Observation of Current IP Network Issues and Innovation Requirements

Dr. Sheng Jiang

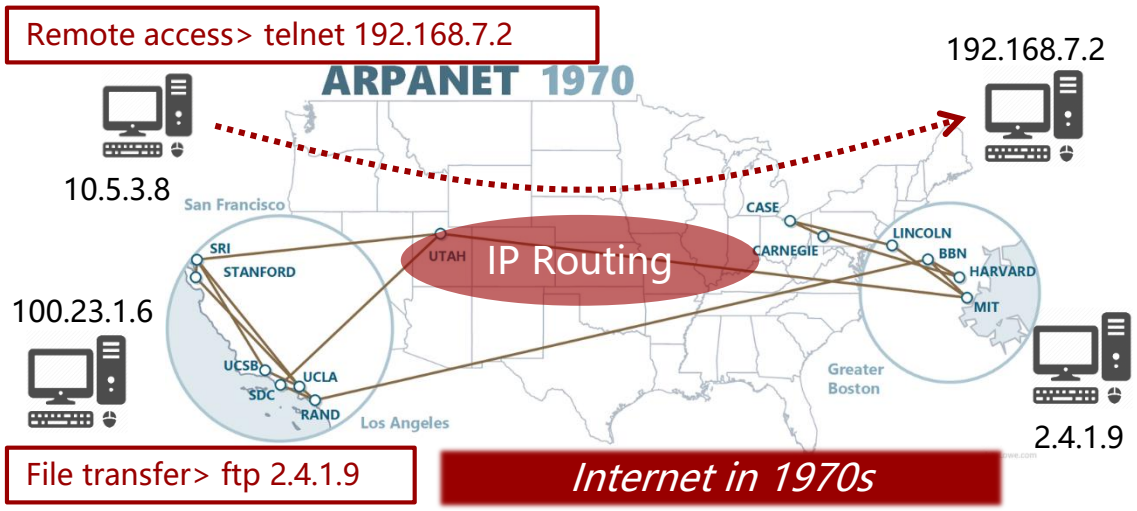
Network Technology Lab

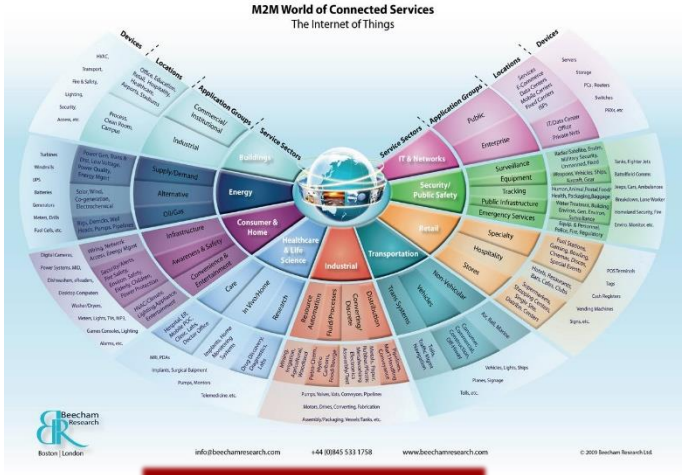
www.huawei.com

HUAWEI TECHNOLOGIES CO., LTD.

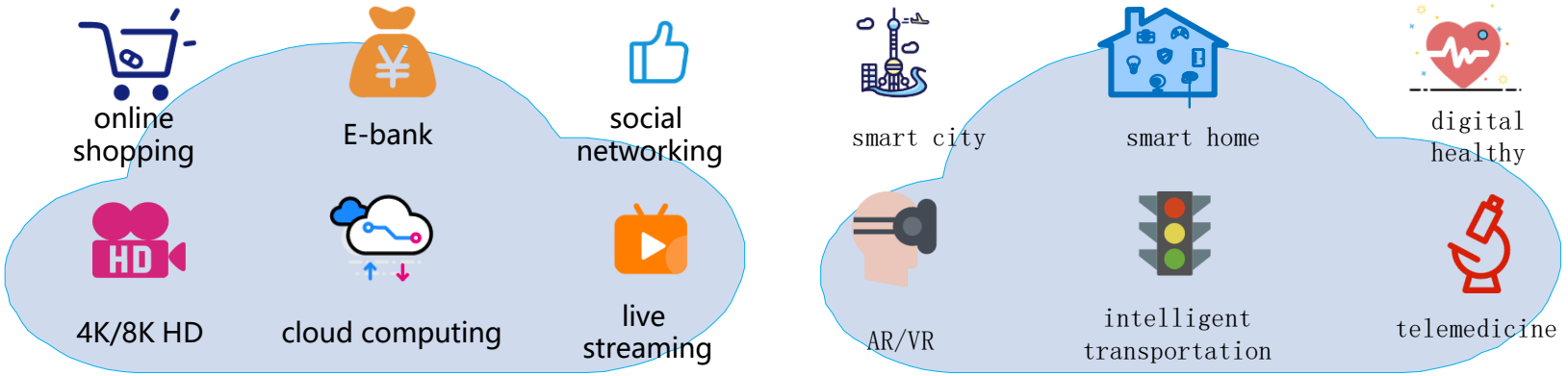
Contents

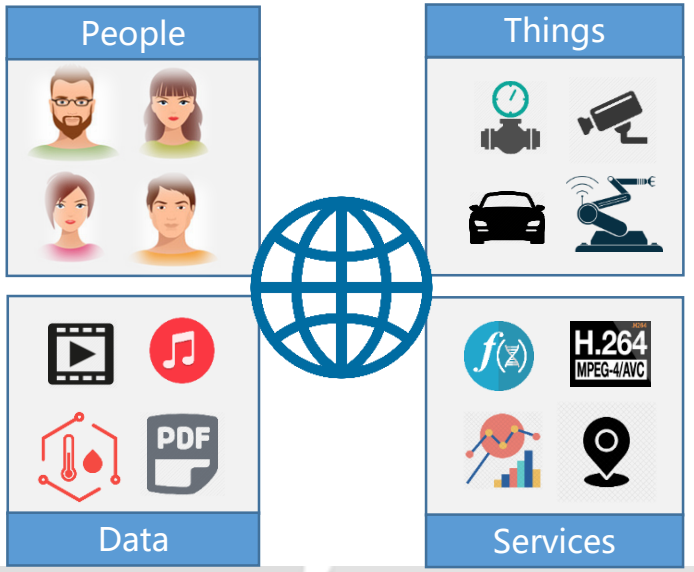
Internet Rapid Development





Source: https://blogs.saphana.com/2015/03/18/part-4-iot-industry-connectivity-standards/





More and more emerging services → Communication entities have been changed

Imagine the Future

Everything
Sensing

DEVICES

Everything
Connected

NETWORKS

Everything
Intelligent

CLOUDS

Larger Scale through Supporting Massive Connections

Future Network

T-bits Applications

Inherent Security

Real-time

Higher Value through Providing Extreme User Experience

Internet Protocol Family has been Continuous Evolving

Provide more **network functions** in the domain, such as mobility, VPN, SFC, etc. besides traffic engineering.

NVGRE
GTP
LISP
PMIP

SR/NSH

Only provide **E2E reachability**; Shortest path forwarding within a domain.

IP

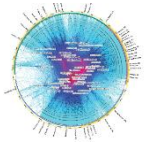
MPLS

MPLS/IP VPN
EVPN
VxLAN
BIER

Provide intra-domain **traffic engineering** capabilities besides IP E2E reachability.

Future Internet Protocol

Security and privacy



Massive connection (>100billion)

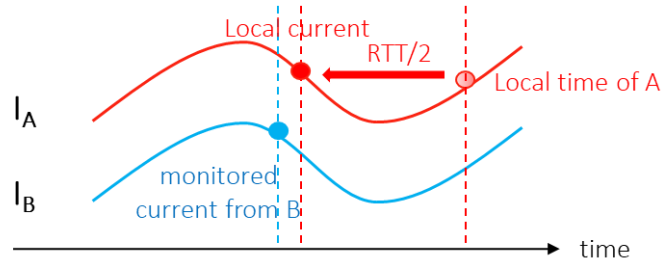
Deterministic ultra low latency (<1us)

Ultra high throughput bandwidth (>1Tbps/person)

Contents

Current IP Network Issues and Innovation Requirements

Numerous Application Scenarios Request Differentiated Services

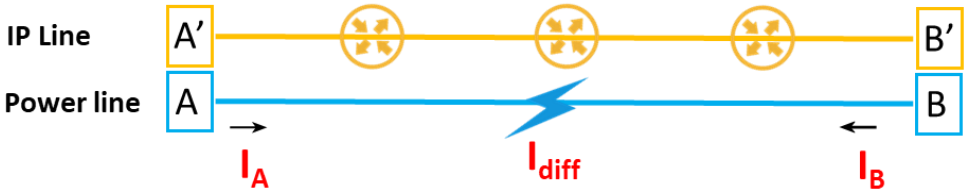


Out-sync sampling leads to false judgement

buy 1 share **every 5 ms** at a price of $1/share (total 5000 shares)

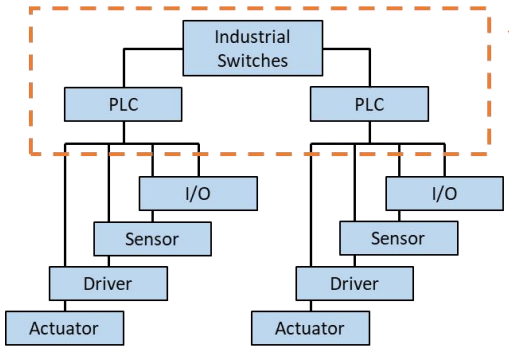
buy 1 share **every 10 ms** at a price of $5/share (total 1000 shares)

Disordered transactions due to non-deterministic forwarding



Differential Protection

Cloud PLC (Programmable Logical Controller)



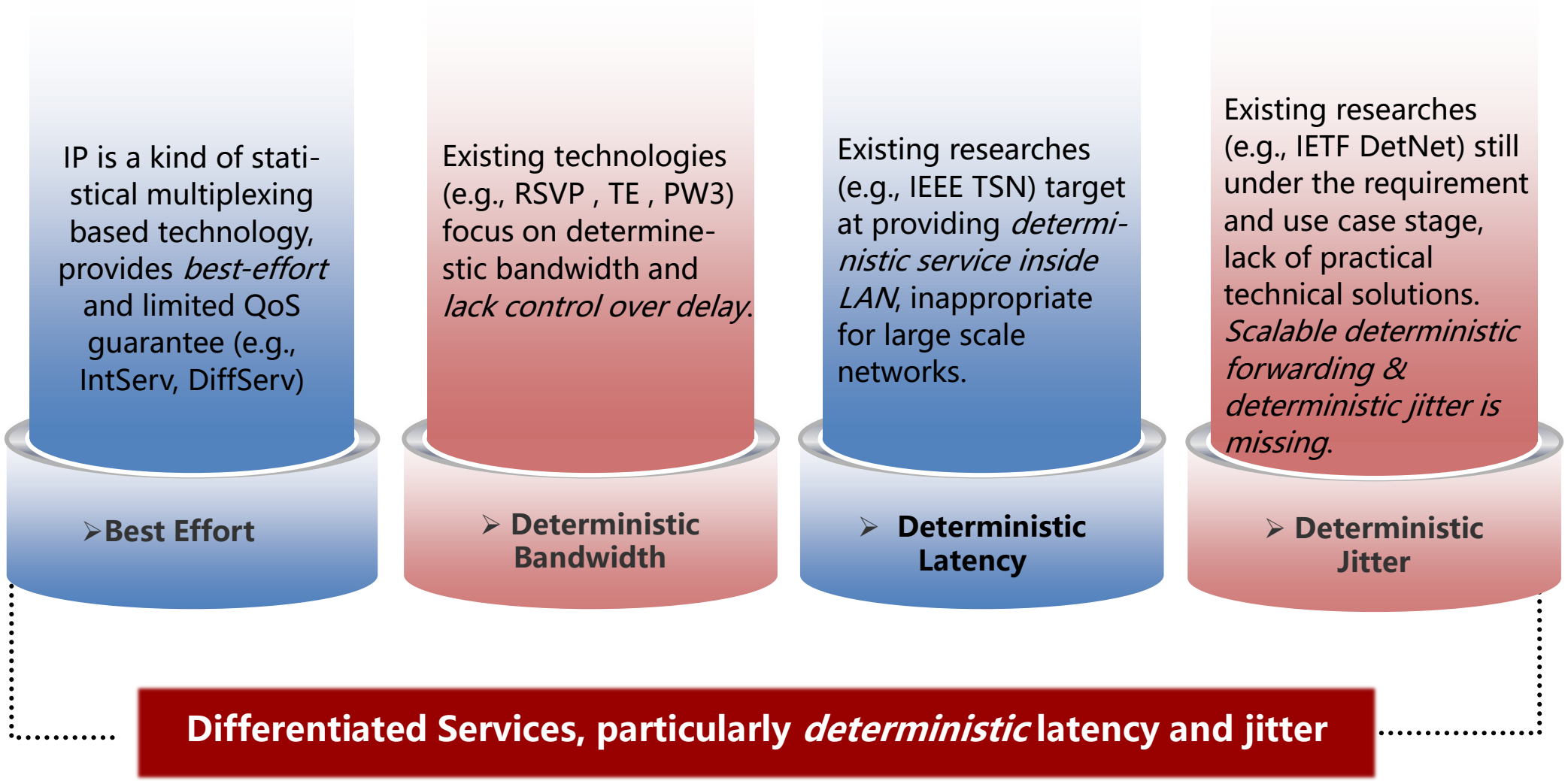
Current factory network

Cloud PLC

Imaginary (one doctor)

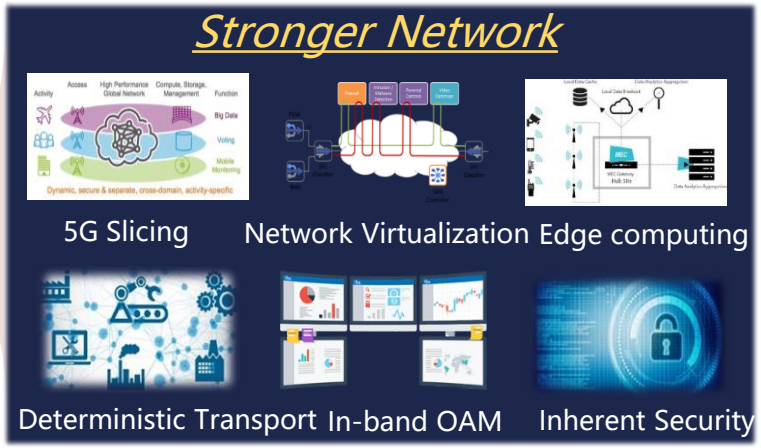
In Fact (multiple doctors and nurses cooperate in a surgery)

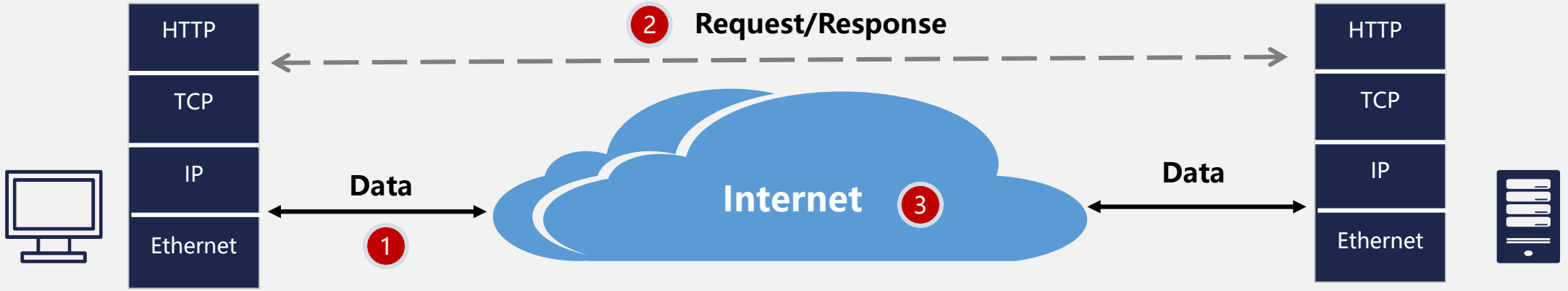
Deterministic Services Beyond the Best Effort



Weak UNI for Network Awareness And Customization

Lack of information on both user → network and network→ user directions

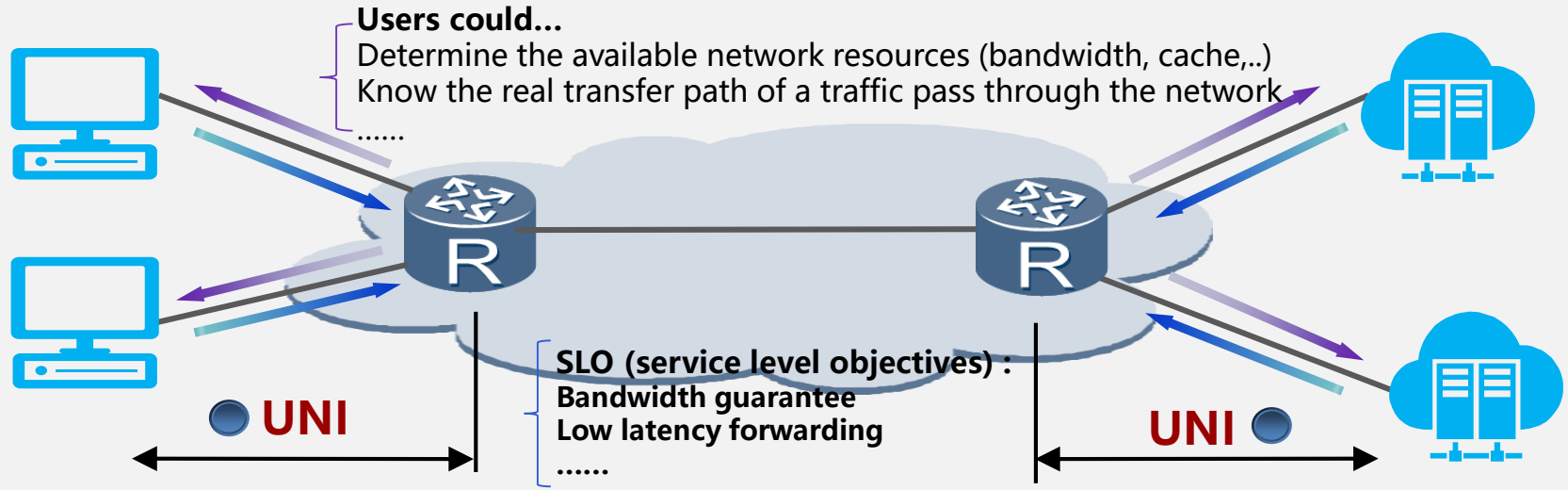




① IP is the only interface between user and network → users cannot express the experience requirements for the network.
② Based on the layer separation principle, TCP/HTTP Request/Response only occurs between terminals.
③ The network is not aware of the needs of the application layer → therefore unable to provide deterministic services.

More Information Exchanging between Users & Network

UNI : Network → User & User → Network



Network → User: Users could be aware of the network status:

- Transfer path of a traffic
- Middle-Box processing information
- Response time
- Network throughput
- Congestion states
- Availability
- Network utilization
- ....

User → Network: users could define the network:

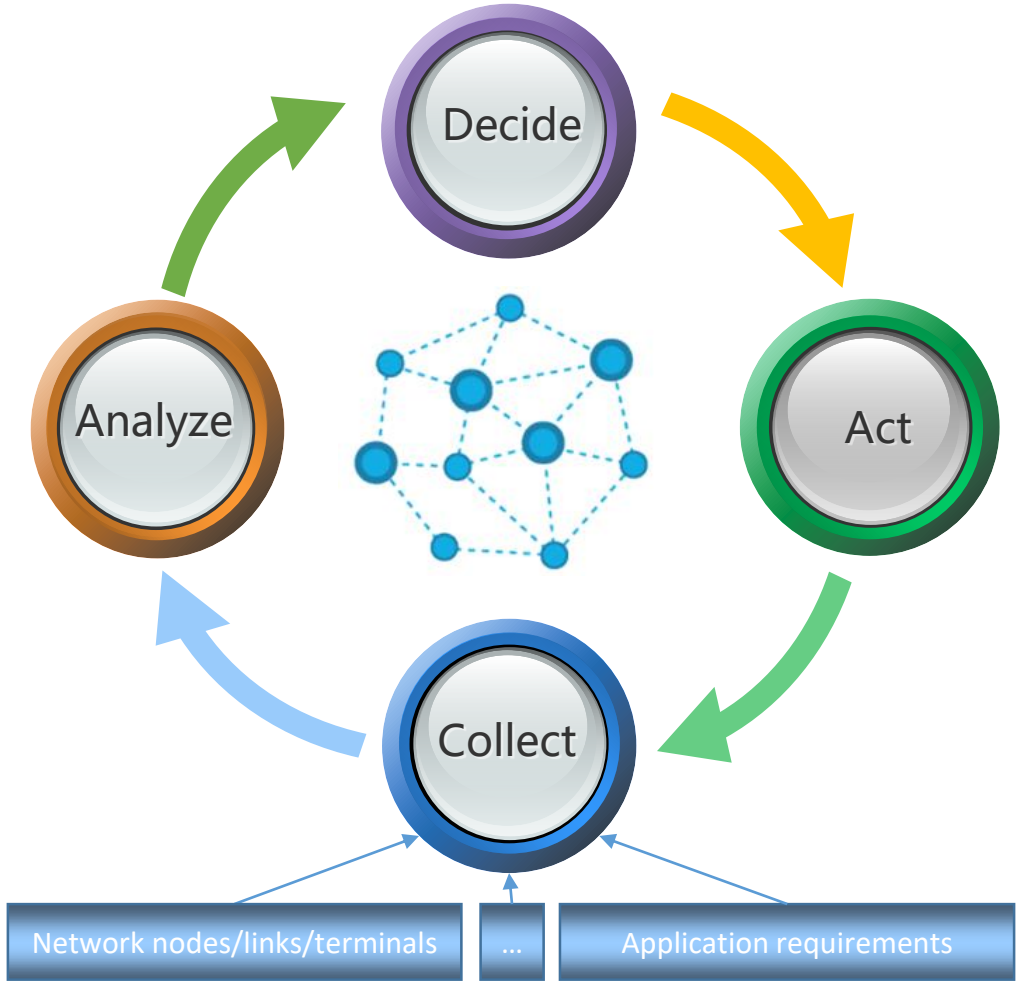
- Low-latency forwarding
- Forward before deadline
- Delay tolerance
- .....
- High-bandwidth forwarding
- Drop notification
- obtain in-band OAM info
- .....

Issues of Current Network Measurement

- **Inaccurate network status data**: Measurement results cannot represent the real path, forwarding performance, failure network elements/links of a service traffic, since paths of different packets going through the network may be different.
  - *Traceroute: Since paths of different packets going through the network may be different, the traceroute result including 1st hop, 2nd hop, …, Nth hop, cannot be regarded as the complete path taken by a service packet.*
- **Non-real-time network status data:** the frequency of collecting network status data is relatively low, most information is reported periodically by network devices.
  - The non-real time measurement data does not meet the current new production requirements, and cannot respond quickly to the network elements/links/terminals' fault, alarm, mis-operation, etc.
- **Incomplete (physical and virtual) network status data:** there are many network aspects that we did not measure. New measurement methods still needs to be newly developed to cover these uncovered, new aspects or new network elements.

Needs for Accurate, Real-time, Complete Network Status Data

- **Network Status data are needed to truly cognize the network**
  - Network availability, utilization, throughput
  - Congestion states (links, routers/switches…)
  - Transfer path of a traffic ……
- **Accurate, real-time, complete network status data could enable more autonomous analysis and best decisions**
  - Optimizing the network performance, user experience.
  - Simplifying the network configuration and management.
  - Quickly responding to the network fault, alarm, mis-operation, etc.
- **Sufficient data could train AI models for faster decision and dynamic adaption capability.**



Problems of Current Network Control and Management

Configuration

Heavy labor cost

- Initial configuration requires highly professional staff to do per-device manual work *(even with assistant of state-of-the-art NMS, manual work couldn't be totally avoided)*
- Run-time configuration adjustment is frequent, according to service requirements

Error prone

- 95% network failure is caused by mistaken manual configuration

Traffic Optimization

Slow response

- Real-time adjustment according to network service's dynamic change is very challenging

High risk

- Sometimes it's very hard to estimate the result of adjustment; failure might happen
- Physical switch might be needed to deal with some serious performance issue, due to the software adjustment limitation

Fault Diagnostic

Heavy time-consuming

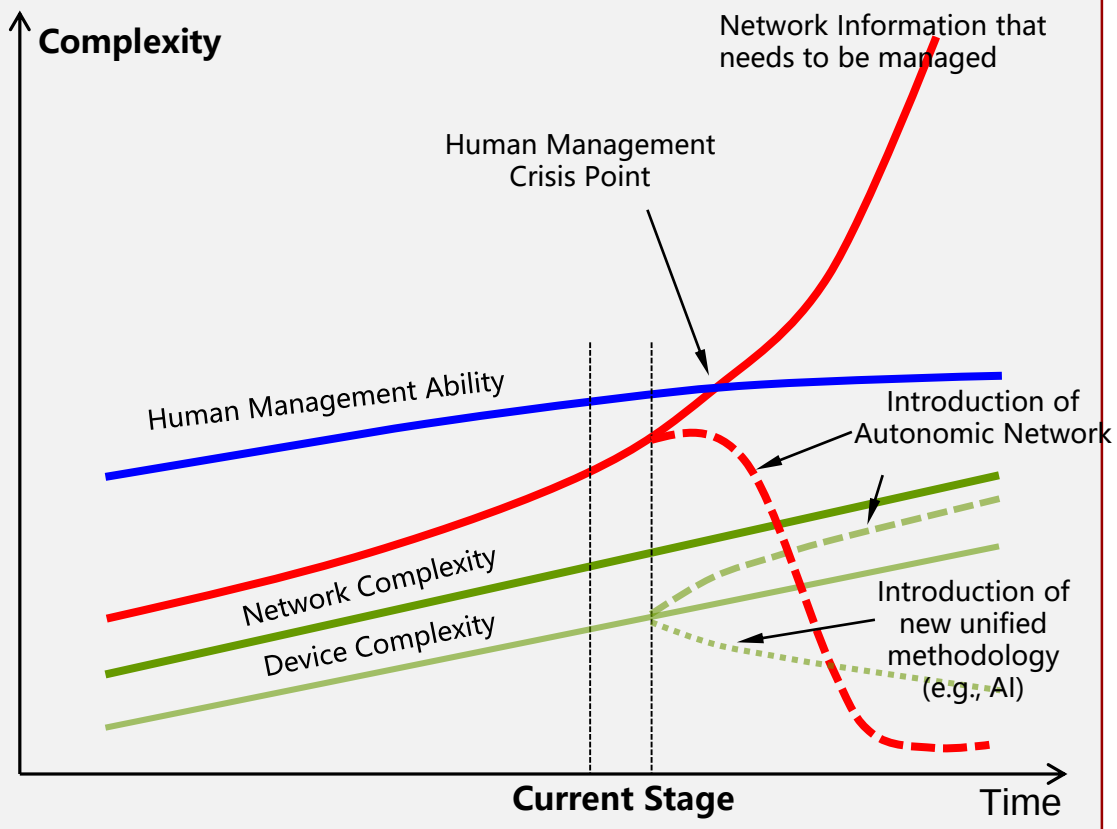
- Manually identifying the network fault tends to always exceed custom's patient

Struggle to deal with

- Sometimes it is very struggle for human to identify the real problems

Network Automation and Intelligence Requirement

Human Cost Issue is Becoming More and More Serious



Network Automation and Intelligence Requirement

- Human-based management cannot handle the more and more complex network.
- Introducing autonomous mechanisms into network could simplify the human management, reduce the human error and the cost of network maintenance, and improve the management efficiency.
- Introducing AI algorithms may have the chance to unify the solutions for various scenarios.

Fixed Address Length makes Extension Hard

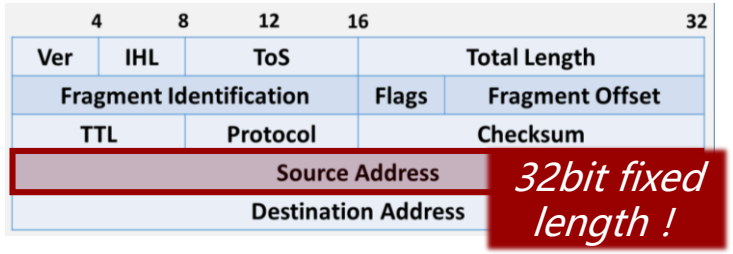
Physical Communication Entities

Smart Terminals	Intelligent Wearing	Smart Home	Smart City	.....
				.....
Global Address	Global Address	Global Address	Global Address	

Virtual Communication Entities

Virtual Machine	Container	Process	Function	.....
				.....
Global Address	Global Address	Global Address	Global Address	

However, the length of network addresses could not be extended as required...



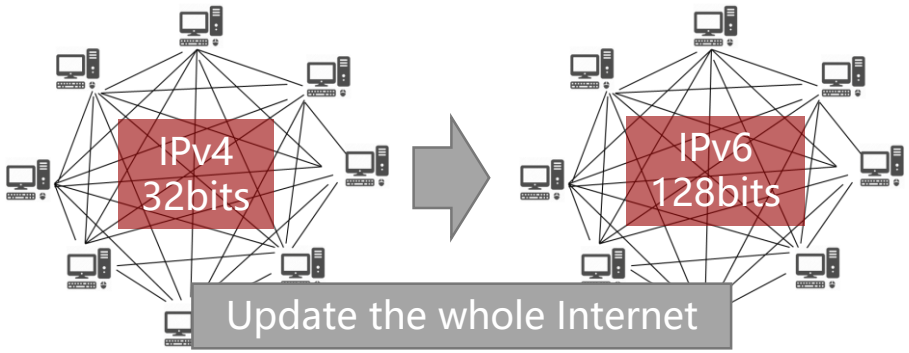
32bit fixed length !

Address Extension = Protocol Update

"If A updates , B must update"

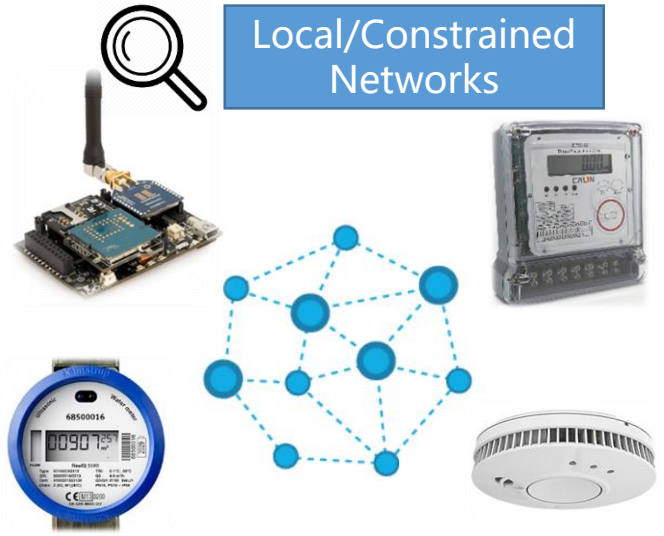
"Updating means nothing for D if C does not update"

"If B updates , C must update"



Fixed length address design makes it extremely hard for address extension!

Short Address are Requested for IoT Scenarios



Local/Constrained Networks

Feature 1

- Local communications are dominant
- Small number of nodes (up to thousands)

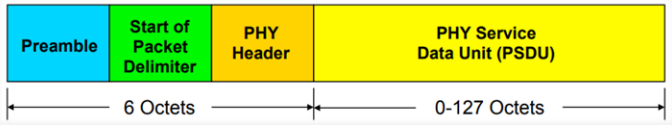
Feature 2

Resource Constrained

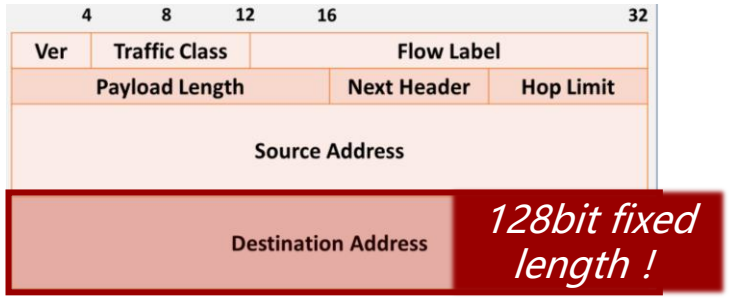
Weak CPU | *Small Memory* | *Lack of Power*

Feature 3

Variety of layer-2 technologies, small MTU : 127B (802.15.4)



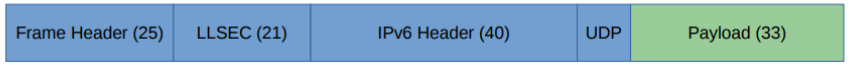
However, the length of network addresses could not be compressed as required...



IPv6 address is too long to fit those scenarios

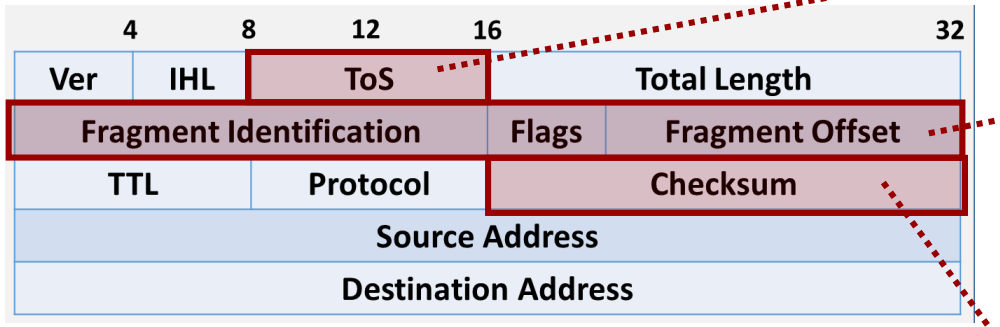
- Maximum frame size in IEEE 802.15.4: 127 byte
- Reduced by the max. frame header (25 byte): 102 byte
- Reduced by highest link-layer security (21 byte): 81 byte
- Reduced by standard IPv6 header (40 byte): 41 byte
- Reduced by standard UDP header (8 byte): 33 byte
- This leaves only **33 byte** for actual payload
- The rest of the space is used by headers

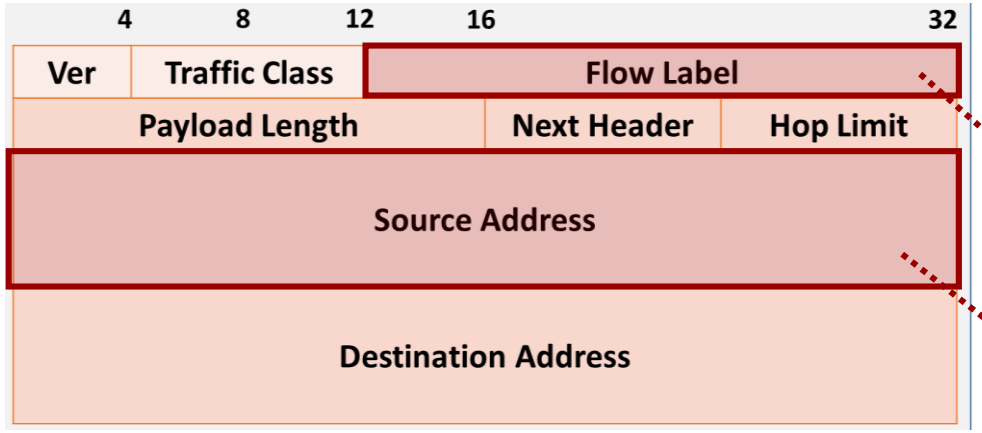
Where 32B (80% space) is used by network addresses !



IEEE 802.15.4 frame format

Fixed Protocol Fields Cause Encapsulation Redundancy





ToS:

Most traffic on the Internet is Best Effort, ToS is usually useless

Fragment ID/Fragment Offset:

These fields are designed for layer-3 fragmentation to adapt different layer-2 technologies (MTUs). However, as Ethernet being prevalent, these fields are no long used because of consistent layer-2 MTUs.

Checksum:

This field becomes useless because layer-1 and layer-2 have their own check method, and most scenarios for IP transfer are fixed network.

Flow Label:

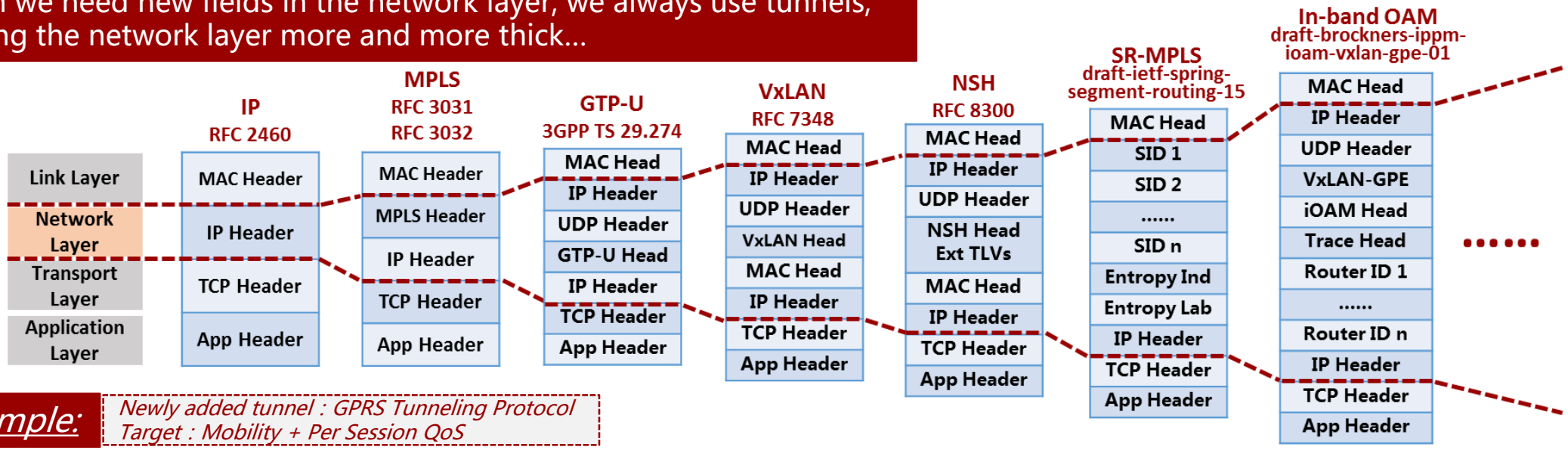
This field is not used in most scenarios

Source Address:

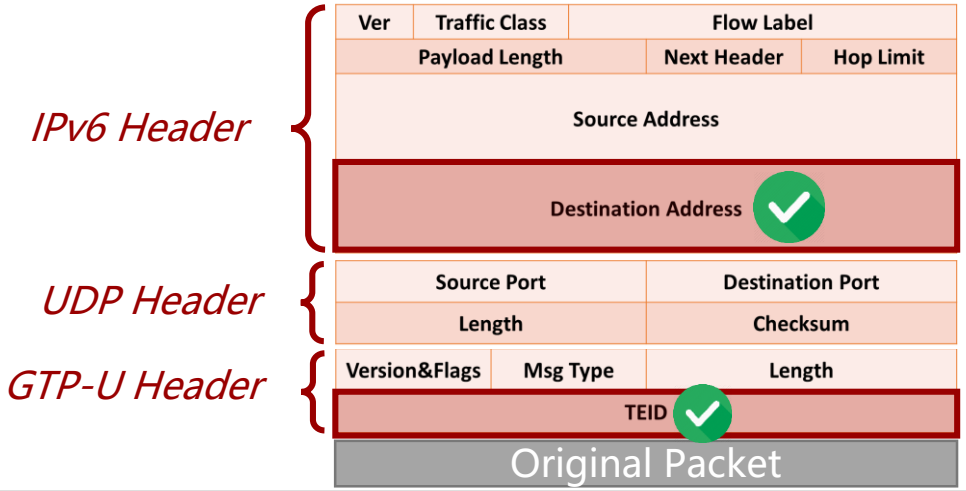
In IoT scenarios, some nodes are in dormancy model until it has some data to report to the gateway. After the reporting, it entries the dormancy model immediately. Therefore, the source address is useless in such scenarios.

Fixed Protocol Fields Lead to Massive Overhead

When we need new fields in the network layer, we always use tunnels, making the network layer more and more thick…



Example: *Newly added tunnel : GPRS Tunneling Protocol*
Target : Mobility + Per Session QoS

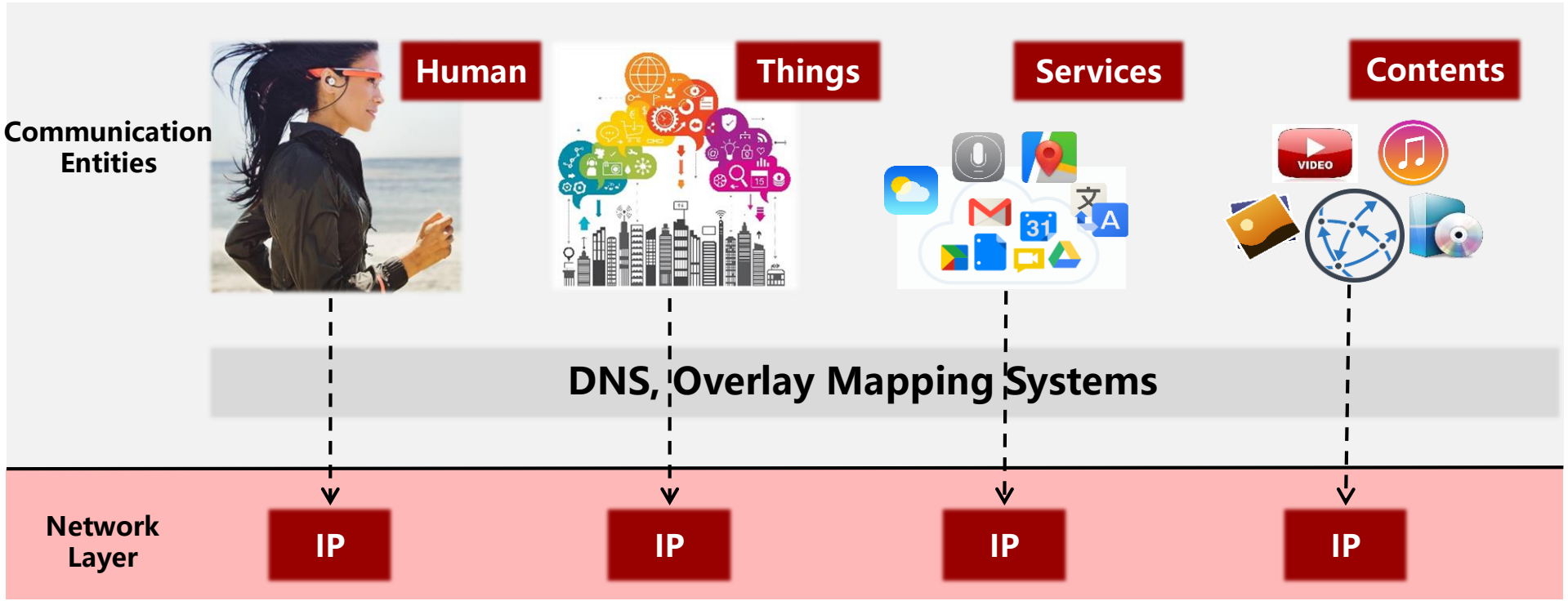


Main issues raised by fixed protocol fields :

- Meaningless data overheads
- Hard for multi-layer cooperation
  e.g., transit nodes do not know which virtual network, session, or service chain the packet belongs to; hard to achieve fine-grain QoS, security policy.

IP Address does not Represent the Real Communication Entitles

A communication entity is mapped to an IP address and not yet identified in the network.

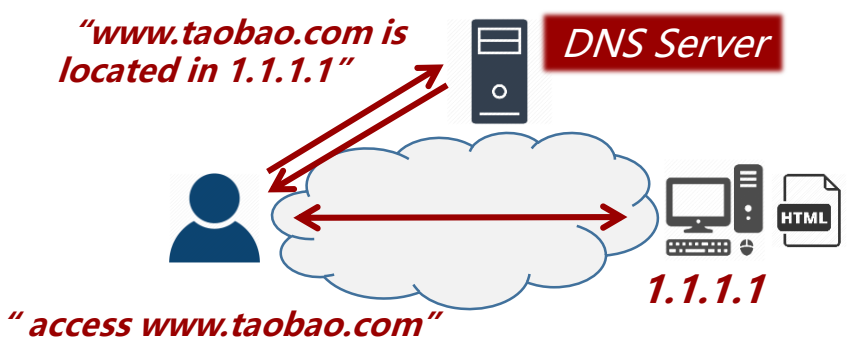


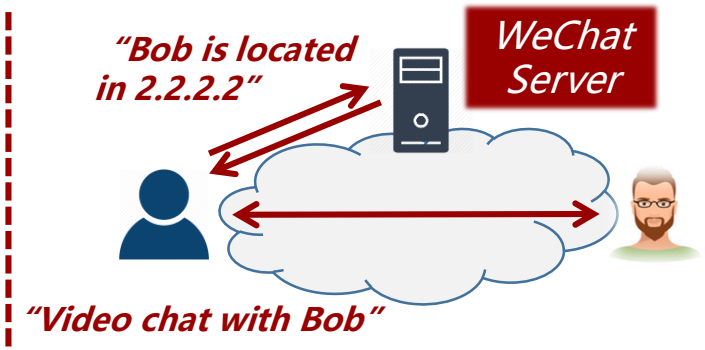
Invisible in the network layer

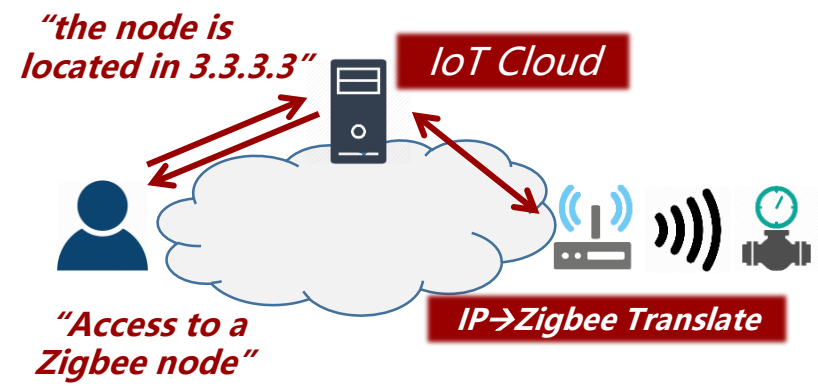
- Mapping between hostname (or other entitle) into IP is time-cost and lost the original meaning
- IP addresses can be forged and mapped wrongly
- Mappings are frequently changed and difficult to trace
  - *The topology semantic of IP address also causes mobility problems*

Locator Semantic of IP Address is hard to Satisfy Different Requests

To satisfy different services, network has to rely on infrastructures (DNS), OTT servers, or gateways...

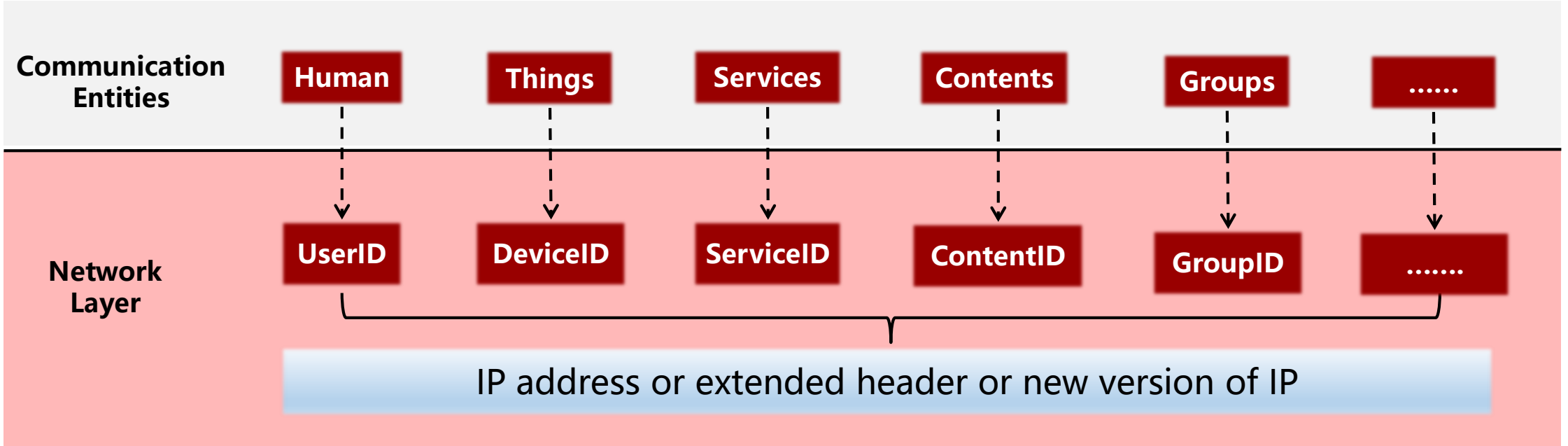






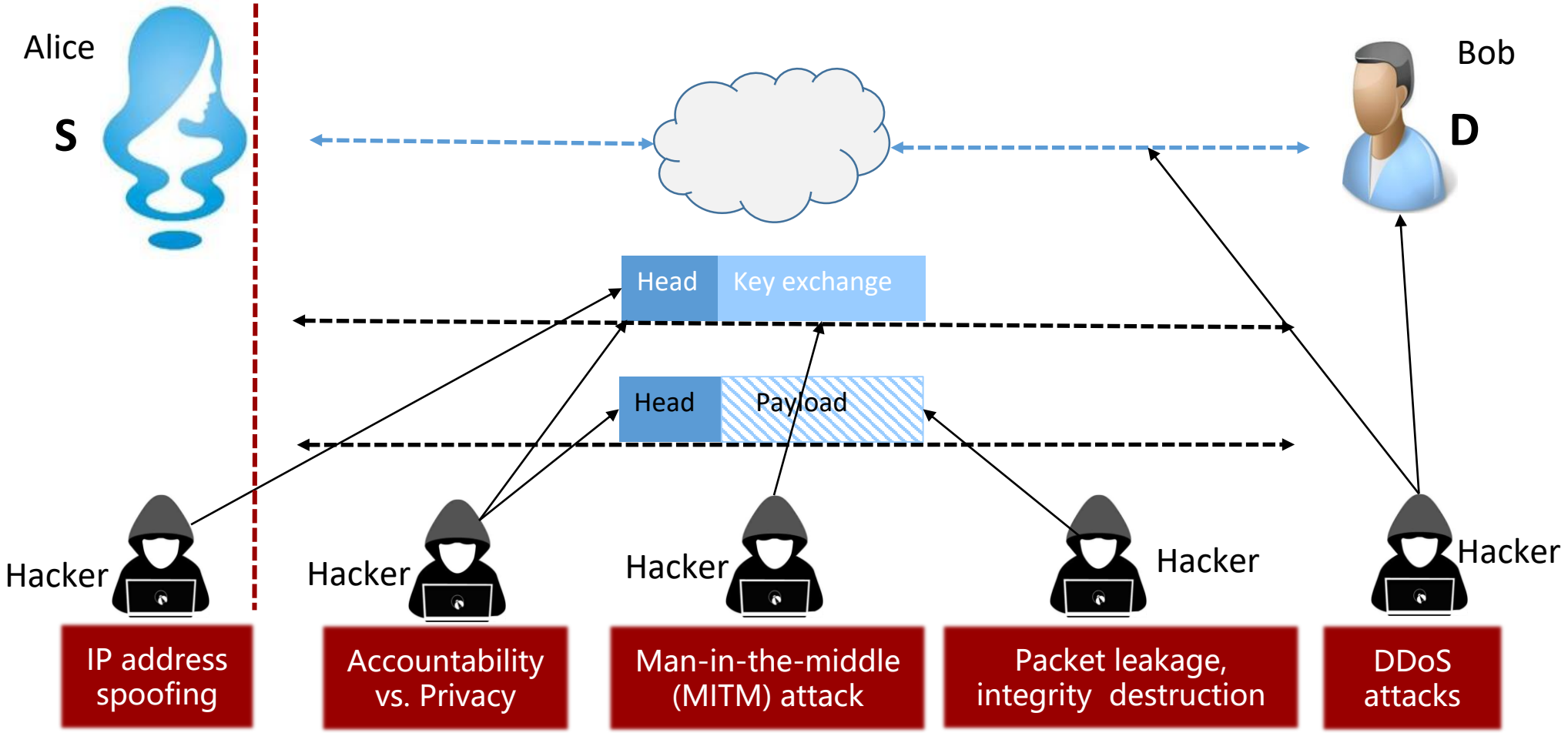
Issue 1 : Path detour causes delay	Issue 2 : Local data being transmitted on Internet causes privacy problems	Issue 3 : Poor reliability	Issue 4 : Dedicated gateway for translation	Issue 5 : Only IP-based QoS, security policy

Using Identifiers in Network Layer to Represent Entities Directly



The network layer could implement routing/forwarding and policies based on real communicating entities. It would be direct and efficient.

Network Vulnerability



E2E Communication Vulnerability Analysis

The trustworthy E2E communication is crucial for building a more secure Internet architecture, which provides a reliable secrete key exchange and DDoS defense capabilities while still balancing the tradeoff between accountability and privacy.

Attacks on End to End communication	*Issues of Current Defense Mechanisms*
DDoS attacks ①***Direct attacks:*** *TCP SYN/ICMP/IGMP Flooding, UDP Flooding* ②***Indirect attacks:*** *Smurf/TCP SYN/ICMP Flooding, DNS/ NTP Flooding*	***The existing methods (e.g., SIFF, TVA, Phalanx) must require source upgrades and also incur new header. Besides, most solutions lacks the considerations for AS deployment.***
Attacking accountability and privacy ① ***Accountability:*** *spoofing attack, mobility, limited NAT transf.* ② ***Privacy:*** *eavesdropping, monitoring and traffic analysis*	***SAVI can only verify the outbound traffic while ignoring the accountability for inbound packets. APIP relies on a centralized delegate that easily becomes an attack target.***

Attacking secrete key exchanging

① ***Dynamic configuration:*** *Diffie-Hellman introduces Man-in-the-Middle attacks*

② ***Pre-configuration:*** *PKI-based key exchange relies on a centralized authority that may be compromised or misconfigured*

Requirements: the existing solutions are fragmented, which mostly patches the partial security issues. In this case, the inherent method is highly required for ensuring the security of E2E communication.

Network Infrastructure Vulnerability Analysis

Border Gateway Protocol (BGP)

- ***Security threats:***
  - *Prefix hijacking, path hijacking, route leaks, inconsistency between routing behaviors and policies.*

Issues of Current Defense Mechanisms

BGPsec is hard to be deployed due to its low efficiency; ROA can address prefix hijacking, lacking defense for path hijacking.

Domain Name System (DNS)

- ***Security threats:***
  - *DNS spoofing, domain name hijacking/redirection/ cache poisoning, DDoS attacks.*

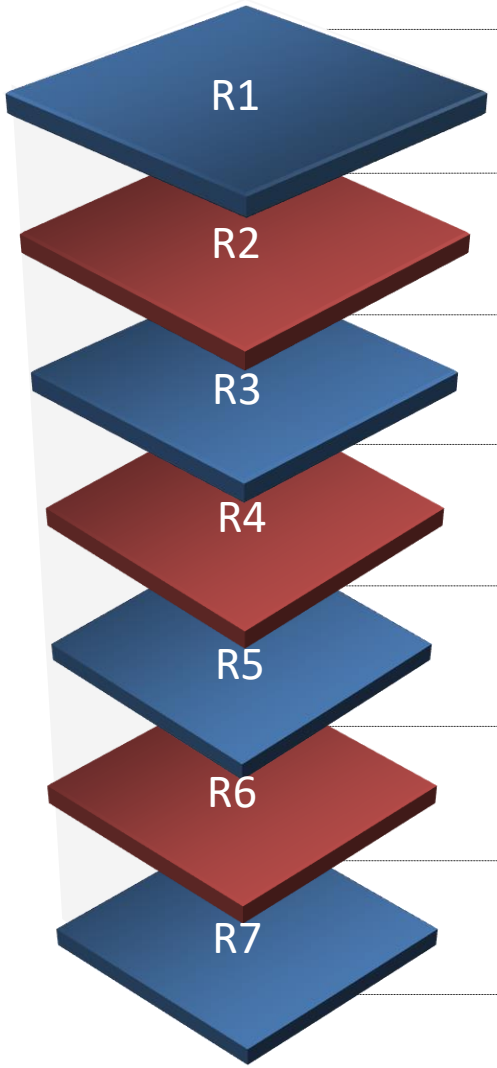
DNSsec can solve most issues but with low efficiency. It works only under the assumption the centralized DNS server is reliable, which may actually be compromised or misconfigured.

The network infrastructure (i.e., BGP and DNS) should be more trustworthy and reliable, which could provide fundamental services for ensuring that the network is working properly in the future.

Without relying on a centralized authority, BGP and DNS can inherently immunize against existing attacks (e.g., BGP route leaks and DNS cache poisoning).

Contents

Conclusion on Requirements



- Differentiated Service – deterministic on latency and jitter
- Better User-Network Interfaces
- More Information from Network Itself
- Network Automation and Intelligence
- IP Address Namespace and Fields Flexibility
- Separating ID Semantic out of IP Address
- Inherent Security & Trustable Infrastructure

Thank you

www.huawei.com