v.2, n.1, 2026

REVISTA CIENTÍFICASOL21
STANDARD OPEN LITERATURE - ISSN 3086-089X

SUPERFICIE DE ATAQUE NO INVENTARIADA Y DENEGACIÓN DE SERVICIO EN LA CAPA DE APLICACIÓN: UN MODELO DE EVALUACIÓN DE LA EXPOSICIÓN DE APIs

SUPERFÍCIE DE ATAQUE NÃO INVENTARIADA E NEGAÇÃO DE SERVIÇO NA CAMADA DE APLICAÇÃO: UM MODELO DE AVALIAÇÃO DA EXPOSIÇÃO DE APIs

UNINVENTORIED ATTACK SURFACE AND APPLICATION-LAYER DENIAL OF SERVICE: AN ASSESSMENT MODEL FOR API EXPOSURE

Marcelo de Melo Marcon[1], Carlos Eduardo Costa Oliveira[2]

e2136
https://doi.org/10.67553/sol21.v.2i1.36

PUBLICADO: 10/2026

RESUMEN

La migración de las arquitecturas digitales hacia modelos orientados a interfaces de programación de aplicaciones (APIs) ha desplazado el vector de denegación de servicio desde la capa de red hacia la capa de aplicación, donde el ataque puede producir efecto con un volumen comparativamente bajo, al explotar la asimetría entre el costo de emitir una petición y el de procesarla. El objetivo es proponer un modelo de evaluación de la exposición de las APIs a esa clase de ataque, aplicable a organizaciones cuya superficie no esté íntegramente inventariada. Se adopta la Design Science Research (DSR), a partir de un levantamiento documental intencional cuyo protocolo de búsqueda, bases, descriptores, período y criterios se explicitan. Los resultados comprenden una taxonomía de cinco clases de agotamiento de recursos, la caracterización de las Shadow APIs y de los endpoints zombis como deuda de superficie de ataque, cuatro supuestos que limitan a las plataformas de protección, y el modelo propuesto, con cinco dimensiones, cuatro niveles ordinales, cinco métricas y una demostración analítica.

PALABRAS CLAVE: Denegación de servicio. Seguridad de APIs. Capa de aplicación. Superficie de ataque. Gobernanza de seguridad de la información. DDoS.

RESUMO

A migração das arquiteturas digitais para modelos orientados a interfaces de programação de aplicações (APIs) deslocou o vetor de negação de serviço da camada de rede para a camada de aplicação, na qual o ataque pode produzir efeito com volume comparativamente baixo, por explorar a assimetria entre o custo de emitir uma requisição e o de processá-la. O objetivo é propor um modelo de avaliação da exposição de APIs a essa classe de ataque, aplicável a organizações cuja superfície não esteja integralmente inventariada. Adota-se a Design Science Research (DSR), a partir de levantamento documental intencional cujo protocolo de busca, bases, descritores, período e critérios são explicitados. Os resultados compreendem uma taxonomia de cinco classes de exaustão de recursos; a caracterização de Shadow APIs e de endpoints zumbis como dívida de

[1] Doutor Honoris Causa em Ciências Exatas pela World Christian University (WCU), Mestre em Direção Estratégica em Tecnologias da Informação pela Universidad Europea del Atlántico, pós-graduado em MBIS – Segurança da Informação e em Engenharia de Redes e Sistemas de Telecomunicações. Possui mais de 20 anos de experiência profissional em Tecnologia da Informação e Comunicação (TIC), com atuação em ambientes críticos nos setores de saúde, governo e telecomunicações. Sua experiência abrange Segurança da Informação, Cibersegurança, Proteção de Dados, Gestão de Riscos, Datacenters, Virtualização, Linux, Redes, Governança, Continuidade de Negócios e Melhoria de Processos. É instrutor da Escola Superior de Guerra – ADESG/Campinas, membro do Comitê Avançado e Altos Estudos de Tecnologia e Inovação da Digital Law Academy, Diretor do Comitê de Segurança da APDADOS e Imortal da Academia Brasileira de Ciências, Artes, História e Literatura (ABRASCI).

[2] Fundador da vBold LLC (EUA) e vBold Ltda (Brasil), empresas que prestam serviços de inovação e segurança em tecnologia a clientes nacionais e internacionais. Graduação em Direito, pós-graduação em Computação em Nuvem e cursando mestrado em Negócios Globais nos EUA. Autor do livro Navegando Pelos Desafios da Segurança da Informação, membro da Mensa Internacional, Vice-Presidente da APDADOS e Imortal da ABRASCI, além de ter recebido o título Honoris Causa em Ciências Exatas pela WCU. Reconhecido por sua liderança e impacto, alia expertise em tecnologia à missão de transformar vidas e gerar oportunidades.

Este artículo se publica en acceso abierto (Open Access) bajo licencia Creative Commons Attribution 4.0 Internacional (CC-BY), que permite uso, distribución y reproducción irrestrictos en cualquier medio, siempre que el autor original y la fuente sean acreditados.

REVISTA CIENTÍFICASOL21
STANDARD OPEN LITERATURE - ISSN 3086-089X

superfície de ataque; quatro pressupostos que limitam as plataformas de proteção; e o modelo proposto, em cinco dimensões e quatro níveis ordinais, com cinco métricas, demonstração analítica e mapeamento normativo. Argumenta-se que a completude do inventário é fator determinante da cobertura efetiva dos demais controles.

PALAVRAS-CHAVE: Negação de serviço. Segurança de APIs. Camada de aplicação. Superfície de ataque. Governança de segurança da informação. DDoS.

ABSTRACT
The migration of digital architectures toward application programming interface (API)-oriented models has shifted the denial-of-service vector from the network layer to the application layer, where attacks can achieve effect with comparatively low volume by exploiting the asymmetry between the cost of issuing a request and the cost of processing it. This article proposes a model for assessing API exposure to this class of attack, applicable to organizations whose attack surface is not fully inventoried. Design Science Research (DSR) is adopted, drawing on a purposive documentary survey whose search protocol, databases, descriptors, period, and inclusion and exclusion criteria are made explicit. The results comprise a taxonomy of five classes of resource exhaustion; the characterization of shadow APIs and zombie endpoints as attack-surface debt; four assumptions that structurally limit protection platforms; and the proposed model, with five dimensions, four proposed ordinal levels, five metrics, an analytical demonstration, and a normative mapping.

KEYWORDS: Denial of service. API security. Application layer. Attack surface. Information security governance. DDoS.

INTRODUCCIÓN

A lo largo de la última década, la arquitectura de las aplicaciones digitales se reorganizó en torno a interfaces de programación de aplicaciones (APIs): la superficie expuesta dejó de ser el conjunto de páginas navegables por un usuario humano y pasó a ser el de endpoints accesibles por un cliente programático. Según Imperva (2026), la actividad automatizada pasó a representar más del 53% del tráfico de la web en 2025, con el 27% de los ataques automatizados dirigidos a endpoints de API.

En paralelo, el vector de denegación de servicio distribuida (DDoS, distributed denial of service) se desplazó. Cloudflare (2026) reportó 29,64 billones de peticiones por protocolo de transferencia de hipertexto (HTTP) asociadas a denegación de servicio en el primer semestre de 2026, y el 90,60% de los ataques concluyó en menos de diez minutos, lo que vuelve inviables las respuestas dependientes de intervención humana. La literatura documenta, además, la migración del ataque hacia la capa de aplicación, donde puede producir efecto con un volumen comparativamente bajo (Praseed y Thilagam, 2019; Deressu et al., 2026).

El mercado respondió con la convergencia de las defensas en plataformas de protección de aplicaciones y APIs, categoría de mercado formulada bajo la sigla WAAP, de web application and API protection. Esa convergencia descansa sobre un supuesto rara vez explicitado: que la organización sabe qué endpoints existen, por dónde se enrutan y cuál es el costo computacional normal de cada uno. Es ese supuesto el que invalidan las Shadow APIs y los endpoints zombis:

Este artículo se publica en acceso abierto (Open Access) bajo licencia Creative Commons Attribution 4.0 Internacional (CC-BY), que permite uso, distribución y reproducción irrestrictos en cualquier medio, siempre que el autor original y la fuente sean acreditados.

rutas de depuración olvidadas en producción y versiones obsoletas pero nunca desactivadas siguen siendo alcanzables y fuera de toda política (OWASP, 2023).

Formulamos la proposición que orienta el trabajo: en entornos orientados a API, la denegación de servicio en la capa de aplicación no es primariamente un problema de caudal, sino de asimetría entre el costo de emitir una petición y el de procesarla, y esa asimetría está gobernada por la completitud del inventario. El adversario no necesita poder volumétrico; necesita localizar el endpoint más barato de alcanzar que dispara la operación más costosa en el backend. Los endpoints no inventariados reúnen, por construcción, las condiciones que elevan esa asimetría, aunque este trabajo no produce evidencia sobre su frecuencia relativa como blanco efectivo.

El objetivo general es proponer un modelo de evaluación de esa exposición, orientado a la gobernanza e independiente de la tecnología de detección; los específicos son construir la taxonomía de las clases aplicables a APIs, caracterizar las Shadow APIs y los endpoints zombis, identificar los supuestos que limitan a las plataformas, proponer el modelo y mapearlo contra referenciales normativos. La investigación se justifica porque, en el corpus examinado, la producción se concentra en la detección: Deressu et al. (2026) sistematizan 47 estudios primarios en ese eje, y Tanveer et al. (2025) organizan las vulnerabilidades de APIs REST sin contemplar el consumo de recursos.

1 - MARCO TEÓRICO

1.1 La capa de aplicación y la denegación de servicio

El modelo de referencia para interconexión de sistemas abiertos (OSI, open systems interconnection), de Zimmermann (1980), estratifica la comunicación en siete capas, de las cuales la séptima corresponde a la aplicación; las APIs basadas en transferencia de estado representacional (REST), en GraphQL o en gRPC operan en esa franja, sobre el protocolo de transferencia de hipertexto (HTTP). El modelo describe la interoperabilidad de protocolos, y no la semántica de negocio: una petición que consulta un saldo y otra que dispara la conciliación de un extracto anual son indistinguibles en el nivel protocolar, aunque difieran en órdenes de magnitud en el consumo de procesamiento.

Mirkovic y Reiher (2004) organizaron los ataques distribuidos según el grado de automatización, la dinámica de tasa y el impacto, y Zargar et al. (2013) sistematizaron los mecanismos de defensa. Praseed y Thilagam (2019) evidenciaron que los ataques en la capa de aplicación pueden conducirse con un volumen comparativamente bajo y con mayor sigilo; Tripathi y Hubballi (2021) destacaron su carácter disimulado; y Deressu et al. (2026) registran entre ellos una clase de ataques asimétricos, que disparan operaciones de servidor costosas a un costo mínimo para el atacante.

Este artículo se publica en acceso abierto (Open Access) bajo licencia Creative Commons Attribution 4.0 Internacional (CC-BY), que permite uso, distribución y reproducción irrestrictos en cualquier medio, siempre que el autor original y la fuente sean acreditados.

REVISTA CIENTÍFICASOL21
STANDARD OPEN LITERATURE - ISSN 3086-089X

De ese conjunto se extrae el concepto central adoptado en este trabajo: la razón de asimetría, entendida como la proporción entre el costo computacional impuesto al servidor y el costo en que incurre el cliente para provocarlo. La formulación converge con la clase de ataques asimétricos de Deressu et al. (2026), y lo que aquí se agrega es su conversión en variable de evaluación, y no de detección. En arquitecturas orientadas a API, esa razón puede alcanzar valores que vuelven irrelevante la discusión sobre ancho de banda.

1.2 Amenazas automatizadas y riesgos específicos de APIs

El proyecto Automated Threats to Web Applications, de OWASP, ofrece una ontología independiente de proveedor con veintiún eventos de amenaza automatizada (OAT, OWASP Automated Threat), de OAT-001 a OAT-021 (OWASP, 2018). Cuatro interesan al objeto: el OAT-015, Denial of Service, que cubre el agotamiento de recursos de la aplicación, de la base de datos o de cuentas; el OAT-021, Denial of Inventory, que describe la depleción de disponibilidad sin que la transacción se complete; el OAT-003, Cost-Inflation Fraud, relativo a la elevación artificial de costos; y el OAT-018, relativo al mapeo automatizado por el cual endpoints no inventariados son localizados por terceros.

La edición de 2023 del OWASP API Security Top 10 organiza los diez riesgos más críticos de APIs (OWASP, 2023), de los cuales tres estructuran el análisis: el API4:2023, relativo a la ausencia de limitación sobre el consumo de procesamiento, memoria, almacenamiento o ancho de banda por cliente; el API6:2023, relativo al acceso automatizado a flujos de negocio cuyo agotamiento perjudica la operación; y el API9:2023, relativo al mantenimiento inadecuado del inventario de APIs activas, de versiones obsoletas y de entornos expuestos.

Dos categorías derivan de ese último riesgo. Una Shadow API es un endpoint en operación que no consta en la especificación, no está registrado en el gateway y no tiene responsable designado. Un endpoint zombi es aquel formalmente obsoleto que permanece alcanzable porque rutas o balanceadores heredados nunca fueron removidos. Cabe registrar una limitación del corpus examinado: en él, casi toda la cuantificación en circulación sobre esas categorías proviene de proveedores de soluciones de protección de APIs, cuyo interés comercial coincide con la magnitud del diagnóstico. Esos datos no se emplean aquí; las categorías se tratan como constructos conceptuales, atestiguados por el reconocimiento normativo en API9:2023. Evidencia indirecta proviene de Santos Filho et al. (2025): entre 1.955 especificaciones OpenAPI analizadas, solo nueve contenían los elementos exigidos.

1.3 Plataformas de protección de aplicaciones y APIs

La sigla WAAP, de web application and API protection, designa la convergencia, en una plataforma única, de cortafuegos de aplicación web, protección contra denegación de servicio, gestión de bots y seguridad de APIs. Se trata de una categoría de mercado, y no de un constructo

Este artículo se publica en acceso abierto (Open Access) bajo licencia Creative Commons Attribution 4.0 Internacional (CC-BY), que permite uso, distribución y reproducción irrestrictos en cualquier medio, siempre que el autor original y la fuente sean acreditados.

académico o normativo, distinción que este artículo mantiene explícita, dado que la literatura sobre el tema es predominantemente producida por proveedores de la propia categoría. Los cortafuegos operan sobre firmas de petición; los gateways aplican límites de tasa, pero no detectan comportamiento anómalo a lo largo de una sesión; las soluciones volumétricas actúan sobre el caudal.

2 - METODOLOGÍA

2.1 Diseño de la investigación

La investigación es cualitativa en cuanto al enfoque, aplicada en cuanto a la naturaleza y documental en cuanto a los procedimientos. Adoptamos la Design Science Research (DSR), diseño apropiado cuando el objetivo no es describir un fenómeno, sino producir un artefacto que resuelva un problema organizacional identificado (Hevner et al., 2004). El recorrido sigue a Peffers et al. (2007), de cuyas actividades se condujeron la identificación del problema, la definición de los objetivos, el diseño y desarrollo y la demostración; la evaluación en entorno de uso no fue ejecutada.

2.2 Protocolo del levantamiento documental

El levantamiento es intencional, y no sistemático: la DSR requiere fundamentación suficiente para diseñar un artefacto, y no una estimación de la distribución de resultados en un campo, finalidad propia de la revisión sistemática. De ahí una consecuencia que explicitamos desde ahora: las afirmaciones de este artículo sobre escasez o ausencia de evidencia se refieren al corpus descrito en esta subsección y en la siguiente, y no autorizan generalización sobre la totalidad de la literatura.

Las bases consultadas fueron de tres tipos. Para literatura revisada por pares, IEEE Xplore, ACM Digital Library, SpringerLink, ScienceDirect y Google Scholar, este último para el rastreo de citas. Para fuentes normativas y de conocimiento abierto, los sitios oficiales de OWASP, de la International Organization for Standardization y del National Institute of Standards and Technology, sin intermediación de base indexadora. Para telemetría, las publicaciones propias de los operadores de infraestructura, en sus dominios institucionales.

Los descriptores se combinaron en inglés y en portugués mediante operadores booleanos, aplicados a título y resumen, con rastreo posterior de las referencias de los trabajos retenidos. En inglés: application layer DDoS, application-layer denial of service, low-rate DoS, API security, API inventory, shadow API, zombie endpoint, unrestricted resource consumption, rate limiting, economic denial of sustainability y web application and API protection; en portugués, los términos correspondientes. El período se definió por estrato: de 2019 a 2026 para la literatura sustantiva; sin límite inferior para obras fundadoras de concepto, lo que explica a Zimmermann (1980) y a Mirkovic

Este artículo se publica en acceso abierto (Open Access) bajo licencia Creative Commons Attribution 4.0 Internacional (CC-BY), que permite uso, distribución y reproducción irrestrictos en cualquier medio, siempre que el autor original y la fuente sean acreditados.

y Reiher (2004); edición vigente en septiembre de 2026 para las normas; y los dos ciclos más recientes de cada operador para la telemetría.

Los criterios de inclusión, acumulativos, fueron cuatro: tratar de denegación de servicio, de consumo de recursos o de la superficie de exposición de APIs; pertenecer a uno de los tres primeros niveles de la jerarquía de la subsección 2.3; estar disponible en texto íntegro en portugués, inglés o español; y permitir rastrear cada dato cuantitativo hasta el documento que lo produjo. Los de exclusión fueron cinco: dato cuya fuente primaria no pudiera alcanzarse; porcentaje atribuido a un tercero sin referencia verificable; material de proveedor como base cuantitativa; preprint cuando existe la versión revisada por pares; y reproducción secundaria de datos ya presentes en la fuente primaria.

Debe registrarse una laguna: no se conservó el registro del número de documentos inicialmente recuperados en cada base antes del cribado. El procedimiento es reconstituible en cuanto a las bases, los descriptores, el período y los criterios, pero no en cuanto al conteo de cribado. El corpus retenido está íntegramente expuesto en la lista de referencias, lo que permite auditar lo que se utilizó, aunque no lo que se descartó.

2.3 Calificación de las fuentes y composición del corpus

Dada la naturaleza del tema, en el cual gran parte de la producción es generada por proveedores de solución, se estableció una jerarquía de calificación de las fuentes en cuatro niveles: normas técnicas y proyectos comunitarios de conocimiento abierto; literatura revisada por pares, incluida la de fundamentación metodológica; informes de telemetría de operadores de infraestructura, admitidos cuando el dato deriva de observación directa de la propia red; y material técnico de proveedores, admitido para caracterización conceptual y nunca como base cuantitativa.

El corpus retenido comprende diecinueve documentos: cuatro en el primer nivel, doce en el segundo, tres en el tercero y ninguno en el cuarto. La aplicación de los criterios tuvo consecuencia sustantiva sobre los resultados: las cuantificaciones en circulación acerca de la prevalencia de endpoints no gobernados fueron íntegramente excluidas, lo que llevó a tratar esas categorías como constructos conceptuales y a convertir la ausencia de evidencia independiente en un hallazgo del trabajo.

2.4 Estrategia de evaluación del artefacto

La evaluación de artefactos en DSR admite posiciones distintas en cuanto al momento y al paradigma. Venable et al. (2016) las organizan en dos dimensiones: la finalidad, de la formativa a la sumativa, y el paradigma, de lo artificial a lo naturalístico. La evaluación artificial recurre a análisis teórico, simulación o experimento y ofrece mayor replicabilidad al costo de abstraer el entorno real; la naturalística examina el artefacto en uso, en una organización real. Distinguen además la evaluación ex ante, sobre el artefacto diseñado, de la ex post.

Este artículo se publica en acceso abierto (Open Access) bajo licencia Creative Commons Attribution 4.0 Internacional (CC-BY), que permite uso, distribución y reproducción irrestrictos en cualquier medio, siempre que el autor original y la fuente sean acreditados.

REVISTA CIENTÍFICASOL21
STANDARD OPEN LITERATURE - ISSN 3086-089X

Situamos este trabajo en ese cuadro. La evaluación conducida es ex ante, artificial y formativa, realizada mediante demostración analítica en un escenario ilustrativo, conforme a la subsección 3.6, y corresponde a la estrategia que Venable et al. (2016) denominan riesgo técnico y eficacia. No se condujo evaluación naturalística ni sumativa. De ahí la restricción observada en todo el texto: el artefacto se designa modelo propuesto, sus niveles son niveles ordinales propuestos y su demostración es analítica; no se afirma eficacia demostrada ni utilidad empíricamente validada.

La evaluación naturalística que completaría el ciclo fue diseñada, aunque no ejecutada, y queda registrada para permitir su conducción por otros investigadores. La unidad de análisis sería la organización que opere una superficie de APIs bajo plataforma de protección contratada. El procedimiento comprendería la aplicación del modelo por un evaluador externo; la conciliación entre el inventario declarado y la alcanzabilidad observada en ejecución; y el contraste entre la lectura de la regla de composición y la de la media aritmética. La regla quedaría desmentida si organizaciones con madurez elevada en D3 a D5 y baja en D1 presentaran cobertura efectiva equivalente a la de aquellas con madurez elevada en D1.

2.5 Limitaciones del estudio

Cuatro limitaciones deben explicitarse. El ciclo de la DSR no fue completado: el artefacto fue demostrado analíticamente, pero no sometido a evaluación naturalística, de modo que su utilidad es argumentada y no demostrada. El levantamiento es intencional y no conserva el registro del conteo de cribado, razón por la cual las afirmaciones sobre escasez de evidencia se restringen al corpus retenido. La ausencia de evidencia independiente sobre la prevalencia de endpoints no gobernados impide calibrar los niveles contra distribuciones observadas. Y la volatilidad del objeto llevó a construir el modelo sobre propiedades de proceso.

2.6 Declaración del uso de inteligencia artificial

Hubo uso de IA: Los autores declaran que utilizaron el asistente Claude, de Anthropic, modelo Claude Opus 5, accedido entre septiembre y octubre de 2026, sin ajuste fino y sin integración a sistemas de los autores, en el levantamiento y cribado de fuentes, la definición del diseño, la redacción, la construcción de la taxonomía, la formulación de las dimensiones, niveles y métricas, la demostración analítica, el mapeo normativo y el formateo. El uso afectó la totalidad del texto, incluidas las Tablas 1 a 6 y las referencias, y abarca la revisión realizada en respuesta a los dictámenes de evaluación por pares. Correspondió exclusivamente a los autores la definición del tema y del recorte, el cambio de diseño del estudio de caso a la proposición de artefacto, los criterios de calificación y exclusión de fuentes, la decisión de proseguir sin las bases cuantitativas sobre endpoints no gobernados, la de mantener en la taxonomía las dos clases que no satisfacen la definición estricta mediante criterio declarado, y la validación final. No se suministraron datos personales, sensibles o protegidos por sigilo, y ninguna organización fue identificada, por lo que no

Este artículo se publica en acceso abierto (Open Access) bajo licencia Creative Commons Attribution 4.0 Internacional (CC-BY), que permite uso, distribución y reproducción irrestrictos en cualquier medio, siempre que el autor original y la fuente sean acreditados.

hubo tratamiento de datos sujetos a la Ley n.º 13.709/2018 de Brasil. Los prompts están archivados por los autores y disponibles a solicitud. Cada referencia fue cotejada contra la fuente primaria, y ningún dato, cita o conclusión fue aceptado sin verificación humana.

3 - RESULTADOS Y DISCUSIÓN

3.1 Taxonomía del agotamiento de recursos en APIs

La primera contribución organiza las clases de agotamiento de recursos aplicables a APIs según el recurso agotado, y no según el protocolo, puesto que la clasificación protocolar no distingue operaciones cuyo costo difiere en órdenes de magnitud.

Tabla 1

Taxonomía de las clases de agotamiento de recursos con efecto de denegación de servicio en APIs

Clase	Mecanismo de agotamiento	Recurso objetivo	Correspondencia
Agotamiento volumétrico	Gran número de peticiones válidas contra endpoints de costo ordinario.	Procesamiento y conexiones	OAT-015; API4:2023
Agotamiento por conexión	Conexiones abiertas por largo tiempo con ancho de banda mínimo, incluidas lectura y escritura lentas.	Descriptores, hilos y pool de conexiones	OAT-015; API4:2023
Amplificación por costo de consulta	Operaciones cuyo costo de procesamiento es desproporcionado al de solicitarlas: consultas anidadas, agregaciones, paginación irrestricta.	Procesamiento, memoria y base de datos	API4:2023
Agotamiento de flujo de negocio	Consumo automatizado de disponibilidad finita sin conclusión de la transacción.	Inventario, reservas, cuotas y posiciones de cola	OAT-021; API6:2023
Inflación de costo, o EDoS	Disparo repetido de operaciones que encadenan servicios tarifados o infraestructura elástica.	Presupuesto de infraestructura y de servicios	OAT-003; API4:2023

Nota. Elaborado por los autores con base en OWASP (2018, 2023), Praseed y Thilagam (2019), Aldhyani y Alkahtani (2022) y Deressu et al. (2026).

La taxonomía admite estratificación en cuanto al propio concepto de denegación de servicio. Las tres primeras clases constituyen denegación de servicio en sentido estricto: producen indisponibilidad por agotamiento de un recurso computacional. La cuarta consume disponibilidad finita sin consumir recurso computacional, y el propio OWASP la designa denegación de inventario al catalogarla como OAT-021. La quinta no produce indisponibilidad: corresponde al constructo que la literatura sobre computación en la nube denomina denegación económica de sostenibilidad (EDoS, economic denial of sustainability), que explota el modelo de pago por uso de manera que el daño se materialice en la factura mientras el servicio permanece disponible (Aldhyani y Alkahtani, 2022).

Mantenemos las dos últimas clases por una razón que conviene declarar, ya que ninguna satisface la definición estricta: ambas comparten con las tres primeras el mecanismo generador, que es la asimetría entre el costo de solicitar y el de atender, y los controles que las mitigan. No las mantenemos por producir indisponibilidad, pues la quinta no la produce. En arquitecturas elásticas,

Este artículo se publica en acceso abierto (Open Access) bajo licencia Creative Commons Attribution 4.0 Internacional (CC-BY), que permite uso, distribución y reproducción irrestrictos en cualquier medio, siempre que el autor original y la fuente sean acreditados.

el dimensionamiento automático que preserva la disponibilidad es precisamente el mecanismo que materializa el perjuicio, de modo que la defensa contra la tercera clase puede agravar la quinta.

3.2 Endpoints no gobernados como deuda de superficie de ataque

La segunda contribución distingue dos categorías frecuentemente tratadas como sinónimas, que tienen orígenes distintos y demandan tratamientos distintos.

Tabla 2

Distinción conceptual entre Shadow APIs y endpoints zombis

Criterio	Shadow API	Endpoint zombi
Origen	Publicación fuera del proceso formal de gobernanza.	Alcanzabilidad mantenida tras la obsolescencia o la sustitución.
Inventario	Nunca registrado.	Registrado como inactivo, sin correspondencia con el estado real.
Riesgo predominante	Ausencia de limitación de consumo, autenticación y monitoreo.	Vulnerabilidades ya corregidas en las versiones vigentes.
Descubrimiento	Análisis de tráfico en ejecución e inspección de artefactos de despliegue.	Conciliación entre inventario declarado y alcanzabilidad observada.

Nota. Elaborado por los autores con base en OWASP (2023).

Las dos categorías difieren también en el ciclo de vida: la Shadow API es activa y en evolución, sin responsable designado; el endpoint zombi es estático y sin correcciones. Designamos el conjunto como deuda de superficie de ataque, por analogía con la deuda técnica: se acumula como subproducto de la velocidad de entrega y devenga intereses, pues cada endpoint no gobernado permanece expuesto a vulnerabilidades ya corregidas en otros puntos.

3.3 Supuestos que limitan a las plataformas de protección

La tercera contribución resulta del análisis de la arquitectura de operación de las plataformas. Se identificaron cuatro supuestos cuya violación degrada la eficacia independientemente de la calidad de la implementación. El de enrutamiento: la plataforma inspecciona solo el tráfico que atraviesa su punto de aplicación, y los endpoints alcanzables por caminos que eluden el gateway quedan fuera de su alcance. El de especificación: la protección basada en modelo positivo compara la petición con el contrato declarado, y los endpoints sin especificación recaen en reglas genéricas —dependencia reconocida por la propia literatura que la propone (Santos Filho et al., 2025)—.

El de línea de base: la detección comportamental depende de un historial que caracterice el uso normal, y los endpoints de bajo tráfico, recién expuestos o intermitentes no lo acumulan. Y el de costo uniforme: los límites expresados en peticiones por unidad de tiempo tratan como equivalentes operaciones cuyo costo difiere en órdenes de magnitud. Los tres primeros son violados precisamente por los endpoints no gobernados; el cuarto, por cualquier endpoint cuyo costo no haya sido caracterizado. De esos enunciados se deriva el diseño del modelo.

Este artículo se publica en acceso abierto (Open Access) bajo licencia Creative Commons Attribution 4.0 Internacional (CC-BY), que permite uso, distribución y reproducción irrestrictos en cualquier medio, siempre que el autor original y la fuente sean acreditados.

3.4 El modelo de evaluación propuesto

El artefacto consiste en un modelo de cinco dimensiones, cada una medida en cuatro niveles ordinales propuestos. Las tres primeras abordan los supuestos de enrutamiento, especificación y línea de base; la cuarta, el de costo uniforme; la quinta trata de la respuesta, cuya oportunidad está condicionada por la duración observada de los ataques.

Tabla 3

Dimensiones del modelo de evaluación de la exposición de APIs

Dimensión	Objeto de evaluación	Evidencia en el nivel gestionado
D1 — Completitud del inventario	Identificar, de forma continua y en ejecución, los endpoints alcanzables.	Descubrimiento automatizado, con conciliación periódica entre lo observado y lo declarado.
D2 — Gobernanza del ciclo de vida	Conducir el endpoint del registro a la desactivación efectiva, con responsable designado.	Política de obsolescencia con plazo y verificación de alcanzabilidad tras la desactivación.
D3 — Caracterización de costo	Conocer el costo computacional de cada operación y de sus dependencias.	Endpoints clasificados por franja de costo, con límites diferenciados por franja.
D4 — Control de consumo	Límites y cuotas proporcionales al costo y al grado de autenticación.	Cuotas por identidad y por costo agregado, con techo para el tráfico no autenticado.
D5 — Detección y respuesta autónoma	Identificar automatización anómala y contener sin intervención humana.	Contención automática por política, con registro auditable y reversión.

Nota. Elaborado por los autores.

Los cuatro niveles son uniformes entre las dimensiones y ordinales: la progresión indica dirección de mejora, y no distancia medida. En el nivel 1, Inicial, la dimensión no es objeto de un proceso definido. En el nivel 2, Definido, hay proceso documentado y responsable designado, pero ejecución periódica y manual. En el nivel 3, Gestionado, el proceso es continuo y automatizado. En el nivel 4, Optimizado, se integra al ciclo de entrega, de modo que la violación de la política impide la promoción del cambio al entorno productivo.

El modelo adopta una regla de composición que expresa la proposición central del trabajo: el nivel global no corresponde a la media de las dimensiones, sino que está limitado por el nivel de D1. La justificación es lógica: los controles de consumo, la caracterización de costo y la detección comportamental se aplican a endpoints conocidos, de modo que, si la organización desconoce parte de su superficie, esos controles operan sobre un subconjunto de tamaño indeterminado y su cobertura agregada es indeterminada en la misma medida. Una organización con madurez máxima en D3, D4 y D5 e inicial en D1 posee protección de extensión desconocida.

3.5 Métricas operativas

Para volver aplicable el modelo, se proponen cinco métricas, calculables a partir de datos que la operación ya produce. La cuarta de ellas desplaza la pregunta de gestión —cuántas peticiones por segundo soporta el sistema— hacia otra, más informativa: cuál es la operación más costosa que un adversario sin credenciales consigue disparar.

Este artículo se publica en acceso abierto (Open Access) bajo licencia Creative Commons Attribution 4.0 Internacional (CC-BY), que permite uso, distribución y reproducción irrestrictos en cualquier medio, siempre que el autor original y la fuente sean acreditados.

Tabla 4

Métricas operativas propuestas

Métrica	Definición operativa	Interpretación
Índice de completitud del inventario	Endpoints observados en ejecución que constan en el inventario, sobre el total observado.	Fracción de la superficie bajo gobernanza; indicador limitante del modelo.
Razón de endpoints no enrutados	Endpoints activos alcanzables por un camino que no atraviesa el punto de aplicación.	Mide la violación del supuesto de enrutamiento.
Latencia de retiro	Tiempo mediano entre la obsolescencia anunciada y la verificación de inalcanzabilidad.	Estima la tasa de acumulación de endpoints zombis.
Costo máximo por petición no autenticada	Mayor costo observado entre las operaciones disparables sin autenticación.	Aproxima el techo de la razón de asimetría disponible para un adversario anónimo.
Cobertura de política de automatización	Endpoints inventariados bajo política de detección y límite de consumo.	Distancia entre la cobertura contratada y la efectiva.

Nota. Elaborado por los autores.

3.6 Demostración analítica del modelo

Para demostrar la aplicabilidad del modelo, lo aplicamos a una configuración ilustrativa, construida para ese fin y que no corresponde a ninguna organización observada: los valores no constituyen dato empírico, sino parámetros elegidos para ejercitar el instrumento. La demostración es, en los términos de la subsección 2.4, una evaluación ex ante, artificial y formativa. La configuración supone una organización de servicios financieros que declara trescientos endpoints en su gateway y presenta trescientos ochenta alcanzables en observación de tráfico; la política de consumo cubre todo el inventario declarado; no hay clasificación por franja de costo; y la conciliación es manual y anual.

Tabla 5

Demostración analítica: configuración ilustrativa y lectura del modelo

Dimensión	Nivel	Evidencia en la configuración ilustrativa	Métrica
D1 — Completitud del inventario	1 — Inicial	Conciliación manual y anual; sin descubrimiento en ejecución.	Completitud = 300/380 = 0,79
D2 — Gobernanza del ciclo de vida	2 — Definido	Política de obsolescencia documentada, verificación manual.	Latencia de retiro = 180 días
D3 — Caracterización de costo	1 — Inicial	Costo de las operaciones no caracterizado.	Costo máximo no autenticado = no determinado
D4 — Control de consumo	3 — Gestionado	Cuotas por identidad automatizadas, sobre todo el inventario declarado.	Cobertura = 1,00 del inventario; 0,79 de lo observado
D5 — Detección y respuesta	3 — Gestionado	Contención automática por política, con registro auditable.	Endpoints no enrutados = 0,21
Lectura global	1 — Inicial	Regla de composición: limitada por el nivel de D1.	Media aritmética = 2,0

Nota. Elaborado por los autores. Configuración ilustrativa, construida con fines de demostración; no corresponde a ninguna organización observada y los valores no constituyen dato empírico.

Este artículo se publica en acceso abierto (Open Access) bajo licencia Creative Commons Attribution 4.0 Internacional (CC-BY), que permite uso, distribución y reproducción irrestrictos en cualquier medio, siempre que el autor original y la fuente sean acreditados.

La demostración produce un resultado que no es evidente antes de aplicar el instrumento. Las dos lecturas divergen: la media aritmética sitúa a la organización en el nivel 2 y sugiere esfuerzo sobre D2; la regla de composición la sitúa en el nivel 1 y lo dirige hacia D1 y D3. La divergencia no es de grado, sino de objeto de inversión. Su origen está en la métrica de cobertura: la política de consumo alcanza todo el inventario declarado y reporta cobertura integral, mientras alcanza el 79% de la superficie observada. El 21% restante no está bajo límite alguno, y la ausencia de caracterización de costo impide estimar el techo de la razón de asimetría disponible para un adversario sin credenciales.

3.7 Mapeo normativo y discusión

Tabla 6

Mapeo de las dimensiones del modelo contra referenciales normativos

Dimensión	OWASP API	ISO/IEC 27001:2022	NIST CSF 2.0
D1 — Completitud del inventario	API9:2023	5.9 Inventario de activos	Identificar
D2 — Gobernanza del ciclo de vida	API9:2023; API8:2023	8.9 Gestión de configuración; 8.32 Gestión de cambios	Gobernar; Proteger
D3 — Caracterización de costo	API4:2023	8.6 Gestión de capacidad	Identificar; Proteger
D4 — Control de consumo	API4:2023; API6:2023	8.26 Requisitos de seguridad de aplicaciones	Proteger
D5 — Detección y respuesta	API4:2023	8.16 Monitoreo; 5.7 Inteligencia de amenazas	Detectar; Responder

Nota. Elaborado por los autores con base en OWASP (2023), International Organization for Standardization (2022) y National Institute of Standards and Technology (2024). Los identificadores de control deben cotejarse contra el texto íntegro de la norma antes de su uso en auditoría.

El modelo dialoga críticamente con la orientación predominante en el campo, que concentra el esfuerzo en la detección, particularmente mediante aprendizaje automático (Praseed y Thilagam, 2019; Tripathi y Hubballi, 2021; Deressu et al., 2026). Esa producción presupone resuelto el problema que este trabajo identifica como anterior: para que un detector observe tráfico anómalo en un endpoint, es necesario que el tráfico de ese endpoint sea observado. La eficacia demostrada en un entorno experimental, donde el conjunto de endpoints es conocido por construcción, no se transfiere a un entorno productivo. Resta una tensión: D1 demanda observabilidad justamente donde ella falta.

4 - CONSIDERACIONES FINALES

Este trabajo propuso un modelo de evaluación de la exposición de las APIs a la denegación de servicio en la capa de aplicación, construido mediante la estrategia de la Design Science Research y destinado a organizaciones cuya superficie de ataque no esté íntegramente inventariada.

Este artículo se publica en acceso abierto (Open Access) bajo licencia Creative Commons Attribution 4.0 Internacional (CC-BY), que permite uso, distribución y reproducción irrestrictos en cualquier medio, siempre que el autor original y la fuente sean acreditados.

Construimos una taxonomía de cinco clases organizadas por el recurso agotado, evidenciándose que solo una corresponde a la concepción volumétrica corriente y que dos se sitúan en el límite del concepto de denegación de servicio. Distinguimos las Shadow APIs de los endpoints zombis, proponiendo la designación del conjunto como deuda de superficie de ataque. Identificamos cuatro supuestos de eficacia de las plataformas, de los cuales los tres primeros son violados por endpoints no gobernados.

La contribución teórica consiste en la formulación de la razón de asimetría como variable explicativa central y en la proposición de que la completitud del inventario es un factor determinante de la cobertura efectiva y de la gobernanza de los demás controles. Conviene precisar su alcance: el estudio sostiene que el inventario delimita la extensión sobre la cual operan los demás controles y que, desconocida esa extensión, la cobertura agregada es indeterminada; no sostiene que la completitud sea condición universal de eficacia. Dada la ausencia de evaluación naturalística, el modelo debe leerse como una propuesta sujeta a examen empírico.

Como agenda futura, sugerimos la evaluación naturalística del modelo conforme al protocolo de la subsección 2.4; la producción de evidencia independiente sobre la prevalencia de endpoints no gobernados; la investigación experimental de la razón de asimetría en interfaces REST y declarativas; y el análisis de las implicaciones regulatorias de la indisponibilidad de servicios mediados por API en sectores regulados en Brasil.

REFERENCIAS

Aldhyani, T. H. H., & Alkahtani, H. (2022). Artificial intelligence algorithm-based economic denial of sustainability attack detection systems: Cloud computing environments. *Sensors, 22*(13), 4685. https://doi.org/10.3390/s22134685

Cloudflare. (2026). *DDoS threat report for 2026 Q2*. Cloudforce One. https://blog.cloudflare.com/ddos-threat-report-2026-h1/

Deressu, T. F., Beyene, A. M., & Gemechu, A. Y. (2026). A survey of application layer distributed denial of service detection: Approaches, models, and datasets. *Artificial Intelligence Review, 59*(6), 139. https://doi.org/10.1007/s10462-026-11528-3

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS Quarterly, 28*(1), 75–105. https://doi.org/10.2307/25148625

Imperva. (2025). *2025 bad bot report*. Thales Group. https://www.imperva.com/resources/wp-content/uploads/sites/6/reports/2025-Bad-Bot-Report.pdf

Imperva. (2026). *Bad bot report 2026: Bad bots in the agentic age*. Thales Group. https://www.imperva.com/blog/bad-bot-report-2026-bots-agentic-age/

International Organization for Standardization. (2022). *ISO/IEC 27001:2022 — Information security, cybersecurity and privacy protection: Information security management systems: Requirements*.

Este artículo se publica en acceso abierto (Open Access) bajo licencia Creative Commons Attribution 4.0 Internacional (CC-BY), que permite uso, distribución y reproducción irrestrictos en cualquier medio, siempre que el autor original y la fuente sean acreditados.

Mirkovic, J., & Reiher, P. (2004). A taxonomy of DDoS attack and DDoS defense mechanisms. *ACM SIGCOMM Computer Communication Review, 34*(2), 39–53. https://doi.org/10.1145/997150.997156

National Institute of Standards and Technology. (2024). *The NIST Cybersecurity Framework (CSF) 2.0* (NIST CSWP 29). U.S. Department of Commerce. https://doi.org/10.6028/NIST.CSWP.29

OWASP. (2018). *OWASP automated threat handbook: Web applications* (Version 1.2). Open Worldwide Application Security Project. https://owasp.org/www-project-automated-threats-to-web-applications/

OWASP. (2023). *OWASP API security top 10 — 2023*. Open Worldwide Application Security Project. https://owasp.org/API-Security/editions/2023/en/0x00-toc/

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of Management Information Systems, 24*(3), 45–77. https://doi.org/10.2753/MIS0742-1222240302

Praseed, A., & Thilagam, P. S. (2019). DDoS attacks at the application layer: Challenges and research perspectives for safeguarding web applications. *IEEE Communications Surveys & Tutorials, 21*(1), 661–685. https://doi.org/10.1109/COMST.2018.2870658

Santos Filho, A., Rodríguez, R. J., & Feitosa, E. L. (2025). Automated broken object-level authorization attack detection in REST APIs through OpenAPI to colored Petri nets transformation. *International Journal of Information Security, 24*(2), 83. https://doi.org/10.1007/s10207-024-00970-5

Tanveer, F., Iradat, F., Iqbal, W., & Ahmad, A. (2025). Towards secure APIs: A survey on RESTful API vulnerability detection. *Computers, Materials & Continua, 84*(3), 4223–4257. https://doi.org/10.32604/cmc.2025.067536

Tripathi, N., & Hubballi, N. (2021). Application layer denial-of-service attacks and defense mechanisms: A survey. *ACM Computing Surveys, 54*(4), 1–33. https://doi.org/10.1145/3448291

Venable, J., Pries-Heje, J., & Baskerville, R. (2016). FEDS: A framework for evaluation in design science research. *European Journal of Information Systems, 25*(1), 77–89. https://doi.org/10.1057/ejis.2014.36

Zargar, S. T., Joshi, J., & Tipper, D. (2013). A survey of defense mechanisms against distributed denial of service (DDoS) flooding attacks. *IEEE Communications Surveys & Tutorials, 15*(4), 2046–2069. https://doi.org/10.1109/SURV.2013.031413.00127

Zimmermann, H. (1980). OSI reference model — The ISO model of architecture for open systems interconnection. *IEEE Transactions on Communications, 28*(4), 425–432. https://doi.org/10.1109/TCOM.1980.1094702

Este artículo se publica en acceso abierto (Open Access) bajo licencia Creative Commons Attribution 4.0 Internacional (CC-BY), que permite uso, distribución y reproducción irrestrictos en cualquier medio, siempre que el autor original y la fuente sean acreditados.

REVISTA CIENTÍFICASOL21
STANDARD OPEN LITERATURE - ISSN 3086-089X

APÉNDICE A — SIGLAS Y TÉRMINOS TÉCNICOS

API, application programming interface; CSF, Cybersecurity Framework; DDoS, distributed denial of service; DSR, Design Science Research; EDoS, economic denial of sustainability; GraphQL, lenguaje de consulta para APIs; gRPC, marco de llamada a procedimiento remoto; HTTP, hypertext transfer protocol; IA, inteligencia artificial; ISO/IEC, International Organization for Standardization e International Electrotechnical Commission; NIST, National Institute of Standards and Technology; OAT, OWASP automated threat; OSI, open systems interconnection; OWASP, Open Worldwide Application Security Project; REST, representational state transfer; WAAP, web application and API protection.

Este artículo se publica en acceso abierto (Open Access) bajo licencia Creative Commons Attribution 4.0 Internacional (CC-BY), que permite uso, distribución y reproducción irrestrictos en cualquier medio, siempre que el autor original y la fuente sean acreditados.