AWS Solutions Architect Associate

Session 402

Networking & CDN: Direct Connect and ENA/EFA on HPC

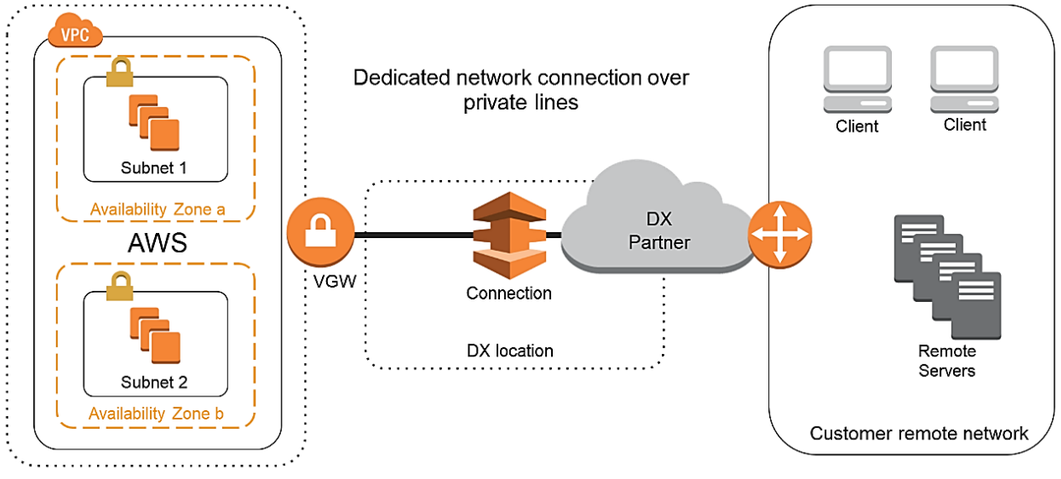
July/2024

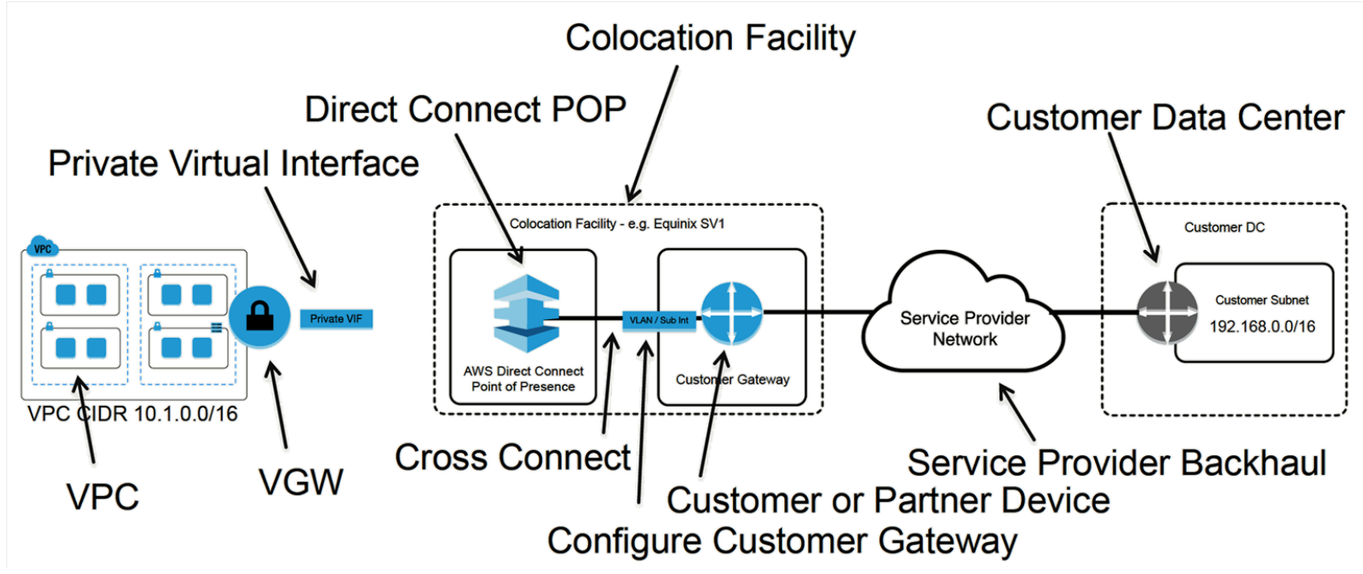
Direct Connect - DxC

Network service that provides an alternative to using the Internet to connect customer's on-premise sites to AWS. No use to connect to Internet. Goes to a Region always.

BENEFITS:
Reduced Bandwidth Costs (* See Pricing: Port and Transfer).
Consistent Network Performance (Latency and Jitter)
AWS Services Compatibility (Integration)
Private Connectivity to AWS VPC (No Encryption)
Elastic (Manual)

COMPONENTS:
Connection
Virtual Interface: Public (i.e. S3), Private (VPC) or Transit (For Transit GW).

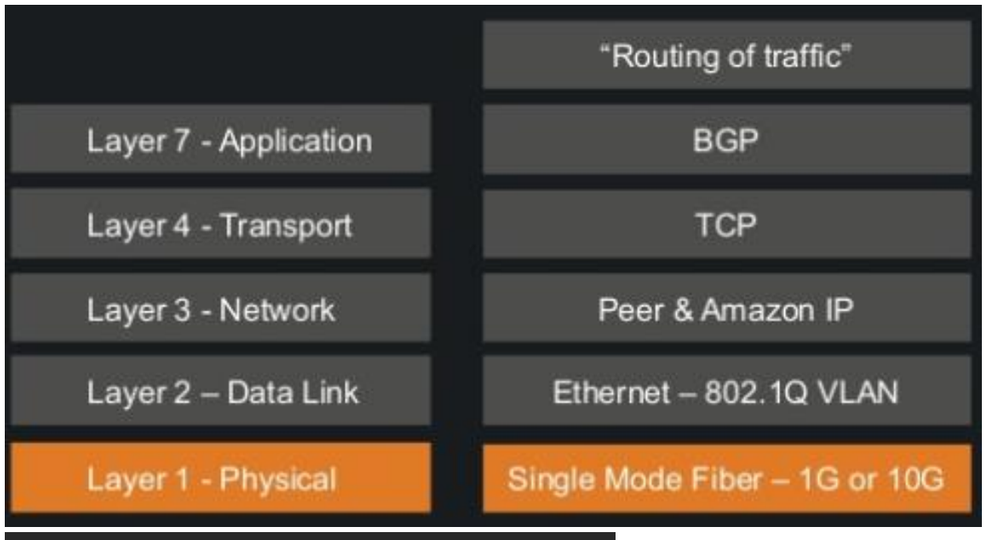


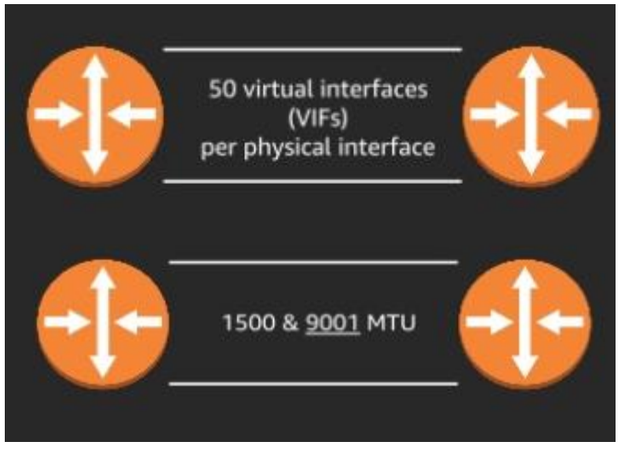


Taken from https://aws.amazon.com/directconnect/faqs/?nc=sn&loc=6 https://jayendrapatil.com/aws-network-connectivity-options/#Direct_Connect https://jayendrapatil.com/aws-direct-connect-dx/ https://aws.amazon.com/premiumsupport/knowledge-center/public-private-interface-dx/ (20/07/2024)

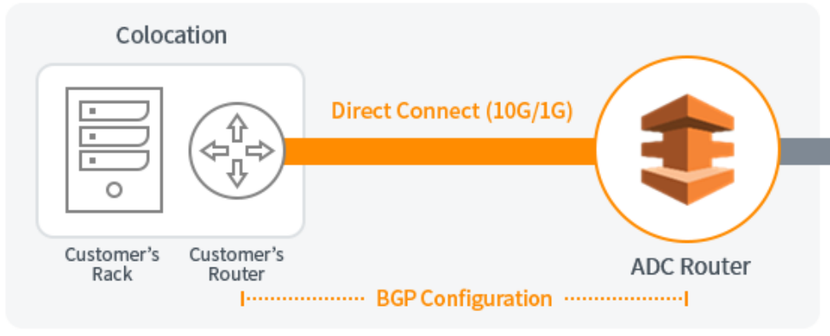
DxC: Features

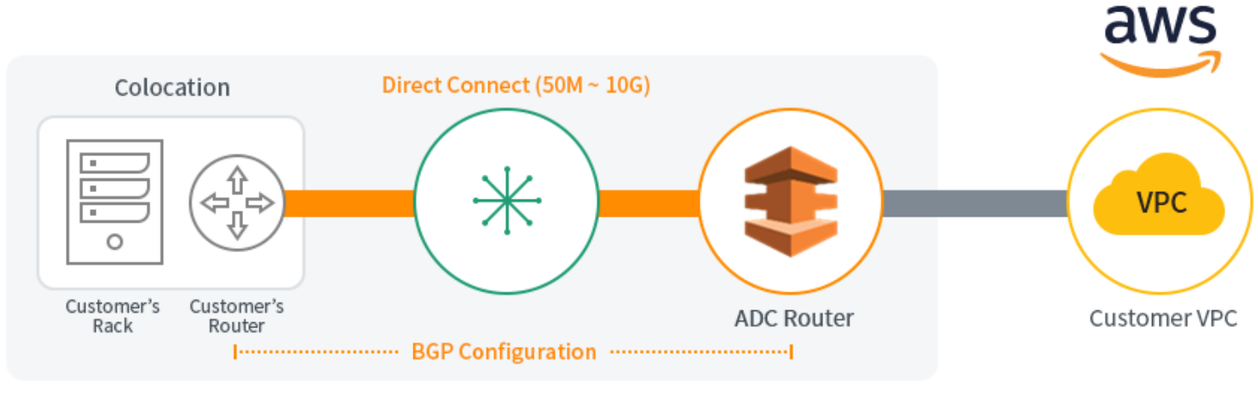
NETWORK REQUIREMENTS (Customer Router):
Single-mode fiber 1000BASE-LX (1310 nm) transceiver for 1 GB Eth or a 10GBASE-LR (1310 nm) transceiver for 10 GB Eth.
Port speed and full-duplex mode must be configured manually (Auto-negotiation must be disabled).
802.1Q VLAN encapsulation must be supported across the entire connection, including intermediate devices.
Your device must support Border Gateway Protocol (BGP) and BGP MD5 authentication.
(Optional) You can configure Bidirectional Forwarding Detection (BFD) on your network. Asynchronous BFD is automatically enabled for AWS Direct Connect virtual interfaces but does not take effect until you configure it on your router.
TYPES:
DEDICATED CONNECTION: 1 or 10GB, LAG upto 40GB.
HOSTED: Trough APN and MPLS Partner. 50 Mbps, 100 Mbps, 200 Mbps, 300 Mbps, 400 Mbps, 500 Mbps, 1 Gbps, 2 Gbps, 5 Gbps and 10 Gbps.





DualStack – IPv4 and IPv6
No In-transit Encryption.
Monitoring with Cloudwatch
AWS Upto 72 h to approve Public VIF





Taken from https://aws.amazon.com/directconnect/faqs/?nc=sn&loc=6
https://jayendrapatil.com/aws-network-connectivity-options/#Direct_Connect (22/07/2024)

DxC: Pricing

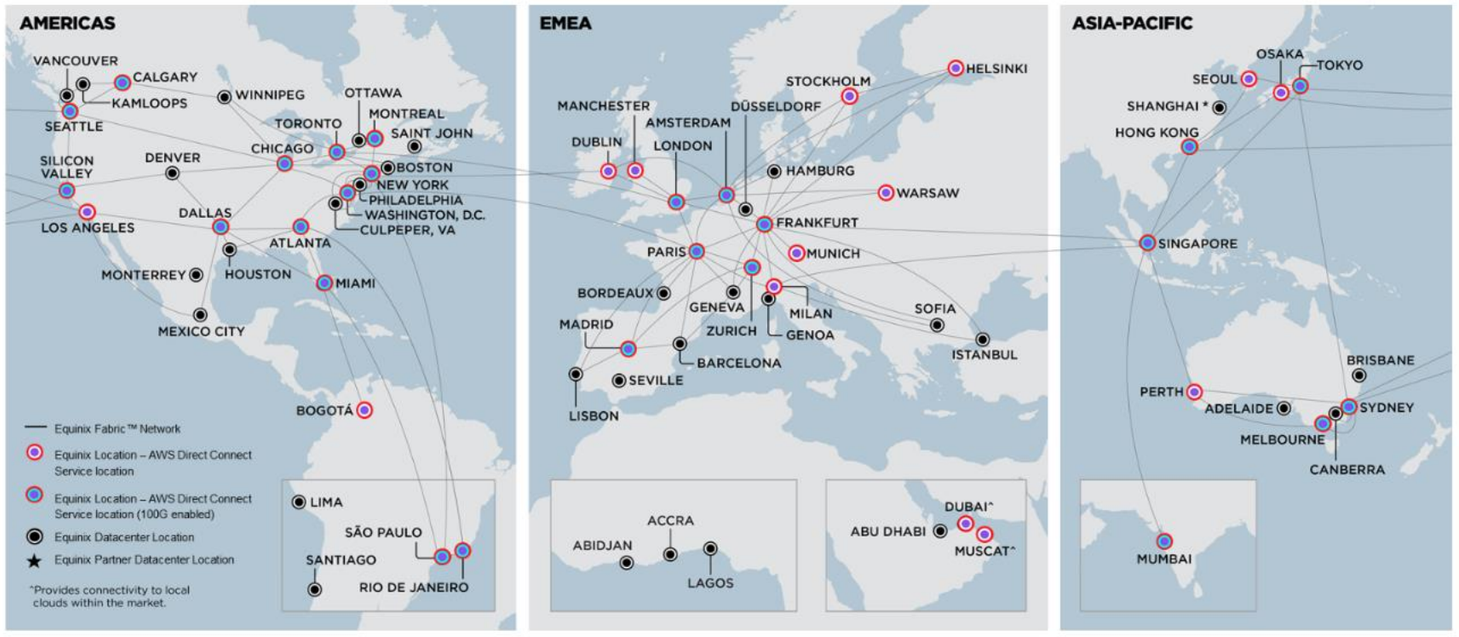
Port Hours: Depend on Bandwith.
Data Transfer IN: Free.
Data Transfer OUT per GB.

Capacity	Port-Hour rate (All AWS Direct Connect locations except in Japan)	Port-hour rate in Japan
1G	$0.30/hour	$0.285/hour
10G	$2.25/hour	$2.142/hour

Capacity	Port-Hour rate (All AWS Direct Connect locations except in Japan)	Port-hour rate in Japan
50M	$0.03/hour	$0.029/hour
100M	$0.06/hour	$0.057/hour
200M	$0.08/hour	$0.076/hour
300M	$0.12/hour	$0.114/hour

Full AWS Direct Connect data transfer pricing

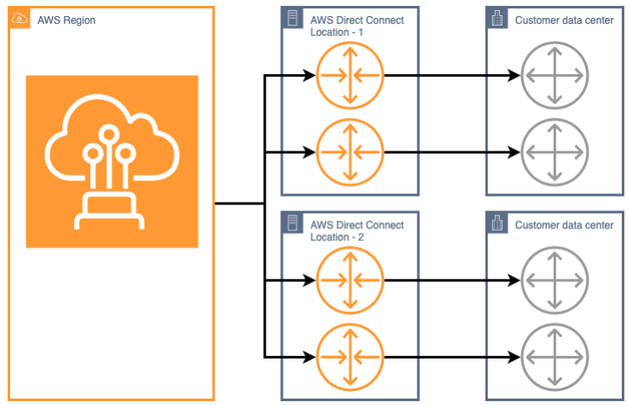
	Data transfer from AWS region									
	US East (Ohio), US East (Virginia), US West (Northern California), US West (Oregon), AWS GovCloud (US-East), AWS GovCloud (US-West)	Canada (Central)	From EU (Frankfurt), EU (Stockholm), EU (Ireland), EU (London), EU (Paris), EU (Milan)	Asia Pacific (Tokyo), Asia Pacific (Osaka-Local)	From Asia Pacific (Seoul), Asia Pacific (Singapore), Asia Pacific (Hong Kong)	Asia Pacific (Mumbai)	South America (Sao Paulo)	Asia Pacific (Sydney)	Middle East (Bahrain)	Africa (Cape Town)
AWS Direct Connect location in:										
United States	$0.0200	$0.0200	$0.0282	$0.0900	$0.0900	$0.0850	$0.1500	$0.1300	$0.1100	$0.1100
Canada	$0.0200	$0.0200	$0.0300	$0.0900	$0.0900	$0.0850	$0.1500	$0.1300	$0.1100	$0.1100
Europe	$0.0200	$0.0300	$0.0200	$0.0600	$0.0900	$0.0850	$0.1107	$0.1300	$0.1000	$0.1100
Japan	$0.0491	$0.0500	$0.0600	$0.0410	$0.0420	$0.1132	$0.1700	$0.1132	$0.1500	$0.1700
Hong Kong SAR, Malaysia, S.Korea, Singapore & Taiwan	$0.0491	$0.0500	$0.0600	$0.0420	$0.0410	$0.10				
India	$0.0600	$0.0600	$0.0625	$0.1132	$0.1107	$0.04				
South America	$0.1107	$0.1107	$0.1107	$0.1700	$0.1107	$0.16				



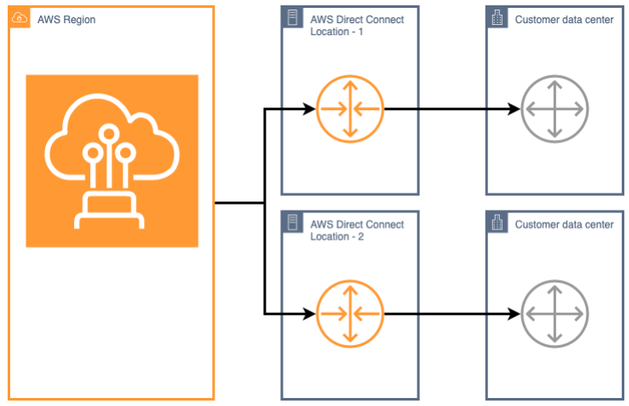
Taken from https://aws.amazon.com/directconnect/pricing/?nc=sn&loc=3 and https://www.equinix.com/partners/aws (20/07/2024)

DxC: Resiliency

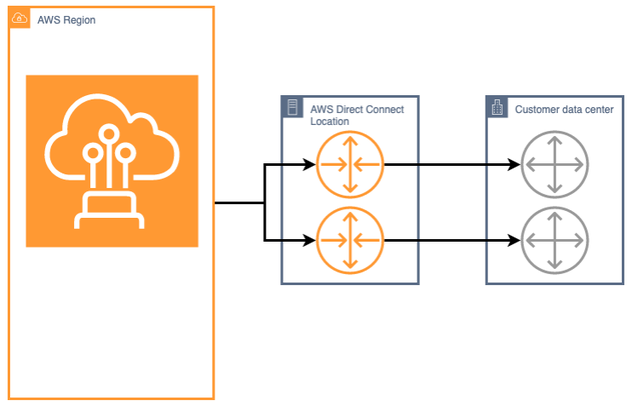
• Maximum Resiliency: This model provides you a way to order dedicated connections to achieve an SLA of 99.99%.



• High Resiliency: This model provides you a way to order dedicated connections to achieve an SLA of 99.9%.



• Development and Test: This model provides you a way to achieve development and test resiliency for non-critical workloads, by using separate connections that terminate on separate devices in one location 99,7%.



• Classic. This model is intended for users that have existing connections and want to add additional connections. This model does not provide an SLA.

Taken from https://docs.aws.amazon.com/directconnect/latest/UserGuide and https://docs.aws.amazon.com/directconnect/latest/UserGuide/resiliency_toolkit.html (20/07/2024)

DxC: Use Case & VPN Comparison

For massive data transmission

· In need of Big Data analysis
· In need of regular back-ups

For fast and reliable data transmission

· In need of Hybrid Cloud (private Cloud+ AWS)
· In need of frequent connections between On-premise resources and AWS

For network cost saving

· In need of cost-saving in AWS transfer price
· In need of frequent transferring between customer service resources and AWS

DxC vs VPN

Site-2-Site

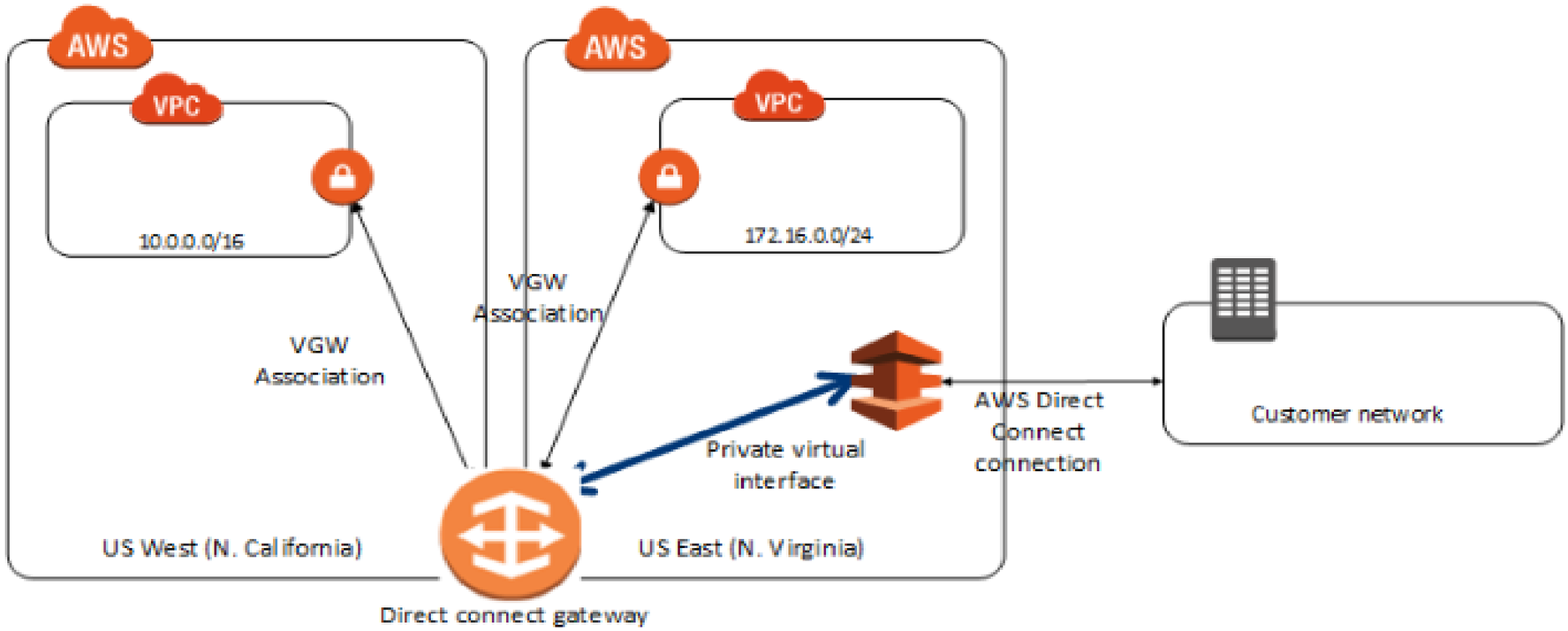
	AWS-Managed VPN	AWS Direct Connect
Performance	<4 GB per VPC	<1 GB, 1 GB, or 10 GB ports Up to 40 GB with Link Aggregation Group (LAG)
Connectivity	1VPN Connection to VPC	2 port connection to multiple VPCs
Resiliency	1 VPN Connection = 2 VPN tunnels	1 AWS router = redundant connectivity to 1 AWS region
Costs	$0.05 per VPN Connection Hour $0.09 per GB data transfer out	$0.2 to $0.3 per GB data transfer out Port hour fees(varies based on port speed)

Taken from https://www.coresite.com/blog/vpn-or-direct-connect-aws-compared (29/11/2020) and more info at https://docs.aws.amazon.com/vpn/latest/s2svpn/VPC_VPN.html (18/07/2024)

DxC Gateway

Direct Connect Gateway allows you to **connect an AWS Direct Connect connection to one or more VPCs in your account that are located in the same or different regions**
Direct Connect gateway can be created in any public region and accessed from all other public regions
Direct Connect gateway **CANNOT** be used to connect to a VPC in another account.
Alternatively, Direct connect locations can also access the public resources in any AWS Region using a public virtual interface.

Amazon Virtual Private Cloud Connectivity Options
AWS Whitepaper

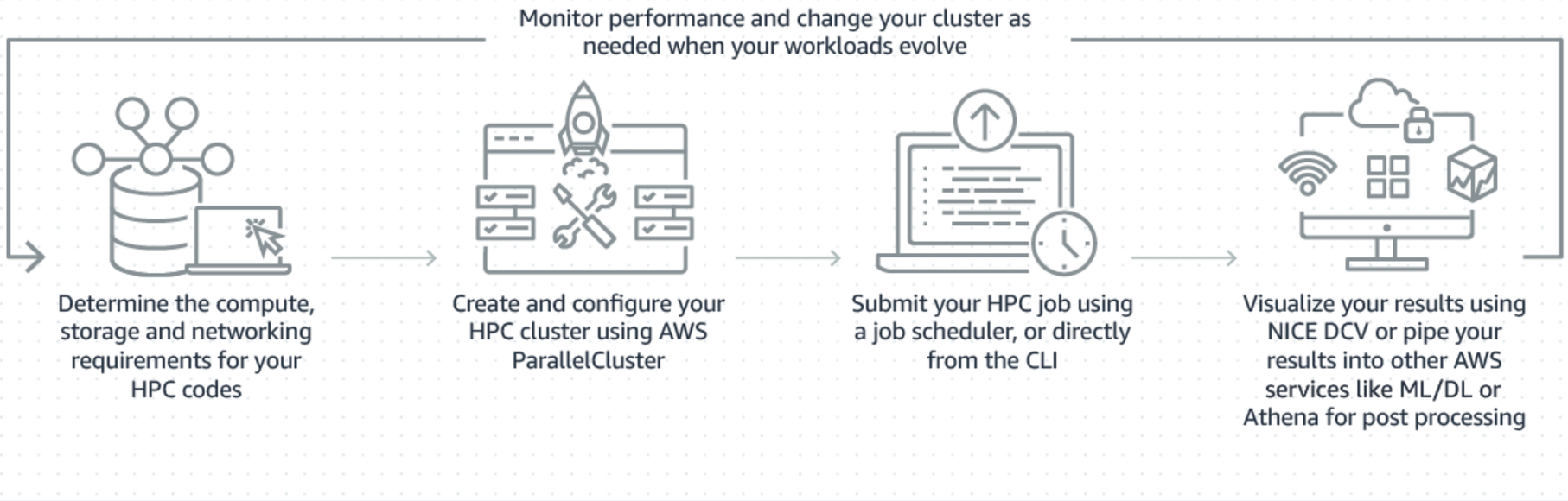


Taken from https://docs.aws.amazon.com/directconnect/latest/UserGuide/direct-connect-gateways-intro.html and https://jayendrapatil.com/aws-network-connectivity-options/#Direct_Connect (20/07/2024)

MANDATORY READ **Amazon Virtual Private Cloud Connectivity Options** at https://docs.aws.amazon.com/whitepapers/latest/aws-vpc-connectivity-options/network-to-amazon-vpc-connectivity-options.html (18/07/2024)

HPC Needs

High-Performance Computing cover a high-demanding needs for ML, DL and other optimized-computing algorithms (Fluid Dynamics, Genetics simulation, Weather, etc) to fulfill in a faster environment for PoC or Production.



Create an Elastic HPC Cluster using AWS ParallelCluster

Access on-demand, scalable resources for your High-Performance Computing (HPC) workloads.

Learn more »

Create an End-to-End HPC Environment

Launch a fully-functional HPC environment, create a sample job submission service, and review important concepts along the way.

Learn more »

Getting Started with Amazon FSx for Lustre

It is easy to sign up for AWS and start using FSx for Lustre to run your compute-intensive workloads on AWS. Here are the simple steps to get started.

Learn more »

Getting started with Elastic Fabric Adapter (EFA)

In this tutorial, you create an EFA-enabled AMI and an EFA-enabled security group, and then launch EFA-enabled instances into a cluster placement group using that AMI and security group

Learn more »

Taken from https://aws.amazon.com/hpc/ (18/07/2024) More information at Best Practices for Running HPC on AWS (https://www.iucc.ac.il/en/blog/best-practices-for-running-hpc-on-aws/) and Optimal Prices for HPC Workloads (https://d1.awsstatic.com/HPC2019/Optimal-Price-Performance-for-HPC-Workloads-Aug2019.pdf) (29/05/2021)

HPC Associated Services

Services to support HPC

Data Management & Transfer

- AWS DataSync
- AWS Snowball
- AWS Snowmobile
- AWS DirectConnect

Compute & Networking

- Amazon EC2 instances (CPU, GPU, FPGA)
- Amazon EC2 Spot
- AWS Auto Scaling
- Placement groups
- Enhanced networking
- Elastic Fabric Adapter
- AWS VPC

Storage

- Amazon EBS with provisioned IOPS
- Amazon FSx for Lustre
- Amazon EFS
- Amazon S3

Automation & Orchestration

- AWS Batch
- AWS ParallelCluster
- NICE EnginFrame

Visualization

- NICE DCV
- Amazon AppStream 2.0

Operations and Management

- Amazon CloudWatch
- AWS Budgets

Security & Compliance

- AWS Identity and Access Managment

AWS high performance computing services

Amazon EC2

Amazon EC2 Hpc7g, Hpc7a, and Hpc6id are HPC-optimized instances purpose built for running HPC workloads at scale on AWS.

Elastic Fabric Adapter

Run HPC applications at scale with Elastic Fabric Adapter (EFA), a network for Amazon EC2 instances with high-level inter-node communications capabilities.

Research and Engineering Studio on AWS

Easily access, manage, and monitor cloud-based research and engineering environments.

AWS ParallelCluster

Quickly build HPC compute environments with AWS ParallelCluster, an open-source tool that simplifies deployment and management of HPC clusters.

AWS Batch

Scale hundreds of thousands of computing jobs across all AWS compute services and features with AWS Batch, a cloud-native batch scheduler.

NICE DCV

Deliver a high performance remote desktop and 3D application graphics with NICE DCV, a bandwidth-efficient and high performance streaming protocol.

Amazon FSx for Lustre

Quickly process massive datasets on-demand and at scale with Amazon FSx for Lustre, a high performance file system with sub-millisecond latencies.

Amazon FSx for OpenZFS

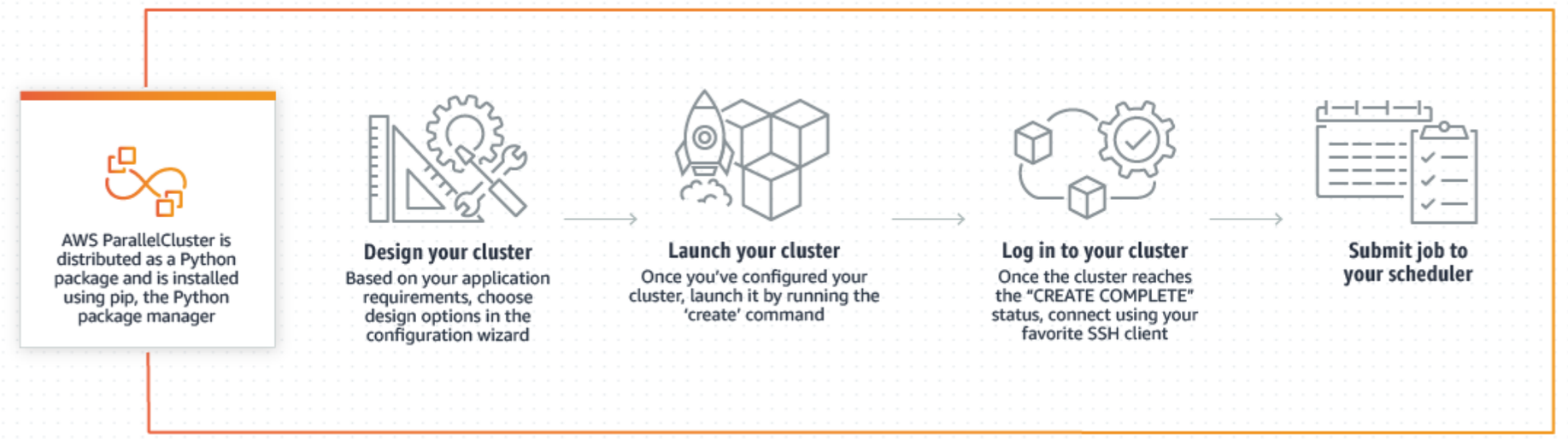
Fully managed shared file storage built on the OpenZFS file system, powered by the AWS Graviton family of processors.

Taken from https://aws.amazon.com/hpc/ (05/10/2020)

AWS Parallel Cluster

Deliver a EC2 cluster for HPC (Open Source Cluster Mgmt Tool). Coming from CfnCluster (Cloudformation Cluster-Automate Net, EC2,) Syntax and evolved. It has integrated services: AWS Batch, Multiple EBS, AutoScaling. In addition, It can be accessed via Private Proxy.
Example time to create a cluster: 8m 33s.
2 schedulers: slurm and awsbatch.
Parallel Cluster don't have a fee, because it looks like a framework for CFn. The cost is associated to required resources.

¿ Parallel Cluster and EC2 Cluster Placement Group ?



Getting started with AWS ParallelCluster

1. Install AWS ParallelCluster
2. Design and launch your cluster
3. Log in to your cluster
4. Submit your job to your scheduler

Taken from https://aws.amazon.com/hpc/parallelcluster/ and https://d1.awsstatic.com/Projects/P4114756/deploy-elastic-hpc-cluster_project.pdf (05/10/2020)
More information at AWS Parallel Cluster (https://aws.amazon.com/hpc/parallelcluster/) Deploy HPC Cluster using Parallel Cluster (https://d1.awsstatic.com/Projects/P4114756/deploy-elastic-hpc-cluster_project.pdf) and HPC Cluster (https://aws.amazon.com/hpc) (29/05/2021)

Network Cards

Virtual network card - 1Gbps

ENI - Elastic Network Interface

Network Cards

Enhanced Networking

- Through Single Root Virtualization IO (SRV-IO)
- All instancia except T2
- Options
  - Elastic Network Adapter - 100Gbps
  - Intel 82599 Virtual Function (VF)- 10Gbps - Old Instances
  - Elastic Fabric Adapter - ENA plus OS-bypass

Use case: [ENA] HPC or DB Workloads, Video Processing. [EFA] HPC, CFD, Weather and ML Apps.

Both as base has a ENI.

ENA on EC2 Instances

Why: Avoid bottleneck for workloads

Advantages: Automatic scaling without modifications due to BW and vCPU count.

Tech: SRV IO, which increase BW, low latency, I/O performance, low CPU utilization and increase Packets per Second.

No additional cost.

ENA for the most instances type, however there are some old instance type that are support by Intel 82599 VF only, even though there are oldest instance type with no option of Enhanced Networking.

- Amazon Linux 2
- Amazon Linux AMI 2018.03
- Ubuntu 14.04 (with `linux-aws` kernel) or later
- Red Hat Enterprise Linux 7.4 or later
- SUSE Linux Enterprise Server 12 SP2 or later
- CentOS 7.4.1708 or later
- FreeBSD 11.1 or later
- Debian GNU/Linux 9 or later

Windows Server 2012 R2, 2016

	EBS only	NVMe EBS	Instance store	Placement group	Enhanced networking
Current generation instances					
• **General purpose:** M5 | M5a | M5ad | M5d | M5dn | M5n | M5zn | M6a | M6g | M6gd | M6i | M6id | M6idn | M6in | M7a | M7g | M7gd | M7i | M7i-flex | Mac1 | Mac2 | Mac2-m1ultra | Mac2-m2 | Mac2-m2pro | T2 | T3 | T3a | T4g					
• **Compute optimized:** C5 | C5a | C5ad | C5d | C5n | C6a | C6g | C6gd | C6gn | C6i | C6id | C6in | C7a | C7g | C7gd | C7gn | C7i | C7i-flex					
• **Memory optimized:** R5 | R5a | R5ad | R5b | R5d | R5dn | R5n | R6a | R6g | R6gd | R6i | R6idn | R6in | R6id | R7a | R7g | R7gd | R7i | R7iz | R8g | U-3tb1 | U-6tb1 | U-9tb1 | U-12tb1 | U-18tb1 | U-24tb1 | U7i-12tb | U7in-16tb | U7in-24tb | U7in-32tb | X1 | X2gd | X2idn | X2iedn | X2iezn | X1e | z1d					
• **Storage optimized:** D2 | D3 | D3en | H1 | I3 | I3en | I4g | I4i | Im4gn | Is4gen					
• **Accelerated computing:** DL1 | DL2q | F1 | G4ad | G4dn | G5 | G5g | G6 | Gr6 | Inf1 | Inf2 | P2 | P3 | P3dn | P4d | P4de | P5 | Trn1 | Trn1n | VT1					
• **High-performance computing:** Hpc6a | Hpc6id | Hpc7a | Hpc7g					
Previous generation instances					
• **General purpose**: A1 | M1 | M2 | M3 | M4 | T1					
• **Compute optimized**: C1 | C3 | C4					
• **Memory optimized**: R3 | R4					
• **Storage optimized**: I2					
• **Accelerated computing**: G3					
M5	Yes	Yes	No	Yes	ENA
M5a	Yes	Yes	No	Yes	ENA

	Instance store	Placement group	Enhanced networking
C3	SSD	Yes	Intel 82599 VF
G2	SSD	Yes	No
I2	SSD	Yes	Intel 82599 VF
M3	SSD	No	No
R3	SSD	Yes	Intel 82599 VF

Taken from https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/instance-types.html#instance-type-summary-table (18/07/2024)

How to use ENA

Supported Instance type (For Intel 82599 VF only for HVM).

Supported OS.

Internet Connectivity.

If you don’t able at launch the instance, you can use AWS CLI or Powershell Tools for AWS to activate ONLY.

The max speed depends on instance type.

Current generation instances

- **General purpose:** M5 | M5a | M5ad | M5d | M5dn | M5n | M5zn | M6a | M6g | M6gd | M6i | M6id | M6idn | M6in | M7a | M7g | M7gd | M7i | M7i-flex | Mac1 | Mac2 | Mac2-m1ultra | Mac2-m2 | Mac2-m2pro | T2 | T3 | T3a | T4g
- **Compute optimized:** C5 | C5a | C5ad | C5d | C5n | C6a | C6g | C6gd | C6gn | C6i | C6id | C6in | C7a | C7g | C7gd | C7gn | C7i | C7i-flex
- **Memory optimized:** R5 | R5a | R5ad | R5b | R5d | R5dn | R5n | R6a | R6g | R6gd | R6i | R6idn | R6in | R6id | R7a | R7g | R7gd | R7i | R7iz | R8g | U-3tb1 | U-6tb1 | U-9tb1 | U-12tb1 | U-18tb1 | U-24tb1 | U7i-12tb | U7in-16tb | U7in-24tb | U7in-32tb | X1 | X2gd | X2idn | X2iedn | X2iezn | X1e | z1d
- **Storage optimized:** D2 | D3 | D3en | H1 | I3 | I3en | I4g | I4i | Im4gn | Is4gen
- **Accelerated computing:** DL1 | DL2q | F1 | G4ad | G4dn | G5 | G5g | G6 | Gr6 | Inf1 | Inf2 | P2 | P3 | P3dn | P4d | P4de | P5 | Trn1 | Trn1n | VT1
- **High-performance computing:** Hpc6a | Hpc6id | Hpc7a | Hpc7g

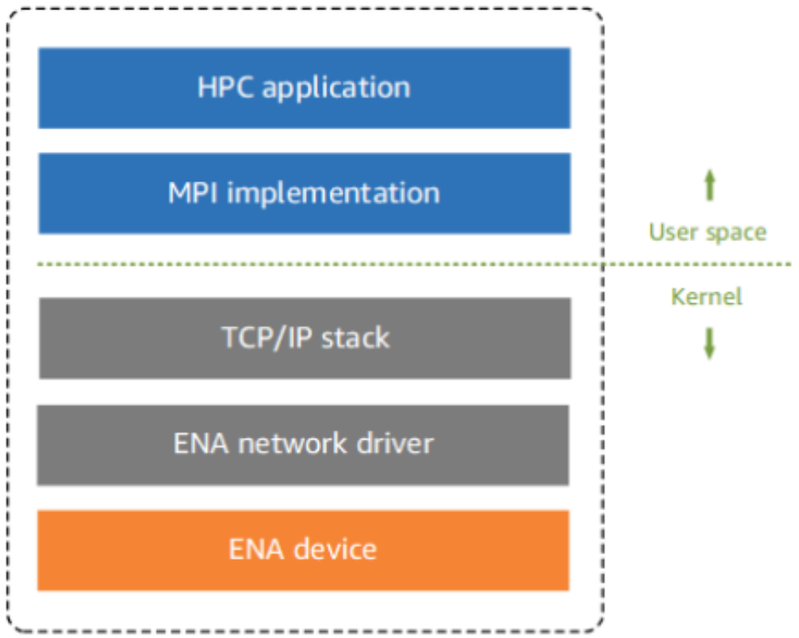
Previous generation instances

- **General purpose**: A1 | M1 | M2 | M3 | M4 | T1
- **Compute optimized**: C1 | C3 | C4
- **Memory optimized**: R3 | R4
- **Storage optimized**: I2
- **Accelerated computing**: G3

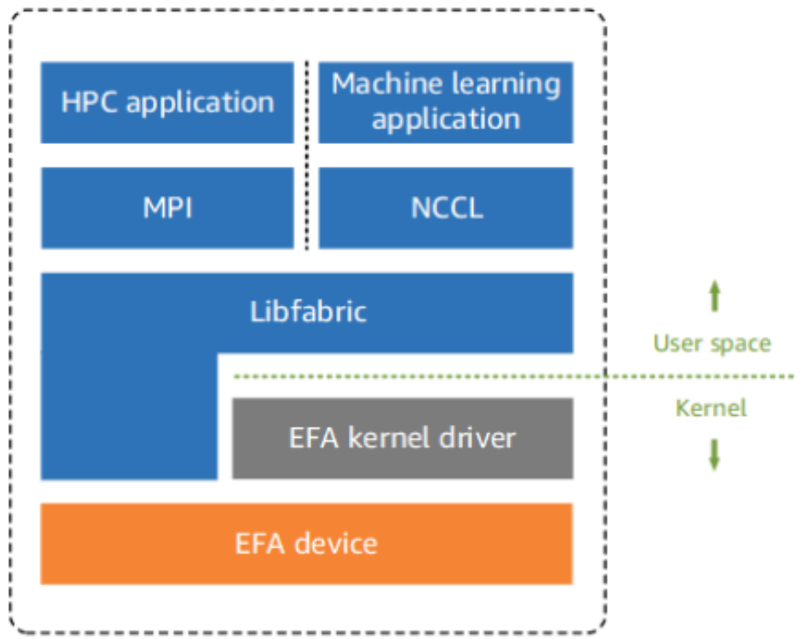
Instance type	Network performance	Enhanced networking
t2.nano | t2.micro | t2.s t3.nano | t3.micro | t3.s t3a.2xlarge | t4g.nano |	Up to 1 Gbps	Not supported
m4.large	Up to 5 Gbps †	ENA
m4.xlarge | m4.2xlarge |	Moderate	Intel 82599 VF
a1.4xlarge and smaller | m6gd.4xlarge and smaller	High	Intel 82599 VF
m4.10xlarge	Up to 10 Gbps †	ENA
m5.8xlarge | m5a.12xlar	10 Gbps	Intel 82599 VF
m5.12xlarge | m5a.16xla	10 Gbps	ENA
m5.16xlarge | m5a.24xla	12 Gbps	ENA
m5dn.4xlarge and smaller	20 Gbps	ENA
m4.16xlarge | m5.24xlar	Up to 25 Gbps †	ENA
m5dn.12xlarge | m5n.12x	25 Gbps	ENA
m5dn.16xlarge | m5n.16x	50 Gbps	ENA
m5dn.24xlarge | m5n.24x	75 Gbps	ENA
	100 Gbps	ENA

Taken from https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/general-purpose-instances.html (05/10/2020)

EFA on Linux Instances



Traditional HPC software stack in EC2



HPC software stack in EC2 with EFA

Advantages: Higher Throughput.

Tech: Low and uniform latency and high throughput than TCP/IP Stack, increase inter-instance communication.

No additional cost.

Modes: Normal IP (ENA) and OS-Bypass.

Some instance type and Linux OS are supported only.

```
aws ec2 describe-instance-types  --region us-east-1  --filters
Name=network-info.efa-supported,Values=true  --query
"InstanceTypes[*].[InstanceType]"  --output text | sort
```

MPI: Message Passing Interface, to support Parallel Programming.
Libfabric: Library to support new datagrams.

Support interface and libraries:
•Open MPI 5 and later
•Open MPI 4.0 or newer is preferred for Graviton
•Intel MPI 2019 Update 5 and later
•NVIDIA Collective Communications Library (NCCL) 2.4.2 and later

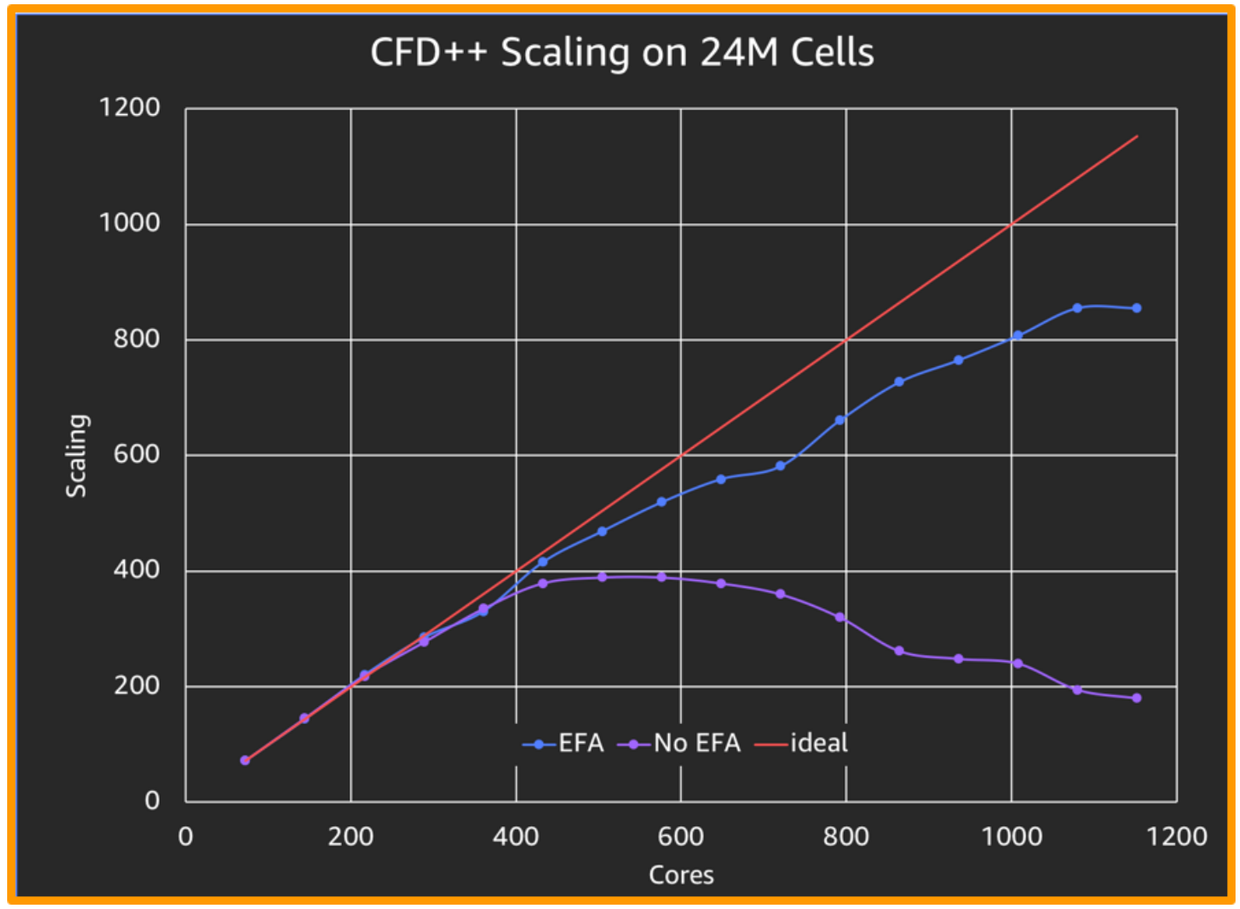
aken from https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/efa.html (18/07/2024)

How to use EFA

1. Instance Support.
2. AMI Support.
3. Install Software components (EFA Kernel Module, Libfabric Network stack, MPI or NCCL Implementation).
4. Apply security group.

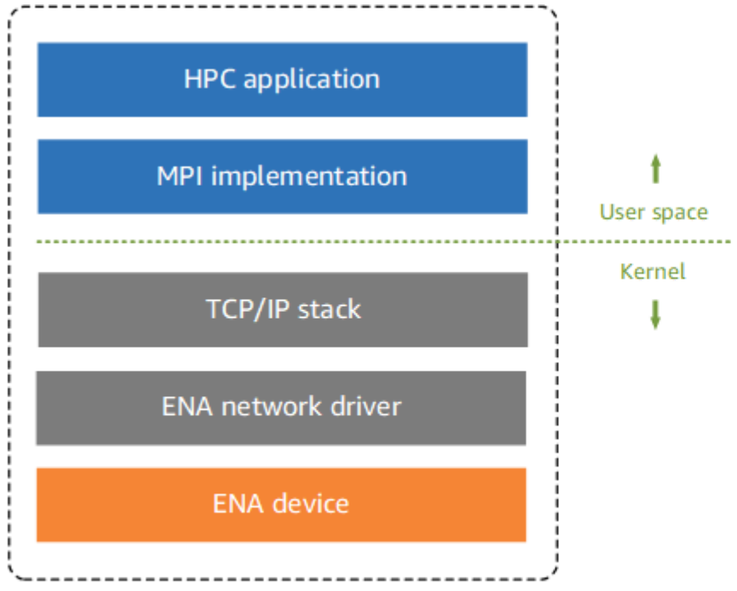
Limits:

On EFA mode, it communicate with a VPC subnet only.
Not hot-attached.
One EFA per instance.
Sec Group ingress and egress only.

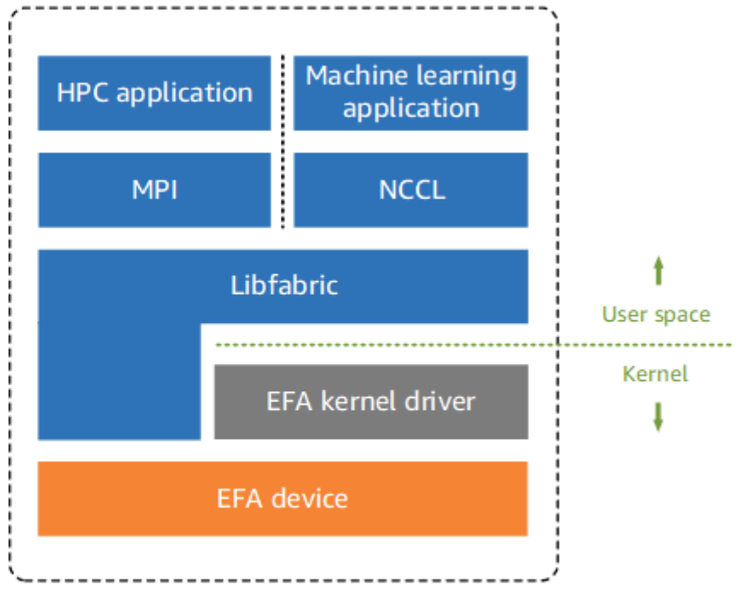


Taken from https://aws.amazon.com/hpc/efa/ and https://aws.amazon.com/blogs/aws/now-available-elastic-fabric-adapter-efa-for-tightly-coupled-hpc-workloads/ (05/10/2020)

ENA vs EFA



Traditional HPC software stack in EC2



HPC software stack in EC2 with EFA

ENA	EFA
	ENA plus OS-Bypass.
Windows and Linux	Linux
More than one ENA per instance	One EFA per instance
SRV IO Tech, up to 100 Gbps.	
Use TCP/IP Stack	Have a mode to overpass OS = OS Bypass to avoid TCP/IP Stack.
Can be hot attach	Cold attach only

More information at ENI Benchmark using Code (https://aws.amazon.com/premiumsupport/knowledge-center/network-throughput-benchmark-linux-ec2/) and Blogs about using EFA for HPC (https://aws.amazon.com/blogs/aws/now-available-elastic-fabric-adapter-efa-for-tightly-coupled-hpc-workloads/) (29/05/2021)