

Enabling Location Based Services in Data Centers via Wireless USB Radios

K. Kant, N. Udar, R. Viswanathan
 krishna.kant@intel.com, udarn@siu.edu, viswa@enr.siu.edu

Abstract—In this paper, we show how the wireless USB (which is based on UWB technology) can enable rich location based services in data centers. The focus of these services is to enable a more effective automated resource management based on precise location of various assets (servers, switches, etc.) in a data center.

Key words: Location estimation, Wireless LAN, Ultrawide-band, RFID, Management services, WSMAN, SML.

I. INTRODUCTION

Major data centers routinely sport several tens of thousands of “assets” (servers, switches, storage bricks, etc.) that usually go into standard slots in a rack or a chassis that fits the rack. Fig. 1 shows a typical row of a data center with the popular “rack mount” assets. These assets may move around for a variety of reasons including replacement, SW patching, manual reorganization, etc. The ease with which assets can be inserted into or removed from the slots makes asset tracking a substantial problem in large data centers and some tracking solutions are beginning to emerge.

Wireless LAN (WLAN) based localization has been extensively explored in the literature [1] and can be implemented easily in software. The root mean square error (RSME) in the best LoS case is 4.1m. The multipath decomposition method in [2] can cut this down to 1.1 meters, which is still very crude to accurately identify even the racks themselves. Ultrasonic or surface acoustic wave (SAW) systems perform localization based on time of flight (TOF) of sound waves. Because of very low speed of sound, SAW systems can measure distance with an accuracy of few centimeters. Unfortunately, SAW systems require substantial infrastructure and uninterrupted sound channels between emitter and receivers [3]. Other potential localization solutions include RFID, but it requires substantial infrastructure and yet rather poor accuracy. HP has developed a solution based on an array of passive RFID tags attached to each server [4]. This is a rather complex system requiring motorized track. In [5], [6], we have explored a localization technique that exploits the ultra wide band (UWB) PHY used by the WUSB solution that is intended to replace a wired USB. Owing to its high bandwidth, UWB provides a much higher resolution than say, WLAN [3]; however, it still falls short of the requirements. In [6], the accuracy is improved by exploiting the regular geometric structure of data centers.

In this paper, a variety of applications of location based services (LBS) in data centers are introduced. Our focus here is not on arming humans with useful information as in traditional LBS, but to allow the middleware to do a better

job of resource management. As a simple example, each rack in a data center has certain capacity for power circuits which cannot be exceeded. Therefore, if the middleware knows the rack membership of servers, it can abide by the rack power limitation when provisioning/migrating applications. However, in a large data center, we need more than just locations – an efficient mechanism to exchange location and other attributes (e.g., server load) are also required. This will enable good provisioning/migration decisions. This is where application of LBS services become vital. We envision the middleware to still be making the final selection of servers based on the appropriate policies; the function of LBS is merely to provide a (hopefully small) set of potential servers to choose from.

This paper makes two significant contributions. First, it introduces the applications of LBS within a data center context. Second, it proposes a scalable architecture based on Wimedia to manage a large number of WUSB servers in a data center. Section II discusses applications of LBS in data centers and IT environments in a bit more detail. In order to illustrate the usefulness of the concepts, this section also shows how better power management can be achieved with LBS. Section III proposes a scalable LBS architecture for large data centers. Finally, section IV concludes the paper.



Fig 1. Snapshot of row in a data center

II. LOCATION BASED SERVICES IN A DATA CENTER

In this section, we first discuss some important data center aspects that are relevant from a LBS perspective. We then discuss some specific examples of LBS for addressing these and other operational issues.

A. Virtualization and Security

Data centers show perennially low average server utilization (5-10% range) and yet ever increasing server count, power consumption, and associated infrastructure and management

costs. The low utilization is attributed not only to unpredictable demands but more importantly to the need for isolation among various applications and activities. Virtualization has recently gained acceptance as a way to increase resource utilization in data centers while still maintaining a level of isolation between various applications and activities. Aggressive virtualization leads to the notion of “utility computing” whereby the entire data center can be viewed simply as a pool of resources (computes, storage, special functions, etc.) which are virtualized. These resource can be allocated dynamically to applications based on the current needs. Virtualization can be viewed as a mechanism to make the physical location of resources irrelevant since any resource can be assigned to any application in this model. While this flexibility brings in several advantages, a location blind resource allocation can lead to anomalies, poor performance and suboptimal resource usage. In other words, a location aware resource management can retain all the advantages of virtualized data center while avoiding its pitfalls.

In order to enforce isolation in a virtualized environment, each application needs to execute on its own “virtual cluster”, defined as a set of virtual machines (or virtual nodes) connected via QoS controlled virtual links[7]. In this way, a problem with one virtual cluster (e.g., virtual machine crash or congestion on virtual pathways) should not affect other applications unless the problem lies in the hypervisor. However, if many virtual clusters were to be mapped to the same set of physical resources (e.g, servers, switches, routers, etc.), it may be difficult to maintain the isolation properties. Furthermore, as data centers increase in size, they become increasingly attractive targets of attacks via viruses, worms, focused traffic (distributed denial of service attacks), etc. At the same time, the evolution of data centers from closed entities toward outsourced utility computing farms not only makes the attacks easier but also virtually eliminates the distinction between “insiders” and “outsiders”. Confining a virtual cluster to a physical region offers advantages in terms of easier containment of attacks and in minimizing the number of applications that are affected by an attack. In this context, the relevant “physical region” is really “network region”, e.g., set of servers served by one or more switches or routers; however, the structure of data center forces a mapping between “network” and “physical” neighborhoods. For example, all blade servers in a chassis share a switch, and all chassis switches are connect to the rack level switch. Thus the location based provisioning and migration can benefit from security/isolation perspective. For essentially the same reasons, a location aware allocation can yield better performance for latency sensitive applications since the reduction in number of switches on the communication paths also reduces the communication latency.

B. Power and Thermal issues

The continued increase in processing power and reduction in physical size has increased power densities in data centers to such an extent that both the power-in (i.e., power drawn) and power-out (i.e., power dissipated as heat) have become serious

problems. For example, most racks in today’s data centers were designed for a maximum of 7 KWHr consumption, but the actual consumption of a fully loaded rack can easily exceed 21 KWHr. As a result, in older data centers, racks are often sparsely populated lest the power circuit capacity exceed resulting in a brownout. In addition, the power and cooling costs are becoming substantial percentage of overall costs. Consequently, an intelligent control over both power consumption and cooling becomes essential. Power/thermal issues are inherently tied to the location of the active assets. For example, cooling can be made more effective and cheaper if the servers with high thermal dissipation are not bunched up [8].

C. Location Based Services

In the following we briefly list several useful services for addressing performance, security, power/thermal, fault location, etc. Needless to say, the list is neither exhaustive nor meant to be prescriptive.

- 1) Inventory Management, i.e., keeping track of which server is inserted where and which ones are “missing” (not plugged in anywhere).
- 2) Logical grouping of assets based on their location in order to simplify allocation, deallocation, migration, etc.
- 3) Power/thermal aware distribution of queries to various servers for e-commerce and database applications.
- 4) Asset specific trouble ticket management, i.e., identify the asset that needs replacement, fixing, SW patching, etc. If the asset itself does not display any trouble lights, the localization must be extremely precise.
- 5) Automatic binding of a mobile device to a specific server where the server is selected by simply placing the device closest to it (and within some maximum absolute distance). This binding may be done to upload or download data to a specific server.
- 6) Quick quarantine of all servers belonging to the same enclosure as the server that detects a DoS or virus attack.
- 7) Automated adjustment of air-flow direction flaps from shared fans in order to maximize cooling of hot spots. This situation is generally applicable to blade chassis which have shared fans. (Racks usually don’t).
- 8) Fan noise balancing within a rack, e.g, if a highly loaded server is running its fan at maximum speed, we may want to take appropriate action to allow fans on adjacent servers to be run slow.

D. Power/Thermal Balance Using LBS

In this section we show that LBS can be used effectively to handle the issues of power-load balance in a data center. Consider a data center having a single row with 2 racks. Each rack has 12 slots and is partially filled with 8 servers. Let us assume that all servers have identical computing capabilities and each rack has maximum power-in of 650 W. The servers in each rack are logically grouped using a rack level switch. Consider, running an application in a data center which demands 320% CPU utilization. We analyze running this application in three different scenarios. (1) No Localization,

i.e., the server locations are not known, (2) We know the rack membership of the server but the exact location in the rack is not known, and (3) The exact location of the server in the rack is known.

It is well known that the power consumption P relates to the CPU utilization U by a non linear relationship. In [9], the authors performed detailed measurements on streaming media servers with several configurations to study the relation between CPU utilization and the power consumption, and obtained the empirical equation $P = D + (M - D)U^{0.5}$ where

D : Power consumed in the idle mode.

M : Power consumed during maximum load, and

U : CPU utilization of the server.

As reported in [9], the CPU in the minimum power save mode consumes 35 W, 69 W in the idle mode and 145 W during the maximum load. We use this data to illustrate the significance of LBS for power-load balance.

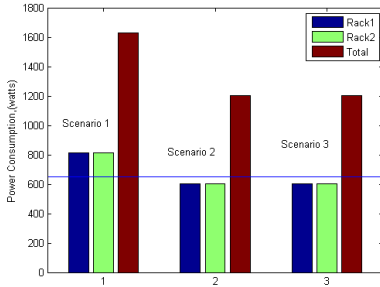


Fig 2. Power-Load Balance in a Data Center

In Scenario1, a conservative approach is to distribute the load equally on all the available servers. Each of the 16 servers in this case share a load of 20% to meet the total load demand of 320%. With equal load sharing, each rack exceeds the maximum power-in for a rack as shown in Fig. 2. In Scenario2, it is known which servers belong to which of the two racks. Therefore, the total load is divided equally between the two racks. Further, within each rack, 4 out of 8 servers share the 40% load and the remaining servers are in the minimum power save mode. The non uniform distribution of load among the available servers leads to power saving as shown in Fig. 2 and also meets the maximum power-in requirement of a rack. Similar power saving is observed in Scenario3; however, in this case, the knowledge of exact location of the server in a rack does not yield any additional power savings. In order to increase power savings, we can try to load 2 servers in each rack maximally, say 80% capacity and leave the other 6 in the minimum power save mode. However, the high heat dissipation in this case would likely cause problems if the two used servers were located next to each other. This is where their precise location can be exploited to balance the thermals as well and allow high server utilizations.

III. SCALABLE ARCHITECTURE FOR DATA A CENTER

It is clear from the above discussion that location based services can form a vital part of effective resource management. The algorithms used for high accuracy localization of WUSB servers in a rack are analyzed in [6] and are not within the scope of this paper. Here, we assume that the each WUSB server is able to determine its position in a rack with high accuracy. However, knowing one's position alone is inadequate to provide LBS. Instead, we need a flexible and scalable

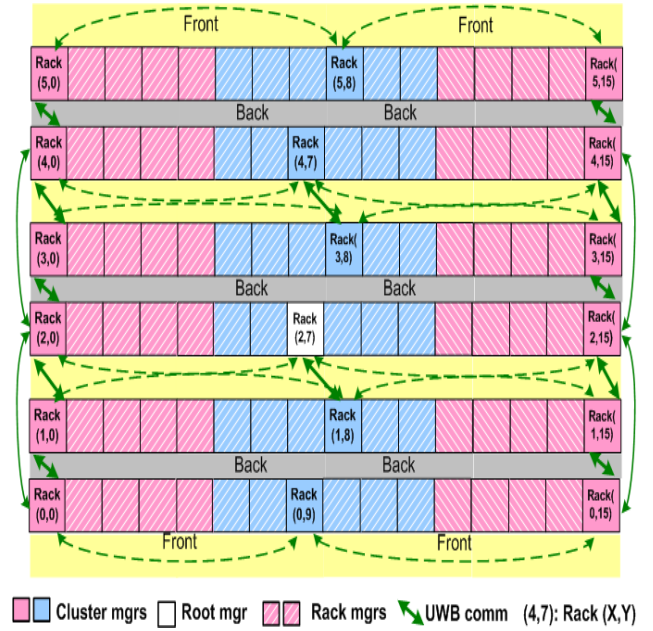


Fig 3. Clustering architecture for 6×16 data center

architecture to exchange location and other relevant data (e.g., power/thermals) so that we can quickly identify the suitable set of servers for allocating new applications or migrating existing ones. We also would like to avoid the need for any additional infrastructure for such communications (e.g., wired communication between certain “leader” nodes), since such an approach will defeat our basic premise of keeping the localization solution self-contained [5].

In order to build a data center wide network of WUSB radios, we surely require a hierarchical structure just to cope with the magnitude of the problem. Furthermore, distributed MAC protocols provide greater flexibility in connecting nodes that are farther apart. Here, we propose a “Cluster-Tree” approach which can be viewed as a special case of peer-to-peer network with a central root node. Cluster tree approach uses MAC protocol based on Wimedia standard. The salient features of this approach are:

- The communication MAC protocol is a domain dependent MAC with a master-slave (Piconet) architecture involving a Piconet controller (PNC) and up to 255 terminals (slaves).
- The PNC maintains global timing using a super frame (SF) structure. The SF consists of 256 slots and each slot has duration of $256\mu s$.
- A SF consists of a beacon period (BP), contention access period (CAP), and contention free period (CFP). The BP is used for PNC to terminal broadcasts, CAP is used by the terminals to communicate with others or to ask PNC for reserved channel time, and CFP is dedicated for individual transmissions over agreed upon time slots.
- WUSB radios based on DMAC protocol operate in 4 band groups of UWB spectrum. In each band group 7 channels are supported.

In the proposed cluster tree approach, four types of nodes are introduced. Root manger (RM) (of which there is only one), cluster manger(CM), rack manager (RkM), and individ-

ual server node (SN). RM is the root node of the tree and is assumed to be connected to the central management console by a wired connection. The CM manages a cluster of racks in physical proximity. The need for the CM level and the number of CM's depends on the size of the data center that we are considering – for very small data centers, this level may be absent altogether. The CM may be connected to another CM to extend the range of the RM. Each rack manager (RkM) controls all the WUSB servers in a rack (henceforth known as end nodes).

To illustrate the organization of cluster tree, let us consider a 6×16 grid of racks shown in Fig. 3. The structure shown is by no means arbitrary; it is mostly dictated by the many constraints that we need to abide by. First, we want a solution involving *only* WUSB radios irrespective of the size of the grid (i.e., no external infracture). Second, data centers typically have alternating rows of front and back facing servers. Assuming that the WUSB radio antennas are located on the front side of the servers, communication across the “back” aisles in Fig. 3 is very challenging due to heavily metallic nature of racks. In particular, with allowed power levels, communication is possible only around the edges as shown. (The green arcs in Fig. 3 all possible UWB communications.) One consequence of edge based communications is that it forces us to define rack clusters as successive racks in a single row, as shown. Finally, the number clients a PNC can talk to is limited by the 256 distinct time slots in a frame, as discussed below.

In Fig. 3, the RM is located in the middle row and can handle 7 CMs simultaneously operating on seven different channels. With frequency reuse more CMs can operate simultaneously. Thus, the cluster tree can easily be scaled to a large network. Since each CM operates on a different channel, the SF time is divided among RkM's connected to each CM. Only 6 RkM's are allowed per CM. The maximum limit of 6 rack nodes is based on the assumption that each rack has 1U rack mount servers and the rack is completely filled. Thus, there are 42 WUSB servers per rack and given that 255 slaves are allowed in the Piconet, a CM can manage at most 6 rack nodes. The CM located in the end of each row acts as router to the CM in the next row.

In order to estimate the time required for a RM to reach every CM, Fig. 3 is redrawn as a cluster tree in Fig. 4. Due to the symmetry of arrangement of rows the cluster tree in Fig 4 considers only rows 2-5. The RM is located in rack(2,7) and has 5 CMs at the first level. The CMs of row 2 act as routers to the CMs of rows 4 and 5 as shown in the level 2 of the tree. At level 3 CMs of a row 4 act as routers to the CMs of row 4 and 5. At level 4, CM in rack (5,0) acts as router to the CM (5,8).

It is clear from Fig. 4 that the minimum height of the tree is 1 and the maximum height is 4. Since the communication between a root manager and cluster managers (along with rack and end nodes connected to the cluster manager is frame synchronized, the minimum time to reach any end node in a 6×16 data center is one frame or 65.536 ms, the maximum time is 4 frames or 0.262 sec. More generally, in a $m \times n$ data center, the maximum height of the tree is $m/2 + \lceil n/6 \rceil - 2$.

It is clear from this discussion that an end-to-end accu-

mulation of location and other attributes (e.g., load, power consumption, etc.) in a large data center could take several seconds. In most circumstances, this is perhaps not a problem since one would expect that most location based decisions could be made based on local information, and only rarely we need to go clear across the data center to accumulate all the needed information. Nevertheless, the analysis here does point to a potential problem in a completely UWB based LBS environment. An alternative is to assume a wired connection between every CM and the central management console that provides a quick short-cut for inter-cluster communications.

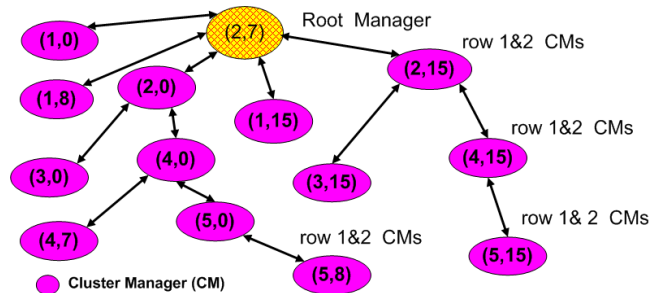


Fig 4. Cluster Tree representation of Fig.3

IV. CONCLUSIONS AND FUTURE WORK

In this paper, we explored services and capabilities that can be enabled by the localization of various “assets” in a data center. The scalable architecture for a large data center was also presented. The future work on the subject includes more careful examination of location based data collection and usage in order to cope with rather large end-to-end delays in a large data center. The other important aspect is to actually build a prototype system along with a few location based services and examine how well the proposed architecture works in practice.

REFERENCES

- [1] Xiang, Z., Song, S., Chan J., Wang, H, et.al, “A Wireless LAN-based Indoor Positioning Technology”, IBM Journal of R&D, Sept-Nov 2004.
- [2] S.A Golden, S.S Bateman , “Sensor Measurements for Wi-Fi Location with Emphasis on Time-of-Arrival Ranging”, IEEE Trans on Mobile Computing, Vol 99, pp1-36, 2007.
- [3] H. Liu, H. Darabi, P. Banerjee, “Survey of Wireless Indoor Positioning Techniques and Systems”, *IEEE Trans. on Systems, Man and Cybernetics*, Part-C, vol. 37, No.6, Nov 2007
- [4] “HP Creates RFID Technology for Tracking Data Center Assets”, PALO ALTO, Calif., Oct. 17, 2006
- [5] N. Udar, K. Kant, R. Viswanathan, D. Cheung, “Characterization of Ultra Wide Band Communications in Data Center Environments”, Proc. of 2007 ICUWB, Singapore, Sept 2007.
- [6] N. Udar, K. Kant, R. Viswanathan, “Asset Localization in Data Center using WUSB Radios”, to appear in IFIP Networking 08.
- [7] K. Kant, “Virtual Link: An Enabler of Enterprise Utility Computing”, Proc of ISPA, Sorrento, Italy, Dec 2006.
- [8] Sharma, R.K., Bash, C.E., et.al, “Balance of Power Dynamic Thermal Management for Internet Data Centers”, *IEEE Internet Computing*, Feb. 2005, pp.42-29.
- [9] C-H Lien, Y-W Bai, M-B Lin, “Estimation by Software for the Power Consumption of Streaming-Media Servers”, IEEE trans on instrumentation & measurement, Vol 56, No 5, Oct. 2007, pp1859-1870