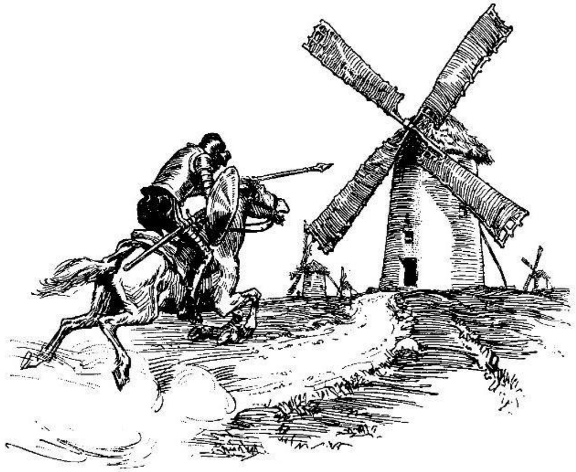
The Challenge of Interdomain Innovation

Ken Calvert

University of Kentucky



Acknowledgments

Thanks to many great collaborators:

- Ilya Baldin
- Bobby Bhattacharjee
- Rudra Dutta
- Jim Griffioen
- Najati Imam
- Billy Mullins
- Anna Nagurney
- Leon Poutievski
- George Rouskas
- Amit Sehgal
- Neil Spring
- James P. G. Sterbenz
- Wen Su
- Tilman Wolf
- Ellen Zegura

Support from NSF, DARPA, Intel, and Cisco is gratefully acknowledged.

None of them are to blame for errors, outlandish ideas, etc.

Agenda

I. BLUF + Background
II. Some network-layer innovations you've never used
III. Some network-layer innovations you likely have used
IV. Success factors for innovation at the network layer
V. Possible way forward

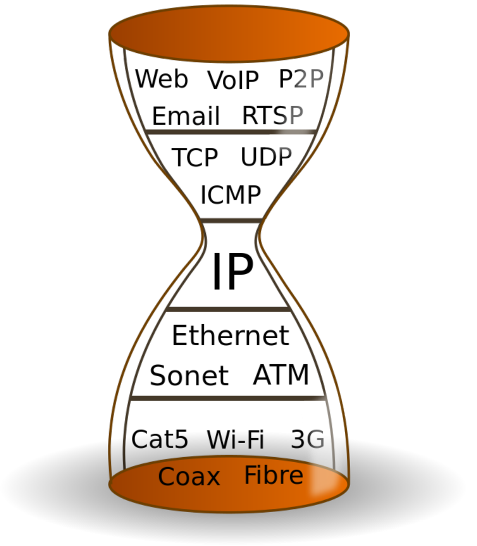
Bottom Line Up Front

The interdomain interface is the "waist of the hourglass".

Changing that interface is hard – for good reasons.

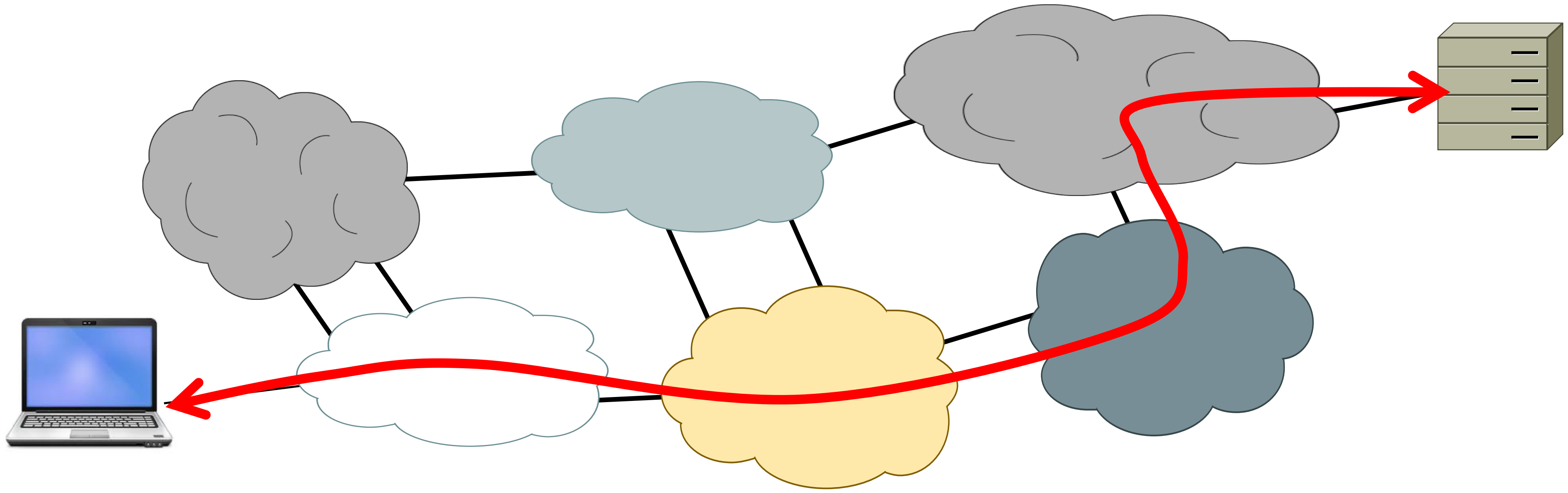
There has been little innovation in that space over the last two decades.

SIGCOMM 2020: 53 papers, 1 paper on inter-domain topics



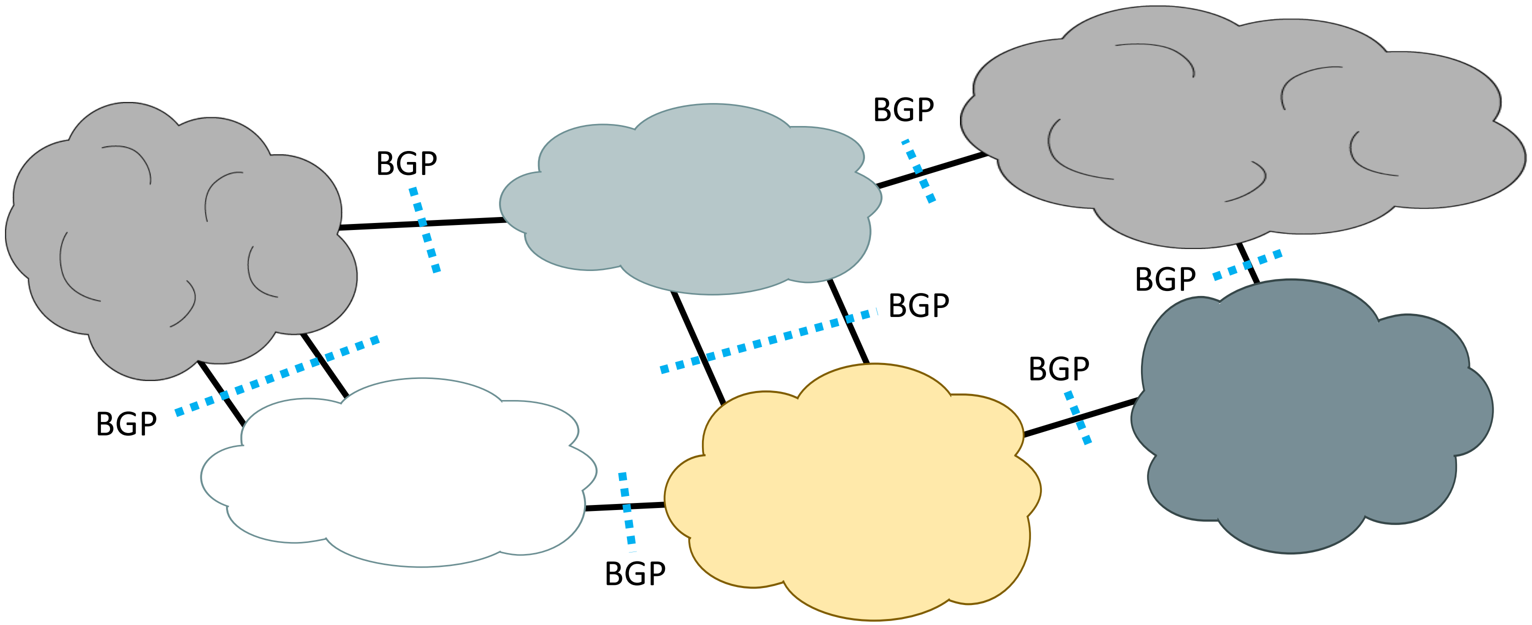
A prerequisite for innovation in end-to-end services is innovation in the ecosystem.

Background: Internet Structure



- Internet: network of networks or Autonomous Systems (AS's).
  - "Access" or "edge" networks: serve end users
  - "Transit" networks: serve other networks
- Virtually all packets traverse multiple AS's going from source to destination.

Background: Internet Structure



- The wire interface between providers is IP and BGP.
- The business interface between providers is outside the architecture.
  - Coarse-grained, slow-changing

Agenda

I. BLUF + Background ✓
II. Some network-layer innovations you've never used
III. Some network-layer innovations you likely have used
IV. Success factors for innovation at the network layer
V. What is needed

Background: Active/Programmable Networking

- DARPA research program, mid-90's – early 2000's
- Motivation: Enable new services by making the network programmable
  - "Put Information Infrastructure on the Technology Curve"
- Research Questions:
  - Programming model?
  - Security??
  - Killer application?

1. Composable Active Network Elements (CANEs) Project Goals

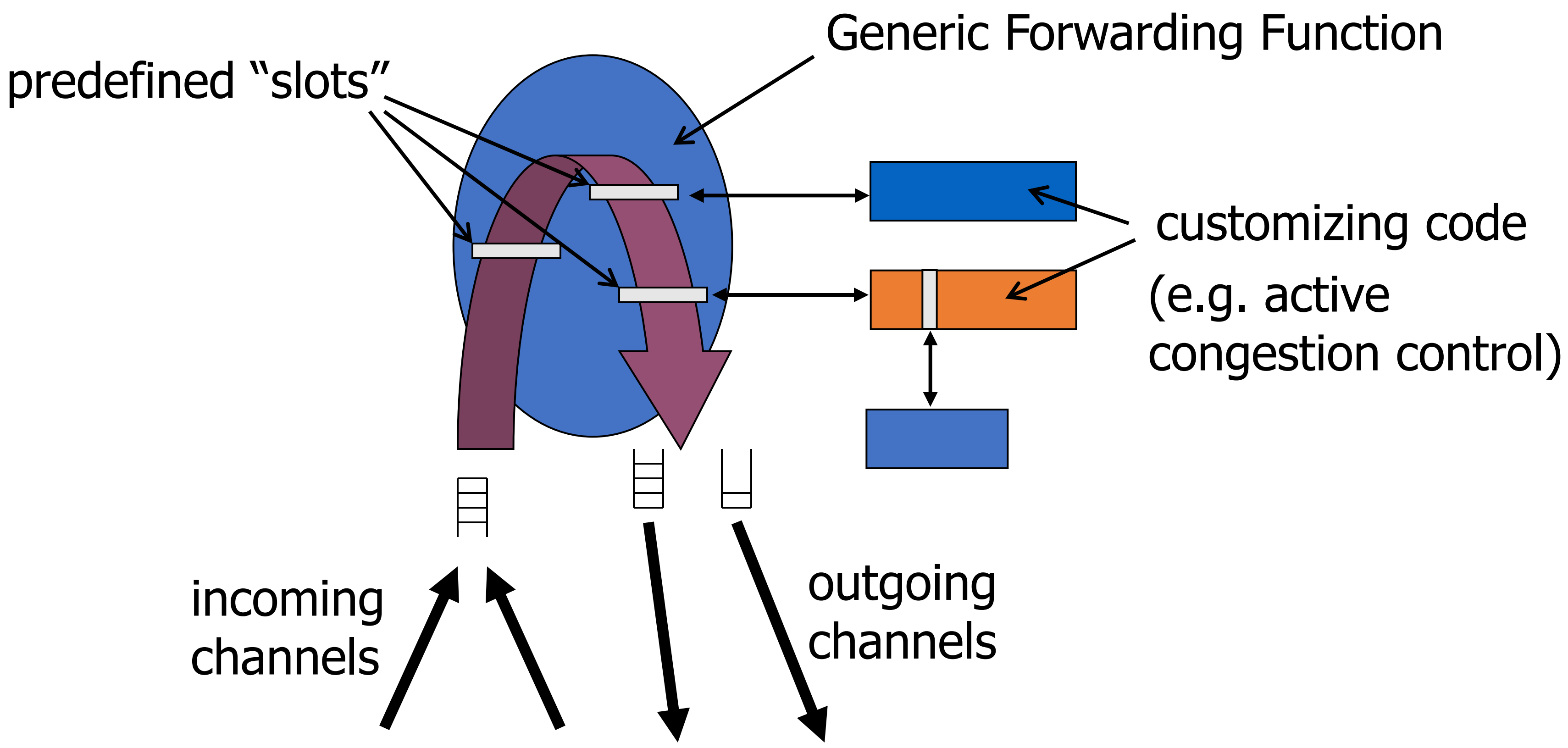
- Show benefits of user-controlled functionality in the net
  - "Bring application knowledge and network knowledge together in space-time"
    - Application-specific adaptation to congestion
    - In-network caching
    - Reliable multicast
    - Mobility

S. Bhattacharjee, K. L. Calvert, E. W. Zegura, "Self-Organizing Wide-Area Network Caches", Proceedings of IEEE INFOCOM '98, San Francisco, April 1998, pp. 600–608.

- Allow for fast forwarding of "plain old vanilla" traffic
  - Generic Forwarding Function that could be implemented in hardware
- Reason formally about composition and resulting global behavior
  - Establish correctness of the underlying fixed functionality
  - Identify sufficient conditions for user-supplied code to preserve that correctness

S. Bhattacharjee, K. L. Calvert, E. W. Zegura, "Reasoning About Active Network Protocols", Proceedings of 1998 IEEE International Conference on Network Protocols (ICNP '98), Austin, Texas, October 14–16, 1998, pp. 31–40.

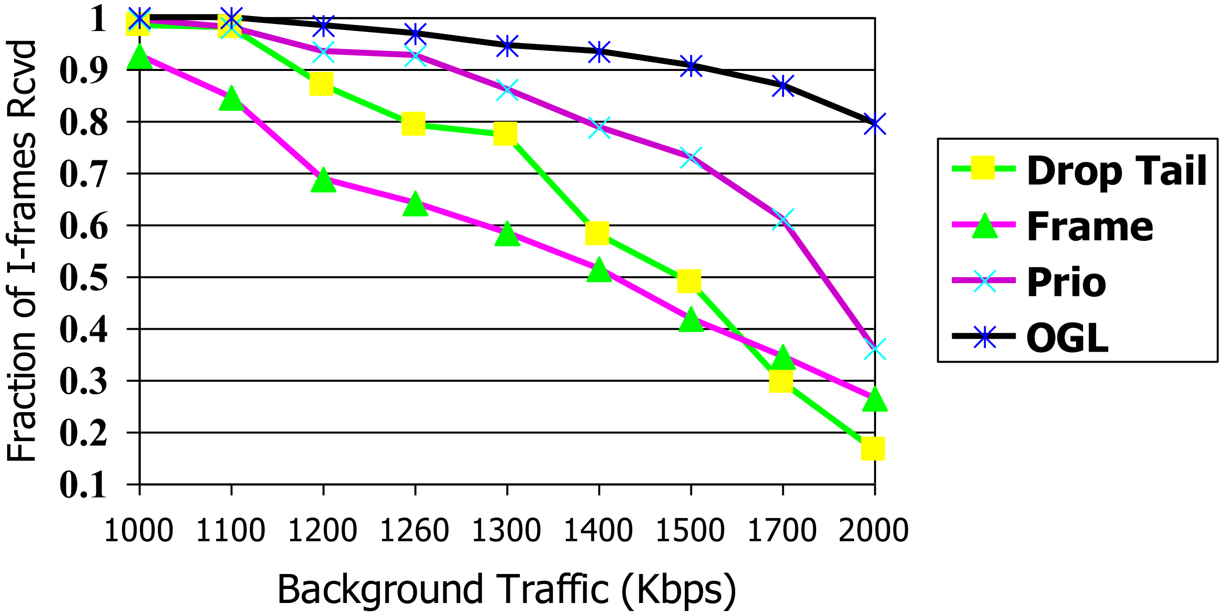
CANEs: Packet-Processing Model



CANEs: Generic Forwarding Function

parse packet to obtain source (S), destination list (D),
   forwarding table identifier (R), selection function (M)
check authorization for R and M
<Slot 0: null>
for each d in D: add Lookup(S,d,R,M) to interface list L;
<Slot 1: null>
if (L is empty) then
      <Slot 2: construct and send notification packet to S>
else for each interface i in L:
      if (i is congested) then <Slot 3: discard>
      else <Slot 4: null>
           enqueue packet for i

Application: Intelligent Discard for MPEG

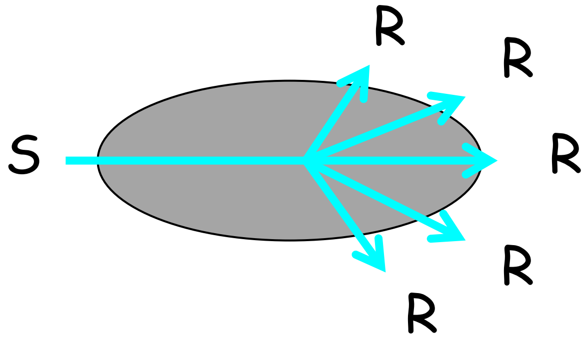


One active router, bottleneck 2Mbps,
MPEG source averages 725 Kbps

2. Concast Service*

Existing Service [at that time]: IP Multicast

- Scalable group communication through abstraction:
  - Single address represents an arbitrary number of receivers
- Network duplicates and delivers message to each receiver

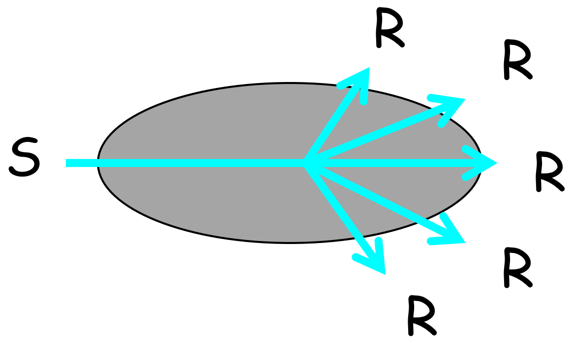


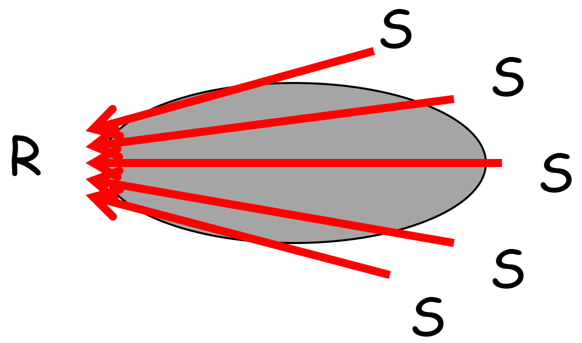
- Benefits both sender and network:
  - One send operation results in many messages received
  - Reduced bandwidth requirements

K. L. Calvert, J. N. Griffioen, B. Mullins, A. Sehgal, and S. Wen, "Concast: Design and Implementation of an Active Network Service", IEEE Journal on Selected Areas in Communications, 19(3), March 2001, pp. 426-437.

Concast: Motivation

- Problem: Many multicast applications need feedback:
  Acknowledgements, retransmission requests, performance statistics (loss rates, delay information)

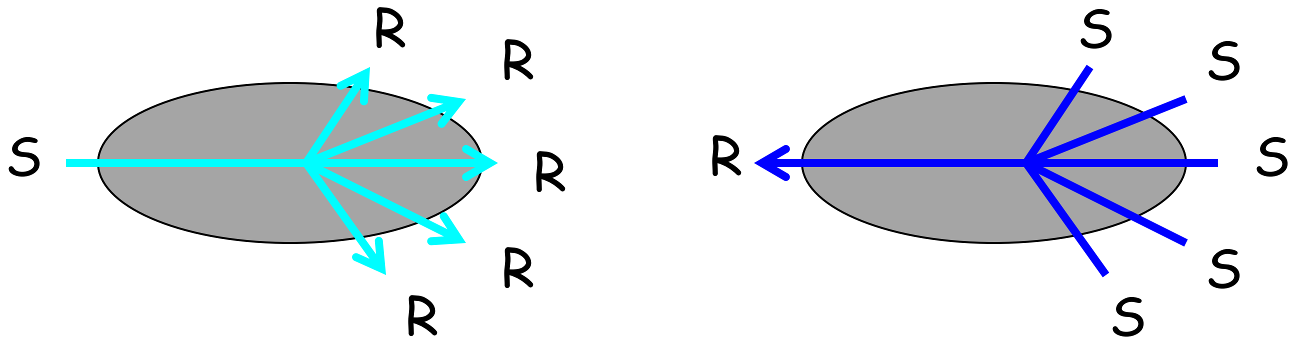




- Feedback must be unicast to the source
  - Sender deals with individual receivers, breaking abstraction
  - Implosion problem limits scalability

Solution: Concast Service

- Principle: scalability through abstraction
  - Single address represents an arbitrary number of senders.



  - Network merges messages from the group
    - According to application specification, defined by a standard 4-function API
  - Multiple sends result in at most one message delivery
- Benefits both receiver and network
  - Reduced bandwidth requirements

Ways to Use Concast

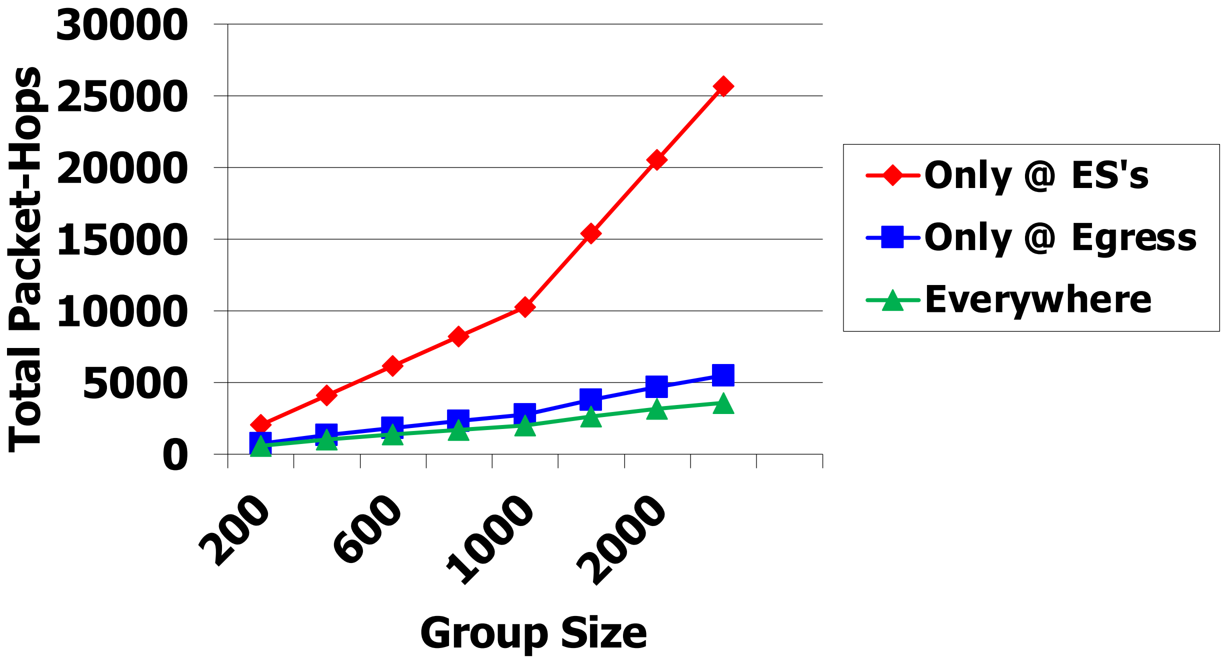
- Application-specific merging
  - Filtering, aggregating telemetry
  - Merging media streams (demonstrated: audio, video)
- Application-independent merging protocols
  - Collecting maximum (or any associative, commutative operator) of group members' sent values
    - E.g. reliable multicast feedback
- Protocol-independent generic services
  - Duplicate suppression (based on hash of IP payload)
  - Aggregation of small packets (TCP acks) [ICNP 2000]

Concast: Partial Deployment Effectiveness

Number of Packet-Hops To Collect a Value

From Every Group Member

4900-node Transit-stub Graphs



3. Ephemeral State Processing

Goal: Simple building-block service

- Packets control simple computations on small amounts of state
  - Middle ground between IETF's problem-specific and active networking approaches
- Alongside, not instead of, IP forwarding

Requirements: IP-like approach

- Bounded resource requirements
- Simple, limited service
  - Responsibility for end-to-end service remains at the edge
- Flexible: Useful for solving multiple real problems
- Deployable without forklift upgrade

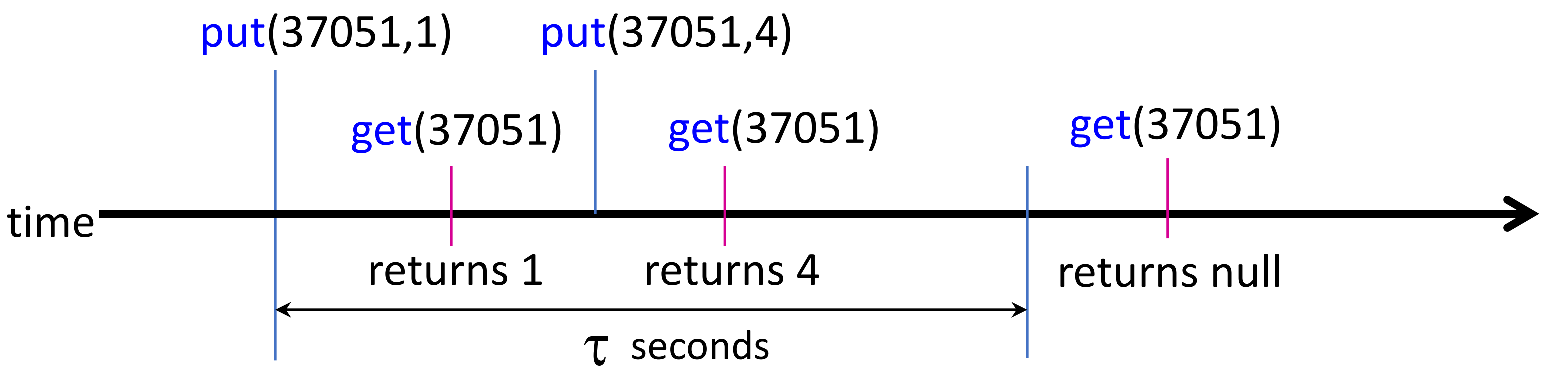
K. L. Calvert, J. N. Griffioen, Wen Su, “Lightweight Network Support for Scalable End-to-End Services”, Proceedings ACM SIGCOMM 2002, Pittsburgh, August 2002, pp. 265–278.

Ephemeral State Processing Components

- Ephemeral State Store (ESS)
  - Time-bounded associative memory = set of (64-bit tag, 64-bit value) pairs
    - One ESS per interface
- Packet-borne "instructions" (one per packet)
  - Each instruction defines a fixed-length computation
    - Operands: values in ESS, packet fields, node-specific values
    - Comparable to machine instructions of general-purpose computer
  - On termination, forward or discard packet
  - Routers support a common instruction set
- Wire protocol
  - ESP instructions carried in payload or shim header
  - Packets recognized and executed hop-by-hop
- End-to-End services
  - Construct by sequencing packets in time and space

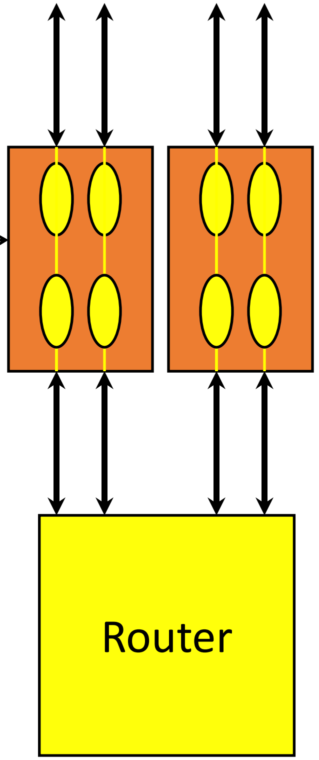
Ephemeral State Store

- Set of pairs (t,v)
  - Local to each router (interface)
  - Pairs persist for τ seconds, then disappear
- Access functions
  - put(t, v): establishes "the set contains (tag, value)"
  - get(t): if ∃ v such that (t,v) is in the set, returns v
    else returns null



Network Processor Implementation

- Per-port ESP facility
  - Transparent to router
- Intel “BridalVeil” IXP1200 board
  - Intercepts packets as they enter/exit the router



Estimated Throughput (Kpps)

Instruction	SRAM	DRAM	SRAM+DRAM
count()	340	259	263
compare()	232	146	188

Uses for ESP

- Controlling packet flow
  - Make drop/forward decisions based on state at node or interface
    Example: Duplicate suppression
- Identifying interior nodes with specific properties
  - Reveal just enough of topology to find what is needed
    Example: Finding multicast branch points
- Processing user data (a la Concast)
  - Simple hierarchical computations scale better
    Example: Aggregating feedback from a multicast group

Other innovations you likely haven't used lately...

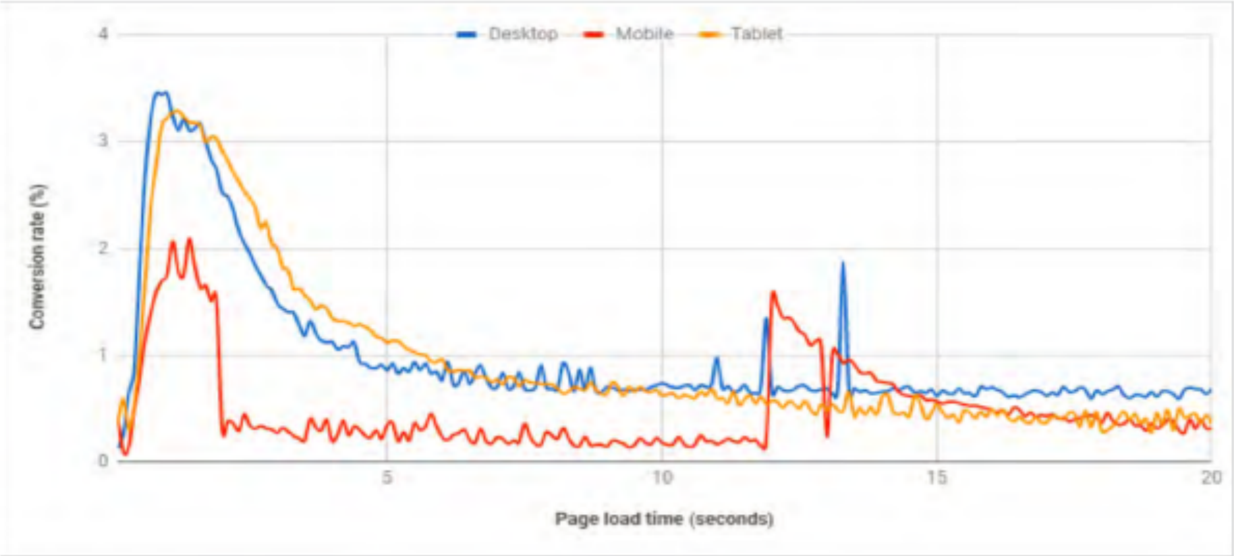
- Postmodern Internet Architecture
- Integrated Services/RSVP
- Global Environment for Network Innovation (GENI)
- Information-Centric Networking
- Future Internet Architectures (NSF Program)
- Secure BGP (Route origin attestations)
- To be determined:
  - IP Multicast
  - SCION

Agenda

I. BLUF + Background ✓
II. Some network-layer innovations you've never used ✓
III. Some network-layer innovations you likely have used
IV. Success factors for innovation at the network layer
V. Possible way forward

Content Delivery Networks

- Motivation:
  - Reduce page load times for web commerce sites
  - Place content servers near "eyeballs"
- Today: many CDN services
  - Traffic analytics
  - Automatic A/B testing
  - DDoS protection
- Akamai:
  - 240K Servers
  - 1700+ networks
  - 3300 locations
  - 80+ Tbps traffic served/day
  - 3 x 10^{12} HTTP requests served/day



Impact of page load times on conversion rates for desktop, mobile, and tablet.

Source: Akamai, State of Online Retail Performance, 2017 Holiday Perspective

Source: Bruce Maggs' Keynote "Economics of Content Delivery" at ACM ICN 2020

Software-Defined Networking (SDN)

Problem: Network innovation controlled by router vendors.

Consequence: Glacially slow innovation of the Internet.

Solution: Network owners and operators write the software to control their networks.

- Nick McKeown, "Software-Defined Networking: How it has transformed networking, and what happens next", 2018

- Idea: Open up router architecture
  - What Industry Standard Architecture did for personal computers 30 years ago
- Motivation:
  - Cost savings!
    - Commodity hardware, open-source software
  - Reduced complexity
  - Faster evolution
  - Network telemetry
- Evolution
  - Programmable control plane (OpenFlow, P4 etc.)
  - Programmable data plane

Agenda

I. BLUF + Background ✓

II. Some network-layer innovations you've never used ✓

III. Some network-layer innovations you likely have used ✓

IV. Success factors for innovation at the network layer

V. Possible way forward

Success Factors

1. Available technology
   – CANEs: way ahead of its time
   – Concast: maybe with today's NFV platforms
   – ESP: Intel IXP 1200/2400 – low-speed links
   + CDN: off-the-shelf servers, network equipment
   + SDN: "merchant silicon", Protocol Independent Switch Architecture

Success Factors

2. "Thrust" = Stakeholders pushing for adoption
   - CANES: advisors & PhD students
   - Concast: presented to Reliable Multicast Research Group in IRTF
     - Then the web killed multicast
   - ESP: academics
   - CDN: Startups!
   - SDN: Startups!

Success Factors

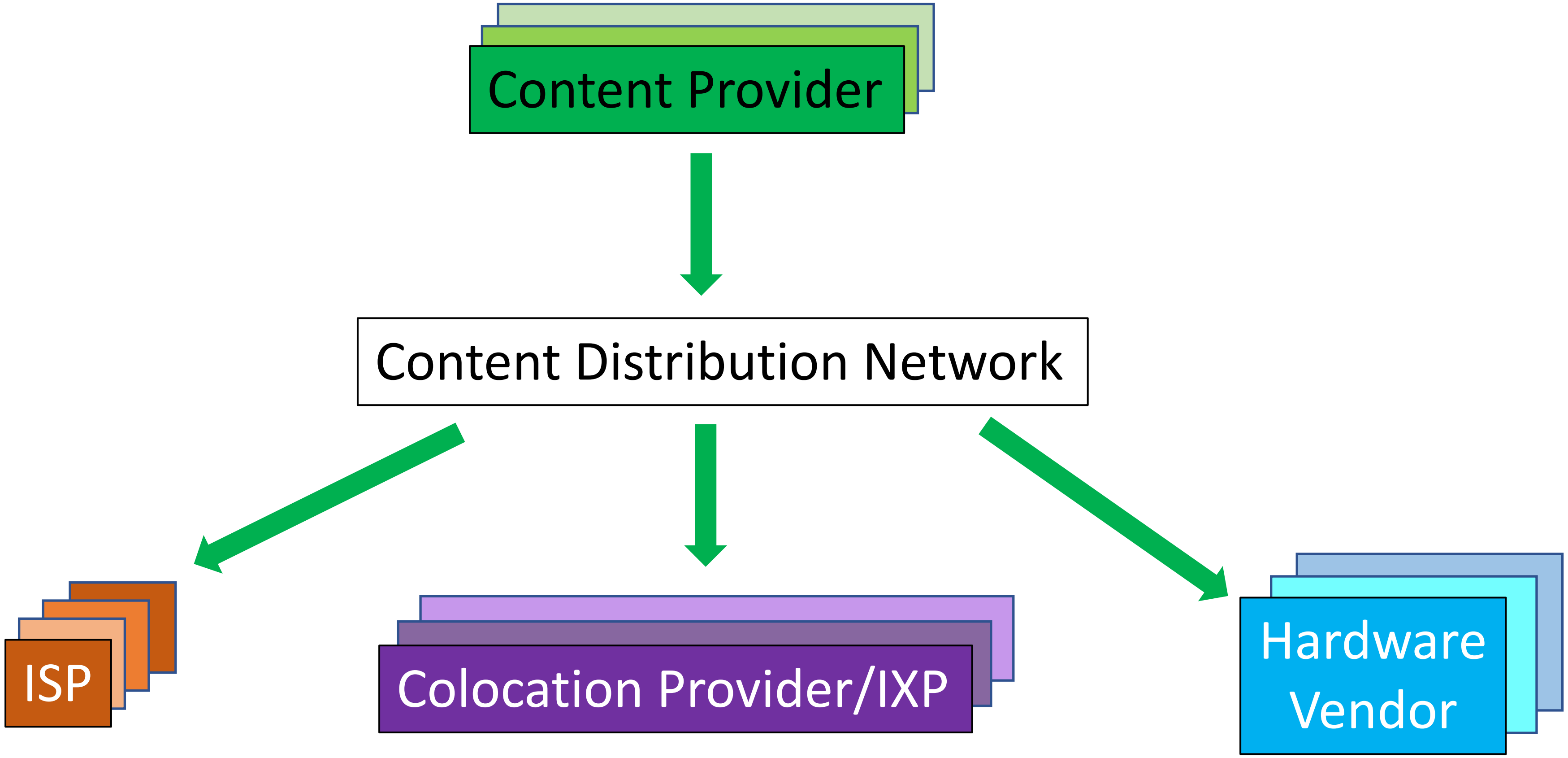
3. Motivation for deployment

   Benefits must be clear, quantifiable, and

   accrue to the one deploying/paying – directly or indirectly.

   – CANEs: How to charge for customized processing? Access control?
   – Concast: Secure signaling protocol for access control
   – ESP: What benefit to the deploying ISP?
   + CDN:
     - Well-known relationship between speed and "conversion rate"
   + SDN: reduced capex (+ reduced opex?)
     - Intra-domain deployment: Operator pays and benefits

CDN Money Flow



Thanks: Bruce Maggs' Keynote "Economics of Content Delivery" ACM ICN 2020

The Importance of Routing & Forwarding Money

A conversation with a Well-Known Economist, recounted by Dave Clark:

WKE:

"The Internet is about routing money. Routing packets is a side effect. You guys screwed up the money routing protocols."

DDC:

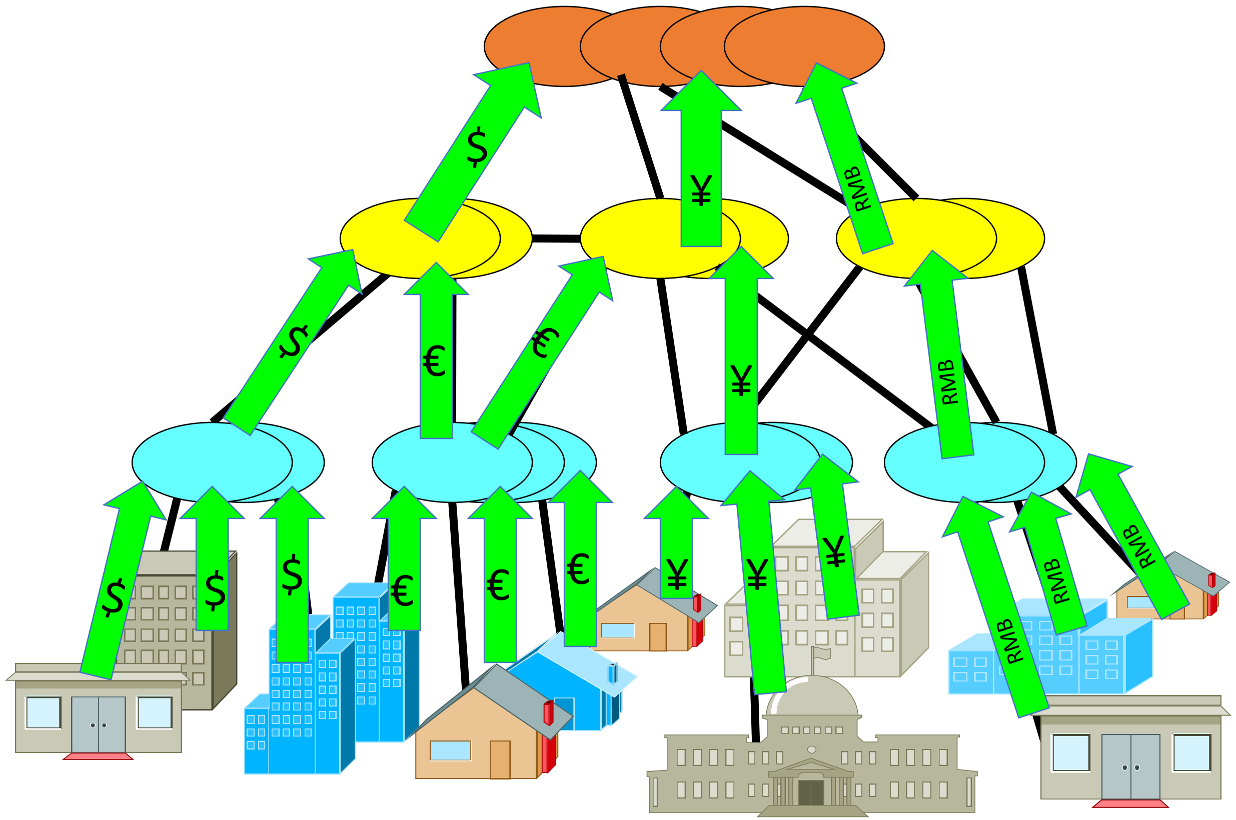
"I did not design any money-routing protocols!"

WKE:

"That's what I said."

– From *Designing an Internet*, David D. Clark, MIT Press, 2018, p. 246

Background: Money Flow in Today's Internet



- Money only enters system at the edge
- Payments flow up the provider hierarchy (only)
- Packets follow money

⇒ Packet flow constrained

⇒ Little competition

⇒ Little innovation visible at the edge

The Crux of the Problem
(In My Humble Opinion)

- ISPs have virtually no incentive to support new E2E services.
  - No advantage to deploying if others do not.
  - No way to choose AS-level routes based on capabilities if some do.
  - No way to recoup costs if they do not directly serve the user.
- As a consequence, E2E innnovation is confined to:
  - Inside domains – data centers, wireless edge
  - New ecosystems (CDNs)
- Needed: Interdomain money-routing/forwarding mechanisms.

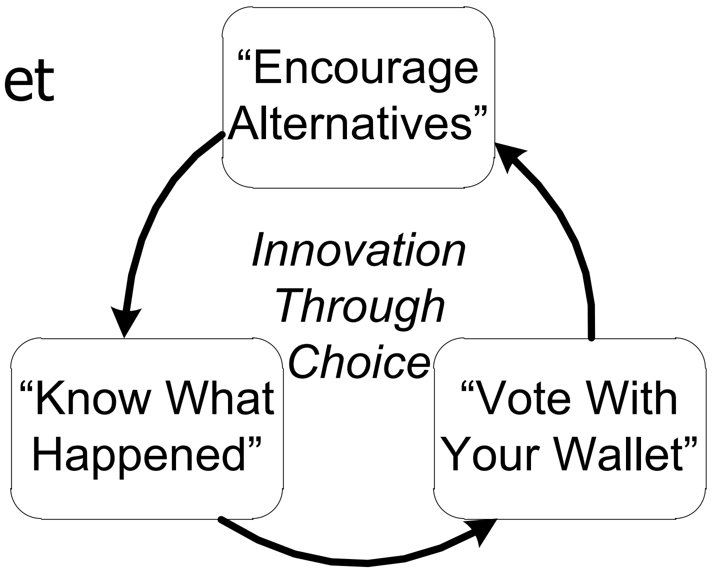
Agenda

I. BLUF + Background ✓
II. Some network-layer innovations you've never used ✓
III. Some network-layer innovations you likely have used ✓
IV. Success factors for innovation at the network layer ✓
V. Possible way forward

ChoiceNet Project

NSF 1111040, 1111088, 1111256, 1111276.

- Motivation
  - End-users/applications have limited choices in the Internet
    - Local network provider(s) only (maybe)
    - No control beyond 1st hop
  - Providers have little incentive to deploy new services
- Goal: build an "economy plane" for Internet
  - Marketplace, where
    - Customers can choose providers
    - Providers have access to all customers
  - Fine-grained contracts for network services
- Hypothesis: Competition between providers will drive innovation



T. Wolf, J. Griffioen, K. Calvert, R. Dutta, G. Rouskas, I. Baldin, and A. Nagurney, "ChoiceNet: Toward an Economy Plane for the Internet", *Computer Communication Review*, 44(3), July 2014, pp. 87–96.

Possible Driver Application: Spot Market for Interdomain Connectivity

- Transit traffic load is more or less predictable.
  - Transit providers have unused capacity available at times.
- Xu and Li: Under certain conditions, transit providers could make additional profit by selling short-term access to their spare capacity*.
  - Customer (edge ISP) motivation: cheaper transit for elastic traffic
- Spot market idea: sell spare capacity cheap

*H. Xu and B. Li, “Spot Transit: Cheaper Internet Transit for Elastic Traffic”, *IEEE Transactions on Service Computing,* vol. 8, issue 5, Sept-Oct 2015, pp. 768–781.

Implementing a Transit Spot Market

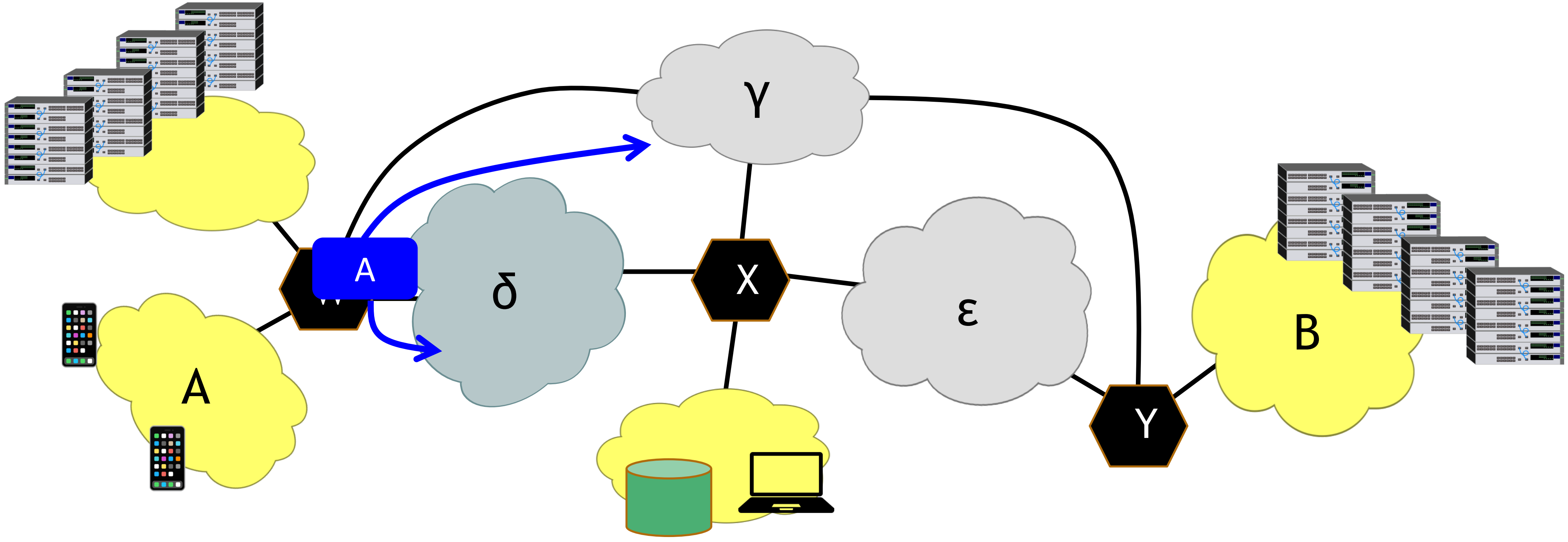
- Transit Providers need to:
  - Determine available capacity, set price
  - Distinguish revenue-generating traffic from other traffic
    - SDN
- Access Providers need to:
  - Determine elastic demand
  - Collect, select offers (including AS-level routing)
- All need:
  - Means of propagating and matching offers
  - Standard interfaces and low transaction costs

K. L. Calvert, J. Griffioen, A. Nagurney and T. Wolf, "A Vision for a Spot Market for Interdomain Connectivity", 39th IEEE ICDCS (Blue Sky Track), July 2019, Dallas, TX.

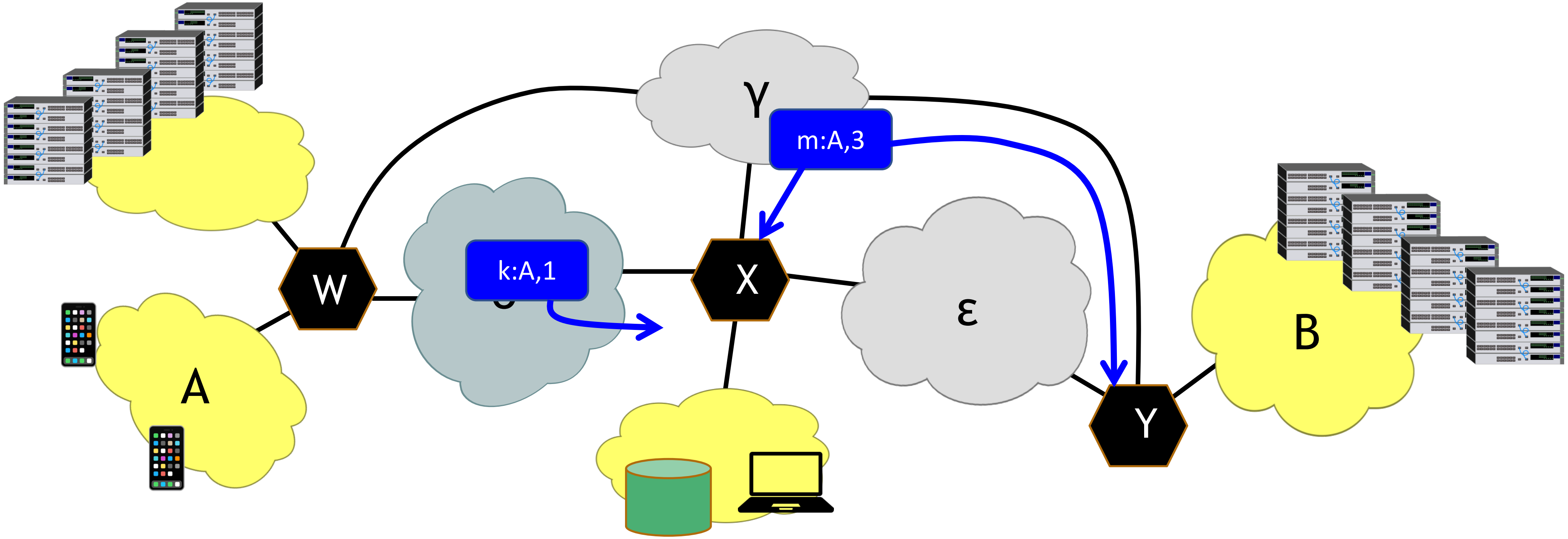
Economic Software-Defined Exchanges (ESDXs)

- Idea: Interdomain eXchange Point (IXP) as neutral third party in provider ecosystem.
  - SDN allows rapid/automated connectivity setup/teardown.
  - Providers interface with ESDXs instead of each other.
  - Standardized interface lowers transaction costs.
- ESDX functions for spot market:
  - Collect and propagate transit connectivity offers
  - Collect access provider acceptance of offers
  - Instantiate forwarding state dynamically as offers are consummated

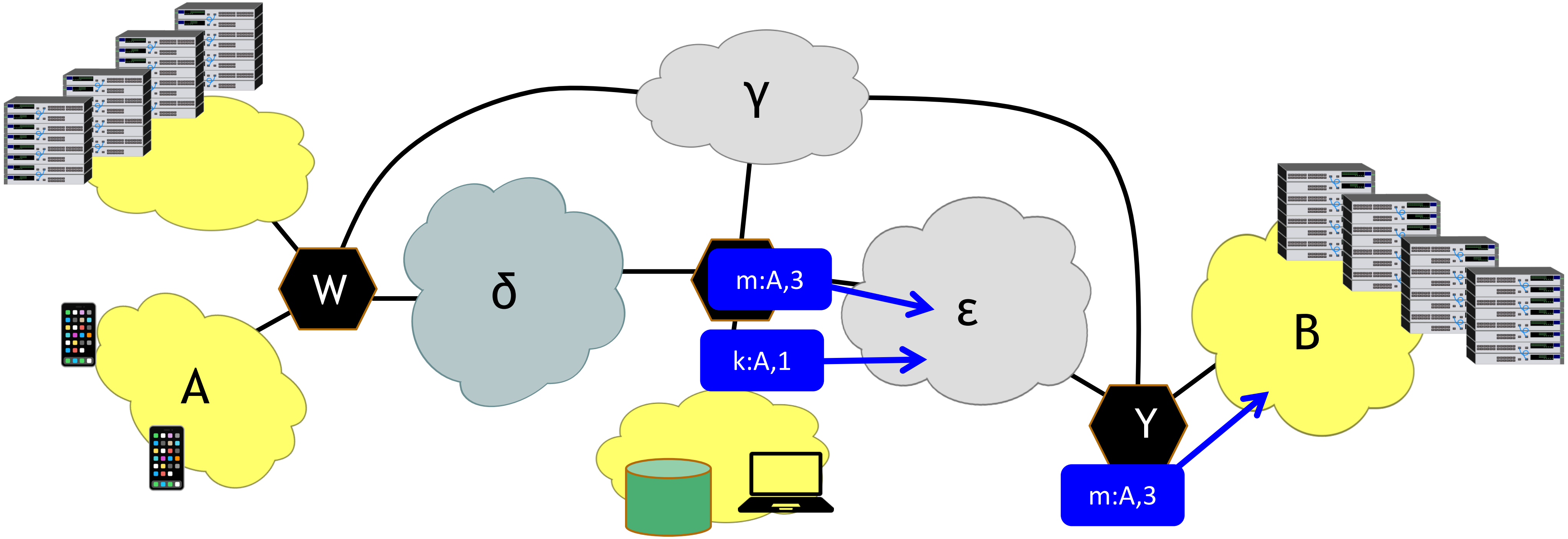
Propagating Offers: BGP-like Approach



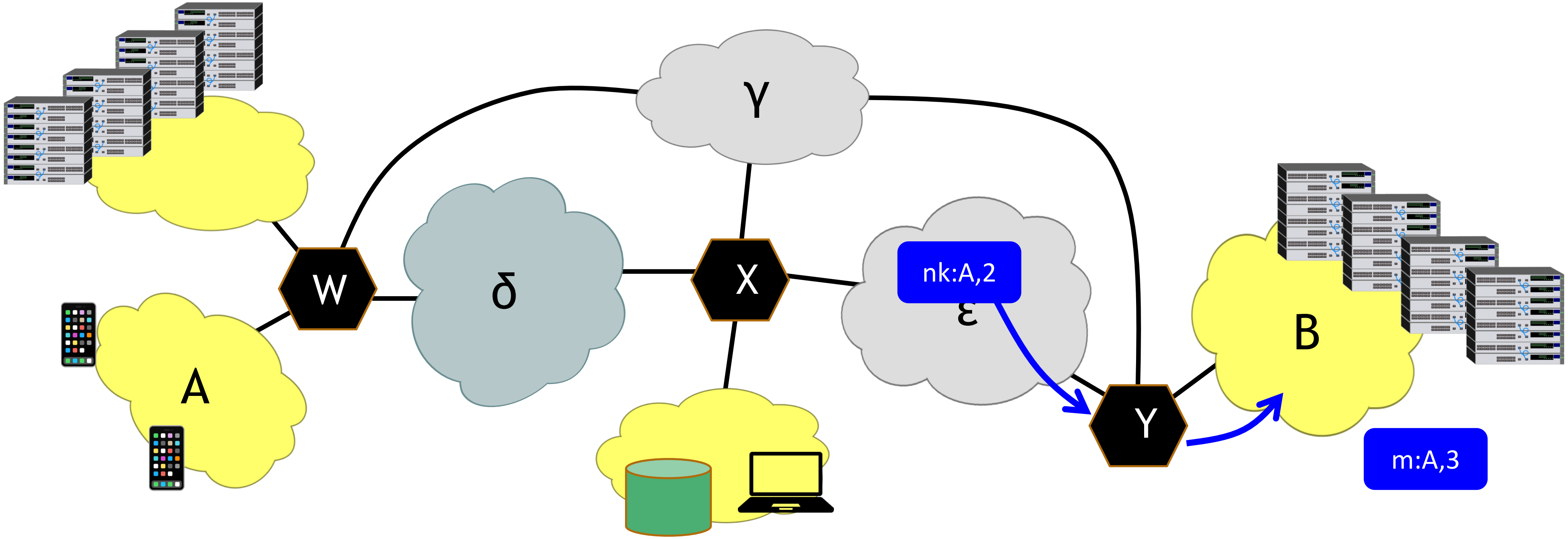
Propagating Offers: BGP-like Approach



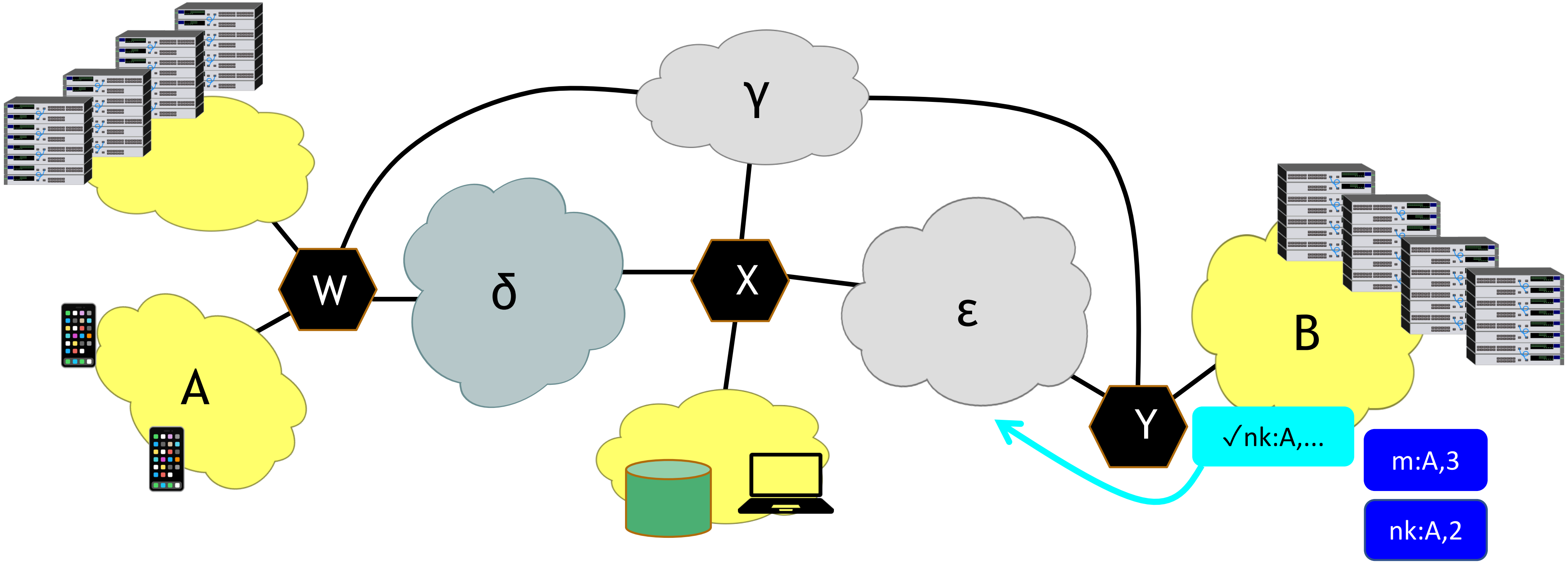
Propagating Offers: BGP-like Approach



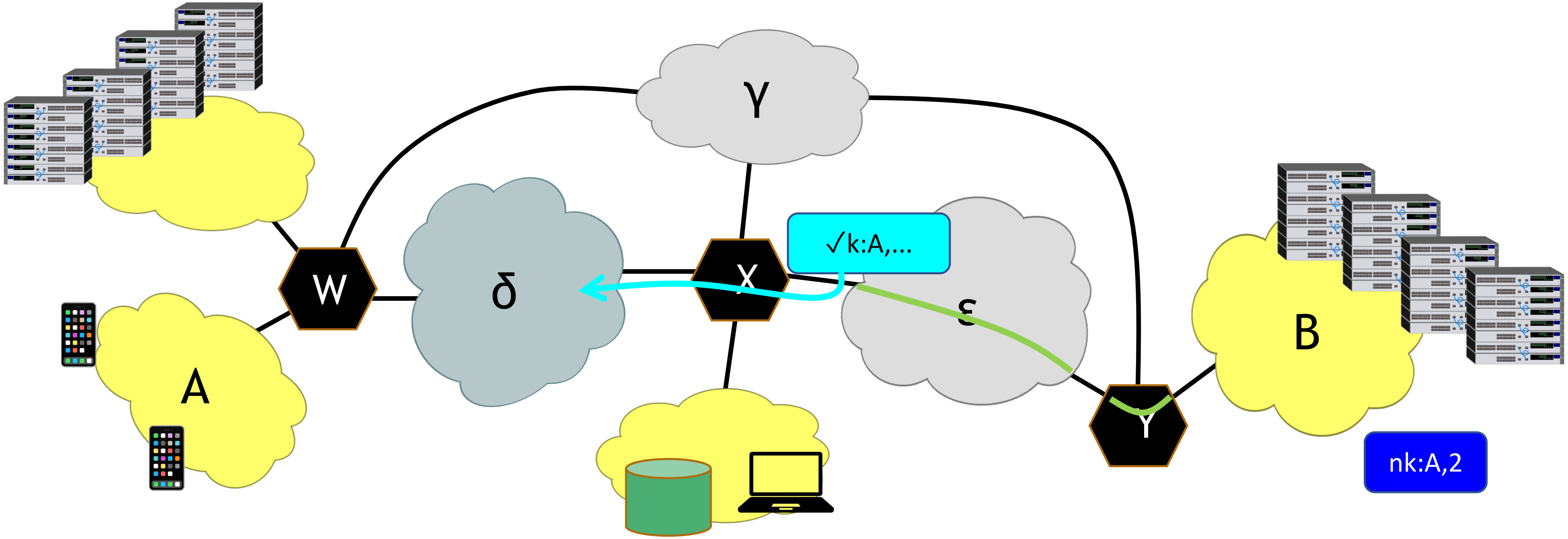
Propagating Offers: BGP-like Approach



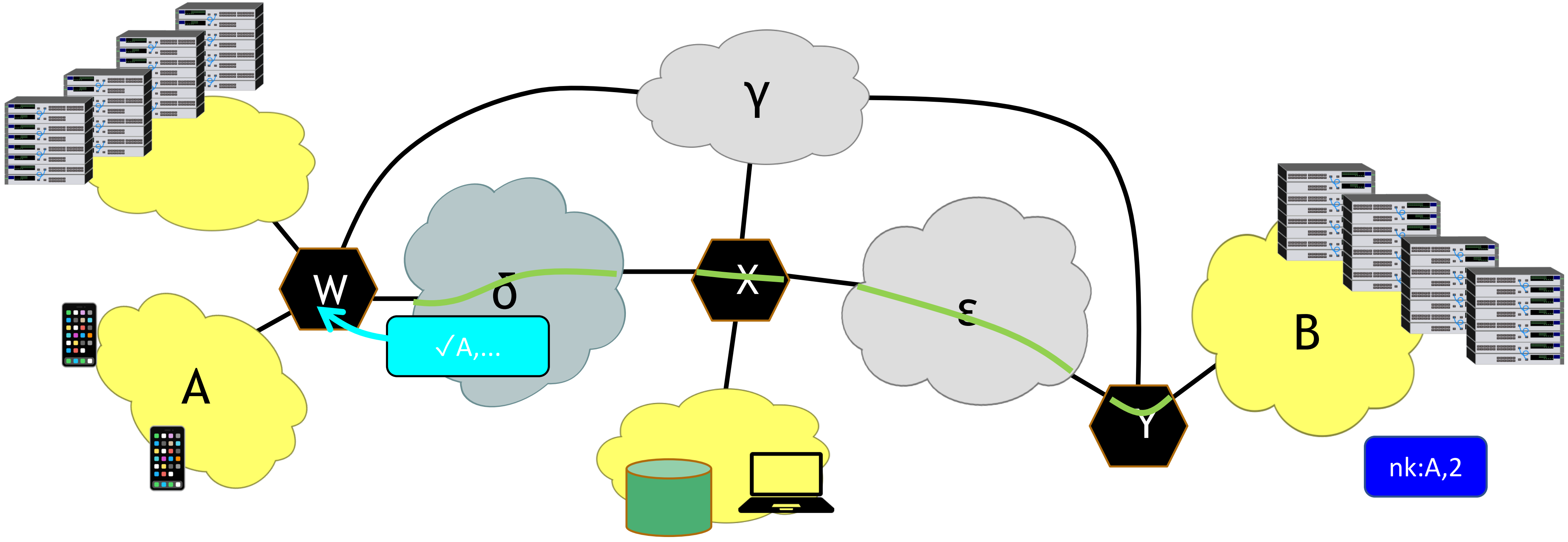
Propagating Offers: BGP-like Approach



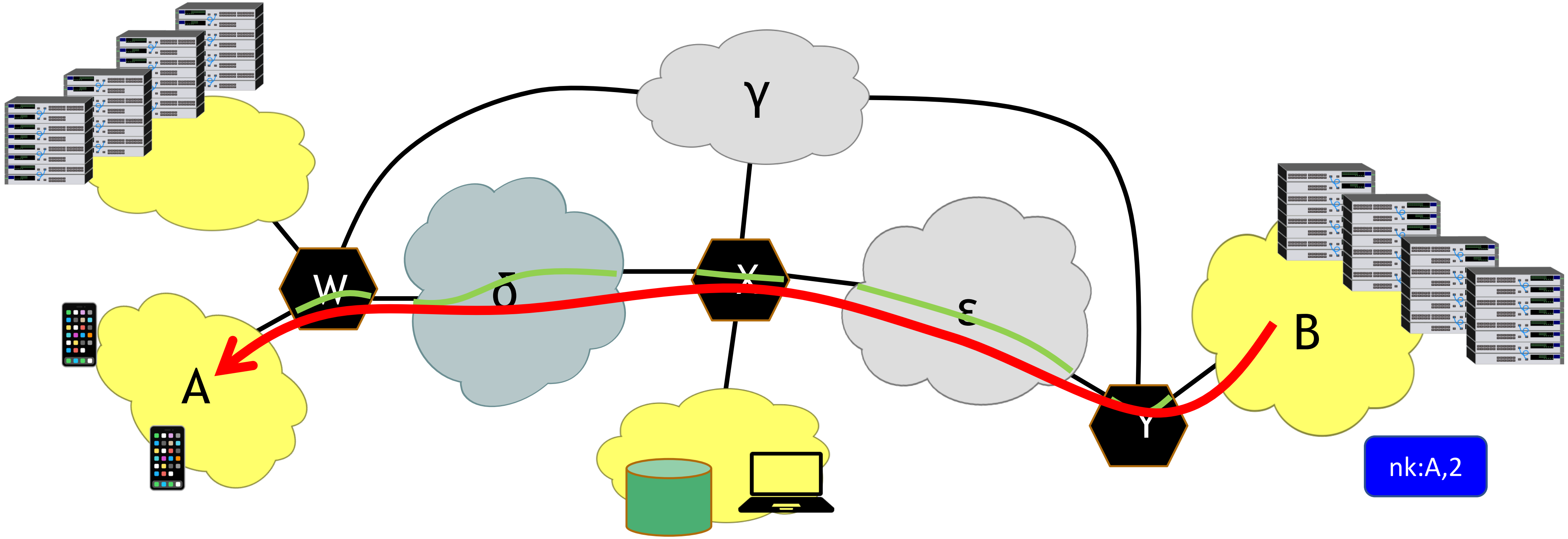
Propagating Offers: BGP-like Approach



Propagating Offers: BGP-like Approach



Propagating Offers: BGP-like Approach

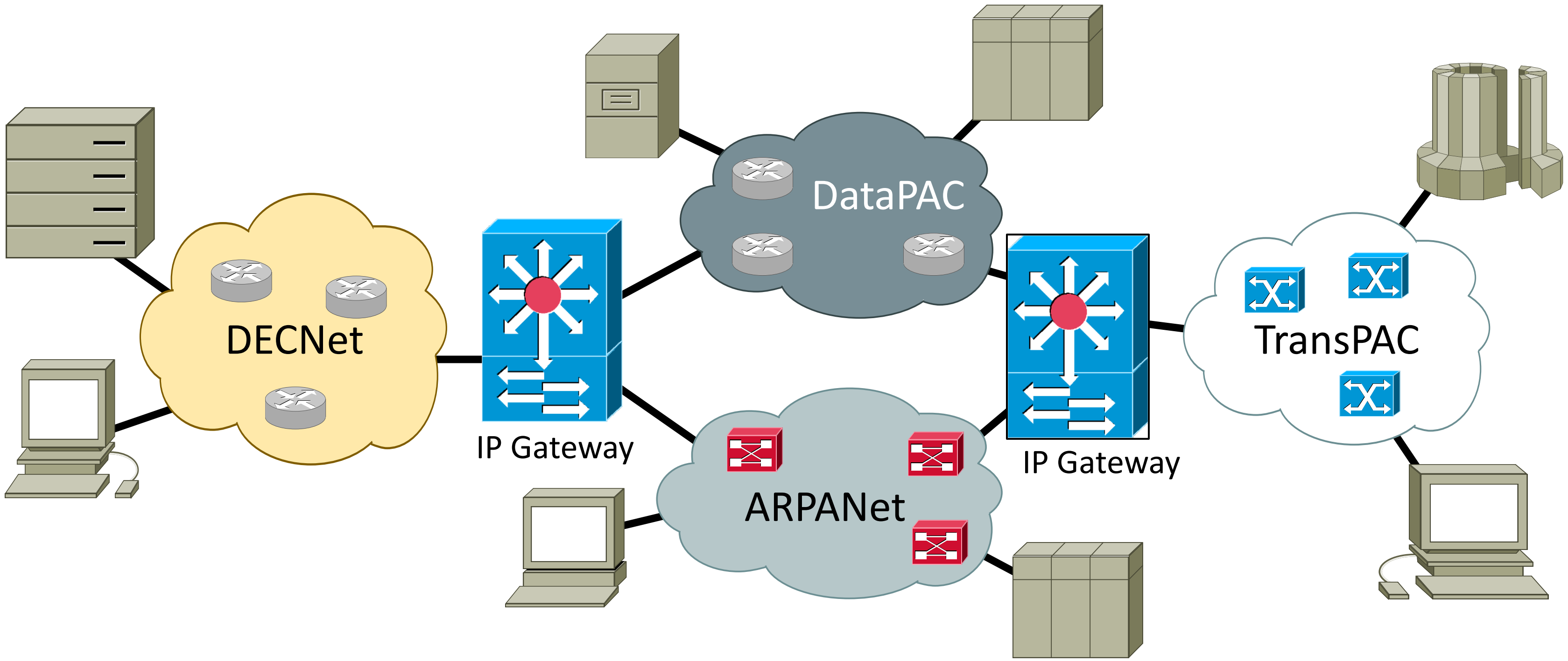


Summary

- Lack of money-routing protocols impedes innovation in the inter-domain space.
  - Just one among several factors
- Standard economic interfaces among provider networks are needed.
  - Economic software-defined exchanges (ESDXs) as possible base
  - Analog of the "shipping container" in multi-modal freight transport
- Such a clearinghouse could be used for various other services.
- Proposition: Inter- ➛ Intra-domain evolution is easier than the reverse.

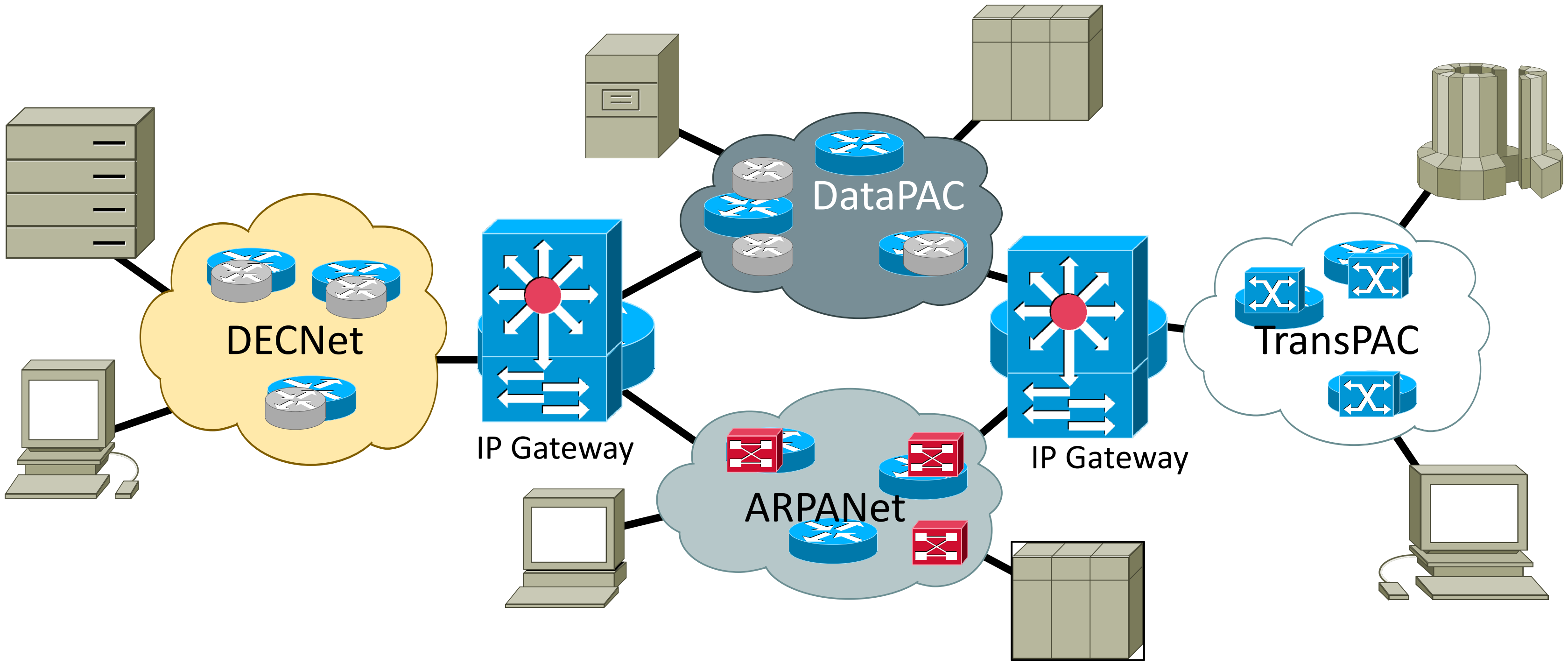
In the Beginning…

Internet Protocol started out as an overlay – designed to connect all kinds of other networks – both LANs and WANs.



As the Technology Matured...

...most of those underlying technologies faded away as IP routers were deployed within domains. IP became "the network layer" (vs. the internet layer).



Thanks!
Questions?