

## 資料提供招請に関する公表

次のとおり物品の導入を予定していますので、当該導入に関して資料等の提供を招請します。

令和8年3月18日

国立研究開発法人情報通信研究機構  
契約担当理事 増山 寛

◎調達機関番号 816 ◎所在地番号 13

○第1号

### 1 調達内容

- (1) 品目分類番号 14
- (2) 導入計画物品及び数量 能動的評価基盤構築のための計算機設備 一式
- (3) 調達方法 購入
- (4) 導入予定時期 令和9年度第1・四半期以降
- (5) 調達に必要とされる基本的な要求要件

本計算機設備は、AI学習用計算システム、AI推論用計算システム、汎用計算システム、制御用システム、大容量ストレージシステム、高速ネットワーク等およびシステムを運用するのに十分な無停電電源装置を含む電源設備と冷却設備から構成される。

- ① AI学習用計算システムは、複数の計算ノードから構成され、各計算ノードは2基以上のCPUおよび8基以上のGPGPUを搭載していること。なお、各計算ノードあたりのGPGPU搭載数は均一でなくても構わない。また、システム全体のGPGPU搭載数は560基以上であること。
- ② AI学習用計算システムの各計算ノードが搭載する各GPGPUは、ECC機能を備えた270 GB以上のメモリを有していること。また、メモリ帯域幅は7.7 TB/s以上、GPU間の相互接続帯域幅は双方向で1.8 TB/s以上であること。
- ③ 単一GPGPUでの理論演算性能は、単精度浮動小数点演算で75 TFLOPS以上であること。
- ④ AI学習用計算システムの各計算ノードの主記憶容量は、単一CPUあたり1 TB以上搭載していること。
- ⑤ AI推論用計算システムは、推論に特化した複数の計算ノードから構成される。AI学習用計算システムの計算ノードと同一のアーキテクチャでも異なるアーキテクチャでも構わない。
- ⑥ 汎用計算システムは、複数の計算ノードで構成され、各計算ノードは2基以上のCPUを搭載していること。また、システム全体のCPU搭載数は100基以上であること。
- ⑦ 汎用計算システムの各計算ノードの主記憶容量は、単一CPUあたり1 TB以上搭載していること。
- ⑧ 制御用システムは、複数の制御用サーバから構成され、システム全体のログインノードやバッチ処理サーバ等を含む、システム全体の制御に必要な機能を有すること。
- ⑨ 大容量ストレージシステムは、100 PiB以上の実効容量を有し、高速かつ高信頼な並列ファイルシステムを提供すること。また、すべてのサーバから高速にアクセス可能である

こと。

- ⑩ 高速ネットワークは、25 Gbps以上の帯域を有しており、すべてのサーバおよびストレージと接続すること。
  - ⑪ オペレーティングシステムはUNIXに準拠し、AI学習用計算システム、AI推論用計算システムおよび汎用計算システムはバッチ処理機能を有すること。
  - ⑫ AI学習用およびAI推論用計算システムは、FORTRAN, C, C++等の言語、MPI等の並列プログラミング環境、開発環境およびAI学習用またはAI推論用の各種ツール群を備えていること。
  - ⑬ 無停電電源装置を含む電源設備および冷却設備を備えていること。
  - ⑭ ハードウェアおよびソフトウェアの保守支援体制を備えること。
- 2 資料及びコメントの提供方法 上記1(2)の物品に関する一般的な参考資料及び同(5)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。
- (1) 資料等の提供期限 令和8年4月27日17時00分まで必着のこと。
  - (2) 提出先 〒619-0289 京都府相楽郡精華町光台3-5 情報通信研究機構ユニバーサルコミュニケーション研究所総合企画室 共通基盤グループ 田中 康司 電話0774-98-6878 E-mail ci-post@khn.nict.go.jp
- 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。
- (1) 交付期間 令和8年3月18日から令和8年4月27日まで。
  - (2) 交付場所 情報通信研究機構ホームページの調達情報よりダウンロードして入手すること。
- 4 説明会の開催 なし
- 5 その他 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
- (1) Classification of the products to be procured: 14
  - (2) Nature and quantity of the products to be purchased: Computer Facilities for the Development of Active LLM Evaluation Platform, 1 set
  - (3) Type of the procurement: Purchase
  - (4) Basic requirements of the procurement: The computing facilities shall consist of an AI training computing system, an AI inference computing system, a general-purpose computing system, a control system, a large-capacity storage system, a high-speed network, as well as power supply and cooling facilities including uninterruptible power supplies (UPS) sufficient for stable system operation.
    - ① The AI training computing system shall

be composed of multiple compute nodes. Each compute node shall be equipped with at least two (2) CPUs and eight (8) or more GPGPUs. The number of GPGPUs per compute node does not need to be uniform. The total number of GPGPUs across the entire system shall be at least 560.

- ② Each GPGPU installed in the compute nodes of the AI training computing system shall have at least 270 GB of memory with ECC functionality. The memory bandwidth shall be at least 7.7 TB/s, and the inter-GPU interconnect bandwidth shall be at least 1.8 TB/s bidirectional.
- ③ The theoretical peak performance of a single GPGPU shall be at least 75 TFLOPS for single-precision floating-point operations.
- ④ Each compute node of the AI training computing system shall be equipped with at least 1 TB of main memory per CPU.
- ⑤ The AI inference computing system shall consist of multiple compute nodes specialized for inference. The architecture may be the same as or different from that of the compute nodes used in the AI training computing system.
- ⑥ The general-purpose computing system shall consist of multiple compute nodes, each equipped with at least two (2) CPUs. The total number of CPUs across the entire system shall be at least 100.
- ⑦ Each compute node of the general-purpose computing system shall be equipped with at least 1 TB of main memory per CPU.
- ⑧ The control system shall consist of multiple control servers and provide the necessary functions for overall system management, including login nodes, batch processing servers, and other system-wide control functions.
- ⑨ The large-capacity storage system shall provide an effective storage capacity of at least 100 PiB and shall offer a high-performance, highly reliable parallel file system. The storage system shall be accessible at high speed from all servers.
- ⑩ The high-speed network shall provide a bandwidth of at least 25 Gbps and shall connect all servers and storage systems.
- ⑪ The operating system shall be compliant with UNIX standards. The AI training computing system, AI inference computing

system, and general-purpose computing system shall support batch processing functionality.

- ⑫ The AI training and AI inference computing systems shall support programming languages such as FORTRAN, C, and C++, parallel programming environments such as MPI, development environments, and a variety of tools for AI training and AI inference.
  - ⑬ The facilities shall include power supply equipment, such as uninterruptible power supplies (UPS), and cooling systems.
  - ⑭ A maintenance and support framework for both hardware and software shall be provided.
- (5) Time limit for the submission of the requested material: 17:00 27 April 2026
- (6) Contact point for the notice: Kouji Tanaka,  
Common Infrastructure Group, General Planning  
Office, Universal Communication Research  
Institute, National Institute of Information  
and Communications Technology, 3-5  
Hikaridai Seika-cho Soraku-gun Kyoto-fu  
619-0289 Japan T E L 0774-98-6878  
E-mail ci-post@khn.nict.go.jp

能動的評価基盤構築のための計算機設備

導入説明書

令和 8 年 3 月

国立研究開発法人情報通信研究機構

## I. 導入の目的等

国立研究開発法人情報通信研究機構（以下「当機構」という。）は、大規模言語モデル（LLM）に関する研究プロジェクトを推進している。民間企業等における AI 開発に必要な大量・高品質な日本語を中心とする学習用言語データを整備して民間企業等へ提供するとともに、進化する AI を評価可能な能動的評価基盤を構築するための世界最高水準の処理性能を有した大規模計算機設備を整備する。

## II. 調達内容

1. 導入計画物品（または調達計画特定役務）および数量  
能動的評価基盤構築のための計算機設備 一式  
（据付、配線、調整等を含む。）
2. 調達方法  
購入
3. 導入予定時期（調達予定時期）  
令和 9 年度 第 1・四半期以降
4. 導入場所  
国立研究開発法人情報通信研究機構  
ユニバーサルコミュニケーション研究所 計算機棟サーバ室（予定）
5. 資料等の提供期限および提供先等
  - 1) 提供期限  
令和 8 年 4 月 27 日（月）17:00（郵送の場合は必着のこと）
  - 2) 提供先（連絡窓口）  
〒619-0289  
京都府相楽郡精華町光台 3-5  
国立研究開発法人情報通信研究機構  
ユニバーサルコミュニケーション研究所 総合企画室 共通基盤グループ  
田中 康司  
TEL：0774-98-6878  
E-mail：ci-post@khn.nict.go.jp
  - 3) その他
    - (1) 資料等を提供する際には、組織の代表者名で、本件招請に対する応募の意思を明確に示す書面で提供すること。
    - (2) 提供資料等に関する照会先を明記すること。
    - (3) 提供資料等は日本語で可能な限り電子ファイルにて提出すること。

6. 追加情報の紹介先  
追加情報の照会または本説明書に関する問合せは、日本語文章により上記 5. 2) の「連絡窓口」へ行うこと。
7. 導入説明会の開催日時および開催場所等  
導入説明会は開催しない。
8. 設置条件
  - 1) 設置スペース
    - (1) 計算機設置スペース (2 階)
      - i. 1 スパンが W17,000 mm×D6,500 mm×H4,100 mm (約 100 m<sup>2</sup>) のスペースが 2 か所および W17,000 mm×D6,000 mm×H4,100 mm (約 102 m<sup>2</sup>) のスペースが 3 か所で 1 つのサーバ室を形成しており、同様のスペースのサーバ室が廊下を挟んで存在する。
      - ii. スラブ床積載荷重は 20,000 N/m<sup>2</sup>であり、H4,100 mm にはフリーアクセスフロアを想定した H600 mm を含んでいる。
      - iii. 1 スパン単位でラックを設置することを考えており、空冷方式の場合はホットアイルコンテインメントを形成すること。
    - (2) 無停電電源装置設置スペース (1 階)
      - i. 1 スパンが W7,500 mm×D6,500 mm×H3,820 mm (約 48 m<sup>2</sup>) が 2 か所および W7,500 mm×D6,000 mm×H3,820 mm (約 45 m<sup>2</sup>) が 4 か所で W15,000 mm×D18,500 mm×H3,820 mm (約 277 m<sup>2</sup>) のスペースを形成し、同様のスペースが基幹無停電電源装置群を挟んで存在する。
      - ii. スラブ床積載荷重は 30,000 N/m<sup>2</sup>である。
      - iii. 上記のサーバ室全体の無停電電源装置設置を賄う必要があり、効率的な配置が求められる。
  - 2) 電源設備・冷却設備
    - (1) 電源設備は、屋外に屋外キュービクルが設置されており、高圧フィーダまでは実装されていることから、高圧フィーダ以降の設備が必要となる。
    - (2) 冷却設備は、室外機等は屋上に設置するが、効率的な配置が求められる。なお、空冷・水冷両方式とも対応可能である。

### III. 基本的要求要件

#### 1. 要求要件の概要

本計算機設備は、AI 学習用計算システム、AI 推論用計算システム、汎用計算システム、制御用システム、大容量ストレージシステム、高速ネットワーク等およ

びシステムを運用するのに十分な無停電電源装置を含む電源設備と冷却設備から構成される。

- 1) AI 学習用計算システムは、複数の計算ノードから構成され、各計算ノードは 2 基以上の CPU および 8 基以上の GPGPU を搭載していること。なお、各計算ノードあたりの GPGPU 搭載数は均一でなくても構わない。また、システム全体の GPGPU 搭載数は 560 基以上であること。
- 2) AI 学習用計算システムの各計算ノードが搭載する各 GPGPU は、ECC 機能を備えた 270 GB 以上のメモリを有していること。また、メモリ帯域幅は 7.7 TB/s 以上、GPU 間の相互接続帯域幅は双方向で 1.8 TB/s 以上であること。
- 3) 単一 GPGPU での理論演算性能は、単精度浮動小数点演算で 75 TFLOPS 以上であること。
- 4) AI 学習用計算システムの各計算ノードの主記憶容量は、単一 CPU あたり 1 TB 以上搭載していること。
- 5) AI 推論用計算システムは、推論に特化した複数の計算ノードから構成される。AI 学習用計算システムの計算ノードと同一のアーキテクチャでも異なるアーキテクチャでも構わない。
- 6) 汎用計算システムは、複数の計算ノードで構成され、各計算ノードは 2 基以上の CPU を搭載していること。また、システム全体の CPU 搭載数は 100 基以上であること。
- 7) 汎用計算システムの各計算ノードの主記憶容量は、単一 CPU あたり 1 TB 以上搭載していること。
- 8) 制御用システムは、複数の制御用サーバから構成され、システム全体のログインノードやバッチ処理サーバ等を含む、システム全体の制御に必要な機能を有すること。
- 9) 大容量ストレージシステムは、100 PiB 以上の実効容量を有し、高速かつ高信頼な並列ファイルシステムを提供すること。また、すべてのサーバから高速にアクセス可能であること。
- 10) 高速ネットワークは、25 Gbps 以上の帯域を有しており、すべてのサーバおよびストレージと接続すること。
- 11) オペレーティングシステムは UNIX に準拠し、AI 学習用計算システム、AI 推論用計算システムおよび汎用計算システムはバッチ処理機能を有すること。
- 12) AI 学習用および AI 推論用計算システムは、FORTRAN, C, C++等の言語、MPI 等の並列プログラミング環境、開発環境および AI 学習用または AI 推論用の各種ツール群を備えていること。
- 13) 無停電電源装置を含む電源設備および冷却設備を備えていること。
- 14) ハードウェアおよびソフトウェアの保守支援体制を備えること。

## 2. 要求要件の具体的説明

- 1) AI 学習用計算システム
  - (1) 複数の計算ノードから構成されること。
  - (2) 各計算ノードは 2 基以上の CPU を搭載していること。

- (3) 各計算ノードは単一 CPU あたり 1 TB 以上の ECC 機能を備えたメモリを搭載していること。
- (4) 各計算ノードはシステムストレージとして十分な容量の SSD を 2 基以上、データストレージとして RAID5 (ホットスペア 1 基) でフォーマットした際の実効容量が 10 TB 以上の SSD を搭載していること。
- (5) 各計算ノードは 8 基以上の GPGPU を搭載していること。なお、各計算ノードあたりの GPGPU 搭載数は均一でなくても構わない。また、システム全体の GPGPU 搭載数は 560 基以上であること。
- (6) 各 GPGPU は ECC 機能を備えた 180 GB 以上のメモリを有していること。また、メモリ帯域幅は 7.7 TB/s 以上、GPU 間の相互接続帯域幅は双方向で 1.8 TB/s 以上であること。
- (7) 単一 GPGPU での理論演算性能は、単精度浮動小数点演算で 75 TFLOPS 以上であること。
- (8) 各計算ノードのネットワークインタフェースは、400 Gbps 以上の帯域を有した RDMA 等の高速通信プロトコルが利用可能なポート、25 Gbps 以上の帯域での通信が可能なポートおよび IPMI を介した制御が可能なポートを有していること。

## 2) AI 推論用計算システム

- (1) 複数の計算ノードから構成されること。
- (2) AI 推論に特化した計算ノードであること。
- (3) AI 学習用計算システムの計算ノードと同一のアーキテクチャでも異なるアーキテクチャでも構わない。

## 3) 汎用計算システム

- (1) 複数の計算ノードから構成されること。
- (2) 各計算ノードは 2 基以上の CPU を搭載していること。また、システム全体の CPU 搭載数は 100 基以上であること。
- (3) 各計算ノードは単一 CPU あたり 1 TB 以上の ECC 機能を備えたメモリを搭載していること。
- (4) 各計算ノードはシステムストレージとして十分な容量の SSD を 2 基以上、データストレージとして RAID5 (ホットスペア 1 基) でフォーマットした際の実効容量が 10 TB 以上の SSD を搭載していること。
- (5) 各計算ノードのネットワークインタフェースは、400 Gbps 以上の帯域を有した RDMA 等の高速通信プロトコルが利用可能なポート、25 Gbps 以上の帯域での通信が可能なポートおよび IPMI を介した制御が可能なポートを有していること。

## 4) 制御用システム

- (1) 複数の制御用サーバから構成されること。

- (2) 本計算機設備全体のログインノードやバッチ処理サーバ等を含む、本計算機設備全体の制御に必要な機能を有すること。
- 5) 大容量ストレージシステム
    - (1) RAID6 相当以上の柔軟な冗長構成に対応し、100 PiB 以上の実効容量を有していること。
    - (2) 高速かつ高信頼な並列ファイルシステムを有し、すべてのサーバから高速にアクセス可能であること。
    - (3) 並列ファイルシステムは POSIX API に準拠していること。また、複数クライアント、プロセスから同一ファイルに対するアクセスに対しロック機能を有し、排他制御を行うこと。
  - 6) 高速ネットワーク
    - (1) すべてのサーバおよび大容量ストレージシステム等と冗長構成にて接続すること。
    - (2) 25 Gbps 以上の帯域を有すること。
    - (3) 当機構内上流ネットワークと 100 Gbps 以上の帯域で接続可能であること。
  - 7) インターコネクトネットワーク
    - (1) 高速ネットワークとは物理的に独立したネットワークであり、すべてのサーバおよび大容量ストレージシステム等と冗長構成にて接続すること。
    - (2) 400 Gbps 以上の帯域を有すること。
    - (3) MPI 通信や大容量ストレージシステム等へのアクセスにおいて、RDMA 等の高速通信プロトコルが利用可能であること。
    - (4) 計算用とストレージアクセス用とに分離独立すること。
  - 8) 管理ネットワーク
    - (1) 本計算機設備全体の管理用ネットワークとして、高速ネットワーク・インターコネクトネットワークとは物理的に独立したネットワークを構成すること。
  - 9) ソフトウェア
    - (1) オペレーティングシステムは UNIX に準拠していること (Linux が望ましい)。
    - (2) AI 学習用計算システム、AI 推論用計算システムおよび汎用計算システムは、バッチ処理機能を有していること。
    - (3) AI 学習用計算システムおよび AI 推論用計算システムは、FORTRAN, C, C++等の言語、MPI 等の並列プログラミング環境、開発環境および AI 学習用または AI 推論用の各種ツール群を備えていること。

10) 無停電電源装置を含む電源設備および冷却設備

- (1) 停電時、停電後2分後より本計算機設備のIT機器（サーバ、ストレージ装置、ネットワーク装置等、各システムを構成する上で必要なすべての機器）が安全に停止できる容量の無停電電源装置を含む電源設備を備えること。
- (2) 本計算機設備を運用するのに十分な冷却設備を備えること。また、冷却設備にも無停電電源装置を備えること。
- (3) 冷却方式は、空冷方式・水冷方式どちらか一方、あるいは併用のいずれでも構わない。

11) その他の要件

- (1) ハードウェアは可能な限り多重化した構成とすること。
- (2) 電源投入、切断、リセット等について遠隔操作が可能であること。
- (3) 各設置スペース（計算機設置スペースおよび無停電電源装置設置スペース）において、以下のモニタリング情報を一元管理可能なシステムを備えること。
  - i. 監視カメラによるリアルタイムな映像収集・蓄積・可視化
  - ii. 電力センサー等によるラック単位・分電盤単位でのリアルタイムな消費電力収集・蓄積・可視化
  - iii. 温湿度センサーによるリアルタイムな温湿度収集・蓄積・可視化
- (4) サーバ、ストレージ装置およびネットワーク装置等はEIA規格に準拠した19インチラックもしくはOCPに準拠した21インチラックに搭載すること。
- (5) システムの稼働を把握するため、サーバ、ストレージ装置およびネットワーク装置等において、消費電力、各種利用状況（CPU、GPGPU、メモリ、ストレージおよびネットワークトラフィック等）を適切な粒度で監視（可視化）できる環境を備えること。
- (6) サーバ単位で利用可能なユーザが異なるため、それらをGUI/CUIにて一元管理可能なシステムを提供すること。
- (7) 本計算機設備は、設置・運用開始後、5年間のハードウェア保守、ソフトウェアのライセンス更新およびシステムの安定的運用を支援する保守作業を検討している。運用への支障を最小限に抑えるため、障害の発生を未然に防ぐための定常的な予防保守体制、および緊急時に迅速な対応を可能とする保守体制を備えること。

3. 総合評価試験および総合評価の概要

1) 総合評価試験の概要

今回の招請資料を参考として最終的に確定した仕様書に基づき、当機構が想定する代表的な処理についての実効性能を評価するために、以下の性能に関する評価試験を用意する予定である。

- (1) AI 学習用および AI 推論用計算システムの実効計算性能
- (2) AI 学習用および AI 推論用計算システムのノード間の実効データ転送速度
- (3) 大容量ストレージシステムの実効データ転送速度

## 2) 総合評価の概要

落札は、入札の技術的要件と入札価格を総合評価して行うが、そのうち導入物品の技術的評価は、以下のような内容を含む予定である。

- (1) AI 学習用計算システムの演算性能、搭載メモリの性能・容量、CPU-GPGPU 間の通信バンド幅、ローカルストレージの性能・容量等
- (2) AI 推論用計算システムの演算性能、搭載メモリの性能・容量等
- (3) 汎用計算システムの演算性能、搭載メモリの性能・容量、ローカルストレージの性能・容量等
- (4) 大容量ストレージシステムの性能・容量・冗長度
- (5) 各ネットワークの構成とバンド幅
- (6) 保守性、管理容易性
- (7) 運用管理・保守体制
- (8) システムの設置方式、冷却方式、専有面積、消費電力

## IV. 落札方式の概要

今回の招請資料等を参考にして策定する「仕様書」を満たす提案をした者のうち、予定価格の制限の範囲内であり、かつ

(除算の場合)

当該競争加入者の申し込みに係る性能等の各評価項目の得点合計を当該競争加入者等の入札価格で除して得た数値の最も高い者をもって落札者とする。

(加算の場合)

当該競争加入者の申し込みに係る入札価格に対する得点配分と、性能等の各評価項目の得点合計を合算して得た数値の最も高い者をもって落札者とする。

## V. 提供招請する資料等

上記の基本的要求要件を満たす提案システムを具体的に説明した資料およびコメントの提供を招請する。提供資料は下記の通りとする。

### 1. 全体構成

システム全体の構成図、配置図、概要説明、各構成機器の名称、型名、数量および性能等を明記すること。特に、保守性、フットプリント、荷重および冷却性能を考慮した具体的なラックの構成を示すこと。地震時の事故対策措置を示すこと。また、ネットワーク等の配線図、配線長を示すこと。

## 2. AI 学習用計算システム

少なくとも以下の項目を含むこと。

- 1) CPUの種類、動作周波数、コア数、キャッシュサイズおよびノード筐体に搭載可能なCPU数・コア数（最小、最大、増設単位）
- 2) メモリの種類、容量、動作周波数およびノード筐体に搭載可能な容量（最小、最大、増設単位）
- 3) GPGPUの種類、動作周波数、コア数、メモリ容量、メモリバンド幅およびノード筐体に搭載可能なGPGPU数（最小、最大、増設単位）
- 4) ローカルストレージの種類、容量、データ転送速度およびノード筐体に搭載可能なストレージ容量・台数（最小、最大、増設単位）
- 5) ネットワークインタフェースの種類、インタフェース数およびデータ転送速度
- 6) 電源の構成と仕様および予測ピーク消費電力
- 7) ノード筐体の仕様

## 3. AI 推論用計算システム

AI 学習用計算システムと同一のアーキテクチャである場合は、上記 AI 学習用計算システムで記載した項目、異なるアーキテクチャである場合は、上記 AI 学習用計算システムで記載した項目に類似する項目を含むこと。

## 4. 汎用計算システム

少なくとも以下の項目を含むこと。

- 1) CPUの種類、動作周波数、コア数、キャッシュサイズおよびノード筐体に搭載可能なCPU数・コア数（最小、最大、増設単位）
- 2) メモリの種類、容量、動作周波数およびノード筐体に搭載可能な容量（最小、最大、増設単位）
- 3) ローカルストレージの種類、容量、データ転送速度およびノード筐体に搭載可能なストレージ容量・台数（最小、最大、増設単位）
- 4) ネットワークインタフェースの種類、インタフェース数およびデータ転送速度
- 5) 電源の構成と仕様および予測ピーク消費電力
- 6) ノード筐体の仕様

## 5. 制御用システム

少なくとも以下の項目を含むこと。

- 1) CPUの種類、動作周波数、コア数、キャッシュサイズおよびノード筐体に搭載可能なCPU数・コア数（最小、最大、増設単位）
- 2) メモリの種類、容量、動作周波数およびノード筐体に搭載可能な容量（最小、最大、増設単位）
- 3) ローカルストレージの種類、容量、データ転送速度およびノード筐体に搭載可能なストレージ容量・台数（最小、最大、増設単位）

- 4) ネットワークインタフェースの種類、インタフェース数およびデータ転送速度
  - 5) 電源の構成と仕様および予測ピーク消費電力
  - 6) ノード筐体の仕様
6. 大容量ストレージシステム  
少なくとも以下の項目を含むこと。
- 1) ファイルサーバの構成と仕様
  - 2) ディスクアレイ装置の構成と仕様およびディスクアレイ装置に搭載可能なストレージ容量・台数（最小、最大、増設単位）
  - 3) ディスクアレイ装置とファイルサーバの接続に関する構成と仕様
  - 4) 並列ファイルシステムの構成と仕様
  - 5) 大容量ストレージシステムの各種予測性能
  - 6) 電源の構成と仕様および予測ピーク消費電力
7. ネットワーク  
少なくとも以下の項目を含むこと。
- 1) 高速ネットワークの構成と仕様および各種予測性能
  - 2) インターコネクトネットワークの構成と仕様および各種予測性能
  - 3) 管理系ネットワークの構成と仕様および各種予測性能
8. ソフトウェア  
少なくとも以下の項目を含むこと。
- 1) オペレーティングシステムの名称、バージョン
  - 2) バッチ処理機能ソフトウェアの名称、バージョン
  - 3) 言語、並列プログラミング環境、開発環境および AI 学習用または AI 推論用のための各種ツール群の名称、バージョン
9. 無停電電源装置を含む電源設備および冷却設備  
少なくとも以下の項目を含むこと。
- 1) 無停電電源装置を含む電源設備の構成と仕様
  - 2) 冷却設備の構成と仕様
  - 3) 1次側受変電設備から各機器までの配電効率
  - 4) 本計算機設備全体の予測ピーク消費電力および予測 PUE 値
10. その他
- 1) モニタリングシステム等、各種管理システムの構成と仕様に関する資料
  - 2) 保守支援体制について具体的内容を明記した資料
  - 3) 提案システムの導入実績に関する資料
  - 4) 提案システムの概算見積額に関する資料

搬入、据付、配線、調整、付帯工事および保守等すべてを含み、個別に額を明記すること。

5) 納入に要する期間に関する資料

11. 上記の他に、参考となるような情報（例えば、本要求性能とは違う GPGPU を搭載したサーバの情報等）があれば是非提供されたい。

※ 全体構成における資料の提供が困難な場合、部分構成における資料の提供でも構わないが、提案システムの保守契約、導入実績、概算見積額および納入に要する期間に関する資料も併せて提供すること。

## VI. その他

提供各社に対し、必要に応じ、本提案資料の記載内容等についてヒアリング等を行う場合がある。この場合において、提案各社は誠実に対応すること。

## I. 件名英訳

Computer Facilities for the Development of Active LLM Evaluation Platform

## II. 要求要件英訳

The computing facilities shall consist of an AI training computing system, an AI inference computing system, a general-purpose computing system, a control system, a large-capacity storage system, a high-speed network, as well as power supply and cooling facilities including uninterruptible power supplies (UPS) sufficient for stable system operation.

1. The AI training computing system shall be composed of multiple compute nodes. Each compute node shall be equipped with at least two (2) CPUs and eight (8) or more GPGPUs. The number of GPGPUs per compute node does not need to be uniform. The total number of GPGPUs across the entire system shall be at least 560.
2. Each GPGPU installed in the compute nodes of the AI training computing system shall have at least 270 GB of memory with ECC functionality. The memory bandwidth shall be at least 7.7 TB/s, and the inter-GPU interconnect bandwidth shall be at least 1.8 TB/s bidirectional.
3. The theoretical peak performance of a single GPGPU shall be at least 75 TFLOPS for single-precision floating-point operations.
4. Each compute node of the AI training computing system shall be equipped with at least 1 TB of main memory per CPU.
5. The AI inference computing system shall consist of multiple compute nodes specialized for inference. The architecture may be the same as or different from that of the compute nodes used in the AI training computing system.
6. The general-purpose computing system shall consist of multiple compute nodes, each equipped with at least two (2) CPUs. The total number of CPUs across the entire system shall be at least 100.
7. Each compute node of the general-purpose computing system shall be equipped with at least 1 TB of main memory per CPU.
8. The control system shall consist of multiple control servers and provide the necessary functions for overall system management, including login nodes, batch processing servers, and other system-wide control functions.
9. The large-capacity storage system shall provide an effective storage capacity of at least 100 PiB and shall offer a high-performance, highly reliable parallel file system. The storage system shall be accessible at high speed from all servers.
10. The high-speed network shall provide a bandwidth of at least 25 Gbps and shall connect all servers and storage systems.

11. The operating system shall be compliant with UNIX standards. The AI training computing system, AI inference computing system, and general-purpose computing system shall support batch processing functionality.
12. The AI training and AI inference computing systems shall support programming languages such as FORTRAN, C, and C++, parallel programming environments such as MPI, development environments, and a variety of tools for AI training and AI inference.
13. The facilities shall include power supply equipment, such as uninterruptible power supplies (UPS), and cooling systems.
14. A maintenance and support framework for both hardware and software shall be provided.