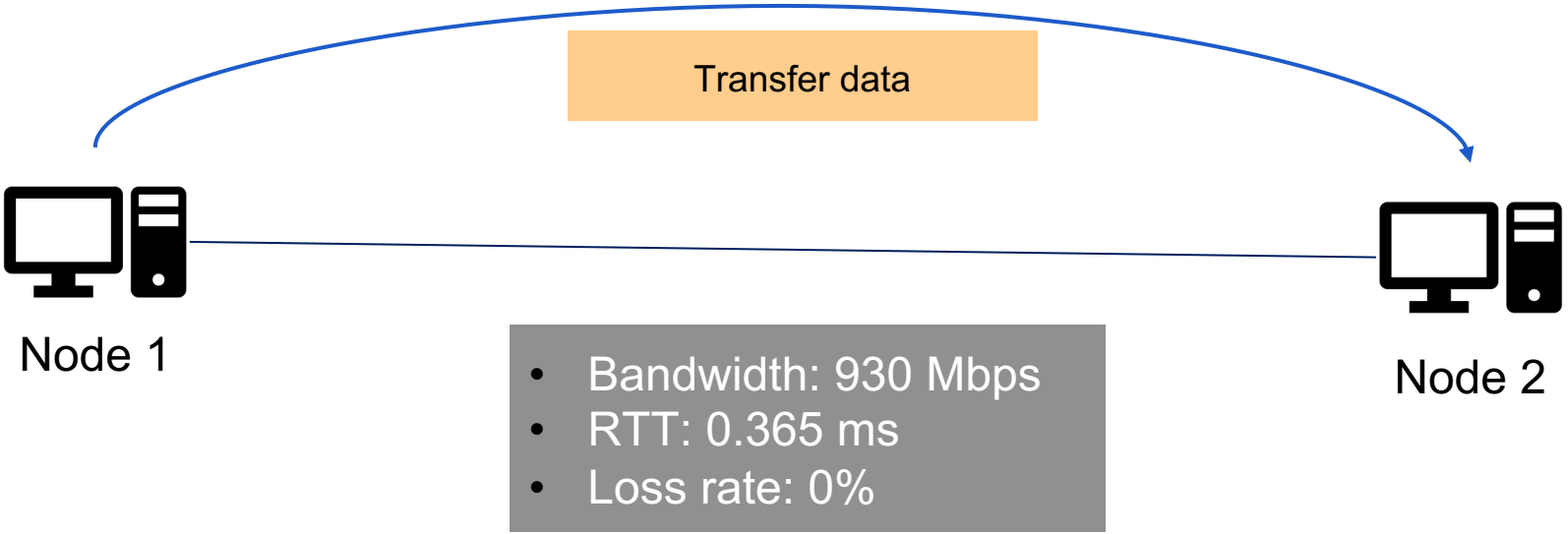
Sample Transfer Optimization with Adaptive Deep Neural Network

Hemanta Sapkota, Md Arifuzzaman, Engin Arslan

University of Nevada - Reno

What is sample transfer?

- Short data transfer to collect network statistics such as available bandwidth, round trip time, loss rate, jitter, etc.

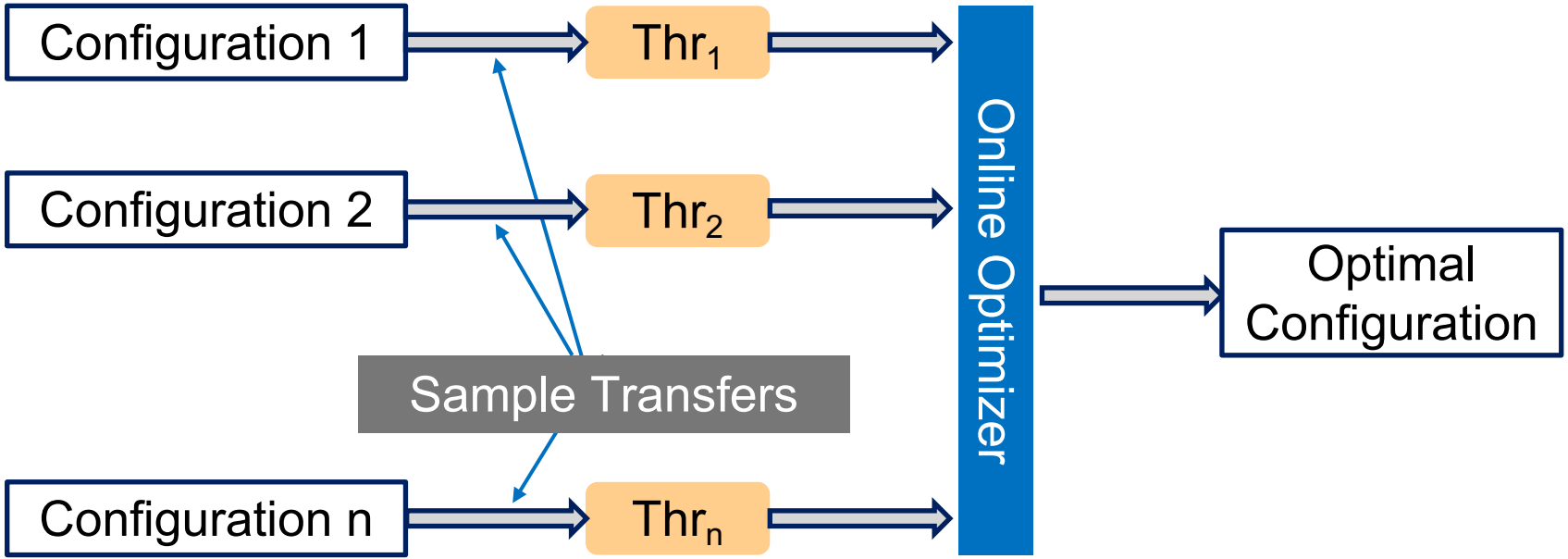


Some examples…

- **Iperf** → Bandwidth estimation
- **PerfSonar** → Bandwidth estimation and anomaly detection
- **Iftop** → Bandwidth estimation

Some examples…

- Transfer Optimization



State-of-the-art

Fixed Duration

Using a fixed duration (e.g., 10 sec) to run sample transfer and measure transferred data size.

Example: Iperf

Disadvantage: Hard to find optimal duration

Fixed Data Size

Using a fixed amount of data (e.g., 10GB) to run sample transfer and measure time

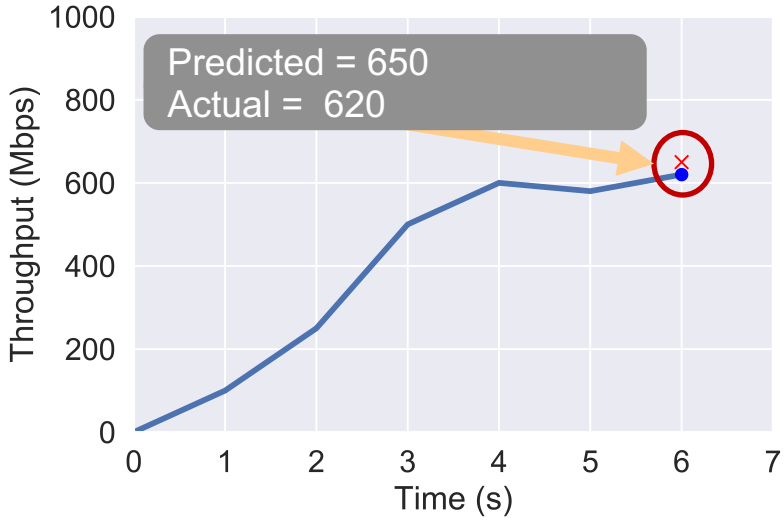
Example: PerfSonar, Yildirim et al.*

Disadvantage: Long transfer time

*_Modeling throughput sampling size for a cloud-hosted data scheduling and optimization service_, E Yildirim, J Kim, T Kosar, Future Generation Computer Systems, 2013

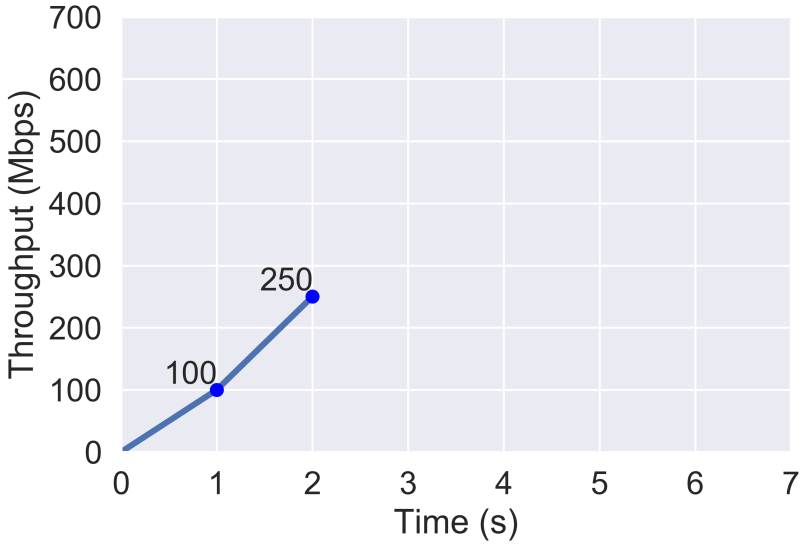
Time Series Analysis

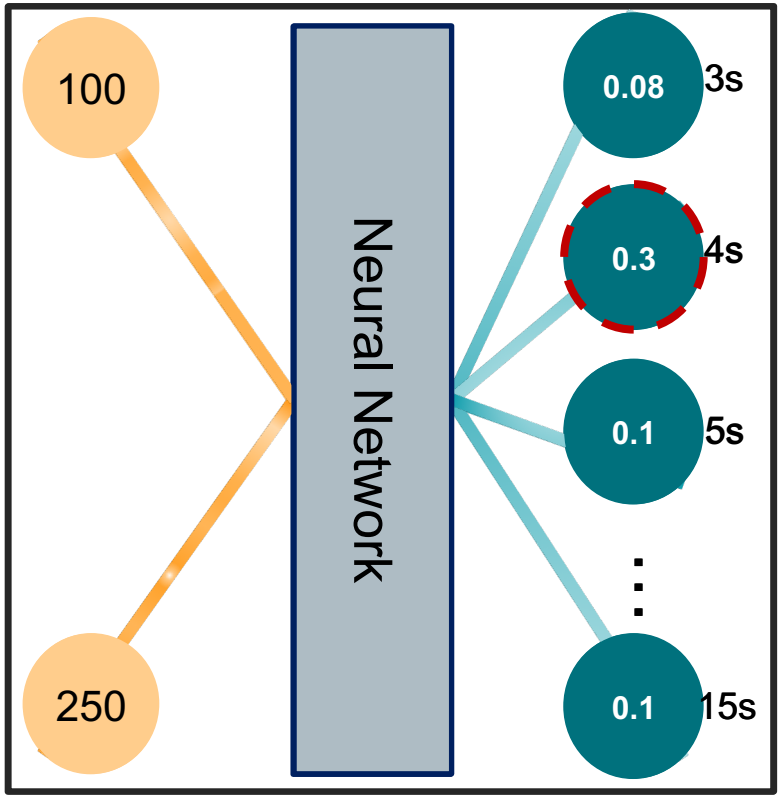
- Use instantaneous throughput values to derive Autoregressive model
- Predict next throughput and compare against actual one to measure its accuracy



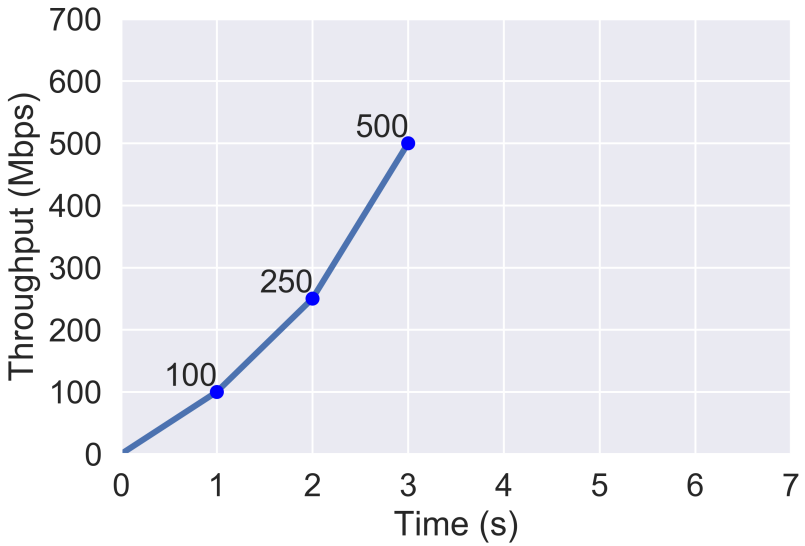
Adaptive Deep Neural Network

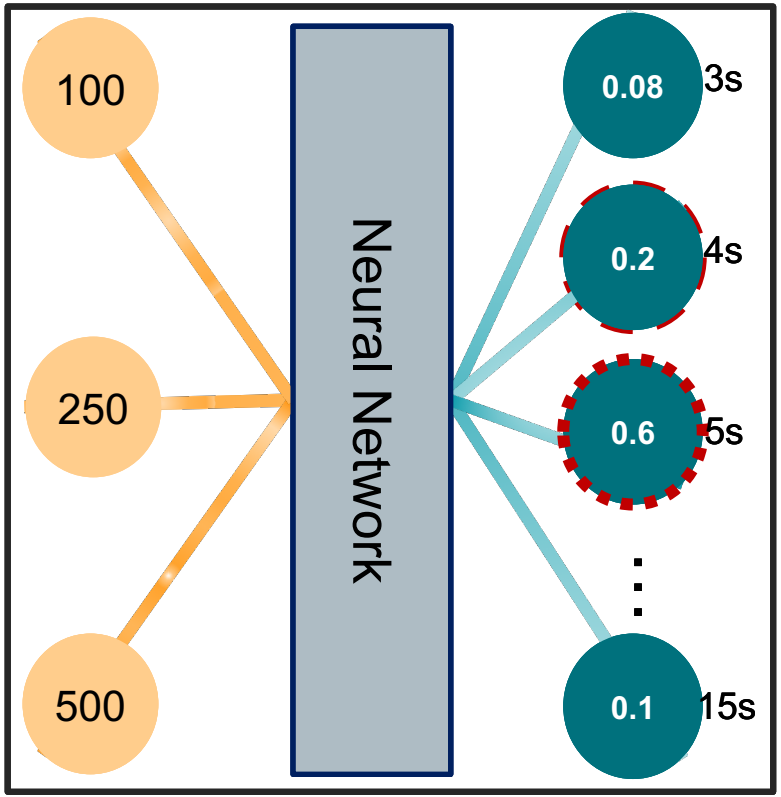
DNN: Convergence time



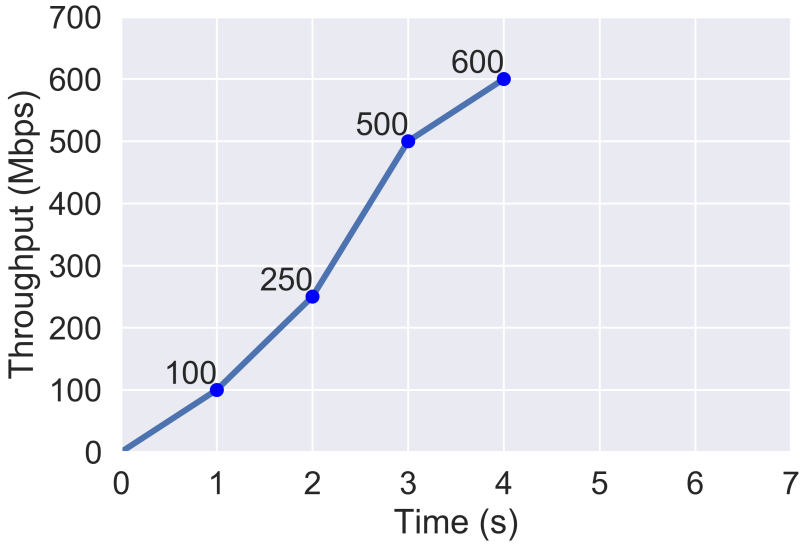


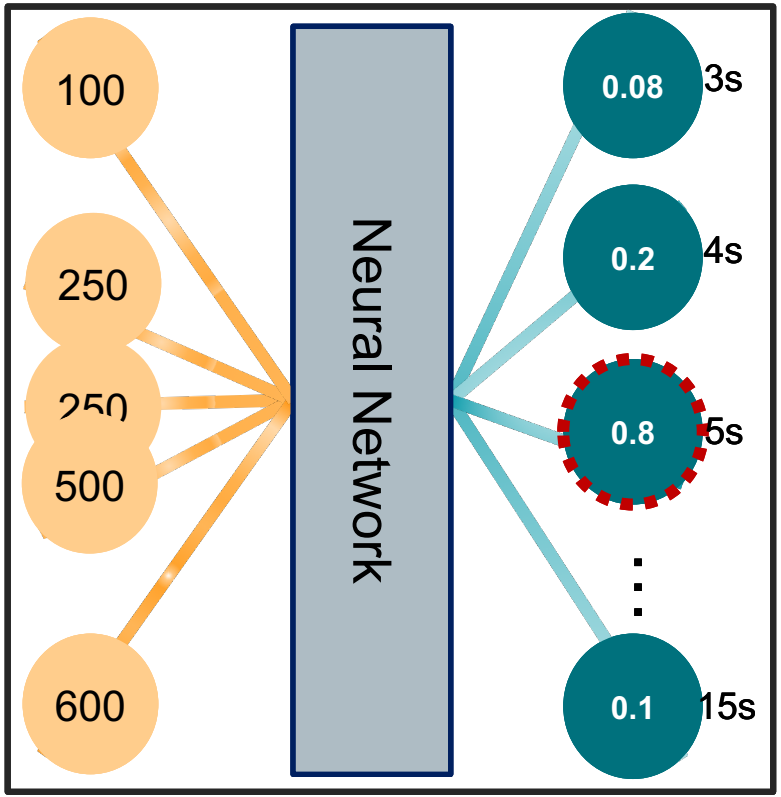
DNN: Convergence time



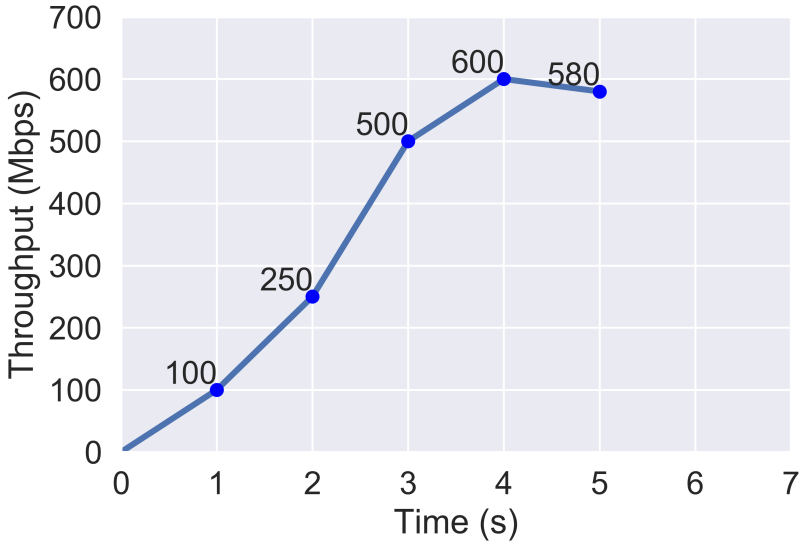


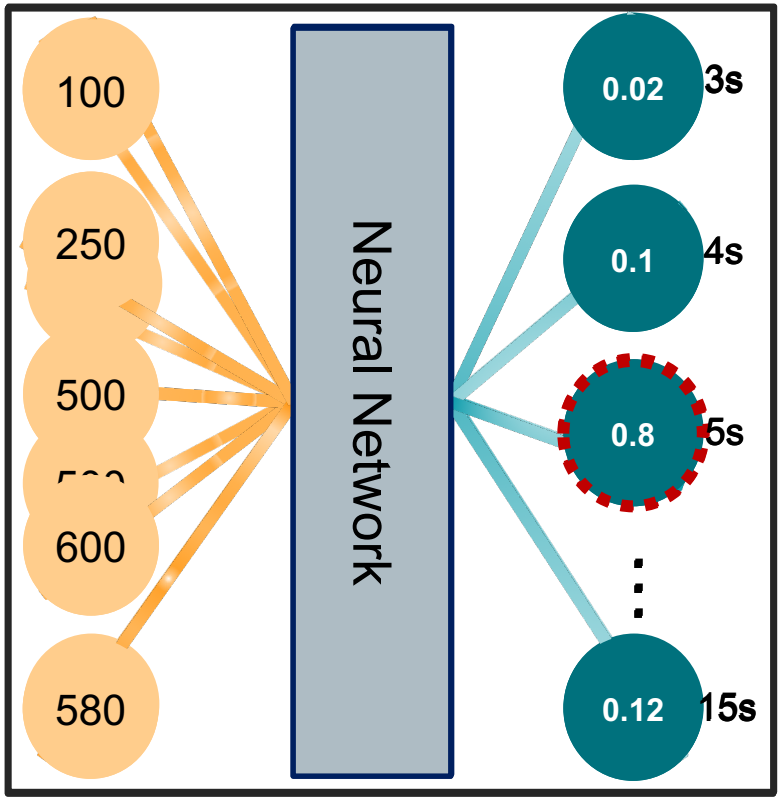
DNN: Convergence time





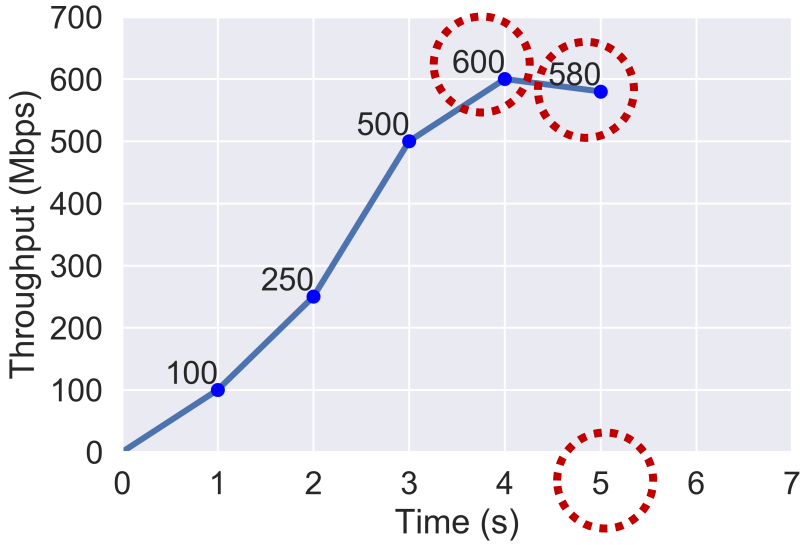
DNN: Convergence time





Throughput Estimation

- Now, since the convergence time (i.e., 5s) is predicted, from neural network, we take **average of previous 2 throughput** from convergence time to predict average throughput.
- Predicted throughput is **590 Mbps**.



$$Predicted\ Throughput = \frac{600 + 580}{2}$$

$$Predicted\ Throughput = \mathbf{590\ Mbps}$$

EXPERIMENTS

System Specs of Experimental Networks

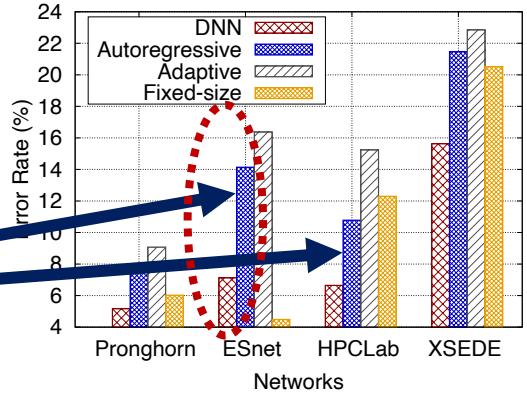
Specs	Storage	Bandwidth (Gbps)	RTT (ms)	Transfer Count
XSEDE	Lustre	10	40	**53,796**
ESnet	RAID-0	100	89	**16,849**
Pronghorn	GPFS	10	0.1	**3,000**
HPCLab	NVMe SSD	40	0.1	**41,768**
Total				**115,413**

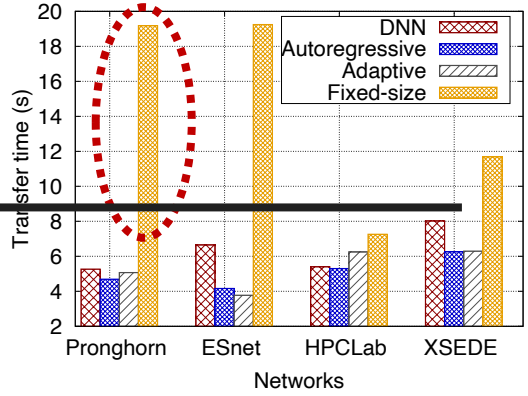
Evaluation Metrics

- **Transfer Time:** The time it takes for a model to predict throughput of sample transfer
- **Error Rate:** Percentage of difference between estimation of a model and actual average throughput

Experimental Results

- Fixed- size approach takes up-to 20 seconds to run
- Autoregressive estimates quickly but causes to high error rate
- Adaptive DNN reduces error rate by up-to 70% compared to Autoregressive
- Adaptive DNN keeps transfer time to less than 8 seconds in all networks

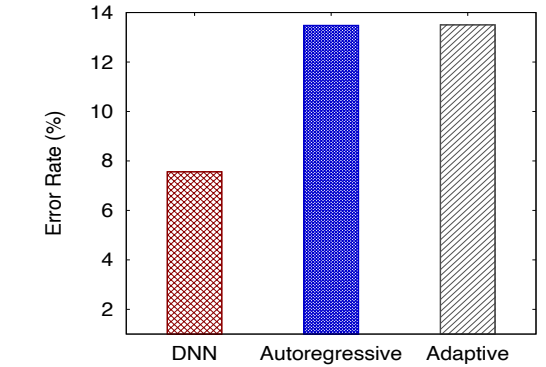


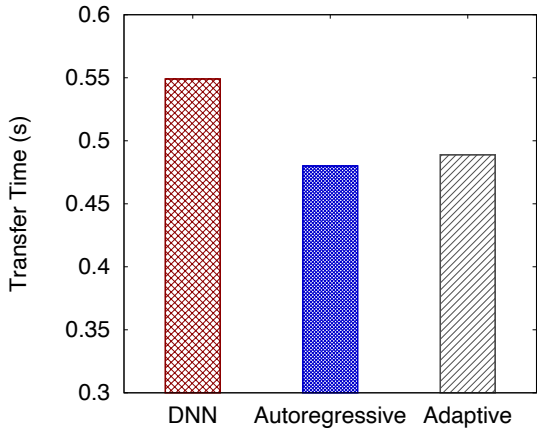


Experimental Results

Increased Data Collection Frequency

- So far, instantaneous throughput is measured once in every second.
- To gather more data points, we measured instantaneous throughput in **every 100 ms**.
- Transfer time **reduced from 5 seconds to 0.55 seconds**.

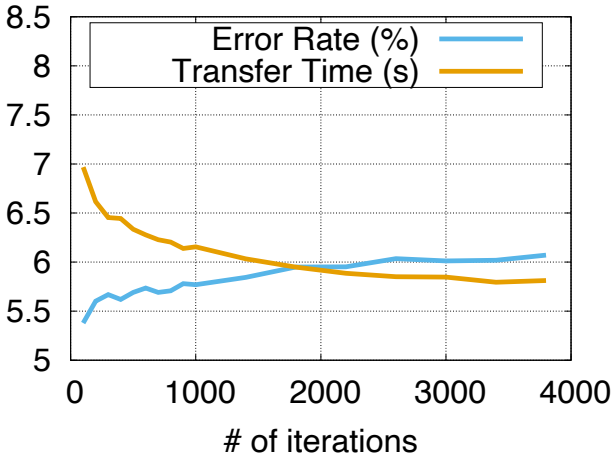




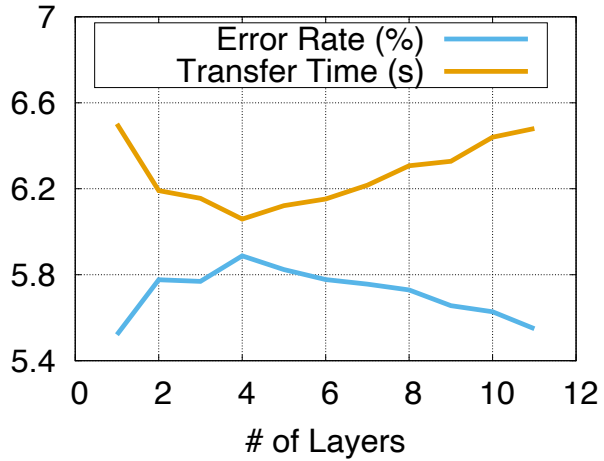
Experimental Results

Impact of Hyperparameters

- Increasing number of iterations **decrease transfer time**, while slightly **increasing error rate**.

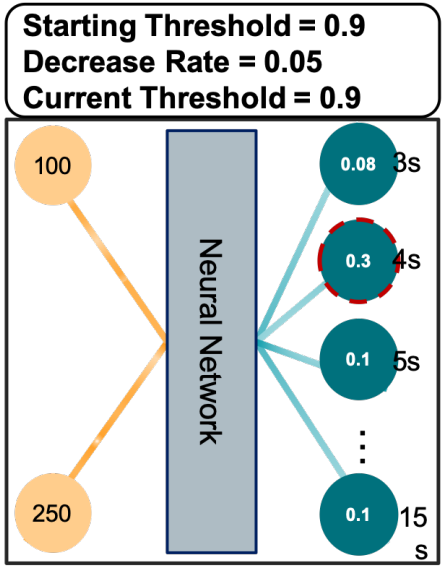


- And increase in number of hidden layers in neural network **decrease error rate**, while **increasing transfer time**.



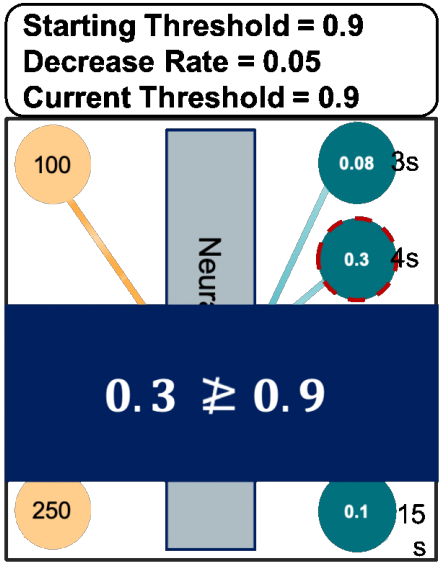
Experimental Results

Impact of probability threshold



Experimental Results

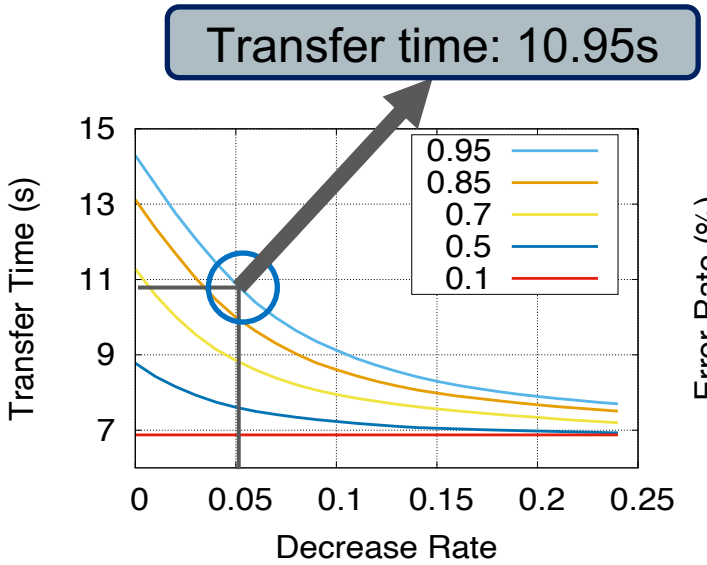
Impact of probability threshold

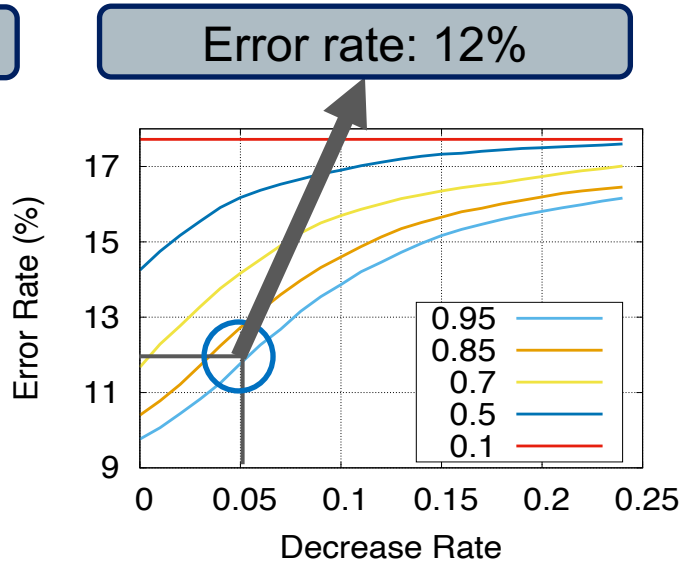


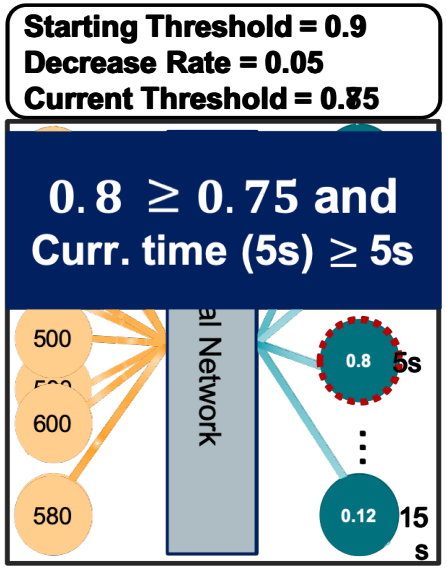
Experimental Results

Impact of probability threshold

- We have **starting threshold** and **decrease rate,** which we can adjust to get the required performance.







Conclusion

- Sample transfers are widely used for various purposes including network measurement and transfer optimization
- Existing approaches causes high error rate (>20%) or long transfer time (~20 seconds)
- Adaptive Deep Neural Network can achieve low error rate by up-to 70% with slight increase in transfer time
- Hyperparameter tuning can help to further reduce error rate or transfer time to meet user/application demand

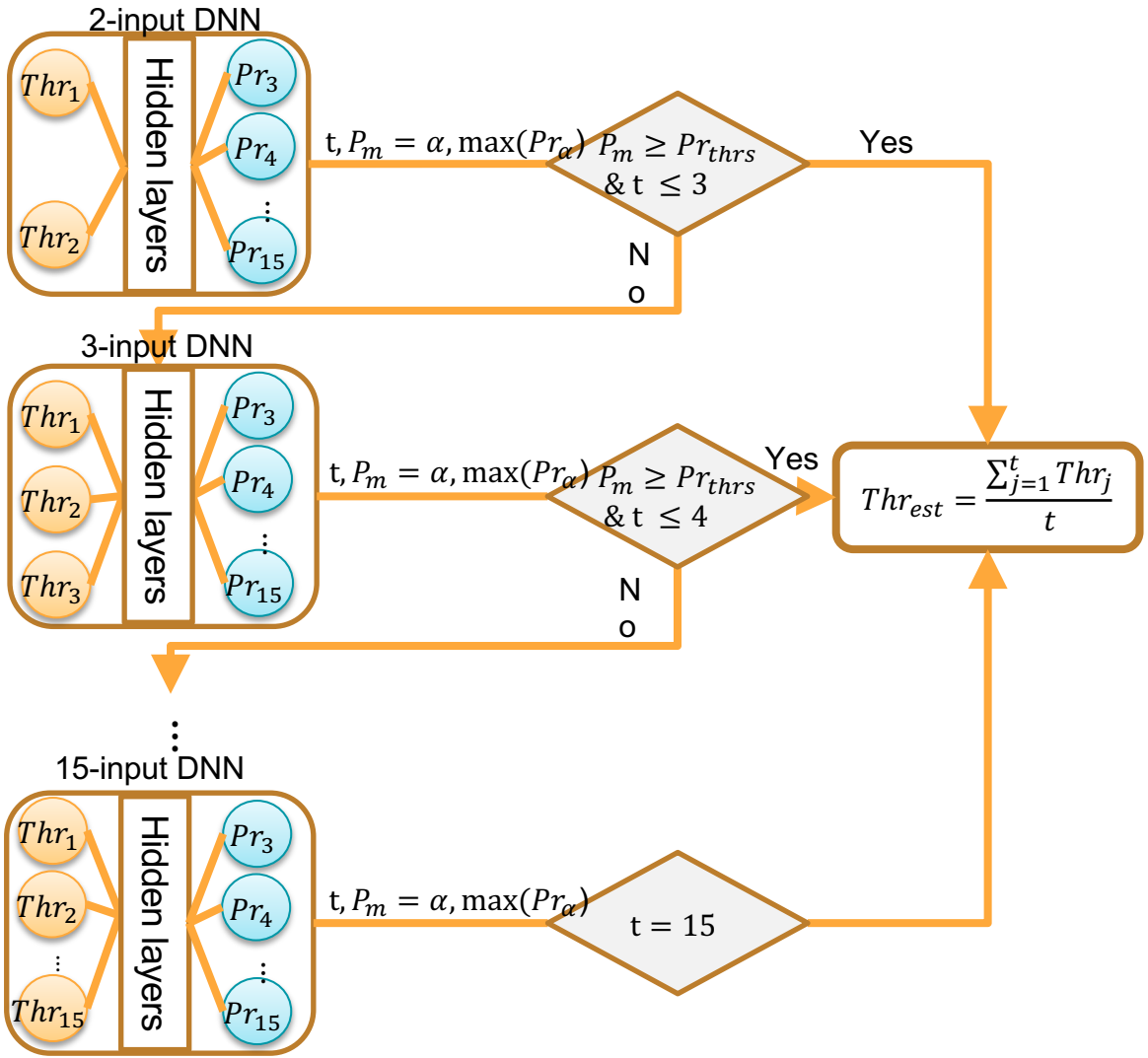
Thank you

Any questions?

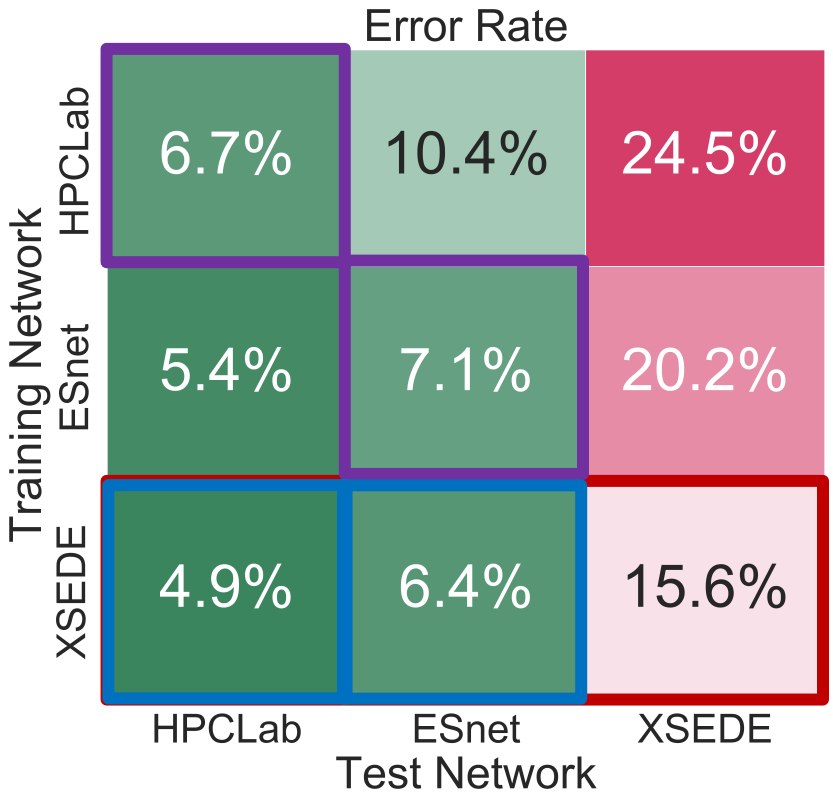
Future works

- Using this model to do parameter tuning and generate optimal configurations with which if transferred data in that particular network, throughput will be optimized.

Deep Neural Network Model



Experiments - Transfer Learning



Experiments - Transfer Learning

- These graph shows the result of training the neural network with different network environment and testing with different environment.
- We can see if we **pick XSEDE as training data** then it performs quite well in terms of error rate with some increase in Convergence time.