

統計数理研究所のスパコンと乱数発生

田村 義保 統計思考院 特任教授

今も昔も統計科学と計算機は切っても切れない関係にある。統計数理研究所においては、常に最先端の計算機を導入し統計科学における計算機利用をリードしてきた。田村は1989年稼働開始のHITAC M682Hから2014年7月稼働のSGI ICE-Xまで、研究所の計算機導入に関係してきた。75周年を記念して統数研の計算機環境を振り返って見ることにした。

年譜にあるように、1956年にFACOM128（リレー計算機）が導入されている。これは「計算機屋かく戦えり」（遠藤諭著、株式会社アスキー、1996年）にあるように、日本の商用計算機の1号機であった。残念ながら、部品等は統数研には残っていないが、富士通の沼津工場には1959年製のFACOM128Bが展示されており、今でも計算可能である。動作可能な状態で計算機を保存・展示されていることに敬意を表したい。統数研では1964年にHIPAC103（名称TSKIII：パラメロン計算機、ALS-1000とのハイブリッド）を導入している。この時の計算素子は、地下の分散コンピュータ博物館（計算機展示室）に展示している。この時に、初代のダイオードの熱雑音を用いる物理乱数発生装置も導入されている。

FACOM128の時は、放射線の計数を利用した発生装置が装着されていたようなので、これから数えると第2代である。

これらの計算機が導入される前後の統計数理研究所彙報には次のような報告がある。記録として、このような報告・解説を残しておくことは重要であると考え。個人的（正しくは概算要求のため）に計算機の歴史をまとめているが、公表はしていない。75周年を迎えるにあたり、何らかの形で記録として公表したいと考えている。このポスターはその一環である。計算機が統計数理の研究に必須のものであることは現在では万人が認めることであろう。統計数理研究所は計算機の利用については常に先進的であったことが次のような記述（原文のまま）からうかがうことができる。

「これらの問題に対して我々は、基礎的、理論的な方面並びに応用的方面より互いに相補いつつ進んで行かなければならないと思う。この為には又、各種計算器をも含めた計算技術の研究も必要となるであらう。」（統計数理研究所彙報 第2巻 第1号 1954 『十周年にあたりて』 松下嘉米男）

「最近、統計数理研究所において新たに乱数作成機がつくられた。これは、放射能のカウント数を計測する方法によっているものである。」（統計数理研究所彙報 第5巻 第2号 1958 石田式乱数作成機のランダム性について 西平美恵子）

「27年9月に佐々木達治郎博士が所長に就任されて初めて、研究所が次第に軌道に乗るようになった。佐々木博士が数値計算の権威であって、リレー式であるが高性能の計算機を試作されたことは、まことに記念すべき業績である。」（統計数理研究所彙報 第7巻 第1号 1959 十五周年記念日を迎えて 末綱恕一）

統計数理、統計科学の研究のためには、最先端の計算機が必須であることは、60年以上前も今も、変わらないということが言える。ビッグデータ、データ中心科学、データサイエンスの重要性が叫ばれる昨今、統計科学の研究における計算機の重要性は増している。田村が最後に執筆した仕様書の一部を示しておく。

昨今、ビッグデータやデータ中心科学の重要性が指摘されている。研究所がこの潮流の中心的な研究機関として研究を推進していく必要がある。このような計算を行うためには、共有記憶型と分散記憶型の計算機が共に必要である。大規模主記憶の共有が可能であることにより初めて計算可能な課題も多い。しかしながら、計算量が莫大となるために、個々のCPUに付属する主記憶が比較的少なくても計算は可能であるが、高い並列度が必要な課題も存在するためには分散記憶型並列計算機が適していると考え。

この仕様書で導入したシステムが統計科学スーパーコンピュータシステム（SGI ICE-Xを中心としたシステム）である。同時に世界最大の共有記憶領域（128TB）を有するデータ同化スーパーコンピュータシステム（SGI UV2000を中心としたシステム）も平成24年度補正予算で導入している。

計算機の年譜を示しておく。

1954	FACOM415A（相関係数専用機）
1956	FACOM128（リレー計算機） 物理乱数発生機（放射能）
1963	HIPAC103（名称TSKIII パラメロン計算機、ALS-1000とのハイブリッド） 物理乱数発生機（ダイオード熱雑音、12,000ビット/秒）
1971	HITAC8500改良型（主記憶524KB）／HITAC8300（主記憶64KB） 物理乱数発生機（ダイオード熱雑音、200KB/秒）
1972	HITAC8700（主記憶524KB）／HITAC8400（主記憶131KB）
1978	HITAC M180（主記憶2MB）
1981	HITAC M200H（主記憶10MB）
1985	HITAC M280H（主記憶24MB）
1989	HITAC M682H（主記憶256MB 内蔵アレプロセッサ） HITAC M660D（主記憶64MB） 物理乱数発生機（ダイオード熱雑音、1.5MB/秒）
1993	HITAC M680H（主記憶256MB） HITAC M680H（主記憶256MB 内蔵ベクトル演算機構）
1994	HITAC M880/180（主記憶512MB） HITAC S3600/120（主記憶256MB 拡張記憶2GB）
1996	HITAC S3800/162（主記憶1GB 拡張記憶4GB 4GFlops）
1996	IBM SP2（48ノード、主記憶12GB）
1999	HITAC SR8000（20ノード 主記憶8GB/ノード 8GFlops/ノード） 物理乱数発生装置（日立製：ダイオード熱雑音、25MB/秒、6枚） HP9000V2250（16CPU 主記憶12GB PA-RISC8500(240MHz)）

2000	Origin2000(64CPU 主記憶48GB R12000(300MHz)) (400MHzにクロックアップ 2000年8月) (2000年1月) 統計計算開発環境支援システムを16CPUにして計80CPUで運用
2004	SGI Altix3700（主記憶1920GB ピーク5.2GFlops×256） 物理乱数発生装置（東芝製：ダイオード熱雑音、133MB/秒、6枚） NEC SX6(主記憶128GB ピーク9GFlops×12) Hitachi SR11000（主記憶128GB ピーク6.8GFlops×64）(2004年1月)
2006	HP XC4000（計算ノード ProLiant DL 145G2 128ノード (Opteron 2.6GHz × 2 主記憶640GB)
2010	Fujitsu SPARC Enterprise 64CPU(256コア) 主記憶2TB ピーク 2949GFlops と 24CPU(96コア) 主記憶1TB ピーク 1105GFlops) Fujitsu PRIMERGY RX200S5(360ノード (2880コア) (2010年1月) 主記憶 48GB×160+24GB×200 ピーク 93.76GFlops×360)
2014	統計科学スーパーコンピュータシステム (SGI ICE-Xを中心としたシステム) (Intel E5-2697v2 2.7GHz/12コア400ノード((2CPU+128GB主記憶) 207TFLOPS 50.0TB(合計) (2014年7月)
2015	統計科学スーパーコンピュータシステム増設 Intel Haswell(コード名) 2.4GHz/14コア120ノード((2CPU+256GB主記憶)/ノード) 129.0TFLOPS 30.0TB (初期との合算336.3TFLOPS)
2018	統計科学スーパーコンピュータシステム(HPE SGI 8600・ 計算ノード(376ノード) CPU: Intel Xeon Gold 6154 x 2 (18core 3.0GHz) Mem: 384GB/Node GPU計算ノード(8ノード) CPU: Intel Xeon Gold 6154 x 2 / node (18core 3.0GHz) GPU: NVIDIA P100 x 4 /node Mem: 384GB/Node システム合計理論性能: 1.49PFlops 総Core数: 13824 core 総主記憶容量: 144TB

若い人は30年前の計算機の主記憶64MBというのを見てなんかの間違いないと思われるかもしれない。まぎれもない事実である。計算機は5年ごとにほぼ同価格で10倍の性能、必要電力ほぼ同じものが調達可能である。100周年の時は今の10万倍の性能の計算機が、あるかもしれない？



パラメロン素子



SGI8600と田村 2018年9月



乱数発生装置
<https://www.ism.ac.jp/ism-tour/>