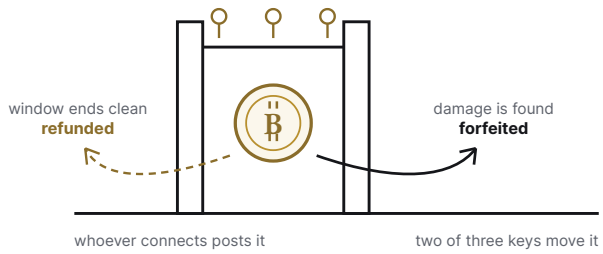
BITCOIN THEMATIC · CYBERSECURITY

AI Makes Attacks Cheap. Can Bitcoin Make Them Expensive?

Since 2009 anyone can take a final payment from an anonymous stranger with no bank in the middle. Charged as a fee, or held as a deposit lost on misuse, it puts a cost on an attacker that he cannot push onto a stolen machine. By the arithmetic it would add value now where attacks mostly fail and the failure shows at once, among users who can pay. Where attacks mostly succeed and stay hidden, which is where the money is, it would, on the mechanism, reduce the damage but deter only under conditions not yet met: to deter today, honest users would post about twice what an attacker could take.

TON618 CAPITAL RESEARCH · AS OF 16 NOVEMBER 2026



A cost the attacker cannot steal.

A puzzle can be solved on someone else's machine.
A deposit has to be the attacker's own money,
and it deters only when it is at least what he stands to gain.

Defending a computer system has always cost the defender more than attacking it costs the attacker. Filters, passwords, monitoring and staff are the defender's bill; the attacker pays for a connection and his time. Artificial intelligence is now cutting the attacker's bill further, and a growing body of reporting from the AI laboratories and from governments says so. One proposed answer has a long history and new advocates: make the attacker pay, at a gate built on Bitcoin. A 2023 thesis from MIT, Jason Lowery's *Softwar*, argued that Bitcoin's real significance is as a way to impose physical cost on adversaries through the internet. In April 2026 the commander of US Indo-Pacific Command told the Senate that Bitcoin, through its proof-of-work, "imposes more costs than just the algorithmic securing of networks". Anthony Scaramucci, the founder of SkyBridge Capital, asked in April what bitcoin would be worth as a global requirement for cybersecurity against AI, and in a post dated 15 November, as reproduced on a third-party account, called it a white swan in the making, writing that AI had made hacking free and that bitcoin would make it expensive. This note sets out how such a gate would work, as a fee or as a deposit the attacker forfeits, and where it would add value: for whom, at what size, against which attacks, and how much bitcoin it would hold if it spread. It states no view on the price of bitcoin.

THE QUESTIONS

What this note answers

1. **Where does the idea come from?** The thirty-year history of charging for access, what *Softwar* proposes, and what it leaves for a builder to supply.

2. **How much cheaper has AI made attacking?** The published dollar costs, per attempt and per success, and the evidence that the fall is overstated.
3. **What already charges for access, and in what?** Twenty systems, from puzzles to fees to deposits, and which of them use bitcoin.
4. **What would a deposit have to do to deter an attacker?** The decision rule, the review window, and what both cost the honest user.
5. **Where does it work?** Six uses, each seen by the operator, the honest user and the attacker.
6. **Why bitcoin and not a dollar token?** Including the security of the ledger the gate stands on.
7. **How much bitcoin could it hold, what would have to be true for that to be large, and what would change this note's answer?**

Holds in part

JUDGED BY WHETHER THE SYSTEM IS BETTER WITH THE GATE: BY THE ARITHMETIC IT WOULD DETER IN TWO OF SIX USES AMONG USERS WHO CAN PAY AND, BY REASONING FROM THE MECHANISM, REDUCE DAMAGE OR LOAD IN THREE AND HELP SLIGHTLY IN ONE. THE USE THAT WOULD HOLD THE MOST BITCOIN, TARGETED INTRUSION, IS ONE WHERE IT ONLY REDUCES

The read. A cost the attacker must pay in his own money is a real defense, and bitcoin is one of the few ways to charge it to a stranger with no bank, card network or token issuer in the middle. The claim holds, in a narrower scope than its proponents give it. The test is whether a system is better with the gate than without it, net of what it costs honest users: an attack that stops paying is the best result, and less damage or load also counts. **Fees would work.** A final payment of a fraction of a cent to a dollar would remove the profit from bulk abuse; it is live in bitcoin on a small scale and in dollar tokens on a larger one. But honest users need a wallet: 10% of American adults used any cryptocurrency in 2025 and 2% paid with it, so a gate open to the general public has to keep a second route or shut out or burden most of them, and a general-public login gated against guessing with no second route would be not better. **Deposits deter when a machine can check the offense.** Where a failed attempt shows at once, a deposit of the prize times the odds of success is enough, and the largest forfeitable deposit system found works this way. **Where a person has to judge the damage, they reduce it.** A deterring deposit must exceed the attacker's whole prize, so an honest user posts about twice the prize in bitcoin to hold it through a review, and the bitcoin exchange that launched with a neutral arbiter's key removed it in 2019. An operator whose users already trust it with their balances can rule on forfeits itself, leaving them the risk of a wrongful seizure by a party that profits from it. Smaller charges still make every attempt cost coins, though not the attacker's own if they are stolen, and make breadth need capital, which squeezes the low-budget operator AI has newly equipped, not one holding stolen coins; and a gate reaches an exploit of a software flaw only where it is hard-coded in front of a private system. **The judged case is where the value is.** No access deposit holding bitcoin was found; the largest bitcoin bond for good behavior found, JoinMarket's at about $30 million, cannot be forfeited; the forfeitable bonds on bitcoin exchanges cover single trades and publish no total; and six conditions, none fully met, stand between that and a large number. The first two findings rest on arithmetic and on systems in use, though no bitcoin fee system publishes its volume; the third rests on few cases and no deployment, which is itself the finding.

$25.46

Mean cost per target scanned in one AI-run campaign against online retailers, from the operator's own records, in one vendor's interim report of September 2026. The strongest measured fact for the view that attacking has become cheap

$82 billion

Ether locked by Ethereum's validators as a deposit that is cut for provable misconduct. 594 validators have ever been slashed. The strongest fact for the view that forfeitable deposits work at scale

2.15×

The least bitcoin a user who counts in dollars must post, as a multiple of what the attacker stands to gain, to hold a deterring deposit through a 30-day review: on bitcoin's record since 2018, if every intrusion found is tied to its session and awarded (1.98× on the record since 2024). Across both records and windows from a week to a year, 1.9× to 2.8×

0

Systems found in production that take a forfeitable deposit from a user for access to a network, API or privileged interface, in bitcoin or in a dollar token. The pattern is documented; the live uses bond trades between users for hours

Terms used throughout, defined once. Proof-of-work: a puzzle that takes real computing effort to solve and almost none to check; Bitcoin uses it to order its ledger, and websites use it to slow down bots. **Denial-of-service:** flooding a system with requests until it cannot serve real users. **Botnet:** a fleet of other people's computers that an attacker controls. **Sybil attack:** one party posing as many, to outvote or out-shout the real users of an open system. **Stablecoin:** a token on a public blockchain that its issuer promises to redeem for one dollar; USDC and USDT are

the largest. **Lightning:** a network built on Bitcoin that settles small payments in about a second. **Hold invoice:** a Lightning payment the receiver can accept and then either keep or hand back. **Multisig:** a bitcoin address that needs several keys to spend; "2-of-3" means any two of three key-holders. **L402 and x402:** two ways for a web server to answer a request with "payment required" and be paid on the spot, the first in bitcoin over Lightning, the second mostly in stablecoins. **Fee:** a payment that is final. **Deposit:** a payment held and returned unless a stated condition is met. **Forfeit:** the deposit is kept. **Review window:** the time a deposit is held before it is returned. **Dwell time:** the days an intruder is inside a network before being found.

BOTH SIDES, FIRST

Both cases, each at full strength

This note's answer cuts against both the widest version of the claim and the flat dismissal of it. Each is set out here in the terms its best advocate would use, before any of it is tested.

The case that this is Bitcoin's next use.

1. **The cost of attacking has fallen where it matters.** One operator, per one vendor's interim report, attacked hundreds of online retailers in late summer 2026, with open-source attack tools driving a commercial AI model, at a mean of $25.46 a target scanned. A frontier laboratory's model built a working exploit for a known flaw for under $1,000 in half a day. The laboratories now write that the economics of attacks have changed, and five governments' cyber agencies said jointly in June 2026 that the timeline "is not years, it is months".
2. **Every existing defense is paid for by the defender.** Filtering, monitoring and identity checks raise the defender's costs, and a puzzle can be solved on machines the attacker stole or rents by the hour. The only cost that lands on the attacker himself is money he has to part with.
3. **Bitcoin is the largest asset that can charge that cost to anyone, with nobody in the middle.** A card payment can be disputed for 120 days, and for some disputes up to 540, and a card hold may not be used as a damage bond. A dollar token has an issuer that can freeze the operator's takings. A bitcoin payment is final and has no issuer.
4. **The gate stands on the ledger with the most computation behind it.** About 985 quintillion hashes a second of computation secure bitcoin's ledger, and no single party holds a key that can change its rules, pause it or freeze a coin. A dollar token's rules can be changed: its issuer holds keys that mint, freeze, pause and upgrade it, and the ledger beneath it is secured by the host chain's stake, not by work. A deposit is only as safe as the ledger it sits on, and a gate meant to stop the best-funded attackers should sit on the one with the longest record.
5. **Forfeitable deposits already secure very large sums.** Ethereum's validators have about $82 billion of ether locked that is cut for provable misconduct. Systems built on Ethereum hold claims for about seven days under bonds that are lost if the claim is proved false. On Bitcoin, peer-to-peer exchanges bond every trade over Lightning.
6. **The value scales with what is guarded.** A fee is spent and resold. A deposit sized to the asset it protects is held, so demand for the coin would grow with the value of the networks behind the gate.
7. **It only has to pay for itself.** No security layer stops every attack. A deposit earns its place if the losses it prevents and the forfeits it recovers exceed what honest users give up while their money is held.
8. **A cost in energy cannot be faked by whoever writes the software.** This is Lowery's own argument. Rules enforced by code, including money put at stake under code, are only as trustworthy as the people who control the code. Proof-of-work is the one cost a program cannot counterfeit, which is why he rejects stake and builds on Bitcoin.
9. **Serious people are looking.** A US combatant commander told the Senate in April 2026 that Bitcoin has "really important computer science applications for cyber security".
10. **Users need not run a wallet of their own.** Wallets built into apps people already use can pay: Cash App's help page says it pays a Lightning invoice scanned with its camera, from its cash balance (with a small processing fee) or a bitcoin one, up to $999 every 7 days. And the first services to gate are those whose users already hold bitcoin.

The case that it is not.

1. **Dollar tokens do the same job, and do one thing bitcoin cannot.** USDC and USDT can be held and sent without an account, and settle finally. The payment standard built for software agents has recorded more than 138 million transactions on one chain alone, in stablecoins. When funds are stolen, Tether freezes them and reissues them to their owners or to law enforcement, about $1.2 billion so far by its own count. A deposit in dollars does not lose a quarter of its value in a bad month.
2. **The controls that exist are cheap and working.** Multi-factor authentication cut the risk of account compromise by 99.22% in Microsoft's own study of 2022 accounts. Filtering a million web requests costs sixty cents. Password attacks fell 26% in a year.
3. **Intrusions are found late.** More than a third of intrusions investigated in 2024 were not found within 30 days, espionage cases took a median of 122 days in 2025, and one survey puts the mean time to identify a breach at 183 days. A deposit that is returned before the damage is found deters nothing.
4. **Nobody does it.** No production system was found that takes a forfeitable deposit for network access. Where forfeiture was designed in, builders have retreated from it: a bitcoin peer-to-peer exchange removed its arbitrator's key in 2019, and a messaging network that specified forfeiture shipped without it.
5. **The cost collapse is overstated.** The United Kingdom's cyber agency judged in November 2026 that human operators "highly likely remain essential". The best model failed seven attempts in ten on an undefended test network, at a derived cost of $8,000 to $42,000 per success. Defenders get the same tools.
6. **Attackers increasingly do not log in, and they are not short of coins.** Exploiting a software flaw is now the most common way into a breached organization, at 31%, and it involves no session to attach a deposit to. Thieves took more than $3.4 billion of cryptocurrency in 2025.
7. **Deposits are lost around the ledger, not on it.** The most recent theft from a bitcoin deposit system on this note's record, 11.59 bitcoin from ten Bisq users in May 2026, came through a missing check in the code that builds the trade's deposit transaction; an earlier one, in April 2020, took about 3 bitcoin and 4,000 monero from seven users through a flaw in the trade protocol. Ledger security does not protect the gate's software, wallet or keys.
8. **Few people can pay.** In 2025, 2% of American adults paid for anything with any cryptocurrency, and 10% used it at all.

Of these, this note tests the cost evidence on both sides, and whether attackers are in fact going around the login and what a gate in front of the whole system changes (§2, §4); where deposits have actually been lost, against Lowery's objection to stake enforced by code (§1, §3, §6); what card rules and token issuers actually permit, and what the ledger's security does and does not protect (§3, §6); the record of twenty systems that charge for access (§3); the size of deposit that deters, the gentler test of whether it pays for itself, and what both cost honest users (§4); each use in turn, including how well the existing controls are working and how many users can pay (§5); and the scale (§7). **It cannot test** whether the fall in attack cost continues, which is a forecast; whether users and operators would accept the friction, because there is no deployment to observe; how far wallets built into apps would close the gap in who can pay, which no primary count measures; how a regulator or court would treat the deposit, which the governing texts leave open; what a state-backed attacker would budget for forfeits, on which no primary source exists; the thesis's larger claims about war and national strategy, which are outside a note about access control; or what Lowery has argued since 2023, which was not found in any primary source.

THE IDEA

An old idea: make the sender pay

1992 the year a computational charge for access was proposed, as a substitute for a payment that did not yet exist

The idea that abuse is a pricing problem is as old as the commercial internet. In 1992 Cynthia Dwork and Moni Naor proposed fighting junk e-mail by making each sender compute "a moderately hard, but not intractable, function" before a message would be accepted. Their paper considered charging money first, and set it aside: usage fees, they wrote, "may

be a deterrent", but they did not want a note between friends to cost as much as a postage stamp. Computation was the stand-in for a payment that could be small, automatic and universal, and no such payment existed. Adam Back built the working version in 1997 and called it hashcash; his 2002 paper is titled "Hashcash - A Denial of Service Counter-Measure". Satoshi Nakamoto's Bitcoin paper cites it, and eight days after Bitcoin's first release, in January 2009, Nakamoto wrote that the software "can already be used for pay-to-send e-mail".

The puzzle version was tested against real data in 2004 and lost. Ben Laurie and Richard Clayton worked out how hard the puzzle would need to be to make spam unprofitable, then how hard it would need to be once spammers "steal processing cycles" from infected home computers, and compared both with the mail a large British internet provider's customers actually sent. Against a million infected machines the puzzle needed to take 346 seconds, and at the difficulty levels that would stop spam, 0.13% to 1.56% of legitimate senders could no longer send their normal mail on a daily count, and 1% to 13% on an hourly one. A 2006 reply conceded the central point, that "a uniform POW mechanism will not work", and rescued the idea only by adding reputation, which is to say identity.

The money version was tried in the same years. Researchers at Michigan and Boston proposed an "attention bond" in 2004: the sender posts a bond with each first message, and the recipient "may, at her sole discretion, seize the bond or release it". A company called IronPort ran one. Its Bonded Sender program had e-mail senders post $250 to $4,000 or more at an escrow bank and debited $20 for each complaint above one per million messages. Microsoft used the list for Hotmail from late 2003. Two features of that program matter for what follows. The senders were identified, certified companies, not strangers. And the forfeits did not go to the people who complained: "In order to avoid conflicts of interest, forfeited bonds are distributed to an independent, disinterested non-profit organization."

What *Softwar* says

Lowery's thesis, submitted at MIT in February 2023, is a 400-page argument that security has only ever come from the ability to impose physical cost, that software has lacked any way to do so, and that Bitcoin's proof-of-work supplies one. Its design proposal takes up a few pages. All that is needed, it says, is to "create a firewall-style application programming interface (API) which utilizes a physical cost function protocol like Bitcoin", a "proof-of-power wall", that rejects any request arriving without proof of work. The cost is tuned "by adjusting the difficulty of the hashing algorithm or requiring more proofs of power". The aim is deterrence, stated exactly: attackers are stopped "because they find it impossible to justify the physical cost of performing an attack". The thesis is candid about its standing. "As of now," it says of its closing theory of Bitcoin as a planetary-scale computer, "these ideas are nothing more than theories". Its policy chapter recommends that the United States treat Bitcoin as a security system first and money second, consider stockpiling it, and consider supporting domestic mining.

What the thesis supplies is the principle and the ledger; what it leaves to a builder is the rest of this note. The cost it specifies at the wall is computation, and the history above shows why computation alone does not carry the idea: a uniform puzzle falls lightly on an attacker with other people's machines, a case the thesis does not take up. It leaves open that the "stamp" presented at the wall could be a coin instead of fresh computation, without saying whether the coin would be paid, held or returned, to whom, or on what condition; no fee, deposit, refund or arbiter appears in it. A deposit forfeited on misuse also goes a step past the book, which argues that money put at stake under software rules cannot do what proof-of-work does (point 8 of the case above). The gate this note builds keeps the principle and the ledger and adds the missing half: a fee or a deposit, paid in bitcoin, with rules for who keeps it.

Two of the thesis's critics made points that any working gate has to answer. Jameson Lopp, a security engineer who runs a hashcash gate on his own website and reports it leaves him with "far fewer" unwanted messages, called the wall "basically hashcash" and asked what would check the proof of work: "If you guessed 'software' then you're correct!" The writer Shinobi made the sharper one in 2024: "A money that anyone can get their hands on does nothing to restrict access to a system to only specific people that you want to access that system." Payment is not authentication. It cannot keep an intruder out; it can only change what an attempt costs him.

Puzzles, meanwhile, went into wide use with no bitcoin in them. Tor added one for its hidden services in August 2023, designed to help defend against a lone attacker and "small botnets"; its designers wrote that a large botnet would need more tools. A tool called Anubis, which makes every visitor's browser solve one, spread across open-source sites in 2025 after AI companies' crawlers overwhelmed them. A security researcher, Tavis Ormandy, priced it: counting one site for every star the tool then had on its code page, solving a week's puzzles for all of them would take a free-tier cloud

machine "about 6 minutes", so "the cost of unrestricted crawler access to the internet for a week is approximately $0". The attacker of 2025 does not need to steal the computers. He owns a data center.

So the principle holds, and computation alone does not deliver it. Money can, and the rest of this note shows what happens when the cost at the gate is a bitcoin payment or deposit.

THE COST OF ATTACKING

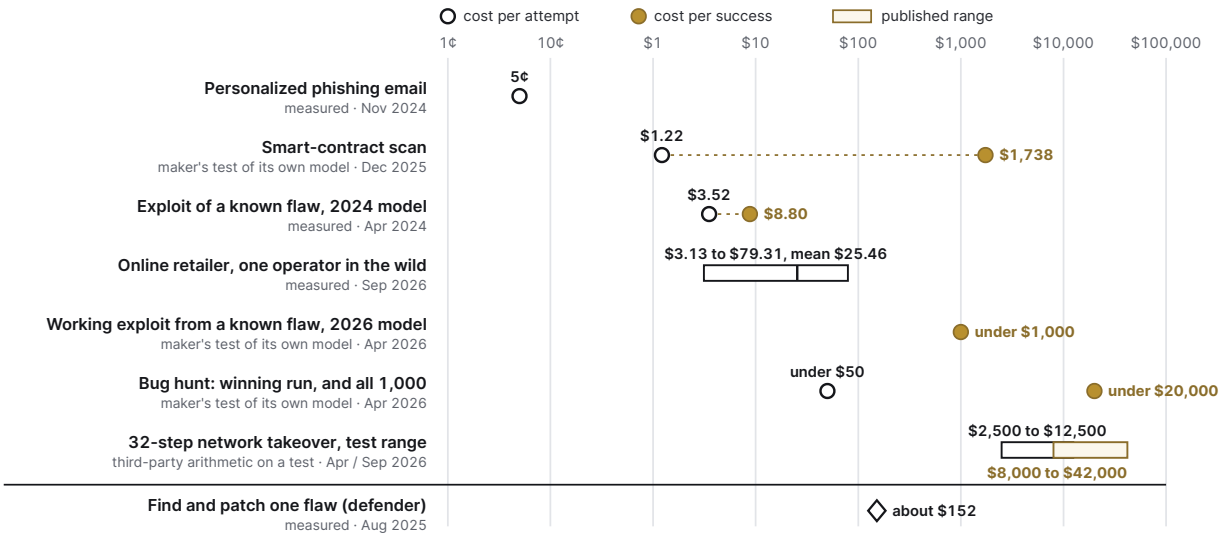
What AI has done to the cost of attacking

2.5× to 1,400× how much more a success costs than an attempt in the studies that report both: 2.5 to 3.3× where three in ten or more attempts succeed, about 1,400× in a laboratory test where successes were rare

The claim to test is that AI is driving the cost of attacking toward zero. The published dollar figures, all of them, are in Exhibit 1.

The direction is not in doubt. A personalized phishing e-mail, researched and written by a model, costs about five cents and, in a study of 101 people, was clicked as often as one written by a human expert, 54% against 12% for a generic one. Scanning a newly deployed smart contract for an exploitable flaw cost $1.22 a run in a laboratory's December 2025 test. In April 2026 the same laboratory reported that its strongest model, which it does not sell openly, turned a known flaw into a working exploit for under $1,000, and ran a thousand bug-hunting passes over an operating system for under $20,000, finding a flaw that had sat there for 27 years. The one figure from the wild comes from a security firm that recovered a criminal's own server in September 2026: the operator's cost review showed a mean of $25.46 for each retailer scanned, card-stealing code was ordered against at least 27 named victims, the firm put the whole campaign's spend at $12,000 to $18,000, and it concluded that "the economics no longer filters anyone out". That report is one vendor's, interim and uncorroborated, and this note treats it as the best single measurement available and no more.

EXHIBIT 1 — WHAT AN AI-RUN ATTACK COSTS, PER ATTEMPT AND PER SUCCESS



How to read it. Each row is one published cost, in US dollars on a scale where each gridline is ten times the last. An open circle is the cost of one attempt; a gold dot is the cost of one success; a box is a published range; a dashed line joins the two where one source gives both. "Under" marks a ceiling the source states. Rows run from April 2024 to September 2026 and are ordered by cost, with the kind of measurement and its date under each label. The diamond at the foot is a defender's cost for comparison. Three rows are a laboratory's tests of its own model, three are measurements by others, and the last attacker row is a third party's arithmetic on a government test. The exhibit shows that attempts cost cents to tens of dollars for phishing, scans and exploits of known flaws, and thousands of dollars on the 32-step test range, and that a success costs from 2.5 to 3.3 times an attempt, where three in ten or more attempts succeed, to about 1,400 times in a laboratory test where successes were rare. In the bug-hunt row the lower figure is the one run that found the flaw, a cost the laboratory says makes sense only in hindsight, so the two are not joined.

"Toward zero" is a different claim, and the same sources do not support it. Every cost on the exhibit is positive, and the cost of a success depends on how often attempts succeed. Where four in ten succeed, as with 2024's exploits of known flaws that the model was told about, a success costs 2.5 times an attempt. Where successes are rare the gap is wide: $1.22 a scan but $1,738 for each vulnerable contract found, about 1,400 times, leaving $109 of profit per exploit. The United Kingdom's AI Security Institute found that the best model took over an undefended 32-step test network in 3 attempts out of 10; a researcher who priced those attempts at published rates put a success at $8,000 to $42,000, and observed that for that money an attacker "could easily hire multiple human operators". Google's threat group wrote in September 2026 that the cost of model access "is one of the primary barriers" for attackers, which is why they steal other people's AI accounts.

Volume is rising in some places and falling in others. One large vendor counts malicious automated traffic at 40% of what it sees in 2025, up from 37% a year earlier, and account-takeover attacks up 70% between July 2024 and July 2025. Cloudflare measured crawling by AI agents acting for users up more than fifteenfold in 2025. A study of one e-mail filter's catch judged 51% of spam to be machine-written by April 2025. Against that, Microsoft's 2026 report says password attacks are "down 26% year-over-year" and credits multi-factor authentication and passkeys. Verizon's study of more than 22,000 breaches finds stolen credentials falling to 13% of the ways in, or 16% on the previous year's definition against 22%, and says the share of breaches that begin with phishing "has barely moved". The share of Cloudflare's traffic it mitigated, as attack traffic or under customers' own rules, was 6.2% in 2025, a quarter of a point lower than in 2024. The United Kingdom's cyber agency, in its review of the year to August 2026 published on 3 November, still judged that AI increases the "speed, scale and efficiency of existing attack methodologies rather than autonomously conducting end-to-end attacks" and that "human operators highly likely remain essential". The same tools serve the defender: a US government competition in 2025 had AI systems finding and patching flaws at an average of about $152 per task.

Two facts in this evidence decide a good deal of what follows. The first is where the change is largest. It is in finding and exploiting software flaws, which Verizon now counts as the most common way into a breached organization, 31% of cases against 20% a year earlier. An attacker who comes in through a flaw never logs in. He meets a payment only where a gate stands in front of the whole system, so that nothing answers before payment (§4). The second is what kind of cost has fallen. It is first the attacker's labor and skill, the cost of trying, and the prize is unchanged. The chance that an attempt succeeds is rising too, from a lower base: the laboratory's newest model produced 181 working exploits of one browser engine where its predecessor managed two in several hundred attempts, and success rates in the studies above run from a fraction of 1% of scans to three or four attempts in ten. The difference between the cost of trying and the odds of success is the hinge of §4.

One more finding, about words. A foreseeable risk with relatively predictable consequences is called a white swan, and Anthony Scaramucci applied the label to this case in a post dated 15 November 2026, as reproduced on a third-party account. An academic writing in November 2025 had called AI security risks "grey swans", rare but more foreseeable events, because "we can see them coming but arguably don't (or won't) prepare for them". The discussion is being had by an investor, and by the laboratories, five governments' agencies and the largest incident-response firms in their own plainer words.

WHAT ALREADY CHARGES

What already charges for access

8 of 20 systems on the ledger that charge in bitcoin: four fees, three bonds or escrows on trades and one time-lock; none guards access to a network

Before asking whether a bitcoin deposit could work, Exhibit 2 sets out what already puts a price on access: twenty systems in three families, charging in computing effort, in fees or in deposits.

EXHIBIT 2 — THE LEDGER OF SYSTEMS THAT CHARGE FOR ACCESS

FAMILY	SYSTEM	WHAT IT CHARGES	IN WHAT	WHO DECIDES A FORFEIT	SCALE	BITCOIN?
Effort	Hashcash; Outlook postmarks (2007)	A puzzle per message	Computer time	No forfeit	Shipped in Outlook and Exchange 2007 as one spam signal	No
Effort	Tor hidden-service defense (2023)	A puzzle per connection, 5 milliseconds to about 1 minute	Computer time	No forfeit	Every Tor release since 0.4.8	No
Effort	Anubis (2025)	A puzzle per visitor, once a week	Computer time	No forfeit	24,200 stars on its code page; used by UNESCO, GNOME and FreeBSD sites	No
Fee	Paid Nostr relay (nostr.wine)	18,888 sats once to join	Bitcoin	No forfeit	Not published	**Yes**
Fee	Stacker News	A fee per post, ten times higher for each repeat within ten minutes	Bitcoin, or credits earned on the site	No forfeit	Not published	**Yes**
Fee	L402	A Lightning payment per request or per access token	Bitcoin	No forfeit	One named user; no volume published	**Yes**
Fee	Machine Payments Protocol, Lightning method	Per request or metered session	Bitcoin (also cards and stablecoins)	No forfeit	No volume published	**Yes**
Fee	x402	Per request	Dollar stablecoins	No forfeit	More than 138 million transactions on one chain; 87% of payments between a hundredth of a cent and ten cents	Specified in September 2026, not built
Fee	Cloudflare pay per crawl (2025)	A flat price per page, set by the site	US dollars, billed by Cloudflare	No forfeit	Closed test, being reworked into payment per use; no payment figure published	No
Deposit	IronPort Bonded Sender (by 2003)	$250 to $4,000 or more, $20 lost per excess complaint	US dollars at an escrow bank	The program's administrators; forfeits went to a non-profit	Used by Hotmail from 2003	No
Deposit	Card pre-authorization (Visa)	A hold of 5 to 30 days, by type of merchant	Dollars	Card rules; an added hold may not cover damage or theft	Universal	No
Deposit	Processor rolling reserve (Stripe)	A share of sales held from 30 days to 180 or more	Dollars	The processor, against disputes	Standard for higher-risk merchants	No
Deposit	Ethereum staking	At least 32 ether per validator	Ether	Nobody: two conflicting signatures prove the offense; part of the stake is burned	44.6 million ether, about $82 billion; 594 validators ever slashed	No
Deposit	Arbitrum and Optimism claim bonds	3,600 ether on Arbitrum; 0.08 ether per claim on Optimism; held about 7 days	Ether	A replay of the disputed computation on Ethereum	Bond sizes only	No
Deposit	UMA optimistic oracle	A bond per claim, held through a challenge period	Tokens	A vote of staked token-holders	99.8% of requests undisputed, by its own count	No
Deposit	Bisq	At least 15% of each trade, from both sides	Bitcoin	Until October 2019 an arbitrator's third key; since then no key-holder	44,795 trades and 1,193 bitcoin in 2026 to 16 November	**Yes**
Deposit	Hodl Hodl	The traded coins, held in a 2-of-3 address	Bitcoin	The platform, which holds the third key	Not published	**Yes**
Deposit	RoboSats	3% of each trade by default, on a hold invoice for the life of the trade; the escrow timer is 1 to 10 hours	Bitcoin	The coordinator; half of a forfeit goes to the other trader	Not published	**Yes**
Deposit	JoinMarket fidelity bonds	Coins locked until a date	Bitcoin	Nobody: the coins are never forfeited	460 bitcoin in 75 bonds on 16 November 2026	**Yes**
Deposit	Waku rate-limit membership	A deposit sized to message rate	Ether or tokens	Designed with forfeiture; deployed without it	Not published	No

How to read it. One row for each system found that makes a party pay, or risk paying, for access. The first column sorts them by what is charged: computing effort, a final fee, or a deposit that can be returned. "Who decides a forfeit" is the column to read down: it separates deposits lost on a fact a machine can check from deposits lost on someone's judgment. Scale is the latest figure found, with its source in §9, and "not published" where none exists; amounts in sats are hundred-millionths of a bitcoin. Bold in the last column marks the eight systems that charge in bitcoin. The ledger shows that fees and deposits both exist in bitcoin, that the large deposit systems forfeit on machine-checked facts, and that no row guards a user's access to a network with a deposit.

Read by family, the ledger says three things.

Effort is free to the well-resourced. The first three rows use no bitcoin and impose no cost that a data center notices. They filter out the laziest automation, and their own builders say so: Anubis's documentation calls its puzzle "a hack" and a placeholder for better detection.

Fees are live, small, and mostly in dollars. Charging per request is the part of the idea that already works. A paid Nostr relay charges 18,888 sats, about $12, to join and markets itself as spam-free. Stacker News raises its posting fee tenfold for each repeat within ten minutes, a rule any forum could copy. But the largest fee system with a published volume, x402, settles in stablecoins. Its governing foundation's premier members include Visa, Mastercard, American Express and Stripe, and its Bitcoin Lightning method has been a written specification since 23 September 2026 with no code in the official toolkits as of this note. Cloudflare, which said in September 2025 that its network sends over a billion "payment required" responses a day, prices pay per crawl in US dollars and bills it through its own accounts. Sixteen months after launch the product is still in closed testing, is being reworked into payment per use, and has published no payment figure.

Deposits work at scale when a machine can check the offense, and stay small or retreat when a person must. The largest forfeitable deposit found is Ethereum's, about $82 billion. A validator is slashed, losing part of its deposit and its place, for signing two conflicting messages, a fact anyone can verify from the signatures themselves; there is no judge, and what is forfeited is destroyed, not paid to a victim. The systems built on Ethereum hold claims for about a week under bond, and a disputed claim is settled by replaying the computation. On the other side of the line are the systems where forfeiture needs a human finding. Bisq, a bitcoin exchange, launched with exactly that design and removed it. "Bisq's new trade protocol moves to 2-of-2 multisig escrow from 2-of-3," its October 2019 release note says. "The third key for arbitrators was removed." Its stated reasons were that a hacked arbitrator could take every locked deposit and a dishonest one could side with a trader to defraud the other, that the role did not scale, and that holding a key made arbitrators look like financial intermediaries. Bonded Sender kept its judges and sent the forfeits to a non-profit so that nobody who ruled on a complaint profited from it. Waku, a messaging network, specified a deposit that would be seized from spammers and shipped a contract whose specification reads: "This specification does not include slashing."

Judged forfeits do survive in two settings. One is where the judge is the venue and the stakes are small: Hodl Hodl still holds a third key to every trade on its own platform, and RoboSats' coordinator decides who loses a bond worth 3% of a retail trade that lasts hours. The other is on Ethereum, where UMA settles disputed claims by a vote of its token-holders, a panel neutral by design whose bond sizes and volume were not found published. No bitcoin system was found in which a neutral party rules, over weeks, on harm done to a system it does not run, and that is what guarding a network with a bitcoin deposit would ask of a judge.

Two rows need a caution, because they are easily read as more than they are. JoinMarket's bonds are the largest pool of bitcoin found locked for good behavior, 460 coins, and they cannot be forfeited at all. Their own documentation says "the valuable thing being sacrificed is the time-value-of-money", and adds that a long-term holder can post one "essentially for free". They make a fake identity expensive; they punish nothing. And the idea of a refundable bitcoin deposit for access is not new with this note. A Lightning software company published it as developer guidance in October 2025: "a user pays a Hold invoice to start a session or request a resource. If they behave, you cancel it with zero cost to them. If they abuse it, you settle the payment as a penalty." A hackathon team built it for code contributions in June 2026, on test coins. What does not exist is anyone running it in production to guard a network.

One more entry belongs here, because it bears on the premise. On 1 May 2026 an attacker took 11.59 bitcoin from ten Bisq users, in trades of bitcoin for other coins, through a flaw in the code that builds the deposit transaction, the second time the escrow plumbing itself was the way in; in April 2020 a flaw in its trade protocol had let an attacker take about 3 bitcoin and 4,000 monero from seven users. Bisq's own review found the flaw with AI assistance "significantly faster than the rest of the team" and said it believed the attacker may have done the same. The precedent for bitcoin deposits may already have met the cost collapse this note is about.

THE DEPOSIT AND ITS RULE

The deposit, and the rule it has to beat

prize ÷ odds of forfeit the deposit that removes an attacker's profit when only successful attacks can be punished

> **How the gate would work.** The gate is hard-coded in front of the system: nothing behind it answers, not even the password prompt, until a bitcoin payment or deposit has arrived from whoever is connecting. A fee is a small final payment per attempt. For a deposit, whoever opens a session posts bitcoin first, to an address that needs two of three keys to spend: the user's, the operator's and an independent arbiter's. The transaction also carries a deadline after which the user can take the coins back alone. If nothing goes wrong, the user reclaims the deposit when the review window ends, or sooner if the operator co-signs. If the operator finds damage traced to that session, it must persuade the arbiter to co-sign a transfer before the deadline. An intruder holding a stolen password has to pay or post from his own wallet before he can type it; he cannot spend the victim's. The aim is to move risk and cost onto whoever misuses the gate, not to earn revenue, so whatever an authenticated user puts up can be returned.

That is the design under test. Its parts are older than they look: the deposit posted by whoever makes contact is the attention bond of 2004, the deadline that returns the coins is an example printed in Bitcoin's own technical standards, and the third key is the one Bisq removed. Two standards apply. The stricter asks whether the attack stops paying, and this section applies it first, because a forfeit that leaves a profit in the attack taxes the attacker without changing his decision. The gentler, which its advocates would set, asks whether the gate pays for itself: whether the losses prevented and forfeits recovered exceed what honest users give up. A third effect sits between them. Across many attackers, a charge removes every one whose margin per attempt it exceeds, so a gate that fails the stricter test against the best-funded can still cut the number of attempts and the load they put on the defender. The puzzles of §1 fit that pattern: Tor designed its puzzle to help against a lone attacker and small botnets, and says a large botnet needs more tools, and a security engineer who runs a hashcash gate on his own site reports "far fewer" unwanted messages. Fewer attempts need not mean less damage, though. A charge leaves the best-funded and those after the largest prizes (§5), so damage can fall by less than the count, and the defender's load then includes the gate itself: its payment path, its refunds and its own code as a target. Exhibit 6 keeps the stricter test and the thinning apart. Its size column is the stricter test; its rating gives credit for thinning, ranked below deterrence, most where the load comes from volume and least where the damage comes from a few.

An attacker decides by comparing what an attempt is worth with what it risks. Write the chance that an attempt succeeds as s, the prize as G, the deposit as D, the chance that a successful attack is found, tied to its session and awarded before the deadline as q, and the chance that a failed attempt is caught and forfeits as f. An attempt pays when $s \times G$ is more than the expected forfeit, $(s \times q + (1 - s) \times f) \times D$. The line leaves out the attacker's cost of capital and the cost of mounting the attempt; both would lower the deposit needed a little, and the second is the cost AI is removing. The deposit that removes the profit is therefore

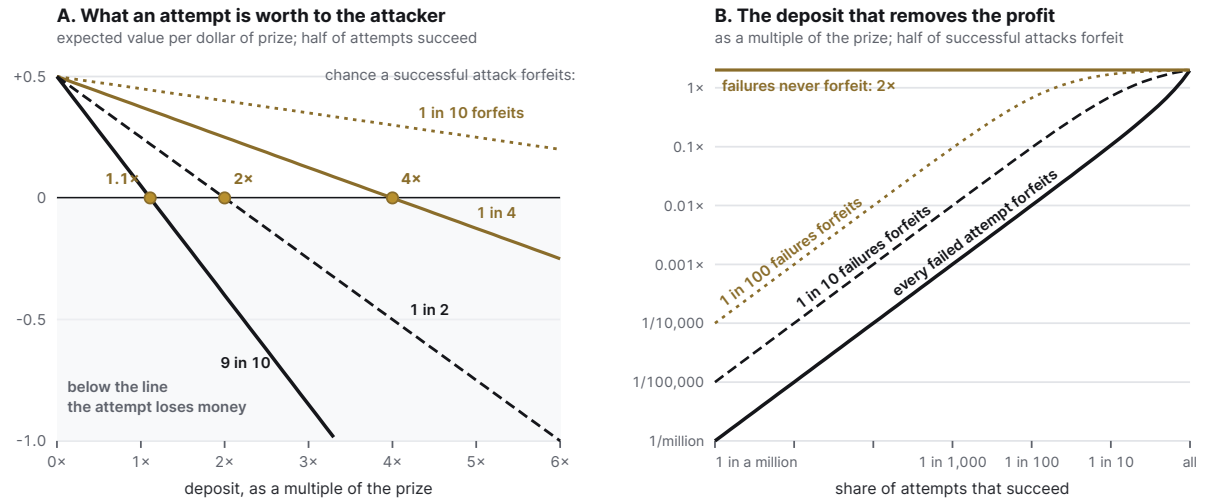
$D = s \times G \div (s \times q + (1 - s) \times f)$.

Two cases fall out of that line, and they are the whole argument.

When only success can be punished, the deposit must exceed the prize. If a failed attempt looks like ordinary traffic, f is zero, the chance of success cancels out of the line, and $D = G \div q$. An attacker who expects to gain $100,000 and to forfeit one time in two is deterred only by a deposit of $200,000. No amount of cleverness in the design changes this: the deposit has to be worth more to him than the attack, and every honest user has to post the same amount, because the gate cannot tell them apart. That is the premise of having a gate.

When failure is punished, the deposit shrinks to the odds. If every failed attempt forfeits, as a wrong password can, D is about $s \times G$. A guess with one chance in a million of opening an account worth $20,699, the average loss in a complaint to the FBI's internet-crime center in 2025, is unprofitable against a forfeit of two cents. At one chance in ten thousand the forfeit needed is $2.07. Honest users, who rarely fail, almost never pay it.

EXHIBIT 3 — THE ATTACKER'S SUM, AND THE DEPOSIT THAT BEATS IT



How to read it. Two panels, both pure arithmetic on the rule in the text, with no data. Panel A: the horizontal axis is the deposit as a multiple of the attacker's prize; the vertical axis is what one attempt is worth to him per dollar of prize when half of attempts succeed and only successes can forfeit. Each line is one chance of forfeiting, from 1 in 10 to 9 in 10; a gold dot marks the deposit at which the attempt stops paying. Panel B: the horizontal axis is the share of attempts that succeed, from one in a million to all, and the vertical axis is the deposit that removes the profit, both on scales where each step is ten times the last. Each line is one rate at which failed attempts forfeit. The exhibit shows that when failures are never punished the deposit must be twice the prize in this example whatever the odds, and that when every failure is punished the deposit can be as small as the odds of success.

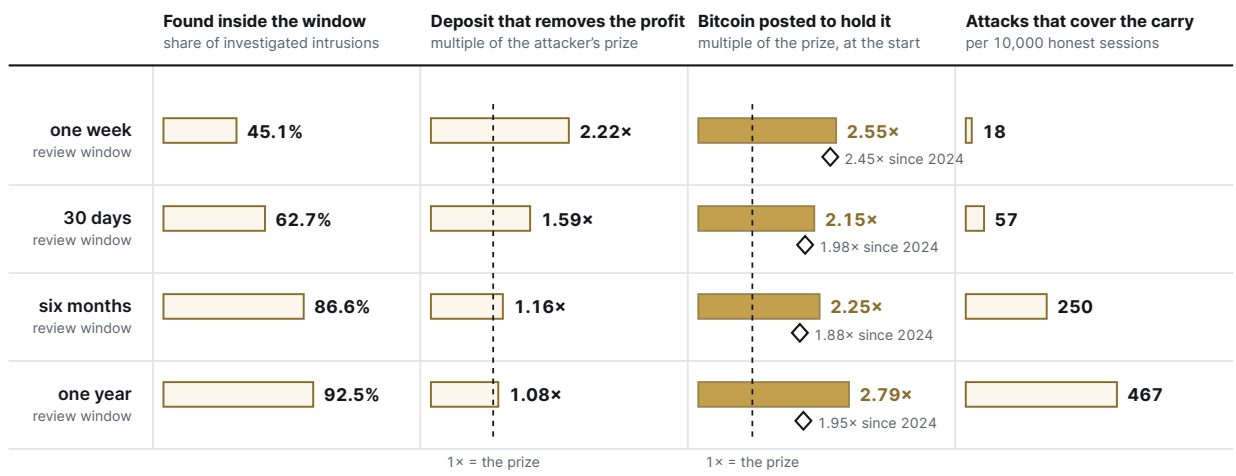
There is a third effect that the rule does not capture, and it favors the deposit. The rule assumes the attacker can fund it. The one campaign measured in the wild ran for about seven weeks on a total spend the investigating firm put at $12,000 to $18,000; in nine of those days its scanner ran against 138 hosts. It scanned public websites for weaknesses, so a session deposit would not in fact have met it; the arithmetic shows what a gate does where every target requires a session. Had each of those 138 targets held a refundable deposit of $1,000 at once, the operator would have needed $138,000 locked, about eight to twelve times his whole budget, before forfeiting a cent. A deposit turns an attack that is cheap to repeat into one that needs working capital in proportion to its breadth. That squeezes the low-budget operator AI has newly equipped. It does not squeeze an attacker holding stolen coins, of which §5 finds no shortage, or one who never passes a gate.

This is why §2 separated the two things AI changes. A falling cost of trying is exactly what a fee or a forfeit replaces: the attacker's labor is free, so charge him money. A rising chance of success is what a deposit cannot answer cheaply. Panel B is that statement as a picture. As attempts go from one-in-a-million guesses toward the laboratory results in Exhibit 1, the deposit that deters climbs from a rounding error toward the whole prize.

The review window

How often a successful attack forfeits depends on how long the deposit is held. The best public data on how quickly intrusions are found is Mandiant's record of the cases it investigates, which reports the share found within a week, a month, six months and a year. Exhibit 4 runs the rule at each of those four windows.

EXHIBIT 4 — THE REVIEW WINDOW: WHAT EACH LENGTH BUYS AND COSTS



How to read it. Four rows, one for each review window at which the dwell-time record has a boundary; four columns of bars, each on its own scale. First column: the share of intrusions investigated in 2024 that were found within the window. Second: the deposit that removes the attacker's profit, as a multiple of his prize, if every intrusion found is also tied to its session and awarded; the dashed line is a deposit equal to the prize. Third, in gold: the bitcoin an honest user must post at the start so that the deposit is still worth that dollar amount in 19 windows of 20, on bitcoin's record since 2018; the open diamond beneath is the same figure on the record since 11 January 2024. Fourth: the number of attacks per 10,000 honest sessions at which recovered forfeits would equal the interest honest users give up on their deposits, at the three-month Treasury yield of 4.32%. The exhibit shows that longer windows catch more and need a smaller dollar deposit but a larger allowance for bitcoin's price, so that the posting stays near or above twice the prize on either record.

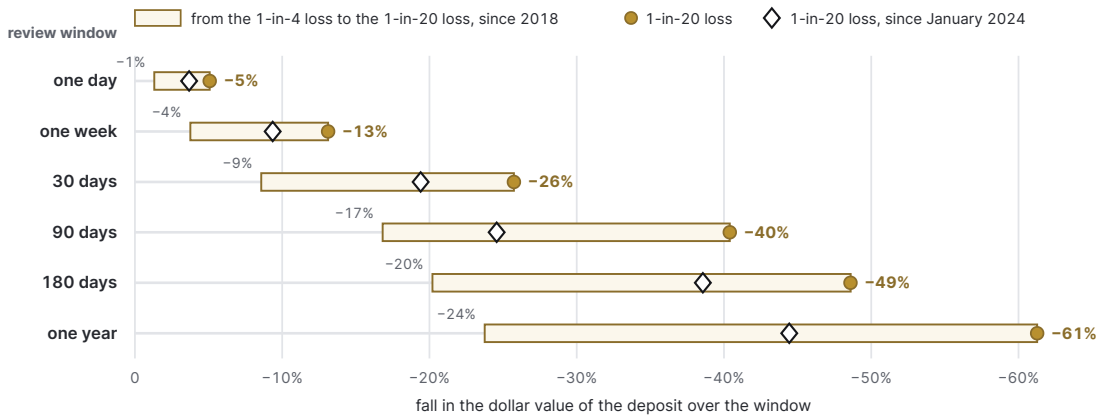
A week's hold catches 45.1% of that population, a month's 62.7%, six months' 86.6% and a year's 92.5%. The dollar deposit that deters falls from 2.22 times the prize to 1.08 times as the window lengthens. But the deposit is in bitcoin, and the longer it is held the further its dollar value can fall, so the bitcoin that must be posted to be safe moves the other way. The two effects nearly cancel. At a week the posting is 2.55 times the prize, at a month 2.15, at six months 2.25, at a year 2.79. On bitcoin's shorter and calmer record since January 2024 the same sums are 2.45, 1.98, 1.88 and 1.95. On either record there is no window at which a user who counts in dollars posts much less than twice what the attacker hopes to steal.

The fourth column is the gentler test. If recovered forfeits were the only benefit, a week's window would cover honest users' forgone interest once attacks reach 18 in every 10,000 sessions, a month's at 57, and a year's at 467. A service under constant attack clears that bar easily. A service where attacks are a far smaller share of sessions does not, and has to rely on the deterrence that the stricter test measures.

Those figures lean in the deposit's favor in three ways, and against it in one. The record counts only intrusions that were eventually found and investigated; those never found are absent. Finding an intrusion is not the same as tying it to one session's deposit and winning the award; if that second step succeeds half the time, every multiple doubles. And the record is moving the wrong way: the median time to find an intrusion rose from 10 days in 2023 to 11 in 2024 and 14 in 2025, with espionage cases at a median of 122 days. An interview survey of breached organizations gives a mean of 183 days to identify a breach; that is a different measure of a different population, and it is slower still. The one way they lean against it: a response firm reports a median of three days across 661 cases, with the fall particularly notable among the clients it monitors continuously, and an operator who logs every session for a deposit gate is closer to such a monitored network than to the average.

What the honest user bears

EXHIBIT 5 — WHAT A BITCOIN DEPOSIT HAS LOST IN DOLLARS OVER A REVIEW WINDOW



How to read it. Each row is one window length, from one day to one year. The horizontal axis is the fall in the dollar value of a deposit fixed in bitcoin between the day it is posted and the day it is returned. The bar spans from the loss exceeded in one window of four to the loss exceeded in one window of twenty, across every overlapping window in bitcoin's daily closes from 1 January 2018 to 15 November 2026. The gold dot and its label mark the one-in-twenty loss; the open diamond marks the same measure since 11 January 2024, when US spot bitcoin funds began trading. The exhibit is a record of losses and says nothing about gains or about what bitcoin will do; it shows that the tail a depositor must allow for grows from 5% over a day to 26% over a month and 61% over a year.

The honest user's costs are four, and only one is small. Interest forgone is small: at 4.32% a year, a $10,000 deposit held 30 days costs about $36. The miner fees for one cycle of deposit and refund were 20 to 39 cents on 16 November 2026, though the same cycle would have cost about $233 at the fee rate of April 2024's peak and today's price. The price risk in Exhibit 5 is the large one for anyone who measures wealth in dollars, and it is why the posting runs to twice the prize; a user who holds bitcoin anyway does not bear it. And the user needs to have the capital at all, which is the cost that excludes people.

The mechanics add limits of their own. The self-executing deadline can be set at most 65,535 blocks ahead, about 455 days. On Lightning, a deposit routed through the network can be held about 14 days at the outside and in practice hours, because the specification caps how long a payment may stay pending and every node along the route has its funds tied up, unpaid, while it does. The team that built the code-contribution bond found this out and wrote it down: "Lightning's timeout mechanics make long-lived review periods difficult to support." They fell back on simply holding the money, which is the arrangement that raises the regulatory question in §5.

No single formula

The rule above is the test a pricing scheme has to pass: over the attempts an attacker plans, what he expects to pay must exceed what he expects to gain. It is not the scheme. The operator can price access in many ways, and combine them.

- **A fee or a forfeit per attempt,** the cases above.
- **A limit in time.** Five attempts an hour on one account allow 43,800 a year, so a guess with one chance in a million takes about 23 years on average to come good, whatever is charged.
- **A price that rises with each attempt.** Start at one cent and double it with each failure on the same account: ten failures cost $10.23, twenty $10,485.75 and thirty $10.7 million. An honest user who mistypes twice puts up three cents, which, if held rather than spent, is refundable once the session proves clean, provided the refund rests on something a guesser lacks, such as a recognized device or a review window that closes with no damage found; a refund on a correct password alone would also refund a guesser who gets in. Long-odds guessing is priced out: a guess with one chance in a thousand needs about a thousand attempts, and the doubling puts that far beyond any prize. Short-odds guesses, such as a common password, stay cheap and need separate handling, for instance refusing common passwords when they are set. And the doubling works against depth, not breadth: a guesser who spreads a few tries across many accounts meets only the first price on each, which still has to pass the rule above.

- **Free at first, then priced.** The first visits from an address, a device or an account cost nothing, and the price rises with use. A public site can run this way: a reader who looks at a few pages may never reach the price, and a scraper that keeps one identity and comes back for more does. One that changes address or device each time avoids it, as below.

Each rule is only as strong as the factor it counts, because the attacker multiplies whatever is cheap to multiply. A new internet address is cheap. Residential addresses that rotate are sold at about $3 to $20 or more a gigabyte, which at an assumed 10 kilobytes an attempt is about a fiftieth of a cent at $20, so a price that rises per address resets for almost nothing. A price that rises per target account, or per resource, cannot be reset that way, because the attacker cannot multiply the account he is after, when he is after one. That has a cost of its own: a price keyed to an account lets an attacker raise what the real owner pays at the next login, so the operator has to spare a recognized device or a correct password from the increase. And no rule of this kind stops a stolen password that works the first time; that is still priced at one entry, or at nothing inside a free allowance.

The purpose of every one of these rules is to transfer risk and cost to the attacker, not to raise revenue, so the operator can return what any authenticated user puts up. The honest user's lasting cost is then having the money and a wallet to pay from, the price risk and interest forgone while it is held, and the miner fees for each cycle; where the operator rules on forfeits, also the risk of a wrongful seizure or of an intruder riding his session (§5). These are worked examples, not observed schemes: the operator chooses and combines a free allowance, a rising price, a limit in time and a deposit, matched to what the gate guards. The arithmetic says which combinations deter, not which one is right.

Hard-coded in front of everything

The strongest form of the gate is not a login page that asks for money. It is a system that cannot be reached at all without it: the first message, before any password prompt, must carry proof of a bitcoin payment or deposit, and anything else is dropped unanswered. Three things follow from the mechanism. A stolen password no longer opens the system by itself: the thief must also pay or post from his own coins. But like a flaw that works the first time, it costs him one entry, so the deposit rule above still sets what deters him. Guessing is priced per guess, the case above where every failure costs and the charge shrinks to the odds. And an exploit of a software flaw behind the gate has to pass the gate first, so the most common way into a breached organization in Verizon's count is reached too, for the systems that sit behind one.

The limits are as plain. The design suits systems that are not meant to answer the public: administrative interfaces, remote access, private interfaces for software, internal networks, custody systems. A public website can answer first visits free and price what follows, which holds against a visitor who keeps one identity and not one who changes address each time; but a flaw or a stolen password that works on a free first visit meets no price, so the gate in front of everything stays a design for private systems. And there is a cheaper way to hide a private system: answer only known parties, through a private network or a list of allowed addresses. The gate adds value where the operator cannot or will not know in advance who is connecting, the same place §5 finds deposits strongest. The gate's own code becomes the thing to attack, so it has to be small: checking a Lightning payment is one hash of the preimage compared with the invoice's payment hash, and a deposit is checked against the bitcoin ledger itself, which is where that ledger's security (§6) enters, because the gate takes its answer from the ledger, not from the party connecting. These consequences are reasoned from the mechanism; no such gate is in use to observe.

Set by value and time, with no judge

A deposit does not need a machine to check the offense, or a neutral arbiter. The operator can set its amount and its hold from the value behind the gate, and keep the decision to forfeit for itself. Both choices change what the gate asks of its users.

Size the deposit to what one session can reach. In the rule, the prize G can be no larger than the value a session can move. An operator who caps that value, as exchanges already cap daily withdrawals, caps the deposit with it, and needs no estimate of any attacker. That caps the prize only where it is money the session can move. A session into an administrative interface or internal network, the systems the hard-coded gate suits, reaches data and a foothold that no withdrawal limit bounds, and there the ransom-sized prizes of §5 apply. Take a session that can move at most $10,000 and a 30-day hold. If the operator sees a loss at once, as it would a withdrawal it did not expect, the deposit equal to the cap deters, and the bitcoin to post is about $13,470 on the record since 2018, or $12,410 since 2024. If the loss is found only as often as Mandiant's cases are found in 30 days, the deposit is $15,950 and the posting $21,480, or

$19,790. The rule is unchanged; what changes is that the operator, not the attacker, sets the size of the prize, where the prize is money.

Let the operator decide. The attention bond of 2004 had no judge: the recipient "may, at her sole discretion, seize the bond or release it". Here the honest user bears the risk of a wrongful seizure, and that is the trade the neutral arbiter was meant to avoid. It is a small trade where the user already trusts the operator with more than the deposit: an exchange's or custodian's customers, whose balances already sit under the operator's control. For them a deposit the operator can keep adds less trust than it would for a stranger, but it adds some: a custodian owes a balance back, while an operator that rules on a forfeit keeps what it rules. Two of the reasons one bitcoin exchange removed its arbiter carry over. A stolen operator key reaches every deposit at once, and a party with a stake in the outcome can rule for itself, which is why the e-mail bond program sent forfeits to a non-profit (§1). A sole power to keep the deposit is also closer to what the model state law calls control (§5).

Or lock it and never forfeit. A deposit that is only locked for a set time costs its holder the interest, about $36 a month on $10,000 at 4.32%, plus the price risk if he counts in dollars. It removes no attacker's profit. What it does is the working-capital effect above: an attacker holding a hundred sessions open at once needs $1 million locked for as long as the gate says. That is a limit on breadth and duration, scaled to the value guarded, and it binds hardest on the low-budget operator.

None of this shortens the time it takes to find an intrusion, which is set by the attacker and the defender's monitoring, not by the gate. It does remove the judge from the cases where the user already trusts the operator, and it makes the deposit a number the operator sets from its own books.

SIX USES, THREE PLAYERS

Six uses, three players each

2 deter · 3 reduce · 1 slight six uses, each rated on whether the system is better with the gate: deterring the attack ranks first, reducing its damage or load second

Each use below is rated on three tests. Can the deposit or fee beat the attacker's sum and still be small enough for honest users? Can a machine check the offense, or must someone judge it? Is a cheaper control already doing the job? The rating then asks what an operator would ask: is the system better with the gate than without it, net of what it costs honest users? DETERS, the best result, means the charge makes the attack stop paying at a size small enough for honest users who can pay it. REDUCES means it does not, but still thins the attacks or cuts their load or damage, which also has value. SLIGHT means the gain is small against what is already in place, or recovers only a small part of the harm. NOT BETTER means honest users' costs and the gate's own running cost would exceed what it removes. No use is rated NOT BETTER as a whole, on reasoning from the mechanism and not measurement, though use 3 would be for a login open to the general public with no second route. REDUCES has no size test of its own, because no gate is deployed to count the attempts it removes; §8 states the count that would test it.

Each use is set out as a proposal, then a test: a box with the proposed design, then the arithmetic and the three players, the limits, how far each limit might be overcome, and the rating. The designs are the house's own, assembled from parts in use elsewhere; none is deployed in the form proposed here.

EXHIBIT 6 — SIX USES, RATED

USE	MECHANISM	BEATS THE ATTACKER'S SUM AT AN AFFORDABLE SIZE?	OFFENSE CHECKABLE BY A MACHINE?	CHEAPER CONTROL IN PLACE?	IS THE SYSTEM BETTER?
1. Spam, fake accounts, bot sign-ups	Fee per message, or a deposit of about a dollar per account	Yes for a fee of a few cents a message: up to 1.1 cents on spam response rates from 2002, about 5 on an assumed 4.5-fold AI lift, taken from clicks in one	Not needed for a fee	Human tests are solved for a tenth to three-tenths of a cent	DETERS · WITH A WALLET

USE	MECHANISM	BEATS THE ATTACKER'S SUM AT AN AFFORDABLE SIZE?	OFFENSE CHECKABLE BY A MACHINE?	CHEAPER CONTROL IN PLACE?	IS THE SYSTEM BETTER?
		study, and under a thousandth of a cent on the one measured conversion, from 2008; a dollar per account is not tested against what an account is worth			
2. Floods, scrapers and software agents	Fee per request past a free allowance	Yes for requests that reach the application	Not needed for a fee	Yes for volume floods: filtering is free below the payment layer	REDUCES
3. Password guessing and probing	Deposit per session; a rising slice kept per failure	Yes while most attempts fail; the size grows with the attacker's success rate	Yes: a failed login is a fact	Yes where identity exists and it is switched on: multi-factor authentication	DETERS · GUESSING · WITH A WALLET
4. Fake identities in open systems	Self-held bond, weighted by size and term	Yes as a cost per identity	Only if nothing is ever forfeited	Identity checks at $1.50, at the price of anonymity	REDUCES
5. Targeted intrusion into high-value networks	Gate in front of the whole private system; modest deposits and fees	No: a deterring deposit is five to six figures per session against a ransom-sized prize; the proposed modest charges thin, not deter	No: damage is a finding, not a fact	The largest way in involves no session unless a gate fronts the whole system	REDUCES
6. Compensation of victims	Forfeits paid to a pool	No: a deterring deposit covers 0.4% to 8.5% of an average breach	No	Insurance covers about 1% of losses	SLIGHT

How to read it. One row per use: three test columns and a rating. The first test is the rule from Exhibit 3, applied with the published figures in the text below and, for guessing, with illustrative success rates. The second asks whether forfeiture can rest on a fact a computer can verify; the third names the control the gate would compete with. The last column says whether the system is better with the gate, net of honest users' costs: "deters" where the attack stops paying, the best result; "reduces" where it does not but still thins attacks or cuts their load or damage; "slight" where the gain is small against what exists or only a small part of the harm is recovered. The ratings are the note's judgment, reasoned and not measured: "reduces" carries no size test in the table (§8 sets the observed fall that would support it), and "not better", unused for any whole use, applies under §5.3 to a general-public login with no second route. Ratings fall from deterring to reducing where the prize is large or the damage needs a judge.

1. Spam, fake accounts and bot sign-ups

Proposed design: a priced front door.

THE PROBLEM	Bulk abuse runs on attempts that cost nothing; the human test meant to stop it is solved for a fraction of a cent.
GATE	At sign-up, and on each message or post above a free allowance.
CHARGE	A deposit of about $1 at sign-up; a fee of one to five cents a message past the allowance (the upper end rests on an assumed AI lift in purchases).
ESCALATION	The allowance and the price are counted per bonded account, never per address, so opening accounts costs a dollar each and rotating addresses buys nothing.
REFUND	The sign-up deposit returns after a few days of clean use; message fees can be credited back to accounts in good standing, since the aim is to charge abusers, not users.
LEGITIMATE USERS	Most never pass the allowance. Users without bitcoin keep the existing route, a human test or an identity check, so the gate adds a choice, not a toll. An abuser can take that route too, so it stays at the human test's price unless it is the identity check. A stranger's first message to someone new carries the higher charge until a reply.
PLATFORM	Spam and fake accounts stop paying; forfeits from accounts its filters or users mark as abusive, under published rules, cover its handling.
ATTACKER BEARS	A dollar locked per fake account, lost on abuse, and a cent or more per message from his own coins.

A service asks for a tiny payment with each message or sign-up, or a deposit of about a dollar returned after a few days of good behavior. Bulk abuse depends on each attempt being nearly free. At the response rates and sale values that Laurie and Clayton cite, a spammer earns between a tenth of a cent and 1.1 cents a message, so a fee of one cent ends the business at three of the four combinations and two cents at all four. Those response rates are from 2002, before AI. If personalized AI-written messages turned readers into buyers 4.5 times as often as generic ones, the multiple by which they were clicked in one 2024 study, which assumes purchases move with clicks, the same arithmetic gives half a cent to

5 cents a message: one cent ends the business at two of the four combinations and five cents at all four. The one measured conversion points the other way: a 2008 study that infiltrated about 1.5% of the Storm botnet and watched its pharmacy spam counted 28 sales from almost 350 million messages sent, under a thousandth of a cent of revenue per message sent (the revenue the sales would have brought, as the study's site took no payments), about 150 times below the lowest 2002 case. A fee of a few cents covers both. The human test that sign-up forms rely on is already for sale: a solving service lists the most common one at $1 to $2.99 per thousand, the same price human solvers charged in 2010, and researchers in Zurich solved 100% of its image tests with standard image-recognition models in 2024. A one-dollar deposit is 334 to 1,000 times that hurdle.

PLAYER	WHAT THEY DO	WHAT IT COSTS THEM	WHAT THEY GAIN OR LOSE	HOW A RATIONAL PLAYER RESPONDS
Operator	Prices messages or sign-ups; returns deposits on a timer	A payment integration; some lost sign-ups	Spam and fake accounts stop paying; fee income	Sets the fee just above the abuser's revenue per attempt, and waives it for users with history
Honest user	Pays a cent, or parks a dollar for a few days	The fee; needing a wallet at all	A service with less junk in it	Pays if the service is worth a dollar; leaves if a rival is free
Attacker	Must fund every attempt	A cent to a dollar each, times millions	Loses the volume model; keeps targeted fraud	Shifts to fewer, better-aimed attempts, or to services that do not charge

The attacker adapts by going narrow: a paid, personalized message to a chosen victim is still worth five cents plus a one-cent fee. The condition on this use is not economic but practical: honest users need a way to pay. A software company told Cloudflare it saw conversions fall by more than half when it asked for a credit card, and a bitcoin wallet is a higher bar than a card for most people. Among users who already hold coins the use is live today. A free allowance that prices only heavy use eases the condition for honest users who stay inside it, but holds against bulk abuse only if it is counted on something the abuser cannot multiply cheaply: an allowance per new account or per address is free volume for a spammer who opens more accounts or rotates addresses (§4).

How the limits might be overcome, and what the tests show. The limit is the honest user's wallet. The nearest primary count is the Federal Reserve's survey of American adults: in 2025, 10% used any cryptocurrency, 9% held it as an investment and 2% used it to buy something or make a payment, shares the survey says its online sample may overstate; bitcoin, and Lightning within it, can only be a part of those shares. A gate with no second route would shut out or burden about nine in ten American adults today, which is why the design keeps one, though that route stays open to the abuser at its old price unless it is an identity check. How far the wallets already built into apps would close the gap is untested. Targeted fraud, the attacker's retreat, can be priced where it starts, with the first message to someone who has never replied costing more than a reply, as the attention bond of 2004 proposed. The price that removes a targeted fraudster's profit is his chance of success times the prize: against the $20,699 average loss in the FBI's 2025 complaints, $2.07 at one success in ten thousand first messages and 21 cents at one in a hundred thousand. A well-aimed message succeeds far more often; at one in a hundred the price is $207. Honest strangers writing for the first time would pay it too, so the design holds only if the charge comes back when the recipient replies, the release the attention bond left to the recipient. On the arithmetic, the priced first message holds against fraud sent at volume and long odds; it weakens where odds are short, and the refund on reply also returns the charge to any fraudster whose target answers.

Is the system better? Deters, for those who can pay. Bulk abuse stops paying at a few cents a message, up to about five on an assumed 4.5-fold AI lift in purchases, taken from one click study. The open condition is that honest users can pay; the allowance and the second route ease it, and the second route stays open to the abuser at its old price unless it is an identity check.

2. Floods, scrapers and software agents

Proposed design: payment required at the door.

THE PROBLEM	Crawlers and software agents fetch by the million at the publisher's expense, and request floods exhaust applications.
GATE	In front of the application: an unpaid request past the allowance gets a price, not a page.
CHARGE	A published price per request for agents and crawlers; free first visits for everyone else.
ESCALATION	Signed-in users and agents with keys get a per-account allowance and, past it, pay a flat rate; anonymous traffic is counted per address, with the price doubling on each reuse past the allowance.
REFUND	A signed-in user's deposit is returned; a crawler's payment is a sale, the one case where the charge is meant as one.
LEGITIMATE USERS	People browsing inside the allowance never pay; anyone behind a shared address can sign in to leave the address count. Signing in trades anonymity for the allowance.
PLATFORM	Machine traffic is paid for; a refusal usually costs far less to send than a page costs to serve. Floods that never ask to be served still need ordinary filtering.
ATTACKER BEARS	A scraper pays per page or is refused; a request flood needs addresses in proportion to its volume or a price that climbs exponentially.

A server answers each request with a price and serves it when paid. For crawlers and software agents that fetch pages by the million, this turns an uninvited cost into a sale, and it is the one use with large companies building it. The defender's alternative is cheap: filtering costs about sixty cents per million requests at list price, and one large provider absorbs volume floods for free. That sets the limit of the use.

PLAYER	WHAT THEY DO	WHAT IT COSTS THEM	WHAT THEY GAIN OR LOSE	HOW A RATIONAL PLAYER RESPONDS
Operator	Quotes a price per request; serves on payment	Integration; pricing each resource	Revenue from machine traffic; less load	Charges where content has value; filters the rest
Honest user	A person browses free; an agent pays per page	Fractions of a cent per request	Access without an account or a contract	Pays where the page is worth more than the price
Attacker	A scraper must pay or be refused; a flooder sends unpaid packets	The price, for the scraper; nothing new, for the flooder	The scraper becomes a customer; the flooder is unaffected by the fee	Scrapes only what is worth paying for, or floods below the payment layer

Floods come in two kinds. A flood of page or interface requests asks to be served, and a price that starts free and climbs on each reused address changes its arithmetic. With ten free requests an hour per address and a price doubling from one cent after that, the botnet Cloudflare called the apex of botnets, an estimated 1 to 4 million infected machines, could send, if each had its own address, about 2,800 to 11,000 requests a second free, about 0.3% to 1.1% of the million a second that Cloudflare counts as hyper-volumetric. That residual is small against the largest floods, not against every server: a few thousand requests a second that each cost work to serve can still overwhelm an application without a large filter in front of it. To go further it must find addresses in proportion to its volume or pay a price that climbs exponentially on each one it reuses: thirty paid requests from one address cost $10.7 million. Renting fresh addresses instead costs about $3 to $20 or more a gigabyte (§4), which at an assumed 10 kilobytes a request is about $108,000 to $720,000 or more for an hour at a million requests a second. This is the same fiftieth of a cent per request that makes a fresh address almost free to a single guesser (§4); a flood pays it millions of times. The gate turns a request the server would have to work to serve into a refusal it can send cheaply, though sending refusals is still work, and honest users who share one address reach the price sooner; they are not authenticated, so the refund in §4 does not reach them. A flood meant only to fill the pipe is the other kind. It never asks to be served, and the reflection and amplification attacks that Cloudflare says are now the center of gravity among attack vectors forge the victim's address as their source, so there is no address to price; those packets still have to be discarded by ordinary means.

How the limits might be overcome, and what the tests show. The free allowance can be tied to something harder to multiply than an address, and one such key already exists. Cloudflare's signed agents, introduced in August 2025, let software agents prove who they are with a cryptographic signature, and the first cohort included ChatGPT agent and Block's Goose. An allowance per signed agent or per signed-in account cannot be reset by rotating addresses; an allowance per address can, at a fiftieth of a cent each (§4). Users behind a shared address escape the address count by signing in. Floods that never ask to be served stay outside any payment gate and with upstream filtering. Reasoned through, the remedy works for agents and people who sign their requests or sign in, and leaves anonymous floods to the address price, which the arithmetic above shows a large botnet can partly absorb.

Is the system better? Reduces. Paid crawling becomes a sale and request floods thin; floods that never ask to be served are untouched.

3. Password guessing and probing

Proposed design: no password prompt without a stake.

THE PROBLEM	Guessing and probing a login cost nothing per try, and multi-factor authentication was missing in 59% of one response firm's cases.
GATE	Hard-coded in front of the login and the reset process: nothing, not even the password prompt, answers without a deposit.
CHARGE	A small deposit per session, with a slice kept for each failed attempt.
ESCALATION	The slice doubles with each failure on the same account, from one cent, and attempts are limited per hour.
REFUND	Everything returns once the user is in from a recognized device or with a second factor; a password alone does not earn the refund.
LEGITIMATE USERS	A mistype costs a cent or two and comes back; failures from elsewhere do not raise the price a recognized device pays. Every user needs a wallet and a deposit before the password prompt appears; one on a new device without a second factor waits out the review window for the refund. This design suits logins whose users already hold bitcoin or whose operator funds the deposit; for a general-public login it needs a second route.
PLATFORM	Guessing becomes a losing trade, and the reset process cannot be used to go around the gate.
ATTACKER BEARS	Every guess from his own coins at a rising price; a stolen password still needs a deposit, held through the review window and not refunded without the device.

A login or an interface holds a small deposit and keeps a slice of it for every failed attempt. This is the cleanest case for the rule, because failure is a fact the machine observes. Against a prize of $20,699, a forfeit of two cents per failure makes one-in-a-million guessing a losing trade, and an honest user who mistypes a password once loses two cents.

PLAYER	WHAT THEY DO	WHAT IT COSTS THEM	WHAT THEY GAIN OR LOSE	HOW A RATIONAL PLAYER RESPONDS
Operator	Holds a small deposit; forfeits a slice per failed attempt	Integration; a refund path	Guessing stops paying; forfeits offset the cost	Sizes the forfeit at the prize times the best current success rate, and raises it as that rate rises
Honest user	Posts a few dollars; loses cents on a mistake	Occasional small forfeits	A login that attackers cannot afford to hammer	Accepts if mistakes are cheap; objects if a forgotten password becomes expensive
Attacker	Funds every guess	The forfeit on nearly every attempt	The numbers game ends	Stops guessing and buys stolen passwords that work the first time, or attacks the reset process

The attacker's adaptation is the limit. A password bought from a thief works on the first try, so it forfeits nothing; the deposit prices guessing and not theft. And where the user is known to the service, multi-factor authentication cut compromise by 99.22% in Microsoft's study of 2022 accounts. That control is far from universal: one response firm found it missing in 59% of the cases it handled in the year to October 2025, and an industry survey in 2026 found 57% of organizations still relying on sign-in methods that can be phished. So the deposit competes with a cheaper control that works where it is used, and it is strongest where the service cannot or will not know its users.

How the limits might be overcome, and what the tests show. Two remedies are already standard. NIST's password rules require a service to refuse passwords on a blocklist of commonly used, expected or compromised values, which closes the short-odds guess, and to stop consecutive failures on one account at no more than 100. The gate adds value in two places those rules leave open. The cap counts failures per account, so an attacker who sprays one guess across many accounts never meets it; there a first-attempt charge has to pass the rule on its own: 21 cents a guess at one success in a hundred thousand against the $20,699 prize, or about $207,000 to spray a million accounts, and $20.70 at one in a thousand, which an honest user on a new device would feel until the refund. And NIST warns that the cap lets an attacker lock the real user out, listing a bot challenge before login among the answers; a priced attempt serves the same aim, since running an account up to its lockout means paying a doubling price for every failure. A stolen password is met, not stopped: the refund needs a recognized device or a second factor, so a thief's deposit stays at risk through the review window, though a small deposit is little against a $20,699 prize. The gate sits alongside multi-factor authentication, not in place of it, and the reset process is gated on the same terms. On the arithmetic, the remedy holds against spraying and lockout; against a stolen password it is a charge, not a barrier.

Is the system better? Deters guessing, among users who can pay. Against a stolen password it charges one entry, from the thief's own coins, held through the review window. Where multi-factor authentication is on, the gain over it is slight;

the rating applies where it is off, as in 59% of one response firm's cases, and rests on illustrative success rates and an average-loss prize, since no measured rate per guess exists. A login open to the general public, gated with no second route, would shut out or burden about nine in ten American adults, who used no cryptocurrency in 2025, and there it would be not better. The rating holds for services whose users already hold bitcoin, such as exchanges and custodians, and for staff logins the operator funds, as in use 5.

4. Fake identities in open systems

Proposed design: a bond that never leaves its owner.

THE PROBLEM	Votes, reputation and selection weighted by headcount can be swamped by fakes that cost nothing; one firm found 149,329 malicious packages on free software registries in a single quarter.
GATE	At registration or publication.
CHARGE	Coins locked to the participant's own key until a date, with no one else holding a key.
ESCALATION	Weight grows with the size and length of the lock, so new identities start light.
REFUND	Always: the lock expires to its owner. Nothing is forfeited, so no judge is needed and no third party holds the money.
LEGITIMATE USERS	Coins never leave their control; a modest lock gives ordinary standing; holders give up time, anyone who counts in dollars bears the interest and the price risk, and those without coins to lock are priced out.
PLATFORM	A fake crowd costs capital in proportion to its size, with nothing to arbitrate.
ATTACKER BEARS	Capital locked for every fake, for the full term, with its interest and price risk.

In a system where anyone may join, and votes, reputation or selection depend on the count of participants, each identity posts a bond. JoinMarket is the working example: makers lock coins until a date, takers favor makers with larger and longer locks, and its designer estimated, for the squared design he proposed in 2019, that swamping the market would take 30,000 to 80,000 bitcoin locked for months, or 45 to 120 destroyed outright, a route the software no longer offers. Nothing is ever taken from anyone, so no judge is needed.

PLAYER	WHAT THEY DO	WHAT IT COSTS THEM	WHAT THEY GAIN OR LOSE	HOW A RATIONAL PLAYER RESPONDS
Operator	Weights participants by bond	Little; the rule is public and automatic	Fake crowds become expensive	Sets the weight so that a real participant's bond is affordable and a thousand fakes are not
Honest user	Locks coins for a period	The use of the coins; nothing if they were being held anyway	Standing in the system	Long-term holders lock gladly; others are priced out
Attacker	Must bond every fake identity	Capital, in proportion to the size of the fake crowd	Can still buy influence, at a price	Rents capital, or concentrates on systems with small bonds

The limit is in the last row. A bond that is only locked, not lost, is a cost that the wealthy bear lightly, and the documentation of the leading example says so. Add forfeiture and the bond deters more, but then someone must rule on misconduct, and the use inherits the problem of the next one. Open registries for software packages are the obvious candidate: one security firm identified 149,329 malicious packages in a single quarter of 2026, on registries where publishing is free. No bond there was found.

How the limits might be overcome, and what the tests show. Weighting meets the attacker who splits his capital and helps the one who locks a single large bond. Belcher's 2019 design for JoinMarket squared the amount locked, and his design notes say this gives rational participants "a strong incentive to lump up all their coin sacrifices together"; the software JoinMarket ships uses a gentler exponent, 1.3 by default. Capital split across a thousand identities carries a thousandth of the weight of one bond when the amount is squared, and about an eighth of it at 1.3, so weighting that rises faster than the amount discourages a fake crowd wherever influence follows weight, more strongly the higher the exponent. It also multiplies the weight of one large bond, favoring the wealthy holder, honest or not, over many small honest ones, and it does nothing where influence follows headcount, one identity one vote. Where misconduct can be shown, as with a package later found to carry malware, a registry can add a separate forfeitable bond under published rules, with its own decision to remove a package as the judge: at $10 a package, the 149,329 malicious packages one firm found in a quarter would have cost their publishers about $1.5 million had all been caught, while honest maintainers lock $10 a package, recover it, and bear the registry's removal errors. It is a charge, not a deterrent, wherever one package's prize exceeds $10 times the chance it is caught. On the arithmetic, weighting that rises faster than the amount

works against splitting and not against wealth; the forfeitable bond charges publishers where malware is found and deters only where the prize is small.

Is the system better? Reduces. Fakes cost capital in proportion to their number, and weighting that rises faster than the amount makes splitting expensive; an attacker willing to lock a great deal can still buy influence.

5. Targeted intrusion into high-value networks

Proposed design: a private system that answers only to payment.

THE PROBLEM	Intrusions into administrative interfaces, remote access, internal networks and custody systems, where ransom prizes run from tens of thousands of dollars to about a million.
GATE	Hard-coded in front of the whole private system (§4): no message is answered without proof of payment or deposit, so a probe or an exploit of a flaw behind the gate has to pay to reach it.
CHARGE	Staff sessions carry a modest deposit, below the deterring size, that the operator funds; outsiders pay a fee for each connection.
ESCALATION	Per account and per resource, with limits in time.
REFUND	Staff deposits are funded by the operator and attached to recognized devices, returned after each clean session; staff never handle bitcoin.
LEGITIMATE USERS	Employees and contractors put up none of their own money, and a known device passes without a rising price. The operator bears the price risk, the deposit a hijacked session puts at stake, and a public-ledger record of its staff's sessions. Outsiders pay final fees.
PLATFORM	Scans and exploit attempts cost coins, breadth needs capital, and every paid session leaves a record.
ATTACKER BEARS	Every probe and every attempt from coins, his own or stolen, and capital in proportion to the number of targets.

This is the use the thesis is about. Run at the size that deters, with a deposit on every session, a review window sized to the asset, an arbiter and forfeiture to the victim, the design fails the first two tests. The proposal above is built around that failure.

It fails on size. Take ransom as the prize. Verizon's 2026 data have a median payment of $139,875 with 31% of victims paying, an expected $43,361 per successful intrusion. Three-quarters of ransomware intrusions are found within 30 days, so the deposit that removes the profit is at least $57,815, and the bitcoin posted to hold it safely at least $77,877, if every intrusion found is awarded, for every session, by every honest user. On a response firm's third-quarter 2026 median payment of $400,000, at Verizon's 31% pay rate, the posting is $222,704; on its average, $570,834. On the lowest published basis, a 2025 median of $59,556 with 28% paying, it is $29,950. And ransom is the kind case, because ransomware announces itself. Against quiet theft, with detection measured in months, the multiples in Exhibit 4 apply to a prize nobody can bound.

It fails on judgment. "Damage traced to this session" is a finding, not a signature. Someone has to make it, the deposit design gives that someone a key, and the bitcoin exchange that began with that arrangement removed it because the key-holder was the weakest point. The alternative is an institutional arbitrator, whose filing fee alone is $2,500 a case at one leading provider. The operator can keep the decision itself (§4), which removes the judge where users already trust it with their balances; it does not shrink the deposit. The table plays out the deterring version.

PLAYER	WHAT THEY DO	WHAT IT COSTS THEM	WHAT THEY GAIN OR LOSE	HOW A RATIONAL PLAYER RESPONDS
Operator	Requires a five- or six-figure deposit per session; must prove damage to an arbiter before the deadline	Lost users; the arbiter; sanctions screening of every deposit; an unsettled licensing question	Deterrence of attackers who log in and expect a bounded prize	Offers it only to counterparties who can post that capital, which means institutions it could have identified anyway
Honest user	Posts about twice the largest plausible prize in bitcoin	Price risk of a quarter of the deposit in a bad month; trust in the arbiter; a record linking sessions on a public ledger	Access	Refuses, unless already a large holder of bitcoin dealing with a service that holds little else
Attacker	Chooses between the gate and a way around it	The deposit, if he logs in	Keeps every route that needs no session	Exploits a flaw instead of logging in; or posts stolen coins; or hijacks an honest user's session so that the victim's deposit is the one at risk

Every adaptation in the last cell is available now. Flaw exploitation is the largest way into the 22,000 breaches Verizon studied, and meets no gate unless one is hard-coded in front of the whole system (§4), which a public-facing system cannot have; in one response firm's 661 cases, guessing passwords by brute force was nearly level with it. Thieves took

over $3.4 billion in cryptocurrency in 2025, $2.02 billion of it by North Korea's state hackers, so coins to post are not scarce, and a deposit funded with stolen coins costs the attacker nothing he earned. Worse for the operator, United States sanctions rules allow civil liability "even without having knowledge" for dealing in property of a blocked person, so a forfeited deposit that traces to a sanctioned thief is something the operator must freeze and report, not keep. And an intruder who takes over a legitimate user's machine rides that user's session; the deposit at stake is then the victim's, and the design has moved risk onto the honest party it was meant to protect.

Whether an operator holding such deposits is a money transmitter under United States law is not settled. The Treasury's 2019 guidance says a party that merely adds a second key to a wallet and lacks "total independent control" is not one, but it describes a co-signer acting at the owner's request, not a key-holder who may end up with the funds. The bill that would have written a safe harbor failed a Senate procedural vote 49 to 50 on 15 September 2026. In the European Union, a regulated crypto business may not keep anonymous accounts from 10 July 2027. Two things point the other way. On 6 October 2026 the Treasury withdrew two proposed rules, one of which would have required banks and money businesses to verify and report their customers' dealings with self-held wallets. And the model state law defines control of a customer's coins as the power to move them alone or block them indefinitely, which one key of three is not. This note flags the question and does not resolve it.

How the limits might be overcome, and what the tests show. Each limit has a partial answer, and each can be sized. The deterring deposit is beyond honest users, so the design does not ask for it: the operator funds its own staff's deposits, and outsiders pay fees, which are final, so the operator holds no one else's money; on the house's reading, which no ruling addresses, that narrows the licensing question. For a thousand staff at $100 each, the operator locks $100,000, gives up about $4,320 a year in interest at 4.32%, and carries a fall of about $25,700 in a bad month, the one-in-twenty 30-day loss. The prize per session shrinks if the network is divided so that no one session reaches everything, and if the money any session can move is capped. Coins linked to sanctioned persons can be screened before a deposit is accepted, which OFAC's 2021 guidance describes as a best practice for virtual currency businesses; a Lightning payment shows no address to screen, so this works for deposits made on the bitcoin ledger itself. Screening catches coins already traced, not coins laundered before they are posted. A hijack from another machine is met by binding sessions to recognized devices and asking again before sensitive actions; an intruder on the recognized device itself is not, and the operator's deposit is then the one at risk. Late detection is the hardest: one response firm's cases, which include monitored clients, had a median dwell time of three days, with the decline most notable in monitored environments. If monitoring brought the median to three days, more than half of intrusions would be found within a week; the deterring deposit would then be under twice the prize and the bitcoin posting under about 2.3 times it on the record since 2018, against 2.55 on Mandiant's record. Sized where the figures allow and reasoned where they do not, the remedies make the gate cheaper to run and an attacker's attempts costlier; none makes a deterring deposit affordable against a ransom-sized prize.

Is the system better? Reduces. By the arithmetic in §4 the use would not deter, because a deposit of the deterring size is beyond honest users; on the mechanism it reduces. A smaller charge at a gate hard-coded in front of a private system would still make every attempt cost coins, though not the attacker's own if they are stolen, and make breadth need working capital (§4), thinning the low-budget attacker and leaving the best-funded and those posting stolen coins. Honest users, or the operator funding them, would pay a smaller charge, not twice the prize, but still bear the hijack, ledger-linkage and sanctions-screening costs above; no such gate is in use to measure by how much.

6. Compensation

Proposed design: forfeits to a pool, not to the operator.

THE PROBLEM	Victims carry most of the loss; about 1% of the world's cyber losses are insured.
GATE	None of its own: it collects what the other gates forfeit.
CHARGE	Forfeits paid into a victims' pool or an insurer under published rules, never kept by the operator who finds the damage.
REFUND	Not applicable; a claim against the pool is capped and open to appeal.
LEGITIMATE USERS	A wrongful claim is limited by the cap and the appeal.
PLATFORM	A gate in place is a signal an insurer can price, as multi-factor authentication is today.
ATTACKER BEARS	The forfeit, where it is upheld.

The last use is the claim that forfeits make victims whole and could let insurers price access. The arithmetic is short. The average breach in the latest industry survey cost $4.99 million. The deposits that deter a ransom-sized prize, computed above, are between 0.4% and 8.5% of that. A deposit large enough to cover the average loss would be $4.99 million per session. Forfeiture is a fine on the attacker. It is compensation only for small harms.

PLAYER	WHAT THEY DO	WHAT IT COSTS THEM	WHAT THEY GAIN OR LOSE	HOW A RATIONAL PLAYER RESPONDS
Operator	Claims the forfeit after a finding	The arbitration; the risk the coins are tainted	A few percent of a serious loss	Treats forfeits as a deterrent and buys insurance for the loss
Honest user	Bears the risk of a wrongful claim	The deposit, if the arbiter errs	Nothing	Demands a neutral arbiter and a cap, which is what an insurance contract is
Attacker	Loses the deposit if caught in time	The deposit	Keeps whatever the attack yielded	Prices the forfeit in, as a cost of doing business, wherever the prize is larger

The insurance market shows how far there is to go: about $15 billion of premium worldwide in 2025, and by the US regulators' own count only about 1% of the world's cyber losses insured. A deposit layer could give an insurer a signal, as multi-factor authentication does today. It does not replace the policy.

How the limits might be overcome, and what the tests show. A single forfeit is small against a breach, but forfeits pooled across many gates and paid out under capped, appealable rules add up, and an insurer can price the gate as a control. Pooling spreads recoveries without raising them above the 0.4% to 8.5% of an average breach computed above: the remedy changes who receives a forfeit, not how much it is, and neither makes forfeiture a substitute for insurance.

Is the system better? Slightly. A forfeit, where upheld, recovers at most about 8.5% of an average breach, and under 1% on the lowest ransom basis, and only with deposits §5.5 finds beyond honest users; net of their cost, the gain is small and not measured. What the deposits in uses 1 and 3 would forfeit into the pool is not sized by any figure here. Of the six, this use is the nearest to "not better".

BITCOIN OR A DOLLAR TOKEN

Bitcoin, or a dollar token?

138 million and $344 million transactions on the stablecoin standard for software agents; and the sum one stablecoin issuer froze in two addresses the day before the government formally listed them

Everything in §4 and §5 could be done with a dollar stablecoin, and for fees it mostly is. Where does the choice of bitcoin matter?

EXHIBIT 7 — WHICH ASSET EACH PROPERTY FAVORS

PROPERTY	BITCOIN	DOLLAR STABLECOIN	MATTERS MOST FOR
Ledger the deposit sits on	Proof-of-work at about 985 quintillion hashes a second; no single party's key can change the rules, pause the ledger or freeze a coin; two inflation bugs in 17 years on the Bitcoin Wiki's vulnerability list, one undone in 2010 when the corrected chain overtook it at block 74691, one patched in 2018 with no known exploit at disclosure	The issuer's contract: keys that mint, freeze, pause and upgrade it, all held by the issuer; beneath it, the host chain's own security	Deposits held against well-funded attackers; operators who must not depend on a key-holder
Payment is final for the receiver	Yes	Yes on-chain; the issuer can still block the receiving address	All uses
No issuer who can freeze the operator's takings or the deposit itself	Yes	No: USDC and USDT issuers both hold that power	Operators outside the issuer's jurisdiction or favor
Stolen funds can be frozen and returned to a victim	No	Yes for USDT in practice: over $7 billion frozen, about $1.2 billion returned to owners and law enforcement, by the issuer's count	Compensation; any operator who fears tainted deposits
Deposit holds its dollar value through the	No: one window in twenty loses 26% over 30 days	Yes	Any deposit held more than a day by a user who counts

PROPERTY	BITCOIN	DOLLAR STABLECOIN	MATTERS MOST FOR
window			in dollars
Escrow can limit the arbiter to "refund" or "release"	Only with added machinery; a third key can sign anything two keys agree	Yes: a published contract does exactly this	Deposits that need a judge
Tiny payments at high speed	Yes, over Lightning	Yes, on low-cost chains	Fees
Tooling and distribution for software agents	Two standards with no published volume; a third specified, not built	The dominant standard, backed by the card networks	Fees
User already holds the asset	Among bitcoin's own users	Among traders and in dollar-short economies	Adoption

How to read it. Each row is one property a fee or deposit system needs. The middle columns state plainly whether each asset has it, with the published figure where one exists. The last column names the uses from Exhibit 6 for which the property is decisive. The table compares the two as they are documented in November 2026 and shows that bitcoin's distinct properties are a ledger secured by work and the absence of an issuer or any key over its rules, and that the dollar token's are a stable value, programmable escrow and recovery of stolen funds.

Bitcoin's distinct property is the ledger itself. A gate is a promise that the money posted at it will be there, unaltered, when the review ends. Bitcoin's ledger is secured by about 985 quintillion hashes a second of computation, and no single party holds a key that can change its rules. Its record is not spotless: in August 2010 a bug let one transaction create more than 184 billion coins, a fixed version of the software was out within five hours, and the transaction was erased when the corrected chain overtook the bad one at block 74691; in 2018 a second inflation bug was found and patched, and the developers knew of no attempt to exploit it. A dollar token's rules sit in its issuer's contract. Circle's own design document lists keys to mint and destroy tokens, to pause all transfers, to block any address and to replace the contract's code, and says "Circle will control the address of all roles"; the ledger beneath is the host chain's, secured by stake, not work. For a deposit, those keys are the issuer's to use, whatever the gate's operator wants. The ledger does not secure the gate itself: the operator's software, wallet and the arbiter's key are attack surfaces of their own, on either asset. Bisq's loss in May 2026 (§3) came through its software.

The second distinct property follows from the first: no issuer stands behind bitcoin. That matters to an operator who cannot rely on an issuer's goodwill: a service in a jurisdiction the issuer's government disfavors, a privacy tool, or any business that has watched Tether freeze $344 million in two addresses on a government's word a day before the formal designation. It also matters, in the other direction, to a depositor who wants collateral that no third party can take. For those parties bitcoin is the largest candidate.

For everyone else the dollar token is the easier instrument, and one cited advantage of it turns out to be weaker than claimed while another is stronger. The weaker one is the freeze as a gatekeeper's tool. Circle's written policy is that it "will not deny access to individual addresses" except for a threat to the token's network, a legal order, or certain urgent law-enforcement, sanctions and national-security requests, so an operator who catches an intruder cannot have his funds frozen on request. The stronger one is recovery. Tether's account of freezing $1.5 billion within 30 minutes of one theft in October 2026, and of returning about $1.2 billion in all, describes a working compensation mechanism, and bitcoin has no equivalent. Card networks, the third candidate, are ruled out by their own rules: Visa's allow a hold of at most 30 days, and that only for hotels, car rental and cruises; say an added hold "must not include any amount covering potential damage, theft"; and give the cardholder, who here is the attacker, 120 days to dispute, and up to 540 for some disputes.

The design question has a twist that favors the dollar chains. The danger of an arbiter with a key is that the key can sign anything. Circle's research arm published a contract in April 2025 that confines the arbiter to two moves, release to the recipient or refund to the payer, with a fixed deadline. Bitcoin can approximate this with pre-signed transactions and an outside party who only attests to an outcome, but the standard for that arrangement says in terms that it "does not offer protection against a party colluding with an oracle".

Put together: for fees, the rails are equivalent and distribution decides, and distribution is with the stablecoin standard. For deposits held longer than a day, stability favors dollars. Bitcoin is decisive where the security of the ledger and the absence of an issuer are the point, and the natural first users are services whose customers already hold it: exchanges, custodians, Lightning and mining businesses. A session deposit in bitcoin, for access to a system whose only contents are bitcoin, has neither the price-risk problem nor the wallet problem.

HOW MUCH BITCOIN

How much bitcoin would it hold?

$30 million and $205 million the two pools of bitcoin locked today for good behavior or payment routing, against a supply worth $1.31 trillion

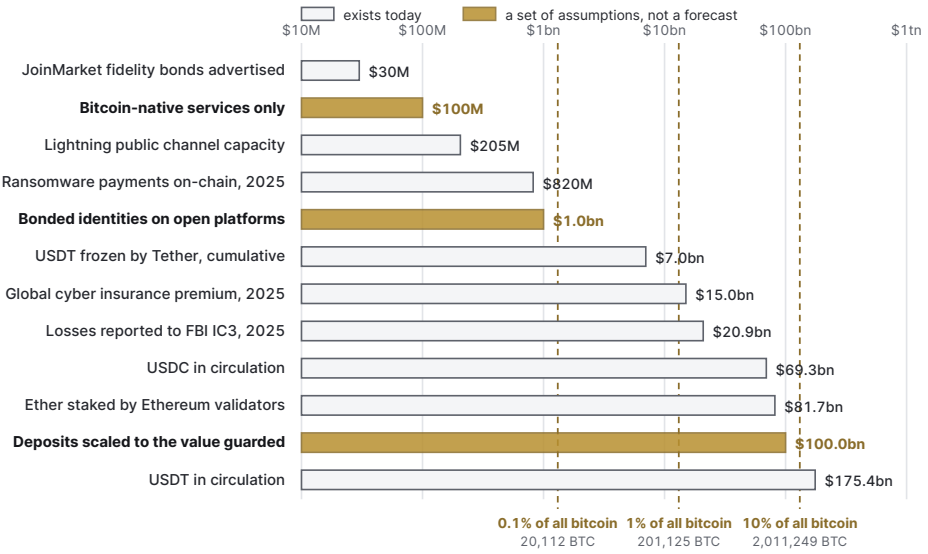
The deposit design's largest claim is that demand for bitcoin would grow with the value of what it guards. Two pieces of arithmetic size it. Neither is a forecast, and nothing here implies a price.

Fees pass through. A fee is received and, by an operator with dollar costs, sold. What it holds is a float: the yearly flow times the days the receiver keeps the coins. If every dollar of the world's cyber insurance premium, about $15 billion a year, were paid instead as bitcoin fees and held for a day, the float would be $41 million, or 630 coins. Held for a month it would be $1.2 billion. Fees are a use, and add little to what is held.

Deposits sit. The bitcoin locked equals the number of deposits outstanding at one moment times their size. Exhibit 8 sets three sets of assumptions beside the pools that exist.

The pools that exist are small. Bitcoin locked as JoinMarket bonds is 460 coins, about $30 million. The whole public capacity of the Lightning network, through which any routed deposit would have to pass, is about 3,151 coins, $205 million, and has fallen from about 4,900 coins in May. Against these, Ethereum's staked ether is about $82 billion and the two large stablecoins $245 billion.

EXHIBIT 8 — BITCOIN LOCKED IN DEPOSITS: THREE SETS OF ASSUMPTIONS BESIDE WHAT EXISTS



How to read it. Each bar is a sum of money in US dollars on a scale where each gridline is ten times the last, sorted from smallest to largest. Grey bars are pools that exist, at their latest published size; bitcoin pools are converted at the 15 November 2026 close of $65,189. Gold bars are three sets of assumptions about deposits, stated in the text, and are arithmetic, not forecasts. The dashed gold lines mark 0.1%, 1% and 10% of all bitcoin in existence, with the coin count beneath. Compare lengths against those lines. The exhibit shows that the two bitcoin pools that exist today are each under two hundredths of 1% of supply, and that only deposits sized to the value guarded reach a share of supply measured in whole percentage points.

The three assumption sets span the range. If 100,000 deposits of $1,000 were outstanding at once among bitcoin-native services, the pool would be $100 million, 1,534 coins, 0.008% of supply and about half of Lightning's capacity. If ten million standing bonds of $100 secured identities on open platforms, it would be $1 billion, 0.08% of supply. Only the third reaches whole percentage points: one million six-figure deposits outstanding at once is $100 billion, 7.6% of supply. To hold 1% of all bitcoin, $13.1 billion, takes 131,112 deposits of $100,000 outstanding at every moment, or 13.1 million of $1,000.

So the claim that demand scales with the value guarded is arithmetically right and depends entirely on the use where §5 finds a deterring deposit unaffordable to honest users. Six things would have to be true together.

EXHIBIT 9 — WHAT WOULD HAVE TO BE TRUE FOR THIS TO BE LARGE

CONDITION	EVIDENCE TODAY	STATUS
Operators adopt deposits, not only fees, for access	None found in production in any asset; the pattern is published and one prototype exists	NOT MET
Deposits are sized to the value guarded	The rule requires it; no user population has been shown to post five or six figures per session	NOT MET
Forfeiture for high-value harm can be decided without a trusted judge	Large deposit systems forfeit only on machine-checked facts; judged forfeits survive where the venue is the judge and sums are small, or on Ethereum by token-holder vote; the bitcoin exchange that began with a neutral key removed it	NOT MET
Bitcoin is chosen over dollar tokens for the deposit	Live bitcoin bonds exist for trades; the agent-payment standard and the published escrow contract are in stablecoins	PARTLY MET
The law lets an operator hold anonymous deposits	US guidance covers a neighboring case; the safe-harbor bill failed a cloture vote (49 to 50, 15 September 2026); an EU prohibition for regulated firms starts July 2027	PARTLY MET
Attack costs keep falling faster than existing controls cut success	Costs are down and automated traffic is up; password attacks are down and credential breaches are down on a changed definition; flaw exploitation is up and needs no login	PARTLY MET

How to read it. Six conditions that the arithmetic in Exhibit 8 needs before the largest assumption set could come about, each with the evidence found for it as of 16 November 2026 and a status. "Met" would mean the condition holds on published evidence, "partly met" that some of the evidence points that way, "not met" that none does. The conditions are joint: the large case needs all six. The exhibit shows three not met and three partly met.

THE VERDICT

The verdict, and what would change it

0 of 6 conditions for the large case fully met; three partly met

What holds. The insight in *Softwar* is sound and is not Lowery's alone: an attack that costs more than it is worth does not happen, and the internet has spent thirty years looking for a way to charge that cost. A final payment that needs no bank is that way, and it did not exist before 2009. Judged by whether the system is better with the gate, net of what it costs honest users, and on reasoning from the mechanism rather than measurement, the gate would be better in each of the six uses examined. Fees just above an abuser's revenue per attempt would end the profit in bulk abuse, and a deposit forfeited on a machine-checked failure would price long-odds guessing out on illustrative success rates, both among users who can pay; fees are live in bitcoin on a small scale. The gate would reduce the damage or load of floods, fake identities and targeted intrusion into private systems behind a hard-coded gate, and may recover a small part of the harm where a forfeit is upheld.

What it still needs. Computation alone, the wall as the thesis specifies it, does not deliver the idea: puzzles fall lightly on attackers with stolen or rented machines, and the puzzles in use today need no bitcoin. Money does, and the deposit design that supplies it builds on parts dating from 2004, 2019 and 2025. It needs users who can pay, or an operator that pays for them. Against a prize of real size it requires honest users to post about twice the prize, it needs a neutral judge of the kind one bitcoin exchange began with and removed, unless its users already trust the operator with their balances, it reaches attackers who never log in only where the gate is hard-coded in front of a private system, and what it recovers is a few percent of the harm.

What it means for bitcoin. As a payment rail for fees, bitcoin is one of several and behind on distribution. As collateral for deposits, it is the largest asset on a ledger secured by work with no issuer or key over its rules, and it carries price risk (Exhibit 5); its natural home is among services and users who already hold it. The holding demand from this use is real, small, and concentrated in the cases where the rule is easiest to meet, not where the value is largest. The claim that it is Bitcoin's next large use rests on six conditions, none fully met.

Qualifications. The detection figures come from one firm's cases and overstate the chance of a forfeit. The one measured attack cost from the wild is a single vendor's interim report. No deployment exists, so the judgments about friction and behavior in §5 are reasoned, not observed. The legal position is unsettled in ways that could move either way. And the evidence on AI's effect is changing by the quarter; the United Kingdom agency's scenario paper of 3 November 2026 says of its own outlook that the prospect for surprise is high.

What to watch

OBSERVABLE	WHERE IT WOULD APPEAR	BY	WHAT IT WOULD SHOW
An exchange, custodian or infrastructure provider requires a forfeitable bitcoin deposit for interface or privileged access	The provider's own terms or documentation	31 December 2027	The first condition met, in the place the note expects it first
An official x402 toolkit ships the Lightning method and a production payment processor settles it	The x402 code repository and processor documentation	30 June 2027	Bitcoin reaches the dominant fee standard's distribution
Any primary source publishes payment volume for L402 or the Lightning method of the Machine Payments Protocol	Lightning Labs, Stripe, Tempo or a named operator	31 December 2027	Whether bitcoin fees are used, not only possible
A deposit for access is deployed in a dollar stablecoin	An operator's terms; adoption of the published escrow contract	31 December 2027	The mechanism spreads without bitcoin
A second independent measurement of the cost per successful AI-run intrusion in the wild	A government agency, a laboratory or an incident-response firm	30 June 2027	Whether the $25-per-target figure generalizes
Median dwell time in the next M-Trends report	Mandiant, spring 2027	30 April 2027	10 days or fewer eases the review-window problem; 14 or more worsens it
Share of US adults who used cryptocurrency to buy something or make a payment (2% in 2025)	Federal Reserve, Economic Well-Being of U.S. Households in 2026	31 May 2027	A rise eases the wallet condition on uses 1 and 3; flat or lower confirms it
Password and credential attacks in the next annual reports	Microsoft Digital Defense Report and Verizon Data Breach Investigations Report, 2027	31 October 2027	Both falling says existing controls are outrunning cheaper attacks; both rising says the reverse
A statutory or regulatory statement on non-custodial deposit-holding	Congress, or the Treasury's Financial Crimes Enforcement Network	31 December 2027	The fifth condition resolved, either way
A published before-and-after count of attempts at a service that adds a bitcoin fee or deposit	The operator's own report or a response firm	31 December 2027	Whether a charge thins attacks in practice, and by how much
A theft from a production bitcoin access gate through its software or keys, with the ledger intact	The operator's incident report or a response firm	31 December 2027	How much the ledger's security protects in practice
Lightning's specification adopts a charge for held payments	The Lightning specification repository	31 December 2027	Short deposits over Lightning stop being a free burden on routing nodes

> ### What would change this note's view
>
> **Toward the wider claim.** The note would move its verdict from "holds in part" toward "holds" on any one of four observations, each by 31 December 2028: a production system that takes a forfeitable bitcoin deposit of $1,000 or more per session for access to a network or interface, and publishes its forfeiture record; bitcoin locked in access deposits exceeding the 460 coins locked in JoinMarket bonds today; a method of proving intrusion damage that a machine can check, in use to trigger forfeiture, which would remove the judge; or published bitcoin fee volume on L402 or the Lightning method of the Machine Payments Protocol of $1 million a day or more.
>
> **Toward dismissal.** The note would move toward "does not hold" on any one of four observations by the same date: a stablecoin deposit for access in production with at least 10,000 deposits or $10 million locked, published by its operator, while no bitcoin one exists; the stablecoin standard's cumulative published transaction count at least tripling from 138 million while bitcoin's standards publish no volume or less than $1 million a day; no production bitcoin access deposit existing and no bitcoin fee volume published; or a production bitcoin gate losing deposits through its own software or keys.
>
> **On the need, in either direction.** Evidence on the attacks themselves moves Exhibit 9's sixth condition, not the verdict on its own. If the next annual reports from Microsoft and Verizon both show password and credential attacks rising, that condition becomes met; if both show them falling, it becomes not met. A second independent measurement of the cost of an AI-run attack in the wild counts the same way: tens of dollars per target supports the condition, thousands or more per target counts against it.

> **On the ratings.** A published before-and-after count at a service that adds a bitcoin fee or deposit moves Exhibit 6, not the headline: attempts down by nine in ten or more with no rise in damage per attack support "deters" for that use, taking the fall in attempts as the observable for an attack that no longer pays, and down by half or more "reduces"; attempts down by less than half move it to "slight"; attempts unchanged, or a gate incident that costs honest users more than the attempts removed, move it to "not better".
>
> **What would not change it.** The price of bitcoin, in either direction. A further official statement that Bitcoin has security applications, without a system named. Another laboratory result on what a model can do in a test.

SOURCES

Sources and Method

Prices and supply. Bitcoin daily closes are the CoinMetrics community reference rate (`PriceUSD` , UTC close), from the first observation in July 2010 to 15 November 2026, pulled 16 November 2026. The 15 November close is $65,189.11. Supply is CoinMetrics `SplyCur` for the same day, 20,112,492 coins; market value is their product, $1.311 trillion. The ether price for the staking comparison is CoinMetrics `PriceUSD` for `eth` on 15 November 2026, $1,831.45. The opportunity cost of a deposit is the 3-month Treasury constant-maturity yield, FRED series `DGS3MO` , 4.32% on 13 November 2026.

Exhibit 5 and the haircut. For each window length the calculation takes every overlapping span of that many calendar days from 1 January 2018 (3,240 one-day spans, 2,876 one-year spans) and from 11 January 2024, computes the simple return, and reports the returns exceeded on the downside in one span of four and one span of twenty, by linear interpolation. No mean or median is computed. The haircut in Exhibit 4 is 1 ÷ (1 − the one-in-twenty loss), about 1.35 at 30 days on the record since 2018; the six-month row uses 180-day spans. The diamonds in Exhibit 4 apply the same sum to the one-in-twenty losses since 11 January 2024 (9.4%, 19.4%, 38.6% and 44.4%).

The decision rule. Exhibit 3 is algebra. One attempt is worth $s \times G - (s \times q + (1 - s) \times f) \times D$ to the attacker, ignoring his cost of capital and the cost of mounting the attempt; the deposit that sets this to zero is the formula in §4. The working-capital example multiplies the 138 hosts scanned between 23 and 31 August 2026, part of a campaign of about seven weeks, by a $1,000 deposit and compares the product with the firm's $12,000 to $18,000 estimate of the whole campaign's spend (7.7 to 11.5 times). Panel A fixes s at one-half and f at zero. Panel B fixes q at one-half and varies f over 1, 0.1, 0.01 and 0. A deposit funded with stolen coins is worth less than face value to the attacker, which raises the deposit needed in proportion; the note states this in words and does not chart it.

Dwell time and Exhibit 4. The distribution is Mandiant's "Global Dwell Time Distribution, 2018-2024" from M-Trends 2025 (published April 2025; investigations of calendar 2024), read from the page image: 45.1% found within one week, 17.6% in 8 to 30 days, 23.9% in 31 days to six months, 5.9% in six months to a year, 7.0% in one to five years, 0.5% longer. Cumulative shares are sums at the bucket boundaries; nothing is interpolated inside a bucket, and "six months" is taken as 183 days. The ransomware row is 56.5%, 18.5%, 19.6%, 2.2%, 3.3%, 0%. The 2025 median of 14 days and the median of 122 days for espionage cases (and for North Korean IT-worker cases) are from the M-Trends 2026 executive edition (March 2026); the 2025 distribution is in the full report, which is not public, so the 2024 distribution is used throughout. The break-even attack rate is 10,000 × (4.32% × window ÷ 365) ÷ the share found, and counts recovered forfeits only, with every found intrusion assumed awarded. The 183-day mean time to identify is from IBM's Cost of a Data Breach Report 2026 (602 organizations breached March 2025 to February 2026), which also gives the $4.99 million average cost. The three-day median, and the 59% of cases without multi-factor authentication, are from Sophos's Active Adversary Report 2026 release (661 response and monitoring cases, November 2024 to October 2025). The 57% of organizations on phishable sign-in is from the FIDO Alliance's 2026 survey, and the 99.22% from a Microsoft study of commercial accounts between April and September 2022.

Use-case arithmetic. Spam revenue per message is response rate × profit per response, using the 0.0023% and 0.013% response rates and the $50 and $85 sale values cited in Laurie and Clayton (2004). The guessing example uses the $20,699 average loss per complaint in the FBI Internet Crime Complaint Center's 2025 report as the prize, with success rates that are illustrative steps, not measurements; no primary source gives a success rate per guess. The intrusion example multiplies a published ransom payment by the share of victims who pay: Verizon's 2026 Data Breach Investigations Report (median $139,875; 69% did not pay), and Coveware's third-quarter 2026 median of $400,000 and average of $1,025,278, for which Verizon's 31% is used because Coveware's rate is published only as a chart. A third basis, Chainalysis's 2025 median of $59,556 with about 28% paid, gives a deposit of $22,234 and a posting of $29,950, and sets the low end of the compensation range. Each is divided by the 75% of ransomware intrusions found within 30 days and multiplied by the 30-day haircut of about 1.35 (1.347 before rounding).

Added checks. The response-rate sensitivity for use 1 applies the 4.5-fold click-through gap between AI-written and generic phishing in Heiding and others (2024) to purchases, an assumption; the measured conversion is from Kanich and others, "Spamalytics" (ACM CCS 2008; Communications of the ACM, September 2009), $2,731.88 of revenue from almost 350 million messages. The share of American adults using cryptocurrency is from the Federal Reserve's Economic Well-Being of U.S. Households in 2025 (May 2026), table 33. Both are in a saved script.

Sizing. Locked value is deposits outstanding × deposit size; fee float is annual flow × days held ÷ 365; shares are of $1.311 trillion. The three assumption sets are stated in §7 and are the note's own. Pools: JoinMarket bonds are the sum of bond outputs advertised by makers online at one public order-book watcher on 16 November 2026 (75 outputs, 460.05 coins), which omits makers who were offline. Lightning capacity is the announced-channel total from one public explorer's series (3,151 coins on 16 November 2026; about 4,900 in May); a second explorer showed 2,251 coins the same day, and private channels are invisible to both, so the figure is an order of magnitude. Staked ether is 44.6 million from a public dashboard of Ethereum's consensus layer on 16 November 2026; the count of 594 slashed validators is a single pull from one public node. USDC ($69.3 billion) and USDT ($175.4 billion) are from the issuers' own data on 16 November 2026. Global cyber premium is Munich Re's estimate for 2025; the 1% insured share is from the National Association of Insurance Commissioners' 2026 report.

Costs on Exhibit 1. Heiding and others, "Evaluating Large Language Models' Capability to Launch Fully Automated Spear Phishing Campaigns" (November 2024); Anthropic, "AI agents find $4.6M in blockchain smart contract exploits" (1 December 2025); Fang and others, "LLM Agents can Autonomously Exploit One-day Vulnerabilities" (April 2024); Gambit Security, interim report of 22 September 2026; Anthropic, "Assessing Claude Mythos Preview's cybersecurity capabilities" (7 April 2026); UK AI Security Institute evaluation (13 April 2026), with dollar figures derived by Marcus Hutchins (3 September 2026) from the token budget and published prices; DARPA AI Cyber Challenge final results (8 August 2025). Laboratory figures are tests by a model's maker of its own model. Volume figures in §2: Thales (Imperva) Bad Bot Report 2026 for automated traffic and account takeover, measured on that vendor's customers; Cloudflare's 2025 Year in Review; Barracuda's June 2025 study with Columbia and Chicago researchers. Other evidence in §2: Microsoft Digital Defense Report 2026; Verizon 2026 Data Breach Investigations Report; UK National Cyber Security Centre Annual Review 2026 (3 November 2026); Google Threat Intelligence Group's AI threat tracker (September 2026); the Five Eyes agencies' statement of 22 June 2026.

The thesis. Jason P. Lowery, *Softwar: A Novel Theory on Power Projection and the National Strategic Significance of Bitcoin*, S.M. thesis, Massachusetts Institute of Technology, February 2023, handle 1721.1/153030. Quotations are from pages 297, 298, 321, 326 and 381 of the MIT file and were checked against it word for word. The statements that the thesis contains no fee, deposit, refund or arbiter and never mentions botnets rest on a search of its full text. The Senate testimony is from the committee's stenographic transcript of the Armed Services hearing of 21 April 2026.

The discussion and the ledger's security. Anthony Scaramucci's post of 15 November 2026 on X was read in a repost, not on his own account, so it is paraphrased, not quoted; his April 2026 post is as reported by Benzinga on 24 April 2026. Bitcoin's hash rate is mempool.space's estimate for 16 November 2026. The 2010 value-overflow incident is from the Bitcoin Wiki's account and its list of vulnerabilities; the 2018 bug is from the Bitcoin Core project's disclosure of 20 September 2018. The stablecoin's controls are from Circle's token design document in its public code repository. The price of rotating residential internet addresses is Zyte's published range, read on 16 November 2026; the 10 kilobytes per attempt and the pricing rules in §4's "No single formula" are illustrative. The remedies in §5 are tested in a saved script, with success rates per attempt and the $10 publishing bond, the $100 staff deposit and the thousand staff illustrative; NIST's rules are from SP 800-63B-4 on its own site, and signed agents from Cloudflare's announcement of August 2025. Botnet size, the hyper-volumetric threshold and the shift to reflection are from Cloudflare's DDoS threat reports for the third quarter of 2025 and the first half of 2026; the flood arithmetic in §5 is in a saved script, with the allowance and the 10 kilobytes illustrative. Tor's design aim for its puzzle is from its proof-of-work specification (spec.torproject.org, "hspow-spec", motivation).

The ledger. Each row of Exhibit 2 is from the system's own documentation, code or published statistics, read on 16 November 2026, except Bonded Sender (archived copies of its 2003 and 2004 pages and a Microsoft press release of 5 May 2004). Bisq's history is from its own release note of 29 October 2019 and its post-mortem of 16 May 2026; its volumes are from its public markets interface. The x402 payment count is Chainalysis's estimate (14 October 2026) for one chain; the status of its Lightning method is from the project's code repository on 16 November 2026. Rows marked "not published" had no figure in any primary source found. The ledger is not a census. It leaves out systems that repeat a row: Mostro and Peach, which bond or escrow trades as RoboSats and Hodl Hodl do; Kleros and Proof of Humanity, juried deposits on Ethereum like UMA's; and Goodmail, a per-message fee in dollars from the 2000s.

Other sources. Anthropic's threat intelligence report of September 2026, for the laboratories' statement that "the economics of the attacks have changed"; Sonatype's open-source malware report for the third quarter of 2026; Tether's own releases on freezes and recoveries; Chainalysis's crypto-theft chapter of December 2025, in its 2026 Crypto Crime

Report; Alby's developer guidance on hold-invoice deposits (October 2025) and the sats-stamp prototype (June 2026); JAMS's fee schedule; 2Captcha's price list (16 November 2026), Plesner and others at ETH Zurich (2024) and Motoyama and others (2010) on solving human tests; Stripe Identity's price list; AWS WAF's price list; Dwork and Naor (1992), Back (1997, 2002), Nakamoto (2008, 2009), Liu and Camp (2006) and Loder, Van Alstyne and Wash (2004) for §1; Jameson Lopp's reviews and contact-form post, Shinobi in Bitcoin Magazine (2024) and Tavis Ormandy on Anubis (2025).

Legal texts. FinCEN guidance FIN-2019-G001 (9 May 2019), section 4.2.2; OFAC, Sanctions Compliance Guidance for the Virtual Currency Industry (October 2021); the Senate record for H.R. 3633 from the Library of Congress; Regulation (EU) 2024/1624, Articles 79 and 90; Visa Core Rules, 24 October 2026 edition; the Federal Register notice of 6 October 2026 withdrawing two proposed rules; the Conference of State Bank Supervisors' model money-transmission law. The note reports what these texts say and where they are silent. It is not legal advice.

Limits. The search for deployments of a deposit for access covered the published payment standards, the main bitcoin and Ethereum deposit systems and a set of web searches; a system that publishes nothing would not have been found. Several vendor reports are available only as summaries, and figures printed only in charts were not used. The attacker's adaptations in §5 are reasoned from the mechanism and from reported attacker behavior, not observed against a deposit gate, because none exists. What Lowery has said or written since 2023 was not found in a primary source, so the note tests the thesis as published.

Earlier TON618 notes relied on. The TAM of AI Using Bitcoin (June 2026), on L402, x402 and agent payments; The Market Bought Traffic. Cloudflare Bills by the Month. (September 2026), on pay per crawl; The Access List (September 2026), on the laboratories' restricted security models; The Referee Has a Book on the Game (September 2026), on cyber insurance. None is needed to follow this note.

Disclosures

Information only. TON618 Capital. This report is for information purposes only. Nothing here is an offer to sell or a solicitation of an offer to buy any security, fund interest, or digital asset, and nothing here is personalized investment advice or a recommendation regarding any instrument.

Publisher's exclusion. All research is published solely as general, impersonal information of regular circulation. It is not tailored to the objectives or circumstances of any individual and is not issued in connection with compensation from any client. The Fund has no clients and distributes all research free of charge. On that basis it publishes in reliance on the publisher's exclusion from the definition of "investment adviser" under the Investment Advisers Act of 1940 (§202(a)(11)(D); cf. *Lowe v. SEC*, 472 U.S. 181 (1985)).

Registration & conflicts. TON618 Capital is not registered as an investment adviser or broker-dealer in any capacity. The Fund is a Bitcoin fund and may hold or transact in the securities or digital assets it discusses. This note discusses bitcoin, the dollar stablecoins USDC and USDT, and products of Cloudflare (NET), Coinbase (COIN), Circle (CRCL), Block (XYZ), Stripe, Visa (V), Microsoft (MSFT), Alphabet (GOOGL), Anthropic and OpenAI; it makes no recommendation on any of them and states no view on the future price of Bitcoin. The Fund receives no compensation from any party in connection with its research.

Use of AI. Artificial intelligence is used in the creation of this research. All methodology and data integrity are reviewed and approved before publication by TON618 Capital's Chief Investment Officer, **Keyth Beck**; errors may nonetheless occur, and readers should verify independently.

CFA. This report was prepared to align with CFA Institute analytical standards (methodology only). CFA® and Chartered Financial Analyst® are registered trademarks owned by CFA Institute. That reference describes the analytical framework applied; it does **not** imply the report was prepared, reviewed, or authored by a CFA charterholder, and the report is **not** issued, reviewed, endorsed, certified, or approved by — nor affiliated with — CFA Institute.

Risk & feedback. Past performance is not indicative of future results. Digital assets and equities are volatile and may result in total loss of capital. Corrections and feedback are welcome — please direct them to CIO **Keyth Beck** at **keyth@ton618capital.com**.