Autonomic Computing in the Cloud: An Overview of Past, Present and Future Trends

Alistair McLean, Roy Sterritt

Presented by **Alistair McLean**
School of Computing
Faculty of Computing, Engineering
and the Built Environment
Ulster University
mclean-a13@ulster.ac.uk

IARIA Congress 2023

Autonomic Computing Systemization of Knowledge Session

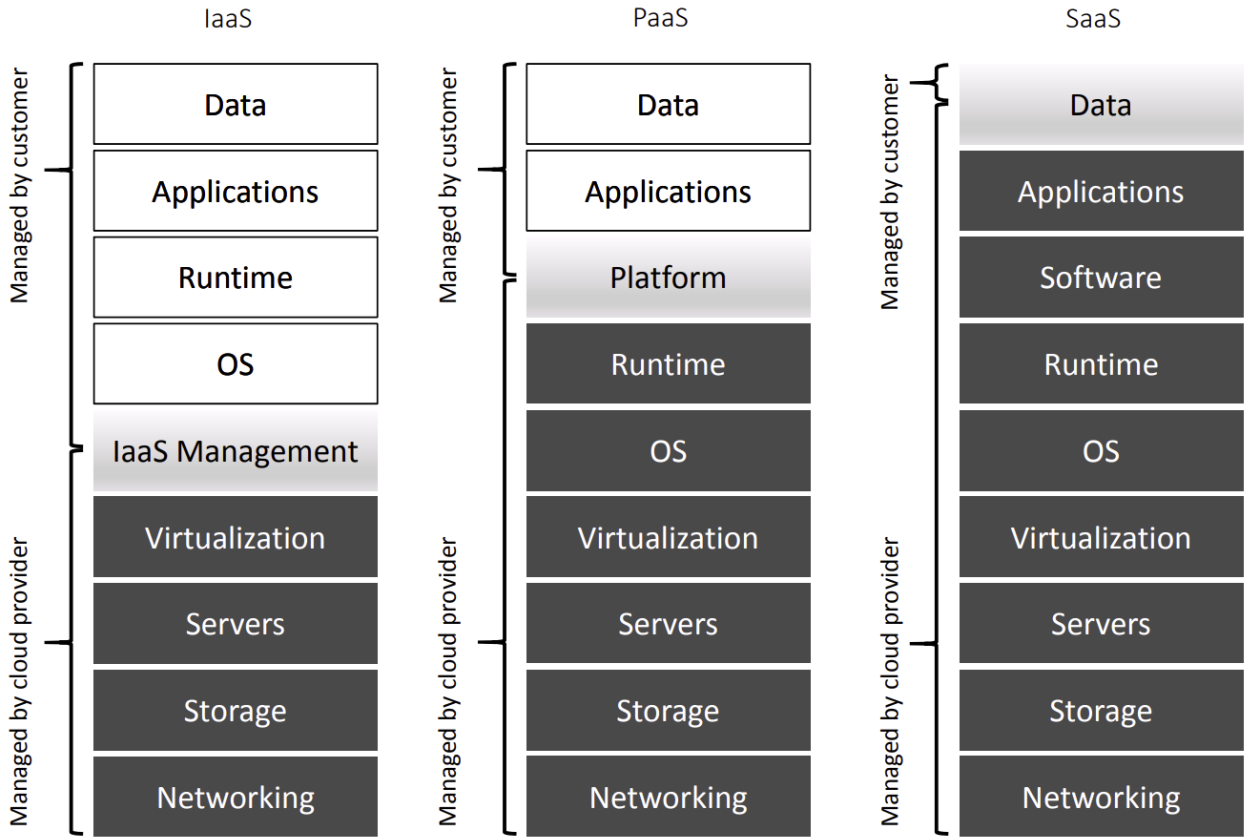
Alistair McLean

- MSc A.I. Candidate at Ulster University
- Autonomic Computing interest from MSc A.I. module: COM760 Autonomic Computing and Robotics
- Experience in software and data engineering, predominantly working in the telecommunications industry
- https://www.linkedin.com/in/alistair-mclean/

Introduction - Cloud Computing

Cloud computing is a method of utilising remote computing resource and capacity provided by means of an internet service [1]. It has become customary in the field of cloud computing to describe cloud service models with the "as a service" (aaS) phrase – prepended by the technology on offer [2]. Although there has been many takes on this "as a service" approach – the major cloud service providers have generally adopted a three-tier approach, each tier representing a distinct level of resource abstraction and control. They include:

1. Software as a Service (SaaS)
2. Platform as a Service (PaaS)
3. Infrastructure as a Service (IaaS)



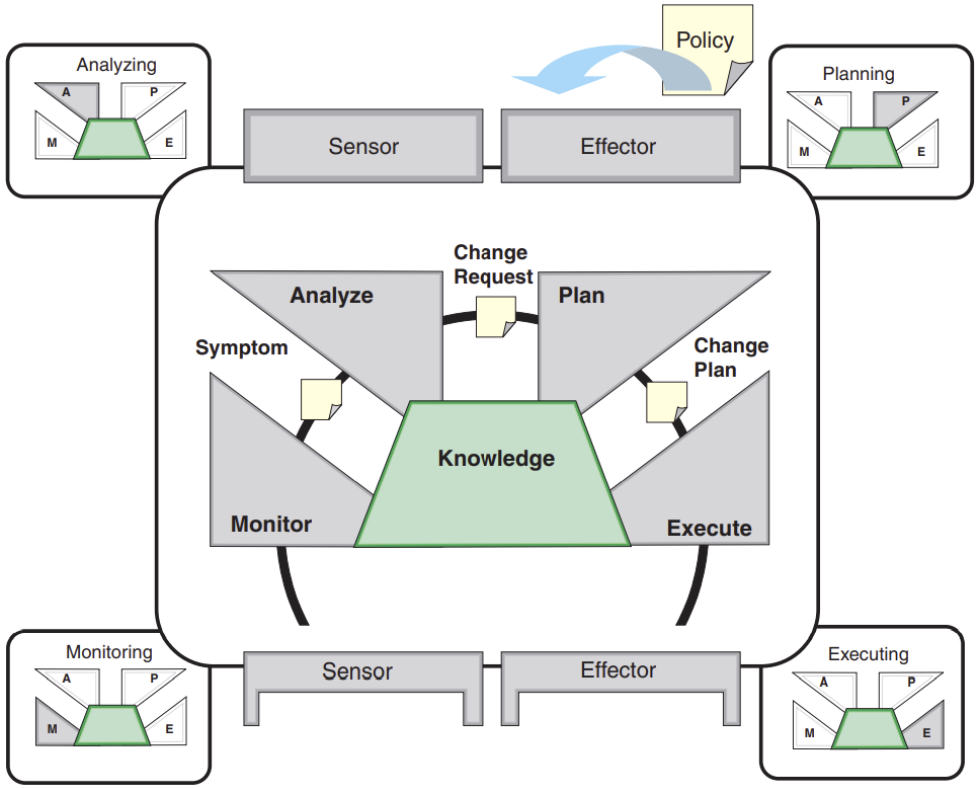
Representation of resource abstraction and level of control for the three main service models [3].

[1] V. Arutyunov, "Cloud computing: Its history of development, modern state, and future considerations", Scientific and Technical Information Processing, 39(3), pp.173-178, 2012.
[2] Y. Duan, G. Fu, N. Zhou, X. Sun, N. Narendra, and B. Hu, "Everything as a Service (XaaS) on the Cloud: Origins, Current and Future Trends", IEEE 8th International Conference on Cloud Computing, pp. 621-628, 2015.
[3] [J. Surbiryala, and C. Rong, "Cloud Computing: History and Overview", IEEE Cloud Summit, pp. 1-7, 2019.

Introduction - Autonomic Computing

Inspired by the autonomic nervous system where bodily functions are unconsciously regulated, autonomic computing is a design model that aims to create computer-based systems that, through self- and environmental awareness, act to self-configure, self-heal, self-optimise, and self-protect (self-CHOP) [1][2].

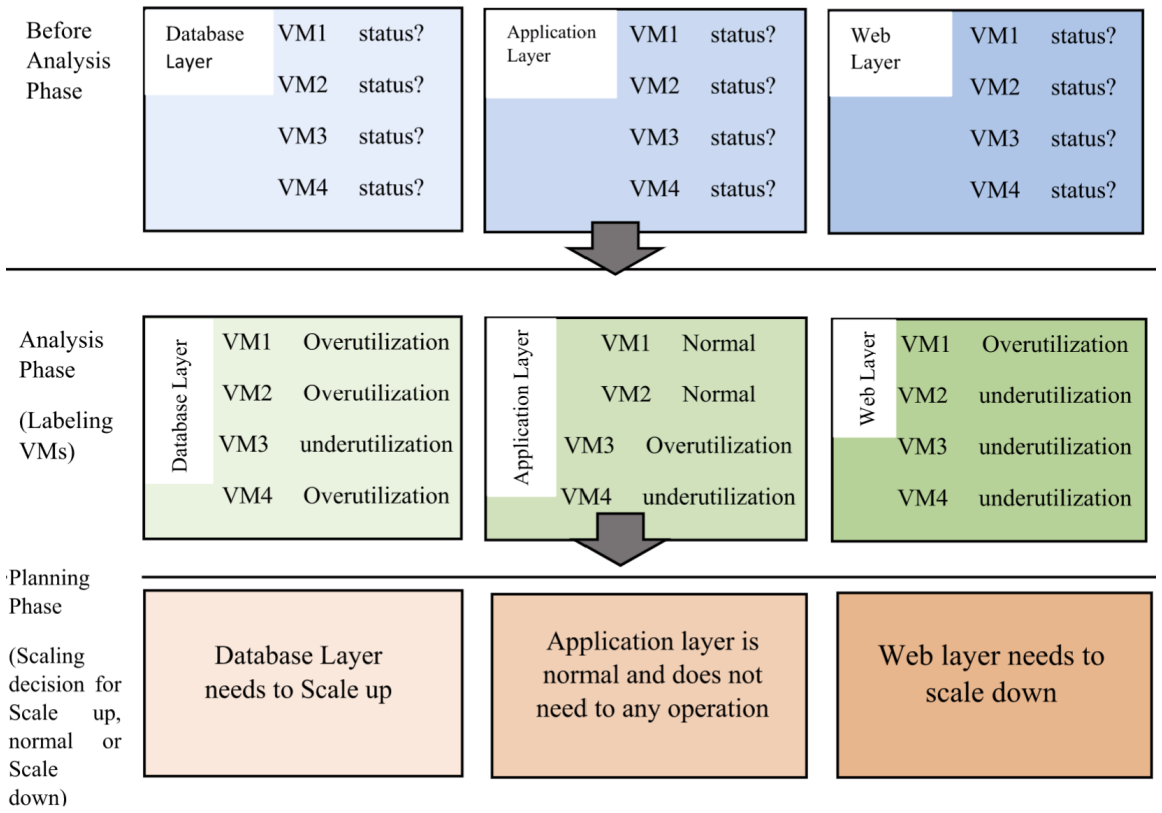
An established method for achieving self- and environmental awareness in computing systems is by use of autonomic managers [3]. An autonomic manager cycles through a four-step control loop entailing Monitoring, Analysing, Planning, and Executing (MAPE) managed elements in a system. Sensors collect information from the autonomic element and from its environment, with effectors able to complete executable tasks to accomplish system adaptation [4]. Figure shows the high-level design of an autonomic manager [3].



[1] R. Sterritt, "Towards Autonomic Computing: Effective Event Management", Proceedings of 27th Annual IEEE/NASA Software Engineering Workshop (SEW), IEEE Computer Society, pp 40-47, 2002.
[2] R. Sterritt, and M. Hinchey, "Why Computer-Based Systems Should Be Autonomic", 12th IEEE International Conference and Workshops on the Engineering of Computer-Based Systems (ECBS'05), pp. 406-412, 2005. doi: 10.1109/ECBS.2005.75.
[3] IBM, "An Architectural Blueprint for Autonomic computing", IBM White Paper 3rd Ed., 2005
[4] R. Kazhamiakin, S. Benbernou, L. Baresi, P. Plebani, M. Uhlig, and O. Barais, "Adaptation of Service-Based Systems", Service Research Challenges and Solutions for the Future Internet, pp.117-156, 2010.

Cloud Computing

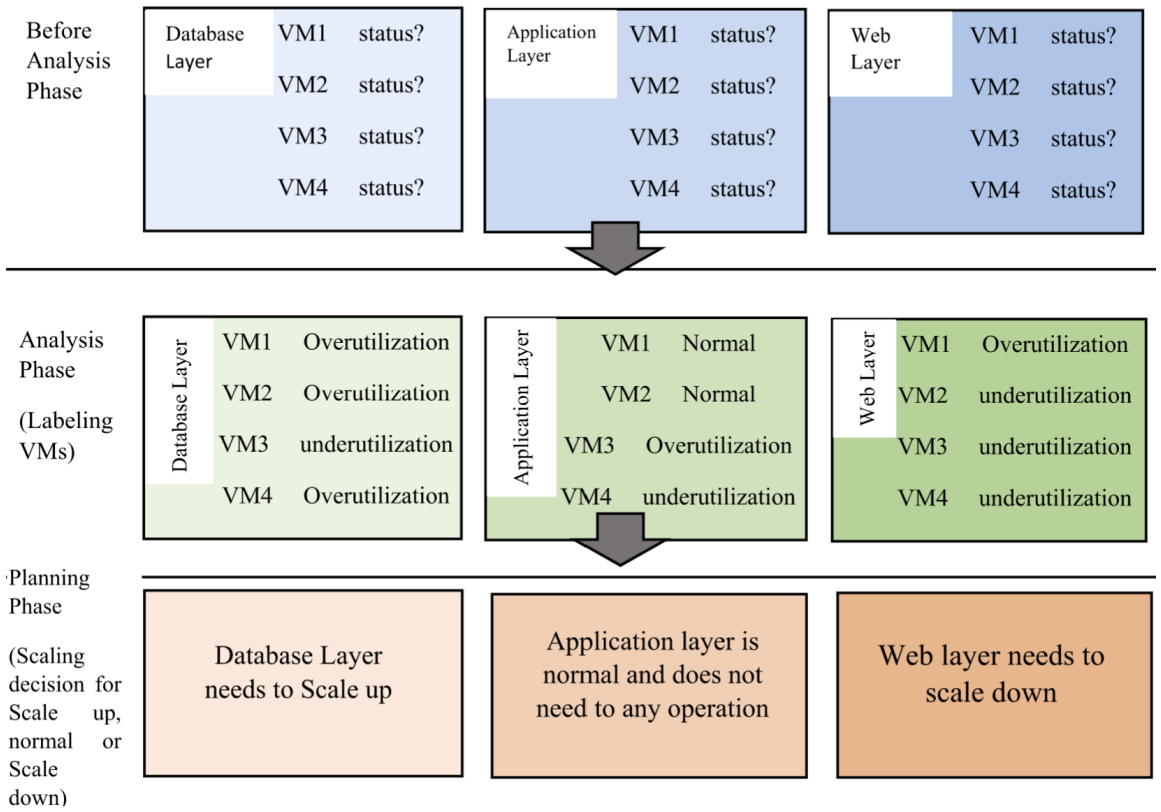
The ability for cloud consumers to scale up their resources when required, and decommission or scale down when demand is reduced, has been one the greatest benefits of cloud computing. This elastic quality reduces the need for vast resource redundancy in preparation for peak demand – the infrastructure can simply scale up its capacity during peak times. This has the benefit of reducing the running costs for cloud providers, with cloud consumers only paying for the resources that are needed to maintain QoS. However, the processes involved with scaling resources up and down take time and have associated costs and therefore research has aimed to optimise this autonomic process. One such paper utilises machine learning techniques to classify Virtual Machines (VM) in a system during the analysis stage of the MAPE-K loop [1]. The VMs are labelled with a status of "Normal", "Underutilized", or "Overutilized" at each layer of the system based on their workload.



[1] A. Mazidi, M. Golsorkhtabaramiri, and M. Tabari, "Autonomic resource provisioning for multilayer cloud applications with K-nearest neighbor resource scaling and priority-based resource allocation", Software: Practice and Experience, 50(8), pp.1600-1625, 2020.

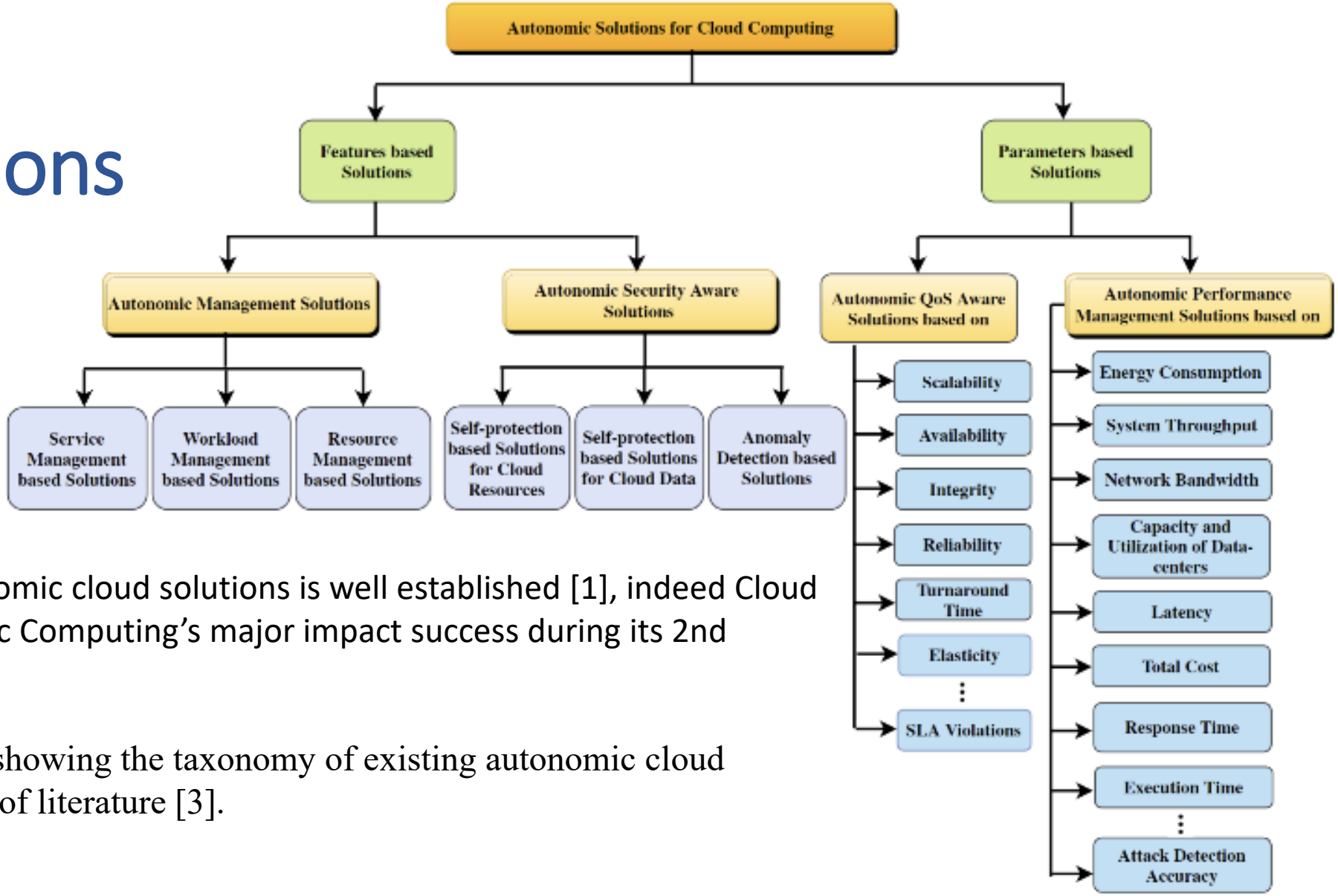
Cloud Computing

Figure shows how labelling the VMs can inform the autoscaling decision process at each layer. Experimental results of simulations using this method discovered benefits including improved VM utilisation, shorter response times for customer requests, and lower operating costs for the cloud consumer [1].



[1] A. Mazidi, M. Golsorkhtabaramiri, and M. Tabari, "Autonomic resource provisioning for multilayer cloud applications with K-nearest neighbor resource scaling and priority-based resource allocation", Software: Practice and Experience, 50(8), pp.1600-1625, 2020.

Autonomic Cloud Solutions



Work in the field of autonomic cloud solutions is well established [1], indeed Cloud Computing was Autonomic Computing's major impact success during its 2nd decade [2].

The figure is an overview showing the taxonomy of existing autonomic cloud solutions based on review of literature [3].

[1] T. Lorimer and R. Sterritt, "Autonomic Management of Cloud Neighborhoods through Pulse Monitoring," 2012 IEEE Fifth International Conference on Utility and Cloud Computing, Chicago, IL, USA, 2012, pp. 295-302, doi: 10.1109/UCC.2012.60

[2] R. Sterritt, "Keynote: 20 Years of Autonomic Computing," in International Conference on Autonomic and Autonomous Systems (ICAS), Online (Covid-19), 2021.

[3] N. Agrawal, "Autonomic cloud computing-based management and security solutions: State-of-the-art, challenges, and opportunities", Transactions on Emerging Telecommunications Technologies, 32(12), 2021.

Potential Future Trends

We have already seen many examples of successful implementation of autonomic cloud solutions and improvements. However, some areas will benefit from continued research and innovation.

1. Cloud Privacy and Security

Relinquishing control of system and user data to the cloud provider will be a concern for many cloud consumers therefore work in the field of security-aware solutions will continue [1]. It has been speculated that the emergent field of "confidential computing" is the future of the cloud [2]. Confidential computing gives cloud consumers full control over their sensitive workloads. It explicitly details the computing components that they must trust, whilst providing strong protection from other components, and preventing attacks from other cloud users. Although still in its initial stages, it is expected that the field will grow rapidly to become as popular as some of the most prevalent privacy mechanisms of today.

[1] S. Gill, and A. Shaghaghi, "Security-Aware Autonomic Allocation of Cloud Resources", Journal of Organizational and End User Computing, 32(3), pp.15-22, 2020.
[2] M. Russinovich, M. Costa, C. Fournet, D. Chisnall, A. Delignat-Lavaud, S. Clebsch, K. Vaswani, and V. Bhatia, "Toward confidential cloud computing", Communications of the ACM, 64(6), pp.54-61, 2021

Potential Future Trends

2. IoT Ecosystems

Although seemingly two independent fields, the IoT and cloud computing are closely linked. IoT "generally refers to scenarios where network connectivity and computing capability extends to objects, sensors and everyday items not normally considered computers, allowing these devices to generate, exchange and consume data with minimal human intervention" [1]. Quite often it is cloud computing services that are facilitating IoT devices, but as the number of connected devices increases, cloud technologies are struggling to sustain real-time demand [2]. It is expected that research will continue to investigate autonomic processes to handle the complexity and demands of IoT systems [3]-[6].

[1] K. Rose, S. Eldridge, and L. Chapin, 2015. The internet of things: An overview, The Internet Society (ISOC), pp. 1–50.
[2] S. Gill, S. Tuli, M. Xu, I. Singh, K. Singh, D. Lindsay, S. Tuli, D. Smirnova, M. Singh, U. Jain, H. Pervaiz, B. Sehgal, S. Kaila, S. Misra, M. Aslanpour, H. Mehta, V. Stankovski, and P. Garraghan, "Transformative effects of IoT, Blockchain and Artificial Intelligence on cloud computing: Evolution, vision, trends and open challenges", Internet of Things, 8, 100118, 2019.
[3] A. Zyane, M. Bahiri, and A. Ghammaz, "IoTScal-H : Hybrid monitoring solution based on cloud computing for autonomic middleware-level scalability management within IoT systems and different SLA traffic requirements", International Journal of Communication Systems, 33(14), 2020.
[4] A. Lam, O. Haugen, and J. Delsing, "Dynamical Orchestration and Configuration Services in Industrial IoT Systems: An Autonomic Approach", IEEE Open Journal of the Industrial Electronics Society, 3, pp.128-145, 2022.
[5] S. Rahman, and G. Jackson, "Autonomic Methods for Mitigating Threats to the Internet of Things (IoT)", International Conference on Computational Science and Computational Intelligence (CSCI), 2017.
[6] E. Mezghani, S. Berlemont, and M. Douet, "Autonomic Coordination of IoT Device Management Platforms", IEEE 29th International Conference on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE), 2020.

Potential Future Trends

3. Energy Consumption and Sustainability

As mentioned, the enormous demand for computational processing and data storage on the cloud means that energy efficiency is a high priority topic. Data centres consume huge amounts of energy with high utilisation of resources, large operating costs, and substantial carbon footprints. In addition to using cleaner energy sources to power data centres, it is paramount that progress continues in the field of computational energy efficiency. We have already described some of the successes in this endeavour, but it is expected that research and development into energy-efficient computation will continue to improve as energy consumption of the cloud grows [1]-[3].

[1] T. Tekreeti, T. Cao, X. Peng, T. Bhattacharya, J. Mao, X. Qin, and W. Ku, “Towards Energy-Efficient and Real-Time Cloud Computing”, IEEE International Conference on Networking, Architecture and Storage (NAS), 2021.
[2] S. Simaiya, V. Gautam, U. Lilhore, A. Garg, P. Ghosh, N. Trivedi, and A. Anand, “EEPSA: Energy Efficiency Priority Scheduling Algorithm for Cloud Computing”, 2nd International Conference on Smart Electronics and Communication (ICOSEC), 2021.
[3] M. Xu, A. Toosi, and R. Buyya, “A Self-Adaptive Approach for Managing Applications and Harnessing Renewable Energy for Sustainable Cloud Computing”, IEEE Transactions on Sustainable Computing, 6(4), pp.544-558, 2021.

Conclusion

Autonomic computing has been a key facilitator in the advancement of cloud computing. With the scale and complexity of cloud computing systems growing, autonomic computing has helped deal with the difficulty of managing these systems. Autonomic computing has shown great advantages, including improved dependability of systems, through the ability to self-configure, self-heal, self-optimise, and self-protect.

Historically, we have seen the challenges and successes of applying autonomic principles to a cloud infrastructure. The ability to manage autonomic elements at runtime in a heterogenous environment was achieved with innovations on the topic of ARM. SLA violations drove advancements in autonomic techniques to create a tailored approach for cloud applications, for example the proposal of an A-MAPE-K control loop within autonomic managers. Furthermore, identifying that the process of resource scaling could be optimised with respect to time and cost saw the introduction of other technologies, such as those used in machine learning, to support decision making.

Conclusion

Evaluation of the current state of the art highlighted new innovative solutions alongside considerable improvements to existing autonomic techniques in the cloud. Autonomic management solutions have been able to drastically reduce SLA violation occurrence rates, increase resource utilisation, and reduce operational costs of resources resulting in increased profit. Development of self-protecting security-aware solutions has expanded on existing security techniques whilst improving the QoS of systems under attack. QoS-aware systems have shown resource management benefits including lower costs, improved latency, and shorter execution times. Performance management research is extremely important for both cloud providers and cloud users, with the aim being to improve energy efficiency of computation. Innovation in this field has shown promise with proposed solutions achieving considerable energy reductions whilst maintaining performance.

The current state of the art in autonomic cloud computing is promising. Further work in the area will likely see optimisations with respect to self-CHOP and MAPE mechanisms in cloud-based computing systems. As the digital age continues, with more and more data generated every day, the importance for cloud providers to handle data in an efficient and secure manner will increase. It is expected that optimisation of cloud security will continue to be an active research topic in the future. Additionally, in the interest of sustainable ethical practices and Corporate Social Responsibility (CSR), cloud providers are becoming increasingly pressured to address the scale of their operational energy consumption. With the vast energy demands of data centres used to provide cloud computing services continually growing, it is expected that research into energy efficient computation will long continue.

Acknowledgements

This paper was produced as part of
COM760
Autonomic Computing & Robotics
on Ulster University's
MSc in Artificial Intelligence.

Thank You

Presented by **Alistair McLean**
School of Computing
Faculty of Computing, Engineering
and the Built Environment
Ulster University
mclean-a13@ulster.ac.uk

IARIA Congress 2023

Autonomic Computing Systemization of Knowledge Session