Fast and Accurate Sample Transfers

Hemanta Sapkota and Engin Arslan
Data Transfer Optimization Techniques

- **Heuristic Approach**
  + Easy to develop
  - Suboptimal performance

- **Historical Data Based Modeling**
  + Good performance
  - Hard to collect data
  - Sensitive to changes

- **Real-time Tuning**
  + Good performance
  + Does not rely on historical data and assumptions
  + Can adapt to changing conditions
Real-Time Transfer Optimizations

- Discovers the optimal setting by comparing different configurations in the run time
- Sample transfer delays affect convergence speed
Sample Transfers

- **Find throughput for a given transfer setting**
  - What would be throughput if I open four TCP streams instead of one?

- **Accuracy and delay is critical for overall performance**
  - 100 samples * 20 sec/sample-transfer = 33 minutes
Current Approaches for Sample Transfers

- **Fixed time duration**
  - Run every configurations for fixed durations (e.g., 5 seconds)
  - Optimal time duration is different on different networks.

- **Adaptive sampling**
  - Tries to identify convergence by comparing throughput of two consecutive times
  - 100, 350, 750, 1200, 1250
Proposed Solution

- Fit a model to instantaneous transfer throughput

\[ Thr = a + \frac{b}{t^2} \]
Evaluation

- 2700 file transfers between XSEDE sites (Comet, Stampede2, Bridges)
- Time it takes for 80% accuracy for Fixed-duration and the Model is 8 seconds vs 2 seconds
- 90% accuracy in 5 seconds whereas Fixed-duration > 20 seconds
Evaluation

- Evaluation of model accuracy for transfer configurations (i.e., parallel streams)

- Model accuracy increases as more TCP streams are used as a result of more stable throughput

- 94% accuracy with as low as 4 data points (i.e. 4 seconds)
Future Work

- Mathematical proof of the model

- Extensive evaluation
  - Test against more transfer settings such as buffer size, block size, the number of concurrent file transfers etc.

- Integration to real-time optimization algorithms
  - How to stop transfer immediately after convergence detection?
  - Can we test multiple configurations settings simultaneously?
Questions?
Extra Slides
Parallelism=1

![Graph showing accuracy over time for different models: Model, Fixed duration, Adaptive, Last Point. The x-axis represents time (s) and the y-axis represents accuracy (%). The graph shows the performance of each model over time, with the Model showing the highest accuracy.]
Parallelism=8