

Network Whisperer

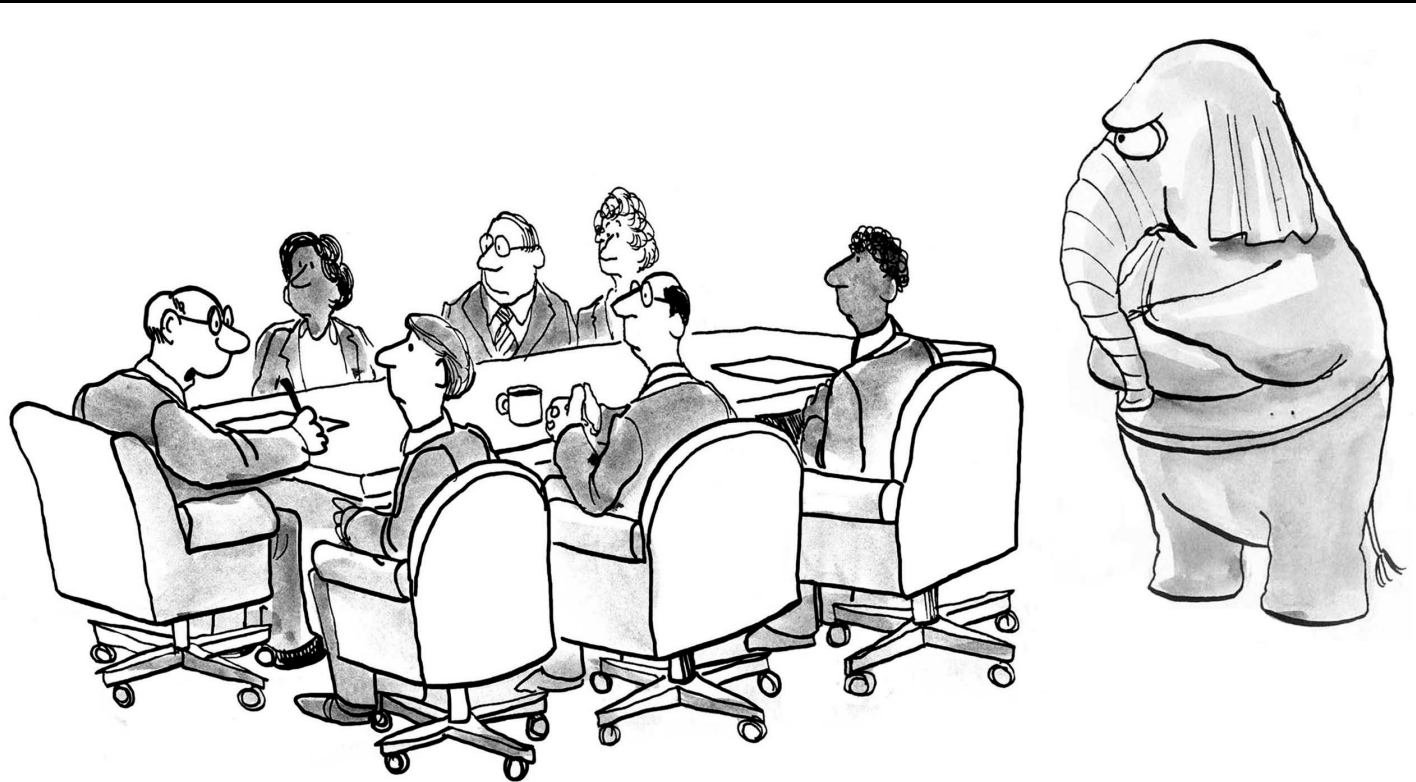
Building the ultimate AI network agent

AUTOCON 2

THE NETWORK AUTOMATION CONFERENCE

Agenda

- AI Hype: Myths vs. Reality
- The Evolution of Network Management
- Building an AI-powered Network Agent
- Real-World Impact and Use Cases
- The Road Ahead



**“I suppose I’ll be the one
to mention the elephant in the room.”**

AI Fatigue



A term used to describe the feeling of being overwhelmed by the rapid growth of artificial intelligence (AI) technologies.

AI Hype: Myth vs Reality

1

Myth: AI Will Replace Network Engineers

Reality: Your skills are needed now more than ever.

2

Myth: AI Will Solve All Network Engineering Problems

Reality: AI excels at pattern recognition and anomaly detection, but don't think you can just give it network data and your problems are solved.

3

Myth: AI-powered Networks are Fully Autonomous

Reality: AI can automate processes, but a human in the loop is still required.

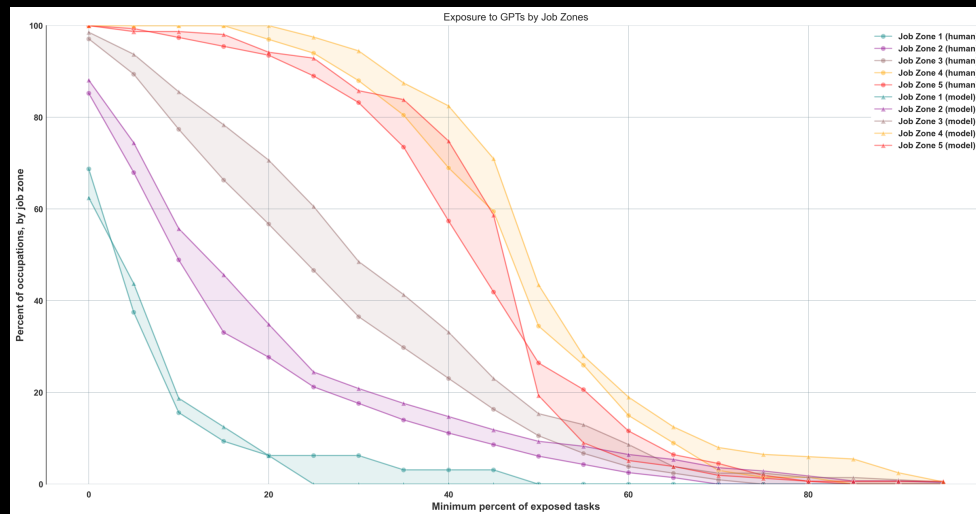
4

Myth: AI in Networking is ALL HYPE

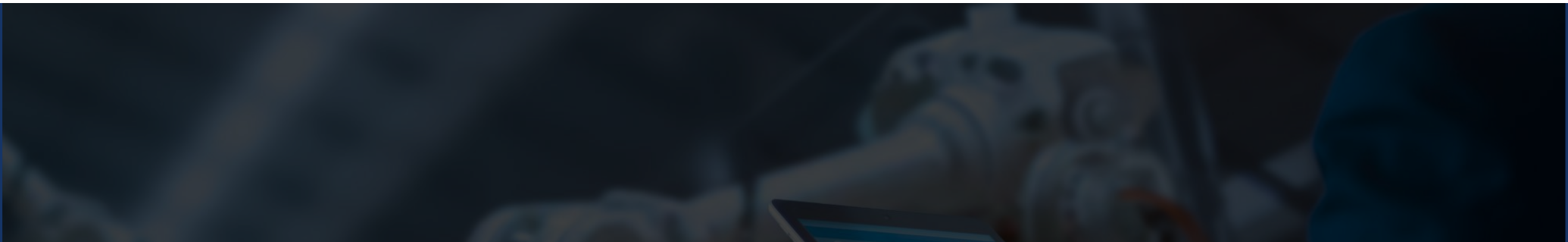
Reality: AI is the next evolution of automation.

AI Hype: Myth vs Reality

Estimated that 80% of workers belong to an occupation with at least 10% of its tasks exposed to LLMs, while 19% of workers are in an occupation where over half of its tasks are labeled as exposed.



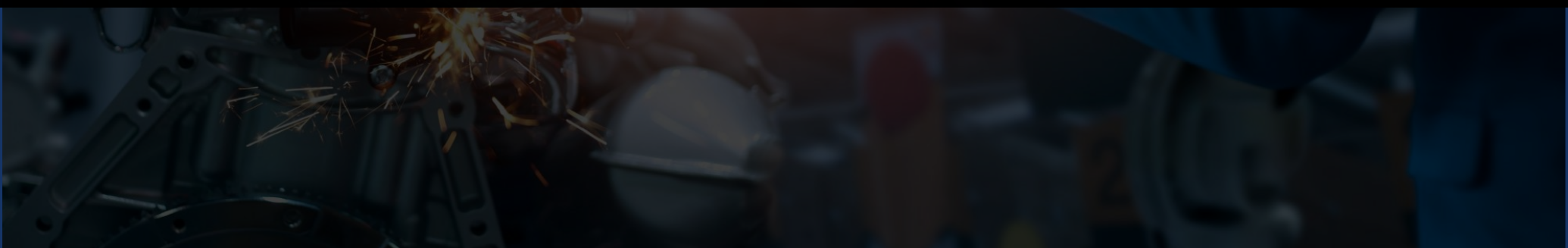
[2023: GPTs are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models](#)



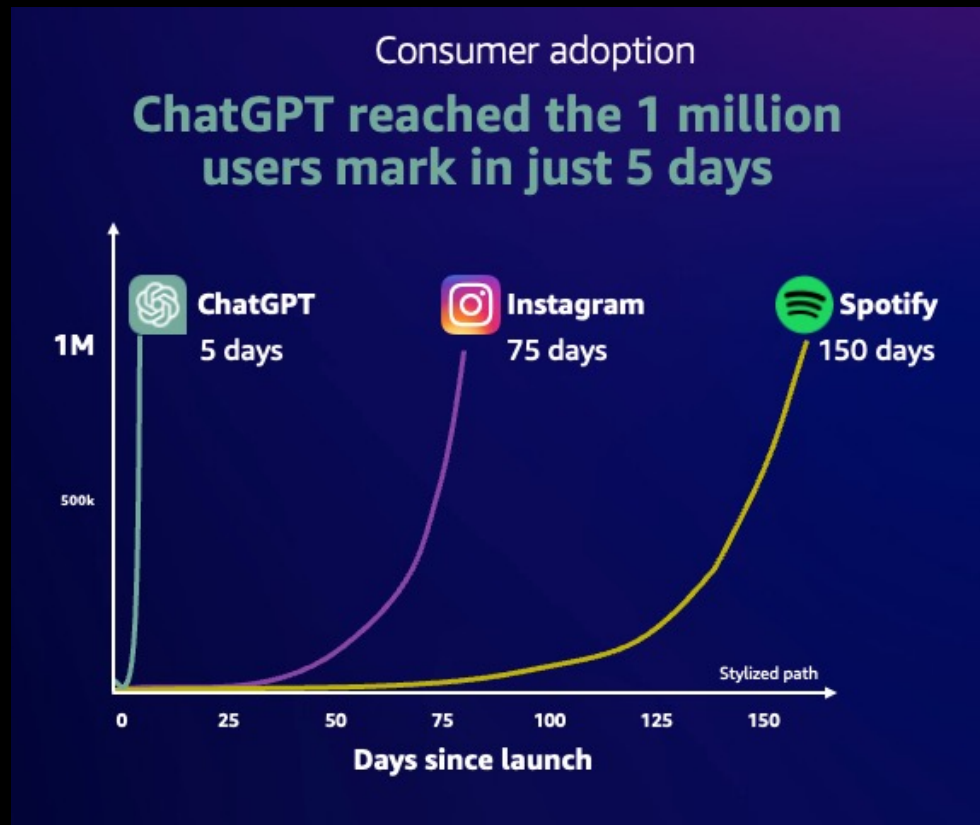
93% of companies believe AI will be a game changer
in driving growth and innovation in the sector

Deloitte, "Survey,"

<https://www2.deloitte.com/cn/en/pages/consumer-industrial-products/articles/ai-manufacturing-application-survey.html>

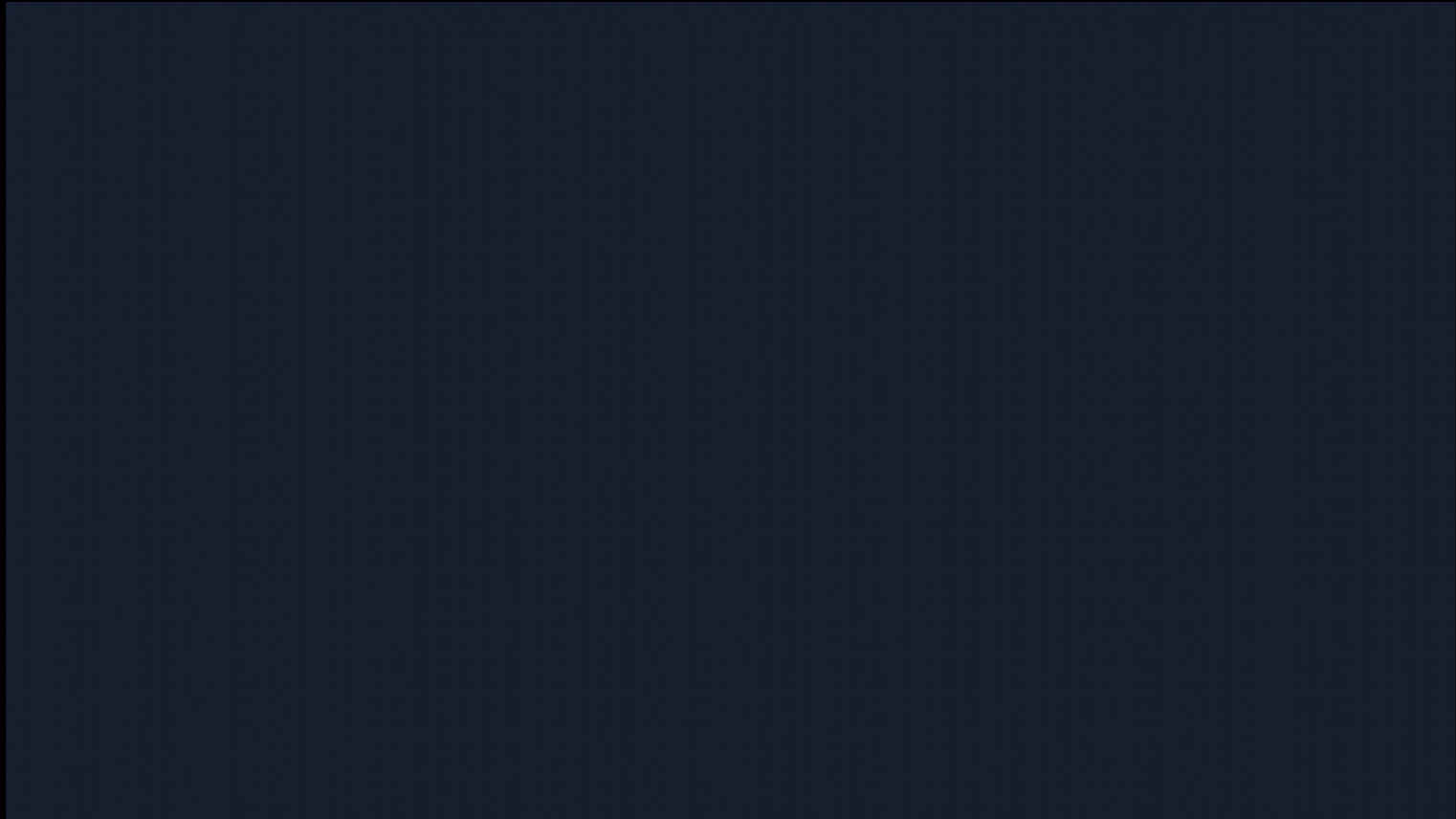


But why?



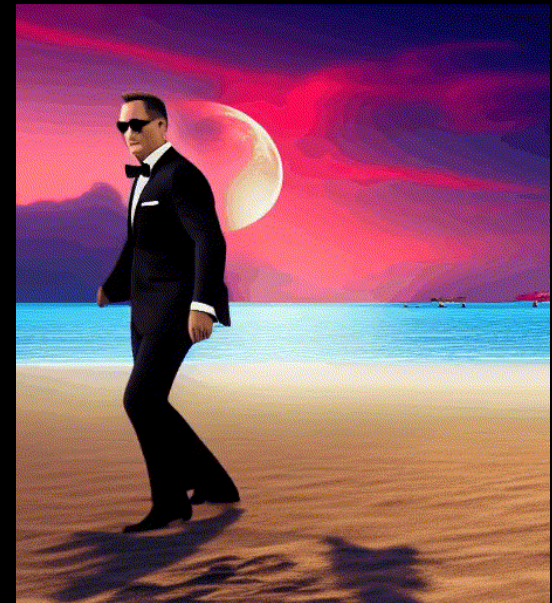
Let's talk about Artificial
Intelligence (AI)

What is generative AI?



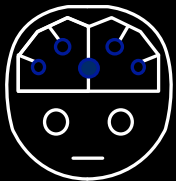
Working with generative AI

- Text, image, other media, and multi-modal models
- Summarization, analogies, translation and localization, personalization, with long memory and conversational capabilities



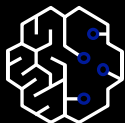
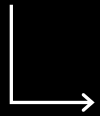
Prompt: "James Bond moonwalk on the beach, animation style"

Where does generative AI fit?



Artificial intelligence (AI)

Any technique that allows computers to mimic human intelligence using logic, if-then statements, and machine learning



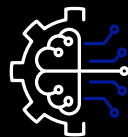
Machine learning (ML)

A subset of AI that uses machines to search for patterns in data to build logic models automatically



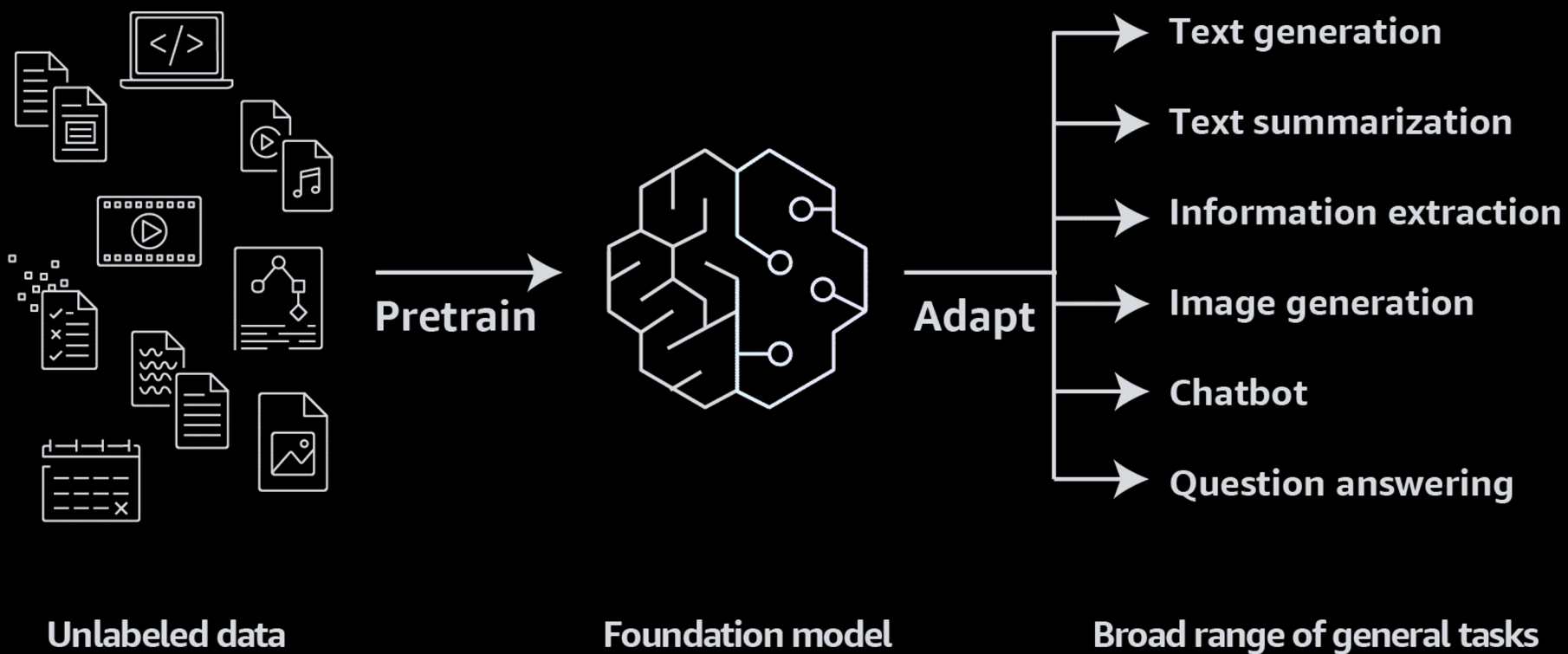
Deep learning (DL)

A subset of ML composed of deeply multi-layered neural networks that perform tasks like speech and image recognition



Generative AI

How does it work?



LLM Network Operations Use cases



Documentation

Configuration
Best Practices
Documentation



Health Checks

Performance
Analysis
Capacity Planning
Optimization



Real-time Monitoring

Alert Analysis
Root Cause Analysis
Fix Generation



Threat Detection

Traffic Analysis
Threat Detection
Policy Updates

“ChatOps”

LLM Challenges

**Lack of Domain
Knowledge**

**Temporal
Awareness**

**Training Data
Cut-off Date**

Hallucinations

**Lack of
Proprietary
Knowledge**

**Privacy
and
Security**

So what does this mean for
networking?

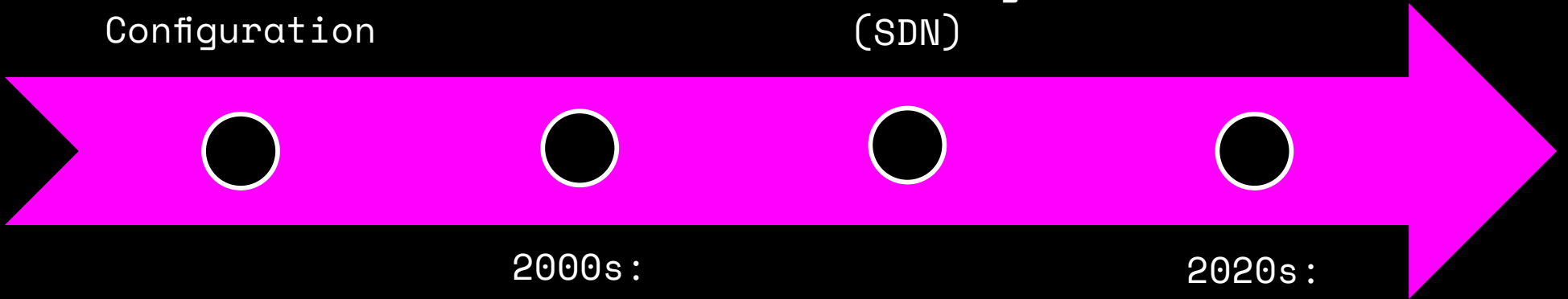
The Evolution of Network Management

1980-1990s: Manual
Configuration

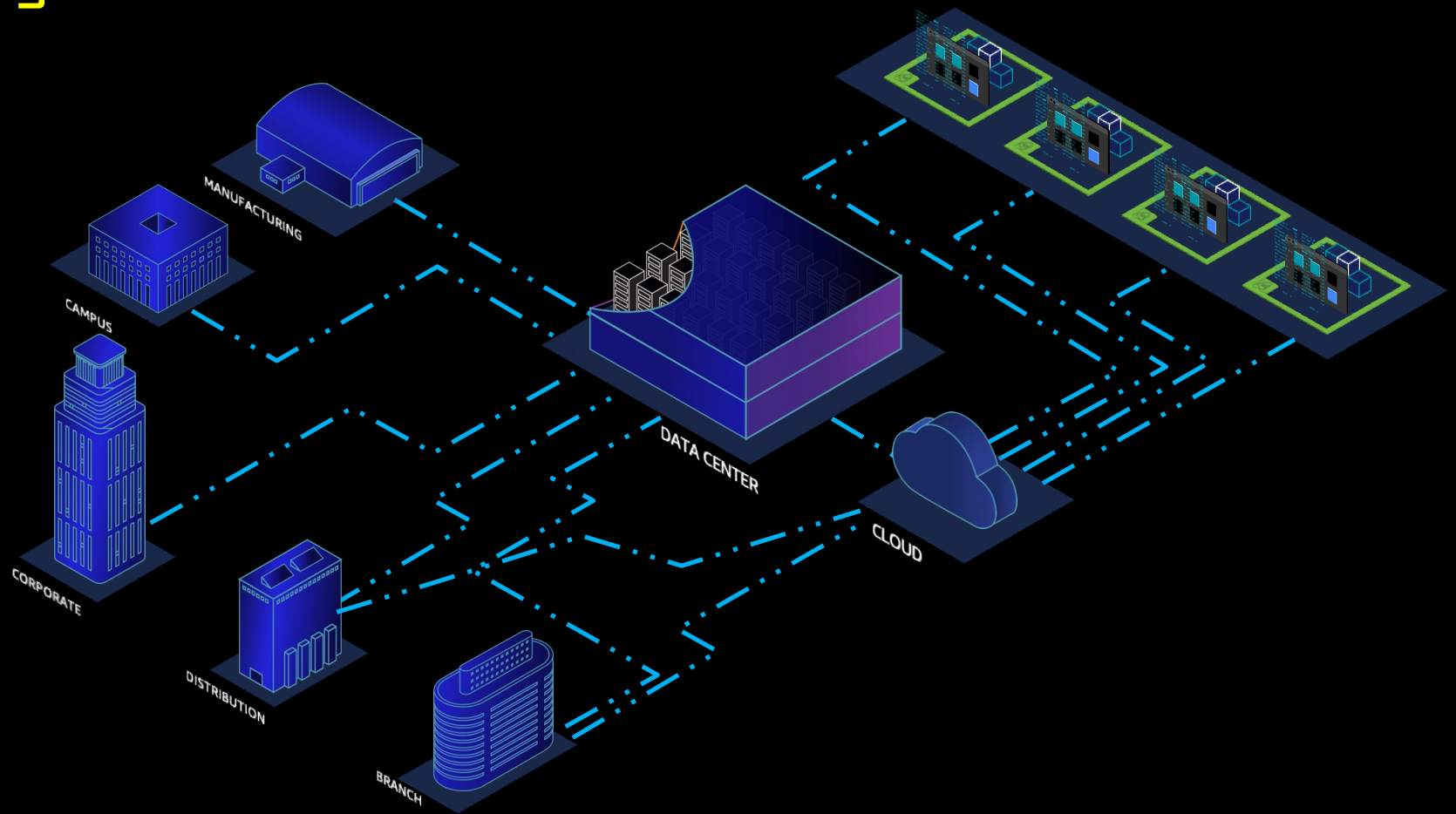
2010s:
Software
Defined
Networking
(SDN)

2000s:
Network
Management
Systems

2020s:
AI-powered
Networking



Today's networks



Application connectivity requirements



Traffic paths

Confined within a VPC
Inter-VPC
Inter-Region
On-premises access



Security requirements

Network segmentation?
Traffic inspection
required?
L7, L4 ?
Traffic initiation
one-way or two-way?
Zero trust
connectivity?



Performance

Bandwidth
Latency



Cost considerations

Data transfer
inter-AZ
Data processing
charges

What would a network agent look like?



What is an AI agent?

An application powered by a large language model (LLM) with a predefined system prompt

Has a set of **tools**, which are self-contained functions designed to perform a specific task



Network Agent Requirements

Must Have

- Data collection
- Single source of truth
- Network telemetry data
- Documentation access

Should Have

- Chat integration (NLP)
 - Historical data
 - Ticketing system
 - Approval workflows

Nice to Have

- Fine-tuned models
- CMDB integration
- Advanced analytics
- Custom dashboards

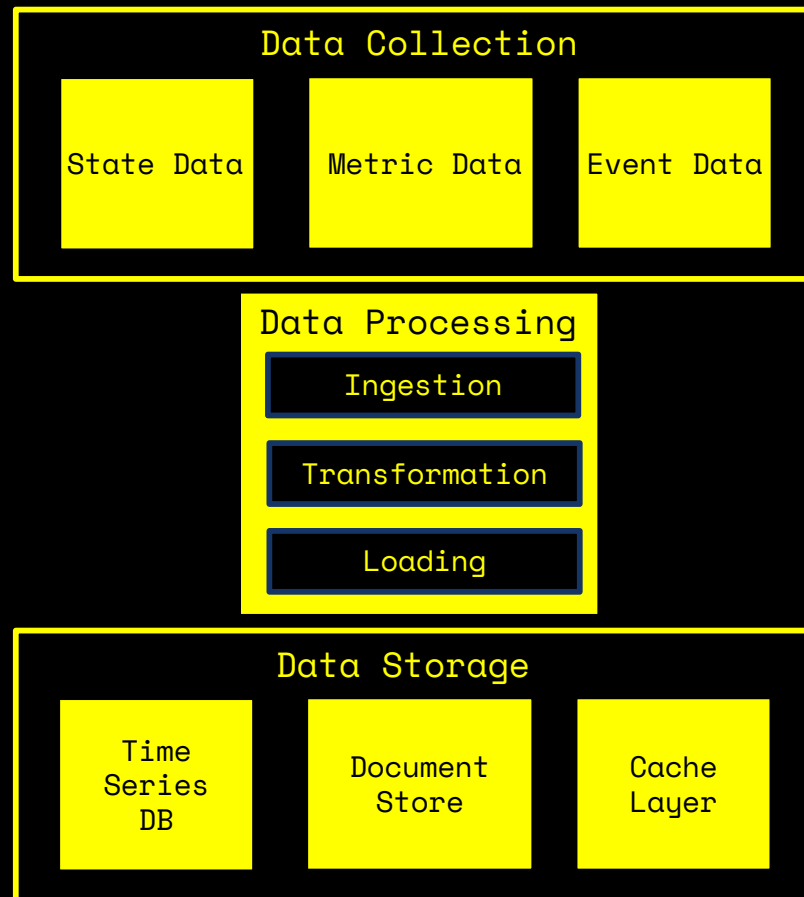
Key Risks and Considerations

- | | | |
|--------------------|--------------------|----------------------|
| • Data sensitivity | • Compliance needs | • Data quality |
| • Access control | • Skills gap | • Legacy integration |

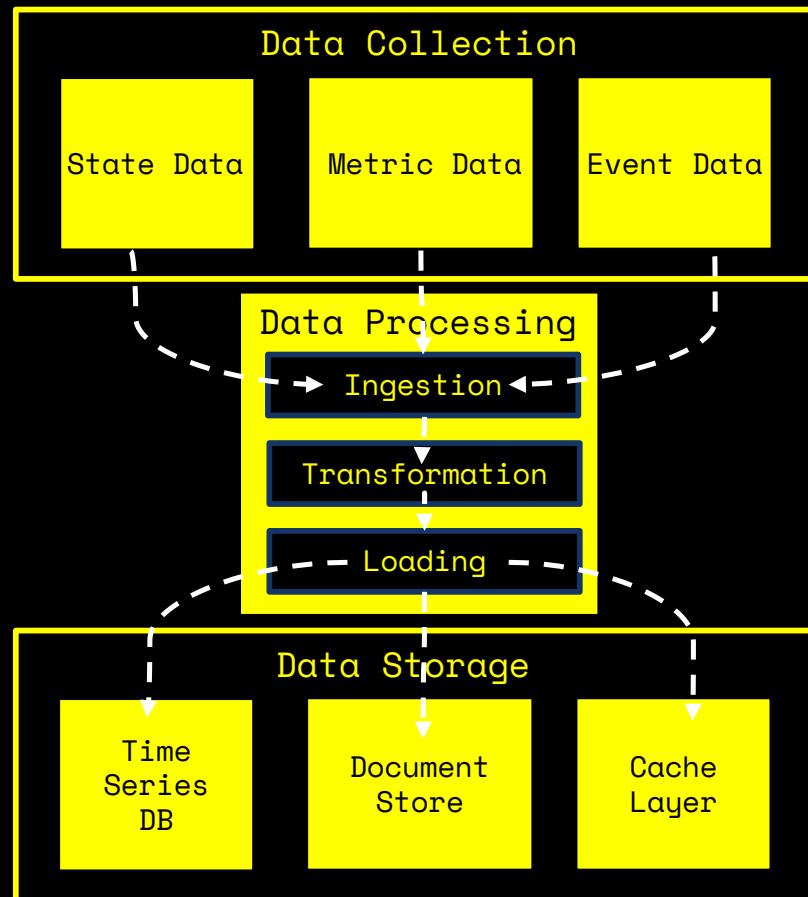
Architecture



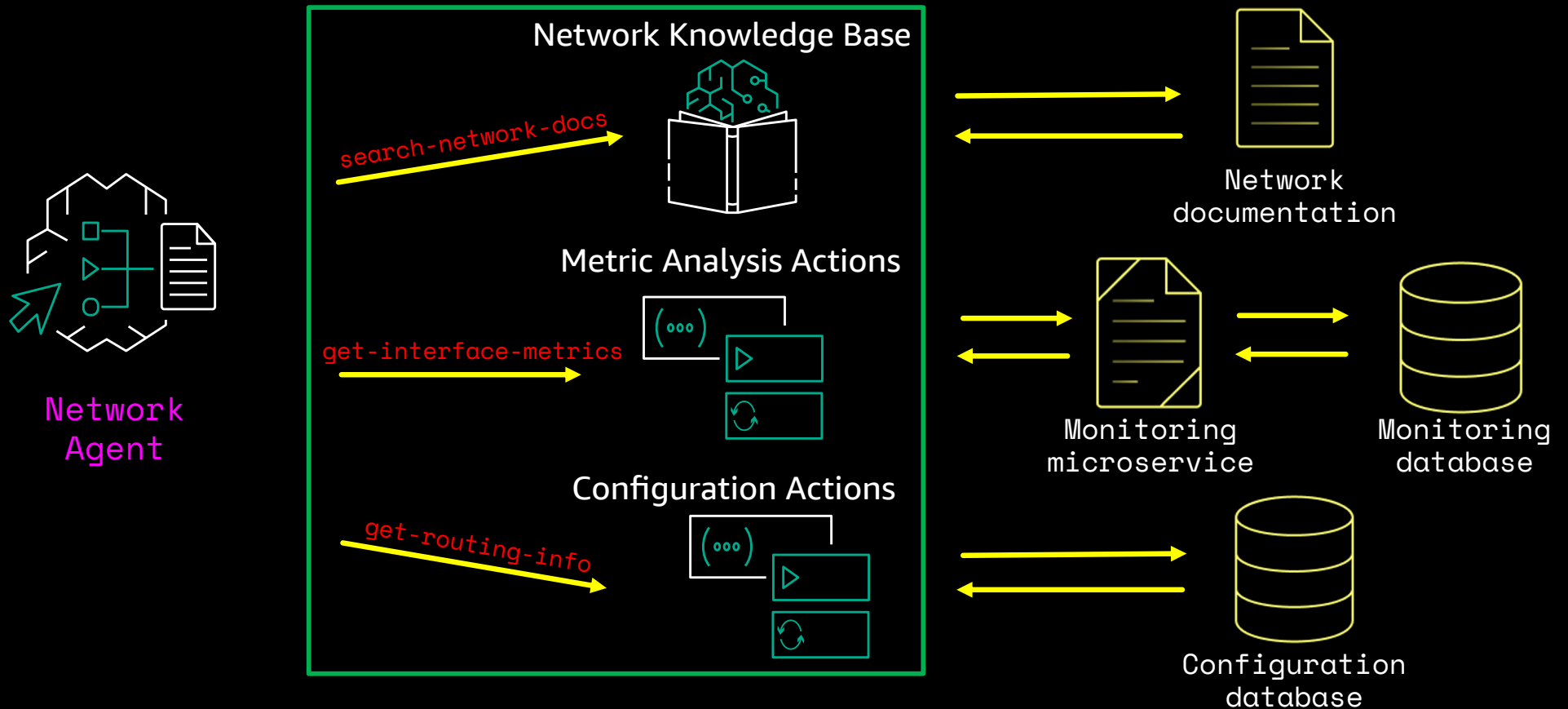
Implementation: Network Layer



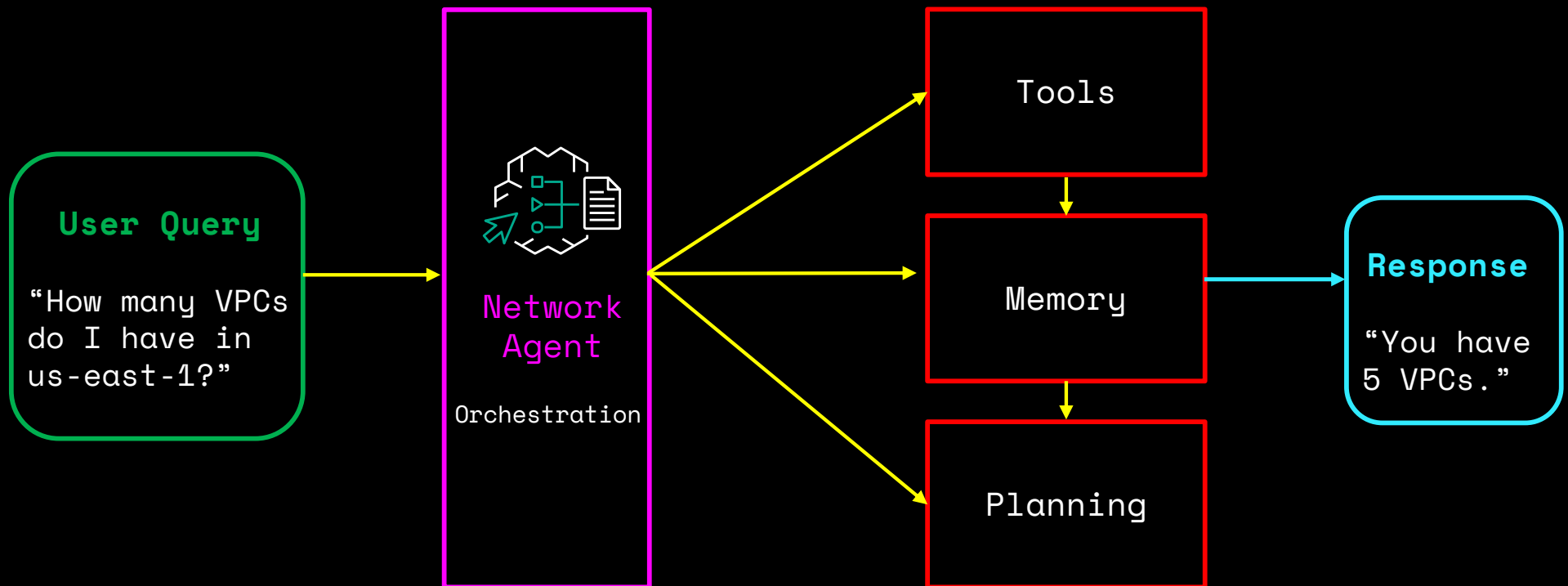
Implementation: Network Layer



Implementation: AI Layer



Implementation: AI Layer



Implementation: AI Layer Tools Example

```
def list_vpcs(region="us-west-2"):
    ec2 = boto3.client('ec2', region_name=region)
    response = ec2.describe_vpcs()
    vpcs = [{'VpcId': vpc['VpcId'], 'CidrBlock': vpc['CidrBlock'],
            'IsDefault': vpc['IsDefault']} for vpc in response['Vpcs']]
    return {
        'vpcs': vpcs,
        "region": region
    }

def check_internet_gateway(vpc_id, region="us-west-2"):
    ec2 = boto3.client('ec2', region_name=region)
    response = ec2.describe_internet_gateways(
        Filters=[
            {
                'Name': 'attachment.vpc-id',
                'Values': [vpc_id]
            }
        ]
    )
    internet_gateways = [
        {
            'InternetGatewayId': ig['InternetGatewayId'],
            'AttachedToVpc': vpc_id in [att['VpcId'] for att in ig['Attachments']]
        } for ig in response['InternetGateways']
    ]
    return {
        'vpc_id': vpc_id,
        'internetGateways': internet_gateways
    }
```

Implementation: AI Layer Tools Example

```
def handle_tool_use(tool_use):
    """
    Handle tool use requests from Claude.

    :param tool_use: Dictionary containing tool use details
    :return: Dictionary with the tool result in the format expected by Claude
    """
    tool_name = tool_use['name']
    input_data = tool_use['input']
    region = input_data.get('region', 'us-west-2')

    if tool_name == "list_vpcs":
        result = list_vpcs(region=region)
    elif tool_name == "check_internet_gateway":
        result = check_internet_gateway(input_data['vpc_id'], region=region)
    elif tool_name == "check_nat_gateway":
        result = check_nat_gateway(input_data['vpc_id'], region=region)
    elif tool_name == "get_route_tables":
        result = get_route_tables(input_data['vpc_id'], region=region)
    elif tool_name == "list_subnets":
        result = list_subnets(input_data['vpc_id'], region=region)
    elif tool_name == "describe_network_acls":
        result = describe_network_acls(input_data['vpc_id'], region=region)
    else:
        result = {"error": f"Unknown tool: {tool_name}"}

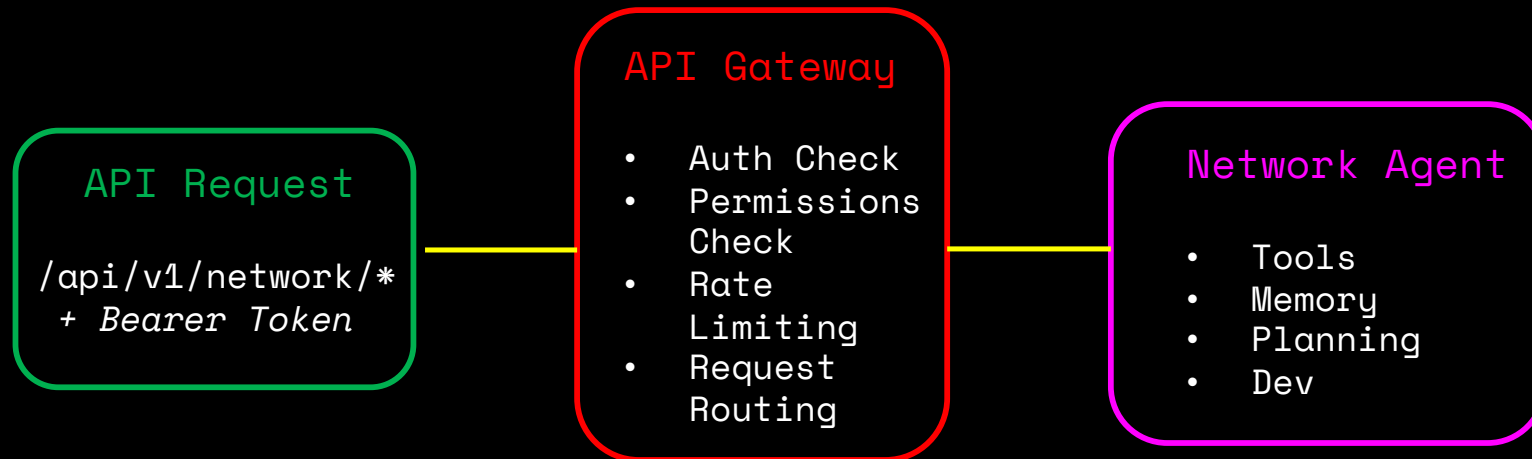
    return {
        "role": "user",
        "content": [
            {
                "toolResult": {
                    "toolUseId": tool_use['toolUseId'],
                    "content": [{"json": result}],
                    "status": "success"
                }
            }
        ]
    }
```

Implementation: AI Layer Tools Example

```
@tool
def get_interface_metrics(interface: str) -> dict:
    """Get interface metrics from monitoring system"""
    try:
        # Query Prometheus/monitoring system
        metrics = monitoring.query(
            f'interface_metrics{{{interface}}}[5m]'
        )

        # Process and format response
        return {
            "status": metrics["status"],
            "throughput": metrics["bytes_per_sec"],
            "errors": metrics["error_count"],
            "timestamp": metrics["last_updated"]
        }
    except Exception as e:
        return {
            "error": f"Failed to get metrics: {str(e)}",
            "status": "error"
        }
```

Implementation: API Integration



```
def create_converse_request(messages: List[Dict[str, Any]], tools: List[Dict[str, Any]],
    max_tokens: int = 500, temperature: float = 0.7, top_p: float = 1, model_key: str = DEFAULT_MODEL
) -> Dict[str, Any]:
    """
    Create a request object for the Bedrock converse API.
    """
    try:
        model_id = AVAILABLE_MODELS.get(model_key)
        if not model_id:
            raise ValueError(f"Invalid model key: {model_key}. Available models are: {' '.join(AVAILABLE_MODELS.keys())}")

        request = {
            "modelId": model_id,
            "messages": messages,
            "system": [{"text": SYSTEM_MESSAGE}],
            "inferenceConfig": {
                "maxTokens": max_tokens,
                "temperature": temperature,
                "topP": top_p
            },
            "toolConfig": {
                "tools": tools,
                "toolChoice": {"auto": {}}
            }
        }
        logger.debug(f"Created converse request for model: {model_id}")
        return request
    except Exception as e:
        logger.error(f"Error creating converse request: {str(e)}")
        raise
```

1. ModelId

2. Message history

3. System Message

4. Hyperparameters

5. List of Tools

Implementation: User Interface

CLI

```
(venv) bedrock-tools python3 main.py
2024-11-11 19:36:39,428 - INFO - Bedrock client initialized for region: us-west-2
Welcome to the AWS Network Assistant. You can ask about VPCs, Internet Gateways, NAT Gateways, Route Tables, and other network components.
Type 'exit', 'quit', or 'bye' to end the conversation.
You: How many VPCs are in us-east-1?
2024-11-11 19:36:54,304 - INFO - Successfully received response from Bedrock using model: anthropic.claude-3-haiku-20240307-v1:0
2024-11-11 19:36:54,304 - INFO - Claude: Okay, let me check the VPCs in the us-east-1 region.
2024-11-11 19:36:54,304 - INFO - Claude is using the list_vpcs tool.
2024-11-11 19:36:55,332 - INFO - Successfully received response from Bedrock using model: anthropic.claude-3-haiku-20240307-v1:0
2024-11-11 19:36:55,332 - INFO - Claude: The list_vpcs tool returned 6 VPCs in the us-east-1 region.
Assistant: The list_vpcs tool returned 6 VPCs in the us-east-1 region.
You: █
```

Implementation: User Interface

Streamlit

AWS Network Assistant

Ask about VPCs, Internet Gateways, NAT Gateways, Route Tables, and other network components.



How many VPCs are in us-west-2?



The output shows that there are 4 VPCs in the us-west-2 region.

Type your question here...



Implementation: User Interface

Custom HTML / React

AWS Network Assistant

Ask about VPCs, Internet Gateways, NAT Gateways, Route Tables, and other network components.

List my VPC IDs that are in us-east-1

The VPC IDs in the us-east-1 region are: - vpc-0f4776d68df31bbcd - vpc-0ea35bceb0040abc1 - vpc-01abbe01460a83df8 - vpc-09e9962bbc417563e - vpc-09f95050d9afcde6a - vpc-0b37885cb9d97d4ac Let me know if you need any clarification or have additional questions!

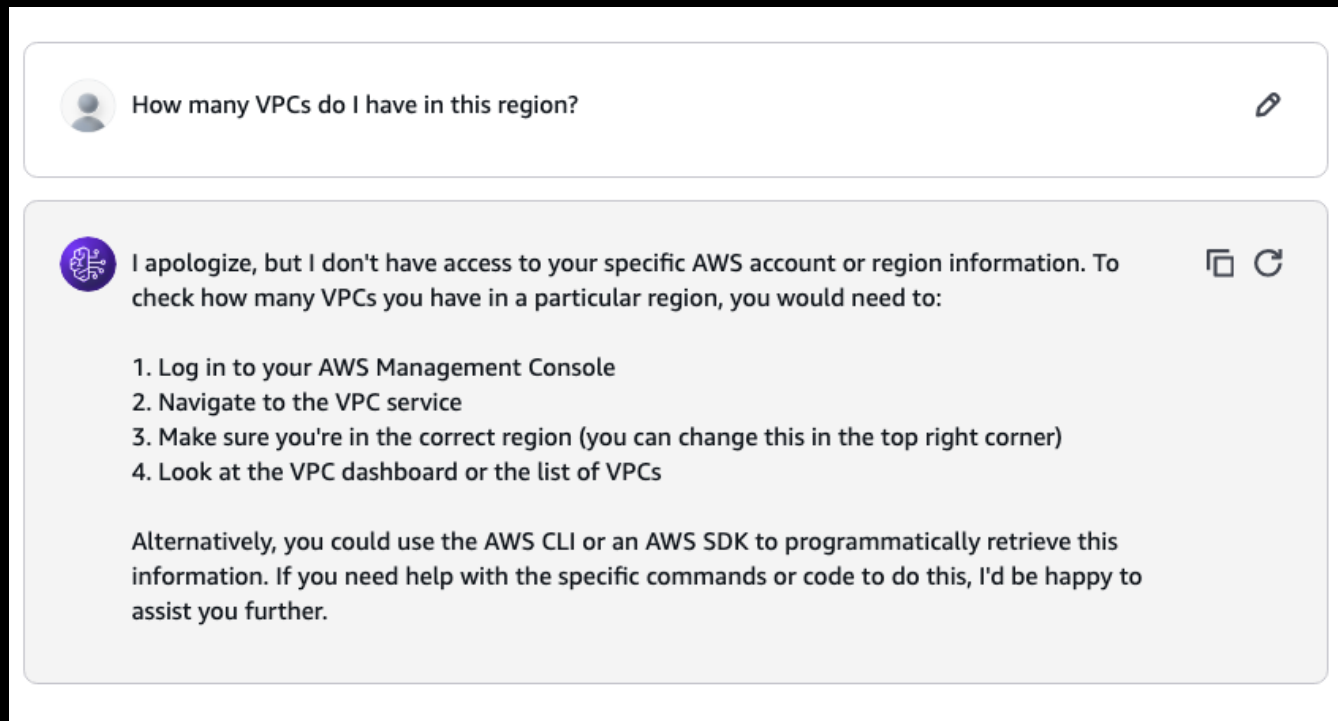
Can you tell me the routes for vpc-0f4776d68df31bbcd

The VPC vpc-0f4776d68df31bbcd has a single route table with the following route: - Destination CIDR block: 10.0.0.0/24 - Gateway ID: local (this is the default route for the VPC) This indicates that all traffic destined for the 10.0.0.0/24 CIDR block will be routed locally within the VPC. Let me know if you need any additional information about the routes for this VPC.

A silhouette of a person wearing a fedora hat and a suit, standing against a dark blue, textured background. The person's face is obscured by the hat and the overall lighting. The text "Network Agent in Action!" is overlaid in the center in a yellow, monospace-style font.

Network Agent in Action!

1. Using a basic network query



Asking Claude about my network 🙋

2. List available tools

```
(venv) bedrock-tools python main.py
14:59 | INFO | Bedrock client initialized for region: us-west-2
Welcome to the AWS Network Assistant. You can ask about VPCs, Internet Gateways, NAT Gateways, Route Tables, and other
network components.
Type 'exit', 'quit', or 'bye' to end the conversation.
You: What tools do you have available?
14:59 | INFO | Successfully received response from Bedrock using model: anthropic.claude-3-haiku-20240307-v1:0
14:59 | INFO | Claude: I have the following tools available:
14:59 | INFO | Claude is using the list_functions tool.
14:59 | INFO | Successfully received response from Bedrock using model: anthropic.claude-3-haiku-20240307-v1:0
Assistant: Apologies, it seems I cannot directly list the available tools. Let me describe them instead:

- `list_vpcs`: This tool allows you to list the VPCs (Virtual Private Clouds) available in a specified AWS region.
- `check_internet_gateway`: This tool checks if there is an Internet Gateway attached to a specified VPC.
- `check_nat_gateway`: This tool checks if there is a NAT Gateway configured for a specified VPC.
- `get_route_tables`: This tool retrieves the route tables associated with a specified VPC.
- `list_subnets`: This tool lists the subnets within a specified VPC.
- `describe_network_acls`: This tool describes the Network ACLs (Access Control Lists) configured for a specified VPC.

Please let me know if you have any other questions about the available tools and how to use them.
You: 
```

Asking our agent about the available tools.

3. Basic Network Agent Query

You: How many VPCs are in this region?

15:19 | INFO | Successfully received response from Bedrock using model: anthropic.claude-3-haiku-20240307-v1:0

15:19 | INFO | Claude: Okay, let me use the `list\_vpcs` tool to check how many VPCs are in the current region.

15:19 | INFO | Claude is using the list_vpcs tool.

15:19 | INFO | Successfully received response from Bedrock using model: anthropic.claude-3-haiku-20240307-v1:0

15:19 | INFO | Claude: Based on the output, there are 4 VPCs in the us-west-2 region.

Assistant: Based on the output, there are 4 VPCs in the us-west-2 region.

Asking the network agent about my network

3. Basic Network Agent Query

You: How many VPCs are in this region?

15:19 | INFO | Successfully received response from Bedrock using model: anthropic.claude-3-haiku-20240307-v1:0

15:19 | INFO | Claude: Okay, let me use the `list\_vpcs` tool to check how many VPCs are in the current region.

15:19 | INFO | Claude is using the list_vpcs tool.

15:19 | INFO | Successfully received response from Bedrock using model: anthropic.claude-3-haiku-20240307-v1:0

15:19 | INFO | Claude: Based on the output, there are 4 VPCs in the us-west-2 region.

Assistant: Based on the output, there are 4 VPCs in the us-west-2 region.

Asking the network agent about my network

Resources by Region

You are using the following Amazon VPC resources

[VPCs](#)

US West [4](#)

► See all regions

It got it right!

4. Complex Network Agent Query

“Find what **region** this **vpc**
<vpc>vpc-01bfb2e935767a000</vpc>
is in and determine if it has
internet connectivity.”

4. Complex Network Agent Query

You: Find what region this vpc <vpc>vpc-01bfb2e935767a000</vpc> is in and determine if it has internet connectivity.
15:30 | INFO | Successfully received response from Bedrock using model: claude_3_haiku
15:30 | INFO | Claude: Okay, let's use the tools provided to find the information you need about that VPC.

First, I'll use the

15:30 | INFO | Claude is using the `check_internet_gateway` tool.

Thinking

15:30 | INFO | Successfully received response from Bedrock using model: claude_3_haiku

15:30 | INFO | Claude: The output shows that this VPC (vpc-01bfb2e935767a000) has an internet gateway attached, which means it has internet connectivity.

Now, to determine the region this VPC is in, I'll use the

15:30 | INFO | Claude is using the `list_vpcs` tool.

15:30 | INFO | Successfully received response from Bedrock using model: claude_3_haiku

15:30 | INFO | Claude: The output shows that the VPC `vpc-01bfb2e935767a000` is located in the `us-west-2` region.

The agent has to think and plan

4. Complex Network Agent Query

You: Find what region this vpc <vpc>vpc-01bfb2e935767a000</vpc> is in and determine if it has internet connectivity.
15:30 | INFO | Successfully received response from Bedrock using model: claude_3_haiku
15:30 | INFO | Claude: Okay, let's use the tools provided to find the information you need about that VPC.

First, I'll use the

Planning

Thinking

15:30 | INFO | Claude is using the check_internet_gateway tool.
15:30 | INFO | Successfully received response from Bedrock using model: claude_3_haiku
15:30 | INFO | Claude: The output shows that this VPC (vpc-01bfb2e935767a000) has an internet gateway attached, which means it has internet connectivity.

Now, to determine the region this VPC is in, I'll use the

Planning

15:30 | INFO | Claude is using the list_vpcs tool.
15:30 | INFO | Successfully received response from Bedrock using model: claude_3_haiku
15:30 | INFO | Claude: The output shows that the VPC `vpc-01bfb2e935767a000` is located in the `us-west-2` region.

The agent has to think and plan

4. Complex Network Agent Query

You: Find what region this vpc <vpc>vpc-01bfb2e935767a000</vpc> is in and determine if it has internet connectivity.

15:30 | INFO | Successfully received response from Bedrock using model: claude_3_haiku

15:30 | INFO | Claude: Okay, let's use the tools provided to find the information you need about that VPC.

First, I'll use the

Planning

Thinking

15:30 | INFO | Claude is using the check_internet_gateway tool.

15:30 | INFO | Successfully received response from Bedrock using model: claude_3_haiku

15:30 | INFO | Claude: The output shows that this VPC (vpc-01bfb2e935767a000) has an internet gateway attached, which means it has internet connectivity.

Now, to determine the region this VPC is in, I'll use the

Planning

15:30 | INFO | Claude is using the list_vpcs tool.

15:30 | INFO | Successfully received response from Bedrock using model: claude_3_haiku

15:30 | INFO | Claude: The output shows that the VPC `vpc-01bfb2e935767a000` is located in the `us-west-2` region.

Assistant: The output shows that the VPC `vpc-01bfb2e935767a000` is located in the `us-west-2` region.

In summary, the VPC `vpc-01bfb2e935767a000` is located in the `us-west-2` region and has internet connectivity through an attached internet gateway.

Response

The agent has to think and plan

4. Complex Network Agent Query

You: Find what region this vpc <vpc>vpc-01bfb2e935767a000</vpc> is in and determine if it has internet connectivity.
15:30 | INFO | Successfully received response from Bedrock using model: claude_3_haiku
15:30 | INFO | Claude: Okay, let's use the tools provided to find the information you need about that VPC.

Resource map [Info](#)

VPC [Show details](#)

Your AWS virtual network

opensearch-privatelink-demo-vpc

Subnets (2)

Subnets within this VPC

us-west-2a

 opensearch-privatelink-demo-pub...

 opensearch-privatelink-demo-priv...

Route tables (3)

Route network traffic to resources

rtb-09fdf16eeeb9b43f2

dev-public-rt

dev Private Routes (AZ1)

Network connections (2)

Connections to other networks

opensearch-privatelink-demo-igw [↗](#)

opensearch-privatelink-demo-nat-gat...

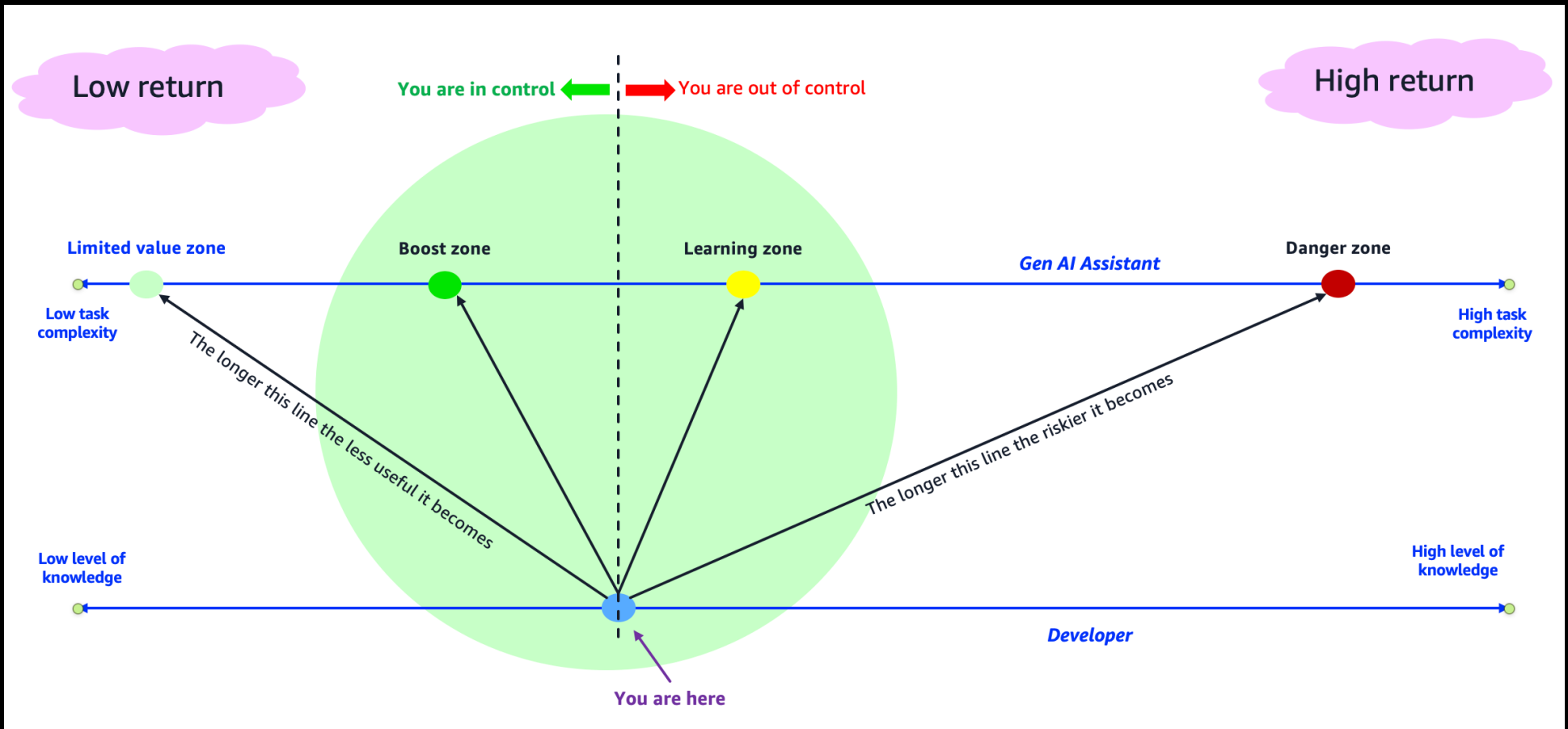
Response

The agent has to think and plan

It got it right again!

Real World Impact

Selecting the right use cases



Considerations



The Road Ahead



Strategies to adapting in the AI era

Start Small

- ✓ Begin with specific use cases
- ✓ Build proof of concepts
- ✓ Document learnings and failures
- ✓ Create feedback loops

Focus on the fundamentals

- ✓ Standardize data collection
- ✓ Establish governance frameworks
- ✓ Define clear security boundaries
- ✓ Create robust documentation

Enhance AI Capabilities

- ✓ Implement advanced NLP interfaces
- ✓ Build domain-specific models or RAG
- ✓ Create test and validation frameworks

Collaborate

- ✓ Mentor and train network engineers
- ✓ Build automation and AI/ML skills
- ✓ Create centers of excellence
- ✓ Foster innovation culture

“AI is not going to replace humans,
but humans with AI are going to
replace humans without AI.”

Karim Lakhani – Professor Harvard University

Thank you!

- @LabEveryday



<https://bit.ly/3UPsYoG>



Questions?