

リレー解説



海外の並列処理研究動向

米国 CMU における並列処理研究†

戸 辺 義 人††

1. はじめに

カーネギーメロン大学 (CMU) における並列処理研究について紹介する。かつて C.mmp¹⁾, Cm²⁾ といった並列処理システムで有名であった CMU の並列処理研究は、現在でも、iWarp³⁾ プロジェクト, Nectar⁴⁾ プロジェクトなどで今なお続けられている。

筆者は、平成2年8月から平成4年2月まで CMU の Department of Electrical and Computer Engineering (ECE) に修士課程学生として在籍するかたわら、School of Computer Science (CS) の徳田研究準教授*をリーダーとする ARTS¹²⁾ グループで分散マルチメディアに関する研究¹⁶⁾を行った。CMU CS は人工知能 (AI) の研究で有名であるが、最近では冒頭に述べた並列処理以外に分散 OS Mach^{9), 10)} を中心とした分散システム、音声認識などへの応用を中心とするニューラルネット、ロボティクスの研究にも力が注がれている。一方の ECE においてもアーキテクチャの研究が行われており、VLIW (Very Long Instruction Word) を拡張した並列処理モデル XIMD¹⁷⁾ を実証するハードウェアを試作している。

本稿では、CS の iWarp および Nectar プロジェクト, ECE の XIMD を中心に CMU の並列処理研究の現状を紹介する。

2. CMU CS の概要

CS は 1965 年に計算機科学科として設立され、研究の幅が広がり規模が拡大したのを契機に 1989 年 1 月には計算機科学部となり、1992 年 9 月には、計算機科学部内の計算機科学科となっ

た。1年間の研究予算はおよそ8億円である。博士課程の学生約150人と教職員約80人で研究活動が行われている。計算機科学部はその組織の中に、ロボット研究所 (Robotics Institute), ソフトウェア工学研究所 (Software Engineering Institute), Center for Machine Translation, Information Technology Center を抱え、CMU 最大のカレッジとなっている。

アメリカの大学にはよくあることではあるが、CMU の教職者、研究者がビジネスに転向したり、他大学へ転出するケースが目につく。

Mach の C-thread¹¹⁾, Nectar の通信プロトコル設計などで貢献があった Cooper 助教授は、ピッツバーグ市内に Fore Systems 社を設立して ATM LAN ビジネスに乗り出し、Mach プロジェクトのリーダーであった Rashid 教授は、マイクロソフト社に移った。

シストリックアレイマシン Warp の生みの親であり、Nectar プロジェクトのリーダーでもあった Kung 教授は、1992 年 1 月にハーバード大学へ移って Nectar Follow-On Project に着手している。



図-1 CMU iWarp プロジェクトのメンバ (Intel 社のメンバは含まれていない。後列左から2人目が Gross 準教授, 後列一番右が O'Hallaron 博士, 後列右から2番目が Webb 博士。)

† Researches on Parallel Processing at CMU by Yoshito TOBE (Toshiba Corp.).

†† (株)東芝重電技術研究所エレクトロニクス技術開発部

* 慶應義塾大学環境情報学部助教授を兼任されている。

Kung 教授は研究に忙しく、筆者が Kung 教授のオフィスを訪れたときも足の踏み場もないほど床が書類で埋め尽くされていたのが印象的である。

1991 年には、転出する教職者が相次ぎ、「次はだれが去るのだろうか。」と話題になった。こうした状況が関連するかもしれないが、CMU は長老教授が数多くいるという雰囲気ではなく、若い研究者たちが独立して伸び伸び研究しているというのが、全体を見渡したときの印象である。

3. シストリックアレイ LSI iWarp

iWarp^{3),4)} はインテル社と共同開発した Warp をベースとする LSI からなるシステムである。iWarp システムのアーキテクチャおよび CMU 内の研究グループの主要なメンバについて簡単に紹介する。

iWarp システムは図-2 に示す相互に通信を行う iWarp セルの集合からなる。一個の iWarp セルは Computation Agent と Communication Agent とで構成される。iWarp システムは Warp をそのまま LSI 化しただけでなく、柔軟に並列処理の形状にできるように工夫がなされている。図-3 に示すように、トーラスとすることもリングとすることも可能である。

Computation Agent は ALU 以外に浮動小数点演算器をもち、20 MFLOPS および 20 MIPS の性能を示す。Communication Agent は 4 組の入出力シリアルポートをもち、ほかの iWarp セルとの間で 40 MBytes/s で通信できる。

並列処理を実現する手段としての iWarp セルどうしの通信に、シストリック通信以外に疎結合のメッセージ通信を設けたのが iWarp システム

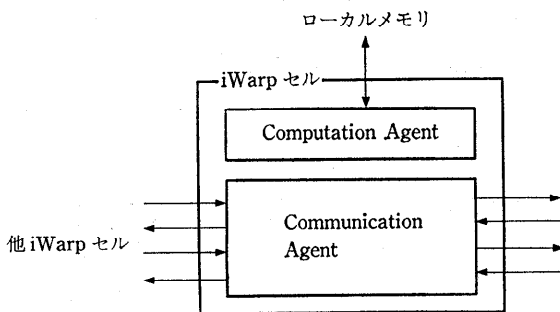
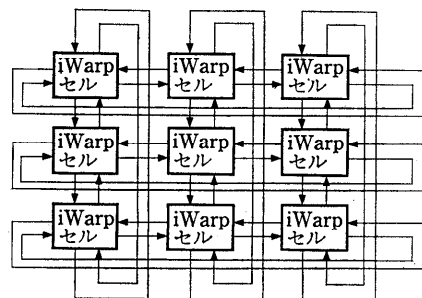


図-2 iWarp セル (文献 4) から引用

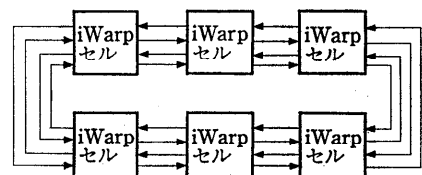
の特徴である。メッセージ通信においては、セル間通信経路は通信実行中に決定される。(コネクションレス通信に相当する。)メッセージのスプーリングは送信部と受信部とで行われ、中間経路上のルーティングはハードウェアで処理される。同期はワードごとに行う。一方の、シストリック通信では、通信元 iWarp セルから通信先 iWarp セルに至るまで、メモリを介さずに演算レベルで同期して動作させる。演算遅れは、FIFO バッファで吸収する。シストリックアレイアルゴリズムはこうしたセル間通信サポート機構により可能となる。

並列処理マシンではソフトウェアサポートが重要である。コンパイラは現在 Stanford 大学にいる Lam 博士が CMU 学生時代 Warp 用コンパイラに適用したソフトウェアパイプラインングの手法を中心に、インテル社が作成した。O'Halleron 博士はメッセージ通信方式をターゲットとした Fortran コンパイラを作成した。Webb 博士は Adapt⁵⁾ という画像処理用メタ言語から C 言語へのトランスレータを完成させて、iWarp システムでの画像処理を容易にした。そのほか、性能評価用ツールとして X ウィンドウを利用した iWarp アレイモニタリングプログラムが用意されている。

現在、Sun ワークステーションをホストとした iWarp システムが三台構築されている。コンパイ



(a) 3×3セルトーラス



(b) 6セルリング

図-3 iWarp システム (文献 4) から引用

ルはほかのどのワークステーションでも可能で、実行時にダウンロードする。現在使用されているアプリケーションとしてはシストリック演算を用いた画像処理が多い。博士学生の1人は、iWarp システムで二次元 FFT を行って医用画像の復元を行っている。ほかにも、ニューラルネット計算への適用が二グループで研究されている。

iWarp プロジェクトは、Gross 準教授をリーダーとして 20 名ほどのメンバで構成されている。すでに各種アプリケーションを動かして実証する段階にあり、これから iWarp システムを評価した報告が発表されると期待できる。

4. 高速計算機結合システム Nectar

Nectar (NEtwork CompuTer ARchitecture) は、計算機を高速ネットワークで結合して、計算機レベルでの並列処理を狙ったシステムである^{6),7)}。

Gross 準教授のほか、Steenkiste 博士も Nectar プロジェクトの中心的役割を果たしている。Nectar システムのプロトタイプはすでに完成して、CMU CS 内で 26 台のワークステーションを結合して実稼働している。現在は米国内 HPCC (High Performance Computer and Communication) プログラム¹⁹⁾で推進される 5 カ所のギガビットネットワークテストベッドの一つとして、Gigabit Nectar を開発中である⁸⁾。図-4 は Nectar プロトタイプのシステム図を示す。同図において、HUB は 16×16、100 Mb/s の光結合クロスバースイッチ、I/O ポート、制御部で構成される。クロスバースイッチを用いるところに、大学としての C. mmp の経験が感じとれる。制御部は回線交換とパケット交換を制御するコマンドを実行する。CAB は通信遅延時間を短縮するためのプロトコルエンジンで、VME バスを介してワークステーションに直結される。CAB を用いると、CAB-CAB 間で 8 KByte のパケットに対して、TCP (Transmission Control Protocol) で 85 Mb/s のスループットを達成することができる。ただし、ホストどうしの間では VME バスの転送速度の制約で 27 Mb/s が限界となっている。応用としては、スイッチ回路シミュ

レータ COSMOS、ロサンゼルス地区の大気汚染シミュレーションなどが実行されている。なかなか興味深い応用ではあったが、ネットワークの性能が大きく実行時間を左右することも分かったという。

現在、Network Systems 社と共同で Gigabit Nectar として 100 MBytes/s のスループット目標のシステムを構築中である。スイッチには P8, P32 HIPPI (High Performance Parallel Interface) スイッチを用いる。iWarp システムおよび DEC 5000 ワークステーション用のネットワークインタフェースにまず取り組んでいる。近年、XTP (eXpress Transfer Protocol)²⁰⁾など、プロトコル全体をハードウェア処理するという方法も提案されている。しかし、Nectar プロジェクトでは、プロトコル処理全部をハード化するとフレキシビリティに欠けるとともに高価なものになると考え、送受

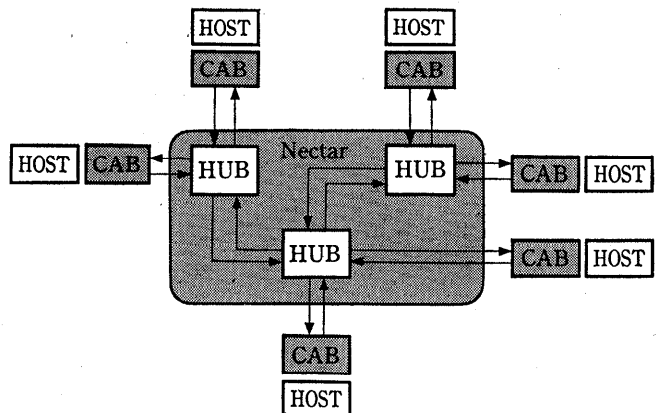


図-4 Nectar プロトタイプ (文献8) から引用

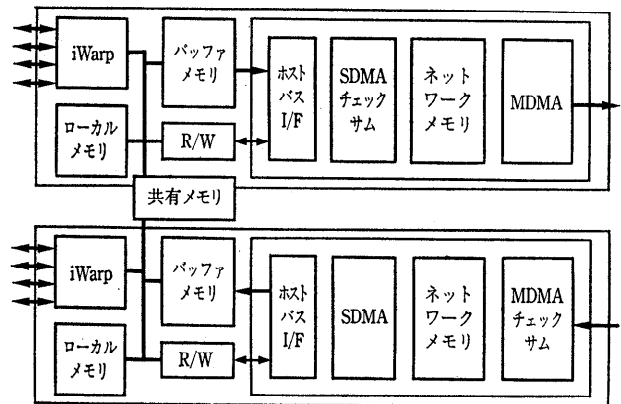


図-5 Gigabit Nectar iWarp ネットワークインタフェース (文献8) から引用

信データのバッファリング、DMA 転送、チェックサム計算だけをハードウェアで処理するようにしている。プロトコル処理そのものよりもデータコピー、チェックサム計算が通信ボトルネックになっているという最近の考え方の趨勢に沿っている。

図-5 は iWarp システムを用いたネットワークインタフェースである。このボードを HIB (HIPPI Interface Board) と呼ぶ。送信用、受信用おのおのに iWarp セルを割り当てる。数多くの iWarp セルの中で特に二個の iWarp セルだけがネットワークとの送受信を受け持つので、その負担が気になる。

HIB はすでに完成し、HIPPI 接続も終了し、HIB-HIB 間で 720 Mb/s のスループットを達成している。Nectar プロジェクトには総勢約 10 名で取り組んでいるが、iWarp プロジェクトを兼任するメンバが何人かいる。

5. X I M D

ECE では、CDS (Center for Dependable Systems) でアーキテクチャの研究が行われている。Shen 教授らは、XIMD¹⁷⁾ という VLIW を拡張した、粒度の低いインストラクションレベルの並列化方式の研究を行っている。

図-6 は並列処理モデル XIMD を Moore マシン²¹⁾ で表現したものである。 $\delta_1, \dots, \delta_n$ は組合せ回路、 S_1, \dots, S_n は次の状態、 $\lambda_1, \dots, \lambda_n$ はデータベースを制御する出力でマイクロ命令メモリからなる。図-7 の MIMD と比較すると、XIMD は各マイクロ命令ストリームの決定がほかのマイクロ命令ストリームの状態にも依存する場合を含むのを特徴とするのが分かる。実装上はシーケンサが複数個存在するということになる。 δ および S が一個しかないときには図-8 の VLIW モデルになり、XIMD は SISD, SIMD, MIMD, VLIW すべてを包含する上位モデルとなる。

Shen 教授のグループは、XIMD より先に White Dwarf¹⁸⁾ という VLIW マシンを構築している。この VLIW マシンを稼働してみても、条件付き分岐が多いコードを効率よく実行させるのが難しかったことが、XIMD を考えるきっかけとなったようだ。XIMD はコードスケジューリングなど、さまざまな研究上、興味がひかれる点があるようである。

現在のところ、XIMD システムの論理設計シミュレーションは完了して、実システムを構築している最中であり、Shen 教授のほか、大学院学生 2 名、研究員 1 名で研究が行われている。

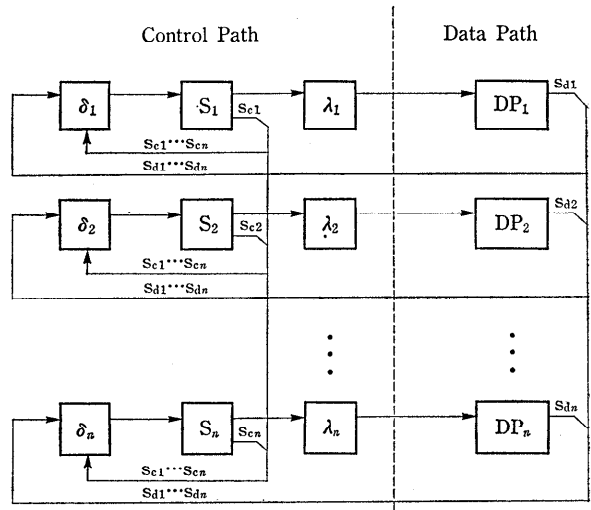


図-6 XIMD モデル (文献 17) から引用)

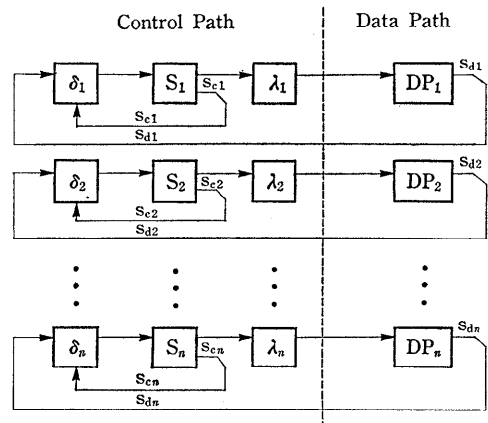


図-7 MIMD モデル (文献 17) から引用)

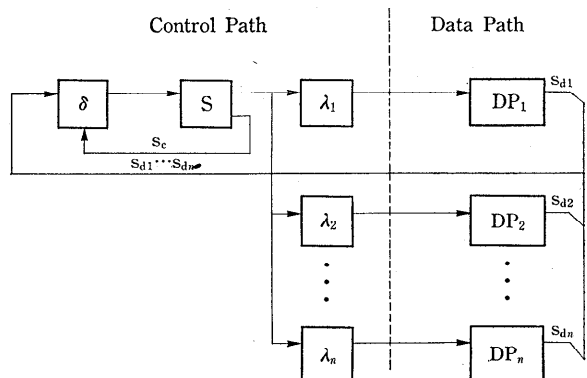


図-8 VLIW モデル (文献 17) から引用)

6. その他のプロジェクト

CMU では、分散 OS の研究も長く行われてきた。ARTS¹²⁾、Accent¹⁵⁾ はその例で、Mach は Accent を発展させたものである。Rashid 教授は、1990 年 9 月に行われた CMU CS 25 周年記念講演の中で、「Mach は突然 (out of the blue), 生まれたわけではない。分散 OS 研究の長い蓄積があったからこそできた。」と述べた。Mach の基本的な原理は文献 9) など多くの書物に記されているので、本稿では割愛させていただく。

Mach プロジェクトの現在のリーダーは Barshad 助教授であり、総勢 15 名ほどで研究が続けられている。Mach のバージョンは MK で数えられるが、現在 MK-73 まで進んでいる。RT-Mach¹³⁾ は Mach にリアルタイム機能を統合したもので、徳田研究準教授のグループによって開発が続けられている。RT-Mach グループでは現在、RT-Mach 上に OS サーバ・RTS (Real Time Server) とネットワークプロトコルサーバ・NPS (Network Protocol Server)¹⁴⁾ を構築している。CMU 以外での Mach の動きとしては、OSF (Open Software Foundation) でセキュリティ機能を強化した Secure Mach の開発が行われている。

CMU CS 内で並列処理に関連するほかの研究を文献 22) から抜き出してみると、Fahlman 博士によるニューラルネットの高速学習、Miller 博士による主として有限要素法に応用する並列アルゴリズム、Blelloch 助教授による Connection Machine 向けの並列アルゴリズム、Clarke 教授による並列プログラミング言語などがあげられる。iWarp, Nectar プロジェクト以外では、主としてソフトウェアでの並列処理が多く研究されている。

7. おわりに

以上、CMU の並列処理研究について紹介した。並列の粒度が、インストラクションレベルの XIMD, プロセッサレベルの iWarp, 計算機レベルの Nectar とさまざまである。企業と共同開発しているプロジェクトが多く、短期間で実システムを実現して活発に活動している。

研究支援環境という点では、LAN, 電子メール・電子掲示板, 文書作成ツールが日本でも普及してきた現在では、さほどアメリカの大学が優れ

ていると感じないかもしれない。しかし、研究を援助する優秀なプログラマや計算機のトラブルをすぐさま解消してくれるスタッフがいるような「研究に専念できる」環境というのが、筆者には大きな魅力であった。

筆者は、日米両方で大学院学生を経験したので、日米の大学院教育の差を感じることも多かった。修士課程における宿題の量もさることながら、博士課程における中間審査が厳格であった点が最も印象に残っている。指導教授は学生に授業料と毎月の生活費を支給しているので、両者は言わば雇用関係にあり、学生は研究成果いかんによっては大学から追い出されることになる。筆者のオフィスメイトであった MIT 修士出身の博士学生も、入学して 1 年目の口頭試験で不合格となって退学になり、寂しい思いをした。こうした厳しい環境で、大学院学生はみな真剣に研究に取り組んでいる。日米大学教育の差異という観点からみると、学部教育ではなく大学院教育でこそ、日本はアメリカに大きく差をつけられていると感じた。

最後に、本稿を書くにあたって、ご協力いただいた、徳田研究準教授、Shen 教授、Steenkiste 博士、Bershad 助教授、Assistant Dean の Ms. Copetas、写真の掲載を許可して下さった Gross 準教授に感謝申しあげる。

参考文献

- 1) Fuller, S. H., Swan, R. and Wulf, W. A.: The Instrumentation of C. mmp: A Multi-Miniprocessor, IEEE Comcon (1973).
- 2) Swan, R. J. et al.: The Implementation of the CM* Multiprocessors, Proc. of NCC, 46, pp. 645-655 (1977).
- 3) Borker, S., Cohn, R., Cox, G., Gross, T., Kung, H. T., Lam, M., Levin, M., Moore, B., Moore, W., Peterson, C., Susman, J., Sutton, J., Urbanski, J. and Webb, J.: Supporting Systolic and Memory Communication in iWarp, IEEE Int. Symp. on Computer Architecture, pp. 70-81. (1990).
- 4) Borker, S., Cohn, R., Cox, G., Gleason, S., Gross, T., Kung, H. T., Lam, M., Moore, B., Peterson, C., Pieper, J., Rankin, L., Tseng, P. S., Sutton, J., Urbanski, J. and Webb, J.: iWarp: An Integrated Solution to High-Speed Parallel Computing, Proc. of Supercomputing '88, pp. 330-339 (1988).
- 5) Webb, J. A.: Steps Toward Architecture-Independent Image Processing, IEEE Computer, pp. 21-31 (1992. 6).

- 6) Arnould, E., Bitz, F., Cooper, E., Kung, H. T., Sansom, R. and Steenkiste, P.: The Design of Nectar: A Network Backplane for Heterogeneous Multicomputers, Proc. of 3rd Intl. Conf. on Architectural Support for Programming Languages and Operating Systems, ACM/IEEE, Boston, pp. 205-216 (1989. 4).
- 7) Brugge, B., Nishikawa, H. and Steenkiste, P.: Computing over Networks: An Illustrated Example, Proc. of the Sixth Distributed Memory Computing Conference, IEEE, pp. 254-257 (1991. 4).
- 8) Steenkiste, P.: The Gigabit Nectar Multicomputer Project, Proc. of Annual Conf. Network Systems User Group, pp. 43-49 (1991. 9).
- 9) Golub, D., Dean, R., Forin, A. and Rashid, R.: Unix as an Application Program, Proc. of Summer USENIX Conf. (1990. 6).
- 10) Tevanian, Jr. A. and Rashid, R. F.: MACH: A Basis for Future UNIX Development, Technical Report, Computer Science Department, Carnegie Mellon University (1987. 7).
- 11) Cooper, E. C. and Draves, R. P.: C Threads, Technical Report, Computer Science Department, Carnegie Mellon University, CMU-CS-88-154 (1987).
- 12) Tokuda, H. and Mercer, C. W.: ARTS: A Distributed Real-Time Kernel, ACM Operating Systems Review, 23 (3) (1990).
- 13) Tokuda, H., Nakajima, T. and Rao, P.: Real-Time Mach: Toward a Predictable Real-Time System, Proc. of USENIX Mach Workshop (1990. 10).
- 14) Tokuda, H., Nakajima, T. and Oikawa, S.: Towards a New Operating System Architecture: Micro Kernel vs. Reflective Architecture, 日本ソフトウェア科学会第9回大会, pp. 345-348 (1992).
- 15) Fitzgerald, R. and Rashid, R. F.: The Integration of Virtual Memory Management and Inter-process Communication in Accent, ACM Trans. on Computer Systems 4(2) (1986. 5).
- 16) Tokuda, H., Tobe, Y., Chou, S. T.-C. and Moura, J. M. F.: Continuous Media Communications with Dynamic QOS Control Using ARTS with an FDDI Network, Proc. of ACM SIGCOMM '92, pp. 88-98 (1992).
- 17) Wolfe, A. L.: XIMD: A Variable Instruction Stream Computer Architecture, Research Report No. CMUCAD-92-04 (1992).
- 18) Wolfe, A. L., Breternitz, Jr. M., Stephens, C., Ting, A. L., Kirk, D. B., Bianchini, Jr. R. P. and Shen, J. P.: The White Dwarf: A High-Performance Application-Specific Processor, Proc. 15th Annual Intl. Symp. on Computer Architecture, IEEE, pp. 212-222 (1988).
- 19) Grand Challenges: High Performance Computing and Communications, The FY 1992 U. S. Research and Development Program.
- 20) Protocol Engines, Inc.: XTP Protocol Definition, Revision 3.5 (1990. 9).
- 21) Hartmanis, J. and Stearns, R. E.: Algebraic Structure Theory of Sequential Machines, PrenticeHall, Englewood Cliffs, NJ (1966).
- 22) Faculty Research Guide 1991-92, School of Computer Science, Carnegie Mellon University (1991).

(平成4年12月15日受付)



戸辺 義人 (正会員)

昭和37年生。昭和59年東京大学工学部電気工学科卒業。昭和61年同大学大学院工学系研究科修士課程修了。同年(株)東芝入社。重電技術研究所勤務。プロセス制御 LAN, 情報通信ネットワークの研究に従事。平成4年カーネギーメロン大学修士課程修了。電子情報通信学会, IEEE, ACM 各会員。