LACP Improvement Proposal

Using LACP for the NNI problem

Rev. 1

Norman Finn

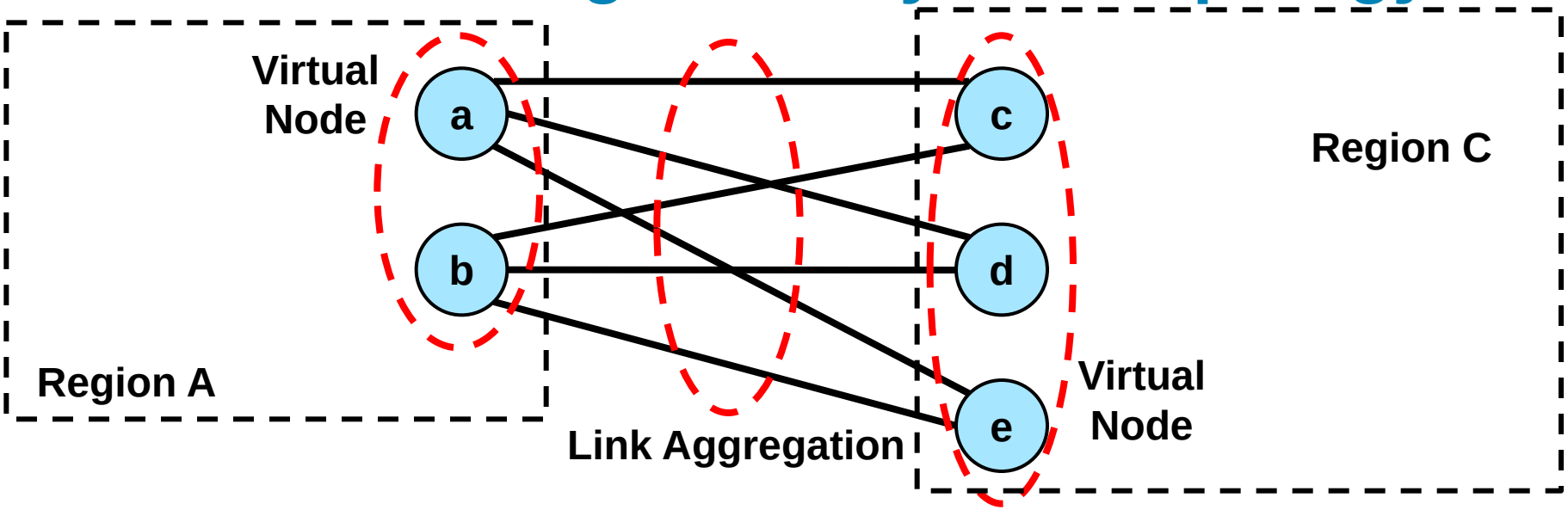
nfinn@cisco.com

Ethernet NNI

- This contribution is available at: new-nfinn-LACP-proposal-0910-v01.pdf.
- A recent related contribution is: new-nfinn-light-nni-0710-v04.pdf.
- The purpose of this contribution is to describe a work program to enhance LACP that will satisfy the need for an Ethernet Network-Network Interface (ENNI).
- Thanks to Don Fedyk for a key idea in this presentation, without delegating to him any responsibility for its accurate representation herein.

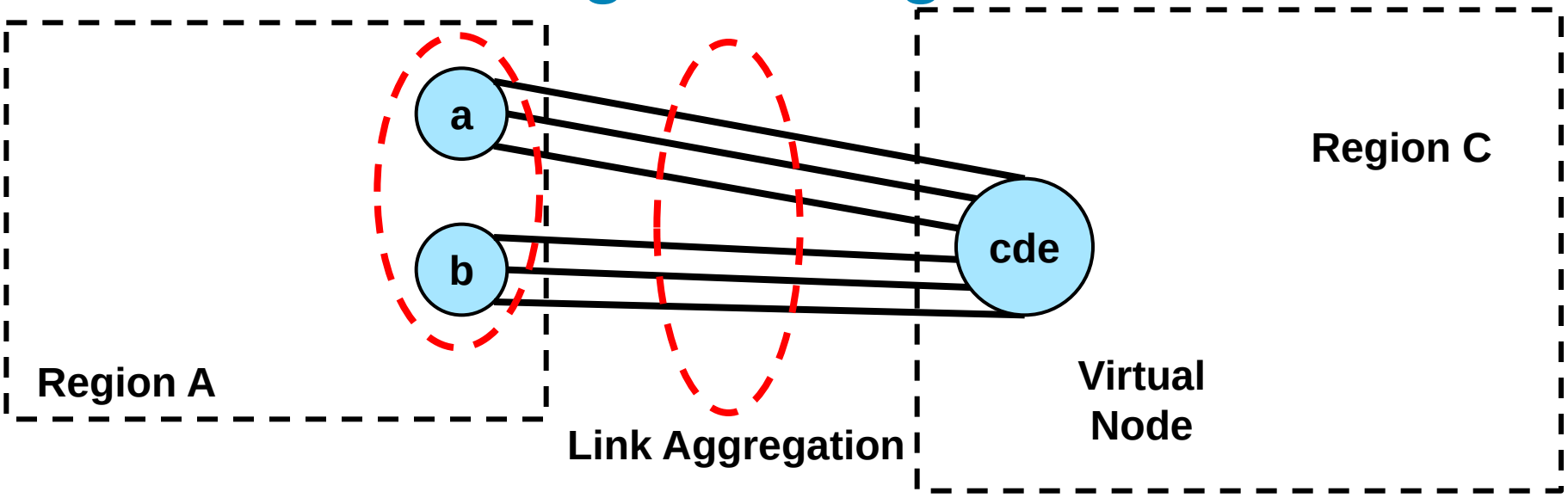
What is a “Virtual Node"

Reference diagram: Physical topology



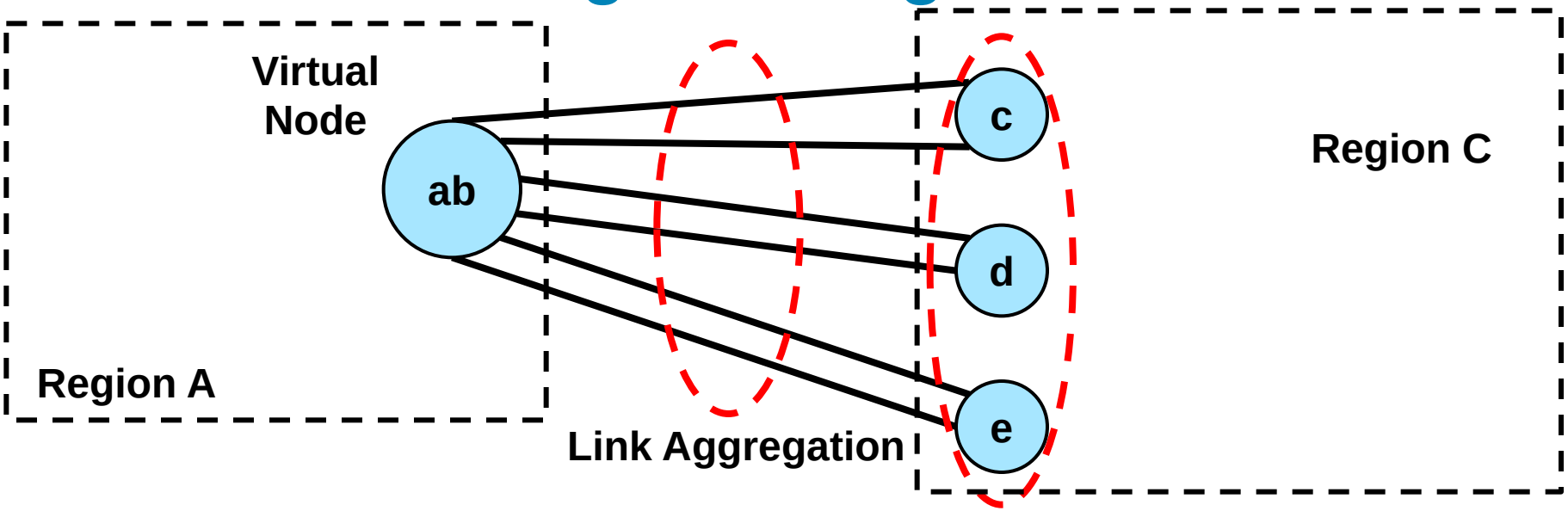
- Let us suppose that this is the physical connectivity between two Regions **A** and **C**.
- We will make no assumptions, for the moment, about the physical or logical connections existing within each Region, e.g. between Nodes **a** and **b**, or among Nodes **c**, **d**, and **e**.

Reference diagram: Region A sees



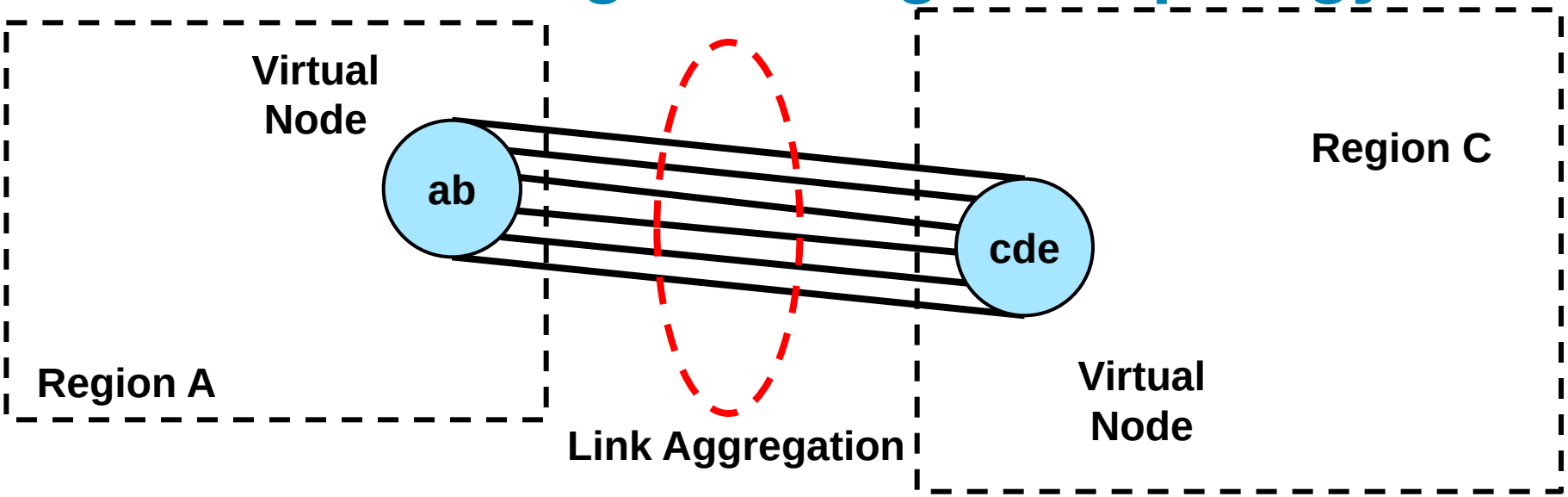
- This is what the interconnect looks like to Region A.
- Nodes **a** and **b** know they are separate systems.
- According to the LACP PDUs received by **a** and **b**, Nodes **c**, **d**, and **e** are a single Node.
- We call this apparent "**cde**" Node a **Virtual Node**.

Reference diagram: Region C sees



- This is what the interconnect looks like to Region C.
- Nodes **c** and **d**, and **e** know they are separate systems.
- According to the LACP PDUs received by **c**, **d**, and **e**, Nodes **a** and **b** are a single Node.
- Again, this apparent "**ab**" Node is a **Virtual Node**.

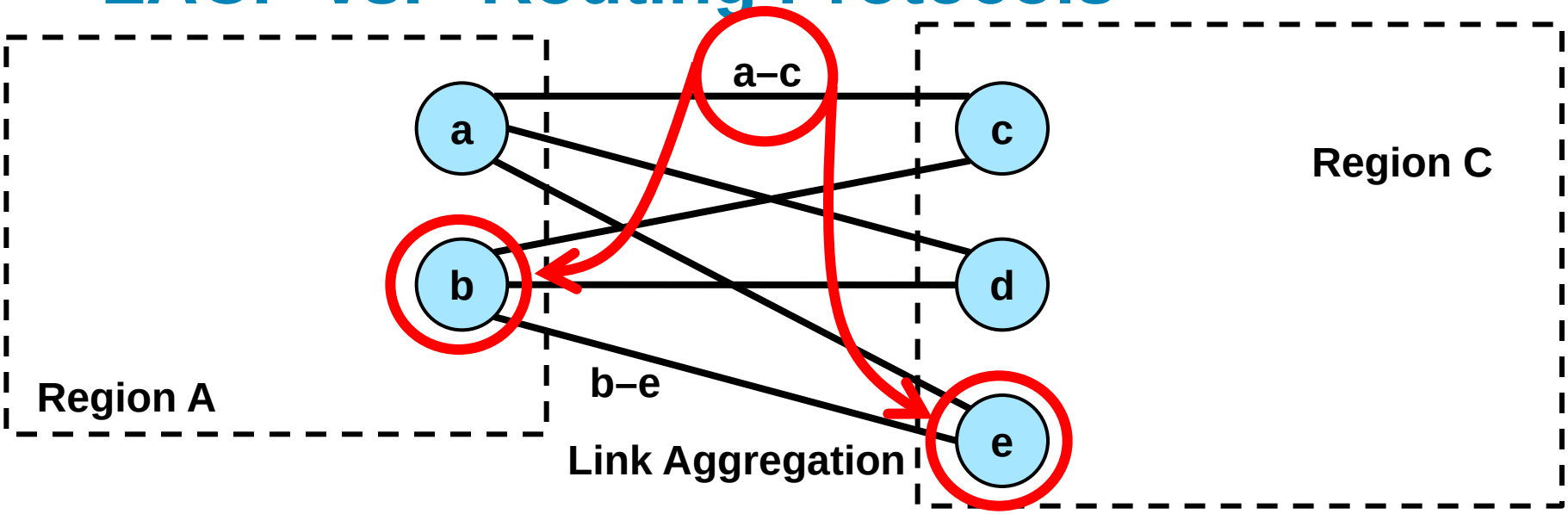
Reference diagram: Logical topology



- Thus, as far as the LACP PDUs are concerned, this is the topology.

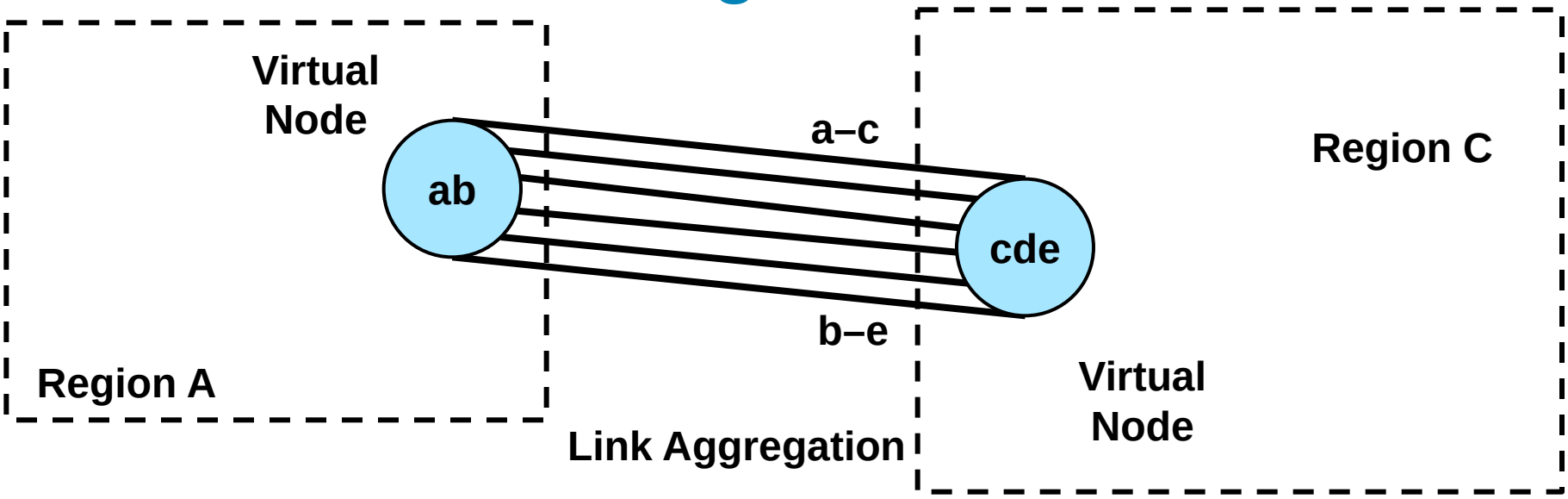
Routing vs. LACP

LACP vs. “Routing Protocols”



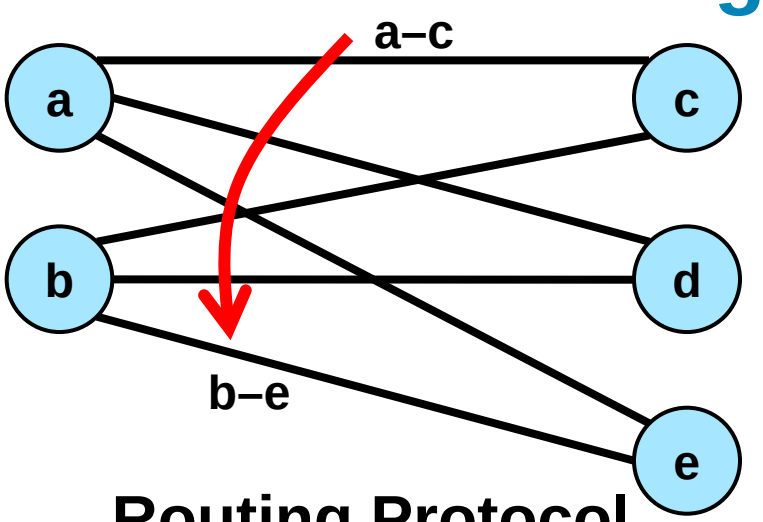
- Let us suppose that link **a–c** is the primary path for some Service, and that **b–e** is the alternate path.
- In order for **b** and **e** to start passing data over Link **b–e**, they must know that Link **a–c** has failed.
- Let us define “**routing protocol**” as the means, whether PDUs or supersonic carrier pigeons, by which knowledge of the state of Link **a–c** reaches **b** and **e**.

LACP vs. “Routing Protocols”

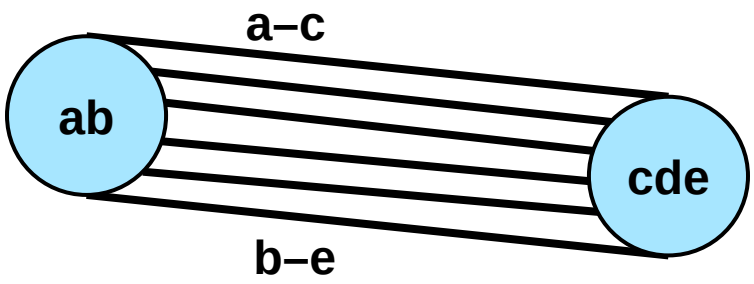


- Thus, as far as the LACP PDUs are concerned, this is the topology.
- Assuming they are using CFM, the two Nodes **ab** and **cde** have direct knowledge of the current state of the topology directly, and they agree on that state within 11 2/3 ms.

LACP vs. “Routing Protocols”



Routing Protocol



LACP

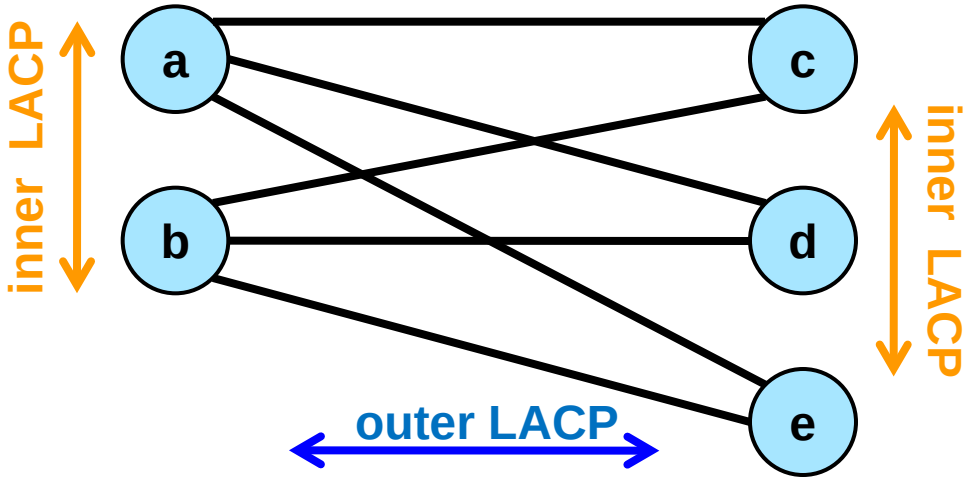
- Using the left diagram, PDUs are required to pass state information from Nodes **a** and/or **c** to Nodes **b** and **e**.
- Using the right diagram, the two Nodes have full knowledge without explicitly passing any Link or Node state from Node to Node.
- That is (this author’s) definition of LACP vs. Routing.

Routing is required. Why LACP?

An LACP solution

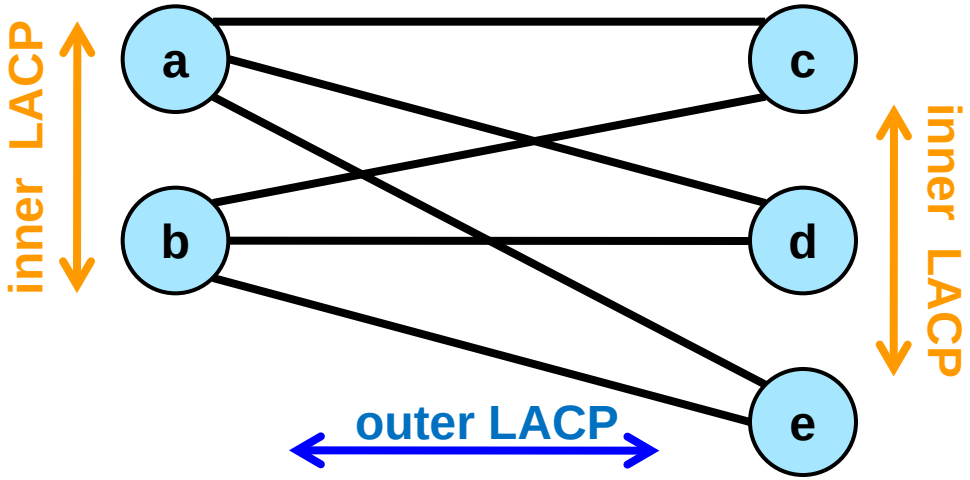
- No amount of abstraction can alter the fundamental fact that the physical topology illustrated (or, in fact, any topology useful to the NNI problem) **demands** a flow of information equivalent to **a routing protocol**.
- One reasonable approach, therefore, is to define (or simply select) some routing protocol suitable to the NNI problem.
- However, there exist implementations that can communicate state information among the members of a Virtual Node (**a–b** or **c–d–e**) much more efficiently than by exchanging protocol PDUs.
- This fact makes the LACP / Virtual Node fiction useful.

An LACP-based solution



- Suppose we divide the routing information flows into two parts, the **outer part** and the **inner part**.
- The **outer part** flows between Virtual Nodes.
- The **inner part** flows inside Virtual Nodes.

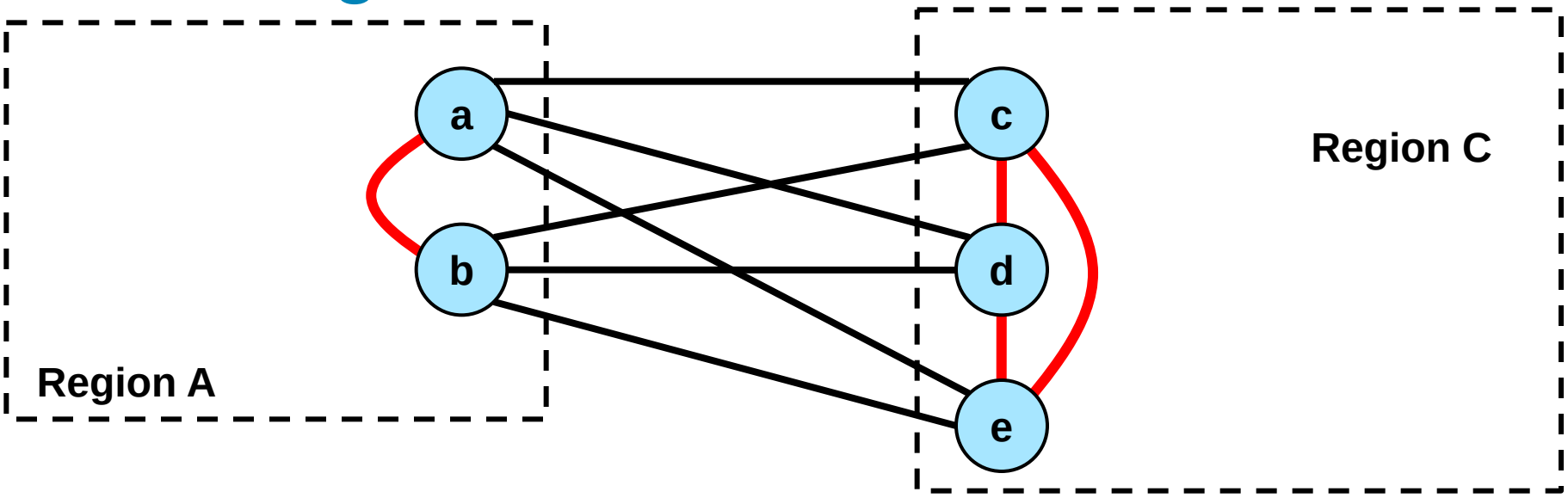
Inner LACP



- If we define the routing protocol in these terms (thanks, Don) then whether the components of a given Virtual Node actually use the inner LACP or some other means to convey the necessary information can be optional.
- That is, the actual use of inner LACP can be optional.

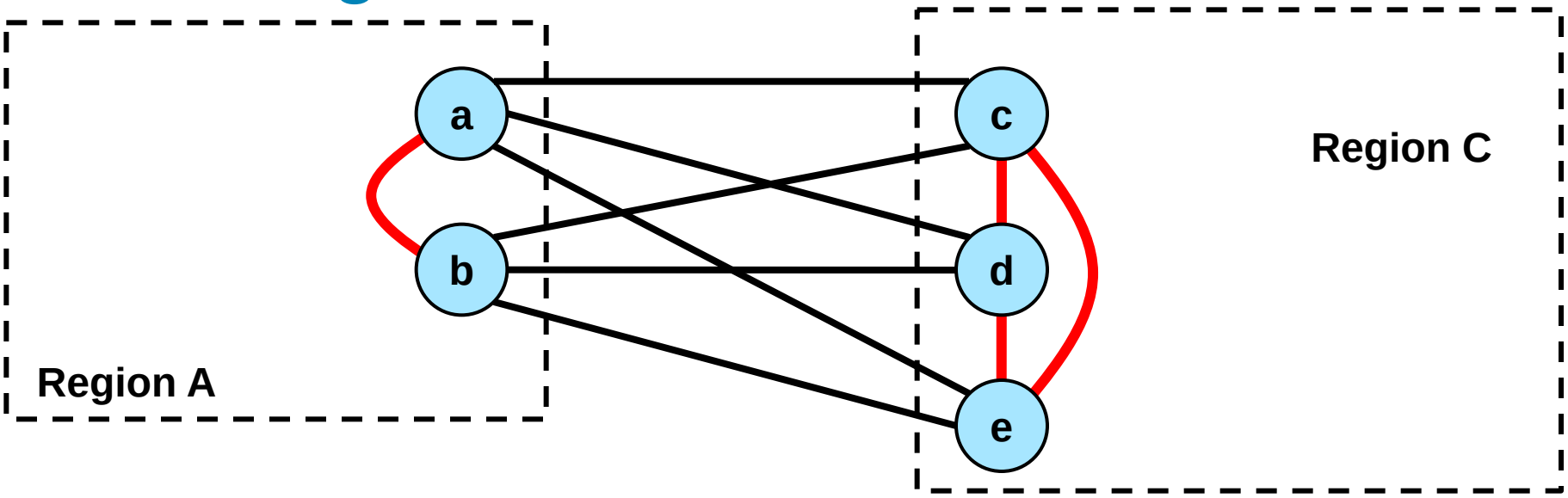
What do we give up by using LACP instead of routing?

Routing vs. LACP



- If the red Links are included in the topology considered by the protocol that is performing fault recovery in the interconnect, then this is routing, not LACP; LACP can consider only the black Links.

Routing vs. LACP



- If only the black links are considered, and Virtual Nodes are not implemented, then a failure in the interconnect seems to force a gateway change.
- E.g., if Link **a**–**b** is the primary Link for a Service, and it fails and Link **b**–**e** takes over, then if **c**, **d**, and **e** are separate Nodes, the Region C Gateway, by definition, has moved from **c** to **e**.
- Is this true, or is there a fix?