KEYSIGHT

Iteration-Based Collective Communication Benchmarking for AI Fabrics

Eric Yu, Product Solutions Manager
Ankur Sheth, Senior Director, Strategic Projects

Agenda

01 Recap: GPU-free emulation with KAI DC Builder

02 Motivation: Why Single-Collective Tests Mislead

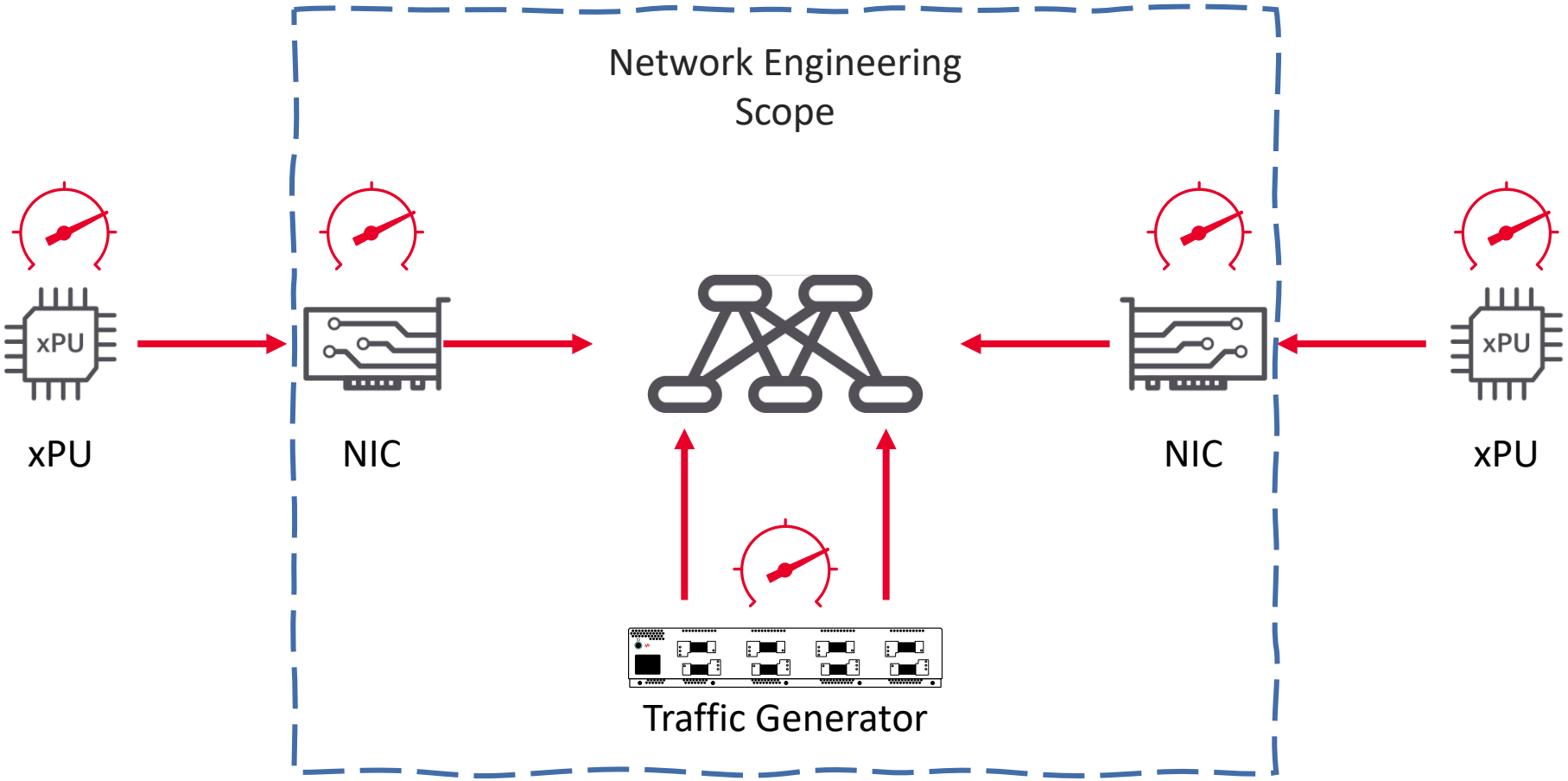
03 Five Layer Validation Framework

04 Iteration Based Methodology: Parameters, Metrics, Results

05 Epoch Level Extension and Real Testbed Findings

06 Limitations and Future Work

Layers of System Benchmarking

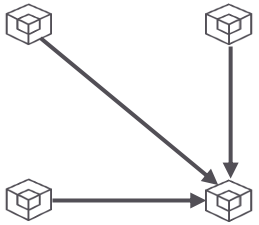


Types and Tools of System Benchmarking

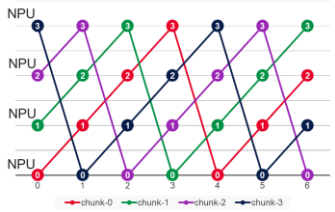
Traditional Network Test Scenarios

P2P

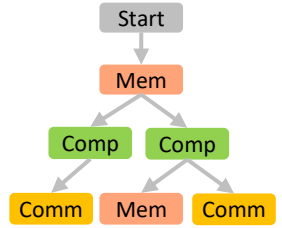
Incast & Fanout



Collective



Model Parallelism

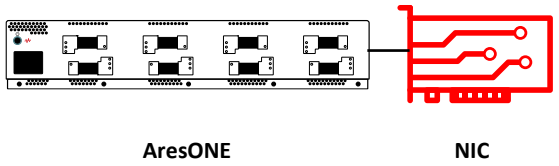


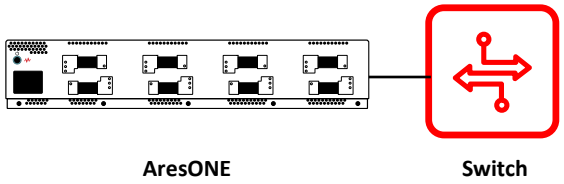
Workload

GPU	gdr_copy ---------> ?	nccl-tests	PARAM	MLPerf
	perftest ---------> ?	OSU Micro-Benchmarks	?	?
	<--------- Traffic Generator --------->	?	?	?

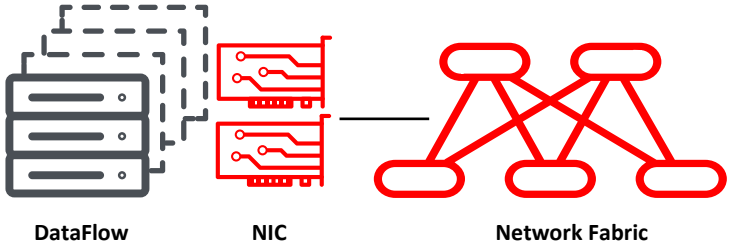
Testbed Topologies

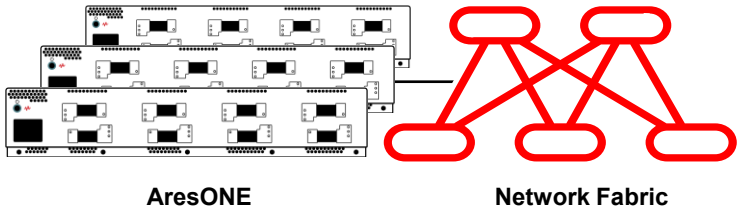
Device level





System level



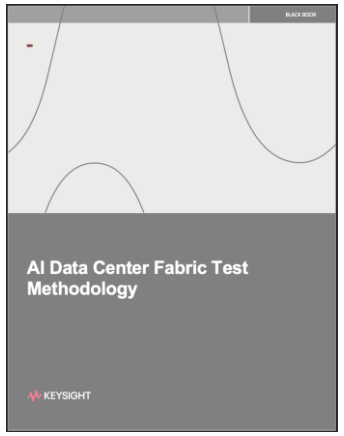


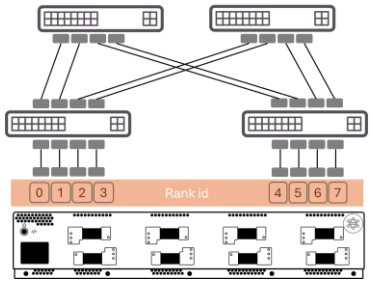
Characterizing Efficiency

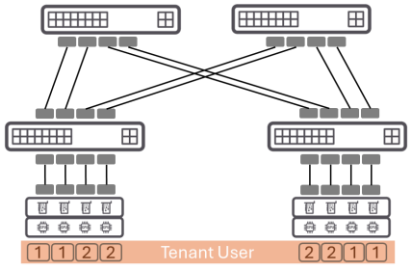
Benchmark	Busbw	Ideal %	Fabric Utilization (%)	P95 FCT Skew (%)
Type of Workload	Algorithm Specific	$\frac{Time_{min}}{Time}$	$\frac{\sum Bytes_{Ingress}}{Ports * Speed * Time}$	$\frac{FCT_{p95} \_ FCTp_{50}}{FCT_{p50}}$
AlltoAll-32 1GB	42.76	86.83	86.91	3.91
AlltoAll-32 1GB	46.67	94.76	94.81	0.66
AlltoAll-8 2GB	47.10	95.64	96.15	2.40

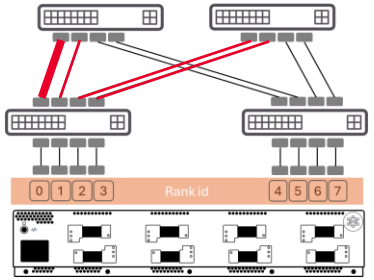
KAI Data Center Fabric Test Methodology

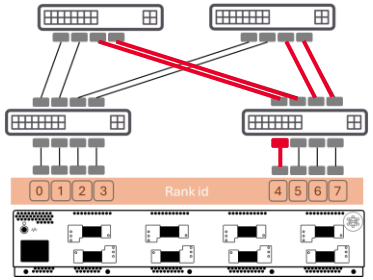
Available for download











Job Completion Time

- Topologies
- Algorithms
- Data sizes
- RDMA message sizes

Performance Isolation

- Noisy neighbors
- Parallel collectives

Load Balancing

- ECMP hashing
- Traffic Engineering
- Q-Pair awareness
- Parallel Q-Pairs
- Dynamic Load Balancing

Congestion Control

- PFC
- ECN/DCQCN
- Fine-tuning

Agenda

01 Recap: GPU-free emulation with KAI DC Builder

02 Motivation: Why Single-Collective Tests Mislead

03 Five Layer Validation Framework

04 Iteration Based Methodology: Parameters, Metrics, Results

05 Epoch Level Extension and Real Testbed Findings

06 Limitations and Future Work

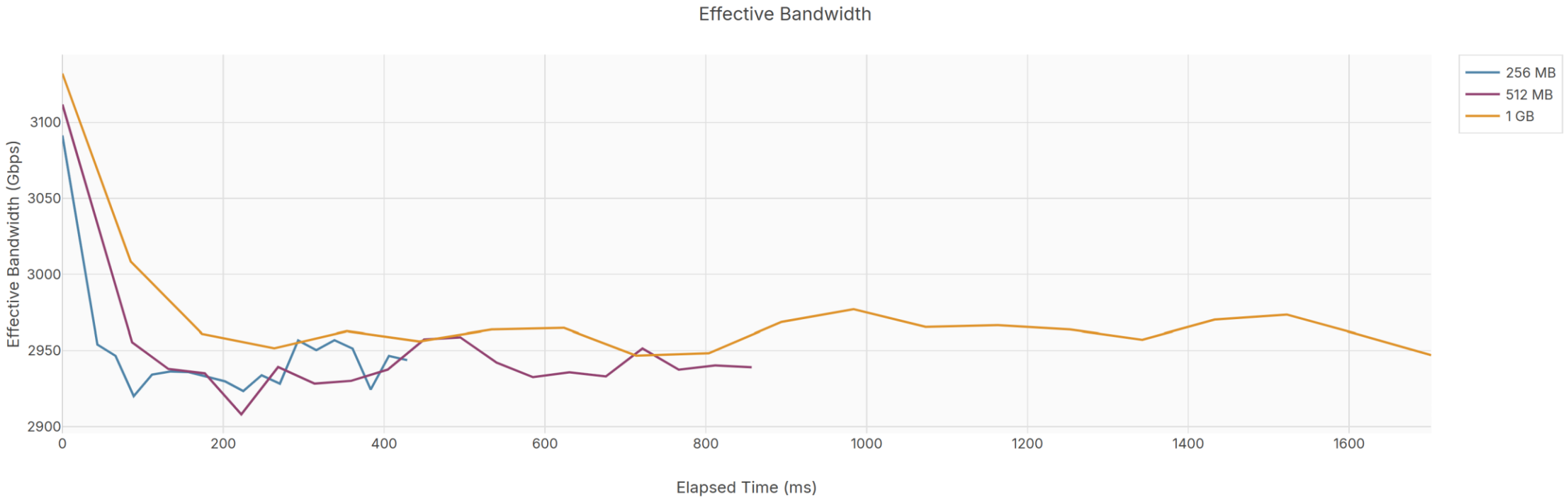
Single-Collective Tests Measure Best Case, Not Production Reality

Four assumptions fail simultaneously in production — not as edge cases

Traditional Assumption	Production Reality
No PFC, No ECN-CE = no congestion	Congestion build up without explicit congestion index
Buffers start empty	Each iteration inherits congestion & queue state from the last
One measurement = steady state	Effective bw on iteration 1 can fall to 70-80% by iteration 5
One collective at a time *	Production workloads interleave multiple collective types/groups - Parallelism

Iteration Observation

32 Ranks All To All Collective GPU-free Emulation



One Collective Iteration Is One Training Step's Communication Phase

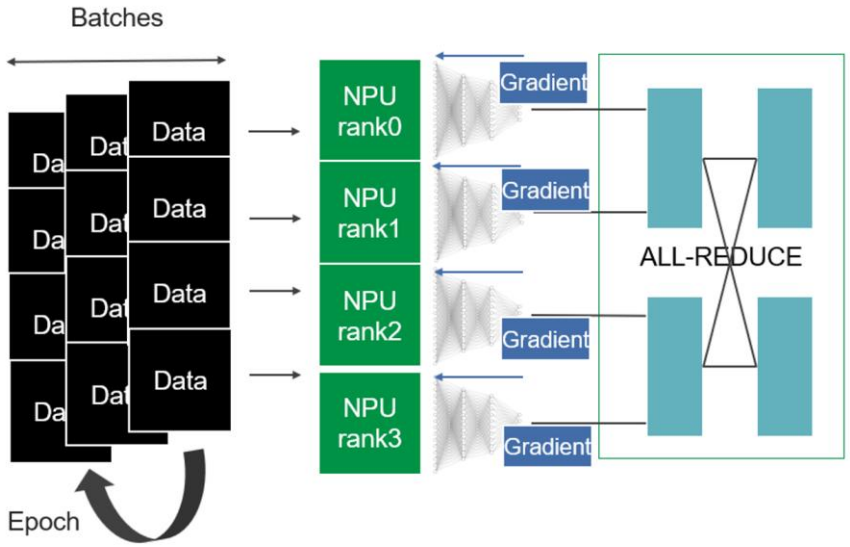
Forward pass → backward pass → AllReduce → weight update

1. Forward + Backward Pass

- Compute gradients = inter_iter_gap

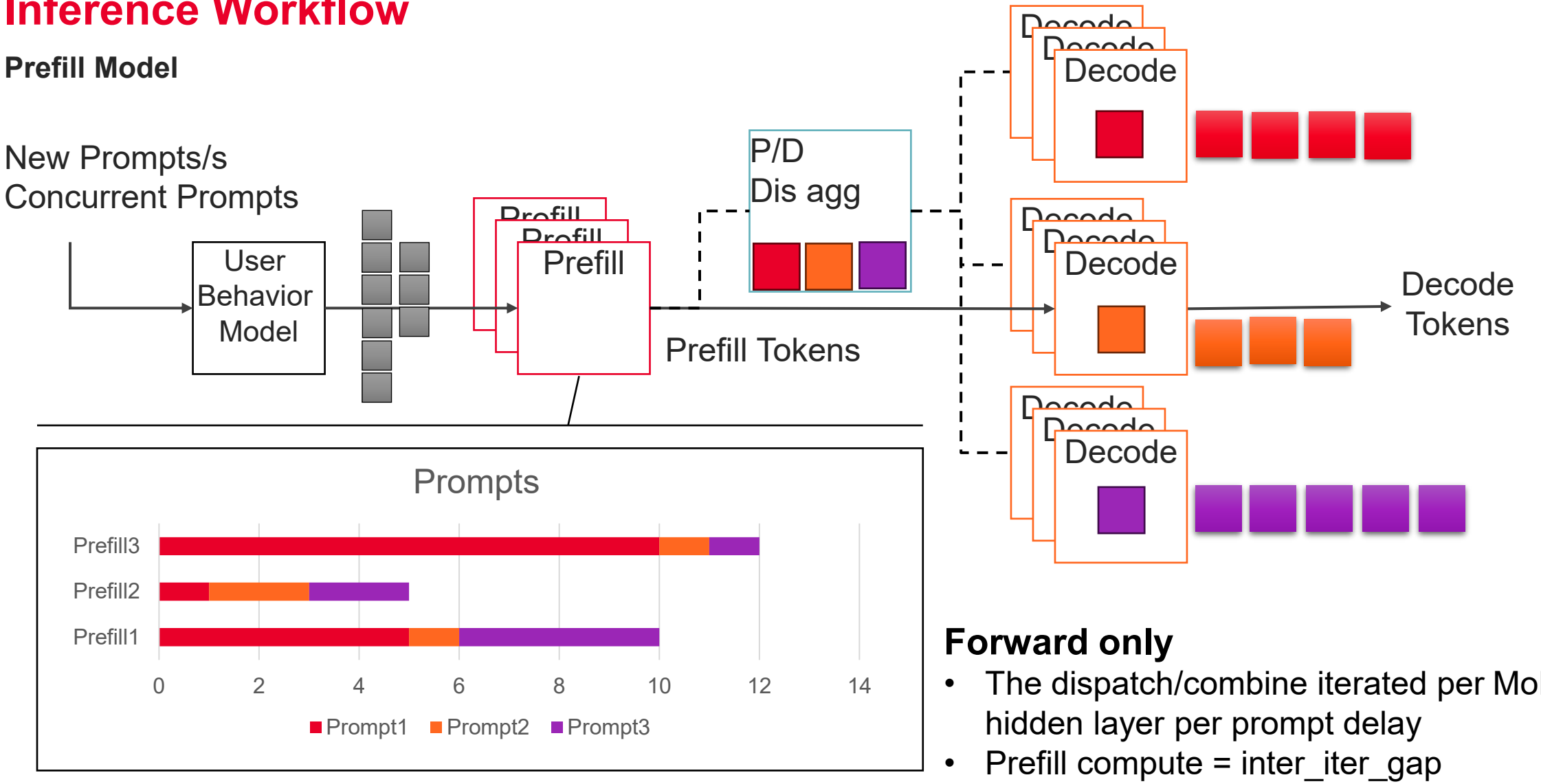
2. All Reduce for Data Parallelism

- The collective iterated per training step
- Next step begins immediately after weight updated
- Epoch start with empty buffer again
- 1 epoch = M iterations
- 1 iteration = dataset_size /global_batch_size



Inference Workflow

Prefill Model

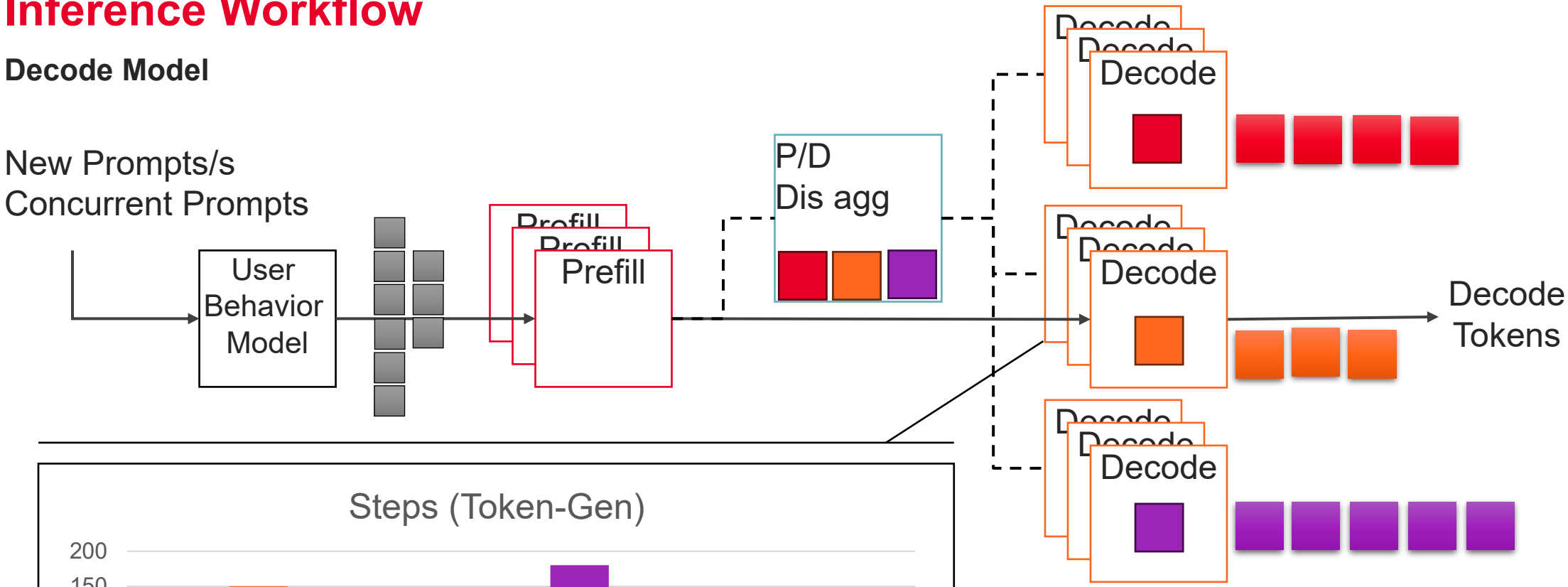


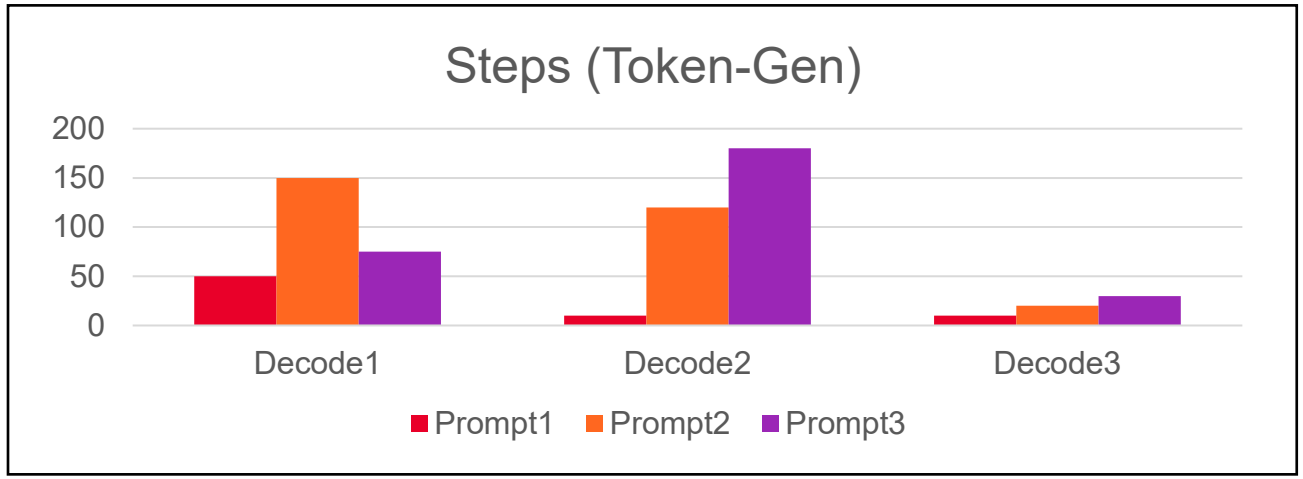
Forward only

- The dispatch/combine iterated per MoE hidden layer per prompt delay
- Prefill compute = inter_iter_gap

Inference Workflow

Decode Model





Forward only

- The dispatch/combine iterated per MoE hidden layer per KV_Cache transfer delay
- Decode compute = inter_iter_gap

Agenda

01 Recap: GPU-free emulation with KAI DC Builder

02 Motivation: Why Single-Collective Tests Mislead

03 Five Layer Validation Framework

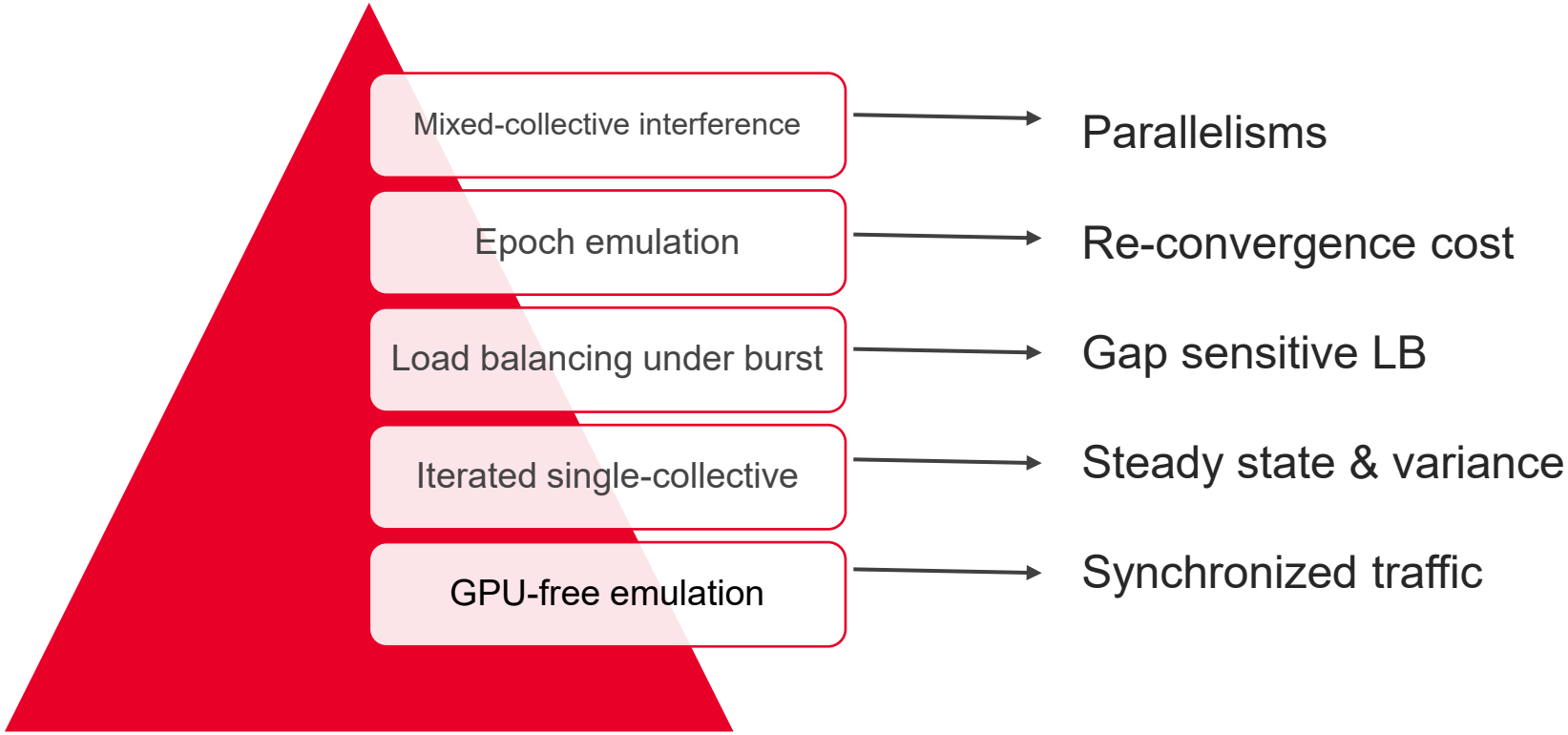
04 Iteration Based Methodology: Parameters, Metrics, Results

05 Epoch Level Extension and Real Testbed Findings

06 Limitations and Future Work

The Bigger Picture — A Five-Layer Validation Framework

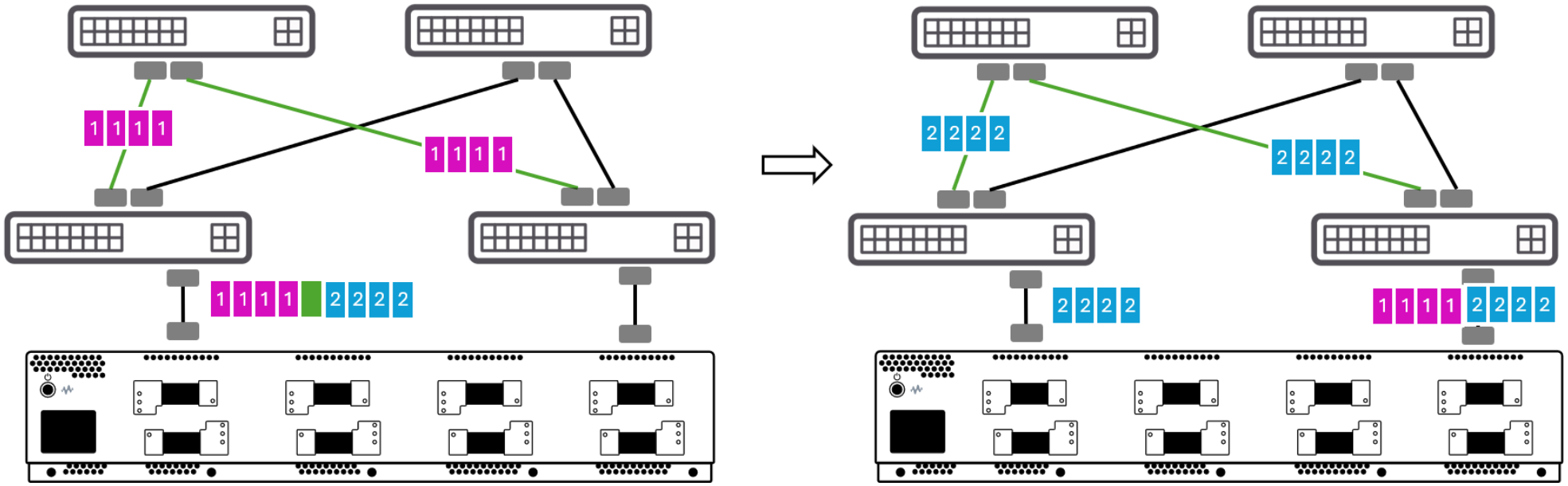
Each layer resolves what the previous layer structurally cannot observe



Load Balancing

Gap sensitive load balancing

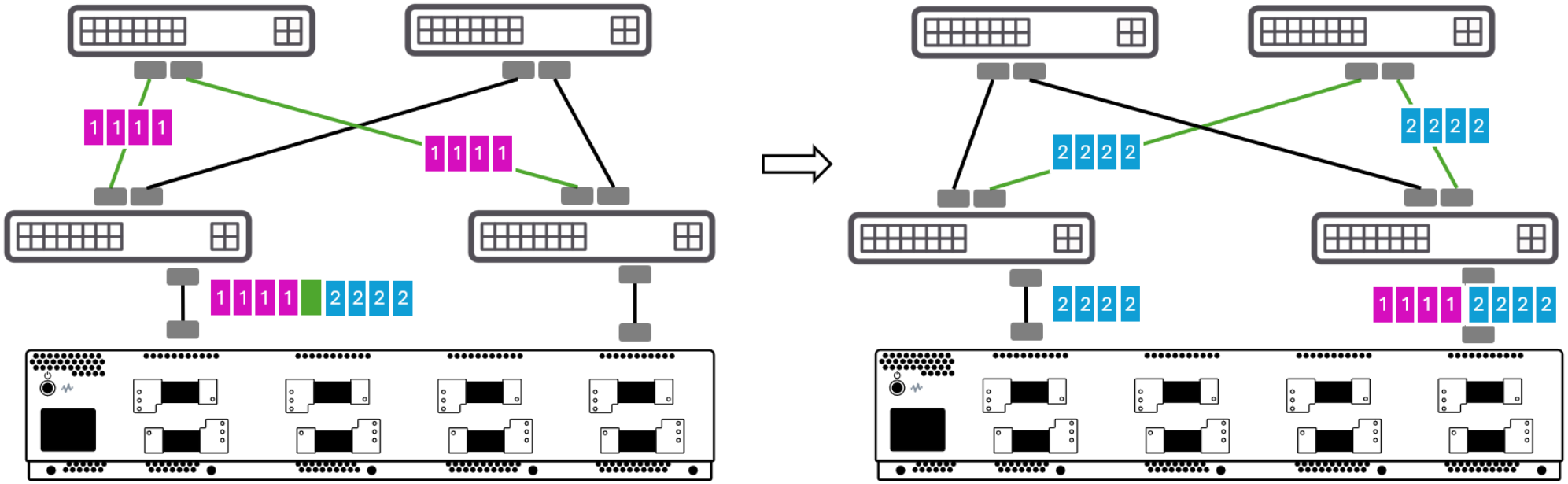
- ECMP
- Flowlet
- Spray



Load Balancing

Gap sensitive load balancing

- ECMP
- Flowlet
- Spray

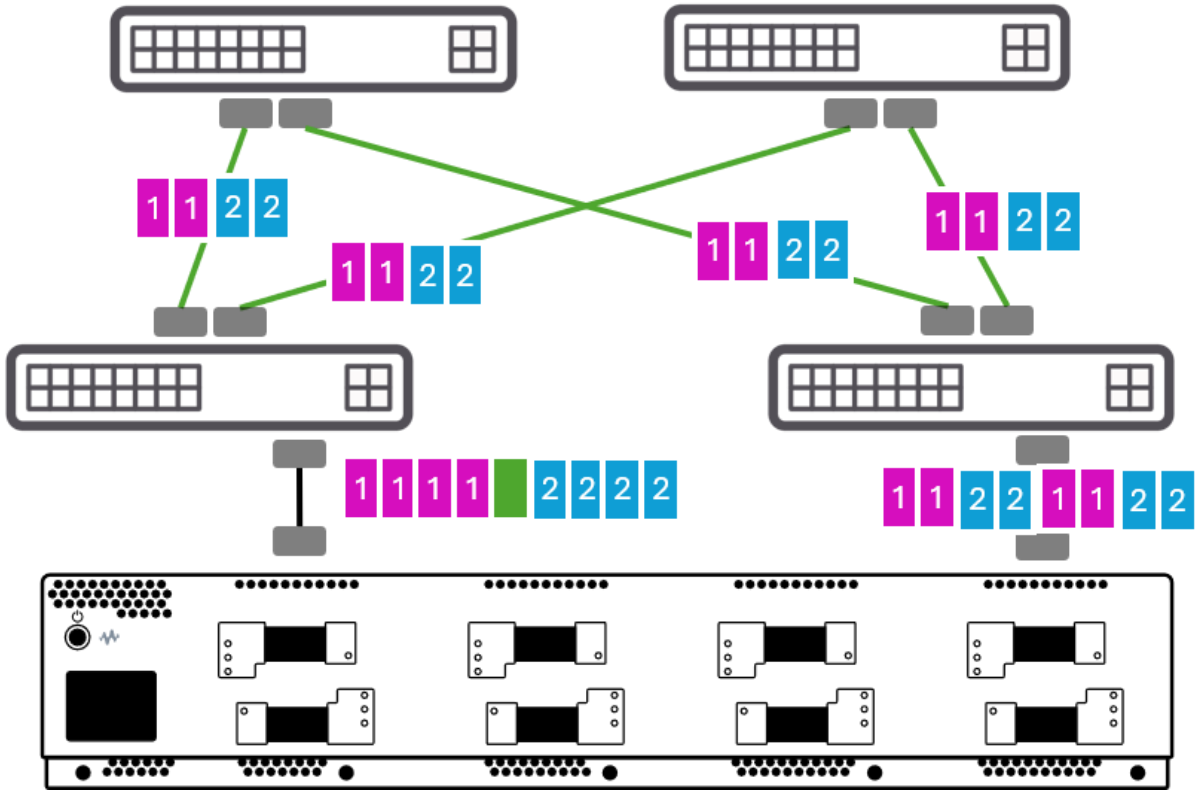


[Flowlet Load Balancing for AI Workloads | Keysight Blogs]

Load Balancing

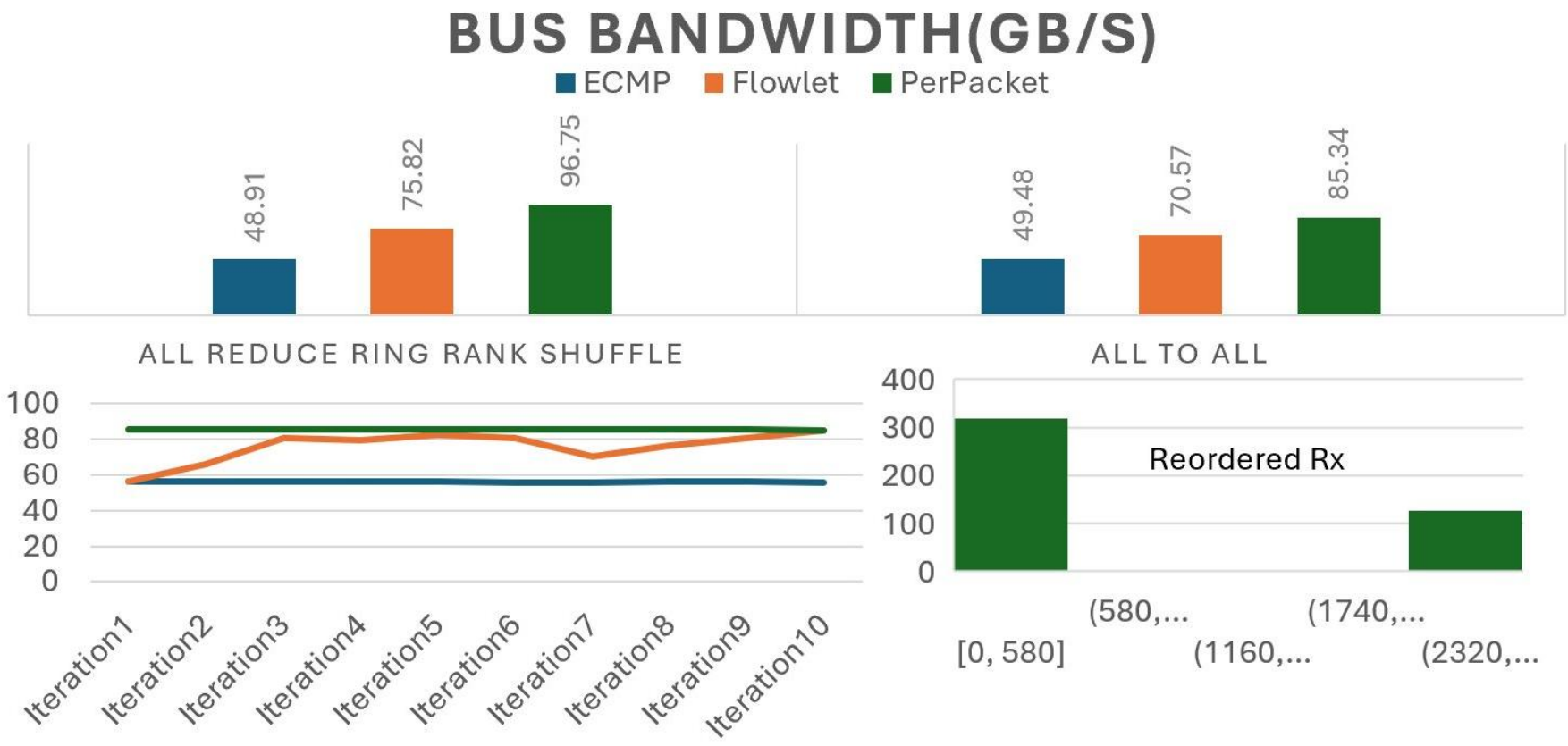
Gap sensitive load balancing

- ECMP
- Flowlet
- Spray



Load Balancing Performance

Iterated benchmark for 800G endpoint



Why Iteration Matters

Congestion, variance, and LB failure only surface over many iterations

- Steady state ≠ transient: iteration 100 inherits iterations 1–99's congestion
- PFC, DCQCN buildup need iterations to manifest
- Variance is signal: poor load balancing causes up to 1.5x oscillation
- Flowlet LB can silently fail when gap = 0 (timeout never fires)

Agenda

01 Recap: GPU-free emulation with KAI DC Builder

02 Motivation: Why Single-Collective Tests Mislead

03 Five Layer Validation Framework

04 Iteration Based Methodology: Parameters, Metrics, Results

05 Epoch Level Extension and Real Testbed Findings

06 Limitations and Future Work

Metrics — Five Numbers That Turn a Time Series Into a Verdict

Replacing a single averaged number with a full stability profile

Metric	Definition	Calculation
CI (Convergence Iteration)	First iteration within ±5% of trailing mean	Per-iteration analysis
SSbw (Steady-State busbw)	Mean busbw after CI	Mean of iterations [CI, N]
P99 busbw	Tail behavior across post-CI iterations	Distribution of [CI, N]
CV (Coefficient of Variation)	std / mean — stability signal	CV < 10% = stable
Degradation rate	Linear-regression slope of busbw vs. iteration	~0 = no systematic drift

Example Result — A Real Convergence Study

16 ranks, 1 GB AllReduce, N = 100, LB = Flowlet (τ = 256 µs)

Metric	Value
Ranks	16
Message size	1 GB AllReduce
Iterations (N)	100
Inter-iteration gap	10 ms
CI (Convergence Iteration)	8
SSbw (Steady-State busbw)	82.4 GB/s (93.5% of 88.1 GB/s baseline)
P99 busbw	79.1 GB/s
CV	4.2%
Degradation rate	−0.012 GB/s / iteration
PFC events	0
Verdict	**PASS**

Agenda

01 Recap: GPU-free emulation with KAI DC Builder

02 Motivation: Why Single-Collective Tests Mislead

03 Five Layer Validation Framework

04 Iteration Based Methodology: Parameters, Metrics, Results

05 Epoch Level Extension and Real Testbed Findings

06 Limitations and Future Work

Extending to Epoch Emulation

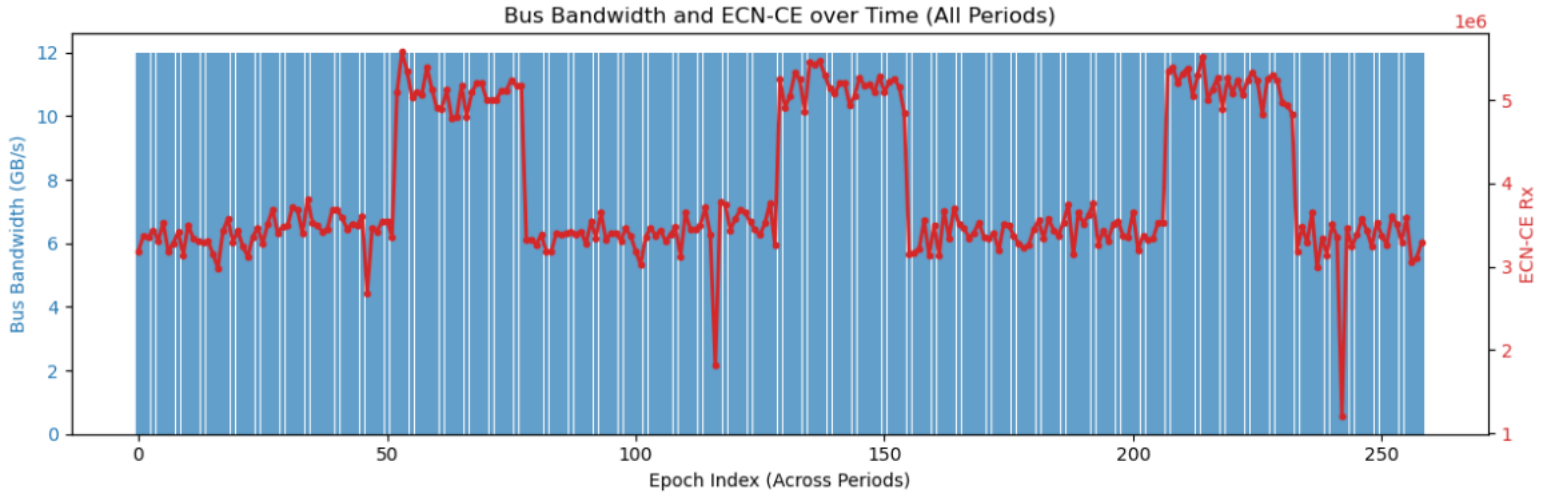
K=3 epochs × N=500 iterations

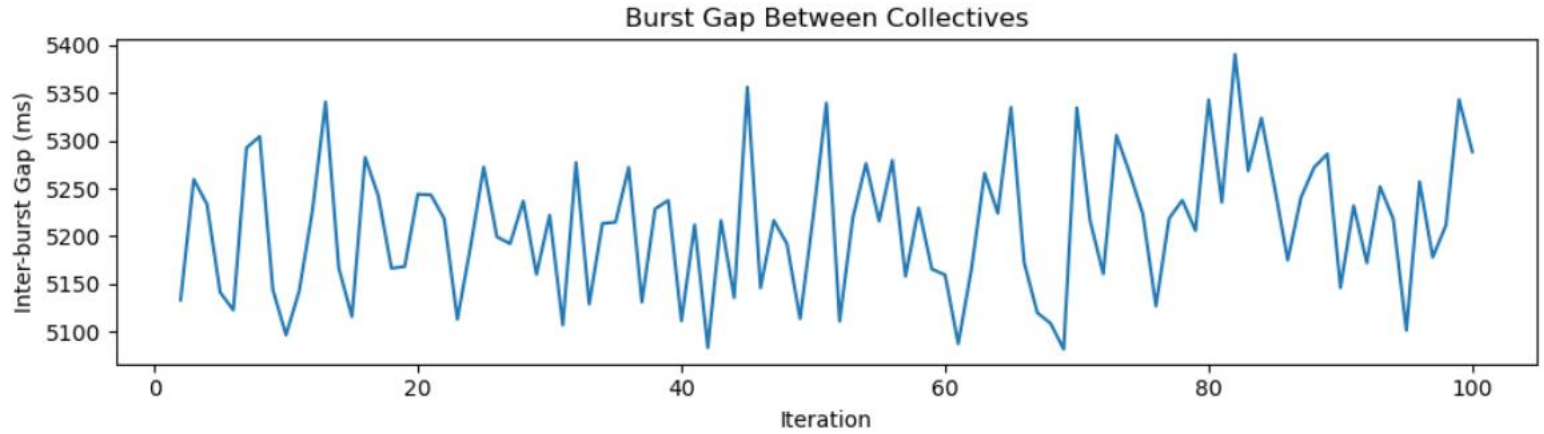
Epoch	CI	SSbw (GB/s)	Note
Epoch 1	7	~84	Initial convergence
Epoch 2	9	~84	Re-converged after boundary gap
Epoch 3	8	~84	Stable, σ = 0.3 across epochs

Real Testbed Observations

A stable mean busbw can still hide long-tail, per-collective oscillation

- 100-iteration run on a real testbed shows a stable busbw plateau
- Periodic ECN-CE spikes indicate transient, recurring congestion
- Start-to-start gap analysis reveals oscillation within one epoch





Agenda

01 Recap: GPU-free emulation with KAI DC Builder

02 Motivation: Why Single-Collective Tests Mislead

03 Five Layer Validation Framework

04 Iteration Based Methodology: Parameters, Metrics, Results

05 Epoch Level Extension and Real Testbed Findings

06 **Limitations and Future Work**

Phase 1 Isolates Iteration Semantics on Purpose — Mixed Collectives Are Next

Scope, limitations, and the future work

- **Phase 1 Scope & Limitations**
- Collective with iterations
- Collective with iteration gap
- Collective iterations with epoch
- One collective at a time
- **Phase 2 Preview**
- Mixed collectives with sync
- Tenants Isolation

Takeaways

Steady-State Visibility

- Iteration-based testing exposes steady-state, variance, and convergence invisible to single-run tests

Epoch-Aware Readiness

- Epoch-aware evaluation, not peak busbw, should be the first-class AI fabric readiness signal

Meet Keysight Team

Dan Mihailescu, Winston Liu

Predictive Performance Modeling and Synthetic Workload Generation for Parallel DNN and LLM Training

MUG ’26
Tue, August 19, 4:00 PM

Visit us in booth C83 for in-person demos

2026 OCP Global Summit

San Jose, California
October 12-15, 2026

KEYSIGHT

Thank you