Scalable Routing with Deep Reinforcement Learning

Ramy E. Ali [1,2] , Bilgehan Erman [2], Ejder Baştuğ [2] and Bruce Cilli [2]

[1]Penn State, [2] IP Networking, ENSA Lab, Nokia Bell Labs

Department Head: Sameer Sharma

October 23, 2019

Unix 50

Unix started it all. Ready for the next computing revolution?

Unix Philosophy

Unix Approach: Do one thing and do it well

Unix Philosophy

Unix Approach: Do one thing and do it well

Write programs to work together

Unix Philosophy

Unix Approach: Do one thing and do it well
Write programs to work together

Unix Success: Shift from centralized mainframes to small decentralized computers*

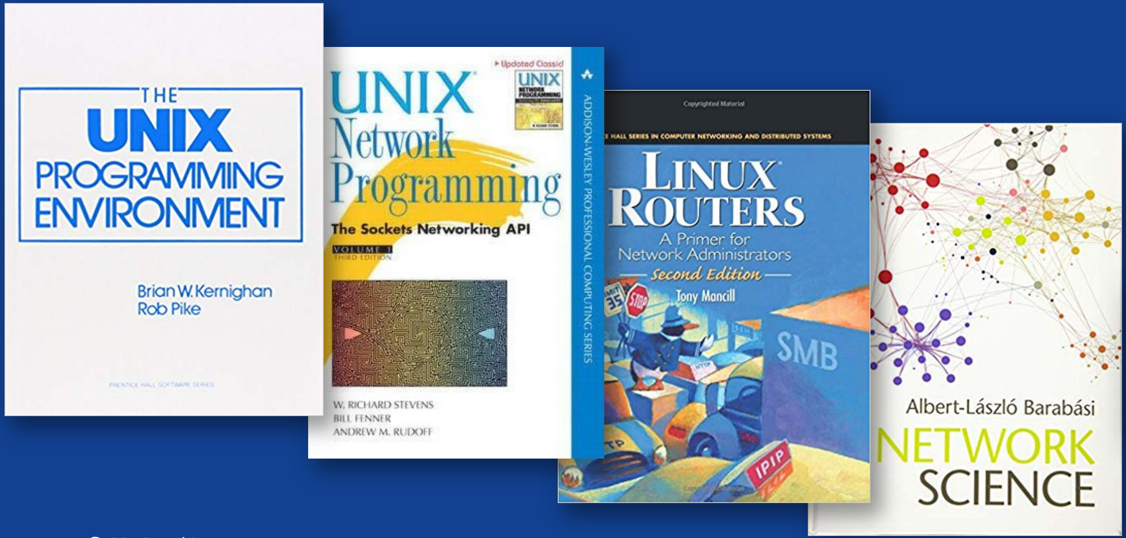
*"Unix and Beyond: An Interview with Ken Thompson"

Unix Philosophy

Unix Approach: Do one thing and do it well

Write programs to work together

Unix Success: Shift from centralized mainframes to small decentralized computers*



Large-Scale Networking

Programmable Network OS (PNOS)

*“Unix and Beyond: An Interview with Ken Thompson”

Routing Challenges in Large-Scale Networks

Route Computation Complexity ⇨ **Scalability**

Routing Challenges in Large-Scale Networks

Route Computation Complexity → Scalability

Service Diversity (Delay, rate, …) → **Large State Space**

Routing Challenges in Large-Scale Networks

Route Computation Complexity	⇨	Scalability
Service Diversity (Delay, rate, …)	⇨	Large State Space
Demand Uncertainty and Dynamism	⇨	**Need for Adaptive Routing**

Prior Art

Distributed Approaches (e.g., BGP)

Prior Art

Distributed Approaches (e.g., BGP)

(No Global Overview)

Prior Art

Distributed Approaches (e.g., BGP)

(No Global Overview)

Centralized Approaches (e.g., PCE)

Prior Art

Distributed Approaches (e.g., BGP)
(No Global Overview)

Centralized Approaches (e.g., PCE)
(Global View)

Prior Art & Our Approach

Distributed Approaches (e.g., BGP)
(No Global Overview)

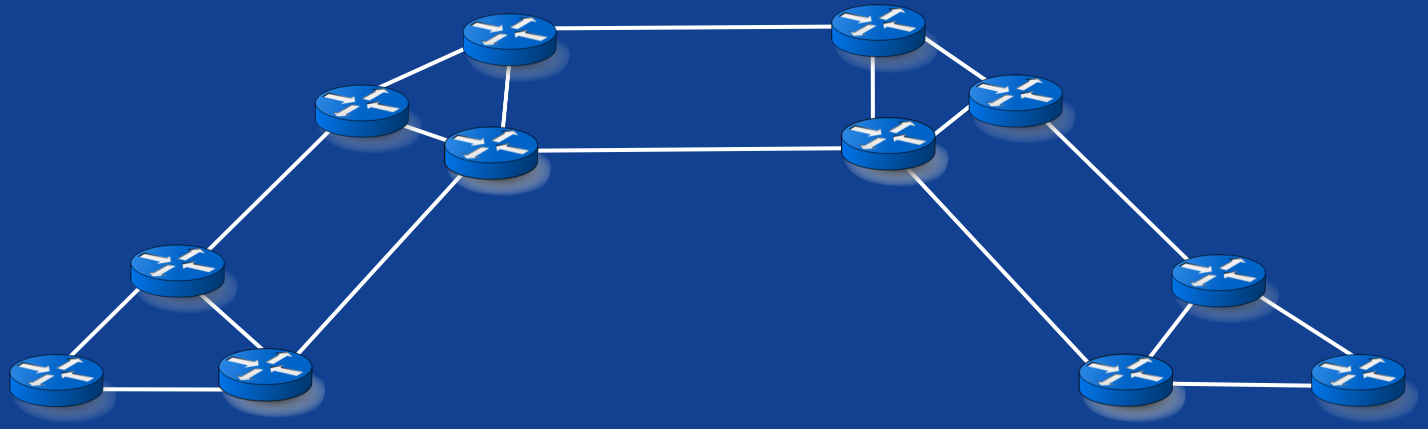
Centralized Approaches (e.g., PCE)
(Global View)

Our Approach: Cluster Oriented Scalable Routing

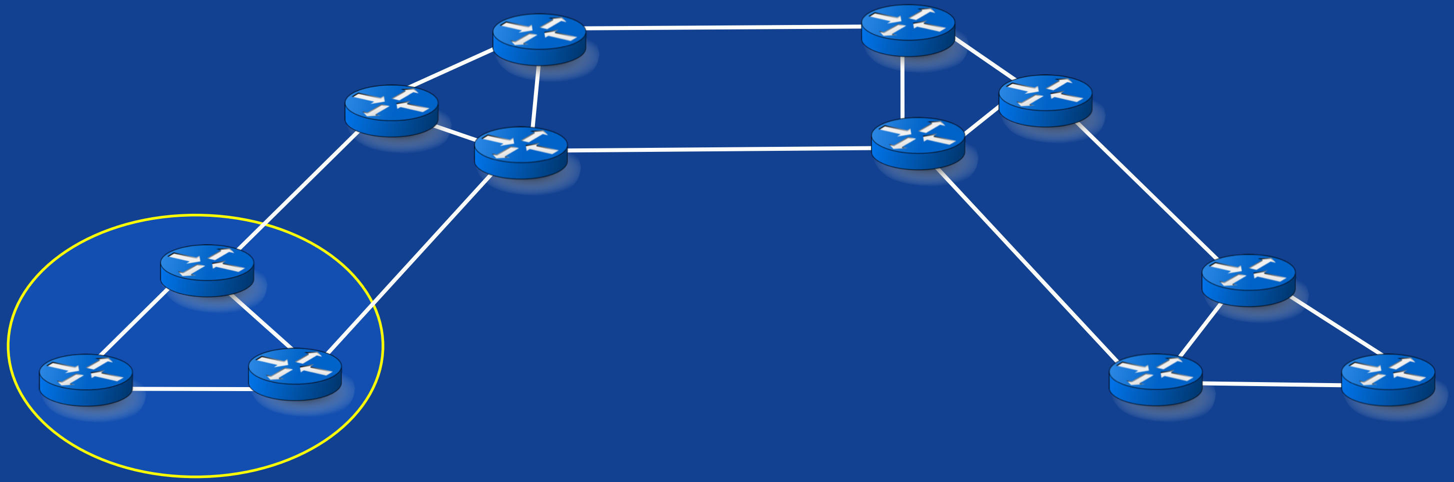
Distributed with Global View

Patent pending

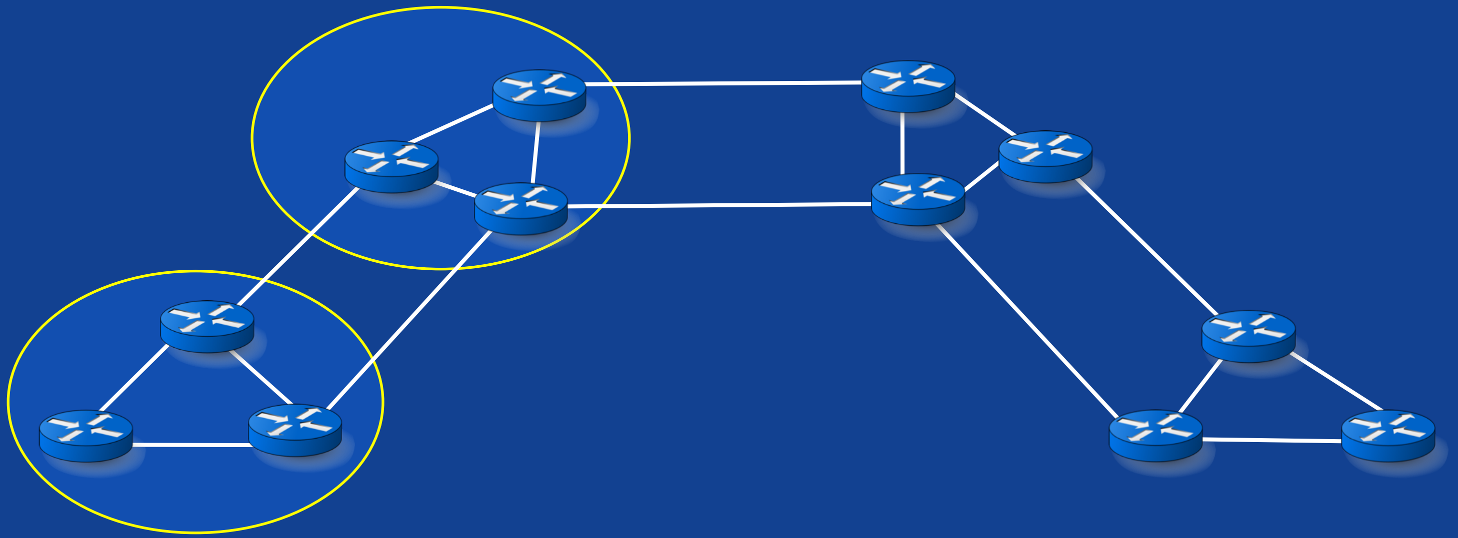
Cluster Oriented Scalable Routing



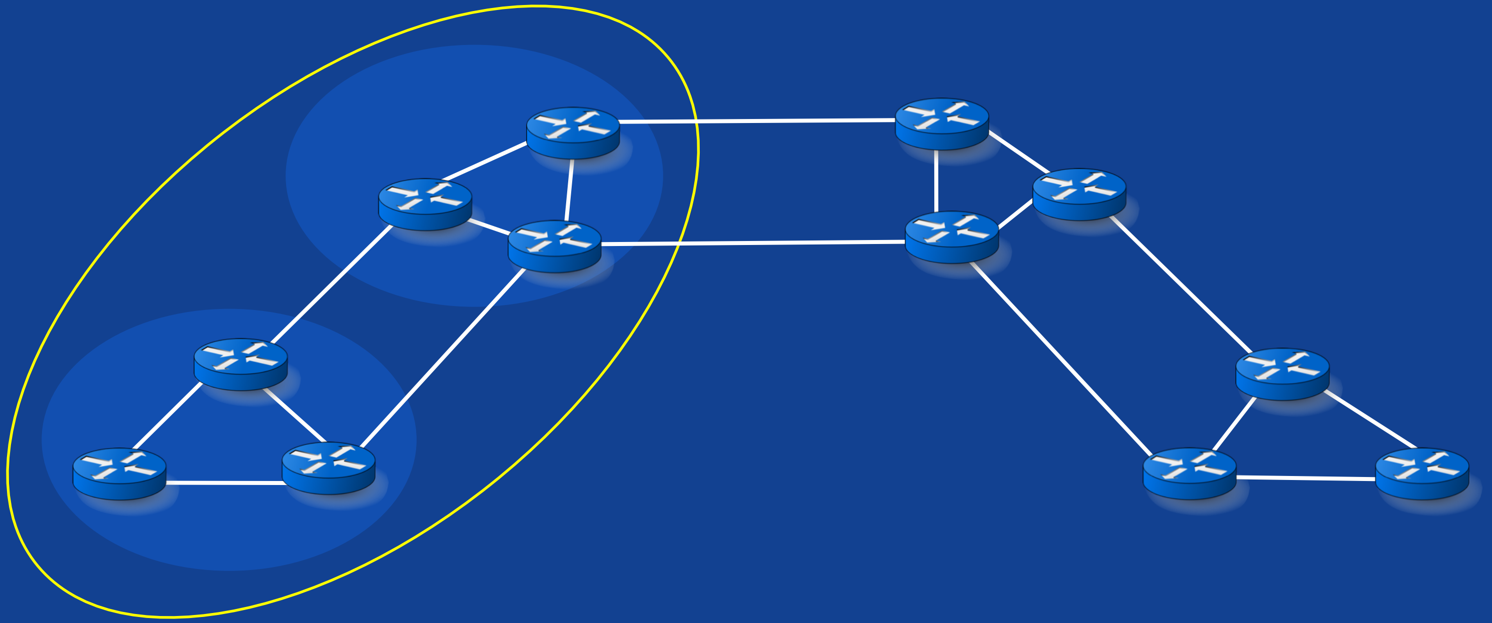
Hierarchical Clusters



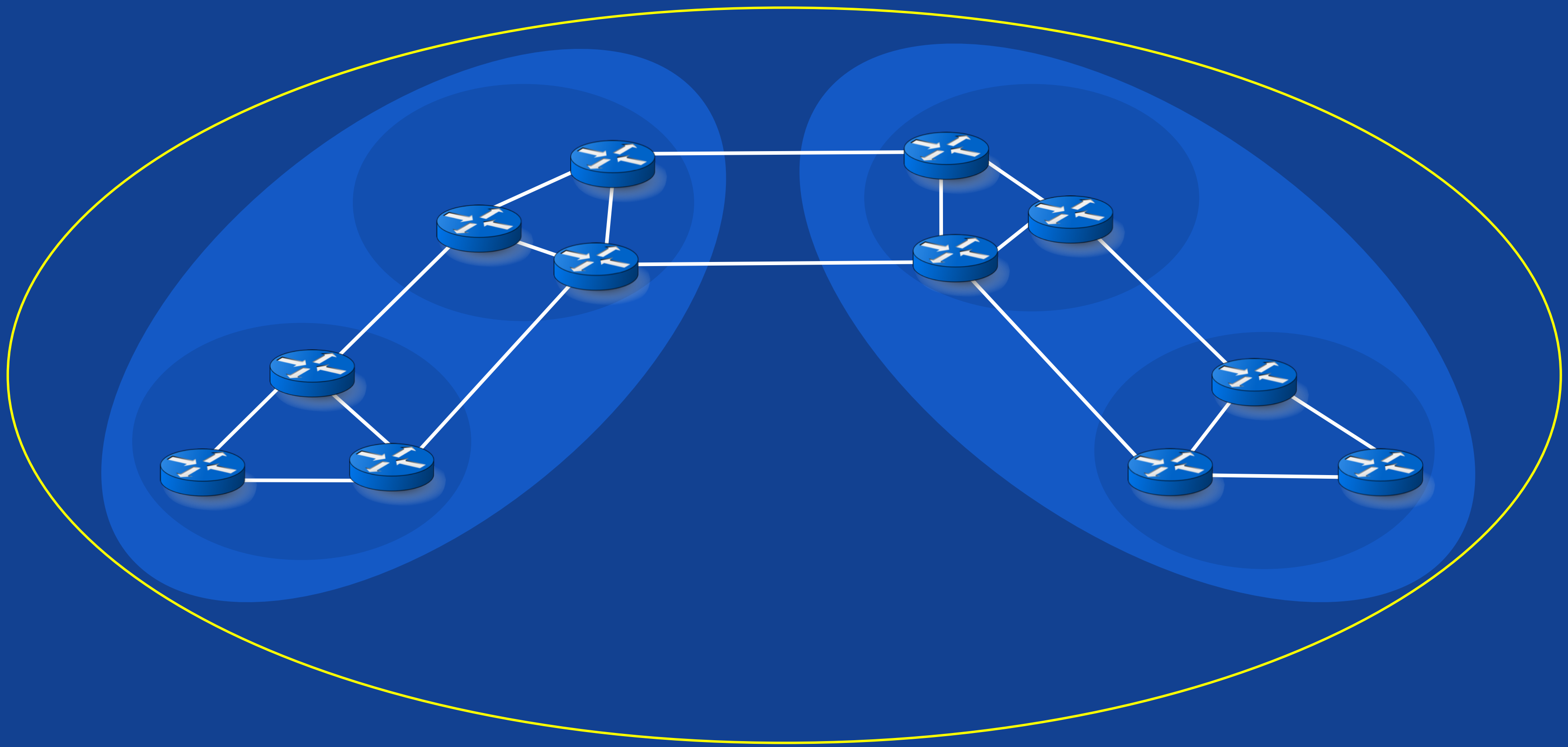
Hierarchical Clusters



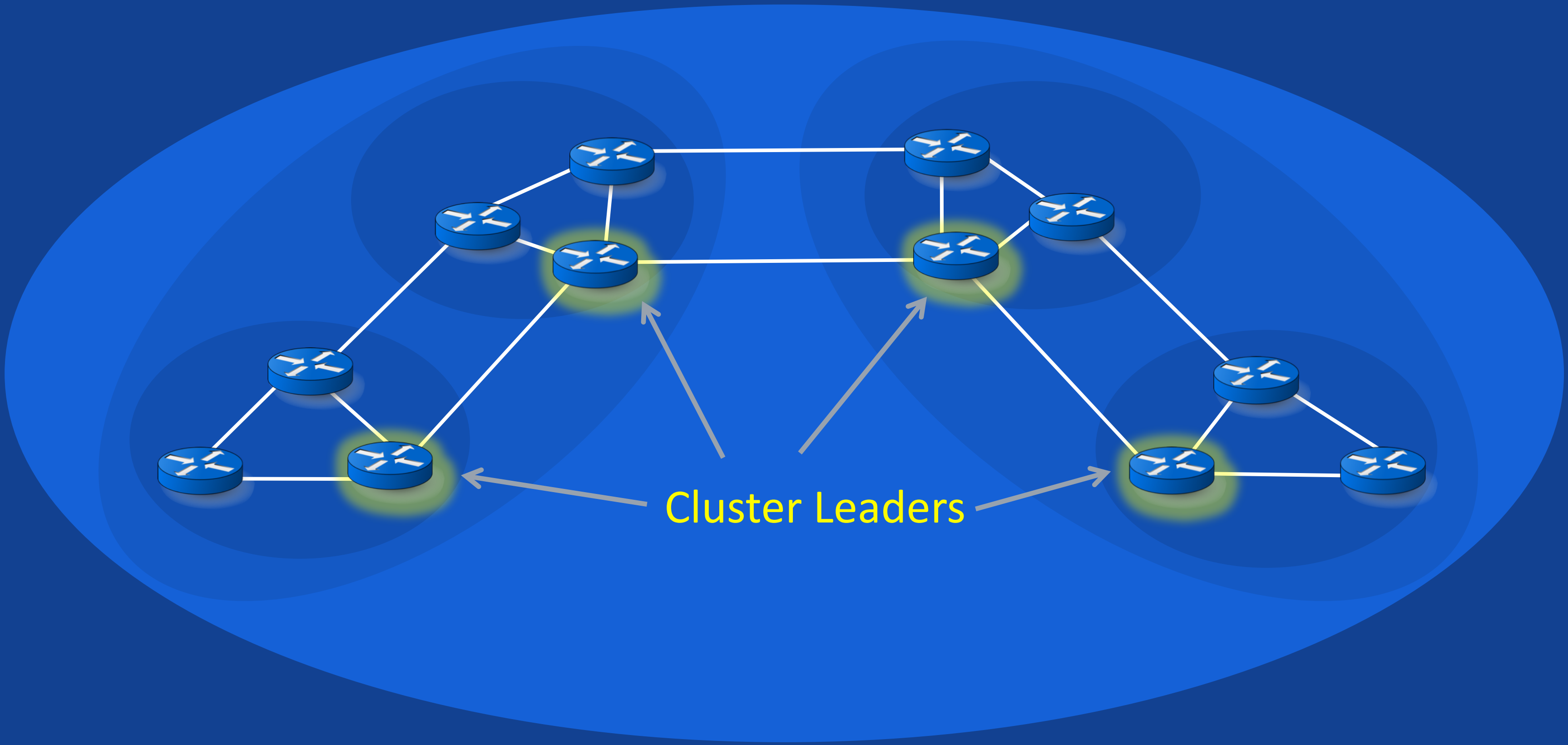
Hierarchical Clusters



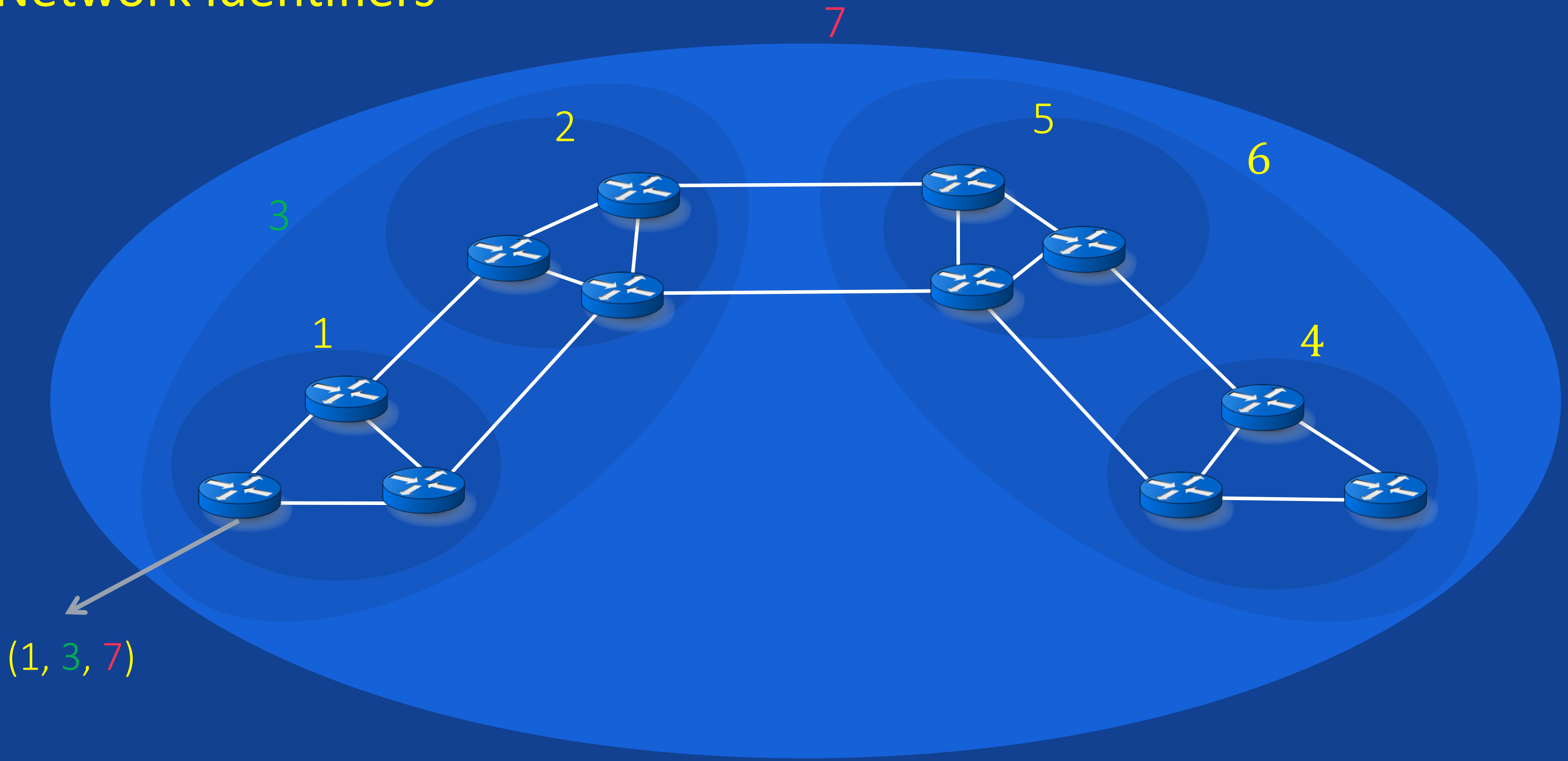
Hierarchical Clusters



Hierarchical Clusters

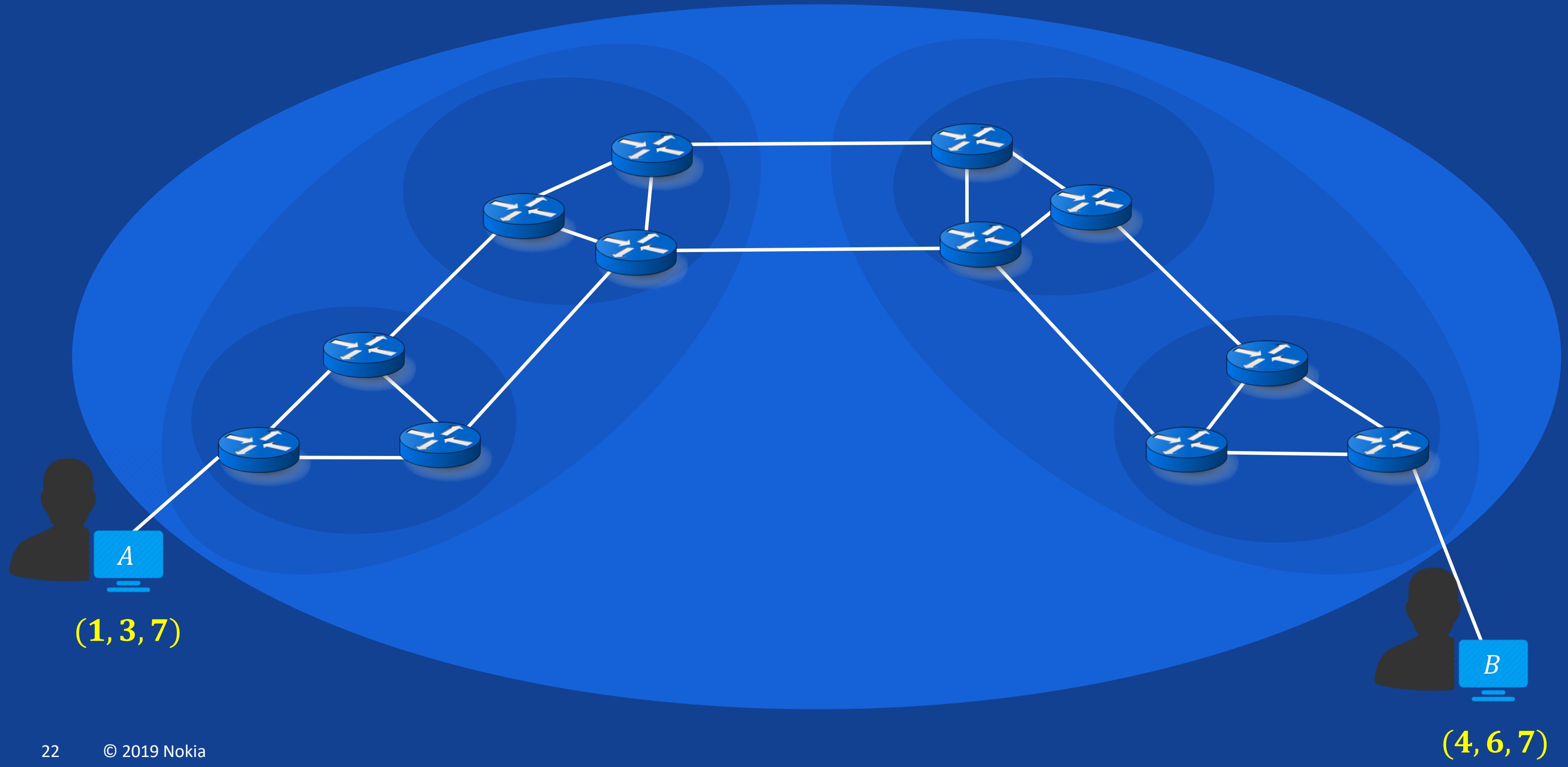


Network Identifiers

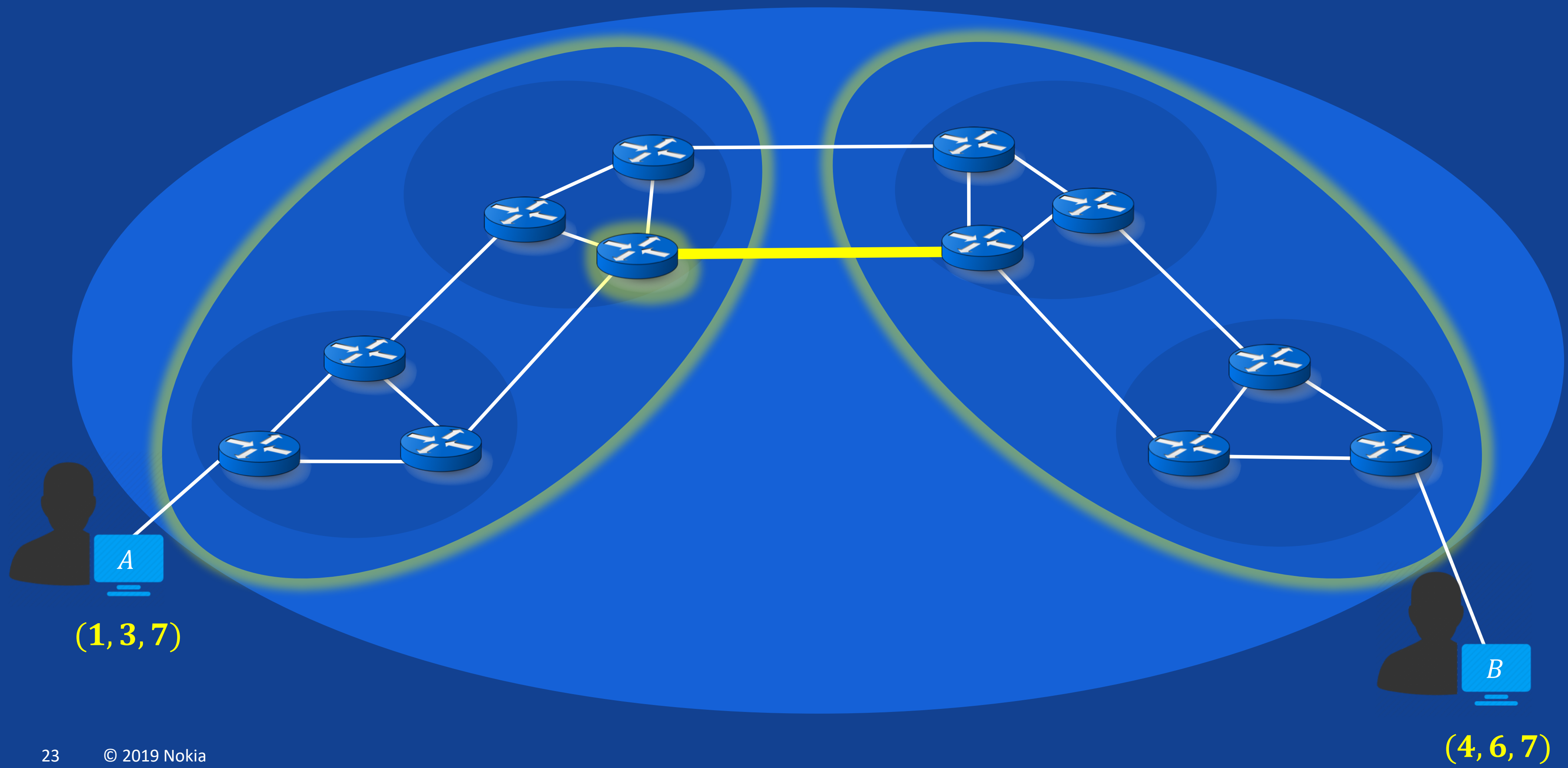


Route Computation

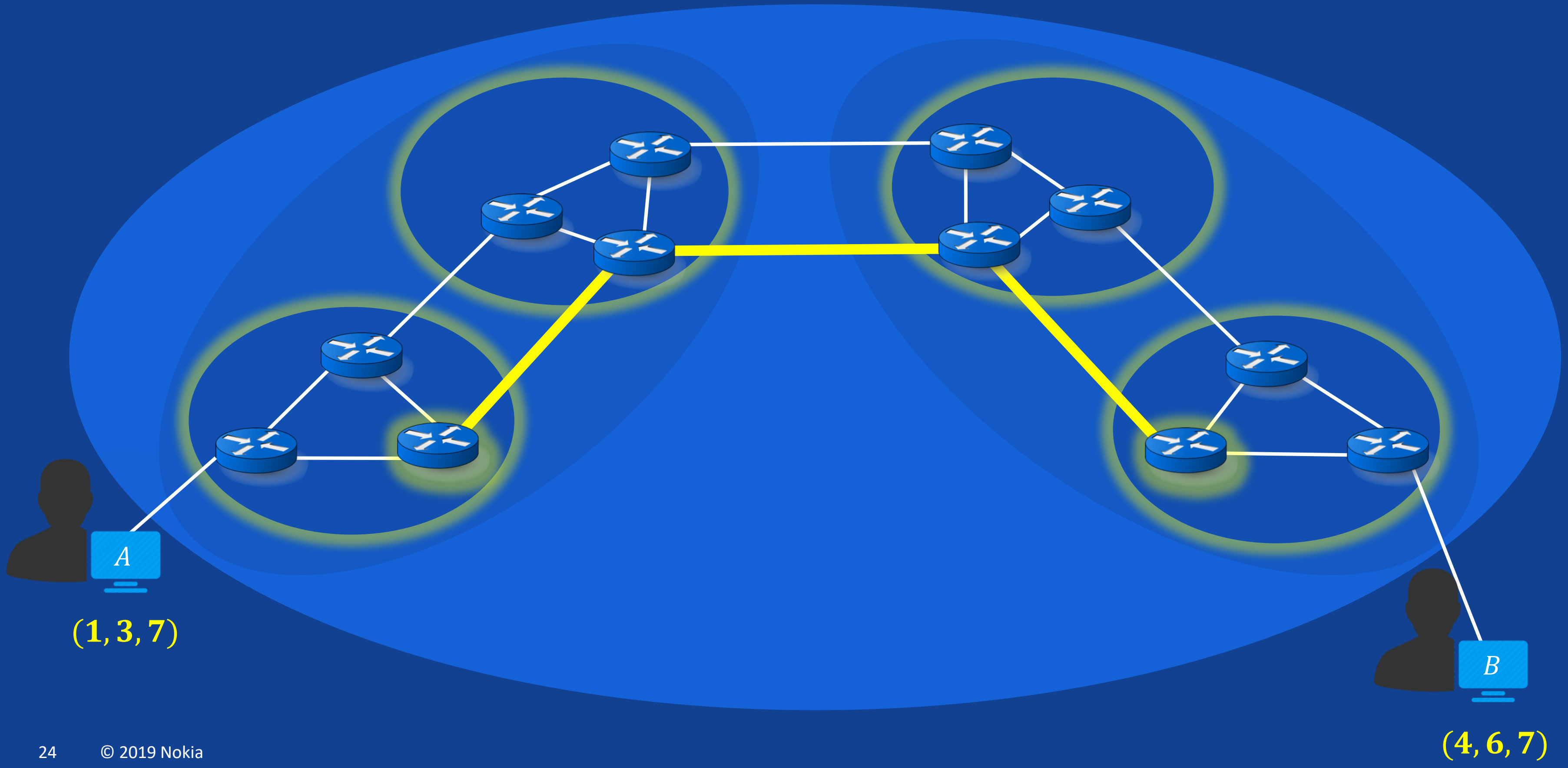
Leader Assistance



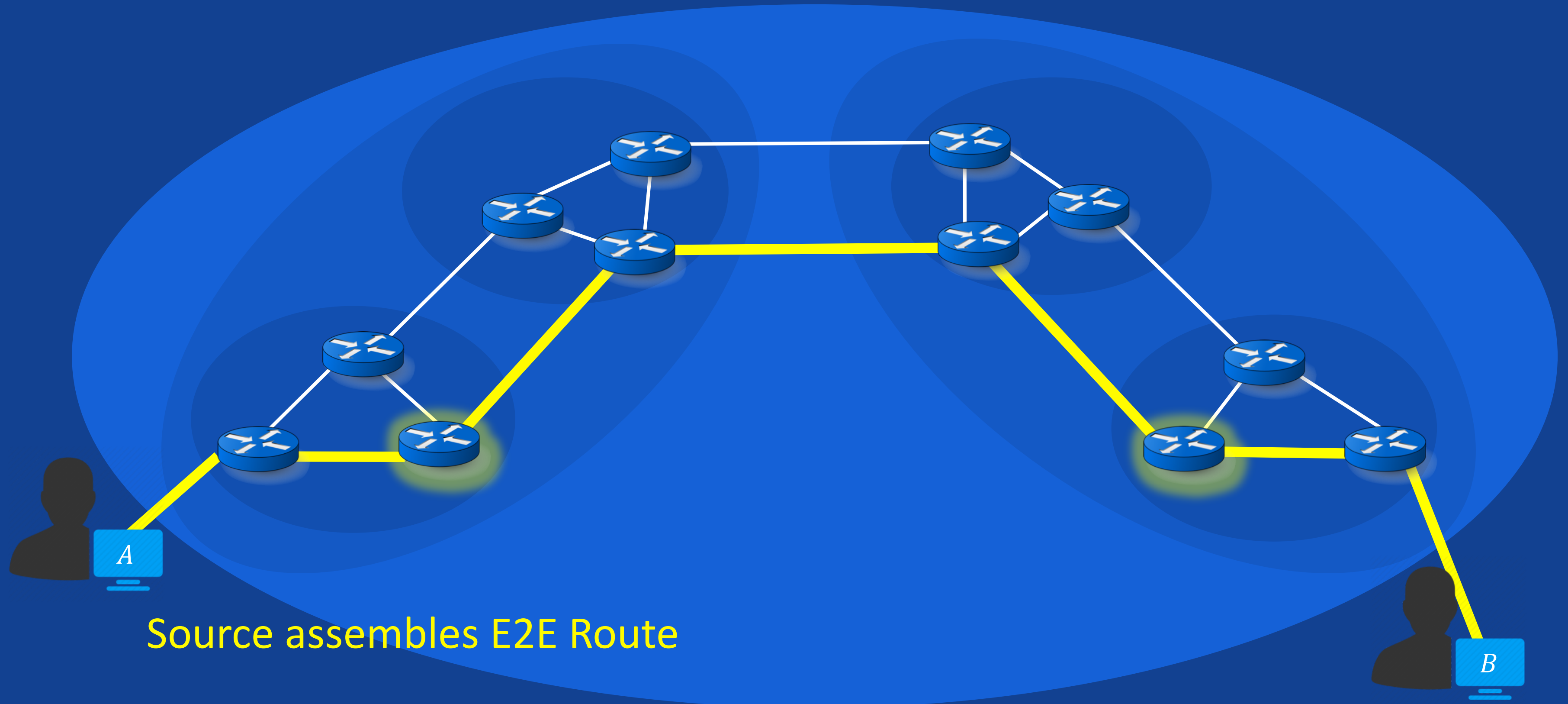
Leader Assistance



Leader Assistance

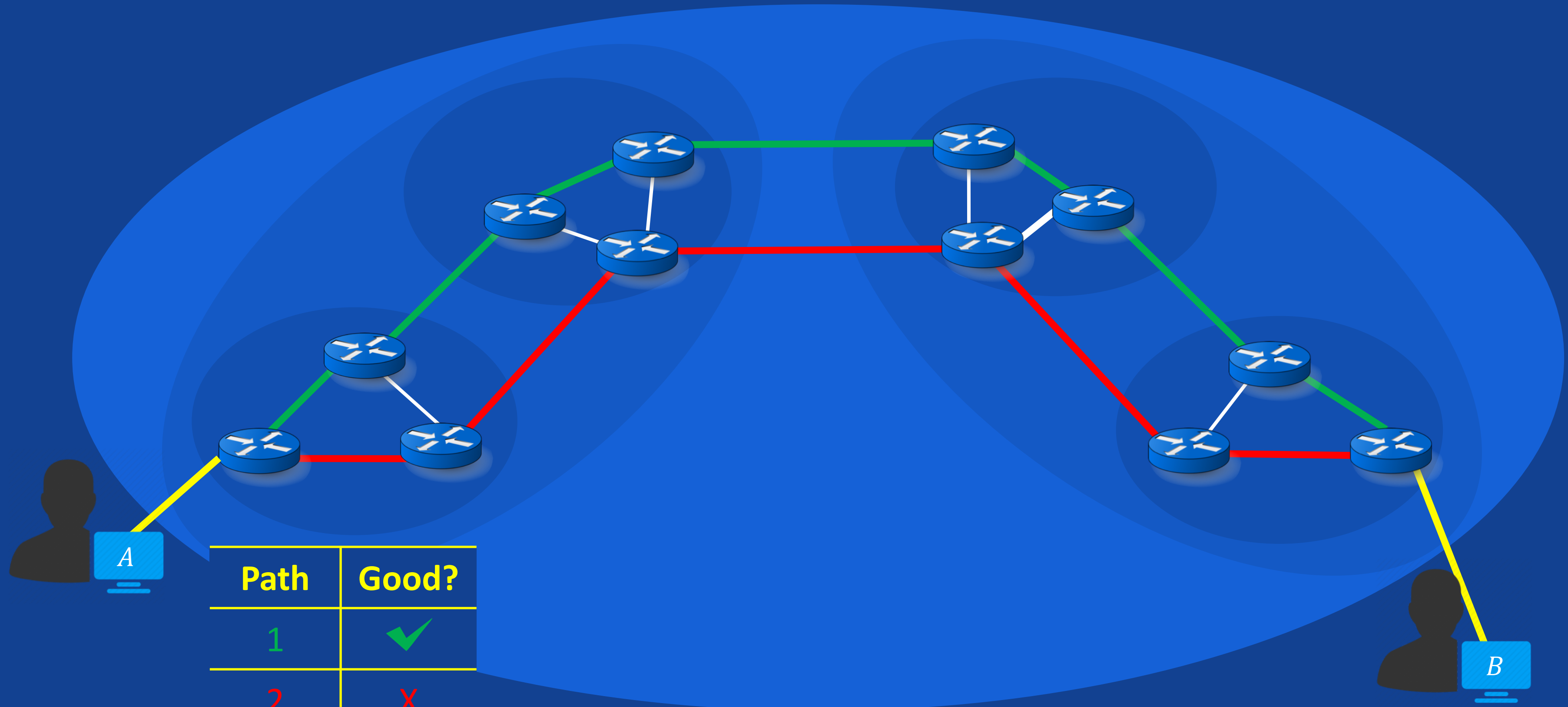


Segment Routing



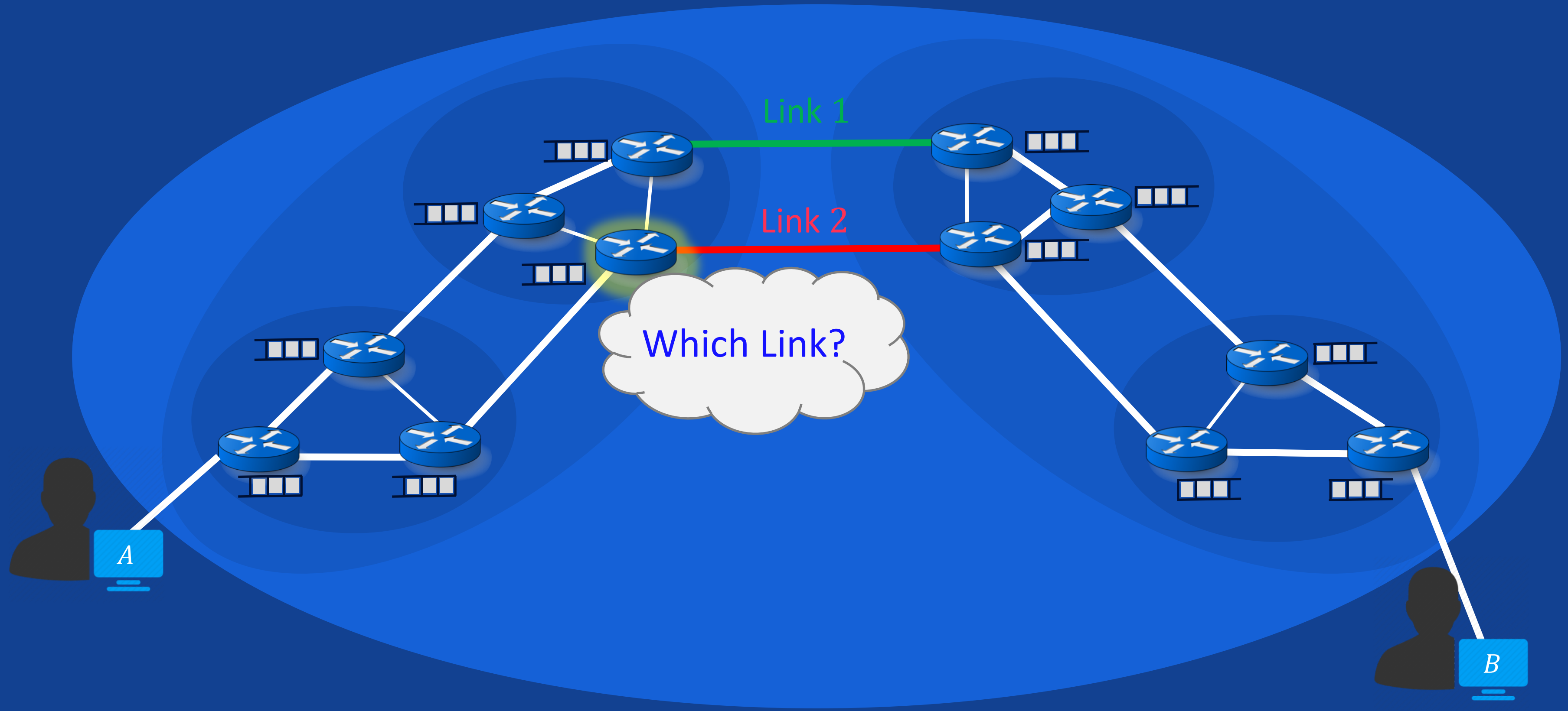
Source assembles E2E Route

Source Intelligence (Global View)

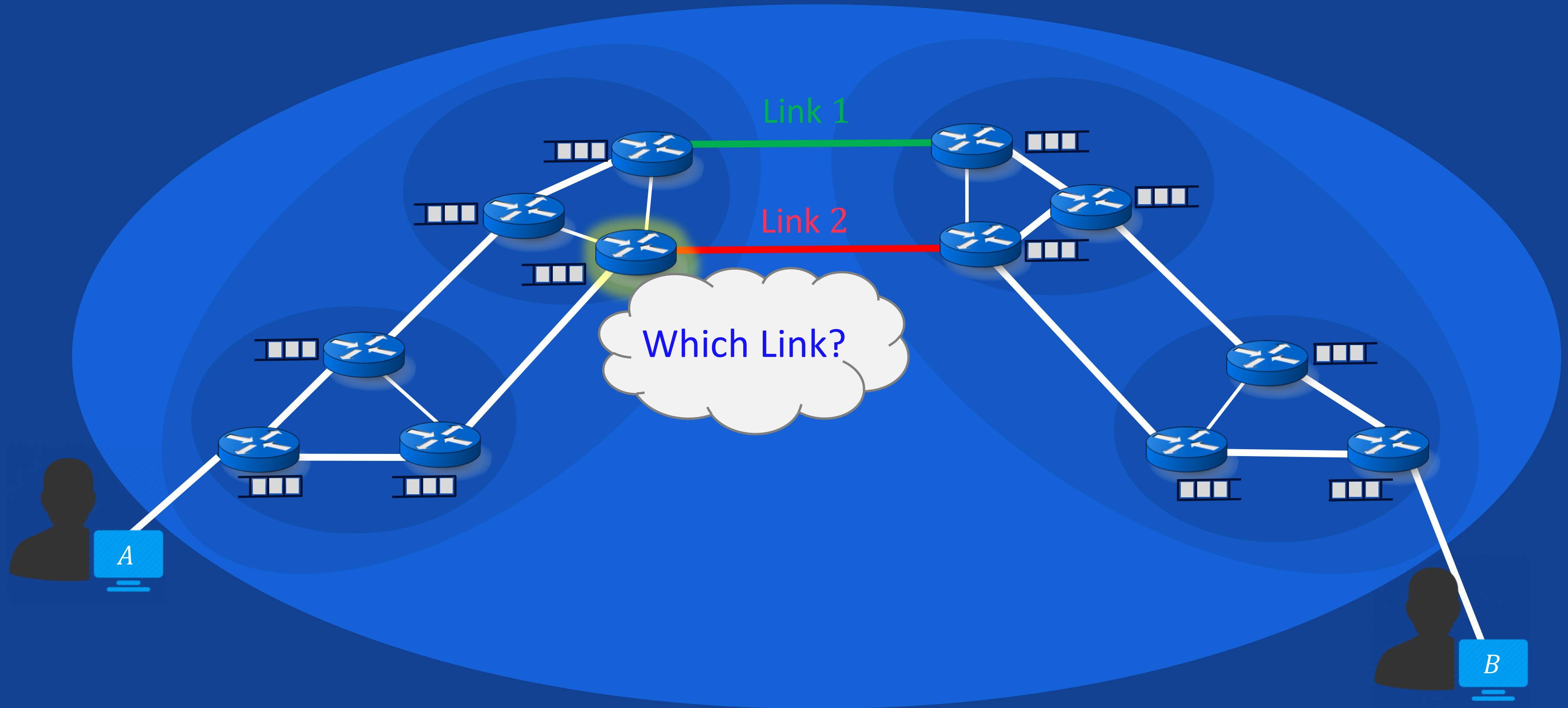


Path	Good?
1	✔
2	X

Based on transmission delays, queuing delays, utilizations, ….



Deep Reinforcement Learning Approach



Reinforcement Learning

Powerful

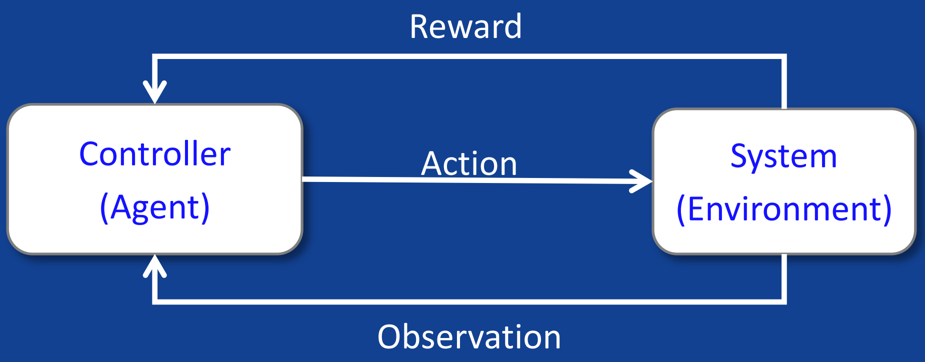
DeepMind's AI beats world's best Go player in latest face-off

Ke Jie, who once boasted he would never be beaten by a computer at the ancient Chinese game, said he had 'horrible experience'

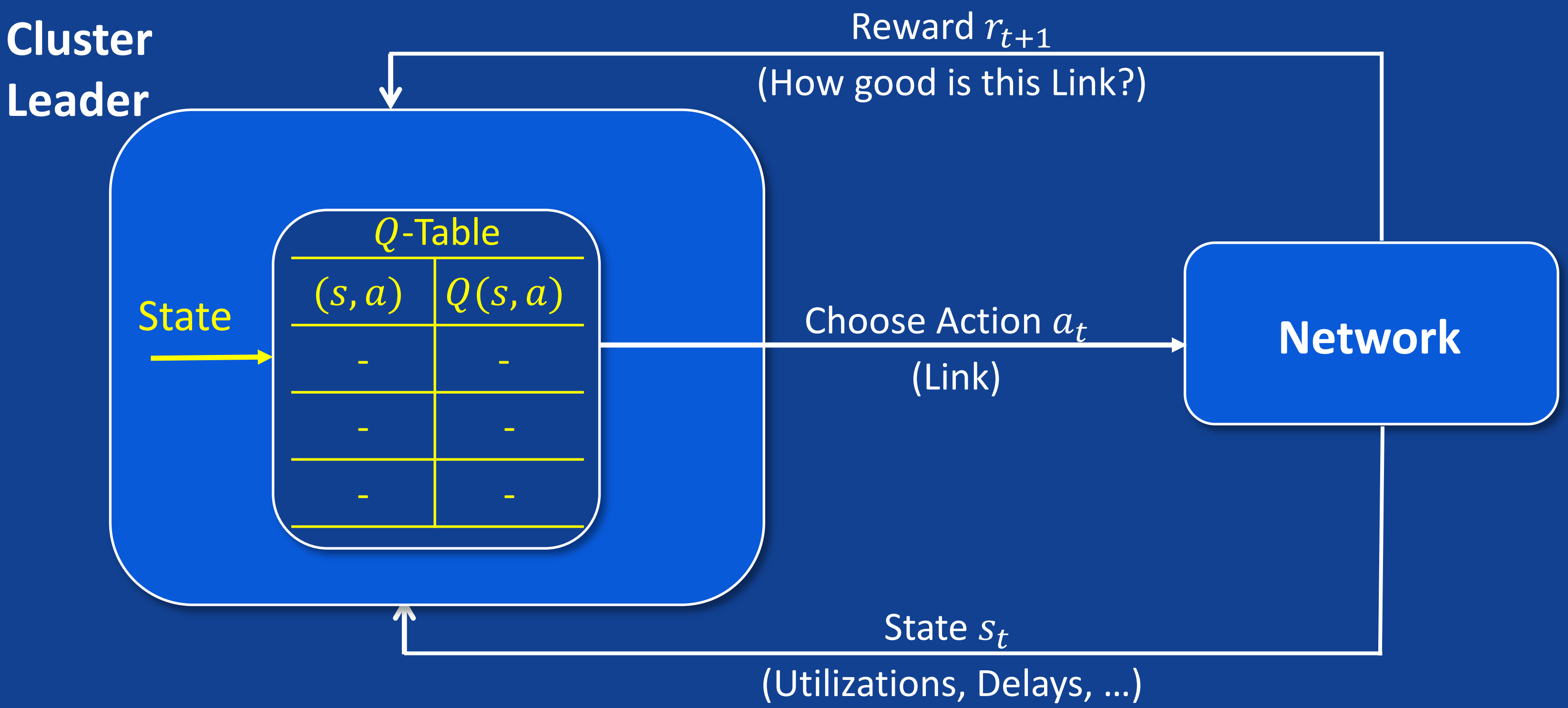
Challenging

Convergence is not universally guaranteed

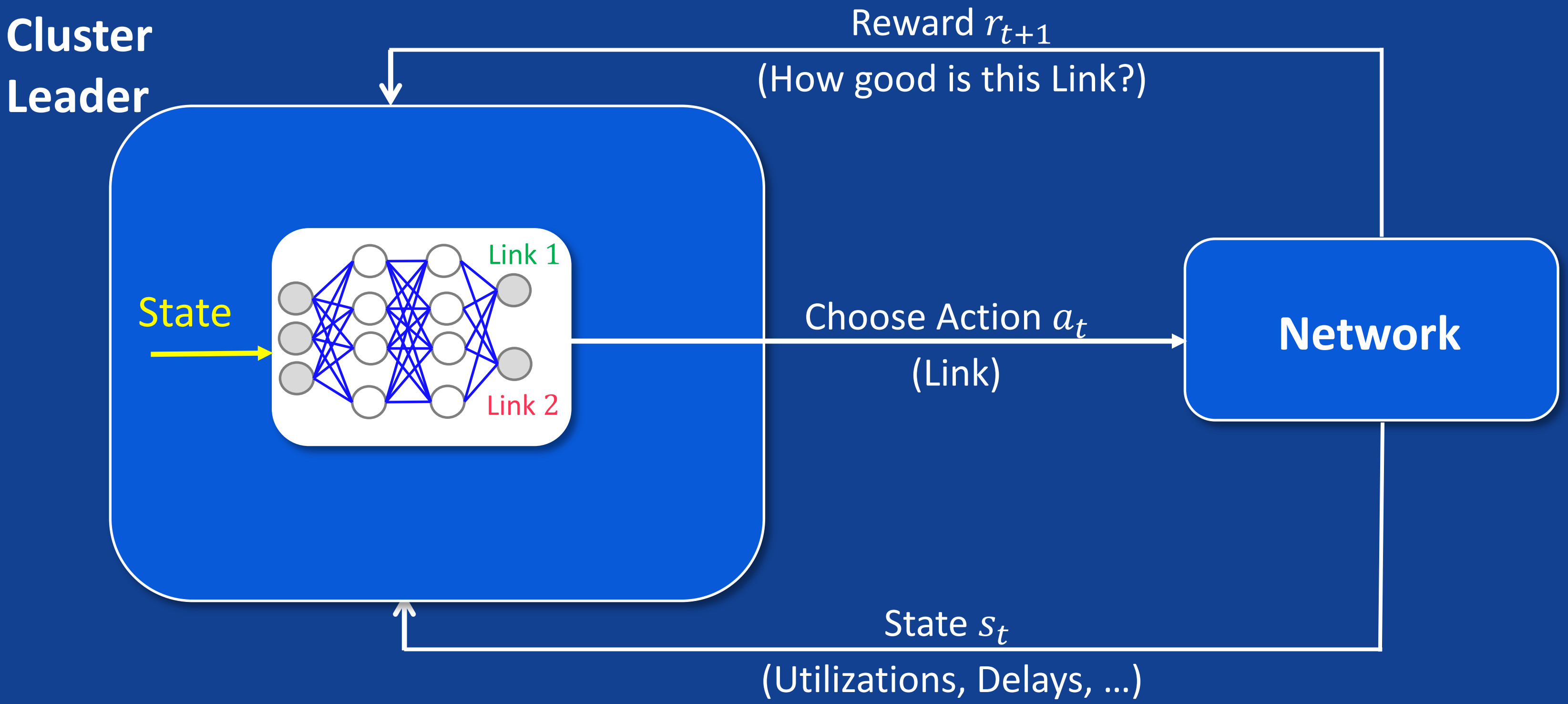
Needs time to experiment – like a human



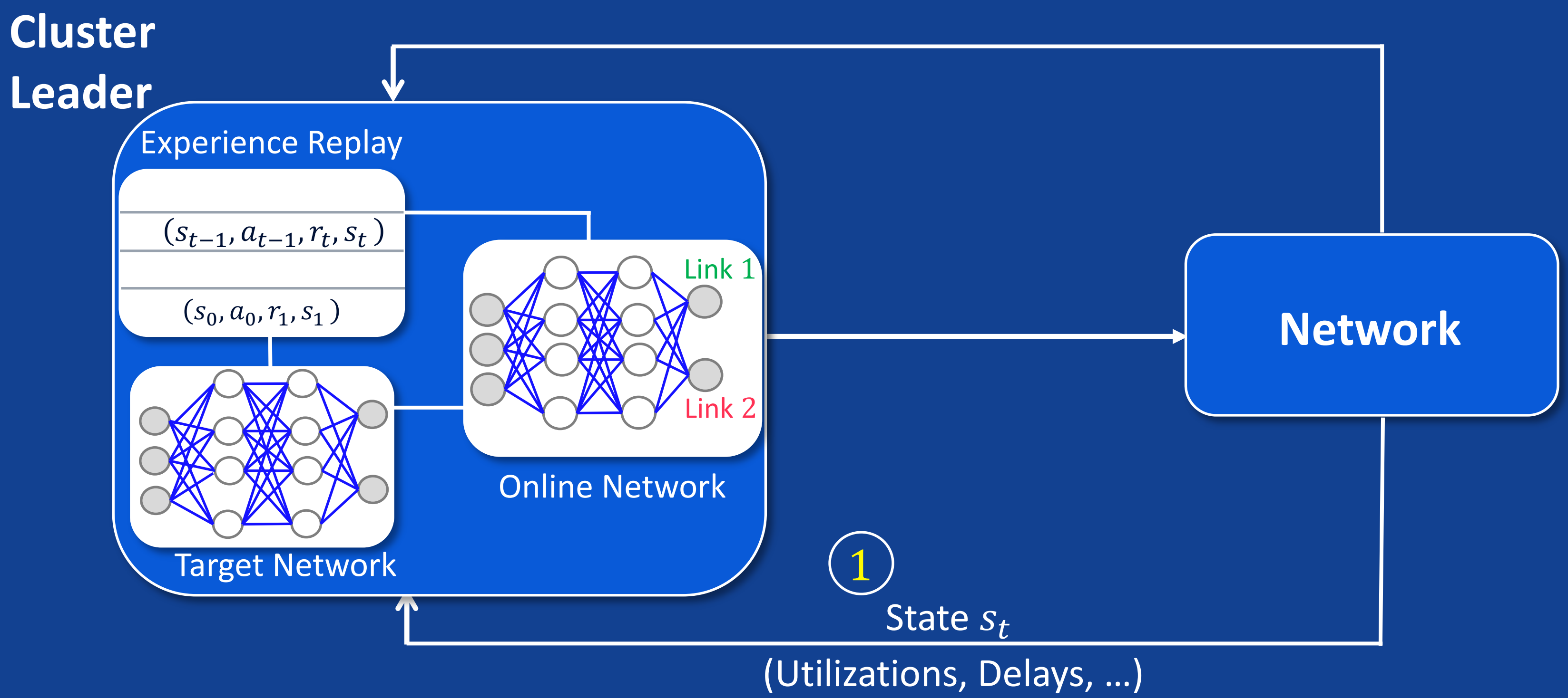
Q-Learning


Cluster
Leader
Q-Table
(s, a)
$Q(s, a)$
-
-
-
-
-
-
State
Reward r_{t+1}
(How good is this Link?)
Choose Action a_t
(Link)
Network
State s_t
(Utilizations, Delays, …)

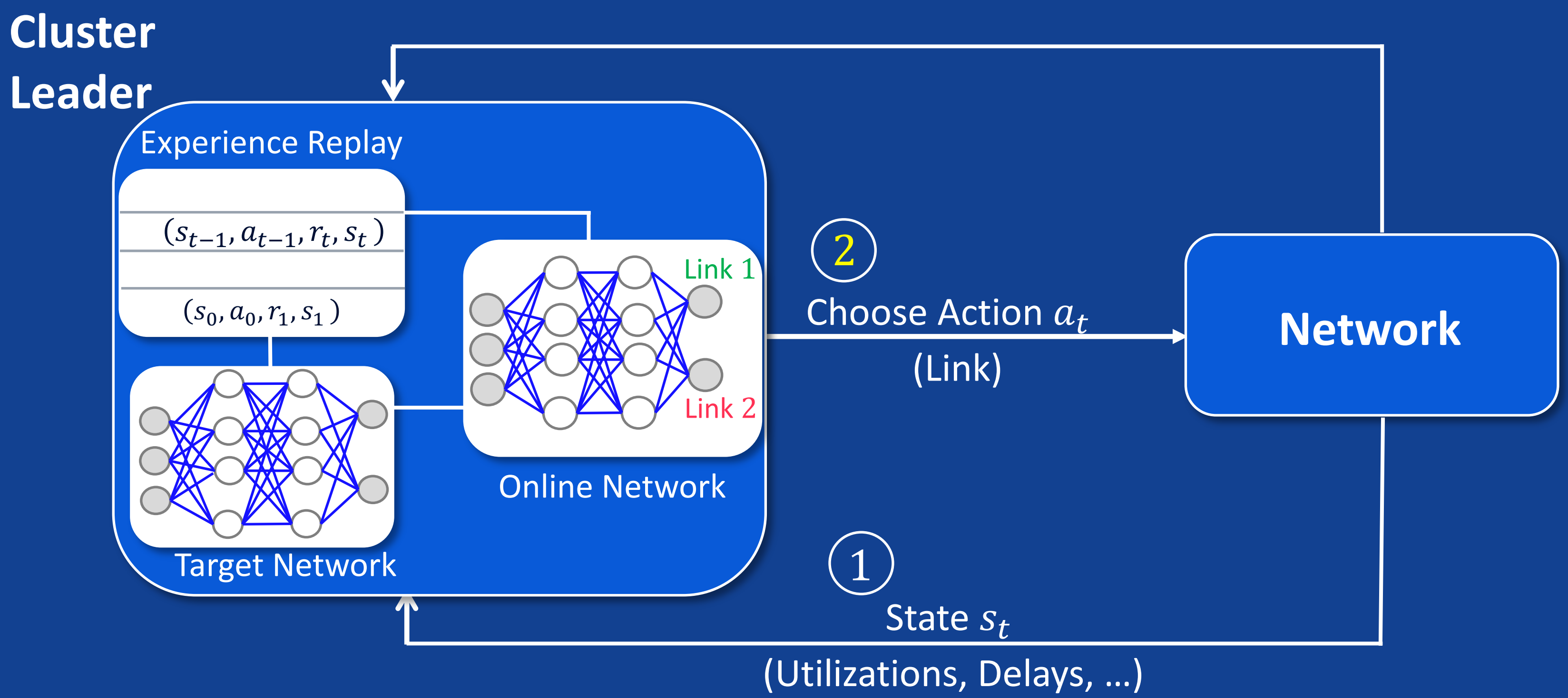
Deep Q-Learning



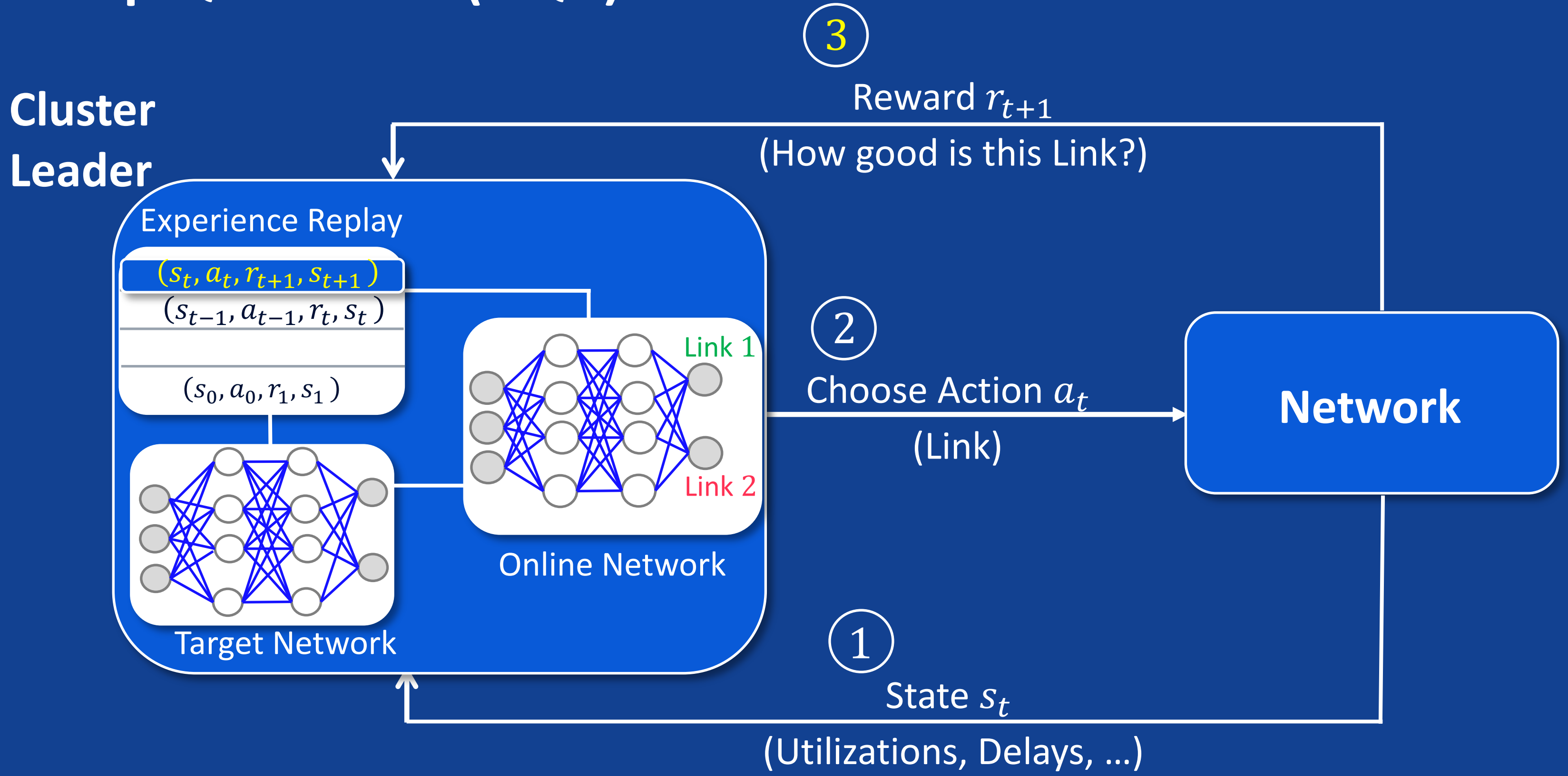
Deep Q-Network (DQN)



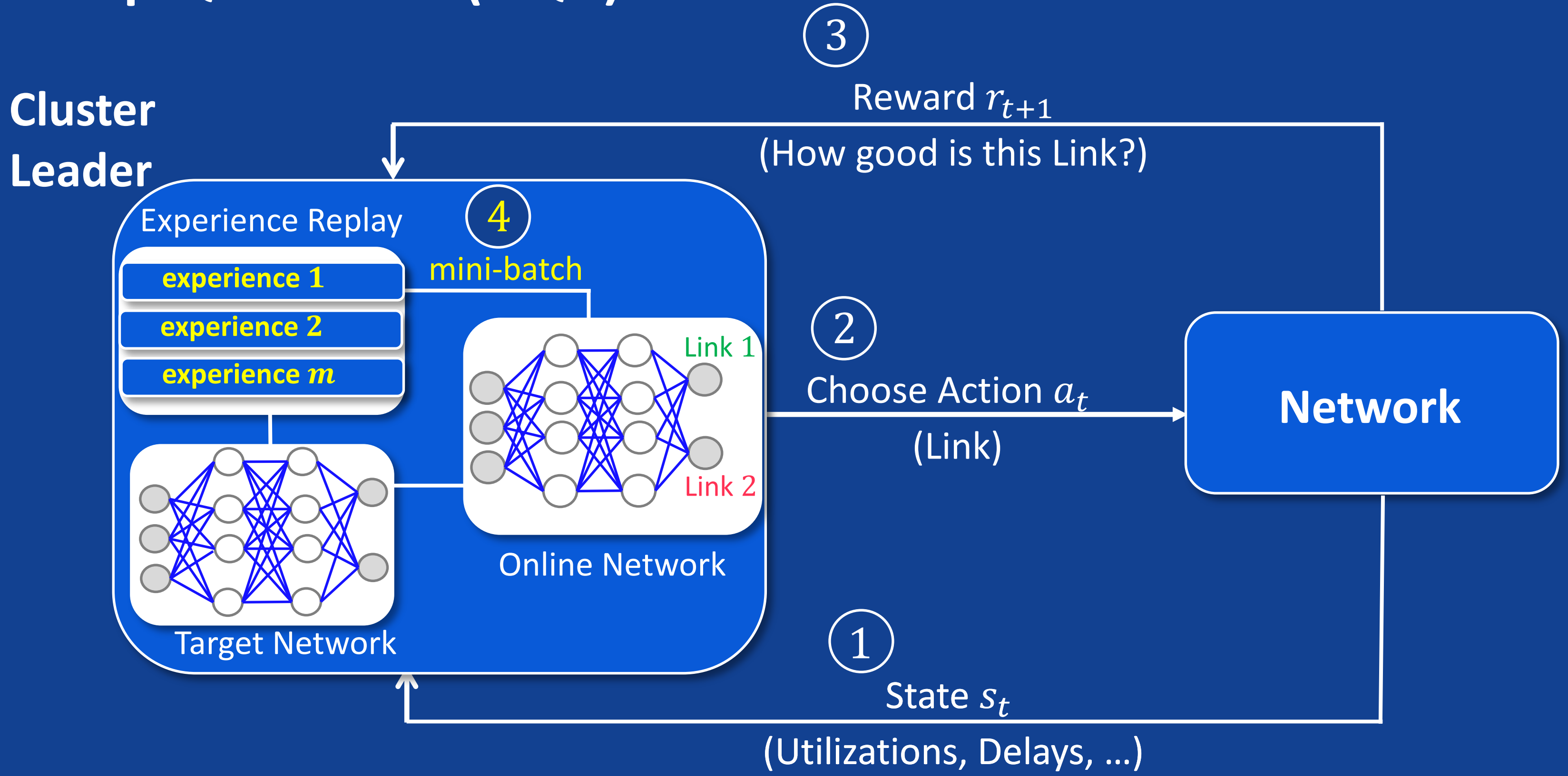
Deep Q-Network (DQN)



Deep Q-Network (DQN)

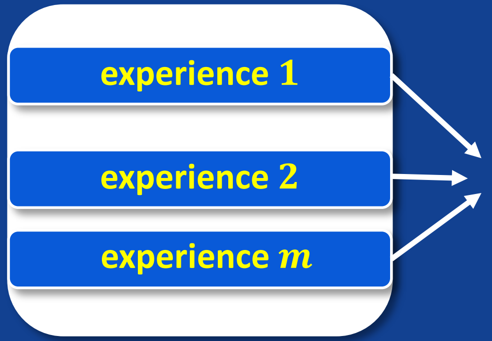


Deep Q-Network (DQN)



Experience Replay & Deep Double Q-Network (DDQN)

Prioritized Experience Replay

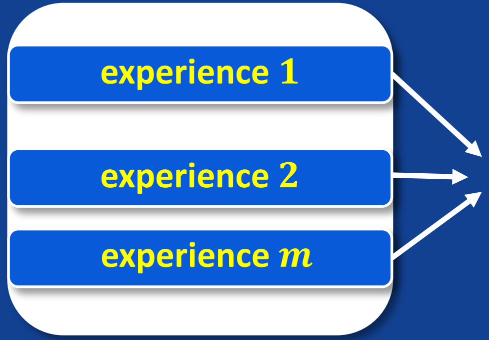


Sample
mini-batch

Priority $\propto$ How surprising?

Experience Replay & Deep Double Q-Network (DDQN)

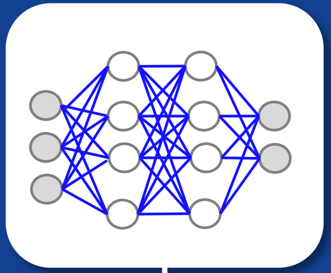
Prioritized Experience Replay



Sample mini-batch

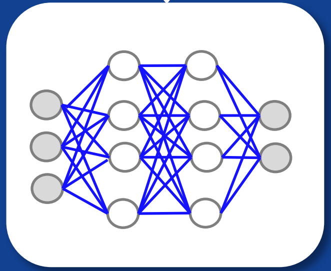
Priority $\propto$ **How surprising?**

Online Network



What's the best link?

Update every τ steps

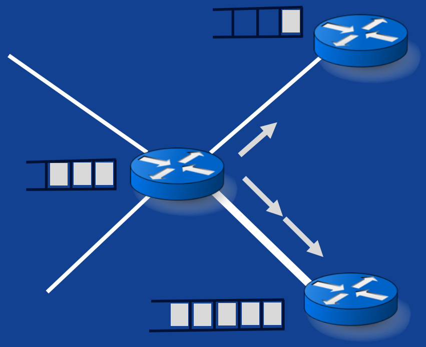


How good is this link?

Target Network

How to design the reward?

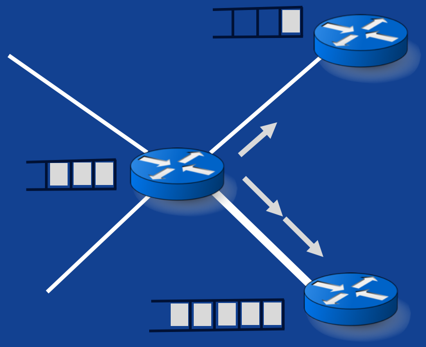
Local Reward

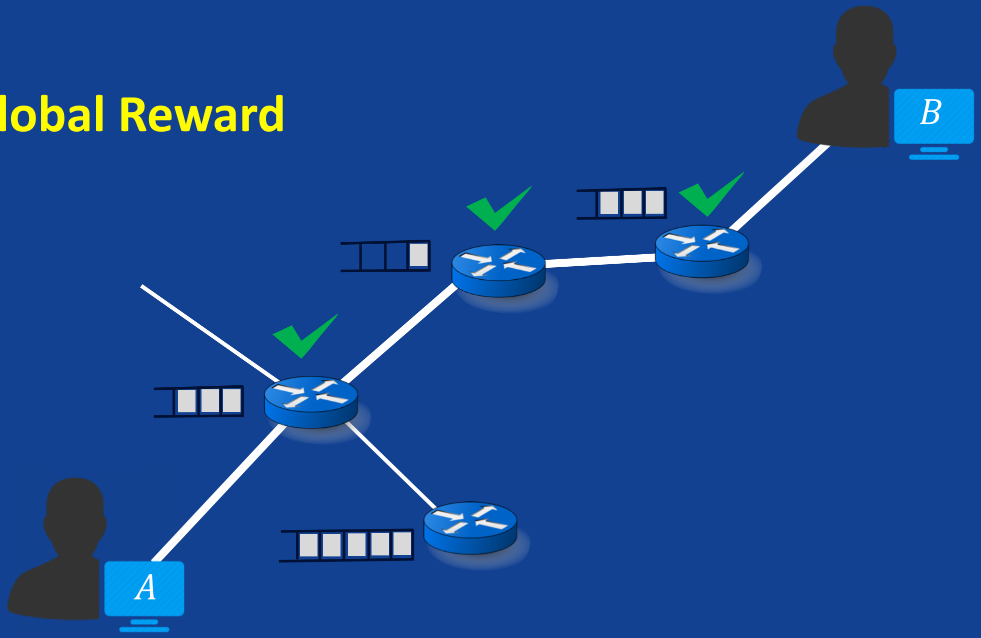


Assigned individually
(Distributing load, Delay, Loss, …)

How to design the reward?

Local Reward + **Global Reward**





Assigned individually
(Distributing load, Delay, Loss, …)

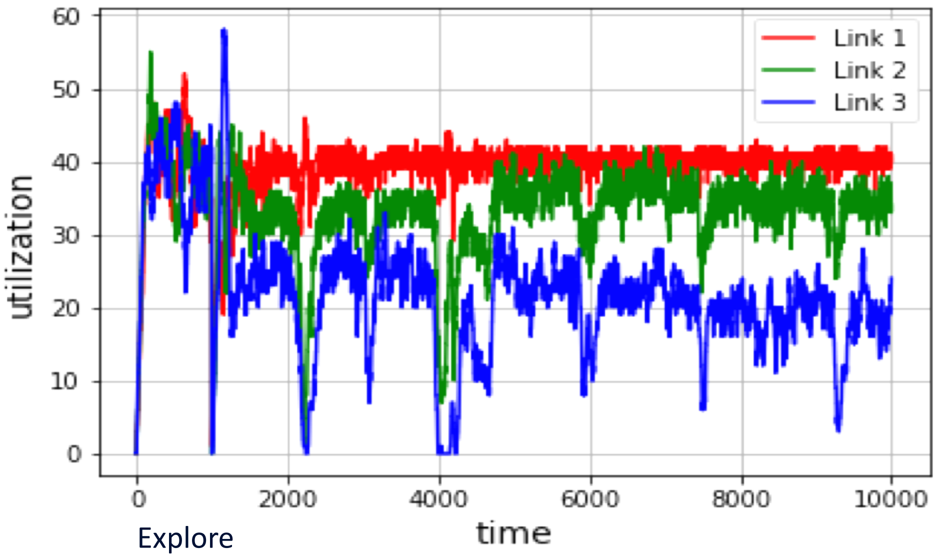
Assigned to all
(End-to-end Delay, …)

Results

Global reward

Path	Good?
1	Best
2	Ok
3	Worst

Two-layer DNN,
Huber Loss,
Batch Size=32,
Replay Memory =1000



Store
Experience

Replay Memory

Learn

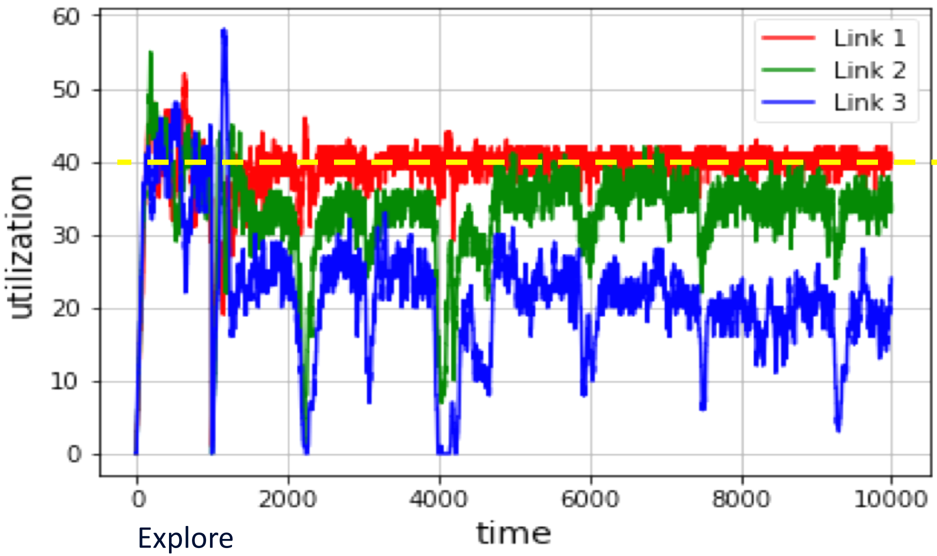
Results

Utilization threshold= 40

(by local reward)

Global reward

Path	Good?
1	Best
2	Ok
3	Worst



Two-layer DNN,
Huber Loss,
Batch Size=32,
Replay Memory =1000

Store Experience

Replay Memory

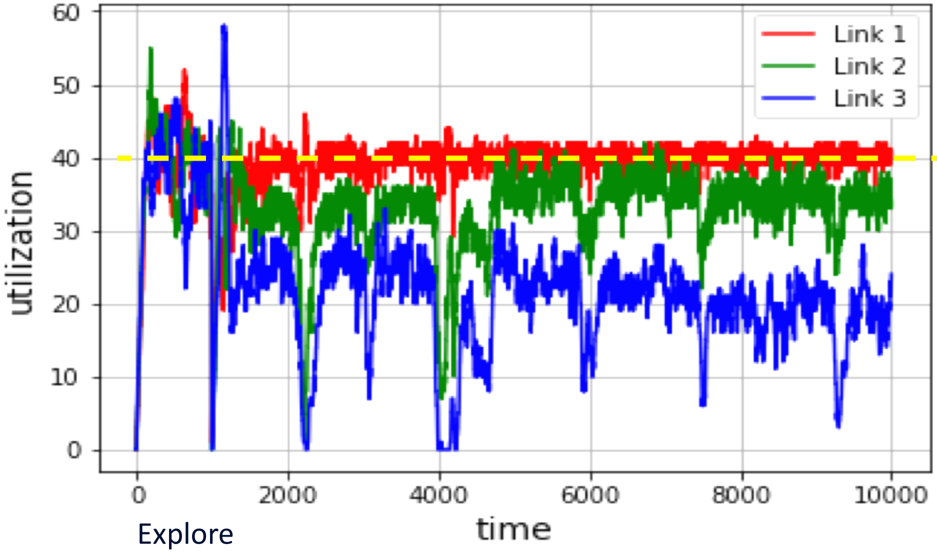
Learn

Results

Utilization threshold= 40
(by local reward)

Global reward

Path	Good?
1	Best
2	Ok
3	Worst



Balance the links
(by local reward)

Two-layer DNN,
Huber Loss,
Batch Size=32,
Replay Memory =1000

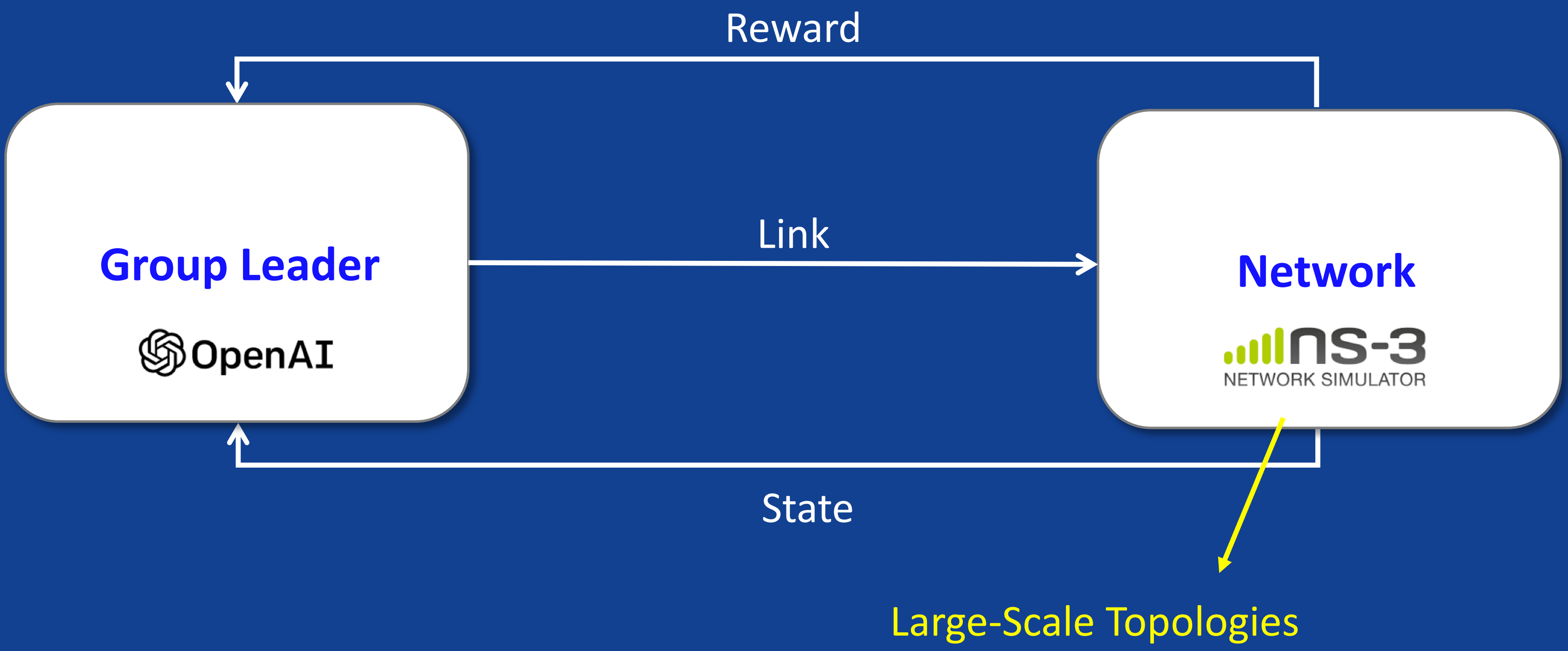
Store
Experience

Replay Memory

Learn

What's Next?

What's Next?



Reward
Group Leader
OpenAI
Link
Network
ns-3
NETWORK SIMULATOR
State
Large-Scale Topologies

Questions?

Thank you

Paper: Hierarchical Deep Double Q-Routing
https://arxiv.org/pdf/1910.04041.pdf

E-mail: ramy.ali@psu.edu

