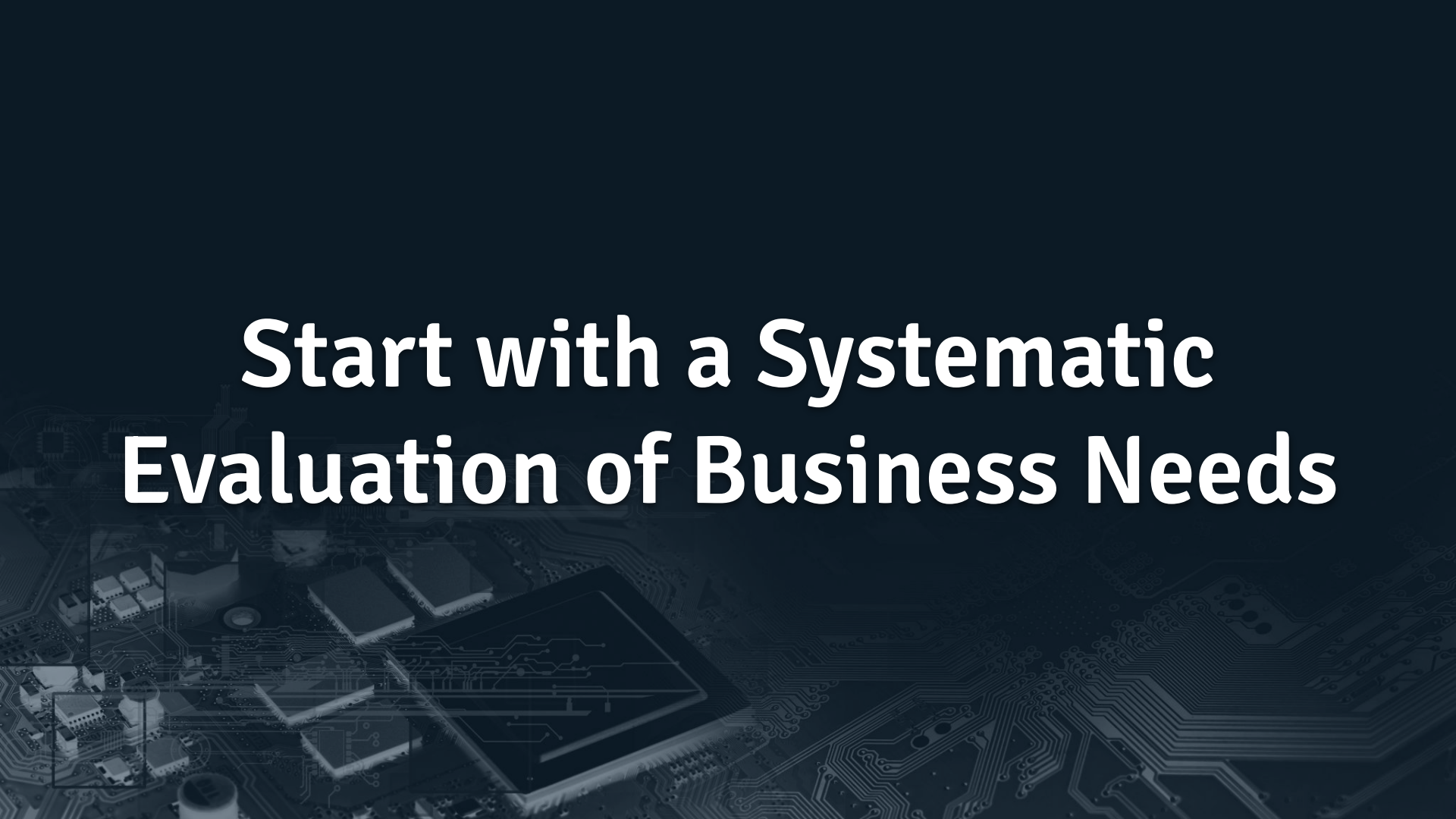
ORBIT: An AI-Powered Assistant for NOC Operations in a Data-Intensive Science Network

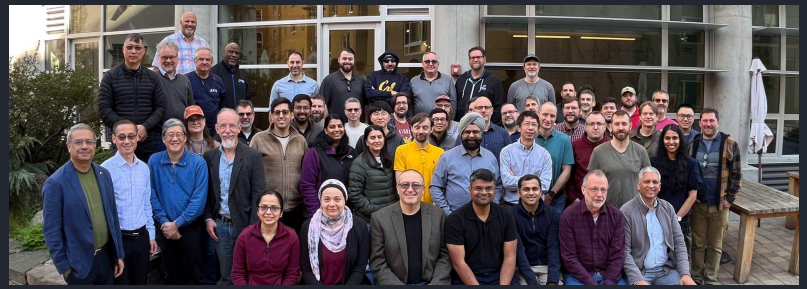
Bin Dong, Sukhada Gholba, Brooklin Gore, Shawn Kwang, David Mitchell, Samuel Oehlert, Garrett Stewart, Brendan White, Luke Baker, Edward Balas, Britt Gathright, Chin Guok, Jon-Paul Herron, John MacAuley, Scott Richmond, Chris Robb, Chris Tracy, Kesheng Wu
LBNL, ESNet, Berkeley, CA, USA

Sunday November 15, 2026

Start with a Systematic Evaluation of Business Needs

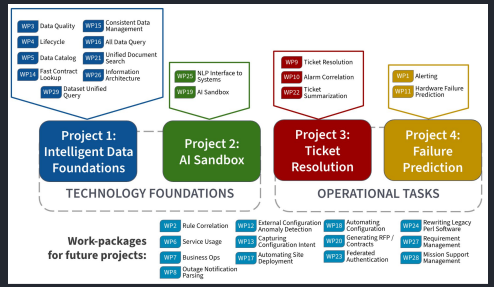


From Workshop Packages to Targeted Project Synthesis



Preparation: February 2025 Workshop

- **Broad Brainstorming:** Produced ~30 work packages spanning telemetry, config, alerting, and retrieval.
- **Published Findings:** Formal technical report established baseline community and lab requirements.
- **Challenge:** Scope was too vast for rapid execution; needed disciplined operational focus.



AI Working Group Synthesis: 4 Pillars

- **Technology Foundations:** Project 1 (Intelligent Data Foundations) + Project 2 (AI Sandbox).
- **Operational Focus:** Project 3 (Ticket Resolution) + Project 4 (Failure Prediction).
- **Strategic Path:** Directly informed narrowing down to the 6 core NOC assistant tasks in ORBIT.

Business Decision: to Focus on NOC Operations



Finding info is slow

Operators search across docs, tickets, and Slack to find procedures — often re-finding things they found last month

Tickets are novels

ServiceNow incidents accumulate pages of text. Understanding status means reading everything.

Handoffs lose context

When tickets move between shifts, the next person starts from scratch

Knowledge is everywhere (and nowhere)

Procedures in Confluence, context in Slack, history in ServiceNow — almost never together

Focus on Six Specific Tasks From the NOC

- *NOC team came up with about 20 common tasks that could use AI help*
- *Six were selected based on available data, resources, and project duration*
- *Common thread: operators already have the answers — they're just buried in five different systems.*

	Activity	Data	Description
1	Summarize incident for hand-off	ServiceNow	When a ticket is handed off to another team or escalated, parse the ticket and generate a history of the actions taken so far in a summary that is easy for the next person to process.
2	Generate ticket timeline	ServiceNow, ESDB	On demand, generate a timeline of the ticket with all actions taken and resolution information for inclusion in an after action report.
3	Suggest alarm procedure	ESDB, ESnet Wiki, Stardust, ServiceNow	Place all the NOC process documentation behind AI and use the alarm data in a given ticket to automatically populate it with relevant potential next steps from our standard processes when the ticket is created. Place this information inline in the ticket with references to the documentation cited.
4	Recommend alarm priority	ESnet Wiki, ESDB, ServiceNow	Use NOC process documentation to recommend any changes to the default alarm priority given the alarm profile and impact.
5	Propose final RFO	ESDB, ESnet Wiki, ServiceNow	Using NOC process documentation and our RFO codes, process the ticket narrative and determine a suggested RFO code and close note summary.
6	NOC chat bot	ESnet Wiki, ESDB	Place all the NOC process documentation behind AI and create a chat bot that enables a user to ask questions of our documentation along with references to the source documentation.

Traditional Automation Limits vs. What AI Delivers

Traditional Approaches (Why They Stall)

1. Better Docs & Wikis: Documentation sprawl always wins. Keyword search fails when symptoms use different terminology.

2. Custom Scripts & Rules: Brittle. Every new alert requires code changes; cannot interpret unstructured notes or messy logs.

3. More Staff & Training: Doesn't scale. Critical institutional knowledge walks out the door with operator turnover.

What AI Enables for Operations

- **Semantic Intent:** Matches intent, not exact words.
- **Cross-Source Synthesis:** Fuses tickets, docs, Slack, and topology into one coherent picture across silos.
- **Historical Pattern Matching:** Instantly links incoming alarms to past incident resolutions and work notes.
- **Speed of Thought:** Surfaces context within seconds instead of 20+ minutes of manual hunting.

Key: AI doesn't replace operator judgment—it removes data silos and enhances decision making.

How We Approached It

ORBIT

Operations Response & Business Intelligence Toolkit
(Yes, we're proud of the acronym.)

ORBIT Schematic Diagram

servicenow

NOC Ticket Test Suite

Ticket Lifecycle

(1) Agentic Actions

- Generate timeline
- Recommend RFO
- Provide bulleted view
- … …

- Suggest procedure
- Recommend priority
- … …

LangChain

(2) ORBIT Server

Dispatch request, Call RAG, Retrieve extra data, …

LibreChat

(3) Chat ⟷ Web Server



Forward requests and results to and from ORBIT Server

Data access capability

(5) Direct Data Accesses



(4) Semantic Search

Vector DB on Docs

Update, reindex, RAG, …

Data Sources

ESDB

ServiceNow

Wiki

The Data Foundation

Useful AI needs the same sources your operators use.



ESDB

Network config
& topology

via MCP



ServiceNow

Tickets & incident
management

via MCP and Stardust Semantic Search



Confluence

Docs & institutional
knowledge

via MCP and Stardust Semantic Search

Plus: Slack conversations, email threads. ServiceNow's Now Assistant only sees ServiceNow data. Our use cases need cross-source synthesis — that's why we built our own.

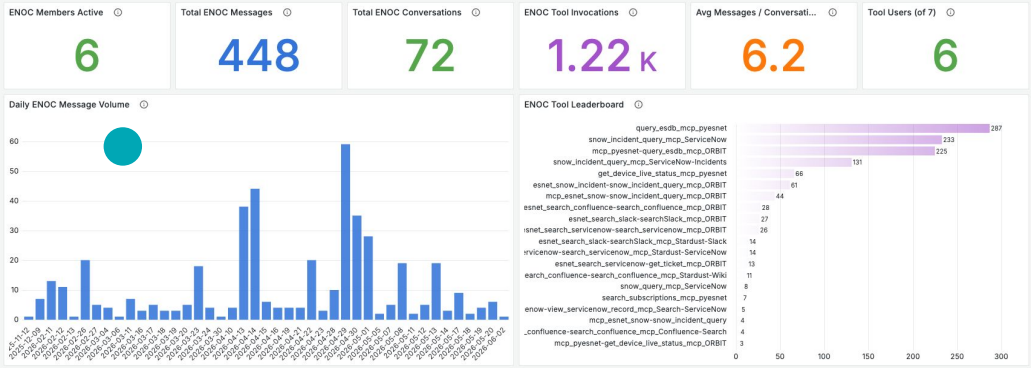
ServiceNow Integration

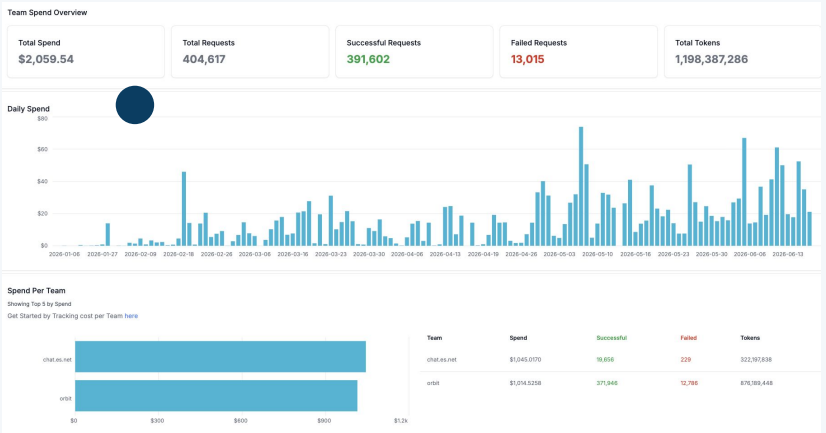
Because nobody wants another tool to check.

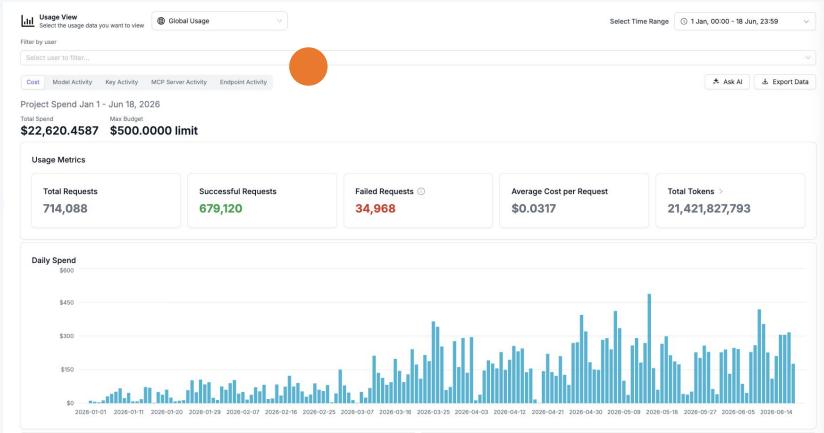
Triggered Automatically Where Appropriate	Business rules can fire on new incident tasks — no operator action needed
Results in Activity Log	AI content appears directly in tickets, visible to the whole team
Model Selection	Claude Opus via LiteLLM — configurable per action type
Operator Always in Control	Accept or reject any suggestion. It's an assistant, not a boss.
Feedback Built In	Quality tracking so we can measure and improve over time

ORBIT Uage Numbers (Jan - Jun, 2026)

- 6 out 7 NOC operators chatted with ORBIT, totaling 1,220 sessions
- ORBIT serviced 404,000 requests and consumed over 1 billion tokens
- LiteLLM model access gateway served over 21 billion tokens, mostly from coding assistants (outside of ORBIT)

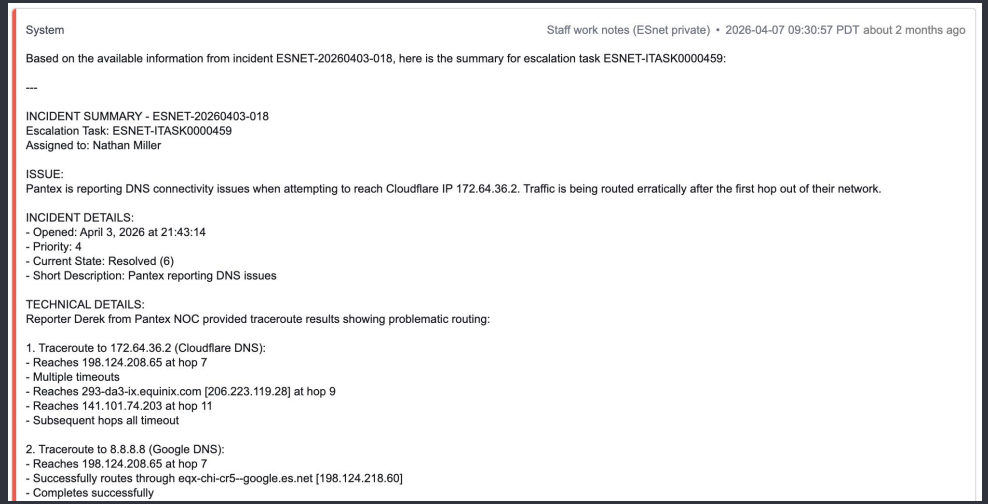






Automated Ticket Summary on Escalation

- Incidents are escalated from the NOC to the Network Engineering on-call through an Incident Task that gets associated with the main incident
  - This keeps incident housekeeping in the NOC and provides the engineer a fresh space to work in
  - Handing off means manually summarizing the ticket and the ask of engineering
- When an Incident Task is created through an escalation, ServiceNow automatically summarizes the Incident work to date and adds it to the Incident Task so the NOC operator has minimal work to do for handoff.



System — Staff work notes (ESnet private) • 2026-04-07 09:30:57 PDT about 2 months ago

Based on the available information from incident ESNET-20260403-018, here is the summary for escalation task ESNET-ITASK0000459:

INCIDENT SUMMARY - ESNET-20260403-018
Escalation Task: ESNET-ITASK0000459
Assigned to: Nathan Miller

ISSUE:
Pantex is reporting DNS connectivity issues when attempting to reach Cloudflare IP 172.64.36.2. Traffic is being routed erratically after the first hop out of their network.

INCIDENT DETAILS:
- Opened: April 3, 2026 at 21:43:14
- Priority: 4
- Current State: Resolved (6)
- Short Description: Pantex reporting DNS issues

TECHNICAL DETAILS:
Reporter Derek from Pantex NOC provided traceroute results showing problematic routing:

1. Traceroute to 172.64.36.2 (Cloudflare DNS):
- Reaches 198.124.208.65 at hop 7
- Multiple timeouts
- Reaches 293-da3-ix.equinix.com [206.223.119.28] at hop 9
- Reaches 141.101.74.203 at hop 11
- Subsequent hops all timeout

2. Traceroute to 8.8.8.8 (Google DNS):
- Reaches 198.124.208.65 at hop 7
- Successfully routes through eqx-chi-cr5--google.es.net [198.124.218.60]
- Completes successfully

ESnet

Four Key Operational Lessons from ORBIT

Insights from deploying agentic assistants and LLM infrastructure in production NOC workflows

Lesson 1: Prompts to Skills

- **Treat Prompts as Code:** Engineered prompts reduced agent actions from 10 to 4 and cut latency by 27%.
- **Organize into Skills:** Structure complex multi-step instructions into modular, versioned agent skills.

Lesson 2: Managing Stochasticity

- **Randomness is Inherent:** Semantic search and LLMs can pick different tools on different runs.
- **Mitigations:** Use strict, descriptive MCP tool names, rich guiding prompts, and specialized task models.

Lesson 3: The Best Use Cases Find You

- **Grassroots Demand:** Metrics meeting change-incident analysis was requested directly by operators.
- **Rapid Delivery:** Built as a ServiceNow UI action in ~1 hour; became one of the most valued tools.

Lesson 4: Organic Platform Adoption

- **Coding Assistants:** LiteLLM gateway served 20B+ tokens (~95%) for organic staff coding demand.
- **Chat Exploration:** 6 of 7 operators generated 1,200+ tool calls exploring ad-hoc questions.

Platform Rule: ***Build reusable infrastructure and tool APIs, then observe where operators derive actual leverage.***

The Biggest Surprise

- AI is so new, most people have difficulty envisioning use cases
  - Most people still view AI as glorified chatbots and aren't keeping up with the field
  - Moving from constraint-based thinking to infinite-possibilities thinking is a challenge without concrete examples
  - Once you give people a taste of what's possible, the ideas begin to flow rapidly
- Don't expect your initial use cases to be your strongest use cases
  - The more infrastructure you build out and the more data sources you add, the more options become available
- AI development encourages rapid iteration. Be prepared!
  - Develop organizational policies for AI usage
  - Focus heavily on security, but don't overcorrect and stifle innovation

Summary

- **Outcome 1: Support Six Core NOC Tasks**
  - **Full Delivery**: implemented all six originally scoped tasks, including Adaptive Incident Summarization (Tasks 1–2), Intelligent Resolution Guidance (Tasks 3–5), Procedure Chat (Task 6)
  - **Workflow Integration**: Embedded directly into ServiceNow, preserving human-in-the-loop decision-making; expanded to include operator-proposed "Caused-By-Change" and “Change-analysis.”
- **Outcome 2: Develop AI Staff Expertise**
  - **Organic Adoption**: 6 of 7 NOC operators adopted the chat interface (chat.es.net) to augment daily work, generating over 1,200 tool invocations.
  - **Infrastructure Mastery**: Served 21 billion tokens via the LiteLLM gateway, primarily driven by organic coding-assistant demand from staff.
  - **Engineering Growth**: Shifted team thinking from "chatbots" to agentic workflows, treating prompts as versioned, testable code.
- **Outcome 3: Performance & Usability Insights for Decision-Making**
  - **Operational Insights**: Both from quantitative and qualitative measures
  - **Strategic Recommendations**: formal adoption of the LiteLLM gateway, chat interface, agentic service, and MCP tool layer as supported organizational services.

ESnet

Interactive Chat Example:

Show me all incidents on ESnet CI ESNET-Z00863 in the past 6 months and whether any were related to scheduled maintenance. Explain the topology and any redundancy found.

🚨 Incidents in the Past 6 Months

🔍 Notable Patterns

- **October 17, 2025 was a bad day** — two incidents (ESNET-20251017-020 and ESNET-20251017-032) hit within ~75 minutes of each other. Incident -020 was a **child of a parent incident** on the `eqxdc4-cr6 ↔ wash-cr6` backbone link, suggesting the ORNL circuit was impacted by a **wider network event** on that date.
- All 6 incidents were **link flaps** — brief, self-recovering outages typically lasting less than 1 minute of actual downtime, even if the ticket remained open for hours.
- **No incidents were linked to scheduled maintenance** (`caused_by` is empty on all records, and no change requests were found associated with this circuit in the past 6 months). The circuit is also currently **not under maintenance** per ESDB.

	032					
5	**ESNET-20251223-015**	2025-12-23 16:02 UTC	2025-12-23 16:14 UTC	~12 min	P4	**"Short flap"** — resolved by Pedro Herrera
6	**ESNET-20260123-009**	2026-01-23 13:28 UTC	2026-01-23 16:22 UTC	~2h 54m	P4	**"Less than 1 minute"** flap

ESnet