

# Reducing Cyber Catastrophe Risk: Diversification and Mitigation in Action

Shaping the Future of Cyber Insurance with Portfolio Manager v6

# Authors



**Edward Asiedu**  
Principal Cyber Risk Modeling

**Doug Fullam**  
Director of Actuarial Modeling

**Jon Laux**  
Vice President of Analytics

**Ethan Spangler**  
Lead Economist

---

Editorial Manager:

**Yvette Essen**

Head of Communications  
& Market Engagement  
at CyberCube

Editorial Design:

**Felix Paula**

Growth Marketing Creative  
at CyberCube



# Executive summary

The cyber insurance market has experienced rapid and sustained growth over the past several years, emerging as a catastrophe (CAT)–exposed and capital-intensive line of business. This trajectory, while promising, heightens the need to understand the role of diversification and risk mitigation—two themes that have been extensively examined in natural catastrophe insurance, but remain comparatively underexplored in cyber.

Our analysis, using CyberCube’s Portfolio Manager v6 (PM v6), highlights several key takeaways for insurance executives:



**Security posture matters greatly** – The cyber hygiene of insured enterprises has material implications for portfolio performance and capital requirements. Strong patch management, network segmentation, and robust backup protocols can decrease modeled tail losses by up to 57% – and greater still when focusing only on widespread ransomware events.



**Portfolio diversification as risk mitigation** – Increasing diversification across geography, revenue, industry, and technology can significantly reduce cyber risk, lowering potential losses by up to 42%.



**Diversification potential is real, but uneven** – Continued risk diversification can be gained if market growth continues beyond the US, with a balance of exposures by company size, industry, and/or technology.



**Exposure remains US-concentrated** – Insured cyber risk is still heavily concentrated in the United States, which accounts for roughly two-thirds of the current global cyber insurance market.



**Single Points of Failure (SPoFs) are still US-centric** – Many of the most significant tail risk events are tied to US-based technologies, particularly operating system providers and major cloud service providers, whose infrastructure remains geographically concentrated.

These findings underscore both the opportunity and challenge facing the market. While geographic and sectoral diversification is gradually improving, critical dependencies remain concentrated, leaving portfolios exposed to correlated, large-scale events. In parallel, improving security standards across the largest insureds could deliver measurable capital efficiency providing another means to reduce both tail risk and overall risk.

The following report examines these themes in detail, exploring the levers of diversification, the nature of cyber peak perils, and the measurable impact of mitigation strategies on portfolio-wide risk. To evaluate the benefits achievable across various dimensions, we constructed portfolios derived from CyberCube's Industry Exposure Database (IED)<sup>1</sup> that varied across the metric of interest, allowing us to fully isolate the impact across diversification and mitigation.



CyberCube's advanced analytics are used by 75% of the top 40 US and European cyber insurance carriers. In July 2025, CyberCube released the latest version of Portfolio Manager, the catastrophe model empowering portfolio-level insights.

Key changes made in Portfolio Manager Version 6 (PM v6) followed extensive research with internal and external cyber experts to understand what will best prepare organizations to avoid the consequences of catastrophic events.

This report reflects the rationale behind the changes. More information on PMv6 can be found [here](#).

---

<sup>1</sup> CyberCube's [Industry Exposure Database \(IED\)](#) provides the foundation for all modeled industry loss estimates.

# Diversification in Cyber Insurance

## The value of diversification

Diversification is a central reason why insurance and risk pooling work successfully. Through the law of large numbers, insurers are able to reduce the variance of potential claims payments relative to the expected (average) level of claims payments and thereby achieve more predictable cashflows. In this regard, Cyber insurance is no different than other lines.

But catastrophe risk brings inherent challenges to the idea of diversification: when many insureds can be affected by a single event and claim at the same time, this does not foster diversification but rather its opposite. For Property risks, geography is a natural aid to diversification and a way that (re)insurers effectively manage the catastrophic potential of the business.

Here is where Cyber insurance is different: with global network connections being the key reason why cyber risk exists at all, geography does not provide a ready path to diversification. And with global Cyber premiums now exceeding \$16 billion and global insured limits nearing \$3 trillion<sup>2</sup>, the potential losses from cyber catastrophes have become too large for (re)insurers to ignore. Yet despite these challenges, (re)insurers that can achieve some level of diversification will more effectively manage their aggregate cyber exposure and grow more efficiently in the process.

## Drivers of diversification

With this in mind, CyberCube has conducted a study to evaluate the diversification potential that exists within the global cyber insurance market. Our analysis considers the common firmographic properties that underwriters use in segmenting portfolios – industry, company size and geography – as well as the Single Point of Failure (SPoF) technology dependencies that are central in creating aggregation potential. CyberCube's PMv6 model update brought specific attention to these levers of diversification and was used to conduct the following analysis.

Our findings are shown in **Exhibit 1**. Here we show the maximum diversification benefit that can be reached, at what point in the loss curve that maximum is achieved, and the diversification benefit at a common “tail” return period, the 1 in 200 occurrence loss. For a full description of how we conducted the analysis, see the end notes of this report. <sup>i,ii</sup>

---

<sup>2</sup> Source: CyberCube Industry Exposure Database

## Exhibit 1: Cyber Diversification Benefits

Diversification Strategy	Maximum Benefit (RP of Benefit)	Benefit at 200yr RP Event
Geography	11% (25yr)	2%
Revenue	17% (65yr)	13%
Industry	27% (140yr)	23%
All Firmographic Strategies (Geography + Revenue + Industry)	38% (170yr)	33%
Technology (Single Point of Failure)	38% (229yr)	37%
All Strategies Combined (Technology + Geography + Revenue + Industry)	42% (242yr)	41%

### Diversifying by Geography

To the degree that Cyber risk is like Property risk, one would expect that a portfolio with policies spread across different regions would have more diversification benefit than a portfolio with all policies from a single region. Indeed, we do find that technology usage and threat actor attack trends vary by geographic region. Yet this variation has its limits as multi-region and global scale events come to dominate the loss curves in the tail. The reasons for these limits are discussed in the following section.

As a result, geographic diversification benefits max out relatively early in the loss curve – the 1 in 25-year point – and their effect is quite muted at the 1 in 200 level.

### Diversifying by Revenue

It's intuitive that small commercial, middle market and large account-focused portfolios would behave somewhat differently; each of these segments will have somewhat different technology reliance, different levels of risk posture, and different patterns of recovery following a catastrophic event.

Segmenting our IED into revenue bands, we generally see a decrease in average annual loss (AAL) and tail loss metrics as we shift from a portfolio concentrated among large firms (\$1bn+ in yearly revenue) to a portfolio that also has a spread of exposure from large firms down to small commercial and “micro” firms (sub-\$10mn in yearly revenue).

This strategy of diversifying by revenue band achieved a maximum of 17% diversification, and 13% at the 1 in 200 loss level.

## Diversifying by Industry

Industries carry varying levels of cyber risk due to differences in attractiveness to threat actors as well as different technology reliance. Not surprisingly, the Information Technology sector shows higher losses than many sectors, while the public sector and Agriculture firms fall on the lower end of the spectrum.

We found industry diversification to be a useful strategy, achieving a maximum benefit of 27% and a 1-in-200 benefit of 23%. It's also worth noting that the maximum industry benefit occurs further out in the tail (1 in 140) than it does for geography and for revenue band.

## Diversification Using Geography, Revenue and Industry

Combining the above three strategies together, we evaluated portfolios that were concentrated among certain geography-revenue band-industry segments versus those writing a diverse number of areas. We observed that portfolios concentrated on large accounts consistently exhibit the highest AAL and EP curves even when controlling for portfolio size. Additionally, the industries previously identified with higher risk metrics continue to show elevated losses compared to other industries within the same size band.

Our results show that by combining these three dimensions of diversification – geography, revenue and industry – insurers can achieve substantial diversification benefits. We see a maximum benefit of 38%, and a sizable 33% at the 1 in 200. These findings illustrate the value of considering multiple dimensions of exposure when constructing portfolios, as doing so can meaningfully reduce expected losses, including in the tail of the distribution. Moreover, these three firmographic dimensions can be readily captured by insurers at the point of underwriting, making such a diversification strategy more transparent and actionable.

## Diversifying by Technology

Since geography, revenue and industry variations each reflect differences in technology dependency, we can also look directly at these dependencies – aka Single Point of Failure (SPoF) technologies – for their efficacy as a diversification strategy. The challenge, though, is with how ubiquitous these technologies are. Florida hurricane risk is recognized as the peak peril among natural catastrophes worldwide, but Florida comprises only 10.8% of US homeowners insurance premiums – and even less when considering non-US premiums as well.<sup>3</sup>

By contrast, SPoF technology usage is extremely concentrated: the largest cloud service provider (Amazon Web Services) holds an estimated 31% global market share, and the largest desktop operating system provider (Microsoft Windows) currently has an estimated 72% market share.<sup>4</sup> These concentrations make it difficult to diversify away from the peak exposures of Cyber.

<sup>3</sup> Source: NAIC, <https://content.naic.org/sites/default/files/publications-key-facts-market-trends-florida.pdf>

<sup>4</sup> Source: CyberCube Portfolio Manager version 6

Nonetheless, when we locate policies that do rely on only one critical service provider or another, we can see that significant diversification is available – for example when combining AWS-only policies with Azure-only policies. Our research found a 38% diversification benefit could be attained from this strategy. In other words, SPoF diversification on its own can produce the same level of benefit as the combination of all firmographic factors: revenue, industry plus geography.

Additional commentary about some limits to SPoF and other diversification strategies are discussed in the following section.

## Diversification Using Geography, Revenue, & Industry + Technology

Combining a technology diversification strategy with the firmographic strategies shown previously we see a maximum benefit of 42%, and a 41% benefit for a 1 in 200 event. It's particularly worth noting that by including technologies (SPoF) in the diversification strategy, the maximum benefit is reached between the 200-year and 250-year points – in other words, at the points that (re)insurers are generally most concerned with for reinsurance, capital and regulatory purposes.



## Limitations to Diversification

The findings above suggest that despite the globally networked nature of cyber risk, (re)insurers can achieve meaningful benefits from thoughtful portfolio construction. We hope the diversification metrics presented provide indicative guidance on the relative role that each factor can play in creating a diversified cyber portfolio. When considering Cyber's diversification potential, insurers should keep several constraints in mind ranging from the structural to the practical.

### Concentration of Insureds

The US is still by far the largest Cyber insurance market in the world, accounting for approximately two-thirds of gross written premium and limits – 10 times the size of the number two and three largest markets, UK and Germany.<sup>5</sup> Today, the concentration amongst US insureds works as a natural constraint on potential diversification, as (re)insurers will find it difficult to attain any sizable premium volume without exposure to US insureds.

This will not always be the case though; over the next few years European and Asian countries are expected to see faster market growth than the US.<sup>6</sup> Insurers with a presence in a variety of markets geographically will likely see further diversification relative to those that remain concentrated on the US.

### Concentration of Single Points of Failure

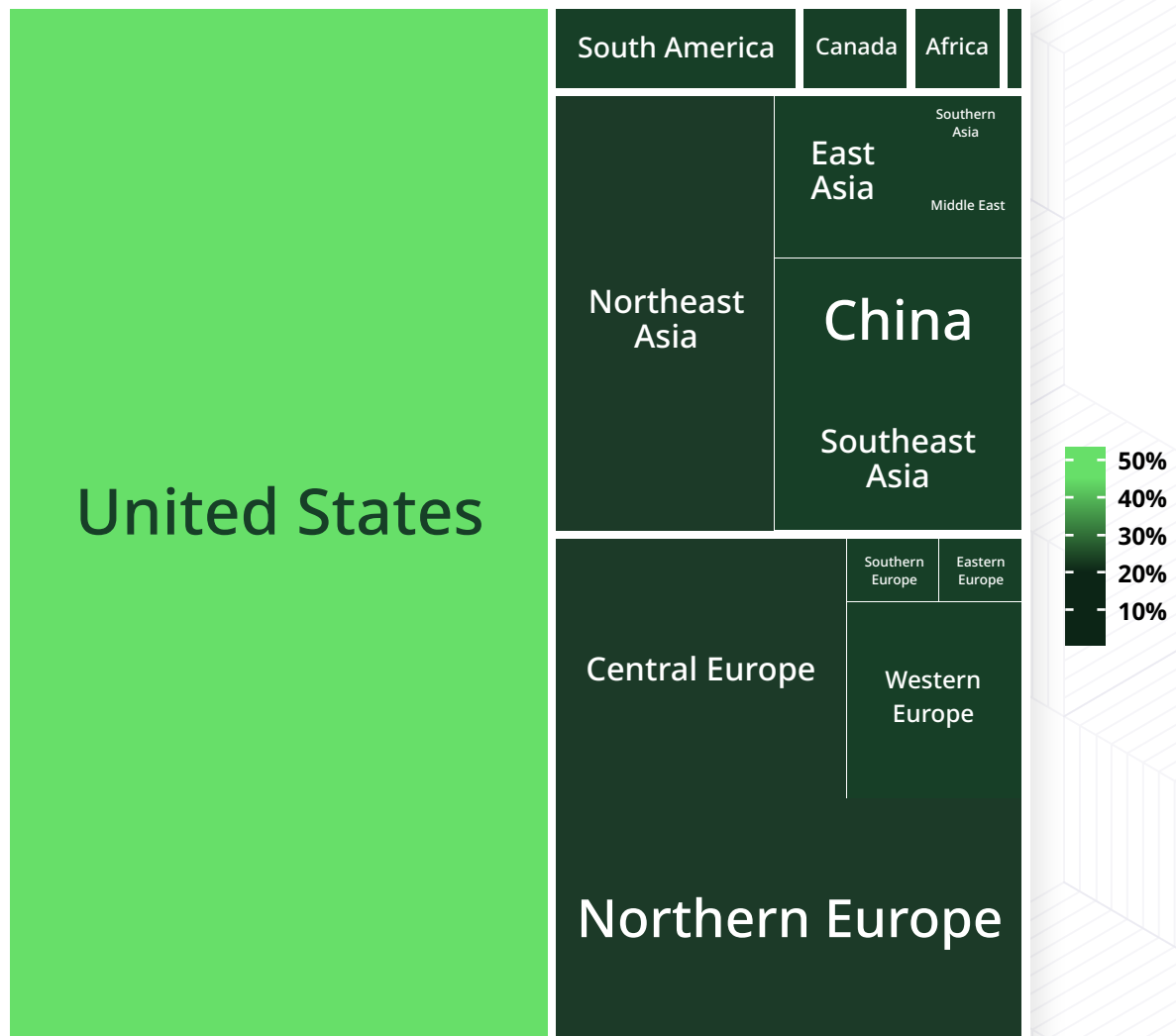
Even if insurance buyers become less geographically concentrated, technology usage itself is concentrated among a very small number of major providers – again based largely in the US. Cloud Service Providers (CSPs) provide an example: collectively, US CSPs make up 75% of the global cloud market. The largest Amazon Web Services region, us-east-1, accounts for 5.5% of global market share on its own. While there are certainly many smaller CSPs based in other countries, these are predominately Tier 2 and Tier 3 providers that often lack infrastructure of their own and are themselves reliant on the “big 3”: AWS, Microsoft Azure, and Google Cloud.

The big three providers have each established a global presence with regions and availability zones all over the world. This serves to store data closer to customers and provide redundancy in case one availability zone or region is offline. But when considering the actual cloud regions that firms rely on, there is again significant US concentration. As shown in **Exhibit 2**, we estimate over 50% of the modeled market share is in US-based regions across the major CSPs. These are not only US-based firms using these cloud regions, but companies from all over the world. The concentrations are most extreme for AWS, Azure, and Google Cloud with roughly half of each's market share stemming from their US cloud regions. Even China-based Tencent and Alibaba have their most-utilized cloud regions in the US.

<sup>5</sup> Source: CyberCube Industry Exposure Database

<sup>6</sup> Source: Munich Re, <https://www.munichre.com/en/insights/cyber/cyber-insurance-risks-and-trends-2025.html>

## Exhibit 2: Geographic Concentration of Cloud Regions



### Practical Limits

Cyber insurers need to balance the potential benefits of diversification with its potential impact on other key factors, including administrative challenges, expense impacts, and their level of underwriting expertise – or lack thereof – in potential new segments. While it may be mathematically desirable for a US-based insurer to establish a mainland European operation, a lack of distribution platform or underwriting expertise in Europe may not make this idea immediately tractable. Such practical constraints exist in other insurance lines as well, of course.

In this regard, reinsurers are likely better positioned to establish diversification strategies than primary carriers, having less administrative burden and being able to scale up or down line sizes at renewal. Such a dynamic should work for the Cyber market, with the diversification benefits gained by reinsurers allowing them to support a variety of cedant programs with greater capital efficiency.

# Risk Mitigation in Cyber Insurance

## Mitigating against cyber catastrophe

Cyber risk is not a threat for firms to face passively; there are tangible security controls that firms can apply both to reduce the likelihood of being affected by an event and to reduce the impact of an event if they are in fact compromised. Such controls form the basis for much of cyber insurance underwriting.

However, little analysis has been done so far in evaluating the efficacy of security controls in mitigating cyber catastrophe risk. While it is the “attritional” claims that will largely contribute to an insurer’s loss ratio in any given year, it is the potential claims from catastrophe events that drive capital allocations, reinsurance decision making, and regulatory reporting. If firms and their insurers can use certain security controls to reduce their cyber catastrophe exposure, this will yield tangible benefits to the bottom line.

CyberCube has been researching catastrophe risk mitigation tactics for several years, with PM v6 now including risk mitigation as a central feature. (Re)insurers can now more fully understand the effects of account-level security posture on portfolio-level loss results and, over time, recognize the ways that helping improve firms’ security posture can reduce the potential cost of catastrophic events as well as lower the catastrophe load included in pricing.

## Drivers of cyber catastrophe risk mitigation

CyberCube’s<sup>7</sup> research found that mitigation factors take prominence in their effectiveness for reducing the likelihood and severity of incidents. These are:

- **Patch management:** the pace and quality that a company identifies, tests, and deploys software updates to fix bugs and close security vulnerabilities
- **Network segmentation:** the extent to which an organization’s IT infrastructure has been subdivided to limit network access and protect critical information assets
- **Data backups:** policies and technologies utilized by a company to provide a fallback in the event that primary network storage become unavailable.

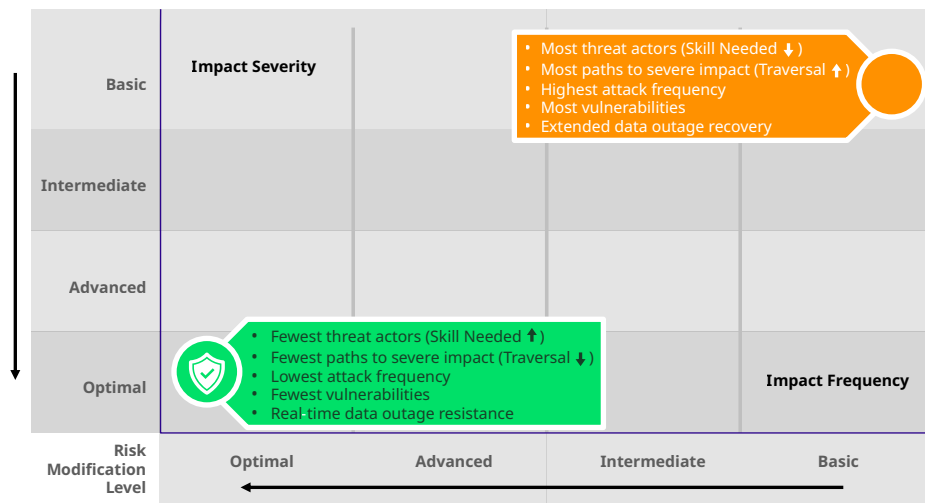
These factors reduce cyber risk by making a firm’s networks more difficult to penetrate in the first place, harder to reach the systems that matter most, and better prepared to recover if indeed the worst does happen. CyberCube has built a framework in PM v6 allowing (re)insurers to use these controls as risk modifiers by selecting one of four risk posture levels: Basic, Intermediate, Advanced, or Optimal.<sup>8</sup>

<sup>7</sup> For more information see CyberCube’s joint report with Munich Re:

<https://insights.cybcube.com/key-insights-into-systemic-cyber-risk>

<sup>8</sup> We have created these four groupings as easy-to-use simplifications that each describe a broad set of policies and practices associated with a certain level of risk mitigation and security maturity. These policies and practices are available to customers in our documentation. Curious souls may also query CIS Critical Security Controls Implementation Groups (<https://www.cisecurity.org/controls/implementation-groups>) which helped inform CyberCube’s framework.

### Exhibit 3: Effects of Risk Mitigation on Frequency & Severity of Impact



As a company improves its posture as reflected in these risk modifiers, the PM v6 model will reflect that improved posture in reduced likelihood and impact from catastrophe events in our catalog. The following table shows the reductions we estimate seeing at the 1-in-200 occurrence loss for a portfolio having all companies with an Advanced risk posture versus a Basic risk posture.

### Exhibit 4: Impact of Risk Modifiers for 1:200 Occurrence Loss By Peril

	Loss Reductions		
Risk Modifier	Widespread Ransomware	Widespread Outage	All Perils
Patch Management	40%	-	16%
Network Segmentation	40%	-	16%
Data Backups	48%	22%	37%
<b>All 3 Combined</b>	<b>80%</b>	<b>22%</b>	<b>57%</b>

Numbers shown reflect the change from Basic to Advanced posture for each risk modifier and for all three combined.



Here we see that Patch Management, Network Segmentation and Data Backups each play an important role in lowering a portfolio's vulnerability to widespread ransomware and wiper malware attacks, with all three combined reducing tail losses by 80% at the 1-in-200 point. Note that this comparison does not assume any companies operating at the Optimal posture level, which would provide even further risk reduction but would be an arduous and costly level of security for most companies to achieve.

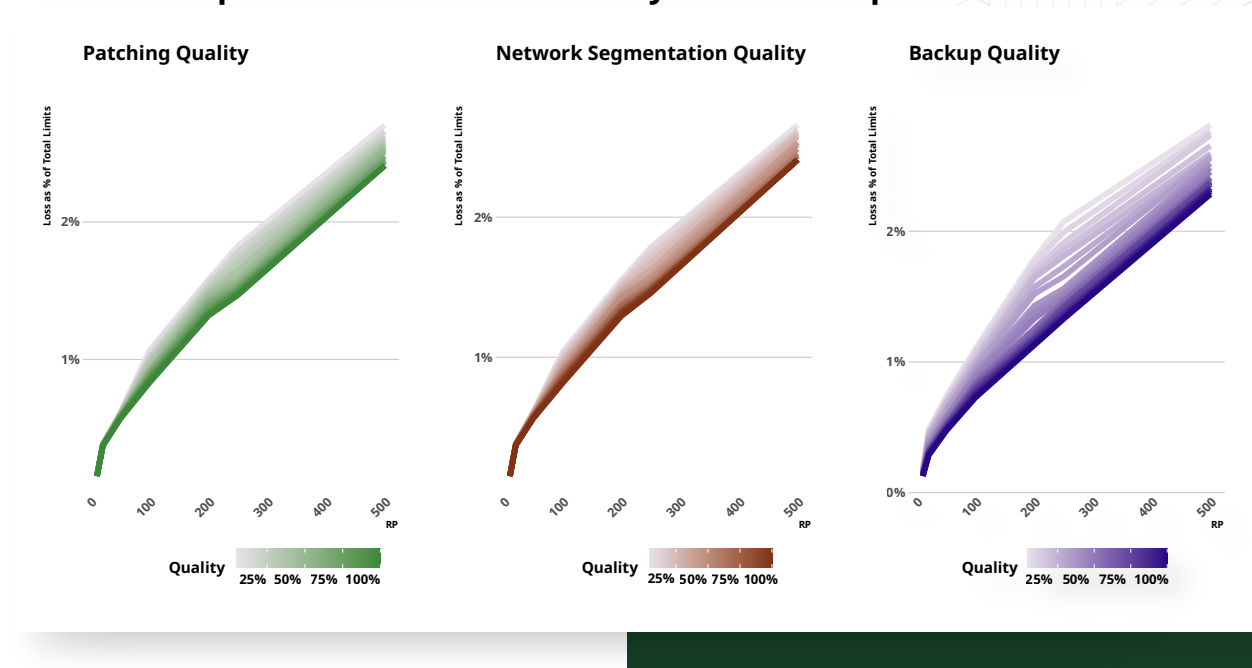
By contrast, these mitigations are less effective in widespread outage events – e.g. a major cloud service provider – since companies have fewer options for recovery when recovery efforts fall mainly upon the service provider that is down. Nonetheless, Data Backups still provide some mitigation in helping companies recover faster from outages.

Across all perils, these three factors can reduce tail risk by **57%** when comparing a portfolio having 'Advanced' posture for all companies versus one with entirely 'Basic' posture.

In practice, shifting an entire portfolio's posture from Basic to Advanced – let alone to Optimal – would be a significant effort. That effort begins with proper data collection, then with taking action. Most cyber insurers are gathering the information during underwriting that is needed to populate these risk modifiers, but few today are capturing that information in a consistent and organized manner that can readily be used to populate catastrophe models – particularly for Network Segmentation. Naturally it will take time for insurers to systematically ascertain the security posture of policyholders, but by doing so they will see benefits especially as they encourage their policyholders to improve in these areas.

**Exhibit 5** shows how losses decrease as the security posture of a portfolio gradually increases.

### Exhibit 5: Impact of Risk Modifiers on Cyber Catastrophe Loss



Quality references the proportion of the portfolio that has Advanced risk modifiers vs Basic.

## Practical steps toward risk mitigation

PM v6 reflects more meaningful variation in results based on the efficacy of risk management practices. We encourage clients to reexamine their portfolios and engage in informed discussions with their cybersecurity, loss control and underwriting teams with the goal of ensuring that policy data related to patch management, network segmentation, and data backup practices is accurate, comprehensive, and consistently interpreted across the organization.

For (re)insurers wanting to take action to recognize the effects of risk mitigation on their catastrophe losses, CyberCube recommends the following steps:

1

**Begin with 'Unknown':** As mentioned above, we recognize that most insurers will need some time to gather risk modifier information at scale and with confidence. PM v6 has been calibrated so that 'Unknown' values for Patch Management, Network Segmentation, and Data Backups provide a reasonable starting point and reflect a typical level of security for companies of all industries and sizes.

2

**Populate data when ready:** For companies where mitigation factors are known with some confidence, (re)insurers can populate PM v6 to reflect them – and generally will see a modest benefit to loss estimates from doing so. CyberCube has also provided benchmark percentages of Basic, Intermediate, Advanced and Optimal mitigation levels from our Industry Exposure Database. While individual portfolios will vary, we expect that overall portfolios will be generally aligned with the benchmarks provided.

In the near term, we recommend clients prioritize deeper analysis of their larger policies, which have a more material impact on catastrophic loss. A realistic target for large (by count) portfolios is to achieve at least 'Intermediate' levels for at least one of the three risk modifiers across a significant portion of policies. In our IED simulations, a large portfolio that moved all three risk modifiers from Basic to Intermediate mitigation saw a 34% decrease in AAL.

3

**Use 'Optimal' with caution:** The 'Optimal' mitigation level is reserved for firms that have gone above and beyond to achieve the absolute best level of protection available today. Such steps are quite costly and impractical for most organizations; often, they are requirements for participating in classified assignments such as national security and defense contracts.

Achieving 'Optimal' levels for all risk modifiers across an entire portfolio is an aspirational goal: it would signify a major advance in industry-wide cyber hygiene. At today's levels, assigning 'Optimal' across all three risk modifiers for a large portion of a portfolio could raise concerns and would be statistically almost impossible unless substantiated with strong evidence.

Insurance works by pooling risk, and portfolios will naturally include varying levels of cybersecurity maturity. Additionally, as attackers evolve along with cyber hygiene, so too will what it means to be 'Optimal'. In time, what were once considered the best security practices will become basic as new techniques and threats emerge.

# Conclusions

As demonstrated, increased diversification and mitigation across a portfolio can have significant impact in decreasing a (re)insurers' exposure to catastrophic cyber risk. Though the full benefits of diversification may be difficult to achieve due to extant market dependency on US policyholders and technology firms, over time this geographic concentration will reduce as other insurance markets expand and mature.

In terms of mitigation, CyberCube has found that Patch Management, Network Segmentation and Data Backups are some of the most important cybersecurity practices when thinking about cyber catastrophe risk. While it will take some time for (re)insurers to gather and act on this information consistently, the benefits for (re)insurers and for policyholders to manage the underlying risk are significant. We are very excited to see how the industry evolves in the coming years as we work together to foster greater cyber resilience.

# Appendix: Technical Notes

<sup>i</sup> **Measuring diversification benefits:** Diversification could be measured in a variety of ways. In this paper we seek to representation diversification as the benefit a portfolio can achieve relative to perfect correlation.

Under perfect correlation ( $\sigma = 1.0$ ), the variance of the whole equals the squared sum of the standard deviations of each segment.  $\text{Var}(X + Y + \dots) = (\sigma_x + \sigma_y + \dots)^2$ . Or put differently, the standard deviation of the whole is the sum of the standard deviations of the segments. Under a perfect correlation scenario, all points along the exceedance probability (EP) curve would be aligned. The 99th percentile event for segment X would also be the 99th percentile event for segment Y.

By contrast, under zero correlation ( $\sigma = 0$ ), the variance of the whole is the sum of the variances of each segment.  $\text{Var}(X + Y + \dots) = \text{Var}(X) + \text{Var}(Y) + \dots$ . Cyber loss correlation will always be between these two scenarios – i.e. bounded within  $[0, 1]$  – as negative correlations do not exist for our purposes. How far below 1.0 the correlation falls determines the diversification benefit.

Let's imagine a combined portfolio comprised of 2 different segments. We calculate the ratio of the combined portfolio probable maximum loss (PML) at a selected return period versus the sum of the PMLs for each segment at that same return period. In other words:

Ratio = [PML for Combined Portfolio] / [Sum of PMLs for each Segment]

Diversification Benefit =  $[1 - \text{Ratio}]$ , stated as a percentage

To provide an example using Geography:

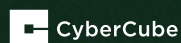
1 in 200 occurrence loss for Geography 1 = 1,000  
 1 in 200 occurrence loss for Geography 2 = 600  
 1 in 200 occurrence loss for all geographies = 1,400  
 Ratio =  $1,400 / (1,000 + 600) = 0.875$   
 Diversification Benefit =  $1 - .875 = 0.125$  or 12.5%

<sup>ii</sup> **About our portfolio testing:** Our findings are presented based on notional portfolios that we created. We sought to ensure these portfolios were realistic, controlling for potential downsides and thinking through typical portfolio constraints.

These metrics can operate as potential rules of thumb, but we did not work to ensure the metrics were the lowest possible. The metrics may also not be fully attainable as we did not consider the execution limiters. One way in which we could have gone further would be to add more controls for the SPoF groups. That said, we thought this would challenge the practicality of the output for two reasons:

- 1) Insurance firms often have difficulty identifying a company's full technology stack. Cloud infrastructure and operating system technologies present some of the greatest potential for widespread impacts which made them a relevant focus for this paper.
- 2) Firms use many technologies; being overly prescriptive in our approach would likely be impractical for the range of situations insurers will encounter.





This document is for general information purpose only and is not and shall not under any circumstance be construed as legal or professional advice. It is not intended to address all or any specific area of the topic in this document. Unless otherwise expressly set out to the contrary, the views and opinions expressed in this document are those of CyberCube and are correct as at the date of publication. Whilst all reasonable care has been taken in the preparation of this document including in ensuring the accuracy of the content of this document, no liability is accepted by CyberCube and its affiliates for any loss or damage suffered as a result of reliance on any statement or opinion, or for any error or omission, or deficiency contained in the document. CyberCube and their affiliates shall not be liable for any action or decisions made on the basis of the content of this document and accordingly, you are advised to seek professional and legal advice before you do so. This document and the information contained herein are CyberCube's proprietary and confidential information and may not be reproduced without CyberCube's prior written consent. Nothing here in shall be construed as conferring on you by implication or otherwise any licence or right to use CyberCube's intellectual property. All CyberCube's rights are reserved. CyberCube is on a mission to deliver the world's leading cyber risk analytics. We help cyber insurance market grow profitably using our world leading cyber risk analytics and products. The combined power of our unique data, multi-disciplinary analytics and cloud-based technology helps with insurance placement, underwriting selection and portfolio management and optimization. Our deep bench strength of experts from data science, security, threat intelligence, actuarial science, software engineering, and insurance helps the global insurance industry by selecting the best sources of data and curating it into datasets to identify trustworthy early indicators.