遅れを取った日本の IoT ハードウェアの対応策

大塚 寛治*, 佐藤 陽一*, 橋本 薫*, 藤井 文明*, 秋山 豊*, 上田 千寿*

To Build a Position of IoT Hardware in which Japan Lags Behind

Kanji OTSUKA*, Yoichi SATO*, Kaoru HASHIMOTO*, Fumiaki FUJII*, Yutaka AKIYAMA*, and Chihiro UEDA*

* 明星大学連携研究センター (〒191-8506 東京都日野市程久保 2-1-1)

* Collaborative Research Center, Meisei University (2-1-1 Hodokubo, Hino-shi, Tokyo 191-8506)

1. はじめに

IoT のメガデータセンターはすべてアメリカの会社が支配している。その強さは過去からの膨大なデータの蓄積があり、これが財産となって、すべてのデータを有機的につなげて利用することで、他の追従を許さない環境となってきた。サービス面でこれを覆すことは今や不可能に近い。

ハードウェアはどうであろうか。彼らにはいくつかの弱点があり、代替技術を彼らに提案は出来そうである。うまくいけば iPhone のように日本のプレゼンスを強めることになろう。メガデータセンターのハードウェアの仕組みが分からないと強める技術とならない。仕組みを理解した上でいくつかの例題を示す。しかし支配体制は変わらない。支配体制を変える新しいアプローチも考えてみよう。その形も示したい。

2. IoT の動向

Economist 2016 年 9 月 17 日¹⁾のレポートから、世界の会社の総資産のランキングベスト 10 位を見よう。IoT 会社のみをピックアップする。1 位 Apple, 2 位 Alphabet (Google), 3 位 Microsoft, 6 位 Amazon, 7 位 Facebook である。10 年前の 2006 年では 4 位 Microsoft のみである。今まで支配していた石油会社、金融会社や自動車会社を押しつけて IoT 会社が席卷し、その伸び率は衰えていない。この中に日本の片鱗すら窺えない。これを利用するばかりであり、ハードウェアがどのようなになっているか関心がない状態である。実は京コンピュータが 2 年前は世界のトップ性能を誇っていた。データセンターの機能的に分解するとほとんど同じ (次章参照) であるにもかかわらず、世界を支配できなかったのはなぜか。世界の演算・通信機能を見渡した時の次のマーケットが「何たるか」に対して関連付けなかっただけである。

Amazon の例を見よう。本の通信販売をアメリカ国内で始め、最短納入を狙うには最短距離配送と在庫管理がキーであると考えた。大きなサーバ基地を適切に配置し、潤沢

な通信網が必要となるので、それを独自で構築し、全世界へと展開した。当然ほかの商品の通信販売も手掛けた。大きく発展したシステムの賃貸しと管理ソフトも売り込めると判断し、今日の IoT の雄になった。Google は検索システムで、いまの Internet 検索をリードし、これをもとに IoT の雄になった。Microsoft はソフトウェアの強みから、大きなイベントを丸抱え管理する事業から展開してきたようで、ロシアのイベント後進をチャンスととらえ、ソチオリンピックを丸抱え管理したことで有名になり、IoT の雄の仲間入りをした。今や 2 番手であり、伸び率が一番高い。

3. IoT のハードウェアの機能

IoT のハードウェアは汎用プロセッサが頭脳であり、幾つかのボードが単位となったプロセッサノードが重なって、ラックが形成され、これが体育館のような大きな建物の中に数百～数千並んだ構成となっている。別に同様構成のルーター・スイッチが配置され、隣接の通信会社とつながっている。画像処理プロセッサは効率よく上記ラックの中に存在し、繰り返し演算を必要とするプログラムが実行される。それぞれのラック内、または隣接ラックに記憶装置である SSD (以前は HDD が主力) が配置された構成である。

一つの構成例を筆者が参画したプロジェクト内容で説明しよう。すなわち、JAMSTIC-NEC-東北大グループのネクスト京プロジェクト調査研究の立案を示すと Fig. 1 のようになる²⁾。IoT のデータセンターもほぼこれと同様の構成であるが、それぞれの会社が独自開発したものであり、過去内容が公表されたものはない。

Fig. 1 で示されたノードの中身を機能的にまとめると Table 1 のようになろう。ハードウェアの要素はデータセンターのエンタープライズサーバーもスーパーコンピュータもエンタープライズ通信機器もほとんど同じであり、異なるところは、汎用プロセッサを除きそれぞれの目的に応じて柔軟性を犠牲にすることで、最大性能を引き出したところにある。

Fig. 1 Next K Preliminary Research Project by JAMSTIC-NEC-Tohoku Univ. Group

Table 1. Functions of processor node

	General processor	Router/Switch	Picture processor
Logical processing	CPU	NPU	GPU
Processing controller	Transfer lookaside buffer	Packet buffer/Search engine	Pixel engine
Processing memory	Main memory	Table memory	Main memory
High-speed bus	PCI Express etc.	XAUI, CAUI	PCI Express etc.
Crossbar switch	CPU (core) connection	NPU connection	GPU (core) connection
High speed treatment	Multicore/Super scaler	Parallel processing	Multicore
Database	SSD (HDD)	SSD	SSD

4. 目的別 IoT データセンター分類

まず、ビッグデータをハンドルするデータ保存の専門機関が存在する。Data Direct Networks (DDN)³⁾ がその代表的な機関で世界の 14% のシェアを占めている。そこには、アメリカのマイナンバー的なデータが入力されているだけでなく、世界の注目すべき人物はほぼリストアップされていて、FBI, CIA も利用している。あらゆる情報が収集の対象であり、医療機関の患者個人情報が多いと言われている。その中の公開情報はすべて一般利用可能である。ユーザーのアクセスデータ量は 200 PB という膨大なもので、ここまでビッグデータを集めると、代行できる機関がないという状態となる。これは今後も脅威となる現象である。Storage Fusion Architecture という多くの SSD を一体に見えるようにし、アクセス速度は Amazon の 10 倍速いと自身が宣伝している。ちなみにこの 2 番手は IBM である。

IoT 基地の巨大なものをメガデータセンターと呼んでいるが 2016 年の初めは世界で 146 基地があり、どこにどの程度の基地があるかは機密となっている。それぞれの分野のリーダを示すと Table 2 のようになる。PaaS は新規・既存のアプリケーションを統合するスケラブルなプラットフォーム、SaaS は、必要な機能を必要な分だけサービスとして利用できるようにしたソフトウェア、IaaS はコンピューターシステムの構築に必要な、ハードウェアなどの

Table 2. Mega data center leaders on each genre

Genre	Giant company (protocol)	Job function
PaaS (Platform as a Service)	Amazon (AWS), Microsoft (Azure), Google (App Engine)	Own hardware and software service
SaaS (Software as a Service)		Divided to ECS, CRM and ERP
ECS (Enterprise Collaboration Software)	Microsoft (Office 365)	Own software based service
CRM (Customer Relationship Management)	Salesforce	User oriented software service
ERP (Enterprise Resource Planning)	Microsoft (365 dynamic)	User resource collaborative software service
IaaS (Infrastructure as a Service)	Amazon (Elastic Compute Cloud (EC2)), Google (Compute Engine), Microsoft (Azure), Oracle (Bare Metal Service)	user based platform and software

基盤（インフラストラクチャー）一式が利用できる環境を、インターネット経由で提供するというサービスで、日本の利用率で示すと SaaS が一番多く、IaaS が続き、PaaS が最後であるが、ほとんど 1/3 の比率である。

各々のリーダはその分野におけるビッグデータを保有し、それを有機的に利用することで強みを発揮している。たとえば Google は猫という画像認識に役立つための画像数が 100 万シーン以上あり、適切分類であるがゆえに目的別に即有効利用できる。ディープラーニングに今後重要な資料となる。

もう一つは通信機関である。メガデータセンター間の通信⁴⁾ は最大のシェアで Equinix がハンドルしている。最近老舗の AT&T が巻き返しを図っているようであるが、Equinix はメガデータセンターの隣に通信基地を設けてサービスしている強みがある。東京でも 8 つほどのメガデータセンターがあり、隣接して Equinix が構えている。大部分が光ケーブル（海底光ケーブル含む）で繋がれており、無線はほとんど使っていないため、セキュリティは万全であるとしている。もし Equinix に政治的意図があれば、世界を変えることができ、中国は早くからその脅威を感じ China Telecom Next Generation Carrier Networkなどを立ち上げ、最近ロシアも対抗策を講じているとのことである。

5. ハードウェア性能

上記 JAMSTIC グループで検討した性能ベンチマークを Fig. 2 に示す。シミュレーションレベルであるが、256 GF のコアを 4 つと 32 MB のキャッシュメモリを Si インターポーザ上に乗せ、これを 4 セットボードに乗せ、1 TF の CPU 性能をつける。そのボードの周辺に TSV 積層の 128-512 GB のメインメモリを配置し、4,096 本の 32 Gbps で伝送可能な高速バスで CPU とつなげ、1-2B/F という驚異的バンド幅で接続する（左上図）。Intel の Xeon（数 10 GF）を 15 チップ集合させた比較例（右上図）は、数 TF の性能である（右上図）。メモリ環境は同じにした。しかし Xeon

Fig. 2 Bandwidth made high performance/reduction power¹⁾

のバス制限から 0.1B/F の接続となる。この B/F に違いが性能に大きな差を生んでいる。ベンチマークソフトによるが 10 倍から 80 倍の性能がバス幅の増大のみで得られる。並列度の高いソフトが 10 倍と少なく、ランダム処理では高性能となる。ここで強調すべきことは CPU の性能でなく、頻度の高いデータ通信が性能を決めるということである。

6. ビッグデータ環境でのシステム性能

Fig. 1 の右図は大量の情報がそれぞれのラックに貯蔵されている。このように分散した情報を有機的につなげるため、よく用いられている方法は 3 次元的なネットワーク (3D Torus) である。分散されている各ユニット間 (ラックでもノードでもよい) のデータ通信がこれで Fig. 3 (左) のようにフラットになるが、効率的ではない。右図のようになればユニット間のデータ通信量は大幅に節約できることになる。この方法が、ビッグデータの中のプロセッシング方式として最も重要で、Fig. 2 のようなメインプロセッシングエンジンの性能以外の要素が大きくなる。

このためにまず重要なことがデータ貯蔵ユニットのスマート化、すなわちデータ貯蔵ユニット単位でプロセッシング機能を持たせることである。具体的な構造として、Fig. 4⁵⁾、Fig. 5⁶⁾ の方式が考えられる。

Fig. 4 はメモリユニット内にプロセッサがあり、メモリユニット内で処理できるものはすべて処理して必要なデータだけメインプロセッサに送り込む方法である。

Fig. 5 は隣接関連データの累積パスのみを送り込むことで、ツリー状にデータ配列をすれば、バンド幅を節約できる。東芝の SSD モジュール⁷⁾が Fig. 5 に近いメモリユニット間の直接通信を行い、高速化を達成している。

いずれにしてもメモリユニット内で有効な情報をグループ化する必要がある。たとえば意味情報をグループ化する例を示すと Fig. 6 のようになる。

この方式にはアプリケーションに応じて多数の手法が考えられている。それを列挙すると、Table 3 のようになる。

Fig. 3 Consideration of storage classified networking

Fig. 4 Traffic quantity reduction by PIM²⁾

Fig. 5 Neighbor Discovery Protocol³⁾

Fig. 6 One example for explicitly classify processing (meaning networking topology)

Table 3. Efficient grouping method

#	Protocol
1	Frame model
2	Object model
3	Production system
4	Message passing interface
5	Persistent file
6	Meaning networks
7	Graph consideration
8	Application file

Table 4. Top100 Super Computer Ranking⁶⁾

Top 100 Super Computer Ranking								
Rank	Name	Actual Performance (PFlops)	Physical Performance (PFlops)	Power Consumption	Number of Cores	Core Name	Cost	Reference
1	Sunway TaihuLight (神威太湖之光)	93.01	125.43	15,371 kW	10,649,600	Original (DEC base?)	27B ¥	1.45 GHz, 28 nm
2	Tianhe-2 (天河二号)	33.86	54.9	17,808 kW	3,120,000	Xeon Phi	10B ¥	
3	Titan	17.59	27.11	8,209 kW	560,640			
4	Sequoia	17.17	20.13	7,890 kW	1,572,864			
5	K (京)	10.51	11.28	12,660 kW	705,024	SPARC	100B ¥	1,400 億 ¥Next 京
15	Tsubame KFC	5.7	17	省エネ 1 位	現在 3 位	Nvidea GPU	3.2B ¥	Green500

Fig. 7 Good performance comparison example by effective classified memory

それぞれがソフトウェア処理であるため、ここでは列挙にとどめる。

結果として、性能が世界のトップになった例を Table 4, Fig. 7 で紹介する。表から見て 1 位と 2 位は中国製である⁸⁾。2 位の天河は Intel Xeon Phi を使用し、並列に処理しているが、Xeon の制限から 0.14B/F しかバンド幅は確保できない。すでに大きなデータ貯蔵の中で、明示的に各ユニットに配分していると思われる。昨年発表された、1 位の神威太湖之光は Fig. 7 (右) のようにさらに貧弱な 0.04B/F となっているにもかかわらず、あらゆるベンチマークテストでトップになっている。データ貯蔵庫へのデータ群の明示的配分とスマート化による工夫が世界のトップを走っていると思われる。ちなみにアメリカ政府の圧力で、中国は Intel の支援が得られなかったといううわさがあり、プロセッサに対して独自開発となった。ただし DEC ベースのアーキテクチャのようである。

もう一つの例はすでに示した DDN である。Amazon の 10 倍の速さで、膨大なデータアクセスが可能になる背景は前述の Storage Fusion Architecture という方式がキーとなっている。Fig. 4 の構造とソフトウェアで多くの SSD ラック間を一体化して取り扱うことで高速アクセス可能となったと思われる。

7. 高速処理技術

いずれにしてもコミュニケーションのバンド幅は高いに越したことはない。Table 2 の IaaS でやっと顔を出している Oracle はデータセンター性能が Amazon より高いと Table 5⁹⁾ のように唱っている。

注目すべき数値が 11.5 倍の IO インターフェース性能である。その 36 コアのチップの IO 仕様はいくつかあるが、その一つに Coherency link lane が 28 Gbps の SerDes バスで

Table 5. Oracle new system comparison by Oracle Open World 2016.9

Metric	Oracle	Amazon highest	Ratio
Cores	36	16	2.25x
D-RAM	512 GB	244 GB	2x
SSD Storage	28.8 TB	6.4 TB	4.5x Capacity 11.5x IOPs
Run Expense	\$5.4/h	\$6.82/hour	1.2x
Bus rate	28 Gbps/lane	?	

用意されている。往復で数えて 12 レーン×7 セットあり、隣接のプロセッサと密結合されている。それでもデータ圧縮と解凍がおこなわれていて、高速性能が実現されている。この性能優位で現在データセンターのシェア拡大を狙っている。

このように高速信号伝送は性能改善に必須の要件であり、筆者グループの取り組みを少しだけ紹介する。

8. 高速信号伝送技術

10 Gbps 以上の高速信号伝達は従来の設計的な延長線では解がないことを示そう。まず、プリント配線板で、通常設計はプレーナランドに対して貫通させる信号 (500 μ m 幅、長さ 5 + 5 mm, FR-4) は Fig. 8 (左上) のようにランドに穴をあけて、その間にビアで通すことがよくおこなわれている。一方、右上図はランドを信号線と同じ幅で這わせて、信号の層間接続はメガネ枠状のランドを通した設計を考えてみよう。シミュレーションベースでそれぞれ比較して S パラメータと 19 GHz (38 Gbps) の電界分布を見たのが中図、下図である。このレベルでは明らかに従来設計が問題となることが判明する¹⁰⁾。

他にもいろいろ経験しているこれらの知識^{11)~13)}をもとにして、デンソーと共同開発した設計案¹⁴⁾を Fig. 9 に示す。絶縁材料は LCP であり、すべてインナービアで配線され、配線幅とビア径は同じでランドの大きさも極力ビア径に近づけた。線幅 100 μ m、厚み 18 μ m、線間隔 80 μ m の差動線であり、隣接差動配線との間隔は 300 μ m とした。これにより 300 mm 長さ (8 層貫通含む) で Fig. 10 のように 30 Gbps という高速伝送が可能となった。2013 年 (発表 2014) 当時世界最高レベルである。2015 年の Xilinx の発表¹⁵⁾で Fig. 11 のように同様構造が重要との指摘がなされている。高速信号系のアプローチはこのような形に収斂さ

Fig. 8 Ground design treatment on high bandwidth

Fig. 9 High speed design feature for PCB

Fig. 10 Random eye pattern (30 Gbps) of 300 mm differential line in LCP board

Fig. 11 Xilinx design for 56 Gbps (PAM4) bus structure improvement

れると思われる。

もう一つ重要な要素があり、それを Fig. 12 で示す。FR-4 基板で相似縮小を行ったストリップ線路特性検討の実験である。5 Gbps という低いレートであるが、210 mm の長さでは一番幅の広い $w = 0.8$ mm の線路のみ信号伝達ができている。微細配線は高周波抵抗損で伝達しないことがわかる。

その経験から、われわれの 30 Gbps レベルの設計ガイドラインは Table 6 のように定めた。微細にすれば短い配線し

Fig. 12 The transmission performance depend on the line width on the FR-4 210 mm micro-strip feature

Table 6. Design guideline for 30 Gbps on 300 mm length of LCP or BCB

Wiring length	Line width
0.3 mm	0.1 μ m
3 mm	1 μ m
30 mm	10 μ m
300 mm	100 μ m

Fig. 13 One of co-design circuit for the simulation with 20 Gbps

か許容できない。チップレベルの 0.1 μ m 配線幅では 0.3 mm 長配線となる。

このような設計を踏まえて、ドライバとレシーバの整合性を考えた Fig. 13 のような Co-design が重要となり、パッケージ、実装レベルだけでは閉じない知識展開が必要となる。筆者研究室は CPU やメモリ設計も可能な総合知識を融合させている。

9. 遅れを取り戻すための施策

8 章で記述した高速信号技術は日本で第一線の研究がおこなわれていて、メガデータセンターの企業は魅力のある存在になっている。特に低損失の、絶縁材料は日本が受け持つことが今でもできる。

しかし、日本が強かった高密度実装技術は Micron や Hynix のスタックメモリ (HMC, HBM), Xilinx や Oracle の Si インターポーザ, TSMC-Apple の InFO (FO WLP), Intel-Altera の EMIB (Embedded) などすでに応用部門とタイアップして、高度実装技術が量産されていて、日本の入り込む

Fig. 14 Von Neumann processor trend

ことが技術的にも価格的にも難しい状態となっている。これらの材料と装置で、多少の強みがうかがえるが、主力となっていない。

日本の致命的な問題は、これらの応用分野の確かなニーズが分からないから、その開発目標が立たないことにある。そこで再びシステムの内容解説から始める。先ず、プロセッサが全てのキーである。その動向を示すと Fig. 14 のように概括できる。

CPU (上図) を1コアではなくマルチコアにする、さらに細かく分割して区分にしてマルチスレッド動作で高速化することがおこなわれているが、動作させる方法がメモリに格納されたプログラムとデータを読み出して、実行するため、その間の通信バンド幅の不足が生じることはすでに5章で述べたとおりである。

6章で述べたようにメモリをスマート化(メモリ単位でCPUを配置する)でバンド幅の制限を少なくした方法がメガデータセンターの方式となっている(下図)。今後はスマートフォンでもメインCPUと多く存在するアプリケーションプロセッサの世界でもこれらをメモリの中に分散させるスマートメモリ化で、バンド幅を節約することがなれると思われる。

しかし一番問題となるのが、CPUの消費電力でFig. 14に示したように効率よく動作させるほど、電力が必要となる。Microsoftが先鞭をつけて低消費電力のFPGA(Altera製)を演算のアクセラレータとして全面的に使用すると2016年初めに表明した。Intelは電力低減の方法論を模索していたので、FPGAの大2シェアを確保しているAlteraを買収した。一方NvidiaのGPUをアクセラレータとして用いる方法が、当面の高性能化の方法となっており、売り上げ額が増加しNvidiaの株価が2016年は急上昇したが、ここに来て電力消費の大きさで、限界を感じてきたユーザーが多くなって、株価が落ち込み始めている。

本章の冒頭で述べたようにIntelはCPU+FPGAをEMIBという手法でシステム化し、電力低減の新方向を世の中の普遍化技術にしようとしている。なぜFPGAは低消費電力であるかをここでひも解く。九九の答えテーブルをFig. 15にその一例を示すが、テーブルはメモリそのもので、XYアドレス(この場合3×8)をヒットするだけで答えが導くことができる。演算がいらないため高速で低消費電力となる。

Y address

X address

Hit X3

Hit Y8

Answered table

Y address \ X address	0	1	2	3	4	5	6	7	8	9
0	0	0	0	0	0	0	0	0	0	0
1	0	1	2	3	4	5	6	7	8	9
2	0	2	4	6	8	10	12	14	16	18
3	0	3	6	9	12	15	18	21	24	27
4	0	4	8	12	16	20	24	28	32	36
5	0	5	10	15	20	25	30	35	40	45
6	0	6	12	18	24	30	36	42	48	54
7	0	7	14	21	28	35	42	49	56	63
8	0	8	16	24	32	40	48	56	64	72
9	0	9	18	27	36	45	54	63	72	81

Fig. 15 An example of LUT

Fig. 16 Typical FPGA structure in case of Xilinx Vertex 5

性能の一例を示す。Open CLの同じ回路規模のベンチマーク¹⁶⁾を見ると、Xilinx Vertex 7vx690tはIntel Haswell CPUと比較して速度は11.5倍、電力は1/5とFPGAが優れている。これらの性能を達成するためのクロック速度も低くでき、CPUの2GHzにたいして、1/10の200MHzで達成している。演算が不要なためである。

ところがXilinx Vertex 5の構造を見ると、Fig. 16のようになる^{17),18)}。LUT(5入力、2出力)を含む論理回路とSRAM(2ポート)がストライプ状に配置されている。しかし、LUTの大きさがあまりにも小さいため、一つの機能を作るためにFig. 16下図に示すように多くのLUTを連結させる必要がある。このためクロスバー的にスイッチブロックが配置されそれらを連結している(Fig. 16右)。この連結回路が冗長である。

しかし、この回路があるため、初期設定で、いくらでも回路構成を変えられる。この操作を再構成(Reconfiguration)と呼んでいて、Field programmable gate arrayといわれる所以である。今後のシステムのニーズにマッチしている。従来のCPUはプログラムでこの再構成を行っているが、FPGAのようにハードウェア処理は圧倒的に速度が速くできる。この基本条件は十分配慮しなければならない要素である。

クロスバースイッチ経由で演算をしてもCPUより高速で、低消費電力となる。もしこのスイッチを経由しない機能が作れるLUT単位で処理できればもっと性能が上がり、

Fig. 17 Efficiency of LUT architecture processing

Fig. 18 Made-up system by 256 w X 8b cluster array

電力も節約できる。筆者の研究室で、検討した結果、Fig. 17のように最適構成があることを突き止めた。小さなFPGAのLUTは2ビット×2⁵からFig. 16で見るように2つのLUTがあるため128bのサイズとなる。これを基準にし、4b×2⁴(=64b)すなわち16w×4b、同様な構成で256w×8b、64kw×16ビットの相対的演算効率を表したもので、256w×8bが最小機能発現構造で、最も自由度の高い構成となる。64kw×16ビットは大きすぎて、無駄な演算が出てくることが判明した。

傍証として、Nvidiaが最近提唱している¹⁹⁾ように64bや32b(通常用いられているプロセッサのビット幅)より16bのGPUの方が効率よく、現在16bのGPUを開発中として、最終的には8bのGPUとする予定となっている。古い昔の8bの有名なIntel8008プロセッサに戻るのには、CISCで膨大になったソフト(命令語の複雑化)をRISCで単純化した歴史と似通っている。

Fig. 17の256w×8bをクラスタとしてタイル状に並べるとFig. 18のようになる²⁰⁾。このクラスタはメモリであり、LUTベースの論理でもある。Fig. 17にあるように周辺に配置したデコーダ(切り替え制御回路)で自在に変化させることができる。一例がFig. 18の右上に書かれたようにある一つのクラスタを論理にし、周辺をキャッシュにすれ

Fig. 19 Data center fabric concept

ば演算回路の出来上がりである。隣接配列のため、高速通信ができる。すなわち、バンド幅制限がなくなる。

ジョブサイズが大きくなれば、同図中央にあるように論理を並べることは自由であり、右下にあるように共通キャッシュを持つマルチタスクもできる。これが動的再構成(Dynamic reconfiguration)が可能になり、深層学習(Deep learning)の必須条件が完成する。もしどこかのクラスタが不良になれば、その時点で底を避ける構成にすればよい。強靱なシステムの出来上がりである。

NvidiaのGPUの8b化もこのようなタイル状(ただしメモリと論理回路は別)にするとと思われる。そこで、未来の実装を考えよう。Fig. 18を重ねればもっと高度なシステムを作れることが分かる。8bGPUでも論理とメモリを多重化すればよいはずである。MicronやHynixのようにDRAMでは価格の高いTSVを使って3D実装している。「TSVを排除する3D実装は如何」と考えればよい。

一つのアイディアを示そう。エピタキシャル層で全てのデバイスを構成できる。すなわち、エピタキシャル層を剥がしてこれを積層すれば多層(三次元)デバイスが可能と考える。Si基板からエピタキシャル層を剥がすプロセスはSOI製作で一つの方法となっている。一例としてSoitec社はSmart cut²¹⁾という手法で、エピタキシャル層を剥がしている。この超薄膜をハンドルするだけでなく検査技術も完備して、積層すればTSVなしの積層システムが出来上がる。ビルドアップ法で、積み重ねるアイディアはビルドアップPCBで経験するように歩留まり問題を解消できない。薄膜積層技術の方が、本質的に利点が多い。

システムの未来の方向が分かれば、実装プロセスの先取りができて、世界をリードできる。実装にかかわる部品材料はそのユーザ顧客に行き、その窓口の言いなりで終わる傾向にあり、全体像が見えないままその要求に応じている。筆者はシステム側で、部品材料の方々にいろいろ注文を付けた経験が長い。そこから発せられる要求はその開発分野に特化したもので、全体傾向の大きな流れを表してい

ないことを痛感している。全体の大きな流れを知る方法は唯一システムの中身を知ることしかない。勉強するつもりがあれば、Web site で十分最新情報取得が可能である。その気持ちを持つことを切に希望するものである。

先行的にシステムの要求を先取りしてプロセスや材料開発ができていれば、iPhone 的な日本の強みを維持できる。しかし、ベガデータセンターの一角を取り込みたいという希望に関して、私見を提案するならば、Fig. 19 の枠取りの部分から進出することであろう。

日本はMEMSやイメージセンサなどの最先端データ取得部分に強い。このセンシング部分をスマートメモリ化して、予備演算をして、必要なデータのみフロントエンドエッジセンターに送り込む。このフロントエンドエッジセンターはそのようなプロセッシングに特化したものでよく、病院、養鶏場、セキュリティセンターなどに設置するものである。このセットアップは特化しているため、独自技術が生きる世界である。この小システムが完成すれば十分市場で競争優位を保持することが可能となる。

これを取りまとめたエンタープライズエッジセンターは一般的なデータセンターと同様であるが、Table 1 のようにメジャーな業務に対して効率化を考えたアプローチを図るのであれば、十分進出できる。このようにして少しずつ本丸に近づくことであろう。

最後に膨大なデータを高速に検索するタスクは重要である。Googleはこのシステムで飛躍した歴史を持っている。筆者も独自検索エンジンを開発した。興味があれば文献22)を参照願いたい。

(2017.3.13- 受理)

文 献

- 1) Economist September 17th 2016, Special Report Companies, <http://www.economist.com/news/special-report/21707048-small-group-giant-companiessome-old-some-neware-once-again-dominating-global>
- 2) H. Kobayashi, et al.: web site, http://www.mext.go.jp/b_menu/shingi/chousa/shinkou/020/shiryo/_icsFiles/afieldfile/2013/04/25/1333805_6.pdf
- 3) Intersect 360 Research 2015, DDN, <http://www.storagenewsletter.com/rubriques/market-reportsresearch/ddn-hpc-storage-leader-in-2015-according-to-intersect360-research/>
- 4) B. Gill: "Eight Trends Will Shape the Colocation Market in 2016," Gartner, ID: G00299639, 20 January 2016
- 5) J.-B. Lee: "Semiconductor Memory Road Map," IEEE Solid State Circuits, Vol. **8**, No. 2, pp. 66–74, Spring 2016
- 6) R. Balasubramonian: "Making the Case for Feature-Rich Memory Systems," IEEE Solid State Circuits, Vol. **8**, No. 2, pp. 57–63, Spring 2016
- 7) 木下敦寛: "ビッグデータ処理を加速するデータ中心アーキテクチャー," 日経エレクトロニクス, pp. 109–117, 2016.12
- 8) Top 100 Super Computer Ranking 2016 <https://www.top500.org/news/china-tops-supercomputer-rankings-with-new-93-petaflop-machine/>
- 9) Oracle Open World 2016.9, <https://oracle.rainfocus.com/scripts/catalog/oow16.jsp?search.event=oracleopenworld>
- 10) K. Otsuka: "How to feed to Greedy IoT Monster," EPTC2016 Keynote 2
- 11) K. Otsuka, F. Fujii, Y. Akiyama, and K. Hashimoto: "IO Interface for over 25 Gbps Operation with Low Power," ICSJ2014, 08-2
- 12) F. Fujii, D. Ogawa, K. Hashimoto, Y. Akiyama, and K. Otsuka: "Co-design Importance for over 20 Gbps I/O Interface," ICSJ2016, 17-1
- 13) K. Otsuka, F. Fujii, and Y. Akiyama: "Design and Analysis of High-Speed and Low Power for High-Performance Systems in Simulation Basis," IMPACT2014, AS-067-1
- 14) R. Kataoka, K. Kondo, J. Akimichi, Y. Akiyama, K. Hashimoto, and K. Otsuka: "300 MM WIRING ENABLING 28 GBPS TRANSMISSION ON LCP BOARD," IMPACT2014 Taipei, AS069-1
- 15) H. Shi, S. Ramalingam, and S. Dong: "Enabling Packaging Technology for Emerging 56 Gbps Lane Rate Transceivers," ECTC2015, pp. 518–522
- 16) Open CL benchmark of CPU and FPGA processing, http://www.electronicproducts.com/Digital_ICs/Standard_and_Programmable_Logic/Compiling_OpenCL_to_FPGAs.aspx
- 17) Xilinx Vertex 5 users guide, https://www.xilinx.com/support/documentation/user_guides/ug191.pdf
- 18) P. B. Minev and V. S. Kukenska: "The Virtex-5 Routing and Logic Architecture," ELECTRONICS-ET pp. 14–17, SEPTEMBER, 2009, <http://kst.tugab.bg/staff/vally/4.pdf#search=%27%2817%29+Xilinx+Vertex+5+structure+and+configuration%27>
- 19) IEEE Spectrum, Top Tech. 2017, "Deeper and Cheaper Machine Learning," pp. 28–39, Jan., 17, 2017
- 20) K. Otsuka and Y. Sato: "High Speed, Flexible, Robust and Low Power Processing Approach," IMPACT2015, AS067-1
- 21) Soitec: <https://www.soitec.com/en/products/smart-cut>
- 22) K. Otsuka, Y. Sato, K. Kobayashi, T. Kouchi, M. Uwai, and M. Nishizawa: "Novel Approach for Search Engine," IMPACT2016, 4-6

著者紹介

大塚寛治 (おおつか かんじ)
 佐藤陽一 (さとう よういち)
 橋本 薫 (はしもと かおる)
 藤井文明 (ふじい ふみあき)
 秋山 豊 (あきやま ゆたか)
 上田千寿 (うえだ ちひろ)