Cloud Computing: Past, Present and Future

Dr. Sanjaya Kumar Panda

IEEE Senior Member and CSI & ACM Distinguished Speaker

Assistant Professor

Department of Computer Science and Engineering
National Institute of Technology, Warangal
(An Institute of National Importance under MHRD, Govt. of India)
Warangal - 506004, Telangana, India
Mobile No.: +91-9861126947
Email: sanjayauce [at] gmail [dot] com
sanjaya [at] nitw [dot] ac [dot] in

Google Scholar DBLP YouTube

Topics to be discussed

Topics to be discussed

1. Public Cloud Computing - Gartner
2. Research Topics
   1. Energy-Efficient Cloud Computing
   2. Renewable Energy-Based Cloud Computing
   3. Vehicular Cloud Computing
3. Conclusion
4. References

Public Cloud Computing - 2 to 5 Years (Grow) and >10 Years (Adaption)

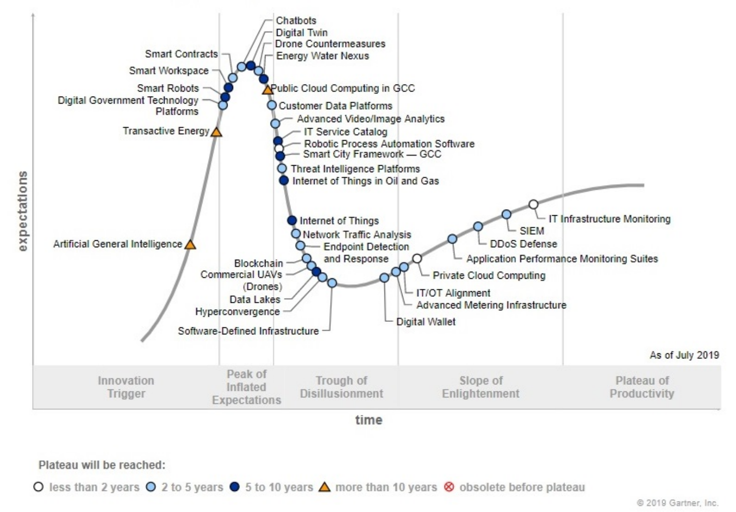


Figure 1: Gartner Hype Cycle for IT in GCC, 2019.

4 Trends Impacting Cloud Adoption in 2020 [2]

Gregor Petri, Vice President Analyst, Gartner

By 2023, the leading **cloud service providers** will have a **distributed ATM**-like presence to serve a subset of their **services**

4 Trends

- **Cost optimization** will drive **cloud adoption**
- **Multicloud** will reduce **vendor lock-in**
- **Insufficient cloud IaaS skills** will **delay migrations**
- **Distributed cloud** will support **expanded service availability**

Energy-Efficient Cloud Computing

- How to **Utilize Resources** and How to Reduce **Power Consumption** [21]?
- Energy Consumption - CPU Utilization [22]
- **Datacenter**: 25,000 Householders [23]
- **ICT Resources**: 8% of Total Energy Consumption (50% in Next Decade) [24]
- **U. S. Electricity**: 66% using Coal and Natural Gas [24]
- **NRDC Report**: 91 billion kilowatt-hours (kWh) of Electricity Consumed by U. S. Datacenters (140 billion kWh by 2020) [25]

CPU Utilization Vs. Power Consumption [26]

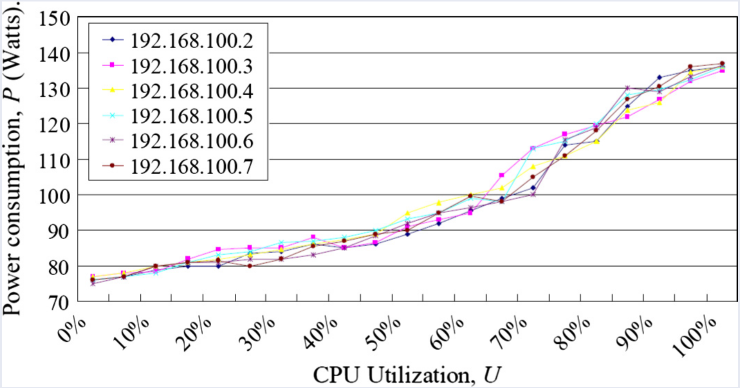


Figure 2: The power consumption of six typical workloads served by a streaming server.

- **CPU utilization Vs. Energy Consumption**: Not Linear

CPU Utilization Vs. Power Consumption [26]

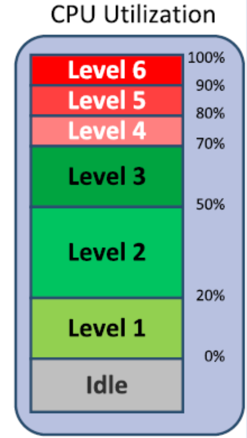
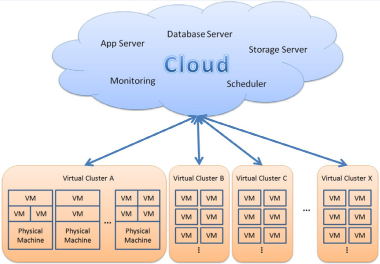


Figure 3: A Suggested Model [21].

A Cloud System Composed of Virtual Clusters (VCs) and Network Bandwidth between VCs



From \ To		Virtual Cluster A	B	C	...
Virtual Cluster	A		250 MB/s	500 MB/s	...
	B	200 MB/s		250 MB/s	...
	C	300 MB/s	300 MB/s		...
	⋮	⋮	⋮	⋮	

Figure 4: A Research Model [21].

Task Consolidation

Table 1: A List of Tasks [27]

Task	Arrival Time	Processing Time	Utilization
0	0	20	40%
1	3	8	50%
2	7	23	20%
3	14	10	40%
4	20	15	70%

Task Consolidation

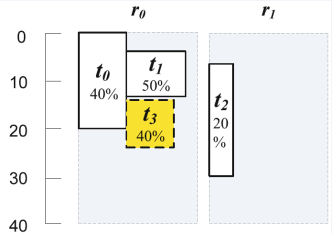


Figure 5: Consolidation Example for Tasks - Choice 1 [27].

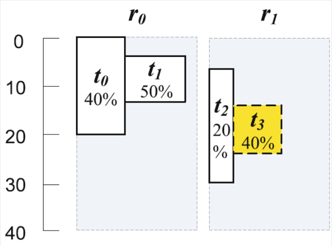


Figure 6: Consolidation Example for Tasks - Choice 2 [27].

Task Consolidation Among VCs - Threshold 70%

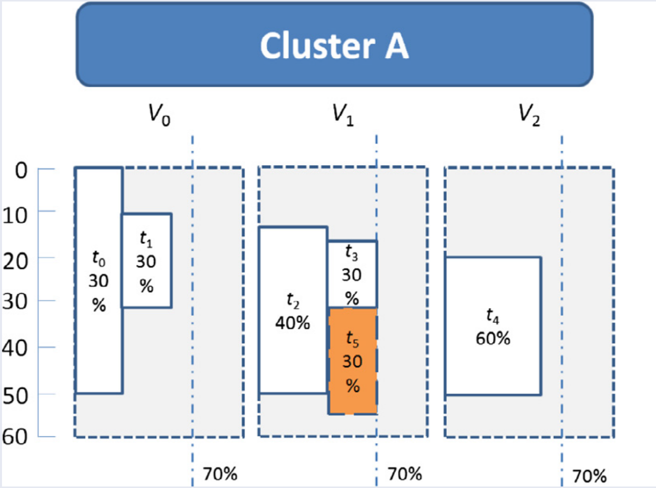


Figure 7: Assigning Tasks in VC_A.

Task Consolidation Among VCs - Threshold 70%

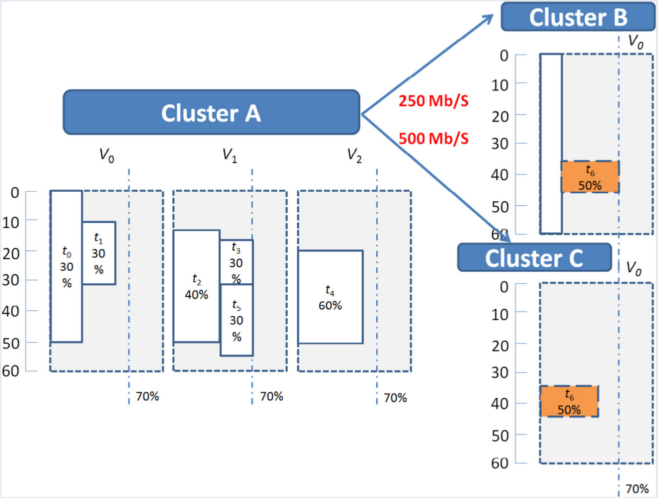


Figure 8: VC_A Asks for Resource Support When Assigning t_6 [21].

Task Consolidation Among VCs - Threshold 70%

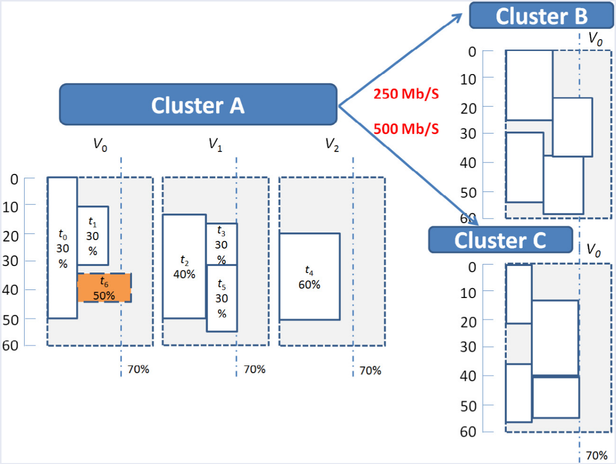


Figure 9: VC_A Assigns t_6 to V_0 Without Conforming to 70% CPU Utility [21].

Task Consolidation Algorithms

- Random [27]
- ECTC [27]
- MaxUtil [27]
- ETC [21]
- ETSA [28]

Renewable Energy-Based Cloud Computing

- Motivation [29]
  - **Datacenters:** 8.6 million Datacenters (3 million in the U. S.)
  - 50,000 to 80,000 Servers - 25 to 30 megawatts
  - **Global Datacenters**: 416 terawatts of Electricity per Year
- Cloud Service Providers [30]
  - **Non-renewable Energy Sources:** Fossil Fuels
    - Coal, Gas, Orimulsion and Petroleum
  - **Carbon Dioxide, Particle and Heat** - Harmful for Environment
  - **Solution:** Renewable Energy Sources (RES)
    - Biomass, Hydropower, Solar and Wind
  - **Google and Microsoft Datacenters**: Fully Powered by RES

Renewable Energy-Based Cloud Computing [30]

- **Renewable Energy Sources**: Not Available Round the Clock
- Both **Non-Renewable** and **Renewable Energy Sources**

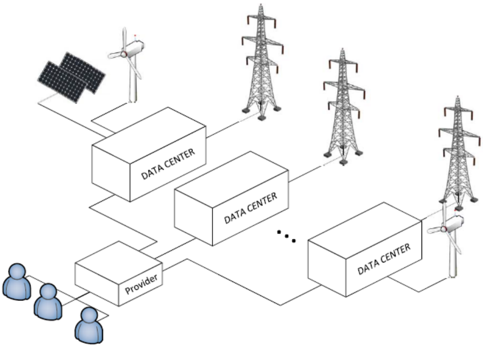
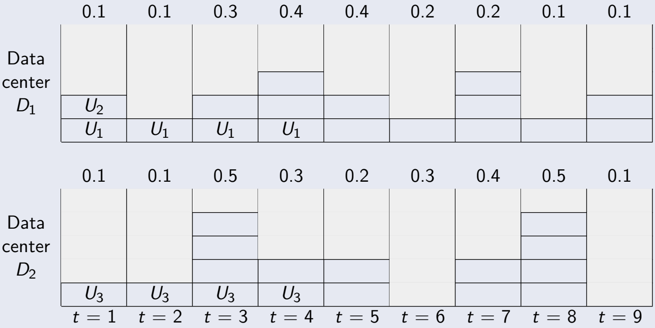


Figure 10: A System Model [30].

Research Topic #2

Nine tasks, U_1 to U_9 (i.e., 3 assigned tasks and 6 unassigned tasks) and two datacenters, D_1 and D_2 (each with 5 nodes) [30]

User Request	U_1	U_2	U_3	U_4	U_5	U_6	U_7	U_8	U_9
Start Time	1	1	1	3	4	5	5	7	8
Nodes	1	1	1	2	1	1	1	2	3
Duration	4	1	4	5	3	2	3	2	2



Future-Aware Best Fit (FABEF) [30]

- Route each request to a datacenter leading to the **lowest cost**

Example - U_4: ST - 3, N - 2 and D - 5

Data center D_1

	0.1	0.1	0.3	0.4	0.4	0.2	0.2	0.1	0.1
			U_4	U_4					
	U_2		U_4	U_4	U_4	U_4	U_4		
	U_1	U_1	U_1	U_1	U_4	U_4	U_4		

Data center D_2

	0.1	0.1	0.5	0.3	0.2	0.3	0.4	0.5	0.1
			U_4	U_4					
			U_4	U_4	U_4	U_4	U_4		
	U_3	U_3	U_3	U_3	U_4	U_4	U_4		
	$t=1$	$t=2$	$t=3$	$t=4$	$t=5$	$t=6$	$t=7$	$t=8$	$t=9$

- Cost of datacenter D_1 for request $U_4 = 0.3 + 0.2 = \mathbf{0.5} \leftarrow$ Lowest cost
- Cost of datacenter D_2 for request $U_4 = 0.3 + 0.3 + 0.3 = 0.9$

Renewable Energy-Based Algorithms

- Future-Aware Best Fit [30]
- Static Cost-Aware Ordering [31]
- Round Robin [30]
- Highest Available Renewable First [30]
- MinBrown [32]
- Fuzzy Logic-Based Load Balancing [30]
- Worst Fit [31]
- MinUtil [33]

Vehicular Cloud Computing [34]

- Motivation
  - Cloud Computing and Vehicular Ad-Hoc Network
  - **Smart Vehicle**: On-Board Unit (GPS, Sensors, Radar Device, Cameras, Digital Map, Processing Unit, Storage Unit, etc.)
  - Smart Vehicle - Host VMs
  - **Underutilized Resources**: Parking Lot, Roadways and Streets
  - Vehicular Clouds

Vehicular Cloud Computing [34]

Table 2: Comparative Study of CC and VCC

Characteristics	**CC**	**VCC**
Mobility of Clouds	No	Yes
Autonomous Cloud Formation	No	Yes
Automatic Cloud Federation	No	Yes
Moving Network Pool	No	Yes
Large Traffic Event Management	Possible	Yes
Planned and Unplanned Disaster Management	Possible	Yes
Corporation as a Service	Possible	Yes

Vehicular Cloud Computing [35]

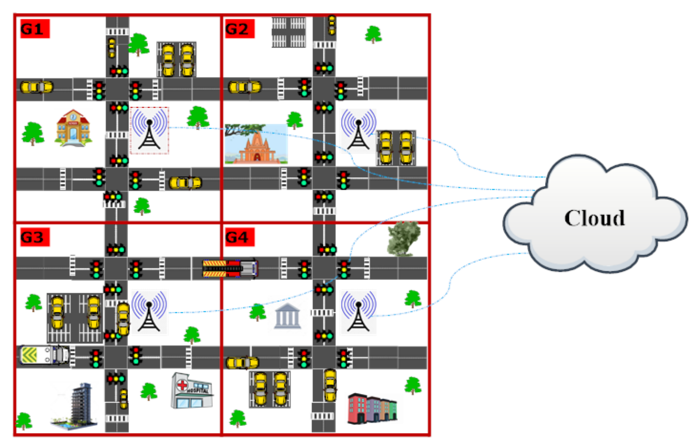


Figure 11: A Large Urban Area With Four Grids.

Vehicular Cloud Computing [35]

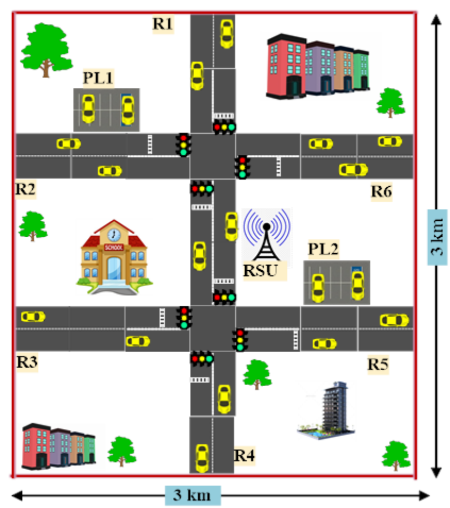


Figure 12: A Sample Grid.

Vehicular Cloud Computing [36]

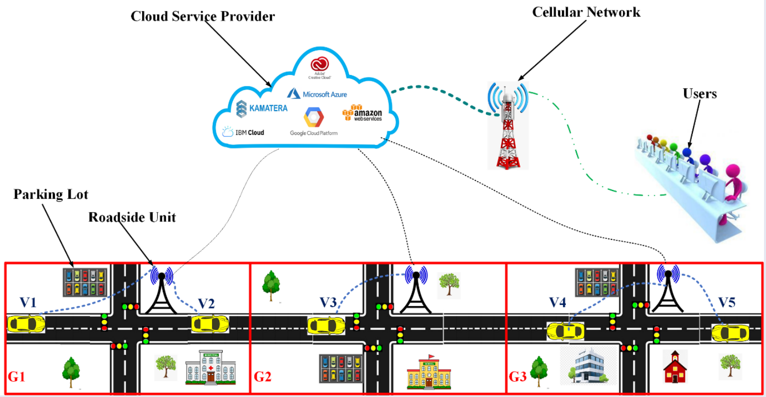


Figure 13: An Overview of VC Model With Three Grids.

Renewable Energy-Based Algorithms

- Vehicular VM Migration-Uniform [37]
- Vehicular VM Migration-Least Workload [37]
- Vehicular VM Migration-Mobility Aware [37]
- Round Robin [38]
- Deficit Weighted RR [39]
- Mobility and Destination Workload Aware Migration [37]
- Dynamic Service Migration [35]
- Smart Cloud Service Management [36]

Conclusion

Public Cloud Computing

- Gartner

Research Topics

1. Energy-Efficient Cloud Computing
2. Renewable Energy-Based Cloud Computing
3. Vehicular Cloud Computing

References

1. Gartner's 2019 Hype Cycle for IT in GCC Indicates Public Cloud Computing Will Transform Businesses, `https://www.gartner.com/en/newsroom/press-releases/2019-10-14-gartner-s-2019-hype-cycle-for-it-in-gcc-indicates-pub`, Accessed on 6^{th} July 2020.
2. 4 Trends Impacting Cloud Adoption in 2020, `https://www.gartner.com/smarterwithgartner/4-trends-impacting-cloud-adoption-in-2020/`, Accessed on 6^{th} July 2020.
3. Gartner Forecasts Worldwide Public Cloud End-User Spending to Grow 23% in 2021, `https://www.gartner.com/en/newsroom/press-releases/2021-04-21-gartner-forecasts-worldwide-public-cloud-end-user-spending-to-grow-23-percent-in-2021` Accessed on 2^{nd} July 2021.
4. Google Trends, `https://trends.google.com/trends/explore?date=today%205-y&q=Public%20Cloud%20Computing`, Accessed on 12^{th} July 2020.
5. A. Mukherjee, "An Efficient Job-Grouping Based Scheduling Algorithm for Fine-Grained Jobs in Computational Grids", M. Tech. Thesis, Department of Computer Science and Engineering, National Institute of Technology, Rourkela, 2011.
6. D. Heap, "Scorpion: Simplifying the Corporate IT Infrastructure", IBM Research White Paper, 2000.
7. A. Chakrabarti, "Grid Computing Security", Springer, 2007.
8. H. Fuchs et al., "Characteristics and Emergency Use of Volume Servers in the United States", Lawrence Berkeley National Laboratory, 2017.
9. Li et al., "Online Optimization for Scheduling Preemptable Tasks on IaaS Cloud Systems", Journal of Parallel and Distributed Computing, Elsevier, Vol. 72, pp. 666-677, 2012.
10. R. Buyya, High Performance Cluster Computing, Pearson Education, 2008.
11. S. Abrishami et al., "Deadline-Constrained Workflow Scheduling Algorithms for Infrastructure as a Service Clouds", Future Generation Computer Systems, Elsevier, Vol. 29, pp. 158-169, 2013.

References

12. T. D. Braun et al., "A Comparison of Eleven Static Heuristics for Mapping a Class of Independent Tasks onto Heterogeneous Distributed Computing Systems", Journal of Parallel and Distributed Computing, Vol. 61, pp. 810-837, 2001.
13. M. Maheswaran et al., "Dynamic Mapping of a Class of Independent Tasks onto Heterogeneous Computing Systems", Journal of Parallel and Distributed Computing, Vol. 59, pp. 107-131, 1999.
14. A. Nathani et al., "Policy Based Resource Allocation in IaaS Cloud", Future Generation Computer Systems, Elsevier, Vol. 28, pp. 94-103, 2012.
15. M. Wu et al., "Segmented Min-Min: A Static Mapping Algorithm for Meta-tasks on Heterogeneous Computing Systems", 9^{th} Heterogeneous Computing Workshop, IEEE, 2000.
16. H. Xiaoshan et al., "QoS Guided Min-Min Heuristic for Grid Task Scheduling", Journal of Computer Science and Technology, Vol. 18, Issue 4, pp. 442-451, 2003.
17. F. Dong et al., "Grid Task Scheduling Algorithm Based on QoS Priority Grouping", Fifth International Conference on Grid and Cooperative Computing, IEEE, 2006.
18. K. Etminani et al., "Min-Min Max-Min Selective Algorithm for Grid Task Scheduling", Third IEEE/IFIP International Conference on Internet, 2007.
19. S. Wang et al., "Towards a Load Balancing in a Three-level Cloud Computing Network", 3^{rd} International Conference on Computer Science and Information Technology, IEEE, pp. 108-113, 2010.
20. H. Chen et al., "User-Priority Guided Min-Min Scheduling Algorithm for Load Balancing in Cloud Computing", National Conference on Parallel Computing Technologies, IEEE, 2013.
21. C. Hsu et al., "Optimizing Energy Consumption with Task Consolidation in Clouds", Information Sciences, Elsevier, Vol. 258, pp. 452-462, 2014.
22. C. H. Lien et al., "Web Server Power Estimation, Modeling and Management", 14^{th} IEEE International Conference on Networks, Vol. 2, pp. 1-6, 2006.
23. A. Beloglazov et al., "Energy-Aware Resource Allocation Heuristics for Efficient Management of Data Centers for Cloud Computing", Future Generation Computer Systems, Elsevier, Vol. 28, pp. 755-768, 2012.
24. M. P. Mills, "The Cloud Begins with Coal: Big Data, Big Networks, Big Infrastructure and Big Power", Technical report, National Mining Association, American Coalition for Clean Coal Electricity, 2013.

References

25. J. Whitney et al., "Data Center Efficiency Assessment", National Resources Defense Council, New York, Issue Paper, 2014.
26. C. H. Lien et al., "Measurement by the Software Design for the Power Consumption of Streaming Media Servers", The IEEE Instrumentation and Measurement Technology Conference, pp. 1597-1602, 2006.
27. Y. C. Lee et al., "Energy Efficient Utilization of Resources in Cloud Computing Systems", Journal of Supercomputing, Springer, Vol. 60, pp. 268-280, 2012.
28. S. K. Panda et al., "An Energy-Efficient Task Scheduling Algorithm for Heterogeneous Cloud Computing Systems", Cluster Computing, Springer, Vol. 22, No. 2, pp. 509-527, 2019.
29. R. Miller, Inside amazon's cloud computing infrastructure, `https://datacenterfrontier.com/inside-amazon-cloud-computing-infrastructure/`. Accessed on 10^{th} April 2020.
30. A. N. Toosi et al., "A Fuzzy Logic-Based Controller for Cost and Energy Efficient Load Balancing in Geo-Distributed Data Centers", 8^{th} International Conference on Utility and Cloud Computing, pp. 186-194. IEEE, 2015.
31. K. Le et al., "Reducing Electricity Cost Through Virtual Machine Placement in High Performance Computing Clouds", International Conference for High Performance Computing, Networking, Storage and Analysis, pp. 22, ACM, 2011.
32. C. Chen et al., "Green-Aware Workload Scheduling in Geographically Distributed Data Centers", 4^{th} IEEE International Conference on Cloud Computing Technology and Science Proceedings, pp. 82-89, 2012.
33. S. K. Nayak et al., "A Renewable Energy-Based Task Consolidation Algorithm for Cloud Computing", Electric Power and Renewable Energy Conference, pp. 1-10, 2020.
34. M Whaiduzzaman et al., "A Survey on Vehicular Cloud Computing", Journal of Network and Computer Applications, Elsevier, Vol. 40, pp. 325-344, 2014.
35. S. K. Pande et al., "Dynamic Service Migration and Resource Management for Vehicular Clouds", Journal of Ambient Intelligence and Humanized Computing, Springer, pp. 1-21, 2020

References

36. S. K. Pande et al., "A Smart Cloud Service Management Algorithm for Vehicular Clouds", IEEE Transactions on on Intelligent Transportation Systems, pp. 1-11, 2020.
37. T. Refaat et al., "Virtual Machine Migration and Management for Vehicular Clouds", Vehicle Communication, Elsevier, Vol. 4, pp. 47-56, 2016.
38. J. Pillmann et al., Car-to-Cloud Communication Traffic Analysis Based on the Common Vehicle Information Model, IEEE 85^{th} Vehicular Technology Conference, pp. 1-5, 2017.
39. T. M. Lim et al., Weighted Deficit Earliest Departure First Scheduling, Computer Communications, Elsevier, Vol. 28, No. 15, pp. 1711-1720, 2005.
40. P. J. Sun, "Security and Privacy Protection in Cloud Computing: Discussions and Challenges", Journal of Network and Computer Applications, Elsevier, 2020.
41. M. Albanese, "Security and Trust in Cloud Application Life-Cycle Management", Future Generation Computer Systems, Elsevier, Vol. 111, pp. 934-936, 2020.
42. Top Emerging Technologies in Cloud Computing, `https://www.datamation.com/cloud-computing/top-six-emerging-technologies-in-cloud-computing.html`, Accessed on 12^{th} July 2020.
43. Kubernetes vs. Docker: What Does It Really Mean?, `https://www.sumologic.com/blog/kubernetes-vs-docker/`, Accessed on 12^{th} July 2020.
44. B. Varghese and R. Buyya, "Next Generation Cloud Computing: New Trends and Research", Future Generation Computer Systems, Elsevier, Vol. 79, pp. 849-861, 2018. directions"
45. Market-Oriented Cloud Architecture, `https://www.researchgate.net/figure/High-level-market-oriented-cloud-architecture_fig3_1771824`, Accessed on 2^{nd} August 2021.

Thank You!