EXPERIENCE FROM L2TP IMPLEMENTATION FOR BITSTREAM

Rafal Szarecki

Date

AGENDA

What is specific for bitstream

L2TP technology recap

Interesting design points

- User identification
- MTU issues.
- Rate enforcement on LAC
- ML-PPP over L2TP and QoS

WHAT IS BITSTREAM

Bitstream Access (w skrócie **BSA**) - termin określający usługę sprzedaży **szerokopasmowej transmisji danych** (np. dostępu do Internetu), najczęściej za pomocą linii telefonicznej (w technologii xDSL). Usługa świadczona jest przez operatora korzystającego (alternatywnego) na rzecz jego klientów, przy wykorzystaniu infrastruktury sieciowej innego operatora (zazwyczaj dominującego na rynku). W rzeczywistości klient końcowy korzysta więc wyłącznie z infrastruktury właściwego operatora sieci, podczas gdy umowę dostawy usługi zawiera z innym. Ten z kolei, na mocy stosownej umowy ramowej, dzierżawi infrastrukturę właściwego operatora sieci.

WHAT IS SPECIFIC FOR BITSTREAM ACCESS

Usługa dizedziczy ograniczenia techniczne wynikajace z infrastruktury operatora właściwego.

- Technologia (e.g. C-Vlan vs. S-VLAN)
- Adresy/identyfikatory sieciowe (VP/VC, VLAN ID)
- Przepływności
- Identyfikowanie abonenta.

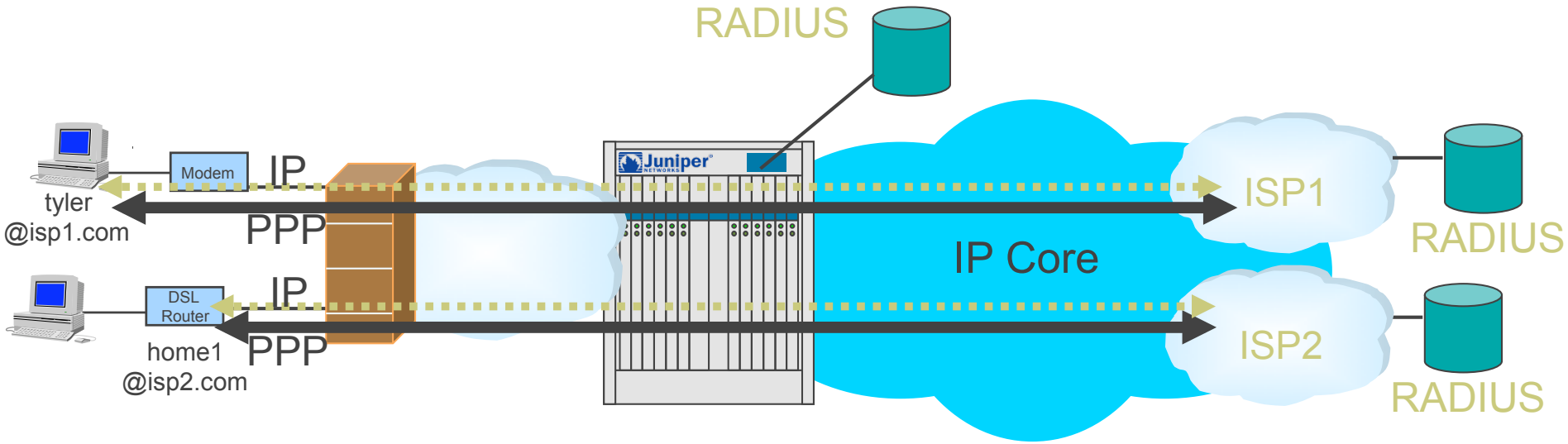
Operatorzy nie lubią współdzielić baz danych

- Aspekty prawne nie są aż tak istotne – identyfikator urzytkownika I jego profil to nie dane osobowe
- Aspekt biznesowy – informacje stat. o abonentach
- Bezpieczeństwo sieciowe baz danych
- Trudności integracyjne:
  - Struktura
  - Wydajność

Szczególnie problematyczne dla L2TP, gdyż brakuje identyfikatora abonenta na w L2 (VLAN, VC)

L2TP TECHNOLOGY RECAP

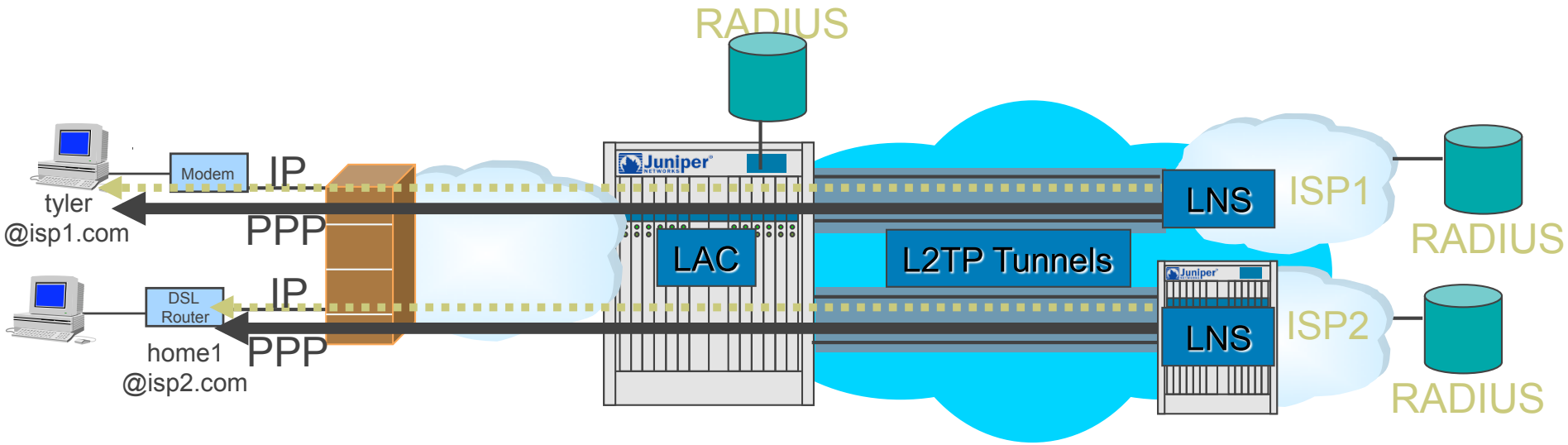
L2TP TERMINATION POINTS



L2TP is a client/server protocol that allows PPP to be tunneled across a network

- Layer 2 connection between user and B-RAS
- PPP session terminated at remote location
- Local or remote RADIUS server provides authentication
- User's IP interface terminated at remote location
- IP core routes L2TP-encapsulated packets

L2TP COMPONENTS



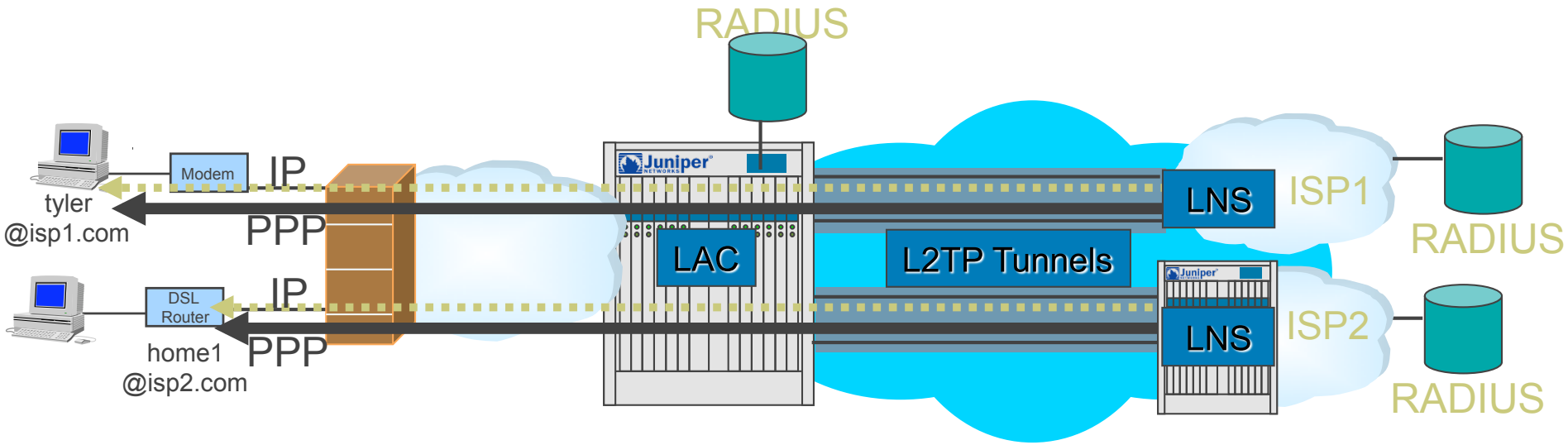
LAC:

- Physical or Layer 2 link termination
- Located at the ISP's point of presence
- Initiates L2TP tunnel and session

LNS:

- Located at the tunnel termination point
- Same provider, different provider, or customer site
- Terminates the PPP session
- Manages the user's IP interface

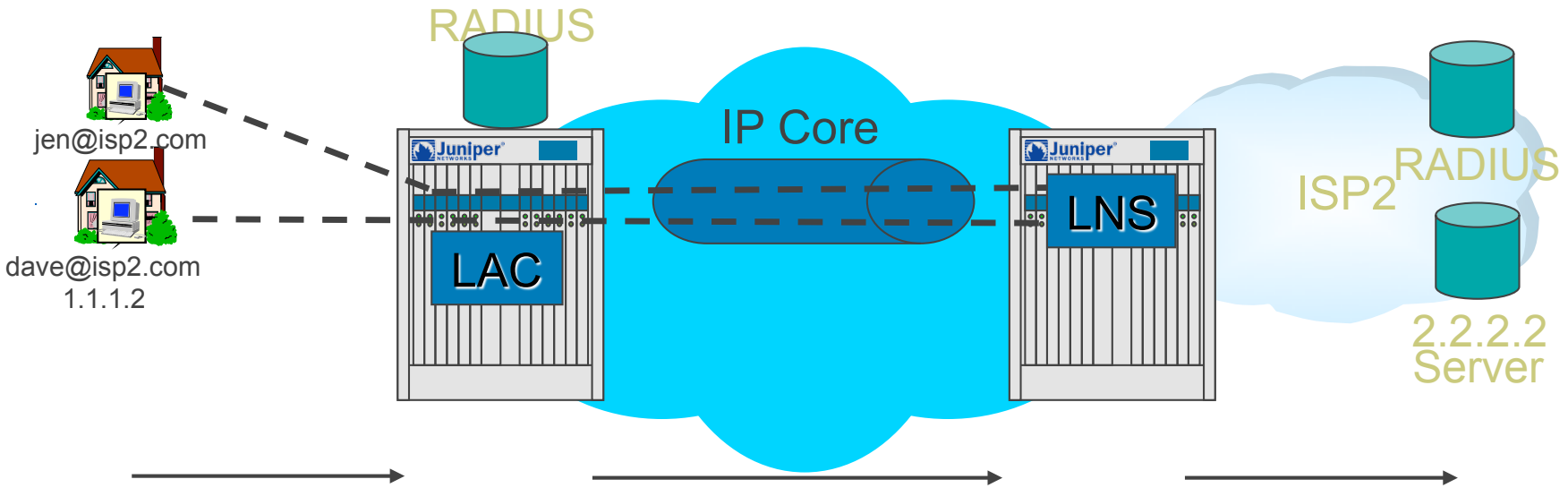
L2TP APPLICATIONS: WHOLESALING



B-RAS wholesaling model:

- Access service provider owns last mile
- Hands off PPP session to different ISPs

L2TP LIFE OF A PACKET

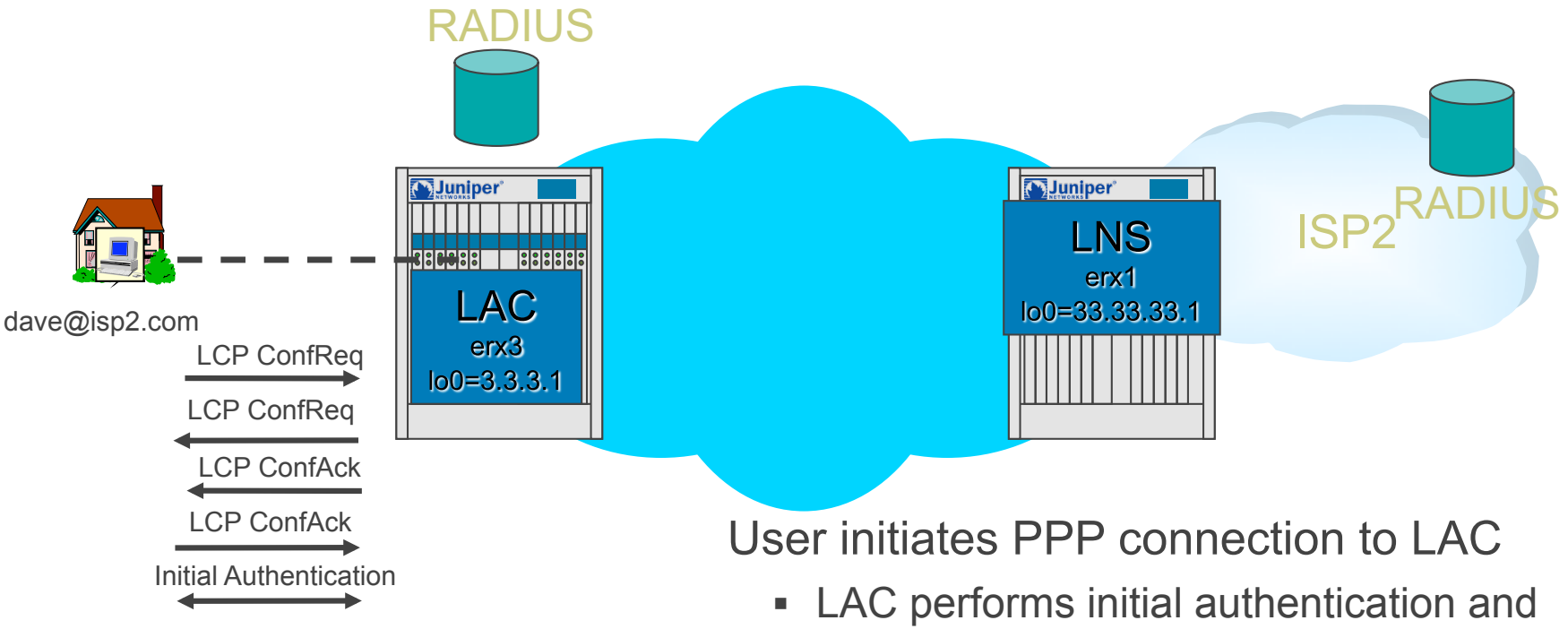


DA IP=2.2.2.2
SA IP=1.1.1.2
PPP Header
ATM Header
RFC 2684
VPI/VCI
Physical

DA IP=2.2.2.2
SA IP=1.1.1.2
PPP Header
L2TP Header
Tunnel ID
Session ID
UDP
DA IP LNS
SA IP LAC
Data Link
Encapsulation
Physical

DA IP=2.2.2.2
SA IP=1.1.1.2
Data Link
Encapsulation
Physical

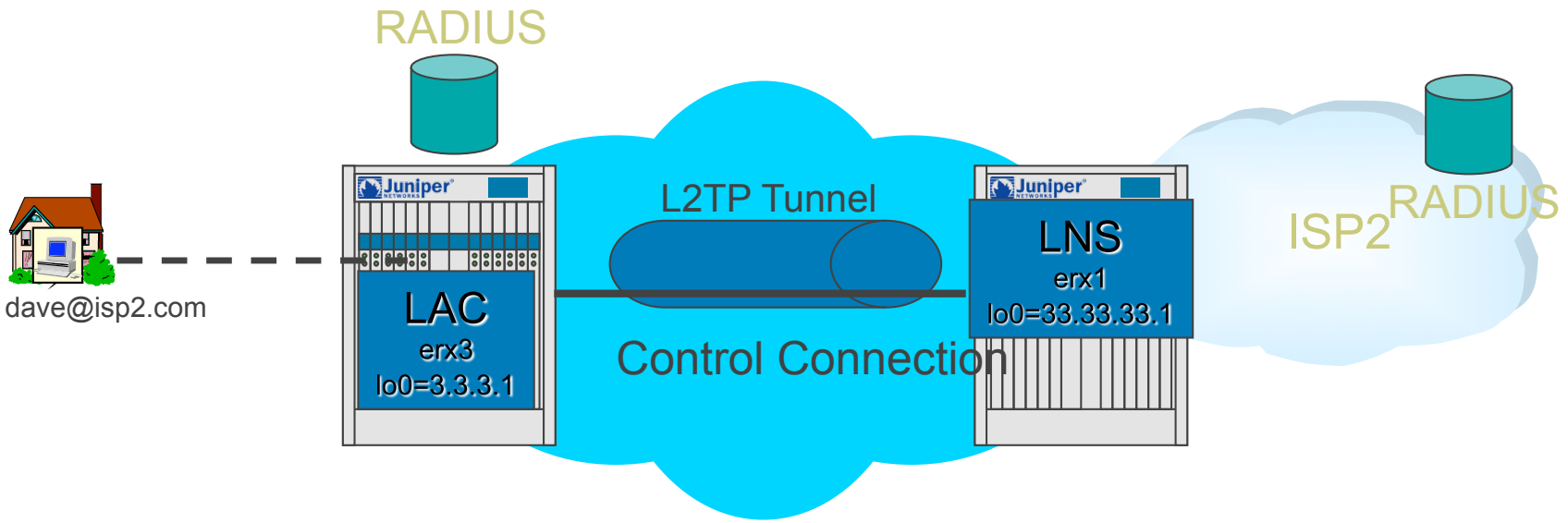
PPP SESSION INITIATION



User initiates PPP connection to LAC

- LAC performs initial authentication and determines whether to:
  - Terminate PPP session locally
  - Tunnel PPP session to an LNS
- Tunnel attributes obtained:
  - Local LAC config
  - RADIUS

L2TP TUNNEL ESTABLISHMENT



Start Control Connection Request (SCCRQ)

Start Control Connection Reply (SCCRP)

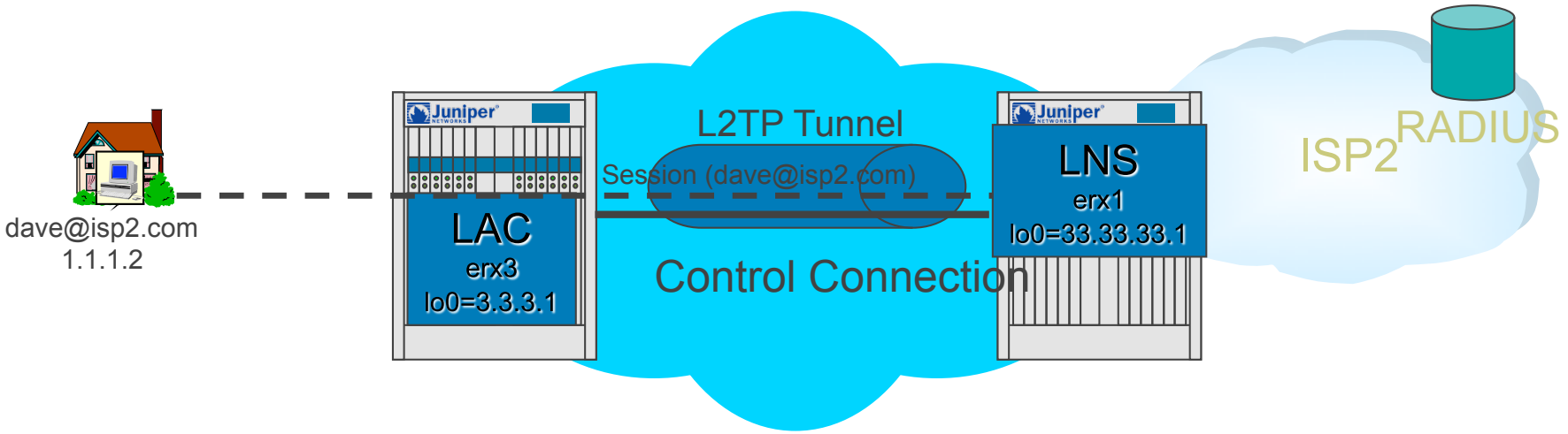
Start Control Connection Connected (SCCCN)

Zero-Length Body (ZLB ACK)

Hello

Hello

L2TP SESSION ESTABLISHMENT



Incoming Call Request (ICRQ)

Incoming Call Reply (ICRP)

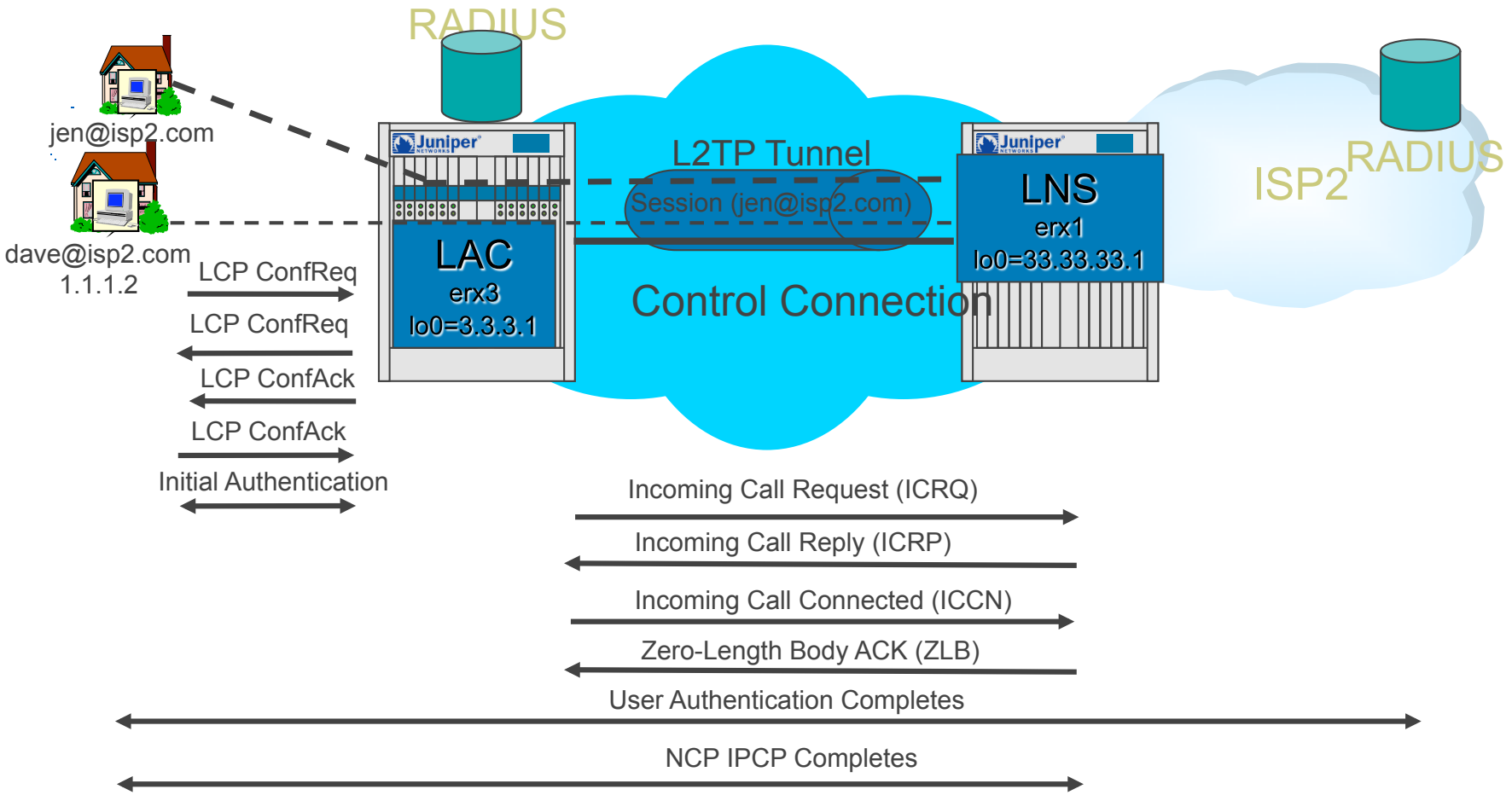
Incoming Call Connected (ICCN)

Zero-Length Body ACK (ZLB)

User Authentication Completes

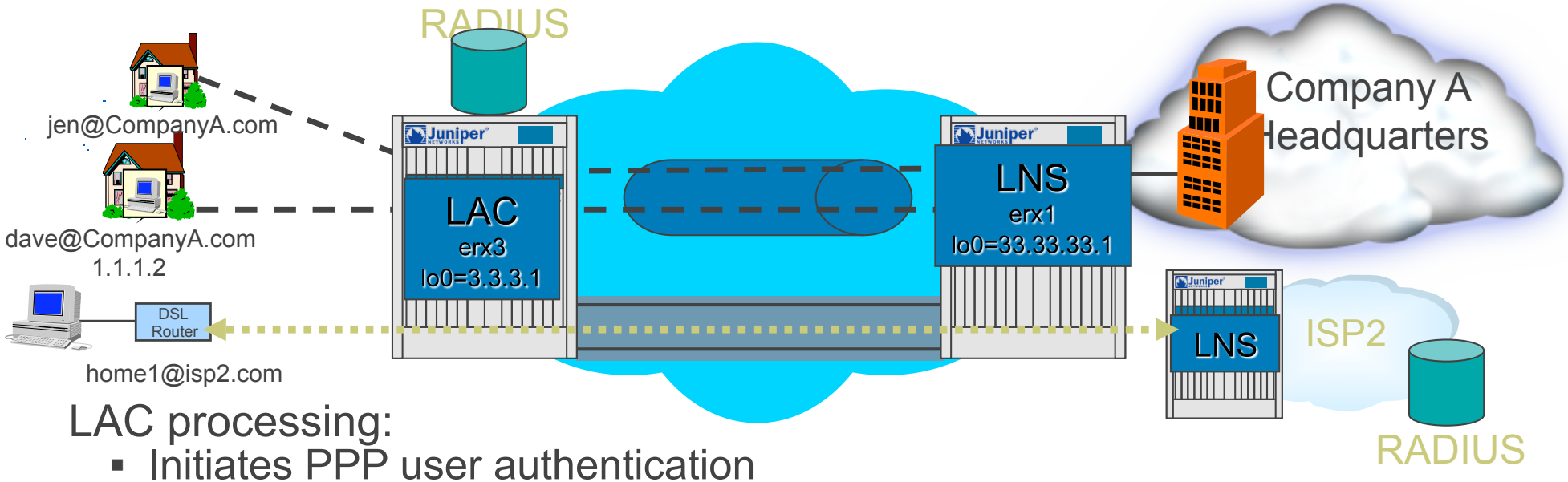
NCP IPCP Completes

ESTABLISHING ADDITIONAL L2TP SESSIONS



Open tunnel already exists—establish additional L2TP sessions

USER AUTHENTICATION



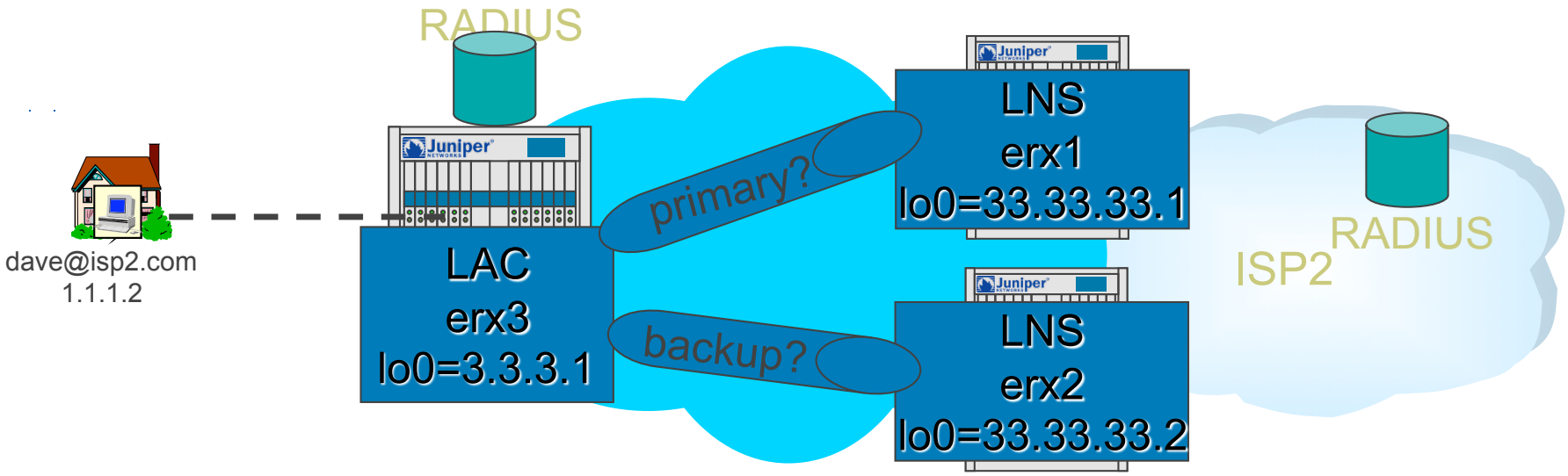
LAC processing:

- Initiates PPP user authentication
- Determines whether to tunnel or terminate PPP session; If tunnel, then which one.
- Sends user authentication response to the LNS

LNS processing:

- Receives user or proxy authentication response
- Determines if it should accept the response or restart the process
- By default, LNS does not use the proxy authentication data
- Proxy authentication allows LNS to accept the proxy data

L2TP TUNNEL SELECTION AT THE LAC



Domain map configuration options:

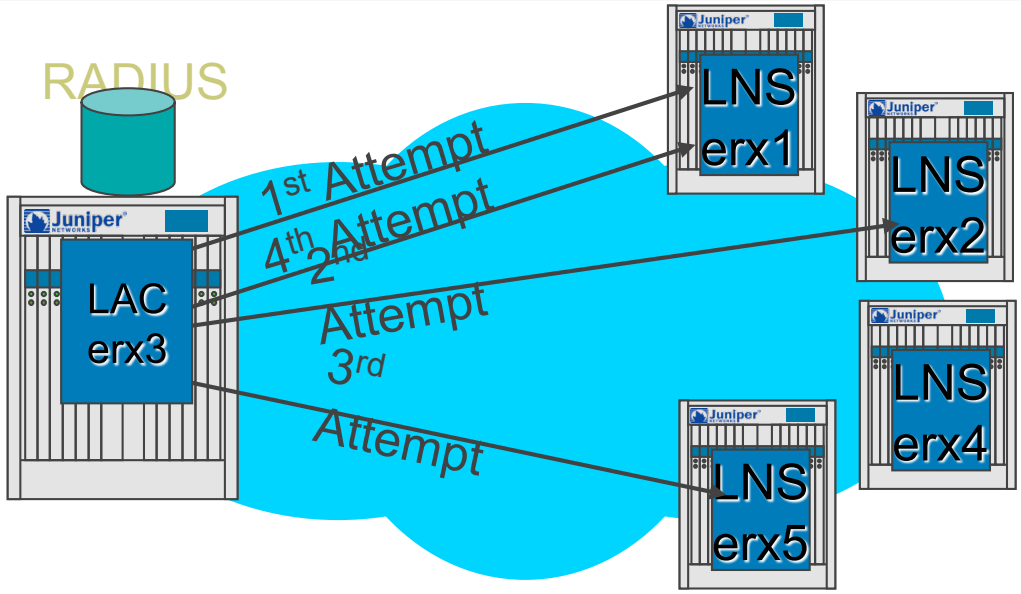
- 31 tunnel destinations per domain
- Up to 8 levels of preference
- Up to 8 destinations per preference level

Tunnel selection configuration options:

- Failover between preference levels
- Failover within a preference level
- Maximum number of sessions per tunnel
- Weighted load balancing

FAILOVER BETWEEN PREFERENCE LEVELS

Domain:	isp1.com	
Tunnel Tag	Server Name	Tunnel Preference
1	erx1	100
2	erx2	100
3	erx1	200
4	erx2	200
5	erx4	200
6	erx1	300
7	erx5	300



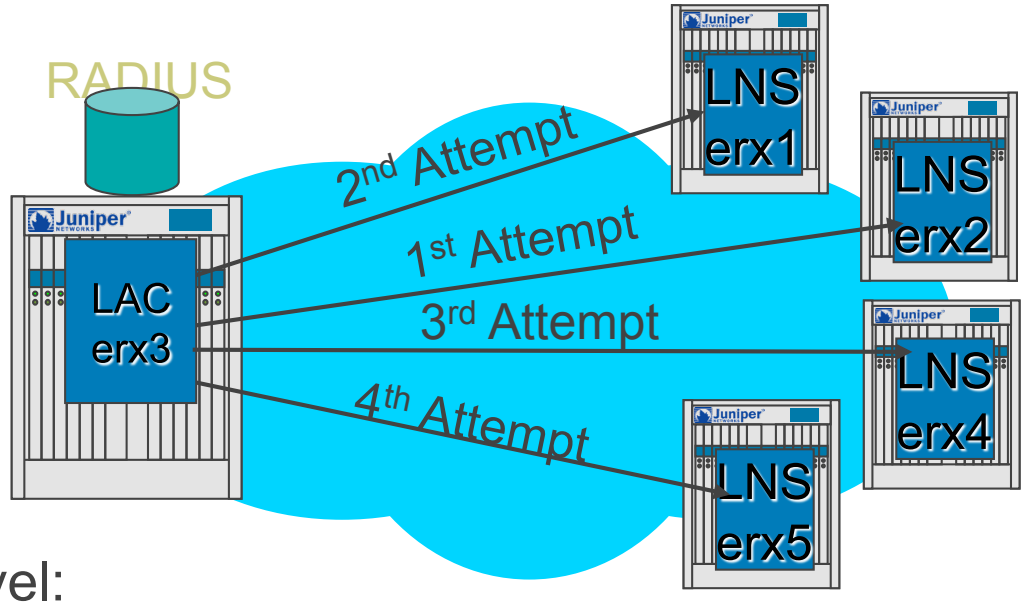
Failover between preference levels:

- Default behavior: LAC selects one destination based on the highest preference
  - The numerically lowest value has the highest precedence
  - Preference level configured in domain map or standard RADIUS attribute
- If preferences are the same, LAC randomly selects one from the group
- If connection attempt fails for that preference level:
  - Destination for all preferences marked unreachable and not retried for 5 minutes
  - No other destinations tried at that same preference level
  - Router selects another reachable destination from the next lower preference level
  - If all destinations marked unreachable, the LAC tries the destination that failed first

FAILOVER WITHIN A PREFERENCE LEVEL

Domain: isp1.com

Tunnel Tag	Server Name	Tunnel Preference
1	erx1	100
2	erx2	100
3	erx4	200
4	erx5	300



Failover within a preference level:

- LAC selects one destination based on the highest preference
- If preferences are the same, LAC randomly selects one destination
- If connection attempt fails for that destination level:
  - Destination marked unreachable and not retried for 5 minutes
  - LAC selects another reachable destination from the same preference level
  - If all destinations at a preference level unreachable, LAC moves to the next lower preference level and repeats the process
  - If all destinations marked unreachable, LAC tries the destination that failed first

USER IDENTIFICATION AND CONTRACT ENFORCEMENT ON LAC

THE PROBLEM STATEMENT

User identification by PPP CHAP/PAP username

- Usually realm used to differentiate retailers by infrastructure owners.
- Know to retailer (alternative ISP) but not by infrastructure.
- How to enforce rates and policies on LAC?
  - Access to retailer DB - Different DB structure.
  - RADIUS proxy. - VSA translation and manipulation.
  - Performance cost (CPS)
- How to avoid frauds (e.g. dual use) by LAC and by LNS.

User identification by PPP IA/Option 82

- Controlled by infrastructure owner but exposed to retailer.
- How to fill retailer DB and link IA to user profile.
- How to avoid frauds (e.g. dual use) by LAC.

User identification by L2 attributes.

- Know to infrastructure owner but not visible to retailer.
- How to link with user profile form retailer DB.

EXPERIENCE – INFRASTRUCTURE PROVIDER

ISP do not agree to share DB, nor use RADIUS proxy.

Use PPP IA to avoid frauds (user allowed to login only from his line/home.

To control or not to control profile (rate) on LAC?

- Why?
  - "Trust but …" – What if retailer gives to customer more the it should?
- Develop multiple domains/realms
  - e.g. 2M_M5.du.ae, 4M_1M.du.ae, …
  - Many options/variants
  - Requires provisioning on infrastructure (LAC or RADIUS) whenever retailer change/expand offer.
  - **Requires CPE configuration change !!!**
- Found too complex operationally.

EXPERIENCE – INFRASTRUCTURE PROVIDER

The infrastructure provider do not differentiate users

- Unlimited bandwidth form LAC PoV
- No DiffServ scheduling on subscriber LAC interface
- Only good thing – SIMPLE

EXPERIENCE – RETAILER PROVIDER (LNS)

Agreed to use agent circuit ID (PPP IA) from LAC in L2TP calling-number AVP, and send to RADIUS as calling-station-id.

- In order to ensure that user log-in from it's home.
- In order to avoid frauds (stolen identity).
- Same procedure used for own PTA customers – unification.

But access network in Infrastructure provider insert something like:
`"GigabitEthernet 13/0/2.380306:38-306#587203013#Agg-SDP-C7609-3/9##pppoe 00:30:0a:84:f5:0f#"`

- Any switch on path in aggregation network append his string to IA string (89B in this case)
- Retailer system can process only 60B.
  - It was sufficient for own PTA
  - Requires substantial changes on BSS system.

BSS enhancement project start, but…

TTM force to launch service without. (TRA involved)

IDENTIFICATION TAKE A WAY

Most probably your partner provider will not give you access to user database or even RADIUS (for proxy).

The remote agent ID and remote circuit ID presented by infrastructure provider may have wired format (for you).

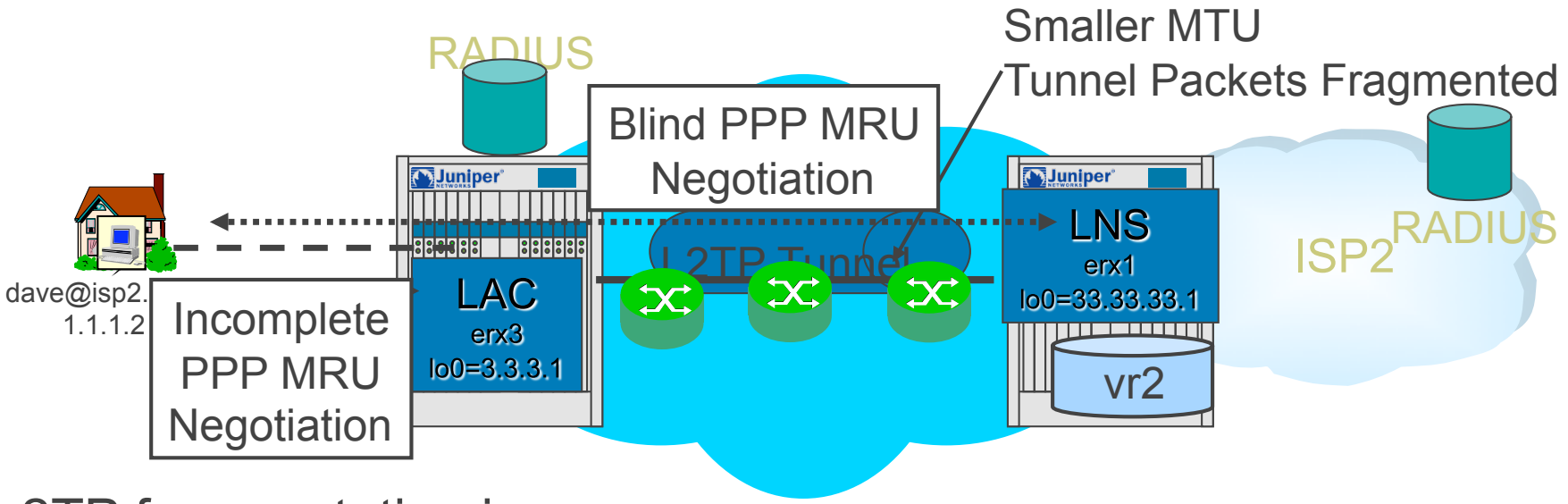
Encoding service characteristic in realm/domain is non-practical, if retailer provides CPE in routed mode. (because service change requires re-programming of CPEs)

Make an audit against required changes is OSS/BSS systems.

- Applicable for infrastructure provider –policy/rate enforcement on LAC
- Applicable for retailer provider – no L2 information.
- Necessary changes
- Cost
- Time
- Performance impact

ISSUES WITH MTU

FRAGMENTATION AND IP REASSEMBLY



L2TP fragmentation issue:

- PPP MRU negotiated between user and LAC or LNS
- No knowledge of links between
- Packets encapsulated at LAC, forwarded into network and routed
- Tunneled IP packets might be fragmented
- Fragmented tunneled packets must be reassembled at endpoint

FRAGMENTATION – REASSEMBLY. SO, WHAT IS A PROBLEM?

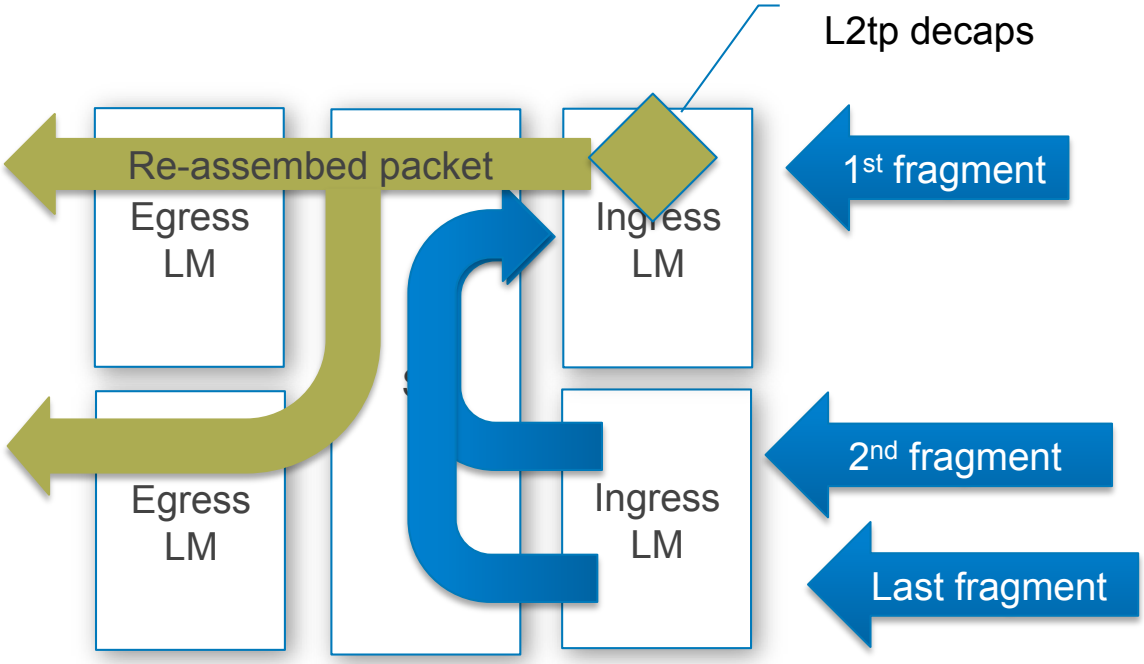
How much traffic is affected?

- A lot 30% of packet generated in internet is 1500B (other popular is 40B and 512B)
- Overhead

Reassembly is complex and costly process

- Requires packet modification in fly BEFORE lookup.
- IP filters and firewalls are not friends of fragmented packets – lack of L4 information in non-first fragments
- Reassembly on distributed systems (e.g. ERX, MX) requires anchor point.
  - Fragment may arrive on different line cards.
  - Extra hardware or Suboptimal use of line modules

REASSEMBLING OF FRAGMENTED L2TP PACKETS IN DISTRIBUTED SYSTEMS



AVOIDING FRAGMENTATION

L2TP fragmentation avoidance:

- Configure an MRU lower than the minimum size between LAC/LNS
- MRU = (Min MTU between LAC/LNS) – (L2TP UDP/IP) – (Max L2TP header)
- Ethernet example:

Minimum link MTU 1500	
L2TP IP header	- 20
L2TP UDP header	- 8
L2TP header	-6
PPP header	-4
MRU size to specify	1462

On LAC:

- Configure the MTU on the access link to take park in initial LCP negotiation

And on LNS:

- Configure PPP MRU within the profile referenced in the L2TP remote host definition

Then CPE and LNS should agree on PPP MRU and calculate IP MTU. BUT …

MRU NEGOTIATION IS GOOD BUT…

Many CPEs ignores LCP MRU negotiation! Be aware.

Are calculation on previous slide was OK? What if:

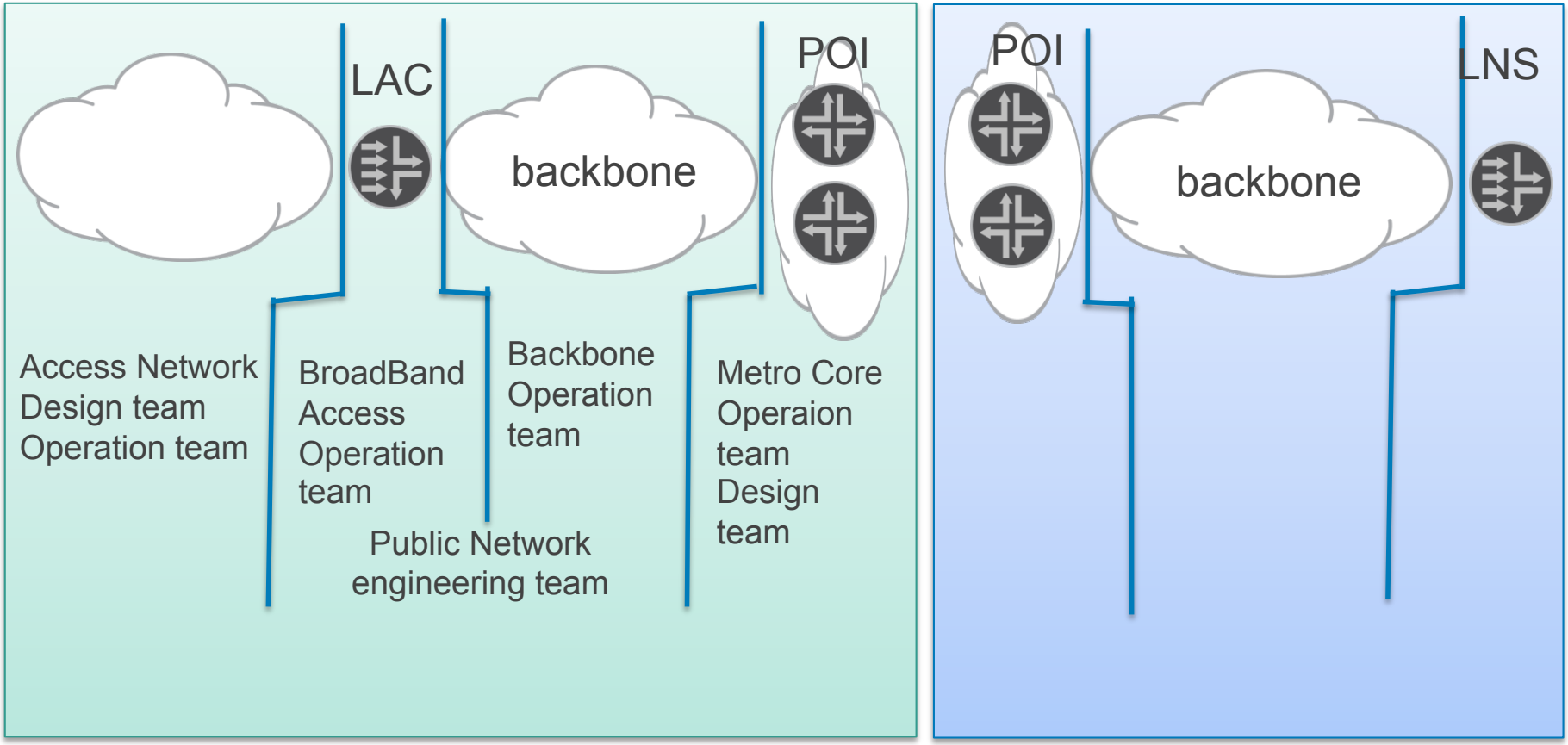
- MPLS is somewhere between LAC and LNS
  - MPLS VPN
  - MPLS FRR (facility backup)
- Network between reroute traffic on another path by IGP (modification, new links, fails)

There is no better way then ensure you know max IP MTU you may send between LAC and LNS without fragmentation.

- As all parties contribute in L2TP bit stream design

The best will be MTU allowing 1500B subscriber traffic to be send all the way w/o fragmentation.

EXPERIENCE:



6 teams, 2 provides, all need to work together to ensure minimum MTU!

CORRECT PPP MRU – NO FRAGMENTATION

Let assume MRU is set to 1454B – 2 MPLS labels considered

L2TP packets are not fragmented – no reassembly on LAC/LNS

BUT…

Subscriber PC is not aware about that – sends 1500B

- CPE has to perform fragmentation

Portal is not aware about this

- LNS has to perform fragmentation

So what?

We observed that some sites/pages/objects are not displayed!!!

FRAGMENTATION OF SUBSCRIBER PACKETS

Performance

- CPU usage
- Overhead
- Small issue nowadays

Some sites (portal) sends packet with DF flag

- God knows why, maybe to compete on performance
- LNS need to ignore DF flag – breaks IETF IS, or
- LNS need to manipulates TCP MSS in in transit traffic. What about non-TCP?
- Otherwise packet is dropped – service/object is not avaliable

Many sites are protected by firewalls

- Non-first fragment has no L4 information – how to classify it to flow?
  - Drop fragmented packets on FW
  - Reassembly on firewall – FW performance issue

AVOID FRAGMENTATION OF SUBSCRIBER PACKETS

Ensure LAC to LNS can handle L2TP packet of 1538B w/o fragmentation in any conditions.

- If MPLS is in place media MTU need to cover also MPLS headers – 1550B is reasonable.
- Well 1550B is not an issue for POS or ATM. For Ethernet is jumbo frame.
- If tuning of Eth MTU needed, why not set to MAX (~9kB).

RULE of TUMB:

For any interface faster ten 10M, always set media MTU as big as possible.

ML-PPP OVER L2TP AND QOS

COLT

ML-PPP IN BITSTREAM

To offerhigher speed then sifle DSL line of infrastructure SP can provide

- Business access
- Service differentiator

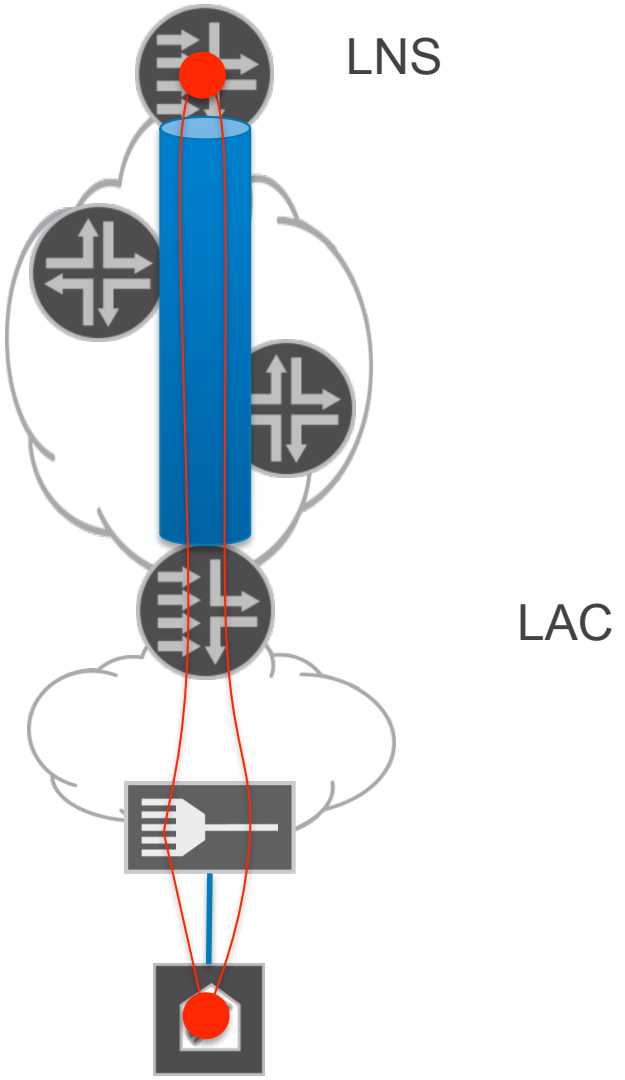
For redundancy

- Same LNS and aggregation network.
- Separate Cu links

QoS is needed for business access.

ML-PPP AND L2TP

It is simple

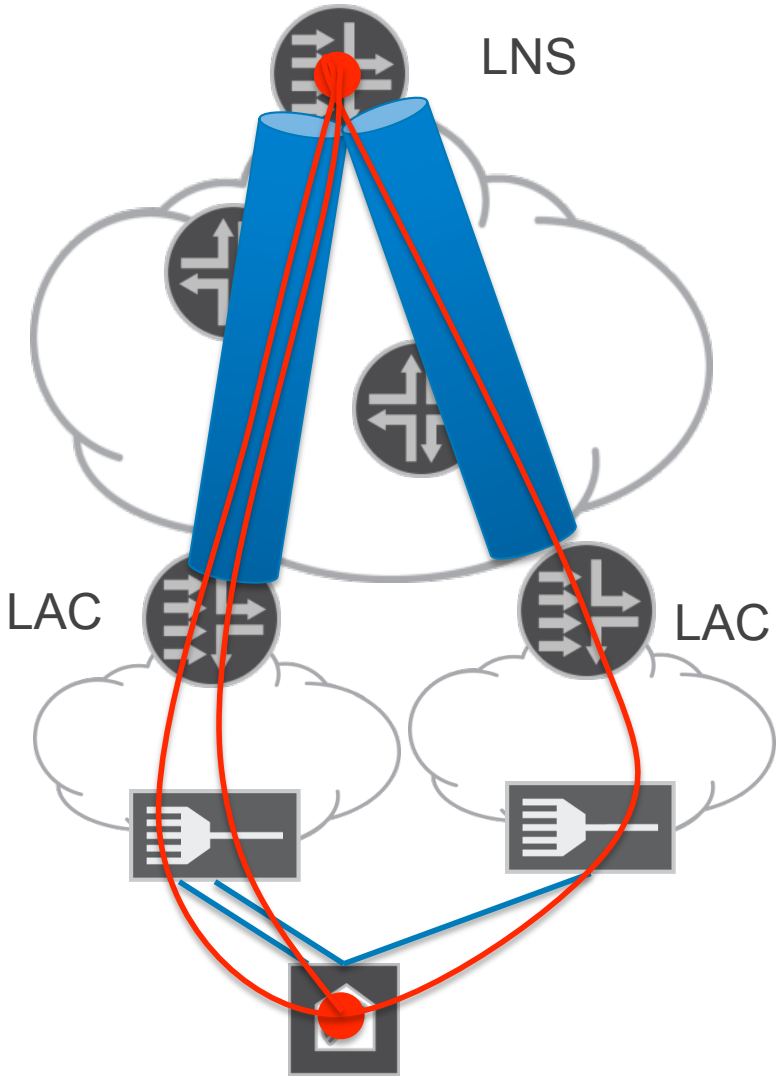


ML-PPP AND L2TP

It is simple

Not really

- Each member ppp session can span two LACs
- Each member ppp session can span two DSLAM
- Each session span two DSL lines
- The only bundling points are LNS and CPE
- LAC may not be aware about subscriber – not aware about rate it should enforce.



QOS ON ML-PPP OVER L2TP

Task

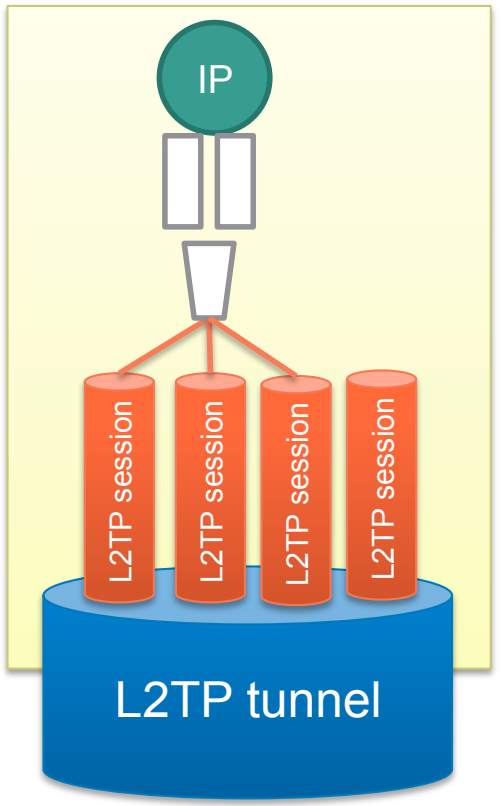
- Shape customer to 3Mbps
- 40% for EF (teal-time, voice)
- Rest for internet

Connectivity

- 3 DSL lines, 1Mbps each

Approach 1

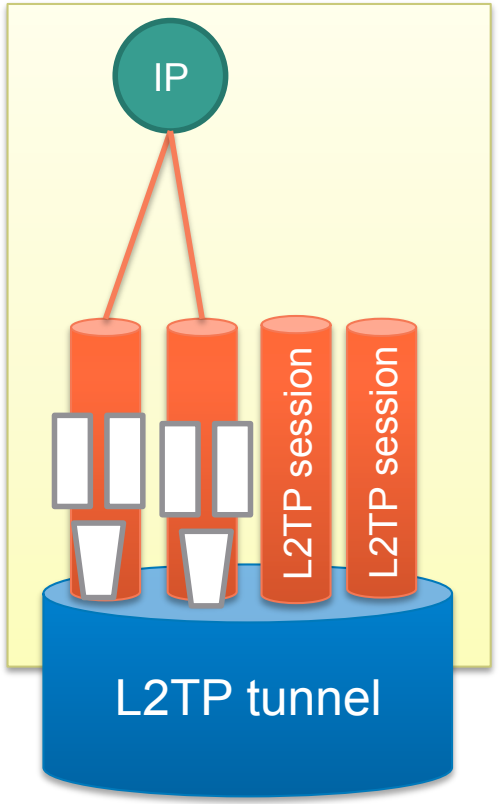
- Apply QoS to bundle (JUNOSe speaking - IP node)
- Shaped to 3M, 40% (1.2M) for EF
- What if one session is lost?
  - 3Mbps is push down to DSLAMs
  - DSLAM(s) has only 2x1Mbps
  - QoS depends on DLAM implementation
  - QoS is out of retailer SP control.



QOS ON ML-PPP OVER L2TP

Approach 2

- Apply QoS to each ppp member session
- Shaped to 1/3 of 3M, 40% (1.2M) for EF
- What if one session is lost?
  - 2Mbps is push down to DSLAMs
  - DSLAM(s) has only 2x1Mbps – no congestion
  - QoS depends on LNS implementation
  - QoS is under retailer SP control.
  - EF is guaranteed to only 0.8Mbps




everywhere