Future Trend: Rain Technology

Dr. Joginder Singh Cheema
HOD, Department of Computer Applications,
Baba Budha College, Bir Sahib, Tarntaran

Abstract: New emerging technology coming to the expansion of the internet named Reliable array of independent nodes. Before this rain technology we can use the cluster technology in which we have number of nodes and it is not easy to maintain the connection of all these nodes but in rain technology we are capable of providing the solution by reducing the number of nodes in the chain linking the client and server in addition to making the current nodes more robust and more autonomous. One implementation done for this was RAIN-Reliable Array Of Independent Nodes developed by the California Institute of Technology, in collaboration with NASA Jet Propulsion Laboratory and the Defense Advanced Research Projects Agency (DARPA).The technology is implemented in a distributed computing architecture, built with inexpensive off-the-shelf components. The RAIN platform involves heterogeneous cluster of nodes linked using many interfaces to networks configured in fault-tolerant topologies.

Keywords: RAIN, NASA, SNOW, RAINWALL

I. INTRODUCTION

RAIN technology originated in a research project at the California Institute of Technology (Caltech), in collaboration with NASA's Jet Propulsion Laboratory and the Defense Advanced Research Projects Agency (DARPA). The name of the original research project was RAIN, which stands for Reliable Array of Independent Nodes. The objective of the RAIN is to recognize and make key building blocks for reliable distributed systems built using reasonably priced off-the-shelf components .RAIN technology also offers the new feature of reinstating an out of order node by a new one thus keeping away from the break in information flow [1][2]. The main purpose of the RAIN project was to identify key software building blocks for creating reliable distributed applications using off-the-shelf hardware. The focus of the research was on high-performance, fault-tolerant and portable clustering technology for space-borne computing. RAIN Technology (Redundant/reliable array of inexpensive/independent nodes) is a heterogeneous collection of nodes called clusters linked through numerous interfaces to networks configured in fault-tolerant topologies. The RAIN technology concentrates on developing high - performance, fault-tolerant, portable clustering technology .RAIN technology was capable to proffer the solution by lessening the number of nodes in the chain connecting the client and server .apart from this it also facilitates in making the current nodes of client-server architecture more robust [3].

II. WHY WE APPLY RAIN TECHNOLOGY?

RAIN technology is implemented to increase fault tolerance in a cluster. The storage clusters can be managed through a centralized management interface. The management software builds a virtual pool of storage devices without requiring the physical presence of network and storage administrators. The RAIN management software automatically detects any new RAIN nodes and allows them to communicate with each other. In case of a node failure, the lost data is replicated among other RAIN nodes in a cluster to avoid immediate replacement of the failed node. RAIN-based grids are more resilient to application workload changes through effective load-balancing features [4].

III. GOALS OF RAIN TECHNOLOGY

The goal of the RAIN project was to identify key software building blocks for creating reliable distributed applications using off-the-shelf hardware.

The focus of the research was on high-performance, fault-tolerant and portable clustering technology for space-borne computing. Two important assumptions were made, and these two assumptions reflect the differentiations between RAIN and a number of existing solutions both in the industry and in academia[5]:

1. The most general share-nothing model is assumed. There is no shared storage accessible from all computing nodes. The only way for the computing nodes to share state is to communicate via a network.
2. The distributed application is not an isolated system. The distributed protocols interact closely with existing networking protocols so that a RAIN cluster is able to interact with the environment.

In short, the RAIN project intended to marry distributed computing with networking protocols. It became obvious that RAIN technology was well-suited for Internet applications. During the RAIN project, key components were built to fulfill this vision.

IV. ADVANTAGES OF RAIN TECHNOLOGY

[6][7] RAIN technology offers various benefits as listed below:

Fault tolerance: RAIN achieves fault tolerance through software implementation. The system tolerates multiple node, link, and switch failures, with no single point of failure .A RAIN cluster is a true distributed computing system that is durable to faults, it works on the principle of graceful degradation[8][9].

Simple to deploy and manage: It is very easy to deploy and administer a RAIN cluster. RAIN technology deals with the scalability problem on the layer where it is happening, without the need to create additional layers in the front. The management software allows the user to monitor and configure the entire cluster by connecting to any one of the nodes. Open and portable: The technology used is open and highly portable. It is compatible with a variety of hardware and software environments. Currently it has been ported to Solaris, NT and Linux.

Supports for heterogeneous environment: It supports a heterogeneous environment as well, where the cluster can consist of nodes of different operating systems with different configurations.

No distance limitation: There is any distance restriction to RAIN technology. It allows clusters of geographically distributed nodes. It can work with many different Internet applications.

Availability: Another advantage of RAIN is its incessant availability. Example as in case of Rainwall ,it detects failures in software and hardware components in real time, shifting traffic from failing gateways to functioning ones without interrupting existing connections.

Scalability: RAIN technology is scalable. There is no limit on the size of a RAIN cluster example Rainwall is scalable to any number of Internet firewall gateways and allows the addition of new gateways into the cluster without service interruption

Load Balancing and Performance: New nodes can be added into the cluster on the spot to take part in load sharing, without deteriorating the network performance as in case of Rainwall. Rainwall keeps track of the total traffic going into each node. When a disproportion is sensed, in the network traffic, it moves one or more of the virtual IPs on the more heavily-loaded node to the more lightly-loaded node. Also new nodes can be added into the cluster to participate in load sharing, without taking down the cluster.

V. GUIDELINES OF RAIN TECHNOLOGY

[10] [11][12] The guiding concepts that shaped the architecture are as follows:

1. Network Applications: The architecture goals for clustering data network applications are different from clustering data storage applications. Similar goals apply in the telecom environment that provides the Internet backbone infrastructure, due to the nature of applications and services being clustered.

2. Shared-Nothing: The shared-storage cluster is the most widely used for database and application servers that store persistent data on disks. This type of cluster typically focuses on the availability of the database or application service, rather than performance. Recovery from failover is generally slow, because restoring application access to disk-based data takes minutes or longer, not seconds. Telecom servers deployed at the edge of the network are often diskless, keeping data in memory for performance reasons, and tolerate low failover time. Therefore, a new type of share-nothing cluster with rapid failure detection and recovery is required. The only way for the shared-nothing cluster to share is to communicate via the network.

3. Scalability: While the high-availability cluster focuses on recovery from unplanned and planned downtimes, this new type of cluster must also be able to maximize I/O performance by load balancing across multiple computing nodes. Linear scalability with network throughput is important. In order to maximize the total throughput, load load-balancing decisions must be made dynamically by measuring the current capacity of each computing node in real-time. Static hashing does not guarantee an even distribution of traffic.

4. Peer-to-Peer: A dispatcher-based, master-slave cluster architecture suffers from scalability by introducing a potential bottleneck. A peer-to-peer cluster architecture is more suitable for latency-sensitive data network applications processing short-lived sessions. Hybrid architecture should be considered to offset the need for more control over resource management. For example, a cluster can assign multiple authoritative computing nodes that process traffic in the round-robin order for each network interface that is clustered to reduce the overhead of traffic forwarding.

VI. ARCHITECTURE OF RAIN TECHNOLOGY

Rain technology is an open architecture approach to storage, which uses inexpensive computing hardware with extremely intelligent management software to make it reliable and efficient. The components of RAIN configuration run in parallel with operating system and network protocols. The fault tolerance is provided by the management software used which is similar to fault tolerance provided by expensive hardware devices [13].

Redundant Array of Inexpensive Nodes surpasses traditional storage architecture by offering data-storage and protection systems that are more distributed, shareable and scalable. A new storage system architecture called Redundant Array of Inexpensive Nodes (RAIN) surpasses this traditional storage architecture by offering data-storage and protection systems that are more distributed, shareable and scalable. RAIN systems also are less expensive than traditional systems. RAIN is an open architecture approach that combines standard, off-the-shelf computing and networking hardware with highly intelligent management software. This combination lets a host of storage and data-protection applications be cost-effectively deployed across a grid of devices that are highly available and self-healing [14][15].

RAIN-based storage and protection systems consist of:

1. RAIN nodes: Data is stored and protected reliably among multiple RAIN nodes instead of within a single storage subsystem with its own redundant power, cooling and hot-swap disk-drive hardware.

2. IP-based internetworking: RAIN nodes are physically interconnected using standard IP-based LANs, metropolitan-area networks (MAN) and/or WANs. This lets administrators create an integrated storage and protection grid of RAIN nodes across multiple data centers. With MAN and WAN connectivity, RAIN nodes can protect local data while offering off-site protection for data created at other data centers.

3. RAIN management software: This software lets RAIN nodes continuously communicate their assets, capacity, performance and health among themselves. RAIN management software automatically can detect the presence of new RAIN nodes on a new network, and these nodes are self-configuring.

4. Information life-cycle management software: This software replaces traditional snapshot, back-up and mirroring data-management tools with innovative virtualization, compression, versioning, encryption, self-healing integrity checking and correcting, retention and replication algorithms. Information life-cycle management software increases the overall reliability of lower-cost SATA disk drives by replicating data among multiple RAIN nodes.

VII. WORKING OF RAIN TECHNOLOGY

A grid of RAIN nodes also can adapt to changing application workloads by load-balancing data across nodes based on utilization or storage capacity. In a RAIN-based storage system, each RAIN node regularly checks all its own files. The combination of hundreds of RAIN nodes forms a powerful parallel data-management grid - one that is much more powerful than today's independent protection architectures. When file corruption is detected, the associated RAIN node initiates a replication request to all other RAIN nodes, which verify their own replicas and work collectively to replace the defective file. Grids of RAIN nodes will replace existing isolated data-storage systems. Low-cost, high-performance disk drives, CPUs and IP networking make this evolution possible. In addition, businesses are demanding simplified, lower-cost, site disaster-recovery systems and faster and more reliable back-up and restore processes. By executing information life-cycle management applications across hundreds of powerful, internetworked storage and compute RAIN nodes, RAIN systems will deliver unprecedented long-term data availability, cost-effective and rapid site disaster recovery, and automated onsite and offsite data back-up protection.

VIII. APPLICATIONS RAIN TECHNOLOGY

Below listed are some of the applications of RAIN[16]:
Video server (RAIN Video), a web server (SNOW), and a distributed check pointing system (RAIN Check) etc. These applications indicate quick failover response, little overhead and near-linear scalability of the Raincore protocols:

SNOW (Strong Network of Web servers): The first application, called SNOW, is a scalable Web server cluster that was developed as part of the RAIN project.

RAINVideo: RAINVideo application is a collection of videos written and encoded to all n nodes in the system with distributed store operations.

RainWall: RainWall is a commercial solution that provides the fault-tolerant and scalable firewall cluster.

RAINCheck: Raincheck is a Distributed Check pointing Mechanism, it implements a checkpoint and rollback/recovery mechanism on the RAIN platform based on the distributed store and retrieve operations.

Distributed Check pointing Mechanism: It is a checkpoint and rollback/recovery mechanism on the RAIN platform based on the distributed store and retrieve operations.

IX. CONCLUSION

The purpose of the RAIN project has been to pave a way to fault-management, communication, and storage in a distributed environment. RAIN technology has been exceedingly advantageous in facilitating resolution of high-availability and load-balancing problems. It is applicable to an extensive range of networking applications, such as firewalls, web servers, IP telephony gateways, application routers, etc. It has very useful in the development of a fully functional distributed computing system. RAIN allows for the grouping of an unlimited number of nodes, which can then function as one single giant node, sharing load or taking over if one or more of the nodes ceases to function correctly.

REFERENCES

[1] http://krititech.in/wordpress/?p=244
[2]http://www.thefreelibrary.com/Latest+network+management+products-a062893683
[3] http://paradise.caltech.edu/papers/etr029.pdf
[4] http://www.seminarprojects.com/Thread-rain-random-array-of-independent-nodes
[5] http://searchdatacenter.techtarget.com/definition/RAIN
[6]http://en.wikipedia.org/wiki/Redundant_Array_of_Inexpensive_Nodes
[7] https://www.networkworld.com/article/2330258/tech-primers/rain-architecture-scales-storage.html
[8] L.E. Moser, P.M. Melliar-Smith, D.A. Agarwal, R.K. Budhia, C.A. Lingley-Papadopoulos, "Totem: A Fault-Tolerant Multicast Group Communication System", *Comm. ACM*, vol. 39, no. 4, pp. 5463, 1996.
[9] "NASA-Funded Software Aids Reliability", *Network World*, no. 51, Dec. 1999.
[10] V. Bohossian et al., "Computing in the RAIN: Reliable Array of Independent Nodes," IEEE Trans. Parallel and Distributed Systems, vol. 12, no. 2, Feb. 2001, pp. 99-114.
[11] T.D. Chandra, S. Toueg, "Unreliable Failure Detectors for Reliable Distributed Systems", J. ACM, vol. 43, no. 2, pp. 225267, 1996.
[12] P.M. Chen, E.K. Lee, G.A. Gibson, R.H. Katz, D.A. Patterson, "RaidHigh-Performance Reliable Secondary Storage", ACM Computing Surveys, vol. 26, no. 2, pp. 145185, 1994.
[13] J. Cotrim Arabe, A. Beguelin, B. Lowekamp, E. Seligman, M. Starkey, P. Stephan, "Dome: Parallel Programming in a Distributed Computing Environment", Proc. IEEE Symp. Parallel and Distributed Processing, pp. 218224, 1996.
[14] A. Singhai, S.-B. Lim, S.R. Radia, "The SunSCALR Framework for Internet Servers", *Proc. IEEE 28th Int'l Symp. Fault-Tolerant Computing*, 1998.
[15] M. Snir, S. Otto, S. Huss-Lederman, D. Walker, J. Dongarra, "MPI: The Complete Reference", 1995.
[16] H.-M. Sun, S.-P. Shieh, "Optimal Information Dispersal for Increasing the Reliability of a Distributed Service", *IEEE Trans. Reliability*, vol. 46, no. 4, pp. 462472, 1997.