Dark Data, The Dark Web, and Your Mainframe Operations

Understanding the threat to your systems, data, and reputation

Prepared by Mark Wilson, Senior Director, Consulting Services,
BMC Mainframe Services by RSM Partners

Table of Contents

Introduction: Are You in Control?

When people talk about dark data, what are they referring to? For many organizations, this may simply be data they have gathered or obtained and then done little or nothing with. In the mainframe world, it can mean something else altogether. In addition to production data, we need to consider all the copies out there, for development and other purposes; and all the versions we don't know about, perhaps further copies made by developers with elevated access, and all those backups and archives? The reality is that a significant number of mainframe practitioners do not bother with data obfuscation (DO), even though tools are readily available. This is possibly due to perceptions about the time and costs involved. Indeed, in times of cost reduction and budget scrutiny, it is often security measures—especially if you believe you already have control—that are hit first. The question is, do we really have the level of control over our data that we think we have?

Are Things Worse Than We Think?

The bad actors don't even need to go after an organization's production data. Why bother, when they can, perhaps far more easily, go after a copy or a backup? It doesn't really matter that it's a few weeks old, there is still more than enough data, and it's all almost certainly unencrypted. As a result, the data ends up being traded on a dark web cryptomarket like Silk Road 3.0.[1] So, while we may be focused on protecting production data, we can actually miss the data theft being committed under our noses. At best, it's on the periphery of our vision. At worst, it's completely outside our sphere of control.

1 Online black market Silk Road was the first modern cryptomarket (aka darknet market), perhaps best known as a platform to sell illegal drugs. Users can browse anonymously and securely without traffic monitoring. The name comes from the historical network of trade routes between Europe, India, China, and other countries.

Silk Road, Tor, and Riffle

The Internet is a far bigger place than most people realize. We all know about Facebook, Amazon, eBay, and the rest, but what lurks beyond them? The surface web is only the tip of an iceberg: the 4 percent we can access via Google. The remaining 96 percent starts with deep web content such as academic databases, legal documents, subscription-only sites, and other similar resources. It then slides into proper illegality with the dark web: drugs, guns, hitmen, computer viruses, and, of course, personal data stolen in hacks and the sale of fake passports, IDs, and credit cards for financial and identity fraud, feeding on the easy availability of all that unencrypted personal and financial data. Use the Tor[2] network and you can find pretty much whatever you want. If Tor isn't fast or efficient enough for you, there's always Riffle.[3]

The dark web is a different world: think The Upside Down in the TV show *Stranger Things*, an alternate dimension in parallel with, and linked to, our own world. In the dark web, website addresses and URLs are not what you'd expect. They consist of random letters and numbers ending in .onion (a top-level domain host suffix designating an anonymous hidden service reachable via the Tor network). Such addresses are not actual DNS names and the .onion is not in the Internet DNS root. But with the appropriate proxy software installed (in most cases, the Tor browser bundle), Internet programs like web browsers can access sites with .onion addresses by sending the request through the Tor network. The purpose is to make both the information provider and the person accessing the information more difficult—if not impossible—to trace by each other, an intermediate network host or an outsider.

While Silk Road was the first and most famous cryptomarket, many similar marketplaces followed after the FBI took down the original site in 2013. Many cryptomarkets sell drugs—more than 55 single-vendor markets are included. When the AlphaBay market was shut down in 2017, it featured more than 40,000 sellers advertising some 250,000 items of illegal drugs and chemicals. But it's not only about drugs. It's extremely easy to find markets selling "high quality CVV and credit cards." Here's an example I use in presentations: "USA fake ID and credit cards plastics (not encoded) from a trusted source with Escrow... High Quality Bank Credit Card Blanks [*sic*]." The pitch names the bank and card provider brands available, "ALL HAVE UV + HOLOGRAMS!" before getting to the prices. The minimum order is 10, at just $10 each. Order 50 or more and the price drops to $7 each. It's a business. And guess what? In the real world you can also buy, legally and cheaply, a magnetic card reader writer encoder for less than $70, including free shipping.

Credit Card Plastics

0.06537656 BTC
This is a tutorial:

I will tell you where to buy USA fake ID and credit card plastics (not encoded) from a trusted source with escrow

What we offer:
High quality bank credit card blanks—Mastercard, VISA, AMEX, Discover.
All have UV + holograms!

Prices:
Min Order 10: $10 each ($100 total)

11-24: $9 each
24-49: $8 each
50+: $7 each

2 Tor ("The Onion Router") is free software that enables anonymous communication. It directs Internet traffic through a free worldwide volunteer overlay network to hide a user's location and usage from network surveillance and traffic analysis.

3 Riffle is an anonymous network that uses a verifiable shuffle. It's claimed to be ten times faster than an onion-based network.

As mainframe professionals, just how worried should we be about dark data and the dark web? As with most things in life, there's good news and bad news. The bad news is, the scale of the threat is huge and increasing in scale. The good news is that we can take focused action, adopt new approaches, and deploy specific tools to better protect our operations and mitigate the risks.

A Target for the Bad Actors

- Mainframes host critical core IT for 92 of the world's top 100 banks, 18 of the world's 25 top retailers, 23 of the 25 top airlines, and all of the world's top 10 insurers—and for 71 percent of Fortune 500 companies
- Mainframes run 30 billion transactions per day, hold 80 percent of the world's business data, and handle 90 percent of all credit card transactions
- Mainframes host more transactions daily than Google (1.3m/second CICS versus 68,542/second Google); including 55 percent of all enterprise transactions

"The Databerg" and Mainframe Operations

The dark web and cryptomarkets give serious value to our data. So why is the mainframe particularly at risk? It's basically a numbers game. Around 80 percent of the world's system of record data already resides on mainframe systems. More commercial transactions are processed on mainframes than on any other platform. And in today's connected Internet of Things (IoT) world, the mainframe no longer lives in splendid isolation. And so much of that mainframe data is being copied and copied and copied...

Only 12 percent of data can be considered business-critical; the visible part of the databerg. Below the surface, things start to get a little shadier. The next 23 percent of data is redundant, obsolete, trivial (Rot), with an estimated cost to global industry of $3.3 billion by 2020. So that just leaves 65 percent as "dark data," hidden within networks, people, and machines. So why does all this dark data exist? The main reason is that no tools exist to capture, unlock, and get rid of it. A significant amount of this data may also be incomplete. Of course, there's also the fact that there is simply too much data and unstructured data.

In the mainframe world, we have that slightly different take on data. We have production data: one copy. Then we have development data: how may copies might that entail? At the same time, developers may have decided to hang on to a particular copy of the dev data, the one they wanted to keep. What about the quality assurance/like-live data? Hopefully, there is just the one copy, but there may be more. Test data? Who knows? It's easy to lose count. And this may not only be about the data: it might be something as simple as a social security number embedded in JCL or SQL. How is that managed?

Do You or Don't You?

How worried should we be? The answer is "not much" if all the copies and versions of the data are obfuscated or masked, as they surely must be. The thing is, how much non-production data is adequately obfuscated or masked? I've heard reports from some quarters of this being on the order of less than 15 percent.

Data Obfuscation (DO)

- A type of data masking—encryption—resulting in unintelligible data, a.k.a. data scrambling and "privacy protection"
- Data is purposely scrambled to prevent unauthorized access to sensitive or confidential material
- Two types: cryptographic (encoding input data before transfer to another encryption schema) and network security (payload attach methods purposely used to avoid detection by network protection systems)

What Does This Mean for Mainframe?

Proper awareness is the start. We should clearly be more concerned about the data under our care, given that 80 percent of all active code runs on the mainframe, more than three-quarters of the world's system of record data resides on mainframes, and more commercial transactions are processed on mainframes than on any other platform. Where do we think our data might end up? With our operations an increasingly desirable target for the bad actors, with today's technologies getting rid of the concept of "mainframe isolation"—with potential access via laptops, tablets, smartphones, printers, even fridges—the dark web seems a highly likely destination for your data. To the bad actors, the mainframe is simply another system to hack into and compromise, while the prizes—in terms of data—are there in abundance.

So, what is that data worth? A huge variety of stolen information is available for sale on the dark web, including financial data and login details. Criminals can also access and buy the tools required to commit identity theft quickly and easily. The last time I looked, you could pick up a standard credit card including user data for as little as $15 each. A premium card with user data would set you back less than $28. And it gets worse. "Fullz" is a slang word used by hackers and data resellers to mean complete packages of an individual's identifying information, sold to identity thieves for use in criminal activity. Fullz typically include someone's name, their social security number, date of birth, account numbers, and other data.

The uncomfortable reality is that existence of the dark web and its cryptomarkets gives value to our mainframe and personal data. That value creates demand. With demand comes motivation. And so, we are under attack. None of us want our mainframe data for sale on the dark web. As a community of practitioners, we need to step up our game in mainframe security. Our systems are just another target to the criminals. We need to get on the front foot and be far more proactive rather than reactive in what we do and how we do it. Our clients, our employers, and our shareholders and regulators are looking to people like us to sort it out. And the tools do exist. Data obfuscation: use it. Then, there is pervasive encryption and the opportunities that it presents. In terms of changing our mindset, I've written elsewhere about adopting a genuinely **data-centric security model**.

Data-Centric Security

Too often in the past, the standard security posture was that everybody has read access to everything. That simply isn't possible in today's complex, connected, cybercrime world. If any individual can read and copy, there are multiple opportunities to exfiltrate data. If you're a bank and an individual can copy the DB2 database for your mainframe banking application and a list of clients with personal details—name, address, birthdate, social security number, credit score—that has real value. And if a hacker can also get a person's credit card 16-digit PAN, CVV, and zip code? Taking an "in -> out" approach rather than "perimeter-in" approach makes a lot of sense to me: start at the middle, then build out through the data, the application layer, and the network.

The first step is to find your data: identify it, understand the risks, (not all data poses the same risks) and then manage how data is handled and stored, using best practice approaches and the latest security and data protection technology (including DO and PE). Evaluate data at rest and in motion, consider any "contamination" to be worth tracking, and classify differentiate your data. You then need to monitor access, which can include using machine learning (ML) and artificial intelligence (AI) to reveal anomalies and possible threats: who is accessing the data, where, and how often? What are the baselines for "normal" activity and who is deviating, where, and when? Step three is to limit access in ways that make sense for your data, your organization, and your people. Carefully manage authorizations; who has access now and who has accessed before?

We can never mitigate all the risks. But what we should be doing is taking far more proactive and assertive steps to tackle the threats posed by dark data, to ensure valued data assets cannot be accessed, compromised, and sold on the dark web, and so avoid hard-won reputations being dragged through the dirt.

For more information

To learn more about Dark Data, The Dark Web, and Your Mainframe Operations, please visit bmc.com

About BMC

From core to cloud to edge, BMC delivers the software and services that enable over 10,000 global customers, including 84% of the Forbes Global 100, to thrive in their ongoing evolution to an Autonomous Digital Enterprise.

BMC – Run and Reinvent **www.bmc.com**

522074