



cesnet
“...”

NERD: **Network Entity Reputation Database**

Václav Bartoš

CyberTIM workshop
Canterbury, UK, 28. 8. 2019



NERD – the idea

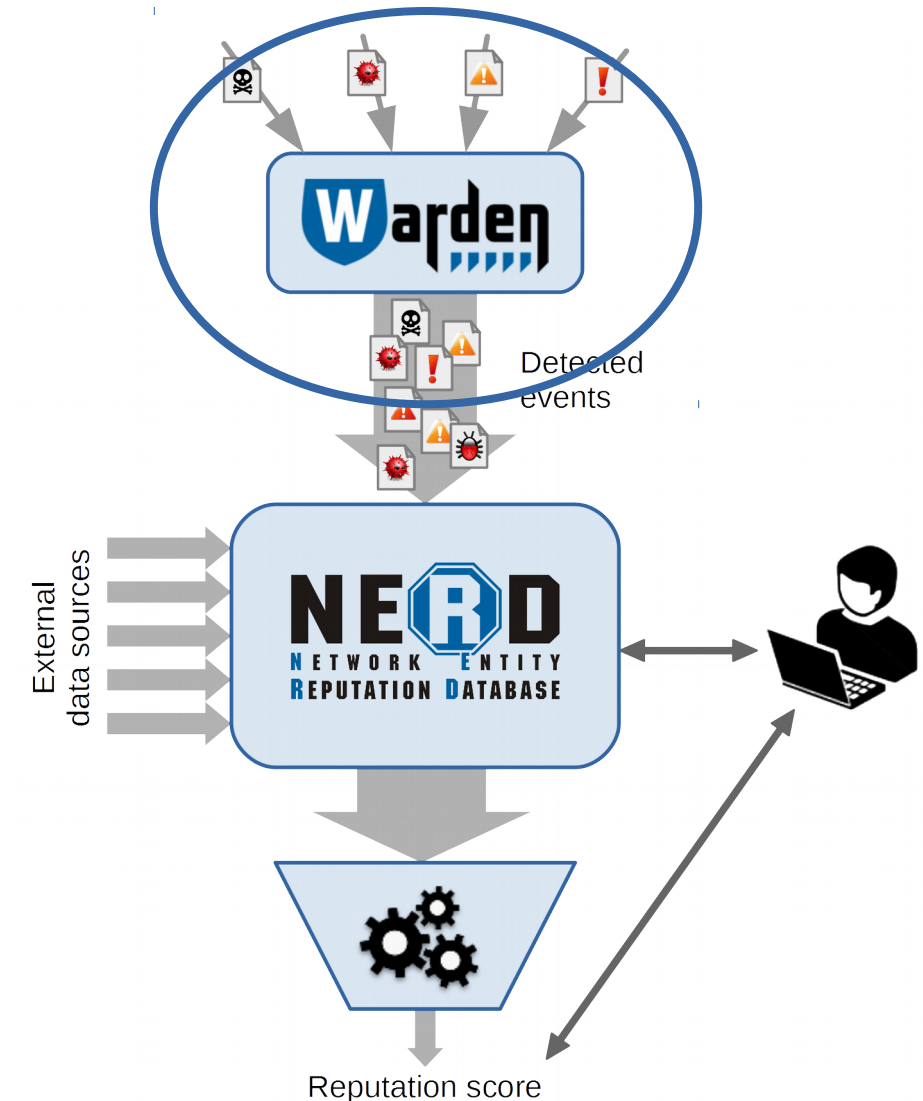
- Alerts from various detection systems (our own and from partners; SABU platform)
- Many external sources of info about malicious IPs
- ➔ Let's build a database of malicious entities (IP addresses, networks, ASNs, domains, ...)
 - More than just a blacklist – detailed info about each entity, data sources, raw alerts, ...
 - Open
 - Share data to broad cyber security community (for free)
 - All data sources and aggregation methods documented
 - Open source codes
- Goals of this presentation
 - Introduce the system
 - Describe what data are there (and plans for near future)
 - Describe how does it work

NERD

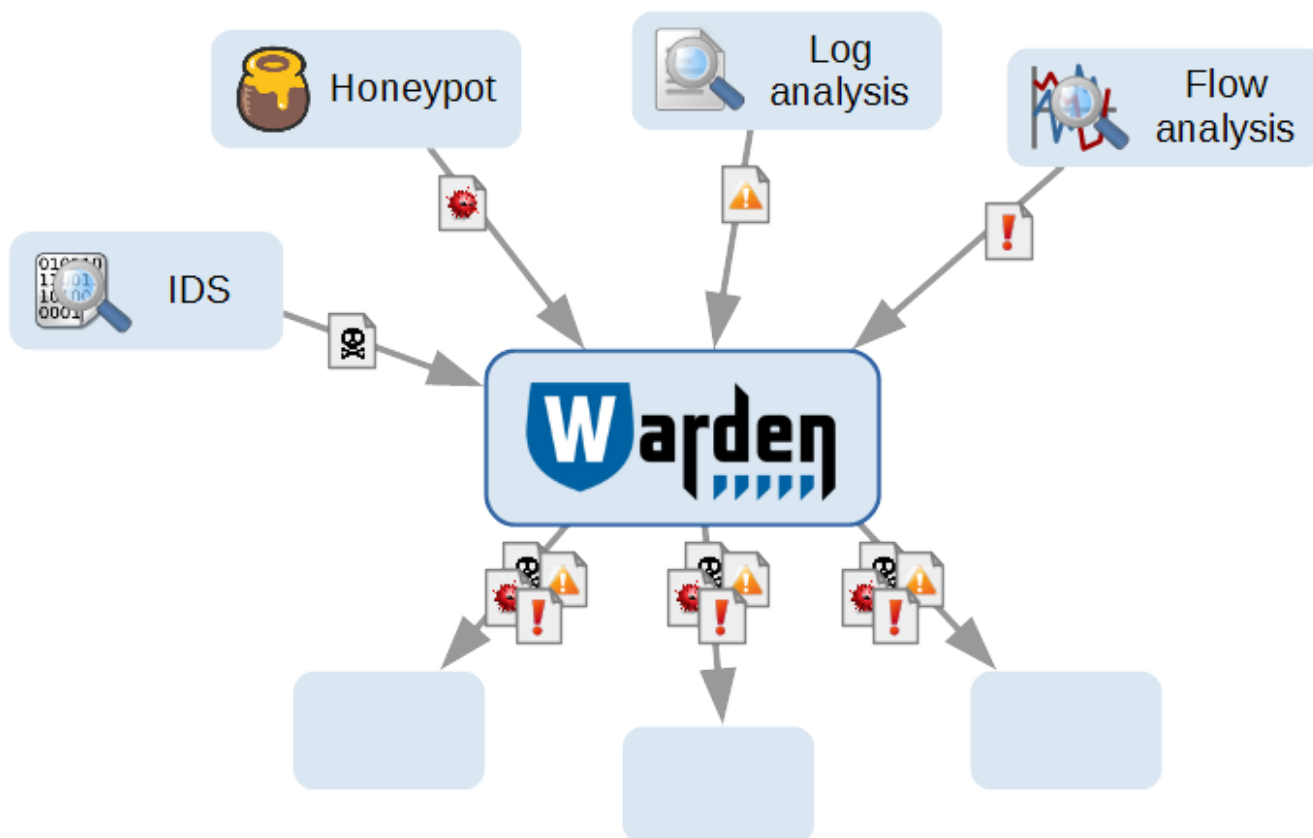
(Network Entity Reputation Database)

cesnet
.....

- Main input – alerts from Warden
 - + MISP events
 - more sources planned
- Each malicious IP has a record
- Enrichment by data from other sources
 - Hostname, ASN, geolocation, ...
 - Presence on blacklists
 - DShield
 - BGP ranking (for ASN)
 - Tor exit nodes
 - ...
- Ranking by „reputation score“
 - Summary of all the information
 - Expresses the level of threat the entity poses



- **Alert sharing system** (Developed and operated by CESNET)
 - Central hub (Warden server)
 - Detectors send alerts to the server
 - Server distributes all data to receiving clients
- **Common data format (IDEA)**
- **Open source:** <https://warden.cesnet.cz/>
- **Reciprocity** – anyone sharing their data can receive data from all others
- **Main server (community) operated by CESNET**
 - Detectors at
 - CESNET and universities in its constituency
 - One commercial ICT operator
 - Pilot deployment at PSNC and RoEduNet

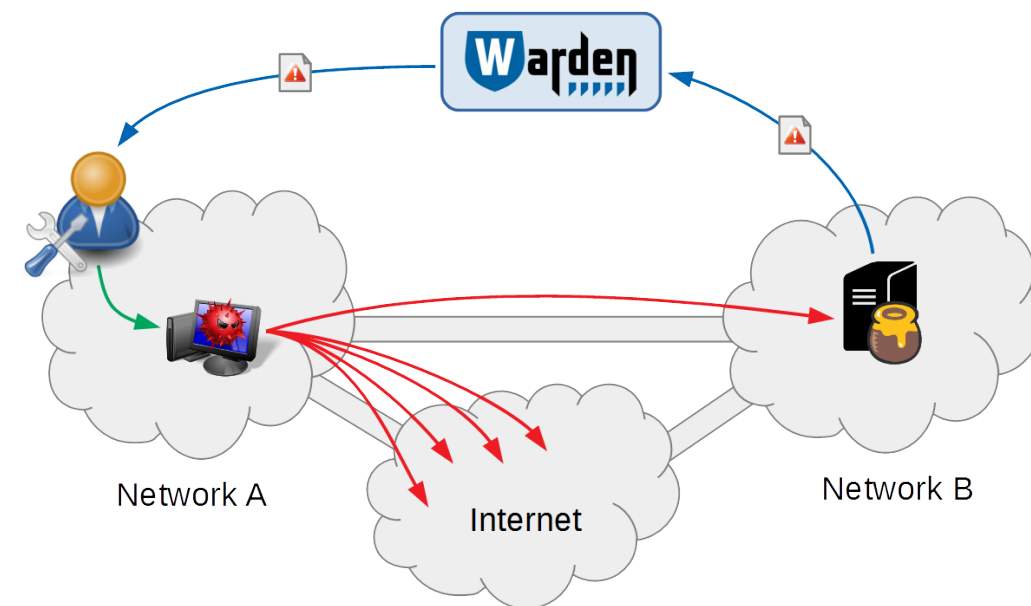


■ Current state:

- ~30 data sources
- ~2.5 mil. alerts per day (29/s)
 - Mostly port scans
 - But also bruteforce logins, exploit attempts, spam, malware communication, DDoS, etc.

■ Typical usage

- Detect & report malicious hosts in participating networks

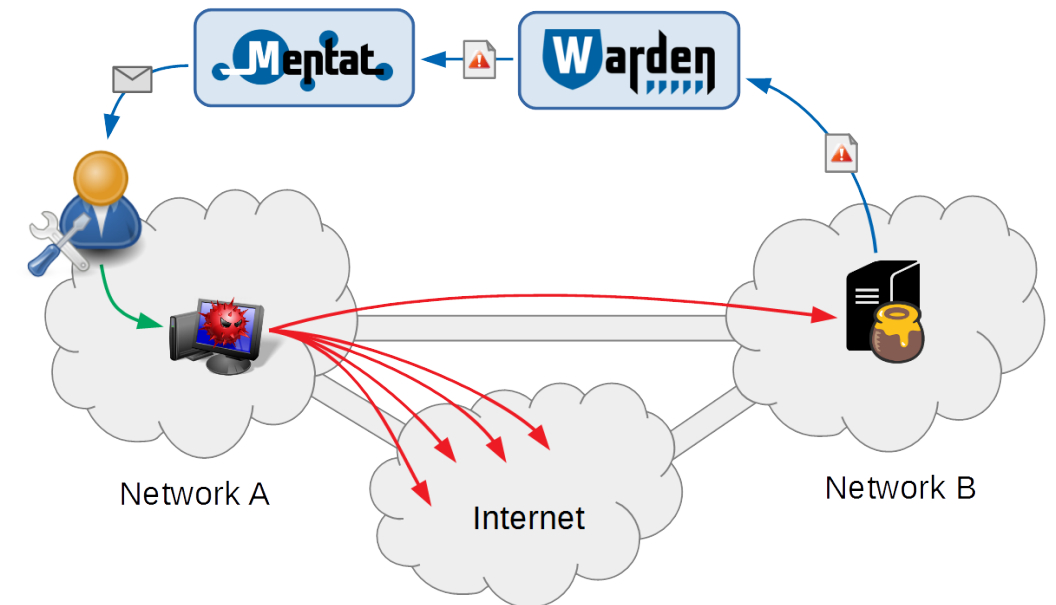


■ Current state:

- ~30 data sources
- ~2.5 mil. alerts per day (29/s)
 - Mostly port scans
 - But also bruteforce logins, exploit attempts, spam, malware communication, DDoS, etc.

■ Typical usage

- Detect & report malicious hosts in participating networks

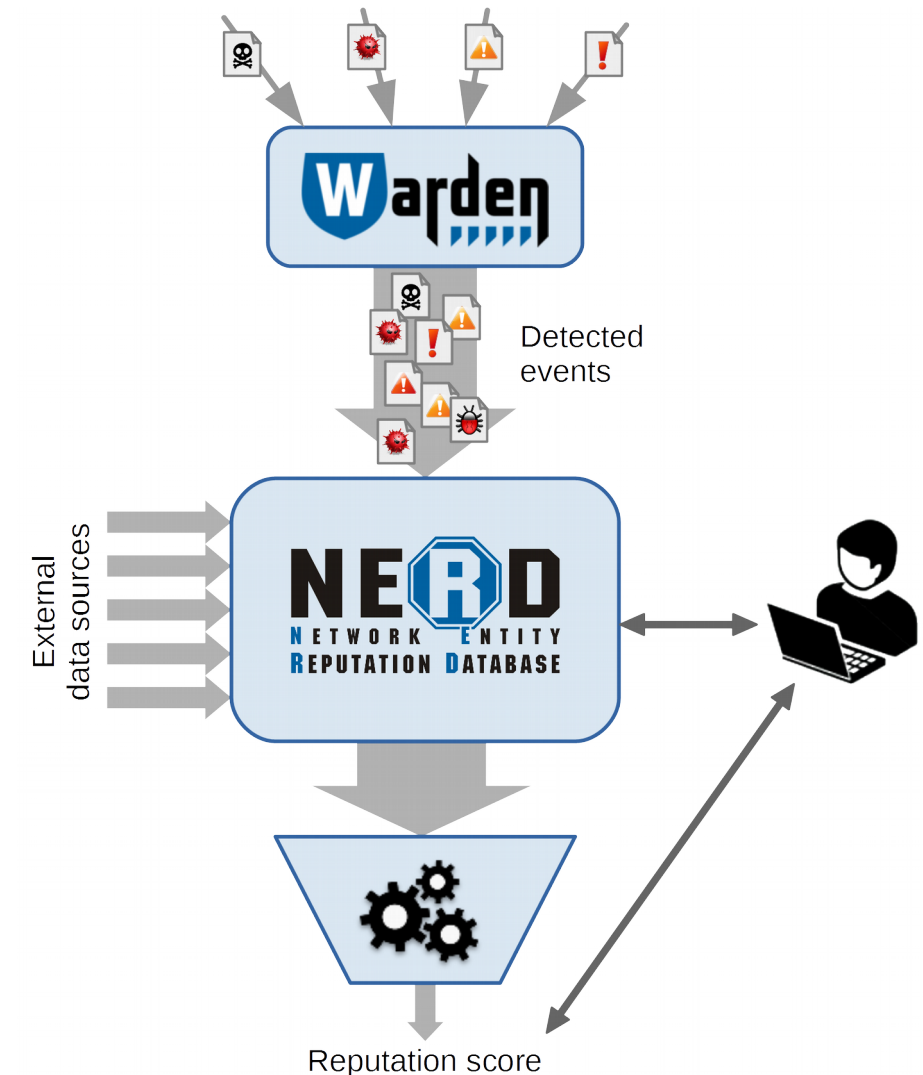


NERD

(Network Entity Reputation Database)

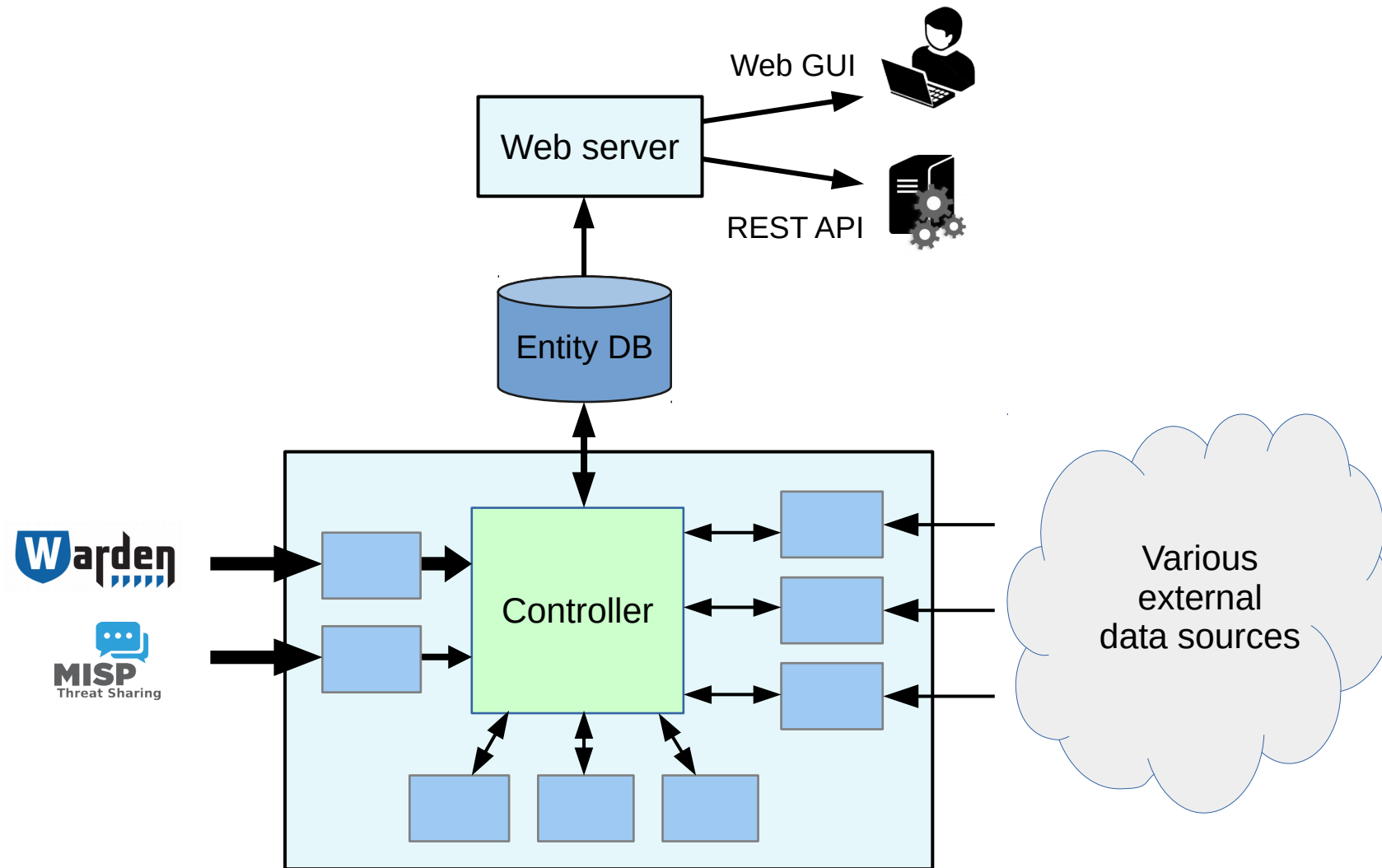
cesnet
.....

- Main input – alerts from Warden
 - + MISP events
 - more sources planned
- Each malicious IP has a record
 - Meta-data about events
- Enrichment by data from other sources
 - Hostname, ASN, geolocation, ...
 - Presence on blacklists
 - DShield
 - BGP ranking (for ASN)
 - Tor exit nodes
 - ...
- Ranking by „reputation score“
 - Summary of all the information
 - Expresses the level of threat the entity poses

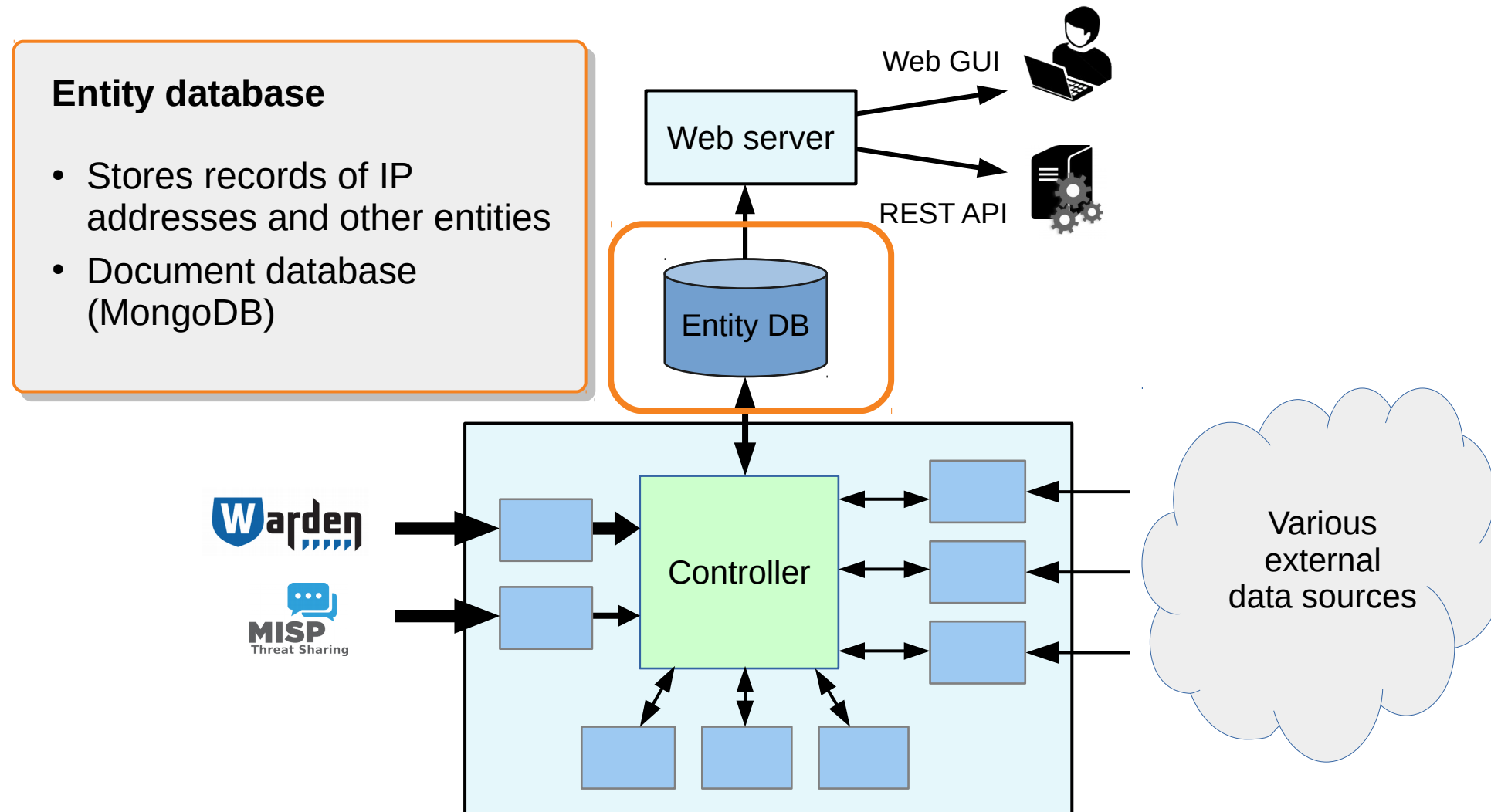


Architecture

cesnet
.....



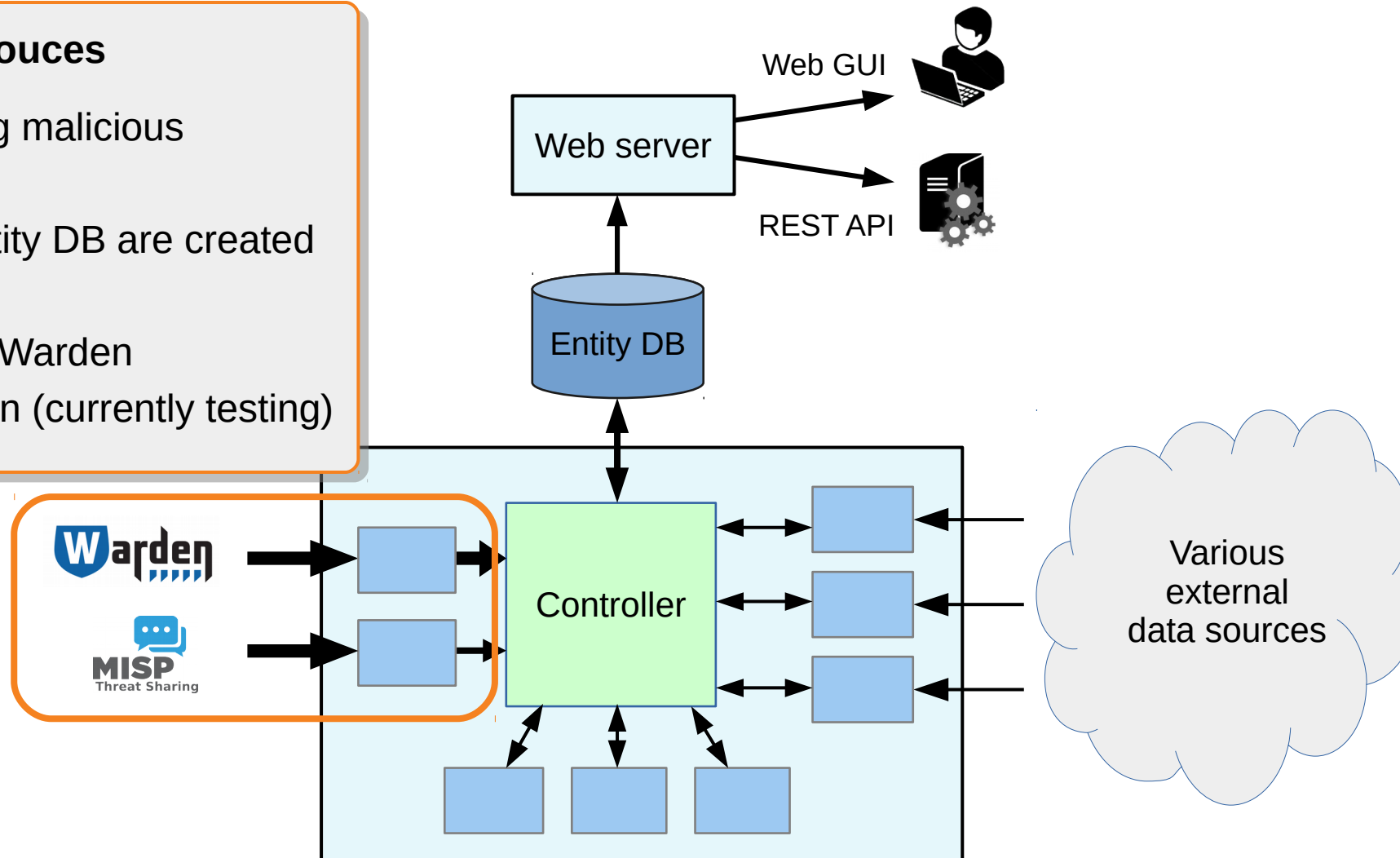
Architecture

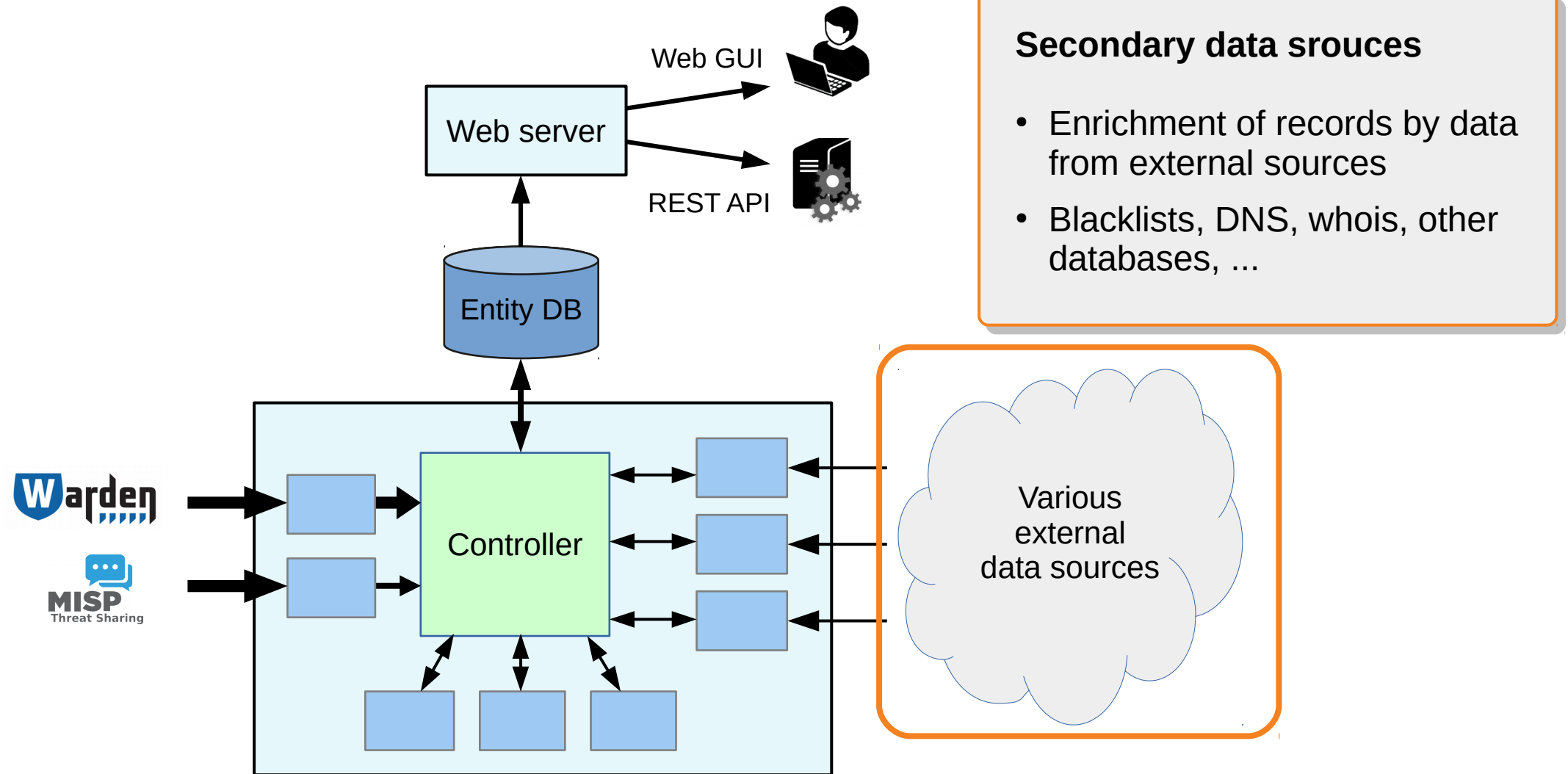


Architecture

Primary data sources

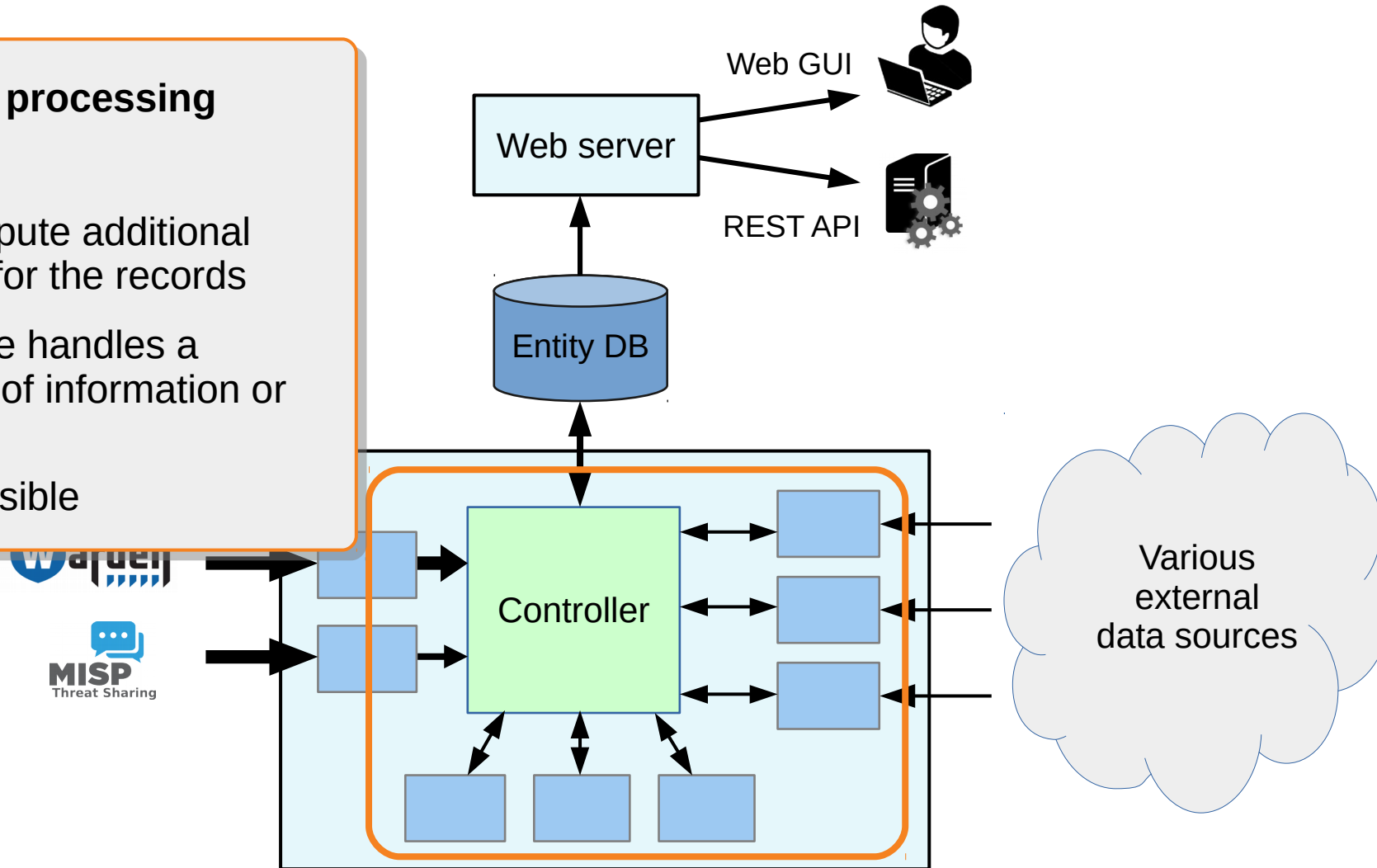
- Alerts reporting malicious addresses
- Records in Entity DB are created based on this
- Currently only Warden
- MISP very soon (currently testing)

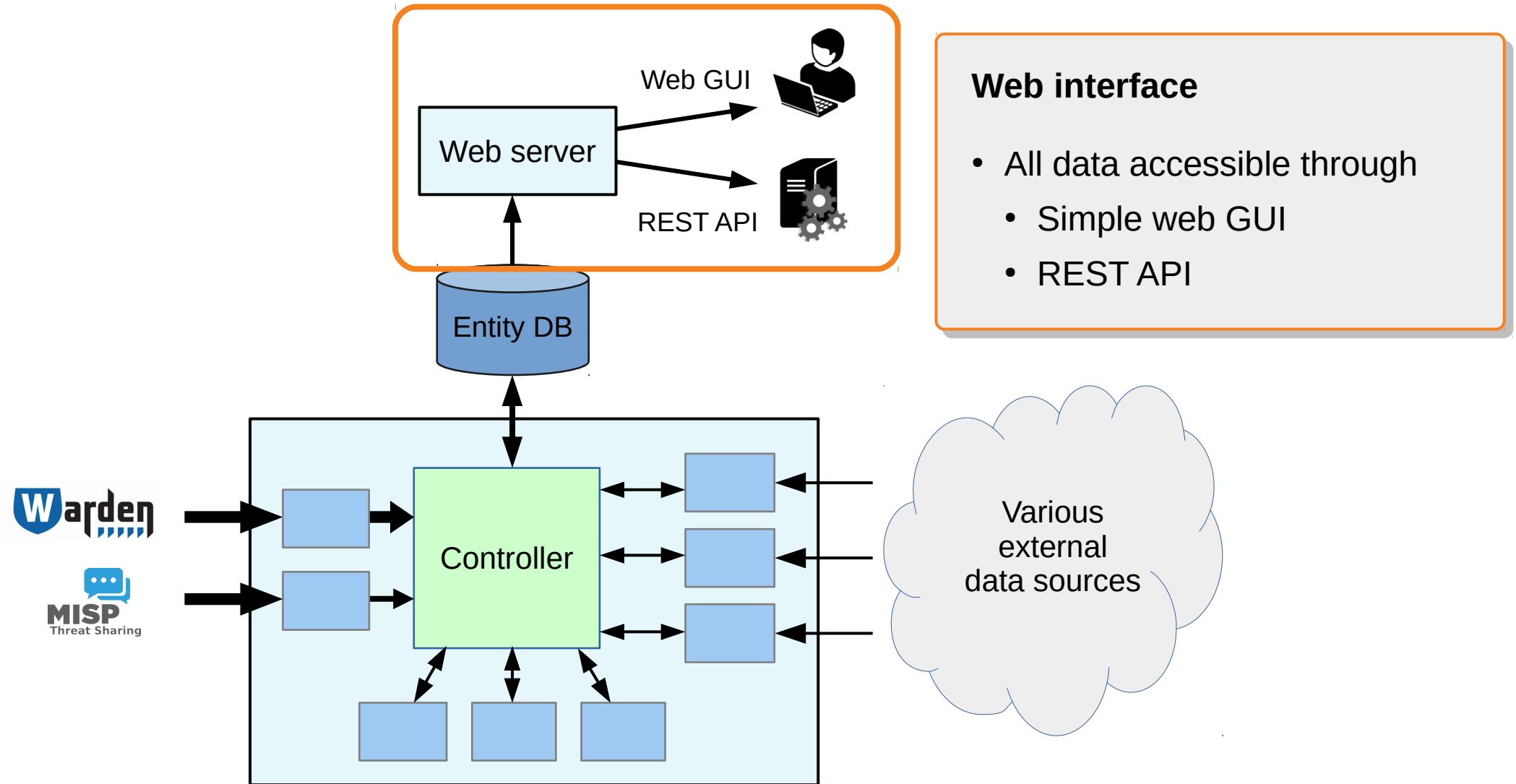




Modular data processing system

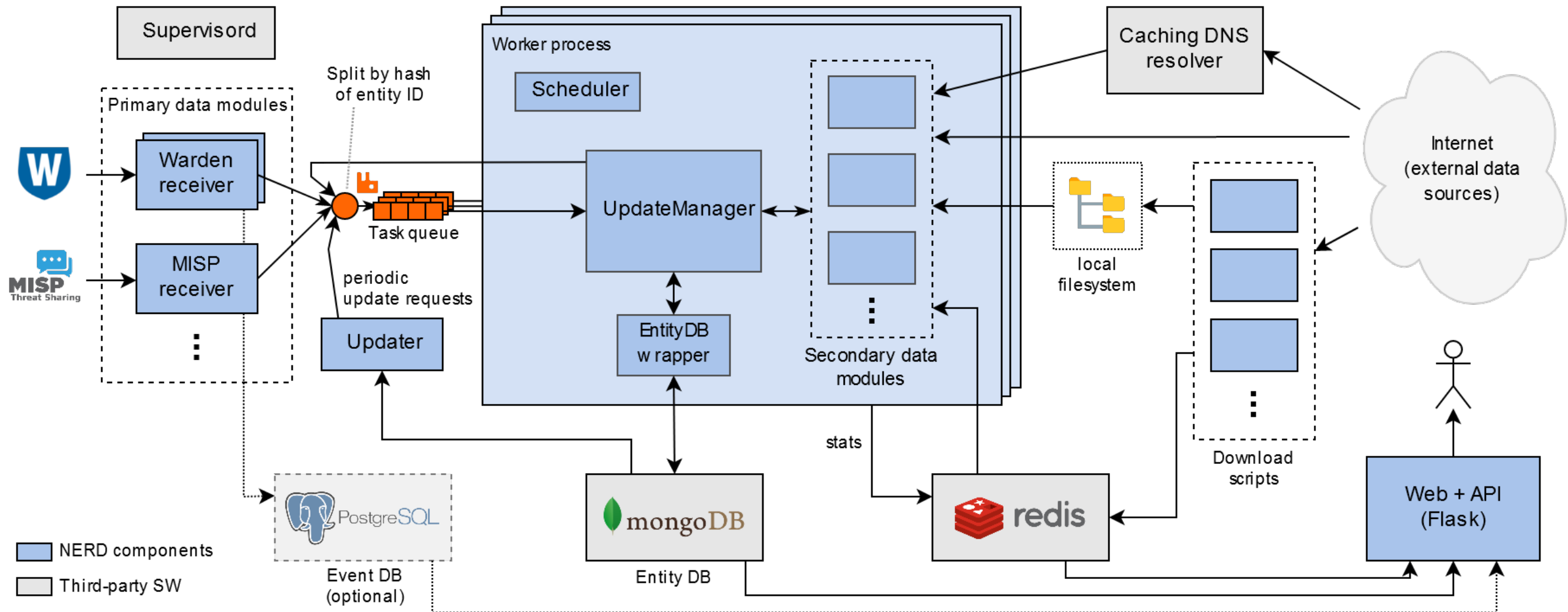
- Find or compute additional information for the records
- Each module handles a certain type of information or source
- Easily extensible





Real architecture (implementation)

cesnet
.....



<https://github.com/CESNET/NERD/wiki/Architecture>

Current data sources

(IP records)

■ **Primary** (used to create records)

- **Warden**
 - around 30 sources
(honeypots, flow analysis, external reports)
 - Alert meta-data stored
- **MISP** (CIRCL's main instance)
 - All IP attributes in events

■ **Primary – planned**

- Selected blacklists
- AlienVault OTX

- Record deleted when no related alert has been received in last 14 days

■ **Secondary** (enrichment of existing records)

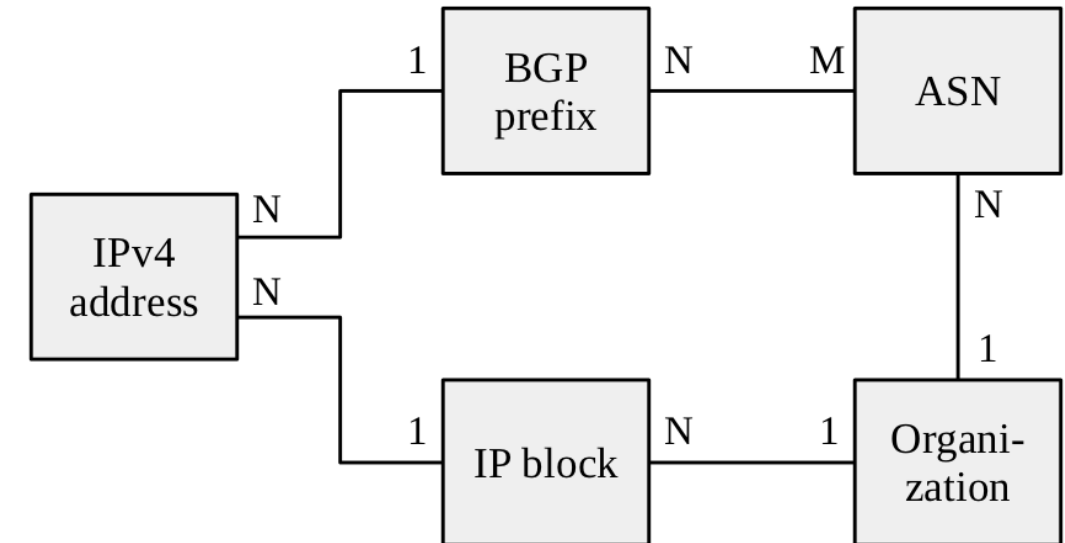
- DNS (hostname)
- Geolocation (MaxMind GeoLite2)
- Blacklists
 - 43 blacklists from 18 providers
 - Mostly IP blacklists, but also domain ones
(IP → passive DNS → hostname → query domain bl.)
 - incl. list of Tor exit nodes
- DShield
- BGP dumps + whois data
- Various tags derived from other information
- FMP score

- All the data regularly updated

Other entity types

■ Simple records also for groups of IP addresses

- BGP prefixes
- Autonomous systems (ASNs)
- IP blocks (whois)
- Organizations (whois)



■ Records contain:

- Basic data from whois databases
- ASN – ranking from external sources (CIRCL BGP ranking)

■ IPv6 – not yet, we observe very little attacks

■ Domain names, URL – not yet

Some numbers ...

- 2 primary data sources
- 10 secondary data sources
 - incl. 43 blacklists
- 18 modules
- Number of entities
(IPs removed after 14 days of inactivity)
 - IP addresses: 1 100 000
 - BGP prefixes: 122 000
 - ASN: 16 000
 - IP blocks: 38 000
 - Organisations: 16 000
- IDEA messages processed
 - 2.5 mil/day
 - 29 / s
- Updates of main database
 - New entity: 0.9/s
 - New event: 29/s
 - Regular update: 13/s
- DB size:
 - 3,5 GB

■ Who

- Limited access to public
 - Info from public sources
 - Aggregated info from Warden
 - Low rate-limits on DB queries and API
- Full access for trusted cybersec community
 - (CSIRT members, security researchers, ...)
 - Also, anyone who shares data to Warden
 - Manually approved on request

■ For free (no paid plans)

■ How

- Web interface
- REST API
- Both allow:
 - See records of individual IP, ASN, IP block, ...
 - Search IPs by given criteria

■ (planned) web browser plugin/script

Known IP addresses

IP prefix & Hostname suffix & Country code & ASN

Event category

▼

Blacklist

▼

 & Tag

▼

 Min tag confidence

Sort by Reputation score

▼

☐ Ascending

Max number of addresses

Odeslat dotaz

Results (≥20)

IP address	Hostname	ASN	Country	Events	Rep.	Other properties	Time added	Last event	Links
193.32.163.89	--	AS201912	RO	7442 7 2	0.992	Scanner Login attempts spamhaus-sbl bruteforceblocker uceprotect spamhaus-xbl-cbl blocklist_de-ssh	2019-04-08 10:52:08	2019-04-23 16:20:00	
88.214.26.47	hostby.fcloud.biz	AS201912	DE	28463 8 4	0.991	Scanner Login attempts	2018-12-28 11:50:01	2019-04-23 16:22:42	
202.29.57.103	--	AS38265	TH	42932 7 1	0.986	Scanner uceprotect spamhaus-sbl	2019-02-02 16:20:05	2019-04-23 16:26:29	
188.246.224.98	--	AS49505	RU	12977 7 2	0.985	Scanner uceprotect	2019-04-11 01:23:43	2019-04-23 16:28:29	
82.202.247.48	--	AS49505	RU	12679 6 2	0.984	Scanner uceprotect	2019-04-10 22:45:47	2019-04-23 16:28:29	
140.143.77.85	--	AS45090	CN	3264 7 4	0.984	Scanner Login attempts spamhaus-pbl uceprotect	2019-01-13 11:54:49	2019-04-23 15:55:00	
89.248.168.51	security.criminalip.com	AS29073 AS202425	NL	12100 6 2	0.973	Scanner uceprotect	2019-04-05 09:29:58	2019-04-23 16:19:29	
96.69.146.41	96-69-146-41-static.hfc.comcastbusiness.net	AS7922	US	10035 7 2	0.970	IP in hostname Static IP Scanner spamhaus-xbl-cbl uceprotect	2019-03-04 14:06:06	2019-04-23 16:25:39	
184.105.247.238	scan-13k.shadowserver.org		US	28639 5 2	0.969	Research scanner Scanner uceprotect	2016-07-12 02:39:45	2019-04-23 16:28:06	
184.105.247.247	scan-14m.shadowserver.org		US	28473 5 2	0.969	Research scanner Scanner uceprotect	2016-07-12 06:34:48	2019-04-23 16:24:01	
216.243.31.2	--		US	58611 5 2	0.969	Scanner uceprotect	2016-07-11 17:57:35	2019-04-23 16:28:31	
184.105.247.252	scan-15n.shadowserver.org		US	30044 5 2	0.969	Research scanner Scanner uceprotect	2016-07-12 03:26:11	2019-04-23 16:20:00	
184.105.139.67	scan-01.shadowserver.org		US	12170 5 2	0.969	Research scanner Scanner uceprotect	2017-04-21 13:04:13	2019-04-23 16:25:21	
94.102.49.193	cloud.census.shodan.io		NL	28251 5 2	0.969	Research scanner Scanner uceprotect	2016-07-11 17:56:31	2019-04-23 16:14:48	
115.238.245.8	--	AS4134	CN	7208 5 2	0.969	Scanner uceprotect	2017-12-11 16:15:06	2019-04-23 15:44:28	
178.73.215.171	--	AS42708	SE	34826 5 1	0.969	Scanner uceprotect	2018-01-09 18:39:07	2019-04-23 16:26:40	
110.249.212.46	--	AS4837	CN	7226 5 2	0.969	Scanner spamhaus-pbl uceprotect spamhaus-xbl-cbl spamhaus-sbl	2018-08-22 07:36:47	2019-04-23 15:25:00	
14.225.3.37	--	AS45899	VN	72405 6 2	0.969	Scanner uceprotect	2018-10-03 21:46:55	2019-04-23 16:28:29	
185.208.208.144	--	AS205406		50756 5 1	0.969	Scanner spamhaus-sbl uceprotect	2018-10-30 13:57:16	2019-04-23 16:26:00	
185.176.26.3	--	AS44901 AS57271		48029 5 1	0.969	Scanner uceprotect spamhaus-sbl spamhaus-drop	2018-11-26 06:45:49	2019-04-23 16:20:00	

https://nerd.cesnet.cz/nerd/



IP address

IP address Odeslat dotaz

.970 **96.69.146.41** 96-69-146-41-static.hfc.comcastbusiness.net

Search at other sites: [🔗](#) [🌐](#) [📄](#) [🌍](#) [🔍](#) [🔖](#)

Origin AS
[AS7922](#) - COMCAST-7922

BGP Prefix
[96.64.0.0/11](#)

IP blacklists
[spamhaus-xbl-cbl](#) [uceprotect](#)

- events (10037)**
- 2019-03-04**
ReconScanning (node.7257ac): 1
 - 2019-03-05**
ReconScanning (node.7257ac): 2
 - 2019-03-07**
ReconScanning (node.7257ac): 1
 - 2019-03-17**
ReconScanning (node.7257ac): 3
AttemptLogin (node.72edc5): 1
 - 2019-03-18**
ReconScanning (node.7257ac): 4
 - 2019-03-19**
ReconScanning (node.7257ac): 10
 - 2019-03-20**
ReconScanning (node.7257ac): 8
AttemptLogin (node.72edc5): 4
 - 2019-03-21**
AttemptLogin (node.63072d): 8

Shodan [\(more info\)](#)

Ports: [23](#) [81](#)

Tags: [scanner](#)

Passive DNS [\(?\)](#)

-- no records --

fmp
{'general': 0.8833602666854858}

geo
[🇺🇸 United States](#), Denver
[🕒 America/Denver](#)

hostname
96-69-146-41-static.hfc.comcastbusiness.net

hostname_class
['ip_in_hostname', 'static']

Address block ('inetnum' or 'NetRange' in whois database)
[96.64.0.0 - 96.95.255.255](#)

rep
0.9702380952380952

tags
[IP in hostname](#) [Static IP](#) [Scanner](#)

<https://nerd.cesnet.cz/nerd/>

Reputation score

■ Idea:

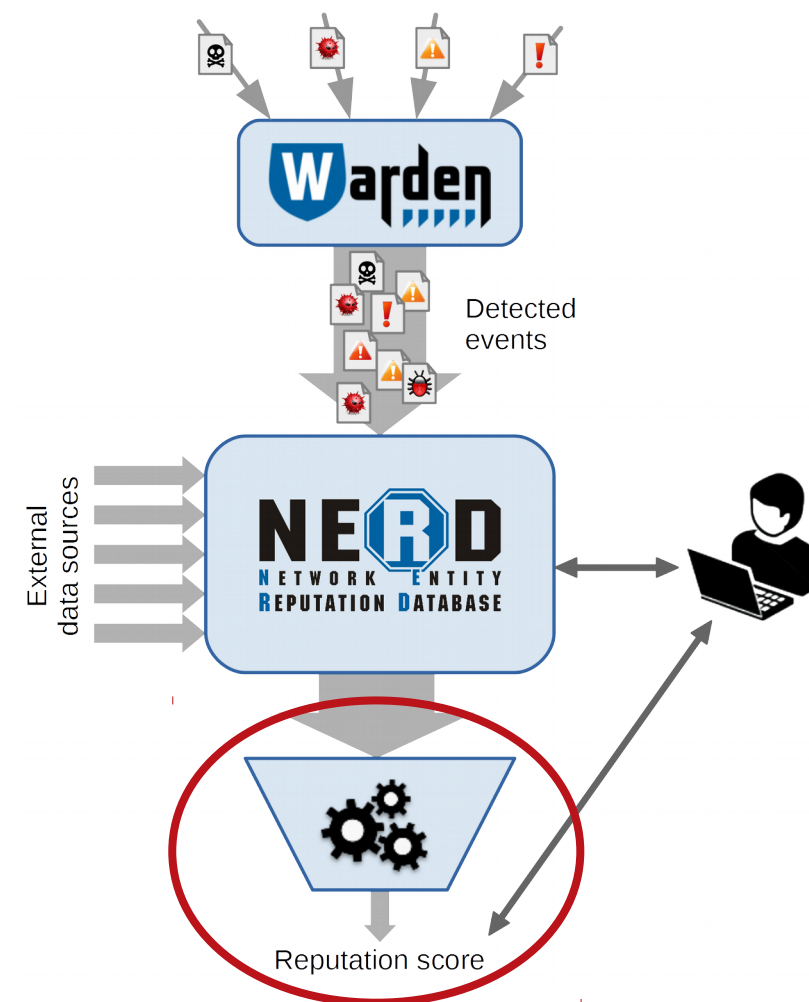
- All the info summarized into a single number (or a small set of numbers)
- "Level of threat the IP address poses"

■ To allow ...

- quick overview
- ranking
- blacklist creation (top-n)

■ **Probability that the IP address will attack in a near future.**

- Prediction based on previous alerts and other data
- Machine learning



Reputation score

■ Idea:

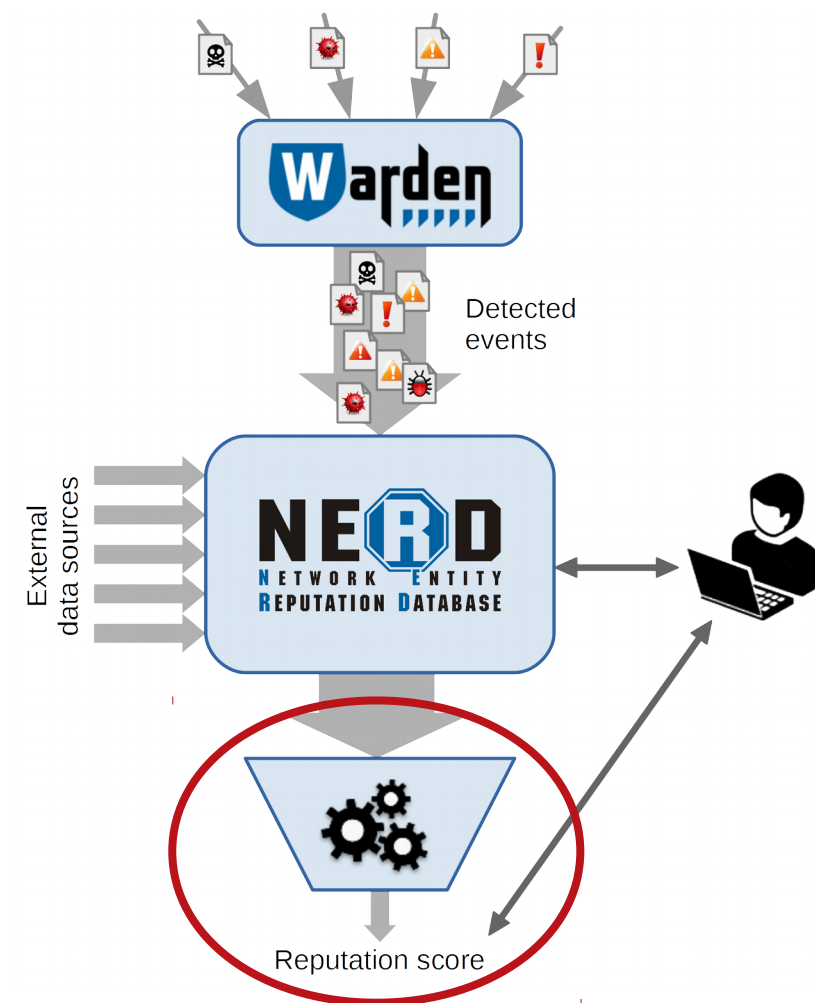
- All the info summarized into a single number (or a small set of numbers)
- "Level of threat the IP address poses"

■ To allow ...

- quick overview
- ranking
- blacklist creation (top-n)

■ **Probability that the IP address will ~~attack~~ be reported in a near future.**

- Prediction based on previous alerts and other data
- Machine learning



Reputation score

■ Idea:

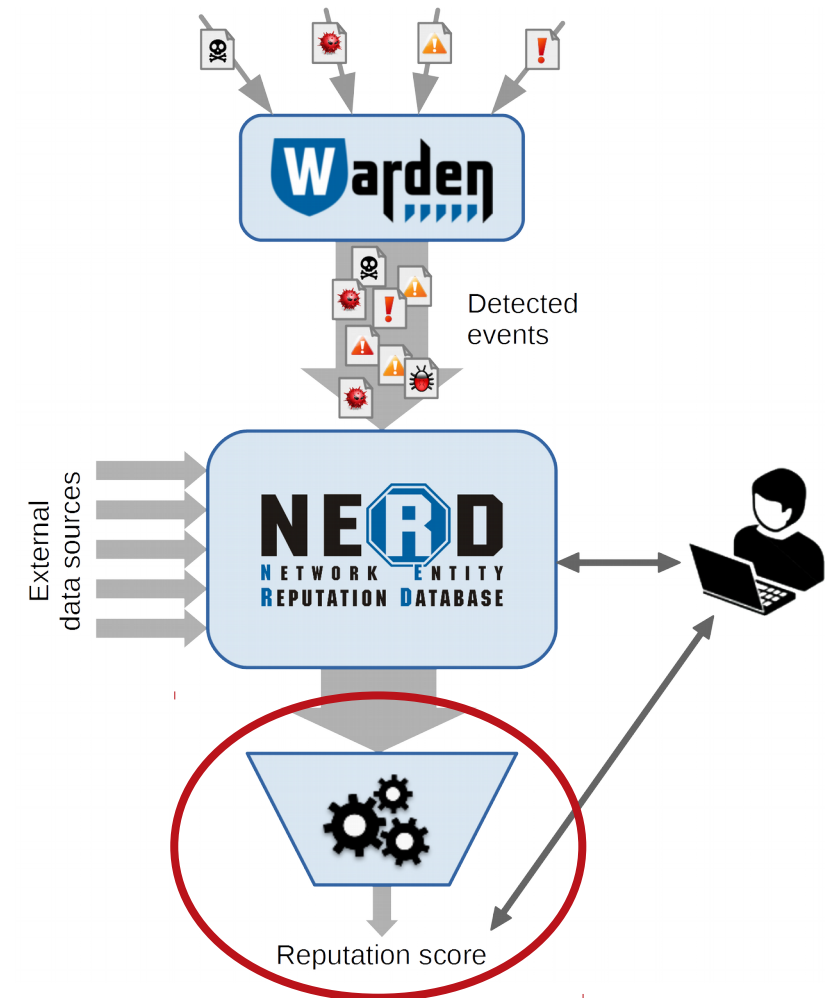
- All the info summarized into a single number (or a small set of numbers)
- "Level of threat the IP address poses"

■ To allow ...

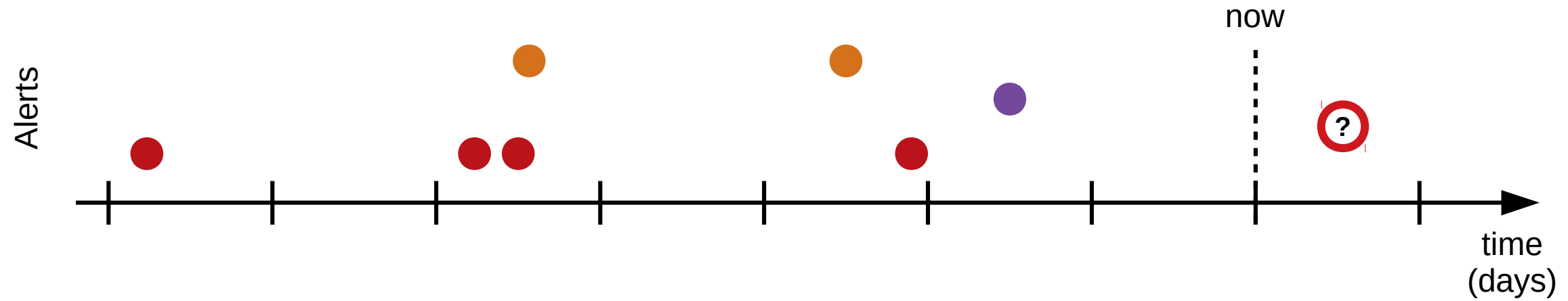
- quick overview
- ranking
- blacklist creation (top-n)

■ **Probability that the IP address will ~~be reported~~ attack in a near future.**

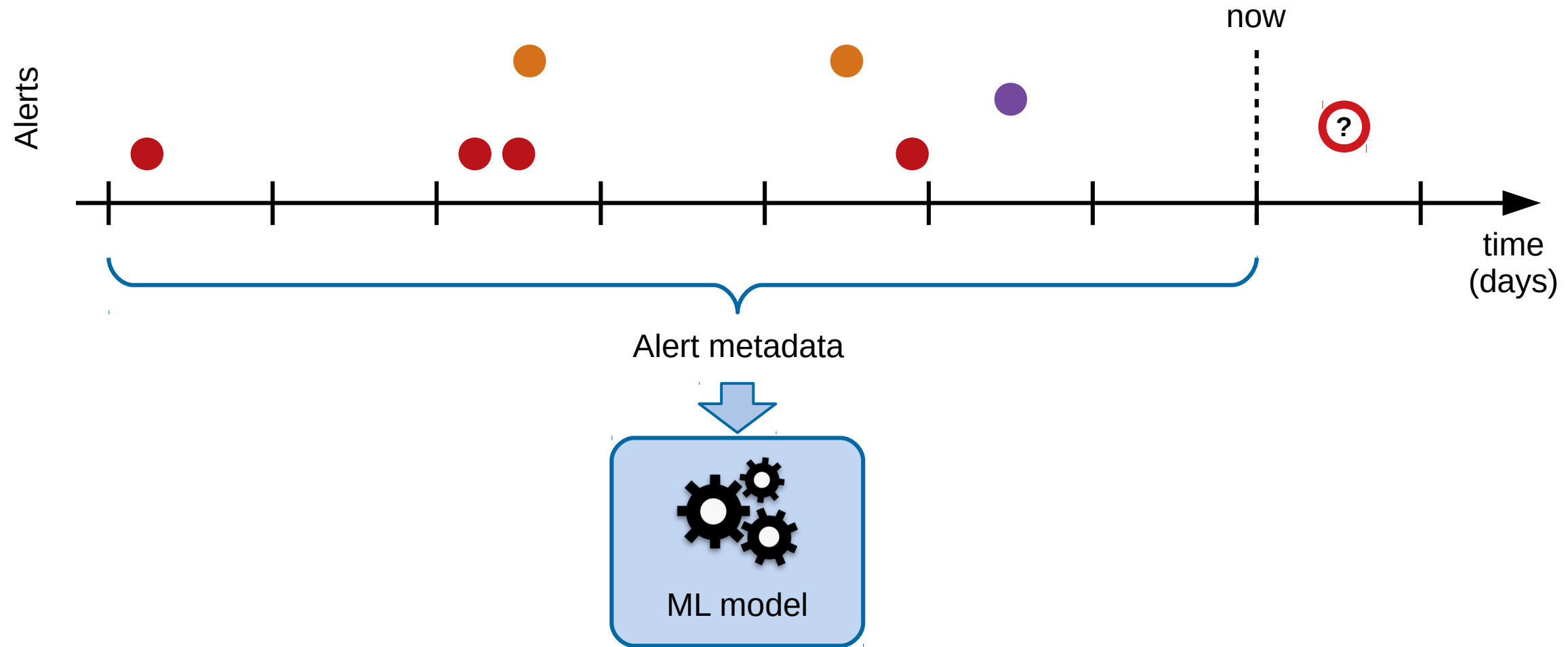
- Prediction based on previous alerts and other data **next 24 hours**
- Machine learning



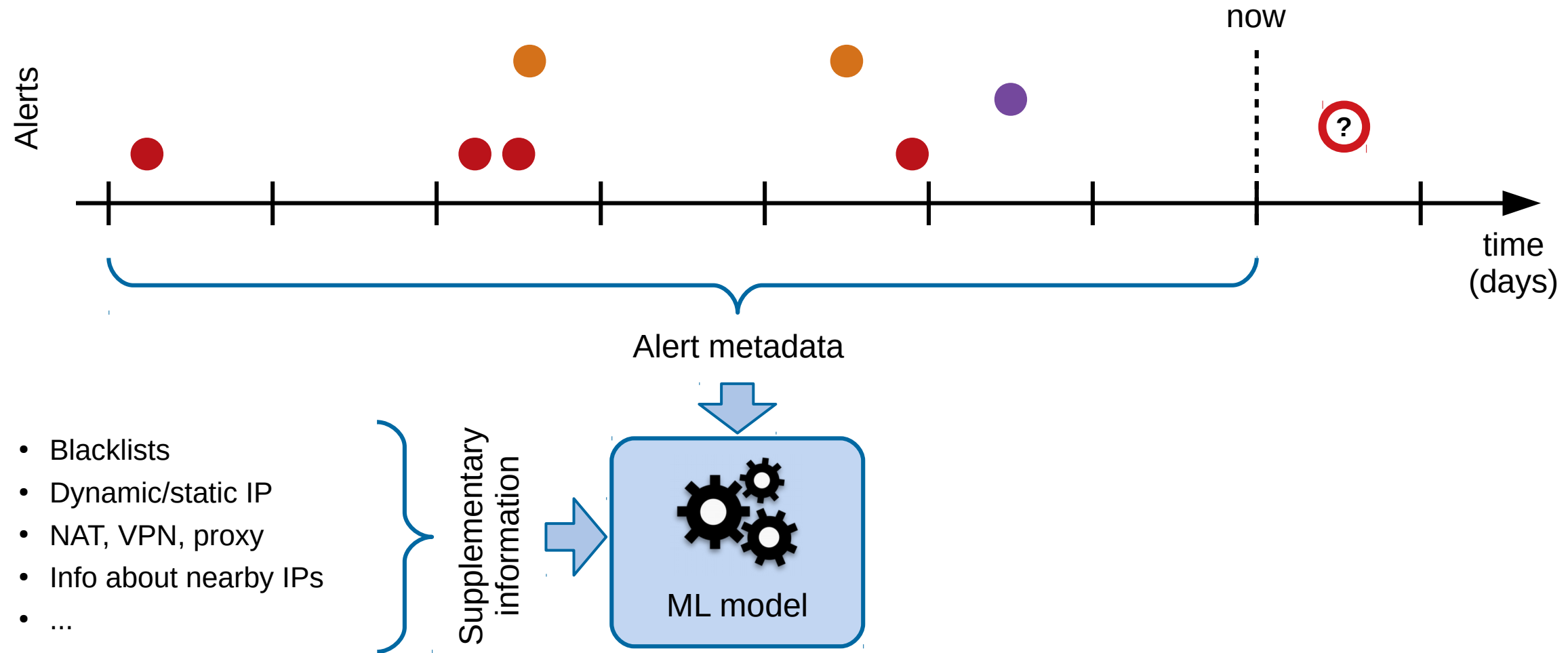
Reputation score - idea



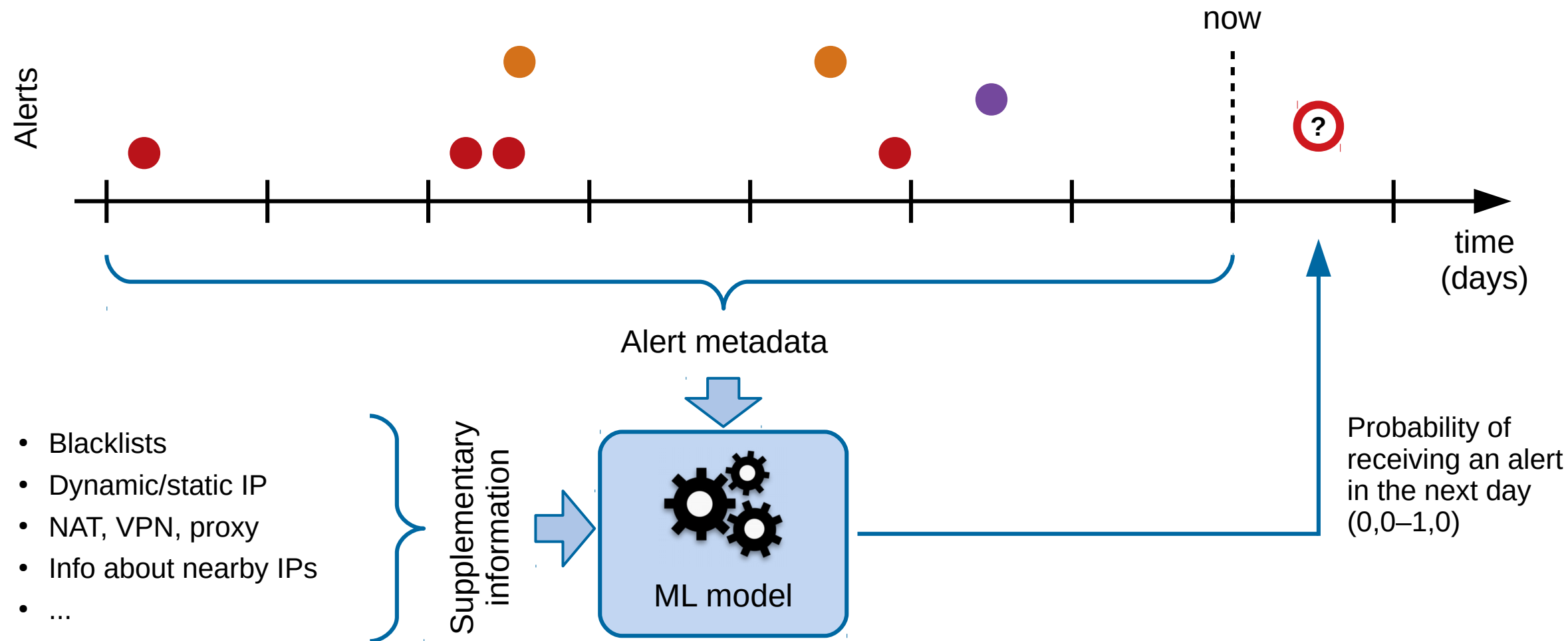
Reputation score - idea



Reputation score - idea



Reputation score - idea



- Prediction method published in:

V. Bartoš, et. al.: Network entity characterization and attack prediction, *Future Generation Computer Systems*, vol. 97, Elsevier, 2019.

- Supervised learning – use historical data

- Data:

- 3 months of alerts from Warden
 - 160 mil. alerts
 - 10 mil. IP addresses
- 58 features for each IP address

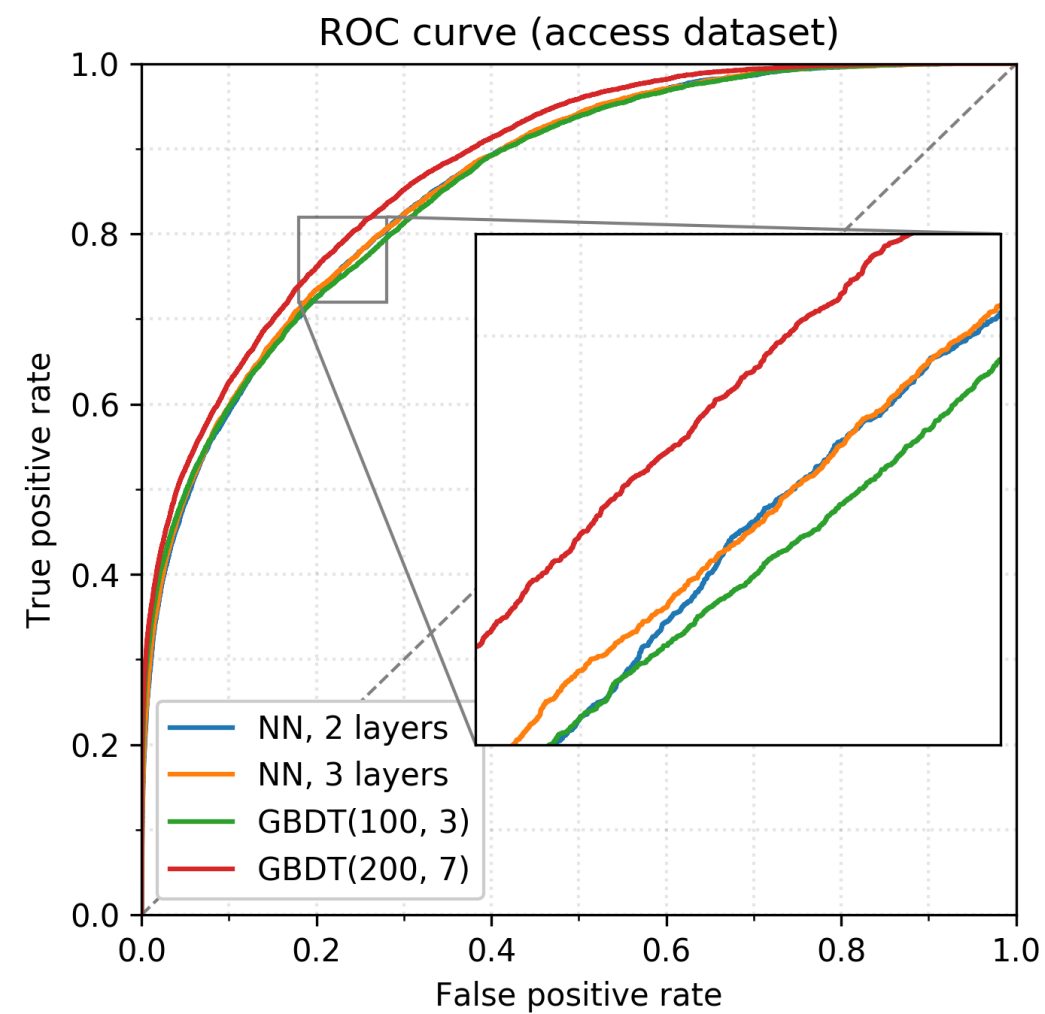
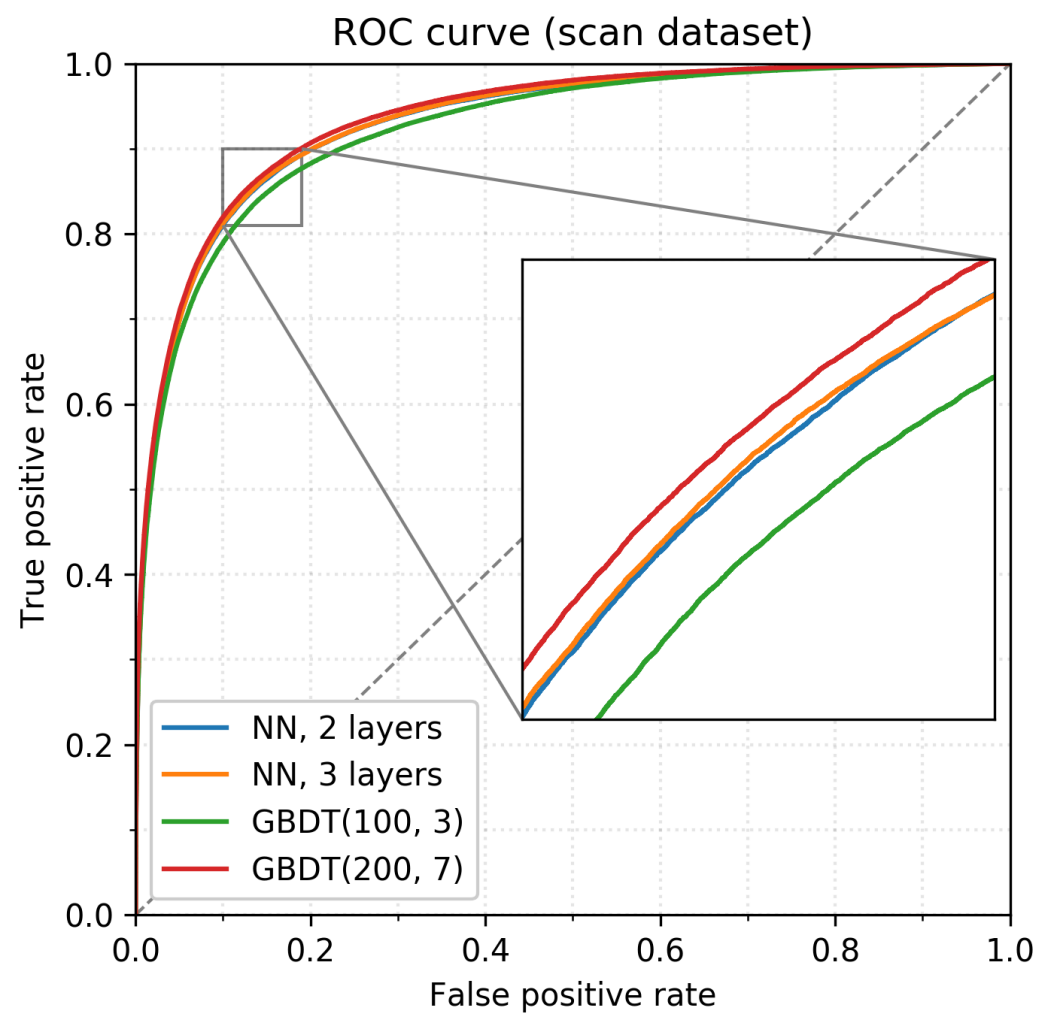
- Two alert categories predicted

- scanning
- Unauthorized access attempt (login, exploit)

- Models

- Neural networks (2-3 hidden layers)
- Decision trees
(Gradient boosted decision trees, xgBoost)

Results



FMP score – implementation in NERD

- Prediction method implemented in NERD (in slightly simpler version for now)
- FMP score computed for each IP address every day
- Plan – include severity/intensity of attacks into the score

Results (≥20)

IP address	Hostname	ASN	Country	Events	Rep.	Other properties
81.22.45.185	--	AS49505	RU	29423 5 2	0.894	Scanner uceprotect spamhaus-sbl spamhaus-drop
172.104.76.217	scan-95.security.ipip.net	AS63949	US	49630 5 2	0.875	Scanner uceprotect
198.108.67.101	scratch-02.sjf.corp.censys.io	AS237	US	52990 5 2	0.937	Scanner uceprotect
122.228.19.79	--	AS134771	CN	43525 5 2	0.810	Scanner uceprotect spamhaus-xbl-cbl spamhaus-sbl-css
198.20.70.114	census3.shodan.io	AS32475	US	66636 4 2	0.937	Research scanner Scanner uceprotect spamhaus-xbl-cbl
94.241.129.44	--	AS16200 AS56466	IR	2192 2 2	0.331	Scanner spamhaus-xbl-cbl spamhaus-pbl-isp
82.208.187.124	romsound-marasesti74-fo.b.astral.ro	AS6830	RO	15155 3 2	0.856	Scanner Login attempts uceprotect spamhaus-xbl-cbl spam
96.76.250.161	96-76-250-161-static.hfc.comcastbusiness.net	AS7922	US	14561 5 2	0.943	P in hostname Static IP Scanner uceprotect spamhaus-xbl
94.241.129.27	--	AS16200 AS56466	IR	950 2 2	0.245	Scanner spamhaus-pbl-isp spamhaus-xbl-cbl
165.227.179.69	--	AS14061	US	10500 6 2	0.937	Scanner
172.93.102.99	--	AS20473	US	1318 5 2	0.355	Scanner uceprotect
139.162.120.76	scan.security.ipip.net	AS63949	US	48820 4 2	0.875	Scanner uceprotect
119.3.215.213	ecs-119-3-215-213.compute.hwclouds-dns.com	AS55990	CN	3349 3 2	0.575	P in hostname Scanner uceprotect
134.209.164.141	--	AS14061	US	1932 6 2	0.446	Scanner spamhaus-xbl-cbl uceprotect

- Web interface:
 - <https://nerd.cesnet.cz/nerd/>
- Source codes:
 - <https://github.com/CESNET/NERD>
 - Docs: <https://github.com/CESNET/NERD/wiki>
- FMP paper:
 - Bartoš, et. al.: "Network entity characterization and attack prediction", FGCS journal, 2019
- The system is constantly being improved ...
 - Use it
 - Provide feedback & feature requests
- If possible, share data
 - Join Warden – share alerts from your IDS, ADS, honeypots, etc.
 - Or provide any other security-related data about IP addresses or other network entities

Contact:
nerd@cesnet.cz

Thank you for you attention

Questions?

Contact:

Václav Bartoš <bartos@cesnet.cz>