であること。非圧縮環境において15TB以上、圧縮時（理論値）の保管容量が60TB以上となるデータを保存できること。一般的なデータベース記述言語による処理が可能であり、かつ、ＯＤＢＣ接続を通じてクライアント側の負荷を最小化にしつつ処理が可能であること。100億件の時系列データを高速、少なくとも数時間以内に入出力が可能であり、全体の集計処理を数分で実施できること。入出力、集計時には特定のインデックスなど事前のデータ構造などに依存しないで高速な集計が可能であること。入出力、集計、結果のグラフ表示までの全てが一貫したＧＵＩ（Graphical User Interface）操作によって完結できること。

② 大規模データ保管用ストレージシステム　ストレージ容量は１PB以上であること。④のネットワークインフラ上に設置された③の計算システムから20Gbyte／s以上（理論値）でデータの入出力が可能であること。Windowsシステムから直接または間接的にデータの入出力が行えること。ＩＯＰＳを最大化するシステム構成が望ましい。

③ 高性能計算機システム　上記①、及び②に蓄積されたデータを深層学習モデルによって分析するための高速計算機システムを構成すること。高性能計算機システムの合算理論ピーク性能は4.5ＰＦＬＯＰＳ以上であることが望ましい。浮動小数数型はＩＥＥＥ754標準に準拠するものとするが、精度は問わない。また、倍精度での合算理論ピーク性能は250ＴＦＬＯＰＳ以上であることが望ましい。搭載メモリの合計容量は6.5TB以上とする。また計算機システムには５台以上の簡易なＡＩ計算が可能となる予備プラットフォームを準備すること。高速計算機システムのＯＳはMicrosoft Windowsを採用することが望ましい。

④ ネットワークインフラ　上記①から③の全てを相互に高い転送バンド幅で接続すること。①と②、③は40GbE以上によって接続すること。②と③は転送バンド幅200Gbps以上で接続すること。①から③の全てを100GbEによって接続し、ＧＡＭＡの外部に設置されたサービスネットワークと接続すること。

⑤ 管理機能　上記①から④の全体をＧＡＭＡ設置サーバルームから100ｍ以下の距離にある２箇所の管理ルームから直接接続し、操作可能な管理システムを有すること。管理ネットワークは④のネットワークのバンド幅に影響を与えないよう独立して敷設すること。

⑥ その他　システムを運用するために十分な給電設備及び冷却設備を構築すること。ただし、最大消費電力は80kW以下とすること。システムを最大限活用し、大規模時系列データ及び画像等多様な大規模データの分析、長期シナリオ分析を実現するソフトウェア群を有すること。また、これらを解析するためのソフトウェア群を有すること。

２　資料及びコメントの提供方法　上記１⑵の物品に関する一般的な参考資料及び同⑹の要求要件等に関するコメント並びに提供可能なライブラリーに関する資料等の提供を招請する。

⑴ 資料等の提出期限　令和２年４月24日17時（郵送の場合は必着のこと。）

⑵ 提出先　〒305―8569 茨城県つくば市小野川16―１　国立研究開発法人産業技術総合研究所　つくば西事業所研究業務推進部会計グループ　大熊　章博　電話029―861―8133

３　説明書の交付　本公表に基づき応募する供給者に対して導入説明書を交付する。

⑴ 交付期間　令和２年３月16日から令和２年４月24日まで。

⑵ 交付場所　上記２⑵に同じ。

４　説明会の開催　本公表に基づく導入説明会を導入説明書交付後別途開催する。

⑴ 開催日時　令和２年３月23日14時

⑵ 開催場所　国立研究開発法人産業技術総合研究所　つくば西事業所　つくば西―１棟1204室

５　その他

⑴ この導入計画の詳細は導入説明書による。なお、本公表内容は予定であり、変更することがあり得る。

6　Summary

⑴ Classification of the products to be procured : 14

⑵ Nature and quantity of the products to be purchased : Collective Intelligence Analysis Platform 1 Set

⑶ Type of the procurement : purchase

⑷ Basic requirements of the procurement : GAMA is composed of the following devices : a data warehouse for time-series data analysis ; a storage system for large-scale data ; a high-performance computer system for AI analysis ; a network infrastructure that these aforementioned devices stably at high speed ; software that makes effective use of these aforementioned devices. All devices and systems are mounted in racks. The requirements of the system are shown below.

A. Data warehouse for time series data analysis A data warehouse system needs to handle more than 10 billion time-series data. It stores data with a storage capacity of 15TB or more in an uncompressed environment ; and a storage capacity of 60TB or more in a compressed environment (theoretical value). Processing is carried out in common database description languages. Processing can be performed while minimizing the load on the client side through ODBC connection. It inputs and outputs 10 billion time-series data at high speed within at least several hours ; and executes the total aggregation processing within minutes. At the time of input／output and aggregation, aggregation is performed at high speed without depending on a specific data structure such as a specific index. The entire process from input／output, totalization, and graph display of results can be completed by a consistent GUI (Graphical User Interface) operation.

B. Storage system for large-scale data storage Storage capacity must be 1PB or more. It inputs and outputs data at a rate of 20 Gbyte／s or more (the oretical value) from the computer system of C installed on the network infrastructure of D. Data can be input and output directly or indirectly from the Windows system. A system configuration that maximizes IOPS is desirable.

C. High performance computer system A high-speed computer system for analyzing the data stored in A and B by a deep learning model is configured. It is desirable that the combined theoretical peak performance of the high-performance computer system is 4.5 PFLOPS or more. The floating-point type conforms to the IEEE 754 standard, but the precision may be depreciated. It is desirable that the combined theoretical peak performance in double precision is 250 TFLOPS or more. The total capacity of the installed memory shall be 6.5TB or more. The computer system is equipped with a spare platform that enables simple AI calculations for five or more units. Using Microsoft Windows as the OS of the high-speed computer system is desirable.

D. Network infrastructure A, B and C are connected to each other with a high transfer bandwidth. A, B and C are connected by 40GbE or more ; while B and C are connected with a transfer bandwidth of 200Gbps or more. A, B and C are connected by 100GbE and connected to a service network installed outside GAMA.

E. Management functions GAMA has an operable management system, in which A, B, C, and D are directly connected from two management rooms at a distance of 100 m or less from GAMA server room. The management network shall be laid independently so as not to affect the bandwidth of the network of D.