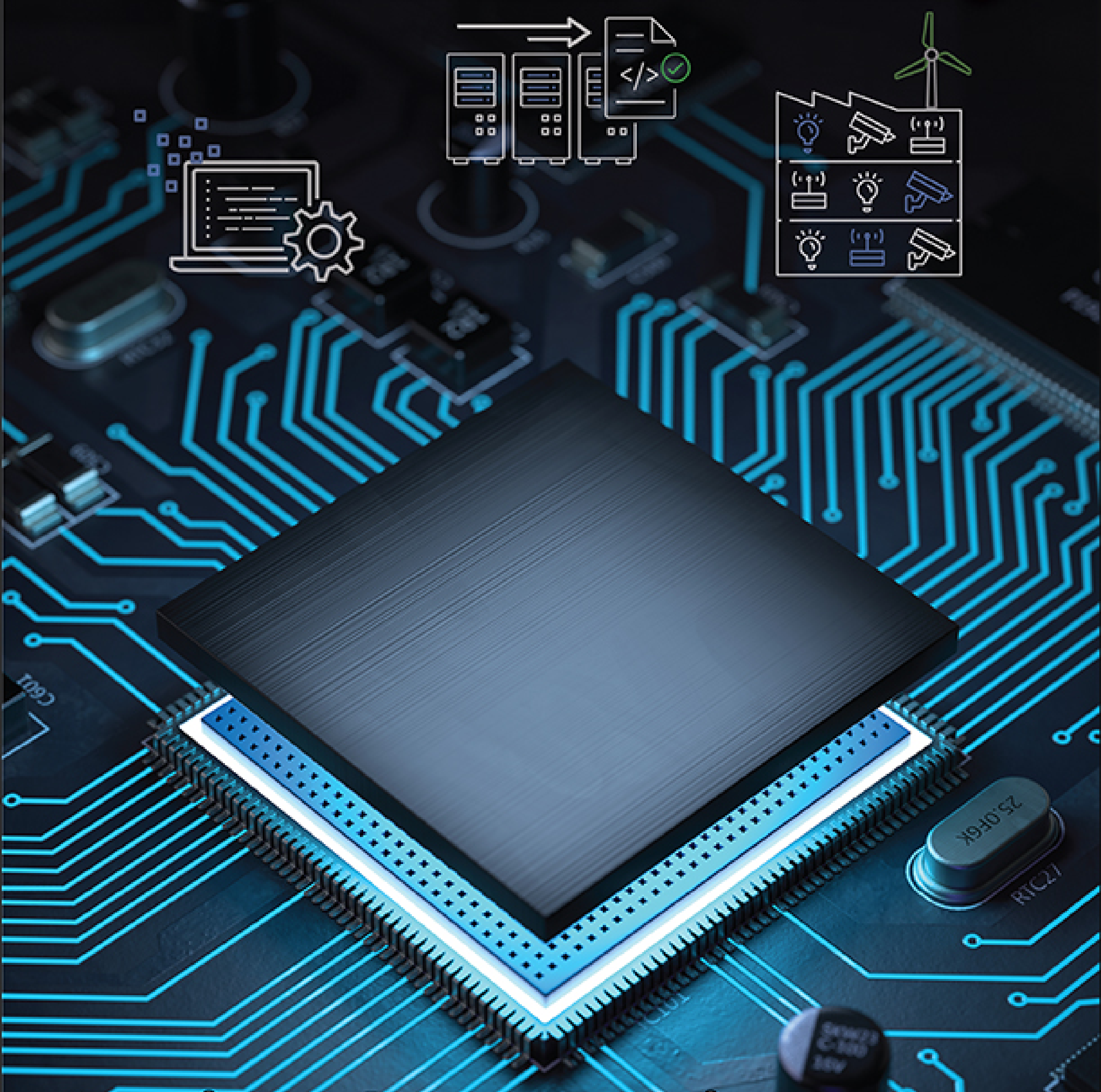
AWS Architecture Monthly



Semiconductor Design

March 2021

Notices

Editor's note

Semiconductor design and manufacturing continues to drive technology. Meeting the design requirements of low power connected devices, which utilize high performant, compact designs can be challenging. AWS provides the infrastructure and services to run the entire semiconductor design workflow. In this issue, our expert Mark Duffield shares innovation, architectural patterns and considerations for Semiconductor Design.

We hope you'll find this edition of Architecture Monthly useful. My team would like your feedback, so please give us a rating and 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.

In The Semiconductor Design issue:

- **Ask an Expert**: Mark Duffield, Tech Leader, Semiconductors, AWS
- **Reference Architecture:** Semiconductor and Electronics on AWS: AWS Services and Data Movement for Semiconductor Design
- **Whitepaper**: Semiconductor Design on AWS
- **Implementation Guide:** Run Semiconductor Design Workflows on AWS
- **Case Study:** Arm Reduces Characterization Turnaround Time and Costs by Using AWS Arm-based Graviton Instances
- **Blog**: Scaling Synopsys Proteus Optical Proximity Correction on AWS
- **Solution**: Scale-Out Computing on AWS (SOCA)
- **Reference Architecture:** Scale-out Computing on AWS
- **Whitepaper**: Using Ellexus Breeze for EDA Workload Migration to AWS
- **Reference Architecture:** Semiconductor and Electronics on AWS: Enabling Collaboration and Innovation from Customer Specification to Silicon
- **Case Study:** proteanTecs Support Tens of Millions of Simulations, Cuts Costs by Up to 60% Using AWS
- **Reference Architecture:** EDA on AWS with IBM Spectrum LSF
- **Blog:** AWS and Arm Help Semiconductor Companies Innovate Faster with Increased Flexibility
- **Related Videos:** AWS for Semiconductor & Electronics, Scaling Semiconductor Design Workflows on AWS, AWS re:Invent 2020: Semiconductor Design on AWS with Qualcomm, AWS Floor28 News – Semiconductor Special Edition

Jane Scolieri, Managing Editor

Ask an Expert

Mark Duffield, Technical Leader for Semiconductors at Amazon Web Services

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

The semiconductor industry is being driven by numerous contemporaneous elements. These have combined to encourage a global and concerted effort in the semiconductor industry to rethink how chips are designed and manufactured. Challenges in the global supply chain, reduced process geometries, and increased consumer demand have resulted in an unforeseen shortfall in semiconductors.

With the AWS Cloud, you can easily scale-out computing resources to tens of thousands of cores. You can quickly meet the exponential growth required for 5nm and smaller chip design workflows. Rather than being limited to a specific number of resources, AWS can help you scale-out workloads in ways not possible in on-premises data centers.

AWS charges by the second for compute resources, which means you pay the same whether you are scaling vertically or horizontally. Horizontal scaling is increasing the number of instances (servers/nodes) and number of cores. By scaling horizontally, jobs can finish in less time. This can dramatically reduce the time-to-market (TTM). Back-end (physical design) workloads typically consume the bulk of computing resources. We have many customers running back-end workloads on AWS to ensure they meet tape-out windows and will be first-to-market.

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

There are several AWS architectures that are being used in the semiconductor industry. Most commonly, customers are launching an environment that is similar to what they are using in their on-premises data center. This allows tool engineers to run jobs in a familiar environment. Chip design workflows can be moved to AWS with little or no customization.

This lift-and-shift approach may not initially take full advantage of the breadth and depth of AWS services. However, it allows for you to quickly see how their tools and applications will run the same as in their on-premises data centers. Another common architecture is enabling secure collaboration with third parties. Most fabless semiconductor companies will rely on their partners and third parties for some part of their design. For example, place and route or regression testing may be out-sourced to a third party. On AWS, an isolated, secure, collaborative environment can be launched in minutes. Previously, access was granted to an existing environment, and security mechanisms were implemented to lock down the environment for third parties. With AWS, you can launch a new environment that has only the required access and tools necessary for the collaborative effort.

Do you see disruption occurring the semiconductor industry?

In the 1980s, we saw the transition from Integrated Device Manufacturers (IDMs) to pure-play foundries and fabless design houses. Today, we are seeing customers abandon legacy infrastructure approaches and the outdated mindset of requiring on-premises infrastructure. Customers are launching new chip designs that are entirely in the AWS Cloud. With the exception of desktops and lab equipment (used for test, firmware, etc.), we have customers that have zero on-premises infrastructure. They have no computing resources, no storage, and no data center costs. They only have the design teams and support teams (IT support, security, etc.) to enable chip design. In addition to rethinking of infrastructure, we also see tool vendors exploring new methodologies that are driven by cloud adoption. Semiconductor design tools have remained fairly static, but we're now seeing new tool companies designing tools specifically built for cloud computing.

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

For most industries, and maybe more so for semiconductor design, data is gravity. Once you have established an AWS account with a secure connection, the next step is to determine what data will be needed for a proof of concept (POC). When building a chip design environment on AWS, we strongly encourage you to run a production workload. This provides comparative results to on-premises runs, and defines an architecture that can be used later for the entire production workflow. If you don't run a workload and use synthetic data and tools not currently in production, the effort becomes purely academic. In addition to what data should be moved, another key architectural consideration is how to access licenses. In the Implementation Guide: Run Semiconductor Design Workflows on AWS, we cover three architectures, which enable license access.

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

On AWS, semiconductor design workflows will initially take advantage of only the necessary services to tape-out the next chip. There are several significant advantages of running on AWS instead of on-premises; we have already mentioned the ability to scale-out quickly on demand. Beyond the necessary infrastructure to run flows, you can use additional AWS services to cost optimize, reduce time to market, and increase your return on investment. Once data and flows have been moved to AWS, you can take advantage of over 200 feature-rich services. For example, you can create a data lake that can be used for centralized storage, and also for an analytics pipeline. This enables data insights that are typically not available in an on-premises environment. AWS offers the broadest and deepest set of machine learning services. These services are being used for wafer defect analysis, reducing job runtimes, license cost, and usage optimization.

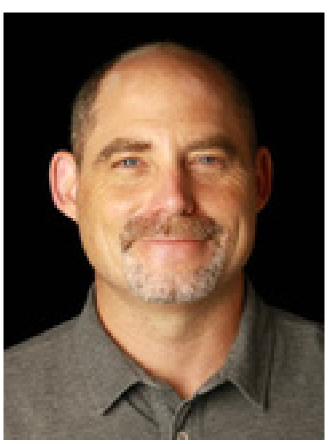
Industry-wide, as customers are moving workflows over to AWS, they are not investing in new on-premises infrastructure. Instead, they are maintaining current data center infrastructure only until end of life. Many are now in the process of moving either one workload or tool, or starting new chip designs on AWS.

Semiconductor Design is an exciting industry. Nearly all technology today is being designed with smart, connected capabilities built-in. Wind turbines, light bulbs, medical equipment, home appliances, and next generation CPUs all benefit from these new chip designs. Many of the challenges for designing future connected devices can be addressed by running on the AWS Cloud.

All of us in the AWS Semiconductor team are builders. Using the AWS Cloud, we are leading the way to a transformative shift that will enable the semiconductor industry to continue to innovate.

Let's start building.

About the expert



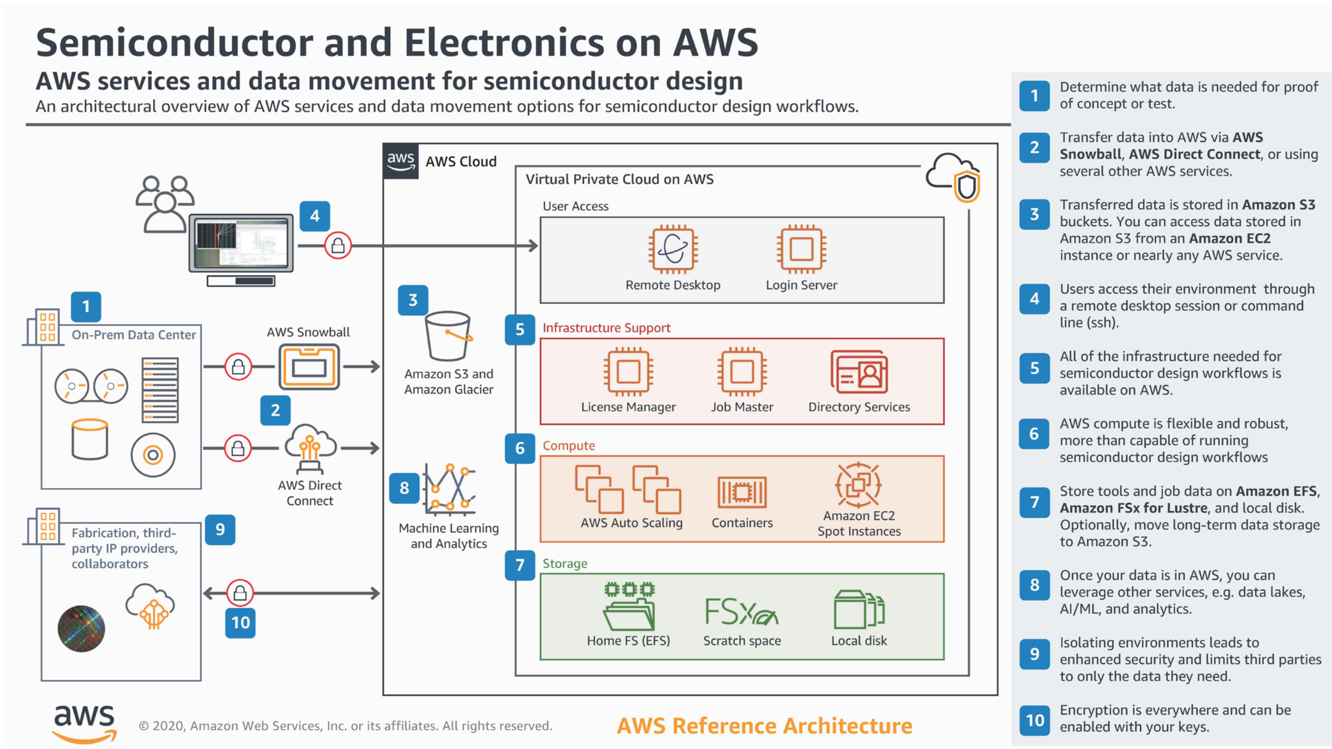
Mark Duffield is the Worldwide Tech Leader for the Semiconductor and Electronics industry at Amazon Web Services. He travels around the world working with semiconductor companies to enable workflows on AWS. He frequently speaks at industry and AWS events discussing architectural designs, market trends, and AWS features and services. He has deep experience with electronic design automation, high performance computing, enterprise software development, and distributed file systems.

Reference Architecture

Semiconductor and Electronics on AWS

AWS services and data movement for semiconductor design

An architectural overview of AWS services and data movement options for semiconductor design workflows



Find Reference Architecture online

Whitepaper

Semiconductor Design on AWS

By David Pellerin, Head WW Semiconductor, AWS, Mark Duffield, WW Tech Leader, Semiconductors, AWS, Allan Carter, WW Solutions Architect, Semiconductors, AWS, and Nafea Bshara, VP/Distinguished Engineer, AWS

Abstract

Semiconductor and electronics companies can significantly accelerate their product development lifecycle and time to market by taking advantage of the near-infinite compute power, storage, and resources available on Amazon Web Services (AWS). This whitepaper presents an overview of the semiconductor design flow, a migration path for moving design and verification workflows to AWS, and the AWS architectural components needed to optimize semiconductor design workloads on AWS.

Introduction

The workflows, applications, and methods used for the design and verification of semiconductors, integrated circuits (ICs), and printed circuit boards (PCBs) have been largely unchanged since the invention of computer-aided engineering (CAE) and electronic design automation (EDA) software. The computing requirements, however, have dramatically increased as device geometries have shrunk and electronics systems and integrated circuits have become more complex. CAE, EDA, and emerging workloads such as computational lithography and metrology have driven the need for massive-scale computing and data management in next-generation connected products.

IT and EDA support organizations face the challenge of providing the infrastructure required to run workflows in a way that meets schedule and budget requirements. They must invest in increasingly large server farms and high-performance storage systems to enable high quality, fast turnaround of workflows. A lot of overall design time is spent verifying components. Workflows like the characterization of intellectual property (IP) cores, functional verification, and timing analysis have a spiky demand and limit engineering productivity. This creates a need to have enough compute capacity to minimize the time that engineers wait for results, but can result in underutilization of resources between runs of the workflows. New and upgraded IC fabrication technologies have increased peak compute and storage requirements, challenging organizations to find ways to meet the needs of silicon development teams while managing costs.

Semiconductor companies today use AWS to deploy infrastructure on an as-needed basis with pay-as-you-go pricing. The massive scale of AWS enables them to run CAE and EDA workflows as quickly as possible with no upfront capital expenditures or long term commitments. They benefit from a more rapid, flexible deployment of CAE and EDA infrastructure to run the complete IC design flow, from register-transfer-level (RTL) design to the delivery of GDSII files to a foundry for chip fabrication. AWS gives them access to the latest compute, storage, network technologies, and higher-level services that enable them to meet the ever-increasing demands of semiconductor design workloads so they can innovate faster.

Companion implementation guide

Although this whitepaper provides some level of detail for running semiconductor design workflows on AWS, it does not dive deep into specific architectural details. Instead, AWS provides a companion guide that gives detailed architectural guidance for building your semiconductor design infrastructure on AWS. The companion implementation guide can be found here: [Run Semiconductor Design Workflows on AWS].

Semiconductor design overview

Semiconductor design workloads comprise workflows and a supporting set of software tools that enable the efficient design of microelectronics, and in particular, semiconductor ICs. Semiconductor design and verification relies on a set of commercial or open-source tools, collectively referred to as EDA software, which expedites and reduces time to silicon tape-out and fabrication. EDA is a highly iterative engineering process that can take months, and in some cases years, to produce a single integrated circuit.

The increasing complexity of integrated circuits has resulted in an increased use of preconfigured or semi-customized hardware components, collectively known as IP cores. These cores (provided by IP developers as generic gate-level netlists) are either designed in-house by a semiconductor company, or purchased from a third-party IP vendor. IP cores themselves require EDA workflows for design and verification, and to characterize performance for specific IC fabrication technologies. These IP cores are used in combination with IC-specific, custom-designed components, to create a complete IC that often includes a complex system-on-chip (SoC), making use of one of more embedded CPUs, standard peripherals, input/output (I/O), and custom analog and/or digital components.

The complete IC itself, with all its IP cores and custom components, then requires large amounts of processing for full-chip verification, including modeling (simulating) all the components within the chip. This modeling, which includes hardware description language (HDL) source-level validation, physical synthesis, and initial verification (for example, using techniques such as formal verification), is known as the *front-end design*.

The physical implementation, which includes floor planning, place and route, timing analysis, design-rule-check (DRC), and final verification, is known as the *back-end design*. When the back-end design is complete, a file is produced in GDSII format. The production of this file is known as *tape-out*. When completed, the file is sent to a fabrication facility (called a *foundry*), which may or may not be operated by the semiconductor company, where a silicon wafer is manufactured. This wafer, containing perhaps thousands of individual ICs, is then inspected, cut into dies that are themselves tested, packaged into chips that are tested again, and assembled onto a board or other system through highly automated manufacturing processes.

All of these steps in the semiconductor and electronics supply chain can benefit from the scalability of the cloud.

Read Full Whitepaper online

Implementation Guide

Run Semiconductor Design Workflows on AWS

By Mark Duffield, Worldwide Tech Leader, Semiconductors, AWS, Matt Morris, Senior HPC Solutions Architect, AWS, David Pellerin, Principal Business Development for Infotech/Semiconductor, AWS and Nafea Bshara, VP/Distinguished Engineer, AWS

Abstract

This implementation guide provides you with information and guidance to run production semiconductor workflows on AWS, from customer specification, to front-end design and verification, back-end fabrication, packaging, and assembly. Additionally, this guide shows you how to build secure chambers to quickly enable third-party collaboration, as well as leverage an analytics pipeline and artificial intelligence/machine learning (AI/ML) services to decrease time-to-market and increase return on investment (ROI). Customers that run semiconductor design workloads on AWS have designed everything from simple ASICs to large SOCs with tens of billions of transistors, at the most advanced process geometries. This guide describes the numerous AWS services involved with these workloads, including compute, storage, networking, and security. Finally, this paper provides guidance on hybrid flows and data transfer methods to enable a seamless hybrid environment between on-premises data centers and AWS.

Migration methodology

When you begin the migration of your semiconductor design workflows to AWS, you will find there are many parallels with managing traditional deployments across multiple sites within your corporate network, whether these sites are distributed engineering locations, or represent an entire data center. Larger organizations in the semiconductor industry typically have multiple data centers that are geographically dispersed because of the distributed nature of their design teams. These organizations typically choose specific workloads to run in specific locations, or replicate and synchronize data to allow for multiple sites to take the load of large-scale, global workflows. Geographically distributed teams may not lend itself to a simple, and straight-forward approach for migrating workflows to AWS. We encourage our customers to look at specific parts of their design flow or even a new project, when considering which workloads to migrate to AWS. Choosing the right tool or workflow will often require, for example, determining what data replication, caching, and license server management will be needed to run the flow on AWS. Most of the same approaches and design decisions related to multiple data centers also apply to the cloud. With AWS, you can build one or more virtual data centers that mirror your existing on-premises electronic design automation (EDA) design environment and data center infrastructure. The foundational technologies that enable compute resources, storage servers, and user workstations are available with just a few keystrokes. This ability to rapidly create new semiconductor design and verification environments in just minutes is a major benefit of deploying on cloud. However, the real power of using the AWS Cloud for semiconductor design comes from the dynamic capabilities and enormous scale provided by AWS, resulting in faster time-to-results, reduced schedule risk, and more efficient utilization of valuable EDA software licenses.

Companion guide

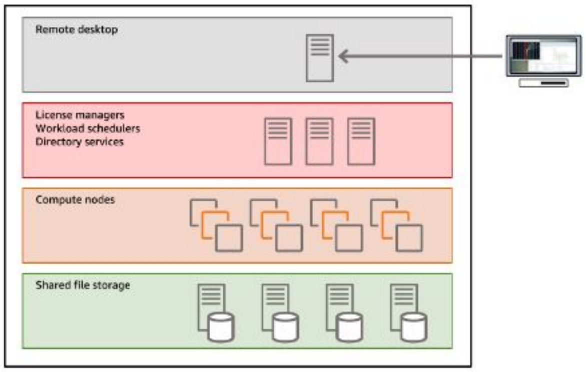
This implementation guide serves as a companion to the Semiconductor Design on AWS whitepaper. Although not necessary, you should consider reading the whitepaper first to ensure you have the building blocks for this guide.

Supplemental GitHub repository

In addition to the whitepaper and this guide, refer to the Semiconductor Design on AWS GitHub repository for low-level commands for operating system tuning and tool optimization.

Components of an on-premises semiconductor design environment

The following figure shows a simple view of a typical on-premises semiconductor and electronics design environment. This diagram has the necessary components to run the entire workflow. As you continue through this guide, make sure to reference this diagram as a comparison to the infrastructure that is launched on AWS. That is, all of the components in this diagram are launched and further optimized on AWS.



Traditional on-premises environment

Migrating and running semiconductor workflows on AWS

Migrating workflows or even one tool to AWS can be a somewhat daunting task, particularly if you are new to AWS. As AWS has well over 200 services, and each service is extremely feature rich, this section narrows the scope and focuses on what you need to know for the semiconductor industry.

Read Full Implementation Guide online

Case Study

Arm Reduces Characterization Turnaround Time and Costs by Using AWS Arm-based Graviton Instances

Arm is a leading technology provider of silicon intellectual property (IP) for intelligent systems-on-chip that power billions of devices. Arm creates IP used by technology partners to develop integrated semiconductor circuits. The company estimates that 70 percent of the world's population uses its technology in their smart devices and electronics.

For many years, Arm relied on an on-premises environment to support electronic design automation (EDA) workloads, resulting in forecast challenges on compute capacity. "The nature of our Physical Design Group business demands a high-dynamic compute environment, and the flexibility to make changes on short notice," says Philippe Moyer, vice president of design enablement for the Arm Physical Design Group. "In the past, the on-premises compute was sometimes sitting idle until the need arose, which is why the scalability and agility of the cloud is a good solution for our business."

Arm was looking for agility improvement to keep development on schedule. "With our on-premises environment, our data center was constrained in terms of scalability, and deployment of additional compute capacity would typically take one month for approvals and at least three months to procure and install hardware," says Vicki Mitchell, vice president of systems engineering for Arm. "We have aggressive deadlines, and waiting that long could make or break a project for us."

> "Using AWS, our EDA workload characterization turnaround time was reduced from a few months to a few weeks."
>
> **Philippe Moyer**
> *Vice President of Design Enablement, Arm*

Moving EDA Workloads to the AWS Cloud

To gain the agility and scalability needed, in 2017 Arm chose to move part of its EDA workload to Amazon Web Services (AWS). "Selecting AWS made sense to us. AWS is a market leader, and it really understands the semiconductor space," says Mitchell. "We were also very impressed with the EDA knowledge of the AWS solution architects we worked with."

Initially, the Arm Physical Design Group ran its EDA workloads on Amazon Elastic Compute Cloud (Amazon EC2) Intel processor–based instances. It also used Amazon Simple Storage Service (Amazon S3), in combination with Amazon Elastic File System (Amazon EFS), for EDA data storage. When AWS announced the availability of Amazon EC2 A1 instances powered by Arm-based Graviton processors, the Arm Physical Design IP team began to run portions of its EDA workloads on A1 instances. “Taking advantage of Graviton instances gives us the opportunity to contribute to the development of the EDA ecosystem on Arm architecture,” says Moyer. In addition, Arm uses Amazon EC2 Spot Instances for all workloads. Spot Instances are spare compute capacity available at up to 90 percent less than On-Demand Instances.

Reducing Characterization Turnaround Time from Months to Weeks

By using AWS, the Arm Physical Design IP team can scale its EDA environment up or down quickly—from 5,000 cores to 30,000 cores—on demand. “This scalability and flexibility brought by AWS translates to a faster turnaround time,” says Moyer. “Using AWS, our EDA workload characterization turnaround time was reduced from a few months to a few weeks.”

Enabling Experimentation and Innovation

With the company’s on-premises environment, Arm engineers sometimes had to wait for compute resources to begin working on projects. By using on-demand compute capacity, engineers are now free to innovate. “It’s much easier for our engineers to prototype and experiment in the cloud,” Mitchell says. “If they’re trying to validate a piece of logic or create a new feature, they can take advantage of Amazon EC2 Spot Instances to submit a job and get instantaneous scheduling without disrupting the project flow. They can move faster as a result.”

Decreasing AWS Costs by 30%

Running its EDA workloads on Arm-based Graviton instances, Arm is lowering its AWS operational costs. “The Graviton processor family enables us to reduce the AWS costs for our logic characterization workload by 30 percent per physical core versus using Intel-powered instances for the same throughput,” says Moyer.

Arm now plans to use the next generation of Amazon EC2 Arm instances, powered by Graviton2 processors with 64-bit Arm Neoverse cores. “The Graviton2 offers even better performance and scalability and caters to a larger number of different EDA workloads,” Moyer says. “We are looking forward to using these AWS processors for better performance and additional cost savings.”

Read Case Study online

Blog

Scaling Synopsys Proteus Optical Proximity Correction on AWS

By Ahmed Elzeftawi, George Bailey, and Richard Paw

Introduction

Photolithography is a key step in the manufacturing of semiconductor chips. Photolithography works by shining light (laser) through a pattern (mask) onto a silicon wafer with photosensitive coating (resist). This changes the properties of the coating allowing manufactures to chemically remove the parts of the coating based on the exposure or lack of exposure to the laser. This forms a pattern on the wafer that matches the pattern on the mask. The manufacturer uses these patterns to selectively add or remove layers onto the wafer to form different devices on silicon.

This process remained largely unchanged until the geometry of the patterns on the chip started to approach the wavelength of the laser used. As these dimensions approached the wavelength of light, basic photolithography techniques could no longer accurately reproduce the patterns on the wafer due to diffraction and other distortions. In order to compensate, optical proximity correction (OPC) calculates the effects of these distortions on the final image and modifies the patterns on the mask. To accomplish this, OPC software, such as Synopsys Proteus, must execute billions of calculations across dozens of mask layers to compensate for the distortions introduced as semiconductor geometries continue to shrink.

This level of complexity makes OPC one of the most computationally demanding workloads in semiconductor manufacturing, often requiring thousands of compute cores running for hours to process a single semiconductor chip. Due to the massive scale of compute required, semiconductor foundries devote a significant portion of their data centers to this single workload.

Scaling Synopsys Proteus OPC on AWS

Synopsys and AWS recognize that as the advances in semiconductor technology continue to push the complexity of each chip, customer data centers face the challenge of keeping up with the growing demand on their resources. OPC is a natural fit to leverage the infinite compute scale of the AWS Cloud as the computation can be parallelized. Synopsys and AWS decided to launch a joint investigation to determine how Synopsys Proteus scales on AWS.

We decided to target scaling across 24,000 compute cores on a single design with a goal of maintaining at least 95% linearity at scale. Validating how a single design scales provides an accurate picture of the benefits customers can achieve as it takes into account the infrastructure and design workload interdependencies as you add workers to the same job. We already know that loosely coupled workloads scale nearly linearly on AWS.

Amazon EC2 Spot Instances are a great way to optimize compute costs for fault tolerant workloads. EC2 Spot Instances use spare Amazon EC2 capacity, which is available for up to a 90% discount over On-Demand Instances. When there is a spike in requests for a particular On-Demand Instance type in a specific Availability Zone (AZ), AWS can reclaim the Spot Instances with a two-minute notification.

We leveraged the official AWS Solution Scale-Out Compute on AWS (SOCA) to quickly create a readily available cloud environment that provides scalability of compute and storage, budget monitoring, job scheduling, etc. For guidance on how to set up SOCA for an EDA workload, check out the blog Scaling EDA Workloads using Scale-Out Computing on AWS.

Proteus runs using a distributed computing architecture. There is a head node that manages and tracks the workload and data while dispatching the individual compute jobs to worker nodes. Each worker node receives the data for a small part of the mask, processes the workload, and returns the finished data back to the head node.

We divided our investigation into two steps – 1) achieve scalability across 24,000 cores using On-Demand Instances and 2) optimize the architecture for cost using EC2 Spot. We started by running the tests with 2000 cores, then 4000, 8000, 10,000, 16,000, and finally with 24,000 cores using On-Demand Instances. For information on the design we used to test, see the table below labeled, Design Details. We were able to successfully scale Proteus to 24,000 and we were able to maintain scaling linearity of over 98% even with 24,000 cores. See the red solid line in figure 1.

Next, we explored running Proteus using Amazon EC2 Spot Instances in order to scale the workload more cost effectively. Starting with version vM-2017.03-9-T-20200117, Synopsys Proteus is now able to leverage the EC2 Spot Instances for the worker nodes. We proceeded to run the same set of tests using EC2 Spot Instances instead of On-Demand Instances. See the black solid line in figure 1. The results show that Spot interruptions did result in some reduction of overall efficiency. However, even with Spot interruptions, Proteus was still able to achieve over 97% scalability when scaling to 24,000 cores on a single design. We took advantage of Spot Fleet, which enables diversification of instance types to minimize the impact of Spot interruptions. Figure 2 shows the Spot instance types used for the duration of the 24,000 cores test.

Design Details

Recipe	EUV OPC
Layout	Synopsys N7 IP block
Node	N7

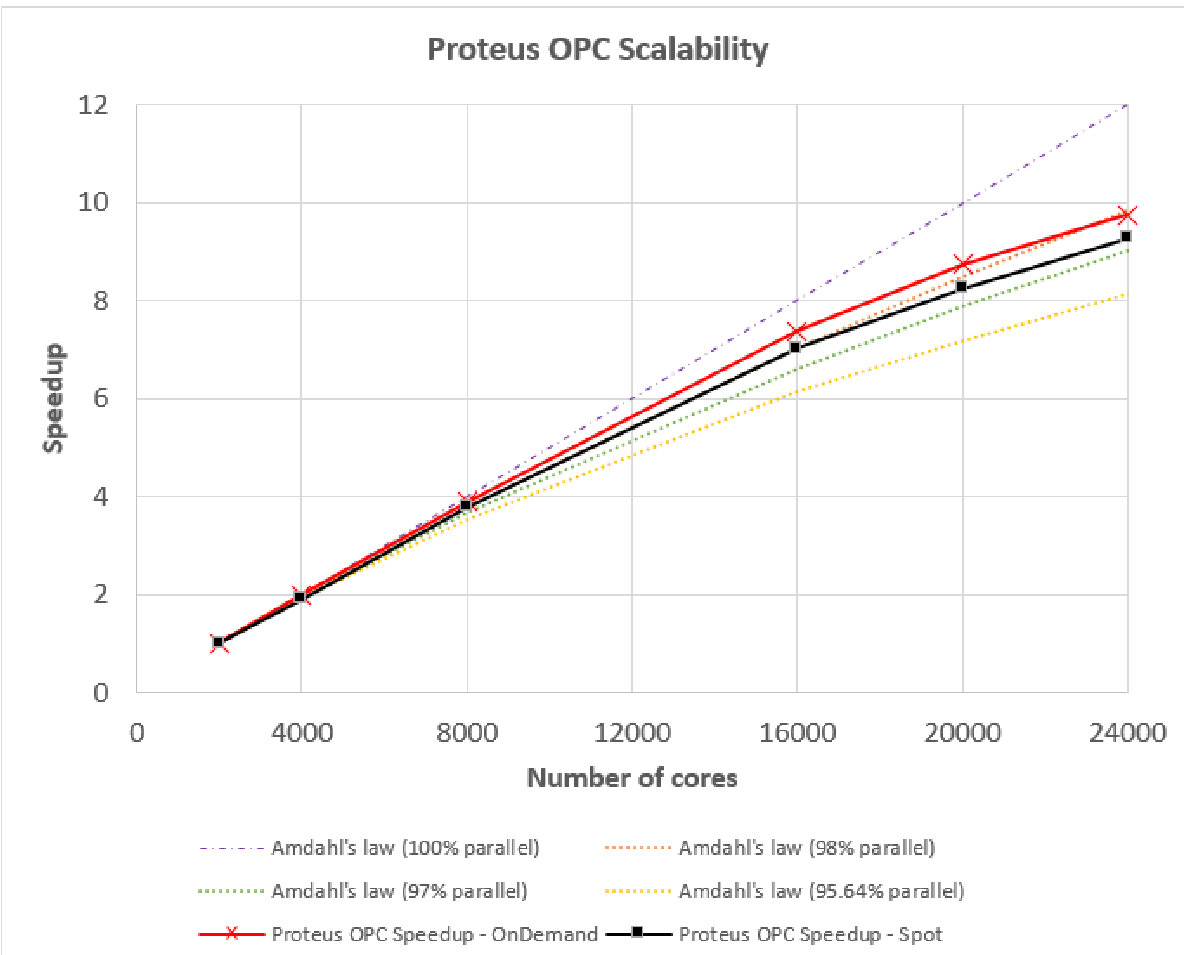


Figure 1: Proteus OPC Scalability on Amazon EC2 On-Demand and Spot Instances

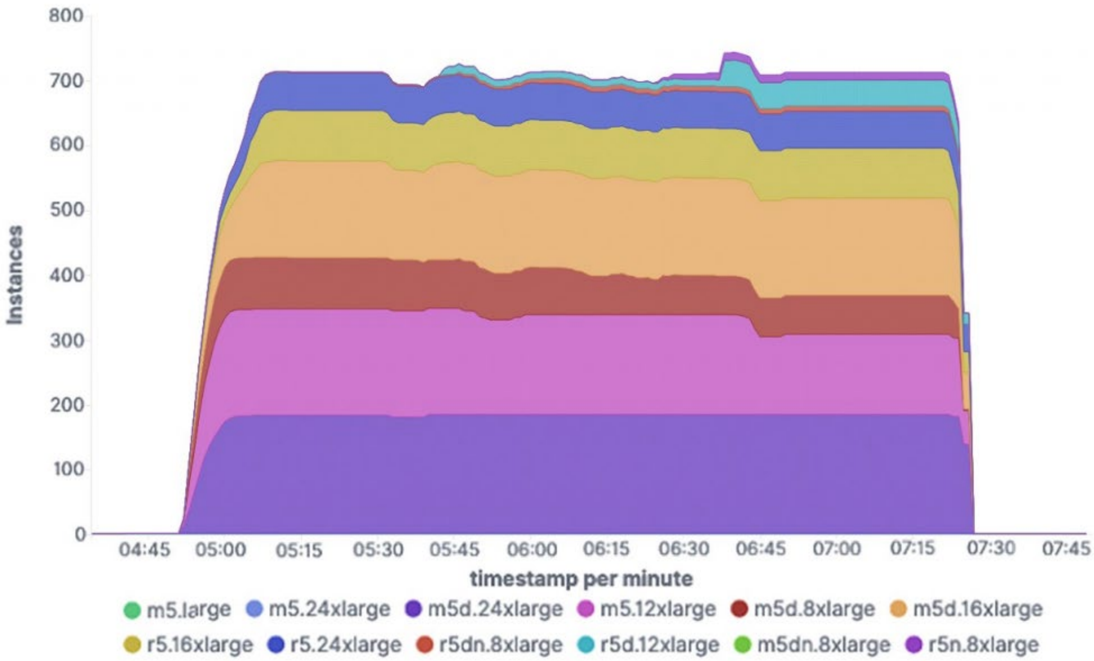


Figure 2: Amazon EC2 Spot instance type distribution for 24,000 cores

Conclusion

OPC is one of the more compute intensive workloads in semiconductor manufacturing, consuming a significant portion of foundry data center capacity. The scale of OPC workloads makes them a natural fit for running on the cloud. The joint Synopsys and AWS team showed that Proteus can successfully scale to the targeted 24,000 cores for a single design while maintaining 98% scalability. Proteus can even successfully leverage the cost savings of Amazon EC2 Spot Instances while achieving over 97% scaling linearity. By leveraging AWS for Synopsys Proteus OPC, foundries and integrated device manufacturers (IDM) have the flexibility and elasticity to scale their OPC workloads across more cores than they could in their own data centers thereby reducing the total turn-around time while still realizing lower total compute cost than running on-premises.

While we achieved the target of this investigation, 24,000 cores on a single design with EC2 Spot, the Synopsys and AWS teams feel that we can go even farther with Synopsys Proteus on AWS. Stay tuned for future updates. For further information about how we scaled this and other EDA workloads on AWS, or just to how to migrate EDA workloads to AWS, reach out to your Synopsys or AWS account teams.

Read Blog post online

Solution

Scale-Out Computing on AWS (SOCA)

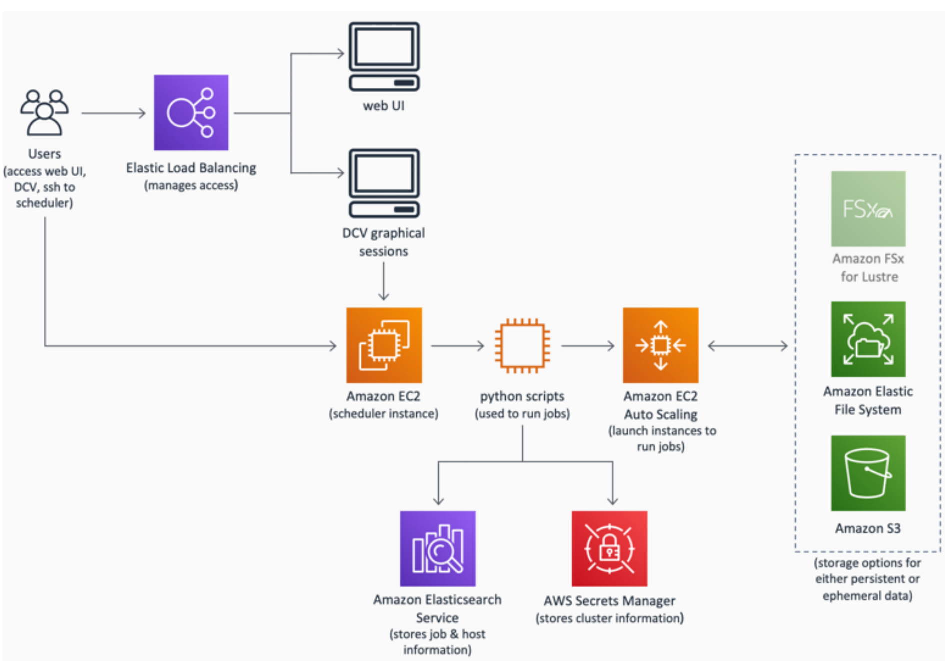
Scale-Out Computing on AWS (SOCA)

What does this AWS Solutions Implementation do?

The Scale-Out Computing on AWS solution helps customers easily deploy and operate a multiuser environment for computationally intensive workflows such as Computer-Aided Engineering (CAE). The solution features a large selection of compute resources, a fast network backbone, unlimited storage, and budget and cost management directly integrated within AWS. This solution deploys a user interface (UI) with cloud workstations, file management, and automation tools that enable you to create your own queues, scheduler resources, [Amazon Machine Images](#) (AMIs), and management functions for user and group permissions.

AWS Solutions Implementation overview

The diagram below presents the architecture you can automatically deploy using the solution's implementation guide and accompanying AWS CloudFormation template.



Scale-Out Computing on AWS architecture

At its core, this solution implements a scheduler Amazon Elastic Compute Cloud (Amazon EC2) instance, which leverages AWS CloudFormation and Amazon EC2 Auto Scaling to automatically provision the resources necessary to execute cluster user tasks, such as scale-out compute jobs and remote visualization sessions.

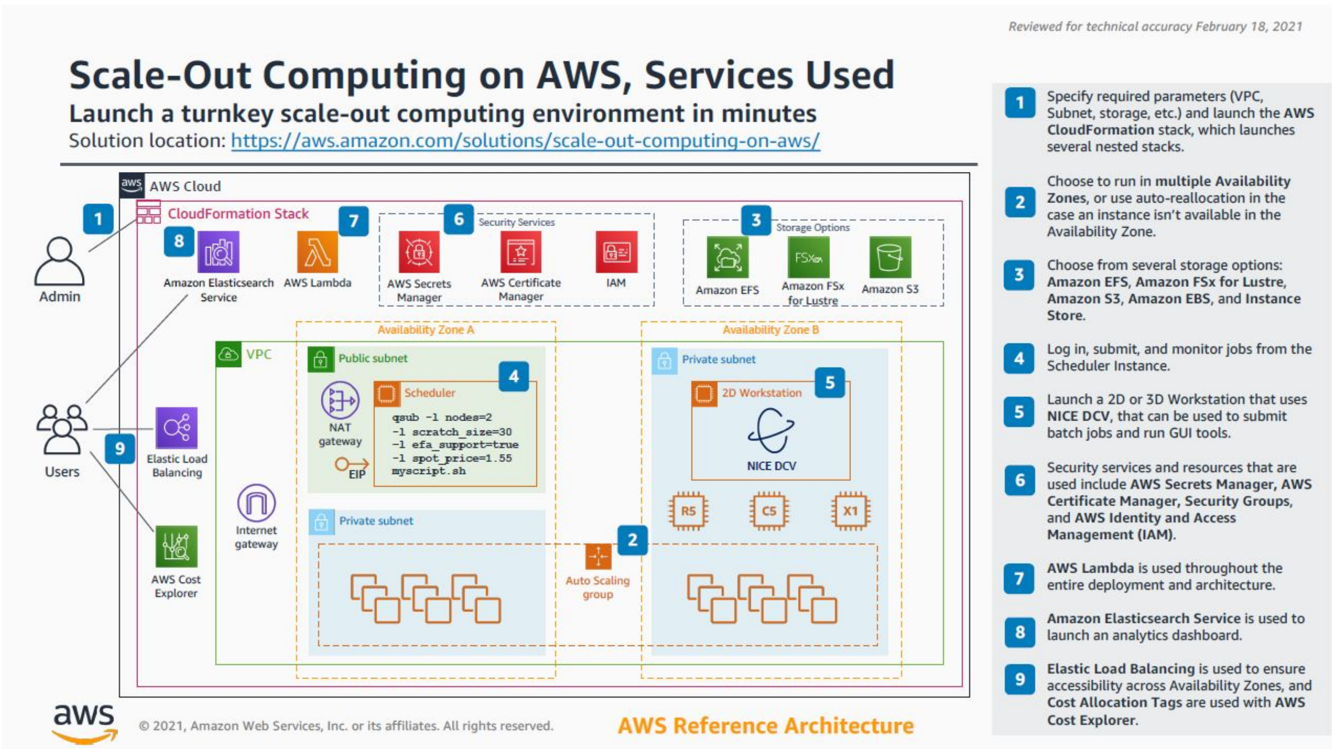
This solution also deploys Amazon Elastic File System (Amazon EFS) for persistent storage; AWS Lambda functions to verify the required prerequisites and create a default signed certificate for an Application Load Balancer (ALB) to manage access to Desktop Cloud Visualization (DCV) workstation sessions; an Amazon Elasticsearch Service (Amazon ES) cluster to store job and host information; and AWS Secrets Manager to store the solution configuration files. The solution also leverages AWS Identity and Access Management (IAM) roles to enforce least privileged access.

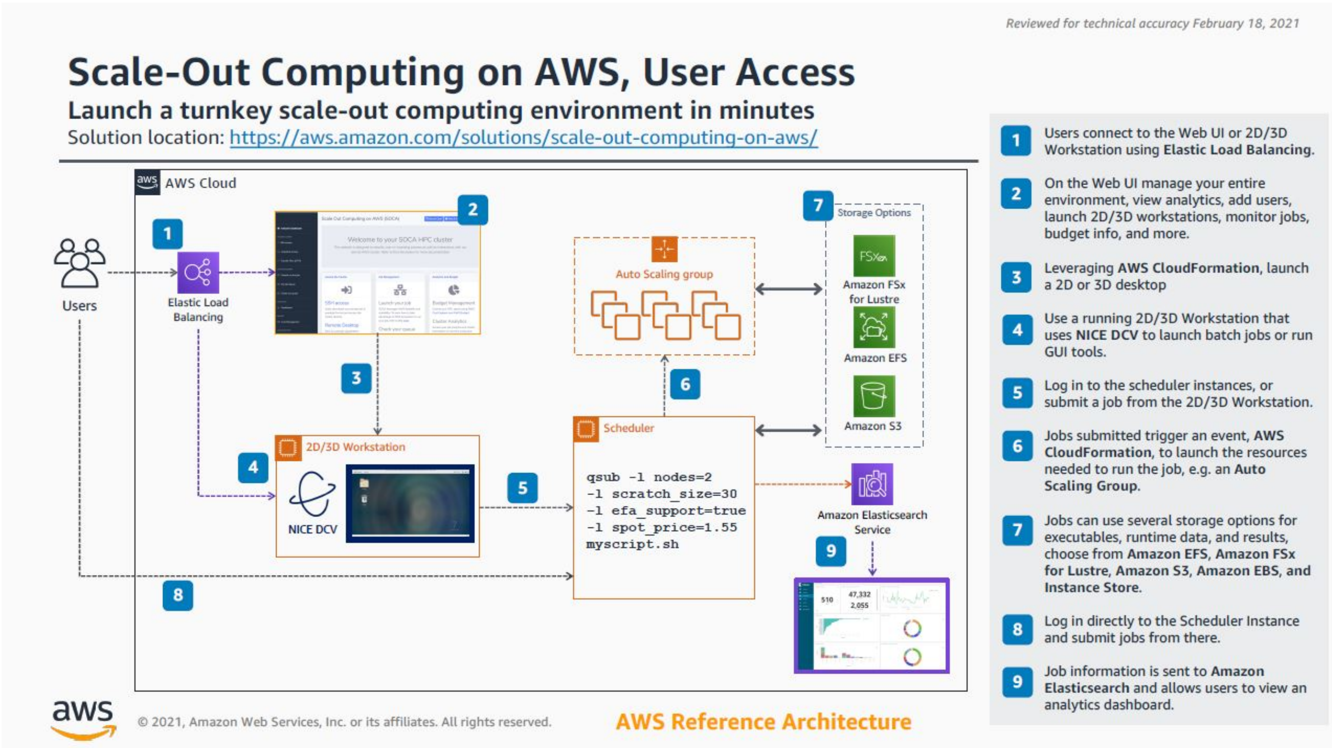
View Implementation Guide online

Reference Architecture

Scale-out Computing on AWS

Launch a turnkey scale-out computing environment in minutes





View Reference Architecture online

Whitepaper

Using Ellexus Breeze for EDA Workload Migration to AWS

By Dr. Rosemary Francis, CEO, Ellexus Ltd, Mark Duffield, Tech Lead for Semiconductor and Electronics, AWS, and Faraz Angabini, Solutions Architect, AWS

Abstract

This whitepaper outlines the best practices for migrating Electronic Design Automation (EDA) workloads to AWS using the I/O profiling and dependency analysis tool suite Breeze from Ellexus. Profiling the EDA tool on premises and in the cloud ensures that the AWS configuration is right sized for the application and that costs are optimized.

Introduction

Ellexus provides I/O profiling solutions to monitor, profile, discover, and debug scientific and engineering workloads and high performance computing (HPC) systems.

Ellexus Breeze is an offline analysis tool that profiles workflows to collect file system dependencies as well as CPU, memory, and I/O requirements. This allows you to analyze and migrate legacy EDA tools, and right-size cloud resources for the entire workflow.

The first challenge when migrating a workload to a new environment is making it work. The second challenge is making it work well. Breeze addresses both of these challenges by giving you the information you need to make informed choices and troubleshoot issues quickly.

Although a deep understanding of Electronic Design Automation (EDA) workflows is not necessary, you should have a basic understanding of EDA tools and their underlying infrastructure.

Addressing migration challenges with Breeze

The following metrics are collected by Breeze, which are vital when migrating workloads to the cloud:

- File and network dependencies
- Process tree and scripted infrastructure
- Program arguments and environments

Customers often use a manual process, and leverage the data from Breeze to run an initial proof of concept (POC) on AWS with an environment that is similar to their existing on-premises infrastructure. After the initial POC is completed, the Breeze Command Line Interface (CLI) can be used to automate the process and migrate the remaining workflows.

Application dependencies and containerization

Often, EDA tools can form legacy workflows that make it difficult to determine which files, libraries, and applications are needed. Breeze addresses this issue by automatically detecting which files, libraries, and applications are needed. This can be as fine grain as listing every file, or a more high-level approach can be taken, determining only the mount points that need to be replicated on AWS.

Once the dependencies have been collected with Breeze, you can containerize the workflow. What to include and exclude from the container depends on your long-term data management strategy, and how the container will be used. If the same rules can be applied to multiple workflows, the containerization process can be automated using the workflow dependency output from the Breeze CLI. More details on dependency output are provided later in this whitepaper.

Debugging licensing issues and other changes in environment

Moving an EDA workflow to new infrastructure usually involves changes to the workflow environment. For example, connecting to a license server on AWS might result in quicker access to licenses, and jobs completing faster. Problems that arise could be related to the new environment, the on-premises IT infrastructure, the EDA tool, or the scripted flow that orchestrates it. Breeze helps to debug these problems by capturing the application and how it runs. For example, a file permissions issue would show up as a failed file open. A missing library would be similarly highlighted, and if the tool is looking in the wrong place for a file or license, Breeze will tell you.

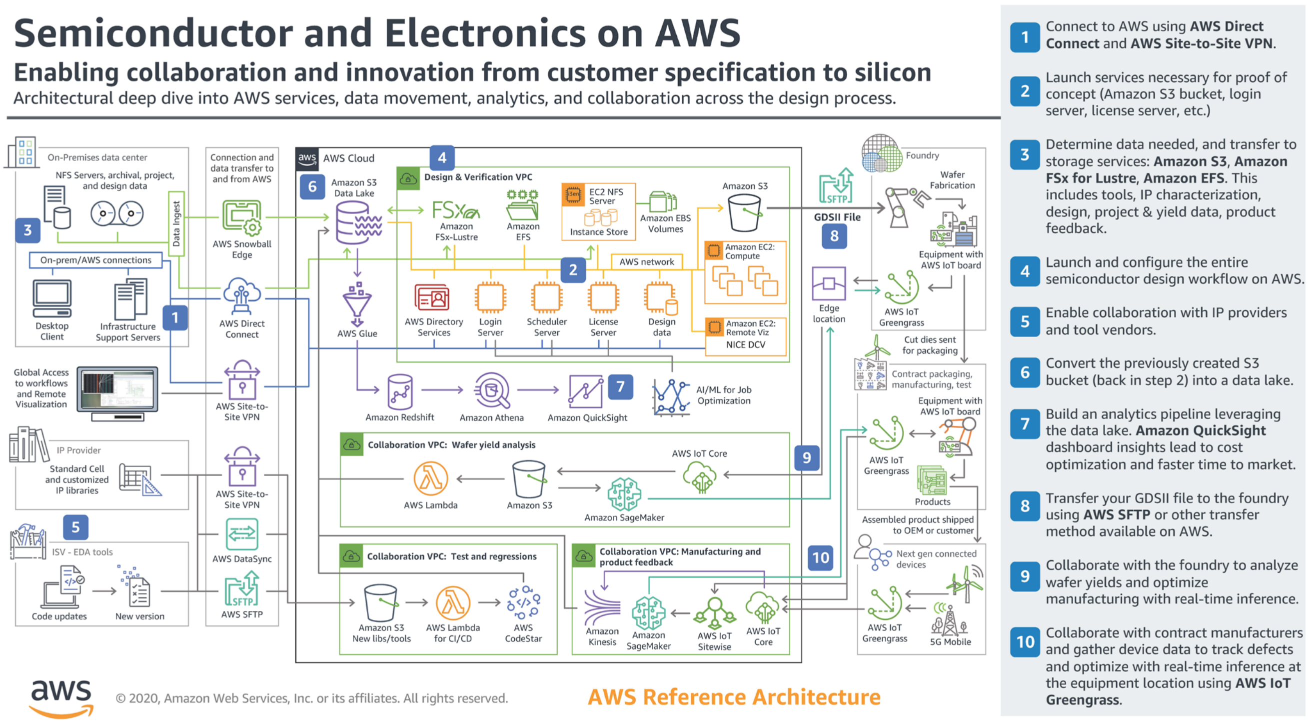
Read Full Whitepaper online

Reference Architecture

Semiconductor and Electronics on AWS

Enabling collaboration and innovation from customer specification to silicon

Architectural deep dive into AWS services, data movement, analytics, and collaboration across the design process.



[View Reference Architecture online]()

Case Study

proteanTecs Support Tens of Millions of Simulations, Cuts Costs by Up to 60% Using AWS

A Startup Seeks Flexibility and Speed

proteanTecs is a startup company and provider of a software solution that uses deep data and machine learning to predict failures in electronics. Microchip manufacturers use the solution to combine data, derived from agents (IPs) embedded in chips, with predictive artificial intelligence (AI) and data analytics to track faults before they become failures, thus increasing quality and reliability of mission-critical electronics. The company also develops electronic design automation (EDA) software tools to support its solution.

As part of its development efforts, proteanTecs needed the flexibility and scalability to support high-performance computing (HPC) workloads that run millions of simulations each day. "We process a large volume of data for each chip and then process machine-learning algorithms from that data," says Yuval Bonen, co-founder and vice president, software for proteanTecs. "We might need to run tens of millions of simulations one week and only a few million the following week," Bonen says. "We really need the flexibility to accommodate our changing requirements." The company must also meet its customers' deadlines for EDA software delivery. "We have to deliver projects on time every time," says Bonen. As a startup, proteanTecs also wanted to reduce its operational costs as much as possible while remaining prepared for rapid business growth.

> **"Based on our customers' requirements, we may need to run tens of millions of simulations in one week and only several million the next week. AWS gives us the flexibility to do that without any effort at all."**
>
> - Yuval Bonen, Co-founder and Vice President, Software, proteanTecs

Running SaaS and EDA Solutions on AWS

proteanTecs conducted a proof of concept (POC), with assistance from Amazon Web Services (AWS) architects, which showed how a migration to the AWS Cloud would increase performance and scalability. "AWS worked closely with us throughout the POC to help us harness the benefits of the cloud in EDA workloads," says Bonen. "We realized that AWS technology was a great fit for us in terms of performance, reliability, and flexibility."

Following the POC, proteanTecs moved its HPC workloads to AWS, running on Intel Xeon processor–powered Amazon Elastic Compute Cloud (Amazon EC2) C5 Instances. The company uses Amazon Relational Database Service (Amazon RDS) to store application data and runs a Kubernetes container orchestration system on Amazon EC2.

To optimize costs, proteanTecs recently began using Amazon EC2 Spot Instances to support its workloads. Spot Instances are spare Amazon EC2 capacity offered at up to a 90 percent discount compared to On-Demand prices.

Performing Tens of Millions of Simulations in Parallel Weekly

By using AWS, proteanTecs can scale to run tens of millions of simulations in parallel every week. “We can easily support our heavy compute and memory requirements by running our HPC workloads on Amazon EC2,” says Bonen. The company also has the flexibility to automatically scale up or down. “Based on our customers’ requirements, we may need to run tens of millions of simulations in one week and only several million the next week,” says Bonen. “AWS gives us the flexibility to do that without any effort at all.”

Speeding Time-to-Market

proteanTecs can quickly provision new compute and storage resources, which speeds time-to-market for delivering its software-as-a-service (SaaS) and EDA solutions to customers. “Running on AWS, we can provision resources within minutes, which helps us accelerate time-to-delivery for our customers,” says Bonen. “For example, we had a customer that was pushing us to deliver a project two weeks earlier than scheduled, and we met that deadline—something we could not have done without AWS. We can scale quickly and consume twice as many compute and storage resources on AWS than we could in an on-premises environment.”

Cutting Costs by 60%

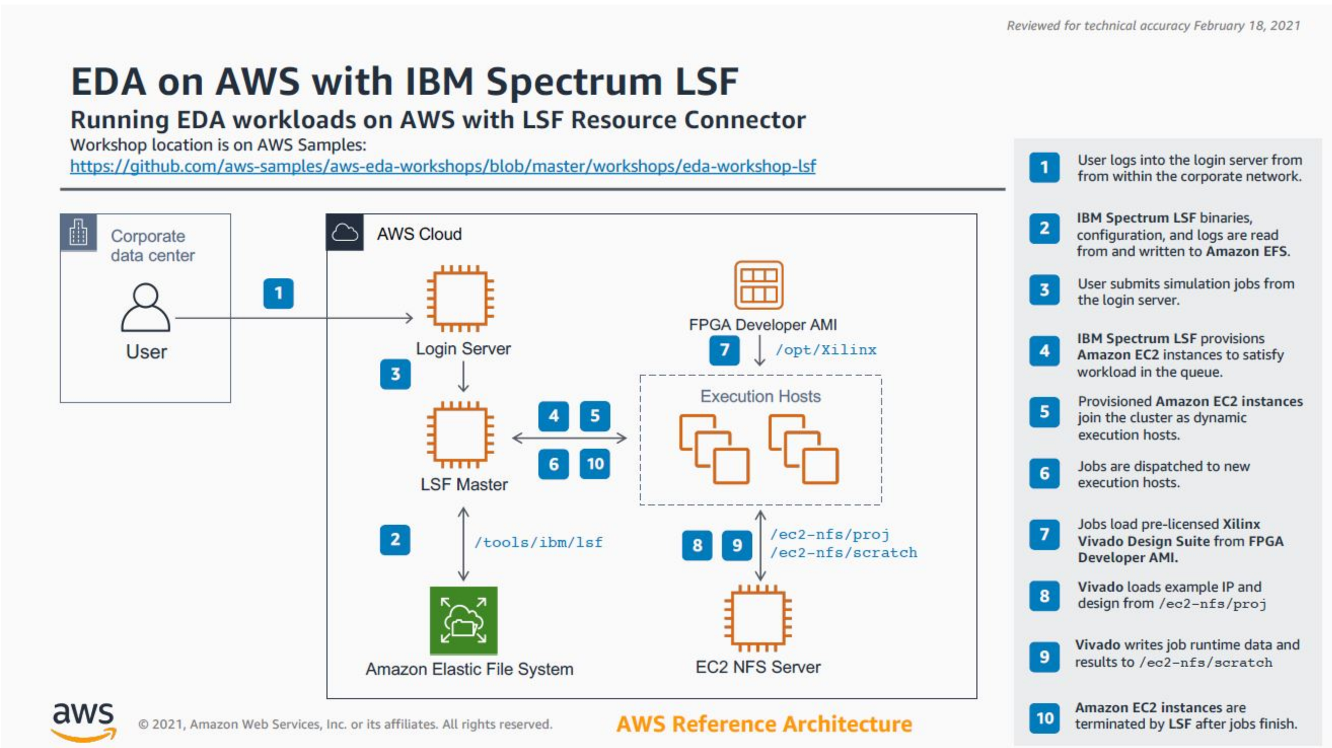
By taking advantage of Amazon EC2 Spot Instances, proteanTecs is running its environment more cost-effectively. “We are reducing our operational costs by up to 60 percent using Amazon EC2 Spot Instances,” says Bonen. “We use those savings to invest more resources into research and development and create new product offerings that will differentiate our company.”

Read Case Study online

Reference Architecture

EDA on AWS with IBM Spectrum LSF

Running EDA workloads on AWS with LSF Resource Connector



[View Reference Architecture online](#)

Blog

AWS and Arm Help Semiconductor Companies Innovate Faster with Increased Flexibility

By Caroline Lawrence and David Pellerin

AWS for Semiconductor Design, Verification, and Fabrication

It’s no secret that speed-to-market is driving a new competitive reality in the semiconductor industry. For both chip manufacturing and design, the pressure to stay ahead of the curve is forcing semiconductor companies to gain greater control of their supply chains, react to changing business realities, and deliver high quality products to customers at a faster rate. The cloud allows semiconductor firms—whether they are fabless, integrated, design services, or IP providers—to accelerate innovation, take more risks, run more experiments, and get the next advanced silicon product into the foundry pipeline faster.

In this new video, we share how Amazon Web Services (AWS) can help companies like Arm increase the pace of innovation, get chips to market faster, and enable rapid chip development, including Electronic Design Automation. Hear from Tim Thornton, Director of Arm-Based Engineering at ARM Limited, and learn how they are working to harness the benefits that cloud computing, including the new Arm-based Graviton2 instances, can offer to silicon design teams in this highly competitive environment.

Arm on Arm on AWS – Digital Verification with Graviton2

Thornton shares how the company is using Arm-based instances on AWS to develop and characterize Arm processor intellectual property (IP) on AWS—or in other words, “Arm on Arm on AWS”—demonstrating how the adoption of this solution was driven by the need for greater flexibility. As Thornton says, “A large amount of compute can be required at a fairly short notice in order to satisfy our customer demand. That means that integrating the flexibility of cloud computing into our current infrastructure is an obvious way of allowing us to deal with such a dynamic workload more efficiently”.

The experiences at Arm represent the benefits that other semiconductor firms are seeing as they migrate EDA and related workloads to cloud. Customers tell us that a major motivation for cloud adoption in semiconductor design and verification is to better optimize and utilize the most valuable resources these companies have: their silicon designers and verification engineers. Customers also tell us that secure collaboration is important, including protecting critical IP when working with external partners such as EDA vendors, IP providers, and foundries. The cloud has proven to be the right solution for secure collaboration, as highlighted in this video.

Read Arm's latest blog on this story – Increasing the Pace of Innovation with Arm and AWS for Silicon Design using EDA software

Why AWS Cloud?

AWS offers a secure, agile, and scalable platform with a comprehensive set of services and solutions for high performance design, verification, and smart manufacturing, supporting EDA and rapid semiconductor innovation in the cloud. Semiconductor companies, including fabless and integrated device manufacturers, and their IP and foundry partners can benefit from the massive scale of AWS infrastructure to design next gen connected products.

- Innovate faster – Prototype, design, and verify complex systems-on-chip, using scalable cloud resources for Electronic Design Automation (EDA).
- Collaborate better – Work seamlessly and securely with third-party partners including IP providers, EDA software vendors, and manufacturing service partners (foundries, OSATS, contract, and original device manufacturers).
- Reduce risk – Advanced Silicon and system verification is becoming increasingly complex. Mistakes can cost millions if not billions of dollars for larger companies.
- Reduce cost – Stop wasting CAPEX on IT and stop wasting valuable engineering time.

Want to learn more? Go to AWS for Semiconductor and Electronics to explore our solutions, download resources, and read customer case studies.

Read Blog post online

Related Videos

AWS for Semiconductor & Electronics

Amazon Web Services (AWS) offers a secure, agile, and scalable platform with a comprehensive set of services and solutions for high performance design, verification, and smart manufacturing, supporting electronic design automation (EDA) and rapid semiconductor innovation in the cloud. Semiconductor companies, including fabless and integrated device manufacturers, and their IP and foundry partners can benefit from the massive scale of AWS infrastructure to design next gen connected products.

Watch Video online

Scaling Semiconductor Design Workflows on AWS

Watch Video online

AWS re:Invent 2020: Semiconductor Design on AWS with Qualcomm

Semiconductor production relies on resources and partners far beyond what is needed for design and verification. This includes working with tool vendors, IP providers, foundries, contract manufacturing, and consumers. Organizations are running advanced node workflows and leveraging the breadth of services on AWS to enable analytics, AI/ML, and yield management. In this session, Qualcomm discusses running its flows on AWS. Also, learn how AWS is working across the entire semiconductor industry to enable innovation and scale workflows.

Watch Video online

AWS Floor28 News – Semiconductor Special Edition

Watch the Semiconductor special edition of AWS Floor28 News!

Watch Video online