NEW NON-ADAPTIVE DISTRIBUTED SYSTEM-LEVEL DIAGNOSIS METHODS FOR COMPUTER NETWORKS

Hiroshi MASUYAMA
Information and Knowledge Engineering, Tottori University
Koyama-cho Minami 4-101, Tottori, 680-8552, Japan

Koji WATANABE
Graduate School, Tottori University, Tottori, 680-8552, Japan

Keywords: Computer networks, System-level diagnosis, Diagnosability, Test graph

Abstract: A hierarchical non-adaptive diagnosis algorithm is presented for testing total N nodes of computer networks. Since general computer networks can be regarded as an N-nodes complete graph, then for the efficient testing, it is essential that the test process be parallelized to enable simultaneous test of multiple nodes. In order to attain this object, we propose a noble test graph enabling to test as many nodes as possible in a network due to a hierarchical architecture of test processes. The amount of test times is evaluated as the diagnosis latency. Optimal diagnosability t is analyzed under clustered fault distribution. In order to reduce the amount of required test times, two revised approaches are discussed and evaluated.

1 INTRODUCTION

There have been significant theoretical researches in the area of **system-level** diagnosis by which every node receives diagnosis. This system-level diagnosis approach was introduced first by Preparata et al. (F.Preparata et al., 1968) where t-diagnosability was introduced. The t**-diagnosability** is the ability to diagnose a fault situation with t or fewer faults given in the network. This means that every node must be tested by more than t other nodes if a network is said to be t-diagnosable. The problems of fault detection (testing) and fault location (diagnosis) have been mostly studied by using testing networks which is reduced to some test graphs, whose vertices denote the nodes and whose an edge or test link p_i, p_j from node p_i to node p_j indicates that p_i tests p_j (C.Feng et al., 1996) ~ (N.H.Vaidya et al., 1994). Since a general graph contains many vertices, one by one test approach requires significant test time.

The fault model of the network characterizes the outcome of test results. The first model of system diagnosis is introduced as **PMC Model** (F.Preparata et al., 1968). In this model, the outcome of a test performed by a fault-free node is correct and equals fault state of the tested node. On the other hand, the outcome of a test performed by a faulty node is unreliable, that is, arbitrary. Classical system-level diagnosis approaches (F.Preparata et al., 1968), (S.L.Hakimi et al., 1974) have a central observer by which all test results are gathered to make a **syndrome** of the network. In the most of these approaches, a **distributed model** is assumed where each node performs independently its own local diagnosis, that is, performs tests of only its definite subset of nodes. If the choice of the next tests, that is, the subset is known in advance, these test approaches are also called a **non-adaptive** test. The central observer uses the results obtained from all test nodes to determine the fault situation, that is, locates the faults in the network.

On condition that a ring can be judged correctly whether the ring has at most one locatable fault or more than one un-locatable faults, a single loop testing (N.H.Vaidya et al., 1994) of one of **adaptive** diagnosis techniques where the choice of the next tests depends on the results of previous tests and not on a fixed pattern, is developed. There exist considerable presented schemes on the condition that the maximum number of faulty nodes distributed in a network is bounded by a predefined limit, and they have been improved to reduce the diagnosis latency (R.P.Bianchini et al., 1992), (E.P.Duarte Jr et al., 1998). However, since test graphs for general computer networks contains

Masuyama H. and Watanabe K. (2004).
NEW NON-ADAPTIVE DISTRIBUTED SYSTEM-LEVEL DIAGNOSIS METHODS FOR COMPUTER NETWORKS.
In *Proceedings of the First International Conference on E-Business and Telecommunication Networks*, pages 161-168
DOI: 10.5220/0001384401610168
Copyright © SciTePress

many vertices, these adaptive diagnosis techniques require significant overhead, that is complex analysis of the test results.

In this paper, we consider a classical system-level diagnosis algorithm in which only the nodes fail because a faulty communication link can be accommodated by treating as a faulty node. And we present a hierarchical non-adaptive diagnosis algorithm for testing total N nodes of computer networks. Since general computer networks can be regarded as an N-nodes complete graph, then for the efficient testing, it is essential that the test process be parallelized to enable simultaneous test of multiple nodes. In order to attain this object, we propose a regular graph of connectivity-$(t+1)$ with N nodes as test graphs. In this test graph, a self-tested node is placed at a key location in a hierarchical structure, and at first the node tests the adjacent nodes. Only adjacent nodes that passed the test can become new monitors and test their adjacent nodes, and so on. This process is propagated to higher levels of the test graph. At each level, all monitors send the announcements of their own test results " I passed the test " when they received a qualification as a monitor first, and in addition send only the test failed results of their test targets when they finish their tests, back to their monitors by which they are tested first. Each monitor also sends data transferred from its test target back to the monitor by which he is tested first. Then all test results are gathered in a host (that is , a central observer) directly connected the original monitor, and then the host can locate all faults in the network. Optimal diagnosability t is analyzed under clustered fault distribution.

Recently, several diagnosis techniques based on this self-testing (F.J.Meyer et al., 1989) have proposed, and achieved a successful diagnosis of a large number of faults. Though most of drawbacks of self-testing are to require many self-testing, papers (L.Zakrevski et al., 1998), and (H.Masuyama et al., 2001) made the drawbacks light by preparing the limited number of monitors, as shown in our approach. However, their target networks are multi-processor networks consisting of homogeneous nodes connected by bi-directional links. Each node can be viewed as a combination of a router and processor along with associated RAM, bus and I/O circuitry, then they differ from us in target networks.

In non-adaptive or even adaptive tests, since each node must performs a certain number of nodes and report to somewhere in the network, then a traffic problem must be cleared. Therefore, not only the time elapsed for testing all nodes and the time complexity of diagnosis algorithm but also the traffic condition are essential to evaluate diagnosis algorithms. In this paper, **diagnosis latency**, that is, the time elapsed for testing all nodes is evaluated as the total number of test times where each test executes in different time. This time is also called as testing round. In order to reduce the amount of required test times, two revised approaches are discussed and evaluated.

2 ALGORITHMS

In this section, we will discuss three algorithms for constructing our test graph, for obtaining necessary test orders, and for test.

2.1 Test graph

For given N and diagnosability t, we will plan to construct a test graph whose connectivity is over t by the following algorithm:

[Algorithm A]

Step 1: Prepare α hypercubes of dimension β independently, and number to these α hypercubes. Each node in a hypercube corresponds to $(\alpha-1)$ nodes in each different hypercubes.

Step 2: For total β sets of α corresponding nodes, connect α corresponding nodes with a completed graph.

Step 3: Select one node as an original monitor arbitrary from N nodes. Set the edges connected with the original monitor and the adjacent nodes as unidirectional edges and all other edges as bidirectional edges.

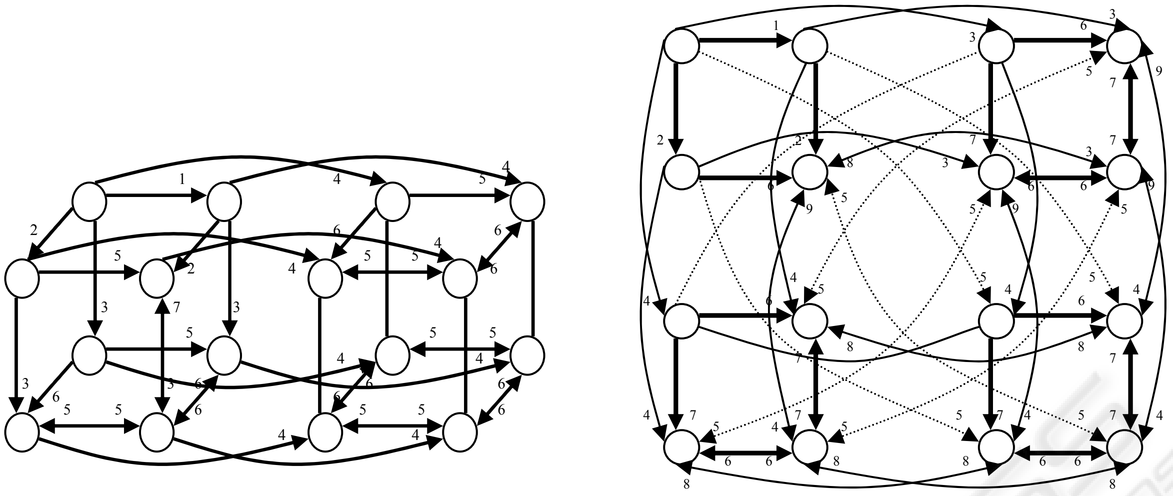
The graph obtained by Algorithm A has $\alpha \cdot 2^{\beta}$ nodes, and the degree of each node is $\alpha+(\beta-1)$. Then, α and β are restricted by given N and t as follows: $N=\alpha \cdot 2^{\beta}$ and $t \leq \alpha+(\beta-2)$. The longest distance d_m from an original monitor is $\beta+1$.

On the strength of algorithm A for constructing test graph, we can give test orders to every adjacent nodes of each node by the following algorithm:

[Algorithm B]

Each node of a β-dimensional hypercube can be indexed 0 to $2^{\beta-1}$, and each of α hypercubes can be numbered 0 to $\alpha-1$. Assume node i is indexed j and hypercube which contain node i is numbered $k(0 \leq k \leq \alpha-1)$. The test orders of each adjacent node of node i are as follows: The adjacent nodes indexed $(j+1)$, $(j+2),\cdots,(j-2),(j-1)$ $(\text{mod}.2^{\beta})$ on hypercube numbered k, the adjacent nodes on hypercubes numbered$(k+1),(k+2),\cdots,(k-2),(k-1)$ $(\text{mod}.\alpha)$.

SCITEPRESS SCIENCE AND TECHNOLOGY PUBLICATIONS



(a) In the case of $\alpha = 1, \beta = 4,$ and $t = 2$

(b) In the case of $\alpha = 4, \beta = 2,$ and $t = 4$

Figure 1: Two test graphs with $N = 16$.

2.2 Test algorithm

On the strength of Algorithms A and B, we can construct a test algorithm for an $N\left(= \alpha \cdot 2^{\beta}\right)$-node network as follows:

[Algorithm C]

First, the monitor tests its adjacent nodes in the test order of the adjacent nodes, and hands a message "faulty node name" to the host if it decides an adjacent node faulty. The monitor hands a qualification as a monitor to its adjacent node if it decides the adjacent node non faulty.

Each node hands first its own test result "I passed the test" to its the first tester when it received a qualification as a monitor. Each node starts testing its adjacent nodes in the test order, and hands a message "faulty node name" to the adjacent node by which it is tested first if it decides its testing adjacent node faulty. It hands a qualification as a monitor to its adjacent node if it decides the adjacent node non faulty. Each node hands messages of "faulty node name" to the adjacent node by which it is tested first if it receives the messages from the adjacent node to which it tested previously.

Then, with Algorithm C all test results can be gathered in a host directly connected the original monitor, and then the host can locate all faults in the network.

Example 1: Figs.1(a) and (b) show two test graphs with N=16 labeled the test orders in the cases of $(\alpha = 1, \beta = 4, t = 2)$ and $(\alpha = 4, \beta = 2, t = 4)$, respectively. Figs.2(a) and (b) show two test graphs with N=32 in the cases of $(\alpha = 2, \beta = 4, t = 4)$ and $(\alpha = 4, \beta = 3, t = 5)$, respectively.

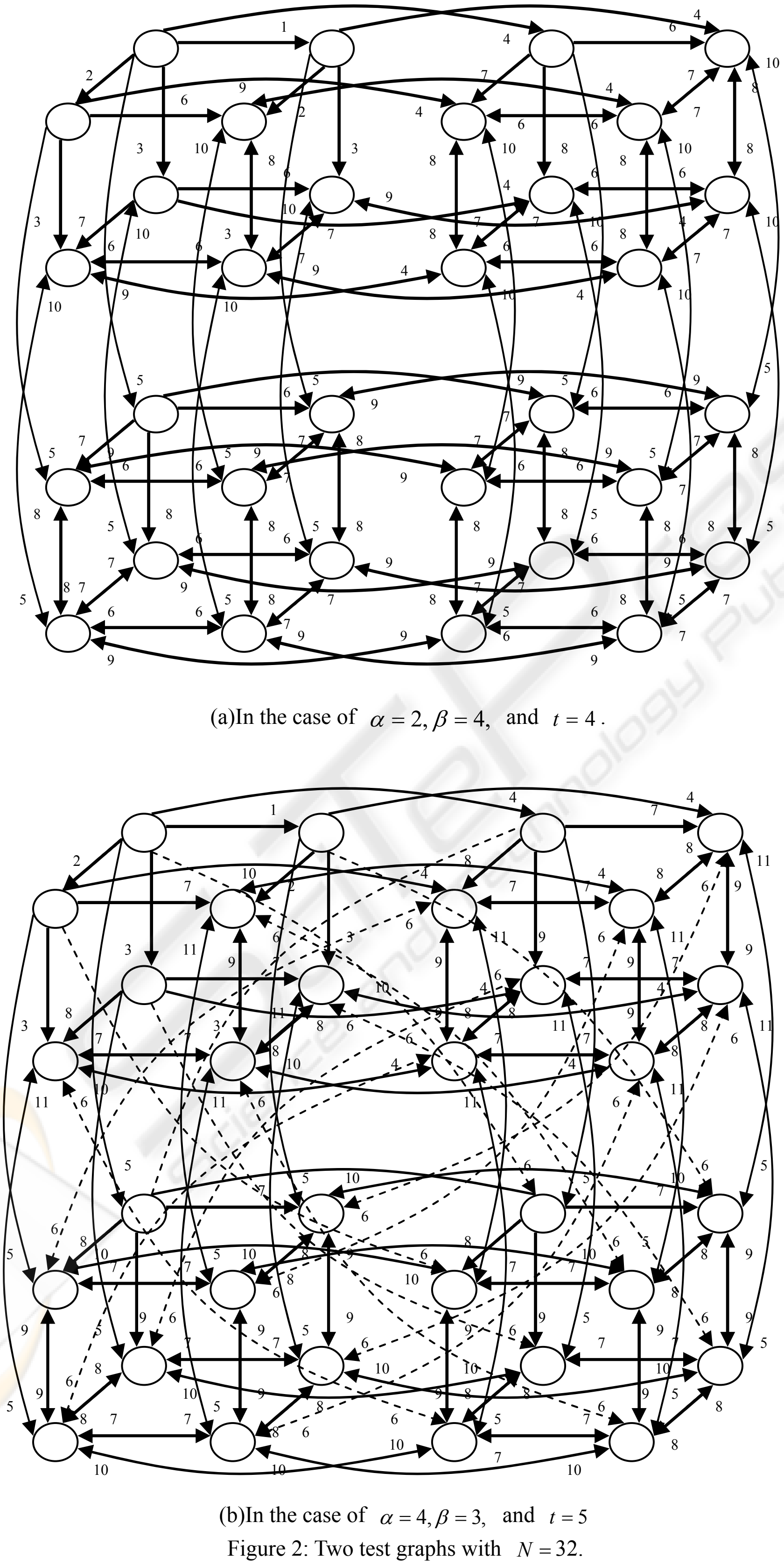
3 EVALUATION

3.1 Number of test times

The total number of edges in a test graph with $N = \alpha \cdot 2^{\beta}$ and $t = \alpha + (\beta - 2)$ is $(N - (\alpha + \beta))(t + 1) + (\alpha + \beta - 1)$, where we count a bidirectional edges as 2 edges. This value becomes close to $N(t+1)$ when N is large. Let the total number of test times where each test executes in different time be T. Since the total number of nodes is N, then the number of tested arcs can increase exponentially up to N by taking test time γ which satisfies $N = 2^{\gamma}$. After the time γ, since the total number of tested arcs is $\sum_{i=0}^{\gamma-1} 2^{i}$,the number of untested arcs is $N(t+1) - \sum_{i=0}^{\gamma-1} 2^{i}$. These $N(t+1) - \sum_{i=0}^{\gamma-1} 2^{i}$ untested arcs can be tested N to N every test time, then it takes total $\left\{N(t-1) - \sum_{i=0}^{\gamma-1} 2^{i}\right\} / N$ times.

Therefore, T is given as follows:

$$\begin{aligned} T &= \gamma + \left\{N(t+1) - \sum_{i=0}^{\gamma-1} 2^{i}\right\} / N \\ &= \gamma + t \\ &= \log N + t \qquad (1) \end{aligned}$$



(a)In the case of $\alpha = 2, \beta = 4,$ and $t = 4$.

(b)In the case of $\alpha = 4, \beta = 3,$ and $t = 5$

Figure 2: Two test graphs with $N = 32$.

SCITEPRESS Science and Technology Publications

Table 1: Probability of correct diagnosis for realistic yield in N= 2^{13}.

t	Yield (%)				
	99.999	99.750	99.500	99.250	99.000
1	1.0000	0.9616	0.8551	0.7031	0.5420
2	1.0000	0.9998	0.9989	0.9951	0.9915
3	1.0000	1.0000	1.0000	0.9999	0.9999
4	1.0000	1.0000	1.0000	1.0000	1.0000
5	1.0000	1.0000	1.0000	1.0000	1.0000
6	1.0000	1.0000	1.0000	1.0000	1.0000

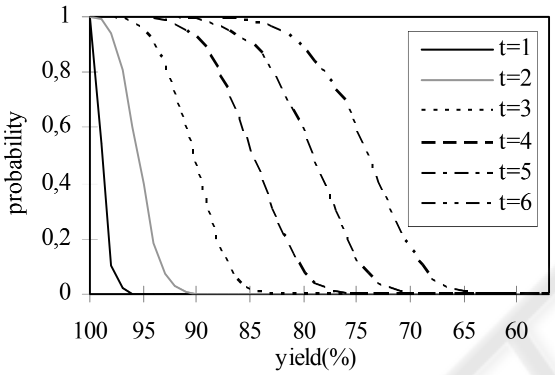


Figure 3: Probability of correct diagnosis for 6 diagnosabilities and N= 2^{13}

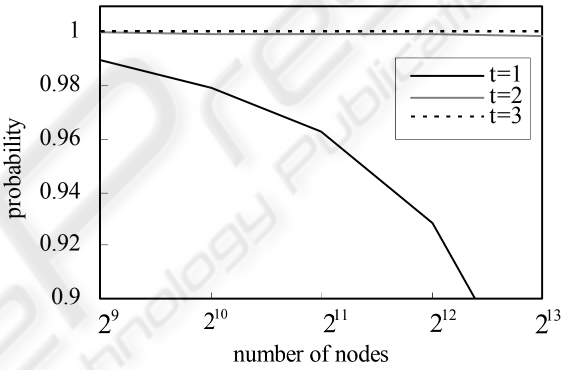


Figure 4: Probability of correct diagnosis for 5 network scales in

3.2 Time complexity of diagnosis algorithm

Each node can test its adjacent nodes asynchronously in the test order which is given automatically by the test graph. Therefore, on the assumption that the time complexity of algorithm to test a node by the adjacent monitor is 1, the time complexity of diagnosis algorithm can be evaluated as the same as T.

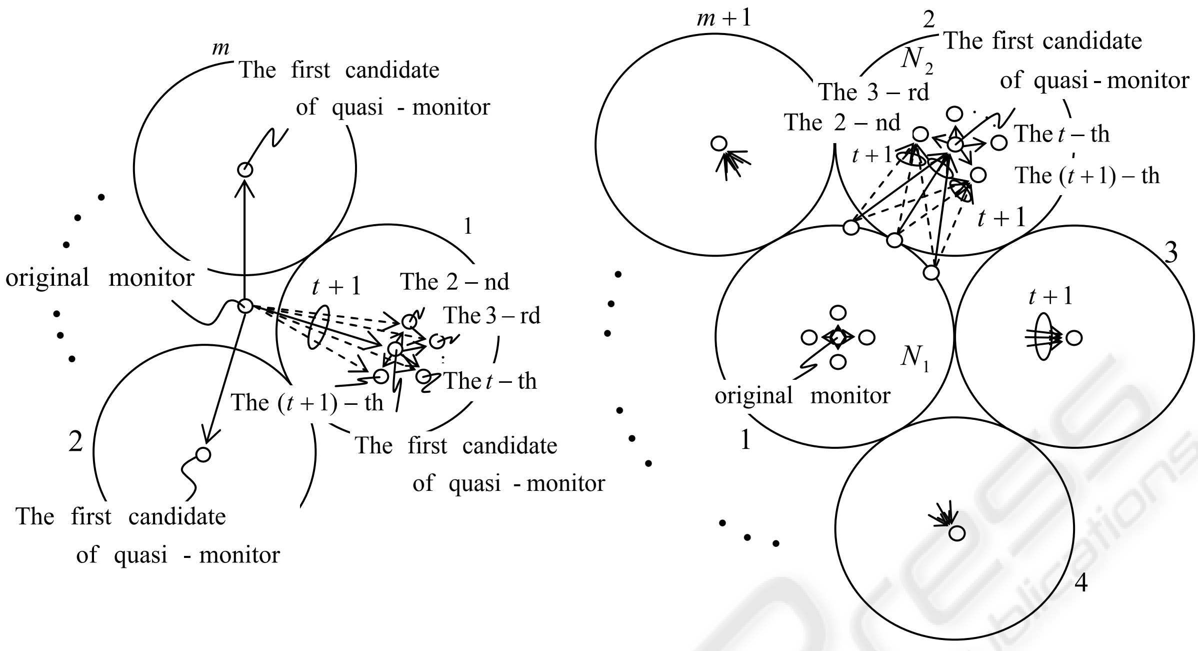
3.3 Amount of transmit messages

Each node hands a message “faulty node name” to the adjacent node by which it is tested first if it decides its testing adjacent node faulty. Then, these messages “faulty node name” pass through at most $t(t+1)d_m$ edges in a test graph. The average amount of transmit messages on an edge is given as $t(t+1)d_m / N(t+1)$, that is td_m / N.

3.4 Analysis of diagnosability t under clustered fault distribution

Extensive simulations were performed for evaluating the diagnosability when faulty nodes are clustered in a system. The examined systems consist of $2^{10} \sim 2^{13}$ nodes. A thousand different configurations of clustered faulty nodes in a system were simulated using negative binominal distributions. The diagnosis algorithm was run on all these configurations. Fig.3 gives the probability of correct diagnosis for the 6 scenarios of diagnosability and N= 2^{13}. It can be observed from Fig.3 that, for any yield Y, the probability of correct diagnosis is higher for higher diagnosability. Thus, the diagnosis with *t*=1 has the least probability of correct diagnosis over all yields, as was expected. What we need to know is the smallest diagnosability by which diagnosis is correctly performed under the limits of realistic circumstances. Table 1 gives the probability within the realistic yield values in

SciTePress Science and Technology Publications



(a)ApproaI

(b)Approach II

Figure 5: Two reduction approaches.

N= 2^{13} . Fig.4 gives the probability for the 5 scenarios of network scale in the case of Y=99.5%. These data show an answer that t=2 is proper.

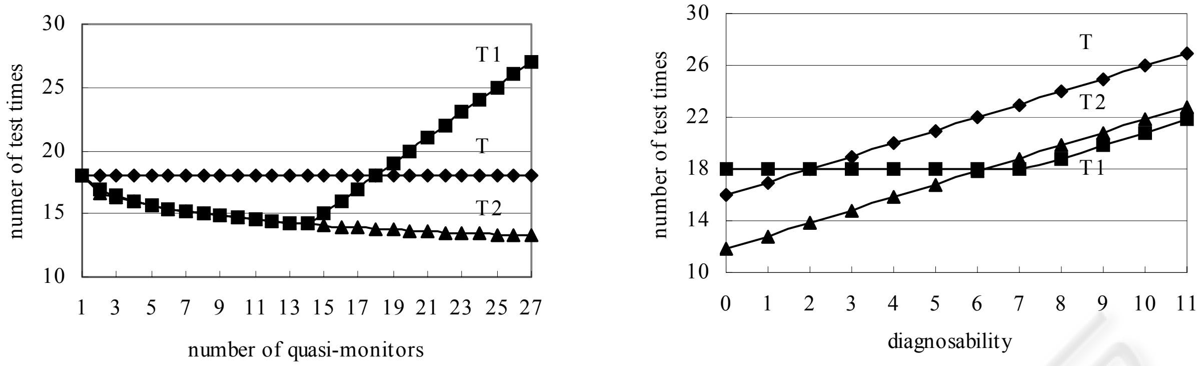
4 REDUCTION OF DIAGNOSIS PROCESS

In this section, we consider a technique to reduce the number of test times. Two approaches can be proposed as follows:

Let us set m quasi-monitors which perform the same test processes as the original monitor's. Since these quasi-monitors are not connected directly with the central observer, the gathered test results (faulty node names with its tester name) are stored temporarily in each quasi-monitor until each quasi-monitor receives a qualification as a monitor. After that, the quasi-monitor hands its test results to its own tester. The tester next hand the test result to the tester's tester, and so on. Finally, the test results is transmitted to the central observer. On this condition, we can consider two approaches to test the quasi-monitors as shown in Fig.5. In Fig.5(a), the original monitor tests only m quasi-monitors, then it does not test any other node. In Fig.5(b), the original monitor does not test any quasi-monitor directly, then each quasi-monitor is tested by the adjacent nodes obtained a qualification as a monitor. This reformed point is that both original and quasi monitors enter for testing simultaneously. The un-inscribed part in each circle in Fig.5 means the same structure as the test graph shown by Algorithm A. Each quasi-monitor hands its stored test results to its tester in order, as mentioned above. Then all test results can be gathered in a host directly connected the original monitor, and then the host can locate all faults in the network.

From the above discussion, we can understand the intention to reduce the number of test times, that is, the test graph can be partitioned into m (in Fig.5(a)) or $m + 1$ (in Fig.5(b)) parts by preparing m quasi-monitors. When the first candidate of quasi-monitor is judged as faulty, the second candidate is next tested, and so on. When an adjacent node of the first candidate of quasi-monitor is judged as non-faulty, the node takes the place of the first candidate of faulty quasi-monitor. The new quasi-monitor begins testing its adjacent nodes from the beginning.

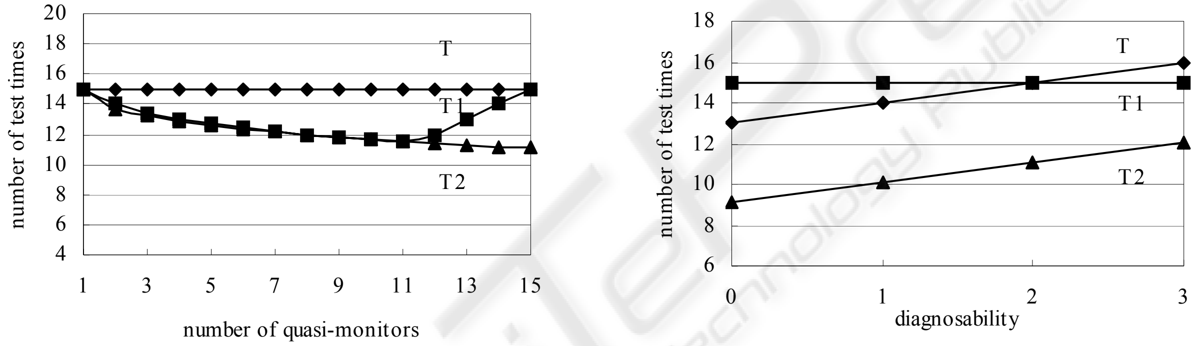
Let us consider the relative merits of the above two approaches in the point of the number of required test times. Let T_1 and T_2 be the numbers of test times required, when all the first candidates of quasi-monitor are non faulty, in the approaches shown in Fig.5 (a) and (b), respectively. That is, T_1 and T_2 are the smallest numbers of test times required in the approaches shown in Fig.5 (a) and (b). We obtain the following two equations from



(a) In the case of $t = 2$

(b) In the case of $m = 18$

Figure 6: The relative merits in the case of $N = 2^{16}$.



(a) In the case of $t = 2$

(b) In the case of $m = 15$

Figure 7: The relative merits in the case of $N = 2^{13}$.

eq.(1):

$$T_1 = \max[m, \log(N/m) + t - 1]$$

$$T_2 = \max[\log N_1 + t + 1, \log N_2 + t]$$

Were N_1 and N_2 are the total numbers of nodes in circles 1 and 2 in Fig.5 (b), respectively. m is restricted by the following relationships:

$$(t+1)m \le N_1$$

$$N_1 + mN_2 = N$$

For simplification, we assume $mN_1 = N_2$, then we obtain T_2 and an inequality for m as

$$T_2 = \max[\log\{N/(m^2+1)\} + t + 1, \log\{mN/(m^2+1)\} + t],$$

$$(t+1)(m^2+1)m \le N \qquad (2)$$

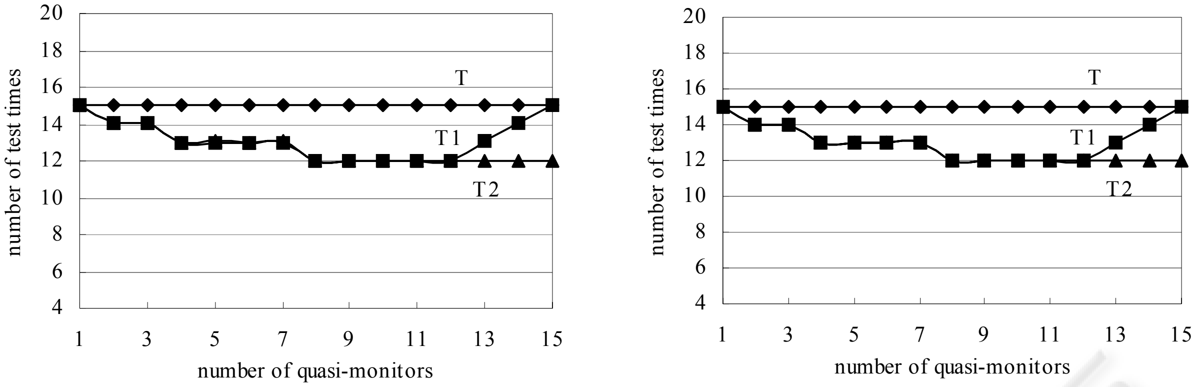
On the other hand, in the worst faulty case, that is, the biggest numbers $T_{1\max}$ and $T_{2\max}$ of test times required in Fig.5 (a) and (b), respectively are as follows:

$$T_{1\max} \cong m + 2t + \log(N/m)$$

$$T_{2\max} \cong 3t + 2\log\{N/(m^2+1)\} + \log m$$

Fig.6(a) shows the relative merits of the above two and original approaches in the case of $N = 2^{16}$ and $t = 2$ under the restriction given by eq.(2). In this case, the boundary line of the relative merits is $m = 18$, that is, the scheme shown in Fig.5(b) is superior to the others. On the other hand, Fig.6(b) shows the merits in the case of $N = 2^{16}$ and $m = 18$ under the same restriction. In this case, the boundary line of the relative merits is $t = 6$, that is, the scheme shown in Fig.5(a) is the best when t is over the boundary. Fig.7 shows relative merits in the case of $N = 2^{13}$, where the results show the same tendency as in the case of $N = 2^{16}$.

Extensive simulations were performed also for evaluating the relationship of the number of test times versus the number of quasi-monitors when faulty nodes are clustered in a system of 2^{13} nodes. A thousand different configuration of clustered faulty nodes in the system were simulated using negative binominal distributions on condition of t=2. Figs.8(a) and (b) show the results in the cases of



(a) In the case of $t = 2$, $N = 2^{13}$ and Y=99.95%

(b) In the case of $t = 2$, $N = 2^{13}$ and Y=99.5%

Figure 8: The relative merits in realistic circumstances of fault pattern.

Y=99.95% and 99.50%, respectively, where Y is the yield of nodes in the system. The same property as mentioned above is proved in realistic circumstances.

5 CONCLUSION

A hierarchical non-adaptive diagnosis algorithm is presented for testing total N nodes of computer networks. We proposed a noble test graph with $(t+1)$-connectivity enabling to test as many nodes as possible in a network due to a hierarchical architecture of test processes. If the maximum number of faulty nodes distributed in a network is bounded by a predefined limit t, our approach is effective. In this approach, an original monitor is placed at a key location in a network, and at first the monitor tests the adjacent nodes. Only adjacent nodes that passed the test can become new monitors and test their adjacent nodes, and so on. This process is propagated to higher levels of the test graph. At each level, every new monitor sends their information as a successful candidate (new monitor) back to a central observer directly connected original monitor through only one route. Monitor sends its test result back to a central observer through only one route if it decides its adjacent node faulty. Consequently, the observer can gather all information of faults in the network. The amount of test times is evaluated as the diagnosis latency. Optimal diagnosability t is analyzed under clustered fault distribution. Two revised approaches to reduce the required test times are discussed and the relative merits of three approaches are evaluated.

REFERENCES

F.Preparata, G.Metze, and R.T.Chien, “On the Connection Assignment Problem of Diagnosable Systems,” IEEE Trans. Electronic Computers, vol.16, pp.848-854, 1968.

C.Feng, L.N.Bhuyan, and F.Lombardi, “Adaptive System-Level Diagnosis for Hypercube Multi-Processors,” IEEE Trans. on Computers, vol.45, no.10, pp.1157-1170, 1996.

C.R.Kime, “System Diagnosiss,” In Fault-Tolerant Computing: Theory and Techniques, vol.2, D.K.Pradhan(ed.), Prentice-Hall, New Jersey, 1986.

D.P.Siewiorek and R.S.Swarz, “Reliable Computer System – Design and Evaluation,” 2nd ed. Digital Press, Bredford, MA, 1992.

N.H.Vaidya and D.K.Pradham, “Safe System Level Diagnosis,” IEEE Trans. Comput. Vol.43, no.3, pp.367-370, 1994.

S.L.Hakimi and A.T.Amin, “Characterization of Connection Assignment of Diagnosable Systems,” IEEE Trans. Comput., no.1, vol.C-23, 1974.

R.P.Bianchini and R.Buskens, “Implementation of On-Line distributed System-Level Diagnosis Theory,” IEEE Trans. Comput. vol.41, no.5, pp.616-626, 1992.

E.P.Duarte Jr. and T.Nanya, “A Hierarchical Adaptive Distributed System-Level Diagnosis Algorithm,” IEEE Trans. Comput. Vol.47, no.1, pp.34-45, 1998.

F.J.Meyer and D.H.Pradhan,”Dynamic Testing strategy for Distributed Systems,” IEEE Trans. Comput., vol.39, no.3, pp.356-365, 1989.

L.Zakrevski and M.G. Karpovsky, ”Fault-Tolerant Message Routing for Multiprocessors.” Parallel and Distributed Proscessing (Edited J.Rolim), Springer, pp.714-731, 1998.

H.Masuyama, Y.Ohashi, and T.Miyoshi, “A Diagnosis Method of Computer Networks.” 2001 Proceedings of IASTED Parallel and Distributed Computing and Systems, pp.474-479, 2001.

SciTePress Science and Technology Publications