Minimum-energy bandwidth management for QoS live migration of virtual machines

D. Amendola; N. Cordeschi; E. Baccarelli

danilo.amendola@uniroma1.it - Computer Networks - 2015,
poster presentation at ComSoc Summer School '16

Abstract

Live virtual machine (VM) migration aims at enabling the dynamic balanced use of the networking/computing physical resources of virtualized datacenters, so to lead to reduced energy consumption. However, the bandwidth consumption and latency of current state-of-the-art live VM migration techniques still reduce the experienced benefits to much less than their potential. Motivated by this consideration, in this paper, we analytically characterize, prototype in software and test through field trials the optimal bandwidth manager for intra-datacenter live migration of VMs. The goal is the minimization of the migration-induced communication energy under service level agreement (SLA)-induced hard constraints on the total migration time, downtime, slowdown of the migrating applications and overall available bandwidth. For this purpose, after recognizing that the resulting (nonconvex) optimization problem is an instance of Geometric Programming, we solve it by resorting to suitably developed adaptive version of the so-called primal-dual gradient-based iterations and, then, we analytically characterize its feasibility conditions.

Introduction and Architecture

Virtualization is an emerging technique that allows running multiple operating systems (OSs) simultaneously on a single server. For this purpose, a special middleware layer, the virtual machine manager (VMM) or hypervisor, abstracts from physical computing/networking resources and provides the so-called virtual machines (VMs), which act like real networked computers with their own virtual resources [?]. In modern virtualized networked datacenters (VNetDCs), live migration allows to move a continuously running VM from one server to another, so to attain multiple goals, including failure tolerance and energy-saving through server consolidation/load balancing [?]. Although live migration is becoming a service primitive function for the resource management of VNetDCs, it may induce slowdown of the application run by the migrating VM, as well as not negligible increments of the networking traffic and the computing-plus-networking energy consumption.

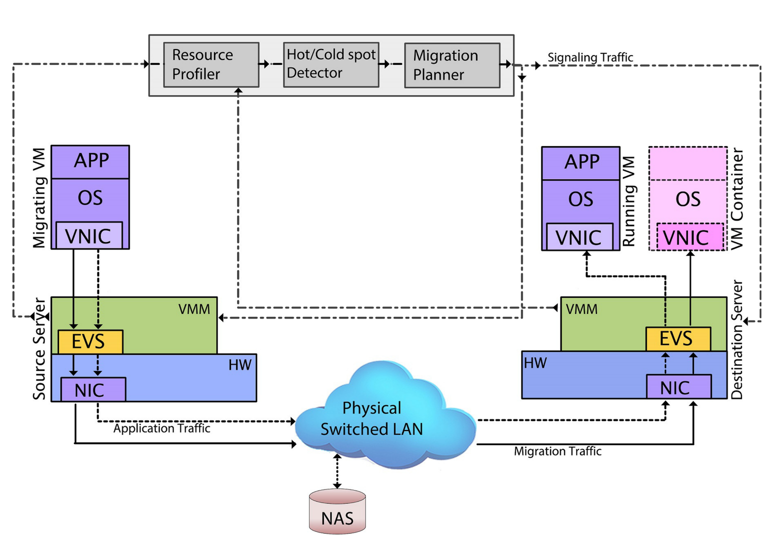


Fig. 1: Reference VNetDC architecture for intra-datacenter live VM migration.

Continue (resp., dotted) arrowed lines denote end-to-end TCP connections conveying migration (resp., application) traffic. Dashed-dotted arrowed lines denote signalling flows for implementing the resource profiling and migration plan. APP: application; VMM: virtual machine manager; OS: operating system; HW: networking/computing hardware; EVS: external virtual switch; VNIC: virtual network interface card; NIC: physical network interface card; NAS: network-attached storage.

Problem setup and optimal resource allocation

R (Mb/s) is the transmission rate.

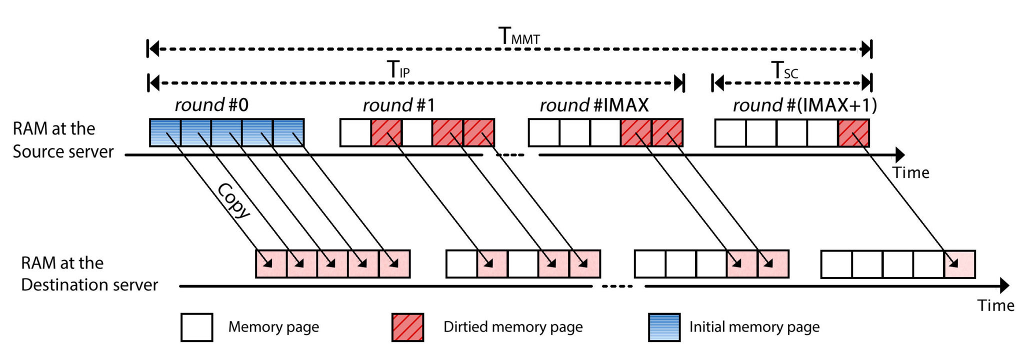


Fig. 2: Time-chart of the PeCM technique.

Memory Migration Time:

$$T_{MMT} \equiv T_{MMT}(R) \triangleq T_{IP}(R) + T_{SC}(R). \quad (1)$$

Total migration time T_{TOT} (s) is the overall duration:

$$T_{TOT} \triangleq T_{PM} + T_{RE} + T_{IP} + T_{SC} + T_{CM} + T_{AT}, \quad (2)$$

Downtime:

$$T_{DT} \triangleq T_{SC} + T_{CM} + T_{AT}, \quad (3)$$

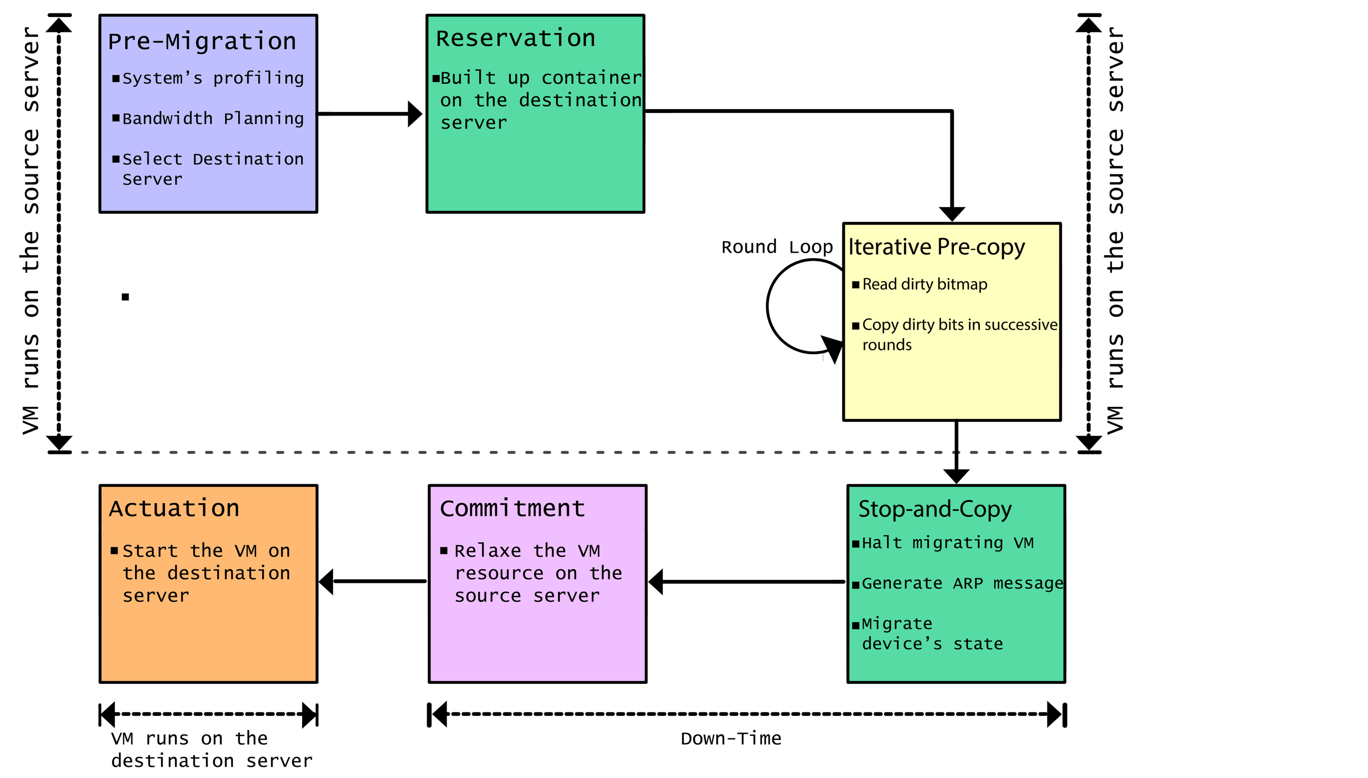
Table 1
Main taxonomy of the paper.

Symbol	Meaning/role
I_{MAX}	Number of migration pre−copy rounds
i	Round index, i=0,...,($I_{MAX}+1$)
$\overline{w}(Mb/s)$	Memory dirty rate of the migrated VM
$R\,(Mb/s)$	Migration bandwidth
$P(R)\,(W)$	Communication power at the migration bandwidth R
$\hat{R}\,(Mb/s)$	Maximum available migration bandwidth
$M_0\,(Mb)$	Memory size of the migrated VM
$\mathcal{E}_{TOT}\,(J)$	Total consumed communication energy
$\Delta_{MMT}\,(s)$	Maximum tolerated memory migration time
$\Delta_{SC}\,(s)$	Maximum tolerated stop−and−copy time
β	Migration speed−up factor
n	Integer-valued iteration index

BroadcomLab Research Group

Keyword: Mobile Cloud Computing; Vehicular Cloud Computing; Big Data; Green Data Center; Communication and Computing Engineering in Cloud and Big Data; LTE; Next Gen-

Pre-copy live VM migration



The PeCM technique involves six stages:

1. *Pre-migration:* resource profiling. This stage spans T_{PM} seconds;
2. *Reservation:* the computing/communication/storage/memory physical resources are reserved at the destination server;
3. *Iterative pre-copy:* this stage is composed by $(I_{MAX}+1)$ rounds and spans T_{IP} seconds. During the initial round, the entire memory content of the migrating VM is sent to the destination server. During the subsequent I_{MAX} rounds, the memory pages modified during the previous round are re-transferred to the destination server (see Fig. 2);
4. *Stop-and-copy:* the migrating VM is halted and a final memory-copy roun is performed (see Fig. 2). This last round spans T_{SC} seconds;
5. *Commitment:* the destination server notifies that it has received successfully. T_{CM} (s) is the duration of this stage;
6. *Re-activation:* the I/O resources and IP address are re-attached to the migrated VM on the destination server. T_{AT} (s) is the needed time.

Tracking capabilities under connection phenomena

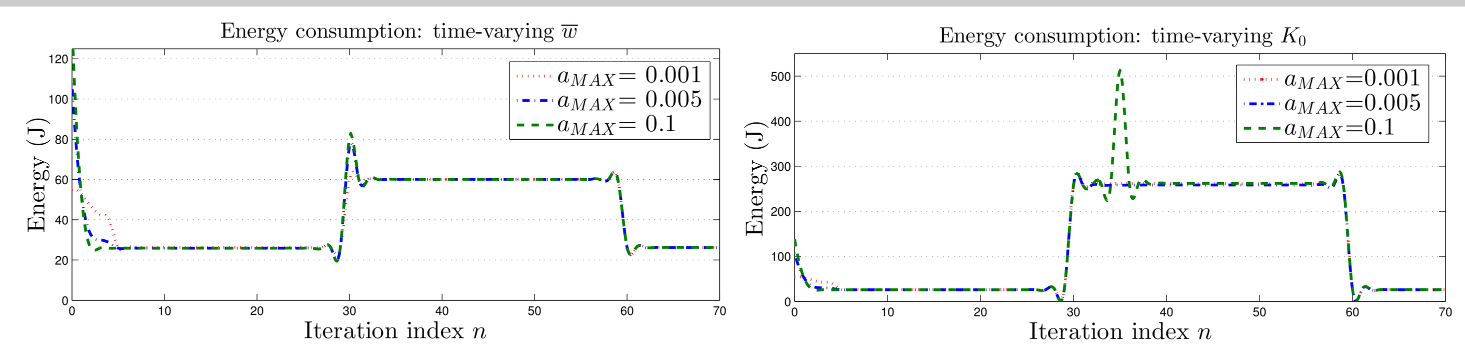


Fig. 6: Time evolutions (in the n index) of the energy consumption of the proposed bandwidth manager at: $\hat{R} = 900$ (Mb/s), $M_0 = 512$ (Mb), $\beta = 1.15$, $\Delta_{MMT} = 13.5$ (s), $\Delta_{SC} = 0.6$ (s), $\widetilde{I}_{MAX} = 3$, and $\gamma = 100$ for the application scenario of Section ??. (a) Case of time-varying $\overline{w}$; (b) Case of time-varying K_0.

Imax formula, validation results

We claim that the optimized setting: $\widetilde{I}_{MAX}$ of I_{MAX} is obtained by computing the value with the equality:

$$\widetilde{I}_{MAX} \equiv \left\lceil \frac{\log(M_0/\Delta_{SC}\hat{R})}{\log(\hat{R}/\overline{w})} - 1 \right\rceil, \text{ for } (\hat{R}/\overline{w}) > 1, \quad (4)$$

We validated this formula with estensive simulations, here we present some results:

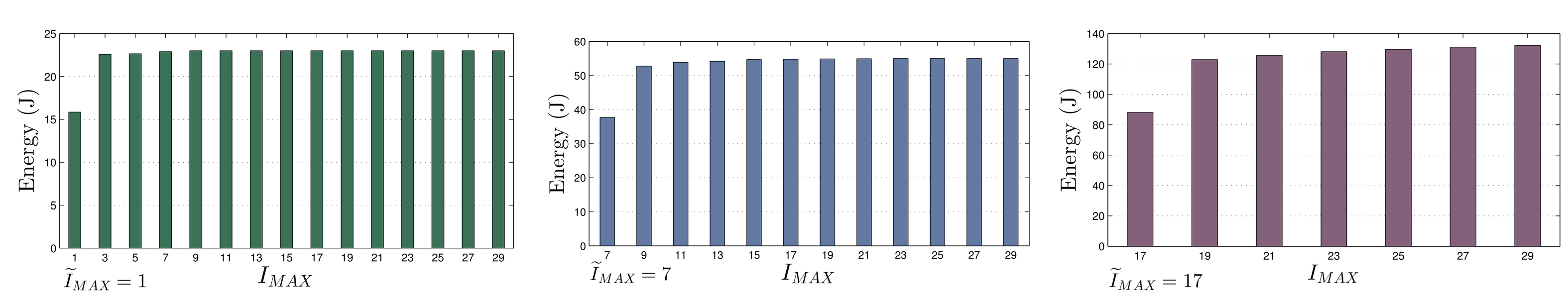


Fig. 7: Time evolutions (in the n index) of the energy consumption of the proposed bandwidth manager at: $\hat{R} = 900$ (Mb/s), $M_0 = 512$ (Mb), $\beta = 1.15$, $\Delta_{MMT} = 13.5$ (s), $\Delta_{SC} = 0.6$ (s), $\widetilde{I}_{MAX} = 3$, and $\gamma = 100$ for the application scenario of Section ??. (a) Case of time-varying $\overline{w}$; (b) Case of time-varying K_0.

Conclusion We developed the optimal bandwidth manager for intra-datacenter live VM migration. It minimizes at run-time the communication energy wasted by the migration of the VM memory under hard QoS constraints on both the migration time and downtime. After implementing it atop a wired test-bed, we measured and compared its energy performance by considering synthetic and real-world workloads, as well as random and ordered migration scheduling disciplines. The carried out field trials highlight that the average energy saving of the proposed bandwidth manager over the corresponding state-of-the-art Xen one is over 40% and approaches 66% under strict constraints on the tolerated downtimes. Interestingly, the measured per-migration CPU slow-down induced by its implementation is, in average, limited up to 1.5−2%, while the measured average stretching of the execution times of the migrated applications is under 20%.