Software Requirements Specification for Minimization Analysis: A program for distributing computational power for data centers

Ning Wang

February 8, 2023

Contents

Revision History

Date	Version	Notes
February 8, 2023	1.0	Initial Release

1 Reference Material

This section records information for easy reference.

1.1 Table of Units

Throughout this document SI (Système International d'Unités) is employed as the unit system. In addition to the basic units, several derived units are used as described below. For each unit, the symbol is given followed by a description of the unit and the SI name.

symbol	unit	SI
m	length	metre
kWh	energy	kilowatt
m^3	volume	cubic meter
s	time	second
°C	temperature	centigrade
HM/s	computational power	hashrate
W	power	watt ($W = J\,s^{-1}$)

Additionally, currency for the total cost, is labelled by unit "CAD" which define Canadian dollars.

1.2 Table of Symbols

The table that follows summarizes the symbols used in this document along with their units. The choice of symbols was made to be consistent with existing documentation for electricity distribution systems. The symbols are listed in alphabetical order.

symbol	unit	description
R_{total}	m^3	carbon dioxide emission
L_{total}	W	total power demanded by the given combination plan
R_{total}	W	total electricity supply by the renewable energy
P_{total}	W	total electricity supply by the distributed energy

1.3 Abbreviations and Acronyms

symbol	description
Ctotal	total cost
Cr	renewable cost
Cp	grid electricity cost
di	Distance between data center and power station
Ai	Computational power for data center
IM	Instance Model
LC	Likely Change
PS	Physical System Description
R	Requirement
SRS	Software Requirements Specification

2 Introduction

The COVID-19 pandemic has clearly shown that rural hospitals are a weak link in regional healthcare systems. They lack the advanced digital services, human and financial resources, and integration with urban hospitals that are needed to provide excellent health care uniformly across a provincial region. Data and Computing Facilities (DCFs) are established in rural hospitals. To meet the challenge of COVID-19 and other public health crises, this model was developed aiming at solving part of the rural hospital's problems. Data centers use about 1%-4% of the electricity in the world[(1)]. Using energy more efficiently is one of the fastest, most cost-effective ways to save money, reduce greenhouse gas emissions, create jobs, and meet growing energy demand.

2.1 Purpose of Document

The purpose of this document is to provide a analytical model of allocate the computing power to each data center with minimize the cost to intended Reader (Section 4.2.1). This document should clearly communicate all necessary background information and detailed system context is described in Section 3.

2.2 Scope of Requirements

The domain of this problem is restricted to given renewable pricing and factory electricity will lose proportional over transmission . In particular, the different areas will have different electricity pricing and different types of renewable resources. However, we can set the constant renewable pricing and minimize the total cost to the final the optimal computing distributed plan for each data center. The domain to one specific type of input data which is distance between data center and the power station in this program, specifically, we will find out the optimization cases of combination under static pricing. Therefore, it has been decided to scope distances as input. Subsequently, since this project is limited to any renewable energy with a given price, any transmission power loss of renewable energy is considered out of scope. Essentially, while this program is intended to eventually be used within the context of specific physical data, it is developed and constructed for application to another area in North America as well.

2.3 Characteristics of Intended Reader

This document assumes the intended reader has familiarity with basic optimization analysis, power system analysis, and electricity distribution analysis. Courses which contribute to background knowledge may be titled Linear Systems (undergrad), Electric Machines and Power Systems (undergrad), Electric Power Transmission, Distribution and Utilization (undergrad), Sustainable Electrical Power Systems (Graduate), and Real Analysis (Graduate). Data constrain and model limitation will be discussed in this document. However, our exposition will only cover the concepts needed for our purposes. For proofs and for a complete exposition of all background materials. Furthermore, this document will explain as entry as possible aiming to appliable for more readers.

2.4 Organization of Document

This document is built on the template recommendations in cas 741 gitlab that seeks to standardize communication tools for software development. The suggested order for reading this SRS document is: Goal Statement, Instance Models, Requirements, Introduction, and Specific System Description.

3 General System Description

This section provides general information about the system. It identifies the interfaces between the system and its environment, describes the user characteristics and lists the system constraints.

3.1 System Context

The following figure depicts a system context view of the computing relocation model.

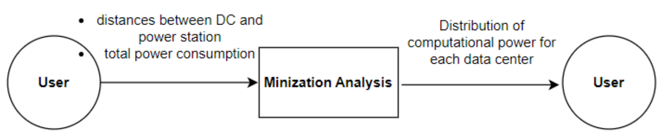

Figure 1: System Context

For each of the entities in the system context diagram its responsibilities should be listed. Whenever possible the system should check for data quality, but for some cases the user will need to assume that responsibility. The list of responsibilites should be about the inputs and outputs only, and they should be abstract. Details should not be presented here. However, the information should not be so abstract as to just say “inputs” and “outputs”. A summarizing phrase can be used to characterize the inputs. For instance, saying “material properties” provides some information, but it stays away from the detail of listing every required property.

- User Responsibilities:
  - keep the unit even, keep every input with SI standard
  - Provide the distances between the power station and data centers and keep input under constraint.

- Change system setting with their unique situations included total power consumption, renewable power rate and grid power rate.

- Minimization AnalysisMinimization Analysis Responsibilities:
  - Keep positive input and less than max limit
  - Detect if the output is capable for a data center to carry
  - Detect data type mismatch, such as relocate data under same unit

3.2 User Characteristics

The front user should keep the unit with SI standard. The end user should have basic knowledge of undergraduate math and physics.

3.3 System Constraints

The method developed in this project is expected to be independent of system constraints at this time.

4 Specific System Description

This section first presents the problem description, which gives a high-level view of the problem to be solved. This is followed by the solution characteristics specification, which presents the assumptions, theories, definitions and finally the instance models. Add any project-specific details that are relevant for the section overview.

4.1 Problem Description

Minimization AnalysisThis model is intended to estimate the minimum cost with different computing distributions of data centers with different distances from electricity power station

4.1.1 Terminology and Definitions

Minimization problem: The function to minimize is called the objective function. The minimum value of the objective function is in the margins of the feasible area delimited by the restrictions of the problem. This value is called the ideal value.

- Construct a graph for feasible region
- find a feasible region and get the optimized combinations of computing distributions

4.1.2 Physical System Description

Physics system will regarding to real world. HVDC transmission losses are quoted as less than 3% per 1,000 km[(2)], a formula was designed to calculate the total power lose along the transmission line. With various of distance as input, the system operation like a net flowing. Each data center will consuming electricity from both renewable resources and grid power station. In the model, only transmission along electricity grid will assume have power lose, renewable resource stations are assuming very close to the data centers the supplied.

The physical system of Minimization Analysis, as shown in Figure2, includes the following elements:

- Data centers are not uniformed in the area.
- Renewable resources stations are very close to the data centers they supplied.

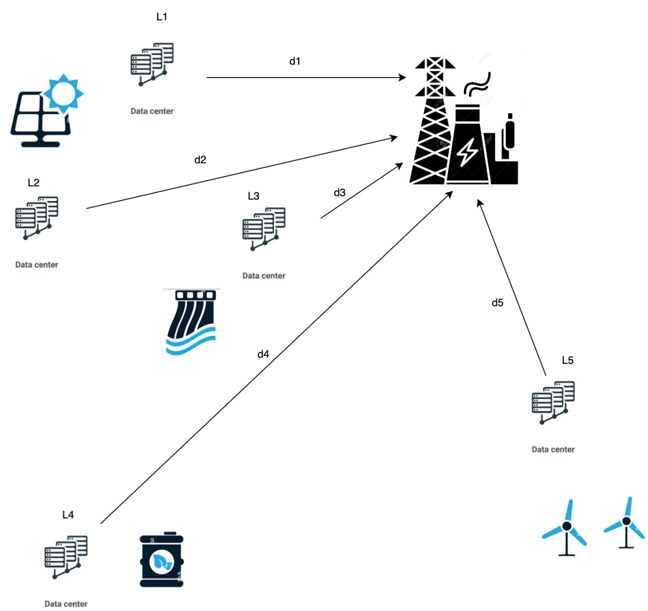

Figure 2: System Context

4.1.3 Goal Statements

Give a bunch of distances between data centers and electric power station as input, total power consumption was set, solving minimization problems:

GS1: This model is intended to estimate the different computing distributions of data centers.

GS2: This model is intended to output final solution under the minimum cost.

GS3: This model possible to output the CO2 emission for reference.

4.2 Solution Characteristics Specification

This section characterizes the attributes of an acceptable solution. Both analysts and stakeholders should agree on these attributes so that the solution can be accepted when the project is complete.

4.2.1 Assumptions

This section simplifies the original problem and helps in developing the theoretical model by filling in the missing information for the physical system. The numbers given in the square brackets refer to the theoretical model [T], general definition [GD], data definition [DD], instance model [IM], or likely change [LC], in which the respective assumption is used.

A1: We know that distances as input should be positive and under reachable constrain.

A2: The software will estimate the distribution mof computing power from a constant number relocated.

A3: The scope of minimization analysis is to compute the optimal plan that also most cost efficiency.

A4: The scope of set power lose with a constant coefficient is to get rid of substations power lose effect.

A5: Depends on the pricing with C_r and C_p, the L_i will distribute with descending or ascending order of di

4.2.2 Theoretical Models

Applying the terminology and definitions from 4.1.1, this section records theorems required to identify a minimum optimal solution.

Consider the total cost for all data centers this is the sum of the total grid power supply cost plus the total renewable supply cost. The computing power distribution will as the same proportion as their electricity consumption[(3)].

RefName: **T1**

Label:

Equation: TM-series-cauchy-condition

Description:

Notes: Minimize the cost

Source: $C_{\text{total}}=\sum_{i=1}^{n} Cri * Ri + CPi * Pi$

Ref. By: The above equation gives the value of the total cost for given Cr and Cp with power consumption as Ri and Pi respectively. When solving a minimizing problem, some constrain of variables must apply to find the feasible region.

Preconditions for T1: None.

Derivation for T1: Not Applicable

https://www.superprof.co.uk/resources/academic/maths/linear-algebra/linear-programming/steps-to-solve-a-linear-programming-problem.html GD2 None

RefName: **T2**

Label: Calculate for Pi

Equation: Pi=Pi-theoretical /(1-(0.03 di / 1000))

Description: The above equation gives the value of power lose with transmission distance varied. The Pi is the actual consumption of energy, Pi-theoretical is the consumption without transmission lose. HVDC transmission losses are quoted as less than 3% per 1,000 km

Notes: None.

Source: https://www.superprof.co.uk/resources/academic/maths/linear-algebra/linear-programming/steps-to-solve-a-linear-programming-problem.html

Ref. By: GD1

Preconditions for T2: None

Derivation for T2: Not Applicable

RefName: **T3**

Label: The power consumption for each data center

Equation: $L_{\text{total}}=\sum_{i=1}^{n} \mathrm{L}_i = R_i + P_i$

Description: The power consumption for each data center will come up with renewable supply and grid power supply.

Notes: None.

Source: https://www.superprof.co.uk/resources/academic/maths/linear-algebra/linear-programming/steps-to-solve-a-linear-programming-problem.html

Ref. By: GD1

Preconditions for T3: None

Derivation for T3: Not Applicable

"Ref. By" is used repeatedly with the different types of information. This stands for Referenced By. It means that the models, definitions and assumptions listed reference the current model, definition or assumption. This information is given for traceability. Ref. By provides a pointer in the opposite direction to what we commonly do. You still need to have a reference in the other direction pointing to the current model, definition or assumption. As an example, if T1 is referenced by G2, that means that G2 will explicitly include a reference to T1.

4.2.3 General Definitions

General Definitions (GDs) are a refinement of one or more TMs, and/or of other GDs. The GDs are less abstract than the TMs. Generally the reduction in abstraction is possible through invoking referencing Assumptions. This section collects the laws and equations that will be used in building the instance models.

Number	GD1
Label	**Power losses along transmission line**
SI Units	kWh
Equation	see description
Description	Transmission and distribution (T&D) loss are amounts that are not paid for by users. T&D Losses= (Energy Input to feeder (Kwh)-Billed Energy to Consumer (Kwh)) / Energy Input kwh x100.
Source	https://www.electricalindia.in/losses-in-distribution-transmission-lines/#:~:text=Transmission%20and%20distribution%20(T%26D,))%20%2F%20Energy%20Input%20kwh%20x100.
Ref. By	T2

Detailed derivation of calculating power transmission losses

It is fact that the unit of electric energy generated by the power station does not match the units distributed to the consumers. Some percentage of the units is lost in the distribution network. This difference in the generated & distributed units is known as transmission and distribution losses. Distribution Sector is considered the weakest link in the entire power sector. Transmission losses are approximately 17% while distribution losses are approximately 50%. In this model, we assume all of the data centers pass through the same number of sectors, so only transmission line losses are calculated.

4.2.4 Data Definitions

The Data Definitions are definitions of symbols and equations that are given for the problem. They are not derived; they are simply used by other models. For instance, if a problem depends on density, there may be a data definition for the equation defining density. The DDs are given information that you can use in your other modules.

All Data Definitions should be used (referenced) by at least one other model.

This section collects and defines all the data needed to build the instance models. The dimension of each quantity is also given. Modify the examples below for your problem, and add additional definitions as appropriate.

Number	DD1
Label	**Level of computing power for each data center**
Symbol	A_p
SI Units	MH/s
Equation	None
Description	Hashrate is a measure of the computational power per second used when mining. More simply, it is the speed of mining. It is measured in units of hash/second, meaning how many calculations per second can be performed. Machines with a high hash power are highly efficient and can process a lot of data in a single second. In the case of Bitcoin, the hashrate indicates the number of times hash values are calculated for PoW every second. Hashrate is usually measured in units of k (kilo, 1,000), M (mega, 1 million), G (giga, 1 billion), or T (tera, 1 trillion). For example, 1 Mhash/s indicates 1 million hash calculations are done every second.
Sources	https://bitflyer.com/en-eu/s/glossary/hashrate
Ref. By	IM1

4.2.5 Instance Models

The motivation for this section is to reduce the problem defined in "Physical System Description" (Section 4.1.2) to one expressed in mathematical terms. The IMs are built by refining the TMs and/or GDs. This section should remain abstract. The SRS should specify the requirements without considering the implementation.

This section transforms the problem defined in Section 4.1 into one which is expressed in mathematical terms. It uses concrete symbols defined in Section 4.2.4 to replace the abstract symbols in the models identified in Sections **??** and 4.2.3.

The goals reference your goals are solved by reference your instance models. other details, with cross-references where appropriate. Modify the examples below for your problem and add additional models as appropriate.

Number	IM1
Label	**electricity consumption distribution to find Computational power distribution**
Input	L_i, A_{total}, L_{total} The input is constrained so that all variables are non-negative
Output	A_1, $A_2 ... A_n$ $L_{\text{total}}/L_i = A_{\text{total}}/A_i$
Description	$L_i$is the power consumption by each data center. L_{total} is the total power consumption for data centers in this model. A_i is the Computational power for each data center. A_{total} is the total Computational power for data center in this model
Sources	https://www.iea.org/reports/data-centres-and-data-transmission-networks
Ref. By	DD1

4.2.6 Input Data Constraints

Table 1 shows the data constraints on the input output variables. The column for physical constraints gives the physical limitations on the range of values that can be taken by the variable. The column for software constraints restricts the range of inputs to reasonable values. The software constraints will be helpful in the design stage for picking suitable algorithms. The constraints are conservative, to give the user of the model the flexibility to experiment with unusual situations. The column of typical values is intended to provide a feel for a common scenario. The uncertainty column provides an estimate of the confidence with which the physical quantities can be measured. This information would be part of the input if one were performing an uncertainty quantification exercise.

The specification parameters in Table 1 are listed in Table 2.

Table 1: Input Variables

Var	Physical Constraints	Software Constraints	Typical Value	Uncertainty
di	$di > 0$	$0 \leq di \leq \text{d}_{\text{max}}$	1000 m	10%
L_{total}	L_{total} ¿ 0	$0 \leq \text{L}_{\text{total}} \leq \text{L}_{\text{max}}$	10 MWh	10%

(*) the system able to convert MWh to KWh

Table 2: Specification Parameter Values

Var	**Value**
C_r	TBD
C_p	TBD

4.2.7 Properties of a Correct Solution

A correct solution must exhibit fill in the details. These properties are in addition to the stated requirements. There is no need to repeat the requirements here. These additional properties may not exist for every problem. Examples include conservation laws (like conservation of energy or mass) and known constraints on outputs, which are usually summarized in tabular form. A sample table is shown in Table 3

Table 3: Output Variables

Var	**Physical Constraints**
A_i	Under data center computational power capable size

5 Requirements

The requirements refine the goal statement. They will make heavy use of references to the instance models.

This section provides the functional requirements, the business tasks that the software is expected to complete, and the nonfunctional requirements, the qualities that the software is expected to exhibit.

5.1 Functional Requirements

R1: Requirements for the inputs that are supplied by the user. This information has to be explicit.

R2: The program shall notify user if an input is out of bounds.

R3: User should keep input validate data type.

R4: The program should able to converting units while minimizing cost.

R5: The program should keep all decision variables under their constrains.

R6: The program output the optimal plan for end users.

R7: The output should allocate the computing power of data centers with their capable size.

5.2 Nonfunctional Requirements

List your nonfunctional requirements. You may consider using a fit criterion to make them verifiable. The goal is for the nonfunctional requirements to be unambiguous, abstract and verifiable. This isn't easy to show succinctly, so a good strategy may be to give a "high level" view of the requirement, but allow for the details to be covered in the Verification and Validation document. An absolute requirement on a quality of the system is rarely needed. For instance, an accuracy of 0.0101 % is likely fine, even if the requirement is for 0.01 % accuracy. Therefore, the emphasis will often be more on describing now well the quality is achieved, through experimentation, and possibly theory, rather than meeting some bar that was defined a priori. You do not need an entry for correctness in your NFRs. The purpose of the SRS is to record the requirements that need to be satisfied for correctness. Any statement of correctness would just be redundant. Rather than discuss correctness, you can characterize how far away from the correct (true) solution you are allowed to be. This is discussed under accuracy.

NFR1: **Accuracy** The accuracy of the computed solutions should meet the level needed for engineering application.

NFR2: **Usability** The program shall not have a user interface but will clearly shows the output and data will easy to copy and read.

NFR3: **Maintainability** The time complexity of this program should be O(n).

NFR4: **Portability** The program should be easily integrated with another software program.

6 Likely Changes

LC1: The likely changes for the program should be adding more renewable supply options.

LC2: The program operating the analysis under minimum cost, it may have potential process analysis under minimum CO2 emission.

LC3: The likely changes for the program should be more flexible for more operating systems.

7 Unlikely Changes

LC4 :The basic design concept of allocate computing instead of power station will unlikely to change.

8 Traceability Matrices and Graphs

The purpose of the traceability matrices is to provide easy references on what has to be additionally modified if a certain component is changed. Every time a component is changed, the items in the column of that component that are marked with an "X" may have to be modified as well. Table 4 shows the dependencies of theoretical models, general definitions, data definitions, and instance models with each other. Table 5 shows the dependencies of theoretical models, general definitions, data definitions, instance models, and likely changes on the assumptions.

	T1	T2	T3	GD1	DD1	IM1
T1		x	x			
T2					x	
T3					x	
GD1		x				
DD1			x			
IM1			x			

Table 4: Traceability Matrix Showing the Connections Between Items of Different Sections

	A1	A2	A3	A4	A5
T1					
T2	x		x		
T3				x	x
GD1		x			
DD1		x			
IM1					

Table 5: Traceability Matrix Showing the Connections Between Assumptions

9 Values of Auxiliary Constants

This section contains nothing at this moment for minimization analysis.

References

[1] Michel Rivier, Ignacio J Pérez-Arriaga, and Luis Olmos. Electricity transmission. *Regulation of the power sector*, pages 251–340, 2013.

[2] Juan Rosellón. Different approaches towards electricity transmission expansion. *Review of network economics*, 2(3), 2003.

[3] Richard P Sedano and Matthew H Brown. Electricity transmission: a primer. *National Council on Electricity Policy, June, Washington, DC*, 2004.