ITRENEW

HARDWARE FOR KUBERNETES

PEELING BACK THE LAYERS

Erik Riedel, PhD
Senior VP, Engineering | ITRenew

@RiedelAtWork

Hyperscale for All: Powering the Circular Data Center

ITRenew delivers maximum financial & sustainability returns from open technology

CIRCULAR CLOUD

Strategic Advisory Services

Infrastructure planning

TCO & Sustainability Modeling

Lifetime value maximization

DECOMMISSIONING

and Data Security

Data center decommissioning services

Teraware data sanitization platform

Value Recovery (*$1B+ TCO to date*)

End-to-end logistics solutions

SESAME BY ITRENEW

Rack-Scale Solutions

Rack-scale solutions for data centers

Open systems, HCI, AI/ML

Breakthrough TCO

EDGE SOLUTIONS

and Components

Edge solutions & building blocks

Server components

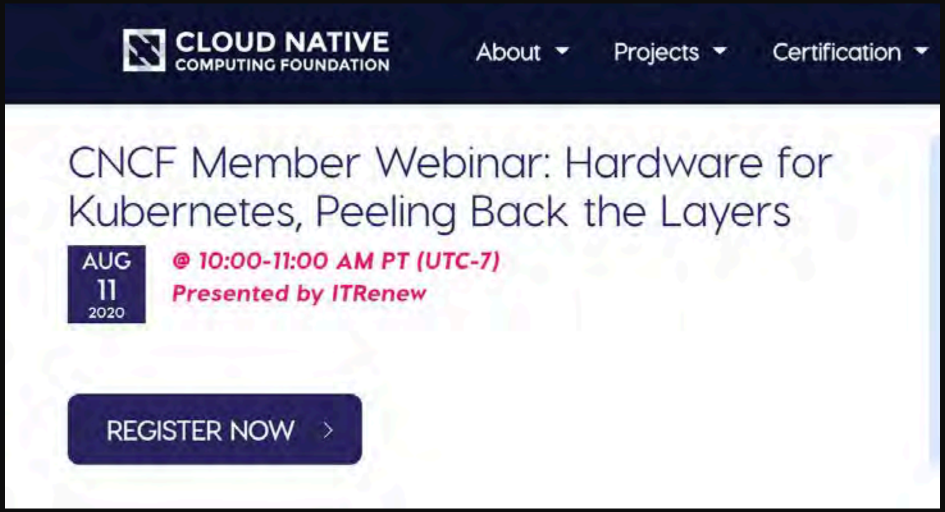
Laptop and PC memory

amazon · Dropbox · ebay · facebook · Google · PayPal · Uber

CNCF Member Webinar

Hardware for Kubernetes, Peeling Back the Layers

August 11, 2020 10:00 AM (America/Los Angeles)



Kubernetes enables developers to deploy and manage applications dynamically, making them more efficient, powerful, and extensible. Many describe the shift away from monolithic stacks on single-purpose machines to cloud native as a “decoupling of applications from infrastructure,” but the reality is that containerized and virtualized software still demands reliable, resilient, and scalable hardware - the “servers” in “serverless”. Because hardware design & performance directly affects the experience that users have with applications and with services, everyone building apps should appreciate the infrastructure layers that live under the work they do every day at least a little.

New models such as hyperscale design and open hardware can be significantly more efficient and cost-effective, making it possible to further stretch and scale users and workloads. Subject matter experts across the industry in servers, storage, networking, power, cooling, and data centers are ensuring that these complex ecosystems work together in harmony and at peak efficiency end-to-end. These systems are designed in the open community and can be matched to run Kubernetes clusters with top performance and scalability from deskside to data center. In this session, we will peel back some of the infrastructure layers that are usually hidden away, demonstrate some of the latest innovations in hyperscale design, and illustrate how to harness the power of the wide and deep hardware ecosystem to realize cloud native applications.

Join Erik Riedel, SVP Engineering at ITRenew, as he draws on 20 years of building hardware for clouds before they had a cool name, to learn how it all works and what it means for you in practice.

Photo by Karolina Grabowska from Pexels

Outline

- layers
- open
- apps + data
- infrastructure
- servers
- circular
- progress
- data
- more stories – disks, BIOS, networks

open

Photo by **Robin Kumar Biswal** from **Pexels**

Open Is Necessary, But Not Sufficient Per Se

design, engineering and product expertise, rigorous commitment to component quality

[software]

requires baseline partnership across hardware & software

[hardware]

design, engineering and product expertise, rigorous commitment to component quality

ORCHESTRATING AN OPEN HARDWARE + OPEN SOFTWARE STACK

devops – collaboration from designers & developers to operations

can't be efficient at hyperscale without the machines making many, most decisions

from capacitors to consensus algorithms to cognitive load

wide dynamic range of skills must be applied & appreciated

automation, analytics, and machine learning for all the things

Platinum

- 2crsi (since 2018)
- 3M (since 2018)
- Alibaba (since 2017)
- Arista Networks (since 2019)
- Inspur (since 2016)
- Intel (since 2011)
- ITRenew (since 2018)
- Microsoft (since 2014)
- ARM (since 2018)
- Asperitas (since 2017)
- ASUS (since 2019)
- AT&T (since 2015)
- MiTAC (since 2017)
- Nokia (since 2015)
- NVIDIA Networking – Mellanox (since 2012)
- Quanta Cloud Technology (since 2012)
- Baidu (since 2019)
- Cumulus Networks (since 2013)
- Delta Electronics (since 2016)
- Deutsche Telekom (since 2016)
- Rackspace (since 2011)
- Rittal (since 2017)
- Schneider (since 2014)
- Silicom (since 2018)
- Edgecore Networks (since 2016)
- Facebook (since 2011)
- Goldman Sachs (since 2011)
- Google (since 2015)
- STORDIS (since 2019)
- Submer (since 2018)
- Tencent (since 2018)
- VeriSilicon (since 2020)
- HPE (since 2015)
- Huawei (since 2018)
- Hyve Solutions (since 2012)
- IBM (since 2013)
- Wiwynn (since 2014)
- Yahoo! Japan (since 2017)

OPEN Compute Project®

Gold

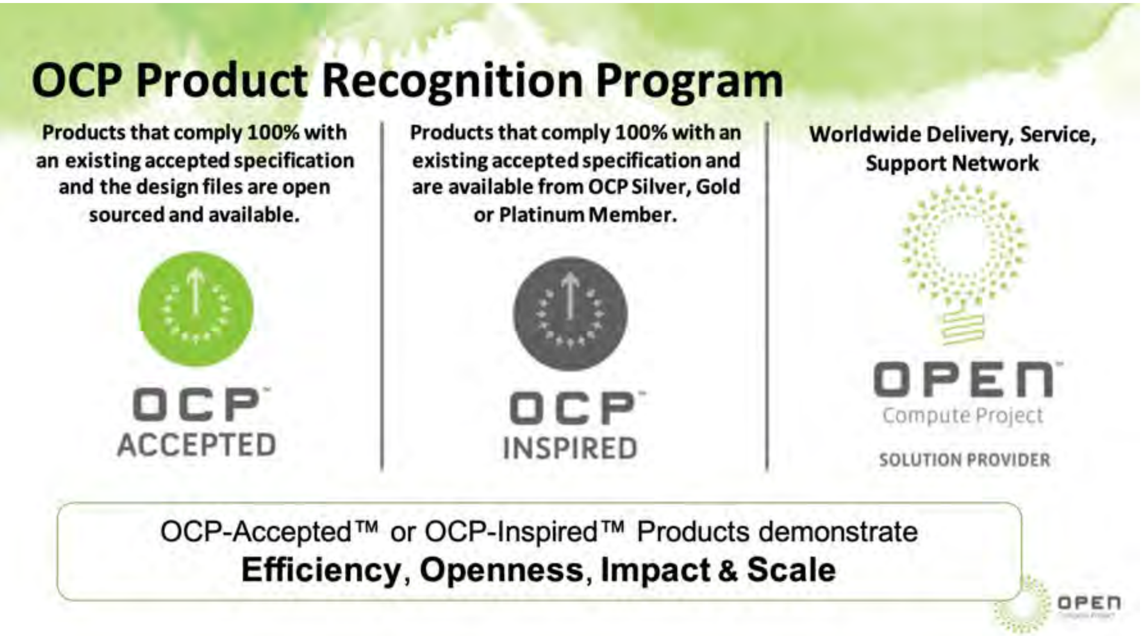
- ITOCHU Techno-Solutions Corporation (since 2014)
- Samsung Electronics (since 2019)
- Seagate (since 2017)
- ZT Systems (since 2019)

Silver

- Circle B (since 2016)
- Cisco (since 2014)
- Inventec (since 2014)
- NVIDIA (since 2017)

Open Compute Summit (NYC) – October 2011

OCP certifications



		location	attendees	companies
April 2011	1st	founding summit	--	--
October 2011	2nd	NYC	300	--
May 2012	3rd	San Antonio	500	--
January 2013	4th	Santa Clara	1,000	--
January 2014	5th	Santa Clara	3,400	--
March 2015	6th	San Jose	2,500	800
March 2016	7th	San Jose	2,400	600
March 2017	8th	Santa Clara	2,800	550
March 2018	9th	San Jose	3,400	800
March 2019	10th	San Jose	3,600	725
May 2020	**11th**	**virtual**	**7,500**	**2,400**

Data Center Facility

Sub-Projects:
Modular Data Center
Critical Facility Operations - Incubation
Advanced Cooling Facility - Incubation

Hardware Management

Sub-Projects:
OpenRMC
Hardware Management Module - Incubation
Hardware Fault Management - Incubation

Networking

Sub-Projects:
ONIE
Open Network Linux
SAI
SONiC

Open System Firmware

Rack & Power

Sub-Projects:
ACS Immersion
ACS Cold Plate
ACS Door Heat Exchange

Security

Server

Sub-Projects:
High Performance Computing - Incubation
Mezz (NIC)
Open Accelerator Infrastructure
Open Domain-Specific Architecture

Storage

Telco

Sub-Projects:
openEDGE

Project

- Server (65)
- Networking (48)
- Rack & Power (36)
- Telco (21)
- Data Center Facility (15)
- Storage (13)
- Security (Incubation) (2)

Show more

Contributor

- Facebook (52)
- Microsoft (35)
- Edgecore Networks (18)
- Intel (7)
- AT&T (6)
- Delta Electronics (6)
- Inspur (6)

Show more

Family

- Network Switch (38)
- OpenRack v2 (24)
- OCS (18)
- OTHER (15)
- Olympus (14)
- Data Center (10)
- Storage (8)
- Telco (8)
- Power (7)
- OpenRack (6)
- SOC Boards (6)
- Server (6)
- 19" Server (5)
- Software (5)
- Accessory (4)
- Optical NW (4)
- ACS (3)
- CG-Openrack-19 (3)
- PCI Card (3)
- Access Point (2)
- Barreleye (2)
- Mezz Card (2)
- OCP Mezzanine (2)
- Security (2)
- uCPE (2)
- Debug Card (1)
- Honey Badger (1)
- Information (1)
- Open Vault Storage (1)

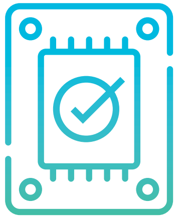
The Benefits of Open Hardware

MORE FLEXIBILITY

Multi-vendor, standards-based hardware for modular solutions to fit your needs

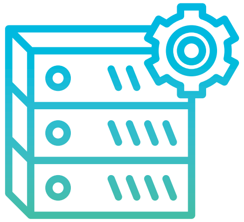
HIGH DENSITY COMPUTING

More server, storage, and network capacity, in less space saves costs



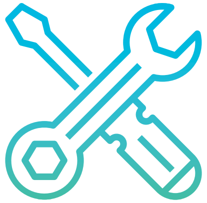
OPTIMIZED POWER

Rack-level power vs. individual server power. More efficient. Less cost. Fewer points of failure



OPTIMIZED COOLING

Rack-level cooling to operate more efficiently. Even more with free-air cooling, if the data centers support it



STREAMLINED MAINTENANCE

Flexible, easy-access design enables faster troubleshooting, updates, and upgrades

The Benefits of Open Software

MORE FLEXIBILITY

Multi-vendor, standards-based software for modular solutions to fit your needs

HIGH DENSITY COMPUTING

More automation, with API-driven scalability, allows more software per silicon in^2

OPTIMIZED POWER

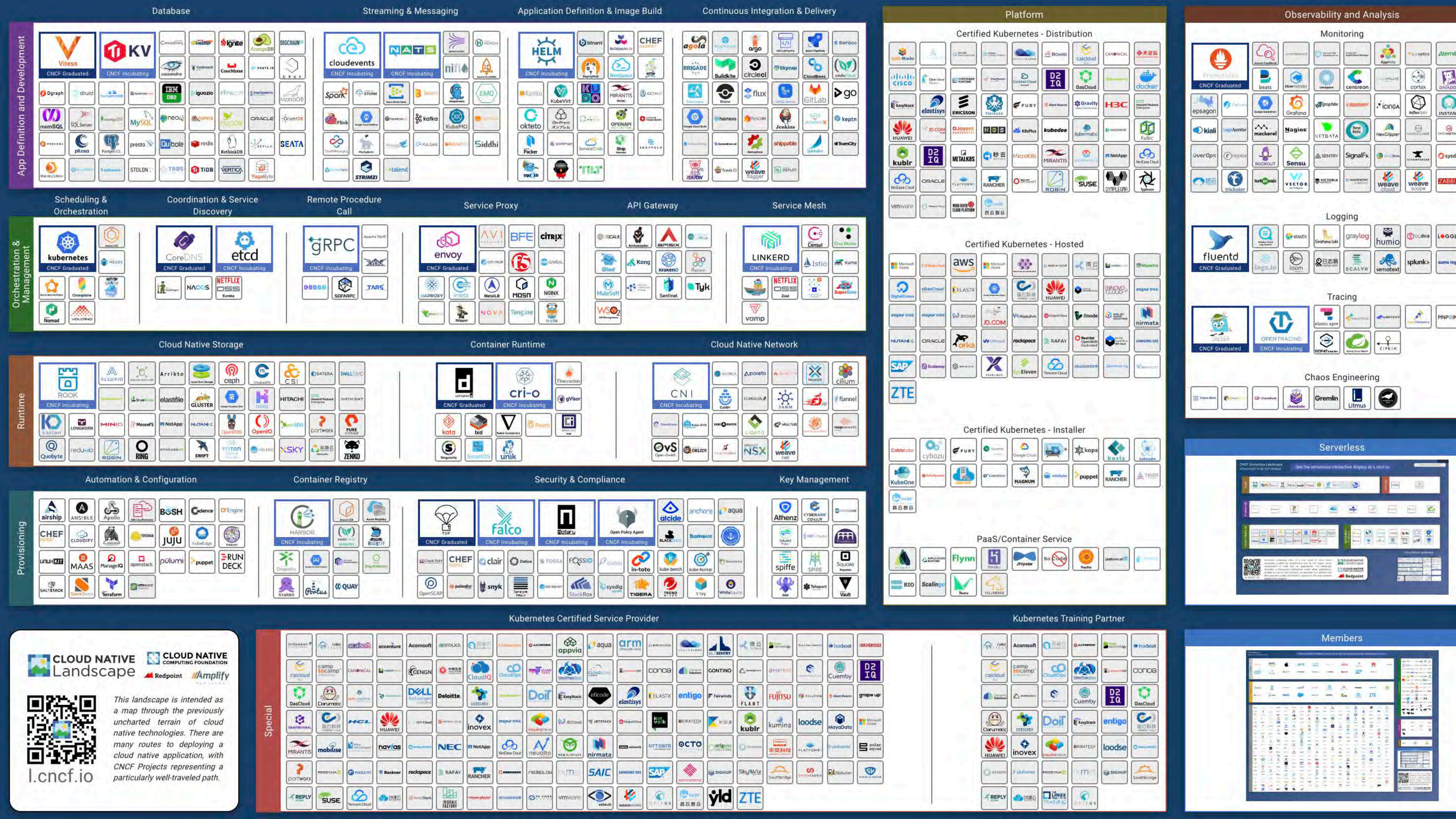
Stack-level power vs. individual packaged software. More efficient. Less cost. Fewer points of failure

OPTIMIZED COOLING

Stack-level continuous integration, continuous deployment (CI/CD) to validate more efficiently. Fewer points of failure in the field

STREAMLINED MAINTENANCE

Flexible, API-based, devops-considered design enables faster troubleshooting, updates, and upgrades

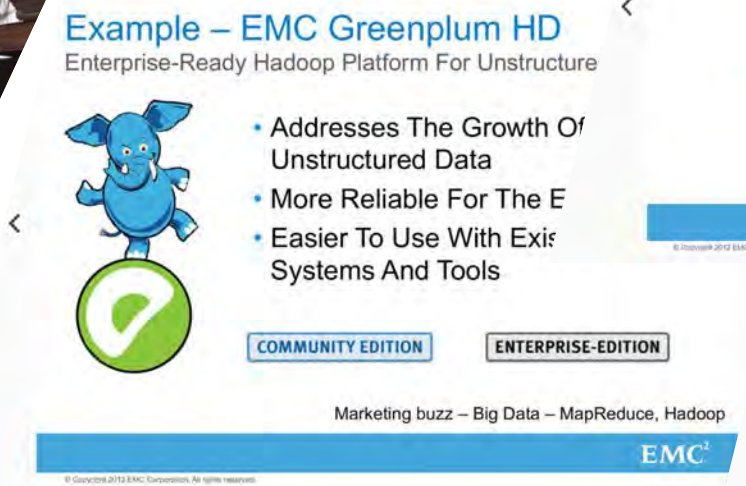


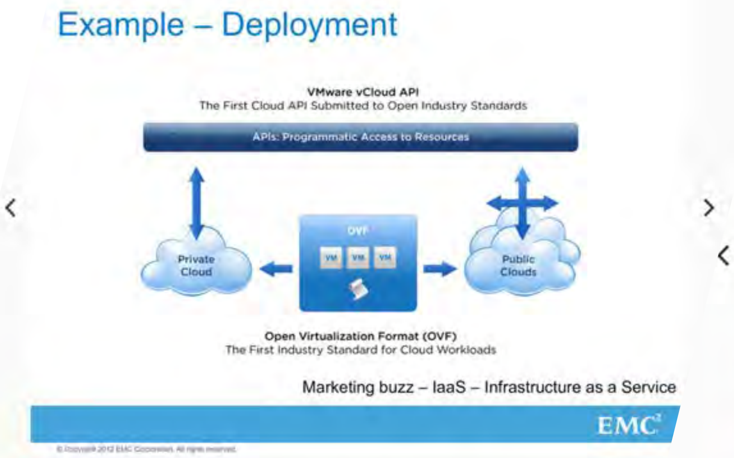
App Definition and Development
Database
Streaming & Messaging
Application Definition & Image Build
Continuous Integration & Delivery
Orchestration & Management
Scheduling & Orchestration
Coordination & Service Discovery
Remote Procedure Call
Service Proxy
API Gateway
Service Mesh
Runtime
Cloud Native Storage
Container Runtime
Cloud Native Network
Provisioning
Automation & Configuration
Container Registry
Security & Compliance
Key Management
Platform
Certified Kubernetes - Distribution
Certified Kubernetes - Hosted
Certified Kubernetes - Installer
PaaS/Container Service
Observability and Analysis
Monitoring
Logging
Tracing
Chaos Engineering
Serverless
Members
Special
Kubernetes Certified Service Provider
Kubernetes Training Partner
CLOUD NATIVE Landscape
CLOUD NATIVE COMPUTING FOUNDATION
Redpoint
Amplify
This landscape is intended as a map through the previously uncharted terrain of cloud native technologies. There are many routes to deploying a cloud native application, with CNCF Projects representing a particularly well-traveled path.
l.cncf.io

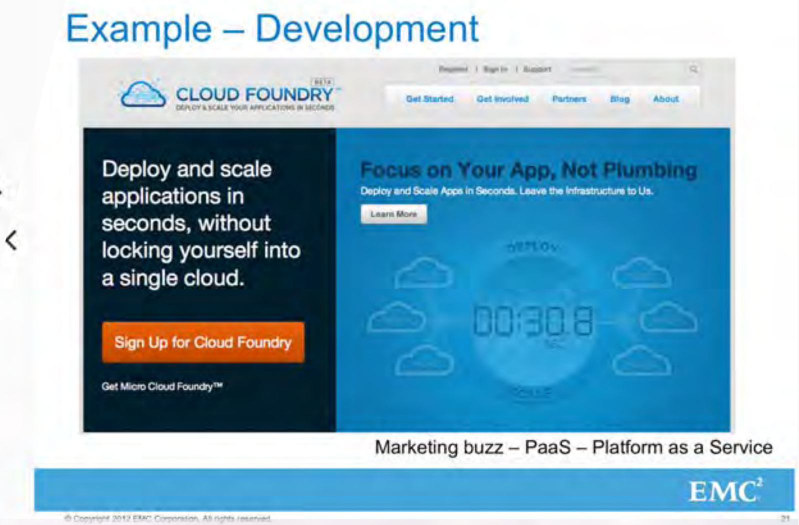
apps + data

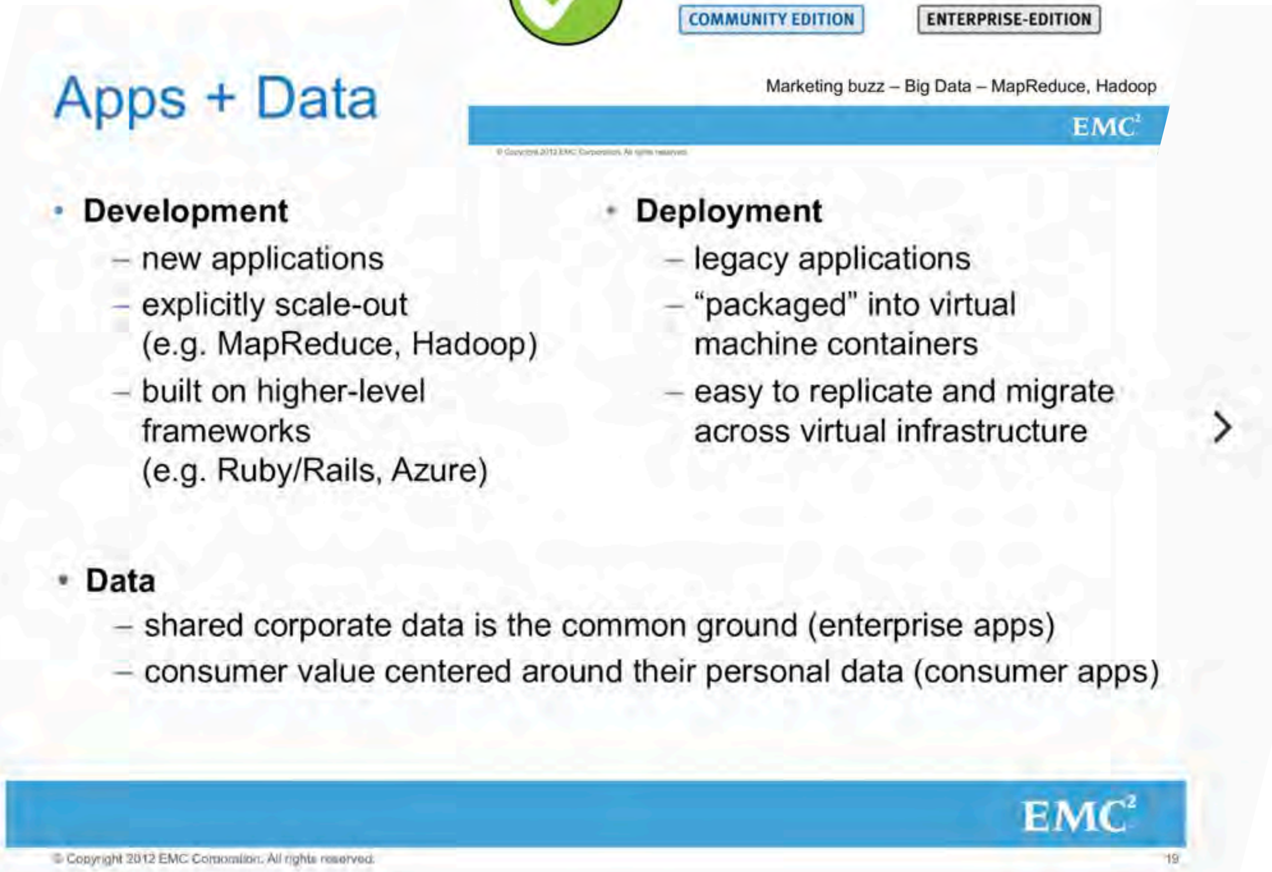
Photo by alleksana from Pexels

looking into the future from 2012





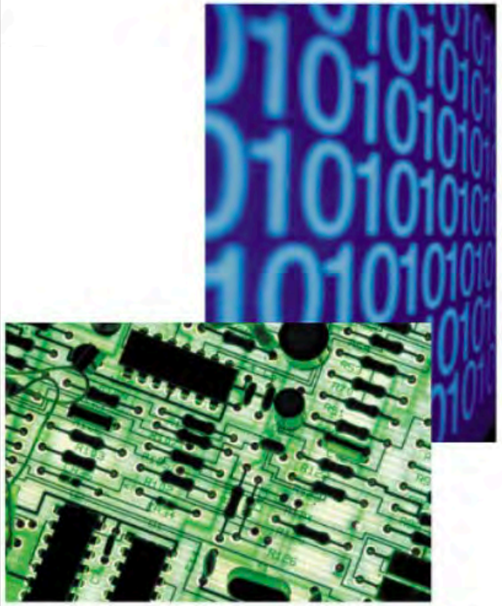


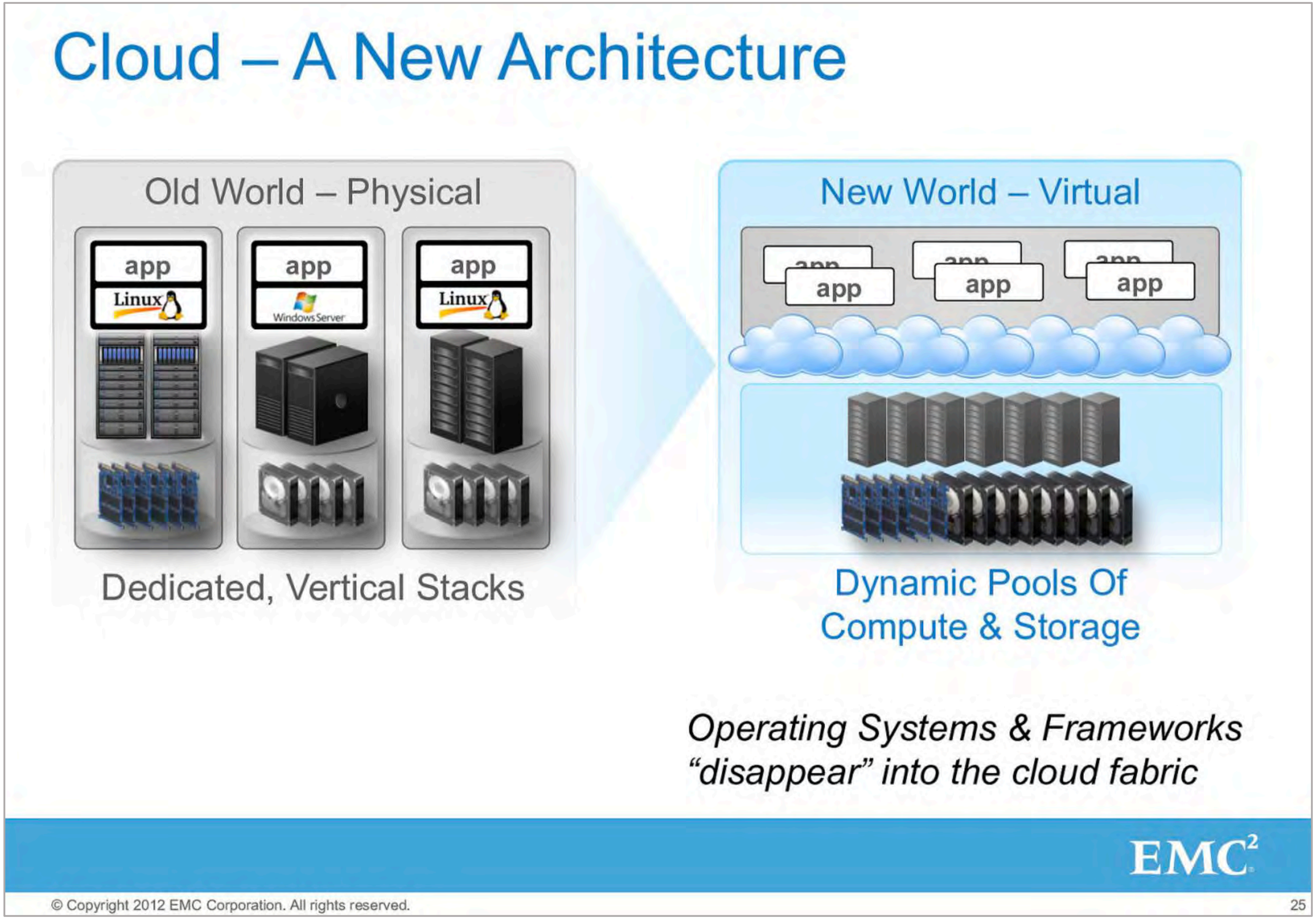


Apps + Data

- **Development**
  - new applications
  - explicitly scale-out (e.g. MapReduce, Hadoop)
  - built on higher-level frameworks (e.g. Ruby/Rails, Azure)
- **Deployment**
  - legacy applications
  - "packaged" into virtual machine containers
  - easy to replicate and migrate across virtual infrastructure
- **Data**
  - shared corporate data is the common ground (enterprise apps)
  - consumer value centered around their personal data (consumer apps)

Looking into the future from 2012



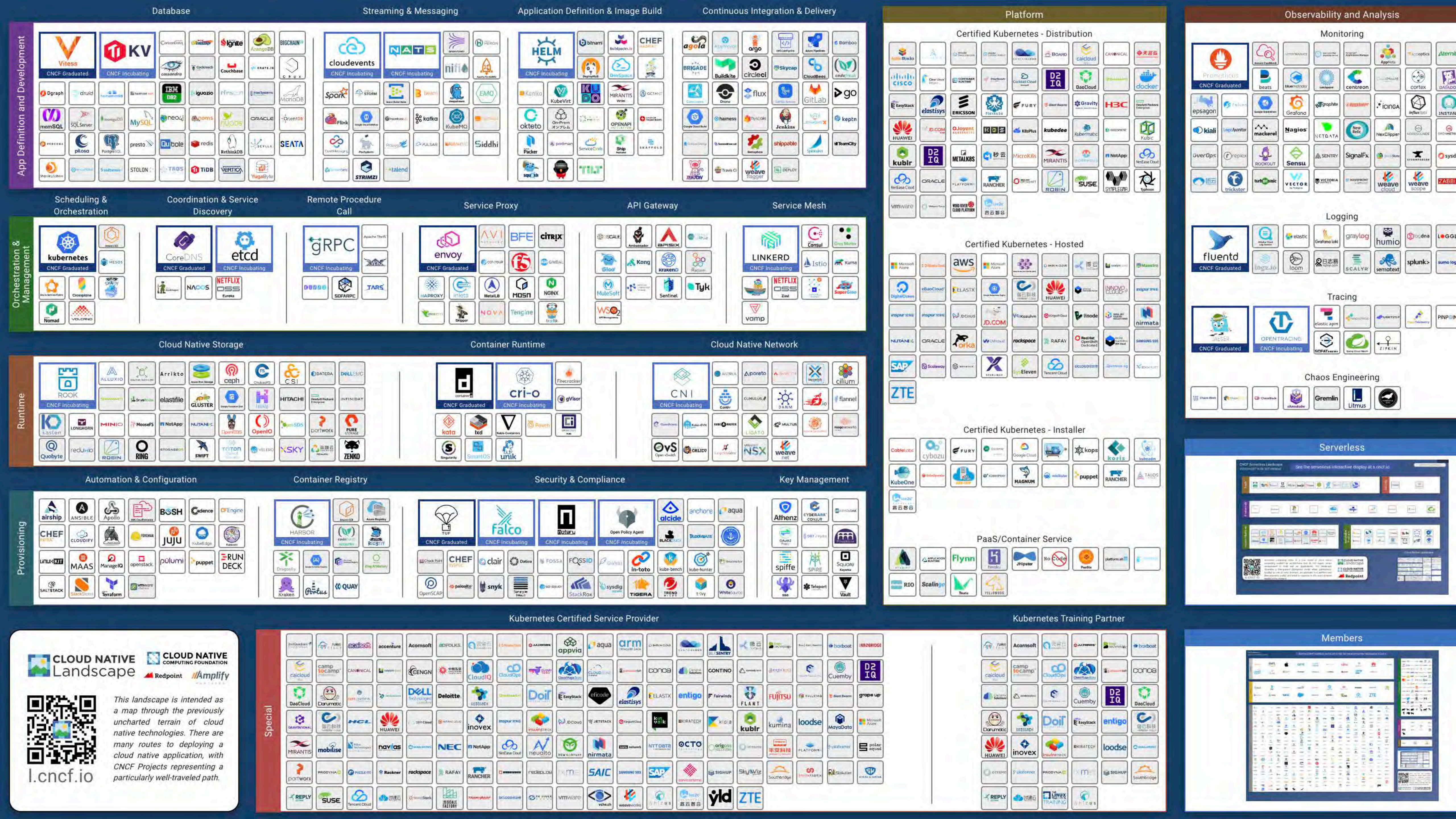


Looking into the future from 2012

fast forward to 2020

Photo by Brayden Law from Pexels


agile
clouds
stacks
FE
FIELD ENGINEER
fieldengineer.com
infrastructure

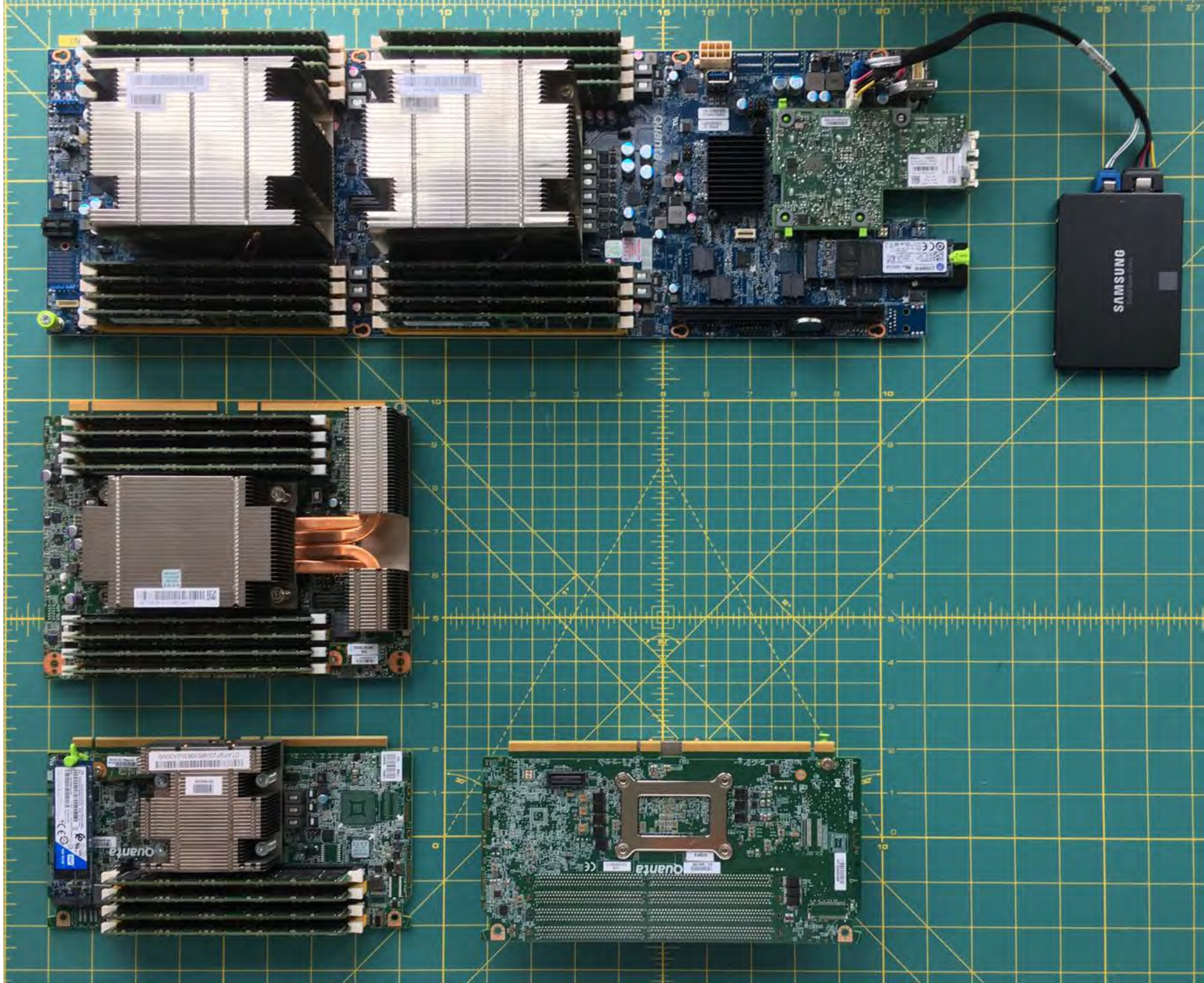

App Definition and Development
Database
Streaming & Messaging
Application Definition & Image Build
Continuous Integration & Delivery
Orchestration & Management
Scheduling & Orchestration
Coordination & Service Discovery
Remote Procedure Call
Service Proxy
API Gateway
Service Mesh
Runtime
Cloud Native Storage
Container Runtime
Cloud Native Network
Provisioning
Automation & Configuration
Container Registry
Security & Compliance
Key Management
Platform
Certified Kubernetes - Distribution
Certified Kubernetes - Hosted
Certified Kubernetes - Installer
PaaS/Container Service
Observability and Analysis
Monitoring
Logging
Tracing
Chaos Engineering
Serverless
Members
Special
Kubernetes Certified Service Provider
Kubernetes Training Partner
CLOUD NATIVE Landscape
CLOUD NATIVE COMPUTING FOUNDATION
Redpoint
Amplify
This landscape is intended as a map through the previously uncharted terrain of cloud native technologies. There are many routes to deploying a cloud native application, with CNCF Projects representing a particularly well-traveled path.
l.cncf.io

infrastructure


threads
Photo by Ksenia Chernaya from Pexels

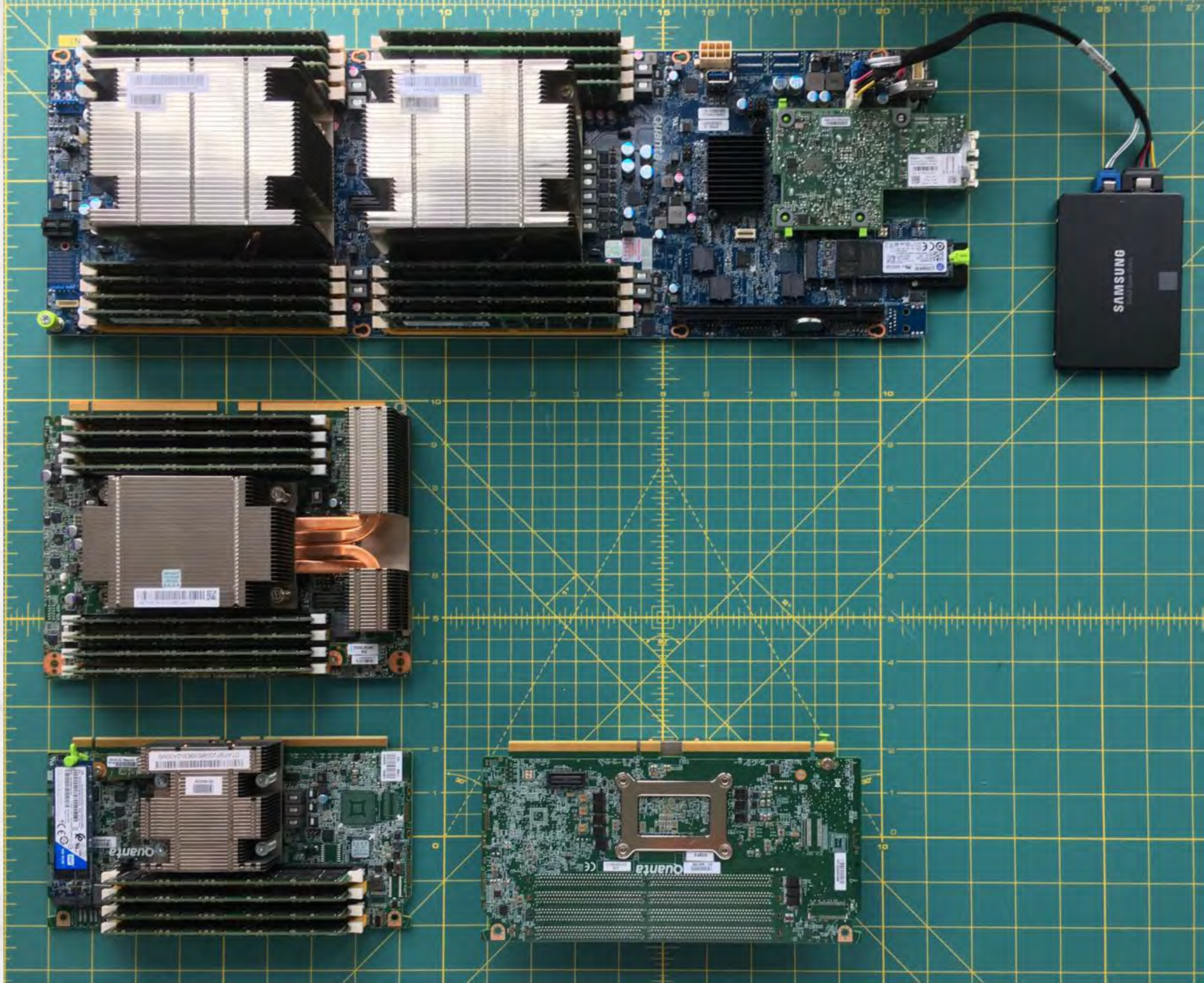

containers
tex
Hapag-Lloyd
AMFCON
CAPITAL

servers



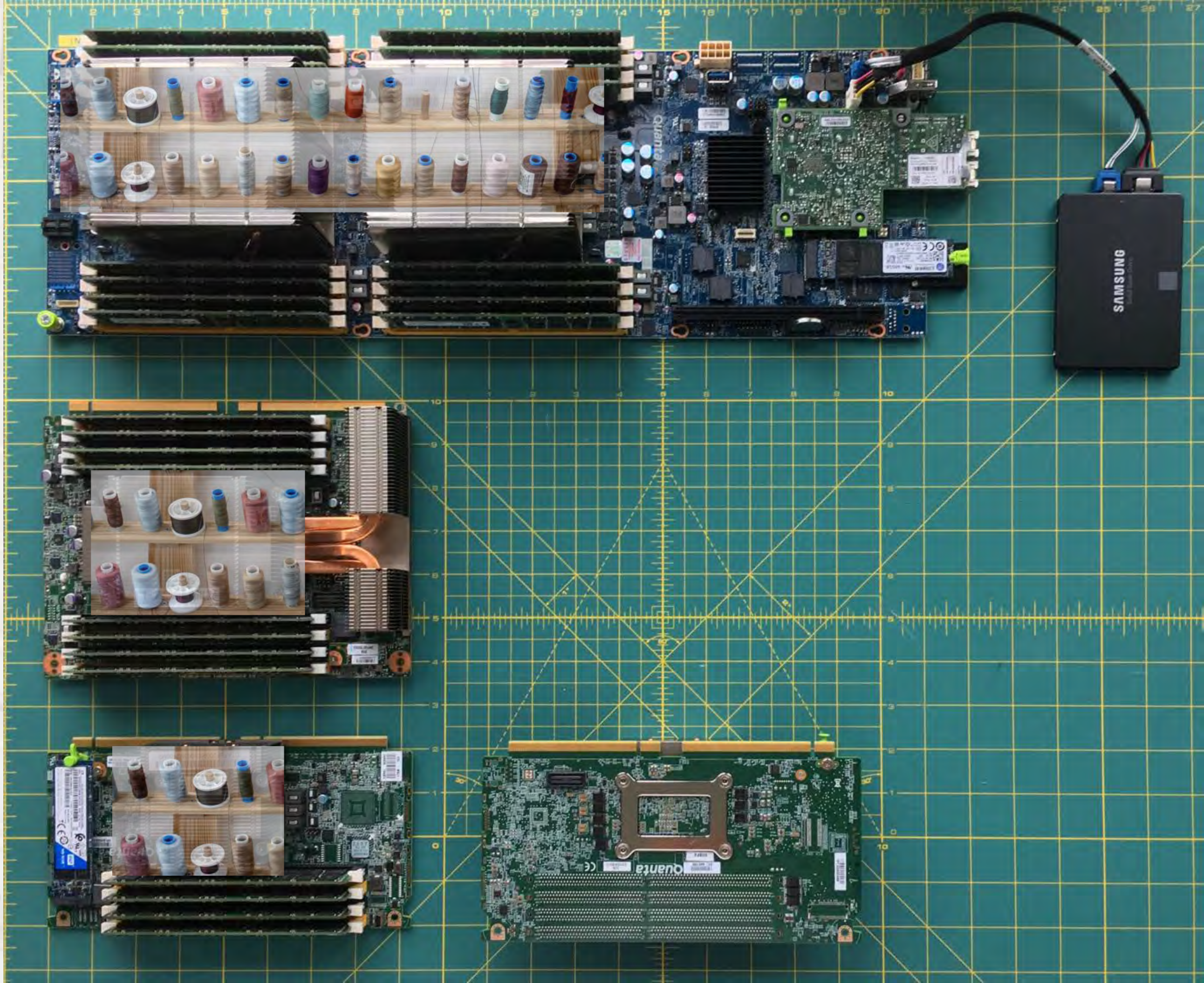
threads

- dual 16-core
  - 32 physical cores
  - 64 virtual cores
- single 18-core
  - 18 physical cores
  - 36 virtual cores
- single 16-core
  - 16 physical cores
  - 32 virtual cores



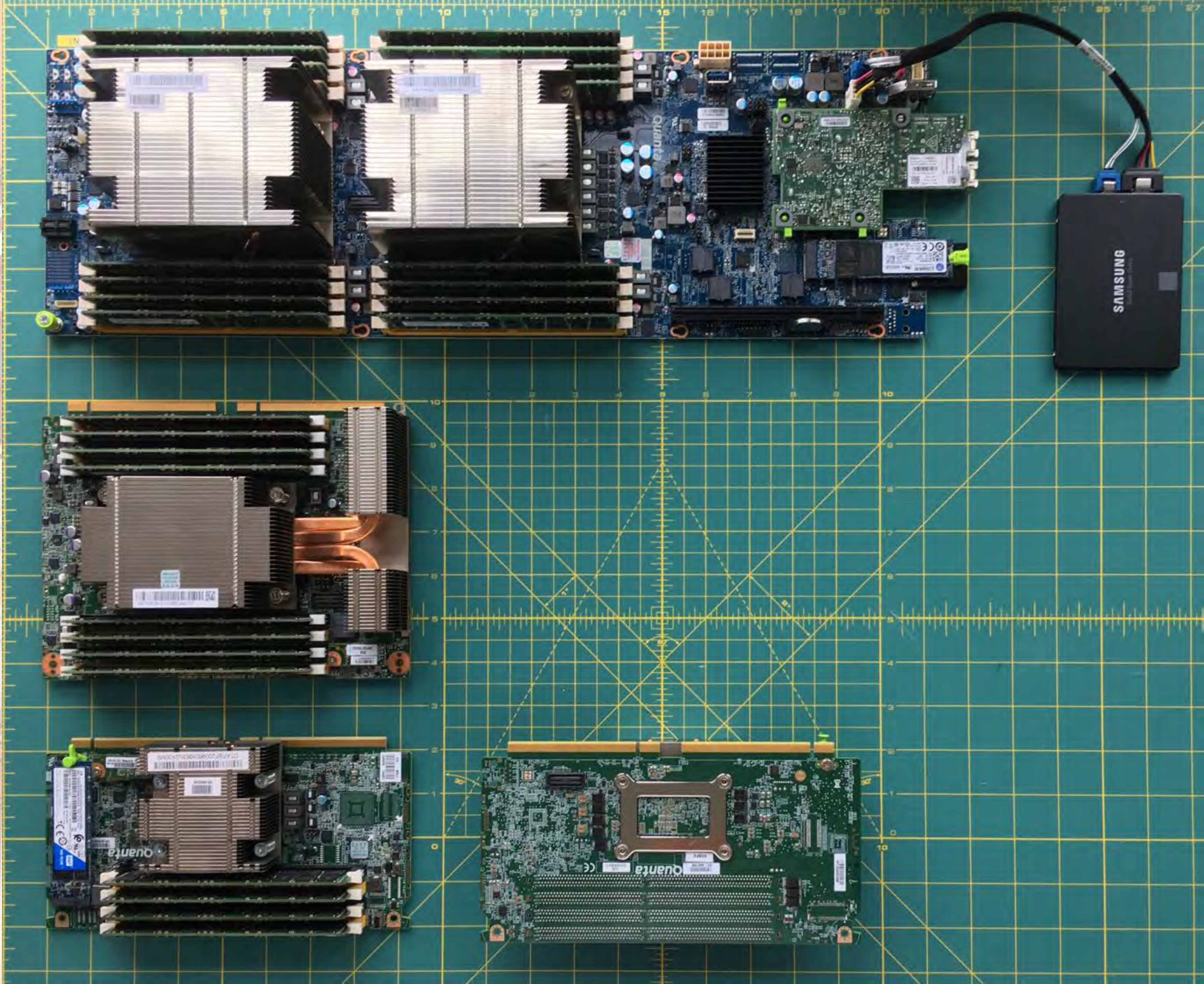
threads

- dual 16-core
  - 32 physical cores
  - 64 virtual cores
- single 18-core
  - 18 physical cores
  - 36 virtual cores
- single 16-core
  - 16 physical cores
  - 32 virtual cores



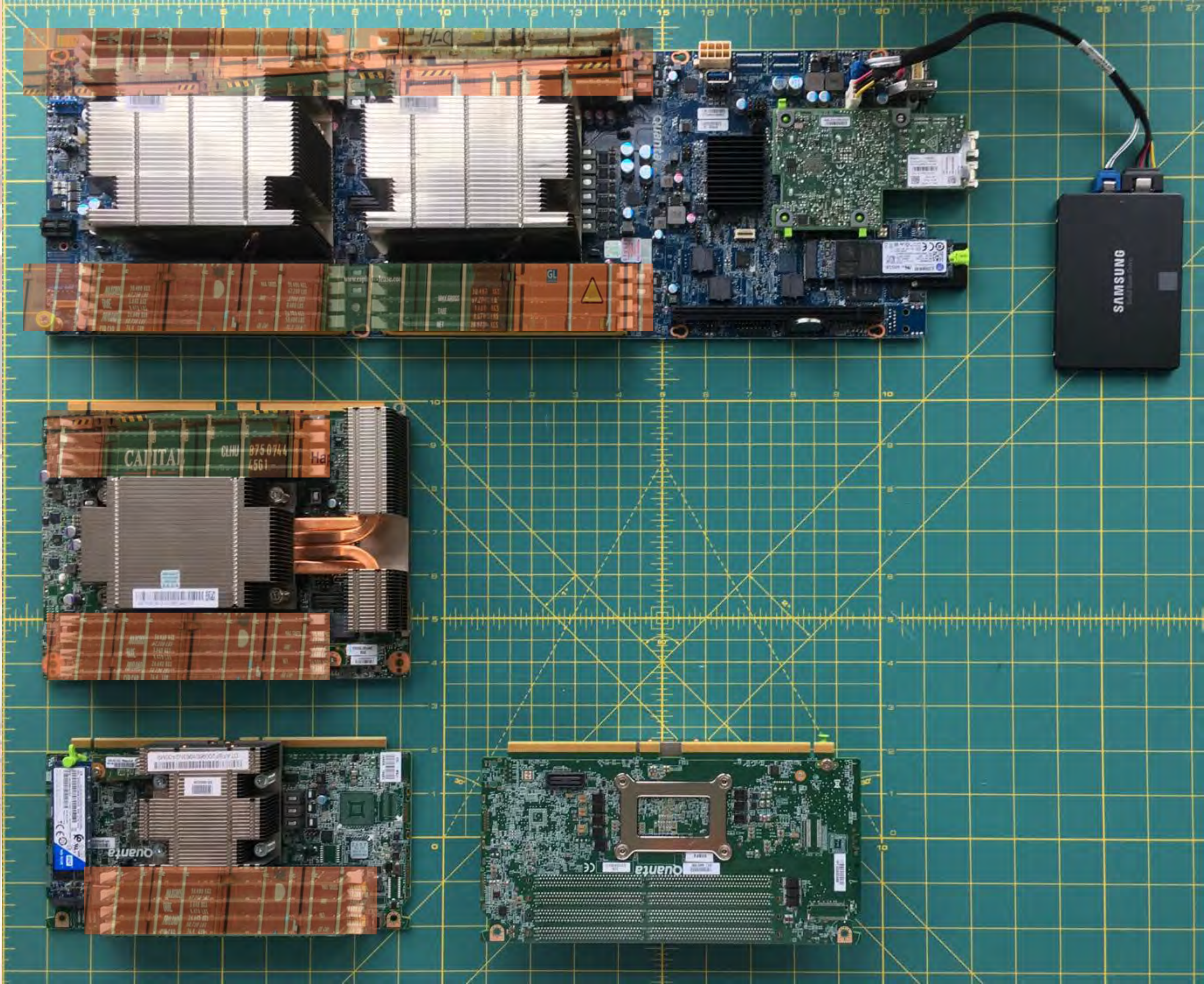
containers

- 512 GB (gigabytes)
  - 16x 32GB dimms
  - 50-100 containers
- 256 GB (gigabytes)
  - 8x 32GB dimms
  - 25-50 containers
- 128 GB (gigabytes)
  - 4x 32GB dimms
  - 12-25 containers



containers

- 512 GB (gigabytes)
  - 16x 32GB dimms
  - 50-100 containers
- 256 GB (gigabytes)
  - 8x 32GB dimms
  - 25-50 containers
- 128 GB (gigabytes)
  - 4x 32GB dimms
  - 12-25 containers



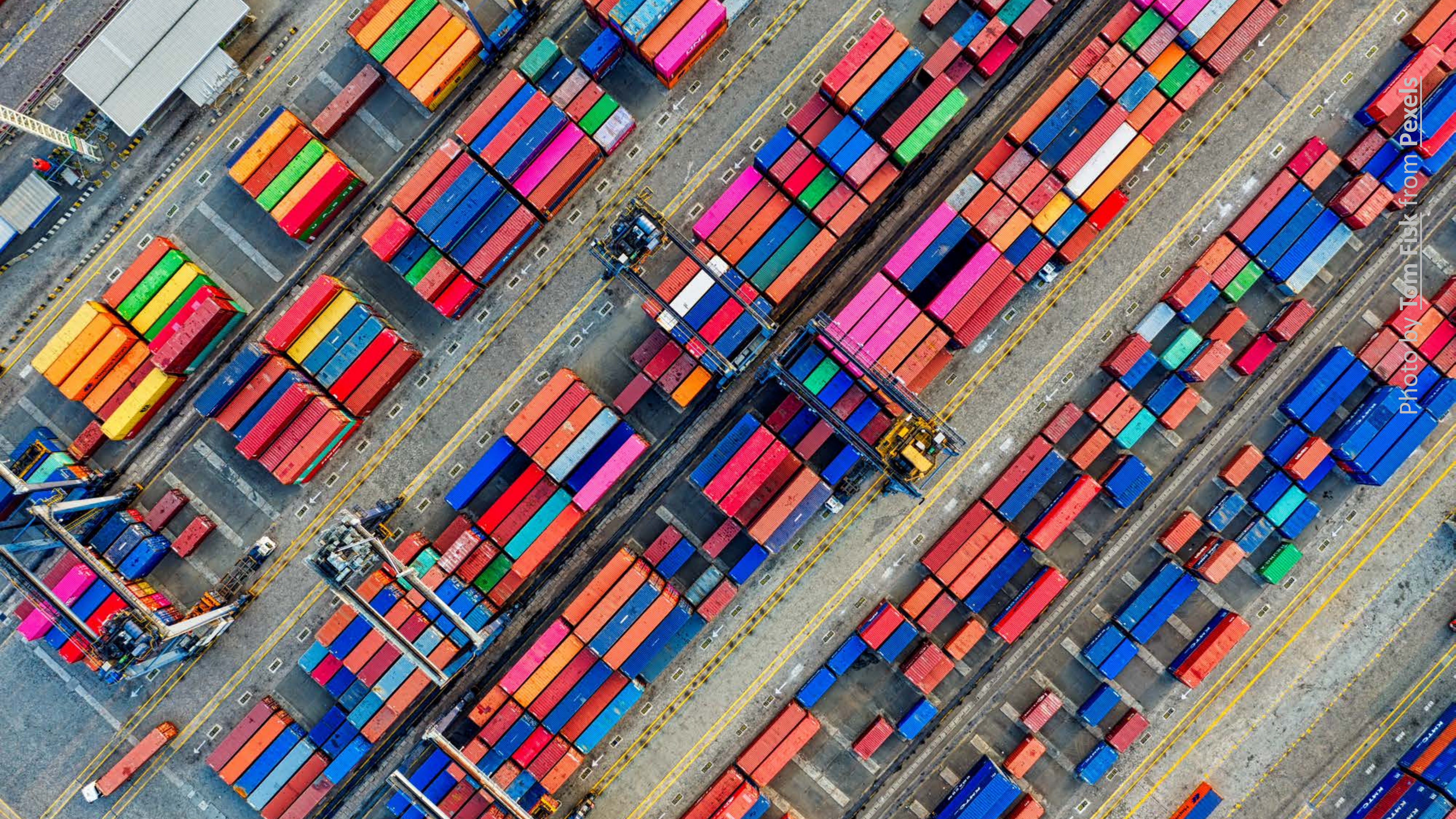


Photo by Tom Fisk from Pexels

threads

- 160 nodes
- 5,120 physical cores
- 10,240 virtual cores

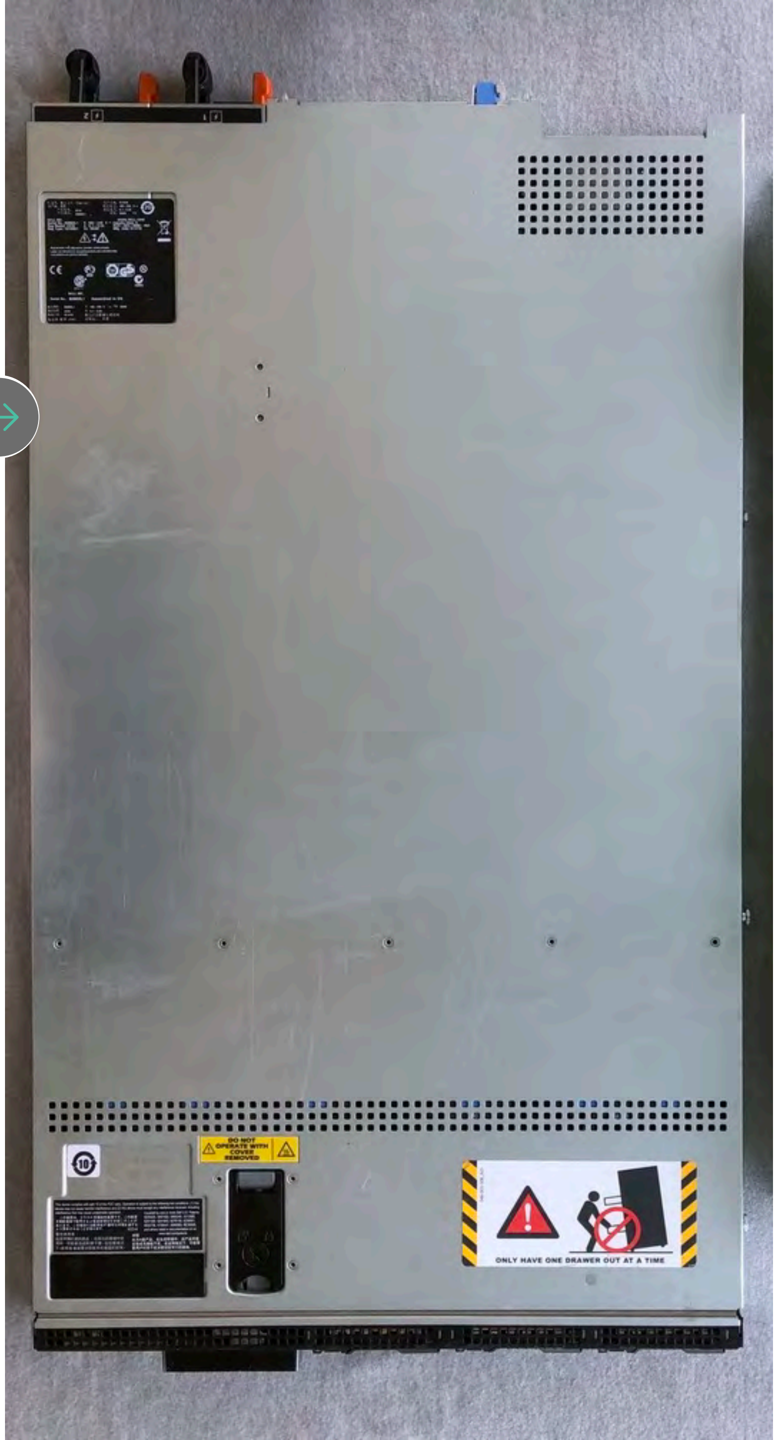
containers

- 160 nodes
- 80 TB (terabytes)
- memory
- 8,000 containers
- to 16,000

servers

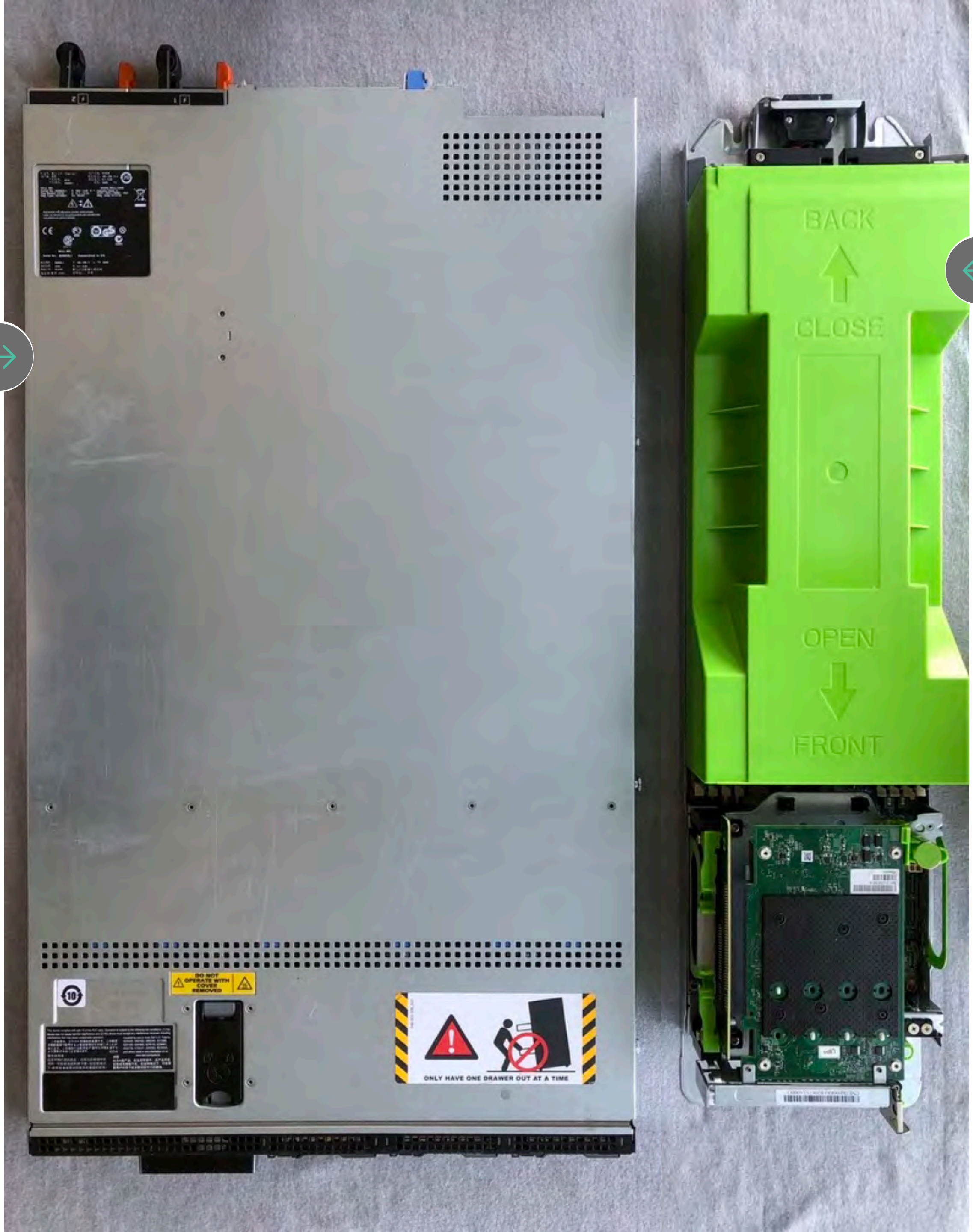
Photo by mali maeder from Pexels

traditional server



traditional server

ocp node



traditional server

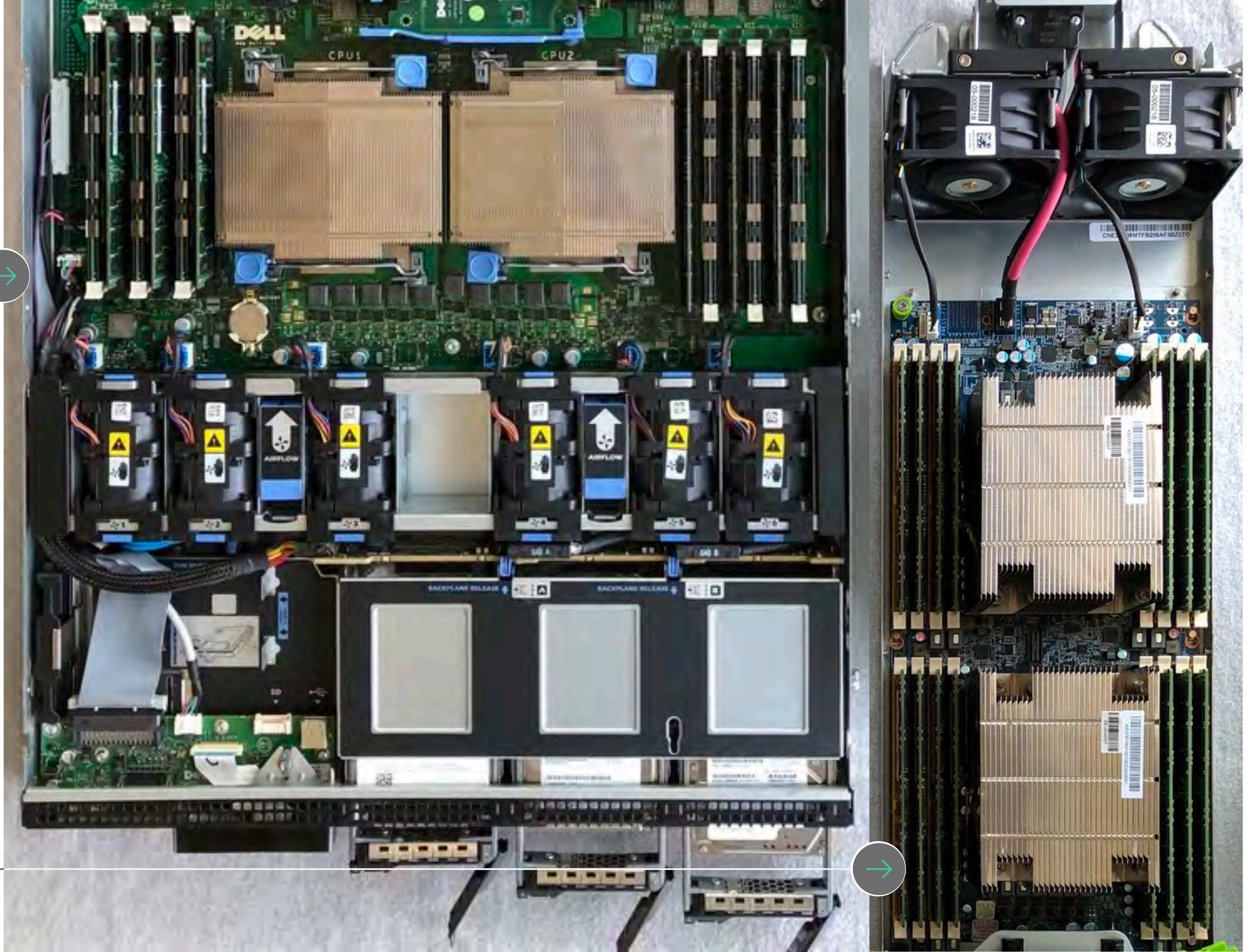
- dual sockets
- two 1U heatsinks
- twelve 1U fans

ocp node

- dual socket
- two 2U heatsinks
- two 2U fans

traditional server

- dual sockets
- two 1U heatsinks
- twelve 1U fans



ocp node

- dual socket
- two 2U heatsinks
- two 2U fans

traditional server

- dual sockets
- two 1U heatsinks
- twelve 1U fans

traditional rack

- 72 cpus
- 72 heatsinks
- 432 fans (1U)

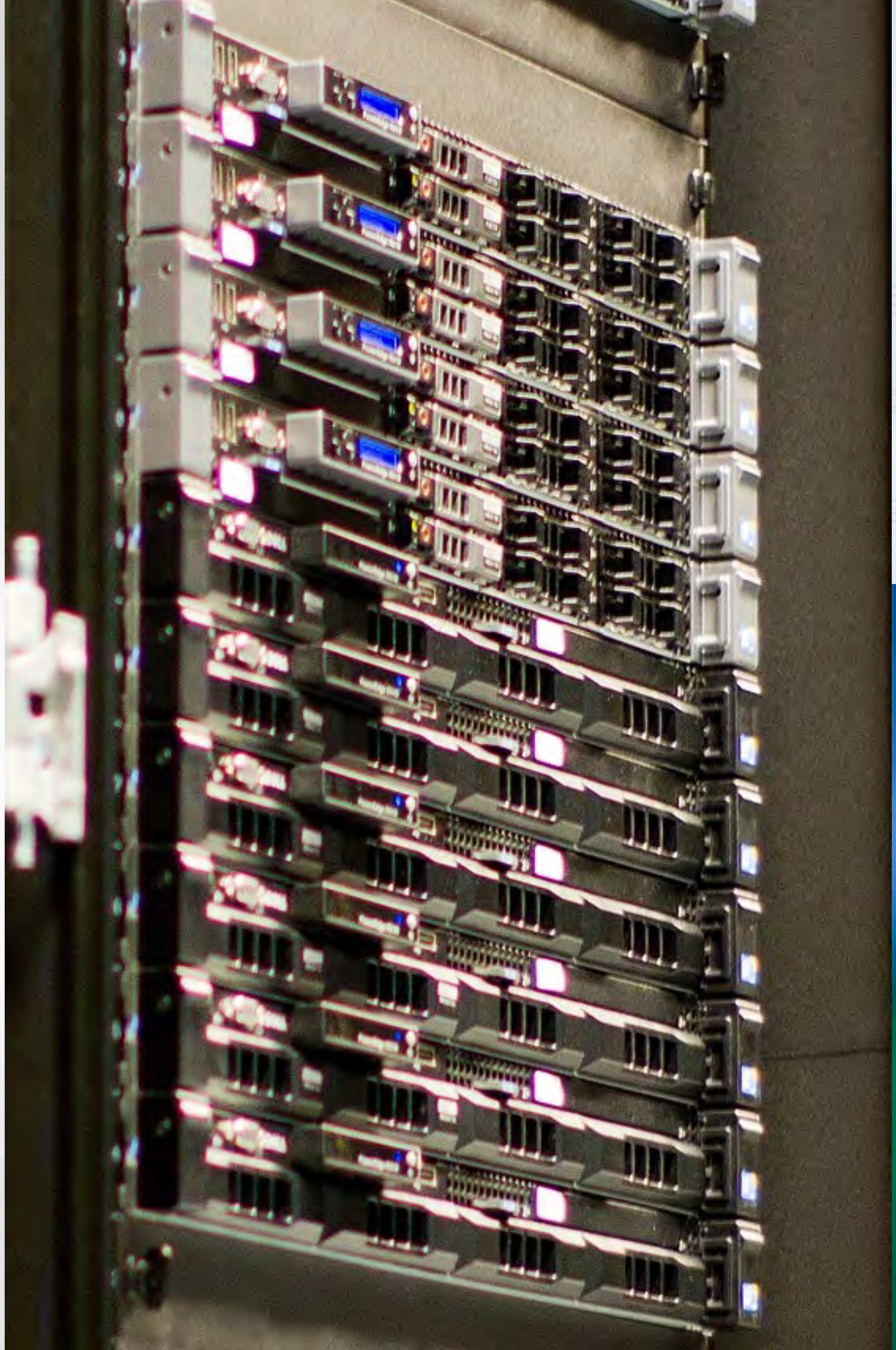
ocp node

- dual socket
- two 2U heatsinks
- two 2U fans

ocp rack

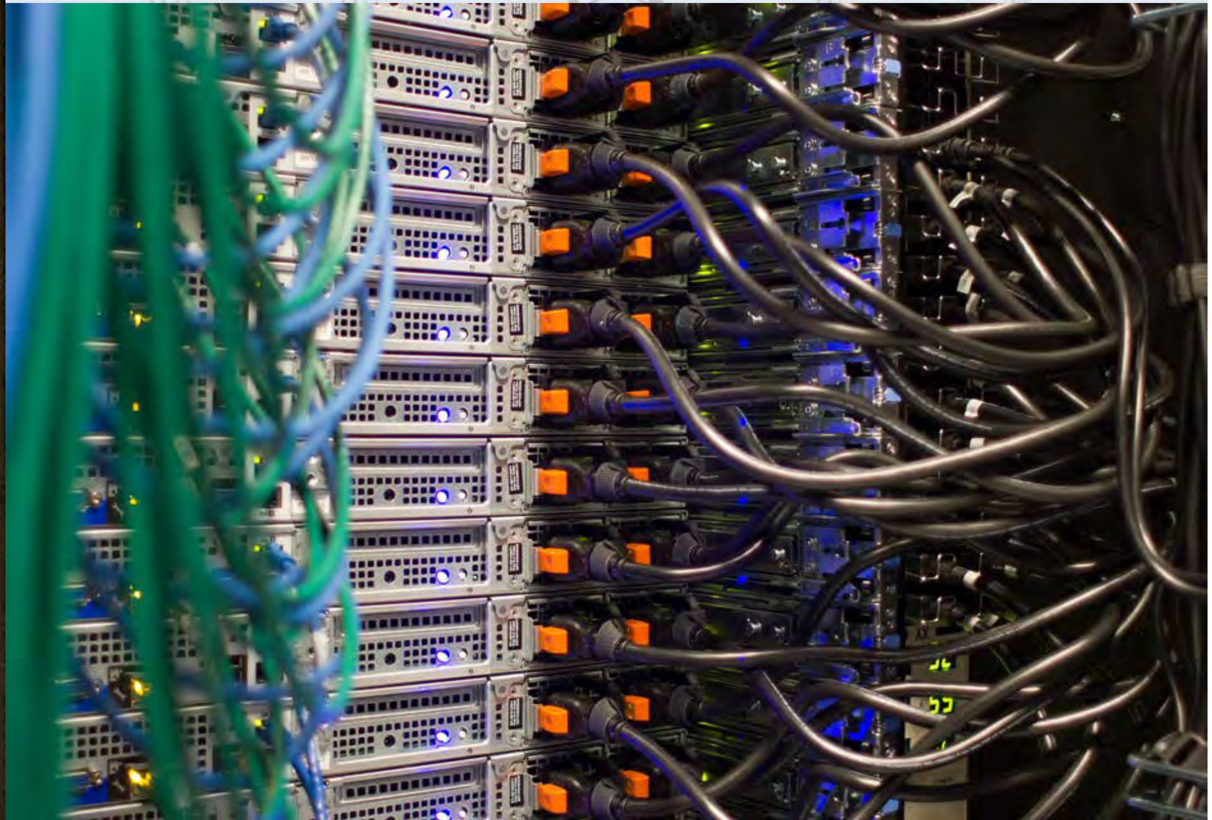
- 96 cpus
- 96 heatsinks
- 96 fans


power
Photo by Polina Tankilevitch from Pexels



traditional rack

- 36 nodes
- 72 cpus
- 72 heatsinks
- 432 fans (1U)

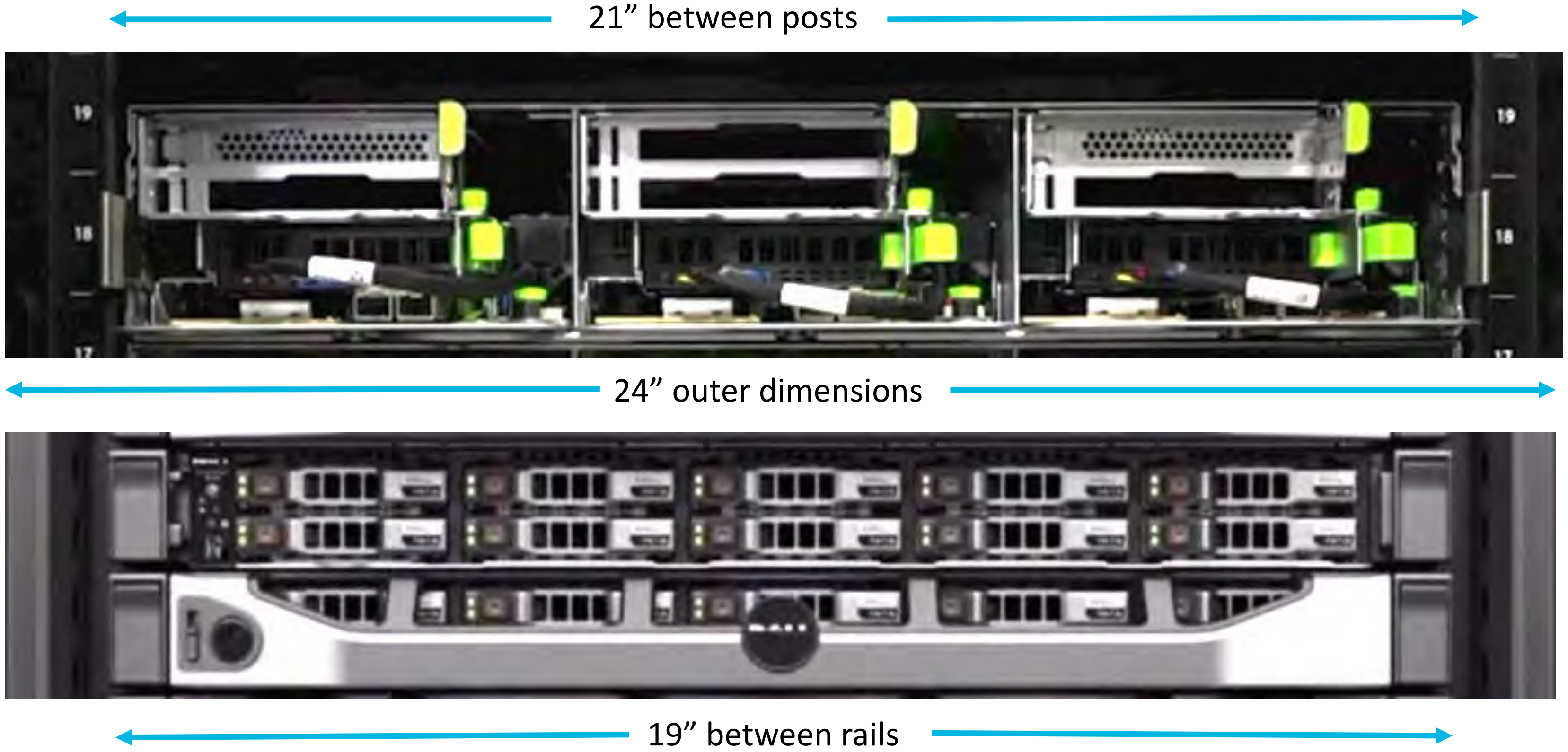


ocp rack

- 48 nodes
- 12 power supplies
- no power cords
- 96 network cables

Open Compute Racks vs Traditional

- all racks are 24" inches wide, outer dimensions (600 mm)
- traditional racks have 19" space between rails. OCP racks have 21" space between the posts
- allows OCP to fit three 2-socket servers instead of two. 50% more servers by optimizing 2" wasted space



OCP is for everyone - no core requirement to redesign data centers or power for OCP racks

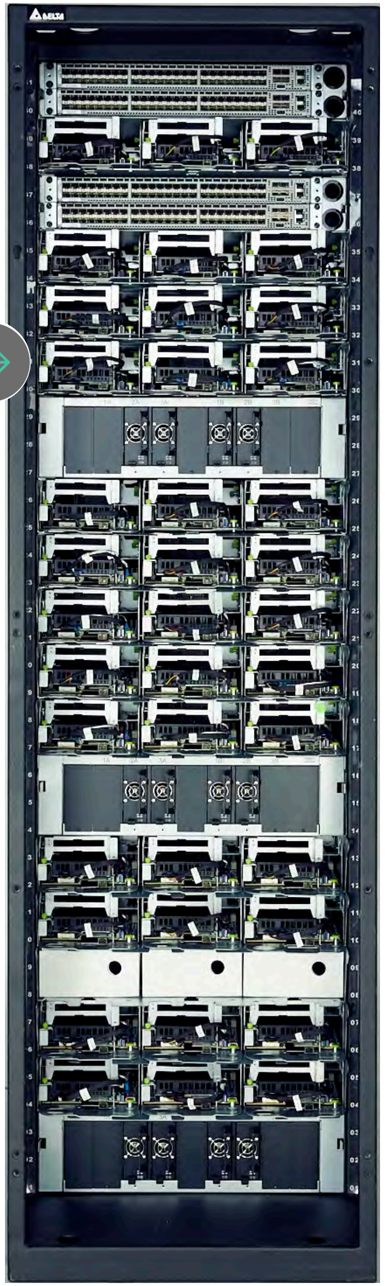
ocp node

- dual socket
- two 2U heatsinks
- two 2U fans

ocp rack

- 96 cpus
- 96 heatsinks
- 12 power supplies
- 96 fans (2U)
- 12 kW

3 nodes / 2 OU

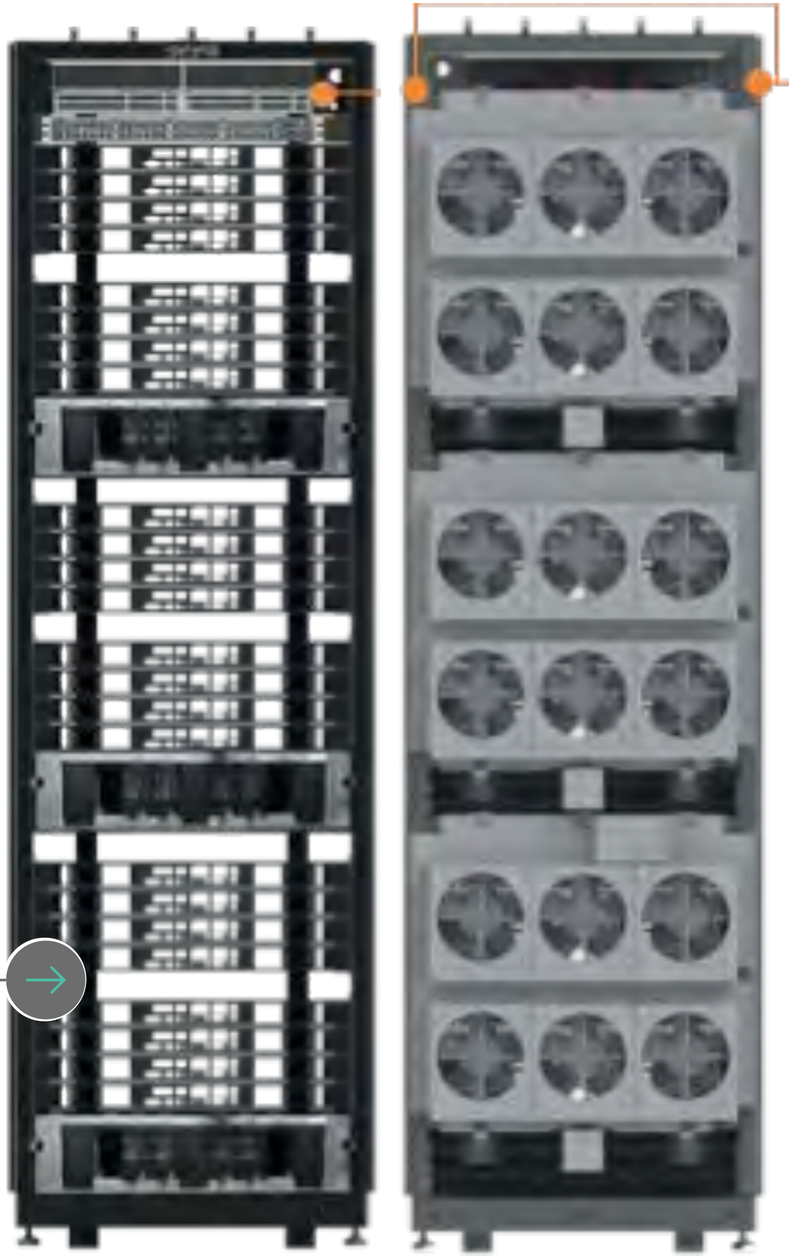


2crsi

3 nodes / 1 OU

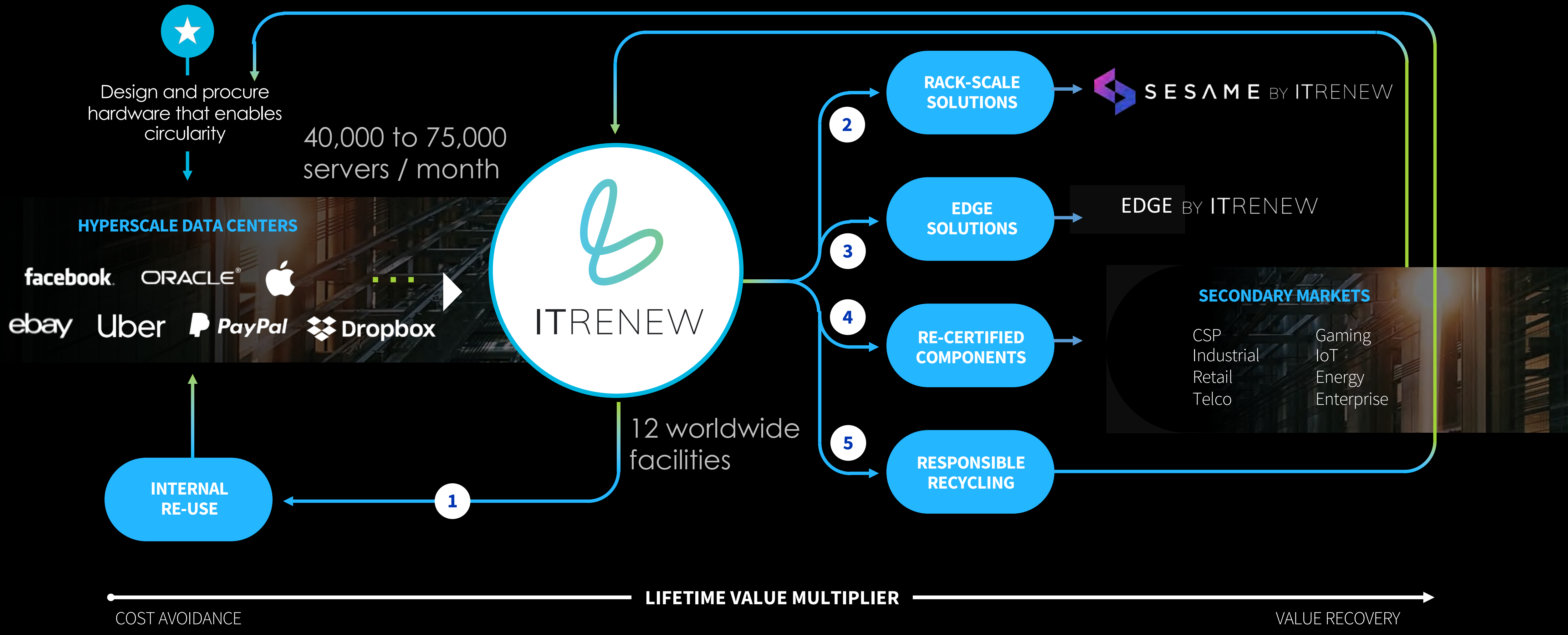
OCtoPus rack

- 180 cpus
- 180 heatsinks
- 18 power supplies
- 23 kW



circular

Circular economy for data center hardware



The circular IT hardware industry opportunity

WHAT IF…

46 million servers

›

31 million tonnes CO_2e

=

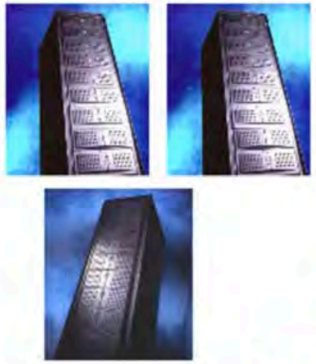
6.7 million cars' annual emissions

progress

Photo by alleksana from Pexels

1999

Network "Appliances" Can Win Today



Dell PowerEdge & PowerVault System

Dell PowerVault 650F	$40,354 x 12 =484,248
512 MB cache, dual link controllers, additional 630F cabinet, 20 x 9 GB FC disks, software support, installation	
Dell PowerEdge 6350	$11,512 x 12 = 138,144
500 MHz PIII, 512 MB RAM, 27 GB disk	
3Com SuperStack II 3800 Switch	7,041
10/100 Ethernet, Layer 3, 24-port	
Rack Space for all that	20,710

NASRaQ System

Cobalt NASRaQ	$1,500 x 240 =360,000	
250 MHz RISC, 32 MB RAM, 2 x 10 GB disks		
Extra Memory (to 128 MB each)	$183 x 360=	65,880
3Com SuperStack II 3800 Switch	$7,041 x 11=	77,451
240/24 = 10 + 1 to connect those 10		
Rack Space (estimate 4x as much as the Dells)		82,840
Installation & Misc		50,000

Comparison

	Dell	Cobalt
Storage	2.1 TB	4.7 TB
Spindles	240	480
Compute	6 GHz	60 GHz
Memory	12.3 GB	30.7 GB
Power	23,122 W	12,098 W
Cost	$650,143	$636,171

From April 1999, Active Disks talk

2019

	Sesame	improvement
Storage	48 TB	10x
Spindles (SSD)	1.6m IOPS	30x
Compute	240 GHz	4x
Memory	24 TB	780x
Power	12,098 W	same
Cost	$96,171	85% less

	Sesame	improvement
Storage (HDD)	9600 TB	2,000x
Spindles (HDD)	130,000 IOPS	2.4x
Power	12,098 W	same

SESAME BY ITRENEW
for Open Systems

SCALE config (48x nodes):
dual 2.5 GHz Xeon, 512 GB RAM, 4x 256GB NVMe disks
Network:
two 32-port 100G, 128-port 25G top-of-rack switches
Rack space (single rack) – deployed in less then 60 minutes

Servers

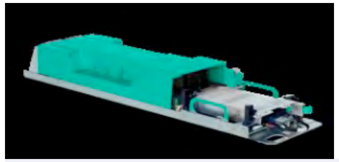
compute
compute
compute
compute
infra
power supply + switch

up to 5 nodes

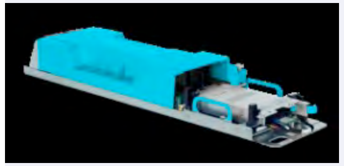
SESAME Fast-Start

external TOR switches (2x)		
ingress	ingress	ingress
internal TOR switches (2x)		
compute	compute	compute
compute	compute	compute
compute	compute	compute
compute	compute	compute
compute	compute	compute
compute	compute	compute
power zone BB		
compute	compute	compute
compute	compute	compute
compute	compute	compute
storage	storage	storage
storage	storage	storage
storage	storage	storage
mgmt	mgmt	mgmt
infra	infra	infra
power zone AA		

up to 48 nodes



single or 2- socket nodes, 25 GbE connectivity



flash-based storage nodes; millions of IOPS and terabytes of capacity

SESAME for Open Systems

data

1,000,000 (million EB)	YOTTABYTES
1,000,000,000 (billion PB)	ZETTABYTES
1,000,000,000 (billion TB)	EXABYTES
1,000,000,000,000,000,000	PETABYTES
1,000,000,000,000,000	TERABYTES
1,000,000,000,000 (trillion)	GIGABYTES
1,000,000,000 (billion)	MEGABYTES
1,000,000 (million)	KILOBYTES
1,000	BYTES

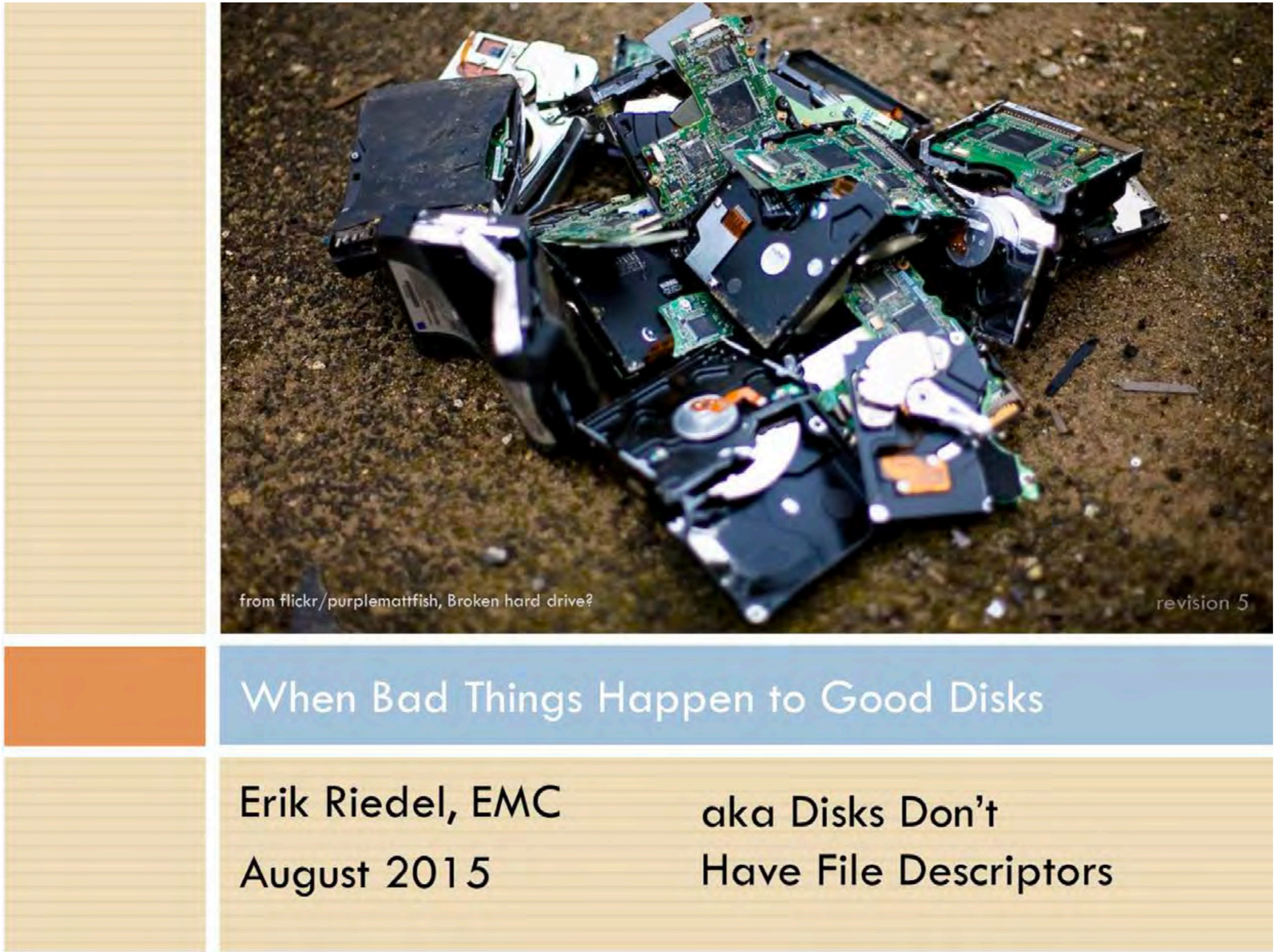
	HELLALOTTABYTES
	LOTTABYTES
1,000,000 (million EB)	YOTTABYTES
1,000,000,000 (billion PB)	ZETTABYTES
1,000,000,000 (billion TB)	EXABYTES
1,000,000,000,000,000,000	PETABYTES
1,000,000,000,000,000	TERABYTES
1,000,000,000,000 (trillion)	GIGABYTES
1,000,000,000 (billion)	MEGABYTES
1,000,000 (million)	KILOBYTES
1,000	BYTES

MYTH

	HELLALOTTABYTES
	LOTTABYTES
1,000,000 (million EB)	YOTTABYTES
1,000,000,000 (billion PB)	ZETTABYTES
1,000,000,000 (billion TB)	EXABYTES
1,000,000,000,000,000,000	PETABYTES
1,000,000,000,000,000	TERABYTES
1,000,000,000,000 (trillion)	GIGABYTES
1,000,000,000 (billion)	MEGABYTES
1,000,000 (million)	KILOBYTES
1,000	BYTES

REALITY

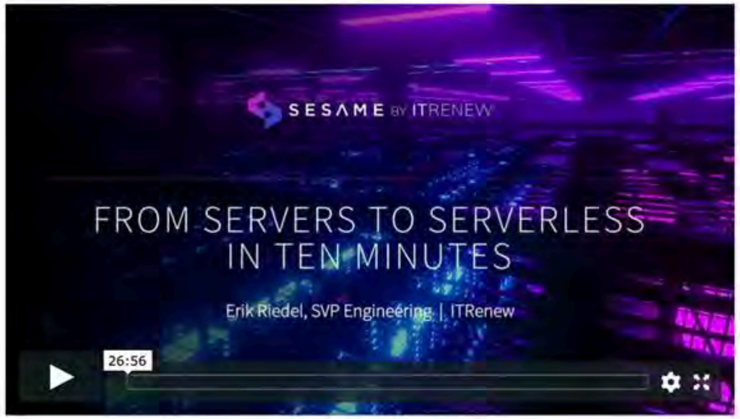
	HELLALOBIBYTES
	LOBIBYTES
1,237,940 (million EB)	YOBIBYTES
1,208,925,819 (billion PB)	ZEBIBYTES
1,180,591,620 (billion TB)	EXBIBYTES
1,152,921,504,606,800,000	PEBIBYTES
1,125,899,906,842,624	TEBIBYTES
1,099,511,627,776 (trillion)	GIBIBYTES
1,073,741,824 (billion)	MEBIBYTES
1,048,576 (million)	KIBIBYTES
1,024	BYTES



https://noti.st/er1p/IPOHXM/when-bad-things-happen-to-good-disks-aka-disks-dont-have-file-descriptors

more stories

For more information: www.itrenew.com/resources



VIDEOS

FROM SERVERS TO SERVERLESS IN TEN MINUTES

Want to deploy and provision a scale-out Kubernetes cluster for running and orchestrating containers & VMs on bare metal in minutes not months? Deploying such clusters on racks of bare metal requires preparation and planning. Learn from ITRenew's evaluation and optimization of various approaches. See how we took a high-density OCP rack with over 1,000 compute cores and 9TB of memory from bare hardware to running a serverless demo app in around 10 minutes.

https://www.itrenew.com/resources/from-servers-to-serverless-in-ten-minutes/

WEBINAR ON-DEMAND

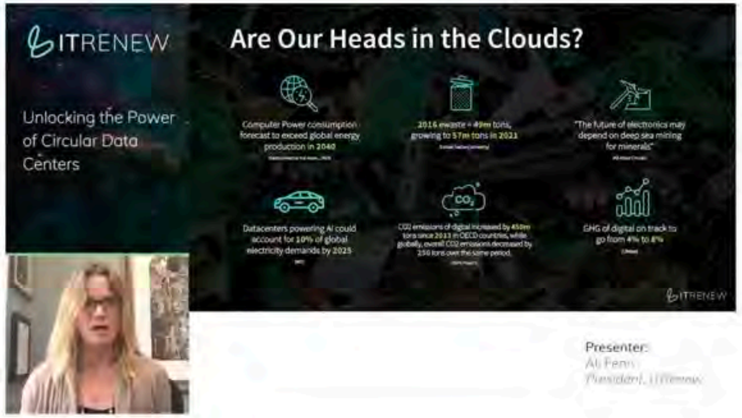
Simultaneous Scaling For User and Application Growth

Is the clock suddenly ticking on your cloud-native and elastic infrastructure initiatives?

Changing market demands and priorities during this global crisis mean businesses can no longer afford to take a multi-year journey to cloud-native. Yet going cloud-native right now means that, overnight, your IT teams must scale capacity up from thousands to millions of users, and scale infrastructure out to support hundreds rather than dozens of apps and workloads. No pressure.

Erik Riedel shares his insights on these trends and addresses why industry leaders worldwide are taking this approach to the multi-dimensional scaling dilemma.

https://www.itrenew.com/webinar-ondemand-full-scaling-for-growth/



VIDEOS

THE TCO OF OCP

The world's largest data center owners leverage open hardware to optimize TCO and refresh cycles, and minimize CO2e impact. Now ITRenew's circular economic model makes the same financial and sustainability opportunity available to broader global markets. Ali Fenn shares the real-world data and is joined by Hydro66 to show the impact of the model in action. Build data centers on the Circular Data Center model to achieve zero waste, lead in energy efficiency, and make a positive impact on the environment – all while lowering your TCO.

Download Slides

https://www.itrenew.com/resources/the-tco-of-ocp/

closing

SESAME BY ITRENEW

QUESTIONS

Call to action

CHECK US OUT ON OUR WEBSITE:

www.itrenew.com/sesame

www.itrenew.com/resources

Watch Video:
Sesame By ITRenew

QUESTIONS OR COMMENTS, REACH US:

@RiedelAtWork

https://github.com/SesameEngineering

ITRENEW

DATA CENTER IMPACT REPORT:
THE FINANCIAL & SUSTAINABILITY CASE FOR CIRCULARITY

Download at itrenew.com