APPS
& INFRA

Microsoft Azure Solutions Business

Running Optical Proximity Correction (OPC) in the cloud

Andy Chan, Sept 2020

Agenda

- Industry perspectives
- Cloud for semiconductor, a recap
- Cloud for semiconductor, use cases
- OPC workload considerations
- Hosted inside of Azure, 1st party support
- QoS eliminates noisy neighbors and maintains swim lanes

Our approach for semiconductors

EDA

Pre-Silicon

Foundry

OEM

Design

- Silicon development & optimization
- Process Efficiencies

Intelligent Supply Chain

- Machine Calibration
- Predictive Maintenance
- Sustainability

Manufacturing

- Process Optimization
- Factory Assistant
- Cognitive Quality
- Supply Chain Visibility & Orchestration
- Track and Trace
- Sourcing, Planning, Demand Forecasting

Sales and Marketing

- Intelligent Customer Engagement
- Partner & Employee Collaboration
- Relationship Selling

Security

Microsoft 365 | Teams | Graph | Security & Compliance

IoT and Edge | Data | AI | Blockchain

Supply Chain Management | Mixed Reality | Power BI | PowerApps

Cloud for semiconductor recap

2019

Curious

- Technology readiness established
- Security no concern
- EDA licensing models changing
- First chip 100% in the cloud
- Identified Impediments
- Winning foundry endorsement

2020

Serious

- Most customers in PoC or production
- Fixed capacity migrations
- New licensing models in the work
- Optimized tooling an expectation

Cloud use case examples

Cloud bursting example

- Generally a single workload (OPC, SPICE, DRC, IR Drop, STA)
- Access to more compute resources when the demand for computing capacity spikes
- Free up on-prem resources by offloading batch friendly workloads to the cloud
- Tested on real world applications vs. synthetic benchmarks
- Single job performance vs 1000 parallel jobs throughput considerations
- Shutdown when done

Startup (born in the cloud) example

- 5/7nm SoC, full tool chain, mixed flow
- No IT staff, small/no existing DC
- From Azure subscription setup to RTL sim pipe cleaning in 2 weeks is doable
- Download foundry's PDK via Azure vNet peering
- Tape out to foundry via Azure vNet peering

Cloud use case examples

Regional data center replacement example

- Data center replacement for a regional India site
- Each engineer assigned 2 to 4 cores for layout work in Azure
- Each engineer on-prem: laptop/light desktop
- Normal office hours usage
- Cross-region data transfer via a single Azure backbone back to Corp DC
- Geo collaboration via Exceed On Demand remote desktop

QA and stress test example

- Greater access to cores and resources to tools providers:
  - QA
  - Stress test
  - troubleshoot (replicating customer's environment)
  - Rapid access to latest gen of hardware
- End users use Azure resources to:
  - Patch management and comparability testing
  - New CPU architecture testing
- New project /product "prototyping"

Pulse of the Industry

press

cadence

Cadence Collaborates With TSMC and Microsoft to Reduce Semiconductor Design Timing Signoff Schedules With the Cloud

Cadence Extends Cloud Leadership With New CloudBurst Platform for Hybrid Cloud Environments

SYNOPSYS®

Synopsys, TSMC and Microsoft Azure Deliver Highly Scalable Timing Signoff Flow in the Cloud

SiFive Selects Synopsys Fusion Design Platform and Verification Continuum Platform to Enable Rapid SoC Design

Mentor®

Mentor's analog/RF/mixed-signal verification tools scale to 10,000 cores on Microsoft Azure

Mentor and AMD verify massive Radeon Instinct Vega20 IC design on AMD EPYC in ~10 hours with ecosystem partners Microsoft Azure and TSMC

tsmc

Microsoft and TSMC announce Joint Innovation Lab to accelerate silicon design on Azure

TSMC Leads the Industry by Hosting the First "TSMC IC Layout Contest" in the Cloud

blogs

Chris Lattner, SiFive

Cloud Accelerated Idea To Silicon

Daniel Payne, SemiWiki

Mentor Adds Circuit Simulators to the Cloud using Azure

Simon Sharwood, The Register

Microsoft cooking Azure instance types just for chip designers

Chad Morgenstern, NetApp

Chip Design and the Azure Cloud: An Azure NetApp Files Story

Omar el-Sewefy, Tech Design Forum,

How cloud computing is now delivering efficiencies for IC design

Semiconductor workloads in Azure, public references

Reference links:

Cadence Collaborates With TSMC and Microsoft to Reduce Semiconductor Design Timing Signoff Schedules With the Cloud. Cadence press release. June 15, 2020.

Synopsys, TSMC and Microsoft Azure Deliver Highly Scalable Timing Signoff Flow in the Cloud. Synopsys press release. June 15, 2020.

New general purpose and memory-optimized Azure Virtual Machines with Intel now available. Microsoft product release, with silicon customer endorsement, June 15, 2020

SiFive Selects Synopsys Fusion Design Platform and Verification Continuum Platform to Enable Rapid SoC Design. Synopsys press release. March 25, 2020.

Cloud Accelerated Idea To Silicon. Blog by Chris Lattner, SiFive, March 25, 2020

Chip Design and the Azure Cloud: An Azure NetApp Files Story. Blog by Chad Morgenstern, NetApp. Mar 17, 2020.

TSMC Leads the Industry by Hosting the First "TSMC IC Layout Contest" in the Cloud. TSMC press release. March 12, 2020.

Mentor Adds Circuit Simulators to the Cloud using Azure. Daniel Payne. SemiWiki.com. November 8, 2019.

Mentor's analog/RF/mixed-signal verification tools scale to 10,000 cores on Microsoft Azure. Mentor press release. October 29, 2019.

TSMC in the Cloud Update #56thDAC 2019. Daniel Nenni. SemiWiki.com. June 13, 2019.

Mentor and AMD verify massive Radeon Instinct Vega20 IC design on AMD EPYC in ~10 hours with ecosystem partners Microsoft Azure and TSMC. Mentor press release. May 30, 2019.

How cloud computing is now delivering efficiencies for IC design. Omar El-Sewefy. Tech Design Forum. May 26, 2019.

TSMC Strengthens OIP Cloud Alliance with New Partner and New Solution Enablement. TSMC press release. April 26, 2019.

Cadence Extends Cloud Leadership With New CloudBurst Platform for Hybrid Cloud Environments. Cadence Design Systems press release. April 4, 2019.

General OPC considerations

Foundry

Manufacturing

- Mask designs are becoming ever increasingly complex with the advancement of Moore's Law
- Every new process node requires more and more aggressive OPC and RET
- Every geometrical device structure on the wafer requires numerous physical layout changes necessary to comprehend optical diffractions and process-chemistry
- Transistor counts on modern devices reaching tens of billions
- Increasing number of process layers that drive up the size of the mask set
- Near exponential increase in complexity is requiring an immense amount of calculation

OPC Infrastructure considerations

Foundry

Manufacturing

- Huge compute farm, north of 40,000 cores
- Heavy shared storage requirements, Unix/Linux NFS centric
- Heavy network bandwidth required, north of 10Gig preferred
- Scalable infrastructure to keep up with compute demands
- A job scheduler for workload management is needed
- Data mobility (before compute, during compute and post compute)

Infrastructure considerations: I/O pattern for optimization

Scheduler Storage (LSF, PBS)

Shared, Small files,

Many small transactions

Tools/Libraries Storage

Large number of files,

Scale via Caching

Heavy read, shared or local

Scratch Space for intermediate writes

Temp space sensitive to write latencies, potentially used multiple times

Merge Storage

Write sensitive

Shared Storage

Source Data

Size and quantity vary depending on the design and workflow

Results

Requirements vary per vertical, heavy write

Archive Tier

A Cheap, Deep Lake of Data

Infrastructure considerations, what cloud offers

High-Performance

High throughput, low latency with scale-out performance

Scale out

To match performance demands, on an as-needed basis

Flexible

Data movement between tiers

Simple

Full managed, enterprise-grade, easy deployment

1st party experience

Cost

Economics at scale

Addition storage considerations for OPC

Performance & Scale

File sizes and quantities	Larger # of small files == greater Ops requirement Throughput vs IOPS Frontend vs Backend
Concurrent access	(# of clients, rate of requests, avoid hotspots)
Latency	Effects on initial load, writes, metadata (e.g., Large # small files + WAN), inode creation speed
Job run time	The whole job, including storage component

Access

- Local (Disk) /tmp
- Shared NAS
- Blob
- Fault Tolerance
- Outside sources

Protocol

- NFS/SMB
- Rest (BLOB)
- Parallel FS

  (GPFS, Lustre, BeeGFS)
- HDFS
- Multi-protocol

Costs

- Right tiering
- Complexity

APPS
& INFRA

Microsoft Azure Solutions Business

Cloud Adoption Steps

Cloud adoption steps

	Do you homework ahead of deployment $$ (IP, licensing, workflow and data hygiene)
1	A dedicated project team Cloud provider / Cloud systems integrator EDA tools provider Engineering IT + IT IT
2	A well-written timetable with step-by-step milestone
3	Defined role and responsibilities
4	“Lift and shift” first
5	Optimization second
	Cloud for OPC is not as simple as other workloads to move to the cloud

Cloud adoption steps

Role	Responsibilities
Overall project lead	High level coordination point
Commercial planner	Pricing quoting Discount requesting Funding considerations
Program project management	Coordinate regular status meetings Designated note taker and communicator on regular status meetings. Keep track of high-level milestones and assignment Manager of documents, (PPT, project task list, etc) centralized in a share point Working-level coordination
Technical project management	Tech communicator: translate technical requirements into Cloud specific solutions Document tech blockers: escalate tech blockers to ISV and/or Cloud provider Coordinate engineering resources from Corp

Cloud adoption steps

Role	Responsibilities
Architect	Overall EDA technical architecture Complete alignment with Technical Project Management and deployment team
Cloud Infra Setup and deployment	Basic IT day-to-day operations, deployment management
Cloud OPC production support	Engineering IT/engineers that are running the OPC themselves

APPS
& INFRA

Microsoft Azure Solutions Business

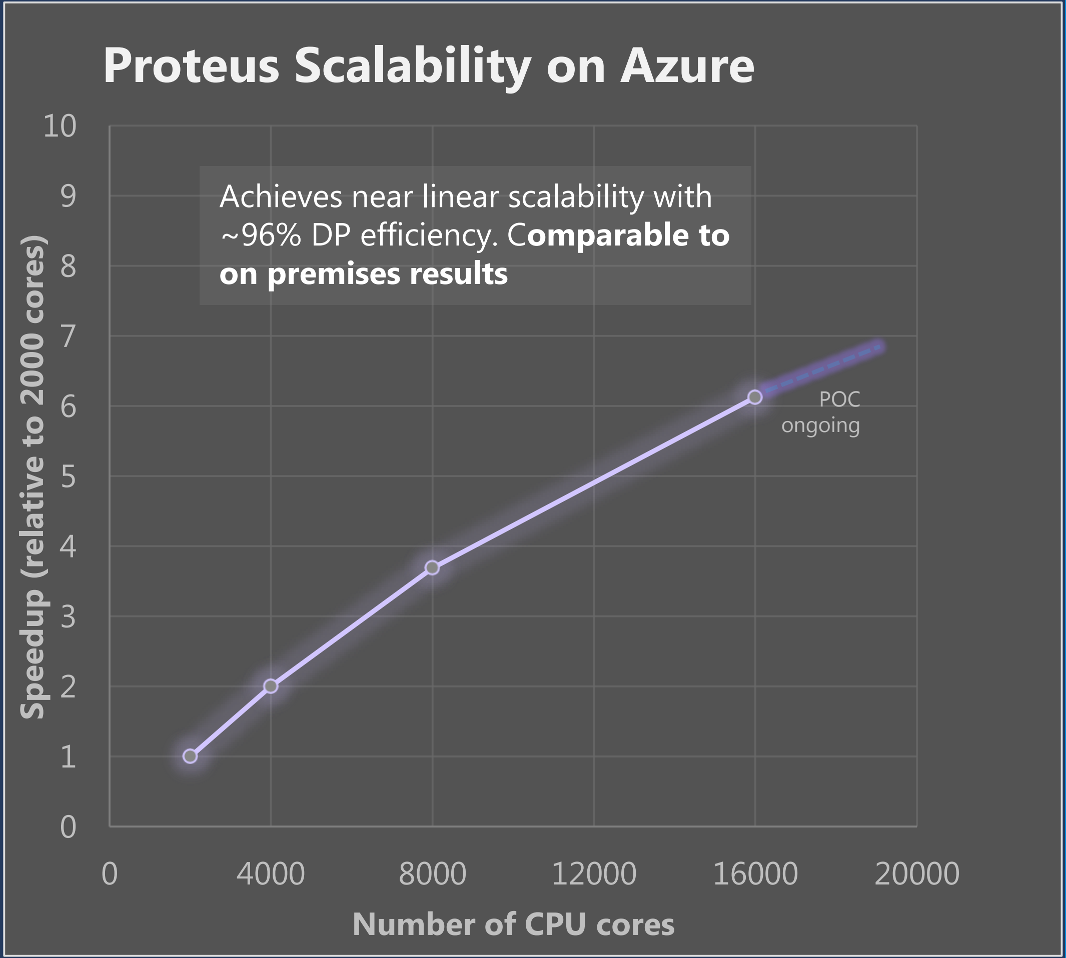
Synopsys Proteus OPC Scales On Azure Example

Synopsys® Proteus OPC Scales On Azure

- Using cloud scaling can reduce OPC time to market from days to hours by accessing highly-scalable instances on-demand

- Improve on-premises hardware utilization and ROI by sizing farms for nominal usage, and leveraging cloud scaling for peak-demand

- Azure's cloud infrastructures provides access the latest compute and storage options to improve through-put and guide on-premises purchasing decisions



ESv4 Intel® Cascade Lake

Azure NFS solution used

High-Performance

Scale out

Flexible

Simple

Azure NetApp Files

Azure NetApp Files (ANF)

Key advantages

- Capacity and performance are dynamically adjustable, readily burstable
- Bare metal NetApp FAS experience
- Excels at IOPS, including small/mixed IO
- Inode scaling, performance scales linearly as volume size grows by TB
- Hosted inside of Azure, 1st party support
- QoS eliminates noisy neighbors and maintains swim lanes

Azure NetApp Files service levels

Service level	Standard	Premium	Ultra
Throupgut	16 MB/s/TB	64 MB/s/TB	128 MB/s/TB
Latency	3 – 6 ms	1-3ms	<1ms
Performance	Standard Performance - 1000 IOPS per TB (16k IO) and 16MB/s/TB	Premium Performance - 4000 IOPS per TB (16k IO) and 64MB/s/TB	Extreme Performance - 8000 IOPs per TB (16k IO) and 128MB/s/TB

VMSS (Virtual Machine Scale Sets)

Offers
The ability to create and manage a group of load balanced VMs
Automatically increase or decrease in response to demand or a defined schedule.
Centrally manage, configure, and update a large number of VMs in a single group or groups
Single event and tasks management
Deploy VMs belonging to different VM series within the same VMSS
Faster provisioning of custom images at scale using Shared Image Gallery

Optimized Azure compute for EDA

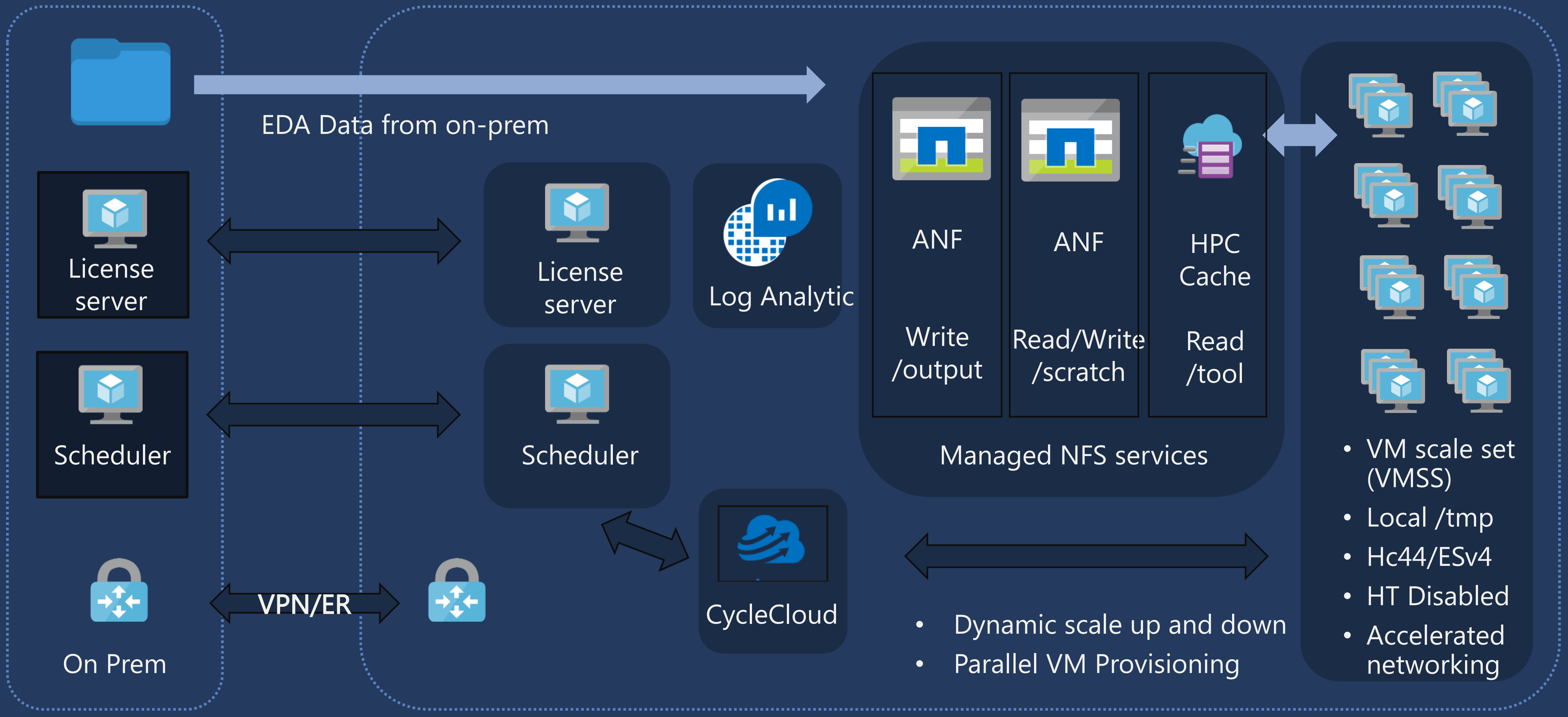
	ESv4 (intel Cascade Lake)	Easv4 (AMD EPYC 7452)	Mv2 (Intel SL Platinum)	Hc44 (Intel SL Platinum)
Status	Public Preview	GA	GA	GA
Clock speed	3.4Ghz turbo	3.35GHz boost	3.8 GHz turbo	3.7 GHz turbo
Max vCore count	64	96	208	44
Max physical core count (Hyperthreading disabled)	32	48	104	No HT
Max memory	432 Gig	672 Gig	5700 Gig	352 Gig
Core to memory	8 Gig per core	8 Gig per core	27.4 Gig per core	8 Gig per core
Core to memory (Hyperthreading disabled)	16 Gig per core	16 Gig per core	54.8 Gig per core	No HT

APPS
& INFRA

Microsoft Azure Solutions Business

Reference architect

A full-production cluster for OPC



APPS
& INFRA

Microsoft Azure Solutions Business

Final remark

Final Remark

Cloud is ready for most intensive OPC flows

OPC can benefit “cloud native” solutions

- Containers
- Architecture of OPC is very scalable, might benefit from Serverless compute
- Spot instance

Cloud for semiconductor, use cases

Adoption AI/ML frameworks