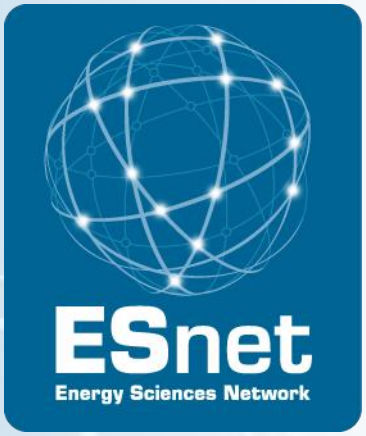


LHCONE Asymmetric Routing Issues

Mike O'Connor, Network Engineer

ESnet Engineering Group

LHCOPN and LHCONE Joint Meeting

Stockholm (SE)

May 3-4, 2012

ESnet Site LHCONE Connection Issues

The ESnet customer sites BNL, FNAL & SLAC have connected to LHCONE.

None of these sites are actively using LHCONE at this time.

Asymmetric routing has been a primary source of connection problems.

The community served by LHCONE is large enough to require a more dynamic approach to route policy implementation than has been required to date with the production LHCOPN.

These service affecting problems can be traced to the combination of packet filtering in **stateful firewalls and asymmetric routing** at the edge of either the local or remote LHCONE participating institution.

Asymmetric Routing

Asymmetric routing in the global Internet is normally an undesirable but acceptable condition.

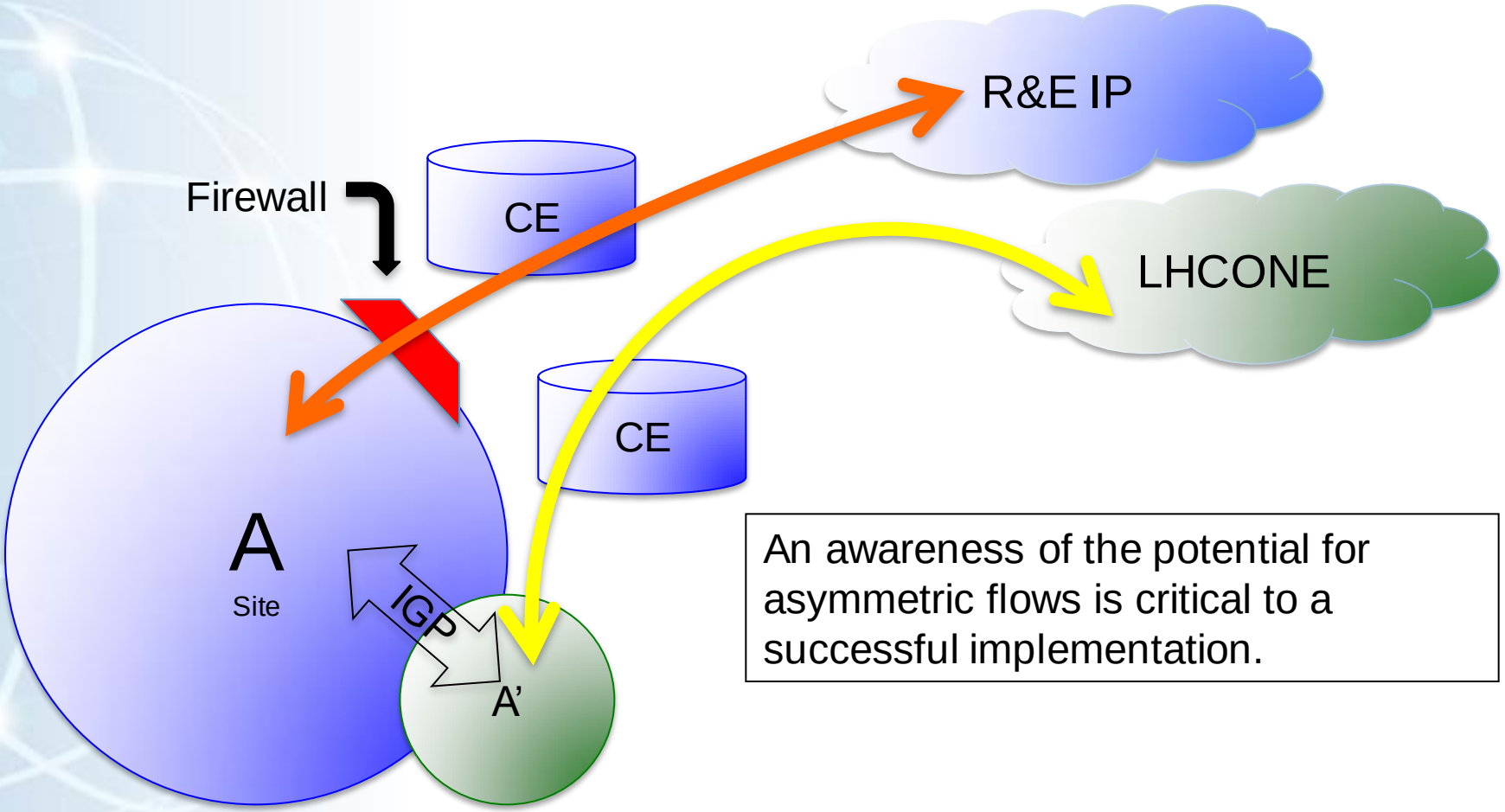
A stateful firewall will only permit connectivity if it is able to observe both directions of each flow.

Routing symmetry must be preserved at the firewall. Asymmetric routing through a state-full firewall will fail in many common connection scenarios.

The Science DMZ architectures in place at these sites provide high performance access controlled path alternatives to some form of perimeter chokepoint, as well as the potential for creating asymmetric flows.

Science DMZ

Provides a high performance path alternative to a common perimeter chokepoint



An awareness of the potential for asymmetric flows is critical to a successful implementation.

Observed Asymmetry

Single day of BNL flow data related to LHCONE

Octets Inbound to BNL

BNL LHCONE Testing

- BNL accepted only Napoli prefixes.
- Other centers that previously connected via R&E IP now transmitted to BNL via LHCONE.
- Replies remained on R&E IP
- Asymmetric flows and service affecting problems.

Napoli flows were symmetric

20% of total traffic.

Policy-based Routing

Policy based routing can be used to address routing asymmetry.

Issues with PBR:

Largely proprietary implementations with vendor and model specific limitations, limiting scalability.

The policy based routing supporting LHCOPN today can be considered a form of static routing. This will not easily scale to the larger more dynamic LHCONE community.

Simple configuration errors can easily create routing loops.

Certain implementations execute in software on the router CPU which impacts performance of the flow as well as the CPU.

The extensible standards based protocols that support routing in the Internet are based on destination routes, PBR is generally used for source based routing which lack common protocols.

Connection Recommendations

to address Asymmetric Routing

1 Define local LAN address ranges that will participate in LHCONE. Advertise these address range prefixes to LHCONE using BGP.

2 Agree to accept all BGP route prefixes advertised by the LHCONE community.

3 Ensure that only hosts in your locally defined LHCONE ranges have the ability to forward packets into the LHCONE network.

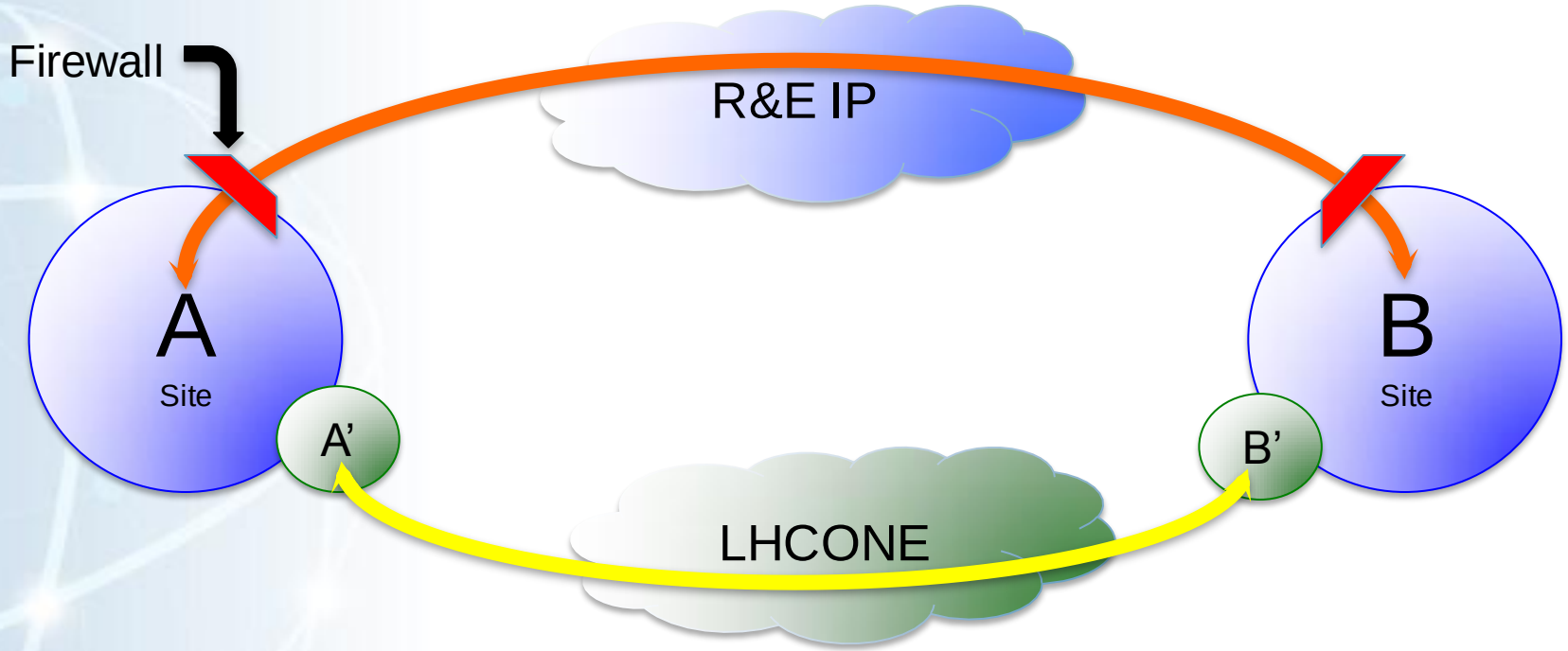
4 Ensure that the LHCONE paths are preferred over general R&E IP paths.

5 End sites should avoid static configuration of packet filters, BGP prefix lists and policy based routing, where possible. RPF filtering is suggested as a dynamic access control method for sites.

Both Sites Connect via General R&E IP

Simple Symmetric Path

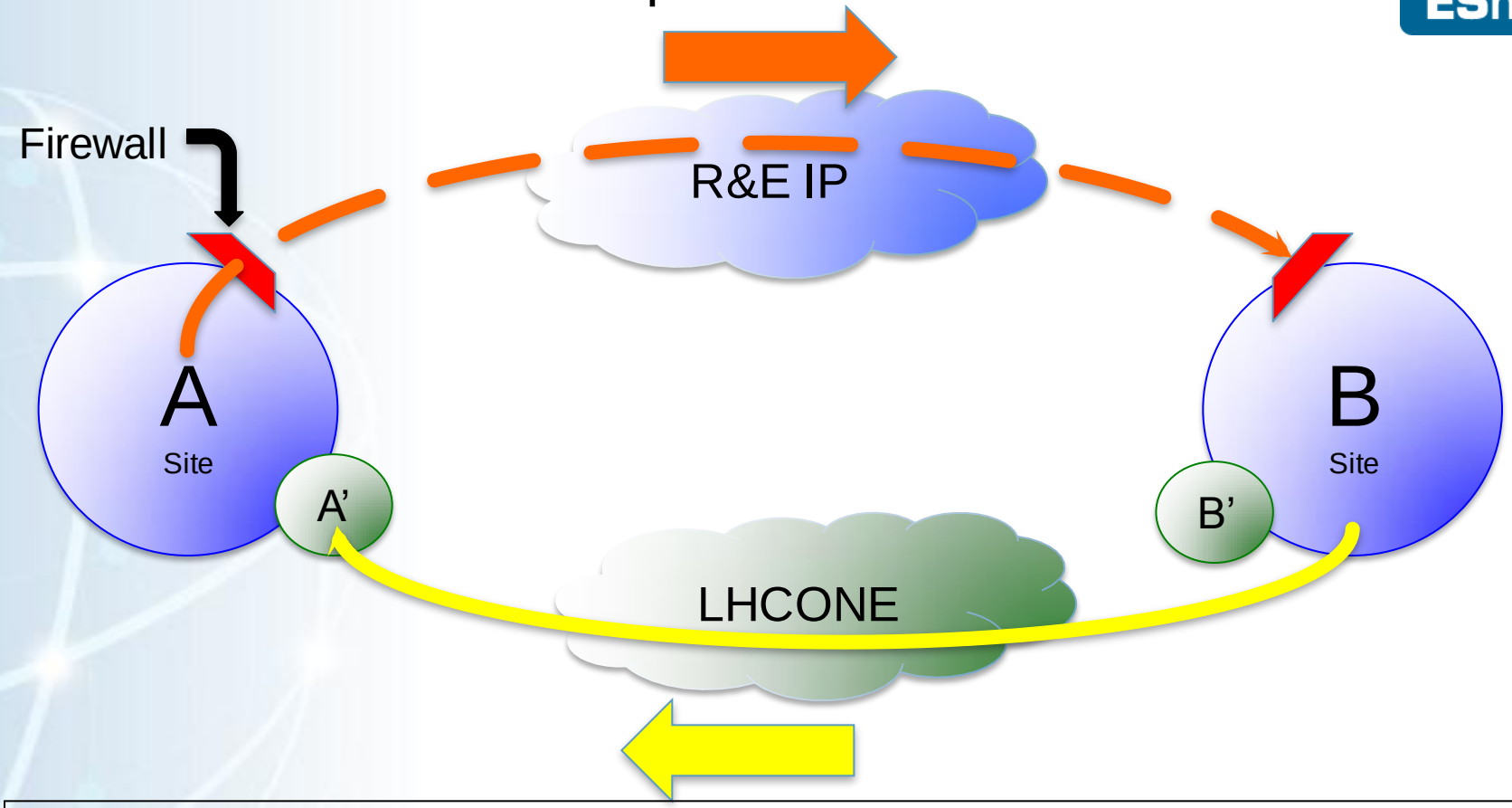
Science DMZ with LHCONE connectivity



A,B - the entire site address range
A',B' - the dedicated Science DMZ affiliated address range

Asymmetry B to A'

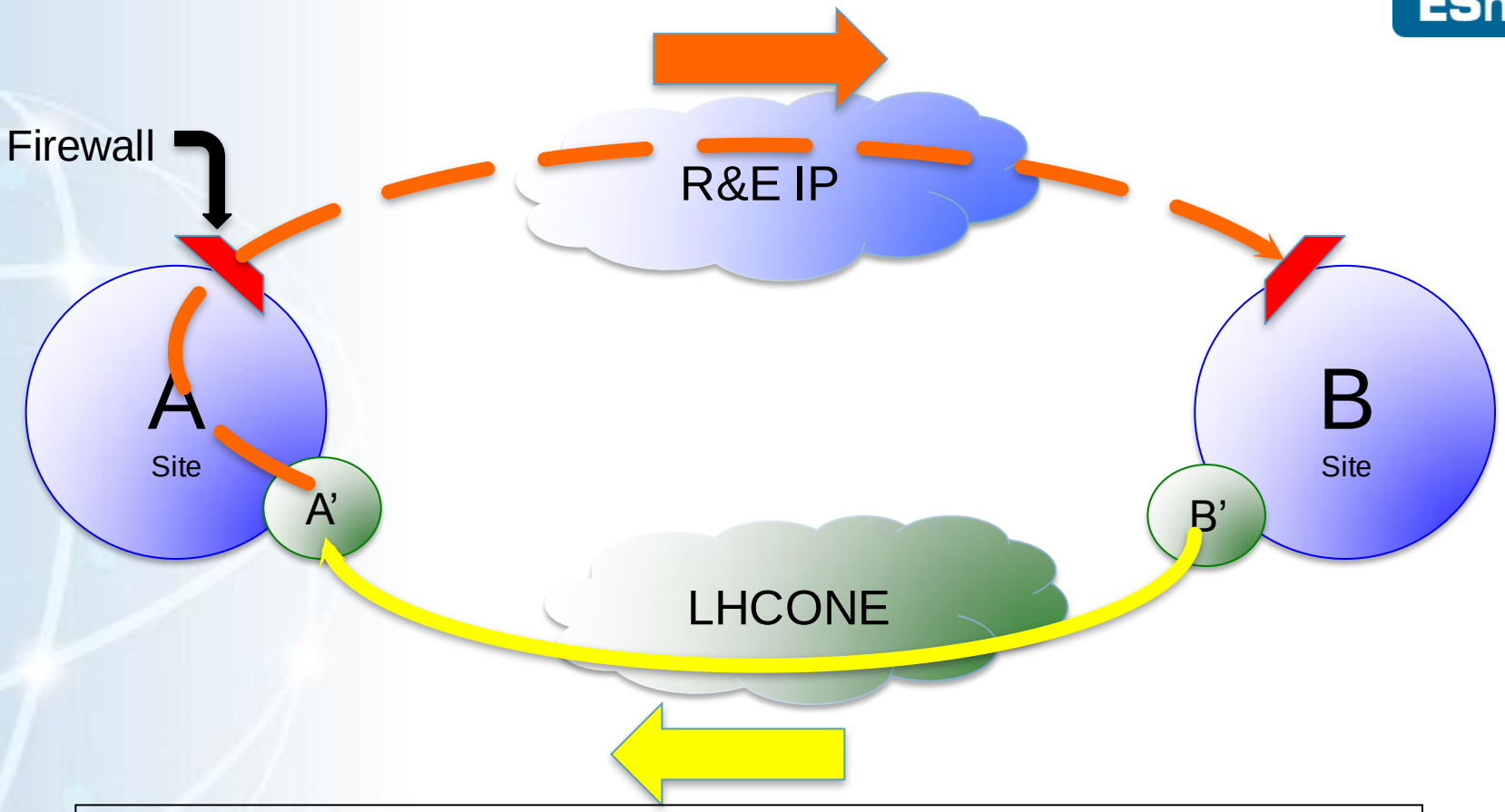
B must not source packets via LHCONE



Ensure that only hosts in your locally defined LHCONE ranges have the ability to forward packets into the LHCONE network.

Asymmetry B' to A'

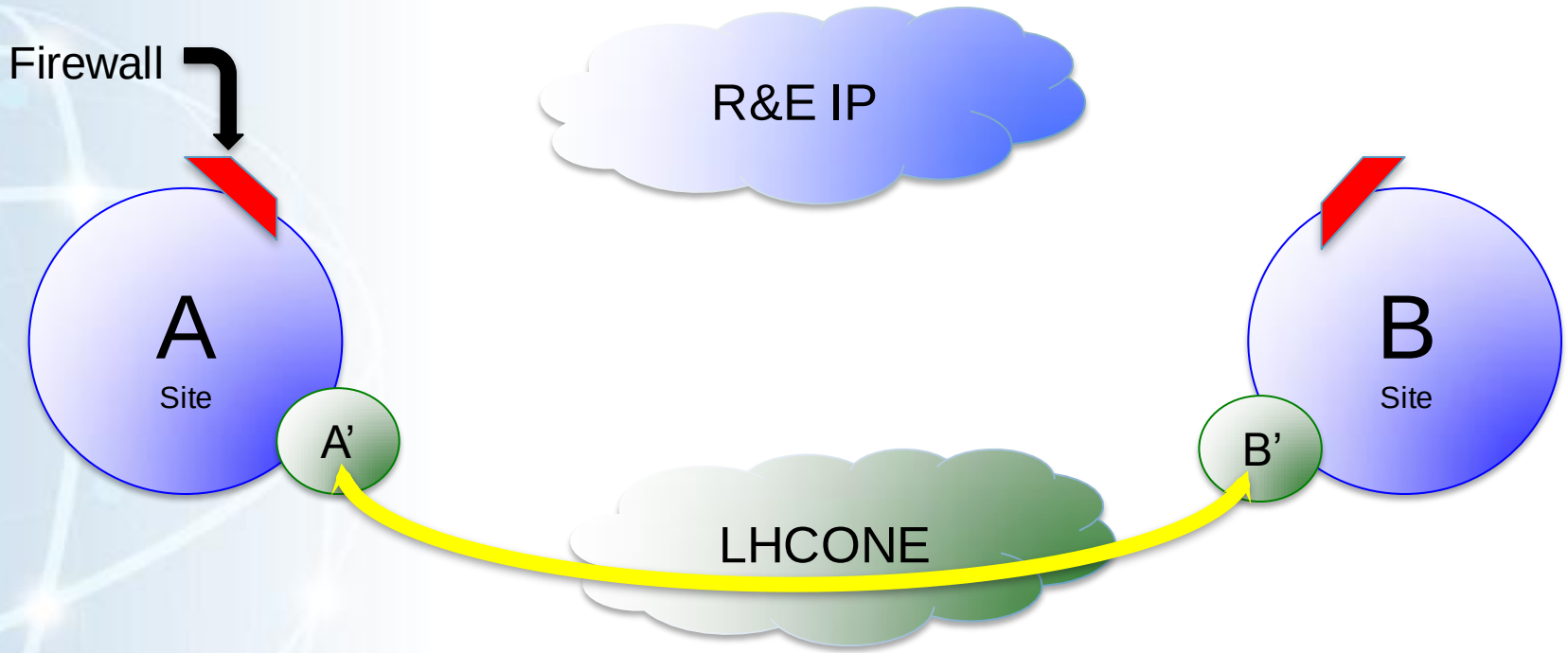
Both LHCONE Routed Prefixes



Site A has blocked the BGP prefix for B'

Agree to accept all BGP route prefixes advertised by the LHCONE community.

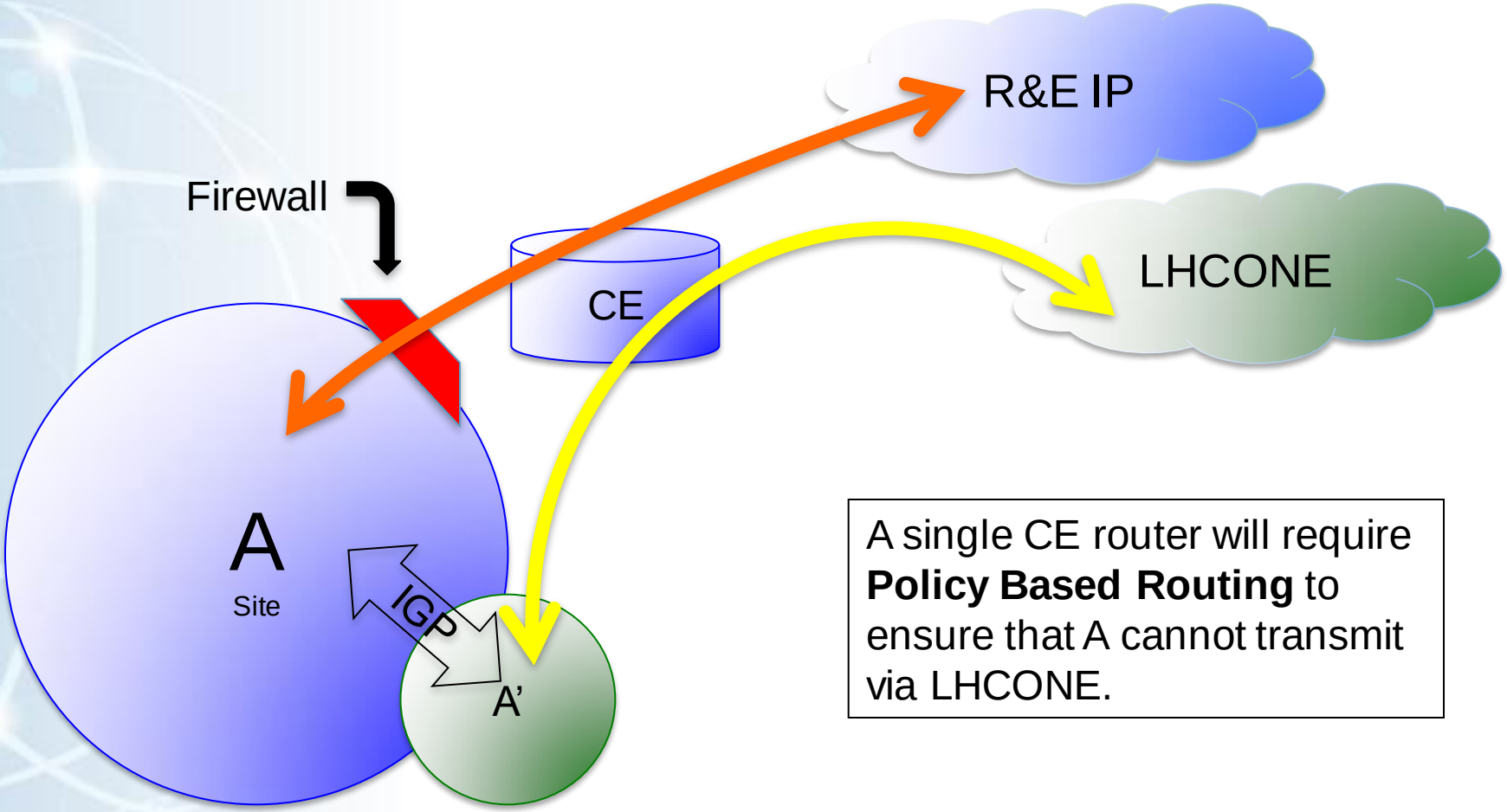
A' to B' must use LHCONE



Ensure that the LHCONE defined ranges are preferred over general R&E IP paths.

Single CE Router

Connecting to Both R&E IP and LHCONE

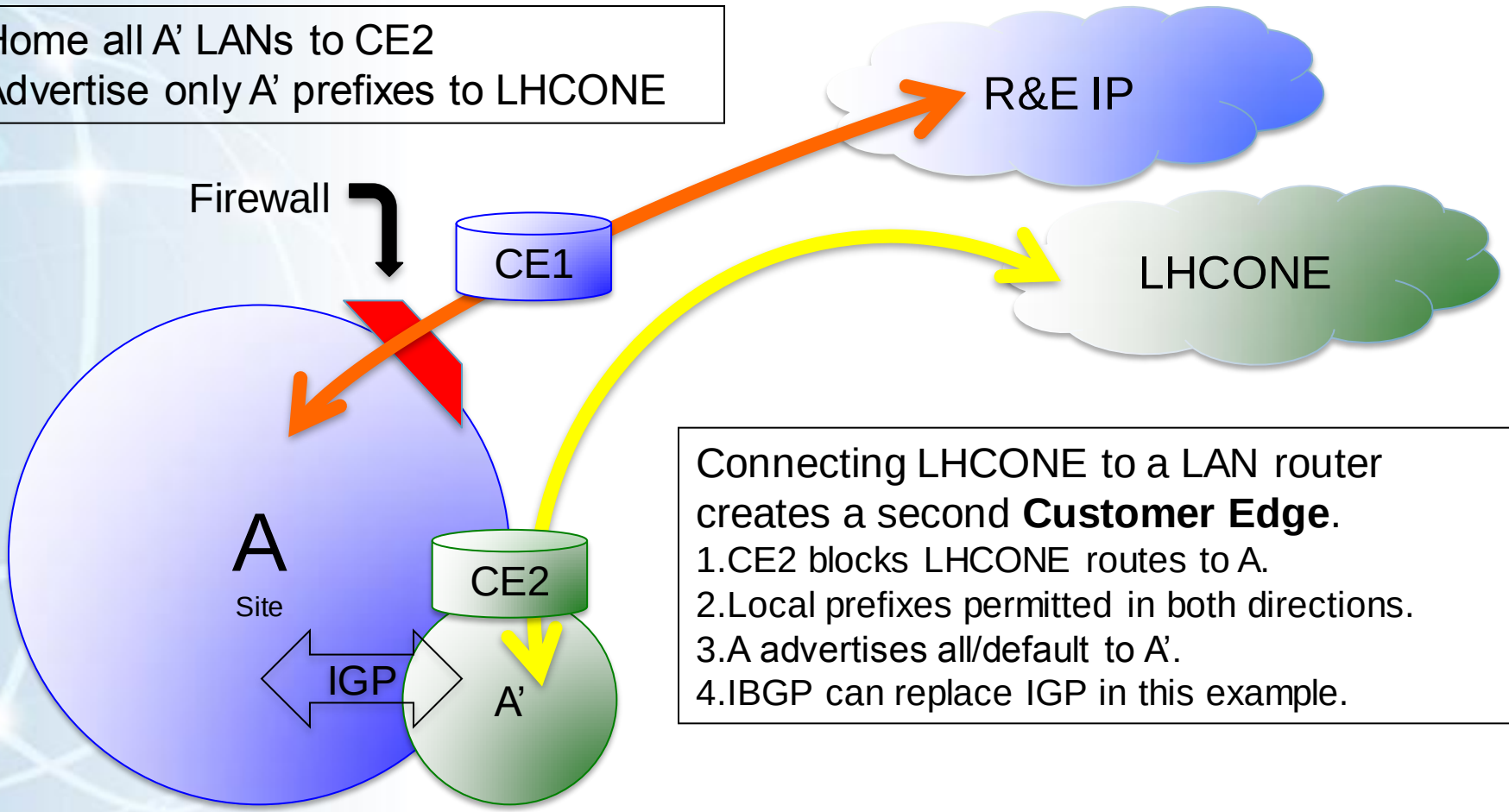


A single CE router will require **Policy Based Routing** to ensure that A cannot transmit via LHCONE.

Dual CE Routers

Maintains Traffic Separation without PBR
An Example Method

Home all A' LANs to CE2
Advertise only A' prefixes to LHCONE



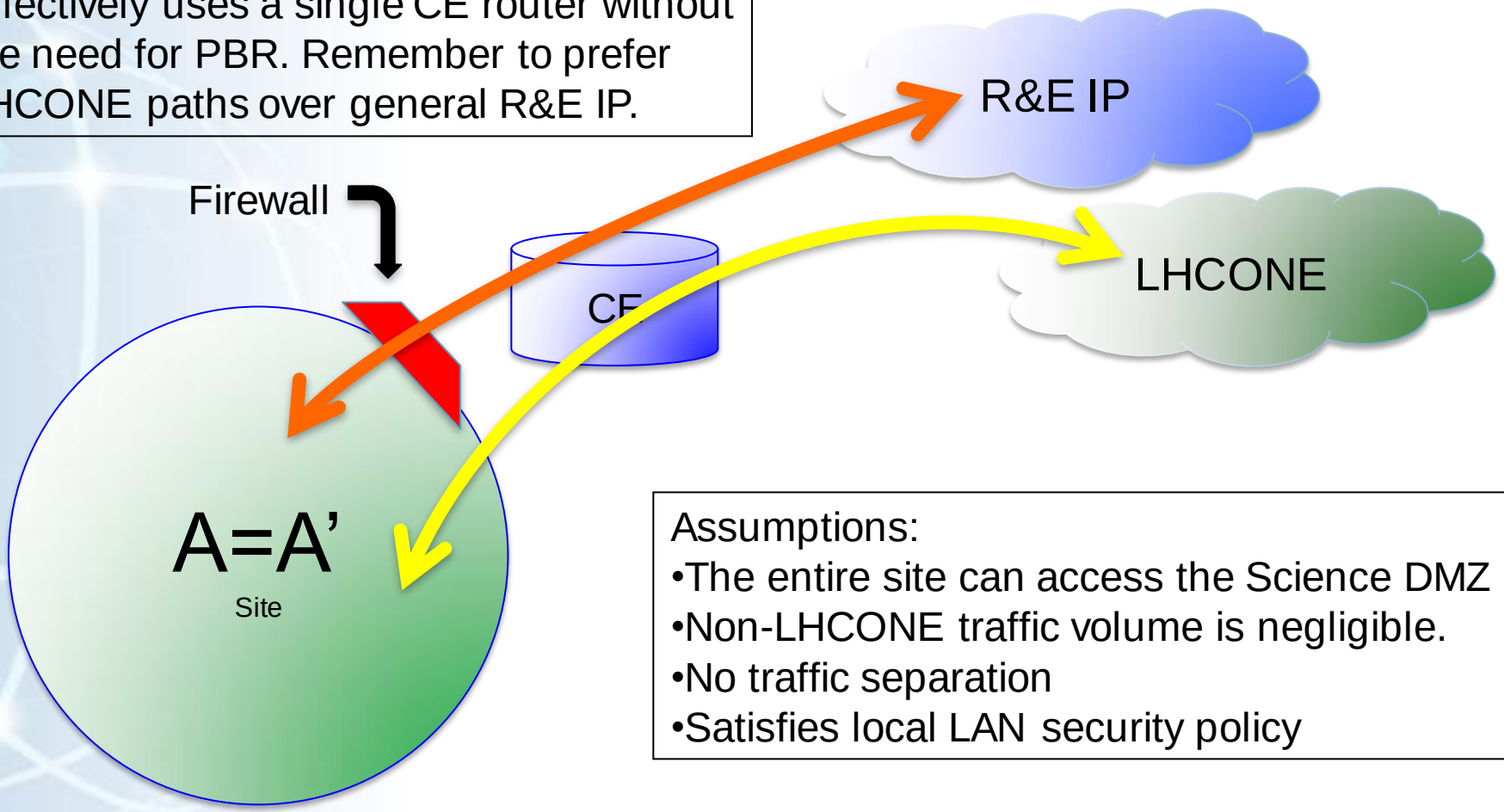
Connecting LHCONE to a LAN router creates a second **Customer Edge**.

1. CE2 blocks LHCONE routes to A.
2. Local prefixes permitted in both directions.
3. A advertises all/default to A'.
4. IBGP can replace IGP in this example.

A = A’

Advertise ALL LAN prefixes to LHCONE

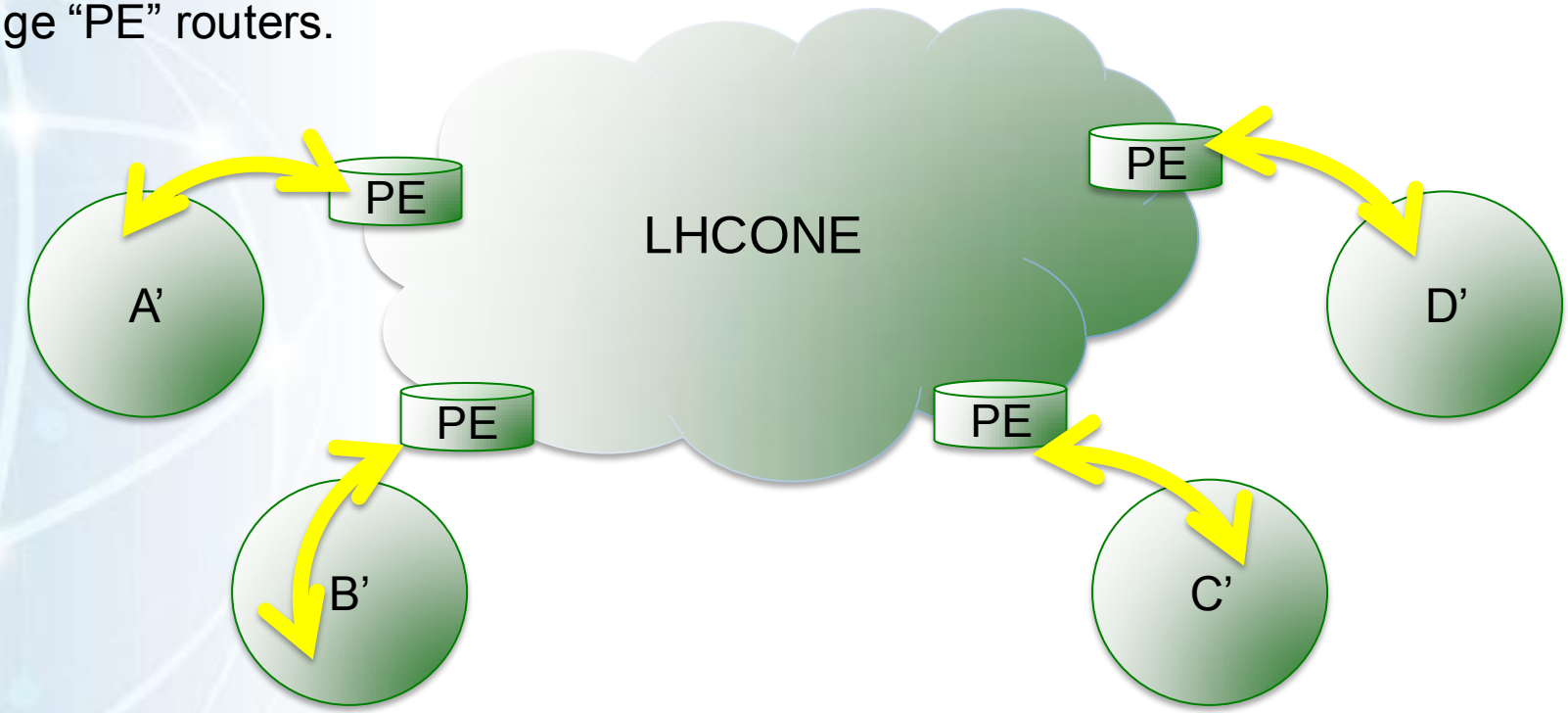
Effectively uses a single CE router without the need for PBR. Remember to prefer LHCONE paths over general R&E IP.



Assumptions:
- The entire site can access the Science DMZ
- Non-LHCONE traffic volume is negligible.
- No traffic separation
- Satisfies local LAN security policy

LHCONE Perimeter Access Control

In order to provide access control into this scalable and dynamic environment, it is critical that the NSPs filter route prefixes at their provider edge “PE” routers.



Filtering packets and prefixes at each customer edge will not accommodate changes to remote collaborating sites without an N^2 out of band negotiation that does not scale and will create random periods of unacceptably long routing asymmetry.

Conclusions

LHCONE connecting sites have experienced service affecting issues.

LAN reconfiguration may be necessary to successfully implement a Science DMZ with LHCONE.

An implicit trust relationship exists between all LHCONE participants and NSPs.

LHCONE NSP's must assume responsibility for filtering both packets and prefixes from their connecting institutes.

Destination based routing protocols should be trusted to to build and distribute the LHCONE routing table in a scalable fashion. Static configuration at the customer edge routers (ie: prefix, packet filters & PBR) should be avoided.

The large number of LHCONE participants will make it difficult to maintain routing symmetry. – *Antonio Ceseracciu SLAC*

Questions?

Michael O'Connor

ESnet Network Engineer

moc@es.net

631 344-7410