Richard Dzina: Four questions on the state of cyber resilience and endpoint security

Remarks by Mr Richard Dzina, Executive Vice President of the Financial Services Group of the Federal Reserve Bank of New York, at the Clearing House and Bank Policy Institute's 2018 Annual Conference, New York City, 26 November 2018.

* * *

As prepared for delivery

I want to thank The Clearing House for this generous invitation and, especially, for assembling such a compelling and topical program.

The relationship between The Clearing House and the Federal Reserve is multi-faceted: across one plane competitor, as CHIPS and the Fedwire Funds Service, and EPN and FedACH, vie for volume across common service lines; across another plane collaborator, as peer infrastructures seeking to maximize the value and minimize the disruption to customers in response to common industry initiatives, with our coordinated plans for adoption of ISO 20022 standard message formats representing a cardinal example; and across a final plane as overseer and overseen. When it comes to the resiliency and security of our infrastructures, however, there is no ambiguity in role or relationship: we rise, or fall, as one.

Standing before an audience like this never fails to elicit a visceral sense of accountability. When the subject is the resiliency and security of the foundational elements of our nation's financial system, on which market function and economic vitality ultimately depend, that sense of accountability resonates even more keenly.

I had the privilege of hearing General Michael Hayden, former head of the NSA and CIA, address a payments symposium hosted by the Federal Reserve Bank of Chicago a few years ago. In that venue, General Hayden spoke of the tectonic shifts affecting the global landscape, and their prospective impact on cyber, terror, and geopolitical threats.

His basic message, in two parts: keep your seat belts on, it promises to be a volatile century (yes, century); and when it comes to our critical financial market infrastructure being subject to attack, it is not a matter of "if" but "when". Responding to the current threat environment represents the challenge of our generation; these are days of preparation.

I will posit that a fundamental distinction exists between cyber resiliency and traditional resiliency: what works for us in conventional resiliency scenarios, namely the automatic and instantaneous replication of data, works against us in cyber scenarios, reflecting risks that a pernicious malware, or a severe data corruption, or even an application failure propagates itself across primary, secondary, and tertiary operating sites, rendering a systemic infrastructure functionally inoperable. In other venues we have therefore elaborated upon the need for a new paradigm for resiliency, one that endeavors to mate traditional precepts of geographic dispersion of infrastructure and human capital with new realities compelling technological diversity, and a platform for recovery, should the core of a financial market infrastructure become corrupted.

We have also expounded upon new realities in endpoint security, borne of painful experience, in which endpoint breaches can have serious implications for public confidence in the integrity of a network as a whole, even in circumstances when an operator's infrastructure, applications, and security perimeter have not been compromised. Historically, market infrastructures have operated on the presumption that endpoint security is principally the responsibility of the endpoint, and infrastructure security is principally the responsibility of the operator. In the present threat environment, that bifurcation in responsibility is no longer so clear: the fates of operator and

endpoint are inextricably linked.

The Wholesale Product Office's efforts to strengthen resiliency and enhance endpoint security remain a consuming preoccupation, inspired meaningfully by international policy guidance promulgated by the Committee on Payment and Market Infrastructures (CPMI) and the Board of the International Organization of Securities Commissions (IOSCO). I want to commend our colleagues in the policy realm for their clarion call and practical guidance which have goaded the industry to action across both of these critical fronts and resulted in demonstrable progress. Rather than provide mere updates on the status of our local resiliency and security initiatives in response to this charge, which have been the focal point of prior remarks, I hope instead to frame a few forward oriented, even provocative, questions to advance the industry discourse and to consider measures to meet the present challenge.

My perspective in pursuing these questions is not as overseer, or as policy maker, but as practitioner. Recall that following the designation by the Financial Stability Oversight Council (FSOC) of eight private sector financial market utilities as systemically important in 2012, the Board of Governors, acting in its oversight capacity (and inestimable wisdom!), committed to hold the wholesale services operated by the Reserve Banks to "as high or higher a standard" as it holds these private sector utilities; I assure you, we are experiencing, appropriately, that "or higher" side of the spectrum.

Like our infrastructure peers we are challenged to maintain the integrity of existing operations while contending with a brave new world; constrained by the demands of cost recovery and limited resources; challenged by the multiplicity of risk and compliance frameworks with which to comply and governance and stakeholder bodies to satisfy; conscious that we have to defend across an extended front while our adversary need only find a single point of entry; and keenly aware that we must be perfect every day, our foe successful but once. These are heavy burdens; I speak not to you, but of you.

With that context these remarks (I assure) reflect my own views and do not necessarily represent those of the Federal Reserve Bank of New York or the Federal Reserve System. (The best aside I ever heard in such a disclaimer came from former Supreme Court Justice Scalia, who was speaking to the Economic Club of New York, coincidentally the same week of his untimely passing. Upon noting that the views expressed do not necessarily represent those of the Supreme Court or the Government of the United States he added, slyly, "they rarely do!" This may be one of those occasions.)

Provocative Question Number 1: To what extent might the 2-Hour recovery time objective (known in industry parlance as the "2-Hour RTO") alter the incentive structure and reaction function of an FMI operator responding to an "extreme but plausible" event, precipitating sub-optimal outcomes?

Encouraging timely recovery following a severe disruption is a difficult thesis with which to argue, and our policy brethren are to be commended for urging financial market infrastructures (FMI) to examine critically their recovery capacities. This supervisory impulse has already yielded rich dividends in encouraging the wholesale services operated by the Reserve Banks to question assumptions, to consider out-of-the-box solutions, and to prioritize the achievement of technical diversity and the assurance of data integrity as anchor elements of our resiliency regime. At the same time, I wonder if specifying a two-hour threshold, notwithstanding its noble intent, may alter the incentives and the reaction function for a financial market infrastructure when responding to an extreme event. Could it, for example, encourage an FMI under duress to invoke prematurely contingency and failover procedures before an accurate diagnosis has been rendered? Alternatively, could it encourage an FMI to endeavor to return to normal operations prior to assurance of complete confidence in and sanitization of a production environment? To put it plainly: could the 2-Hour RTO unwittingly exacerbate rather than assuage the impact of a severe

disruption and/or impede prospects for recovery?

Reflecting upon these important questions, allow me to offer the following practitioner's interpretation, which I hope maximizes the value of the intent of the 2-Hour RTO while minimizing the downside risks of its mis-application. Rather than spending valuable time and industry resources re-litigating established supervisory guidance, we should consider the 2-Hour RTO as a vital organizing objective for planning purposes in peacetime, an aspiration to mobilize our best critical thinking to achieve, rather than an inviolable requirement that must be met in all circumstances recognizing the fog of war. A common understanding of this nuance across policy makers, supervisors, and operators strikes me as an important step forward both to focus our resiliency preparations in advance of an extreme event and to promote sound decision making in its remediation.

Provocative Question Number 2: To what extent does the proliferation of risk management and information security frameworks to which FMIs must comply risk creating a "check the box" mentality to security and resilience that diverts resources and management attention from other critical initiatives?

In framing this question, I asked my team to identify the cascading elements of supervisory guidance, oversight requirements, compliance regimes, and maturity frameworks with which we must contend related to risk management and information security. Their list in reply included 17 elements, admittedly some of which are self-induced and at different states of maturity and implementation. All are well intended, all are well designed, and all seek to address different and meaningful dimensions of the control environment. At the same time, compliance across this waterfront has become a consuming pre-occupation, and arguably one that if we are not careful can intrude upon the more tangible actions, and commitments, to increase the resiliency of our infrastructures and enhance the security of our networks.

Make no mistake: some form of risk management and control framework and corresponding policy and oversight guidance is absolutely essential to ensure a necessary discipline and accountability. This is not a call for a collective policy/supervisory/industry confab to craft a consolidated and harmonized uber-framework applicable to all parties in every circumstance—that would be as naïve as it is impractical. Indeed, the proliferation of frameworks can serve as a useful prism to ensure that no stone is left unturned in the development of an FMI's resiliency and security regime.

It is an appeal for awareness among standards bodies, supervisors, directors, and other organizational hierarchies that their individual actions, as rational and defensible as they may be, can have a suffocating effect in the aggregate. It is also an appeal for application of a time-honored managerial principle: greater autonomy and discretion of local management to develop a tailored approach that balances risk, complexity, and effectiveness appropriate for the needs of a particular institution, and a corresponding accountability upon local management for the results. Absent this sensibility I will state my fear plainly: beware the risk management industrial complex.

Provocative Question Number 3: To what extent might the invocation of manual procedures in "protracted outage" scenarios exacerbate rather than assuage from a financial stability perspective?

As you know, the CPMI-IOSCO cyber guidance exhorts financial market infrastructures to consider alternatives for processing critical transactions in the event automated means of recovery are not successful. We are therefore compelled to consider not merely "star wars" resiliency but also "stone age" contingency.

Locally, this work proceeds across multiple fronts, including analyzing and parsing our transaction flow to identify systemic activity, developing procedures and tools to facilitate

operations in a manual mode, and conducting simulation exercises with systemically important customers and interfacing financial market infrastructures to test our hypotheses and procedures.

I would be remiss, especially in recognition of our Clearing House venue, not to herald our partnership to support each other in the event either the Fedwire Funds Service or CHIPS experiences an operational disruption from which it cannot recover on a same day basis. While the two services are not perfect substitutes, these preparations include, for example, ensuring that the Fedwire Funds Service has sufficient capacity to accommodate the full complement of Chips activity, even on peak volume days, and that CHIPS can support all CHIPS-eligible Fedwire Funds traffic.

Arguably we have made greater progress advancing this body of work than in prior iterations, for which policy makers and supervisors deserve great credit. The application of critical thinking related to protracted outage has also yielded rich dividends in greater understanding of the end-to-end transaction flow and identification of potential points of vulnerability. These benefits are not to be dismissed.

But let us not delude ourselves: no matter how mature our framework for responding to protracted outage scenarios, no matter how sound our procedures, no matter how tested our protocol, we never want to rely on these measures from a contingency perspective. The ultimate lesson for any market infrastructure (and for their directors, governing bodies, and supervisors), is to so invest in resiliency and security in the core of its operation that it never finds itself in such a state.

The unintended consequences of a systemic infrastructure invoking protracted outage procedures in the event of a severe disruption are unknowable, and potentially raise reasonable questions as to whether the remedy may prove more toxic than the disease. Might it be preferable in select cases for a systemic infrastructure to declare a settlement holiday, unpalatable as that may be, or to work with relevant industry associations to declare that select markets are closed? Are the operational challenges of conducting manually activity that is fundamentally dependent upon automation a bridge too far? Do reconciliation challenges that become exponentially more complex as a disruption event extends in duration risk prospects for an orderly return home? Might industry dialogue to agree on the accepted circumstances and corresponding practices for dealing with a settlement holiday in a pre-meditated manner prove to be a superior course? These questions are not easy to answer, but it is very important in the present evolution of industry discourse that they are asked.

Provocative Question Number 4: To what extent do measures to address endpoint security risks, such as anomaly detection and assurance regimes, compel a new paradigm for the attribution of liability in the end-to-end payment transaction flow?

The reality that endpoint security breaches can undermine confidence in a network even when an operator's infrastructure has not been violated has spawned a novel body of work, including initiatives to enhance assurance that external endpoints are operating in adherence with an operator's strict security requirements, and the development of application tools such as anomaly detection to enhance a participant's management of fraud risk. As this work has progressed, it has been hard to ignore the concern that these initiatives and tools might shift the attribution of liability between participant and operator in a manner that could challenge an operator's ability to take appropriate precautionary measures to protect the network.

For example, if an operator becomes aware of security deficiencies at a participant as part of an assurance program, and the participant subsequently experiences a breach, does the operator incur some measure of liability for any loss? In such a circumstance, what degree of latitude does an operator realistically possess to deny a participant access to services that are essential

to the participant's ongoing viability? Moreover, if an operator detects potentially anomalous transactions coming from a participant, and does not intervene to prevent a transfer that is in fact fraudulent, does the operator assume a measure of liability for the resulting loss?

One way in which we have endeavored to manage these concerns is to provide tools and services to participants to manage their endpoint security risks locally, and in a manner that does not inappropriately shift liability from participant to operator, but admittedly these are difficult questions that compel novel thinking to ensure that the legal framework supporting the payment system is sufficiently robust and responsive to the escalating threat environment.

An additional, but not incidental, point on endpoint security: I observe a reluctance across the industry to share information about endpoint security experiences for fear that such transparency may expose an individual institution to concerns about its internal information security control environment. This is an understandable but parochial response which, collectively, has the potential to expose a network to considerable risk, especially if it inhibits socialization of known threat vectors and techniques. A solution to this vexing dynamic could be industry-led, operator-led, or supervisory-led. The impulse does not particularly matter; developing a venue for the coordinated sharing of endpoint intelligence and experience across the industry matters profoundly, and represents a fundamental mutual interest, compelling both leadership and collaboration.

Conclusion

I do not presume answers to these questions, and merely observe that they are important ones to inspire reflection in the collective pursuit of resilience and security for the anchor elements of our nation's financial system. What better venue than The Clearing House's payments symposium and annual conference to continue this discussion?

In response to this generational challenge, let us be sure not to construct an inflexible Maginot line whose rigidities are easily subverted by a creative and nimble adversary. Let us instead develop a coherent and integrated system that relies upon the classical elements of defense, but none of them exclusively: perimeter security to keep the adversary outside of the environment; defense in depth to safeguard our most critical assets; sophisticated intelligence to understand the adversary's tactics; robust surveillance to monitor for intrusion and ensure environmental integrity; rapid response to fend off attack; effective collaboration with allies to enhance information sharing and collective security; and a strategic reserve to respond deftly in the event of loss.

Make no mistake: as it relates to the wholesale services of the Federal Reserve Banks, we aspire not merely to a commercial standard of resiliency, or even a supervisory standard, but something approaching national security grade. We proceed on this trajectory from a position of strength, reflecting a record of experience that has endured the most severe terror attack in our nation's history, a financial crisis of historic proportion, and an array of extreme weather events. But in this sphere, either intentionally we are progressing, or inevitably we are regressing: there is no idleness.

Thank you for your generous attention today. I'd be pleased to respond to a few questions.