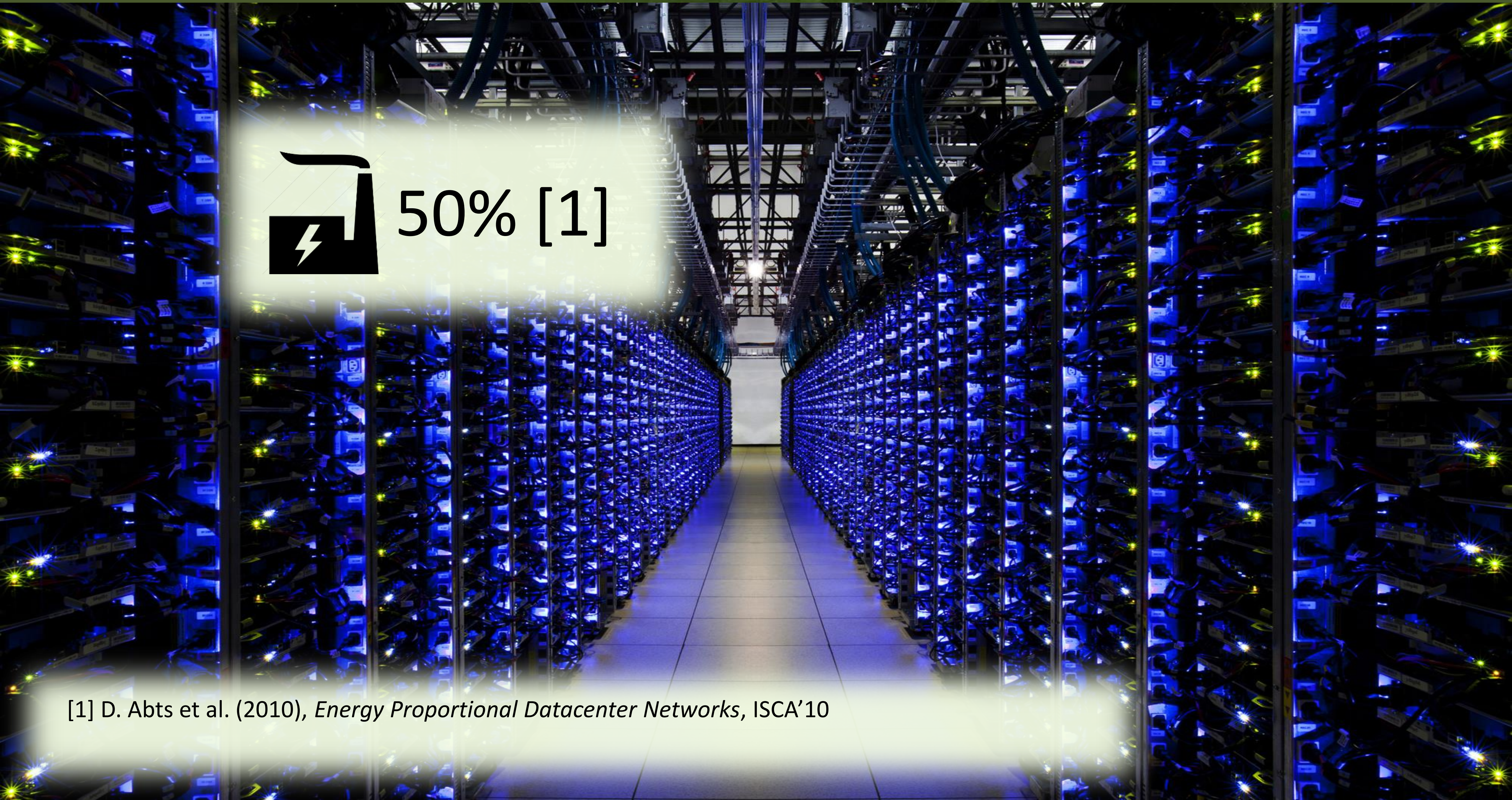
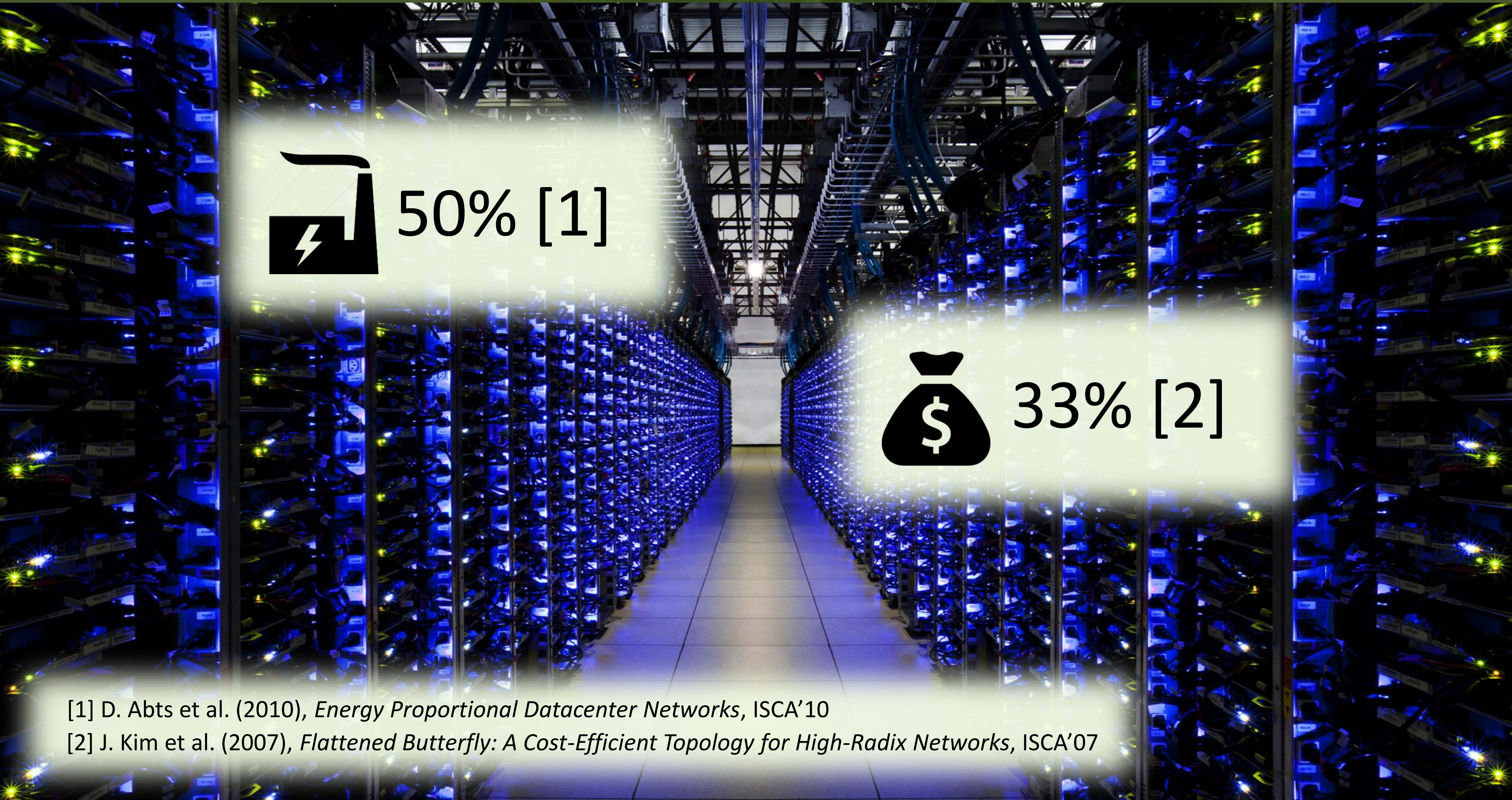
Lowering Diameter Enables Cost-Effective and High-Performance Networks

MACIEJ BESTA, ERIK HENRIKSSON, TORSTEN HOEFLER



50% [1]

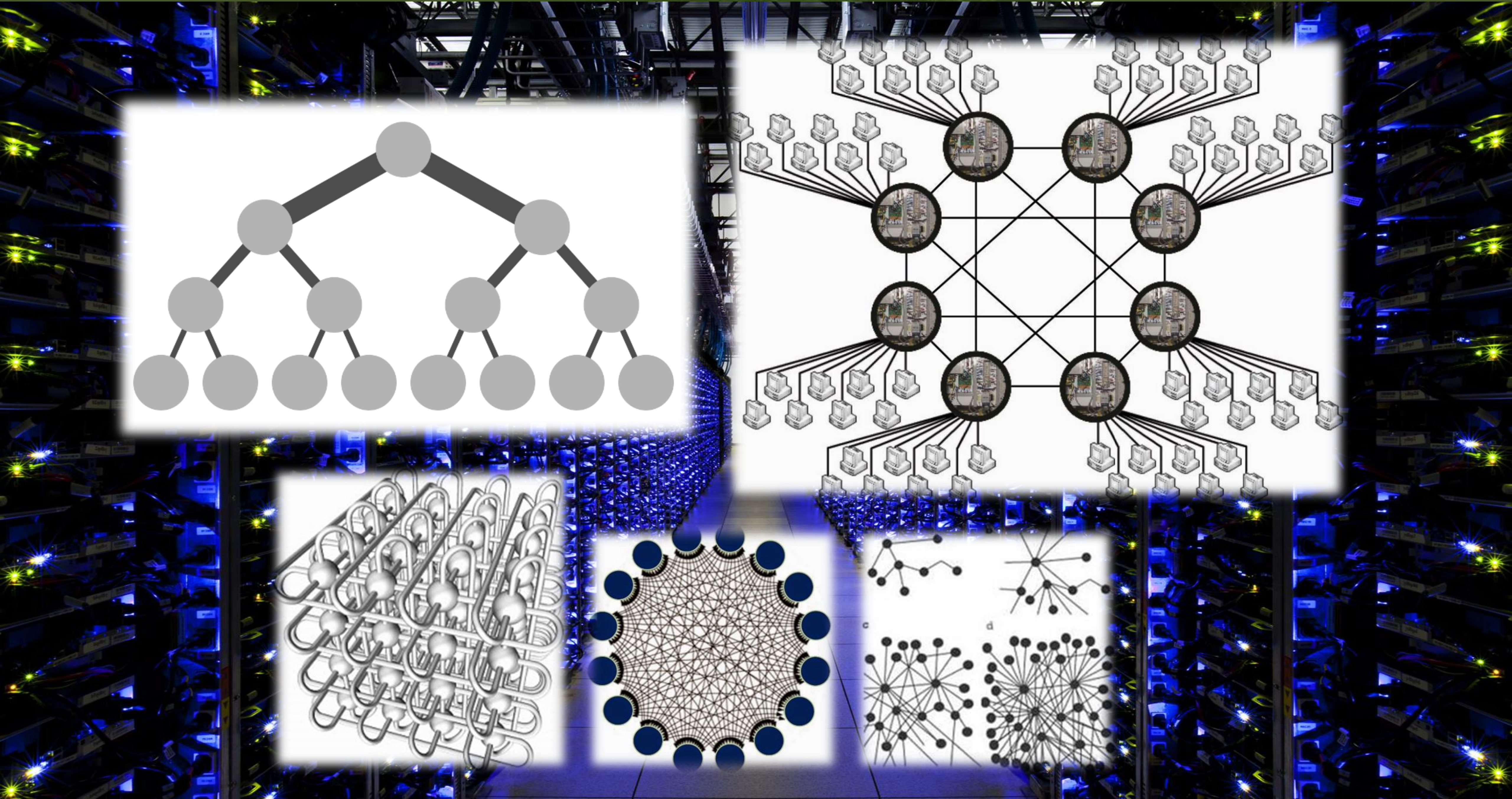
[1] D. Abts et al. (2010), *Energy Proportional Datacenter Networks*, ISCA'10

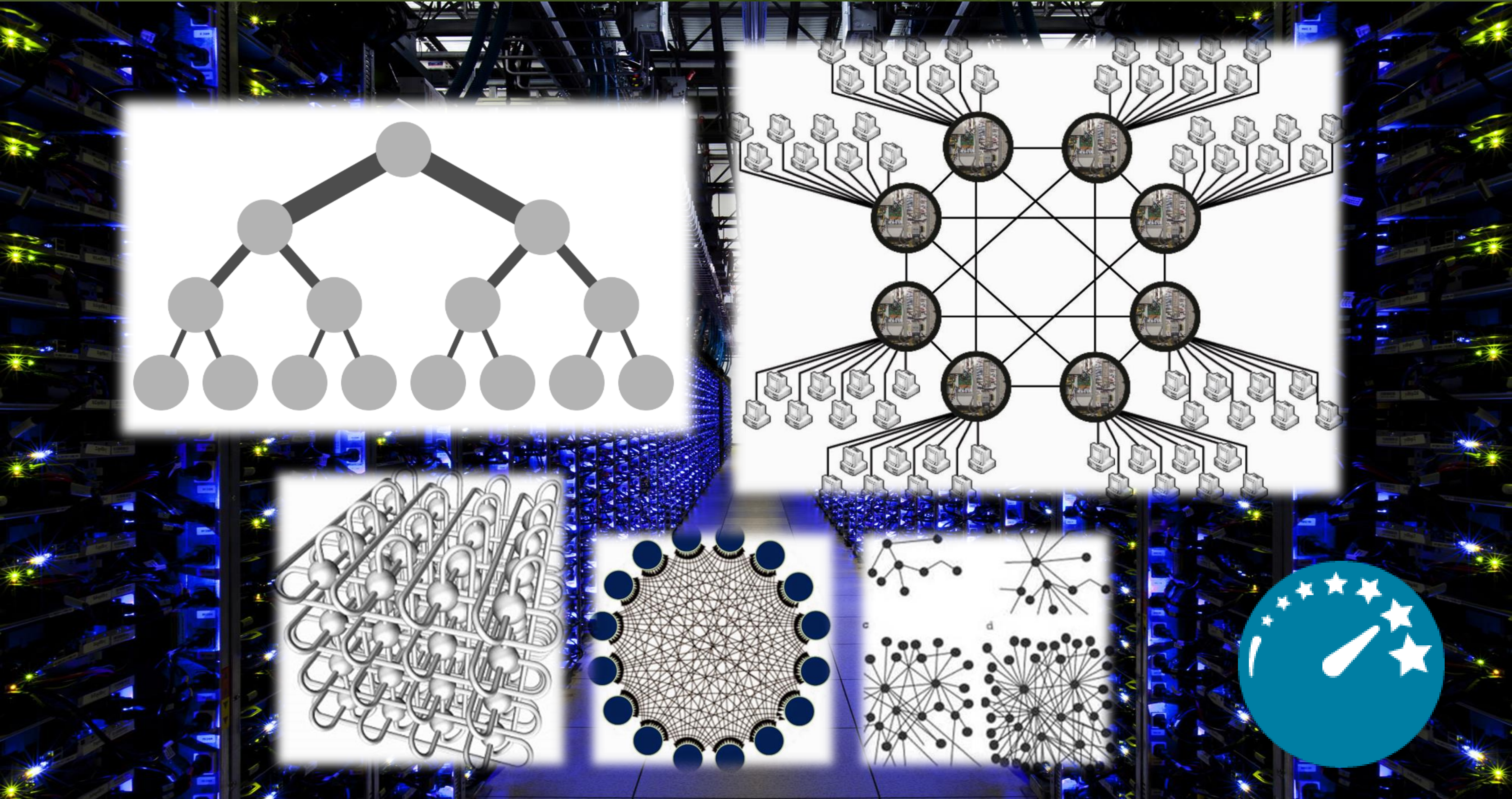


50% [1]

33% [2]

[1] D. Abts et al. (2010), *Energy Proportional Datacenter Networks*, ISCA'10
[2] J. Kim et al. (2007), *Flattened Butterfly: A Cost-Efficient Topology for High-Radix Networks*, ISCA'07





Designing An Efficient Network Topology

Designing An Efficient Network Topology

Key idea

Designing An Efficient Network Topology

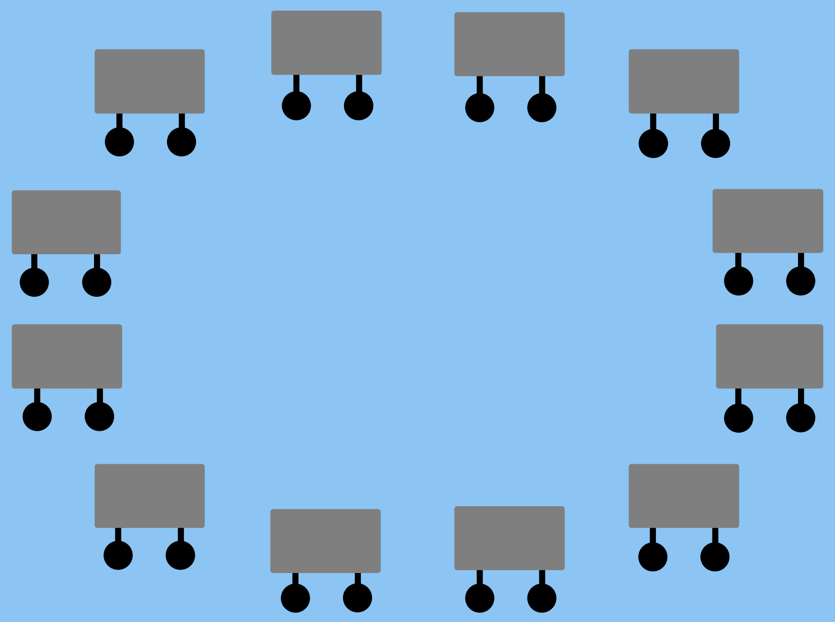
Key idea

Lower diameter and thus average path length: fewer cables and routers necessary.

Designing An Efficient Network Topology

Key idea

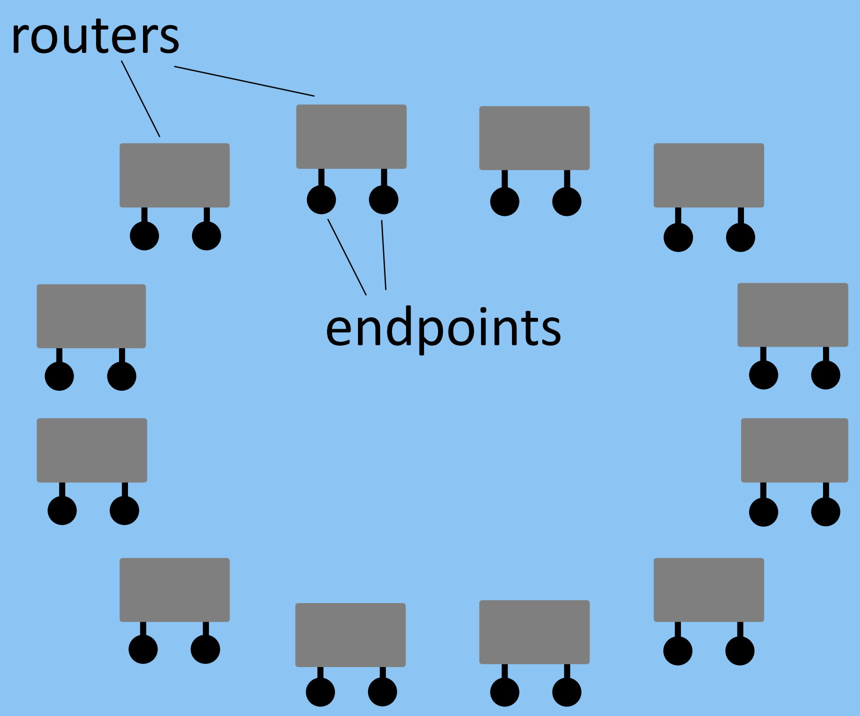
Lower diameter and thus average path length: fewer cables and routers necessary.



Designing An Efficient Network Topology

Key idea

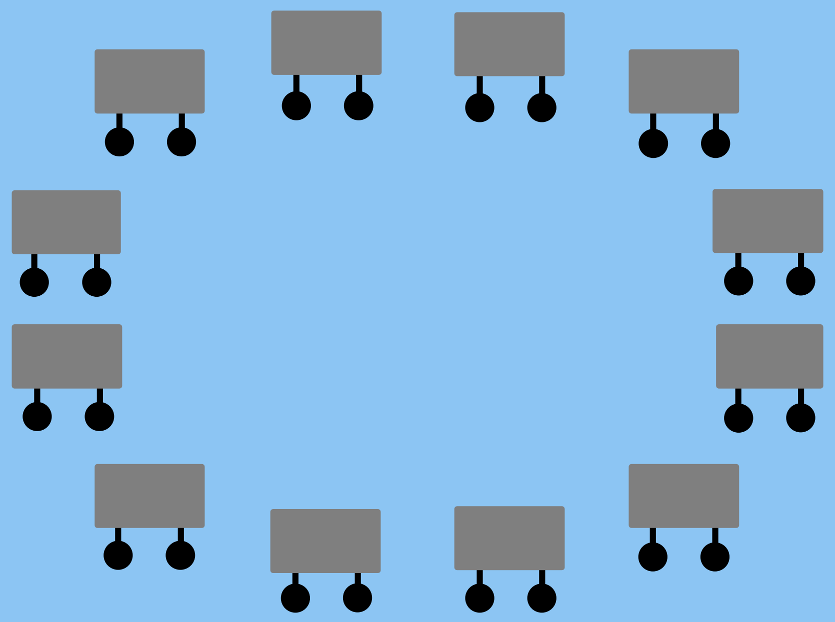
Lower diameter and thus average path length: fewer cables and routers necessary.



Designing An Efficient Network Topology

Key idea

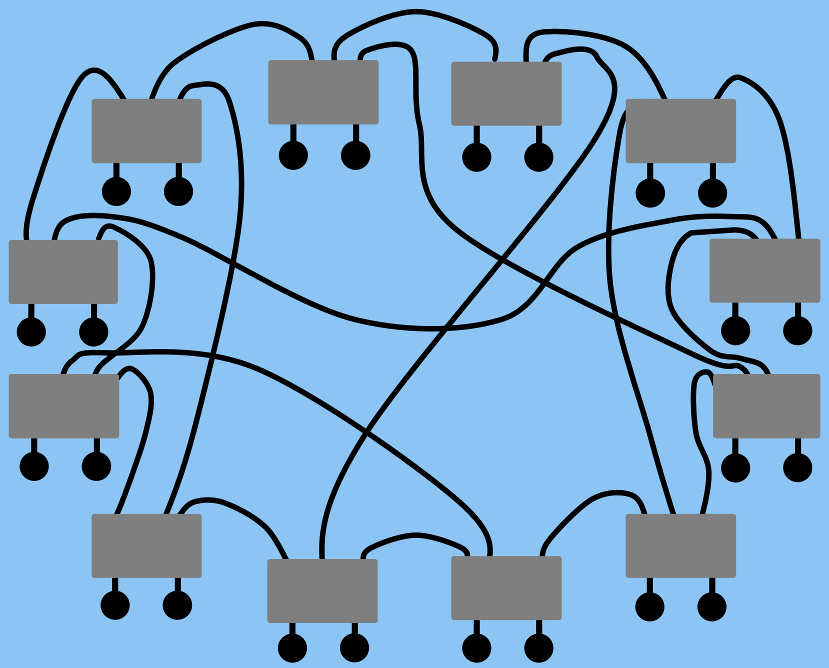
Lower diameter and thus average path length: fewer cables and routers necessary.



Designing An Efficient Network Topology

Key idea

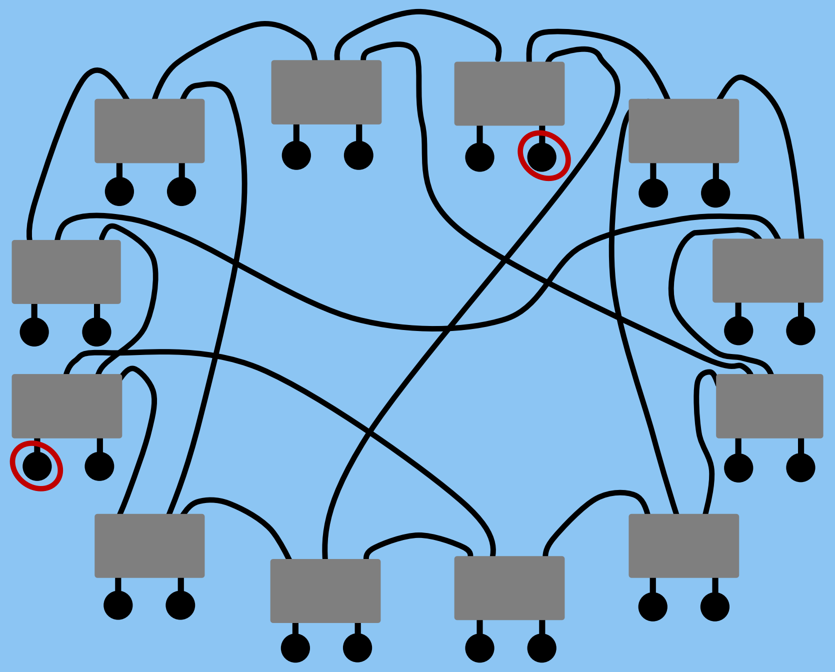
Lower diameter and thus average path length: fewer cables and routers necessary.



Designing An Efficient Network Topology

Key idea

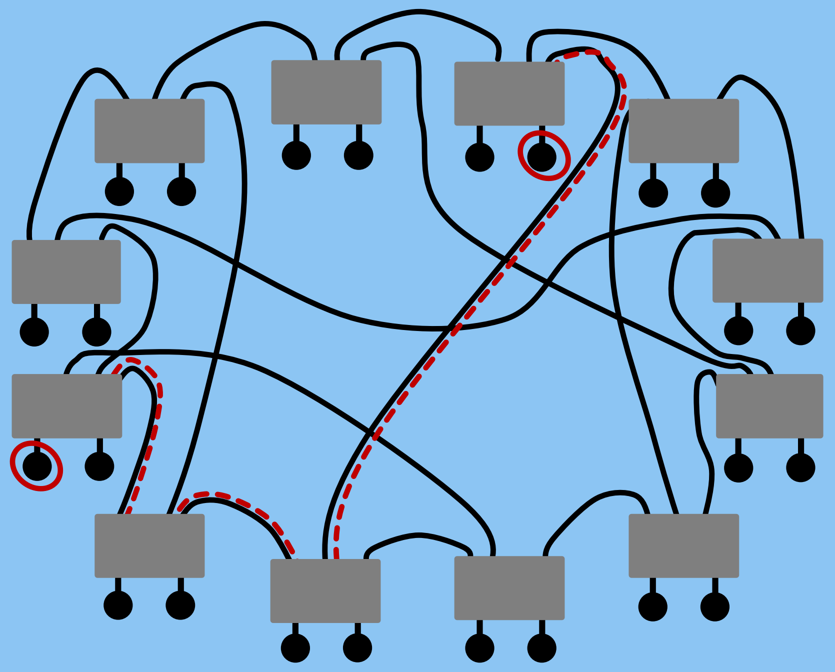
Lower diameter and thus average path length: fewer cables and routers necessary.



Designing An Efficient Network Topology

Key idea

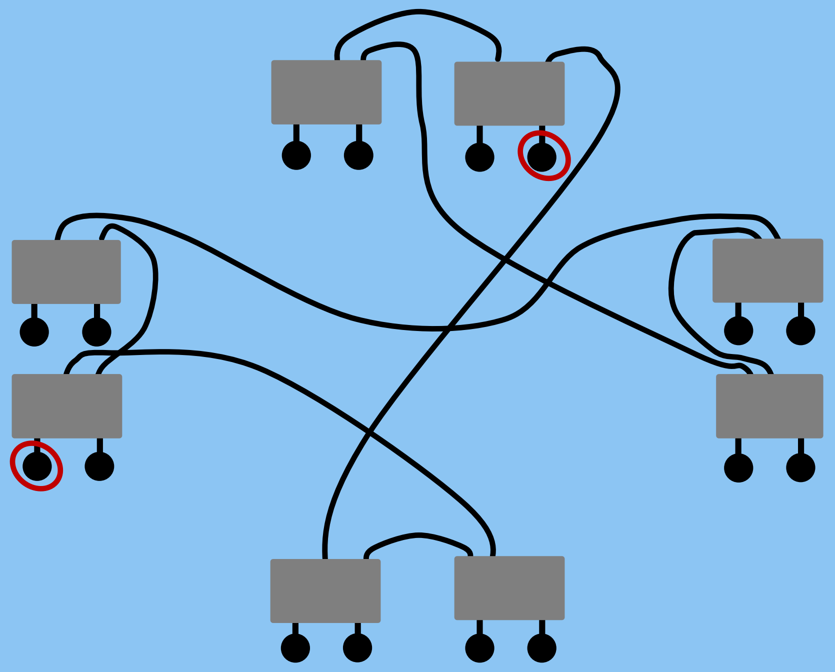
Lower diameter and thus average path length: fewer cables and routers necessary.



Designing an Efficient Network Topology

Key idea

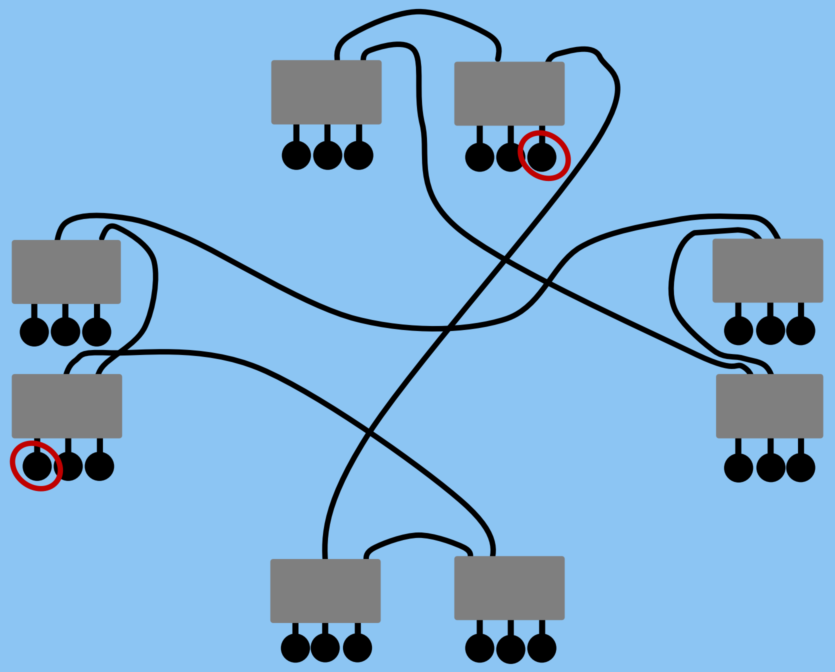
Lower diameter and thus average path length: fewer cables and routers necessary.



Designing An Efficient Network Topology

Key idea

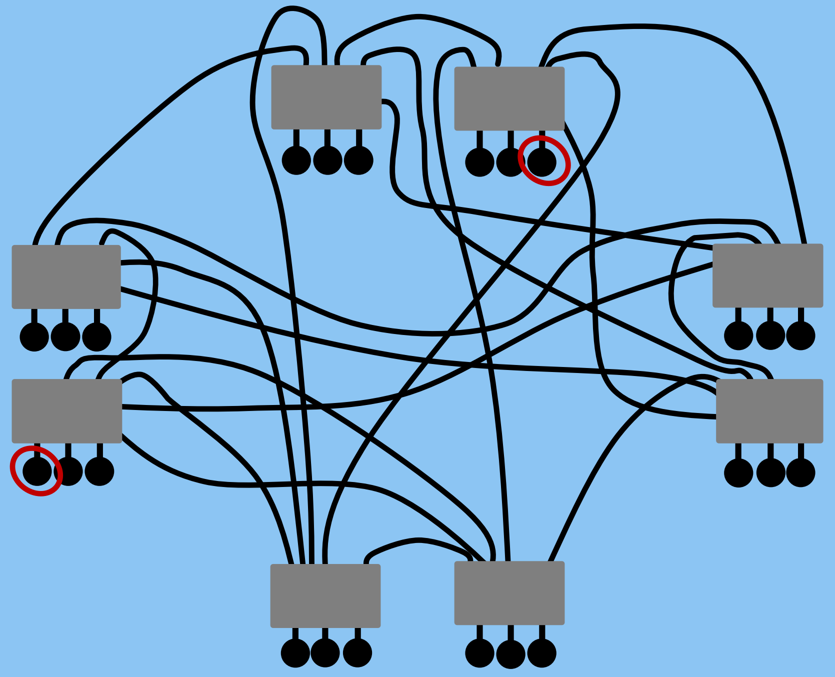
Lower diameter and thus average path length: fewer cables and routers necessary.



Designing An Efficient Network Topology

Key idea

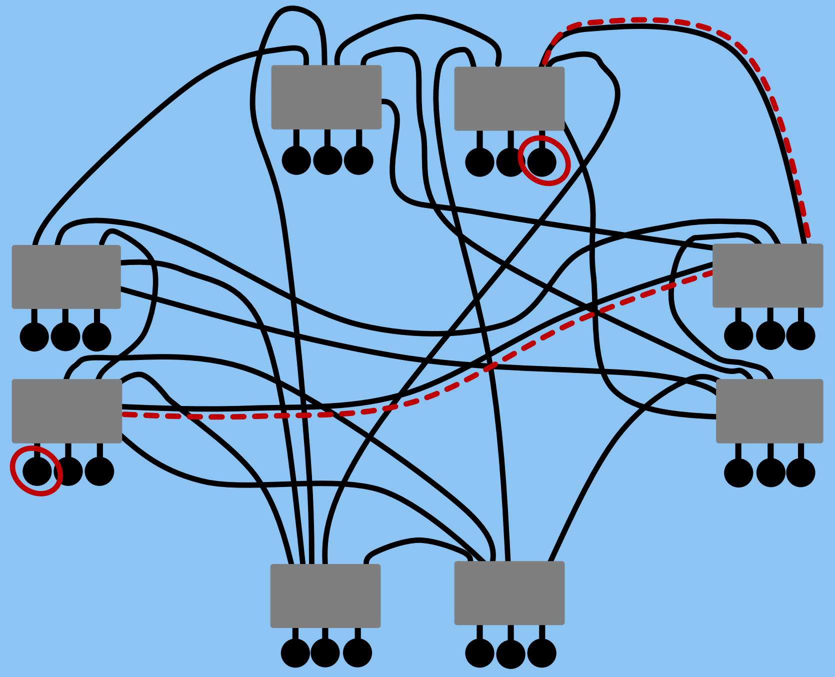
Lower diameter and thus average path length: fewer cables and routers necessary.



Designing An Efficient Network Topology

Key idea

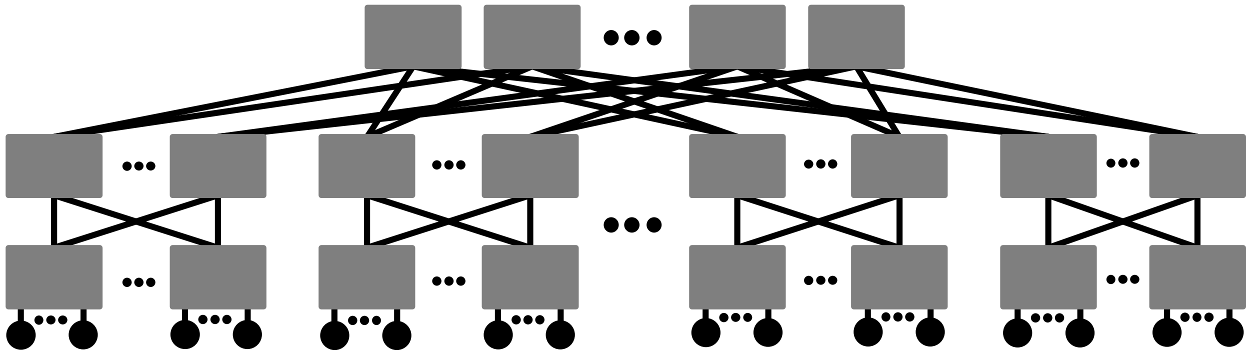
Lower diameter and thus average path length: fewer cables and routers necessary.



Designing An Efficient Network Topology

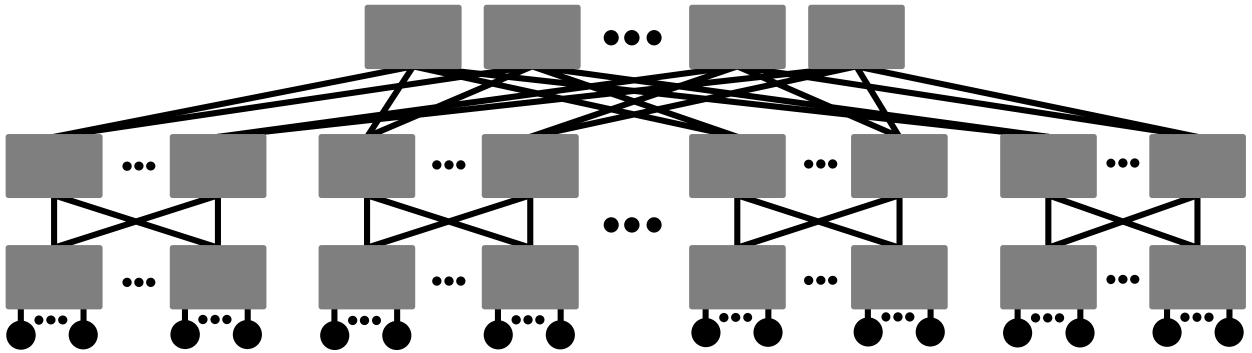
Designing An Efficient Network Topology

3-level fat tree [1]:



DESIGNING AN EFFICIENT NETWORK TOPOLOGY

3-level fat tree [1]:

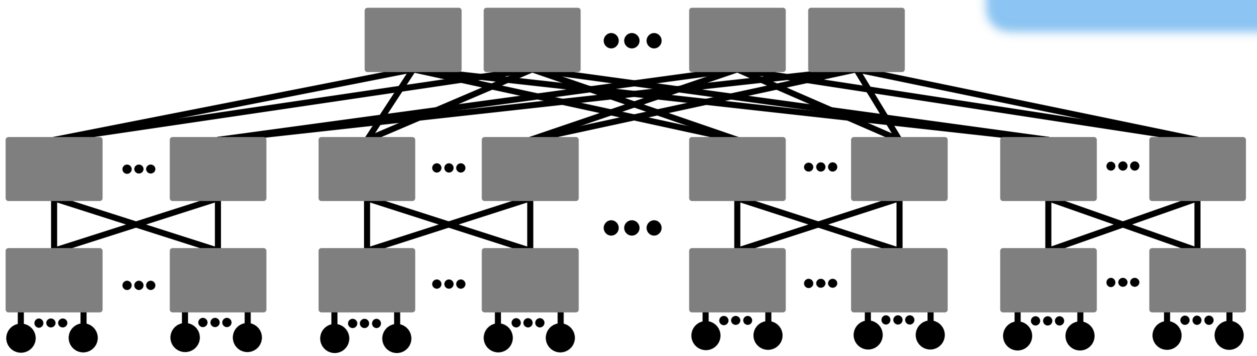


TSUBAME2.0

Designing An Efficient Network Topology

3-level fat tree [1]:

diameter = 4

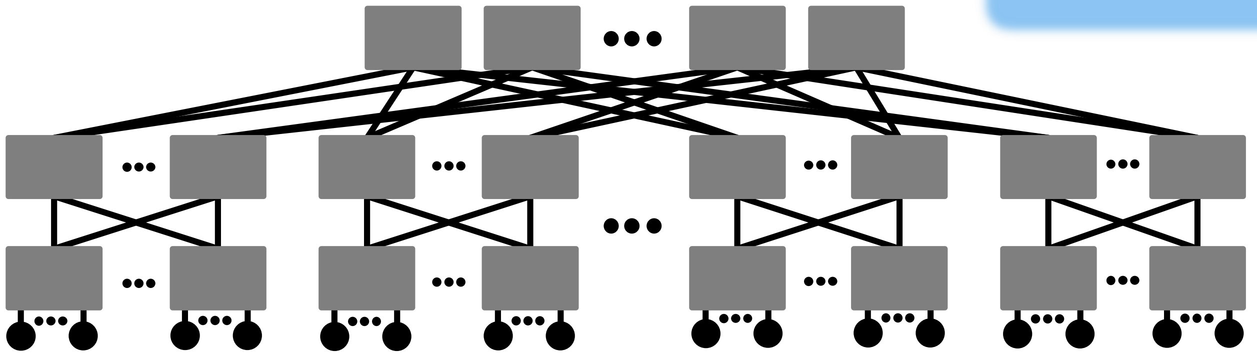


TSUBAME2.0

Designing An Efficient Network Topology

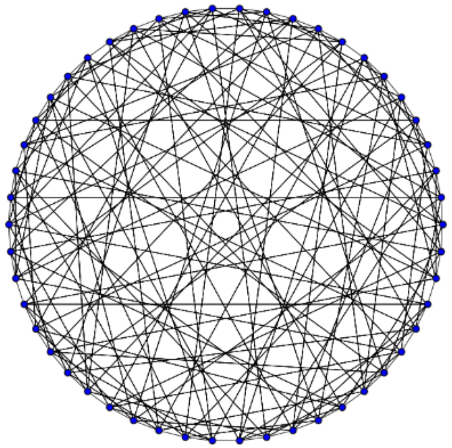
3-level fat tree [1]:

diameter = 4



TSUBAME2.0

Slim Fly [2] based on the Hoffman-Singleton Graph [3]:

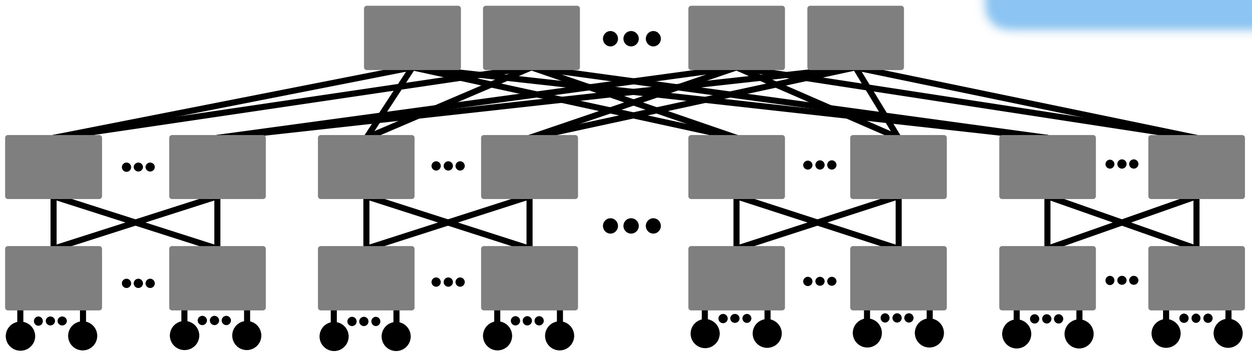


[1] C. E. Leiserson. Fat-trees: Universal Networks for Hardware-Efficient Supercomputing. IEEE Transactions on Computers. 1985.
[2] M. Besta and T. Hoefler. Slim Fly: A Cost Effective Low-Diameter Network Topology. SC14.
[3] Hoffman, Alan J.; Singleton, Robert R. (1960), *Moore graphs with diameter 2 and 3*, IBM Journal of Research and Development

Designing An Efficient Network Topology

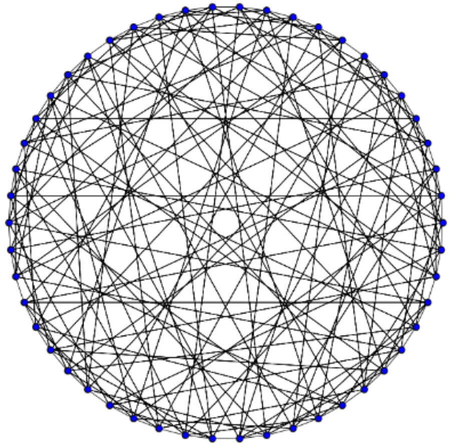
3-level fat tree [1]:

diameter = 4



TSUBAME2.0

Slim Fly [2] based on the Hoffman-Singleton Graph [3]:



diameter = 2
> ~50% fewer routers
> ~30% fewer cables

[1] C. E. Leiserson. Fat-trees: Universal Networks for Hardware-Efficient Supercomputing. IEEE Transactions on Computers. 1985.
[2] M. Besta and T. Hoefler. Slim Fly: A Cost Effective Low-Diameter Network Topology. SC14.
[3] Hoffman, Alan J.; Singleton, Robert R. (1960), *Moore graphs with diameter 2 and 3*, IBM Journal of Research and Development

Designing An Efficient Network Topology

Designing An Efficient Network Topology

Key method

Designing An Efficient Network Topology

Key method

Optimize towards the Moore Bound [1]:
the upper bound on the *number of vertices* in a graph with given *diameter D* and *radix k.*

[1] M. Miller, J. Siráň. Moore graphs and beyond: A survey of the degree/diameter problem, Electronic Journal of Combinatorics, 2005.

Designing An Efficient Network Topology

Key method

Optimize towards the Moore Bound [1]:
the upper bound on the *number of vertices* in a graph with given *diameter D* and *radix k.*

$$MB(D, k)$$

[1] M. Miller, J. Siráň. Moore graphs and beyond: A survey of the degree/diameter problem, Electronic Journal of Combinatorics, 2005.

DESIGNING AN EFFICIENT NETWORK TOPOLOGY

Key method

Optimize towards the Moore Bound [1]:
the upper bound on the *number of vertices* in a graph with given *diameter D* and *radix k.*

$$MB(D, k)$$

[1] M. Miller, J. Siráň. Moore graphs and beyond: A survey of the degree/diameter problem, Electronic Journal of Combinatorics, 2005.

Designing An Efficient Network Topology

Key method

Optimize towards the Moore Bound [1]:
the upper bound on the *number of vertices* in a graph with given *diameter D* and *radix k.*

$$MB(D, k) = 1$$

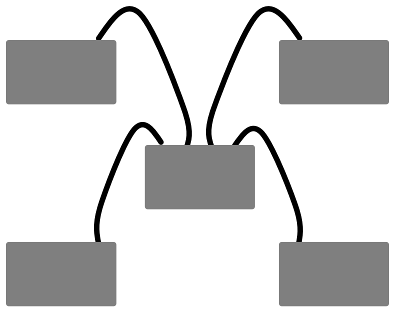
[1] M. Miller, J. Siráň. Moore graphs and beyond: A survey of the degree/diameter problem, Electronic Journal of Combinatorics, 2005.

DESIGNING AN EFFICIENT NETWORK TOPOLOGY

Key method

Optimize towards the Moore Bound [1]: the upper bound on the *number of vertices* in a graph with given *diameter D* and *radix k.*

$$MB(D,k) = 1$$



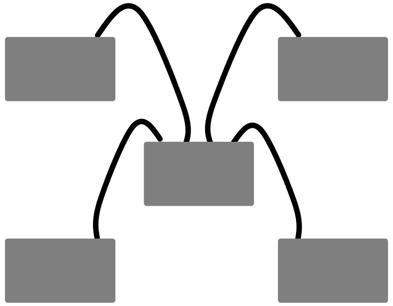
[1] M. Miller, J. Siráň. Moore graphs and beyond: A survey of the degree/diameter problem, Electronic Journal of Combinatorics, 2005.

DESIGNING AN EFFICIENT NETWORK TOPOLOGY

Key method

Optimize towards the Moore Bound [1]: the upper bound on the *number of vertices* in a graph with given *diameter D* and *radix k.*

$$MB(D,k) = 1 \ \ + k$$



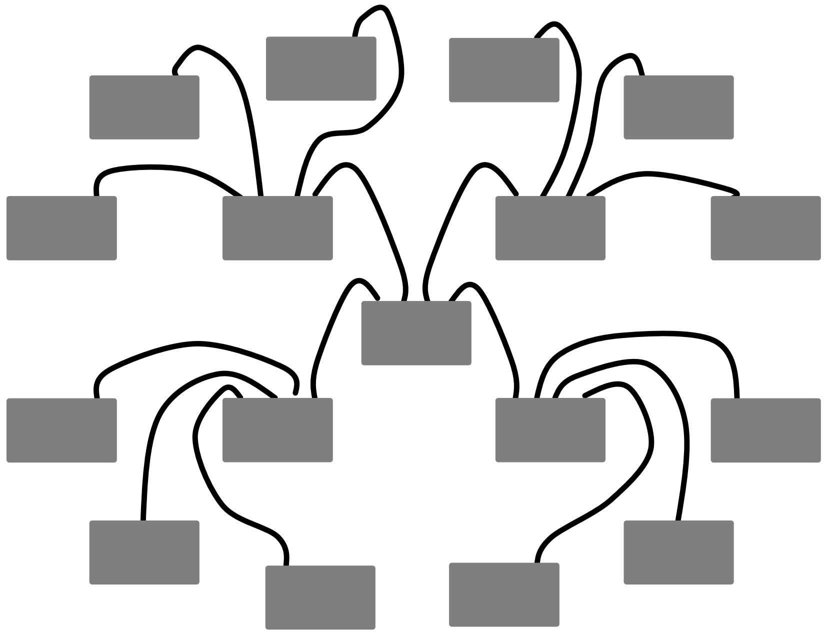
[1] M. Miller, J. Siráň. Moore graphs and beyond: A survey of the degree/diameter problem, Electronic Journal of Combinatorics, 2005.

Designing An Efficient Network Topology

Key method

Optimize towards the Moore Bound [1]:
the upper bound on the *number of vertices* in a graph with given *diameter D* and *radix k.*

$$MB(D,k) = 1 \quad + k$$



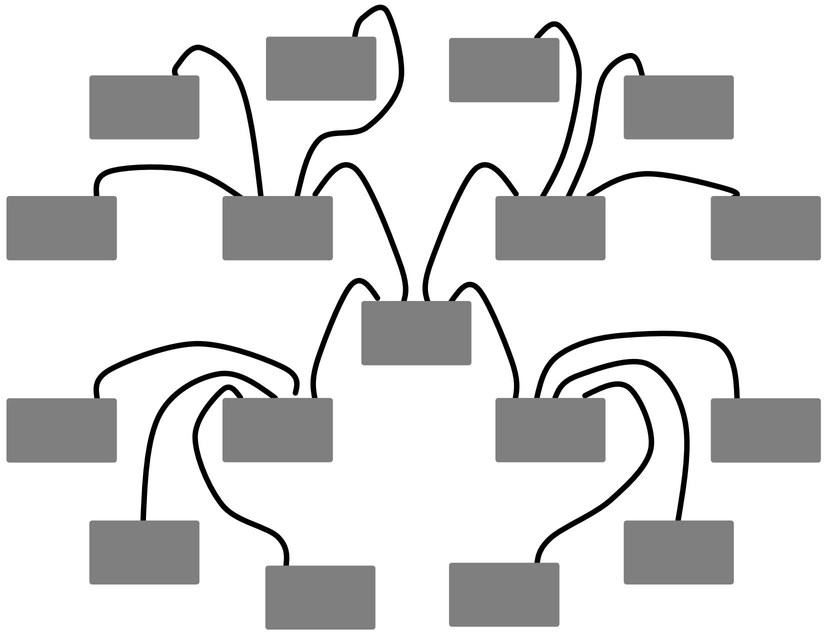
[1] M. Miller, J. Siráň. Moore graphs and beyond: A survey of the degree/diameter problem, Electronic Journal of Combinatorics, 2005.

DESIGNING AN EFFICIENT NETWORK TOPOLOGY

Key method

Optimize towards the Moore Bound [1]:
the upper bound on the *number of vertices* in a graph with given *diameter D* and *radix k.*

$$MB(D,k) = 1 \ + k \ + k(k-1)$$



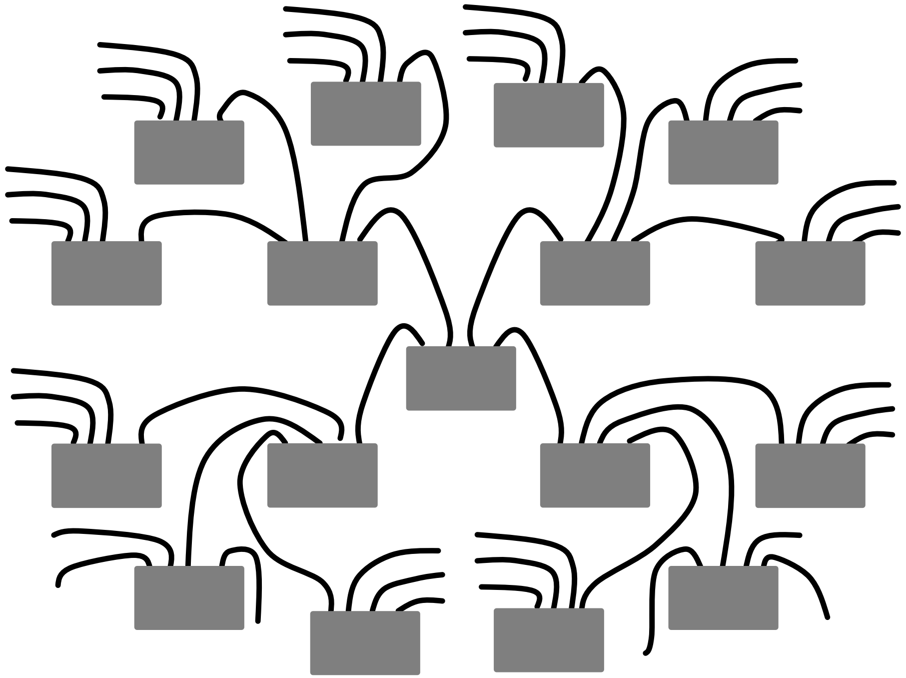
[1] M. Miller, J. Siráň. Moore graphs and beyond: A survey of the degree/diameter problem, Electronic Journal of Combinatorics, 2005.

DESIGNING AN EFFICIENT NETWORK TOPOLOGY

Key method

Optimize towards the Moore Bound [1]: the upper bound on the *number of vertices* in a graph with given *diameter D* and *radix k.*

$$MB(D,k) = 1 + k + k(k-1)$$



[1] M. Miller, J. Siráň. Moore graphs and beyond: A survey of the degree/diameter problem, Electronic Journal of Combinatorics, 2005.

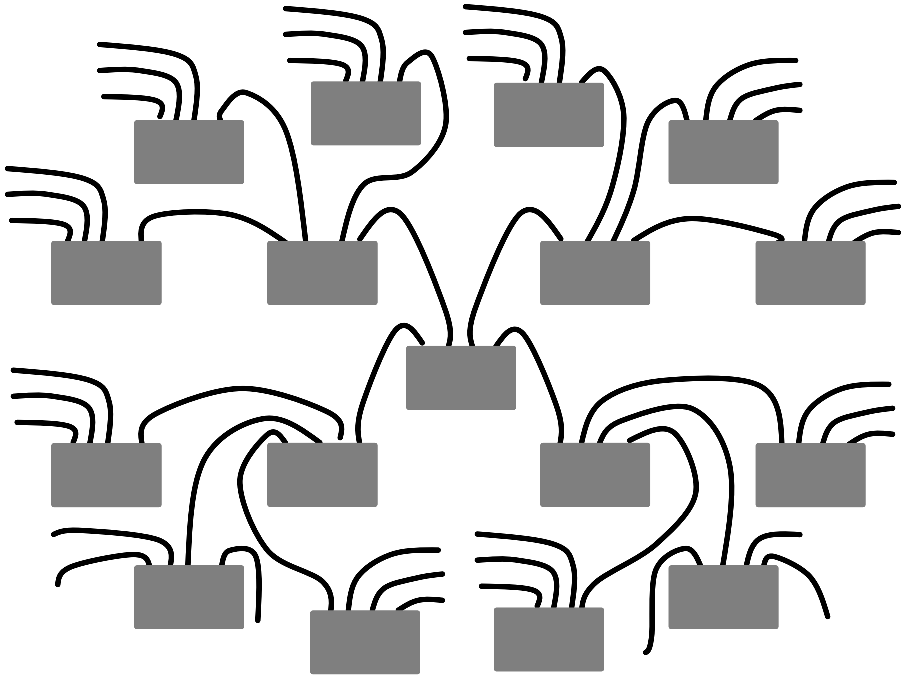
Designing An Efficient Network Topology

Key method

Optimize towards the Moore Bound [1]:
the upper bound on the *number of vertices* in a graph with given *diameter D* and *radix k*,

$$MB(D,k) = 1 + k + k(k-1) + k(k-1)^2 + \cdots$$

$$= 1 + k\sum_{i=0}^{D-1}(k-1)^i$$



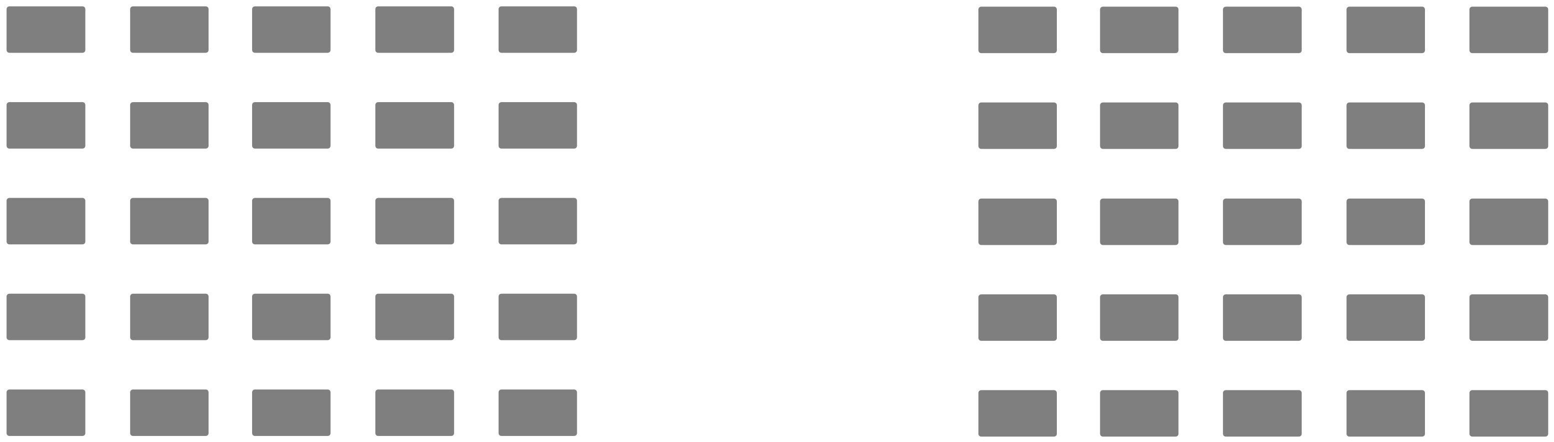
[1] M. Miller, J. Siráň. Moore graphs and beyond: A survey of the degree/diameter problem, Electronic Journal of Combinatorics, 2005.

DIAMETER-2 SLIM FLY
STRUCTURE INTUITION

- Example design for *diameter* = 2

Diameter-2 Slim Fly
Structure Intuition

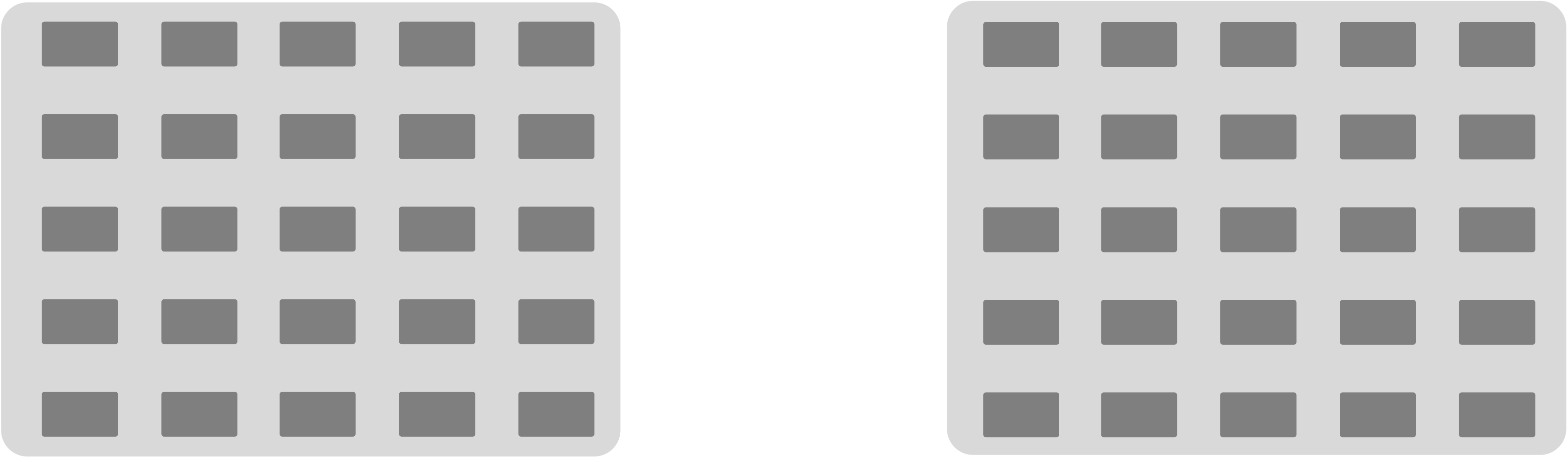
- Example design for *diameter* = 2



Diameter-2 Slim Fly
Structure Intuition

- Example design for *diameter* = 2

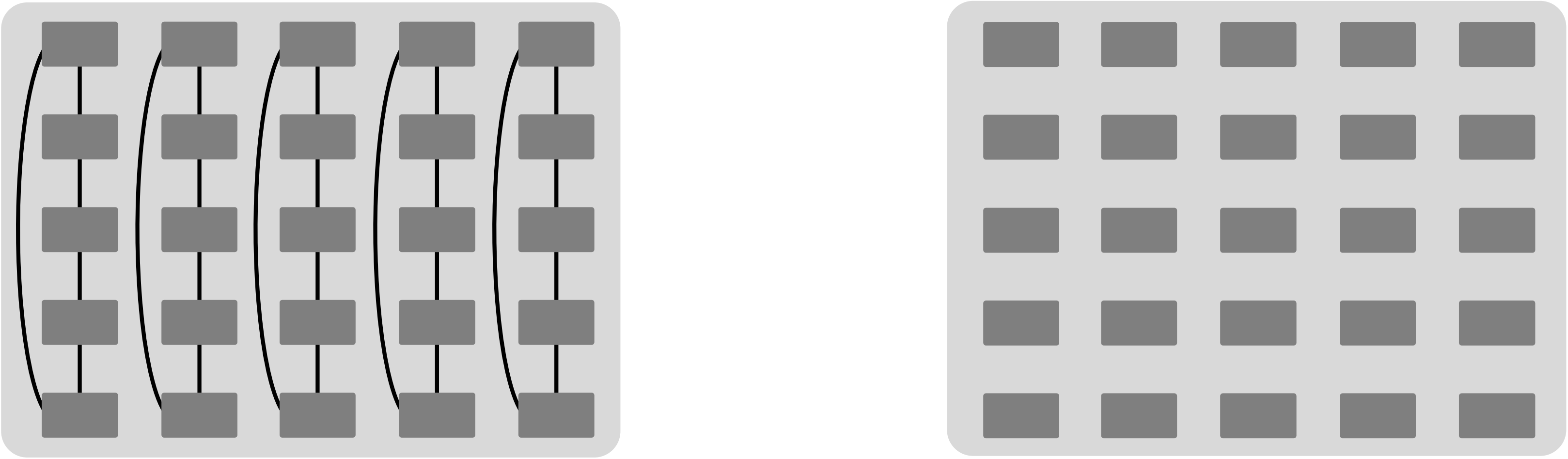
A subgraph with
identical groups of routers



Diameter-2 Slim Fly
Structure Intuition

- Example design for *diameter* = 2

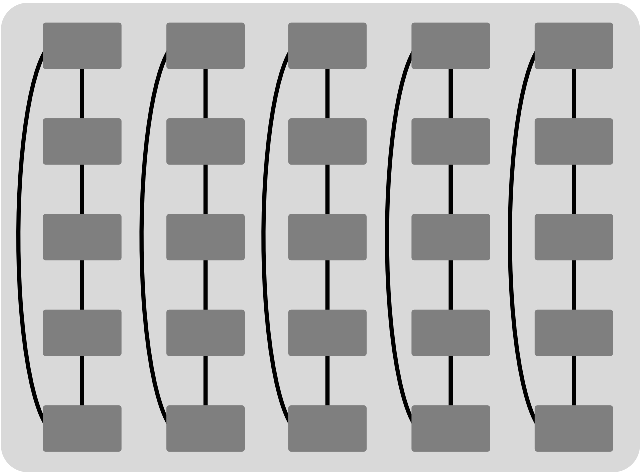
A subgraph with
identical groups of routers



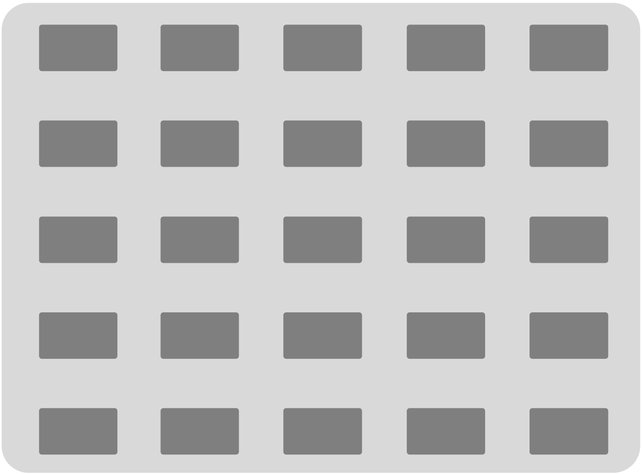
Diameter-2 Slim Fly
Structure Intuition

- Example design for *diameter* = 2

A subgraph with
identical groups of routers



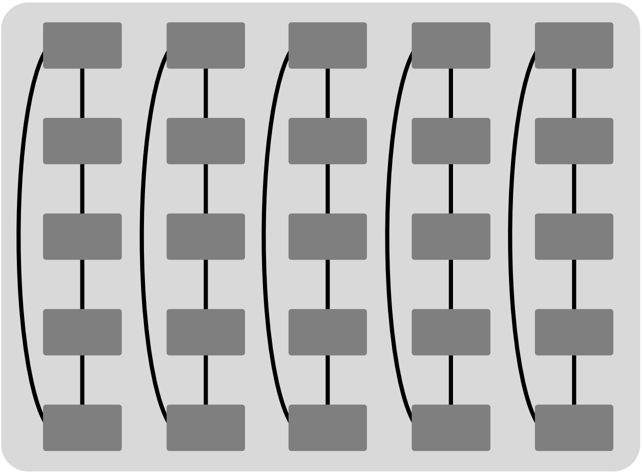
A subgraph with
identical groups of routers



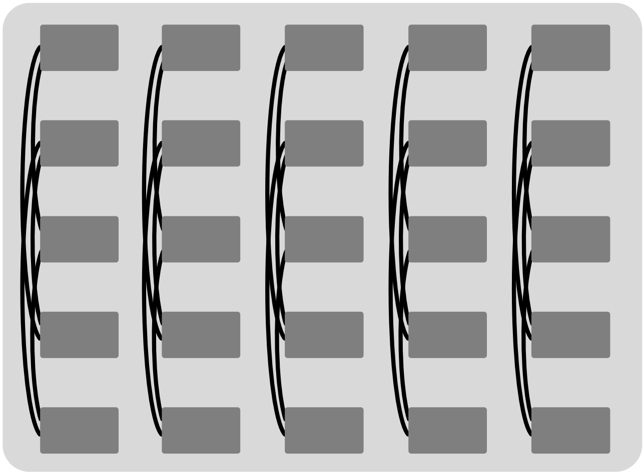
Diameter-2 Slim Fly
Structure Intuition

- Example design for *diameter* = 2

A subgraph with
identical groups of routers



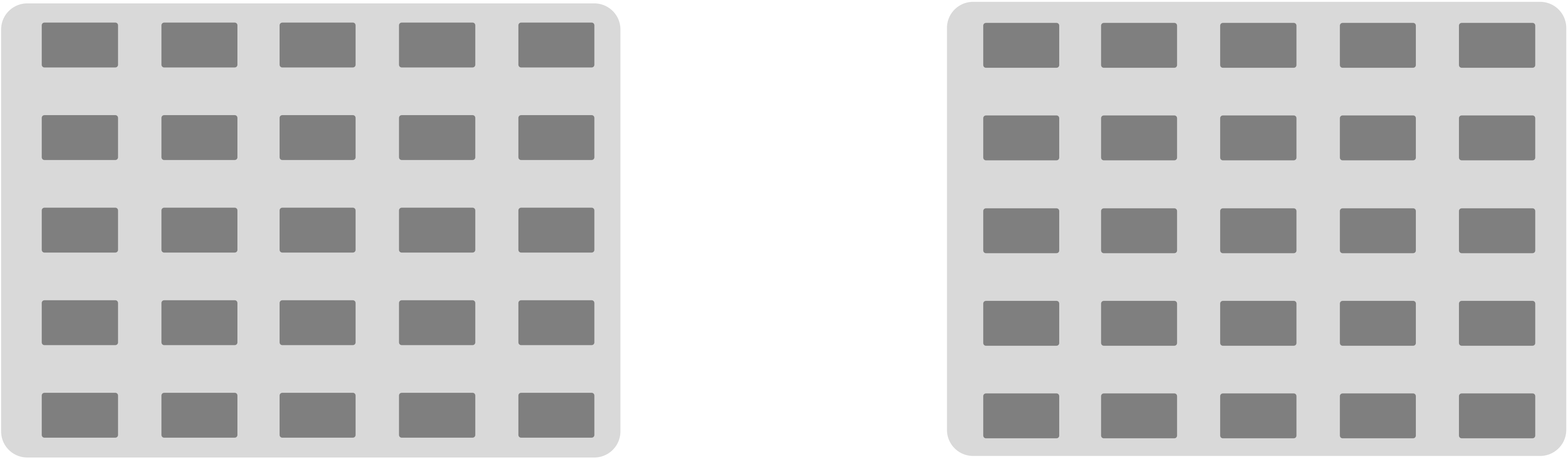
A subgraph with
identical groups of routers



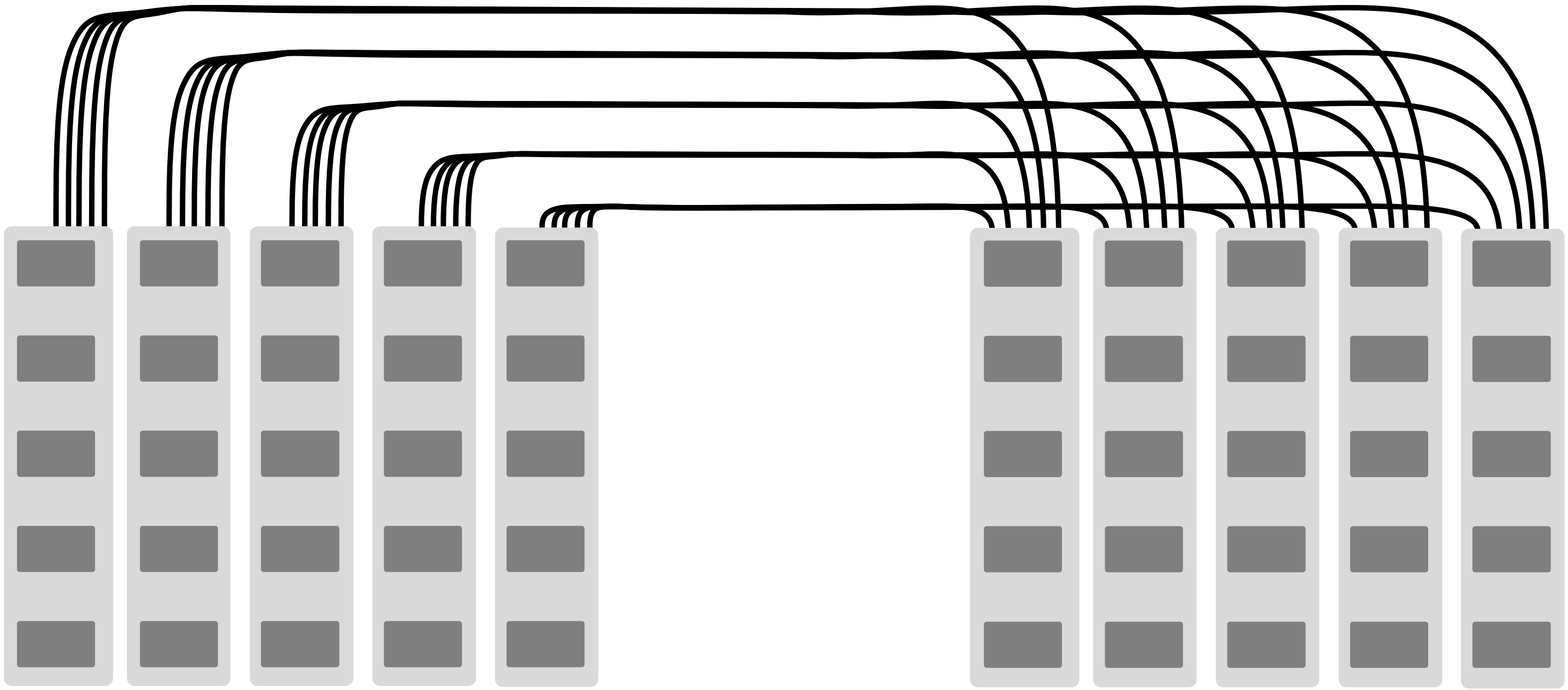
Diameter-2 Slim Fly

Structure Intuition

- Example design for *diameter* = 2



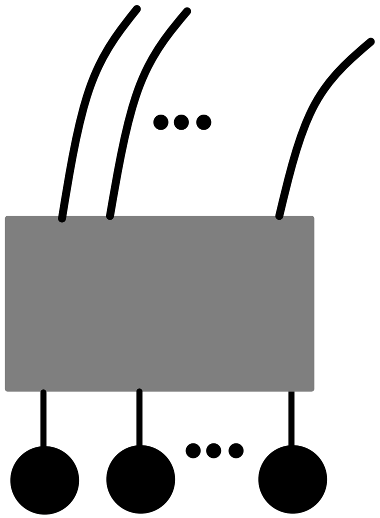
Diameter-2 Slim Fly
Structure Intuition



Groups form a fully-connected bipartite graph

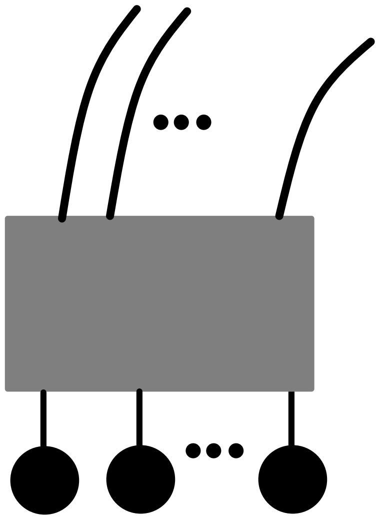
Diameter-2 Slim Fly

Attaching endpoints: Diameter 2



Diameter-2 Slim Fly
Attaching endpoints: Diameter 2

1 Get load / per router-router channel (average number of routes per channel)

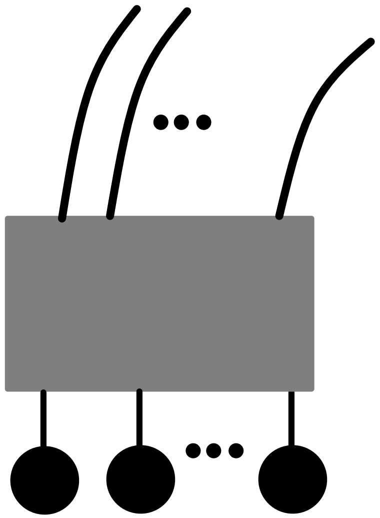


Diameter-2 Slim Fly

Attaching endpoints: Diameter 2

1 Get load l per router-router channel (average number of routes per channel)

$$l = \frac{total\ number\ of\ routes}{total\ number\ of\ channels}$$

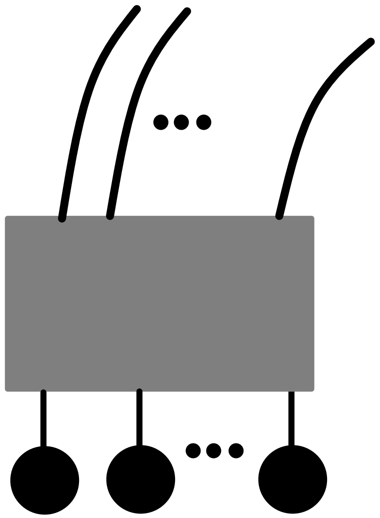


Diameter-2 Slim Fly
Attaching endpoints: Diameter 2

1 Get load l per router-router channel (average number of routes per channel)

$$l = \frac{total\ number\ of\ routes}{total\ number\ of\ channels}$$

2

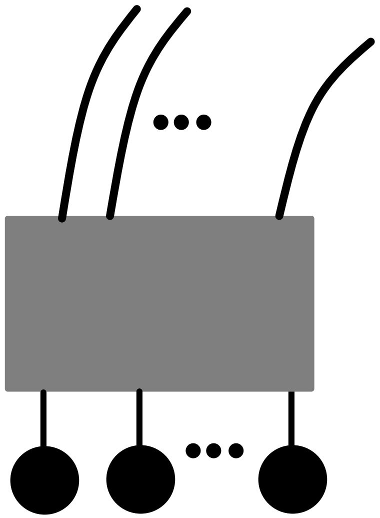


Diameter-2 Slim Fly
Attaching endpoints: Diameter 2

1 Get load l per router-router channel (average number of routes per channel)

$$l = \frac{total\ number\ of\ routes}{total\ number\ of\ channels}$$

2 Make the network balanced, i.e.,:



Diameter-2 Slim Fly

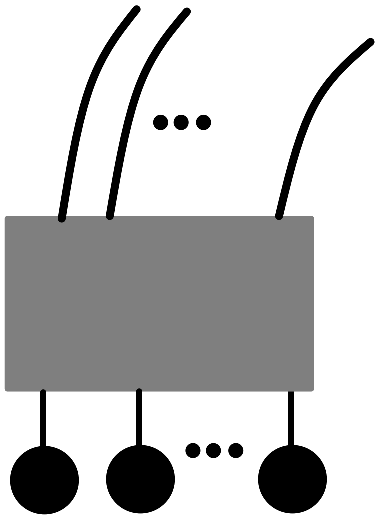
Attaching endpoints: Diameter 2

1 Get load l per router-router channel (average number of routes per channel)

$$l = \frac{total\ number\ of\ routes}{total\ number\ of\ channels}$$

2 Make the network balanced, i.e.,:

each endpoint can inject at full capacity



Diameter-2 Slim Fly
Attaching endpoints: Diameter 2

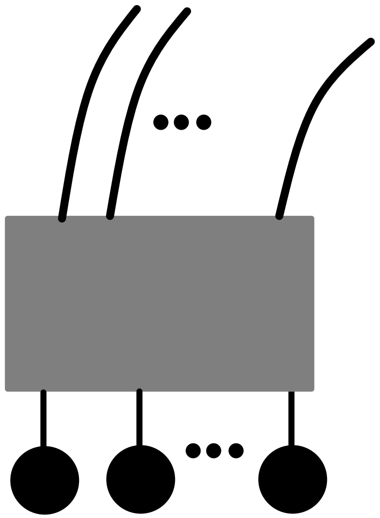
1 Get load l per router-router channel (average number of routes per channel)

$$l = \frac{total\ number\ of\ routes}{total\ number\ of\ channels}$$

2 Make the network balanced, i.e.,:

each endpoint can inject at full capacity

$$local\ uplink\ load = number\ of\ endpoints$$



Diameter-2 Slim Fly
Attaching endpoints: Diameter 2

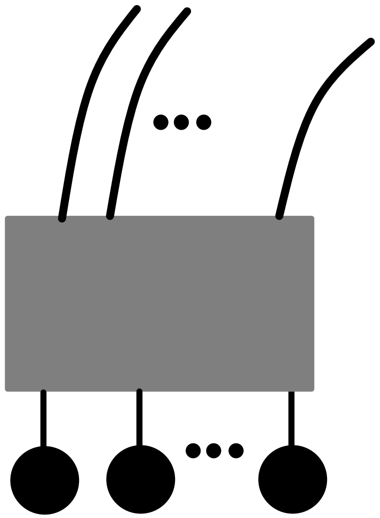
❶ Get load *l* per router-router channel (average number of routes per channel)

$$l = \frac{total\ number\ of\ routes}{total\ number\ of\ channels}$$

❷ Make the network balanced, i.e.,:

each endpoint can inject at full capacity

$$local\ uplink\ load = number\ of\ endpoints = l$$



Diameter-2 Slim Fly
Attaching endpoints: Diameter 2

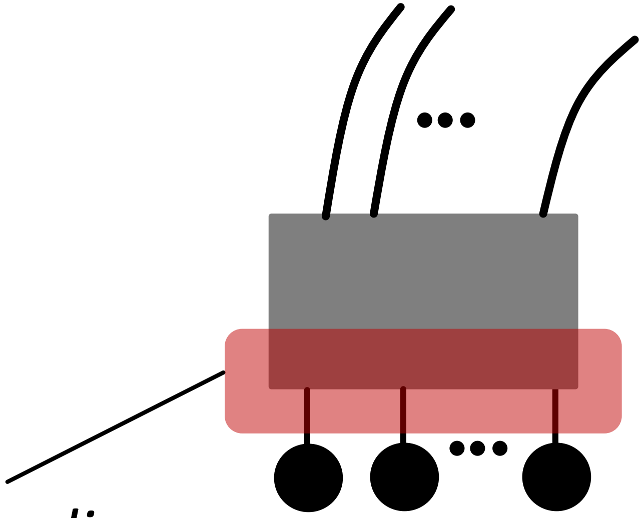
1 Get load l per router-router channel (average number of routes per channel)

$$l = \frac{total\ number\ of\ routes}{total\ number\ of\ channels}$$

2 Make the network balanced, i.e.,:

each endpoint can inject at full capacity

$$local\ uplink\ load = number\ of\ endpoints = l$$



concentration = 33% of router radix

Diameter-2 Slim Fly
Attaching endpoints: Diameter 2

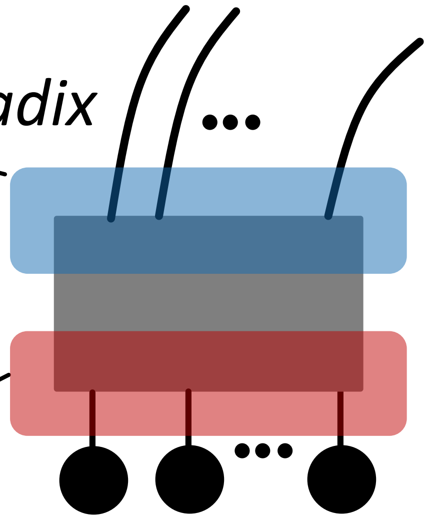
❶ Get load l per router-router channel (average number of routes per channel)

$$l = \frac{total\ number\ of\ routes}{total\ number\ of\ channels}$$

❷ Make the network balanced, i.e.,:

each endpoint can inject at full capacity

$$local\ uplink\ load = number\ of\ endpoints = l$$

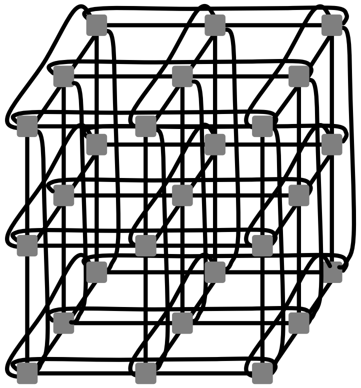


Comparison Targets

Low-radix topologies

Comparison Targets
Low-radix topologies

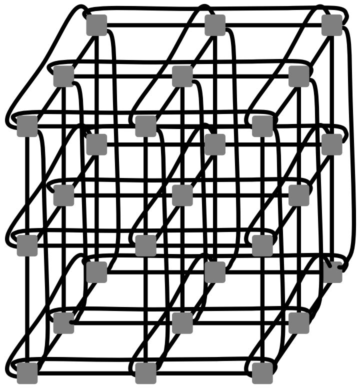
Torus 3D



Cray XE6

Comparison Targets
Low-radix topologies

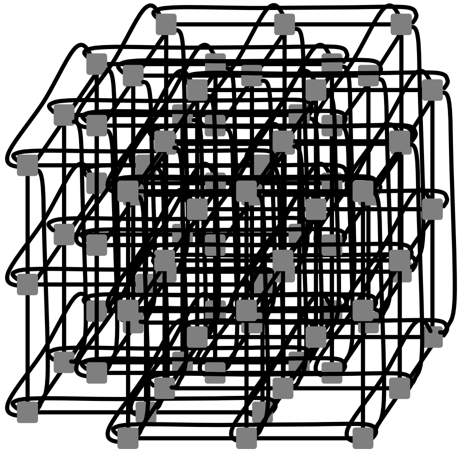
Torus 3D



Cray XE6

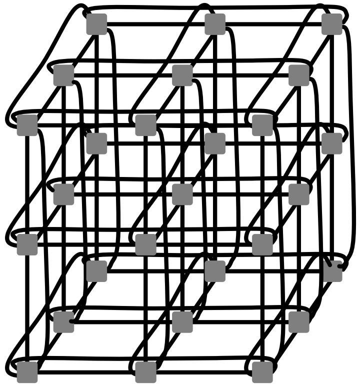
IBM BG/Q

Torus 5D



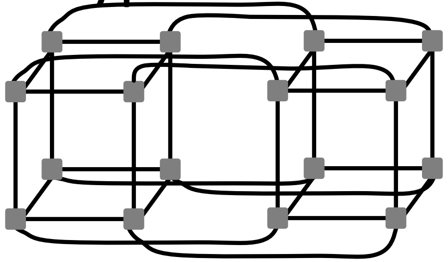
Comparison Targets
Low-radix topologies

Torus 3D



Cray XE6

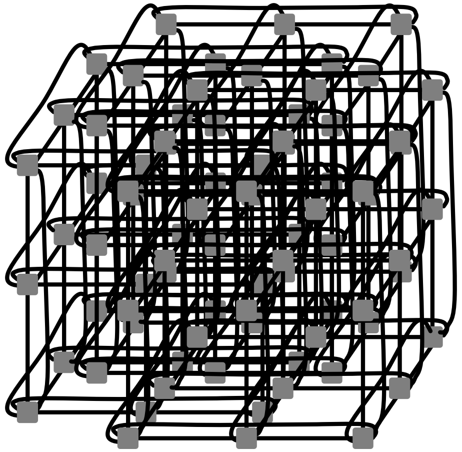
Hypercube



NASA Pleiades

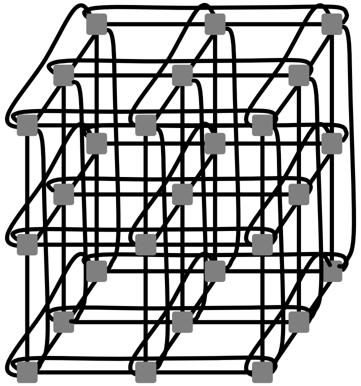
IBM BG/Q

Torus 5D



Comparison Targets
Low-radix topologies

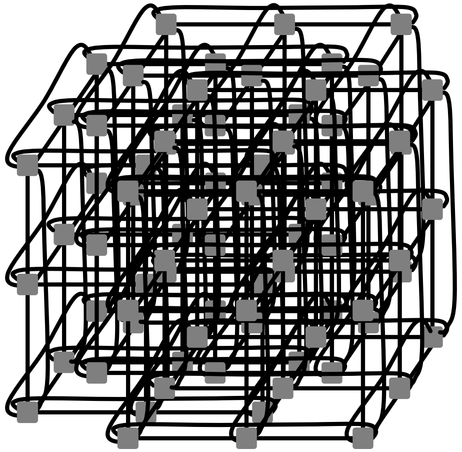
Torus 3D



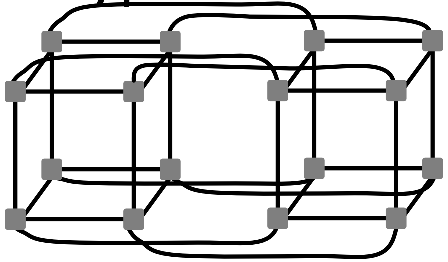
Cray XE6

IBM BG/Q

Torus 5D



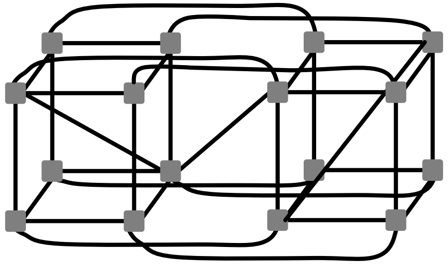
Hypercube



NASA Pleiades

Infinetics

Long Hop [1]

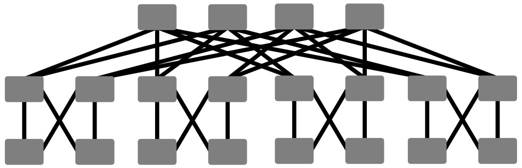


[1] Tomic, Ratko V. Optimal networks from error correcting codes. 2013 ACM/IEEE Symposium on Architectures for Networking and Communications Systems (ANCS)

Comparison Targets
High-radix topologies

Comparison Targets
High-radix topologies

Fat tree [1]

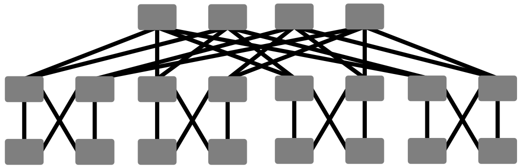


TSUBAME2.0

[1] C. E. Leiserson. Fat-trees: universal networks for hardware-efficient supercomputing. IEEE Transactions on Computers. 1985

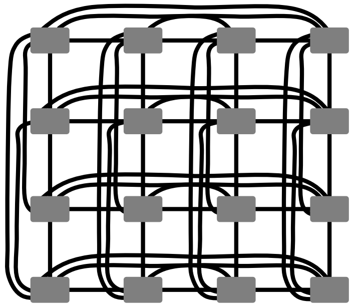
COMPARISON TARGETS
HIGH-RADIX TOPOLOGIES

Fat tree [1]



TSUBAME2.0

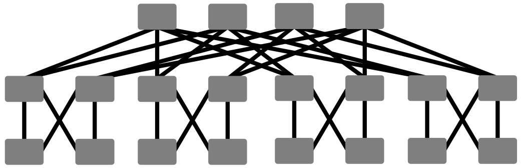
Flattened Butterfly [2]



[1] C. E. Leiserson. Fat-trees: universal networks for hardware-efficient supercomputing. IEEE Transactions on Computers. 1985
[2] J. Kim, W. J. Dally, D. Abts. Flattened butterfly: a cost-efficient topology for high-radix networks. ISCA'07

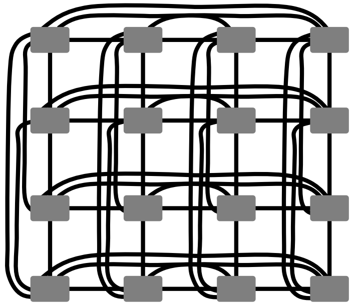
Comparison Targets
High-radix topologies

Fat tree [1]

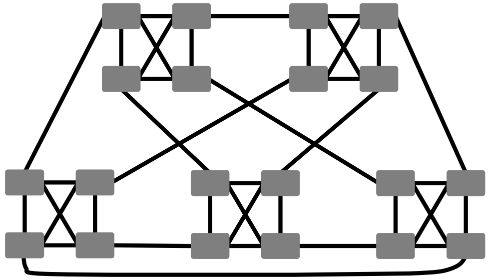


TSUBAME2.0

Flattened Butterfly [2]



Dragonfly [3]

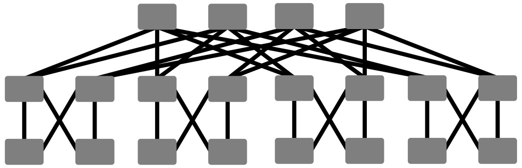


Cray Cascade

[1] C. E. Leiserson. Fat-trees: universal networks for hardware-efficient supercomputing. IEEE Transactions on Computers. 1985
[2] J. Kim, W. J. Dally, D. Abts. Flattened butterfly: a cost-efficient topology for high-radix networks. ISCA'07
[3] J. Kim, W. J. Dally, S. Scott, D. Abts. Technology-Driven, Highly-Scalable Dragonfly Topology. ISCA'08

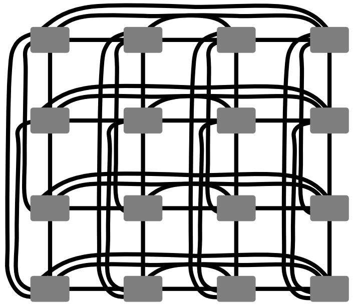
Comparison Targets
High-radix topologies

Fat tree [1]

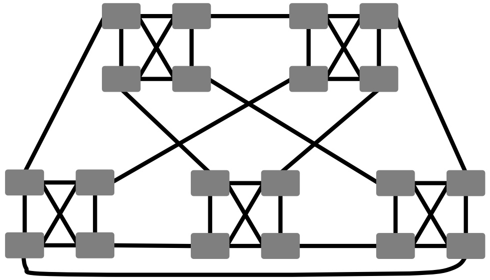


TSUBAME2.0

Flattened Butterfly [2]

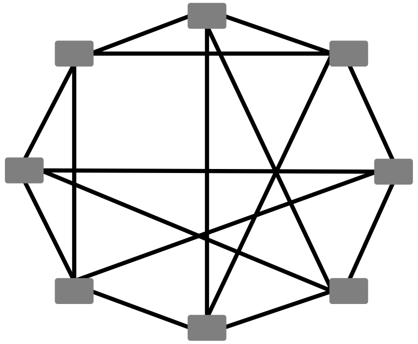


Dragonfly [3]



Cray Cascade

Random Topologies [4]



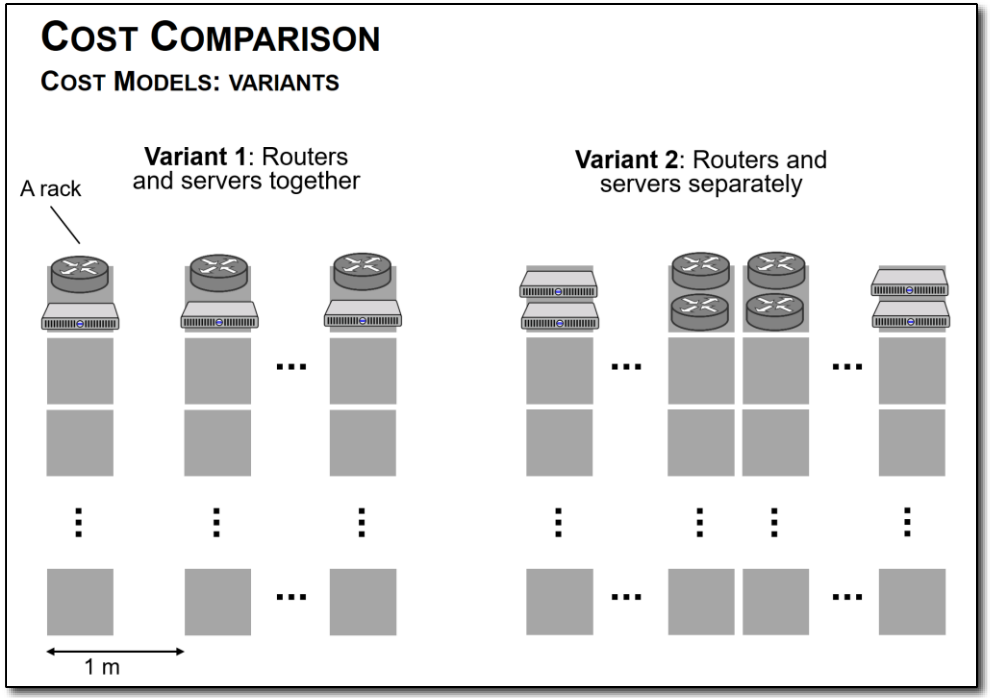
[1] C. E. Leiserson. Fat-trees: universal networks for hardware-efficient supercomputing. IEEE Transactions on Computers. 1985
[2] J. Kim, W. J. Dally, D. Abts. Flattened butterfly: a cost-efficient topology for high-radix networks. ISCA'07
[3] J. Kim, W. J. Dally, S. Scott, D. Abts. Technology-Driven, Highly-Scalable Dragonfly Topology. ISCA'08
[4] M. Koibuchi, H. Matsutani, H. Amano, D. F. Hsu, H. Casanova. A case for random shortcut topologies for HPC interconnects. ISCA'12

Cost of Network Construction

Models, Variants

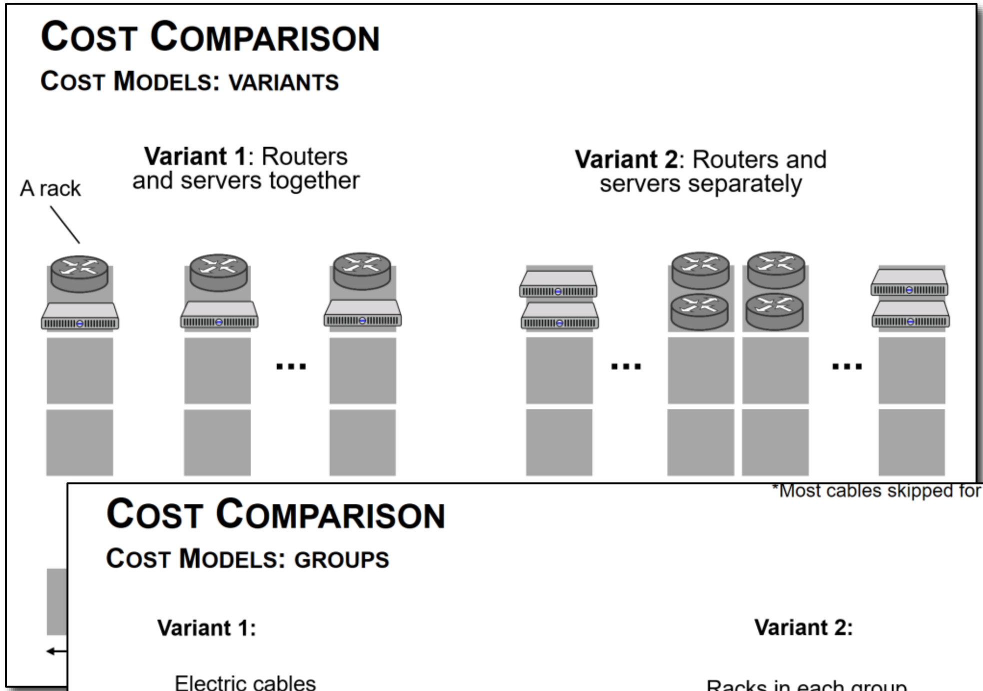
Cost of Network Construction
Models, Variants

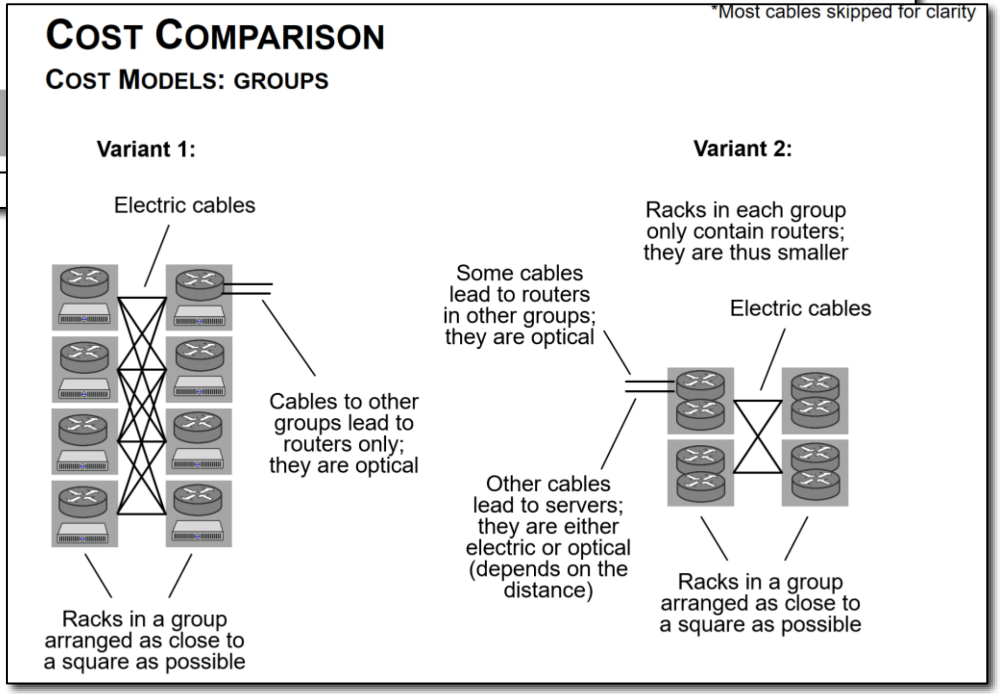
Cluster structure



COST OF NETWORK CONSTRUCTION
MODELS, VARIANTS

Cluster structure

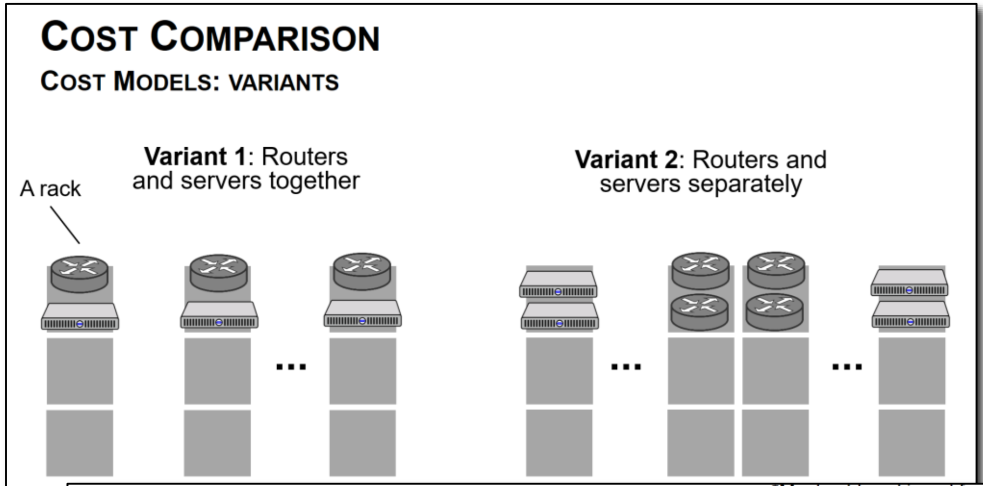




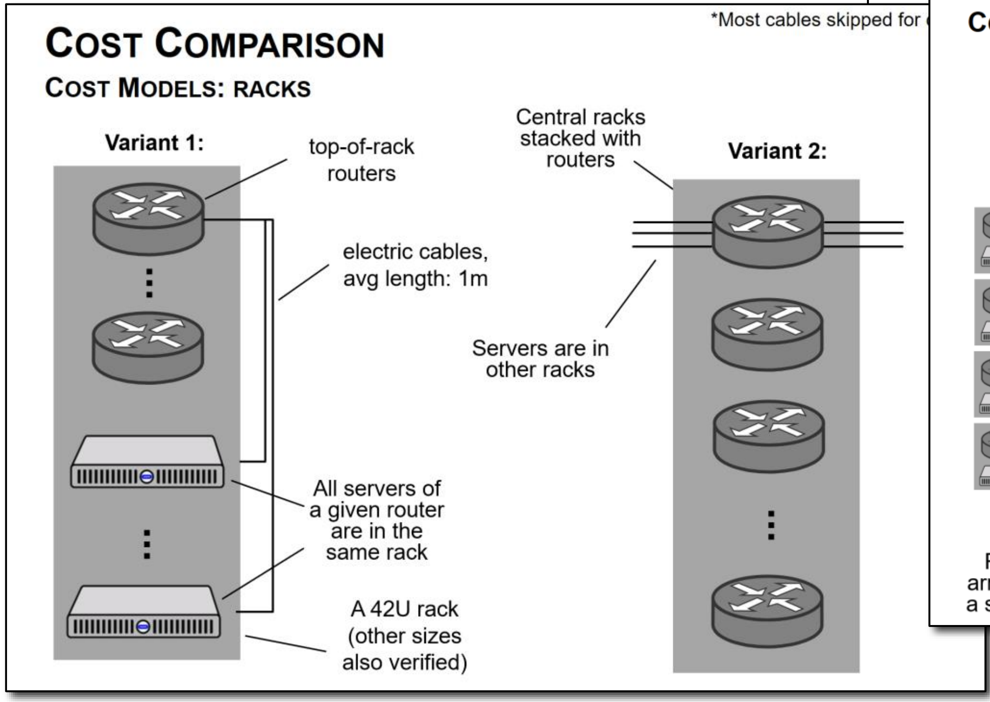
Structure of router groups

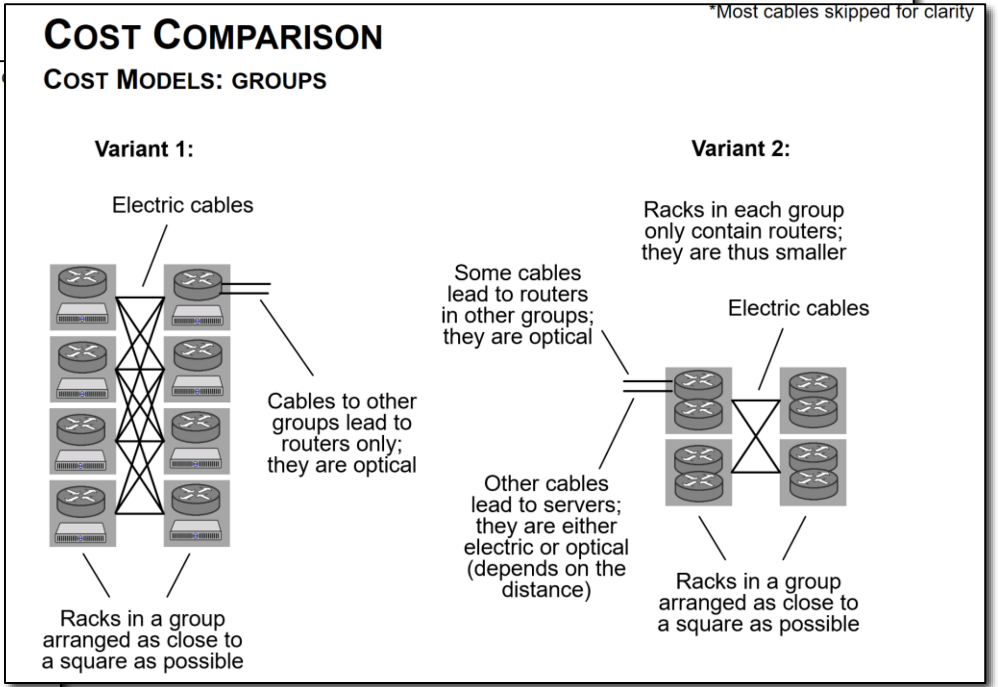
Cost of Network Construction
Models, Variants

Cluster structure



Rack structure

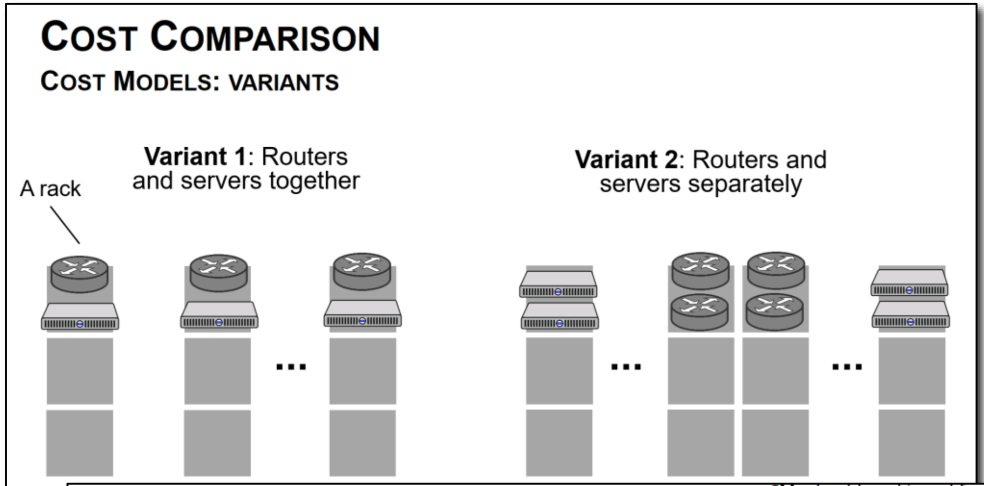




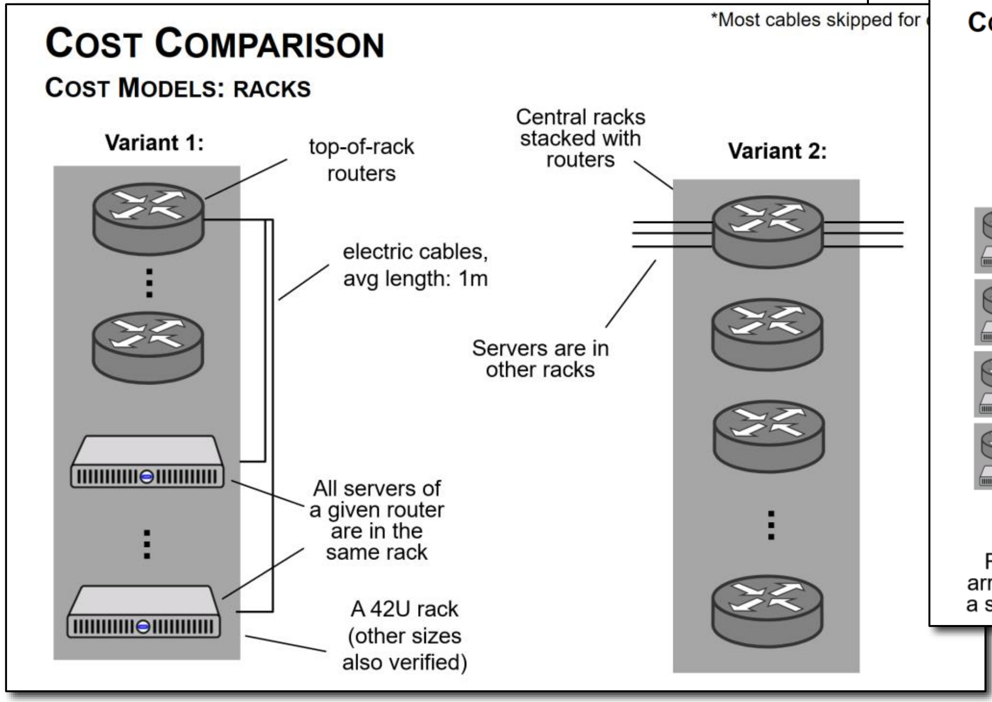
Structure of router groups

Cost of Network Construction
Models, Variants

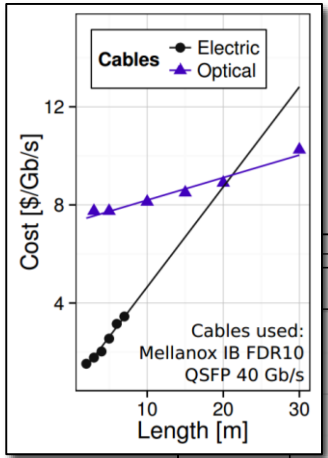
Cluster structure

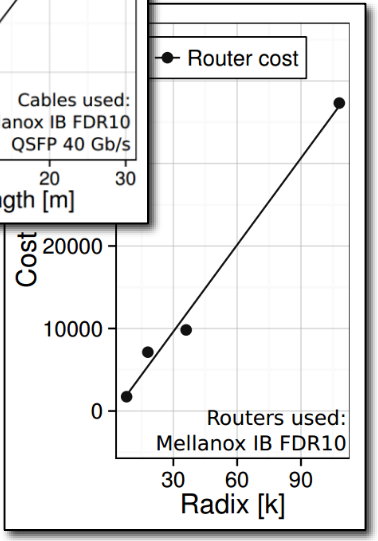


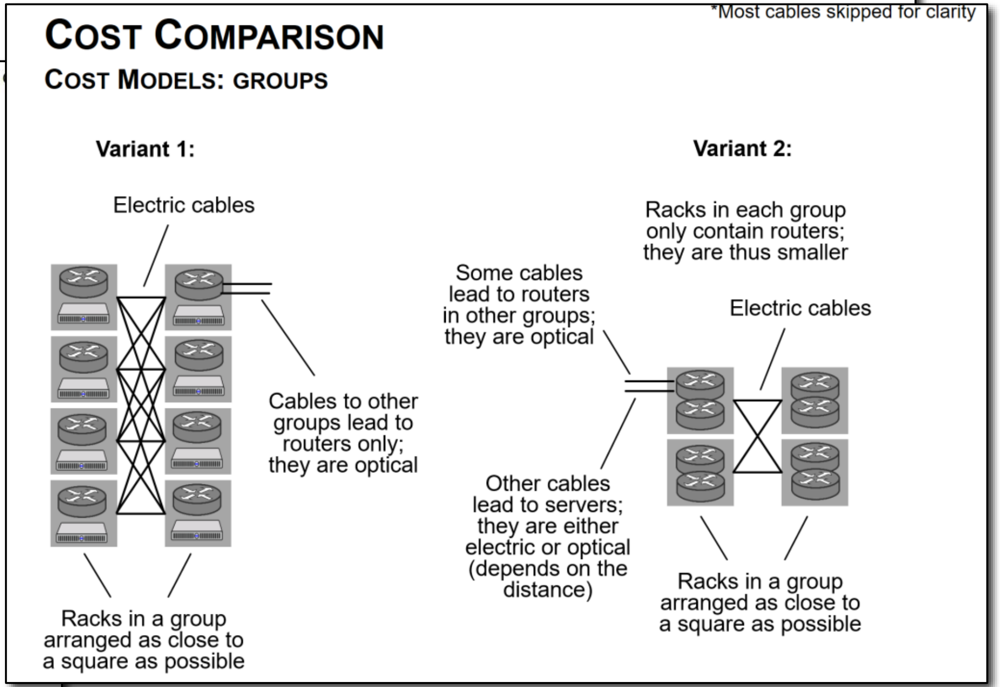
Rack structure



Models





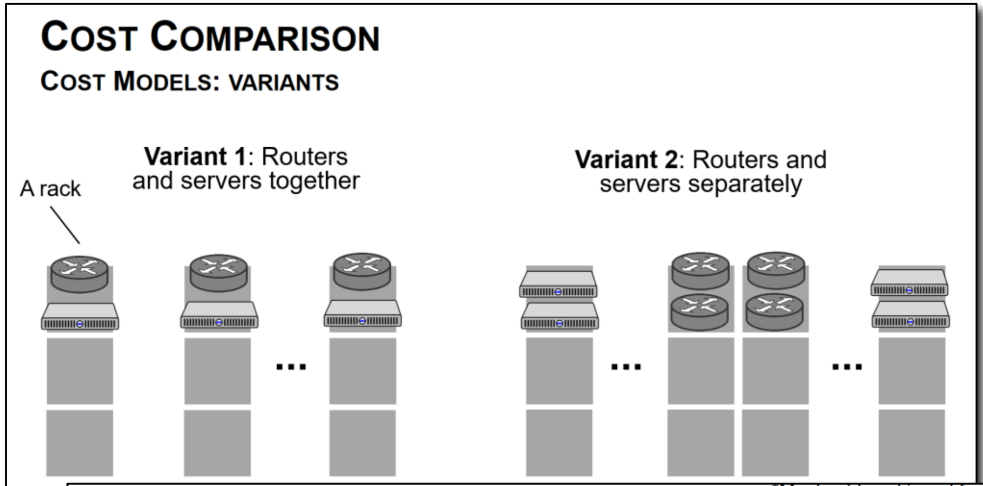


Structure of router groups

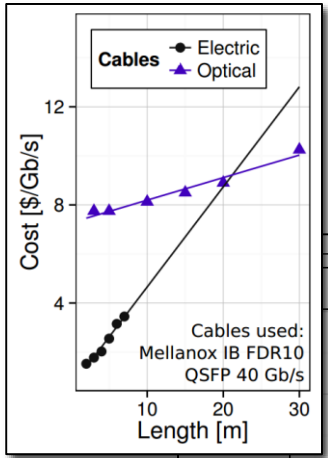
Cost of Network Construction
Models, Variants

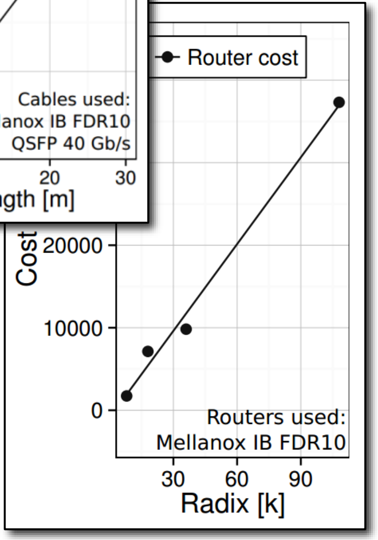
Routers

Cluster structure

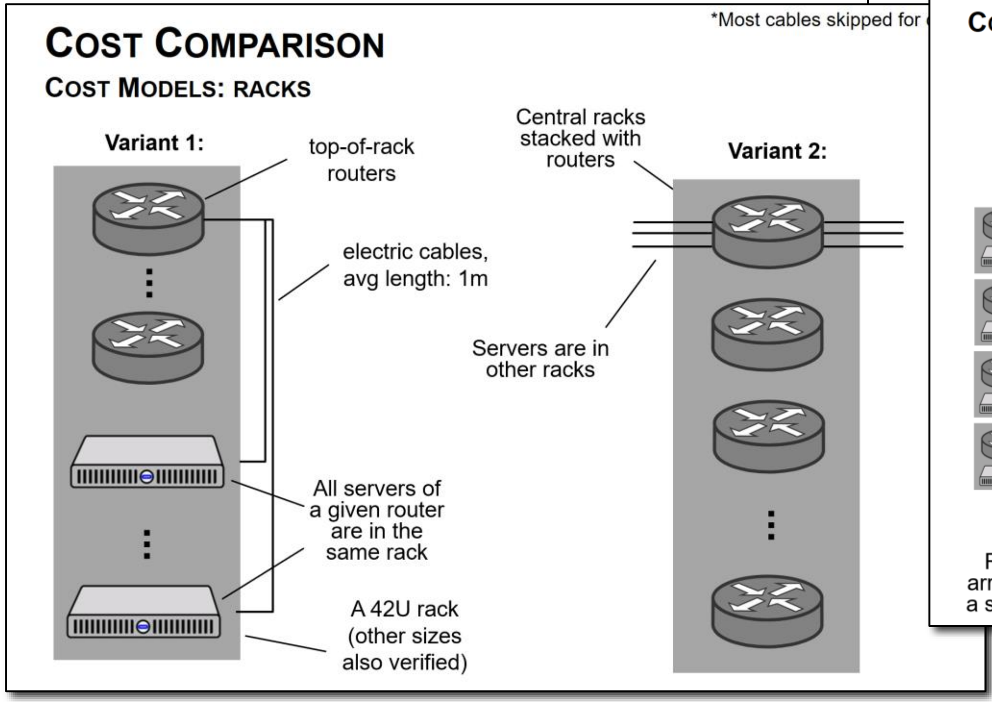


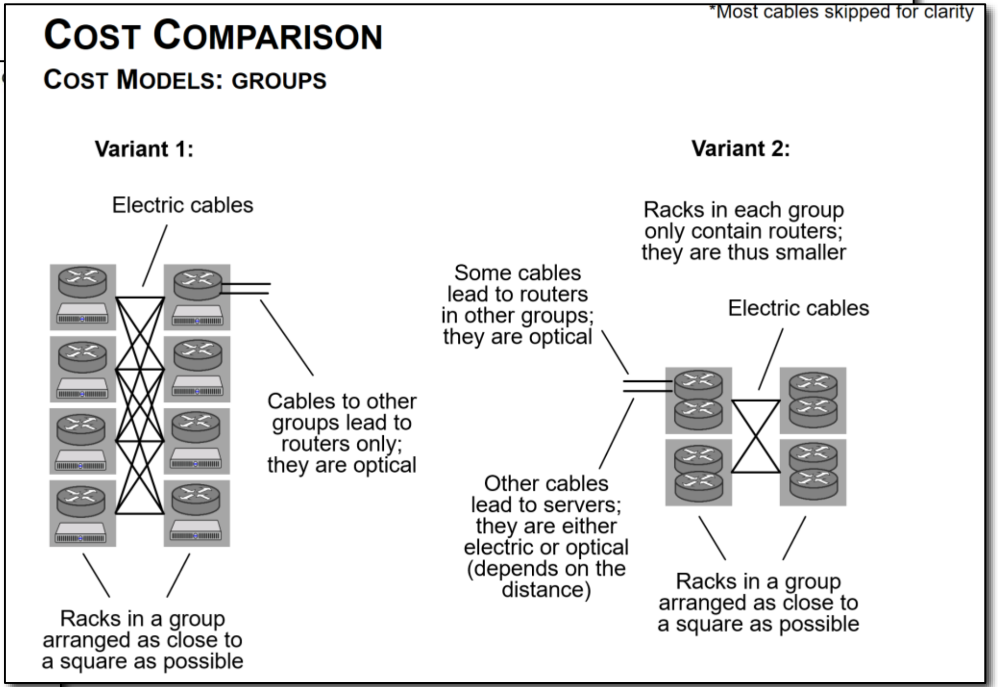
Models





Rack structure



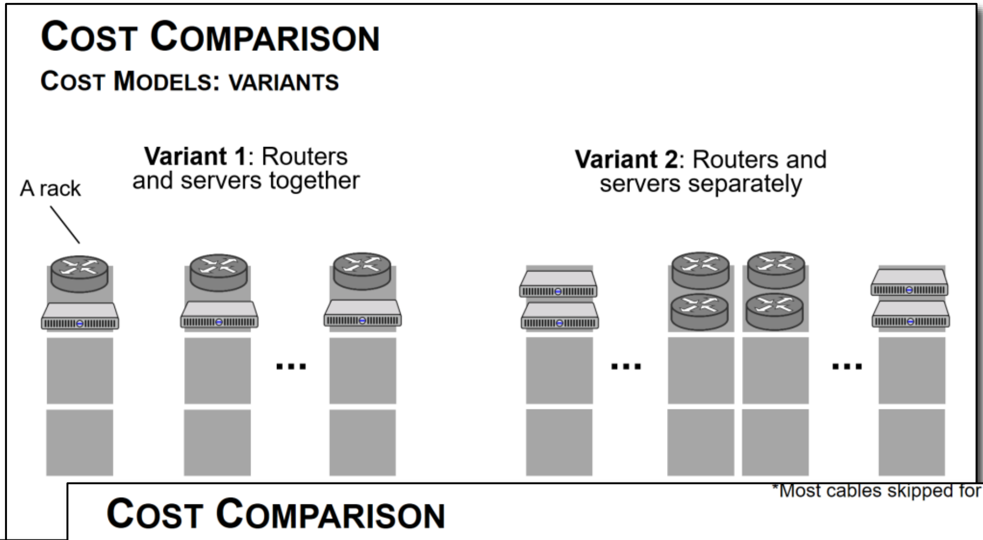


Structure of router groups

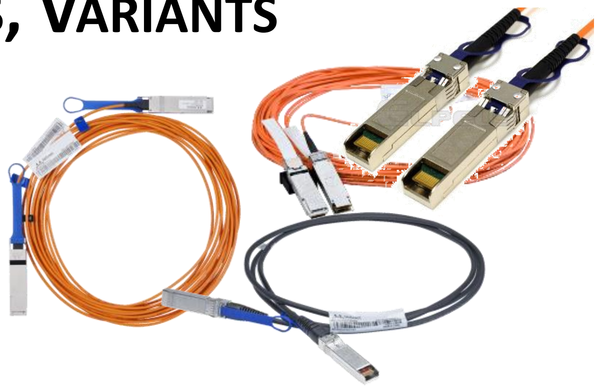
Cost of Network Construction Models, Variants

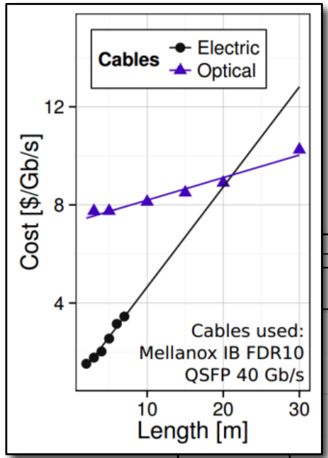
Routers

Cluster structure

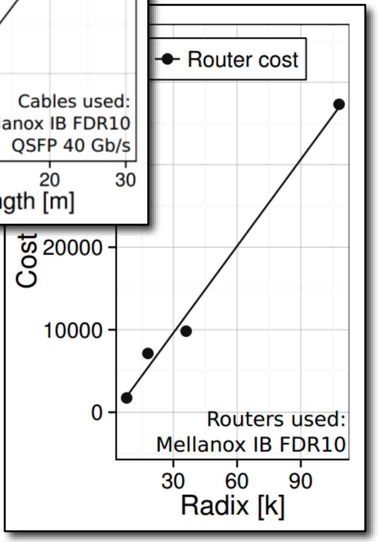


Cables

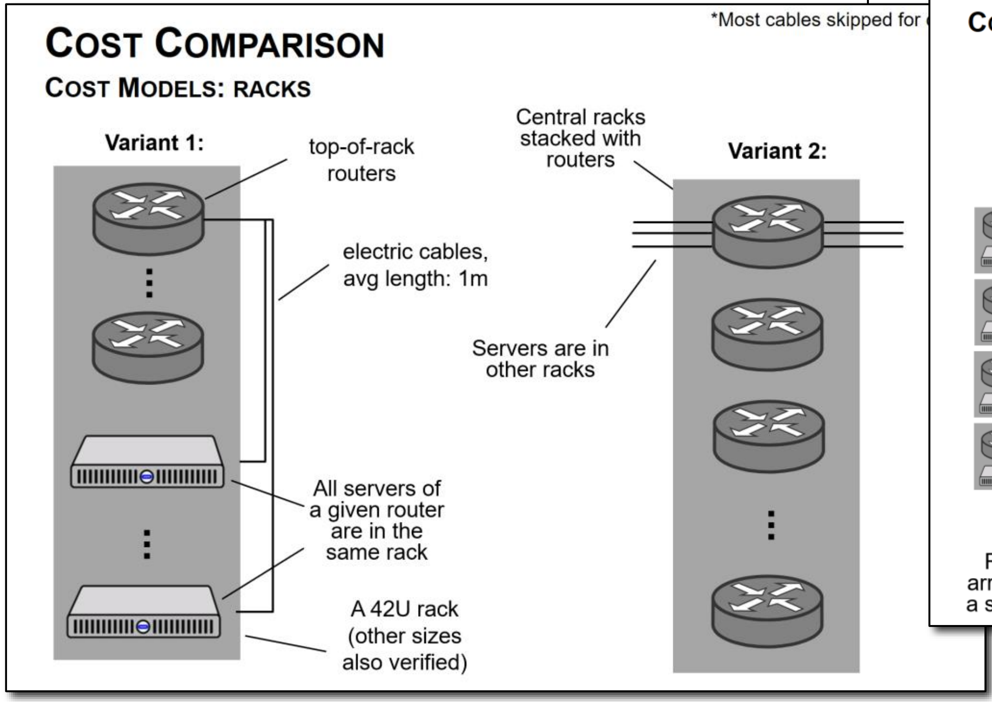


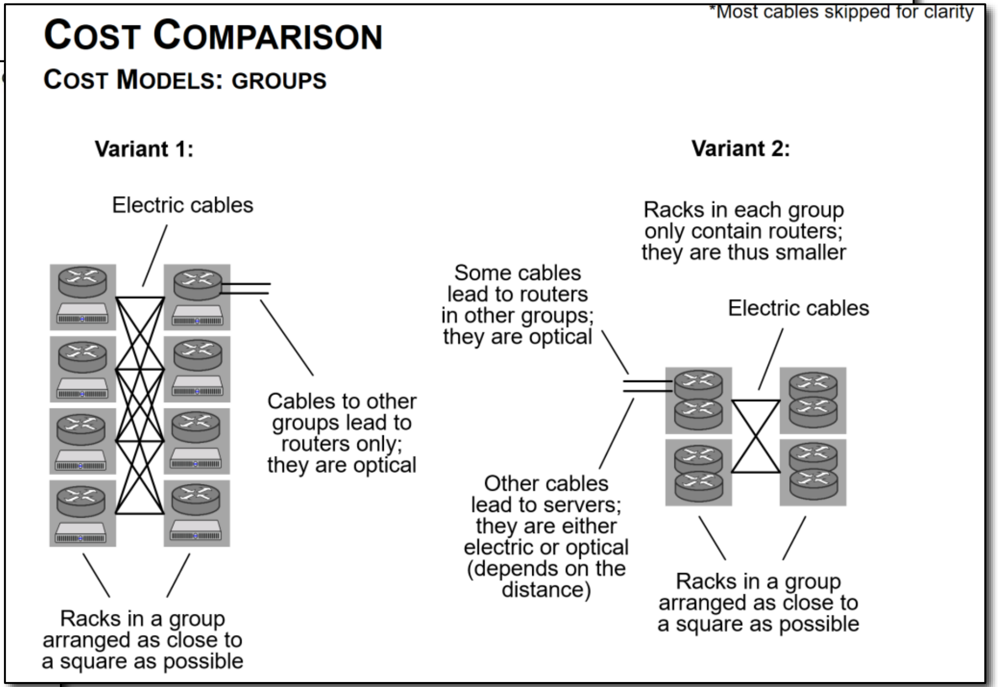


Models



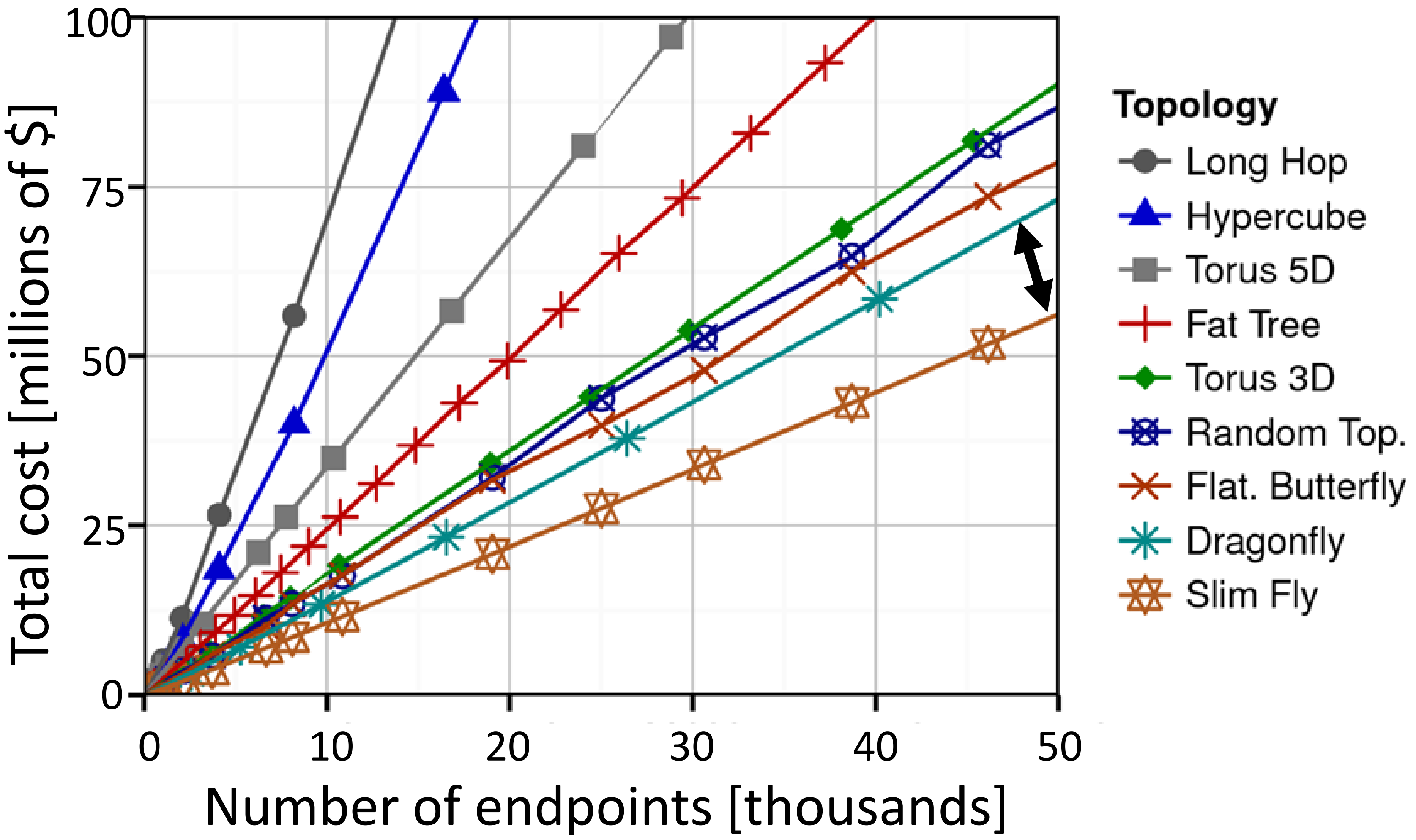
Rack structure



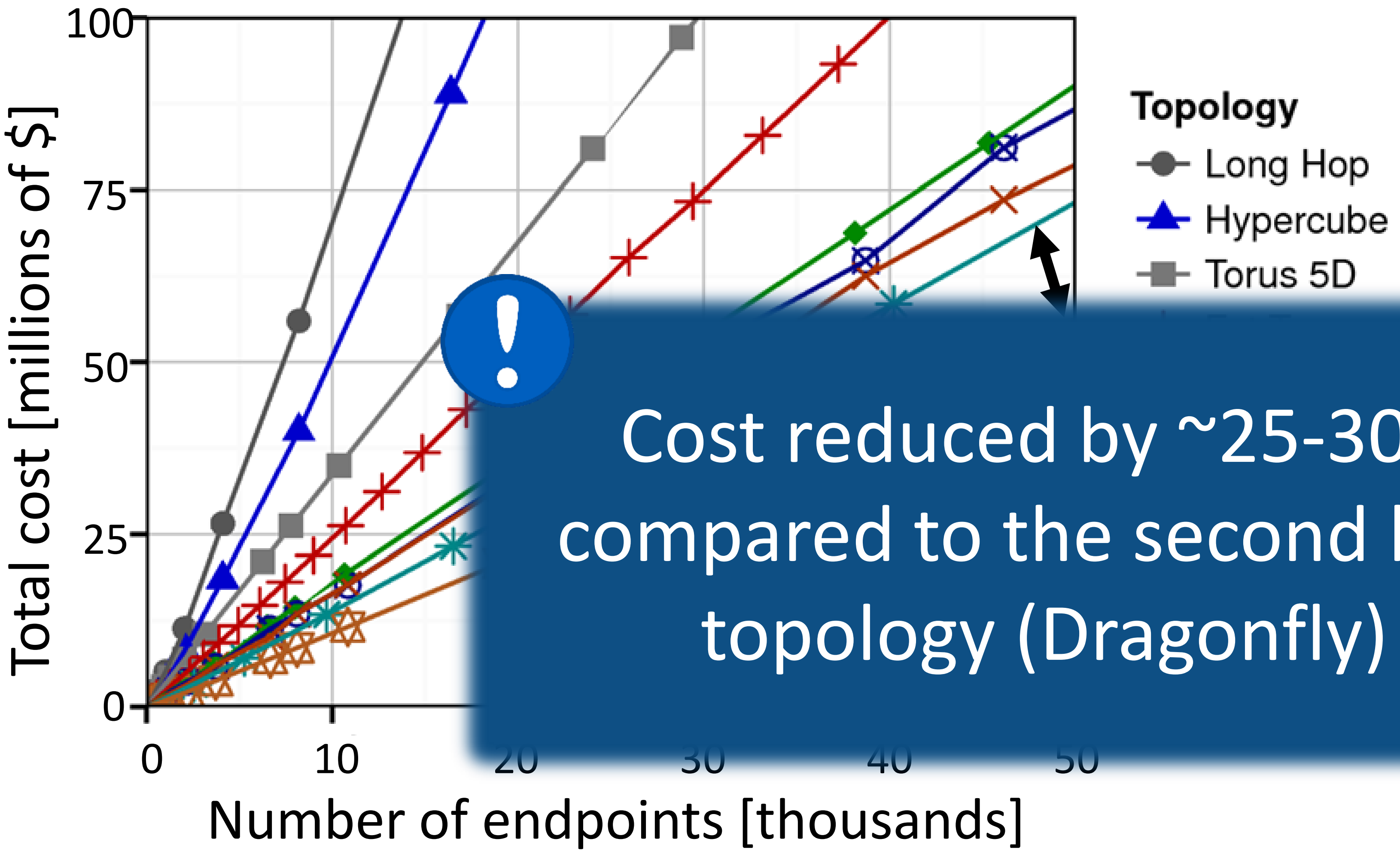


Structure of router groups

Results: Cost of Network Construction



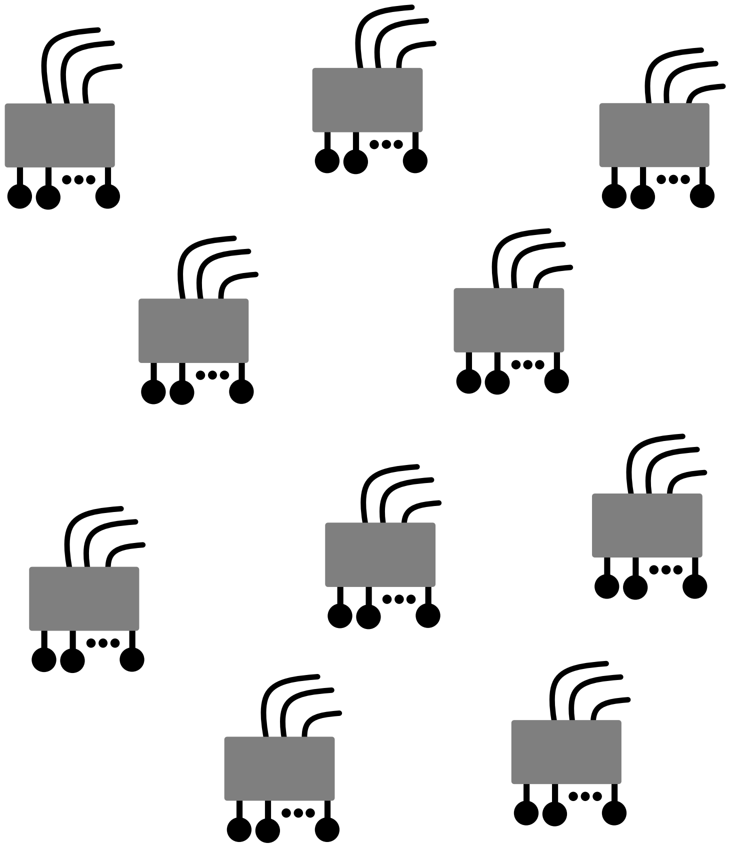
RESULTS: COST OF NETWORK CONSTRUCTION



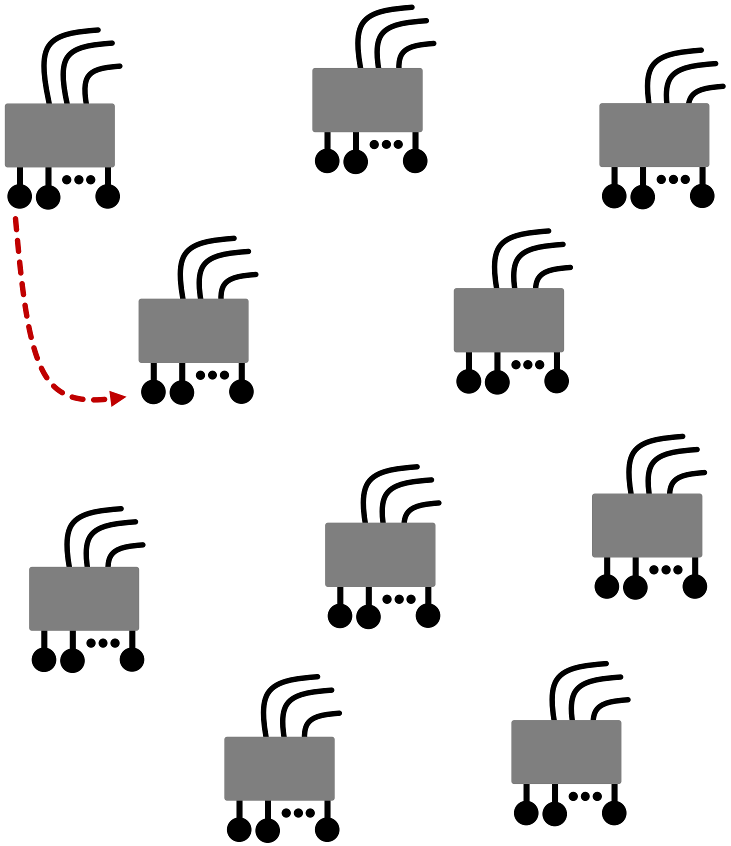
Cost reduced by ~25-30% compared to the second best topology (Dragonfly)

Results: Performance

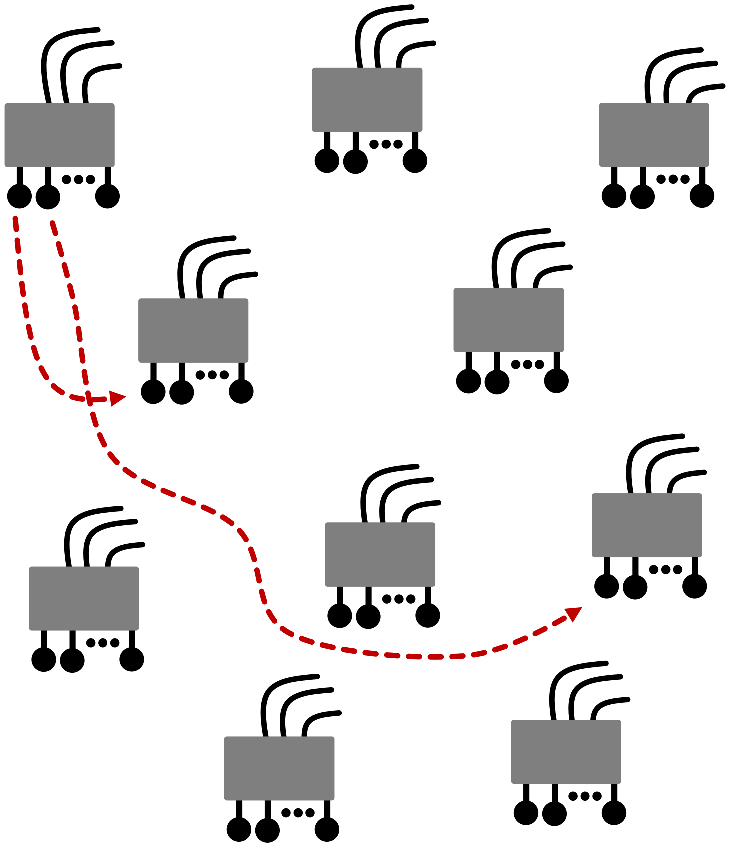
Results: Performance



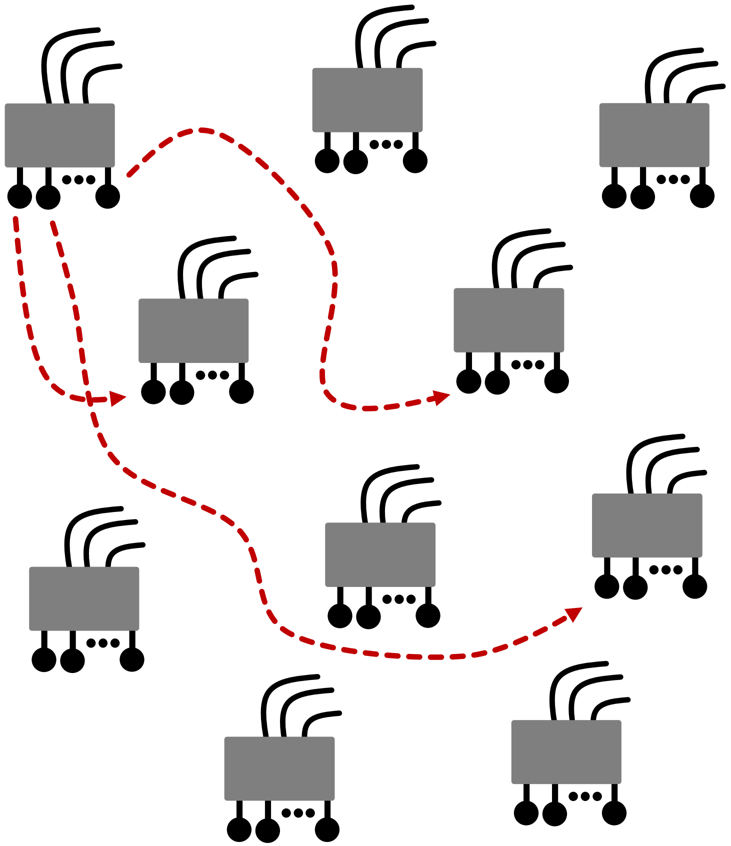
Results: Performance



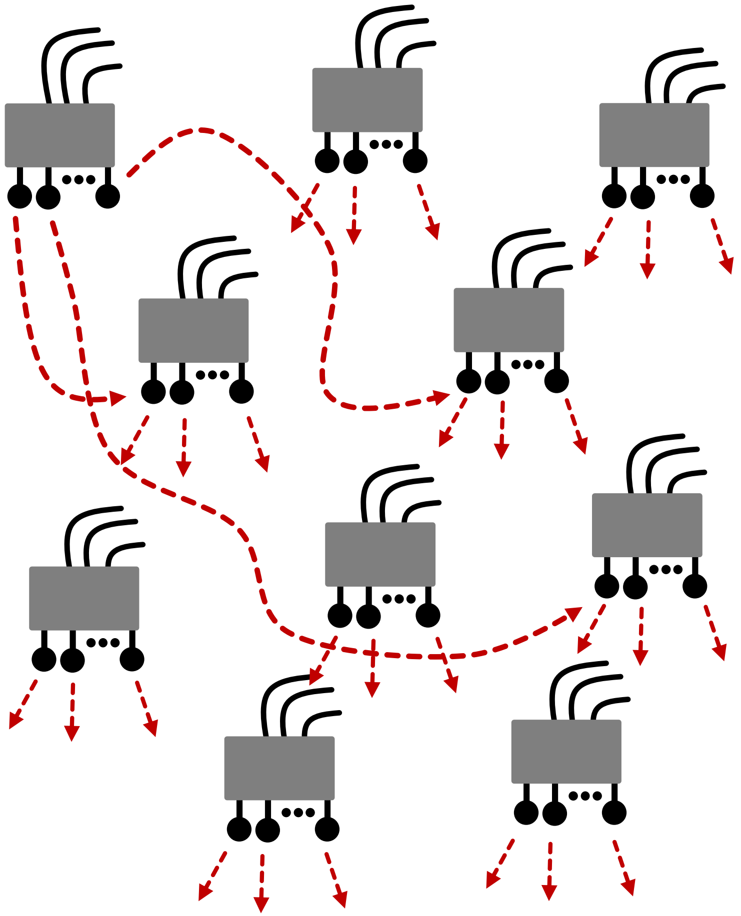
Results: Performance



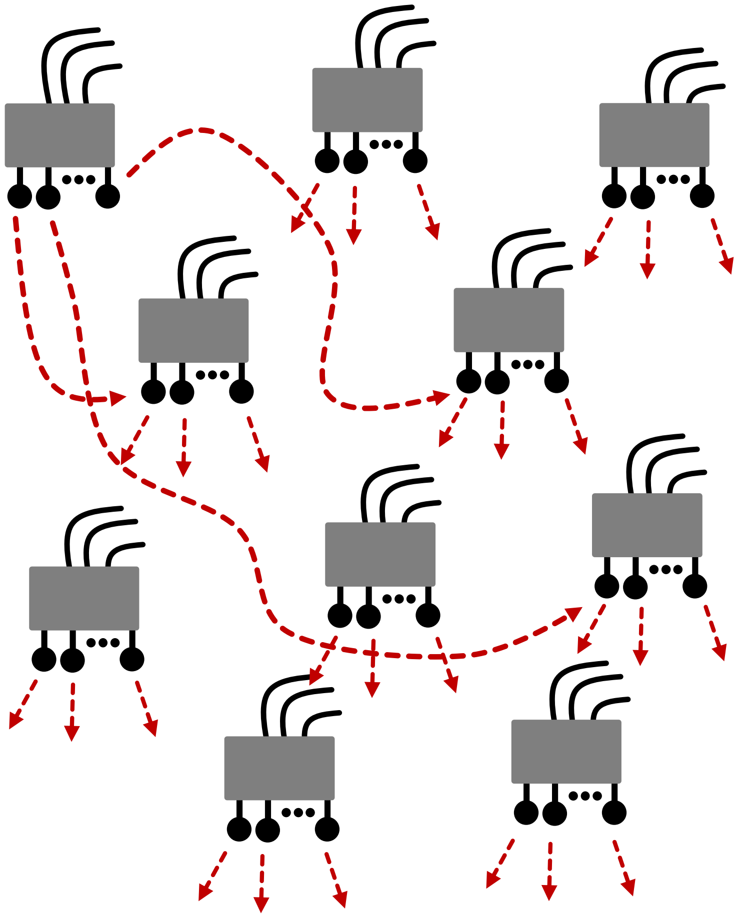
Results: Performance

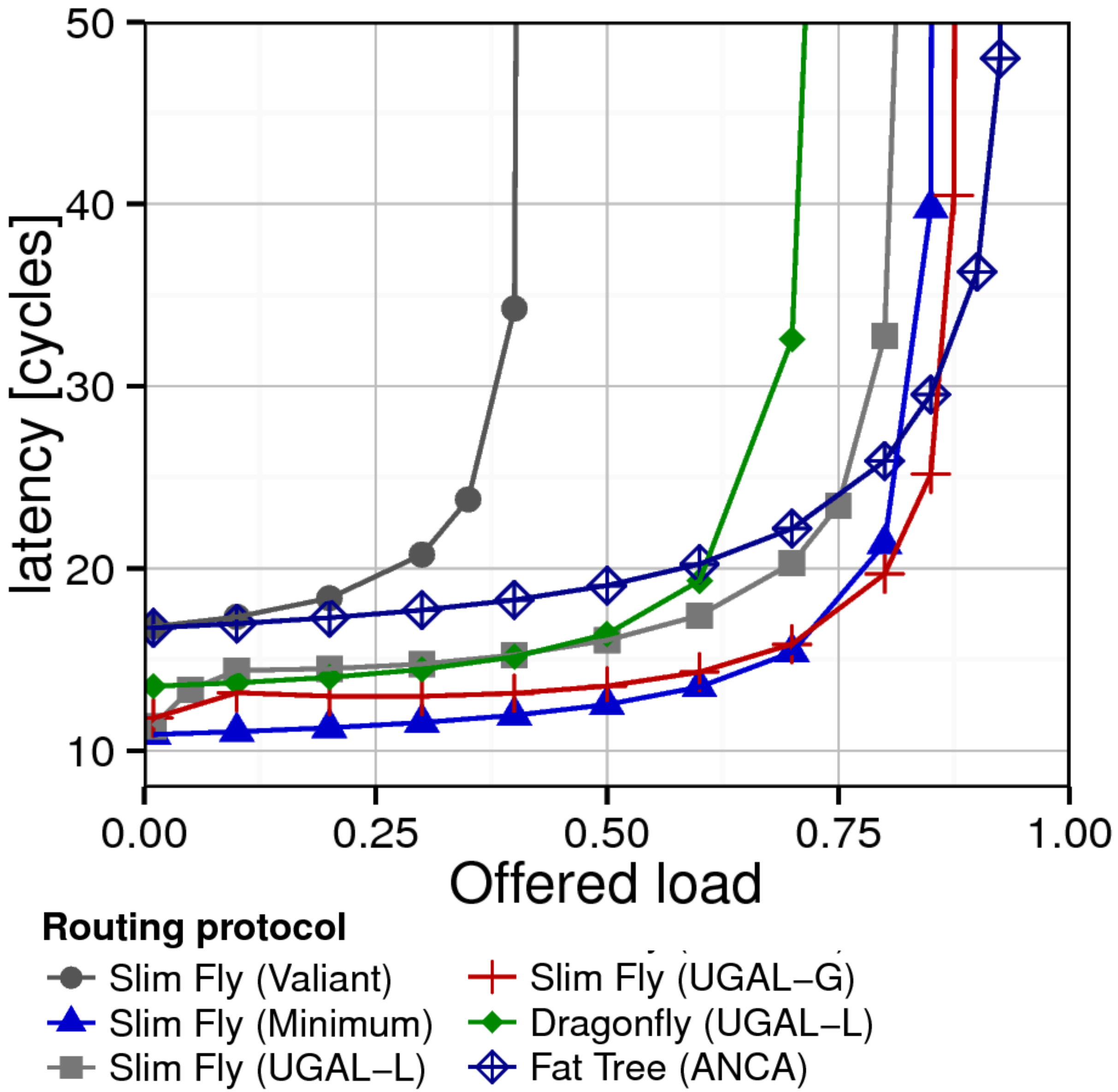


Results: Performance

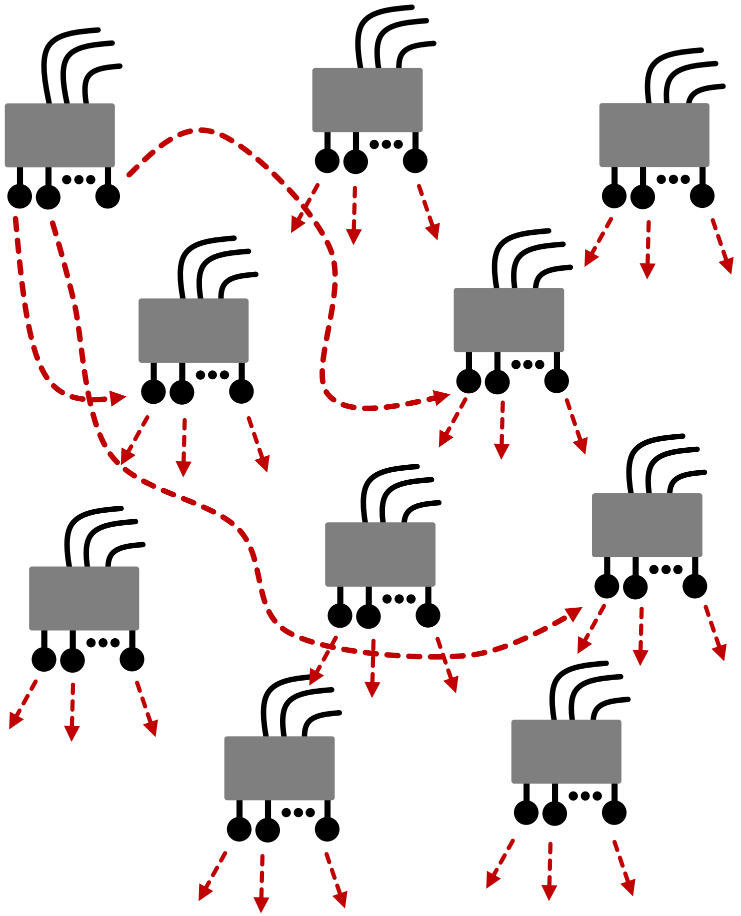


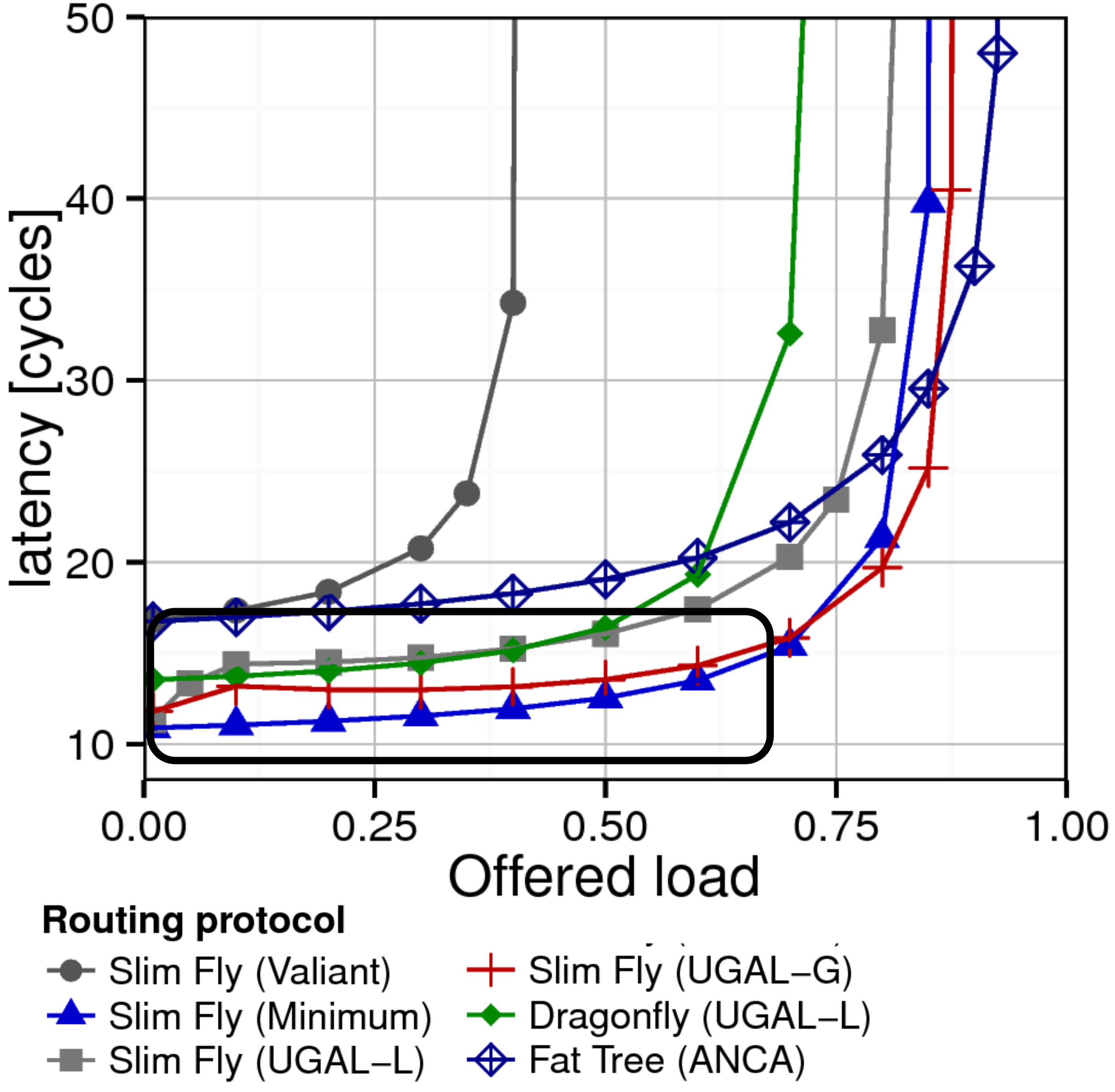
RESULTS: PERFORMANCE



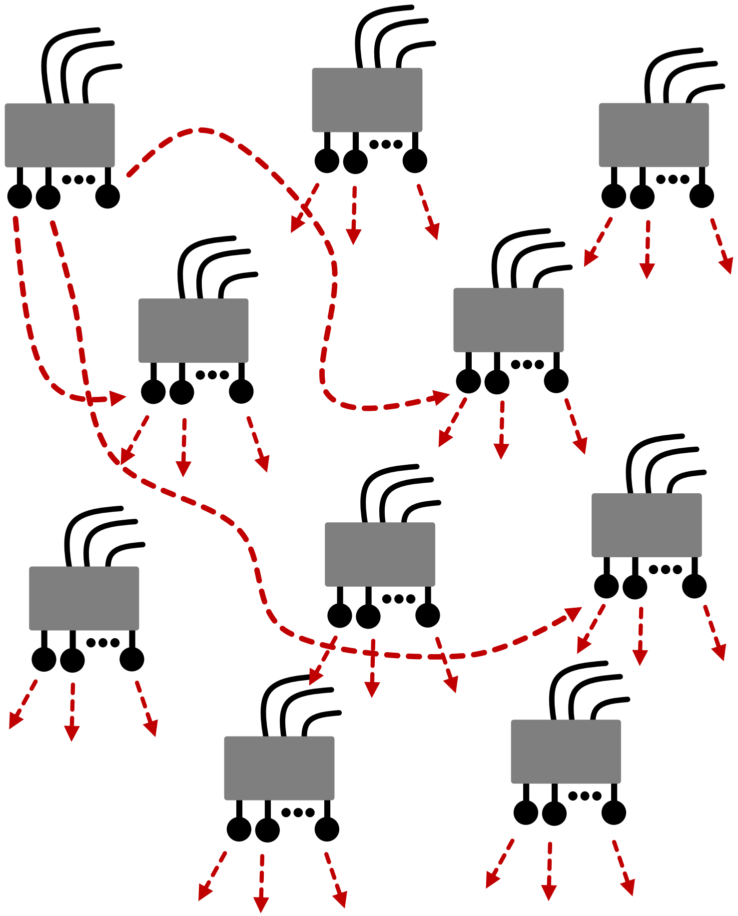


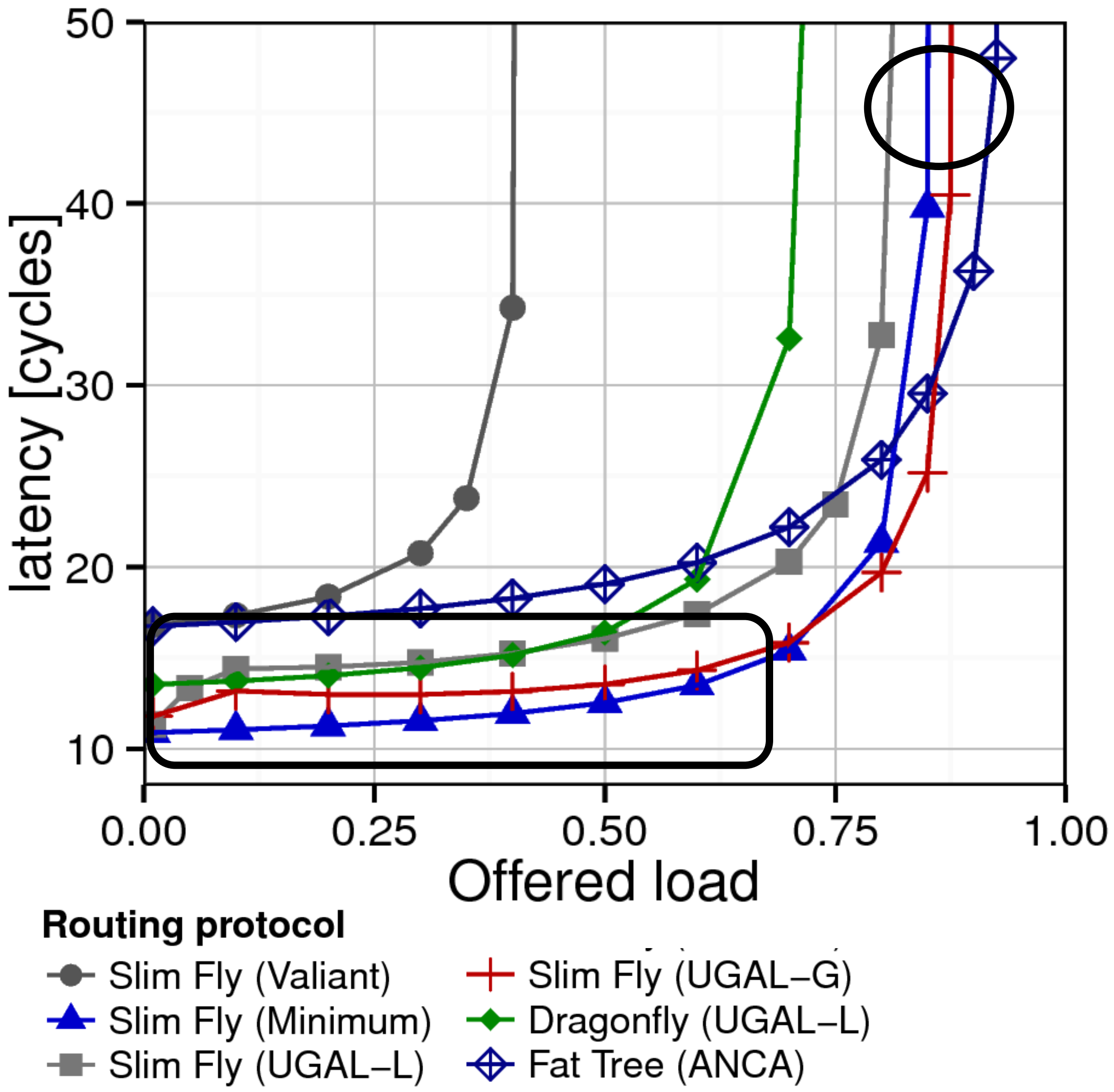
RESULTS: PERFORMANCE



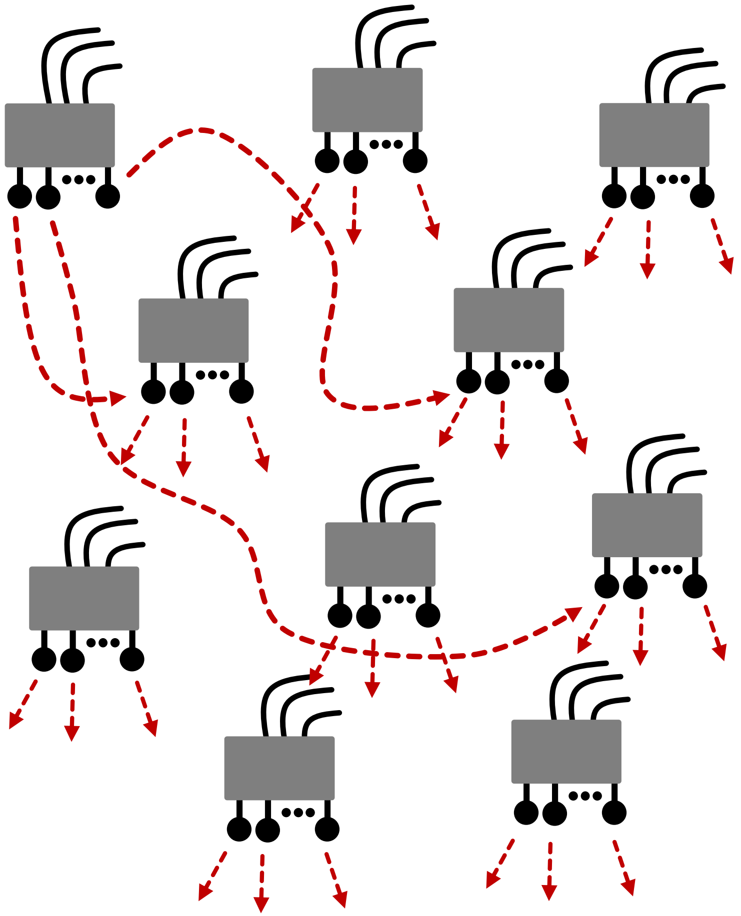


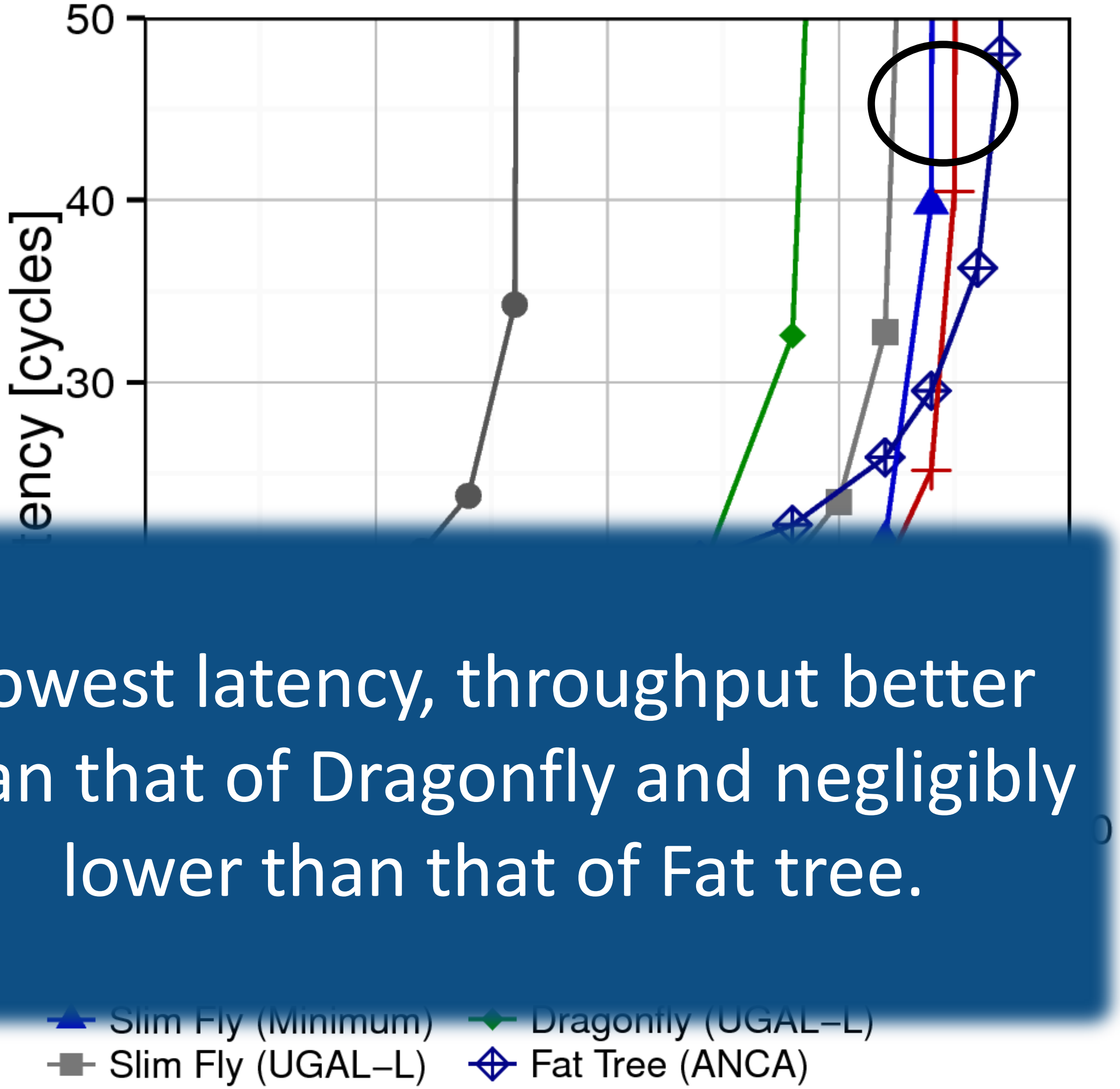
RESULTS: PERFORMANCE





Results: Performance





Lowest latency, throughput better than that of Dragonfly and negligibly lower than that of Fat tree.

Results: A Case Study

Results: A Case Study

- A Slim Fly with;
  - N = 10,830
  - k = 43
  - $N_r = 722$

RESULTS: A CASE STUDY

Topology					
Endpoints (N)	19,876	40,200	20,736	58,806	**10,830**
Routers (N_r)	2,311	4,020	1,728	5,346	**722**
Radix (k)	**43**	**43**	**43**	**43**	**43**
Electric cables	19,414	32,488	9,504	56,133	**6,669**
Fiber cables	40,215	33,842	20,736	29,524	**6,869**
Cost per node [$]	2,346	1,743	1,570	1,438	**1,033**
Power per node [W]	14.0	12.04	10.8	10.9	**8.02**

- A Slim Fly with;
  - N = 10,830
  - k = 43
  - N_r = 722

RESULTS: A CASE STUDY

Topology					Slim Fly
Endpoints (N)	19,876	40,200	20,736	58,806	**10,830**
Routers (N_r)	2,311	4,020	1,728	5,346	**722**
Radix (k)	**43**	**43**	**43**	**43**	**43**
Electric cables	19,414	32,488	9,504	56,133	**6,669**
Fiber cables	40,215	33,842	20,736	29,524	**6,869**
Cost per node [$]	2,346	1,743	1,570	1,438	**1,033**
Power per node [W]	14.0	12.04	10.8	10.9	**8.02**

- A Slim Fly with;
  - N = 10,830
  - k = 43
  - N_r = 722

Results: A Case Study

Topology	Fat tree	Random	Flat. Butterfly	Dragonfly	**Slim Fly**
Endpoints (N)	19,876	40,200	20,736	58,806	**10,830**
Routers (N_r)	2,311	4,020	1,728	5,346	**722**
Radix (k)	**43**	**43**	**43**	**43**	**43**
Electric cables	19,414	32,488	9,504	56,133	**6,669**
Fiber cables	40,215	33,842	20,736	29,524	**6,869**
Cost per node [$]	2,346	1,743	1,570	1,438	**1,033**
Power per node [W]	14.0	12.04	10.8	10.9	**8.02**

- A Slim Fly with;
  - N = 10,830
  - k = 43
  - N_r = 722

Results: A Case Study

Topology	Fat tree	Random	Flat. Butterfly	Dragonfly	**Slim Fly**
Endpoints (N)	19,876	40,200	20,736	58,806	**10,830**
Routers (N_r)	2,311	4,020	1,728	5,346	**722**
Radix (k)	**43**	**43**	**43**	**43**	**43**
Electric cables	19,414	32,488	9,504	56,133	**6,669**
Fiber cables	40,215	33,842	20,736	29,524	**6,869**
Cost per node [$]	2,346	1,743	1,570	1,438	**1,033**
Power per node [W]	14.0	12.04	10.8	10.9	**8.02**

- A Slim Fly with;
  - N = 10,830
  - k = 43
  - N_r = 722

Results: A Case Study

Topology	Fat tree	Random	Flat. Butterfly	Dragonfly	Slim Fly
Endpoints (N)	19,876	40,200	20,736	58,806	**10,830**
Routers (N_r)	2,311	4,020	1,728	5,346	**722**
Radix (k)	**43**	**43**	**43**	**43**	**43**
Electric cables	19,414	32,488	9,504	56,133	**6,669**
Fiber cables	40,215	33,842	20,736	29,524	**6,869**
Cost per node [$]	2,346	1,743	1,570	1,438	**1,033**
Power per node [W]	14.0	12.04	10.8	10.9	**8.02**

- A Slim Fly with;
  - N = 10,830
  - k = 43
  - N_r = 722

Results: A Case Study

- A Slim Fly with;
  - N = 10,830
  - k = 43
  - N_r = 722

Topology	Fat tree	Random	Flat. Butterfly	Dragonfly	Slim Fly
Endpoints (N)	19,876	40,200	20,736	58,806	**10,830**
Routers (N_r)	2,311	4,020	1,728	5,346	**722**
Radix (k)	**43**	**43**	**43**	**43**	**43**
Electric cables	19,414	32,488	9,504	56,133	**6,669**
Fiber cables	40,215	33,842	20,736	29,524	**6,869**
Cost per node [$]	2,346	1,743	1,570	1,438	**1,033**
Power per node [W]	14.0	12.04	10.8	10.9	**8.02**

Topology					
Endpoints (N)	**10,718**	**9,702**	**10,000**	**9,702**	**10,830**
Routers (N_r)	1,531	1,386	1,000	1,386	**722**
Radix (k)	35	28	33	27	**43**
Electric cables	7,350	6,837	4,500	9,009	**6,669**
Fiber cables	24,806	7,716	10,000	4,900	**6,869**
Cost per node [$]	2,315	1,566	1,535	1,342	**1,033**
Power per node [W]	14.0	11.2	10.8	10.8	**8.02**

Results: A Case Study

- A Slim Fly with;
  - N = 10,830
  - k = 43
  - N_r = 722

Topology	Fat tree	Random	Flat. Butterfly	Dragonfly	Slim Fly
Endpoints (N)	19,876	40,200	20,736	58,806	**10,830**
Routers (N_r)	2,311	4,020	1,728	5,346	**722**
Radix (k)	**43**	**43**	**43**	**43**	**43**
Electric cables	19,414	32,488	9,504	56,133	**6,669**
Fiber cables	40,215	33,842	20,736	29,524	**6,869**
Cost per node [$]	2,346	1,743	1,570	1,438	**1,033**
Power per node [W]	14.0	12.04	10.8	10.9	**8.02**

Topology					Slim Fly
Endpoints (N)	**10,718**	**9,702**	**10,000**	**9,702**	**10,830**
Routers (N_r)	1,531	1,386	1,000	1,386	**722**
Radix (k)	35	28	33	27	**43**
Electric cables	7,350	6,837	4,500	9,009	**6,669**
Fiber cables	24,806	7,716	10,000	4,900	**6,869**
Cost per node [$]	2,315	1,566	1,535	1,342	**1,033**
Power per node [W]	14.0	11.2	10.8	10.8	**8.02**

Results: A Case Study

- A Slim Fly with;
  - N = 10,830
  - k = 43
  - N_r = 722

Topology	Fat tree	Random	Flat. Butterfly	Dragonfly	**Slim Fly**
Endpoints (N)	19,876	40,200	20,736	58,806	**10,830**
Routers (N_r)	2,311	4,020	1,728	5,346	**722**
Radix (k)	**43**	**43**	**43**	**43**	**43**
Electric cables	19,414	32,488	9,504	56,133	**6,669**
Fiber cables	40,215	33,842	20,736	29,524	**6,869**
Cost per node [$]	2,346	1,743	1,570	1,438	**1,033**
Power per node [W]	14.0	12.04	10.8	10.9	**8.02**

Topology	Fat tree	Random	Flat. Butterfly	Dragonfly	**Slim Fly**
Endpoints (N)	**10,718**	**9,702**	**10,000**	**9,702**	**10,830**
Routers (N_r)	1,531	1,386	1,000	1,386	**722**
Radix (k)	35	28	33	27	**43**
Electric cables	7,350	6,837	4,500	9,009	**6,669**
Fiber cables	24,806	7,716	10,000	4,900	**6,869**
Cost per node [$]	2,315	1,566	1,535	1,342	**1,033**
Power per node [W]	14.0	11.2	10.8	10.8	**8.02**

Results: A Case Study

- A Slim Fly with;
  - N = 10,830
  - k = 43
  - N_r = 722

Topology	Fat tree	Random	Flat. Butterfly	Dragonfly	**Slim Fly**
Endpoints (N)	19,876	40,200	20,736	58,806	**10,830**
Routers (N_r)	2,311	4,020	1,728	5,346	**722**
Radix (k)	**43**	**43**	**43**	**43**	**43**
Electric cables	19,414	32,488	9,504	56,133	**6,669**
Fiber cables	40,215	33,842	20,736	29,524	**6,869**
Cost per node [$]	2,346	1,743	1,570	1,438	**1,033**
Power per node [W]	14.0	12.04	10.8	10.9	**8.02**

Topology	Fat tree	Random	Flat. Butterfly	Dragonfly	**Slim Fly**
Endpoints (N)	**10,718**	**9,702**	**10,000**	**9,702**	**10,830**
Routers (N_r)	1,531	1,386	1,000	1,386	**722**
Radix (k)	35	28	33	27	**43**
Electric cables	7,350	6,837	4,500	9,009	**6,669**
Fiber cables	24,806	7,716	10,000	4,900	**6,869**
Cost per node [$]	2,315	1,566	1,535	1,342	**1,033**
Power per node [W]	14.0	11.2	10.8	10.8	**8.02**

Results: A Case Study

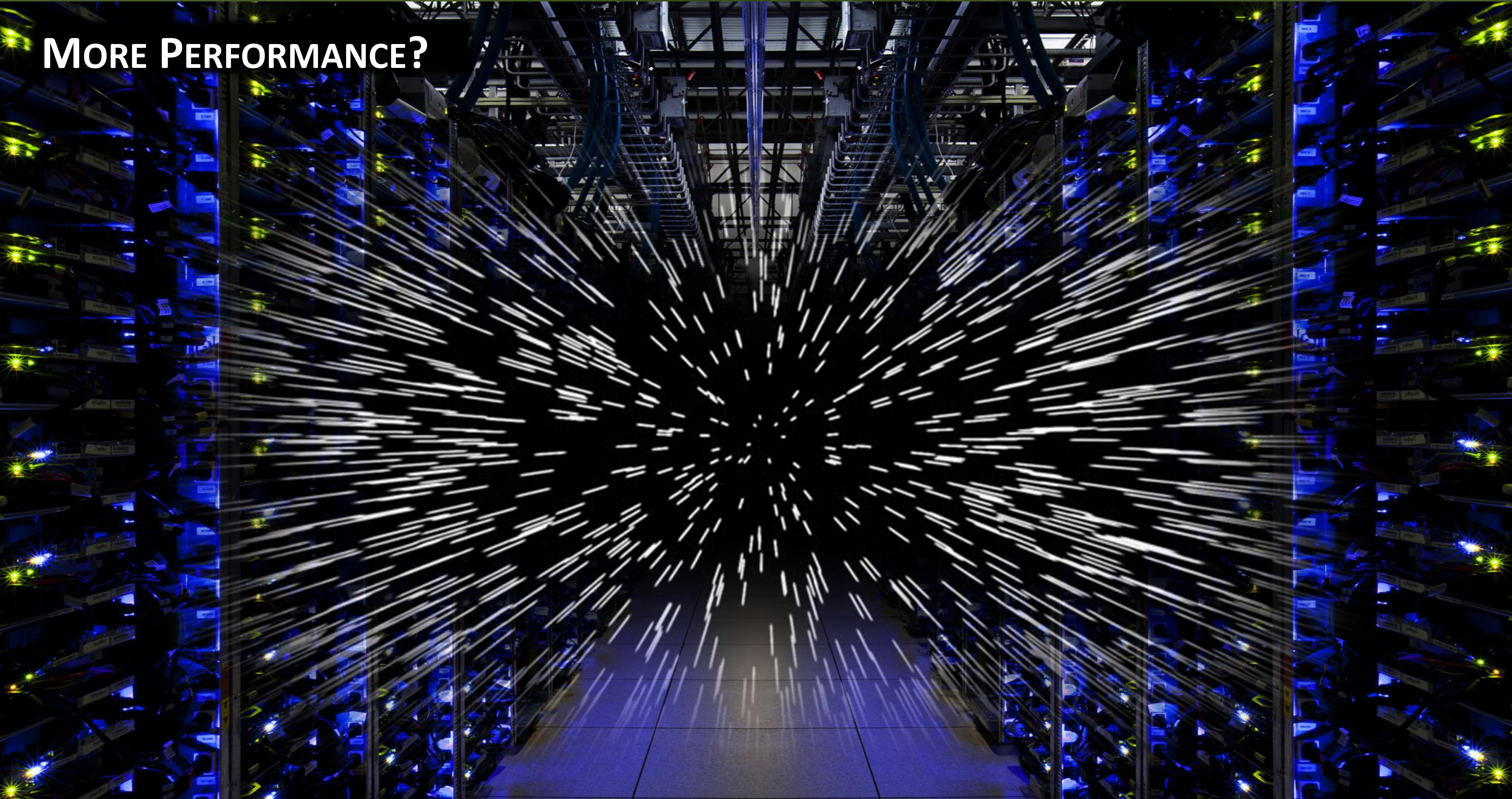
- A Slim Fly with;
  - N = 10,830
  - k = 43
  - N_r = 722

Topology	Fat tree	Random	Flat. Butterfly	Dragonfly	**Slim Fly**
Endpoints (N)	19,876	40,200	20,736	58,806	**10,830**
Routers (N_r)	2,311	4,020	1,728	5,346	**722**
Radix (k)	**43**	**43**	**43**	**43**	**43**
Electric cables	19,414	32,488	9,504	56,133	**6,669**
Fiber cables	40,215	33,842	20,736	29,524	**6,869**
Cost per node [$]	2,346	1,743	1,570	1,438	**1,033**
Power per node [W]	14.0	12.04	10.8	10.9	**8.02**

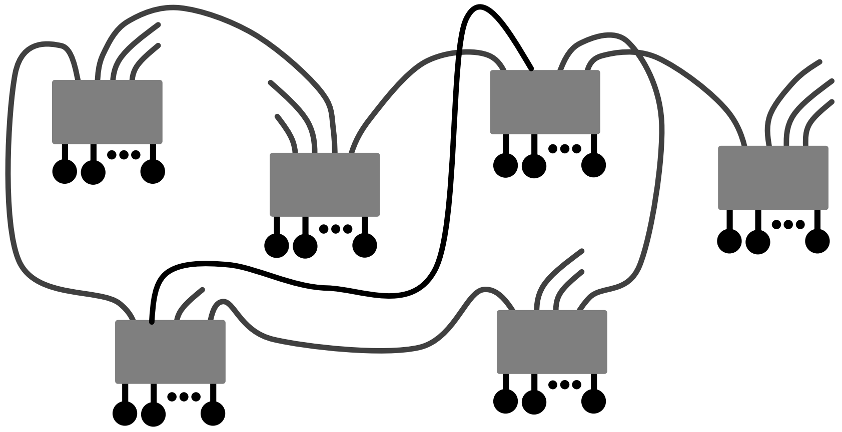
Topology	Fat tree	Random	Flat. Butterfly	Dragonfly	**Slim Fly**
Endpoints (N)	**10,718**	**9,702**	**10,000**	**9,702**	**10,830**
Routers (N_r)	1,531	1,386	1,000	1,386	**722**
Radix (k)	35	28	33	27	**43**
Electric cables	7,350	6,837	4,500	9,009	**6,669**
Fiber cables	24,806	7,716	10,000	4,900	**6,869**
Cost per node [$]	2,315	1,566	1,535	1,342	**1,033**
Power per node [W]	14.0	11.2	10.8	10.8	**8.02**

More Performance?

More Performance?

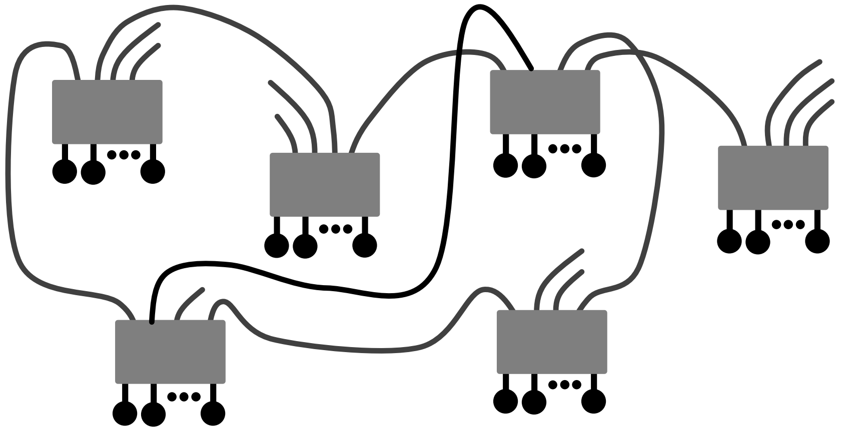


Multi Pathing?



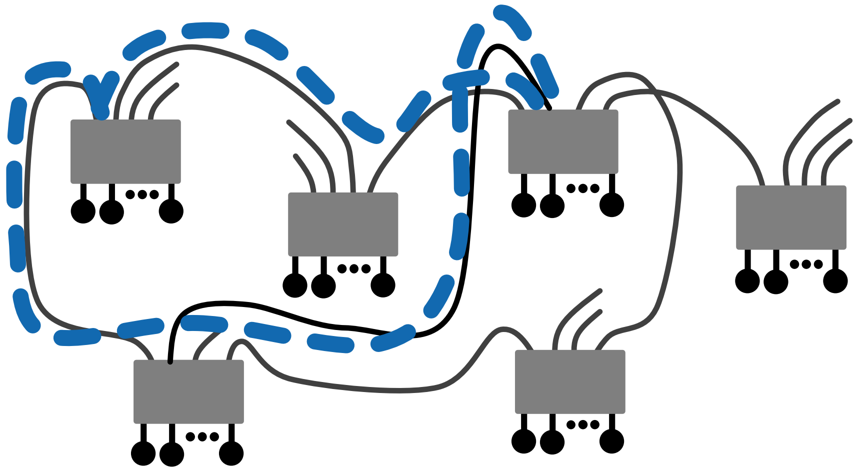
MULTI PATHING?

Minimal paths?



MULTI PATHING?

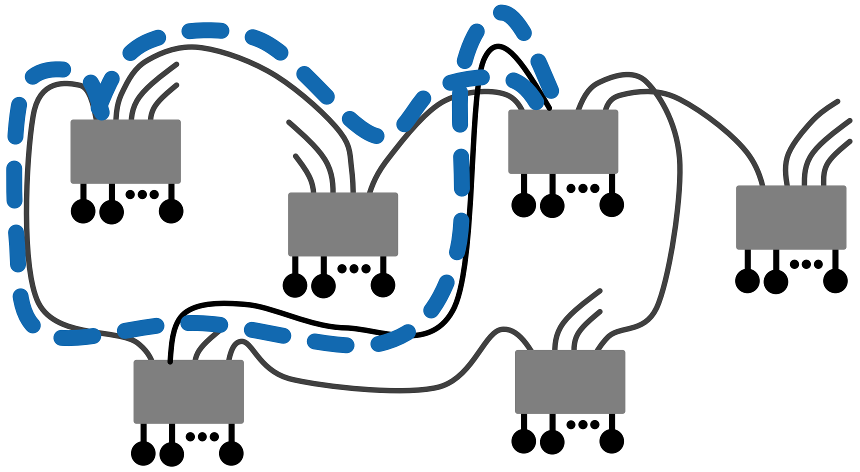
Minimal paths?



Multi Pathing?

Minimal paths?

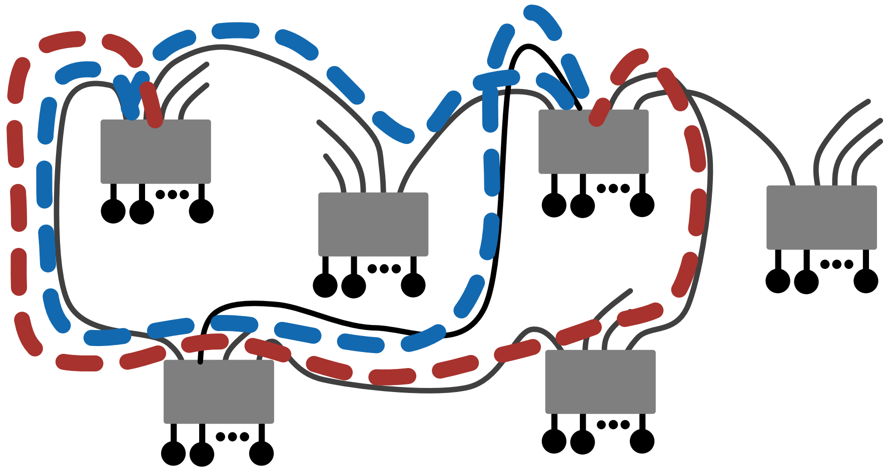
Non-minimal paths?



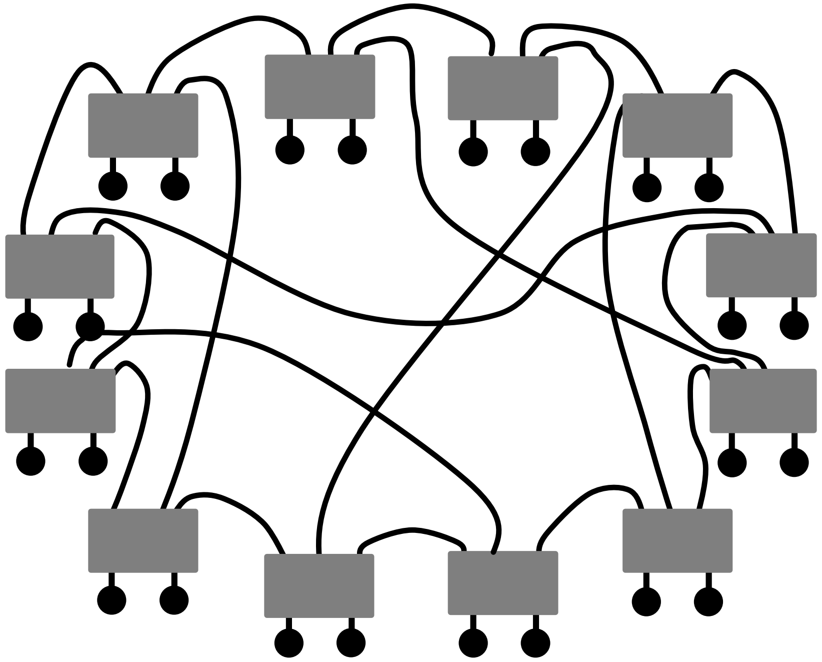
Multi Pathing?

Minimal paths?

Non-minimal paths?

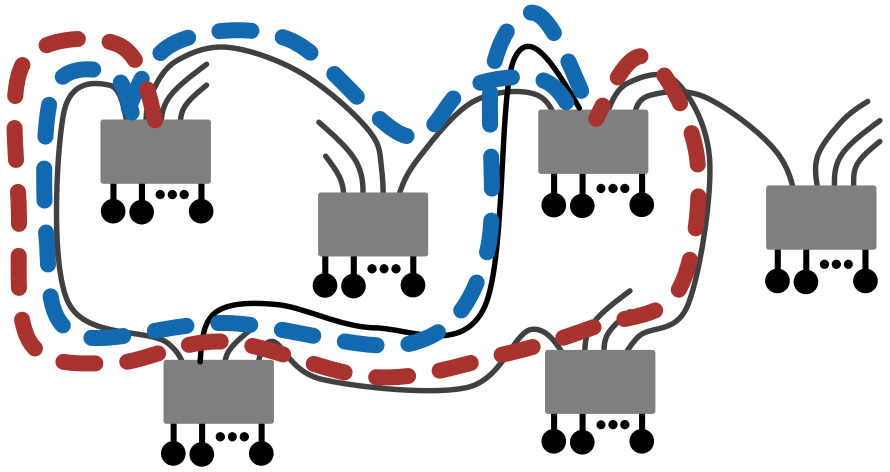


Multi Pathing?



Minimal paths?

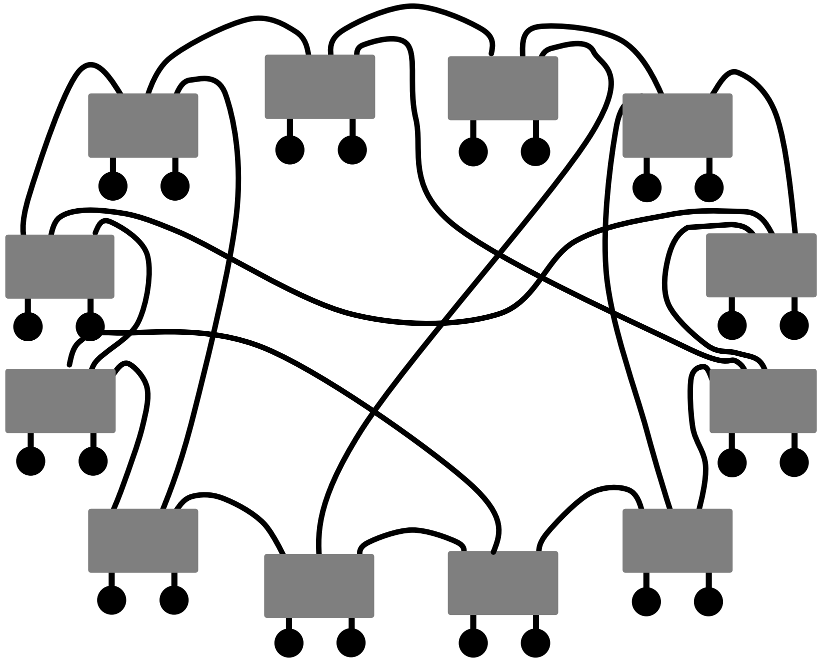
Non-minimal paths?

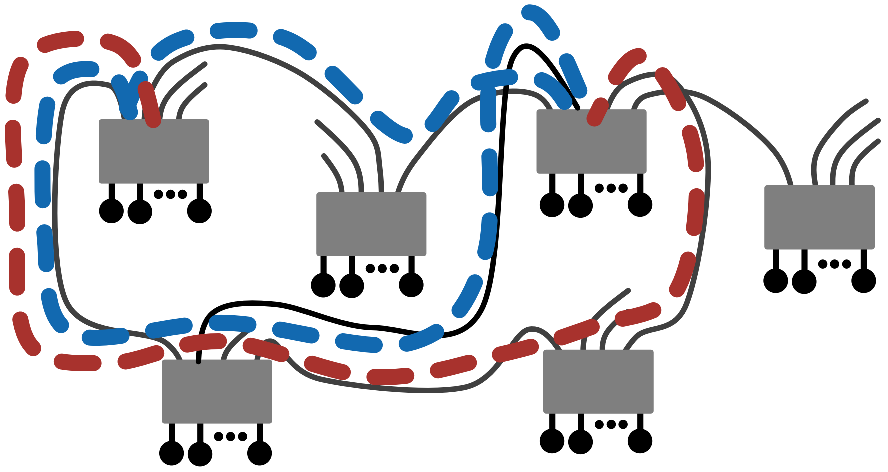


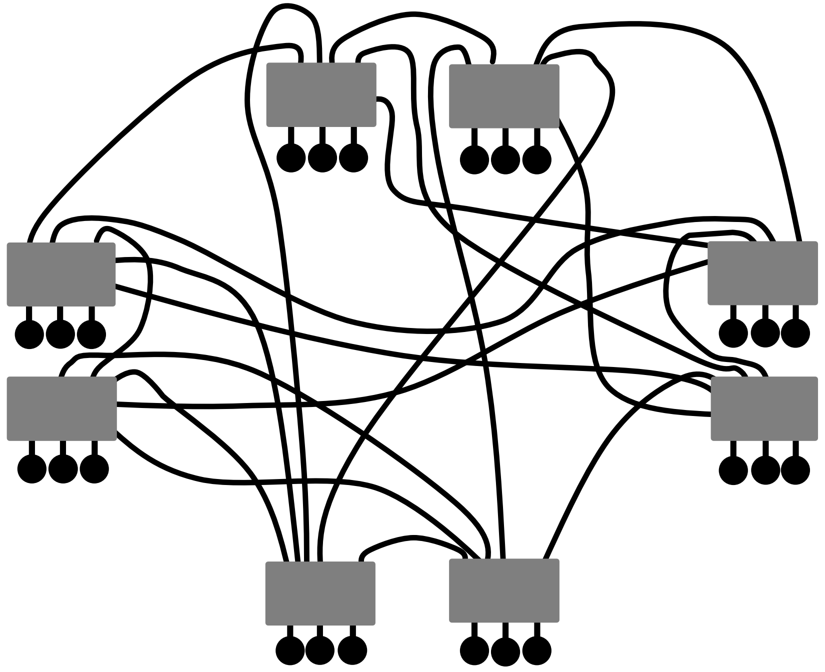
Multi Pathing?

Minimal paths?

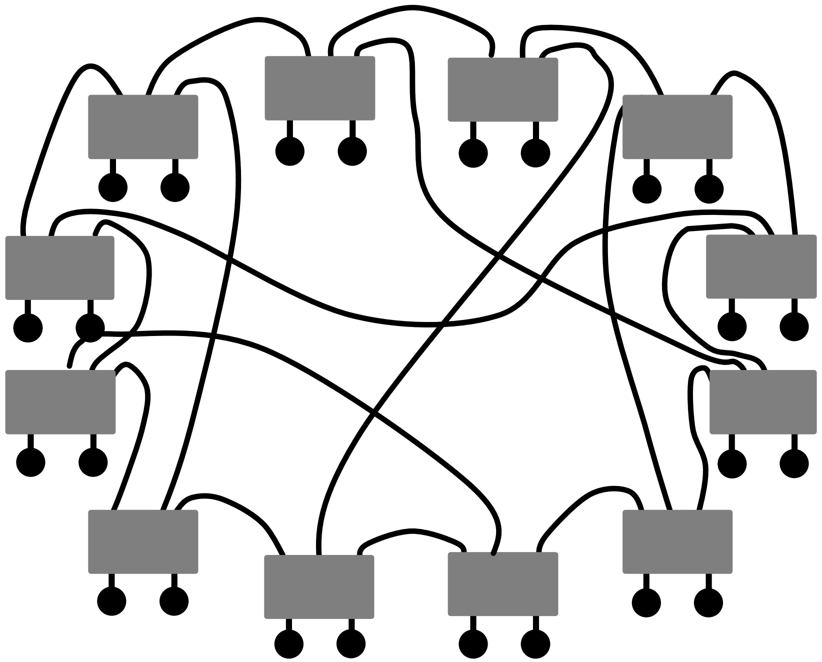
Non-minimal paths?





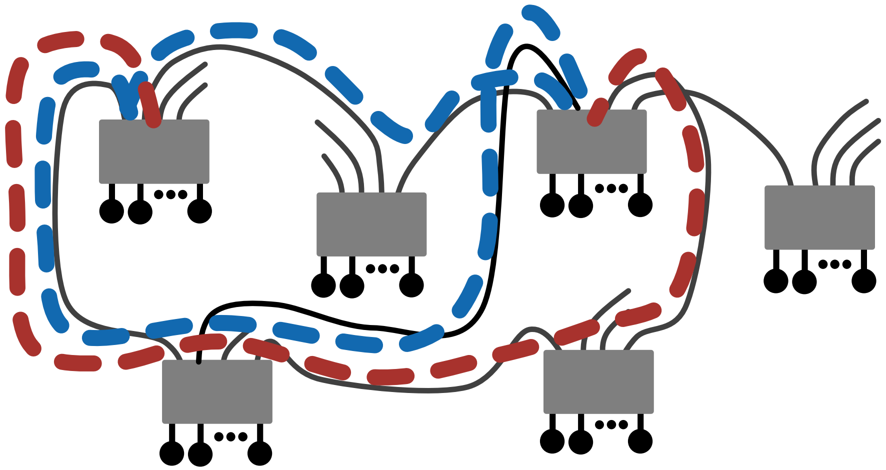


MULTI PATHING?

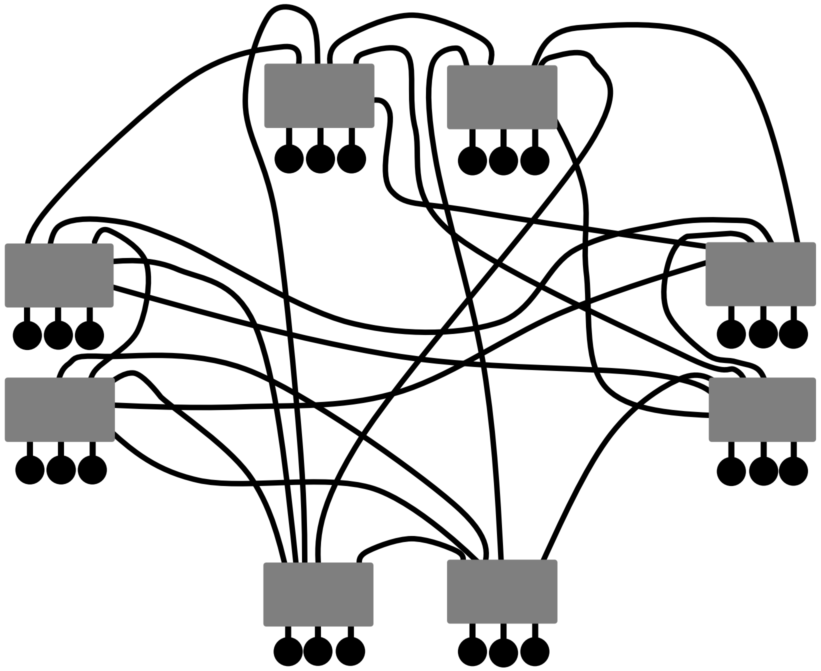


Minimal paths?

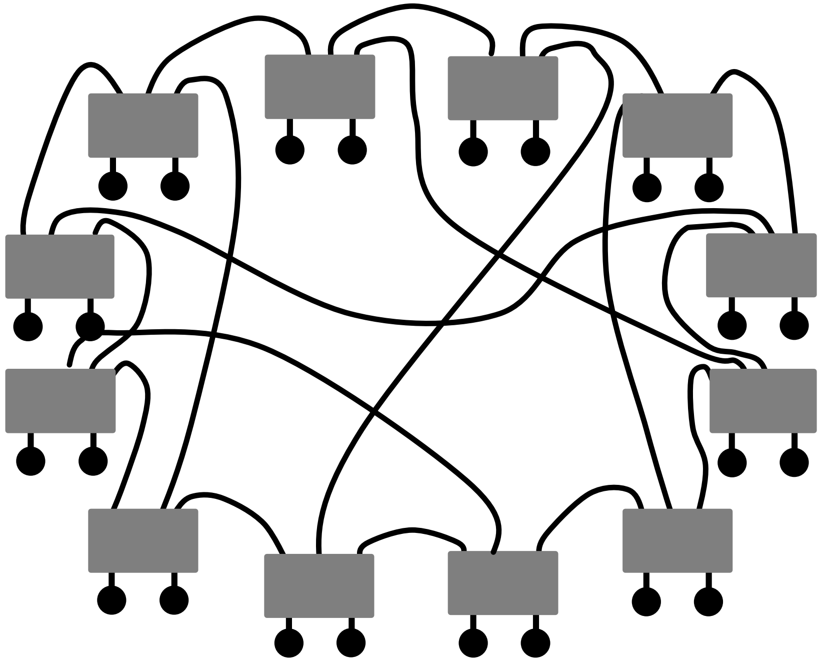
Non-minimal paths?



Expanders are sparse…

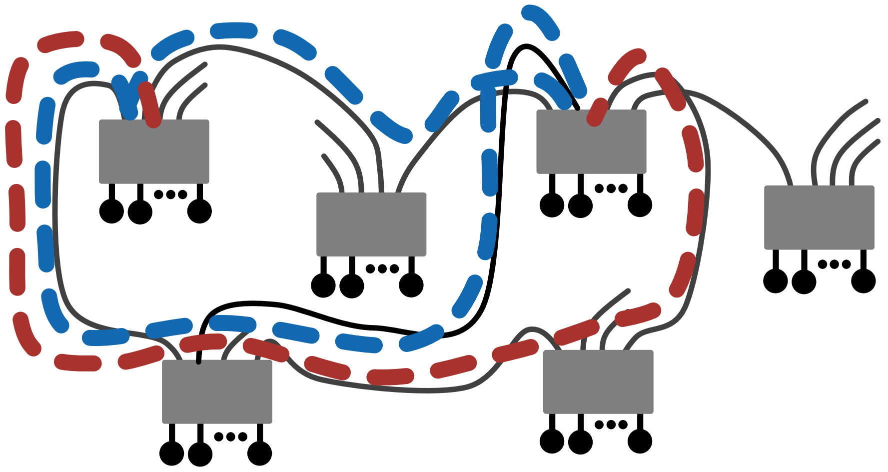


Multi Pathing?

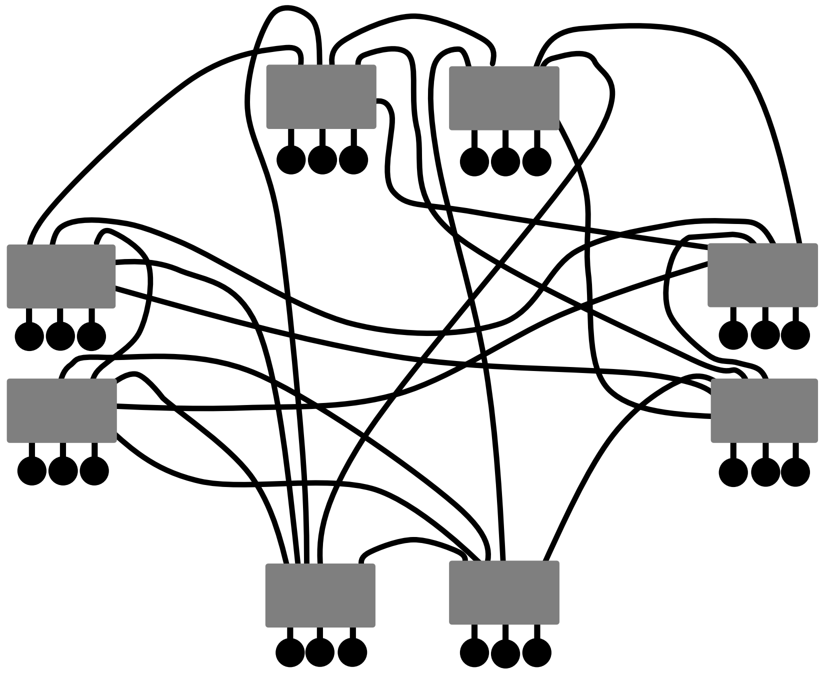


Minimal paths?

Non-minimal paths?

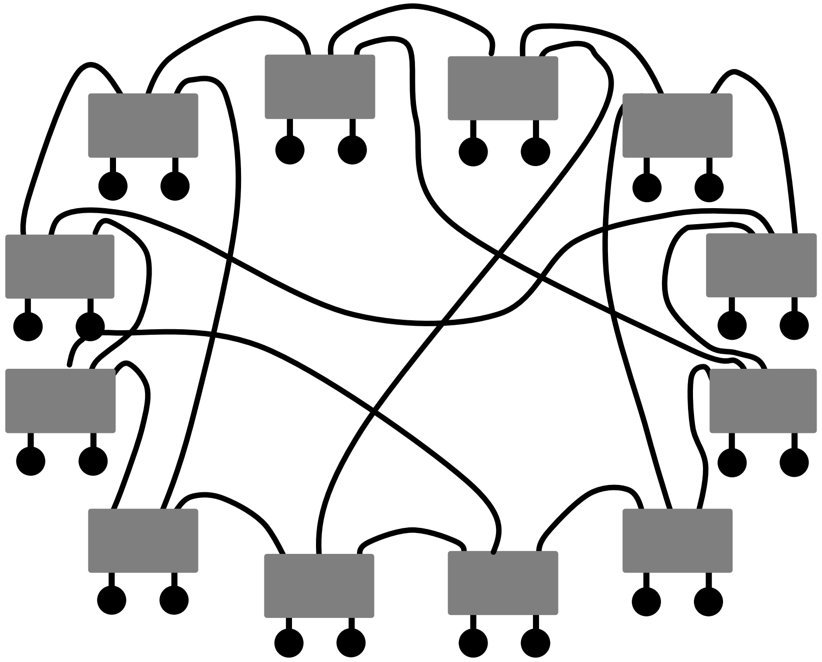


Expanders are sparse…



Is there enough path diversity?

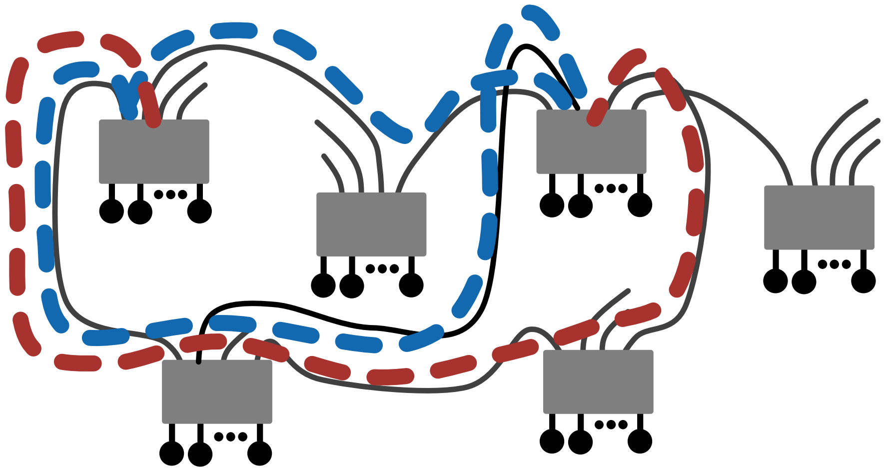
MULTI PATHING?



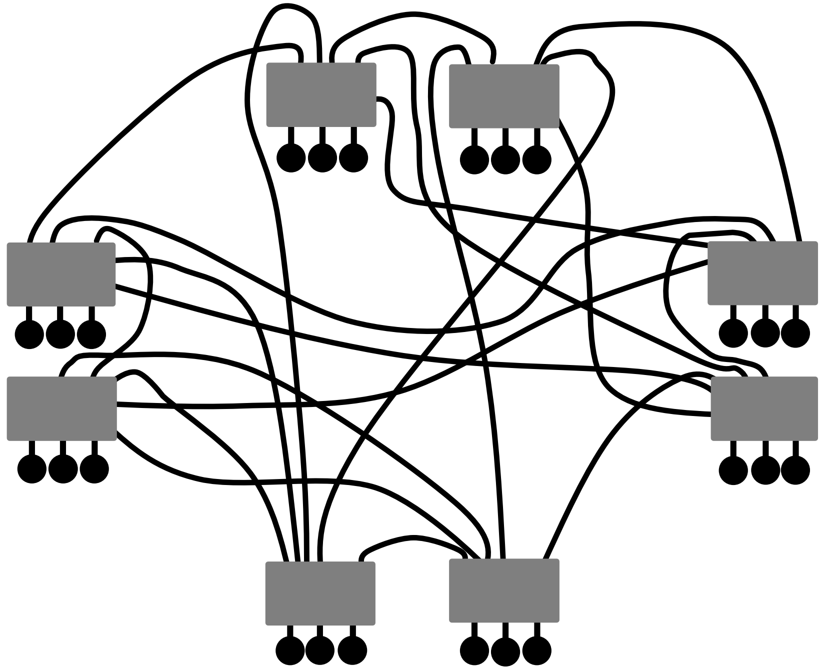
Minimal paths?

Non-minimal paths?

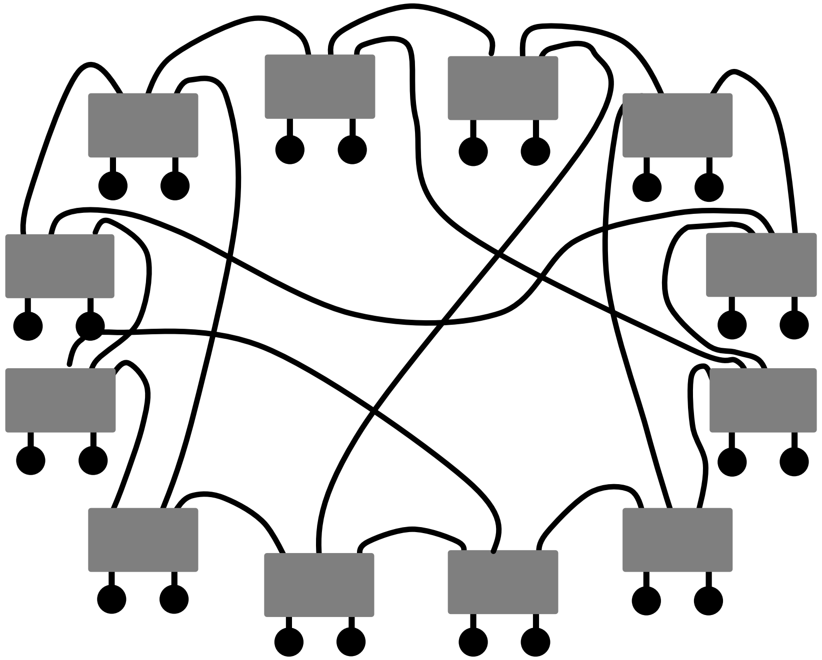
Expanders are sparse…



IS THERE ENOUGH PATH DIVERSITY?



MULTI PATHING?

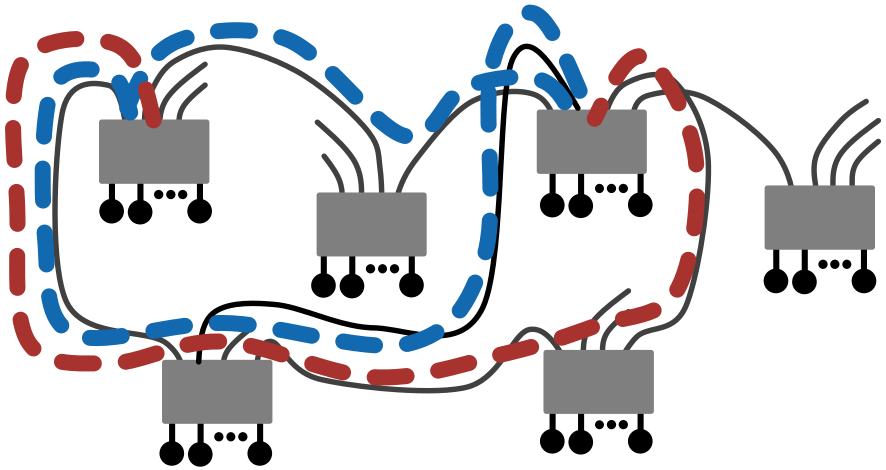


HOW TO USE MULTIPATHING?

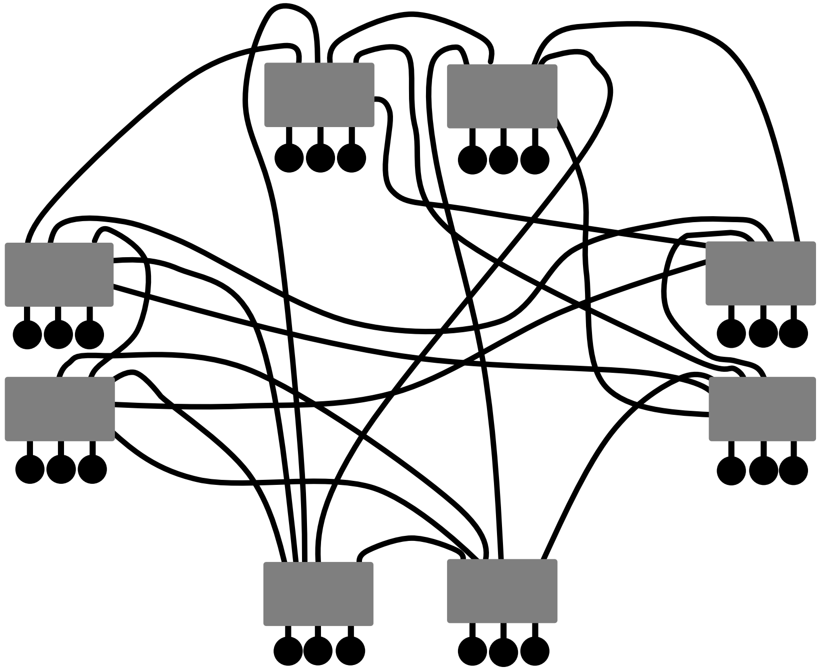
Minimal paths?

Non-minimal paths?

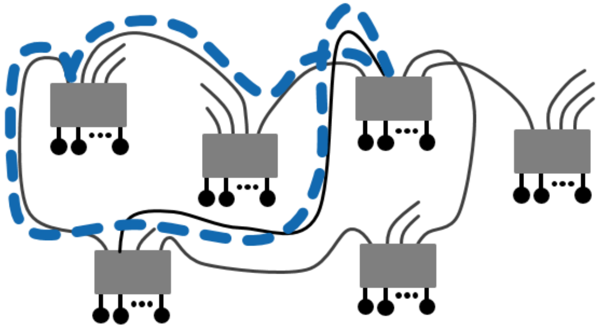
Expanders are sparse…



IS THERE ENOUGH PATH DIVERSITY?

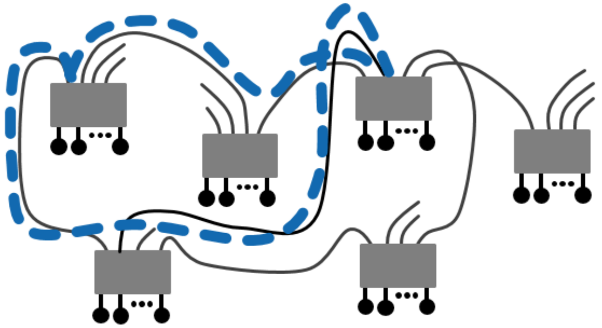


RESULTS: MINIMAL PATHS

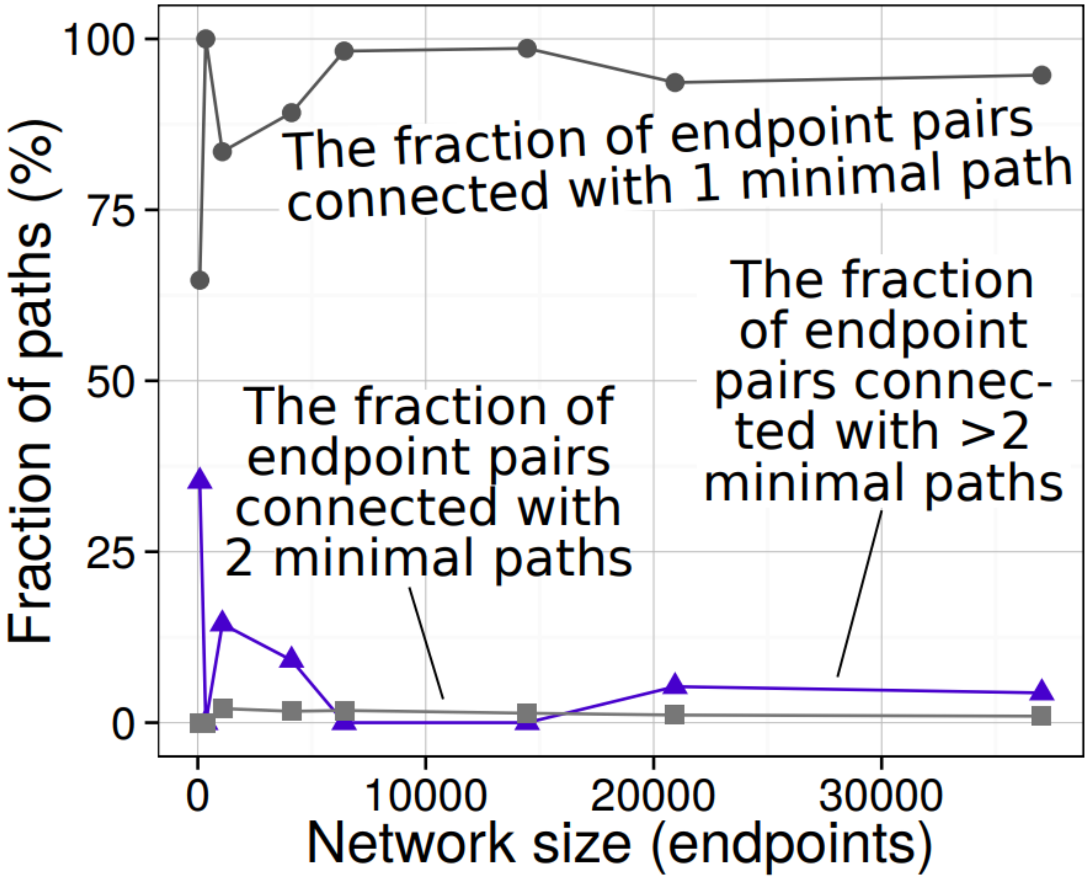


Slim Fly

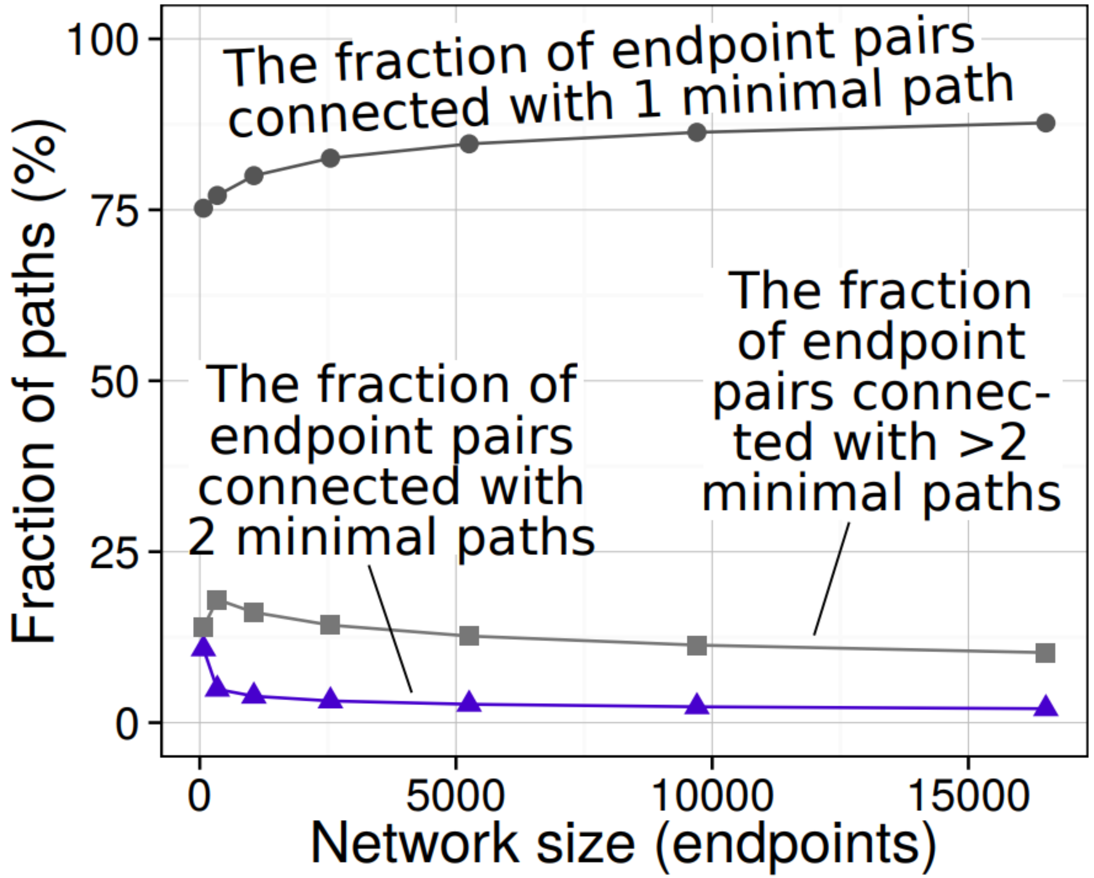
Results: Minimal Paths



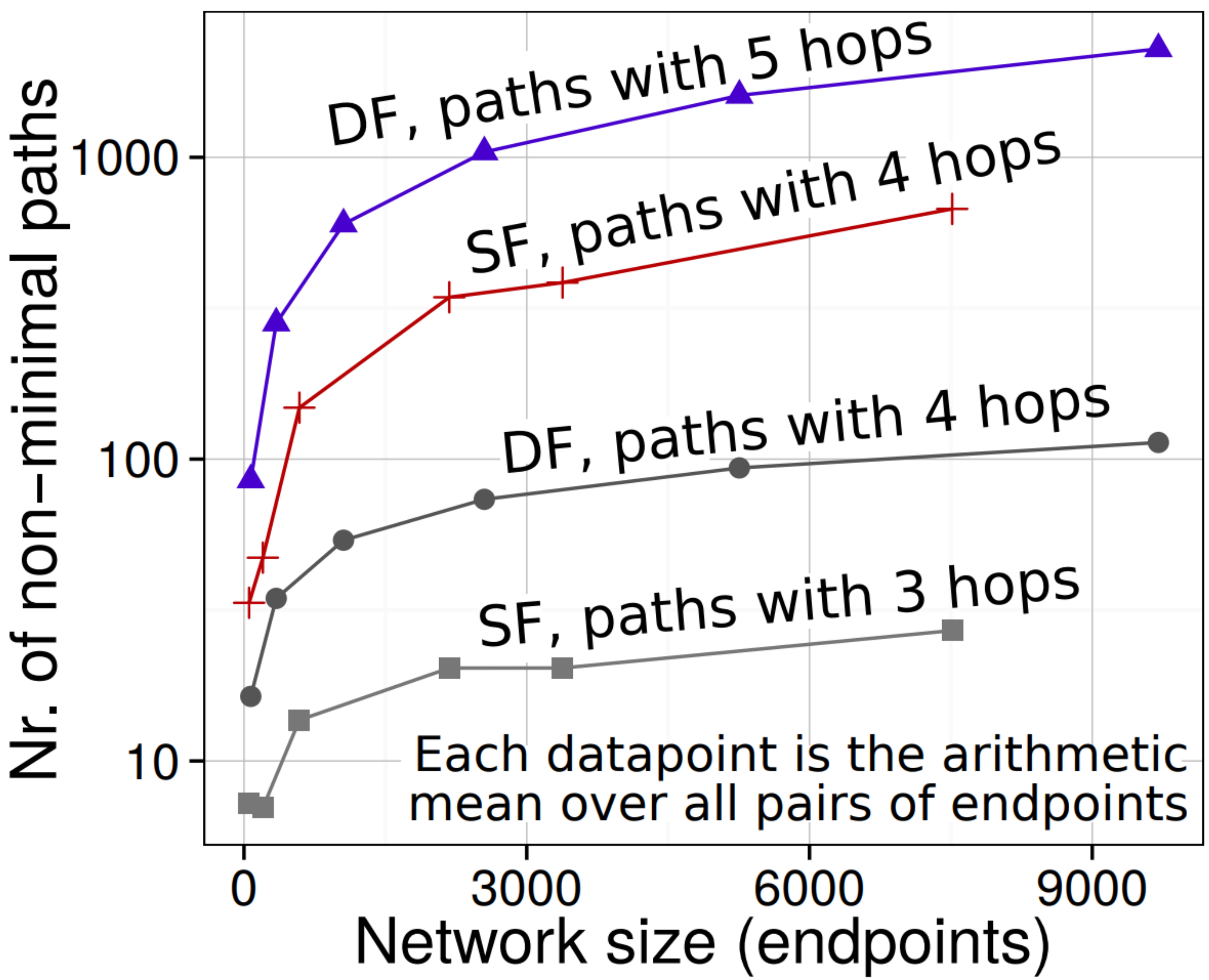
Slim Fly

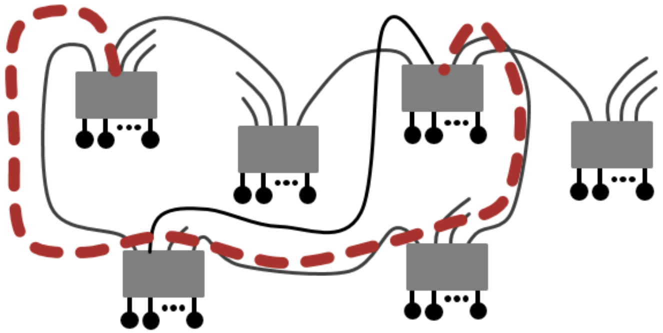


Dragonfly

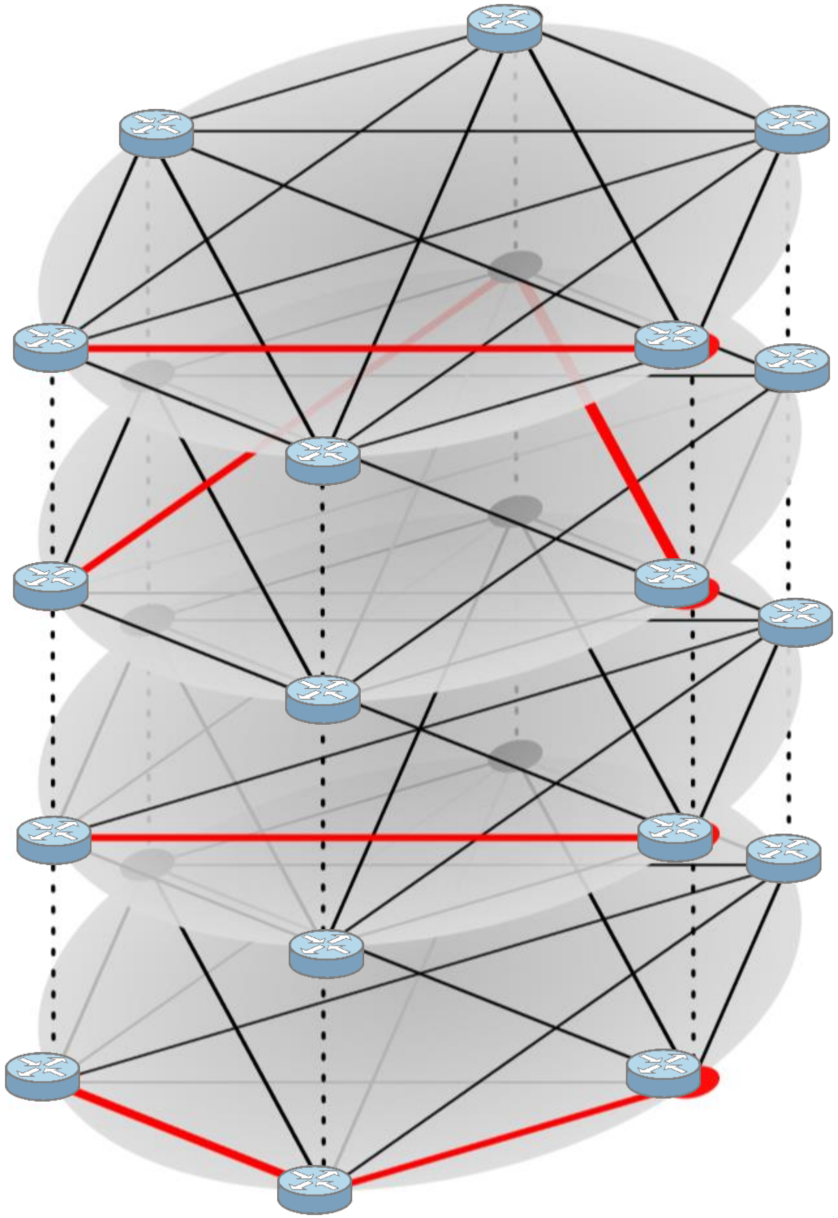


RESULTS: NON-MINIMAL PATHS



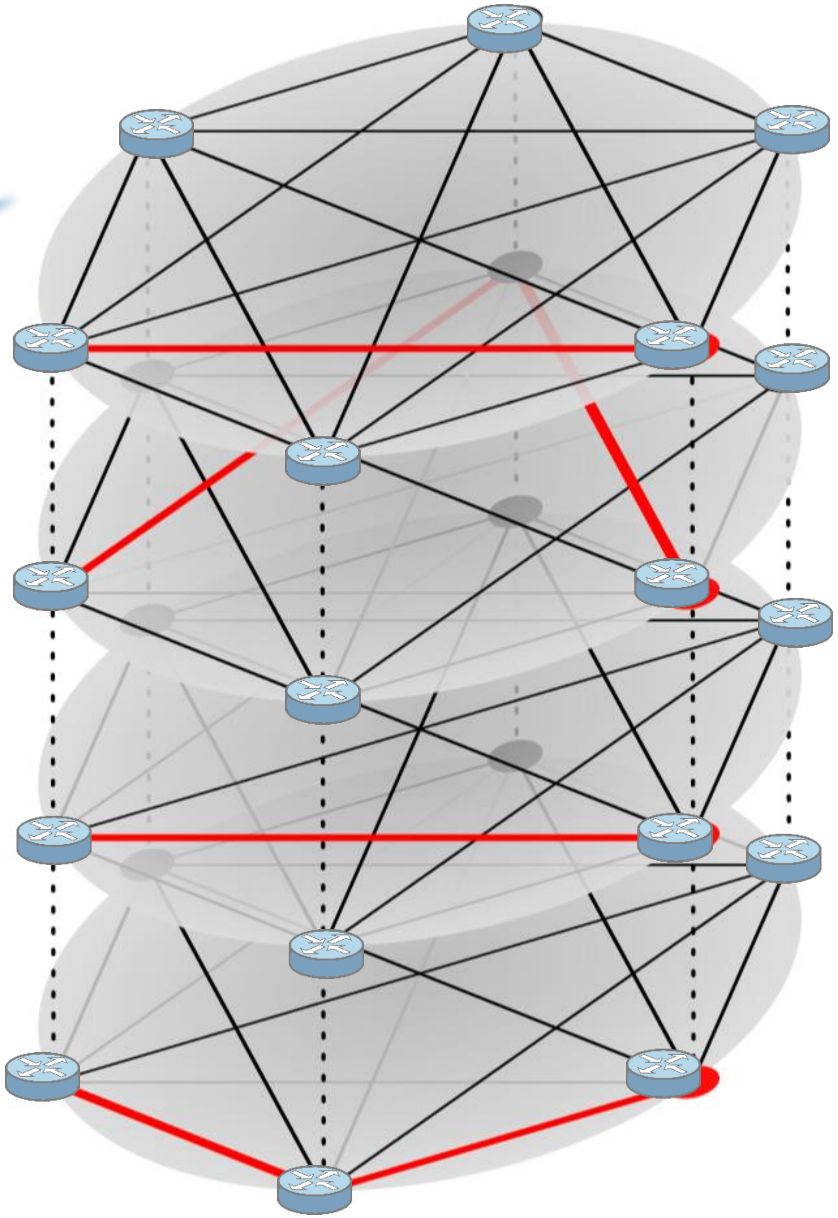


Deployment: Layers of "Almost" Minimal Paths



Deployment: Layers of "Almost" Minimal Paths

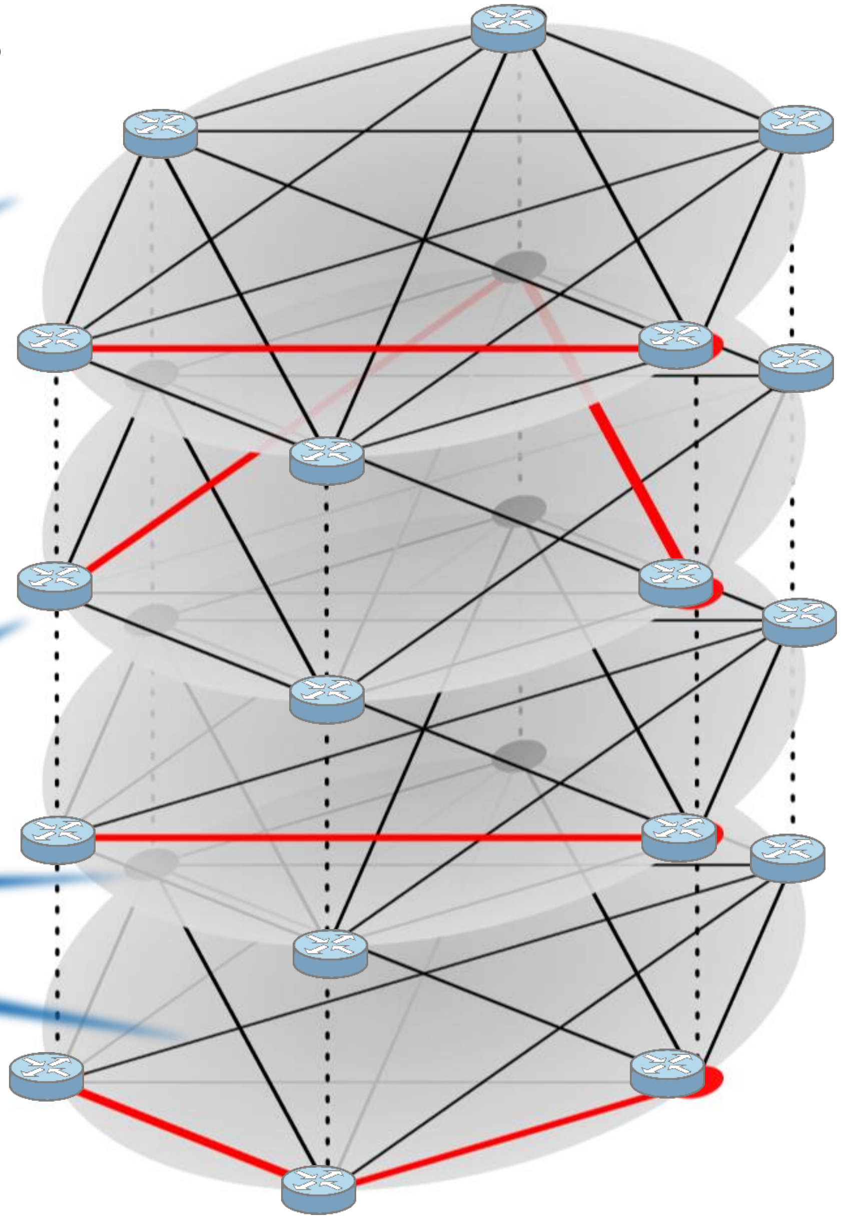
Layer 1: include all links and route minimally



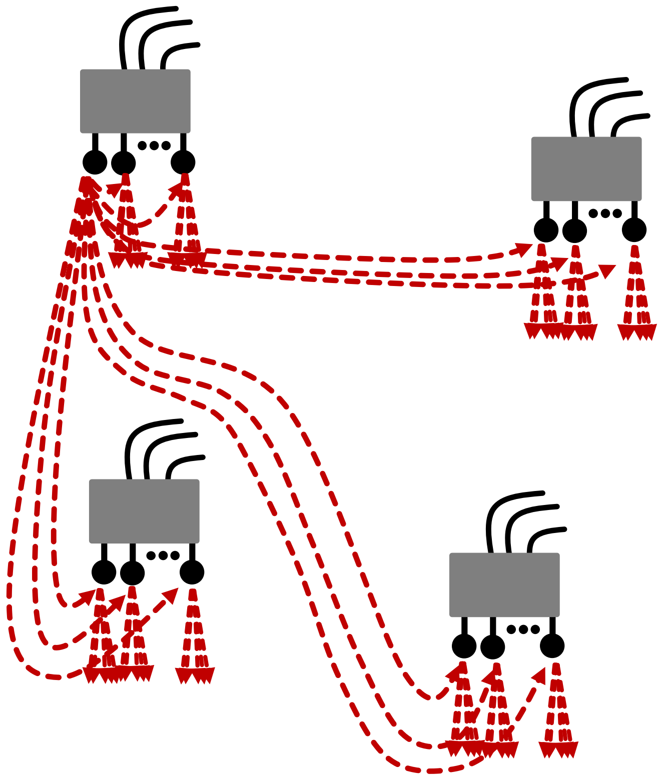
DEPLOYMENT: LAYERS OF "ALMOST" MINIMAL PATHS

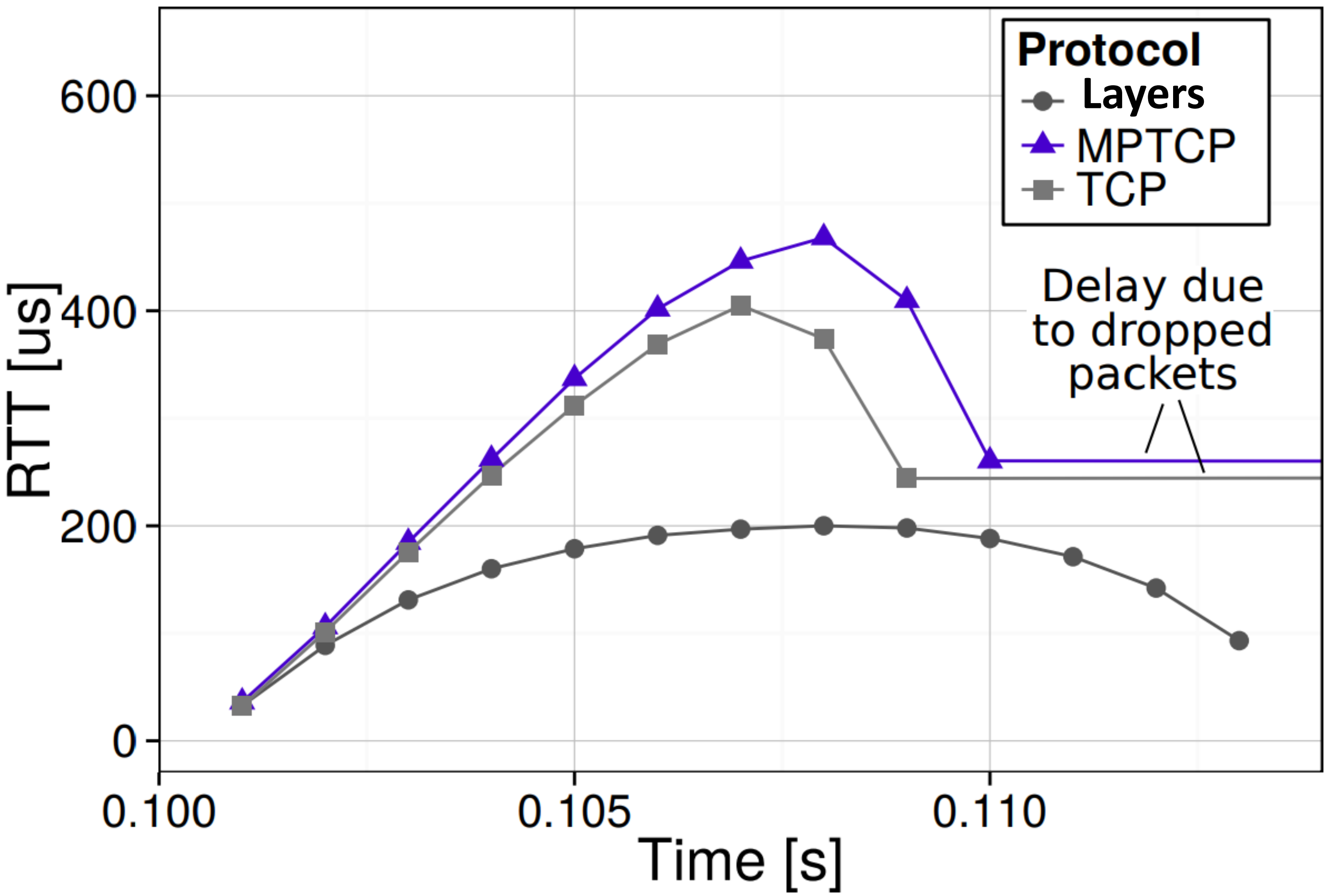
Layer 1: include all links and route minimally

Layers 2-n: remove some links (e.g., select uniformly at random) and route minimally

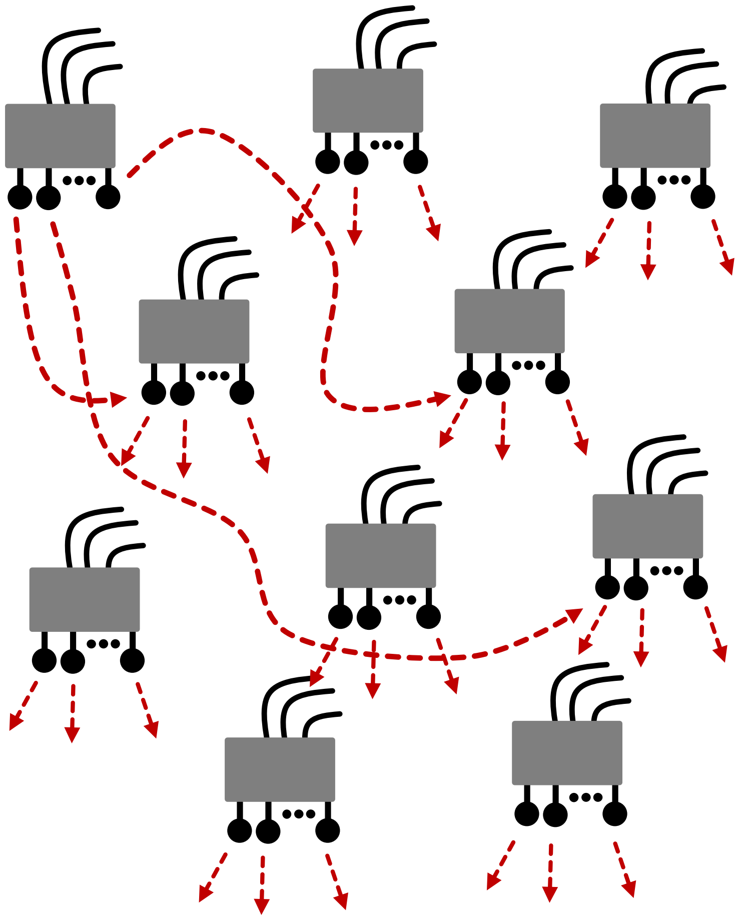


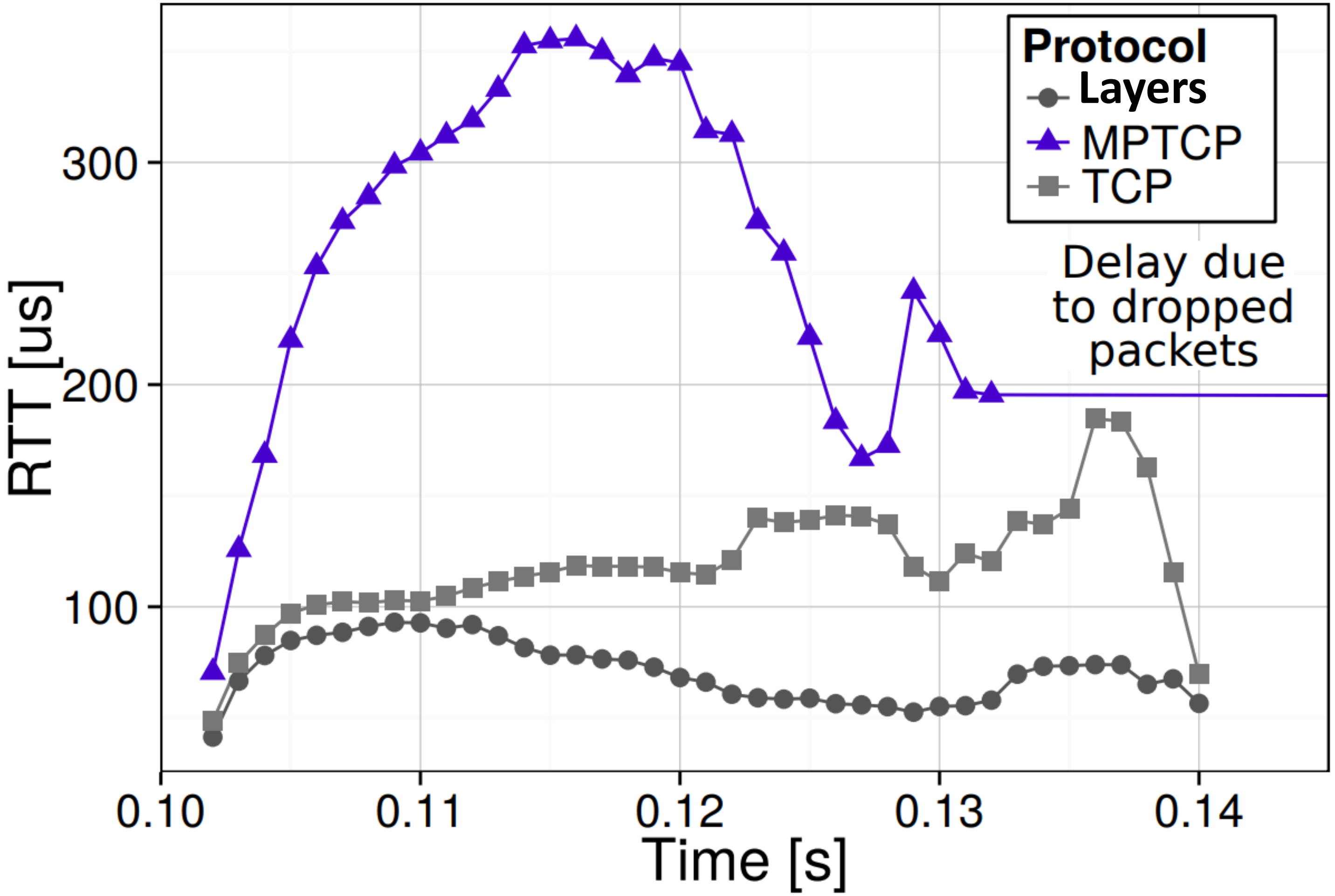
RESULTS: ALL-TO-ALL





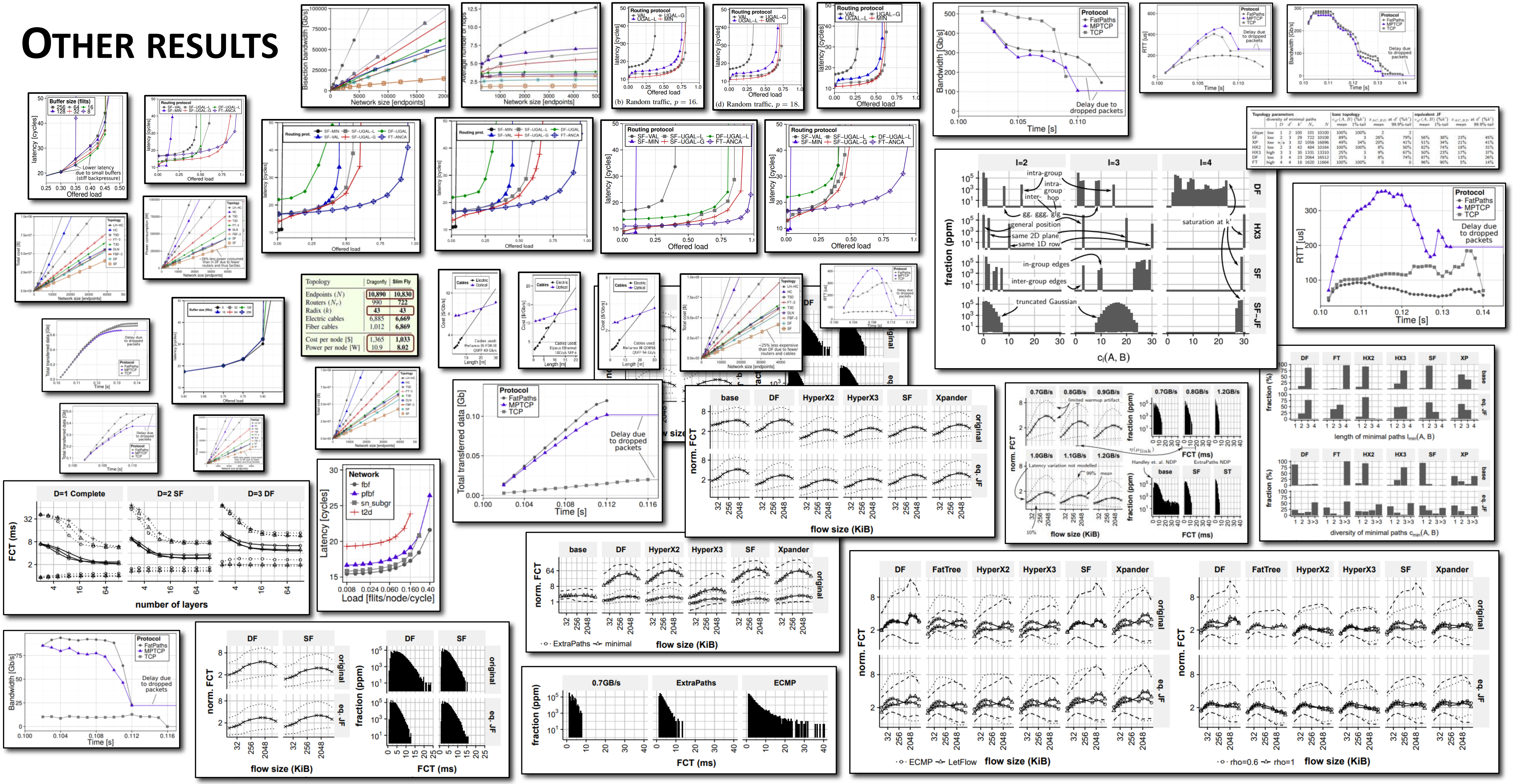
Results: Random Uniform



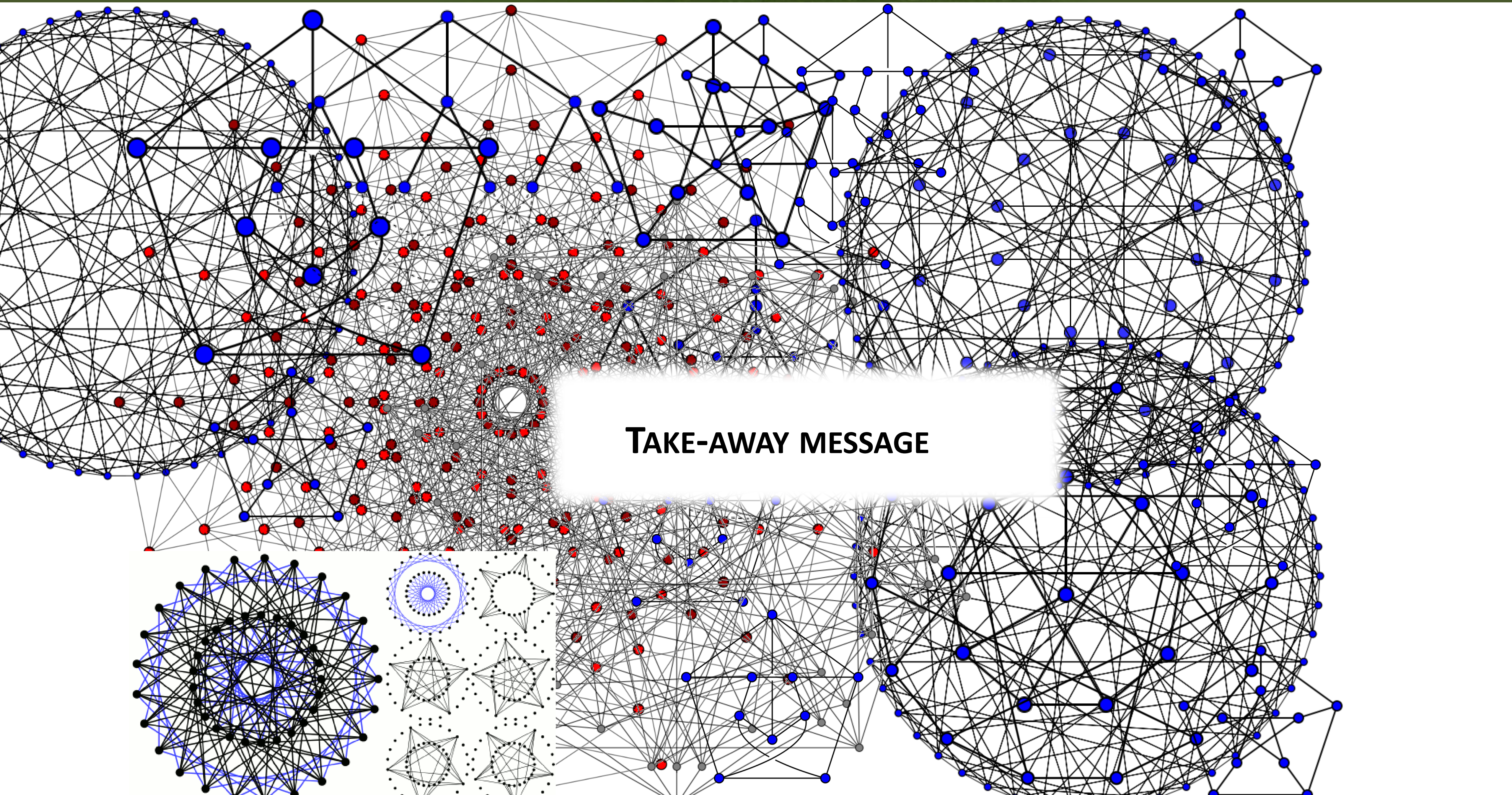


OTHER RESULTS

OTHER RESULTS

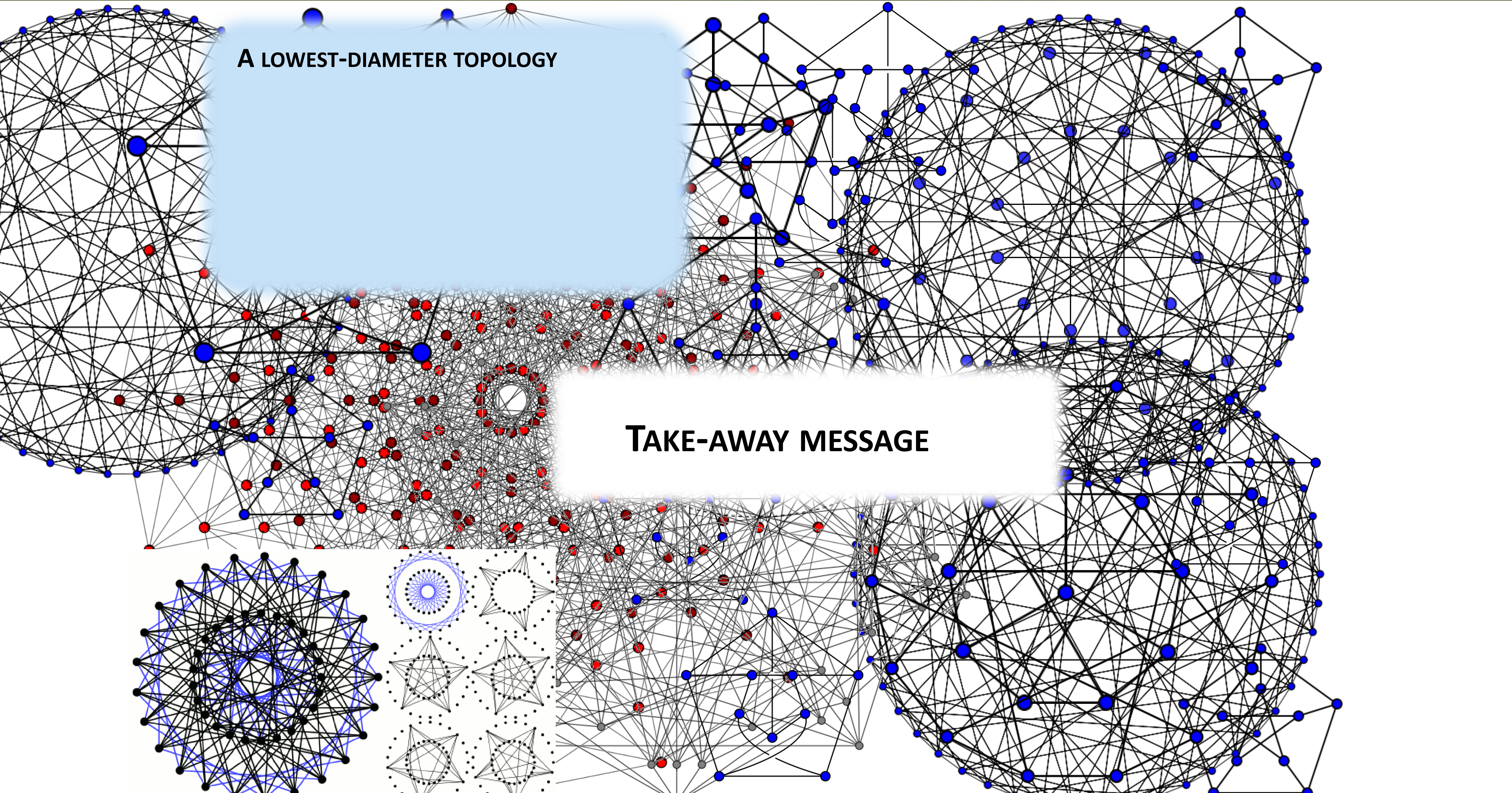


Topology	Dragonfly	Slim Fly
Endpoints (N)	10,890	10,830
Routers (N_r)	990	722
Radix (k)	43	43
Electric cables	6,885	6,669
Fiber cables	1,012	6,869
Cost per node [$]	1,365	1,033
Power per node [W]	10.9	8.02



TAKE-AWAY MESSAGE

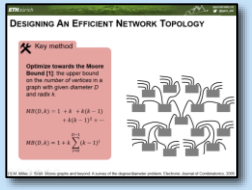
A LOWEST-DIAMETER TOPOLOGY

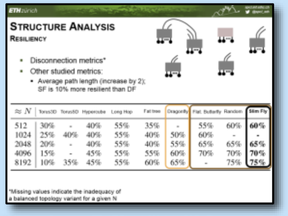


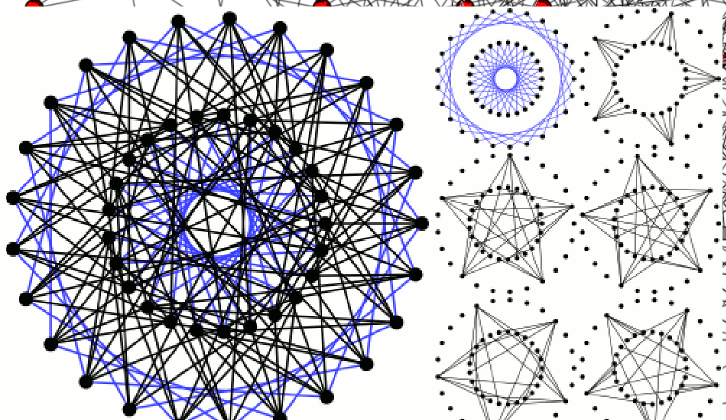
TAKE-AWAY MESSAGE

A lowest-diameter topology

- → Approaching the Moore Bound
- → Resilient



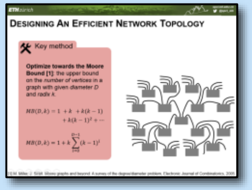


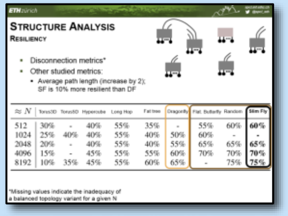


Take-away message

A lowest-diameter topology

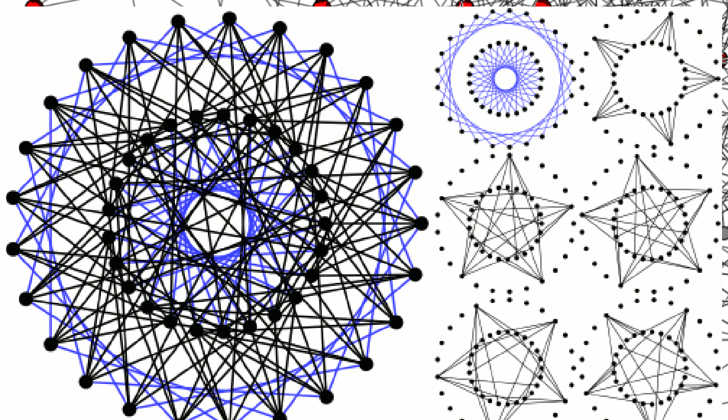
- → Approaching the Moore Bound
- → Resilient





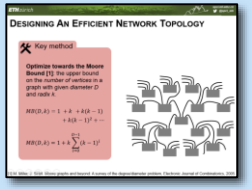
A cost & power effective topology

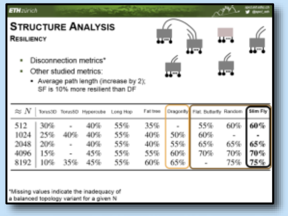
Take-away message



A lowest-diameter topology

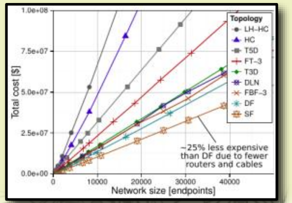
- → Approaching the Moore Bound
- → Resilient



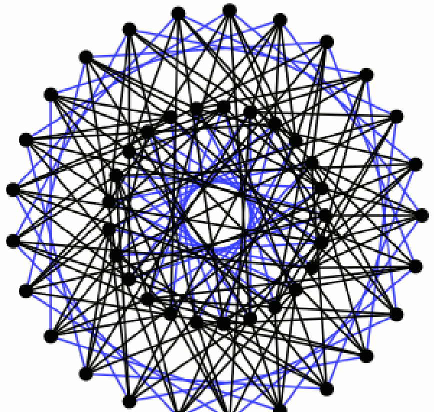


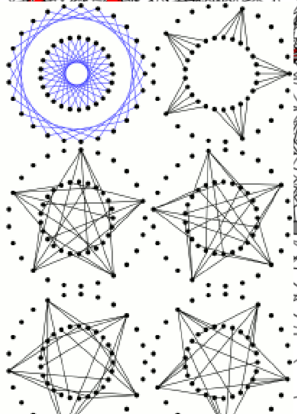
A cost & power effective topology

- → 25% less expensive than Dragonfly,
- → 26% less power-hungry than Dragonfly



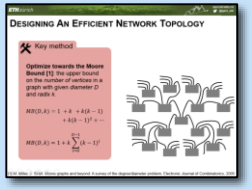
Take-away message

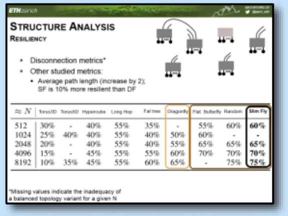




A LOWEST-DIAMETER TOPOLOGY

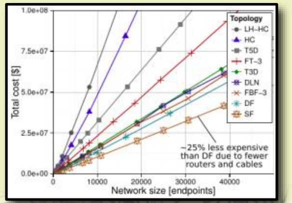
- → Approaching the Moore Bound
- → Resilient





A COST & POWER EFFECTIVE TOPOLOGY

- → 25% less expensive than Dragonfly,
- → 26% less power-hungry than Dragonfly



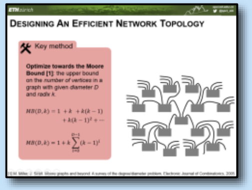
TAKE-AWAY MESSAGE

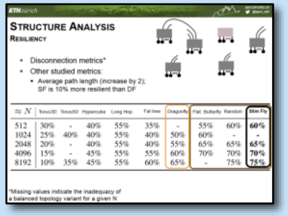
A HIGH-PERFORMANCE TOPOLOGY

A lowest-diameter topology

→ Approaching the Moore Bound

→ Resilient

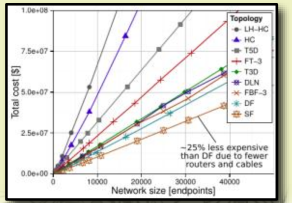




A cost & power effective topology

→ 25% less expensive than Dragonfly,

→ 26% less power-hungry than Dragonfly



Take-away message

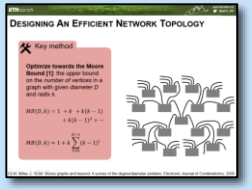
A high-performance topology

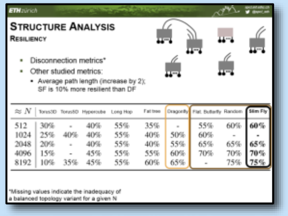
→ Lowest latency

→ Full global bandwidth

A lowest-diameter topology

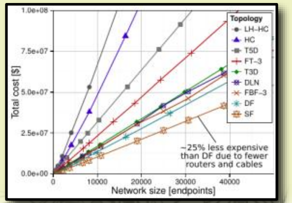
- → Approaching the Moore Bound
- → Resilient





A cost & power effective topology

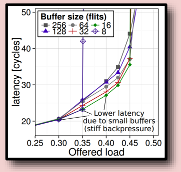
- → 25% less expensive than Dragonfly,
- → 26% less power-hungry than Dragonfly

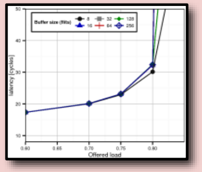


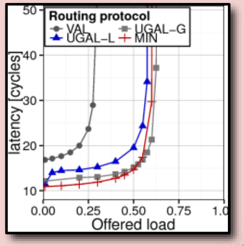
Take-away message

A high-performance topology

- → Lowest latency
- → Full global bandwidth





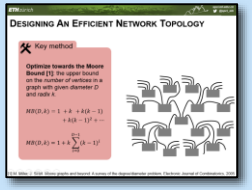


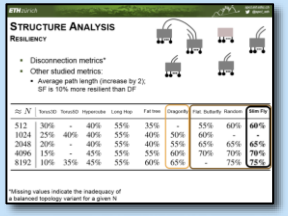
TAKE-AWAY MESSAGE

A LOWEST-DIAMETER TOPOLOGY

→ Approaching the Moore Bound

→ Resilient

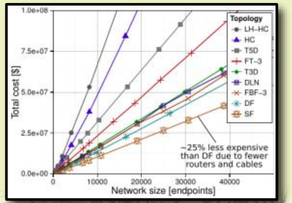




A COST & POWER EFFECTIVE TOPOLOGY

→ 25% less expensive than Dragonfly,

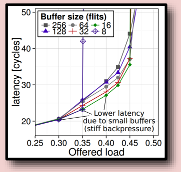
→ 26% less power-hungry than Dragonfly

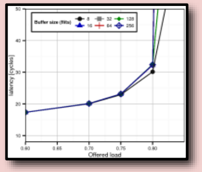


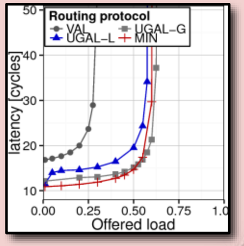
A HIGH-PERFORMANCE TOPOLOGY

→ Lowest latency

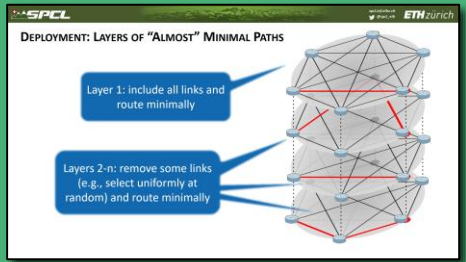
→ Full global bandwidth





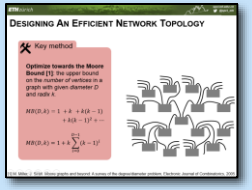


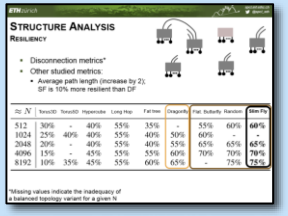
POTENTIAL FOR NOVEL SOLUTIONS



A LOWEST-DIAMETER TOPOLOGY

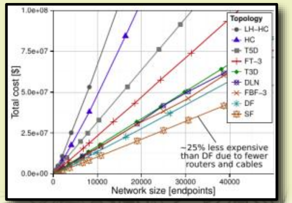
- → Approaching the Moore Bound
- → Resilient





A COST & POWER EFFECTIVE TOPOLOGY

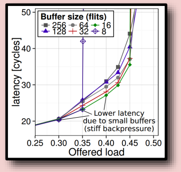
- → 25% less expensive than Dragonfly,
- → 26% less power-hungry than Dragonfly

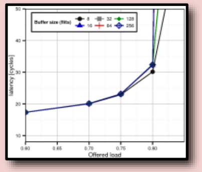


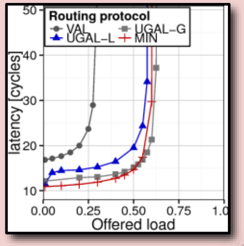
TAKE-AWAY MESSAGE

A HIGH-PERFORMANCE TOPOLOGY

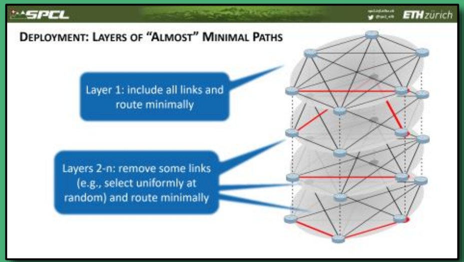
- → Lowest latency
- → Full global bandwidth







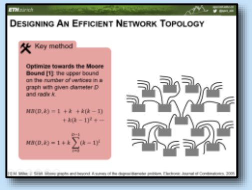
POTENTIAL FOR NOVEL SOLUTIONS

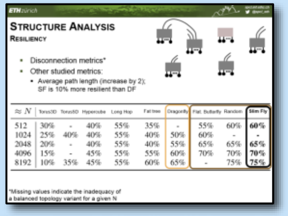


A LOWEST-DIAMETER TOPOLOGY

→ Approaching the Moore Bound

→ Resilient

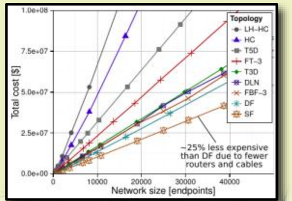




A COST & POWER EFFECTIVE TOPOLOGY

→ 25% less expensive than Dragonfly,

→ 26% less power-hungry than Dragonfly

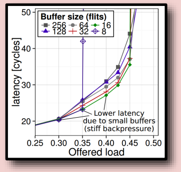


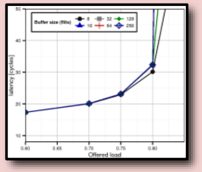
TAKE-AWAY MESSAGE

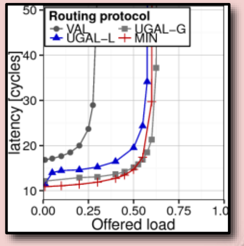
A HIGH-PERFORMANCE TOPOLOGY

→ Lowest latency

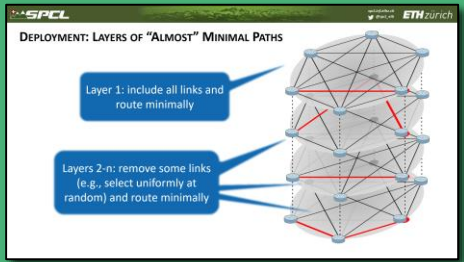
→ Full global bandwidth







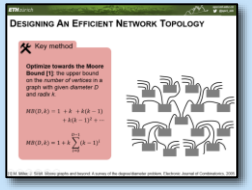
POTENTIAL FOR NOVEL SOLUTIONS

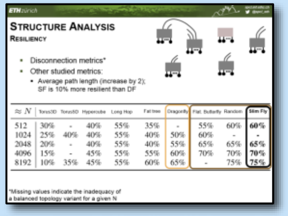


A lowest-diameter topology

→ Approaching the Moore Bound

→ Resilient

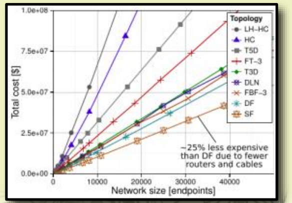




A cost & power effective topology

→ 25% less expensive than Dragonfly,

→ 26% less power-hungry than Dragonfly

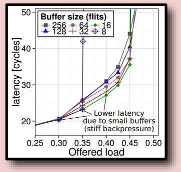


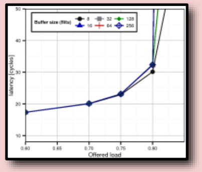
Thank you for your attention

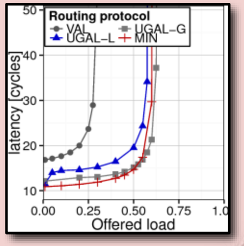
A high-performance topology

→ Lowest latency

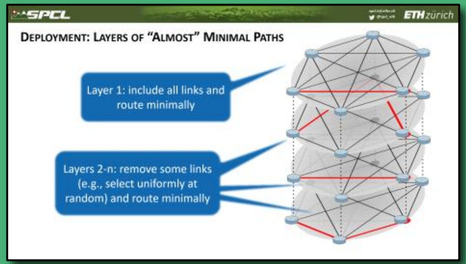
→ Full global bandwidth





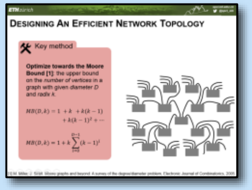


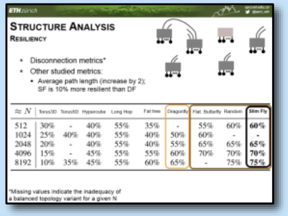
Potential for novel solutions



A LOWEST-DIAMETER TOPOLOGY

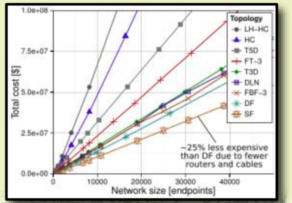
- → Approaching the Moore Bound
- → Resilient



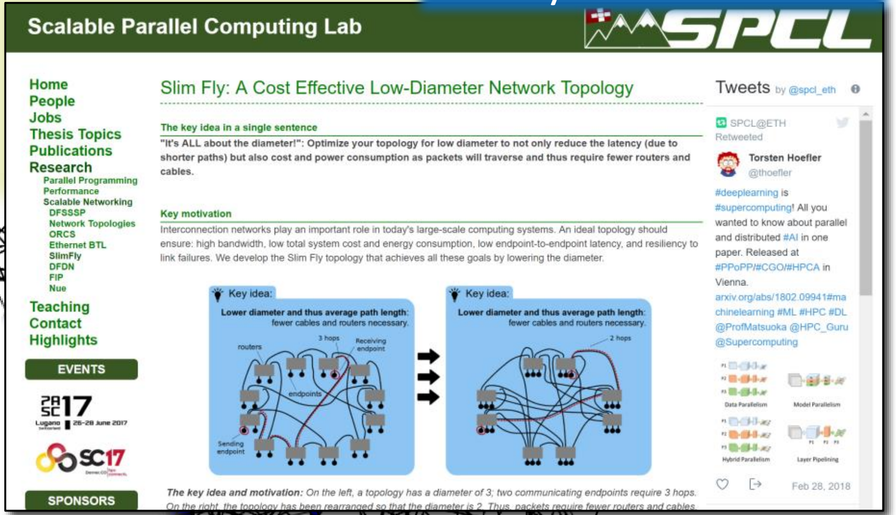


A COST & POWER EFFECTIVE TOPOLOGY

- → 25% less expensive than Dragonfly,
- → 26% less power-hungry than Dragonfly



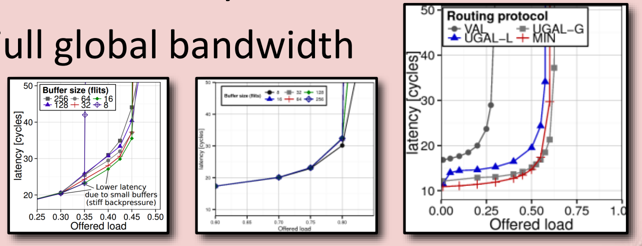
http://spcl.inf.ethz.ch/Research/Scalable_Networking/SlimFly



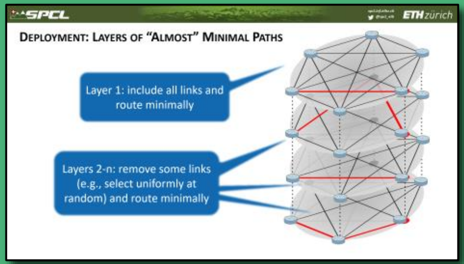
Thank you for your attention

A HIGH-PERFORMANCE TOPOLOGY

- → Lowest latency
- → Full global bandwidth



POTENTIAL FOR NOVEL SOLUTIONS



Lowering Diameter Enables Cost-Effective and High-Performance Networks

MACIEJ BESTA, ERIK HENRIKSSON, TORSTEN HOEFLER

DIAMETER-2 SLIM FLY

1 *Select a prime power q*

Diameter-2 Slim Fly

1 *Select a prime power q*

$$q = 4w + \delta;$$

$$w \in \mathbb{N} \quad \delta \in \{-1,0,1\},$$

DIAMETER-2 SLIM FLY

1 *Select a prime power q*

$$q = 4w + \delta;$$

$w \in \mathbb{N} \quad \delta \in \{-1, 0, 1\},$

A Slim Fly based on q :

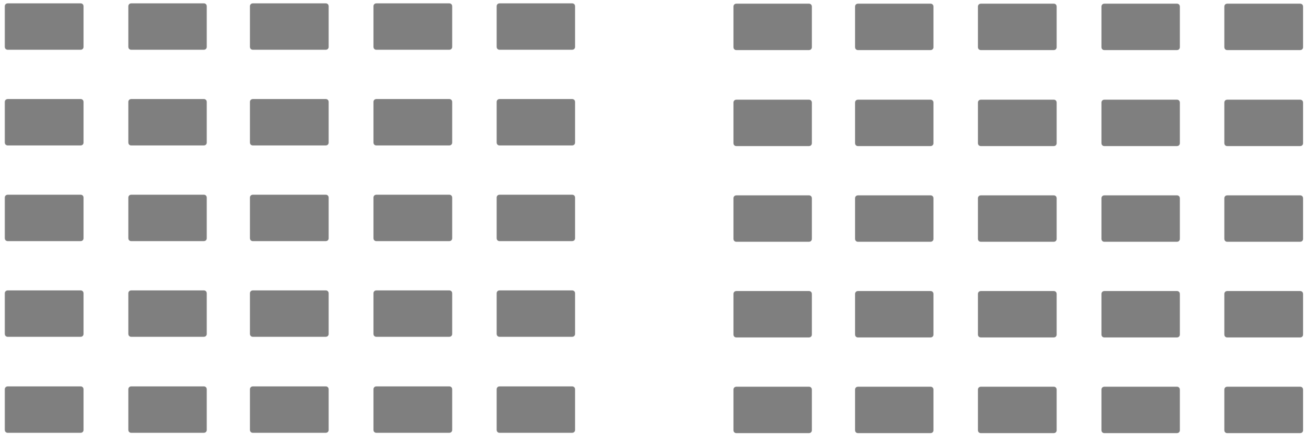
Diameter-2 Slim Fly

1 *Select a prime power q*

$$q = 4w + \delta;$$

$$w \in \mathbb{N} \quad \delta \in \{-1,0,1\},$$

A Slim Fly based on q :



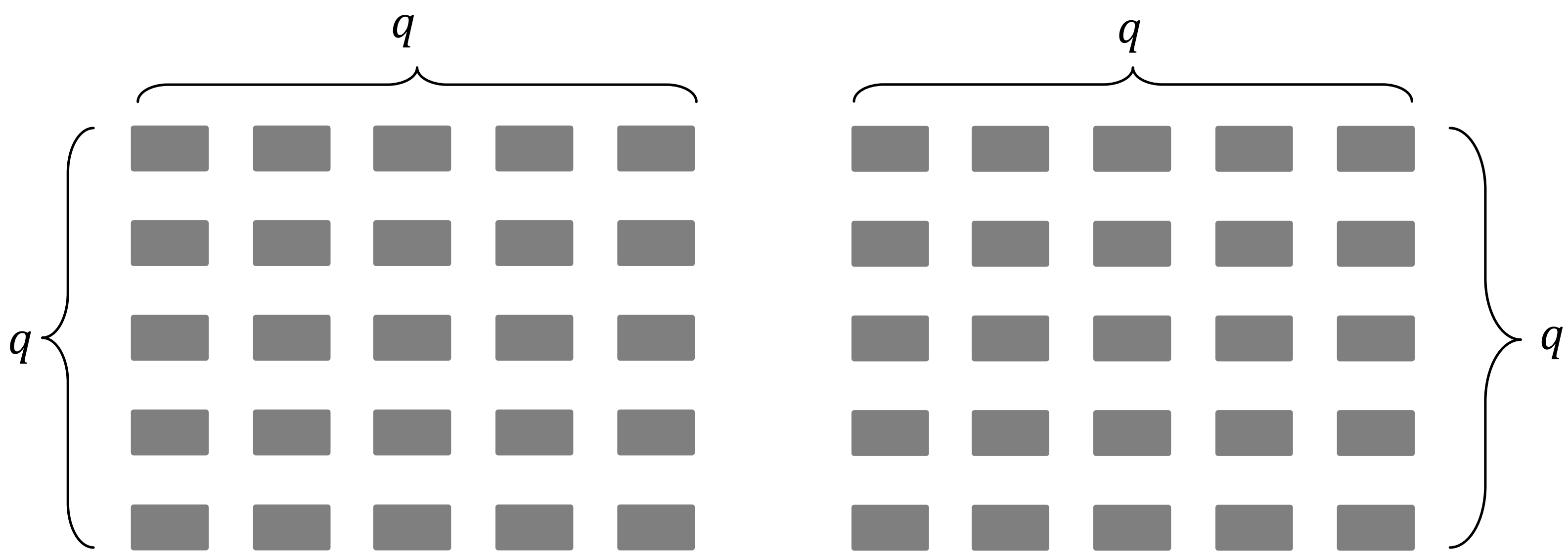
Diameter-2 Slim Fly

1 *Select a prime power q*

$$q = 4w + \delta;$$

$w \in \mathbb{N} \quad \delta \in \{-1,0,1\},$

A Slim Fly based on q :



Diameter-2 Slim Fly

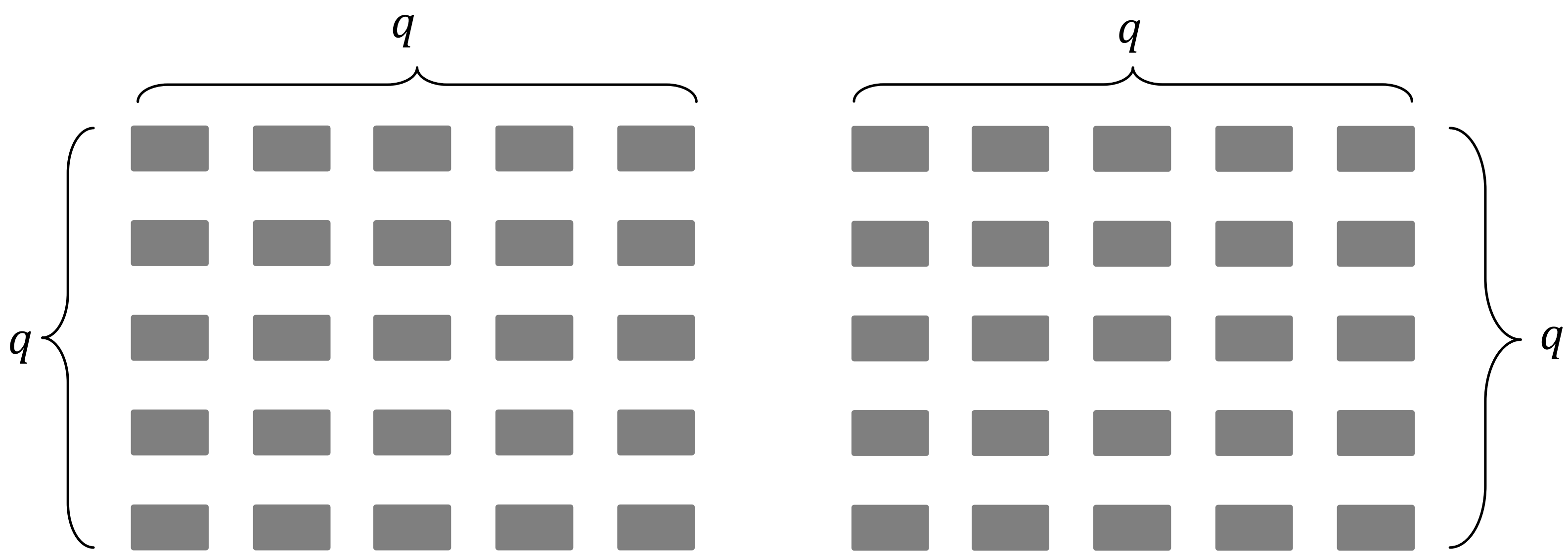
1 *Select a prime power q*

$$q = 4w + \delta;$$

$w \in \mathbb{N} \quad \delta \in \{-1, 0, 1\},$

A Slim Fly based on q:

Number of routers: $2q^2$



DIAMETER-2 SLIM FLY

1 *Select a prime power q*

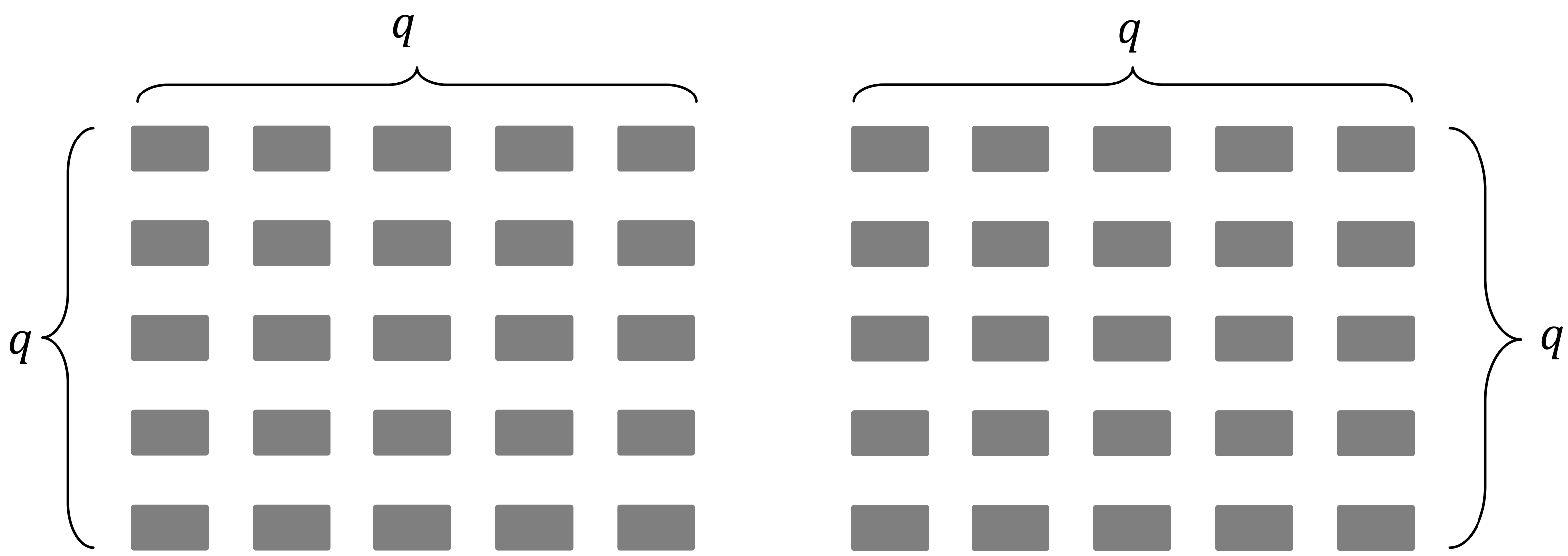
$$q = 4w + \delta;$$

$w \in \mathbb{N} \quad \delta \in \{-1,0,1\},$

A Slim Fly based on q:

Number of routers: $2q^2$

Network radix: $(3q - \delta)/2$



Diameter-2 Slim Fly

1 *Select a prime power q*

$$q = 4w + \delta;$$

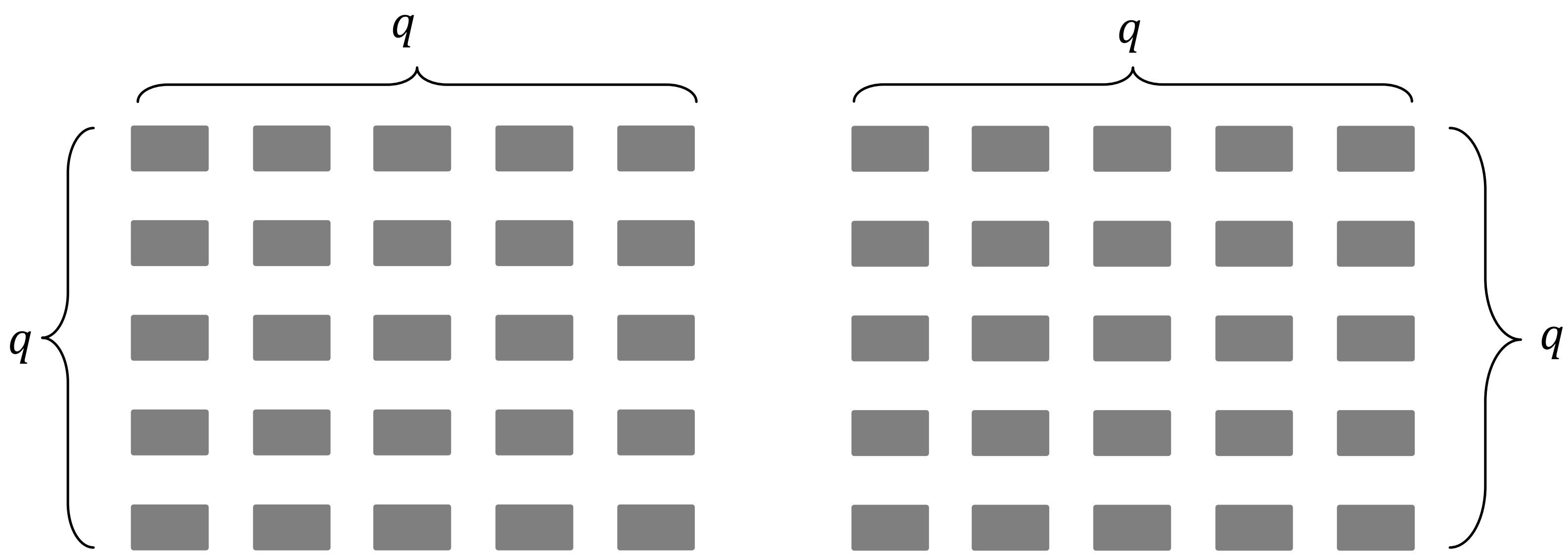
$$w \in \mathbb{N} \quad \delta \in \{-1,0,1\},$$

A Slim Fly based on q:

Number of routers: $2q^2$

Network radix: $(3q - \delta)/2$

2 *Construct a finite field $\mathcal{F}_q$.*



Diameter-2 Slim Fly

1 *Select a prime power q*

$$q = 4w + \delta;$$

$w \in \mathbb{N} \quad \delta \in \{-1,0,1\},$

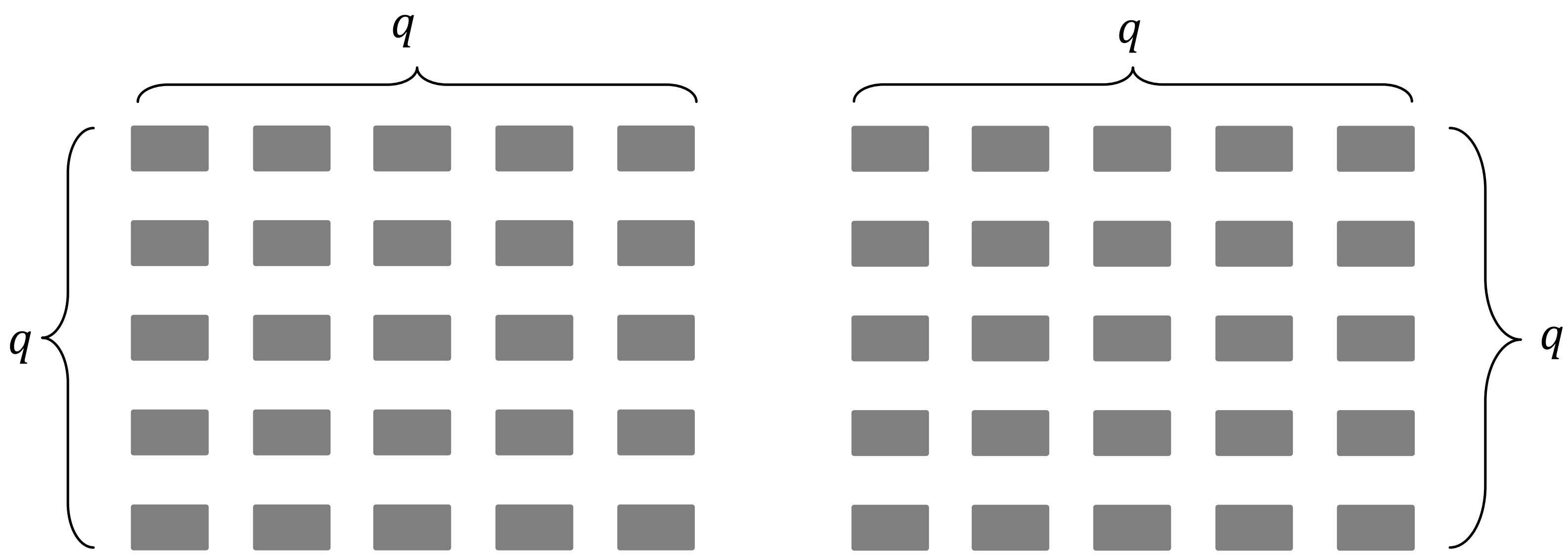
A Slim Fly based on q:

Number of routers: $2q^2$

Network radix: $(3q - \delta)/2$

2 *Construct a finite field $\mathcal{F}_q$.*

Assuming q is prime:



DIAMETER-2 SLIM FLY

1 *Select a prime power q*

$$q = 4w + \delta;$$

$w \in \mathbb{N} \quad \delta \in \{-1,0,1\}$,

A Slim Fly based on q:

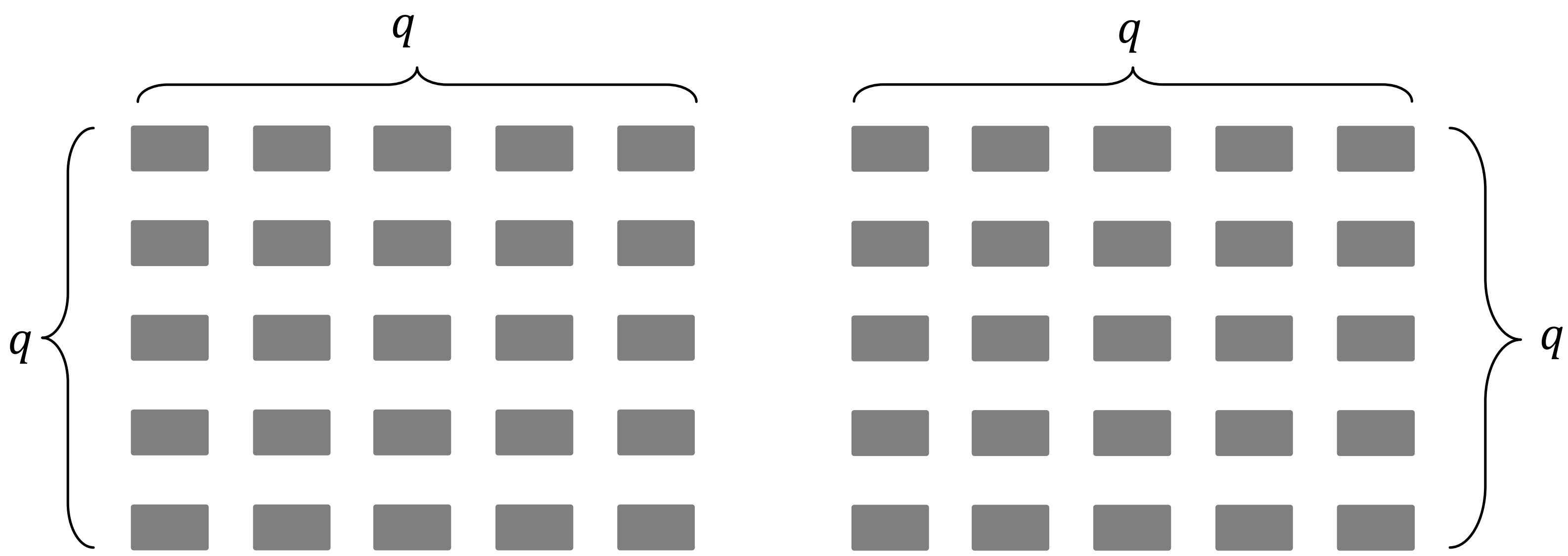
Number of routers: $2q^2$

Network radix: $(3q - \delta)/2$

2 *Construct a finite field $\mathcal{F}_q$.*

Assuming q is prime:

$\mathcal{F}_q = \mathbb{Z}/q\mathbb{Z}$



Diameter-2 Slim Fly

1 *Select a prime power q*

$$q = 4w + \delta;$$

$w \in \mathbb{N} \quad \delta \in \{-1,0,1\},$

A Slim Fly based on q:

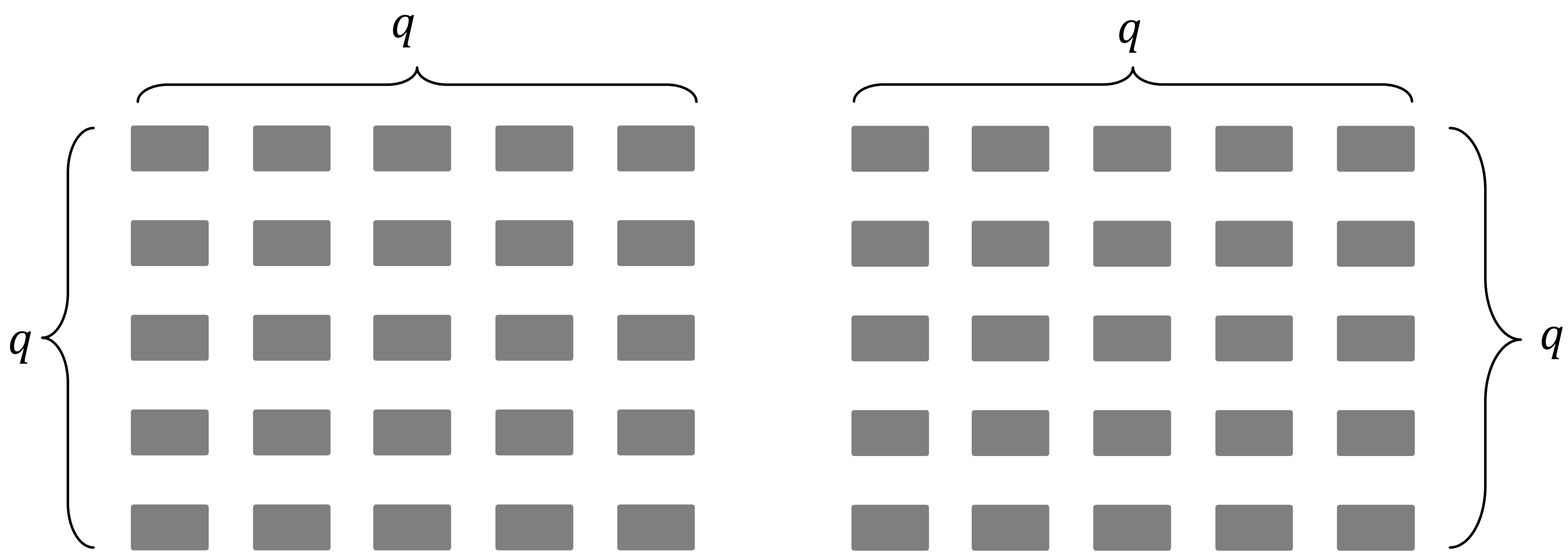
Number of routers: $2q^2$

Network radix: $(3q-\delta)/2$

2 *Construct a finite field $\mathcal{F}_q$.*

Assuming q is prime:

$$\mathcal{F}_q = \mathbb{Z}/q\mathbb{Z} = \{0,1,\dots,q-1\}$$



Diameter-2 Slim Fly

1 *Select a prime power q*

$$q = 4w + \delta;$$

$w \in \mathbb{N} \quad \delta \in \{-1,0,1\},$

A Slim Fly based on q:

Number of routers: $2q^2$

Network radix: $(3q - \delta)/2$

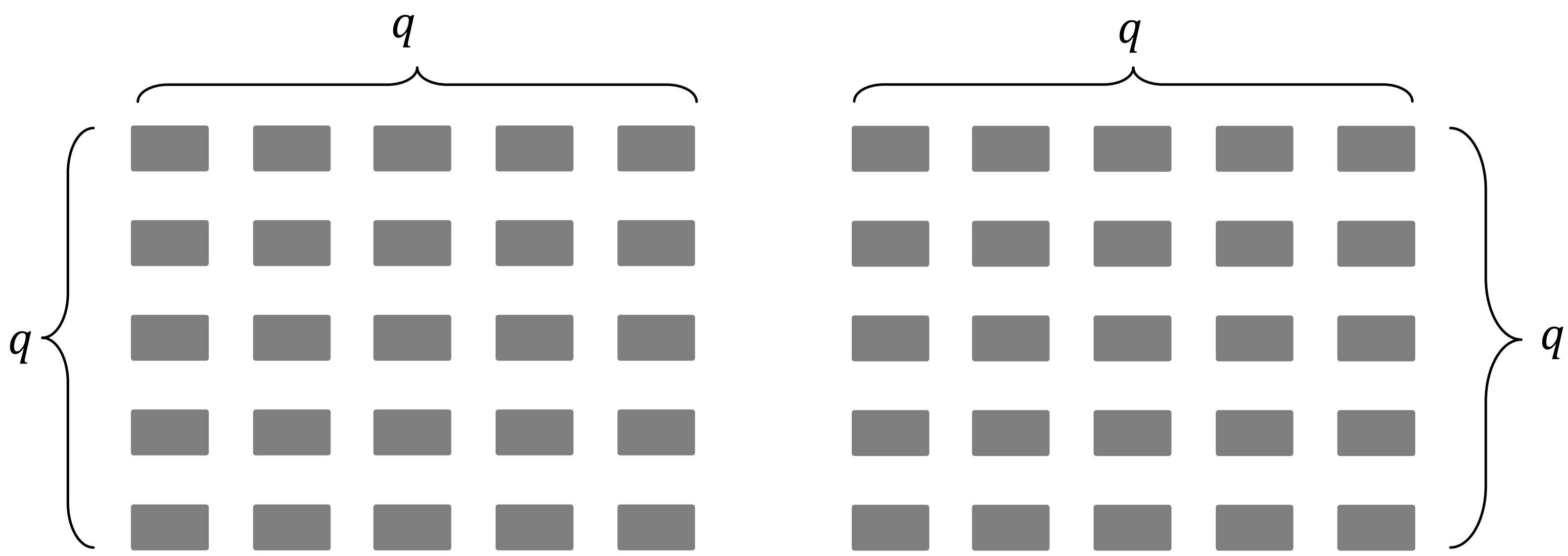
2 *Construct a finite field $\mathcal{F}_q$.*

Assuming *q* is prime:

$$\mathcal{F}_q = \mathbb{Z}/q\mathbb{Z} = \{0,1,\dots,q-1\}$$

with modular arithmetic.

E *Example*: $q = 5$



Diameter-2 Slim Fly

1 *Select a prime power q*

$$q = 4w + \delta;$$

$w \in \mathbb{N} \quad \delta \in \{-1,0,1\}$,

A Slim Fly based on q:

Number of routers: $2q^2$

Network radix: $(3q - \delta)/2$

2 *Construct a finite field $\mathcal{F}_q$.*

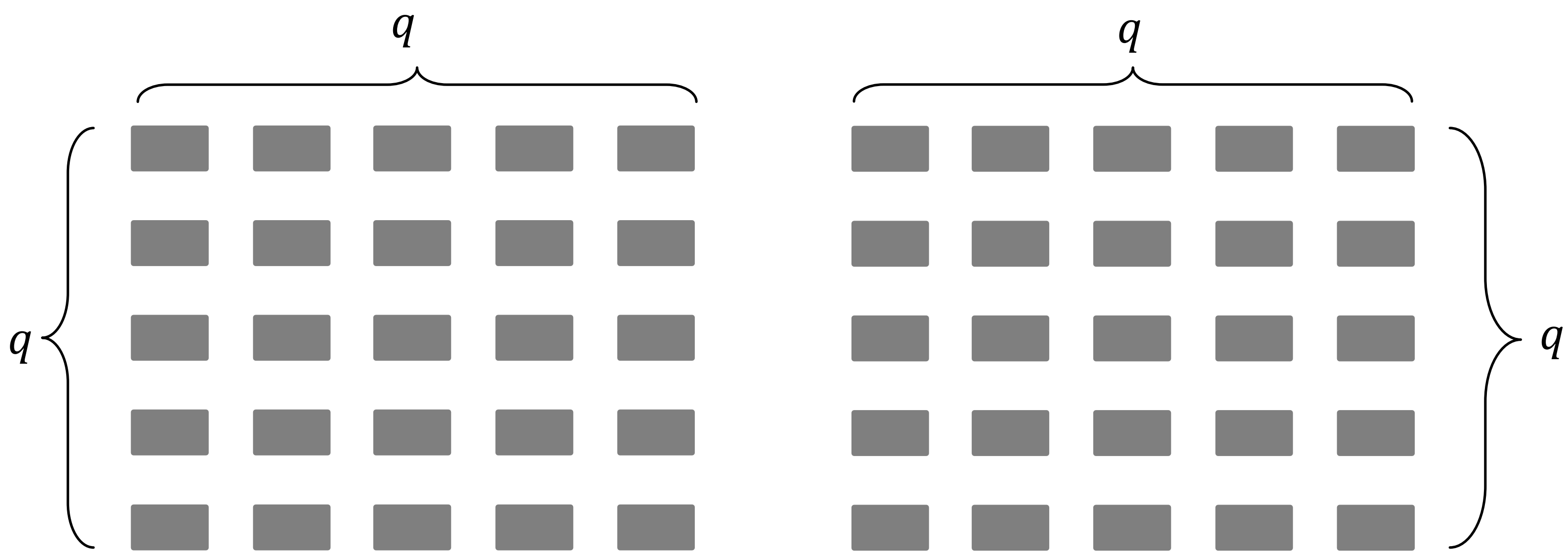
Assuming *q* is prime:

$$\mathcal{F}_q = \mathbb{Z}/q\mathbb{Z} = \{0,1,\dots,q-1\}$$

with modular arithmetic.

E *Example*: $q = 5$

50 routers
network radix*:* 7



Diameter-2 Slim Fly

1 *Select a prime power q*

$$q = 4w + \delta;$$

$w \in \mathbb{N} \quad \delta \in \{-1,0,1\},$

A Slim Fly based on q:

Number of routers: $2q^2$

Network radix: $(3q - \delta)/2$

2 *Construct a finite field* $\mathcal{F}_q$.

Assuming q is prime:

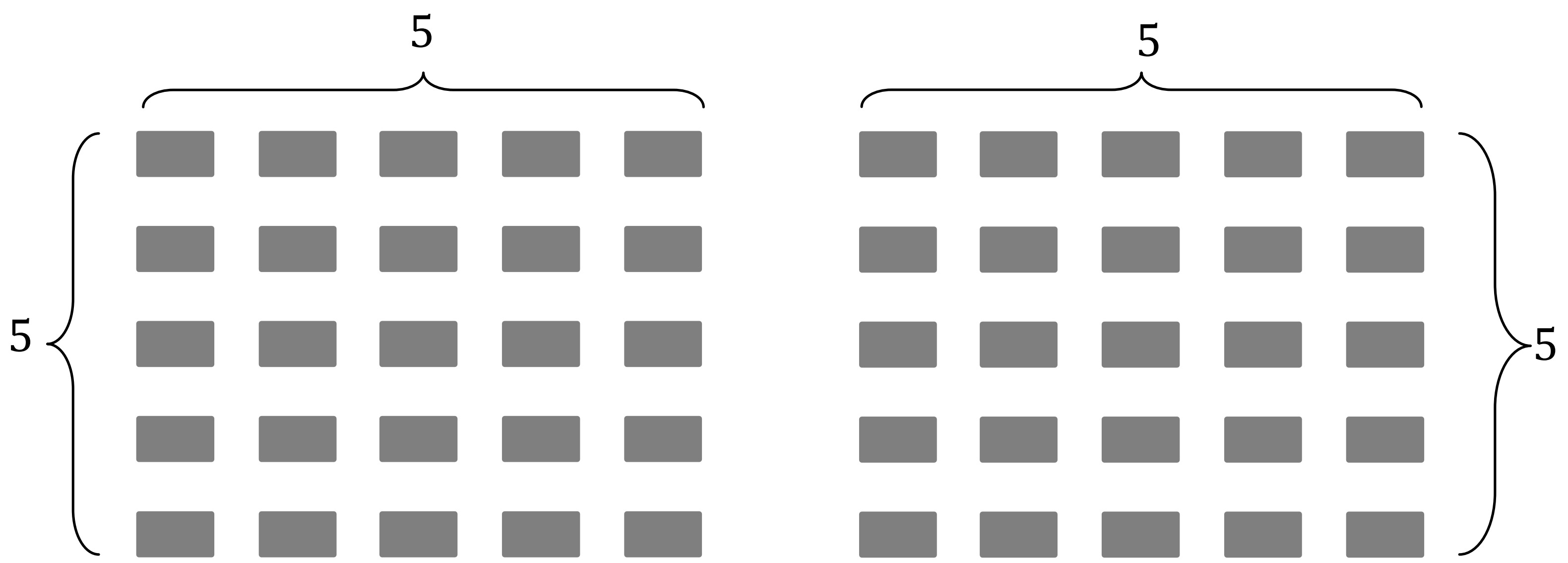
$$\mathcal{F}_q = \mathbb{Z}/q\mathbb{Z} = \{0,1,\dots,q-1\}$$

with modular arithmetic.

E *Example*: $q = 5$

50 routers

network radix: 7



Diameter-2 Slim Fly

1 *Select a prime power q*

$$q = 4w + \delta;$$

$w \in \mathbb{N} \quad \delta \in \{-1,0,1\},$

A Slim Fly based on q:

Number of routers: $2q^2$

Network radix: $(3q - \delta)/2$

2 *Construct a finite field $\mathcal{F}_q$.*

Assuming *q* is prime:

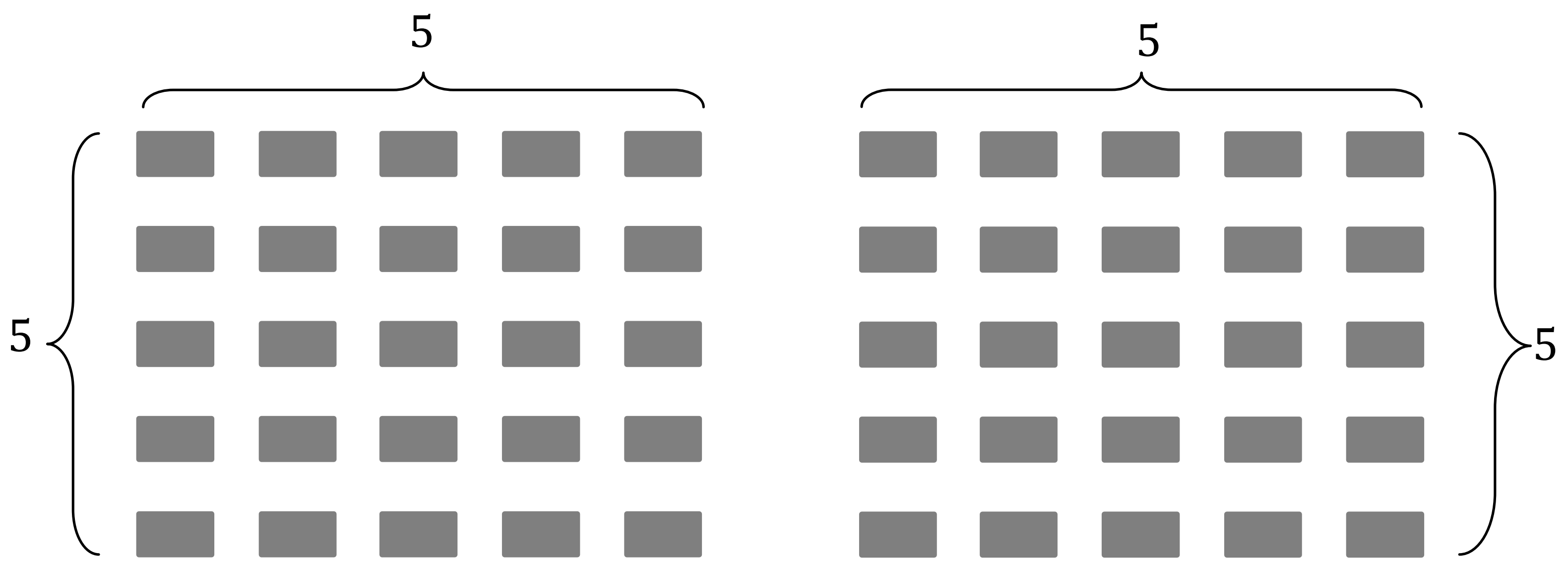
$$\mathcal{F}_q = \mathbb{Z}/q\mathbb{Z} = \{0,1,\ldots,q-1\}$$

with modular arithmetic.

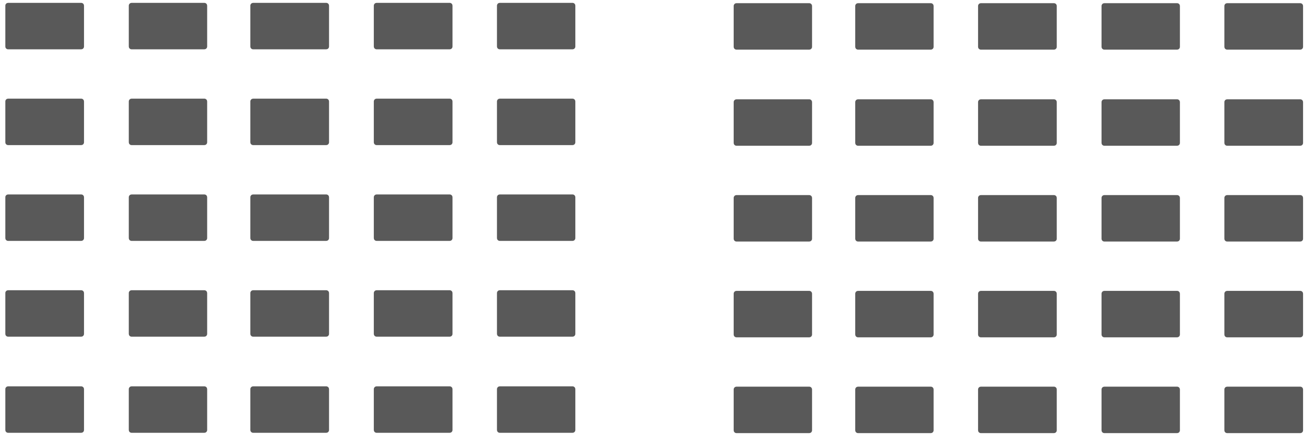
E *Example*: $q = 5$

50 routers
network radix: 7

$\mathcal{F}_5 = \{0,1,2,3,4\}$

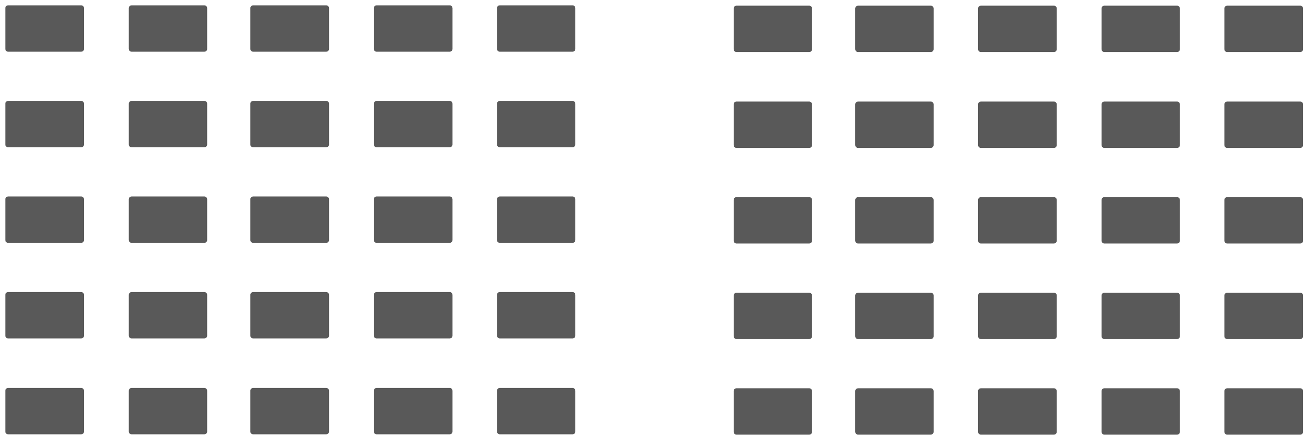


Diameter-2 Slim Fly



Diameter-2 Slim Fly

3 *Label the routers*

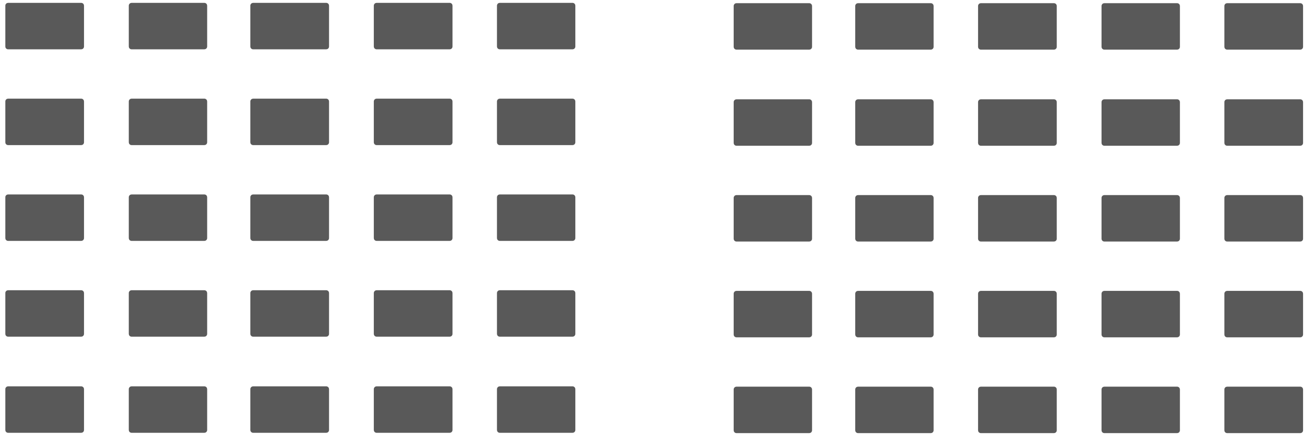


Diameter-2 Slim Fly

3 *Label the routers*

Set of routers:

$$\{0,1\} \times \mathcal{F}_q \times \mathcal{F}_q$$



Diameter-2 Slim Fly

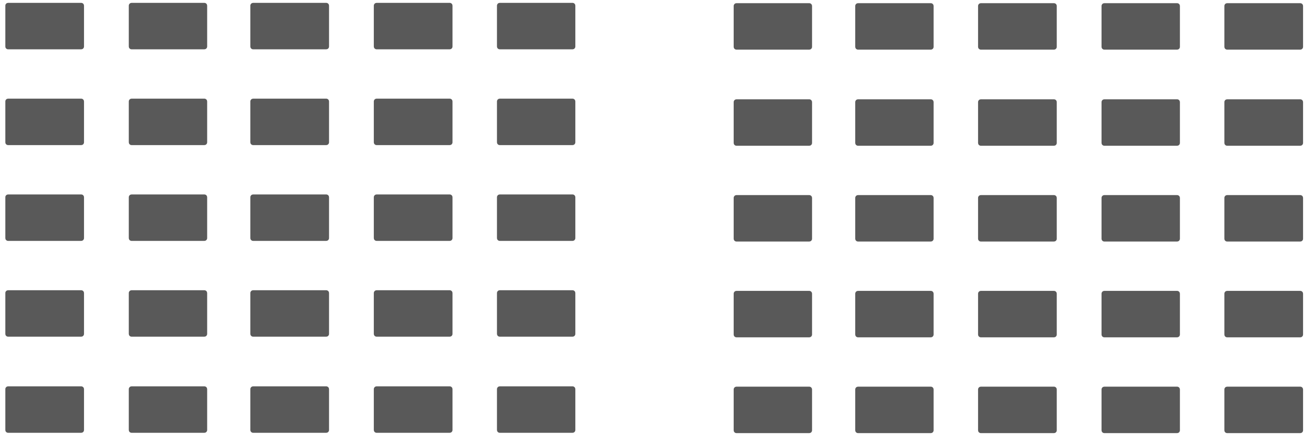
3 *Label the routers*

Set of routers:

$$\{0,1\} \times \mathcal{F}_q \times \mathcal{F}_q$$

E *Example:* $q = 5$

…



Diameter-2 Slim Fly

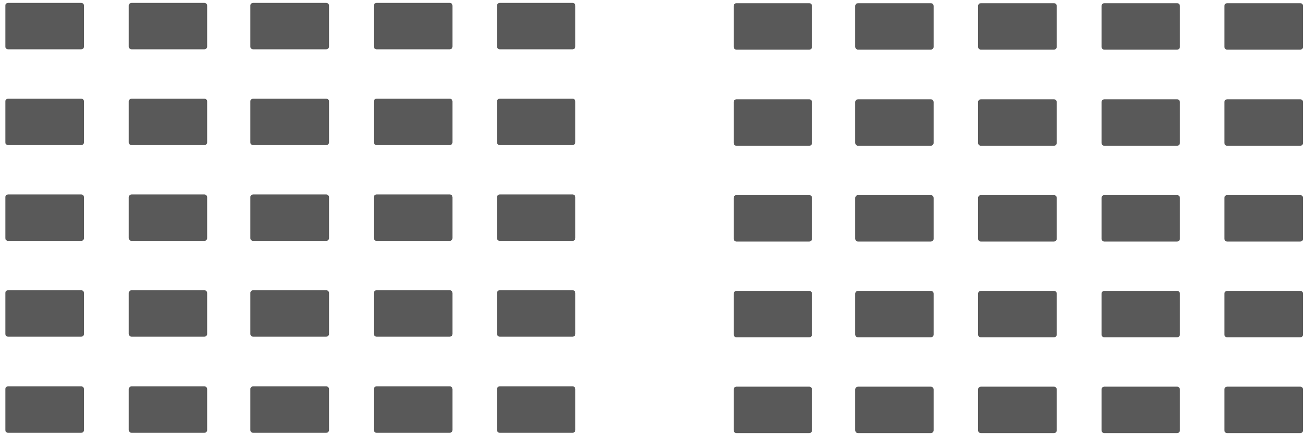
3 *Label the routers*

Set of routers:

$\{0,1\} \times \mathcal{F}_q \times \mathcal{F}_q$

E *Example:* $q = 5$

…



Diameter-2 Slim Fly

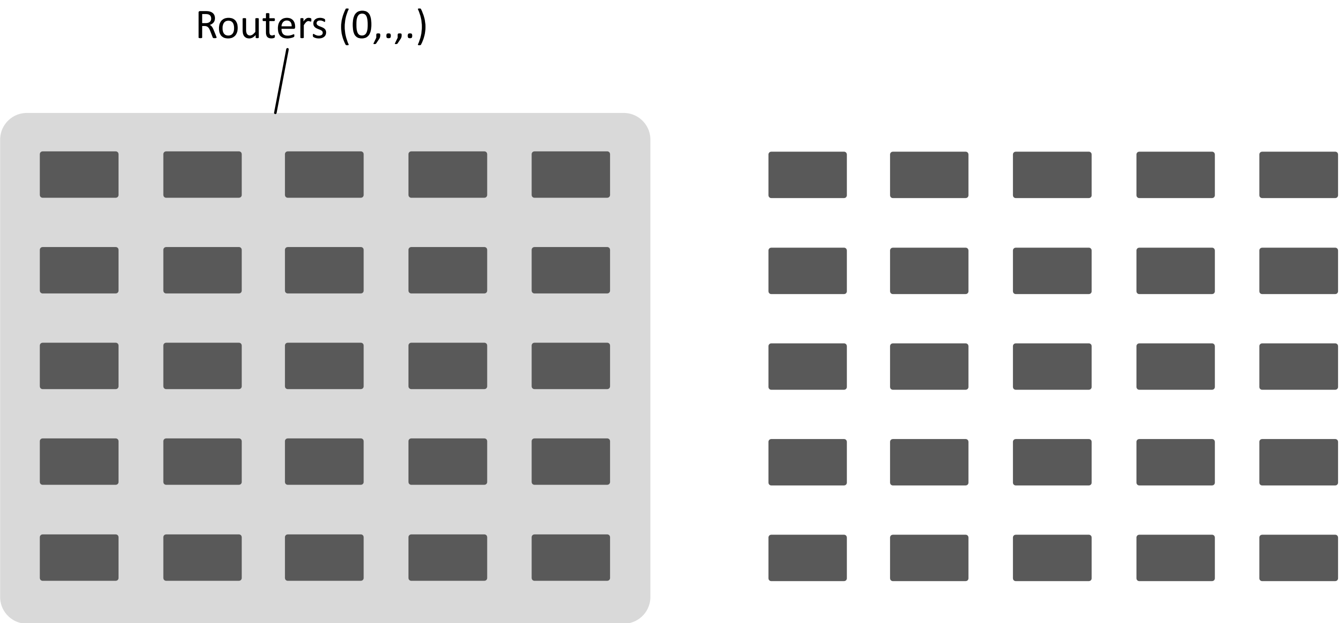
3 *Label the routers*

Set of routers:

$\{0,1\} \times \mathcal{F}_q \times \mathcal{F}_q$

E *Example:* $q = 5$

...



Diameter-2 Slim Fly

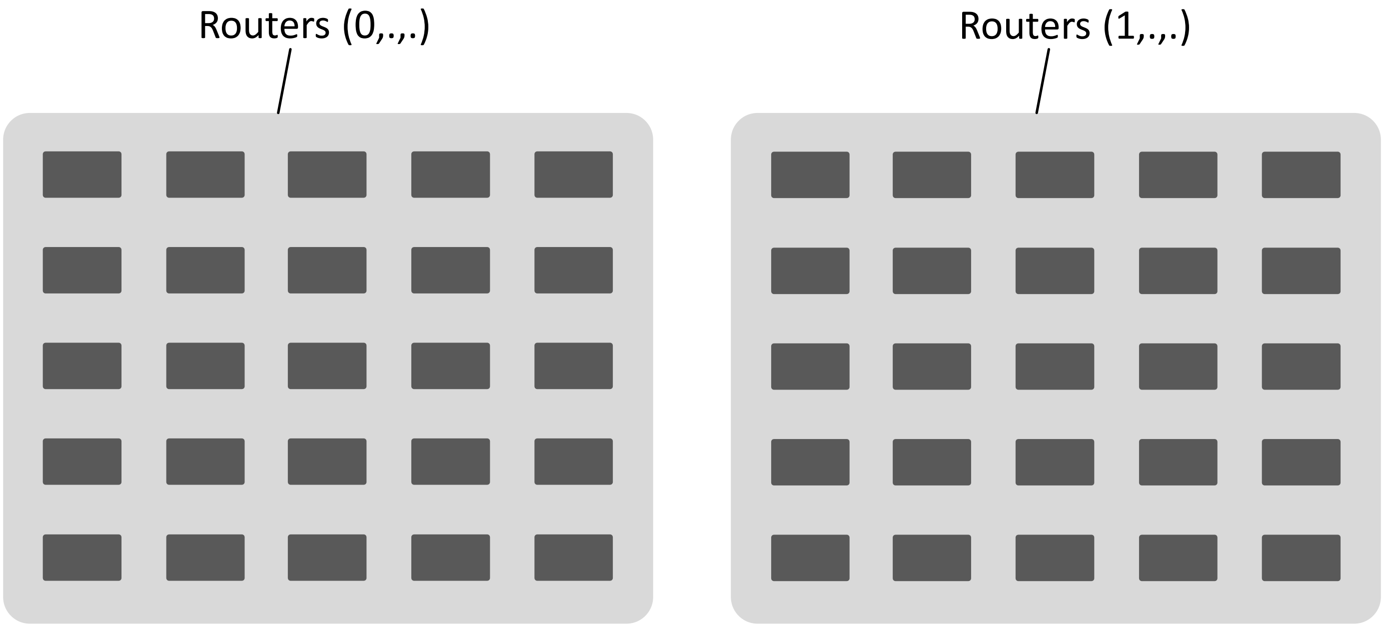
3 *Label the routers*

Set of routers:

$$\{0,1\} \times \mathcal{F}_q \times \mathcal{F}_q$$

E *Example:* $q = 5$

…



Diameter-2 Slim Fly

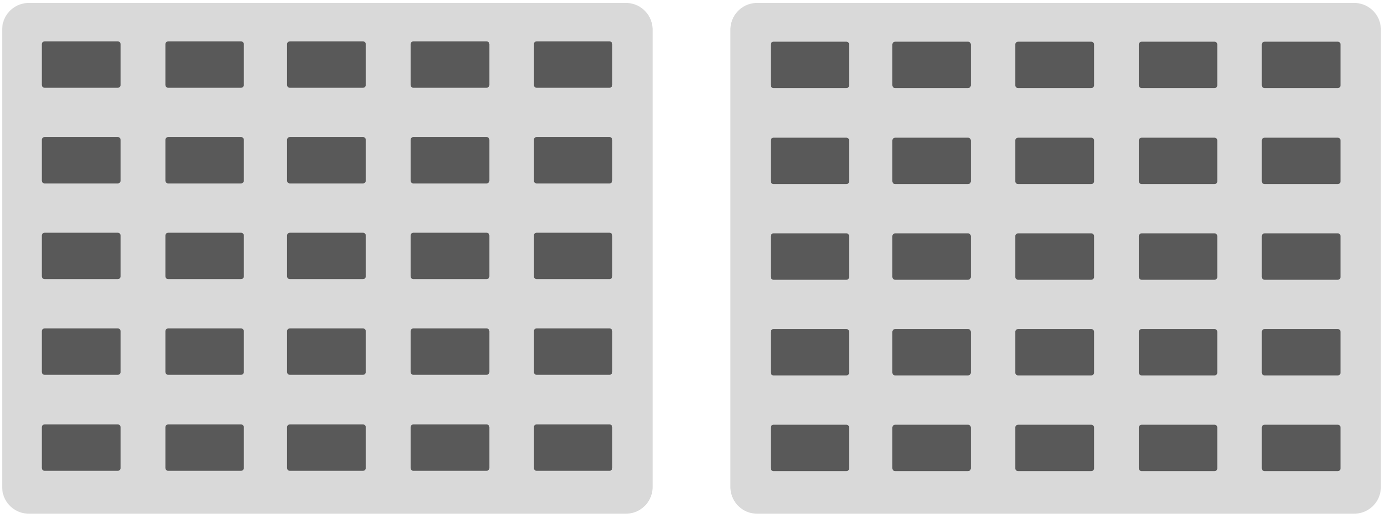
3 *Label the routers*

Set of routers:

$\{0,1\} \times \mathcal{F}_q \times \mathcal{F}_q$

E *Example:* $q = 5$

...



Diameter-2 Slim Fly

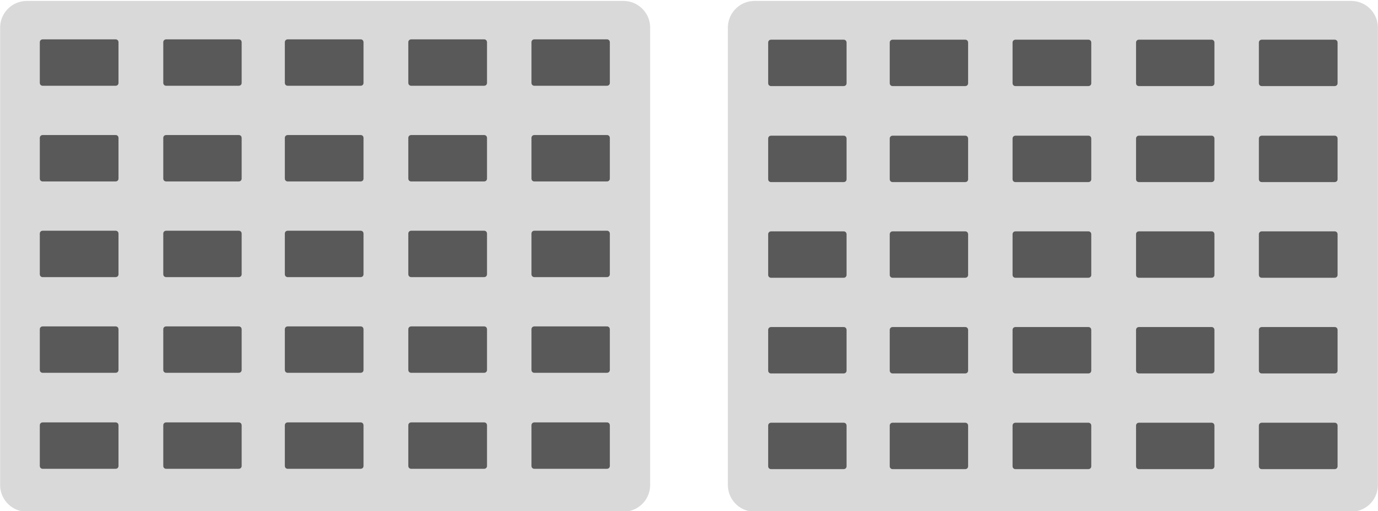
3 *Label the routers*

Set of routers:

$\{0,1\} \times \mathcal{F}_q \times \mathcal{F}_q$

E *Example:* $q = 5$

...



Diameter-2 Slim Fly

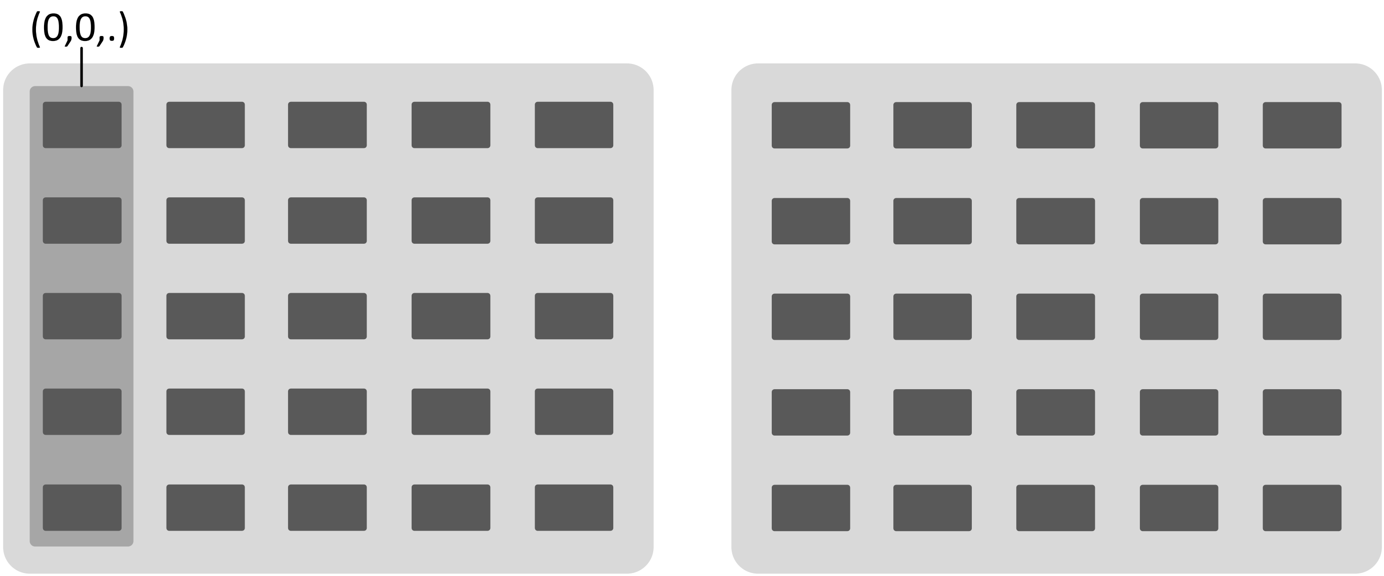
3 *Label the routers*

Set of routers:

$\{0,1\} \times \mathcal{F}_q \times \mathcal{F}_q$

E *Example:* $q = 5$

...



Diameter-2 Slim Fly

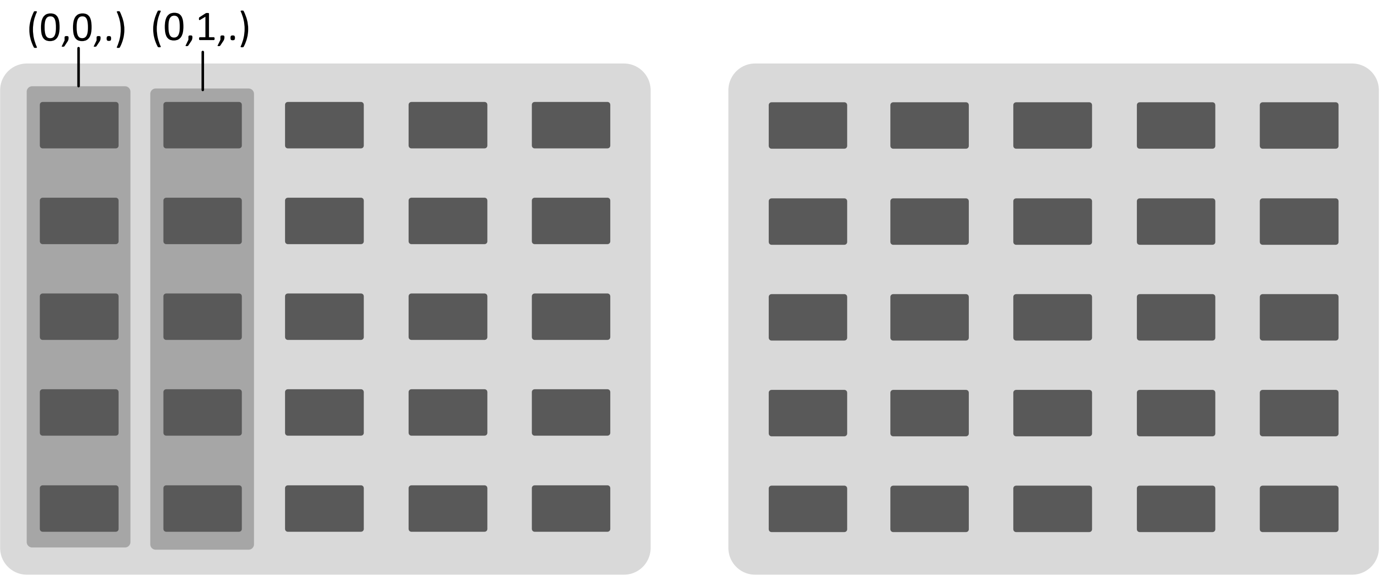
3 *Label the routers*

Set of routers:

$\{0,1\} \times \mathcal{F}_q \times \mathcal{F}_q$

E *Example:* $q = 5$

...



Diameter-2 Slim Fly

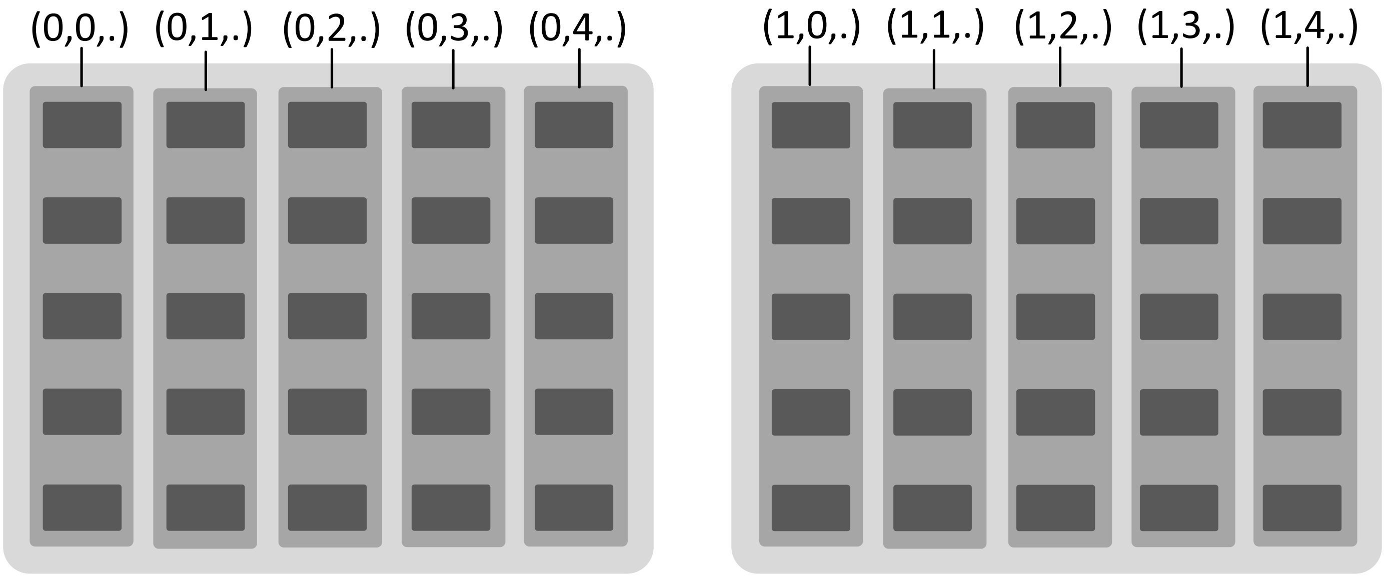
3 *Label the routers*

Set of routers:

$\{0,1\} \times \mathcal{F}_q \times \mathcal{F}_q$

E *Example:* $q = 5$

...



Diameter-2 Slim Fly

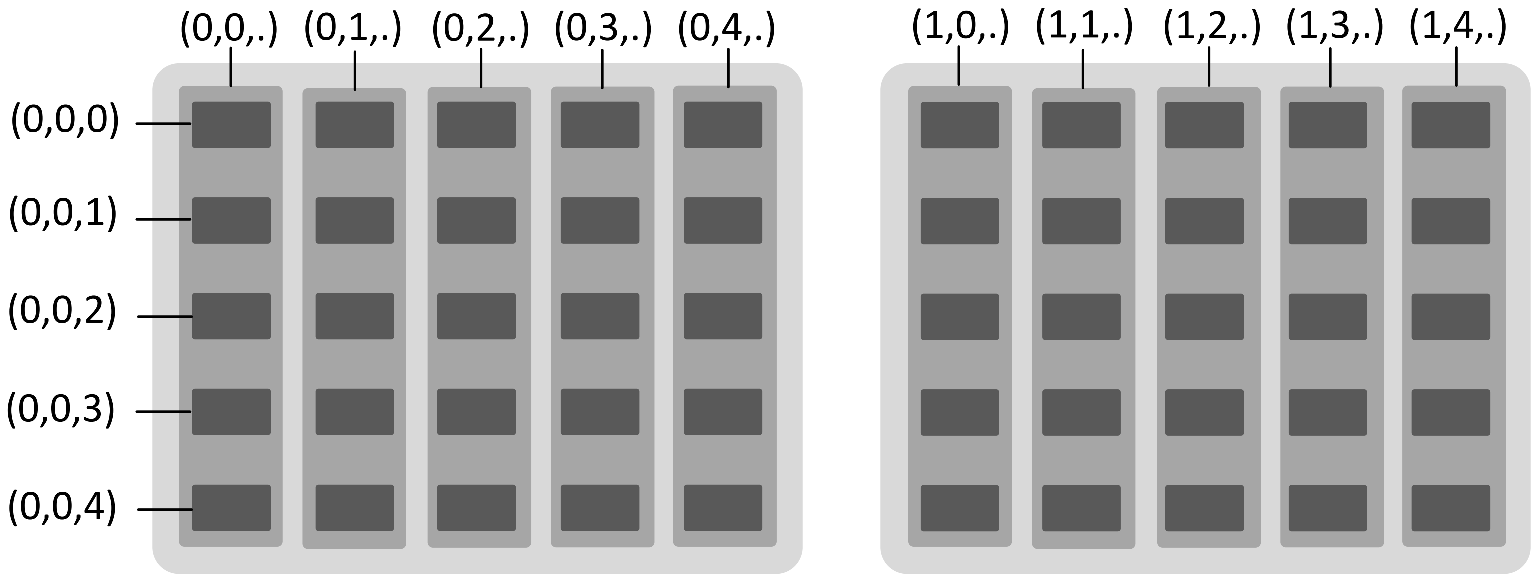
3 *Label the routers*

Set of routers:

$\{0,1\} \times \mathcal{F}_q \times \mathcal{F}_q$

E *Example:* $q = 5$

...



Diameter-2 Slim Fly

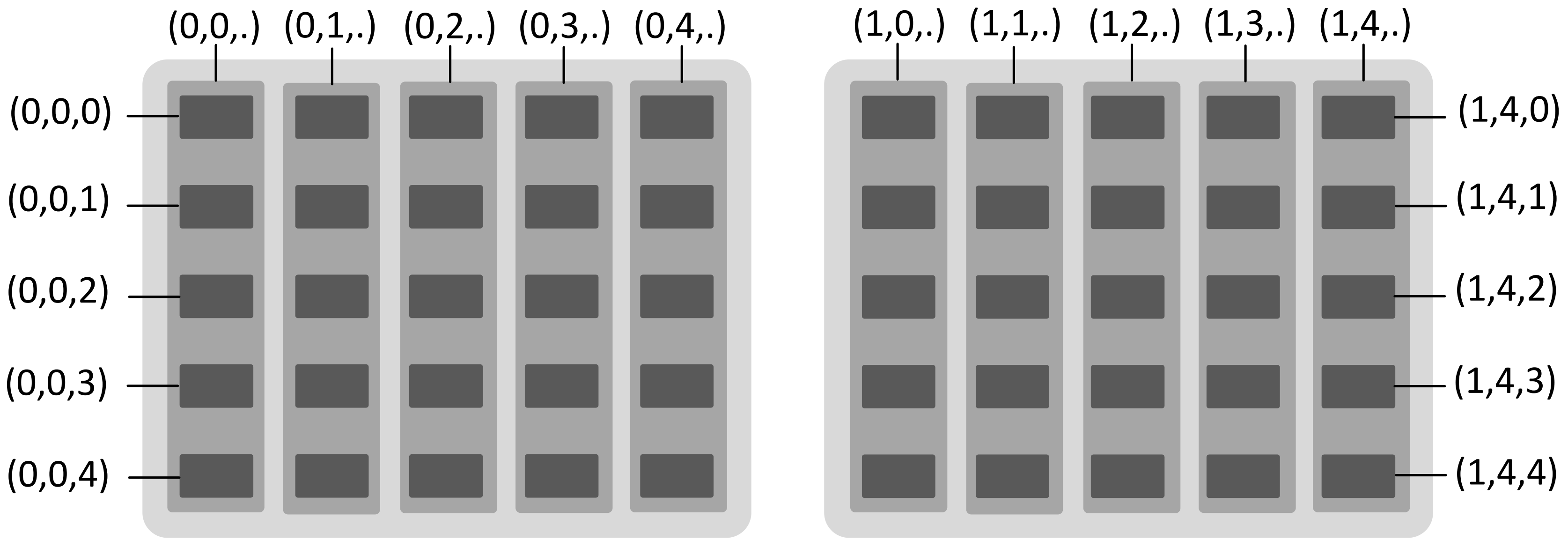
3 *Label the routers*

Set of routers:

$$\{0,1\} \times \mathcal{F}_q \times \mathcal{F}_q$$

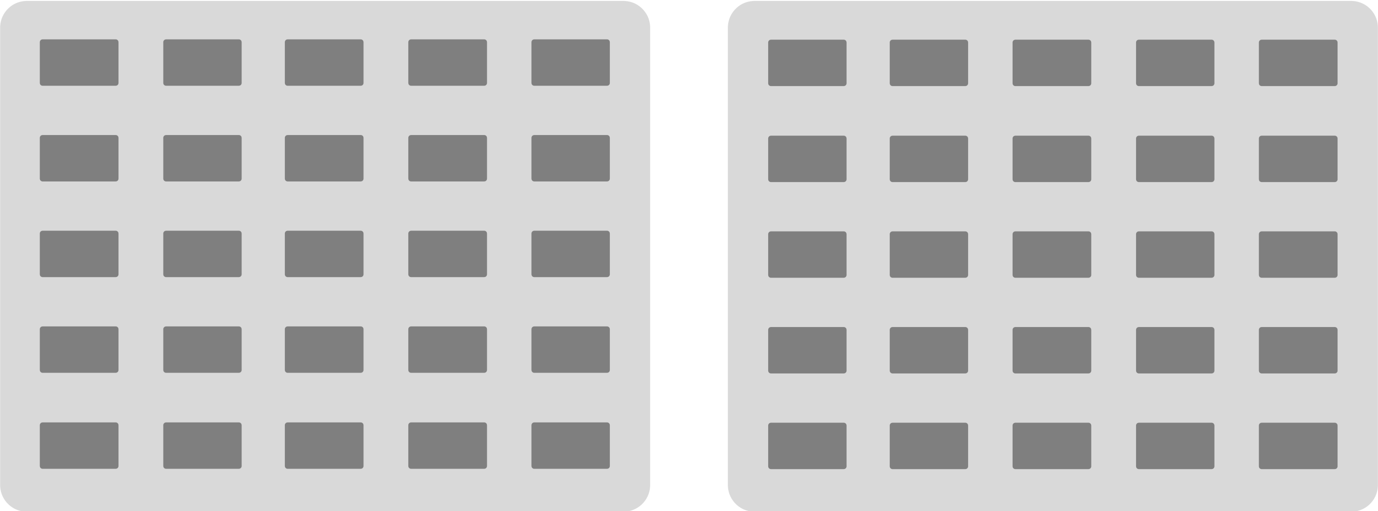
E *Example:* $q = 5$

…



Diameter-2 Slim Fly

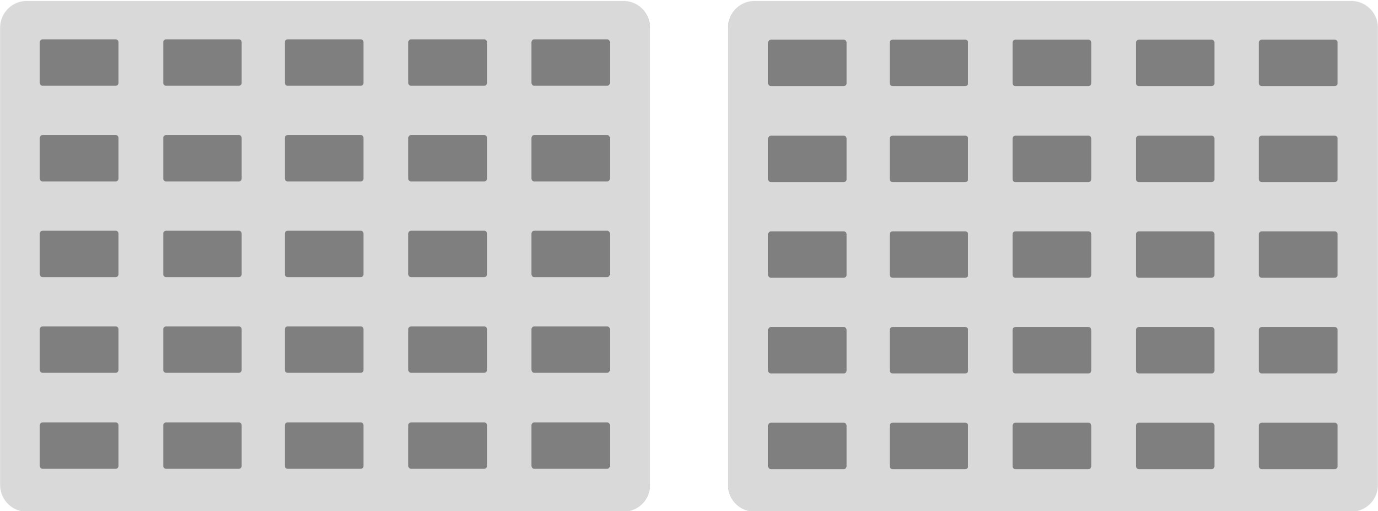
4 *Find primitive element* ξ



Diameter-2 Slim Fly

4 *Find primitive element* ξ

$\xi \in \mathcal{F}_q$ generates $\mathcal{F}_q$:

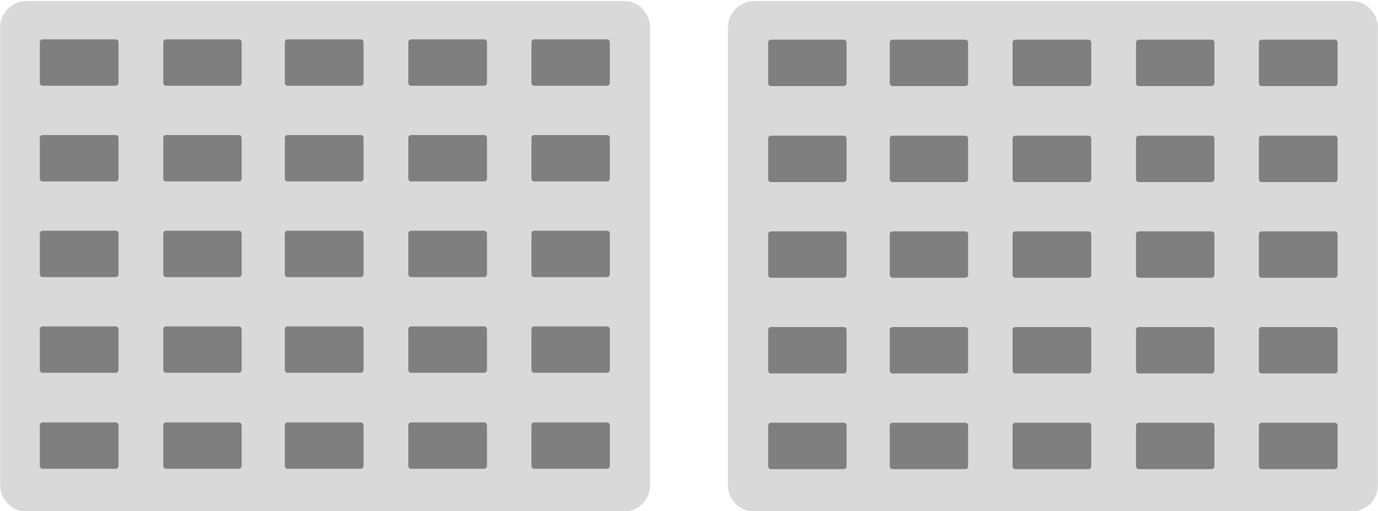


Diameter-2 Slim Fly

4 *Find primitive element* ξ

$\xi \in \mathcal{F}_q$ generates $\mathcal{F}_q$:

All non-zero elements of $\mathcal{F}_q$ can be written as ξ^i; $i \in \mathbb{N}$



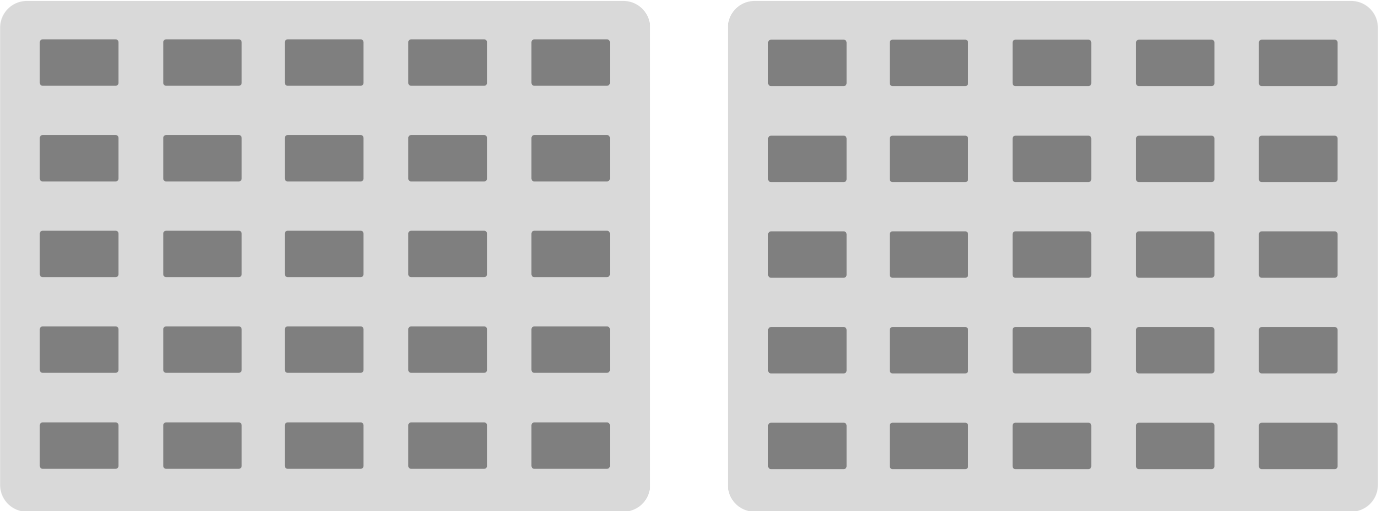
Diameter-2 Slim Fly

4 *Find primitive element* ξ

$\xi \in \mathcal{F}_q$ generates $\mathcal{F}_q$:

All non-zero elements of $\mathcal{F}_q$ can be written as ξ^i; $i \in \mathbb{N}$

E *Example:* $q = 5$



Diameter-2 Slim Fly

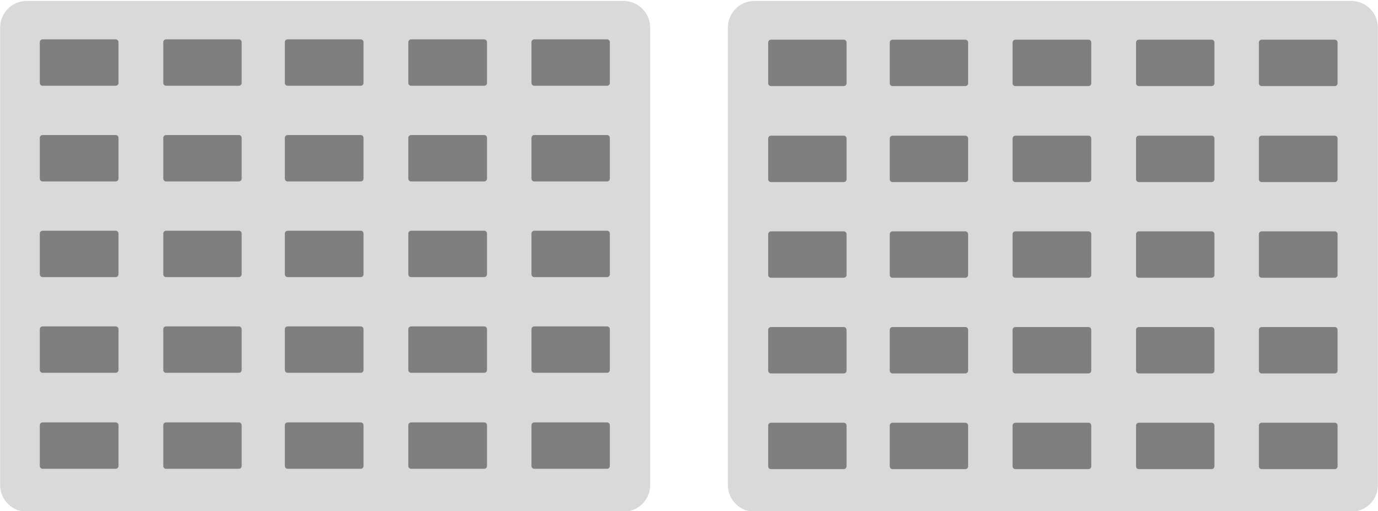
4 *Find primitive element* ξ

$\xi \in \mathcal{F}_q$ generates $\mathcal{F}_q$:

All non-zero elements of $\mathcal{F}_q$ can be written as $\xi^i;\ i \in \mathbb{N}$

E *Example:* $q = 5$

$\mathcal{F}_5 = \{0,1,2,3,4\}$



Diameter-2 Slim Fly

4 *Find primitive element* ξ

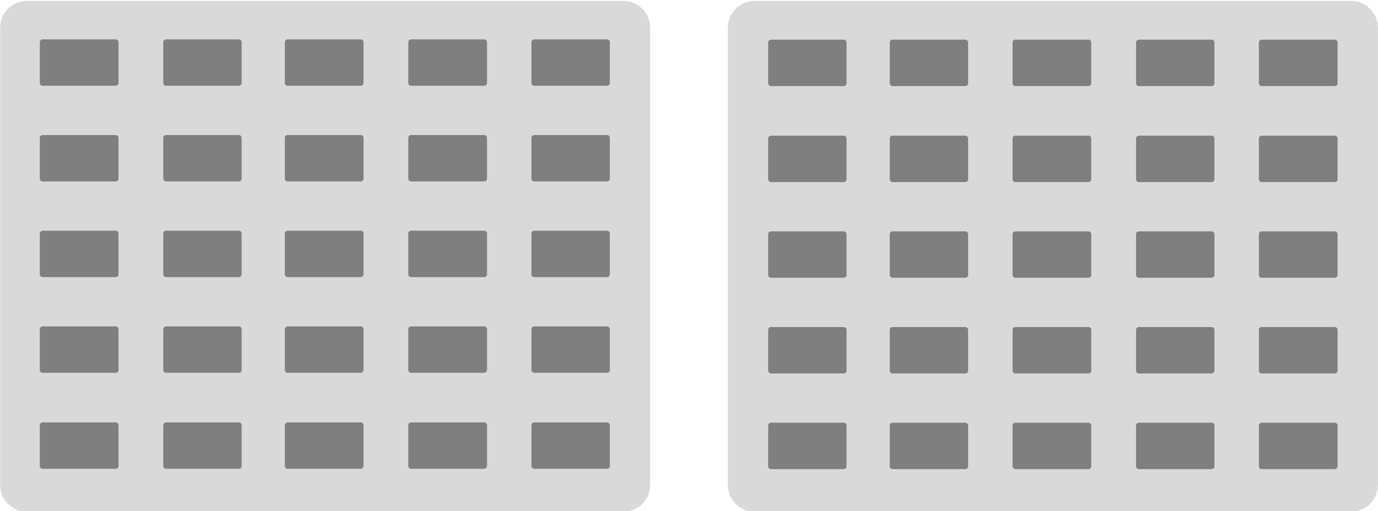
$\xi \in \mathcal{F}_q$ generates $\mathcal{F}_q$:

All non-zero elements of $\mathcal{F}_q$ can be written as ξ^i; $i \in \mathbb{N}$

E *Example:* $q = 5$

$\mathcal{F}_5 = \{0,1,2,3,4\}$

$\xi = 2$



Diameter-2 Slim Fly

4 *Find primitive element* ξ

$\xi \in \mathcal{F}_q$ generates $\mathcal{F}_q$:

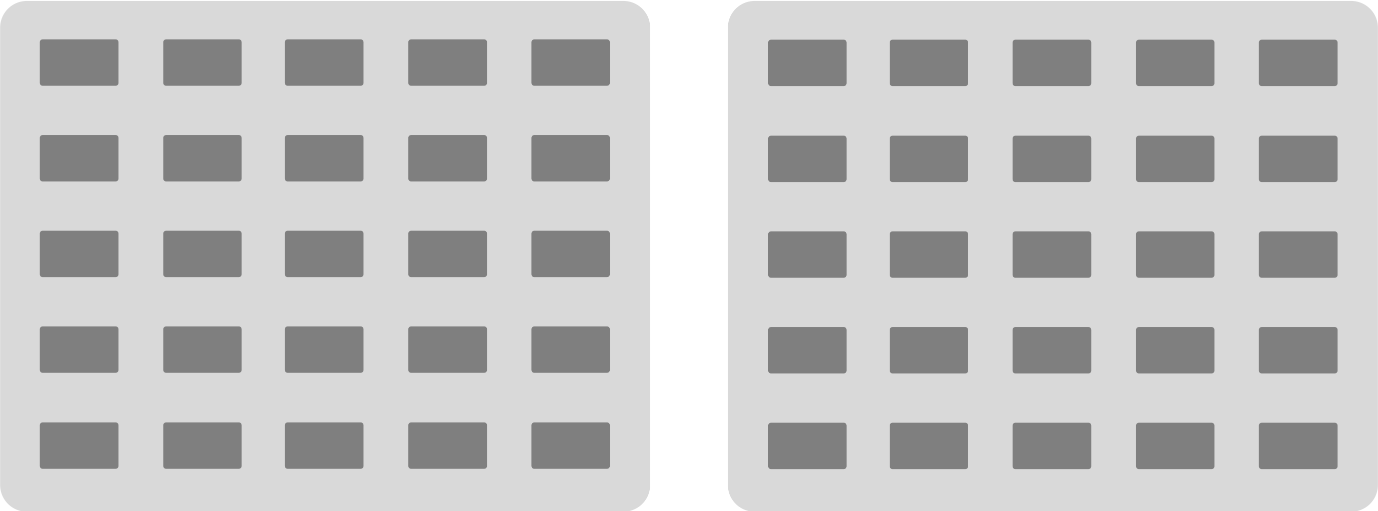
All non-zero elements of $\mathcal{F}_q$ can be written as $\xi^i;\ i \in \mathbb{N}$

E *Example:* $q = 5$

$\mathcal{F}_5 = \{0,1,2,3,4\}$

$\xi = 2$

$1 = \xi^4 \; mod \; 5 = 2^4 \; mod \; 5 = 16 \; mod \; 5$



Diameter-2 Slim Fly

4 *Find primitive element* ξ

$\xi \in \mathcal{F}_q$ generates $\mathcal{F}_q$:

All non-zero elements of $\mathcal{F}_q$ can be written as ξ^i; $i \in \mathbb{N}$

5 *Build Generator Sets*

$$X = \{1, \xi^2, \dots, \xi^{q-3}\}$$

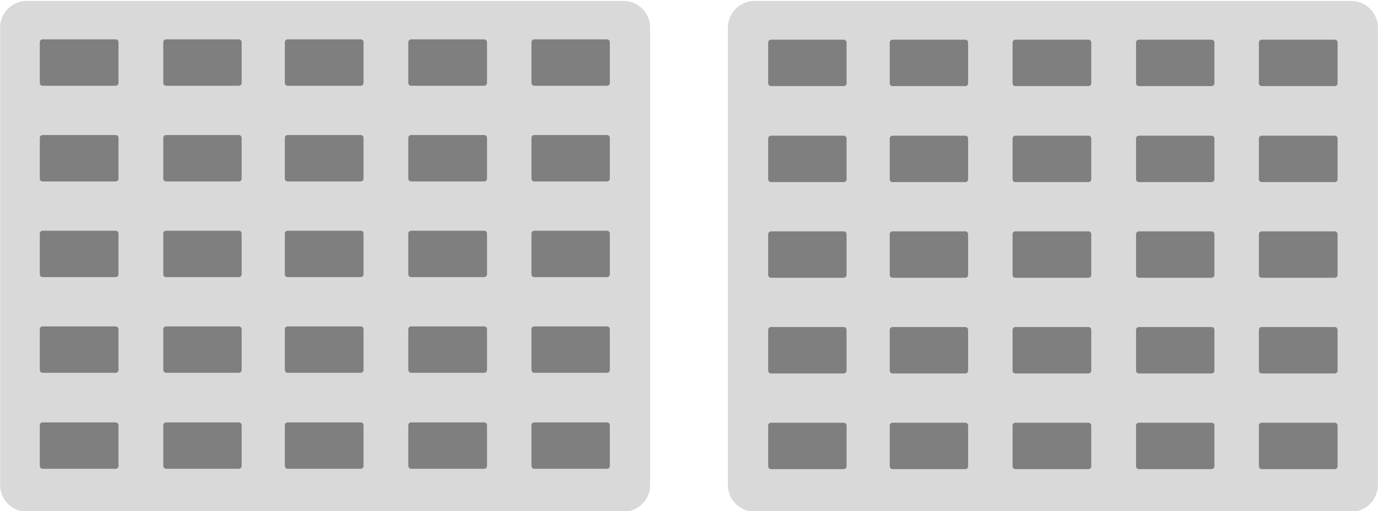
$$X' = \{\xi, \xi^3, \dots, \xi^{q-2}\}$$

E *Example:* $q = 5$

$$\mathcal{F}_5 = \{0,1,2,3,4\}$$

$$\xi = 2$$

$$1 = \xi^4 \; mod \; 5 = 2^4 \; mod \; 5 = 16 \; mod \; 5$$



Diameter-2 Slim Fly

4 *Find primitive element* ξ

$\xi \in \mathcal{F}_q$ generates $\mathcal{F}_q$:

All non-zero elements of $\mathcal{F}_q$ can be written as $\xi^i;\ i \in \mathbb{N}$

5 *Build Generator Sets*

$$X = \{1, \xi^2, \ldots, \xi^{q-3}\}$$

$$X' = \{\xi, \xi^3, \ldots, \xi^{q-2}\}$$

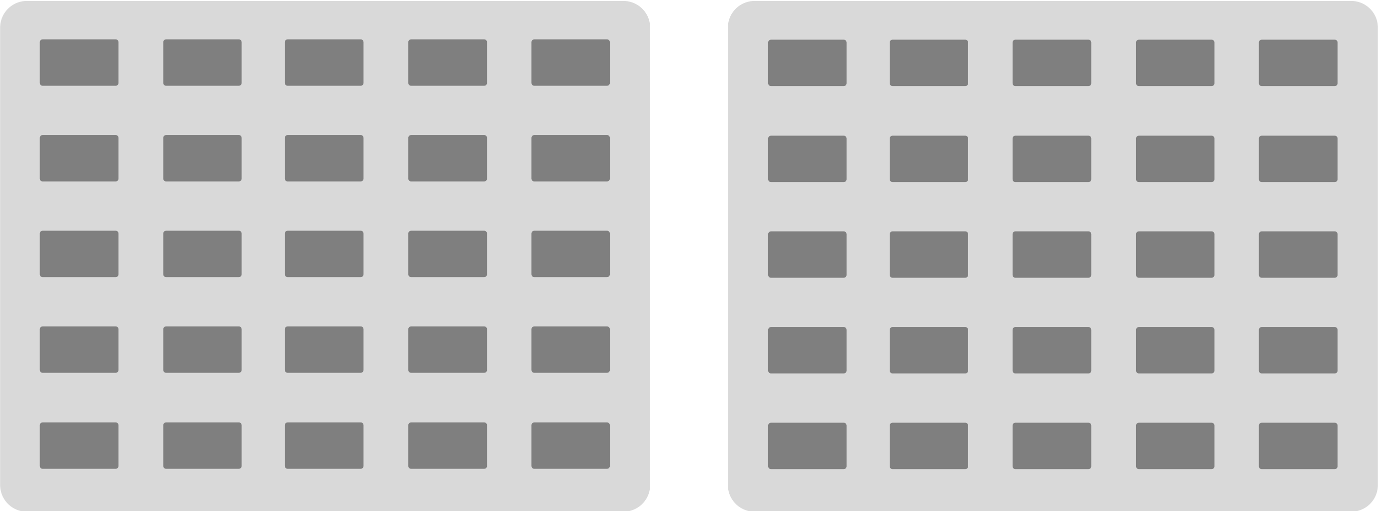
E *Example:* $q = 5$

$$\mathcal{F}_5 = \{0,1,2,3,4\}$$

$$\xi = 2$$

$$1 = \xi^4 \ mod \ 5 = 2^4 \ mod \ 5 = 16 \ mod \ 5$$

$$X = \{1,4\}$$



Diameter-2 Slim Fly

4 *Find primitive element* ξ

$\xi \in \mathcal{F}_q$ generates $\mathcal{F}_q$:

All non-zero elements of $\mathcal{F}_q$ can be written as ξ^i; $i \in \mathbb{N}$

5 *Build Generator Sets*

$$X = \{1, \xi^2, \ldots, \xi^{q-3}\}$$

$$X' = \{\xi, \xi^3, \ldots, \xi^{q-2}\}$$

E *Example:* $q = 5$

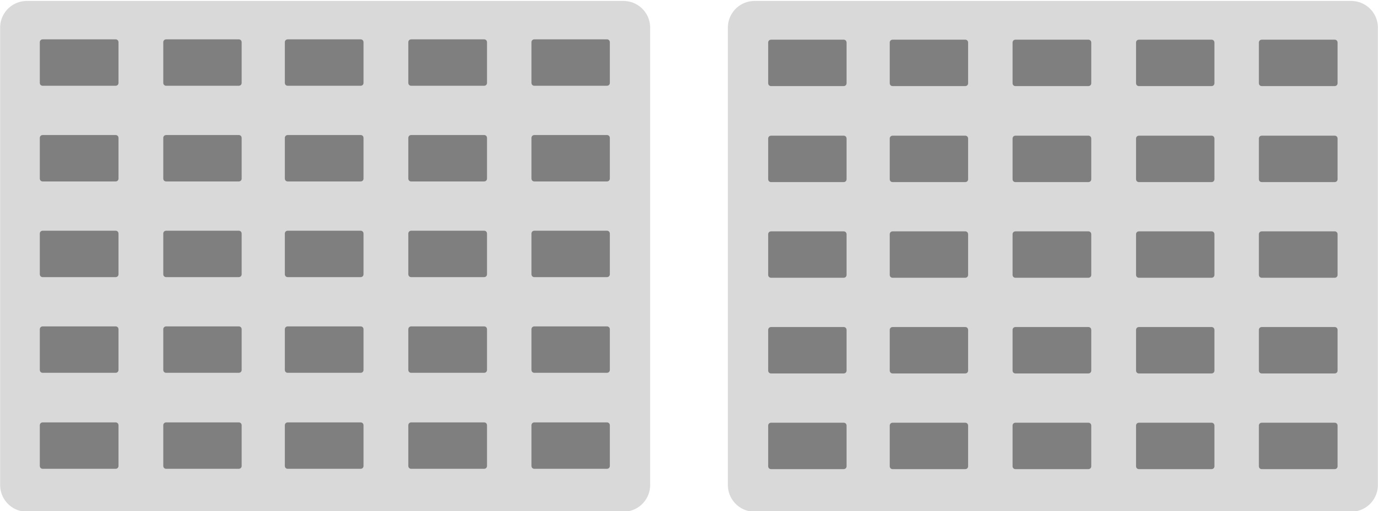
$$\mathcal{F}_5 = \{0,1,2,3,4\}$$

$$\xi = 2$$

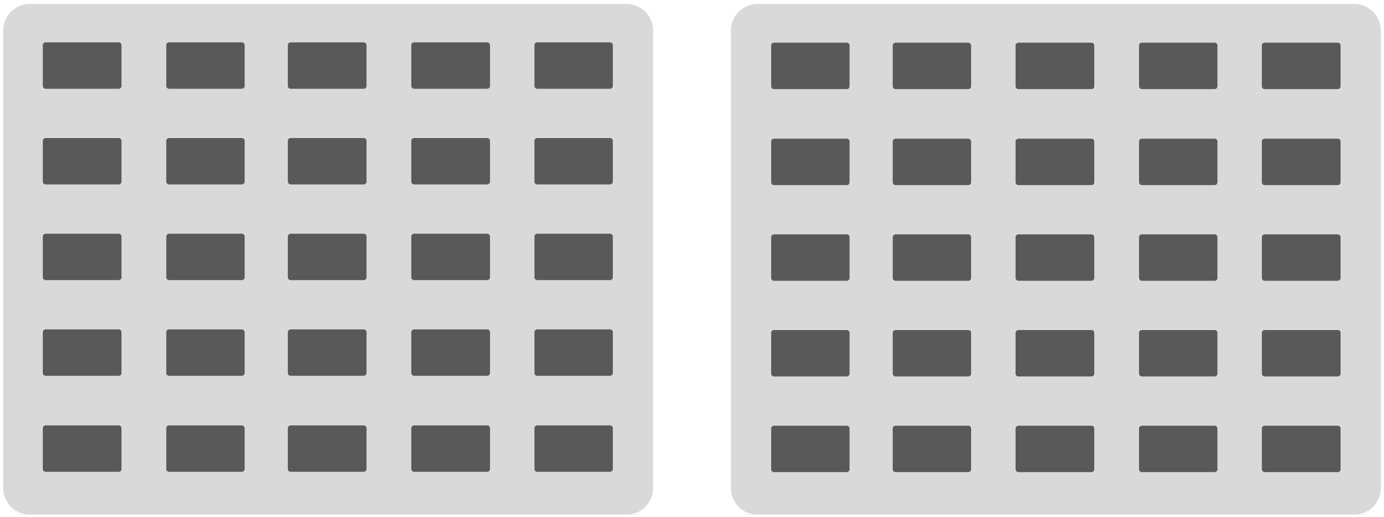
$$1 = \xi^4 \ mod\ 5 = 2^4 \ mod\ 5 = 16 \ mod\ 5$$

$$X = \{1,4\}$$

$$X' = \{2,3\}$$

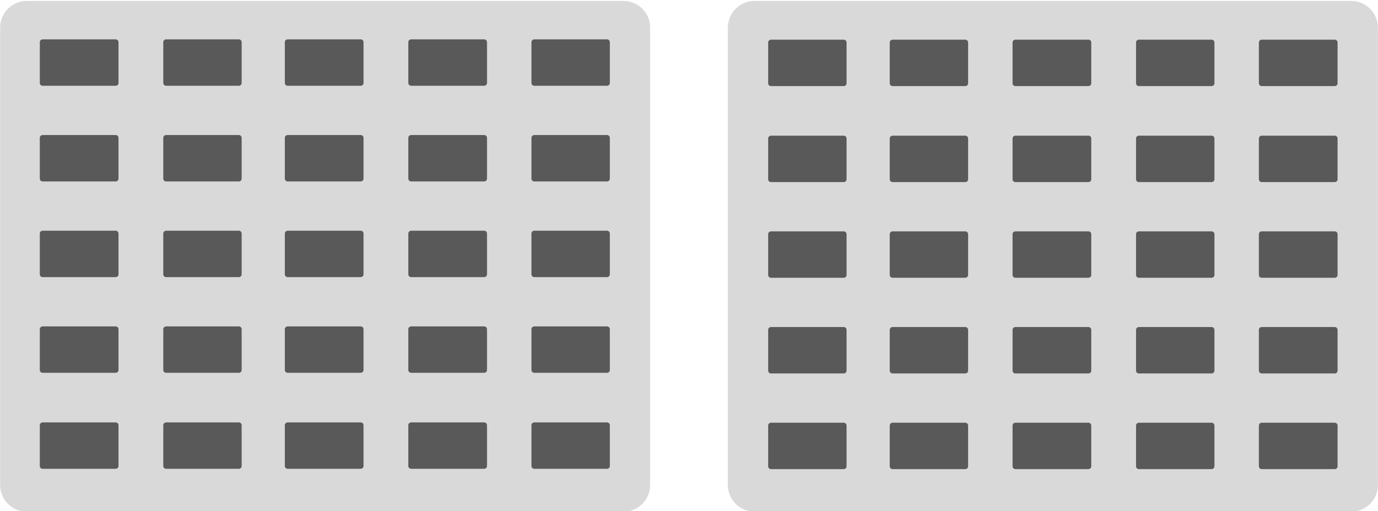


Diameter-2 Slim Fly



Diameter-2 Slim Fly

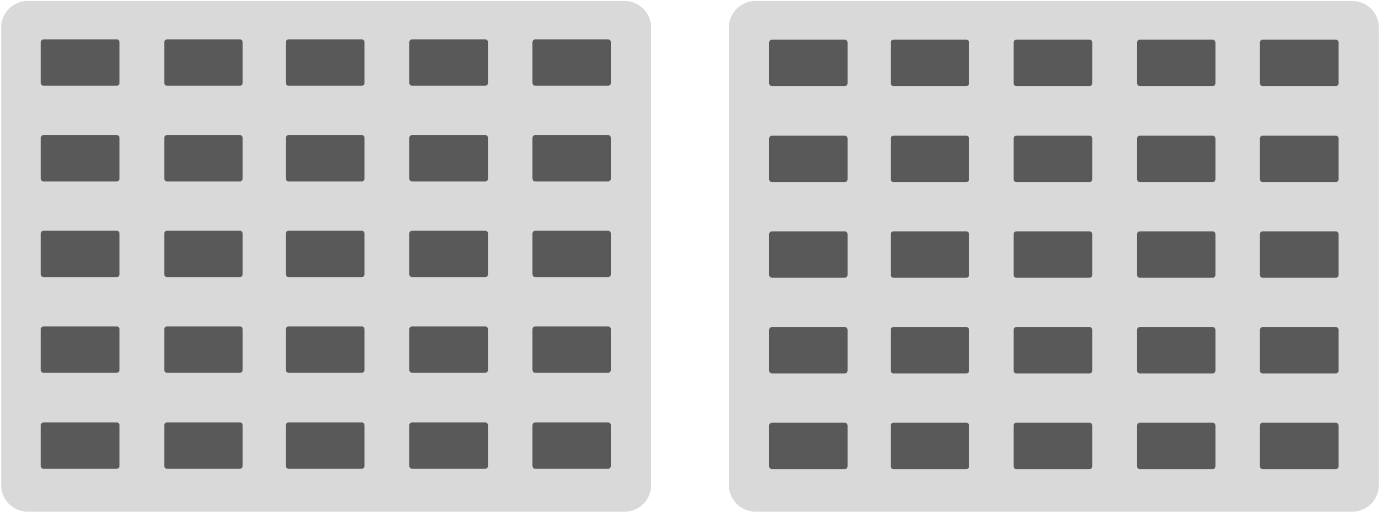
6 *Intra-group connections*



Diameter-2 Slim Fly

6 *Intra-group connections*

Two routers in one group are connected iff their "vertical Manhattan distance" is an element from:

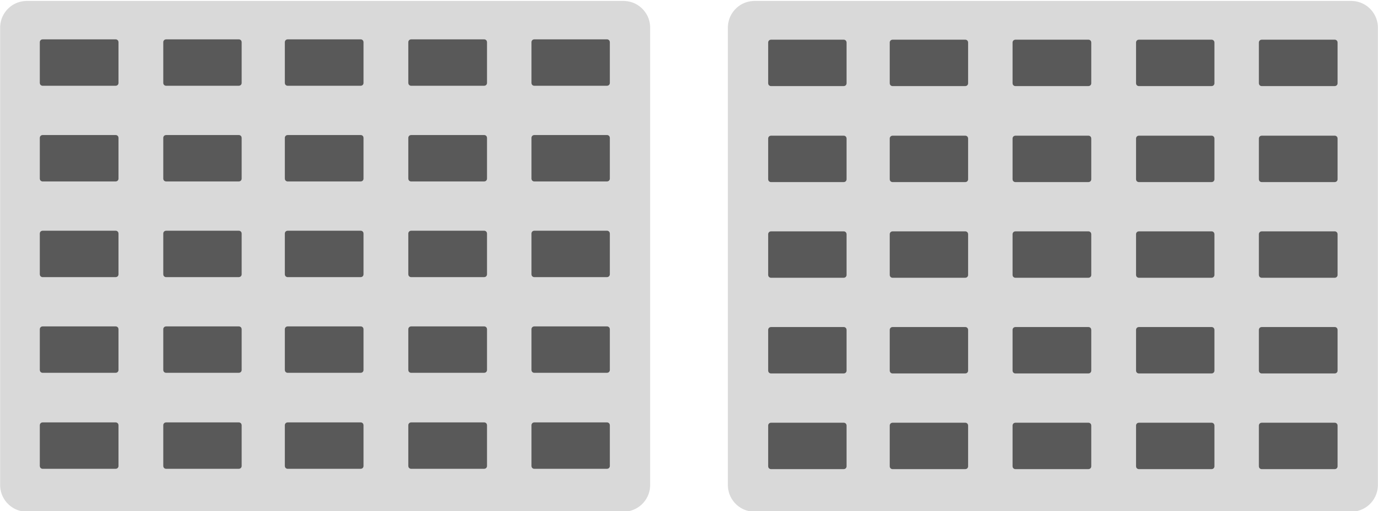


Diameter-2 Slim Fly

6 *Intra-group connections*

Two routers in one group are connected iff their “vertical Manhattan distance” is an element from:

$$X = \{1, \xi^2, \dots, \xi^{q-3}\}$$ (for subgraph 0)



Diameter-2 Slim Fly

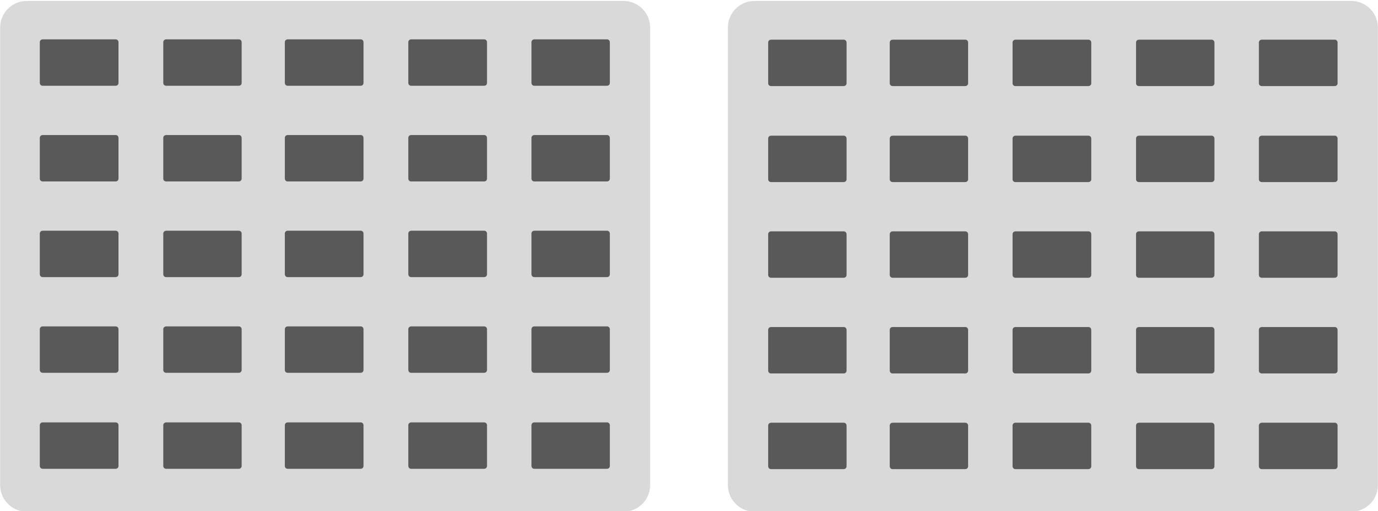
6 *Intra-group connections*

Two routers in one group are connected iff their “vertical Manhattan distance” is an element from:

$$X = \{1, \xi^2, \dots, \xi^{q-3}\} \text{ (for subgraph 0)}$$

$$X' = \{\xi, \xi^3, \dots, \xi^{q-2}\} \text{ (for subgraph 1)}$$

E *Example:* $q = 5$



Diameter-2 Slim Fly

6 *Intra-group connections*

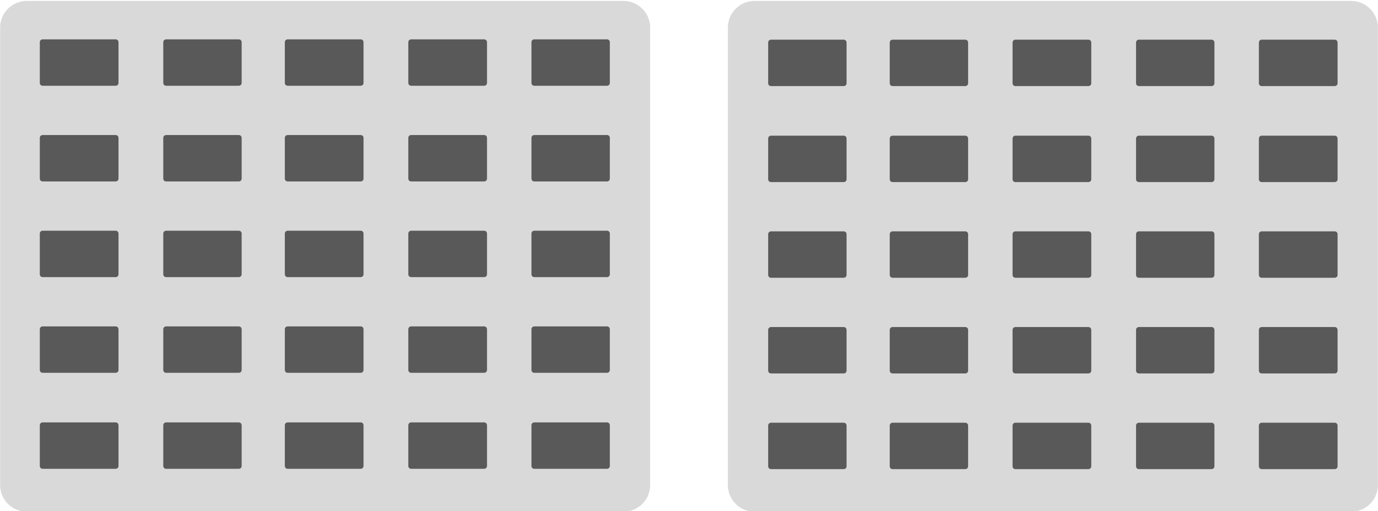
Two routers in one group are connected iff their "vertical Manhattan distance" is an element from:

$X = \{1, \xi^2, \ldots, \xi^{q-3}\}$ (for subgraph 0)

$X' = \{\xi, \xi^3, \ldots, \xi^{q-2}\}$ (for subgraph 1)

E *Example:* $q = 5$

Take Routers $(0,0,.)$



Diameter-2 Slim Fly

6 *Intra-group connections*

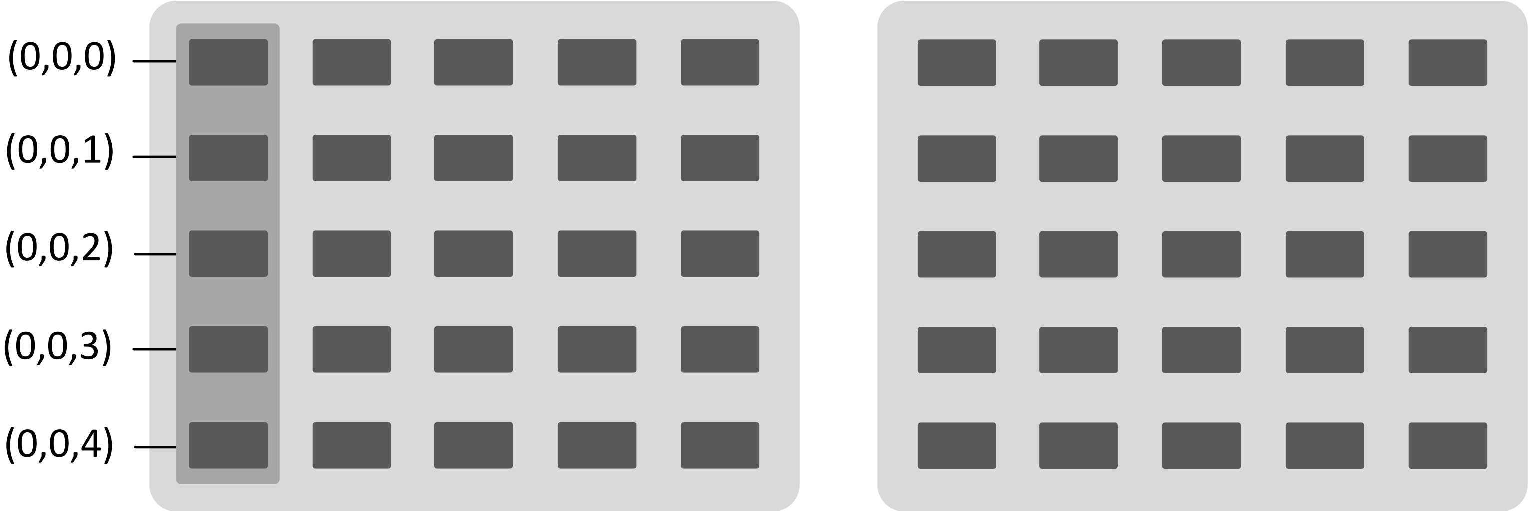
Two routers in one group are connected iff their "vertical Manhattan distance" is an element from:

$$X = \{1, \xi^2, \ldots, \xi^{q-3}\} \text{ (for subgraph 0)}$$

$$X' = \{\xi, \xi^3, \ldots, \xi^{q-2}\} \text{ (for subgraph 1)}$$

E *Example:* $q = 5$

Take Routers $(0,0,.)$



DIAMETER-2 SLIM FLY

6 *Intra-group connections*

Two routers in one group are connected iff their "vertical Manhattan distance" is an element from:

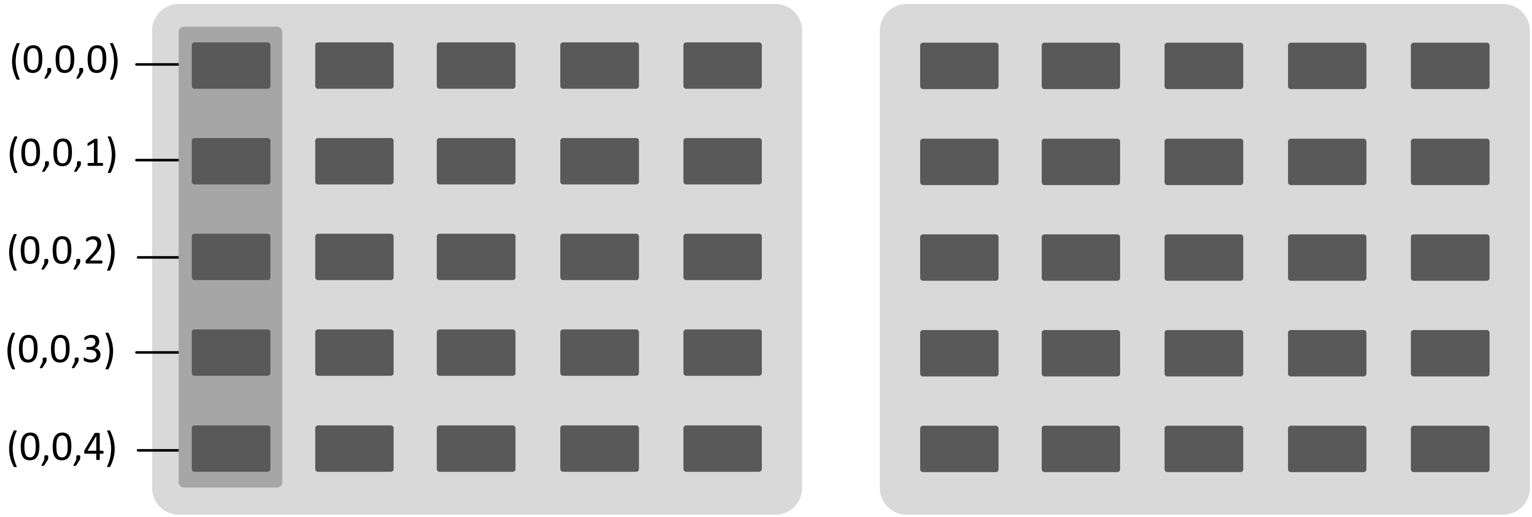
$X = \{1, \xi^2, \ldots, \xi^{q-3}\}$ (for subgraph 0)

$X' = \{\xi, \xi^3, \ldots, \xi^{q-2}\}$ (for subgraph 1)

E *Example:* $q = 5$

Take Routers $(0,0,.)$

$X = \{1,4\}$



Diameter-2 Slim Fly

6 *Intra-group connections*

Two routers in one group are connected iff their "vertical Manhattan distance" is an element from:

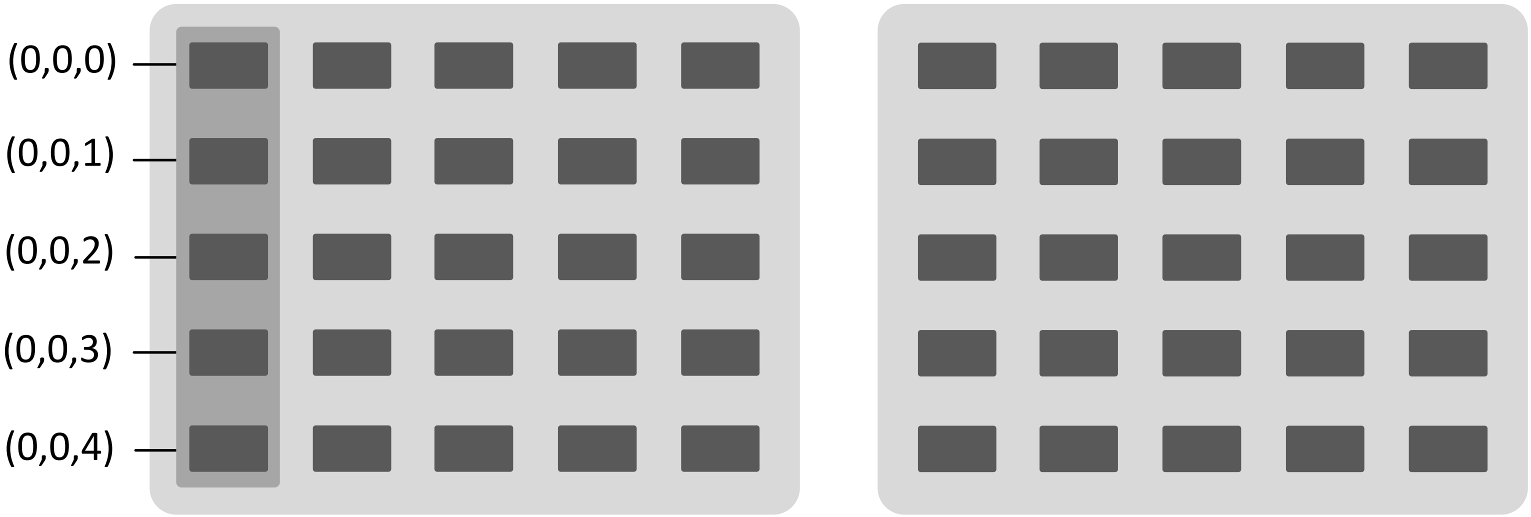
$$X = \{1, \xi^2, \ldots, \xi^{q-3}\} \text{ (for subgraph 0)}$$

$$X' = \{\xi, \xi^3, \ldots, \xi^{q-2}\} \text{ (for subgraph 1)}$$

E *Example:* $q = 5$

Take Routers $(0,0,.)$

$$X = \{1, 4\}$$



DIAMETER-2 SLIM FLY

6 *Intra-group connections*

Two routers in one group are connected iff their "vertical Manhattan distance" is an element from:

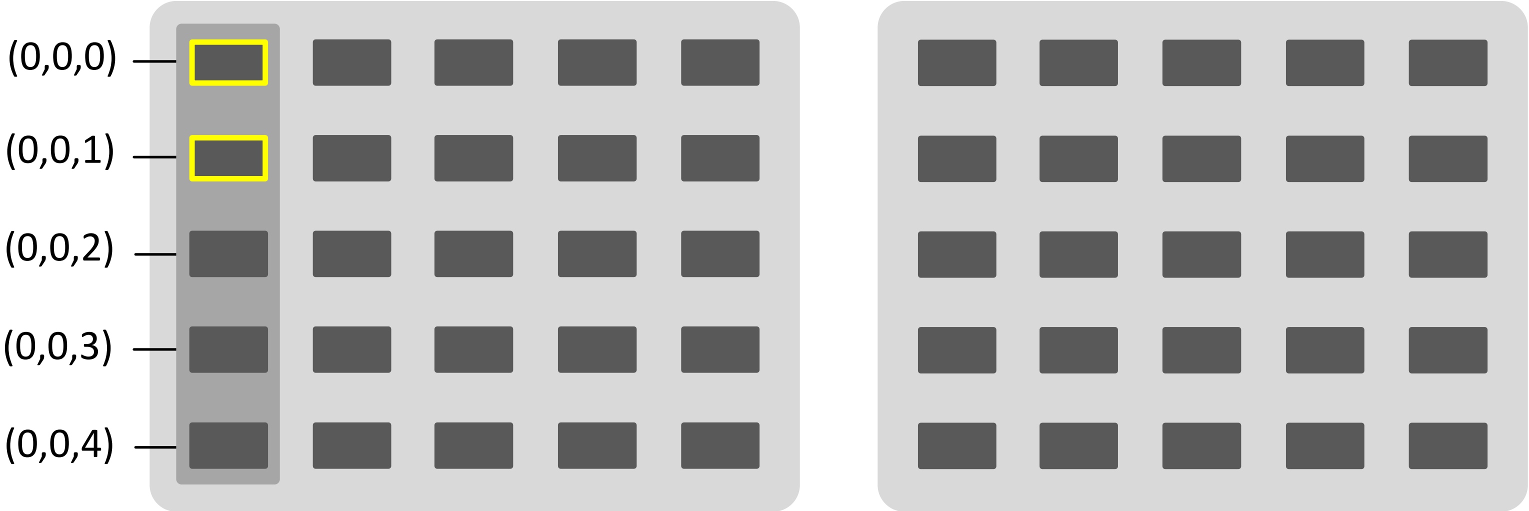
$$X = \{1, \xi^2, \dots, \xi^{q-3}\} \text{ (for subgraph 0)}$$

$$X' = \{\xi, \xi^3, \dots, \xi^{q-2}\} \text{ (for subgraph 1)}$$

E *Example:* $q = 5$

Take Routers $(0,0,.)$

$X = \{1, 4\}$



Diameter-2 Slim Fly

6 *Intra-group connections*

Two routers in one group are connected iff their "vertical Manhattan distance" is an element from:

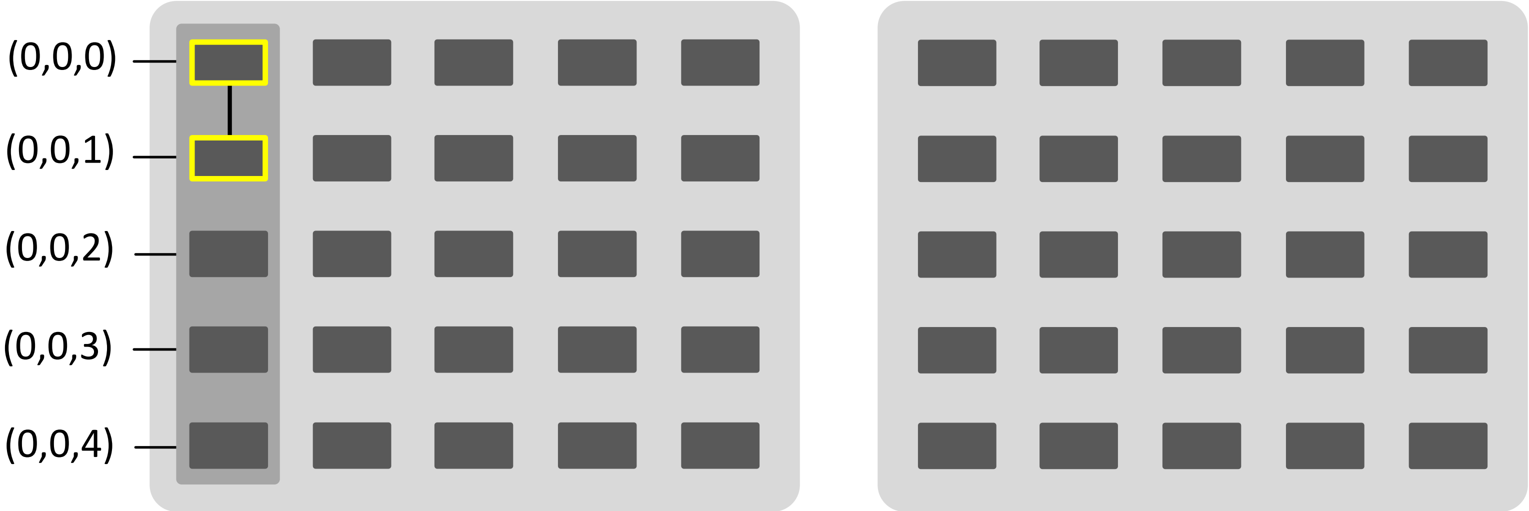
$$X = \{1, \xi^2, \ldots, \xi^{q-3}\} \text{ (for subgraph 0)}$$

$$X' = \{\xi, \xi^3, \ldots, \xi^{q-2}\} \text{ (for subgraph 1)}$$

E *Example:* $q = 5$

Take Routers $(0,0,.)$

$$X = \{1, 4\}$$



DIAMETER-2 SLIM FLY

6 *Intra-group connections*

Two routers in one group are connected iff their "vertical Manhattan distance" is an element from:

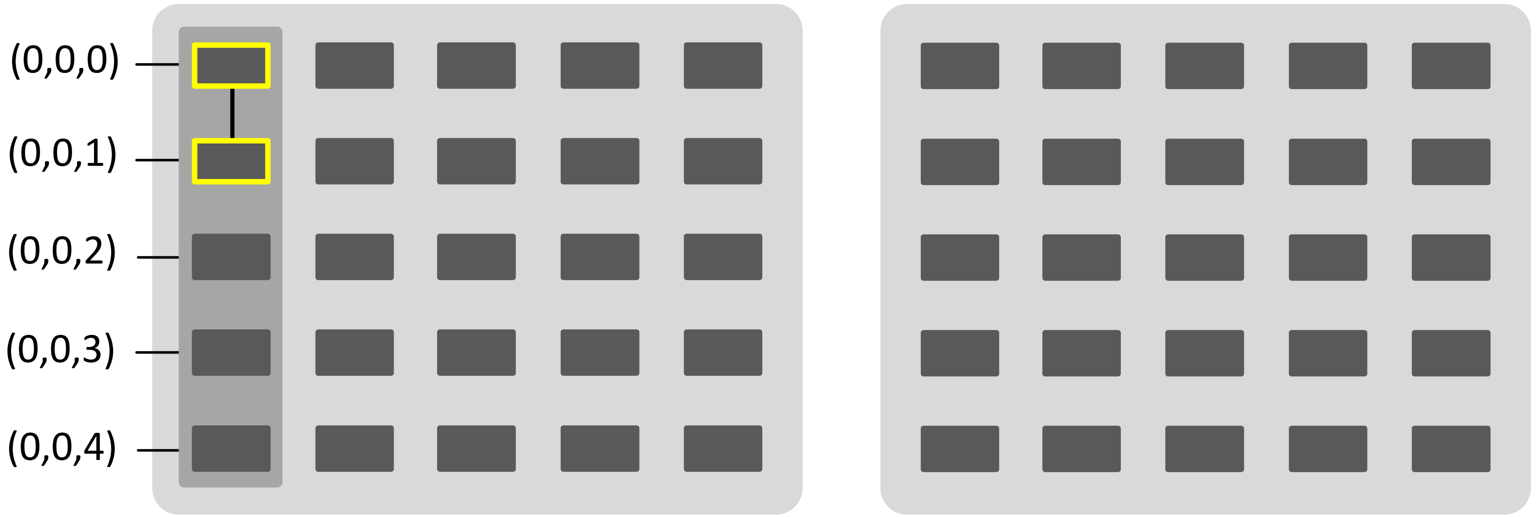
$$X = \{1, \xi^2, \dots, \xi^{q-3}\} \text{ (for subgraph 0)}$$

$$X' = \{\xi, \xi^3, \dots, \xi^{q-2}\} \text{ (for subgraph 1)}$$

E *Example:* $q = 5$

Take Routers $(0,0,.)$

$X = \{1, 4\}$



Diameter-2 Slim Fly

6 *Intra-group connections*

Two routers in one group are connected iff their "vertical Manhattan distance" is an element from:

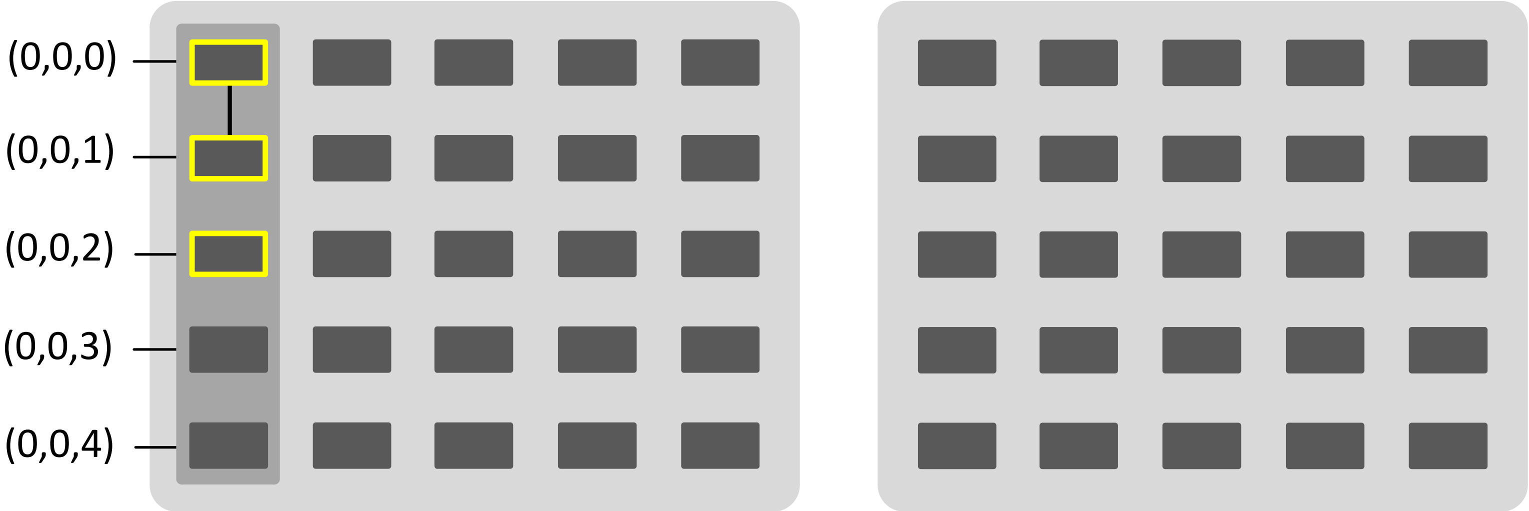
$X = \{1, \xi^2, \dots, \xi^{q-3}\}$ (for subgraph 0)

$X' = \{\xi, \xi^3, \dots, \xi^{q-2}\}$ (for subgraph 1)

E *Example:* $q = 5$

Take Routers $(0,0,.)$

$X = \{1, 4\}$



Diameter-2 Slim Fly

6 *Intra-group connections*

Two routers in one group are connected iff their "vertical Manhattan distance" is an element from:

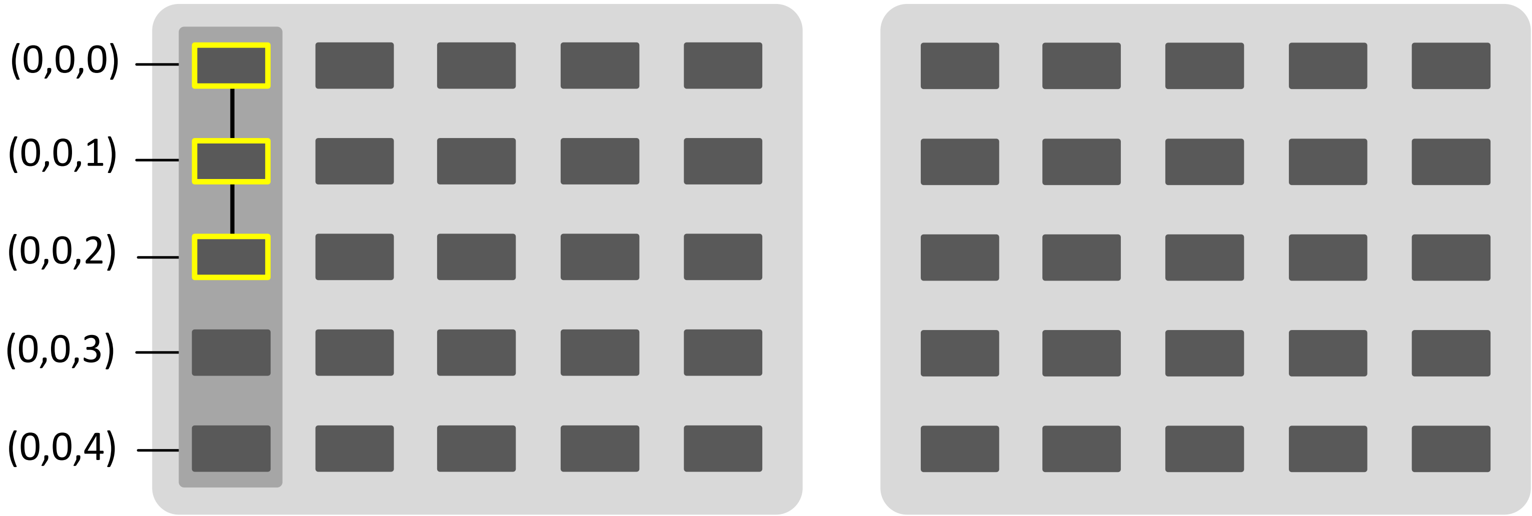
$$X = \{1, \xi^2, \dots, \xi^{q-3}\} \text{ (for subgraph 0)}$$

$$X' = \{\xi, \xi^3, \dots, \xi^{q-2}\} \text{ (for subgraph 1)}$$

E *Example:* $q = 5$

Take Routers $(0,0,.)$

$$X = \{1, 4\}$$



Diameter-2 Slim Fly

6 *Intra-group connections*

Two routers in one group are connected iff their "vertical Manhattan distance" is an element from:

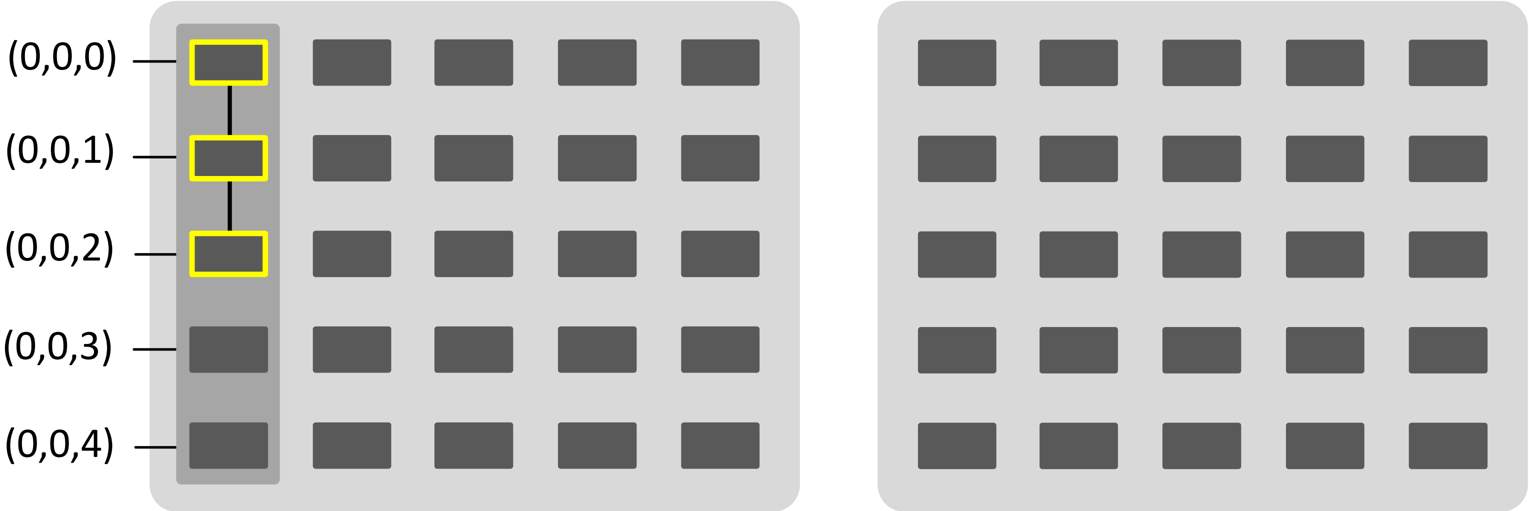
$$X = \{1, \xi^2, \ldots, \xi^{q-3}\} \text{ (for subgraph 0)}$$

$$X' = \{\xi, \xi^3, \ldots, \xi^{q-2}\} \text{ (for subgraph 1)}$$

E *Example:* $q = 5$

Take Routers $(0,0,.)$

$$X = \{1, 4\}$$



Diameter-2 Slim Fly

6 *Intra-group connections*

Two routers in one group are connected iff their “vertical Manhattan distance” is an element from:

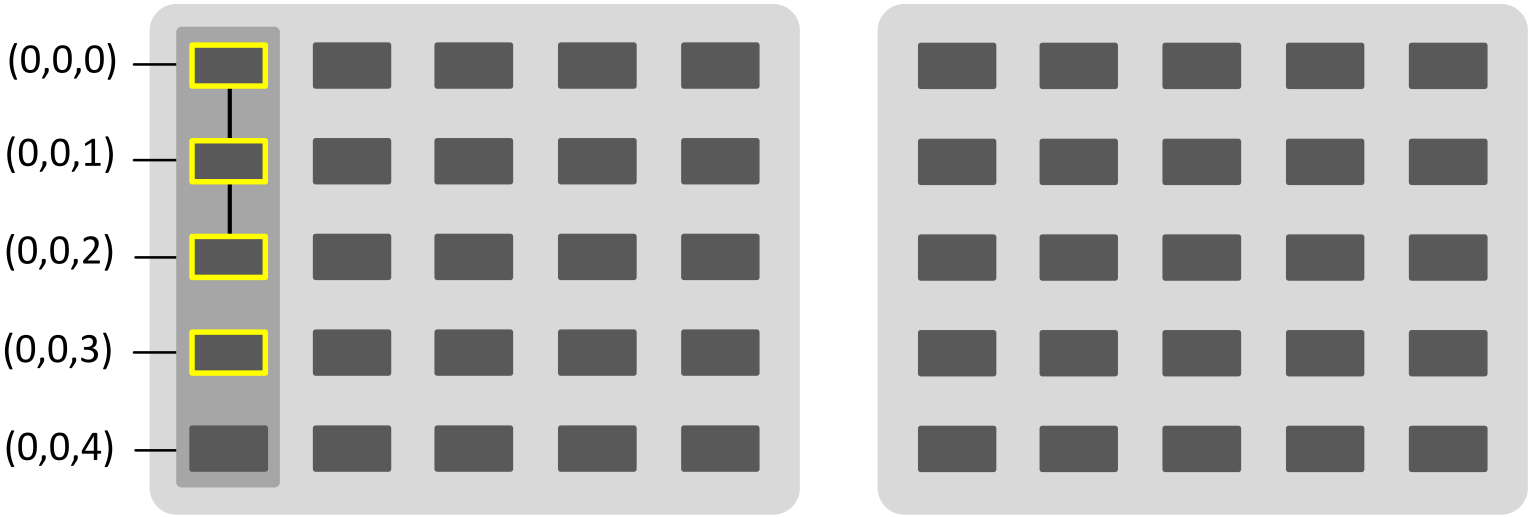
$$X = \{1, \xi^2, \dots, \xi^{q-3}\} \text{ (for subgraph 0)}$$

$$X' = \{\xi, \xi^3, \dots, \xi^{q-2}\} \text{ (for subgraph 1)}$$

E *Example:* $q = 5$

Take Routers $(0,0,.)$

$X = \{1, 4\}$



Diameter-2 Slim Fly

6 *Intra-group connections*

Two routers in one group are connected iff their "vertical Manhattan distance" is an element from:

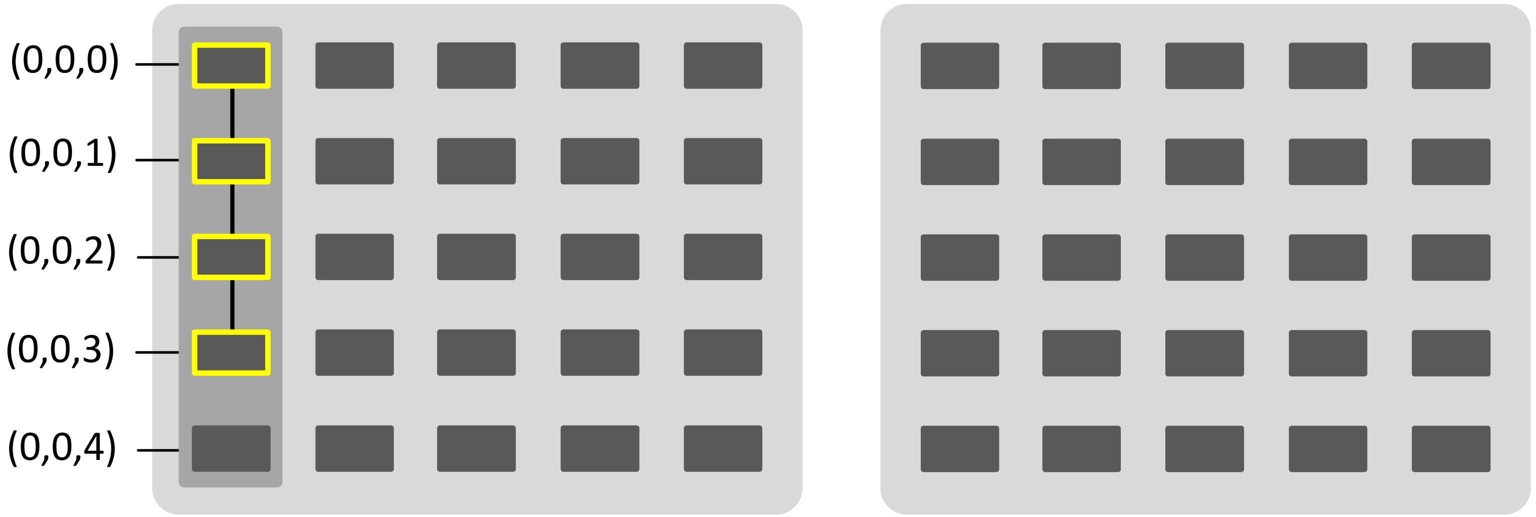
$$X = \{1, \xi^2, \ldots, \xi^{q-3}\}$$ (for subgraph 0)

$$X' = \{\xi, \xi^3, \ldots, \xi^{q-2}\}$$ (for subgraph 1)

E *Example:* $q = 5$

Take Routers $(0,0,.)$

$$X = \{1, 4\}$$



Diameter-2 Slim Fly

6 *Intra-group connections*

Two routers in one group are connected iff their "vertical Manhattan distance" is an element from:

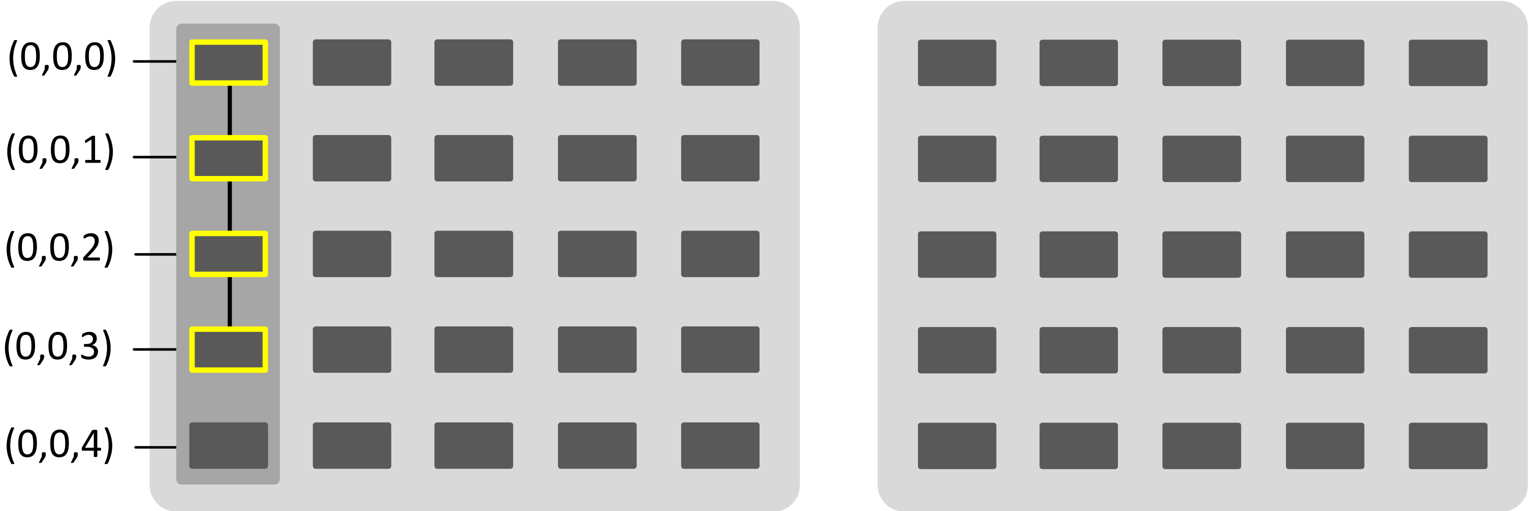
$$X = \{1, \xi^2, \ldots, \xi^{q-3}\} \text{ (for subgraph 0)}$$

$$X' = \{\xi, \xi^3, \ldots, \xi^{q-2}\} \text{ (for subgraph 1)}$$

E *Example:* $q = 5$

Take Routers $(0,0,.)$

$X = \{1, 4\}$



Diameter-2 Slim Fly

6 *Intra-group connections*

Two routers in one group are connected iff their "vertical Manhattan distance" is an element from:

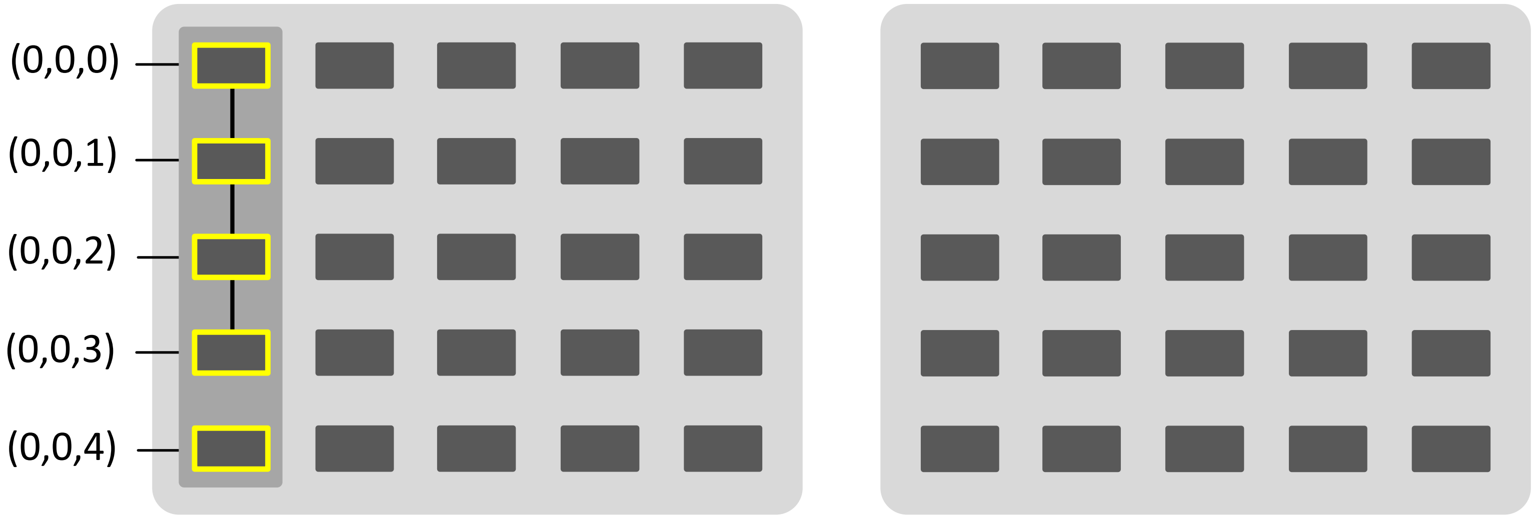
$X = \{1, \xi^2, \dots, \xi^{q-3}\}$ (for subgraph 0)

$X' = \{\xi, \xi^3, \dots, \xi^{q-2}\}$ (for subgraph 1)

E *Example:* $q = 5$

Take Routers $(0,0,.)$

$X = \{1, 4\}$



Diameter-2 Slim Fly

6 *Intra-group connections*

Two routers in one group are connected iff their "vertical Manhattan distance" is an element from:

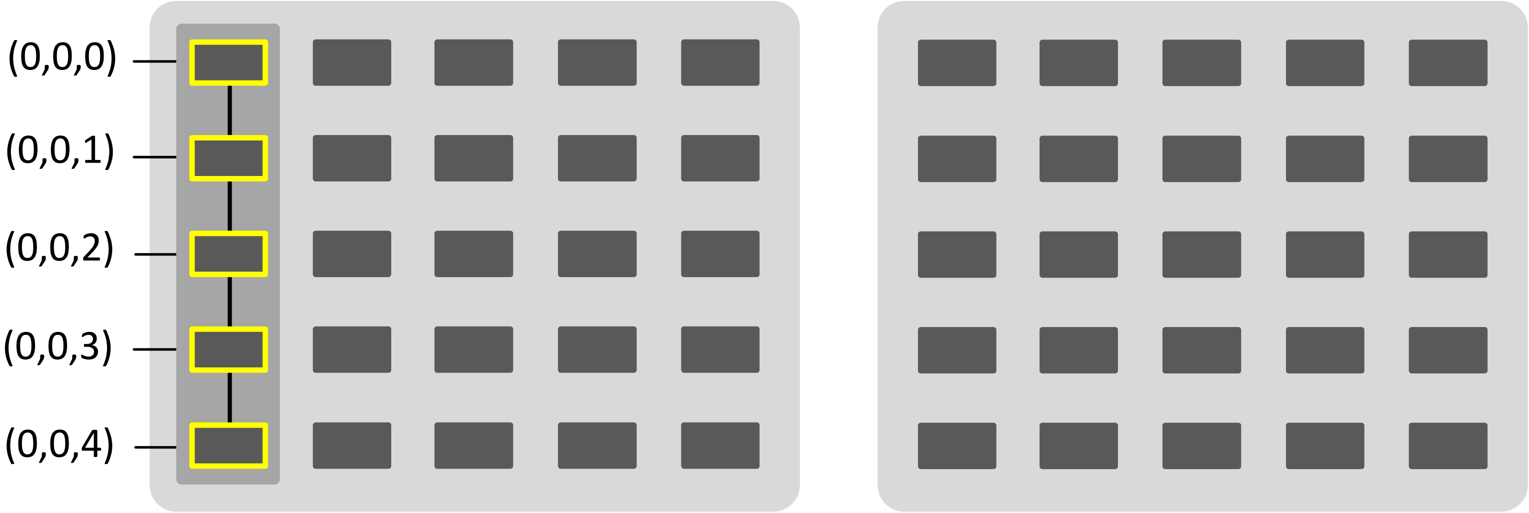
$X = \{1, \xi^2, \dots, \xi^{q-3}\}$ (for subgraph 0)

$X' = \{\xi, \xi^3, \dots, \xi^{q-2}\}$ (for subgraph 1)

E *Example:* $q = 5$

Take Routers $(0,0,.)$

$X = \{1, 4\}$



Diameter-2 Slim Fly

6 *Intra-group connections*

Two routers in one group are connected iff their "vertical Manhattan distance" is an element from:

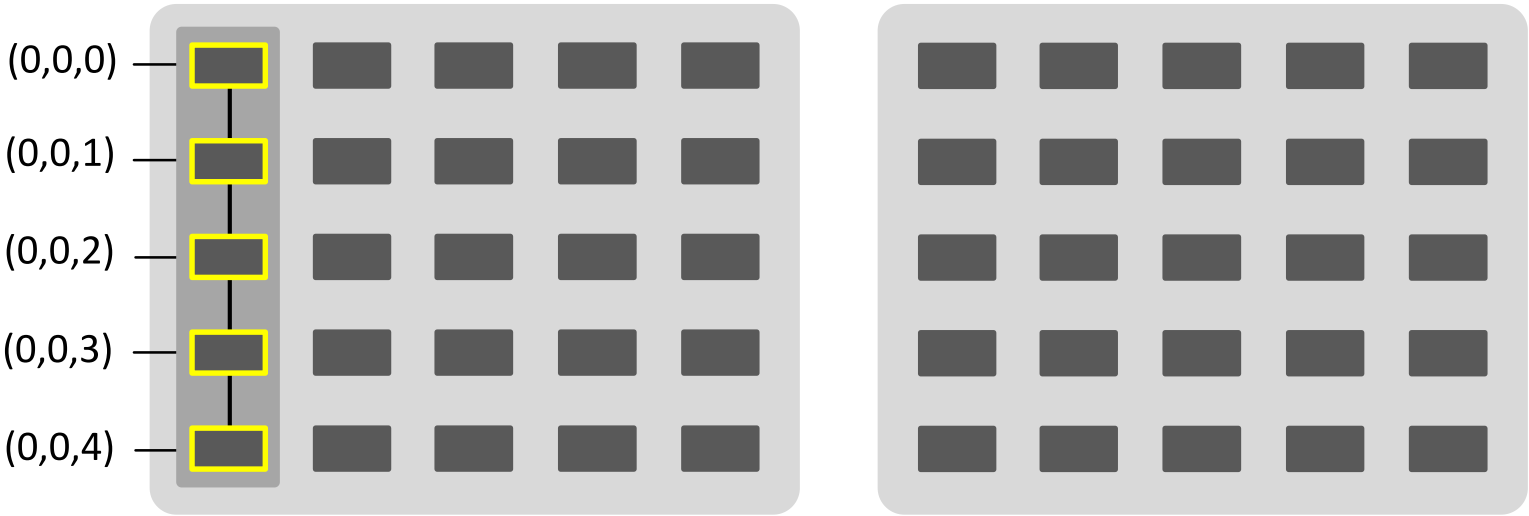
$$X = \{1, \xi^2, \dots, \xi^{q-3}\} \text{ (for subgraph 0)}$$

$$X' = \{\xi, \xi^3, \dots, \xi^{q-2}\} \text{ (for subgraph 1)}$$

E *Example:* $q = 5$

Take Routers $(0,0,.)$

$X = \{1, 4\}$



Diameter-2 Slim Fly

6 *Intra-group connections*

Two routers in one group are connected iff their "vertical Manhattan distance" is an element from:

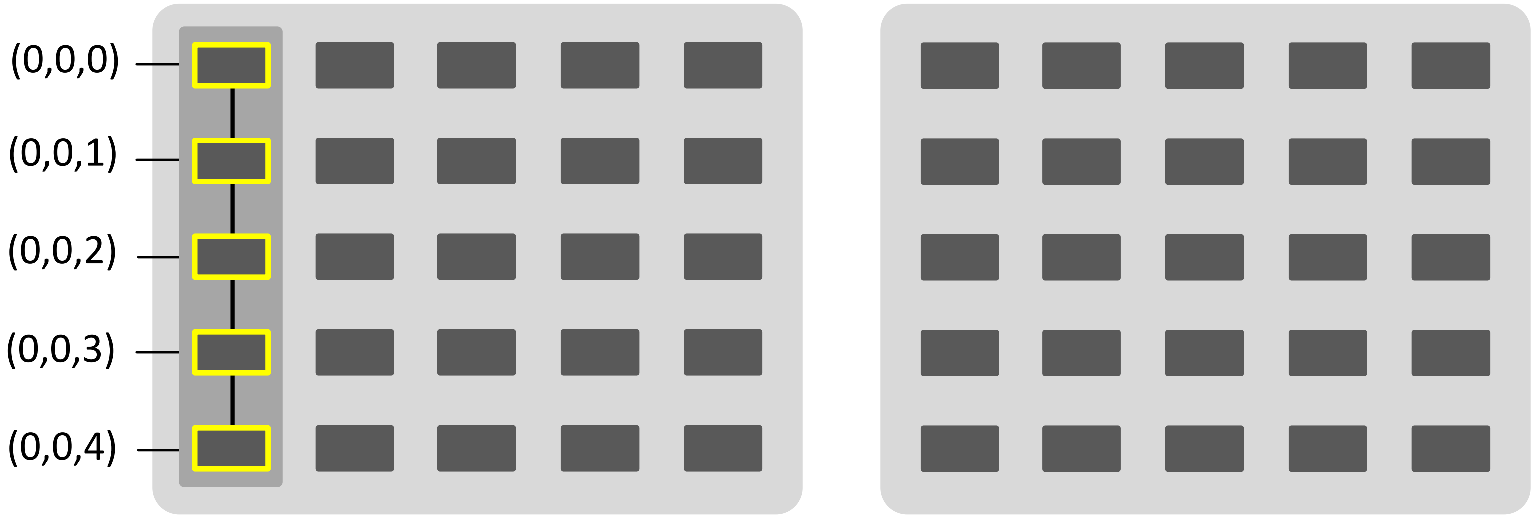
$X = \{1, \xi^2, \dots, \xi^{q-3}\}$ (for subgraph 0)

$X' = \{\xi, \xi^3, \dots, \xi^{q-2}\}$ (for subgraph 1)

E *Example:* $q = 5$

Take Routers $(0,0,.)$

$X = \{1,4\}$



Diameter-2 Slim Fly

6 *Intra-group connections*

Two routers in one group are connected iff their "vertical Manhattan distance" is an element from:

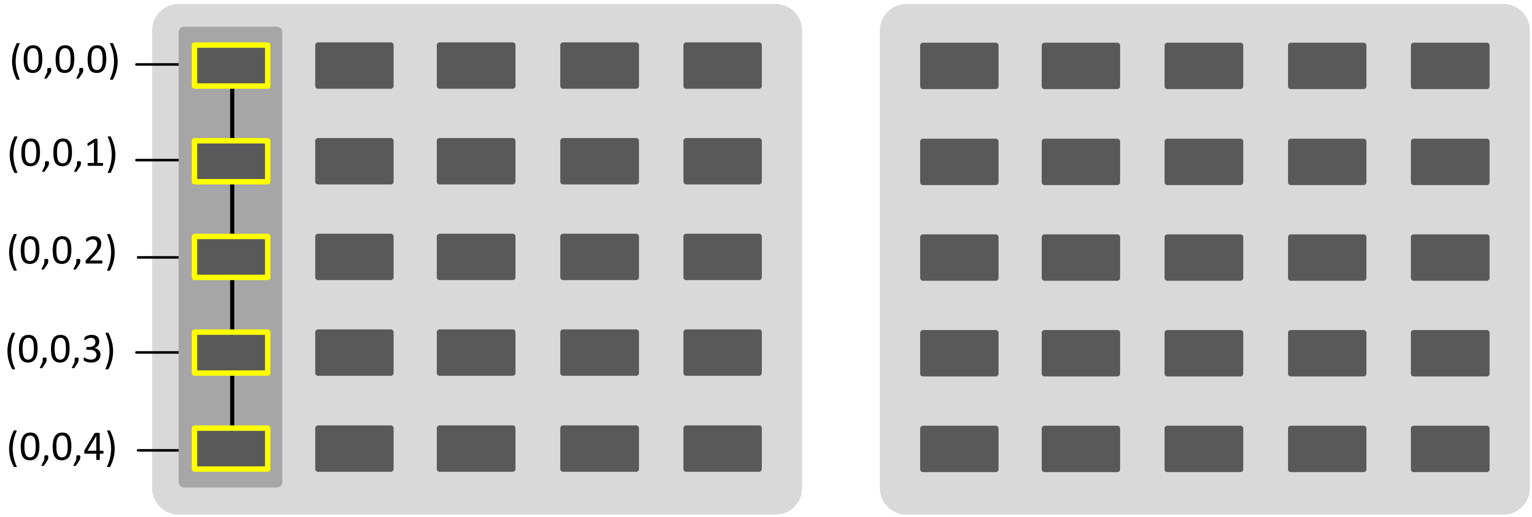
$X = \{1, \xi^2, \ldots, \xi^{q-3}\}$ (for subgraph 0)

$X' = \{\xi, \xi^3, \ldots, \xi^{q-2}\}$ (for subgraph 1)

E *Example:* $q = 5$

Take Routers $(0,0,.)$

$X = \{1, 4\}$



Diameter-2 Slim Fly

6 *Intra-group connections*

Two routers in one group are connected iff their "vertical Manhattan distance" is an element from:

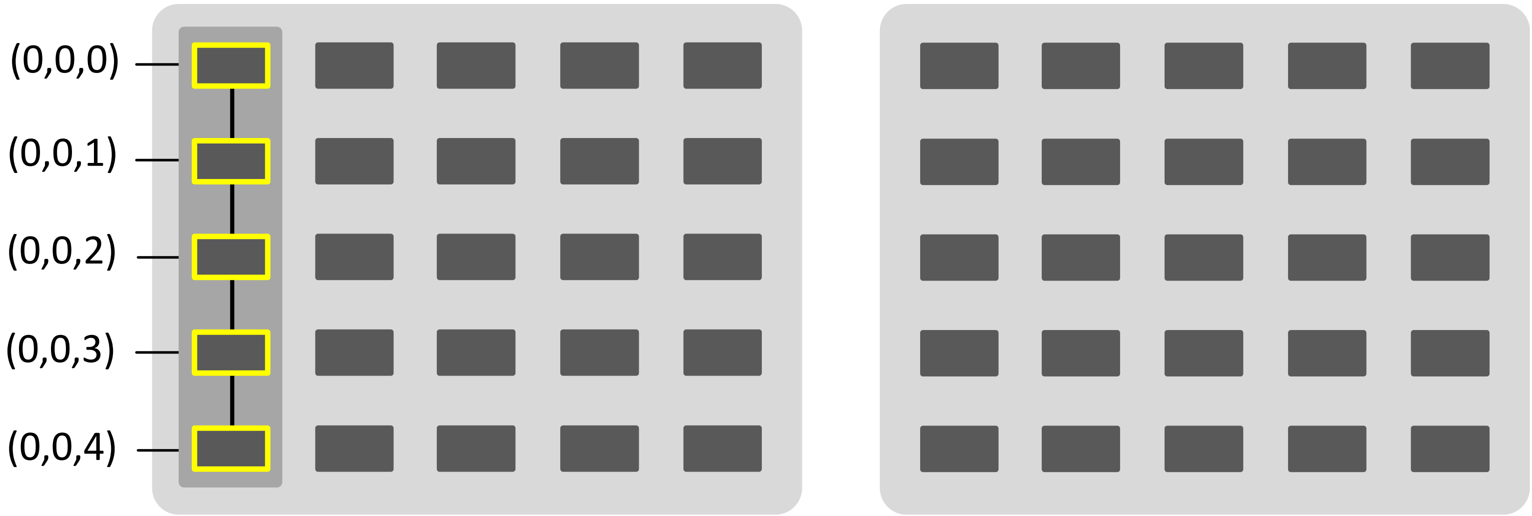
$$X = \{1, \xi^2, \dots, \xi^{q-3}\} \text{ (for subgraph 0)}$$

$$X' = \{\xi, \xi^3, \dots, \xi^{q-2}\} \text{ (for subgraph 1)}$$

E *Example:* $q = 5$

Take Routers $(0,0,.)$

$$X = \{1, 4\}$$



Diameter-2 Slim Fly

6 *Intra-group connections*

Two routers in one group are connected iff their "vertical Manhattan distance" is an element from:

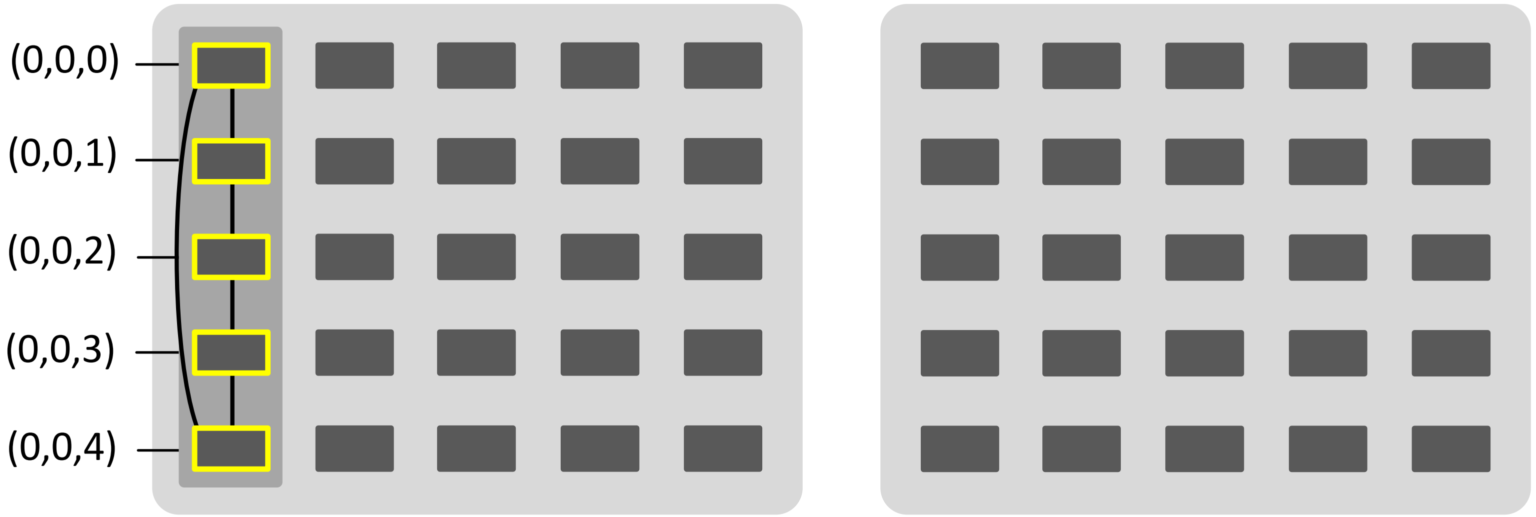
$X = \{1, \xi^2, \ldots, \xi^{q-3}\}$ (for subgraph 0)

$X' = \{\xi, \xi^3, \ldots, \xi^{q-2}\}$ (for subgraph 1)

E *Example:* $q = 5$

Take Routers $(0,0,.)$

$X = \{1, 4\}$



Diameter-2 Slim Fly

6 *Intra-group connections*

Two routers in one group are connected iff their "vertical Manhattan distance" is an element from:

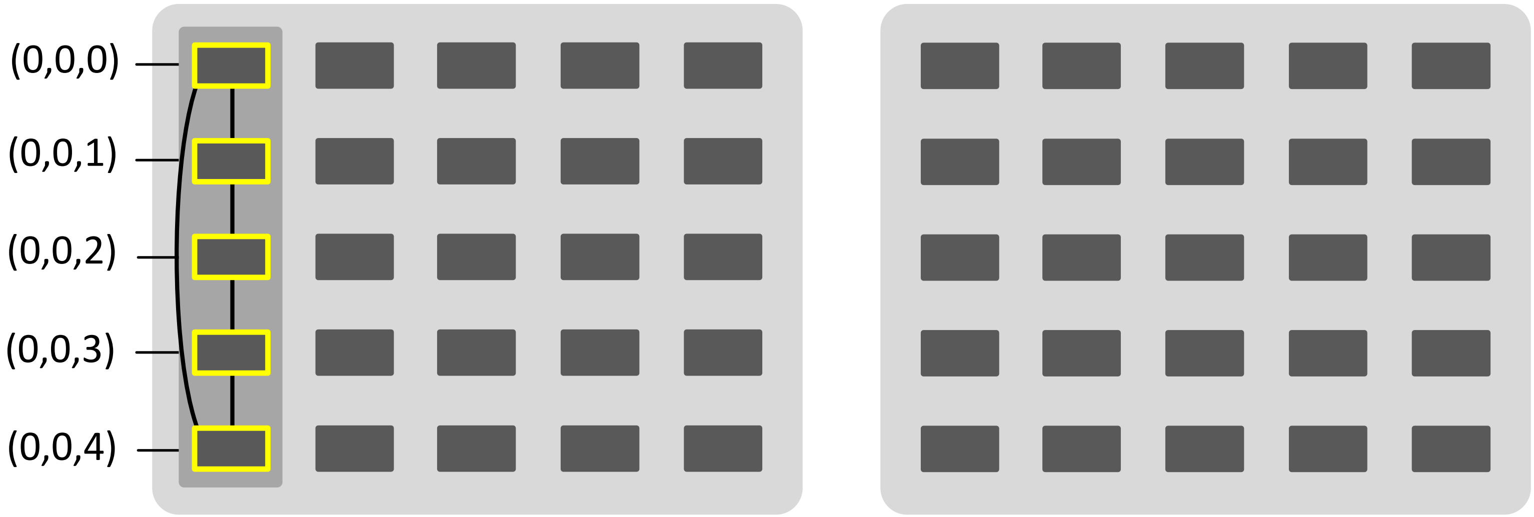
$X = \{1, \xi^2, \ldots, \xi^{q-3}\}$ (for subgraph 0)

$X' = \{\xi, \xi^3, \ldots, \xi^{q-2}\}$ (for subgraph 1)

E *Example:* $q = 5$

Take Routers $(0,0,.)$

$X = \{1, 4\}$



Diameter-2 Slim Fly

6 *Intra-group connections*

Two routers in one group are connected iff their "vertical Manhattan distance" is an element from:

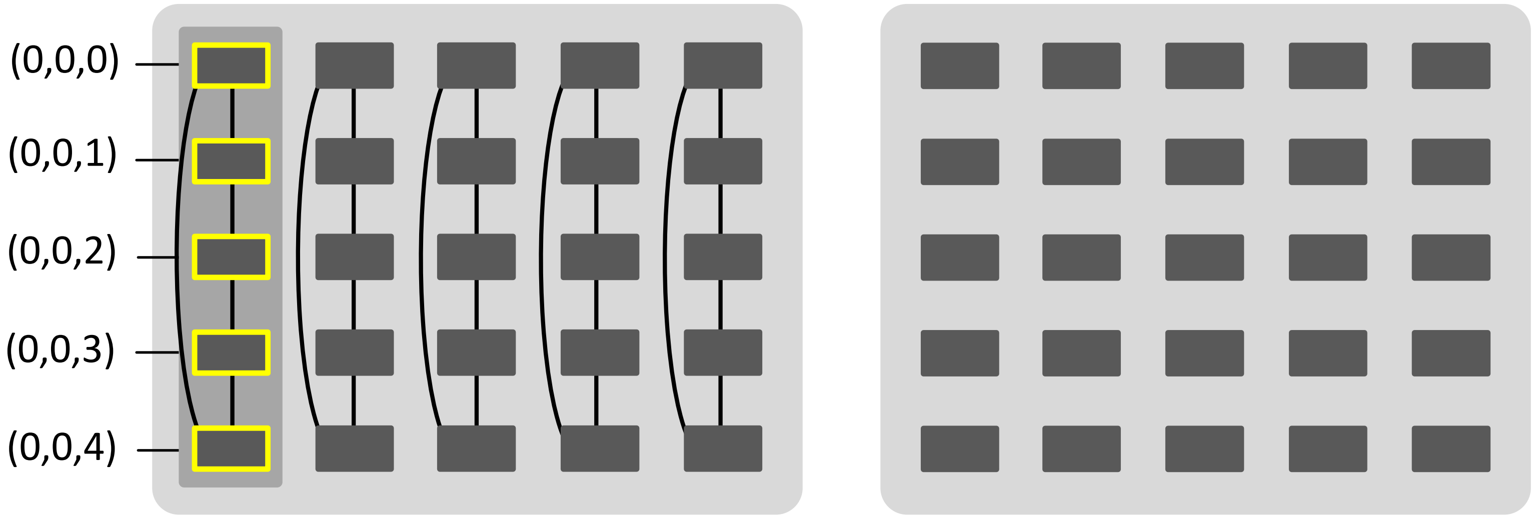
$$X = \{1, \xi^2, \dots, \xi^{q-3}\} \text{ (for subgraph 0)}$$

$$X' = \{\xi, \xi^3, \dots, \xi^{q-2}\} \text{ (for subgraph 1)}$$

E *Example:* $q = 5$

Take Routers $(0,0,.)$

$X = \{1, 4\}$



Diameter-2 Slim Fly

6 *Intra-group connections*

Two routers in one group are connected iff their "vertical Manhattan distance" is an element from:

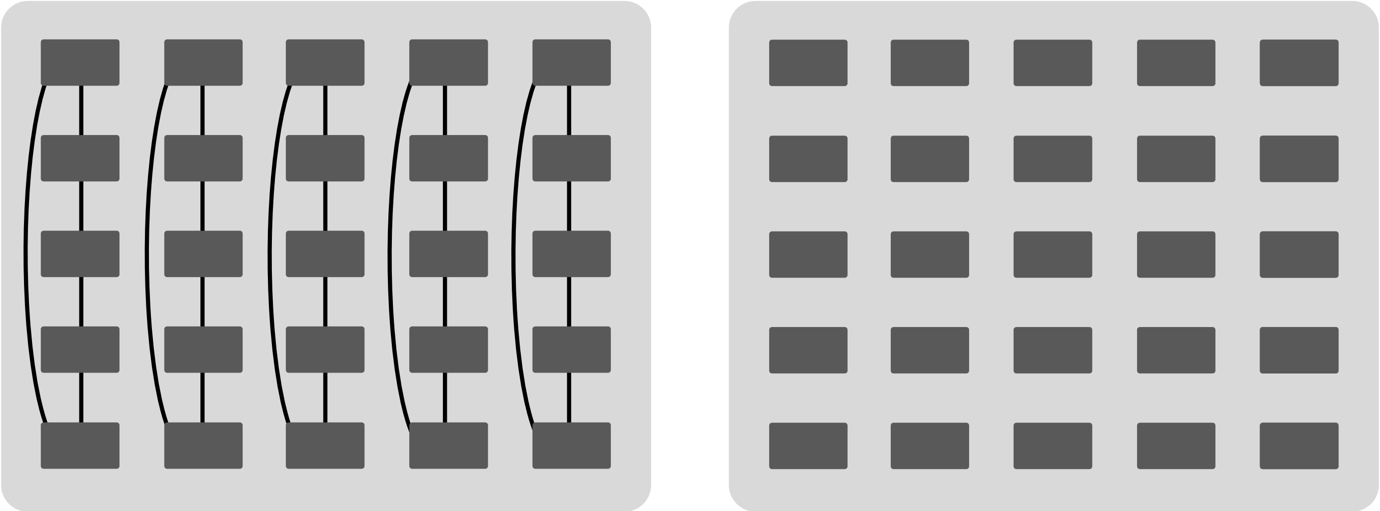
$$X = \{1, \xi^2, \dots, \xi^{q-3}\} \text{ (for subgraph 0)}$$

$$X' = \{\xi, \xi^3, \dots, \xi^{q-2}\} \text{ (for subgraph 1)}$$

E *Example:* $q = 5$

Take Routers $(1,4,.)$

$X' = \{2,3\}$



Diameter-2 Slim Fly

6 *Intra-group connections*

Two routers in one group are connected iff their “vertical Manhattan distance” is an element from:

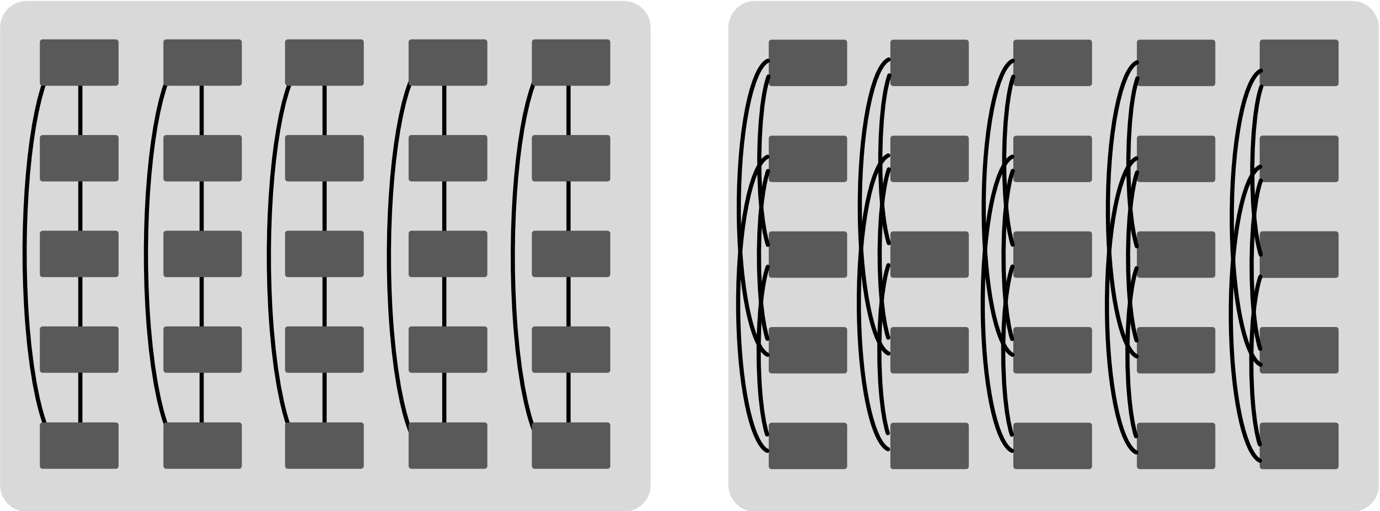
$X = \{1, \xi^2, \dots, \xi^{q-3}\}$ (for subgraph 0)

$X' = \{\xi, \xi^3, \dots, \xi^{q-2}\}$ (for subgraph 1)

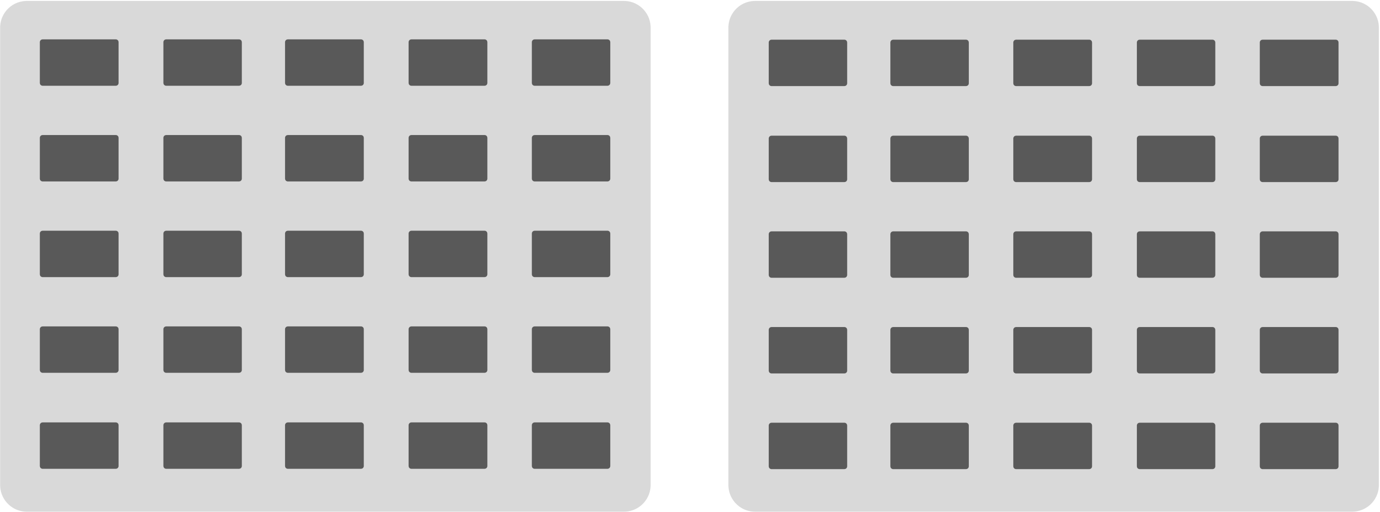
E *Example:* $q = 5$

Take Routers $(1,4,.)$

$X' = \{2,3\}$

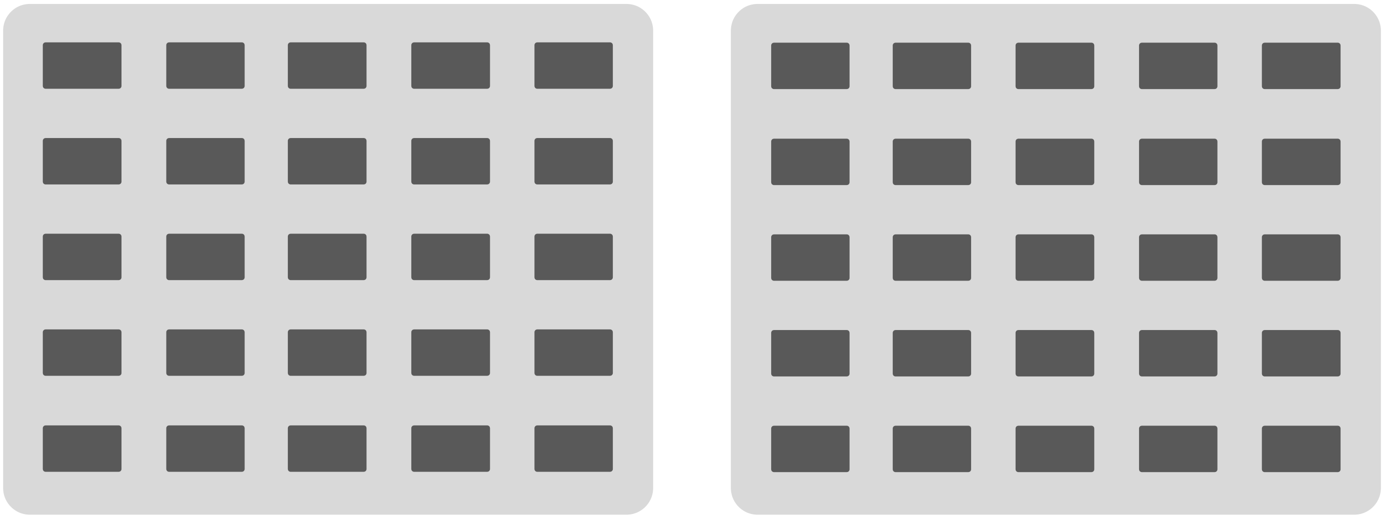


DIAMETER-2 SLIM FLY



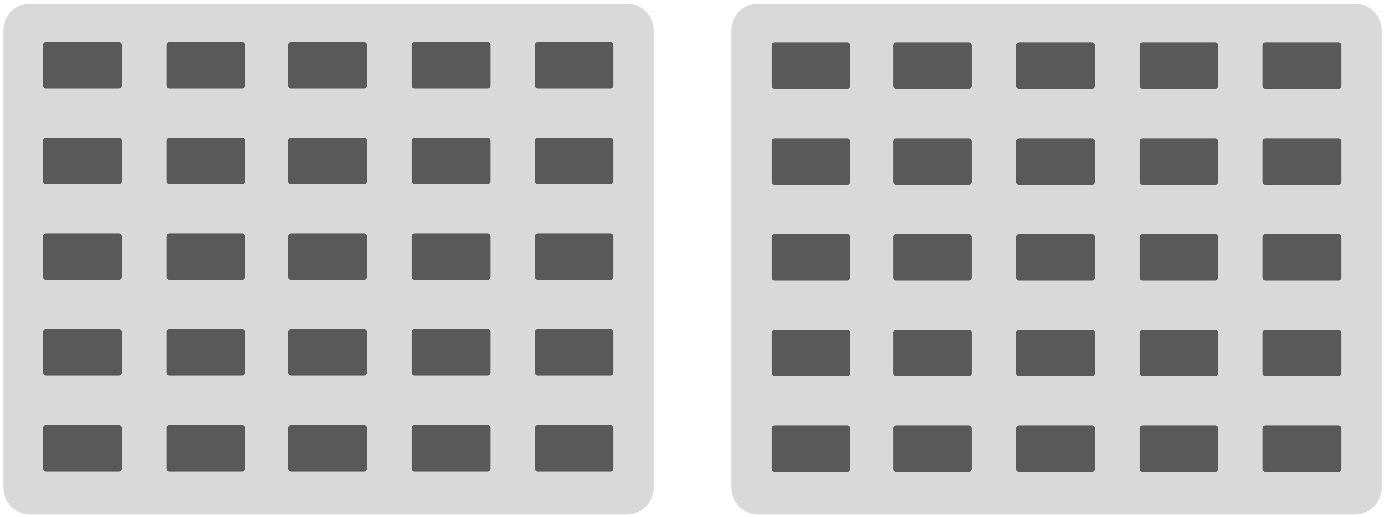
DIAMETER-2 SLIM FLY

7 *Inter-group connections*



DIAMETER-2 SLIM FLY

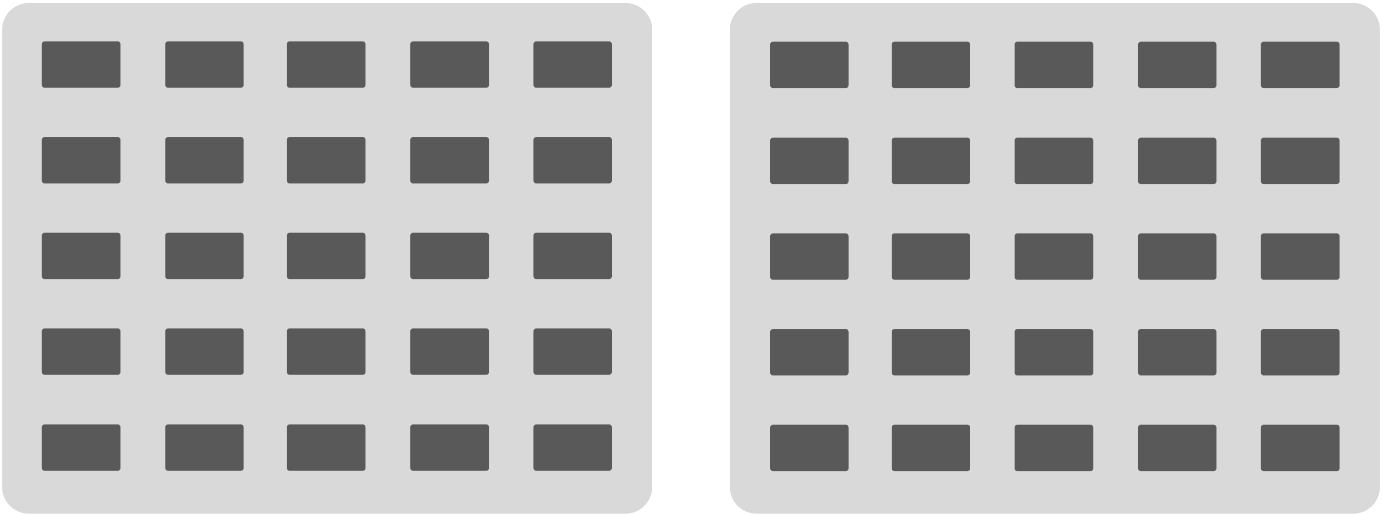
7 *Inter-group connections*



Diameter-2 Slim Fly

7 *Inter-group connections*

Router $(0, x, y) \leftrightarrow (1, m, c)$

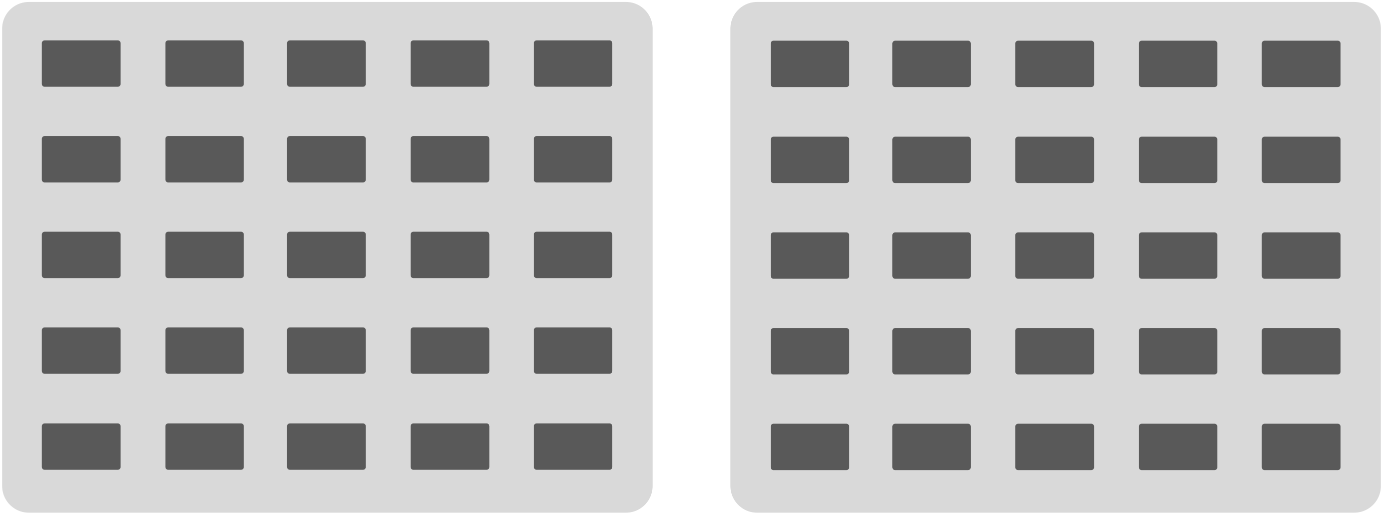


Diameter-2 Slim Fly

7 *Inter-group connections*

Router $(0, x, y) \leftrightarrow (1, m, c)$

iff $y = mx + c$



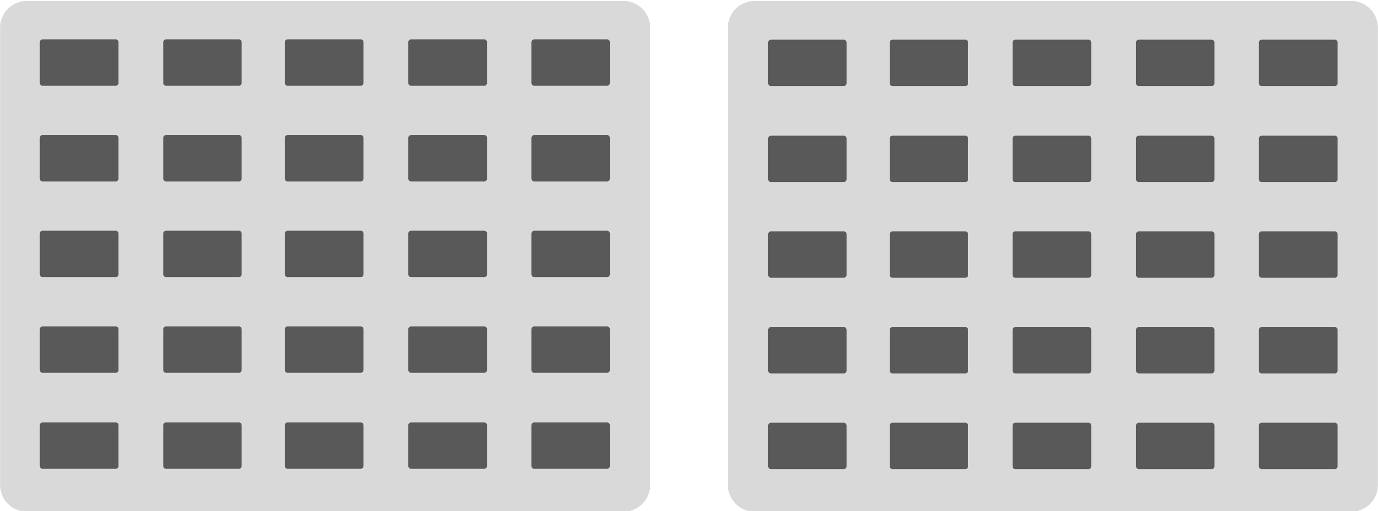
Diameter-2 Slim Fly

7 *Inter-group connections*

Router $(0, x, y) \leftrightarrow (1, m, c)$

iff $y = mx + c$

E *Example:* $q = 5$



Diameter-2 Slim Fly

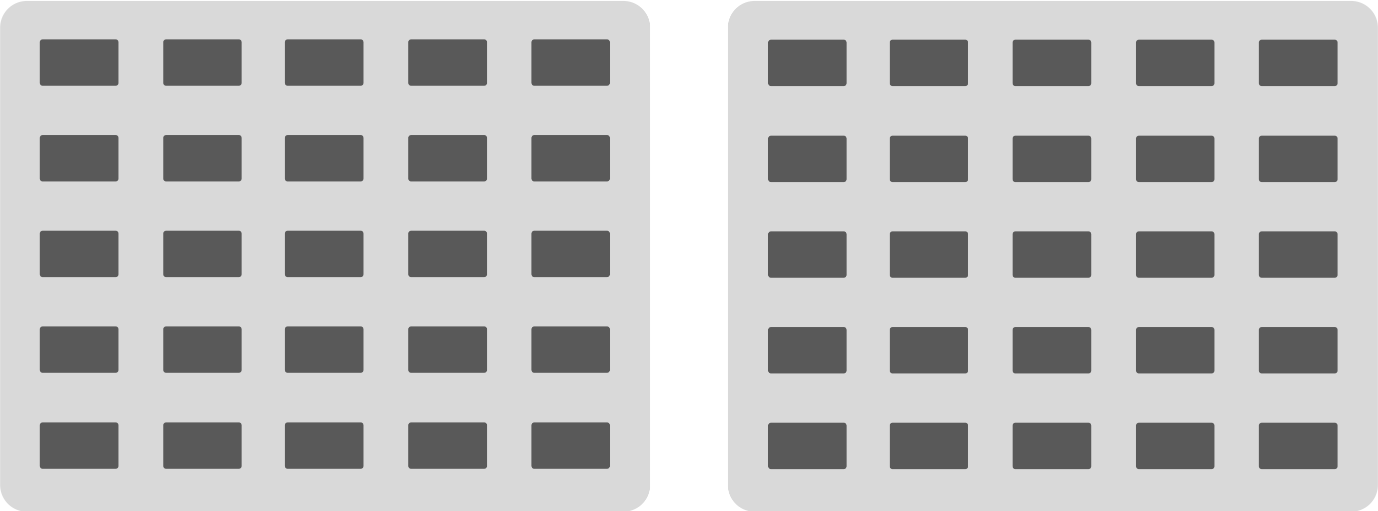
7 *Inter-group connections*

Router $(0, x, y) \leftrightarrow (1, m, c)$

iff $y = mx + c$

E *Example:* $q = 5$

Take Router $(1,0,0)$



Diameter-2 Slim Fly

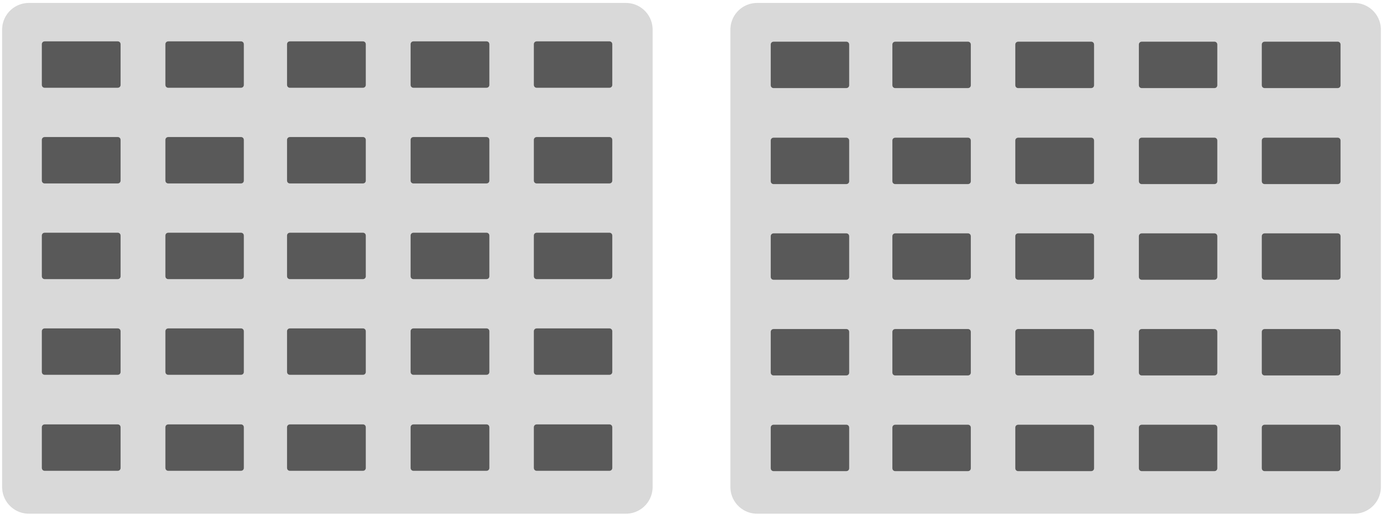
7 *Inter-group connections*

Router $(0, x, y) \leftrightarrow (1, m, c)$

iff $y = mx + c$

E *Example:* $q = 5$

Take Router $(1,0,0)$



Diameter-2 Slim Fly

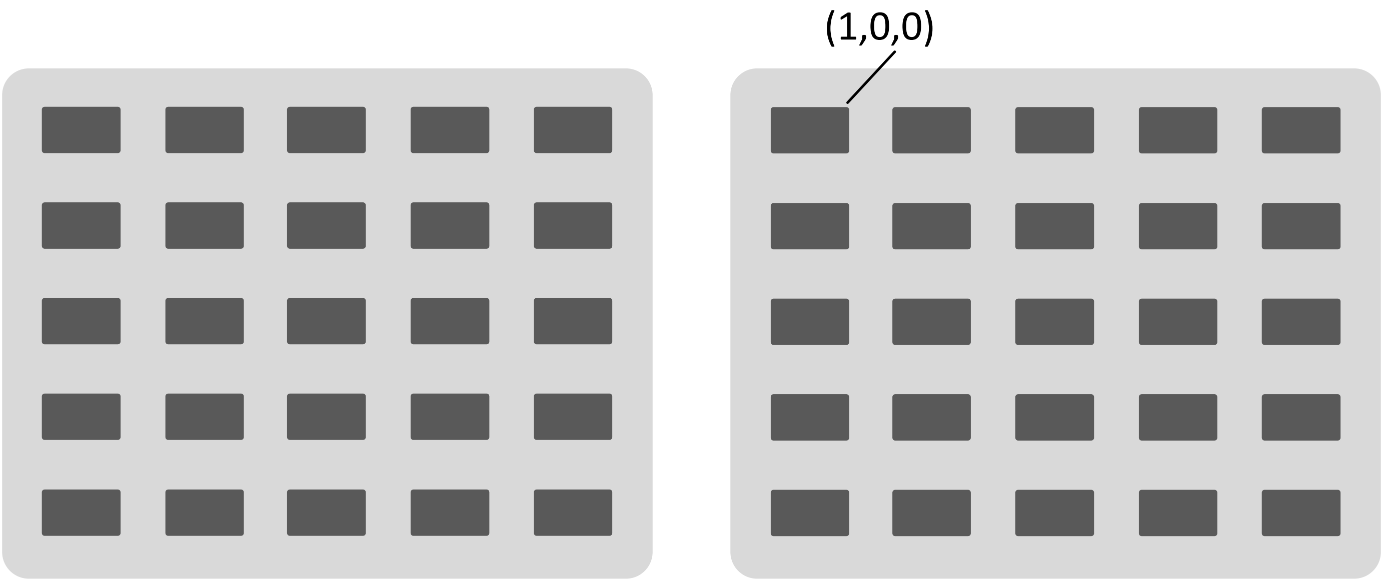
E *Example:* $q = 5$

Take Router $(1,0,0)$

7 *Inter-group connections*

Router $(0, x, y) \leftrightarrow (1, m, c)$

iff $y = mx + c$



Diameter-2 Slim Fly

7 *Inter-group connections*

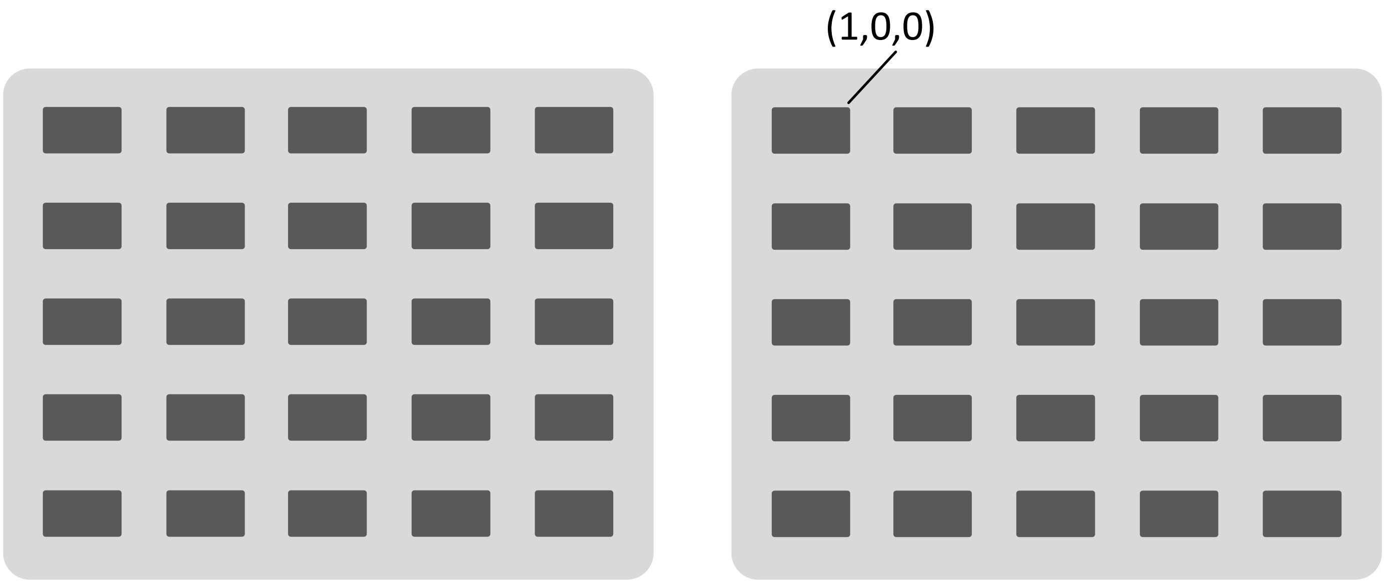
Router $(0, x, y) \leftrightarrow (1, m, c)$

iff $y = mx + c$

E *Example:* $q = 5$

$m = 0, c = 0$

Take Router $(1,0,0)$



Diameter-2 Slim Fly

7 *Inter-group connections*

Router $(0, x, y) \leftrightarrow (1, m, c)$

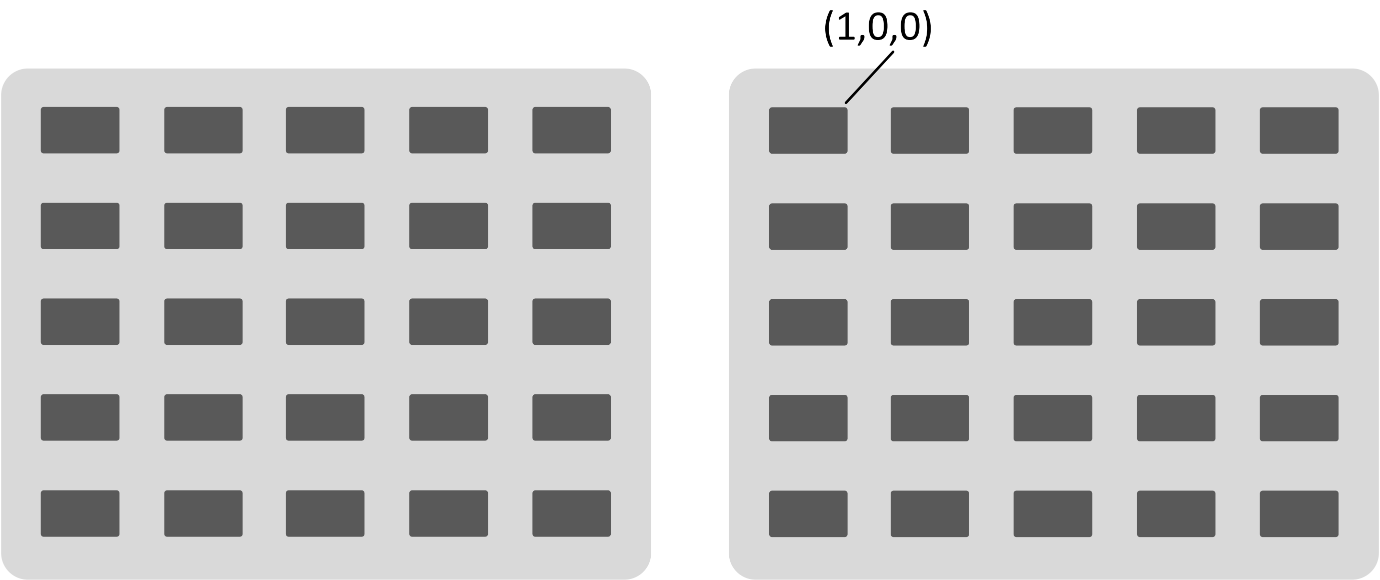
iff $y = mx + c$

E *Example:* $q = 5$

$m = 0, c = 0$

Take Router $(1,0,0)$

$(1,0,0) \leftrightarrow (0, x, 0)$



Diameter-2 Slim Fly

7 *Inter-group connections*

Router $(0, x, y) \leftrightarrow (1, m, c)$

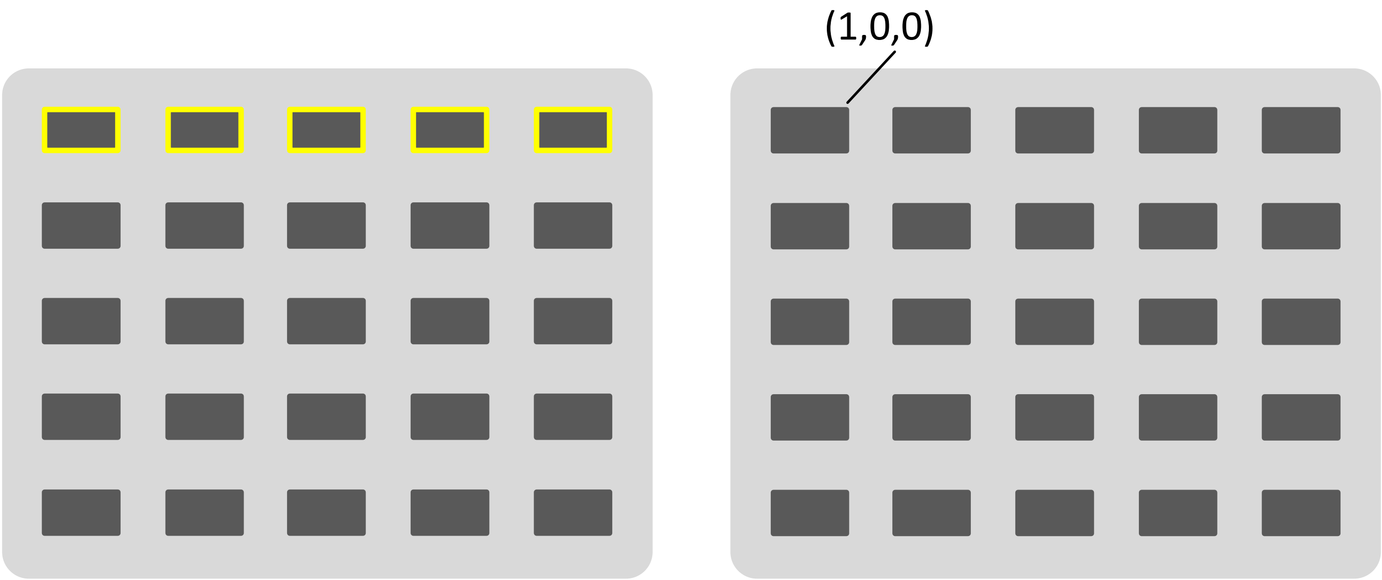
iff $y = mx + c$

E *Example:* $q = 5$

$m = 0, c = 0$

Take Router $(1,0,0)$

$(1,0,0) \leftrightarrow (0, x, 0)$



Diameter-2 Slim Fly

7 *Inter-group connections*

Router $(0, x, y) \leftrightarrow (1, m, c)$

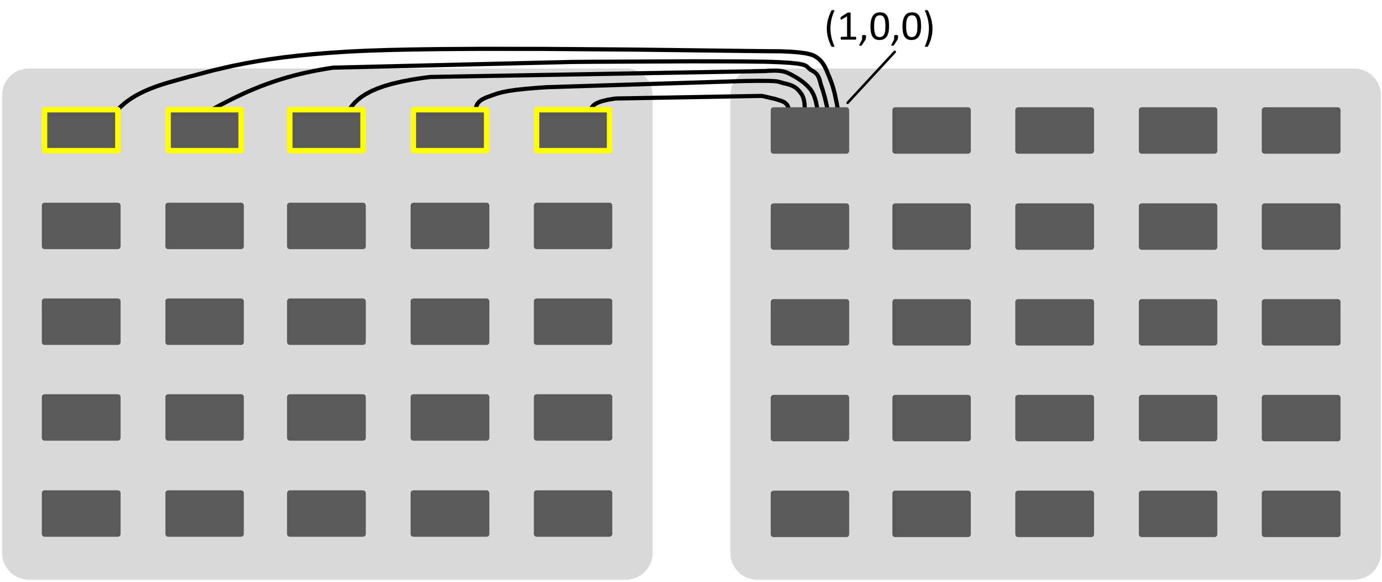
iff $y = mx + c$

E *Example:* $q = 5$

$m = 0, c = 0$

Take Router $(1{,}0{,}0)$

$(1{,}0{,}0) \leftrightarrow (0, x, 0)$



Diameter-2 Slim Fly

7 *Inter-group connections*

Router $(0, x, y) \leftrightarrow (1, m, c)$

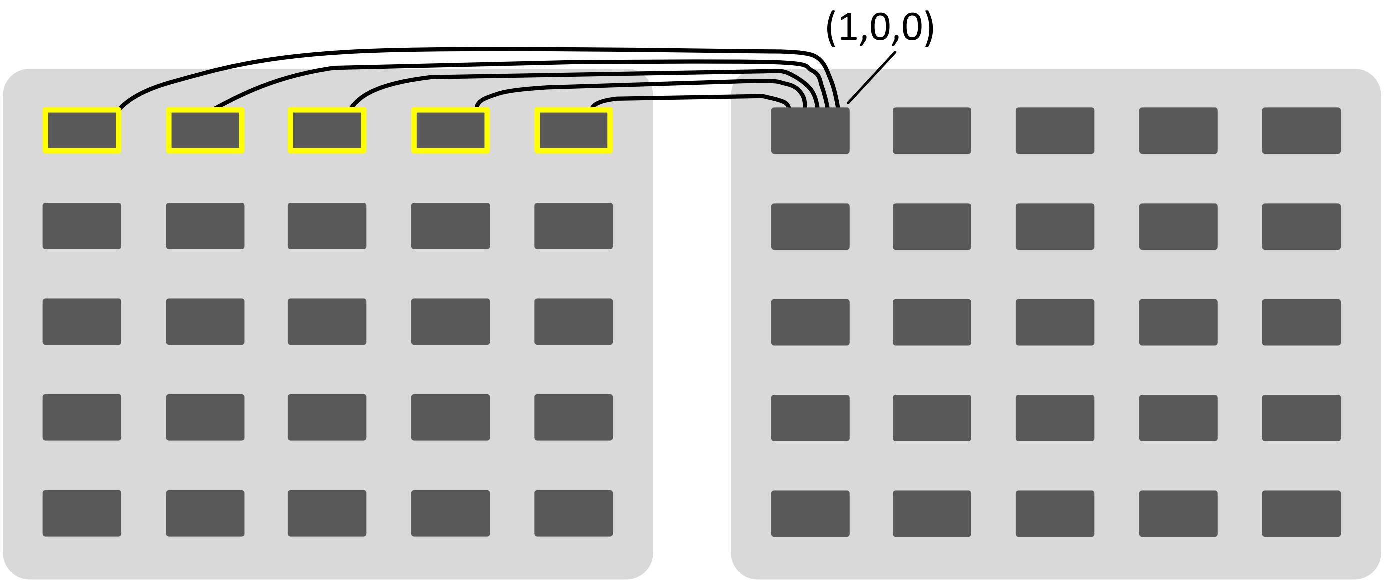
iff $y = mx + c$

E *Example:* $q = 5$

$m = 0, c = 0$

Take Router $(1,0,0)$

$(1,0,0) \leftrightarrow (0, x, 0)$



Diameter-2 Slim Fly

7 *Inter-group connections*

Router $(0, x, y) \leftrightarrow (1, m, c)$

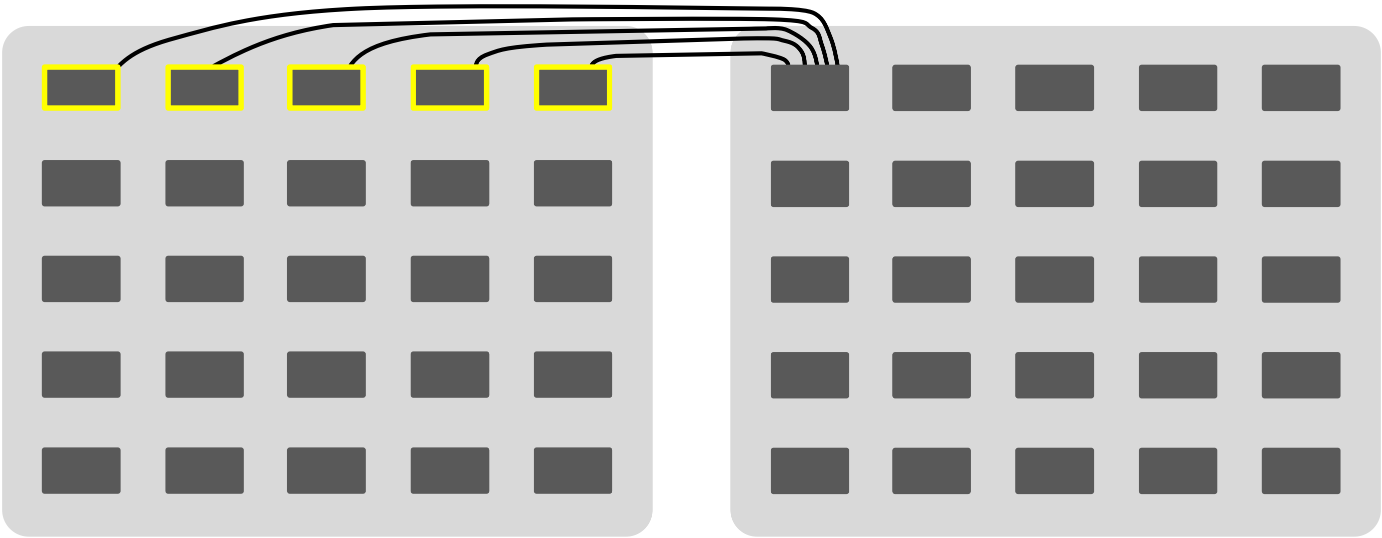
iff $y = mx + c$

E *Example:* $q = 5$

$m = 0, c = 0$

Take Router $(1,0,0)$

$(1,0,0) \leftrightarrow (0, x, 0)$



DIAMETER-2 SLIM FLY

7 *Inter-group connections*

Router $(0, x, y) \leftrightarrow (1, m, c)$

iff $y = mx + c$

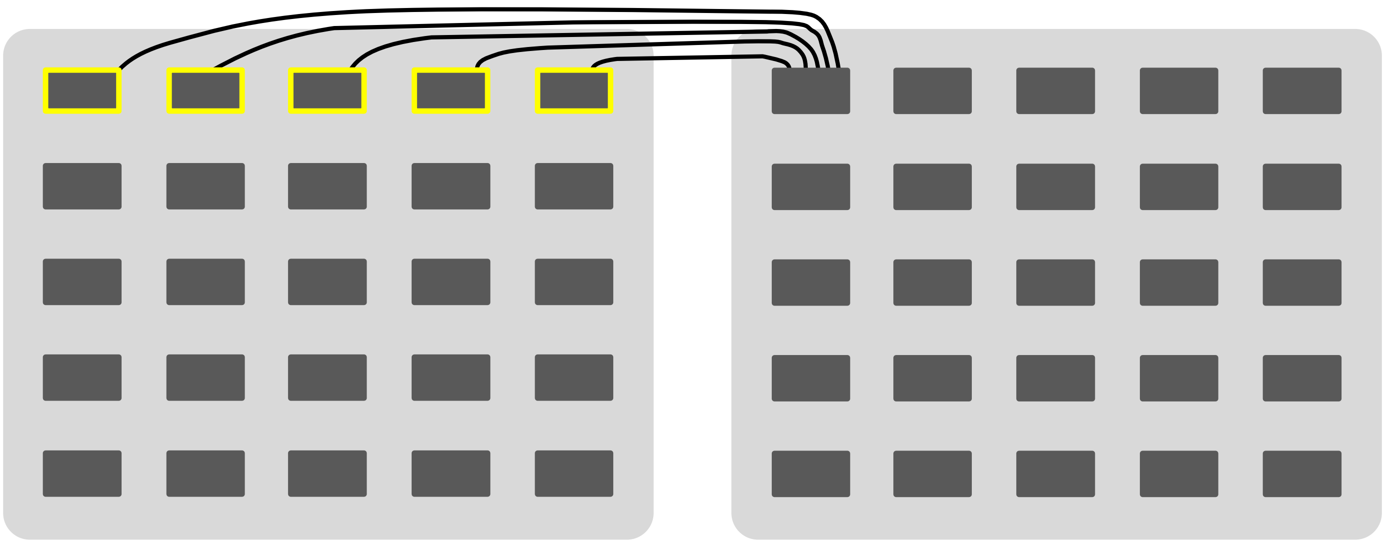
E *Example:* $q = 5$

$m = 0, c = 0$

Take Router $(1,0,0)$

$(1,0,0) \leftrightarrow (0, x, 0)$

Take Router $(1,1,0)$



Diameter-2 Slim Fly

7 *Inter-group connections*

Router $(0, x, y) \leftrightarrow (1, m, c)$

iff $y = mx + c$

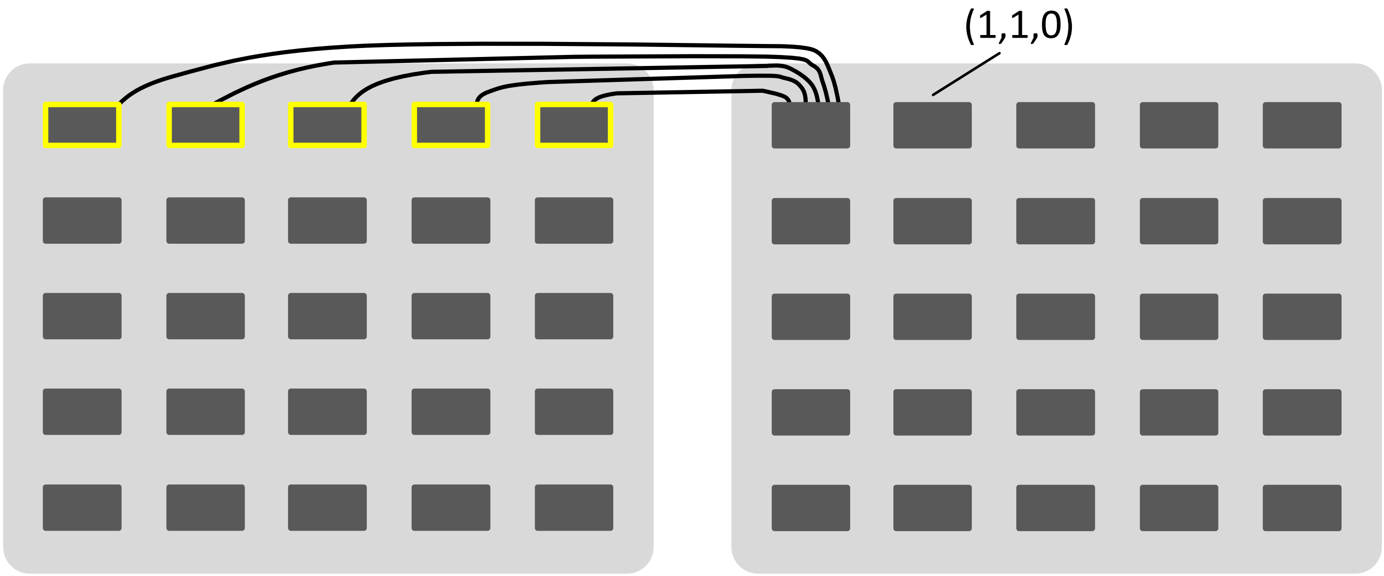
E *Example:* $q = 5$

$m = 0, c = 0$

Take Router $(1,0,0)$

$(1,0,0) \leftrightarrow (0, x, 0)$

Take Router $(1,1,0)$



Diameter-2 Slim Fly

7 *Inter-group connections*

Router $(0, x, y) \leftrightarrow (1, m, c)$

iff $y = mx + c$

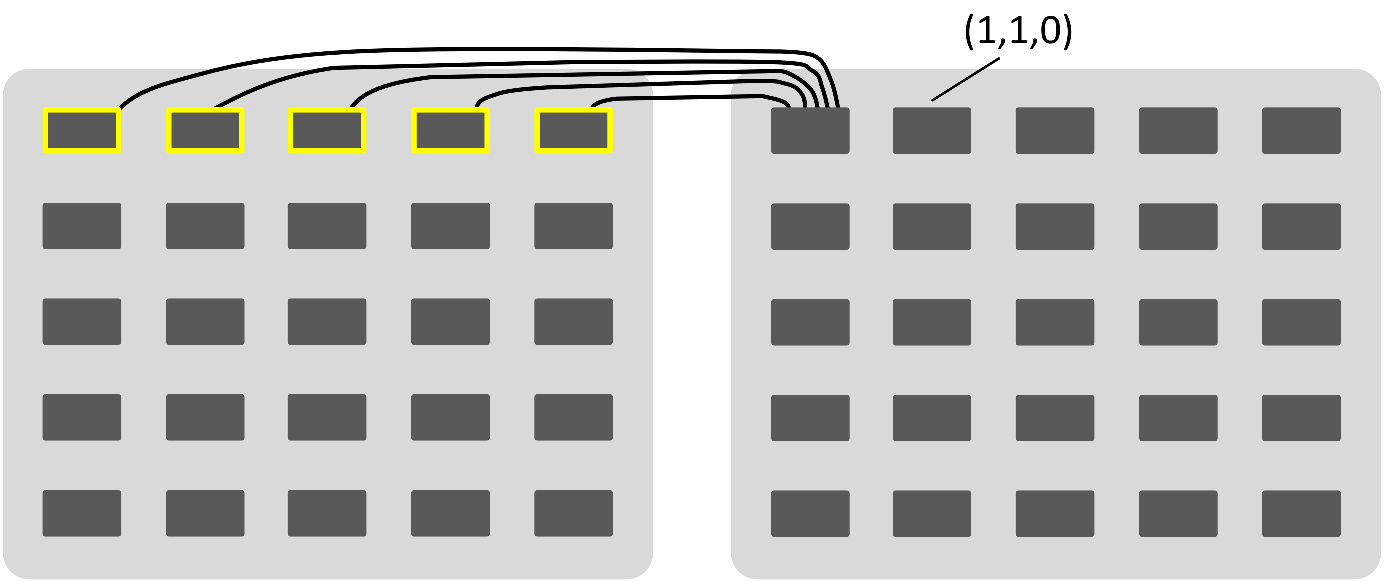
E *Example:* $q = 5$

$m = 0, c = 0$

Take Router $(1,0,0)$

$(1,0,0) \leftrightarrow (0, x, 0)$

Take Router $(1,1,0)$



Diameter-2 Slim Fly

7 *Inter-group connections*

Router $(0, x, y) \leftrightarrow (1, m, c)$

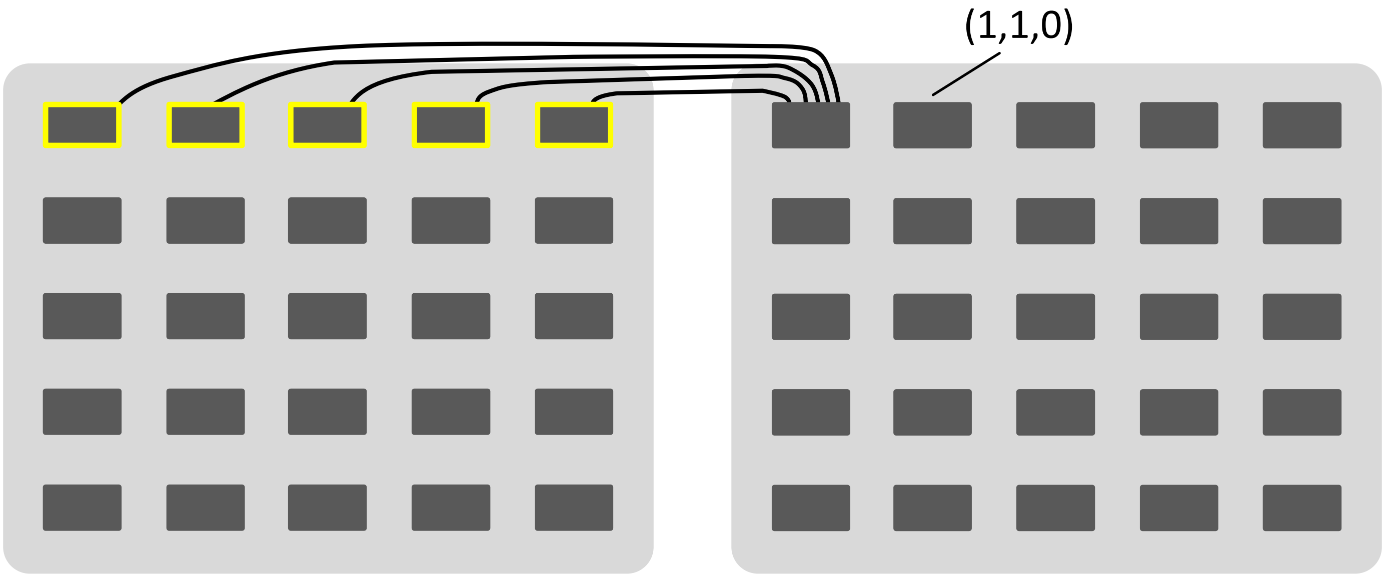
iff $y = mx + c$

E *Example:* $q = 5$

Take Router $(1,0,0)$ $m = 0, c = 0$

$(1,0,0) \leftrightarrow (0, x, 0)$

Take Router $(1,1,0)$ $m = 1, c = 0$



Diameter-2 Slim Fly

7 *Inter-group connections*

Router $(0, x, y) \leftrightarrow (1, m, c)$

iff $y = mx + c$

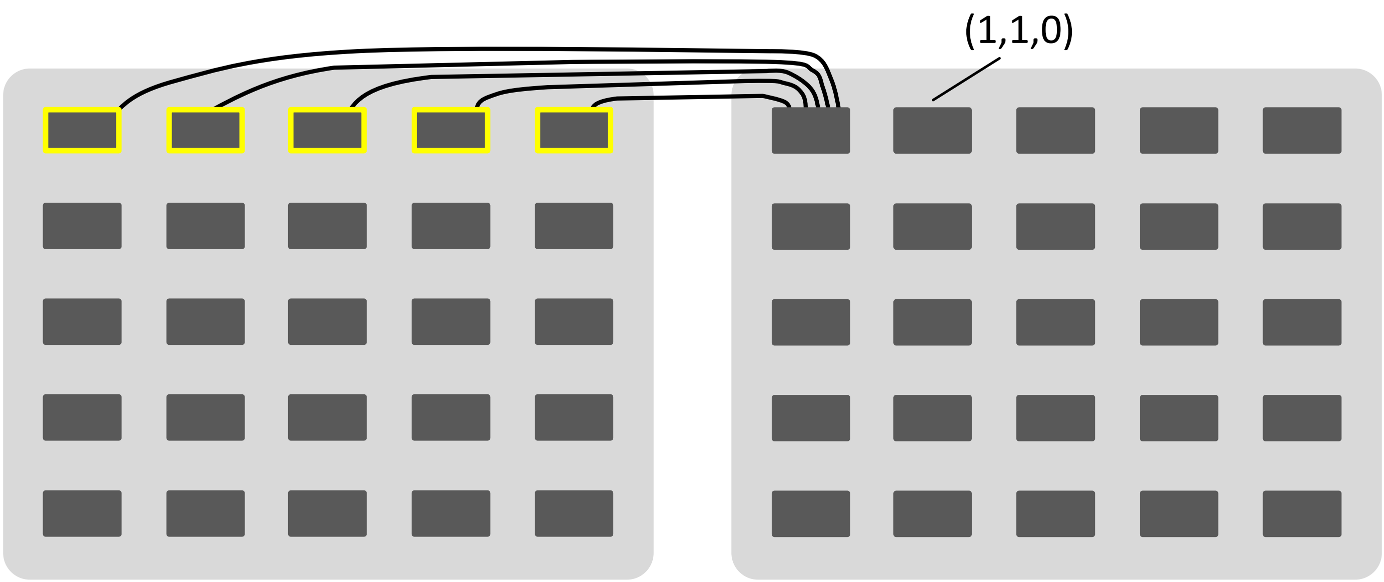
E *Example:* $q = 5$

Take Router $(1,0,0)$ $m = 0, c = 0$

$(1,0,0) \leftrightarrow (0, x, 0)$

Take Router $(1,1,0)$ $m = 1, c = 0$

$(1,0,0) \leftrightarrow (0, x, x)$



Diameter-2 Slim Fly

7 *Inter-group connections*

Router $(0, x, y) \leftrightarrow (1, m, c)$

iff $y = mx + c$

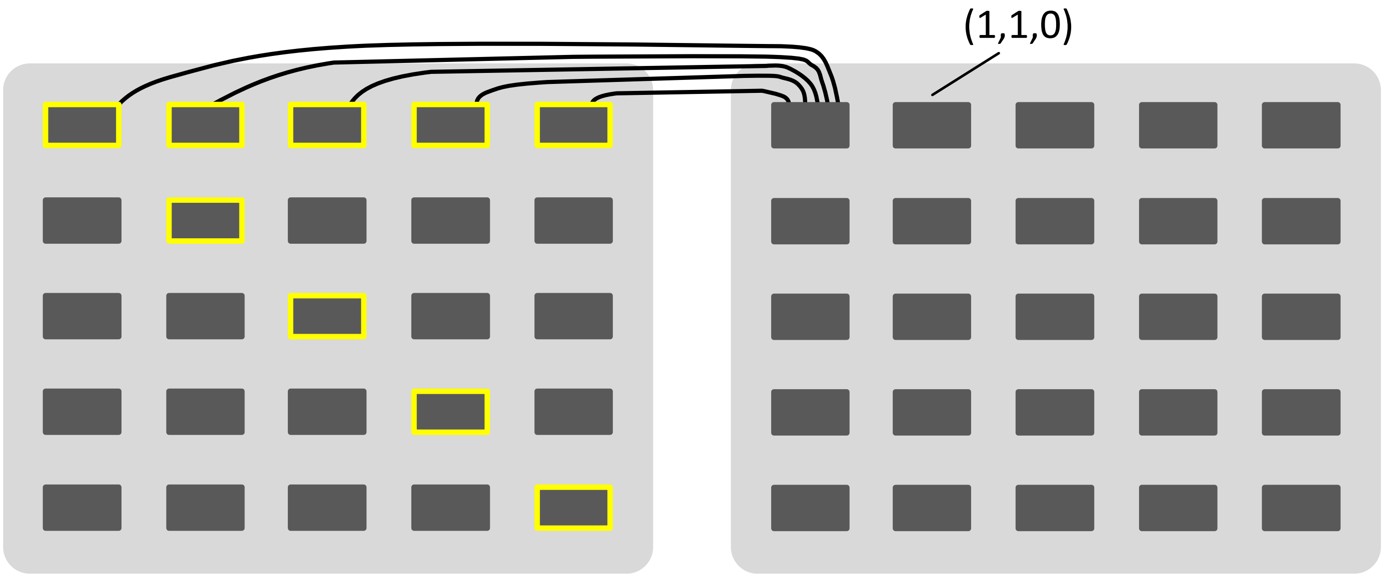
E *Example:* $q = 5$

Take Router $(1,0,0)$ $m = 0, c = 0$

$(1,0,0) \leftrightarrow (0, x, 0)$

Take Router $(1,1,0)$ $m = 1, c = 0$

$(1,0,0) \leftrightarrow (0, x, x)$



Diameter-2 Slim Fly

7 *Inter-group connections*

Router $(0, x, y) \leftrightarrow (1, m, c)$

iff $y = mx + c$

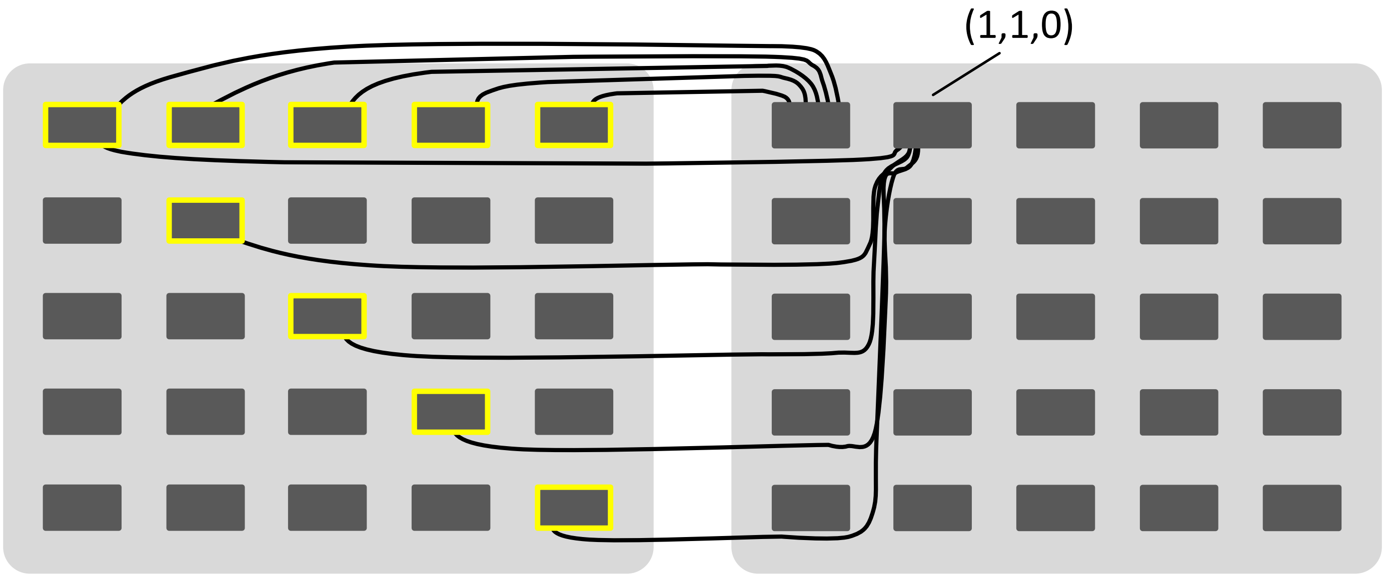
E *Example:* $q = 5$

Take Router $(1,0,0)$ $m = 0, c = 0$

$(1,0,0) \leftrightarrow (0, x, 0)$

Take Router $(1,1,0)$ $m = 1, c = 0$

$(1,0,0) \leftrightarrow (0, x, x)$



Diameter-2 Slim Fly

7 *Inter-group connections*

Router $(0, x, y) \leftrightarrow (1, m, c)$

iff $y = mx + c$

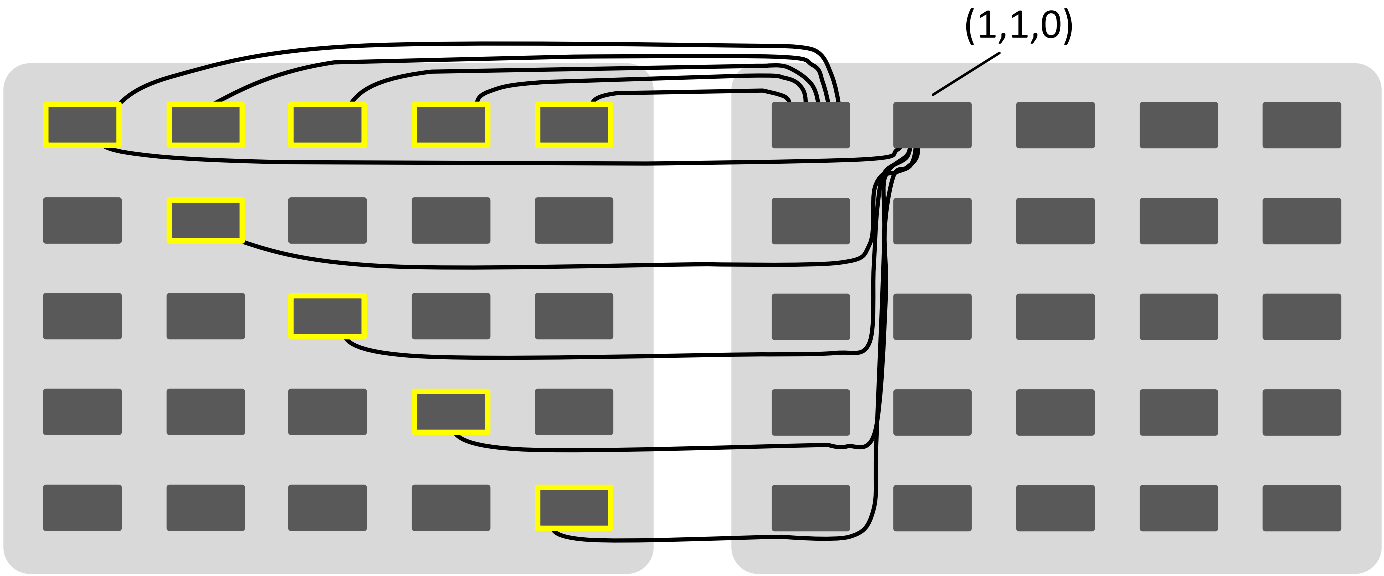
E *Example:* $q = 5$

Take Router $(1,0,0)$ $m = 0, c = 0$

$(1,0,0) \leftrightarrow (0, x, 0)$

Take Router $(1,1,0)$ $m = 1, c = 0$

$(1,0,0) \leftrightarrow (0, x, x)$



Diameter-2 Slim Fly

7 *Inter-group connections*

Router $(0, x, y) \leftrightarrow (1, m, c)$

iff $y = mx + c$

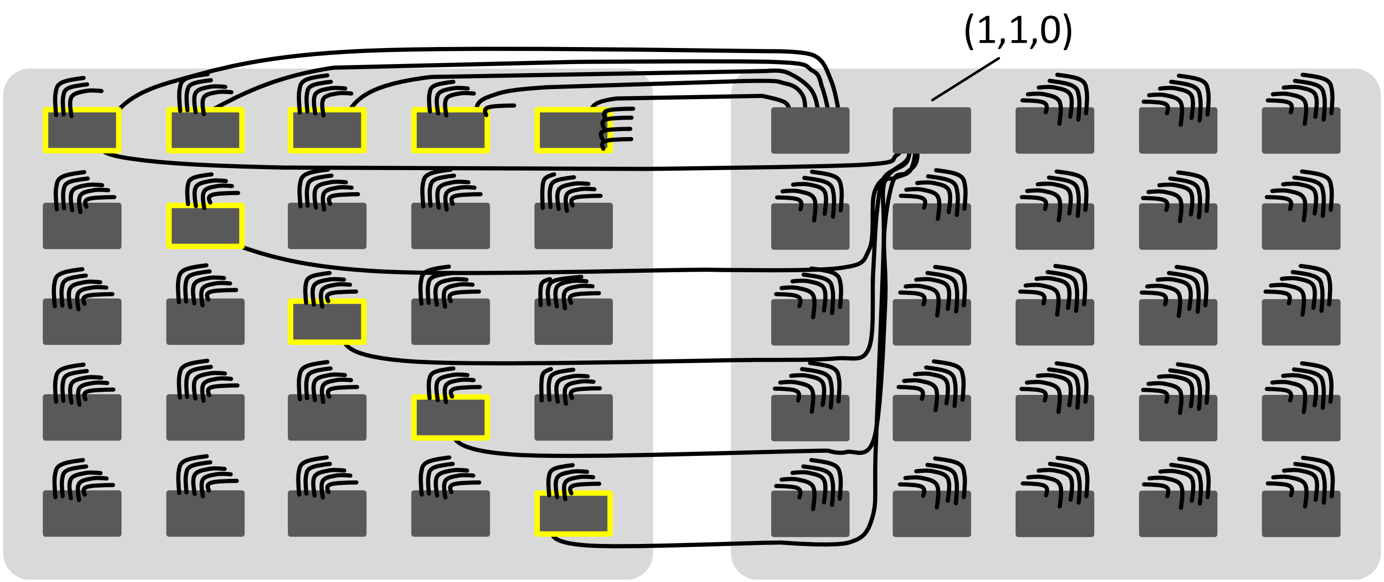
E *Example:* $q = 5$

Take Router $(1,0,0)$ $m = 0, c = 0$

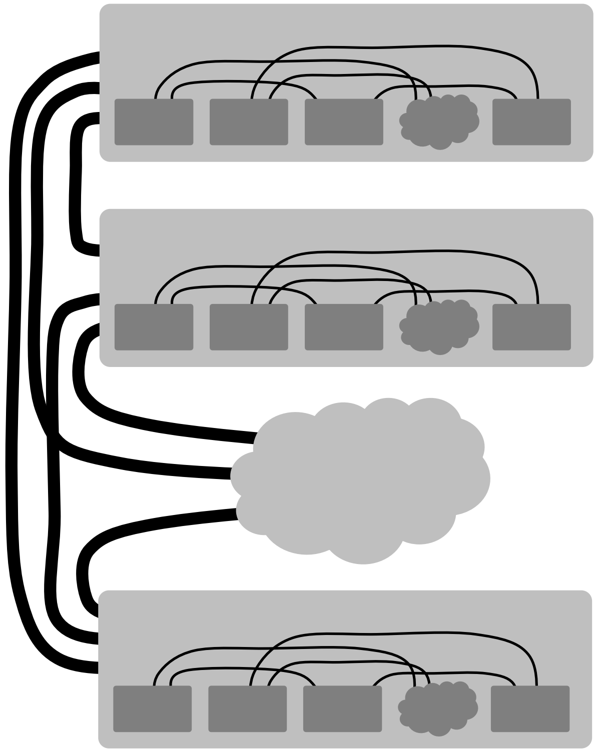
$(1,0,0) \leftrightarrow (0, x, 0)$

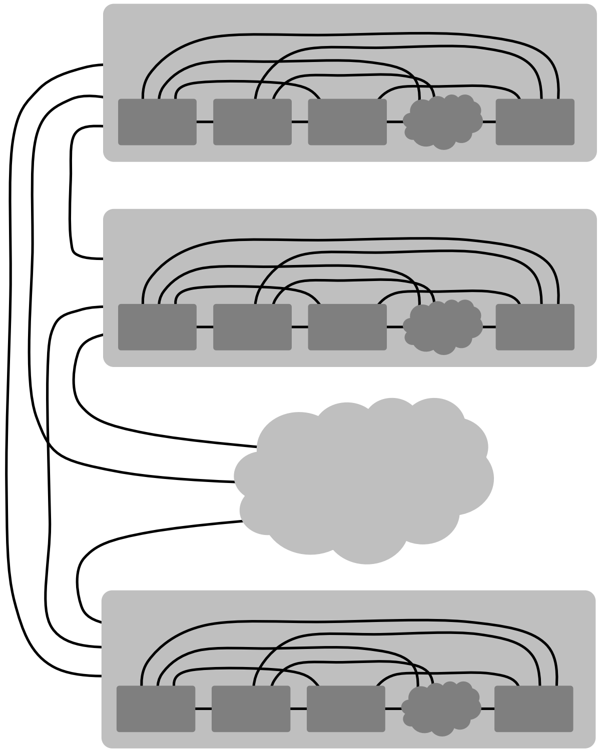
Take Router $(1,1,0)$ $m = 1, c = 0$

$(1,0,0) \leftrightarrow (0, x, x)$



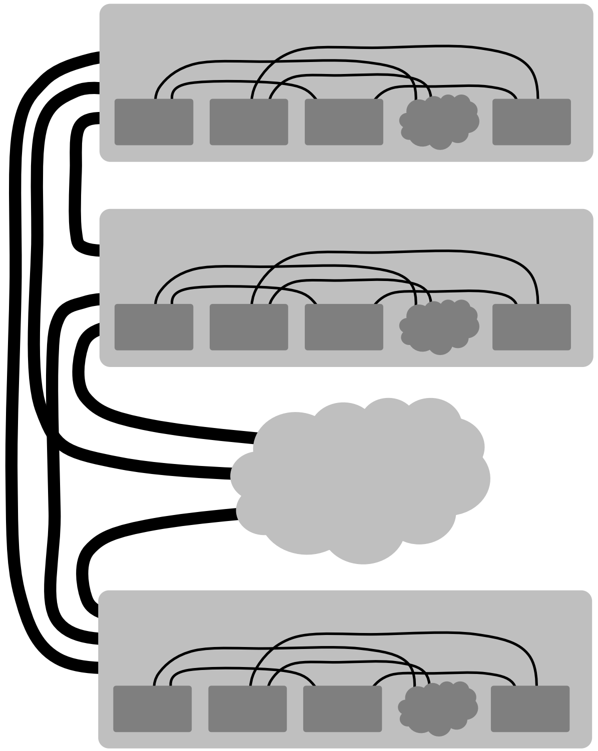
Design Intuition



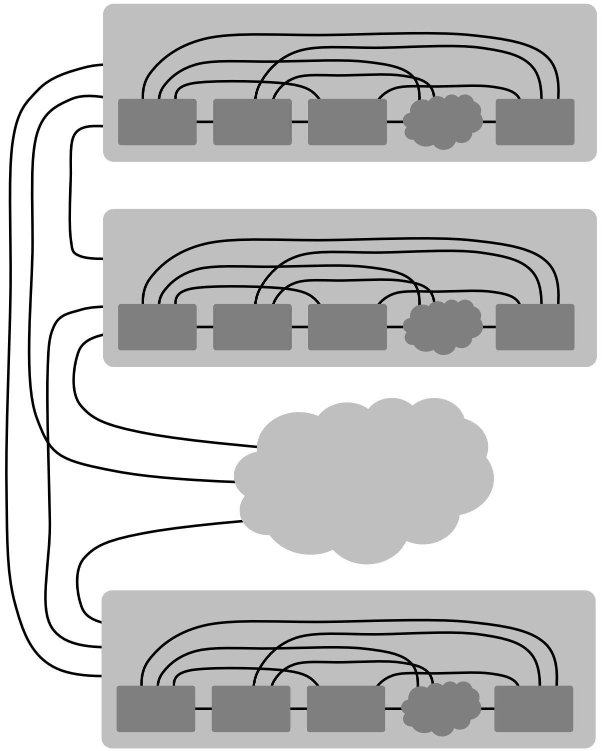


Design Intuition

SlimFly:

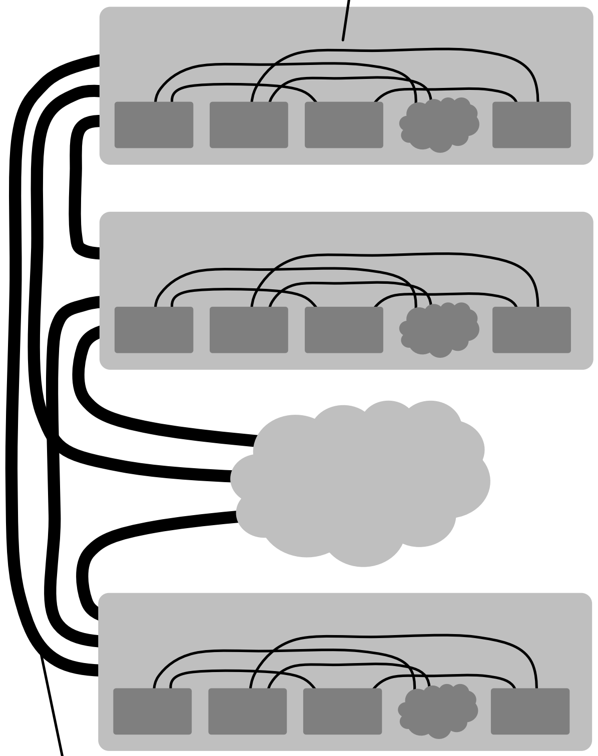


Dragonfly:

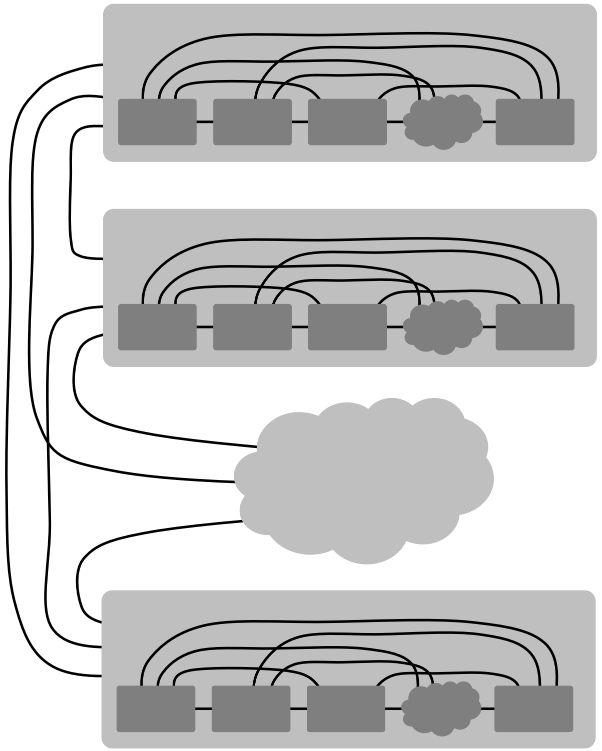


Design Intuition

SlimFly:



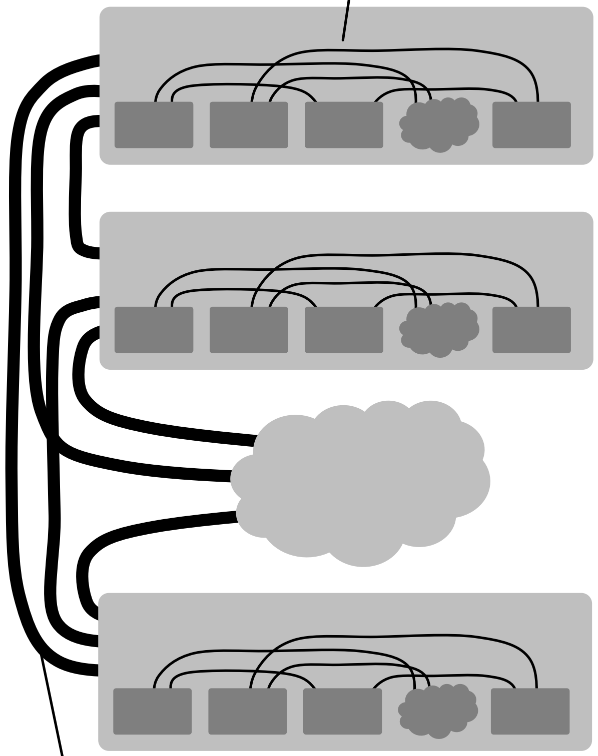
Dragonfly:



DESIGN INTUITION

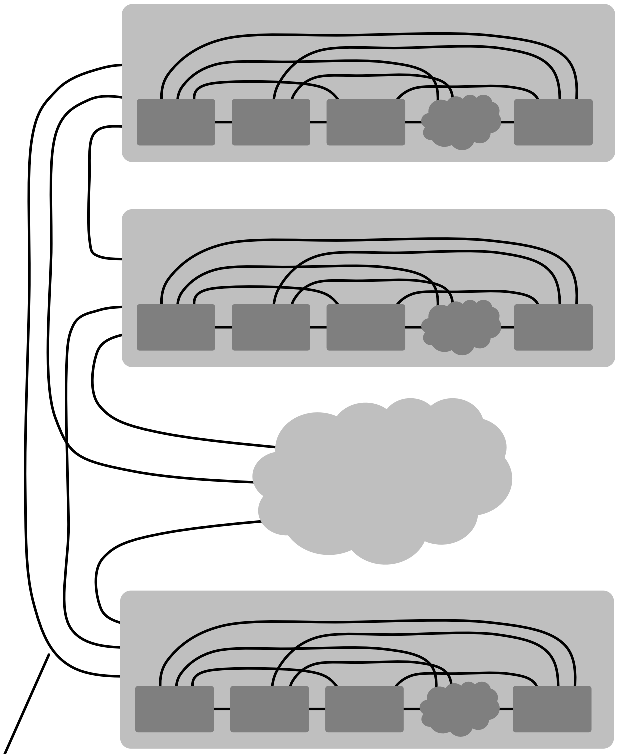
SlimFly:

~50% fewer
intra-group cables



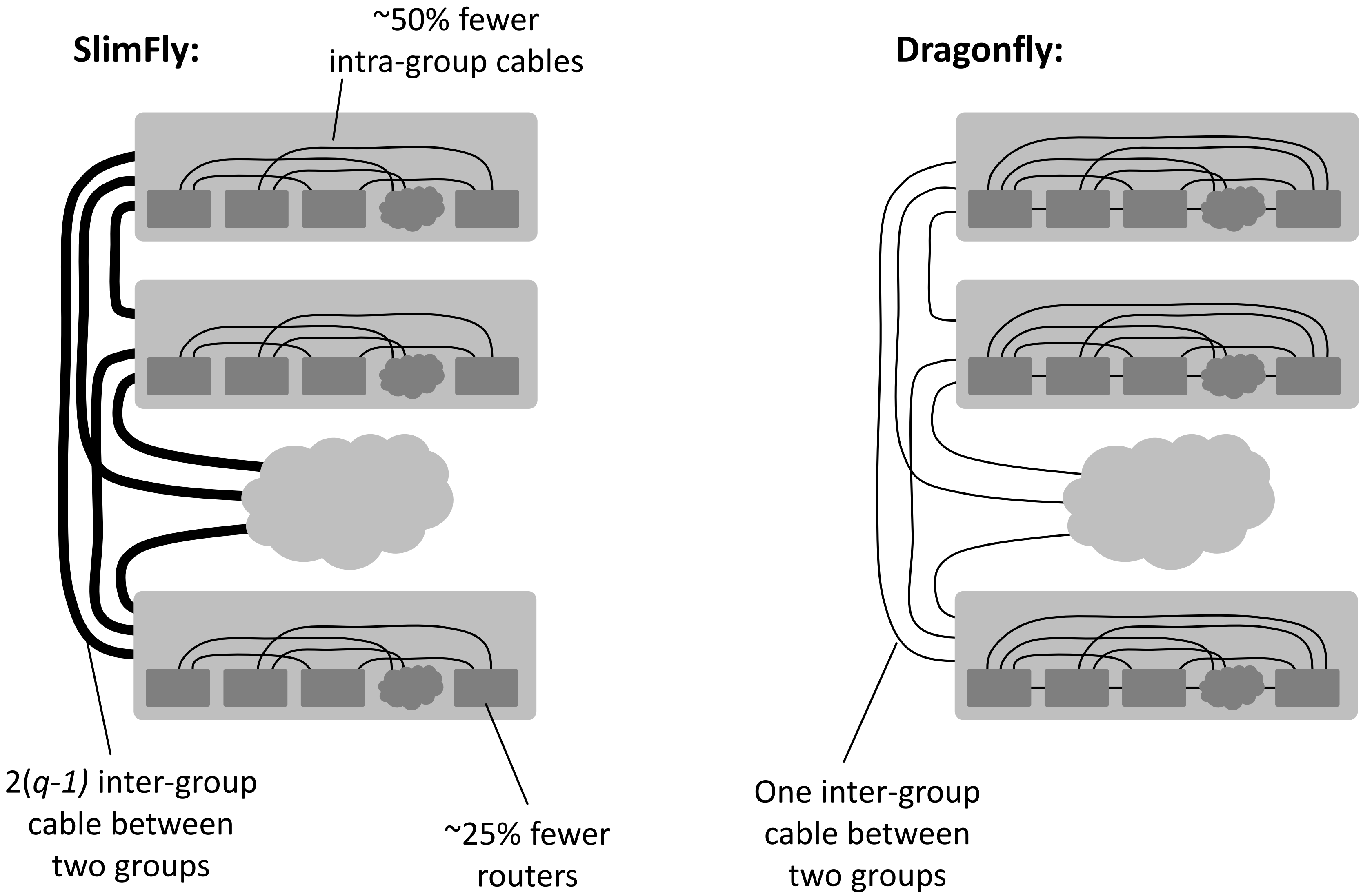
2(*q*-1) inter-group
cable between
two groups

Dragonfly:

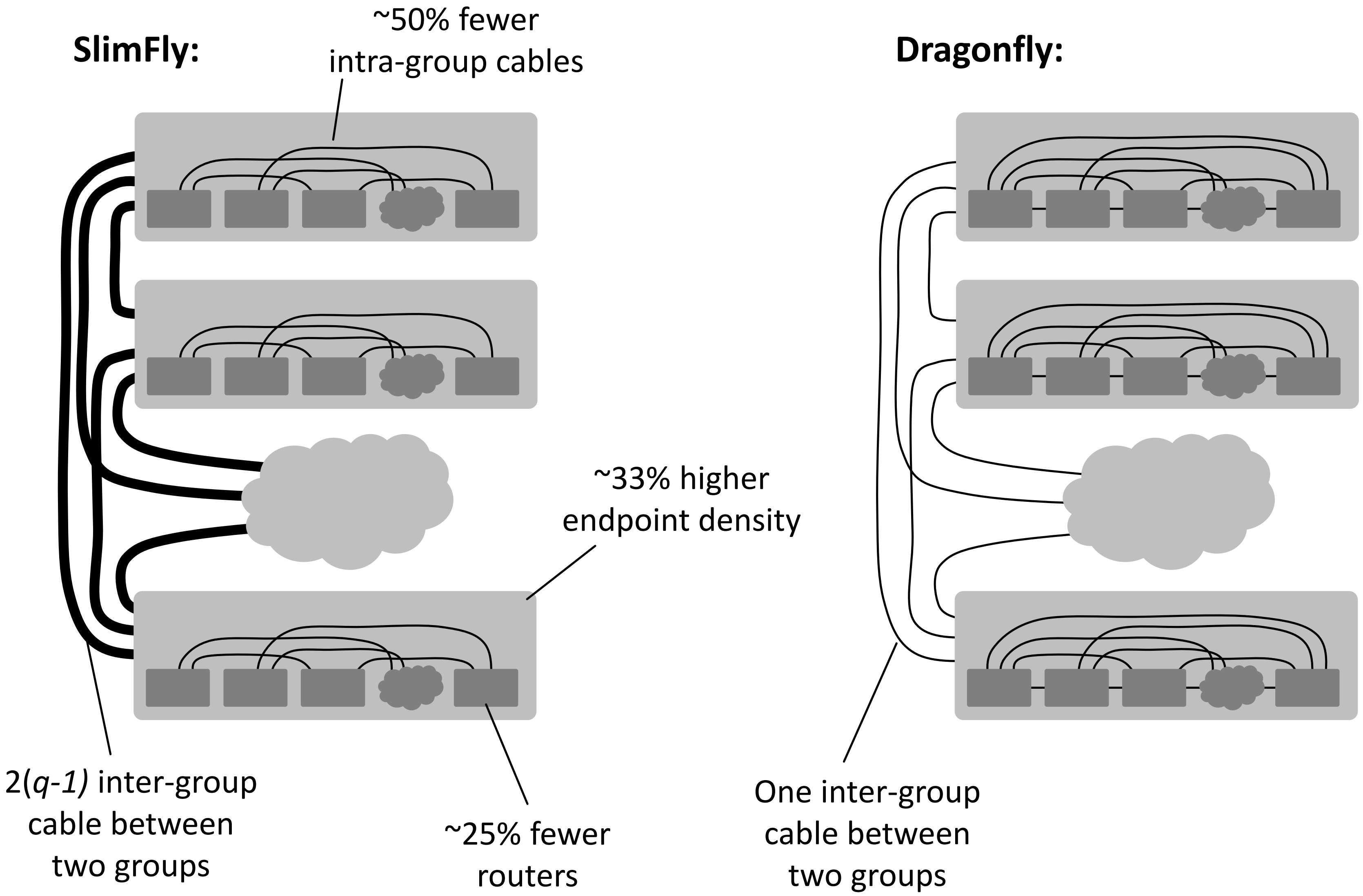


One inter-group
cable between
two groups

Design Intuition



DESIGN INTUITION



Deadlock Freedom

Minimum Static Routing

Deadlock Freedom

Minimum Static Routing

- Assign two virtual channels (VC0 and VC1) to each link

Deadlock Freedom

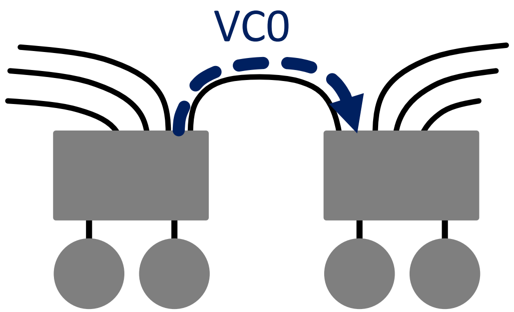
Minimum Static Routing

- Assign two virtual channels (VC0 and VC1) to each link
- For a 1-hop path use VC0

Deadlock Freedom

Minimum Static Routing

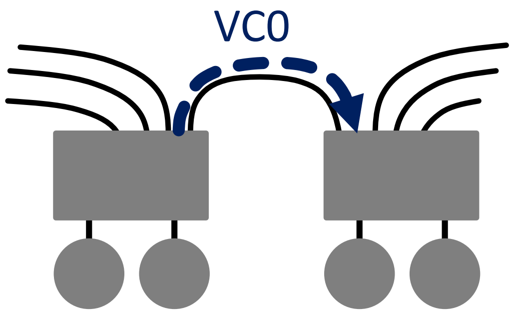
- Assign two virtual channels (VC0 and VC1) to each link
- For a 1-hop path use VC0



Deadlock Freedom

Minimum Static Routing

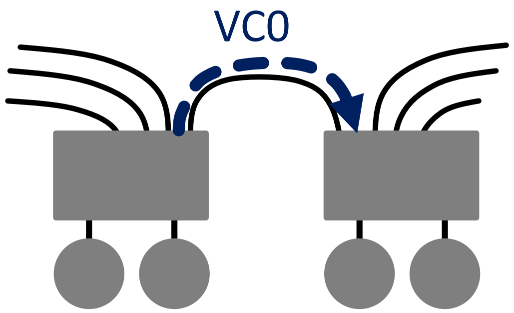
- Assign two virtual channels (VC0 and VC1) to each link
- For a 1-hop path use VC0
- For a 2-hop path use VC0 (hop 1) and VC1 (hop 2)

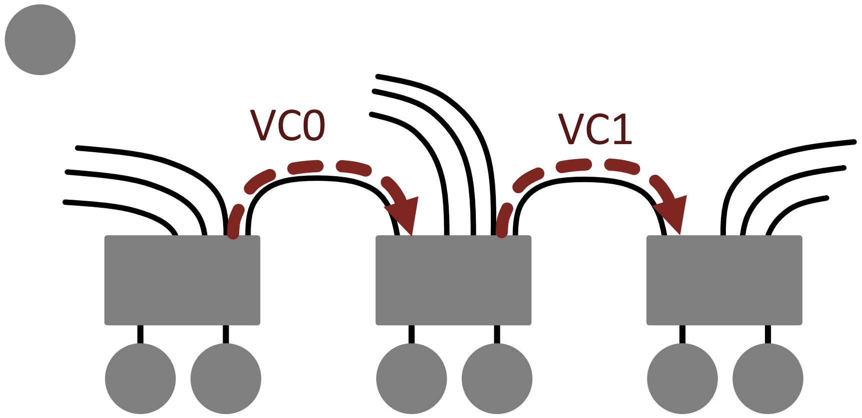


Deadlock Freedom

Minimum Static Routing

- Assign two virtual channels (VC0 and VC1) to each link
- For a 1-hop path use VC0
- For a 2-hop path use VC0 (hop 1) and VC1 (hop 2)

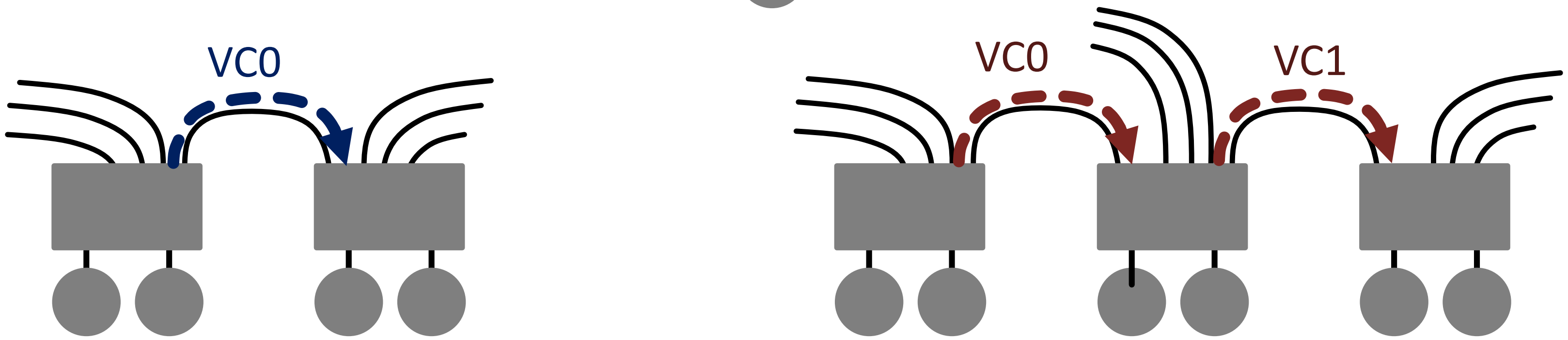




Deadlock Freedom

Minimum Static Routing

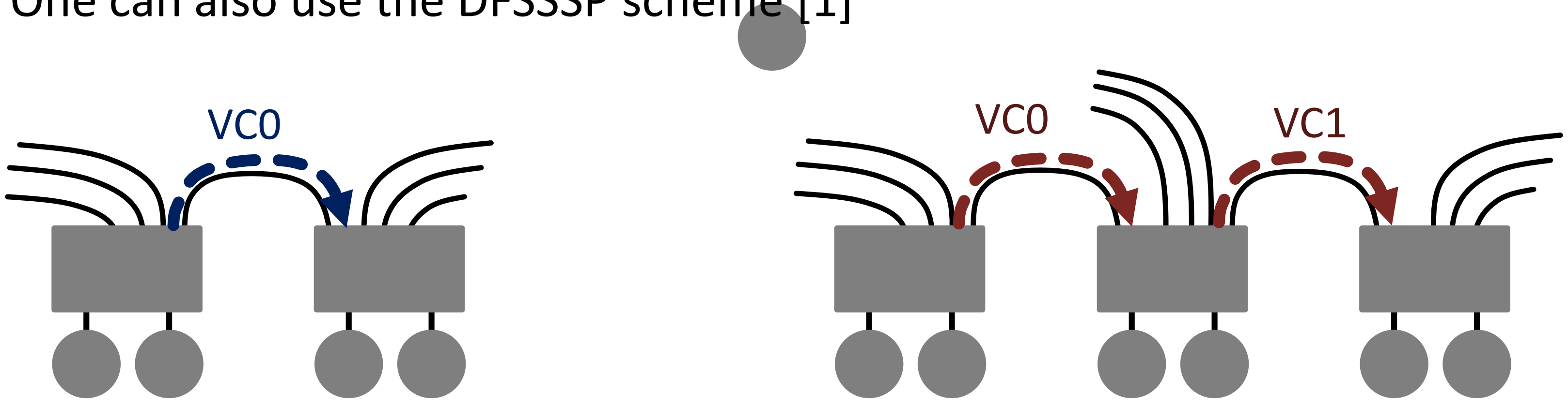
- Assign two virtual channels (VC0 and VC1) to each link
- For a 1-hop path use VC0
- For a 2-hop path use VC0 (hop 1) and VC1 (hop 2)
- One can also use the DFSSSP scheme [1]



Deadlock Freedom

Minimum Static Routing

- Assign two virtual channels (VC0 and VC1) to each link
- For a 1-hop path use VC0
- For a 2-hop path use VC0 (hop 1) and VC1 (hop 2)
- One can also use the DFSSSP scheme [1]



[1] J. Domke, T. Hoefler, and W. Nagel. Deadlock-Free Oblivious Routing for Arbitrary Topologies. IPDPS'11

Deadlock Freedom

Adaptive Routing

Deadlock Freedom

Adaptive Routing

- Simple generalization of the previous scheme

Deadlock Freedom

Adaptive Routing

- Simple generalization of the previous scheme
- Assign four virtual channels (VC0 – VC3) to each link

Deadlock Freedom

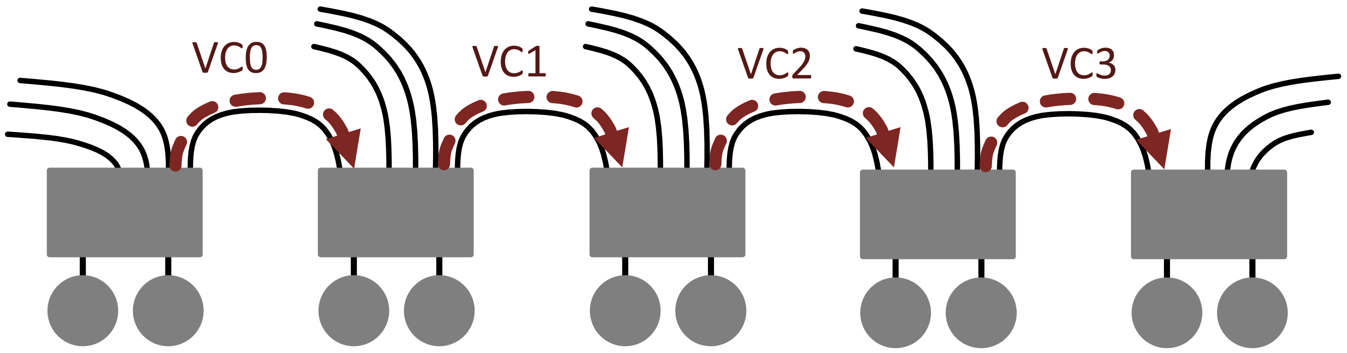
Adaptive Routing

- Simple generalization of the previous scheme
- Assign four virtual channels (VC0 – VC3) to each link
- For hop *k* path use VC*k*, 0 <= k <= 3

Deadlock Freedom

Adaptive Routing

- Simple generalization of the previous scheme
- Assign four virtual channels (VC0 – VC3) to each link
- For hop *k* path use VC*k*, 0 <= k <= 3



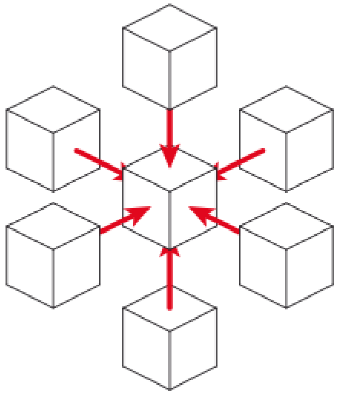
PERFORMANCE

PERFORMANCE

- Bit permutation traffic

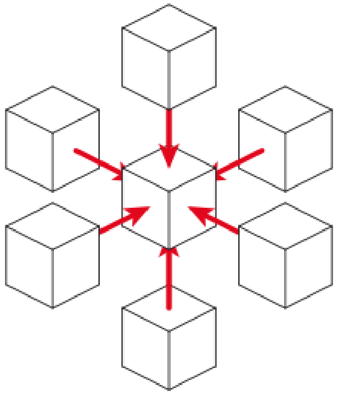
PERFORMANCE

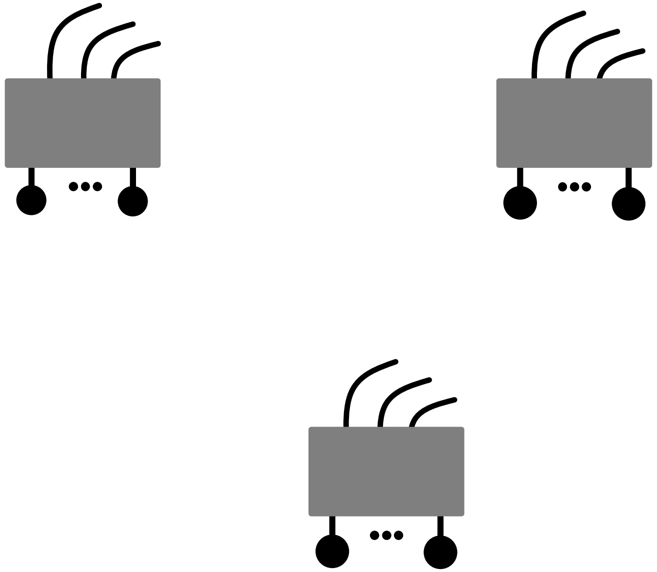
- Bit permutation traffic



PERFORMANCE

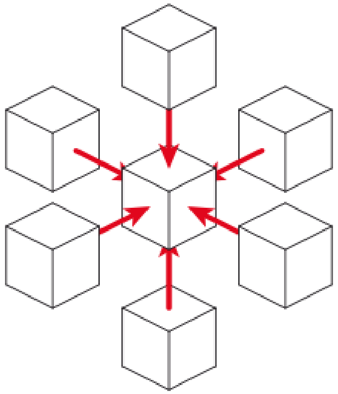
- Bit permutation traffic

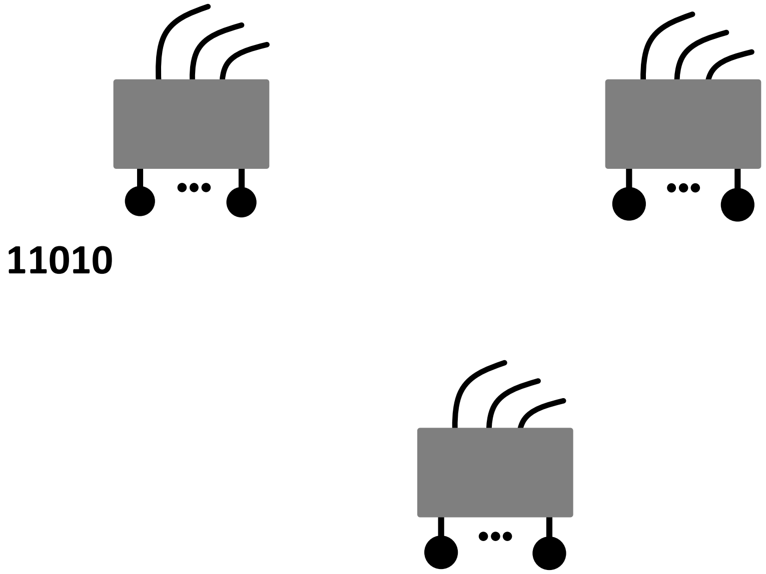




PERFORMANCE

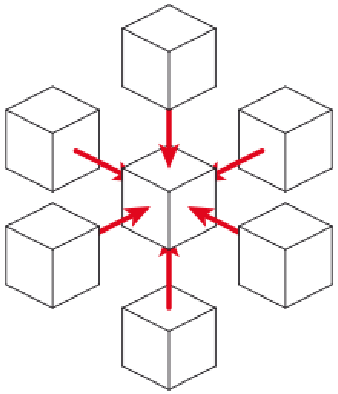
- Bit permutation traffic

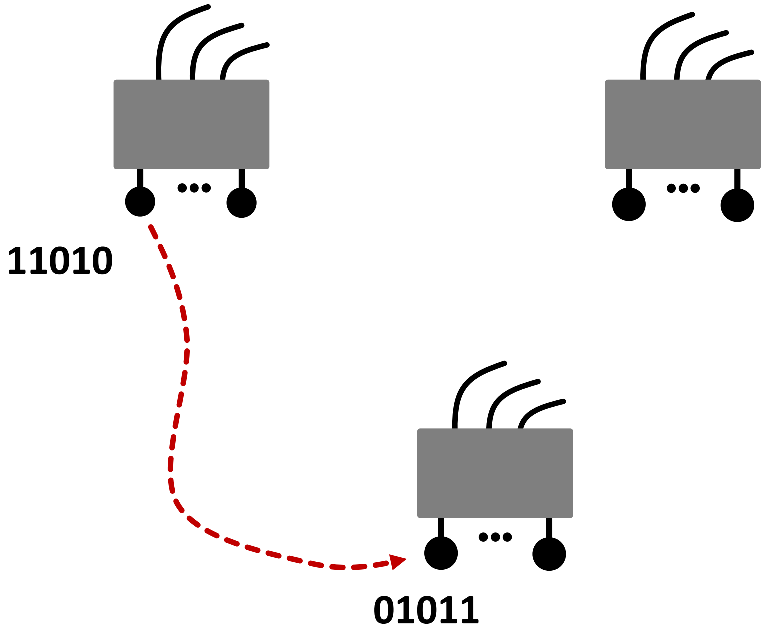




Performance

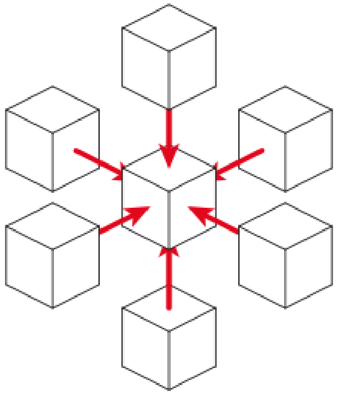
- Bit permutation traffic

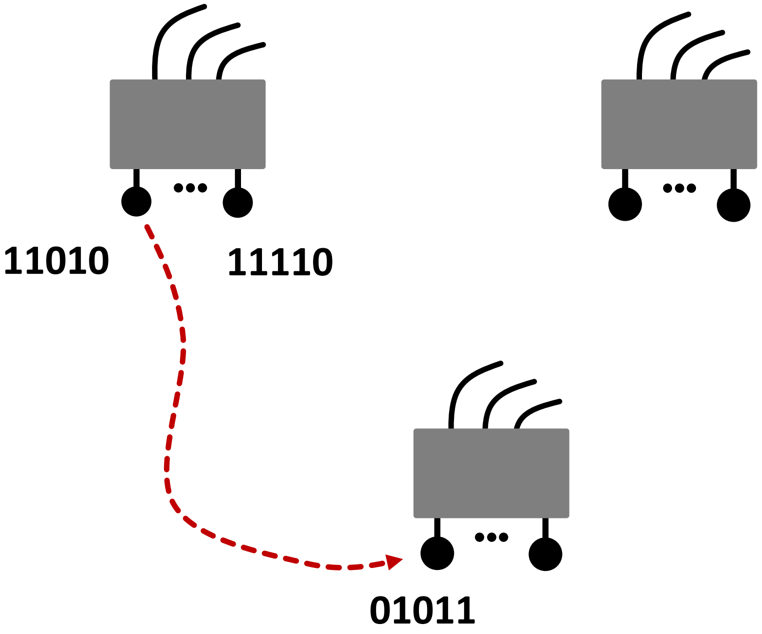




PERFORMANCE

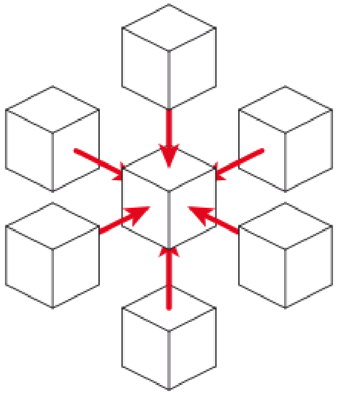
- Bit permutation traffic

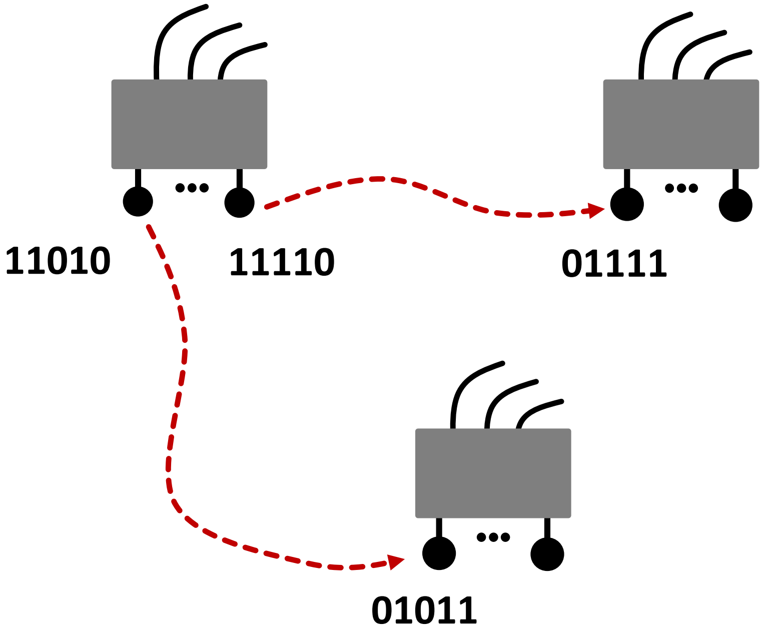




PERFORMANCE

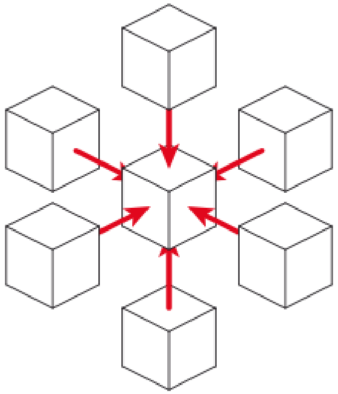
- Bit permutation traffic

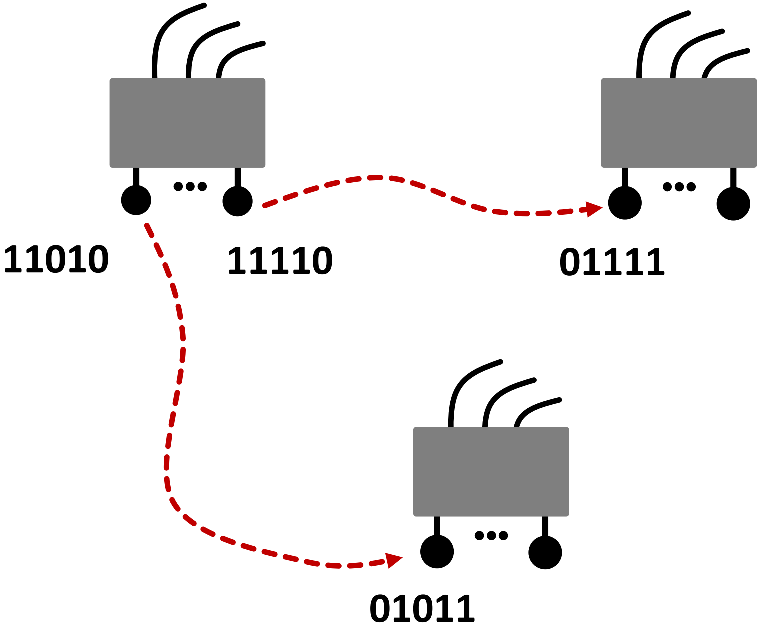


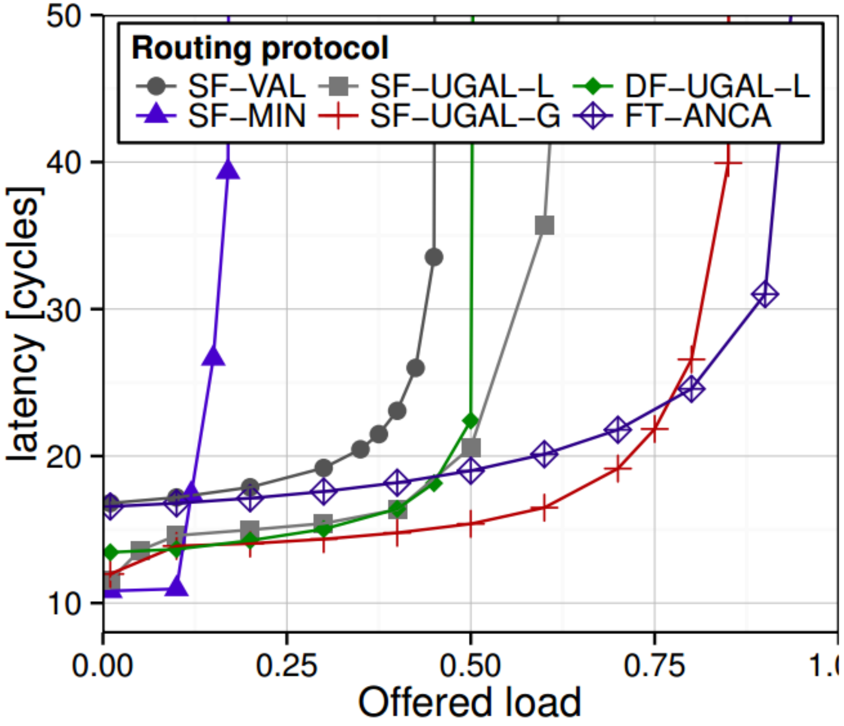


Performance

- Bit permutation traffic

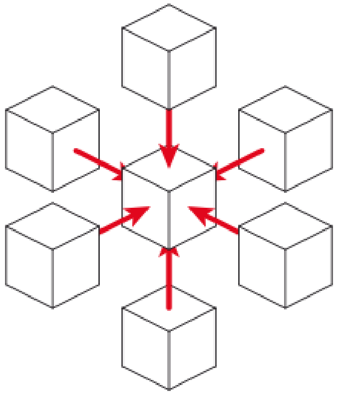


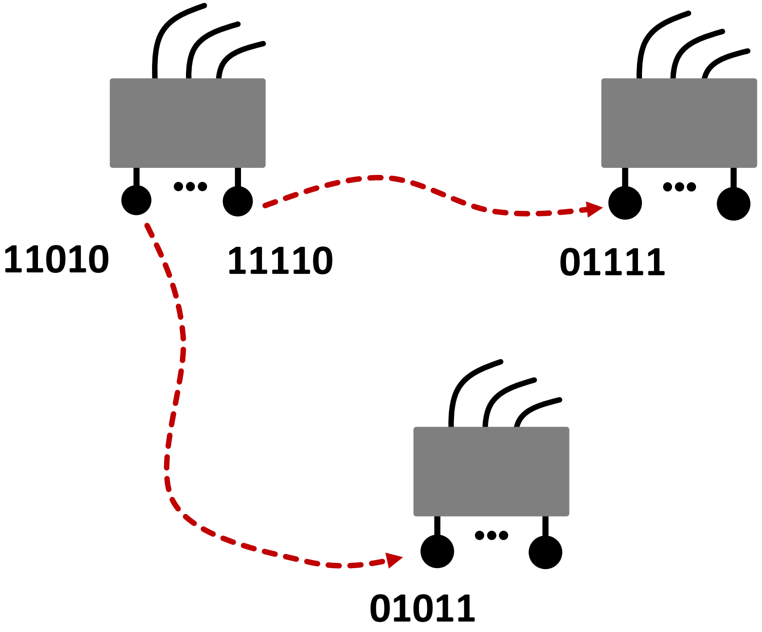


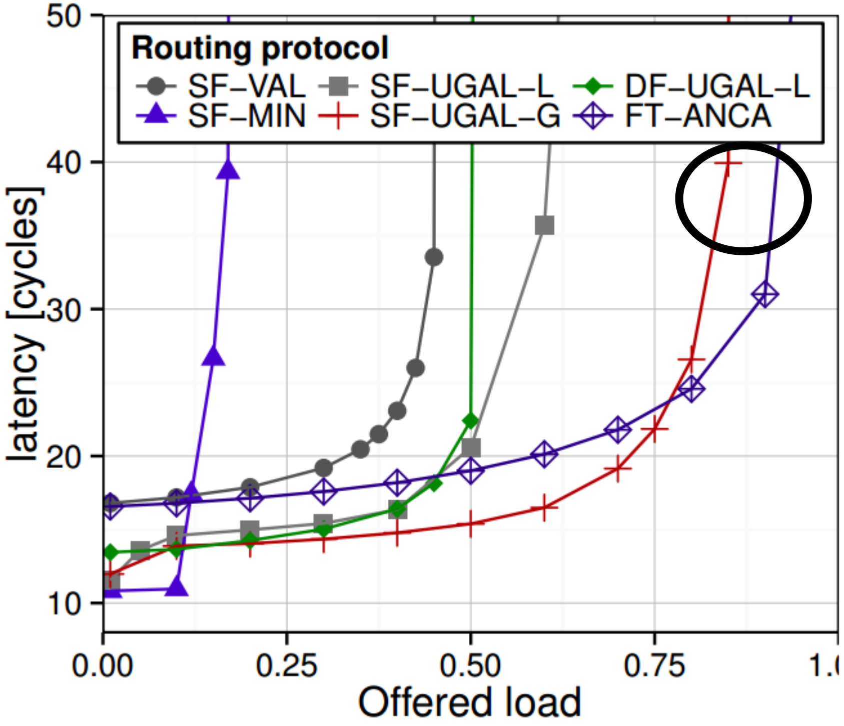


PERFORMANCE

- Bit permutation traffic

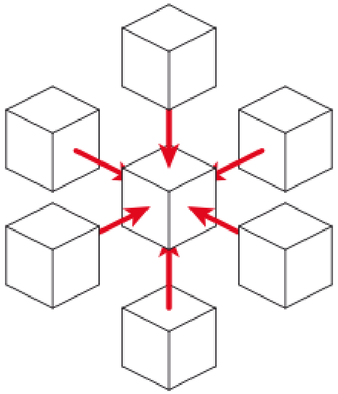


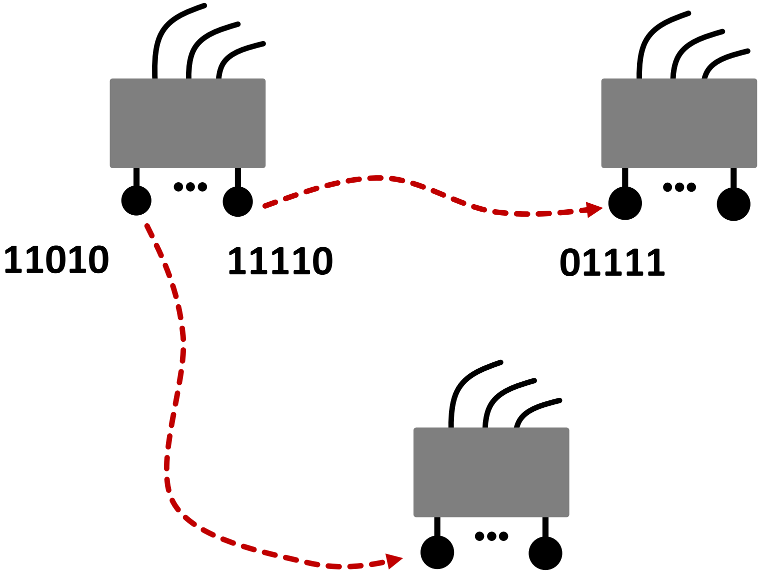


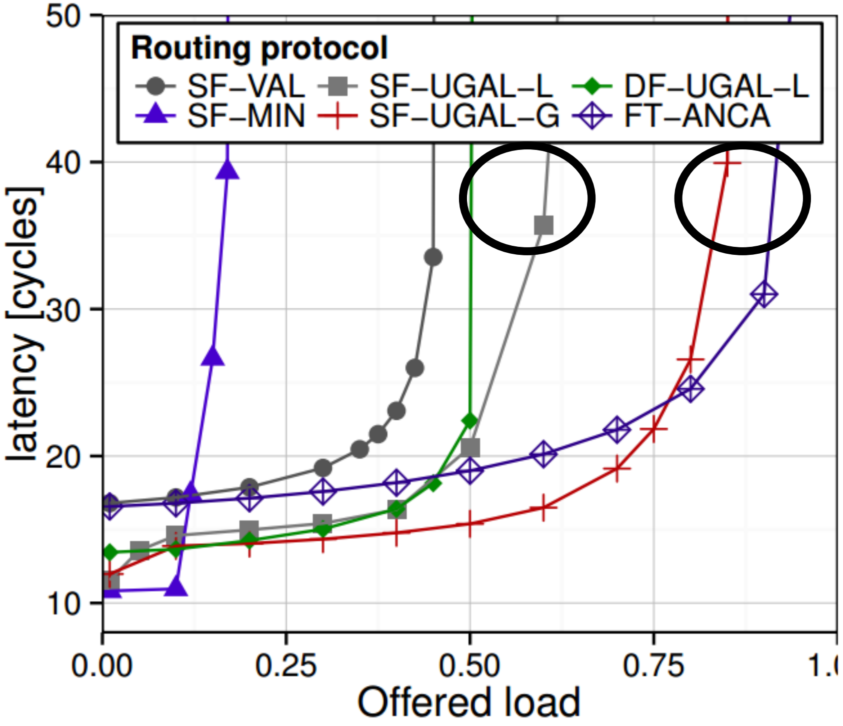


PERFORMANCE

- Bit permutation traffic







PERFORMANCE

PERFORMANCE

- Shift traffic

PERFORMANCE

- Shift traffic

dest id

$d =$

$d =$

PERFORMANCE

- Shift traffic

dest id

source id

$$d = \left(s \; mod \frac{N}{2}\right) + \frac{N}{2}$$

$$d =$$

Performance

- Shift traffic

dest id

source id

$$d = \left(s \bmod \frac{N}{2}\right) + \frac{N}{2}$$

$$d = s \bmod \frac{N}{2}$$

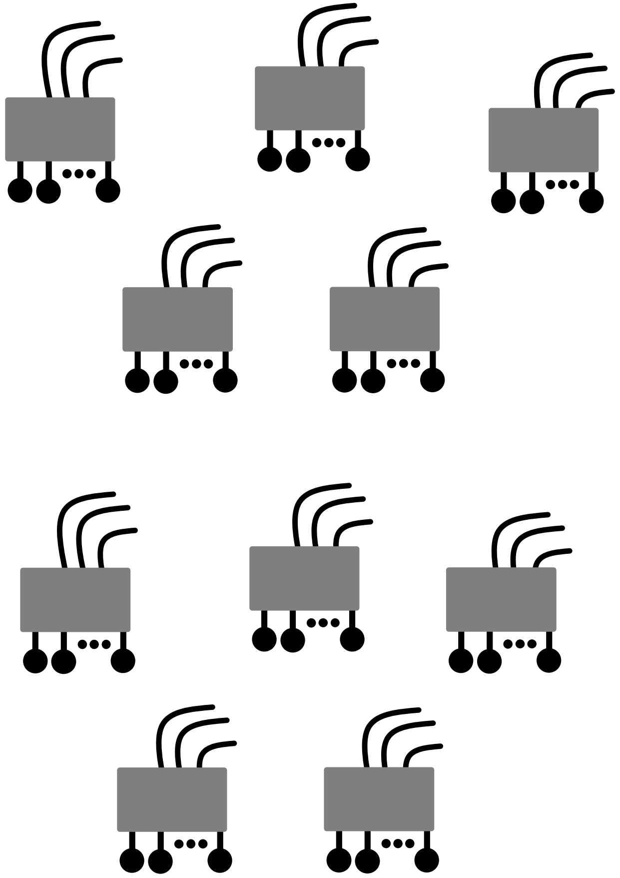
PERFORMANCE

- Shift traffic

dest id
source id

$$d = \left(s \bmod \frac{N}{2}\right) + \frac{N}{2}$$

$$d = s \bmod \frac{N}{2}$$



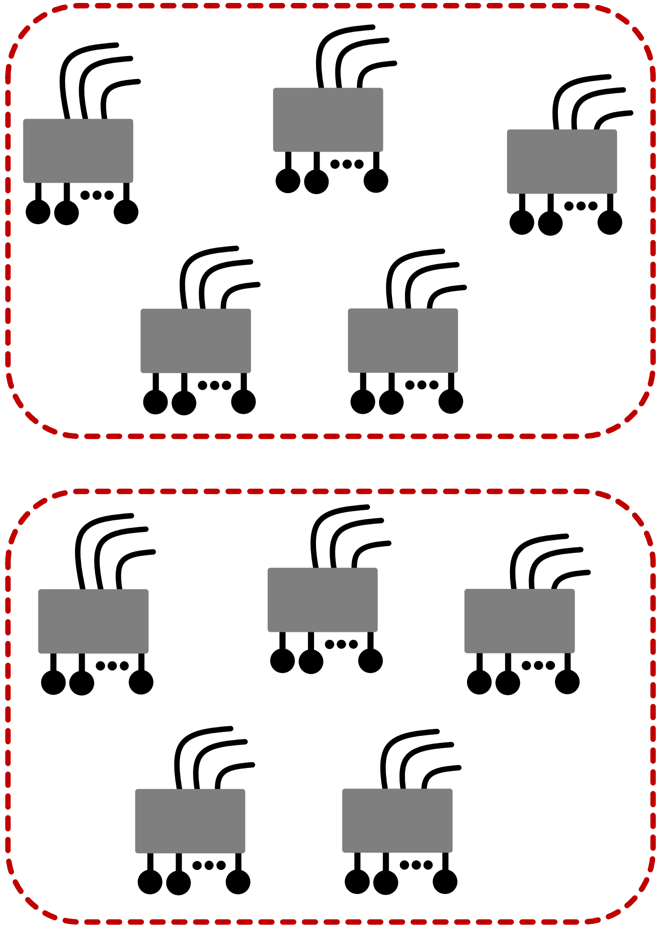
Performance

- Shift traffic

dest id — source id

$$d = \left(s \bmod \frac{N}{2}\right) + \frac{N}{2}$$

$$d = s \bmod \frac{N}{2}$$



Performance

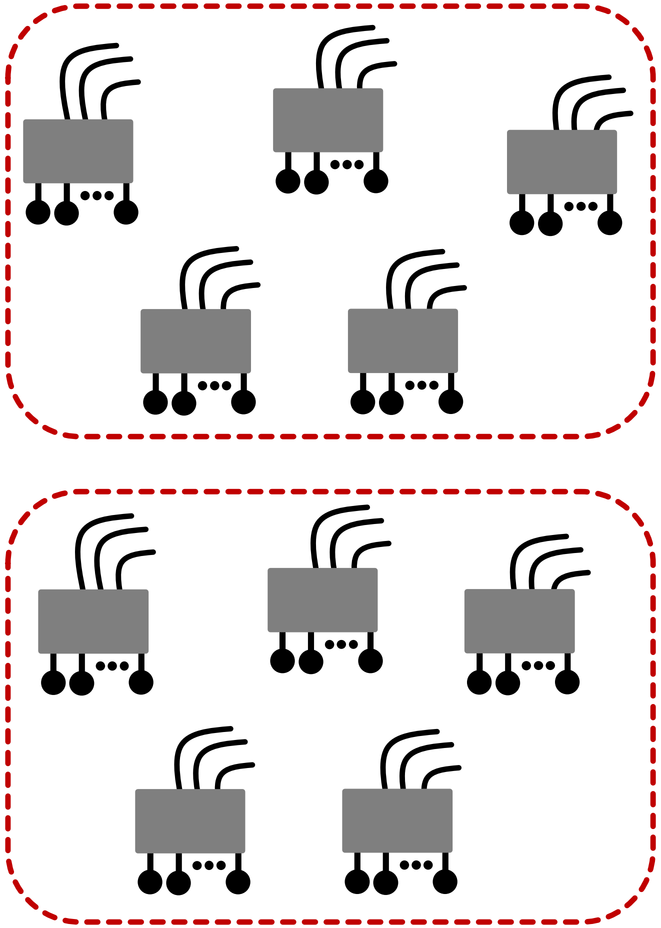
- Shift traffic

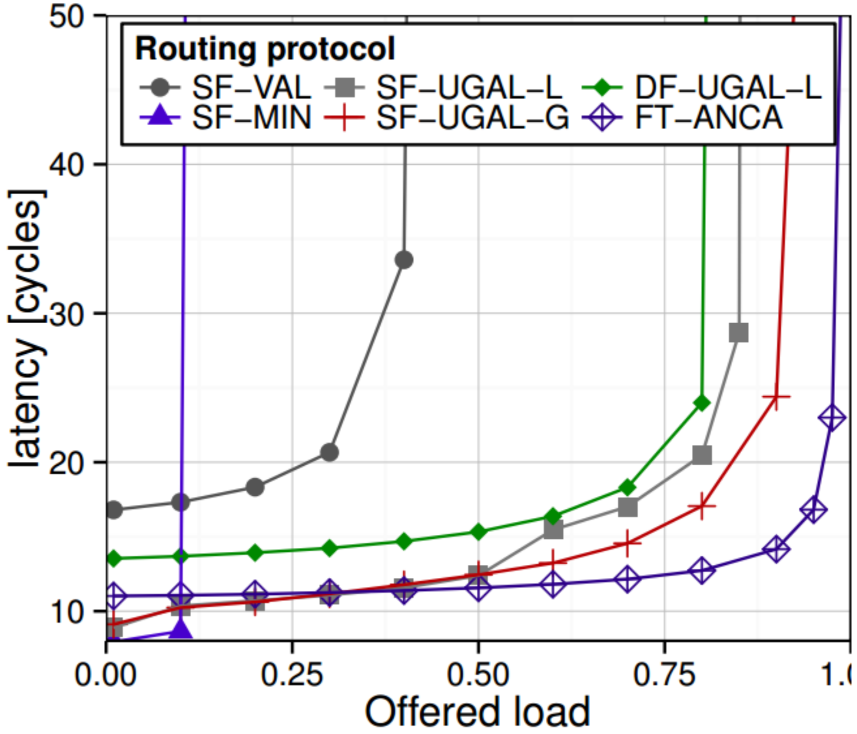
dest id

source id

$$d = \left(s \bmod \frac{N}{2}\right) + \frac{N}{2}$$

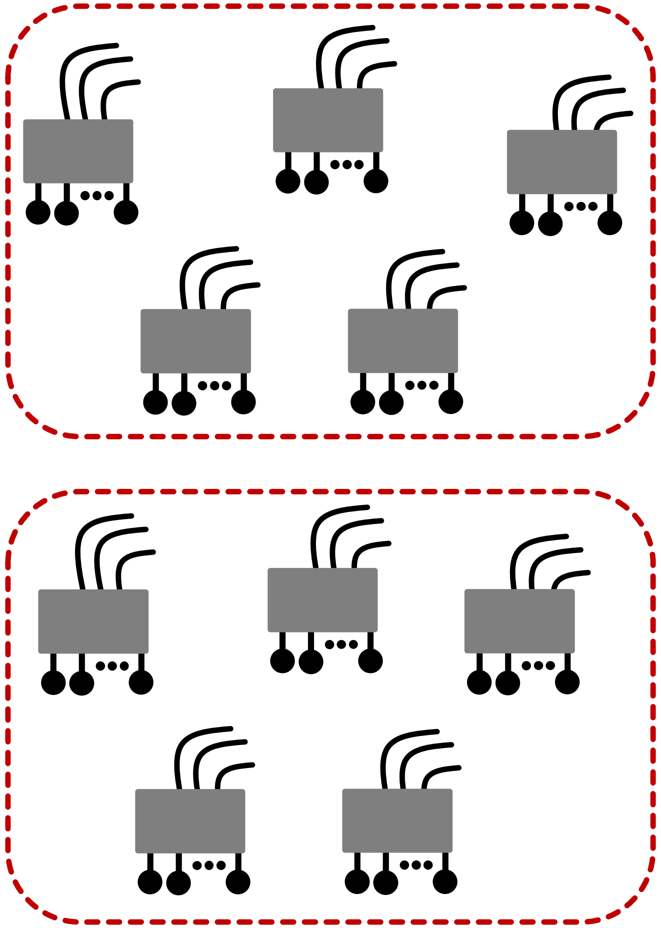
$$d = s \bmod \frac{N}{2}$$





PERFORMANCE

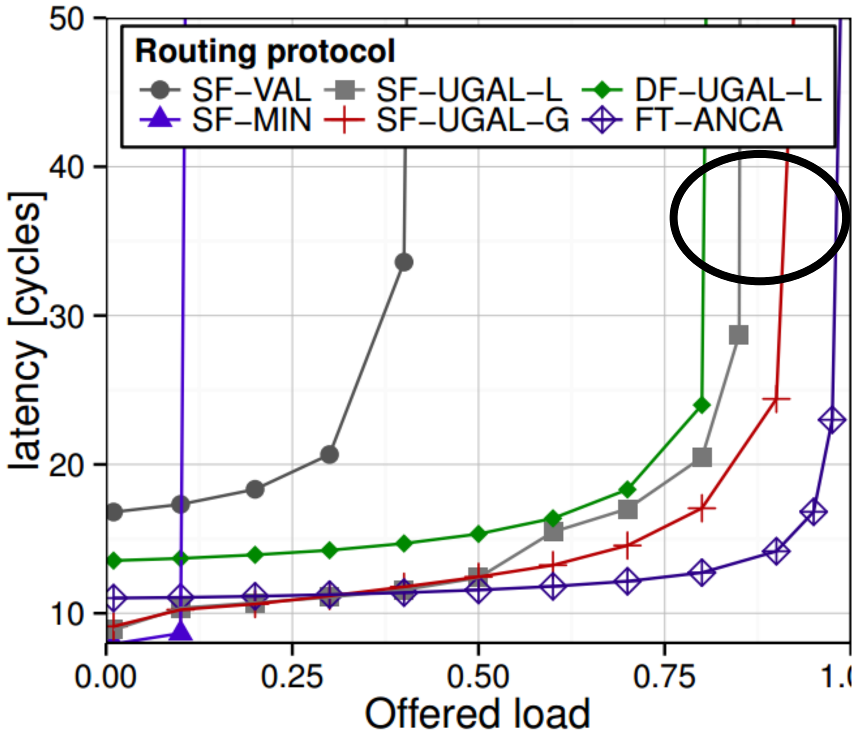
- Shift traffic



dest id
source id

$$d = \left(s \bmod \frac{N}{2}\right) + \frac{N}{2}$$

$$d = s \bmod \frac{N}{2}$$



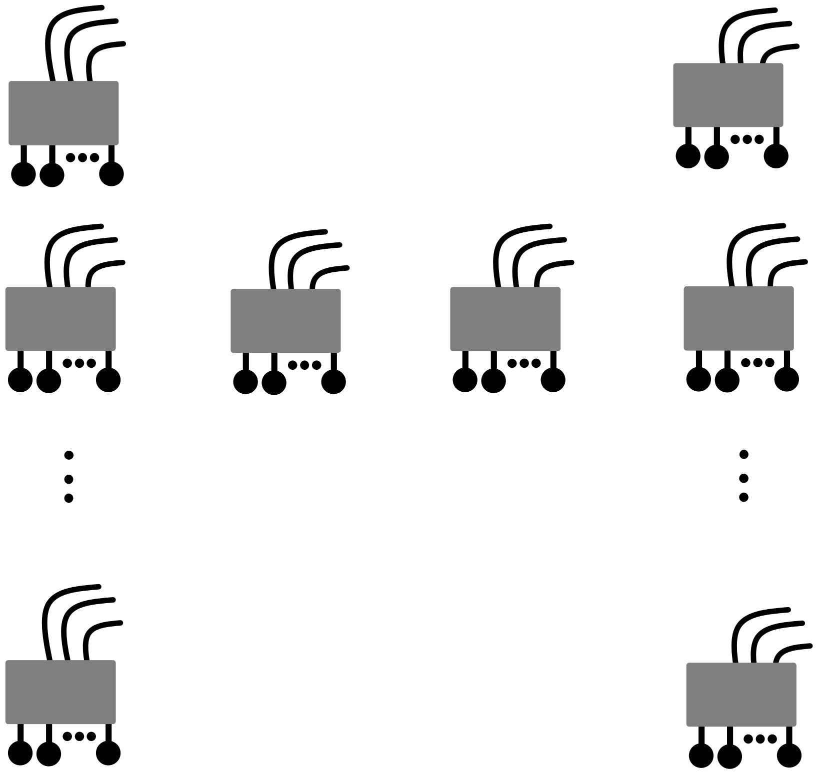
PERFORMANCE

PERFORMANCE

- Worst-case traffic

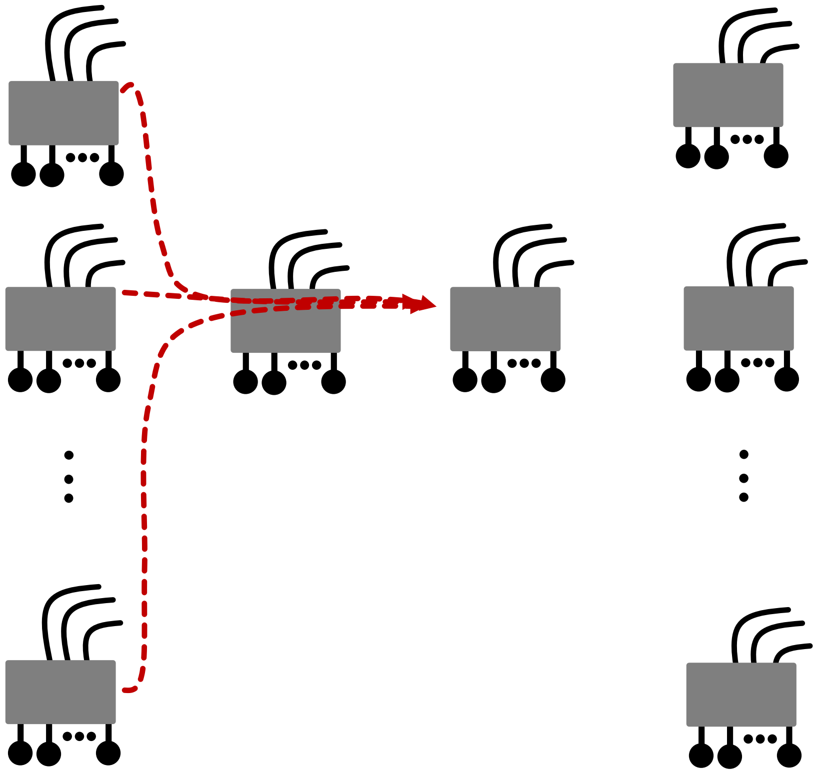
PERFORMANCE

- Worst-case traffic



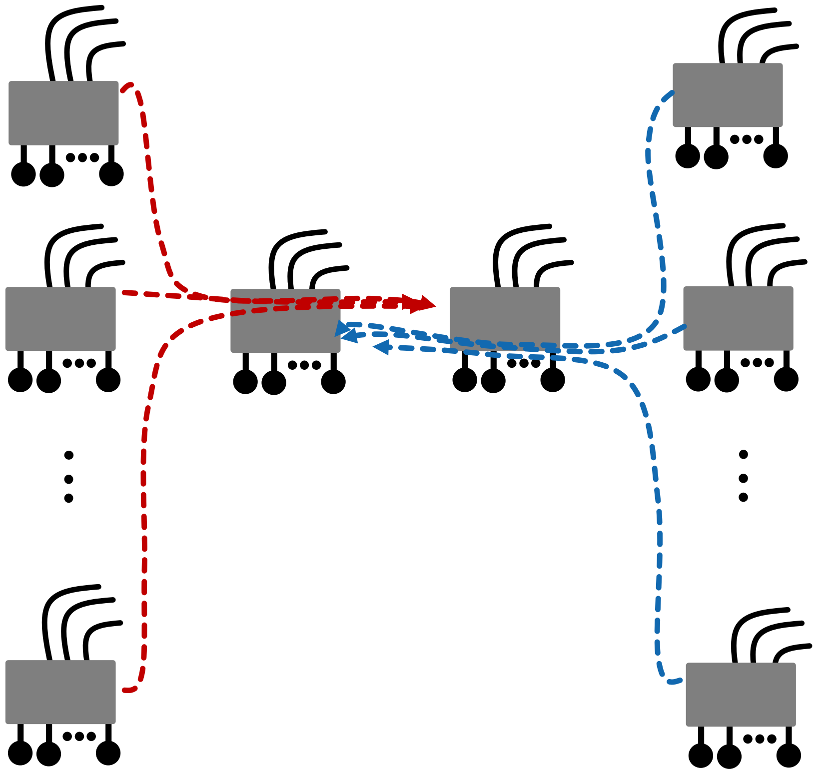
Performance

- Worst-case traffic



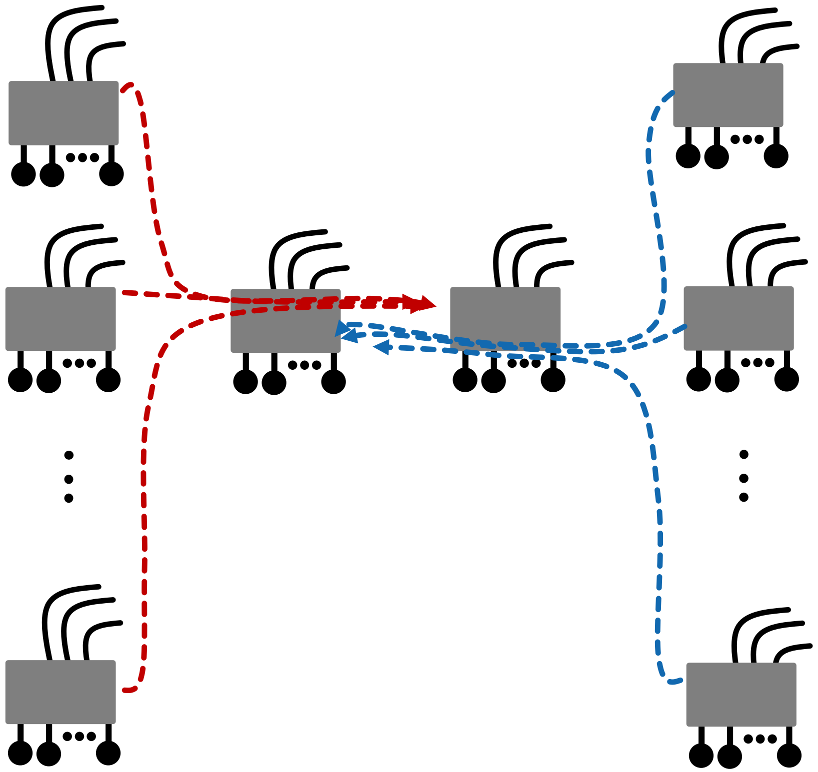
Performance

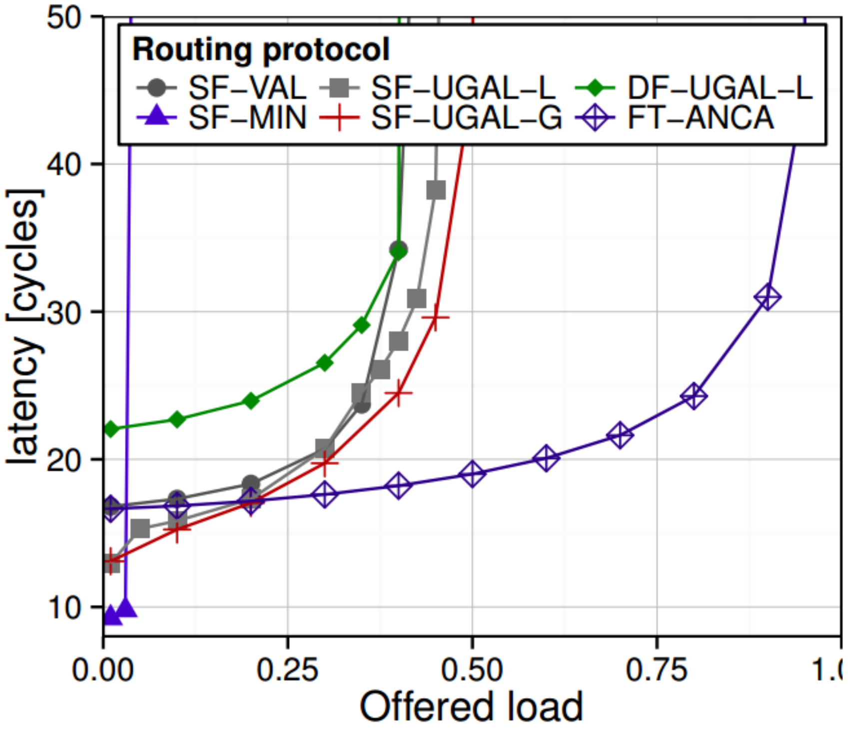
- Worst-case traffic



PERFORMANCE

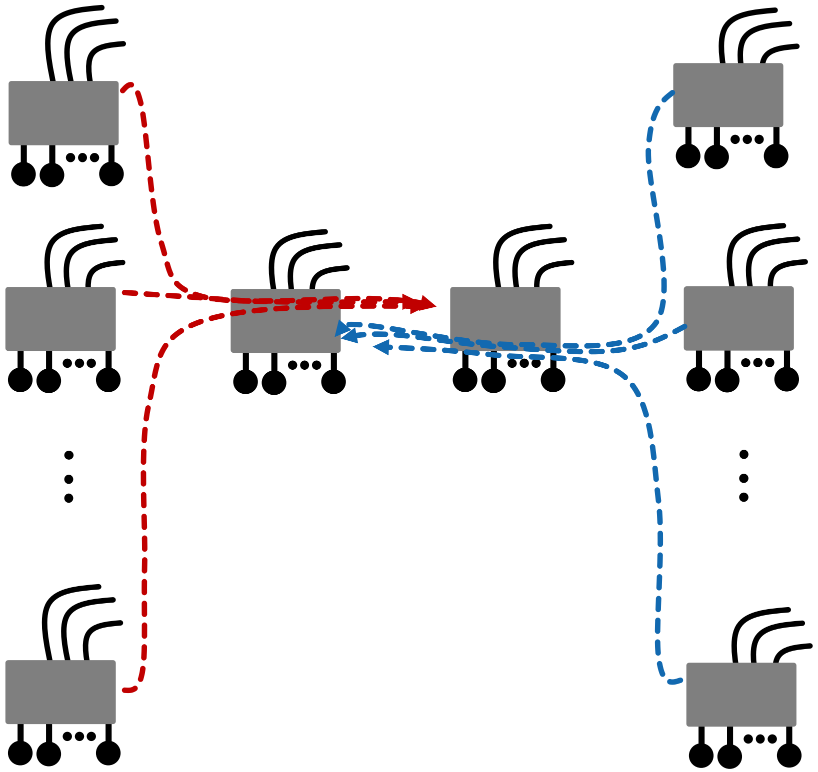
- Worst-case traffic

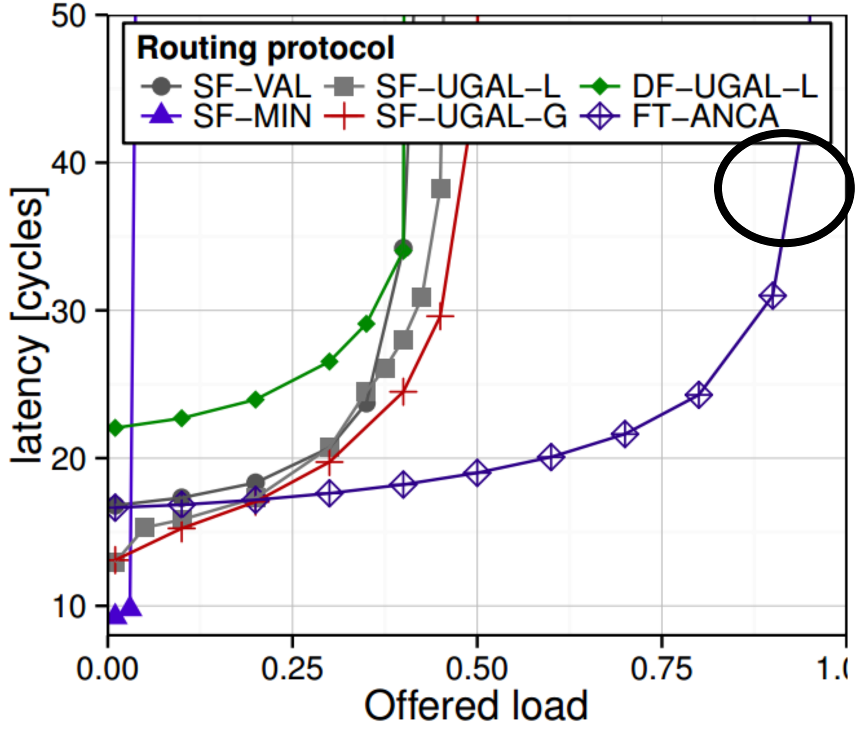




Performance

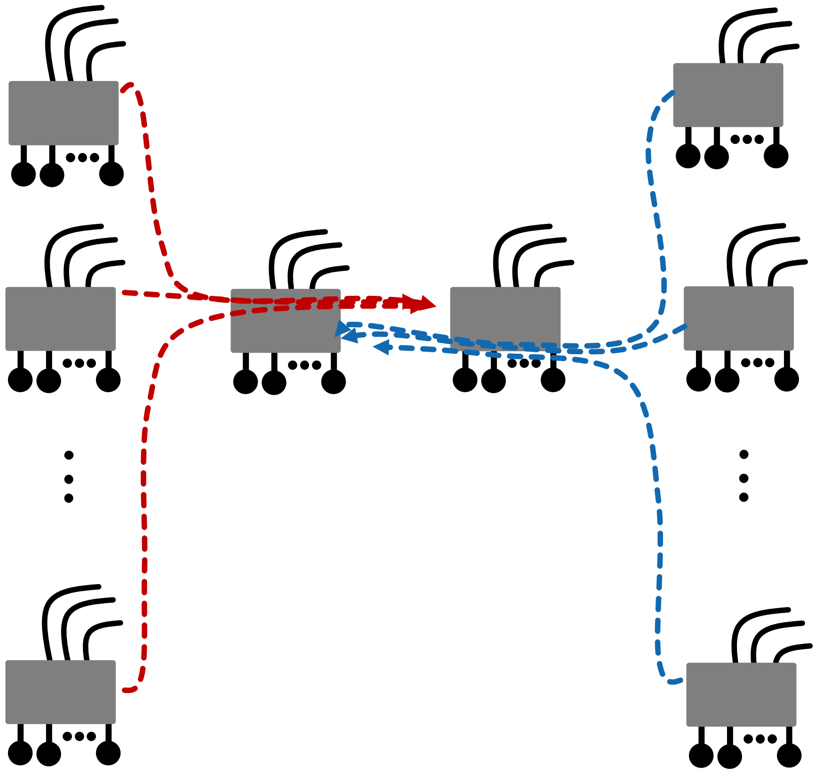
- Worst-case traffic

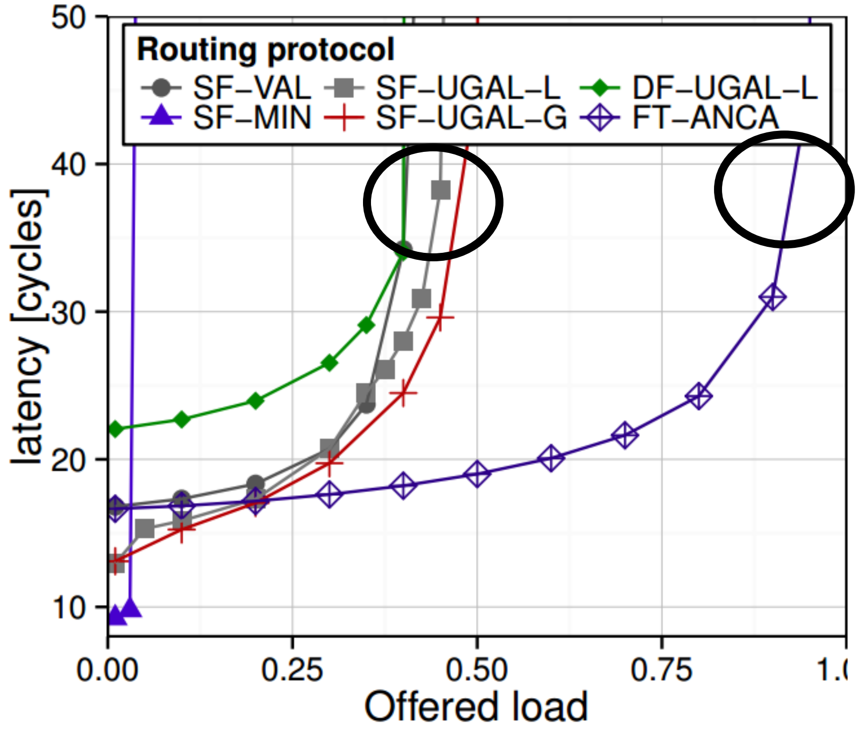




Performance

- Worst-case traffic





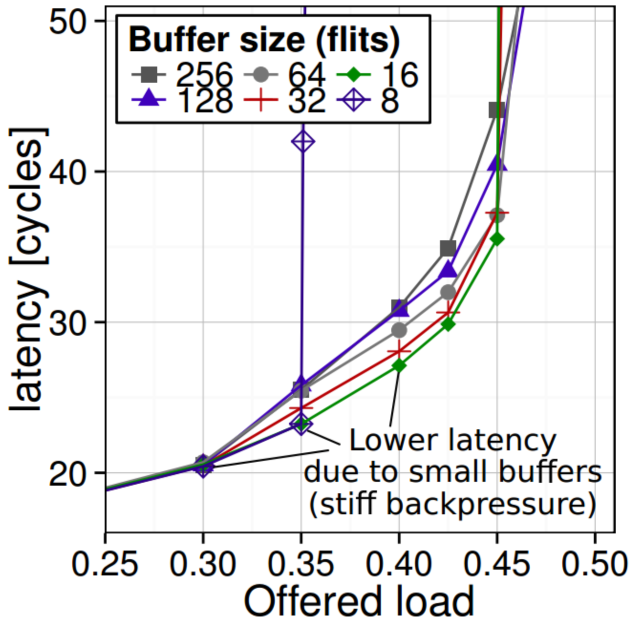
PERFORMANCE

Performance

- Buffer sizes (UGAL-L, worst-case traffic)

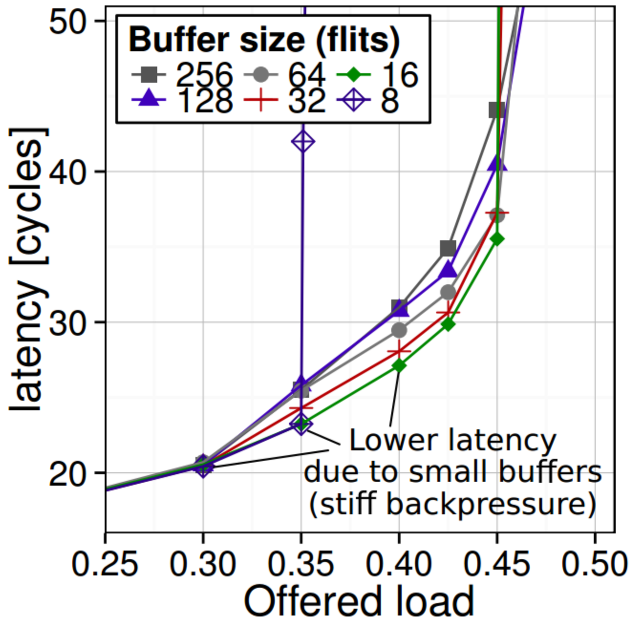
PERFORMANCE

- Buffer sizes (UGAL-L, worst-case traffic)



Performance

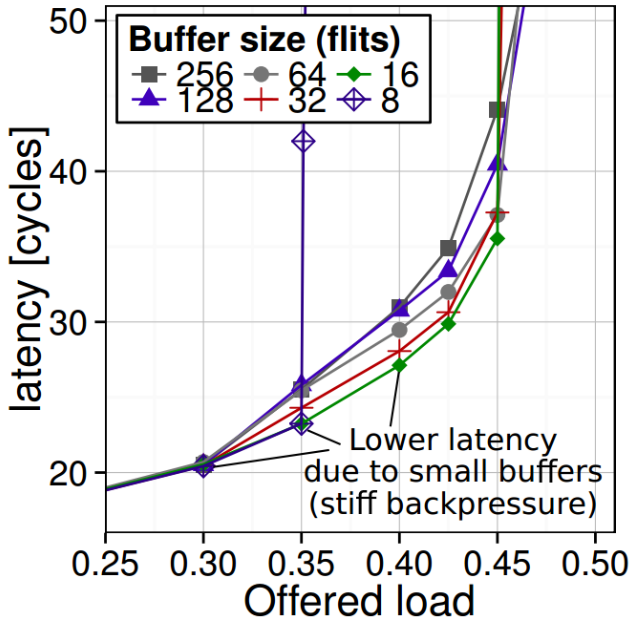
- Buffer sizes (UGAL-L, worst-case traffic)
- Oversubscription (64 flits)



Performance

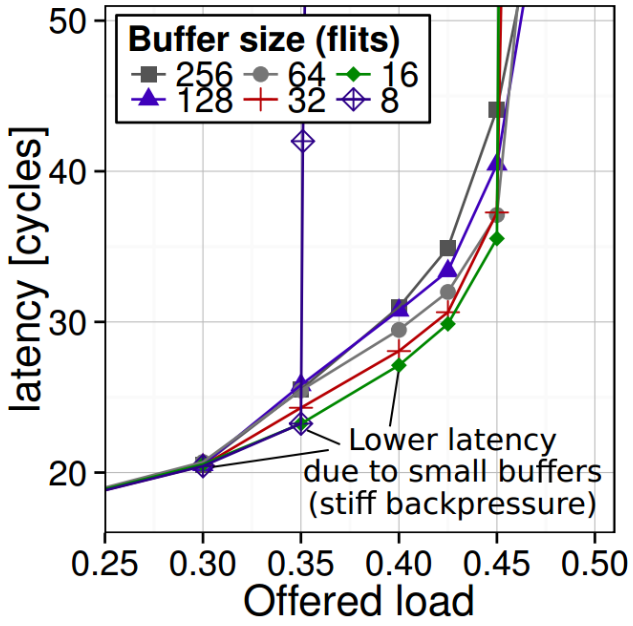
- Buffer sizes (UGAL-L, worst-case traffic)
- Oversubscription (64 flits)

p=16

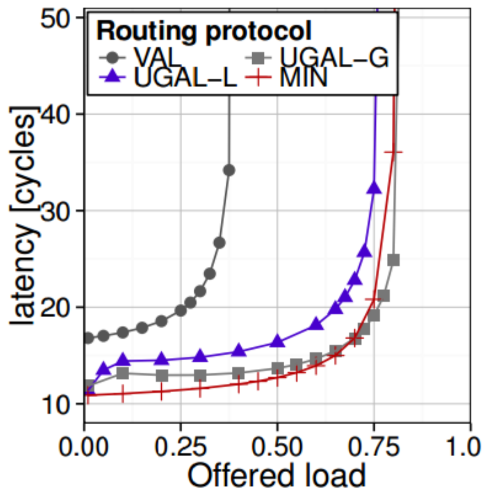


Performance

- Buffer sizes (UGAL-L, worst-case traffic)
- Oversubscription (64 flits)



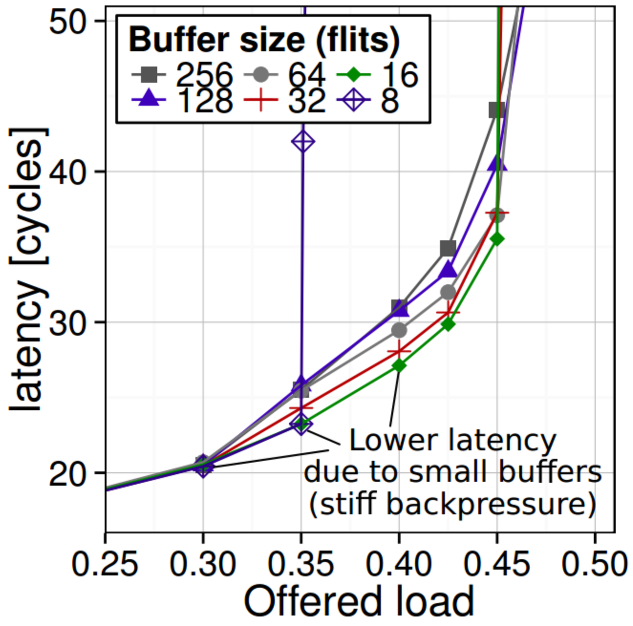
p=16



(b) Random traffic, $p = 16$.

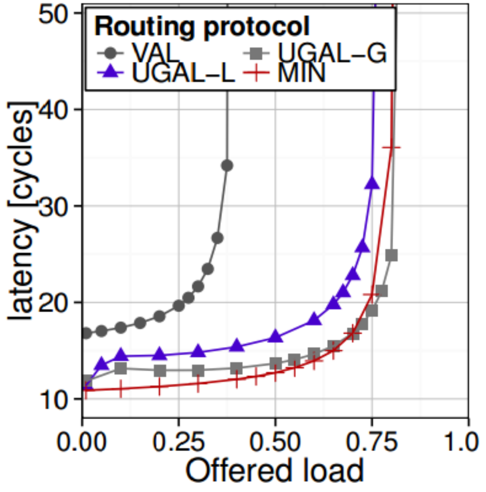
PERFORMANCE

- Buffer sizes (UGAL-L, worst-case traffic)
- Oversubscription (64 flits)



p=16

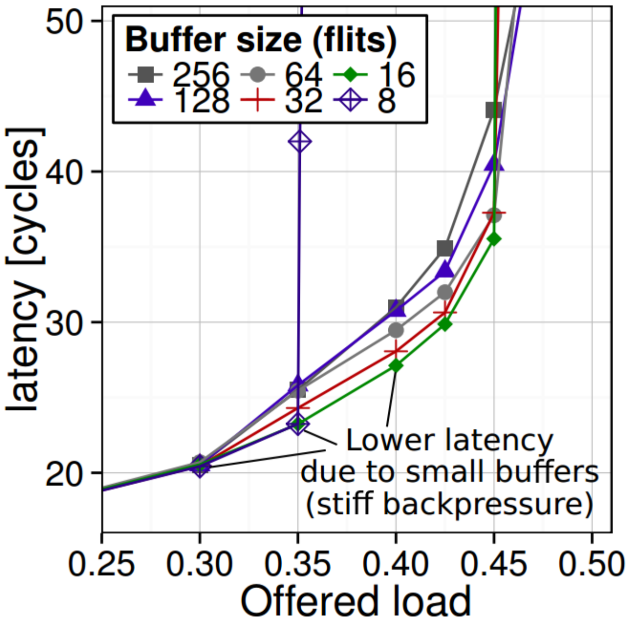
p=18



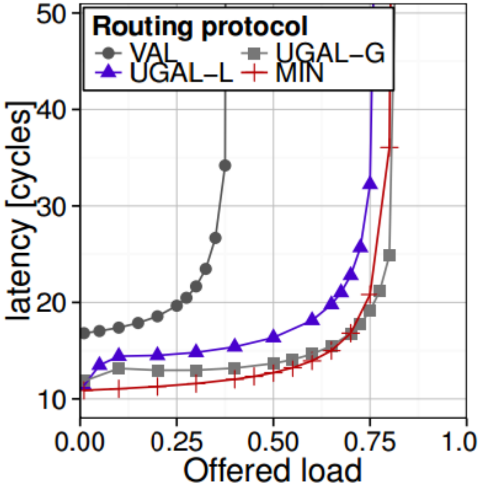
(b) Random traffic, $p = 16$.

Performance

- Buffer sizes (UGAL-L, worst-case traffic)
- Oversubscription (64 flits)

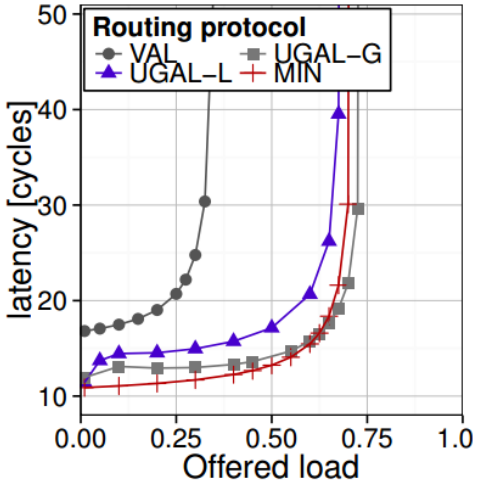


p=16



(b) Random traffic, $p = 16$.

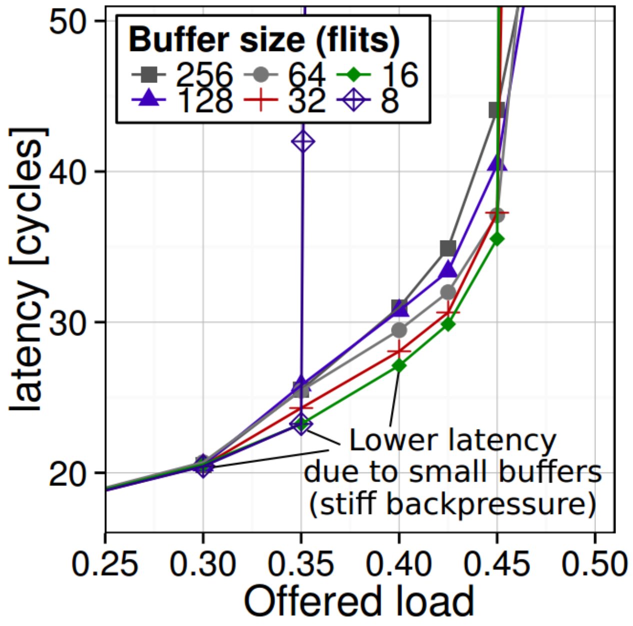
p=18



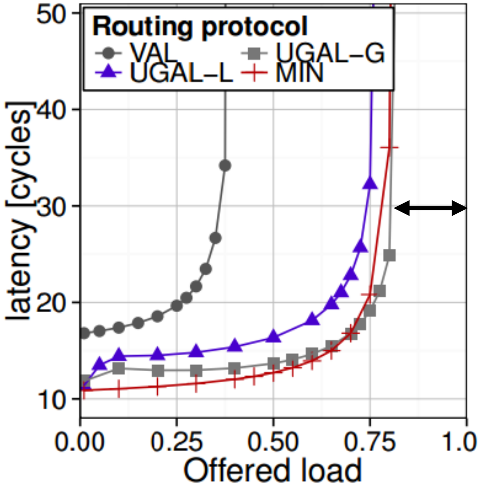
(d) Random traffic, $p = 18$.

PERFORMANCE

- Buffer sizes (UGAL-L, worst-case traffic)
- Oversubscription (64 flits)

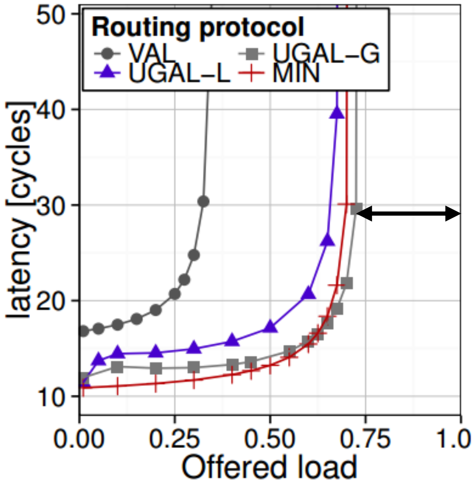


p=16



(b) Random traffic, $p = 16$.

p=18



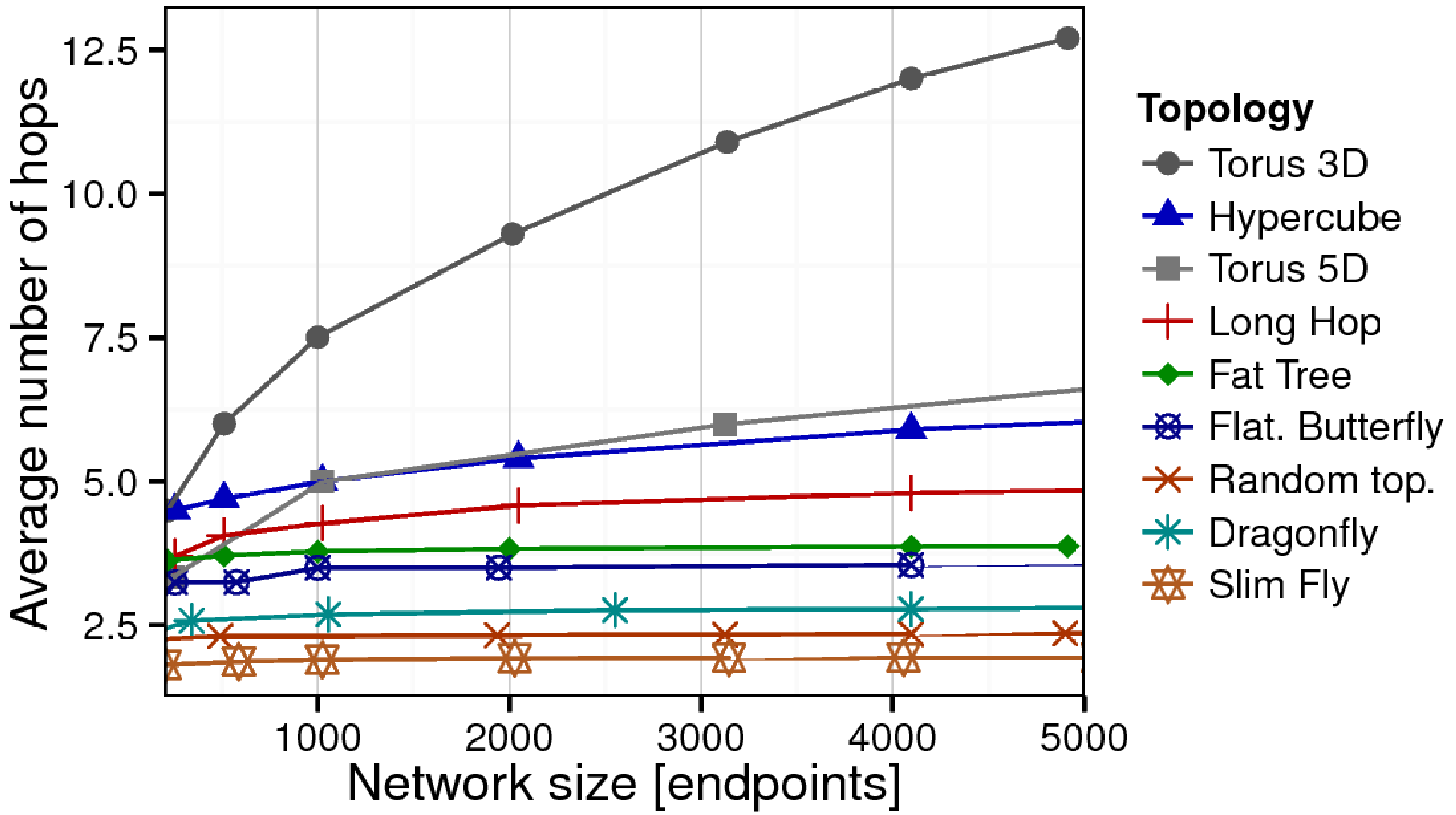
(d) Random traffic, $p = 18$.

STRUCTURE ANALYSIS

AVERAGE DISTANCE

Structure Analysis

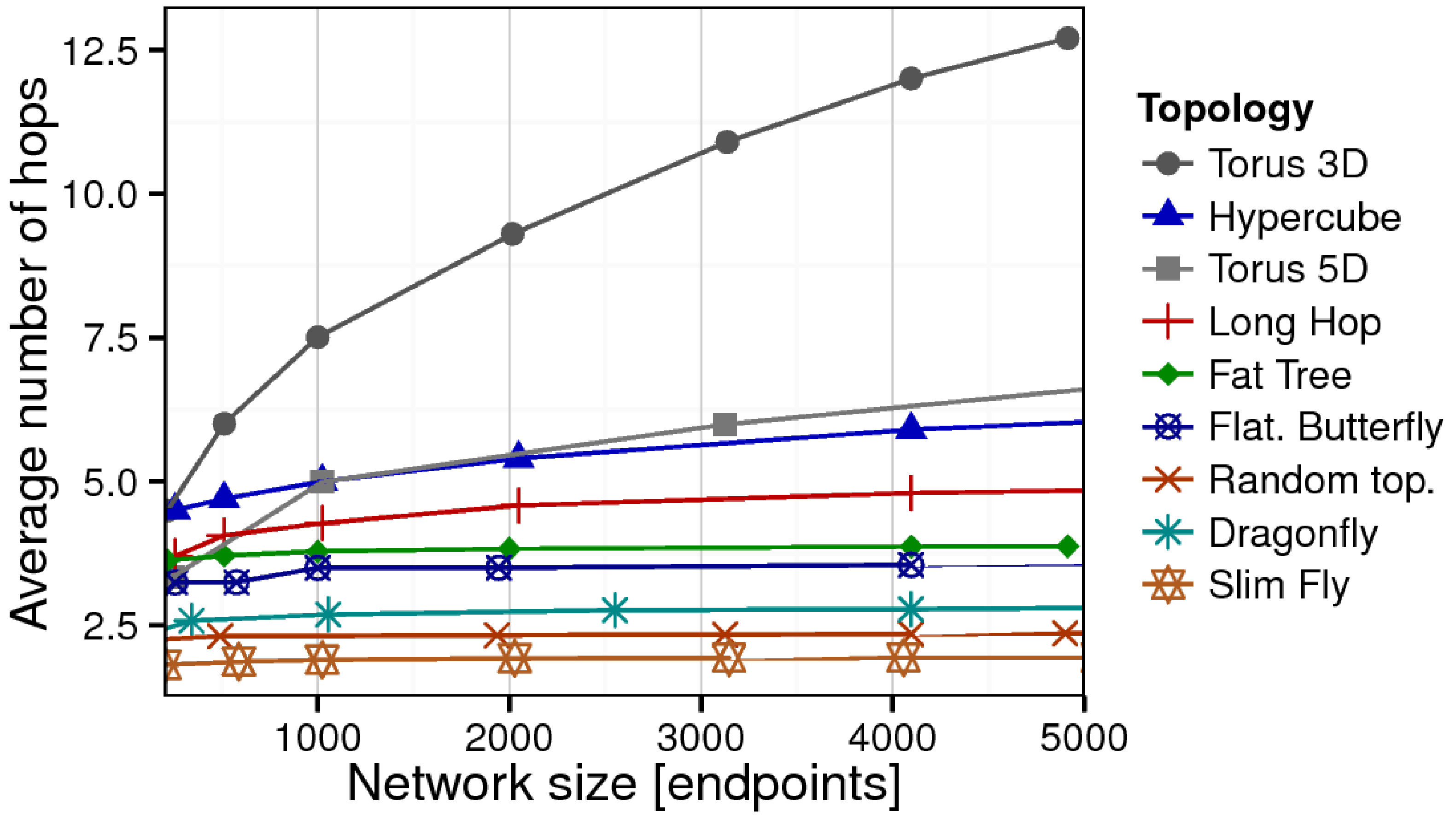
Average Distance



Structure Analysis

Average Distance

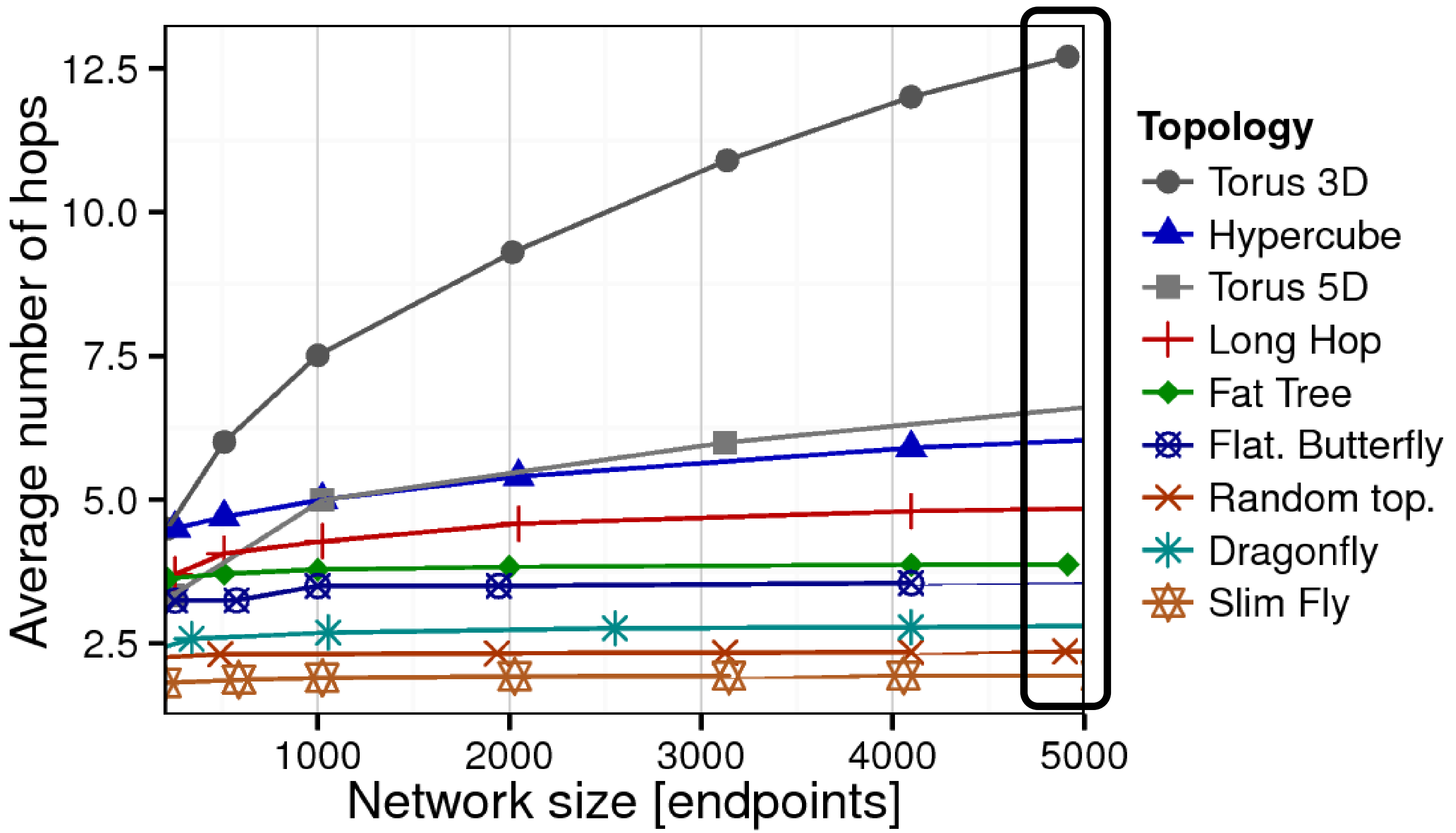
Random uniform traffic
using minimum path routing



Structure Analysis

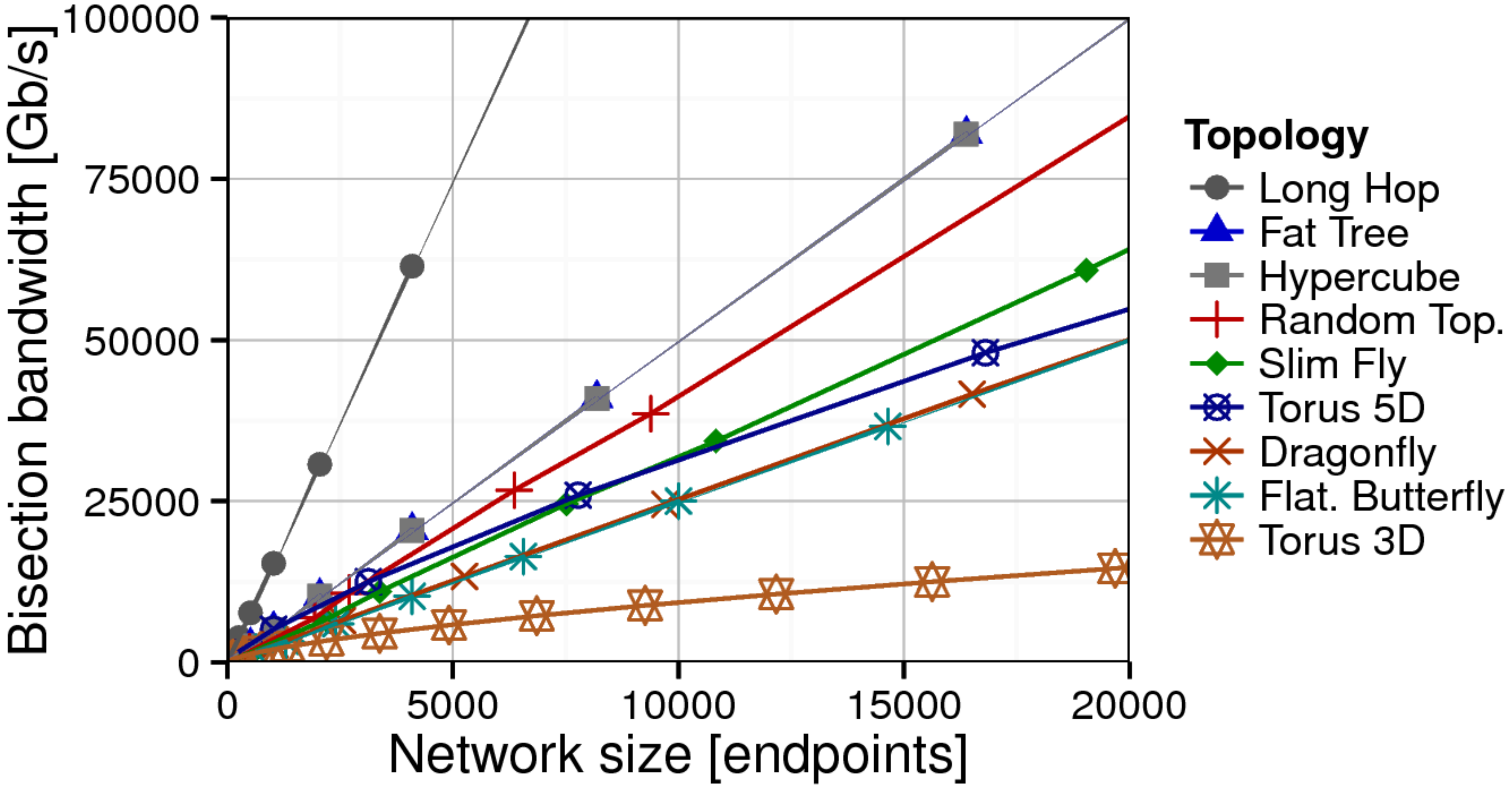
Average Distance

Random uniform traffic
using minimum path routing



Structure Analysis

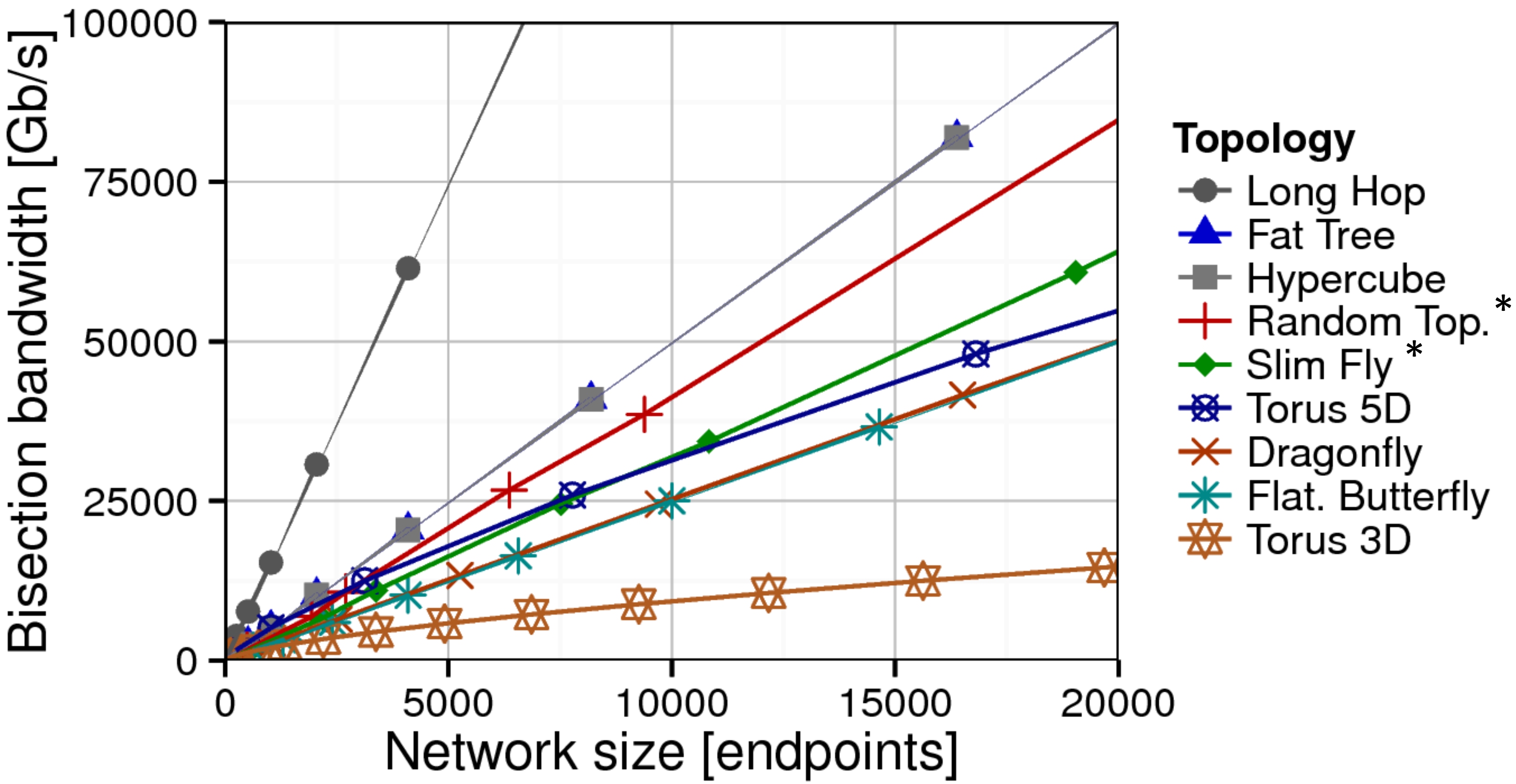
Bisection Bandwidth (BB)



Structure Analysis

Bisection Bandwidth (BB)

*BB approximated with the Metis partitioner [1]

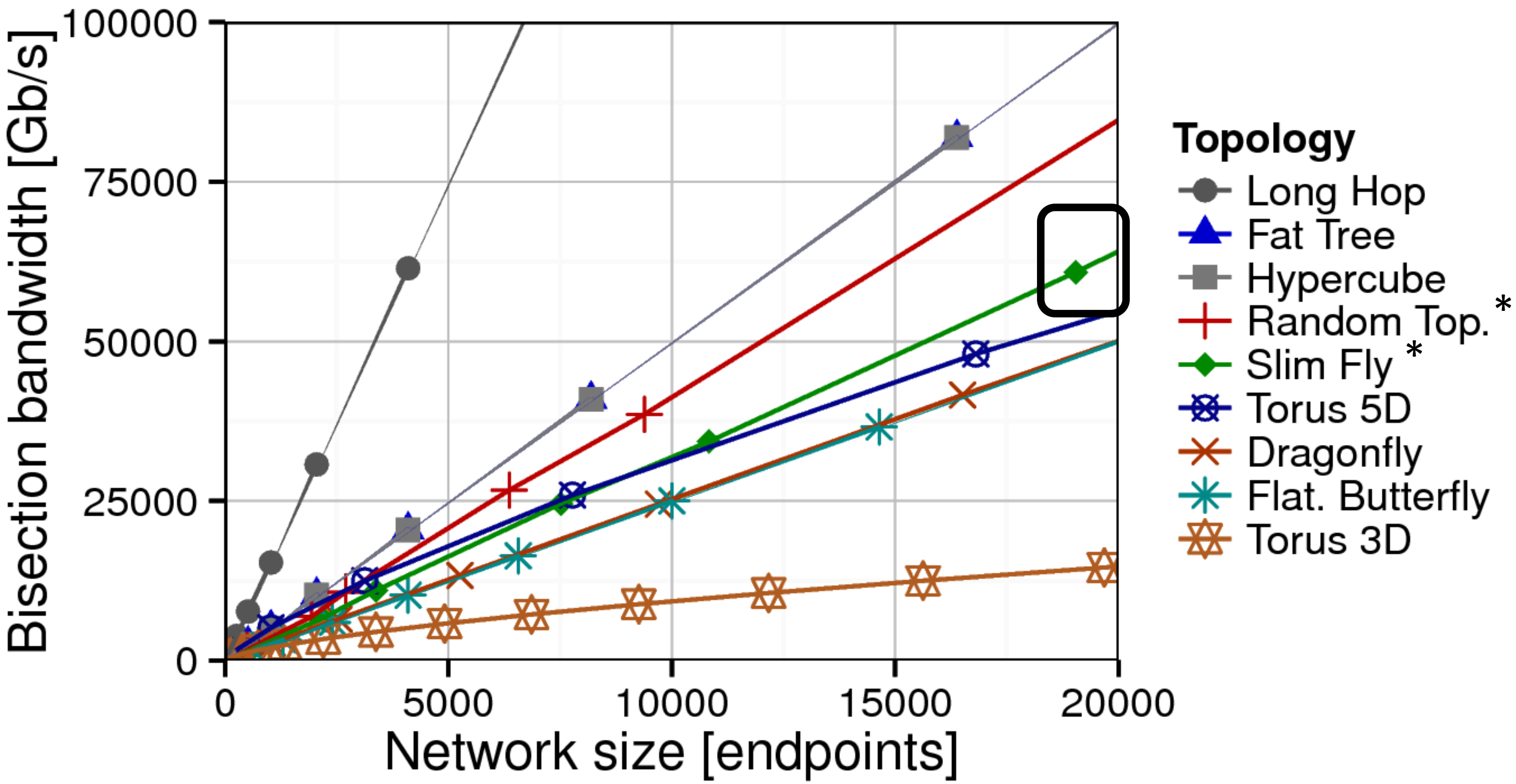


[1] G. Karypis, V. Kumar. A fast and high quality multilevel scheme for partitioning irregular graphs. ICPP'95

STRUCTURE ANALYSIS

BISECTION BANDWIDTH (BB)

*BB approximated with the Metis partitioner [1]



[1] G. Karypis, V. Kumar. A fast and high quality multilevel scheme for partitioning irregular graphs. ICPP'95