Perceptive Routing

Tianran Zhou

zhoutianran@huawei.com

Qinru Shi

shiqinru@huawei.com

Lookback: IETF Activities to Meet the Cloud DC Requirement

Cloud DC Requirements

- **Virtualization**: Virtualization allows multiple virtual machines to run on a single physical server, increasing efficiency.
- **VM migration**: VM migration enables the seamless movement of virtual machines between servers for load balancing and maintenance.
- **Massive number of tenants**: Cloud providers host numerous tenants, each requiring isolated and secure network environments.

IETF Activities

- **NVO3 Working Group**
  - **Solutions like VXLAN**: VXLAN helps extend Layer 2 networks across Layer 3 boundaries using overlay tunnels.
  - **Overlay tunnels over UDP**: These tunnels encapsulate Ethernet frames within UDP packets.
  - **Provides Virtual Network Identifiers (VNIs)**: VNIs uniquely identify and isolate tenant networks within shared infrastructure.
- **RIFT and LSVR Working Groups**
  - **Focused on fat-tree topologies**: Fat-tree topologies provide high bandwidth and low latency in modern data centers.
  - **Fast and efficient routing computation**: These groups develop protocols for quick and robust routing in complex networks.

Several Changes since AI Training

- **New Topologies**
  - Dragonfly+
  - Multi-rail fat tree
  - 3D Torus
- **New Traffic Models**
  - **Large Packets**: AI workloads typically involve the transmission of large data packets (4KB), unlike traditional data center traffic which may consist of smaller, more frequent packets.
  - **High Volume Traffic**: The nature of AI tasks, such as training machine learning models, generates significant amounts of data (X GB data), resulting in high traffic volumes.
  - **Fewer Flows**: Compared to traditional data centers, AI data centers handle fewer but larger data flows, reflecting the intensive and concentrated data transfer needs of AI applications.
- **New Characteristics**
  - Predictable Topology
  - Predictable Traffic

→ **More room for optimization**

New Requirements for AI DC

Key Challenges in AI Data Center Networks:

- **High-Performance Interconnect**
  - High Bandwidth: 200Gbps~400Gbps
  - High Scale-out Capability: thousands of GPUs
- **High Availability**
  - Load Balancing: avoid hash collision
  - Congestion Control
  - Fast Convergence
- **Operations and Maintenance**
  - Refined Monitoring
  - One-click Fault Diagnosis
  - Self-healing Mechanism
- **Security**
  - Data Confidentiality and Integrity: especially in multi-tenant environments

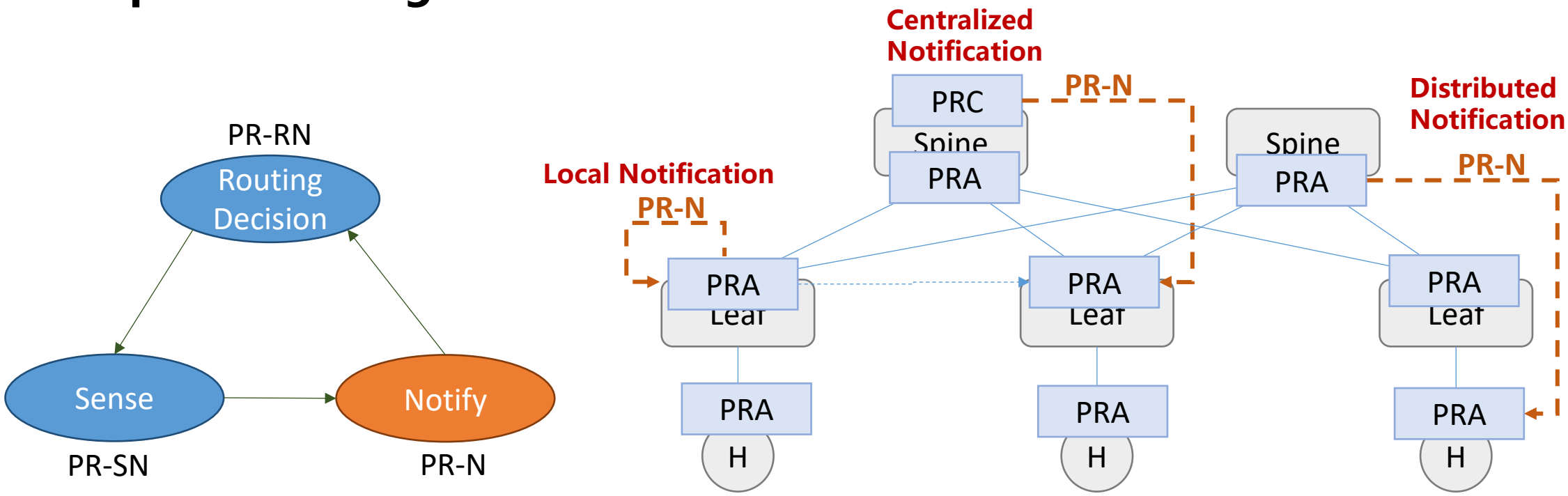
Proposal: Perceptive Routing

- **Current Gaps**:
  - **More Information:** Devices need more info to make more informed routing decision.
  - **Faster Reaction:** Need faster reactions to congestions or failures.
  - **Easier Setup:** Need easier development and setup of new routing mechanisms.

- **Proposal:** Set up **a standard notification framework** to facilitate the multi-dimensional sensory information.
  - Maximize utilization of existing sensory tools

- By enhancing the network's awareness capability, it can **make informed routing decisions** to improve efficiency, reliability, and scalability.
  - Failure-Aware
  - Congestion-Aware
  - Service-Aware
  - Tenant-Aware
  - …

Perceptive Routing Framework



PR-SN: PR Sensing Node, percept local and network information for routing decisions.

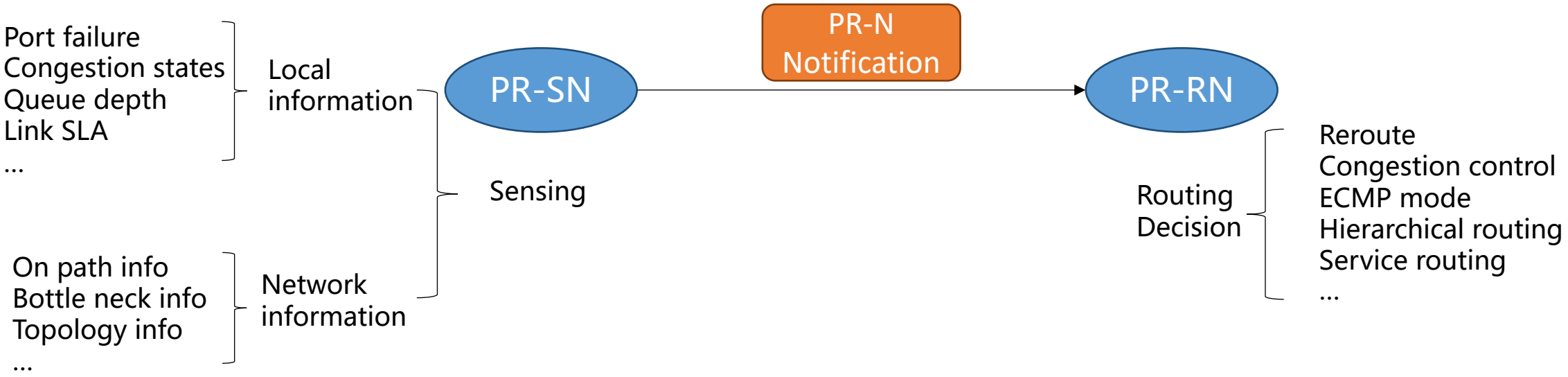
PR-RN: PR Routing Node, use multi-dimensional sensory information to make routing decisions, including reroute, adjust speed, load balance, etc.

PR-N: PR Notification, the message from PR-SN to PR-RN.

PRA: PR Agent

PRC: PR Controller

Work on Various Notifications



Encapsulations:
BGP, IGP, L2, UDP, New msg, PHY

PR-N:

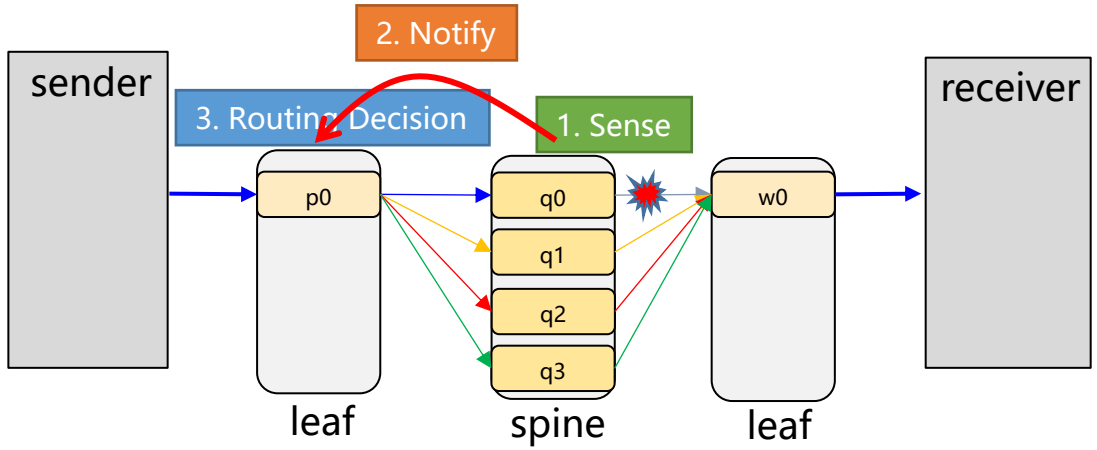
Information Model:
Need to develop

Failure Awareness for Fast Reroute

UC：Distributed training tasks require frequent, high-volume, and efficient communication. Failure-aware nodes must quickly notify and converge routes to minimize impact of network failures and prevent application performance degradation due to packet loss.

Procedure:

1. Continuously **Sense** state changes, such as link or port failures (down)
2. When state changes exceed thresholds, devices send immediate **Notifications** to upstream devices.
3. Utilize failure-specific information to dynamically **adjust routing decisions** (switching flows to healthy paths).



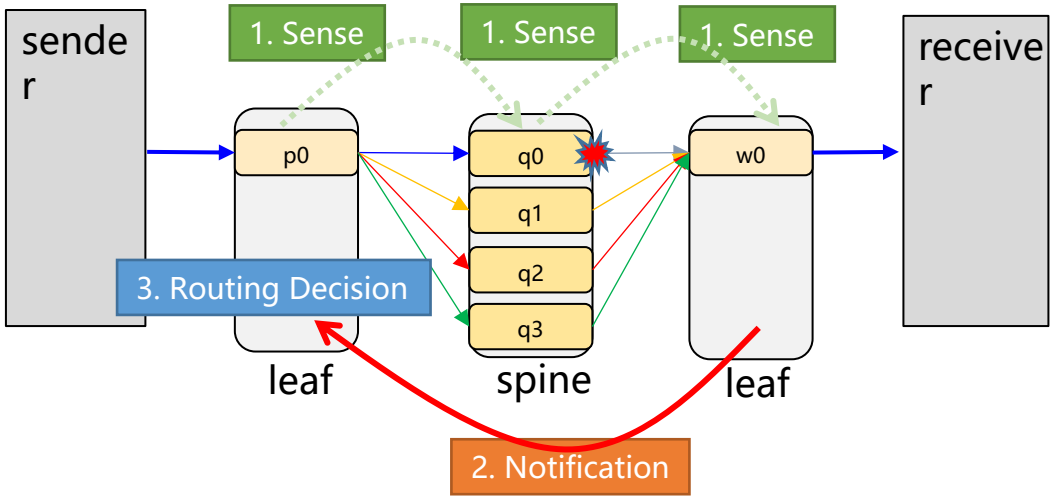
IETF drafts
draft-cheng-rtgwg-ai-network-reliability-problem
draft-liu-rtgwg-path-aware-remote-protection-01
draft-wh-rtgwg-adaptive-routing-arn

Congestion Awareness for Load Balancing

UC：The high volume and simultaneous point-to-point communications can easily cause congestion. While network devices can quickly detect congestion using various existing technologies (e.g., CSIG/INT/IOAM), there is a lack of standardized methods to notify the ingress/source, allowing subsequent traffic to avoid congested points and leverage the multipath capability of networks to alleviate congestion.

Procedure：

1. **Sense** congestion information: Continuously gather congestion data (e.g., IOAM, CSIG)
2. Immediate **notification** right after sensing congestion: Quickly notify upstream network devices or source nodes to react to the congestion.
3. Upstream devices or source nodes promptly **reroute** (switch the flows to alternate paths to avoid congested points or apply proactive rate limiting).



IETF drafts
draft-wh-rtgwg-adaptive-routing-arn
draft-lyu-rtgwg-coordinated-cm
draft-xu-idr-fare
draft-xu-lsr-fare
draft-wang-rtgwg-dragonfly-routing-problem
draft-agt-rtgwg-dragonfly-routing
draft-cheng-rtgwg-adaptive-routing-framework
draft-liu-rtgwg-adaptive-routing-notification

Service Awareness for In-Network Computing

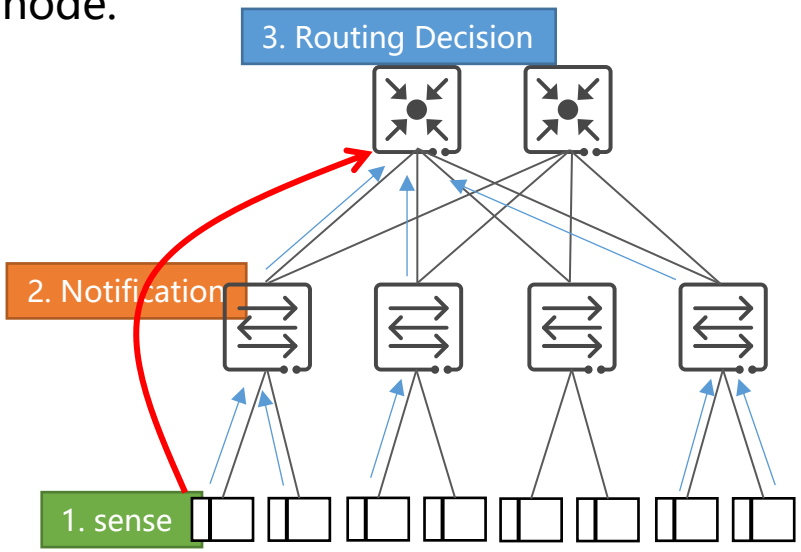
UC: Switches naturally sit at the center of a network, and thus can aggregate information with less data transmission. Offloading collective communication operations such as AllReduce to the network can significantly improve the efficiency of AI training.

Current in-network computing method:

- xPUs that need to perform a communication operation send request to switches or controller;
- Switches or controller build an aggregation tree based on the request(s) and allocate resources;
- Root switch or controller notify the xPUs to send data;
- INC switches aggregate data from child nodes and send to parent node.
- Root switch distribute the aggregated result to all relevant xPUs.

Analysis:

- **Sense**: xPUs sense job information and resource requirement.
- **Notification**: xPUs send aggregation requests to switches or controller.
- **Route**: INC switches build aggregation trees based on the request, and then aggregate and route the data packets based on the structure of the aggregation tree.



IETF drafts
draft-lou-rtgwg-sinc
draft-liu-nfsv4-rocev2

Service Awareness for Routing method differentiation

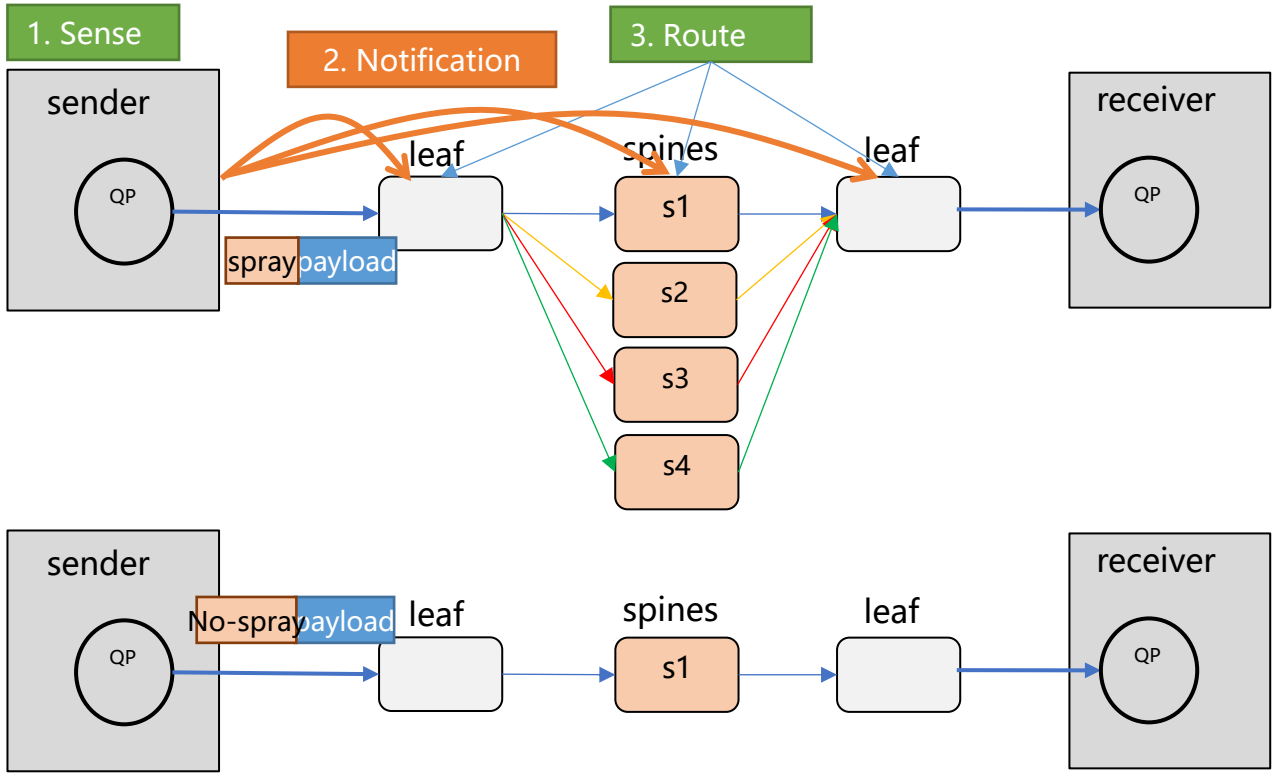
UC: A network could use multiple strategies to send a flow. For example, the network could use packet spraying to achieve better load balance at the cost of reordering; Or the network could send all packets of a flow on the same path to preserve packet order. For different types of flows, the optimal strategy may be different. However, network clusters usually could only employ one routing strategy at a time, as the strategy need to be configured on each device.

Proposal:

- Packets carry info of preferred routing strategy
- Network devices employ corresponding strategy

Procedure:

- **Sense:** Sender or TOR senses the characteristics of a flow and decide the best routing strategy.
- **Notification:** Routing strategy info is carried in data packets throughout the network path.
- **Route:** Network devices then route packets based on the routing strategy information, allowing the network to employ different strategies for different types of flows.



IETF drafts

draft-lyu-rtgwg-coordinated-cm

Next Steps

- Collaborations:
  - Use cases
  - The information model for the notification
  - New messages
  - Encapsulations in related working groups: IDR, 6MAN, etc
- BoF at IETF121?

Thanks