Technische
Universität
Braunschweig

Institute of Operating Systems
and Computer Networks

Temporality a NVRAM-based virtualization platform

Vasily A. Sartakov, Arthur Martens, Rüdiger Kapitza

34th IEEE Symposium on Reliable Distributed Systems (SRDS '15)
Montréal, Canada

Agenda

- Motivation
- Hardware and Assumptions
- Architecture and Implementation
- Evaluation
- Conclusion

Motivation

Common failures (55%) in data centers are related to power outages*.

- Partial: 56 minutes
- Complete unplanned: 119 minutes

Cost: $8k per minutes but also indirect costs:

- Loss of revenue
- Decreased performance
- Loss of temporary data
- Recovery causes high load on background storage

*Source: 2013 Cost of Data Center Outages, Ponemon Institute, 2013.

Motivation

Additional, expensive, fragile <..> hardware:

- UPS
- Backup power supplies
- Additional hardware – replications, mirroring, etc.

NVRAM-based way

- Decrease recovery time
- Don’t loose performance

Non-volatile random-access memory

NVRAM

- New technologies: STT-RAM, PCM, MRAM, FeRAM
- Retain data without external source of power
- DRAM-like performance (not all)
- Prospective technologies provide more than DRAM density
- Byte-addressable access

Main memory can be persistent and storage-sized

Hardware and Assumptions

Battery-backed RAM (NVDIMMs)

- Ordered and atomic writings
- DIMMs equal performance

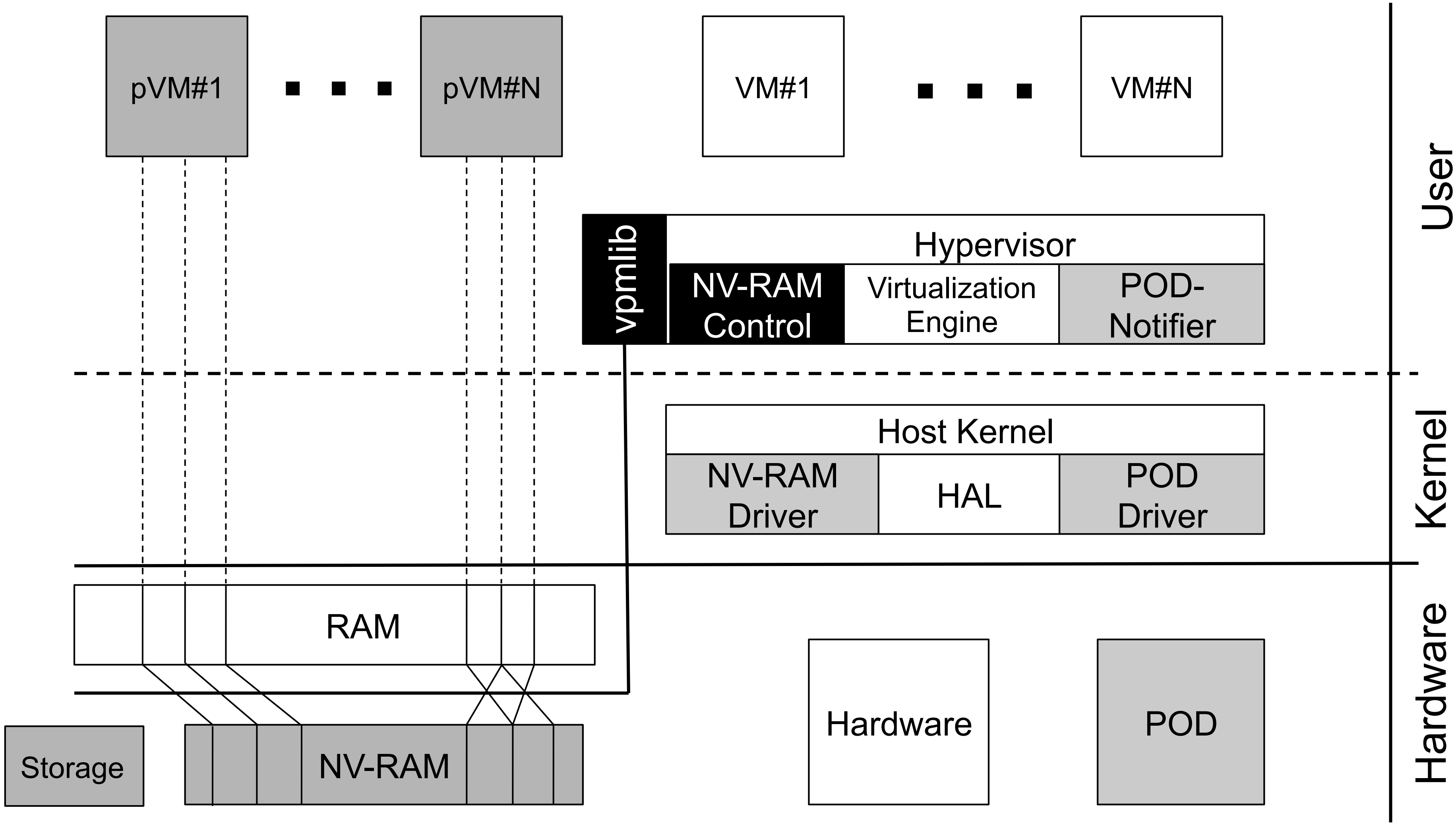
Volatile CPU, Volatile devices

Residual energy

Power Outage Device

G. Heiser, E. Le Sueur, A. Danis, A. Budzynowski, T.- I. Salomie, and G. Alonso, “RapiLog: reducing system complexity through verification,” in *Proc. of the 8th ACM European Conference on Computer Systems*, 2013.

Architecture



Components

Life cycle:

- Normal boot, creation, execution
- Power outage: save the context of pVM, flush caches
- Full restart of host system (except NVRAM), reallocate pVMs
- Continue execution of pVM after power recovery

Name-based allocator

- Control sequence of allocation/reallocation

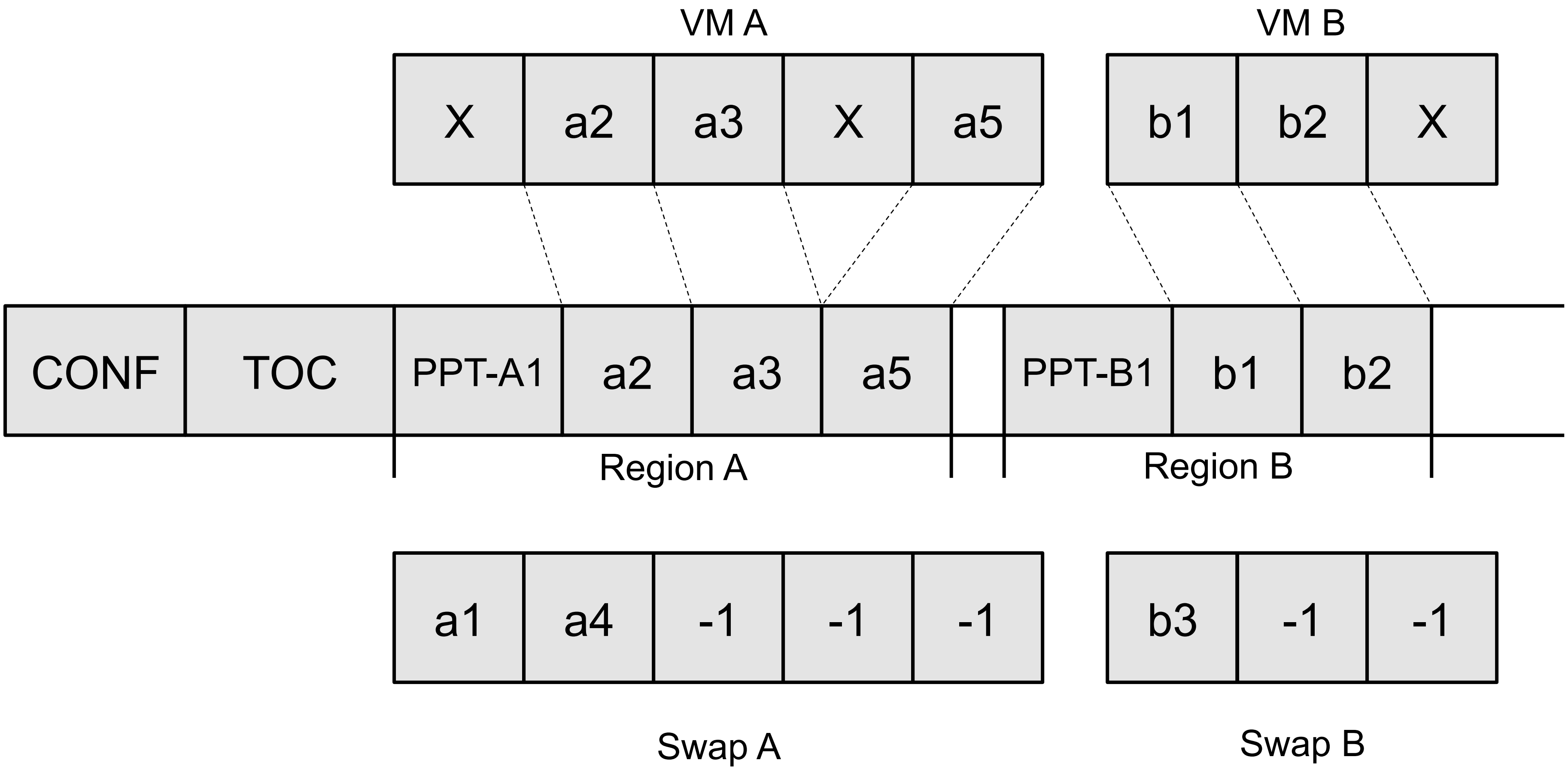
Virtual memory support

- Supports multiple pVM use

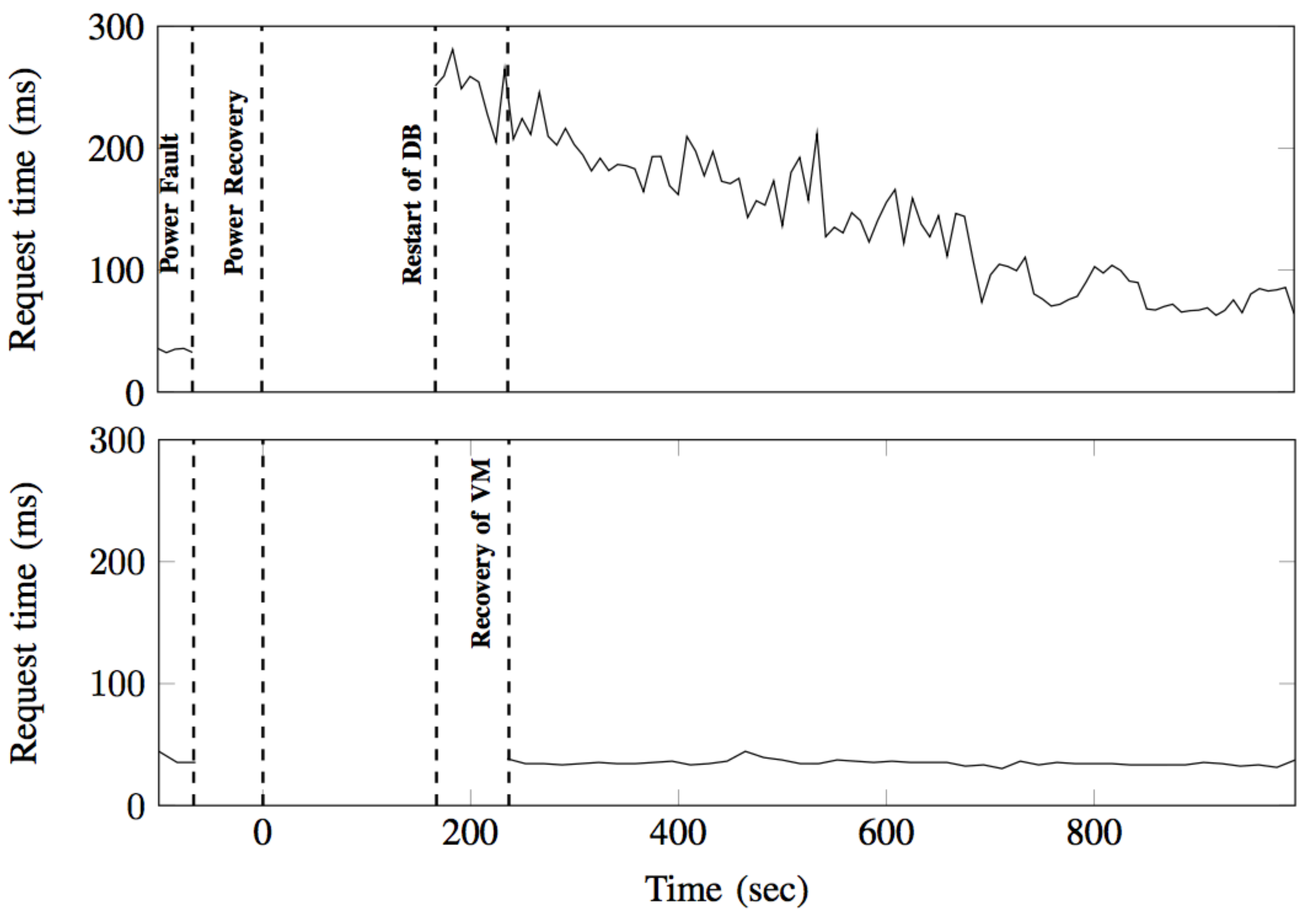
Persistent VMs:

- Device state (should be “fixated” in a moment of power outage)
- Memory image (Persistent memory and swap)

Name-based allocator with swapping

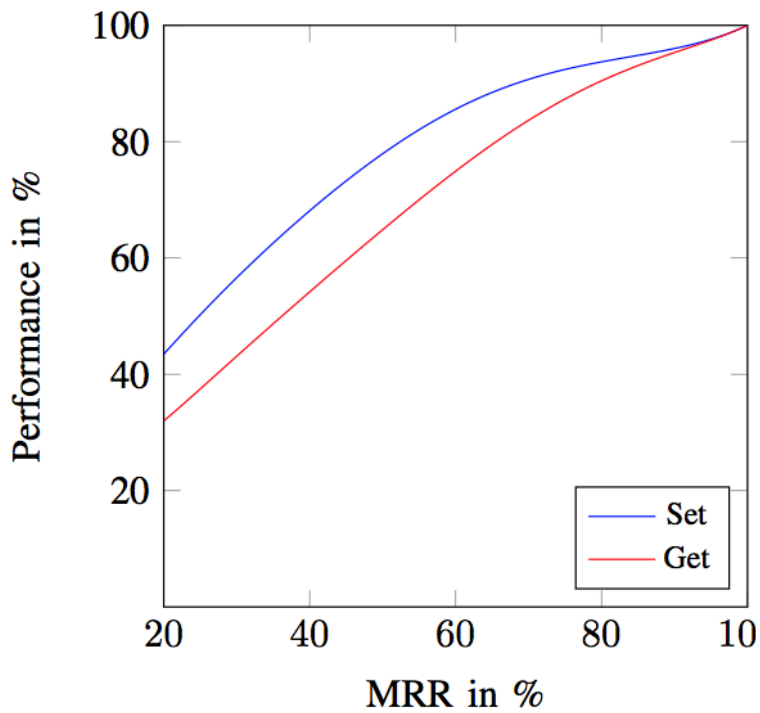


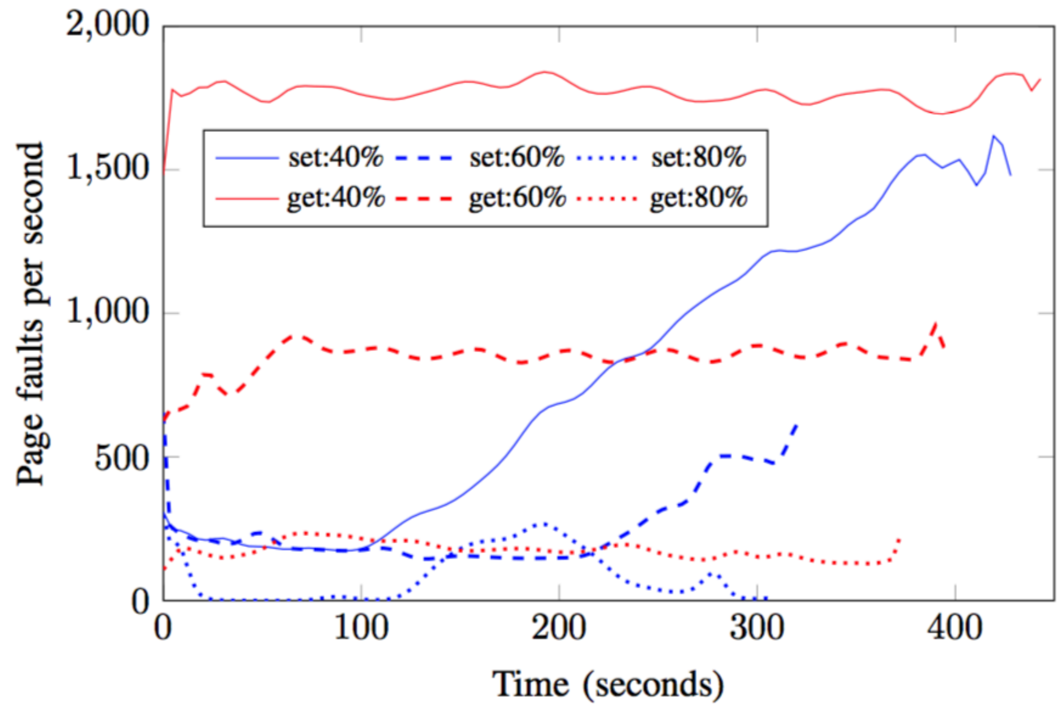
Performance recovery



V. A. Sartakov and R. Kapitza, “Nv-hypervisor: Hypervisor- based persistence for virtual machines,” in *Dependable Systems and Networks (DSN)*, 2014.

Memory Residence Ratio and Pagefaults per second





MRR = InMemory/(InMemory + InSwap)∗100%

Conclusion

- Temporality provides transparent persistence for commodity virtual machines
- It offers support for multiple VMs including support for virtual memory
- Using Temporality recovery time is decreased and temporary data is protected
- Validation uses market-available hardware

Thank you

Vasily A. Sartakov
Sartakov@ibr.cs.tu-bs.de