FSPOS

Finansiella Sektorns Privat-Offentliga Samverkan

Former och metoder för test av cyberresiliens

Version: 2018-09-06
FSPOS FG Cybersäkerhet

Inledning

Övning och test är en framgångsfaktor för att organisationer bättre ska kunna hantera cyberrelaterade risker. Sådana aktiviteter möjliggör verifiering och utvärdering av hur effektivt organisationens ramverk för cyberresiliens är, och är ett värdefullt verktyg för att identifiera sårbarheter och förbättringsområden.

CPMI (Committee on Payments and Market Infrastructures) och IOSCO (International Organization of Securities Commissions) har utvecklat särskilda riktlinjer, som bland annat riktar in sig på hur finansiella aktörer kan använda övningar och tester för att stärka sig cyberresiliens. Samma riktlinjer definierar cyberresiliens som förmågan att förutse, motstå, begränsa och snabbt återhämta från en cyberattack.[1] Under 2016 upprättade FSPOS en särskild projektgrupp som tog fram en rapport om tolkning och tillämpning av riktlinjerna. Under 2017 och 2018 har gruppens arbete fortsatt i form av en fokusgrupp, FG Cybersäkerhet. Ett område från 2016 års analys som bedömts som särskilt viktigt och utmanande är just test och övning.

Detta dokument beskriver former och metoder för att testa organisationers förmåga inom cyberresiliens, såsom sårbarhetsanalyser, scenariotester, penetrationstester och så kallade Red team-tester. Innehållet är en lägesbeskrivning inom ett område som utvecklas snabbt eftersom behovet av information ökar och förändras som en följd av att hotbilden och riskerna förändras. FSPOS är ett nätverk för samverkan mellan organisationer i finansiell sektor och har bland annat till uppgift att sprida kunskap om hur operativa risker kan hanteras för att samhällsviktiga finansiella tjänster ska fungera. Detta dokument är ett led i denna kunskapsspridning. Dokumentet kan även utgöra en grund för fortsatt analys inom området, exempelvis gällande behov av att utveckla befintliga tester och övningar inom den finansiella sektorn.

Syfte

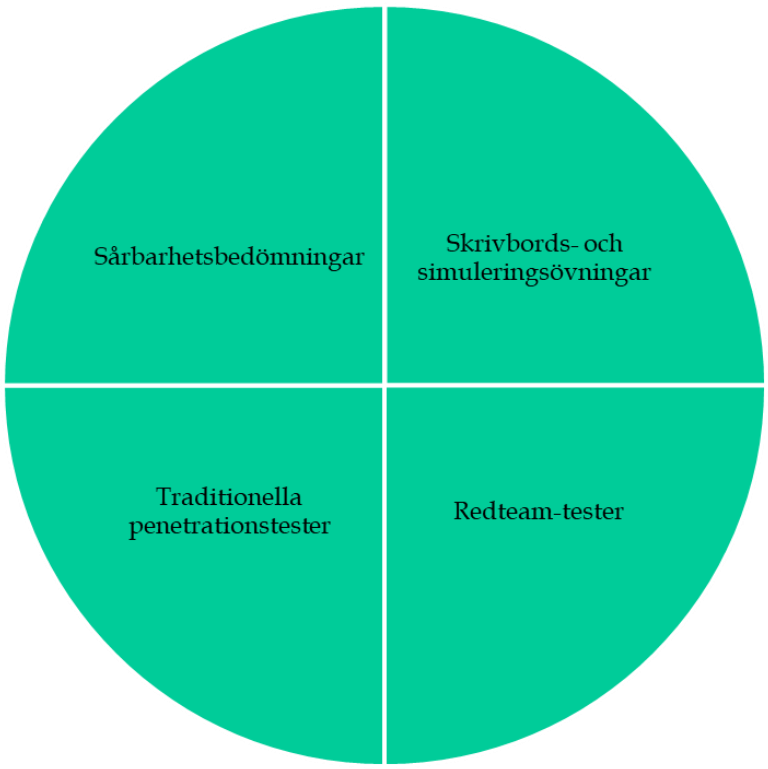
Syftet med detta dokument är att stärka den finansiella sektorns förmåga att utvärdera effektiviteten i sitt arbete med cyberresiliens med hjälp av test och övning. Målet är att beskriva olika former och metoder för tester, med utgångspunkt i god praxis. Beskrivningarna bygger dels på en genomgång av relevanta rapporter och vägledningar, dels på diskussioner med experter som arbetar med berörda frågor inom sina respektive organisationer. Ett inledande avsnitt ger en översikt av vilka test- och övningsalternativ som finns och används. I det efterföljande avsnittet ligger tyngdpunkten på en av dessa övningsformer, så kallade Red team-tester. Red team-tester är riktade penetrationstester som bygger på specifik information om den övade organisationens verksamhet och realistiska hot, såsom angripares metoder. Testerna utförs ofta men inte alltid av en extern tredje part.

---

[1] CPMI/IOSCO (2016), *Guidance on cyber resilience for financial market infrastructures*.

Översikt av olika former för övning och test av cyberresiliens

Det finns ett flertal olika övnings- och test-alternativ som aktörer i den finansiella sektorn kan använda för att utvärdera effektiviteten i sitt arbete med cyberresiliens. Dessa kan med fördel komplettera varandra som en del av ett samlat test-program. I vissa fall kan de olika formerna även kombineras och överlappa varandra. Enligt riktlinjer från CPMI/IOSCO, så bör test av cyberresiliens omfatta *sårbarhetsbedömningar*, scenariobaserade tester (*skrivbords- och simuleringsövningar*), *traditionella penetrationstester* och *redteam-tester*.



Figur 1. Olika typer av test och övning av cyberresiliens.

1.1 Sårbarhetsbedömning

En sårbarhetsbedömning är processen att definiera, identifiera, klassificera och prioritera sårbarheter i datorsystem, applikationer och nätverksinfrastrukturer i syfte att ge organisationen en bedömning av relevanta sårbarheter. Bedömningen ger organisationen nödvändig kunskap och medvetenhet för att förstå cyberhot och hantera dem på ett lämpligt sätt.[2] [3]

Sårbarhetsbedömningar använder ofta automatiserade testverktyg som skannar organisationens IT-miljö och klassificerar sårbarheter genom att matcha dem mot redan kända systembrister. Systembristerna kan röra sig om exempelvis saknade patchar och föråldrade protokoll, certifikat och tjänster. Sårbarhetsscanningarna kan både utföras internt och av externa leverantörer. Sårbarhetsbedömning syftar vanligtvis till att validera en miniminivå för den säkerhetsnivån som ska tillämpas. Sådana bedömningar ses därför ofta som en förutsättning och som en mer grundläggande nivå än mer specialiserade penetrationstester eller komplexa övningar.[4]

National Institute of Standards and Technology (NIST) definierar sårbarhetsbedömning som en systematisk granskning av ett informationssystem eller en produkt för att fastställa säkerhetsåtgärdernas tillräcklighet, identifiera säkerhetsbrister, tillhandahålla data för att förutsäga effektiviteten av de föreslagna säkerhetsåtgärderna och bekräfta att sådana åtgärder är tillräckliga efter genomförandet.[5] *CIS Critical Security Controls for Effective Cyber Defense* rekommenderar löpande sårbarhetsmätningar som en del av ett effektivt cybersäkerhetsarbete.[6]

Ett sätt att klassificera sårbarhetsmätningar är att dela in dem i *Autentiserade skanningar* och *Icke-autentiserade skanningar*. Autentiserade skanningverktyg tillåts samma systemåtkomst som privilegierade systemanvändare, vilket gör att skannern typiskt sett får möjlighet att komma djupare in i organisationens system. Därigenom kan också djupare sårbarheter identifieras, exempelvis förekomst av skadlig programvara, undermåliga lösenord, konfigurationsproblem. Genom att autentiserade skanningar efterliknar vad en användare har åtkomst till, så ger det en bild av vilken potentiell skada som kan uppstå via systemanvändaren och tilldelade privilegier.

---

[2] Techtarget (2018), "Vulnerability assessment (vulnerability analysis)", https://searchsecurity.techtarget.com/definition/vulnerability-assessment-vulnerability-analysis [Tillgänglig 2018-05-11].

[3] Enligt riktlinjer från CPMI/IOSCO är sårbarhet en svaghet, mottaglighet eller fel i ett system som en angripare kan komma åt och utnyttja för att kompromettera systemets säkerhet. Sårbarhet härrör från tre delar: förekomsten av en mottaglighet eller fel i systemet; En angripares tillgång till den felaktigheten och en angripares förmåga att utnyttja felet. CPMI/IOSCO (2016), *Guidance on cyber resilience for financial market infrastructures,* sida 26.

[4] CREST (2017), *A guide for running an effective Penetration Testing programme*.

[5] NIST (2013), *Glossary of Key Information Security Terms*.

[6] CIS Controls: https://www.cisecurity.org/controls/

Vid icke-autentiserade skanningar tillåts inte sådana privilegier. Dessa skanningar söker istället sårbarheter baserat på vilka system som är tillgängliga från utsidan av organisationen. Jämfört med autentiserade skanningar, så ger dessa tester typiskt sett en lägre detaljgrad avseende organisationens sårbarheter. Testen kan dock göra nytta genom att identifiera potentiella externa vägar in i systemmiljön.[7]

Ytterligare ett sätt att klassificera sårbarhetsmätningar är att dela in dem i *network-based scans*, *host-based scans*, *wireless network scans*, *application scans* och *database scans*.[8]

Network-based scans används för att identifiera eventuella nätverkssäkerhetsattacker. Host-based scans används för att lokalisera och identifiera sårbarheter i servrar, arbetsstationer eller andra nätverkshanterare. Bedömningen granskar exempelvis portar och tjänster som även kan vara synliga för nätverksbaserade skanningar, men host-based scans ger större insyn i konfigurationsinställningarna och patchhistoriken för skannade system. Wireless network scans av organisationens Wi-Fi-nätverk fokuserar på attacker i den trådlösa nätverksinfrastrukturen. Förutom att identifiera sårbara åtkomstpunkter kan bedömningen även fokusera på verifikation av att nätverket är konfigurerat säkert. Application scans kan användas för att testa webbplatser för att upptäcka sårbarheter och felaktiga konfigurationer i nätverks- eller webbapplikationer. Database scans kan användas för att identifiera svaga punkter i en databas för att förhindra skadliga attacker, såsom SQL-injektionsattacker.[9]

Det finns en god erfarenhet från aktörer i den finansiella sektorn av att använda sig av olika sorters skanning- och rapportverktyg för att stärka sin cyberresiliens. Ett exempel på sådana mätningar är genomförandet av veckovisa skanningar av organisationens nätverk, påslagna servrar samt arbetsstationer, m.m. Mätningarna görs ofta utifrån standardiserade koder (Common Vulnerabilities and Exposures, CVE) av kända sårbarheter, som laddas upp till skanningverktyget via en global leverantör. Skanningarna kan exempelvis leta efter missade patchningar och installationer som har uteblivit. Vissa organisationer använder en kombination av mer koncernövergripande skanningar och mer specifika skanningar på systemnivå. Organisationsövergripande mätningar har ett bredare fokus och ”skannar allt”, exempelvis på databas- eller servernivå. De mer specifika mätningarna inriktar sig istället på utpekade kritiska tillgångar, såsom organisationens kritiska system.

Sårbarhetsmätningen ger i sin tur en rapport eller flera rapporter som tillställs organisationen vid ett senare tillfälle, alltifrån några dagar senare till månadsvis. Sådana rapporter beskriver bland annat identifierade sårbarheter, som kan klassificeras enligt en given skala, exempelvis från P1 till P5, där P1 utgör mer

---

[7] ”Sårbarhetsscanning och penetrationstester – när, hur och varför ska metoderna användas?” https://www.sentor.se/nyheter/sarbarhetsscanning-och-penetrationstester/

[8] ISACA (2017), Vulnerability Assessment.

[9] Techtarget (2018), “Vulnerability assessment (vulnerability analysis)”, https://searchsecurity.techtarget.com/definition/vulnerability-assessment-vulnerability-analysis [Tillgänglig 2018-05-11].

allvarliga och prioriterade sårbarheter. Ett annat exempel på klassificering är enligt nivåerna low, medium, critical. Kopplat till klassificeringen finns ofta en fördefinierad åtgärdstid som kan vara baserad på organisationens patchstandard, exempelvis att de mest prioriterade sårbarheterna ska mitigeras inom två veckor.

I många fall utgör åtgärdstiden inte ett problem, exempelvis om en patchning behöver göras inom en befintlig systemversion. I andra fall kan det vara betydligt mer utmanande, till exempel om den berörda programvaran inte är kompatibel och behöver uppgraderas. I det senare fallet finns dock annan hantering att tillgå för att kortsiktigt mitigera sårbarheten, exempelvis att ändra behörigheter eller genomföra justeringar i nätverket för att minska attackvektorn.

I och med att mätningar görs på olika teknisk och organisatorisk nivå, så har rapporterna även olika syften, format och målgrupper. Organisationsövergripande mätningar kan ha ett internationellt/globalt perspektiv, med rapporter i excelformat som exempelvis mottas av ett koncernövergripande skanningteam. Det som identifieras av mätningarna kan dock behöva skickas vidare till och följas upp med berörda team i organisationen. Övergripande mätningar rapporteras i regel även till hög nivå inom organisationen. Rapporter på systemnivå är typiskt sett mer lättillgängliga och vanligen i PDF-format, bland annat då målgruppen för rapporterna är såväl riskägare från verksamheten och tekniska systemägare. Rapporterna kan även inkludera historik och lösningsförslag, exempelvis genom att CVE-koder anges. Både verksamhet och IT involveras i utformningen av rapporterna och vad man vill få ut av dem. Systemägare behöver ofta hantera och mitigera riskerna, medan verksamhetens riskägare är de som använder systemen. Verksamheten behöver därför förstå riskbilden och kunna följa upp och kravställa mot systemägare. Att involvera rätt målgrupper i utformningen av rapporterna är en framgångsfaktor i arbetet med sårbarhetsmätningar, men kan även vara utmanande om verktygen inte är flexibla. En nyckel till att få ut så mycket som möjligt av sårbarhetsmätningar är att kunna skära resultat på rätt sätt för att få fram olika detaljeringsgrad över sårbarheter och statistik. En framgångsfaktor för att uppnå detta är mätningarna inte bara blir en tekniskt fråga för systemägare, utan att även riskägare och VD- och styrelsenivå involveras i utformning och som mottagare av resultat. I större organisationer kan det även finnas en stor nytta av att det finns en ytterligare nivå däremellan, exempelvis genom CISO-rollen, som kan fånga upp och lyfta resultaten till ledningsnivå. Resultat och slutsatser som bör anpassas till olika målgrupper i organisationen, exempelvis så att VD och styrelse ges en relativt mer övergripande och mindre detaljerad översikt jämfört med målgrupper på teknisk nivå. En framgångsfaktor är att säkerställa informationsflödet mellan olika delar av organisationen, t.ex. mellan IT, verksamhet och ledningsnivå. Ett sätt att uppnå detta är att arbetet hålls ihop av en projektledare och att exempelvis CISO och representant från systemägarteam finns med från början av arbetet för att ge fokus och prioritering.

1.2 Skrivbords- och simuleringsövningar

Till skillnad från mer tekniska och skarpa penetrationstester, så kan organisationen använda sig av olika former av scenarioövningar, som bygger på deltagarnas diskussioner och agerande utifrån fiktiva scenarier. FSPOS Fokusgrupp Cybersäkerhet har tidigare konstaterat att övning med fokus på cyberresiliens medför utmaningar, inte minst med att hitta relevanta övningsformer och scenarier. Arbete med övning och uppföljning kan också bli resurskrävande.

I viss utsträckning övar redan aktörer i den finansiella sektorn med fokus på cyberresiliens, exempelvis avseende integritetsintrång. Sådana övningar har dock främst haft ett processfokus och genomförts som skrivbordsövningar. Dessutom har övningar som genomförts mer fokuserat på åtkomst och mindre på specifika felaktigheter. Simuleringsövningar och med fokus på system är svårare och mer kostsamt, bland annat då det kräver en lämplig testmiljö. Det bedöms därför som utmanande att öva skarpt hur man hanterar scenarier där angripare kommer åt skyddsvärda tillgångar. Att öva ett skarpt scenario med avvikande transaktioner, gå tillbaka till en ”golden point” och skicka om transaktionerna är komplext.

För att testa organisationens cyberresiliens, så bör övningarna behandla ett lämpligt brett spektrum av scenarier, inklusive simulering av extrema men troliga cyberattacker, och bör utformas för att utmana organisationens om respons, återupptagning och återhämtning. Underrättelseinformation, såsom information om verkliga angripares metoder och mål, kan med fördel användas som ingångsvärde till scenarioutvecklingen. Övningarna behöver inte inskränka sig till tekniska aspekter utan bör även testa deras personal och processers förmåga att hantera scenarierna.[10]

Några aspekter som kan testas är exempelvis organisationens reaktionsförmåga vid cyberrelaterade händelser, medvetandenivån om kundpåverkan, samverkan med tredjepart, intern och extern kommunikation vid cyberrelaterade incidenter, beslutsfattande baserat på begränsad information, m.m. Övningarna kan även skilja sig baserat på målgrupp, exempelvis om de riktar sig mot en mer strategisk ledningsnivå, en taktisk koordinerande nivå, eller mer operativ responsteam-nivå.[11]

Ett exempel gällande övning av cyberresliens är FOI:s övningsmiljö CRATE.[12] Miljön gör det bl.a. möjligt att öva simulerade attacker i samband med utbildningar där man samtidigt har möjlighet att tillämpa sina egna processer. Ett annat exempel är övningarna CAPP som arrangeras av FS-ISAC.[13] Övningarna genomförs under två

---

[10] CPMI/IOSCO (2016), *Guidance on cyber resilience for financial market infrastructures,* sida 18.

[11] EY (2017), *Cybersecurity incident simulation exercises – Is simply waiting for a security breach the right strategy*.

[12] CRATE, Cyber Range And Training Environment, har byggts upp av och förvaltas av FOI (Totalförsvarets forskningsinstitut).

[13] CAPP (Cyber Attack against Payment Processes) genomförs av FS-ISAC (The Financial Services Information Sharing and Analysis Center), som är en icke-vinstdrivande global organisation för finansiell sektor, som ägs av privata aktörer och som utformar gemensamma strategier, verktyg och samordning.

dagar, där de övade deltar 1-1,5 timme per dag. Ett ytterligare övningsformat är DECIDE, som är utvecklat av bl.a. amerikanska banker. Övningsplattformen, för framförallt skrivbordsövningar, har tagits fram av NUARI på Norwich University i USA och på uppdrag av Department of Homeland Security (DHS) som är MSBs motsvarighet i USA.[14]

För en mer detaljerad beskrivning av hur övningar kan planeras, genomföras och utvärderas, så rekommenderas läsaren FSPOS vägledning *6 steg till bättre övningar*. Generellt kan övningar delas in i två format – *skrivbordsövningar* och *simuleringsövningar*. De två övningsformaten har sina särskilda förutsättningar och fördelar. Övningar kan variera från att vara mycket enkla till mer omfattande och komplexa. En övning som ska vara realistisk och verklighetstrogen kräver betydligt större resurser än en diskussionsbaserad övning. Vilket övningsformat som väljs ska alltid utgå från övningens syfte och mål. Övningsformatet väljs bland annat med hänsyn till organisationens mognad, övningskunskap, storlek, komplexitet och nisch inom den finansiella sektorn. Det är viktigt att organisationen utvecklar sin övningserfarenhet innan avancerade övningar genomförs. En övning som är för svår för organisationen kan leda till svårtolkade eller bristande resultat. Utöver organisationens förutsättningar är det också viktigt att basera valet av övningsformat på det syfte och de mål som övningen avser uppnå.

En skrivbordsövning är en relativt enkel form av övning som ofta inte kräver omfattande förberedelser. Skrivbordsövningar genomförs med fördel för mindre grupper, ibland i syfte att förbereda deltagarna för mer komplexa övningar. Deltagarna samlas för en diskussion kring ett eller flera scenarier som utifrån givna frågeställningar ska diskuteras och/eller besvaras.

En simuleringsövning innebär att övningsorganisationen försöker skapa en situation som är så lik en verklig händelse som möjligt. De övade utsätts för ett stort antal händelser i realtid och får agera precis som de skulle gjort i en skarp situation. De övade agerar kontinuerligt, exempelvis genom att samverka med relevanta aktörer, kommunicera externt och internt, upprätta lägesbild eller liknande, baserat på den information de ges allteftersom övningen fortskrider. Verifieringsgraden på simuleringsövningar är hög, då övningen är mer testande och möjliggör kvalitetssäkring av exempelvis roller och ansvar eller av rutiner. Övningsformatet är samtidigt mer komplext och kräver mer omfattande förberedelser än en skrivbordsövning. Simuleringsövningar har ett lärande och testande fokus. De övade får agera och finna lösningar, testa hur det är att agera under stress, analysera situationen och utifrån det ta beslut.[15]

---

[14] Se https://decideplatform.com/

[15] FSPOS (2017), *6 Steg till bättre övningar*, sida 22-24.

1.3 *Traditionella penetrationstester*

Genom penetrationstester kan tekniska experter simulera realistiska attacker för att identifiera metoder för att kringgå säkerhetsfunktionerna hos ett program, system eller nätverk, men även i processer och hos organisationens personal. Ett penetrationstest är således en ”etisk attack” som syftar till att visa eller validera effektiviteten av säkerhetskontroller i en viss miljö. Testerna syftar till att idenfiera och påvisa sårbarheter inom en avgränsad miljö, exempelvis för en särskild applikation eller ett nätverk.[16]

Penetrationstest är generellt uppbyggda enligt en manuell testprocess som är utformade för att kunna gå djupare än generiska svar och automatiserade skanningar. Penetrationstest kan även inkludera en blandning mellan manuella och automatiserade inslag. Till skillnad från sårbarhetsmätningar, som är ofta är helt automatiserade, så kräver penetrationstest ofta personell expertis eftersom testet syftar till att identifiera sårbarheter och ingångar i systemmiljön som inte nödvändigtvis är kända sedan innan.[17]

Testerna kan exempelvis identifiera hur väl ett system tolererar riktiga angreppsmönster, hur sofistikerat ett angrepp behöver vara för att kunna kompromettera systemet, eller vilka motåtgärder som kan motverka hot mot systemet. En målsättning med penetrationstester kan även vara att fastställa den egna organisationens förmåga att upptäcka attacker och hantera dem på ett lämpligt sätt.

I praktiken så innebär penetrationstester att man genomför riktiga attacker mot organisationens verkliga system och med verktyg och tekniker som vanligtvis används av angripare. De flesta penetrationstester innebär att man utnyttjar kombinationer av sårbarheter i ett eller flera system för att få mer åtkomst än vad som skulle kunna uppnås genom en enda sårbarhet. Värt att betona är att penetrationstester ofta även inkluderar icke-tekniska metoder, såsom att bryta mot fysiska säkerhetsskydd, stjäla utrustning, m.m.

Två vanliga former av penetrationstester är *applikationspenetrationstest* (vanligen avseende webbapplikationer) som identifierar tekniska sårbarheter och *infrastrukturpenetrationstest*, som granskar servrar, brandväggar och annan hårdvara för säkerhetsbrister. Andra former är penetrationstester för mobilapplikationer, klient-server, enheter (arbetsstationer, laptops, mobiltelefoner, etc.), telefoni/VoIP eller trådlös infrastruktur. Olika moment i penetrationstesten inkluderar ofta efterforskning, identifiering av sårbarheter, utnyttjande av svagheter, rapportering av resultat, samt åtgärdande av brister.[18]

---

[16] NCC Group, “Penetration testing: Thinking in scenario“, https://www.nccgroup.trust/us/about-us/newsroom-and-events/blog/2017/july/penetration-testing-thinking-in-scenarios/ [Tillgänglig 2018-05-28].

[17] CREST (2017), *A guide for running an effective Penetration Testing programme*.

[18] CREST (2017), *A guide for running an effective Penetration Testing programme*.

Givet att faktiska attacker genomförs mot riktiga system, så finns risk att penetrationstestet leder till att system skadas eller på annat sätt blir oanvändbara eller leder till oönskade konsekvenser. Penetrationstester behöver därför föregås av noggrann analys och överväganden för att minimera sådana risker.[19]

1.4 Redteam-tester

Med redteam-tester menas tester där en grupp testar organisationens kontrollfunktioner och rutiner för potentiella sårbarheter. Ett redteam kan bestå av såväl egna anställda som utomstående experter. Syftet med redteam-tester är att identifiera och beskriva potentiella angripares attackväg.

Namnet har sitt ursprung i militärens test av stridsberedskap. Övningsstrukturen består vanligtvis av två grupper som är ställda mot varandra. Det röda laget består av professionella "etiska" hackare (från den egna organisationen eller från en anlitad tredje part) som attackerar det blå (defensiva) lagets informationssystem. Det blå laget känner inte till ett Red team-test på förhand och har till uppgift att skydda sina system och upptäcka intrång och allvarliga incidenter eller incidenter med potential att eskalera till större händelser. Ett redteam utgör på så vis en ”sparring-partner” åt den egna organisationen, som möjliggör realistiska förberedelser för riktiga angrepp. Inom den testade organisationen, känner endast ett fåtal personer till testet, medlemmarna i ett så kallat White team.[20]

Redteaming simulerar en verklig bild av vad en angripare skulle göra för att genomföra ett angrepp. Testen fokuserar typiskt sett inte bara på tekniska aspekter, såsom nätverksinfrastruktur eller webbapplikationer, utan även på potentiella svaga punkter relaterade till processer och personal. Det gör testerna till mer komplexa och djupgående än andra traditionella testformer. Ytterligare djup ges typiskt sett genom att testets målsättningar motsvarar troliga mål för angripare, exempelvis: Är det möjligt att komma åt information om kunders kortbetalningar? Är det möjligt att komma åt personuppgifter med påverkan på efterlevnaden av GDPR? Nivån på redteamtesterna bör anpassas till verksamhetens mognad och behov. Ett simulerat angrepp från ett redteam, kan omfatta alltifrån generiska phishing-kampanjer till kommersiellt tillgängliga verktyg och mer avancerade fjärråtkomstverktyg (remote access tools, RATs).[21]

Under de senaste åren har flera ramverk utvecklats för så kallad Threat Intelligence-based Ethical Red Teaming (TIBER), där scenarierna baseras på verkliga angripares tekniker och metoder (tactics, techniques and procedures, TTPs). Flera av testkoncepten bygger på att de testade organisationerna anlitar externa leverantörer,

---

[19] NIST (2008), *Technical Guide to Information Security Testing and Assessment*.

[20] CATS (2011), *Handbook for planning, running and evaluating information technology and cyber security exercises*.

[21] BSI (2017), *Enhancing your penetration testing regime – A whitepaper.*

som uppfyller särskilda ackrediteringskrav, av underrättelseinformation för scenario (Threat Intelligence) och leverantörer av genomförande av penetrationstest.

1.5 Analys och jämförelse mellan övningsformerna

Samtliga testformer kan utvecklas och användas av såväl den egna organisationen som av en extern leverantör. En fördel med att använda egna resurser är att den organisationsspecifika kunskapen är högre och att man kan vilja undvika tredjepartsrisker. Externa leverantörer kan samtidigt bidra med erfarenhet, nya perspektiv och effektivitet. En extern part kan även ge en mer objektiv bild av nätverksmiljön och undvika intressekonflikter som kan uppstå inom den egna organisationen. Flera beskrivningar av testkoncept framhåller att sårbarhetsmätningar, traditionella penetrationstester, skrivbords- och simuleringsövningar, samt redteamtester bör ses som komplement till varandra, snarare än substitut.[22]

Sårbarhetsmätningar är ofta automatiserade och är därför inte lika personalintensiva som övriga testformer. En styrka är samtidigt att automatiserade lösningar kan genomföras oftare, vilket kan vara motiverat från ett effektivitetsperspektiv. Enklare sårbarhetsmätningar kan vara relativt lätta att utföra, medan penetrationstester kräver mer erfarenhet och teknisk expertis. Erfarenhet från aktörer i den finansiella sektorn understryker att det krävs rätt kompetens för att sätta upp, tolka och underhålla särbarhetsmätningarna. Arbetet kräver att rätt saker skannas och att man behåller fokus. Både ingångsvärden och resultatet från mätningarna är färskvara. Sårbarheter kan utvecklas över tid, exempelvis att en sårbarhet som har klassificerats på en viss nivå ändras till en högre eller lägre klassificeringsnivå. Därutöver behöver åtgärder följas upp löpande. En utmaning som har nämnts av aktörer i den finansiella sektorn är att förstå och sätta upp mätningarna på rätt sätt. Då icke-autentiserade mätningar skannar på system som är tillgängliga utan användares privilegier, så fångar dessa i regel upp färre sårbarheter än autentiserade mätningar. Om organisationen använder sig av en större andel icke-autentiserade mätningar, så finns en risk för att resultaten undervärderar sårbarheterna om resultaten inte tolkas på rätt sätt.

Penetrationstester behöver inte genomföras lika ofta som sårbarhetsmätningar. En tumregel kan dock vara att genomföra penetrationstester årligen eller vid implementering av nya applikationer och system. En utmaning med sårbarhetsmätningar är förekomsten av så kallade ”falska positiva” (false positives) i skannerns automatgenererade rapporter, med andra ord att en skanner varnar för brister som egentligen inte finns. Risken är då att resurser används till att minska sårbarheter och rätta till brister i onödan.[23] Sårbarhetsmätningar utgår ofta från en mer generell grund än riktade penetrationsattacker. Sårbarhetsmätningarna utnyttjar således inte de sårbarheter som kan identifieras för organisationen för att replikera och simulera en riktig attack. Jämfört med skrivbords- och simuleringsövningar, eller med

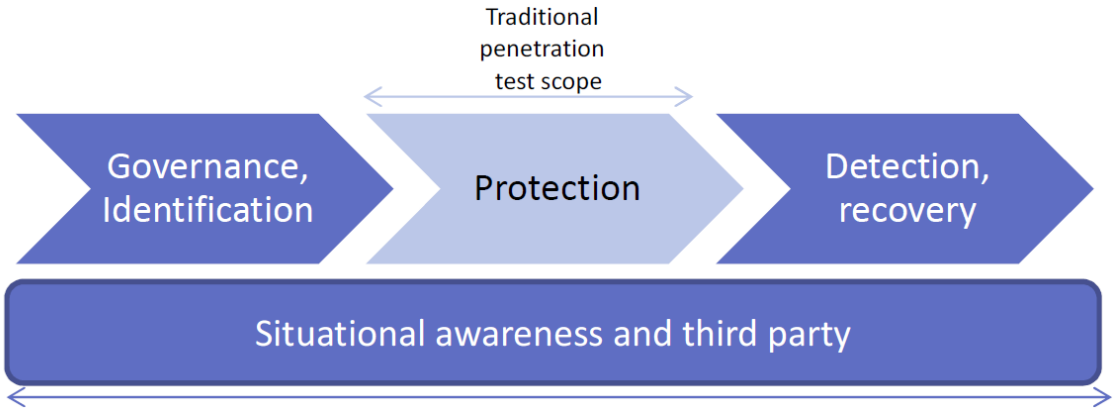
22 TIBER-NL (2017), *How to conduct the TIBER-NL test.*

23 Computer Sweden (2013), “Sårbarhetsskanning är inte penetrationstester”.

mer avancerade redteam-tester, så beaktar sårbarhetsmätningar heller inte de övergripande säkerhetsprocesser, rutiner och personella risker som påverkar cyberresiliensen för organisationen.

Traditionella penetrationstester kan ge värdefulla resultat, exempelvis om sårbarheter, ofta kopplat till enskilda system och för isolerade miljöer. Däremot täcker traditionella penetrationstester sällan hela den risk- och hotbild som finansiella aktörer utsätts för och omfattar inte processrisker och personella risker. För dessa kan skrivbords- och simuleringsövningar ge ett kompletterande perspektiv. Även mer avancerade redteam-tester kan ta hänsyn till en bredare hotskala och uppnå ett större djup, genom att testet är mer komplext och sofistikerat.

I flera sammanhang poängteras även att traditionella penetrationstester inte fullt ut möjliggör mätning av organisationens förmåga att identifiera och bemöta angrepp, något som emellertid utgör en central del av så kallade *Threat Intelligence-based Ethical Red Teaming*. Sådana redteam-testkoncept är mer avancerade, med realistiska och aktuella metoder som motsvarar de som används av faktiska angripare, och innehåller i några fall även fördefinierade indikatorer för att mäta förmågor.[24] Ett sätt att särskilja traditionella penetrationstester från Threat Intelligence-based Ethical Red Teaming ges i nedanstående figur.[25]



Figur 2. Traditionella penetrationstester i relation till Threat Intelligence-based Ethical Red Teaming.

Traditionella penetrationstester fokuserar således mer på skydd mot angrepp, medan Threat Intelligence-based Ethical Red Teaming även omfattar områden som identifiering, återställning, styrning, situationsmedvetenhet och tredjepartsberoenden.[26]

---

[24] ECB (2018), *TIBER-EU FRAMEWORK – How to implement the European framework for Threat Intelligence-based Ethical Red Teaming*.

[25] MKMA (2016), *CYBER RESILIENCE ASSESSMENT FRAMEWORK, CONSULTATION DRAFT*.

[26] MKMA (2016), *CYBER RESILIENCE ASSESSMENT FRAMEWORK, CONSULTATION DRAFT*.

Ramverk för Threat Intelligence-based Ethical Red Teaming

Detta kapitel beskriver olika ramverk för särskilda redteam-tester, så kallade Threat Intelligence-based Ethical Red Teaming, då flera initiativ sådana testformat är under utveckling och har tagits fram under de senaste åren. Redteam-tester har även varit ett särskilt föremål för diskussioner i FSPOS Fokusgrupp Cybersäkerhet inom ramen för tidigare års aktiviteter, exempelvis vid genomgången av riktlinjerna *Guidance on cyber resilience for financial market infrastructures* från CPMI/IOSCO.

De testkoncept som är mest utförliga gällande dessa tester är det brittiska konceptet CBEST, det holländska TIBER, Hong-Kong-baserade iCAST, samt ramverket TIBER-EU som har utvecklats av ECB. Utöver de mer mogna koncepten, så pågår även utveckling av liknande format i Danmark. Det första Red-teamtestet inom ramen för det danska initiativet förväntas vara klart under 2019 och formatet kommer att baseras på de i Storbritannien, Holland och på ECB:s ramverk.[27] I nuläget är tillgängliga beskrivningar om detta arbete begränsat, varför det danska formatet därför inte beskrivs mer ingående i denna rapport.

2.1 CBEST (UK)

Det av flera liknande test-ramverk som först etablerades är det brittiska CBEST. Konceptet initierades 2013 av finansiella myndigheter (Bank of England, Her Majesty's Treasury och Financial Conduct Authority), med konsultation av aktörer i den finansiella sektorn och leverantörer av underrättelsetjänster (threat intelligence) och penetrationstester. Den externa leverantören av penetrationstest utgör ett redteam för respektive testtillfälle. Bank of England är drivande och sammanhållande för implementeringen av CBEST.

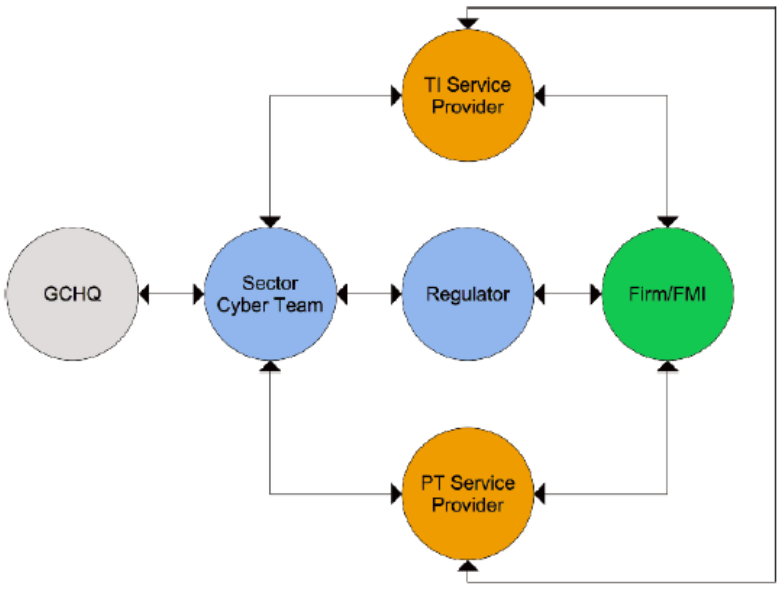
CBEST är ett ramverk för att leverera kontrollerade, skräddarsydda, cyber-säkerhetstester, som är baserad på relevanta och aktuell scenarioinformation. Testerna replikerar riktiga hotaktörers metoder och är därigenom baserade på verkliga hot mot systemviktiga finansiella institut. För CBEST har man i Storbritannien identifierat ett 30-tal kritiska aktörer som har genomfört testet. Resultatet har i stor utsträckning hållits inom Bank of England, även om informationsdelning sker med exempelvis Financial Conduct Authority (FCA).

CBEST använder särskilda indikatorer för att mäta förmåga och mognad och ger på vis möjlighet till benchmark för sektorn och tillsynsmyndigheterna. En tydlig benchmark möjliggör i sin tur att sektorns och dess aktörer kan stärka förmågan att skydda sig mot

[27] FSOR (2017), *Årsberetning 2017*.

cyberattacker och att upptäcka, besvara och återhämta sig från dem på ett lämpligt sätt.[28]

De deltagande aktörerna i ett CBEST-test illustreras i nedanstående figur. Förutom den finansiella aktör som testas (Firm/FMI[29]), så ingår tillsynsmyndigheter (Regulator), leverantörer av underrättelsetjänster (TI Service Provider) respektive av penetrationstest (PT Service Provider), Bank of Englands Sector Cyber Team och Government Communications Headquarters (GCHQ).



Genomförandet kan sammanfattas i fyra huvudsakliga faser: *Initieringsfas, Underrättelsefas, Penetrationstestfas, Avslutande fas*.[30] I initieringsfasen startas CBEST-testet formellt, omfattningen/avgränsning etableras och leverantörer av underrättelsetjänster och penetrationstester upphandlas. Endast leverantörer som har ackrediterats av ett särskilt råd kan användas.[31] I underrättelsefasen produceras den huvudsakliga underrättelseinformationen, såsom relevanta hot, angripares mål och metoder. I fasen utvecklas testets scenarier och ett utkast till penetrationstestplan. Kontroll över testplaneringen överlämnas till leverantören av penetrationstest. I penetrationstestfasen genomförs penetrationstest mot den testade organisationens utpekade system och tjänster stödjer kritiska funktioner. Den testade organisationens förmåga att detektera och respondera mot angreppen utvärderas. I den avslutande fasen utvecklas en testrapport som inkluderar bedömningar av organisationens

[28] CREST (2014), *An introduction to CBEST*.

[29] FMI, Financial Market Infrastructure.

[30] *Initiation Phase, Threat Intelligence Phase, Penetration Testing Phase, Closure Phase*. Se BoE (2016), *CBEST Implementation Guide*.

[31] *Council for Registered Ethical Security Testers (CREST)*.

förmågor. I fasen utvecklas även en åtgärdsplan, vars implementering övervakas av tillsynsmyndigheterna.[32]

2.2 TIBER-NL (NL)

Det holländska ramverket drivs sedan 2016 av De Nederlandsche Bank och benämns Threat Intelligence-based Ethical Red teaming (TIBER). TIBER inriktar sig, liksom CBEST, i första hand på kritiska aktörer i den finansiella sektorn (Financial Core Infrastructure, FCI). På ett liknande sätt, så anlitas även externa leverantörer för kontrollerade attacker mot faktiska system hos den testade organisationen. Testscenarier baseras på kommersiellt tillgänglig underrättelseinformation om verkliga hot och angriparens metoder, som vid behov kompletteras med myndighetsinformation om hot och angripares metoder.

Ett TIBER-test kan beskrivas som ett sofistikerat red teaming-test med samma taktik, tekniker och förfaranden (TTP) som verkliga motståndare, mot kritisk produktionsinfrastruktur och utan förhandskännedom för organisationens försvarande Blue Team (BT). BT är således inte medvetet om TIBER-NL-testet.

Vid testet, så utvärderas befintliga förmågor, kontroller, förebyggande åtgärder och säkerhetsdetektering och responskapacitet mot avancerade angrepp. Testet leder även till att eventuella svagheter, fel eller andra säkerhetsproblem identifieras på ett kontrollerat sätt. Testfasen efterföljs av ett fullständigt avslöjande och genomgång av testet. Under testet kommer ett White team, bestående av säkerhets- och affärsexperter från den övade organisationen, att övervaka testet och ingripa vid behov. Ingripanden kan exempelvis vara nödvändiga när testet riskerar att leda till kritisk påverkan, till en nivå som har fastställt i förväg.[33]

Liksom för CBEST, så beskrivs TIBER-NL översiktligt i en särskild implementeringsguide – *How to conduct the TIBER-NL test*, som sammanfattar testets målgrupper, faser, aktiviteter och önskade resultat. De centrala aktörerna som är involverade i testet är den finansiella aktör som övas (inklusive ett Blue Team och ett White Team), ett TIBER-NL Cyber Sector Team (TCST) från den holländska centralbanken, samt leverantörer av Red Team och underrättelsetjänster.

Genomförandet kan sammanfattas i fyra huvudsakliga faser: *Generisk scenarioinformation om hot, förberedelsefas, testfas* och *avslutande fas*.[34] Generisk scenarioinformation tillhandahålls från myndigheter och inkluderar information om de mest avancerade hotaktörerna som är relevanta för de nederländska finansinstituten, samt information om dessa aktörers mål. I förberedelsefasen påbörjas TIBER-NL-testet formellt, White Team etableras, omfattning/avgränsning fastställs, och en Red Team-

---

[32] BoE (2016), *CBEST Implementation Guide*.

[33] TIBER-NL (2017), *How to conduct the TIBER-NL test*.

[34] Generic Threat Intelligence, Preparation Phase, Test Phase, Closure Phase

leverantör upphandlas. Eventuellt upphandlas även en leverantör av underrättelsetjänster. I testfasen fördjupas och berikas den generiska underrättelseinformationen till mer fullödiga scenarier och en testplan. Därefter genomförs ett redteam-test mot utpekade mål, såsom system och tjänster som stödjer kritiska funktioner. I den avslutande fasen genomförs en genomgång av de genomförda scenarierna mellan Blue Team och Red Team. I fasen utvecklas den testade organisationsens åtgärdsplan. Samverkan och erfarenhetsdelning sker med såväl andra branschkollegor som med tillsynsaktörer.[35]

2.3 *iCAST (Hong Kong)*

iCAST (intelligence-led cyber attack simulation testing) är en del av ett initiativ från Hong Kong Monetary Authority (HKMA). Initiativet har sin bakgrund i ett behov av att höja nivån på penetrationstester för finansiella aktörer, bland annat med mer relevant Threat Intelligence (t.ex. riktade attacker mot sårbara punkter såsom ”people” och ”processes”). iCAST tillhandahåller även KPI:er som hjälper till att benchmarka deltagares förmåga att upptäcka och svara på attackerna i testet.

I implementationsguiden till iCAST, så framhålls att traditionella penetrationstest har gett stor nytta genom att tekniska sårbarheter identifiera, dock ofta inom ett enda system eller en isolerad miljö. Samtidigt understryks att bredden av relevanta och sannolika scenarier för en riktade attack mot en finansiell aktör (inklusive människor och processer samt angriparens tekniker och metoder) ofta inte helt omfattas av traditionella penetrationstester. Därutöver möjliggör inte traditionella penetrationstester lika hög grad av utvärdering av förmåga att hantera cyberangrepp och ger heller inte konkreta mätvärden eller indikatorer (KPI) för att mäta effektiviteten av den egna organisationens cybersäkerhetsprogram.

Genomförandefaser i iCAST är *avgränsning*, *utveckling av Threat Intelligence*, *utveckling av testscenarios*, *test*, samt *rapportering*. Innan iCAST initieras, så upprättar den testade organisationen en kontrollgrupp bestående av personal på senior nivå. Andra kollegor ska inte informeras om testet. I avgränsningsfasen, så identifieras vilka verksamhets- och stödfunktioner som testet bör fokusera på. Därefter fastställs relevanta kritiska system och processer och typer av hot-händelser för testet formuleras. I ett nästa steg analyseras relevant underrättelseinformation och sammanställs i en särskild rapport, antingen med generisk underrättelseinformation om verkliga hot eller sådan som är anpassad till den testade organisationen. I nästa fas utvecklas med fullödiga scenarier, i form av redogörelser för mål, metoder, tidpunkter m.m. avseende respektive angrepp. Dessa scenarier används sedermera i testfasen. Avslutningsvis sammanställs ett antal rapporter, från såväl den egna organisationen som från tjänsteleverantörer av underrättelseinformation och redteam-test.[36] En del av ramverket från HKMA är även

---

[35] TIBER-NL (2017), *How to conduct the TIBER-NL test.*

[36] MKMA (2016), *CYBER RESILIENCE ASSESSMENT FRAMEWORK, CONSULTATION DRAFT.*

utvecklingen av testare inom ramen för ett utbildnings- och certifieringsprogram (Professional Development Programme, PDP).

2.4 TIBER-EU (EU)

Europeiska centralbanken (ECB) har utvecklat ett ramverk för underrättelsebaserade redteam-tester, som benämns TIBER-EU. Ramverket påminner i hög grad om tidigare och liknande koncept och bygger på införande av nationella TIBER inom EU. ECB är samtidigt sammanhållande för ramverket och kommer exempelvis att etablera ett så kallat TIBER-EU Knowledge Center (TKC), i syfte att stöjda nationell och europeisk implementation samt för att bidra till samarbete och informationsdelning. Målgruppen för tester inom TIBER-EU beror på hur ramverket implementeras i de respektive medlemsländerna. Vilka finansiella aktörer som deltar beror därför på hur ländernas myndigheter väljer att tillämpa ramverket, även om målgruppen ska inkludera aktörer som är viktiga för landets finansiella stabilitet.

ECB:s ramverk ämnar åstadkomma en EU-övergripande standardisering och ömsesidigt erkännande av cybertester. Detta skulle bland annat medföra att finansiella aktörer kan undvika det dubbelarbete som följer av att man testas i/av varje EU-medlemsstat där aktören är verksam. Lösningen på detta är antingen ett gemensamt arbete med myndigheter från länder där den finansiella aktören är verksam, alternativt att en myndighet (lead authority) leder testet och att myndigheter från övriga länder erkänner testet som legitimt och gällande för sina länder, givet att TIBER-testet uppfyller formerna för TIBER-EU.

Många finansiella aktörer använder sig redan av interna redteams, men ECB framhåller samtidigt nyttan av externa redteams, som utgör en central del av TIBER-EU. Externa leverantörer möjliggör nya och oberoende perspektiv, vilket är svårt att uppnå om testen planeras och genomförs av personal från den egna organisationen. Därutöver kan externa aktörer bidra med mer resurser och uppdaterad kunskap inom området.

TIBER-EU kan beskrivas i form av tre faser: förberedelsefas, testfas, och avslutande fas.[37] Förberedelsefasen inkluderar upprättande av organisation, avgränsning av testet och upphandling av leverantörer av tjänster för underrättelseinformation och redteamtest. I testfasen utvecklas testets scenarier och själva testet genomförs. I den avslutande fasen sammanställs en testrapport av redteam-leverantören, som beskriver genomförandet av testet och dess resultat. Rapporten inkluderar även eventuella rekommendationer och nödvändiga åtgärder för organisationen som har testats, såsom tekniska kontroller, policies, rutiner, utbildning och medvetandehöjande insatser. Aktörerna som har deltagit i testet går gemensamt igenom testscenarierna och diskuterar problem, risker och utmaningar som har identifierats vid testet. Den testade

[37] Preparation phase, test phase, closure phase. Se ECB (2018), *TIBER-EU FRAMEWORK How to implement the European framework for Threat Intelligence-based Ethical Red Teaming*.

organisationen omsätter testets resultat och rekommendationer och sammanställer, i samråd med sin tillsynsmyndighet, en egen åtgärdsplan.

2.5 *Analys och jämförelse mellan ramverken*

Det kan konstateras att testkoncepten liknar varandra i väldigt stor utsträckning. I vägledningar till testerna, så använder många avsnitt till och med samma formuleringar. Med största sannolikhet har CBEST varit utgångspunkten för övriga beskrivningar, givet att det brittiska formatet etablerades först.

De olika testramverken beskriver planering, genomförande och uppföljning av testerna i termer av faser, aktiviteter och resultat. Även deltagare och informationsflöden illustreras och beskrivs i vägledningar och implementationsguider till testerna. Samtliga av testkoncepten bygger även på att de testade organisationerna anlitar externa leverantörer av Threat Intelligence och Penetrationstest, som uppfyller särskilda ackrediteringskrav. En generell utmaning som har nämnts inom ramen för FSPOS Fokusgrupp Cybersäkerhet är hur test-koncepten kan ge jämförbarhet, givet att olika scenarier och olika leverantörer kan användas vid testerna. Ackrediteringskrav och ett EU-gemensamt ramverk ger sannolikt en viss enhetlighet och förutsägbarhet. Trots detta, kan skillnader kvarstå mellan länder och leverantörer.

TIBER-EU och TIBER-NL är mest lika varandra, exempelvis uttryckt i steg och aktiviteter i processen för planering och genomförande. Ett exempel på hur dessa två skiljer sig från CBEST och iCAST är tillförandet inom TIBER-EU och TIBER-NL av generisk information om hot och verkliga angripares mål och metoder som ett inledande steg inom TIBER-EU och TIBER-NL. Inom CBEST och iCAST genomförs avgränsningen innan underrättelseinformation om angripares metoder inkluderas i arbetsprocessen.

Deltagandet från nationella myndigheter – i planering, genomförande och utvärdering av tester – är tydligare beskrivet och mer påtaglig inom de europeiska ramverken än inom iCAST. TIBER-EU förordar att deltagande länders relevanta myndigheter upprättar ett nationellt TIBER Cyber Team (TCT). I det holländska ramverket finns exempelvis ett TIBER-NL Cyber Sector Team (TCST) från De Nederlandsche Bank (DNB). Ett liknande Sector Cyber Team (SCT) från Bank of England finns inom ramen för CBEST.

En skillnad mellan ramverken är att CBEST är frivilligt medan iCAST är obligatoriskt för kritiska finansiella aktörer. I beskrivningen av TIBER-EU, så konstateras även att de landspecifika TIBER-programmen kan vara antingen frivilliga eller obligatoriska, beroende på vad som beslutas av respektive lands myndigheter.[38]

En annan skillnad mellan ramverken är att Hong-Kong-baserade iCAST utgör en delmängd i ett bredare utvärderingsramverk för cyberresiliens, där finansiella aktörer

[38] BIS (2017), *FSI Insights, on policy implementation No 2 Regulatory approaches to enhance banks' cyber-security frameworks*, samt ECB (2018), *TIBER-EU FRAMEWORK How to implement the European framework for Threat Intelligence-based Ethical Red Teaming*.

även genomför riskbedömningar och mognadsmätningar. Motsvarande delar finns inte inkluderade i övriga ramverk. En skillnad mellan TIBER-EU och övriga ramverk är ECBs tonvikt på enhetlighet inom EU-området och lösningar för ömsesidigt erkännande av cybertester mellan EU-länder.

CBEST har krav på att ackrediterade leverantörer bör utföra testerna, för att säkerställa tillräcklig tillförlitlighet och kompetens. Utvecklingsprogrammet inom HKMAs ramverk har en liknande målsättning, även om utförare av tester inte behöver komma från externa leverantörer utan även kan komma från den egna organisationen. Där medges exempelvis andra experter, så länge de har motsvarande kompetens som de från HKMAs Professional Development Programme (PDP).[39] Ett motsvarande ackrediterings- eller utvecklingsprogram finns inte beskrivet i implementationsguiden för TIBER-NL. Däremot anges i beskrivningen av TIBER-EU att externa leverantörer av tester och underrättelsetjänster ska ha genomgått en certifierings- och ackrediteringsprocess. På ett liknande sätt som för CBEST, så beskrivs kraven i ett särskilt dokument, *TIBER-EU Services Procurement Guidelines*.[40]

[39] HKMA (2016), *Cybersecurity Fortification Initiative*.

[40] ECB (2018), *TIBER-EU FRAMEWORK How to implement the European framework for Threat Intelligence-based Ethical Red Teaming*.

Avslutande reflektioner

Test och övning är ett kraftfullt verktyg för att stärka organisationens förmåga att stå emot och hantera cyberangrepp. Olika testformer är en naturlig och nödvändig del av finansiella organisationers ramverk för cyberresiliens. De olika testformerna har olika förutsättningar, styrkor och utmaningar. Ofta innebär detta att flera sorts tester, såväl automatiserade skanningar som mer djupgående och avancerade tester, kompletterar varandra och behöver vara en del av organisationens testprogram. I takt med att kraven på finansiella aktörers cyberresiliens ökar, utgör s.k. Threat Intelligence-based Ethical Red Teaming ett framväxande och alltmer uppmärksammat format.

I och med utvecklingen av fler initiativ inom området, inte minst genom ECBs TIBER-EU, är det sannolikt att svenska finansiella aktörer kommer att beröras och delta i tester inom en snar framtid. Lärdomar från dessa tester kan med fördel även överföras till andra delar av den finansiella sektorn, exempelvis inom samverkan i FSPOS.