What happened to IPMulticast?
SANOG 7

Greg Shepherd

shep@cisco.com

Agenda

- **What happened to IPMulticast?**
  - *Is it still alive?*
  - *Where?*
  - *Who's using it and why?*
  - *What are their biggest challenges today?*
  - *Why has it not been deployed as rapidly as we all hoped..or Why don't I have it at home?*
  - *Is the dream of global multicast deployment still alive?*
  - *If so, what are the current barriers to deployment today?*
  - *Does it have a future?*

What happened to IPMulticast?

- **A question with a past…**
  - *Steve Deering*
  - *Making a routed network look like a LAN*
  - *Global MBONE built with DVMRP tunnels*
    - *Hub-and-spoke*
  - *Early applications were conferencing tools*
    - *Vic, Vat, WB, SDR*
  - *Some early "visionaries" pinned-up radio stations*
  - *Source discovery was expected of the network*

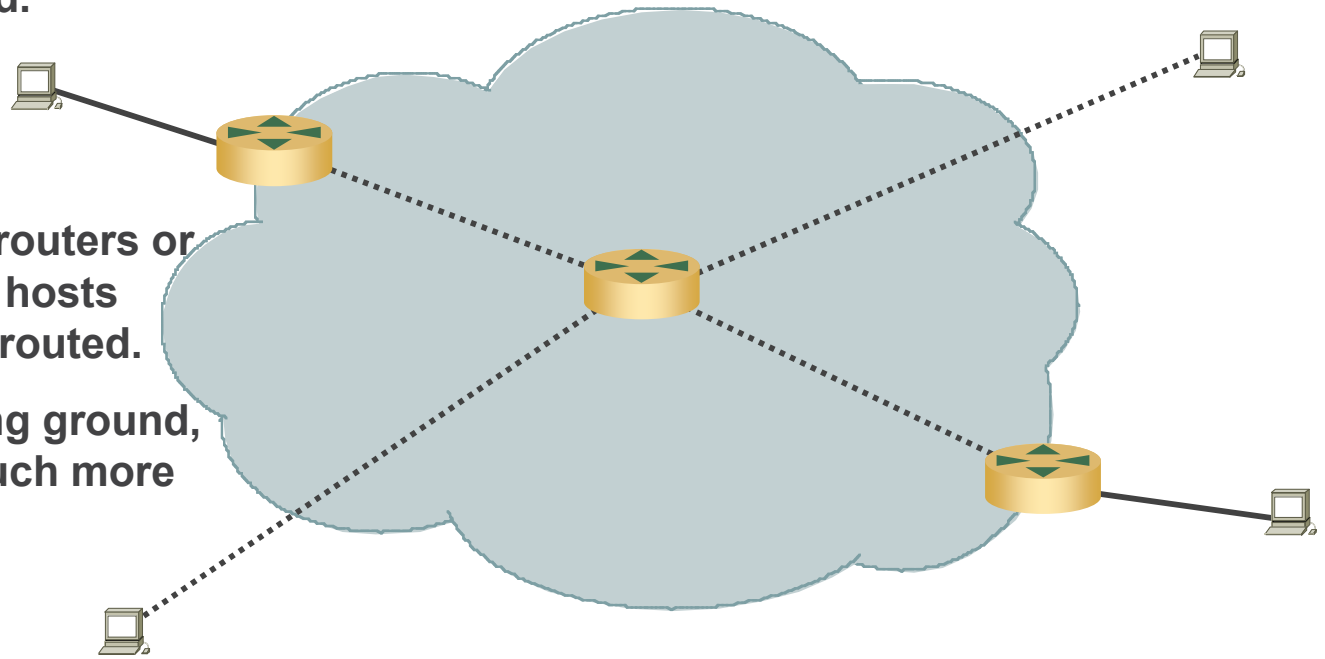
MBONE evolution…

- **Central DVMRP global architecture**
  - *MBONE – a flat world*
- **MBGP/PIM transit – preMSDP**
- **MSDP/MBGP/PIM-SM**

MBONE - the world is flat

DVMRP tunnels to a central router.

A flat world.



From site routers or from local hosts running mrouted.

Nice testing ground, but not much more than a toy.

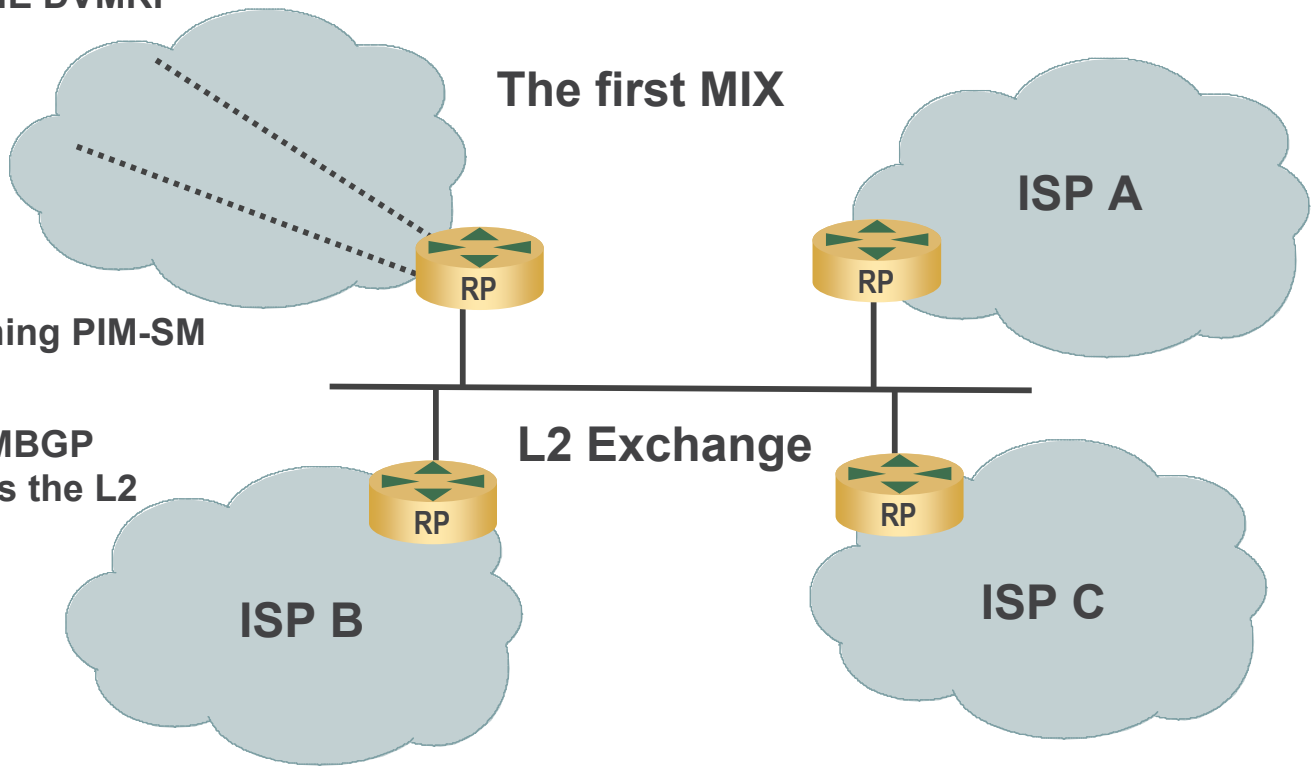
"Hey, this is like running RIP across the Internet!" – yep!

NoBone
MBGP/PIM transit – preMSDP

Legacy MBONE DVMRP tunnels

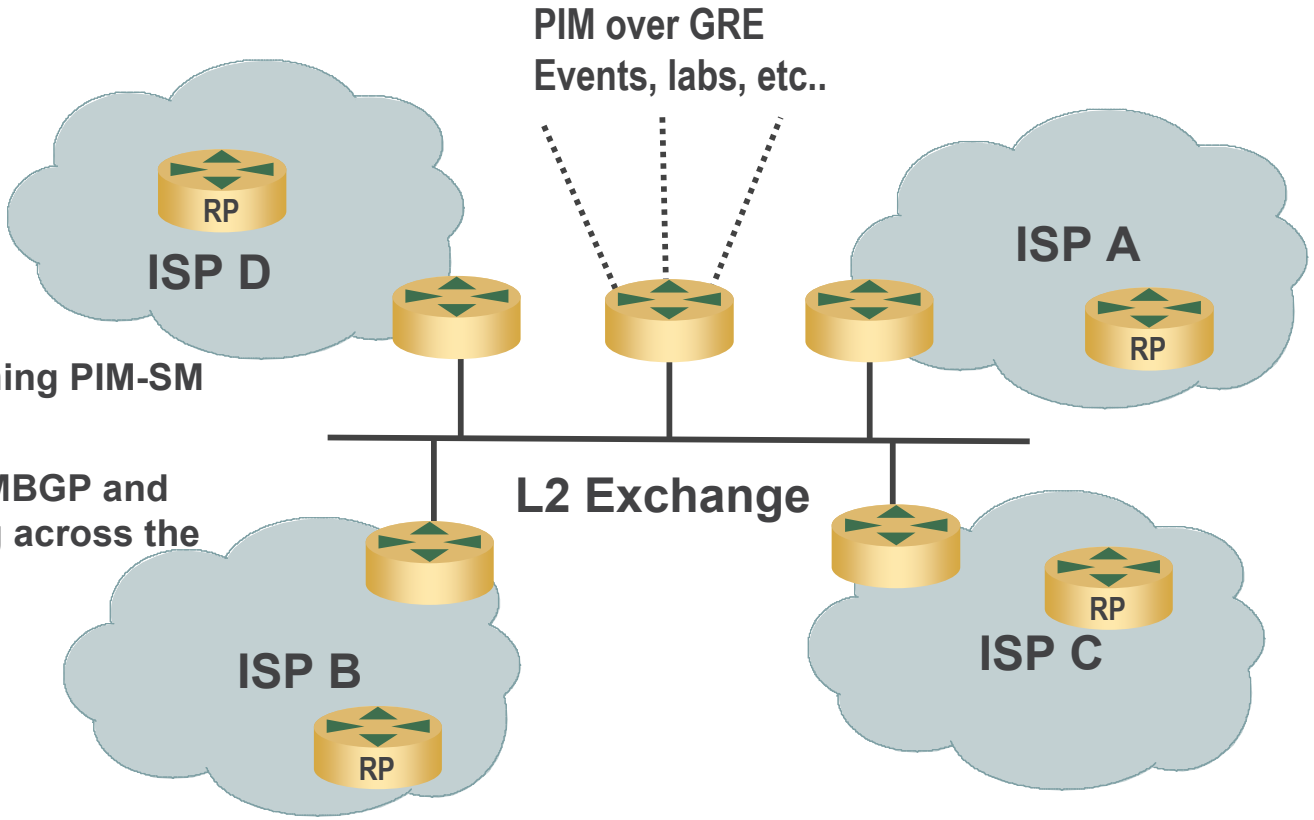
Each ISP running PIM-SM internally.

Full mesh of MBGP peering across the L2 exchange.



BUT, RPs MUST be on the exchange, and PIM-DM run on the MIX interfaces to flood S,G info to all other RPs

NoBone
MSDP/MBGP/PIM-SM – Today.. Well 1999/2000



Each ISP running PIM-SM internally.

Full mesh of MBGP and MSDP peering across the L2 exchange.

Each ISP PIM domain learns of external sources via MSDP.

RPs can be anywhere inside each PIM domain.

NoBone
MSDP/MBGP/PIM-SM – Today.. Well 1999/2000

We're done!

WRONG

What's Wrong?

- **Multicast in the Internet is an all-or-nothing solution**
  - *EVERY ROUTER on EARTH needs to be mcast enabled. *sigh**
- **Even Mcast-aware content owners resorted to unicast streams to gain audience size**
  - *Broadcast.com - early visionaries, too early.*
- **Unicast doesn't scale**
  - *Splitters/Caches just distribute the problem*
    - *Still has a cost-per-user*
  - *As receiver BW increases, problem gets worse.*
  - *Creates a non-functional business model*
    - *Maybe we need a new business model*

Meanwhile…

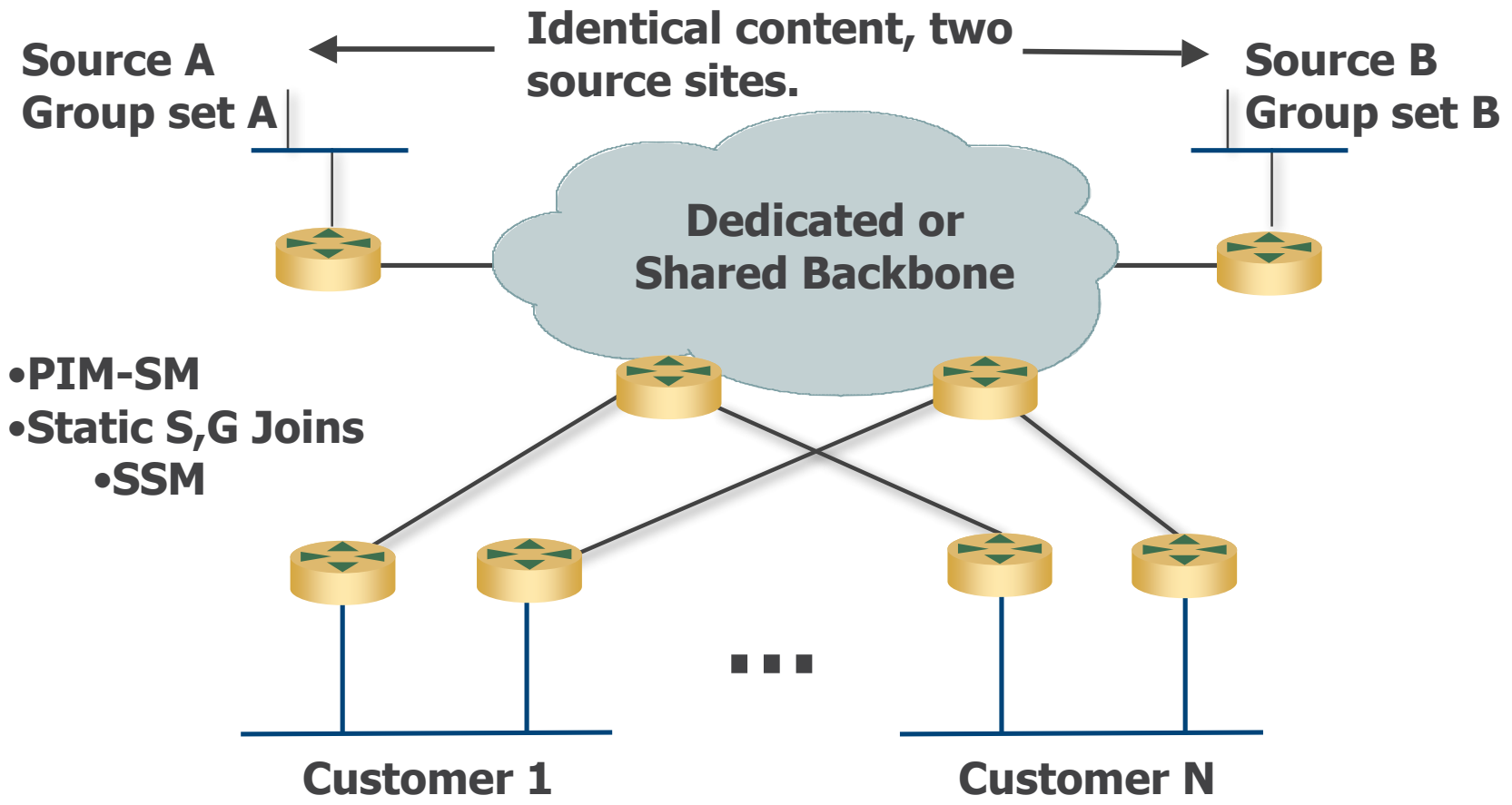
..while the Internet Multicast Peacenics were holding hands at a drum circle dreaming of global multicast deployment, working late-nights helping engineers around the world turn-on PIM and turn-up multicast transit and peering, and listening to classical music from Oregon with Vat…

Multicast was being deployed! Islands of deployment

- **Financial Networks**
  - *Security Exchanges*
    - *NASDAQ, NYSE, AMSE, HKSE, etc..*
  - *Securities Trading Enterprises*
- **Enterprises**
- **Service Providers**
  - *MVPNs*
  - *Walled-Gardens*
    - *DSL, Cable*

How?
Security Exchanges

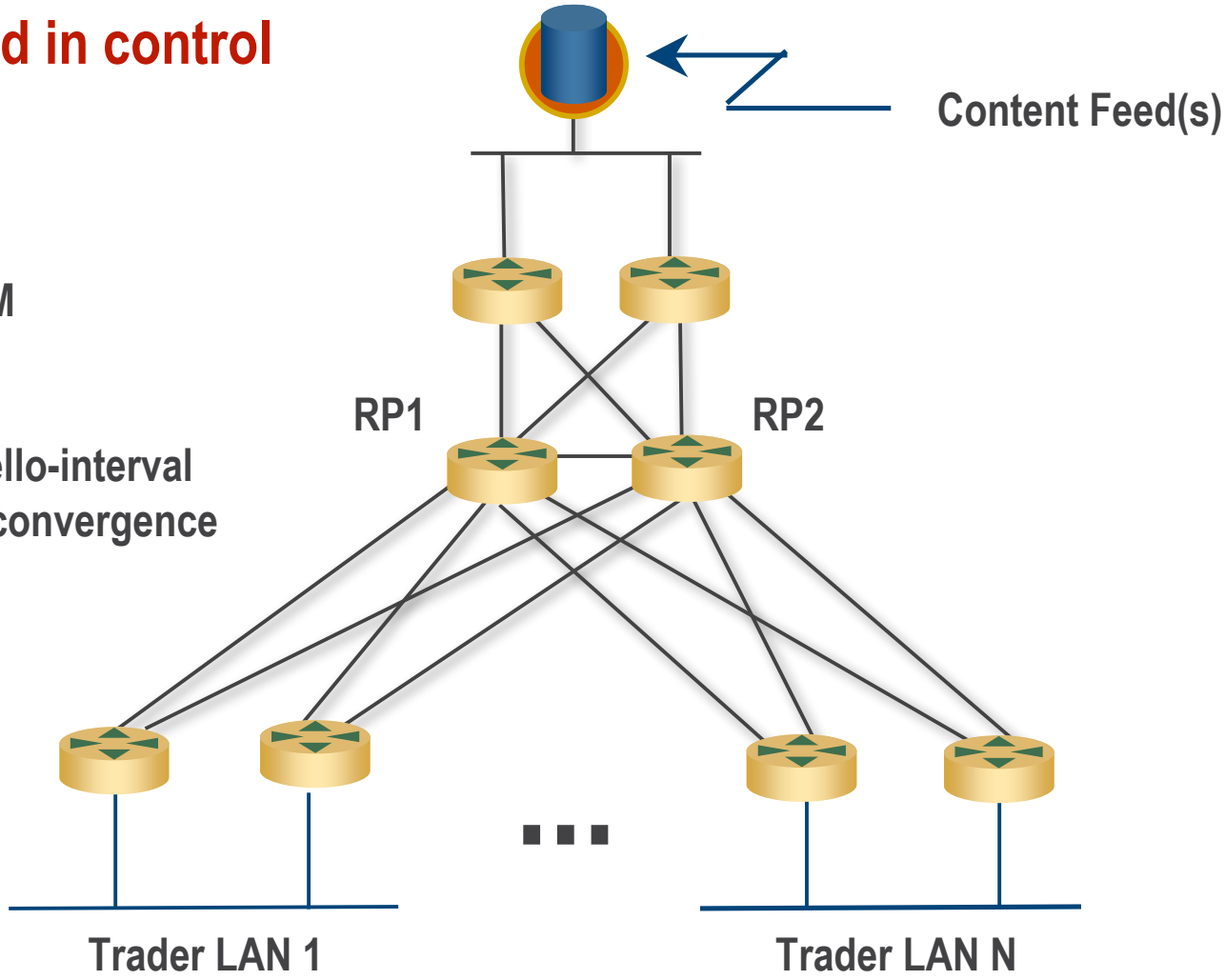
End-to-end in control (mostly…)



How?
Security Trading Enterprises

End-to-end in control

- PIM-SM
  - BiDir-PIM
- Anycast-RP
  - MSDP
- Tuned PIM Hello-interval
- Needs FAST convergence



Security Trading Enterprises Phase I

PIM-SM

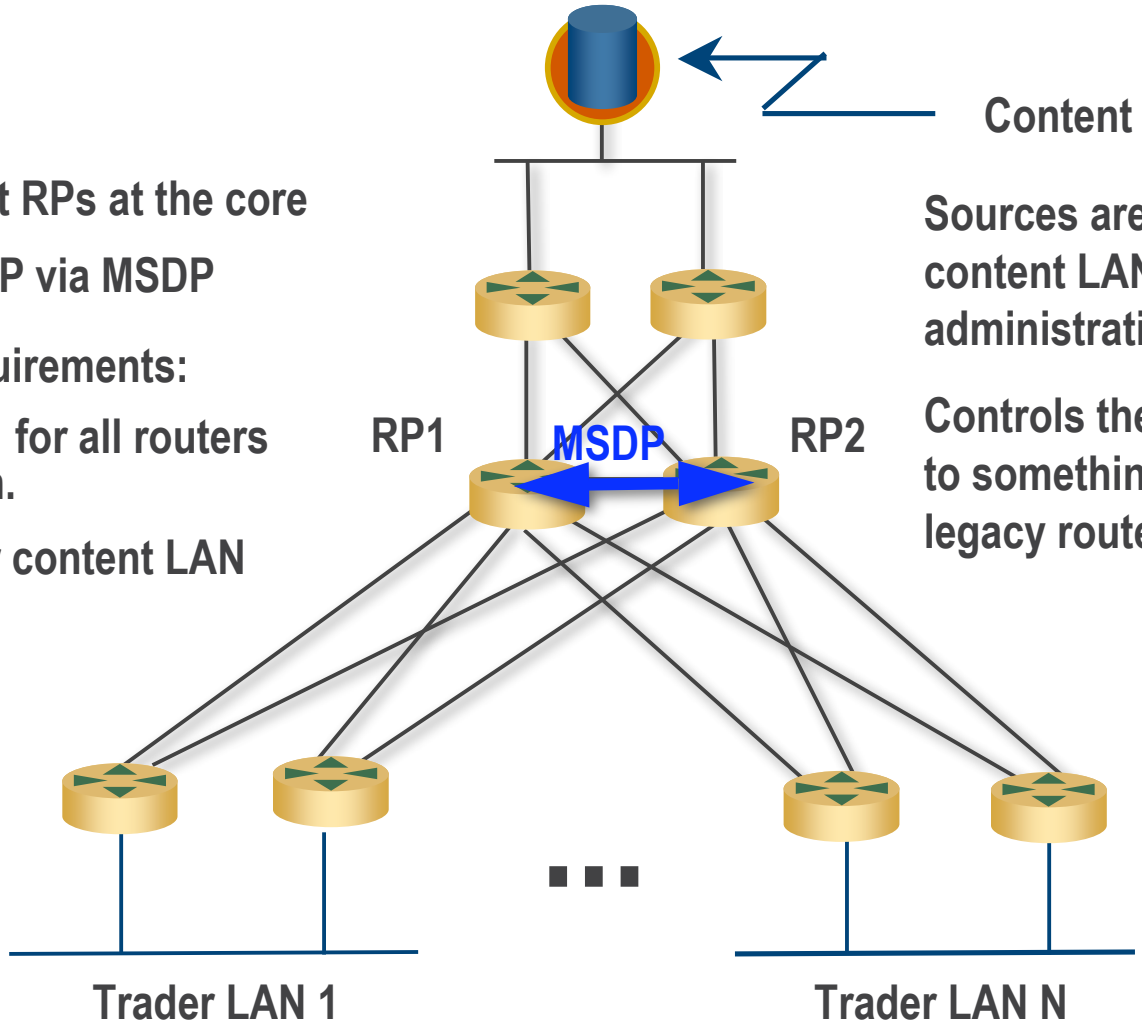
Redundant RPs at the core

Anycast-RP via MSDP

State Requirements:

(S)x(G) + 1 for all routers in the path.

(S)x(G) for content LAN routers



Sources are confined to the content LAN, and G is administratively controlled.

Controls the amount of state to something manageable in legacy routers.

Security Trading Enterprises Phase II

PIM-SM
Redundant RPs at each content LAN (multiple content LANS)

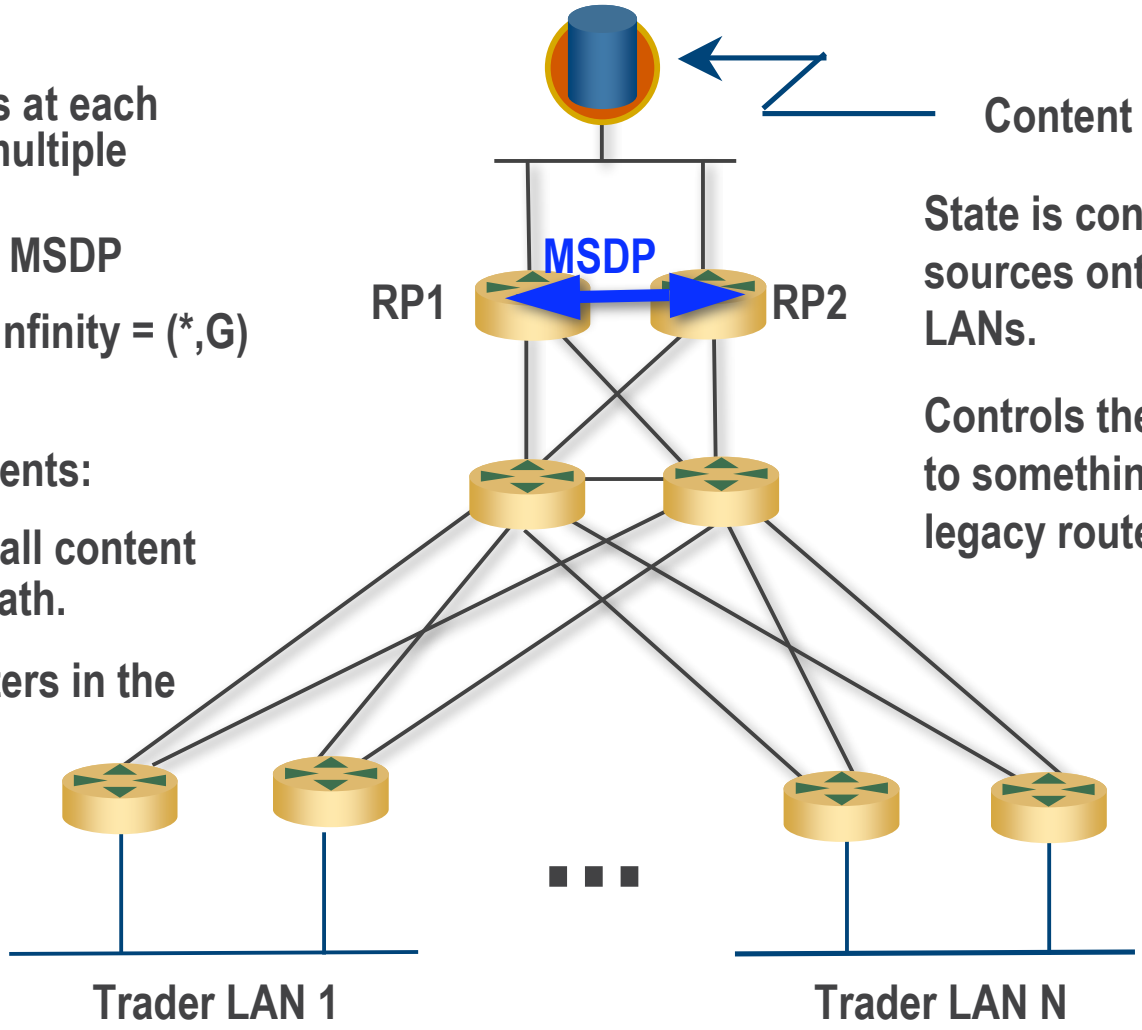
Anycast-RP via MSDP

SPT threshold infinity = (*,G) forwarding

State Requirements:

$(S_L)x(G) + 1$ for all content routers in the path.

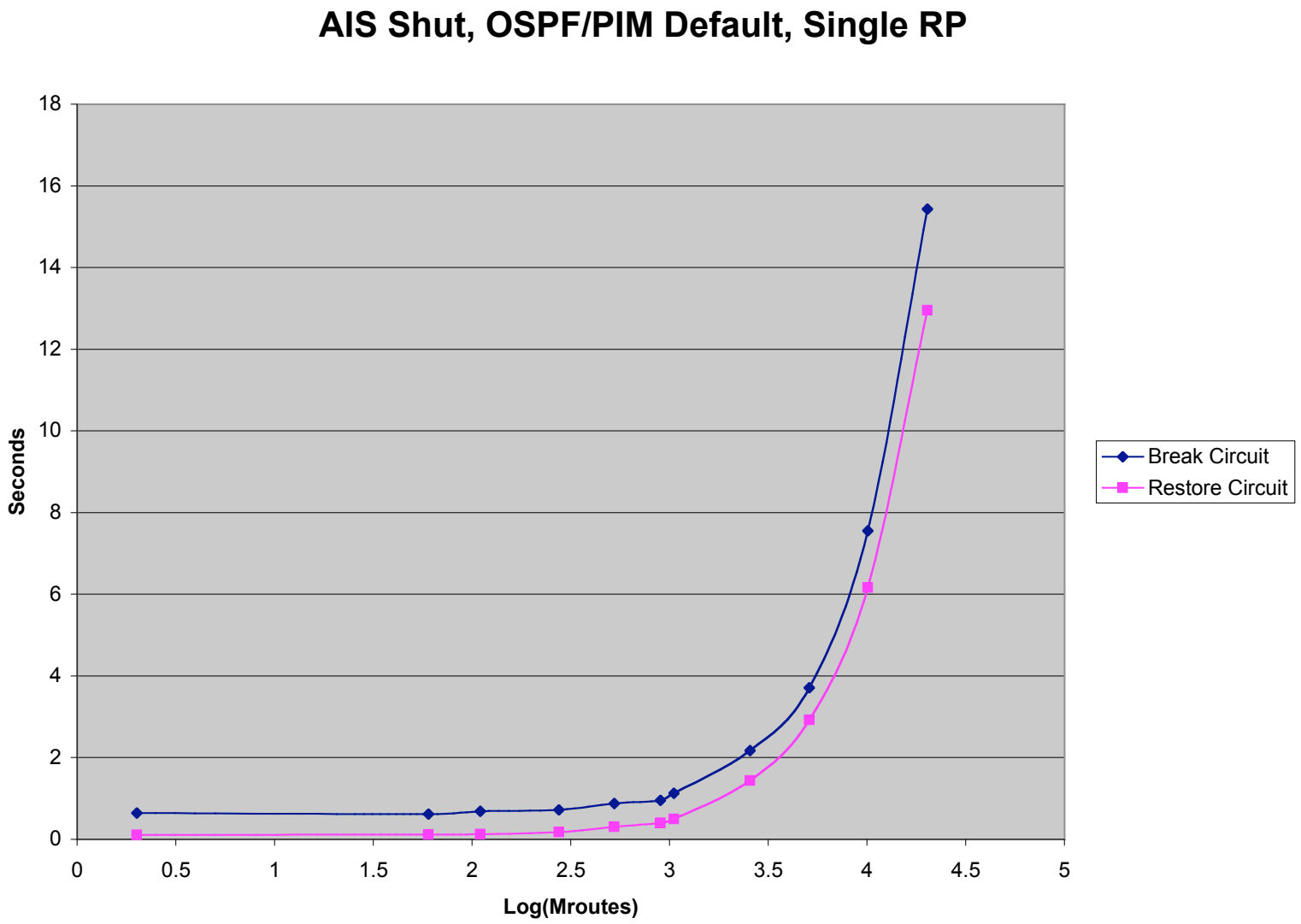
G for other routers in the path



State is controlled by splitting sources onto separate content LANs.

Controls the amount of state to something manageable in legacy routers.

Convergence Time
tightly coupled to mroute state



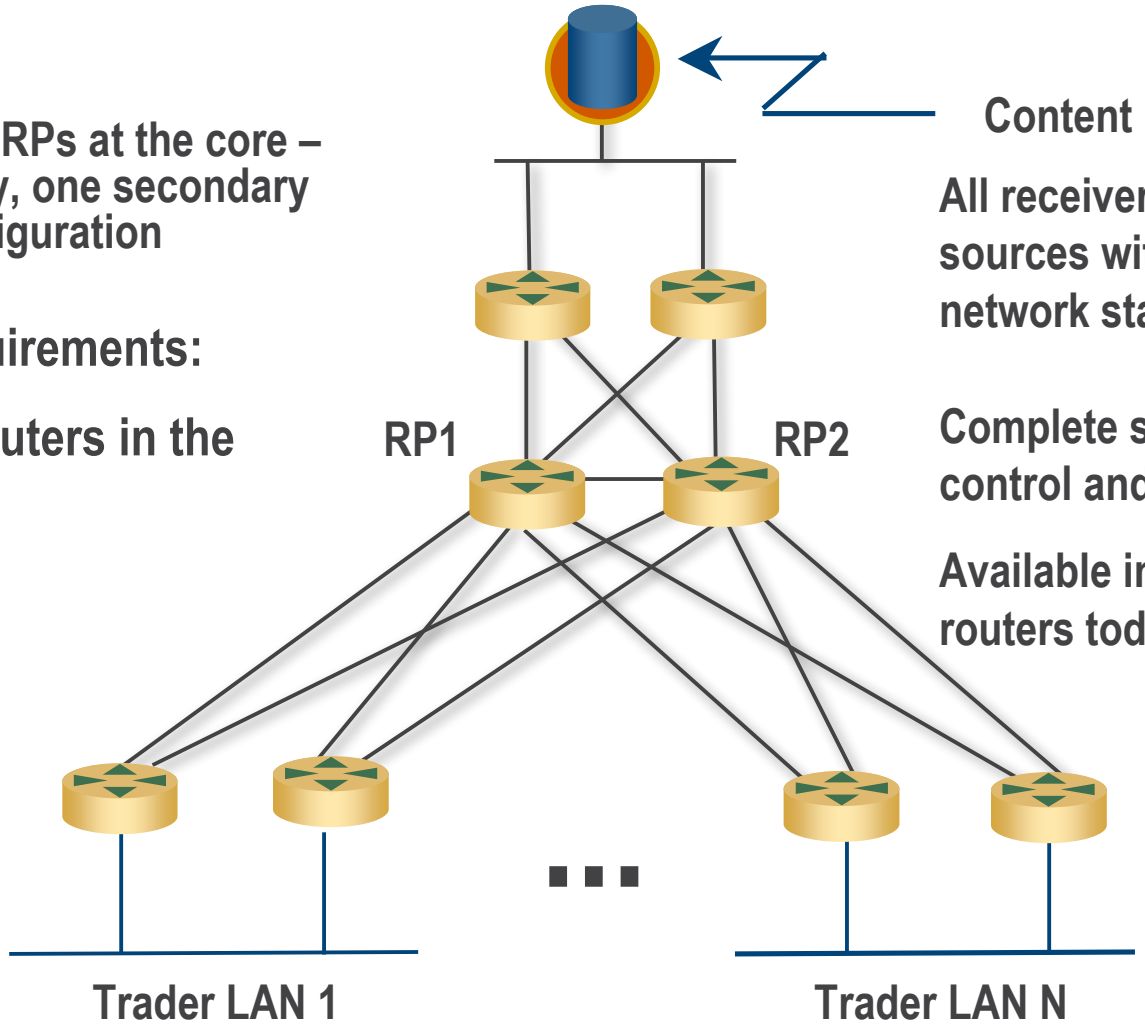
Security Trading Enterprises Phase III

BiDir PIM

Redundant RPs at the core – one primary, one secondary as per configuration

State Requirements:

G for all routers in the path



All receivers can also be sources with no impact on network state requirements

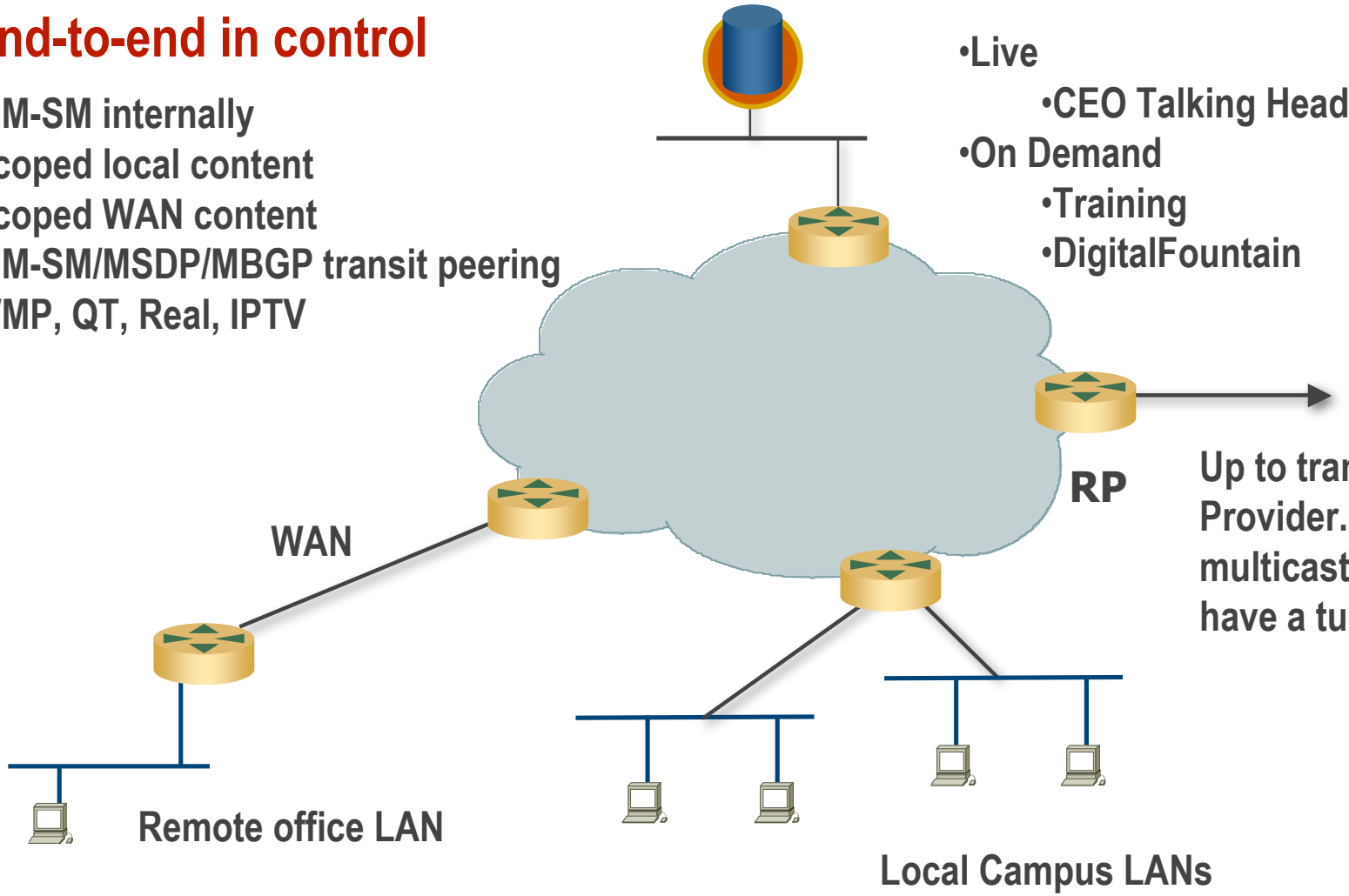
Complete separation of control and data plane

Available in high-performance routers today

How?
Enterprises

End-to-end in control

- PIM-SM internally
- Scoped local content
- Scoped WAN content
- PIM-SM/MSDP/MBGP transit peering
- WMP, QT, Real, IPTV

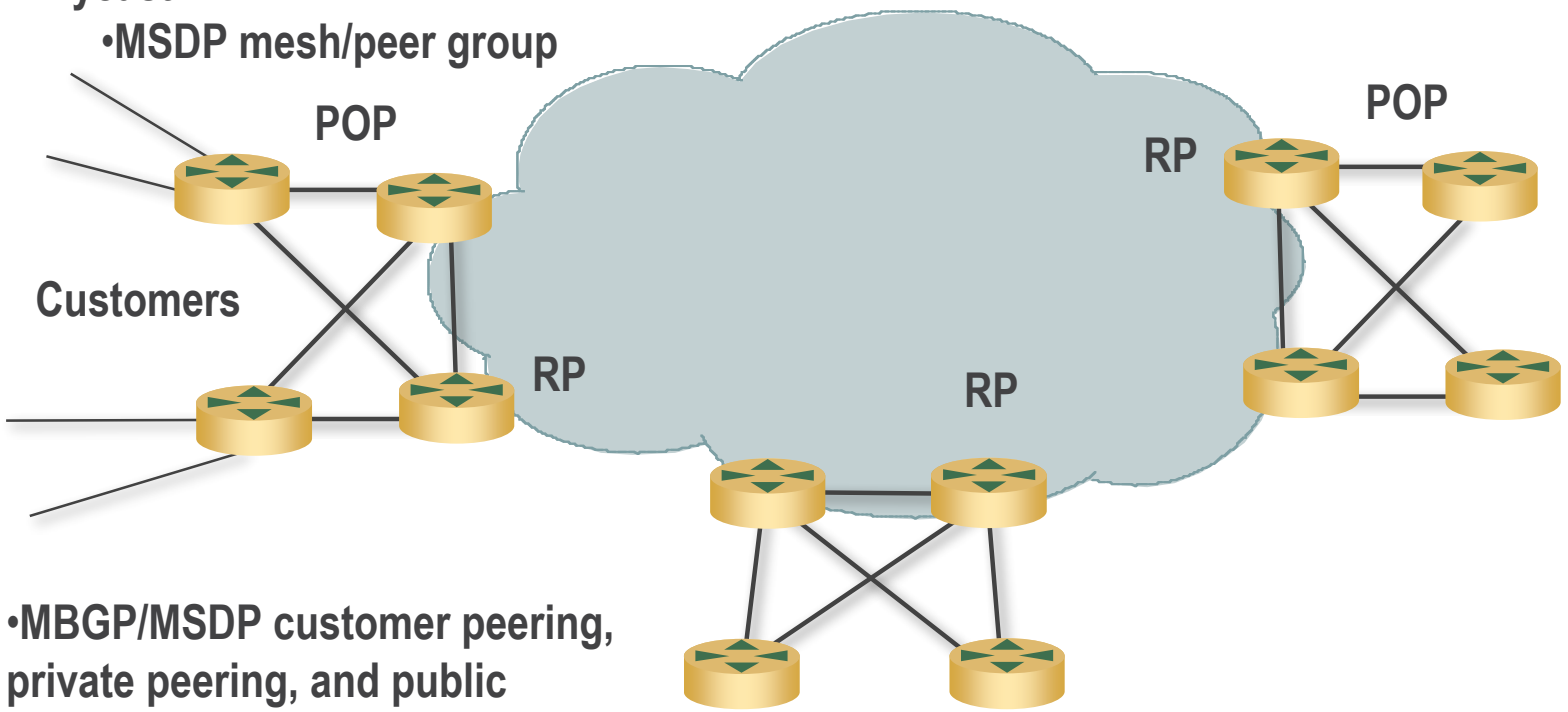


How?
Service Providers (IP)

Internal control only

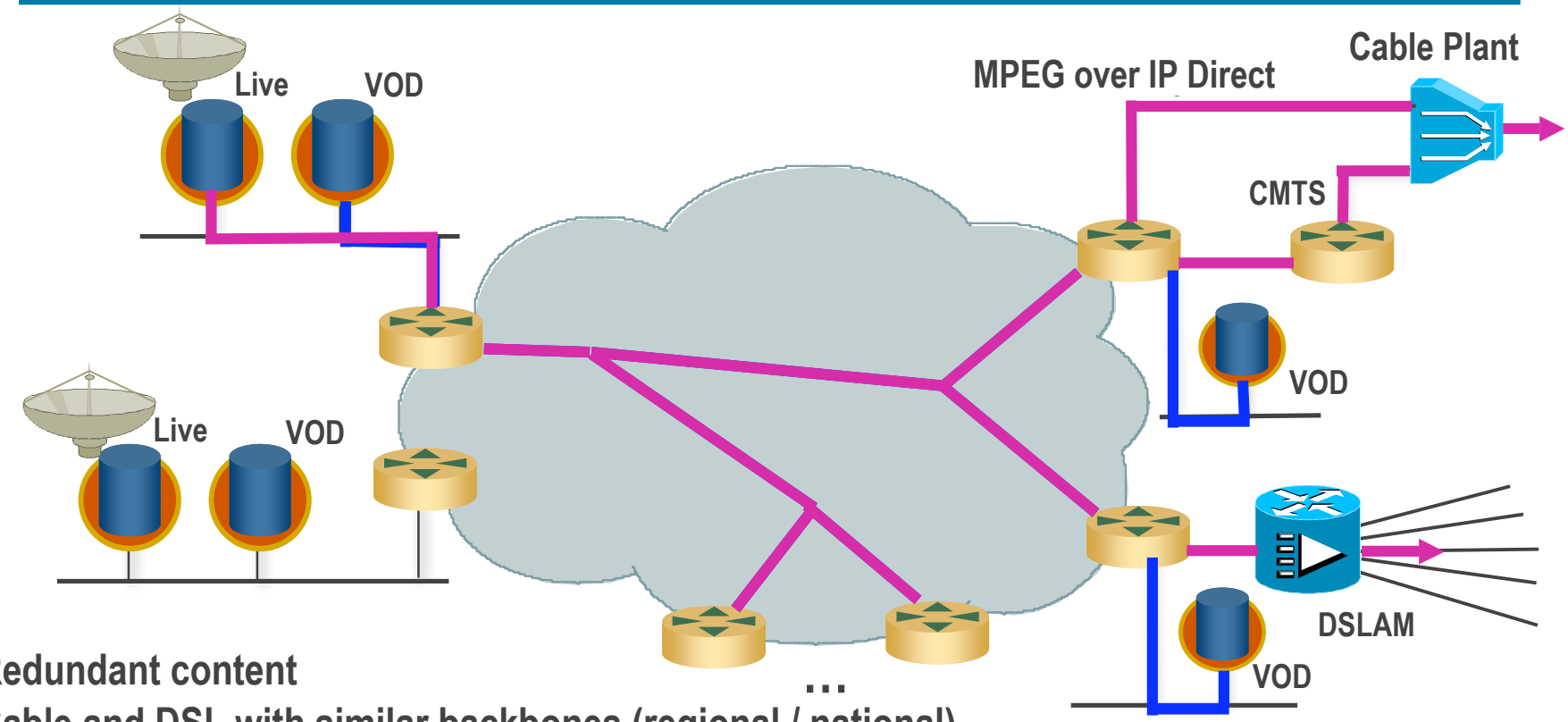
- PIM-SM internally
  - One PIM domain
- Anycast-RP
  - MSDP mesh/peer group

- May have sourcing policy
- Some content
- Peering with ISC :)



- MBGP/MSDP customer peering, private peering, and public exchange.

Content Service Networks
Cable/DSL



Redundant content

Cable and DSL with similar backbones (regional / national)

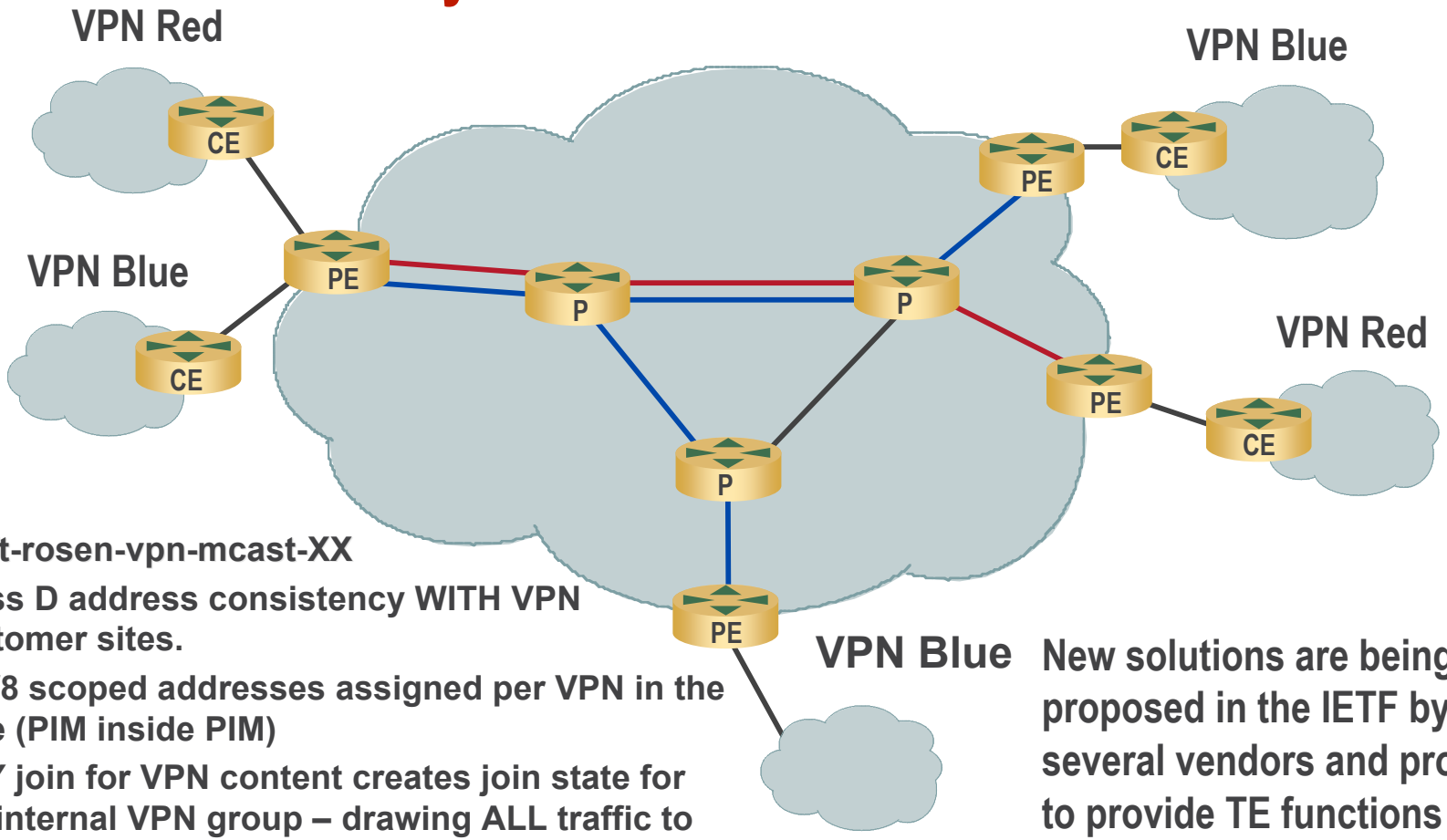
Numerous customer aggregation sites

Populate local VOD servers via multicast over the backbone

Live video over IP Multicast

How?
VPN Providers – Rosen opt2 (2547)

Internal control only



draft-rosen-vpn-mcast-XX

Class D address consistency WITH VPN customer sites.

239/8 scoped addresses assigned per VPN in the core (PIM inside PIM)

ANY join for VPN content creates join state for the internal VPN group – drawing ALL traffic to the PE. Unwanted group traffic (once decap'd by the PE) must be dropped.

New solutions are being proposed in the IETF by several vendors and providers to provide TE functions for MVPNs.

But what happened to the dream?

- ISC Multicast Project
- Mroute Aggregation Network
- Multicast.isc.org

ISC Multicast Project
What?

- **Mroute aggregation network**
  - *Routers at EQX, PAIX, ORIX, 200 Paul*
  - *Transit tunnels to other IX routers in Italy, Germany, Asia*
- **Native peering at MIXs**
- **Direct P2P transit/peering**
- **Transit tunnels to remote MIXs**
- **Transit tunnels for remote networks**
  - *Campus, office, lab, edge networks, content networks*
- **Transit tunnels for events (NANOG, IETF, etc)**
- **Configuration assistance, troubleshooting, training**
  - *http://www.nsrc.org*

ISC Multicast Project
Why?

- Internet Multicast Peacenics are still out there
- Multicast is still not ubiquitous
- Promote connectivity in support of interests, applications, and content
- “updated MBONE” to encourage continued usefulness of MCAST
- Develop multicast services at remote IXs
- Sandbox for new multicast solutions

ISC Multicast Project
How?

- **Native L2 peering (transit) at the MIXs**
  - *EQX, PAIX, FIXW, ORIX, 200 Paul*
  - *PIM, MBGP, MSDP*
- **PNI peering / transit**
- **Tunnels**
  - *Remote IXs, networks, and events*
  - *GRE, IPIP*
- **Remote RP services**
- **Default route for receiver-only networks**

ISC Multicast Project
Success example

- **USP (Brazil)**
  - *Campus eng contacts me for transit tunnel*
  - *Assist with config, and campus mcast deployment planning*
  - *Local content sourcing, receiver application testing*
  - *Help set requirements for upstream provider*
  - *Upstream turns-up native transit in SE US.*
  - *USP Campus rolls over to native upstream transit*
  - *Tunnel is turned down*

ISC Multicast Project
Some coolness

- 5900 prefixes today
- 25 peers
- 600 PIM joined streams
- ISC has the largest mcast footprint (prefixes)
  - *True global transit service offering*
- Bitrate
  - *Background: 5-10Mbps*
  - *Peak: 40Mbps*
  - *Internet2 raw HDTV streams becoming available*
- Production-dependent connectors today
- AUP-free access to Internet2 content and connectors

ISC Multicast Project Future

- **Peering / transit always available**
- **Looking for**
  - *More holes to fill*
  - *More islands to bridge*
  - *Interesting content, applications, solutions, etc*
- **M6Bone**
- **AMT – auto-tunnels – Last Mile Solution (the missing piece)**
  - *draft-ietf-mboned-auto-multicast-xx*
- **Is Multicast Growing?**
  - *YES – but from the edges-in*
  - *Large walled-garden projects worldwide*

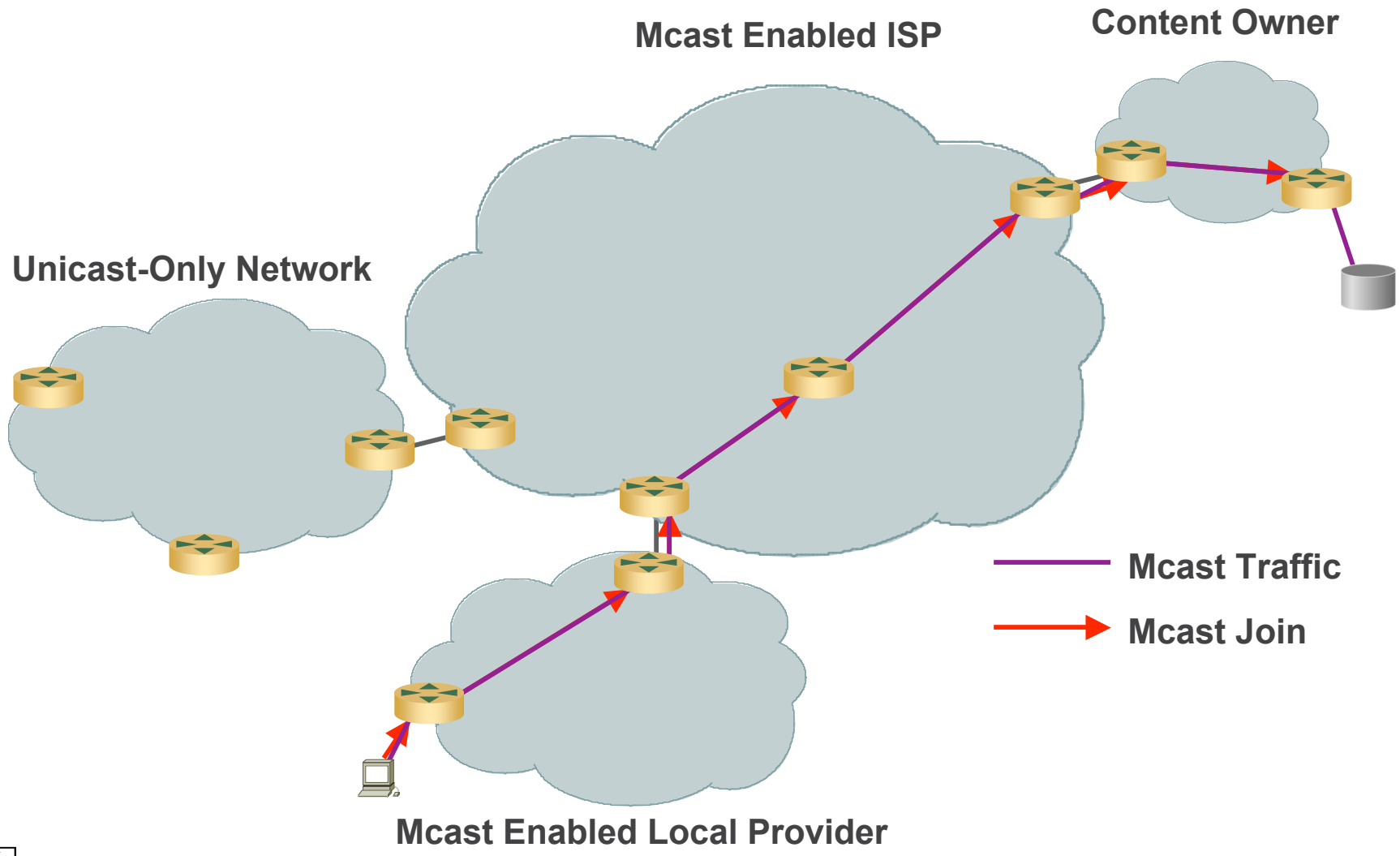
ISC Multicast Project
Thanks to all who have helped!!

Cisco
DigitalFountain
EQX
ISC
Juniper
NASA / FIXW
ORIX / UofO
PAIX / S&D
Procket
Sprint
Verio

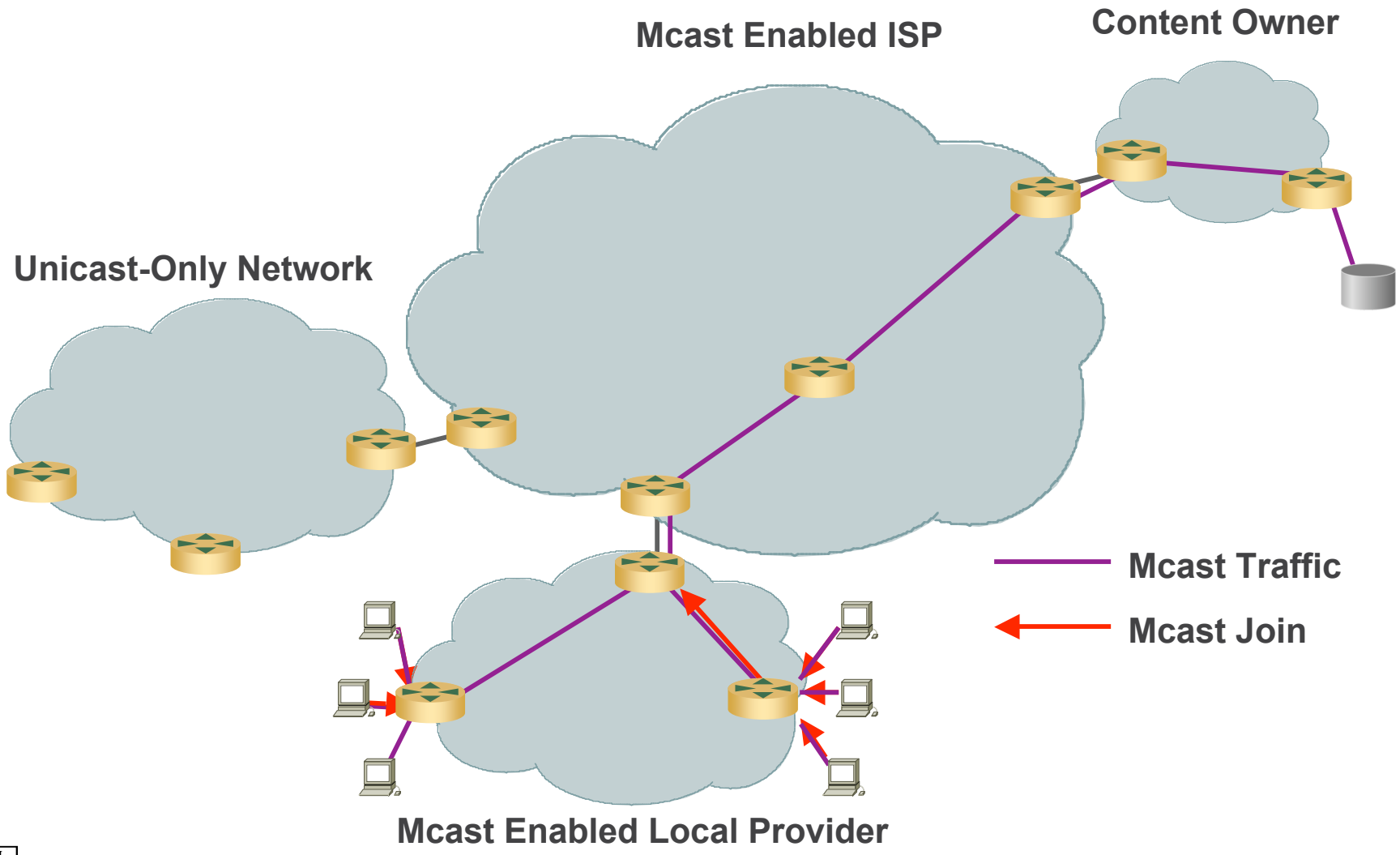
Going Forward…

So what's right and wrong with multicast today?
What's being done to fix it?

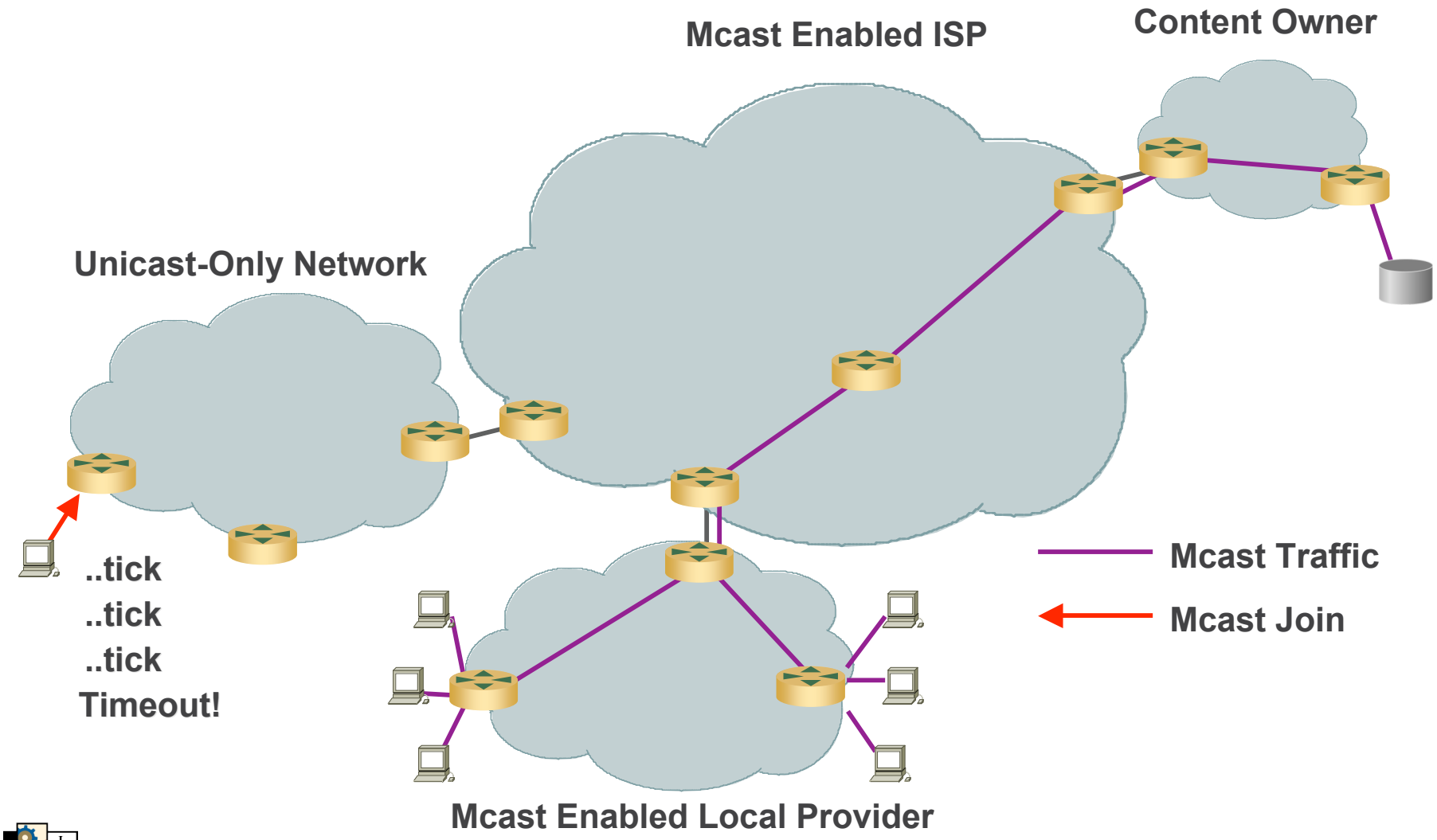
What Worked?



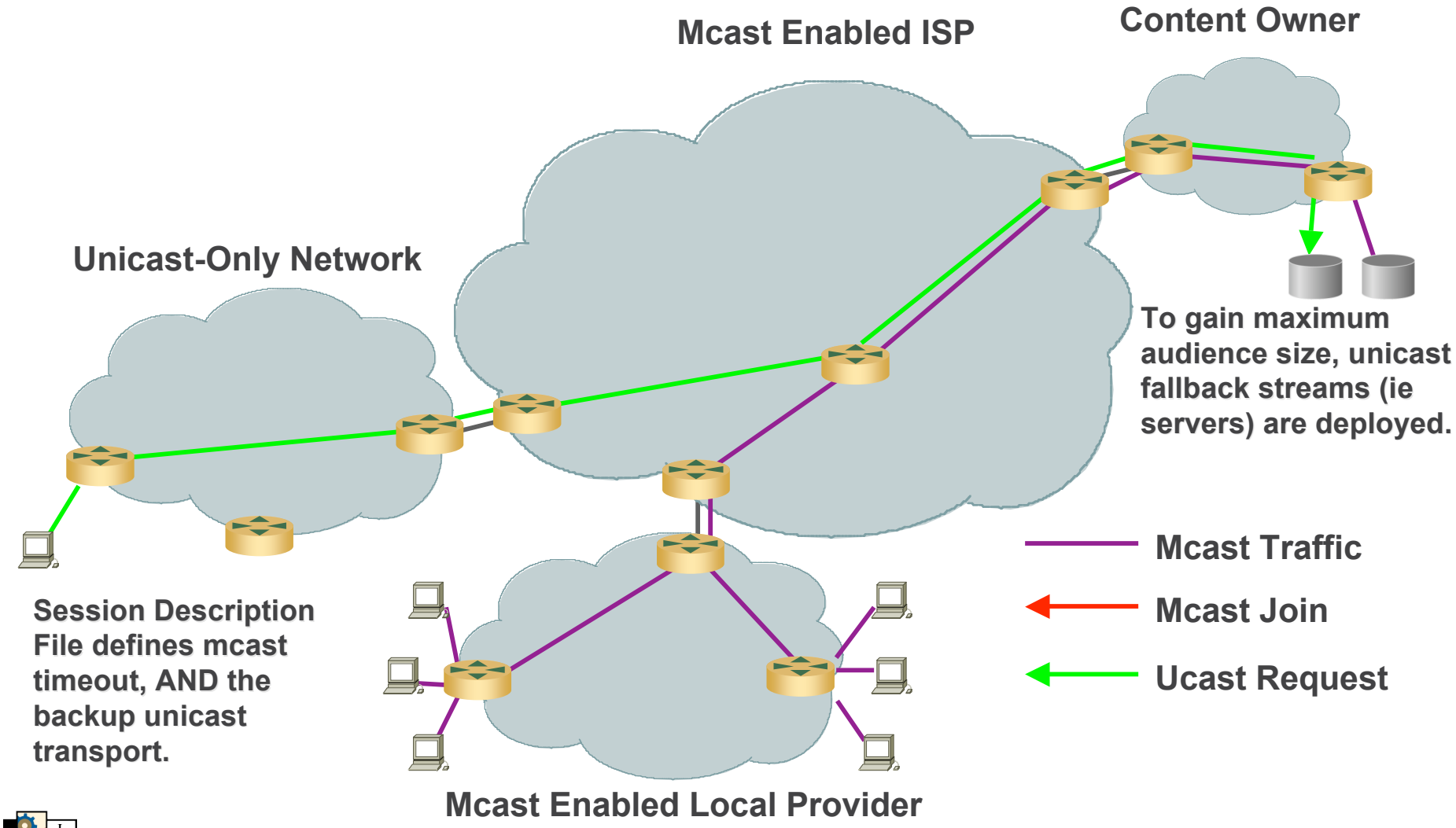
What Worked?



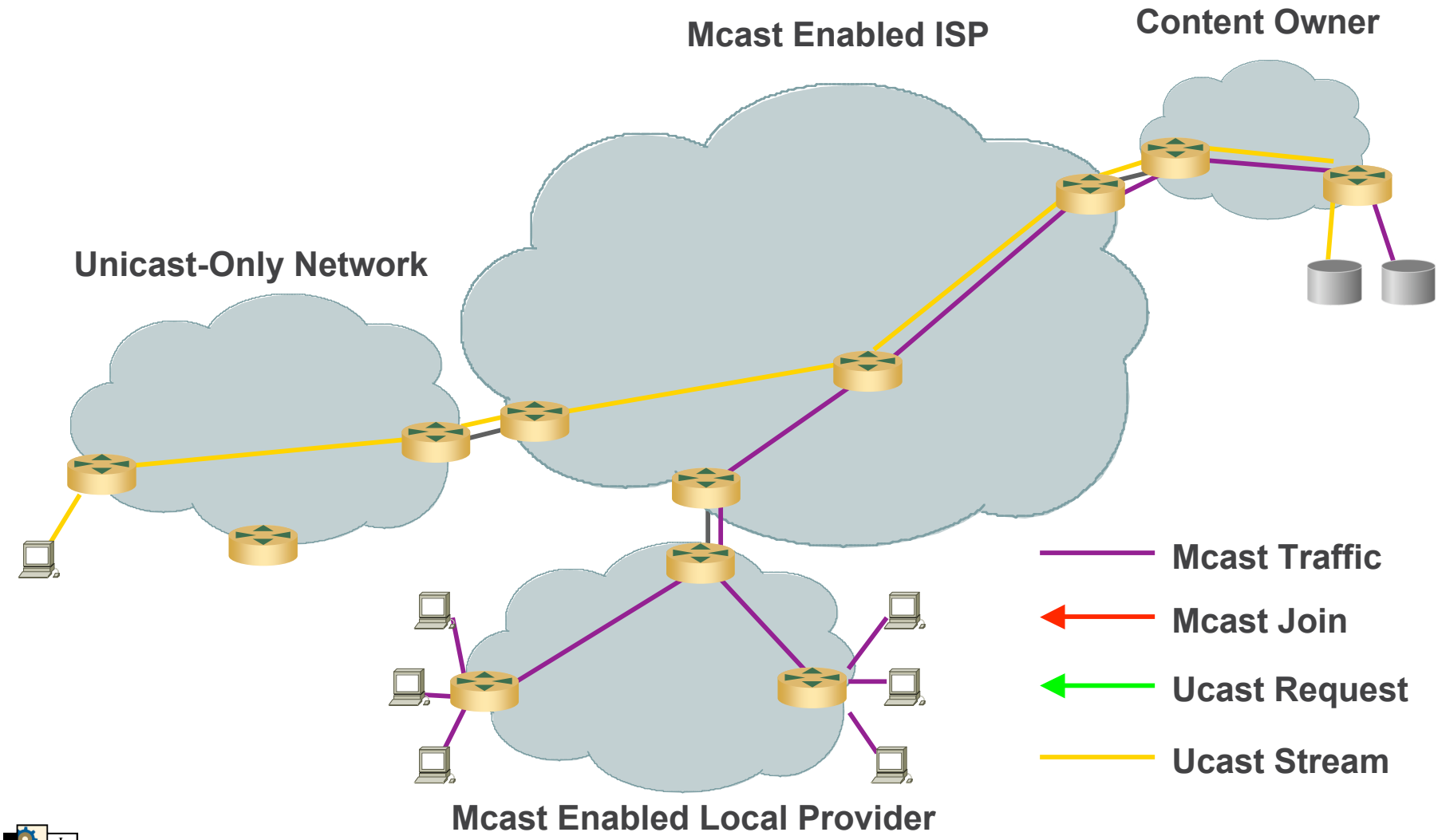
What Worked?



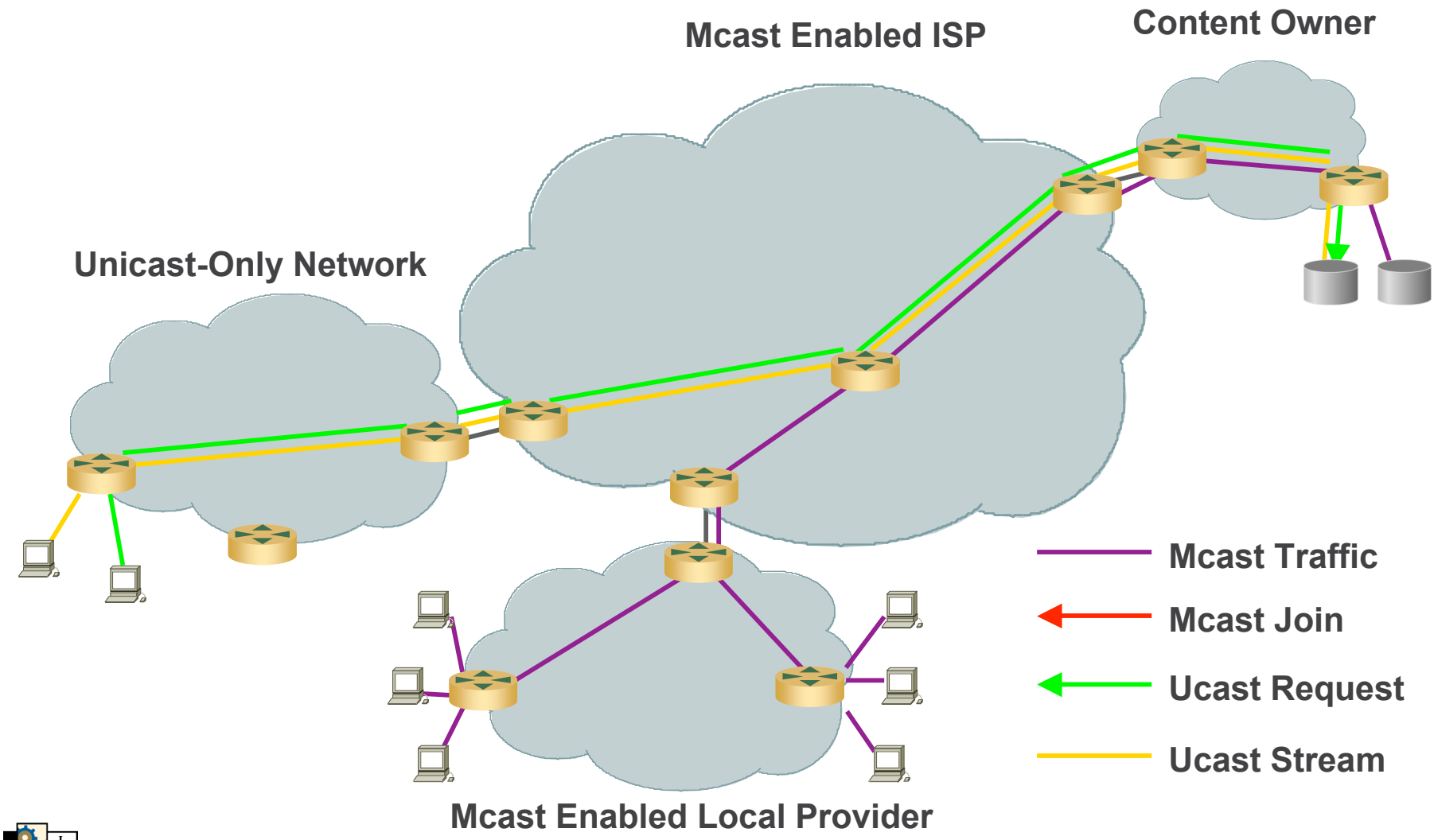
What Didn't?



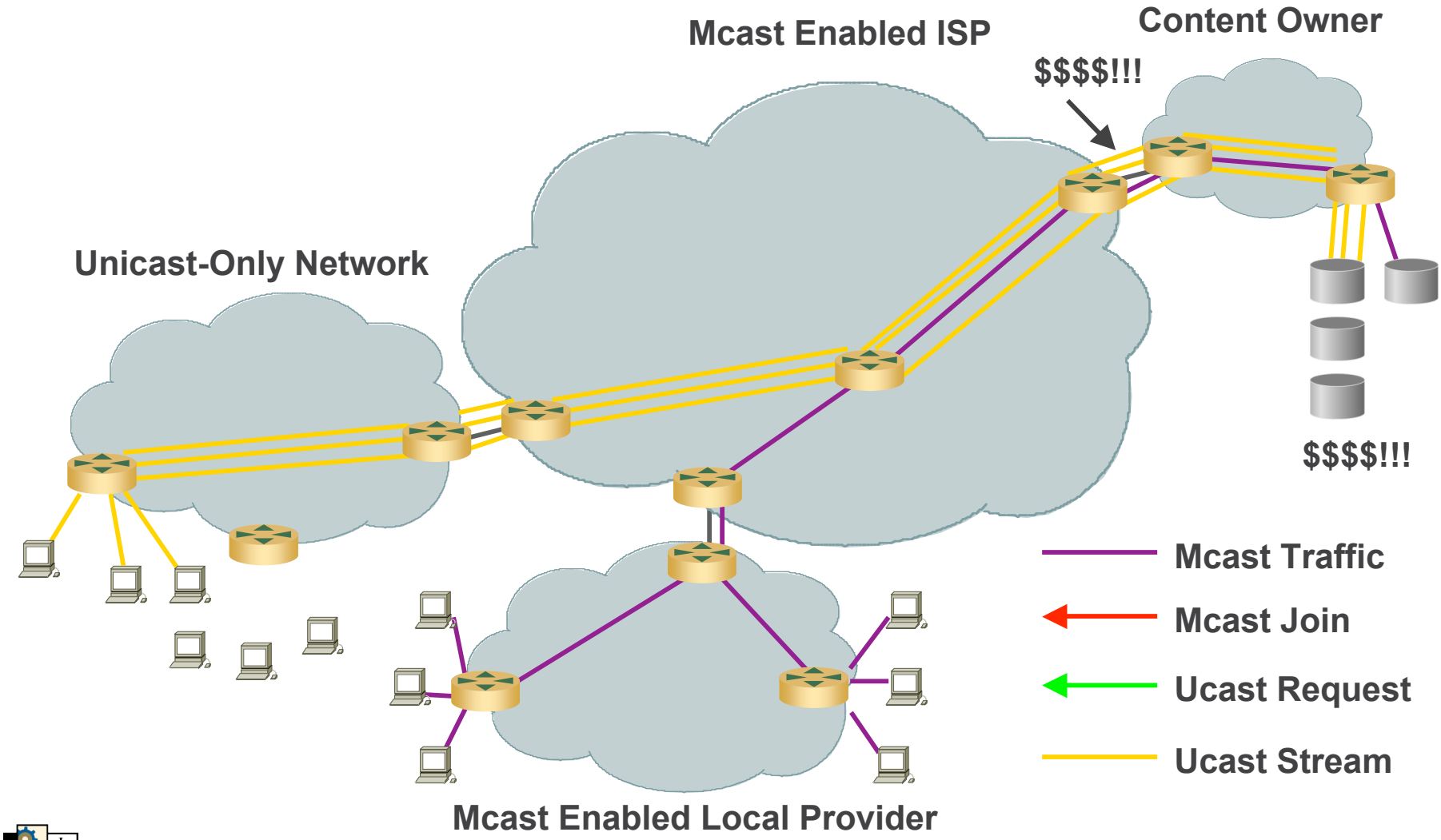
What Didn't?



What's Wrong?



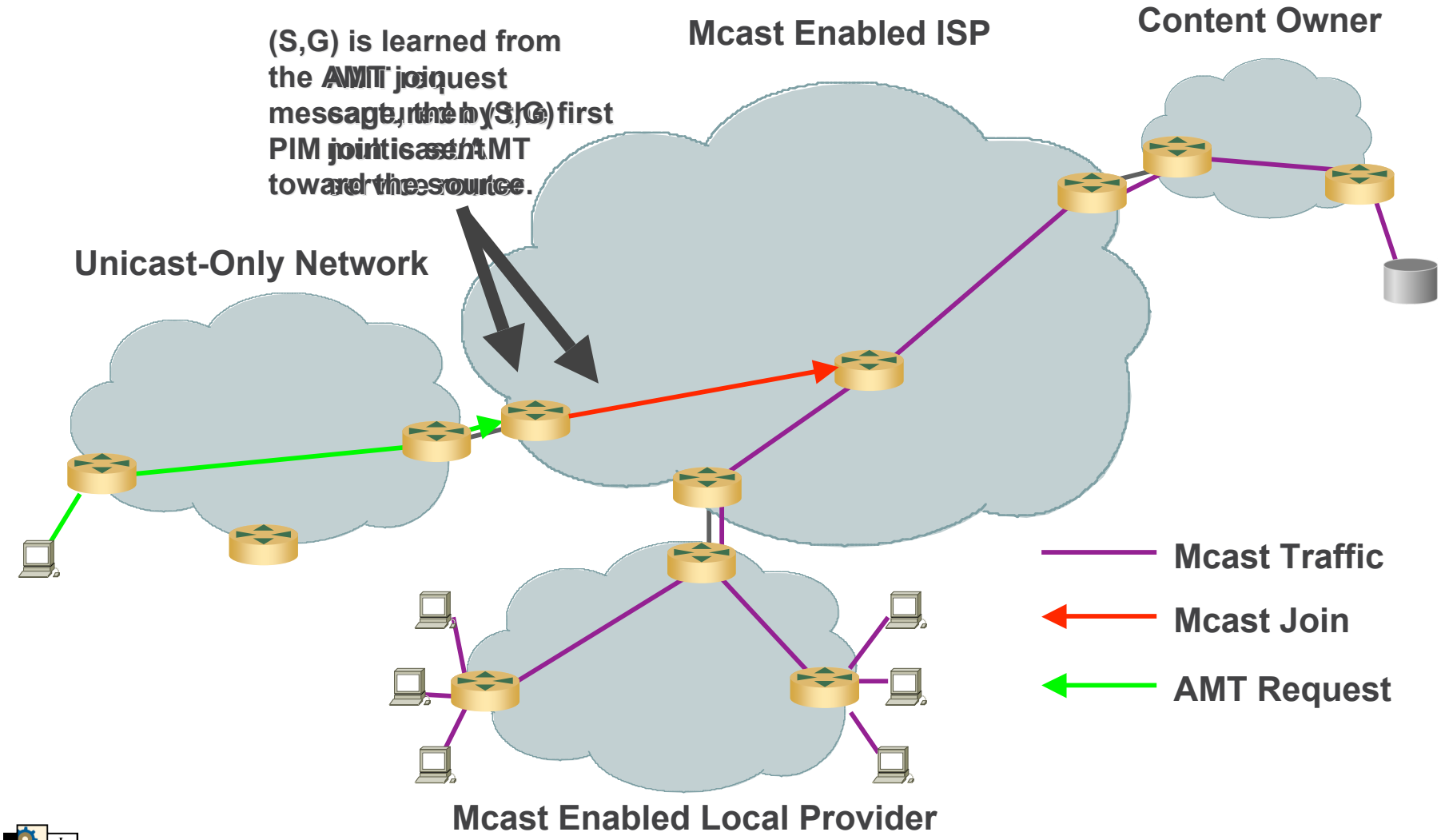
What's Wrong?



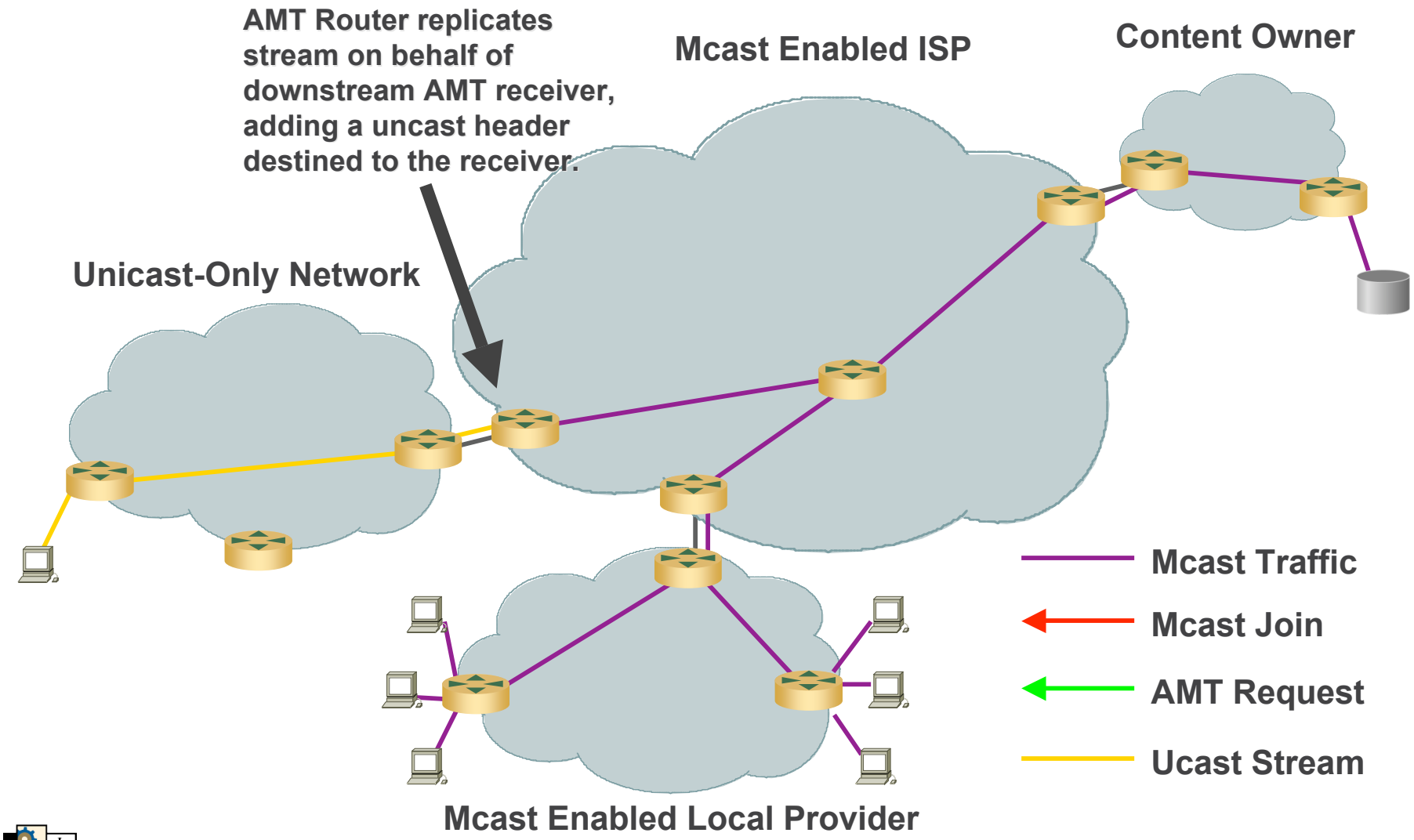
What's Wrong?

- Multicast in the Internet is an all-or-nothing solution
- Is there a way to provide a multicast-only solution for content owners?
- Do they really need multicast today for global distribution?

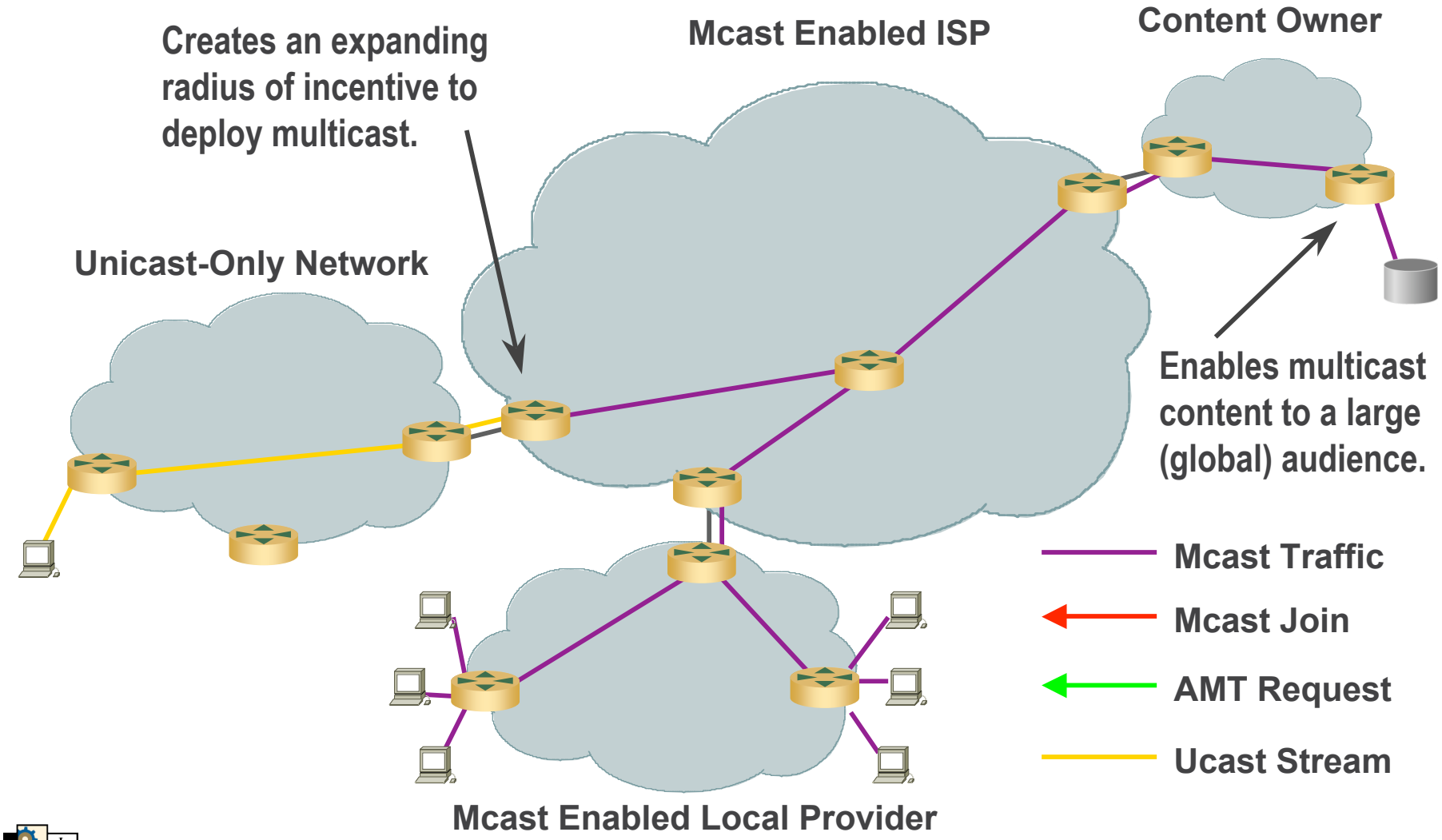
One Solution: AMT



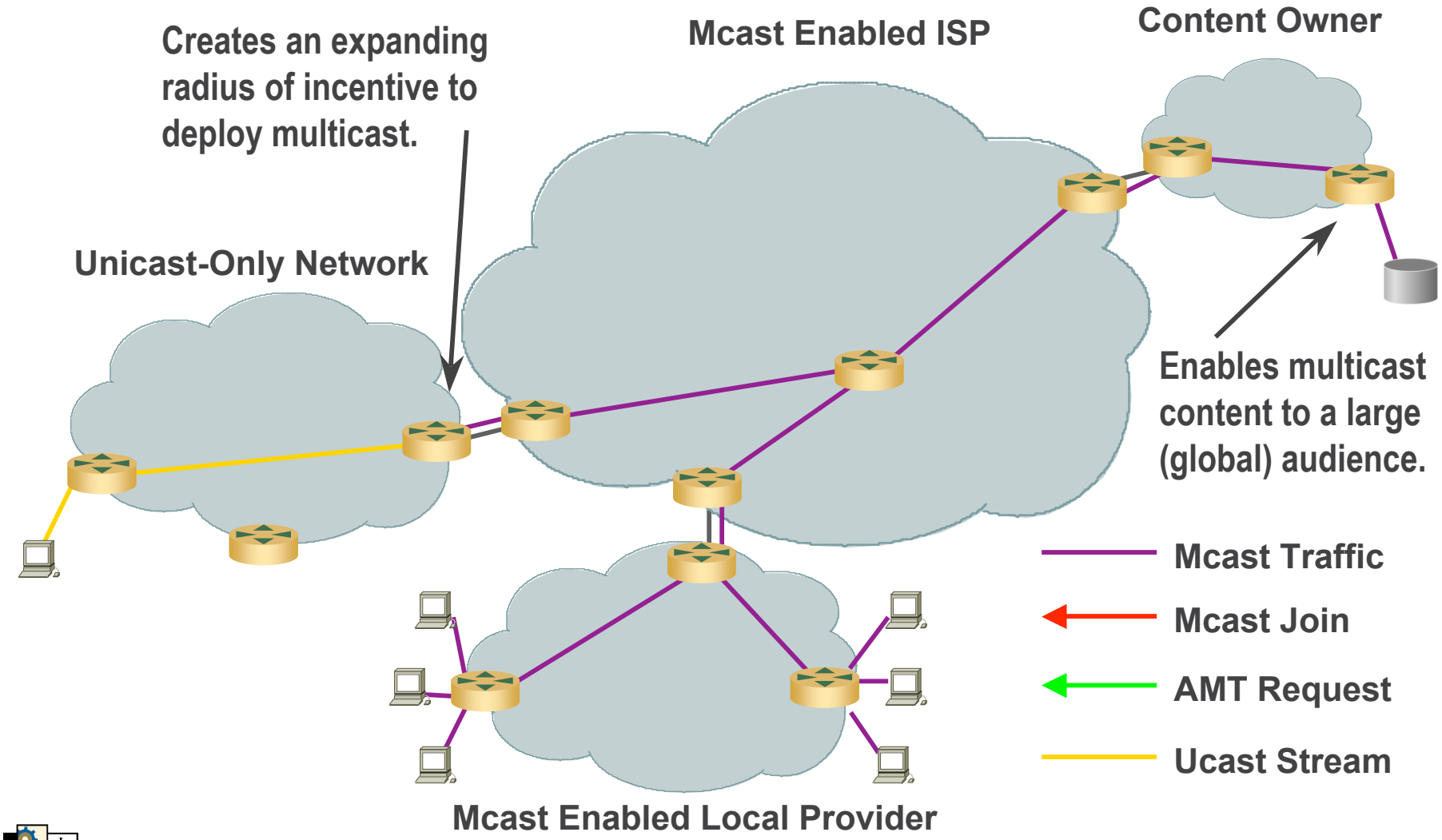
One Solution: AMT



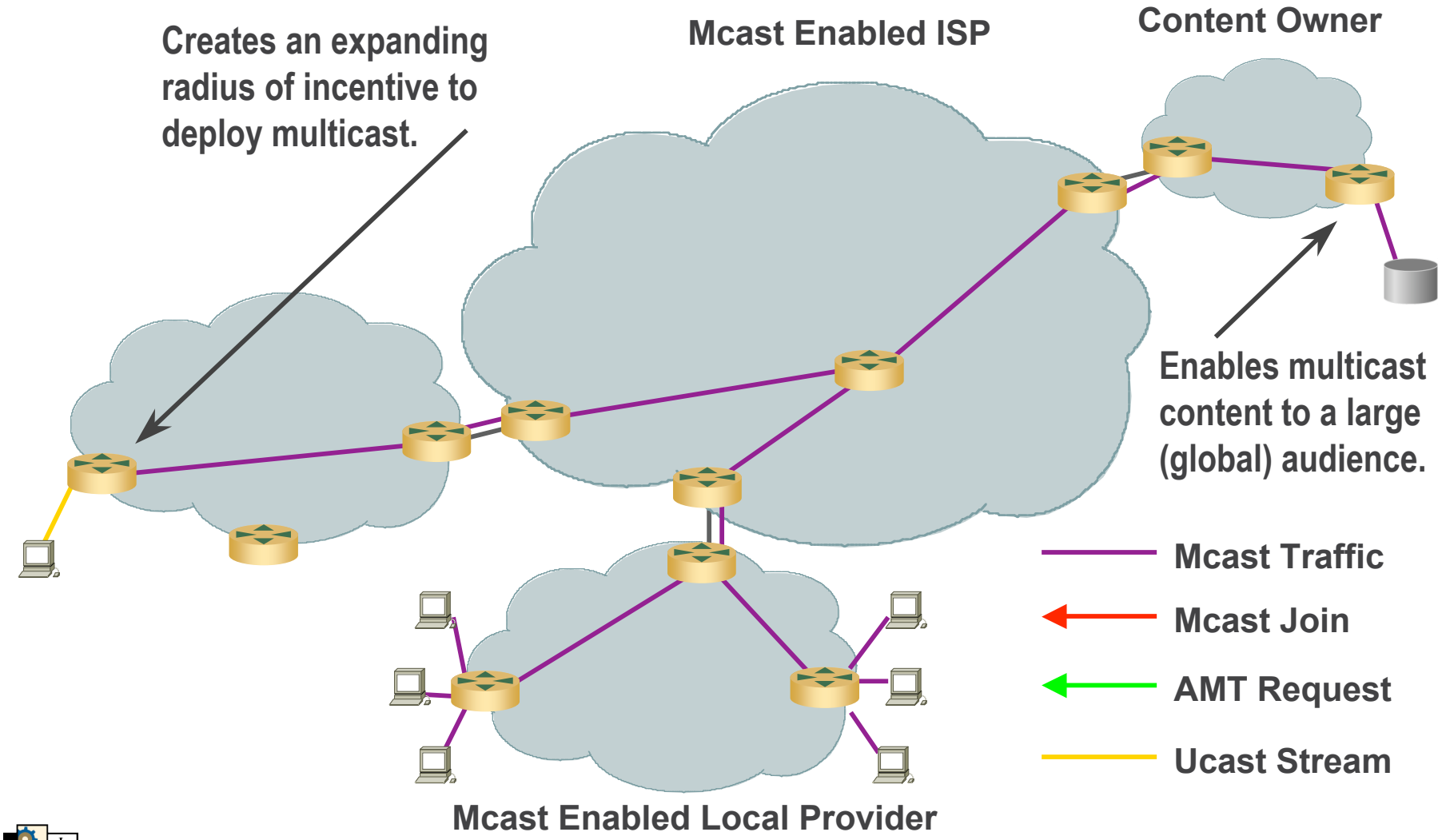
One Solution: AMT



One Solution: AMT



One Solution: AMT



While Multicast Sleeps…

- **Perfect for one-to-many live content**
  - *But without a large enough audience it was a non-starter in the Internet*
  - *"It's become an enterprise application" …heard during the break.*
- **Content "distribution" has been dominated with P2P**
  - *But no live p2p solution today*
- **A subscriber-distributed sharing network with enough storage and bandwidth could provide VOD on a massive scale.**
- **A few docs on the scaling and distribution characteristics of P2P networks:**
  - *http://www.americafree.tv/rankings/vc_usage.total_hours.post_fit.rank_shift_8.png*
  - *http://www.isi.edu/~annc/papers/ChervenakThesisNov94.ps.Z*
  - *http://www.isa.its.tudelft.nl/~pouwelse/Bittorrent_Measurements_6pages.pdf*

Multicast today

- **Walled Gardens are taking-off globally**
  - *Triple-play networks*
  - *Server Farms*
  - *Financial networks*
- **Global multicast connectivity**
  - *Still growing, but at a much slower-rate*
  - *Still needs to overcome the all-or-nothing barrier*
  - *SSM a must*
  - *ISC will continue to spread the love.*

Thank you!

shep@cisco.com

www.cisco.com

www.isc.org