매크로-스타 그래프에서의 일-대-다 방송 알고리즘

이형옥○, 류광택

한국전산원

{oklee,ktryu}@nca.or.kr

요 약

대규모 병렬 컴퓨터에서 메시지를 가진 한 노드에서 다른 모든 노드들로 그 메시지를 전달하는 방송은 데이터의 복제, 신호 처리와 같은 다양한 응용프로그램에서 이용되는 중요한 통신 패턴이다. 매크로-스타 그래프는 스타 그래프를 기본 모듈로 가지면서 스타 그래프가 갖는 노드 대칭성, 최대 고장 허용도, 계층적 분할 성질을 갖고, 스타 그래프보다 망 비용이 개선된 상호 연결망으로 최근에 제안되었다. 본 논문에서는 매크로-스타 그래프의 계층적 분할 성질과 기본 모듈을 이용한 매크로-스타 그래프에서의 일-대-다 방송알고리즘을 제안한다.

1. 서론

최근 반도체 기술의 발달과 높은 성능을 요구하는 응용분야의 증대로 고성능 컴퓨터에 대한 관심이 증대하고 있다. 고성능을 얻기 위한 방법으로 병렬처리에 대한 필요성이 크게 증가하여 병렬컴퓨터 대한 연구가 많이 진행되고 있다. 병렬컴퓨터는 크게 공유 기억 장치를 갖는 다중프로세서(multiprocessor) 시스템과 분산 기억 장치를 사용하는 다중컴퓨터(multicomputer) 시스템으로 분류한다. 다중컴퓨터 시스템은 각각의 프로세서들이 자신의 기억 장치를 갖고, 각 프로세서는 상호 연결망에 의해 연결되어 있으며 프로세서간의 통신은 상호 연결망을 통하여 메시지 전송 방식으로 이루어진다. 다중컴퓨터에서 상호 연결망은 전체 시스템의 성능과 시스템의 확장성에 큰 영향을 미친다. 상호 연결망의 평가 척도중 하드웨어의 비용과 관련된 분지수(degree)와 연결망에서의 망 처리량을 결정하는 지름(diameter)은 중요한 평가 기준으로써, 분지수와 지름의 곱으로 정의된 망 비용(network cost)[1]이 있다. 널리 알려진 상호 연결망으로 메쉬, 하이퍼큐브, 스타(star) 그래프[4], 매크로-스타(Macro-Star) 그래프[5]등이 있다. 매크로-스타 그래프는 스타 그래프를 기본 모듈로 하면서 스타 그래프와 같은 수의 노드를 가질 때 스타 그래프보다 적은 분지수를 가지면서 노드 대칭성, 계층적 구조, 최대 고장 허용도를 가지면서 스타 그래프보다 망 비용(network cost)이 개선된 상호 연결망이다[5].

상호 연결망의 방송(broadcasting)은 메시지를 갖고 있는 한 노드에서 다른 모든 노드로 메시지를 전송하는 것이다. 스타 그래프에서는 아래의 조건을 만족하는 최적 방송 알고리즘이 [4]에서 제시되었다. 방송은 다음과 같은 제약 하에서 통신 링크를 통해 메시지를 전송하는 과정인 호출을 연속해서 수행함으로써 이루어진다[2,4]. (1) 각 호출에는 오직 두 노드만이 관련된다. (2) 각 호출을 끝내는데는 한 단위 시간이 소요된다. (3) 각 노드는 각 단위 시간에 하나의 호출에만 참여할 수 있다. (4) 각 노드는 인접한 노드만을 호출할 수 있다.

본 논문에서는 스타(star) 그래프보다 망 비용(network cost)이 개선된 매크로-스타 그래프에서의 방송(Broadcasting) 알고리즘을 제안한다. 논문구성은 2장에서 매크로-스타 그래프를 정의하고, 3장에서 제안하는 방송 알고리즘을 알아보고, 마지막으로 결론과 추후 연구방향을 제시한다.

2. 매크로-스타 그래프의 정의 및 성질

매크로-스타 그래프 $MS(l,n)$은 $(nl+1)!$개의 노드와 $(nl+1)!*(n+l-1)$개의 에지로 구성된 연결망이다. 각 노드의 주소는 $k(=nl+1)$개의 서로 다른 심벌의 순열로 표현되고, 노드 v와 w의 연결관계는 아래의 2가지 에지 발생기 T_j, S_i를 적용하여 생성된 순열들 사이에 에지가 존재한다. k개의 서로 다른 심벌 집합을 $\langle k \rangle=\{1,2,..,k\}$이라 하고, $\langle k \rangle$에 대한 심벌의 순열을 $U=u_{1:k}=u_1u_2...u_i...u_k$, $u_i \in \langle k \rangle$이라 할 때 매크로-스타 그래프 $MS(l,n)$은 다음과 같이 정의된다[5].

$$V(MS(l,n))=\{U=u_{1:k} \mid u_i,u_j \in \langle k \rangle,\ u_i \neq u_j,\ i \neq j,\ 1 \leq i,j \leq k\},$$

$$E(MS(l,n))=\{(U,V) \mid U,V \in V(MS(l,n)) \text{ satisfying } U=T_j(V) \text{ or } U=S_i(V),\ 2 \leq j \leq n+1,\ 2 \leq i \leq l\}$$

에지 발생기 $T_j(U)=u_ju_{2:j-1}u_1u_{j+1:k}$는 순열 $U=u_{1:k}$에서 첫 번째 심벌 u_1과 u_j를 서로 교환하여 생성된 순열을 연결하는 에지이고, 에지 발생기 $S_{n,i}(u_{1:k})=u_1u_{(i-1)n+2:in+1}u_{n+2:(i-1)n+1}u_{2:n+1}u_{in+2:k}$는 순열 $U=u_{1:k}$에서 심벌 시퀀스 $u_{2:n+1}$과 $u_{(i-1)n+2:in+1}$를 서로 교환하여 생성된 순열을 연결하는 에지이다. 에지 발생기 $S_{n,i}$를 간단히 S_i라 한다. 여기서 심벌 시퀀스 $u_{(i-1)n+2:in+1}$를 클러스터라 하고, 한 클러스터를 구성하는 원소의 개수는 $MS(l,n)$에서 n개이다.

매크로-스타 그래프 $MS(l,n)$의 노드에서 l-번째 클러스터의 마지막 심벌이 $p(l \leq p \leq k)$를 갖는 노드들로 구성된 그래프를 super-cluster(l,p)로 정의한다. 그리고 super-cluster(l,p)에서 l-번째

클러스터를 구성하는 노드들의 심벌이 가장 우측부터 $p_n, p_{n-1}, p_{n-2}, \ldots, p_1$로 구성된 노드들을 super-cluster($l, p_1p_2p_3 \ldots p_n$)이라 하자. super-cluster(l,p)에서 p를 구성하는 심벌이 n개이면 super-cluster(l,p)를 구성하는 노드들은 매크로-스타 그래프 $MS(l,n)$에서 l-번째 클러스터를 기준으로 한 그래프의 서브그래프이다. <그림 1>은 매크로-스타 그래프 $MS(2,2)$를 두 번째 클러스터를 기준으로 구성된 노드를 표현했다. <그림 1>의 우측 큰 원은 두 번째 클러스터의 심벌이 23으로 구성된 노드들로써, 145로 표현된 노드의 순열을 모두 표현하면 14523이고, 순열 14523에서 에지 발생기 $T_2(14523)$에 의해 생성된 순열은 41523이고, 에지 발생기 $T_3(14523)$에 의해 생성된 순열은 54123이다. 그리고 어떤 순열 P에 에지 발생기 T_j와 S_i를 순차적으로 적용한 경우 $S_i(T_j(P))$로 표현한다. 예를 들어 $S_2(T_2(14523))$에 의해 생성된 순열은 먼저 순열 14523에 에지 발생기 T_2에 의해 41523을 생성하고, 순열 41523에 에지 발생기 S_2를 적용하여 42315 순열에 이르게 된다.

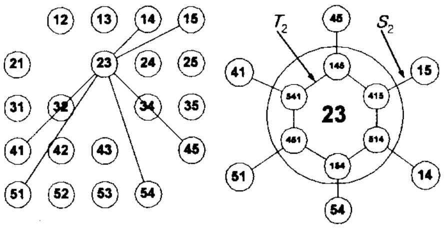

그림 1 매크로-스타 그래프 $MS(2,2)$

3. 방송 알고리즘

본 논문의 방송 알고리즘에서 메시지를 가진 노드 p에 에지 발생기 시퀀스 $\langle S_2, S_3 \rangle$을 적용한다는 것은, 방송의 첫 번째 단위 시간에는 노드 p에서 에지 발생기 시퀀스의 첫 번째 에지 발생기 S_2에 의해 연결된 노드 $S_2(p)$로 메시지가 전송되고, 두 번째 단위 시간에는 메시지를 갖고 있는 2개의 노드 p와 $S_2(p)$에서 에지 발생기 시퀀스의 두 번째 에지 발생기 S_3에 의해 연결된 노드 $S_3(p)$와 $S_3(S_2(p))$로 메시지가 각각 전송됨을 의미한다.

방송 알고리즘의 개요는 다음과 같다. 매크로-스타 그래프 $MS(l,n)$을 계층적 분할 성질을 이용하여 $k*k-1*\cdots*k-(n-1)$개의 매크로-스타 그래프 $MS(l-1,n)$로 분할한다. 그리고 매크로-스타 그래프 $MS(l,n)$의 메시지를 가진 한 노드에 에지 발생기 시퀀스 $\langle S_2, S_3, \ldots, S_l, T_2, T_3, \ldots, T_{n+1}, S_l, T_{n-m+1}, S_l \rangle$을 순차적으로 적용하여 분할된 각각의 매크로-스타 그래프 $MS(l-1,n)$의 한 노드로 메시지를 전송한다.

보조정리1 매크로-스타 그래프 $MS(l,n)$에서 메시지를 가진 한 노드에 에지 발생기 시퀀스 $\langle S_2, S_3, \ldots, S_l, T_2, T_3, \ldots, T_{n+1}, S_l \rangle$을 순차적으로 적용하면 매크로-스타 그래프 $MS(l,n)$의 노드에서 가장 우측 심벌이 p로 구성된 k개의 super-cluster(l,p) 각각에서 적어도 한 노드에 메시지를 전송할 수 있다($1 \le p \le k$).

증명 매크로-스타 그래프 $MS(l,n)$의 한 노드 U의 순열을 $U = u_1u_{2:n+1}u_{n+2:2n+1}u_{2n+2:3n+1} \ldots u_{(i-1)n+2:in+1} \ldots u_{(l-1)n+2:k}$이라 하자. 노드 U의 순열에 에지 발생기 시퀀스 $\langle S_2 \rangle$를 적용한 노드 $S_2(U)$의 순열은 $S_2(U) = u_1u_{n+2:2n+1}u_{2:n+1} \ldots u_{(l-1)n+2:k}$이고, 노드 U의 순열에 에지 발생기 시퀀스 $\langle S_3 \rangle$를 적용한 노드 $S_3(U)$의 순열은 $S_3(U) = u_1u_{2n+2:3n+1}u_{n+2:2n+1}u_{2:n+1} \ldots u_{(l-1)n+2:k}$이다. 이와 유사한 방법으로 노드 U의 순열에 에지 발생기 시퀀스 $\langle S_i \rangle$를 적용한 노드 $S_i(U)$의 순열은 노드 U의 첫 번째 클러스터와 i-번째 클러스터가 서로 교환된 순열 $S_i(U) = u_1u_{(i-1)n+2:in+1}u_{n+2:2n+1} \ldots u_{2:n+1} \ldots u_{(l-1)n+2:k}$이다. 노드 U의 순열에 에지 발생기 시퀀스 $\langle S_2, S_3, \ldots, S_i, \ldots, S_l \rangle$을 각각 적용한 순열들의 두 번째 심벌 즉, 노드 U의 순열에 에지 발생기 시퀀스 $\langle S_2, S_3, \ldots, S_i, \ldots, S_l \rangle$을 각각 적용한 순열의 첫 번째 클러스터의 시작하는 심벌들은 u_{n+2}, u_{2n+2}, u_{3n+2}, ..., $u_{(i-1)n+2}$, ..., $u_{(l-1)n+2}$이다. 노드 U, 그리고 노드 U의 순열에 에지 발생기 시퀀스 $\langle S_2, S_3, \ldots, S_i, \ldots, S_l \rangle$을 각각 적용한 순열들에 에지 발생기 시퀀스 $\langle T_2 \rangle$를 적용하게 되면 메시지를 가진 순열의 첫 번째 심벌 u_1과 각 순열의 두 번째 심벌 즉, u_2, u_{n+2}, u_{2n+2}, u_{3n+2}, ..., $u_{(i-1)n+2}$, ..., $u_{(l-1)n+2}$이 서로 교환된 순열들이 메시지를 전달받게 된다. 노드 U에 에지 발생기 시퀀스 $\langle T_2 \rangle$를 적용하여 메시지를 전달받은 노드 $T_2(U)$의 순열은 $u_2u_1u_{3:n+1}u_{n+2:2n+1} \ldots u_{(l-1)n+2:k}$이고, 노드 U의 순열에 에지 발생기 시퀀스 $\langle S_2 \rangle$를 적용한 노드 $S_2(U)$의 순열에 에지 발생기 시퀀스 $\langle T_2 \rangle$를 적용하여 메시지를 전달받은 노드 $T_2(S_2(U))$의 순열은 $u_{n+2}u_1u_{n+3:2n+1}u_{2:n+1}u_{2n+2:3n+1} \ldots u_{(l-1)n+2:k}$이고, 노드 $S_i(U)$에 에지 발생기 시퀀스 $\langle T_2 \rangle$를 적용하여 메시지를 전달받은 노드 $T_2(S_i(U))$의 순열은 $u_{(i-1)n+2}u_1u_{(i-1)n+3:in+1}u_{n+2:2n+1} \ldots u_{2:n+1} \ldots u_{(l-1)n+2:k}$이다. 이와 유사하게 메시지를 가진 노드들에 에지 발생기 시퀀스 $\langle T_3, \ldots, T_{n+1} \rangle$을 순차적으로 적용하면 메시지를 가진 노드들의 순열에서 첫 번째 클러스터의 n-번째 심벌들이 u_1, u_2, u_3, ..., u_i, ..., u_k를 갖는 순열들이 적어도 하나씩 존재한다. 메시지를 가진 노드들에 에지 발생기 시퀀스 $\langle S_l \rangle$을 적용하면 첫 번째 클러스터와 l-번째 클러스터를 서로 교환하게 되어 메시지를 가진 노드들의 k-번째 위치에 심벌 u_1, u_2, u_3, ..., u_i, ..., u_k를 갖는 순열들이 적어도 하나씩 존재한다. 따라서 매크로-스타 그래프 $MS(l,n)$에서 메시지를 가진 한 노드에 에지 발생기 시퀀스

$<S_2,S_3,...,S_l,T_2,T_3,...,T_{n+1},S_l>$을 순차적으로 적용하면 매크로-스타 그래프 $MS(l,n)$의 노드에서 가장 우측 심벌이 p로 구성된 k개의 super-cluster(l,p) 각각에서 적어도 한 노드는 메시지를 갖고 있다 $(1 \le p \le k)$. □

보조정리2 매크로-스타 그래프 $MS(l,n)$에서 l-번째 클러스터를 구성하는 노드들의 심벌이 가장 우측부터 $p_m,p_{m-1},...,p_1$로 구성된 super-cluster$(l,p_1p_2...p_m)$의 한 노드에 에지 발생기 시퀀스 $<S_2,S_3,...,S_l,T_2,T_3,...,T_{n+1},S_l,T_{n-m+1},S_l>$을 순차적으로 적용하면 super-cluster$(l,pp_1p_2...p_m)$ 각각에서 적어도 한 노드는 메시지를 전송 받는다$(1 \le m < n,\ 1 \le p \le k,\ p_i \ne p, 1 \le i \le m)$.

증명 매크로-스타 그래프 $MS(l,n)$의 super-cluster$(l,p_1p_2...p_m)$를 구성하는 노드들은 노드를 나타내는 순열에서 l-번째 클러스터를 구성하는 심벌이 가장 우측부터 $p_m,p_{m-1},...,p_1$로 구성된 노드이다. 이러한 노드들에서 임의의 한 노드를 $U=u_1u_{2:n+1}u_{n+2:2n+1}u_{2n+2:3n+1}\ ...\ u_{(i-1)n+2:in+1}\ ...\ u_{(l-1)n+2:k}$ 라 할 때, 노드 U에 에지 발생기 시퀀스 $<S_2,S_3,...,S_l,T_2,T_3,...,T_{n+1},S_l,T_{n-m+1},S_l>$을 순차적으로 적용하면 l-번째 클러스터를 구성하는 노드들의 심벌이 가장 우측부터 $p_m,p_{m-1},...,p_1,p$로 구성된 super-cluster$(l,pp_1p_2...p_m)$를 구성하는 노드들에서 적어도 한 노드가 메시지를 전달받음을 보인다. 노드 U에 에지 발생기 시퀀스 $<S_2,S_3,...,S_l,T_2,T_3,...,T_{n+1}>$을 순차적으로 적용하면 보조정리 1에서와 유사하게 메시지를 가진 노드들의 순열에서 첫 번째 심벌이 $u_1, u_2, u_3, ..., u_i, ..., u_k$를 갖는 순열들이 적어도 하나씩 존재한다. 메시지를 가진 노드들에 에지 발생기 시퀀스 $<S_l>$을 적용하면 첫 번째 클러스터와 l-번째 클러스터를 서로 교환하게 되어, 에지 발생기 시퀀스 $<S_l>$를 적용한 노드들의 첫 번째 클러스터의 심벌들은 가장 우측부터 $p_m,p_{m-1},...,p_j,...,p_1$로 구성되어 있고, 노드의 첫 번째 심벌은 $u_1, u_2, u_3, ..., u_i, ..., u_k$를 갖는다$(p_j \ne u_i)$. 이제 메시지를 가진 노드들에 에지 발생기 시퀀스 $<T_{n-m+1}>$를 적용하면 메시지를 가진 노드들의 첫 번째 심벌과 첫 번째 클러스터의 가장 우측부터 $m+1$번째 위치의 심벌이 서로 교환된다. 메시지를 가진 노드들에는 첫 번째 클러스터의 가장 우측부터 $m+1$번째 위치에 $u_1, u_2, u_3, ..., u_i, ..., u_k$를 갖는 노드들이 적어도 하나씩 존재하게된다. 에지 발생기 시퀀스 $<S_l>$를 적용하면 첫 번째 클러스터와 l-번째 클러스터가 서로 교환된 순열들에 메시지를 전달하게 되어 super-cluster$(l,pp_1p_2...p_m)$를 구성하는 노드들에서 적어도 한 노드가 메시지를 전달받음을 알 수 있다. □

[One-to-All Broadcasting Algorithm]

Begin

[단계1] 매크로-스타 그래프의 메시지를 가진 노드에서 super-cluster(l,p)의 한 노드로 메시지를 전송한다$(1 \le p \le k)$.

[단계2] 메시지를 가진 super-cluster(l,p)의 노드에서 super-cluster$(l,p_1p_2...p_m)$의 모든 노드로 메시지를 전송한다$(1 \le m < n,\ p_i \ne p, 1 \le i \le m)$.

[단계3] 위의 단계1과 단계2 과정을 l-번째 클러스터부터 2-번째 클러스터까지 반복한다.

[단계4] 매크로-스타 그래프 $MS(l,n)$의 기본 모듈인 스타 그래프 S_{n+1}은 스타 그래프의 방송 알고리즘을 적용한다.

EndBegin

4. 결론

병렬 컴퓨터의 위상으로 잘 알려진 스타 그래프는 짧은 지름, 노드 대칭성, 계층적 구조와 최대 고장 허용도 등을 갖는 상호 연결망이다. 최근에 제안된 매크로-스타 그래프는 이러한 스타 그래프를 기본 모듈로 하면서 스타 그래프보다 망 비용이 개선된 상호 연결망이다.

본 논문에서는 매크로-스타 그래프의 계층적 분할 성질과 이미 제안된 스타 그래프에서의 방송 알고리즘을 이용한 매크로-스타 그래프에서의 방송 알고리즘을 제안하였다.

참고문헌

[1] S. B. Akers and B. Krishnamurthy, "A Group-Theoretic Model for Symmertric Interconnection Network," IEEE Trans. Comput., Vol. 38, No. 4, pp. 555-565, 1989.

[2] S. M. Hedetniemi, S. T. Hedetniemi, and A. L. Liestman, "A Survey of Gossiping and Broadcasting in Communication Networks," Networks, Vol. 18, pp. 319-349, 1988.

[3] J-H. Park, Circulant Graphs and Their Application to Communication Networks, Ph.D. Thesis, Dept. of Computer Science, KAIST, Taejon Korea, 1992.

[4] V. E. Mendia and D. Sarkar, "Optimal Broadcasting on the Star Graph," IEEE Trans. Parallel and Distributed Systems, Vol. 3, No. 4, pp. 389-396, July 1992.

[5] C. H. Yeh and E. A. Varvarigos, "Macro-Star Networks: Efficient Low-Degree Alternatives to Star Graphs," IEEE Trans. Parallel and Distributed Systems, Vol. 9, No. 10, pp. 987-1003, October 1998.