

# 分布式指挥控制系统网络结构的研究与设计

张伟

(中国科学院长春光学精密机械研究所, 长春 130022)

**摘要** 本文系统地论述了分布式控制系统的概念、原理及构成,研究了各种典型分布式互连网络的性能和适用条件,提出了针对于指挥控制系统的设计思想及重构策略。在此基础之上,完成了一种高可靠性的分布式指挥控制系统的结构设计,并给出了该系统的通讯性能和可靠性能分析,最后对在可修系统作了可用性能分析。

**关键词:** 分布式计算机系统; 可靠性; 互连网络; 重构; 可用性

## 1 前言

近十几年来,微型机的发展很快。但是,单台微机的资源及功能毕竟很有限。尤其对于某些要求较高的实时处理系统就显得力不从心。于是人们打破了传统的单处理机观念,从分布计算和并行处理的概念出发,组织计算机系统。有些问题如矩阵、迭代等本身存在着并行性,计算机系统内部的某些操作如运算、输入/输出、资源调度和分配等也存在着同时性。这种并行性和同时性就是设计分布式体系结构的基础。

实际上分布式系统是一个并行处理系统,而有人也逐步认识到并行处理系统必须要分布化,于是现在形成了分布式控制的概念。把并行处理作为手段,把分布式控制作为形式,是分布式控制系统模式的主要指导思想。

分布式控制系统包括的内容十分广泛,在此我们只作定性说明。分布式控制系统具有以下特性:

- 具有资源分散性,包括物理资源和逻辑资源;
- 结构模块性,任一结点可为一完整模块;
- 工作并行性,在其全部资源上,存在真正的物理并行处理;
- 协作自治性,各结点间协作自治,进程间通过共享信道进行有双边协议的消息通信;
- 高可用性,局部出现故障可通过重构,适度降级继续工作。

## 2 分布式控制系统互连网络的拓扑结构

分布式控制系统模型中的物理层是由若干个处理机用通信线路连接而成的一个计算机网

络。网络的拓扑结构如何,对整个系统的性能有很大的影响。

网络一般是由接口机、通信线路连接的通信子网、各接口机上连接的主机或集中器构成的。与网络性能相关的参数很多,主要有通信延迟时间、传输能力、通路数目、故障影响和瓶颈等几个参数。下面是几种典型的分布式网络结构。

(1)公用总线式结构:这种结构是将网上的结点都排列在一条或多条公用母线的周围,各结点类似网络中的功能部件,各结点之间没有直接数据链路。所有数据都通过公共总线传送。

(2)环型网络模式:在环形网中,每个结点在物理上和另外两个结点连接,为单向环模型。在单向结构中,一个结点只能把信息传给它的顺序相邻结点,所有的结点都必须以一个方向传递信息。

(3)星形网络结构:在星形网络中,只有一个结点与其余所有结点相连接,这个结点叫中心结点。除中心结点外,其余所有结点间均不相互连接,我们称这些结点为卫星结点。

(4)树形结构模式:这种结构比较适合于分级控制或社团计算机网。在顶层的计算机要求有通用性,任务是控制和协调整个系统。在底层的结点一般和应用有关,具有专用性,用于完成事先规定好的专门任务。

(5)全互连网络:这种网络中的任何每两个结点之间都有一条直接的通信通道,把所有分布在不同地点,负责不同任务的主机经链路互相连接起来。

### 3 系统设计思想

(1)对主处理机的考虑:选用高速度、少数量的计算机组成分布式指挥控制系统可使系统达到更高的效率。

在分布式系统中,并不是处理机越多则系统性能越高。实际上,由于多个处理机的协作造成的通信通路的冲突和调配负载平衡所消耗的机时,使得多台处理机的实际加速倍数远远小于理想值。

在不考虑一些额外冲突情况下,假设用单处理机可以在单位时间  $T_1=1$  内执行完任务。令  $f_i$  为该任务分配给  $i$  台处理机并使每台处理机工作平均负荷为  $d_i=1/i$  的概率。另外,假定每个操作方式使用  $i$  台处理机的概率相等,即对于  $n$  个操作方式有  $f_i=1/n, n=1, 2, \dots, n$ 。则可得在这  $n$  个处理机上完成这个任务所需的平均时间:

$$T_n = \sum f_i * d_i = \frac{\sum 1/i}{n} \tag{1}$$

式中的求和表示有  $N$  种操作方式。则平均加速倍数可从  $T_1$  与  $T_n$  的比得出:

$$S = T_1/T_n = \frac{n}{\sum 1/i} \leq \frac{n}{LN_n} \tag{2}$$

从图上可看出,处理机的数量越多,其加速的趋势越小,当有 1024 台处理机协作时,其平均加速倍数只有理想值的 1/7。

(2)分布式计算机系统的可靠性设计原则:在混联模型中选用并串联模型。

串并联的数学模型为:

$$R_{sp}(t) = 1 - [1 - R^n(t)]^N \tag{3}$$

并串联模型的数学模型为:

$$R_{ps}(t) = [1 - (1 - R^N(t))]^n \tag{4}$$

当  $n=2, N=2$  时, 根据上面的公式, 串并联与并串联系系统可靠度如图 2。

### 4 系统网络拓扑的结构设计

分布式指挥控制系统的网络设计与其它应用网络的设计不太相同。由于指挥控制系统对实时性处理的要求非常高, 而且其数据传输量很大, 因此, 它的网络要求有更短的时延, 更大的吞吐量和更高的可靠性。

在本指挥系统中:

处理机的分布属于任务分布型, 处理机结点之间的地理分布十分接近, 因此通信线路带来的费用问题可以放在次要位置。

处理机数量较少, 因此增加冗余链路不会造成技术困难。

通过对本指挥系统信息流程的分析, 发现跟瞄预处理机与中央综合处理机之间, 告警予处理机与中央综合处理机之间, 中央综合处理机与显示处理机之间的通信最忙。根据上述考虑, 我们可简略得出系统网络的设计方案。即连接上述三条链路, 暂定名为方案一。这种连接方案属于星形网络, 它的优点是延时小, 简易, 扩充性好。但它具有一个最大的弱点, 即可靠性不高。而对于一个指挥系统来说, 可靠性是系统的最重要因素之一。因此我们不能采取这个方案。

由于整个系统的网络中结点数很少, 且相互之间的距离又很近, 因此我们倾向于采用全互连网络形式, 虽然只增加了三条链路的连接量, 却使网络性能得到了很大的提高。同时在设计中我们发现, 作为指挥系统的信息接收部分的跟瞄处理机和告警处理机是一个薄弱环节, 因此在设计时为保证其可靠性和便于系统重构, 增加了链路冗余。

由于整个系统的网络中结点数很少, 且相互之间的距离又很近, 因此我们倾向于采用全互连网络形式, 虽然只增加了三条链路的连接量, 却使网络性能得到了很大的提高。同时在设计中我们发现, 作为指挥系统的信息接收部分的跟瞄处理机和告警处理机是一个薄弱环节, 因此在设计时为保证其可靠性和便于系统重构, 增加了链路冗余。

### 5 网络通信性能研究

#### (1) 通信性能参数

对于分布式控制系统的设计来说, 最重要的是提高系统中主处理机的使用效率。也就是说要最大限度地减少数据处理之外的开销诸如主处理机相互等待和通信费用。

通信费用主要有以下几个参数:

- 网络的链路总数  $M$
- 结点间的平均距离  $L$ : 结点间的平均距离是指网路内任意两个结点之间最短距离之和除以路径总条数。假设网络中共有  $N$  个结点, 用  $d_{ij}$  表示结点  $i$  与结点  $j$  之间的最短距离, 则根据

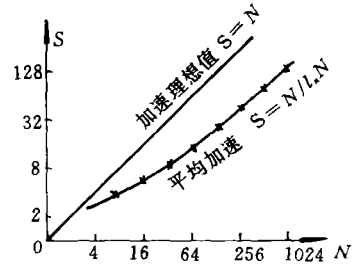


图1 分布式多计算机系统处理速度加倍情况

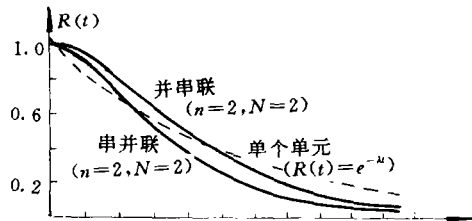


图2 串并联与并串联系系统可靠度曲线

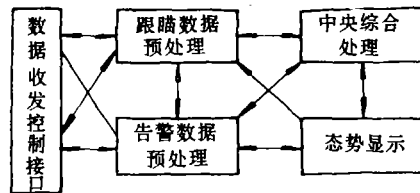


图3 全互连网络结构

定义, 可得 
$$L = \frac{1}{N} \sum \frac{1}{N-1} \sum d_{ij} \quad (5)$$

●网络的报文包传送时间  $\Delta t$ : 这里我们可以定义  $\Delta t$  为一个报文包通过链路距离所用的时间。

现在假设网络中的一次通信中, 网内每个结点都向系统中其余结点发送报文包, 则在这一过程中, 网络上共通过了  $X$  个报文包。

$$X = N(N - 1) \quad (6)$$

这样, 网络的一次通信时间为  $t_c$ 。

$$t_c = \frac{LX\Delta t}{M} \quad (7)$$

## (2) 系统建模

分析本指挥控制系统中各结点上的工作情况, 一个任务可以分成  $K$  个并行步完成, 而每个并行步的一般模式为:

- 接收数据
- 处理数据
- 发送结果

设每个结点处理数据时间为  $t_p$ , 则该并行步所用时间  $t$  为:

$$t = t_o + t_p + t_o \quad (8)$$

则该任务的所用总的时间  $T$  为

$$T = kt = k(t_p + 2t_o) \quad (9)$$

在该任务中总的通信时间为  $T_c$ 。

$$T_c = k * 2t_o \quad (10)$$

由以上各式可知, 结点间的通信时间  $T_c$  与全部时间开销  $T$  之比:

$$\begin{aligned} T_c/T &= \frac{2kt_o}{k(t_p + 2t_o)} \\ &= \frac{2}{2 + [M/(L * N(N - 1))] * (t_p/\Delta t)} \end{aligned} \quad (11)$$

公式 11 中的  $M/(L * N(N - 1))$  为网络结构参数, 由此可以看出, 网络的结点平均距离越小, 平均连接数越高, 则网络中通信开销占总开销的百分比就越小。

从软件方面可得出:

相对处理数据时间, 结点间发送的信息越少, 报文包越小, 则可提高  $t_p/\Delta t$  之比, 可以减少  $T_c/T$  的百分比。这就要求软件设计人员更好地利用任务的独立性, 减少通信的频率和数据量。

根据网络设计数据, 可得如下比较图。

## 6 网络重构策略

分布式实时系统与传统的集中控制系统相比较, 具有许多优越性, 其中之一就是分布式系统具有极高的可靠性。在某些环节出现故障时, 系统能够进行重构以维持主要工作的连续。

重构的基本任务有

在系统网络内有链路出现故障时, 与其相关的处理机能够报告及自动采用迂回路由实现通信。

在系统网络中有通信控制器出现故障时,与其相关的处理机能自动关闭该接口(使接口处于高阻状态),并采用迂回路由通信。

在系统中如有主处理机发生故障,则由事先预定的处理机接替其工作,并自动切断与该处理机的通信联系。

故障的管理与重构系统可工作在系统中的数台主处理机中,从硬件和软件两方面入手,采用自测与互测相结合的故障检策技术。自测即对本处理机及本处理机上的通信控制器的输出进行检查,一旦发现处理机内某部分出现故障可向其它处理机及管理人员报警,自测工作也可防止本身出现错误后,误以为其它处理机有故障,而导致系统的误动作。互测为每台处理机对与它通讯的处理机的状态进行监督,并对接收的数据进行校验,一旦发现某一处理机与本机通信不规律,或接收的数据始终有校验错误,则可把情况通知其它正常处理机,进行综合表决判断。

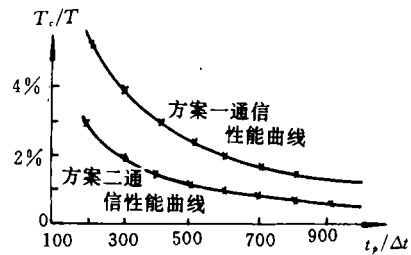


图4 不同网络形式通信占用机时百分比

## 7 网络可靠性研究

### (1) 方案一的可靠性分析

方案一中的网络连接属于星形结构,对其工作过程的分析可得出其任务可靠性框图如图6。

在整个任务中,这三台主处理机缺一不可,且无并一旁联措施,因此可以认为是一串联系统。这种系统的可靠性极低,只要有任一条链路或任一个结点出现故障,整个系统都会陷入瘫痪。

此时系统可靠度

$$R_s(t) = \prod R_i(t) \quad (12)$$

### (2) 方案二的可靠性分析

网络设计方案二的结构属于全互连网络类型,这是一种高度连通的网络。因而可提供较多的迂回路由并能够组成较好的任务可靠性模型。

这样,通过对信息流程及系统工作过程的分析,我们可

由于并联情况下  $R_s(t) = 1 - \prod (1 - R_i(t))$

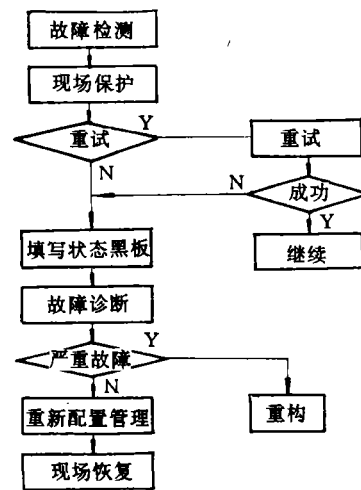


图5 重构策略

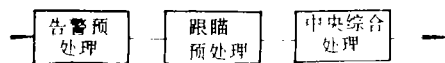


图6 方案一网络可靠性模型

## 8 可用性研究

一个系统圆满实现其规定功能的能力通常称为系统的可用性。指挥控制系统总是要求具有很高的可用性,其性能取决于系统的可靠性以及部件自动切换、修复的时间。

可用性的度量指标是“可用度”,用  $A$  表示。可用度定义为在任一随机时刻,当任务需要时,系统在任务开始时刻处于可投入使用状态的概率。

基本数学模型为:

$$A = \frac{MTBF}{MTBF + MCT} \quad (13)$$

其中  $MCT$  为平均维修时间。

在利用马尔可夫过程方法建立可用度模型时,为简便起见,作下列假设:

由于系统各部件都是工作在偶然故障期间内故其寿命分布为指数分布,即故障率  $\lambda$  恒定。并设修理工的维修率  $\mu$  是常数,系统配一名修理工。

在时刻  $(t, t + \Delta t)$  发生故障的条件概率是  $\lambda \Delta t$ 。

在时刻  $(t, t + \Delta t)$  完成修复的条件概率是  $\mu \Delta t$ 。

在  $\Delta t$  时间内出现两次或两次以上故障/修复事件的概率为零。

每次故障或修复事件是独立的,与其他事件无关。

下面,对系统状态进行分析:

首先,让我们看一下本系统的任务可靠性模型。我们只讨论系统网络上的四个主要结点,其中 A 型结点是由一台主处理机和两块通信处理板组成, B 型结点是由一台处理机和一块通信处理板组成。

我们假设 A 功能块的故障率为  $\lambda$ , 维修率为  $\mu$ , B 组功能模块的故障率为  $\lambda_b$ , 维修率为  $\mu_b$ 。当部件发生故障时,由一个修理工负责。

用  $x(t)$  表示  $t$  时刻系统可能的状态。

0: 两组四个功能模块都完好,系统状态为工作。

1: A 组模块完好, B 组模块有一个故障,系统状态为工作。

2: A 组模块完好, B 组模块二个都故障,系统待修。

3: A 组模块一个故障, B 组模块完好,系统状态为工作。

4: A 组模块一个故障, B 组模块一个故障,系统状态为工作。

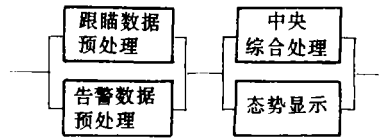


图7 方案二可靠性模型

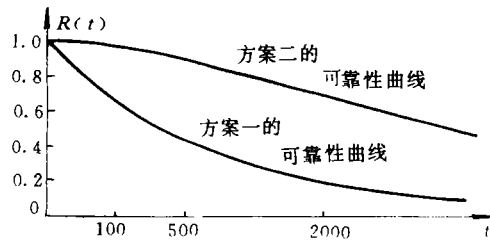


图8 根据参考数据得出的两种方案可靠度曲线

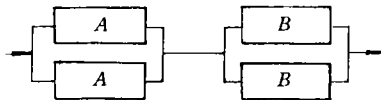


图9 方案二系统可靠性模型

5: A 组模块一个故障, B 组模块二个故障, 系统状态为待修。

6: A 组模块二个故障, B 组模块完好, 系统状态为待修。

7: A 组模块二个故障, B 组模块一个故障, 系统状态为待修。

通过解状态方程可得出系统的稳态可用度:

$$\begin{aligned}
 A &= A_0 + A_1 + A_3 + A_4 \\
 &= \frac{\mu^1 + 2\mu^3\lambda + 2\mu^2\lambda^2}{\mu^4 + 2\mu^3\lambda + 6\mu^2\lambda^2 + 8\mu\lambda^3}
 \end{aligned}
 \tag{14}$$

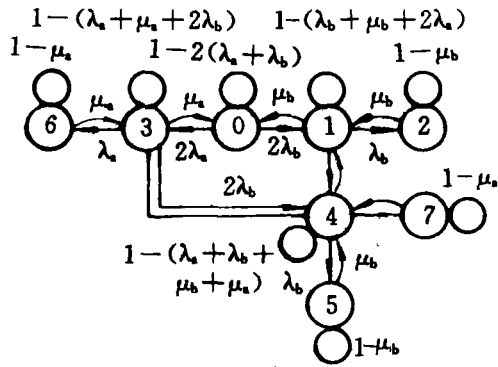


图 10 方案二的系统状态转移图

## 9 结 束 语

对分布式控制系统的研究,基本可以分为两部分:一是多处理机的互连网络结构,一是软件的并行性。研究的总的目标是完善各处理机之间的协作,充分发掘过程的并行性,最大限度地减少通讯时间,提高网络利用效率。

从我们对系统设计的分析结果来看,地理分布距离较近的系统,有目的地增加网络的辅助链路和进行重构,可以在费用增加不多的情况下,大大提高系统的综合性能。

### 参 考 文 献

- [1] 黄铨,《计算机结构与并行处理》. 科学出版社,1990
- [2] 陈瑞源,《通信指挥和控制系统导论》. 国防工业出版社,1980
- [3] 刘键,《分布式计算机系统》. 人民邮电出版社,1990
- [4] D. R. Mcglynn 《Distributed processing and data communication》. John Wiley & Sons, INC. 1978
- [5] S. M. Rezaul Islam, Hang H. Ammar, Performability Analysis of Distributed real-time system. IEEE, Transactions on computers, 40(11)
- [6] Richard R. Muntz, Ambuj Goyal, Bounding Availability of Repairable Computer System. IEEE, Transactions on computer, 38(12)
- [7] Hermann Kopetz, wilhelm Ochsenreiter, Clock Synchronization in Distributed Real-time Systes. IEEE, Transactions on computer, 36(8)
- [8] Sheldon S. L. Chang, multiple-Read Single-Write Memory and Its Applications. IEEE, Transaction on computers, 29(8)
- [9] Michael Flynn, John Hennessy, Parallelism and Representation Problems in Distributed Systems. IEEE, Transactions on computers, 29(12)

## **A Study and Design on the Structures of Distributed Computer Network**

Zhang Wei

*(Changchun Institute of Optics and Fine Mechanics, Chinese  
Academy of Sciences, Changchun 130022)*

### **Abstract**

Based on the structures and principles of distributed computer systems, this paper puts forward a interconnection network designing method suitable to C<sup>3</sup>I systems and its reconstructure tactics. It also presents a methodology for the availability and communication analysis of repairable distributed computer systems.

**Key words:** Distributed computer system, Reliability, Interconnection network, Reconstruct, Availability