Knowledge Based Life Cycle Management for Applications & Infrastructure Components in Scaled Hyper Converged Software Defined Data Center

DISSERTATION

By

Ravi Ranjan

2018AB04501

Under the supervision of

Ravi kottapalli

Staff Customer Adoption Solution Engineer

BIRLA INSTITUTE OF TECHNOLOGY & SCIENCE, PILANI
Work Integrated Learning Programmes Division
I SEMESTER 2020-21

DSE CL ZG628T **DISSERTATION**

Knowledge Based Life Cycle Management for Applications & Infrastructure Components in Scaled Hyper Converged Software Defined Data Center

Submitted in partial fulfilment of the requirements of the

M. Tech. Data Science and Engineering Degree programme

By

Ravi Ranjan
2018AB04501

Under the supervision of

Ravi kottapalli
Staff Customer Adoption Solution Engineer

BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE
PILANI (RAJASTHAN)

August, 2021

Acknowledgements

The satisfaction that accompany of successful completion of this task had made it possible with help of my BITS Class mates whose constantly given guidance and efforts to success. I am very thankful and express gratitude with respect to all those who helped me in the completion of this project.

Very thankful to **Mr. Ravi Kottapalli**, Staff Customer Adoption Solution Engineer, **VMware India Pvt. Ltd**., Bangalore for many all discussions on design regularly meeting over zoom and slack.

I would also like to thank VMware India Pvt. Ltd, Bangalore for providing all the necessary infrastructure, software for accomplishing this project and a special thanks to Birla Institute of Technology for opening a door to completion of my Post- Graduation.

Ravi Ranjan

BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE, PILANI

CERTIFICATE

This is to certify that the Dissertation entitled **Knowledge Based Life Cycle Management for Applications & Infrastructure Components in Scaled Hyper Converged Software Defined Data Center**

and submitted by Mr./Ms. **Ravi Ranjan** ID No. 2018AB04501

in partial fulfilment of the requirements of DSE CL ZG628T Dissertation, embodies the work

done by him/her under my supervision.

Place: Bangalore

Signature of the Supervisor

Date: 4 August 2021

Abstract

Proposing a Knowledge Based Life Cycle Management method for the Infrastructure Components and the Applications in the Scaled Hyper Converged Software Defined Data Center is a project to manage life cycle Management (LCM) software which will look for "**Best Window time**" or "best time" to perform **LCM Operations**. So that the Administrator can perform LCM whenever upgrade or patch is available for any entity (Application or Infrastructure component) . This will help avoid any risk because of ongoing upgrades or patches.

The Machine Learning Predictive Technique that we will be using here is:

Analytic Engine which runs Machine Learning Technics. Predictive Machine Learning Technics runs against the Log Database which contains the details of each application usage with respect to time and data. The predictive data samples will be taken for a year date or more. More Historical data will give more accurate predictive results.

Signature of Supervisor

Signature of student

List of Symbols & Abbreviations used

SDDC - Software-Defined Data Center

LCM – Lifecycle Management

VCF – VMware cloud Foundation

WLD – Workload domain

E1, E2 – Entities 1, Entities 2

Mgr - Manager

OLA – Optimal LCM Agent

List of Tables

Table 1. OLA Agent Table.....12

Table 2. Entity LCM Table.....15

List of Figures

Fig 1. Idea overview..... 11
Fig 2. Architecture diagram.....13

Table of Contents

DISSERTATION *1*

Acknowledgements *3*

CERTIFICATE *4*

Abstract *5*

List of Symbols & Abbreviations used *6*

List of Tables *7*

List of Figures *8*

Chapter 1 ..... *10*

Fig 1 . Idea Overview **11**

Invention Theory and Operation **11**

Table 1. OLA Agent Table **12**

Machine Learning Technic - **13**

Fig 2. Architecture diagram **13**

Chapter 2 ..... *14*

Details Workflow: **14**

Novelty Aspects: **15**

Table 3. Entity LCM Table **15**

Conclusions / Recommendations *17*

Directions for future work *18*

Bibliography/References *19*

Appendices *20*

List of Publications/Conference Presentations *21*

Duly Completed Checklist *22*

Final Evaluation Sheet *23*

EVALUATION DETAILS *24*

Chapter 1

Now a days If any upgrade is available for any entity, it was just showing as “upgrade available”, but there is no mechanism exists today to propose a best time to trigger the available upgrade along with best proposal time to trigger the upgrade. It will be good and optimized mechanism if any entity shown with “available upgrade along with best time to apply the upgrade and Risk Analysis”. At present there is no sophisticated method or mechanism available to us in modern Software Defined Data center to handle this situation.

I am taking above issues or drawbacks as problem statement and coming up with a Method that will dynamically handle the situation in scaled environments and the proposing innovation that leveraging the capabilities of Machine Learning Techniques to predict the best window time or best possible time for any entity to go for LCM operation.

The innovation predominantly focusing on the following key points

1. Based on Administrator/Customer request, Predict the best possible time to go for an LCM operation for any given Entity in SDDC.

2. Dynamically identify the Entities that has “Available Upgrades” in SDDC and predict & propose a best possible time to trigger the Available Upgrades and also it will give the Risk Analysis with respect proposed time. Example – If Administrator triggers LCM with the ML proposed time, then how many users may feel performance impact or how many people may be users at the proposed time.

3. Dynamic Time based LCM - Dynamically associate the Application that is triggered by Administrator to LCM when Administrator chosen the proposed or predicted time suggested by this invention.

4. Adoptability of this innovation to VMware new architecture for VMware Cloud Foundation.

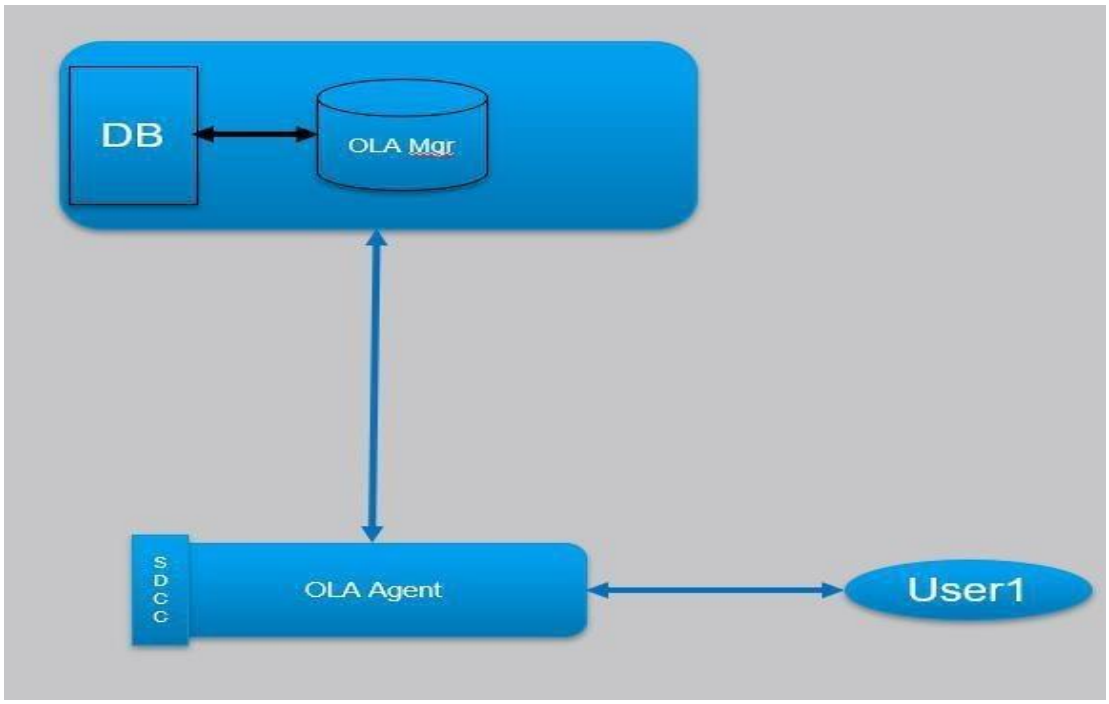


Fig 1 . Idea Overview

The above diagram represents a high-level summary of the invention. I request reviewers please have a look at the Invention Theory and Operations Section to get full essence of the Invention. Whenever User or Administrator requested for to find best optimal time for a particular Entity, the request will be proxied to the OLA Manager through OLA Agent. OLA Manager handover the request to the Analytics Engine which is part of the OLA Manager will run Machine Learning Technics on the Log Database to find the optimal time and return the results to Entity through OLA Manager and OLA Agent.

Invention Theory and Operation

Before Jump on to the Invention Theory and Operation, it is worth to focus on few terminologies that tied to this invention.

Entity – It can be an Application or an Infrastructure Component. VMware Infrastructure Components - are PSC, vCenter, NSX, vRLI, vROPS, vRA.

Applications – A software application that runs on the Software Defined data center Cluster - Example Email Server, Database or many more.

Optimal Window Time – Optimal Window time is the time interval or period in which Entity usage will be less in compared with the other intervals.

Optimal LCM Agent (OLA) – It is a software module sit in each Software Defined Data center Central Entity. As per VMware VCF is considered it will be placed in SDDC Manager. The main responsibility of this module is – to detect any upgrades available for any Entity in the Data center. The upgrade availability may be with manual or automatic.

For example – As soon as Administrator uploaded the image for Infrastructure Components like PSC, VC, NSX then these Entity statuses will show as "Upgrade Available". For other entities like Outlook, Database will get automatic upgrade availability option as and when new image or patch available for those from internet. OLA can detect the "Upgrade Availability" status for each Entity and inform to the OLA Manager which is placed in the Pantheon (Will explain more about the Pantheon in the upcoming section) Monitoring and Logging Module. If administrator wants to scope out any particular Entity from this OLA vicinity, he can create Exclude Entity and provide the list to OLA Agent. Those Exclude List agent will not considered any more for Optimal LCM.

Apart from the above functionality, OLA Agent will generate unique Id for each entity in SDDC and will be maintained in separate table called "OLA Agent Table".

Application	Unique Id for Application	SDDC Id
App1	aaaaa	XXYYCC
App2	bbbbb	XXYYCC
App3	ccccc	XXYYCC
APp4	ddddd	XXYYCC

Table 1. OLA Agent Table

This module also responsible for to carry Administrator requests to find the optimal window time for entity to OLA Manager which is sitting at the Pantheon Monitor and Logging module. When the OLA Manager return the best possible time to an Entity, that will be mediated or facilitated or communicated by OLA Agent to the Entity.

OLA Manager – It is a software module, it will control and coordinates the functions of OLA Agents sitting in the various SDDC Managers at SDDCs.

It receives the requests from the OLA Agents and process the requests for to find the Best Optimal Time for LCM operation to an entity. OLA Manager Southbound connectivity is with Analytical Platform that will have connectivity intern to the Database in which all the logs for all the data

center will be maintained. Machine Learning Techniques will be running in the OLA Manager Analytic engine. As soon as OLA Manager receives the request from OLA Agent, the request will he handover to Analytic Platform, there Machine Learning Algorithms will run on the Database based on the input query parameter given by OLA Manager. The result will be sent back to the OLA Manager and Manager intern proxy the request to respective OLA Agent. Detail work flow will be explained later.

Machine Learning Technic -

The Machine Learning Predictive Technic that we are using here is –

Whenever upgrade is available for an Entity or Administrator manually trigger an event for an Entity to find an optimal time for LCM, then OLA Agent send the request with Application Name, Application Id and Data center Id to the OLA Manager. OLA Manager process the request and handover the request to Analytic Engine which runs Machine Learning Technics. Predictive Machine Learning Technics runs against the Log Database which contains the details of each application usage with respect to time and data. The predictive data samples will be taken for a year date or more. More Historical data will give more accurate predictive results.

One more important aspect is exporting the logs and maintaining that in log server DB. Log Server DB will maintain Each Entity logs for each SDDC.

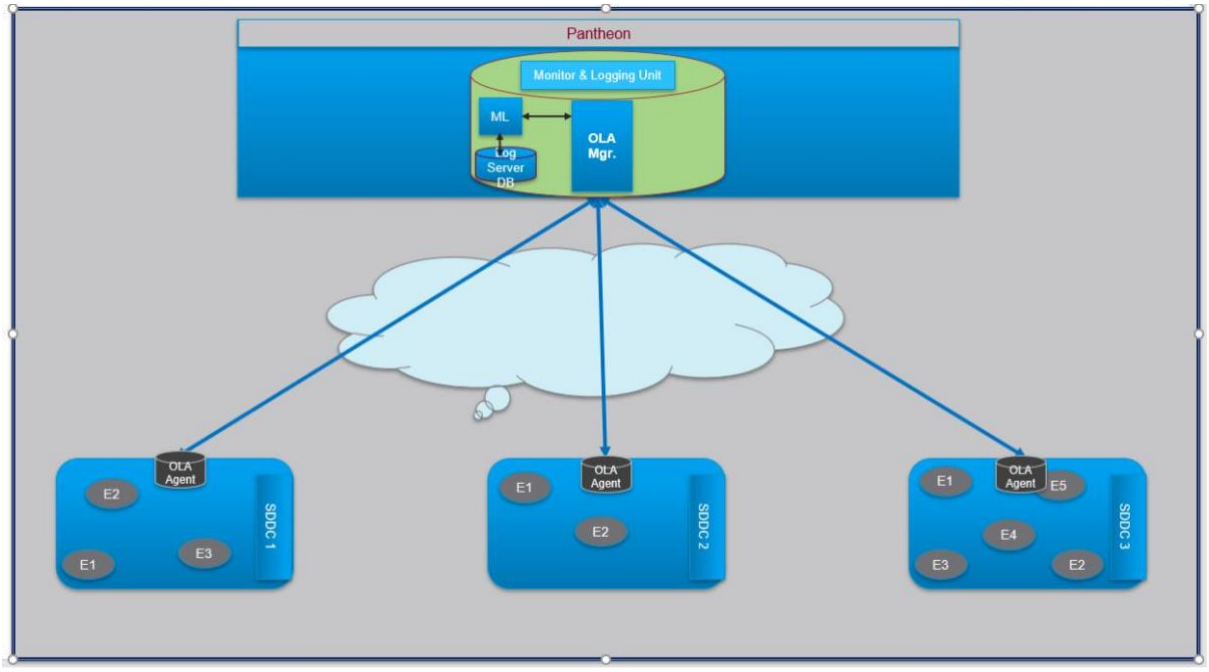


Fig 2. Architecture diagram

Chapter 2

Details Workflow:

1. For each application in the SDDC , OLA Agent will generate unique Id and associate with the application. This data will be maintained in the table called “OLA Agent Table”.

2. case 1 - Whenever any application gets “Upgrade Availability”, immediately OLA Agent form a request to the OLA Agent with Application Name, Application Id and SDDC Id.

3. This request will be proxied to the OLA Manager.

4. OLA Manager form a search string with the details of Application Id and search Period [Max 1 year] and handover the request to Machine Learning Module [Analytic Module].

5. With the inputs given by OLA Manager, ML Technic will run against the Logs Server data with search criteria. ML Technics will predict the best window time for the application to trigger LCM operation and also it will predict how may users may be users for that Application or Entity in that predicted Optimal Window time.

6. The ML will return the Optimal proposed time along with the Estimated Risk in the proposed time to the OLA Manager.

7. OLA Manager proxy the information to OLA agent and OLA agent will provide the details to the Administrator or user in the form of display text or this can be maintained in separate table called “Entity LCM Table” in the management node of SDDC.

8. Entity LCM table is nothing but a simple table hold the information of Application Id, Name, Proposed LCM Time, Risk in the Proposed time, like wise it will maintain for all the applications.

9. Case 2 – If there is no “Upgrade Available” for an entity but Administrator wish to find out the best time for an upgrade, then he can manually trigger an event that will set a Bit in the buffer, and Agent Proxy will read the bit and immediately follow the step 3 to 8.

Note : Implementor have an option to integrate the upgrade code based on the invention estimated time – meaning if user or administrator opted the invention or Agent Manager proposed time, then at the proposed time Upgrade will be automatically triggers without Administrator

intervention. This can be achieved if Administrator intergrade LCM code to the “Entity LCM Table”. This Entity LCM Table has an option even to trigger the LCM based on Administrator choice.

Apart from this there is another important point associated with the invention that is ..At any point of time, Administrator Risk Associated information for his chosen time for LCM.

Example – If Administrator wants to trigger the Upgrade with his chosen time, then before triggering Upgrade for an application, he can generate a request to the OLA Agent and Manager to get the Risk Associated information with the Administrator chosen time.

Novelty Aspects:

1. Based on Administrator/Customer request, Predict the best possible time to go for an LCM operation for any given Entity in SDDC.

2. Dynamically identify the Entities that has “Available Upgrades” in SDDC and predict & propose a best possible time to trigger the Available Upgrades and also it will get the Risk Associated with respect proposed time. Example – If Administrator triggers LCM with the ML proposed time, then how many users may feel performance impact or how many people may be users at the proposed time.

3. Dynamic Time-based LCM - Dynamically associate the Application that is triggered by Administrator to LCM when Administrator chosen the proposed or predicted time suggested by this invention.

4. Administrator Dynamically can get Risk Associated with the any given time to LCM for an Entity.

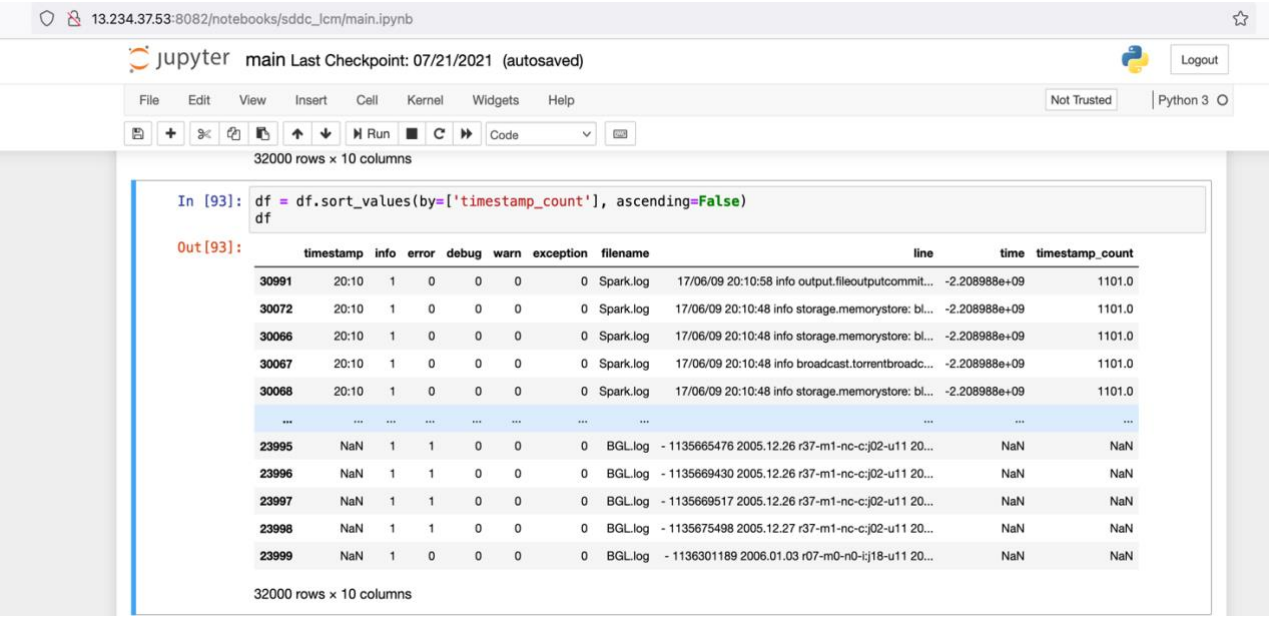
   Example – If Administrator wish to find the Risks Analysis for a specific time, then he can simply send a request to the OLA Manager. ML Algorithms will give the Risk Associated with that particular time that administrator chosen and send the details to Administrator.

Application Id	Name	Proposed LCM Time	Risk in the Proposed time
aaaaa	App1	17/06/09 20:10	8%
bbbbb	App2	17/06/09 20:10	10%
ccccc	App3	17/06/09 20:10	2%
ddddd	App4	17/06/09 20:10	5%

Table 3. Entity LCM Table

Code Implementation : https://github.com/rvirjn/sddc_lcm

Jupyter Notebook screen shot : http://13.234.37.53:8082/notebooks/sddc_lcm/main.ipynb



Conclusions / Recommendations

This Invention best suitable for VMware new upcoming architecture called Pantheon.

Directions for future work

Extend to BOTS.

Artificial intelligence AI bot software.

Bibliography/References

VMware VCF:

https://infohub.delltechnologies.com/l/architecture-guide-vmware-cloud-foundation-4-2-on-vxrail-7-0-1/cloud-management-8

https://m.softchoice.com/web/newsite/documents/partners/dell/hci/vmware_cloud_foundation_on_vxrail_architecture_guide.pdf

Appendices

VMware Cloud Provider Platform: Architectural Guidelines Powered by VMware Cloud Foundation 4.0

List of Publications/Conference Presentations

Presented to VMware vSphere solution org to take forward and get approval for implementation on company wider.

Duly Completed Checklist

a.	Is the Cover page in proper format?	Y
b.	Is the Title page in proper format?	Y
c.	Is the Certificate from the Supervisor in proper format? Has it been signed?	Y
d.	Is Abstract included in the Report? Is it properly written?	Y
e.	Does the Table of Contents page include chapter page numbers?	Y
f.	Does the Report contain a summary of the literature survey?	Y
g.	Are the Pages numbered properly?	Y
h.	Are the Figures numbered properly?	Y
i.	Are the Tables numbered properly?	Y
j.	Are the Captions for the Figures and Tables proper?	Y
k.	Are the Appendices numbered?	Y
l.	Does the Report have Conclusion / Recommendations of the work?	Y
m.	Are References/Bibliography given in the Report?	Y
n.	Have the References been cited in the Report?	Y
o.	Is the citation of References / Bibliography in proper format?	Y

Signature of Supervisor

Signature of student

BIRLA INSTITUTE OF TECHNOLOGY & SCIENCE, PILANI
Work Integrated Learning Programmes Division
I SEMESTER 2020-21

DSE CL ZG628T **DISSERTATION**

Final Evaluation Sheet

NAME OF THE STUDENT : **Ravi Ranjan**

ID NO. : **2018AB04501**

EMAIL : **2018ab04501@wilp.bits-pilani.ac.in**

NAME OF THE SUPERVISOR : **Ravi kottapalli**

PROJECT TITLE : Knowledge Based Life Cycle Management for Applications & Infrastructure Components in Scaled Hyper Converged Software Defined Data Center.

S.No.	Criteria	Excellent	Good	Fair	Poor
1	Work Progress and Achievements	✓			
2	Technical/Professional Competence	✓			
3	Documentation and expression	✓			
4	Initiative and originality	✓			
5	Punctuality	✓			
6	Reliability	✓			
	Recommended Final Grade	✓			

EVALUATION DETAILS

EC No.	Component	Weightage	Marks Awarded
1	Dissertation Outline	10%	10%
2	Mid-Sem Progress		
	Seminar	10%	10%
	Viva	5%	5%
	Work Progress	15%	15%
3	Final Seminar/Viva	20%	20%
4	Final Report	40%	35%
	Total out of	100%	95%

	Supervisor	Additional Examiner
Name	Ravi kottapalli	
Qualification	MS	
Designation & Address	Staff Customer Adoption Solution Engineer, Bangalore	
Email Address	talk2me.ravi@gmail.com	
Signature		
Date	8 August 2021	

Address:

Athulyam 687, Ground Floor 3B main road
Shanti Niketan layout
BENGALURU, KARNATAKA 560076
Pin Code : 560076