https://doi.org/10.5194/sp-2024-37
Preprint. Discussion started: 30 September 2024

Distributed Environments for Ocean Forecasting: the role of Cloud Computing

Stefania Ciliberti[1], Gianpaolo Coro[2]

[1]Nologin Oceanic Weather Systems, Santiago de Compostela, Spain
[2]Istituto di Scienza e Tecnologie dell'Informazione "Alessandro Faedo", Area della Ricerca CNR di Pisa, Pisa, Italy

Correspondence to: stefania.ciliberti@nowsystems.eu

Abstract. Cloud computing offers an opportunity to innovate traditional methods for provisioning of scalable and measurable computed resources as needed by operational forecasting systems. It offers solutions for more flexible and adaptable computing architecture, for developing and running models, for managing and disseminating data to finally deploy services and applications. The review discussed on the key characteristic of cloud computing related on on-demand self-service, network access, resource pooling, elasticity and measured services. Additionally, it provides an overview of existing service models and deployments methods (e.g., private cloud, public cloud, community cloud, and hybrid cloud). A series of examples from the weather and ocean community is also briefly outlined, demonstrating how specific tasks can be mapped on specific cloud patterns and which methods are needed to be implemented depending on the specific adopted service model.

1 Introduction

Cloud computing is a specialized form of distributed computing that introduces utilization models for remotely provisioning scalable and measured computing resources (e.g., networks, servers, storage, applications, and services) (Mahmood et al., 2013), offering organizations different benefits for their business services and applications: scalability, cost savings, flexibility and agility, reliability and availability, collaboration and accessibility, innovation and experimentation, and sustainability.

The term originated as a metaphor for the Internet which is, in essence, a network of networks providing remote access to a set of decentralized IT resources. The idea of computing in the cloud traces back to 1961 when the computer scientist John McCarthy publicly proposed: “If computers of the kind I have advocated become the computers of the future, then computing may someday be organized as a public utility just as the telephone system is a public utility.… The computer utility could become the basis of a new and important industry”. While the general public has been leveraging forms of Internet-based computer utilities since the mid-1990s as form of search engines, e-mail services, social media platforms, etc., it wasn’t until 2006 that the term cloud computing emerged, when Amazon launched its Simple Storage Service (Amazon S3) followed by the Elastic Compute Cloud (Amazon EC2) service, enabling organizations to lease computing capacity and storage to run their business applications.

https://doi.org/10.5194/sp-2024-37
Preprint. Discussion started: 30 September 2024

The currently most widely accepted definition of cloud computing is provided by the North American National Institute for Standard and Technology (NIST) (Mell and Grance, 2011): Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources that can be rapidly provisioned and released with minimal management effort or service provider interaction.

NIST further elaborates on cloud computing providing a Cloud Computing Reference Architecture based on five essential characteristics, three service models, and four deployment models.

Cloud computing essential characteristics defined by NIST are reported in Table 1. Cloud computing service models define the types of services that are provided by cloud service providers: Table 2 illustrates the NIST definitions with the corresponding use cases.

Table 1: NIST Cloud Computing Essential Characteristics.

Essential Characteristic	***Description***
On-demand self-service	A consumer can unilaterally provision computing capabilities, such as server time and network storage, as needed automatically without requiring human interaction with each service provider
Broad network access	Capabilities are available over the network and accessed through standard mechanisms that promote use by heterogeneous thin or thick client platforms (Thin client relies mainly on network connection and cloud services to do any processing (Chromebooks are a sort of thin client). Thick clients can function independently - e.g. a traditional PC -).
Resource pooling	The provider's computing resources are pooled to serve multiple consumers using a multi-tenant model with different physical and virtual resources dynamically assigned and reassigned according to consumer demand. Multi-tenancy means that multiple users share the same resources of a cloud provider (e.g. one physical server is used to provide multiple virtual machines belonging to different users). A multi-tenant model guarantees that resources can be shared among multiple users and instances of each tenant are isolated and secured from the other tenants.
Rapid elasticity	Capabilities can be elastically provisioned and released, in some cases automatically, to scale rapidly outward and inward commensurate with demand.
Measured service	Cloud systems automatically control and optimize resource use by leveraging a metering capability

https://doi.org/10.5194/sp-2024-37
Preprint. Discussion started: 30 September 2024

Table 2: NIST Cloud Computing Service Models

Service Model	***Description***	***Use Cases***
Infrastructure as a Service (IaaS)	The capability provided to the consumer is to provision processing, storage, networks, and other fundamental computing resources where the consumer is able to deploy and run arbitrary software	Suitable for organizations that want full control over their infrastructure resources (virtual machines, networks, storage) that want their flexibility in customizing software stack and applications.
Platform as a Service (PaaS)	The capability provided to the consumer is to deploy onto the cloud infrastructure consumer-created or acquired applications created using programming languages, libraries, services, and tools supported by the provider.	Suitable for developers and organizations that want to develop, deploy and maintain applications without the burden of managing the underlying infrastructure (virtual machines, network and storage), that is provisioned and deployed by the providers.
Software as a Service (SaaS)	The capability provided to the consumer is to use the provider's applications running on a cloud infrastructure.	It enables organizations to focus on their core business activities while relying on the expertise and infrastructure provided by the SaaS provider.

Beside the NIST definitions, similar to PaaS another service model is the Serverless model (or Function as a Service - FaaS), that is the capability provided to the user to abstract infrastructure concerns away from applications, where developers can implement application functionality as invokable functions/services whilst providers automatically provision, deploy, and scale these services based on a range of criteria, including efficiency, cost, load balancing, etc. Examples of Serverless/FaaS services are AWS Lambda[1] and Fargate[2], Microsoft Azure Functions[3], Google Cloud Functions[4], Scaleway Serverless Functions[5].

Cloud computing deployment models can be based on different approaches offering organizations options for workload placement, application development, and resource allocation to optimize their cloud strategy based on their needs, cost considerations, performance requirements, compliance regulations and desired level of control. The four cloud computing deployment models identified by NIST are reported in Table 3 with a description and some examples-

[1] https://aws.amazon.com/lambda
[2] https://aws.amazon.com/fargate
[3] https://azure.microsoft.com/en-us/products/functions
[4] https://cloud.google.com/functions
[5] https://www.scaleway.com/en/serverless-functions

https://doi.org/10.5194/sp-2024-37
Preprint. Discussion started: 30 September 2024

Table 3: NIST Cloud Computing Deployment Models.

Deployment Model	*Description*	*Examples*
Private Cloud	Deployment of cloud infrastructure and services exclusively for a single organization or entity. In a private cloud, the computing resources, such as servers, storage, networking, and virtualization technologies, are dedicated to and managed by the organization itself. The infrastructure can be hosted on-premises within the organization's own data centers or in a dedicated off-site facility.	Open source software solutions such as CloudStack[6], OpenNebula[7], Openstack[8], allow organizations to build their own private cloud computing solutions.
Public Cloud	Use of cloud services provided by third-party vendors over the internet. The infrastructure and resources in the public cloud are shared among multiple customers and the cloud service provider is responsible for managing and maintaining the underlying hardware, software, and infrastructure. Users can access and utilize the services on a pay-as-you-go basis, typically through a subscription or usage-based pricing model.	Examples of Public cloud providers are Alibaba[9], Amazon Web Services[10], Google Cloud Platform[11], Hetzner[12], Microsoft Azure[13], Scaleway[14].
Community Cloud	Cloud infrastructure and resources are shared among organizations with common interests, such	EGI[15] is a federation of different European Data Centers providing a cloud infrastructure for research

[6] https://cloudstack.apache.org
[7] https://opennebula.io
[8] https://www.openstack.org
[9] https://www.alibabacloud.com
[10] https://aws.amazon.com
[11] https://cloud.google.com
[12] https://www.hetzner.com/cloud
[13] https://azure.microsoft.com
[14] https://www.scaleway.com/en
[15] https://www.egi.eu

https://doi.org/10.5194/sp-2024-37
Preprint. Discussion started: 30 September 2024

	as industry-specific regulations, security requirements, or collaborative projects. In a community cloud, the infrastructure is designed and managed for the specific needs of the community members, and it allows organizations within the community to share costs, resources, and expertise while maintaining a higher level of control and customization compared to public cloud services.	communities. The European Open Science Cloud (EOSC[16]) is an environment for hosting and processing research data to support EU science, built on top of EGI cloud infrastructure. The European Weather Cloud [17] will deliver data access and cloud-based processing capabilities for the European Meteorological Infrastructure (EMI) and their users.
Hybrid Cloud	Combines both public and private cloud environments to create a unified computing infrastructure, allowing organizations to host some applications or data in a private cloud (i.e. greater control, security and compliance), while utilizing public cloud services for other applications or workloads (i.e. scalability, cost-effectiveness and flexibility for workload burst/on-demand peaks). The hybrid approach provides the ability to address specific requirements, such as regulatory compliance or data sovereignty, by keeping sensitive data within a private infrastructure while utilizing the public cloud for less sensitive workloads.	Netflix[18] uses a hybrid cloud storage solution in order to store and move assets across Amazon AWS S3 and multiple on-premises storage systems.

Beside the cloud deployment models identified by NIST, there are few other approaches that are worth mentioning that provide further capabilities to the organizations that decide to embrace cloud technology.
Multi-cloud computing refers to the strategy of using multiple cloud service providers, allowing organizations leveraging the services of two or more public cloud providers, combining their offerings to build and manage their applications and infrastructure. This approach allows businesses to take advantage of the strengths and capabilities of different cloud providers,

[16] https://eosc.eu
[17] https://www.europeanweather.cloud
[18] https://aws.amazon.com/solutions/case-studies/netflix-storage-reinvent22

https://doi.org/10.5194/sp-2024-37
Preprint. Discussion started: 30 September 2024

such as cost-effectiveness, performance, geographic coverage, or specialized services. It also offers increased flexibility, redundancy, and mitigates the risk of vendor lock-in.

Distributed cloud-edge computing, one of the main innovation streams for cloud computing, combines elements of cloud computing with edge computing, extending the capabilities of the traditional centralized cloud infrastructure by distributing cloud services closer to the edge of the network, where data is generated and consumed, rather than relying solely on centralized data centers. By moving cloud services closer to where data is generated, latency is reduced, and real-time or time-sensitive applications can benefit from faster response times and improved performance. This is especially crucial for applications requiring immediate data processing and low latency. Recently, public cloud providers started to offer pre-configured appliances (e.g. AWS Outpost, Azure Stack) that brings the power of the public cloud to the private and edge cloud, and have defined collaborations with telcos (e.g. AWS and Vodafone, Google and ATT) to create 5G edge services. Furthermore, the main open source cloud management platforms provide extensions (OpenNebula ONEedge, OpenStack StarlingX, Kubernetes KubeEdge) for enhancing private clouds with capabilities for automated provisioning of compute, storage and networking resources and/or orchestrate virtualized and containerized application on the edge.

As of today, cloud computing has been established as a new paradigm for the development and delivery of business and enterprise applications. Cloud-native applications - i.e., building, running, and maintaining apps based on techniques and technologies for cloud computing - provide abstraction from underlying infrastructure and use agile and DevOps methods thanks to the flexibility and agile infrastructure provided by cloud computing technologies. Cloud-native application development is driven by new software models, such as microservices and serverless, and is made possible through technologies such as Linux containers (i.e., Docker) and container orchestration tools like Kubernetes, that are becoming the de facto leading standards for packaging, deployment, scaling and management of enterprise and business applications on cloud computing infrastructures.

Following the rise of Linux container use in enterprise environments, the adoption of container technologies has gained momentum in technical and scientific computing, i.e., high-performance computing (HPC). Containers can address many HPC problems (Mancini and Aloisio, 2015), but unfortunately the mainstream container engine such as Docker was not able to make its mark in HPC centers due to some technical challenges. Several container platforms have been created to address the needs of the HPC community such as Shifter (Jacobsen and Canon, 2015), Singularity (Kurtzer et al., 2017) (now Apptainer), Charliecloud (Priedhorsky and Randles, 2017) and Sarus (Benedicic et al., 2019). Recently, Podman has been analyzed to investigate its suitability in the context of HPC (Gantikow et al., 2020), showing some promise in bringing a standard-based, multi-architecture enabled container engine to HPC.

2 Cloud Computing in Operational Ocean Forecasting

Cloud computing presents an opportunity to rethink traditional approaches used in Operational Ocean Forecasting (Vance et al, 2016), since it can enable a more flexible and adaptable computing architecture for observations and predictions, offering

https://doi.org/10.5194/sp-2024-37
Preprint. Discussion started: 30 September 2024

new ways for scientists to observe and predict the ocean's state. Cloud computing can provide a powerful and agile platform for development and running operational models, for management and dissemination of data, for building and deploy services to downstream business and applications, and finally for analyses and visualization of oceanographic products, enabling researchers to tackle larger and more complex problems without burden of building and maintaining computing and storage infrastructures.

For example, high-performance resources to run operational models can be leased as needed without the necessity of creating and supporting infrastructure, enabling the possibility to collaborate and use the same resources from geographically diverse locations. Large scale datasets related to forecast and observational oceanographic products can be stored in cloud-native storages (e.g., S3 Object Storage) and accessed from any location with public connectivity, enabling data proximate computations (Ramamurthy, 2018) for analysis of large datasets using remote resources (close to data) rather than downloading vast amounts of data locally and needing a local infrastructure in support.

Vance et al. (2019) analyze uses of the cloud for management and analysis of observational data and model results, describe the workflows for running models and streaming observational data, based on the cloud patterns for Atmospheric and Ocean Research identified Butler and Merati (2016). The different cloud patterns are reported in Table 4.

Table 4: Cloud Computing Patterns for Operational Oceanography.

Cloud Pattern	Intent
Cloud-Based Management of Scientific Data – Getting Data From the Cloud	Explores integrating the use of cloud-based data and how scientists can access large volumes of diverse, current and authoritative data, addresses the problem of locating and using large amounts of scientific data.
Cloud-Based Management of Scientific Data – Storing Data in the Cloud	Explores storing and managing data in the cloud. Addresses the problem of ever increasing data quantities with decreasing budgets for data management. Explores the ways scientific projects can meet data access and dissemination requirements
Computing Infrastructure for Scientific Research	Explores the ways in which cloud computing, in the form of PaaS or IaaS could be used as part of a research program and for teaching. It addresses the need for larger computational capabilities, especially under constrained budgets.
Analysis in the Cloud	Explores conducting analyses in the cloud. Addresses the problem of wanting to perform analyses on ever larger datasets and on datasets from multiple sources. Explores the secondary question of ways scientific projects can standardize analysis tools among geographically distributed researchers.

https://doi.org/10.5194/sp-2024-37
Preprint. Discussion started: 30 September 2024

Visualization in the Cloud	Explores creating visualizations using cloud-based tools and making the visualizations available via the cloud. Addresses the need to visualize larger amounts of data.
Results Dissemination in Real Time/Storytelling/Outreach	Explores ways in which cloud-based platforms and tools can be used to reach new audiences. Addresses the need to make research results rapidly available and relevant to a wide variety of audiences—scientific and non-scientific.
Running HPC numerical simulation in the Cloud	Explores ways in which cloud-based platforms and tools can be used to run numerical models used in forecasting services.

In the following, some US and EU programmes, initiatives and projects are reported as examples on how cloud computing technologies and patterns have been used to provide services to the oceanographic and scientific community in general.

2.1 NOAA Open Data Dissemination & Big Data Program

NOAA's Open Data Dissemination (NODD[19]) Program is designed to facilitate public use of key environmental datasets by providing copies of NOAA's information in the Cloud, allowing users to do analyses of data and extract information without having to transfer and store these massive datasets themselves. NODD started out as the Big Data Project in April 2015 (and then later became Big Data Program); NODD currently works with three IaaS providers (Amazon Web Services (AWS), Google Cloud Platform, and Microsoft Azure) to broaden access to NOAA's data resources. These partnerships are designed to not only facilitate full and open data access at no net cost to the taxpayer, but also foster innovation by bringing together the tools necessary to make NOAA's data more readily accessible. There are over 220+ NOAA datasets on the Cloud Service Providers (CSPs) platforms. The datasets are organized by the NOAA organization who generated the original dataset (https://www.noaa.gov/nodd/datasets).

2.2 Copernicus Service and Data and Information Access Services

Copernicus (https://www.copernicus.eu) is the Earth Observation component of the EU Space programme, looking at earth and its environment to benefit all European citizens. Copernicus is generating on a yearly basis petabyte of data and information that draw from satellite Earth Observation and in-situ (non-space) data. The up-to-date information provided by the core services (Atmosphere, Climate Change, Marine, Land, Security and Emergency) are free and openly accessible to users. As the data archives grow, it becomes more convenient and efficient not to download the data anymore but to analyze them where they are originally stored.

[19] https://www.noaa.gov/nodd

https://doi.org/10.5194/sp-2024-37
Preprint. Discussion started: 30 September 2024

To facilitate and standardize access to data, the European Commission has funded the deployment of five cloud-based platforms (CreoDIAS, Mundi, Onda, Sobloo, Wekeo), known as DIAS – Data and Information Access Services - that provide centralized access to Copernicus data and information, as well as to processing tools. The DIAS provides users with a large choice of options to benefit from the data generated by Copernicus: to search, visualize and further process the Copernicus data and information through a fully maintained software environment while still having the possibility to download the data to their own computing infrastructure. All DIAS platforms provide access to Copernicus Sentinel data, as well as to the information products from the six operational services of Copernicus, together with cloud-based tools (open source and/or on a pay-per-use basis). Thanks to a single access point for the entire Copernicus data and information, DIAS allows the users to develop and host their own applications in the cloud, while removing the need to download bulky files from several access points and process them locally.

2.3 Blue-Cloud

The European Open Science Cloud (EOSC) provides a virtual environment with open and seamless access to services for storage, management, analysis and re-use of research data, across borders and disciplines. Blue-Cloud aims at developing a marine thematic EOSC to explore and demonstrate the potential of cloud -based open science for better understanding and managing the many aspects of ocean sustainability (https://blue-cloud.org/news/blue-clouds-position-paper-eosc). The Blue-Cloud platform, federating European blue data management infrastructures (SeaDataNet, EurOBIS, Euro-Argo, Argo GDAC, EMODnet, ELIXIR-ENA, EuroBioImaging, CMEMS, C3S, and ICOS-Marine) and horizontal e-infrastructures (EUDAT, DIAS, D4Science), provides FAIR access to multidisciplinary data, analytical tools and computing and storage facilities that support research. Blue Cloud provides Services through pilot Demonstrators for oceans, seas and fresh water bodies for ecosystems research, conservation, forecasting and innovation in the Blue Economy, and accelerates cross-discipline science, making innovative use of seamless access to multidisciplinary data, algorithms, and computing resources.

References

Benedicic, L., Cruz, F.A., Madonna, A., Mariotti, K.: Sarus: Highly Scalable Docker Containers for HPC Systems. In: Weiland, M., Juckeland, G., Alam, S., Jagode, H. (eds) High Performance Computing. ISC High Performance 2019. Lecture Notes in Computer Science(), vol 11887. Springer, Cham. https://doi.org/10.1007/978-3-030-34356-9_5, 2019.

Butler, K., and Merati, N.: Chapter 2 - Analysis patterns for cloud centric atmospheric and ocean research. Cloud Computing in Ocean and Atmospheric Sciences, 15-34. https://doi.org/10.1016/B978-0-12-803192-6.00002-5, 2016.

Gantikow, H., Walter, S., Reich, C.: Rootless Containers with Podman for HPC. In: Jagode, H., Anzt, H., Juckeland, G., Ltaief, H. (eds) High Performance Computing. ISC High Performance 2020. Lecture Notes in Computer Science, 12321. Springer, Cham. https://doi.org/10.1007/978-3-030-59851-8_23, 2020.

Jacobsen, D.M., and Canon, R.S.: Contain This, Unleashing Docker for HPC. Cray User Group 2015, 14. Available at https://www.nersc.gov/assets/Uploads/cug2015udi.pdf, 2015 (last access: 29/07/2024).

Kurtzer, G.M., Sochat, V., Bauer, M.W.: Singularity: Scientific containers for mobility of compute. PLOS ONE 12(5), 1–20. https://doi.org/10.1371/journal.pone.0177459, 2017.

Mahmood Z., Puttini, R., Erl, T.: Cloud Computing: Concepts, Technology & Architecture. Prentice Hall Press, USA, 1st edition, 2013.

Mancini, M., and Aloisio, G.: How advanced cloud technologies can impact and change HPC environments for simulation. 2015 International Conference on High Performance Computing & Simulation (HPCS), Amsterdam, Netherlands, 667-668. DOI: 10.1109/HPCSim.2015.7237116, 2015.

Mell, P., and Grance, T.: The NIST Definition of Cloud Computing. Gaithersburg, MD: National Institute of Standards and Technology. doi: 10.6028/NIST.SP.800-145, 2011.

Priedhorsky, R., and Randles, T.: Charliecloud: Unprivileged Containers for User-Defined Software Stacks in HPC. SC17: International Conference for High Performance Computing, Networking, Storage and Analysis, Denver, CO, USA, 2017, pp. 1-10, 2017.

Ramamurthy, M. K.: Data-Proximate Computing, Analytics, and Visualization Using Cloud-Hosted Workflows and Data Services. American Meteorological Society, 98th Annual Meeting, Austin, 2018. Available at: https://ams.confex.com/ams/98Annual/webprogram/Paper337167.html, 2018 (last access: 29/07/2024).

Vance T. C., Wengren, M., Burger, E., Hernandez, D., Kearns, T., Medina-Lopez, E., Merati, N., O'Brien, K., O'Neil, J., Potemra, J.T., Signell, R.P., and Wilcox, K.: From the Oceans to the Cloud: Opportunities and Challenges for Data, Models, Computation and Workflows. Frontiers in Marine Science, 6. https://doi.org/10.3389/fmars.2019.00211, 2019.

Vance, T. C. , Merati, N., Yang, C., Yuan, M.: Cloud Computing in Ocean and Atmospheric Sciences, Academic Press. https://doi.org/10.1016/C2014-0-04015-4, 2016.

Competing interests

The contact author has declared that none of the authors has any competing interests.

Acknowledgements

The Authors are grateful to Dr. Marco Mancini for sharing his expert view in support of this contribution.