Sandia National Laboratories

Exceptional service in the national interest

QuESt Grid Planning Toolbox Enhancements:

LARGE LOADS, RELIABILITY-CONSTRAINED PLANNING, & INTEGRATED PLANNING FOR ENERGY STORAGE

Dr. Cody J. Newlun

Electric Power Systems Research Department

DOE OE Energy Storage Peer Review & Annual Meeting, 2026

Presentation ID #: 504

SAND2026-23904C

OVERVIEW & TEAM

- Motivation
- **Large Load Modeling** in QuESt Planning
- **Reliability**-based Capacity Expansion Planning (CEP)
- **QuESt Grid Planning Toolbox** updates
- Project Impact and next steps

Team:

- ***Dr. Dilip Pandit***
- ***Dr. Atri Bera***
- ***Dr, Tu Nguyen***
- ***Andres Lopez*** (PhD Student at Florida International University)
- ***Gustavo Cuello-Polo*** (PhD student at Iowa State University)
- ***Chandan Chaudhary*** (PhD student at Michigan State)
- ***Veronica Diaz Pacheco*** (PhD student at North Carolina State)

QuESt Planning

progress

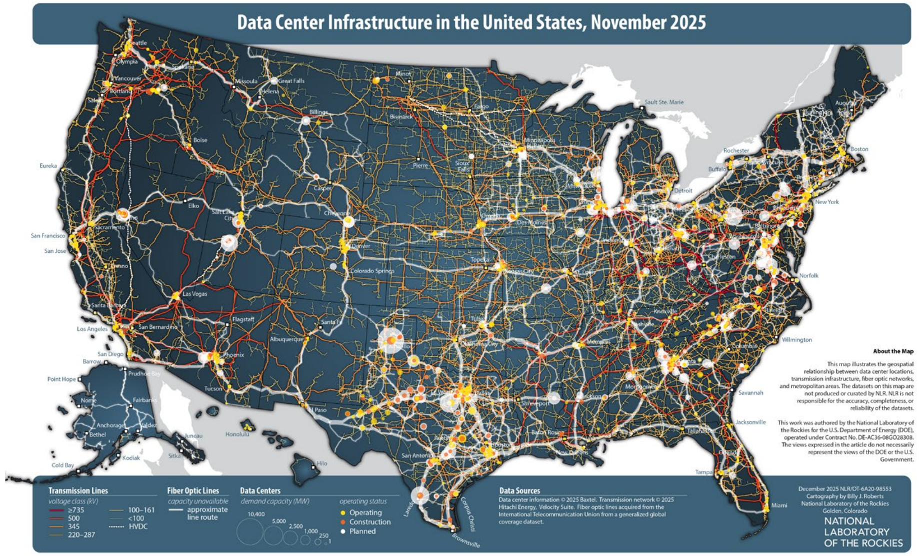
QuESt PCM

PROJECT OVERVIEW

- **Goal**: Develop models and novel frameworks to accurately evaluate energy storage technologies in open-source grid planning tools
- **Current Practice**: Existing capacity expansion planning models and grid planning frameworks lack large load modeling and enhanced ESS models. Furthermore, they are not well coordinated with other tools such as production cost modeling and resource adequacy.
- **Why Sandia**: Sandia has SMEs in power systems, energy storage, optimization, and software development to produce reproducible and impactful studies.
- **Innovation**: Developing an open-source capacity expansion planning model to better represent ESS and co-optimize grid investments to meet reliability and affordability targets.
- **Duration**: FY24 - FY26
- **Impact**: The open-source tools benefit utilities, researchers, and grid planners to evaluate ESS technologies and make reliable and affordable investments
- **Vertical/Investment Area**: Analytics
- **Alignment/Strategic Pathway**: Optimize grid investments to expand the U.S. energy infrastructure to ensure grid security, reliability, and affordability.

MOTIVATION

- **Problem**: Rapid load growth from AI data centers pose significant economic and reliability risks to the U.S. power system.
- Must ensure that the U.S. power system is **reliable**, **affordable**, and **secure**.
- ESS are deployed for multiple services that enhance the reliability, stability, resilience, and security of the grid

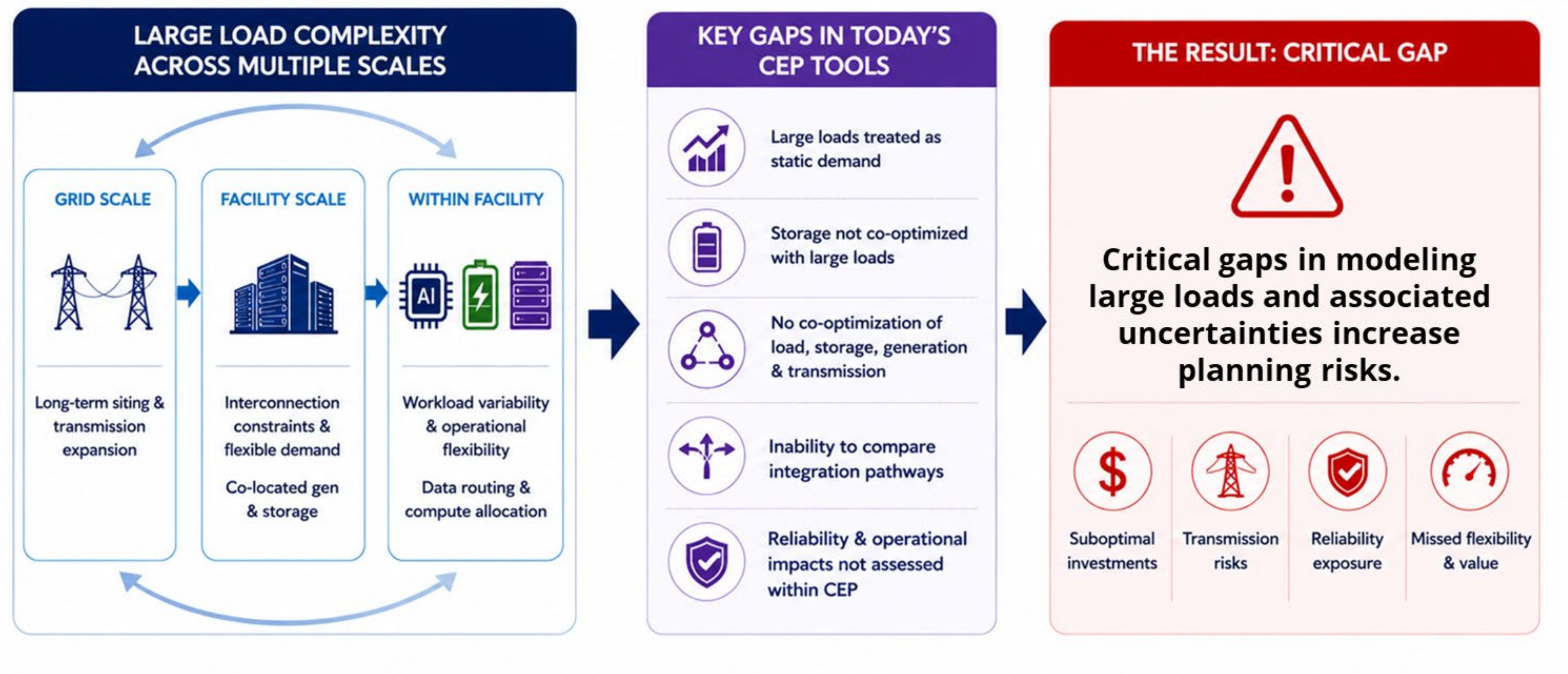


https://docs.nlr.gov/docs/gen/fy26/98553.jpg

Requires ***coordinated*** planning tools to capture ***detailed operations***, address long-term ***uncertainties***, and ensure ***grid reliability*** to make well-informed ESS investment decisions

CAPACITY EXPANSION PLANNING WITH LARGE LOADS: KEY GAPS

- With large loads, expansion planning complexity spans system to module levels; key gaps exist in today's grid planning tools:



Failure to improve grid planning tools will result in sub-optimal grid investments and pose reliability risks to the grid

FLEXIBLE LARGE LOAD AND STORAGE CO-OPTIMIZATION IN CEP

Large electric loads are modeled as flexible, grid interactive resources with co-optimized generation, storage, and transmission

Large Load Characteristics

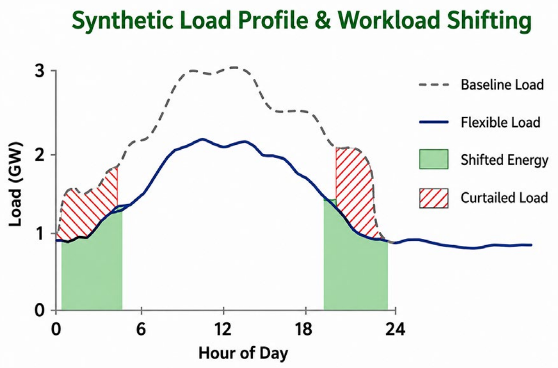
Peak demand

Load type (AI data center, crypto, etc.)

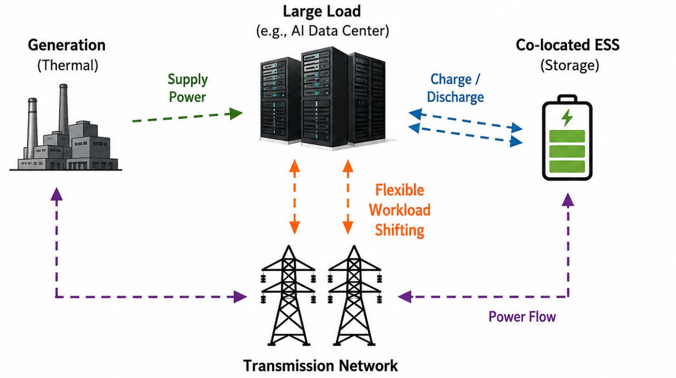
Deployment year

Flexible vs. critical workloads

Temporal variability



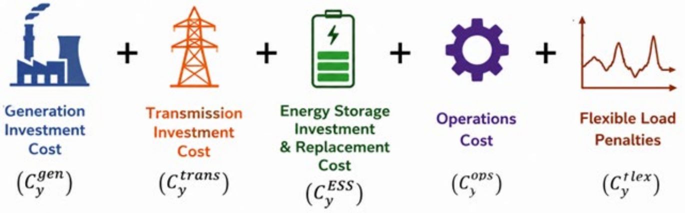
Co-optimization



Enhanced CEP Objective

Minimize total system cost

$$\min \sum_{y} \left(C_y^{gen} + C_y^{trans} + C_y^{ESS} + C_y^{ops} + C_y^{flex} \right)$$



- Incorporate flexibility and reliability costs in objective
- Co-optimize storage, generation, transmission, large load operations, and co-located generation/storage

FLEXIBLE LARGE LOAD OPERATIONS

1. WORKLOAD SHIFTING (Energy Neutrality)	$\sum_{t \in T} S_t = 0$	$\forall t$
2. RAMP RATE LIMITS (Flexible Load)	$-R^{down} \leq L_t - L_{t-1} \leq R^{up}$	$\forall t$
3. MINIMUM UPTIME (Critical Workloads)	$U_t \geq U^{min}$	$\forall t$
4. CURTAILMENT / INTERRUPTIBILITY	$0 \leq Curt_t \leq \alpha^{max} L_t$	$\forall t$
5. FLEXIBLE LOAD BALANCE (Net Load Composition)	$L_t = L_t^{crit} + L_t^{flex} - Curt_t - S_t$	$\forall t$

L_t : Total large load
L_t^{crit}: Critical (inflexible) workload
L_t^{flex}: Flexible (shiftable) workload
S_t : Net shifted load (positive = shifted out, negative = shifted in)
$Curt_t$: Curtailed / interrupted load
U_t : Uptime (1 = on, 0 = off)
R^{up}, R^{down} : Ramp up / ramp down limits
α^{max} : Maximum curtailment fraction

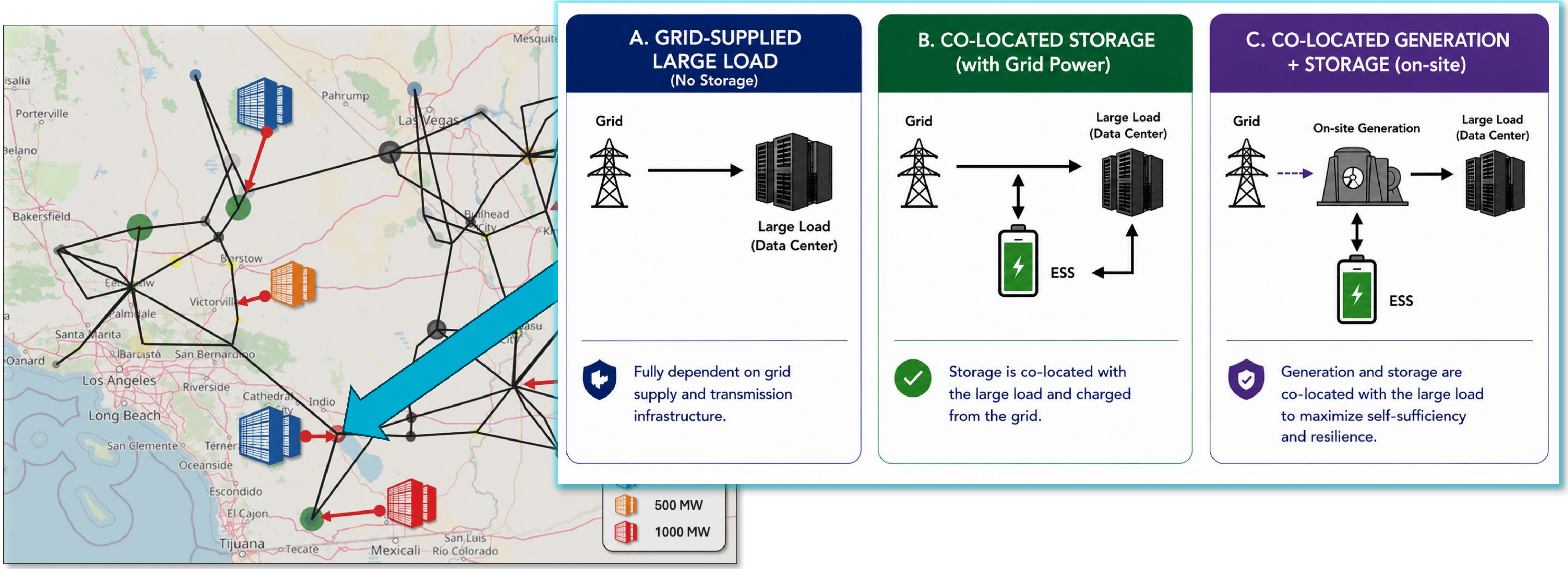
KEY
All constraints apply for every time period t.

Modeling ***flexibility*** allows for optimal sizing and siting of ESS

Key Research Contributions

- ***Endogenously*** model large loads in CEP with representation of flexibility and reliability
- Captures ***multi-timescale coupling*** long-term investments and hourly operations
- ***Co-optimization*** of large loads, co-located storage, generation, and transmission

LARGE LOAD CONFIGURATIONS AND TRADEOFF ANALYSIS IN THE CEP



Role of Energy Storage

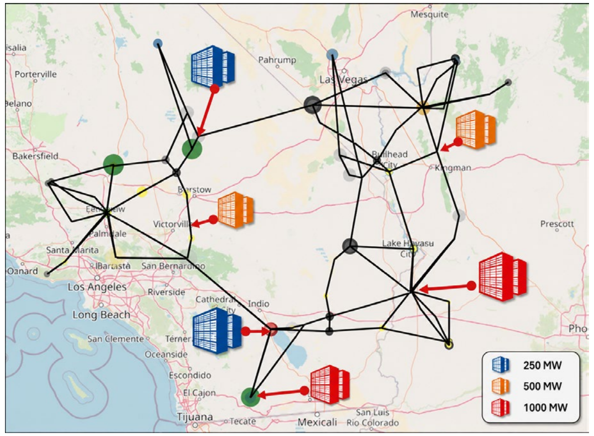
- When co-located, storage can provide flexibility needed to shift, shape, and meet large load demands
- Supports the bulk power system and enables transmission deferment
- Co-located generation and storage enhances energy independence and provides highest reliability for critical large loads

Tradeoff Analyses

- ***Costs vs. Reliability (Load Curtailment)***
- **Transmission expansion vs. Operational flexibility**
- **Implementation configuration vs. Economic Benefit**

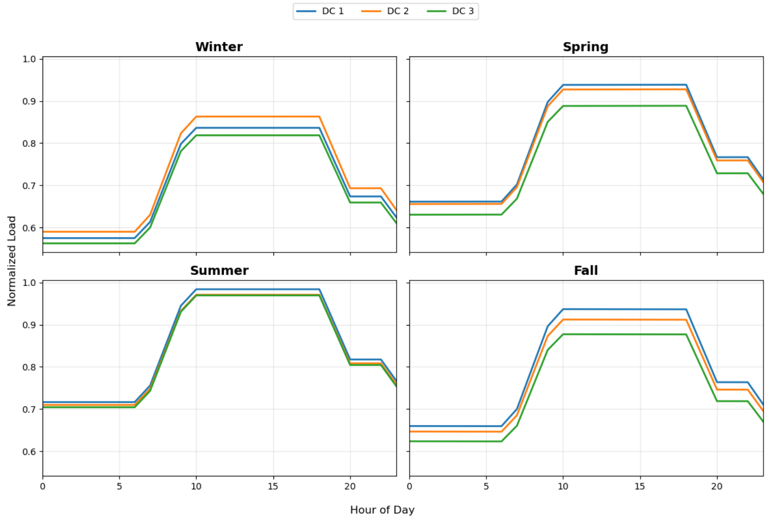
LARGE LOAD CASE STUDY

Model Parameters	Assumption
System	RTS-GMLC nodal model
Large Load Size	500MW & 1000MW
Flexible Portion	30%
ESS Duration	4-hour
ESS Candidate Size	0–1 GW
Planning Horizon	3 investment years (15 years)
Temporal Resolution	3-linked representative days per season & peak day (hourly)
Capital Costs	Assumed from NLR ATB[1]
System-wide annual load growth	3%



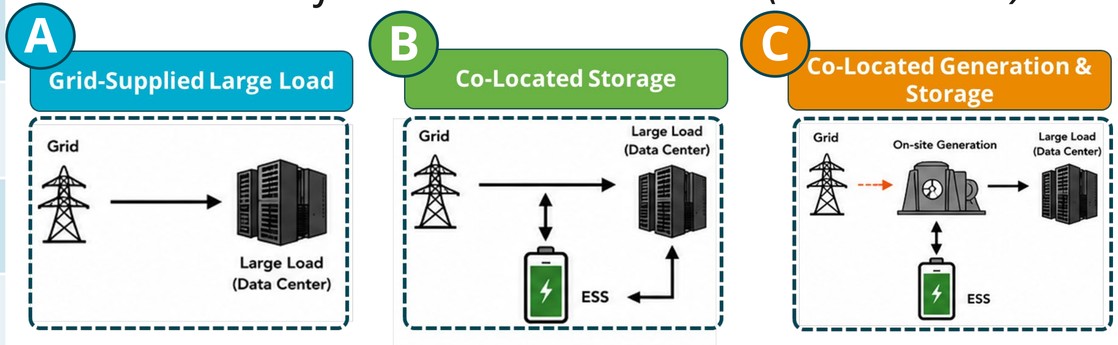
Large loads are deterministically sited onto the RTS GMLC grid

Seasonal daily large load shapes – scaled deterministically



Scenario	Large Load Commissioning	Large Load Configurations	Large Load Flexibility
Baseline + Flex	Base (1.5 GW) (year 1 & 5)	A	✓
Aggressive Growth + ESS	Aggr (3.5 GW) (year 1, 8, 15)	A, B	✓
Aggressive Growth + No Flex	Aggr (3.5 GW) (year 1, 8, 15)	A, B, C	X
Aggressive Growth + Flex	Aggr (3.5 GW) (year 1, 8, 15)	A, B, C	✓

***Large load flexibility**: able to curtail up to 25% and account for curtailment costs ($200/MWh)*

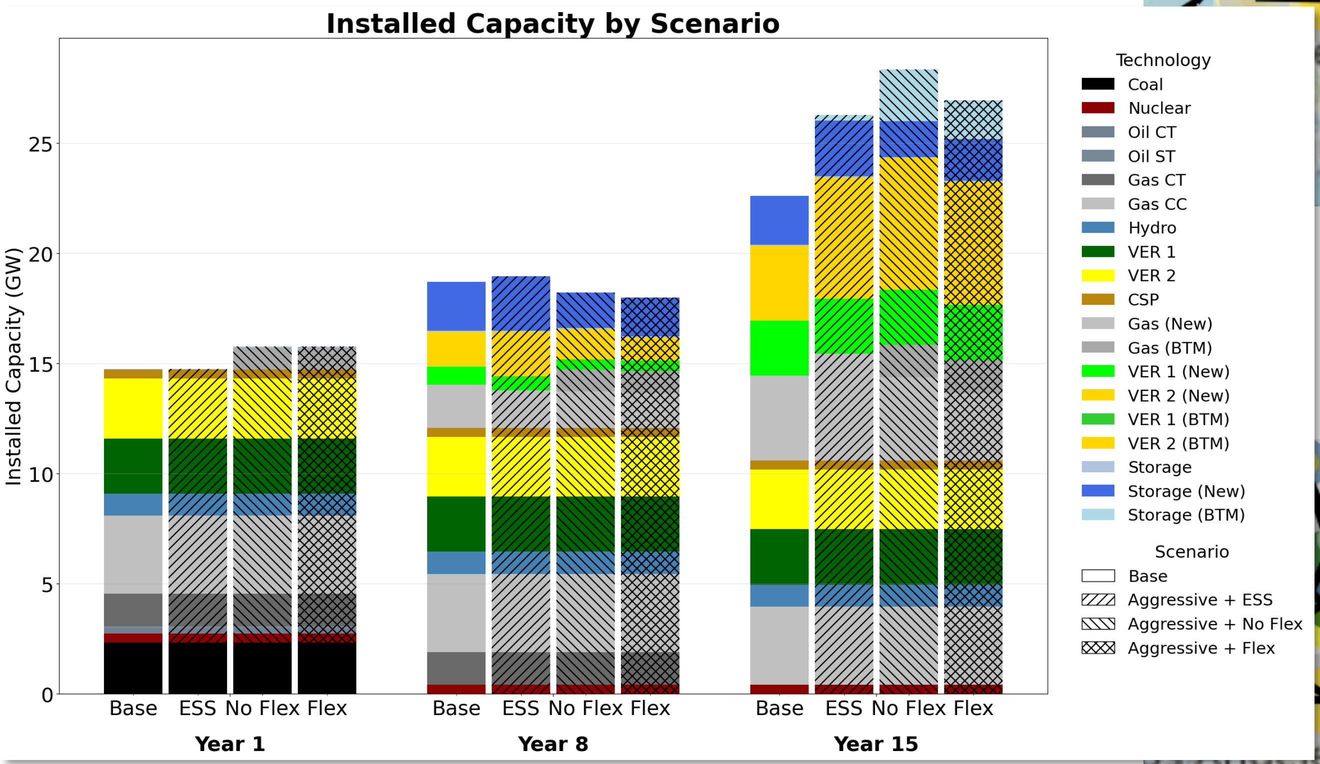


Large load configuration options

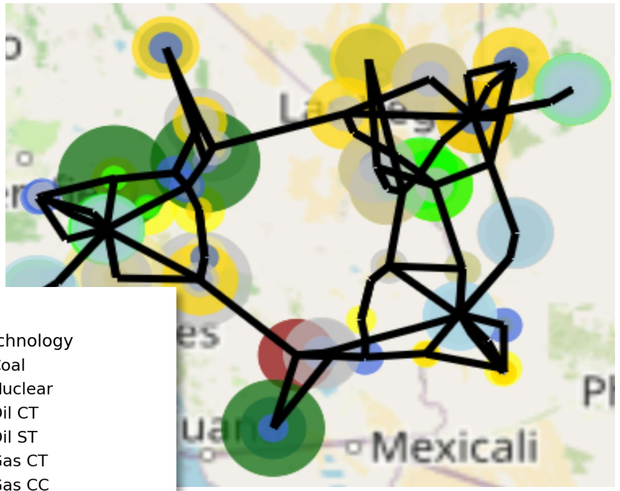
LARGE LOAD CASE STUDY: EXPANSION RESULTS

Key takeaways:

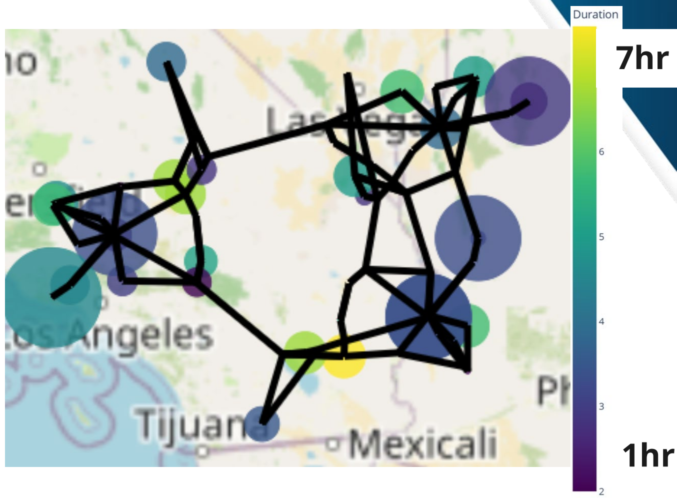
- ESS investments are impacted when implementing flexibility constraints
- Siting and sizing of ESS and generation are impacted when implementing different configurations and flexibility options



Aggressive Growth + No Flex

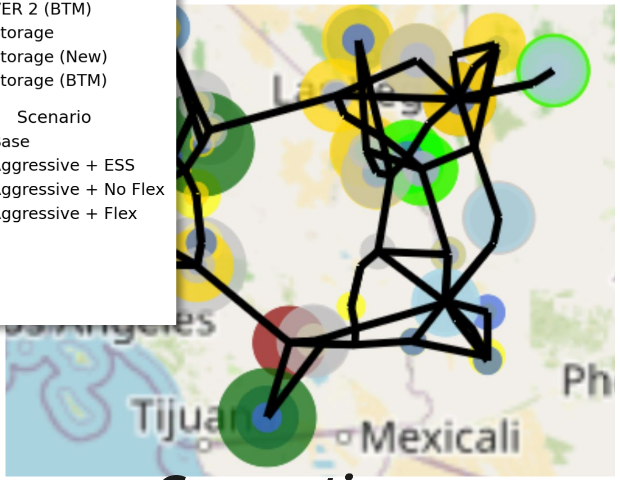


Generation

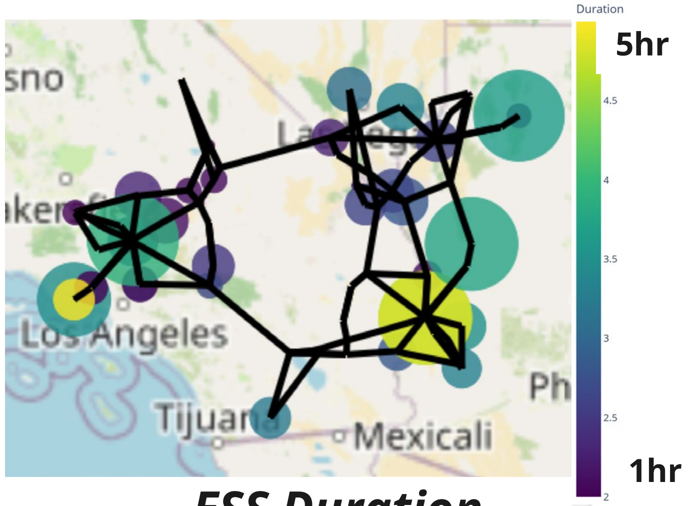


ESS Duration

Aggressive Growth + Flex



Generation

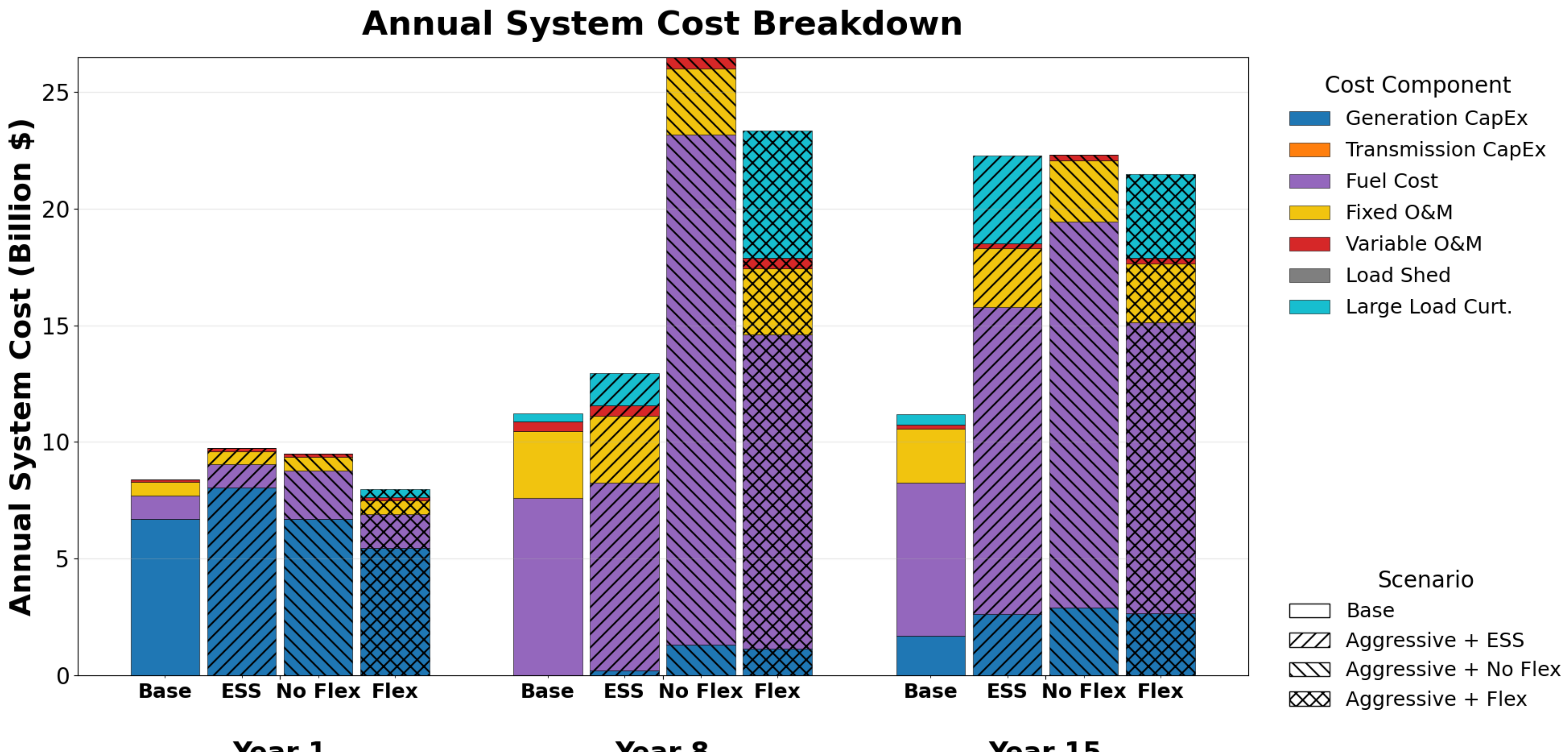


ESS Duration

LARGE LOAD CASE STUDY: RESULTS

Key takeaways:

- Capital and operating costs vary across planning assumptions – fuel costs impacted with large load flexibility
- Accounting for flexibility and various configurations unlocks more implementation pathways for large loads

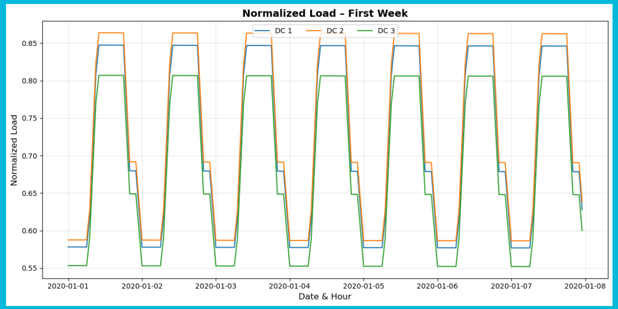


DEVELOPMENT OF LARGE LOAD INTEGRATION FEATURE IN QUEST PLANNING

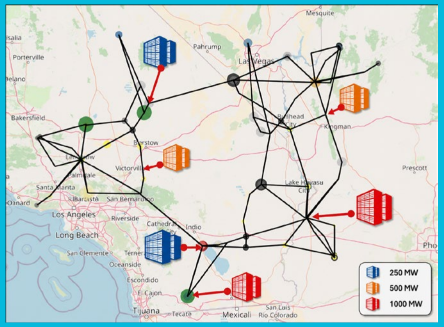
- Allow users to specify size and location of planned large loads with custom load profiles, flexible operations options, and co-located generation options
- ***Goal***: Develop a large load integration feature into QuESt Planning such that users can evaluate ***economic tradeoffs*** and assess the ***reliability and operations*** of future power systems with rapid integration of large loads via the QuESt Planning toolbox

Large Loads Integration Tool


```yaml
#***************Large Load Nodal Adjustment*****************
large_load_nodal_loads_option: True
large_loads:
  - id: LL_1
    bus: 3
    capacity_mw: 500
    deploy_year: 2028
    profile_file: data_center_profile_normalized_from_100MW.
    profile_type: normalized   # "normalized" or "absolute"
    profile_timezone: local
    baseline_pue: 1.15
    pue_sensitivity_alpha: 0.02
    flexibility:
      curtailment:
        enabled: true
        max_fraction: 0.1
        curtail_cost: 1000
      shift:
        enabled: true
        max_fraction: 0.1
        max_window_hours: 6
      onsite_resources:
        pv:
          candidate: true
          max_capacity_mw: 200
        bess:
          candidate: true
          max_energy_mwh: 800
          max_power_mw: 200
```

Synthetic load profile

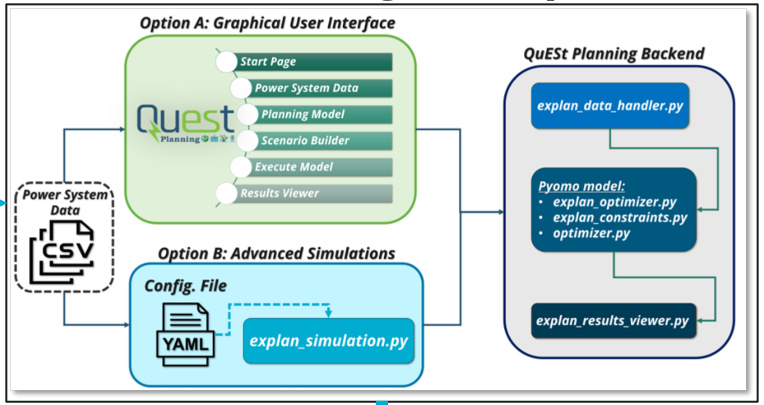


1 *Construct yaml configuration file to specify large load parameters*

2 *Define hourly load profile*

3 *Select buses on transmission system for integration*

QuESt Planning CEP Optimizer



Evaluate tradeoffs & Assess grid reliability

RELIABILITY-BASED CEP & QUEST GRID PLANNING TOOLBOX UPDATES

RELIABILITY-BASED CEP

- **Problem:**
  - Load growth poses significant challenges for system reliability
  - Ensuring long-term adequacy amid load growth calls for reliability-aware capacity expansion models
- **Goal:** Develop a reliability-constrained capacity expansion planning approach
- **Proposed method** [1]:
  - Quest Planning generates incumbent investment decisions
  - Analytical reliability framework performs a quick adequacy assessment for all incumbent generation mixes
  - For any year with reliability violations, present a reliability cut back into Quest Planning using persistent solver interface
  - Repeat the process until no violations are present

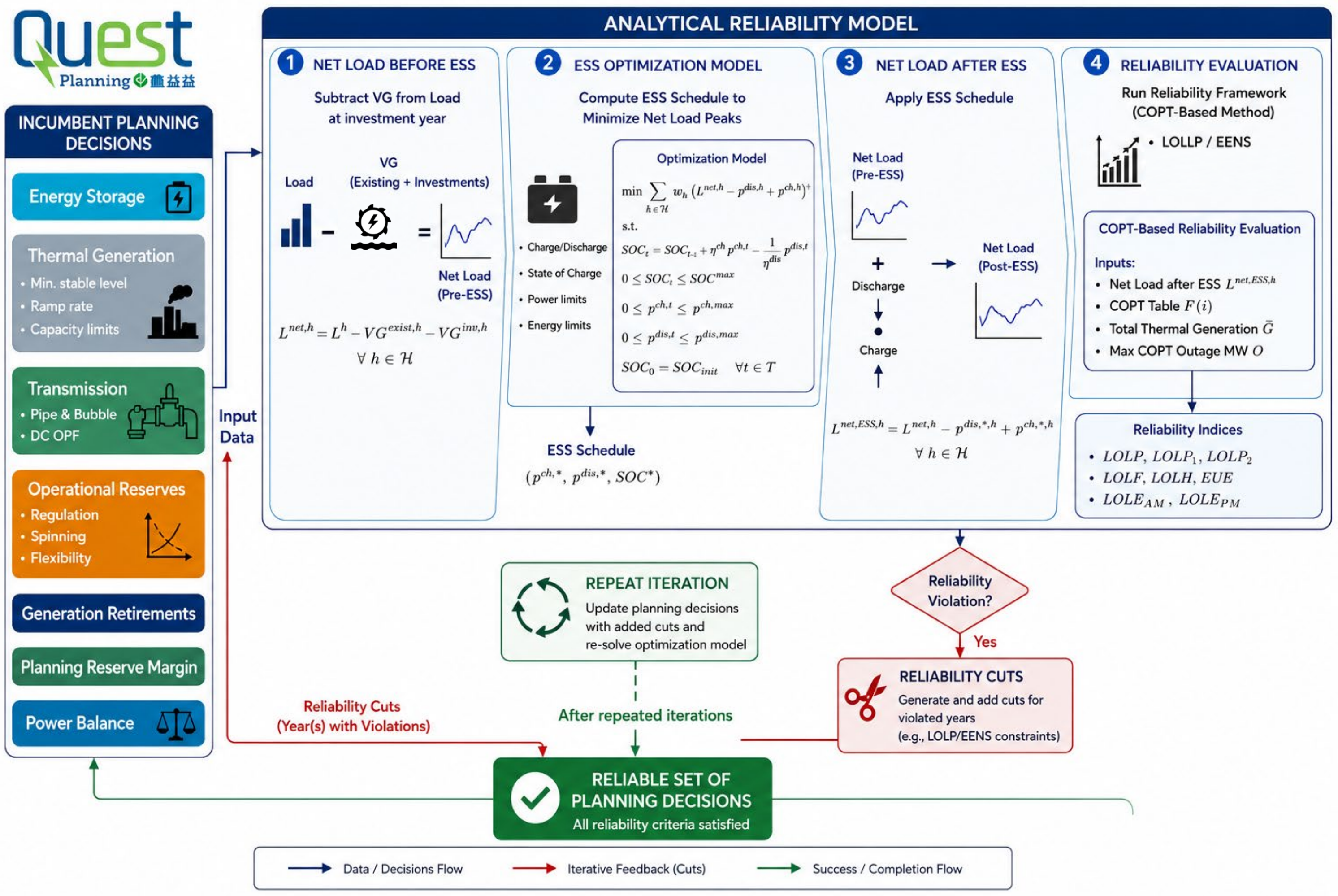
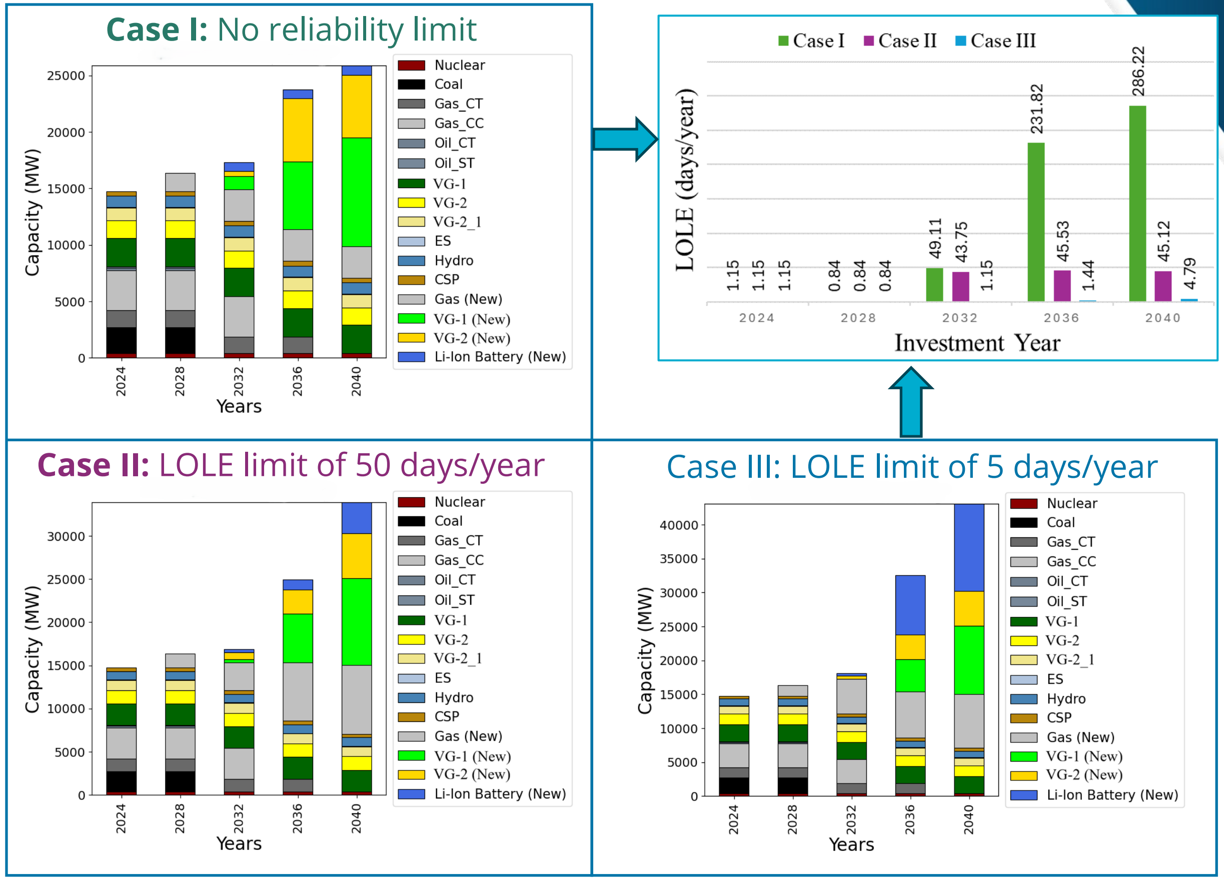


Fig: Overview of the proposed approach

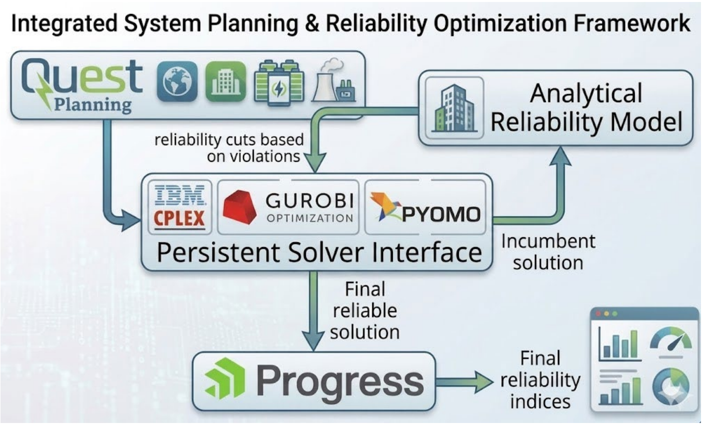
[1] Pandit, D., Newlun, C. J., Bera, A., Cuello-Polo, G., Nguyen, T., & Byrne, R. (2026). A Reliability-Constrained Generation and Storage Capacity Expansion Planning Framework. Paper accepted for presentation at IEEE PMAPS 2026.

RELIABILITY-BASED CEP CASE STUDY

- Case studies performed on the RTS-GMLC system
- Key parameters:
  - **Investment years:** {2024, 2028, 2032, 2036, 2040}
  - **Temporal scope:** Seasonal load blocks
  - **Candidates:** Gas CT (20 candidates), VG-1 (73 candidates), VG-2 (73 candidates), 2h Li-ion (71 candidates), each rated 400 MW
  - **Retirements:** Coal (2032), Oil (2032), Existing Gas CT (2036)
- Case studies show that:
  - **Reliability Drives Storage:** Increasing LOLE limits directly accelerate storage investments
  - **Approach is robust:** Solutions remain tight to the prescribed probability or energy indices



Ongoing Work



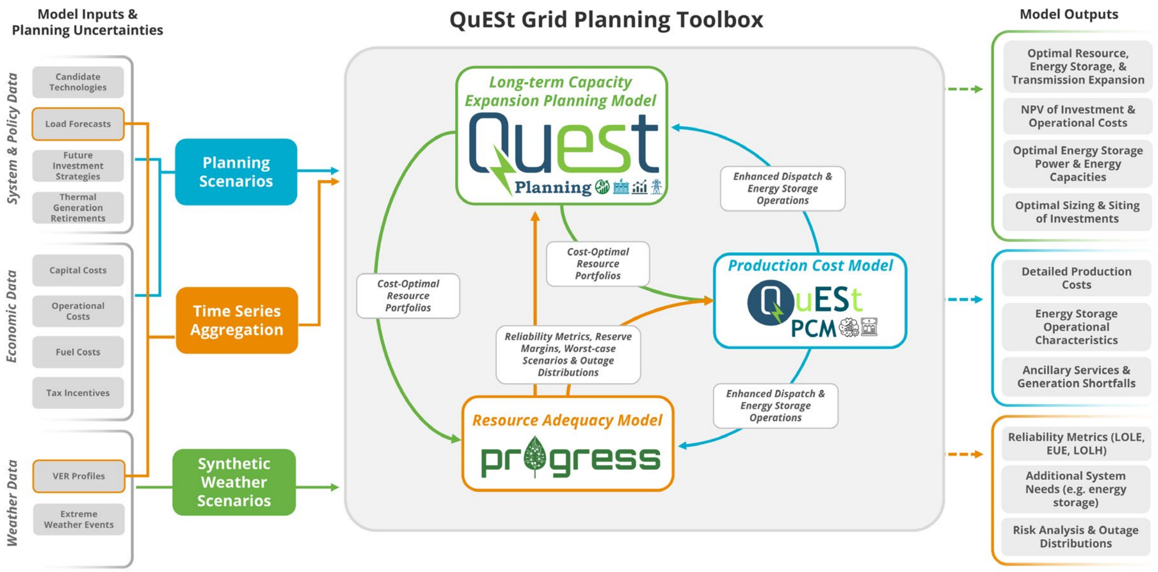
Testing robustness of the approach to different reliability criteria

Test 1: LOLE limit of 5 days/year.
Test 2: Normalized EUE (nEUE) limit of 1%.
Test 3: LOLH limit of 1000 hours/year.

Year 2040 reliability indices and investment profile

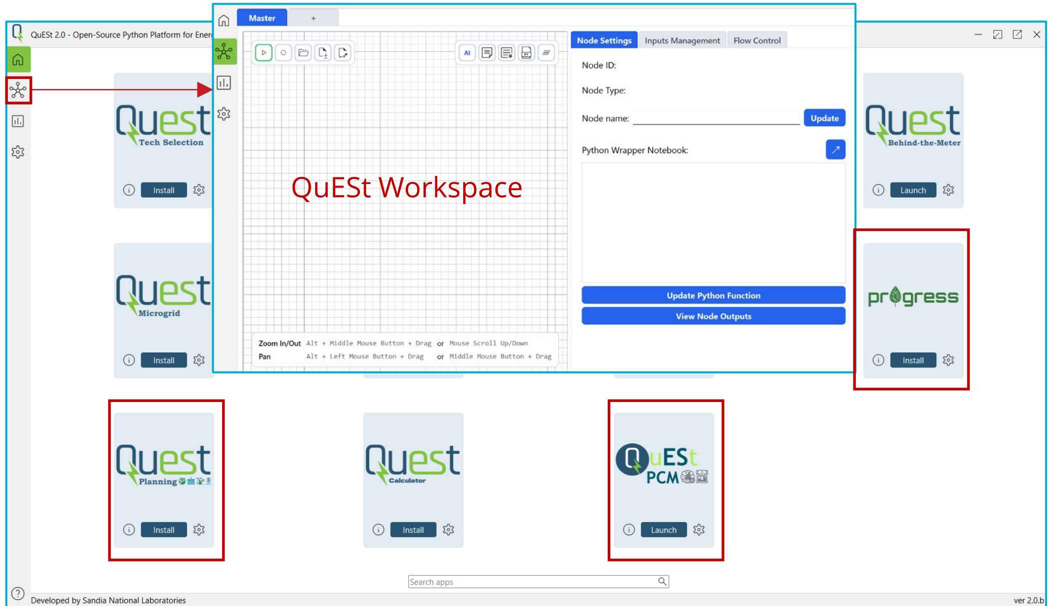
Case	Reliability Indices				Candidates invested		
	LOLE (days/yr)	LOLH (hours/yr)	LOLF (events/yr)	nEUE (%)	Gas GW	VG GW	ESS GW
Base	286.2	2749	392.42	9.419	0	15.2	0.8
Test 1	4.787	34.98	4.787	0.0153	8.0	15.2	12.4
Test 2	100.5	473.4	100.5	0.9011	7.6	15.2	0.4
Test 3	186.3	995.2	191.5	2.5643	5.6	16.8	0.4

QUEST GRID PLANNING TOOLBOX



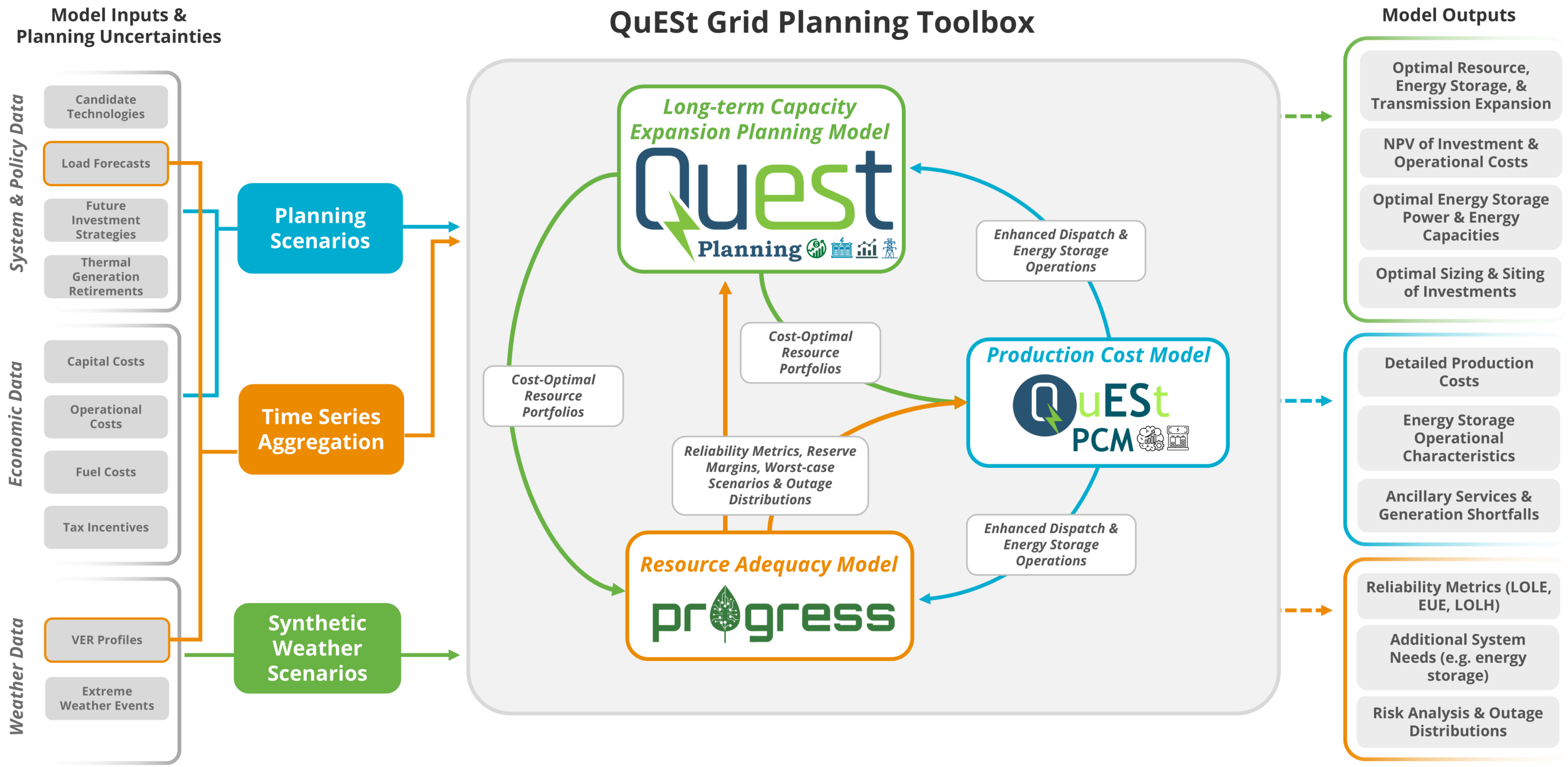
- **QuESt Grid Planning Toolbox:**
  - **QuESt Planning**: optimal generation, ESS, and transmission investments
  - **ProGRESS**: assess system reliability under evolving grid conditions
  - **QuESt Production Cost Model (PCM)**: evaluate detailed system operations and role of ESS – **released to GitHub January 2026**

QuESt Workspace and ***QuESt AI agent*** facilitate effective integration of these tools



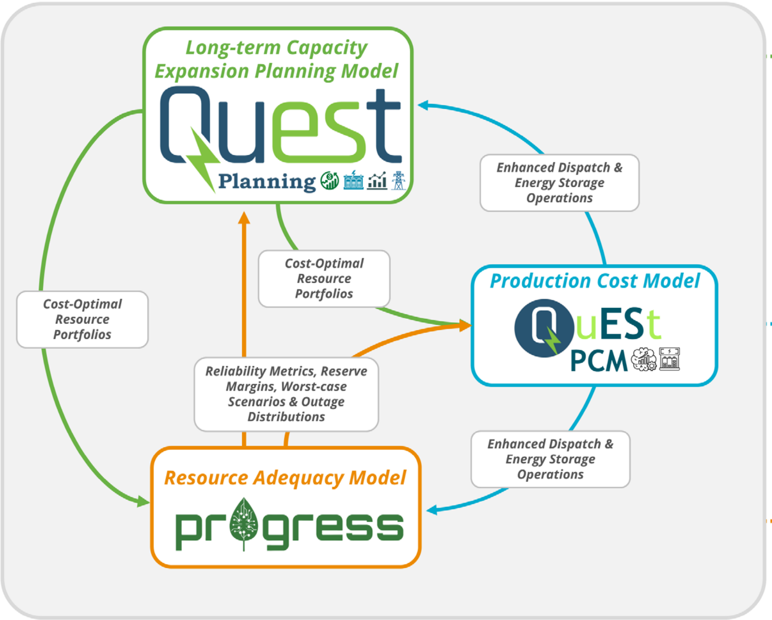
QUEST GRID PLANNING TOOLBOX COORDINATION

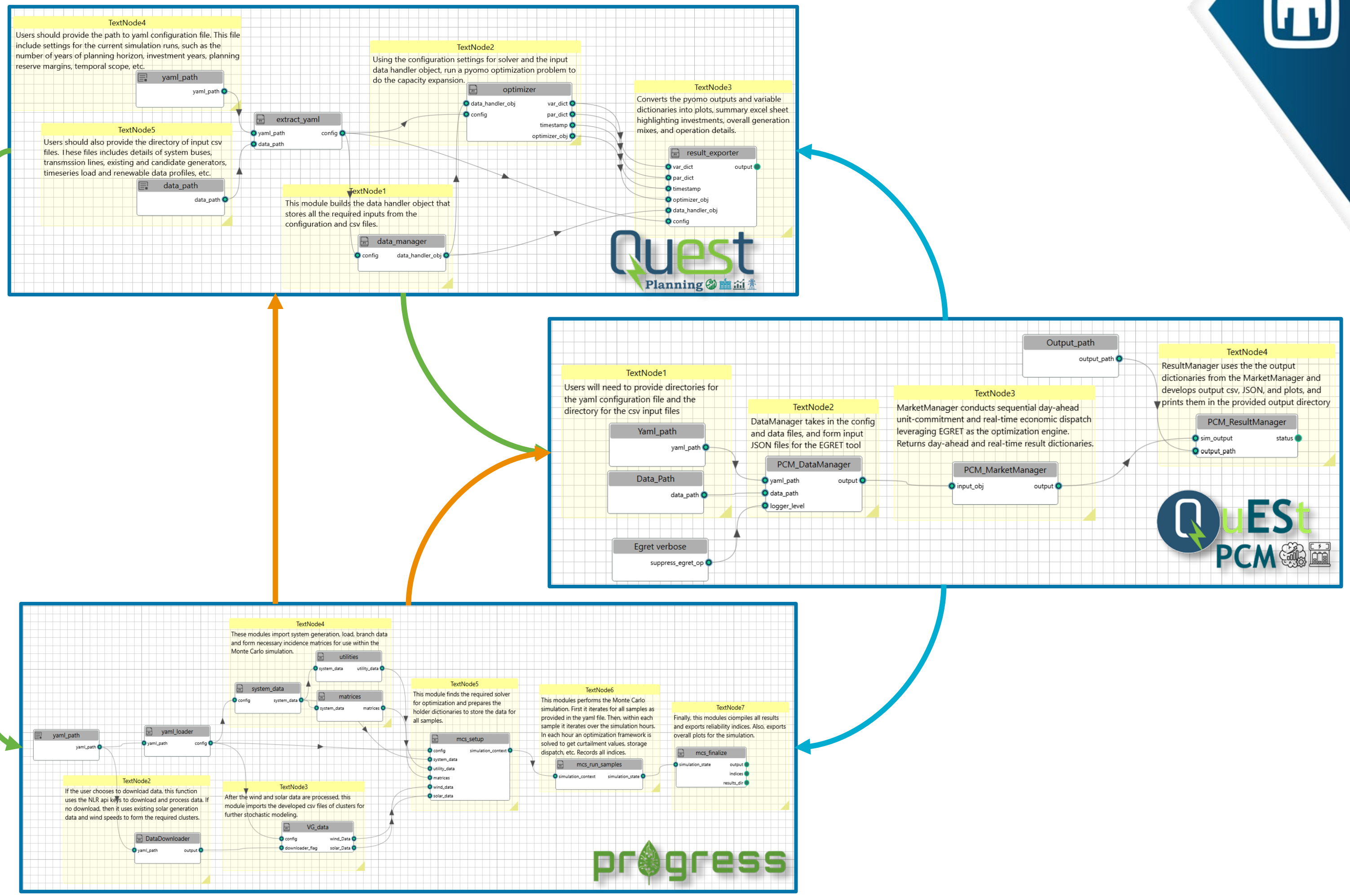
Efficient coordination of grid planning tools is required to optimize the grid and ensure reliability



QUEST WORKSPACE FOR GRID PLANNING

QuESt Grid Planning Toolbox





Tool Integration Challenges

- Tools run in different environments
- Simulation fidelity varies: decades for planning, yearly for adequacy, and sub-hourly for PCM
- Different data requirements for each tool

Why QuESt Workspace & Agent

- Integration of different Python environments
- Visual representation of modules, processes
- Enhanced workflow development, runtime debugging, and result interpretation using QuESt Agent

EXAMPLE LEVERAGING QUEST AGENT FOR CO-PLANNING

1. **QuESt Agent responds to user prompt and matches closest skill**

create workflow to do a power system capacity expansion and use it's outputs to perform reliability assessment on all investment years

10:12 AM

Build path: Reuse Skill Workflow JSON

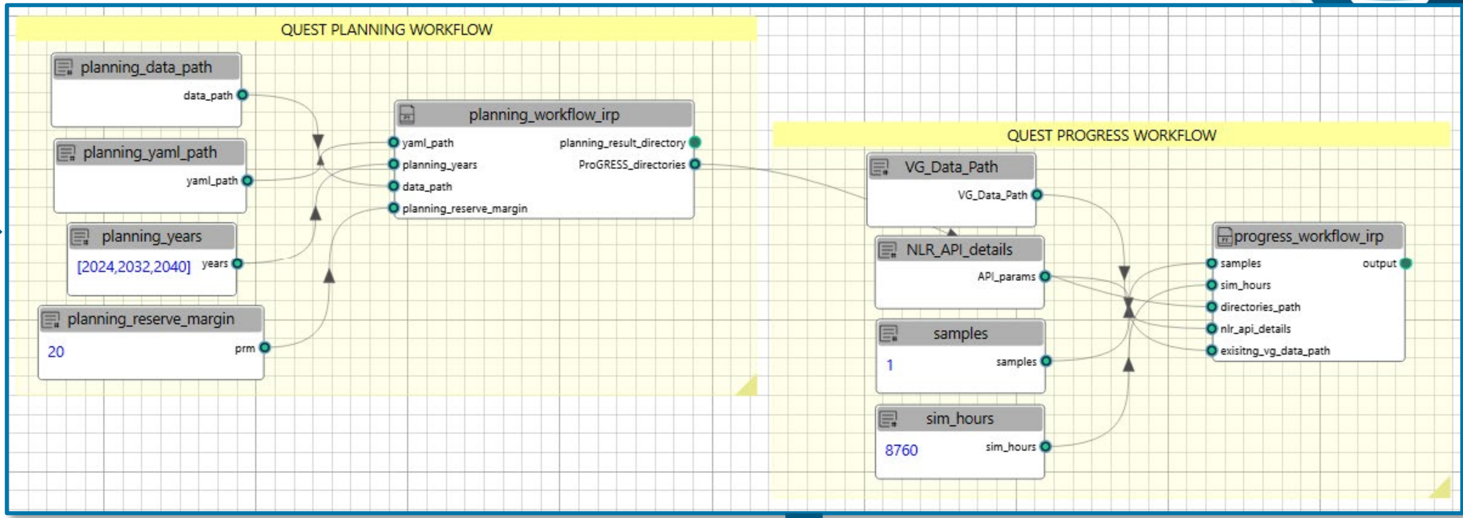
Reason: Matched skill template 'IRP master flow for planning plus resource adequacy assessment' is an exact-enough fit, the canvas is empty, and no explicit differences were requested, so direct workflow reuse is the easiest path.

Proposed canvas plan:

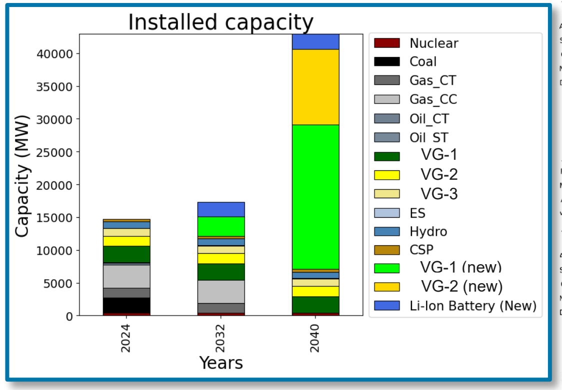
1. Load the matched workflow template from skill 'irp-master-flow-for-planning-plus-resource-adequacy-assessment'.
2. Validate current flow and prepare the next fix-up plan if anything is still missing.

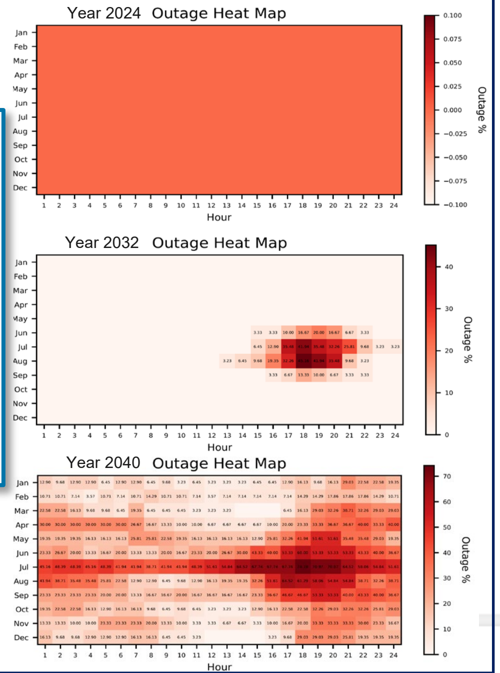
Next step (1/2): Load the matched workflow template from skill 'irp-master-flow-for-planning-plus-resource-adequacy-assessment'. Use the actions below to execute, review, or cancel this plan. You can still send a revised request to change it.

Next | Show Plan | Cancel



2. **QuESt Agent loads workflow**





3. **Uses makes desired changes and runs the workflow**

4. **Agent provides deeper insights into results**

WHAT THE RESULTS SAY NOW

2024: Planning keeps a broad conventional fleet online; Progress shows no adequacy issues.

2032: Planning has retired much coal and oil and adds more variable generation.

2040: Planning builds a very large variable generation system with some storage; conventional thermal is almost gone.

In the exported 2040 Progress system, only 6 thermal generators remain plus storage.

LOLP 0.2, LOLH 20, EUE ~27,172, LOLF 3, LOLE 3.

HOW TO INTERPRET THE INVESTMENT–RELIABILITY LINK

- Planning finds a least-cost expansion under its own assumptions: seasonal blocks, reserve constraints, transmission-style representation, and no annual load shedding cost.
- The least-cost solution is dominated by variable generation by 2040, with very little firm thermal left.
- Progress stress-tests the exported system hour by hour with outages and variable generation variability.
- **Result:** The 2040 least-cost Planning portfolio is likely not firm enough under the current Progress setup, once outages and variability are considered.

WHY THE 2040 RESULT IS SUSPICIOUS BUT INFORMATIVE

It is informative because it indicates a real risk: the 2040 plan may be over-dependent on variable generation and too light on firm capacity.

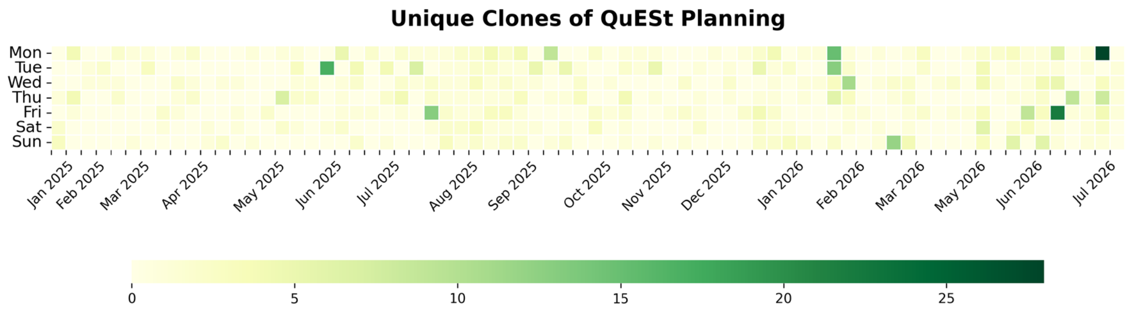
It is also suspicious because the two tools use different system models:

- Planning uses tx_model = Transportation with reserves enabled.
- Progress is forced to model = Copper Sheet.
- Planning solves on seasonal representative structure; Progress simulates chronological stochastic conditions.
- The exported Progress system is simplified CSVs, not the full Planning formulation.

SUMMARY, IMPACTS, & NEXT STEPS

Project contributions in FY26:

- ✓ **Large load integration feature** implemented into QuESt Planning to enable robust analysis for immediate risks on the US power system
  - ✓ Incorporated flexibility constraints to evaluate evolving large load impacts on the grid
  - ✓ Co-optimizing large loads, storage, generation, and transmission
  - ✓ Enabled tradeoff analysis of large load configurations
- ✓ Developed and implemented **reliability-based CEP framework**
- ✓ **QuESt Grid Planning Toolbox:**
  - ✓ **QuESt PCM released January 2026**
  - ✓ AI-enabled workflow developed in **QuESt Workspace via QuESt Agent**



~500 total unique clones

Accomplishments & Impact:

- **QuESt Planning update** with large load modeling features
- Papers presented at ***IEEE EESAT 2026***, ***IEEE PES General Meeting*** and ***PMAPS*** *(accepted)*
- Presented posters at ***IEEE PES Energy and Policy Forum*** – Washington D.C.
- Moderated panel at ***IEEE PES T&D*** (Chicago): "The Role of Energy Storage Systems in Grid Planning & Operations to Support Large Electric Load Growth" with panelists from El Paso Electric, Michigan State University, Columbia University, and Elevate Energy Consulting
- Presented at the ***2026 FERC Software Conference*** on the QuESt Grid Planning Toolbox
- Presented in panel *at* ***IEEE PES GM (Montreal)*** on energy storage and large load impacts
- **Integration with other grid planning tools** like *ProGRESS* and *QuESt PCM* using *QuESt Workspace* allows users comprehensive grid planning options

THANK YOU!

Cody Newlun
cjnewlu@sandia.gov

This material is based upon work supported by the U.S. Department of Energy, Office of Electricity (OE), Energy Storage Division.

U.S. DEPARTMENT of ENERGY

Quest Planning

progress

QuESt PCM

REFERENCES

- Accepted and submitted papers:

  - **C.J. Newlun**, A. Bera, D. Pandit, G. Cuello-Pollo, A. Lopez, Y.J. Pomeroy, T. Nguyen, "QuESt Planning: A Long-term Capacity Expansion Planning Tool for Energy Storage Systems & Modeling Flexibility" 2026 IEEE Electrical Energy Storage Applications and Technologies Conference (EESAT), Tucson, AZ, USA, 2026.
  - **A. Lopez**, C. J. Newlun, D. Pandit, T. Wilcox, A. Bera, and T. Nguyen, "LLM-driven Hyperparameter Optimization Framework for Capacity Expansion Planning Models," in Proc. 2026 IEEE Power & Energy Society General Meeting (PESGM), Montreal, QC, Canada, 2026.
  - **V. Diaz Pacheco**, D. Pandit, C. J. Newlun, A. Bera, T. Nguyen, and J. Kern, "Rolling Horizon Unit Commitment via Polyak-Based Lagrangian Relaxation," in Proc. 2026 IEEE Power & Energy Society General Meeting (PESGM), Montreal, QC, Canada, 2026.
  - **D. Pandit**, C. J. Newlun, A. Bera, G. Cuello-Polo, T. Nguyen, and R. Byrne, "A reliability-constrained generation and storage capacity expansion planning framework," Accepted for presentation at IEEE PMAPS 2026.

BACKUP

AI DATA CENTER LOAD MODEL

- Stochastic data-center reliability model developed using semi-Markov process
- Proposed model:
  - Overall load for data-center i is:

$$P_i(t) = P_i^{base}(h) + P_i^{IT}(t)$$

  - **Base Load** = HVAC + Chiller + Auxiliaries (Driven by ambient Temperature)

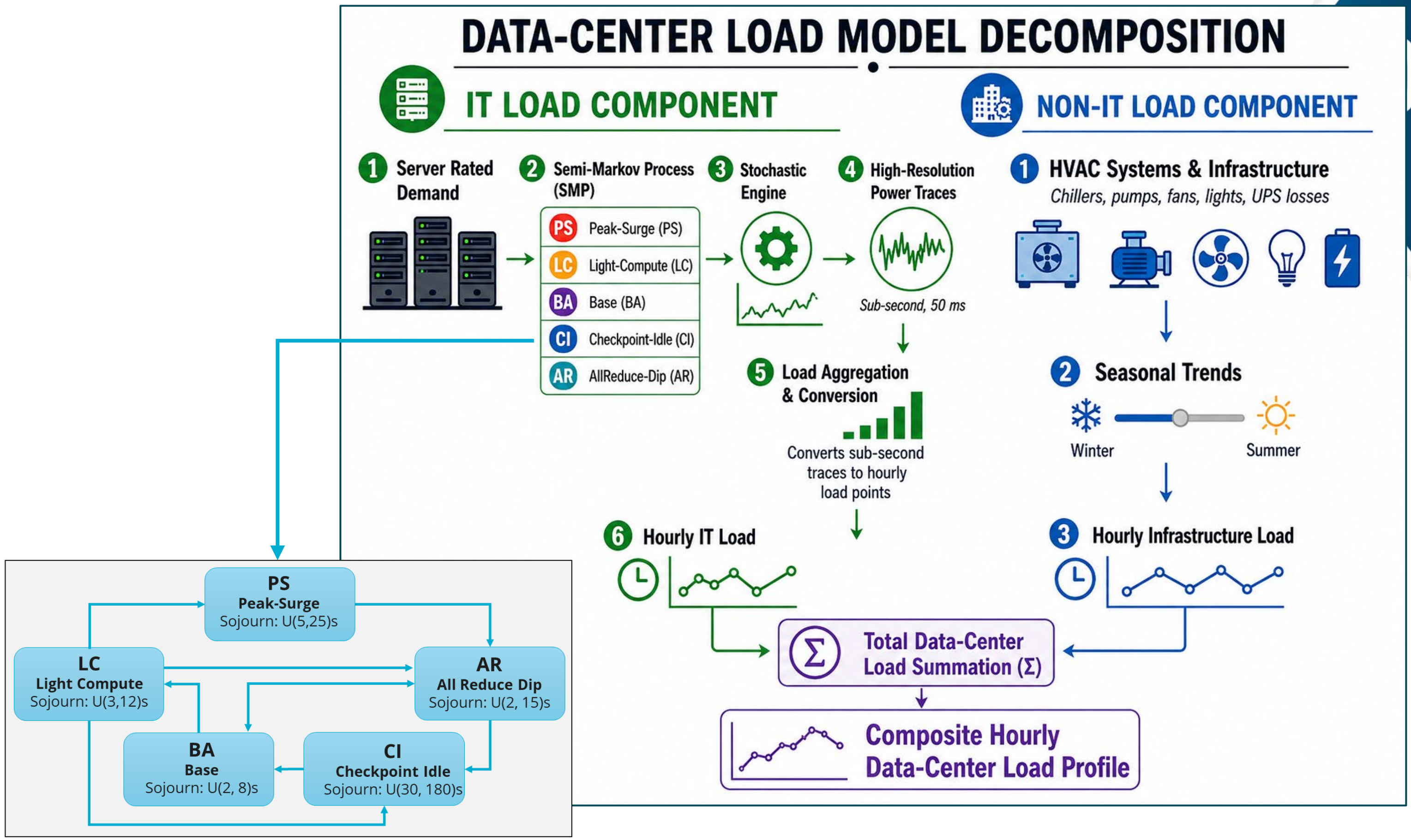
$$P_i^{base}(h) = [\alpha_{cl}(T_{amb}(h) + \alpha_{aux})]P_i^{rated}$$

  - **IT Load** = Fast Stochastic AI Workload *(5 State Semi-Markov Model)* Shaped by diurnal envelope of AI usage, Potential correlation between facilities
    - Nominal hourly IT load is:

$$\overline{P_i^{IT}}(h) = \alpha_{IT} P_i^{rated} \mu_D(h)$$

    - IT load after semi-Markov process is:

$$P_i^{IT}(t) = \overline{P_i^{IT}}(h)(1 \pm \delta_{state} + \sigma_{state})$$



Indices:

α_{cl} = Cooling fraction	α_{IT} = IT load fraction
T_{amb} = Ambient temperature	$\mu_D(h)$ = Diurnal utilization envelope
α_{aux} = Auxiliary fraction	δ_{state} = Power fraction from Markov state
P_i^{rated}= Rated data-center power	σ_{state} = Noise corresponding to the state

Five State Semi Markov IT Load Model

Peak Surge: $\delta_{PS} \sim U(0.25, 0.40)$
- High power surge, Peak tensor-compute burst
- +25 to +40% above nominal load

Light Compute: $\delta_{PS} \sim U(0.25, 0.40)$
- Moderate compute
- +5 to +15% above nominal IT load

Base: $\delta_{BA} = 0$
- Short inter-iteration baseline gap
- Nominal Load

Checkpoint Idle: $\delta_{PS} \sim U(0.15, 0.25)$
- Low Power Idle point
- -15% to -25% of Nominal Load

AllReduce Dip: $\delta_{PS} \sim U(0.35, 0.50)$
- Communication synchronization dip
- -35% to -50% of Nominal Load

AI DATA CENTER LOAD MODEL

- IT load obtained using $P_i^{IT}(t)$ generates sub-second power traces
- Converted to hourly demand for reliability assessment using the peak state encountered during that hour:

$$P_i^{IT,peak}(h) = \max_{state} P_i^{IT}(t) \quad \forall t \in h$$

- Correlation among data center facilities:
  - Shared orchestration or synchronized AI jobs
  - Same-region operational similarity and common environment drivers
- Modeling correlation:
  - ✓ Modeled as a likelihood of two or more facilities operating in same state
  - ✓ Embedded via and workload synchronization (hourly states)