A Transformação das Telecomunicações na Era da Inteligência Artificial

Ivana Lemos & Francisco Sant’Anna

Advisor, Portfolio and Solutions Marketing

A Transformação das Telecomunicações na Era da IA

1. Conceitos Fundamentais de IA

2. Impacto da IA nos Requisitos de Rede

3. AIOps: Operação de Redes com IA

Conceitos Fundamentais de IA

ARTIFICIAL INTELIGENCE

Applications:
- Computer Vision
- Speech recognition
- Natural language processing

By objective:
- Generative
- Discriminative
- Predictive
- Prescriptive

By functionality:
- Narrow AI
- General AI
- Super AI
- Conscious AI

By capability:
- Reactive
- Limited memory
- Theory of mind
- Self-awareness

Conceitos Fundamentais de IA

ARTIFICIAL INTELIGENCE

Applications:
- Computer Vision
- Speech recognition
- Natural language processing

By objective:
- Discriminative
- Predictive
- Prescriptive
- Generative

By functionality:
- Narrow AI
- General AI
- Super AI
- Conscious AI

By capability:
- Reactive
- Limited memory
- Theory of mind
- Self-awareness

Learning:
- Supervised
- Unsupervised
- Reinforced
- Continual
- Transfer
- Active

MACHINE LEARNING

Algorithms

Datasets

Predictors (Input)

Phases of model development:
- Data preparation
- Training
- Fine tuning
- Inference
- Evaluation

Conceitos Fundamentais de IA

ARTIFICIAL INTELIGENCE

Applications:
- Computer Vision
- Speech recognition
- Natural language processing

By objective:
- Discriminative
- Predictive
- Prescriptive
- Generative

By functionality:
- Narrow AI
- General AI
- Super AI
- Conscious AI

By capability:
- Reactive
- Limited memory
- Theory of mind
- Self-awareness

Learning:
- Supervised
- Unsupervised
- Reinforced
- Continual
- Transfer
- Active

MACHINE LEARNING

Phases of model development:
- Data preparation
- Training
- Fine tuning
- Inference
- Evaluation

NEURAL NETWORKS

- Convolutional (CNN)
- Recurrent NN (RNN)

Conceitos Fundamentais de IA

ARTIFICIAL INTELIGENCE

Applications:
- Computer Vision
- Speech recognition
- Natural language processing

By objective:
- Discriminative
- Predictive
- Prescriptive
- Generative

By functionality:
- Narrow AI
- General AI
- Super AI
- Conscious AI

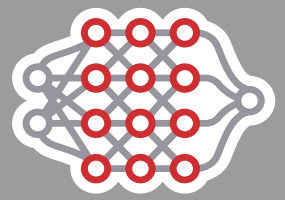
By capability:
- Reactive
- Limited memory
- Theory of mind
- Self-awareness

MACHINE LEARNING

Learning:
- Supervised
- Unsupervised
- Reinforced
- Continual
- Transfer
- Active

NEURAL NETWORKS

DEEP LEARNING



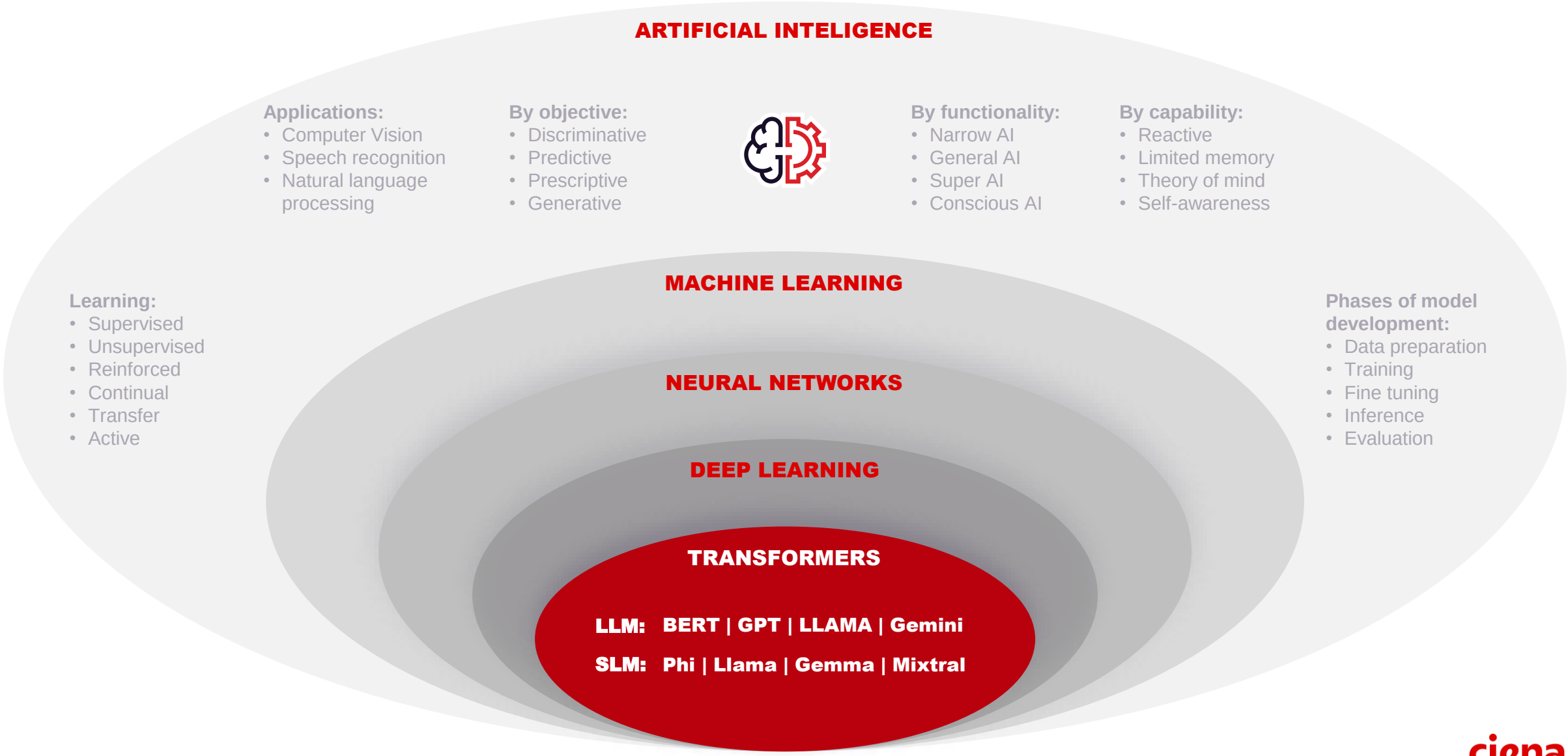
Phases of model development:
- Data preparation
- Training
- Fine tuning
- Inference
- Evaluation

Deep Learning Segundo Sam Altman

Deep learning worked, got predictably better with scale, and we dedicated increasing resources to it.

—Sam Altman

Conceitos Fundamentais de IA

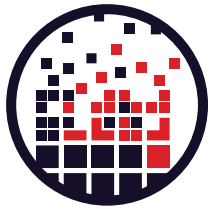
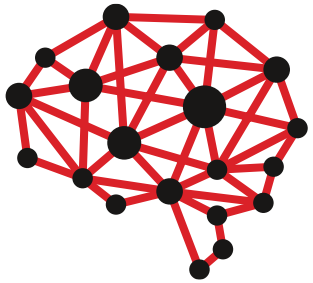
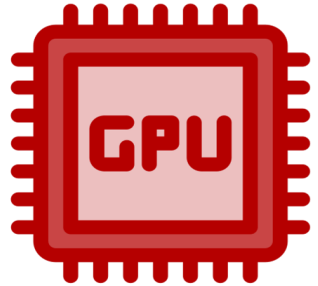


Credit: https://pixabay.com/photos/success-man-mountain-climb-top-5941539/

...Ops! Maybe I don’t know what I don’t know

Credit:
https://pixabay.com/photos/s%C3%A4ntis-switzerland-swiss-alps-2801963/

Construindo IAs Mais Inteligentes: Os 3 Pilares

Data	Algorithms	Compute
		

These inputs drive network demand and challenges!

O Desafio dos Dados

Novos modelos podem precisar de mais dados do que os humanos podem criar

Each new model requires more data

➔ Trillions of tokens

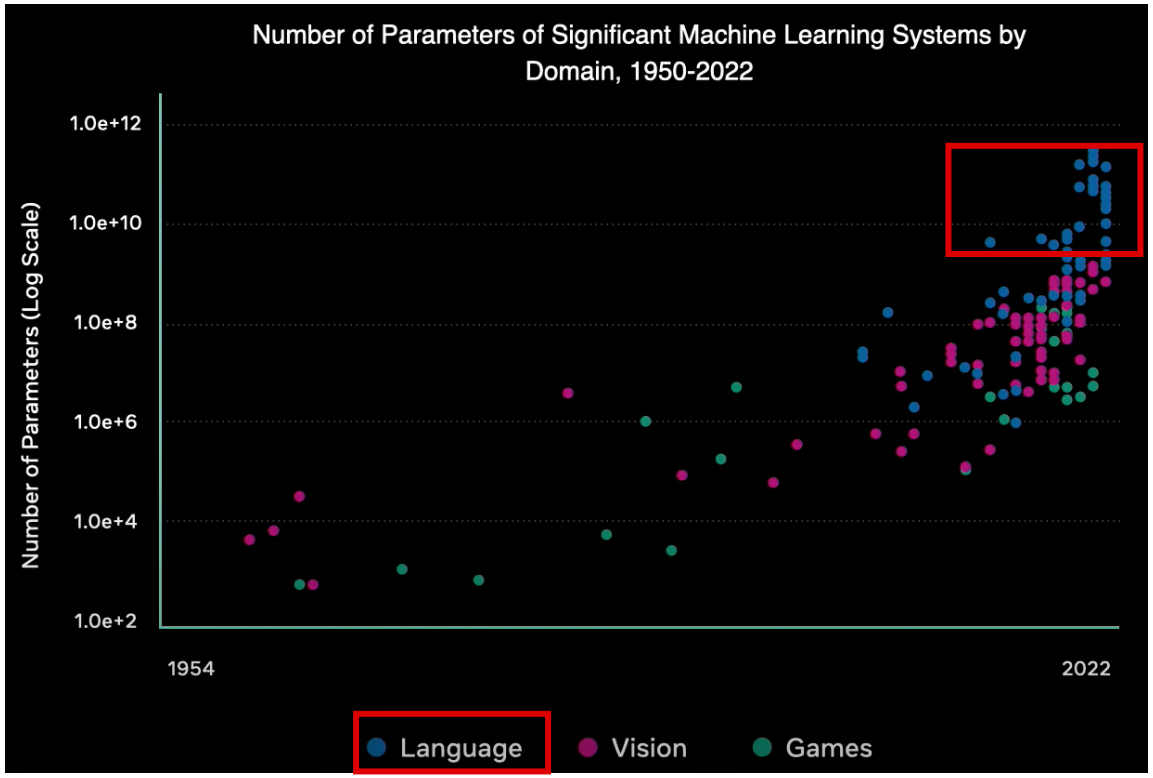
There is a growing recognition that data might become the bottleneck in next few years

➔ Synthetic data (~2028 timeframe)

The industry continues to work on how to get more quality insights from the data

➔ Algorithms

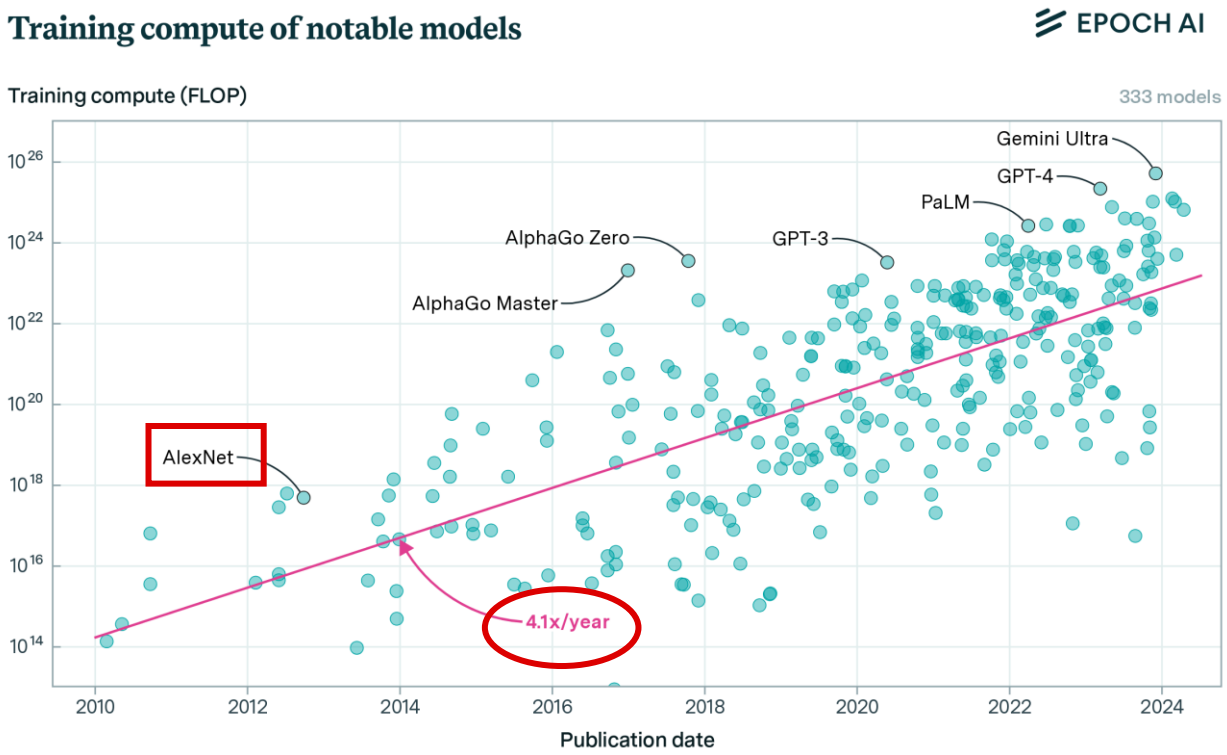
O Desafio dos Algoritmos



More data isn't always the answer – the industry is exploring different algorithms

Source: *Aparna Ramani discusses the future of AI infrastructure - Engineering at Meta*

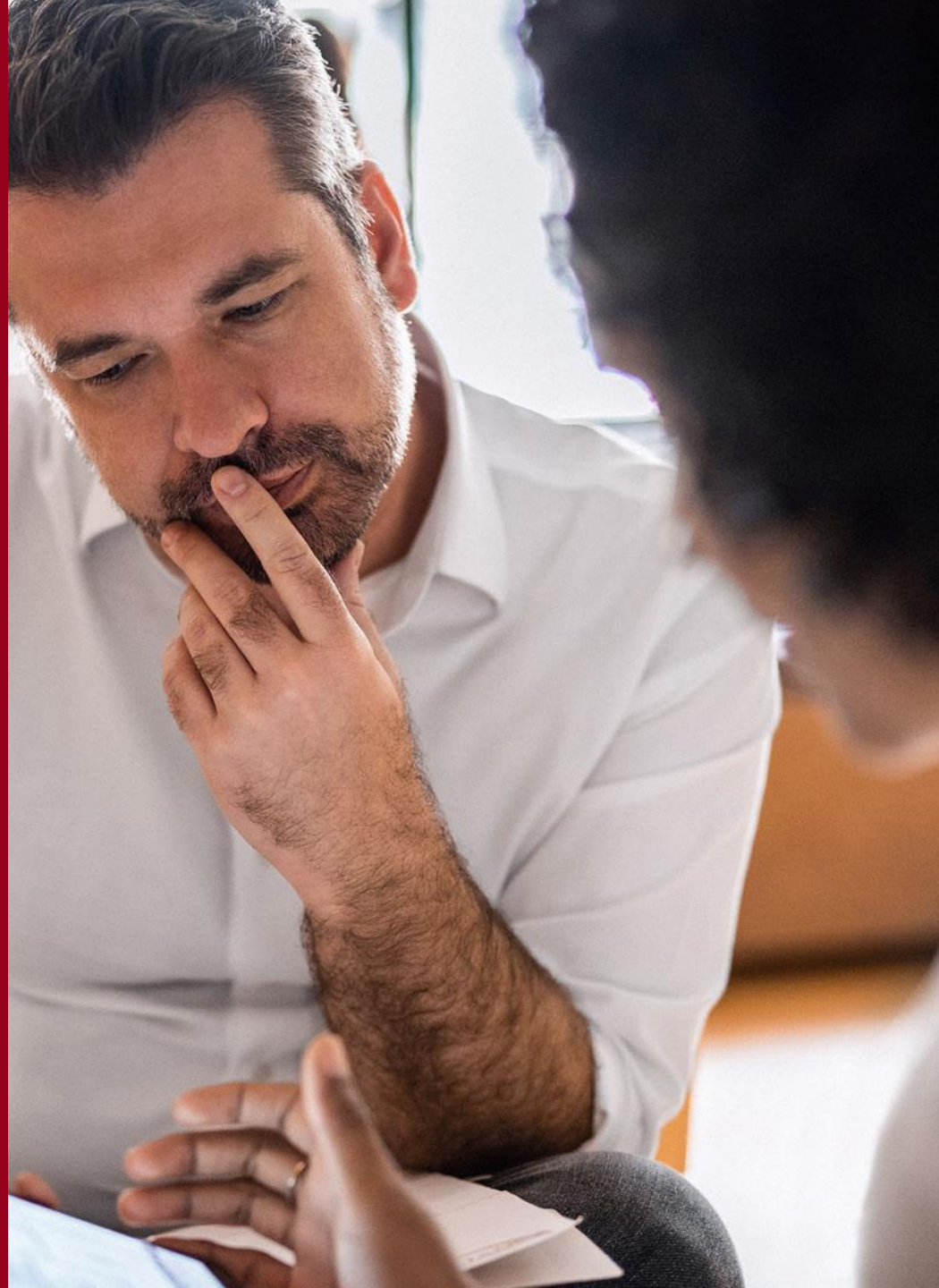
O Desafio do Poder Computacional



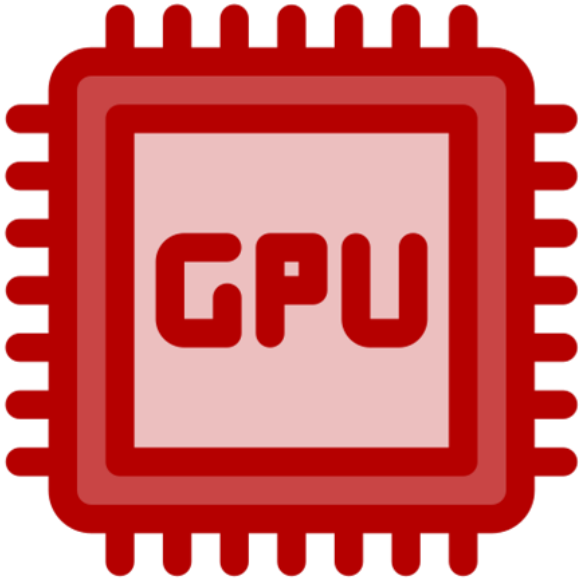
Conservative estimates put the growth of training at 4x a year for large models

Source: EPOCH AI, at https://epochai.org/blog/training-compute-of-frontier-ai-models-grows-by-4-5x-per-year

Impacto da IA nos Requisitos de Rede

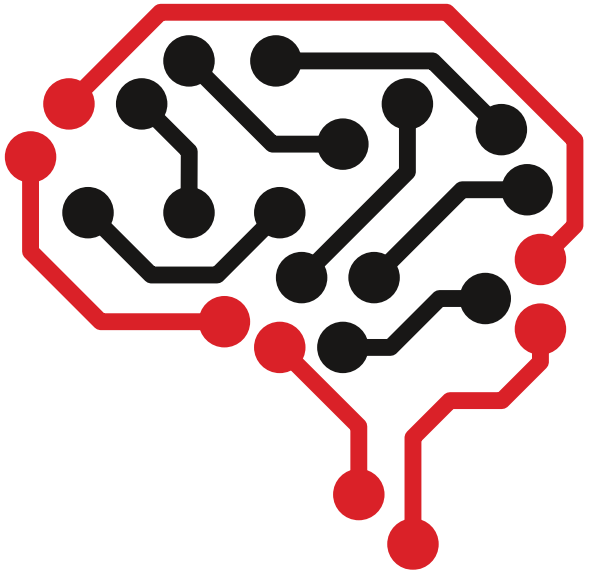


Consumo de Energia: GPU vs. Cérebro Humano



GPU

0.5 to 1 Kilowatt

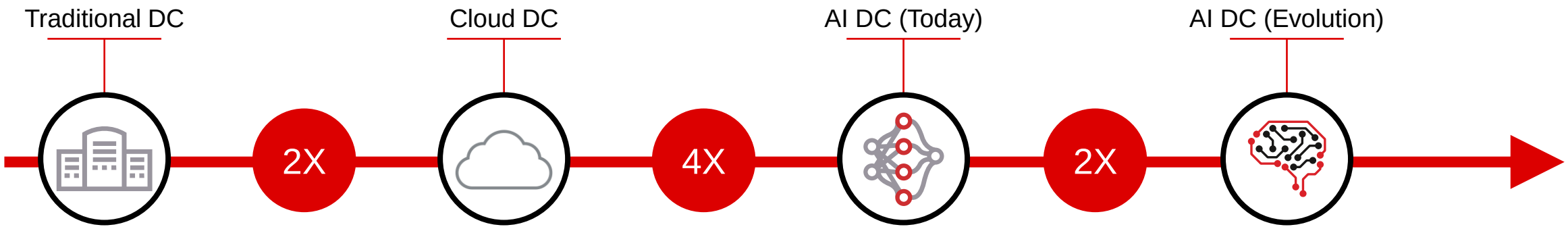


Human brain

25 Watts

(20 – 40x more energy-efficient)

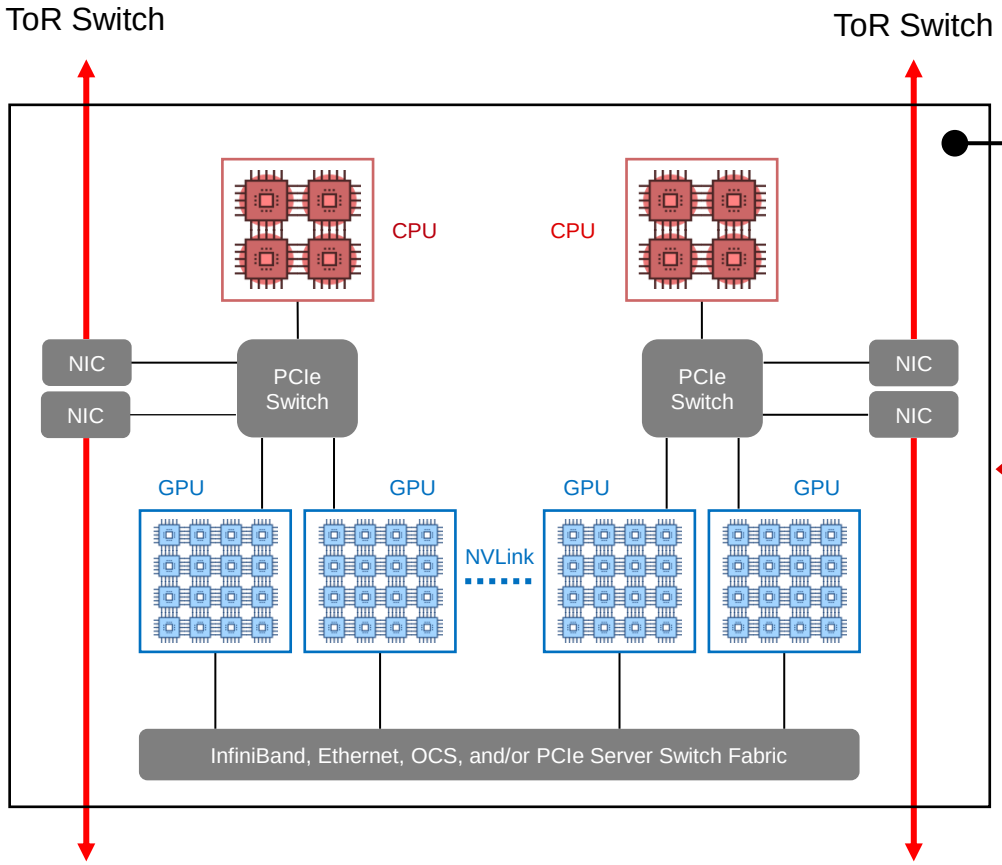
Energia: Possível Gargalo Para o Crescimento da IA



- 3-7 kW/rack
- 10-14 kW/rack
- 40-75 kW/rack
- 120 kW/rack

Dentro de um Rack: Os Componentes de um Servidor de IA Típico

Optical Ethernet
(Frontend Network: AI servers to Data Center Interconnect)



InfiniBand, OCS, and/or Ethernet (optical or copper)
(Backend Network: AI server cluster interconnect)

Optical network connections

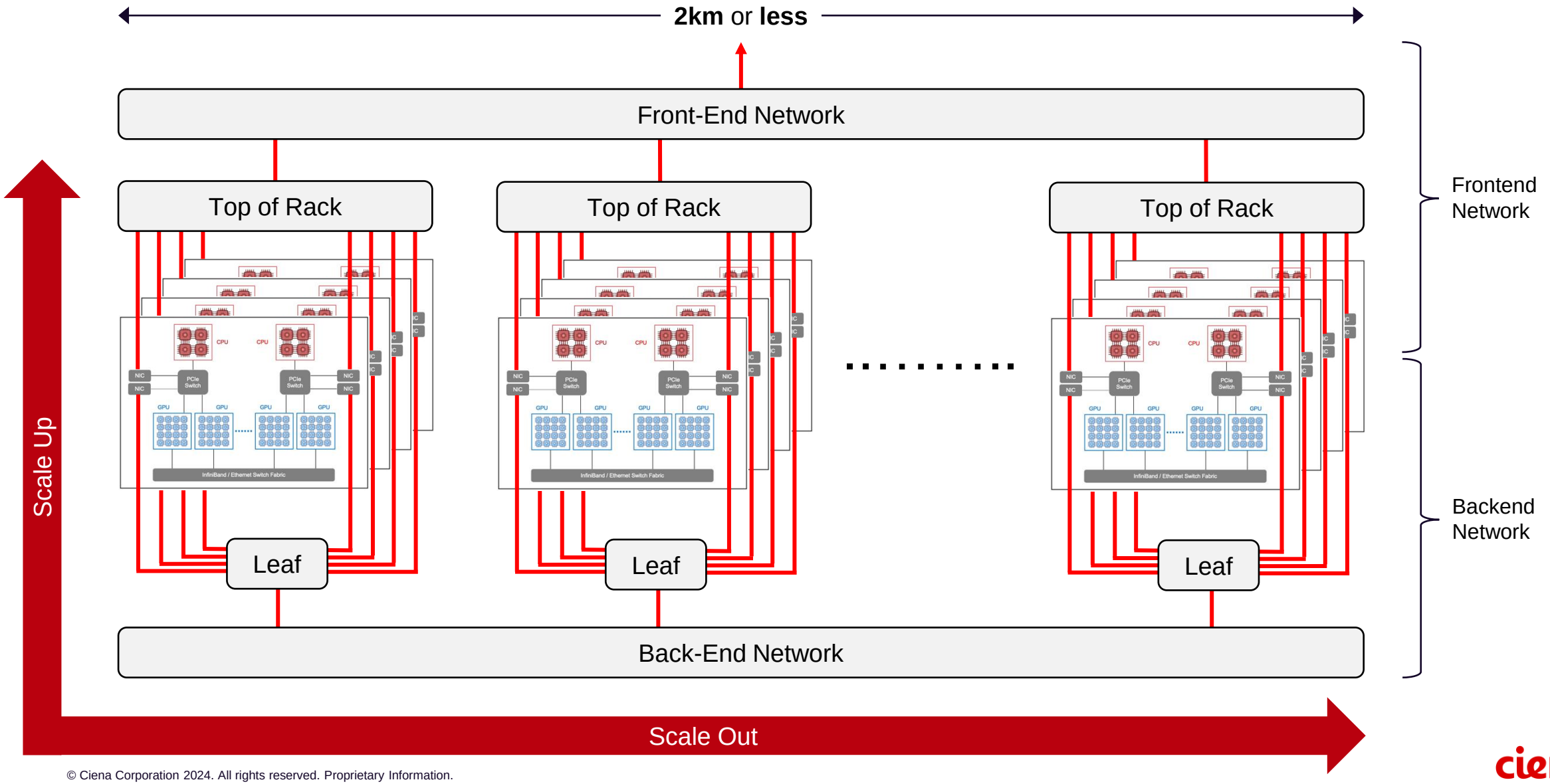
Intra-server networking uses mainly copper connections, as it meets performance and power consumption expectations

Customized and optimized building block of AI clusters

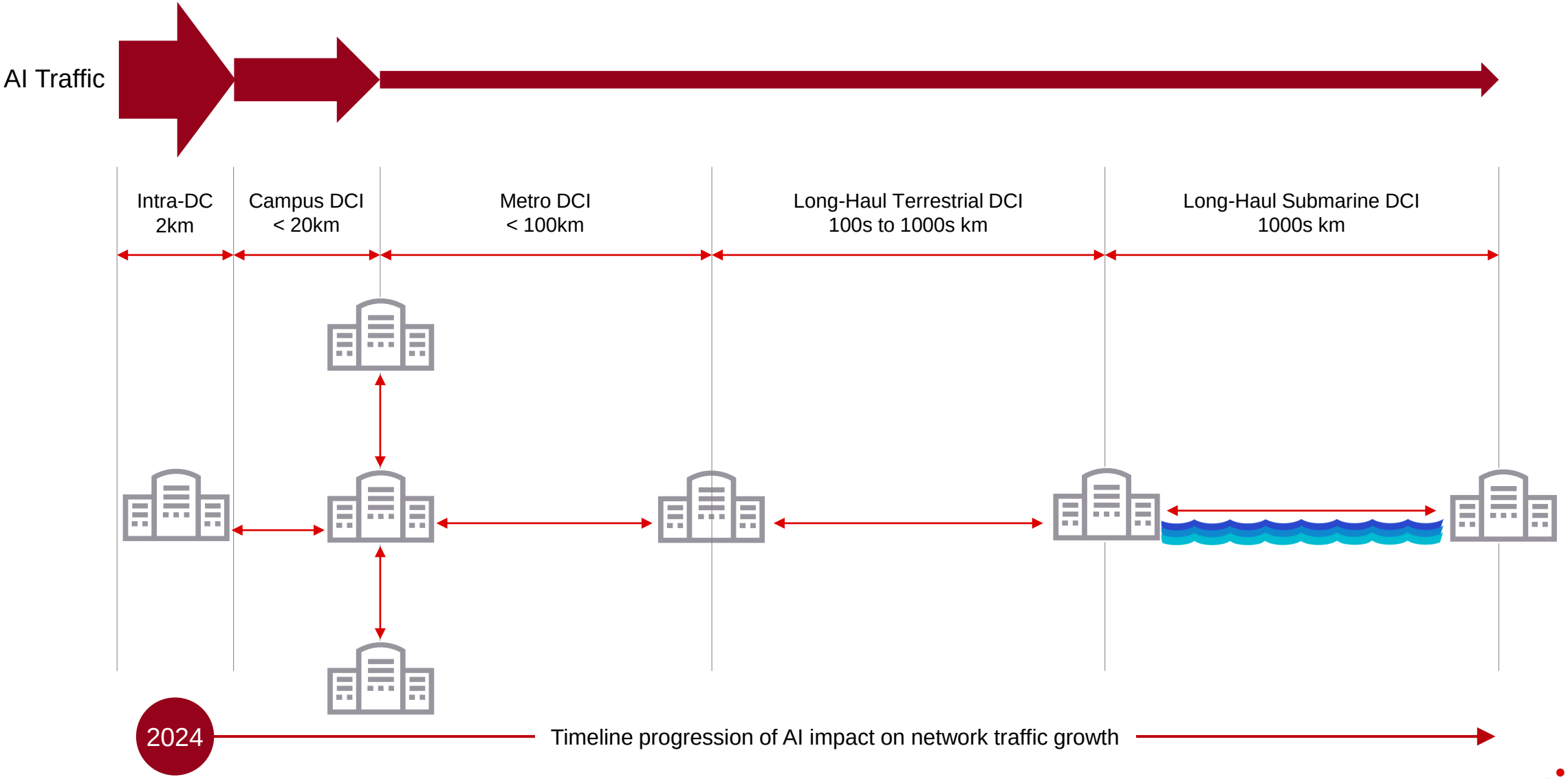
Inter-server networking uses copper and/or optical connections, based on reach and performance requirements of AI clusters

OCS (Optical Circuit Switch)

Dentro do Data Center: Infraestrutura de um Cluster de IA

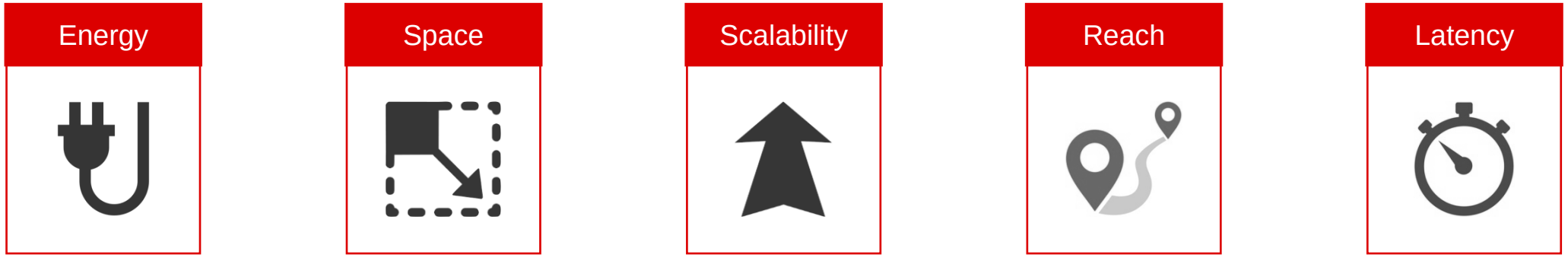


Entre Data Centers: Campus e Metro DCI como Resposta à Demanda de Recursos em Data Centers



A IA Exigirá Redes Ópticas Dentro e Ao Redor dos Data Centers

Distinct design requirements to ensure optimal AI workload performance inside and between data centers:



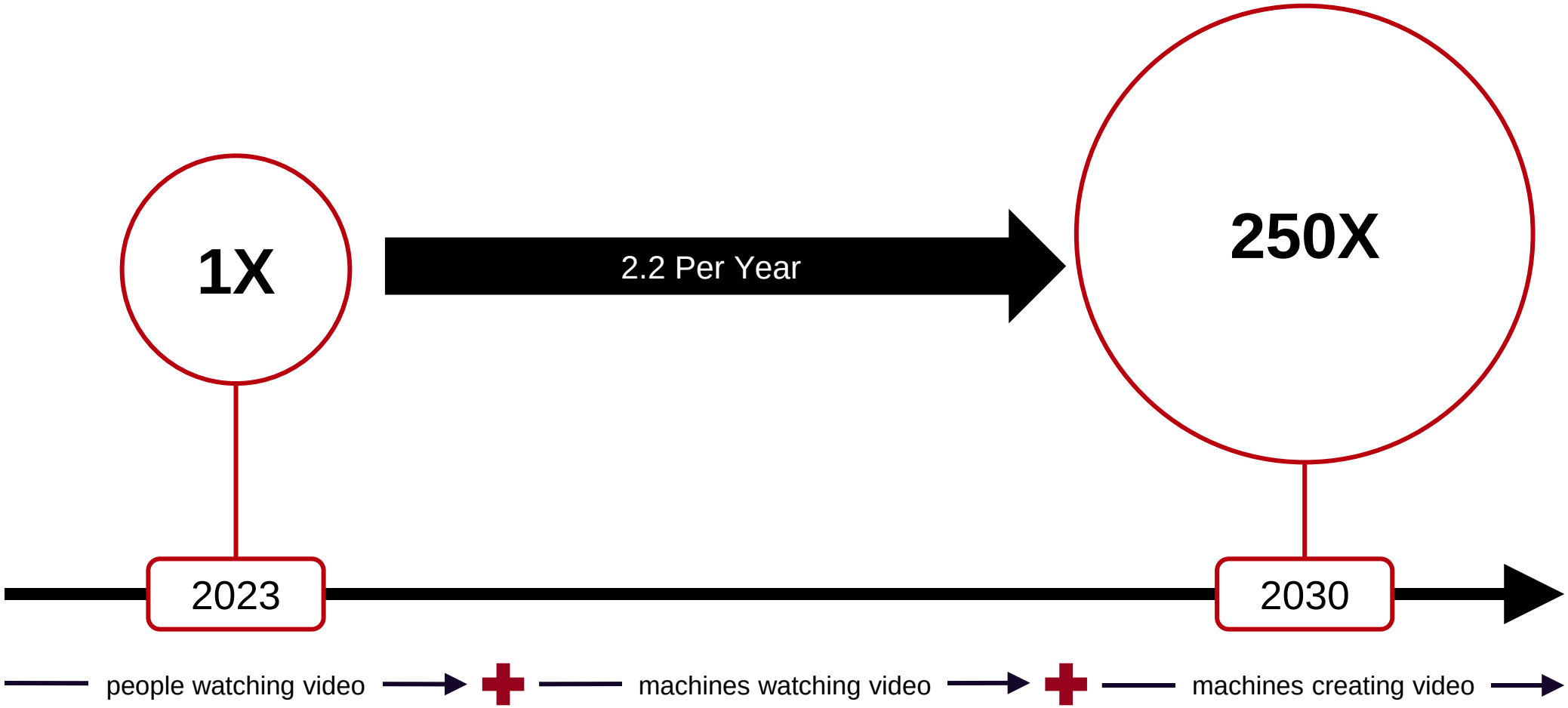
	Distance (km)	400Gb/s	800Gb/s	1600Gb/s	3200Gb/s
Submarine	10,000	Coherent	Coherent	Coherent	Coherent
Long Haul	100s to 1000s	Coherent	Coherent	Coherent	Coherent
Metro DCI	< 100	Coherent ZR	Coherent ZR	Coherent ZR	Coherent ZR
Campus	< 20	IMDD	IMDD, Coherent Lite	IMDD, Coherent Lite	Coherent Lite
Fabric	2	IMDD	IMDD	IMDD	IMDD, Coherent Lite
Cluster Optical	< 0.5	IMDD	IMDD	IMDD	IMDD, Coherent Lite
Cluster Copper	< 0.1	IMDD	IMDD/Parallel	Electrical	Electrical

→ One size does not fit all AI will require different optics, form-factors, and platforms

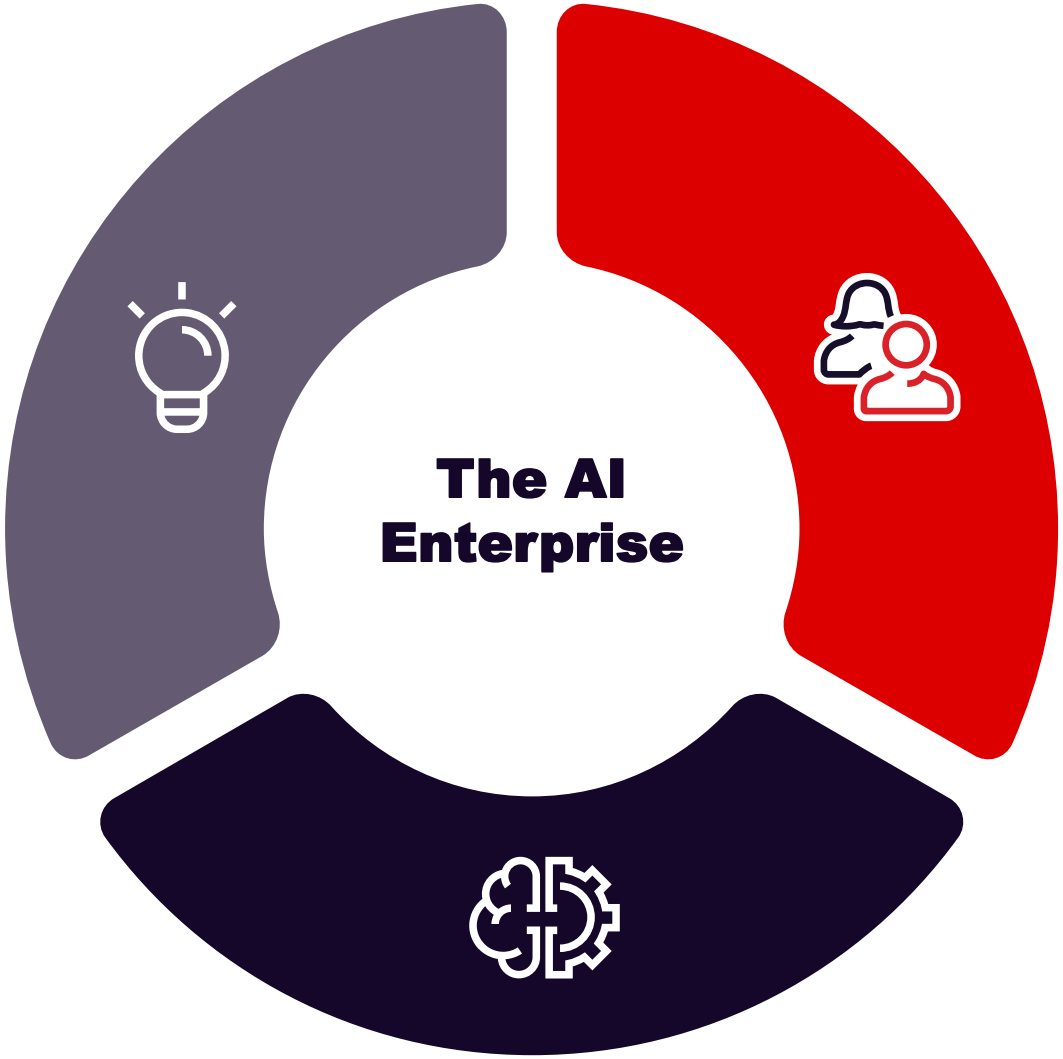
IMDD = Intensity Modulation Direct Detection

ciena

A Perspectiva do Usuário Final: O Impacto da IA no Crescimento da Internet



Para o corporativo, uma nova era de transformação digital impulsionada por IA



1

AI Agents

AI boosts staff productivity with conversational tools and virtual assistants.

2

Automation

Intelligent automation across business processes.

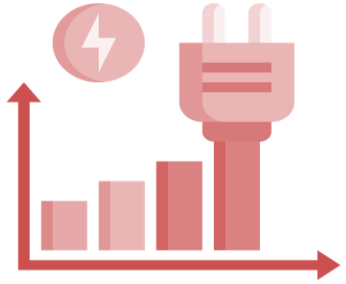
3

Data-driven

Real-time insights and predictive analytics

Demandas crescentes da IA trazem novos requisitos de conectividade para interconexão, Internet e corporativo

Connect AI training farms

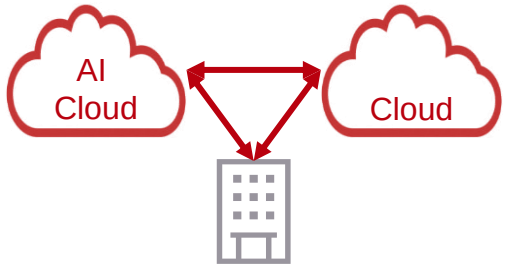


Distributed Training: Powering the New GWatt AI Data Center Campus

Inference everywhere

Inference gets more complex, multimodal data, data sovereignty

Enterprise datasets



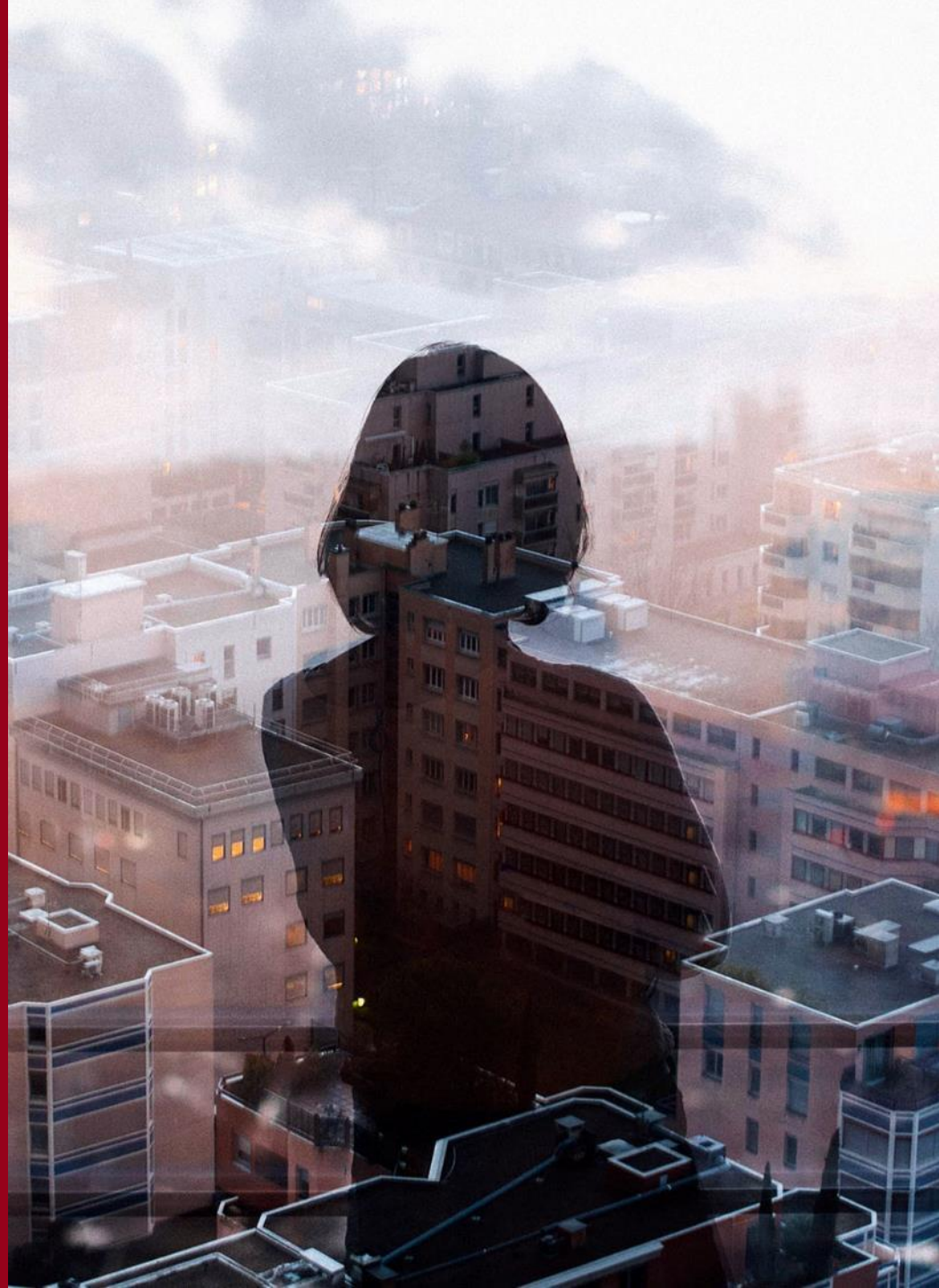
Enterprise AI initiatives stressing cloud exchange and onramp

New high-capacity DCI

Rapidly evolving inference compute and connectivity requirements

A new era of enterprise digital transformation

AIOps: Operação de Redes com IA



AIOps: Exemplos de Casos de Uso de Sucesso

Proactive failure detection and mitigation

Silent Fault Detection

Anomalies & Predictions

Fiber QoS Fingerprint Analysis

Fiber Health Assurance

Service assurance and performance optimization

Spectrum Defragmentation

Network operations workflow automation

Conversational help

DevOps acceleration

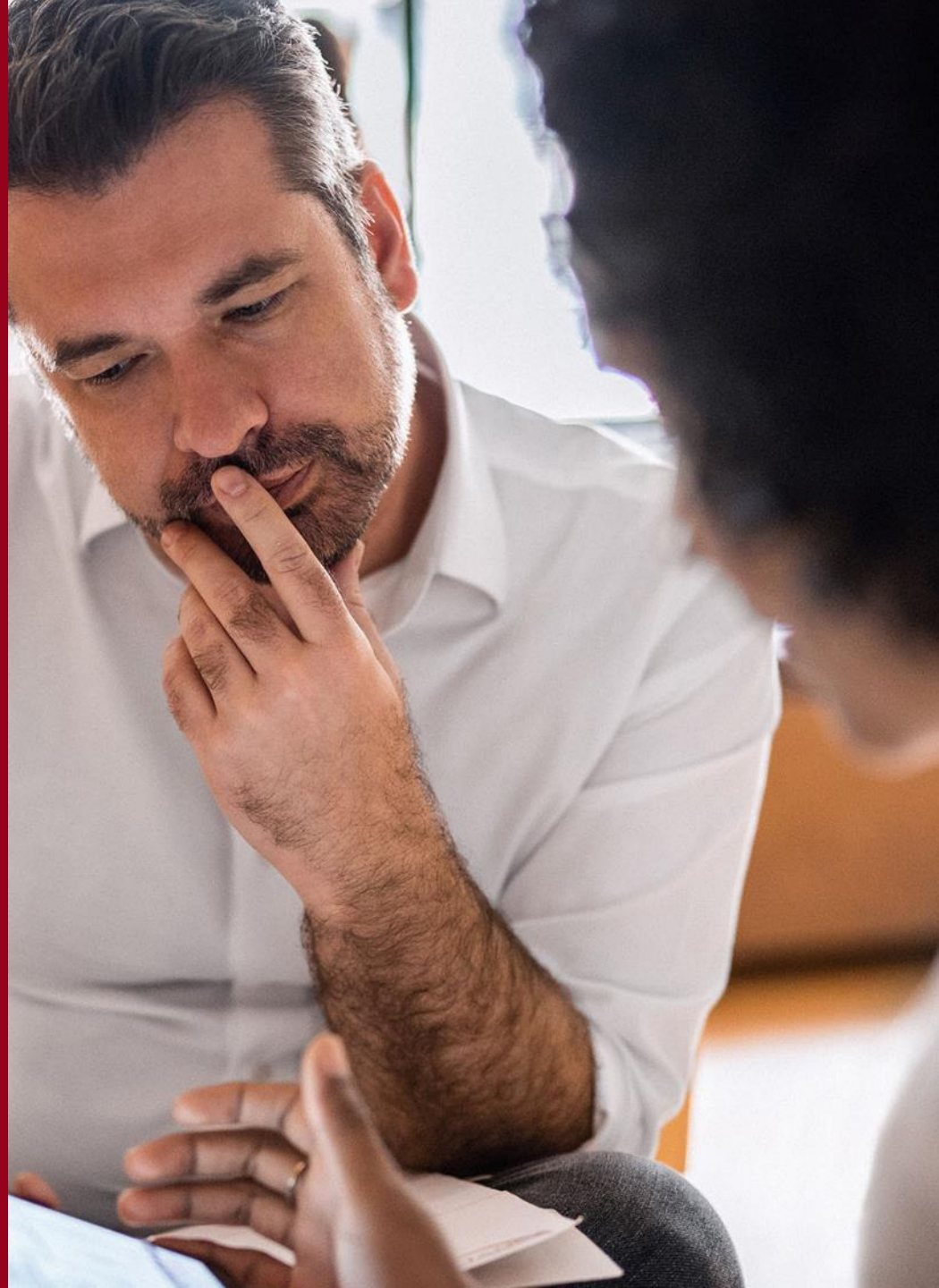
Ferramentas de AIOps: Os Princípios Tecnológicos Essenciais

1. Accuracy of AI responses will vary across a spectrum from ‘does not really matter’ to ‘absolutely has to be correct’. Network applications tend to the latter characterization which imposes significant test and verification requirements.

2. Tool chaining – combining AI with search and other tools in a ‘chain’ can improve accuracy of responses.

3. Retrieval Augmented Generation – training standard LLMs with specific content (like product documentation) improves accuracy and usefulness.

Considerações Finais



Considerações Finais

1. The AI race is a developing story, but there is no “turning back”

2. **Near-term connectivity opportunities (AI traffic):** Examine your current high-speed connectivity portfolio and infrastructure, and identify potential enhancements in DCI and wave services

3. **AIOps** use cases are available today (service agility, multi-layer visibility, and programmability)

Thank You