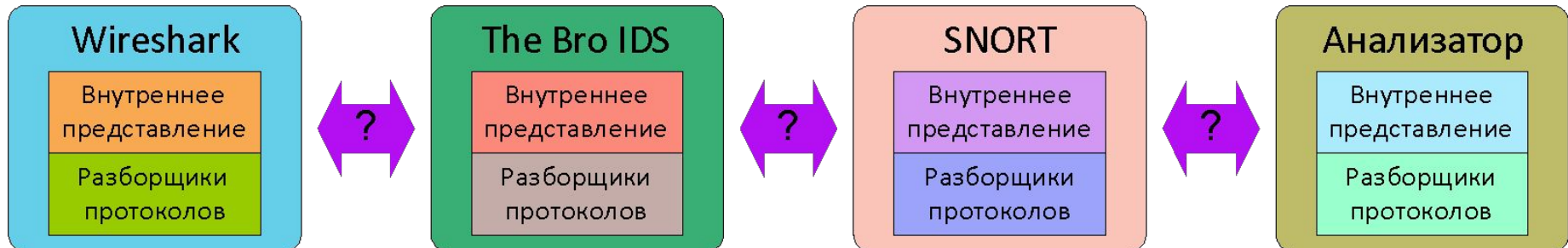
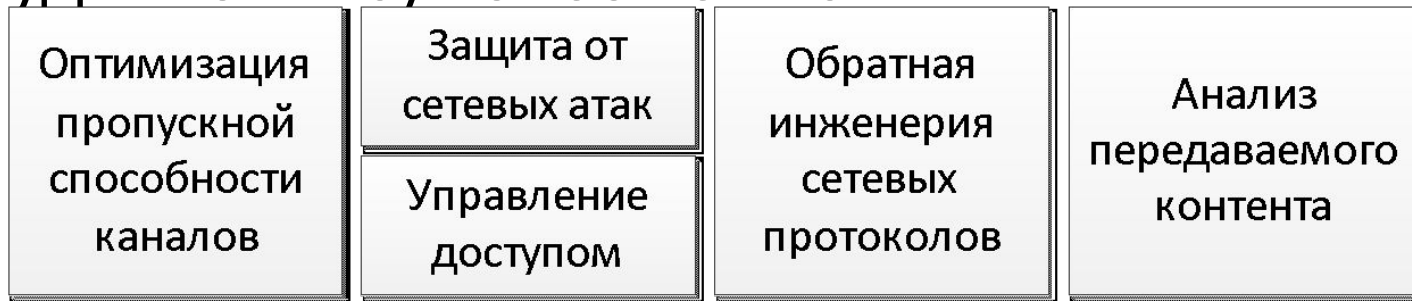


ПРОТОСФЕРА:
программная инфраструктура для
глубокого разбора сетевого трафика

Анализ сетевого трафика:

практические задачи и инструменты для их решения

- Спектр инструментов для решения практических задач анализа трафика очень широк
- Интеграция инструментов затруднительна/невозможна



Типовые действия при разборе сетевого трафика

1. Распознавание протокола

- Определение прикладного протокола:
 - по номеру порта
 - посредством поиска характерных шаблонов в полезной нагрузке пакетов

2. Разбор пакетов

- Задаются форматы сообщений, разделяющие пакеты на служебные поля и полезную нагрузку

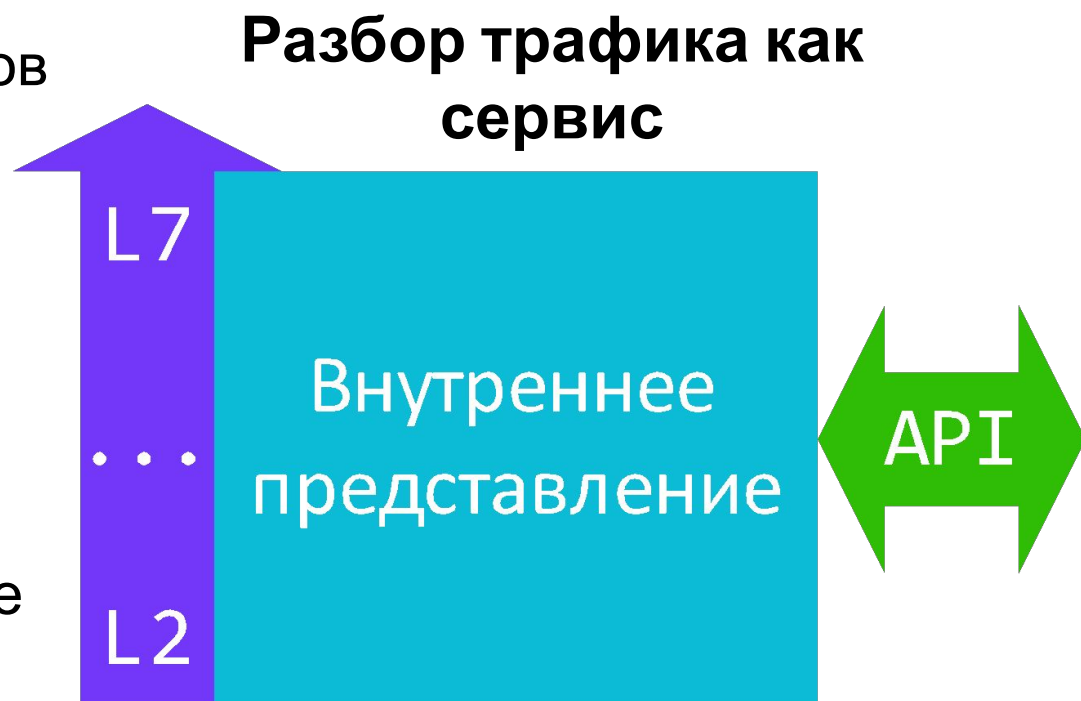
3. Сборка потока данных, передаваемого посредством протокола следующего уровня

- Используется упрощенный автомат протокола

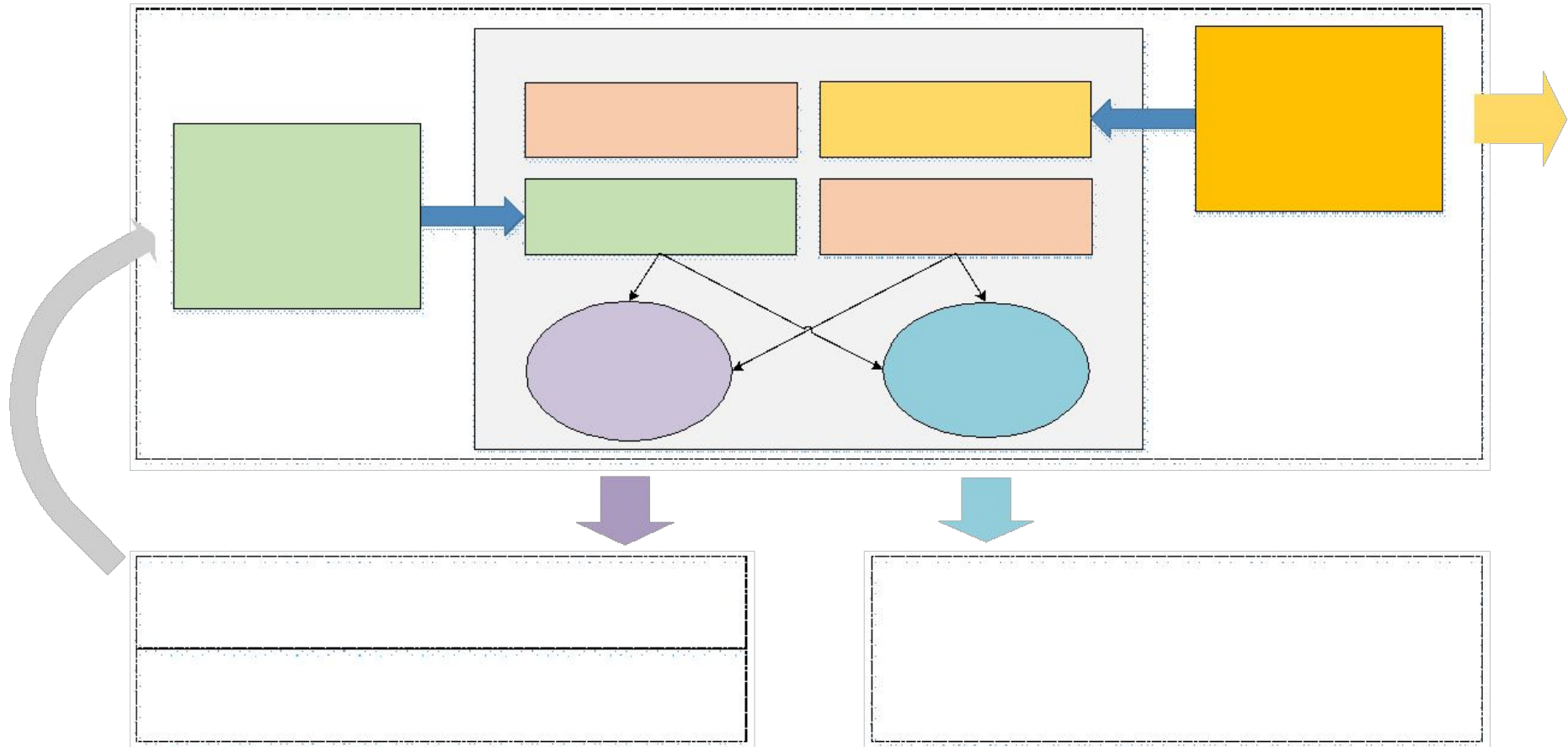


Требования к системе

- Открытый интерфейс (API)
 - добавление разборщиков
 - получение результатов разбора
- Разбор протоколов произвольного сетевого стека
 - сборка потоков
 - зашифрованные данные
 - связанные потоки
- Отладка системы на трафике, при разборе которого возникли ошибки



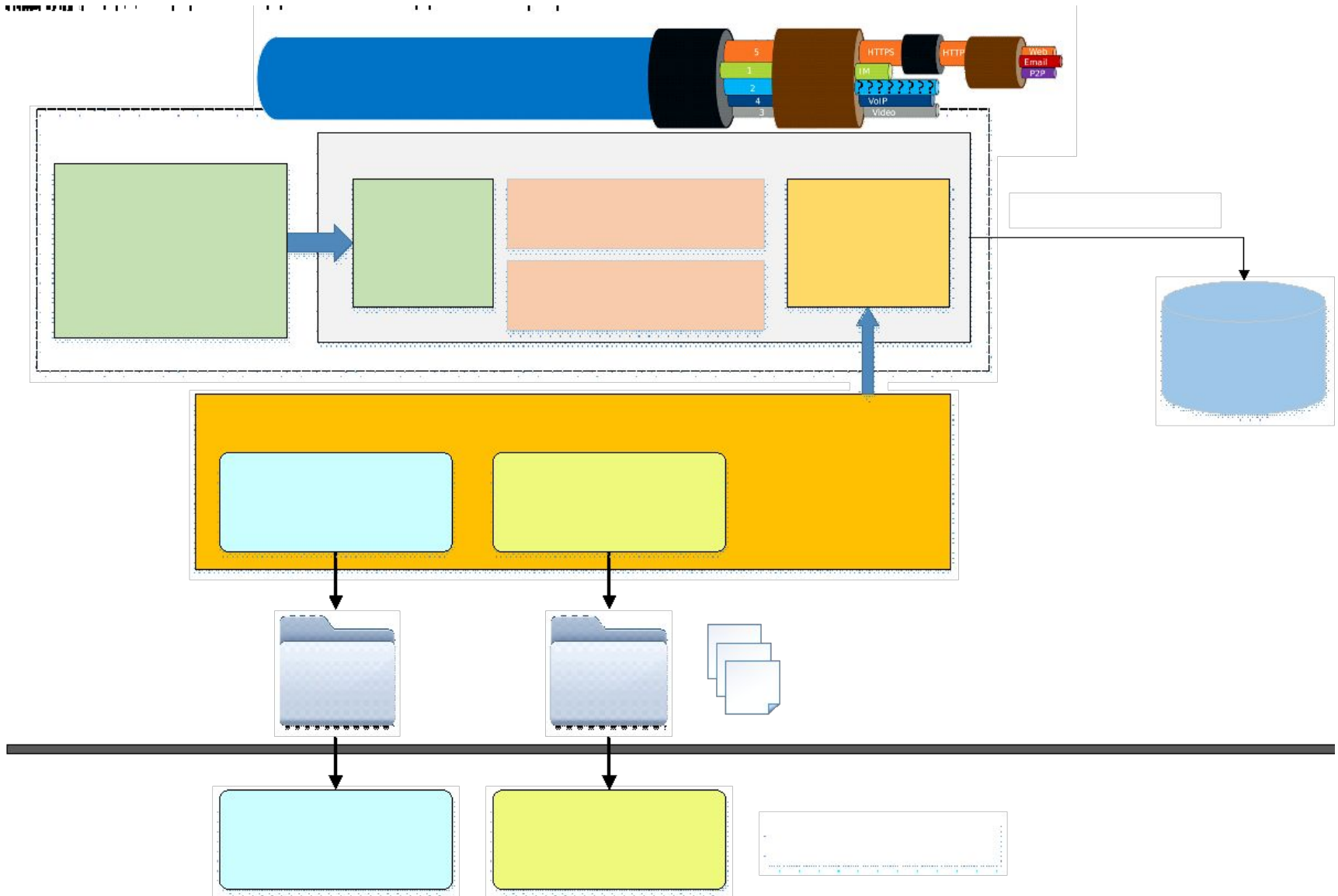
Архитектура системы



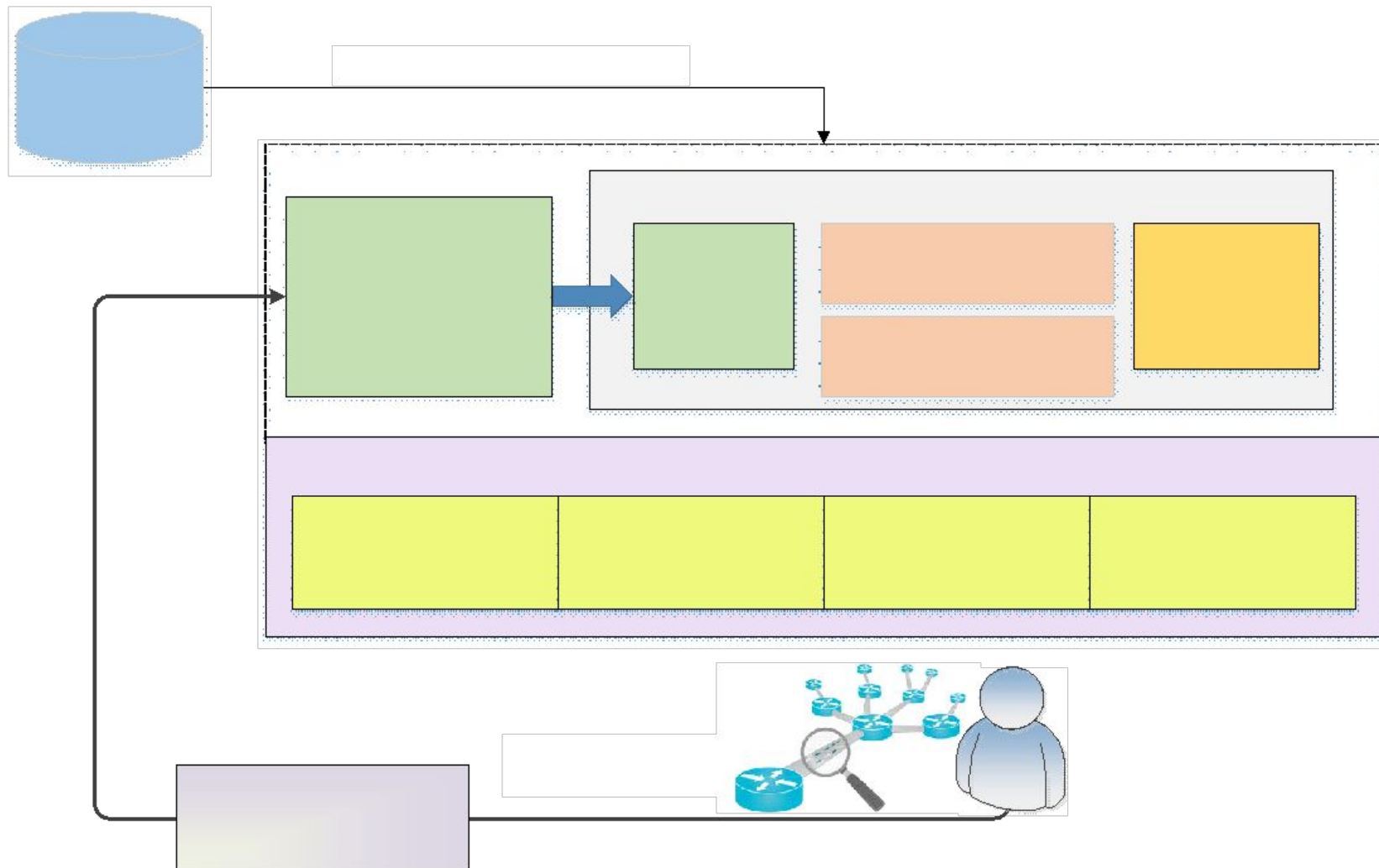
Модель представления данных

- Реализация типовых действий при разборе сетевого трафика:
 - интерфейс для подключения «распознавателей» протоколов (автоматическое распознавание)
 - представление пакета в виде дерева разбора
 - сборка потоков (отдельные буферы) для протоколов произвольного уровня сетевого стека с учетом потери и переупорядочивания пакетов при передаче
 - поддержка состояний для stateful-протоколов с привязкой к сетевому стеку
 - возможность разбора зашифрованных данных
 - выявление связанных потоков

Пример №1: DSI на потоке



Пример №2: доработка разборщиков



Отладка разборщика: журнал событий

Timestamp	Topic	Content
13:29:10.100	HTTP	Found HTTP-client.
13:29:10.100	HTTP	Found HTTP-server.
13:29:10.111	HTTP	Gif file found.
13:29:10.122	EthLog	Can't recognize parse type.
13:29:10.123	IPv4	Can't recognize parse type.
13:29:10.123	IPv6	Unsupported IPv6 next header
13:29:10.123	IPv6	Can't recognize parse type.
13:29:10.149	EthLog	Can't recognize parse type.
13:29:10.382	SSL	Can't decrypt data from server
13:29:10.386	SSL	Can't decrypt data from server

5841, Packet(PcapPacket), [0x4267C0:0x66]
Data(Eth), Src: 40:61:86:62:d9:a1, Dst: 33:33:00:00:00:01
Data(Ip6), From: 100:0:600:0:78fb:100::, To: ff02::1
Version(IpVersion), 6
Class(TrafficClass), 0
Label(FlowLabel), 0
PayloadSize(BeUint16), 0x0020
NextHeader(Uint8)
HopLimit(Uint8)
Src(Ip6Addr), 100:0:600:0:78fb:100::
Dst(Ip6Addr), ff02::1
Data(Undefined)
5842, Packet(PcapPacket), [0x426826:0x6C]
Data(Eth), Src: 8c:89:a5:1a:14:ff, Dst: ff:ff:ff:ff:ff:ff (broadcast)
Data(Ip4), From: 10.10.14.102, To: 10.10.14.255
Data(Udp), SrcPort: 137, DstPort: 137
5843, Packet(PcapPacket), [0x426892:0x7E]
Data(Eth), Src: 8c:89:a5:1a:14:ff, Dst: 00:1c:7f:30:da:ed
Data(Ip4), From: 10.10.14.102, To: 10.10.12.2
Data(Udp), SrcPort: 123, DstPort: 123

004267F0:	01 00 00 00 00 00 FF 02 00 00 00 00 00 00 00
00426800:	00 00 00 00 00 01 3A 00 05 02 00 00 00 00 82 00
00426810:	F9 C5 03 EB 00 00 00 00 00 00 00 00 00 00 00
00426820:	00 00 00 00 00 00 78 24 D0 50 13 33 0E 00 5C 00
00426830:	00 00 5C 00 00 00 FF FF FF FF FF FF 8C 89 A5 1A
00426840:	14 FF 08 00 45 00 00 4E 22 F4 00 00 80 11 E6 32
00426850:	0A 0A 0E 66 0A 0A 0E FF 00 89 00 89 00 3A F0 86
00426860:	9A FA 01 10 00 01 00 00 00 00 00 00 20 46 48 46

Выявление связанных потоков на примере SIP

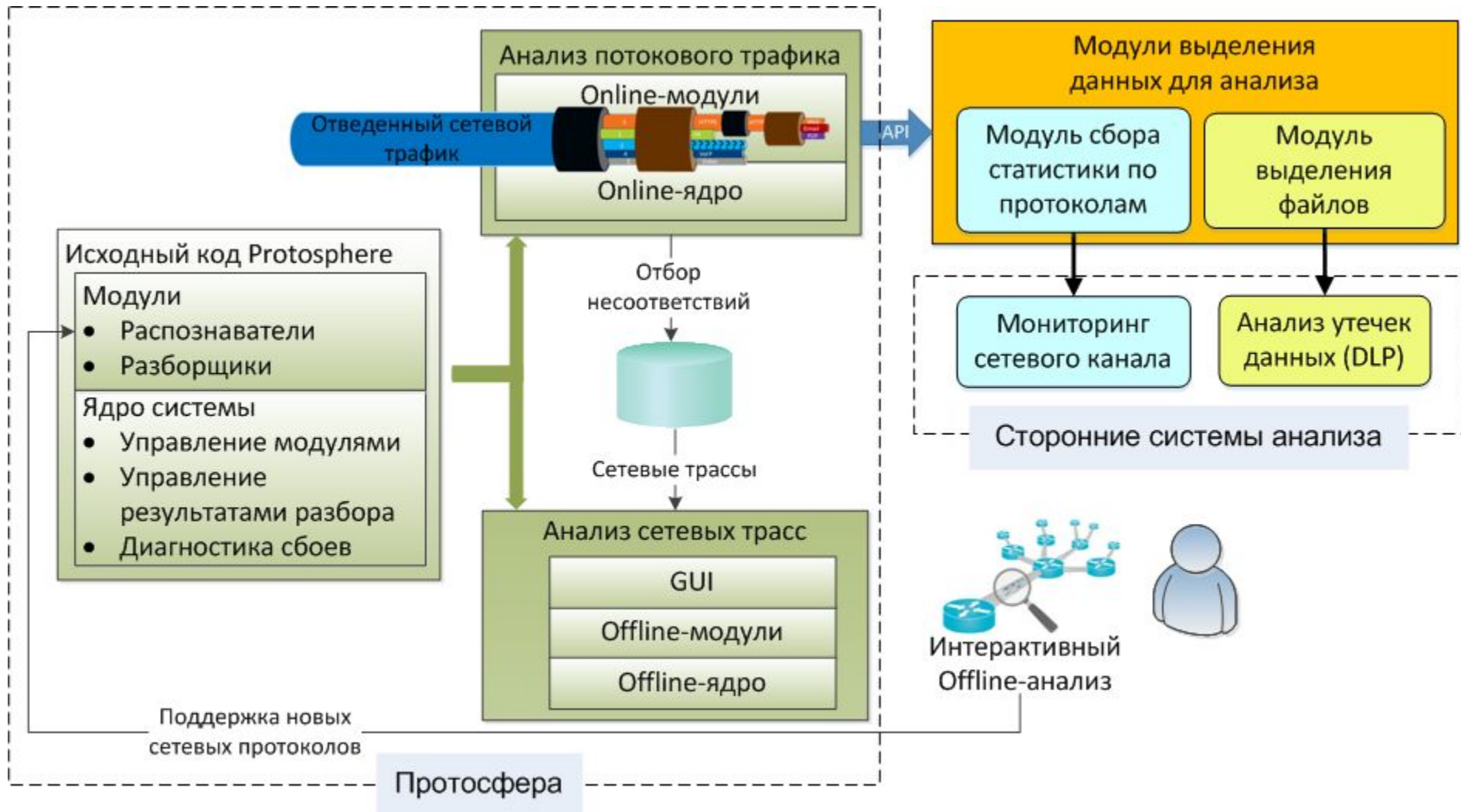
The screenshot displays a network analysis tool interface with several panels:

- Block tree:** Shows a hierarchical view of captured packets. Packet 497 (Data(Tcp)) and packet 498 (Data(Eth)) are highlighted, showing their respective source and destination addresses and ports.
- Data:** A hex dump of the selected packet data, showing hexadecimal values and their corresponding ASCII characters.
- Contexts:** A tree view showing the context of the selected packet, including IP4, UDP, and SIP contexts.
- Find string (4):** A search window showing the results of a search for the string "sip" in the selected packet, listing four context instances.
- Causality table:** A table showing the relationship between causes and results in the network flow.

	Cause	Result	Reason
1	Sip. Udp: [{"Src": "5060"}, {"Dst": "5061"}]. Ip4: [{"Src": "200.57.7.195"}, {"Dst": "200.57.7.204"}].	Udp: [{"Src": "8000"}, {"Dst": "40376"}]. Ip4: [{"Src": "200.57.7.204"}, {"Dst": "200.57.7.196"}].	Discovered SIP session parameters=(address=200.57.7.196, port=40376)
2	Sip. Udp: [{"Src": "5061"}, {"Dst": "5060"}]. Ip4: [{"Src": "200.57.7.204"}, {"Dst": "200.57.7.195"}].	Udp: [{"Src": "8000"}, {"Dst": "40376"}]. Ip4: [{"Src": "200.57.7.204"}, {"Dst": "200.57.7.196"}].	Discovered SIP session parameters=(address=200.57.7.204, port=8000)
3	Sip. Udp: [{"Src": "5060"}, {"Dst": "5061"}]. Ip4: [{"Src": "200.57.7.195"}, {"Dst": "200.57.7.204"}].	Udp: [{"Src": "40376"}, {"Dst": "8000"}]. Ip4: [{"Src": "200.57.7.196"}, {"Dst": "200.57.7.204"}].	Discovered SIP session parameters=(address=200.57.7.196, port=40376)
4	Sip. Udp: [{"Src": "5061"}, {"Dst": "5060"}]. Ip4: [{"Src": "200.57.7.204"}, {"Dst": "200.57.7.195"}].	Udp: [{"Src": "40376"}, {"Dst": "8000"}]. Ip4: [{"Src": "200.57.7.196"}, {"Dst": "200.57.7.204"}].	Discovered SIP session parameters=(address=200.57.7.204, port=8000)

Программный комплекс Протосфера

DPI как сервис



DPI

Веб

1. Выделение **пользовательских** веб-запросов для оценки посещаемости отдельных сайтов
2. Тематическая группировка посещаемых сайтов
 - На основе краулинга ресурсов классификации (Alexa, SimWeb, Trendmicro)
 - На основе анализа текстового содержимого страниц (Texterra)
3. Выявление веб-технологий и фреймворков применяемых на сайте (аналог Wappalyzer, ускорение за счёт hyperscan x5)

Шифрованный трафик

1. Анализ качества передаваемых видео-потокков на основе машинного обучения
2. Идентификация клиентского приложения по TLS-рукопожатию
3. Классификация трафика по протоколам прикладного уровня на основе машинного обучения по потоковым данным

Общие задачи

1. Поиск большого числа регулярных выражений в данных произвольного протокола/файла на с использованием библиотеки hyperscan

Дальнейшее развитие

- Интеграция со средствами анализа бинарного кода:
 - выделение структуры и визуализация данных в памяти
- Интеграция с системой скоростного анализа трафика:
 - web-интерфейс управления и визуализации результатов
 - распараллеливание обработки
 - предварительная классификация по протоколам L7
- Единая база форматов данных на декларативном языке. Примеры:
 - Kaitai Struct, Peach-Pit, ...
- Интеграция с фаззером с использованием ИЗВЕСТНЫХ :
 - автоматизированная генерация reach-pit на основе знаний о структуре

Спасибо за внимание