On Virtual Data Centers and Their Operating Environments

Vadim Kotov
Computer Systems and Technology Laboratory
HP Laboratories Palo Alto
HPL-2001-44
March 8th , 2001*

E-mail: kotov@hpl.hp.com

data center, virtual data center, operating environment, e-service, services, servers, service monitoring and management

The business and enterprise computing infrastructures are consolidating into data centers in order to improve the access control, increase security and reduce the maintenance costs. Big corporations or e-service providers span their data and applications over several data centers, combining their resources into virtual data centers. There may be many layers of virtualization in general case. At the bottom are platform centers that provide "physical" computing resources and logistics, such as computer platforms, basic software, databases, and network capabilities. At the top are e-service centers that face end customers and are specialized in providing specific e-services to the customers. Virtual centers that serve as intermediaries between the first two layers help to split design and management concerns, in particular, to balance customer satisfaction by services with efficiency of using platform resources. Each lower layer provides to a higher layer an operating environment, in which terms the higher level sees the lower layer and maps its objects onto the objects of the environment. The report outlines the basic high-level concepts, issues, and main problems, such virtual (core) services and virtual resources (servers), configuring of a virtual center inside a platform center or over several platform centers, deployment of services among servers of the virtual center, dynamic service monitoring and management.

* Internal Accession Date Only Approved for External Publication
© Copyright Hewlett-Packard Company 2001

TABLE OF CONTENTS

1 INTRODUCTION

E-business and enterprise computing infrastructures are consolidating into ***data centers***, as both enterprises and businesses seek to re-centralize or outsource their data processing and storage functions in order to improve the access control, increase security and reduce the maintenance costs [1,2]. There exist other networking computing paradigms, like the ***peer-to-peer computing*** [3,4]. However, they should be considered not as negation of the data centers, but as complementary concepts filling inescapable holes in the main trend. Simultaneously, the existing variety of electronically supported businesses, services, and commerce is converging into electronic services, or ***e-services***, that will be provided by networks of e-service providers.

Consolidation of the IT infrastructures into the data centers is accompanied by the growth of data centers scale, by increasing diversification of e-services, and by the dynamism of the ways in which the data centers are used by businesses and customers. Thus, ***scale***, ***diversification***, and ***dynamism*** are the keywords characterizing the future data centers.

Large customers, such as big corporations or e-service providers, profit from or are required spanning of their data and applications over several data centers, combining the resources of these centers. Smaller customers tend to use only part of the resources and services of a data center to which they outsourced their IT infrastructure. Combination of these options may occur in general case. In any case, the customers form their private ***virtual data centers*** that provide for them resource consolidation and location independence. The virtual centers also extend the customer's capability to optimize the center for their business needs and for particular pattern of workload. It hides the data center actual platform architecture providing instead for customers a much more simple and convenient "high-level" virtual architecture.

The virtual centers are built on the top of one or several ***base data centers***. They hide the base data centers actual architecture providing instead for customers a more simple and convenient "higher-level" virtual architecture.

In general case, a virtual center may be a base center for the next level of virtualization in the chain that may look as follows:

Hosting Data Centers -> ASP Provider -> Enterprise Data Center → Enterprise Branch Data Center.

In this chain, an owner of a data center is on the left side of an arrow and its customer is on the right side of the arrow, though this customer may be, in its turn, the owner of the right virtual center.

We will refer to the center that starts the chain as to ***a platform center*** (sometime also called physical center) and to the center that terminates the chain as to an ***e-service center***. The e-service centers face end customers and are usually specialized in providing specific consumer or business services to the customers.

We assume that the platform centers are not just co-location centers; they provide for hosted virtual centers computing resources and logistics, such as computer platforms, basic software and utilities, databases, network capabilities, and even middleware.

Virtualization is a way to kill two birds by one stone: one bird is the customer satisfaction by services and another is the efficient use of data center resources. Virtualization is required by diversification of e-services: web hosting services, application services, storage services, streaming media services, wireless access services, and others. The right strategy is to consolidate the base centers and diversify at the virtual levels. The more flexible and efficient are virtualization mechanisms, the larger variety of customer-oriented centers can be built (and in a more simple way).

Growing complexity and dimensions of data centers require some systematic view of the data center architecture. To help to navigate in this complex environment, this document introduces some new, higher-level architectural lexicon with concepts related to different aspects of the design, use, and management of data centers. These notions are arranged into a taxonomy of layers and hierarchies that allows us to discuss the data center architecture and behavior in a more systematic way.

Each next layer of virtualization is implemented on the top the preceding on. Its structure, functions and processes are mapped onto the structure, functions and processes of the underlying base layer. Inside each data center we distinguish two types of components:

- ***services*** provided by the center, and
- ***resources*** owned by the data center.

With some degree of simplification, we may say that services of a base data center are resources for its hosted virtual centers.

Both services and resources can be aggregated and represent hierarchies of component granularity: applications, application suites, services, service suites, computers, servers, clusters (and tiers), centers and networks of centers, etc.

Operating environment maps services onto resources and provides a framework and common logistics for the data center ***administration*** and ***management***. It provides an external view of a center to the higher layer virtual centers. In other words, each center sees its base centers in terms of operating environment of the lower layer.

The operating environment of a data center fully determines its efficiency and quality of service. If we view a platform data center as a (super) computer then its operating environment is an analogue of the kernel of an operating system.

Each of the operating environments deals with a large number of the administration and management tasks. We are focusing at the operating environment of virtual centers and discuss those issues that are important for the scalability, performance and manageability, such as:

- (static) configuring of a virtual center inside a platform center or over several platform centers,
- deployment/redeployment of services among servers of the virtual center,
- dynamic monitoring and management (utilization, inter- and intra-traffic control, load balancing, caching,…) of virtual data centers.

This document is the first step in studying the operating environments of virtual centers. It just outlines the basic high-level concepts, issues, and problems. The subsequent documents will refine specific subjects, such as

- more rigorous definition of services, servers, and related concepts, including service level protocols for virtual center,
- analysis of typical e-service architectures,
- analysis of typical platform architectures,
- analysis of typical virtual center configurations,
- customization of virtual data centers,
- trade-offs in distribution of the scale and functionality,
- operating environment schemes for different types of virtual centers.

2 EVOLUTION OF DATA CENTERS

2.1 Diversification of Data Centers

Internet is expanding not only in the customer base, but also in the spectrum of provided services, which may differ from each other significantly by their requirements of resources and operating modes. The current assortment of data centers, which is presented mostly by Internet data centers serving as Internet service providers, will expand into a broad spectrum of centers that will differ significantly by their scale and by their specialization, starting from application service providers to those centers dedicated to particular E-business segments and even to very specific services.

For example, the amount of data being stored is growing at an exponential rate, quadrupling every three years. Centralized data storage enables organizations to put essential and frequently accessed data in a single location and make it available to customers anywhere anytime quickly and efficiently. So, ***storage centers*** will form the high end of data centers. They will provide high-capacity storage capabilities and specialize at hosting of large “back-end” databases, data warehouses, and large information libraries. The storage centers connected by a network form decentralized storage that nevertheless can provide a single-point view of information for customers through virtualization of these centers.

Mobile data centers provide a provision of back-up systems to the customers in the case of unforeseen circumstances that could interrupt their business operations for an extended period of time.

Streaming media centers provide the instantaneous delivery of live and on-demand audio, video or other multimedia content over the Internet or enterprise intranets. They contain presentation servers, archives of past events, data reporting tools, distribution systems that direct audience members to the closest regional center.

Mobile Internet is unique and powerful field for further Internet growth. But it requires seamless integration of multiple wireless networks and mobile devices. ***Wireless*** (sometimes also called mobile) ***data centers*** act as the wireless providers for companies that don't want to layout the millions it would cost to build wireless centers for themselves. These centers focus on gateway solutions for mobile Internet, converting existing different wireless protocols into others and enabling wireless content providers not to offer their services based on all these protocols.

Internet services, application services, and storage services form a sort of three macro-tiers in the data centers space. Different combinations of their capabilities will produce “functionally complete” ***full service centers*** that will target different E-service areas and customers (enterprise, electronic delivery services, education, community, etc.).

2.2 Storage Centers

Information and storage management strategies require special considerations and planning for the following aspects: continuous volume expansion, backup, archiving, data availability, data integrity and security, disaster protection and recoverability, cross-platform data sharing, data manageability.

The storage capabilities are implemented either as a storage center or as a back-end tier in a data center. In both cases, this is a collection of storage servers that form a storage virtual architecture, on the top of which a storage service infrastructure is implemented. Modularity is very important from much consideration, primarily due to the volume expansion and the possibility to easy map the services onto the data center architecture.

Storage Area Network (SAN) is a basic paradigm to be developed into the storage centers.

Main research problems are:

- Scalable, high-performance, distributed architecture solutions specifically for large storage centers.
- High performance parallel architectures for storage servers and clusters of servers.
- True information sharing with direct, seamless access to the most current information, regardless of the platform: UNIX, Linux, or NT systems.
- Tracing, measurement and analysis of huge amount of diverse data (text, images, video, and voice) to learn about the storage center access patterns.
- Query results caching and load balancing in storage centers.
- Data partition, replication, and placement methods to optimize the data deployment for maximum performance.

2.3 Virtual Data Centers

The success of a data center will be depend on how well the data center architecture supports the basic customer requirements:

- expected end-to-end performance,
- low total cost of services,
- openness, scalability, flexibility to business changes,
- accessibility, availability and security.

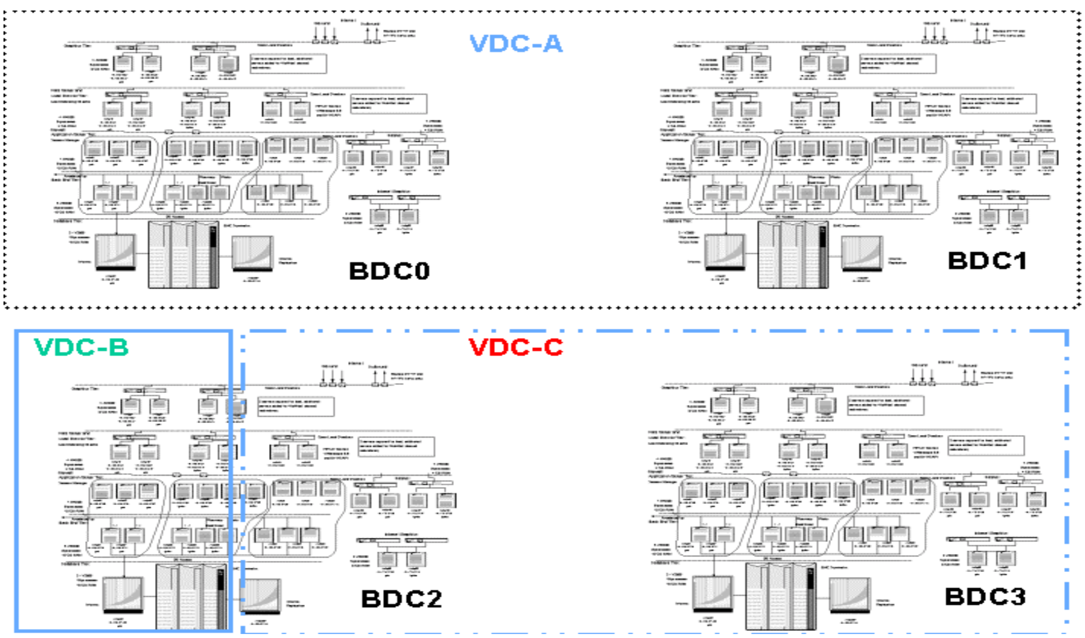

Figure 1. Virtual centers over platform data centers.

Different customers may have different (including partially contradictory) compositions of requirements. To meet them effectively, a platform data center hosts one or several ***virtual data centers***. The virtual data centers, being primary customers of the platform data centers, offer their services directly to end users or to the secondary (e-service) centers that actually provide end e-services. For example, geographically distributed platform data centers may host a virtual data center of a company XYZ that is their primary customer. A service center that works directly with clients, subscribed for the services provided by the company XYZ, have actually access to the virtual data center of this company.

A virtual data center provides a completely independent and isolated infrastructure for its owner, which is managed exclusively by the owner.

Figure 1 shows examples of three virtual data center built over four platform data centers. The first virtual center VDC-A is built on the top of two platform centers BDC-0 and BDC-1 using all resources of both centers. The second virtual center VDC-B uses only part of the resources of the platform center BDC-2. Finally, the last virtual center VDC-C combines part of the resources of the base center BDC-2 and all resources of the platform center BDC-3.

3 ARCHITECTURE OF DATA CENTERS

3.1 Template Architecture

First of all, we want to outline those features of data centers that are common for all layers and present them as a sort of template architecture of data centers. As it was mentioned in Introduction, we want to distinguish in any center such objects as services and resources.

Service is a work done by a data center for a hosted virtual data center or for an end customer. Services may be aggregated into hierarchical ***service suites*** and consist of subservices. Some other additional relations between the services may accompany the hierarchy relation.

The nature of these relations depends on the specific character of the services. For example, they may reflect actual ***service dependency***: a service may not be completed without requesting the help of another service. Or a service may be called by another service if some special situation is present (an extra service for extra money,). The relations between services force the services to communicate and exchange data.

The frequency of communication and the volume of the exchanged data are important quantitative characterization of the relations. They define the service ***workload pattern.*** The set of all services and all relations between them (together with their quantitative parameters) plus the service workload pattern form the service environment.

Resource is a computing facility (may be a virtual one) that hosts a service or several services and provides the services with everything needed to execute them. (As we noted in Introduction, we may say that services of a base data center are resources for the virtual center that it hosts).

Operating environment maps services onto resources, administer and manage them.

Figure 2 displays a sort of a map of basics notions that we use in the context of the template architecture. Figure 3 instantiate the template architecture for three layers:

platform, virtual and e-service layers. Figure 4 shows how different types of data centers can be built using combinations of layers: one virtual center can be built on the top of two platform centers; two virtual centers can be built on the top of one platform center; one e-service center on the right side may combine all layers.

The overall strength and quality of a data center architectural decision will depend mostly on how efficiently its services are mapped onto its resources.

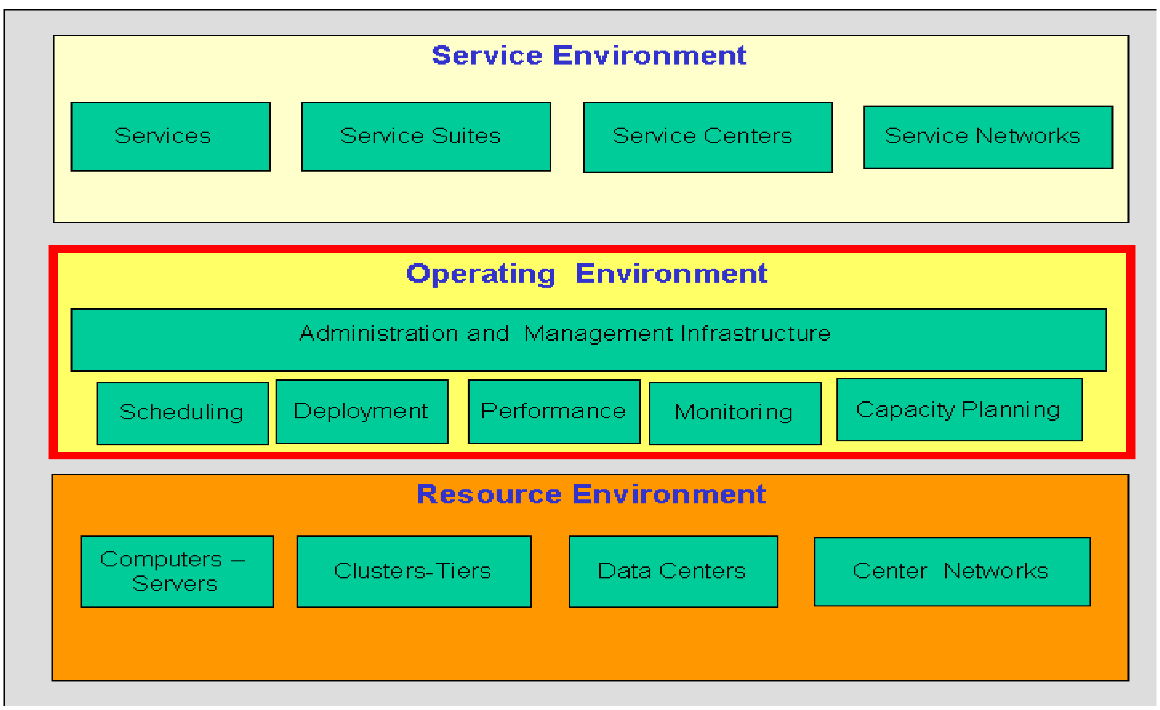

Figure 2.Template data center architecture

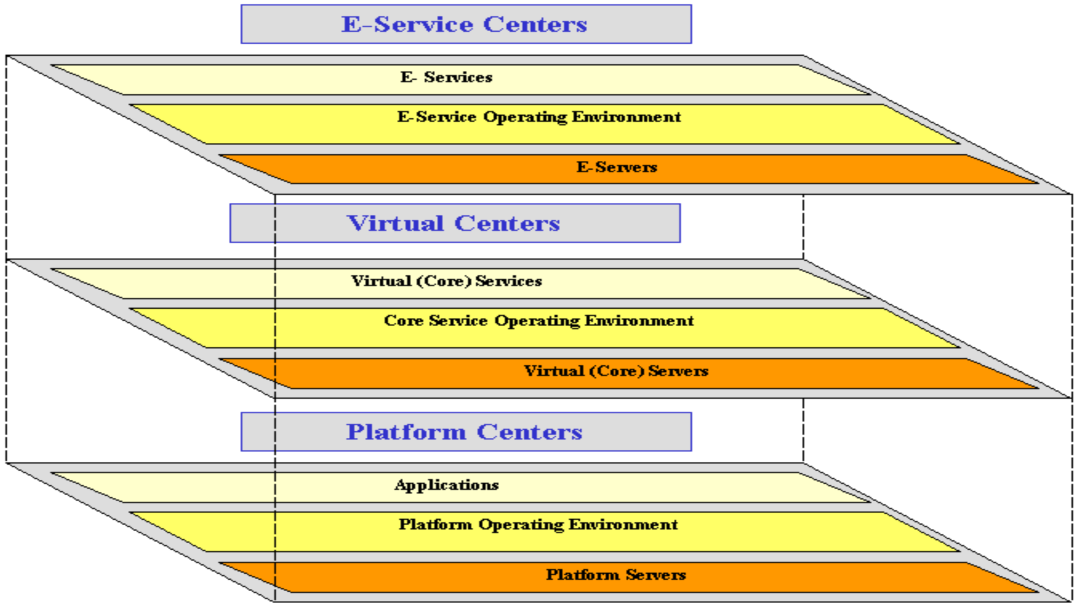

Figure 3. Three data center layers.

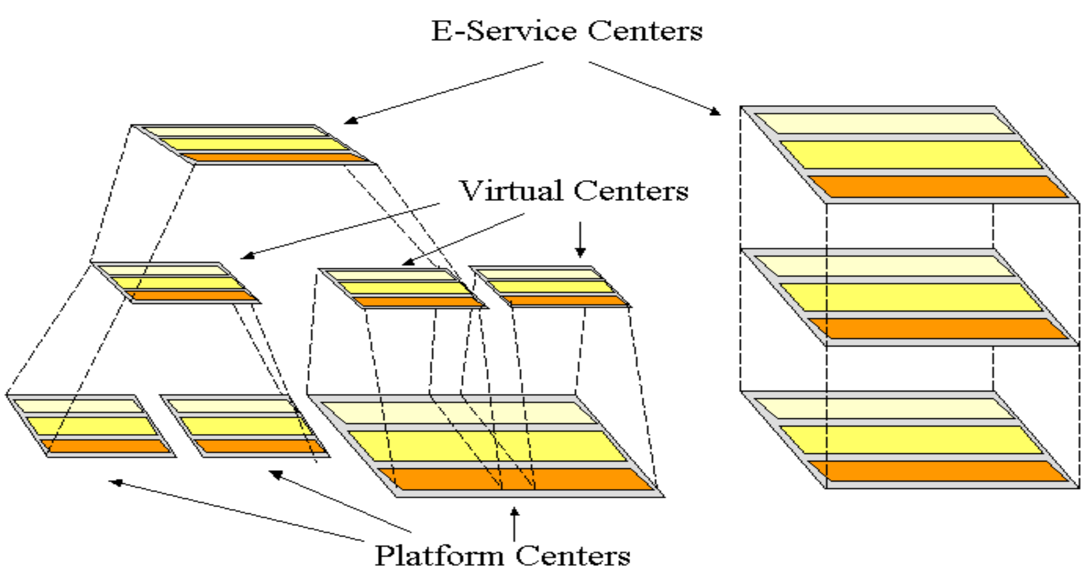

Figure 4. Data centers combined of different layers.

3.2 Platform Architecture

Virtual data centers generally do not own the hardware or basic software. Instead, they rent space for services housed on the equipment of their base platform centers. A virtual center may pay fee, generally depending on the resource usage, to use the equipment of the base center, which the latter cares for and monitors. The platform center is responsible for the operations of the system, starting from the facilities infrastructure (UPS, generators, environment, and security), to installation and startup services for software and hardware, and up to full-service LAN administration. Often, its physical hardware is shared by several virtual data centers.

Typical platform architecture includes applications that are services of this layer and computing resources that are computers of different types, often combined into ***clusters*** or/and ***tiers***. The clusters are related to scalability, the tiers are related to functionality of data centers. Switches, interconnects and routers combine computes, clusters and tiers into the data center platform. External networks connect geographically distributed data centers.

Components of a platform data center are typically characterized by the following parameters (aggregated for aggregated components):

- processing speed,
- storage space,
- interface bandwidth,
- utilization,
- reliability.

Scalability incorporates concepts and principles at which a platform data center may be significantly extended in the number of nodes, preserving the high end-to-end performance.

To accommodate the growth in the number of customers and the growths in their businesses, the data centers should be scalable. With the data centers with tens of thousands nodes in mind, the platform architecture should focus on clustering and connectivity as the main issues of scalability.

Clustering merges the computers into groups that are more tightly coupled among themselves than with other computers. It is possible that there will be a multi-level hierarchy of clusters. Such a partitioning may be induced either by the interconnect bandwidth or by the mapping of the virtual architecture onto the platform architecture (or both).

Connectivity is a crucial issue. Its importance grows with the scale of a center. Examples of problems in the platform architecture related to scalability and connectivity:

- Technical, content, and other criteria and constrains to combine nodes into clusters.

- Homogeneous clusters versus clusters with nodes that functionally complement each other.
- The number of the clustering levels and the number of servers or clusters at each level as a function of the data center scale and workload.
- Smoothing proximity-dependent latency differences of the clustered architecture.
- Requirements to connectivity at each clustering level:
  - Requested interconnect bandwidth and latency.
  - How intelligent should be the level interconnect?
  - Which protocols and connectivity topology are optimal for the level?

3.3 E-Service Architecture

Those electronic services that help businesses to provide additional values and convenience to their customers are referred to as ***e-services*** [5]. E-services can be used by people, businesses, and other e-services and can be accessed via a wide range of devices from "anywhere anytime". In some sense, the e-service architecture is the "most virtual" architecture, as it just represents an e-service framework that includes:

- dependencies between e-services,
- e-service communication and collaboration rules and patterns,
- organizational and/or geographical distribution,
- mechanisms to advertise and discover e-services, e-service brokering, and similar aspects.

E-service architecture is dynamic in that sense that e-services may emerge and disappear; the number and nature of customers and e-service providers may be not fully known; radically new types of e-services may appear.

3.4 Virtual Architecture

As it was noted in Introduction, there may be many layers of virtuality. However, we will consider here, for the sake of simplicity, only one layer that is hosted by one or several platform centers and host one or several e-service "centers". (We use quotes here, as there may be no special e-service center; the virtual center may play the role of such center.)

The main artifacts of a virtual data center are ***virtual services,*** which we call also ***core system service*** as they considered to be the core concept of the service-centric system organization, and ***virtual resources,*** which we call ***servers***. (Both core services and servers are subjects of a special technical report [6] to be published soon.) Servers and may be aggregated into hierarchical server clusters (tiers are clusters too).

A virtual service is a high-level system concept denoting a set of computing activities that are activated by some ***service request*** from a client (customer) of the e-service layer or by another service and result in a ***service response*** (or responses) forwarded back to the originators. The main difference between an e-service and a core service is the following. The former supports and regulates interactions between clients and services, such as service search, domain protection, service brokerage, etc. Thus, e-services are abstractions of the customer services; they mostly address the templates of interaction between customers and service providers. The latter servers to organize, deploy and manage e-services, that is, the applications and data that participate in delivering e-services, in data center context and in an efficient way.

E-services are primary (but not the only) originators of service requests. virtual services generate secondary requests to other virtual services if they need help in executing their requests.

Thus, we define the virtual service from the system point of view and use it as a pure system term. Services provided by virtual data centers may differ very much from each other by their contents and purposes. But they may be classified and arranged into some content-independent entities that are characterized by their resource requirements to the center and its servers.

The core services may be characterized, for example, by the processing power they require, by the size of the storage space they need to access, by the volume of data they transfer to served clients or between servers, by the maximally allowed response time, etc. For large-scale systems, it is better to have more simple resource metrics that is obtained by converting multiple resource characteristics into one “service share”.

We may regard services and related notions introduced here as forming one more level (***service level***) on the top of the networking application level. XML is the base notation for the definition of services and their interactions. Assuming that the services interact using XML, there is yet more to define: the service level structures and protocols are need to be agreed upon, in the same way as the Web has agreed on HTTP. Protocols for interoperation in different e-service areas exist. However, these protocols are disparate and have relatively high implementation costs while we need a simple format, easy to use and easy to understand. The solution resides in building our service level protocol on the top of one simple, lightweight protocol: a protocol like the Simple Object Access Protocol (SOAP) that is flexible, platform neutral, and text-based.

The design of service architecture is determined by the businesses and IT processes to be supported by a virtual data center and includes several steps:

- goals and objectives of the design are analyzed, QoS requirements, the virtual data center scale and budget are studied,
- characterization of services requirements and profiling of a potential workload are made,
- users special requirements are analyzed and classified,

- an assortment of services is identified together with dependencies between them, communication patterns and requirements to resources.

A full-service virtual center consists, as a rule, of three or four tiers, for example: access tier, web tier, application tier, and storage tier. An Internet and storage data centers have less tiers. The former is shifted to the front-end tiers; the latter is shifted to the back-end tiers.

The platform computers host servers. Servers host services. Thus, the architecture of a virtual data center provides a bridge between a given e-service architecture and a given platform architecture, mapping in the end the e-service architecture onto the platform one.

In general, there is no one-to-one relationship between computers, servers, and services. A server is either a computer or a computer cluster. A server may provide either one specific service or several services.

A virtual architecture is a solution of an equation with two variables: one is e-service architecture and the second is platform architecture. The problem is that this equation has too many solutions and finding a right one requires in depth analysis of how the requirements of the service architecture can be met by a given platform architecture.

Another problem is that both platform and e-service architectures are not static: they are changing due to an increased number of e-services or/and equipment, introducing new type of applications and computers, new types of networks, etc. The virtual architecture should be “dynamically optimal”, that is to be reactive, flexible and resistant to negative side effects of changes.

Scalability of a virtual data center that is hosted by only one base center is put to the test when it expands beyond one base center. So, the architecture of a virtual center should be both ***intra-center*** and ***inter-center scalable***, that is, easily expandable to geographically distributed platform architecture without significant remodeling of the virtual architecture

It seems that inter-center scalability is more difficult to achieve than the intra-center one. Expandability that represents just a replication of identical data centers seems to be easier done than combining different specialized centers into one virtual center.

4 OPERATING ENVIRONMENTS OF DATA CENTERS

The life cycle of a data center includes: its deployment; upgrades, modifications, and re-engineering; runtime management. An operating environment of a data center should support all the phases of its life cycle and technically is:

- an architecture, including a set of guidelines and standards that define what are the center components (services and resources) and how they fit together and interact,
- a definition for acceptable specifications, interfaces, and protocols both for intra- and inter-center communication,

- administrating of mappings of the center services onto its resources,
- a runtime management environment that facilitate establishing clashless, secure, reliable, and efficient execution of the higher-level tasks and, at the end, the services requested by customers.

Specification of the architectures, interfaces and tasks at all layers should be based on standards (expressed, for example, in XML or some XML-based specifications and protocols).

4.1 Platform Operating Environment

The main administrative task of the platform operating environment are:

- capacity planning,
- installation and startup,
- resource allocation/reallocation.

Capacity planning is made for virtual data centers hosted by a given platform center, specifically ***capacity on demand***, an incremental planning, when an initial capacity is provided on the customers demand and then is added as needed.

Installation and startup services for software and hardware, as well as software upgrading, versioning, package dependency control, legacy management, are currently costly and time-consuming procedures. Quick deployment of ready-to-use systems and software management (distribution, assignment, …) will define the quality of services provided by platform data centers.

A platform center that put its resources at disposal of several virtual centers should be able to carry out rapid reconfiguration if some change in its customer population or in customers' demand occurs. Resource allocation includes resource planning, allocation and optimization of CPU time, storage, I/O, bandwidth, … .

At the platform level, there are individual or specific ***management tasks,*** solutions for which are either available currently or will be available soon, such as:

- network management (connectivity, traffic monitoring and control, diagnostics),
- backup, failure, fail-over and recovery management,
- system evolvement (scaling, rapid capacity re-planning, …),
- environment and infrastructure management (power supply, external networks…),
- monitoring of a large number of system events and dynamic component parameters.

4.2 Virtual Operating Environment

An important factor of a virtual center performance is a smart organization and control of message and data traffic between the center and its customers and among the center services and servers. An important factor of the traffic efficiency is ***partition, placement,*** and ***replication*** of services among servers, as it creates a “primary” traffic of requests and responses. The latter generates a “secondary” traffic between services that generate a “ternary” traffic between applications and system programs, and so on. A bad organization of the primary traffic is the main cause of the bad secondary and subsequent traffics.

Main administration problems are:

- Analysis of the e-service parameters (including the e-service workload) in the businesses that will be owners or customers of a virtual data center.
- Analysis of the virtual servers parameters (their capabilities to host services).
- Efficient and secure partitioning (including replication) of services among servers providing an efficient average traffic for statistically dominating workload.
- Automatic (centralized/decentralized) deployment of services on the basis of the partitioning.
- Optimal repartitioning of services onto servers based on the data center statistics.

The dynamic management tasks that are addressed now:

- monitoring of large number of system events and dynamic parameters,
- the non-intrusion measurement of traffic to/from services and servers,
- quality management (availability, guaranties),
- accounting management,
- security and access management.

New management task are related to the service management and to dynamic management of the system load and traffic:

- dynamic repartitioning of data, applications, and services in case of servers failure, uneven load, bottlenecks,
- caching, traffic routing, and load balancing.

4.3 Operating Environment Structure

It is important to have a proper distribution of the management task between the platform and virtual operating environments. This may be a special problem to resolve, as there is no clear line of demarcation and it may migrate for different types of data centers and in time when some functions of the virtual level will move to the platform level. Being clearly separated, both environments may communicate with each other in a secure way.

Though different in their tasks and administered objects, both operating environments may have similar infrastructure, as well as unified administration and management principles involved.

As data centers are large and complex system, it is extremely important to streamline and consolidate both administration and management tasks in the operating environments:

- reduce the number of data monitored and displayed for the system administrators and managers,
- make all the management tasks feel and look alike, have a common console for the management
- replace gradually the operator-based management by a reactive automatic management of events, load, and traffic,
- decentralize the operating environment; find a right proportion of centralized/decentralized management of the distributed service.

Functions of a dynamic operating environment fall into three categories:

- monitoring and measurement,
- decision making, and
- controlling.

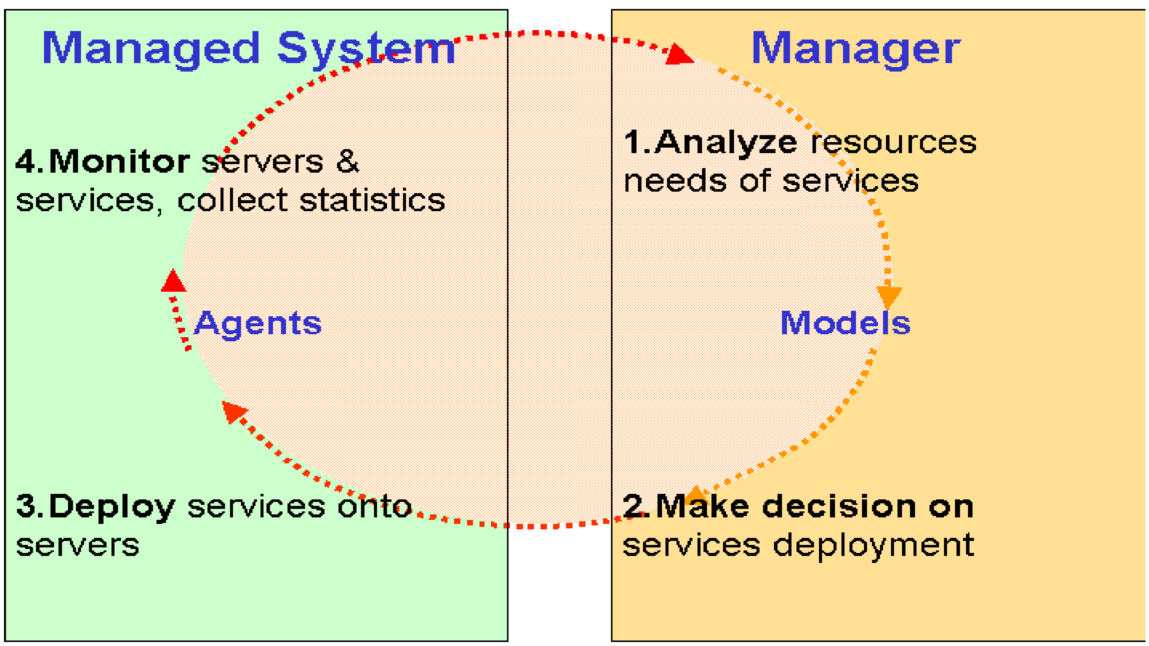

Figure 5. Management loop.

Managers and ***agents*** implement these functions. Agents do monitoring and measurement and send the monitoring results to managers. The managers make decisions and send orders to agents to make changes in the center structure or/and behavior. To make intelligent decisions, the managers may use data center models that contain information about the center that is necessary and sufficient for finding optimal management solutions. The management loop formed by managers and agents is shown in Figure 5.

Decentralized operating environment of large data centers requires the distribution of management control and responsibility for certain domains to responsible managers. There may be a hierarchy of managers with managers of lower levels serving also as agents for the higher-level managers. (We call them ***managents***).

5 CUSTOMIZABLE VIRTUAL DATA CENTERS

Customers first! The data center design should start from the analysis of customer businesses and services. E-business processes in different industries and corporations have both common and specific features. The whole space of e-business processes may be partitioned into ***business segments*** of related business models. In the same way, the space of e-service infrastructures that support different business models may be ramified into correspondent ***e-service segments***. Finally, implementation of the e-service segments in virtual data centers results into a branching structure of typical virtual center solutions. Such a ramification may have, in general case, several virtual layers, each layer being a refinement of the previous one.

After we have structured virtual data centers according to e-service infrastructures, we need to identify, what are their common features and what features are specific for a given virtual center or a group of virtual centers. These features are, for example:

- the number and assortment of services,
- required number and assortment of servers of different types,
- requirements to base data centers (geographical distribution, security, …), and similar.

Knowing the common features, we can construct a ***generic virtual center*** for a given e-business segment. Knowing specific features of a center for a specific e-service segment, we can customize the generic virtual center into a custom one.

This ramification procedure is valid also for the data center operating environments: both common and specific objects and tasks are identified for both administration and management parts of the environments.

6 CONCLUSION

The Internet-enabled computing is evolving into a global environment of very large scale, complex connectivity, divers usage, and great dynamism. Data centers, and specifically virtual data centers, will became the piers of this environment. Such complex systems are new, cutting edge products that create completely new research problems. Among them deployment, administration, and management are the most important and difficult issues, as they are the basics of highly effective operating environments of data centers. Their solution will bring a big advantage to those data center providers that implement them.

7 ACKNOWLEDGEMENT

Many thanks to Sven Graupner and Holger Trinks for many inspiring discussions and useful suggestions.

8 REFERENCES

1. Goodyear, M. (Editor). Enterprise System Architectures. CRC Press, 1999.
2. Rajput, W. E-Commerce Systems Architecture and Applications. Artech House, 2000.
3. Chen, Q., Hsu, M., Dayal, U. Peer-to-Peer Collaborative Internet Business Servers. HP Laboratories Technical Report HPL-2001-14, 2001.
4. Kan, G. (Editor). Peer-to-Peer: Harnessing the Power of Disruptive Technologies. O'Reilly & Associates, 2001.
5. Peraire, C., Coleman, D. Modeling for E-Service Creation. Technical Report. System Design Laboratory, SRI International, July 2000.
6. Kotov, V. Towards Service-Based System Organization. . HP Laboratories Technical Report to be published.