AWS Big Science Best Practices

Dario Rivera – AWS Solutions Architect

Brendan Bouffler – AWS BDM Scientific Computing

AWS becoming Valuable partner for Cloud Scientific Computing

- High Performance Computing (HPC) for Engineering and Simulation
- High Throughput Computing (HTC) for Data-Intensive Analytics
  - Hybrid Supercomputing centers
- Collaborative Research Environments
- Science-as-a-Service

Schrodinger & Cycle Computing: Computational Chemistry for Better Solar Power

Simulation by Mark Thompson of the University of Southern California to see which of 205,000 organic compounds could be used for photovoltaic cells for solar panel material.

Estimated computation time 264 years completed in 18 hours.

SCHRÖDINGER.

CYCLECOMPUTING

Loosely Coupled

- 156,314 core cluster, 8 regions
- 1.21 petaFLOPS (Rpeak)
- $33,000 or 16¢ per molecule

Schrodinger & Cycle Computing: Computational Chemistry for Better Solar Power

Simulation by Mark Thompson of the University of Southern California to see which of 205,000 organic compounds could be used for photovoltaic cells for solar panel material.

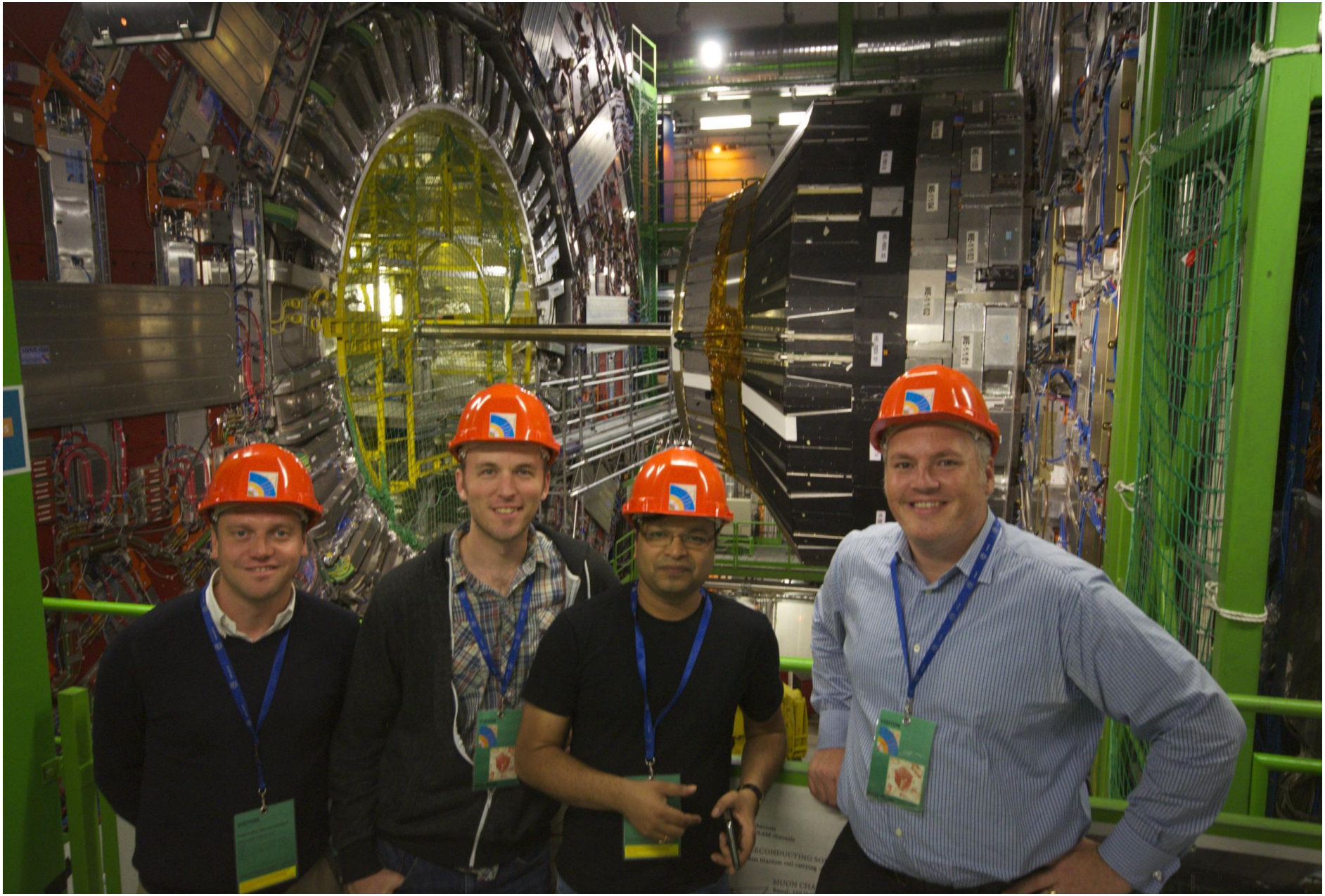
Estimated computation time 264 years completed in 18 hours.

SCHRÖDINGER

CYCLECOMPUTING

Loosely Coupled

- 156,314 core cluster, 8 regions
- 1.21 petaFLOPS (Rpeak)
- $33,000 or 16¢ per molecule



ATLAS Program Challenges helped by AWS

Challenge 1: Leverage the Network for Scale

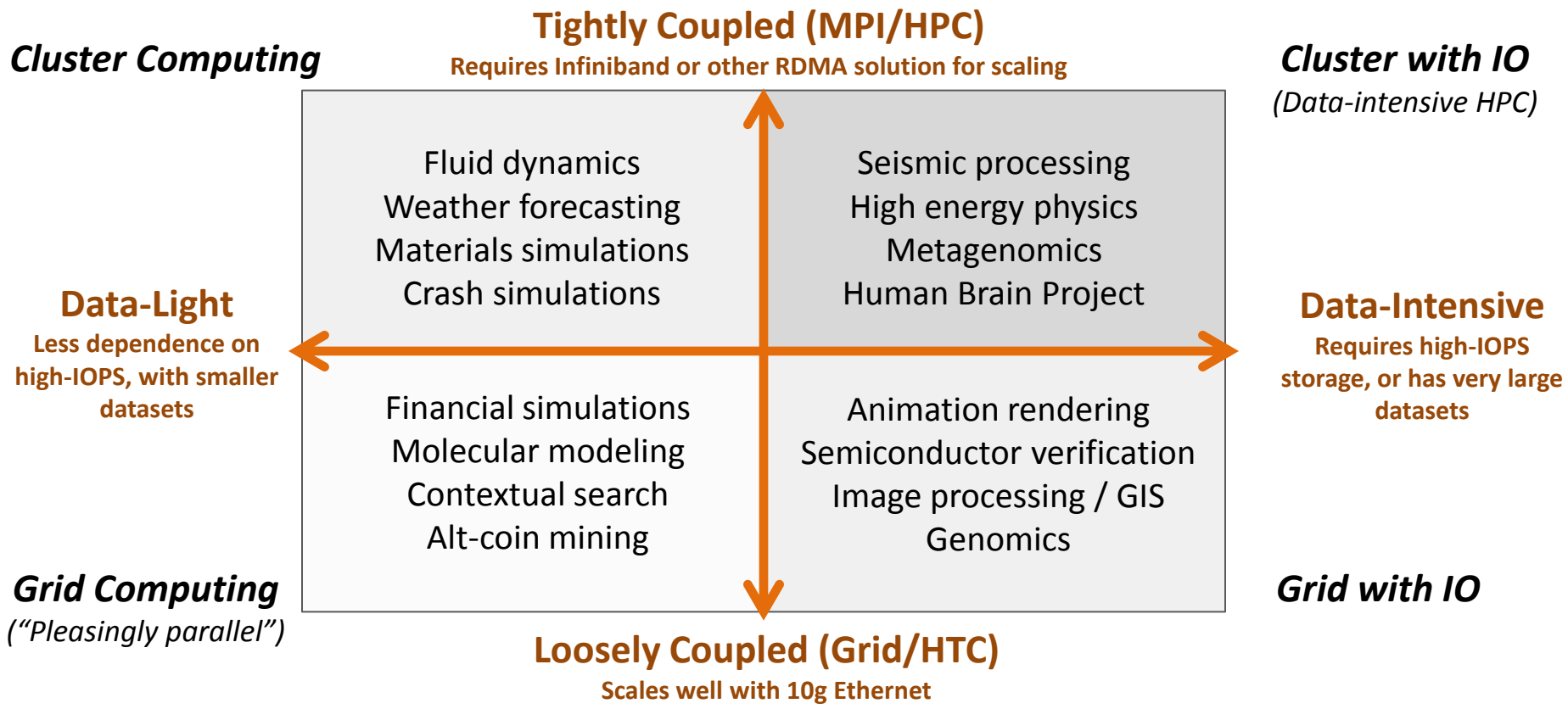
- Challenge:
  - Hybrid Architecture running On-premises and on AWS with minimal Latency absolutely necessary
  - Provide dedicated bandwidth to allow large scale data transfer from On-Prem Storage to Elastically Scaled Cloud Storage.
- Answer:
  - DirectConnect Service through AWS
    - AWS DirectConnect (DX) lets you establish a dedicated network connection between your network and one of the AWS Direct Connect locations. Using industry standard 802.1q VLANs. DX can be provisioned for 1G, 10G or multiples there of via Equal Cost Multipath binding. (Hashing)

DirectConnect Best Practices

- Leverage your AWS SciCo Team Rep as much as possible (more on this later)
- Although AWS provisions DX fairly automated, there are still some Manual steps which require human interaction… engaging AWS support sooner rather than later to help finalize connection.
- DX Getting Started:

http://docs.aws.amazon.com/directconnect/latest/UserGuide/getstarted.html

Mapping HPC Applications



Challenge 2: Provisioning Cheap Flexible Compute

- Once connectivity is established with your Virtual Private Cloud (VPC) provisioning cheap compute offered at High Flexibility within AWS

- Challenge 2: Contrary to popular perception, there is NO cloud that has unlimited compute. How can compute be provisioned elastically, securely, and cheaply for HIGH Data workloads

Answer 2: EC2 Spot Market

- EC2 Spot Market - Spot Instances allow you to name your own price for Amazon EC2 computing capacity. You simply bid on spare Amazon EC2 instances and run them whenever your bid exceeds the current Spot Price, which varies in real-time based on supply and demand.
- EC2 SPOT Market actually is over 20 different markets (1 per AZ)
- Bidding strategies can vary based on workload

Spot Market Best Practices

- Capacity can vary depending on the competitiveness of the market (AZ) for the given instance type
- Bid on the instance type based on how much the per hour use of that instance is worth to your budget
- Use Termination notices when possible: https://aws.amazon.com/blogs/aws/new-ec2-spot-instance-termination-notices/

Other Spot Best Practices

Source: http://www.ijecse.org/wp-content/uploads/2012/09/Volume-1Number-4PP-2395-2397.pdf

A: Save Work Frequently via Work Splits or Checkpoints (or both)

B: Test Application for Fault Tolerance and Interruption handling

C: Minimize Group instance Launches

- Launch Groups – minimize group size as larger groups will be harder to fulfill
- AZ Groups – if possible avoid in order to increase chance of fulfilment across Azs within a given region

D: Use Persistent Requests for Continuous Tasks

- A one-time request will only be satisfied once; a persistent request will remain in consideration after each instance termination.

E: Track when Spot Instances Start and Stop

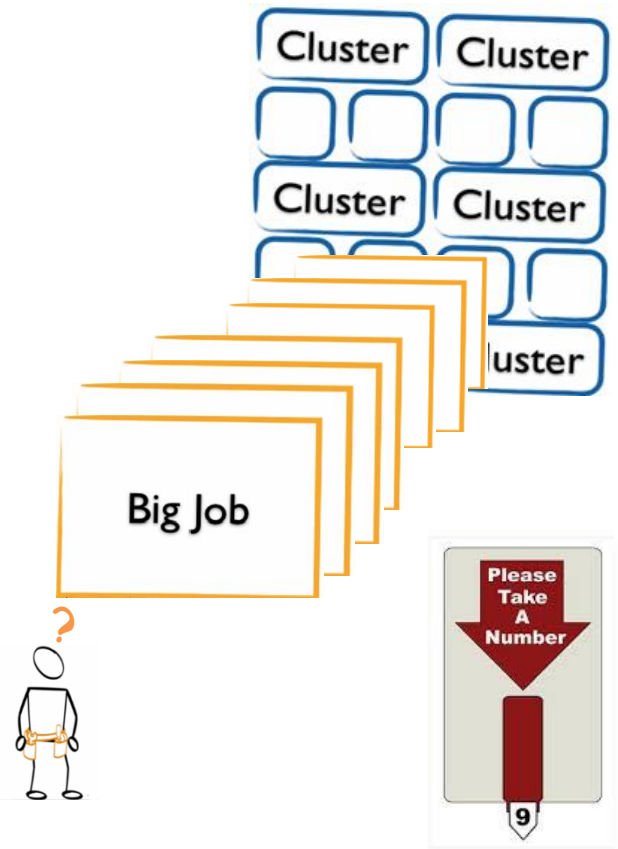
- Mature workloads can optimize by looking for state changes to spot requests which help further understand which markets have availability at what intervals

F: Access Large Pools of Compute Capacity

- If customer needs are urgent, customer can specify a high maximum price which will raise customer request's relative priority and allow customer to gain access to as much immediate capacity as possible

Challenge 3: Getting the most performance out of your Compute

i.e. – HPC Job Queues are Evil!



Conflicting goals

- HPC users seek fastest possible time-to-results
- IT support team seeks highest possible utilization

Result:

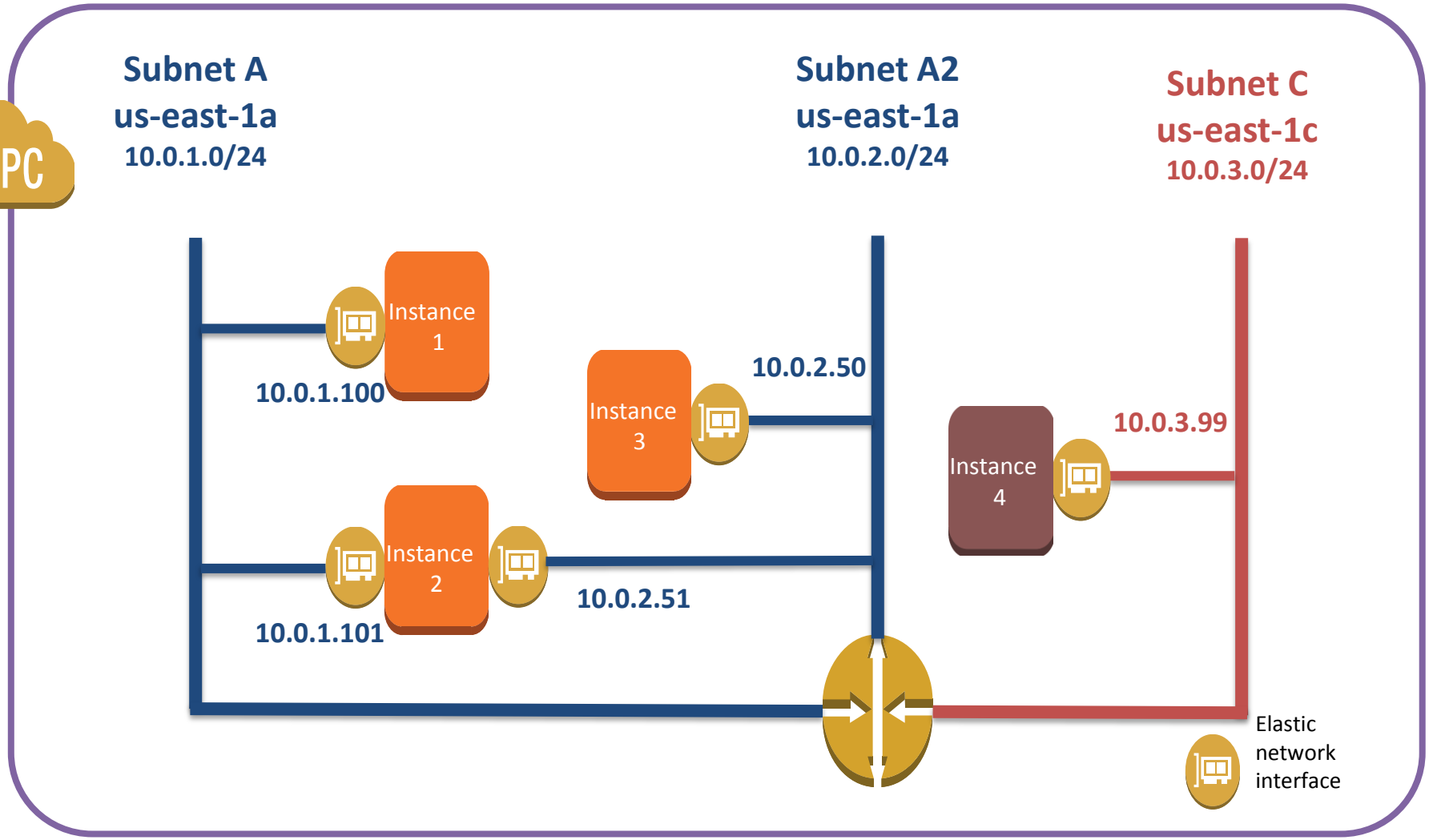
- The job queue becomes the capacity buffer, and there is little or no scalability
- Users are frustrated and run fewer jobs
- Money is being saved, but for the wrong reasons!

Answer: Utilizing the right EC2 Tools, Configuration and Instance Types for your Workload

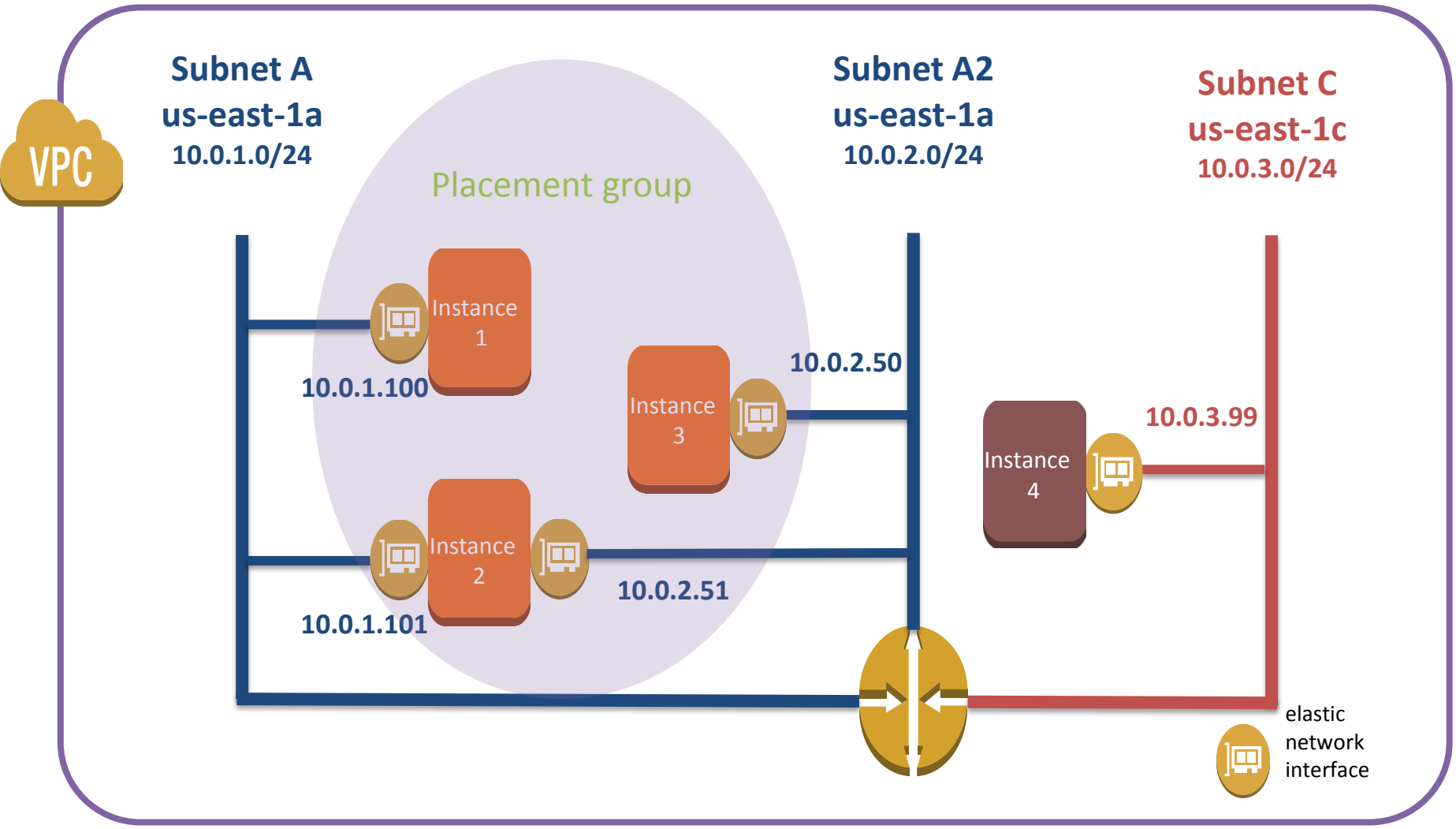
Source: https://www.youtube.com/watch?v=JUw8y_pqD_Y

- Recognize the Tools you have in your AWS Toolbox for compute
  - Placement Groups (i.e. Shared Storage Algorithms)
  - Elastic Network Interface (ENI) and Elastic IP
  - SR-IOV
- Cloudwatch Monitoring, other performance monitoring tools

Placement groups

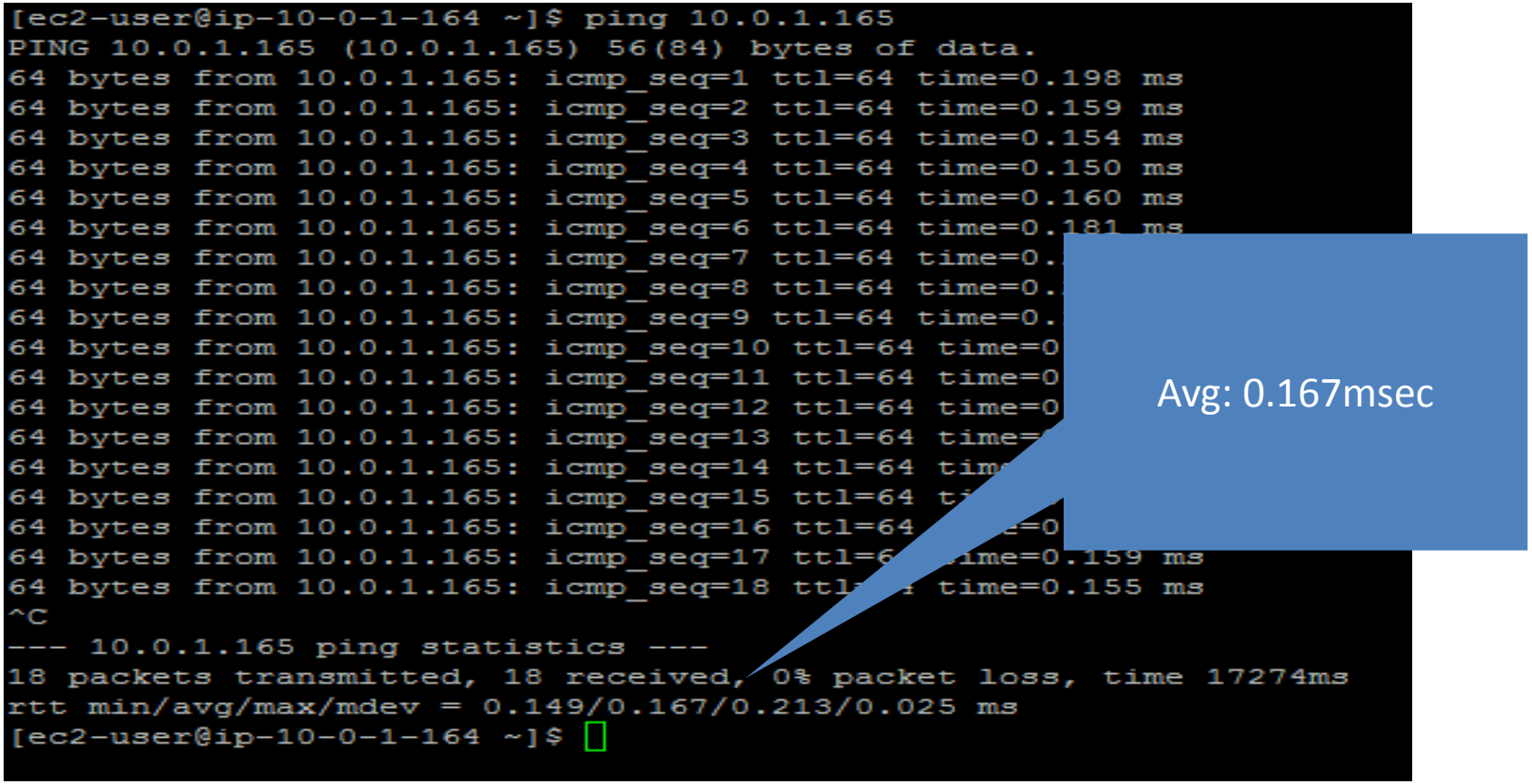

VPC
Subnet A
us-east-1a
10.0.1.0/24
Subnet A2
us-east-1a
10.0.2.0/24
Subnet C
us-east-1c
10.0.3.0/24
Instance 1
10.0.1.100
Instance 3
10.0.2.50
Instance 4
10.0.3.99
Instance 2
10.0.2.51
10.0.1.101
Elastic network interface

A VPC

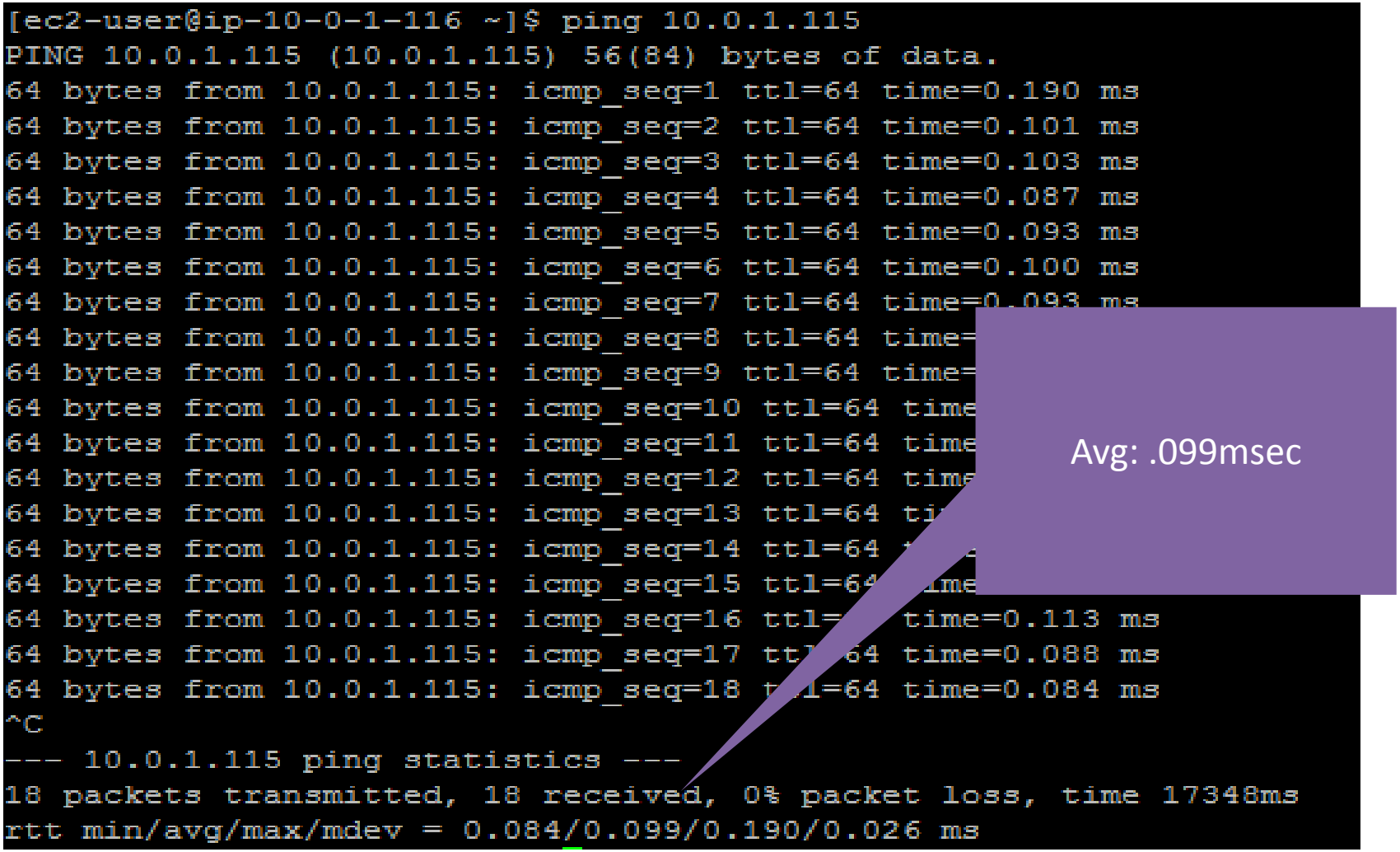


A VPC with a placement group

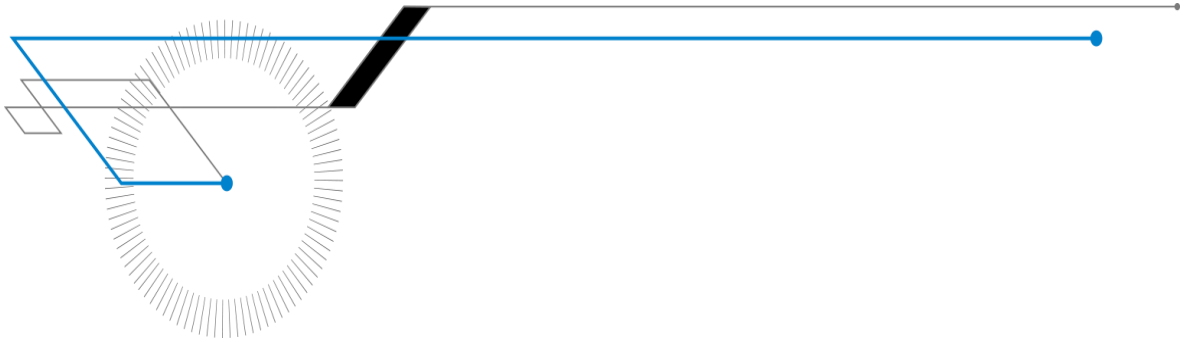
Inter-instance ping: No placement group



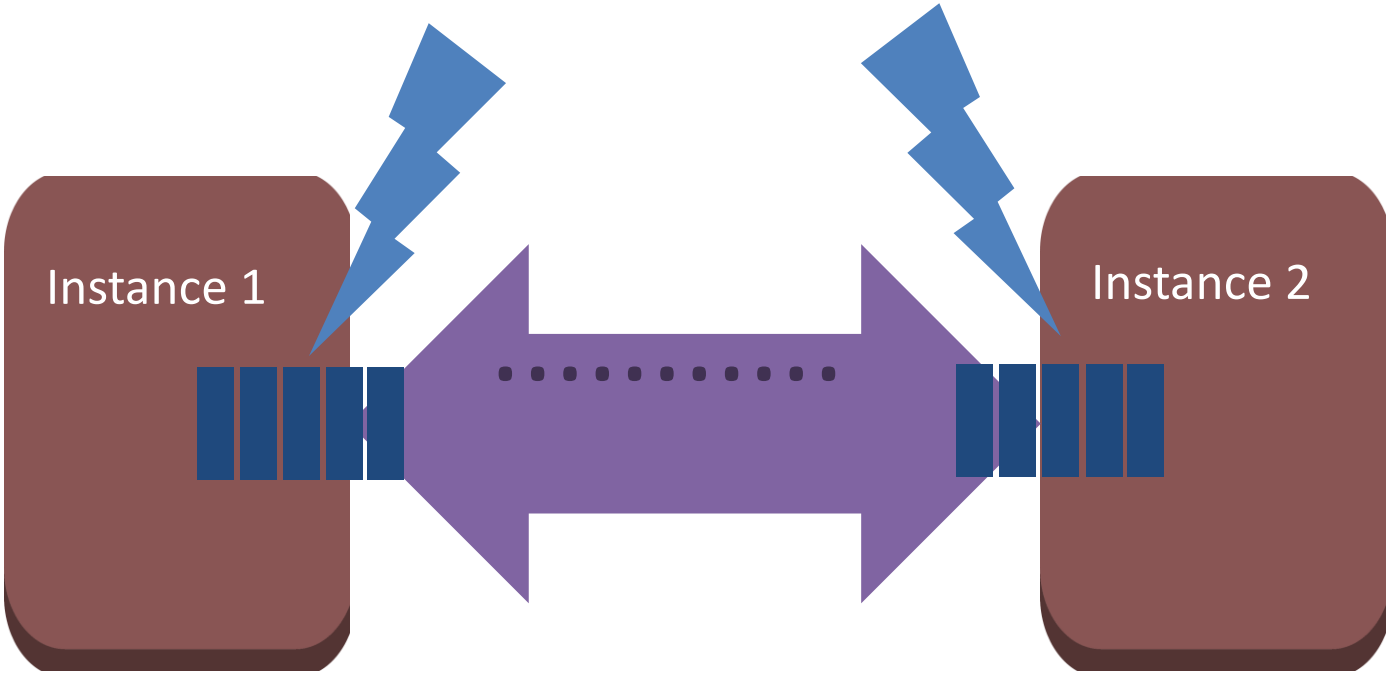
Inter-instance ping: Placement group



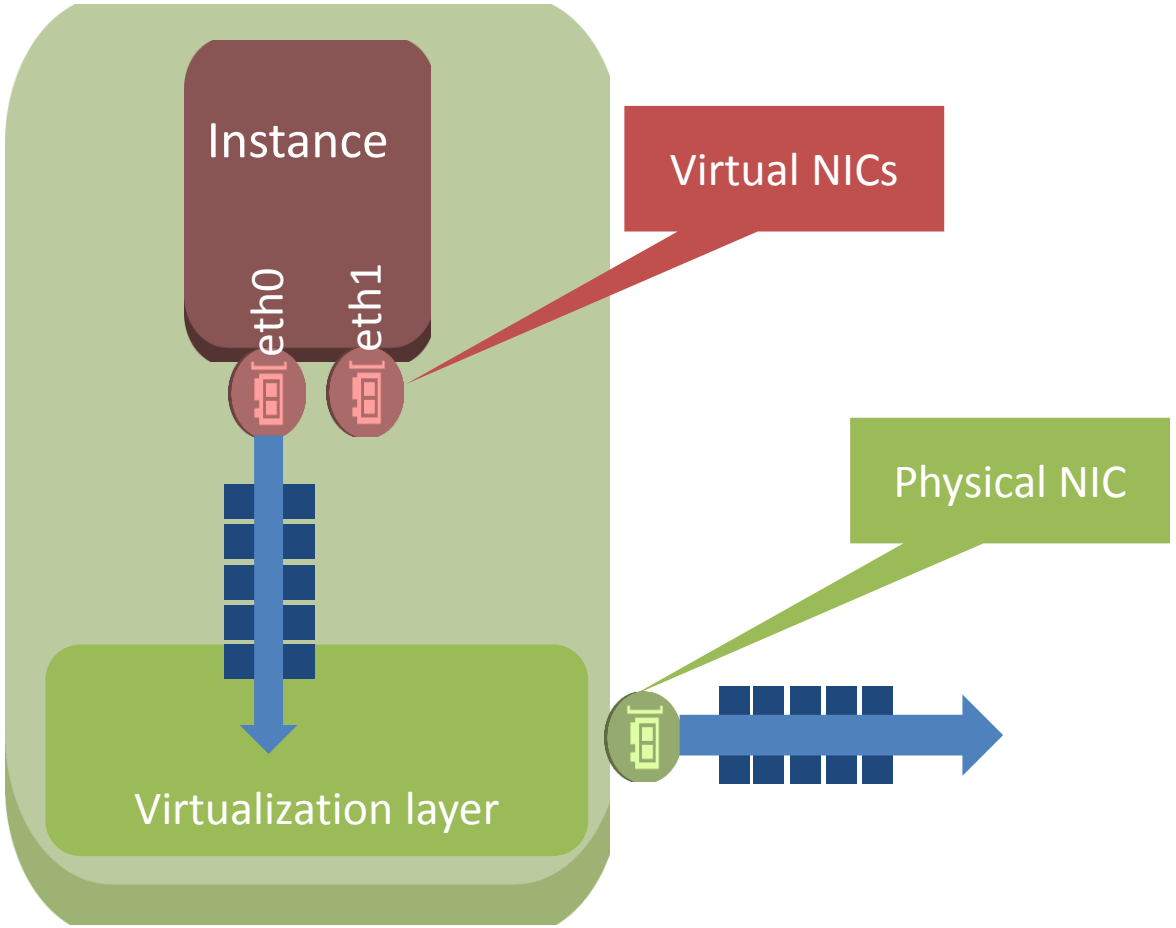
Enhanced Networking



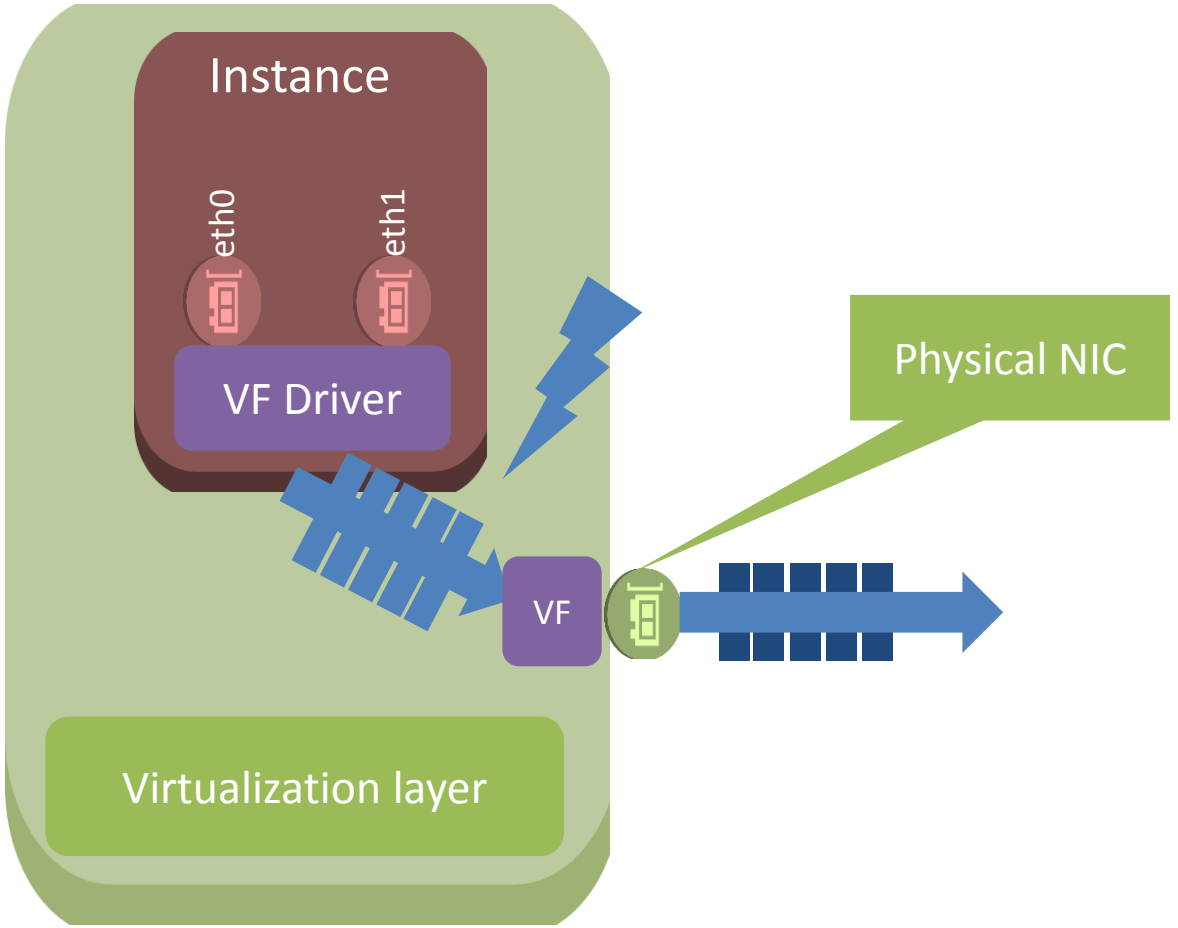
Latency: Packets per second



Packet processing in Amazon EC2: VIF



Packet processing in Amazon EC2: SR-IOV



SR-IOV: Is this thing on? It may already be!

For many newer AMIs, Enhanced Networking is already on:

- Newest Amazon Linux AMIs
- Windows Server 2012 R2 AMI

No need to configure

SRIOV: Is this thing on? (Linux)

No

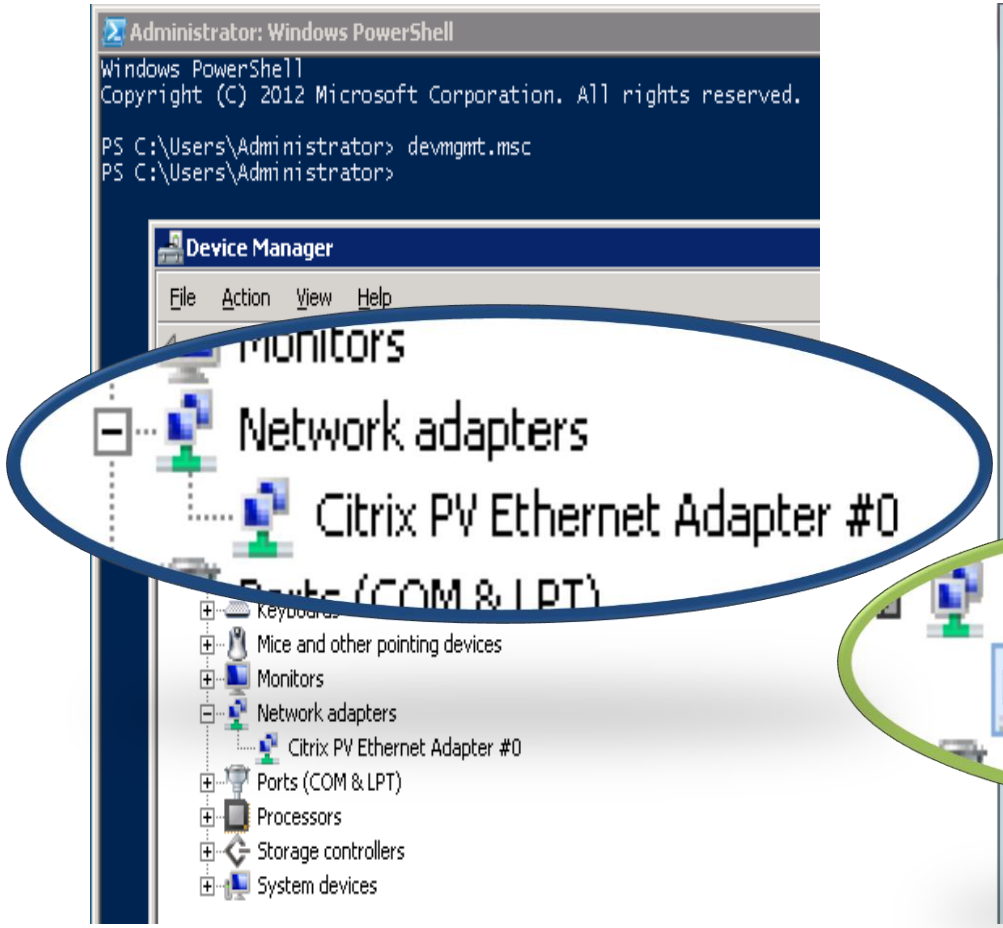
```
[ec2-user@ip-10-0-3-70
~]$ ethtool -i eth0

driver: vif
version:
firmware-version:
bus-info: vif-0
…
```

Yes!

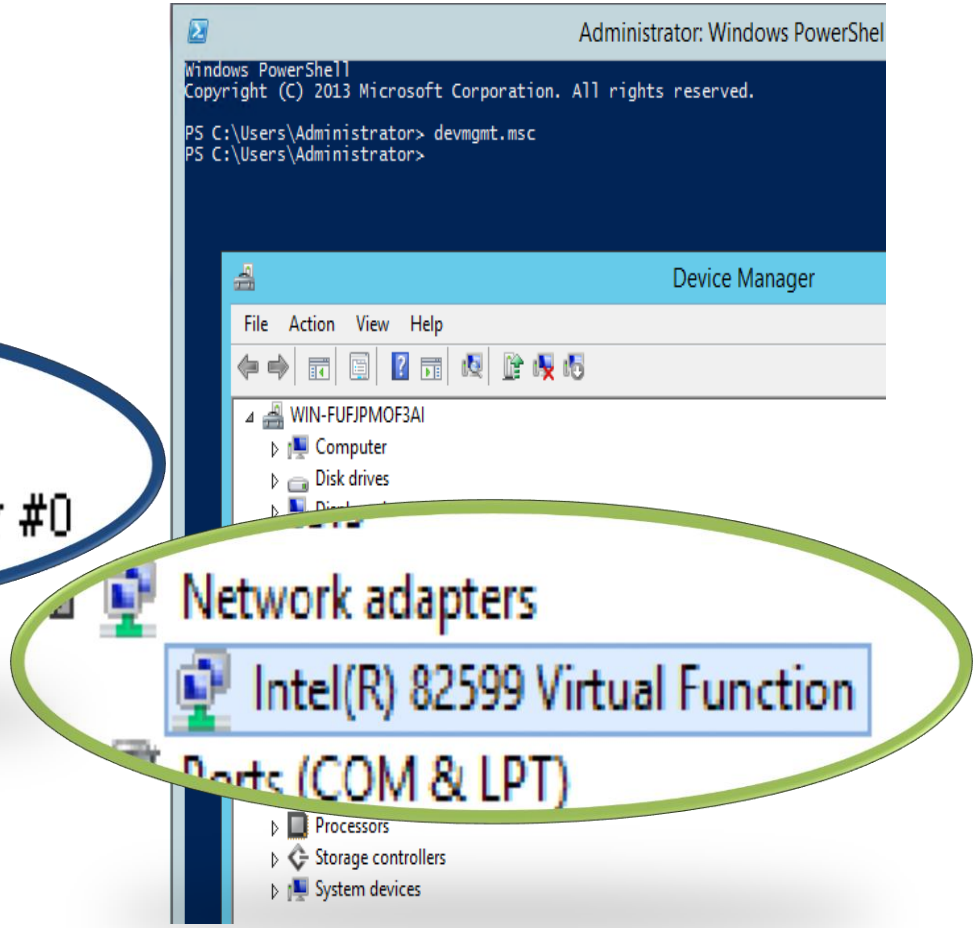
```
[ec2-user@ip-10-0-3-70 ~]$
ethtool -i eth0

driver: ixgbevf
version: 2.14.2+
firmware-version: N/A
bus-info: 0000:00:03.0
…
```

SRIOV: Is this thing on? (Windows)

No



Yes!



AMI/instance support for SR-IOV

- C3, C4, R3, I3 instance families: 18 instance types
- HVM virtualization type
- Required kernel version
  - Linux: 2.6.32+
  - Windows: Server 2008 R2+
- Appropriate VF driver
  - Linux: ixgbevf 2.14.2+ module
  - Windows: Intel® 82599 Virtual Function driver

Challenge 3: Finding Highest Storage Capacity for Best Performance at Cheapest Cost

- Answer: Understand your options:
  - S3 (~.03/GB) – Unlimited Capacity, Webscale Accessibility, available via DX and Public Interface
  - EBS (~.10/GB) – Make use of automation to have SPOT connect to EBS for persistent check pointing
  - Instance Storage (Free) – fastest performance disk, ephemeral – lost after instance stop, reboot or termination
  - Many others: DynamoDB, RDS, Aurora, Redshift, etc

Block-Based Storage Best Practice

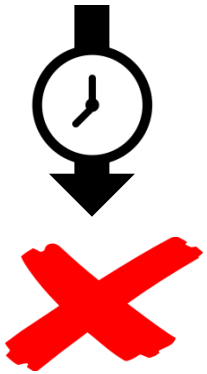
- Be Sure to Pre-Warm!!
  - Best for Short running Jobs
  - Long running job will take the affect of cold disk over the life of job
  - Analyze job times for sweet spot of prewarming over not
- Source: http://docs.aws.amazon.com/AWSEC2/latest/UserGuide/ebs-prewarm.html

S3 Best Practices

Source: https://www.youtube.com/watch?v=2DpOS0zu8O0

Tip #1: Versioning + Life Cycle Policies

Versioning → Recycle bin



Lifecycle policies → Automatic cleaning

S3 Best Practices

Tip #2: Maximizing performance

- Use a key naming scheme with randomness at the beginning for high TPS
  - Most important if you regularly exceed 100 TPS on a bucket
  - Avoid starting with a date
  - Consider adding a hash or reversed timestamp (ssmmhhddmmyy)
- Multipart upload provides parallelism
  - Allows you to upload a single object as a set of parts
  - Enables pausing and resuming, and beginning before the total size is known
  - Encouraged for objects larger than 100MB; required above 5GB
- Source: http://docs.aws.amazon.com/AmazonS3/latest/dev/request-rate-perf-considerations.html

AWS SciCo Team Engagement

I'm not psycho... I just like SciCo-tic things. – Gerard Way

Why Did AWS Create SciCo?

In order to make it easy to find, discover and use AWS for scientific computing at any scale. Specifically, SciCo helps AWS Users:

- More effectively support global "Big Science" Collaborations
- Develop a solution-centric focus for engaging with the global scientific and engineering communities
- Accelerate the development of a scientific computing ecosystem on AWS
- Educate and Evangelize AWS' role in Scientific Computing

Introducing the SciCo Team

Jamie Kinney (jkinney@) – Sr. Manager for SciCo (Seattle)

Michael Kokorowski (mkoko@) –Technical Program Manager (Seattle)

Traci Ruthkoski (ruthkosk@) - Technical BDM, Earth Sciences (D.C.)

Angel Pizarro (angel@) – Technical BDM, Genomics & HCLS (Philadelphia)

Brendan “Boof” Bouffler (bouffler@) – Technical BDM, Hybrid Supercomputing Centers, SciCo in APAC, and the SKA (Sydney)

Kevin Jorissen (jorissen@) - Technical BDM, Materials Science and Long Tail solutions (Seattle)

The AWS Research Grants Program

AWS provides millions of dollars in free AWS usage annually to researchers. The program was initially designed to help AWS understand the needs of the academic and research communities. We now focus 90% of our grants on driving specific outcomes, including:

- Creation of AMIs, white papers, financial/technical benchmarks, and other long-tail solutions
- Accelerating the development of Science-as-a-Service applications
- Training new AWS users by providing resources for hands-on labs and workshops
- Accelerating the launch of new projects by strategic customers through support for pilot projects

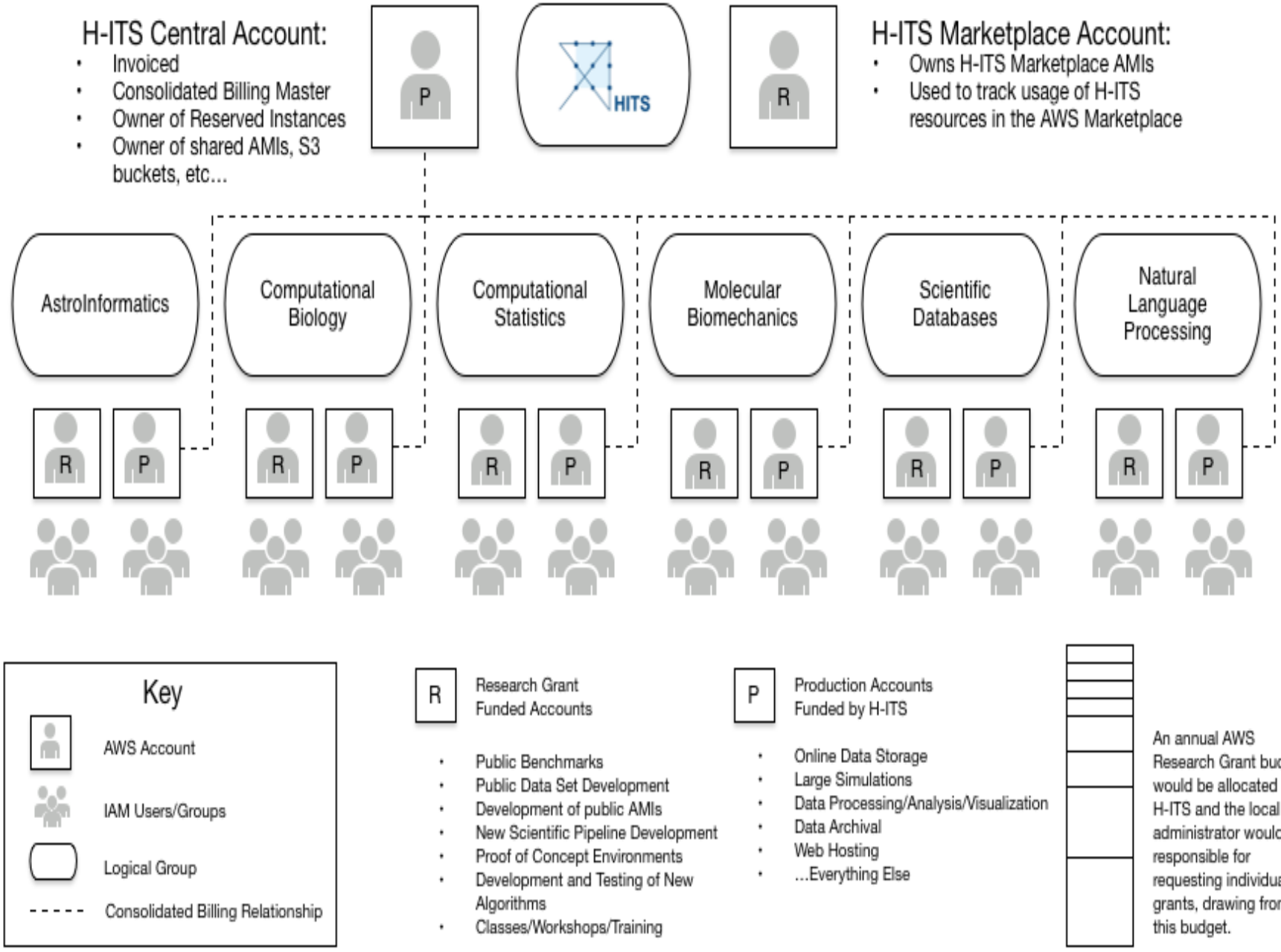
SciCo Engagement Example: HITS

The Heidelberg Institute for Theoretical Studies is a non-profit research organization funded by Claus Tschira (co-founder of SAP). HITS conducts basic research in natural sciences, mathematics and computer science to process, structure, and organize large amounts of data. The areas covered by HITS research range from astrophysics to cell biology.

As a SciCo long-tail solutions partner, HITS is actively developing AMIs, science-as-a-service platforms and is actively promoting the benefits of using AWS for scientific computing.

HITS

AWS Account Structure for the Heidelberg Institute for Theoretical Studies



Thank you!