Governo do Estado do Rio de Janeiro
Gabinete de Segurança Institucional
Subsecretaria de Avaliação de Cenários e Inteligência Estratégica

TERMO DE REFERÊNCIA

1. INTRODUÇÃO (art. 6º, inc. XXIII, da Lei n.º 14.133/2021 c/c art. 16 do Decreto n.º 48.816/2023)

1.1. Propósito: AQUISIÇÃO DE BENS COMUNS, a saber:

ITEM	BREVE DESCRIÇÃO	QUANT
1	**GPUs com 48GB GDDR6** Passive GPU ECC, para o DataCenter GSI, compatíveis com os servidores Lenovo Thinksystem já adquiridos	3
2	**Módulos de 64GB** (sessenta e quatro gigabits) de **Memória TruRAM DDR4 3200Mhz,** totalizando 512 GB (quinhentos e doze gigabits), **para servidores Lenovo ThinkSystem SR650 V2 Type 7Z73 CT01WW**	8
3	12 (doze) **placas de rede 25GbE para servidores Lenovo ThinkSystem SR650 V2 Type 7Z73 CT01WW**	12
4	24 (vinte e quatro) **cabos de cobre** passivos, **5 metros** de comprimento, **com conexão direta (DAC) SFP28 25GbE para DataCenter**	24

1.2. Justificativa da necessidade da contratação (art. 6º, inc. XXIII, alínea "b" c/c art. 17, inc. I do Decreto n.º 48.816/2023)

Trata-se de demanda apresentada por esta Coordenação, que necessita de equipamentos para implementação, melhoria e perenidade das ações ora desempenhadas, razão pela qual delineamos abaixo os critérios norteadores que viabilizam a aquisição dos bens comuns ora relacionados, imprescindíveis ao bom desenvolvimento das atividades da Subsecretaria de Inteligência Estratégica de Estado (GSI/SSIE), que agregarão recursos, contribuirão e estarão disponíveis para fruição por parte de toda a Secretaria de Estado do Gabinete de Segurança Institucional.

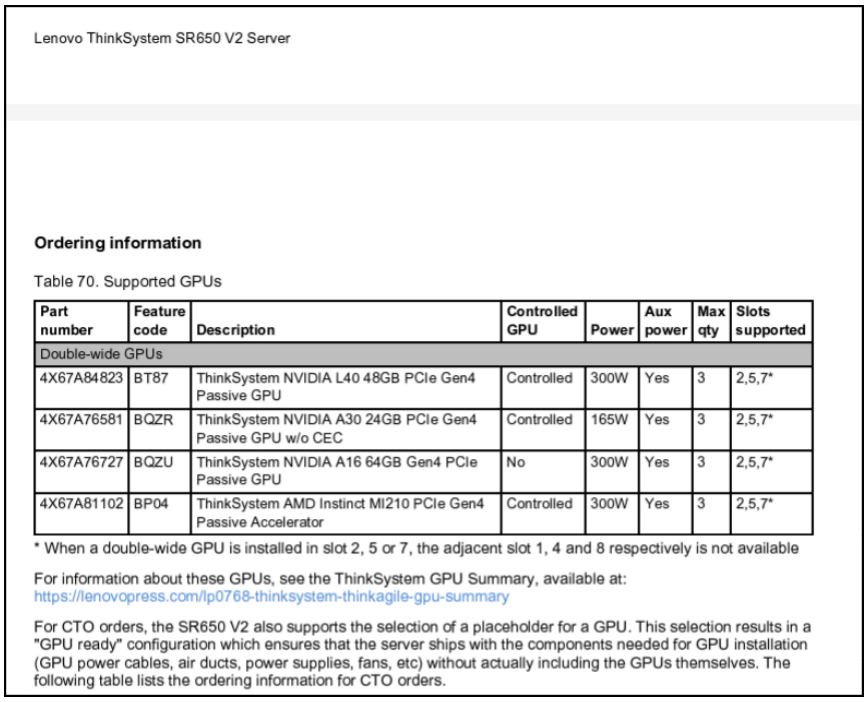
- **Item 1 - 03 (três) GPU's 48GB GDDR6 (para servidores Lenovo ThinkSystem SR650 V2 Type 7Z73) – Previsto no PCA 2025**

A aquisição de 03 (três) GPU's para 01 (um) dos servidores que compõem o DataCenter do GSI/RJ deve, antes de tudo, respeitar um **pré-requisito**, qual seja, ser **absolutamente compatível com** a marca e modelo dos servidores de aplicação recentemente adquiridos por esta Secretaria de Estado, a saber: **Servidores Lenovo ThinkSystem SR650 V2 Type 7Z73-CT01WW**

Segundo consta da Tabela 70 (*Table 70. Supported GPUs*), pág. 84 do Manual do Produto (*Product Guide*, acessível por meio do sítio eletrônico https://lenovopress.lenovo.com/lp1392.pdf), **existem apenas 4 (quatro) modelos de GPUs** *(Fig.1 – Quadro informativo)* **homologados para o servidor** em questão - que suporta somente 3 (três) GPUs do tipo double-wide (que ocupam 2 slots PCIe) em cada máquina-, reduzindo sobremaneira o espectro de possibilidades, notadamente quando se estima a **necessidade de**:

- **Mínimo de 128GB** de memória VRAM GDDR6 **com as 3 GPUs somadas**
- Barramento de memória de **256bit ou superior** (em cada GPU)
- **Mínimo de 256 TFLOPS** (teraflops) de desempenho de ponto flutuante **com as 3 GPUs somadas**

Em síntese, a aquisição dessas GPUs é uma decisão estratégica que fortalece a capacidade da agência de inteligência em realizar operações avançadas de análise de dados, predição e modelagem de cenários. Esses equipamentos não apenas garantem a eficácia operacional, mas também asseguram a segurança, a confiabilidade e a conformidade com os mais altos padrões de qualidade, alinhando-se às demandas complexas e dinâmicas da inteligência estratégica moderna.



Lenovo ThinkSystem SR650 V2 Server

Ordering information

Table 70. Supported GPUs

Part number	Feature code	Description	Controlled GPU	Power	Aux power	Max qty	Slots supported
Double-wide GPUs							
4X67A84823	BT87	ThinkSystem NVIDIA L40 48GB PCIe Gen4 Passive GPU	Controlled	300W	Yes	3	2,5,7*
4X67A76581	BQZR	ThinkSystem NVIDIA A30 24GB PCIe Gen4 Passive GPU w/o CEC	Controlled	165W	Yes	3	2,5,7*
4X67A76727	BQZU	ThinkSystem NVIDIA A16 64GB Gen4 PCIe Passive GPU	No	300W	Yes	3	2,5,7*
4X67A81102	BP04	ThinkSystem AMD Instinct MI210 PCIe Gen4 Passive Accelerator	Controlled	300W	Yes	3	2,5,7*

* When a double-wide GPU is installed in slot 2, 5 or 7, the adjacent slot 1, 4 and 8 respectively is not available

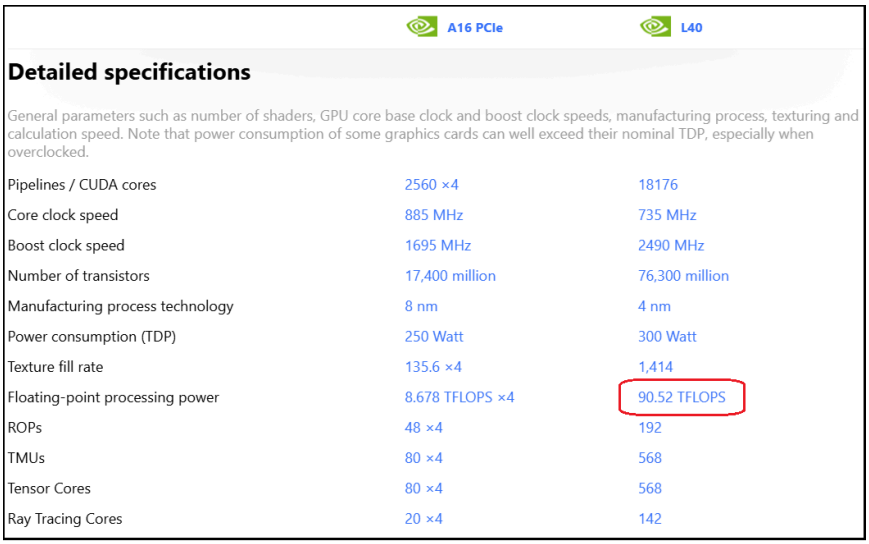
For information about these GPUs, see the ThinkSystem GPU Summary, available at:
https://lenovopress.com/lp0768-thinksystem-thinkagile-gpu-summary

For CTO orders, the SR650 V2 also supports the selection of a placeholder for a GPU. This selection results in a "GPU ready" configuration which ensures that the server ships with the components needed for GPU installation (GPU power cables, air ducts, power supplies, fans, etc) without actually including the GPUs themselves. The following table lists the ordering information for CTO orders.

Fig.1 – Quadro informativo

Adicionalmente, as GPUs devem ser equipadas com tecnologia de resfriamento passivo/controlado e correção de erros (ECC), sendo esta uma necessidade estratégica e operacional da Secretaria de Estado do GSI, que atua em um ambiente de alta complexidade e demanda por processamento de dados em larga escala.

Esses equipamentos são fundamentais para suportar operações avançadas de Machine Learning (ML), Deep Learning (DL), Federated Learning (FL), modelos de linguagem de grande escala (LLMs), Inteligência Artificial Generativa, Big Data, Processamento de Linguagem Natural (NLP), redes neurais, ciência de dados, predição e análise de dados. A justificativa para essa aquisição baseia-se em uma série de fatores técnicos, operacionais e estratégicos que se interconectam para garantir a eficácia das operações de inteligência.



Detailed specifications

General parameters such as number of shaders, GPU core base clock and boost clock speeds, manufacturing process, texturing and calculation speed. Note that power consumption of some graphics cards can well exceed their nominal TDP, especially when overclocked.

	A16 PCIe	L40
Pipelines / CUDA cores	2560 ×4	18176
Core clock speed	885 MHz	735 MHz
Boost clock speed	1695 MHz	2490 MHz
Number of transistors	17,400 million	76,300 million
Manufacturing process technology	8 nm	4 nm
Power consumption (TDP)	250 Watt	300 Watt
Texture fill rate	135.6 ×4	1,414
Floating-point processing power	8.678 TFLOPS ×4	90.52 TFLOPS
ROPs	48 ×4	192
TMUs	80 ×4	568
Tensor Cores	80 ×4	568
Ray Tracing Cores	20 ×4	142

Fig. 2 - Quadro comparativo

Em um cenário onde a quantidade e a complexidade dos dados aumentam exponencialmente, a multiplicidade de GPUs com pelo menos 48GB de memória GDDR6 destacam-se pela capacidade de processamento paralelo, essencial para lidar com cargas de trabalho intensivas. As unidades em questão permitem o treinamento de modelos de ML e DL em grandes conjuntos de dados, reduzindo significativamente o tempo necessário para tarefas como reconhecimento de padrões, análise preditiva e detecção de anomalias. A capacidade de processar grandes volumes de dados de forma rápida e eficiente é crucial para operações de inteligência que dependem de insights precisos e em tempo real, especialmente quando se trata de monitorar ameaças emergentes ou analisar informações provenientes de múltiplas fontes, como interceptações de comunicações, imagens de satélite e redes sociais.

A demanda por modelos de IA generativa e LLMs, como os baseados em arquiteturas Transformer, exige recursos computacionais robustos. Esses modelos são utilizados para gerar relatórios automatizados, tradução de idiomas, análise de sentimentos e simulação de cenários, tarefas que são cada vez mais críticas para a produção de conhecimento estratégico. As GPUs com 48GB de memória são ideais para treinar e executar esses modelos, pois permitem o armazenamento e processamento de grandes matrizes de dados necessários para operações complexas.

Além disso, a tecnologia ECC garante a integridade dos dados durante operações prolongadas, minimizando erros que poderiam comprometer a precisão e a confiabilidade dos resultados.

No contexto do Federated Learning (FL), que permite o treinamento de modelos de ML de forma descentralizada sem a necessidade de centralizar dados sensíveis, as GPUs com alta capacidade de memória são essenciais. Essa abordagem é particularmente relevante para agências de inteligência, que devem garantir a conformidade com leis de proteção de dados e a segurança das informações. A capacidade dessas GPUs de suportar processamento distribuído em servidores Lenovo Thinksystem permite que a agência treine modelos de forma segura e eficiente, mantendo a privacidade dos dados e a integridade das operações.

A análise de Big Data em tempo real é outra área onde as GPUs de alto desempenho se mostram indispensáveis. A capacidade de processar fluxos contínuos de dados e executar algoritmos complexos em tempo real é crucial para a identificação de ameaças emergentes, o monitoramento de infraestruturas críticas e o suporte à tomada de decisões estratégicas. Além disso, o Processamento de Linguagem Natural (NLP), utilizado para análise de textos, tradução automática e detecção de sentimentos, beneficia-se significativamente da capacidade de processamento paralelo e da alta capacidade de memória dessas GPUs. Modelos avançados de NLP, como os baseados em transformers, exigem recursos computacionais substanciais, e a tecnologia de resfriamento passivo/controlado garante a estabilidade térmica necessária para operações prolongadas.

No campo das redes neurais e do Deep Learning, as GPUs com 48GB de memória GDDR6 são otimizadas para manipular grandes volumes de dados e executar operações matriciais complexas. Essas capacidades são essenciais para tarefas como reconhecimento de imagens, análise de vídeos e detecção de objetos, que são críticas para a inteligência de defesa e segurança pública. A tecnologia ECC, por sua vez, assegura a precisão dos cálculos, um fator indispensável para modelos que exigem alta confiabilidade e precisão.

A análise preditiva, utilizada para antecipar eventos como ataques cibernéticos, movimentos de grupos hostis ou crises humanitárias, também se beneficia do poder de processamento dessas GPUs. A capacidade de executar algoritmos de predição baseados em ML e DL de forma rápida e eficiente permite que a agência identifique padrões e tendências em grandes conjuntos de dados históricos e em tempo real, fortalecendo sua capacidade de prevenção e resposta.

A proteção de infraestruturas críticas, como redes de energia, sistemas de transporte e comunicações, depende da análise contínua de dados e da detecção de ameaças. As GPUs permitem a execução de modelos de detecção de anomalias e simulações de cenários de ataque, essenciais para garantir a segurança e a resiliência dessas infraestruturas. Além disso, a conformidade com requisitos de segurança e confiabilidade é um aspecto crítico para agências de inteligência. A tecnologia ECC das GPUs garante a correção de erros de memória, assegurando a integridade dos dados durante operações prolongadas e de alta precisão.

Por fim, a eficiência energética e a sustentabilidade são fatores que não podem ser negligenciados. As GPUs com tecnologia de resfriamento passivo/controlado reduzem o consumo de energia e os custos operacionais, ao mesmo tempo que mantêm o desempenho necessário para tarefas de alta complexidade. Essa eficiência é particularmente importante em operações contínuas, onde a estabilidade e a redução de custos são prioridades.

- **Item 2 - 8 (oito) Módulos de 64GB** (sessenta e quatro gigabits) de **Memória TruRAM DDR4 3200Mhz,** totalizando 512 GB (quinhentos e doze gigabits), **para servidores Lenovo ThinkSystem SR650 V2 Type 7Z73 CT01WW**

A recente aquisição de servidores de aplicação próprios por parte do Gabinete de Segurança Institucional permitiu ao órgão construir uma infraestrutura robusta para processamento de dados, execução de softwares sensíveis e fornecimento de acesso seguro a recursos de Tecnologia da Informação e Comunicação (TIC) não apenas para a Agência de Inteligência Estratégica de Estado (GSI/SSIE), mas também para todos os demais setores e profissionais que compõem esta Secretaria de Estado.

Este avanço proporcionou um aumento significativo na qualidade, eficiência e capacidade operacional da GSI/SSIE e, de forma subsidiária, de todos os demais setores do GSI, contudo, a utilização intensiva desses servidores evidenciou que a memória RAM inicialmente instalada (64GB) tornou-se insuficiente para atender à crescente demanda.

A quantidade de memória RAM originária/inicial (64GB) mostrou-se adequada somente durante a fase de implementação dos servidores, contudo, ante à adesão em massa e a utilização intensiva dos recursos pelos profissionais do GSI, rapidamente houve s**aturação da memória disponível,** especialmente picos de operação, forçando o sistema a recorrer a memória virtual (swap), o que degrada o desempenho, **impacto no tempo de resposta dos serviços,** com aumento na latência de execução de aplicativos e no processamento de dados sensíveis, além do **comprometimento da estabilidade do sistema e de suas operações,** aumentando o risco de falhas críticas, especialmente em sistemas sensíveis.

- **Item 3 -** 12 (doze) **placas de rede 25GbE para servidores Lenovo ThinkSystem SR650 V2 Type 7Z73 CT01WW**

A substituição das placas de rede atuais (1Gbps) por modelos de maior capacidade (25Gbps) é uma medida essencial para garantir a escalabilidade, desempenho e resiliência da infraestrutura de TI em um datacenter moderno. Essa necessidade é fundamentada em aspectos técnicos e operacionais como o aumento da demanda por Largura de Banda, a Redução de Latência para aplicações críticas, o suporte a tecnologias emergentes (como NVMe over Fabrics - NVMe-oF, Software-Defined Networking - SDN, Armazenamento em rede - NAS/SAN, além de Machine Learning e Inteligência Artificial), otimizando custos e future-froofing, trazendo maior eficiência energética e densidade.

A evolução das aplicações corporativas, especialmente em ambientes de cloud computing, microsserviços e processamento de dados em tempo real, exige taxas de transferência mais elevadas. Sistemas como bancos de dados distribuídos (ex: Cassandra, MongoDB), plataformas de análise de big data (ex: Hadoop, Spark) e soluções de virtualização (ex: VMware, Kubernetes), como as que se pretende implementar nesta Secretaria de Estado do Gabinete de Segurança Institucional, demandam conexões de rede ultrarrápidas para evitar gargalos.

Segundo consta da Tabela 69 (*Table 69. Supported PCI-e Network Adapters*), pág. 82 do Manual do Produto (*Product Guide)*, acessível por meio do sítio eletrônico https://lenovopress.lenovo.com/lp1392.pdf, **existem apenas 8 (oito) modelos de Placas de Rede** *(Fig.3 – Quadro informativo)* **homologados para o servidor** em questão, trabalhando na velocidade desejada de 25 Gbps (25 GbE), reduzindo sobremaneira o espectro de possibilidades, notadamente quando se estipula a **necessidade de**:

- **Mínimo de 25gbps de velocidade**
- Barramento PCI-e (PCI Express) Ethernet
- **Conectividade SFP28**
- **No mínimo duas portas (do tipo SFP28)**
- **Compatibilidade integral com os servidores de aplicação do DataCenter do GSI**

25 Gb Ethernet					
4XC7A84827	BUQK	ThinkSystem AMD X3522 10/25GbE DSFP28 2-Port PCIe Ethernet Adapter (Low Latency)	x8	8	All slots
4XC7A08238	B5T0	ThinkSystem Broadcom 57414 10/25GbE SFP28 2-port PCIe Ethernet Adapter V2	x8	8	All slots
4XC7A08316	BD49	ThinkSystem Broadcom 57454 10/25GbE SFP28 4-port PCIe Ethernet Adapter V2	x16	6	Any 6 slots
4XC7A80566	BNWM	ThinkSystem Broadcom 57504 10/25GbE SFP28 4-port PCIe Ethernet Adapter	x16	6	Any 6 slots
4XC7A08295	BCD6	ThinkSystem Intel E810-DA2 10/25GbE SFP28 2-Port PCIe Ethernet Adapter	x8	8	All slots
4XC7A80267	BP8M	ThinkSystem Intel E810-DA4 10/25GbE SFP28 4-Port PCIe Ethernet Adapter	x16	6	Any 6 slots
4XC7A62580	BE4U	ThinkSystem Mellanox ConnectX-6 Lx 10/25GbE SFP28 2-port PCIe Ethernet Adapter	x8	8	All slots
4XC7A62581	BHE2	ThinkSystem Solarflare X2522-Plus 10/25GbE SFP28 2-Port PCIe Ethernet Adapter	x8	8	All slots
100 Gb Ethernet					

Fig.3 - Quadro informativo

Item 4 - 24 (vinte e quatro) **cabos de cobre** passivos, **5 metros** de comprimento, **com conexão direta (DAC) SFP28 25GbE para DataCenter**

A aquisição de cabos com conexão direta SFP28 25GbE é **indispensável para garantir a alta disponibilidade, desempenho escalável e eficiência energética da infraestrutura de rede do datacenter.** Esses cabos são fundamentais para interligar servidores, switches e sistemas de armazenamento em uma topologia de alta velocidade e baixa latência, atendendo às demandas crescentes de computação em nuvem, virtualização, big data e aplicações críticas em tempo real.

A migração para uma infraestrutura 25GbE (25 Gigabits por segundo) é impulsionada por fatores críticos, tais como o aumento exponencial do tráfego de dados devido à adoção de microsserviços, contêineres (Docker/Kubernetes) e arquiteturas distribuídas, a demanda por baixa latência em aplicações financeiras, sistemas de banco de dados (Oracle RAC, SQL Server AlwaysOn) e plataformas de machine learning, além da eliminação de gargalos de rede em ambientes de hyperconvergência (HCI) e storage definido por software (Ceph, VMware vSAN). Os cabos SFP28 DAC (Direct Attach Copper) oferecem vantagens superiores em relação a soluções tradicionais (cabos de fibra óptica com transceivers separados), como menor latência (nanossegundos vs. microssegundos em transceivers ópticos), custo reduzido (elimina a necessidade de transceivers ópticos caros), consumo energético inferior (em média, 0,5W por cabo vs. 1,5W em transceivers SFP+), e compatibilidade com switches 25GbE/100GbE (future-proofing da infraestrutura).

A adoção de **cabos SFP28 25GbE** proporcionará maior throughput para storage em rede (iSCSI, NVMe-oF) – Evitando congestionamentos em backups e réplicas, Suporte a SDN (Software-Defined Networking) – Permitindo gerência dinâmica de largura de banda, e preparação para futura infraestrutura de 100GbE – Via agregação de links (4x25G = 100G). Esta solução está alinhada com as normas IEEE 802.3by (25GbE) e ANSI/TIA-568 (para cabeamento estruturado), com as recomendações do NIST (SP 800-193) para segurança e resiliência de datacenters, e com a Lei 14.133/2021, no que tange à modernização de infraestrutura pública visando à economicidade.

1.3.Alinhamento com Plano de Contratações Anuais, Plano Estratégico, Lei Orçamentária anual e outros (art. 6º, inc. XXIII, alínea "j", da Lei n.º 14.133/2021 c/c art. 16, Parágrafo Único do Decreto n.º 48.816/2023)

- **Item 1 - 03 (três) GPU's 48GB GDDR6 (para servidores Lenovo ThinkSystem SR650 V2 Type 7Z73) –** Em consulta ao documento SEI de índice 76657809, anexado em 12/06/2024 ao processo SEI-390004/000205/2024, que trata do PCA GSI para o ano de 2025, verificou-se que **há previsão da demanda**

- **Item 2 - 8 (oito) Módulos de 64GB** (sessenta e quatro gigabits) de **Memória TruRAM DDR4 3200Mhz,** totalizando 512 GB (quinhentos e doze gigabits), **para servidores Lenovo ThinkSystem SR650 V2 Type 7Z73 CT01WW** - Em consulta ao documento SEI de índice 76657809, anexado em 12/06/2024 ao processo SEI-390004/000205/2024, que trata do PCA GSI para o ano de 2025, verificou-se que **há previsão da demanda**

- **Item 3** - 12 (doze) **placas de rede 25GbE para servidores Lenovo ThinkSystem SR650 V2 Type 7Z73 CT01WW -** Em consulta ao documento SEI de índice 76657809, anexado em 12/06/2024 ao processo SEI-390004/000205/2024, que trata do PCA GSI para o ano de 2025, verificou-se que **não há** previsão da demanda

- **Item 4** - 24 (vinte e quatro) **cabos de cobre** passivos, **5 metros** de comprimento, **com conexão direta (DAC) SFP28 25GbE para DataCenter -** Em consulta ao documento SEI de índice 76657809, anexado em 12/06/2024 ao processo SEI-390004/000205/2024, que trata do PCA GSI para o ano de 2025, verificou-se que **não há** previsão da demanda

1.4 Resultados Esperados (art. 6º, inc. XXIII, alínea "c" da Lei n.º 14.133/2021 c/c art. 17, inc. I do Decreto n.º 48.816/2023)

Item 1 - 03 (três) GPU's 48GB GDDR6 (para servidores Lenovo ThinkSystem SR650 V2 Type 7Z73)

Problema: Carência de processamento e análise de grandes volumes de dados sensíveis, necessários a operações estratégicas, com a execução de modelos de Large Language Models (LLMs) e outras aplicações de Inteligência Artificial (IA) de modo isolado da rede mundial de computadores, a fim de garantir o sigilo e a compartimentação de dados inerentes às Agências de Inteligência, como a GFI/SSIE

Necessidade: GPUs de alto desempenho com capacidade mínima de 32GB de memória GDDR6 (sendo recomendável a aquisição de GPUs com os 48GB de memória que se mostram como limite dos servidores de aplicação existentes), e que sejam compatíveis com servidores Lenovo Thinksystem Modelo SR650 V2 Type 7Z73.

Resultado Esperado: Processamento local de LLM's, eliminando a necessidade de transferir dados sensíveis para ambientes externos, reduzindo riscos de vazamentos ou interceptações; aumento da eficiência operacional; execução de LLMs e modelos de IA localmente, reduzindo latência e aumentando a velocidade de análise, permitindo respostas rápidas e precisas em cenários críticos; capacidade de Análise de Dados Complexos, realizando análises preditivas, processamento de linguagem natural (NLP) e treinamento de modelos de IA de forma mais eficaz, gerando insights estratégicos com maior precisão; Fortalecimento da autonomia tecnológica da agência, trazendo a capacidade de processar dados localmente, garantindo controle sobre as operações, alinhando-se às diretrizes de segurança e soberania tecnológica da agência.

Item 2 - 8 (oito) Módulos de 64GB (sessenta e quatro gigabits) de **Memória TruRAM DDR4 3200Mhz,** totalizando 512 GB (quinhentos e doze gigabits), **para servidores Lenovo ThinkSystem SR650 V2 Type 7Z73 CT01WW**

Problema: Carência de memória RAM nos servidores de aplicação recentemente adquiridos pelo GSI, que necessitam de upgrade para que possam suportar operações avançadas de Machine Learning (ML), Deep Learning (DL), Federated Learning (FL), modelos de linguagem de grande escala (LLMs), Inteligência Artificial Generativa, Big Data, Processamento de Linguagem Natural (NLP), redes neurais, ciência de dados, predição e análise de dados

Necessidade: No mínimo mais 512GB de memória RAM para os servidores Lenovo recentemente adquiridos

Resultado Esperado: Implementação de um sistema de Macrodados (Big Data), modelos de aprendizado profundo e de linguagem ampla generativa (LLM, MLP e RNN), potencializando os recursos disponíveis e maximizando a produtividade dos funcionários lotados nesta Secretaria de Estado.

Item 3 - 12 (doze) **placas de rede 25GbE para servidores Lenovo ThinkSystem SR650 V2 Type 7Z73 CT01WW**

Problema: Gargalos de largura de banda, latência elevada em aplicações sensíveis, escalabilidade comprometida e ineficiência energética

Necessidade: **Placas de rede 25GbE para servidores Lenovo ThinkSystem SR650 V2 Type 7Z73 CT01WW**

Resultado Esperado: Aumento imediato do throughput, reduzindo o tempo de ransferência em aproximadamente 70%, redução da latência em aplicações críticas, com melhoria de até 50% no tempo de resposta, preparação para arquiteturas modernas como NVMe-oF, SDN e Cloud híbrida, redução no consumo de energía, e caminho aberto para upgrades futuros de infraestrutura de rede na casa dos 100GBPS.

Indicador	Baseline (1Gbps)	Meta (25Gbps)	Ganho Esperado	Impacto Operacional
Largura de Banda Máxima	1 Gbps	25 Gbps	**25x maior**	Elimina gargalos em backups, migração de VMs e transferência de grandes volumes de dados.
Latência Média	800 µs (0.8 ms)	≤150 µs (0.15 ms)	**~80% menor**	Melhora desempenho em transações em tempo real (bancos de dados, trading financeiro).
Throughput Efetivo*	~950 Mbps	~23 Gbps	**24x maior**	Acelera processamento distribuído (Hadoop, Spark) e replicação de storage.
Consumo Energético	3.5W/Gbps	0.6W/Gbps	**83% mais eficiente**	Reduz custos operacionais em larga escala.
Capacidade de Conexões Simultâneas	10-20 VMs/servidor	50-100 VMs/servidor	**5x maior**	Suporta alta densidade de virtualização e microsserviços.

Item 4 - 24 (vinte e quatro) **cabos de cobre** passivos, **5 metros** de comprimento, **com conexão direta (DAC) SFP28 25GbE para DataCenter**

Problema: Garlados de largura de banda, latência elevada, custos operacionais excessivos pelo uso de soluções ópticas com transceivers e fibras, falta de escalabilidade future-proofing

Necessidade: Cabos de cobre passivos, 5 metros de comprimento, com conexão direta (DAC) SFP28 25GbE para DataCenter

Resultado Esperado: Melhoria imediata no desempenho, redução no tempo de transferência, conectividade plug-and-play, redução de 30% no CAPEX, já que cabos DAC custam cerca de metade das demais soluções ópticas equivalentes, além da compatibilidade por meio do suporte SFP28 nativo e da preparação para redes 100GBPS.

2. DESCRITIVO DO OBJETO (art. 6º, inc. XXIII, alíneas "a" e "c" da Lei n.º 14.133/2021 c/c art. 17, inc. III do Decreto n.º 48.816/2023)

2.1. Descrição Resumida do Objeto (art. 17, inc. III, alínea "b" do Decreto n.º 48.816/2023)

Item 1 - **03 (três) GPU's 48GB GDDR6 (para servidores Lenovo ThinkSystem SR650 V2 Type 7Z73)**

Item 2 - **8 (oito) Módulos de 64GB** (sessenta e quatro gigabits) de **Memória TruRAM DDR4 3200Mhz,** totalizando 512 GB (quinhentos e doze gigabits), **para servidores Lenovo ThinkSystem SR650 V2 Type 7Z73 CT01WW**

Item 3 - 12 (doze) **placas de rede 25GbE para servidores Lenovo ThinkSystem SR650 V2 Type 7Z73 CT01WW**

Item 4 - 24 (vinte e quatro) **cabos de cobre** passivos, **5 metros** de comprimento, **com conexão direta (DAC) SFP28 25GbE para DataCenter**

2.2. Identificação dos itens, quantidades e unidades (art. 6º, inc. XXIII, alíneas "a" e "i", da Lei n.º 14.133/2021 c/c art. 17, inc. III, alínea "b" do Decreto n.º 48.816/2023)

Item	Código do objeto	ID	Descrição	Unidade de fornecimento	Quantidade
01		188974	**03 (três) GPU's 48GB GDDR6 (para servidores Lenovo ThinkSystem SR650 V2 Type 7Z73)**	unidade	03
02		188772	**Módulos de 64GB** (sessenta e quatro gigabits) de **Memória TruRAM DDR4 3200Mhz,** totalizando 512 GB (quinhentos e doze gigabits), **para servidores Lenovo ThinkSystem SR650 V2 Type 7Z73 CT01WW**	unidade	08
03		192270	**12 (doze) placas de rede 25GbE para servidores Lenovo ThinkSystem SR650 V2 Type 7Z73 CT01WW**	unidade	12
04		192268	24 (vinte e quatro) **cabos de cobre** passivos, **5 metros** de comprimento, **com conexão direta (DAC) SFP28 25GbE para DataCenter**	unidade	24

2.3. Especificações detalhadas do objeto

2.3.1 ***Detalhamento do descritivo do item 1***

a. Memória: 48GB RAM GDDR6 com ECC
b. Interface: PCIe
c. Segmento: Para DataCenter
d. Compatibilidade: Servidores Lenovo ThinkSystem SR650 V2 Type 7Z73-CT01WW
e. Emprego: Inteligência Artificial, Ciência de Dados, Inferência, Machine Learning, Deep Learning, Big Data, redes neurais, predição.
f. Desempenho do Tensor FP32 : mínimo 80 TFLOPS por placa
g. Saídas de monitor: Mínimo 2 saídas HDMI ou DisplayPort

h. Largura de banda: Mínimo 800 GB/s
i. Embalagem original do fabricante, com garantia mínima de 01 ano

2.3.2 ***Detalhamento do descritivo do item 2***

a. Módulos de 64Gb de memória RAM
b. Memórias TruDDR4
c. Compatíveis com servidores de aplicação Lenovo Thinksystem SR650 V2 Type 7Z73-CT01WW
d. Velocidade 3200Mhz
e. Embalagem original do fabricante, com garantia mínima de 01 ano

2.3.3 ***Detalhamento do descritivo do item 3***

a. Duas Portas 10/25 Gigabit Ethernet
b. Conexão PCI-E
c. Compatíveis com servidores de aplicação Lenovo Thinksystem SR650 V2 Type 7Z73-CT01WW
d. Totalmente compatível com o padrão SFF-8402
e. Compatível com x8 PCI Express v3.0
f. Mecanismo de processamento de fluxo TruFlow
g. Descarregamentos de IPv4 e IPv6
h. Interface de mídia plugável de porta dupla, que pode ser equipado com 25G SFP28 ou 10G SFP + óptico transceptor ou com cabo de conexão direta de cobre.
i. Embalagem original do fabricante, com garantia mínima de 01 ano

2.3.4 ***Detalhamento do descritivo do item 4***

a. 5 metros
b. Formato SFP28 - SFP28 de conexão automática
c. Cabos de cobre passivos para conexão direta 25GbE
d. Compatíveis com servidores de aplicação Lenovo Thinksystem SR650 V2 Type 7Z73-CT01WW
e. Conectores macho
f. Embalagem original do fabricante, com garantia mínima de 01 ano

3 . MODELAGEM / DESENHO DA CONTRATAÇÃO (art. 6º, inc. XXIII, alíneas “a”, “c” e “h”, da Lei n.º 14.133/21 c/c art. 17 do Decreto n.º 48.816/2023)

3.1. Regime de contratação

Está contratação será regida pela Lei n.º 14.133/2021 e suas regulamentações.

3.2. Forma e critério de seleção do fornecedor (arts. 23, 28, 33, 58, 74, 75 da Lei n.º 14.133/2021 c/c art. 17, inc. VI do Decreto n.º 48.816/2023)

Sugere-se, s.m.j., a contratação dos objetos ora pretendidos por meio de **PREGÃO ELETRÔNICO**, com adoção do critério de julgamento pelo **MENOR PREÇO**,de acordo com o art. 6º, XLI e art. 28, I, ambos da Lei 14.133/2021. O modo de disputa mais adequado é o modo aberto, tendo em vista que o objeto a ser licitado não é de alta complexidade, não tendo um mercado restritivo e o modo aberto traria maior disputa e economicidade à contratação. O orçamento estimado da licitação deverá ser divulgado no edital, quando da publicação do aviso. Portanto, acredita-se que tais parâmetros, em conjunto, são capazes de garantir uma combinação adequada e eficiente para a obtenção da proposta mais vantajosa à Administração Pública, considerado todo o ciclo de vida do objeto.

3.3.Regime de execução para serviço (art. 6º, incs. XXVII a XXXIV da Lei n.º 14.133/2021 c/c art.17. inc. IV do Decreto n.º 48.816/2023)

Fornecimento, por meio da entrega direta, aos cuidados da equipe de fiscalização do contrato, dos itens, quantidades e unidades descritos no tópico 3.2 deste ETP, exaurindo-se a contratação com a entrega e aceite definitivo.

3.4.Forma de execução do contrato (art. 6º, inc. XXIII, alínea "e"; art. 40, § 1º Lei n.º 14.133/2021 c/c art. 17, inc. IV do Decreto n.º 48.816/2023)

O bem será considerado adquirido no ato de conclusão do Pregão Eletrônico, devendo a entrega ser realizada em até 30 dias do encerramento do citada etapa licitatória, sendo a demanda dada por encerrada somente após Recebimento Definitivo, atestado pela equipe de fiscalização contratual.

3.5. Amostras/Exame de conformidade/Prova de conceito/Prova de qualidade similar/Laudo/certificação (art. 6º, inc. XXIII, alínea "d"; art. 17, § 3º; art. 41, inc. II; art. 42; art. 47, § 2º, da Lei n.º 14.133/2021)

Não se aplica.

3.6. Garantia da proposta (art. 58 da Lei n.º 14.133/2021 c/c art. 24 do Decreto n.º 48.778/2023 c/c art. 47, inc. XIII do Decreto n.º 48.816/2023)

Não se aplica, por tratar-se da aquisição de bens comuns, cujos padrões de desempenho e qualidade podem ser objetivamente definidos pelo edital, por meio de especificações usuais do mercado.

3.7. Âmbito da licitação

Esta licitação será nacional.

3.8. Habilitação Técnica (art. 67 da Lei n.º 14.133/2021 c/c art. 17, inc. V, alínea "d" do Decreto n.º 48.816/2023)

Não se aplica, por tratar-se da aquisição de bens comuns, cujos padrões de desempenho e qualidade podem ser objetivamente definidos pelo edital, por meio de especificações usuais do mercado.

3.8.1.Visita Técnica

Não se aplica

4. INFORMAÇÕES CONTRATUAIS (art. 6º, inc. XXIII, alíneas "a" e "c", da Lei n.º 14.133/2021)

4.1. Vigência do contrato

A contratação se encerrará com a entrega e aceite definitivo exarado pela equipe de fiscalização, devendo, contudo, honrar a garantia contratual

4.2. Reajuste de Preços (art. 6º, incs. LVIII e LIX; art. 25, § 7º; art. 92, § 3º; art. 135; art. 136 da Lei n.º 14.133/2021)

Não se aplica, por tatar-se de contratação que se exaure com a pronta entrega e aceite definitivo do bem

4.3. Garantia contratual (art. 96 da Lei n.º 14.133/2021 c/c art. 17, inc. V, alínea "a" do Decreto n.º 48.816/2023)

A garantia dos bens adquiridos deverá ser de, no mínimo, 12 meses de vigência, a contar da data da entrega dos bens à equipe de fiscalização.

4.4. Transferência de conhecimento, tecnologia e técnicas empregadas e Transição Contratual (art. 7º, Parágrafo Único, inc. II do Decreto n.º 48.816/2023)

Não se aplica, por tratar-se da aquisição de bens comuns, cujos padrões de desempenho e qualidade podem ser objetivamente definidos pelo edital, por meio de especificações usuais do mercado.

5. Práticas de sustentabilidade (art. 18, inc. XII do § 1º, da Lei n.º 14.133/2021 c/c art. 8º do Decreto n.º 48.816/2023)

Não se aplica, por tratar-se da aquisição de bens comuns, cujos padrões de desempenho e qualidade podem ser objetivamente definidos pelo edital, por meio de especificações usuais do mercado.

5.1. Possibilidade de subcontratação (art. 122, § 2º da Lei n.º 14.133/2021 c/c art. 17, inc. VI, alínea "c" do Decreto n.º 48.816/2023)

Não se aplica, por tratar-se da aquisição de bens comuns, cujos padrões de desempenho e qualidade podem ser objetivamente definidos pelo edital, por meio de especificações usuais do mercado.

5.2.Possibilidade de participação de Consórcio (art. 17, inc. VI, alínea "b" do Decreto n.º 48.816/2023)

Não se aplica, por tratar-se da aquisição de bens comuns, cujos padrões de desempenho e qualidade podem ser objetivamente definidos pelo edital, por meio de especificações usuais do mercado, além de não ser compatível com o objeto em apreço.

5.3. Possibilidade de participação de Cooperativa (art. 16, inc. IV da Lei n.º 14.133/2021)

Não se aplica, por tratar-se da aquisição de bens comuns, cujos padrões de desempenho e qualidade podem ser objetivamente definidos pelo edital, por meio de especificações usuais do mercado, além de não ser compatível com o objeto em apreço.

5.4. Programa de Integridade (art. 1º da Lei Estadual n.º 7.753/2017 c/c art. 25, § 4º, da Lei n.º 14.133/2021)

Não será exigido Programa de Integridade da empresa a ser contratada, tendo em vista que valor do contrato será menor que R$ 1.430.000,00 (um milhão quatrocentos e trinta mil reais) e com prazo acima de 180 dias, conforme estabelecido na Lei Estadual n.° 7.753/2017.

5.5. Obrigações das partes

5.5.1.Obrigações da contratante

O CONTRATANTE será responsável, por:

a) Acompanhar e fiscalizar a execução do contrato, bem como atestar nas faturas / notas fiscais a efetiva entrega do objeto deste Termo de Referência;

b) Efetuar os pagamentos à CONTRATADA em até 30 (trinta) dias após a emissão do aceite definitivo.

c) Realizar o cadastro para permitir a entrada dos funcionários da Contratada na sede do GSI

d) Aplicar à CONTRATADA as sanções regulamentares e contratuais.

5.5.2. Responsabilidades da contratada

A CONTRATADA será responsável por:

a) Apresentar a fatura / nota, para ateste e conferência do Fiscal do Contrato;

b) Responder por danos causados pelos seus agentes quando da execução do instrumento de contrato;

c) Permitir o acompanhamento na realização do serviço contratado por servidor público indicado pelo CONTRATANTE como Fiscal de Contrato ou outro profissional ou empresa que a CONTRATADA julgar pertinente;

d) Instruir seus empregados e contratados a tratar os funcionários do Gabinete de Segurança Institucional - GSI com urbanidade e respeito;

e) Pagar todos os tributos, contribuições fiscais e para-fiscais que incidam ou venham a incidir, direta e indiretamente, sobre os produtos vendidos, bem como eventual custo de frete na entrega;

f) Fornecer todos os esclarecimentos e as informações técnicas que venham a ser solicitadas pelo CONTRATANTE sobre os bens fornecidos;

g) Manter um representante, durante toda a execução do contrato, que deverá se reportar diretamente ao Fiscal do Contrato, para ser responsabilizar pela entrega dos bens adjudicados;

h) Comunicar imediatamente ao CONTRATANTE qualquer alteração ocorrida no endereço, conta bancária e outros julgáveis necessário para recebimento de correspondências;

6. REQUISITOS PARA A EXECUÇÃO CONTRATUAL (Art. 6º, inc. XXIII, alínea "d", da Lei n.º 14.133/2021)

6.1. Amostras / Exame de conformidade / Prova de conceito / Prova de qualidade similar / Laudo / certificação (Art.17, § 3º; art. 41, inc. II; art. 42 da Lei n.º 14.133/2021)

Não se aplica, por tratar-se da aquisição de bens comuns, cujos padrões de desempenho e qualidade podem ser objetivamente definidos pelo edital, por meio de especificações usuais do mercado, além de não ser compatível com o objeto em apreço.

6.2. Garantia da proposta (Art. 58 da Lei n.º 14.133/2021 c/c art. 24 do Decreto n.º 48.778/2023)

Não será exigida, haja vista o diminuto valor da contratação, além de tratar-se de entrega de bem, exaurindo-se em ato único o contrato.

6.3. Habilitação Técnica (Art. 67 da Lei n.º 14.133/21 c/c art. 17, inc. V, alínea "d" do Decreto n.º 48.816/2023)

Não se aplica, por tratar-se da aquisição de bens comuns, cujos padrões de desempenho e qualidade podem ser objetivamente definidos pelo edital, por meio de especificações usuais do mercado

6.4. Habilitação Econômico-financeira (Art. 69 da Lei n.º 14.133/21 c/c art. 17, inc. V, alínea "d" do Decreto n.º 48.816/2023)

A Empresa deverá apresentar certidões negativas de falências e recuperação judicial e extrajudicial expedidas pelos distribuidores da sede da pessoa jurídica, ou de execução patrimonial, expedida no domicílio da pessoa física. Sendo a mesma localizada fora da Comarca da Capital do Estado do Rio de Janeiro, as certidões deverão vir acompanhadas de declaração oficial da autoridade judiciária competente, relacionando os distribuidores que, na Comarca de sua sede, tenham atribuição para expedir certidões negativas de falências e recuperação judicial, ou de execução patrimonial.

6.5. Condições e restrições institucionais e legais (Art. 18, inc. III do § 1º da Lei n.º 14.133/2021 c/c art. 17, inc. V, alínea "b" do Decreto n.º 48.816/2023)

Não se aplica, por tratar-se da aquisição de bens comuns, cujos padrões de desempenho e qualidade podem ser objetivamente definidos pelo edital, por meio de especificações usuais do mercado

6.6. Licenças Necessárias (Art. 17, inc. V, alínea "d" do Decreto n.º 48.816/2023)

Todas as autorizações e licenças referidas deverão ser mantidas durante todo o prazo da contratação, cabendo às empresas contratadas as renovações, substituições e demais providências relacionadas à sua atuação regular, competindo ao órgão contratante a sua adequada fiscalização.

6.7. Fornecimento de Material

O fornecimento de qualquer meio, recurso ou material necessário ao integral adimplemento desta contratação ficará a cargo da empresa Contratada.

7. MODELO DE GESTÃO E FISCALIZAÇÃO DO CONTRATO (Art. 6º, inc. XXIII, alíneas "f" e "g" da Lei n.º 14.133/2021 c/c art. 17, inc. V, alínea "c" do Decreto n.º 48.816/2023)

7.1. Servidores indicados para gestão e fiscalização do contrato

Antônio Carlos Rabelo Ensa Junior	GSI/DIVMIT 4326740-8	(21) 2334-3328	divmit@subsecmilitar.rj.gov.br
Reinaldo de Carvalho Andrade	GSI/DIVMIT 4208187-4	(21) 2334-3328	divmit@subsecmilitar.rj.gov.br

7.2.Formas de comunicação entre as partes

Na fiscalização, o Fiscal do Contrato deverá fazer anotações e registros de todas as ocorrências, determinando o que for necessário à regularização das falhas ou defeitos observados.

Toda comunicação será formalizada por e-mail ou SEI, entre o preposto da contratada e a comissão fiscal do contrato

7.3. Recebimento provisório e definitivo (Art. 40, § 1º, inc. II; art. 92; art. 113; art. 140 da Lei n.º 14.133/2021 c/c art. 17, inc. IV, alínea "c" do Decreto n.º 48.816/2023);

Na forma e nos prazos constantes do Cronograma de execução.

7.4. Cronograma de execução (Art. 18, inc. VII do § 1º, da Lei n.º 14.133/2021 c/c art. 17, inc. IV, alínea "b" do Decreto n.º 48.816/2023)

Data-limite	Atividades	Responsável
30 dias da assinatura do contrato	Entrega dos bens adjudicados	Antônio Carlos Rabelo Ensa Junior e Reinaldo de Carvalho Andrade
No ato entrega dos bens	Termo de Recebimento Provisório	Antônio Carlos Rabelo Ensa Junior e Reinaldo de Carvalho Andrade
7 dias do encerramento da etapa anterior	Termo de Recebimento Definitivo	Antônio Carlos Rabelo Ensa Junior e Reinaldo de Carvalho Andrade
30 dias do encerramento da etapa anterior	Pagamento de 100% do valor à contratada	DGAF

7.5. Pagamento (Art. 6º, inc. XXIII, alínea "g" c/c art. 17, inc. IV, alínea "f" do Decreto n.º 48.816/2023)

À vista, no prazo constante do Cronograma de execução.

8. ASSINATURA DOS RESPONSÁVEIS PELA ELABORAÇÃO

Antônio Carlos Rabelo Ensa Junior
Id. Func. 4326740-8
GSI/DIVMIT

Reinaldo de Carvalho Andrade
Id. Func. 4208187-4
GSI/DIVMIT

Rio de Janeiro, 15 abril de 2025

Documento assinado eletronicamente por **Claudio Gomes de Brito**, **Coordenador**, em 22/05/2025, às 16:36, conforme horário oficial de Brasília, com fundamento nos art. 28º e 29º do Decreto nº 48.209, de 19 de setembro de 2022 e no art. 4º do Decreto nº 48.013, de 04 de abril de 2022.

Documento assinado eletronicamente por **Miguel Angelo Duarte Ticom**, **Superintendente**, em 22/05/2025, às 16:41, conforme horário oficial de Brasília, com fundamento nos art. 28º e 29º do Decreto nº 48.209, de 19 de setembro de 2022 e no art. 4º do Decreto nº 48.013, de 04 de abril de 2022.

Documento assinado eletronicamente por **Antonio Carlos Rabelo Ensá Junior**, **Primeiro Tenente**, em 22/05/2025, às 17:14, conforme horário oficial de Brasília, com fundamento nos art. 28º e 29º do Decreto nº 48.209, de 19 de setembro de 2022 e no art. 4º do Decreto nº 48.013, de 04 de abril de 2022.

Documento assinado eletronicamente por **Reinaldo de Carvalho Andrade**, **Administrador de Redes**, em 22/05/2025, às 17:16, conforme horário oficial de Brasília, com fundamento nos art. 28º e 29º do Decreto nº 48.209, de 19 de setembro de 2022 e no art. 4º do Decreto nº 48.013, de 04 de abril de 2022.

A autenticidade deste documento pode ser conferida no site http://sei.rj.gov.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=6, informando o código verificador **100795267** e o código CRC **FE18B710**.

Rua Pinheiro Machado, S/Nº - Laranjeiras - Bairro Laranjeiras, Rio de Janeiro/RJ, CEP 22231-901
Telefone: