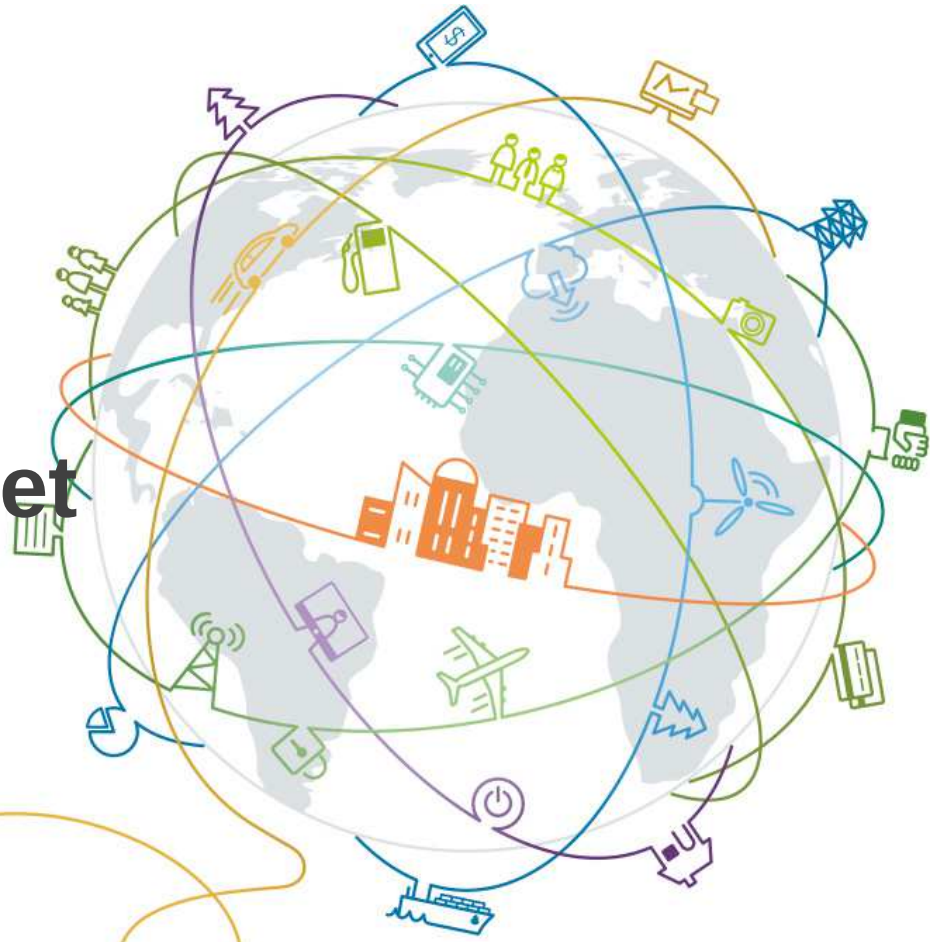
IP 2020:
On The Road Towards Next Generation Internet



Yingzhen Qu
Senior Research Scientist, Future Networks
Huawei Technologies, US Research Center

October 2017

Huawei at a Glance

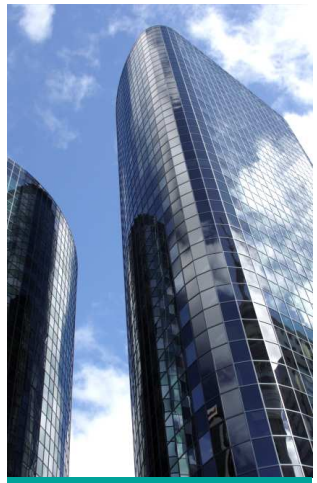
180,000
Employees

80,000
R&D employees

170+
Countries

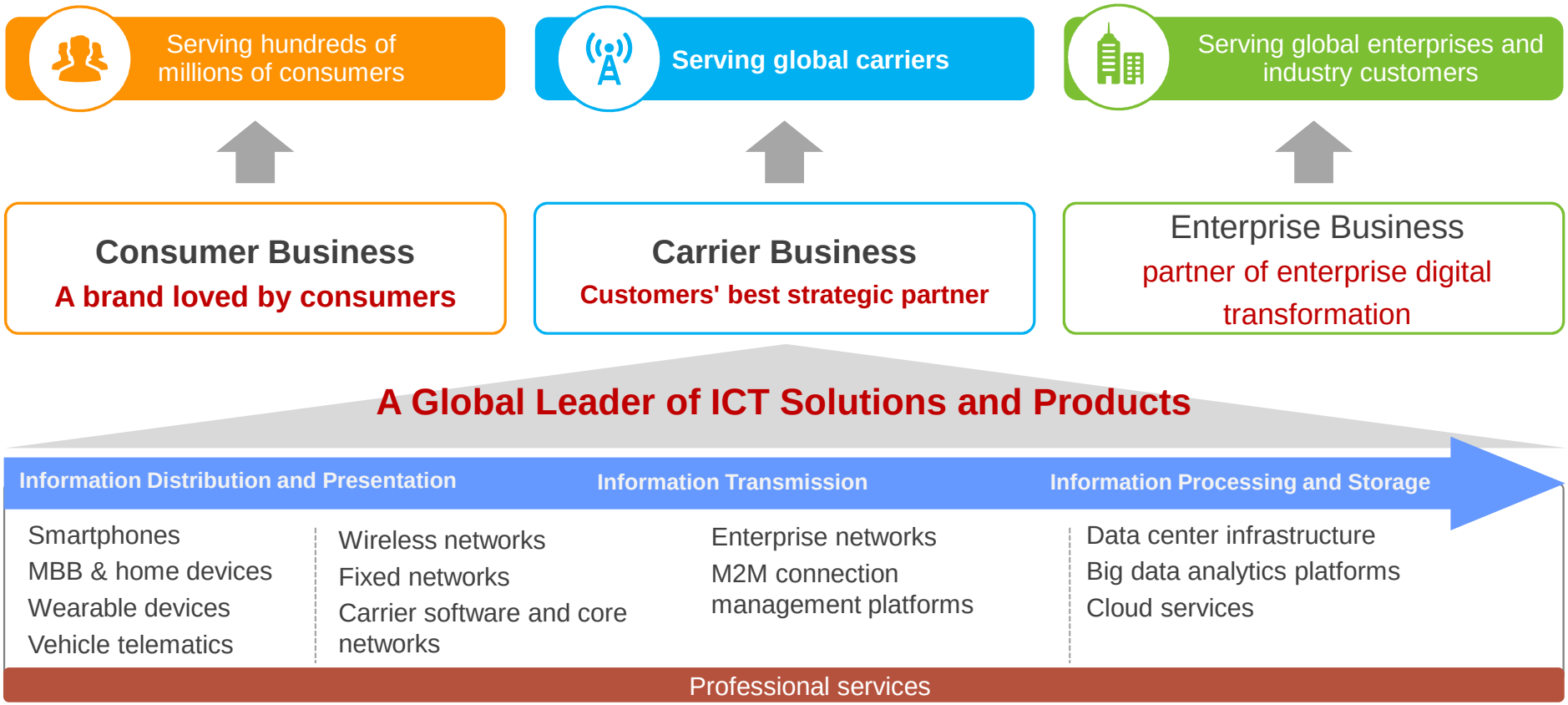
15
R&D institute and centers

No. 72
Interbrand's Top 100 Best Global Brands

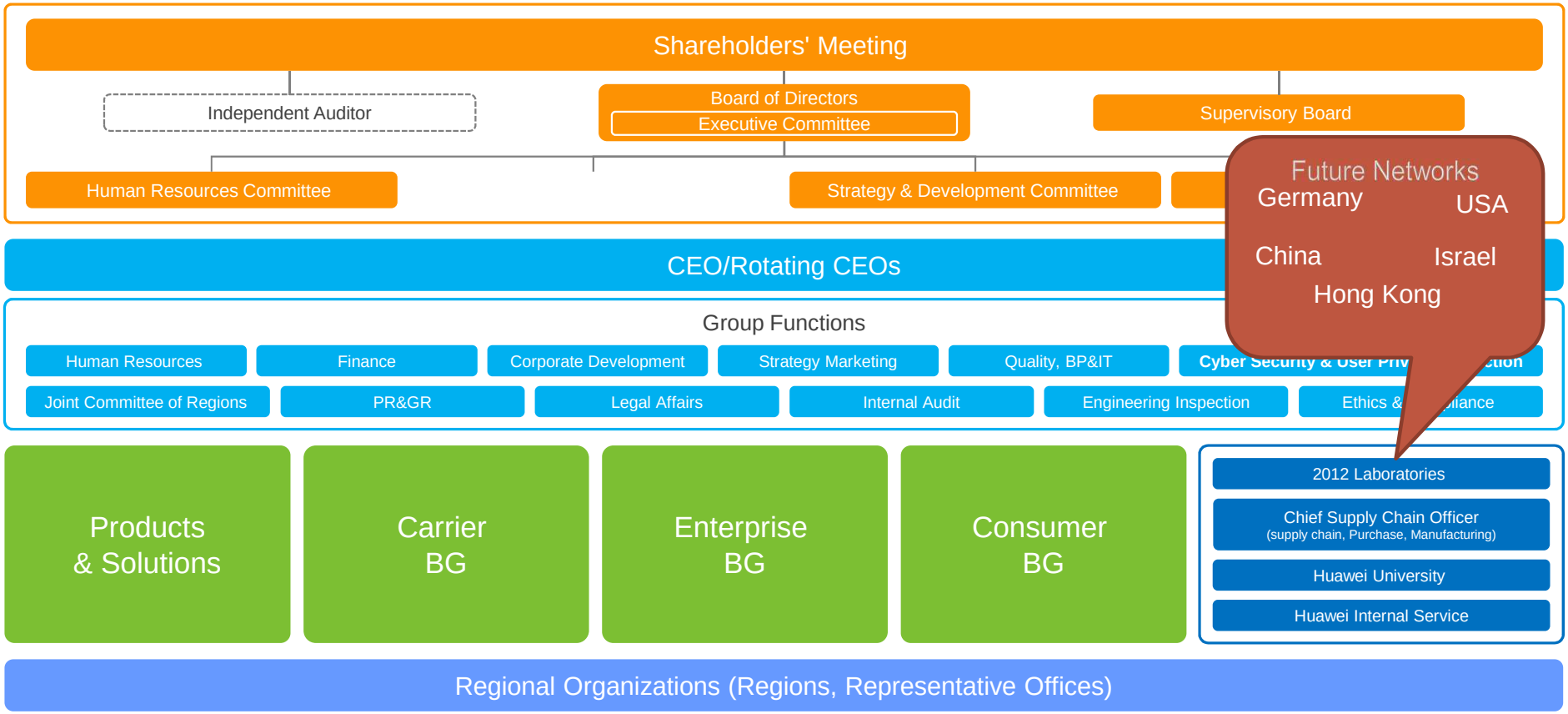


No. 129
Fortune Global 500

Provide ICT Solutions and Services for Three Customer Groups



Corporate Governance Structure



Research at Huawei's

- Huawei has consistently invested over 10% of its revenue in R&D every year.
- In 2016, approximately 8000 employees, 45% of total workforce were engaged in R&D.
- The Huawei Innovation Research Program (HIRP)
  - An Open program that offers funding opportunities leading universities and research institutes
  - Conducting innovative research in the field of communication technologies and computer science
  - http://innovationresearch.huawei.com/IPD/hirp/portal/index.html
- Active Collaboration with many universities in the US

University Of Maryland

Harvard

Winlab @Rutgers

MIT

University Of California Riverside

University Of California Berkley

Stanford

Agenda

- A review of Internetworking Technologies and Problems
- IP2020
  - ID Enabled Networks
  - Deterministic Transport
  - Intelligent Data Center Networking
- Summary

A Review of Internetworking Technologies and Problems

IP Has Been So Successful, But Where Did It Come From?

Paul Baran, RAND, US

Inventor Packet Switching

1961

Frank Heart &Team,

BBN IMP Spec

1968

Queen Elizabeth II

Sends her first email.

1976

TCP/IP Standardized

1980

ARPANET ceased

1990

Conceptual — Experimental — Standardization — Explosion

1965

Inventor Packet Switching

Don Davis, NPL, UK

1974

Cyclades at INRIA, France, 1971-1979,

Louis Pouzin

1979

IEEE, "A Protocol for Packet Network Intercommunication" , 1974

Vinton Cerf and Bob Kahn

1989

WWW, CERN

Tim Berners-Lee

INWG Standards

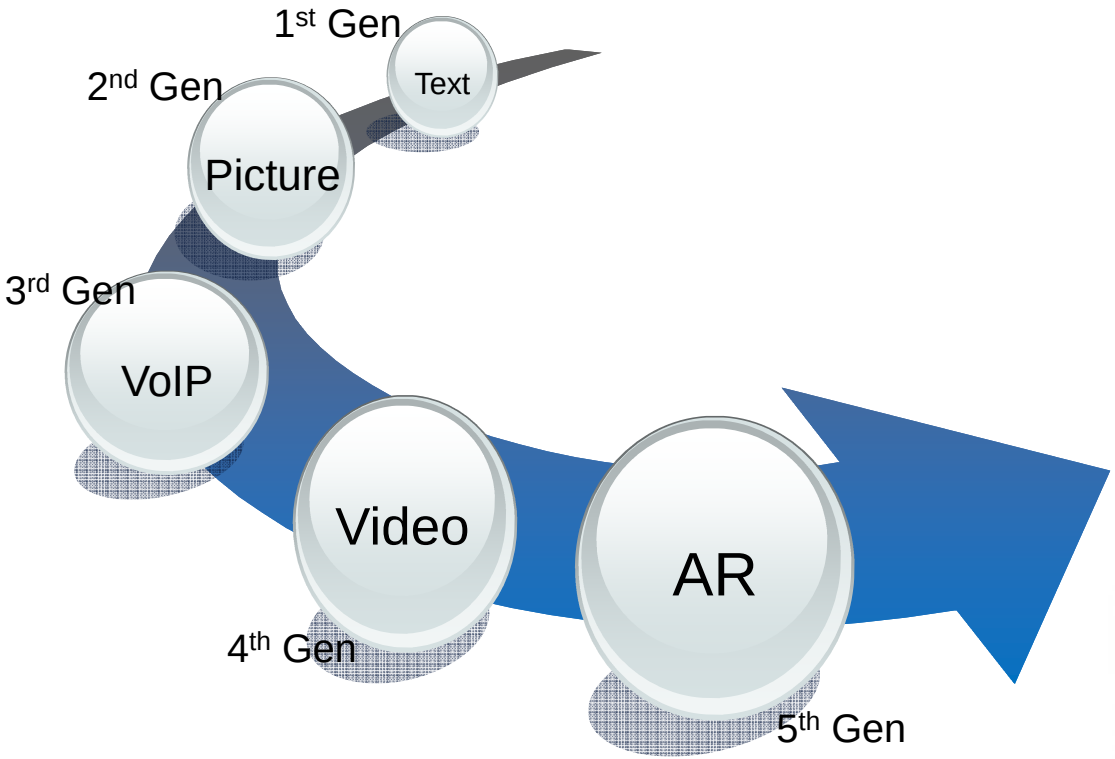
1973 — Host & Process Level protocol — INWG39

1974 — Standard HOST/HOST Protocol — Hubert Zemmermann — INWG61

1976 — Cerf (INWG 39 and ARPA), Zimmermann (INWG 61 and Cyclades), Roger Scantlebury (NPL and EIN) — Alex McKenzie — INWG 91

Packet Switching | Landmark Projects | Commercialization

Media Evolution: It's all about experience



Ultra-high Throughput

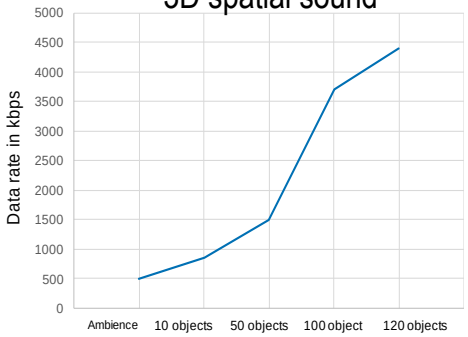
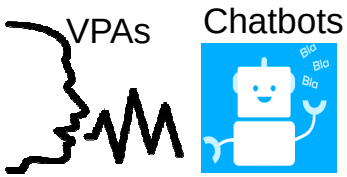
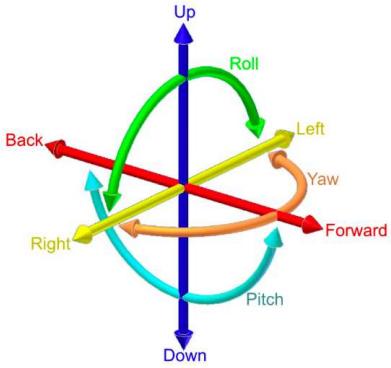
360 vision needs 3~5x than single vision

Panoramic VR needs 25Mbps to 5Gbps

MTP Latency < 20ms

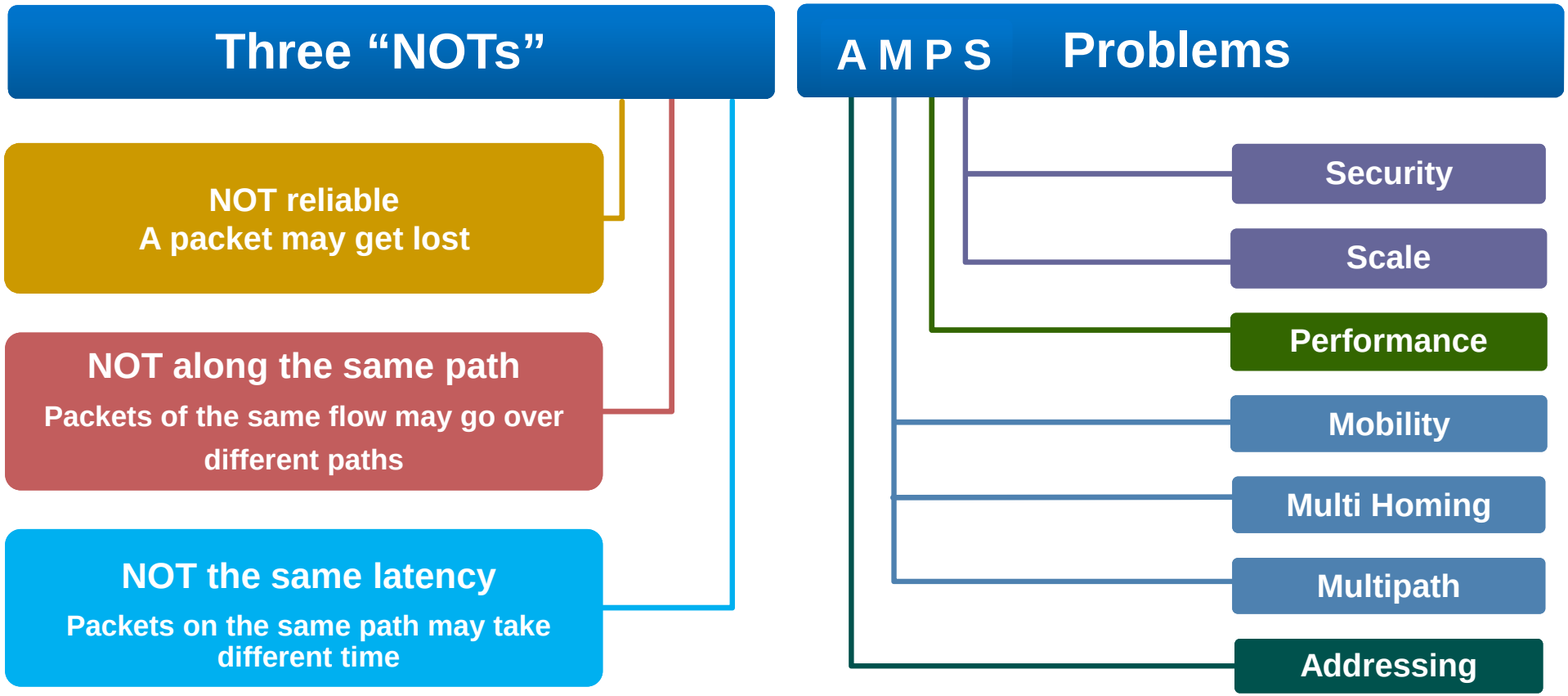
Motion-To-Photon Latency < 20ms

Immersive content consumption is cool! But demanding

Sound	Speech	Video	Touch
5,644,8 kbps Super Audio 6,144 kbps AC3 9.6 Mbps DVD-Audio	50% - Voice based search 85% customer service - chatbots	30 fps, 100 Mbps - Basic VR 60 fps, 400 Mbps – Adv. VR 120 fps, 1000 Mbps – Ult. VR	Provide the medium for transporting touch and actuation in real-time
3D spatial sound	VPAs Chatbots	1: MTP < 20ms 2: Throughput > Gbps	Sampled at 1 kHz leading to 1000 packets per second
			Joystick, Haptic wearables, vibrations
Object Based Representation (Dolby)	Speech to text translations	Tradeoffs Between Six Degrees Of Freedom Latency 5-20ms – 400-600 Mbps Latency 1- 5ms – 100-200 Mbps	

Source Qualcomm-AR-VR

Observations on IP and Statistical Multiplexing

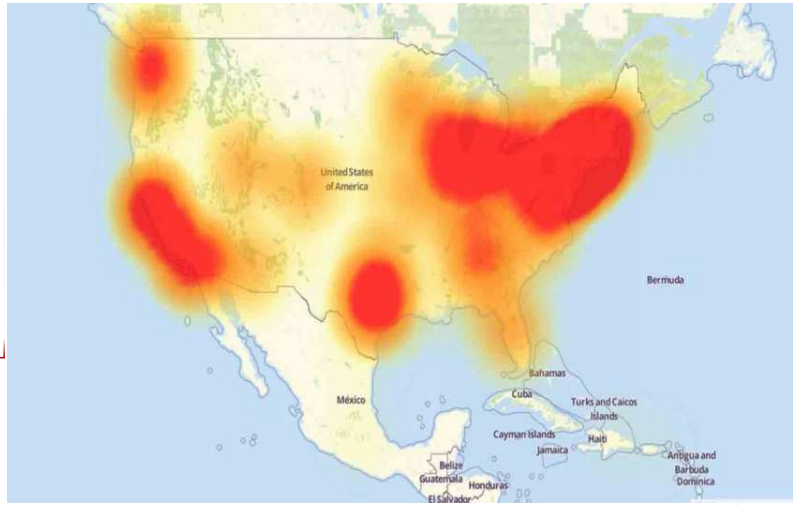


Missing – Security. Internet Is Fragile

Uncontrollable Malware Spread At The Scale Of IoT

21 Oct 2016 DDoS Attach at Dyn.
Up to 10,000 IoT Devices involved

First [7 AM]
Second [noon]
Third [4 PM][4]



Massive Outages Due To Configuration Errors

Amazon Outage of 28th Feb 2017 (Typo Error)[5]

"Unfortunately, one of the inputs to the command was entered incorrectly and a larger set of servers was removed than intended," the Amazon note states.

Identity Thefts and Data Breaches

Between 2013 and 2016 Billion accounts were hacked – thrice.[6]

Yahoo hit in worst hack ever, 500 million accounts swiped

4:[DDoS] : http://money.cnn.com/2016/10/21/technology/ddos-attack-popular-sites/

5. http://money.cnn.com/2017/03/02/technology/amazon-s3-outage-human-error/

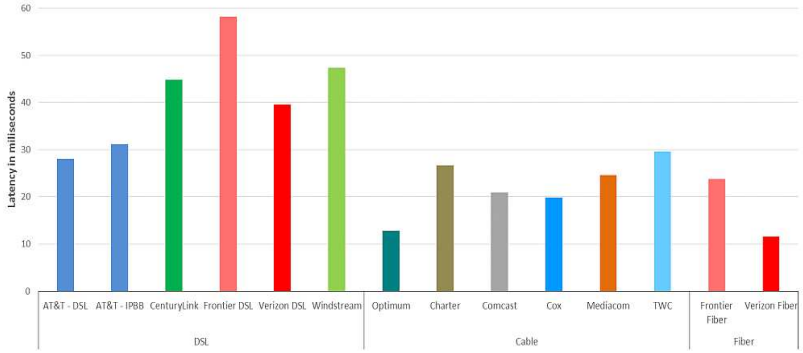
6 http://www.cnbc.com/2017/02/15/yahoo-sends-new-warning-to-customers-about-data-breach.html

Missing – Service guarantees

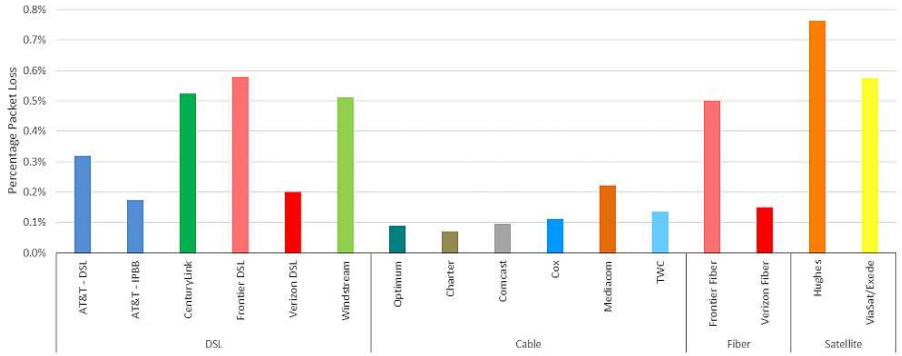
Non-existent Service Level Agreements for Residential Users

Residential Services have no SLA [REPORT][3]

Shared bandwidth with other customers that may degrade some application performance



packet loss = latency exceeds 3 second. Cable: 0.1%



Latency : 12ms-58ms

Effects of Over Subscription upon congestion

- Saving Cost: Divert traffic on already optimally used paths ➔ may cause congestions on existing flow
- Heavy Investments: Fully redundant systems.

3. [REPORT] https://www.fcc.gov/reports-research/reports/measuring-broadband-america/measuring-fixed-broadband-report-2016

Goal of IP2020

Our networks have to be **Open, Smart and Deterministic**

Mobility
- Across different accesses with continuity

Multihoming, Multipath
- Always reachable and discoverable
- Same device different paths per flow basis or load balanced

Addressing
- Favorable to diverse category of end points

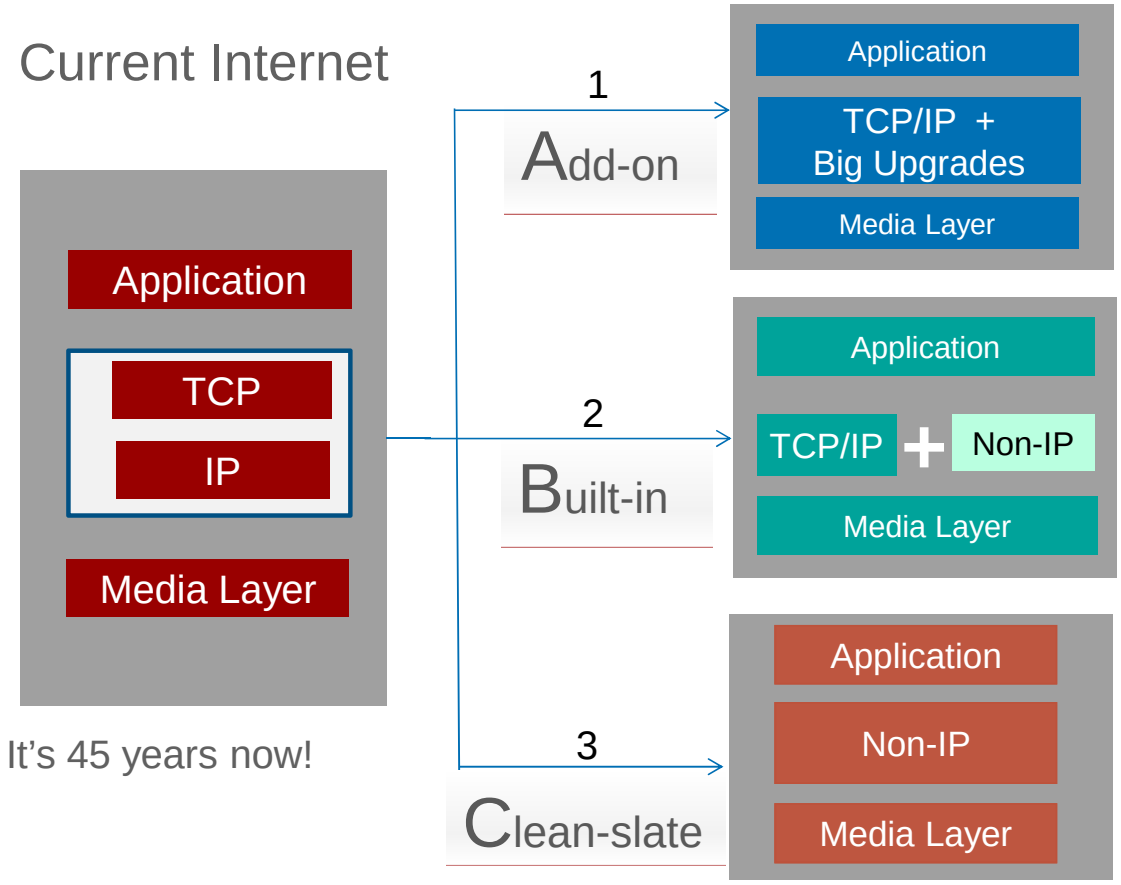
Scale & Security

Path Consistency
Economy of path taken
Eg. BGP path distribution

Reliability

Latency
Predictable & Measurable

How to make it open, smart and deterministic in Data Plane



Example: IAB IP Stack Evolution

- QUIC
- MPTCP
- L4S, PLUS

Example: Vendor/US NSF

- IP 2020
- MobilityFirst *
- NDN *, CCN, ICN,
- XIA *

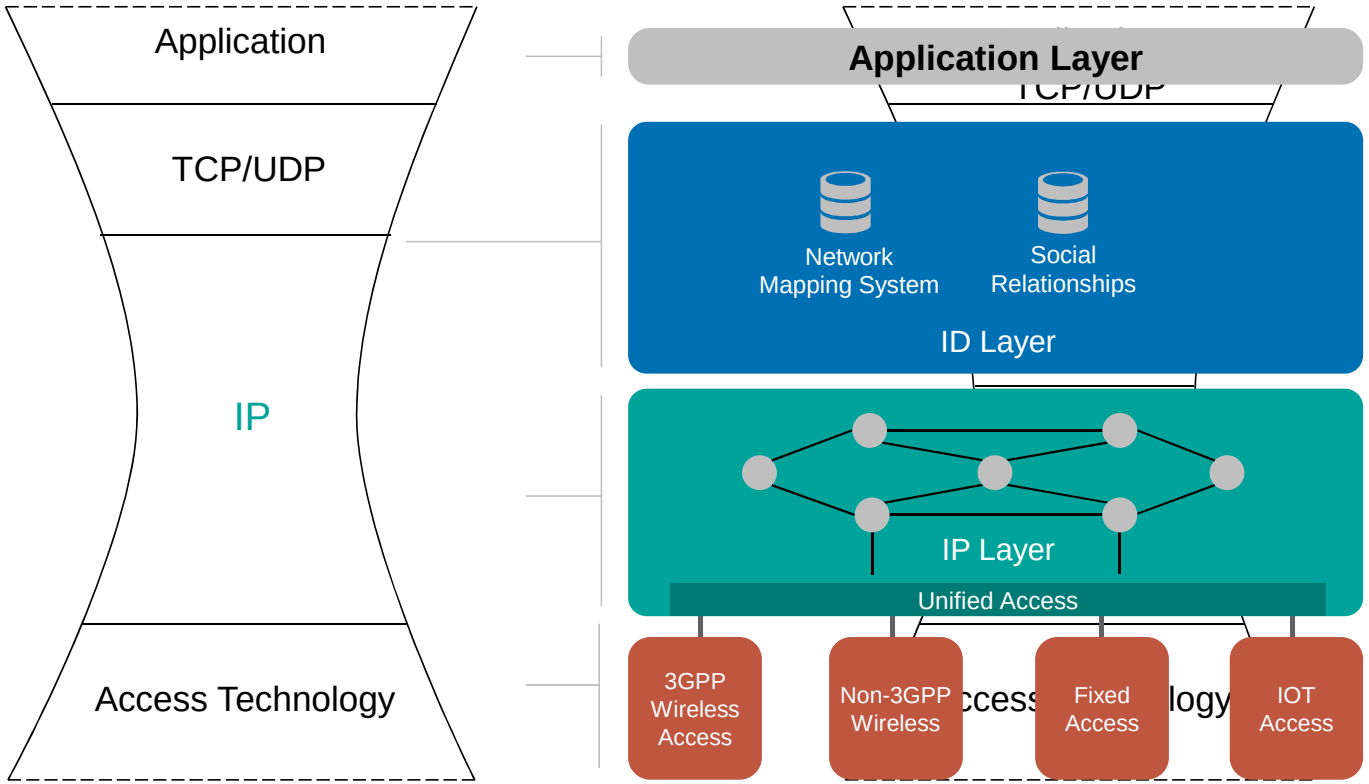
* FIA: Future Internet Architecture

Example: Mostly Academia or history

- ICN, Scion
- RINA
- ATM, Frame Relay

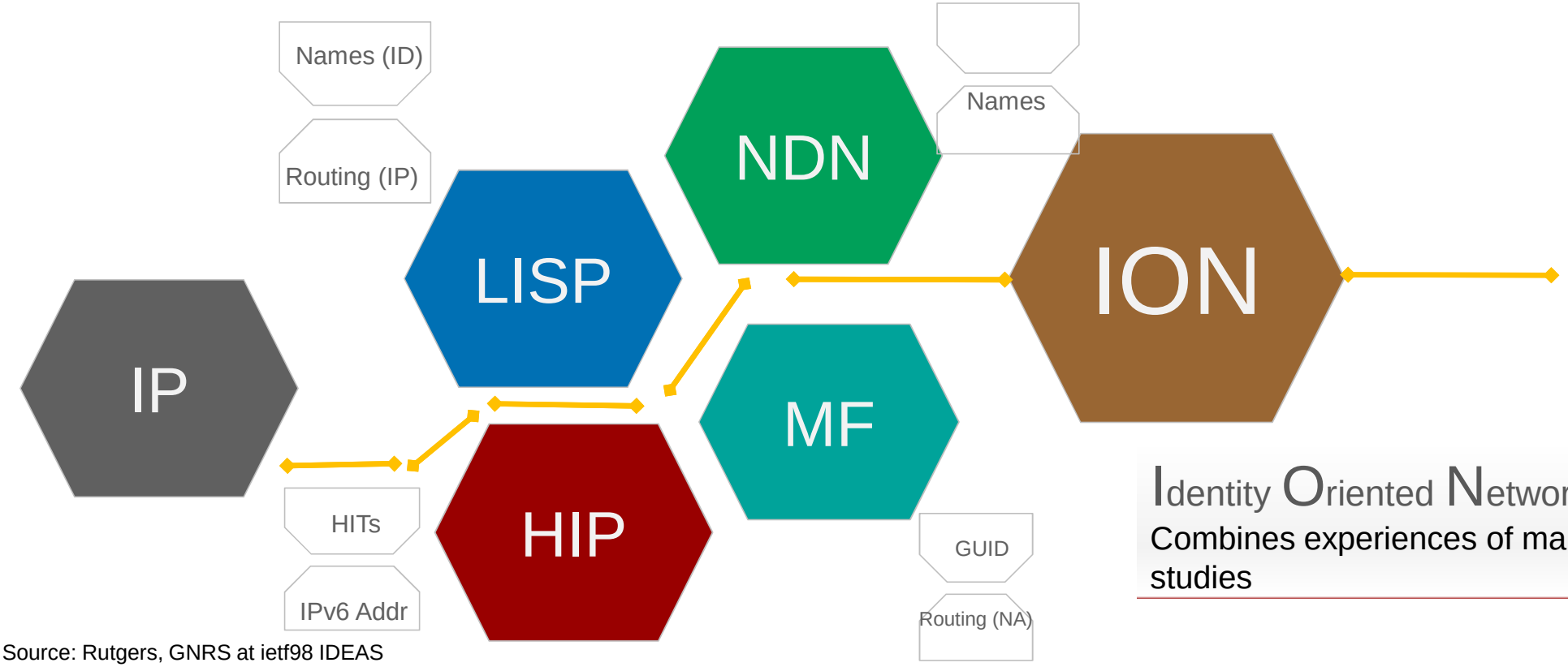
ID Enabled Networks

ID Oriented Networking (ION) Paradigm



ID Oriented Networks (ION) And Architectures

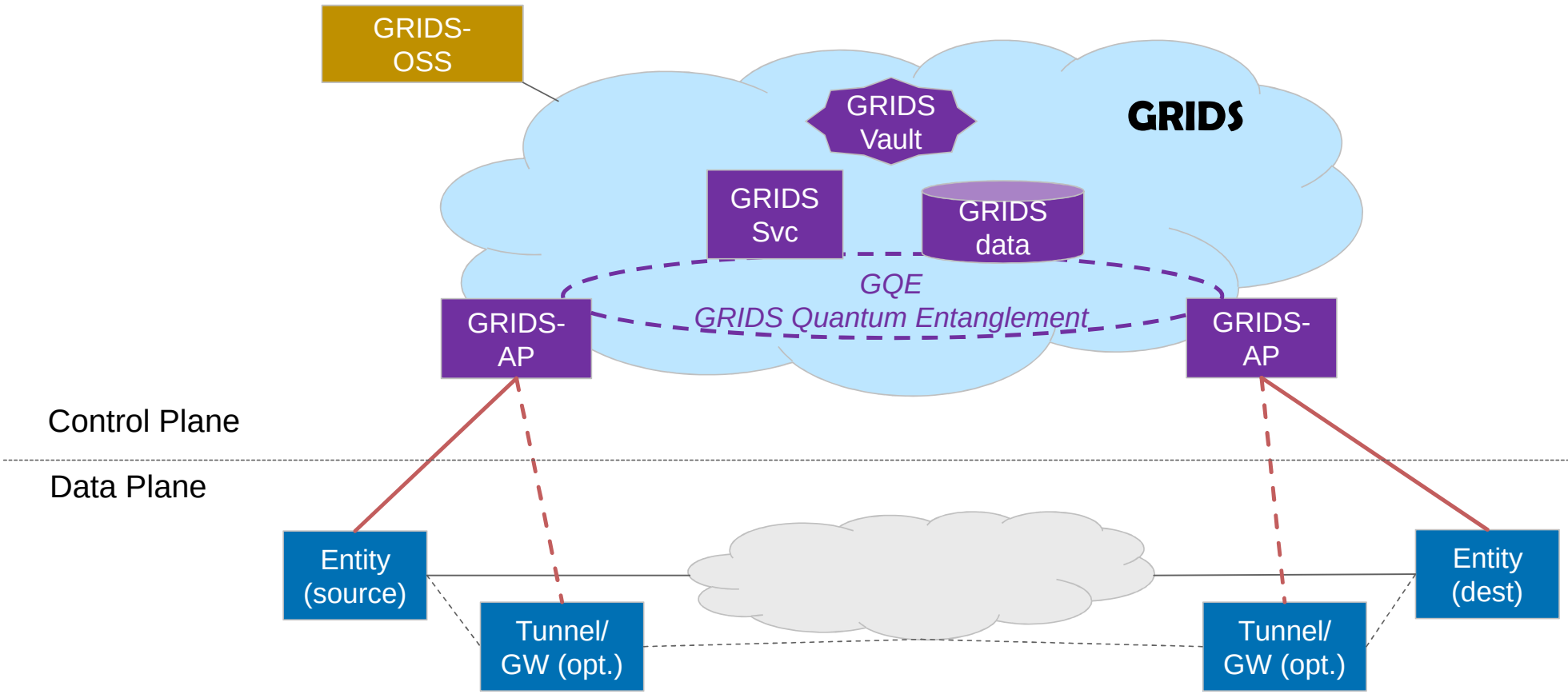
Current Internet | Research and Experiments | Industrial Awareness and Adopts



Identity Oriented Networks
Combines experiences of many studies

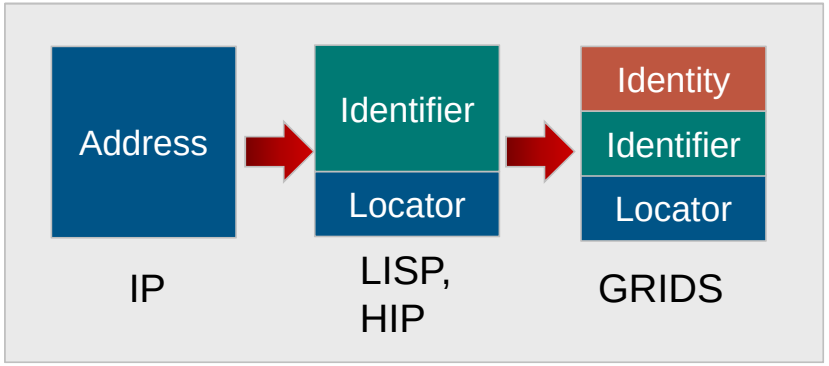
Source: Rutgers, GNRS at ietf98 IDEAS
https://drive.google.com/drive/folders/0BwYx7u1T_20RODdLaWpIdk9feHc?usp=sharing

High-Level GRIDS Architecture



What does GRIDS provide

- **Common Control Infrastructure and Services to enable separation of Identity (Idy) and Identifier (Idf) in addition to separation of Identifier and Locator**
- **Separate <u>how an endpoint is referenced</u> from <u>who the endpoint actually is</u>**

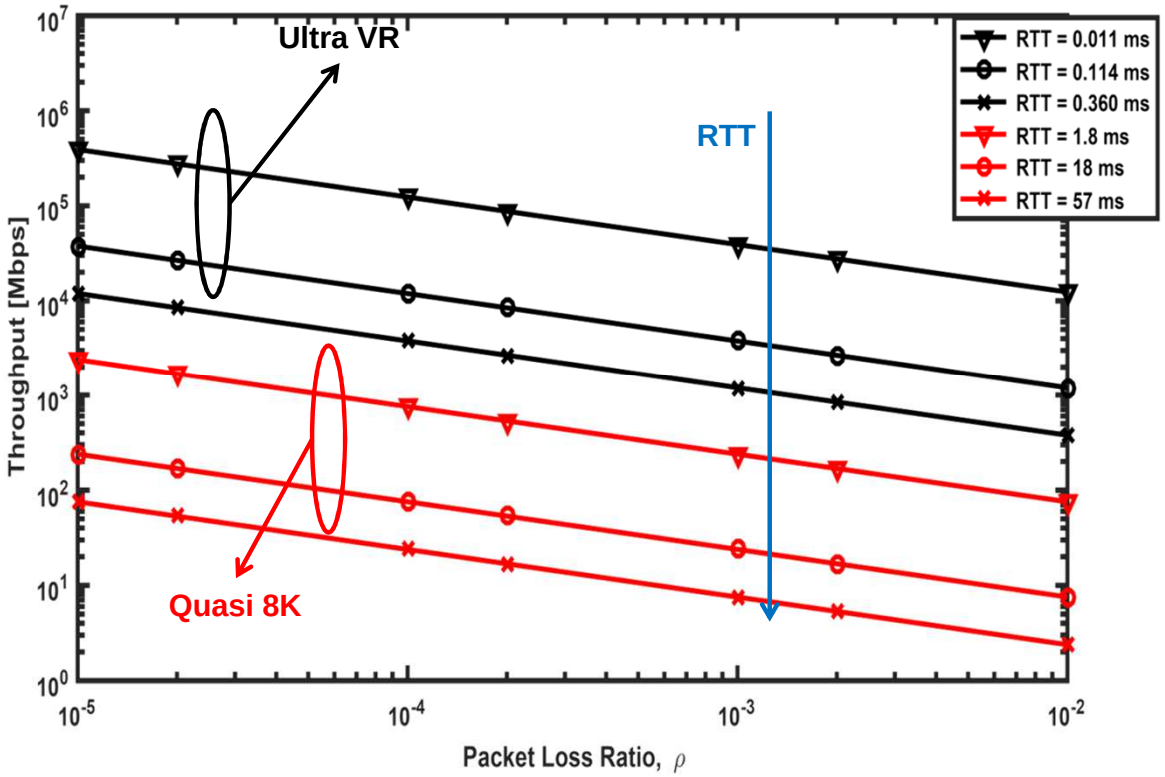


- **Control access to personal communications data on need-to-know basis (vs. all-or-none today)**
  - Balance concerns for privacy and desire to retain control of personal data with need to keep communications secure
  - Facilitate management + enforcement of communication policies while respecting privacy and ensuring security of communication peers
  - Leverage metadata about communicating endpoints for additional degrees of security and control

Deterministic Transport

TCP Throughput Law

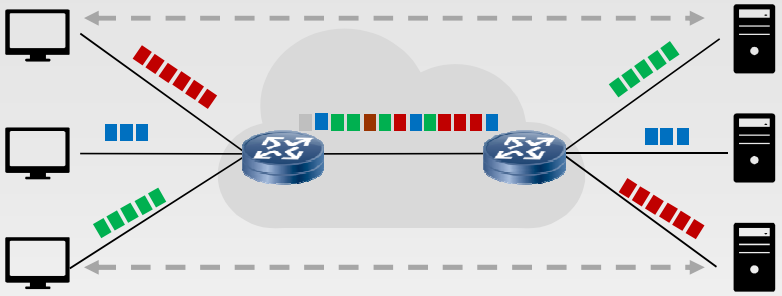
Relationship between Throughput, Packet Loss and Delay



$$\text{TCP Throughput} \le \min\left(\text{BW}, \frac{\text{WindowSize}}{\text{RTT}}, \frac{\text{MSS}}{\text{RTT}} \times \frac{\text{C}}{\sqrt{\rho}}\right)$$

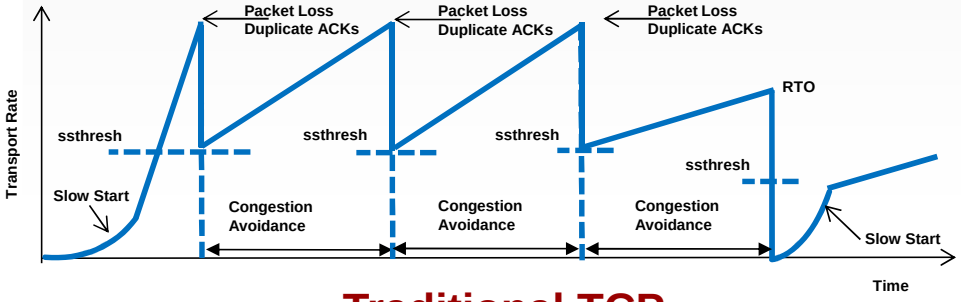
- The TCP throughput is inversely proportional to its packet loss ratio and round-trip time delay.
- Example: For throughput 12 gbps and packet loss ratio 1/10,000, the end-to-end delay is 0.114 ms.

Deterministic TCP (DTCP)

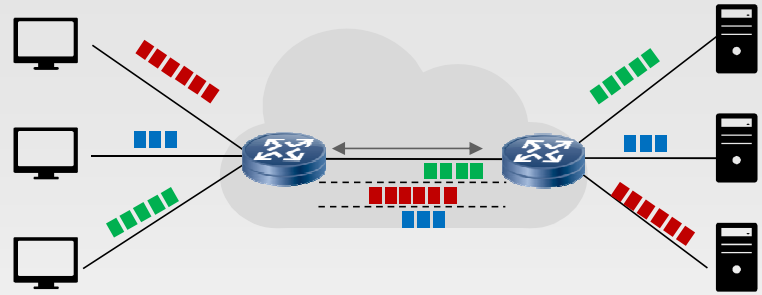


TCP Requirement:
- Avoid congestion
- Higher bandwidth utilization

Method:
- Sliding window mechanism



Traditional TCP

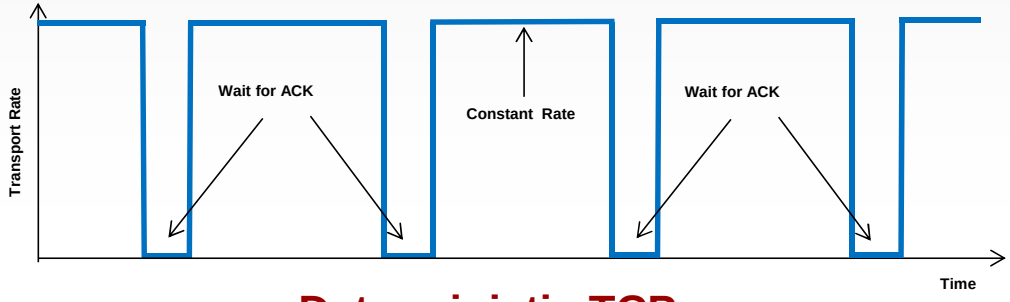


Flow control for DTCP：
- Leaky bucket：rate limit
- Token bucket：allow burs

Retransmission mechanism:
- FIFO

DTCP：
- Signaling by TCP
- Provide guaranteed network resource

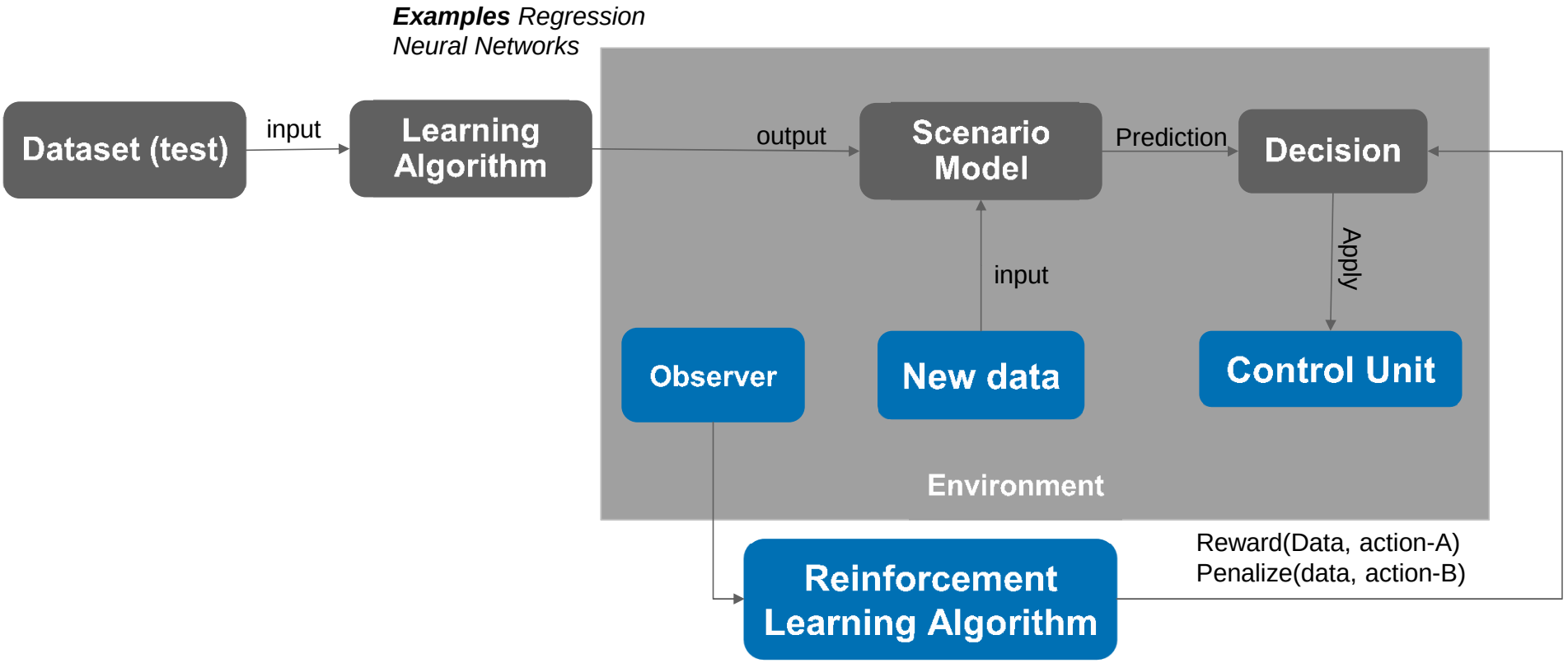
Transport benefits：
- TCP does not need to use different congestion avoid mechanism to reach target rate
- Higher bandwidth utilization due to no packet loss signaling



Deterministic TCP

Intelligent Data Center Networking

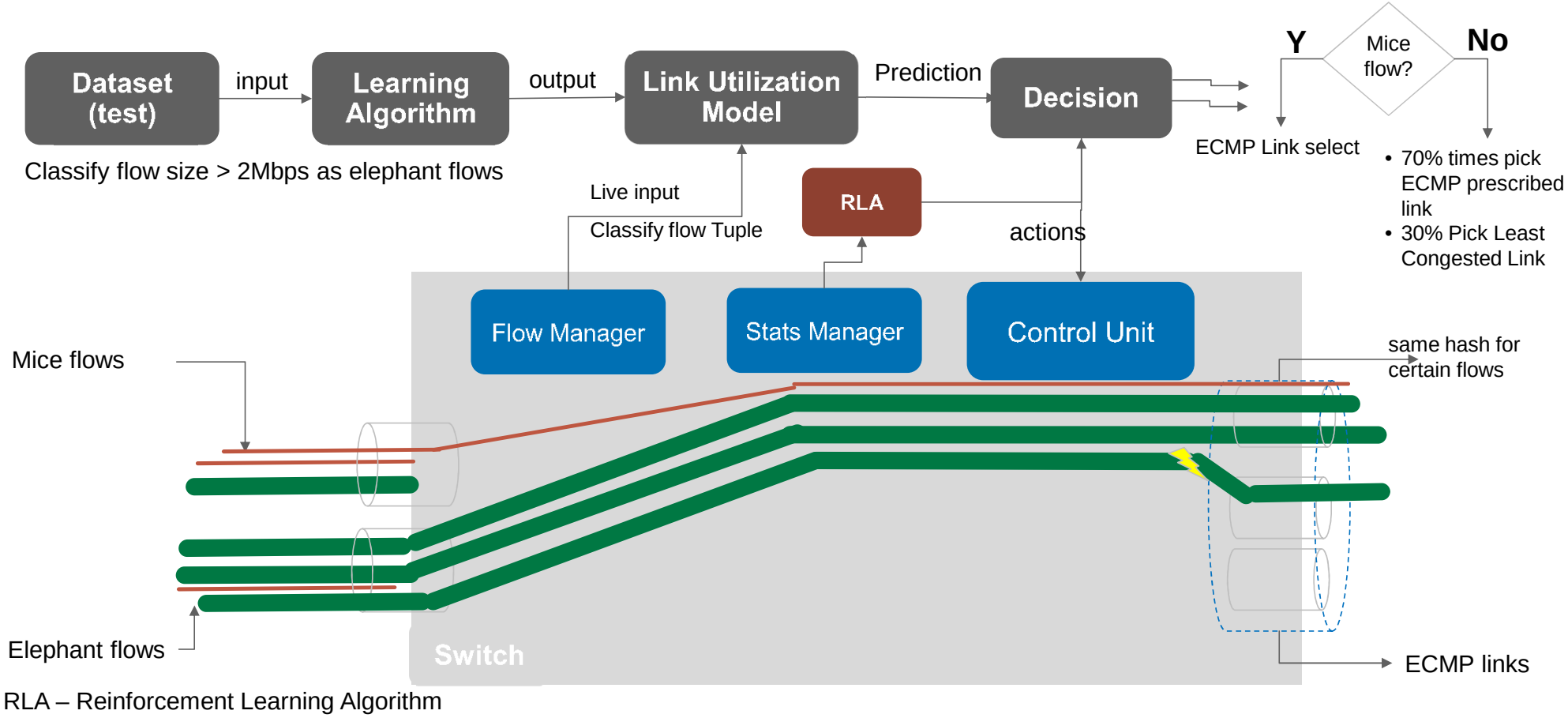
A Generalized Machine Learning Loop



An Example: ECMP Based Link Utilization Problem in a Switch

- **Massive Scale DCs use fixed spine-leaf topology**
- **ECMP distributes traffic across multiple paths**
- **ECMP uses Hash computation to balance similar flows over multiple links**
- **However, the flows are not evenly balanced**
  - Low-bandwidth (Mice) flows: Majority of flows are short-lived and latency sensitive.
    - Example: Web, chat applications
  - High-bandwidth (Elephant) flows consume majority bandwidth and are long-lived.
    - Example Storage-intensive big-data, data-replication and backup applications
- **Problem**
  - Variance in the amount of bandwidth used between long-lived vs short-lived flows does not ensure that traffic is balanced across all the links.
  - Increase in Mean-time-to completion for mice flows
  - Reduced data-rate for elephant flows due to congestion control

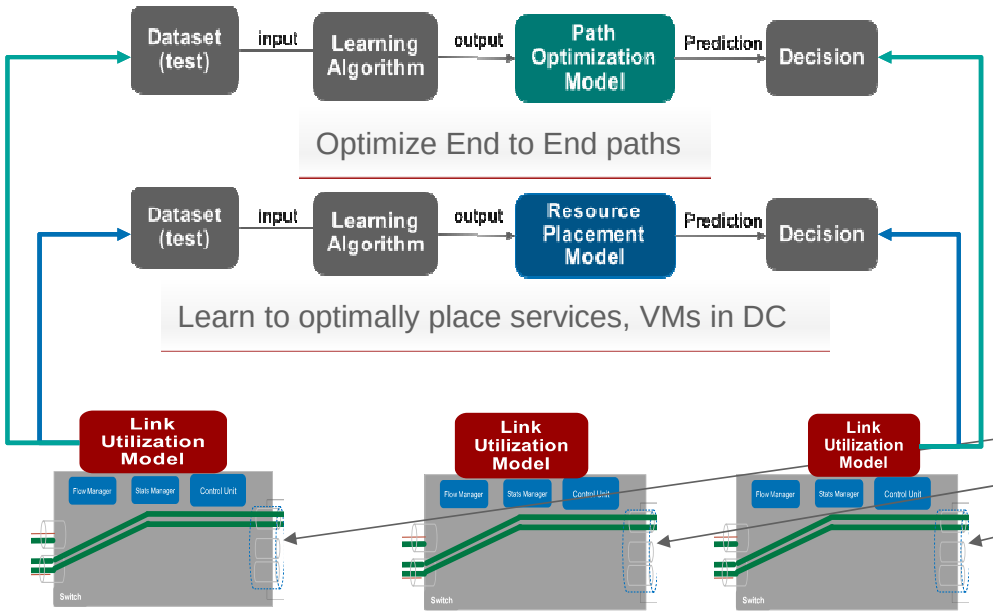
Machine Learning for ECMP Link Utilization in a Switch

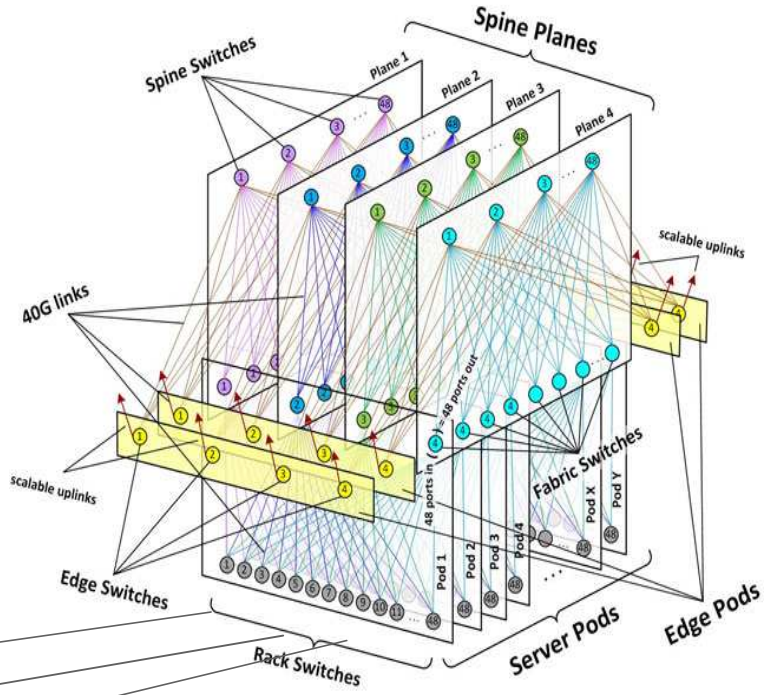


RLA – Reinforcement Learning Algorithm

Intelligence Driven Networking – DC Scenarios with Global Scope

- Extend to wider scoped learning - Global models across multiple switches
- Different Learning models for different scenarios together





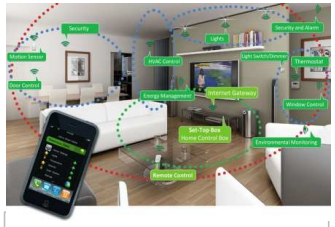
Src: Facebook data center

Summing Up - IP2020 Delivers Next Generation Networks

Identity **O**riented **N**etworking
(ION)

Intelligence **D**riven **N**etworks
(IDN)

Wearable WSN



Smart Home

Next-**G**en **T**ransport
(NG-T)

Autonomous Driving

Smart Cities

Thank you

www.huawei.com