ESnet
ENERGY SCIENCES NETWORK

ESnet Site Resilience Program

Resilience: the capacity to withstand or to recover quickly from difficulties!

Joe Metzger
Energy Sciences Network
Lawrence Berkeley National Laboratory

Internet2 Technology Exchange
September 21, 2023
Minneapolis, Minnesota

U.S. DEPARTMENT OF ENERGY
Office of Science

BERKELEY LAB

Challenge Question

Will you experience an outage next month?

What will be the impact?

ESnet

Challenge Question

What is the basis of your answer?

- A hopeful wish?
- An educated guess?
- A robust risk based systems analysis?

Are you comfortable with your answer?

Would your stakeholders be comfortable with your answer?

ESnet

Overview

1. ESnet Site Resilience Program (SRP)
2. Resilience Requirements
3. Estimating Nines

ESnet

ESnet Site Resilience Program - SRP

- SRP is a "Program" within ESnet to bring structure and a framework to our efforts to improve network service quality for our user community.
- SRP is trying to encourage collaborative risk management to feed into appropriate investments to best support the DOE mission.
- SRP is concerned about end-to-end network quality, with a focus on the impact of connections between ESnet and our users networks.

SRP is not just an ESnet program, it is a shared journey to improve the way we think about, talk about, and manage network risks!

ESnet

SRP Approach:
Communication, Analysis, Continuous Improvement!

- Improve how our community is communicating about resilience.
  - Estimating, documenting, and sharing the probability and impactics of network disruption.
  - Provide a framework to help define and quantify network dependencies, and expectations between organizations
- Analysis
  - Identify reasonable expectations for different design patterns.
  - Identify situations where expectations are not aligned.
- Continuous Improvement
  - Measure what is important
  - Identify, quantify, and evaluate potential improvements.
  - Focus resources (from all involved organizations) on the important problems.

ESnet

SRP Vision

- Our users **understand the risks** and benefits of different network connection resilience models.
  - They know what level of availability they need.
  - They know what level of availability they should expect from their current network connection model.
  - They are confident the current services meet their reliability requirements, or they know specifics about what can be done to improve their network resilience.
- Reduction in Single Points of Failure
  - Most network components should be able to be serviced (or fail) during normal business hours without adversely impacting users.

ESnet

Modeling is a key component of SRP

- What is the current design pattern for connections to each site?
- What level of availability should a particular design pattern provide?
- What should the availability target be for each site?
- What would it take to implement the target availability for each/all sites?
  - Who would need to make changes?
  - What would it cost?
  - Who could/should/would fund it?
- How do you quantify the benefits of a change?
  - What is the reduction in probability of an outage?
  - What is the dollar value of an outage?
  - What is the relationship between the groups reaping the benefits and the source of funds?

ESnet

SRP Milestones

FY22

- ✓ Select 3 resilience projects to start in FY23 (INL, ORNL, Y12)
- ✓ Develop & start executing SRP communications plan
- ✓ Create internal dashboard for tracking resilience projects
- ✓ Develop & refine SRP budget strategy
- ✓ Create a report of the Connection Model implemented at each site

FY23

- ✓ Select at least 4 resilience projects to start in FY24
- ✓ Automate creation of Connection Model report
- ✓ Continue executing SRP communications plan
- ✓ Continue refining SRP budget strategy

FY24 & Future Years

- Select at least 4 resilience projects to start next year.
- Add resilience information to the ESnet portal
- Continue executing SRP communications plan
- Continue refining SRP budget strategy

3-4 resilience projects are started each year based solely on importance and impact.
Other resilience projects triggered by unique opportunities, or urgent requirements will also be worked in parallel.

ESnet

End-to-End resilience planning is a shared activity!

- Each of us is responsible for different pieces of the puzzle.
  - The resiliency in each domain is different.
  - What really matters to the users is the end-to-end paths crossing multiple domains.
  - We need to work together to leverage each other's strengths, and compensate for each other's weaknesses if possible, not compound them.
- One of the SRP goals is to improve how our whole community is communicating about resilience.
  - That's why I am talking to you today!

ESnet

Mental Frameworks I have been using for thinking about Site Connection Resilience

1. Risk Management Framework
   - Identify, define and quantify **Risks**
   - Identify, define and quantify potential **Mitigations**
   - Do the math and implement the mitigations that make sense…
2. Quality Improvement Cycle
   - Plan, Do, Check, Act
3. Infrastructure Investment Framework
   - What network resilience improvements will allow us to collectively provide a **better value** to our community helping them advance their mission?

Are these just different aspects of the same problem?

No matter which mental model we use, we need to quantify the same things!

ESnet

1. Risks Reminder

Risks are expressed as:

If <risk event happens> ,

then <consequence will occur>

which will <impact an objective>.

Example:

If all circuits to a site go down for 30 minutes

then there will be a network service outage for 35 minutes

which will prevent users at the site from doing their work for 40 minutes.

ESnet

1. Managing Risks

How do we quantify the **probability** that risk will occur?

- MTBF
- Cut Rate Per KM for fiber or circuits
- Other?

How do we identify and quantify the **consequences**?

- Hard failures?
- Soft failures?

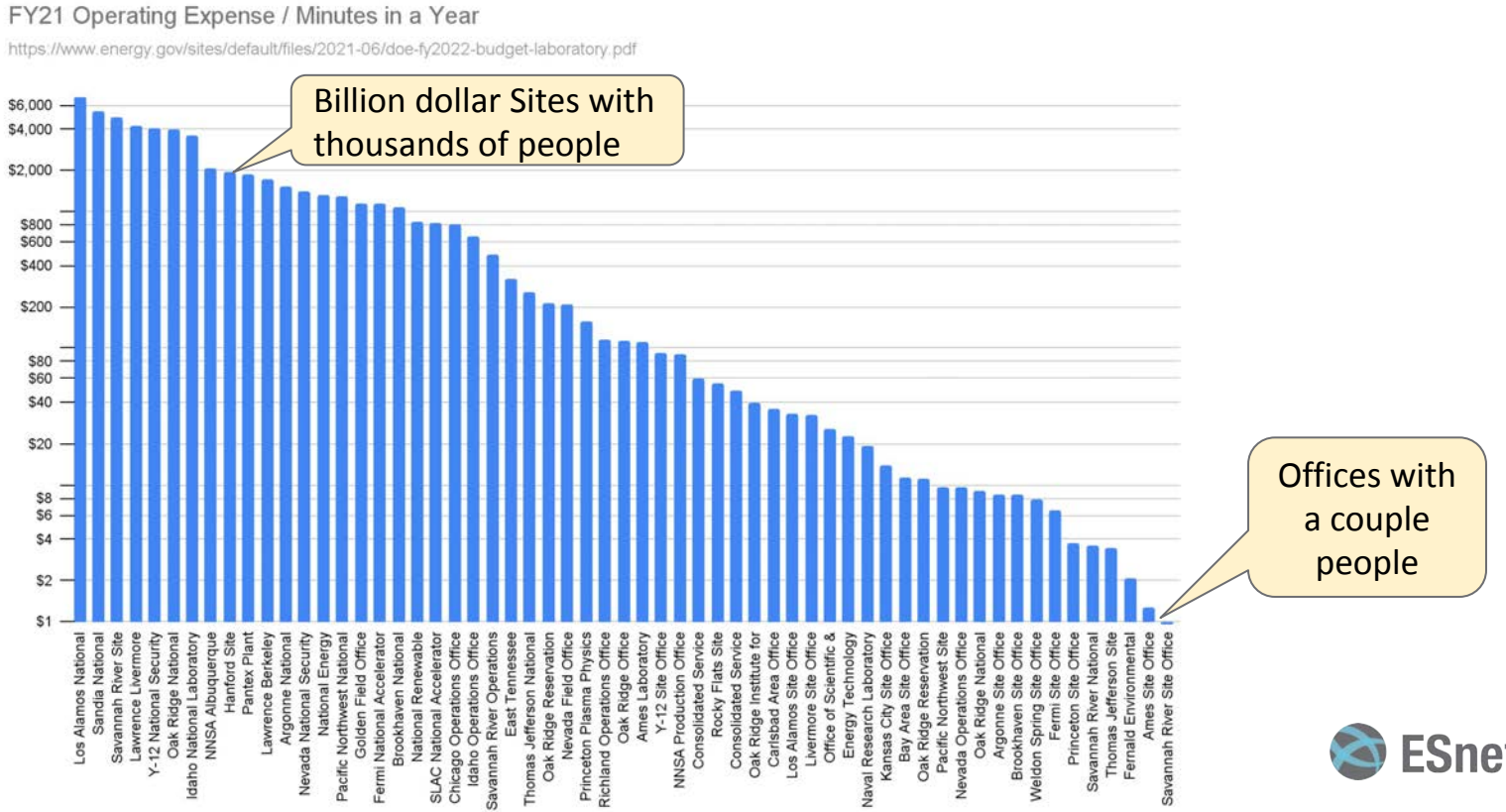
How do we quantify the **impact**?

- What is the value lost?
- Cost per minute of outage?

ESnet

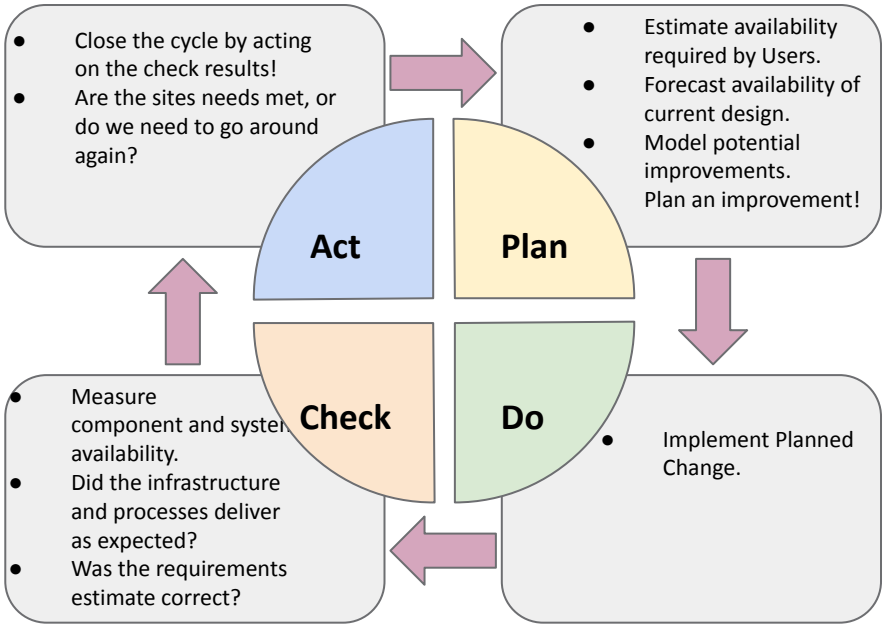
Quick & Dirty Impact estimation exercise!
Is operating cost per minute a useful measure of impact?

Notice Log Scale



ESnet

2. Availability is a measure of Quality.
So, we can use a Quality Improvement Cycle!

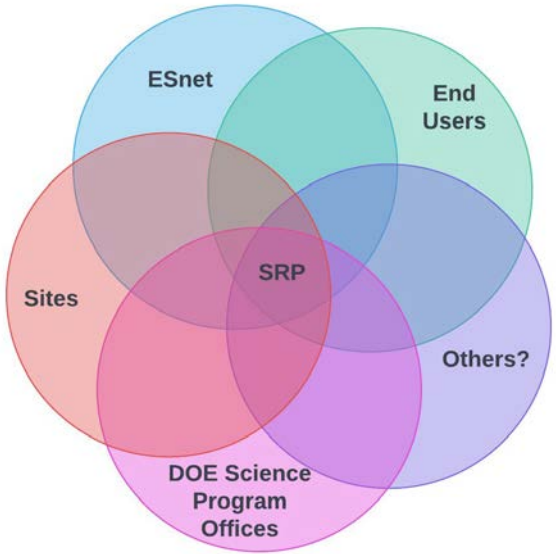


3. Investment Value = Cost / Benefit Ratio of Resilience Investments

Benefits - the Value of:

- Work not missed due to downtime
- Sustained or Improved Reputation
- Enabling activities not possible in less reliable and deterministic environments.
- Etc.

Stakeholders



Costs - NRC & ARC for:

- Equipment
- Services
- People's Time
- Etc.

ESnet

ESnet
ENERGY SCIENCES NETWORK

Resilience Requirements

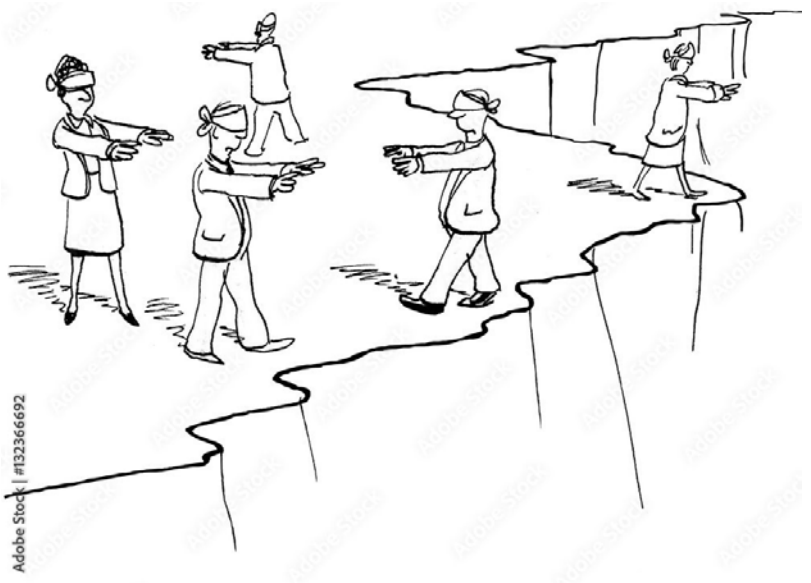
U.S. DEPARTMENT OF ENERGY
Office of Science

BERKELEY LAB

Communication about Availability Requirements

An important component of the SRP is improving our communities ability to communicate about resilience and availability.

- How much availability does a site need?
  - How do we figure this out?
- How should we describe and measure it?



ESnet

30 seconds on FISMA

The Federal Information Security Modernization Act of 2014 (FISMA) requires federal agencies to report the status of their information security programs to OMB and requires Inspectors General (IG) to conduct annual independent assessments of those programs.

Many of the ESnet sites *get to* participate in a FISMA compliance program where they rate their compliance on objectives around Confidentiality, Integrity, and Availability as "**low**", "**moderate**", or "**high**".

ESnet is currently FISMA Low/Low/Low.

Is it possible that an organization with FISMA moderate requirements relies solely on a FISMA low availability network for services, and that is OK? **Yes**, in some situations…

Is this the best methodology for characterizing and communicating network availability requirements? **Definitely Not!**

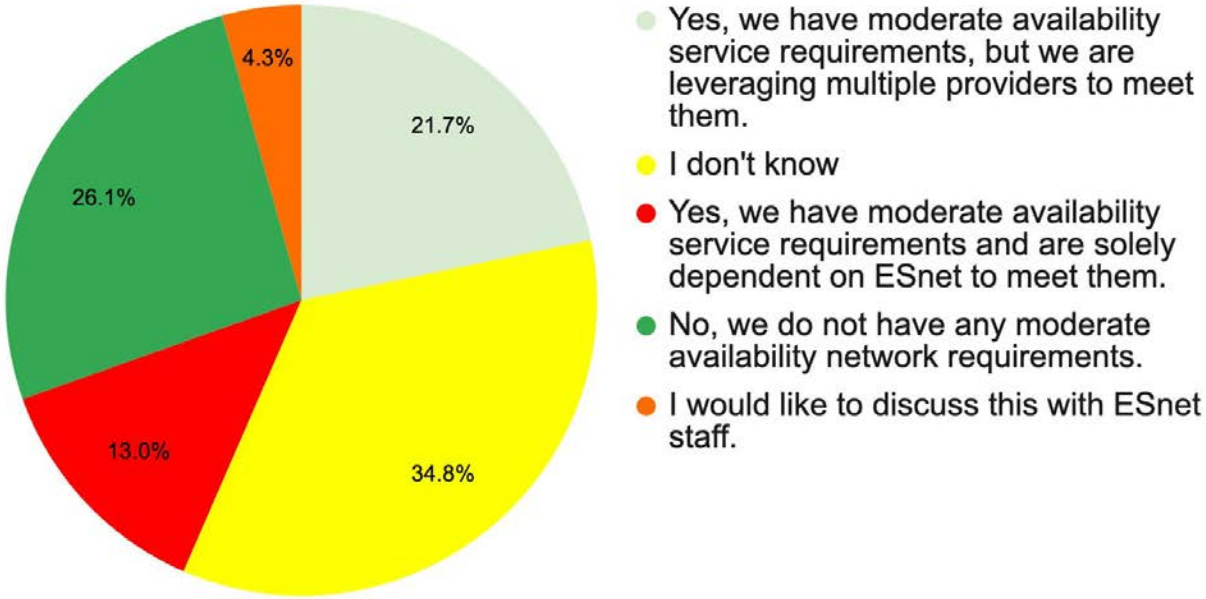
Is classifying Network Availability as Low, Moderate or High sufficient?

No, but it is data I thought we should have.

ESnet

January 2023 Survey of the ESnet Sites

ESnet is a FISMA Low Availability Facility. Are there any FISMA Moderate Availability or higher enclaves or DOE/DHS designated "High Value Assets" at your site which are relying solely on ESnet for external network services?



How should we quantify Availability or Resilience Requirements?

Should requirements be specified as:

- A level of **Redundancy**?
- A **Design Pattern**?
  - The "ESnet Resilience Score" is essentially a ranked set of design patterns.
- **Mean Availability** (ie the Number of Nines?)

Nines	Availability	Downtime / Year
2	99%	3.65 days
3	99.9%	8.76 hours
4	99.99%	52.56 minutes
5	99.999%	5.26 minutes

ESnet

ESnet
ENERGY SCIENCES NETWORK

SRP Engineering: Estimating 9's

Engineering is the application of science and math to solve problems.

Wikipedia:

Engineering is "The creative application of scientific principles to design or develop structures, machines, apparatus, or manufacturing processes, or works utilizing them singly or in combination; or to construct or operate the same with full cognizance of their design; or to forecast their behavior under specific operating conditions; all as respects an intended function, economics of operation and safety to life and property"

NASA:

Engineers are problem solvers. They combine the principles of science and math with a sense of creativity and innovation, improving society by putting STEM into action.

If we are 'network engineering', then we need to do more math. 

ESnet

Initial Assumptions

- Starting with a **very simple** model and refine each cycle.
  - Simplifying assumptions **we *will* want to revisit**:
    - All components have 99.9% availability (all circuits, routers, etc)
    - Parallel circuits between the same nodes are ignored because they have significant shared fate, and we want to keep it simple.
    - Port/Card diversity within a device is ignored
    - Everything multi-layered is ignored. (IE completely ignoring OLS. OLS provided circuits are just like any other.)
    - Not differentiating between outages due to ESnet vs ESnet Subcontractors vs Site vs Site Subcontractors.
    - Only looking at the portion of the network starting at the sites edge router, to the first ESnet core router.
- Look at relative numbers this iteration, don't stress on absolute numbers!
- We can improve our assumptions each time we run through a cycle.

ESnet

Computing System Availability

Availability of Simple Systems:

- 2 Components in Series = $A_1 * A_2$
- 2 Components in Parallel = $1 - (1-A_1)(1-A_2)$

Complex systems can be calculated using:

- Analytical Combinatorial models
- Fault Tree Analysis
- Markov Models
- Monte Carlo Simulations Models*
- A good lecture is here and 40 page primer is here!

Many of our site connection systems are complex, not simple!

- I am using Isograph NAP ($,$$$) to model and analyse small networks using using multiple methods.
- Also played with Relyence RBD ($$$)
  - Works for very simple network patterns like the following slide.
  - But difficult to analyze complex networks.
- Have not found any open-source solutions.
  - But in theory, it shouldn't be too hard…

ESnet

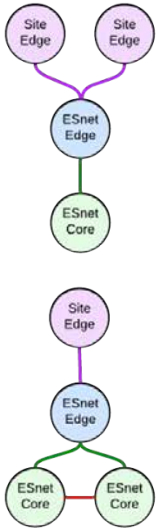
ESnet Connection Model Output:

Model **assumes** all components start at 99.9%

Resilience Score 0

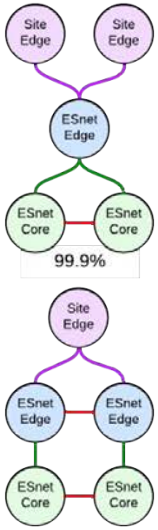
5 Single Points of Failure
99.5%

Resilience Score 1



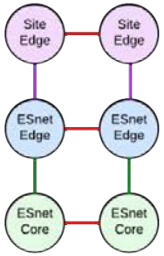
3 Single Points of Failure
99.7%

Resilience Score 2



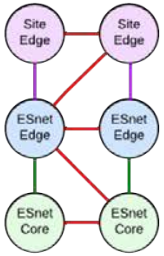
1 Single Point of Failure
99.8988%

Resilience Score 3



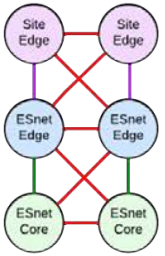
No Single Points of Failure
99.9983%

Resilience Score 3.5



Survives Most Dual Failures
99.9987%

Resilience Score 4



Survives Many Triple Failures
99.9997%

OK, we have the model output. Now what?

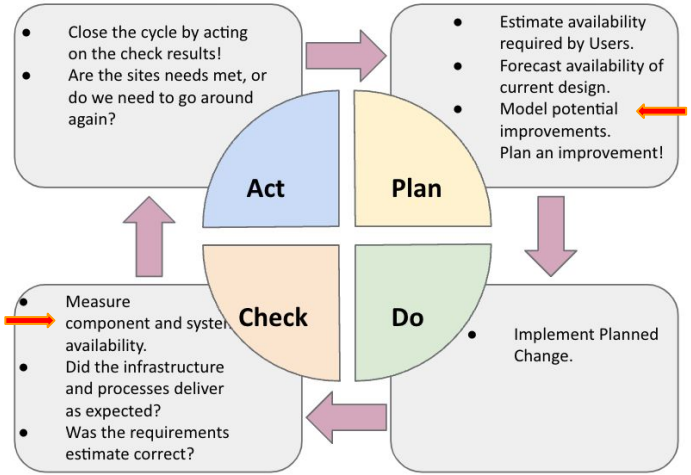
Fun Quotes:

- George Box: “All models are wrong, some are useful”.
- Idit Cohen: "The exponential distribution is the most widely used distribution in system engineering applications. It is mainly because it is the simplest function to work with, not because it is a correct one."

Don't forget the assumptions!

- Garbage in, garbage out!
  - My input wasn't garbage, but it wasn't right either...
- The results forecast **average availability** over time.
  - **There should be long periods that are much better, and short intervals that are much worse!**

Implement a Closed Quality Engineering Loop!



ESnet

Ideas for putting this into practice.

- Encourage more explicit communication about:
  - Resilience requirements & risk tolerances
  - Availability expectations
- Explore more rigorous management practices
  - Closed loop Quality Management Framework (Plan, Do, Check, Act)
  - Comprehensive risk management
- Do more modeling before picking solutions for complex changes, or changes in complex situations.

ESnet

Questions...

ESnet