AWS Architecture Monthly

Sustainability
August 2021

Editor's note

In 2019, Amazon and Global Optimism co-founded The Climate Pledge, a commitment to reach the Paris Agreement 10 years early and be net-zero carbon by 2040. Now 100+ organizations have signed The Climate Pledge, sending an important signal that demand will rapidly grow for products and services that help reduce carbon emissions. We are on a path to powering our operations with 100% renewable energy by 2025, five years ahead of its original goal of 2030. This is a commitment and investment in sustainability because it's good for the planet, for business, for our customers, and for our communities. Read more at https://www.aboutamazon.com/planet.

This issue of the magazine includes articles that span across industries, reference architectures to help you build sustainably, and videos and case studies to inspire. We hope you find it useful.

We'd like to thank our experts, Margaret O'Toole, Worldwide Tech Leader for Environmental Sustainability, and Joseph Beer, Worldwide Tech Leader - Power and Utilities, for their contributions. Thanks also to Elise Greve, Principal Product Marketing, AWS Sustainability, and our AWS Media Series Technical Editor, Bonnie McClure.

Please give us your feedback! Include your comments on the Amazon Kindle page. You can view past issues and reach out to aws-architecture-monthly@amazon.com anytime with your questions and comments.

Jane Scolieri, Managing Editor

Table of Contents:

- **Executive Summary**: Elise Greve, Principal Product Marketing, AWS Sustainability
- **Ask an Expert**: Margaret O'Toole, WW Tech Leader for Environmental Sustainability, and Joseph Beer, WW Tech Leader, Power and Utilities
- **Case Study**: ENGIE Builds the Common Data Hub on AWS, Accelerates Zero-Carbon Transition
- **Whitepaper**: Blockchain for Sustainability
- **Blog**: Intertrust uses AWS to track climate change
- **Reference Architecture**: Meter Data Analytics Platform for Utilities
- **Quick Start**: Utility Meter Data Analytics on AWS
- **Blog**: How the cloud is helping us better understand and manage the oceans
- **Case Study**: Saildrone Explores the Oceans with Machine Learning
- **Reference Architecture**: DERMS on AWS for Utilities
- **Blog**: Forecasting energy usage using Amazon machine learning and data lakes
- **Case Study**: CropX Runs Soil Sensor Application on AWS to Help Global Farmers Enable Sustainability
- **Whitepaper**: The Carbon Reduction Opportunity of Moving to Amazon Web Services
- **Blog**: Open data on AWS supports sustainable agricultural practices and crop optimization
- **Additional Resources**
- **Videos**: Building Sustainably, Carbon footprinting, Carbon Lighthouse, Alexa energy dashboard, PayGo, CropIn

Notices

Executive Summary

Further and Faster, Together

The focus of this month's issue of Architecture Monthly is sustainability, which follows the publishing of our third Amazon Sustainability report. As outlined by Kara Hurst—Amazon Vice President Worldwide Sustainability—the work of sustainability can be accomplished with passion and partnership as we go "Further and Faster, Together." To create lasting and systemic change on our planet, we are focused on driving change in partnership with others.

In that spirit, we bring you this edition of Architecture Monthly. It emphasizes resources for your organization to learn, implement, and collaborate on methods and insights to sustainability. Start with the Building Sustainably on AWS video presented by Luke Hargreaves, AWS Senior Solutions Architect. Luke shares our science-based, customer-centric approach to integrate sustainability throughout Amazon business operations, reviews how utilizing cloud resources vs. on-premises can improve your carbon footprint, and how to integrate sustainability into workloads through optimization and informed architecture patterns. Don't miss reading through the Ask an Expert section for sustainability solutions in addition to tools to implement sustainability practices into IT design.

Next, dive into examples of AWS customers integrating sustainability into their businesses. This includes power and utility companies building systems to supply renewable energy and an agricultural analytics company addressing the challenge of water conservation and fertilizer distribution to crops.

Go deeper into the topic and explore our Blockchain for Sustainability whitepaper or review reference architectures for power and utility meter data analytics and management.

Lastly, we know that data is a significant factor in an organization's sustainability journey, whether it's collecting your own data or access-

ing other data resources for modeling or benchmarking. We include customer stories that use the Amazon Sustainability Data Initiative (ASDI), which seeks to accelerate sustainability research and innovation by minimizing the cost and time required to acquire and analyze large sustainability datasets. ASDI offers datasets from scientific and research organizations such as the National Oceanic and Atmospheric Administration (NOAA), National Aeronautics and Space Administration (NASA), and the UK Meteorological Office (Met Office) to identify, host, and deploy key datasets on AWS Cloud, which are 100% free to access. We also offer additional datasets through AWS Data Exchange, which aims to support sustainability by making it easy to find, subscribe to, and use third-party sustainability-related data in the cloud.

Our hope is that this edition deepens your knowledge of managing and architecting for sustainability as we go ***further and faster, together.***

About the Author

Elise Greve is a Principal Product Marketer for AWS Sustainability. She focuses on helping customers identify, implement, and deploy AWS technology solutions to drive their corporate sustainability journey. Elise has spent the last 15+ years bringing new products to market and invests her energy and experience in innovation so people can work better, enjoy life more, and be thrilled by new products and services. She's based in Seattle and is an avid diver and fish geek.

Ask an Expert:

Margaret O'Toole, Worldwide Tech Leader - Environmental Sustainability
Joseph Beer, Worldwide Tech Leader - Power and Utilities

When putting together an AWS architecture to solve business problems specifically for sustainability-focused customers, what are some of the considerations?

A core idea of sustainability comes down to efficiency: how can we do the most work with the fewest number of resources? In this case, you want efficiency when you design and build the solution and also when you apply and operate it.

In broad strokes, there are two main things to consider. First, you want to optimize technology usage to reduce impact. Second, you want to find and use the best mix of technology to support sustainability. These objectives must also delight your customers, constituents, and stakeholders as you meet your business objectives in the most cost effective and expeditious way possible.

However, to be successful in combining technology and sustainability, you must consider the culture change of the sustainability transformation. Sustainability must become part of each person's job function. When it comes to responsibility around sustainability at AWS, we think about it through two lenses.

First, we have the sustainability OF the AWS Cloud, which is our responsibility at AWS. This covers the work we do around purchasing renewable energy, operating efficiently, reducing water consumption in the data centers, and so on. There is more information on sustainability of the AWS Cloud on our sustainability page.

Then, there's sustainability IN the cloud, which focuses on customers and their AWS usage. This is again focused on efficiency, mostly how to optimize existing patterns of user consumption, data access, software and development patterns, and hardware utilization.

In a related but slightly different vein, we also talk about sustainability THROUGH the cloud. This is how our customers use AWS to work on sustainability projects that help them meet their sustainability goals. This can include anything from carbon tracking or accounting to route optimization for fleets to using machine learning (ML) to reduce packaging and anything in between.

What are the general architecture pattern trends for sustainability in the cloud?

Solutions designed with sustainability in mind aim to be highly efficient. An architect wanting to optimize for sustainability looks for opportunities within user patterns, software patterns, development/test patterns, hardware patterns, and data patterns.

There is no one-size-fits-all way to optimize for sustainability, but the core themes are maximizing utilization and reducing waste or duplication. Most customers start with relatively easy things to accomplish. These typically include things like using the AWS Instance Scheduler to turn off compute when it will not be used or comparing cost and utilization reports to find hot spots to reduce utilization.

Another way to optimize for sustainability is to incorporate services managed by AWS (for example, Amazon Managed Streaming for Apache Kafka (Amazon MSK), Amazon EMR, and Amazon Connect) and serverless (for example, AWS Lambda and Amazon API Gateway services as much as possible. These services not only increase the speed and efficiency of your design and build time and lower your overhead to run, but they also include automatic scaling as part of the service, which increases compute

efficiency. Where these services are not applicable, you can often configure automatic scaling into the solutions themselves. Automate everything, including your continuous integration and continuous delivery/deployment (CI/CD) code pipeline, data analytics and artificial intelligence (AI)/ML pipelines, and infrastructure builds where you are not using AMS.

And finally, include ongoing AWS Well-Architected reviews and continuously review and optimize your usage of AMS and the size and mix of your compute and storage in your standard operating procedures.

What are the key AWS-based sustainability solutions you are seeing customers ask for across industries and unique to specific industries?

Almost all industries have a set of shared challenges. This generally includes things like facilities or building management, design or optimization, and carbon tracking/footprinting. To help with this, customers must first understand the impact of their facilities, operations, or supply chain. Many customers use AWS services for ingestion, aggregation, and transformation of their real-world data. Once the data is collected and customers understand their relative impact, this data can be used to form models, which act as the basis for optimization. Technologies such as AWS IoT Core, Amazon Managed Blockchain, and AWS Lake Formation are crucial here.

For industries like power and utilities, there are more targeted solutions. Many of these are aimed at supporting grid decarbonization and the transition to electric vehicles (EVs). Smart EV charging, for example, uses the power of AWS Cloud and AI/ML to lessen the aggregate impact to the grid that may occur because of EV charging peaks and ramp ups. This helps avoid running natural gas generation at peak times. Amazon Forecast, a fully managed service that delivers highly accurate forecasts, can be useful in the case of short-term electric load forecasting. The AWS Cloud enables running frequent grid voltage optimization (aka Volt-VAR Optimization (VVO)) calculations quickly and affordably, which allows utilities to lower their carbon footprint.

Within supply chains, customers use AWS to support traceability and carbon dashboarding to nudge suppliers toward greener energy. Customers commonly look for ways to track and trace throughout their supply chains, either to measure and reduce scope 3 emissions or to optimize their logistics network.

What's your outlook for sustainability, and what role will the cloud play in future development efforts?

The cloud is critical to solving sustainability challenges that businesses and governments are being challenged with right now. It gives you the flexibility to use resources only when you need them, coupled with immense computing power. Thus, the cloud will be an essential tool in solving many data challenges like reporting and measuring and predicting and analyzing trends.

Migration to the cloud is essential to optimizing workloads and handling massive amounts of data. We can see this directly in how Boom used AWS HPC to support the creation of the world's fastest and most sustainable aircraft. Additionally, FLSmidth is pursuing sustainable, technology-driven productivity under MissionZero. This initiative is working to achieve zero emissions and zero waste in cement production and mining by 2030 with the help of AWS high performance computing (HPC).

Do you see different trends in sustainability in the cloud versus on-premises?

The usage pattern is different. With the cloud you can use what you want, whenever you want, which allows for customers to drive up a high utilization. This type of efficiency is critical. It's why 451 Research found that the same task can be completed on AWS with an 88% lower carbon footprint compared to the average surveyed US enterprise data center.

The cloud offers technology that wouldn't be available on premises, such as large GPU-backed instances capable of processing huge amounts of data in hours that would take weeks on premises. It can also ingest massive streams of data from energy- and resource-consuming and producing assets to optimize their performance and environmental impact in near-real-time.

With the cloud, you have the flexibility and the power to move quickly through research and development to solve sustainability challenges. You can accelerate the development process of new ideas and solutions, which will be essential for the transformation to a carbon neutral, climate positive economy.

About the experts

Margaret O'Toole is the tech leader for environmental sustainability at AWS. She oversees customer-facing technical initiatives in sustainability and acts as liaison between the AWS Service and Infrastructure teams and the technical field community. Her goal is to make it easier for customers of any size to solve challenges in their sustainability transformations by using the AWS Cloud. Margaret has been with AWS since 2017.

Joseph (Joe) Beer is the AWS Worldwide Technology Leader for the Power and Utilities vertical. He is responsible for advising senior leadership within our customers and partners on how AWS Cloud can enable their digital transformation strategies to help them achieve their business goals. He also guides the development of business-outcome driven reusable IT, OT, Customer Engagement, Data, and Asset Management AWS-based solutions for the Utility industry.

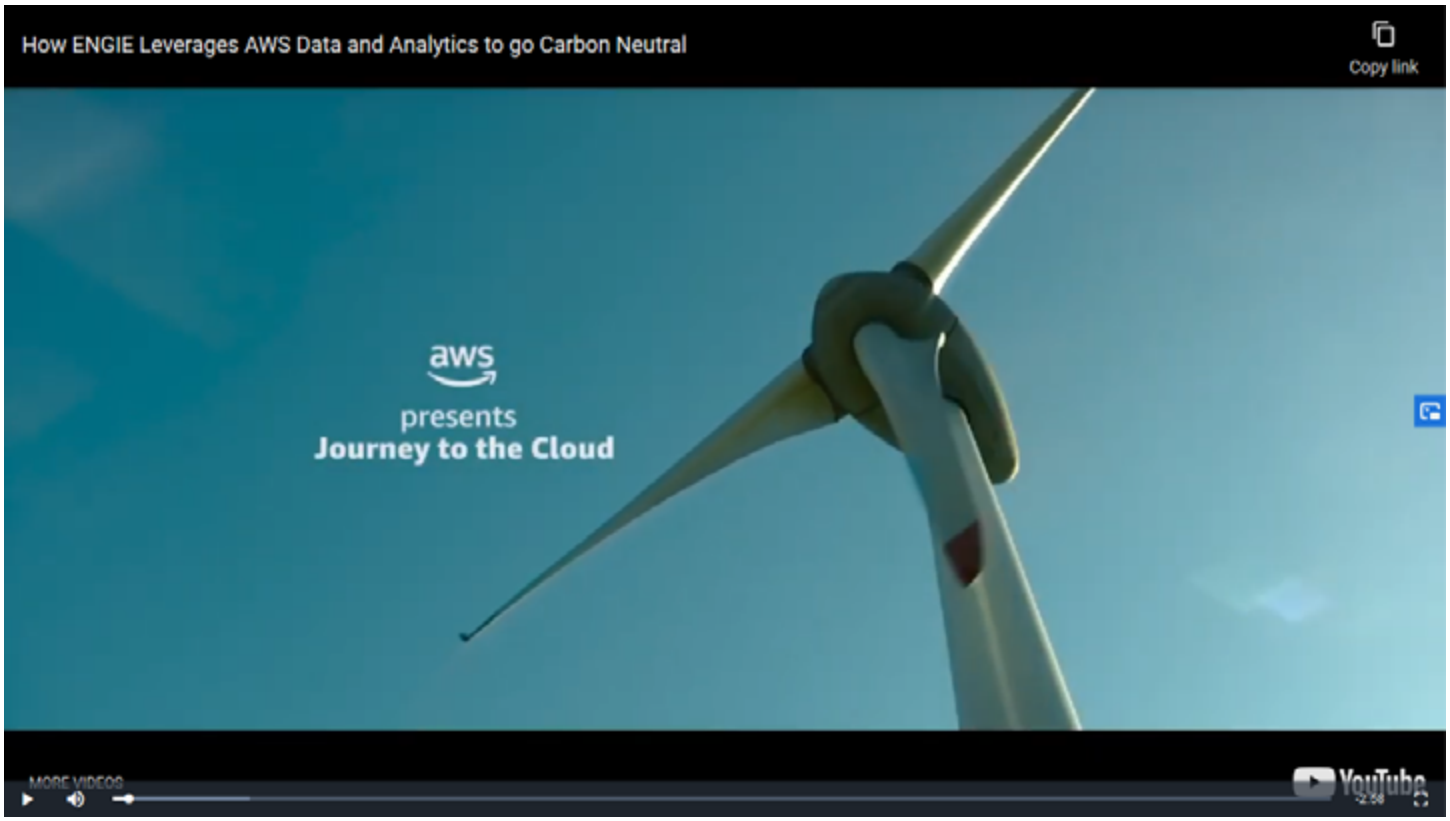
Case Study

ENGIE Builds the Common Data Hub on AWS, Accelerates Zero-Carbon Transition

2021

ENGIE—one of the largest utility providers in France and a global player in the zero-carbon energy transition—produces, transports, and deals electricity, gas, and energy services. With 160,000 employees worldwide, ENGIE is a decentralized organization and operates 25 business units with a high level of delegation and empowerment. ENGIE's decentralized global customer base had accumulated lots of data, and it required a smarter, unique approach and solution to align its initiatives and provide data that is ingestible, organizable, governable, sharable, and actionable across its global business units.

In 2018, ENGIE decided to accelerate its digital transformation through data and innovation by becoming a data-driven company. First, ENGIE wanted to build an enterprise data repository named the Common Data Hub to align its customers and business units around the same solution. The Common Data Hub helped ENGIE's business units easily ingest, store, share, and consume datasets through a unified platform and highly secured environment, ultimately enabling the company to increase productivity, make accurate energy-production predictions, and bring new services to customers.



ENGIE used Amazon Web Services (AWS) to create the Common Data Hub, a custom solution built using a globally distributed data lake and analytics solutions on AWS. The Common Data Hub empowers teams to innovate by simplifying data access and delivering a comprehensive set of analytics

tools. AWS Professional Services supported ENGIE in designing and implementing the solution and putting in place an internal service team (called the data@ENGIE team) that is in charge of evolving and operating the Common Data Hub platform.

Identifying a Need for Smarter Data

ENGIE's customers vary widely, from cities to retail customers to large companies and beyond. The company increasingly supports customers' ability to generate their own electricity with decentralized assets, including solar panels and wind farms. As ENGIE trended further toward decentralization, it found that its Systems, Applications, and Products in Data Processing enterprise resource planning software needed updating. The company needed a uniform method of collecting and analyzing data to help customers manage their value chains. "We need to use data in order to measure consumption and anticipate the volumes of electricity production depending on the weather forecast, for instance," says Gregory Wolowiec, the technology team leader who guides the development and delivery of ENGIE's data programs. Wolowiec also cites problems of isolation and inconsistencies across countries: "All the solutions were different from one country to another; there was no sharing between the different parts of the organization. It became very important for us to be able to collect and share the data in a streamlined way everywhere on the planet."

Yves Le Gélard, chief digital officer at ENGIE, explains the company's purpose: "Sustainability for ENGIE is the alpha and the omega of everything. This is our raison d'être. We help large corporations and the biggest cities on earth in their attempts to transition to zero carbon as quickly as possible because it is actually the number one question for humanity today."

ENGIE's group chief data officer, Gérard Guinamand, adds, "Our strategy when it comes to data is actually directly linked to our purpose. If you want to drive and execute on a zero-carbon transition, you need first to gather data on what's happening. That includes data on how much carbon dioxide you burn, where you burn it, and how it all correlates with environmental topics like the weather, the temperature, and the number of people. All this data needs to be stored, gathered, and computed so that you can measure progress and follow a road map."

Whatever method ENGIE adopted needed to have a high level of security and be compliant with regulations all over the world. As the company put together a proof of concept, it explored a variety of solutions with local and global cloud providers. "We were convinced that AWS was a good solution for many reasons, including the cost model—and especially in terms of data storage," says Wolowiec. So ENGIE began undertaking its large data project on AWS in mid-2018.

Developing the Common Data Hub and Deploying It around the World

Wolowiec describes the Common Data Hub as "the collaborative and distributed data lake that enables ENGIE to store data, share data, and create value with data." It was built using Amazon Simple Storage Service (Amazon S3), an object storage service that offers industry-leading scalability, data availability, security, and performance. The solution also uses Amazon Redshift, a fully managed, petabyte-scale data warehouse service in the cloud that can query semistructured data in the Amazon S3 data lake, demonstrating the lake house approach to data warehousing.

Because the solution uses Amazon Redshift, customers can deploy data warehouses securely in their Common Data Hub environments and make use of analytics. The company's business unit administrators, by managing Amazon Redshift clusters on the Common Data Hub, can be added to Common

Data Hub projects and easily access datasets in the Amazon S3 data lake, as well as build valuable insights from the Common Data Hub's rich datasets catalog. The Common Data Hub uses Amazon Redshift for two different types of data access. It uses Amazon Redshift Spectrum for direct query to the Common Data Hub's S3 buckets and uses Amazon Redshift as a provisioned data warehouse with its own internal storage. "We rely on the Amazon Redshift Spectrum feature to make the link between the Amazon S3 data lake managed by the Common Data Hub and the Amazon Redshift data warehouse," says Wolowiec.

Other AWS services involved in the Common Data Hub include Amazon Kinesis Data Streams (Amazon KDS), a massively scalable and durable real-time data streaming service. Using Amazon KDS enables ENGIE to easily collect, process, and analyze streaming data from Internet of Things devices in real time, meaning ENGIE can quickly gather the information it uses to develop insights. AWS Glue—a fully managed extract, load, transform (ETL) service for a metadata repository—further helps transfer and clean the data. Amazon Athena, an interactive query service that makes it easy to analyze data in Amazon S3 using standard SQL, lets ENGIE business units view data. And to glean further insights from data, ENGIE relies on Amazon SageMaker, a fully managed service that provides every developer and data scientist with the ability to build, train, and deploy machine learning models quickly.

To facilitate smooth and easy adoption of the Common Data Hub all around the world, ENGIE provided acceleration templates and documentation to help its business unit administrators see the value of the data they collect and access the data in the distributed data lake. The Common Data Hub also enables high levels of data governance and security. Data producers can share and control access to datasets and workflows, and consumers can request access to and consume data.

The integration of AWS services provided a secure, agile, and scalable solution for ENGIE. Now different business units can use the framework in the ways they need without sacrificing anything

essential to operations. Ease of use and automation let ENGIE business units increase productivity quickly after building the Common Data Hub solution on AWS. There are also positive environmental impacts: ENGIE uses data to obtain the maximum amount of energy possible from wind farms, helping increase the efficacy of this important renewable energy source. "We provide the right tools so that entities can focus on value creation instead of taking time to deal with technical issues," says Wolowiec. As of July 2020, ENGIE had collected 95 TB of data set up in the Common Data Hub.

Facilitating a Standardized Top-Down Approach

The Common Data Hub forms the backbone of ENGIE's data-driven strategy by enabling data community between information technology and business users, accelerating increased data literacy at every level of ENGIE and helping optimize internal processes or create new data-driven services. All business units are now empowered with a unique solution to build data-driven applications faster. ENGIE currently has more than 351 projects set up on the Common Data Hub going on across the world. The Common Data Hub offers a truly cohesive solution since it eliminates silos and enables every department to benefit from equal access to the common framework.

With its new method of collecting and sharing data, ENGIE sees an opportunity to change how it does business—and the company is in the process of building a vertical data hub to do just that. ENGIE has historically had a bottom-up approach, with its business units providing services for customers in their respective regions. However, since many of its energy services are the same, this results in an unnecessary duplication of work. "Our electricity generation activity and especially our renewable energy generation are basically the same everywhere," says Wolowiec, "We can use the Common Data Hub to build common use cases around the world. Next for us is to introduce more and more top-down approaches—especially for wind farms."

ENGIE discovered significant value by using AWS services to build its Common Data Hub, enabling its global business units to collect, share, and analyze data in more productive ways. ENGIE's business units still retain autonomy, but they can now benefit from the strengths of centralized data, garnering important insights from similar use cases as they discover newer, more efficient ways to power the world.

Read case study online

About ENGIE

ENGIE is a global energy company with 25 business units operating worldwide. The company powers millions of customers and develops integrated solutions throughout the value chain to support corporations' and local authorities' zero-carbon transition.

Benefits of AWS

- Collected 95 TB of data across 351 projects
- Automated energy predictions
- Increased business unit productivity
- Maximized wind farm energy production

Whitepaper

Blockchain for Sustainability

Exploring the application of Amazon Managed Blockchain to improve transparency in supply chains, enable real-time auditability, and reduce carbon footprints

December 22, 2020

Introduction

Sustainability challenges span social, economic, and environmental considerations. Both public and private organizations increasingly seek to provide a more transparent view into their sustainability practices for employees, shareholders, regulators, and consumers.

To meet sustainability commitments, for example carbon emission targets, organizations must measure the impact of their entire supply chain. However, challenges with transparency and data reconciliation in supply chains make this difficult. In a typical supply chain, individual participants are able to see one level upstream and downstream, but not beyond that.

Efforts to improve product and facility audits and certifications can be vulnerable to tampering, and often increase overhead. In addition, many audits are conducted annually, providing limited, snapshot-in-time insight into suppliers' practices.

Although premium commodities, such as organic materials, can be tracked through the supply chain, tracking is typically done at a mass balance level, and not at an individual product level. This type of tracking leaves room for double spending of premium materials, as well as an inability to make certain product claims about the materials it contains.

Blockchain enables visibility into the entire supply chain, allowing organizations to:

- accurately track the provenance and authenticity of materials and goods at the product level.
- gain insight into the sustainability practices of participants.

In this document, we'll share some benefits of blockchain and explore how it can contribute to sustainability goals in the context of three use cases:

- Transparent supply chains and the circular economy.
- Auditing and certifications at scale.
- Carbon emission tracking and marketplaces.

Benefits of blockchain

A blockchain is a decentralized network where governance and operation are maintained by a loosely connected network of members. Blockchains enable parties to transact without the need of an intermediary providing safeguarding services such as maintaining escrow. This approach is enabled by cryptographic algorithms that ensure data can never be tampered with or deleted once committed to the blockchain. Data can be verified at any time for accuracy and authenticity, and with applications designed to run on a blockchain, process integrity can be verified.

Data is stored in a ledger in an append-only manner, ensuring it can't be modified or deleted. All blockchain activity is recorded to the ledger within blocks of data, enabling anyone to replay the transactions and verify that the contents of the ledger have not been manipulated. This feature is a tremendous benefit during audits, as the full ledger history can be provided to an auditor to verify the transactional integrity.

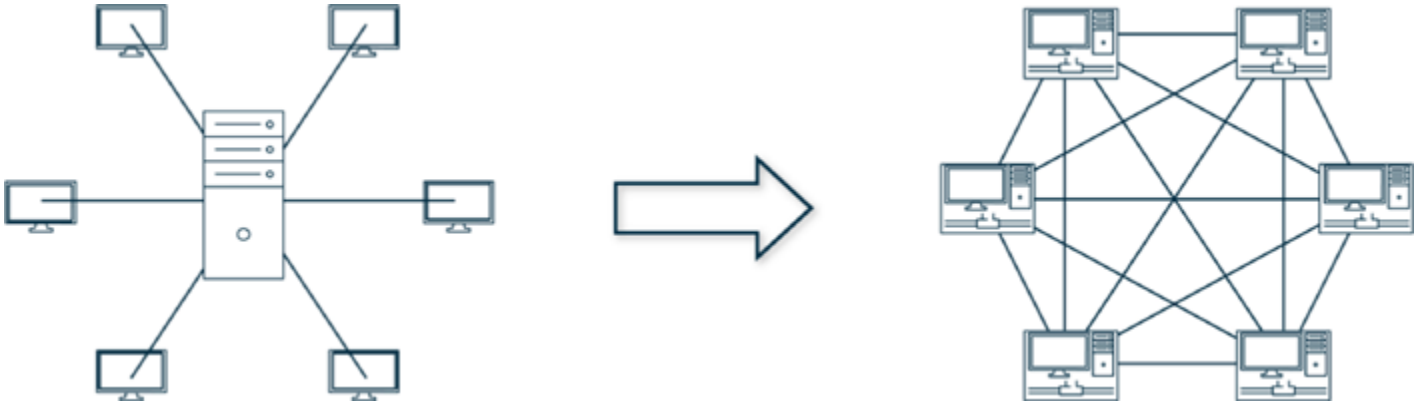


Figure 1 – Blockchain enables transactions without a centralized intermediary

These same foundational benefits underpin the motivation to use blockchain across a variety of industries and use cases. For example, using blockchain to trace a raw commodity in a supply chain provides a complete, transparent, and trusted history of its journey as it moves through the supply chain. As the commodity is transferred between counterparties, additional data can be captured on the blockchain, such as who owned the commodity at various times. Additional data can also include environmental data captured from sensors, cameras, and Internet of Things (IoT) devices.

Bitcoin was the first blockchain network, launched in 2008 as a way to exchange monetary value in a decentralized environment with no intermediary required. The Ethereum blockchain launched in 2015 and included support for smart contracts, something that wasn't available on Bitcoin. Smart contracts are applications containing business rules (for example, terms and agreements) that are deployed and executed on the blockchain. From its inception, Ethereum has had a vibrant and active developer ecosystem contributing improvements and new concepts, including tokenization, also discussed in this document.

Public blockchain network basics

A *public blockchain network*, also referred to as a *permissionless network*, can be used by anyone. Although there are many public blockchain networks available, the Bitcoin and Ethereum blockchain networks are widely considered as the leading public blockchains, based on percentage of global blockchain transactions. Blockchain members running the network collect fees for processing transactions and minting new blocks on the blockchain.

Public blockchain networks enable low barriers to onboarding because no central authority is required to grant network access. This allows counterparties to quickly begin using the network to share data and deploy smart contracts. Users pay for their transactions as they create them, with members operating the network collecting the transaction fees. The cost of transacting on public Ethereum fluctuates greatly, as it strongly correlates with other activity occurring on the network. For example, the cost of executing a simple transaction on Ethereum fluctuated by a factor of 10x between July and September 2020.

Given the low barrier to onboarding, along with the low transaction throughput, public blockchains are ideally suited for storing data that is updated infrequently, and read many times. An annual facility audit report is a good example. An auditor would publish an audit report and put the report on the public blockchain. Once there, anyone can access this report. Public key infrastructure (PKI) can provide even stronger verification, allowing a reader to verify the

authenticity of the publisher.

Challenges with public blockchain

Public networks have significant limitations making them unsuitable for many enterprise uses:

- They are unable to process transactions at high speed, limiting concurrent usage on the network.
- Fluctuation and risk in the cost of issuing transactions make it difficult to project costs.
- Public availability of all data on a public blockchain cause data privacy challenges. It's possible to privatize the data via encryption, but that skews toward a centralized model where requests for decrypting data must pass through a trusted party that owns the decryption key.

The Bitcoin and Ethereum networks are powered by miners who are rewarded for ensuring the integrity of the network, and solving an extremely difficult math problem. The math problem requires significant computing power, leading to greatly increased energy consumption.

There are alternatives to the math problem approach, including incentivizing good behavior to maintain the integrity of the network. Ethereum has been working on such a transition away from high energy consumption with an algorithm called Proof of Stake, which it began beta testing in 2020. Once fully transitioned, Ethereum's eco-footprint will be reduced by 99%.

Understanding private blockchain networks

Private blockchain networks were developed to provide the benefits of blockchain while meeting enterprise requirements. Private blockchains typically have native constructs for data privacy, and they support significantly higher transaction throughput compared to public blockchain networks.

Data privacy enables a subset of members of a blockchain network to transact with one another without revealing details of their transaction to other network members. All members, including those not involved in the private transaction, are still involved in endorsing the transactions and enforcing the business rules within the smart contracts, providing a high level of trust in the network, independent of the privacy level of individual transactions.

Private blockchains enable transaction throughputs on the scale of hundreds to thousands per second, compared to dozens per second on a public blockchain on the high end. This increase in throughput can be attributed to the difference in number of blockchain members needed to achieve consensus. Public blockchain networks can have thousands of members who need consensus on their shared state, while private blockchain membership numbers typically range on a scale from a handful to dozens, leading to significantly shortened times for all members achieving consensus.

Given the higher throughput than public blockchain networks, and their native support for data privacy, private blockchain networks are ideally suited for powering transparent supply chains. Some of the relevant use cases include:

- Tracking premium materials (for example, organic cotton) from their source and throughout the supply chain, ensuring quality in finished products.
- Product safety and recall by tracking goods as they move through a supply chain, providing the ability to trace a faulty product upstream to proactively identify other potentially faulty products.

Read full whitepaper online

Blog

Intertrust uses AWS to track climate change

by Brandon Zupancic

Higher temperatures, wilder weather, melting glaciers, rising sea levels, and thawing permafrost represent just a few of the shifting conditions on Earth that signal climate change. Carbon dioxide (CO2) or greenhouse gas emissions from daily human activities are major contributing factor to climate change. Between transportation, electricity production, industry, agriculture, forestry, and more, the US alone emits an estimated 6,677 million metric tons of CO2 equivalent a year.

In response, companies around the world are pledging to implement more sustainable practices. For example, Amazon co-founded The Climate Pledge in 2019, a demonstration of its commitment to achieve net zero carbon across Amazon's business by 2040, ten years ahead of the Paris Agreement. International software technology company Intertrust, an AWS Partner, is working to help the world better understand the impacts of CO2 emissions and how subtle changes in daily activities could lead to positive change . Using Amazon Web Services (AWS), it's developed View 2020, a tool designed to help companies, communities, scientists, and countries to track global climatology trends and gauge air quality on a granular level.

View 2020 is a set of online dashboards that track and monitor a range of climate and air quality data over time. Using interactive maps and graphs, the View 2020 Air Quality Dashboard captures fluctuations in global air quality over the last four years. These visualizations effectively show the stark impact that the COVID-19 pandemic has had on global air quality levels. The climate team at Intertrust will soon be updating the Air Quality Dashboard to include fine grained station-based air quality data. In addition to air quality, the View 2020 Climatology Dashboard charts global climatology dating back to the 1970s. Users can filter by date, geography and climate variable to effectively understand how our climate has changed over time, both globally and on a country-by-country basis.

"View 2020 is a data-driven application that provides a realistic view of our changing climate. This insight is sorely needed in a world where the climate crisis is increasingly politicized and sensationalized by the media. The tool is deliberately easy-to-use, making it comprehensible for people regardless of their level of climate knowledge," said Kezia Wright, Head of Climate and Sustainability at Intertrust.

View 2020 is built on Intertrust's Planet OS Datahub, a climate and weather data catalog. Intertrust's Planet OS Datahub uses Amazon Simple Storage Service (Amazon S3) for data storage, which provides the reliability and scalability to grow the platform in technical complexity and in bare data volumes, according to Andres Luhamaa, Data Integration Engineer and Climate Data Expert at Intertrust. View 2020 uses data from Planet OS's Datahub to populate the Dashboards. Both Planet OS Datahub and View 2020 are ultimately designed to make climate data easier to access and use.

Planet OS Datahub has revolutionized how companies access climate data. Previously, an organization looking to access a range of weather variables would have to jump through hoops to obtain it. For example, if an energy company wanted to know the wave height and wind speed across an international network of wind farms, the company might approach individual data providers for each geographical location in question, then piece together the data. Planet OS Datahub and View 2020 eliminate these steps by providing a curated catalog of more than 2,000 climate and weather variables from data providers including Copernicus (European Union's Earth observation program), NASA, NOAA, Global Systems Laboratory and Meteo-France. Data is made available via a single application programming interface (API), and is searchable by geography, time frame, and climate variable.

Australian agricultural technology developer The Yield, uses data from Intertrust's Planet OS DataHub catalog to develop tools that help farmers increase crop yield and aid environmental sustainability, while also reducing waste and mitigating the risk and costs associated with bad weather. Planet OS Datahub is also used by energy companies who need access to historical and forecast weather

data. For example, RWE Renewables' Gwynt y Môr, one of the world's largest offshore wind farms, has been able to streamline operations and maintenance using data from the Planet OS Datahub, in combination with a cloud-based tool that delivers an interactive view of the wind turbines. With these technologies, RWE Renewables can acquire and visualize weather data from multiple operational systems including IoT data, turbine SCADA, and observed and predicted weather conditions.

Intertrust also collaborates closely with the AWS Open Data team to mine relevant datasets from the AWS Open Data archives as it curates open datasets across View 2020 and Planet OS Datahub. Symbiotically, the Planet OS Datahub dataset is mirrored on the AWS Open Data archive. Together, the two envisioned one of the most valuable, in-demand weather datasets today dubbed ERA5, which they've highlighted through open code examples, notebooks and blog posts.

"Our collaboration with AWS is essential to both Planet OS Datahub and View 2020. We relay feedback from our users to the AWS Open Data team where relevant, and also accept insights from AWS users as well, all which help drive our platforms forward," said Eneli Toodu, Data Integration Engineer at Planet OS. "By engaging with AWS Open Data users, we're able to learn more about their data usage patterns and harness those insights to optimize datasets for more efficient consumption, like providing an alternative to NetCDF – ZARR formatting which that fits cloud-based access models better."

As Amazon strives to build a more sustainable business for its customers and the planet, the data that Intertrust has made available via the Planet OS Datahub and View 2020 will be instrumental. As Kezia concludes, "The data aggregated by Planet OS and View2020 has the potential to inform a global response to the climate and biodiversity crises. But data is only one part of the picture, we also need action – urgent action. So, we encourage people to use Planet OS to develop useful and creative solutions to global environmental issues."

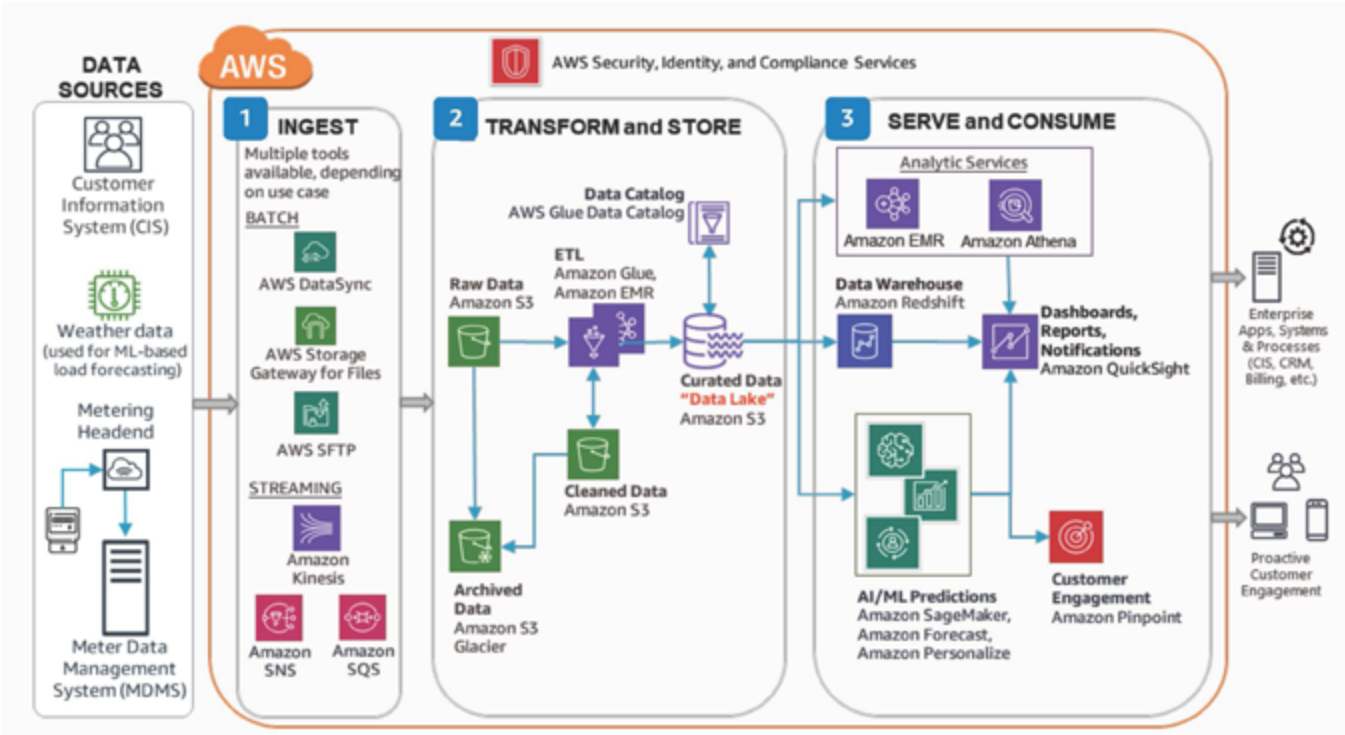
To learn more about Amazon S3, please visit https://aws.amazon.com/s3/

Read blog post online

Reference Architecture

Meter Data Analytics Platform for Utilities

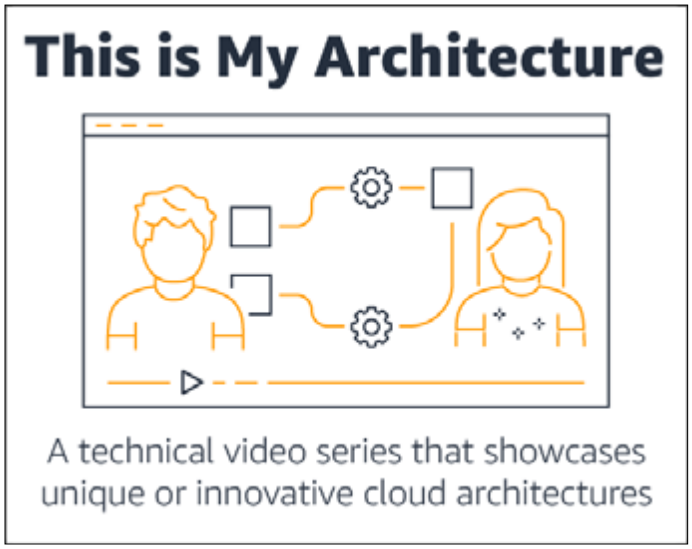
Build a modern data platform and democratize your meter data to gain operational and customer insights. Unlock your data silos and use the right data stores, analytics and AI/ML tools for every job. Detect meter and distribution circuit anomalies, run circuit balancing, thwart energy theft, predict demand run circuit balancing and enhance customer engagement with proactive and meaningful analytics and ML-based forecasts and predictions.

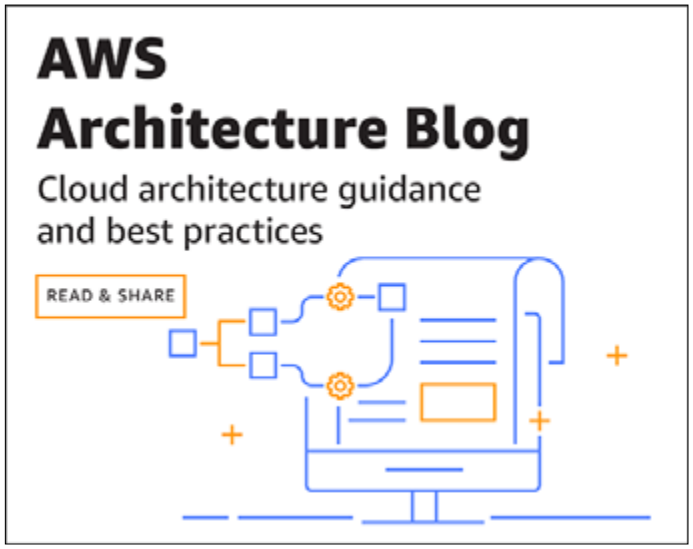


1. Use a variety of AWS tools and services to ingest customer and meter data – both batch and streaming, depending on use case.
2. With **AWS Glue**, automate ETL processes such as file transforms and deduplication, or value add processing such as running meter data VEE process and creating billing determinants. Use **Amazon S3 Glacier** as low cost storage for archival copies and retention compliance. Your final curated data sets are stored in an S3 bucket; your **Data Lake** serves as the "single source of truth" for downstream analytics and ML work. **AWS Glue** also automates the process of data schema discovery and metadata tagging to create a metadata catalog that makes all data visible and searchable. A multitude of **AWS Security, Identity, and Compliance** services are available to you to keep your data safe and secure.
3. Query petabytes of structured, time-series and semi-structured data across your data warehouse and your data lake using standard SQL with **Amazon Redshift**. Perform complex analytics with **Amazon EMR** and ad-hoc data discovery and query against your lake and warehouse with **Amazon Athena**. Use **Amazon SageMaker** AI/ML services to detect grid anomalies, forecast energy usage, predict equipment failures, and more. Create and publish interactive dashboards that include AI/ML insights with

Amazon QuickSight. Dashboards can then be accessed from any device, and embedded into your apps and websites. Proactively communicate with customers using **Amazon Pinpoint** and measure customer engagement across multiple channels including email, SMS, and mobile push notifications. Use analytics and ML outputs with **Amazon Pinpoint** to create personalized customer target segments and campaigns.

View reference architecture online





Quick Start

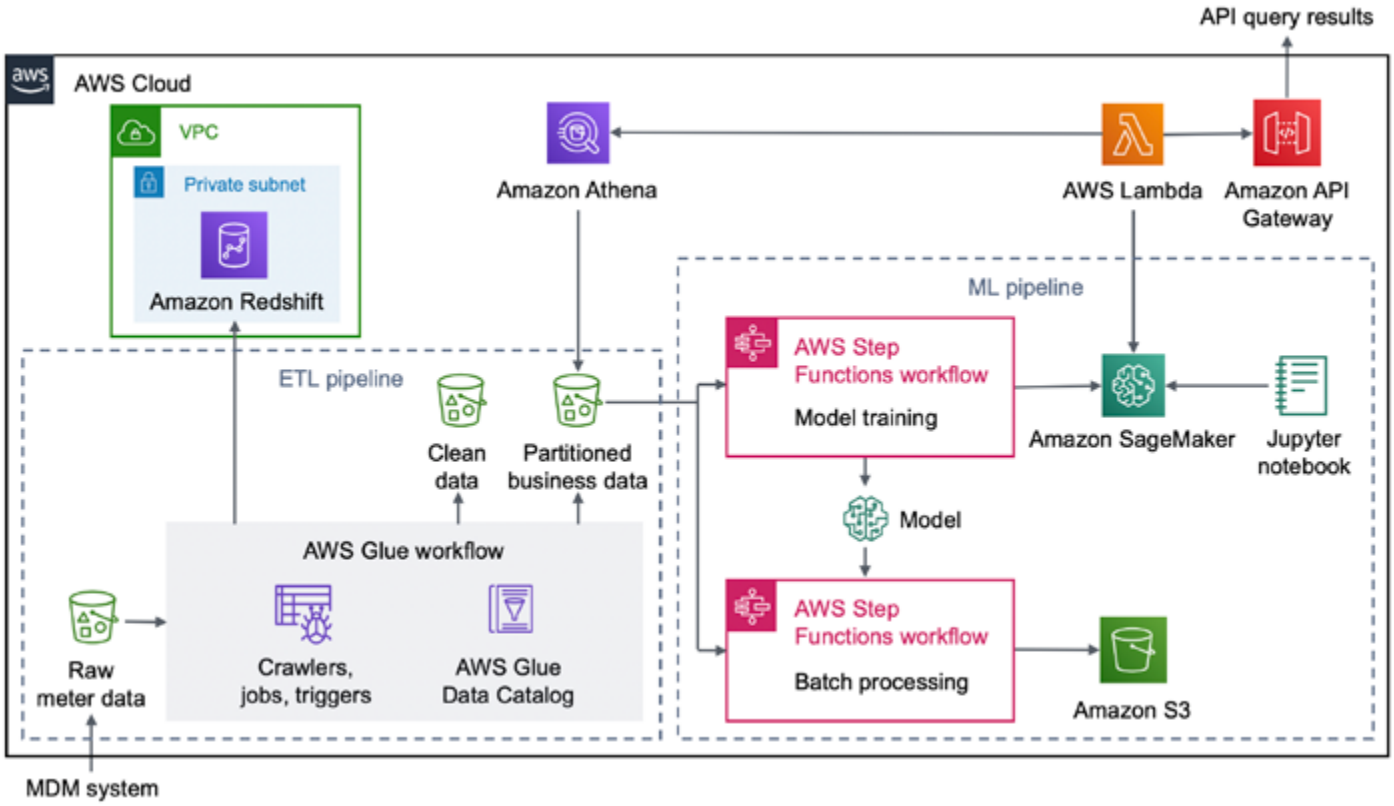
Utility Meter Data Analytics on AWS

Gain insights from data from smart utility meters

This Quick Start deploys a platform that uses machine learning (ML) to help you analyze data from smart utility meters.

Built by Amazon Web Services (AWS), this Quick Start is for utility companies and other organizations that are looking to gain insights from smart-meter data. This data comes from meter data management (MDM) or similar systems. Insights include unusual energy usage, energy-usage forecasts, and meter-outage details.

Except for Amazon Redshift, this is a serverless architecture.



The Quick Start sets up the following:

- A virtual private cloud (VPC) configured with a private subnet, according to AWS best practices, to provide you with your own virtual network on AWS.*
- In the private subnet, an Amazon Redshift cluster that stores business data for analysis, visualization, and dashboards.
- An extract, transform, load (ETL) pipeline:
  - Amazon Simple Storage Service (Amazon S3) buckets to store data from an MDM or similar

system. Raw meter data, clean data, and partitioned business data are stored in separate S3 buckets.
    - An AWS Glue workflow:
        - Crawlers, jobs, and triggers to crawl, transform, and convert incoming raw meter data into clean data in the desired format and partitioned business data.
        - AWS Glue Data Catalog to store metadata and source information about the meter data.
- An ML pipeline:
    - Two AWS Step Functions workflows:
        - Model training, which uses the partitioned business data to build an ML model.
        - Batch processing, which uses the partitioned business data and the data from the model as a basis for forecasting.
    - Amazon S3 for storing the processed data.
    - Amazon SageMaker for real-time forecasting of energy usage.
    - A Jupyter notebook with sample code to perform data science and data visualization.
- AWS Lambda to query the partitioned business data through Amazon Athena or invoke SageMaker to provide application programming interface (API) query results.
- Amazon API Gateway to deliver API query results for energy usage, anomalies, and meter outages.

* The template that deploys the Quick Start into an existing VPC skips the components marked by asterisks and prompts you for your existing VPC configuration.

View quick start online

Blog

How the cloud is helping us better understand and manage the oceans

by Ana Pinheiro Privette and Zac Flamig

World Ocean Day recognizes the crucial role that the oceans play in our lives. Oceans produce at least 50 percent of the planet's oxygen, absorb about 30 percent of carbon dioxide produced by humans*—buffering the impacts of global warming—and are home to most of earth's biodiversity. In addition, the oceans are key to our economy, being the main source of protein for more than a billion people around the world, and the* United Nations estimates *that 40 million people will be employed in ocean-based industries by 2030.*

To protect and preserve the oceans, we need to extensively understand its systems, and data is at the core of that process. The Amazon Sustainability Data Initiative (ASDI) *is committed to enabling better access to the foundational data that can help researchers, businesses, and policy makers better monitor and manage the ocean's valuable resources.*

Democratizing access to ocean data on AWS

The world's waters are largely unknown, with vast areas still unmapped. Only 46 percent of US ocean, coastal, and Great Lakes waters have been mapped to modern standards. Filling these data gaps is vital to human and marine health, safe navigation, and national security. In addition to generating new data, making existing oceanographic data and information more accessible to users can enhance the decision-making process, allowing mariners to improve safety at sea, optimize routes, and save fuel.

Through a collaboration with the National Oceanic and Atmospheric Administration (NOAA) and its Big Data Program (BDP), ASDI is making available petabytes of ocean-related data on Amazon Web Services (AWS) to help map, monitor, and manage ocean resources and activities. The collaboration leverages the AWS Open Data Sponsorship program to host foundational data on AWS, covering the storage and egress costs for qualifying datasets and allowing data providers to retain complete control and ownership of their data hosted on AWS. NOAA established BDP to provide public access to NOAA's open data and foster discoverability and usability.

NOAA ocean data currently hosted on AWS through this program includes the Crowdsourced Bathymetry and the NOAA National Bathymetric Source Data. These datasets support the creation of next-generation nautical charts and efforts to advance ocean-related science, and back industry and regulatory needs. Another dataset hosted on AWS is the NOAA World Ocean Database (WOD), the largest uniformly formatted and quality-controlled historical subsurface ocean profile database. The dataset provides global aggregation of ocean variables including temperature, salinity, oxygen, nutrients, and more. This dataset allows for the study and understanding of the changing physical and chemical state of the ocean.

NOAA BDP and ASDI also host ocean forecast data as guidance information to help mariners safely navigate their local waters. Operational nowcast (analyses of near present conditions) and forecast guidance (out to 48 hours) are available on AWS, providing access to parameters such as water levels, water temperature, salinity, and currents. For example, the NOAA Global Extratropical Surge and Tide Operational Forecast System (Global ESTOFS) offers users nowcasts and forecasts of water level conditions for the entire globe.

In addition to working with government agencies, ASDI also works with academia, nongovernmental organizations (NGOs), and private sector groups to make ocean-related data resources more accessible. Through the Farallon Institute, the Multi-Scale Ultra High Resolution (MUR) Sea Surface Temperature (SST) dataset is now accessible to anyone on AWS. This is a global, gap-free, gridded, daily, one kilometer sea surface temperature (SST) dataset created by merging several satellite derived sea surface temperature measurements. SST is an indicator of warming trends and critical to studying climate change.

Enabling better, faster, and cheaper solutions

Hosting these foundational large datasets on AWS means scientists can bring their code to the data, which removes many of the barriers that prevent quick exploration of new ideas and promotes collaboration. "This opens the door to additional users creating a more diverse and inclusive scientific community" said Dr. Chelle Gentemann, senior scientist at Farallon Institute, "…and helps us address an additional challenge: science's reproducibility problem. When programs like ASDI stage sustainability-related foundational datasets on the cloud and make it available at no cost, it makes it easier for researchers to create cloud-based scientific analysis and make the analysis code available adjacent to the data. This enables anyone to test the reproducibility of science results, which is important for transparency reasons and it allows scientists to build on each other's results and move the field forward."

The AWS Cloud is also enabling better estimates of hurricane wind speeds allowing for improved decisions around evacuations and general hurricane response planning, saving both lives and property. Hurricane wind speed estimates are currently made using the manual Dvorak technique. The National Hurricane Center releases them every three to six hours. Artificial intelligence (AI) experts with the IMPACT team at NASA's Marshall Space Flight Center and Development Seed created the Deep Learning-Based

Hurricane Intensity Estimator (on AWS) to automate this process and provide wind speed estimates with minute level cadence.

The private sector is also benefiting from open data and cloud technology to improve their services. nauricAI uses Internet of Things (IoT) devices deployed on vessels to track motion information, which when combined with ship location and weather data (available on AWS) provides customers with information on how wind and waves affect on-board navigation. Having the large weather and ocean data co-located in the cloud improves the efficiency of the underlying workload. These solutions enable prediction and optimization of navigation, and ultimately helps increase route efficiency, reduce emissions, and reduce environmental impacts.

Innovating for ocean protection and preservation

"Ninety-five percent of the ocean remains unexplored, and the lack of data for those areas is detrimental to ocean conservation efforts," said Fernanda Ubatuba, president and chief operating officer (COO) at OCEARCH. Her organization's mission is to uncover the history and behavior of sharks as key indicators of ocean health. By using AWS, OCEARCH is enabling large teams of collaborating scientists and experienced fishermen to collect critical data on shark movements and make these findings available to the public. The collaborative environment enabled by AWS provides customers access to previously hard to obtain data to help determine overall ocean health, inform fisheries management for sustainability, and support public safety efforts.

Groups like Saildrone are leveraging the AWS Cloud to better map the oceans. Saildrone builds and operates a fleet of unmanned surface vehicles (USVs) designed to collect high-resolution oceanographic and atmospheric data in remote oceans. Known as saildrones, each vehicle can stay at sea for up to 12 months, transmitting real-time data via satellite. The data collected is used to inform climate models and extreme weather prediction, maritime domain awareness, maps and charts, and sustainable management of resources.

Learn more about ASDI and listen to our recent Fix This podcast episode.

ASDI acknowledges contributions from the NOAA Big Data Program's Adrienne Simonson, Patrick Keown, Jena Kent and Jenny Dissen. Learn more about BDP here.

Read blog post online

Case Study

Saildrone Explores the Oceans with Machine Learning

2020

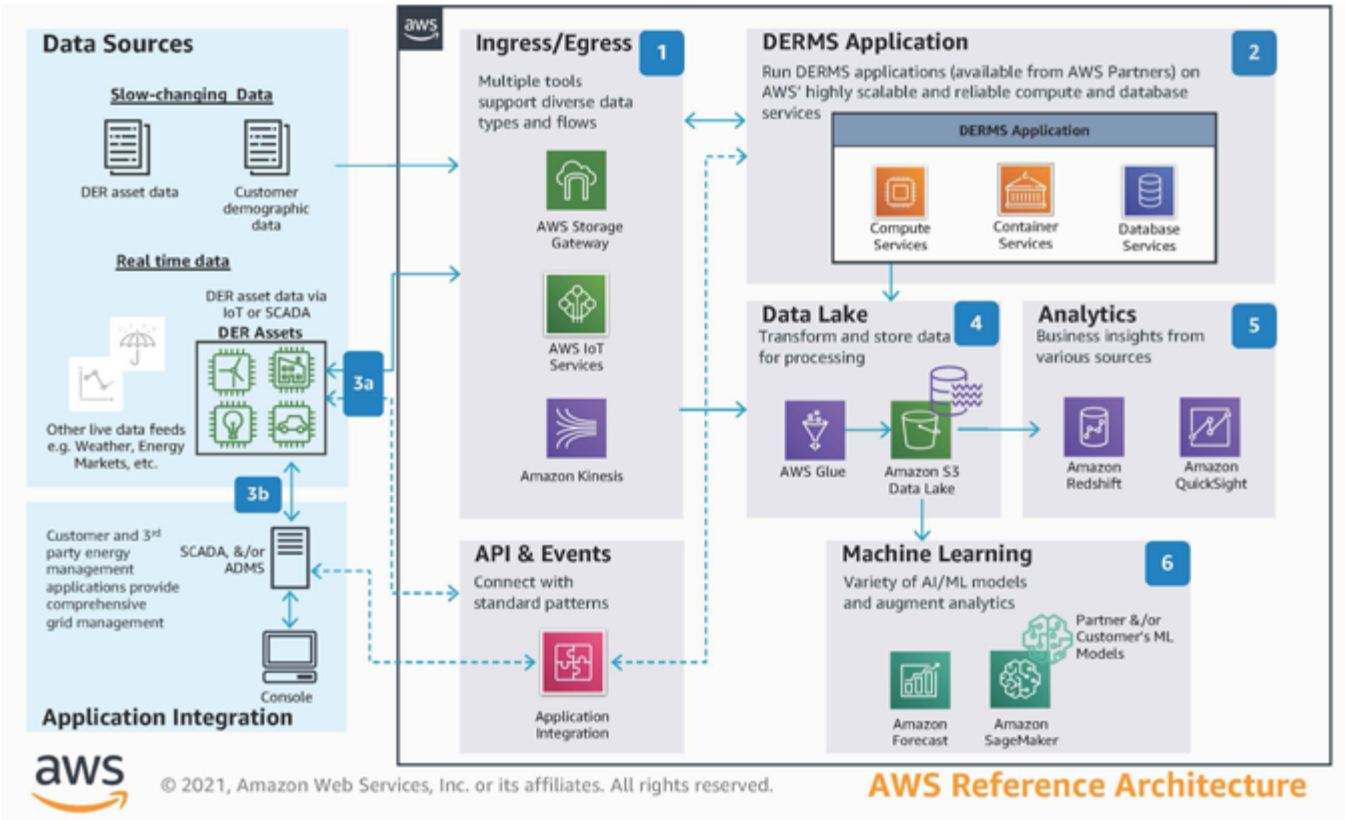
Saildrone collects global environment data with its wind-powered ocean drones. The company uses machine learning on AWS to better quantify the behavior and trends of major fish stocks and their predators, such as sharks and seals, to help with conservation efforts and sustainable fishery management. Saildrone has recently completed the first successful autonomous circumnavigation of Antarctica, producing key insights into ocean and climate processes, and used machine learning on AWS to identify the risk of collision with icebergs.

Read case study online

Reference Architecture

DERMS on AWS for Utilities

Run your **Distributed Energy Resource Management Systems (DERMS)** on AWS to meet increasing demand to monitor, forecast, and control your wind, solar, battery, and emerging "green" DERs. Deploy a scalable solution to effectively match generation with load, improve grid resiliency and reliability. Use the right compute, data store, and analytics tools to gain operational and customer insights, and run proactive and meaningful analytics using machine learning (ML).



1. Streaming and batch data is collected from various data sources, such as customer and grid data, meter data, DER measurements, market prices, and weather. Integration to DERMS is via one or more of AWS File Transfer, IoT, streaming, or API services.
2. Vendor DERMS applications can be deployed using different compute services, such as **Amazon EC2**, **AWS Lambda**, or AWS container services. Their associated transactional and operational data sets can be optimally stored in purpose-built AWS databases based on their data structure and access patterns.
3. A DERMS may control DER assets directly, via IoT or APIs (3a) or make recommendations to the ADMS/SCADA, and let it execute the actual control (3b). ADMS and other systems integrate with DERMS to ensure a balanced grid using application integration services such as **Amazon API Gateway**, **Amazon SNS**, **Amazon SQS** or **Amazon EventBridge**.

4. Automate ETL processes such as transform and deduplication with AWS Glue. Use Amazon S3 as low cost storage for raw and curated data, with archival copies for retention and compliance. Your **Data Lake** serves as the single source of truth for downstream analytics and ML work. **AWS Glue** can also automate the process of data schema discovery and metadata tagging to create a metadata catalog that makes all data visible and searchable.
5. Query petabytes of structured and semi-structured data across your data warehouse and your data lake using standard SQL with **Amazon Redshift**. Create and publish interactive dashboards with **Amazon QuickSight**, which can be accessed from any device, and embedded into your apps and websites.
6. Harness the power of pre-trained AI models, such as **Amazon Forecast**, to detect grid anomalies, forecast energy usage, and predict equipment failures. Build, train and deploy your own ML model quickly with **Amazon SageMaker**.

View reference architecture online

Blog

Forecasting energy usage using Amazon machine learning and data lakes

by Joseph Beer

Executives within utilities and energy providers of all types and sizes have multiple ongoing needs to forecast energy usage. For example, as chief customer officer, your teams can use energy forecasts at the household level to proactively engage homeowners with high bill alerts and predict pre-pay or month-end energy charges. As the head of energy efficiency and commercial energy programs, your teams can use forecast to predict potential savings when applying different energy efficiency measures, as well as recommend the best measures to use. You can forecast for each home owner, or in aggregate across multiple commercial buildings for each campus or property manager. As VP of operations you can deploy Amazon Forecast in many different ways, from predicting demand at the circuit or substation level to predicting substation or distribution transformer failures. You can also combine Amazon Forecast with customer demographic data to predict which neighborhoods and streets are most likely to see a rise in solar, battery, and electric vehicle installations.

Most utilities have the data they need to make these types of forecasts including metering, SCADA and customer data, but it's often locked away in separate incompatible data silos. Utilities often have a variety of tools deployed for forecasting, the most widely being Excel. Usually only a handful of staff are using advanced analytic tools though, and even fewer utilities have data scientists who may have used machine learning (ML) to build artificial intelligence (AI) models for the utility.

Several of our AWS power and utility customers have overcome these issues by leveraging Amazon's breadth and depth of AI/ML and Big Data & Analytics services to consolidate data into a data lake. These are used to produce analytics insights, predictions, and forecasts in a timely and affordable

manner. Utilities have done this without hard-to-find data scientists, using just the skills of their existing IT developers, data engineers, and analytics professionals. One recent AWS customer success story in this area is that of Xylem, a leading global utility technology company.

Our customers can do this because Amazon has taken over a decade of experience in using AI/ML to ensure the smooth running of Amazon's supply chain and customer engagement processes. That knowledge has been built into a set of easy-to use ML tools, such as Amazon SageMaker and a cadre of fully managed AI/ML services, for functions ranging from image recognition and natural language chat bots, to fraud detection, personalized recommendations, time series data forecasting, and more.

Amazon Forecast is the specific service which utilities would use to predict energy consumption. It's a fully managed AWS service that uses machine learning (ML) against time series data (exactly what SCADA and metering systems generate) to produce highly accurate forecasts, without requiring any prior ML experience.

Further reading

This detailed technical blog explores how utilities can use Amazon S3 Data Lakes and the Amazon Forecast ML service to predict energy usage by combining historical interval meter read data with external weather data. Puget Sound Energy used this approach to predict electrical and gas consumption at a typical residence.

Read more about PSE's proof of concept here.

Click here to learn more about AWS for Power & Utilities.

Read blog post online

Case Study

CropX Runs Soil Sensor Application on AWS to Help Global Farmers Enable Sustainability

2020

The global population is experiencing major growth, with the United Nations projecting an increase from 7.8 billion today to 9.8 billion by 2050. To meet this growth, agriculture must produce more food with less water. In fact, farmers already use 70 percent of available freshwater and 50 percent of habitable land. Lacking accurate irrigation and fertilization management capabilities, many farmers do not have comprehensive insight into soil conditions. As a result, farmers often provide either too much irrigation and fertilizer or too little, which wastes valuable resources and does not maximize yield potential.

CropX, an agricultural analytics company, is addressing this challenge by helping farmers integrate soil data with numerous above-ground data layers. "Most agricultural companies rely on above-ground data such as satellite imagery, and less than 10 percent of companies get data from within the soil, which is where the most valuable data is," says Matan Rahav, director of business development at CropX. "By the time there are visible signs of crop stress detected from space, the damage is already done. We provide real-time data on soil moisture, salinity, and temperature, so farmers can be predictive and preventive and can determine whether a plant has received enough fertilizer even before the plant knows."

However, many state-of-the-art soil sensors require lengthy and complicated installation and calibration, which contributes to rising costs and hinders scalability. "Sensors also create large amounts of raw data that farmers can't easily interpret," Rahav says. "The only way to process this data in real time, while integrating additional data layers, is in the cloud. As a startup company, we realized the cloud was our best option for quickly building and scaling our agricultural analytics application."

Using AWS to Support a Big Data, In-Soil Sensor Application

To solve the problem, the company chose to run its in-soil sensor data solution on Amazon Web Services (AWS). "AWS provides easy-to-deploy and scalable technology, managed with a simple console," says Sagi Briteman, vice president of research and development for CropX. "It is also easy to establish the architecture. We started with building blocks, with uncertainty about our physical deployment and how much we needed. Using AWS, it was easy to create instances as we grew."

The CropX big data solution captures data from thousands of global soil sensors and sends it to a centralized platform running on hundreds of Amazon Elastic Compute Cloud (Amazon EC2) instances. The solution analyzes and saves satellite imagery data to Amazon Simple Storage Service (Amazon S3) buckets and uses Amazon Relational Database Service (Amazon RDS) as its primary database and the core for storing all agronomical data and insights. Additionally, CropX relies on Elastic Load Balancing to support growing traffic from web and mobile devices, and it uses Amazon Simple Queue Service (Amazon SQS) to transfer messages between different services.

The CropX platform combines above-ground datasets with in-soil data measured by sensors that transmit that data to the AWS-based solution. The data is integrated with imaging,

weather, topography, and soil data, as well as crop models, hydraulic models, and user inputs. It is then analyzed by AI-based algorithms to provide analytical insights via the CropX web or mobile app. Farmers, breeding companies, agrochemical companies, crop insurers, and irrigation system manufacturers use the app to receive irrigation and crop nutrition recommendations for their crops. "Our customers can use the app to view the updated status of soil readings, moisture levels, and root zones," says Briteman. "We also provide alerts and notifications if their fields are too dry or too wet."

Scaling to Support Fourfold Data Growth

CropX can now scale its in-soil data analytics platform to support four times more data than the company had two years ago. As a result, the company can ingest millions of data points per day, including soil sensor data such as soil moisture, salinity, and temperature. CropX can also scale to ingest more satellite imagery and real-time weather data to store in Amazon S3 buckets. "AWS enables us to collect raw data, process it in real time, and provide real-time recommendations that are specific to the crops, the geography, and the soil," says Briteman. "We would not be able to support a big data analytics platform without the cloud."

> ***"AWS enables us to collect raw data, process it in real time, and provide real-time recommendations that are specific to the crops, the geography, and the soil."***
>
> Sagi Briteman
> *Vice President of Research and Development, CropX*

Enabling Agricultural Sustainability

By running on AWS, CropX is helping its farming customers contribute to increased sustainability by using less energy and fewer resources. For example, during irrigation experiments using its technology, CropX has demonstrated more than 40 percent water savings across different crop types, with a 10 percent yield increase. "We have seen similar savings on fertilizers, energy, and labor because farmers spend fewer resources on travel and equipment," says Rahav.

Farmers can also make better decisions about how much to irrigate their crops. "If you don't irrigate enough, you don't maximize your yield potential and you expose the crop to constant stress, so you're not as efficient, which is a sustainability issue," says Rahav. "Conversely, over-irrigation can cause fertilizer to reach groundwater, which creates pollution. It can also expose crops to disease through oversaturated soil. Using the soil data insights from our application, customers can maximize their yield potential and better control irrigation."

Deploying 9,000 Sensors across the Globe

CropX has fueled significant business growth by running on AWS. The company now serves more than 1,200 farming customers worldwide and has almost 9,000 sensors deployed. "We quadrupled our sales in the past year and we have expanded to dozens of countries because we can rely on AWS to manage the backend environment for our application," says Rahav. "We don't need to travel to new areas for new projects or installations. As a startup, we wouldn't

be where we are now if we weren't running on AWS."

In addition, CropX acquired CropMetrics, which provides cloud-based, precision irrigation tools. Through this acquisition, CropX is giving its customers access to in-soil data, advanced farm management analytics, and tools to support decision making. "We are exploring the use of AWS services such as AWS IoT to drive our expansion," says Briteman. "AWS gives us the scalability we need to fuel our growth and help enable sustainable agriculture."

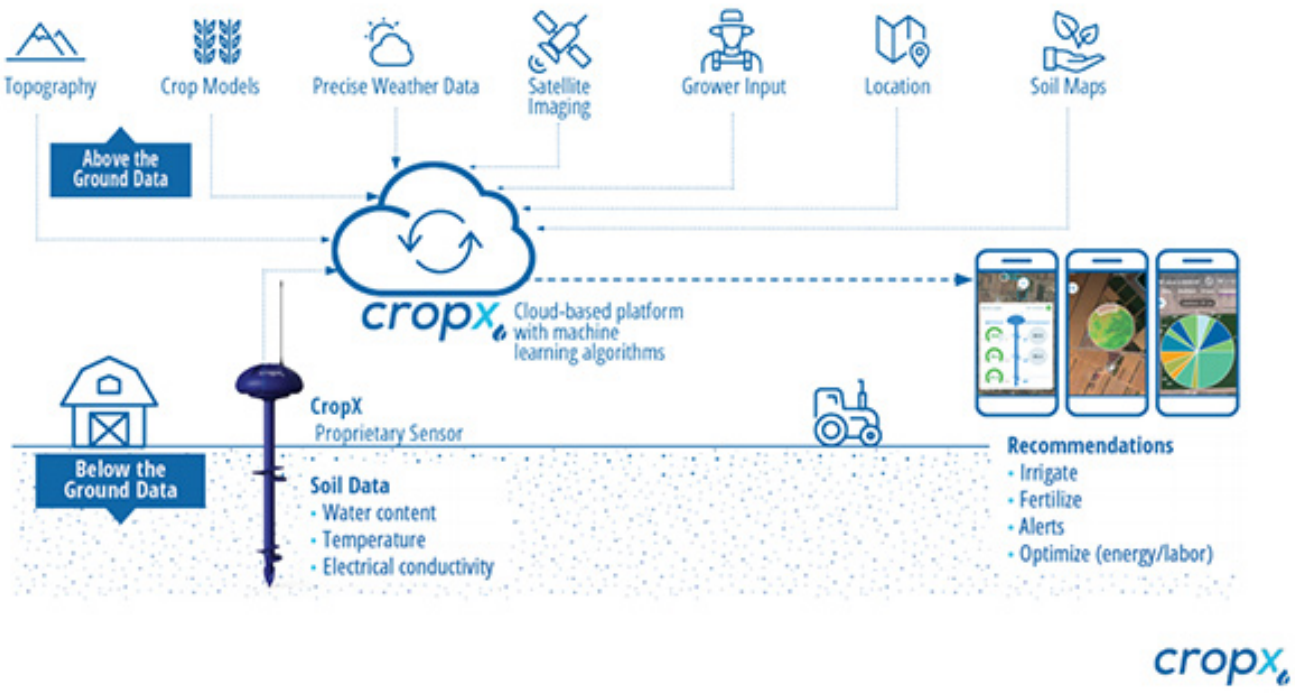
To learn more, visit aws.amazon.com/big-data/datalakes-and-analytics.

Read case study online

About CropX
CropX is an agricultural analytics company that develops cloud-based software solutions integrated with wireless sensors, which helps farmers boost crop yield. The company offers an advanced, adaptive farm management software service designed to deliver water and energy cost savings while conserving the environment.

Benefits of AWS

- Enables agricultural sustainability
- Scales to support a fourfold increase in soil sensor data
- Quadruples annual sales
- Supports 9,000 global sensors
- Gives farmers new insights into their crops



Whitepaper

The Carbon Reduction Opportunity of Moving to Amazon Web Services

October 2019

Introduction

Enterprises use public cloud infrastructure for various reasons, but faster time-to-market and the ability to flexibly scale resources to track workload demand are the top two drivers, according to 451 Research's Voice of the Enterprise survey on cloud, hosting and managed services[1]. This is particularly true for organizations further down the line in their digital transformation, which shows enterprises increasingly value the agility and scalability of cloud as they grow more acquainted with it.

Moving workloads to public cloud infrastructure also presents enterprises with the opportunity to dramatically reduce the environmental footprint of their IT operations. Climate change and its effects direct more attention at resource efficiency as a key part of sustainability responsibility, which is of growing importance to businesses. In 2018, 86% of the companies in the S&P 500 index published a sustainability report, up from only 20% in 2011, notes the Governance and Accountability Institute[2].

Enterprises want to be seen as responsible corporate citizens, and many have made sustainability commitments and achieved progress in multiple areas of their operations. Yet even with an emphasis on sustainability, running data centers and IT is not a core competency of most enterprises, many of which lack the expertise and resources to make major investments in infrastructure efficiency. Similarly, most enterprises are not prepared for the effects of climate change (such as extreme weather conditions, drought or floods) on their data center operations.

---

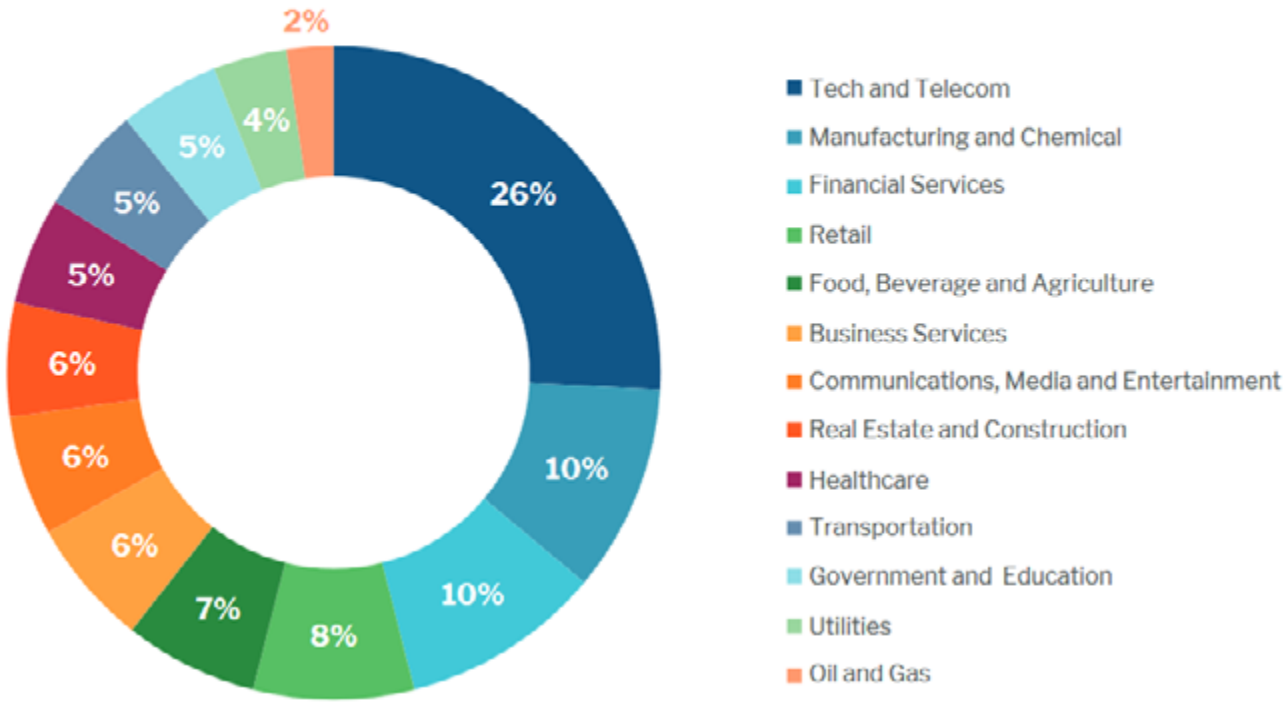
1. Voice of the Enterprise: Cloud, Hosting and Managed Services, Workloads and Key Projects - Quarterly Advisory Report, 451 Research, June 28, 2019
2. Flash Report, Governance and Accountability Institute, a corporate risk & sustainability consultancy, May 16, 2019

To estimate the environmental benefits for enterprises moving to its public cloud infrastructure, Amazon Web Services commissioned 451 Research, a technology market research and advisory firm, to conduct a study on the energy and carbon efficiency of enterprise data centers and server infrastructure.

For this study, we focused on US enterprises with revenue between $10m and $1bn, businesses that face IT challenges similar to those of larger enterprises, often with smaller budgets and less IT expertise. Even when enterprises invest in newer infrastructure, they still have to provision for peak demand, and many energy-efficiency measures are only viable at the scale of thousands of servers and when applying advanced design techniques (e.g., efficiency-optimized custom servers, wide temperature bands and indirect evaporative cooling). This means that moving workloads to the cloud can help enterprises steeply reduce energy consumption and their carbon footprint compared to their internal operations.

We surveyed senior stakeholders at over 300 companies – typically with the titles of CIO, CTO, IT director and data center manager – seeking information about their infrastructure operations at their largest US data center location (on-premises or colocation). Participating companies came from a broad cross section of the US economy representing over 20 sectors, such as IT and telecommunications, media and entertainment, financial services, manufacturing and healthcare. We also conducted 10 in-depth interviews to get more insight into the thinking and challenges of enterprise infrastructure operations.

Figure 1: Demographics of surveyed US enterprises by major industry groupings (n=302)
Source: 451 Research



451 Research devised a carbon efficiency model that offers a grid-to-chip view of efficiency. Considering the power intensity of their operations, servers account for the majority of the enterprise infrastructure carbon footprint and are indicative of relative efficiencies. Also, servers run a growing variety of data storage and network services, with the rise of trends like hyperconvergence and software-defined infrastructure. The model includes data center facility overhead as captured by the widely used power-usage effectiveness (PUE) ratio and server energy efficiency. We used survey results, inputs from Amazon Web Services on its operations in the US, and third-party industry data (including from data center design and operations authority Uptime Institute, a 451 Group company) to populate the model.

We did not include the carbon footprint of water or direct emissions of carbon and other greenhouse gases from sources such as emergency power generators in our carbon model. Water does contribute to the environmental footprint of data centers and is a scarce resource in some regions, but its effect on carbon emissions is much lower than electricity usage. Similarly, direct (Scope 1) emissions in data centers have marginal carbon impact compared to indirect (Scope 2) emissions attributed to energy. Future studies may consider these views as well as embodied emissions (Scope 3) in buildings and hardware for a more complete picture.

Executive Summary

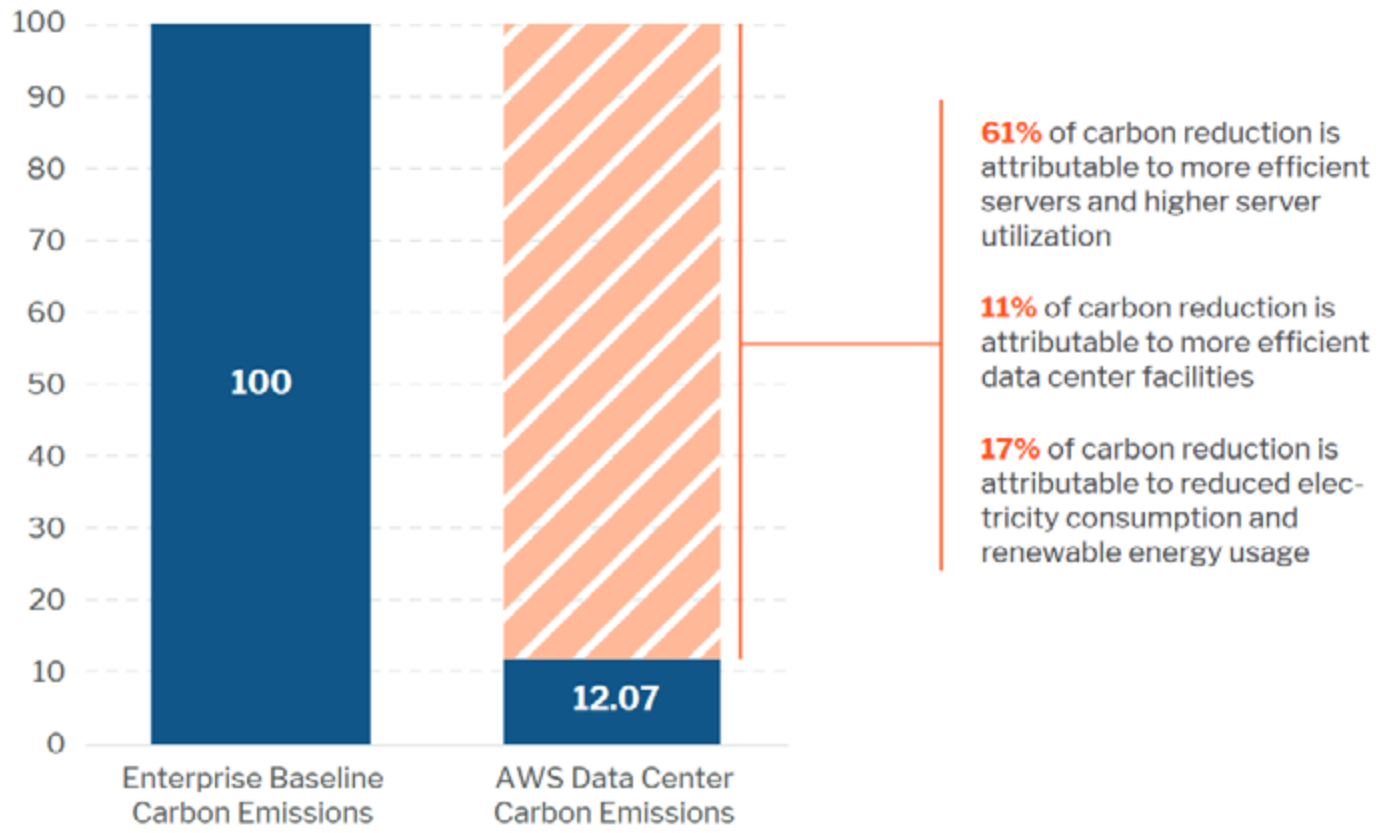
The success of public cloud services has resulted in the creation of cloud data center campuses much larger than enterprise sites. While hyperscale campuses have attracted scrutiny for their energy usage, they are much more efficient and offer a workload carbon footprint that is a fraction of what enterprises typically produce in their on-premises or colocation data centers. There are several technical components to this carbon advantage, but what really makes public cloud inherently more efficient is structural. Cloud operators make the entire technical organization work in unison to attain high infrastructure efficiency by design, while the cloud business model of a shared and monetized infrastructure drives server utilization well above what is possible for enterprises.

This is contrasted with enterprise infrastructures, which remain fragmented in their management. While efficiency does factor into technology and product choices, efforts are mostly limited to the existing infrastructure and organizational framework, as opposed to fundamental changes to assumptions in engineering and operations. A prime example of this approach is temperature controls, which remain tight at enterprise data centers and are energy-intensive and expensive to maintain year-round. Enterprises also tend to run underutilized or idle servers in significant numbers - the Uptime Institute estimated in a study that roughly 20% of racked enterprise servers are completely unused and abandoned by application administrators, owing to insufficient monitoring of the infrastructure and the lack of a rigorous decommissioning process.

Our results show that AWS's infrastructure is 3.6 times more energy efficient than the median of the surveyed US enterprise data centers. More than two-thirds of this advantage is attributable to the combination of a more energy efficient server population and much higher server utilization. AWS data centers are also more energy efficient than enterprise sites due to comprehensive efficiency programs that touch every facet of the facility.

When we factor in the carbon intensity of consumed electricity and renewable energy purchases, which reduce associated carbon emissions, AWS performs the same task with an 88% lower carbon footprint. (Figure 2).

Figure 2: Carbon efficiency of AWS infrastructure compared to surveyed US enterprises
Source: 451 Research



Even when compared to the top 10% most efficient organizations surveyed, moving to AWS would deliver a 72% reduction in carbon footprint on average. The results suggest that moving to the cloud would reduce workload carbon footprint for virtually any US enterprise in this cohort of companies with revenues between $10m and $1bn.

451 Research expects this carbon benefit to grow in the coming years. Our modeling suggests that AWS should be able to improve its efficiency faster than enterprises using on-premises infrastructure based on currently known server technology roadmaps and expected growth rates in public cloud infrastructure.

Read full whitepaper online

Find out more about how the AWS Cloud can help enterprises and public sector organizations in Asia Pacific lower their carbon footprint. Read the 451 report for Asia Pacific: The Carbon Reduction Opportunity of Moving to the Cloud for APAC.

https://amzn.to/AWS-Sust-451-APAC

Blog

Open data on AWS supports sustainable agricultural practices and crop optimization

by Christopher Goodman, Adrienne Simonson, Jared Oyler, and Jenny Dissen

Weather and climate data are important building blocks for digital farming solutions and critical for informing the numerous strategic and operational decisions that farmers must make daily. From selecting the right seed and crop production purchases, spray location and timing, or determining when to plant and harvest, having access to timely and accurate weather information is vital. However, most weather and climate data products available don't reflect field-specific conditions and are therefore not directly relevant to the spatial and temporal scales at which farmers make their decisions. This is a particularly challenging problem for developing effective digital solutions that are applicable globally. To address this challenge, BASF Digital Farming is leveraging both National Oceanic and Atmospheric Administration (NOAA) weather data hosted on Amazon Web Services (AWS) and commercial weather data to build hyperlocal agriculture-specific weather and climate solutions that scale globally.

As part of the Amazon Sustainability Data Initiative (ASDI), we invited Christopher Goodman (BASF), Jared Oyler (BASF), and Jenny Dissen (NOAA BDP/CISESS) to share how AWS Cloud technology and open data are supporting BASF Digital Farming's efforts to make agricultural practices more sustainable and efficient.

An advanced digital farming platform for crop optimization

With a rapidly growing population, the world is increasingly dependent on the ability to develop and maintain sustainable agriculture and healthy environments. Working with farmers, agricultural professionals, pest management experts, and others, BASF Digital Farming is developing digital farming solutions to support agriculture and crop optimization. **xarvio™ Digital Farming Solutions**, hosted on AWS, provides growers with valuable information on how to make better farm management decisions. xarvio™ products provide tools to help farmers effectively monitor and manage their fields, including the ability to identify crop diseases, weeds, and pests, and recommends action to mitigate these in-field issues. These digital farming tools help drive farm sustainability by enabling farmers to use crop applications more efficiently and precisely only where and when needed.

Building this platform on the secure AWS Cloud enabled xarvio™ to offer its digital farming solutions for agriculture and crop optimization to support farmers worldwide. The platform currently supports more than 4 million farmers and advisors in more than 100 countries, allowing them to consistently, closely monitor their crops. The xarvio™ SCOUTING smartphone application (app), for example, which is free to download and use, utilizes image recognition technology. It lets a farmer use their smartphone to identify early-growth weeds, diseases, and pests; map their location for sharing; keep a history; and act to mitigate.

NOAA's weather and climate information drive farmer's decisions

Weather science is at the forefront of BASF's Digital Farming's tools. Within xarvio™ Digital Farming Solutions' products, weather and climate information is provided to growers to help them make business and strategy decisions, such as whether or not to apply an herbicide, and how to increase the effectiveness of the crop production application while minimizing cost and environmental impact. BASF Digital Farming uses the NOAA public Global Historical Climatology Network Daily (GHCN-D) and Integrated Surface Temperature Dataset (ISD) station observation datasets hosted on AWS to benchmark weather and climate data product performance and train new agrometeorological models on AWS. Such models support the farmer in targeting their planting, crop protection applications, harvest, and other operational and strategic decisions.

Other analysis-ready datasets freely available on AWS, like Landsat and Sentinel-2, are also used to create efficient research and development (R&D) workflows and support the development of BASF Digital Farming's solutions. Its goal is to combine foundational public open data with commercially sourced information to help farmers improve yields and adopt sustainability and environmentally friendly practices.

Weather and climate data on AWS allow for easier, faster, and cheaper data analysis

By hosting foundational data on the AWS Cloud, the NOAA Big Data Program (BDP), in collaboration with ASDI, seeks to remove some of the redundant efforts that a customer like BASF Digital Farming has to endure when discovering, accessing, and using NOAA data to support its workloads. Most of a data scientist's effort is often associated with time consuming dataset acquisition and wrangling tasks rather than model development and knowledge creation. NOAA data staged on AWS addresses this issue by enabling both more efficient data access and analyses. BASF Digital Farming's teams no longer need to acquire and maintain the data, and in addition, can leverage AWS scalable computing and services right next to the data.

However, because NOAA datasets are usually provided in their legacy formats (e.g., NetCDF, GRIB, CSV), the user may need additional investment to integrate datasets and conduct analysis in the cloud. Making the data available in more cloud optimized formats like Zarr or Parquet allows for users to slice and dice the data and conduct spatial temporal queries in the cloud rather than having to download each file prior to analysis. To facilitate analysis of the NOAA data on AWS, BASF Digital Farming is converting NOAA Global Forecast System (GFS) data to cloud optimized formats and analyzing the data directly with basic services like Amazon Simple Storage Service (Amazon S3) and Amazon Elastic Compute Cloud (Amazon EC2).

ASDI is working with groups like Pangeo to support data providers with further exploration on how to create cloud data optimized formats. Hosting datasets in analysis ready formats in the cloud will allow BASF Digital Farming and others to conduct research faster and cheaper, and to deliver better solutions.

The NOAA data hosted on AWS through ASDI leverages the Open Data Sponsorship Program. For more information about the Amazon Sustainability Data Initiative, visit here. Visit here for more information on NOAA Big Data Program, or email NOAA.BDP@NOAA.GOV.

Read blog post online

Additional Resources

How Cloud Enables Utility Decarbonization and Flexibility

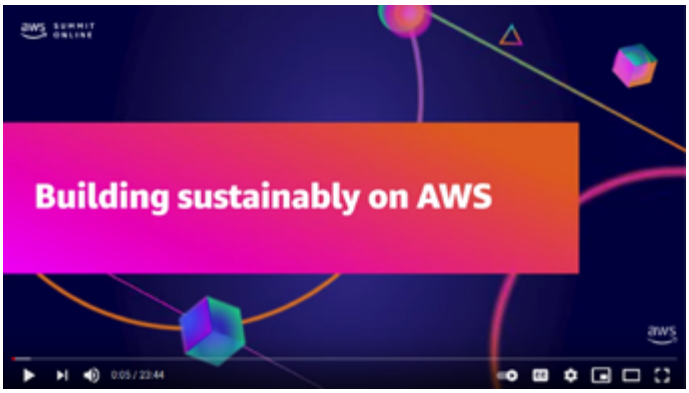
AWS Data Exchange Sustainability

Amazon Sustainability Data Initiative

Videos

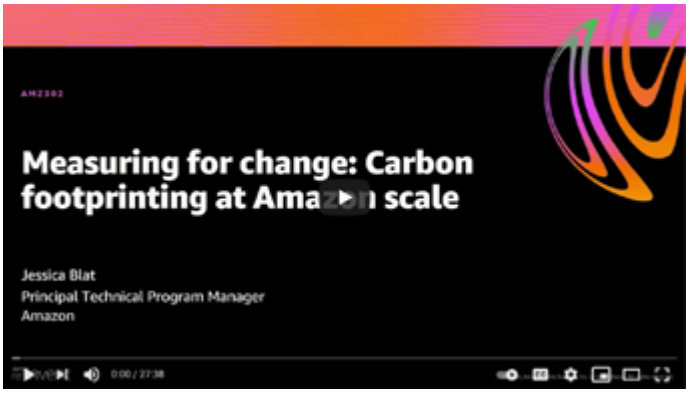
Building Sustainably on AWS

Did you know AWS infrastructure is 3.6 times more energy efficient than traditional data centres? AWS performs the same task with an 88% lower carbon footprint and covers non-renewable energy used in certain regions with a goal to be 100% renewable by 2025. In this session, we share Amazon's science-based, customer-centric approach to integrate sustainability throughout its business operations, and provide internal and external customer examples.



AWS re:Invent 2020: Measuring for change: Carbon footprinting at Amazon scale

A critical step in Amazon's Climate Pledge strategy is creating an accurate and useful carbon footprint of the company's activities. In this session, you learn how Amazon has constructed its carbon footprint and ways that you can think about your own footprint. This session summarizes the technologies that went into the Amazon solution, the challenges solved and those that still remain, and what the future holds for sustainability. Additionally, this session covers the specific AWS services that were applied to build an energy-efficient and secure solution.



Carbon Lighthouse Tackles CO2 Emissions with Machine Learning

Commercial buildings are responsible for 40% of U.S. emissions. Learn how Carbon Lighthouse uses machine learning on AWS to develop insights that deliver energy savings and decrease CO2 emissions in commercial real estate.

Learn how Alexa's new energy dashboard can track, manage, and save energy usage.

Starting later this year, Alexa can help you manage the energy consumed by compatible smart devices in your home. You can track energy usage, see personal insights, and get helpful tips—right in the Alexa app. Soon, try an exciting new Hunches feature that can help you save energy without even thinking about it. Now, if Alexa has a hunch that you forgot to turn off a light and no one is home or everyone went to bed, Alexa can automatically turn it off for you.

PayGo: Leveraging Smart Meters to Help Customers Conserve Energy!

On this episode, learn how PayGo leverages smart meters technology to enhance sustainability and increase customer satisfaction by helping them conserve energy. Chad Gates from PayGo explains how they built a highly available mobile enabled services on AWS to help customer use bar code technology to pay for their utilities and integrates it to their home with smart grid technology to reduce operational costs for the utility.

All in the Field: AWS Agriculture Live | S2 E7 | Smart & Sustainable Farm Solutions: Exploring AI with CropIn