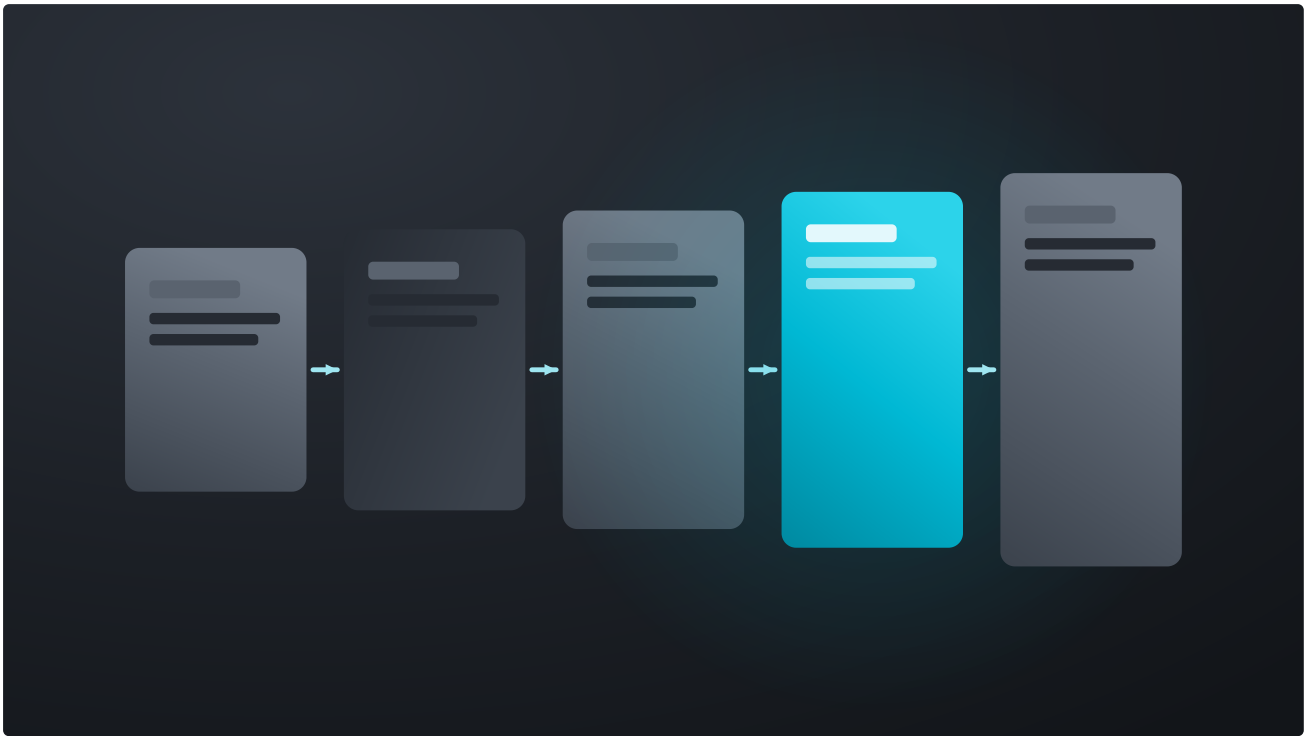
Black Shard

TECHNICAL WHITEPAPER

Jetstone: the agent, detection, response and undo model

Jetstone will watch Windows and Mac devices and Microsoft 365 or Google Workspace sign-ins, act on its own against high-confidence threats on a published list, and let every automatic action be undone.



Entity	Black Shard Pty Ltd · ABN 66 696 910 773
Document ID	BS-RPT-04
Classification	Public
Version	1.2
First published	25 October 2026
This edition	29 October 2026
Figures	Current at 25 Oct 2026

Executive summary

Jetstone is a cyber defence agent for organisations with 1 to 200 devices. It will watch Windows and Mac devices and Microsoft 365 or Google Workspace sign-ins, and it is built in Australia and will be hosted there.

A lightweight agent for Windows 10 and 11 and macOS 13 and later. Detection will run in Australia, with rules mapped to MITRE ATT&CK. On a high-confidence detection, Jetstone is designed to act on its own against threats on a published list: end the process, quarantine the file, isolate the device, or sign the user out of Microsoft 365 or Google Workspace. Every automatic action will be undoable in one tap.

Three design rules carry the product. Detection will be deterministic, so every finding can be traced to the rule that matched it. The plain-English write-up will explain an incident and never decide what happens to it. And every automatic action will be reversible, so a business can turn protection on without fearing the false alarm.

- An agent on each device, reading what the operating system records.
- Detection and storage in Australia.
- Automatic response only against a published list, only on high confidence, inside fixed guardrails.

Method

Jetstone has not yet launched. Every capability below is described as it is designed, in the future tense.

Prices, the count of detection rules and detection rates are held back. A detection rate means something only beside the test that produced it, and we will publish figures with their tests.

The model draws on the work our practice already does in security monitoring, incident response and threat advisories. Our security monitoring analysed its first 430 million security events from 1 October to 26 October 2026, counted at 00:45 AEST.

The model

FROM THE DEVICE TO THE UNDO

Jetstone's model in five steps, as it is designed to run. Source: Jetstone product design.

Jetstone will work in five steps. The agent on the device will collect what the operating system records. Detection will run in Australia, matching that telemetry against rules mapped to MITRE ATT&CK. Findings will be grouped into an incident and written up. On a high-confidence detection of a threat on the published list, the response will act on its own. And every action will be reversible from the notification that reports it.

The agent

A lightweight agent for Windows 10 and 11 and macOS 13 and later. The agent will read operating system telemetry on the device: processes, image loads, file, registry, network and DNS activity, PowerShell script blocks, and the Windows Security, System and Defender logs.

Local containment. A local trigger will suspend a process tree that starts renaming or encrypting files in bulk, on the device and without a round trip to the cloud, and then report it.

Posture. Every check-in will report the device's posture: installed software, pending updates, local administrators, disk encryption, antivirus state, remote desktop, the firewall and screen lock. The posture will feed a health score in the console; it is designed to show what to fix, and it will not be a vulnerability scanner.

Device and identity. The agent will report which account is signed in, so a risky sign-in can be matched to a managed device or flagged as coming from somewhere else.

Detection

Detection will run in Australia, with rules mapped to MITRE ATT&CK.

Each rule will be a predicate over the telemetry, written to catch one attacker technique and named for it, so a finding will always say which technique it matched and on what evidence. A scoring step will weigh findings by confidence. Rules will be added with each release, and each will be tested against the technique it names before it ships.

Findings will be grouped into one incident per device or per user, with the evidence behind them kept together.

Grouping matters to a small business. Many findings from one intrusion are one problem, and the console will show them as one incident with one account of what happened.

The write-up

Every alert will be written up in plain English. Each write-up will be drawn from the incident's evidence only; a write-up that names anything the evidence does not hold will be rejected and replaced by a plain template.

The write-up will never choose the response: the rules and the preset will.

Each write-up will say what happened, what Jetstone did and what is waiting on the owner. The prompts, the outputs and the decisions will be logged with the incident, so an investigator or an insurer can see how the account was produced.

Response and undo

On a high-confidence detection, Jetstone is designed to act on its own against threats on a published list: end the process, quarantine the file, isolate the device, or sign the user out of Microsoft 365 or Google Workspace.

Three policy presets, Watch, Guard and Lockdown, with fixed isolation guardrails. Watch will detect and notify, and will take no action on its own. Guard, the default preset, will act on its own on high-confidence detections on the published list and turn everything else into a card to approve or dismiss. Lockdown will be the strictest of the three.

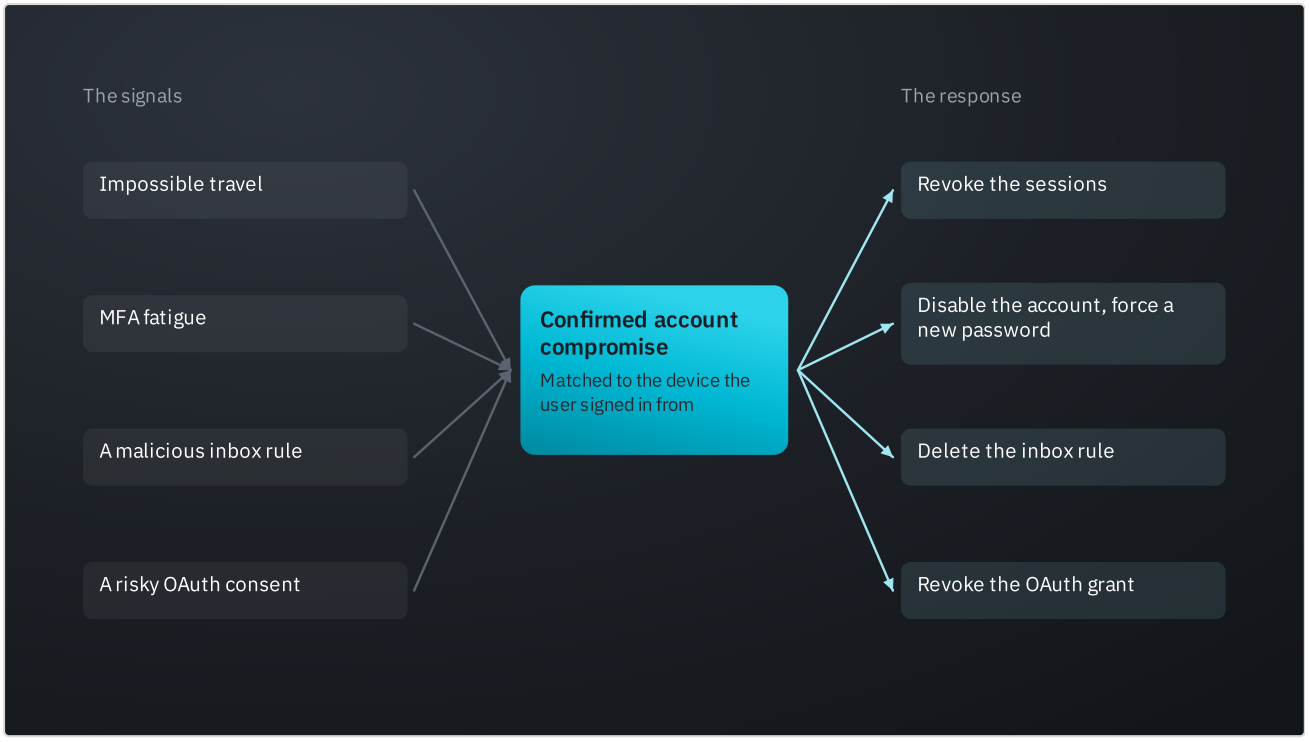
ACTION	WHAT IT WILL DO	HOW IT WILL BE UNDONE
End the process	Stop the process tree behind the detection	Allow the program to run again
Quarantine the file	Move the file out of reach	Restore the file to where it was
Isolate the device	Cut the device off the network, keeping its link to Jetstone	Release the isolation
Disable a task or service	Switch off what the attack created to come back	Enable it again
Sign the user out	End the user's sessions in Microsoft 365 or Google Workspace	The user signs in again

Every action will be logged, and every automatic action will be reversible from the notification that reports it. Every automatic action will be undoable in one tap.

Undo is what makes automatic response usable on a small fleet. A business without a security team cannot afford to have a payroll laptop cut off on a false alarm and wait until Monday for someone to release it, so release will sit in the same notification as the action.

Identity

IDENTITY: FROM THE SIGNAL TO THE RESPONSE



Signing the user out is the first step of the response, never all of it. Source: Jetstone product design.

Identity watch will flag impossible travel, MFA fatigue, malicious inbox rules and risky OAuth consents. On a confirmed account compromise, the response will revoke the user's sessions, disable the account and force a password change, delete the malicious inbox rule and revoke the risky OAuth grant.

Signing a user out ends the sessions an attacker holds, and it will be the first step only: a malicious inbox rule keeps forwarding mail after a password change, and an OAuth grant keeps reading files after a sign-out. The response will remove each of them.

Identity coverage will be switched on with one consent by the owner. The console will state which coverage the tenant has, because what Microsoft 365 reports, and how quickly, depends on the licences the business holds.

Guardrails

Automatic isolation will run inside set limits, and past them Jetstone will wait for a person.

The limits will be fixed in code, so no preset and no setting will be able to remove them.

Residency, console and licensing

Built in Australia. Data will be held in Australia. A phone-sized console, a Monday summary and an incident export for a broker or insurer.

Jetstone will be sold per device per year as software. Designed for businesses with 1 to 200 devices.

Jetstone is designed to need no IT team: the agent will install from one package, the presets will carry sensible defaults, and the Monday summary will tell the owner what happened in a week in a few lines.

What to do

Questions to put to any detection and response product before it goes on your devices:

1. Where is the telemetry processed and stored, and does any of it leave that region?
2. Which actions can the product take on its own, and is that list published?
3. Can every automatic action be undone, by whom, and how quickly?
4. Does anything other than a rule decide what action to take?
5. What evidence is a write-up drawn from, and what stops it describing something that did not happen?
6. What limits stop a misfiring rule from isolating the whole fleet?
7. How does the product work with the antivirus already on the device?

The Jetstone waitlist is open. Register at blackshard.com.au/contact to hear first when it opens to new businesses.

Our record

Black Shard builds and defends the systems Australian organisations run on. Jetstone is being built by the same practice that runs our own security monitoring and our incident response.

The record is as held at 29 October 2026, the date this edition was revised.

- SMB1001:2026 Gold, issued by CyberCert, valid 16 Jun 2026 to 17 Jun 2027.
- Essential Eight Maturity Level 3, assessed September 2026.
- Prequalified supplier to Queensland councils under Local Buy arrangement LB308, ICT Solutions, Products, Services and New Technologies.
- Prequalified supplier under the NSW Government ICT Services Scheme (SCM0020).
- Holder of Victorian Fair Jobs Code pre-assessment certificate FJC-261023-11119, required for Victorian Government threshold and high-value procurement.
- Our CISO holds the Exemplar Global ISO/IEC 27001:2022 Lead Auditor credential.
- Professional indemnity, public liability, cyber and workers' compensation insurance, all current.
- More than 65,000 tests across our repositories, counted on 26 October 2026.
- More than 1,500 deploys and releases shipped since 1 January 2026.

> 24/7 incident line: 1800 370 270.

Sources

- Jetstone product design and engineering specification, Black Shard
- MITRE ATT&CK: https://attack.mitre.org/