String Figure: A Scalable and Elastic Memory Network Architecture

Matheus Ogleari

Ye Yu, Chen Qian, Ethan Miller, Jishen Zhao

HPCA 2019

UC San Diego

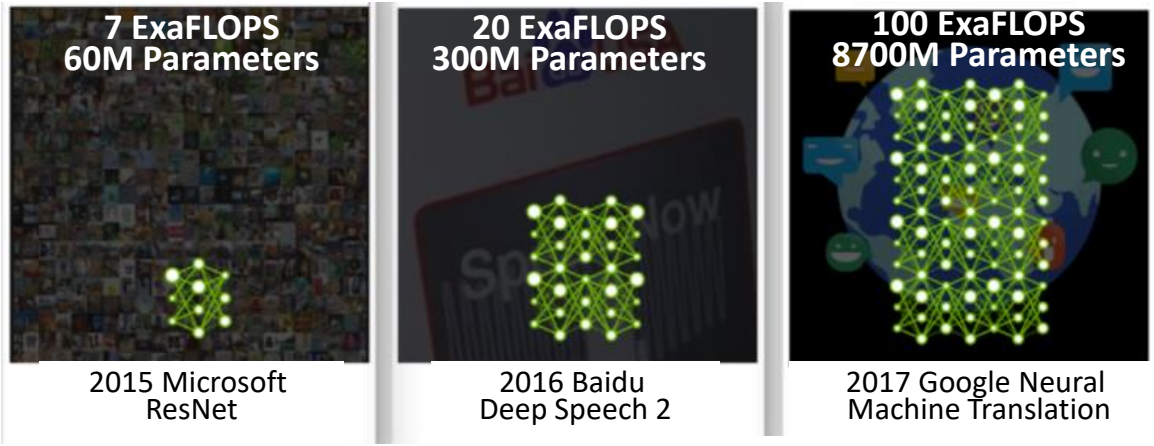
UNIVERSITY OF CALIFORNIA SANTA CRUZ

PURESTORAGE

ssrc STORAGE SYSTEMS RESEARCH CENTER

Issues with Memory Capacity Bottleneck

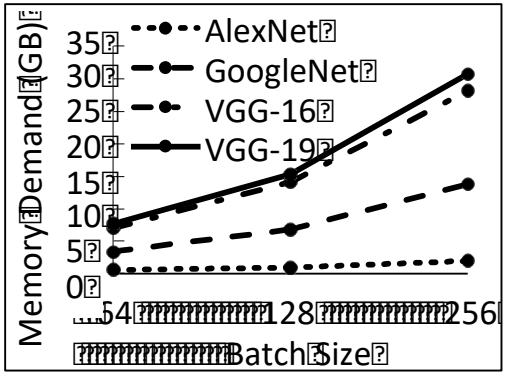
Training neural networks is data intensive, and increasingly so.

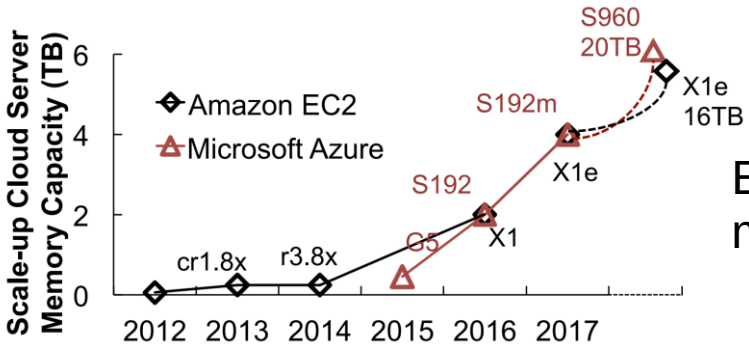


2015 Microsoft ResNet

2016 Baidu Deep Speech 2

2017 Google Neural Machine Translation

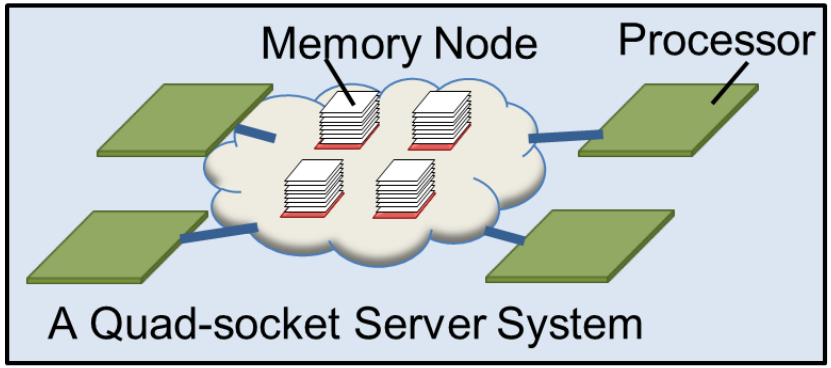




Expanding server memory demand

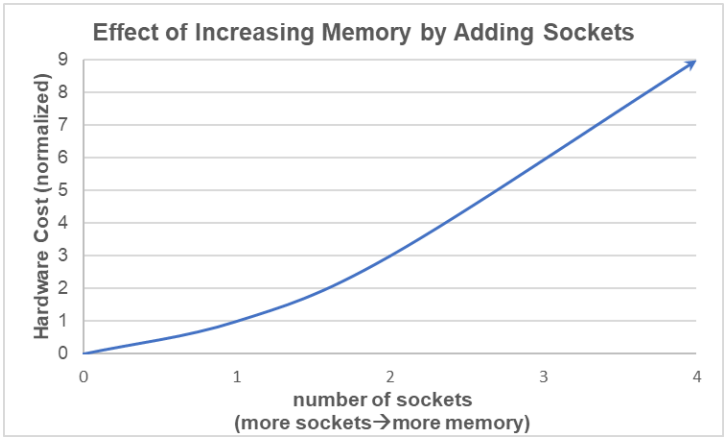
Why Memory Network

- **Memory Networks**

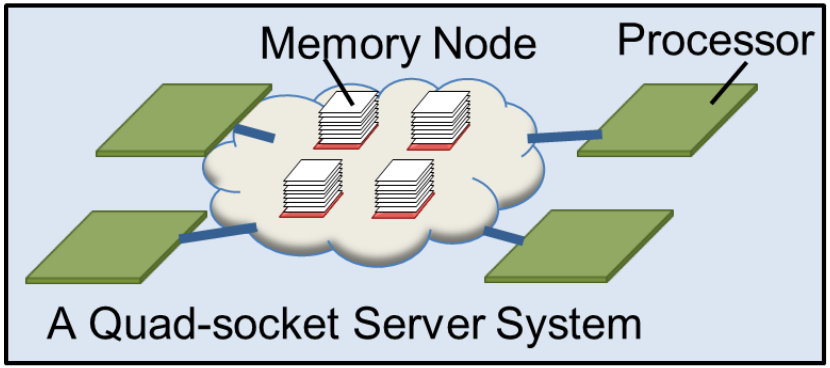


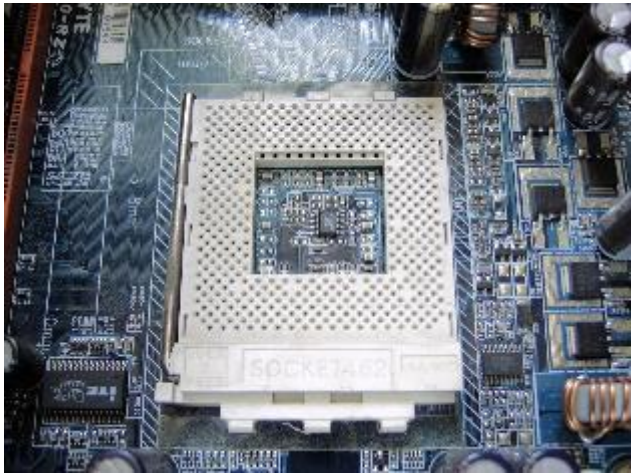
Why Memory Network

- **Memory Networks**
- **More sockets, more problems**

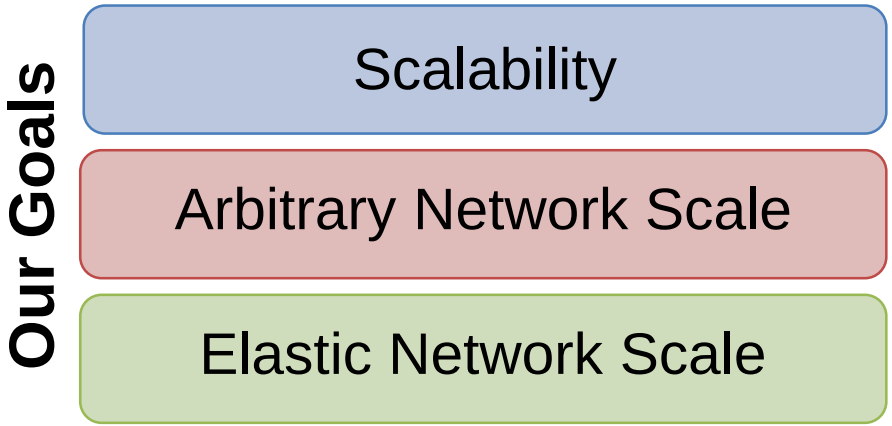


[Zhao, TACO 2015]





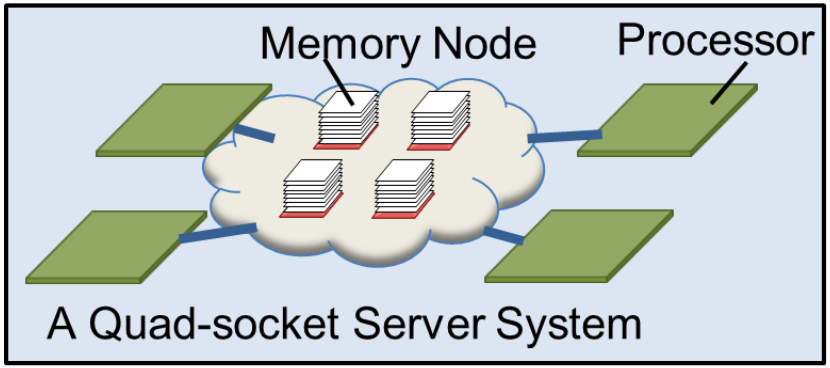
Design Goals



Design Goals

Our Goals

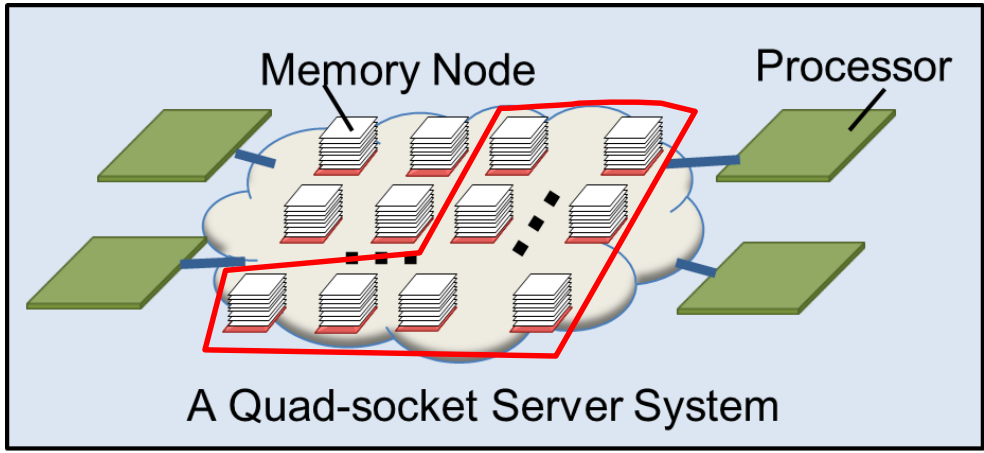
- Scalability
- Arbitrary Network Scale
- Elastic Network Scale



Design Goals

Our Goals

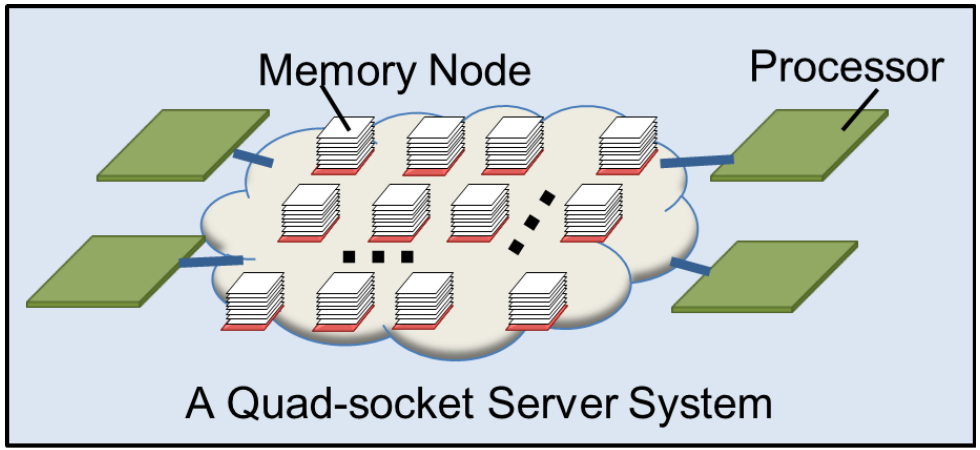
- Scalability
- Arbitrary Network Scale
- Elastic Network Scale



Design Goals

Our Goals

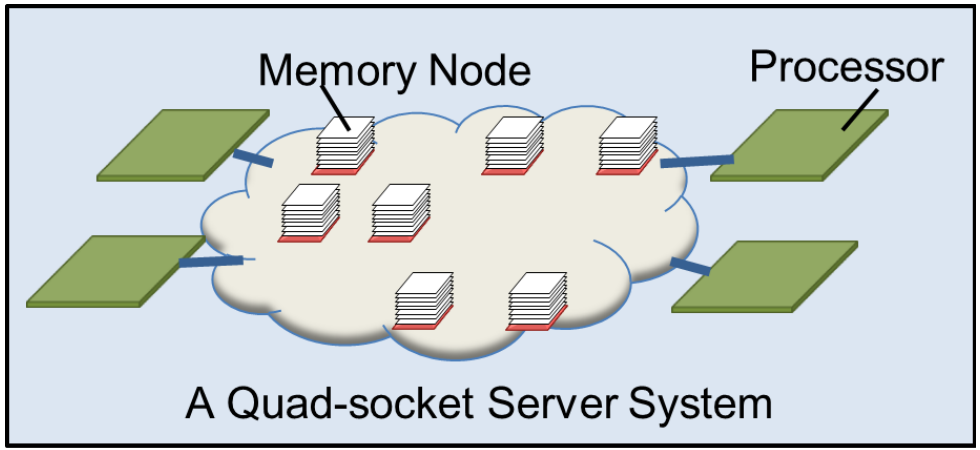
- Scalability
- **Arbitrary Network Scale**
- Elastic Network Scale



Design Goals

Our Goals

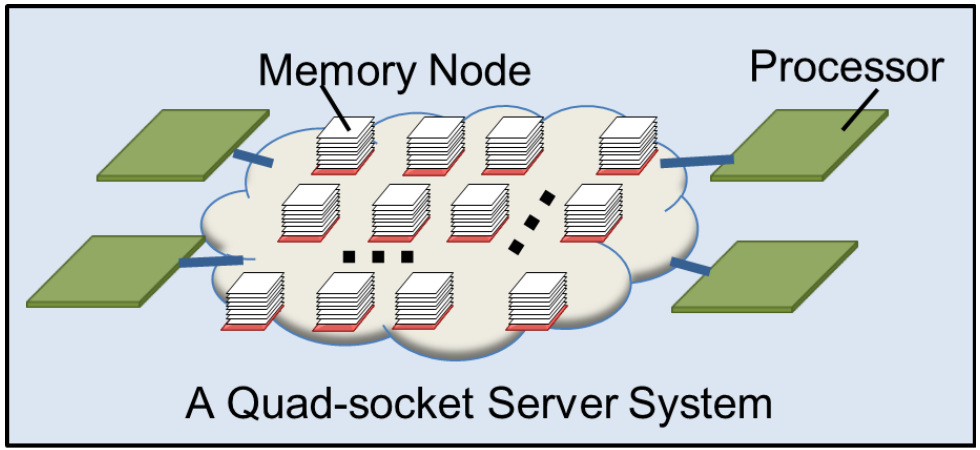
- Scalability
- **Arbitrary Network Scale**
- Elastic Network Scale



Design Goals

Our Goals

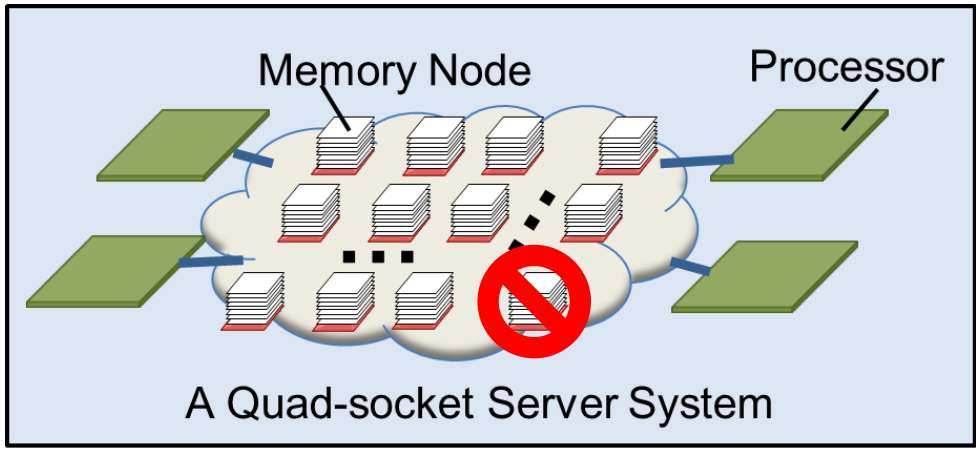
- Scalability
- Arbitrary Network Scale
- Elastic Network Scale



Design Goals

Our Goals

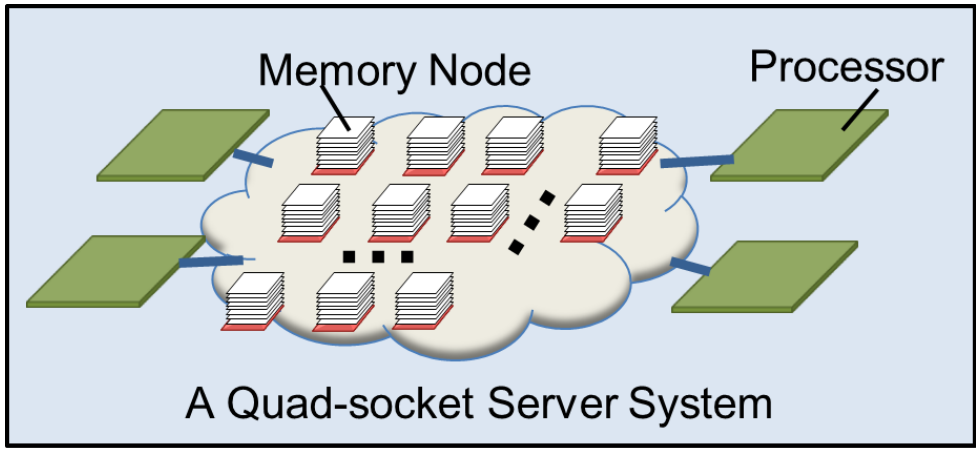
- Scalability
- Arbitrary Network Scale
- Elastic Network Scale



Design Goals

Our Goals

- Scalability
- Arbitrary Network Scale
- Elastic Network Scale

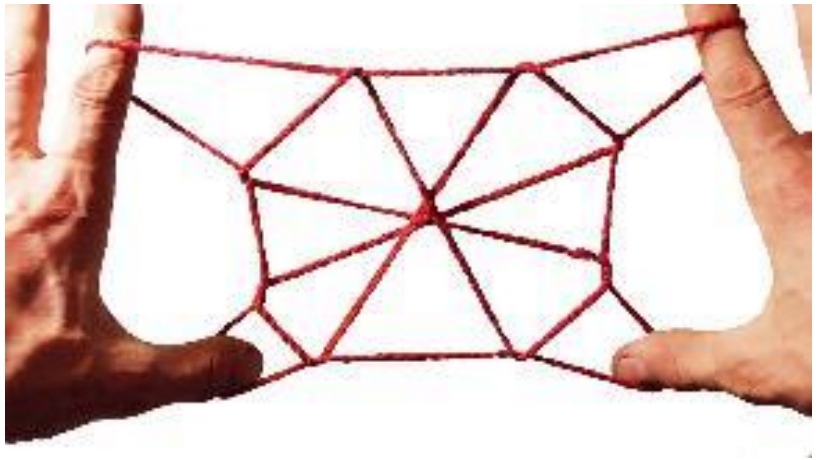


String Figure:

A Scalable and Elastic Memory Network Design

Our Goals

- Scalability
- Arbitrary Network Scale
- Elastic Network Scale



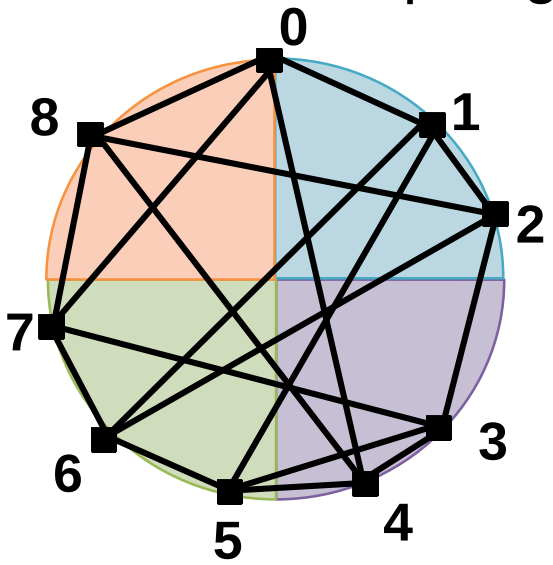
Topology design

Topology is a function of number of router ports

Our Goals

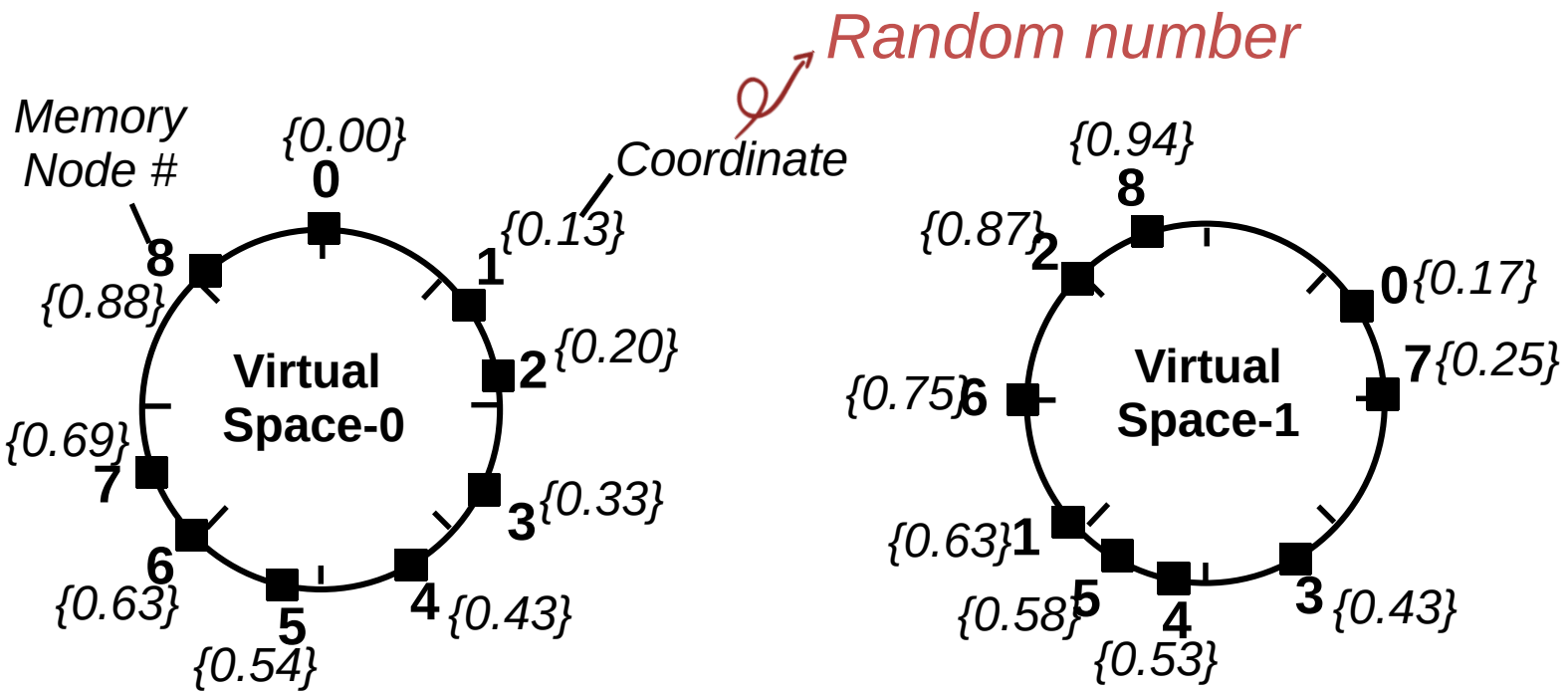
- Scalability
- Arbitrary Network Scale
- Elastic Network Scale

Random Topology



Random topology generation

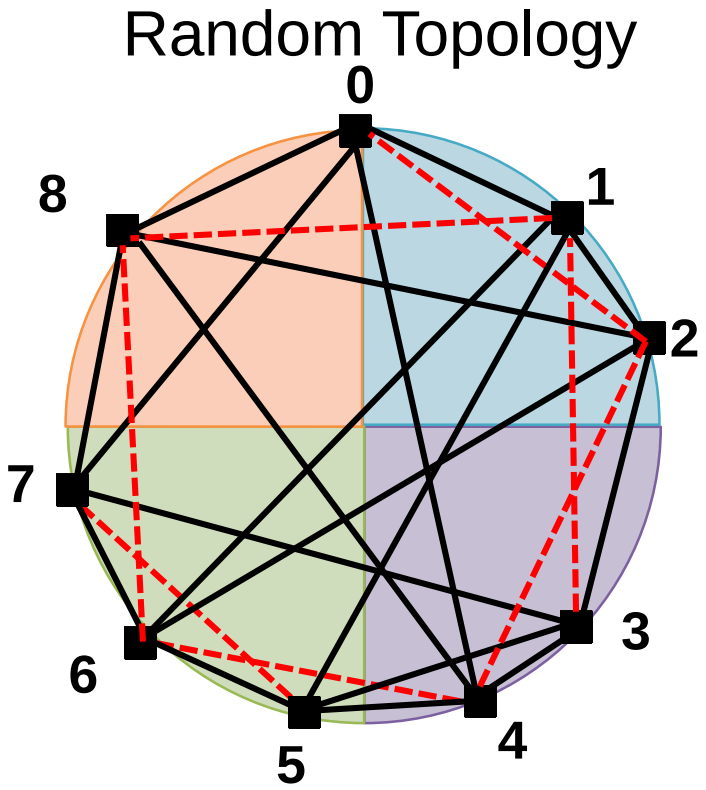
Number of spaces = Number of ports ÷ 2



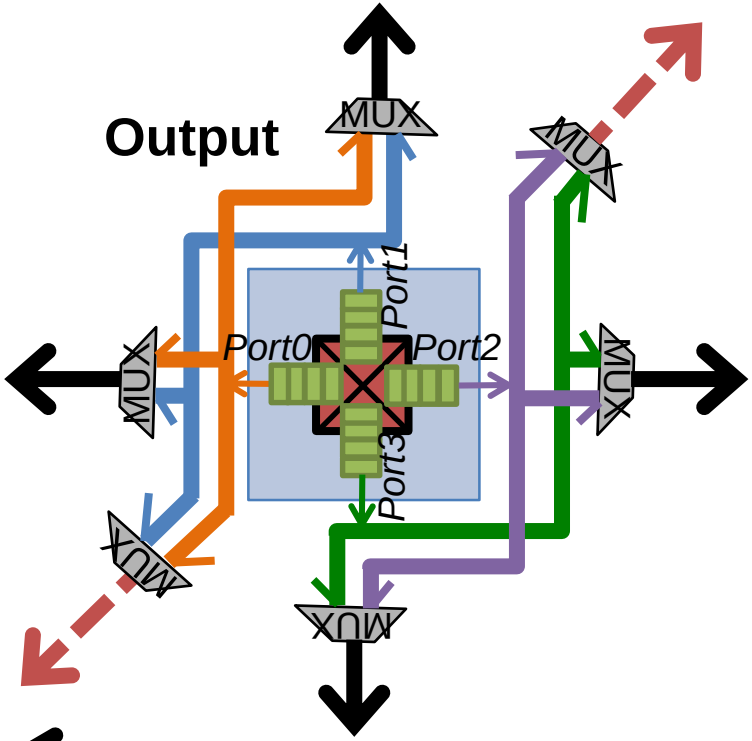
Short cuts

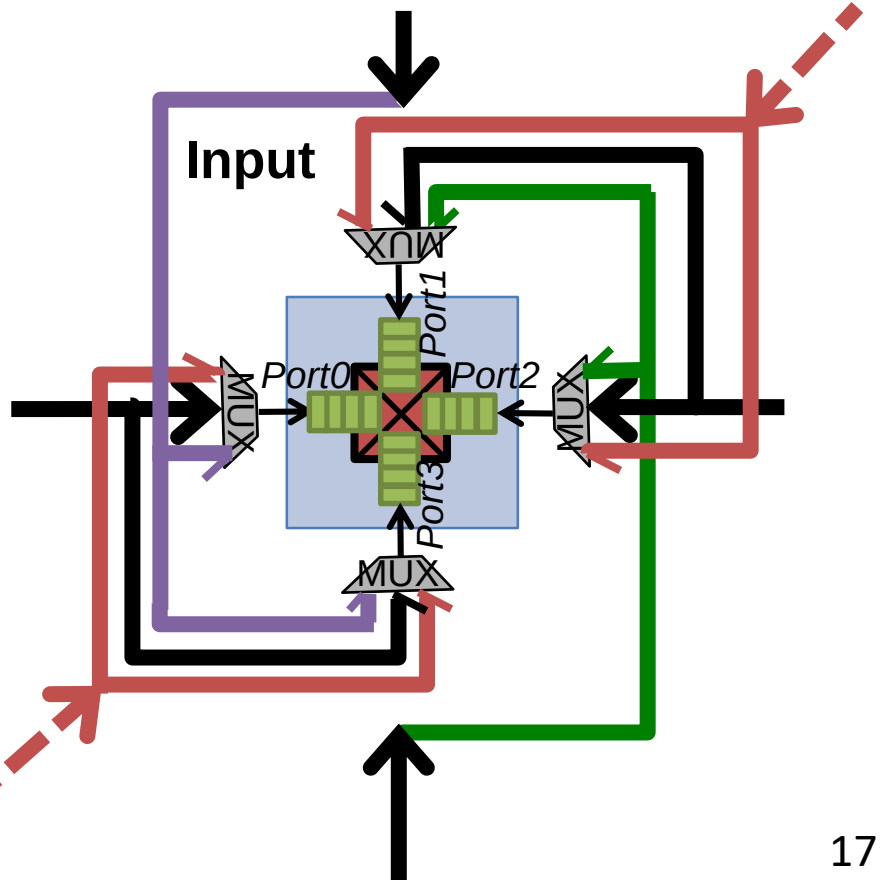
Our Goals

- Scalability
- Arbitrary Network Scale
- Elastic Network Scale



Reconfigurable router





Random Connections
Short Cuts

Greedy routing protocol

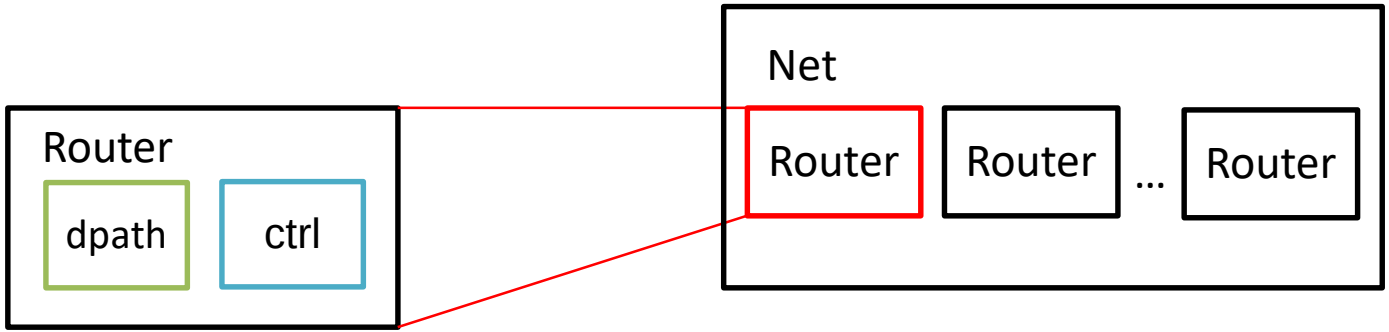
Node #	D in Space-0	D in Space-1	MD
7	0.49	0.62	0.49
0	0.20	0.70	0.20
3	0.13	0.44	0.13
6	0.43	0.12	0.12
8	0.68	0.07	0.07

Minimum distance (MD) to Node-2 from Node-7 and Node-7's neighbors

Node#	Block.	Valid	Hop	Space#	Coordi.
0	1	1	0	0	0.00
2	1	1	0	0	0.20
...					
5	1	1	0	1	0.58
6	1	1	0	1	0.75
...					
3	1	1	1	0	0.33
8	1	1	1	0	0.88
...					

Routing table entries

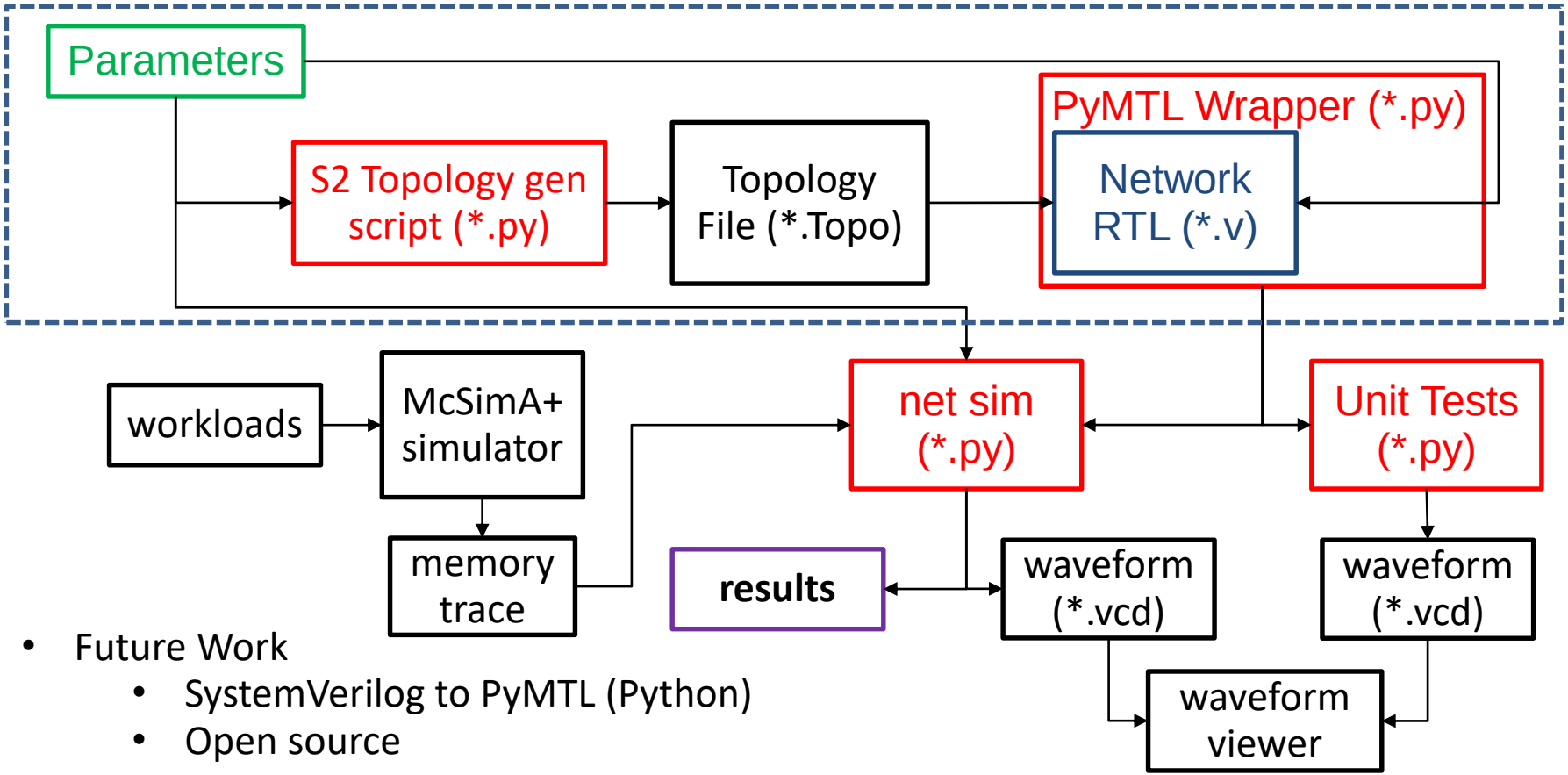
Implementation



Design Parameters:

- Payload Size (bits)
- # of net ports
- # of routers
- # of router ports

Simulation



- Future Work
  - SystemVerilog to PyMTL (Python)
  - Open source

Evaluation

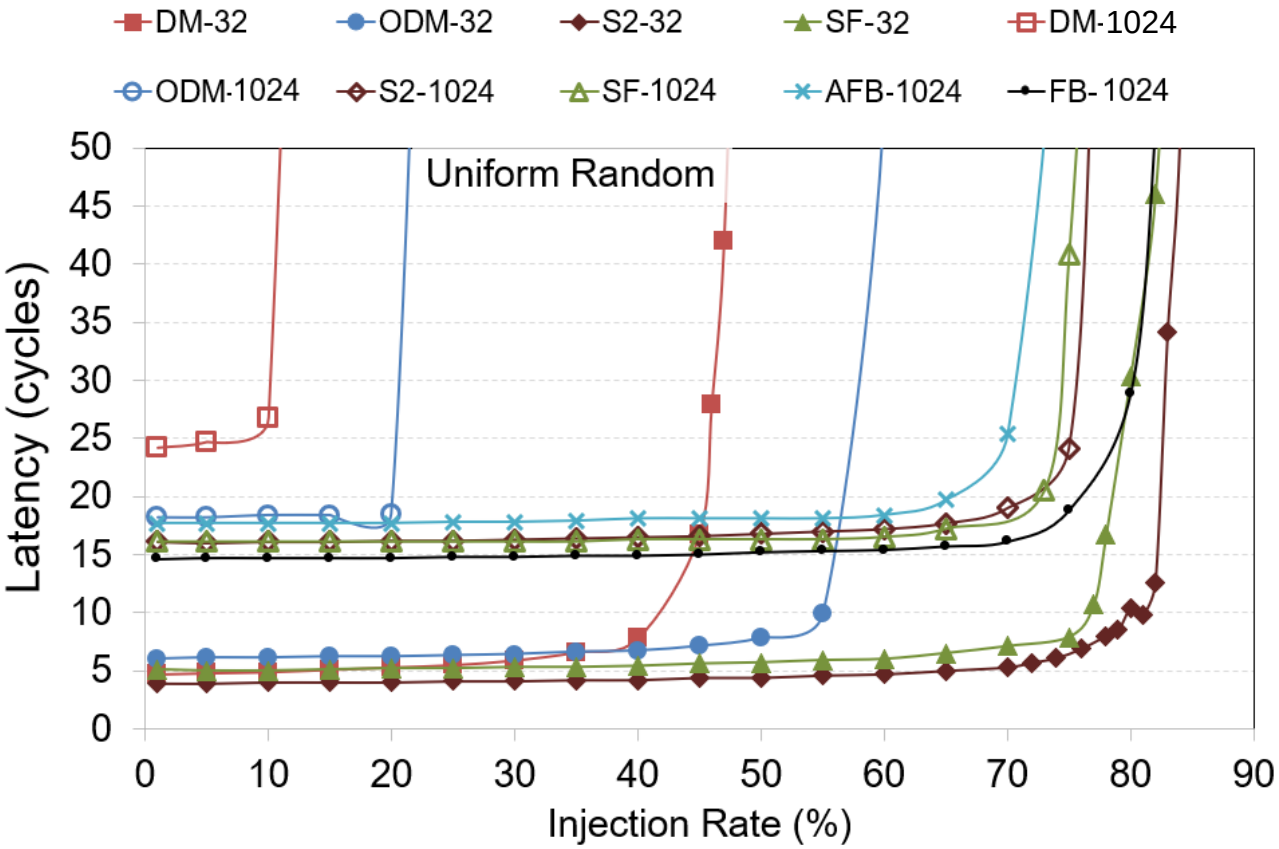
Topology	Number of Nodes (N), Number of Ports per Router (p)											Routing Scheme
	N	**16**	**17**	**32**	**61**	**64**	**113**	**128**	**256**	**512**	**1024**	**1296**	
Distributed-Mesh (DM) / Optimized DM (ODM)	**p**	4	n/a	4	n/a	4	n/a	4	4	4	4	4	Greedy + adaptive
Flattened Butterfly (FB)	**p**								20	24	31	33	minimal + adaptive
Adaptive FB (AFB)	**p**								13	17	23	25	minimal + adaptive
Space Shuffle Ideal (S2-ideal)	**p**	4	4	4	4	4	4	4	8	8	8	8	look-up table (LUT)
String Figure (SF)	**p**	4	4	4	4	4	4	4	8	8	8	8	LUT + greedy + adaptive

Experiment Setup

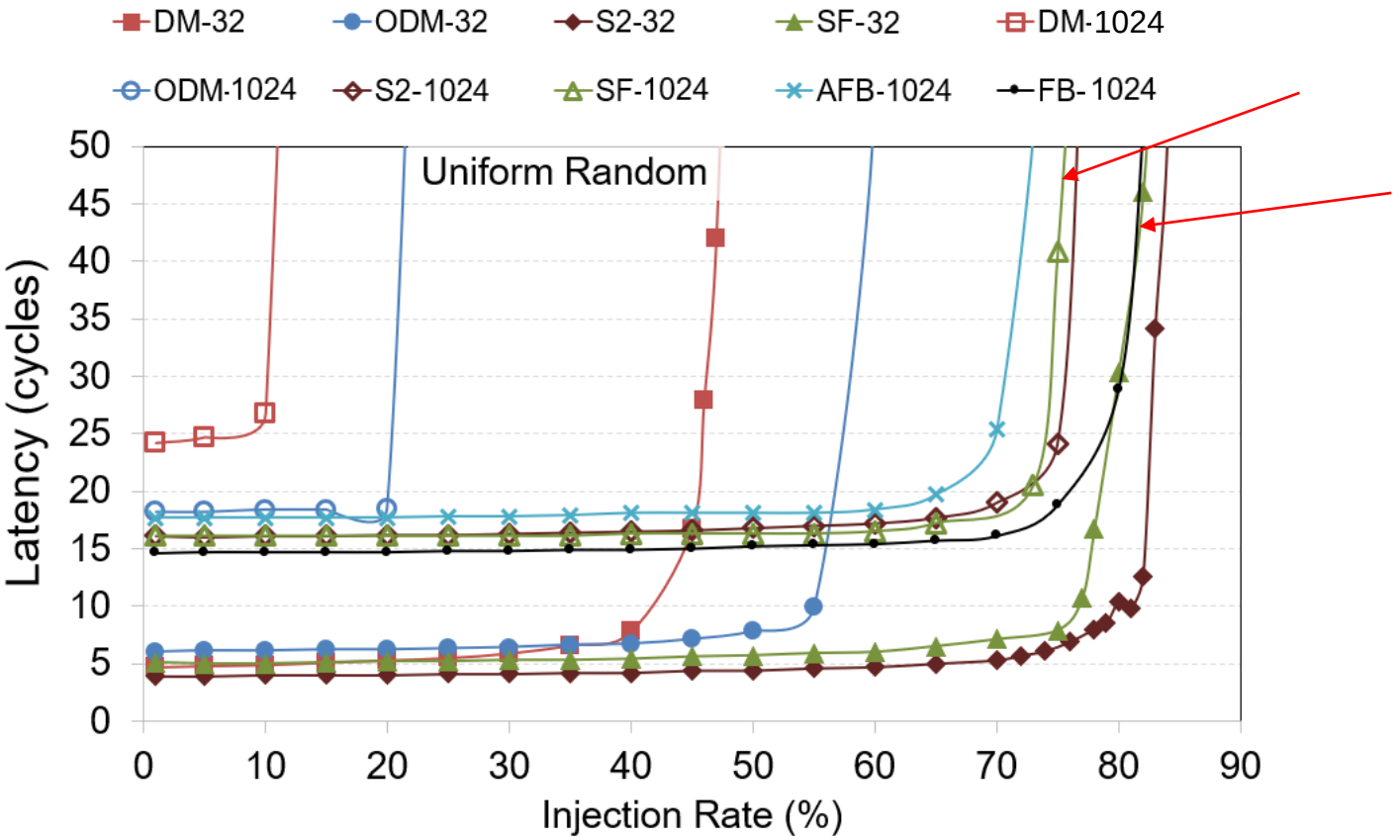
- **Traffic Patterns** – specific routing behaviors in networks
  - Uniform Random
  - Tornado
  - Hotspot
  - Opposite
  - Nearest Neighbor
  - Complement
  - Partition

- **Workloads** – real-world applications for memory networks
  - Spark-wordcount
  - Spark-grep
  - Spark-sort
  - Pagerank
  - Redis
  - Memcached
  - Matrix Multiply
  - Kmeans

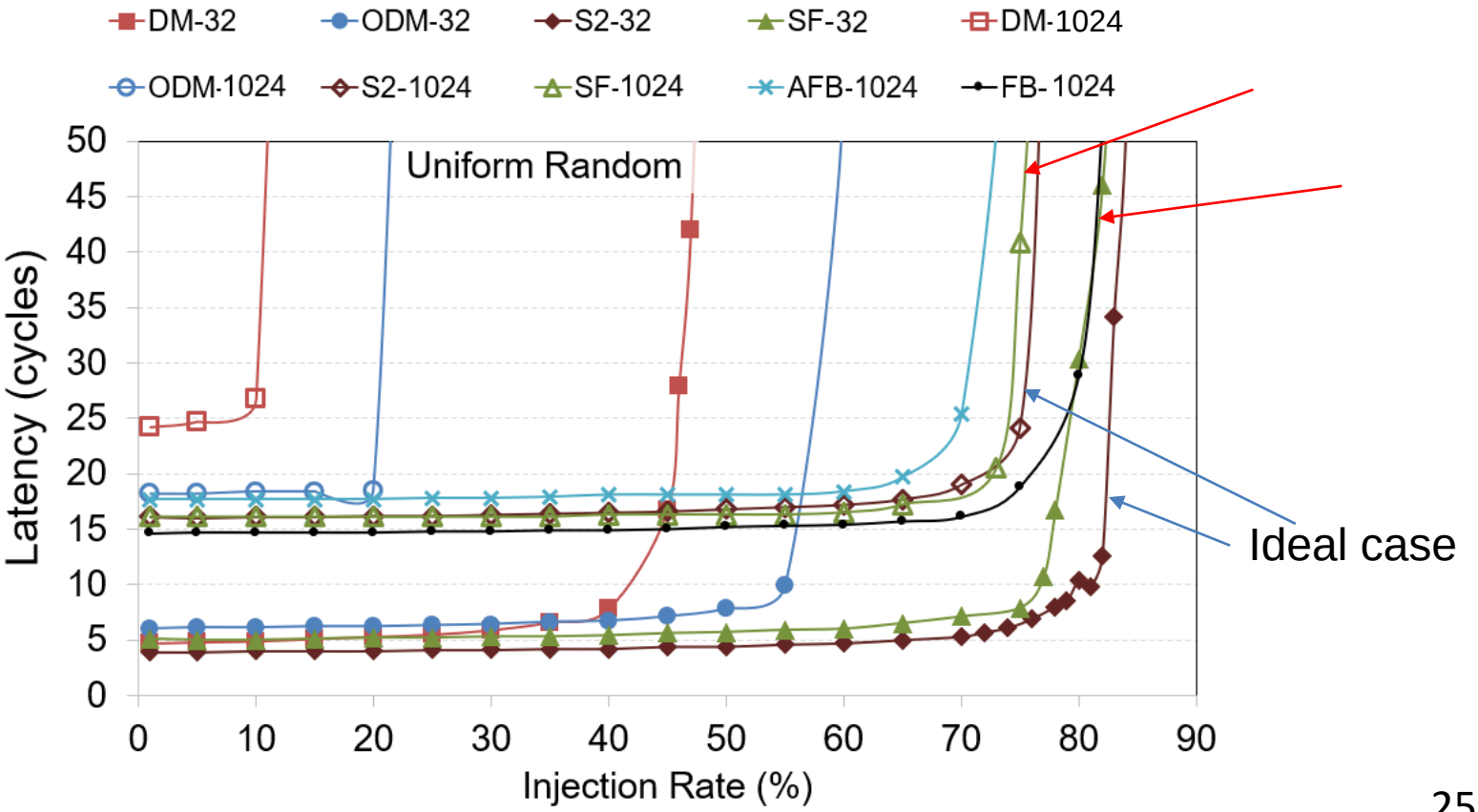
Results


DM-32
ODM-32
S2-32
SF-32
DM-1024
ODM-1024
S2-1024
SF-1024
AFB-1024
FB-1024
Uniform Random
Latency (cycles)
Injection Rate (%)
0
5
10
15
20
25
30
35
40
45
50
0
10
20
30
40
50
60
70
80
90

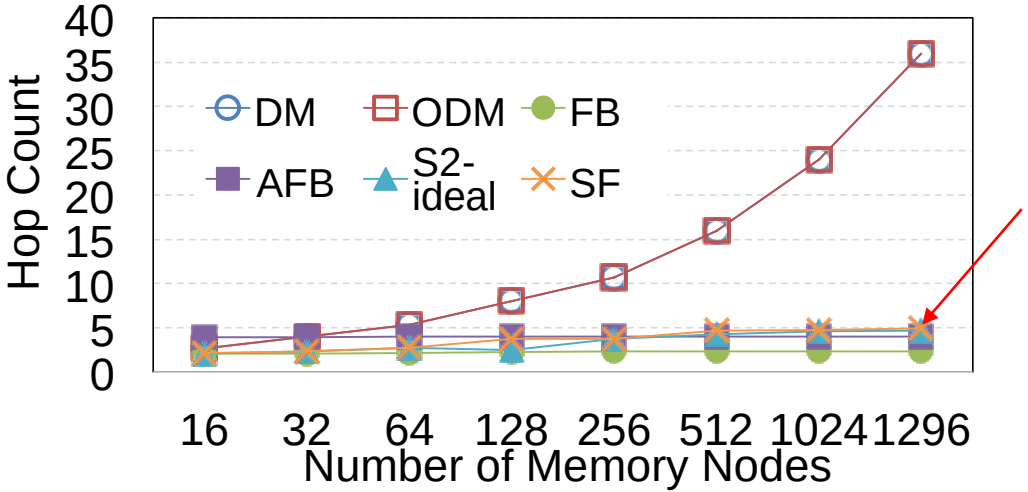
Results


DM-32
ODM-32
S2-32
SF-32
DM-1024
ODM-1024
S2-1024
SF-1024
AFB-1024
FB-1024
Uniform Random
Latency (cycles)
Injection Rate (%)
0
5
10
15
20
25
30
35
40
45
50
0
10
20
30
40
50
60
70
80
90

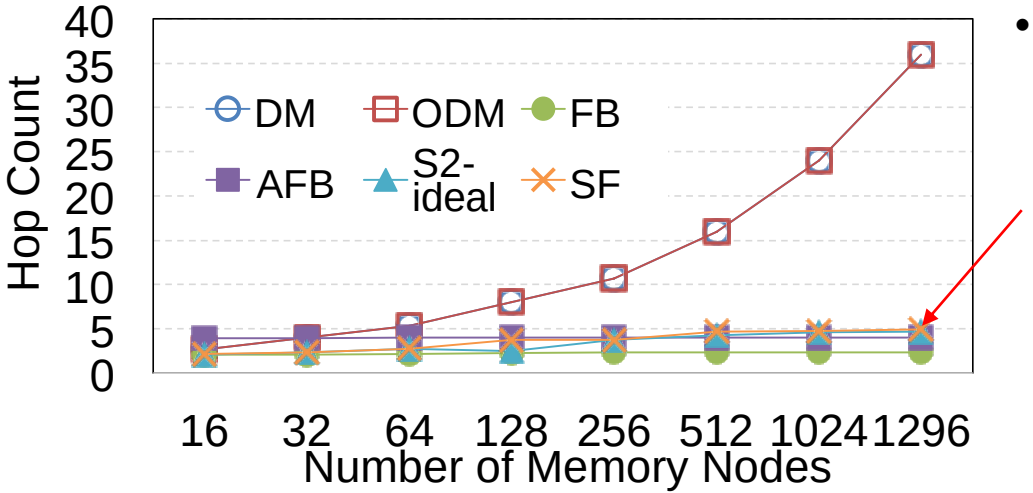
Results


DM-32
ODM-32
S2-32
SF-32
DM-1024
ODM-1024
S2-1024
SF-1024
AFB-1024
FB-1024
Uniform Random
Latency (cycles)
Injection Rate (%)
0
5
10
15
20
25
30
35
40
45
50
0
10
20
30
40
50
60
70
80
90
Ideal case

Results


Hop Count
40
35
30
25
20
15
10
5
0
DM
ODM
FB
AFB
S2-ideal
SF
16
32
64
128
256
512
1024
1296
Number of Memory Nodes

Results

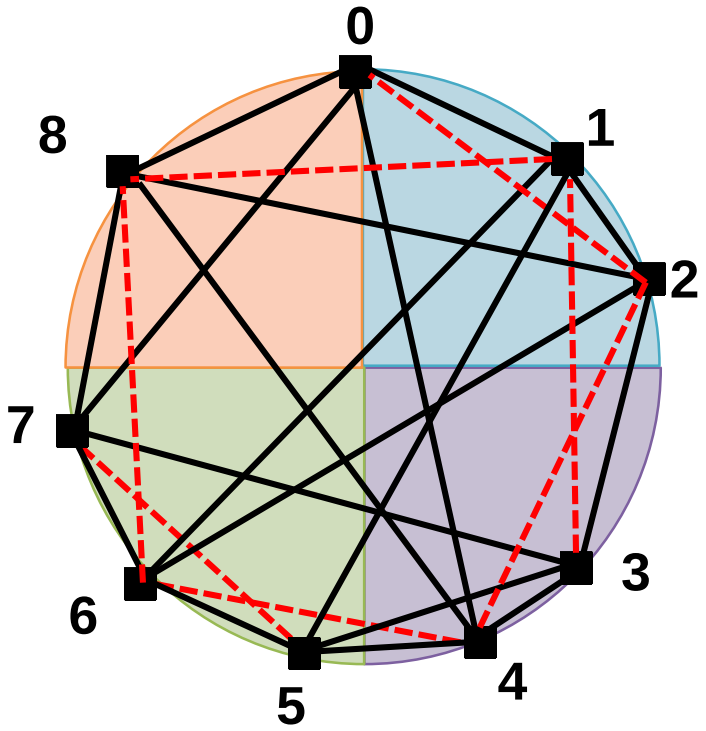


- **More in the paper!**
  - Traffic pattern latencies
  - Real workload performance
  - Network saturation
  - Energy-Delay Product (EDP)
  - Network scaling
  - Deadlock avoidance

Summary of String Figure

Benefits

- Scalability
- Arbitrary Network Scale
- Elastic Network Scale



String Figure: A Scalable and Elastic Memory Network Architecture

Matheus Ogleari

Ye Yu, Chen Qian, Ethan Miller, Jishen Zhao

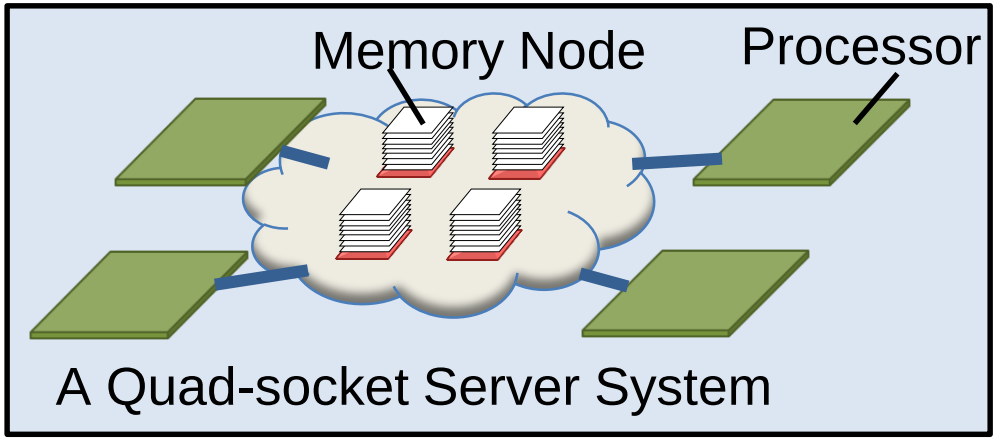
HPCA 2019

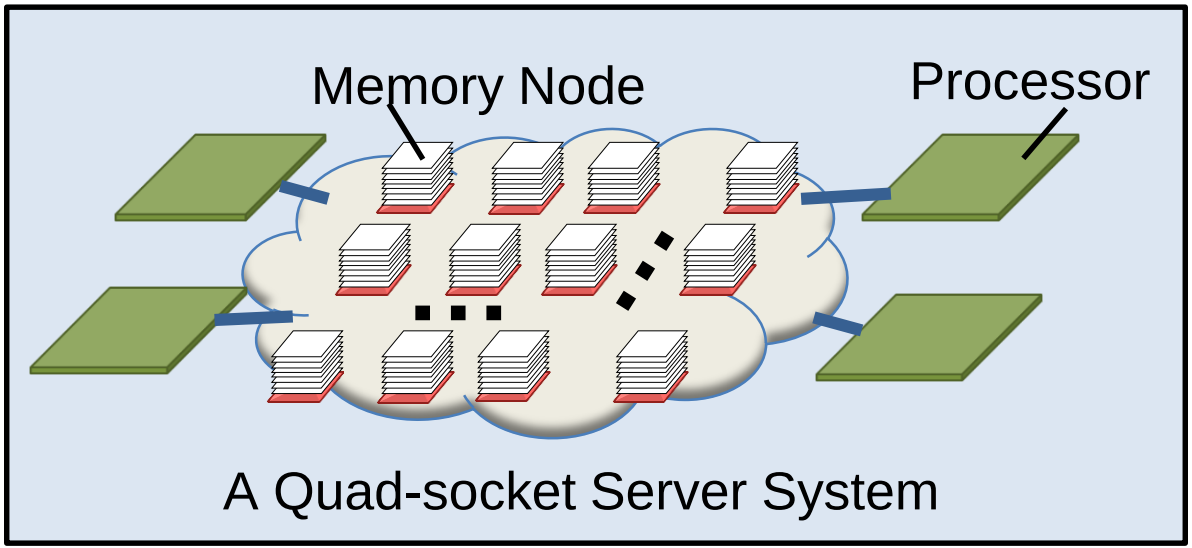
UC San Diego

UNIVERSITY OF CALIFORNIA SANTA CRUZ

PURESTORAGE

ssrc STORAGE SYSTEMS RESEARCH CENTER


Memory Node
Processor
A Quad-socket Server System


Memory Node
Processor
A Quad-socket Server System