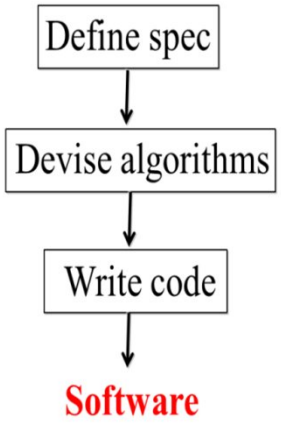
工業技術研究院
Industrial Technology
Research Institute

# DNN Training As a Service

# Introduction to Deep Neural Network

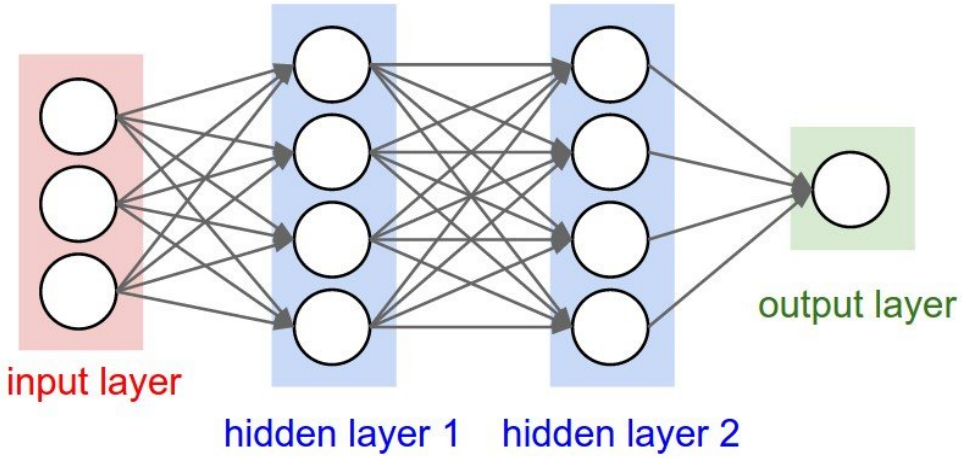
- AI ~ Deep Neural Network (DNN)-based Machine Learning
  - Algorithmic breakthrough that enables training of deep neural network
- AI system: a system with analysis and synthesis capabilities powered by DNN-based machine learning or Deep Learning
  - Autonomous driving vehicle, drone, robot, personal virtual assistant, etc.
- Machine learning: a universal algorithm for building a functional mapping between sample inputs and associated outputs
  - A new paradigm of software development
  - Learn from many normal people

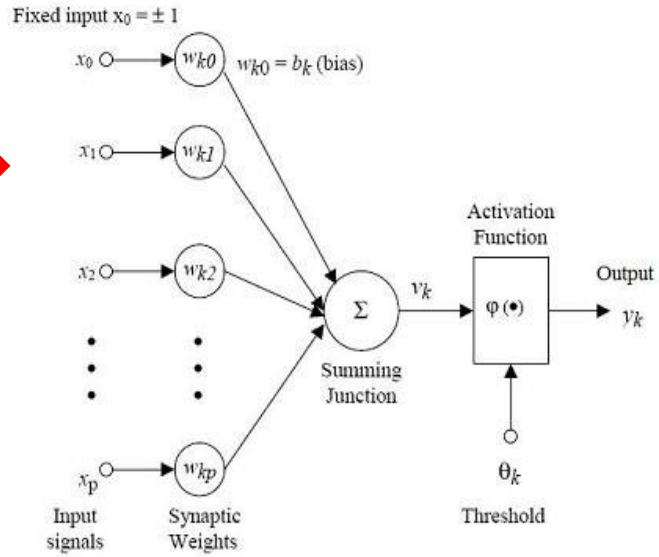
    vs. Design with few gifted experts



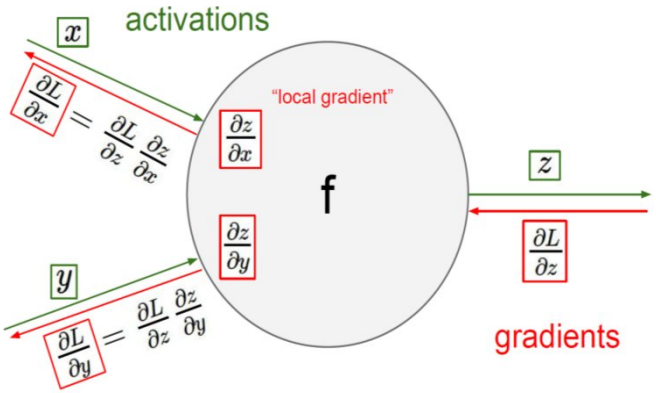
# Training and Inference in DNN

Training: Forward/Backward Propagation ⬅➡




Inference: Forward Propagation ➔



# Key Challenges in DNN Systems

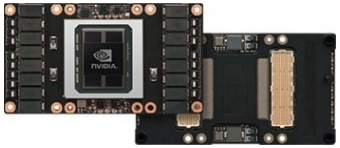
- Training of DNN model
  - Quality: how to acquire high-quality training data set
    - Label correctness and diversity
    - Semi-automatic training data collection and labeling
  - Speed:
    - Reduce the number of rounds required in the training process
      - Round ➔ Epoch ➔ Batch
    - Reduce the computation overhead associated with each training round
- Speed and power consumption of applying DNN model (inference)
  - Real-time: autonomous driving
  - Embedded system: low power and low cost
- Explainability of learned DNN models

工業技術研究院
Industrial Technology
Research Institute

# NVIDIA DGX-1 Appliance

DGX-1 with P100 at USD$129,000*
DGX-1 with V100 at USD$149,000*

* Manufacturer's Suggested Retail Price (MSRP) as established by NVIDIA. Taxes, shipping and other fees not included. Support plan is required and must be purchased separately. All pricing is subject to change without notice.

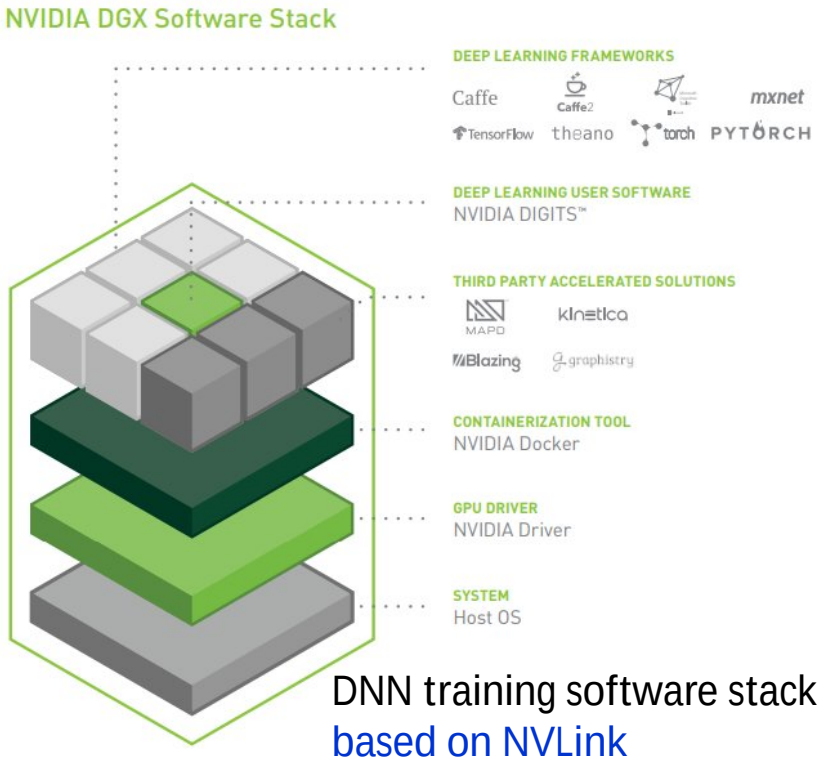
**SYSTEM SPECIFICATIONS**

| | | |
|---|---|---|
| GPUs | 8X Tesla V100 | 8X Tesla P100 |
| TFLOPS (GPU FP16) | 960 | 170 |
| GPU Memory | 128 GB total system | |
| CPU | Dual 20-Core Intel Xeon E5-2698 v4 2.2 GHz | |
| NVIDIA CUDA® Cores | 40,960 | 28,672 |
| NVIDIA Tensor Cores (on V100 based systems) | 5,120 | N/A |
| Maximum Power Requirements | 3,200 W | |
| System Memory | 512 GB 2,133 MHz DDR4 LRDIMM | |
| Storage | 4X 1.92 TB SSD RAID 0 | |
| Network | Dual 10 GbE, 4 IB EDR | |
| Software | Ubuntu Linux Host OS See Software Stack for Details | |
| System Weight | 134 lbs | |
| System Dimensions | 866 D x 444 W x 131 H (mm) | |
| Packing Dimensions | 1,180 D x 730 W x 284 H (mm) | |
| Operating Temperature Range | 10–35 °C | |

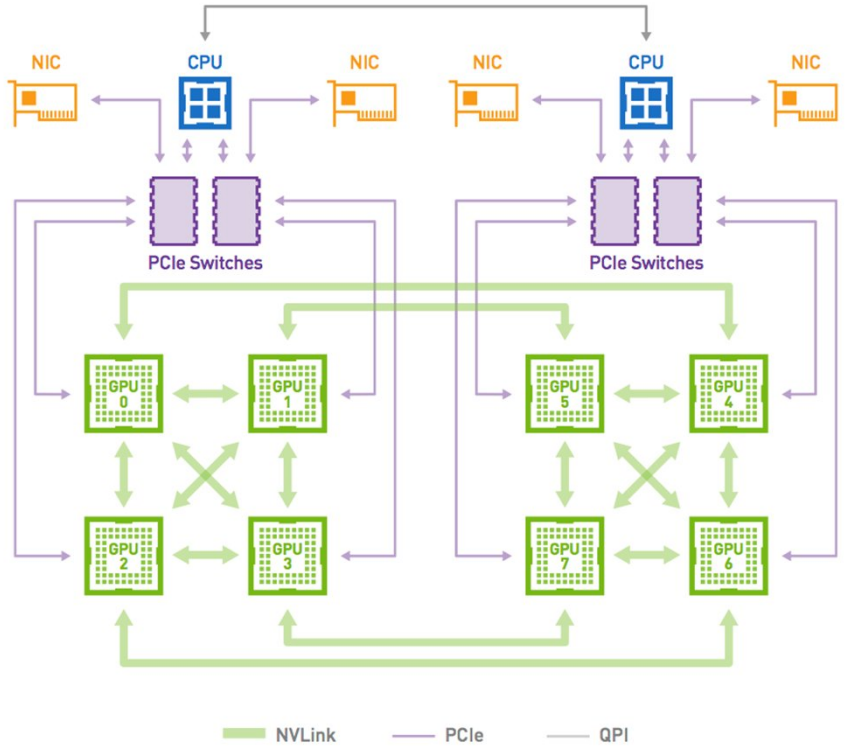
https://blogs.nvidia.com/blog/2016/08/15/first-ai-supercomputer-openai-elon-musk-deep-learning/
https://www.nvidia.com/content/dam/en-zz/Solutions/Data-Center/dgx-1/dgx-1-ai-supercomputer-datasheet-v4.pdf

# DGX-1's System Architecture

NVIDIA Tesla P100 GPU SXM2(for NVLink)
NVIDIA Tesla V100 GPU SXM2(for NVLink)




DNN training software stack
based on NVLink



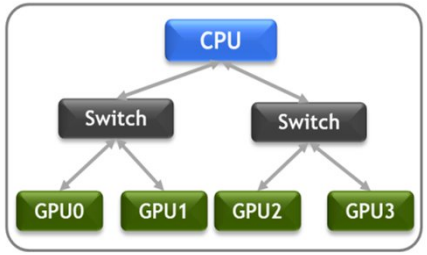

NVLink
Extraordinary bandwidth for multi-GPU connectivity

| | PCIe 3.0 x16 | NVLink 1.0 | NVLink 2.0 |
|---|---|---|---|
| Sub-Link Data Rate (per data direction) | 16 GB/s | 20 GB/s | 25 GB/s |

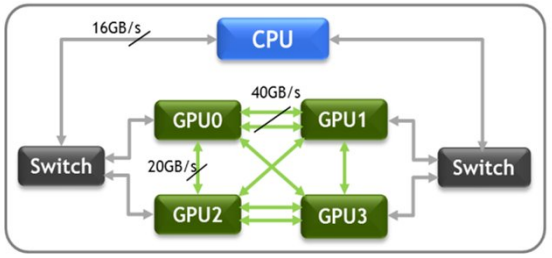
https://www.nvidia.com/content/dam/en-zz/Solutions/Data-Center/dgx-1/dgx-1-ai-supercomputer-datasheet-v4.pdf

# ITRI DNN Training Appliance

- Objective: 3/4 of the sustained performance at 1/3 of the price
- Hardware configuration
  - Processor Type:
    - Nvidia's Tesla P100 and V100 (12GB, 4.7TFLOPs, $5899)
    - Nvidia's GeForce GTX 1080Ti (11GB, 11.3TFLOPs of FP32, $699)
    - AMD's Radeon RX-500 and RX Vega
    - Intel's Knights Mill (KNM)
    - FPGA
  - Number of "GPU"s: 16+
  - System Interconnect:
    - Sufficient cross-section bandwidth
    - Meshed PCIe network
  - Cooling



4 GPUs with PCIe



4 GPUs with NVLink

# Software for ITRI DNN Training Appliance

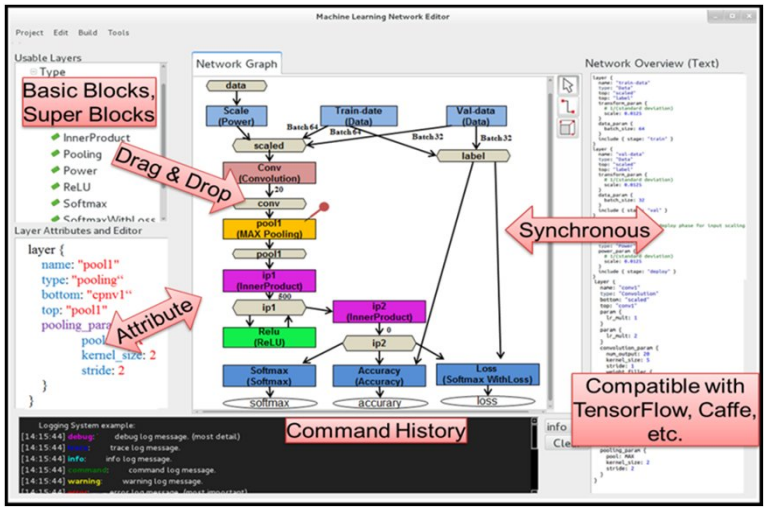
- Supported graphics driver APIs
  - CUDA and OpenCL
- Supported DNN training frameworks
  - Caffe, NVCaffe, Caffe2 and TensorFlow
- Performance Optimizations
  - How to minimize the performance impacts of absence of NVlink
    - Glitch-less data movement from storage, via main memory, to GPU memory
    - Using larger mini-batch size to reduce inter-CPU synchronization overhead
    - Cache memory access optimization via tiling and blocking
  - Amplify the effective memory size of a GPU card via transparent compression/decompression for data stored in GPU memory
- DNN Integrated Development Environment (IDE)
- DNN model compression

# Integrated Development Environment for DNN Training

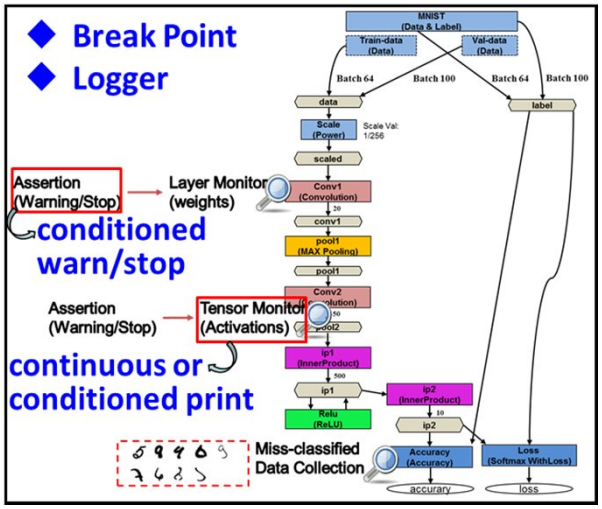
- Coding vs. Training
  - Caters to pattern recognition/classification programs
- Transfer learning of existing DNN models: library function reuse
  - How to enable a user to describe/characterize his DNN training problem so that the system could choose a model in the Model Zoo as a starting point?
  - What transformations should be made to the starting-point model?
- Automated/guided hyper-parameter tuning
  - Human-in-the-loop pruning and steering of a large number of concurrent training rounds, each corresponding to a distinct hyper-parameter combination
  - Automatic tuning of critical hyper-parameters such as learning rate, regularization constant, dropout probability, network depth/breadth, etc.
- Debugging support for weakness analysis of current DNN model
  - Development of informed intermediate metrics and their visualization that reflect a DNN model's overall effectiveness
  - Analysis and categorization of training samples on which predictions are incorrect: mislabeled, noise, difficult/easy to classify, etc.

# DeepMAD: Monitor, Analysis and Debug

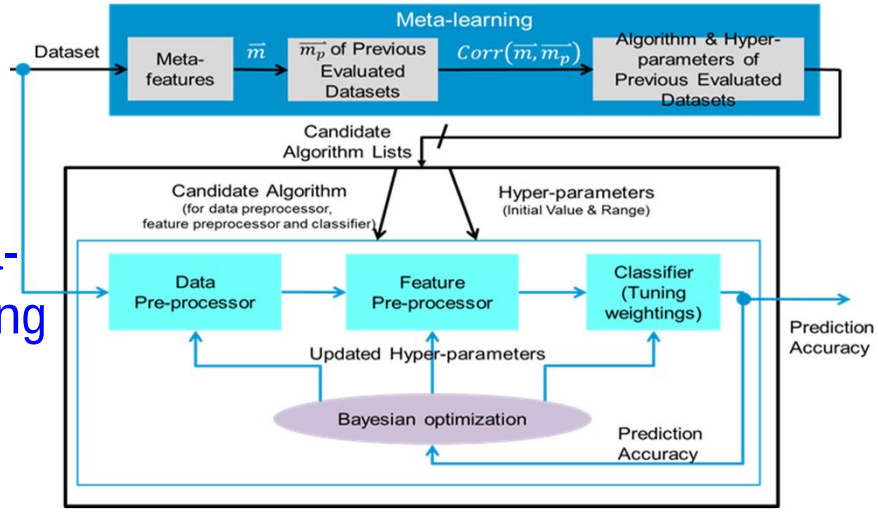
Neural Network Editor




Debugger

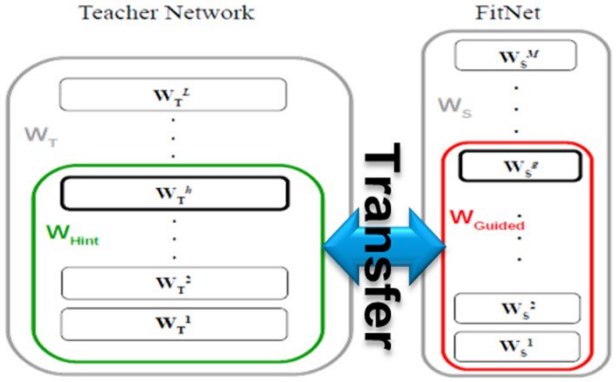



Meta-Learning




Transfer Learning

- FitNet knowledge transfer
- Net2Net network expanding



工業技術研究院
Industrial Technology
Research Institute

# Hardware As a Service (HaaS)

# Introduction to HaaS

- HaaS focuses on leasing of physical machines
  - Each tenant gets a physical data center instance (PDCI), which consists of a set of physical servers, a physical network connecting them, and a set of local/remote storage volumes accessible to the servers.
  - User could remotely configure, monitor, and install OS and applications on servers in its PDCI.
- Why HaaS?
  - Big data/DNN training/HPC: efficient utilization of HW resources is critical
  - Container-based virtualization is sufficient.
- Comparison among service models:

| Model | Rental Unit | IT HW Ownership | HW Management |
|---|---|---|---|
| IaaS | Virtual machine | Service provider | Service provider |
| HaaS | Physical machine | Service provider | User |
| Colocation | Rack space | User | user |

# Technical Building Blocks for ITRI HaaS

- Deployment time
  - Server provisioning: Bare Metal Provisioning from ITRI
  - Storage provisioning: local storage vs. shared storage
  - Network provisioning:
    - Agentless and scalable multi-tenancy network isolation: One HaaS tenant's virtual network is isolated from other HaaS tenants
      - Support up to hundreds of thousands of IP subnets
    - Multi-resolution tenancy-aware network traffic analysis
      - Physical link ➔ Tenant ➔ host-pair ➔ Network connection
    - Load balancing among physical network links
- Run time
  - IT hardware inventory and maintenance: HaaS provider
  - System monitoring and administration: HaaS provider and HaaS tenant

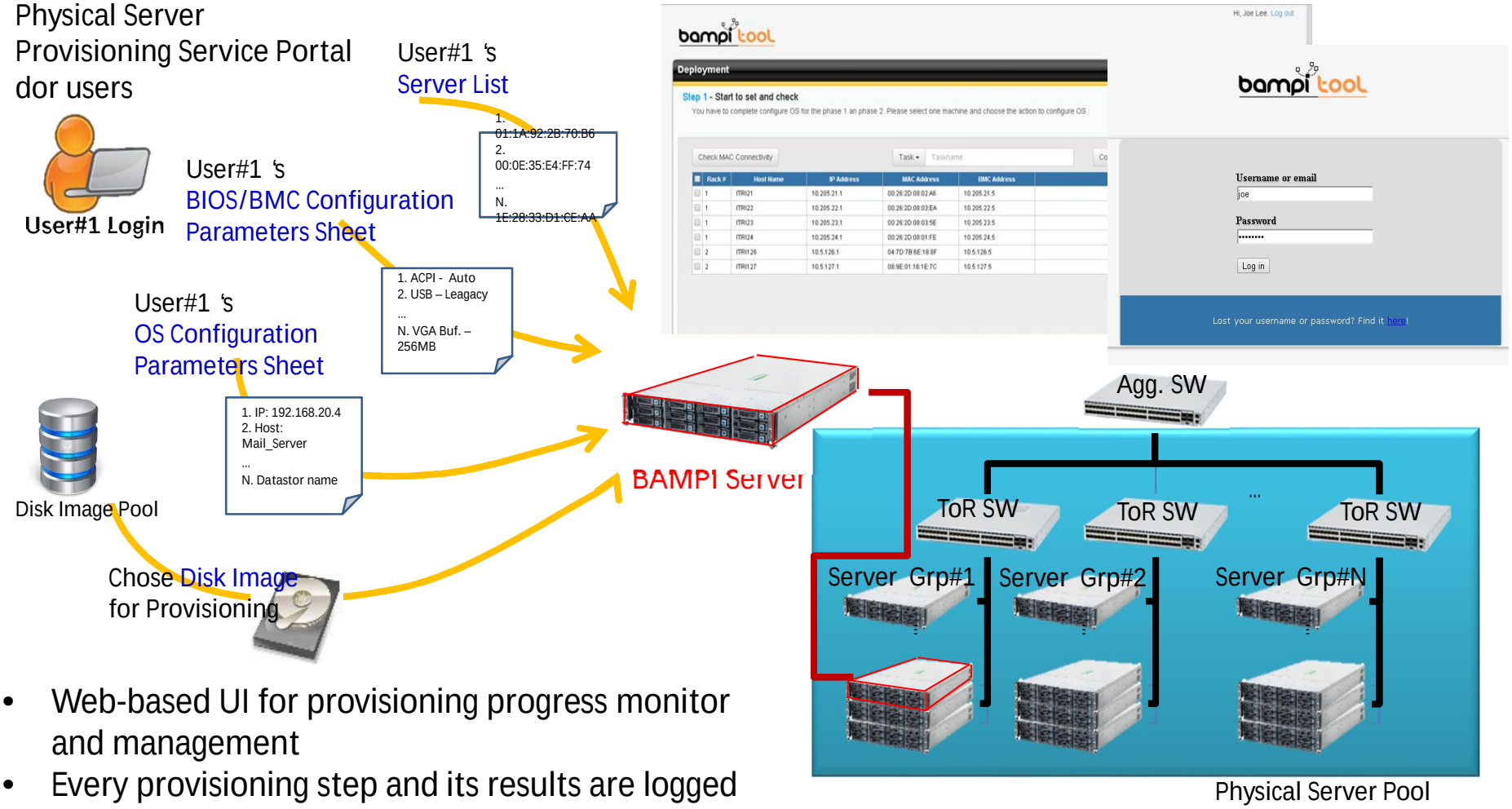
工業技術研究院
Industrial Technology
Research Institute

# BAMPI: Bare-Metal Provisioning from ITRI

- Problem: setting up hardware, network, firmware and first OS on fresh servers correctly is a pain point
- Vision: The only thing a HW vendor needs to do is setting MAC addresses for each server's (BMC/work) NICs, and automates the rest:
  - Extract the physical connectivity between hosts and switches
  - Configure the IP addresses for the work NICs according to a user-specified policy, e.g. machine 13 connected to slot 14 of switch 2 has IP address X
  - Configure the BMC and upgrade its firmware
  - Configure the BIOS and upgrade its firmware
  - Configure NIC and RAID card, and upgrade their firmware
  - Install a local OS (VMware, KVM, Linux) on each server and configure this OS
  - Scalable to thousands of machines residing in multiple L3 subnets
    - DHCP relay + ARP proxy + IP multicast
- Benefits: reduced errors, lower cost and speedier deployment

工業技術研究院
Industrial Technology
Research Institute

# User View of BAMPI




- Web-based UI for provisioning progress monitor and management
- Every provisioning step and its results are logged
- Robust error handling: cancel, clear, and retry
- Remote bare-meta provisioning

# Peregrine: HaaS Network Provisoning

- Physical network connectivity discovery
- VLAN-based inter-tenant network isolation
  - VLAN tree calculation for physical network load balancing
    - Spread the VLANs ' underlying tress across the physical network as much as possible
  - VLAN configuration on  Cisco, HP, Dell and Edgecore switches
  - Scale up the number of VLANs beyond 4096
    - On-demand VxLAN creation
    - VLAN ID reuse
- Tenant-aware multi-resolution network traffic analysis
- Packaged into a Neutron plugin for OpenStack
- Integration with OpenStack's Ironic API is on-going.

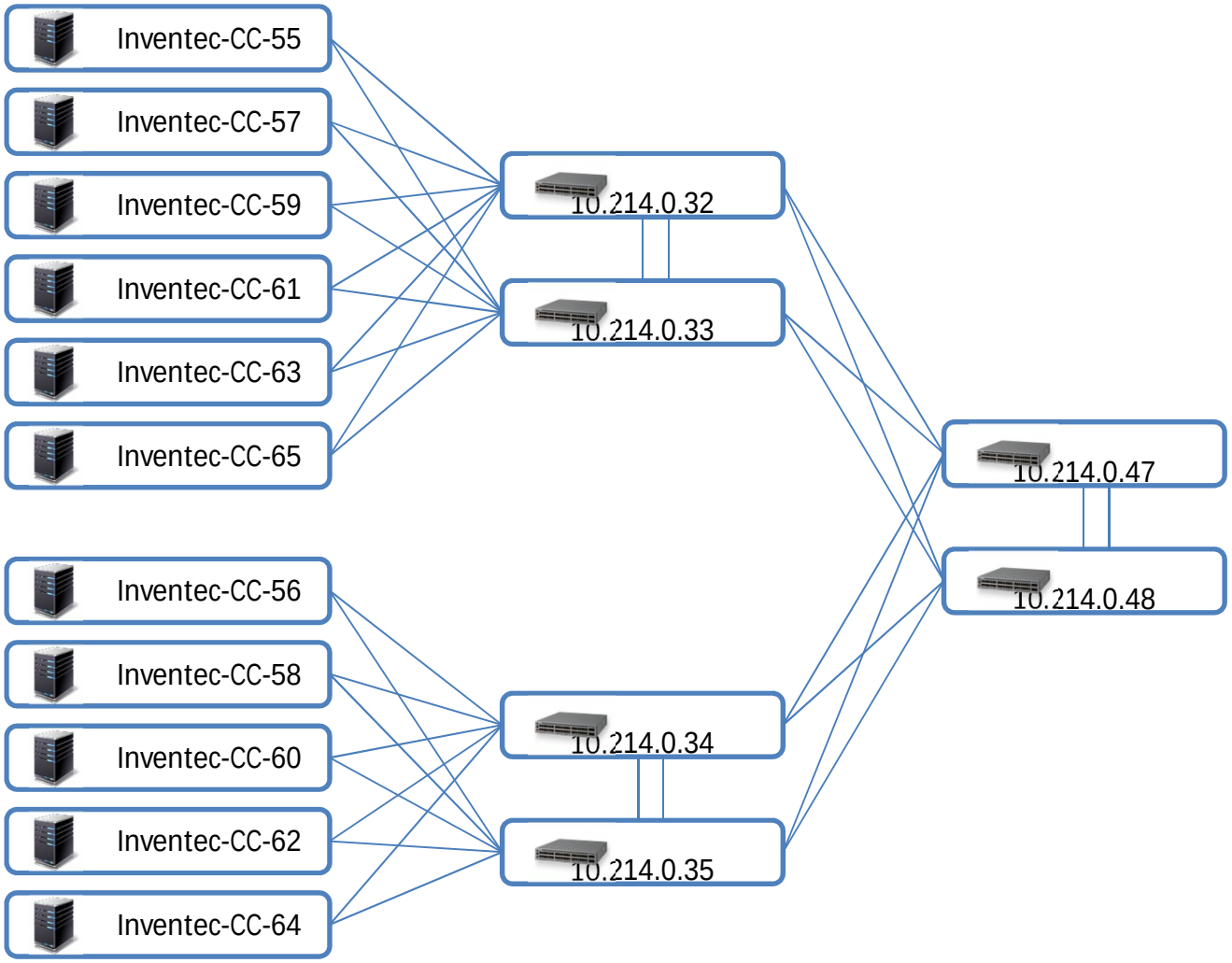
# Physical Connectivity Discovery

- Pain point: How to physically identify a problematic NIC on a physical server for maintenance purpose?
  - No apparent ordering of physical NICs on a server
- Assumptions:
  - A physical server has multiple NICs, each of which has a MAC address and is abstracted into an Eth_x interface
  - A physical server is connected to multiple switches via its NICs
- Technology:
  - Eth_x interface ←→ MAC address ←→ <switch, port>
- Results:
  - Locate a NIC based on the <switch, port> to which it is connected
  - Assign an IP address to a NIC according to its physical connectivity

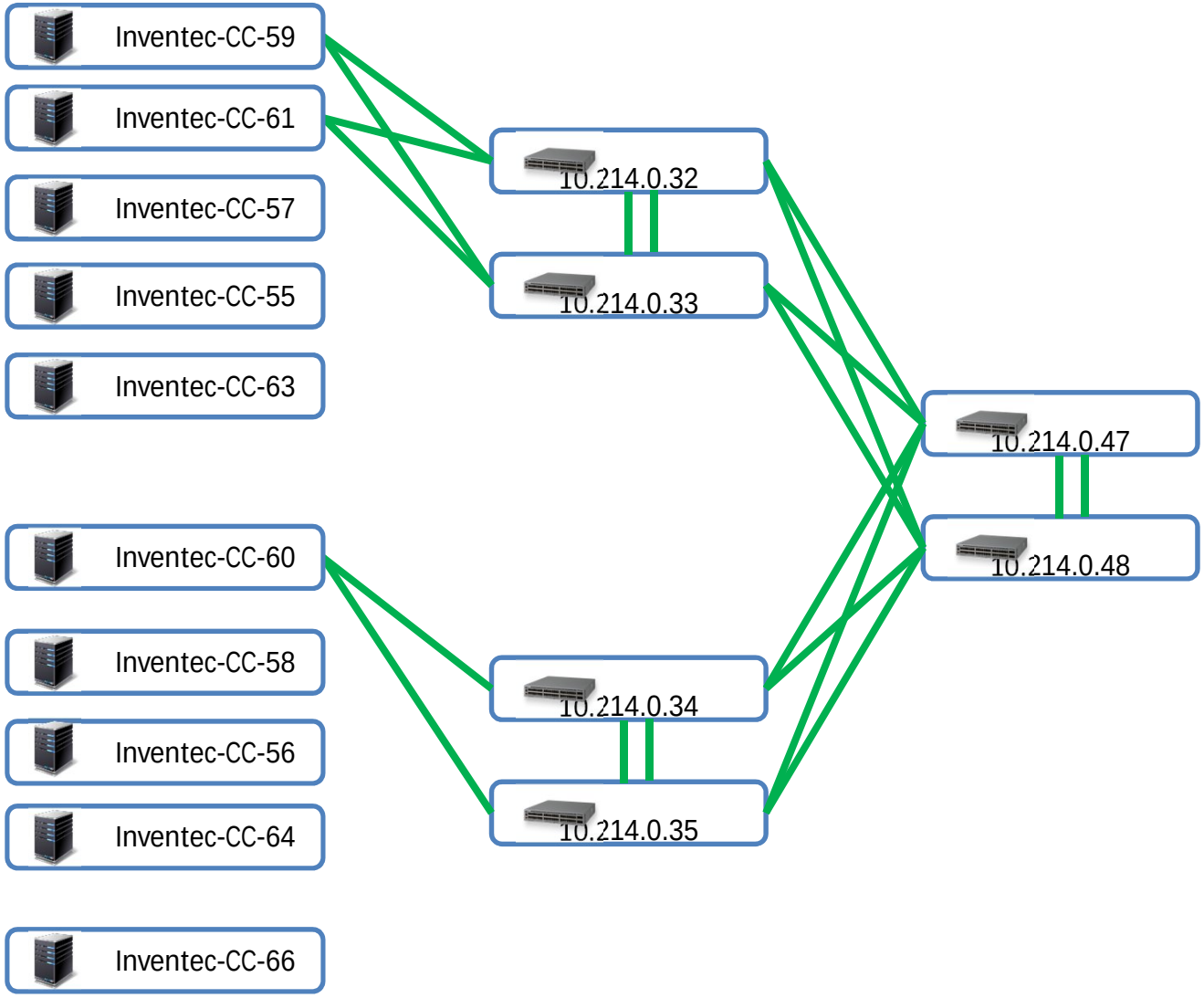
# Tenant-Aware Network Traffic Analysis

- Use case: A HaaS tenant calls to complain that the performance of its virtual network is slow, and the HaaS operator is supposed to find out the root cause, fix it, and get back to the tenant within 30 minutes.
- How?
  - Virtual to physical network mapping
  - Health monitoring of physical network devices
  - Multi-resolution network resource usage attribution
    - Which physical network links?
    - Which VLANs?
    - Which server pairs?
    - Which network connections?
    - What do these network connections do?

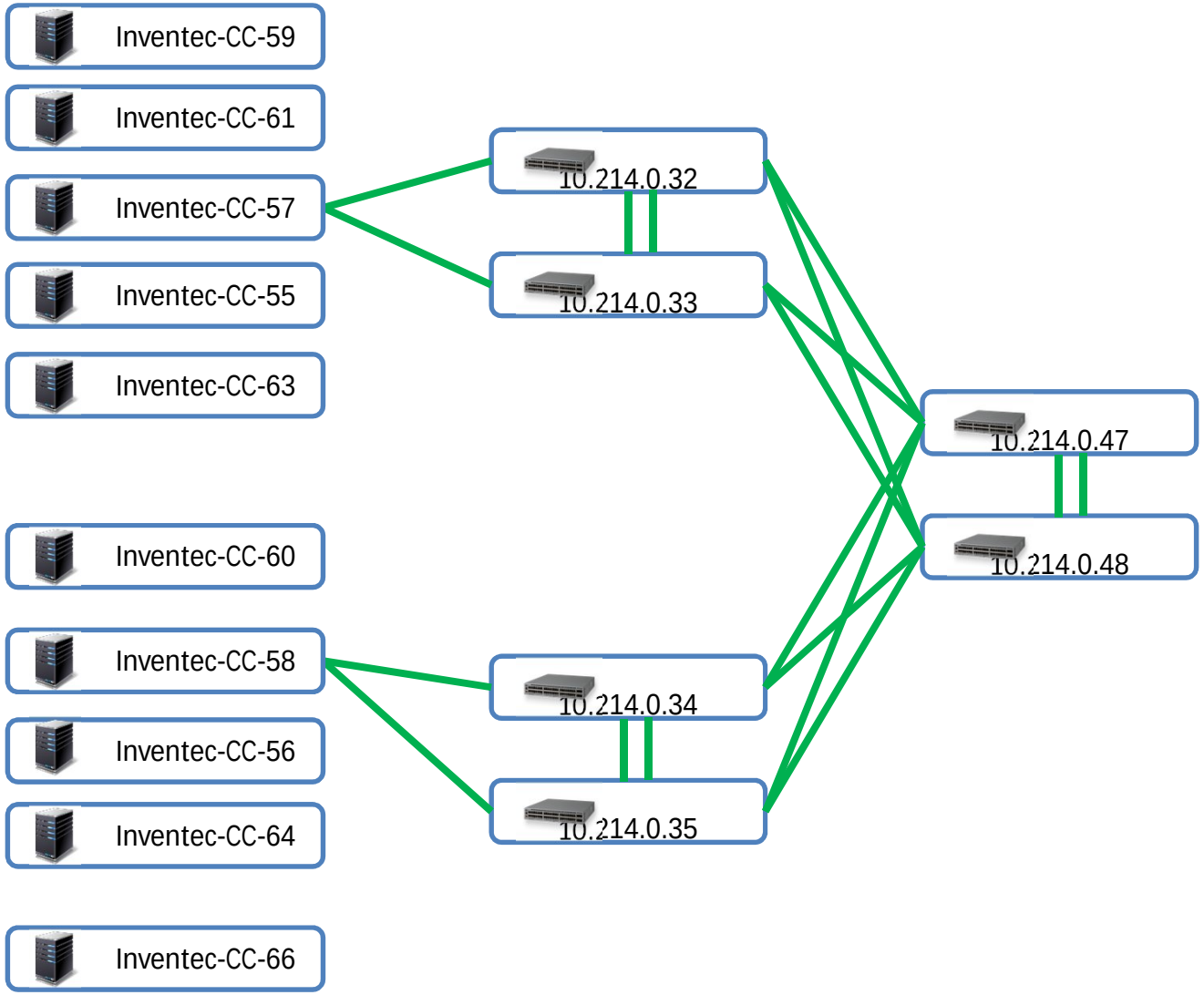
# Underlying Physical Network Topology



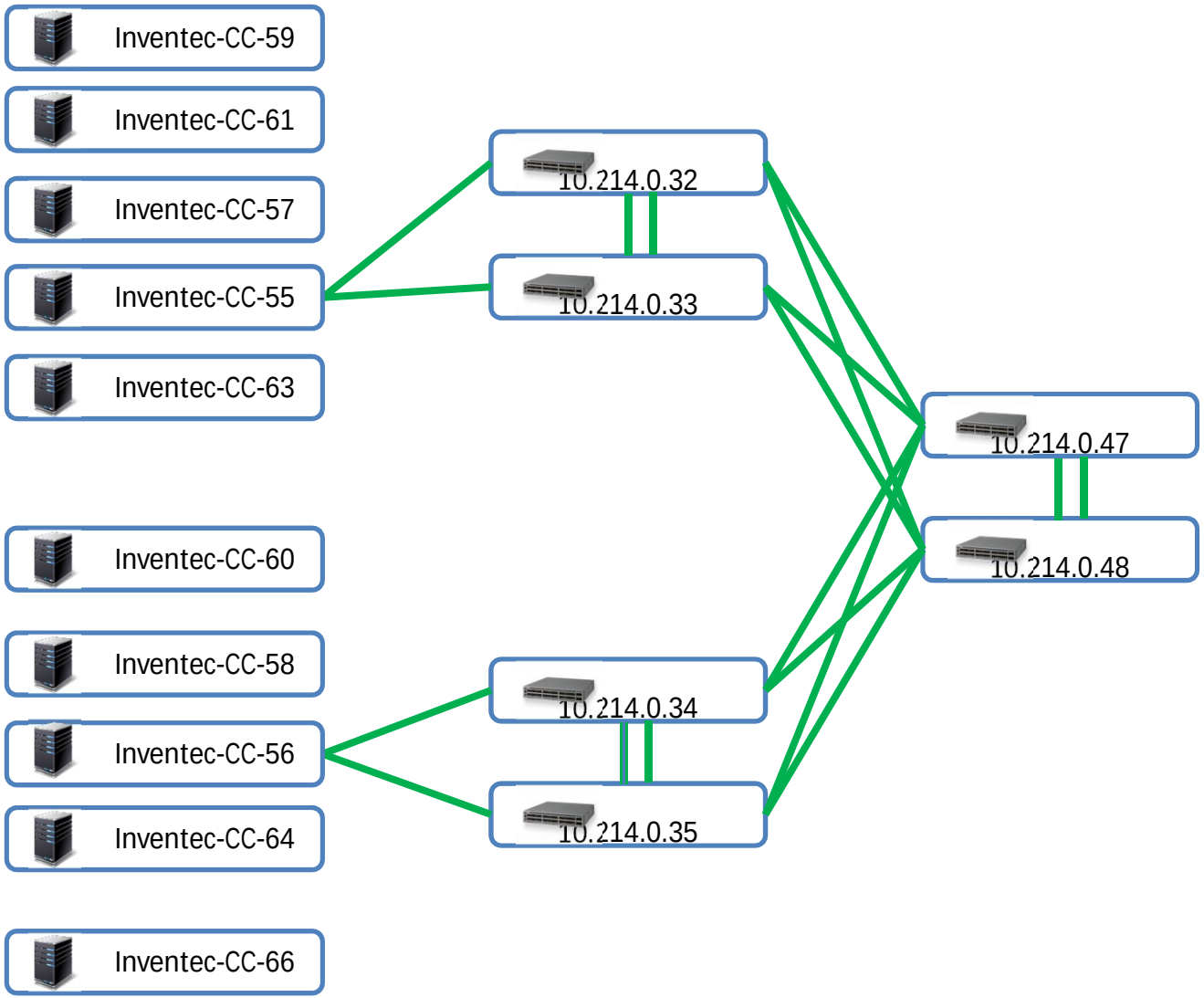
# Topology for Tenant A's Virtual Network



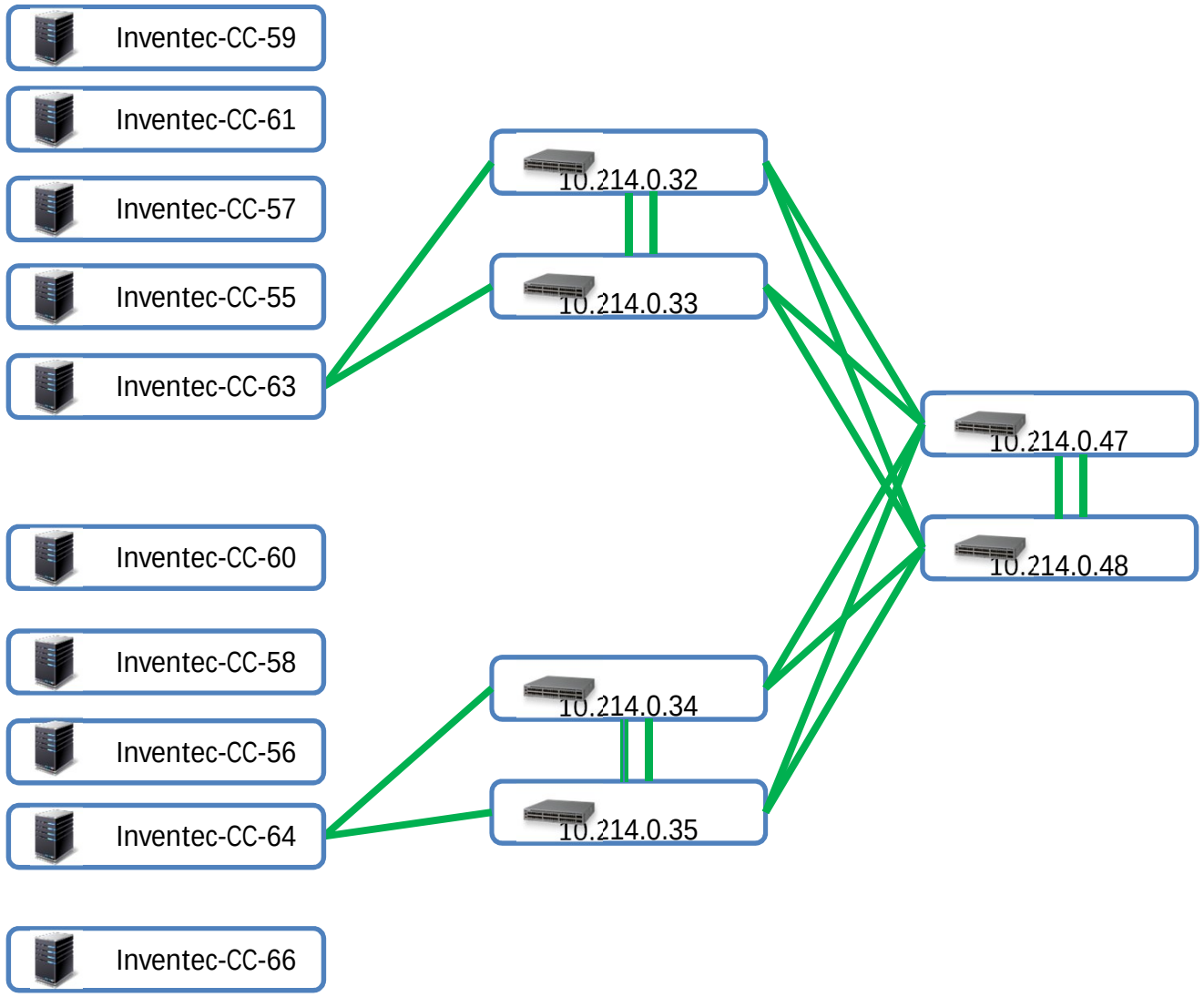
# Topology for Tenant B's Virtual Network



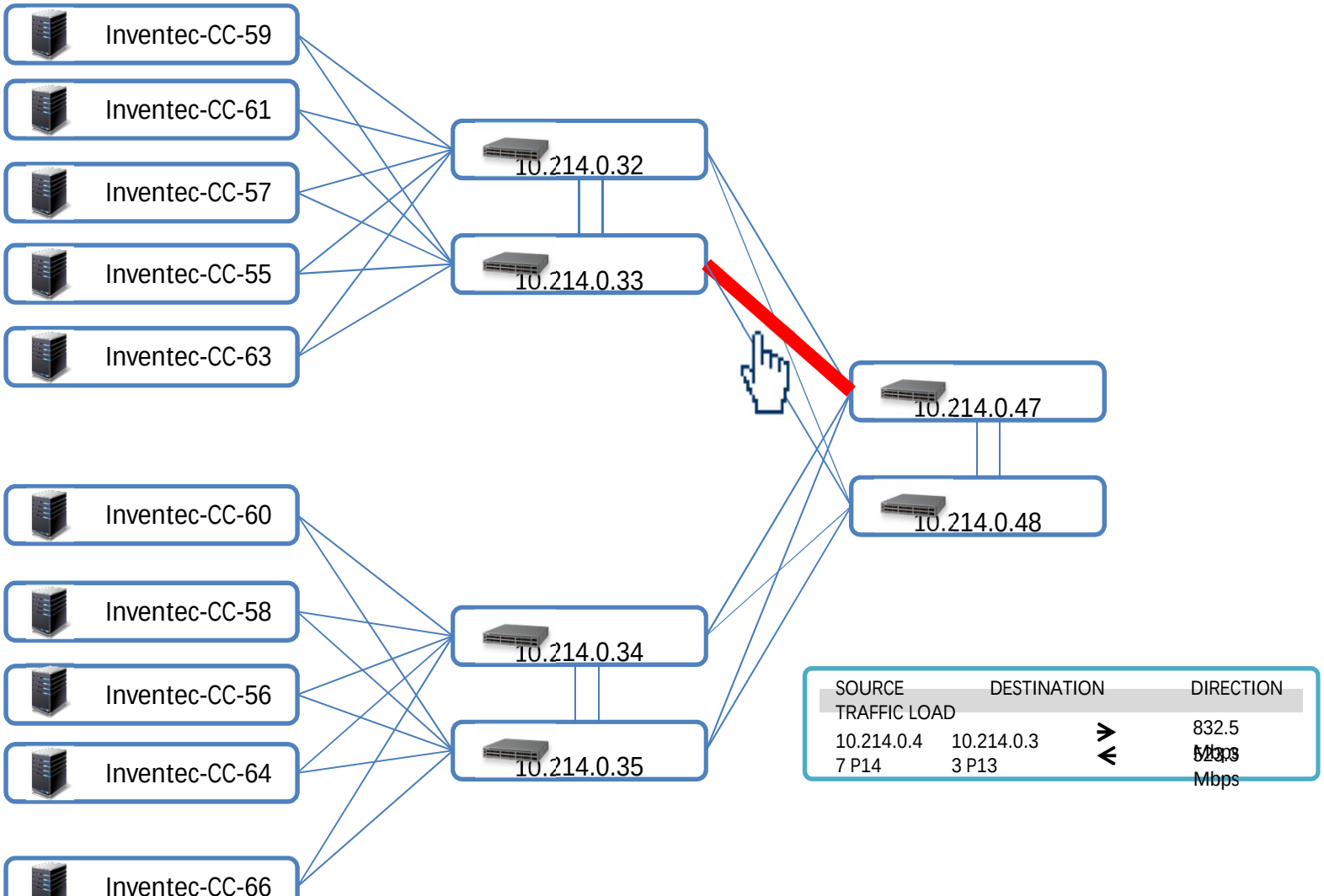
# Topology for Tenant C's Virtual Network



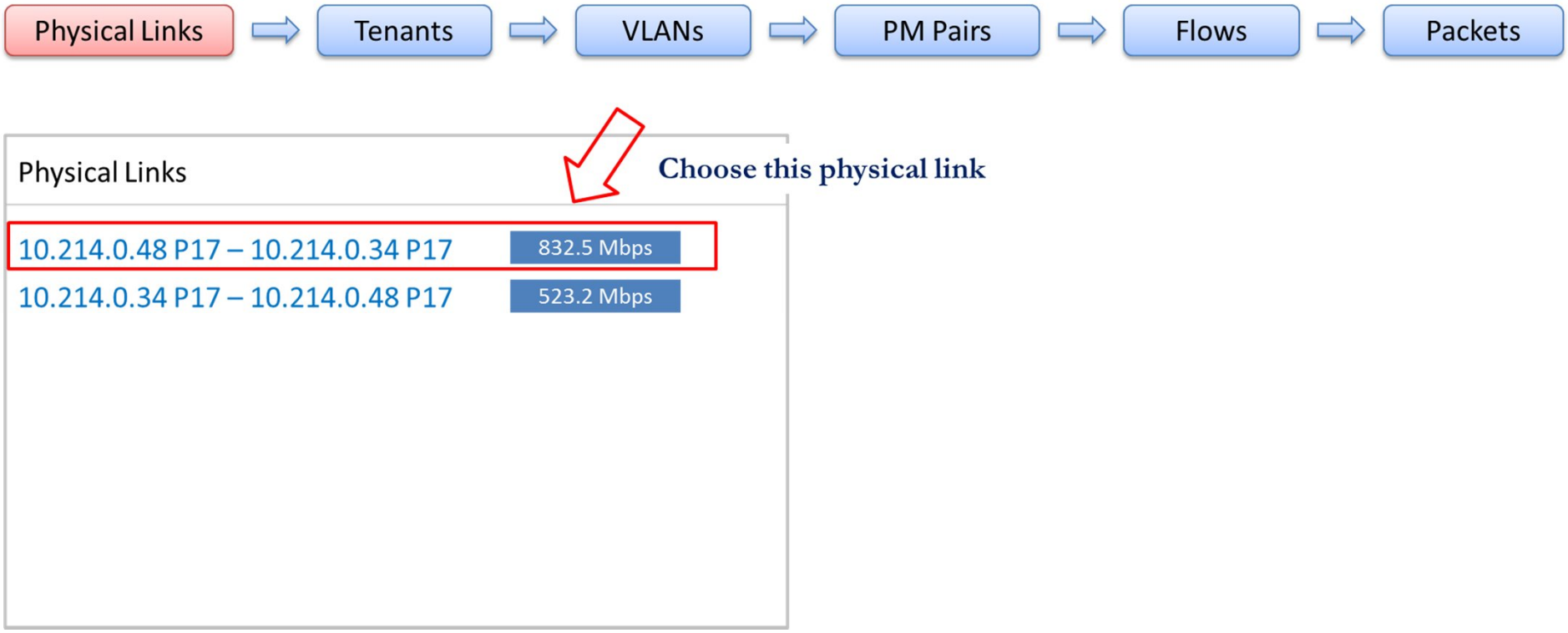
# Topology for Tenant D’s Virtual Network



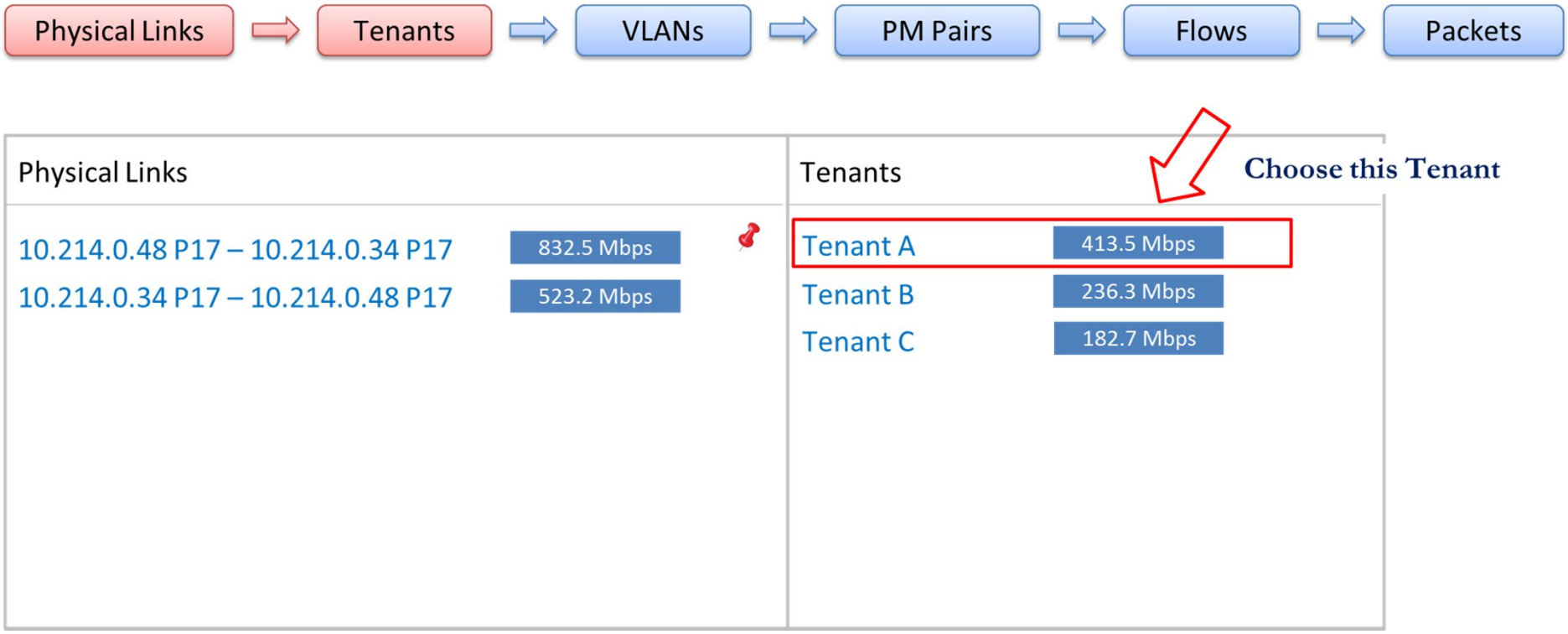
# When an Alert Is Raised



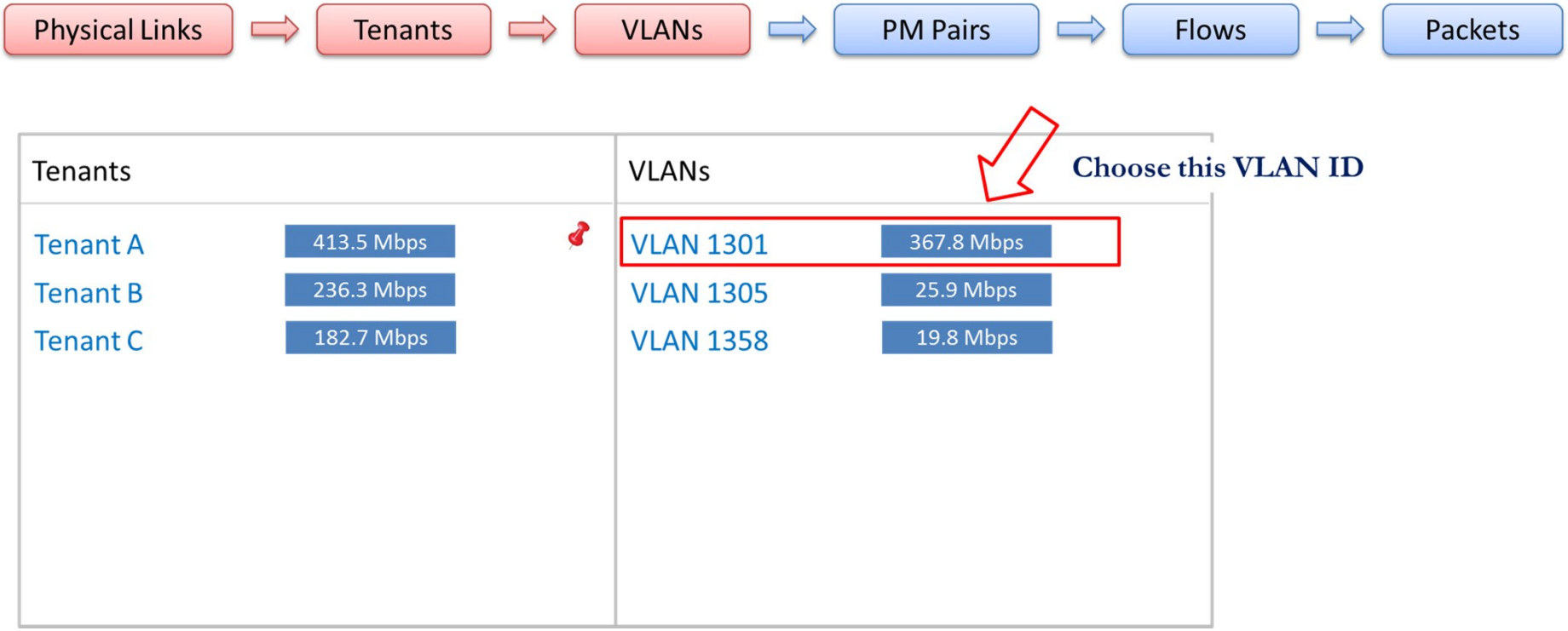
# Identify the Problematic Physical Link



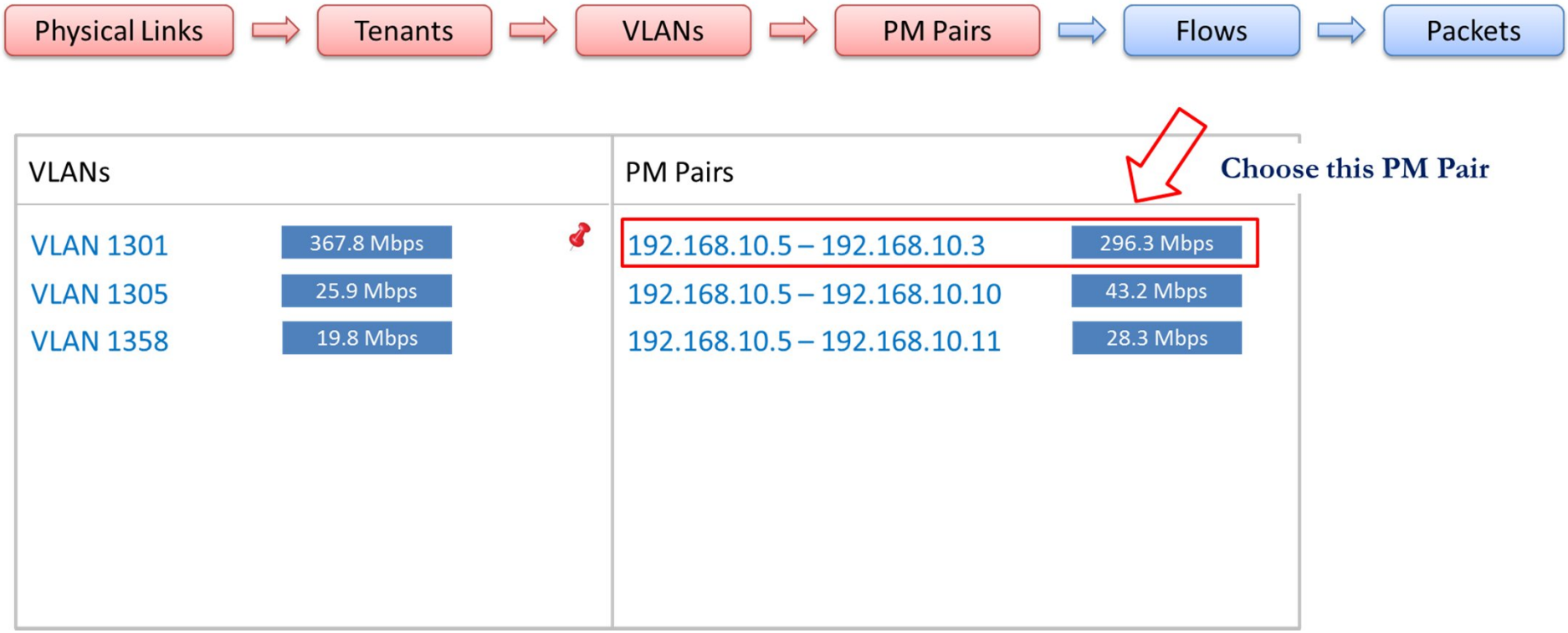
# Identify the Problematic Virtual Link



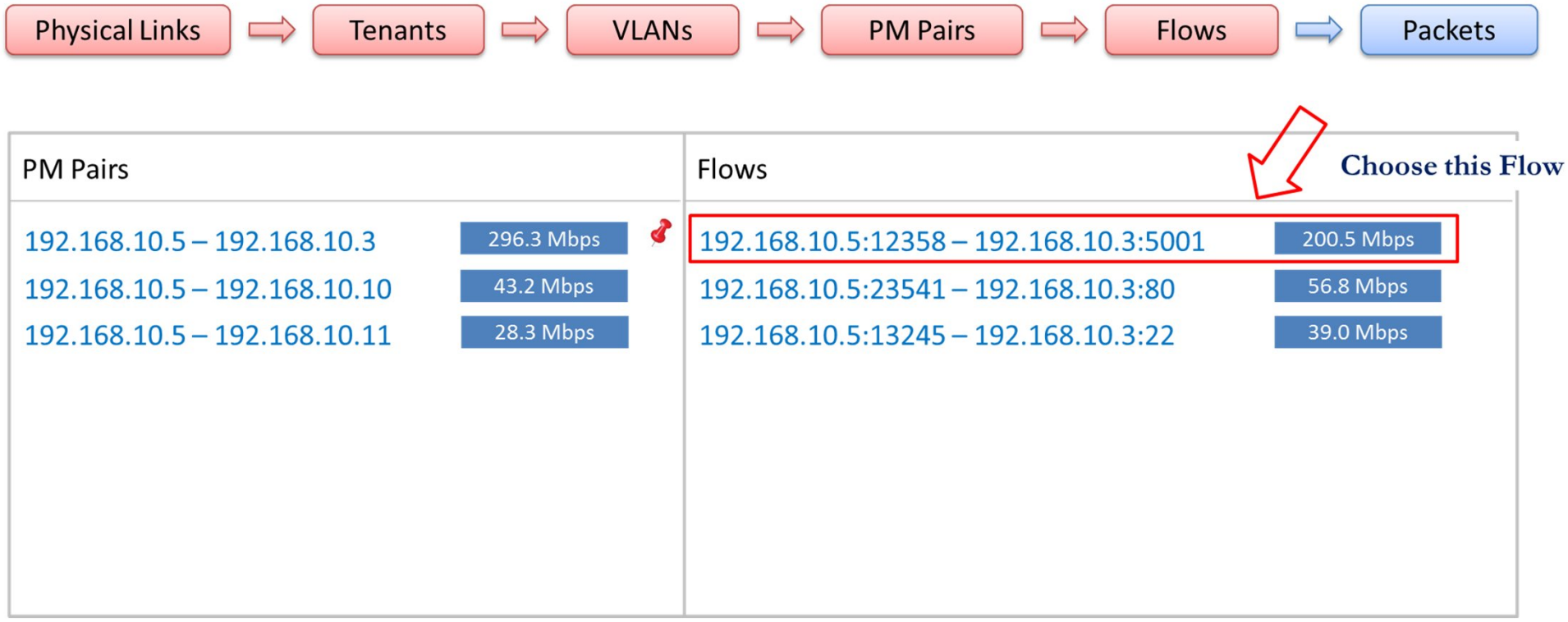
# Identify the Problematic VLAN



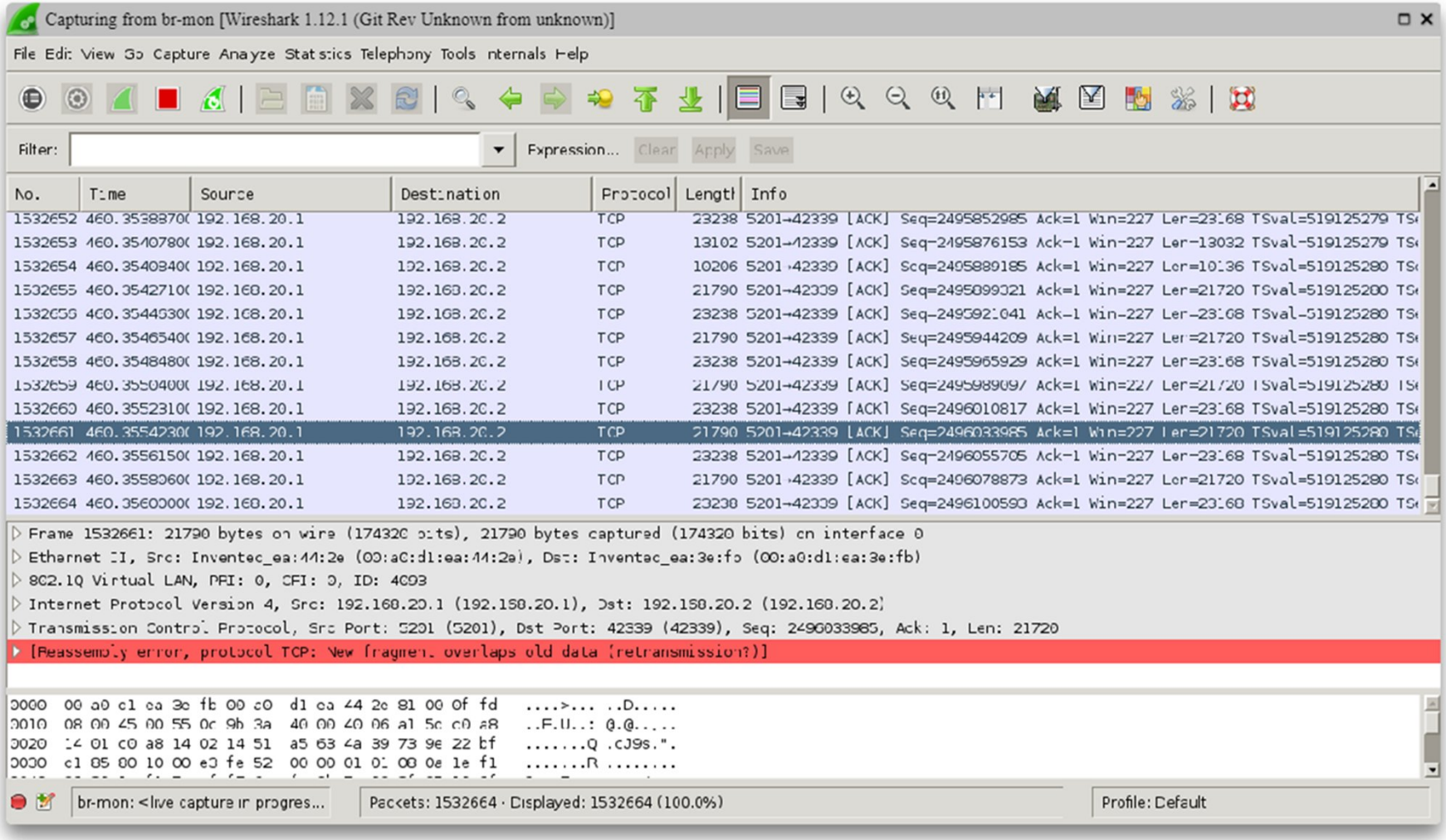
# Identify the Problematic Host Pair



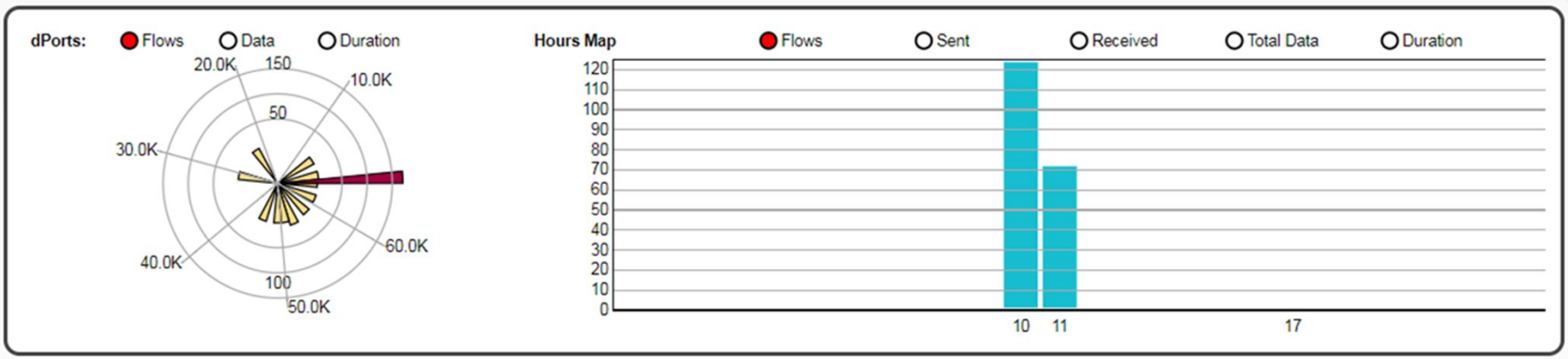
# Identify the Problematic Network Flow

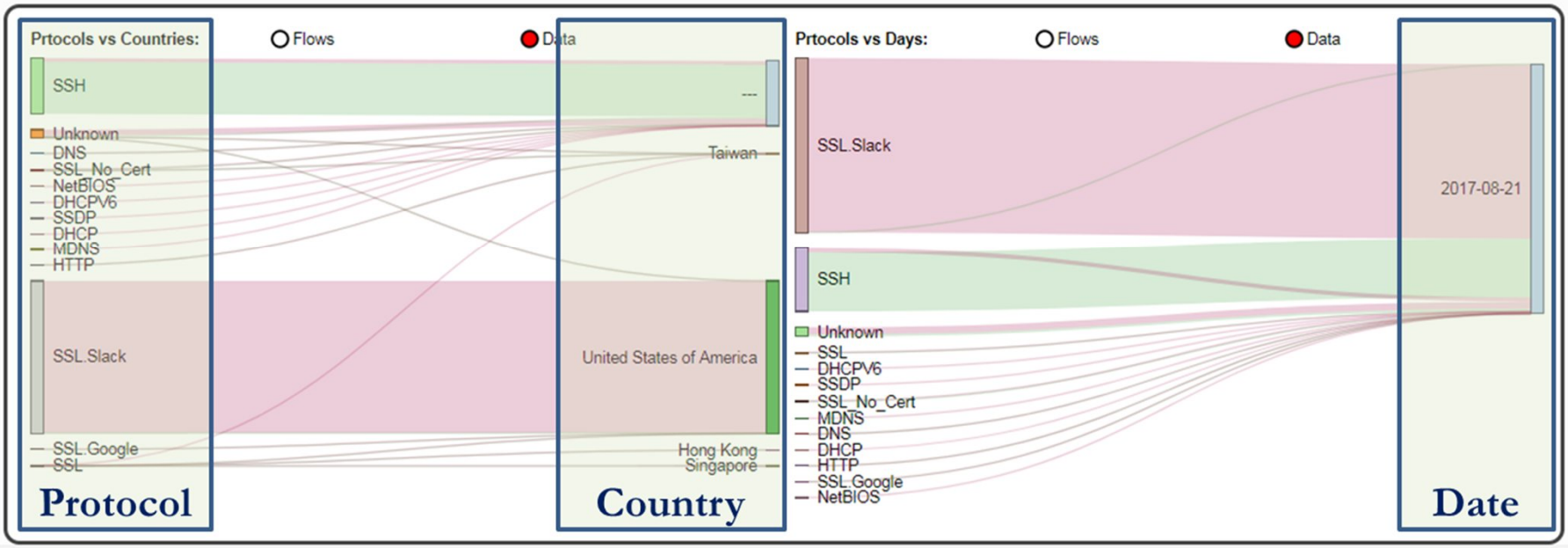


# Zoom into Problematic Network Flow



# Apply Deep Packet Inspection



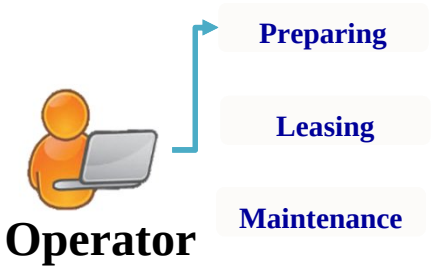



工業技術研究院
Industrial Technology
Research Institute

# Operational Support for HaaS

## HW Preparation

- A hardware usage portal for a HaaS provider operator
  - Maintenance of server HW specifications
  - Entry of hardware asset inventory list




## HaaS Operation

- A service portal for a HaaS provider operator
  - Hardware asset health status monitoring
  - Hardware asset usage by each HaaS tenant
  - How to isolate PDCIs used by multiple tenants
    - Images, packages and archived data volumes are made accessible to authorized tenants
    - PDCI preparation when released to users, and clean-up upon return
    - Root cause diagnostics of reported or detected HW problems: Alerts ➔ Problematic HW assets
    - Lest disruptive server HW replacement: replace the problematic hardware and automatically re-do the associated provisioning

# Comparison

| Metric | ITRI HaaS | IBM Softlayer | OpenStack Ironic |
|---|---|---|---|
| **Network Isolation** | Y | Y | Y |
| **Network Traffic Analysis** | Y | N | N |
| **Physical Network Load Balancing** | Y | N | N |
| **Agentless Implementation** | Y | Y | Y |
| **Interoperability with multiple switch types** | Y | Unknown | N |
| **Compatibility with OpenStack** | Y | N | Y |

# Summary

- DNN is expected to play a critical role in the emerging AI revolution across a wide range of industry segments.
- DNN training will become a major workload on future private and public cloud computing systems.
- DNN training appliance is a key enabling building block for corporations to apply DNN to make sense of their data and/or improve their businesses.
- DNN training as a service = DNN training appliance + HaaS

工業技術研究院
Industrial Technology
Research Institute

Thank You!

Questions and Comments?

tcc@itri.org.tw