Université de Liège

UNIVERSITÉ DE STRASBOURG

November 17, 2015

MPLS Under the Microscope: Revealing Actual Transit Path Diversity

Yves VANAUBEL

Pascal MÉRINDOL

Jean-Jacques PANSIOT

Benoit DONNET

Plane

Agenda

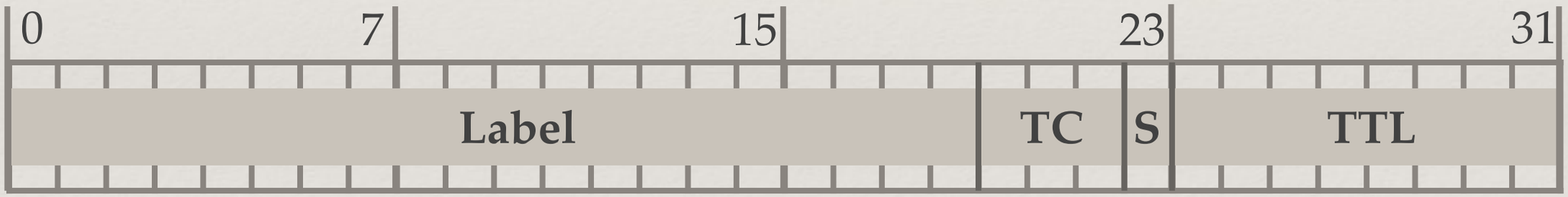
- MPLS background
- Revealing MPLS tunnels
- MPLS Tunnels Diversity
- Conclusions

Agenda

- **MPLS Background**
  - Label Stack Entries
  - MPLS Network
  - Label Distribution
- Revealing MPLS tunnels
- MPLS Tunnels Diversity
- Conclusions

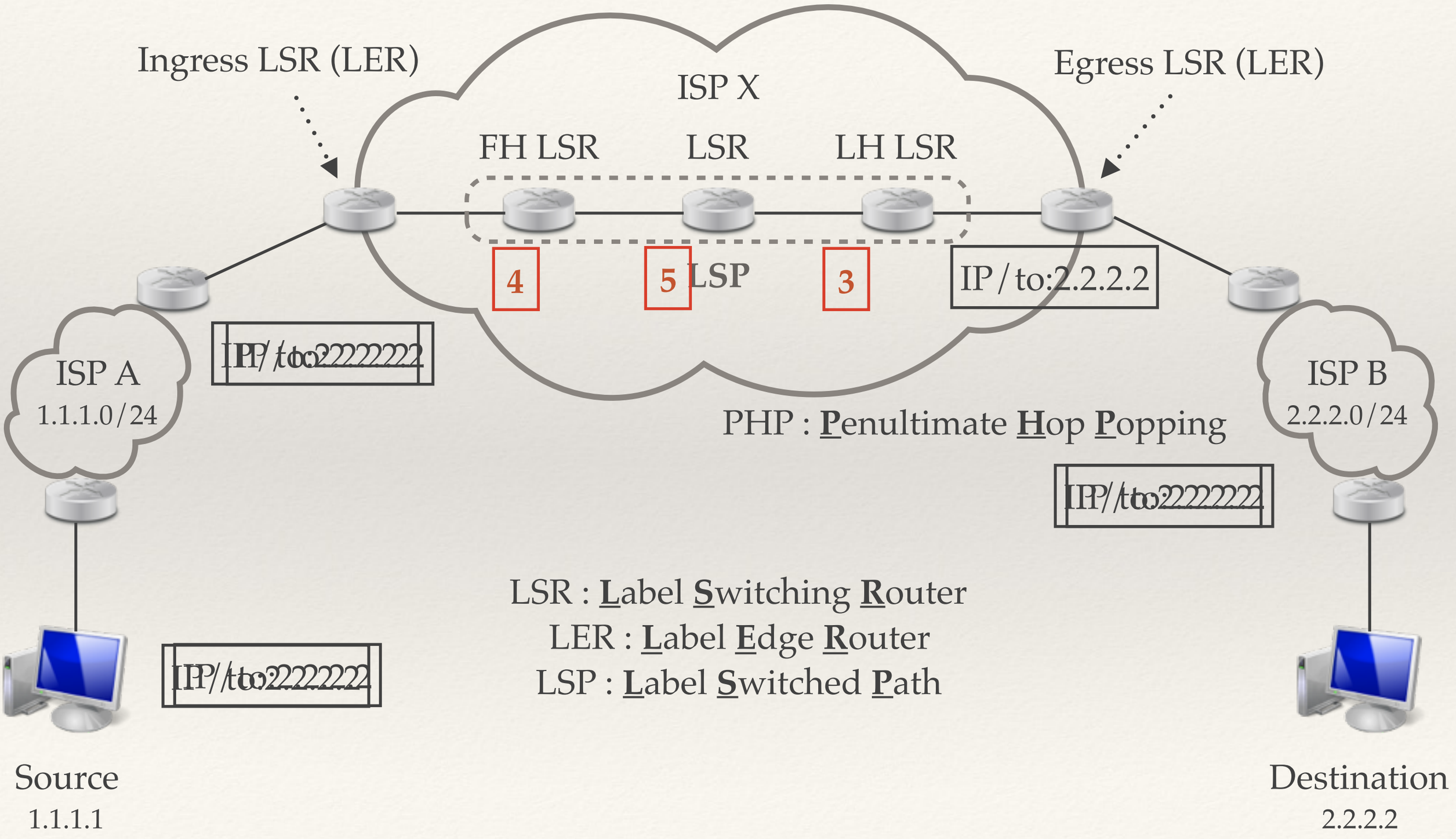
MPLS Label Stack Entries

- **L**abel **S**tack **E**ntries (LSE) :
  - 32 bits
  - Inserted between the MAC and the IP layer



- Label : Label value, 20 bits
- TC: **T**raffic **C**lass field, 3
- S: Bottom of stack, 1 bit
- TTL: **T**ime **T**o **L**ive, 8

MPLS Network



Label Distribution

- **L**abel **D**istribution **P**rotocol (LDP) [RFC5036]
  - Allows distribution FEC-to-label bindings among LSRs
  - In this case, FECs are prefixes in IGP routing tables
  - Downstream operation
  - Messages follow IP route
- **R**esource Re**S**er**V**ation **P**rotocol - **T**raffic **E**ngineering (RSVP-TE) [RFC3209]
  - Allows resource reservation on the path
  - **E**xplicit **R**oute **O**bject (ERO) extension : Allows the source to pre-calculate the LSP (not necessary the IP route)

Agenda

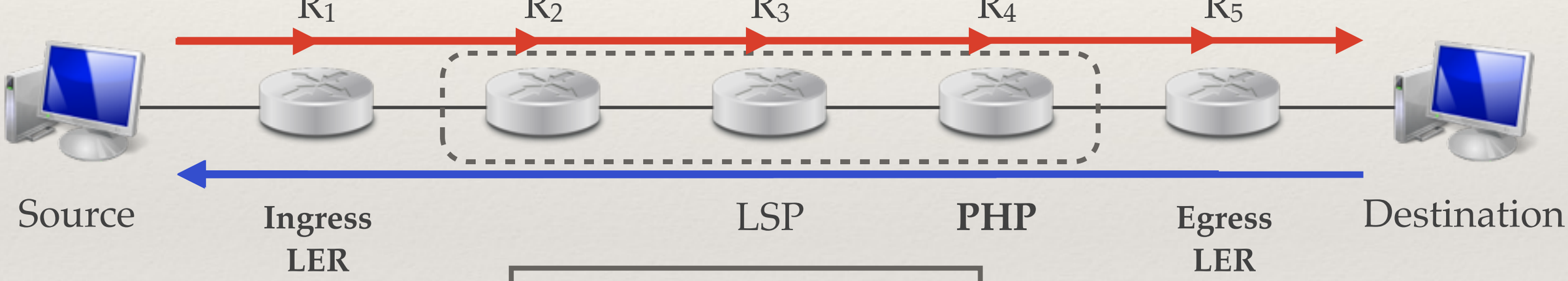
- MPLS Background
- **Revealing MPLS tunnels**
  - Measurement Technique
  - Explicit MPLS Tunnels
- MPLS Tunnels Diversity
- Conclusions

Measurement Technique

- MPLS tunnels discovery is based on standard active measurement tools (`traceroute`)
- Two options are required:
  - **ICMP extension** ([RFC4950]):
    - ✓ If an MPLS router must forge an ICMP *time exceeded* message, it should quote the MPLS LSE into it.
  - **TTL propagation** ([RFC3443]):
    - ✓ The ingress router of an MPLS tunnel should initialize the LSE-TTL with the value inside the IP-TTL field (iTTL).
    - ✓ The opposite operation is done by the egress LER (oTTL).

Explicit Tunnels

- The two options are activated
- This kind of tunnel is perfectly visible with `traceroute`



Traceroute output:

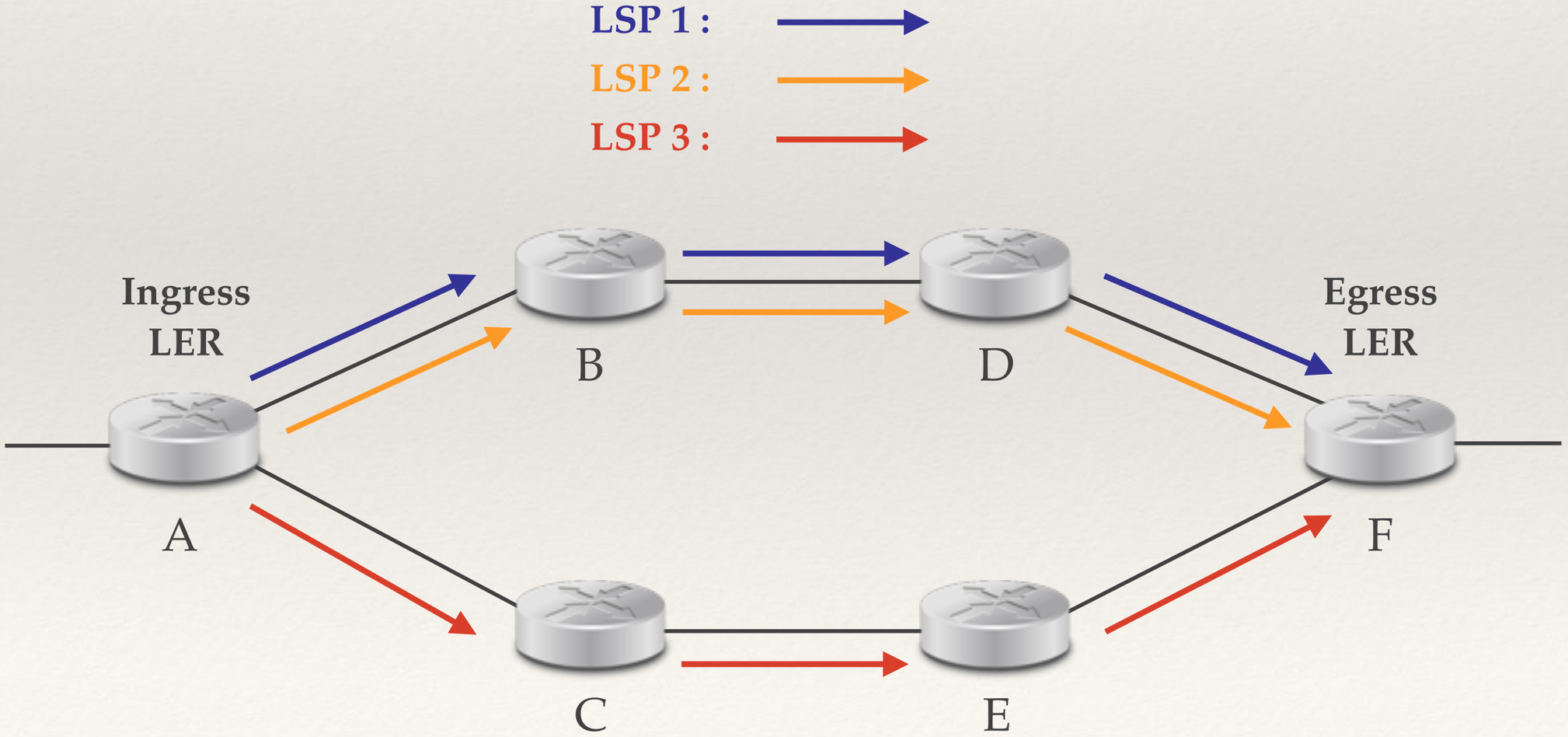
1. R_1
2. R_2 - *MPLS tag*
3. R_3 - *MPLS tag*
4. R_4 - *MPLS tag*
5. R_5
6. Destination

Agenda

- MPLS Background
- Revealing MPLS tunnels
- **MPLS Tunnels Diversity**
  - Motivations
  - Label Pattern Recognition Algorithm
  - Data Collection
  - Results
- Conclusions

Motivations

- Several LSPs may exist for a given <Ingress, Egress> LER pair



LPR

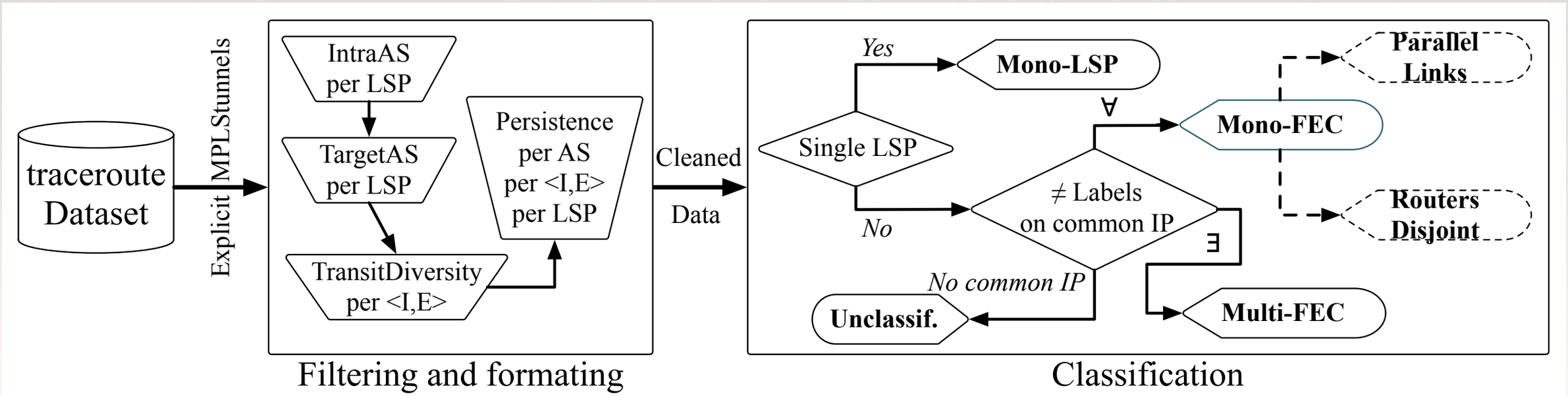
- **L**abel **P**attern **R**ecognition algorithm
- Allows to distinguish multi-FEC from IP load balancing in transit MPLS tunnels
- Passive classification method (offline)
  - Requires no additional probing than `traceroute`
- Recognizes behaviors of LDP vs. RSVP-TE based on MPLS labels distribution

LPR

- IOTP (**I**n-**O**ut **T**ransit **P**air): <Ingress, Egress> LER pair, i.e. set of explicit MPLS tunnels having the same IP entry and exit points
- LPR classifies each IOTP based on IP addresses and LSP labels
- Four output classes
  - **Mono-LSP**
  - **Multi-FEC**
  - **Mono-FEC** (ECMP)
  - **Unclassified**

LPR - Global View

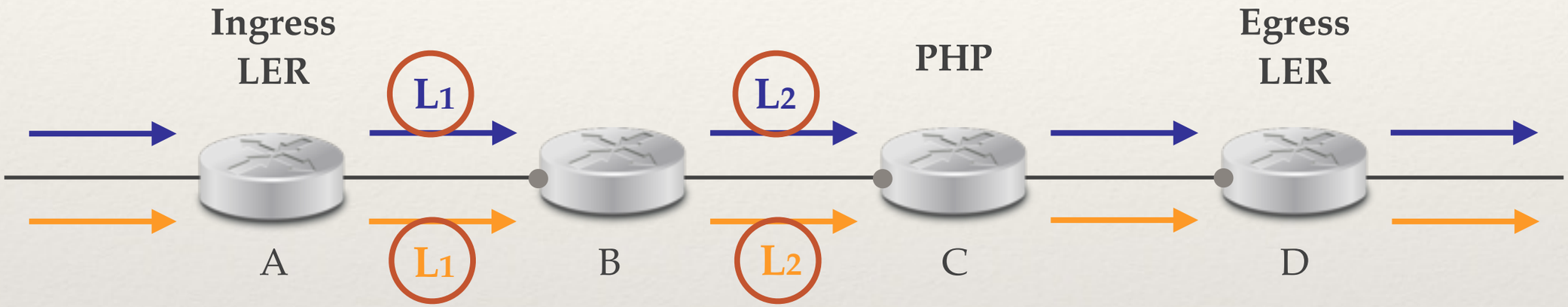
- Two main steps
  - Data filtering and formatting
  - Classification
- Per AS study



LPR - Classification

- Class 1: **Mono-LSP**

- means an interface



Trace LSP1:

1. A
2. B - Label L1
3. C - Label L2
4. D

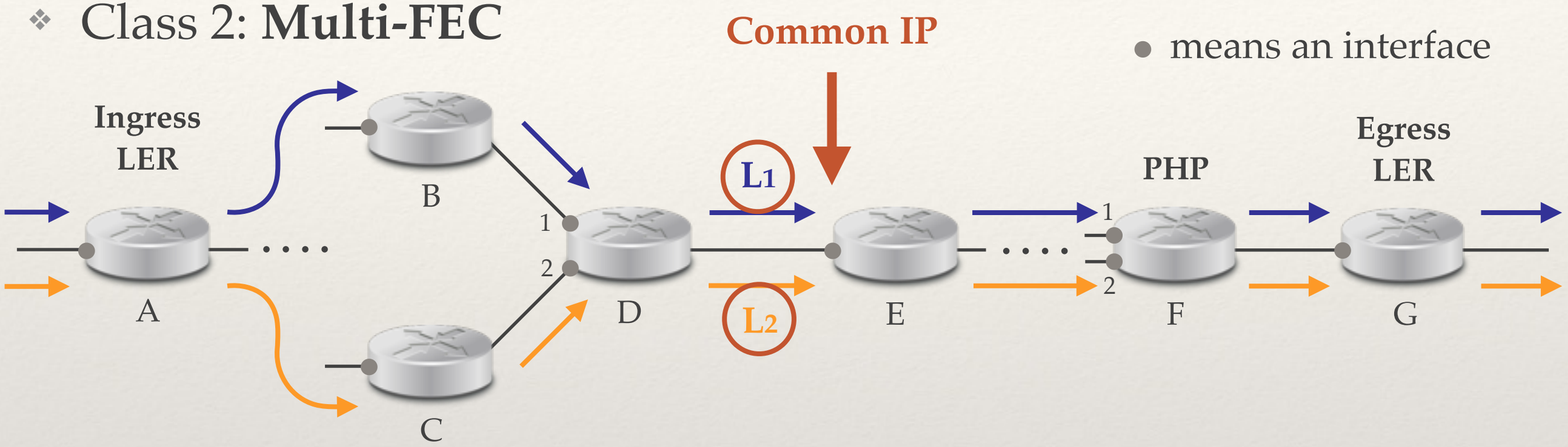
Trace LSP2:

1. A
2. B - Label L1
3. C - Label L2
4. D

Same IPs and same Labels

LPR - Classification

- Class 2: **Multi-FEC**



Trace LSP1:

1. A
2. …
3. B - Label
4. D1 - Label
5. E - Label L1
6. …
7. F1 - Label
8. G

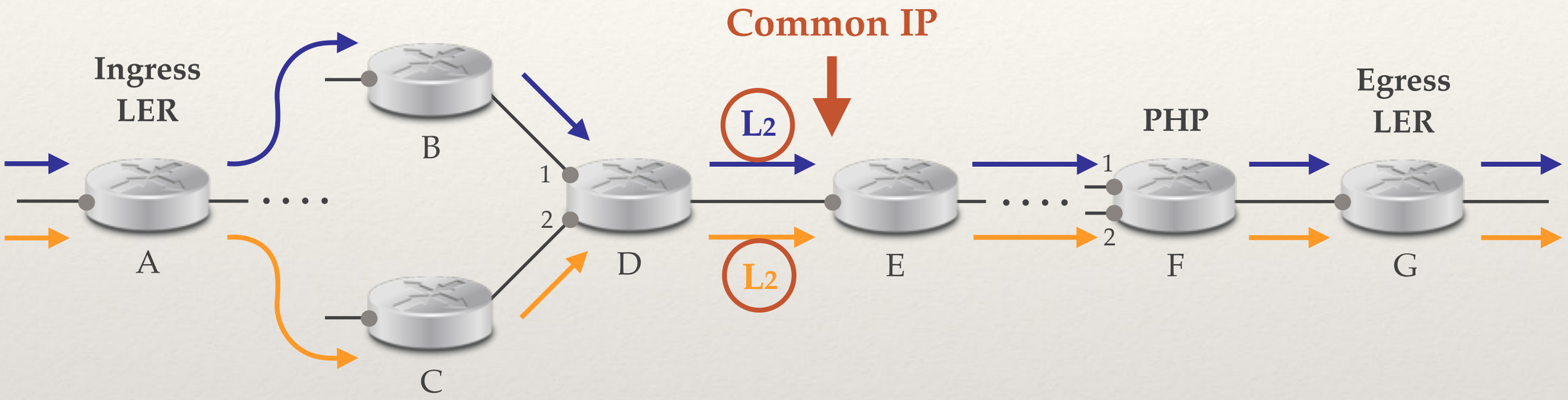
Trace LSP2:

1. A
2. …
3. C - Label
4. D2 - Label
5. E - Label L2
6. …
7. F2 - Label
8. G

Different labels for at least 1 common IP

LPR - Classification

- Class 3: **ECMP Mono-FEC (Disjoint Routers)**



Trace LSP1:

1. A
2. …
3. B - Label
4. D1 - Label L1
5. E - Label L2
6. …
7. F1 - Label
8. G

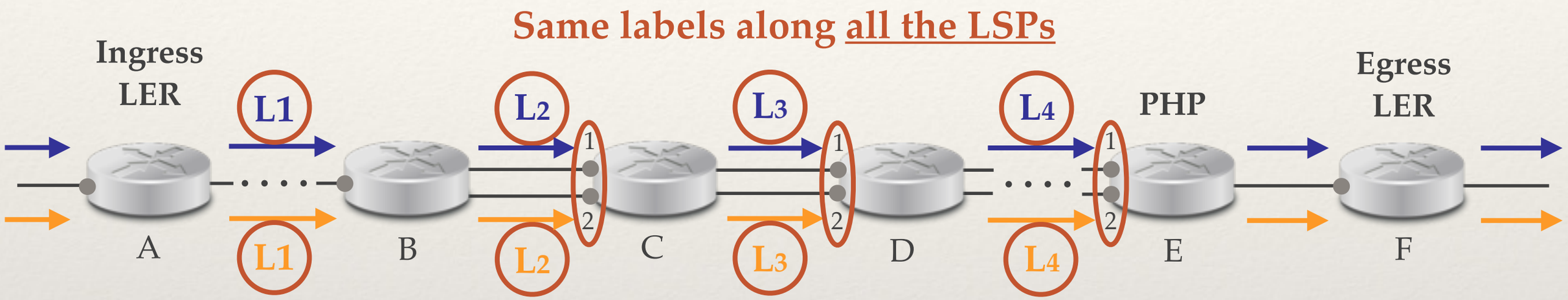
Trace LSP2:

1. A
2. …
3. C - Label
4. D2 - Label L1
5. E - Label L2
6. …
7. F2 - Label
8. G

Same label ∀ common IPs

LPR - Classification

- Class 3: **ECMP Mono-FEC (Parallel Links)**



Different IPs are aliases!

Trace LSP1:

1. A
2. …
3. B - Label L_1
4. C_1 - Label L_2
5. D_1 - Label L_3
6. …
7. E_1 - Label L_4
8. F

Trace LSP2:

1. A
2. …
3. B - Label L_1
4. C_2 - Label L_2
5. D_2 - Label L_3
6. …
7. E_2 - Label L_4
8. F

LPR - Classification

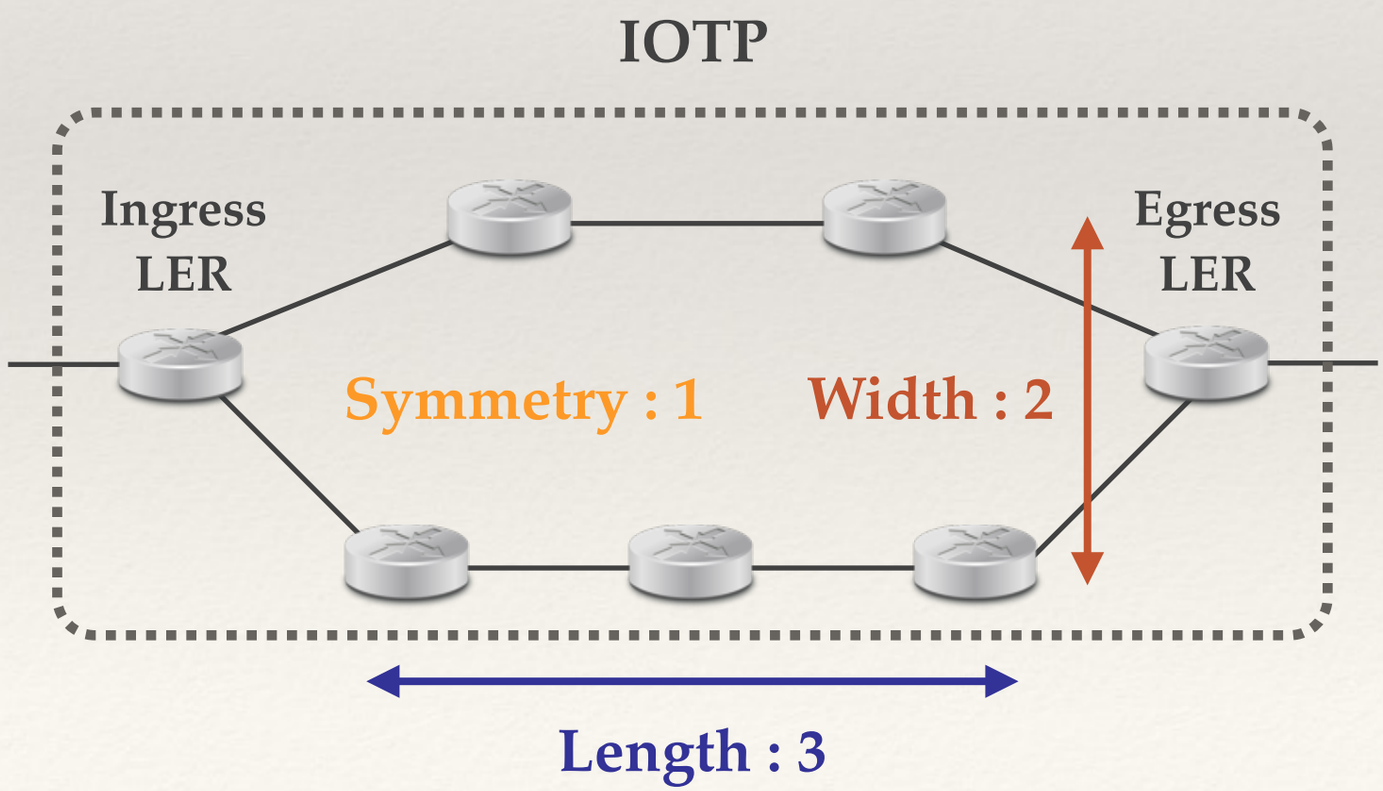
- Class 4: **Unclassified**
  - If PHP is used, the Egress LER does not exhibit any label
  - It may happen that LSPs do not intersect on a common IP address
  - In this case, the IOTP is arbitrarily tagged as **unclassified**

Data Collection

- *Archipelago* platform
  - More than 100 monitors scattered all around the world, divided into 3 teams
  - `paris-traceroute` to all routed /24 prefixes
- Work on data collected between January 2010 and December 2014
  - Cycle: First monthly run of each team
  - 60 cycles
- For each cycle:
  - IP2AS mapping using Routeviews data
  - MPLS explicit tunnels extraction

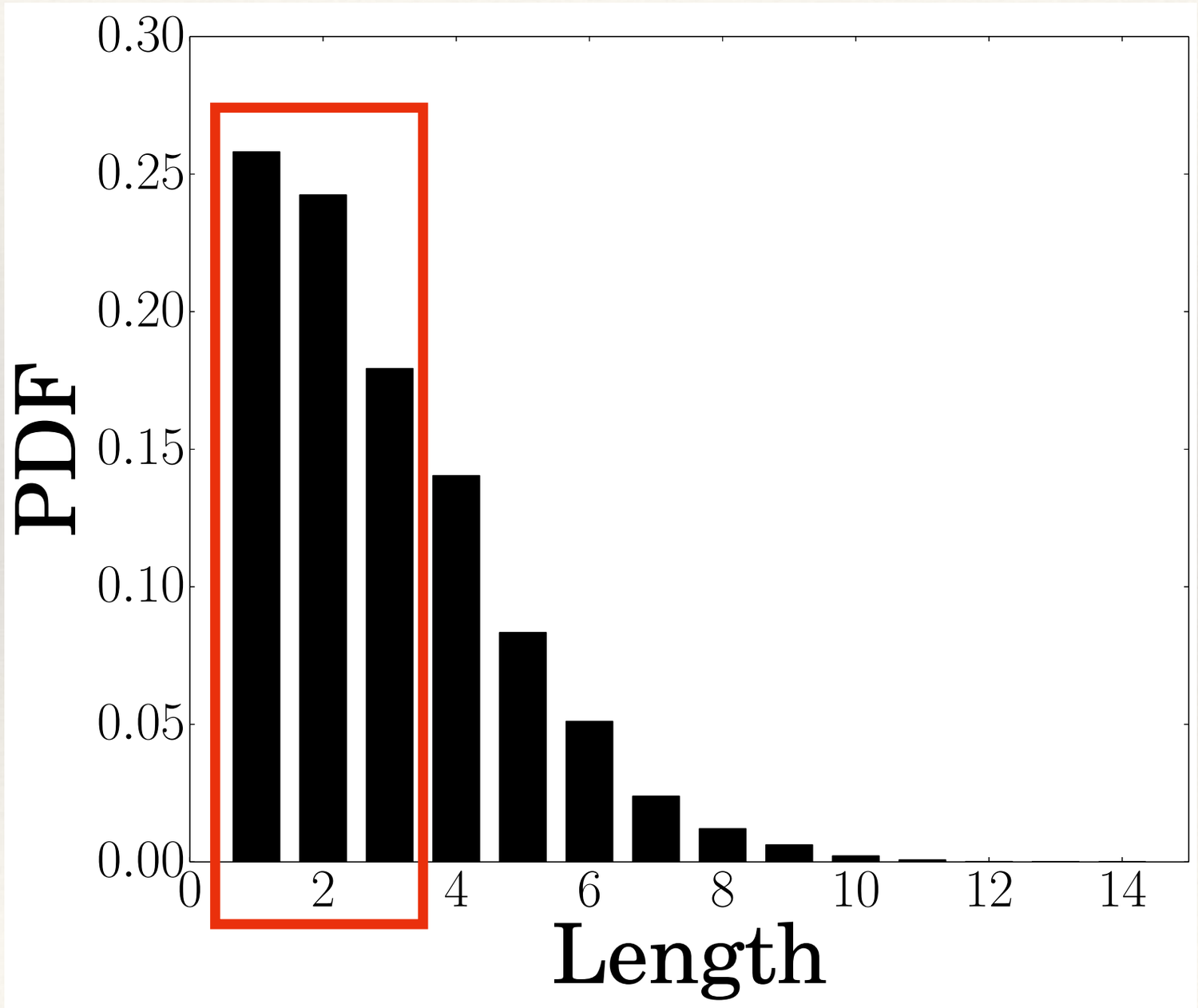
LPR - Results

- IOTP basic metrics
  - **Length** : Number of LSRs in the longest LSP
  - **Width** : Number of branches in the IOTP (logical or physical)
  - **Symmetry** : Difference between IOTP length and number of LSRs in the shortest LSP



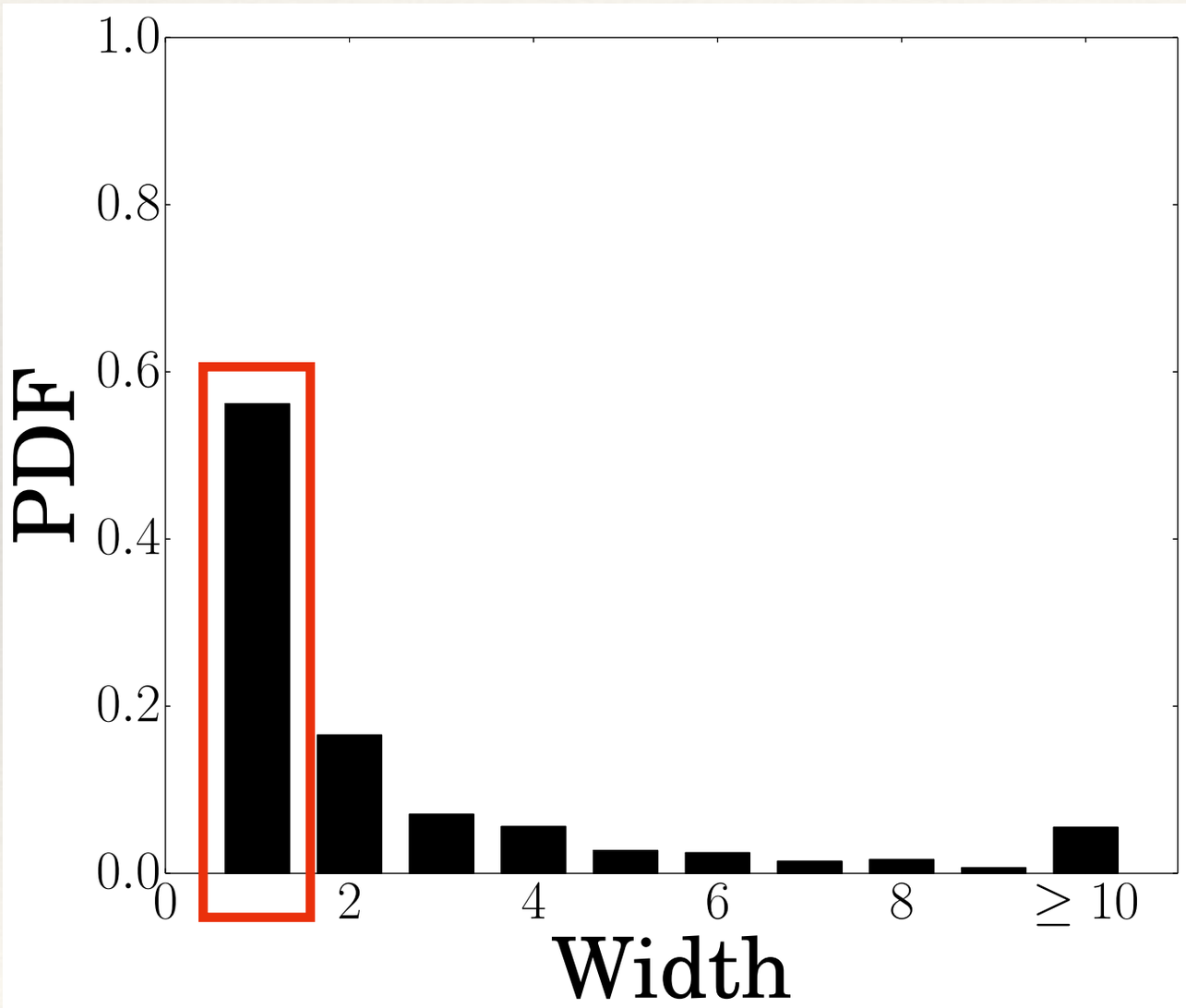
LPR - Results

- IOTPs length for December 2014

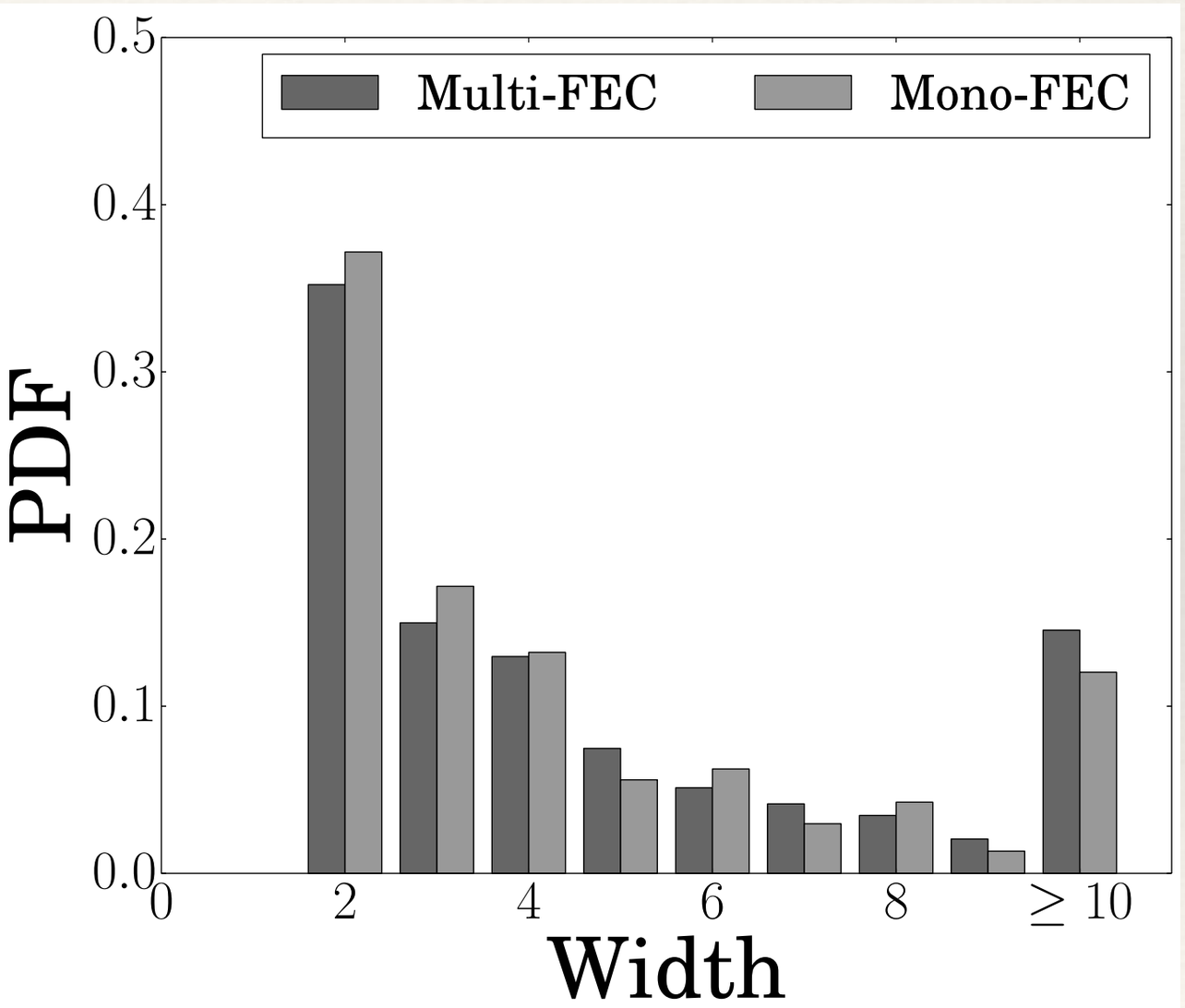


LPR - Results

- IOTPs width for December 2014



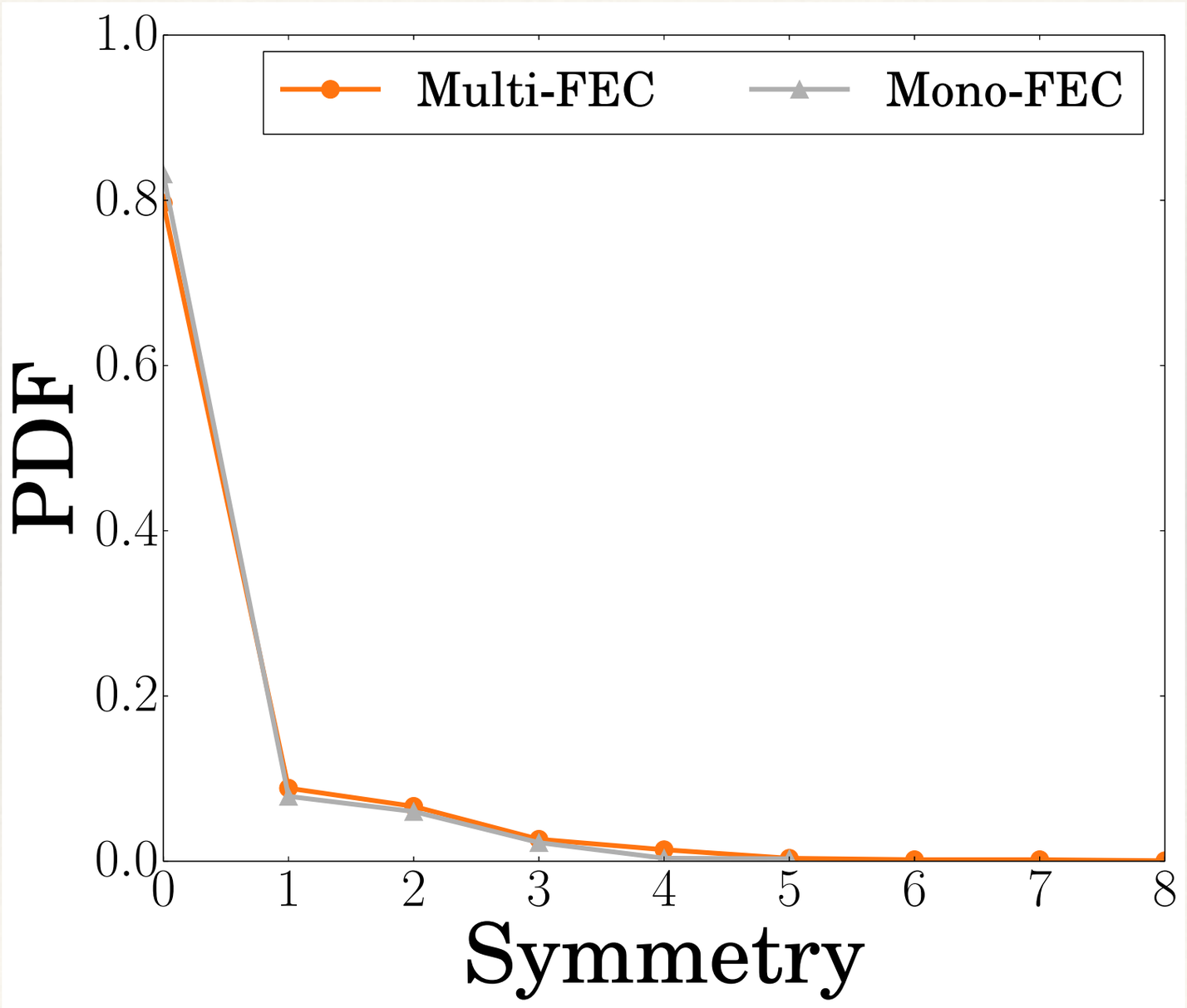
With Mono-LSP class



Without Mono-LSP class

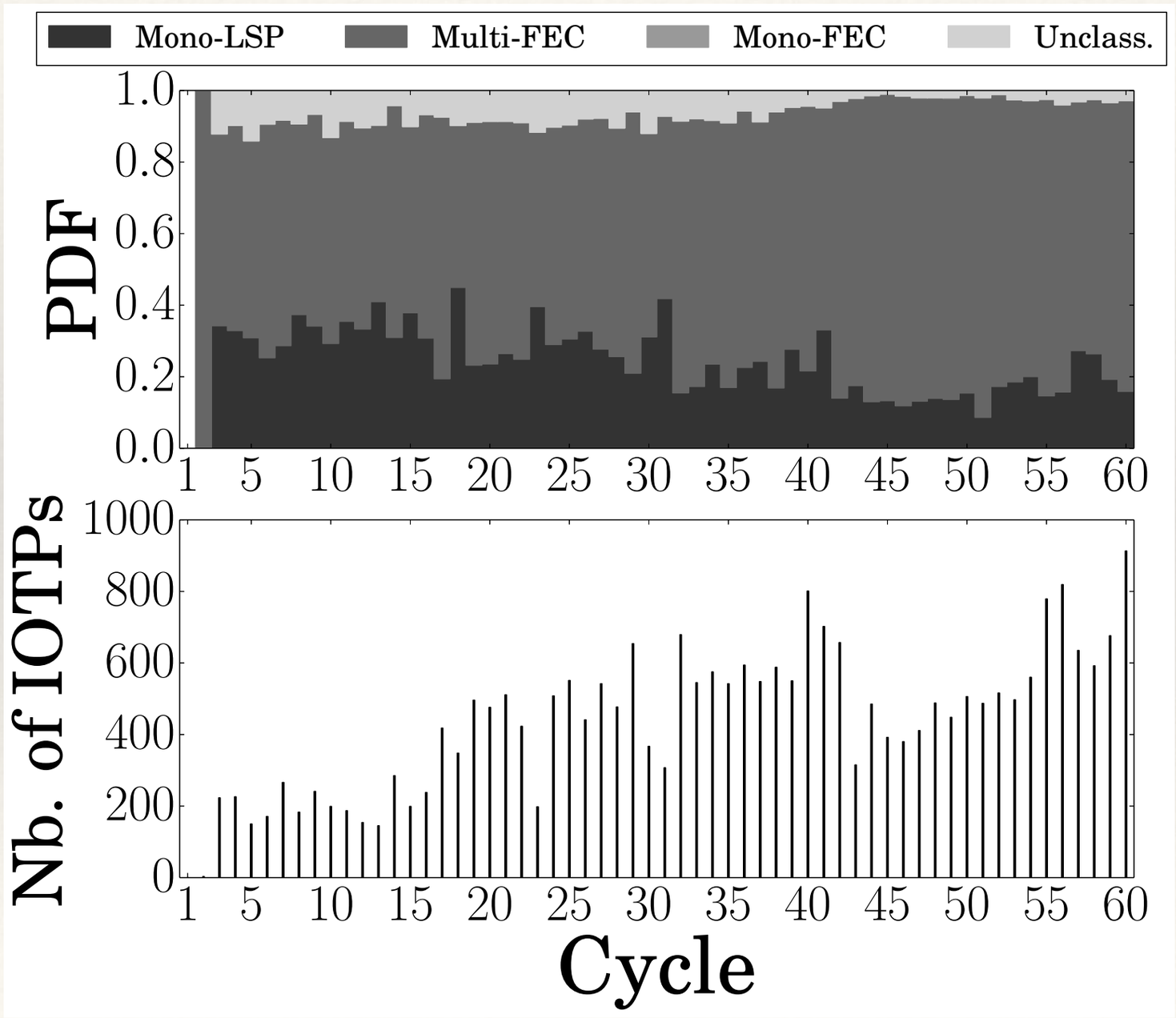
LPR - Results

- IOTPs symmetry for December 2014



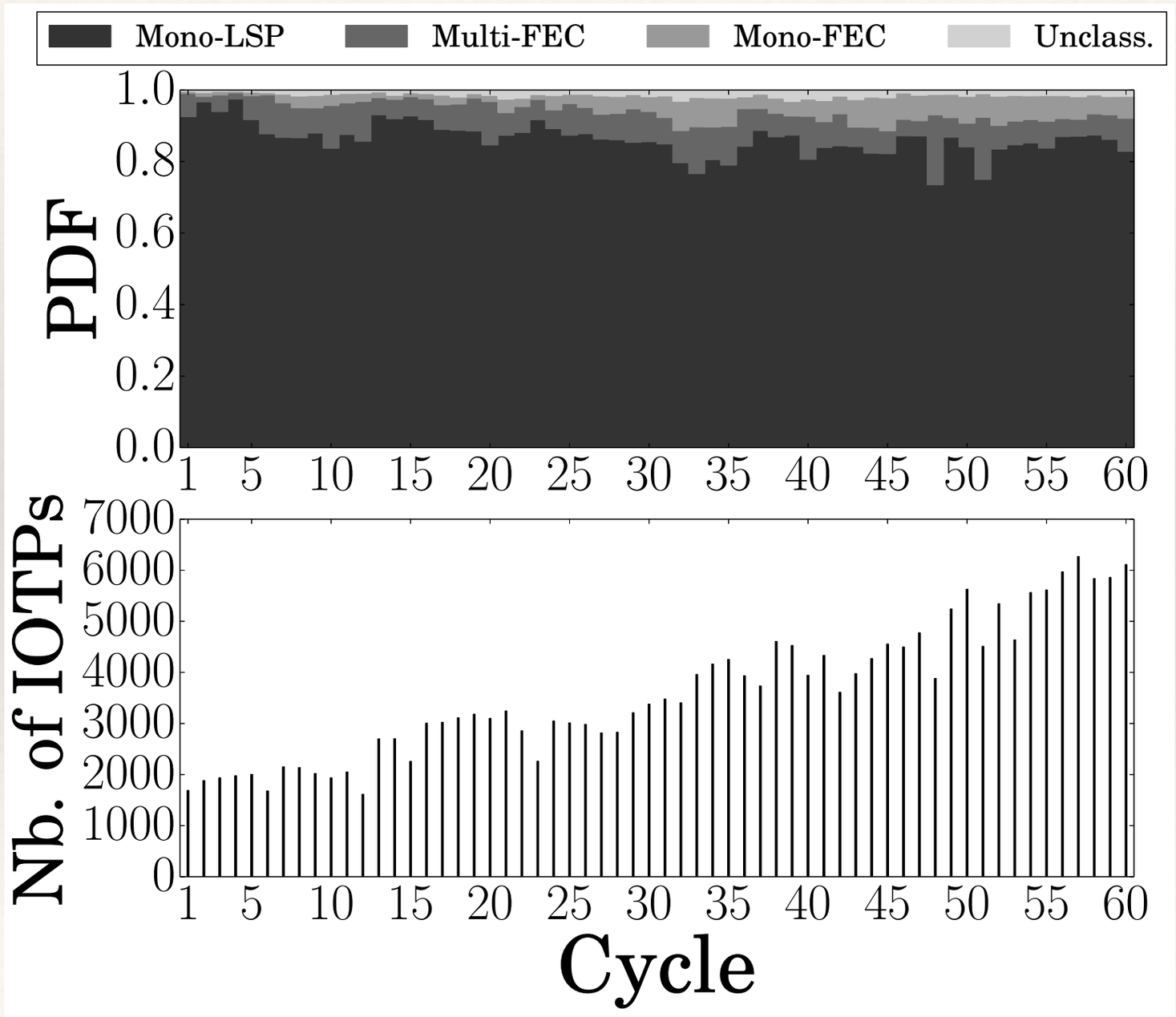
LPR - Results

- Classification for AS1273 (*Vodafone*)



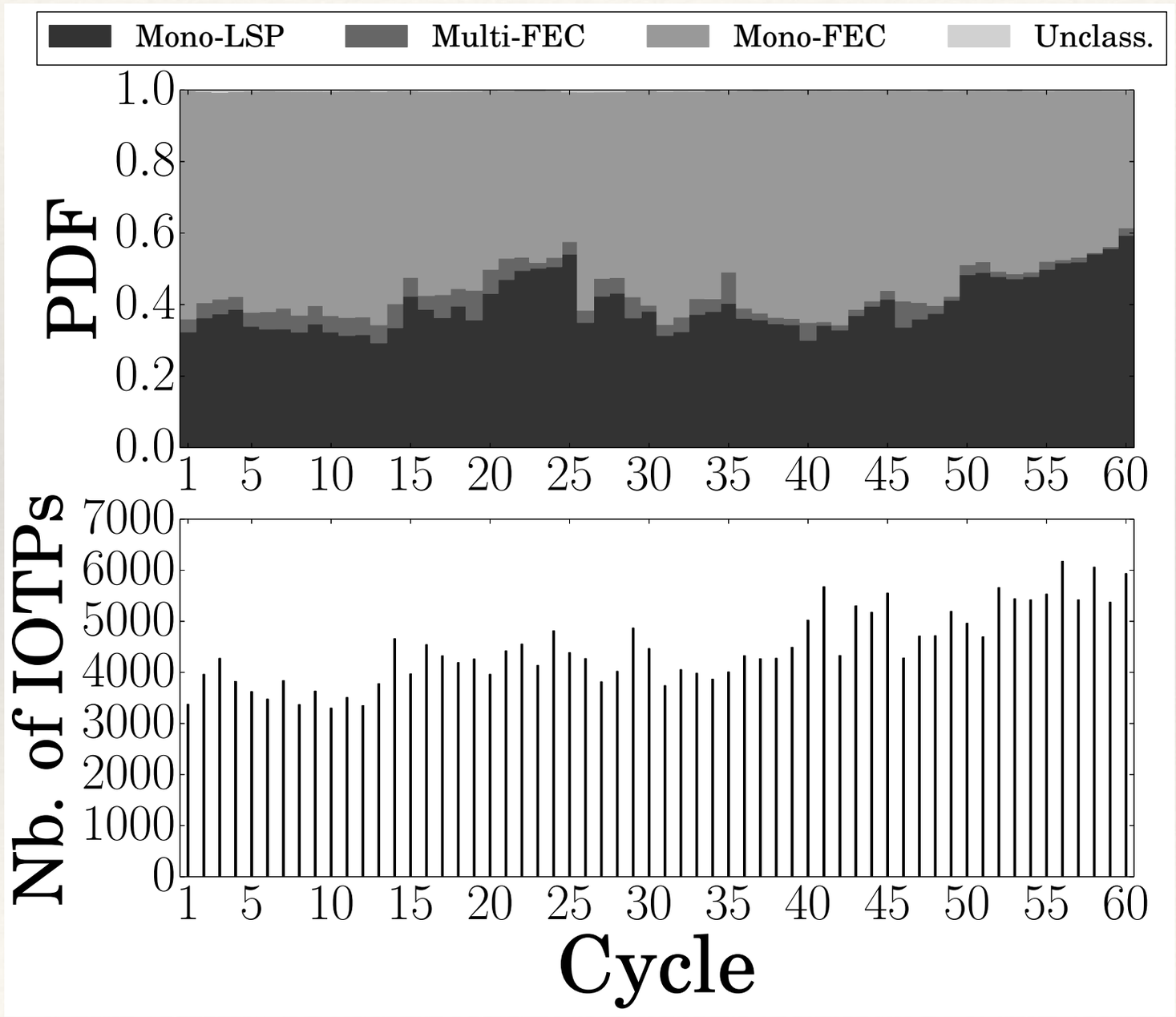
LPR - Results

- Classification for AS2914 (*NTT*)



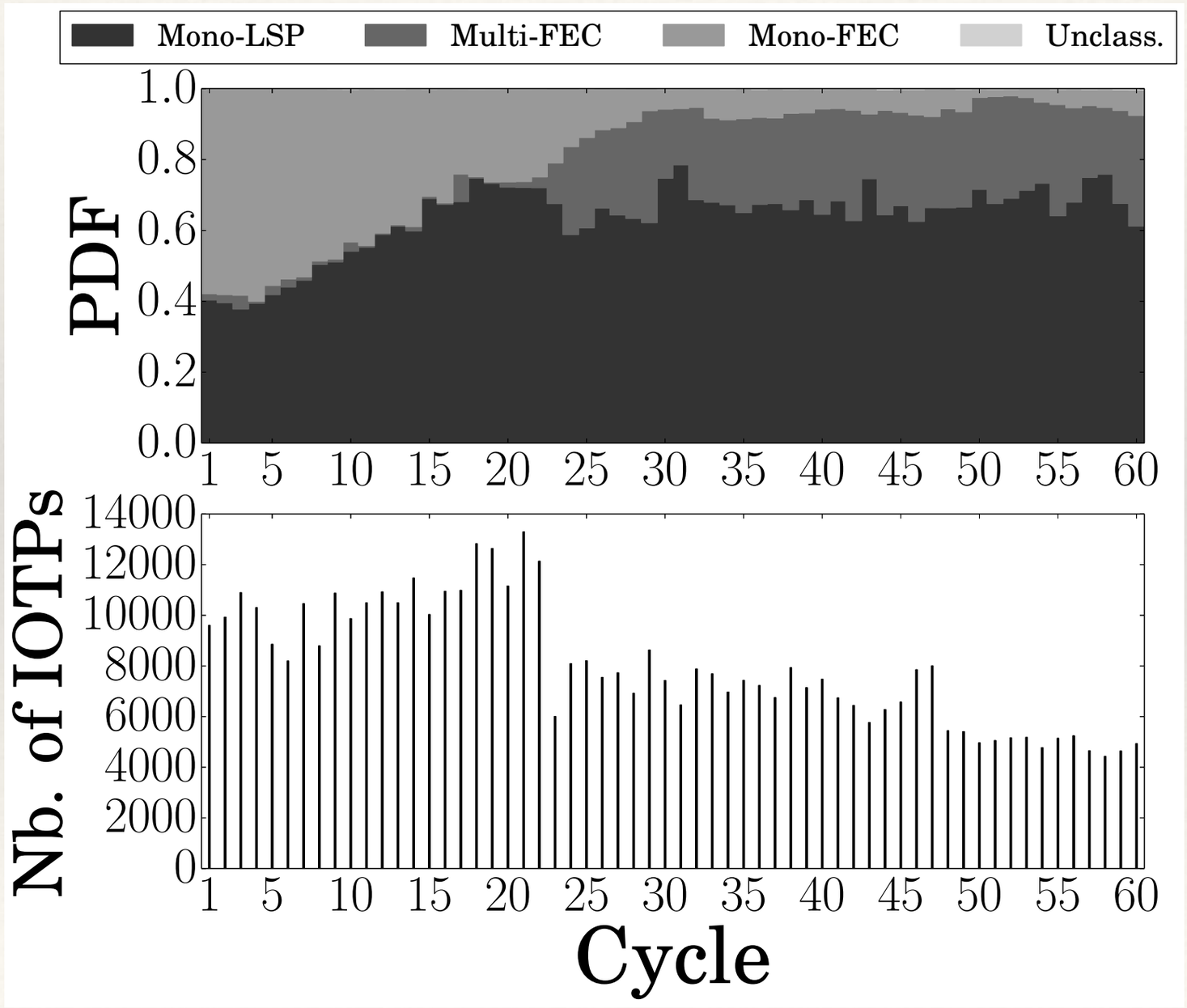
LPR - Results

- Classification for AS6453 (*Tata Communications*)



LPR - Results

- Classification for AS7018 (*AT&T*)



Agenda

- MPLS background
- Revealing MPLS tunnels
- MPLS Tunnels Diversity
- **Conclusions**

Conclusion

- New algorithm (LPR)
  - Reveals MPLS usage in ASes :
    - ✓ Basic usage for performance issue (*Mono-LSP*)
    - ✓ Traffic Engineering (*Multi-FEC*)
    - ✓ ECMP load balancing, with parallel links or disjoint routers (*ECMP Mono-FEC*)
  - Studies the evolution of MPLS usage

Conclusion

- Usage of MPLS increases over time
- Most operators seem to deploy MPLS
  - Usage depends on operator
    - ✓ Basic most of the time (Mono-LSP or ECMP)
    - ✓ Traffic engineering less common but well represented in some ASes
- In case of ECMP, the parallel links architecture seems predominant
- When TE is deployed, in many cases, different LSPs between the same endpoints take the same IP path
  - Bandwidth sufficiently abundant for allowing all LSPs on the same route

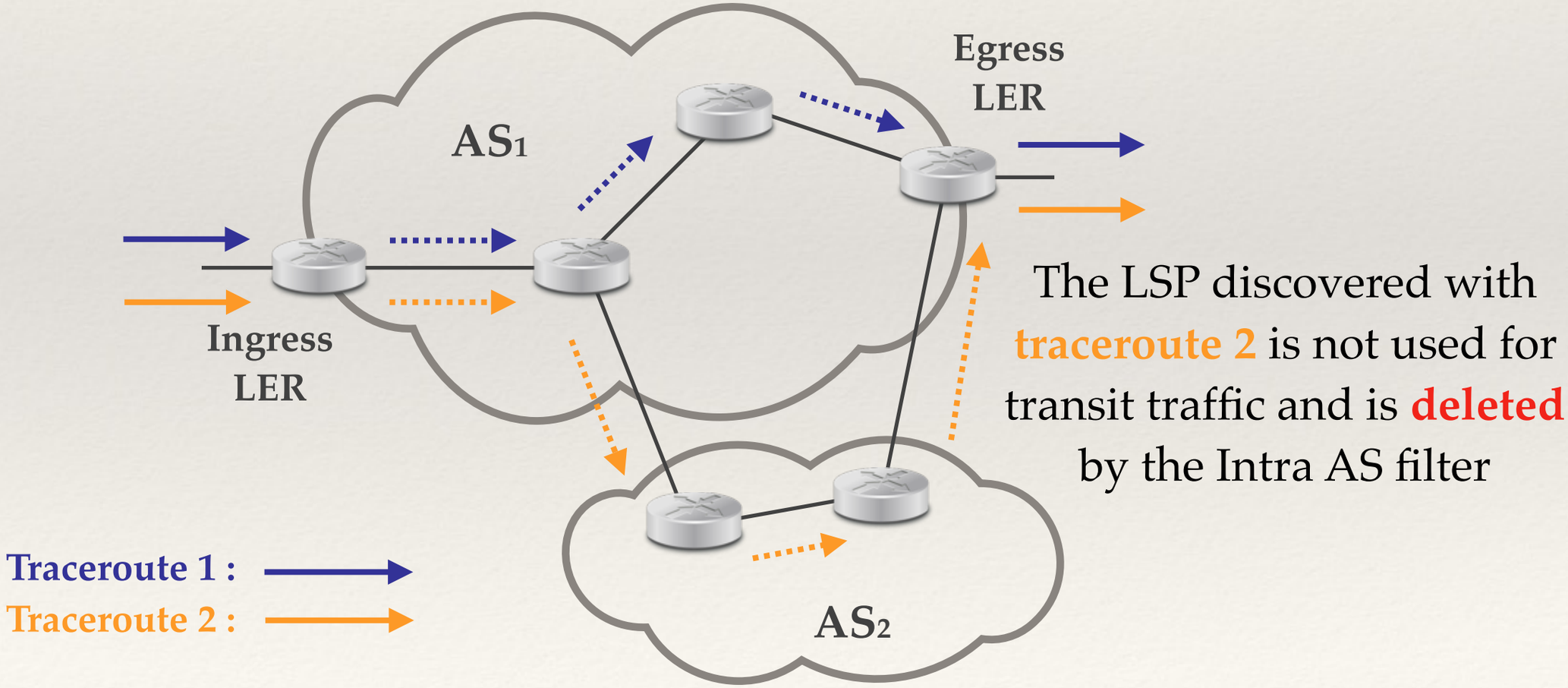
LPR - Static Filters

- Three static filters applied sequentially:

1. **Intra AS**
2. **Target AS**
3. **Transit Diversity**

LPR - Static Filters

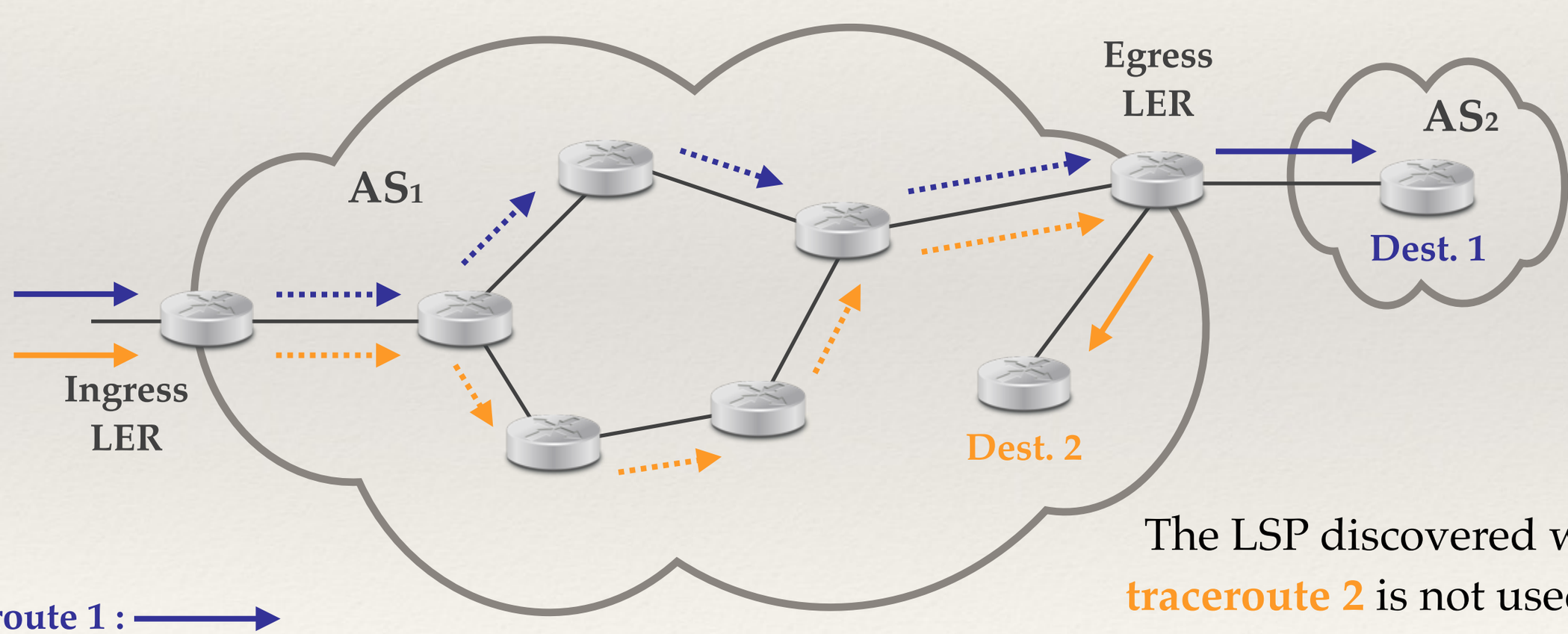
- **Intra AS filter**
  - IP addresses involved in an LSP must belong to the same AS



The LSP discovered with **traceroute 2** is not used for transit traffic and is **deleted** by the Intra AS filter

LPR - Static Filters

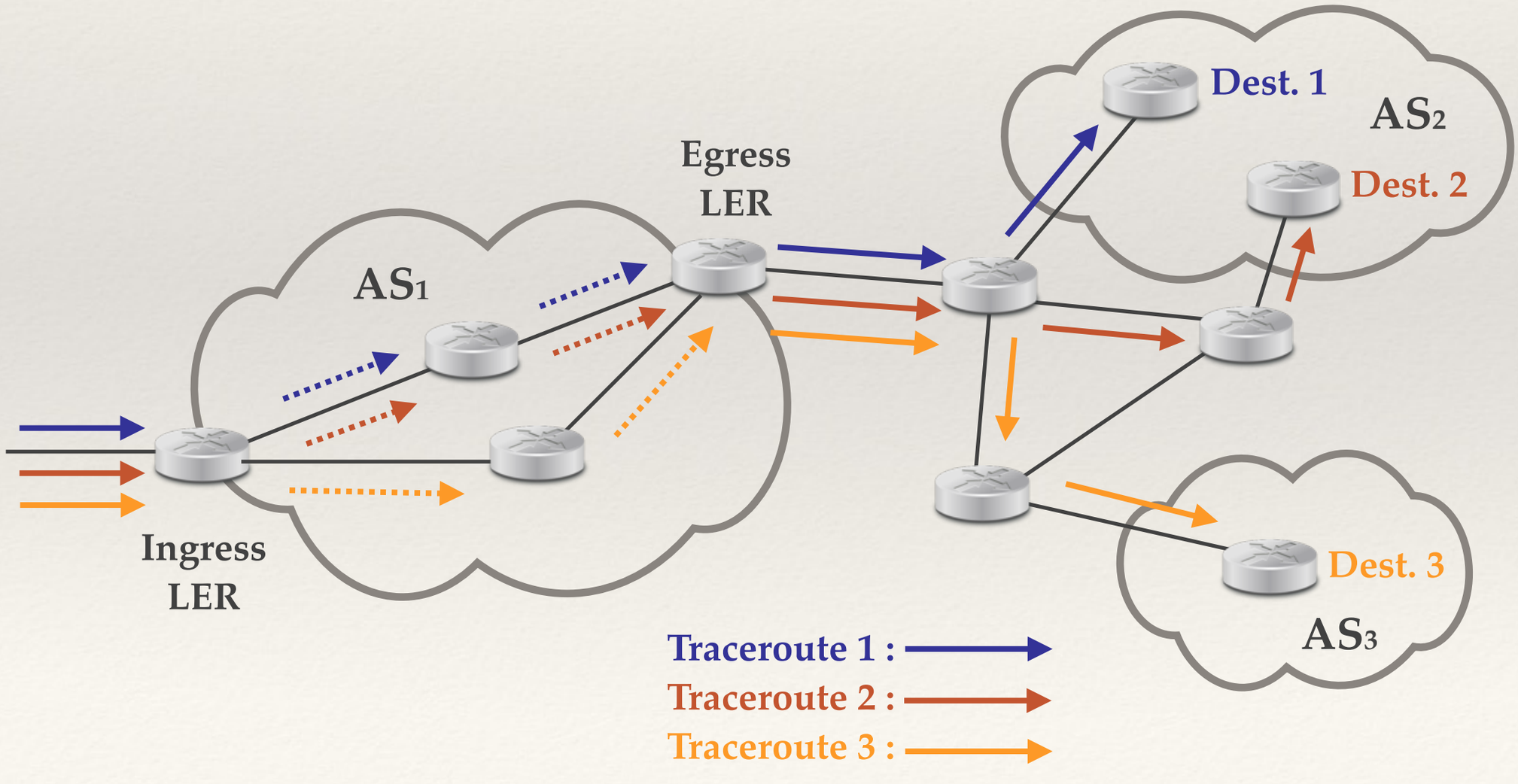
- **Target AS filter**
  - Destination of `traceroute` must be in a different AS than the LSP



The LSP discovered with **traceroute 2** is not used for transit traffic and is **deleted** by the Target AS filter

LPR - Static Filters

- **Transit Diversity filter**
  - The destinations of the `traceroutes` passing through a given IOTP must belong to at least two different ASes



LPR - Temporal Filter

- A temporal filter:
  - **Persistence** : Keep an LSP in measurement cycle X if it was also seen in cycle X + 1 or X + 2
  - Deletes noise due to routing changes

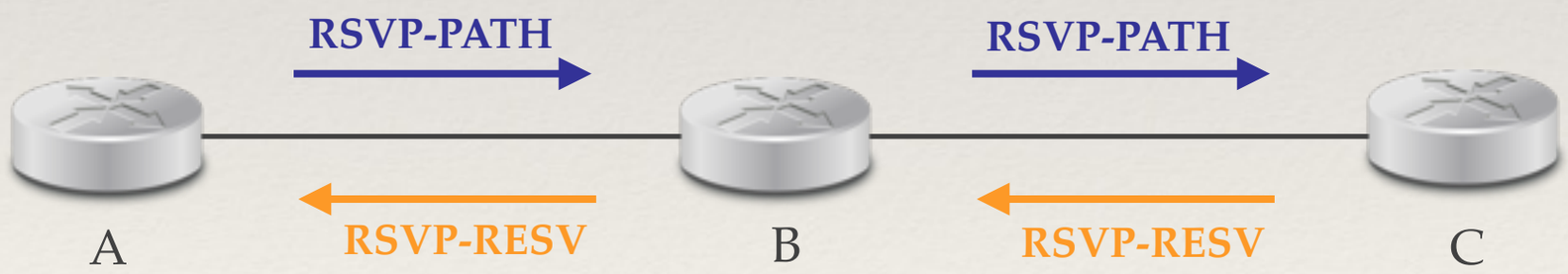
LPR - Results

- About 14M of LSPs on average per cycle
- Filtering impact:

1. **Intra AS**: Removes ~1% of LSPs
2. **Target AS**: Removes ~13% of remaining LSPs
3. **Transit Diversity**: Removes ~7% of remaining LSPs
4. **Temporal**: Removes ~10% of remaining LSPs

RSVP

- **R**esource Re**S**er**V**ation **P**rotocol (RSVP) [RFC2205]
- Used to allocate resources on the path
- Messages follow IP route



RSVP - TE

- **R**esource Re**S**er**V**ation **P**rotocol - **T**raffic **E**ngineering (TE) [RFC3209]
- *RSVP-RESV* can piggyback MPLS labels
- **E**xplicit **R**oute **O**bject (ERO) extension :
  - Allows the source to pre-calculate the LSP (not necessary the IP route)