Message Platform:
These messages all aim at selling the Cisco products that stack to create and/or streamline an organization's AI infrastructure. They are broken apart to sell what I think are compelling arguments for adoption. For example, the full functionality of AI is something you need to stay competitive in your industry [e.g., Thousand Eyes, Flashstack, etc.], which means you need safe infrastructure that can be built quickly [e.g,. Splunk Flexpod]; however, this also needs to be built sustainably. Cisco/ConRes gives you all three.

Cisco Validated Design ties together the functionality (1) and infrastructure (2) arguments. **If the focus should be pure infrastructure, then combine these two and think about adding a ConRes-specific sell [ConRes: ConRes engineers can help you determine what you need].**

**The best claim from the Readiness Index for these two sells that I could find is: "An organization adopting AI requires a robust digital infrastructure encompassing specialized hardware (like GPUs), the right level of latency and throughput, software platforms, data storage and management solutions, and enhanced cybersecurity measures."

[Functionality Sell] Short-Term (1): *Simple-yet-powerful AI systems*.
- Thousand Eyes: Streamline Network Visibility and Management
- Nvidia: GPUs that Power AI-Workloads.
    - Flashstack

[Infrastructure Sell] Short-Term (2): *Build AI systems quickly and safely in your data center.*
- Splunk: AI Management and Security
- Cisco Ethernet Networking
- Nvidia: GPUs that Power AI-Workloads.
    - Flexpod
- Cisco Security Cloud or Secure Access

[Sustainability Sell] Long-Term: *Modernize your on-prem data center for a hybrid workforce.*
- Cisco's Investments in Acquisitions and Development

**One claim that really stood out when reading through the Cisco Readiness Index was the following: "When creating a strategic plan for AI adoption, organizations need to prioritize short-term actions like infrastructure development, talent training, and pilot projects, while planning for long-term goals such as scalability, continuous learning, and robust governance." It seems to me that this is a place where we can really hammer home Cisco's investments. Not only do these organizations need to solidify their AI infrastructure, but they also need to ensure that it isn't outclassed/outdated in a few years. Cisco's continued investment in AI can help them guard against this.

Relevant Stat Pulls from Readiness Index:
- 97% say the urgency to deploy AI-powered technologies has increased in their company within the past six months.
- 86% of companies across the globe are not fully prepared to leverage AI and AI-powered technologies to the fullest potential.
- 95% of respondents recognize that AI will increase infrastructure workloads.

- 61% of respondents anticipate having just one year or less to implement their AI strategy before incurring significant negative business impacts from falling behind.
- 76% require further data center graphics processing units (GPUs) to support future AI workloads.
- 81% of respondents admitted that their data exists in silos across the organization.

Notes on Specific Cisco Products:

Cisco Ethernet Networking: If customers want to improve their purchasing power, they need to move to multi-source components. Cisco Ethernet Networking offers advantages of both economy and scale while providing a broadly adopted standard. "We make it possible to standardize switching and routing with a single silicon architecture — Cisco Silicon One." This reduces the complexity of an organization's network architectures by managing one experience across its entire network/network functions. Cisco also produces a full range of silicon, systems and optics to eliminate the premiums introduced when adding additional vendors to a value chain as well as drive end-to-end innovations.

Cisco Security Cloud: "If it's connected, you're protected." This Security Cloud is also powered by AI, which allows for more efficient and effective response actions. Additionally, the Security Cloud is built on zero-trust principles, which secures access for in-person, hybrid and remote workers without compromising protection. This is a simplified solution that nonetheless delivers remarkable agility and flexibility. It also allows organizations to avoid vendor lock-in by featuring unified management for multiple networks and cloud providers. Finally, the Security Cloud can scale up or down to accommodate organizational needs while also offering open APIs for easy third-party integrations.

Flashstack: A holistic stack that incorporates accelerated computing, essential AI software and pre-trained models to simplify the deployment of AI models across diverse applications. Flashstack provides a foundational reference architecture for Generative AI; it also enables organizations to deploy Large Language Models and other Generative AI models. The key benefit here is that FlashStack components are integrated, so organizations can deploy the solution quickly and economically while eliminating many of the risks associated with researching, designing, building, and deploying similar solutions from the foundation. There is an inherent ability here to maintain consistency at scale.

Flexpod: This validated end-to-end AI solution for full stack provides faster time to value, better scalability, easier system upgrades, better reliability, improved staff productivity, and a longer lifespan for the platform. What does this ultimately mean? Reduced AI infrastructure complexity with cloud-based operations that deliver global visibility, consistency, and control while simultaneously enhancing operational efficiency and agility. Furthermore, it also enhances security with a holistic approach including secure separation, device hardening, micro-segmentation, encryption, and zero-trust architecture for data protection and threat mitigation.

Splunk: This AI-powered security and observability solution accelerates detection, investigation and responsiveness for organizational networks. It is supported by an enterprise-class platform that enables shared data, context and workflows. Now that Splunk is partnered with Cisco, organizations will be further empowered to deliver seamless, secure customer and employee experiences across the physical, digital, and AI-powered worlds. This is due to Splunk's ability to monitor, investigate and respond rapidly at scale for organizations of all sizes, utilizing cloud, network, and endpoint traffic for superior visibility.

Thousand Eyes: Offers a combination of active and passive monitoring techniques along with real-time internet outage detection. The result is a deeper insight into the user experience delivered via your organizational applications and services. Thousand Eyes integrates monitoring and visualization of device health, end-to-end network paths and the performance of your internally hosted and cloud applications. This allows organizations to identify critical dependencies in their internal networks and monitor how device health impacts their application performance.

Challenge: Companies Face Barriers Implementing Infrastructure for On-Premises AI

Traditional enterprise on-premises infrastructure was not designed to be able to handle the power consumption and scheduler-driven workloads introduced by Generative AI. For many organizations, this creates complex obstacles, particularly when integrating with outdated architectures. However, the urgency is clear: more than 60% of companies recognize that they have just one year or less to implement their AI strategy before incurring significant negative business impacts from falling behind.

To implement a robust AI strategy, companies need to understand the key challenges they face:

1. **Cost-Effectiveness**: While investing in robust AI infrastructure is essential, the need to optimize hardware choices and minimize energy consumption will often dictate how much capital on-prem infrastructure requires. Furthermore, on-prem AI systems require substantial computational resources for intensive AI processes, which can make it cumbersome for organizations to identify the best solution for their needs. There are additional maintenance costs that can affect an organization's ability to maximize its on-prem infrastructure.

2. **Performance**: Organizations are faced with prioritizing both hardware acceleration (via GPUs) and efficient software frameworks to maximize their return on performance. The difficulty is identifying the optimal balance of GPUs and CPUs to provide outstanding performance while still accounting for the cost effectiveness of your AI deployment. In other words: investing in high-performance hardware does not solve your organization's performance challenge. This is also because this hardware must be repurposed for other tasks when your AI is not in demand.

3. **Scalability**: As AI workloads grow, scalability becomes crucial. Your high-performance AI system needs to perform as well as a single workstation as it does an AI cluster. By regularly auditing your data center, you might be able to eliminate applications and systems you no longer need. However, there is also a software challenge: it is critical that organizations integrate their AI services with their existing technology stacks. Otherwise, they risk not only service disruptions but also compatibility issues during deployment.

4. **Security**: To manage sensitive data, organizations require robust security measures that can safeguard their data from unauthorized access, breaches, and cyber threats. On the one hand, on-prem infrastructure will allow companies to maximize control over their data and models; on the other hand, the software, hardware and security protocols must be kept consistently up to date to avoid vulnerabilities. Furthermore, as we have already observed with scalability, strengthened on-prem security requires an increased amount of space and resources.

Many organizations will find that the customization options available with on-prem infrastructure go hand-in-hand with the greater scalability offered by cloud-based solutions. Moreover, a hybrid approach enables informed decision making that allows organizations to choose the best of both environments: the ability to fine-tune local resilience with robust fault tolerance. This ensures that AI systems meet business needs while protecting your data against external threats.

However, no matter how your company decides to implement its infrastructure, time is of the essence. The question your organization faces is: how can we deploy our AI infrastructure faster?

Enter: Cisco and Nvidia

Cisco and Nvidia are transforming the way companies deploy their cloud-based and on-premises AI infrastructure. With Cisco Validated Designs and NVIDIA AI Enterprise software, your organization will experience an increase in your AI performance output with strengthened security and scalability. This will allow you to build and deploy advanced AI and generative AI workloads faster than ever before.

Cisco Story—Cisco offers the scalability and security required to quickly deploy AI infrastructure. This is achieved through the Cisco Unified Computing System, as evidenced via Flexpod and Flashstack. However, the unifying feature of the Cisco x Nvidia partnership are Cisco Validated Designs (CVDs). They allow for the quick deployment and management of AI clusters at any scale through detailed blueprints that reduce deployment time by up to 60%. These step-by-step instructions offer the best practices for deploying AI infrastructure to simply the deployment process. By building on the success of proven capabilities and existing IT processes, CVDs bolster AI infrastructure without adding unnecessary risk or new operations silos.

Flexpod (On-Prem): One of the crucial benefits of the FlexPod Datacenter is your organization's ability to mainstream your AI infrastructure. By combining Cisco computing and networking with NetApp storage, organizations now have access to a scalable and secure foundation for AI workloads that allows them to unify silos of infrastructure and operations. Moreover, FlexPod seamlessly integrates with dynamic AI software stacks like as NVAIE, which allows companies to harness the power of NVIDIA GPUs for AI workloads. Whether it's generative AI inferencing or GPU-intensive applications, FlexPod ensures that your AI infrastructure deployment follows industry standards and proven configurations.

F**lashstack (Cloud):** The flexible, proven infrastructure of Flashstack AI reduces delivers global visibility, consistency and control. This modern AI architecture combines the Cisco UCS X-Series Modular System, the Cisco Data Center Networking as well as Pure Storage (which provides scalability for large training data sets and model serving). By supporting a wide range of AI frameworks, Flashstack allows companies to choose the GPU and CPU options that best fit their needs, increasing their AI infrastructure performance. Additionally, Flashstack for AI incorporates Cisco's and Pure Storage's best practices to prevent unauthorized access and protect your data.

Nvidia Story—Nvidia offers the high-performance you need to maximize your AI infrastructure. This is why, on the Nividia side, the unifying feature of the Cisco x Nvidia partnership are Nvidia graphic processing units (GPUs). Because they are much more powerful than traditional CPUs, GPUs not only integrate with Flashstack and Flashpod, but they also serve as the foundation for Nvidia's offerings. In short, the Cisco Validated Designs will allow you to deploy your AI infrastructure as quickly as possible, and Nvidia GPUs will power your generative AI.

GPUs: Because NVIDIA GPUs are designed for high-performance computing, they are ideal for AI workloads. This allows companies to reduce the overall amount of time it takes to train their AI models. And with the scalability of NVIDIA GPUs, companies can add additional units as their workloads increase, instead of purchasing units in bulk beforehand. Similarly to CVDs, these graphics processing units allow companies to fine-tune pre-trained models quickly, which significantly speeds up AI infrastructure deployment. This means that organizations are not forced to create their own training models, saving them both time and resources.

NVEAI: This end-to-end, cloud-native software platform streamlines the deployment of generative AI applications. Through hybrid-cloud providers, organizations are provided immediate access to Nvidia-powered infrastructure as well as generative AI for enterprises. This means that each AI application deployment is manageable for your organization. In addition, your organization can be assured that its data and intellectual property is protected via ongoing monitoring for system security vulnerabilities whether you deploy in the cloud or on-prem.

NIM: These microservices allow companies to deploy generative AI models on Nvidia GPUs across a myriad of infrastructures, including cloud environments and on-premises datacenters. This allows for 1) companies to optimize their AI infrastructure for maximum efficiency and cost-effectiveness, and 2) developers to integrate NIM into their existing infrastructure and applications without the need for extensive customization or specialized expertise. Critically, NIM also ensures scalability, security and reliability by empowering organizations to retain full control over their applications and data—no matter where these AI models are deployed.

1. Pillars + Proposed Content (14 potential pieces)

- Case for On-Prem & Hybrid AI Infrastructure (2 pieces)
  - FAQ on On-Prem & Hybrid AI
    - Traditional enterprise on-premises infrastructure wasn't designed to handle Generative AI demands; the cloud doesn't provide the same control or security as on-prem. Organizations require both for maximum ROI.
  - 10 Reasons to Invest in On-Prem
    - E.g., control, security, better performance, scalability, etc.
- Considerations for Building On-Prem & Hybrid AI Infrastructure (2 pieces)
  - Webinar(?)
    - Of these five pillars, this is the one that benefits most from an interactive expert, given that many of these considerations (e.g., talent, support) are company-specific and not Cisco/Nvidia-specific. Which means there isn't truly an explicit tie-in but something more interactive could get us to scalability, etc.
  - Checklist Infographic
    - A list of everything that an organization needs to consider before investing in on-prem or hybrid AI infrastructure.
- Barriers Faced when Building On-Prem & Hybrid AI Infrastructure (1 piece)
  - Barriers/Solutions One-Pager
    - There are several key barriers (e.g., software frameworks, scalability, robust security measures) that can be solved with Cisco and Nvidia offerings—this is a natural link from issue to product.
- Cisco/Nvidia/ConRes as an Infrastructure Solution (4 pieces)
  - Cisco On-Prem/Hybrid Infographic
    - Outlining Flexpod and Flashstack as the two key offerings for Cisco's on-prem and hybrid AI infrastructure support.
  - Cisco Product-Specific One-Pager
    - Cisco Validated Designs as the focus: Cisco reduces deployment times and bolsters AI infrastructure without adding unnecessary risks or operations silos.
  - Nvidia On-Prem/Hybrid Infographic
    - Outlining NVEAI and NIM as the two key Nvidia offerings for AI deployment.
  - Nvidia Product-Specific One-Pager
    - Nvidia GPUs as the focus: designed for high-performance computing, they are ideal for AI workloads and provide a boost to your organizational functioning.
- Cisco/Nvidia/ConRes as a Competitive Advantage (2 pieces)
  - Strength-of-Partnership One-Pager
    - Discussing how the combination of Ethernet networking/GPU technology as well as the Nvidia Tensor Core GPUs in Cisco Servers allows for both speed of deployment (Cisco) and powerful generative AI (Nvidia).
  - Cisco & Nvidia Integration Infographic
    - Highlighting how each respective offering integrates with the other to provide a seamless generative AI experience without the need for another vendor.
- Timeline Infographic (?)
    - Assuming that focusing on the "We have a year or less to implement AI strategy" statistic works, this infographic would be a general overview of the timeline to implementation, starting with AI infrastructure deployment.
- Timeline ConRes Expert Video (?)

        - Following from the infographic, it would be helpful to highlight the role that ConRes can play in both AI strategy as well as AI infrastructure deployment. This not only positions ConRes as a thought leader but also as the go-to source for AI infrastructure for these companies.
- Short-Hitting Video/Audio [Email Campaign]
        - My preference is short audio over video; however, either would work. Instead of a full interview, we take two minute-or-less clips from a larger conversation addressing these five pillars and centering the email messaging around them.

2. Overarching Themes

Powerful Generative AI in Less than a Year

The Quick and Dynamic Path to Transformational AI

Rapidly Scale Your AI Capabilities with a Robust On-Prem Infrastructure

3. Essential Story

According to the Cisco AI Readiness Index, more than 60% of companies recognize that they have just one year or less to implement their AI strategy before incurring significant negative business impacts from falling behind. Your organization can meet this deadline, but it requires **quickly deploying powerful On-Prem/Hybrid AI Infrastructure**. ConRes has the Cisco/Nvidia offerings to get you there.

Pillars as Bullets:

- There are enormous consequences for waiting to shift your organization toward AI.
- Your pre-existing on-prem or cloud infrastructure can't do the job.
- Cisco/Nvidia expedites your timeline with quickly deployable on-prem/hybrid AI infrastructure.
- Through offerings like CVDs and GPUs, your generative AI is both powerful and scalable.
- ConRes can help you reach your AI goals, beginning with infrastructure implementation.

4. Landing Page:

ConRes with Cisco & Nvidia: Expedite Your AI Timeline
ALT: Leverage the Full Power of Generative AI
ALT: Invest in Industry-Leading AI Infrastructure

(Intro description of story/themes)

Timeline Expert Video
Building On-Prem AI Infrastructure for Your Generative AI Goals

Timeline Infographic

The ConRes Advantage	Why Choose On-Prem/ Hybrid AI Infrastructure	What to Consider	Overcoming Barriers	On-Prem/Hybrid Infrastructure Solution	The Cisco & Nvidia Partnership

ConRes Understands Implementation Challenges for Generative AI
(Intro)

1. Performance
2. Scalability
3. Security
4. Cost-Effectiveness

Why Choose On-Prem/Hybrid AI Infrastructure?
(Intro)

- FAQ on On-Prem & Hybrid AI
- 10 Reasons to Invest in On-Prem

What to Consider when Investing in AI Infrastructure
(Intro)

- Webinar (?)
- Checklist Infographic

Barriers to On-Prem/Hybrid AI Infrastructure
(Intro)

- Barriers/Solutions One-Pager

Quickly Deployable, Powerful AI Infrastructure
(Intro)

- Cisco On-Prem/Hybrid Infographic
- Cisco Product-Specific One-Pager
- Nvidia/Hybrid Infographic
- Nvidia Product-Specific One-Pager

The ConRes + Cisco & Nvidia Advantage
(Intro)

- Barriers Strength-of-Partnership One-Pager
- Cisco & Nvidia Integration Infographic

The Clock Starts Now (CTA)