



AI-Ready Data Center

CIO GOV Brasil - 2026

Alessandro Knobloch

Sr. Solutions Architect



Sobre a NTT DATA

NTT DATA

NTT DATA Value Prop: Supporting stats

\$30B

Annual Revenue
(and no debt)

190K+

Employees

50+

Countries

33

Countries recognized
as top employer

Top 1%

of IT service
providers in S&P
Global Sustainability
Rating

2nd

Applications &
Infrastructure
globally

3rd

Largest Data Center
provider in the world

4th

IT & Managed
Services globally

4th

Largest IP backbone
globally

Top 10

Most valuable IT
services brand

NTT Group by the numbers

\$100B

Annual Revenue

330K+

Employees

150+

Years in business

\$3.6B

Annual investment in
R&D

4th

Largest telecom
company globally



Some of the reasons why 65% of the Fortune 500 and 75% of the Fortune 100 are our clients.

Credenciais e parcerias NTT DATA AI

Credenciais

17,000+

Data & AI practitioners ao redor do globo

7

Centro de inovações de dados e IA

6,000+

Cientistas de dados

1

Escritório global de GenAI

100%

Força de trabalho focada

50

Países

Elogios

Gartner nomeou NTT DATA como leader

Generative AI Consultoria e Implantação

HFS Horizons

Generative Enterprise™ Market Líder

World Intellectual Property Organization*

#12 como empresa detentora de patentes de IA globalmente

Alianças & parcerias

Parcerias globais com os principais fabricantes**:



430+ novos parceiros de GenAI visando otimizar nossas entregas fim a fim:



* World Intellectual Property Organization is the United Nations agency for innovation and creation

** Not an exhaustive list of partners and alliances

Líder no Gartner

A NTT DATA foi nomeada Emerging Leader in the Gartner® Innovation Guide for GenAI Consulting & Implementation Services.

Generative AI Consulting and Implementation Services





Desafios de adoção de IA

Desafios ao escalar IA

Por que as empresas precisam de uma nova abordagem para a entrega de IA?



Dívida tecnológica com
infraestrutura legada



Falta de **habilidades GenAI**
em escala



Estouro de custos,
planejamento e orçamento
ruins



Complexidade de **integração**
e compatibilidade

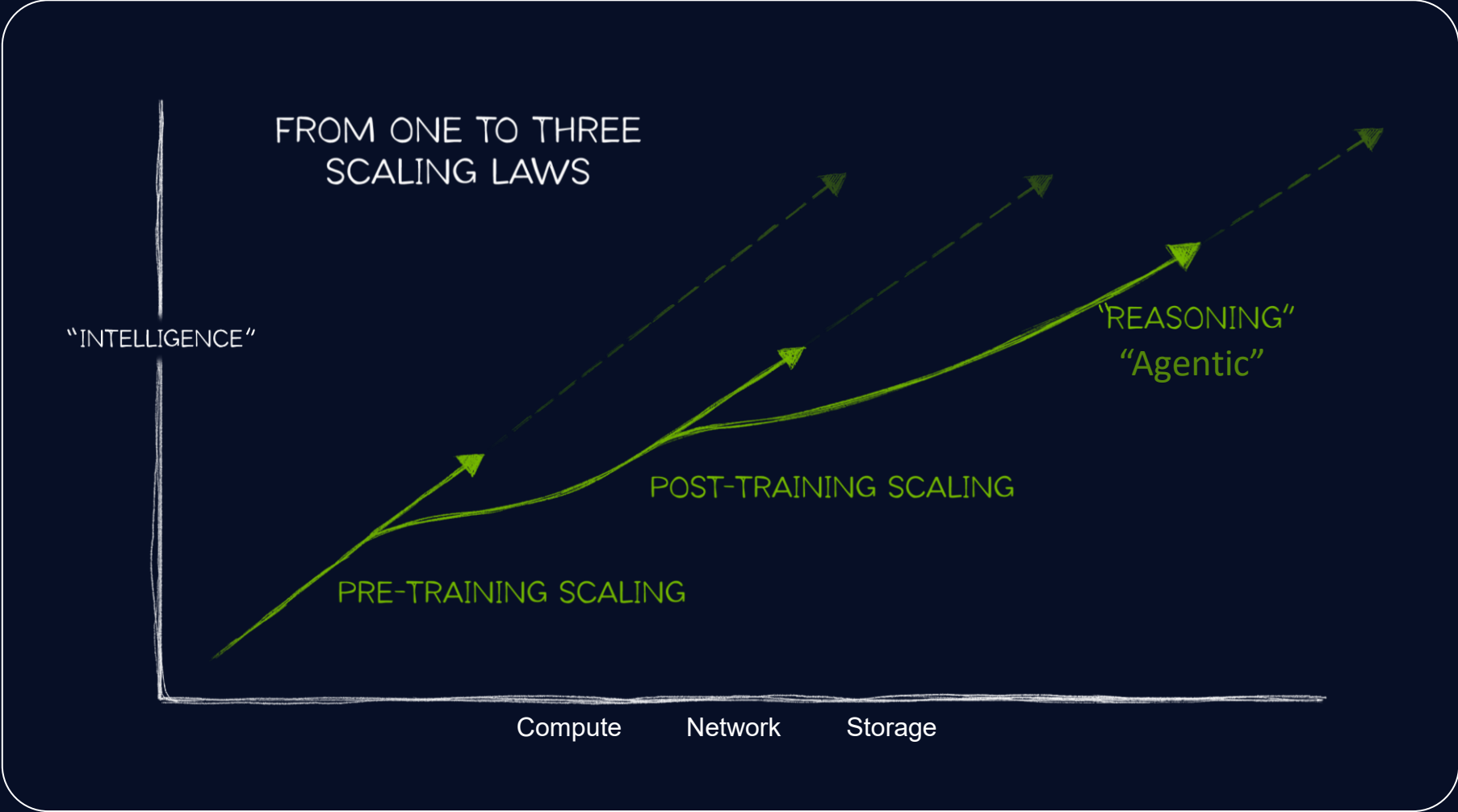


Governança, **segurança** e
riscos de conformidade



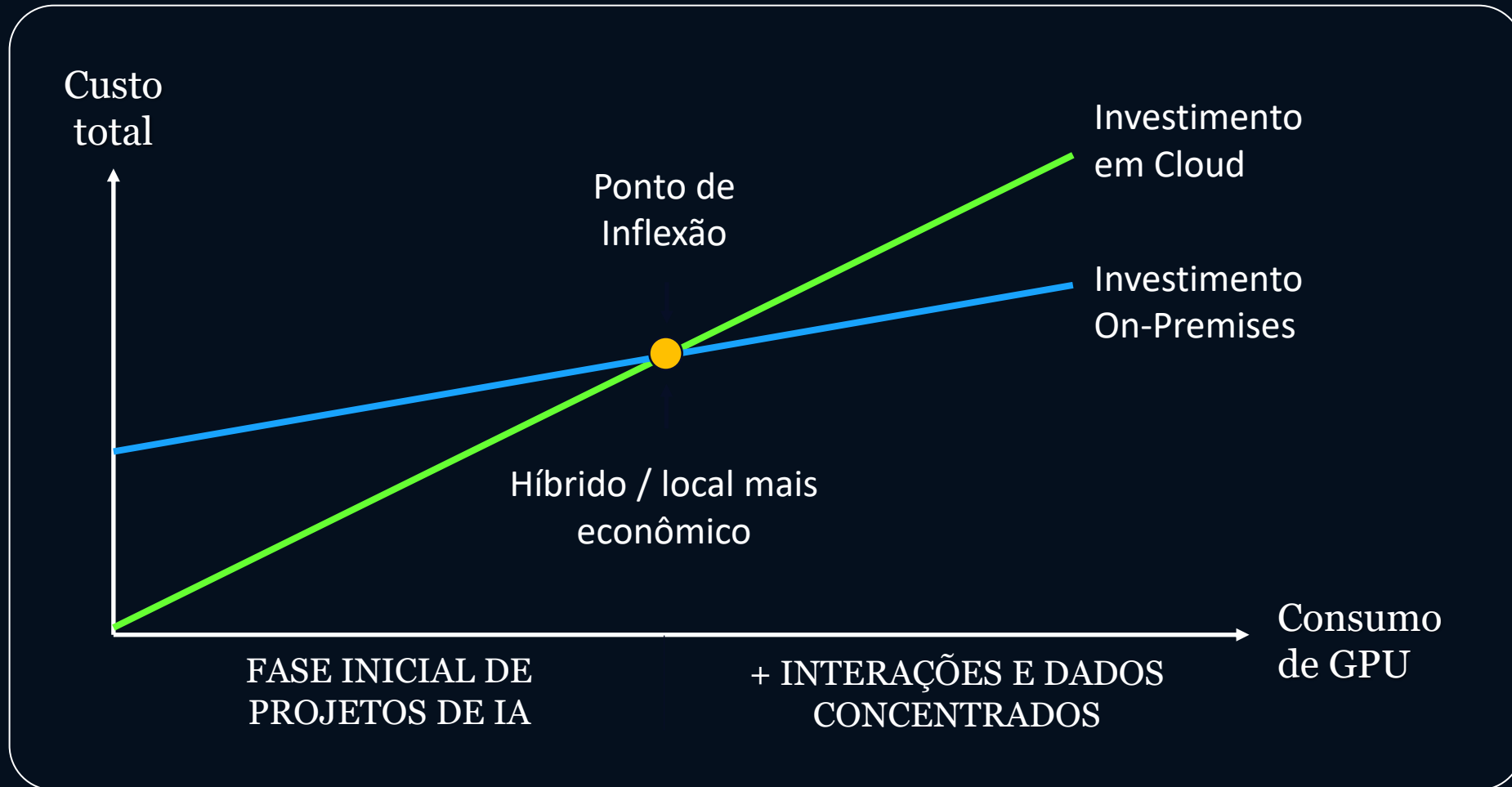
Falta de **gerenciamento**
de **dados** e evolução de
agentes além da POC

Raciocínio de modelos e evolução da arquitetura



ON-PREMISES + *CLOUD*

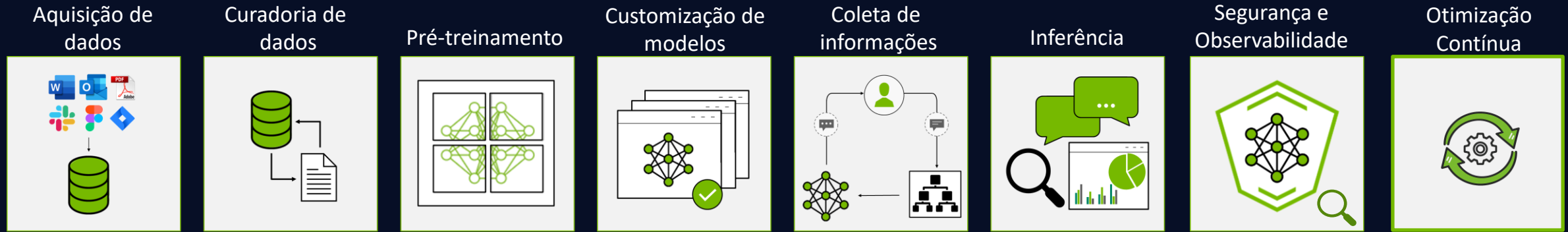
Qual você deve escolher?





Como **podemos ajudar?**

Onde seu negócio está nessa jornada?



Preparação dos dados

Treinamento e customização

Deployment/Day-2

Aumentar a qualidade dos dados

Focando no domínio de uso

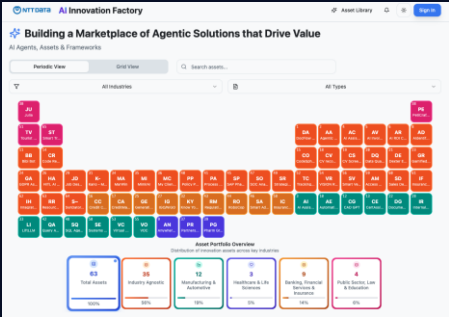
Implementação segura e em escala

*Jornada de IA
fim a fim*

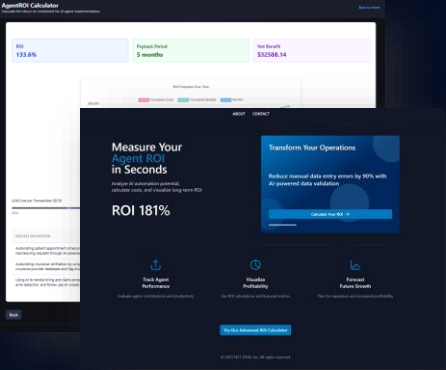
Aceleradores que otimizam sua experiência



AI Assessment e Jumpstart



”Tabela periódica” de Agentes / AI Marketplace

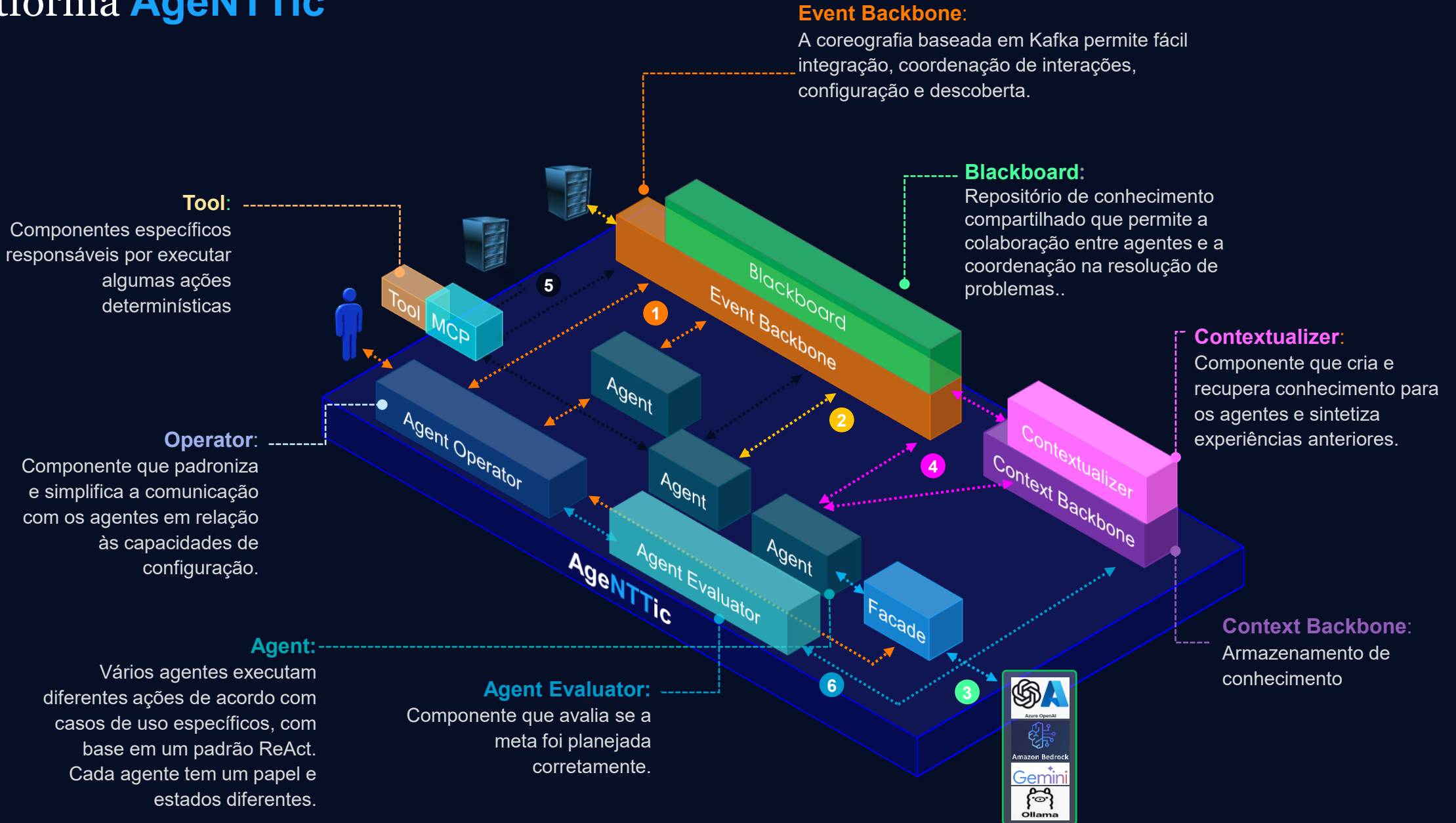


Calculadora de ROI



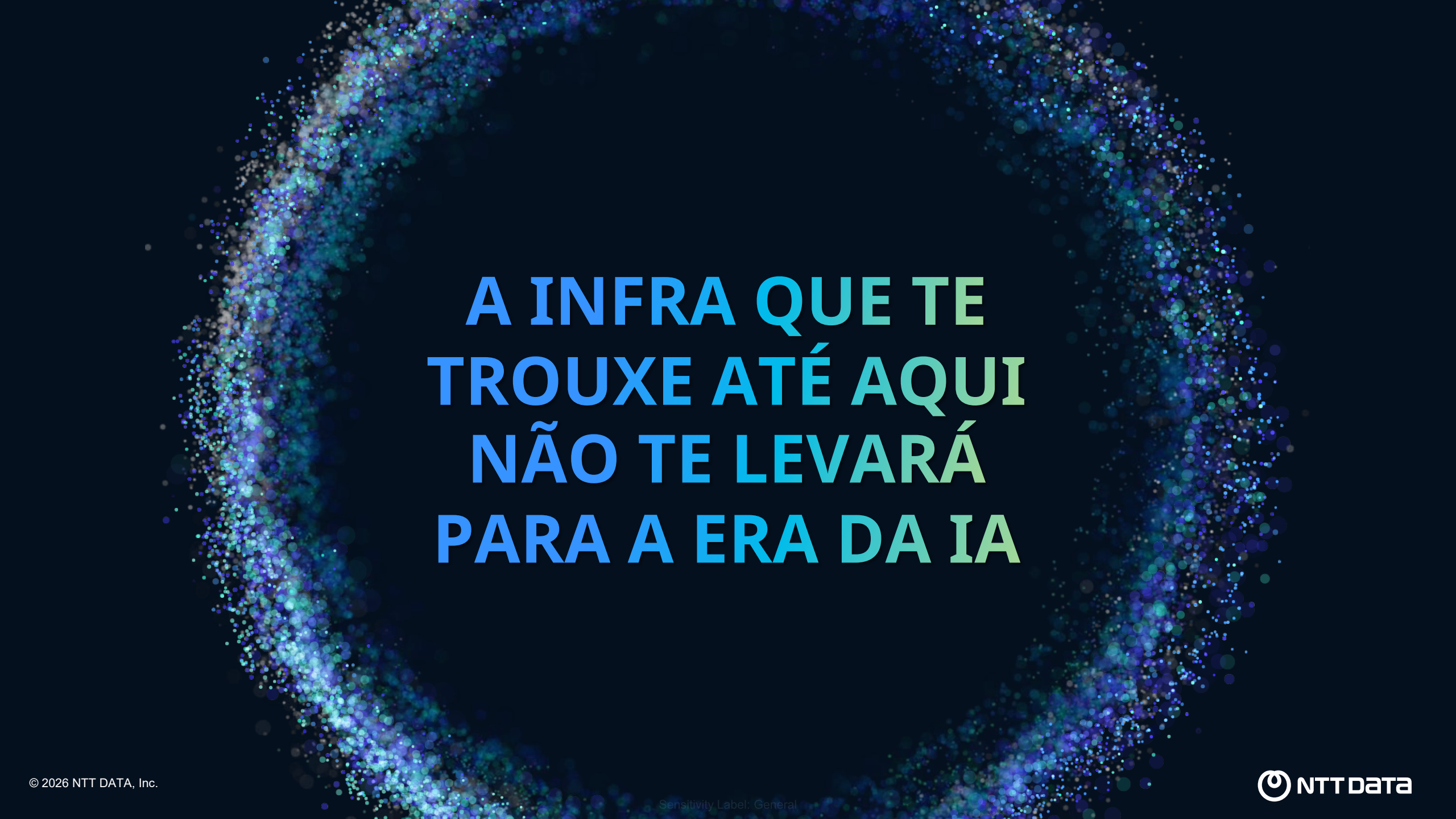
Plataforma AgeNTTic

Plataforma AgeNTTic



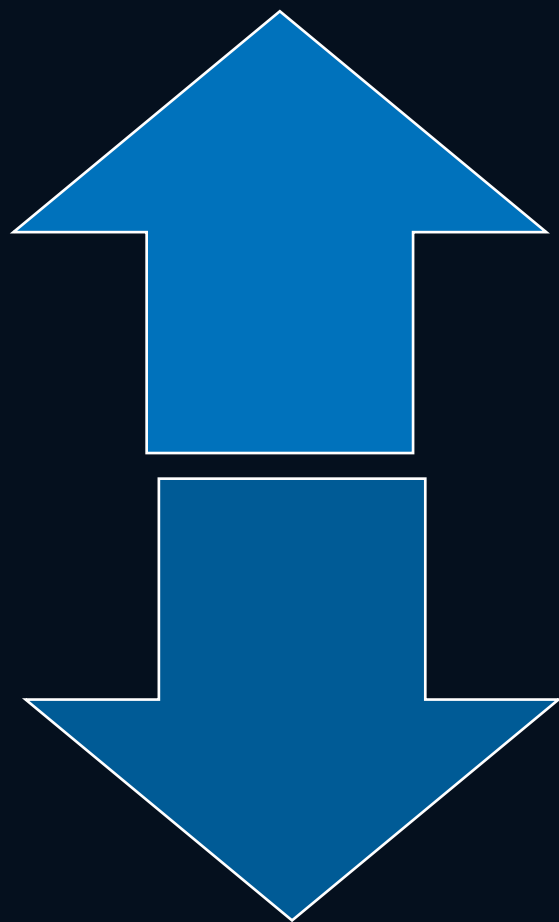


Data Center pronto para IA



**A INFRA QUE TE
TROUXE ATÉ AQUI
NÃO TE LEVARÁ
PARA A ERA DA IA**

Treinamento X Inferência

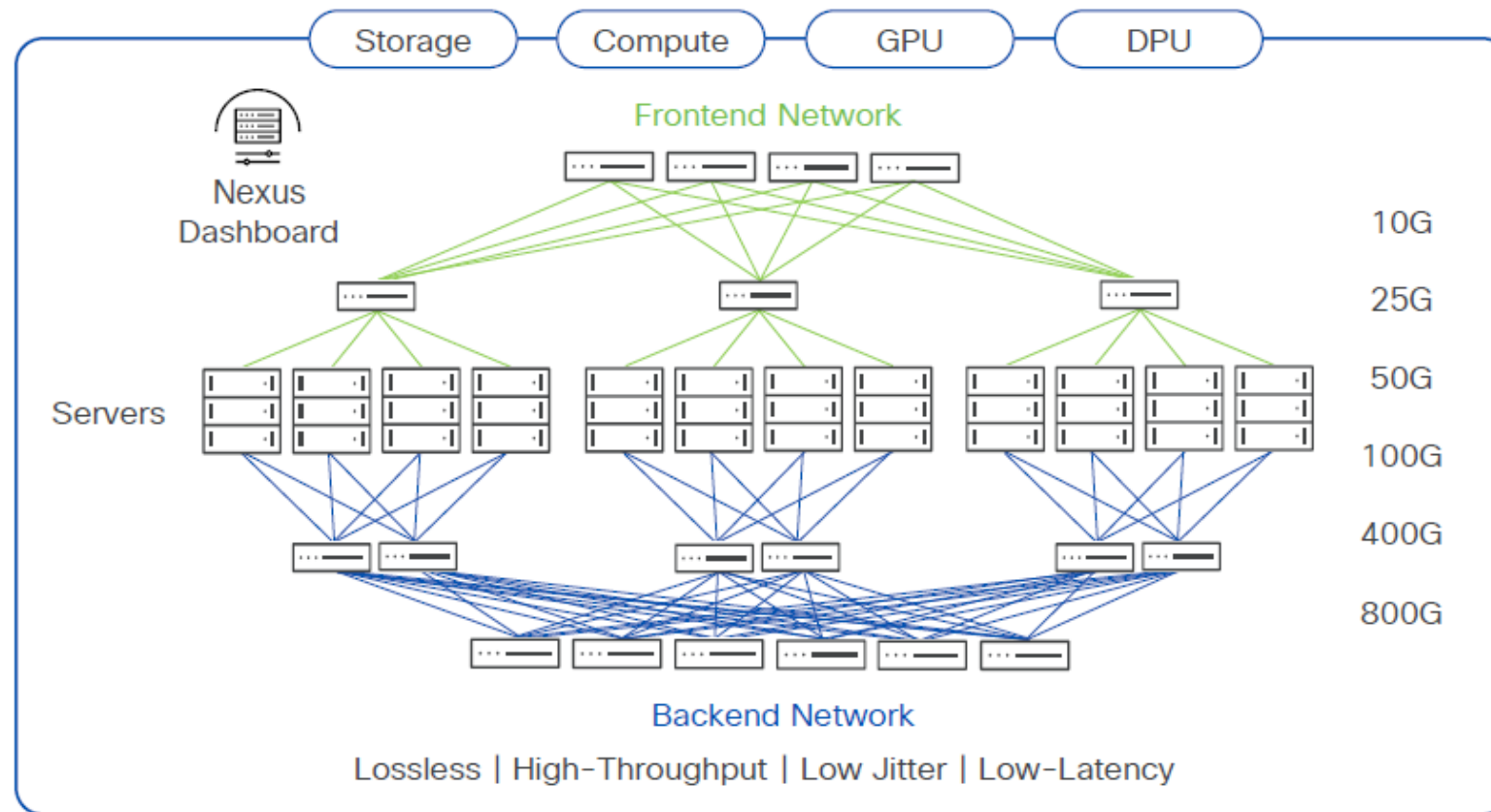


Treinar um modelo avançado de IA, como o LLaMA da Meta, exige milhares de GPUs funcionando ao mesmo tempo (por exemplo, 2.000 GPUs).


Para usar o modelo já treinado e gerar respostas (inferência), são necessárias muito menos GPUs — geralmente cerca de 16.

O impacto da IA nas redes

Treinamento e Fine Tuning




Data center


Colocation

O impacto da IA nas redes

200 usuários - Inferência

```
ubuntu@ai-11: ~  
Cisco UCSC-C240-M5SX computing node (hostname: ai-11)  
CPU: 2 x Intel(R) Xeon(R) Gold 6248R CPU @ 3.00GHz with 24 cores  
GPU: 6 x Tesla T4  
  
      Use      Memory Use  
CPU   6.17%    18Gi/1.5Ti  
  
GPU1  83%       13.2/15.0Gi  
GPU2  90%       12.4/15.0Gi  
GPU3  86%       12.4/15.0Gi  
GPU4  87%       12.4/15.0Gi  
GPU5  74%       12.4/15.0Gi  
GPU6  74%       12.4/15.0Gi  
  
NIC1 tx: 6.53 Gbps, rx: 6.91 Gbps (eno5)  
LLM: 728.05 tokens/s [API up]
```

```
ubuntu@ai-12: ~  
Cisco UCSC-C240-M5SX computing node (hostname: ai-12)  
CPU: 2 x Intel(R) Xeon(R) Gold 6248R CPU @ 3.00GHz with 24 cores  
GPU: 6 x Tesla T4  
  
      Use      Memory Use  
CPU   3.52%    6.1Gi/1.5Ti  
  
GPU1  76%       12.4/15.0Gi  
GPU2  76%       12.4/15.0Gi  
GPU3  0%        0.0/15.0Gi  
GPU4  0%        0.0/15.0Gi  
GPU5  0%        0.0/15.0Gi  
GPU6  0%        0.0/15.0Gi  
  
NIC1 tx: 6.75 Gbps, rx: 6.41 Gbps (eno5)  
LLM: 728.05 tokens/s [API up]
```

- 8B Llama 3
- 2 Servidores UCS com 8 GPUs
- 200 usuários
- **Consumo de rede ao redor de 6+Gbps**
- ~700 tokens/sec



AI Monitor


<https://github.com/pl247/ai-monitor>


Visibilidade de redes em tempo real




Cisco AI PODs


Cisco CVD, NVIDIA ERA



 Projeto validado pela Cisco





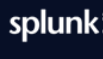
 Pós Venda Cisco
(produtos de terceiros)

 Tamanhos
P, M, G e Scale-Out

Extend To


Cisco® Security
AI Defense Hypershield
Firewall/Nexus® Smartswitch


Observability
 **Observability Cloud**
Open Telemetry Extensions



Workload management & ops
   

Treinamento






Otimização


Inferência


Cisco AI PODs





Operação **Automação** **AI software**



 **INTERSIGHT® & NEXUS DASHBOARD**     **NIM Operator NeMo CUDA**

Kubernetes  **Red Hat OpenShift**  **ubuntu RANCHER BY SUSE**

Computação Acelerada  **UCS®**

Redes de Alta Performance  **NEXUS**

Ecosistema estendido  **VAST**  **NetApp**  **PURE STORAGE**

 **NUTANIX**  **HITACHI**

Serviços Avançados

Cisco
CX Customer Experience

Como **NTT DATA e Cisco** podem ajudar?

Implantar de forma ágil com designs validados



Tomar decisões mais rápidas e com mais confiança



Evitar riscos e falhas



Simplificar e integrar ambientes complexos



Ajustar recursos conforme o crescimento do seu projeto








Case COP30

Garantimos a conectividade segura na área **Blue Zone da COP30** no Brasil

Para a NTT DATA, não foi só um evento

Era uma cidade conectada.

-  1.500 pontos de acesso
-  8.000 conexões físicas
-  180 switches em operação.

Até 50.000 pessoas
conectarias por dia.
Ao mesmo tempo.

A Blue Zone da COP30
operou como um
micro-pais por 15 dias.

Mais de 60.000 m²

O tamanho de um
aeroporto de medio porte.

Por trás de tudo isso,
uma infraestrutura
invisível.

E absolutamente essencial.

NTT DATA na espinha dorsal
da conectividade da COP30.



**Escaneie o QRCode e veja
o conteúdo completo.**

