

Towards Database Virtualization for Database as a Service

Aaron J. Elmore[‡]

Carlo Curino[§]

Divyakant Agrawal[‡]

Amr El Abbadi[‡]

[‡]University of California, Santa Barbara
Santa Barbara, CA, USA
{aelmore, agrawal, amr}@cs.ucsb.edu

[§]Microsoft CISL
Mountain View, CA, USA
ccurino@microsoft.com

ABSTRACT

Advances in operating system and storage-level virtualization technologies have enabled the effective consolidation of heterogeneous applications in a shared cloud infrastructure. Novel research challenges arising from this new shared environment include load balancing, workload estimation, resource isolation, machine replication, live migration, and an emergent need of automation to handle large scale operations with minimal manual intervention. Given that databases are at the core of most applications that are deployed in the cloud, database management systems (DBMSs) represent a very important technology component that needs to be virtualized in order to realize the benefits of virtualization from automatic management of data-intensive applications in large scale data-centers. The goal of this tutorial is to survey the techniques used in providing elasticity in virtual machine systems, shared storage systems, and survey database research on multitenant architectures and elasticity primitives. This foundation of core Database as a Service advances, together with a primer of important related topics in OS and storage-level virtualization, are central for anyone that wants to operate in this area of research.

1. INTRODUCTION

Modern enterprises are often faced with managing a large number of databases. Consolidating databases into a unified service can reduce the high costs associated with running dedicated servers and maintenance by multiple database administrators. This is referred to as Database as a Service (*DaaS*). Traditional DBMS technology has focused on fully utilizing a single server to meet the needs of a specific application. Building a Database as a Service infrastructure requires several key challenges to be addressed to transform traditional DBMSs into a scalable, elastic, and autonomic database platform [2]. The scale of applications hosted within a database platform calls for solutions to self-manage resource allocation, load-balancing, replication and other issues not supported by commercial DBMS vendors.

Several technological advances have been produced by the systems and storage communities to face similar virtualization and

consolidation needs. Virtualization technology has been a key enabler in building consolidated infrastructure services, by allowing physical hardware to be multiplexed into many virtual machines (*VMs*). Consolidated services, such as Amazon's public compute cloud EC2, use virtualization to share resources of large commodity servers between many smaller compute instances, which often can operate without dedicated resources. In virtualized environments, a *shared storage abstraction* can be used to separate the physical storage of virtual disks from the virtual machine. Shared storage enables mobility of VMs by requiring only a machine's memory state be migrated between hosts, as the larger persistent image is stored and mounted from an external network service. For virtualized computers and storage, the systems must respond to over-utilized servers by using automated resource isolation and workload placement to ensure applications receive adequate resources. A large part of success in cloud computing is driven by such advances in virtualization technology that have resulted in huge economies of scale in large-scale datacenter operations.

The database research community and key industrial players have embraced the idea of virtualization, and contributed several approaches to tackle this problem ranging from direct application of OS and shared-storage virtualization techniques (e.g., Amazon EC2's DBMS-in-a-VM approach), re-interpretations of the same ideas in the DBMS space, to completely novel technical solutions. The common underlying intuitions are that: 1) databases represent an important enough subclass of applications to be tackled in a special way, and 2) databases offer a narrow and semantically rich API towards the applications, that can be leveraged by system builders to devise specialized optimization. This is a rapidly evolving field with advances both in academic and industrial context and therefore in this tutorial we try to provide a snapshot of the current state-of-the-art in what we label *Database Virtualization*, and conclude by identifying research trends in this direction.

2. TUTORIAL OUTLINE

This tutorial is intended to span three hours, and is divided into three sections. The first hour and a half will be spent on the first two sections. We now provide an outline of the topics covered in each section, and highlight how the subjects relate to challenges in building a Database as a Service infrastructure.

2.1 Process Elasticity

With an effective ability to consolidate operating systems into a single machine, virtualization is a key component in providing cloud offerings. Despite the presence of a controller (hypervisor) to share a system's limited physical resources, colocated VM instances do compete for the same limited resource capacity. To

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Articles from this volume were invited to present their results at The 39th International Conference on Very Large Data Bases, August 26th - 30th 2013, Riva del Garda, Trento, Italy.
Proceedings of the VLDB Endowment, Vol. 6, No. 11
Copyright 2013 VLDB Endowment 2150-8097/13/09... \$ 10.00.

mitigate the impact of resource starvation, two techniques are often utilized. First, load-balancing the placement of instances can remove resource bottlenecks, but requires the ability to migrate a VM instance between machines [1]. Second, resource allocation, or isolation, can be enabled by the hypervisor to ensure fair resource access between VMs [7]. Both solutions require the ability to characterize a workload and predict the impact of colocation.

This section of the tutorial will introduce the audience to virtualization and motivate embedding virtualization into a database system. Building on this foundation we will explore the use of virtualization techniques in databases, including hosting database servers within a VM and how virtualization techniques have influenced database migration, consolidation, and resource allocation.

2.2 Storage Elasticity

Another important challenge is efficient, elastic sharing, and access to storage resources. Significant industrial development has been focused in providing an automated, decoupled, storage infrastructure which controls the placement and load-balancing of IO workloads across a set of storage devices. These systems enable a VM to mount a virtual disk, which is hosted on one or more servers within the same network. The decoupling of storage from computation enables lightweight elasticity, consolidation, availability, and fault-tolerance.

By managing virtual disks for tens to hundreds of datastores, these systems highlight the need for automated solutions in placement and load-balancing. To manage these issues, a system must make decisions based on workload characterization [3], device models, and analytical or sampling-based formulations [4]. While this section of the tutorial will focus on these systems and the challenges in using shared storage abstractions, we will also touch on other storage system issues related to building a database platform, including dynamic resource allocation, scalable load balancing, workload characterization, and service level objectives.

2.3 Database Elasticity

With an understanding of existing techniques to enable load-balancing, resource-isolation, consolidation, and related issues in building consolidated services, the remainder of the tutorial will focus on recent advances in database platforms. This includes current architectures to enable multitenancy, how virtualization solutions work for databases, and identify the short-comings of prior solutions when applied in a database system. The exploration of this design space raises research questions this tutorial will cover.

Within the various multitenant systems, recent research has focused on bringing virtualization to database platforms. A critical issue in these systems is ensuring that each tenant has resources to serve well-formed requests within a certain time period, alternatively a service level object (SLO). As with other consolidated systems, several approaches can ensure that SLOs are met, even in the presence of dynamic workload. Resource isolation can ensure that each tenant is guaranteed a fixed or minimal amount of underlying system resources [6], or intelligent tenant placement can ensure SLOs are met [5]. Both approaches require the ability to predict the impact of colocation, model workloads, and predict resource consumption. For solutions that rely on tenant placement to ensure SLOs, a migration primitive must be available to move tenants between servers. Similar goals and approaches as VM migration have been proposed to address live database migration. We will review recent publications in the area of database migration. Beyond maintaining SLOs, we will discuss workload driven partitioning techniques for systems that allow tenants to span servers, using virtualization for high-availability, and using shared-storage to achieve

fault-tolerance. We will conclude the tutorial with thoughts on future problems to consider in building database platforms, such as data sharing, pricing and privacy/security concerns.

Acknowledgements This work is partly funded by NSF grants III 1018637, IIS 0905084, CNS 1053594, and NEC Labs America.

3. BIOGRAPHICAL SKETCHES

Aaron J. Elmore is currently a PhD candidate at the University of California, Santa Barbara. He has a MS in computer science from the University of Chicago. His research interests involve cloud computing, multitenant databases, and ecoinformatics.

Carlo Curino received a PhD from Politecnico di Milano, and spent two years as Post Doc Associate at CSAIL MIT leading the relational cloud project. He worked at Yahoo! Research as Research Scientist focusing on mobile/cloud platforms and entity deduplication at scale. Carlo is currently a Senior Scientist at Microsoft in the recently formed Cloud and Information Services Lab (CISL) where he is working on big-data platforms and cloud computing.

Divyakant Agrawal is a Professor of Computer Science and the Director of Engineering Computing Infrastructure at the University of California at Santa Barbara. His research expertise is in the areas of database systems, distributed computing, data warehousing, and large-scale information systems. He currently serves as the Editor-in-Chief of Distributed and Parallel Databases and is on the editorial boards of the ACM Transactions on Database Systems and IEEE Transactions of Knowledge and Data Engineering. Prof. Agrawal is a Fellow of ACM and a Fellow of IEEE.

Amr El Abbadi is a Professor of Computer Science at the University of California, Santa Barbara. He received his B. Eng. from Alexandria University, Egypt, and his Ph.D. from Cornell University. Prof. El Abbadi is an ACM Fellow, an AAAS Fellow and was Chair of the Computer Science Department at UCSB from 2007 to 2011. He has been Program Chair for multiple database and distributed systems conferences, most recently SIGSPATIAL GIS 2010 and ACM Symposium on Cloud Computing (SoCC) 2011, COMAD India 2012 and ACM COSN (Conference On Social Networks) 2013.

4. REFERENCES

- [1] C. Clark, K. Fraser, S. Hand, J. G. Hansen, E. Jul, C. Limpach, I. Pratt, and A. Warfield. Live migration of virtual machines. In *NSDI*, pages 273–286, 2005.
- [2] C. Curino, E. Jones, R. Popa, N. Malviya, E. Wu, S. Madden, H. Balakrishnan, and N. Zeldovich. Relational Cloud: A Database Service for the Cloud. In *CIDR*, pages 235–240, 2011.
- [3] A. Gulati, C. Kumar, and I. Ahmad. Storage Workload Characterization and Consolidation in Virtualized Environments. In *VPACT*, 2009.
- [4] A. Gulati, G. Shanmugathan, I. Ahamad, C. Waldspurger, and M. Uysal. Pesto: Online Storage Performance Management in Virtualized Datacenters. In *SoCC*, pages 19:1–19:14, 2011.
- [5] Z. Liu, H. Hacigümüs, H. J. Moon, Y. Chi, and W.-P. Hsiung. Pmax: tenant placement in multitenant databases for profit maximization. In *EDBT*, pages 442–453, 2013.
- [6] V. Narasayya, S. Das, M. Syamala, B. Chandramouli, and S. Chaudhuri. Sqlvm: Performance isolation in multi-tenant relational database-as-a-service. In *CIDR*, 2013.
- [7] C. A. Waldspurger. Memory resource management in vmware esx server. In *OSDI*, pages 181–194, 2002.