”multiONE” with BGP communities

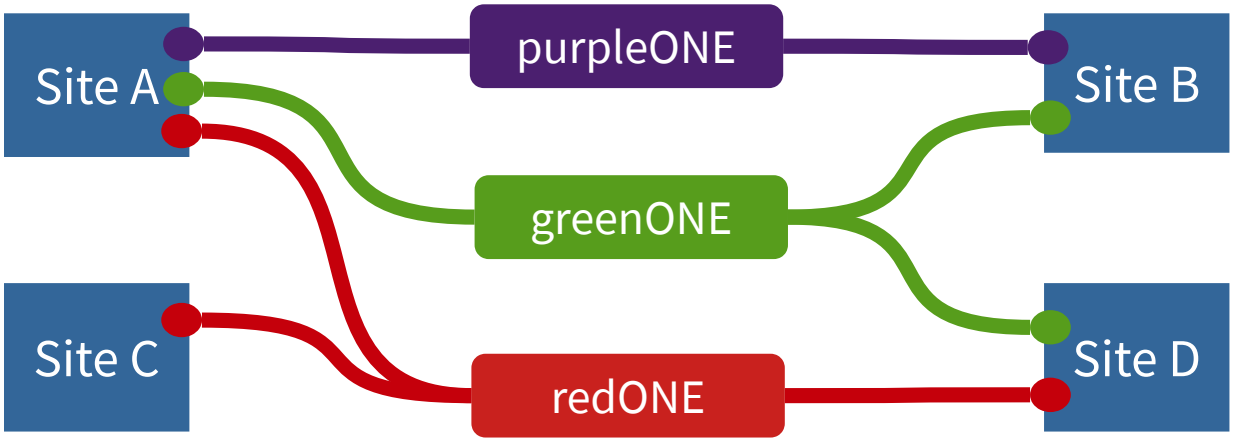
LHCONE meeting #52
10th April 2024 – Catania IT
edoardo.martelli@cern.ch

Recap: multiple “LHCONEs”

Each site joins only the VPNs of the groups it is collaborating with (e.g. ATLAS-ONE, CMS-ONE, DUNE-ONE, BelleII-ONE...)

- **Major Benefit**: reduced exposure of data-centre/Science-DMZ to other sites

- **Major Challenge**: how to correctly route traffic into VPNs at sites that join several of them? Operational complexity



What happened so far?

Several proposals made, but all difficult to implement

Main problems:
- traffic separation not easy to achieve at sites
- configuring multiple VPNs add complexity to network operations at connected sites and NRENs

New proposal

Don’t add any additional VPN (or maybe just another one for “Other-Big Sciences-beside-LHC”?)

Each prefix announced to LHCONE is tagged with BGP communities that identify the collaborations served by the site

The tagging is done by the sites. Or by the connecting REN if the site is unable to do it

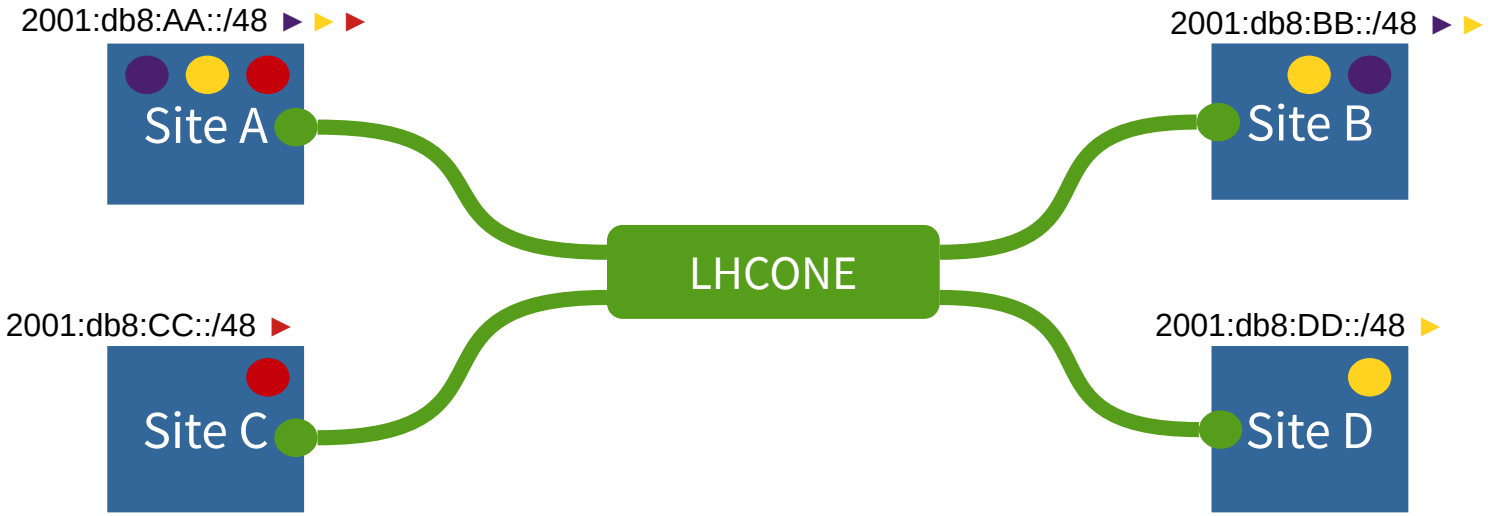
Sites can/should accept only the prefixes of the collaboration they are working with

In addition/alternative, RENs could announce to a given site only the prefixes of the collaborations related to the site

Practical example

Tagging

Each LHCONE site:
- tags its prefixes announced to LHCONE with all the BGP communities that identify the collaborations the site is participating in

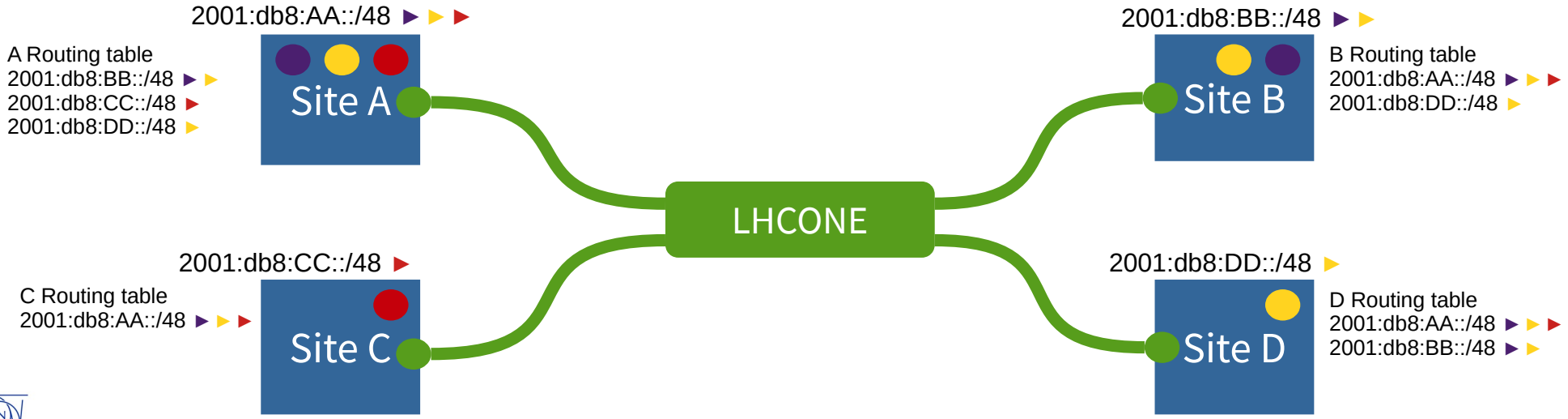


Practical example

Filtering

Each LHCONE site:

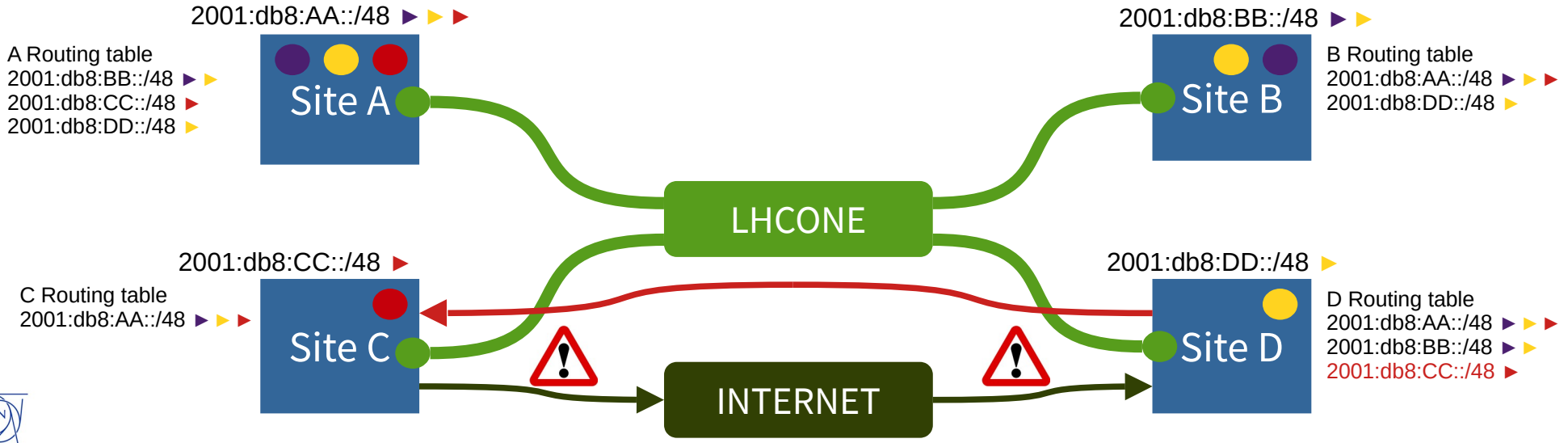
- accepts only the prefixes tagged with the BGP communities of its own collaborations



What could go wrong?

Site D (▶) doesn't filter out not-relevant prefixes (▶ ▶)

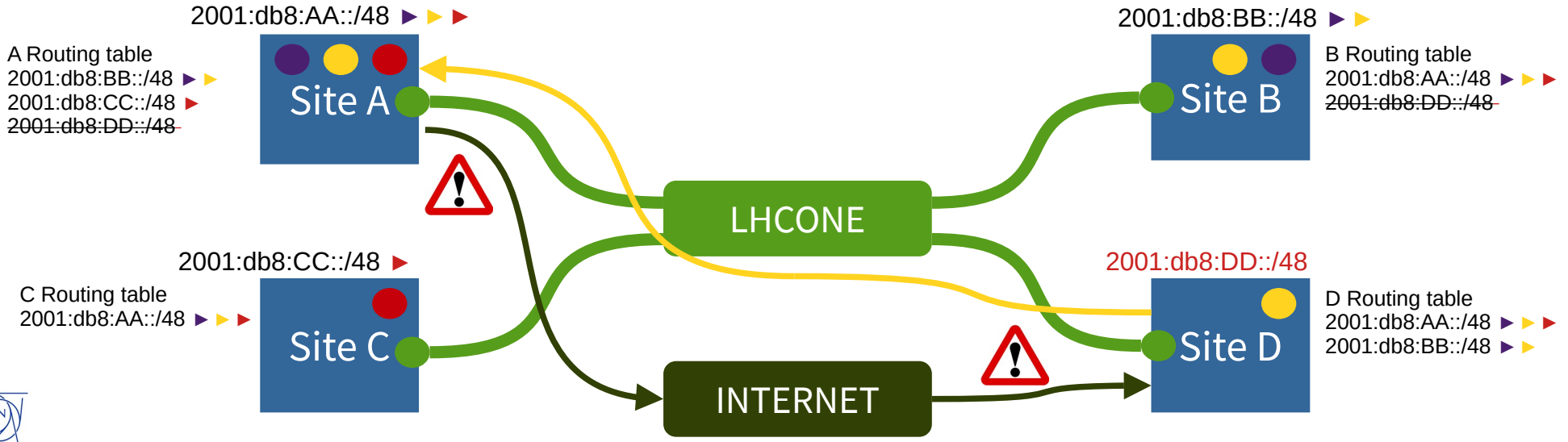
- traffic from D to any site is sent to LHCONE but comes back via the Internet
- asymmetry may cause connections drop (by statefull firewalls)
- ..but anyway there should be no traffic between those sites, and the fix is easy to apply (add filtering at site D)



What could go wrong?

Site D (▶) doesn't tag its prefixes (▶)

- all sites (including the ▶ ones) will drop the D's untagged prefixes
- traffic from D to any site goes via LHCONE but comes back via the Internet
- asymmetry may cause connections drop
- Site D will immediately realize the mistake because LHCONE doesn't work for it



Benefits

- Reduced exposures of sites
- No additional VPNs to configure
- No changes at sites when a new site connects to LHCONE
- Changes will be necessary when a new Collaboration joins LHCONE; all interested sites will have to accept the new tag
- Communication errors will be an incentive to adhere and will highlight already existing implementation errors and weaknesses

Limitations

- Security: a malicious sites can tag its prefixes with all the existing tags and get the prefix accepted (it could be contained if RENs do the tagging, or filter according to CRIC)

Implementation proposal

- Define BGP communities for the different LHCONE collaborations

- Implement prefix tagging at sites
    - if sites can’t do it, their LHCONE providers will do for them

- Verify all prefixes in the LHCONE routing tables are tagged

- Gradually implement filtering at sites

Existing LHCONE BGP communities

BGP communities are already in use in LHCONE for traffic engineering

Community	Type	Meaning	Notes
65001:XXXX	Operational	prepend 1x to ASxxxx	Mandatory
65002:XXXX	Operational	prepend 2x to ASxxxx	Mandatory
65003:XXXX	Operational	prepend 3x to ASxxxx	Mandatory
65010:XXXX	Operational	do not announce to ASxxxx	Mandatory

https://twiki.cern.ch/twiki/bin/view/LHCONE/LhcOneVRF#BGP_communities

Existing LHCONE BGP communities

Actually, we already had such communities, but were not implemented

Community	Type	Meaning	Notes
(tierx-org-as):65201	Informational	Tier-X supporting ALICE	Optional, set by the Tier-X
(tierx-org-as):65202	Informational	Tier-X supporting ATLAS	Optional, set by the Tier-X
(tierx-org-as):65203	Informational	Tier-X supporting CMS	Optional, set by the Tier-X
(tierx-org-as):65204	Informational	Tier-X supporting LHCb	Optional, set by the Tier-X

https://twiki.cern.ch/twiki/bin/view/LHCONE/LhcOneVRF#BGP_communities

LHCONE BGP communities - definition

BGP communities to identify the Collaboration served by the tagged IP prefix could be defined this way:

ASN:Exp-ID

Autonoums System Number could be:
- AS of the site originating the prefix (may become difficult to filter)
- Private AS (risk of clashes, but we are already using them)
- An AS number to be assigned (possibly 16bits)
- An AS number already in use, e.g. AS61339 which is used for the LHCONE looking-glass

Experiment-Identifier
- An unique number per collaboration. We could use the ExpId already defined for SciTags

LHCONE BGP communities - proposal

ASN = 61339 (LHCONE looking glass)
ExpId = SciTags ExpId values

Collaboration	*BGP Community (AS:ExpId)*
ALICE	61339:**5**
ATLAS	61339:**2**
BelleII	61339:**6**
CMS	61339:**3**
DUNE	61339:**8**
JUNO	61339:**12**
LHCb	61339:**4**
NOvA	61339:**13**
Pierre Auger Observatory	61339:**11**
XENON	61339:**14**

Implementation - Proposal

Agree on the project

Reach out all the LHCONE sites and request to implement the tagging, while reviewing their prefix declarations in CRIC

Monitor the progress of the tagging in the LHCONE routing tables

Target: all prefixes tagged by LHCONE meeting #54 (Spring 2025)

Implement filtering at WLCG sites during year 1 of LHC LS3 (2026)

Questions? Comments?

edoardo.martelli@cern.ch