



Developing Real World Applications

Using P4 based Architecture

Krishna Doddapaneni

AMD 
together we advance_

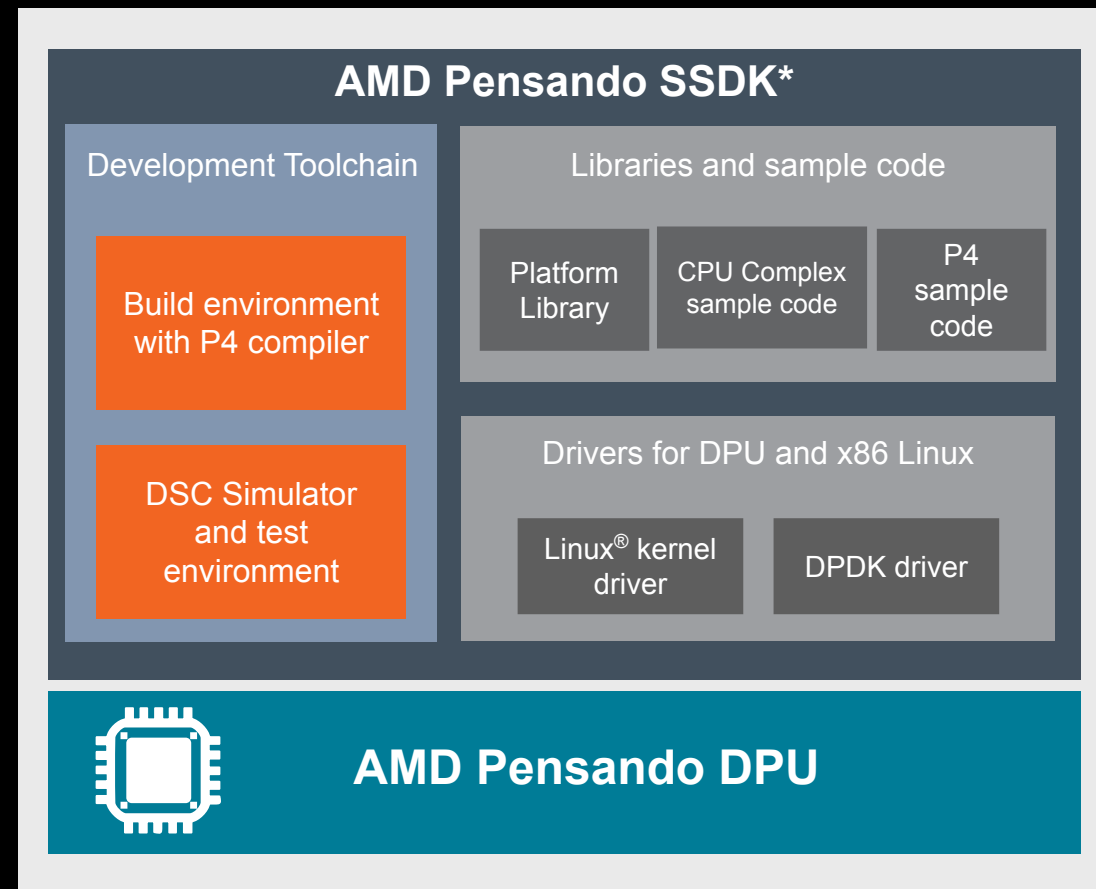
AMD Pensando DPUs and Projects

- Shipping 1st generation DPU in 2019, 2nd generation shipped 2021
- Full stack solution, native P4 hardware with a full P4 centric software
- Many years of DPUs in production at Cloud and Enterprise customers
- P4 Applications in Production
 - Multiple SDN Stacks in Clouds
 - Enterprise DPU Distributed Firewall
 - Storage Target Offload
 - Storage NVMeoF Initiator
 - SDN Disaggregation
 - VPN/NAT/Cloud GW
- Deployed in Various Physical Form Factors
 - PCIe card in server
 - Smartswitch
 - SDN Network Accelerator Appliance
 - Storage Target



A Fully Programmable System Enables Technology Evolution

- P4 SDN Data Path Programmability
 - Parser/De-Parser/Match-Action
- P4 Host Programmability
 - Host Interfaces
 - DPDK, CNIC, RDMA (Queue Pairs)
 - Device Emulation
 - VMXNet3, NVMe Virtualization and NVMe/TCP
 - Stateful Transport Applications
- P4 & simulator toolchain
- Rest of the software stack
 - Software Data Path
 - Control Plane
 - Management Plane



*Software in Silicon Development Kit

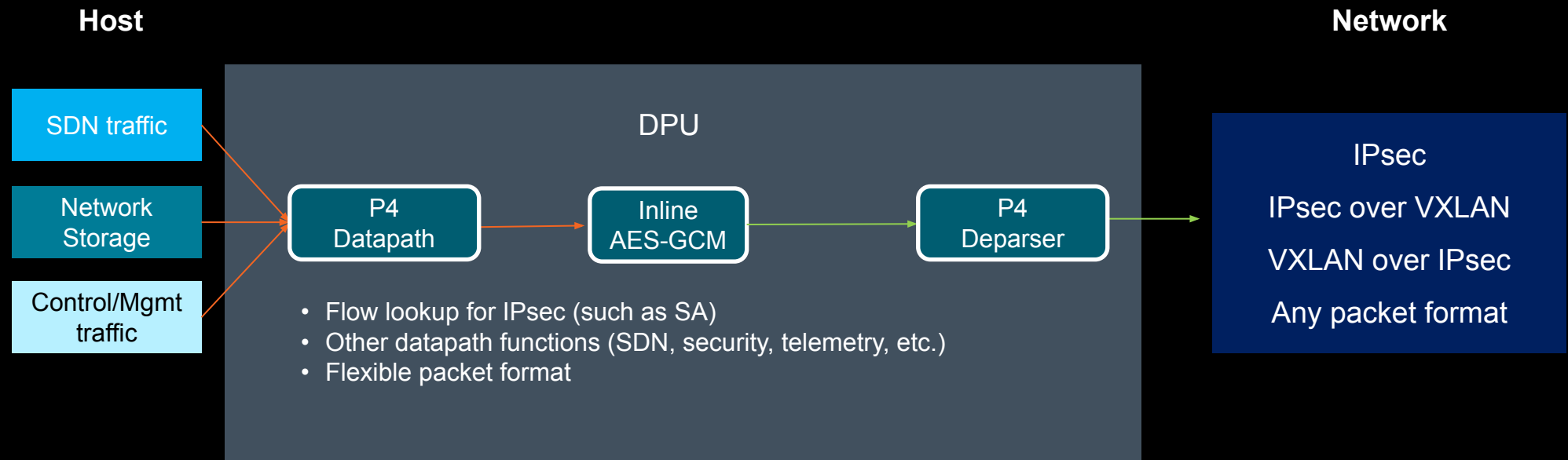
What does P4 architecture give you ?

- Flexibility
 - No Public cloud SDN is the same
 - No Public cloud network encryption format is the same
 - Brownfield Insertion into Public Clouds
 - Tune the P4 resources for the customer specific use case to get the best performance
- Agility
 - Weekly releases on additional requirements
 - Incremental feature deployment
 - Solve requirements that you and your customers don't know
 - Compatibility to future generation P4 ASIC's
- Better operational experience
 - Customization for telemetry and debuggability metrics

P4 Beyond Packet Transformation: Examples

- IPsec inline encryption with low latency @ line rate
- Stateful services acceleration; P4 assisted policy evaluation
- HA with redundant DPUs using P4 for real-time inline flow replication
- Seamless upgrades in production
- PCIe Device Emulation: presents devices to host and/or DPU CPU complex
- Custom telemetry – IPFIX with custom TLV's, Flow Logs, etc..
- Service chaining for Storage Offloads
- HW assisted dirty page tracking for Live Migration

Security Encryption Pipeline – Inline Programmable Encryption



- Line rate throughput with low latency
- Encryption key in any granularity
- Programmable encryption at different layers: overlay vs underlay
- IPsec control plane-agnostic

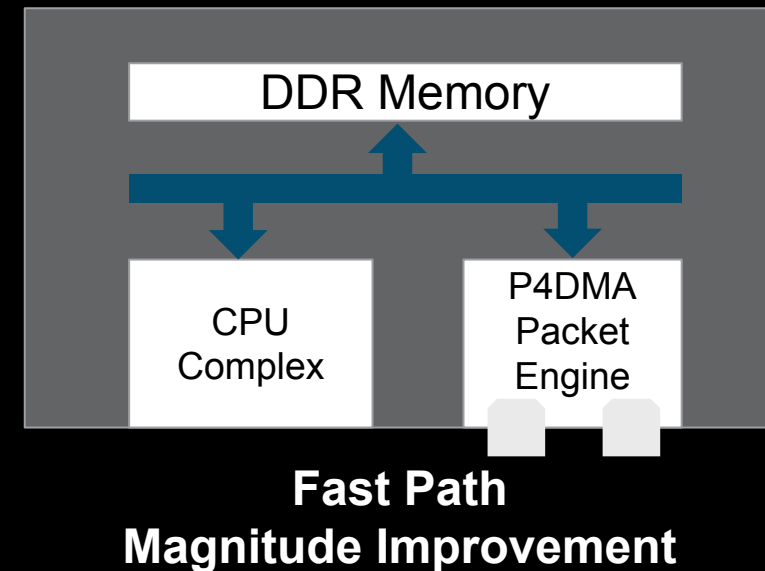
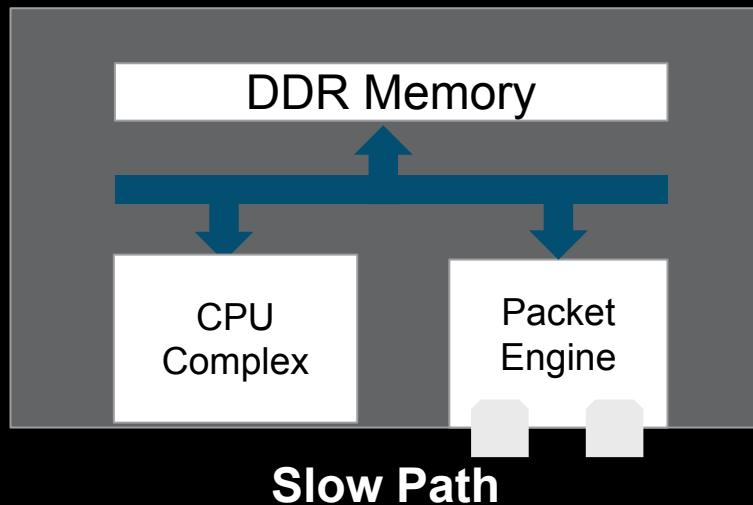
Accelerate CPS with P4

DPU CPU connection setup/tracking

1. New flow packet received
2. DPU CPU checks policy
3. DPU CPU performs table lookup
 - Next hop, metering, telemetry, NAT, LB etc.
4. DPU CPU create flow entry
5. Packet Forwarded
6. Flow timeout checked periodically by Arm[®] CPU

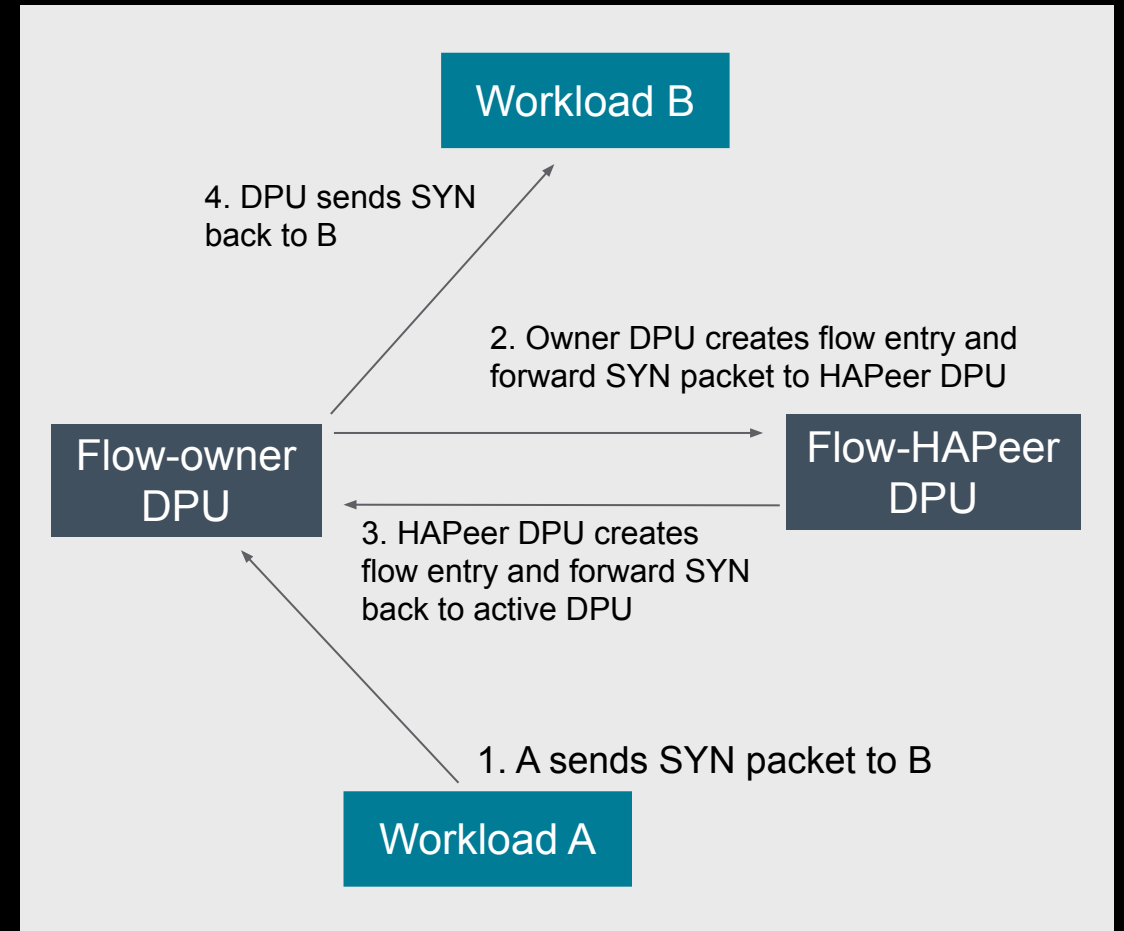
DPU P4 connection setup/tracking

1. New flow packet received
2. P4 performs table lookup
 - Next hop, metering, telemetry, NAT, LB etc..
3. DPU CPU creates flow entry
4. Packet Forwarded
5. Flow timeout checked periodically by P4



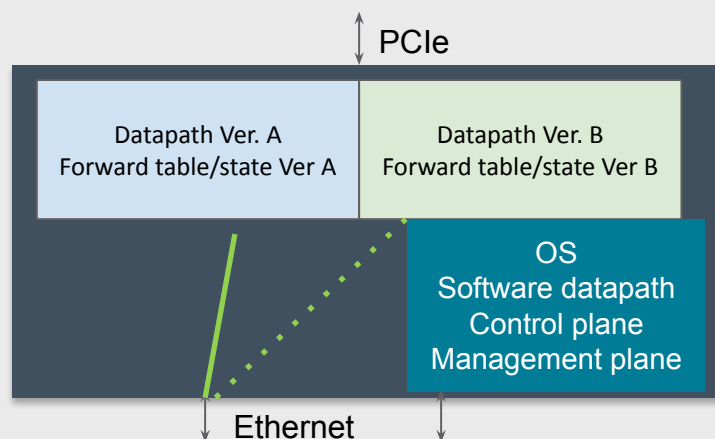
HA – Real-Time Inline Flow Replication

- Replication with P4
 - Flow is set up/cleaned up in the logical DPU before forwarding packets to endpoint
 - Achieve high CPS and high session scale with minimum overhead to DPU CPU complex
 - Minimum-to-no packet loss during failover
- Soft-Data Plane (Exception Plane) for bulk flow sync
 - Bulk sync when DPU comes back online



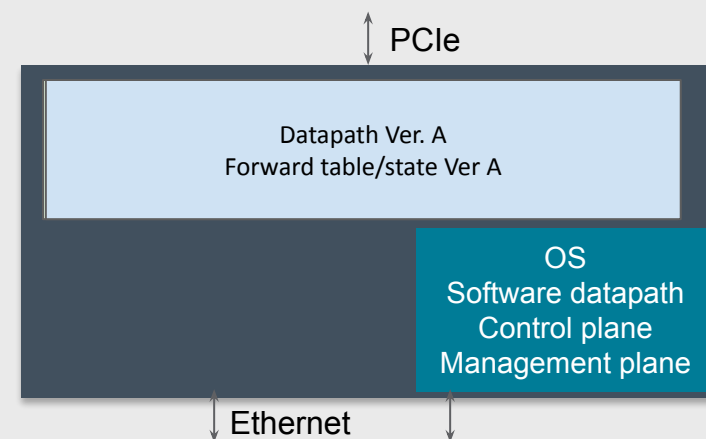
Seamless Upgrades in Production without Host Reboot

Hitless Software Upgrade



- All links remain up
- Minimum to no traffic impact
- Build new P4 program in partition B
- Build table in partition B and sync flow state from A
- Continue data forwarding with A until B is fully built
- Need extra memory for Partition B

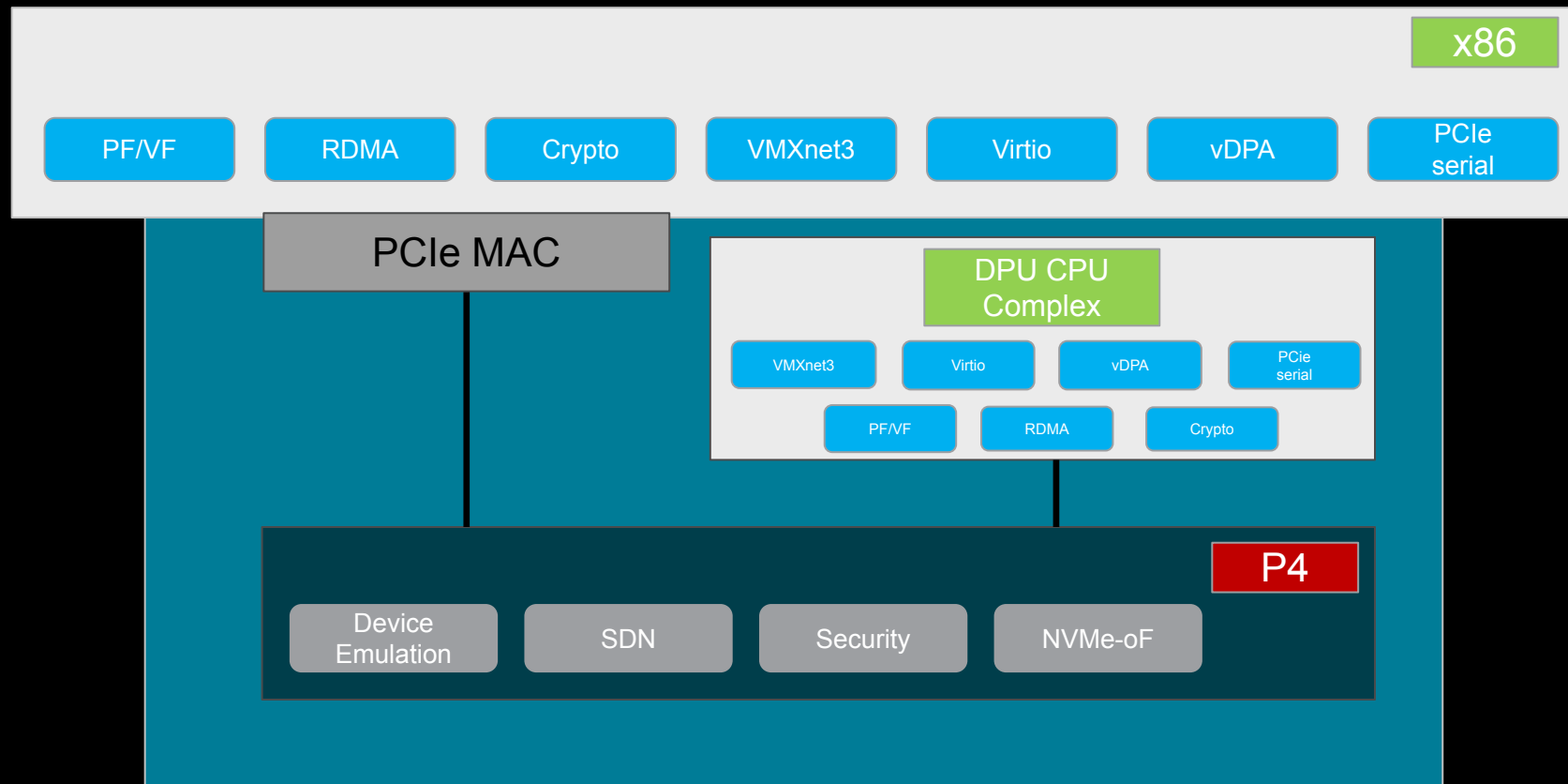
Graceful Software Upgrade



- PCIe link remains up during upgrade
- Upgrade datapath program
- Rebuild forwarding table and relearn session state (if needed)
- No restriction on kernel upgrade
- No extra memory reserved for upgrade

DPU Device Emulation in P4

- Supports different type of devices and drivers via P4 programmable PCIe and host interfaces
- All devices can be presented on PCIe to the host or as a platform device to the CPU complex of DPU
- Device emulation plus other services



Summary / Future

- Fully programmable DPU with native P4 architecture provides
 - Flexibility and agility in developing/deploying SDN/SDS/Security Services
 - Power in the hands of Software Developers and Research community to Innovate on new ideas
- Future
 - Explore new transport layers using P4 ?
 - Multi-pathing, LB etc.
 - More standardization ?
 - SDN Disaggregation - DASH, OPI
 - How to make it easy for new users to use P4 arch based systems ?
 - And make it interoperable ?

Questions?

Disclaimer

The information presented in this document is for informational purposes only and may contain technical inaccuracies, omissions, and typographical errors. The information contained herein is subject to change and may be rendered inaccurate for many reasons, including but not limited to product and roadmap changes, component and motherboard version changes, new model and/or product releases, product differences between differing manufacturers, software changes, BIOS flashes, firmware upgrades, or the like. Any computer system has risks of security vulnerabilities that cannot be completely prevented or mitigated. AMD assumes no obligation to update or otherwise correct or revise this information. However, AMD reserves the right to revise this information and to make changes from time to time to the content hereof without obligation of AMD to notify any person of such revisions or changes.

THIS INFORMATION IS PROVIDED 'AS IS.' AMD MAKES NO REPRESENTATIONS OR WARRANTIES WITH RESPECT TO THE CONTENTS HEREOF AND ASSUMES NO RESPONSIBILITY FOR ANY INACCURACIES, ERRORS, OR OMISSIONS THAT MAY APPEAR IN THIS INFORMATION. AMD SPECIFICALLY DISCLAIMS ANY IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY, OR FITNESS FOR ANY PARTICULAR PURPOSE. IN NO EVENT WILL AMD BE LIABLE TO ANY PERSON FOR ANY RELIANCE, DIRECT, INDIRECT, SPECIAL, OR OTHER CONSEQUENTIAL DAMAGES ARISING FROM THE USE OF ANY INFORMATION CONTAINED HEREIN, EVEN IF AMD IS EXPRESSLY ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

AMD, the AMD Arrow logo, Pensando and combinations thereof are trademarks of Advanced Micro Devices, Inc. ARM[®] trademark(s) is/are the registered trademark(s) of ARM Limited in the EU and other countries. PCI Express, PCIe[®] are trademarks of PCI-SIG Corporation. Linux[®] is a trademark of Linus Torvalds. Other product names used in this publication are for identification purposes only and may be trademarks of their respective companies.

AMD 