Becoming Disconnected

Tom Daly
tjd@q7.io

Acknowledgements

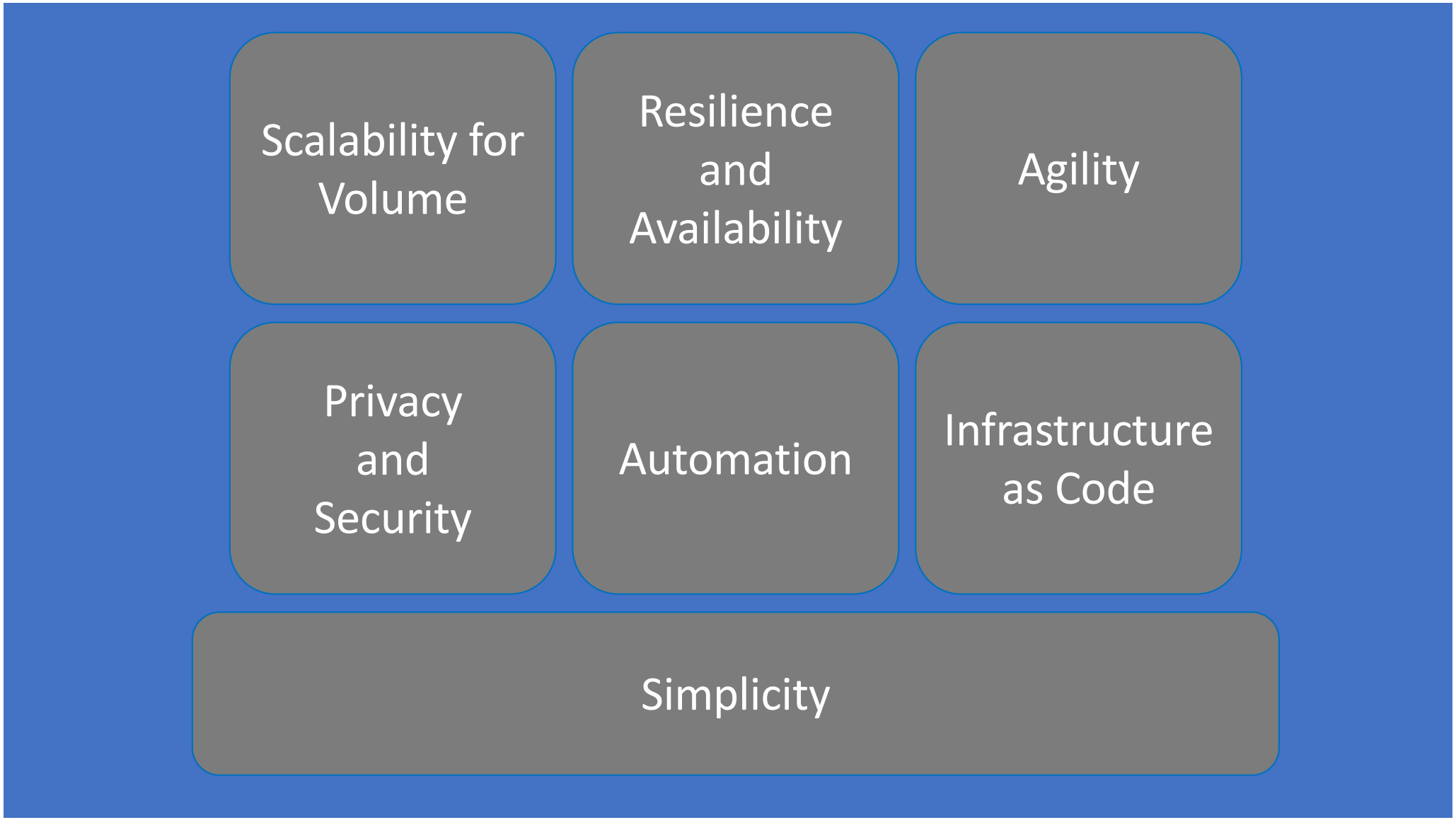
- Jana Iyengar, Distinguished Engineer, Fastly; Editor, IETF QUIC WG
- Mark Nottingham, Principal Engineer, Fastly; Chair, QUIC WG
- Ryan Landry, Technical Operations, Fastly
- Ian Swett, Google
- Tim Hoffman, Edge Reach Consulting
- Aaron Hughes, 6Connect
- Job Snijdners, NTT

Perspective

- Building Content Delivery and Edge Compute at Fastly

Disclaimer

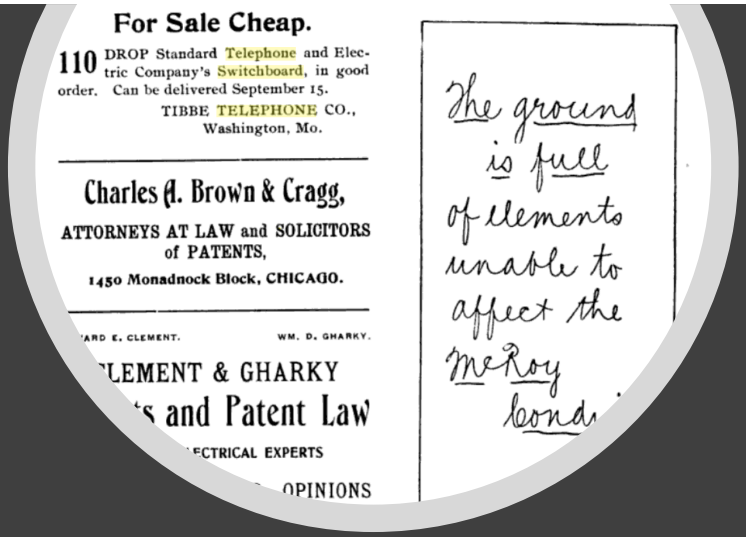
- The views expressed in this talk do not express the viewpoints of Fastly, Inc.


Scalability for Volume
Resilience and Availability
Agility
Privacy and Security
Automation
Infrastructure as Code
Simplicity

INNOVATION

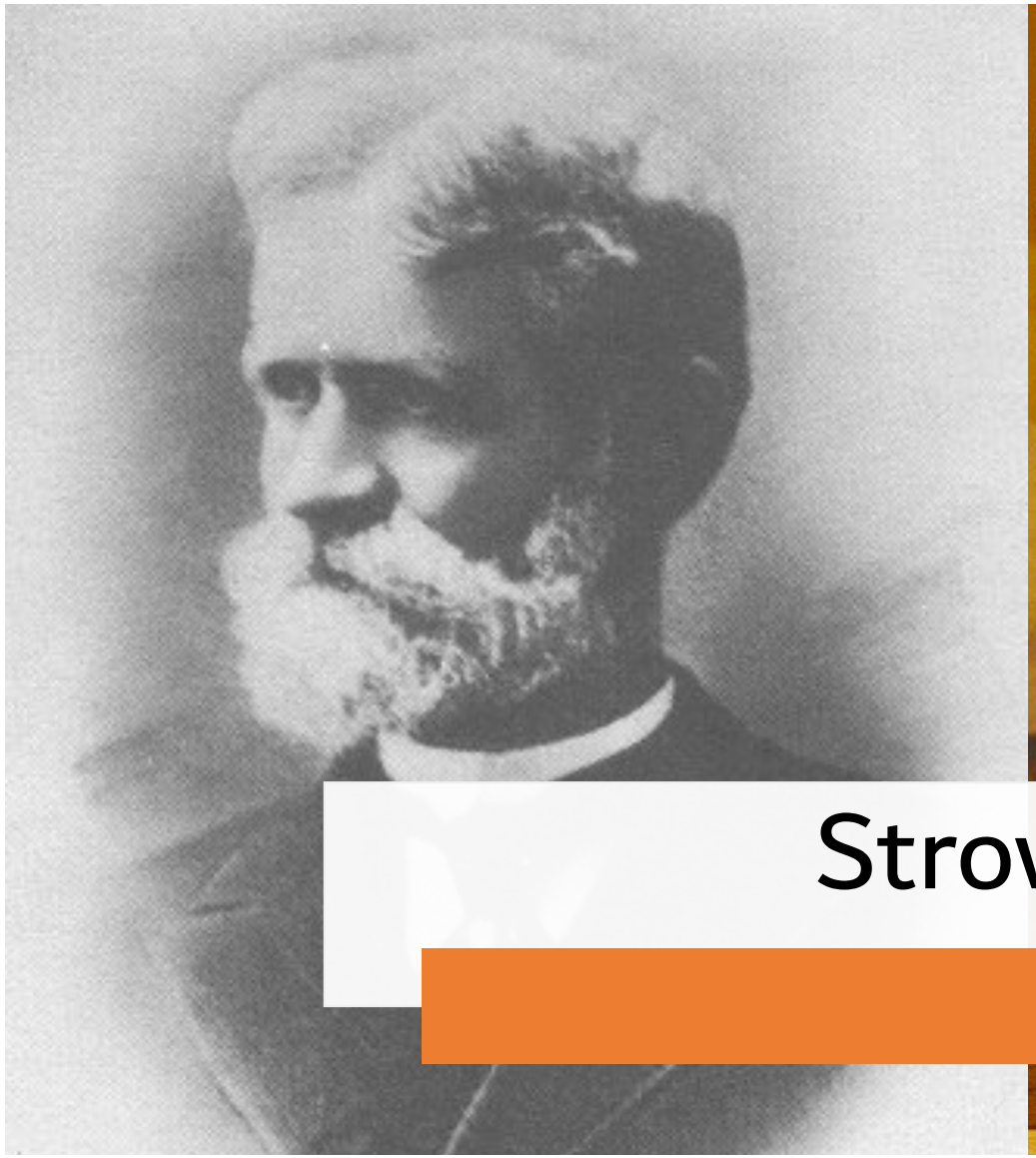
INNOVATION

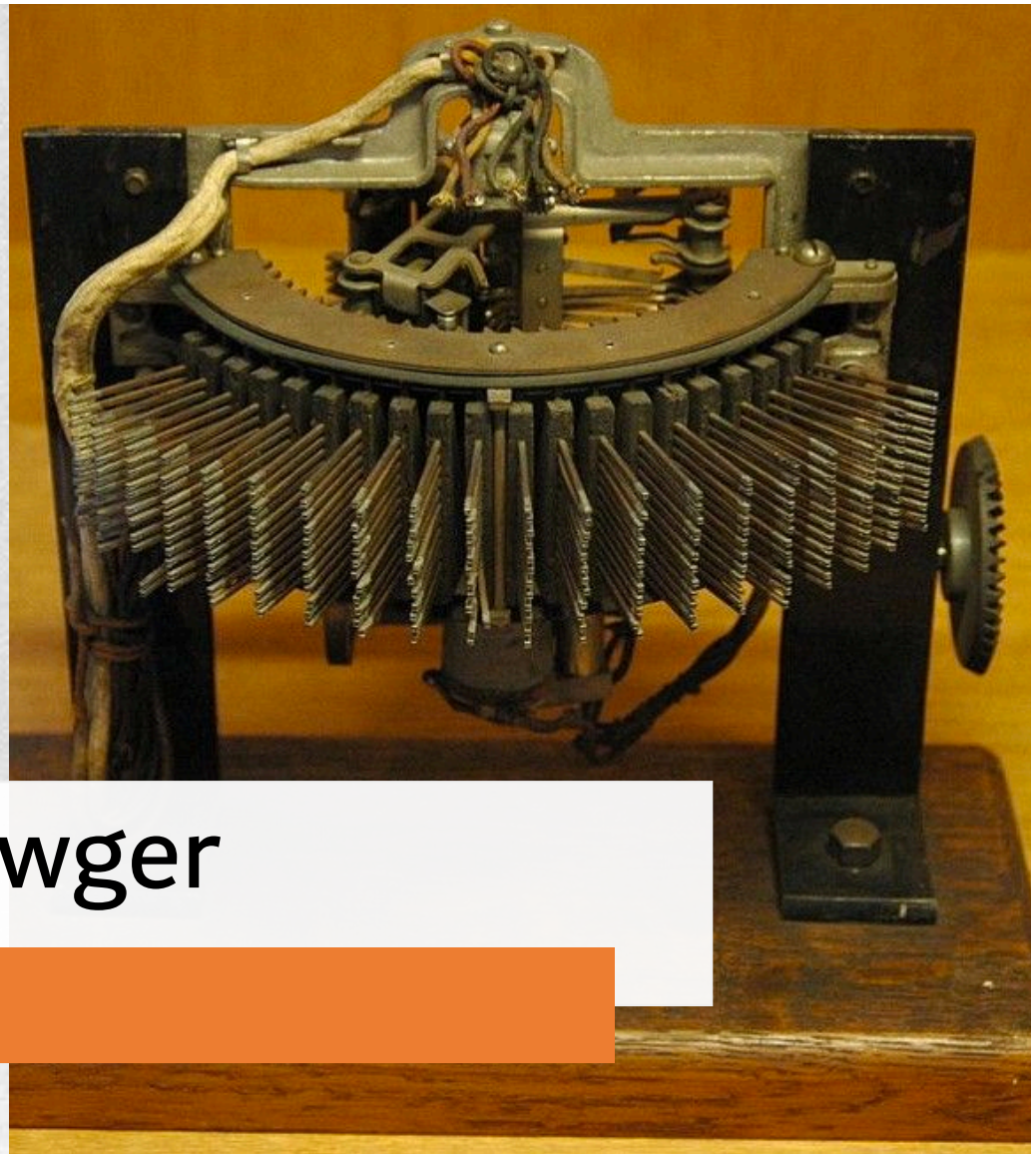
- React to Changes
- Discover New Opportunities
- To Create Competitive Advantage
- To Scale Delivery
- To Drive Efficiency
- To Attract Top Talent
- Serendipity



110 Lines @ 9.6kBps

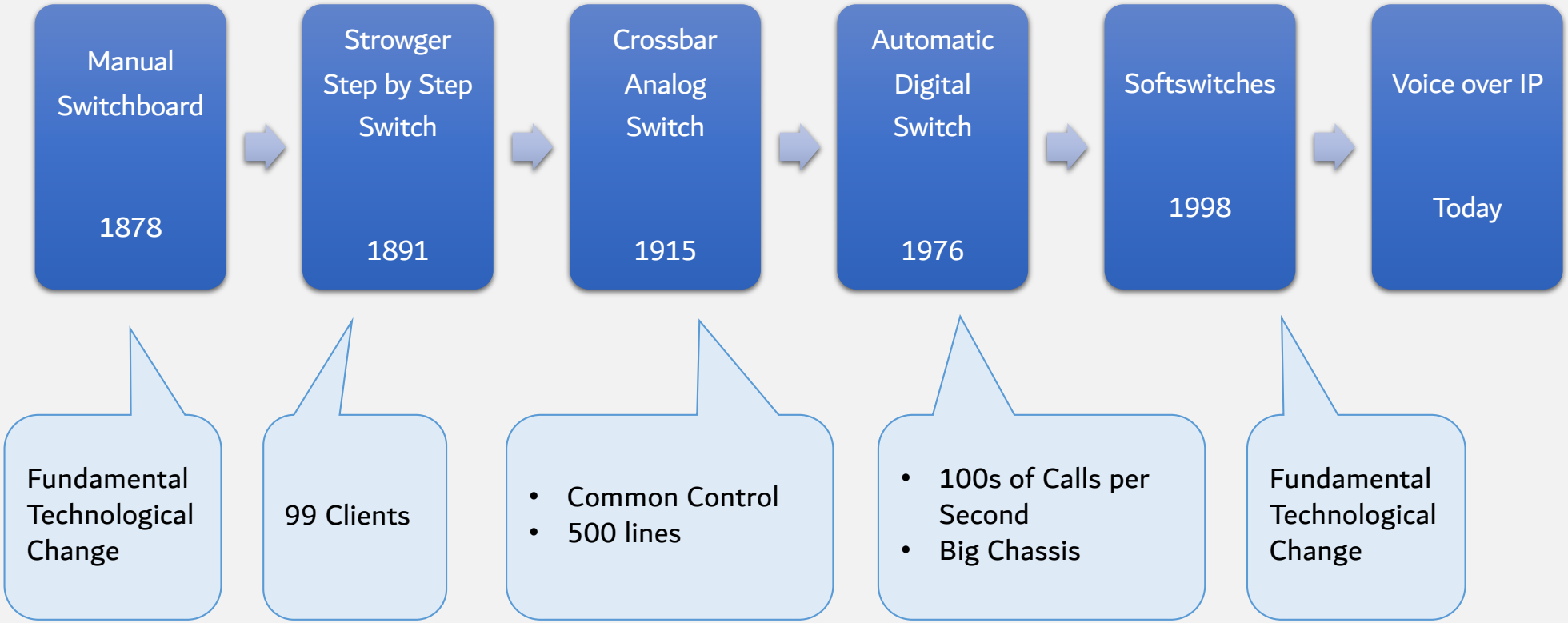
Telephone Magazine August 1900

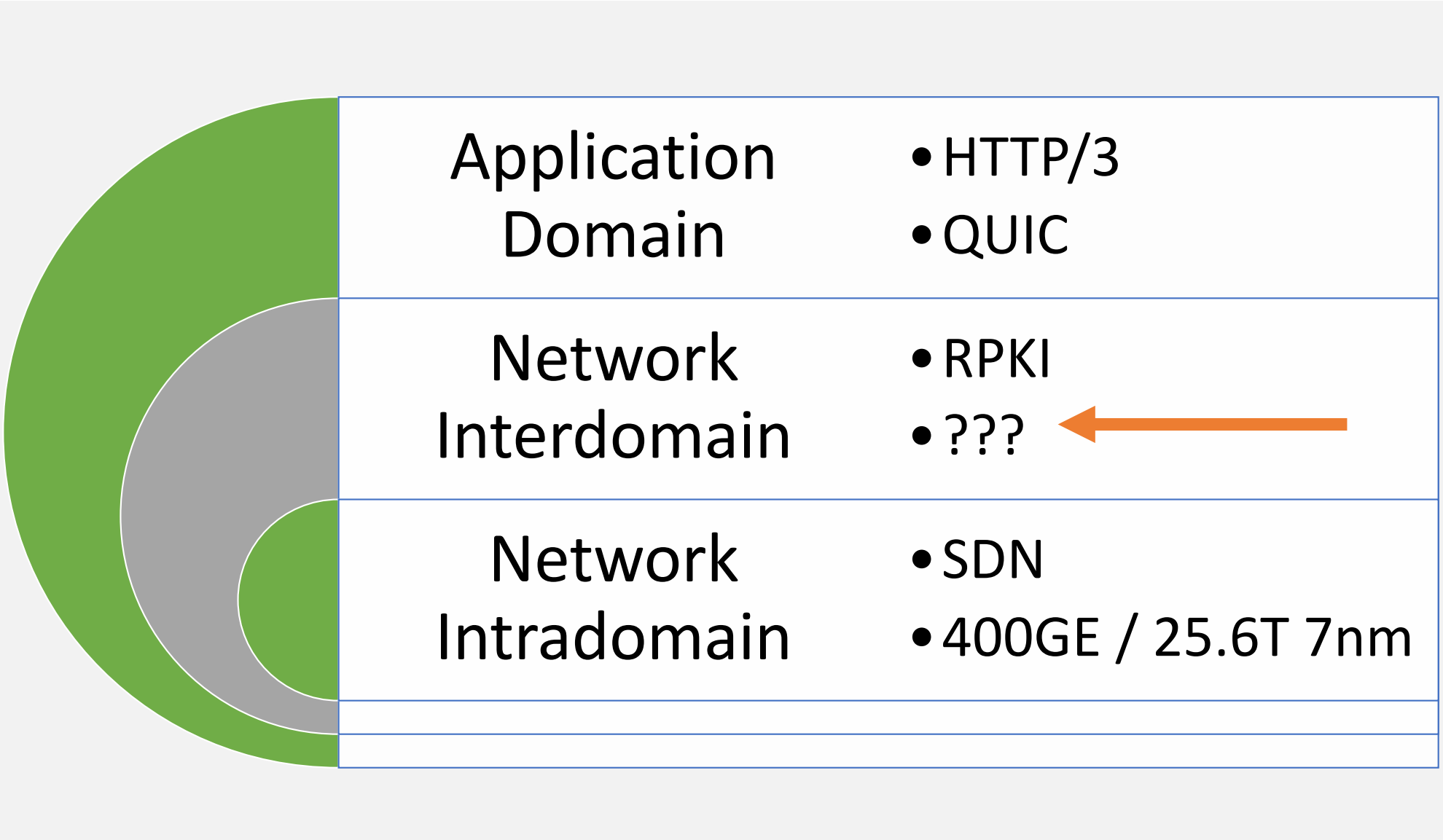




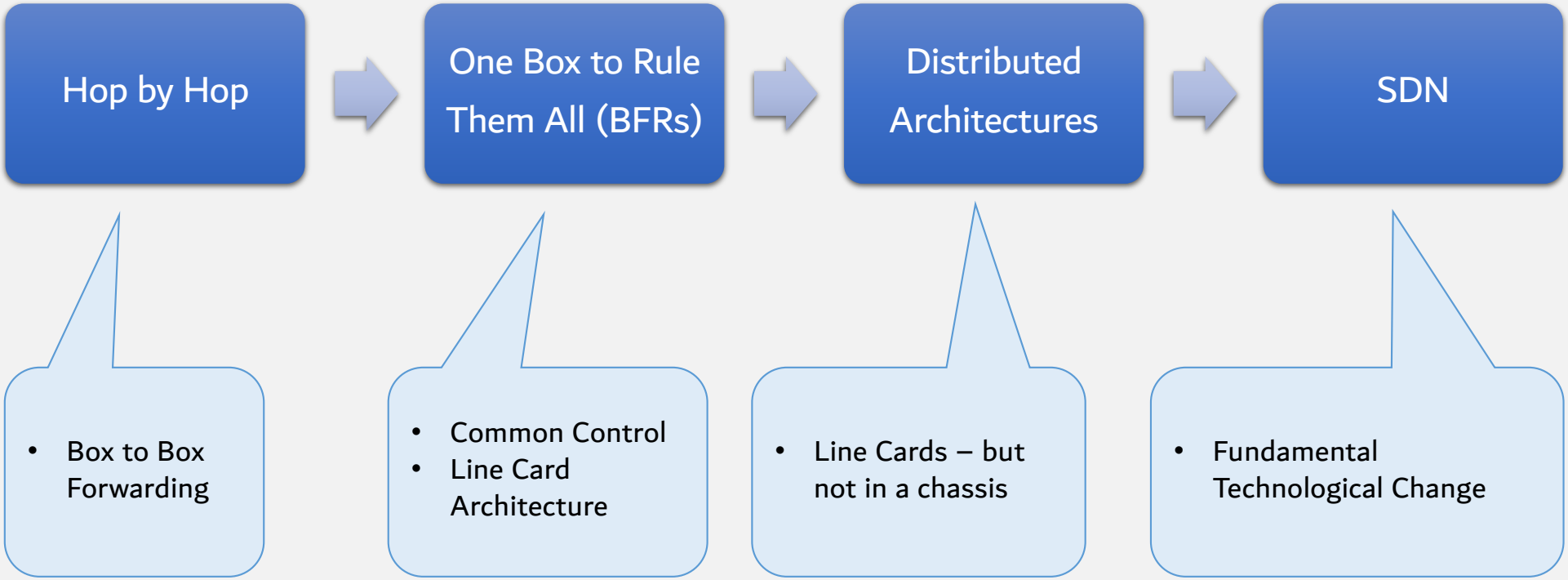
Strowger

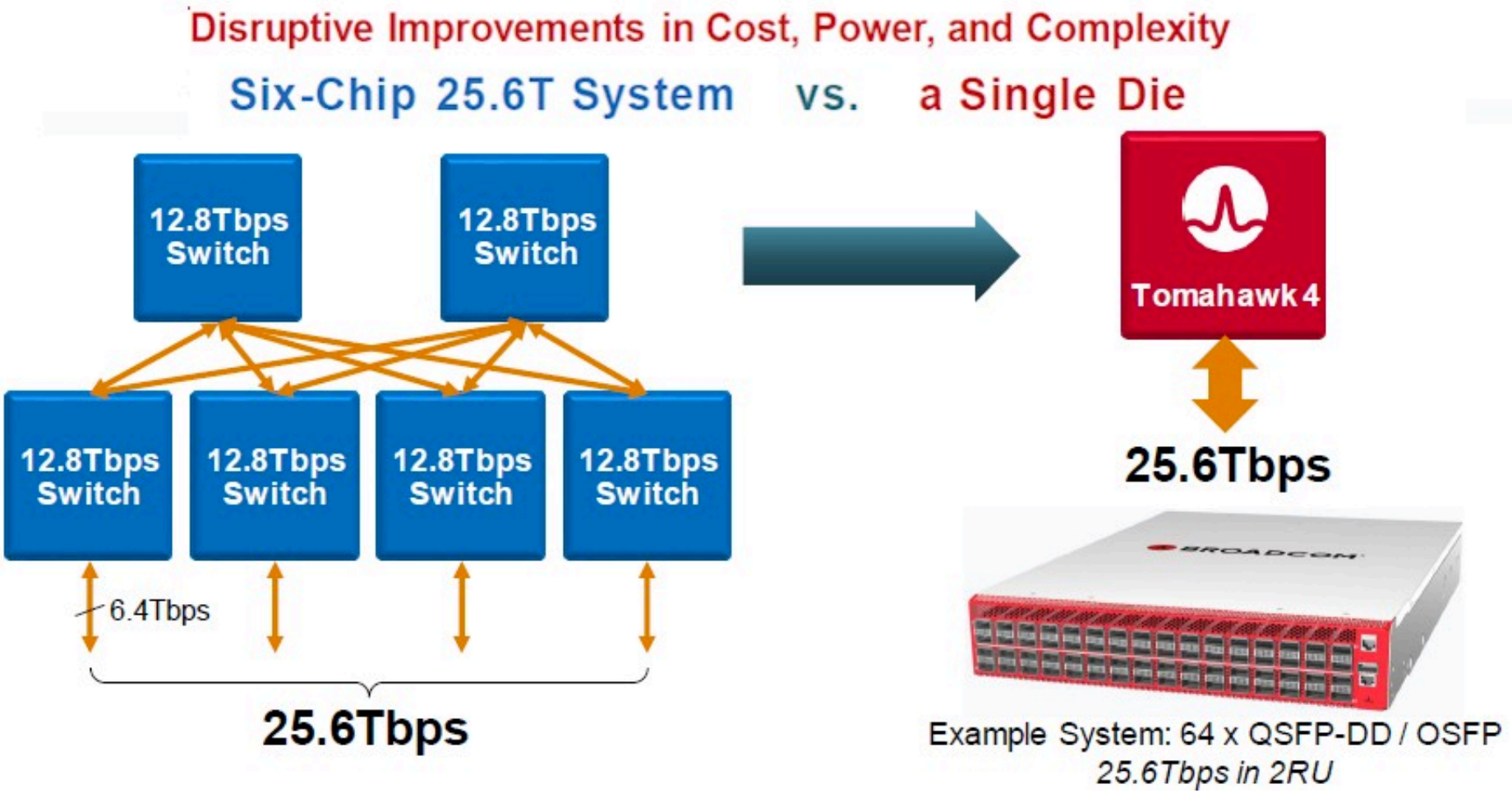
Where did we go: Voice Switching




Application Domain
• HTTP/3
• QUIC
Network Interdomain
• RPKI
• ???
Network Intradomain
• SDN
• 400GE / 25.6T 7nm

Where did we go: IP Routing

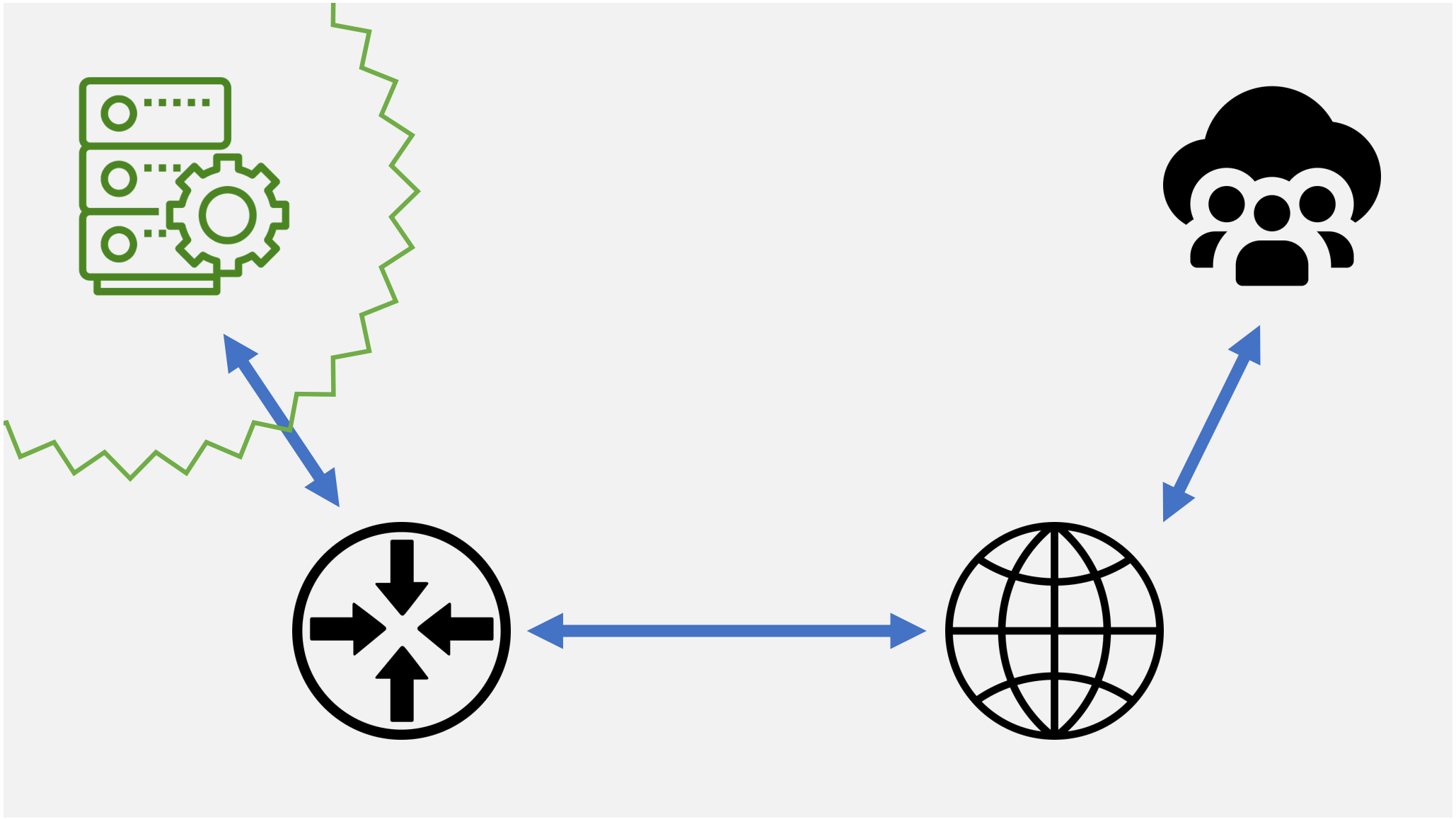


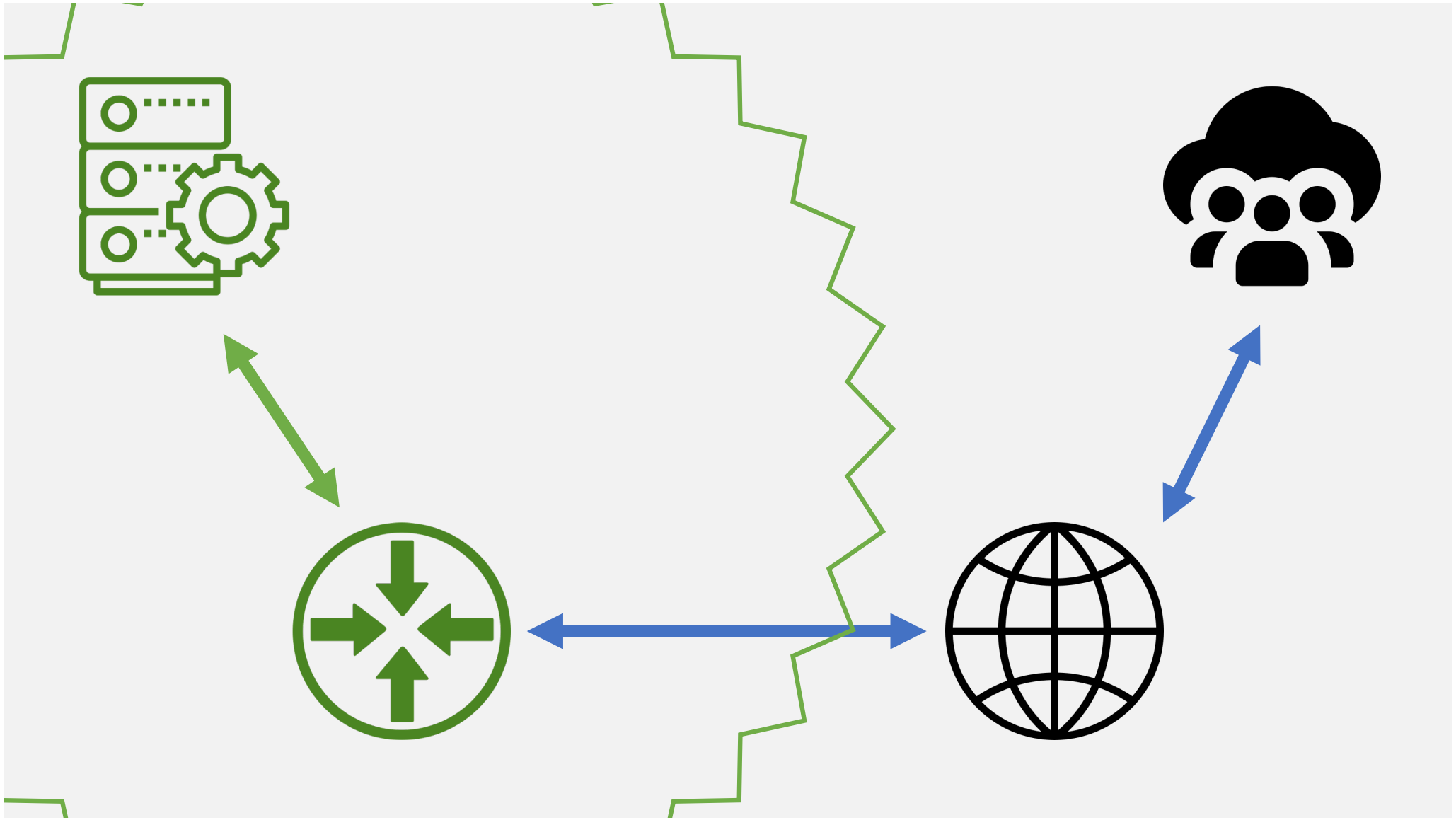


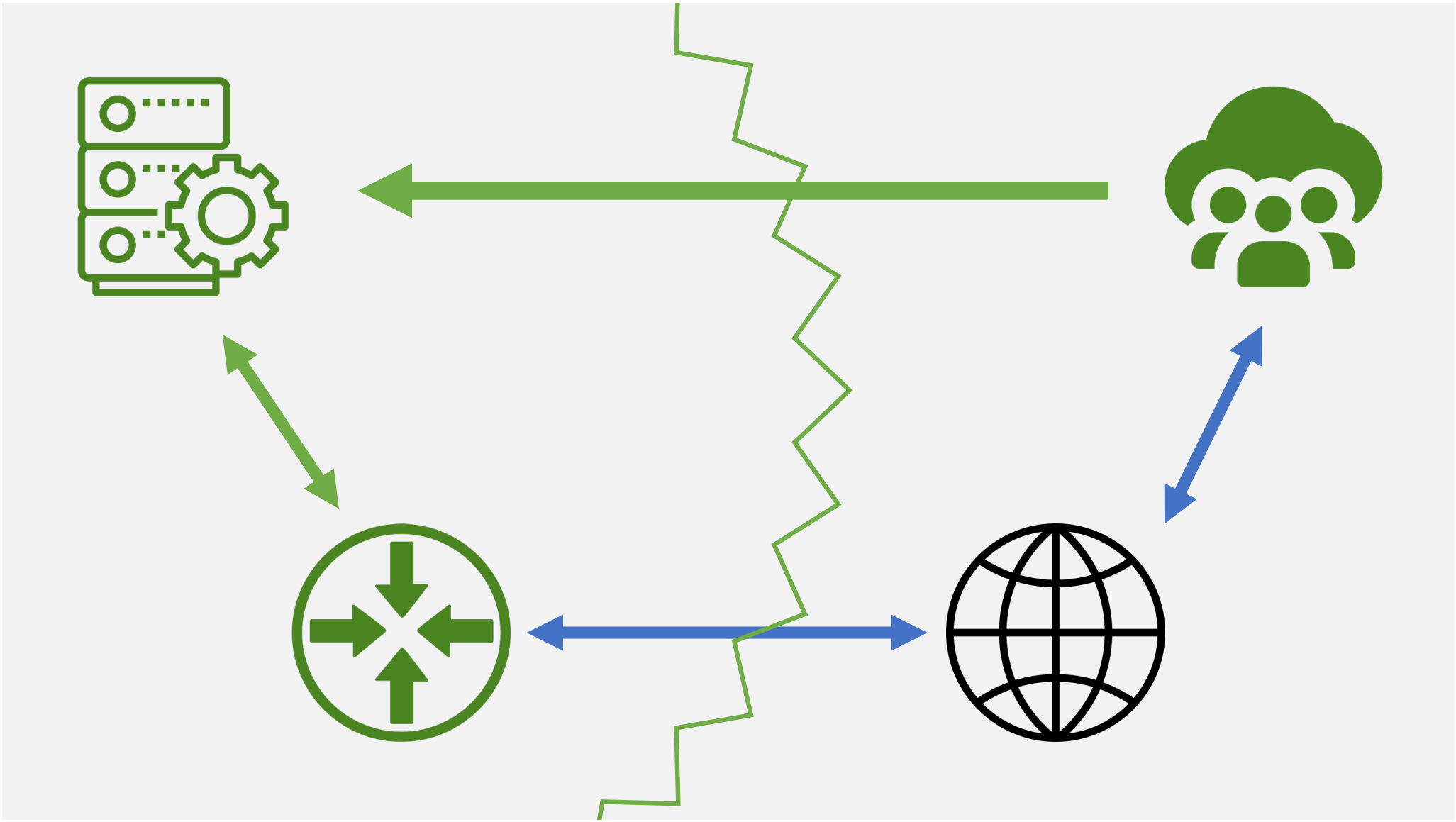
Credit: Broadcom / TheNextPlatform.com

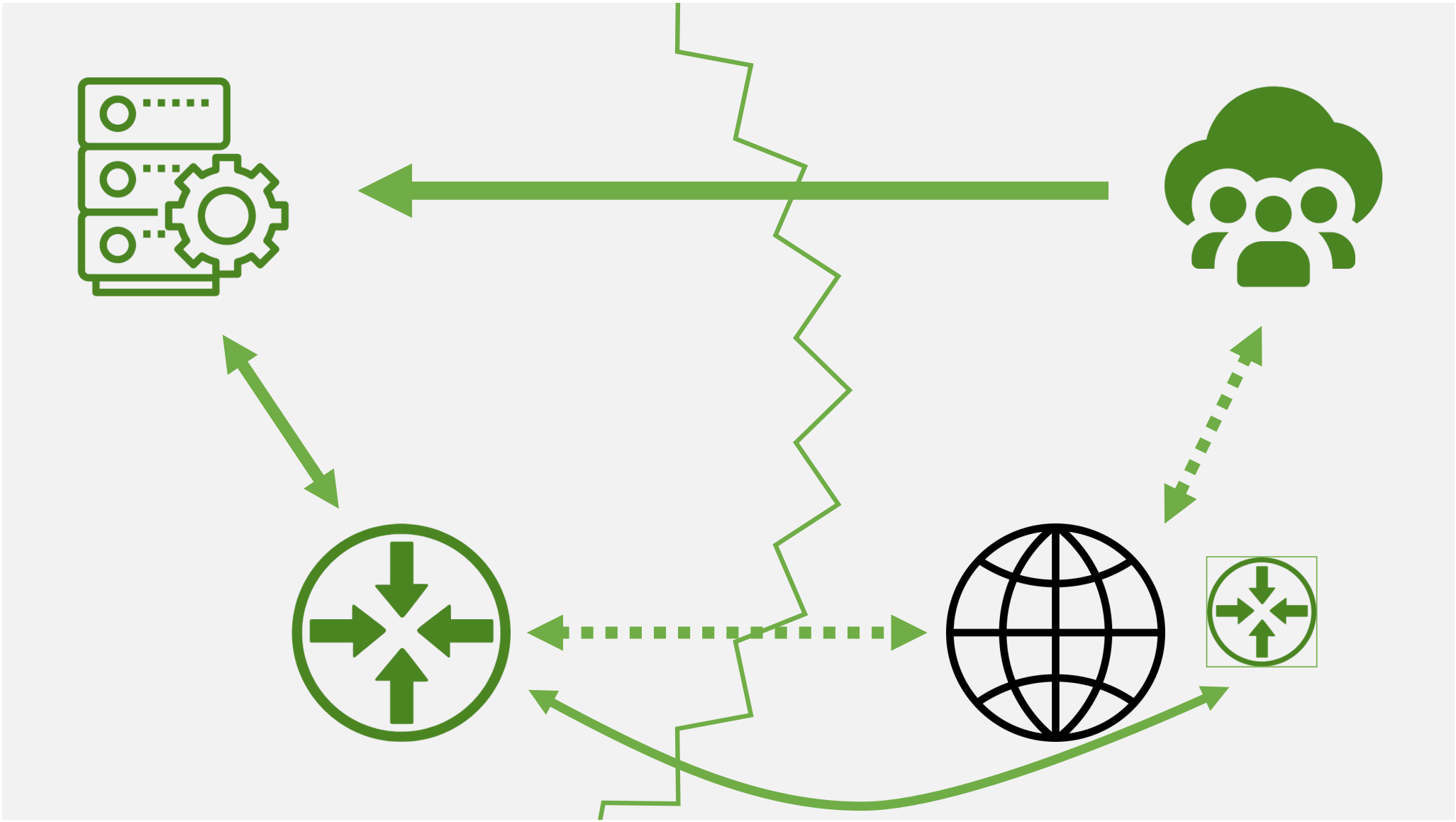
SDN: Prevalence of Single Sided Intelligence

- Abstracted traditional inter-domain routing protocols.
- Each autonomous system can continue to form its own world view of the Internet, as we historically have.
- However, the intermediate ASNs participating in remain opaque to the transaction.
- Therefore, the natural path forward is to extend the edge as far as possible to reduce intermediate ASN participation.
- Server / Client Out of Band telemetry enriches our view of the network.







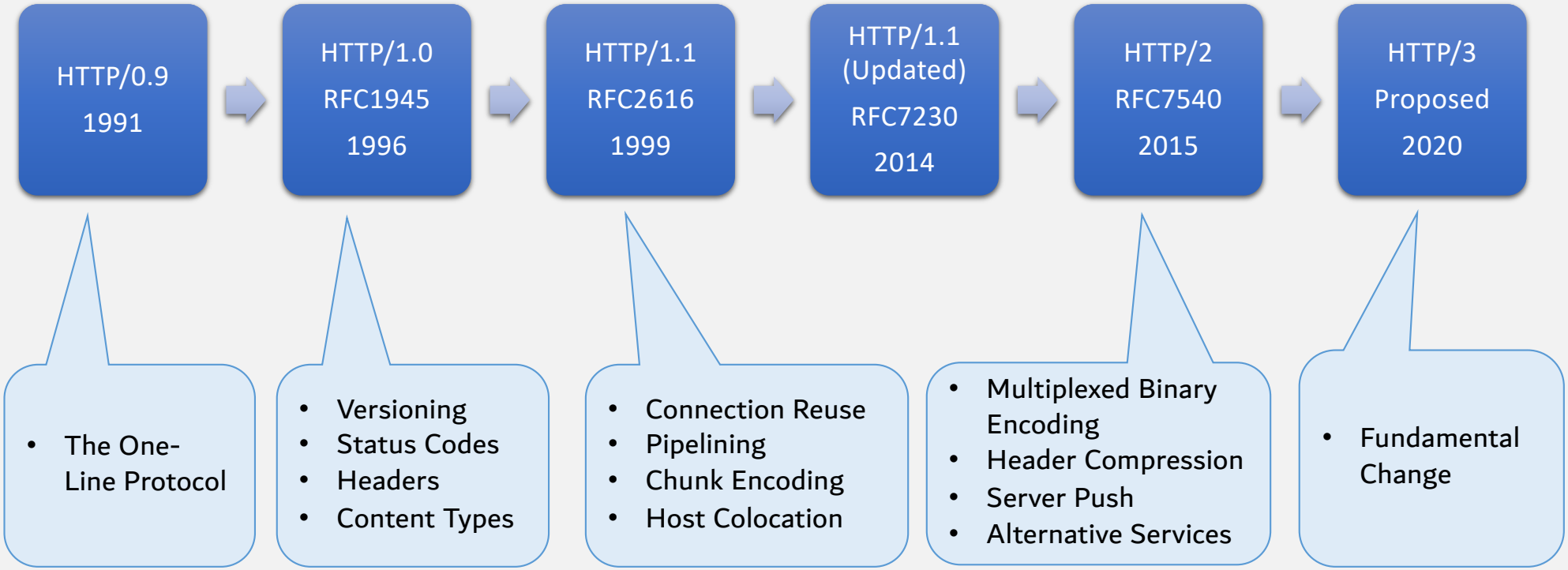


Interdomain Information Asymmetry

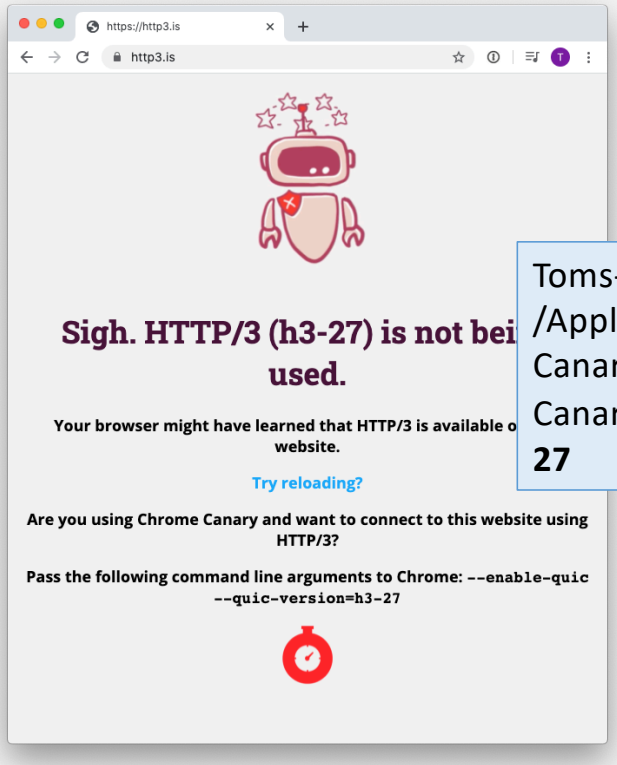
- *"For markets to function **efficiently**, participants must have reasonably good information about things such as prices, quality, available technologies, and the risks associated with working in certain jobs or consuming certain products. When participants in the market have **incomplete information** about such thing, the result will be **inefficiencies** in the input use and in firms' output." **
- IOW: We each know our own network better than anyone else can know our network.
  - Policy
  - Economics
  - Constraints
  - Preferences

** Baer, Prince; Managerial Economics and Business Strategy; pp. 542*

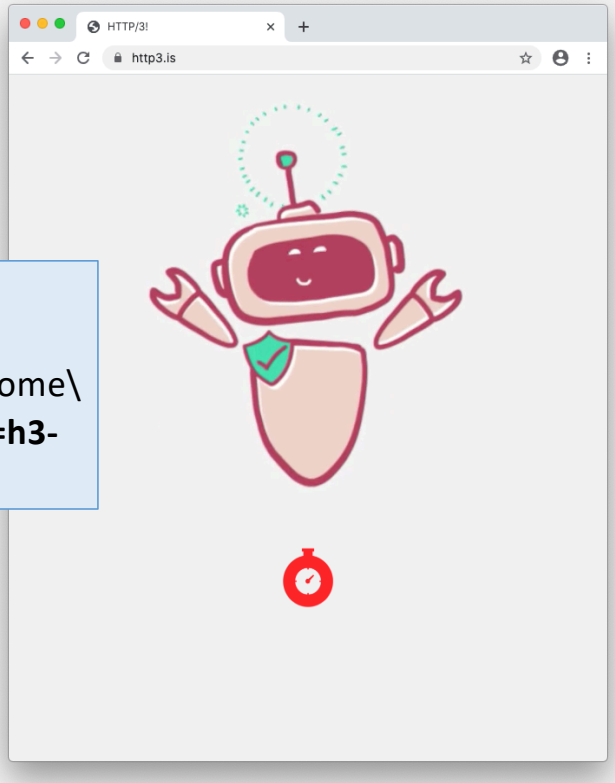
Abridged History of HTTP



HTTP/3 is coming – demo at http3.is



Toms-MacBook-Pro:~ tjd$ /Applications/Google\ Chrome\ Canary.app/Contents/MacOS/Google\ Chrome\ Canary --args **--enable-quic --quic-version=h3-27**



HTTP/3

HTTP/3 is the upcoming version of Hypertext Transfer Protocol.

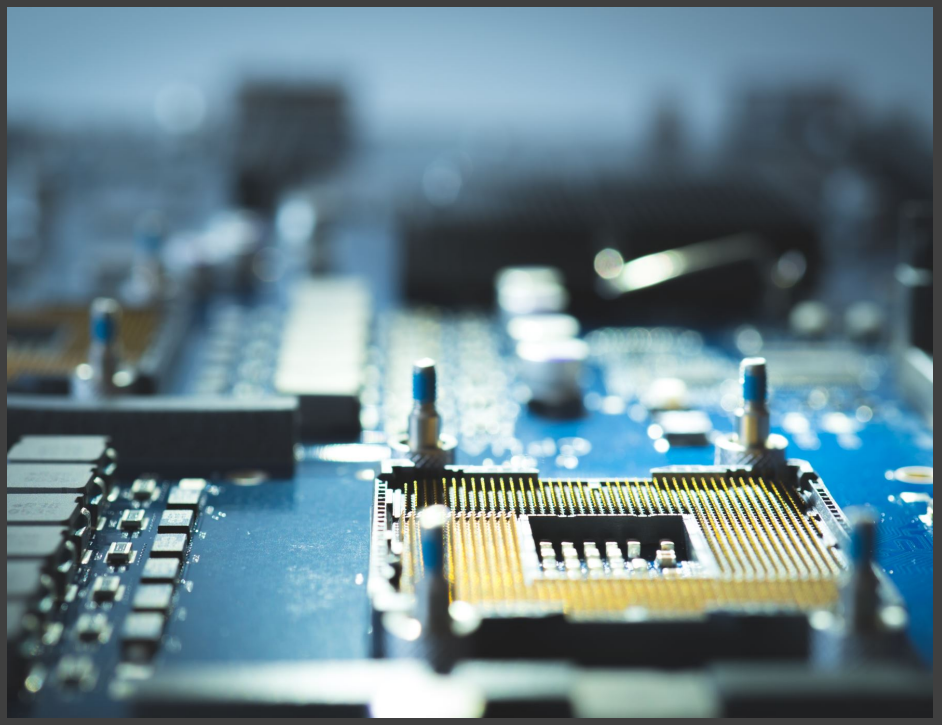
Multiple draft server / client implementations available.

Feature parity from HTTP/2

Carried over QUIC; a new transport protocol.

HTTP/2 over TCP

- Binary Protocol
- Multiplexed to remove order delivery dependency and head of line blocking
- Compresses headers natively
- Populate client data caches using server push
- Alt-Svc allows “DNS CNAMEs inside of HTTP”



HTTP Alt-Svc

Problem: HTTP/2 connections are meant to be long lived.

BGP Anycast topology changes break TCP connections.

DNS GSLB changes require new lookups.

Solution: Allow servers to direct clients to an alternative URI to fetch the same content.

Security concerns aside – this works.

HTTP/3 over QUIC over UDP

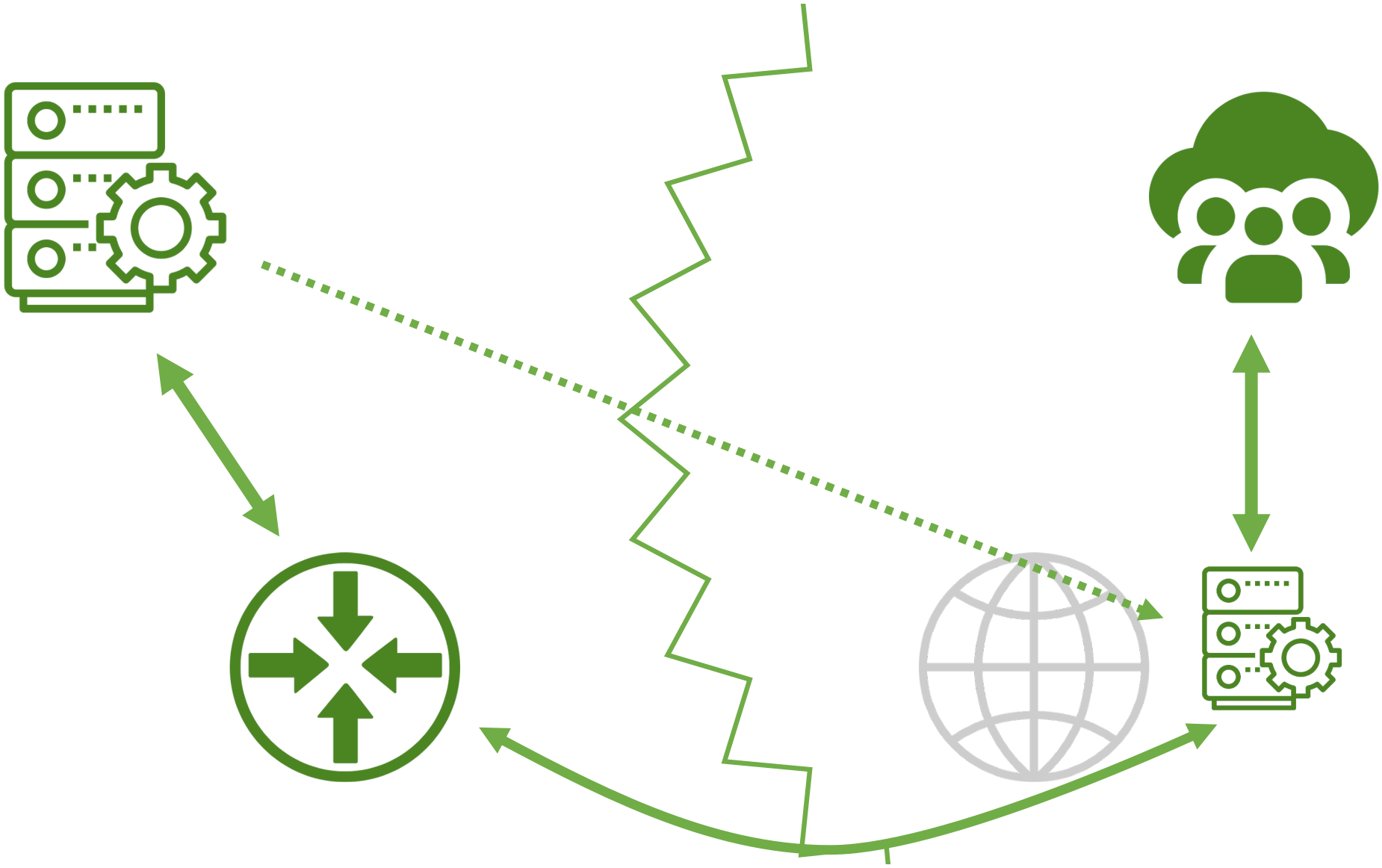
- Performance
  - Resolves head of line blocking
  - Allows for more performance over high latency or high packet loss networks.
  - 20% reduction in YouTube rebuffering over 150+ms / 5% packet loss networks. *
- Connections
  - Defined as part of QUIC, not by the traditional 5-tuple
- Low latency handshakes
  - When combined with TLS 1.3
  - <1.5 RTTs to encrypted communications vs. 2+ RTTs with TLS 1.2 over TCP
  - 0-RTT resumption for repeat handshakes.
- Encryption
  - Other than UDP header and QUIC destination connection ID, the payload is encrypted; reduced metadata leakage

* Source: Google

Mobility with QUIC

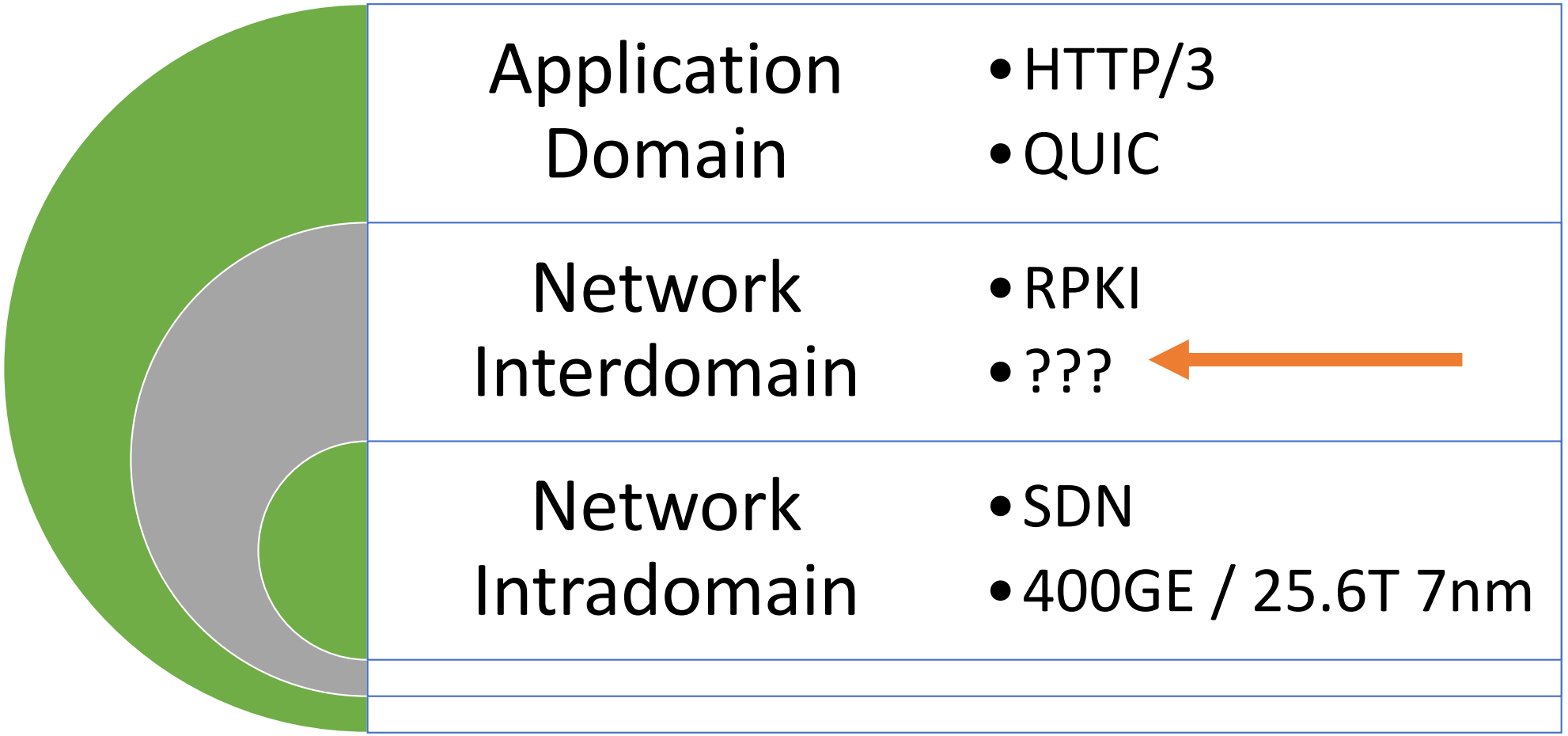
- Optimize serving stack location through HTTP Alt-Svc
  - Obviates need for Anycast or DNS
- Move QUIC Connections around network in a DSR-like manner as long as session state can be migrated.
- Think Kubernetes Containers or Serverless Functions along the Edge.

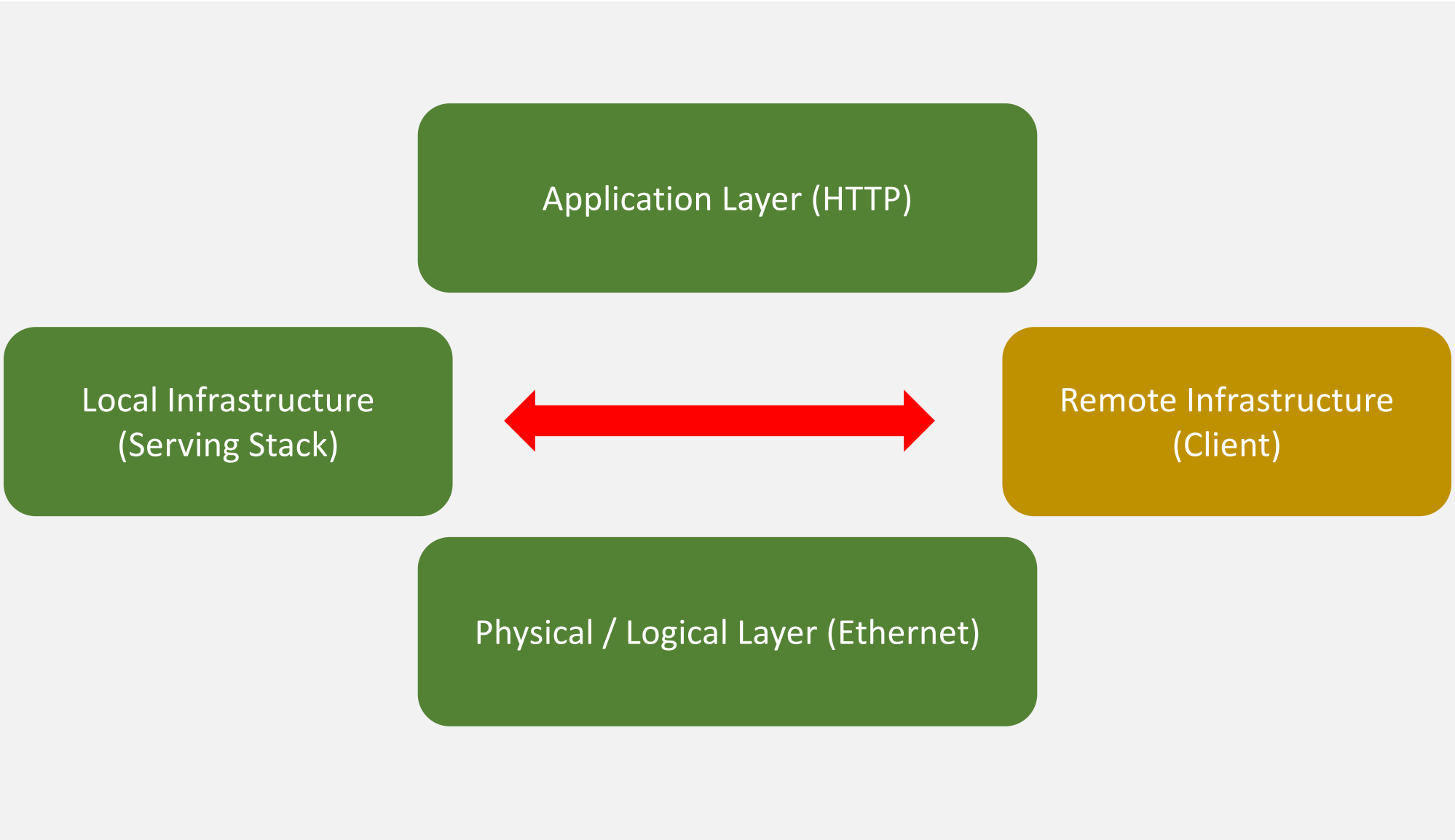
- Solve the “Parking Lot” handoff problem.
- Connect to HTTP Alt-Svc endpoints depending upon local network conditions.



Fundamental Changes for Network Operators

- Network Middleboxes can't touch QUIC
  - interference results in connection resets.
- Routing decisions are moved upward in the protocol stack
  - HTTP Alternative Services (Alt-Svc)
- DNS over HTTP (DOH)
  - Localized Effects of "My Internet" is not "Your Internet…"
- Reduced observability; less traffic classification; perhaps less opportunity for value creation?
- READ: Manageability of the QUIC Transport Protocol (draft-ietf-quic-manageability-06)


Application
Domain
• HTTP/3
• QUIC
Network
Interdomain
• RPKI
• ???
Network
Intradomain
• SDN
• 400GE / 25.6T 7nm


Application Layer (HTTP)
Local Infrastructure
(Serving Stack)
Remote Infrastructure
(Client)
Physical / Logical Layer (Ethernet)

The “Remote” Network is Disconnected from Developers

Guessgineering in the Application Stack

- Developers have far more access to control the network behavior than they did 10 years ago.
  - Developers can write more client code than ever before.
  - Single companies under common mission control server, network, and client code.
  - Advent of ”apps” remove constraints of browsers as clients.
  - HTTP/3 and Alt-Svc provides the same in a browser.
- The software can just understand the conditions of the network.
- Developers can iterate faster; e.g. QUIC stack in userspace

Information != Control

“Suggestive” Signaling?

Preferred backup paths for a peer; e.g.:

- 7018:191:2914: AT&T desires primary overflow traffic via NTT
- 7018:192:3356: AT&T desires secondary overflow traffic via CenturyLink
- 7018:193:174: AT&T desires tertiary overflow traffic via Cogent

Alternative entry points to a network; e.g.:

- 7018:18:1: Congestion in Dallas; please handoff in Houston
- 7018:18:2: Congestion in Dallas; please handoff in Atlanta
- 7018:18:3: Congestion in Dallas; please handoff in Ashburn

Traffic classification by content VIP; e.g.:

- 54113:200:1: This prefix carries real time traffic.
- 54113:200:2: This prefix carries website traffic.
- 54113:200:3: This prefix carries bulk download traffic.

Bilateral Session Establishment at an IXP; e.g.:

- 7018:99:54113: We see you via the Route Server and are ready to peer via BLPA sessions.

Information
drives Simplicity

Reducing Waste?

Overprovisioning the network because we don't know where to best route traffic?

Better yet; alternative paths when congestion or outages occur?

Moving workloads based upon network conditions?

Human energy spent debugging, reconfiguring, engineering to find ideal paths when computers do it better anyways?

We decided this was poor in the Network Intradomain when we decided to invest in SDN.

Objections?

- ”My network topology and current network conditions are my competitive edge.”
  - Too late: Telemetry already tells us the current state of your network.
- “My router can’t run more processes...”
  - SDN and RPKI have already taught us we can sideload data into the network.
  - My routers don’t run my vendor’s BGP implementation anymore.
- “You’re forcing the network to be a least common denominator”
  - I’m trying to encourage efficient allocation of resources
- ”This information allows for network arbitrage”
  - It’s already happening...

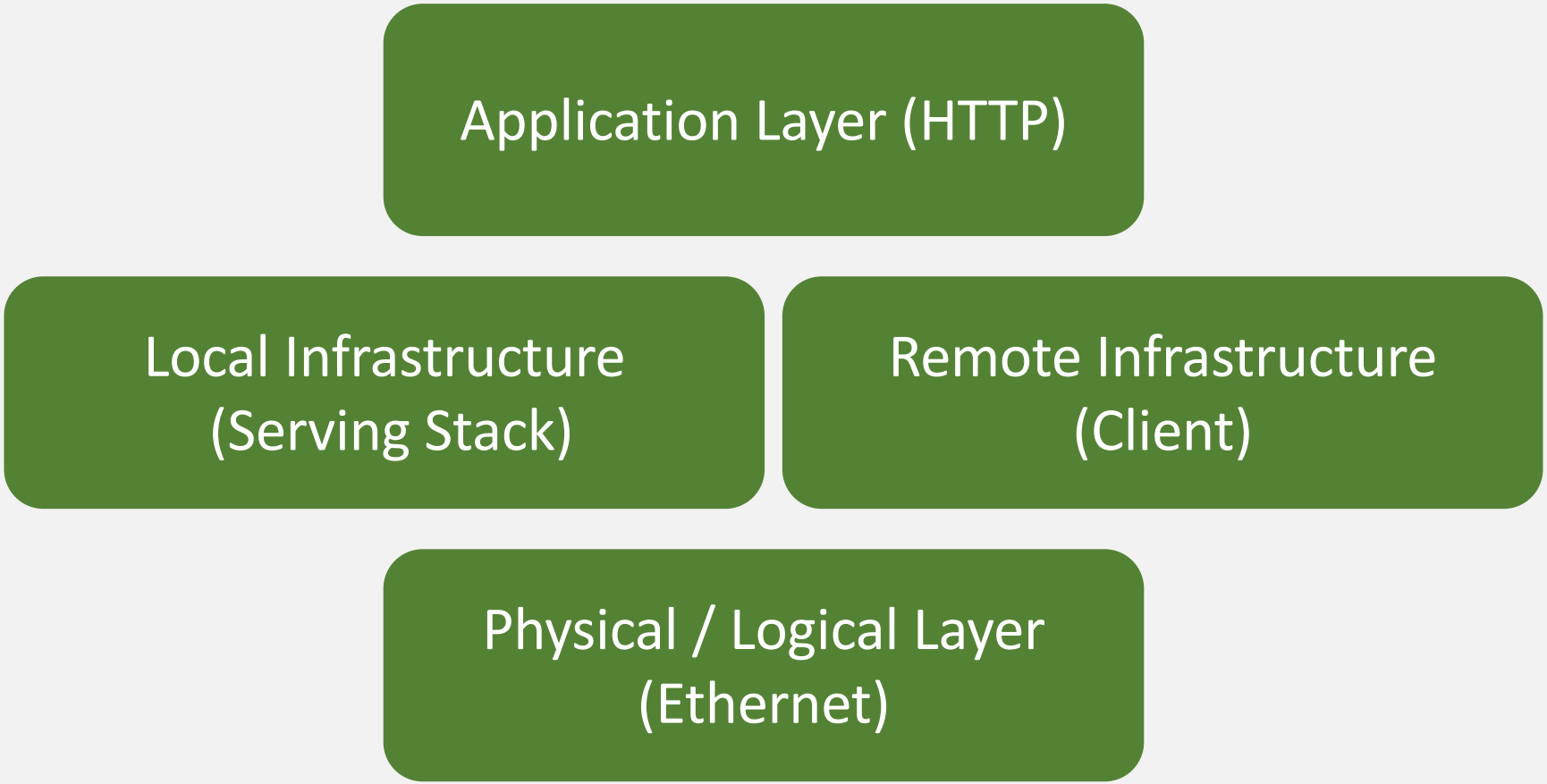
Where can we go?

Fundamental Technological Change

Information Sharing

Human Networking

Fundamental Technological Change?


Application Layer (HTTP)
Local Infrastructure
(Serving Stack)
Remote Infrastructure
(Client)
Physical / Logical Layer
(Ethernet)

Thank You / Questions

Tom Daly
tjd@q7.io