Navigating the Noise: A Journey with Zeek's ICS/OT Logs

Vince Stoffer
Field CTO
Corelight

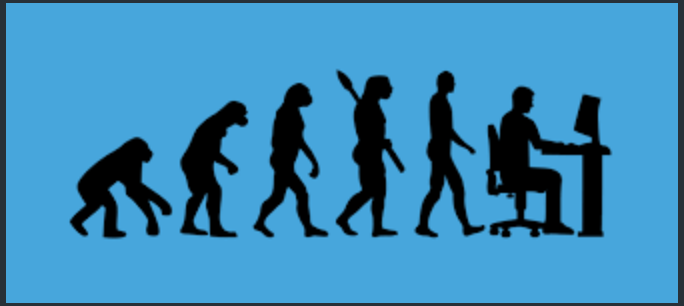
Opinions ahead

My own, not that of Corelight, Zeek or anyone else

Zeek's native ICS capabilities

- Modbus
  - https://github.com/zeek/zeek/tree/master/src/analyzer/protocol/modbus
- DNP3
  - https://github.com/zeek/zeek/tree/master/src/analyzer/protocol/dnp3
- MQTT (IoT pub/sub - arguably not ICS/OT)
  - https://github.com/zeek/zeek/tree/master/src/analyzer/protocol/mqtt

Amazon -> ICSNPP evolution



Amazon created a few analyzers in 2019?

- https://github.com/amzn/zeek-plugin-enip
- https://github.com/amzn/zeek-plugin-profinet
- https://github.com/amzn/zeek-plugin-bacnet
- https://github.com/amzn/zeek-plugin-tds
- https://github.com/amzn/zeek-plugin-s7comm

Then CISA created the ICSNPP project

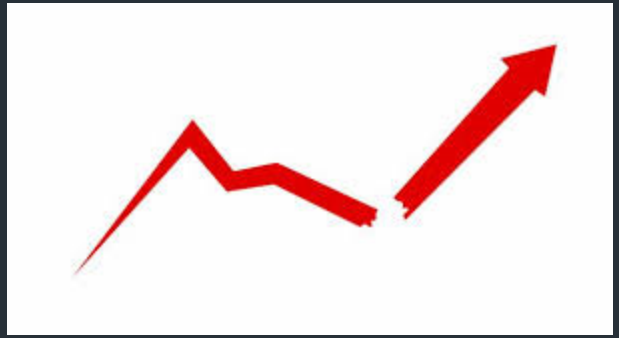
https://github.com/cisagov/ICSNPP

Which now has 14 new analyzers/script modifications for data:

bacnet bsap dnp3 enip ethercat ge-srtp genisys hart-ip modbus omron-fins opcua-binary profinet-io-cm s7comm synchrophasor

Corelight's approach

- We integrated these ICS/OT analyzers and made it easy to enable them
- BUT we started seeing huge spikes in log volume
- In our testing, we examined the logs…



Let’s talk logs

- What should a log provide?
  - Transaction details
  - Deep protocol details
  - Debugging information
  - Security Value
  - Alerts / Notices
  - Ability to be easily machine-parsed

What else?!

Zeek logs provide all of those things

Set aside notices for now (that's another talk)

For most protocol logs, the focus should be on
Security Value

Meaning: An analyst should be able to gather relevant details for the connection from a single line, without extraneous detail, and the most important components should be exposed as fields that are easily searchable
AND
Efforts should be made to reduce the noise of chatty conversations using summarization, aggregation, and other techniques

1st example: IEC-104

(IEC 60870-5-104)
ICS/SCADA protocol widely used in European energy systems

Jan Grashoefer has been working on this analyzer for Corelight

Graciously shared some test data, thanks Jan!

But normally logs just start as protocol vomit (test PCAP)

ts	uid				is_orig	send_seq	recv_seq	asdu.type_id
	asdu.num_objects	asdu.sq	asdu.cause			asdu.conf	asdu.test	
XXXXXXXXX.XXXXXX	CHhAvVGS1DHFjwGM9		T	0	0	IEC_104::InfoObjectType_C_IC_NA_1		1
	IEC_104::CauseOfTransmission_Activation		F	F	0			
XXXXXXXXX.XXXXXX	CHhAvVGS1DHFjwGM9		F	0	0	IEC_104::InfoObjectType_M_EI_NA_1		1
	IEC_104::CauseOfTransmission_Initialized		F	F	0			
XXXXXXXXX.XXXXXX	CHhAvVGS1DHFjwGM9		T	1	1	IEC_104::InfoObjectType_C_IC_NA_1		1
	IEC_104::CauseOfTransmission_Activation		F	F	0			
XXXXXXXXX.XXXXXX	CHhAvVGS1DHFjwGM9		F	1	1	IEC_104::InfoObjectType_C_IC_NA_1		1
	IEC_104::CauseOfTransmission_ActivationConfirmation			F	F			
XXXXXXXXX.XXXXXX	CHhAvVGS1DHFjwGM9		F	2	1	IEC_104::InfoObjectType_M_SP_NA_1		4
	IEC_104::CauseOfTransmission_InterrogatedGeneral			F	F			
XXXXXXXXX.XXXXXX	CHhAvVGS1DHFjwGM9		F	3	1	IEC_104::InfoObjectType_M_DP_NA_1		4
	IEC_104::CauseOfTransmission_InterrogatedGeneral			F	F			
XXXXXXXXX.XXXXXX	CHhAvVGS1DHFjwGM9		F	4	2	IEC_104::InfoObjectType_M_ST_NA_1		4
	IEC_104::CauseOfTransmission_InterrogatedGeneral			F	F			
XXXXXXXXX.XXXXXX	CHhAvVGS1DHFjwGM9		F	5	2	IEC_104::InfoObjectType_M_BO_NA_1		4
	IEC_104::CauseOfTransmission_InterrogatedGeneral			F	F			
XXXXXXXXX.XXXXXX	CHhAvVGS1DHFjwGM9		F	6	2	IEC_104::InfoObjectType_M_ME_NA_1		4
	IEC_104::CauseOfTransmission_InterrogatedGeneral			F	F			
XXXXXXXXX.XXXXXX	CHhAvVGS1DHFjwGM9		F	7	2	IEC_104::InfoObjectType_M_ME_NB_1		4
	IEC_104::CauseOfTransmission_InterrogatedGeneral			F	F			
XXXXXXXXX.XXXXXX	CHhAvVGS1DHFjwGM9		F	8	2	IEC_104::InfoObjectType_M_ME_NC_1		4
	IEC_104::CauseOfTransmission_InterrogatedGeneral			F	F			
XXXXXXXXX.XXXXXX	CHhAvVGS1DHFjwGM9		F	9	2	IEC_104::InfoObjectType_M_SP_TB_1		4
	IEC_104::CauseOfTransmission_InterrogatedGeneral			F	F			
XXXXXXXXX.XXXXXX	CHhAvVGS1DHFjwGM9		F	10	2	IEC_104::InfoObjectType_M_DP_TB_1		4
	IEC_104::CauseOfTransmission_InterrogatedGeneral			F	F			
XXXXXXXXX.XXXXXX	CHhAvVGS1DHFjwGM9		F	11	2	IEC_104::InfoObjectType_M_ST_TB_1		4
	IEC_104::CauseOfTransmission_InterrogatedGeneral			F	F			
XXXXXXXXX.XXXXXX	CHhAvVGS1DHFjwGM9		F	12	2	IEC_104::InfoObjectType_M_BO_TB_1		4
	IEC_104::CauseOfTransmission_InterrogatedGeneral			F	F			
XXXXXXXXX.XXXXXX	CHhAvVGS1DHFjwGM9		F	13	2	IEC_104::InfoObjectType_M_ME_TD_1		4
	IEC_104::CauseOfTransmission_InterrogatedGeneral			F	F			
XXXXXXXXX.XXXXXX	CHhAvVGS1DHFjwGM9		F	14	2	IEC_104::InfoObjectType_M_ME_TE_1		4
	IEC_104::CauseOfTransmission_InterrogatedGeneral			F	F			

A filtered approach (limited to specific command types, same PCAP)

XXXXXXXXX.XXXXX	CHhAvVGS1DHFjwGM9	F	5	2	IEC_104::InfoObjectType_M_BO_NA_1	4	F
	IEC_104::CauseOfTransmission_InterrogatedGeneral		F				
XXXXXXXXX.XXXXX	CHhAvVGS1DHFjwGM9	F	11	2	IEC_104::InfoObjectType_M_ST_TB_1	4	F
	IEC_104::CauseOfTransmission_InterrogatedGeneral		F				
XXXXXXXXX.XXXXX	CHhAvVGS1DHFjwGM9	F	21	2	IEC_104::InfoObjectType_M_BO_NA_1	4	F
	IEC_104::CauseOfTransmission_InterrogatedGeneral		F				
XXXXXXXXX.XXXXX	CHhAvVGS1DHFjwGM9	F	27	2	IEC_104::InfoObjectType_M_ST_TB_1	4	F
	IEC_104::CauseOfTransmission_InterrogatedGeneral		F				
XXXXXXXXX.XXXXX	CHhAvVGS1DHFjwGM9	F	50	8	IEC_104::InfoObjectType_M_ST_TB_1	1	F
	IEC_104::CauseOfTransmission_Spontaneous		F				
XXXXXXXXX.XXXXX	CHhAvVGS1DHFjwGM9	F	53	9	IEC_104::InfoObjectType_M_BO_NA_1	1	F
	IEC_104::CauseOfTransmission_Spontaneous		F				

A sessionized approach (same PCAP)

XXXXXXXXXX.XXXXXX	CHhAvVGS1DHFjwGM9 C_DC_NA_1,C_SC_NA_1	10.20.102.1	46413	10.20.100.108	2404	IEC_104::Started	16	75

2nd example: ICSNPP analyzers

Issues (Ethercat, Genesys, OPCUA, …)

- No DPD signatures or overly broad matching
  - Only port based matching
  - Ethernet w/o further signature (e.g. matching on all ARP)
- Writing out separate fields for all src, dst
  - Concerned that sessionization might miss protocol nuance
- Logging was creating massive volume problems
  - Several customers had ICS logs 2-4x the size of their conn logs!

Proposal discussion

- Need to have debugging level output for sponsor and devs
- Need to have a better concise option for security analysts

Solution:

Split logging!
debug/summary logs (default to debug off)
simple approach, extensible to multiple protocols

Corelight's VPN analyzers do something similar
- known entities example

ICS/OT 12

- BACnet
- BSAP
- DNP3
- ENIP
- Ethercat
- Genisys
- Modbus
- OPCUA Binary
- PROFINET
- S7comm
- TDS
- MQTT

Analyzers 10

- AMQP Analyzer
- GENA
- IPsec
- LDAP
- OpenVPN
- SSDP
- STUN
- Telnet+TN3270 Analyzer
- Wireguard
- SMB

Other analyzers

But writing and debugging analyzers is hard

Often the devs doing the work are NOT security analysts

This approach can be broadly applied to many new protocols

Telnet, new SMB logs, LDAP

and projects: MITRE's ACID

How to teach the (Zeek) way?

The re-write of the logs for Zeek 2 was an example of moving to security value

We need to drive expansion of the project which means focus on simplicity and broader appeal to a less network-centric security analyst (and maybe AI?)

We should use examples (like IEC-104) to demonstrate the strength of summarization and security value

Maybe part of events + training?

E.G.
Log aggregation


```
{
  "_path": "dns_agg",
  "_system_name": "bsec-ap3k-dev-aldn",
  "_write_ts": "2025-01-10T20:46:50.111498Z",
  "count": 20,
  "id.orig_h": "[redacted]0.62",
  "id.resp_p": 53,
  "qtype": 28,
  "qtype_name": "AAAA",
  "query": "collective.borg.[redacted]",
  "rejected": [
    false
  ],
  "ts": "2025-01-10T20:45:46.618876Z",
  "ts_last": "2025-01-10T20:46:39.414433Z",
  "uids": [
    "CxdKHZ1Fj9qTAgWJcd",
    "CbMqzuP5fB2CTsIDd",
    "CYLWzP2mSEIH5pDVi3",
    "CbVV3912XznBOfoXc3",
    "CMW6Ty3vFkNwMPBH1g",
    "CGWvDK3IZG8h0C2qN4",
    "Cu3eEmtiMG09VHzJi",
    "Ccbs2x4eT3t8h2c5Lg",
    "CavKUd1Z0nlTBDZGi8",
    "CSuoRR1x8nNkALcdE5",
    "CAdWMO3wQwCkQaBjeh",
    "C9NLwt4oE5DROTgOvk",
    "C321ifjDcSNg732f2",
    "CNaf8Z3iVSl6CTygBe",
    "Ci6Iec3Q7eJRFt59wi",
    "CnuziA4lz3O3I07dhd",
    "CRQsOG2pKkFYNM5s18",
    "Cyw0sn5PreX7ctnm3",
    "CsFoY4yCihLfRadmj",
    "CibhSONYGjmrdVCGh"
  ]
}
```

Summary

- Creating high value security logs is hard
- Reducing the noise of new analyzers is hard
- It’s worth it for the project and the community
- Consider using summarization, aggregation, and split logging
- How do we codify and teach the Zeek way?

Any other questions? Please ask Johanna!

Thank you!

vince@corelight.com