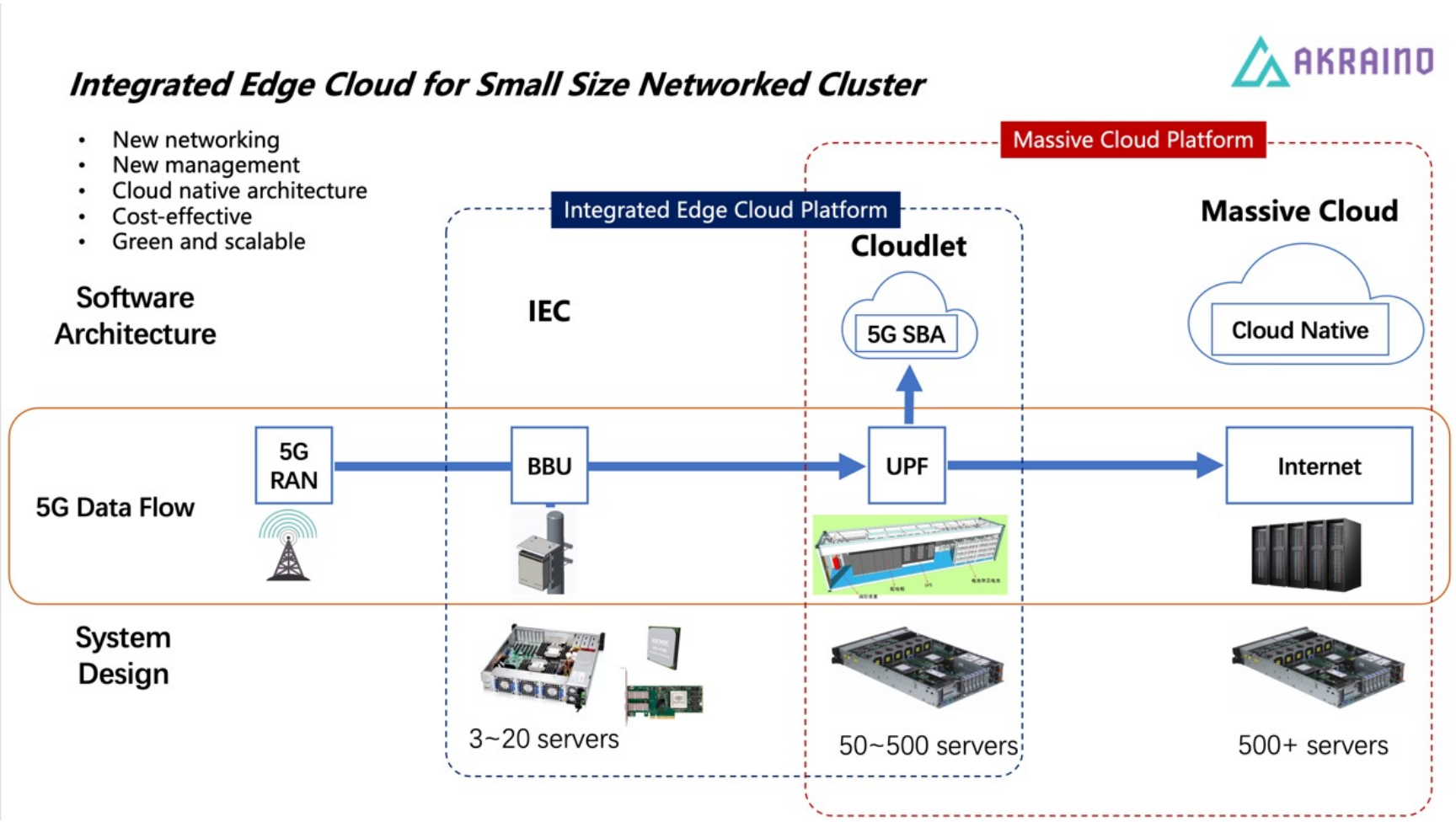
Akraino Release 6 Review of Integrated Edge Cloud Type 5

LEO
li@socnoc.ai

2022.5.19

[Release6] Integrated Edge Cloud Overall Introduction

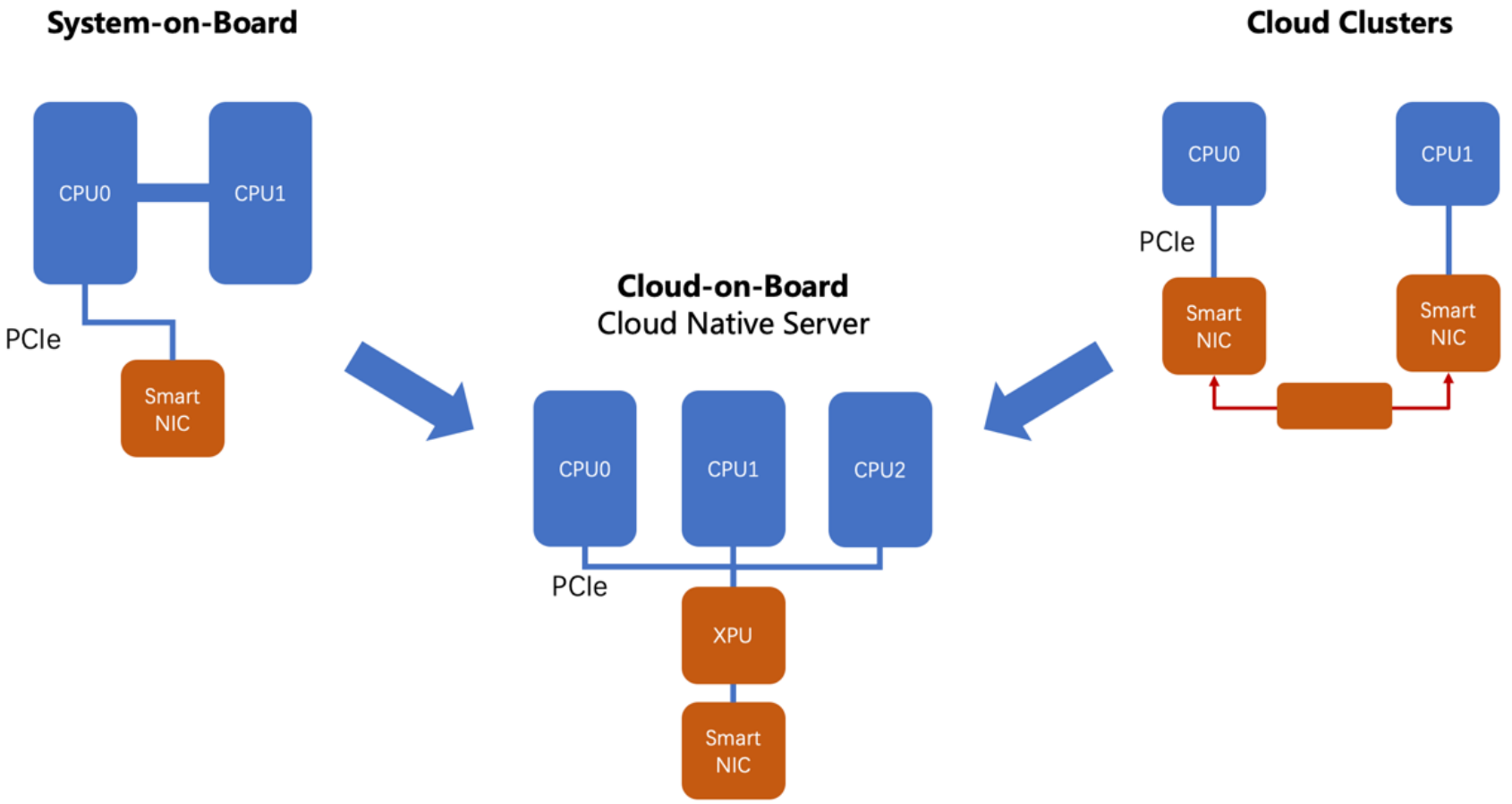


In R6, we introduce an innovative networking architecture based on PCIe data fabric to lower both the cost (CAPEX) and power consumption (OPEX) in small clusters for edge cloud computing. Based-on innovative data processor (DPU and XPU), the next-generation networking features with:

- New networking architecture to lower the TCO of edge infrastructure
- TCP/IP compatible and cloud native for develops and developers
- Green to protect the environment for lasting development
- Scalable and composable to meet the dynamical workload

Goal: An innovative architecture for small-size edge cloud using data processor

Vision: Merging SoB and Clusters into CoB

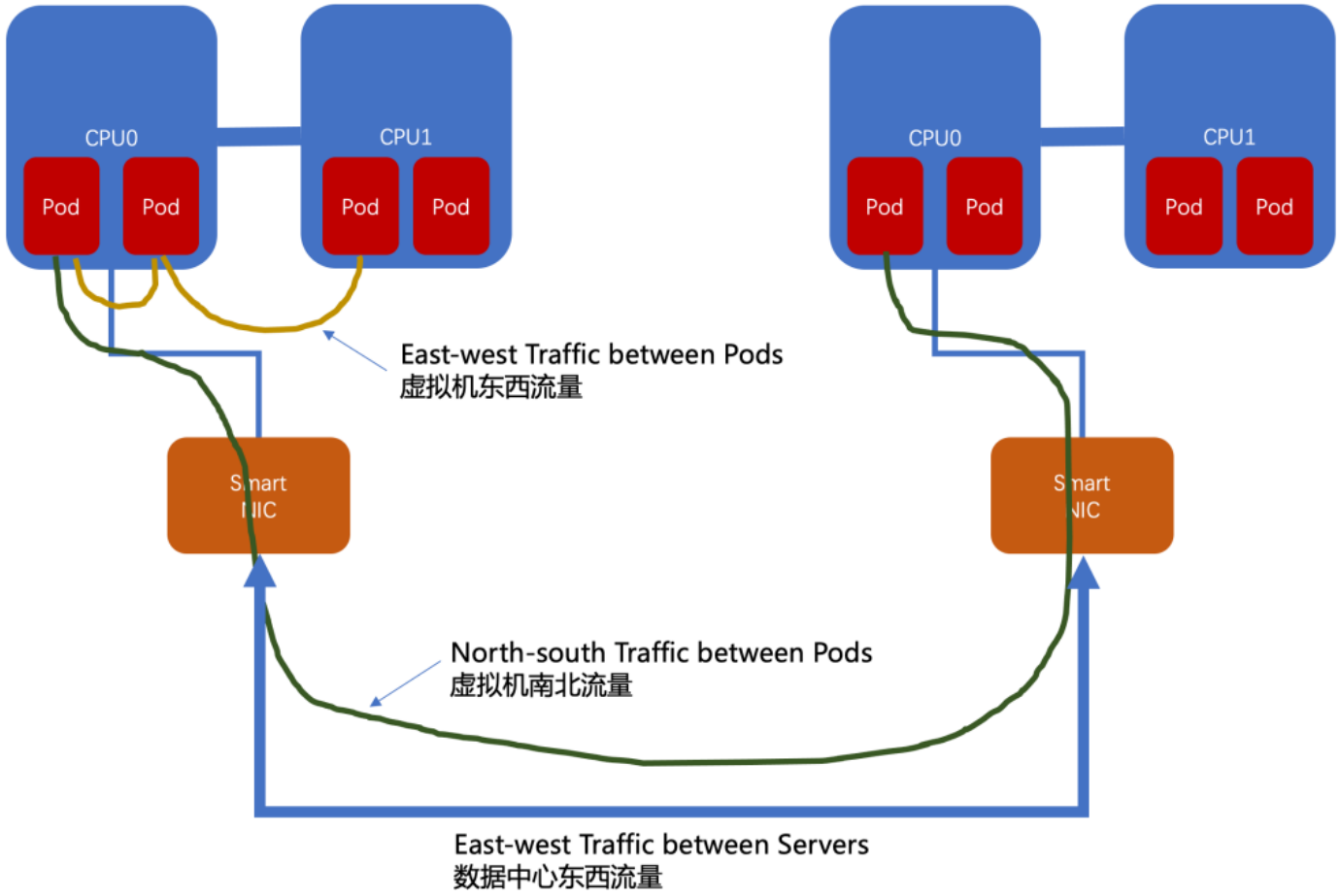


By the advantage of PCIe networking, we can unified the system-on-board (SoB) connection and the cloud cluster topologies into one single and simple architecture, which we named as Cloud-on-Board (CoB) Architecture.

- Less moving parts, playable
- Low power consumption
- Low cost
- High bandwidth
- High performance

Goal: An innovative architecture for small-size edge cloud using data processor

Paradigm Shift: Intra-core to inter-pod communication



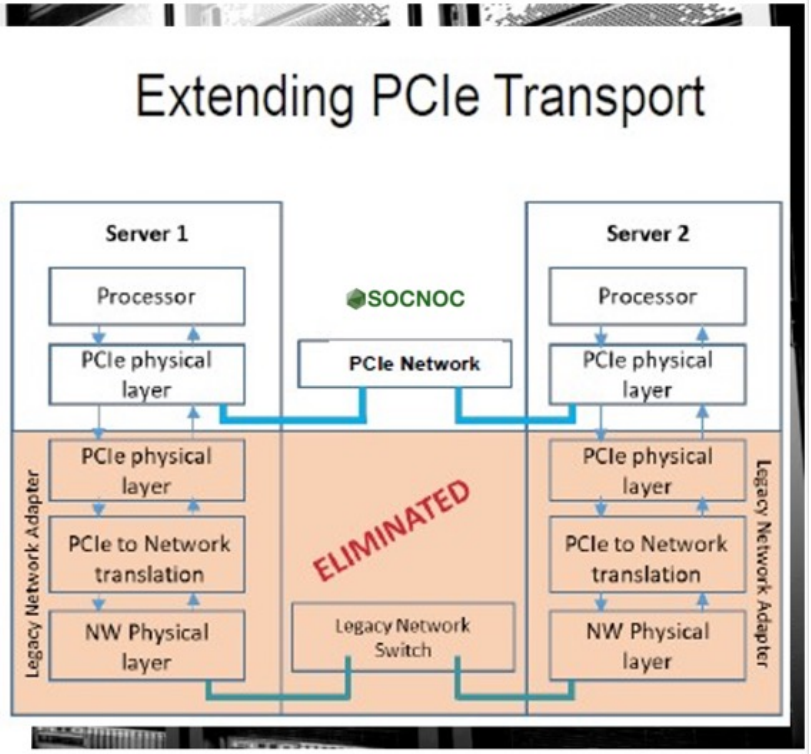
Goal: An innovative architecture for small-size edge cloud using data processor

Technology: Extending PCIe Transport to PCIe Net

SOCNOC

Architecture Advantage: Less is More

https://www.nextplatform.com/2019/10/02/a-new-twist-on-pci-express-switching-for-the-datacenter/



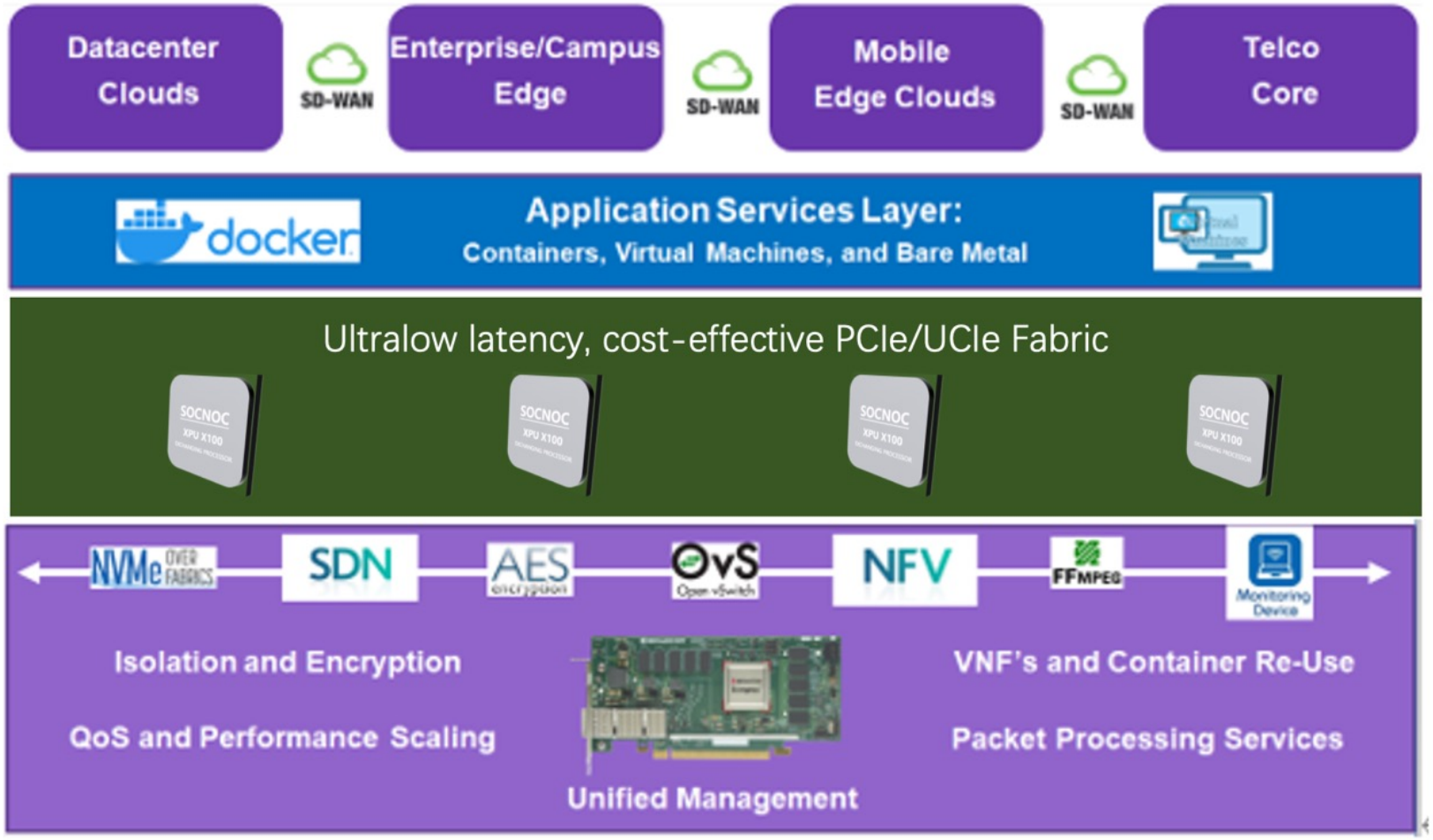
AKRAINO | WeOpen | NextArch FOUNDATION

Extending PCIe Transport to Net

- NTB based, universal
- No additional adapters needed
- Rich roadmap to PCIe 7.0
- Extendable with CXL/UCIe

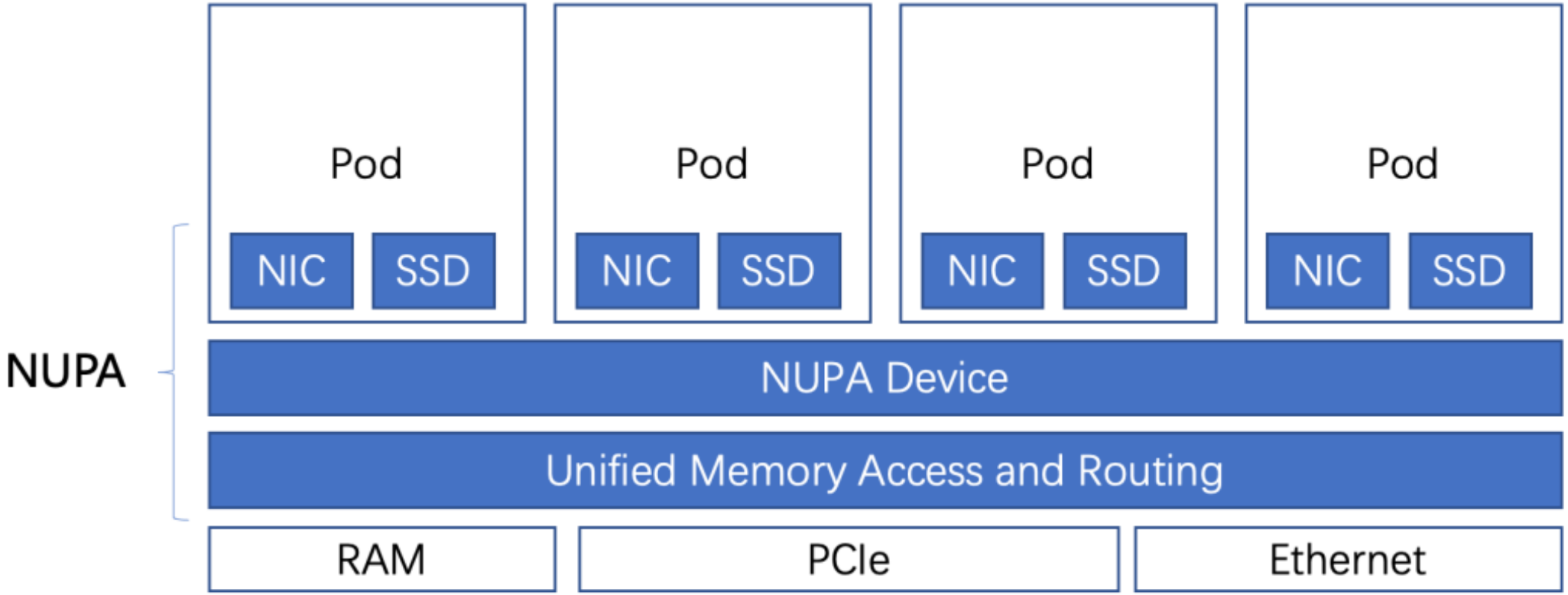
Goal: An innovative architecture for small-size edge cloud using data processor

Architecture: PCIe-fabric based Cloud Native System



Since DPU is PCIe-compatible device, we can further combine DPU and PCIe Networking together. In R6, we introduce a hardware layer or physical link/fabric layer between the DPU and the CPUs as below. With this layer, we extend the DPU cluster size and also use the DPU management features as well.

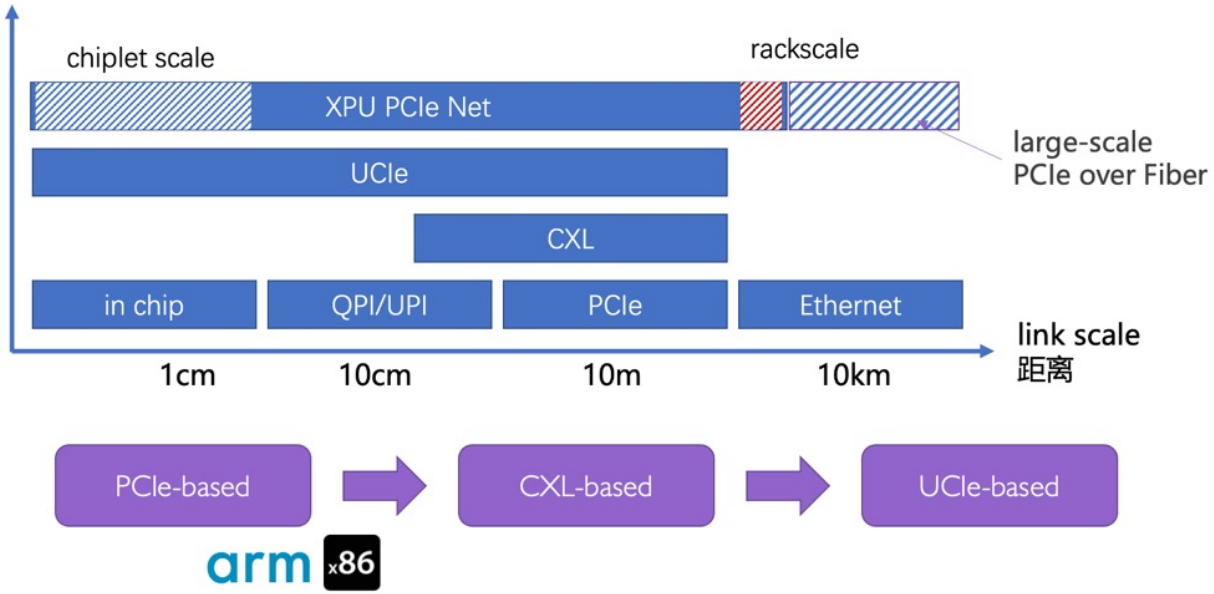
NUPA: Unified Networking Protocol Architecture



Roadmap: from PCIe to UCIe

SOCNOC

Roadmap and Ecosystem



AKRAINO | WeOpen | NextArch FOUNDATION

PCI Specification	Spec Released	Data Rate	Encoding	Frequency	Maximum Server Slot Bandwidth	Maximum Server Slot Type
PCI	1992	1.06 Gb/sec	*32b/34b*	33 MHz	133 MB/sec	32-bit Simplex
PCI 2.0	1993	4.26 Gb/sec	*64b/66b*	66 MHz	533 MB/sec	64-bit Simplex
PCI-X	1999	8.5 Gb/sec	*64b/66b*	133 MHz	1.06 GB/sec	64-bit Simplex
PCI-X 2.0	2002	17 Gb/sec	64b/66b	266 MHz	2.13 GB/sec	64-bit Simplex
PCI-Express 1.X	2003	2.5 Gb/sec	8b/10b	2.5 GT/sec	8 GB/sec	x16 Duplex
PCI-Express 2.X	2007	5 Gb/sec	8b/10b	5 GT/sec	16 GB/sec	x16 Duplex
PCI-Express 3.X	2010	8 Gb/sec	128b/130b	8 GT/sec	32 GB/sec	x16 Duplex
PCI-Express 4.0	2017	16 Gb/sec	128b/130b	16 GT/sec	64 GB/sec	x16 Duplex
PCI-Express 5.0	2019	32 Gb/sec	128b/130b	32 GT/sec	128 GB/sec	x16 Duplex
PCI-Express 6.0	2021	64 Gb/sec	PAM-4, FLIT	64 GT/sec	256 GB/sec	x16 Duplex
PCI-Express 7.0	*2023*	*128 Gb/sec*	*PAM-16, FLIT*	*128 GT/sec*	*512 GB/sec*	*x16 Duplex*
PCI-Express 8.0	*2025*	*???*	*???*	*256 GT/sec*	*1 TB/sec*	*x16 Duplex*
PCI-Express 9.0	*2027*	*???*	*???*	*512 GT/sec*	*2 TB/sec*	*x16 Duplex*
PCI-Express 10.0	*2029*	*???*	*???*	*1 TT/sec*	*4 TB/sec*	*x16 Duplex*

Comparison: Ethernet, InfiniBand and Aquila

	Ethernet	InfiniBand	GNet	PCIe Net
Contributor	-	Mellanox	Google	Socnoc
Application	Everywhere	HPC	HPC and Datacenter	Edge and Datacenter
Adapter	NIC	HCA	TiN	Not needed
Protocol	TCP/IP	TCP/IP, RDMA	TCP/IP, RMA	TCP/IP, UMA
Cost ($) per 1Gbps	30-50	50-100	20~30?	3-5
TDP(watts) per 1Gbps	0.3-2	0.5-3	?	0.05 – 0.1

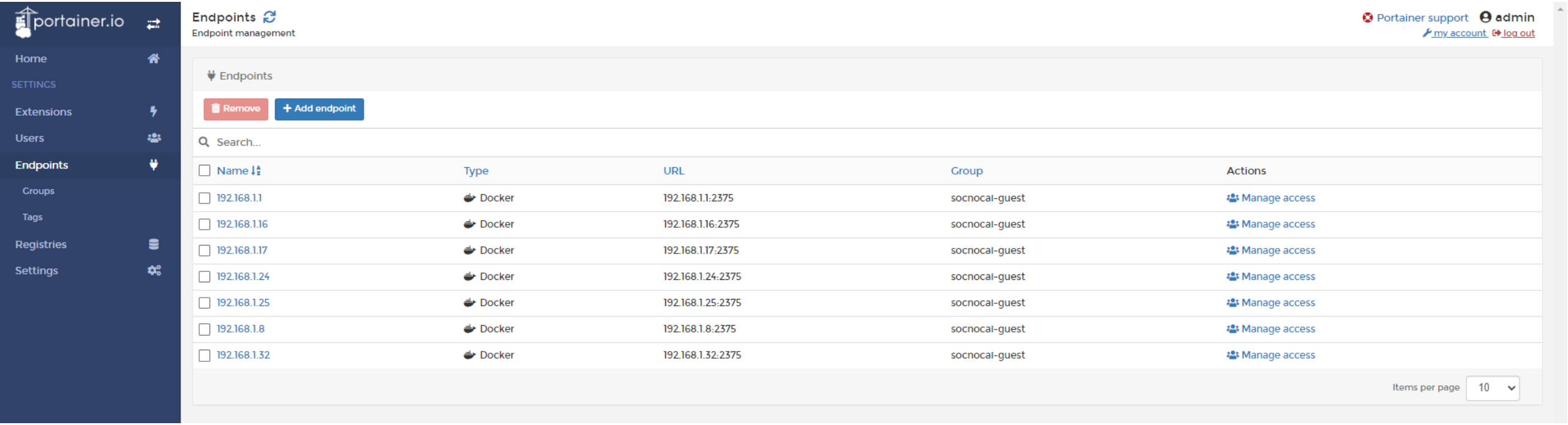
[Release6] Demo Application

- Based on this architecture, the following three use cases are implemented:
  - Use case1: High-density Edge Server/Cluster/Brick
  - Use case2: Edge Database All-in-One
  - Use case3: Edge Android Brick
- Next
  - Use case4: Tri-socket Cloud Server, High-performance cluster

[1] High Density Edge Cluster/Brick

- ¼ Standard 2U space formfactor
- **150 Watts per brick**
- **16 Gbps interconnection**
- 10-20 Gbps configurable outbound connection
- 168 cores @1.0GHZ A53
- AI card supported

[1] Docker Swarm PaaS with GUI



With MAC-in-MAC over PCIe, we can extend the VP/FP to 1000+ per node

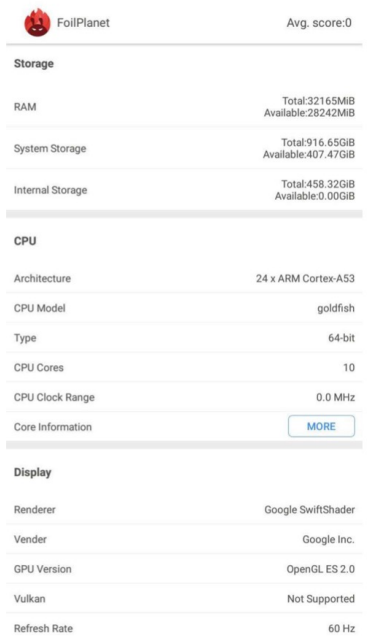
[2] Edge Databrick: PolarDB and TiDB supported

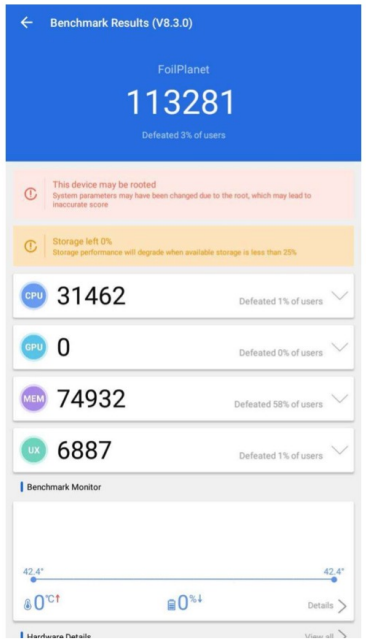
```
[1902926.143882] veth6f13209: renamed from eth0
[1902926.241432] docker0: port 2(veth90d2d35) entered disabled state
[1902926.277682] device veth90d2d35 left promiscuous mode
[1902926.282875] docker0: port 2(veth90d2d35) entered disabled state
[1904264.591426] docker0: port 2(vethfd791da) entered blocking state
[1904264.597576] docker0: port 2(vethfd791da) entered disabled state
[1904264.603870] device vethfd791da entered promiscuous mode
[1904264.609659] IPv6: ADDRCONF(NETDEV_UP): vethfd791da: link is not ready
[1904265.178860] eth0: renamed from vetha285947
[1904265.213756] IPv6: ADDRCONF(NETDEV_CHANGE): vethfd791da: link becomes ready
[1904265.221006] docker0: port 2(vethfd791da) entered blocking state
[1904265.227116] docker0: port 2(vethfd791da) entered forwarding state
[1904265.701099] docker0: port 2(vethfd791da) entered disabled state
[1904265.707334] vetha285947: renamed from eth0
[1904265.809211] docker0: port 2(vethfd791da) entered disabled state
[1904265.845497] device vethfd791da left promiscuous mode
[1904265.850690] docker0: port 2(vethfd791da) entered disabled state
[1905093.586259] docker0: port 2(veth5bab49b) entered blocking state
[1905093.592389] docker0: port 2(veth5bab49b) entered disabled state
[1905093.598694] device veth5bab49b entered promiscuous mode
[1905093.604509] IPv6: ADDRCONF(NETDEV_UP): veth5bab49b: link is not ready
[1905094.178399] eth0: renamed from veth0e959c3
[1905094.209270] IPv6: ADDRCONF(NETDEV_CHANGE): veth5bab49b: link becomes ready
[1905094.216504] docker0: port 2(veth5bab49b) entered blocking state
[1905094.222614] docker0: port 2(veth5bab49b) entered forwarding state
[1905094.708029] docker0: port 2(veth5bab49b) entered disabled state
[1905094.714275] veth0e959c3: renamed from eth0
[1905094.823988] docker0: port 2(veth5bab49b) entered disabled state
[1905094.860252] device veth5bab49b left promiscuous mode
[1905094.865430] docker0: port 2(veth5bab49b) entered disabled state
[1907665.272865] docker0: port 2(vethbbaa260) entered blocking state
[1907665.279008] docker0: port 2(vethbbaa260) entered disabled state
[1907665.285333] device vethbbaa260 entered promiscuous mode
[1907665.291168] IPv6: ADDRCONF(NETDEV_UP): vethbbaa260: link is not ready
```

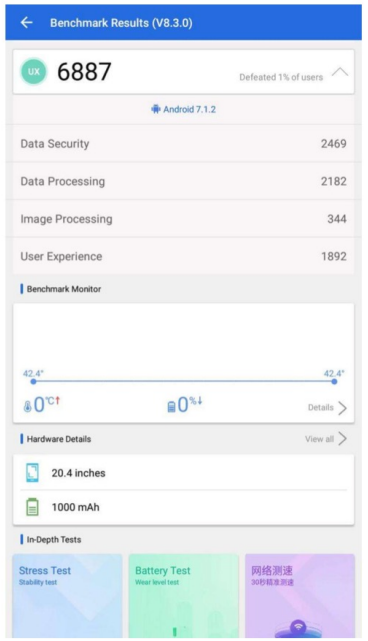
```
$HOME/tmp_basedir_polardb_pg_1100_bld/bin/psql \
    -p 5432 -h 127.0.0.1 -c 'select version();'
            version
------------------------------
 PostgreSQL 11.9 (POLARDB 11.9)
(1 row)
```

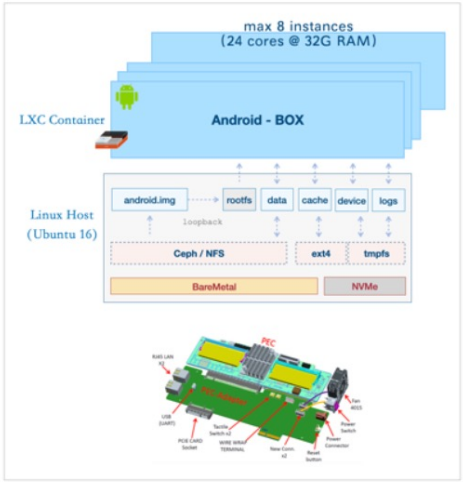
	Huawei Taishan 200	X86 Server	Feiteng Server	QE100-testbed
Data source	*2-华为鲲鹏-陈龙.pdf*		Testbed	Testbed
CPU	Kunpeng920-5250 (48 cores @ 2.6Ghz)	Xeon® E5-2630 v4 (20 vcores @ 2.2Ghz)	FT2000+ (64 cores @ 2.2GHz)	QE100 (24 cores @ 1.0Ghz)
CPUs each node/chasis	4	2	1	32
RAM		128GB	128GB	32GB
DISK		**Intel Optane SSD P4800X 375G * 1**	Intel SSD SATA 3 S4510 960G * 8	Samsung SSD NVMe PM961 500G * 1
NIC	**100GE**	10Gb Ethernet	10Gb Ethernet	10Gb DDT
Nodes/Chasis	**6**	3	4	1
threads		**3 * 128**	4 * 120	**32 * 60**
oltp_point_select	1,621,668	208,190	91,850	244,091
oltp_update_index	141,331	20,447	1,890	38,003
oltp_insert	171,412		2,095	7,554
oltp_read_write (TPS)	25,733		645	2,510
oltp_read_write (QPS)			13,089	50,205
oltp_read_only (TPS)		6,684	1,203	7,109
oltp_read_only (QPS)		106,947	19,242	113,751

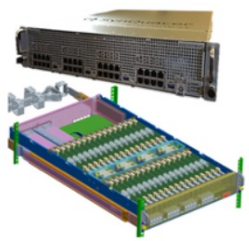
[3] Edge Android Cloud Brick

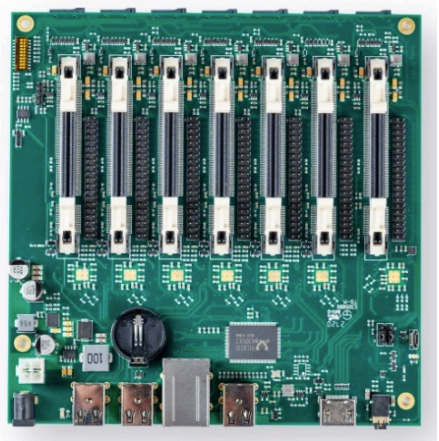








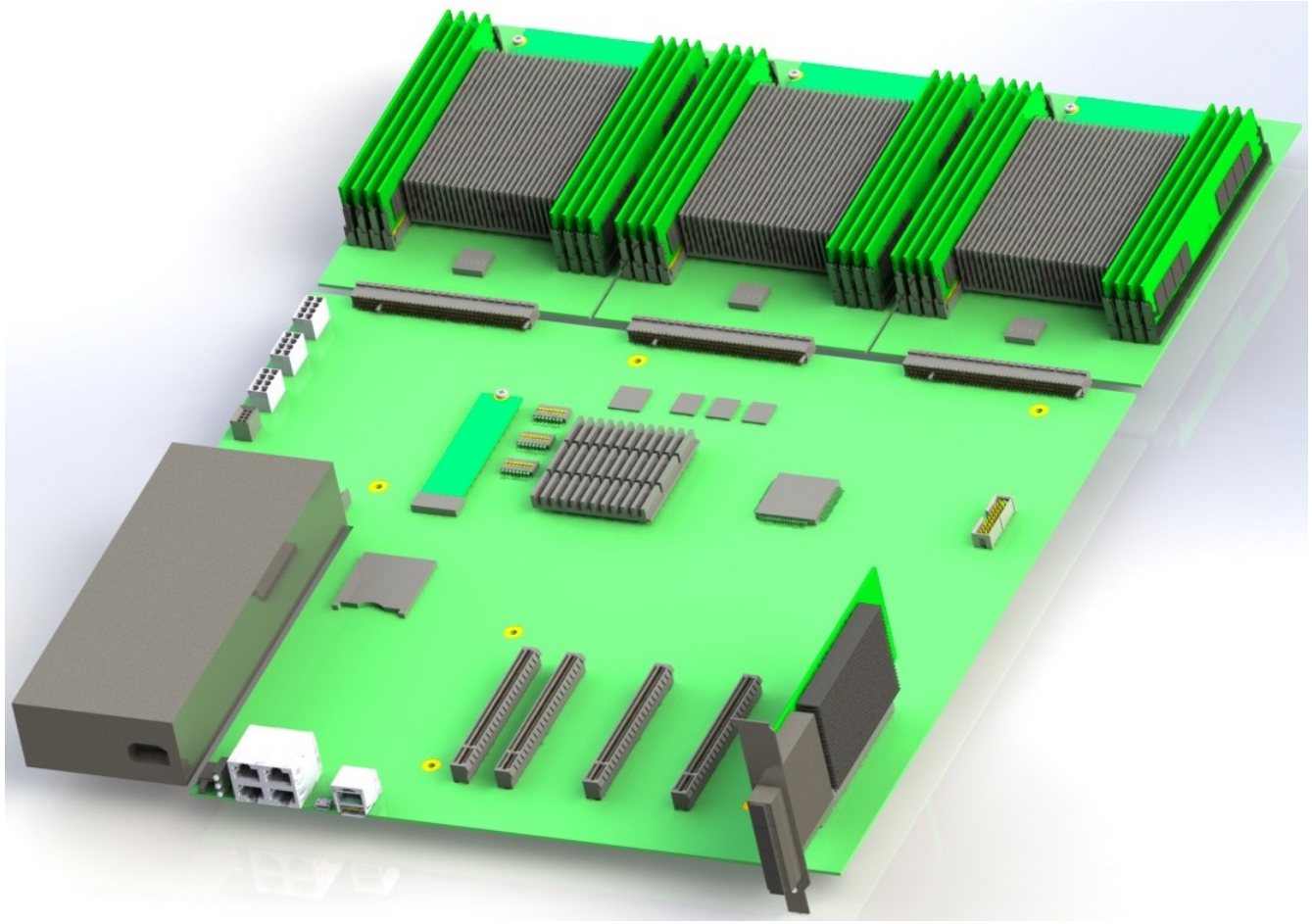




Increase bandwidth by x10

Reduce network cost by 80%

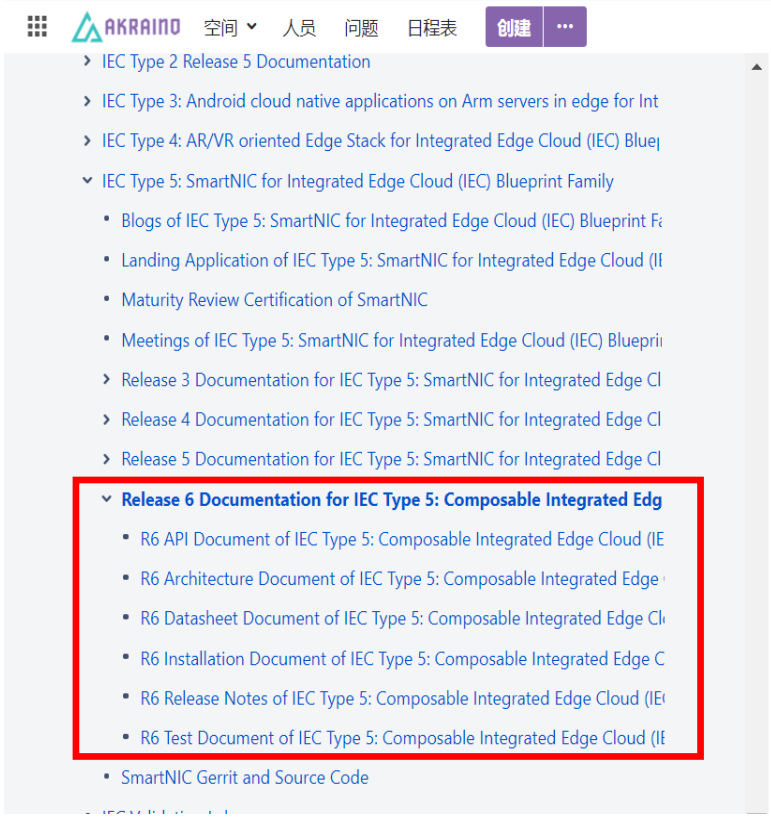
[4] High Performance Cloud-on-Board Server

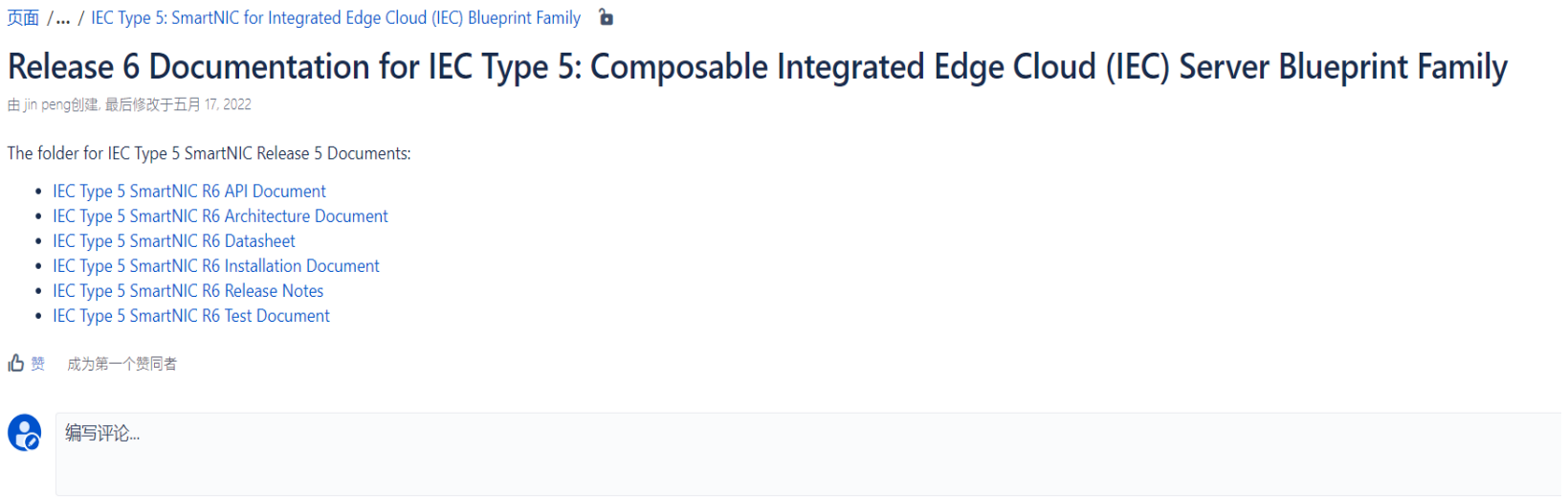


- Standard 2U formfactor
- **600 Watts per server**
- **32 Gbps interconnection**
- 200+ cores @3.0GHZ
- AI card supported

[Release6] Documentations

https://wiki.akraino.org/display/AK/Release+6+Documentation+for+IEC+Type+5%3A+Composable+Integrated+Edge+Cloud+%28IEC%29+Server+Blueprint+Family





[Release6] Release Status

15	IEC Type 3: Android cloud native applications on Arm servers in edge for Integrated Edge Cloud (IEC) Blueprint Family @hanyu ding @Rajeev Gadgil @Davy Zhang		No	Incubation	https://nexus.akraino.org/content/sites/logs/ysemi/job/v1/upload/iec-tox-verify-master_317/	https://wiki.akraino.org/download/attachments/24084647/IEC%20Release3-IEC%20Type3-datasheet.docx?version=5&modificationDate=1591272863000&api=v2	API form uploaded by Davy Zhang 2022-4-28 , scheduled for review at API subcommittee meeting 2022-4-29 Approved by API subcommittee at weekly meeting 2022-4-29	https://nexus.akraino.org/content/sites/logs/ysemi/job/v1/validation_results_v4/ 2022-5-12 https://nexus.akraino.org/content/sites/logs/ysemi/job/v1/validation_results_v5/ 2022-5-13 https://nexus.akraino.org/content/sites/logs/ysemi/job/v1/validation_results_v6/ 2022-5-17	2022-5-17 **lynis results:** Accepted **vuls results:** **CVE-2017-12194: failed in scan; Fix is partially available for "spice" package according to the Ubuntu CVE database. Please update and list the versions of the "spice" packages on the exception request page.** **CVE-2019-19948: failed in scan: Fix is available according to the Ubuntu CVE database.** **CVE-2019-19949: failed in scan: Fix is available according to the Ubuntu CVE database.** **kube-hunter results:** Accepted	Approved
16	IEC Type 5: SmartNIC for Integrated Edge Cloud (IEC) Blueprint Family @Leo Li @jin peng		No	Incubation	https://dev.socnoc.cn/#/home	SOCNOC Release 6 One pag - Akraino - Akraino Confluence	No API changes expected from R5, per Leo Li in TSC meeting 2022-4-7 . Waiting for e-mail from Leo to confirm this Leo confirmed by mail 2022-5-11	socnoc - Akraino - Akraino Confluence		
17	Enterprise Applications on Lightweight 5G Telco Edge @Gaurav Agrawal		No	Incubation	https://nexus.akraino.org/content/sites/logs/huawei/job/Ealtedge-aio-log/15/	EALTEDGE Release 6 Datasheet	No API changes from R5, per e-mail from Khemendra Kumar 2022-4-28 Info for EALTEdge APIs: https://wiki.akraino.org/pages/viewpage.action?pageId=53478299	https://nexus.akraino.org/content/sites/logs/huawei/job/Ealt-edge-security-test/26/results/	2022-5-16 **lynis results:** Accepted **vuls results:** Accepted **kube-hunter results:** ***pod:*KHV043 - Cluster Health Disclosure** Disable --enable-debugging-handlers kubelet flag.	Approved

Thank You!