Research @ Databases and Distributed Systems Group

TECHNISCHE UNIVERSITÄT DARMSTADT
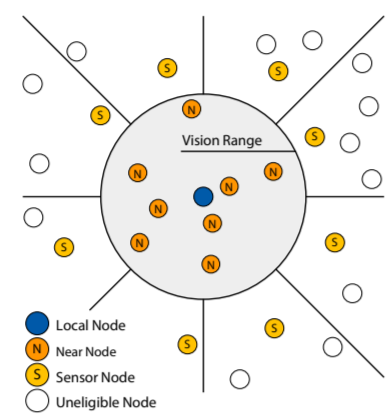
DVS

Planet PI4
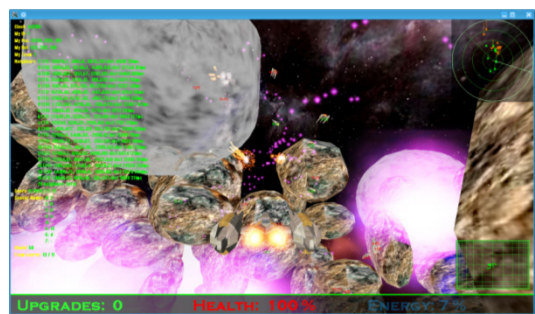
- Massively multiplayer online game prototype for the evaluation of (P2P) overlays
- Parameterizable space scenario
- Configurable bots for workload generation

pSense

- Dynamic P2P overlay for virtual environemnts
- Low-latency communication within vision range

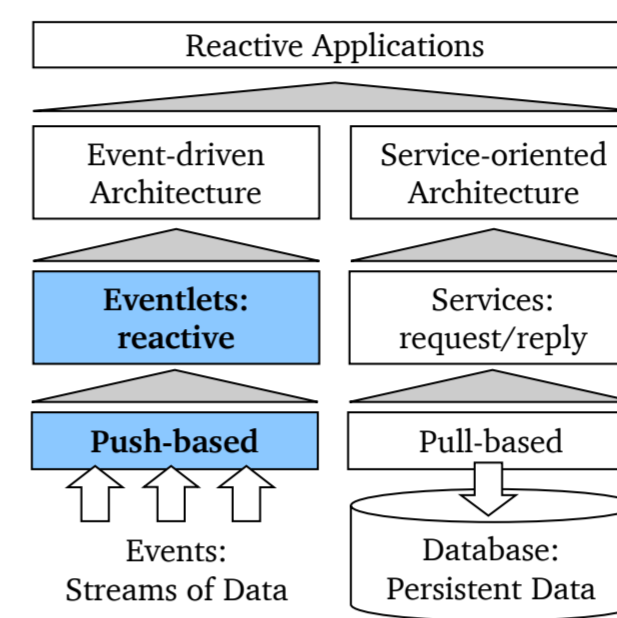Legend: Local Node, Near Node, Sensor Node, Unreliable Node

BubbleStorm

- P2P information management system
- Rendezvous search
- Different replication modes
- Store/retrieve, publish/subscribe, and more complex usage patterns

BubbleStorm
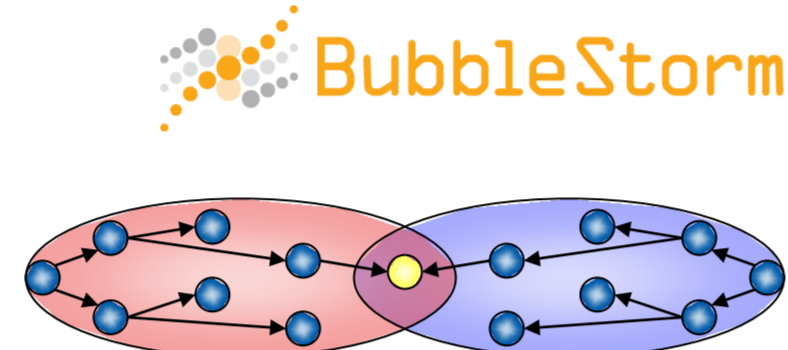
Reactive Enterprise Software Systems

- Integration of events and event streams with business applications
- Support for event streams from business process modeling to business process execution
- *Eventlets* as push-based equivalent to pull-based services

Reactive Applications

| Event-driven Architecture	Service-oriented Architecture
Eventlets: reactive	Services: request/reply
Push-based	Pull-based
Events: Streams of Data	Database: Persistent Data

SPECjms2007

- Standard benchmark for performance evaluation of message-oriented middleware that implements the Java Message Service (JMS)
- Jms2009-PS: Modified version to test publish/subscribe communication
- Realistic workload: supply chain management of supermarkets

spec
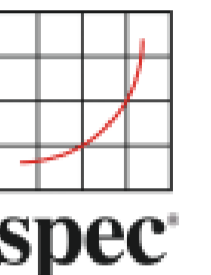
Novel Transport Protocols - CUSP

- Generic transport protocol optimized for the demand of P2P applications
- Variable number of potentially short-lived, unidirectional streams on a common channel
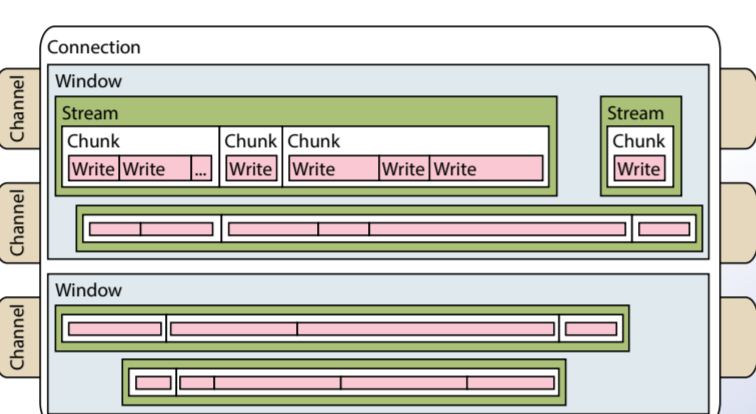
Contextualization

Events need to be put into context to be usable:

- Dynamic integration of software systems across enterprises
- Transformation of events to cope with different semantics and data formats (ACTrESS)
- Contextualize by adding external information on demand
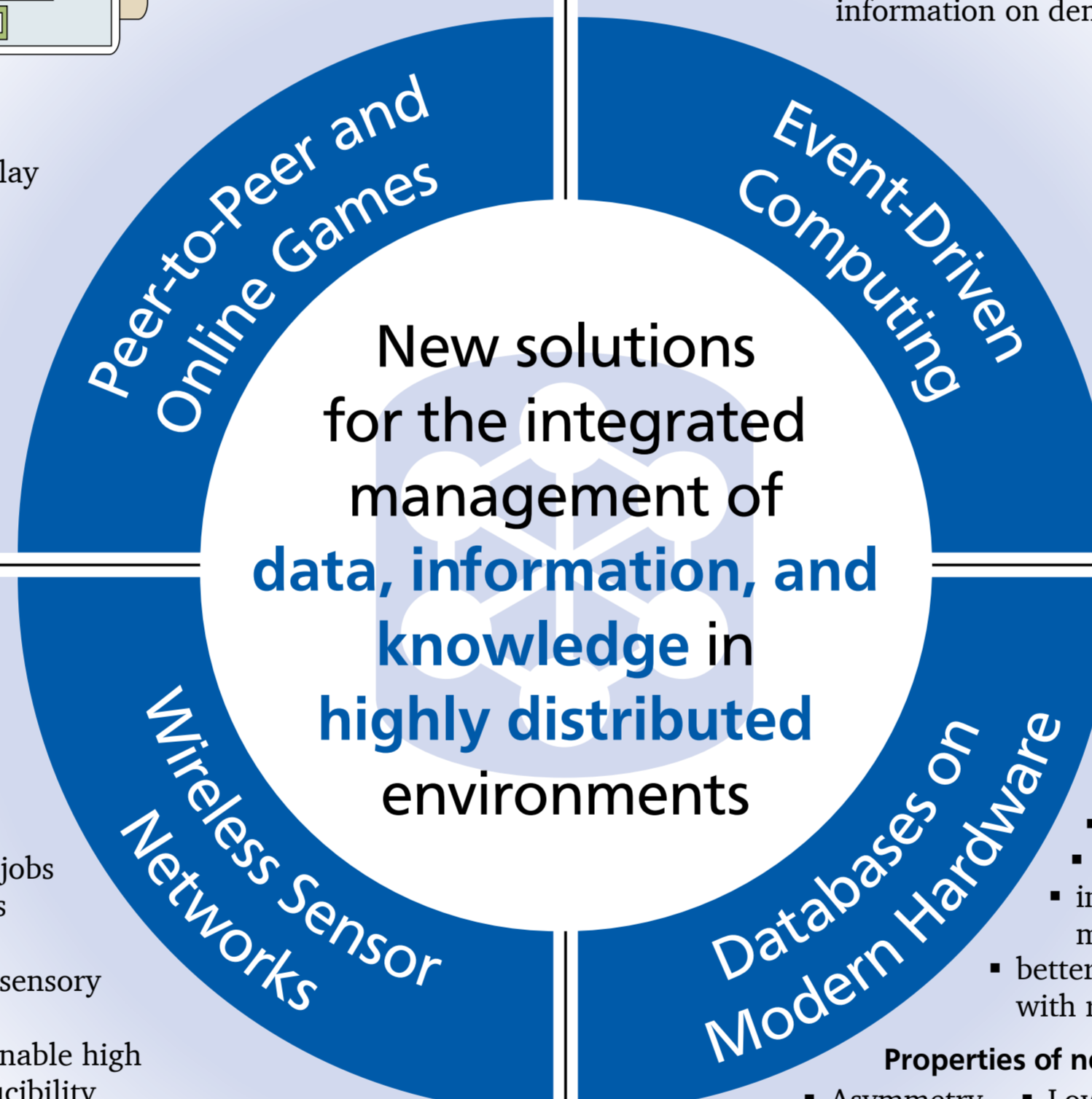
Aggregation for Pub/Sub Systems

- Lightweight support for in-network aggregation of information in content-based publish/subscribe systems, e.g., number of subscribers
- International cooperation between TU Darmstadt, Purdue University, Imperial College London, University of Otago, and Vrije Universiteit Amsterdam
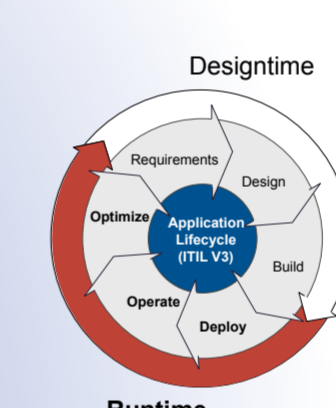
ASIA

Overlay Network Testbed
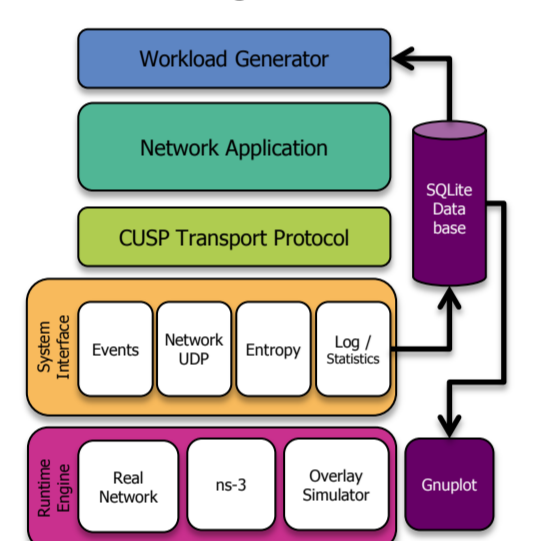
- Testbed for the evaluation of overlay networks and their applictions
- Deterministic simulation and real network mode using the same application code
- Configurable workloads
- Central database of results, analysis tools

Workload Generator
Network Application
CUSP Transport Protocol
System, Events, Network UDP, Entropy, Log / Statistics
Real Network, ns-3, Overlay Simulator
SQLite Data base
Gnuplot

Peer-to-Peer and Online Games

Event-Driven Computing

Wireless Sensor Networks

Databases on Modern Hardware

New solutions for the integrated management of **data, information, and knowledge** in **highly distributed** environments

Governance in Event-based Systems

- Service level agreements (SLA) are an essential mechanism in service-oriented architectures
- Integration of event-based components in enterprise applications requires SLA concepts suitable for the push-based nature of event integration in enterprises

Designtime, Requirements, Design, Application Lifecycle (ITIL V3), Build, Deploy, Operate, Optimize, Runtime

myHealthAssistant

- Smartphone middleware for body sensor networks (BSNs)
- Handles sensor communication
- Provides data abstraction to application
- Transforms sensor data into required data abstraction
- (Preventive) health care applications built on-top for vital sign, fitness workout and daily activity monitoring

myHealthAssistant
ukuFlow
Scopes
A3ME CEACH
TUDμNet
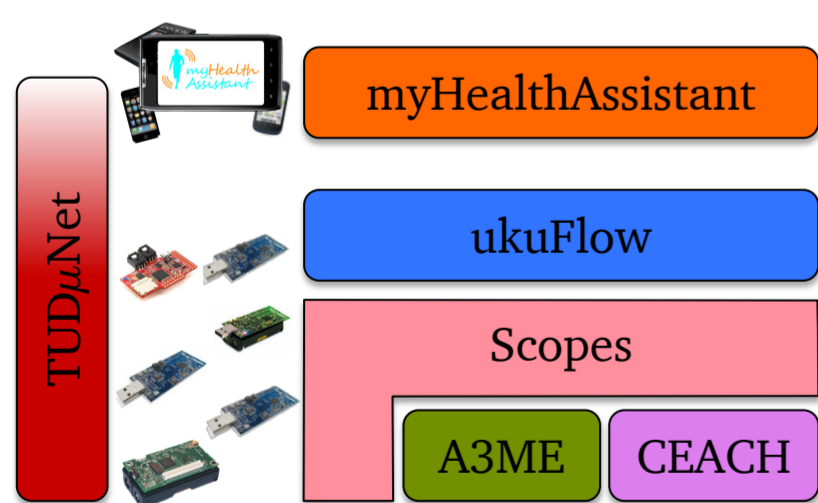
TUDμNet

- Federation of multiple, autonomous WSN testbeds:
  - multiple, application-specific domains
  - centralized coordination, hierarchical zones, parallel jobs
- Combine best from simulators and real-world deployments
  - real platform, wireless and sensory environment
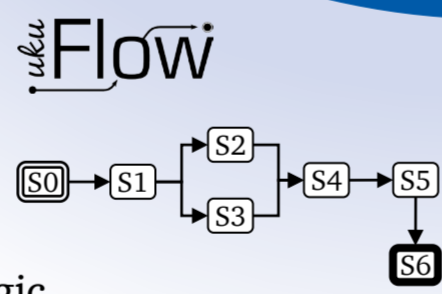 - simplify experimentation, enable high level of experiment reproducibility

ukuFlow

ukuFlow

- **Workflows** as mechanism to define WSAN application logic
- Extensions to Business Process Modeling Notation (BPMN) v2.0 to define app. logic:
  - **imperative** and **event**-driven macro-programming
  - abstracts away internal complexities of routing, grouping, data collection, event detection and action execution, among other
- entirely **in-network** workflow engine for low power, 8/16-bit embedded microcontrollers
- virtual machine (interpreted)

S0 → S1 → S2 → S4 → S5 → S6 / S3

Scopes

- Multiple, parallel apps. on sensor networks
- Declarative creation of node groups, called **scopes**:
  ```
  CREATE SCOPE Room6Motion
  (EXISTS SENSOR Motion)
  AS SUBSCOPE OF Room22;
  ```
- Bidirectional comm. link between members and root node
- Automated maintenance: support for network dynamics (node churn)

A3ME

- Enable **interoperability** between devices: sensor nodes, mobile phones, etc.
- **Device agents** to model entities
- Common, extensible agent **ontology** to classify device agent properties
- Query language, encoded in ASN.1 Packet Encoding Rules

CEACH

- Clustering service for WSN traffic
- Considers different aspects for cluster creation:
  - dynamic properties (e.g., residual energy)
  - link quality metrics (both HW and SW)
  - node degree (size of 1-hop neighborhood)
  - network aspects (hop distance)
  - use EWMA for cluster quality assessment
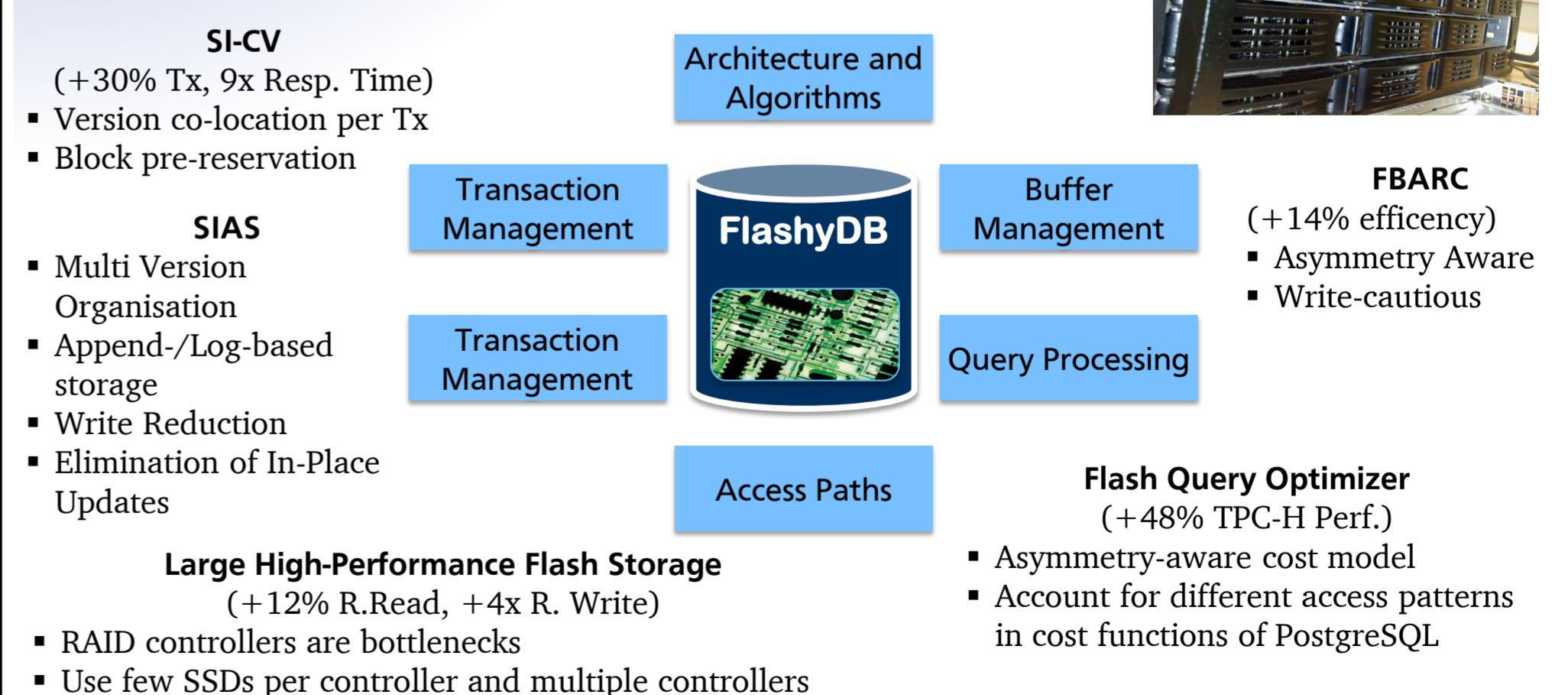- Adaptive node switching between clusters

Databases on Modern Hardware

The architecture and algorithms of database systems are designed around the characteristics of legacy hardware:

- Access gap | Low Memory/Disk volume ratios | few CPUs
- Spinning Disks - HDD

Trends in Hardware [1]:

- many-core CPUs - 1000 cores per chip by 2022
- growing RAM volumes - more than 128 TB RAM per server by 2022
- increasing interconnect bandwidth memory: 2.5 TB/s, I/O: 250 GB/s by 2022
- better I/O technologies - fast 1TB Flash chips and Non-Volatile Memories with read performance similar to that of DRAM

Properties of new IO Technologies:

- Asymmetry
- Low latency → shrinking access gap
- Endurance
- Parallelism
- Addressability (Byteadress vs. Blockaddress)
- Access Patterns (random, sequential | read, write)

What is the impact on data-intensive systems?

SI-CV (+30% Tx, 9x Resp. Time)
- Version co-location per Tx
- Block pre-reservation

SIAS
- Multi Version Organisation
- Append-/Log-based storage
- Write Reduction
- Elimination of In-Place Updates

Architecture and Algorithms
Transaction Management
FlashyDB
Buffer Management
Transaction Management
Query Processing
Access Paths

FBARC (+14% Efficency)
- Asymmetry Aware
- Write-cautious

Flash Query Optimizer (+48% TPC-H Perf.)
- Asymmetry-aware cost model
- Account for different access patterns in cost functions of PostgreSQL

Large High-Performance Flash Storage (+12% R.Read, +4x R. Write)
- RAID controllers are bottlenecks
- Use few SSDs per controller and multiple controllers

Efficient use of new hardware is a dominating trend in data-intensive computing, influencing the architecture, and algorithms

[1] Andreas von Bechtolsheim. Technologies for Data-Intensive Computing. HTPS 2009.

QuaP2P DFG FlashyDB topologie der technik spec ADiWa cocoon CASED PLANETLAB Software-Cluster Dynamo PLV software campus