

Title	エクサスケールコンピューティングのためのエネルギー効率を考慮した新しい階層型相互結合網
Author(s)	Faisal, Faiz Al
Citation	
Issue Date	2018-09
Type	Thesis or Dissertation
Text version	ETD
URL	http://hdl.handle.net/10119/15533
Rights	
Description	Supervisor:井口 寧, 情報科学研究科, 博士

氏 名	FAISAL, Faiz Al		
学 位 の 種 類	博士(情報科学)		
学 位 記 番 号	博情第 401 号		
学 位 授 与 年 月 日	平成 30 年 9 月 21 日		
論 文 題 目	A New Hierarchical Interconnection Network with considering Efficient Energy Usage for Exa-scale Supercomputing		
論 文 審 査 委 員	主査	井口 寧	北陸先端科学技術大学院大学 教授
		金子 峰雄	同 教授
		田中 清史	同 准教授
		リム 勇仁	同 准教授
		三浦 康之	湘南工科大学 工学部 教授

論文の内容の要旨

The requirement of high processing power is enormous; one of the high desire of next generation HPC systems. Since every computer chip has limited processing power, sequential processors can't be the suitable choice. For example, an Intel Core i7-3630QM processor (4 cores, 22nm fabrication process) can achieve about 76.8GFlops through the requirement of 45W electrical power usage. However, the requirement for exa-scale computing will require about 13 million of connecting such processors. Today's most powerful supercomputer Sunway Taihulight System has already achieved about 93 petaFlops performance with 10,649,600 cores requiring about 15.3MW electrical power using low degree 2D Mesh interconnect (2D Mesh has the performance constraints and faster saturation rate). Moreover, a high degree network like- Tofu (6D Mesh/Torus) interconnect used in K-computer has achieved 10.51 petaFlops performance (88,128 SPARC64 VIIIfx processors, Tofu interconnect with 10 cores to cores connectivity) by requiring 12.6MW of electrical power. Hence, to build an exa-flops Tofu system, it will require near about 1260MW of electrical power with the current advancements. These observations confirm that conventional structures are not feasible for the next generation networks due to the high power usage for the high degree core to core connectivity (Tofu interconnect) and also shows poor network performances (2D Mesh interconnect). Hence, the possible solution to reach the next generation exa-scale performance is to redesign the "Interconnection Network".

Exa-scale supercomputing requires network scalability over millions of cores and the performance constraint affects heavily for the large system along with the total power usage. In considering those constraints our focus resides on the "Hierarchical Interconnection Networks (HIN)". HINs possess the features like- constant node degree, small average distance, better bisection width, small number of wires and low network latency with high throughput. Constant node degree ensures the fixed router cost throughout the entire system, small average distance eliminates the possibility of the performance degradation, better bisection width ensures the

network traffic handling capability, wiring complexity is effective for reducing the network power usage and finally, network latency ensures the packet reachability with the requested traffic load. This research also considers a new parameter of "Network Energy Usage" to ensure the high performance and the low power usage. Moreover, we have considered two possible network configuration of 65K cores and 1M cores analysis to ensure the superiority of our network for the exa-scale system.

Keywords: Interconnection Network, Hierarchical Flattened Butterfly Network, Estimation of Power Consumption, Dynamic Communication Performance, Energy Usage.

論文審査の結果の要旨

本研究は近年非常な大規模化が進展している超並列計算機の相互結合網に関して新しい提言をおこなうものである。近年では大規模なスーパーコンピュータは多数のマイクロプロセッサを結合し構成されているが、その規模は年々増大している。例えば京では 8 コアのチップが 82,944 ノード結合され、全体で 66 万コアのシステムとなっている。規模の他に、結合の物理的媒体が多岐にわたる点も課題となっている。チップ内、PCB ボード上のチップ間の銅線結合、キャビネット間は光ファイバーによる結合など、物理階層ごとに様々な結合手段が用いられている。さらに、京では消費電力が 15kW と膨大であり、次世代システムでは消費電力の低減が大きな課題である。これらを統合的に評価し、かつ性能あたりの消費電力が少ない総合結合網が求められている。

本論文では、最初に相互結合網に関する状況や指標について俯瞰した。重要な指標として、消費電力×平均遅延として、性能あたりの消費電力の指標を導入した。階層型の考えを導入し、最下位層(チップ内結合に相当)を Basic Module(BM)と定義した。BM の最適なトポロジを実装性も考慮しつつ検討し、2 次元コーダリングとも言える Flatten Butterfly 網 が性能あたり消費電力最低となることを示した。次に上位階層の結合について検討し、同様に 2 次元トーラス網が上位階層に適している事を示した。これら Flatten Butterfly 網と 2 次元トーラス網の組み合わせとして HFBN (Hierarchical Flatten Butterfly Network) を定義した。HFBN は各階層の大きさや、階層の高さ、階層間結合の結合度に応じて複数の構成が考えられるが、その構成について検討を行った。また、上位階層への取り出しポートの定義など、HFBN を定義付ける構成アルゴリズムを示した。さらにデッドロックフリールーティングアルゴリズムなど、HFBN を実用的に利用できる方式も示した。

次に提案した HFBN を他の幾つかの相互結合網と比較した。2D メッシュ・トーラス、既存網の MMN や TESH、実用的な網として TOFU などと比較し、大規模な結合の場合は有意に性能当たり消費電力が小さいことが示された。本論文での一つの功績は、従来は単一の物理条件、つまり全てのノードが同一の条件で測定されていたのに対し、光ファイバー

による結合や銅線での結合など，異なる物理条件でのシミュレーションを仔細に行った点である．理論的な側面だけでなく，現実の大規模システムを構築する際に近い条件でのシミュレーション結果が示されている．

これら一連の業績は，次世代の大規模並列システムの相互結合網に求められる性能指標を評価するという点で有意な実用上の方向性を示したものである．博士学位論文として必要な水準を満たしていると認める．