

Proteus: taming heterogeneity through virtualization & JIT adaptivity

Data-Intensive Applications and Systems Laboratory

Catching up with an Evolving Landscape



Data, Hardware & Workload Virtualization



Reporting Server Spreadsheet

Next generation systems must adapt

Virtualize and JIT specialize to remove overheads

Elasticity & Isolation in Heterogeneous HW

Adaptively Navigating the HTAP Spectrum





HW- to performance-isolation via logical partitioning

Reconcile fresh data movement with HW efficiency

Heterogeneous Hardware

Heterogeneous Data



Efficient execution via Accelerator-level Parallelism

Adapt access paths to data & queries

Fast Analytics on Fresh Data through JIT Code Generation & GPU-acceleration

Code Generation GPU-accelerated

In-memory speed, fresh data

- Multi-CPU, multi-GPU execution
- Freshness-aware access paths
- JIT access & execution

More than in-memory data

- High-bandwidth NVMe arrays
- Hardware-conscious, proportional caching
- Optimize memory utilization

Faster than a scan

- Approximate to exceed hardware limits
- Line-speed sampling, reuse & cache
- Interactive data exploration





Reduce volume

Approximate

Full data: exact result

Sample: fast & accurate result



Hardware-specialized algorithms & mappings





