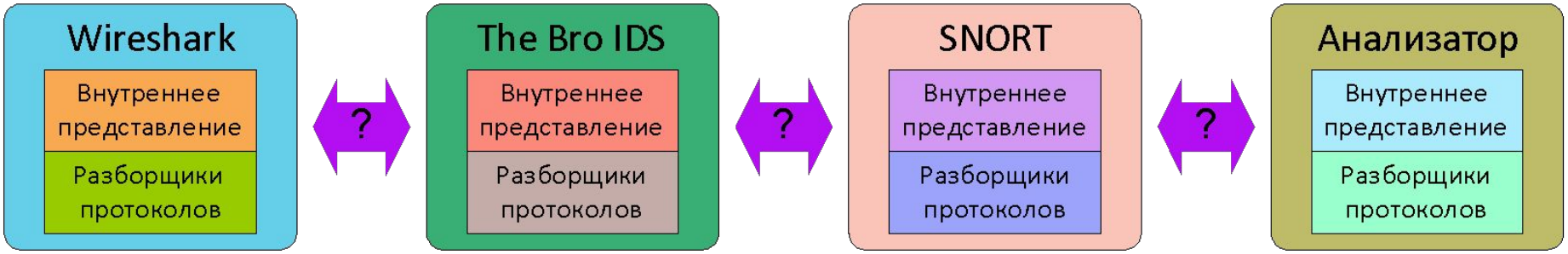
ИСП РАН

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

Оптимизация пропускной способности каналов	Защита от сетевых атак Управление доступом	Обратная инженерия сетевых протоколов	Анализ передаваемого контента



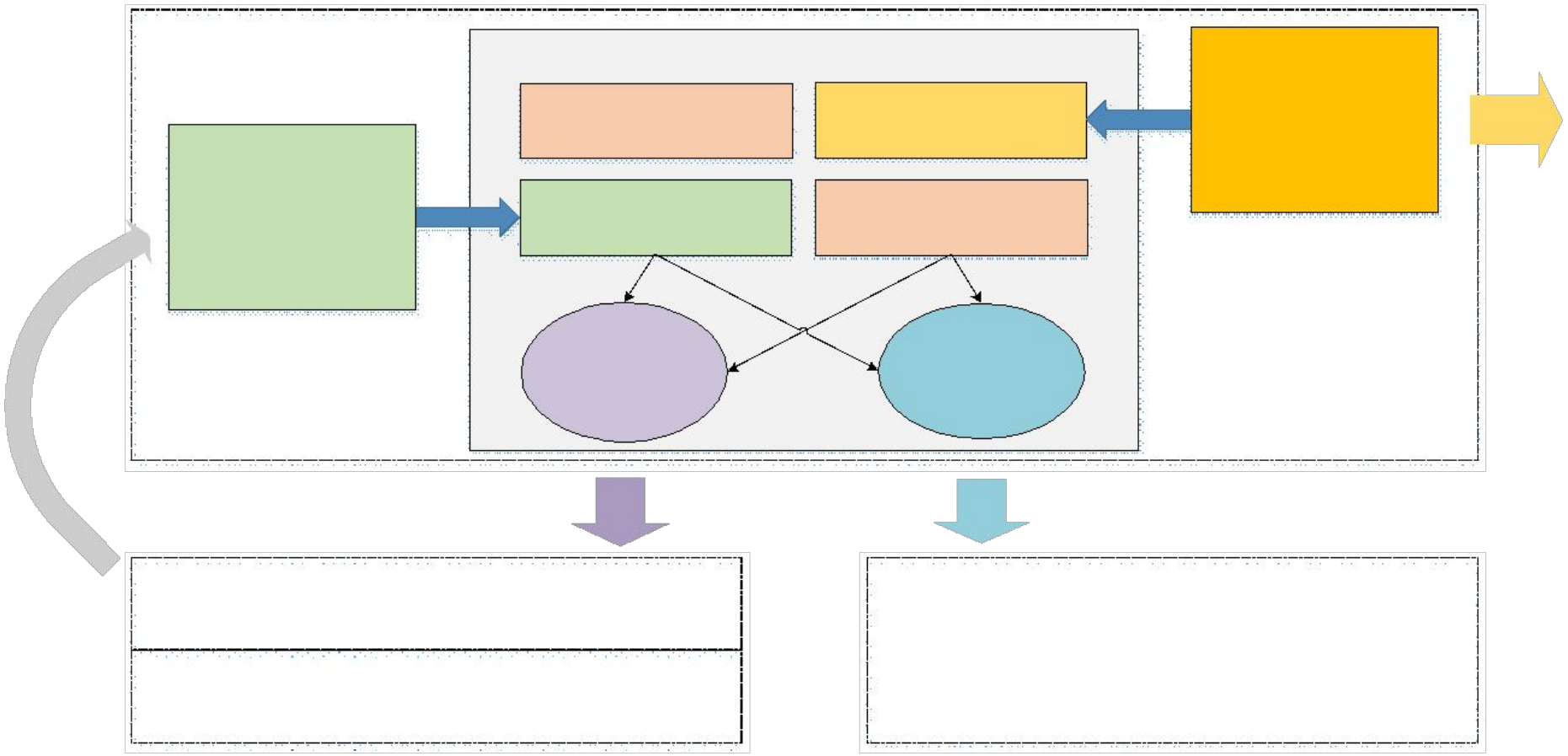
Типовые действия при разборе сетевого трафика

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

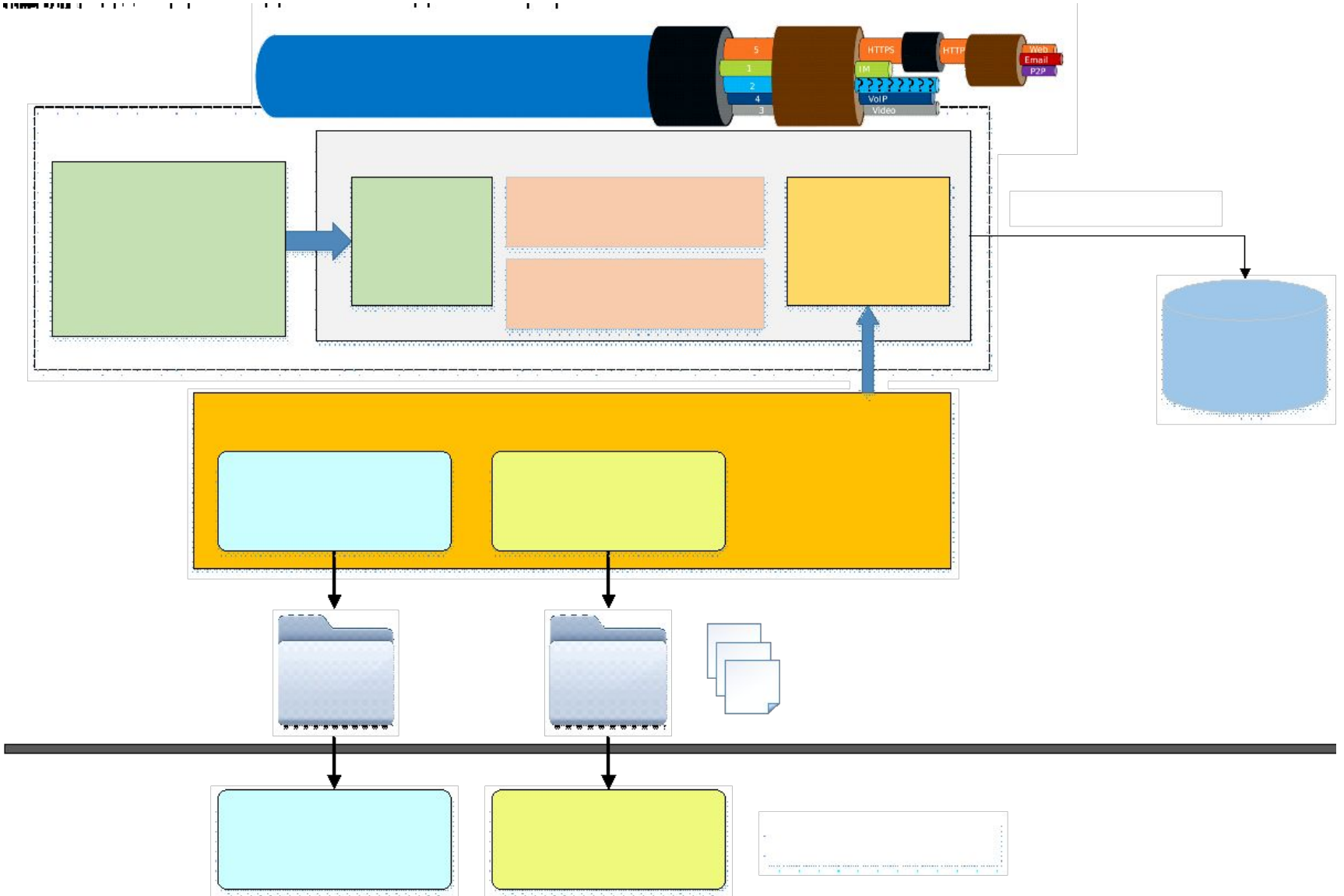
Архитектура системы



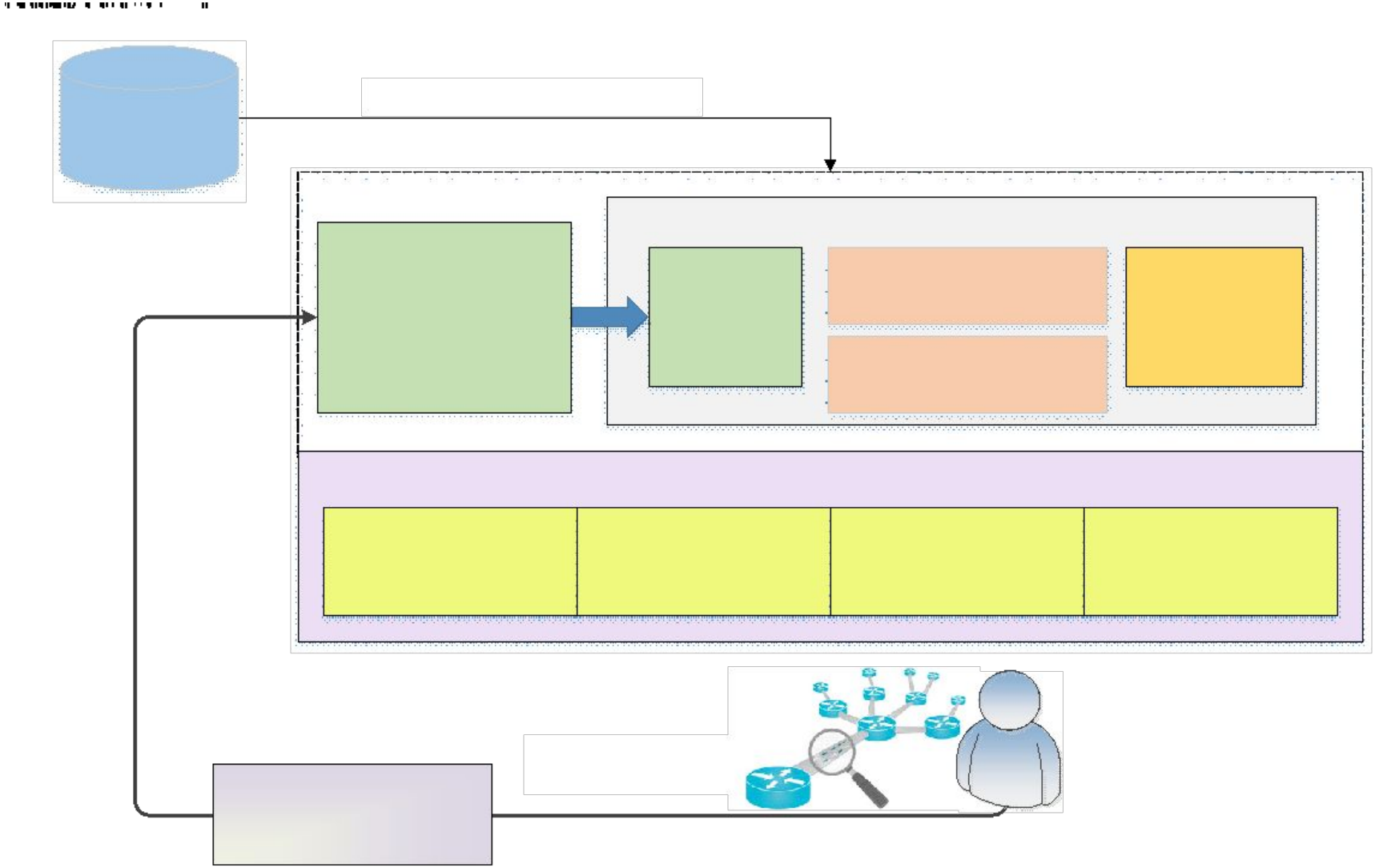
Модель представления данных

- 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

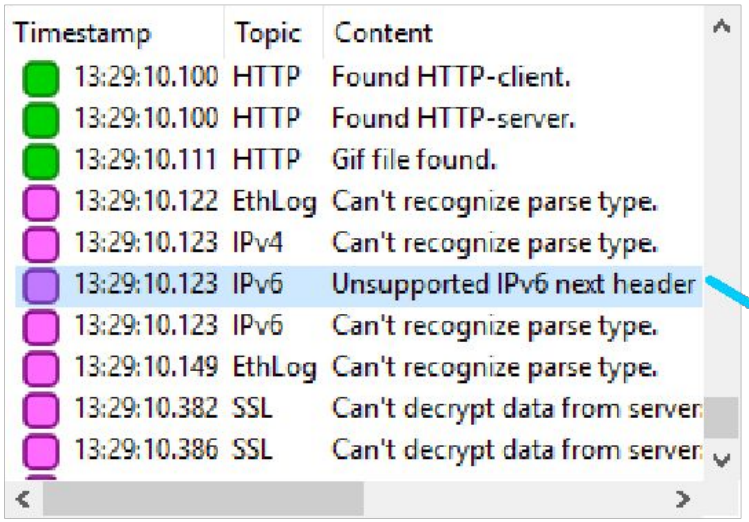
Пример №1: DCI на потоке

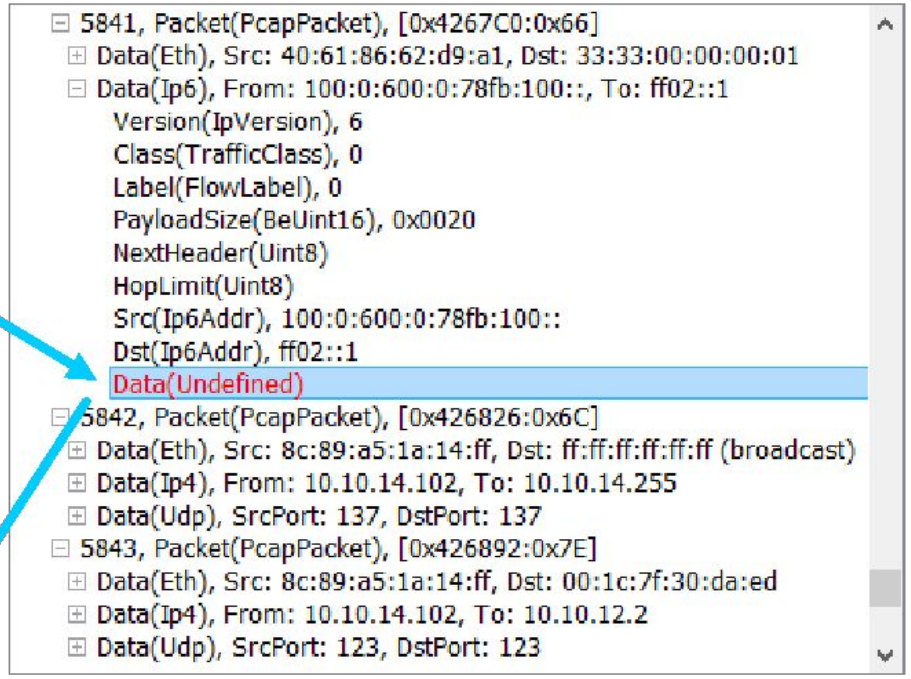


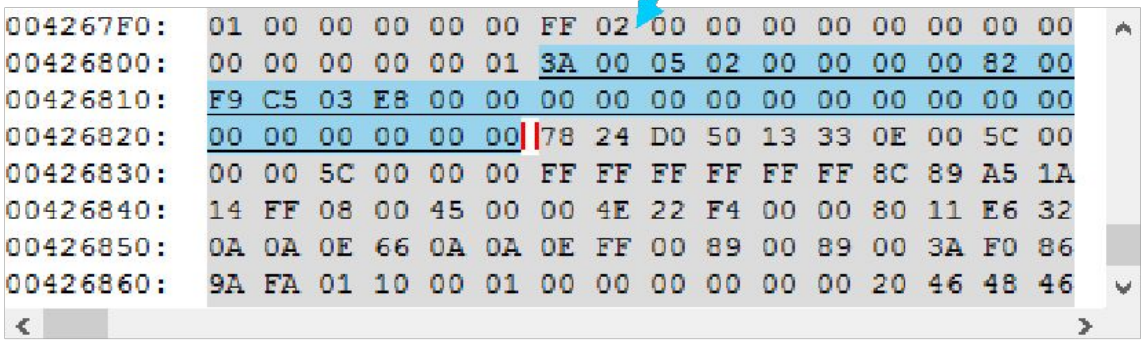
Пример №2: доработка разборщиков



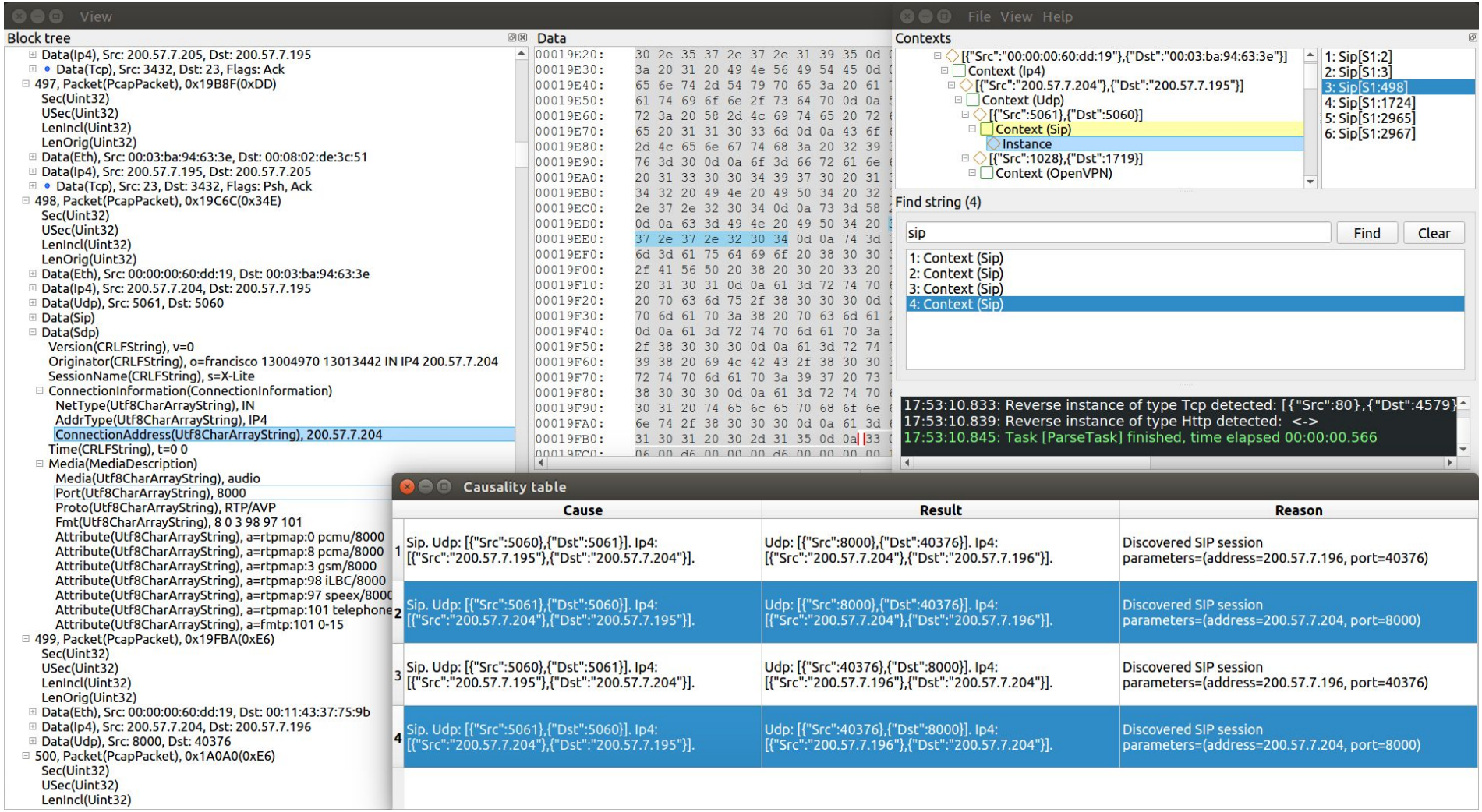
Отладка разборщика: журнал событий



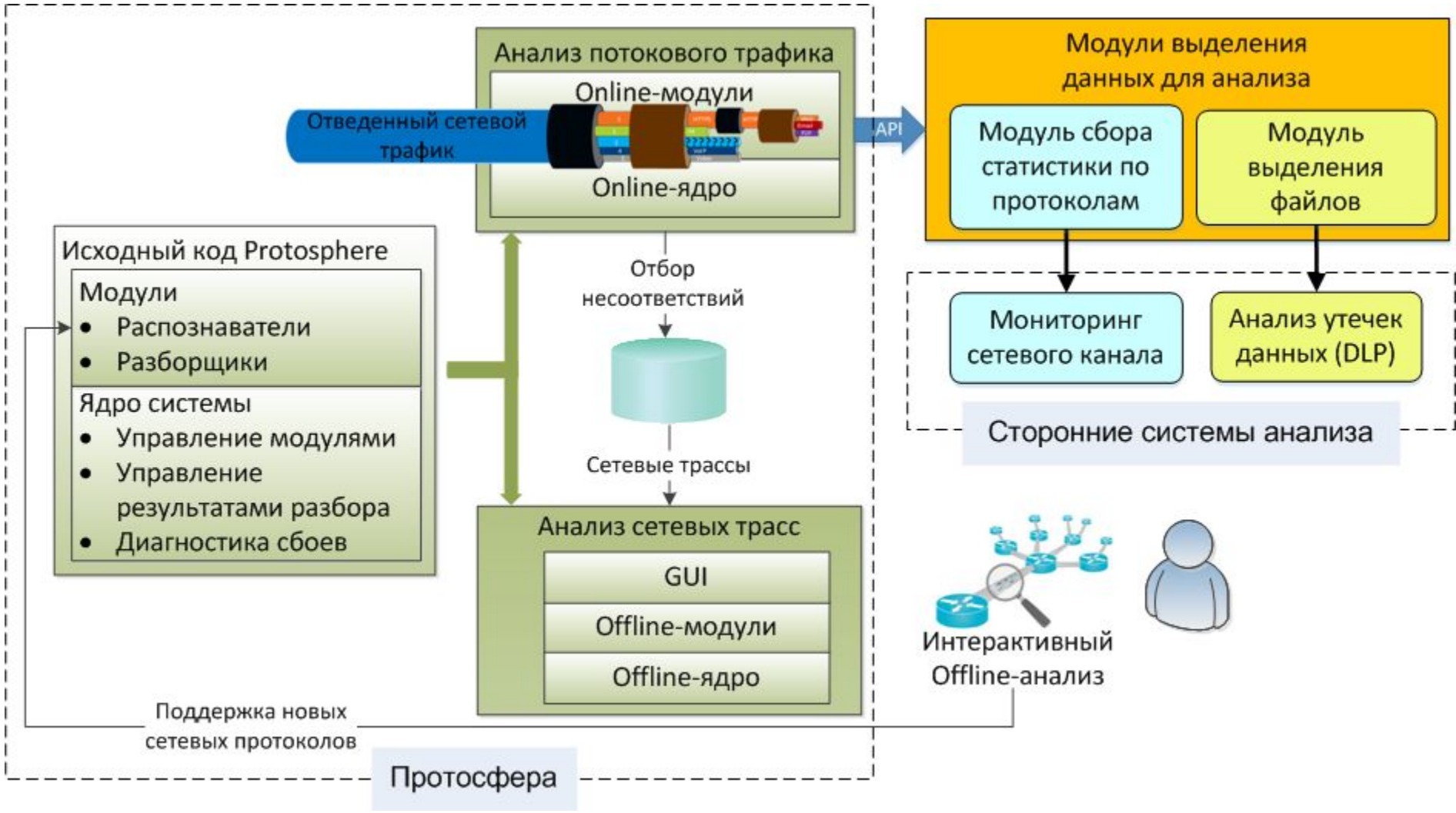




Выявление связанных потоков на примере SIP



Программный комплекс Протосфера
DPI как сервис



Решение прикладных задач DPI

ИСП РАН

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-потоков 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

ИСП РАН

- 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 peach-pit на основе знаний о структуре

ИСП РАН

Спасибо за внимание