KORE-TEK STRATEGIC PERSPECTIVES, NO. 6

The Attention Tax

How Tool Sprawl Turns Every Alert Into a Hidden Operating Cost

A framework for CFOs, CIOs, CTOs, and COOs to price human attention in network operations, model it over three years, and choose among the three remedies that reduce it.

OCTOBER 2026

Executive Summary

Disclosure. Kore-Tek sells co-managed NOC services, one of the three remedies compared here. Every figure comes from a stated formula and stated assumptions, and we will send the working model to any reader who asks, so the comparison can be checked without relying on our judgment.

Most organizations treat monitoring alerts as a technical nuisance. They are better understood as transactions. Each alert that reaches a person spends a measurable quantity of skilled attention, and that spend is rarely itemized on any budget line.

Network monitoring estates have grown one tool at a time. Enterprise Management Associates (EMA) found in 2024 that the typical network team still runs anywhere from 3 to 15 tools, even after some recent consolidation. The same study found that fewer than 29% of network alerts were actionable, down from 43% in 2020.[1] Much of the remainder still has to be opened, read, and dismissed by someone.

This paper defines a three-stage causal chain, gives a formula that converts alert volume into annual cost, and applies it to an illustrative 24x7 environment. At 100 to 300 human-handled alerts per day, the model produces **$0.38 million to $1.14 million a year** in attention cost, or **2.5 to 7.4 engineer-equivalents**. It is mostly time already on payroll. It becomes cash only when it lets you retire a tool, avoid a hire, or move a staffed seat; otherwise it is capacity, worth what you do with it.

The paper then evaluates three remedies that work on different variables. AIOps correlation reduces how often people are interrupted. Tool rationalization reduces how much each interruption costs. A co-managed NOC moves the first touch to a contracted provider. Each is shown gross, with a break-even: the most you could pay for it each year before the attention it recovers no longer covers its cost. Under our planning ranges, break-even runs from $0.18M a year (rationalization) to $0.59M (co-managed); compare real quotes against those.

The chain

Tool sprawl creates alert proliferation. Alert proliferation forces context switching. Context switching is where the cost hides.

The price

Attention can be measured in minutes, priced at loaded labor rates, and modeled over a contract term like any other operating cost.

The levers

How often people are interrupted, what each interruption costs, and who carries the first touch. The remedies can be combined.

> *"A wealth of information creates a poverty of attention."*
>
> HERBERT A. SIMON, 1971 (NOBEL LAUREATE IN ECONOMICS, 1978)

Simon wrote this about the design of organizations and their information systems.[2] His point was that information consumes the attention of its recipients. In network operations, the producers are monitoring tools that usually cost nothing extra per alert. The recipients are engineers who cost the organization by the minute.

01 The Problem Hierarchy

Tool consolidation is usually argued on license cost. That argument is real but, in our modeling, smaller than the labor cost, and it is the wrong unit of analysis. The larger cost sits two steps downstream. Others have described tool sprawl as a "cognitive tax" on service desks.[3] This paper's contribution is narrower and more practical: a formula that prices that tax in dollars, and a way to choose among remedies by the variable each one moves.

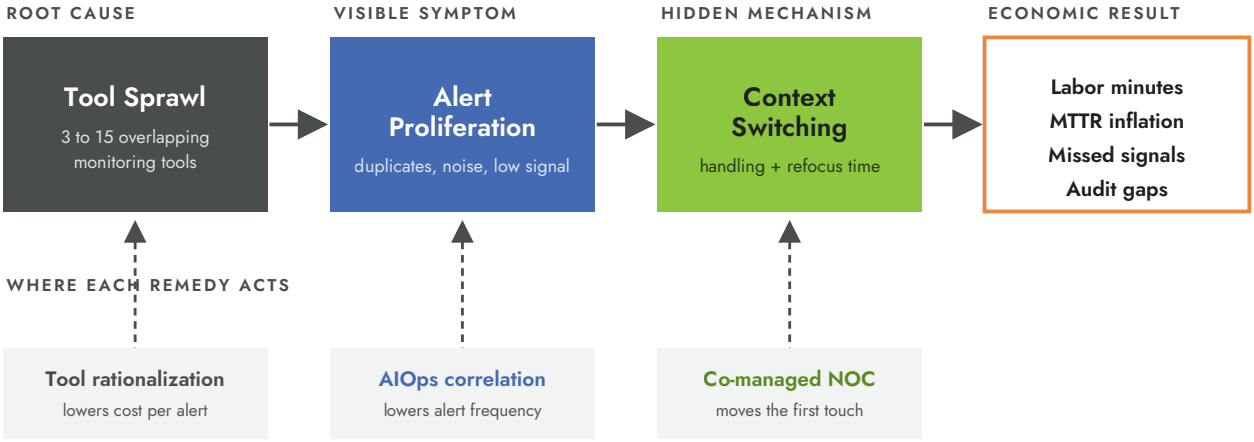

Figure 1. The attention cost chain and the point at which each remedy intervenes.

Tool sprawl (root cause)

Each tool was bought to close a specific visibility gap: device health, traffic flow, optical layer, cloud, configuration. Few were retired when the next arrived. The result is overlapping coverage, with several tools watching the same failure from different angles and each raising its own alarm.

Alert proliferation (visible symptom)

Overlap multiplies alerts without multiplying information. In a carrier network, one fiber cut can produce hundreds of events within minutes: loss of signal and AIS on every affected optical channel, protection switches, and routing adjacency flaps, each raised by whichever system watches that layer. Element managers and alarm hierarchies suppress some of this at the source. What gets through still lands on several consoles. Splunk's *State of Observability 2025*, a survey of 1,855 ITOps and engineering professionals in nine countries, found that 73% had experienced outages due to ignored or suppressed alerts, and 43% said they spend too much time responding to alerts.[4]

Context switching (hidden mechanism)

The cost of an alert is not only the minutes spent on it. It includes the disruption to whatever the engineer was doing before it arrived. Laboratory work by Rubinstein, Meyer, and Evans showed that switching between tasks carries a time cost that grows with task complexity; in their experiments it was measured in fractions of a second per switch.[5] Leroy found that people who switch away from an unfinished task carry "attention residue" into the next one and perform worse on it, an effect that was smaller when the first task had been finished under time pressure.[6] Both are controlled studies, mostly with students. They establish the mechanism, not the size of the cost in a NOC.

02 The Economics of Attention

To be useful to a finance function, attention has to be expressed in units that can be priced. The model separates two kinds of work. **First touch** is the Tier 1 handling of every alert that reaches a person: opening it, pulling context, deciding, closing or escalating. **Escalation** is the share of alerts that go to a senior engineer, who is usually doing something else when the page arrives. Only the escalation carries **refocus minutes**, the time lost re-entering the interrupted work. A dedicated console seat pays no refocus cost. For that seat, the alert is the work.

> THE ATTENTION COST FORMULA
>
> **Annual cost = 365 × Alerts per day × (First-touch cost + Escalation cost)**
>
> First-touch cost = First-touch min ÷ 60 × Tier 1 rate
>
> Escalation cost = Escalation share × (Senior min + Refocus min) ÷ 60 × Senior rate
>
> Use 365 days for any 24x7 operation, 260 only where alerts are genuinely handled on business days alone. Engineer-equivalents = annual hours ÷ 1,850 productive hours per FTE (2,080 paid hours less leave, holidays, and training).

Setting the inputs

Loaded rates. For first touch we use the Bureau of Labor Statistics May 2025 median for computer network support specialists, $76,220 a year or $36.64 an hour at 2,080 hours. For escalations we use the median for computer network architects, $64.45 an hour.[7] In June 2026, wages were 68.5% of total compensation for private-industry management, professional, and related workers.[8] Dividing by that share and adding a 25% Kore-Tek planning allowance for tooling seats, facilities, and supervision gives **$65 an hour for Tier 1 and $120 for senior engineers**, each rounded to the nearest $5. A single blended rate built the same way from the network and computer systems administrators median ($47.66) comes to $85.

Volume and handling. Planning assumptions, to be replaced with local data: 300 alerts a day reach a person, each takes 5 minutes at first touch, 10% are escalated, and an escalation takes a senior engineer 15 minutes. In Kore-Tek's experience, 300 a day fits a regional carrier, while a county or a 25-site enterprise may see far fewer, so the sensitivity table starts at 100.

Refocus minutes. This is the input most often overstated. A widely repeated figure says it takes 23 minutes to return to interrupted work. That number comes from a 2008 magazine interview with Gloria Mark.[9] Her 2005 field study of 24 information workers found that interrupted work was resumed after an average of 25 minutes 26 seconds, but people worked on about two other tasks in between.[10] That is elapsed time, not lost time. A separate lab study by Mark and colleagues, with 48 participants who were mostly university students, found that interrupted people finished an email task faster, at the price of more stress, frustration, and effort.[11] The model uses **10 refocus minutes per escalation**, a Kore-Tek assumption, and shows 0 and 20 as well.

What the formula produces

At 300 alerts a day, first touch takes 25 engineer-hours a day and escalations 12.5, about 13,700 hours a year. That is **$1.14 million a year**, about $3.4 million over a three-year term, or 7.4 engineer-equivalents (4.9 Tier 1, 2.5 senior). The effective blended rate is $83 an hour. Rates are per paid hour, so pricing worked hours at

them understates cost by about 11%; we leave it that way as a conservative choice. Per engineer-equivalent at 1,850 productive hours, that is about $154,000, against about $173,000 of loaded pay per head.

Sensitivity (baseline annual cost, 24x7)	Low	Base	High
Alerts reaching a person per day (100 / 300 / 600)	$0.38M	$1.14M	$2.28M
First-touch minutes per alert (3 / 5 / 12)	$0.90M	$1.14M	$1.97M
Escalation share (5% / 10% / 20%)	$0.87M	$1.14M	$1.69M
Refocus minutes per escalation (0 / 10 / 20)	$0.92M	$1.14M	$1.36M
Both loaded rates (minus 20% / base / plus 20%)	$0.91M	$1.14M	$1.37M

One input varied at a time; all others held at base. Where no one sits at a dedicated console and every first touch interrupts an engineer doing other work, adding 5 refocus minutes to each first touch raises the base case to $1.73M. At 100 alerts a day, the same case is about $0.58M.

WHY THE ARITHMETIC MATTERS

A common shortcut multiplies 300 alerts by 5 minutes by one $85 rate over a 260-day year. That gives $0.55M, which is reasonable for a business-hours operation. For a 24x7 operation, 365 days alone raises it to $0.78M. Repricing first touch at the Tier 1 rate and adding escalations at the senior rate, with refocus, brings it to $1.14M. Applied to a 24x7 NOC, the shortcut understates the cost by about half.

Cash, capacity, and risk

Attention hours are capacity consumed, not headcount that can be cut. A 24x7 first-line seat takes about 4.7 FTE to staff (8,760 hours ÷ 1,850). Halving its alert load does not remove the seat. Any business case built on this model should keep three kinds of value apart.

Cash

Licenses retired, hires avoided, a staffed seat moved to a provider. Only these change the budget. They come from your contracts and staffing plan, not from this model.

Capacity

The hours this model prices. Real, but already on payroll. Worth what the recovered time is redeployed to: change work, projects, documentation.

Risk

Downtime, missed alerts, audit findings. Real and hard to price. Kept separate from the figures above and never added to them.

> *"Every alert your team touches spends engineering time you already paid for. Most organizations never see the invoice."*
>
> RYAN YOUNG, CEO, KORE-TEK

03 Where the Cost Hides

The formula captures the labor directly consumed by alerts. Three further costs sit around it.

Error and rework. Fatigue changes behavior. Teams under alert overload tune thresholds aggressively, close tickets prematurely, or suppress categories outright. Each saves minutes today and creates rework or blind spots later.

Downtime amplification. When a real incident arrives inside a stream of noise, detection and diagnosis slow down. ITIC's 2024 survey of more than 1,000 firms found that an hour of downtime costs more than $300,000 for over 90% of mid-size and large enterprises.[12] That is a threshold most respondents exceed, not a floor, and it is self-reported, enterprise-wide, and excludes litigation, penalties, and remediation. Use your own figure. As an illustration only: six customer-impacting hours a year at $300,000 is $1.8 million of exposure, and a 20% cut in mean time to repair (MTTR) would reduce it by $360,000. The 20% is an assumption, not a measured result. This belongs in the risk bucket.

Governance risk. An alert that was seen, judged, and dismissed leaves a record only if someone wrote one. Under overload, people often do not. When an auditor or a customer asks why a degradation went unaddressed, the organization may not be able to show that its monitoring was acted on.

Hardware, software, energy, and maintenance costs are covered in Kore-Tek's companion paper, *The Economics of Modern NOC: A CFO Playbook*.[13] In Kore-Tek's view, attention is the component most often left out of TCO models.

04 Three Remedies, Three Levers

The formula has three controllable variables: how many alerts reach a person, how long each takes, and whose people make the first touch. The ranges below are Kore-Tek planning assumptions, not published benchmarks. We found no neutral study that measures these effects on human-handled alerts. Vendor figures for noise reduction usually describe raw event compression, which is a different and much larger number. Treat each as a hypothesis for a pilot.

AIOps correlation: reduce how often people are interrupted

Correlation platforms ingest events from every monitoring source, group related events into one incident, suppress duplicates and known noise, and attach topology and change context before a person sees anything. We model a 50% to 70% cut in items reaching a person. Three adjustments keep that honest. A correlated incident is fatter than a single alert, so first touch takes 30% longer. Escalations fall at half the rate of alerts, because escalations are mostly real problems that correlation does not remove. Year one delivers half of the steady-state benefit while the platform is tuned.

When it pays: high alert volumes, overlapping sources, and a topology that can be modeled. *When it does not:* low volumes, no clean inventory or topology data, and no capacity to tune it for several months. Correlation built on a poor inventory correlates the wrong things, and aggressive suppression can hide the one alert that mattered.

Tool rationalization: reduce what each interruption costs

Rationalization removes tools that duplicate coverage and gives each role a single primary view. Its main effect is on handling minutes, because engineers stop hopping between consoles to assemble context. We model 15% to 30% fewer handling minutes and 10% to 20% fewer alerts. It is also the remedy most likely to produce cash, through retired licenses.

When it pays: at a license renewal, when several tools overlap heavily, or when one tool is already the system of record. *When it does not:* when each tool covers a distinct domain (optical, IP, wireless, cloud) that no single product handles well. Forcing consolidation there trades attention cost for blind spots.

Co-managed NOC: move the first touch into a contracted service

In a co-managed model, an external NOC carries 24x7 first-line monitoring, triage, and correlation, while the internal team keeps architecture, change, and escalation decisions. The attention cost does not disappear. It moves to a provider at a contracted price. We assume the provider's correlation matches the AIOps range and that it resolves 30% to 60% of what would otherwise escalate. This is the assumption the co-managed row is most sensitive to: if the provider resolves none of them, its conservative break-even falls from $0.59M to $0.47M a year. Ask for measured resolution rates. What reaches the internal team is harder: each residual escalation takes 20 minutes plus refocus, and someone spends one to two hours a day managing the provider. Where the contract requires it, the provider's ticket record can serve as audit evidence, but only if records ownership, retention, and access at exit are written down.

When it pays: when 24x7 coverage is required but a full internal shift rotation is not economic, when turnover keeps resetting tribal knowledge, or when a first-line seat is open. *When it does not:* when the organization already runs a mature, well-staffed 24x7 NOC with low turnover; when the environment is so bespoke that onboarding costs more than it saves; when regulation or labor agreements keep monitoring in-house; or when the provider will not agree to clear escalation criteria and exit terms. In those cases, correlation and rationalization are the better spend.

CONSTRAINTS BY SECTOR: CHECK THESE BEFORE MODELING ANY REMEDY

- **Healthcare.** A provider that can see protected health information in logs or tickets will likely need a HIPAA business associate agreement before monitoring starts.
- **County, city, and DOT.** Networks carrying criminal justice information fall under the FBI CJIS Security Policy, including personnel screening. Tickets may be public records with retention schedules. Contracting out work done by civil service or union staff may require meet-and-confer. Procurement rules and annual appropriations shape any multi-year term.
- **Oil and gas, utilities.** TSA pipeline security directives and, for bulk electric assets, NERC CIP govern who may access operational monitoring and how.
- **Carriers.** FCC Part 4 outage reporting runs on deadlines from discovery. The contract must say who detects, who notifies, and how fast.

A checklist, not legal advice. Confirm applicability with counsel.

QUESTIONS TO ASK ANY PROVIDER, INCLUDING US

1. What reduction in *human-handled* alerts have you measured, as opposed to raw events, and how?
2. How will you count missed actionable alerts during the pilot and after?
3. What exactly escalates to our team, and is that written into the contract?
4. Who owns tickets and logs, how long are they kept, and how do we get them at exit?
5. What do onboarding and tuning cost, and how long until steady state?
6. What are the exit terms, notice period, and transition support?
7. Will you sign the agreements our sector requires and show staff screening?
8. Will you run this model on our data and give us the working model?

05 Comparative Outcomes Over Three Years

The table applies each planning range to the illustrative environment. All figures are gross capacity: internal attention hours priced at loaded rates, before any platform, implementation, or service fee. None of them is net, and none is cash until it is redeployed or released.

Scenario	Primary lever	Internal attention cost / yr	Capacity recovered / yr	Recovered over 3 yrs (yr-1 ramp)	Break-even remedy cost / yr
Baseline	**None**	**$1.14M**	**n/a**	**n/a**	**n/a**
Tool rationalization	Cost per alert	$0.74M to $0.93M	$0.21M to $0.40M	$0.53M to $1.01M	$0.18M
AIOps correlation	Alert frequency	$0.59M to $0.80M	$0.34M to $0.55M	$0.86M to $1.38M	$0.29M
AIOps + rationalization	Frequency and cost	$0.45M to $0.70M	$0.44M to $0.69M	$1.10M to $1.72M	$0.37M
Co-managed NOC	First touch	$0.21M to $0.43M	$0.71M to $0.93M	$1.77M to $2.32M	$0.59M

Assumptions. Baseline as in Section 02: 300 alerts/day; 5 first-touch minutes at $65/hour; 10% escalated at 15 minutes plus 10 refocus at $120/hour; 365 days. Ranges run low to high; in the cost column the low figure is the favorable case, in the recovery columns it is the conservative case. Remedy assumptions as described in Section 04. Year 1 delivers 50% of steady-state benefit. **Break-even** is the average annual recovered capacity over three years at the conservative end: the most a remedy could cost each year, all in, and still pay for itself on attention alone. It values recovered hours at full loaded cost; if the hours cannot be redeployed, the true break-even is lower.

Three conclusions follow. First, under these planning ranges, remedies that act on frequency recover more than rationalization alone, so correlation is worth piloting wherever volumes are high. Second, co-managed recovers the most because it moves the most work. Its fee is likely the largest, and its $0.59M break-even is close to the cost of the internal first-line seat it would replace ($0.57M priced at 8,760 hours x $65; about $0.64M as 4.7 FTE of loaded annual pay). Third, the decision turns on real quotes: set each remedy's all-in annual cost against its break-even and prefer the widest margin at the conservative end.

Illustrative model: a 40-site regional operator

A regional operator with 40 sites and five monitoring tools staffs one first-line console seat around the clock (about 4.7 FTE). Senior engineers take escalations by day and on call overnight. At 300 alerts a day, first-line load is 25 hours a day, so the seat is overloaded: 25 hours of work a day for a 24-hour seat, so alerts queue or get skimmed. Overnight, about a third of the 30 daily escalations wake the on-call engineer; 10 refocus minutes is a floor for those. Modeled, not drawn from a client engagement.

Metric	Before	After AIOps + rationalization (conservative, steady state)
Items reaching a person per day	300	150
First-touch minutes per item	5.0	5.5
Escalations per day	30	22.5
First-line hours per day	25.0	13.8
Senior hours per day	12.5	8.5
Engineer-equivalents (1,850 hours)	7.4	4.4
Annual attention cost	$1.14M	$0.70M

The operator recovers about 3.0 engineer-equivalents and $0.44 million a year of capacity at steady state. Look at where it lands. About 2.2 of those engineer-equivalents sit on the console seat, which still has to be staffed around the clock; that time becomes room for change verification and documentation, not a budget cut. About 0.8 goes back to senior engineers, worth $0.17 million a year in project time. That split, not the total, is what a CFO should test.

06 Measuring Attention and Running a Pilot

None of the figures in this paper should be accepted on faith, including ours. Every input can be measured locally within a few weeks.

SEVEN MEASUREMENTS FOR A BASELINE

1. Alerts reaching a person, per day and per engineer per shift.
2. Share of those alerts closed with no action taken.
3. Average first-touch minutes per alert, sampled from ticket timestamps.
4. Escalation rate to senior engineers, and minutes per escalation.
5. Average number of tools opened to resolve one incident.
6. MTTR for customer-impacting incidents, with detection and diagnosis time separated.
7. Missed actionable alerts: incidents found by customers or field staff before monitoring flagged them.

A pilot should cover one domain or region for 60 to 90 days. Capture four weeks of baseline, introduce one remedy, and measure the same seven again. The seventh matters most: a remedy that cuts noise by suppressing a real alert has made things worse, and only a false-negative count will show it. The output is a local version of the comparative table, with real inputs in place of planning ranges. That table, not a vendor's case study, is what should go to the finance committee.

07 Common Objections

"Our team is coping. Why change what works?"

Coping is what the model measures. Mark's lab study found that interrupted people keep pace by working faster under more stress.[11] It did not study NOCs or turnover. In Kore-Tek's view, the same pattern in a NOC shows up later as turnover, suppressed alerts, and slower response to the incident that matters.

"The team will resist this."

In our experience, engineers rarely defend alert noise. They defend control. Remedies that keep architecture and escalation decisions internal and remove only the repetitive first touch are easier to accept. In the public sector, involve labor relations before the pilot, not after.

"We are committed to our current tools."

Correlation and co-managed models work on top of existing tools, and rationalization can wait for renewal. None requires a forced migration.

"We don't trust a modeled number."

Neither should anyone else. That is why the formula uses inputs every NOC already records, and why the recommended next step is a measured baseline rather than a purchase.

08 Conclusion

Tool sprawl is usually presented as a licensing problem. Its larger cost is paid in engineering attention, one alert at a time, and it is invisible because few organizations add it up. Bringing it under control takes three things: measure it with a formula a finance team can audit, choose remedies by the variable each one moves, and hold every option to its break-even before committing. The question for the executive team is not whether the attention tax exists. It is whether anyone in the organization knows what it costs, and how much of it is cash.

Request a 20-minute briefing.

Send us four weeks of alert and ticket counts and timestamps; no ticket content is needed, and none should be sent before the agreements your sector requires are in place. We will run this model on it, show which remedies clear their break-even, and give you the working model to keep, whether or not you buy anything from us. kore-tek.com/service/network-operations-center

Sources

1. Enterprise Management Associates, *Network Management Megatrends 2024: Skills Gaps, Hybrid and Multi-Cloud, SASE, and AI-Driven Operations* (n=406), Key Findings and Figure 8. Copy hosted by Broadcom: docs.broadcom.com.
2. Simon, H. A. (1971). "Designing Organizations for an Information-Rich World." In M. Greenberger (Ed.), *Computers, Communications, and the Public Interest*, pp. 40-41. Baltimore: Johns Hopkins Press. The sentence continues: "and a need to allocate that attention efficiently..."
3. Gorelo, "MSP Tool Sprawl: The Hidden Cognitive Tax Slowing MSP Service Desks," January 2026. gorelo.io.
4. Splunk, "State of Observability 2025," October 21, 2025 (n=1,855, nine countries). splunk.com.
5. Rubinstein, J. S., Meyer, D. E., & Evans, J. E. (2001). "Executive Control of Cognitive Processes in Task Switching." *Journal of Experimental Psychology: Human Perception and Performance*, 27(4), 763-797.
6. Leroy, S. (2009). "Why Is It So Hard to Do My Work? The Challenge of Attention Residue When Switching Between Work Tasks." *Organizational Behavior and Human Decision Processes*, 109(2), 168-181.
7. U.S. Bureau of Labor Statistics, *Occupational Outlook Handbook*, median pay May 2025: computer network support specialists; computer network architects; network and computer systems administrators.
8. U.S. Bureau of Labor Statistics, *Employer Costs for Employee Compensation*, June 2026, Table 4 (private industry, management, professional, and related occupations). bls.gov.
9. Pattison, K. (2008). "Worker, Interrupted: The Cost of Task Switching." Interview with Gloria Mark. *Fast Company*, July 28, 2008. fastcompany.com.
10. Mark, G., Gonzalez, V. M., & Harris, J. (2005). "No Task Left Behind? Examining the Nature of Fragmented Work." *Proceedings of ACM CHI 2005*, p. 326. PDF.
11. Mark, G., Gudith, D., & Klocke, U. (2008). "The Cost of Interrupted Work: More Speed and Stress." *Proceedings of ACM CHI 2008*. PDF.
12. ITIC, *2024 Hourly Cost of Downtime Survey*. Copy hosted by Calyptix: survey results PDF.
13. Kore-Tek, *The Economics of Modern NOC: A CFO Playbook*, June 2026. kore-tek.com.

All modeled figures are illustrative, derived from the stated assumptions in a model described in Sections 02 and 05, and are not drawn from any client engagement. Remedy ranges, overhead, refocus, escalation, and MTTR figures are Kore-Tek planning assumptions, not guarantees or published benchmarks.

ABOUT THIS PAPER

The Attention Tax is No. 6 in the Kore-Tek Strategic Perspectives series, framework-level analysis for senior leaders responsible for network operations. Earlier papers include No. 1, The AIOps and Managed NOC Convergence; No. 2, The Co-Managed NOC; No. 3, The Outsourcing Inflection; No. 4, NOC Is the New Boardroom; and The Economics of Modern NOC: A CFO Playbook. The full library is at kore-tek.com/insights.

Kore-Tek provides network operations center, field engineering, and professional services to carriers, enterprises, and public-sector organizations. 225 South 6th Street, Suite 3900, Minneapolis, MN 55402. kore-tek.com | (800) 866-9176