Enhancing Live Migration Process for CPU and/or memory intensive VMs running Enterprise applications

Benoit Hudzia
CEC Belfast / SAP Research
08/2011
With the contribution of Aidan Shribman and Petter Svard

SAP RESEARCH

Agenda

- **Background: Enterprise Applications and Live Migration**
- **Warm Up**
- **Delta Compression**
- **Page Priority**
- **Future Works**

SAP RESEARCH

© 2011 SAP AG. All rights reserved.

Background

Migrating Enterprise Class applications

SAP

SAP RESEARCH

Enterprise application and Live Migration

Issues

- **Enterprise class application:**
  - Bigger than average resource requirement
  - Average SAP ERP 16GB + per VM with 32 GB of swap more than common
  - OLTP system such as ERP are very sensitive to time variation.
  - Rely heavily on precise scheduling capabilities, triggers, timers and on the ACID compliance of the underlying

- **Challenge when migrating such application:**
  - Disconnection of services:
    - Gigabit Ethernet timeout ≈ 5 seconds (>500 MB memory left in stop and copy phase)
    - Downtime is workload dependent
  - Disruption of services:
    - Migration progressively increasing the amount of resource dedicated to itself => gradually degrade performance of the coexisting systems / VMs.
  - Difficulty to maintain consistency and transparency
  - Unpredictability and rigidity

© 2011 SAP AG. All rights reserved.

Warm Up for Live Migration

Increasing the flexibility of Live Migration

SAP RESEARCH

Warm Up

Increasing flexibility

Extended adaptive Pre-copy phase without triggering actual migration

Increased flexibility :

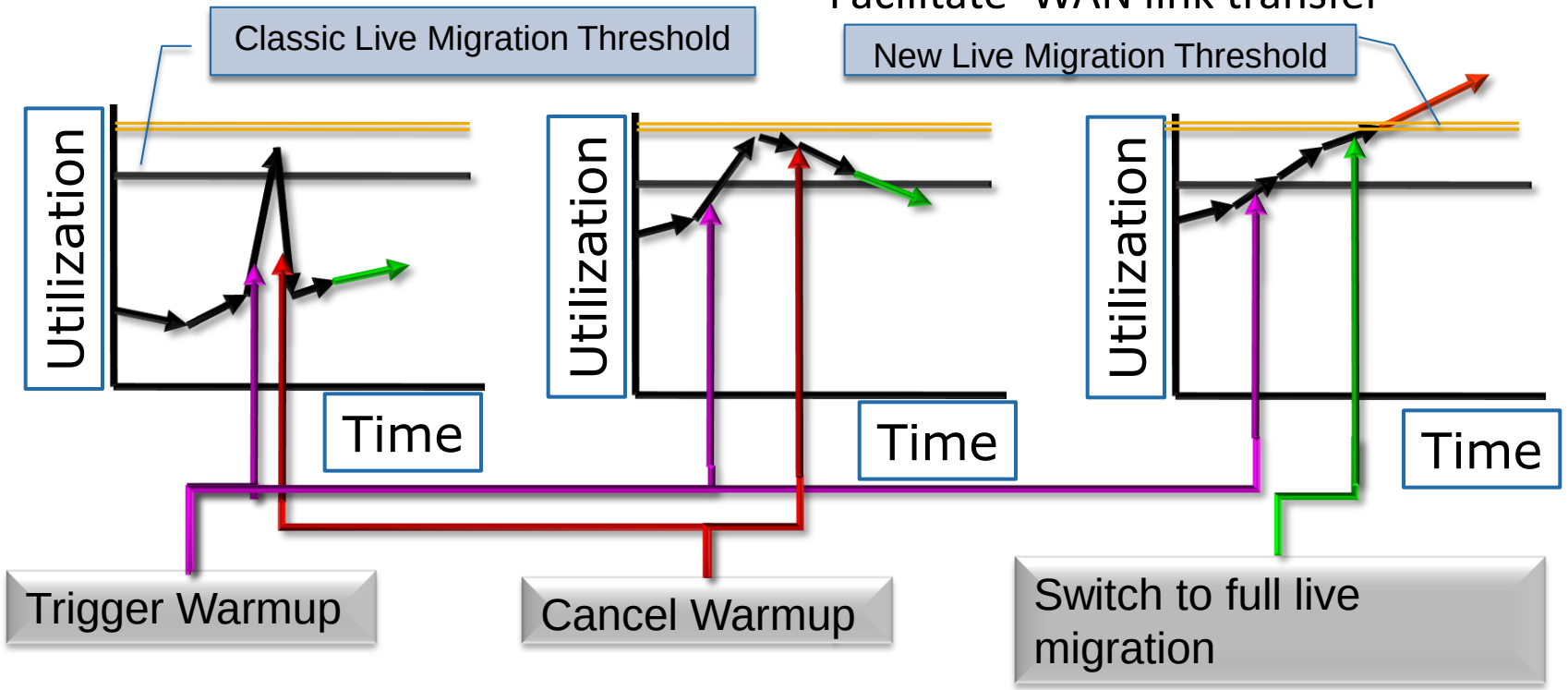
- "just in time" triggering of live migration
- Reduce down time

Dynamic adaptive bandwidth allocation

- Manual and automatic

Allow "hot standby"

Facilitate WAN link transfer

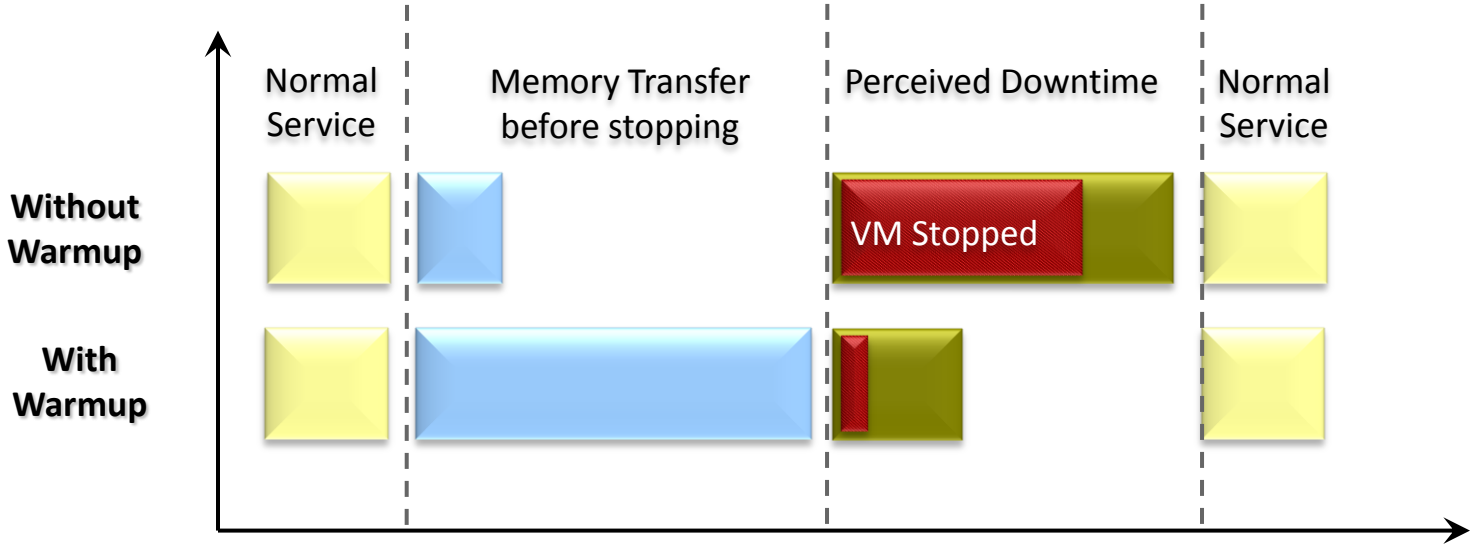


SAP RESEARCH

© 2011 SAP AG. All rights reserved.

Experimental Results: Warm-up Summary

SAP Sales and Distribution Benchmark



VM size : 4GB

SMP : 2 vCPU

Users : 150

Load ~= 80%

	CPU	Avg Response Time
Baseline	60%	2.18 sec
Warm-up	73%	2.16 sec

Downtime under load: <1 sec
Success ratio : ~99%

SAP RESEARCH

© 2011 SAP AG. All rights reserved.

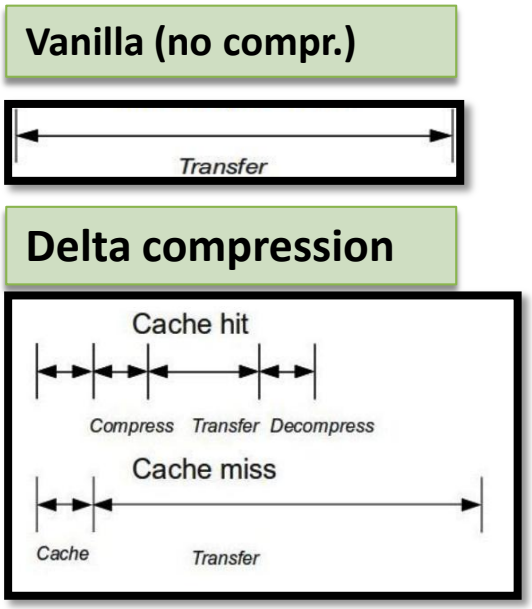
Delta Compression of Page

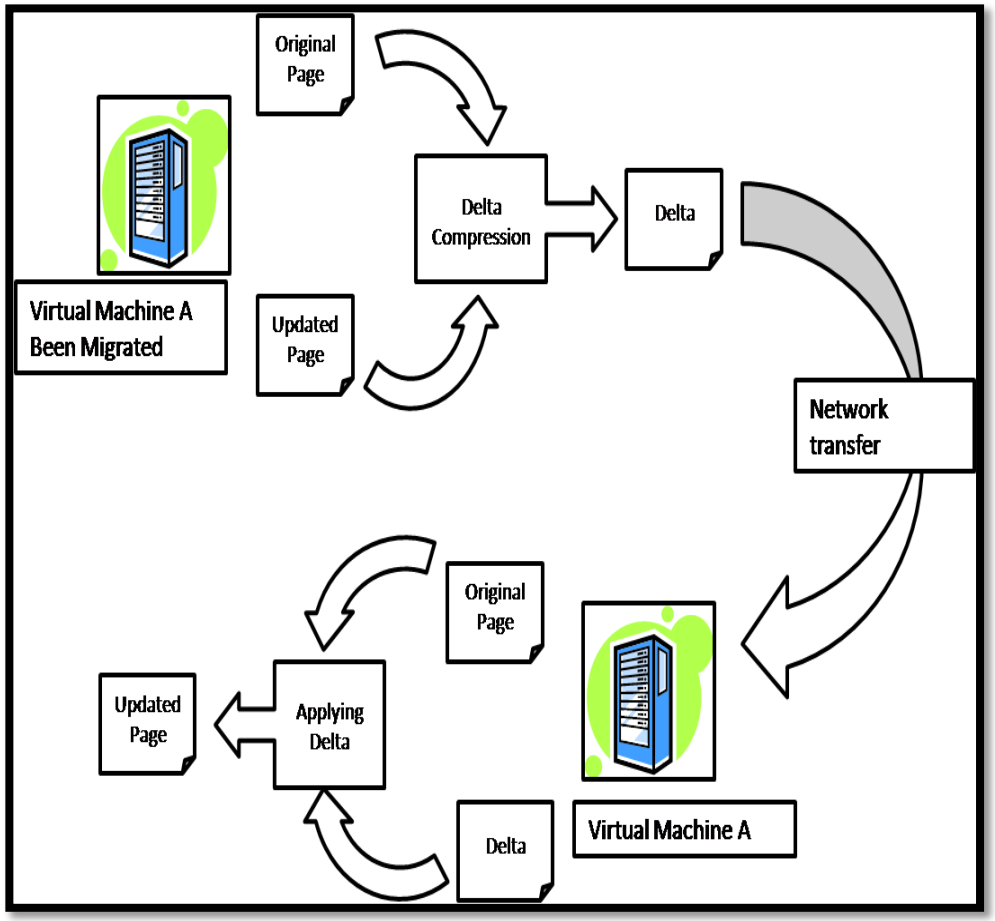
Limiting the impact of resending Page

SAP RESEARCH

Dirty Page Delta Compression

- Cache page with highest dirtying rate during send operation
- Compression Algorithm:
  - XBRLE : XOR +binary run length encoding





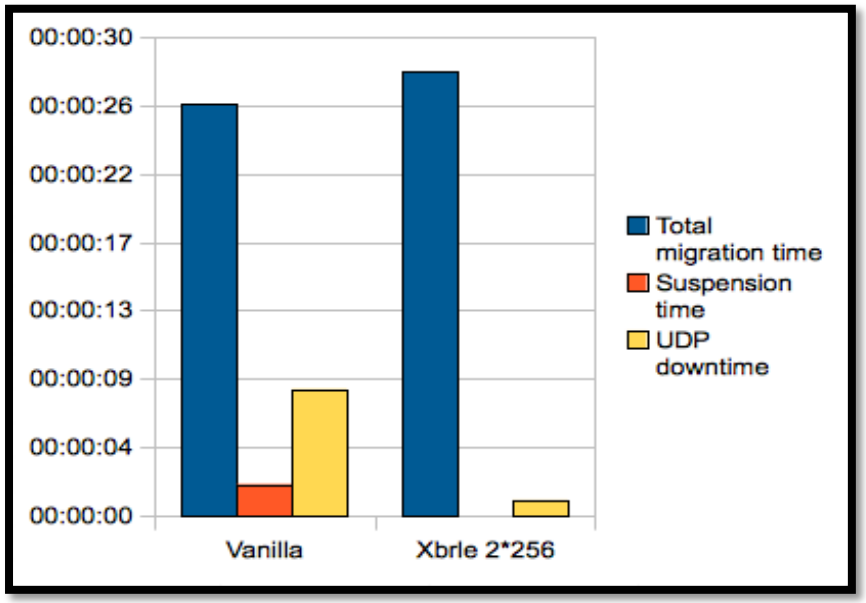
SAP RESEARCH

© 2011 SAP AG. All rights reserved.

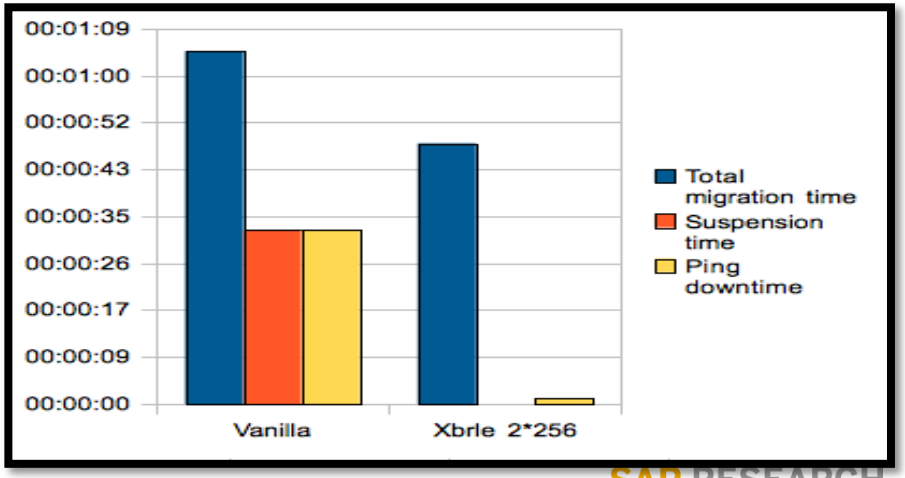
Evaluation

Benchmark

- **Memory write benchmark (lm_bench)**
  - 1 GB RAM, 1 vcpu VM
  - Near ideal case
  - Downtime reduced by a factor of 100
  - Throughput increased by 63 %



- **Transcoded HD Video (VLC)**
  - 1 GB RAM, 1 vcpu VM
  - Real-world, non-ideal case
  - UDP downtime reduced from 8 s to 1
  - Migration is transparent using XBRLE
  - 31% faster, 51% less data sent



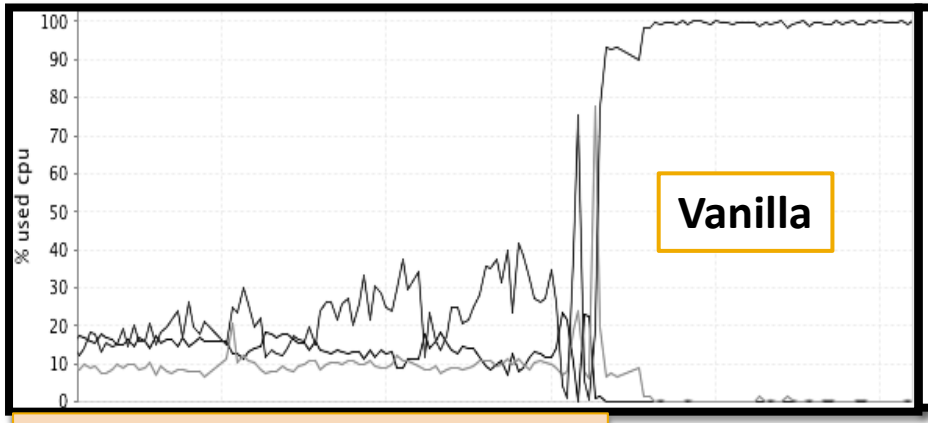
SAP RESEARCH

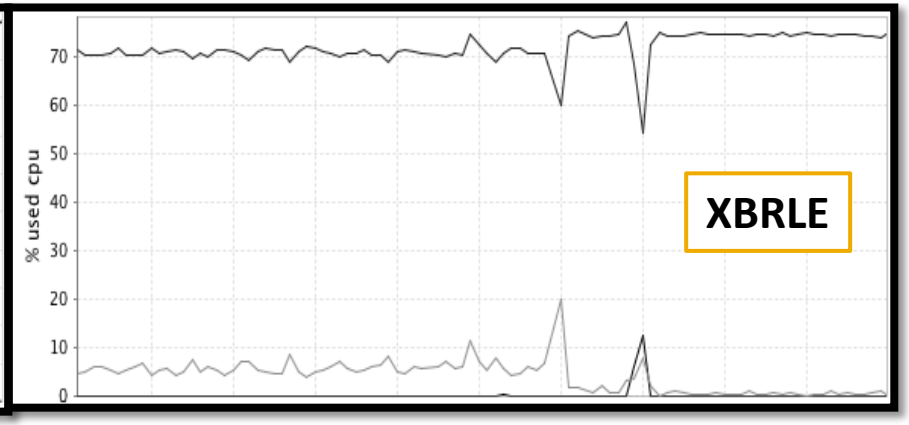
© 2011 SAP AG. All rights reserved.

Evaluation- SAP ERP

Sales and Distribution benchmark, load 100%

- Non-responsive on resume with vanilla algorithm
- Survived using XBRLE
- >0.5s of downtime = risk of damaging the system
- Measured downtime was 0.2s for XBRLE and 2s for vanilla
- Live Migration Cpu usage directly impact (limit) the available resource for the ERP





HW:4x 3,0GHz Xeon dual-core 32GB RAM
16TB Raid 5, 6Gbits/s trunked NFS server
1000Mbit/s Network

VM:8 GB RAM, 4 vcpus VM
App: SAP ERP 7.0 / S&D Benchmark

SAP RESEARCH

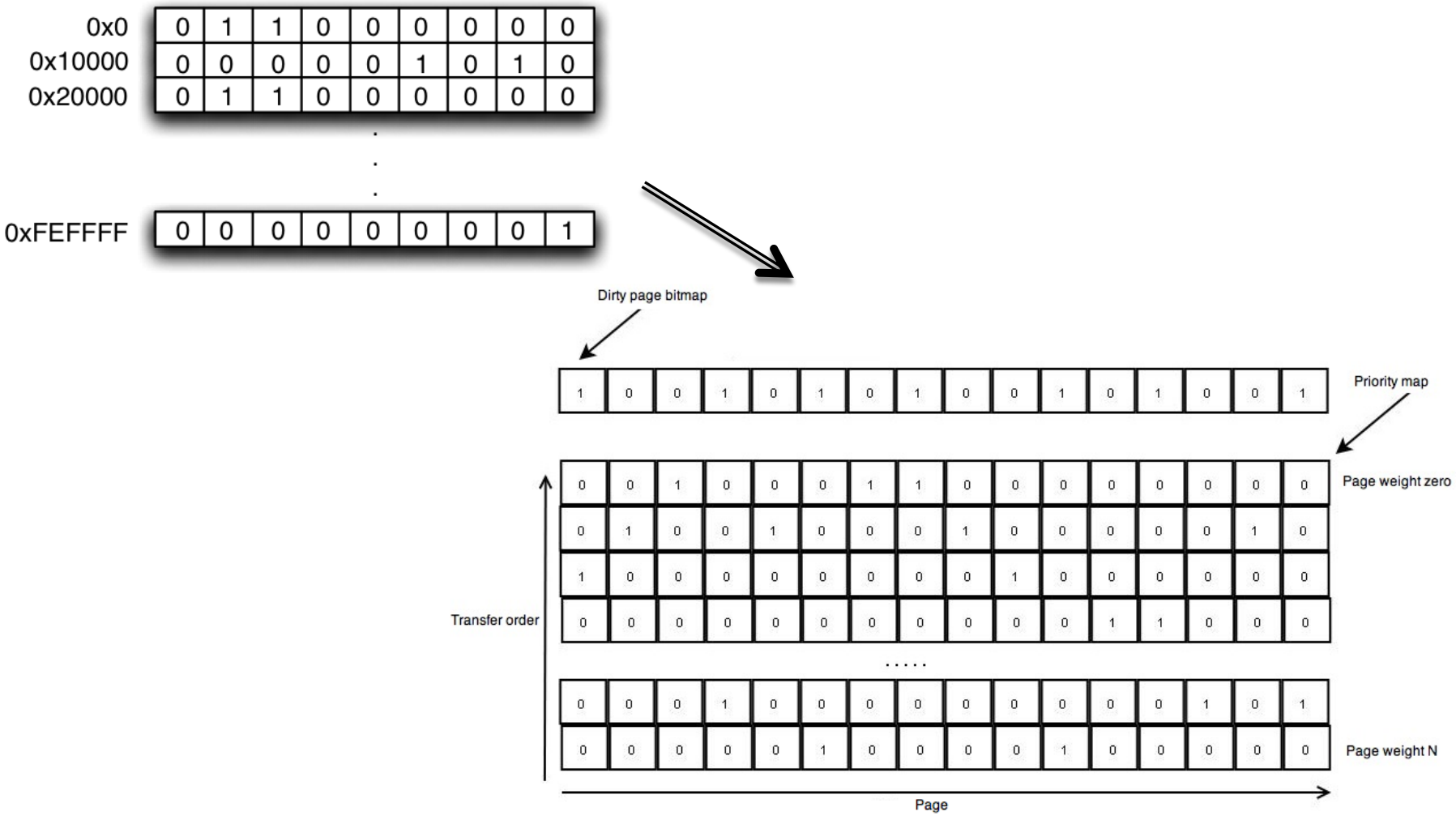
© 2011 SAP AG. All rights reserved.

Page Prioritization

Dynamic page transfer reordering

Dynamic page transfer reordering

Prioritizing page sends (similar to writable working set concept in Xen)



SAP RESEARCH

© 2011 SAP AG. All rights reserved.

Dynamic page transfer reordering

Prioritizing page sends

- Streaming HD video migration

	Total migration time	Transferred data
Vanilla	22.1 s	459 MB
PRIO	15.4 s	225 MB

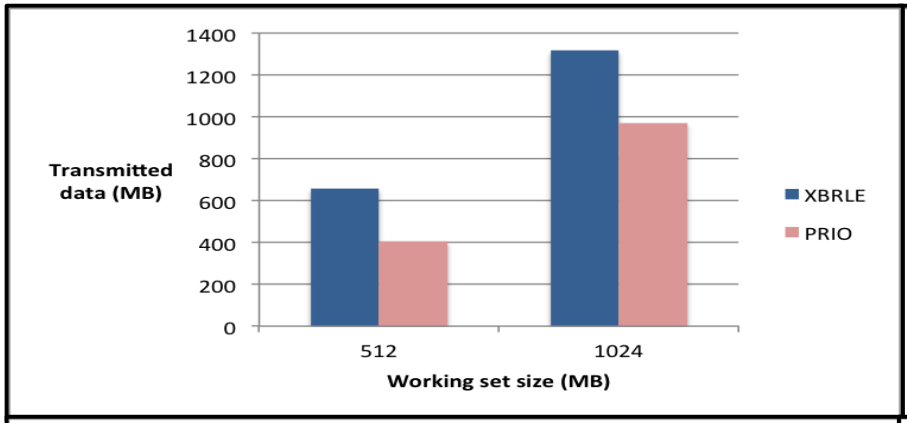
- 31% faster, 51% less data sent

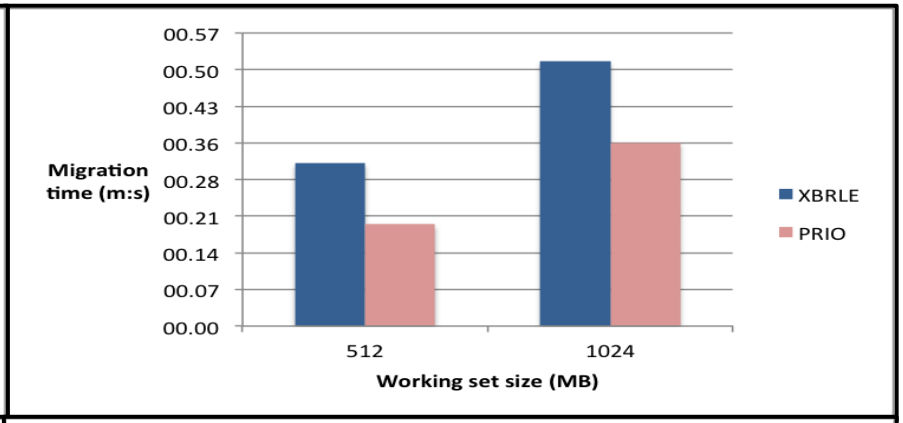
SAP RESEARCH

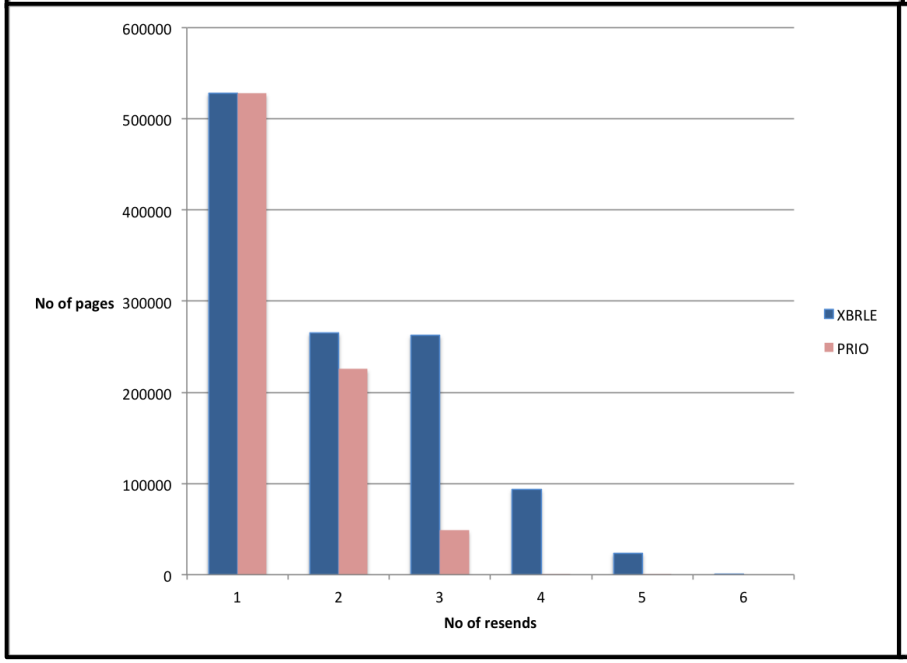
© 2011 SAP AG. All rights reserved.

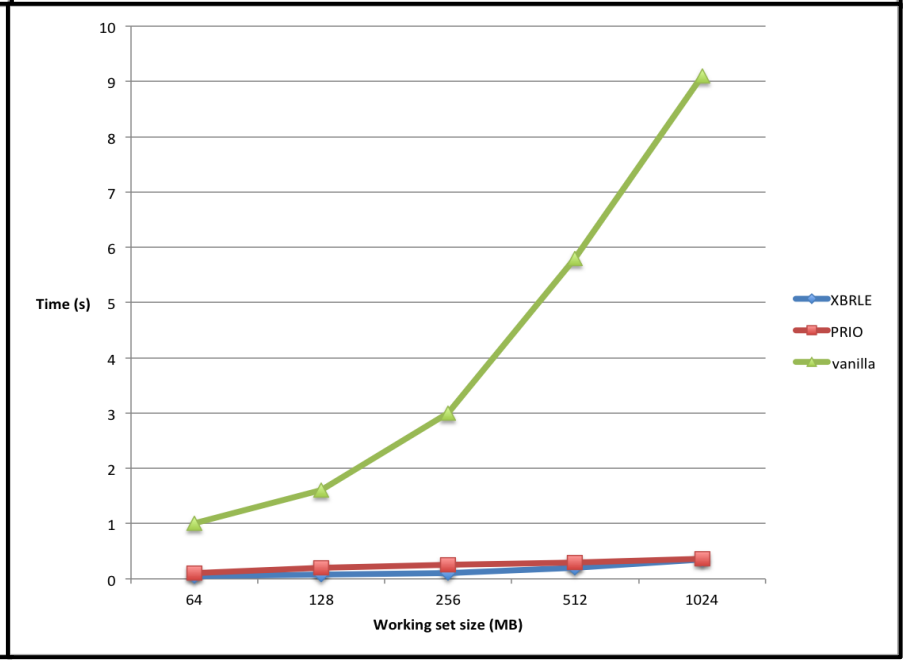
Evaluation

Prio vs XBRLE : reveal Cache miss and compression efficiency Issue









SAP RESEARCH

© 2011 SAP AG. All rights reserved.

Optimizing Compression

Making XBRLE more efficient

SAP RESEARCH

XBZRLE

Increase compression speed /efficiency

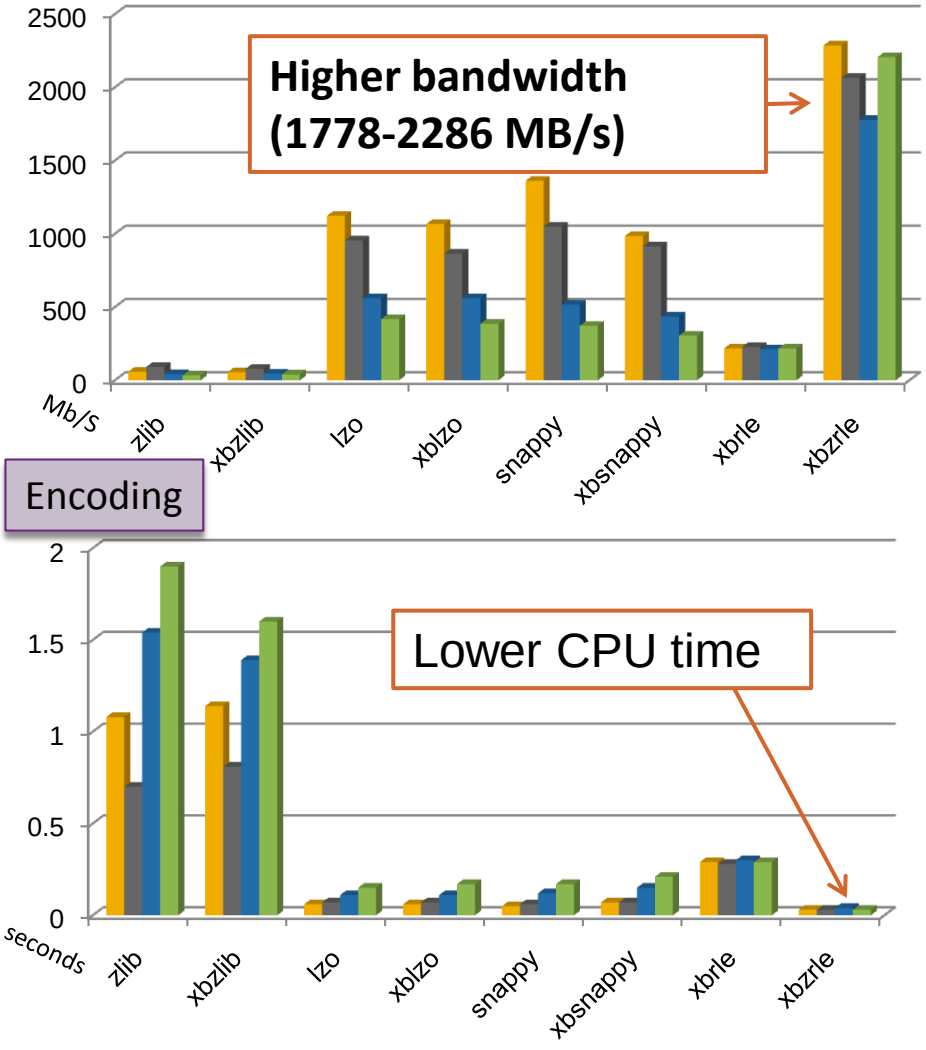
- Only compress unmodified data using word aligned encoding and only encodes runs of zeros
- For encoding page diffs XBZRLE is:
  - Compression :
    - 20% more efficient than XBRLE
    - 20% less efficient than LZO/Snappy.
  - Speed:
    - Overall 2.5x-5x faster than XOR + LZO/Snappy
    - 11x-9x faster than the original XBRLE
- Doesn’t solve the impact of cache miss

SAP RESEARCH

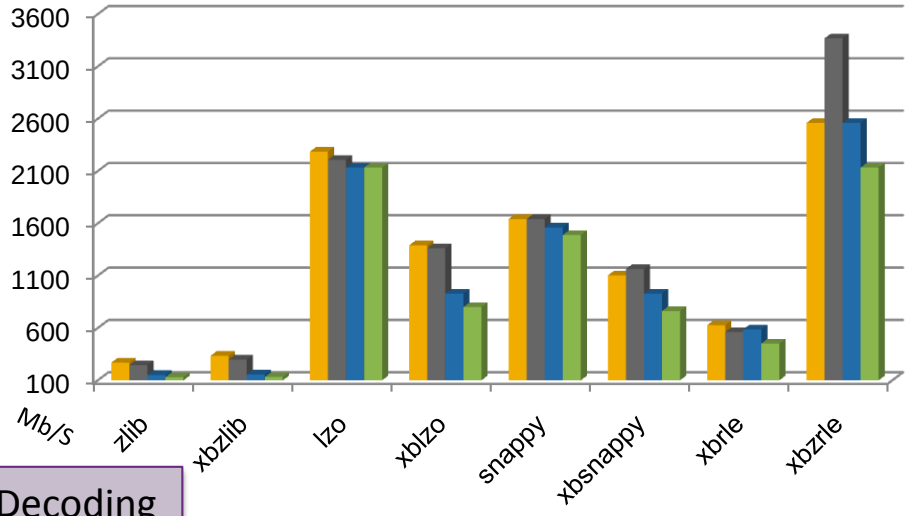
© 2011 SAP AG. All rights reserved.

Performance comparison

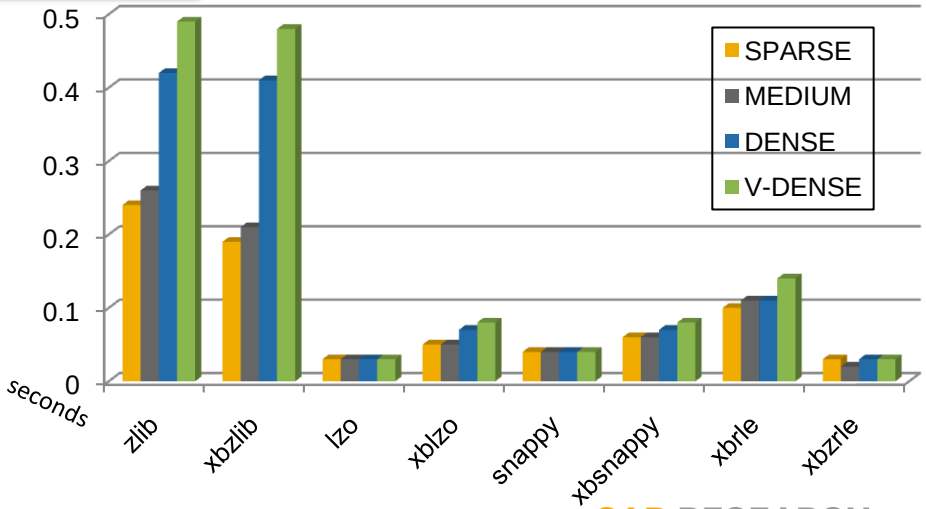
Synthetic benchmark representing enterprise workload



Encoding



Decoding

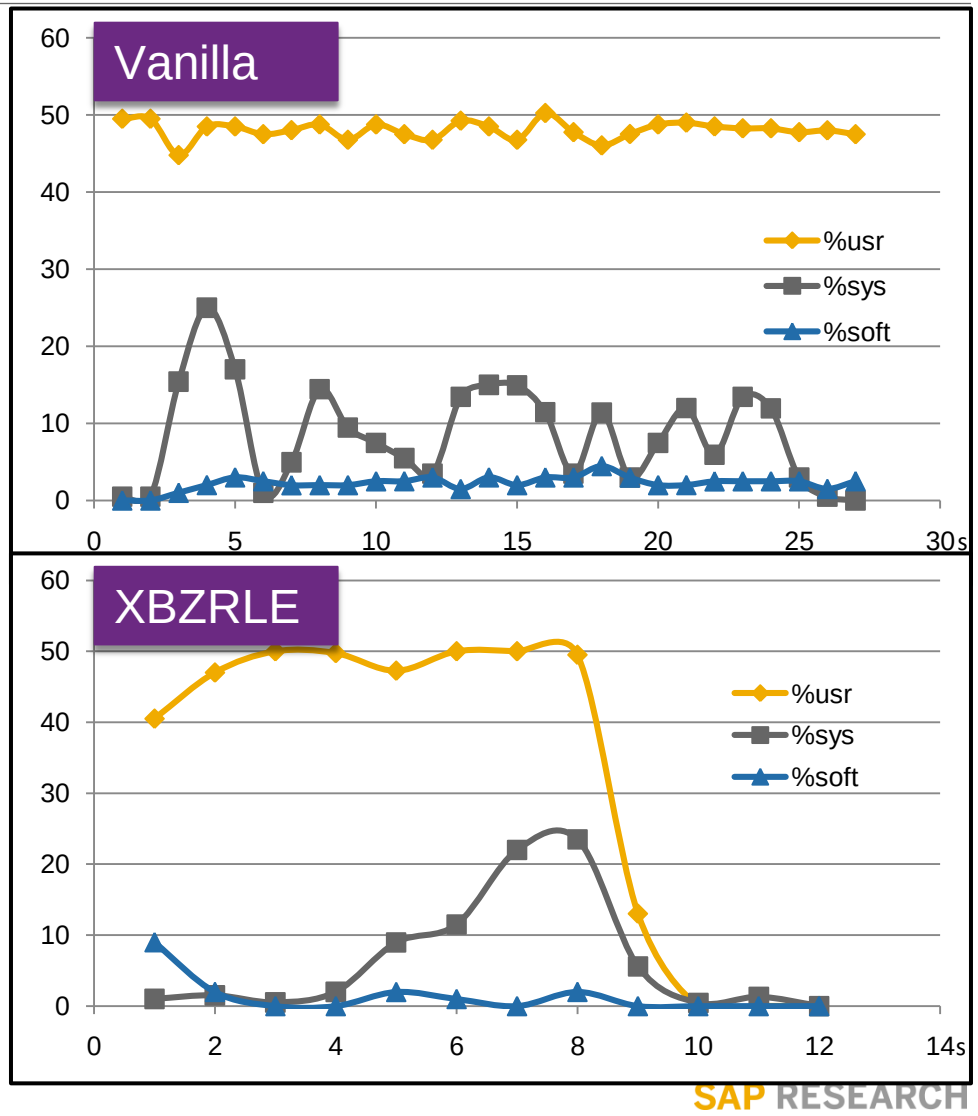


SAP RESEARCH

© 2011 SAP AG. All rights reserved.

Performance comparison
Live Migration Benchmark

- Compute capacity used for live migration :
  - **xbzrle** : 50%
  - **vanilla**: between 30%-60%
- Live Migration:
  - **xbzrle** : terminate in seconds
  - **Vanilla** :not able to complete in the allocated time



SAP RESEARCH

© 2011 SAP AG. All rights reserved.

Future Work

Future Works

- **Dynamically disable XBZRLE algorithm if the cache miss ratio is to important**
- **Combine Page priority algorithm and XBZRLE:**
  - Cache page with highest dirtying rate
  - Eliminate unnecessary cache check
  - Eliminate page compression with low potential return

SAP RESEARCH

© 2011 SAP AG. All rights reserved.

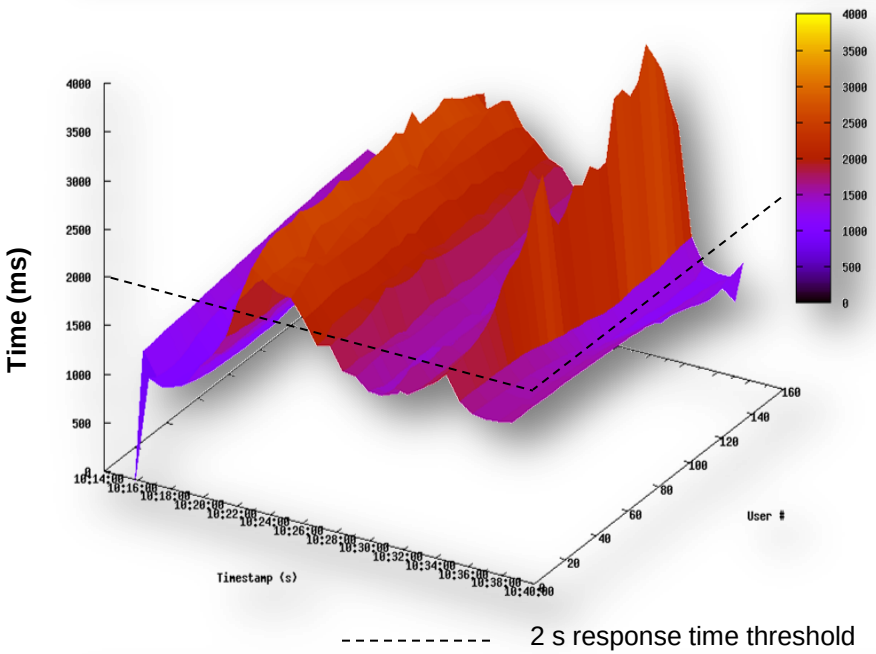
Thank You!

Contact information:

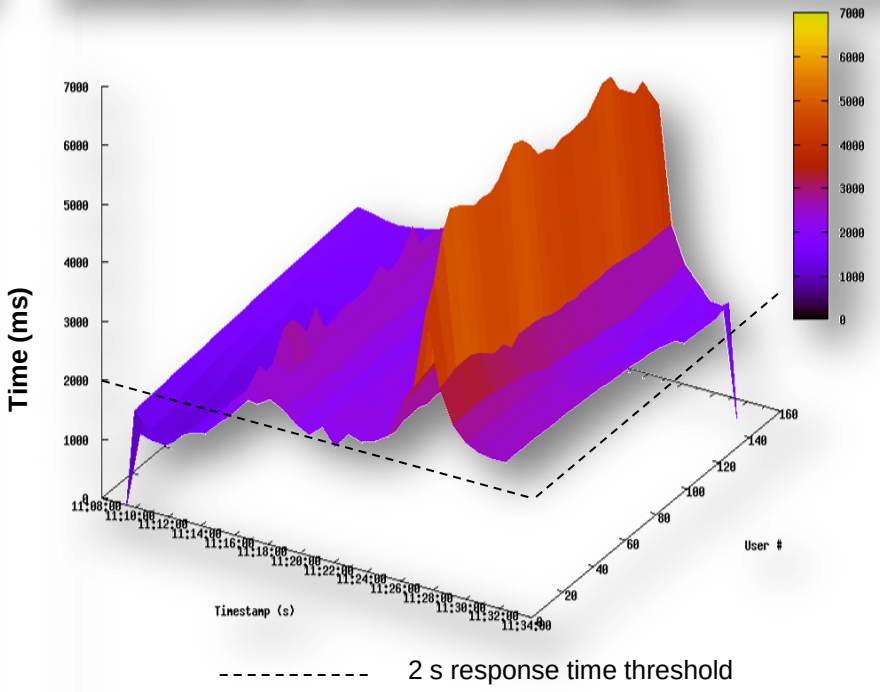
Dr. Benoit Hudzia
Senior Researcher
benoit.hudzia@sap.com

Experimentations Results: S&D Benchmark with/out warm-up

Response Time (baseline)



Response Time (warm-up)



VM size : 4GB

SMP : 2 vCPU

Users : 150

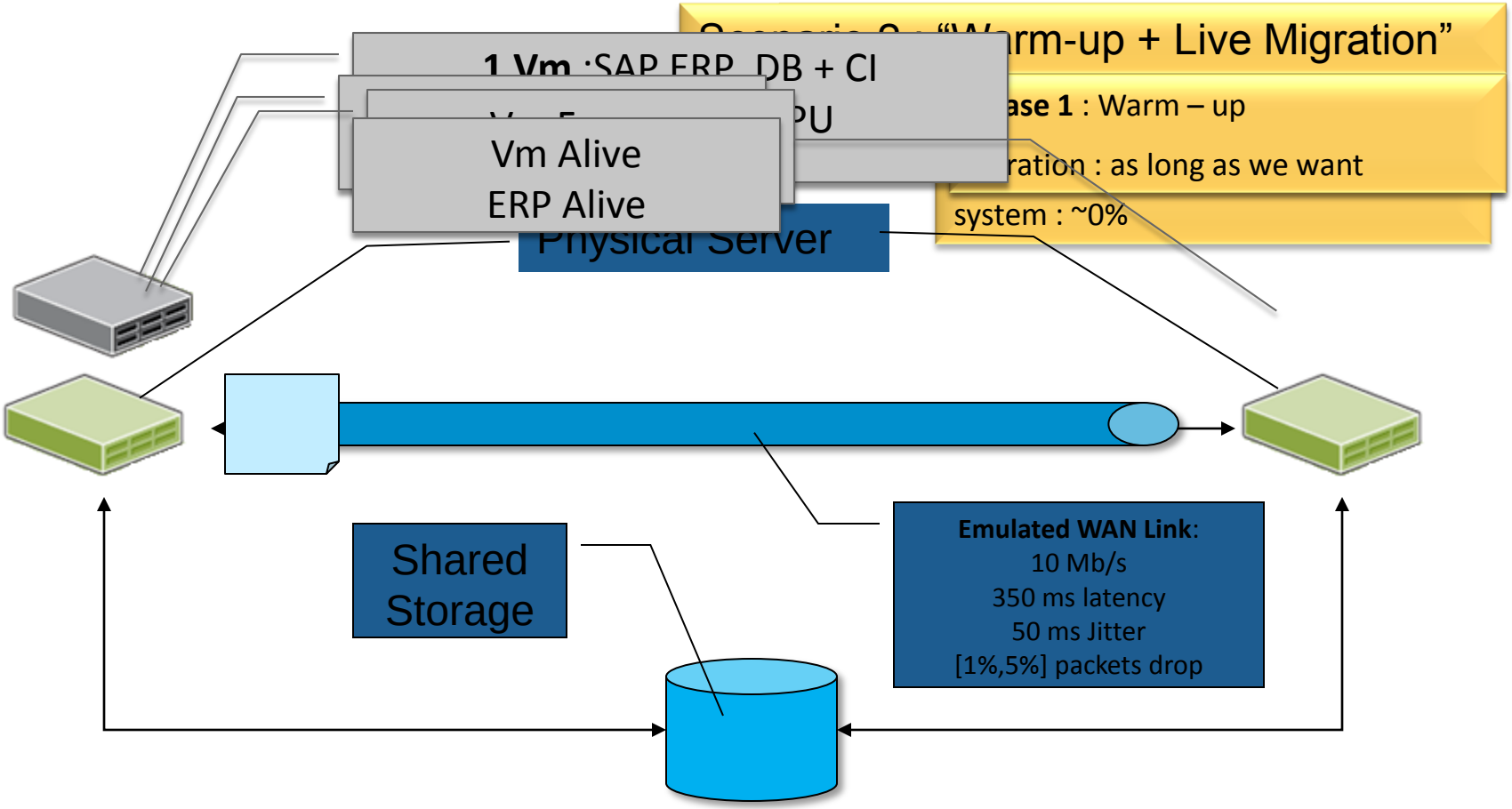
	CPU	Avg Response Time
Baseline	60%	2.18 sec
Warm-up	73%	2.16 sec

Downtime under load: <1 sec
Success ratio : ~99%

SAP RESEARCH

© 2011 SAP AG. All rights reserved.

Live Migration over emulated WAN Link



SAP RESEARCH

© 2011 SAP AG. All rights reserved.