

# Design Tips for a More Environmentally Sustainable Database Architecture

# Table of Contents

Introduction	3
An SLO for Planet Earth	4
Hosting on a Greener Infrastructure	4
The Sustainability of Scale	5
Elastic Resources for the Win	6
Modeling Data for Efficiency	7
Tools for Sustainability	7
Conclusion	8
More Information	8

# Introduction

If you've traveled by air the last few years, you may have noticed the green leaf logos that indicate which flights have lower carbon emissions and are more environmentally sustainable.

Even if you haven't traveled, you've probably noticed how the air travel of celebrities has come under scrutiny for their environmental impact. In all facets of our economy, increased awareness of climate change has put industries in the spotlight when their environmental records don't align with popular sentiment or scientific consensus. While it may not be as high profile as celebrities in private jets, carbon emissions for the Information Communications Technology (ICT) industry are on par with the aviation industry. [According to one estimate](#), ICT is expected to contribute to up to 23% of the globally released greenhouse gas emissions by 2030. As software developers, DBAs, project managers, and product managers, we are key constituents and have tremendous influence on the ICT industry. How we design our systems and architectures can have a huge impact on our carbon footprint.

As part of our commitment to environmental, social, and governance (ESG) issues, MongoDB has a duty to conserve natural resources and practice

good environmental stewardship. We believe that environmentally responsible operating practices will serve to benefit stockholders, partners, customers, and employees alike. We strive to incorporate sustainability into all areas of the business, from product development to our workspaces. We also have a responsibility to consider the impact on the environment as a result of the use of our product. The convenient truth about MongoDB Atlas, our [developer data platform](#), is that the same practices that help our users achieve faster, more performant, and cost-effective database outcomes also serve to reduce CPU utilization, energy consumption, and environmental impact. It's important to keep in mind that every CPU core and gigabyte of network bandwidth has a cost and corresponding carbon footprint that increases every second it's in use. This white paper discusses environmentally sustainable architecture best practices at the data layer. It also highlights how organizations can leverage MongoDB Atlas to design and deploy more environmentally sustainable applications.



# An SLO for Planet Earth

If you consider the average global temperature of the Earth as a service level objective (SLO), average global temperature rise should not exceed 1.5°C. above pre-industrial levels.<sup>1</sup> At the rate of current average global temperature rise, we are on pace to breach the 1.5 °C SLO within eight years, a tipping point past which the equilibrium of the planet could be at risk. If we're to meet the SLO for the planet, the ICT industry needs to reduce global greenhouse gas emissions by 45%.<sup>2</sup> Data and data platforms are key to addressing this global challenge.

## Hosting on a Greener Infrastructure

The number one way most organizations can reduce their greenhouse gas emissions is by hosting their systems on greener infrastructure. Public cloud services from the major hyperscalers have less impact on the environment than most private data centers. On average, migrations from on-premises infrastructure to the public cloud result in an 84% reduction in greenhouse gas emissions.<sup>3</sup> All three of the major public cloud providers play a key role in our sustainability efforts. We are pleased to partner with suppliers that are dedicated to renewable energy. Our primary cloud providers are on track to offer more data centers powered fully by renewables. Google has 100% renewable-powered data centers, while AWS and Microsoft Azure will reach their target of fully renewable-powered data centers in 2025.

It's not that private data centers can't be more sustainable. It's that running and managing a data center isn't a core business objective for most organizations. Even if they were to invest in a more energy-efficient on-premises data center, doing so is not likely to help them gain competitive advantage or increase market share. For major public cloud providers like AWS, Google, and Microsoft, running their data centers as efficiently as possible is a core objective for the business. So it's worth it for them to invest the time and resources to optimize energy consumption and efficiency.

Switching to public cloud infrastructure offers several advantages that lead to significantly reduced resource utilization and carbon footprint:

- **Power and cooling.** Cloud providers use more efficient electrical distribution, custom server designs, and evaporative cooling to achieve world-class energy-efficiency results.
- **Compute utilization.** The shared infrastructure business model combined with techniques such as dynamic allocation of compute resources and high density servers drives higher server utilization.
- **Hardware efficiency.** Cloud providers use higher performance processors because they can monetize them more effectively, resulting in lower carbon overhead for a unit of performance.
- **Carbon-free energy.** Cloud providers have greater renewable energy mixes, minimizing data center carbon footprints through renewable energy.

Many of the major public cloud regions from AWS, Google, and Microsoft are powered by renewable energy, low-carbon sources.<sup>5</sup> [MongoDB Atlas](#) provides easy access to these low-carbon global cloud locations and seamless migration between more than 100 different cloud regions as their emission profiles change over time. Atlas is the only multi-cloud operational database

1. Intergovernmental Panel on Climate Change, 2. International Telecommunication Union, 3. Accenture Newsroom,

5. Sustainable cloud regions: [AWS](#), [Google Cloud](#), [Microsoft Azure](#)



platform that offers the ability to run across all three major cloud platforms. So it provides the most flexibility in choosing a location that uses renewable energy and enables you to build a more sustainable architecture.

Businesses can now consider and act upon the environmental impact of growing their operations by choosing carbon-friendly cloud regions to deploy workloads. As businesses scale up and out, they can deploy Atlas closer to where workloads are needed and where they can reduce the environmental impact of IT operations.

Database operations can be extremely resource intensive, especially when you consider how geographically distributed data, applications, and users are in today's digital economy. In most cases, storing your data closer to where it's being generated and used is more efficient and less resource-intensive. In a global deployment, getting the data as close to the users as possible is paramount to managing emissions.

Dedicating clusters for specific workloads like analytics is another option for designing a more environmentally sustainable architecture. For

“IaaS migrations can reduce carbon emissions by more than 84% compared with conventional infrastructure.”<sup>4</sup>

time-sensitive transactional workloads, it may not always be possible to choose a cloud region based on sustainability criteria. High velocity read and write workloads require faster resources with the least amount of latency. However, you can look to offload other components of your workload in a more sustainable region. Analytics, machine learning (ML), and BI are long-running processes that are not as time-sensitive, can tolerate higher latency, and don't usually require write access to the database. ML and BI tools can therefore run in isolation against secondary nodes. Another benefit of workload isolation is that if a particular workload is bigger or smaller than the transactional workload, you can size each node according to its workload size to reduce compute resources.

4. Accenture, [The green behind the cloud](#)

## The Sustainability of Scale

The software as a service (SaaS) consumption model is inherently more environmentally sustainable than the traditional model of software development, distribution, and consumption. SaaS solutions such as MongoDB Atlas are developed once and used many times. As the vendor, we build and maintain all the underlying engineering of Atlas, including high availability, backups, auto upgrades, and more. That work is leveraged by millions of users around the world.

If each business or organization started developing technology solutions from scratch instead of using SaaS solutions, it could lead to more resource consumption as many of these solutions will not be as optimized in terms of how they consume the underlying resources. For example, building your own auto-scaling and spinning up replica sets could consume far more resources than just using Atlas.



# Elastic Resources for the Win

Around 40% of cloud instances are provisioned to be one size larger than required.<sup>6</sup> There's often a good reason for this. Sometimes it's hard to predict the size of a workload ahead of time. Other times, over-provisioning is a way to ensure resources are available if activity in the application spikes for a period of time. But overprovisioning to this extent leads to roughly double the compute cost and emissions.

MongoDB Atlas includes a wide range of scaling options that ensure the most carbon efficient instance size is always in operation. Scaling events, both up and down, occur in the background without incurring downtime or impacting performance, and can be triggered on demand by a click in the GUI or an API call.

One way to avoid needless overprovisioning is with cluster and storage autoscaling. [The way this works in Atlas](#) is that the service tracks how much resources an application is using over time (RAM and CPU) and [automatically scales](#) up if usage exceeds 75% utilization for a sustained period of time, or down if usage remains below 50% utilization, without any manual intervention. The same process works for storage as well. For more predictable workloads, scaling can be [scheduled using Atlas triggers](#) to ensure adequate resources are available during the times of day or year when you expect activity to spike.

Another way to avoid overprovisioning is by deploying a [serverless](#) database. With Atlas serverless instances, the instance seamlessly scales up and down based upon the workload demand. If the database is not being used, it will scale down to zero. If there is a sudden spike in activity, it will instantly scale up, making serverless a great choice for highly variable or unpredictable workloads.

Many organizations are embracing serverless architecture because of the cost savings from eliminating idle resources. According to Forrester's most recent developer survey, 24% of full-time developers have reported they or their teams already develop with serverless approaches, and 15% of decision-makers cited embracing serverless as the cloud initiative they are most likely to adopt over the next 12 months.<sup>7</sup> But serverless is a case where what's good for cost reduction is also good for carbon reduction, a win-win for businesses and the environment. In fact the environmental benefits of serverless scale with greater adoption. So the more businesses embrace serverless architecture, the more environmentally sustainable the ICT industry will become. In some cases, as with some large-scale production workloads, it may not be possible to adopt serverless architecture, at least not right away. But non-production environments, which sit idle for 76% of the week<sup>6</sup>, can operate far more efficiently, at lower cost, and with a smaller carbon footprint deployed on serverless instances.

6. DevOps.com, [The Cloud is Booming but so is Cloud Waste](#), 7. Forrester, [Demystifying Serverless Development](#)



# Modeling Data for Efficiency

[There are ways to model your data](#) in MongoDB that can increase the efficiency of database operations and reduce costs. In the document model, data that is accessed together is stored together. This results in fewer server round trips, less data movement across networks and hence, less resource consumption. It also reduces the need for costly and resource-intensive joins. The document model provides a lot of flexibility in how developers model data. That flexibility can be harnessed through the application's data access patterns. By modeling data according to data access patterns of the application, you can reduce the instance size necessary to handle workloads.

Schema design in MongoDB also has a tremendous impact on efficiency, and thus performance, of your application. In fact, performance issues can frequently be traced to poor schema design. The additional use of indexes, read concerns, and write preferences can improve query performance and reduce query resource consumption. With relational databases, queries may result in a high number of joins, higher time complexity and in turn, more infrastructure required to support the workload. By co-locating data that's accessed together through data access patterns in MongoDB, you can reduce many of these expensive and resource-intensive database operations.

## Tools for Sustainability

MongoDB Atlas is a [developer data platform](#) that uniquely provides a tightly integrated collection of data and application infrastructure capabilities, allowing development teams to quickly and easily address a wide variety of different application requirements. It supports transactional, search, and analytical use cases across mobile, distributed, event-driven, and serverless architectures, all while reducing complexity and helping developers design more environmentally sustainable architectures. Atlas also includes tools that enable organizations to monitor storage size, resource consumption, data transfer, and CPU utilization as a way to control costs and to measure sustainability practices for ESG reporting.

While it is important to monitor metrics related to compute, storage, data transfer, and hardware utilization, it's even more important to do so in relation to non-functional requirements, which commonly include cost, availability, quality of results, and response time among others. Atlas exposes more than 100 KPIs in our monitoring metrics dashboards, giving greater visibility into these metrics and enabling decision-makers to be more informed. MongoDB Atlas makes it easy to visualize these key metrics directly in our platform with [Atlas Charts](#).

In order to meet today's challenge of a 1.5°C SLO for Planet Earth, we should add sustainability to the list of those non-functional requirements, and start making appropriate tradeoffs keeping sustainability in mind. For example, if your application does not require real-time responsiveness, it might be okay to incur a slight latency on queries in order to be able to host the database in a physically farther but carbon-neutral cloud region.

Additional efficiencies might also be found by archiving historical data that's not accessed very much on cheaper, more energy-efficient object storage like Amazon S3. Storing hot data generates more carbon than warm or cold data. [Atlas Online Archive](#) automates the process of managing data tiering and therefore carbon tiering. With Online Archive all data remains queryable through a single MongoDB interface. With [Atlas Data Federation](#), developers can seamlessly run federated queries across data sources without having to move or copy all the data, reducing data duplication and complexity, making the developer experience of working with data more seamless.



## Conclusion

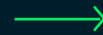
More organizations are seeing the benefits of simpler, more efficient, and more environmentally sustainable architectures. Sustainable architecture decisions can lead to lower costs, higher efficiency, higher developer productivity, and long-term resilience of applications, all of which are goals that align with overall business goals.

According to a report by Accenture, businesses that leverage the twin transformers of digital technologies and sustainability are 2.5 times more likely to be among tomorrow's strongest-performing businesses.<sup>8</sup> Even more pressing than business results is the need to reduce our impact on the planet. We have less than eight years to

stop the rise in global temperatures from reaching the 1.5°C SLO for Planet Earth. Sustainability by itself is enough to justify the work involved. As developers, we view sustainability as an important design principle and an essential part of the software development lifecycle.

## More Information

[Contact us](#) for more information about using MongoDB Atlas to incorporate more environmentally sustainable practices into your application development efforts.



8. Accenture, [Reimagining the Agenda](#)