



COLO: COarse-grain LOck-stepping Virtual Machine for Non-stop Service

Eddie Dong, Will Auld

Legal Disclaimer



INFORMATION IN THIS DOCUMENT IS PROVIDED IN CONNECTION WITH INTEL® PRODUCTS. NO LICENSE, EXPRESS OR IMPLIED, BY ESTOPPEL OR OTHERWISE, TO ANY INTELLECTUAL PROPERTY RIGHTS IS GRANTED BY THIS DOCUMENT. EXCEPT AS PROVIDED IN INTEL'S TERMS AND CONDITIONS OF SALE FOR SUCH PRODUCTS, INTEL ASSUMES NO LIABILITY WHATSOEVER, AND INTEL DISCLAIMS ANY EXPRESS OR IMPLIED WARRANTY, RELATING TO SALE AND/OR USE OF INTEL® PRODUCTS INCLUDING LIABILITY OR WARRANTIES RELATING TO FITNESS FOR A PARTICULAR PURPOSE, MERCHANTABILITY, OR INFRINGEMENT OF ANY PATENT, COPYRIGHT OR OTHER INTELLECTUAL PROPERTY RIGHT. INTEL PRODUCTS ARE NOT INTENDED FOR USE IN MEDICAL, LIFE SAVING, OR LIFE SUSTAINING APPLICATIONS.

Intel may make changes to specifications and product descriptions at any time, without notice.

All products, dates, and figures specified are preliminary based on current expectations, and are subject to change without notice.

Intel, processors, chipsets, and desktop boards may contain design defects or errors known as errata, which may cause the product to deviate from published specifications. Current characterized errata are available on request.

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

*Other names and brands may be claimed as the property of others.

Copyright © 2012 Intel Corporation.

Non-Stop Service with VM Replication

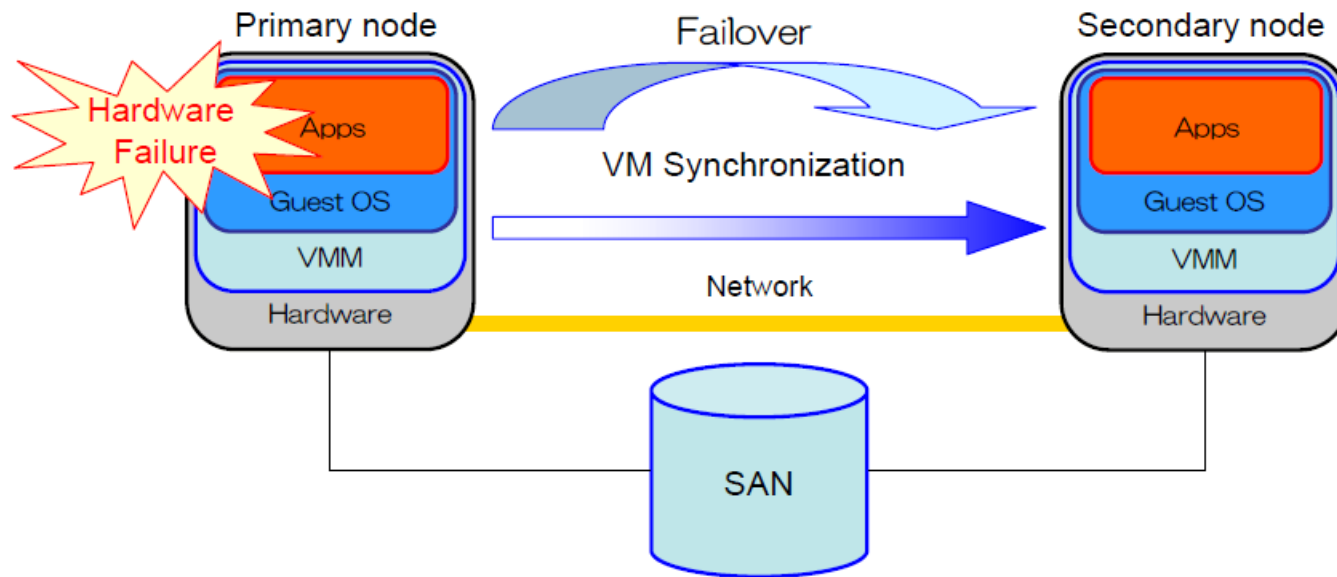


Client server

Xen state of the art:

Ramus - buffers checkpoints per epic

VM Replication



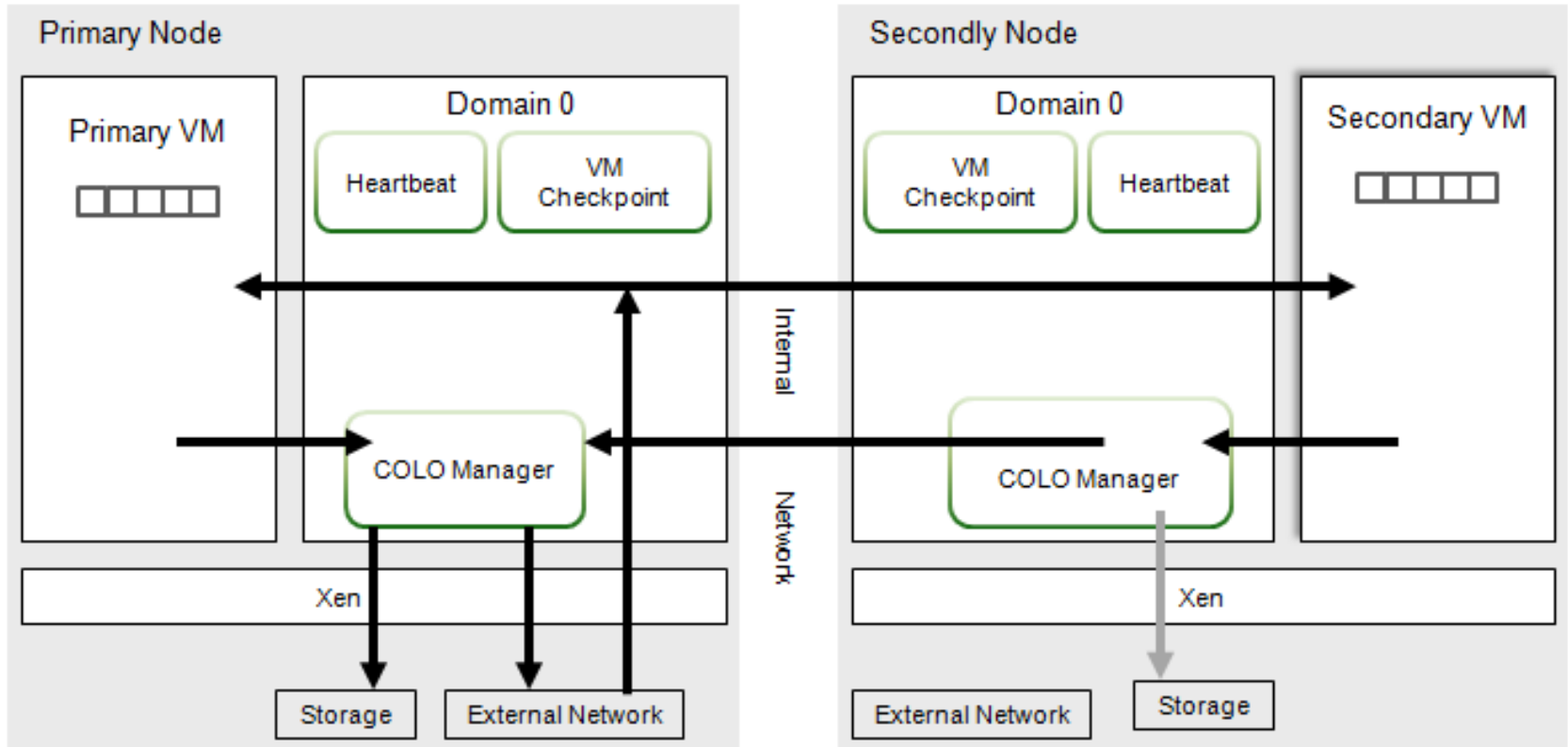
Why COarse-grain LOck-stepping (COLO)



VM Replication is an overly strong condition
Coarse-grain lock-stepping VMs
Secondary VM is a replica, as if it can generate
same response with primary so far

Non-stop service focus on server response, not internal
machine state!

Architecture of COLO



COarse-grain LOck-stepping Virtual Machine for Non-stop Service

Improving Response Similarity



Minor Modification to Guest TCP/IP Stack

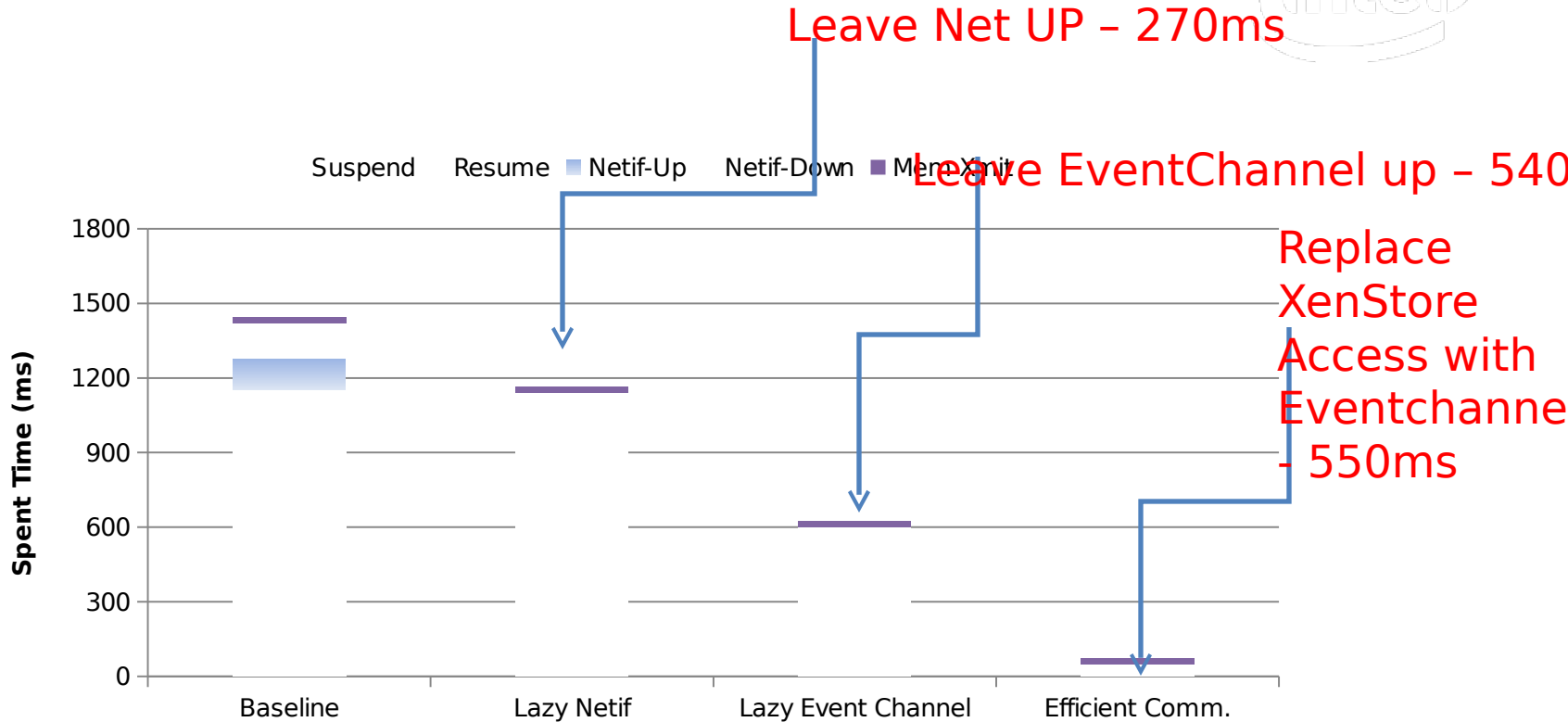
Coarse Grain Time Stamp

Highly-deterministic ACK mechanism

Coarse Grain Notification Window Size

Per-Connection Comparison

Checkpoint Cost with Optimizations



Final cost: 74ms/checkpoint: (1/3 on page transmission, 2/3 on suspend/resume)