

**CHENNAI – PONDICHERRY**

**Shadow-Routing Based Dynamic Algorithms for Virtual Machine Placement in a Network Cloud**

**Abstract:**

We consider a shadow routing based approach to the problem of real-time adaptive placement of virtual machines (VM) in large data centers (DC) within a network cloud. Such placement in particular has to respect vector packing constraints on the allocation of VMs to host physical machines (PM) within a DC, because each PM can potentially serve multiple VMs simultaneously. Shadow routing is attractive in that it allows a large variety of system objectives and/or constraints to be treated within a common framework (as long as the underlying optimization problem is convex). Perhaps even more attractivefeature is that the corresponding algorithm is very simple to implement, it runs continuously, and adapts automatically to changes in the VM demand rates, changes in system parameters, etc., without the need to re-solve the underlying optimization problem “from scratch”. In this paper we focus on the minmax- DC-load problem. Namely, we propose a combined VM-to- DC routing and VM-to-PM assignment algorithm, referred to as Shadow scheme, which minimizes the maximum of appropriately defined DC utilizations. We prove that the Shadow scheme is asymptotically optimal (as one of its parameters goes to 0). Simulation confirms good performance and high adaptivity of the algorithm. Favorable performance is also demonstrated in comparison with a baseline algorithm based on Vmware implementation. We also propose a simplified – “more distributed” – version of the Shadow scheme, which performs almost as well in simulations.

**Existing System:**

The VM placement problem to multiple data centers (DCs) or server clusters. We consider multiple types of Virtual machines (VMs) and physical machines (PMs), where a VM type is characterized by the vector of the required resource amounts; and a PM type is characterized by the vector of the resource amounts it possesses. Different DCs may consist of different PM types.

Not all resources are associated with individual PMs – some of them (e.g. disk storage) exist as one resource pool associated with a DC. The placement problem consists of two parts (layers): (i) routing layer decides which DC a given VM should be routed to; and (ii) DC layer assigns an “arriving” VM to a specific PM within the DC – this assignment, of course, has to respect VM-to-PM “packing constraints”, i.e. the total resource requirements of all VMs assigned to a PM cannot exceed the resource amounts at the PM (see (1) below). The latter constraints are sometimes called vector packing (see e.g. [5]). Packing constraints substantially complicate the VM placement problem, and require a sufficiently intelligent strategy to efficiently utilize the physical resources.

**Proposed System:**

We propose a shadow routing based placement algorithm in this paper. This means that, roughly speaking, each arriving VM first “arrives” into a specially designed virtual queueing system, where it is routed to one of the (virtual) queues. The “service” in the virtual system is performed by a “superserver”, whose “service rate vectors” are feasible “packing configurations” of VMs into PMs. The advantage of shadow routing approach is that it is simple and adaptive: no need to know a priori, or explicitly measure, the VM arrival rates; if the arrival rates change, the algorithm adapt automatically. Yet, as we show, the algorithm is asymptotically optimal. (The virtual system in our case is an instance of a general model of and therefore our algorithm definition and its optimality are derived from the results. All these features are confirmed by our simulations, where we compare our algorithm, referred to as Shadow scheme, to a baseline scheme, which is along the lines of some of the currently used algorithms.

.