Running DRP Tabletop Exercises

Josh Simon *(he/him, jss@umich.edu),*
Senior Systems Administrator
University of Michigan
College of Literature, Science, & the Arts (LSA)

***SREcon25 Americas** / March 26, 2025*

LSA TECHNOLOGY SERVICES
UNIVERSITY OF MICHIGAN

Agenda

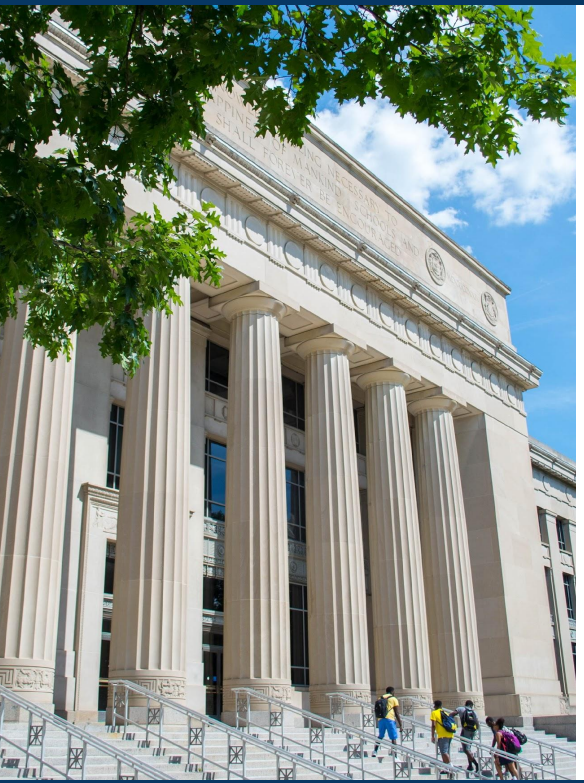
1. Overview
2. Problems
3. DRP contents
4. Tabletop exercises
5. Measuring success
6. Next steps
7. Additional considerations

Overview

What is our environment like?

Who we are

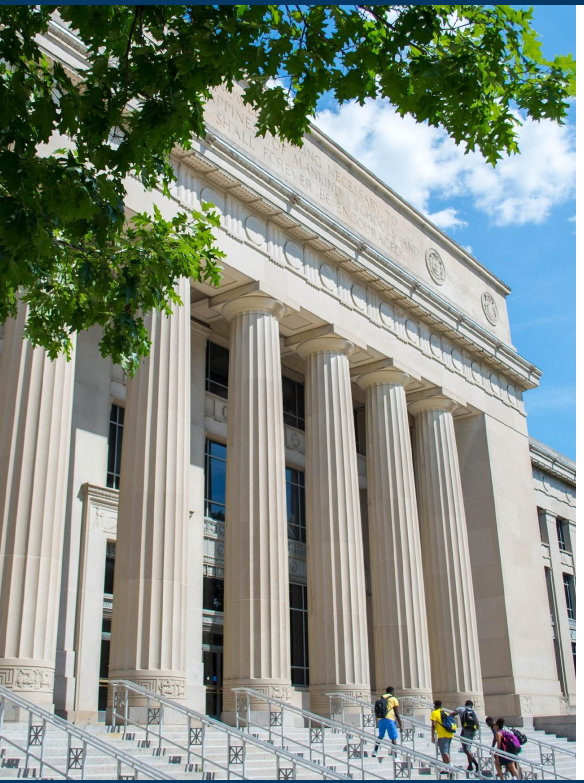
The university and the college:



- **The University of Michigan (U-M)**, founded in 1817, is the oldest institution of higher education in the state, with:
  - Over 52,800 enrolled students
  - Around 9,200 faculty and 47,900 staff
  - A main campus in Ann Arbor with 19 schools & colleges
  - Two regional campuses in Dearborn and Flint
  - A hospital system with many statewide general and specialty clinics
- **The College of Literature, Science, and the Arts (LSA)** is the largest and most academically diverse of UM-Ann Arbor's 19 schools and colleges with approximately 22,000 students and 3,600 faculty and staff.

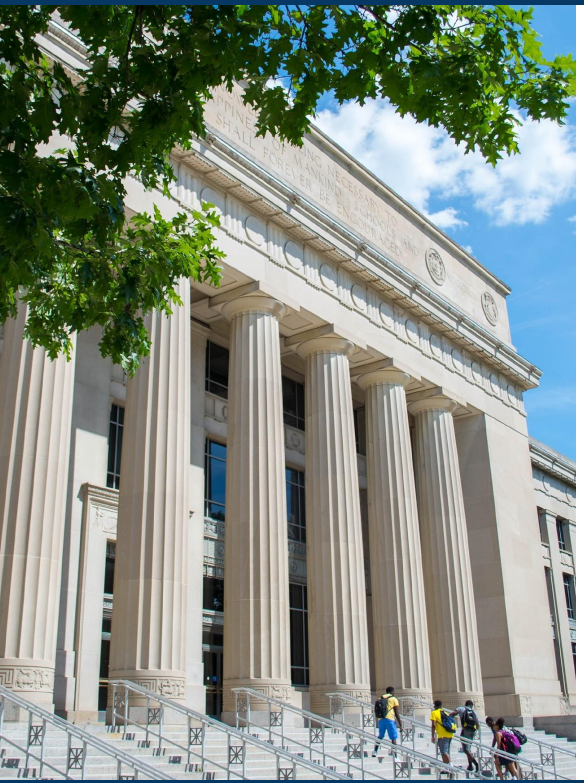
Who we are

The unit and the team:



- **LSA Technology Services (LSA TS)** had about 180 full-time staff to manage all the administrative and classroom technology in nearly 30 central campus buildings and several remote facilities.
- The LSA TS **Infrastructure & Security** team had 13 staff and two managers who provide or coordinate over 70 foundational IT services for the college and the university.

Innovation Day and tabletop exercises:



- The Infrastructure & Security team holds a monthly “Innovation Day” where we can focus on a specific project or service **outside** of our regular daily operations.
- We ran tabletop exercises of some of our disaster recovery plans (DRP) on our Innovation Days in February 2024, March 2024, February 2025, and March 2025 (last Friday!).
- This presentation is about our 2024 exercises.

Problems

What were some of the problems we identified?

Problems

We identified some problems with our existing plans:

- We've had disaster recovery plans for at least our major customer-, patron-, or user-facing services since at least 2008, but:
  - Most only assumed natural disasters such as earthquake or tornado.
  - Only some (but not all) assumed service, server, or component failure.
  - Few specified recovery steps beyond "Rebuild from scratch" and "Restore from tape."
  - Most services didn't have a current build guide to use as a basis for rebuilding.

Problems

We identified some problems with our existing plans (continued):

- Not all of our existing production services had DRPs.
- When spinning up a new service, the plan author would often make a copy of someone else's plan and edit it… even if that plan was missing one or more sections from the original template.
- In 2022 we standardized our DRP template and included embedded guidance about using it.
- We tested each disaster recovery plan at least annually, typically during one of our mid-semester maintenance weekends (in Feb, Jul, and Oct).

Problems

We identified some additional problems with our process:

- Testing could be performed differently depending on the plan owner:
  - Some would perform a (solo) thought experiment.
  - Some would build a new VM, stand up the service using its build guide, restore from backups, and compare it to the production VM.
- Plan authors often worked in isolation and could have different assumptions than other authors.
- The original author of the plan was likely the service owner, service manager, or technical lead when it was first written… but what about staff turnover and service reassignments?

Level setting

We wanted all plan owners to:

- Operate with a common set of assumptions.
- Identify all service dependencies up and down the stack.
- Write and test their plans the same way.

DRP contents

What goes into a disaster recovery plan (DRP)?

DRP contents

Our DRP template contains these sections:

- **Description** — A high-level overview of the service itself
- **Dependencies** — A list of specific dependencies on which this service depends and a list of which other services depends on this service
- **Hardware** — A list of physical and virtual hardware for the service:
  - Host names
  - IP addresses
  - Model and serial number (physical) –or– hypervisor information (virtual)
  - CPU count
  - Memory
  - Disk
  - Location (e.g., data center, row, rack, and position)
  - ITAM inventory assetID

DRP contents

Our DRP template contains these sections (continued):

- **Software** — A list of any specific software for the service (with specific version numbers if applicable)
- **Configuration** — Any specific configuration information, such as service account names or database user accounts
- **Security and facilities** — Details about the physical security for the hardware and the online security of the servers, appliances, databases, and so on
- **Backups** — Details about how the service components (such as databases or servers) are backed up

DRP contents

Our DRP template contains these sections (continued):

- **Operations** — Information about the service operations:
  - **Annual testing** — When and how the DRP will be tested
  - **Minimal operational standard** — The minimal operating condition for the service, such as “read only” or “single server”
  - **Fully operational standard** — The fully operational condition for the service, such as “read-write” and “redundant servers”
  - **Documentation** — The location of operational documentation like architecture diagrams, build guide, run books, administrators guide, users guide, and so on
  - **Scheduled tasks** — A list of what runs when as whom

DRP contents

Our DRP template contains these sections (continued):

- **Contacts** — Contact information for:
  - **Service management** — The service owner, service manager, and technical leads for the service
  - **Customers** — Any service-specific customer email lists for notification purposes
  - **Partners** — Any partner organizations who may provide dependencies or be dependent (such as Facilities or Plant Operations)
  - **Vendors** — Any hardware, software, or service vendors involved in providing the service or its components

DRP contents

Our DRP template contains these sections (continued):

- **Service levels and impact** — Information about the impact of the service not being available, including but not limited to:
  - Service Level Agreement (SLA)
  - Service Level Expectations (SLE)
  - Operating Level Agreement (OLA)
  - Statement of impact
  - Severity of impact
  - Impact timelines (for the university, college, department, and project)
  - Maximum acceptable downtime

Tabletop exercises

We used two consecutive Innovation Days to run tabletop exercises to level-set our assumptions, dependencies, and expectations.

Tabletop exercises

The first four of the five W's:

- **Who** — Our entire team of 15 (playing different roles)
- **What** — A guided tabletop exercise of [some of] our disaster recovery plans
- **Where** — A big conference room
- **When** — Our February and March 2024 Innovation Days

Tabletop exercises

The fifth W:

- **Why:**
  - Improve our understanding of how we'd actually implement our DRPs.
  - Consider what we may need to add to, change in, or remove from our environment (both college-wide and service-specific).
  - Identify and fill any gaps in our DRPs.
  - Identify any blind spots to remediate before our next disaster.

Tabletop exercises

Roles

There are generally four roles in **formal** DRP tabletop exercises:

- **Facilitators** provide situation updates and moderate discussions. They also provide additional information or resolve questions as required. They may also assist with facilitation as subject matter experts (SMEs) during the exercise.
- **Players** are personnel who have an active role in discussing or performing their regular roles and responsibilities during the exercise. Players discuss or initiate actions in response to the simulated emergency. They respond to the situation presented based on current plans, policies, and procedures. *(Service owners, service managers, and technical leads)*

Roles (continued)

- **Observers** do not directly participate in the exercise; however they may support the development of player responses to the situation during the discussion by asking relevant questions or providing subject matter expertise. *(Management and 1–2 volunteers who aren't players)*
- **Data collectors** observe and record the discussions during the exercise, participate in data analysis, and assist with drafting the After-Action Report (AAR). *(2–3 volunteers who aren't players)*

Tabletop exercises

Process: What we did

- We **informally** walked through some guided scenarios using three selected services' DRPs:
  - Our VM hypervisor infrastructure
  - Our web hosting environment
  - Our on-campus data center
- The service manager and technical leads played as their own roles. Others played as other parties, such as our own management, the Dean's Office, central IT's Identity and Access Management (IAM) team, the network team, and so on.

Tabletop exercises

Process: What we did (continued)

- We tried to get everyone to participate as a player at least once.
- We asked 3–4 different people to act as **observers** and **data collectors** in each exercise.

Process: What we didn't do

- We did **not** hide the facilitator behind a scrim.
- We did **not** use dice to determine actions' success or failure.
- We did **not** draft a formal AAR document.
- We did **not** post to any Slack channels or social media.
- We did **not** actually notify our customers, patrons, or users, by email or otherwise.

Exercise one summary

Our first exercise was for our own VM hypervisor environment:

- We posited an attacker had gotten into the hypervisor software itself on one node in the cluster.
- Subsequent discussion implied they had access to all nodes in the cluster as well as the underlying storage… which included our (sometimes unencrypted) disk-level backups.
- If they had been sufficiently patient to wait for our server backups to expire, we would have been unable to rely on any of those.

Exercise one summary (continued)

End result: Were this a real event, we would have to rebuild our entire virtualization environment from scratch, and then rebuild every server and service that relied on it from scratch, all without trusted backups.

Exercise two summary

Our second exercise was for our web hosting environment:

- We posited an attacker had gained access to an individual web server which hosted 136 websites.
- Subsequent discussion implied all 136 sites could have been compromised.
- **Mitigation:** Because we use host-unique root passwords and disallow ssh as root, knowledge of this server's password doesn't grant access to any other servers.

Exercise two summary (continued)

End result: Were this a real event, we would need to rebuild the server, reinstall the hosting environment software, restore all websites from backups both made before the attack and stored on a different server (where possible), regenerate all SSL keys and certificates, reset every website's administrative passwords, and request site owners to reset their software's administrative passwords.

Exercise three summary

Our third exercise was for our on-campus data center:

- We posited an attacker had gained physical access to that data center.
- **Mitigation:** We now have a lot fewer physical servers and services either in or dependent on the data center (which predates the existence of the south-of-campus data center and both our and central IT's virtualization services).

Exercise three summary (continued)

- We discussed several possibilities:
  - Did they **steal** something (one or more servers, disks, and so on)?
  - Did they **damage** something (such as servers, racks, cables, or HVAC)?
  - Did they damage something to **cover up** that they stole something?
  - Did they attempt to **remove** UPS batteries while hot?
- Subsequent discussion identified some unique constraints and concerns. For example, we are contractually obligated to have a specific server and its data in a specific locked rack in that data center.

Exercise three summary (continued)

- We have additional areas to look into as takeaways:
  - Are all servers' disks encrypted at rest?
  - What *really* happens when the EPO button is pressed?

End result: Were this a real event, we know (and have documented) what servers and services are in the data center and the priority order for bringing them back up. We also know (and have documented) who to work with in LSA Facilities, the Office of the Vice President and General Counsel (OGC), Plant Operations, and Risk Management, depending on the nature of the event.

Measuring success

How did we define and measure whether these exercises were successful?

Success criteria

How did we know if we were successful?

- Did we have an improved understanding about how we'd implement our DRPs in reality?
- Did we consider both college-wide and service-specific what to:
  - Add to our environment?
  - Change in our environment?
  - Remove from our environment?
- Did we identify any more gaps in our plans? If so, did we have or make plans to fill those gaps?
- Did we identify any more blind spots to remediate before our next disaster?

Post-event surveys

We asked our participants!

- We sent out post-event surveys after both days that asked:
  - How well did today go?
  - What did you like most or find most helpful?
  - What did you like least or find least helpful?
  - What could have gone better?
- The second time, we also asked if the first time was better, the second time was better, or if they were about the same.

First survey

February 2024 survey results

In the first survey we asked:

- **How well it went** — On a 5-point scale (1=awful to 5=great) we averaged 4.30 (n=10 (71.43%), min=3, max=5).
- **What went well** — It was engaging. We learned a lot about the DRPs and team members' thought processes and priorities. Some felt the roundtable discussion was helpful. It was a useful brainstorming session for our virtualization service, even for people who don't work on it. People identified holes in their own plans, new things to think about, perspectives to change, and a deeper understanding of the service itself.

First survey

February 2024 survey results (continued)

- **What could have gone better**:
  - Some thought that we got lost in the weeds a bit.
  - Some thought that the problem wasn't as clearly defined as would have been helpful.
  - Some thought that including our Major Incident (MI) process, which is mostly about communications and coordination, was an unnecessary complication.

We identified areas for improvement that we implemented for the second pass.

Favorite responses

What were some of our favorite text responses to the first survey?

- **What went well** — “The format was engaging, evaluating next steps given limited information helped contextualize the DRPs, and it was cool to have everyone’s expertise in the room for additional details on impact.”
- **What needed improvement** — “It’s easy to go down some pretty deep rabbit holes; there’s a balance here, because those rabbit holes sometimes dig up value, but sometimes they don’t.”
- **Other comments** — “This exercise was terrifying.”

Second survey

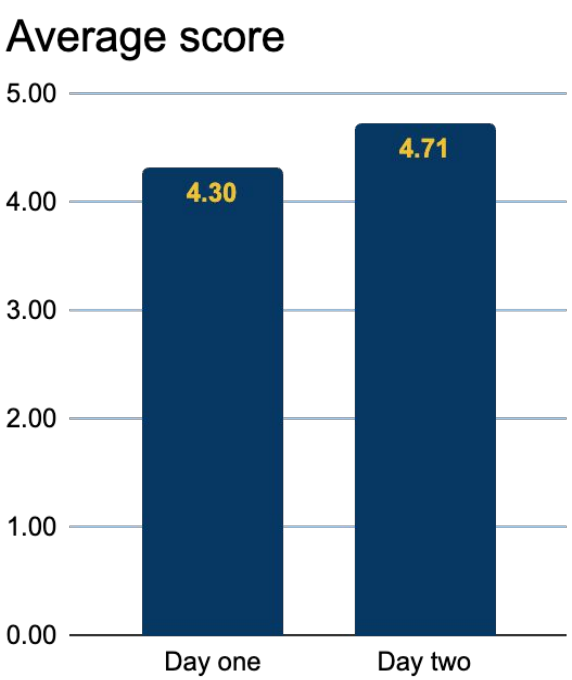
March 2024 survey results

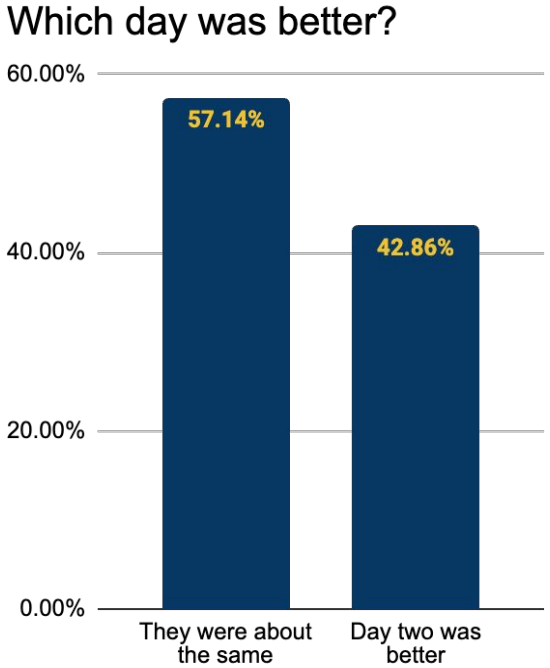
The second time through:

- **How well it went** — On the same 5-point scale we averaged 4.71 (n=7 (63.64%), min=4, max=5).
- **Which was better** — Everyone said the second session was either as good as (57.1%) or better than (42.9%) the first session.

Second survey

March 2024 survey results (continued)





Second survey

March 2024 survey results (continued)

- **What went well** — The more-focused topics were easier and the open discussion was still helpful. We focused more on technical actions than communication actions.
- **What could have gone better** — We still spent too much time going down rabbit holes. The scope may have been too narrow for everyone to feel like they were contributing.

We identified more areas for improvement that we planned to implement next time.

Next steps

What did we need to do after the exercises?

For participants

Participants were asked to:

- Update their DRPs again now that we:
  - Have an improved understanding about their implementation.
  - Considered college-wide and service-specific additions, changes, and deletions.
  - Identified our common dependencies.
  - Identified any gaps and blind spots.
  - Identified restoration time.
- Create or update any architecture, build, or design documents to assist in service recovery.

Common dependencies

What are those common dependencies?

We identified and documented many common dependencies, some of which were originally considered “too obvious to list:”

- Power and cooling infrastructure (the data center environment/s)
- Networking services (such as DNS, firewalls, load balancers, and the university-provided core network)
- Centrally-provided university-wide backup, database, server, and storage services, and other college- or department-local equivalents

Common dependencies

What are those common dependencies (continued)?

- Centrally-provided university-wide authentication services:
  - Active Directory, Kerberos, and LDAP
  - Two-factor authentication (2FA)
  - Shibboleth or other federated authentication
  - Password management solutions
- Centrally-provided university-wide time services (NTP) on which the authentication services depend
- Notification channels (such as email, chat, and social media)
- Documentation storage (such as an external service provider or your Knowledge Base or wiki)

For participants

Participants were also asked to:

- Compare actual restoration time to maximum acceptable downtime.
- Implement any remaining takeaways from the group discussions, including:
  - Add any missing software licenses to the appropriate inventory management system.
  - Define and implement any virtual machine groups, affinity rules, and anti-affinity rules.
  - Ensure all data is encrypted at rest.
  - Update what we're monitoring.
- Join the university-wide Disaster Recovery & Business Continuity Community of Practice if they were so inclined.

For facilitators

What were facilitators' next steps?

- **Internally for the team:**
  - Analyze the survey results and feedback.
  - Discuss them at subsequent team meetings.
  - Decide on a cadence for repeating the exercises (for example, annually in February and March and for our major services first).
  - Develop a training module and documentation templates to train additional facilitators.
- **Externally for the university** — Reach out to the DR/BC COP leads about presenting our exercises and results.
- **Externally for the community** — Present this case study at SREcon25 Americas. (Oh, look!)

Additional considerations

What else should you consider when writing and testing your own plans?

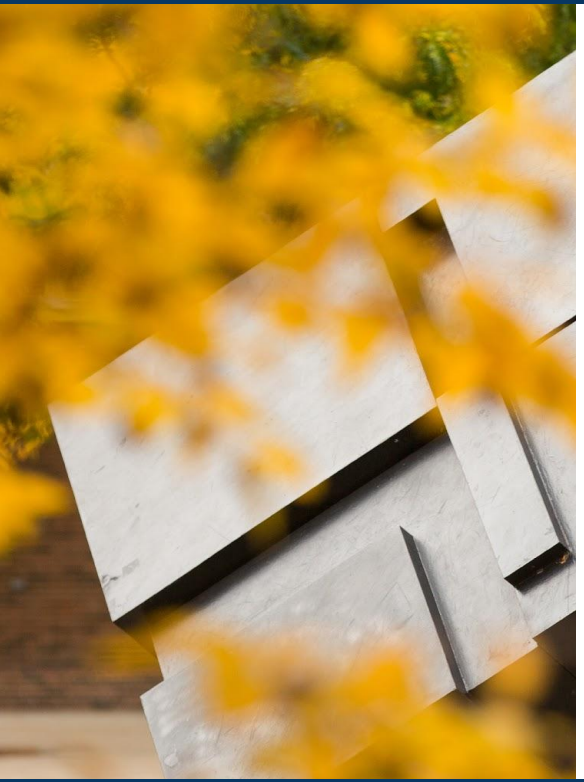
Considerations when writing

What should you consider when writing your plans?

- What if “Rebuild from zero without having usable backups” is where you end up?
- Does your service have a build guide or architecture diagrams?
    - Do you keep them up to date as the environment changes?
    - Do you have a change log to aid in getting from “built as documented” to “recovered”?
- A DRP can link to an external resource (such as a Knowledge Base article or vendor URL), but what if that resource is unavailable at implementation- or recovery-time?

Considerations when testing

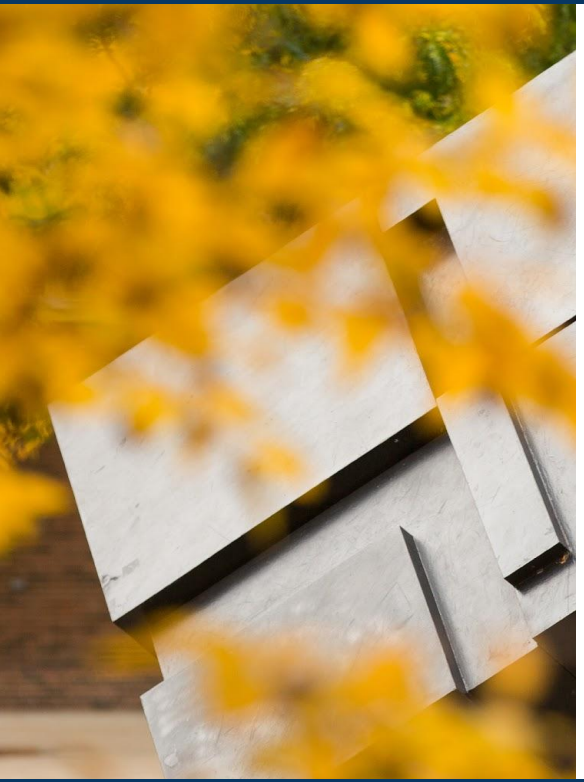
What should you consider when testing your plans?



- What criteria are you using to assess the success of their processes or the reliability of their backups?
- How in-depth are you going to test?
  - Is a thought experiment or tabletop exercise sufficient?
  - Are you restoring the database and file systems and comparing them to the development, test, staging, or production copies?
  - Are you rebuilding an entire new test version of the service?
  - Will you destroy anything you build? If so, when?

Considerations when testing

What should you consider when testing your plans (continued)?



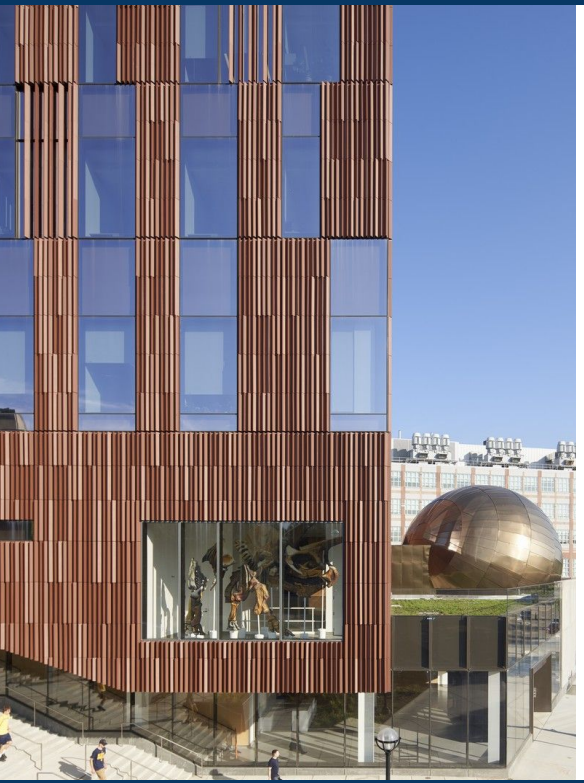
- Do you already have a test environment you can destroy and rebuild?
- Do you want to do a thought experiment annually and a full service rebuild every 3–5 years?

Resources

What resources might be helpful for writing and testing your own plans or for planning your own exercises?

Resources

We found these resources helpful:



- Our project management page for the tabletop exercises
- The “Disaster Recovery Management” page at the U-M “Safe Computing” website
- U.S. Cybersecurity and Infrastructure Security Agency (CISA) Tabletop Exercise Package (CTEP), especially:
  - Critical Infrastructure Tabletop Exercise Program (26-page PDF)
  - Exercise Planner Handbook (40-page PDF)
- Backdoors & Breaches, an incident response card game

Questions?

Thank you!

Running DRP Tabletop Exercises

Josh Simon *(he/him, jss@umich.edu),*
Senior Systems Administrator
University of Michigan
College of Literature, Science, & the Arts (LSA)

***SREcon25 Americas** / March 26, 2025*

LSA TECHNOLOGY SERVICES
UNIVERSITY OF MICHIGAN