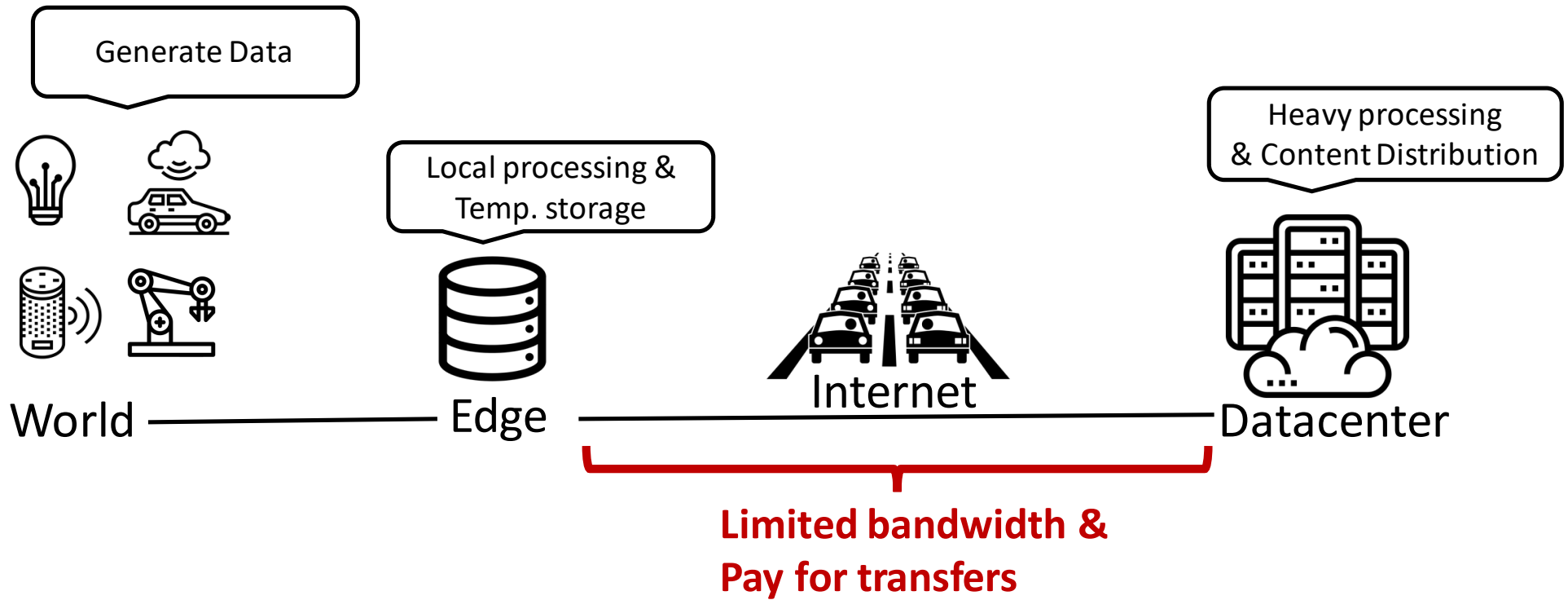


Edge Replication Strategies for Wide-Area Distributed Processing

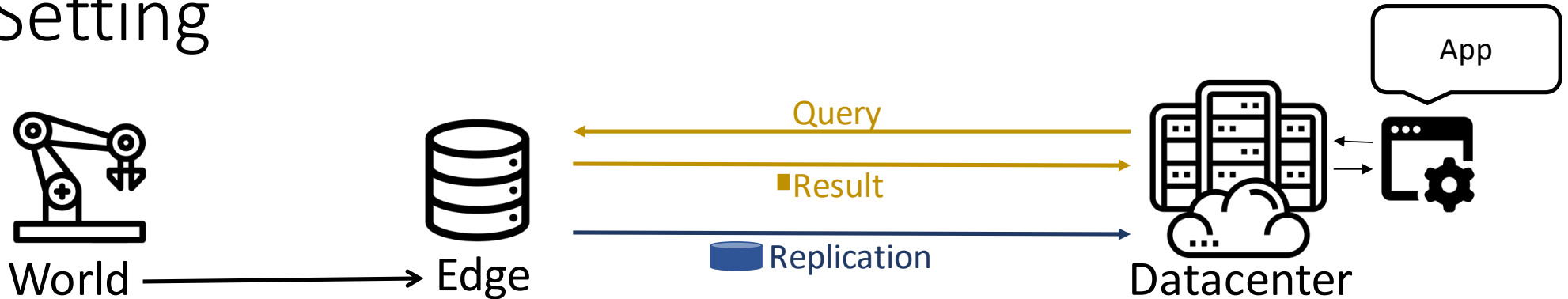
Niklas Semmler, Matthias Rost, Georgios Smaragdakis, Anja Feldmann





How do we reduce the transferred data volume?

Setting



Option A:

Transfer query results.

Cost

Per-query-result
(cumulative)

Good for ...

Few small non-
overlapping results

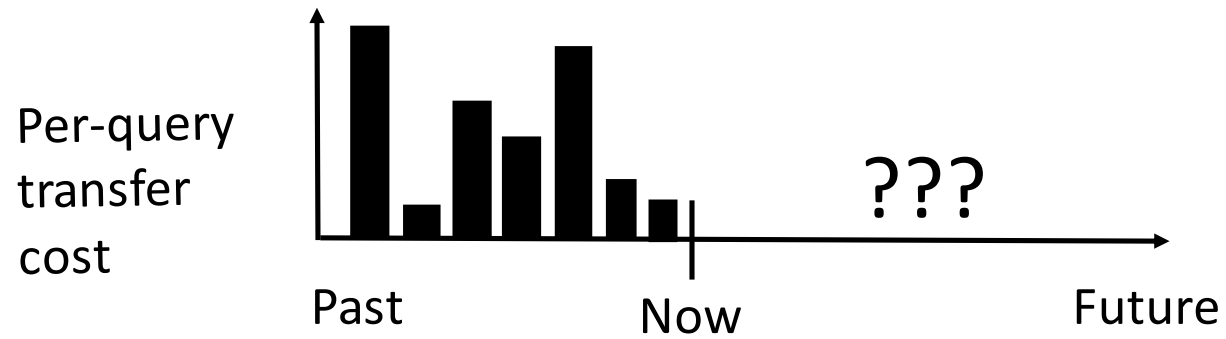
Option B:

Replicate raw data.

Replication cost
(one time)

Many large
overlapping results

Problem



Future demand is not known in advance!

Replication strategy

Strategy determines when data is replicated given a record of its past accesses.

Naïve

- Replicate immediately.
- Replicate never.



Data-dependent

Optimal Offline

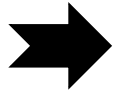
- Replicate immediately, if future demand is larger than replication cost.



Requires knowledge of future

Can we do better?

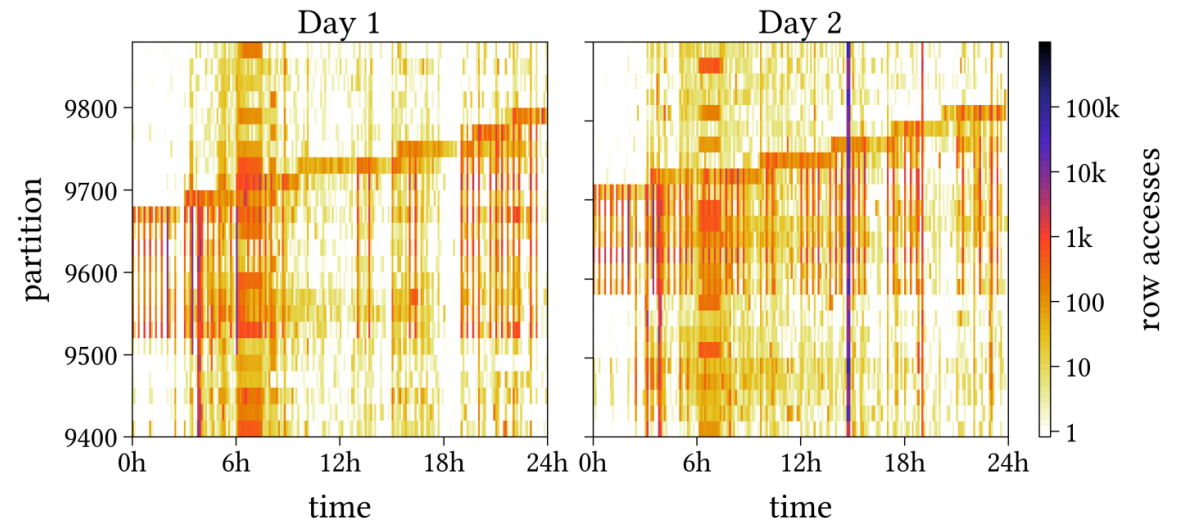
Data Organization: Partition



- Data is immutable.
 - e.g., machine logs
- Data is partitioned.
 - Space: e.g., by machine, by location, etc.
- A partition is accessible for a time window.
 - then removed or archived.

Dataset

- Trace of an ERP database of a Global 2000 company.
- Accesses at row-level.
- Partition := 10k rows
- Time window := 1 day



Note: logarithmic color-scale!

Potential reduction

- Cumulative cost :=
 - Sum of query result sizes sent over time window
- Replication cost :=
 - Partition size x replication cost factor

Replication cost factor depends on compression, overhead, ...

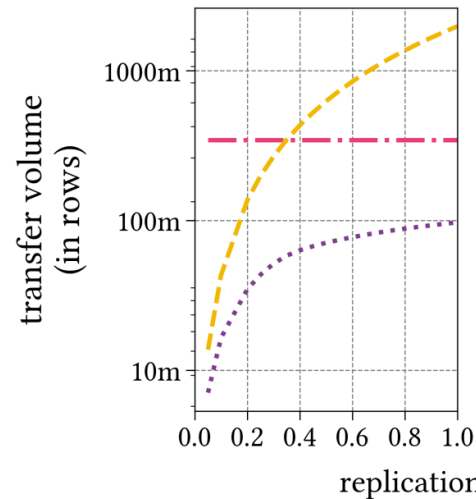
Cheap replication



Costly replication



*>50%
potential
reduction*



Replication Strategies

I. Competitive

- Guaranteed worst-case performance.

II. Heuristic

- Exploit access traces.

III. Hybrid

- Combination of above.

Strategies: Competitive

Competitive Strategy

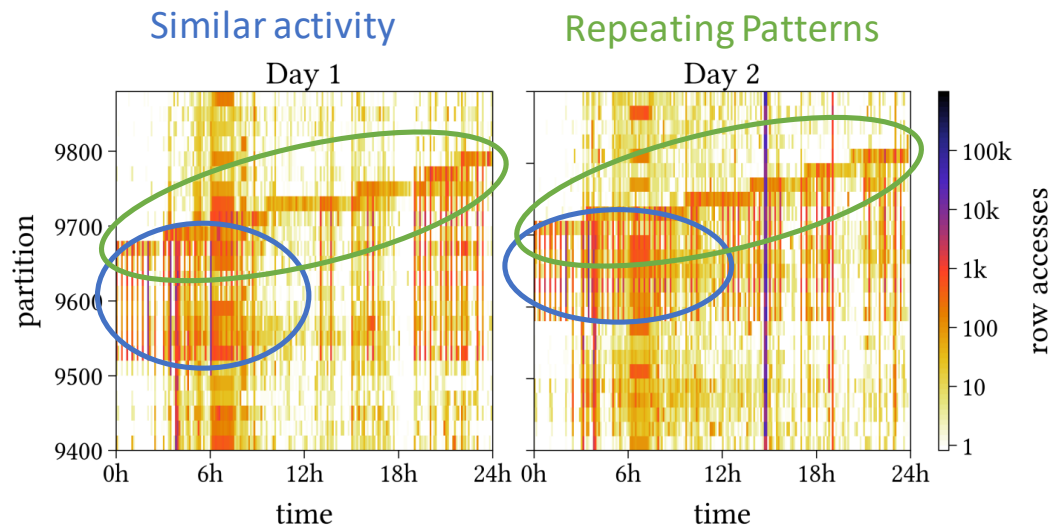
A strategy that has a bounded worst-case performance in comparison to the optimal offline strategy.

Ski-rental (Karlin et al.)

- Use threshold to decide replication.
 - If past transfer cost $>$ replication cost: replicate!
- 2-competitive algorithm.
 - Provably best worst-case bound.

Why do we need more than this?

Dataset Insights

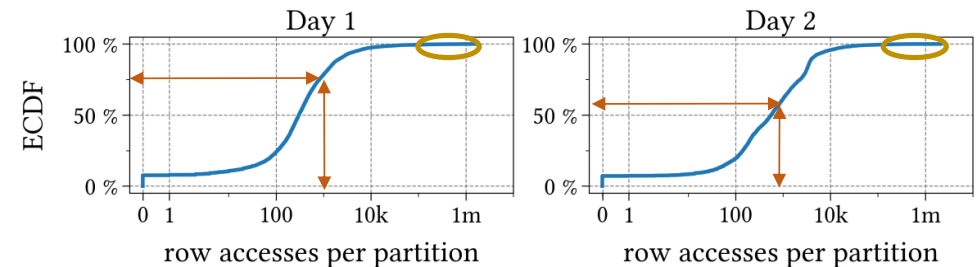


Does popularity depend on location?

Do popular partitions exhibit patterns of activity?

> 50% partitions have < 1k accesses

< 1% partitions have > 100k accesses



Skewed distribution:

Accessed partition is more likely to be accessed in the future than not.
Ski-rental does not use this!

Strategies: Heuristics

- Last-partition
 - Replicate if partition in previous time window exceeded replication cost.
- Last-threshold
 - Compute best threshold over partitions in past time window.
- Machine learning classifier (Random Forest)
 - Classify patterns into exceeding/not exceeding replication cost.
 - Replicate if accesses pattern match.

Strategies: Hybrid

- Replicate if either Ski-rental OR Classifier replicate.
 - Configure ML to be conservative.
- Goal: Replicate earlier than pure Ski rental → avoid transfers.

Replication Strategies

I. Competitive

- Ski-rental

II. Heuristic

- Last-partition
- Classifier
- Last-threshold

III. Hybrid

- Ski-rental *OR* Classifier

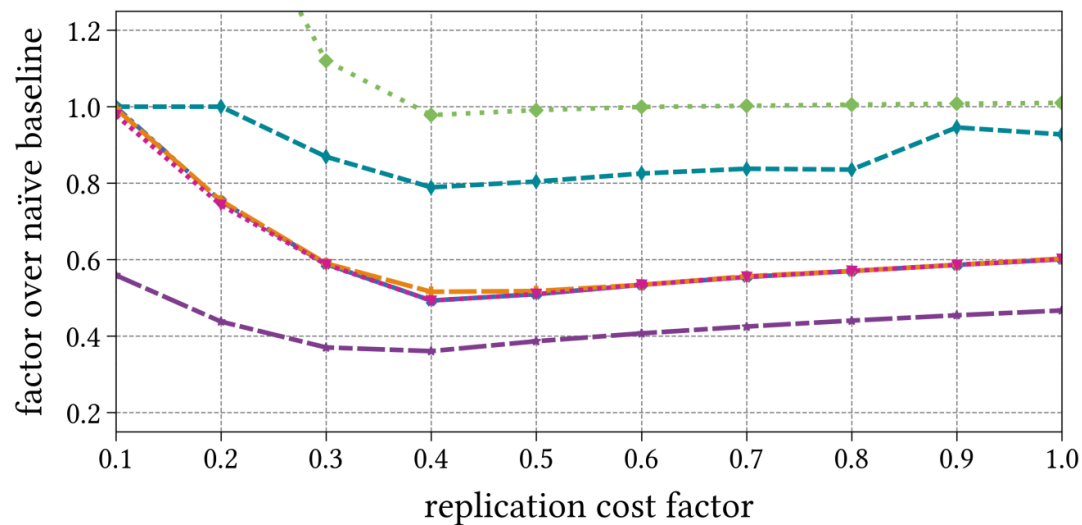
VS

Naïve Baseline

$\min(\text{Replicate-all}, \text{Replicate-nothing})$

Optimal Offline

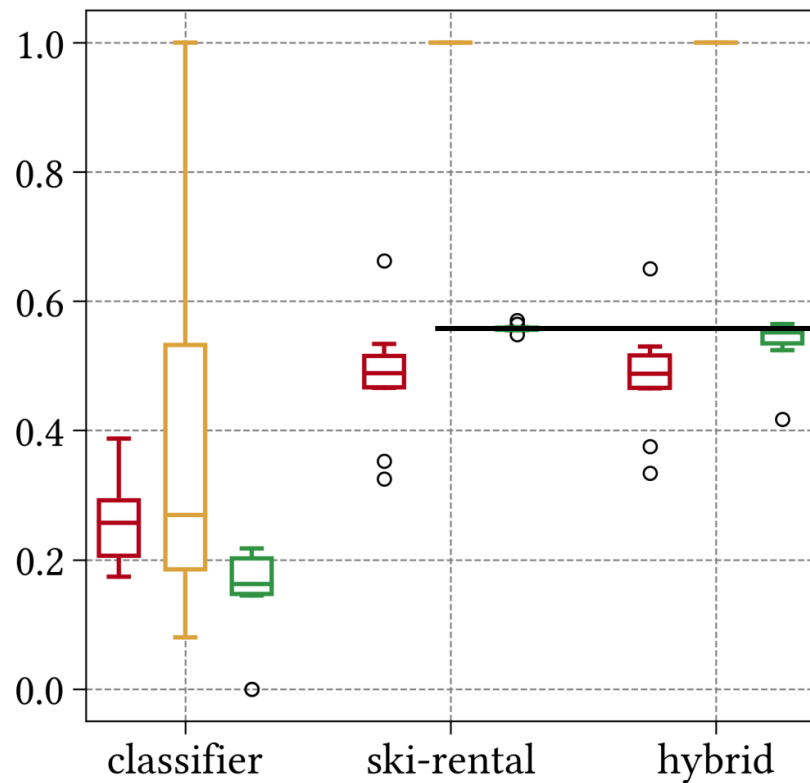
Transfer Cost Reduction



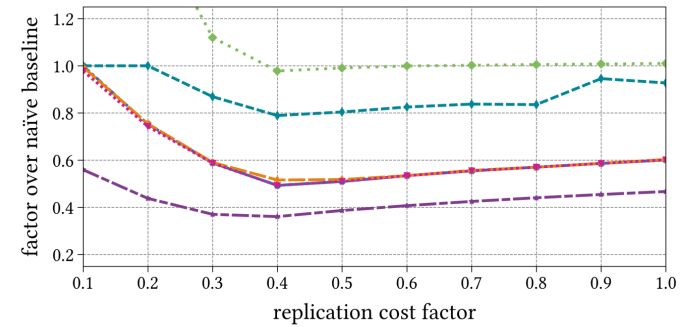
Insights

1. **Ski-rental** achieves 38% reduction on average. Up to 50% for some cases.
2. **Last-partition** performs poorly.
3. **Last-threshold** close to **ski-rental**.
4. **Classifier** worse than **ski-rental**.
5. **Hybrid**: Small improvement.

Transfer Cost Reduction



— precision
— recall
— mean time of replication



— optimal offline — ski-rental — last-threshold
— last-partition — classifier — hybrid

Hybrid: Slight improvement in replication timing.

Conclusion

- Introduced replication strategies.
 - Ski-rental reduces transfers by 22%/50% on average/best-case.
 - Hybrid strategy improves performance by 25%/51%.
- } Both traces

Ongoing work

- Improve machine learning.
- Include other cost factors (storage, etc.)

*Interested in the performance on your data?
Please contact us: niklas.semmler@sap.com*

Thank you!