DISTRIBUTED DATABASES WITH ACCESS MECHANISMS , SECURITY, CONSISTENCY AND REDUNDANCY

Marin Lungu, Dan – Ovidiu Andrei, Gabriel Toma – Tumbar, Lucian- Florentin Barbulescu

University of Craiova, Faculty of Automation, Computers and Electronics,
Department of Computer Systems and Communication, Romania

Abstract: In the modern organization of a database, the logical functions accomplished by servers, storage units and connecting networks are separated from the physical units that implement these functionalities (the allocation of the physical components is being done dynamically, in a transparent way from the point if view of the user). When taking into consideration the data storage requirements, a good performance and allocation of the resources it becomes necessary to envision an abstraction from the application point of view on the storage units. The correspondence is no longer a direct one-to-one mapping between an application requirements and a physical storage; the data exact location becomes irrelevant from an application point of view. It is the consequences of such a change in perspective that we are investigating in the presented article.

Keywords: Distributed Databases, Direct Attached Storage, Storage Area Network, Network Attached Storage, Databases' Security.

1. INTRODUCTION

The exponential growth of data accumulation within an organization lead to the development of new technologies and approaches for data storage. Among these, specific ones deal with new requirements in terms of information security, fast and concurrent access to data and cost reduction. To address part of these needs, physical storage units such as hard-drives, optical units have become virtual storage devices for data, being dynamically allocated to applications as a function or necessity.

2. DATABASES' SECURITY: UP-TO-DAY SOLUTIONS

Among the various solutions offered by companies for safe storage and continuous access to data we will describe the following: NAS (Network Attached Storage), SAN (Storage Area Network), DAS (Direct Attached Storage) and Backup. Complex solutions provided are mostly based on standard software and hardware components as opposed to proprietary ones being thus flexible and able to include equipments and software from various sources and producers. An important requirement is scalability in terms of making possible an increase in storage capacity as a function of budget. As an example we mention the following solutions: DELL PowerEdge 1800 – 3750 on Intel XEON EMT64 and Intel Itanium 2-64bit platforms, as well as IBM x-Series and e-Series systems on the same platforms.

State of the art solutions accomplish the traditional functionality of file systems as well as that of back-up servers which leads to a significant increase in bandwidth required to access data through SCSI (Small Computer System Interface) controllers or Fiber Channel.

The flexibility and scalability required from the new storage platforms is accomplished by using modular extension units this technique leading to ensuring storage capacities of over 16TB. The modular architecture permits designing and building a base unit to which new components can be added during functioning. In this strategy switches have an essential role. Furthermore, all components are redundant, being replaceable while in use (hot-

swappable) and include: power sources, hard-drives, RAID controllers (Redundant Array of Independent Disks), I/O modules. The mentioned devices are NEBS certified (Network Equipment Building Standards), ETSI certified (European Telecommunications Standards Institute) as well as MIL-STD-810F certified (an American Army Standard).

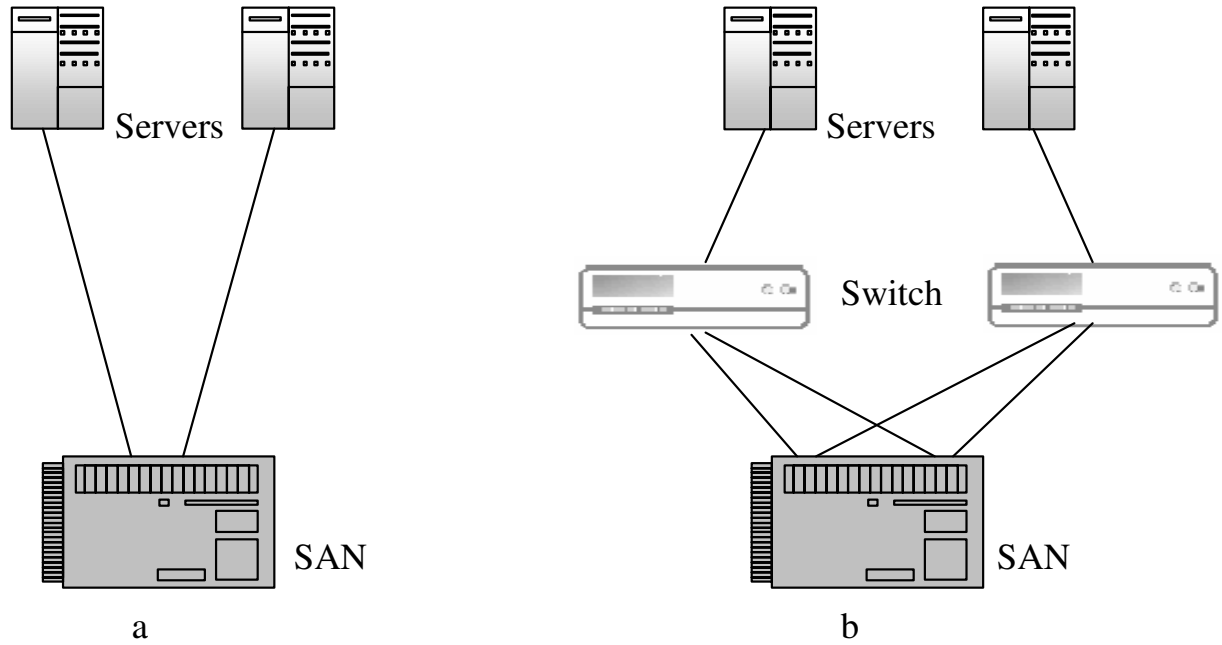

Fig. 1. (a) SAN without switches and (b) SAN with switches

One such architecture is the one presented in Figure 1(b). While analyzing the proposed architecture it is important to consider:

- The requirements of ports for an individual network node
- Grouping of users on nodes with which they communicate most often on the same segment
- Techniques for avoiding network bottleneck by using high speed up-links
- Iterative distribution of users belonging to different segments to nodes with target of using all nodes less that 1/3 of maximum capacity

3. STORAGE SYSTEM SECURITY

One of the most important implications of information security is risk management. One way to understand the risk in the IT domain is as a trade-off/report between two components: the necessity of having access to information at any moment and the involved exposure of the same information to outside agents. The risk is thus generated by the need of information availability.

As the Internet evolved and expanded to did the implications for information security and risk management. The technologies used for the transmission and storage of information grew in complexity while in the same time an entire community of malicious users emerged in the IT word.

Ensuring physical and logical security of computer networks imposed applying protection techniques for the information systems from attacks generated both from the outside and the inside of an organization. In the hardware and software that compose an information system one of the most vulnerable component is the data storage.

Preserving the security of private information is a very important component of any computer network and must be carefully analyzed for several reasons:

- As personal computers can be connected to large networks from homes a variety of activities can be accomplished remotely by individual persons. We must consider the type of data that somebody can have access to read, the persons with which they can communicate, the programms to which they have access to run, etc.
- More and more information stored in separate files becomes available at the same time, making large correlations possible. These correlations may affect the private character of a lot of data that were safely anonimous taken individually.
- Information is vulnerable to an attack in any point of the network, from its input into the network until the point of its final destination. However, information is particularly susceptible to an attack when it passes network nodes and communication environments.

3.1 Vulnerability: tests, identification and analysis

Until not a long time ago the architecture of storage systems was designed so as to have a single access point between the storage device and the application machine. The user, in turn was ensured access to the application machine through a username and a password. In these settings, security as ensured by physical, logical isolation or both.

However, as the separation between data storage and user application involved more layers and becomes more complex the demands from security mechanisms increases as well. Our recommendation is for vulnerabilities analysis and penetration tests to be done periodically, taking into account any changes and vulnerability points that appear over time.

These tests can be done either internally inside an organization (but not by the authors of an application) or by an exterior consultant. The second alternative ensures the objectivity of the entire assessment but also involves following certain guidelines in selecting the security consultant. Some of the aspects that need to be taken into account are:

- The methods employed in the security assessment should not imply only an entirely automatic scan of the application but also a detailed analysis of the vulnerabilities and a set of recommendations for solving the problem.
- The methods and procedures used by the security consultant should respect international standards for testing IT systems:
    - National Institute of Standards and Technology – NIST
    - Open Source Security Testing Methodology – OSSTM
    - Open Information Systems Security Group
    - ISSAF
    - Information Systems Audit and Control Association – ISACA

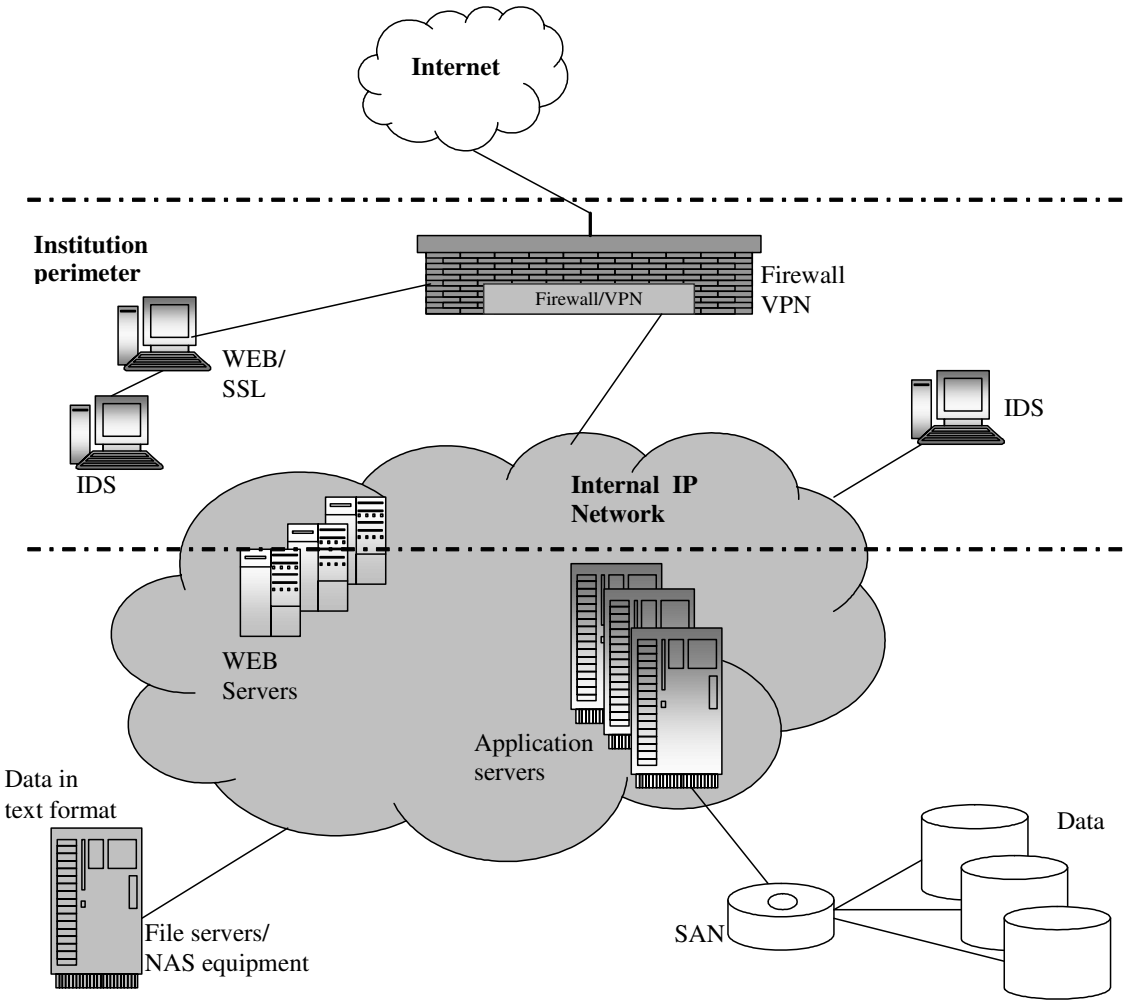

Fig. 2. Traditional architecture with security elements for a computer network

Today's storage systems have grown to include complex topologies and this fact has changed the way the data is accessed and used. The data providers no longer need to be located on a single host within the local LAN, instead the data can be stored on many host systems or special SAN (Storage Area Network) located through the world .

This extended data availability also introduced specific storage network risks.

The attacks targeted at the storage systems can be either direct, when an intruder tries to gain unauthorized access to the systems or it can be in the form of DoS attack (Denial of Service) which renders the systems unusable for the rest of the users

The security policies and procedures, which address this specific area of information security must include the following security measures:

- Network resource integrity and availability must be enforced and thsi leads to HA (High Availability) and failover design in order to reduce the network downtime to an acceptable level . In addition to the hardware resource availability, the security policies must also provide means to reduce the risk of unauthorized access, data tampering and data integrity violation
- The data confidentiality must be also enforced and these policies and procedures will regulate the process of aquiring and manipulation of personal data.

The attack threat has become difficult to handle because of the growing attack complexity, the high availability of automated attack tools and the increasing computing power of today's computers.

The attackers that we face today are either outsiders who try to gain access to remote systems or they are insiders, who are trying to escalate the privileges they already own. No matter which case, their goal is to gain access to confidential data and to be able to control the remote systems which can be used in further criminal actions.

In addition to the direct attacks, there is another type of attacks, which are equally destructive: DoS attacks. The goal is to render a system or service unusable thus preventing legitimate users to access it. It can be initiated from a single host (DoS) or it can be launched simultaneously from many compromised systems (DDoS – Distributed Denial of Service) .

A DoS or DDoS attack is carried out by overloading the target system or service so that no more legitimate request can be processed. The target resources that can be affected include: bandwidth, protocol implementations (eg. number of open TCP connections), CPU or memory. The countermeasures must include policies and procedures to limit the resource access for every user. No single user or service can take over the whole system resources

3.2 Vulnerability assessment

The main causes that lead to vulnerabilities in the distributed applications are:

- Business strategies to release early and cheap also known as fast time-to-market.
- The user's acceptance of software bugs and patches
- The lack of internal software QA and poor testing procedures (if existent!)
- Poor security and secure coding awareness among the software development teams
- Massive outsourcing of software projects without proper QA and testing of the delivered code.

One common vulnerability of distributed applications is the arbitrary code injection. This vulnerability affects most of the applications that rely on a code interpreter such as PERL or PHP.

Every application provides a way to input data from the user. Exploiting this vulnerability takes advantage of improper input data sent to the application and it permits the attacker to execute arbitrary calls such as:

- Operating system calls
- Shell commands issued within the security context of the application server
- Database calls also known as SQL injection

The countermeasures must include:

- Efficient patching and security updates management for the application server
- Defining the external calls which are permitted (starting with a deny all default policy)
- Proper input sanitization in order to block invalid data sent by the user

The data available to an application can be either static (data-at-rest) or in transit (data-in-flight) .

The application's security policy must handle all data, regardless of it's state and it must provide confidentiality, integrity and availability. Stationary data includes data stored on servers, SANs, backup tapes and any other storage medium and tranzit data represents the communication data within the LAN or WAN.

The high-risk components of a storage system are:

- Host servers: The most effective countermeasure for host servers' attack is proper access control. Correct network

segmentation and server cloaking is good but not enough.

- Access control, also known as authorization represents the process of issuing proper credentials in order to access certain resources. One of the most vulnerable sections of a distributed application is the administration section. A well written application presents different user profiles, but, once authenticated, a user can escalate his/hers privileges and gain unauthorized access to special sections. An efficient countermeasure for this attack is mapping a clear matrix of user profiles and resource availability. This matrix is a core component of the application security policy and it will be used to test privilege escalation
- Storage media. This component's vulnerability is related to the system's physical security. Unauthorized access to server rooms, storage facilities can lead to security breaches.
- Local network and the Internet. The data communication is vulnerable to sniffing and man-in-the-middle attacks. Sometimes it is impossible to impose ACLs on all the network devices (eg. routers) that will carry the data. In this case the data must be sent trough secured tunnels and encrypted using either symmetric algorithms (DES, 3DES, CAST, AES, Blowfish) or asymmetric algorithms (RSA, DSA, ECDSA).

The user's interface to an application allows him/her to send requests and parameters to the backend application server. This attack vector can be used to send malicious data, which can crash the application, or it can exploit the way the application handles the user input data internally.

In order to cope with this vulnerability, there are best practices, which impose that:

- The application should perform a syntactic and semantic check on all user supplied data.
- All input validation should be done server-side because client side validation can be easily bypassed.

Storage systems security best practices recommend that that all application's processes and sections should be covered by a clear security policy that will be imposed trough well defined security procedures.

It is mandatory that the security controls defined within the security procedures should include at least: antivirus, firewall, proxy server, IDS (intrusion detection system).

Using username and passwords in the authentication process is only the beginning of application security. Authorization should be implemented trough a series of well-defined roles and user profiles.
Data transport services should operate within encrypted tunnels based on technologies such as SSL and VPN.

Application configuration and user credentials should be stored in secure locations because a security breach on this level can compromise the whole system.

Application security is a process, not a product. It has to include vendor specific vulnerability reports and security updates reported by well-known organizations such as SANS, CERT.

REFERENCES

Koziol J. (2003). Intrusion Detection with Snort

Manual PostgreSQL, http://www.postgresql.org/docs/manuals/

Rehman R (2004). Intrusion Detection with SNORT: Advanced IDS Techniques Using SNORT, Apache, MySQL, PHP, and ACID

Slackware Linux, www.slackware.com

Snort, www.snort.org

Strassberg K, G. Rollie and R Gondek (2002). Firewalls: The Complete Reference, Osborne/McGraw-Hill, ISBN: 0072195673

Symantec Security Response, http://www.symantec.com/avcenter/

Weinstabl P. (2004). PostgreSQL, m. CD-ROM, C & l Computer- U. Literaturverlag, Noiembrie 2004

Ziegler R. (2001). Linux Firewalls (2nd Edition), Pearson Education; 2 edition, ISBN: 0735710996