

网页防爬虫保护系统 产品白皮书

版本：V2.0

发布日期：2026年09月

出品单位：北京国隐科技有限公司

一、产品概述

网页防爬虫保护系统是面向政企内网业务平台、政务公开网站、行业信息系统打造的**智能化网页数据防窃取防护产品**。针对恶意爬虫批量爬取页面数据、批量盗采公开信息、接口刷取、内容镜像复制等安全风险，解决传统防火墙、WAF无法精准区分正常访问与恶意爬虫的行业短板。

产品采用无感人机识别与动态风控机制，无需改造原有网站代码，在保障正常用户浏览、办公访问不受影响的前提下，精准拦截批量爬虫、模拟访问、脚本刷取等恶意行为，实现网站数据防盗采、防镜像、防批量外泄，保障网页数据资产安全可控。

二、核心价值

- 精准人机识别，零误拦截**：智能区分人工正常访问与机器爬虫行为，合法用户无感通行，规避传统防护误拦截、卡顿问题。
- 全维度爬虫拦截**：覆盖普通脚本爬虫、高频刷取、多IP轮询爬取、页面镜像、接口批量盗采各类攻击方式。
- 零改造快速部署**：旁路接入现有网站与业务系统，无需源码改造、无需业务停机，极速上线防护。
- 保障数据资产安全**：杜绝网页涉密信息、政务数据、行业资料被批量爬取倒卖、违规扩散。
- 稳保业务稳定运行**：抵御爬虫高频访问造成的服务器卡顿、带宽挤占、系统过载瘫痪风险。

三、典型应用场景

- 政务公开网站防护**：政府公示信息、公告文件、公开数据防批量爬取，规范公开数据使用边界。
- 内网业务平台防护**：内网业务系统、查询平台、统计页面防恶意脚本爬取，杜绝内部数据外泄。
- 行业信息平台防护**：行业公示系统、交易平台、数据查询网站防数据盗采、镜像建站抄袭。
- 国企资讯系统防护**：企业公开资讯、招投标信息、公示数据防批量爬取，规避数据滥用风险。

四、产品整体架构

系统采用**旁路智能风控+行为分析+动态拦截**轻量化架构，适配各类网页与Web业务系统。



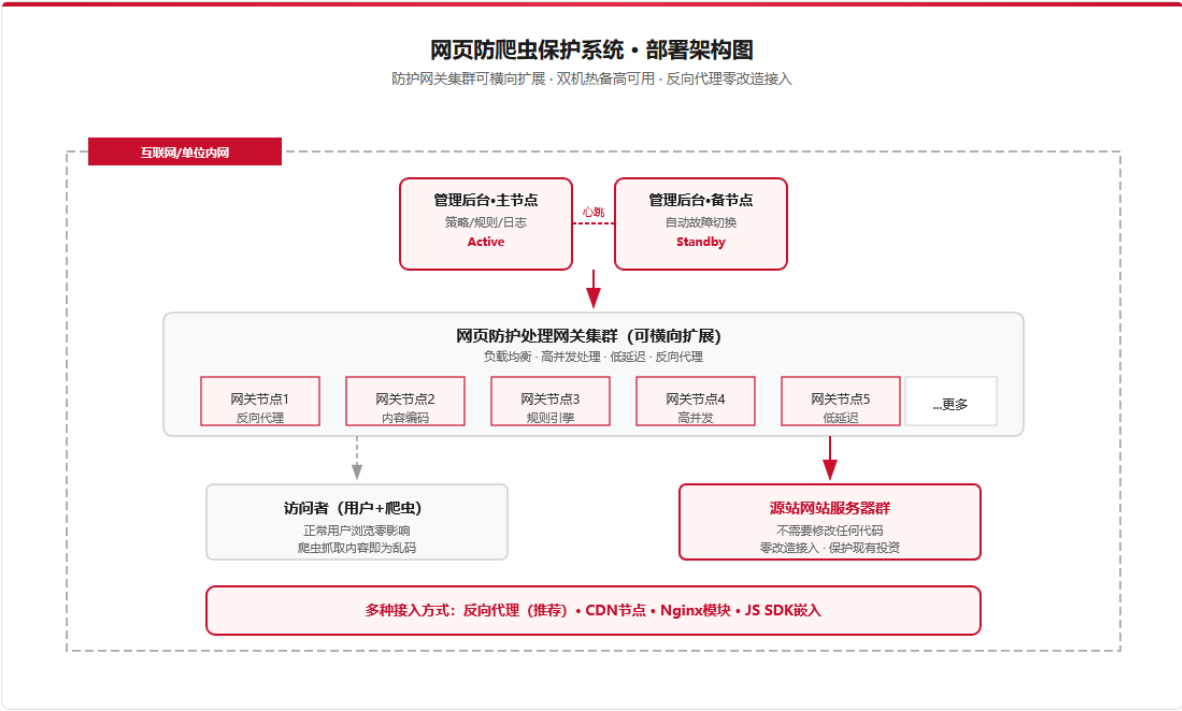
- **风控管理后台**：统一配置防护策略、访问规则、黑白名单，统计爬虫攻击日志、输出安全审计报告。
- **人机行为分析引擎**：基于访问频率、操作轨迹、设备特征智能识别爬虫行为，动态判别风险等级。
- **动态拦截溯源模块**：对恶意爬虫自动限流、拦截、告警，留存攻击轨迹，支撑安全溯源取证。

五、核心能力

- 智能人机识别，精准区分正常访问与机器爬虫访问
- 高频访问、批量请求、多IP轮询爬虫行为精准拦截
- 页面内容爬取、接口盗刷、网站镜像复制行为防护
- 按URL、访问时段、IP网段、访问权限精细化策略管控
- 恶意攻击实时告警、自动拦截、流量清洗，保障业务稳定
- 全量访问与攻击日志留存，支持安全审计与溯源分析
- 兼容内外网网站，适配静态、动态全类型网页业务

六、部署方案

产品采用轻量化旁路部署，不侵入网站业务架构，不影响网页访问速度，适配各级政企网站规模化防护。



6.1 部署模式

- **单机部署**：适配中小型官网、单站点业务系统，快速部署、即装即用。
- **双机热备（推荐）**：7×24小时高可用防护，适配政务、国企核心门户网站。
- **集群部署**：适配多站点集群、大型门户平台、高并发访问场景。

6.2 基础硬件要求

服务端：4核CPU、8G内存、500G硬盘、千兆内网；部署轻量化，对原有网站服务器零性能损耗。

七、信创全栈适配

产品完成全栈国产化适配，可在纯信创服务器与内网环境稳定运行，满足国产化替代合规要求。

适配类别	主流适配产品
桌面/服务器系统	银河麒麟、统信UOS、深度Deepin、openEuler
国产CPU架构	飞腾、鲲鹏、龙芯、海光、兆芯、申威
Web服务环境	国产Web容器、内外网门户系统、业务Web平台
数据库/中间件	达梦、人大金仓、东方通、金蝶天燕

注：以上为已实测适配的主流信创产品，具体适配可按需定制。

八、产品优势

- **智能精准防护**：区别于传统粗暴拦截模式，行为级智能识别，安全与体验双向兼顾
- **零业务侵入落地快**：旁路无感接入，无需改代码、无需停机，适配全部存量Web业务
- **攻防适配性强**：针对新型动态爬虫、分布式爬取具备持续对抗防护能力
- **信创生态全覆盖**：纯国产化环境适配，满足政企信创建设与数据安全合规要求

关于我们 | 联系方式

企业名称：北京国隐科技有限公司

地址：北京市海淀区丹棱街18号创富大厦九层910室

联系电话：13552814357

商务邮箱：xufang@hiding.com.cn

官方网站：www.gohiding.cn

本白皮书内容仅供参考，产品功能以实际交付版本为准