

探索アルゴリズムと並列計算 Search algorithms and Parallel computing

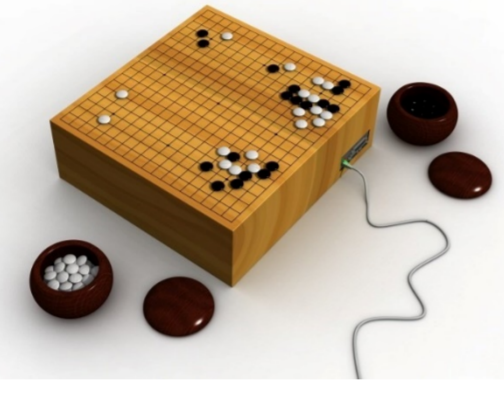
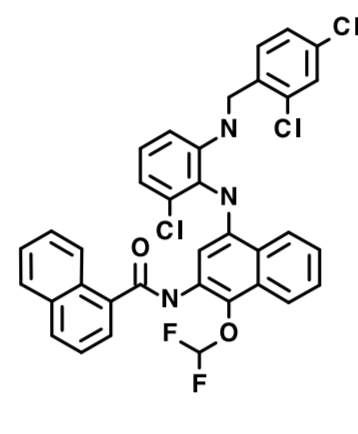
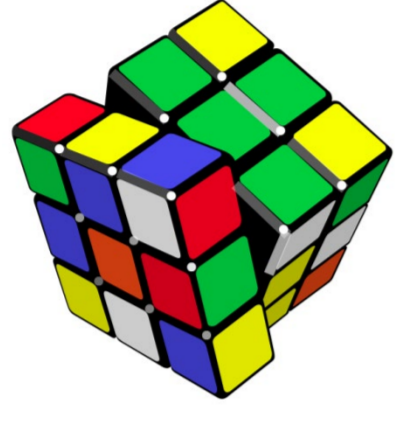
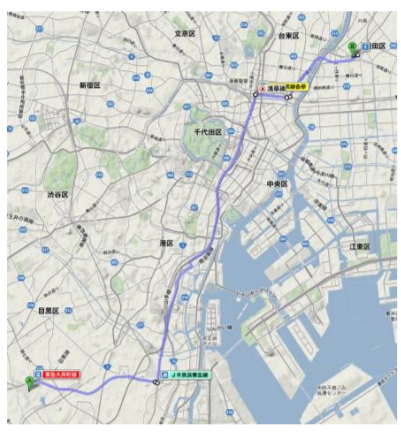
ここで言う探索とは

グラフから特定の節点又は経路
を発見すること

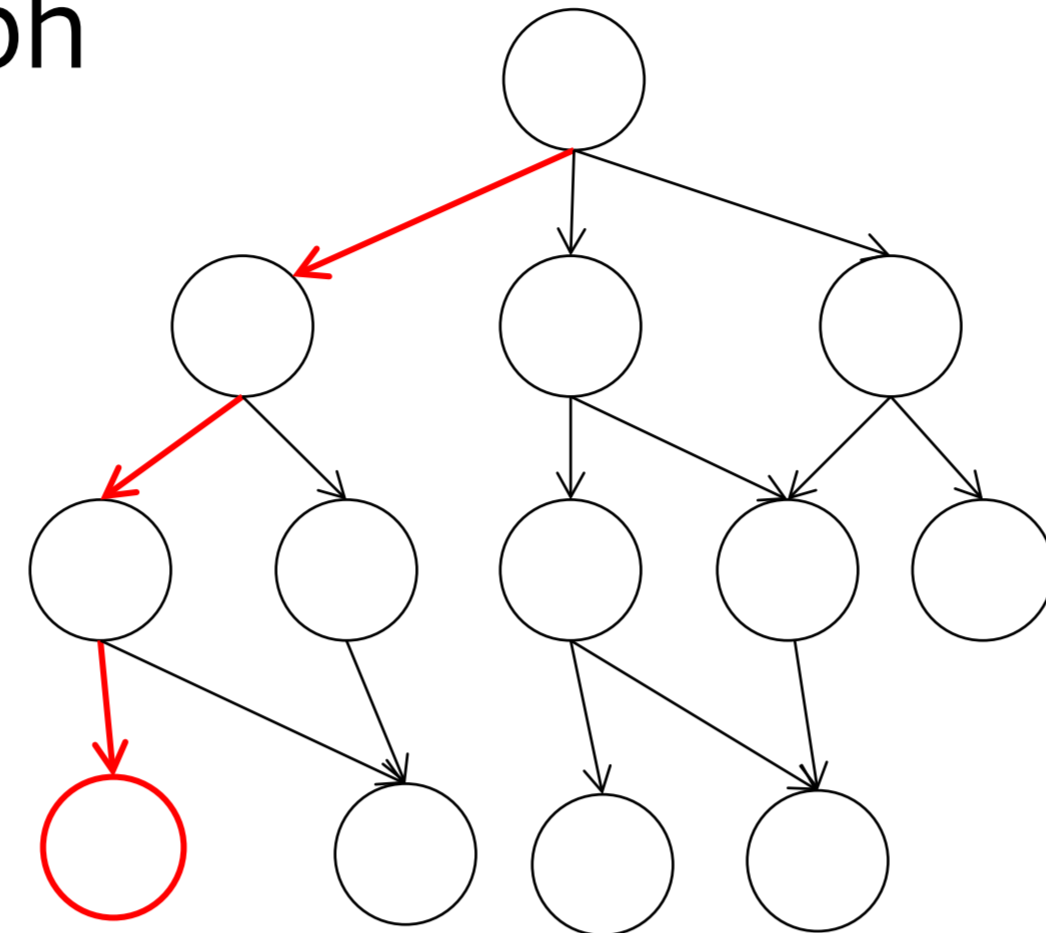
here "search" means

finding specific **node** or **path**
from given graph

「最短経路」
「最善の組合せ」
「最善手」などに対応



which corresponds to
"shortest path"
"optimal combination"
"best move (play)"



材料化学への応用 1/2

Application to Material Science 1/2

昨年度から継続して

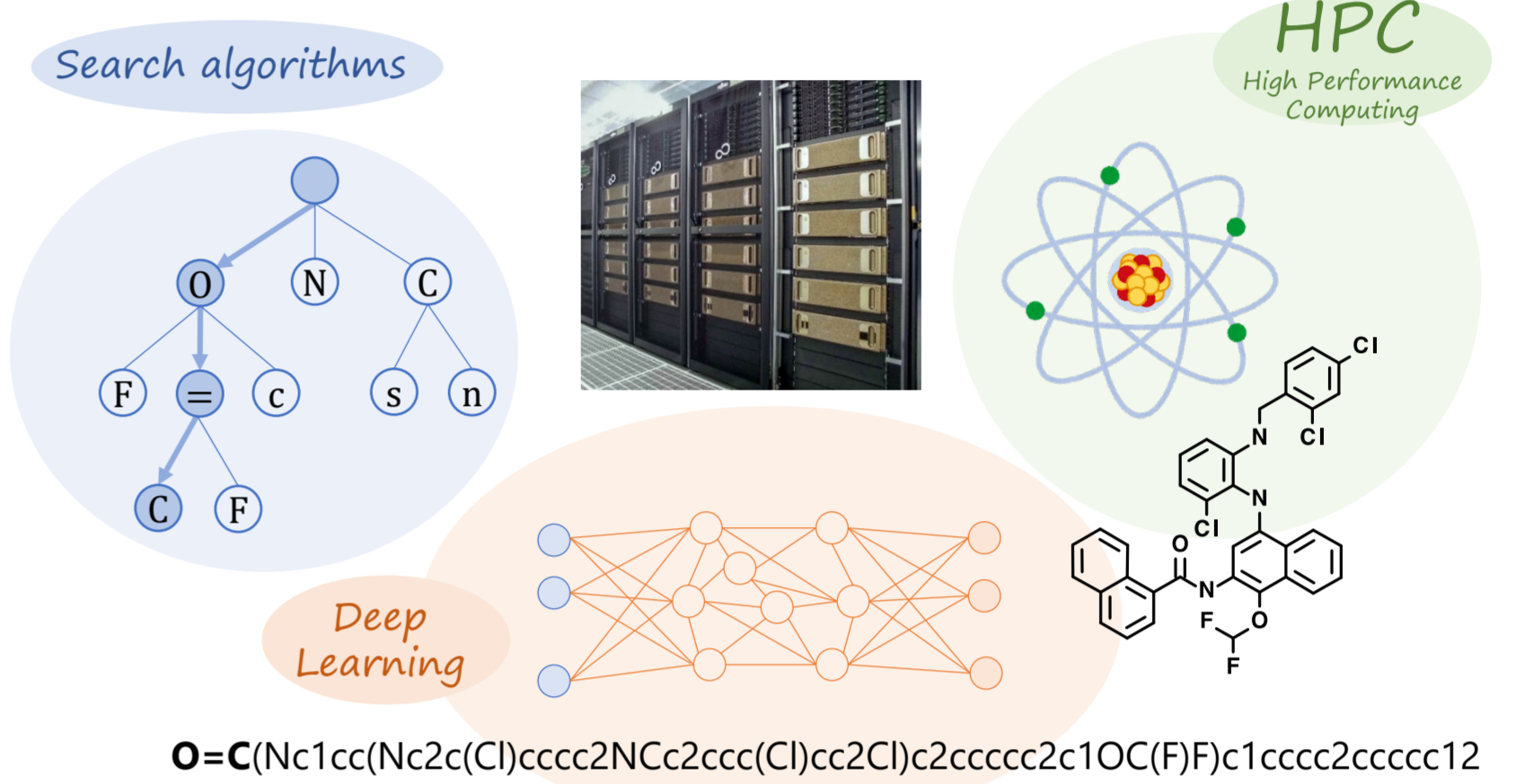
深層学習 + モンテカルロ木探索 + 計算化学シミュレーションによる新規化合物の探索

今年度は並列化により大規模に

Continuing work from last fiscal year

Finding new molecules using DL + MCTS + computational chemistry simulator.

This year, applying large scale parallel search



多重検定への応用

Application to Multiple testing contribution to biology/medical science

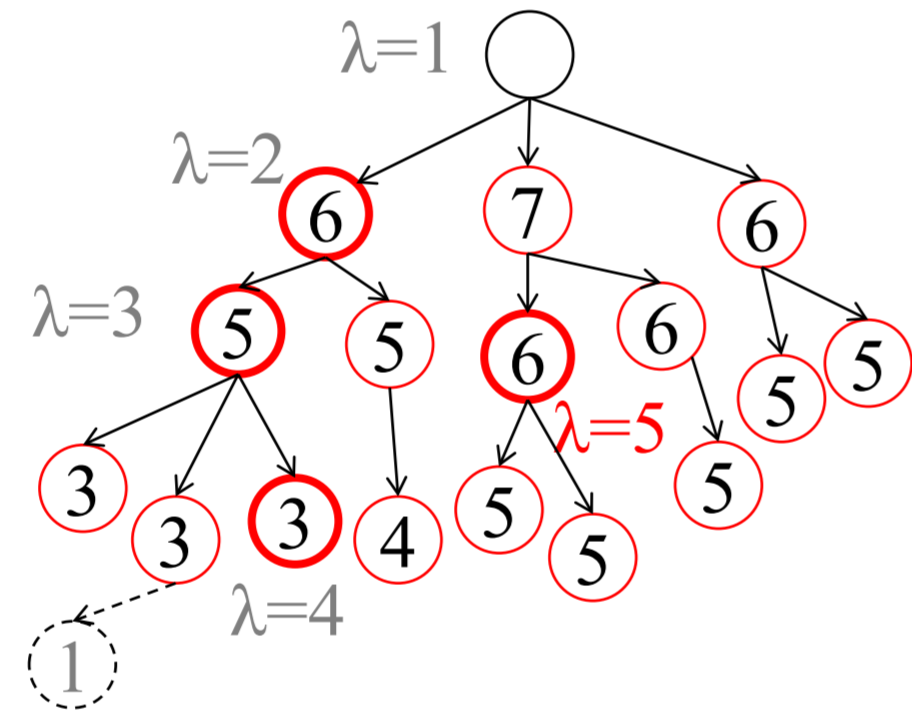
大規模並列統計的パターンマイニング

適用先を広げ、さらに京コンピュータ等でより大規模な並列化

Massive Parallel statistical pattern mining

new applications and larger scale parallelization on supercomputers including K-computer

昨年度まで主に遺伝病の解析などに適用 (SNP解析, GWAS) 今年度は診断データ等に適用対象を拡大
Previously this method was applied to genomic disease analysis (SNP analysis, GWAS). This year applied to other data including diagnostic data.

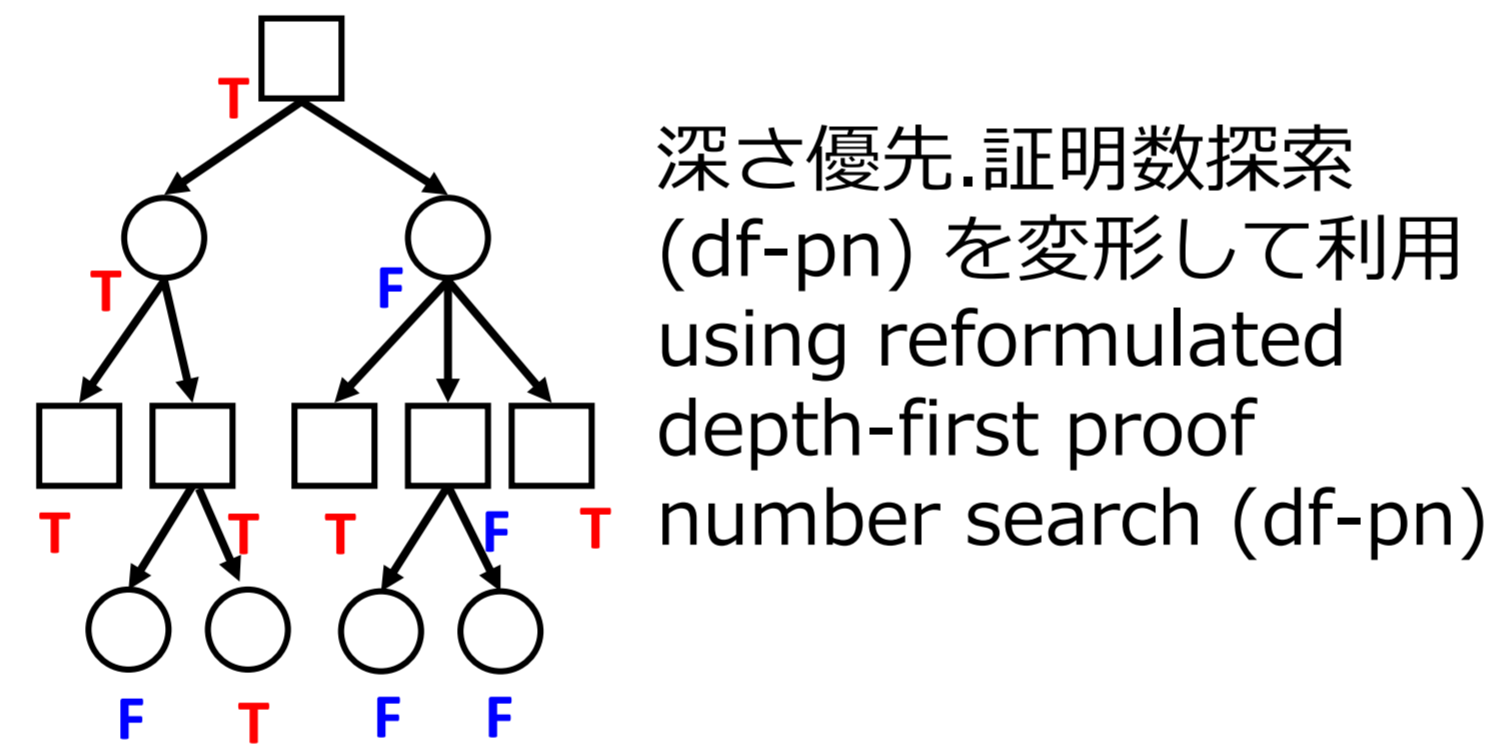
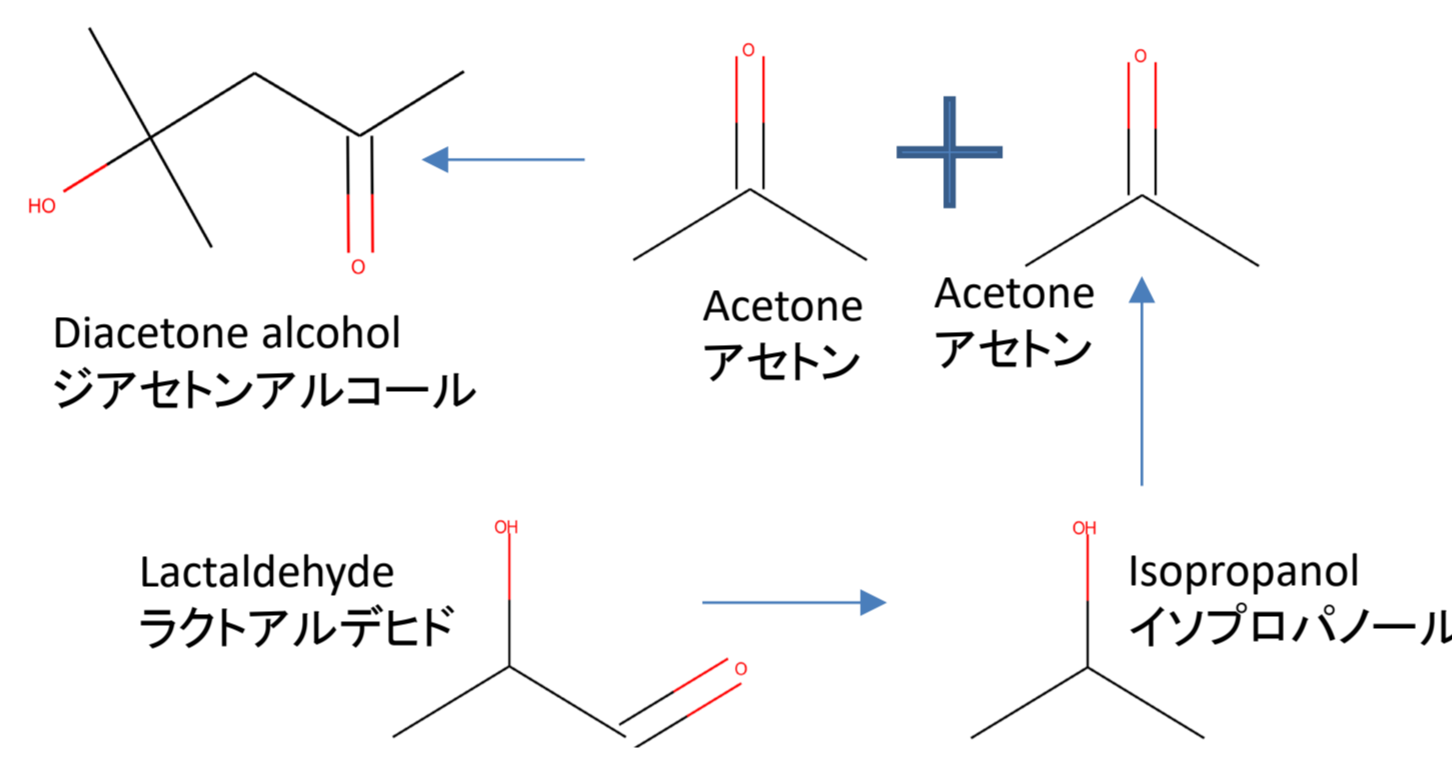


AND-OR木探索を用いた

化合物の逆合成解析

AND-OR tree based

Retrosynthesis algorithm



深さ優先. 証明数探索 (df-pn) を変形して利用 using reformulated depth-first proof number search (df-pn)

その他の取り組み

- 大規模並列グラフアルゴリズム
- 交通流予測、災害時の交通対策

Other ongoing research

- Large scale parallel graph algorithms
- application to traffic prediction, disaster time traffic manipulation

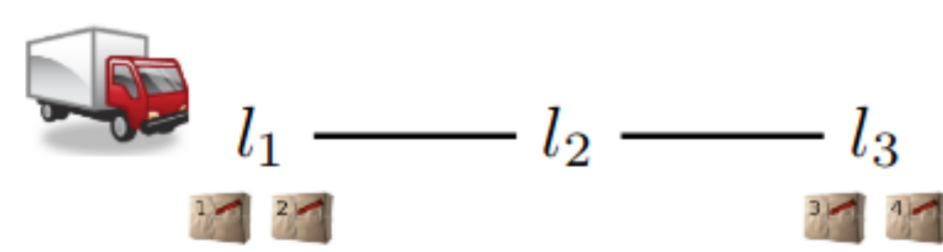
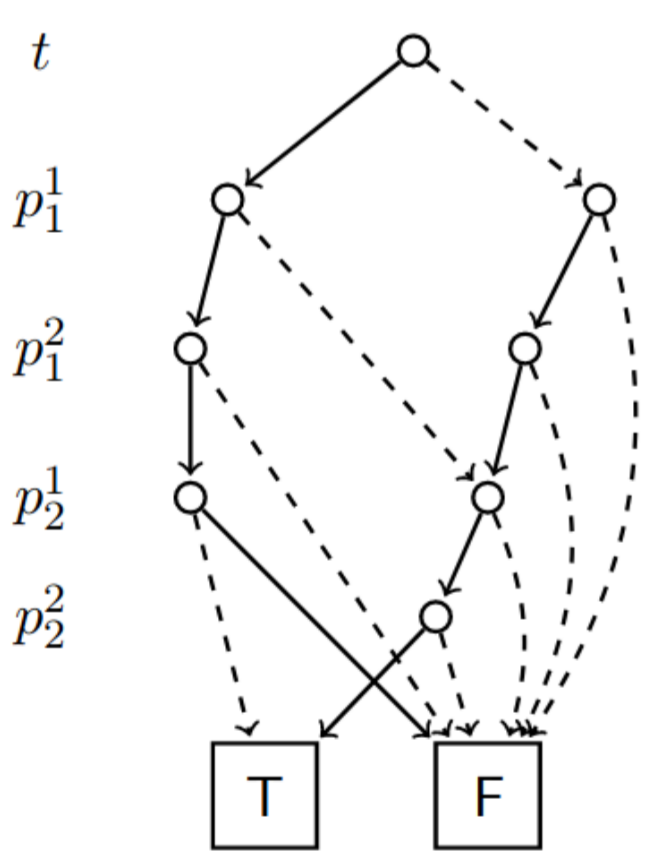
主な研究目的

- 探索アルゴリズム自体の研究
- 機械学習 (深層学習) 等を用いて応用範囲の拡大
- 主に並列化による高速化

Main research goals

- Improve search algorithms
- Apply to new domains using ML/DL, enumeration
- High speed algorithm mainly based on parallelization

combining search and BDD (binary decision diagram) と探索の組合せ



AI研究用計算機システムの整備

RAIDENの企画・運用

理研AIPのスーパーコンピュータ (RAIDEN)のソフト/ハードウェアの選定・調達・保守・ユーザ対応などを担当

Dockerコンテナへの対応

従来のスーパーコンピュータシステムで用いられてきたバッチジョブシステムを踏襲しつつ、研究のスピードを向上させるコンテナ技術にも対応 (dockerコンテナ)



Tesla V100 GPU

多彩な研究ニーズに応えるシステム構成

- 昨今のAI研究でのスタンダードとなっている GPU (NVIDIA Volta GPU) 搭載の計算機
- 大規模並列演算を可能にするPCクラスタ
- 大容量かつ高速化を実現した分散ファイルシステム

高性能・省エネルギー

2017年3月

2018年3月

NVIDIA DGX-1 Tesla P100 24 node

NVIDIA DGX-1 Tesla V100 54 node

理論性能 (FP16精度) 4PFLOPS

理論性能 (FP16精度) 54PFLOPS

Green500 (June 2017) 4位
10.602 GFLOPS/W (FP64精度)

Green500 (June 2018) 10位
11.363 GFLOPS/W (FP64精度)

さらに NVIDIA DGX-2 (Tesla V100 32GB x 16) を 6 node, および高速ストレージを増設予定 (2019年3月中)