

92-73, 陈

关于链路故障的分布式故障诊断

蔡兵 周六工 陈延槐

(重庆大学计算机研究所, 重庆630044)

TP338.8

摘要 In the distributed fault diagnosis, besides diagnosing the fault of nodes, the fault of links should be thought. This paper discusses the relationship of diagnosability T , edge connected degree K , bus number L , and node number N , given the diagnosability T and the node number N , how to determine the edge connected degree K and bus number L , and puts forward a distributed link fault diagnosis algorithm.

关键词 Link fault, Distributed fault diagnosis, Edge connected degree, Diagnosability, Bus number.

一、引言

在分布式系统中,故障诊断应包括故障结点机的诊断和故障链路的诊断,故障链路的诊断在分布式系统级故障诊断中占有重要地位,它不仅是分布式系统级故障诊断的目标,而且也是分布式系统级故障诊断的前提,因为一个系统的链路发生故障,则将会对故障结点机的诊断带来困难。在现有的诊断算法中,并不是所有的诊断算法都考虑了链路故障。

在分布式系统中,虽然任两个结点机之间都能进行通讯,也即都能进行互测,因而实际分布式系统的测试图的拓扑结构可以是任意结构的,但我们一般只在实际分布式系统的拓扑结构是分布式的或多总线的情形下才考虑链路故障。因为在单总线系统中,链路故障的出现将是致命的,会使整个系统的故障诊断成为不可能,因此一般不考虑链路故障。

在现有的几个著名的分布式故障诊断算法中,NEW-SELF 和 EVENT-SELF 算法考虑了链路故障,但它们采用的方法在单总线时是无效的,自适应 DSD 算法未考虑链路故障,本文主要是在自适应 DSD 算法已有的基础上考虑链路故障。

二、考虑链路故障的分布式系统

在考虑链路故障的分布式系统级故障诊断中,实际分布式系统的物理结构是欲考虑的重要因素,在物理拓扑结构已确定的情形下,这个系统的边连通度是一个重要决定因素,边连通度越大,这个系统的链路容错性越好,而在物理结构未确定的分布式系统中,怎样设计物理拓扑结构,使得在总线条数一定的情形下,边连通度 K 最大。

在分布式系统中,只要当故障链路数小于边连通度时,这个系统才是可诊断的系统,因为如果故障链路数大于或等于边连通度 K ,那么在一个特殊的情况下,删去它使系统不连通的 K 条链路恰好发生故障,将会使这个系统变成不可诊断系统。

我们可以考虑用图来表示一个分布式系统,即 $G=(V, E)$ 表示一个分布式系统 S ,用 V 表示系统中的结点机集合, E 表示系统中的线路集合。为设计一个容错度较高的分布式系统,可以采用多总线结构或分布式结构,本文研究多总线结构,一个这样的系统如图1所示。这里对容错度来讲,存在一个瓶颈,即每个结点只与两条总线相连,使系统的容错度达不到期望的目标,克服的方法是使每个结点与每条总线都相连。

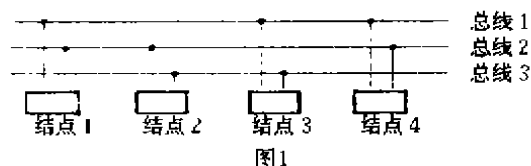


图1

一个有三条总线,四个结点机的高链路容错度系统如图1所示,它的抽象拓扑结构有多种,实际上等价于图2:

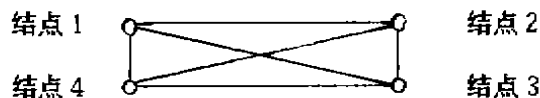


图2

这是一个有四个结点的完全图,它的边连通度在四个顶点的所有图中是最大的,因此它的链路容

错度最大。对于 N 个顶点的图, 它的边连通度最大可以达 $N-1$, 因此所需的总线数也将达 $N-1$, 这在实际应用中仍难以实现, 我们的目标是找到一个折衷, 让边连通度尽可能大, 而总线条数不至太多。

在分布式系统中, 它的诊断性数 T 可以只是指故障结点机数, 也可以是故障结点机数与故障链路数之和。前一种情况较简单, 它将链路故障另外加以考虑; 而后一种情况较复杂, 本文不考虑。一个系统是一个 T -可诊断系统, 是指当一个系统中同时出现的故障结点机数不超过 T 个, 且最大可能出现的故障链路数小于系统的边连通度 K 时, 这个系统的故障结点机与故障链路是能被唯一确定的。关于如何确定故障结点机已经研究得很多, 我们这里的工作是确定故障链路。

三、链路故障的诊断

在有 k 条总线的系统中, 当故障链路数小于 $k-1$ 时, 系统中结点机故障诊断能够正确完成, 而当故障链路数大于或等于 k 时, 系统将不能正常工作, 更谈不上进行诊断。当在系统中, 一个结点机诊断另一个结点机为无故障时, 这两个结点机间的链路无故障, 而当在一个系统中, 一个结点机诊断另一个结点机为故障时, 则它们之间的链路可能有故障。对于这种故障的情形就复杂了。在集中式故障诊断中, 可以由外部观察者, 假设用 A 表示, 对任一结点机 B 进行测试得到有关链路的状态, 如 A 测 B 为故障, 那么存在二种情况, 要么是 B 故障, 要么是它们之间的链路有故障。如果第一次测试用的是总线 1, 第二次用总线 2, ... 等等。直到通过一条总线测得 B 为无故障或所有总线试完。第一种情况是总线 1 到总线 i 中有故障, 第二种情况是总线无故障, B 为故障。这样一个系统如图 3 所示。

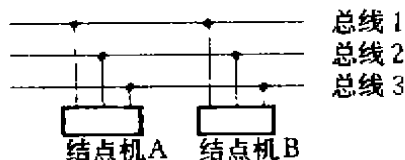


图3

在分布式故障诊断中, 没有了外部观察者, 对系统中结点机故障和链路故障的诊断是由各个结点机分别独立进行的, 为了对除自己外的其它结点机进行诊断, 必须先对各条链路进行诊断, 然后选择一条无故障的链路对结点机进行诊断。对于链路的瞬态故障(如受到不可预测的短暂干扰等), 可以采用循

环校验码等技术来发现(这是通信上经常采用的保证通信正常的方法, 这里不详述); 对于链路的永久性故障, 采用如下的方法, 首先在每个结点机处设置一个数组 $SL_i[j]$, 其中 $0 \leq i \leq N-1, 0 \leq j \leq L-1$, 用这个数组来表示各个结点机分别得到的关于各条链路的状态, $SL_i[j] = 0$ 的含义是结点机 i 认为第 j 条链路是无故障(故障) 1。在给出算法前, 先给出保证该算法正确运行的二个条件, 一是系统应是可诊断系统, 且 $N \geq 2T+1$, 其中 N 为系统中结点机个数, T 为系统中可能同时出现的故障结点机的最大数; 二是无故障结点机测试链路的状态能得到正确的测试结果, 故障结点机测试链路所得的测试结果无法预测。算法如下:

/* 此过程在每个结点机 i 处执行 */

```
1. For  $j=0$  To  $L-1$  Do  $SL_i[j]=0$ ;
2. For  $j=0$  To  $L-1$  Do {
    结点机  $i$  测试第  $j$  条链路的状态, 若无故障, 则
     $SL_i[j]$  不变; 若故障, 则  $SL_i[j]=1$ ;
};
3. For  $j=0$  To  $L-1$  Do { 将  $SL_i[j]$  数组在一条结点机  $i$ 
    认为无故障的链路上广播, 使  $SL_m[j]=SL_m[j]$ 
    +  $SL_i[j]$ , 其中  $0 \leq m \leq N-1$ , 且  $m \neq i$ 
};
```

当这个算法运行完后, 系统中每个无故障结点机的 SL 数组元素的值 v 都是一样的, 且有二种情况: $0 \leq v \leq T$ 或 $v > T$, T 为系统诊断性数, 值大于 T 的链路肯定是故障的链路; 而值为 $0 \leq v \leq T$ 的链路肯定是无故障链路。这个算法是一个有效算法, 能够很容易地得出它的时间复杂性为 $O(L)$ 。

在一个分布式系统中, 当其中的结点机数 N 给定的情况下, 为了使整个系统的链路容错度尽可能的大(在系统资源冗余度适当的情况下), 应设置较多的链路条数。在一个系统中, 全互连时的边连通度最大(为 $N-1$), 但所需的链路条数也最多(达 $N * (N-1)/2$ 条); 而采用多总线的结构时, 为了使边连通度达到 $N-1$, 只需 $N-1$ 条总线, 这大大减少了链路条数的规模, 也使系统资源浪费减少。在多总线结构中, 边连通度 K 是等于总线条数 L , 因此系统的容错度是与总线条数密切相关的, 为了达到一定的容错度, 需要在系统开销与总线条数之间进行折衷。

为了在系统开销与链路条数之间达成折衷, 可以将一个大系统划分为几个子系统, 每个子系统中结点机数 $> 2T+1$ (T 为原系统的诊断性数), 这样, 按自适应 DSD 算出, 每个子系统仍然是 T -可诊断系统, 每个子系统的总线条数 $\geq 2T$, 那么子系统个数 $\leq \lceil N/(2T+1) \rceil$, 子系统的组成原则为: 由地理上邻近的结点或频繁更换数据的结点组成。子系统之间由一条总线连接, 这样整个系统的总线条数减少

了。图4为有10个结点机的2-可诊断系统。
这个系统可以看成是由2个全互连的子系统构

成,这样的结构既可以提高系统的链路容错度,又可以使整个系统的总线条数不至过高。

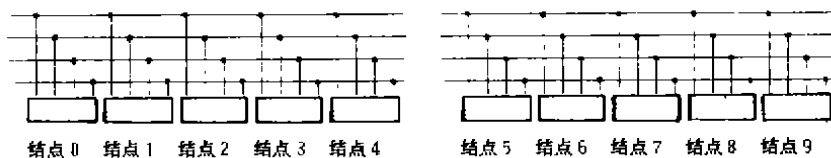


图4

四、性能分析

在单总线系统中,链路的故障会使整个系统的诊断功能全部失效,根本谈不上诊断结点故障,在一般的多总线系统中(即没有划分子系统的多总线系统),由于存在多条总线,所以当至少有一条总线未失效时,整个系统仍然可以正常运行,仍然可以进行故障诊断处理(包括结点机故障与链路故障诊断)。

单总线系统与多总线系统各有长处与不足,单总线系统的系统开销较小,但链路容错度为0,而多总线系统的开销虽大,但链路容错度很高(随着总线条数的增加而增高)。本文提出的多总线结构与一般的多总线结构不同,它由若干子系统组成,各子系统内为一般意义上的多总线结构,各子系统间为单总线结构,它是单总线结构与一般多总线结构的一个折衷。在一个有 N 个结点机的系统中,为了在物理结构上达到全互连,需 $N * (N-1)/2$ 条线路,而在一般多总线结构中,为了在物理上达到全互连需 $N-1$ 条线路。由此可见,为了达到同样的目的,所需的线路条数会大大减少,特别是当系统越大,即系统中结点机个数越多时,效果更明显。在本文的结构中,当系统中的结点机个数为 N 时(诊断性数为 T),子系统内的结点机数应 $\geq 2T+1$,子系统的总线条数 $\geq 2T$,与一般多总线结构相比,线路条数再次减少。另外,由于

各子系统内的结点机数 $\geq 2T+1$,所以每个子系统中总是至少有一个结点机是无故障的,这样只需采用适当的诊断模型及诊断算法,既可以把系统看作一个整体来进行诊断,也可以分别对各个子系统分别进行诊断,于是为诊断提供了很大的灵活性。

五、结束语

链路的故障对系统的正常运行有着异常重大的影响,将链路的故障诊断出来,然后使链路状态恢复正常。为了使系统具有链路诊断故障的功能以及较高的容错度,必须使系统的边连通度 $K>1$,而系统采用多总线结构是能达到这种要求的一类较简单的方法。本文介绍的这种多总线结构与一般意义上的多总线结构略有不同,它是一种介于单总线结构与一般多总线结构之间的一种结构,它的系统开销比一般多总线结构小,而它的容错度比单总线结构大。

在具有诊断链路故障及提供较高容错度的系统中,系统除了采用多总线结构外,还可以采用其它结构,如分布式结构,其目的都是使其边连通度提高,而线路数不至太大。在一般的考虑链路故障的故障诊断中,都是将结点机故障与链路故障分开考虑,也许将结点机故障与链路故障结合起来考虑会取得更好的效果。(参考文献共7篇略)

鸣哀:本文作者之一——周六丁先生因心脏病突发于美国逝世,特此哀悼。

计 算 机 科 学

(1974年1月创刊)

第21卷第4期(双月刊)

1994年8月23日出版

国内统一刊号:CN51-1239/TP

定价:3.00元 国外定价:5美元

邮发代号:78-68

编辑者:中国科学技术信息研究所重庆分所

顾问:《计算机科学》审编委员会

出版者:中国科学技术信息研究所重庆分所

重庆市市中区胜利路132号

邮政编码:630013 电话:(0811)3850828

印刷者:中国科学技术信息研究所重庆分所印刷厂

总发行处:四川省重庆市邮政局

订购处:全国各地邮政局