A failure detection method of remote disaster recovery and backup system

Min Liang*, *Fei* Gao, and *Min* Shi

National Key Laboratory of Astronautic Dynamics, Xi'an Shaanxi, China

Abstract. Failure detection is one of the basic functions of building a reliable disaster recovery backup system. Aiming at the application-level disaster recovery backup failure detection problem, this paper analyzes the remote disaster recovery center architecture and failure detection hierarchy, and predicts the arrival time of cross-domain heartbeat information through the back propagation neural network based on particle swarm optimization (PSO-BP). When the actual timeout is reached, the active Auxiliary Detection (AD) is used to improve the correctness of failure detection, and finally the effectiveness of method PSO-BP-AD is verified through simulation.

Keywords: Disaster recovery, Failure detection, Back propagation neural network, Particle swarm optimization.

1 Introduction

With the continuous development of computer technology, the degree of informatization in all walks of life continues to increase, and more and more businesses rely on computer systems. Computer systems and users are interconnected through the network, which significantly improves work efficiency and reduces production costs. Computer systems in the fields of finance, securities, insurance, power, aerospace, etc., have become an important national infrastructure and need to operate continuously to provide safe and reliable services to the outside world. In recent years, the business continuity requirements of computer systems have become increasingly prominent. Factors such as natural disasters, social unrest, network failures, equipment failures, and human operation errors affect system security and business continuity to varying degrees. Even a short-term system shutdown will cause huge losses.

Disaster recovery refers to the ability to restore computer system functions after a disaster, ensuring business reliability and availability. Disaster here refers to all events that cause abnormal shutdown or failure of the system. Disaster recovery can be divided into data level, system level and application level [1]. Data-level disaster recovery aims to protect data and ensure the availability of system data after a disaster. System-level disaster recovery protects the operating environment of the system, such as the operating system and communication network, to ensure that the system software can continue to run or

* Corresponding author: lm7186345@163.com

quickly recover after a disaster occurs. Application-level disaster recovery is based on the data-level and system-level, ensuring that upper-level applications can quickly switch to the backup system during disasters and continue to provide services. This is the current research hotspot and is also the background studied in this article.

Failure detection is a prerequisite for disaster recovery and backup, and the combination of the two ensures the continuity of the system's business. The two main indicators of business continuity are the recovery time objective, which indicates how long the system can resume operation after a disaster occurs; the recovery point objective, which indicates the degree of data loss. The ideal requirement is RTO=0, RPO=0. Only by detecting the failure of the main production system in time when a disaster occurs, and enabling the backup system, can the RTO value be minimized. Therefore, comprehensive real-time monitoring of the main production business system and timely detection of the overall failure status is of great significance for the rapid and smooth switch to the backup system and maintaining business continuity.

The purpose of failure detection is to identify the failure of the production system in time and enable the disaster recovery and backup system to take over the business. The existing failure detection strategy uses an adaptive estimation algorithm, usually using history statistical heartbeat information or sliding windows to calculate the expected arrival of the heartbeat information, so as to predict the arrival time of the next heartbeat, while attaching a safety margin to ensure the detection accuracy. The heartbeat contains the status information of the detected system, and its network delay is affected by many factors such as the communication protocol, operating environment, and physical link quality. The failure detector needs to identify the state of the detected system under acceptable time constraints to minimize the probability of misjudgment.

This paper first introduces the current development of failure detection technology, analyzes the remote disaster recovery architecture and hierarchical failure detection structure, and proposes a new cross-domain failure detection method which uses BP neural network based on particle swarm optimization (PSO) (Back propagation neural network, BPNN) to predict the arrival time of heartbeat information, and uses auxiliary active detection to avoid misjudgment and improve the accuracy of detection. Finally, simulation proves the effectiveness of the method.

2 Related rork

Failure detection technology is the basic technology to achieve high reliability of disaster recovery system. It monitors the operating status of the system in real time, detects disasters in time, and sends out warning messages to trigger disaster preparedness emergency measures. In the study of early failure detection methods, the heartbeat message sending interval and timeout value are adjusted to adapt to changes in the network environment, so as to improve the detection accuracy under a certain detection speed constraint. Falai et al. [2] summarized the timeout value prediction methods, and selected five timeout value prediction methods: LAST, MEAN, WINMEAN (N), LPF (b) and ARIMA (p, d, q), respectively. Use different security boundaries (SM_{CI} and SM_{JAC}) and evaluate its performance through experimental methods. The results show that the ARIMA (p, d, q) method has higher prediction accuracy, while the LAST method with lower algorithmic complexity combined with the SM_{JAC} security boundary obtains the fastest detection time and better detection accuracy.

Early researches on QoS are still in the qualitative stage, and cannot accurately describe the QoS that failure detection can meet, and users can't put forward quantitative QoS requirements for failure detection. In response to this situation, Chen et al. [3] proposed a complete set of failure detection QoS evaluation index system, which is a measure of

detection speed and accuracy, and achieved a quantitative evaluation of failure detection capabilities. At the same time, based on the above QoS indicators, the adaptive failure detector NFD-E proposed by Chen et al. can automatically adjust detection parameters according to the QoS requirements given by users or applications, so as to obtain quantitative QoS indicators with minimal detection load. Aiming at the problem of detection speed of NFD-E detector, Bertier et al. [4] proposed a failure detector based on dynamic security boundary, which uses Jacobson's round-trip time (RTT) to estimate the security boundary. It can dynamically change with the changes of the network environment, reduce the timeout value, and increase the detection speed, but the detection error rate is higher. Tomsic et al. [5] proposed a dual-window-based failure detector 2W-FD, which uses two sliding windows of different sizes to store historical heartbeat messages, calculates the timeout value separately, and selects the larger value from the result as the final timeout value captures more heartbeat messages, thereby improving detection accuracy.

In response to multiple types of different QoS problems, Défago et al. [6] proposed an Accrual failure detection model, which can simultaneously meet the different QoS requirements of multiple applications in large-scale distributed systems. Accrual separates the monitoring from the interpretation, and is only responsible for detecting nodes. At the same time, it also outputs a continuous value *slqp(t)* that represents the degree of suspicion of the detected node, instead of directly outputting the binary detection results of believed or suspected [7], but its accuracy and timeliness needs to be improved.

3 Remote disaster recovery failure detection architecture

The purpose of the disaster recovery system is to prevent the information system from causing the system service stop or data loss in case of a disaster. The implementation method is mainly through the maintenance of a backup system equivalent to the production system, and to improve the disaster recovery capacity of the key business system through the geographical dispersion. A common disaster recovery system architecture is shown in Fig.1. The remote disaster recovery system is consistent with the computer architecture of the local production system. Each subsystem module of the main production system has a local backup function, which is high availability (HA). This paper focuses on failure detection between cross-domain active and standby centers. The business center computer system does not exist in isolation on the network, the applications need to interact with external entities (such as users, sensors, etc.), and the communication link is relatively separated from the disaster recovery link.

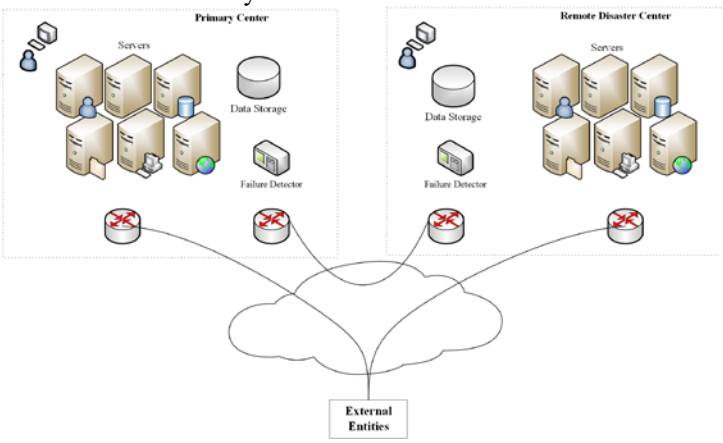

Fig. 1. Disaster recovery system architecture.

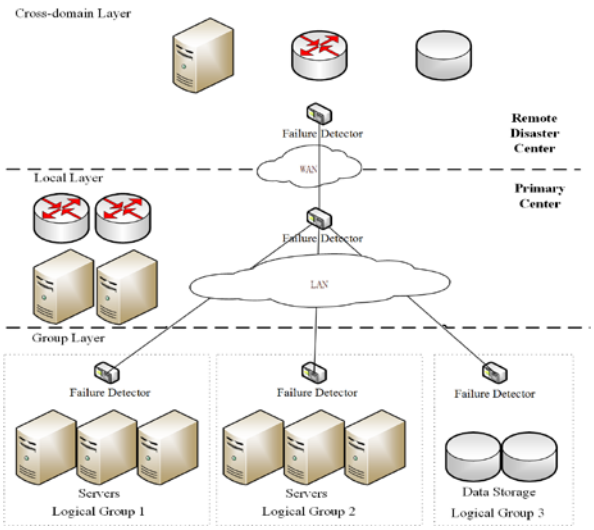

Fig. 2. Failure detection hierarchy.

Disaster failure detection structures can be divided into three levels, as shown in Fig.2.

The first level is the disaster recovery grouping layer, which divides the local system into multiple groups according to standards such as IP subnets, functional modules, or device types, and selects a master node in the group to deploy a failure detection module, which is responsible for detecting other nodes in the group.

The second level is the local disaster recovery layer. The root node summarizes and analyzes the failure detection results of each group to obtain the local production system status information and publish it to the public. The nodes of the first and second layers are connected through the local area network, which is characterized by making full use of LAN low latency, and high bandwidth, increasing the detection of node performance and service quality, in the node business or node failure, using cluster high available technology implement local switching to reduce the business pause time, improve the stability of the main production system business.

The third level is the cross-domain disaster recovery layer, including the main production system and the disaster recovery system. It can generate disaster recovery strategies and global resource views to provide a basis for the implementation of correct disaster recovery decisions. When the main production system fails, the disaster recovery backup can be activated in time to take over the business. Communication interaction entities such as users, sensor terminals, and data sources outside the business center system can be used as assistants for failure detection to improve the effectiveness of detection.

4 Failure detection method

4.1 Failure model

The main production system contains multiple functional modules, which are connected by physical links within and between modules. However, the actual physical link cannot fully guarantee the quality of information transmission, especially cross-domain transmission. According to the behaviour of the system after failure, the failure is usually divided into 3 types: Fail-stop model, Byzantine model and Timing-Failure model [8][9].

When a failure occurs in the Fail-Stop model, the node will stop normal work, and there will be no output. When a failure occurs in the Byzantine model, the node does not completely stop working, but in a random state; the failure detection process on the node does not stop working, but its output is unreliable. This failure is generally caused by logic errors or malicious attacks. In the Timing-Failure model, the process cannot complete the specified service within the specified time. This failure is generally caused by excessive network delay or excessive node load.

In a real disaster recovery scenario, the more common failure model is the Fail-Stop type. However, due to the uncertainty of the network delay and the unreliability of the communication link, it is difficult for the system to distinguish between the failure of the Fail-Stop and the Timing-Failure models, leading to the miscalculation. Failure detection needs to consider this issue.

4.2 Process

The failure detector (FD) provides status information of the main production system, and can issue a failure alarm to the disaster recovery backup system. The failure detector may take a long time to suspect a node that has crashed, or it may mistakenly suspect a normal node (due to loss of information, network delay or interruption). In order to improve the detection accuracy and timeliness, this paper proposes a new failure detection method PSO-BP-AD, which predicts the network delay THB of the heartbeat information through the PSO-BP neural network algorithm; adds a fixed safety time margin α to form a timeout period Δt. According to the relationship between the predicted arrival time and the actual arrival time, set the active detection timing, and ensure the effectiveness of the algorithm through auxiliary detection (AD). Based on the reference to the PSO-BP algorithm [10], the specific steps are formed as follows:

Step 1 Normalize 1000 sets of heartbeat information delay data under each network state and divide them into 800 sets of sample data and 200 sets of test data.

Step 2 Initialize the BP neural network. Determine the neural network structure, neuron processing functions of each layer, and assign initial values to each parameter.

Step 3 Initialize the PSO algorithm. Select particles in group H that meet the initial weights and threshold requirements of the BP neural network randomly and assign initial values to other parameters in the PSO algorithm; the sum of the square of the difference between the positive and expected output value of the test data through the BP neural network is used as a function of the adaptive value of the PSO algorithm.

Step 4 Calculate the fitness value of the particles.

Step 5 Determine the number of iterations of the particle. If the number of iterations of the particle does not reach the set value, update the speed and position of the particle, and increase the number of iterations by 1, and then return the updated particle to step 4 to recalculate the fitness value of the particle; if the number of iterations reaches the set value, stop the iteration and output the optimal position Gb of the group in the PSO algorithm.

Step 6 Obtain the optimal weight and threshold. Assign the optimal position Gb of the group in the PSO algorithm to the BP neural network as the initial weight and threshold.

Step 7 Calculate the error. The BP neural network with the initial weight and threshold optimized by the PSO algorithm is trained on 800 sets of sample data. The sum of squares of the difference between the output value of the BP neural network and the expected value is selected as the error function.

Step 8 Determine that the conditions are met. Compare the error value with the set value. If the error value does not reach the set accuracy, use the error reflection propagation channel to update the weights and thresholds of the BP neural network; if the error value

reaches the set accuracy, output the BP neural network output-heartbeat information delay T_{HB}.

Step 9 Generate timeout time $\Delta t = T_{HB} + \alpha$.

Step 10: Failure detection judgment. If the actual time delay of the heartbeat information $T_{hb} < \Delta t$, the system state is judged according to the content of the heartbeat information; if $T_{hb} > \Delta t$, the heartbeat message was not received during Δt, the auxiliary active detection performed to further verify the system state.

4.3 Active detection

In order to improve the accuracy of failure detection, PULL(Fig.3) active detection is performed with the help of detection assistants to obtain the status of the detected system. Previously, the main production system used PUSH model (Fig.4) to pushed heartbeat information of the failure detector deployed in the disaster recovery system, simultaneously pushed the heartbeat information of the system status to the assistants continuously. When $T_{hb} > \Delta t$, the failure detector does not issue an alarm immediately, but actively learns the status of the main production system from the auxiliary to make sure if the network between the main and standby center is available. If the auxiliary receives the main production system the "normal" heartbeat information indicates that the heartbeat information of the failure detector is timed out due to the network between the active and standby centers, but it does not affect the normal operation of the business. The network failure should be eliminated immediately instead of enabling the disaster recovery system to take over the business. On the contrary, if the auxiliary does not receive or receives the "failed" heartbeat information, it means that there is a problem with the main production system and business network, and the external business is interrupted. A disaster recovery plan should be implemented.

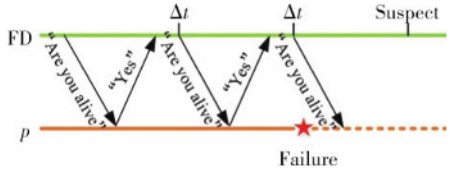

Fig. 3. PULL model.

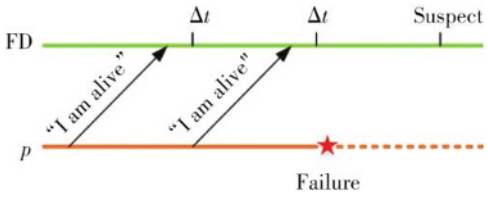

Fig. 4. PUSH model.

5 Simulation analysis

5.1 Simulation parameter setting

The PSO-BP neural network algorithm requires parameter settings [10]. The number of neurons in the input layer of the BP neural network is 2, and the number of neurons in the output layer is 1; the number of hidden layers is 1, The number of hidden layer neurons is

selected by $l = \sqrt{m+n} + a$, and the value is 10; the maximum number of trainings is set to 1000, the error accuracy is 0.01, and the learning rate is 0.5. The particle dimension in the PSO algorithm is taken as N=41, and the learning factor c_1=c_2=1.49; the position and velocity of the particles in the PSO algorithm are limited to the range of [-0.5,0.5]; the inertia weight ω_{start}=0.9,ω_{end}=0.4.

Fixed safety time margin α=500ms. Failure detection interval is 5000ms.

5.2 Simulation comparison

In the same environment, the performance of PSO-BP-AD and Chen failure detection algorithms were tested respectively, as shown in Fig.5 and Fig.6.

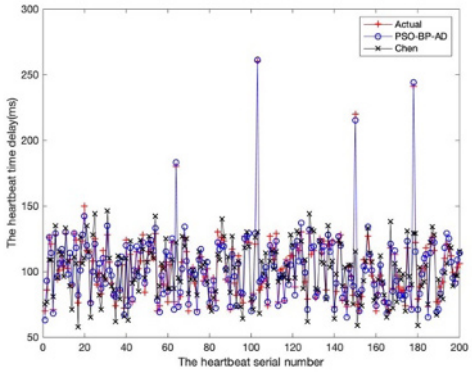

Fig. 5. Comparison of heartbeat time delay.

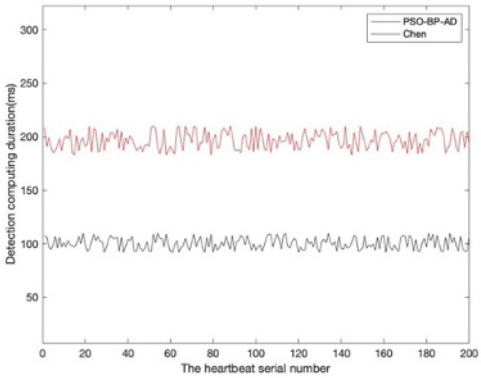

Fig. 6. Comparison of failure detection duration.

The prediction time of using Chen's failure detection algorithm is relatively short, but when the heartbeat information is lost, or the network between the main and standby systems is faulty, Chen's failure detection algorithm will make a misjudgment. The PSO-BP-AD algorithm is more accurate in failure detection and predicts network delays more accurately. It can adapt to network packet loss and link failures, triggers the assistant to actively detect, and eliminates misjudgments caused by network status changes. The PSO-BP-AD algorithm design is effective and can meet the needs of disaster recovery failure detection.

6 Conclusion

In the disaster recovery backup system, it is necessary to detect the collapse of the main production system in real time. Aiming at environmental characteristics such as the uncertainty of the cross-domain connection network between the remote disaster recovery center system and the main production center, a failure detection method PSO-BP-AD is designed. This method uses the PSO-BP algorithm to predict the arrival time of the heartbeat information. When the actual arrival time exceeds the predicted arrival time, the auxiliary detection (AD) method is used to actively verify the status of the main production system and improve the effectiveness of failure detection. The simulation analysis results show that the algorithm design is effective and can meet the needs of failure detection.

References

1. Charilaos Stais, George Xylomenos, Giannis F. Marisa. Sink Controlled Reliable Transport for Disaster Recovery. *PETRA'14*, Greece, ACM, 2014.
2. FALAI L, BONDAVALLI A. Experimental evaluation of the QoS of failure detectors on wide area network. *Int. Conf. on Dependable Systems and Networks*. Yokohama, Japan: IEEE Press, 2005: 624 - 633.
3. CHEN Wei, TOUEG S, AGUILERA M K. On the quality of service of failure detectors *IEEE Transactions on Computers*, 2002, 51(1): 13 - 32.
4. BERTIER M, MARIN O, SENS P. Implementation and performance evaluation of an adaptable failure detector *DSN '02: Int. Conf. on Dependable Systems and Networks*. Washington, DC, USA: IEEE Computer Society, 2002: 354 - 363.
5. TOMSIC A, SENS P, GARCIA J, et al. 2W - FD: A failure detector algorithm with QoS *Int. Conf. on Parallel and Distributed Processing Symposium (IPDPS2015)*. Hyderabad, India: IEEE Press, 2015: 885 - 893.
6. DÉFAGO X, URBAN P, HAYASHIBARA N et al. Definition and specification of accrual failure detectors *Int. Conf. on Dependable Systems and Networks*. Yokohama, Japan: IEEE Computer Society, 2005: 206 - 215.
7. HAYASHIBARA N, DÉFAGO X, YARED R, et al. The Φ accrual failure detector *Int. Conf. on Reliable Distributed Systems*. Florianopolis, Brazil: IEEE Press, 2004: 66 - 78.
8. C. Fetzer, F. Cristian. Fail-awareness in timed asynchronous systems *Int. Conf. on* 15th *Annual ACM Symposium on Principles of Distributed Computing*, 314-321,1996.
9. Bertier, M. Martin, o. Sens. Performance Analysis of a Hierarchical Failure Detector. *Int. Conf. on Dependable Systems and Networks*, 2003:354-364.
10. SHI Wei-guo, LEI he-fen. Algorithm Prediction of Network Delay Using BP Neural Network Based on Particle Swarm Optimization [J]. *Automation & Instrumentation*.2020, 35(07):1-5.