State Data Center: Top Ten Design Parameters

Ajay Ahuja[1]

ABSTRACT

State Data Centers (SDC) are one of the key elements of the e-Governance Initiatives of the Government Of India. All the major States across the country will be hosting Citizen Services out of these State Data Centers. Many States are in the conception stage of designing and architecting these Data Centers. Many factors and design parameters need to be evaluated and considered while designing theses Data Centers. These design parameters are critical for the successful implementation of these State Data Centers. This paper discusses Top Ten, Most important Design parameters, which should be considered while designing State Data Centers.

Keywords: State Data Centers, Design parameters, Architecture

1. Introduction

As a part of e-Governance initiative, to provide IT enabled Citizen Services to the common man, Government of India has started setting up State Wide Area Network (SWAN) connectivity across various states. This is going to be followed by setting up of Citizen Service Centers (CSCs) as front-end access kiosks and State Data Centers (SDCs) at the backend. The SDCs will host the Services and the data. Data Center Design and implementations have been complex in past and have been continuously changing over time with the advancement and changes in technology and infrastructure components. Today, vendors have come out with Virtualized Data Centers, Mobile Data Centers and even Data Centers in a Standard shipping Container.

Every State shall be investing heavily in the State Data Center initiative and needs to evaluate various parameters before getting into actual implementation of State Data Centers. Various Design parameters need to be evaluated and considered before designing and architecting the State Data Centers.

Following "Top Ten Design Parameters" are the foundation for the design of any Data Center. These parameters are a must to be considered and evaluated for State Data Center implementation and design. They are the building blocks for the State Data Center design. (Figure 1)

The Top Ten Design Parameters for the SDC are:

- Reliability
- Availability
- Serviceability
- Scalability
- Manageability

[1]Sun Microsystems India Pvt. Ltd, Munirka, Olof Palme Marg, New Delhi 110067, India, (E-mail: ajay.ahuja@sun.com)

- Virtualization and Consolidation
- SWaP (Space, Watts/Power and Performance)
- Modularity
- Open Systems
- Security

This paper discusses each of these parameters and gives possible recommendations and suggestions against each of them. These recommendations can be useful to the States while evaluating and selecting various technologies during the SDC implementation.

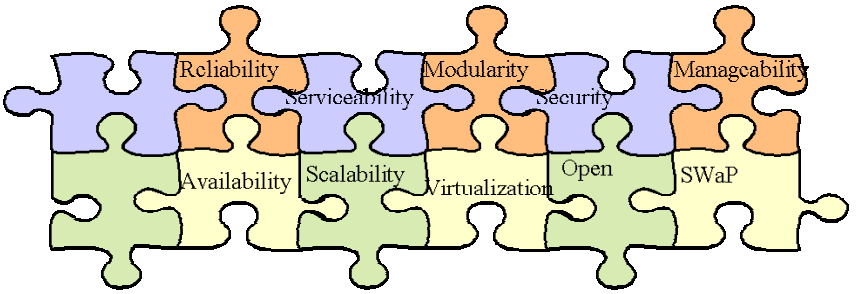

Figure 1: Design Parameters as Building Blocks for SDC

2. State Data Center: Top Ten Design Parameters

Reliability

As defined in wikipedia (http://en.wikipedia.org/wiki/), In general, reliability (systemic def.) is the ability of a person or system to perform and maintain its functions in routine circumstances, as well as hostile or unexpected circumstances. The IEEE defines it as " . . . the ability of a system or component to perform its required functions under stated conditions for a specified period of time."

Reliability is quantified as Mean Time Between Failures (MTBF) for repairable components and Mean Time To Failure (MTTF) for non-repairable components. (http://www.vicr.com/documents/quality/Rel_MTBF.pdf)

While implementing SDCs, the Government organizations should ensure selection of reliable systems, reliable Storage and reliable network components. The individual components within the systems including Power Supplies, System Memory Boards, Hard Disks etc should be highly reliable and the organizations may want the vendors to submit tested MTBF and MTTF values for individual components and or Systems. There should be no compromise on selecting most reliable equipment for the back end data center equipment like Servers, Storage and Network components.

Availability

As defined in wikipedia (http://en.wikipedia.org/wiki/) , the term availability has the following meanings:
1. The degree to which a system, subsystem, or equipment is operable and in a committable state at the start of a mission, when the mission is called for at an unknown, *i.e.,* a random, time. Simply put, availability is the proportion of time a system is in a functioning condition.
Note 1: The conditions determining operability and committability must be specified.
Note 2: Expressed mathematically, availability is 1 minus the unavailability.

2. The ratio of (a) the total time a functional unit is capable of being used during a given interval to (b) the length of the interval.
Note 1: An example of availability is 100/168 if the unit is capable of being used for 100 hours in a week.
Note 2: Typical availability objectives are specified either in decimal fractions, such as 0.9998, or sometimes in a logarithmic unit called nines, which corresponds roughly to a number of nines following the decimal point, such as "five nines" for 0.99999 reliability.

The most simple representation for availability is as a ratio of the expected value of the uptime of a system to the aggregate of the expected values of up and down time, or

$$A = \frac{E[\text{Uptime}]}{E[\text{Uptime}] + E[\text{Downtime}]}$$

While designing and implementing State Data Centers, The Government organizations have to consider this parameter, in term of High Availability of Servers, Storage, Network, and Application. Ideally, The Application should be available to the citizen's 24 Hors a day, 365 days a year. To achieve this kind of Availability, the design should incorporate reliable systems in High Availability Clustered environment, so that any failure in any component on one system should in no way affect the Citizen Services.

The Government departments should ensure high availability of individual systems, by having every component in the system as redundant, be it CPU, Memory, Disk Sub System, Network cards, Interconnect or even System Clock. There are systems available in the market with Main Frame class features, which can be configured in Highly Available architecture. The design should ensure High Availability Fail over Cluster at the System level so that a failure in one system does not cause failure of services. This can be accompanied by using Cluster Software's like Sun High Availability Cluster, Veritas Cluster etc. Load Shared and Load Balanced application architecture, should also be provisioned at application level to achieve application level availability. Oracle Real Application Cluster is an example of Database level Highly Available architecture.

Availability is generally expressed as Uptime in a given year. Uptime and availability are not same. A System may be Up and running but may not be available to the users due to network problem. So SDCs should be designed with an uptime parameter. Availability for highly available system is calculated based on uptime and downtime of system or services. Since the Services should be available to the citizens at all times, the Availability should be calculated based on Services, using the above mentioned formula.

Common values for highly available Systems are (http://en.wikipedia.org/wiki/):

- 99.9% ≡ 43.8 minutes/month or 8.76 hours/year
- 99.99% ≡ 4.38 minutes/month or 52.6 minutes/year
- 99.999% ≡ 0.44 minutes/month or 5.26 minutes/year

The State Data Centers should be designed for at least 99.99% Service Availability. This parameter should be kept in mind in selecting the hardware infrastructure in implementing the SDCs.

Serviceability
As defined in wikipedia (http://en.wikipedia.org/wiki/), in software engineering and hardware engineering, serviceability is also known as supportability, and is one of the-ilities or aspects. It refers to the ability of technical support personnel to debug or perform root cause analysis in pursuit of solving a problem with a product.

While selecting the infrastructure for the SDCs, Government departments should ensure that the Systems and Storage is serviceable online. For example to replace a failed CPU, Memory or I/O component, one does not have to bring down the system, this should be possible online. Any failure in any component should be visible through LEDs, Logs, alarms etc. The system should have provision for online upgradation and online Dynamic Reconfiguration. This should allow adding or removing faulty or healthy components on the fly.

The underlying Operating system should have features to predictively detecting and correcting faults. Solaris Operating System, which is an open source Operating System, provides these features and runs across heterogeneous platforms. The same is strongly recommended for the SDC environment.

The combination of above three design parameters is called as RAS (Reliability, Availability and Security). State Data Centers should ensure RAS infrastructure.

Scalability

As defined in wikipedia (http://en.wikipedia.org/wiki/), scalability is a desirable property of a system, a network, or a process, which indicates its ability to either handle growing amounts of work in a graceful manner, or to be readily enlarged. For example, it can refer to the capability of a system to increase total throughput under an increased load when resources (typically hardware) are added. An analogous meaning is implied when the word is used in a commercial context, where scalability of a company implies that the underlying business model offers the potential for economic growth within the company.

Scalability, as a property of systems, is generally difficult to define and in any particular case it is necessary to define the specific requirements for scalability on those dimensions, which are deemed important. It is a highly significant issue in electronics systems, database, routers, and networking. A system, whose performance improves after adding hardware, proportionally to the capacity added, is said to be a scalable system. An algorithm, design, networking protocol, program, or other system is said to scale if it is suitably efficient and practical when applied to large situations (e.g. a large input data set or large number of participating nodes in the case of a distributed system).

Scalability is one of the key parameters in designing State Data Centers. Since most of the States are in initial stages of IT enablement, the requirements in terms of infrastructure and resources after next two to five years is not known. A Particular state may require to cater to twice the present load after 2 years and there may be another one which may require only half the proposed infrastructure, due to social, political or other reasons. There is no well-defined definite growth parameter. Therefore there is no well-defined fixed scalability parameter as of today. For this reason, the systems selected today, for designing and implementing the State Data Centers should be highly Scalable and flexible.

The Servers, Storage and Networking infrastructure, selected today, should have provision for at least twice the scalability. Scalability can be Horizontal or Vertical Scalability. Horizontal Scalability implies adding more and more separate independent systems to service the increased load. Vertical Scalability is adding resources within the box to cater to increased workload. The Back end systems, like Data base Servers, Data Warehousing systems, Centralized Storage should all be Vertically Scalable, to at least twice the present capacity. Whereas the front-end systems, like Application Servers, Web Tier systems, Security Systems etc should be horizontally scalable.

It is also strongly recommended to have similar Operating System across the back-end, and the Front-end Systems. Open Source Solaris Operating System can be a good choice for maximum Scalability for back-end systems and horizontal seamless scalability across the front-end Systems.

Manageability

State Data Centers will involve many types of equipment as a part of the infrastructure. This will include Servers of various categories including back-end servers, Tier0 or front-end servers, and Storage, which may be Fiber based Storage area network (SAN), Network Attached Storage (NAS) or Direct Attached Storage (DAS). The system will be connected using Local Area Network and connected to Citizen Services Kiosks or the outside world over Wide Area Network.

It is extremely important to consider the manageability aspect of these components. Effective, Easy and Efficient management of these components is one of the major design issues to be considered while designing and architecting the SDCs.

The design should ensure out of band and in band management of systems and storage. This implies that there should be provision to manage the infrastructure components, while they are up and running and even when they are not operational but just powered on. There should be provision to remotely manage the servers and the storage, in a secure manner.

The selection of equipment should be done keeping in mind the ease of installing, configuring, patching, upgrading, monitoring and diagnosing the equipment from remote consoles. There are many management solutions available from various vendors to achieve this functionality. SDCs should ensure that the hardware supports remote in-bound and out-bound management and corresponding software's to achieve desired functionality are a part of Manageability Solution.

As defined at wikipedia, Manageability solutions refer to a range of software available to monitor and maintain a computer system in good working condition. This software can be divided into hardware monitoring and software monitoring solutions. Solutions that monitor the computer hardware, like temperature and power variations, are hardware monitoring systems. These applications have a high interaction with the hardware and are usually low-level systems. Software monitoring systems interact mainly with the operating systems. i.e., they are used to manage the operating system rather than the actual hardware.

WBEM is an industry-level standard for computer management solutions.

Some of the companies providing manageability solutions are (http://en.wikipedia.org/wiki/) :

- Hewlett Packard provides SAM or System Administration Homepage for HP-UX. In addition, HP has a wide range of enterprise management suites also, like OpenView. These are aimed at HP server customers.
- IBM provides management software for AIX, like SMIT.
- Sun Microsystems provides Solaris Management Console for Solaris.
- Webmin is a manageability solution for Linux and most any flavors of Unix.
- Microsoft provides its own control panel for windows administration.

Most of the Servers available today come with a separate Remote management processor for remote management and Lights Out Management. The servers selected for the SDCs` should have this functionality.

Virtualization and Consolidation

Virtualization is becoming extremely popular and customers are realizing the benefits of the same. Virtualization at Server, Storage and Network levels is getting popular. With the Virtualized Data Center design Government organizations can look at their Data Center infrastructure as a pool of shared resources,

which can be provisioned, re-provisioned, dynamically designed and re-designed based on present and future application needs (Gupta S K, et al, 2007)

Virtualization is possible both at hardware and software levels. Various vendors have different Virtualization techniques.

Virtualization Techniques can be broadly classified into following categories:

- Server Virtualization
- Operating System Virtualization
- Storage Virtualization
- Network Virtualization
- Virtual Access Software (Client Level Virtualization)
- Virtual Application Environment
- Virtualization for Provisioning and Management

In computing, Virtualization is a broad term that refers to the abstraction of computer resources (http://en.wikipedia.org/wiki/) One useful definition is "a technique for hiding the physical characteristics of computing resources from the way in which other systems, applications, or end users interact with those resources. This includes making a single physical resource (such as a server, an operating system, an application, or storage device) appear to function as multiple logical resources; or it can include making multiple physical resources (such as storage devices or servers) appear as a single logical resource."

Since the most of the states are starting the State Data Center initiative, they may not be able to predict the growth and exact infrastructure requirement for next couple of years. The design should be flexible and modular, so as to allow easy expansion and growth, as on need basis. Virtualization at all levels should be kept in mind while designing and architecting the State Data Centers.

The systems should have capability of multiple hardware and software partitioning so as to distribute the CPU, Memory and I/O resources within the server for various applications and create fault isolated logical servers. There should be features like online and dynamic creation of partitions. The hardware partitioning, which gives better flexibility and highest level of fault isolation is also recommended.

The Operating System should support virtualization by allocating resources to various applications, without any performance degradation. Open Source Solaris is the best choice for this kind of functionality.

Virtual Machine is another technique of Virtualization where in a software layer sits between the Operating system and the hardware, enabling users to run different kinds of Operating systems and application in isolation from each other on the same set of hardware. Vmware, Xen hypervisor, Sun's Logical Domaining and HPs Virtualization on Integrity Virtual machines uses this approach.

Storage Virtualization should also be kept in mind while designing SDCs. Storage Virtualization ensures that various kinds of Storages from heterogeneous vendors can be viewed as storage pool of resources. The Virtualization software ensures that the relevant data on the storage resources be available to the users transparent of its location.

Consolidation is another technique, which should be strongly considered while designing SDCs. Consolidation ensures efficient use of resources, increasing System Utilization, lowering Total Cost of Ownership (TCO) and improving profitability. Consolidating many small servers into a single Enterprise class server ensures easy management, better serviceability and higher efficiency.

SWaP

The New metric of Space, Watts and Performance is getting popular and addresses the practical challenges of Today's data center. SWaP should be strongly considered while designing the SDCs. Evaluating a new server for your data center is no longer simply a matter of measuring raw performance. With today's increasing demands, you also need to consider how much power, air conditioning and space a server consumes. While traditional metrics are good for calculating throughput, they don't consider these new power and space demands in the equation. (http://www.sun.com/servers/coolthreads/swap/).

Swap is defined as

$$SWaP = \frac{Performance}{Space \times Power\ Consumption}$$

Let us understand SWaP using an example, let there be two servers, Server A and B, with a similar performance metric, as an example similar Spec_int_rate2006 benchmark numbers. Spec_int_rate2006 are most common benchmark statistics available from independent third party bodies and are most widely used performance measurement benchmarks used by the various enterprises across the world. Let us assume that Server A and B both have a spec_int_rate2006 of 85.

Let us assume, that Server A takes two rack unit space and Server B takes four rack unit space, and Server A consumes 300W of power compared to Server B which consumes 800 Watts, which server should we select?

The answer is in spite of both servers having same performance numbers, one should select Server A, which takes lesser real estate space and consumes lesser power. This example is shown in table 1.

It is suggested to consider Space and Power along with Performance (SWaP), while deciding on the servers for the SDC design.

Table 1: SWaP Example

	Server A	**Server B**	**Server A Difference**
Performance	Spec_int_rate2006 = 85	Spec_int_rate2006 = 85	Same
Space	2 RU	4 RU	x 2 Smaller
Power	300 Watts	800 Watts	x 2.7 Times less Power
SWaP	0.141666667	0.0265625	5.3 times Better

Modularity

As defined in wikipedia (http://en.wikipedia.org/wiki/) Modularity **i**n computer science, Modularity (programming) is the property of computer programs that measures the extent to which they have been composed out of separate parts called modules. In design, modularity refers to development of a complex product (or process) from smaller subsystems that can be designed independently.

While designing the SDCs, Modularity and Flexibility should be kept in mind. The selection of systems, Storage, networking and other components should be based on flexible and modular design. It should be possible to modularly grow the Server or the Storage, by just adding resources like CPU, Memory or disks. The possible upgrades in future should involve using the existing system with forklift upgrades. This will ensure a better Return on Investment.

Open Systems

Open systems (http://en.wikipedia.org/wiki/Open_system_%28computing%29) are computer systems that provide some combination of interoperability, portability, and open software standards. (It can also mean

systems configured to allow unrestricted access by people and/or other computers; this article only discusses the first meaning.)

Although computer users today are used to a high degree of both hardware and software interoperability, prior to the year 2000 open systems could be promoted by Unix vendors as a significant competitive advantage. IBM and other companies resisted the trend for decades, exemplified by a now-famous warning in 1991 by an IBM account executive that one should be "careful about getting locked into open systems". (http://en.wikipedia.org/wiki/Open_system_%28computing%29).

However, in the first part of the 21st century many of these same legacy system vendors, particularly IBM and Hewlett-Packard, began to adopt Linux as part of their overall sales strategy, with "open source" marketed as trumping "open system". Consequently an IBM mainframe with open source Linux on zSeries is marketed as being more of an open system than servers using closed-source Microsoft Windows—or even those using Unix, despite its open systems heritage. In response, more companies are opening the source code to their products, with a notable example being Sun Microsystems and their creation of the Openoffice.org and OpenSolaris projects, based on their formerly closed-source Star Office and Solaris software products. (http://en.wikipedia.org/wiki/Open_system_%28computing%29).

While designing and architecting SDCs, one should ensure application of Open Source and Open Standards. The design should accommodate and interoperate with variety of heterogeneous systems. The foundation of the SDC should be based on open source and open standards based operating environment like OpenSolaris. There are features in OpenSolaris which are unmatchable and which are a must to be implemented in the SDCs.

Security

Security is a key parameter for any enterprise and computer security in the State Data Centers is one of the major design criteria.

While designing the SDCs, one should ensure that the Security should be Systemic, it should be built-in the design of a State Data Center and not bolt on. Every layer of the Architectural stack should have Security bundled into it and following parameters should be ensured while designing the SDCs. (Ahuja A, 2007), (Brunette G, 2006)

- Multi Tier Architecture separated with Secure Firewalls and Intrusion detection and prevention systems
- Antivirus Software's, along with encryption techniques should be deployed.
- Each System should be Secured and hardened individually
- Secure Operating Environment, like the open source Solaris Operating System, as the foundation of secure design
- Shared Secure Application Access using Identity based access should be used
- Secure Presentation layer using Identity Manager and Portal should be used
- Secure Desktop Access using Ultra Thin Clients should be used

3. Concluding Remarks

State Data Centers are the backbone of Government of India's e-Governance initiative; of making IT enabled Citizen Services available to the common man. While designing, architecting and implementing the State Data Centers following design parameters should be carefully examined and considered as the top design parameters. The selection of Servers, Storage and Network infrastructure components should be based on the top ten parameters which are – Reliability, Availability, Serviceability, Scalability,

Manageability, Virtualization and Consolidation, SWaP (Space, Watts/Power and Performance), Modularity, Open Systems, and Security

References

1. Gupta S.K, Ahuja A, Bhattacharya J, Singh S.P (2007), Virtualization and Consolidation: Design criteria for Government Data Center, Compendium for the 10'Th National Conference on e-Governance, 29-35.
2. Ahuja A, (2007), Security a Key Design Parameter for State Data Centers, Compendium for the 10'Th National Conference on e-Governance, 58-64.
3. Brunette G (2006), Towards Systemically Secure IT Architectures, *Sun Blueprints*
4. Ahuja A (2006), State Data Center: Dynamic Design and Components, Proceedings of 4th International Conference on e-Governance, *Technology in Government*, 46-57.
5. Vicor Reliability Engineering, Scott Speaks Paper, Reliability and MTBF Overview, available at http://www.vicr.com/documents/quality/Rel_MTBF.pdf, accessed June 2007
6. URL wikipedia definitions, http://en.wikipedia.org/wiki/ accessed June 2007
7. URL Calculating SwaP, http://www.sun.com/servers/coolthreads/swap/ accessed May 2007
8. URL Open Systems, http://en.wikipedia.org/wiki/Open_system_%28computing%29 accessed June 2007

About the Author

Ajay Ahuja works as Senior IT Architect and Technologist with Sun Microsystems India Pvt. Ltd. He has Post Graduate qualifications in Engineering and Management and has over 18 years of IT industry experience in various capacities. He has been involved in Technology enablement initiatives along with designing and architecting solutions for various enterprises across Government, Education and Defence segments. He has participated and presented at many National and International conferences on e-Governance.