Data Center Scheduling Optimization Problem

Data Science Track

Prepared by

Farid Ahmed Farid
Mohamed Ahmed Abdo Elhamamsy
Mohamed Abdelghani Mohamed

Under supervision of:

Dr. Basma Mostafa

Contents

Business Problem

As a cloud provider organization, there is a necessity to set up a schedule for turning the hardware on or off to minimize power and cooling cost, in addition to minimizing greenhouse gas emission which leads to a fundamental tradeoff between performance and energy usage while supplying a variant demand over time.

Business Objectives

- **Who:** IT solution Architects, Finance department, R&D department.
- **What:** Put a switch on/off schedule for the data center racks.
- **Why:** Minimizing the power and cooling cost in addition to greenhouse gas emission cost

---

Decision Variable:

x_{jk}: The contribution of the k_{th} cluster in the j_{th} period.

Note that an assumption that each rack must be OFF for two periods a day Was made.

Supply	X1	X2	X3	X4	X5	X6
T1	x_{11}	0	0	x_{14}	x_{15}	x_{16}
T2	x_{21}	x_{22}	0	0	x_{25}	x_{26}
T3	x_{31}	x_{32}	x_{33}	0	0	x_{36}
T4	x_{41}	x_{42}	x_{43}	x_{44}	0	0
T5	0	x_{52}	x_{53}	x_{54}	x_{55}	0
T6	0	0	x_{63}	x_{64}	x_{65}	x_{66}

Constraints:

1. Demand Constraints

The total contribution of all racks at each period must be equal to the demand in the same period.

Periods	Demand GB/4hrs
T1	30000
T2	30000
T3	40000
T4	40000
T5	35000
T6	35000

- $\sum_{i=1}^{6} x_{1i} * p_i = 30000$
- $\sum_{i=2}^{6} x_{2i} * p_i = 30000$
- $\sum_{i=3}^{6} x_{3i} * p_i = 40000$
- $\sum_{i=4}^{6} x_{4i} * p_i = 40000$
- $\sum_{i=5}^{6} x_{5i} * p_i = 35000$
- $\sum_{i=6}^{6} x_{6i} * p_i = 35000$

---

2. Max Supply Per Rack Constraints

The total contribution of a given rack throughout the period must be less than or equal to the max supply the rack can contribute with, in any given period.

	C1	C2	C3	C4	C5	C6
Max Supply GB/period	7500	7500	5000	10000	10000	15000

- $\sum_{j=1}^{6} \sum_{i=1}^{6} x_{ij} <= 4$

---

3. Binary Constraint

$x_{jk} = \{0,1\}, \ \forall \ j,k \in \{1,2,3,4,5,6\}$

Objective Function:

- $\sum_{k=1}^{6} \sum_{j=1}^{6} x_{jk} * c_k$ (Minimize)

  Where C_k is the combined cost of the k^{th} rack per GB per period.

Cost+CO2	C1	C2	C3	C4	C5	C6
On Cost + CO2 $/GB/Period	29.475	29.475	14.3	50.2	57	106.2

Classification with respect to the Optimization model

- **Continuous Optimization**

An assumption was made that any x_{jk} can take 0 or 1.

- **Constrained Optimization**

The optimization model is constrained by the imposed business constraints.

- **One Objective**

The model has one and only one objective which is to minimize the total combination cost considering the cost of providing the service plus the equivalent CO2 emissions costs.

- **Deterministic Optimization**

The model uses a Deterministic approach to find the optimal solution, in Other words, the objective function is convex leading the search Algorithm to reach a global optimum.

Assumptions & Approximations

- Any X_{JK} variables can take 0 or 1.
- Demand is covered within six periods throughout the day, four hours each, for instance, T1 period demand is 7500 GB/4hr.
- Each rack must be OFF for at least two periods a day.
- Demand is the transfer rate in MB/S.
- Each rack consists of a different number of devices, for instance routers, servers, and firewalls.
- Each rack can serve a different max transfer rate, for example, c1 can serve up to 100MB/s.

One $/GB costs differently for each rack, for instance, c1 rock has 15.225 $/GB.

- Each rack contributes differently into the Co2 emission cost, for instance C3 costs 10.3 $/GB for 6gm Co2 emission.

Racks Specs

Rack 1, 2 (7500GB/4h):
Rack 1 can serve 100MB/s,
6GB/m.
2 UPS systems
router
firewall
2 switches
30 workstations

Rack 3 (7500GB/4h):
1 UPS systems
router
firewall
1 switch
15 workstations

Rack 4 (10000GB/4h):
3 UPS systems
2 routers
firewall
3 switches
45 workstations

Rack 5 (10000GB/4h):
4 UPS systems
2 routers
firewall
4 switches
60 workstations

Rack 6 (15000GB/4h):

5 UPS systems

2 routers

firewall

5 switches

75 workstations

System Implementation using JuMP/Julia and Excel

- Source Code is available in another file with the name Cloud Provider.

Solution of the model using Julia

```
 Schedule:

[1.0, 1.0, 1.0, 1.0, 0.0, 0.0]
[0.0, 0.0, 1.0, 0.0, 1.0, 1.0]
[1.0, 1.0, 0.0, 1.0, 0.0, 1.0]
[1.0, 1.0, 1.0, 1.0, 1.0, 0.0]
[0.0, 0.0, 0.0, 1.0, 1.0, 1.0]
[1.0, 1.0, 1.0, 0.0, 0.0, 1.0]

-> What is the minimum cost?
1089.6000000000001

julia>
```

Solution of the model using Excel

			Supply	X1	X2	X3	X4	X5	X6	Total		
			T1	1	1	1	1	0	0	30000	=	30000
			T2	0	0	1	0	1	1	30000	=	30000
			T3	0	0	1	1	1	1	40000	=	40000
			T4	1	1	0	1	0	1	40000	=	40000
			T5	1	1	0	1	1	0	35000	=	35000
			T6	1	1	1	0	0	1	35000	=	35000
			Total	4	4	4	4	3	4			
				^	^	^	^	^	^			
				4	4	4	4	4	4			
Running Cost	1089.6											

Model Validation

The model has provided quite good results that make sense, for instance: -

- The total supply across each period is typically equal to the demand.
- The model has met all the constraints set, for example, the racks supply doesn't exceed its maximum threshold.
- The result of the running cost provided by the model according to the inputs provided is very reasonable.