



Assistant IA Ask Me

Information Security for Workload Factory

NetApp
September 16, 2026

Sommaire

- Assistant IA Ask Me 1
 - Architecture de sécurité Ask Me pour NetApp Workload Factory 1
 - Architecture RAG dans Ask Me 1
 - Sécurité multicouche des invites 1
 - Isolation et confidentialité des données du modèle 1
 - Utilisation d'outils et sécurité agentique 2
 - Contrôle d'accès Ask Me pour NetApp Workload Factory 2
 - Authentification 2
 - Protection des identifiants et gestion des secrets 2
 - Sécurité des infrastructures 2
 - Modes d'exécution Ask Me pour NetApp Workload Factory 3
 - Garde-fous pour l'exécution des outils 3
 - Contrôles de sortie 3
 - Demandez-moi la confidentialité et la disponibilité de NetApp Workload Factory 3
 - Gestion des données et confidentialité 4
 - Contrôles de disponibilité et d'isolation 4

Assistant IA Ask Me

Architecture de sécurité Ask Me pour NetApp Workload Factory

Ask Me est un assistant génératif basé sur l'IA, intégré à NetApp Workload Factory. Il offre une assistance conversationnelle en temps réel pour Amazon FSx for NetApp ONTAP et les sujets Workload Factory. Les réponses s'appuient sur une documentation NetApp vérifiée, et Ask Me ne peut pas accéder à Internet ni utiliser les données client pour entraîner le modèle. Plusieurs couches de sécurité bloquent les requêtes malveillantes, limitent les réponses aux domaines pris en charge et empêchent que les saisies utilisateur ne remplacent les instructions système. Lorsque Ask Me utilise des outils, les limites d'autorisation, l'application de la lecture seule pour les requêtes, un environnement de test éphémère et la journalisation des audits permettent de restreindre et de surveiller l'exécution.

Vous pouvez désactiver Ask Me à tout moment via les paramètres **Administration** de Workload Factory, ce qui désactive l'assistant pour votre compte utilisateur. ["En savoir plus sur la gestion des fonctionnalités GenAI."](#)

Architecture RAG dans Ask Me

Ask Me utilise la génération augmentée par récupération (RAG).

- Base de connaissances organisée : les réponses s'appuient sur un corpus géré de documentation NetApp vérifiée et publiée.
- Évaluation et filtrage de la recherche : seul le contenu suffisamment pertinent est inclus dans le contexte du modèle.
- Aucun accès à Internet : le modèle ne peut pas récupérer, parcourir ou extraire de contenu externe.
- Attribution des sources : les réponses incluent des références aux documents sources utilisés.

Sécurité multicouche des invites

- Couche 1 : Le classificateur de sécurité (LLM-as-a-judge) bloque les requêtes malveillantes (injection de prompt, élévation de privilèges, exfiltration de données, ingénierie sociale, violations de politiques).

Le système enregistre les requêtes bloquées, et le modèle génératif ne les voit jamais.

- Couche 2 : Le renforcement des invites système empêche les entrées de l'utilisateur de prendre le pas sur les instructions du système.
- Couche 3 : La portée du domaine limite les réponses aux sujets FSx ONTAP et NetApp Workload Factory.
- Couche 4 : La gouvernance de l'utilisation des outils valide les paramètres et exécute les requêtes dans un environnement sécurisé ; le modèle ne peut pas exécuter d'opérations arbitraires.

Isolation et confidentialité des données du modèle

- Aucun entraînement du modèle sur les données client
- Isolation au niveau des requêtes (aucune fuite de données entre locataires)

- Contexte de conversation à l'échelle de la session avec un historique limité
- Aucune mémoire de modèle persistante
- Infrastructure d'inférence gérée

Utilisation d'outils et sécurité agentique

- Outils à autorisation limitée (limites des permissions utilisateur appliquées)
- Délais d'exécution stricts des outils
- Bac à sable éphémère, limité à la session, pour les données récupérées
- Application des requêtes en lecture seule avec liste d'autorisation
- Auditabilité des appels d'outils avec nettoyage des paramètres sensibles

Contrôle d'accès Ask Me pour NetApp Workload Factory

Dans NetApp Workload Factory, la sécurité d'accès Ask Me repose sur des sessions authentifiées, la responsabilité au niveau de l'utilisateur et la gestion protégée des secrets lors des opérations d'exécution. Les contrôles d'authentification valident chaque session et associent les interactions à un utilisateur spécifique. Les informations d'identification sensibles sont récupérées depuis un coffre-fort de secrets géré lors de l'exécution et supprimées des journaux. Le service applique également des contrôles d'infrastructure, notamment l'exécution de conteneurs sans privilèges root et une liste d'autorisation CORS restreinte.

Authentification

- Authentification JWT signée avec validation cryptographique (y compris les vérifications de l'audience, de l'émetteur et de l'algorithme telles que RS256)
- Authentification appliquée par le middleware à la frontière du service
- Définition du périmètre d'identité de l'utilisateur afin que les interactions soient attribuables à un utilisateur spécifique

Protection des identifiants et gestion des secrets

- Stockage sécurisé des identifiants : les identifiants sensibles utilisés par le service sont stockés dans un coffre-fort de secrets géré et récupérés lors de l'exécution. Aucun secret n'est stocké dans le code de l'application, les fichiers de configuration ou les images de conteneur.
- Nettoyage des journaux : les valeurs sensibles (identifiants, jetons, mots de passe, clés d'accès, certificats) sont supprimées des journaux avant leur enregistrement.

Sécurité des infrastructures

- Exécution de conteneur non root
- CORS restreint à une liste blanche explicite de domaines de confiance NetApp

Modes d'exécution Ask Me pour NetApp Workload Factory

NetApp Workload Factory propose plusieurs modes d'exécution pour Ask Me, chacun présentant des caractéristiques de sécurité et de confidentialité différentes. Lorsqu'Ask Me utilise des intégrations d'outils, par exemple pour interroger les ressources Workload Factory du client, l'exécution des outils est contrôlée par des dispositifs de protection multicouches. L'exécution avec outils s'effectue dans le cadre des autorisations de l'utilisateur authentifié et applique des limites à la durée et à la portée de chaque opération. Les données récupérées restent dans un environnement de session éphémère qui bloque tout accès externe et sont détruites à la fin de la session. Des contrôles indépendants des invites et du code imposent des requêtes en lecture seule, tandis que les journaux d'audit enregistrent l'activité des outils après suppression des valeurs sensibles.

Garde-fous pour l'exécution des outils

- Les outils à portée d'autorisation fonctionnent dans le cadre des autorisations de l'utilisateur authentifié.
- Le délai d'expiration de l'exécution de l'outil limite les opérations de longue durée.
- Le bac à sable de données éphémères stocke les données récupérées par l'outil dans le cadre de la session, bloque l'accès externe, les E/S de fichiers, les appels réseau et le chargement d'extensions au niveau du moteur, et est détruit à la fin de la session.
- L'application du contrôle des requêtes en lecture seule valide les requêtes générées via une liste d'autorisation stricte.
- La sécurité appliquée par invite et la sécurité appliquée par code sont appliquées indépendamment.
- L'audit des appels d'outils enregistre le nom de l'outil, ses paramètres, les horodatages et l'état du résultat, avec un filtrage des valeurs sensibles.

Contrôles de sortie

- Application des limites de jetons
- Sortie déterministe pour les opérations critiques de classification/sécurité (température 0)
- Diffusion en continu avec arrêt anticipé et nettoyage

Demandez-moi la confidentialité et la disponibilité de NetApp Workload Factory

Dans NetApp Workload Factory, les contrôles de confidentialité Ask Me limitent l'exposition des données, maintiennent l'isolation du traitement des sessions et préservent les frontières de confidentialité pour les interactions utilisateur. Vos entrées sont traitées par un service d'IA géré qui ne les utilise pas pour entraîner ou améliorer les modèles de base. Les données opérationnelles récupérées restent dans un stockage éphémère en mémoire pendant la durée de la session et ne sont ni conservées ni partagées. Une chaîne de repli multi-modèles répartie sur plusieurs régions cloud garantit la disponibilité lorsque le modèle principal est limité ou indisponible, tandis que les

mêmes contrôles de sécurité s'appliquent à tous les modèles de repli.

Gestion des données et confidentialité

- Données utilisateur dans les invites : traitées par un service d'IA géré qui n'utilise pas vos entrées pour entraîner ou améliorer les modèles de base.
- Contenu de la base de connaissances : documentation NetApp sélectionnée et publiée uniquement ; vos données ne sont pas incluses.
- Données opérationnelles : récupérées uniquement lorsqu'un utilisateur interroge ses propres ressources ; stockées de manière éphémère en mémoire pendant la durée de la session et non conservées ni partagées.
- Journalisation des audits : les métadonnées des requêtes (identifiant utilisateur anonymisé, horodatages, durées) sont enregistrées, avec les champs sensibles supprimés.
- Données de retour d'information : stockées séparément et liées à un identifiant de requête, non à des informations permettant d'identifier une personne au-delà de l'identifiant utilisateur anonymisé.

Contrôles de disponibilité et d'isolation

Ask Me utilise une chaîne de repli multi-modèles répartie sur plusieurs régions cloud.

Si le modèle principal est limité ou indisponible, le système peut basculer vers des modèles alternatifs.

Tous les modèles sont soumis aux mêmes contrôles de sécurité, au même renforcement du système de sécurité et aux mêmes garanties d'isolation.

Informations sur le copyright

Copyright © 2026 NetApp, Inc. Tous droits réservés. Imprimé aux États-Unis. Aucune partie de ce document protégé par copyright ne peut être reproduite sous quelque forme que ce soit ou selon quelque méthode que ce soit (graphique, électronique ou mécanique, notamment par photocopie, enregistrement ou stockage dans un système de récupération électronique) sans l'autorisation écrite préalable du détenteur du droit de copyright.

Les logiciels dérivés des éléments NetApp protégés par copyright sont soumis à la licence et à l'avis de non-responsabilité suivants :

CE LOGICIEL EST FOURNI PAR NETAPP « EN L'ÉTAT » ET SANS GARANTIES EXPRESSES OU TACITES, Y COMPRIS LES GARANTIES TACITES DE QUALITÉ MARCHANDE ET D'ADÉQUATION À UN USAGE PARTICULIER, QUI SONT EXCLUES PAR LES PRÉSENTES. EN AUCUN CAS NETAPP NE SERA TENU POUR RESPONSABLE DE DOMMAGES DIRECTS, INDIRECTS, ACCESSOIRES, PARTICULIERS OU EXEMPLAIRES (Y COMPRIS L'ACHAT DE BIENS ET DE SERVICES DE SUBSTITUTION, LA PERTE DE JOUISSANCE, DE DONNÉES OU DE PROFITS, OU L'INTERRUPTION D'ACTIVITÉ), QUELLES QU'EN SOIENT LA CAUSE ET LA DOCTRINE DE RESPONSABILITÉ, QU'IL S'AGISSE DE RESPONSABILITÉ CONTRACTUELLE, STRICTE OU DÉLICTELLE (Y COMPRIS LA NÉGLIGENCE OU AUTRE) DÉCOULANT DE L'UTILISATION DE CE LOGICIEL, MÊME SI LA SOCIÉTÉ A ÉTÉ INFORMÉE DE LA POSSIBILITÉ DE TELS DOMMAGES.

NetApp se réserve le droit de modifier les produits décrits dans le présent document à tout moment et sans préavis. NetApp décline toute responsabilité découlant de l'utilisation des produits décrits dans le présent document, sauf accord explicite écrit de NetApp. L'utilisation ou l'achat de ce produit ne concède pas de licence dans le cadre de droits de brevet, de droits de marque commerciale ou de tout autre droit de propriété intellectuelle de NetApp.

Le produit décrit dans ce manuel peut être protégé par un ou plusieurs brevets américains, étrangers ou par une demande en attente.

LÉGENDE DE RESTRICTION DES DROITS : L'utilisation, la duplication ou la divulgation par le gouvernement sont sujettes aux restrictions énoncées dans le sous-paragraphe (b)(3) de la clause Rights in Technical Data-Noncommercial Items du DFARS 252.227-7013 (février 2014) et du FAR 52.227-19 (décembre 2007).

Les données contenues dans les présentes se rapportent à un produit et/ou service commercial (tel que défini par la clause FAR 2.101). Il s'agit de données propriétaires de NetApp, Inc. Toutes les données techniques et tous les logiciels fournis par NetApp en vertu du présent Accord sont à caractère commercial et ont été exclusivement développés à l'aide de fonds privés. Le gouvernement des États-Unis dispose d'une licence limitée irrévocable, non exclusive, non cessible, non transférable et mondiale. Cette licence lui permet d'utiliser uniquement les données relatives au contrat du gouvernement des États-Unis d'après lequel les données lui ont été fournies ou celles qui sont nécessaires à son exécution. Sauf dispositions contraires énoncées dans les présentes, l'utilisation, la divulgation, la reproduction, la modification, l'exécution, l'affichage des données sont interdits sans avoir obtenu le consentement écrit préalable de NetApp, Inc. Les droits de licences du Département de la Défense du gouvernement des États-Unis se limitent aux droits identifiés par la clause 252.227-7015(b) du DFARS (février 2014).

Informations sur les marques commerciales

NETAPP, le logo NETAPP et les marques citées sur le site <http://www.netapp.com/TM> sont des marques déposées ou des marques commerciales de NetApp, Inc. Les autres noms de marques et de produits sont des marques commerciales de leurs propriétaires respectifs.