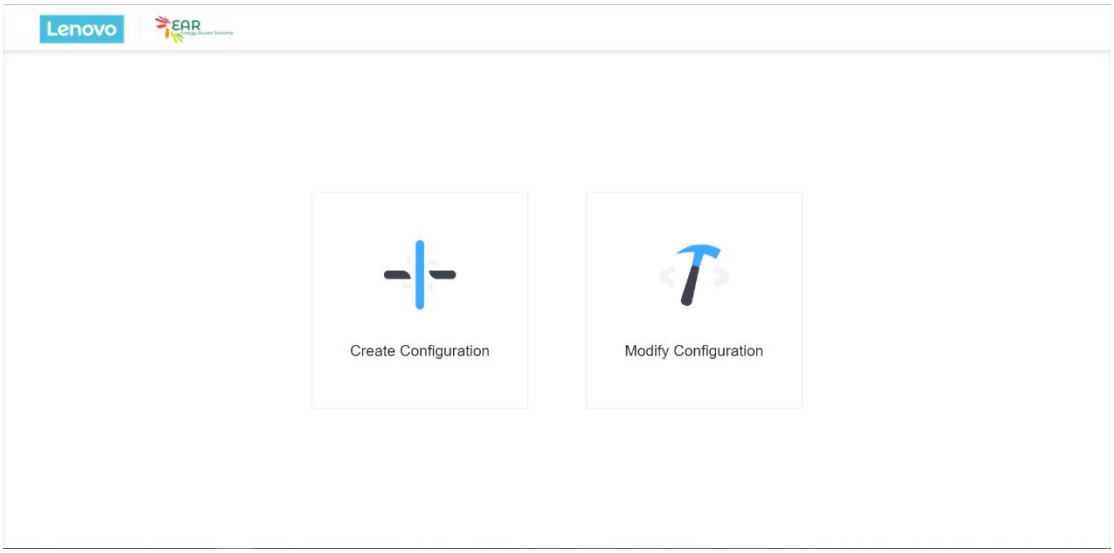
Open your web browser and access URL:

http://<host_ip>:<port>/

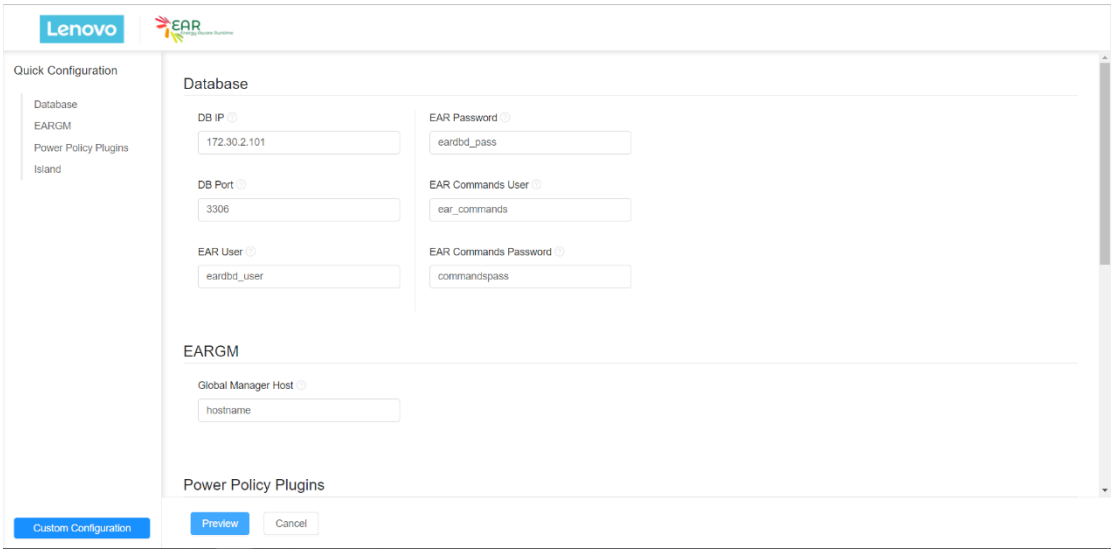
3 Usage

3.1 Create a new configuration file

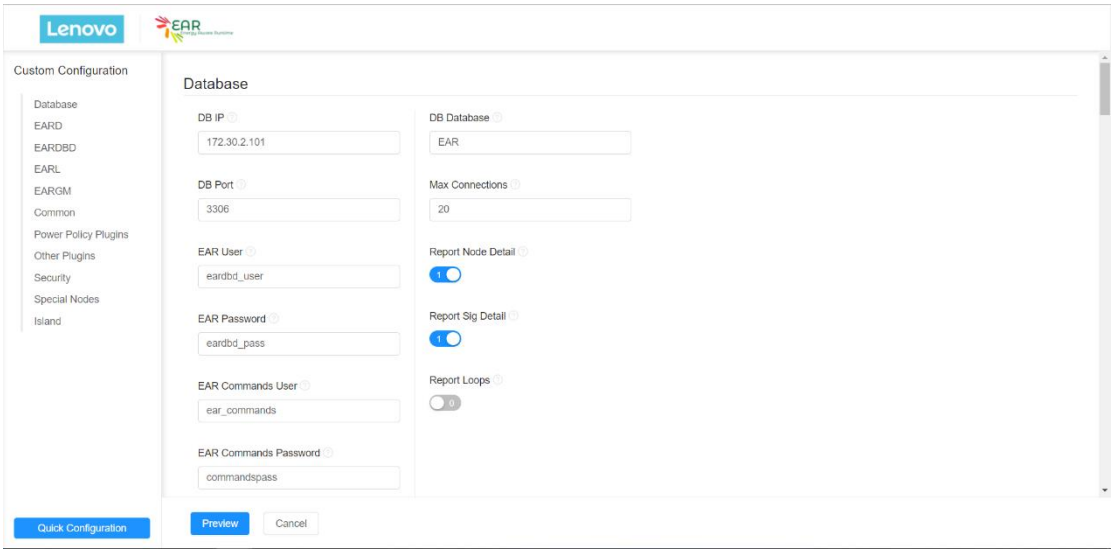
1. Click **Create Configuration** on the entrance page:



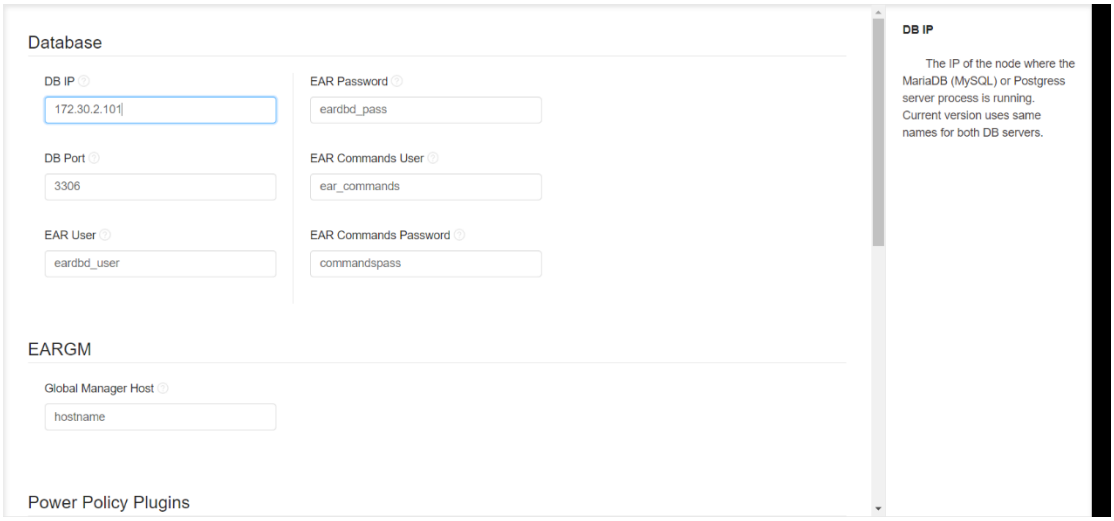
2. Input the setting values on the configuration page:



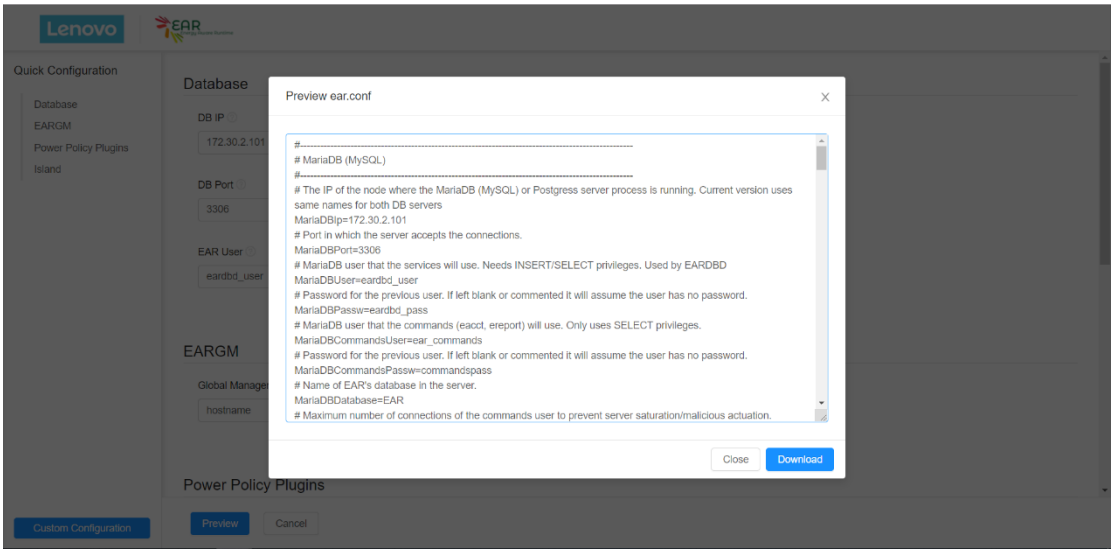
3. There is a navigation bar on the left of the page. Using it can quickly switch between different configuration sections.
4. By default, the configuration page is in quick mode. In the quick mode, the advanced setting items are hidden. Click **Custom Configuration** in the bottom of the navigation bar, the mode can be shift to custom mode. In custom mode, all the available setting items are shown.



5. Click the icon after every label, you can view the help message of this item:



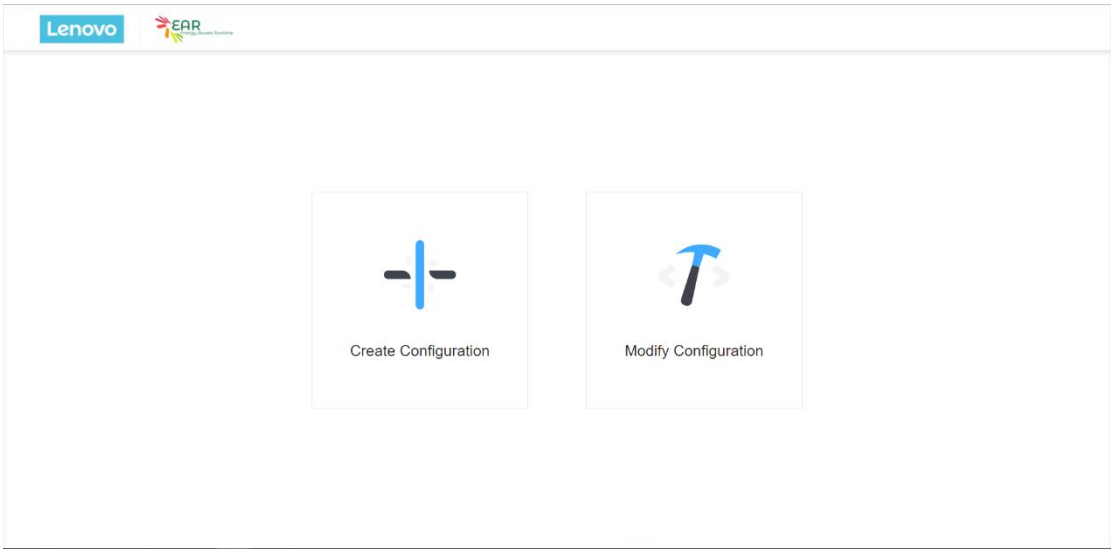
6. After you finish the configuration, click the **Preview** on the bottom of the page. You can view the content of the ear.conf on the popup dialog:



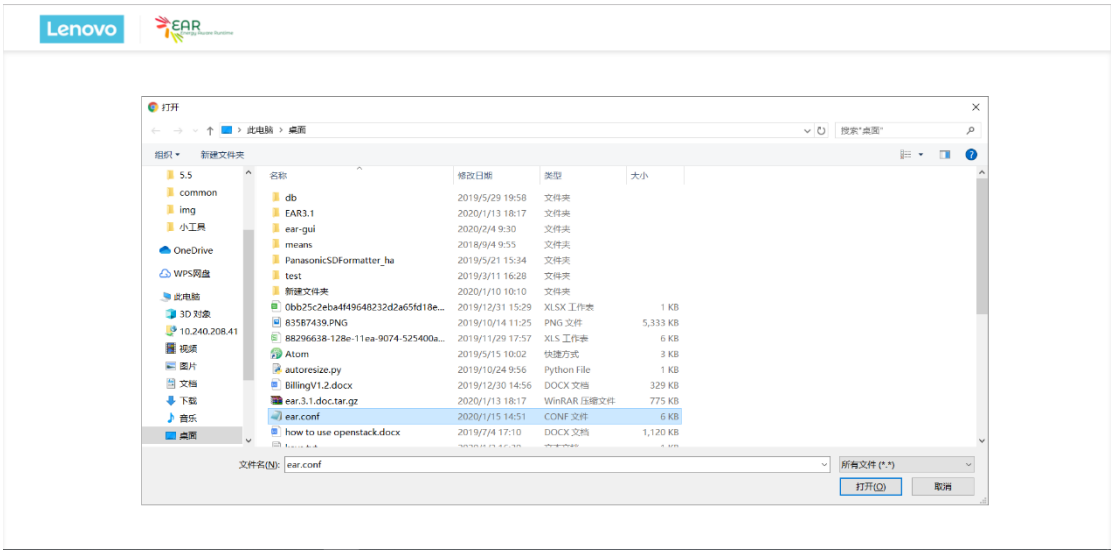
7. You can click **Download** to save the content to your local storage.

3.2 Modify an existed configuration file

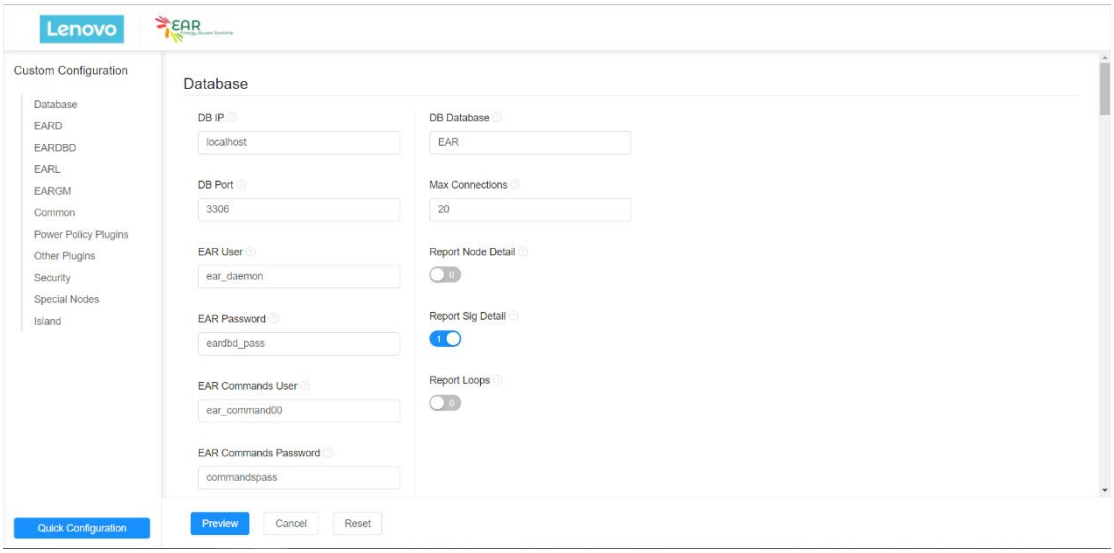
1. Click **Modify Configuration** on the entrance page:



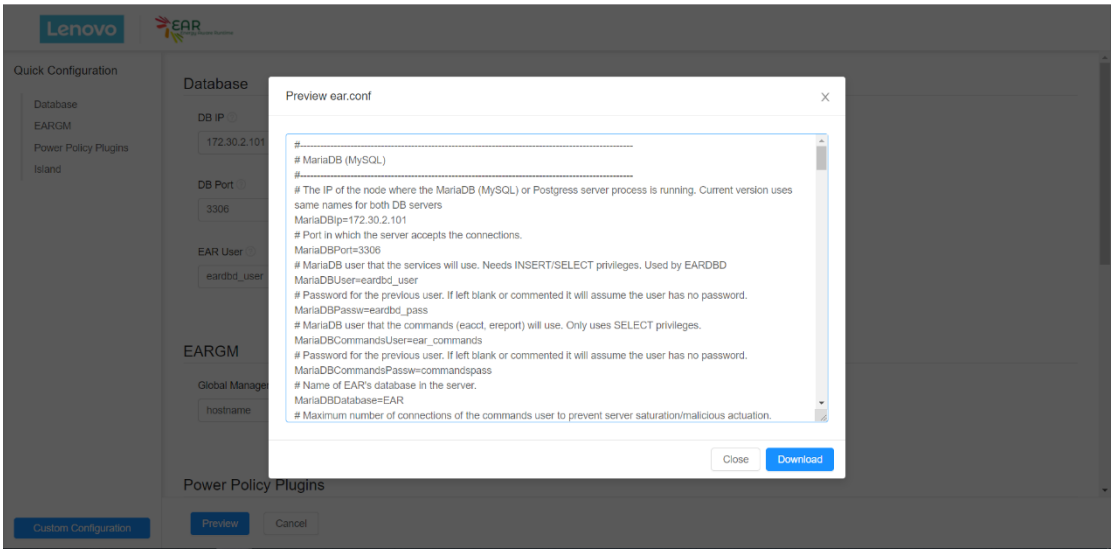
2. On the popup file browser dialog, select an existed EAR configuration file on your local storage:



3. Modify the setting values on the configuration page, by default it is in custom mode. If you want to discard all the changes, you can click the **Reset** on the bottom of the page:



4. After you finish the configuration, click the **Preview** on the bottom of the page. You can view the content of the ear.conf on the popup dialog:



5. You can click **Download** to save the content to your local storage.

4 Appendix

4.1 Help Messages

Database

Item	Message
DB IP	The IP of the node where the MariaDB (MySQL) or Postgress server process is running. Current version uses same names for both DB servers.

DB Port	Port in which the server accepts the connections.
DB User	MariaDB user that the services will use. Needs INSERT/SELECT privileges. Used by EARDBD.
DB Password	Password for the previous user. If left blank or commented it will assume the user has no password.
DB Commands User	MariaDB user that the commands (eacct, ereport) will use. Only uses SELECT privileges.
DB Commands Password	Password for the previous user. If left blank or commented it will assume the user has no password.
DB Database	Name of EAR's database in the server.
Max Connections	Maximum number of connections of the commands user to prevent server saturation/malicious actuation. Applies to MariaDBCommandsUser.
Report Node Detail	Extended node information reported to database (added: temperature and avg_freq in power monitoring).
Report Sig Detail	Extended signature hardware counters reported to database.
Report Loops	Set to 1 if you want Loop signatures to be reported to database.

EARD

Item	Message
Daemon Port	The port where the EARD will be listening.
Power Monitoring Freqency	Frequency used by power monitoring service, in seconds.
Max Pstate	Maximum supported frequency (1 means nominal, no turbo).
Turbo Frequency	Enable (1) or disable (0) the turbo frequency.
Use DB	Enables the use of the database.
Use EAR DBD	Inserts data to MySQL by sending that data to the EARDBD (1) or directly (0).
Force Frequencies	'1' means EAR is controlling frequencies at all times (targeted to production systems) and 0 means EAR will not change the frequencies when users are not using EAR library (targeted to benchmarking systems).
Verbose	The verbosity level [0..4].
Use Log	When set to 1, the output is saved in '$EAR_TMP'/eard.log (common configuration) as a log file.Otherwsie, stderr is used.
Min Time Performance Accuracy	Minimum time between two energy readings for performance accuracy.

EARDBD

Item	Message
Daemon Port	Port where the EARDBD server is listening.
Mirrow Daemon Port	Port where the EARDBD mirror is listening.
Synchronize Port	Port is used to synchronize the server and mirror.
Aggregation Time	In seconds, interval of time of accumulating data to generate an energy aggregation.
Insertion Time	In seconds, time between inserts of the buffered data.
Memory Size	Memory allocated per process. This allocations is used for buffering the data sent to the database by EARD or other components. If there is a server and mirror in a node a double of that value will be allocated. It is expressed in MegaBytes.
Memory Size Per Type	The percentage of the memory buffer used by the previous field, by each type. These types are: mpi, non-mpi and learning applications, loops, energy metrics and aggregations and events, in that order. If a type gets 0% of space, this metric is discarded and not saved into the database.
Use Log	When set to 1, eardbd uses a '$EAR_TMP'/eardbd.log file as a log file.

EARL

Item	Message
Coefficients Directory	Path where coefficients are installed, usually $EAR_ETC/ear/coeffs.
DynAIS Levels	Number of levels used by DynAIS algorithm.
DynAIS Window Size	Windows size used by DynAIS, the higher the size the higher the overhead.
Dynais Timeout	Maximum time in seconds that EAR will wait until a signature is computed. After this value, if no signature is computed, EAR will go to periodic mode.
Library Period	Time in seconds to compute every application signature when the EAR goes to periodic mode.
Check EAR Mode Period	Number of MPI calls whether EAR must go to periodic mode or not.

EARGM

Item	Message
Global Manager Host	The IP or hostname of the node where the EARGMD demon is running.
Global Manager Port	Port where EARGMD will be listening.

Global Manager Use Aggregated	Use '1' or not '0' aggregated metrics to compute total energy.
Global Manager Period T1	Period T1 and period T2 are specified in seconds. T1 must be less than T2. Global manager updates the information every T1 seconds and uses the energy/power in T2 period to estimate energy/power constraints.
Global Manager Period T2	
Global Manager Units	Units field, Can be '-' (Joules), 'K' KiloJoules or 'M' MegaJoules.
Global Manager Energy Limit	This limit means the maximum energy allowed in 259200 seconds in 550000 KJoules.
Global Manager Mode	Global manager modes. Two modes are supported '0' (manual) or '1' (automatic). Manual means Gobal Manager is only monitoring energy&power and reporting to the DB . Automatic means it takes actions to guarantee energy limits.
Global Manager Mail	A mail can be sent reporting the warning level (and the action taken in automatic mode). 'nomail' means no mail is sent. This option is independent of the node.
Global Manager Warnings Percentage	Percentage of accumulated energy to start the warning DEFCON level L4, L3 and L2.
Global Manager Grace Periods	Number of 'grace' T1 periods before doing a new re-evaluation. After a warning, EARGM will wait T1xGlobalManagerGracePeriods seconds until it raises a new warning.
Global Manager Verbose	Verbose level.
Global Manager Use Log	When set to 1, the output is saved in '$EAR_TMP'/eargmd.log (common configuration) as a log file.

Common

Item	Message
Verbose	Default verbose level.
TMP Directory	Path used for communication files, shared memory, etc. It must be PRIVATE per compute node and with read/write permissions. $EAR_TMP.
ETC Directory	Path where coefficients and configuration are stored. It must be readable in all compute nodes. $EAR_ETC.
Inst Directory	
Database Path	Path where metrics are generated in text files when no database is installed. A suffix is included.

Power Policy Plugins

Item	Message
Default Power Policy	Policy names must be exactly file names for policies installed in the system at /path/to/inst/lib/plugins/policies (without the extension .so).
Custom Policy	Example of the definition of 3 policies with different configurations: It must be included policy name, default frequency, privileged means whether normal users can execute or nor. Settings depends on policy.

Other Plugins

Item	Message
Plugin Energy	Energy reading plugin (without the extension). Allows to use different system components to read the energy of the node. In this case, this plugin reads the energy of the system using Intel Node Manager. look at /path/to/inst/lib/plugins/energy folder to see the list of installed energy plugins.
Plugin Power Model	Power model plugin (without the extension). The power model plugin is used to predict the power and energy consumption of the next iteration of the executing application.

Security

Item	Message
Authorized Users	Authorized users that are allowed to change policies, thresholds and frequencies are supposed to be administrators. A list of users, Linux groups, and/or SLURM accounts can be provided to allow normal users to perform that actions. Only normal Authorized users can execute the learning phase.
Authorized Accounts	
Authorized Groups	
Energy Tag	Energy tags are pre-defined configurations for some applications (EAR library is not loaded). This energy tags accept a user ids, groups and SLURM accounts of users allowed to use that tag.

Special Nodes

Describes nodes with some special characteristic such as different default P_STATEs, default coefficients file and/or policy thresholds.

Island

This section is mandatory since it is used for cluster description. Normally nodes are grouped in islands that share the same hardware characteristics as well as its database managers

(EARDBDS). Each line describes an island, and every node must be in an island.

Remember that there are two kinds of database daemons. One called 'server' and other one called 'mirror'. Both performs the metrics buffering process, but just one performs the insert. The mirror will do that insert in case the 'server' process crashes or the node fails.

It is recommended for all islands to have symmetry. For example, if the island I0 and I1 have the server N0 and the mirror N1, the next island would have to point the same N0 and N1 or point to new ones N2 and N3.

Multiple EARDBDs are supported in the same island, so more than one line per island is required, but the condition of symmetry have to be met.

It is recommended that for a island to the server and the mirror running in different nodes. However, the EARDBD program could be both server and mirror at the same time. This means that the islands I0 and I1 could have the N0 server and the N2 mirror, and the islands I2 and I3 the N2 server and N0 mirror, fulfilling the symmetry requirements.

The min_power, max_power and max_temp are threshold values that determine if the metrics read might be invalid, and a warning message to syslog will be reported if the values are outside of said thresholds. error_power is a more extreme value that if a metric surpasses it, said metric will not be reported to database.