

Arista veröffentlicht entscheidende Neuentwicklungen für KI-Netzwerke

EOS Smart AI Suite liefert höchste Performance für KI-Workloads

Santa Clara (USA), 13.03.2025 – Arista Networks (NYSE: ANET), ein führender Innovator für Cloud- und [KI-Netzwerke](#), präsentiert entscheidende Funktionen zur Steigerung der Performance und Effizienz von KI-Clustern. Cluster Load Balancing (CLB) in [Arista EOS®](#) maximiert die Performance von KI-Workloads durch konsistente und latenzarme Netzwerkdatenströme, und [Arista CloudVision®](#) Universal Network Observability™ ([CV UNO™](#)) bietet jetzt eine KI-Job-zentrierte Observability für eine verbesserte Fehlerbehebung und eine schnelle Problemerkennung, um die Zuverlässigkeit der Job-Abwicklung auch im großen Umfang sicherzustellen.

Intelligentes KI-Networking

Die Arista EOS Smart AI Suite wurde für die Stabilität und den Schutz von KI-Anwendungen entwickelt und unterstützt KI-Cluster mit der Neuerung „Cluster Load Balancing“ - einem Ethernet basierten KI-Load-Balancing, die auf RDMA- Queue-Paaren basiert und eine hohe Bandbreitennutzung zwischen Spines und Leaves ermöglicht. KI-Cluster haben in der Regel eine geringe Anzahl von Datenströmen mit großen Bandbreiten.

Herkömmliche Load-Balancing-Methoden sind für KI-Workloads daher oft ineffizient und führen zu einer ungleichmäßigen Verteilung des Traffics und einer erhöhten Tail-Latenz. CLB behebt dieses Hindernis durch die Verwendung von RDMA-fähiger Flow-Platzierung, um eine einheitlich hohe Leistung für alle Datenströme zu gewährleisten und gleichzeitig die Tail-Latenz gering zu halten. CLB verfolgt einen ganzheitlichen Ansatz und optimiert den Datenfluss in beide Richtungen, von Leaf zu Spine und von Spine zu Leaf, so dass eine gleichmäßige Auslastung und gleichbleibend niedrige Latenzzeiten erreicht werden.

„Da Oracle seine KI-Infrastruktur mit Arista-Switches weiter ausbaut, benötigen wir hochentwickelte Load Balancing-Methoden, um Flow-Probleme zu vermeiden und den Durchsatz in ML-Netzwerken zu erhöhen“, sagt Jag Brar, Vice President & Distinguished Engineer, Oracle Cloud Infrastructure. „Das Cluster Load Balancing Feature von Arista hilft dabei.“

Ganzheitliche KI-Observability

CV UNO, die KI-gesteuerte 3600 Network Observability Plattform auf Basis von Arista AVA™, bietet nahtlose, durchgängige KI-Job-Transparenz, indem sie Netzwerk-, System- und KI-Job-Daten innerhalb des Arista Network Data Lake ([NetDL™](#)) vereint. EOS NetDL Streamer ist ein Echtzeit-Telemetrie-Framework, das kontinuierlich granulare Netzwerkdaten von Arista Switches in NetDL streamt.

Im Gegensatz zu herkömmlichem SNMP-Polling, das auf periodischen Abfragen beruht und kritische Updates verpassen kann, bietet der EOS NetDL Streamer ereignisgesteuerte Einblicke in die Netzwerkleistung mit niedriger Latenz und hoher Frequenz. Dies ist der Schlüssel zur Optimierung großer KI-Trainings- und Inferencing-Infrastrukturen.

Der Streamer wurde für KI-Beschleunigercluster entwickelt und ermöglicht eine schnellere Analyse der Leistungsdaten, eine genauere Lokalisierung von Problemen und eine schnellere Behebung von Problemen, so dass die Zeit für die Fertigstellung von Jobs minimiert wird.

Zu den wichtigsten Vorteilen gehören:

- **AI-Job-Monitoring** - Ermöglicht einen umfassenden Überblick über den Zustand von AI-Jobs, einschließlich Job- Fertigstellungszeiten, Überlastungsindikatoren (ECN-markierte Pakete, PFC- Pausenframes, Paketfehler) und Puffer-/Link-Auslastung für Echtzeit-Einblicke.
- **Deep-Dive-Analytik** - Entdeckt kritische jobspezifische Informationen durch die Analyse von Netzwerkgeräten, Server-NICs (z. B. PFC-Outs-of-Sync-Ereignisse, RDMA-Fehler, PCIe-Fatal-Fehler) und zugehörigen Flows, um Performance-Engpässe präzise zu identifizieren.
- **Flow-Visualisierung** - Nutzt die Leistungsfähigkeit des CV-Topologie-Mappings, um in Echtzeit einen unmittelbaren Einblick in die Abläufe von KI-Jobs mit einer Granularität von Mikrosekunden zu erhalten - und beschleunigt so die Erkennung und Lösung von Problemen.
- **Proaktive Fehlerbehebung** - Erkennt Anomalien frühzeitig und korreliert die Netzwerk- und Rechenleistung innerhalb von NetDL - für eine unterbrechungsfreie, hocheffiziente KI-Workload-Ausführung.

Arista AI-Centers gesteuert durch AVA

Die Etherlink™ KI-Plattformen von Arista liefern extrem leistungsstarke, standardbasierte Ethernet-Systeme für KI-Netzwerke der nächsten Generation. Mit festen, modularen und verteilten 800G/400G-Plattformen, die mit dem Ultra Ethernet Consortium (UEC) kompatibel sind, skaliert Etherlink von kleinen KI-Clustern bis hin zu komplexen Implementierungen mit über 100.000 Beschleunigern. Arista bietet den AI Analyzer, powered by Arista AVA, der hochauflösende Trafficdaten in 100-Mikrosekunden-Intervallen liefert und so eine präzise Performance-Optimierung und Fehlersuche ermöglicht.

Netzwerkadministratoren können so die Leistung optimieren, Probleme schnell beheben und fundierte Entscheidungen für KI-gesteuerte Netzwerke treffen. Arista AVA steuert auch einen Remote EOS AI Agent, der Telemetriedaten von SuperNICs oder Servern an NetDL streamt und so eine nahtlose Netzwerküberwachung, Fehlersuche und QoS-Konsistenz über den gesamten Stack hinweg gewährleistet.

Verfügbarkeit

- CLB
- Verfügbar ab sofort für die Plattformen 7260X3, 7280R3, 7500R3 und 7800R3.
- Unterstützung für die Plattformen 7060X6 und 7060X5 geplant für Q2 2025
- Unterstützung für 7800R4 geplant für das 2. Halbjahr 2025
- CV UNO ist ab sofort verfügbar. Die Verbesserungen der Observability für AI befinden sich in aktiven Kundentests, die allgemeine Verfügbarkeit ist für Q2 2025 geplant.

Hier finden Sie Reaktionen aus der Branche zu dieser Ankündigung:

<https://solutions.arista.com/industry-support-03122025>

Melden Sie sich hier für ein Webinar am 10. April 2025 an, um mehr über die EOS Smart AI-Suite zu erfahren: <https://events.arista.com/ai-eos-cloudvision-webinar>

Hier finden Sie einen Blog-Beitrag von Praful Bhaidasna, Director Product Management bei Arista: <https://blogs.arista.com/blog/modern-ai>

Über Arista

Arista Networks ist ein branchenführender Anbieter von datengesteuerten Client-to-Cloud-Netzwerken für große Rechenzentrums-, Campus- und Routing-Umgebungen. Die mehrfach ausgezeichneten Plattformen von Arista bieten Verfügbarkeit, Agilität, Automatisierung, Analyse und Sicherheit durch CloudVision® und das fortschrittliche Netzwerkbetriebssystem Arista EOS®.

Weitere Informationen finden Sie unter: www.arista.com.

ARISTA, EOS, CloudVision, NetDL und AVA gehören zu den eingetragenen und nicht eingetragenen Marken von Arista Networks, Inc. in allen Ländern der Welt. Andere Firmennamen oder Produktnamen können Marken ihrer jeweiligen Eigentümer sein. Weitere Informationen und Materialien finden Sie unter www.arista.com.

Diese Pressemitteilung enthält zukunftsgerichtete Aussagen, einschließlich, aber nicht beschränkt auf, Aussagen zu Kosteneinsparungen, Leistung, Funktionen und Sicherheit. Alle Aussagen, die sich nicht auf historische Fakten beziehen, sind Aussagen, die als zukunftsgerichtete Aussagen betrachtet werden können. Zukunftsgerichtete Aussagen unterliegen Risiken und Ungewissheiten, die dazu führen können, dass die tatsächliche Leistung oder die Ergebnisse wesentlich von den in den zukunftsgerichteten Aussagen zum Ausdruck gebrachten abweichen, einschließlich des raschen Technologie- und Marktwandels, der Kundenanforderungen und der Industriestandards sowie anderer Risiken, die in unseren bei der SEC eingereichten Unterlagen aufgeführt sind, die auf der Website von Arista unter www.arista.com und auf der Website der SEC unter www.sec.gov verfügbar sind. Arista lehnt jede Verpflichtung ab, zukunftsgerichtete Aussagen öffentlich zu aktualisieren oder zu revidieren, um Ereignisse oder Umstände widerzuspiegeln, die nach dem Datum, an dem sie gemacht wurden, eintreten.

Pressekontakt

Amanda Jaramillo
Corporate Communications
Tel: (408) 547-5798
amanda@arista.com

Investorenkontakt

Rudolph Araujo
Rod Hall
Investor Relations
Tel: (408) 547-5885
ir@arista.com