

ВЫЧИСЛИТЕЛЬНЫЕ ПОТОКОВЫЕ АРХИТЕКТУРЫ: ИСТОРИЯ И ПЕРСПЕКТИВЫ РЕАЛИЗАЦИИ

Захаров В.Н., Степченков Ю.А.,
Хилько Д.В., Дьяченко Ю.Г.



Федеральный Исследовательский Центр
«Информатика и Управление»
Российской Академии Наук

СОДЕРЖАНИЕ

- **Специфика потоковых архитектур**
- **Недостатки потоковых архитектур общего назначения**
- **Особенность потоковых ЦСП**
- **Состояние дел в России**
- **Потоковая рекуррентная архитектура**
- **Заключение**

СПЕКТР ПОТОКОВЫХ АРХИТЕКТУР

(НАЧАЛО 1980-Х - СЕРЕДИНА 2000-Х ГОДОВ)

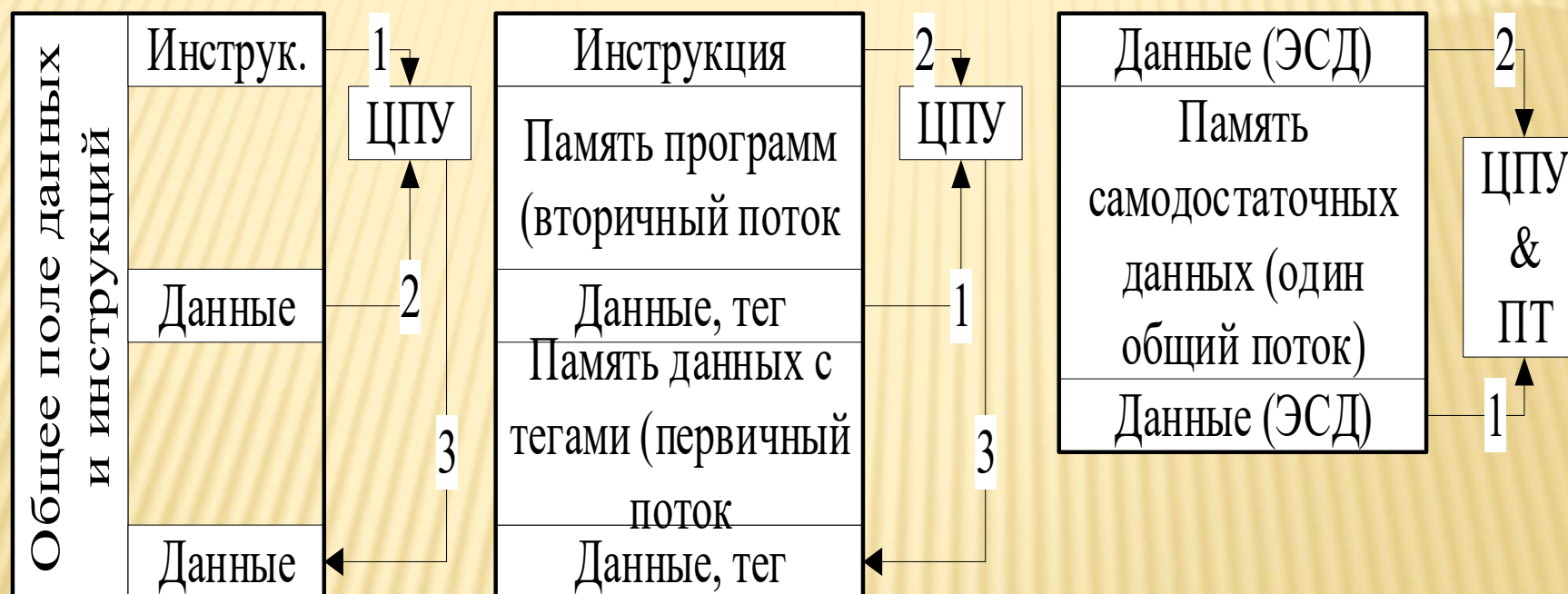
- статические потоковые архитектуры
- динамические потоковые архитектуры:
с помеченными токенами и с явно
адресуемыми токенами
- редукционные потоковые архитектуры
- комбинированные (гибридные)
потоково-фон-Неймановские
архитектуры

КЛЮЧЕВЫЕ ОТЛИЧИЯ АРХИТЕКТУР: ИНИЦИАЦИЯ ВЫЧИСЛЕНИЙ И ОРГАНИЗАЦИЯ ПАМЯТИ

Традиционная (CF/S)

Потоковая (DF/S)

МПРА (DF/D)



- ✗ Основной критерий - по организации памяти: *Control-Flow/Static* (CF/S), *Data-Flow/Static* (DF/S), *Data-Flow/Dynamic* (DF/D)

КЛЮЧЕВЫЕ ОТЛИЧИЯ АРХИТЕКТУР:

ЧИСЛО ВЫЧИСЛИТЕЛЬНЫХ ШАГОВ

Традиционная (CF/S)	Выборка инстр	Дешифр инстр	Чтение данных	Выполн инстр	Запись результ
	1	2	3	4	5
Потоковая (DF/S)	Сравнение тегов	Выборка инстр	Дешифр инстр	Выполн инстр	Запись результ
	1	2	3	4	
РПА (DF/D)	Сравнение тегов, Дешифр. Инстр.	Выполн инстр. Преобраз. тегов.	Запись результ		
	1	2	3		

НЕДОСТАТКИ ПОТОКОВЫХ АРХИТЕКТУР (1)

- ❖ Аппаратная сложность и время сравнения тегированных маркеров в Памяти сравнения достаточно высоки
- ❖ Последовательные участки программного кода исполняются неэффективно
- ❖ Отсутствие ограничения числа одновременно выполняемых итераций цикла и контекста процедуры ведет к увеличению размера тегов, потерь времени на их коммуникацию по сети и объема аппаратуры.
- ❖ Невозможно использовать регистровые структуры с такой же эффективностью, как в обычной многопроцессорной системе

НЕДОСТАТКИ ПОТОКОВЫХ ЦСП АРХИТЕКТУР

- ❖ **Динамическое распределение данных для большинства ЦОС применений излишне**
- ❖ **Большой размер теговых полей - превышает в десятки раз размер собственно обрабатываемых данных**
- ❖ **Использование принципа однократного присваивания - избыточно**
- ❖ **Высокая цена многопроцессорных реализаций и потери производительности при внекристальной реализации**

Основной недостаток – стоимостной фактор

ИТОГИ РАЗРАБОТКИ

Исторический опыт разработки компьютерных архитектур свидетельствует, что новая архитектура может потерпеть неудачу по следующим причинам:

- ❖ если она базируется на чрезвычайно сложных аппаратных механизмах для поддержки предназначенной ей модели программирования
- ❖ если она игнорирует совместимость с существующими вычислительными средами
- ❖ если она игнорирует проблему программируемости

СОСТОЯНИЕ ЗА РУБЕЖОМ

**В НАСТОЯЩЕЕ ВРЕМЯ ЗА РУБЕЖОМ ВОПРОСАМИ
РАЗРАБОТКИ ПОТОКОВЫХ АРХИТЕКТУР ВЕДУТСЯ
ТОЛЬКО ОТДЕЛЬНЫМИ ИССЛЕДОВАТЕЛЯМИ**

Sundararajan Sriram. Minimizing communication and synchronization overhead in multiprocessors for digital signal processing // Phd, University of California at Berkeley, Computer Science Division 571 Evans Hall Berkeley, CA, United States, Order Number: UMI Order No. GAX96-21374, 187 p.

В диссертации исследованы методы минимизации затрат на межпроцессорную связь в статически запланированных мультипроцессорах для DSP на основе DF. Основная идея - обмен данными в статически запланированном оборудовании достаточно предсказуемы, и ее можно использовать для параллельной реализации с низкими накладными расходами при низкой стоимости оборудования на базе самосинхронизации.

СОСТОЯНИЕ В РОССИИ

На постоянной основе вопросами разработки потоковых архитектур занимаются три группы исследователей

- ❖ **Серьезные результаты в области разработки ВС массового параллелизма были достигнуты коллективом под руководством академика Бурцева В.С. (ИПИ РАН), после кончины которого работы были продолжены в ИППМ РАН (пока на модельном уровне)**
- ❖ **Мультиклеточные ядра компании Международный деловой центр "Микрон" (Екатеринбург) обладающей свойством естественной реализации параллелизма (без решения задачи распараллеливания). Мультиклеточная архитектура позволяет эффективно реализовать вычислительный процесс, что дает увеличение производительности ядра при одновременном снижении энергопотребления**
- ❖ **Многоядерная потоковая рекуррентной архитектуры, которая изначально разрабатывается как специализированная, предназначенная для реализации параллельных вычислительных процессов обработки сигналов в реальном времени**

ПРИЧИНЫ МЕДЛЕННОГО РАЗВИТИЯ ПОТОКОВЫХ АРХИТЕКТУР (МНЕНИЕ СПЕЦИАЛИСТОВ ИППМ РАН)

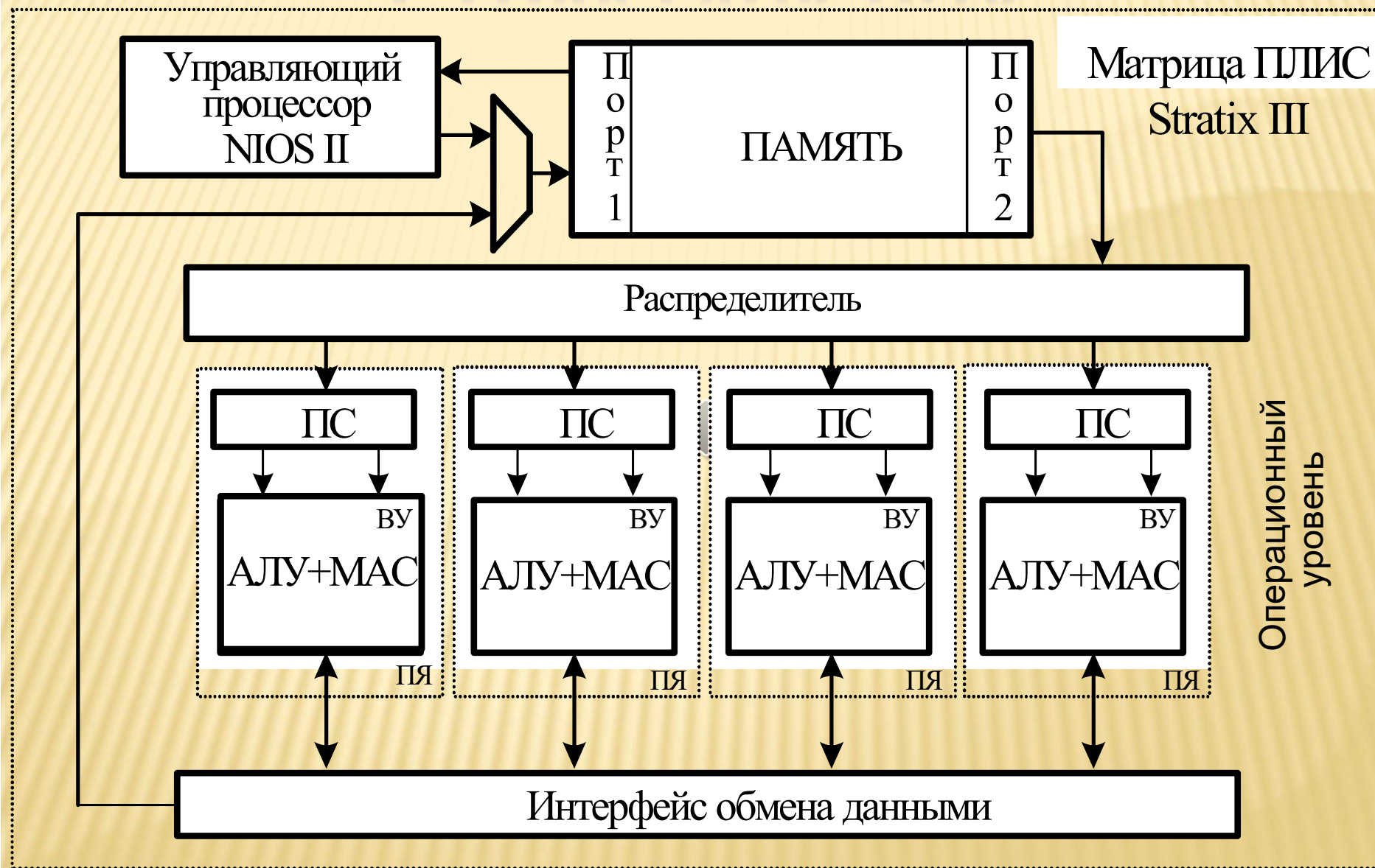
- ❖ **Ориентация разработчиков на господствующую методологию проектирования (парадигму “сбора”) привела в конечном счете к утрате конкурентных преимуществ потоковых архитектур и к ‘поражению’ в «соревновании» с традиционной моделью программирования на основе адресуемой памяти и к фактическому сворачиванию всех работ по данному направлению”**
- ❖ **Известным потоковым процессорам требуется в 2 - 3 раза больше команд для выполнения программы по отношению к процессорам традиционной архитектуры**

ОСНОВНЫЕ ОСОБЕННОСТИ ПОТОКОВОЙ РЕКУРРЕНТНОЙ АРХИТЕКТУРЫ (РПА)

Учет вышеизложенных факторов (причин неуспеха потоковой концепции) требует разработки особого подхода для внедрения потоковой парадигмы в область DSP

- В РПА оба потока сливаются в один общий поток самодостаточных данных, в котором, помимо собственно обрабатываемых данных, содержится и необходимая для их обработки управляющая и служебная информация
- тегируемые данные являются самодостаточными, т.к. они содержат всю информацию по их обработке: информацию о компоненте пары по обработке; о команде, которая должна быть выполнена над парой, и о месте назначения результата операции и не только на текущем шаге обработки

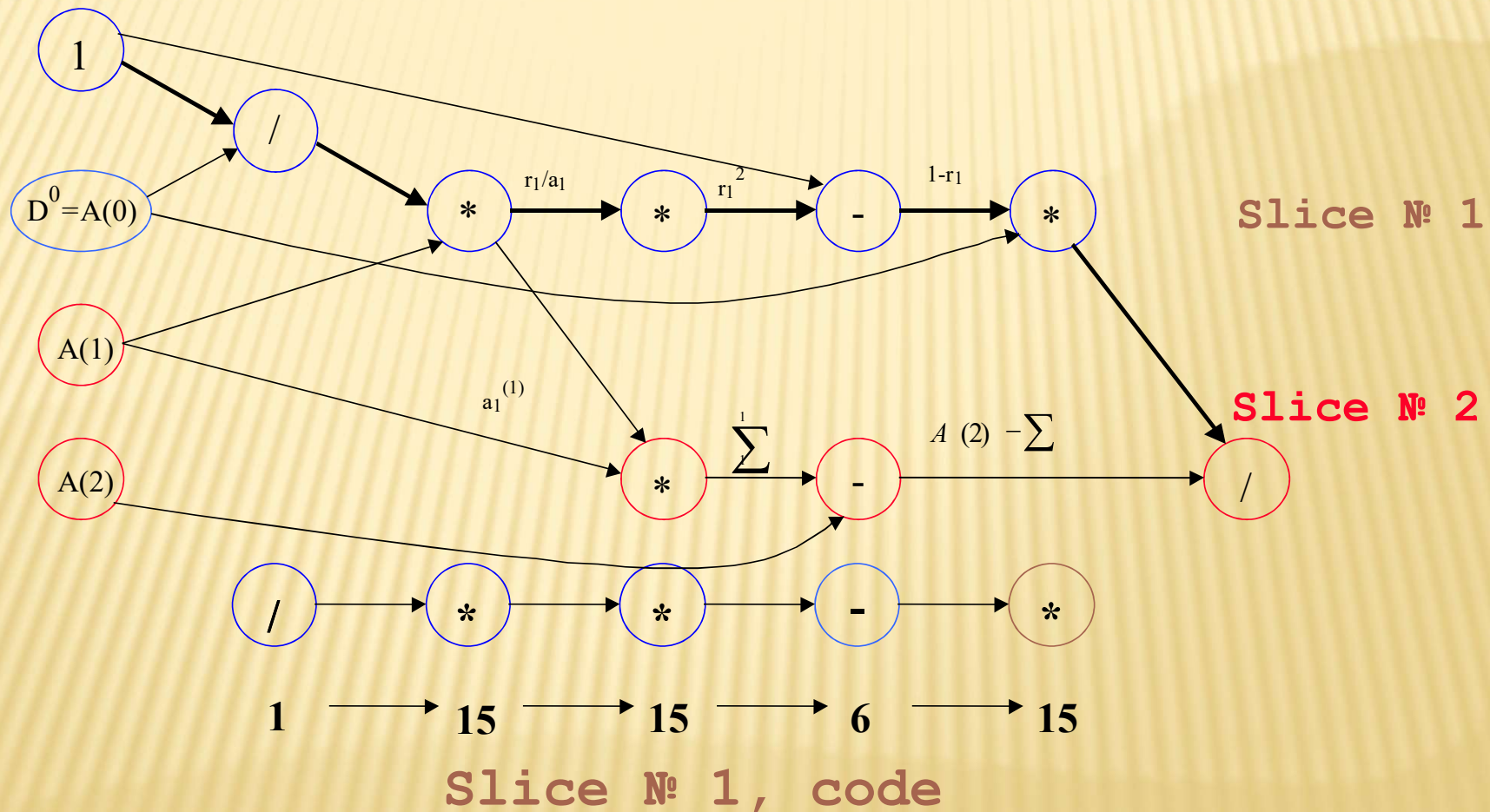
СТРУКТУРА РПА



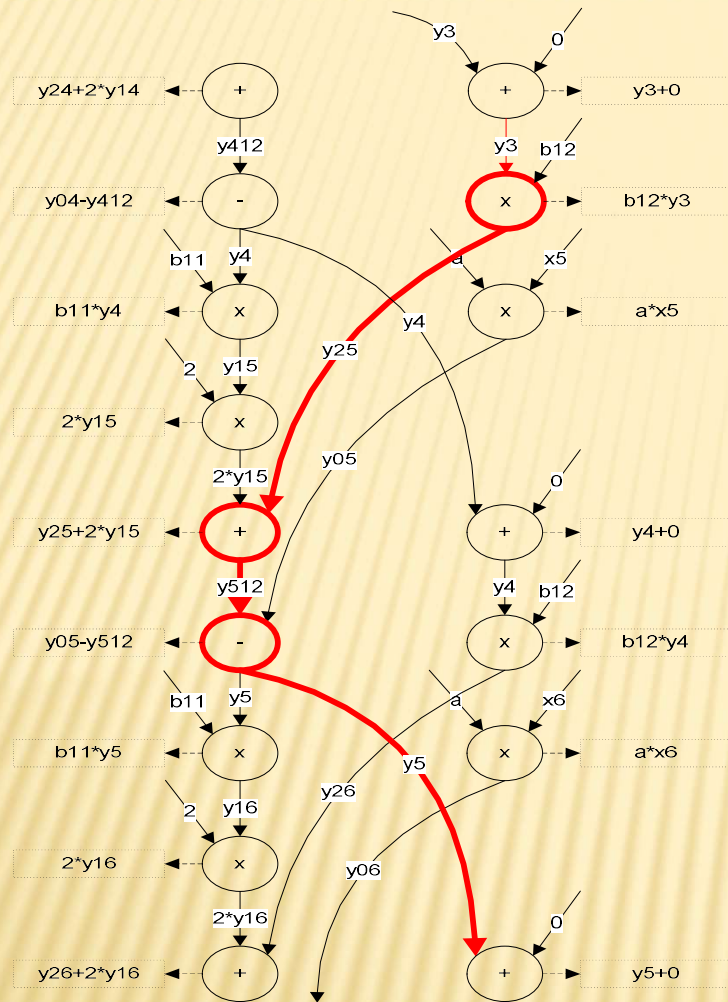
ПРИНЦИП РЕКУРРЕНТНОЙ РАЗВЕРТКИ АЛГОРИТМОВ (1)

Алгоритм Дурбина

$$r_i = \frac{1}{D(i-1)} \left[A(i) - \sum_{j=1}^{i-1} a_j^{(i-1)} A(i-j) \right]$$



ПРИНЦИП РЕКУРРЕНТНОЙ РАЗВЕРТКИ АЛГОРИТМОВ (2)

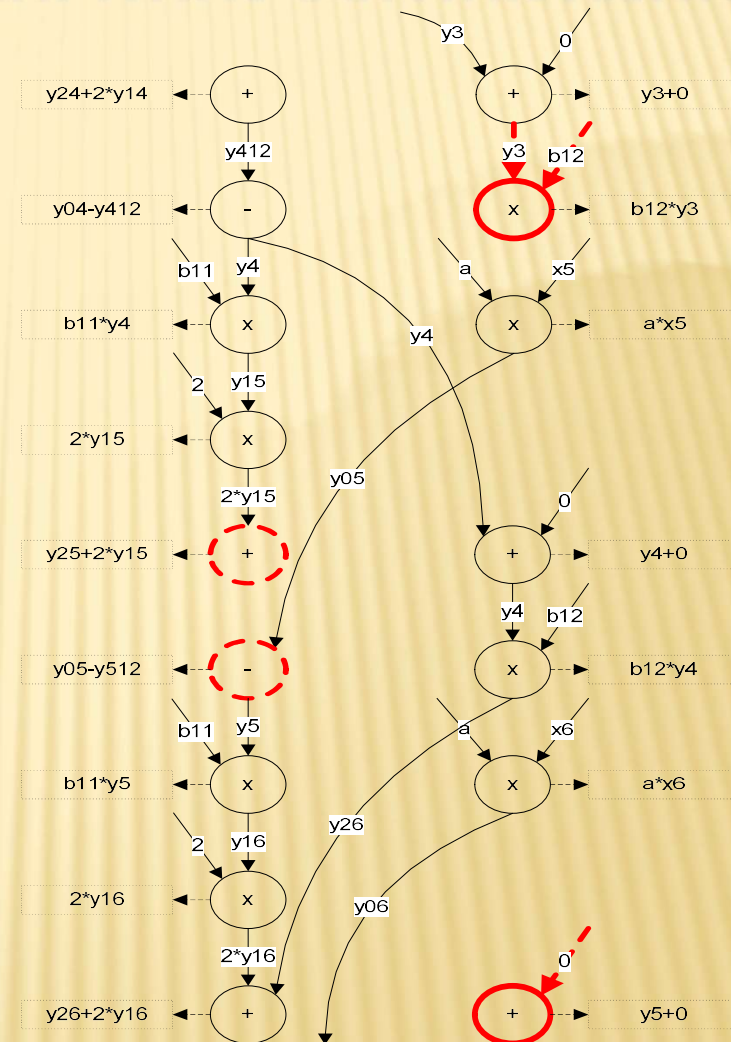


Развернутый граф

— сворачиваемая цепочка

ФИЦ ИУ РАН

= >



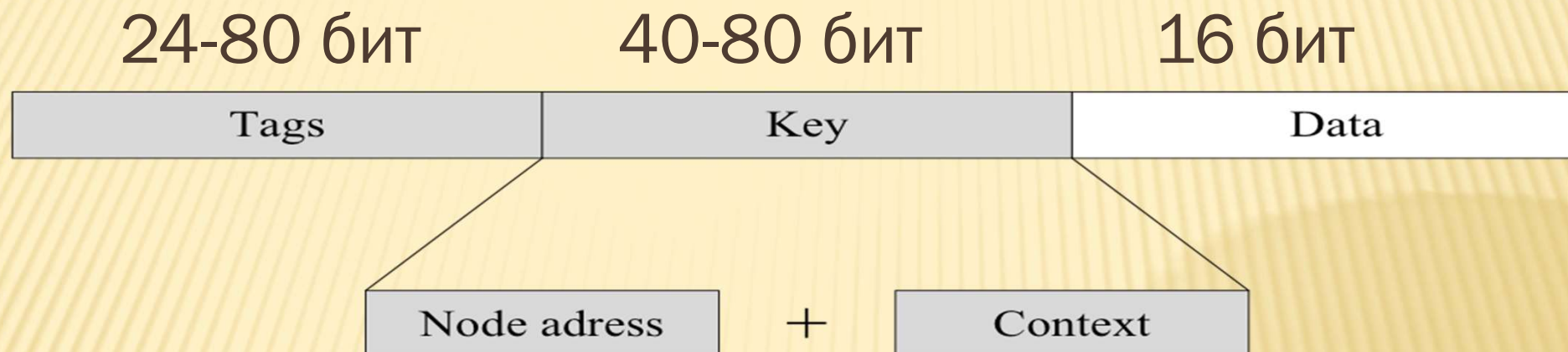
Динамический свернутый граф

динамически порождаемые вершин

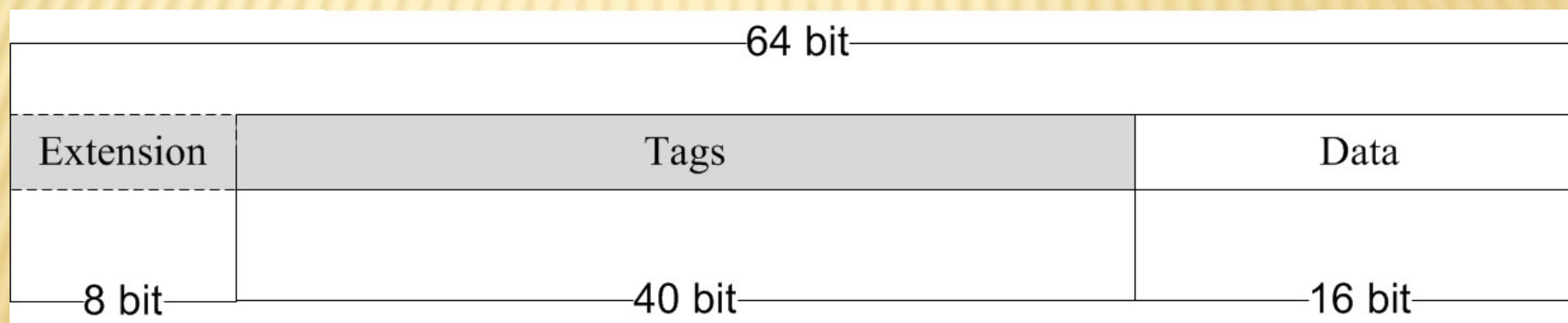
МММЭК-2021

15 из 21

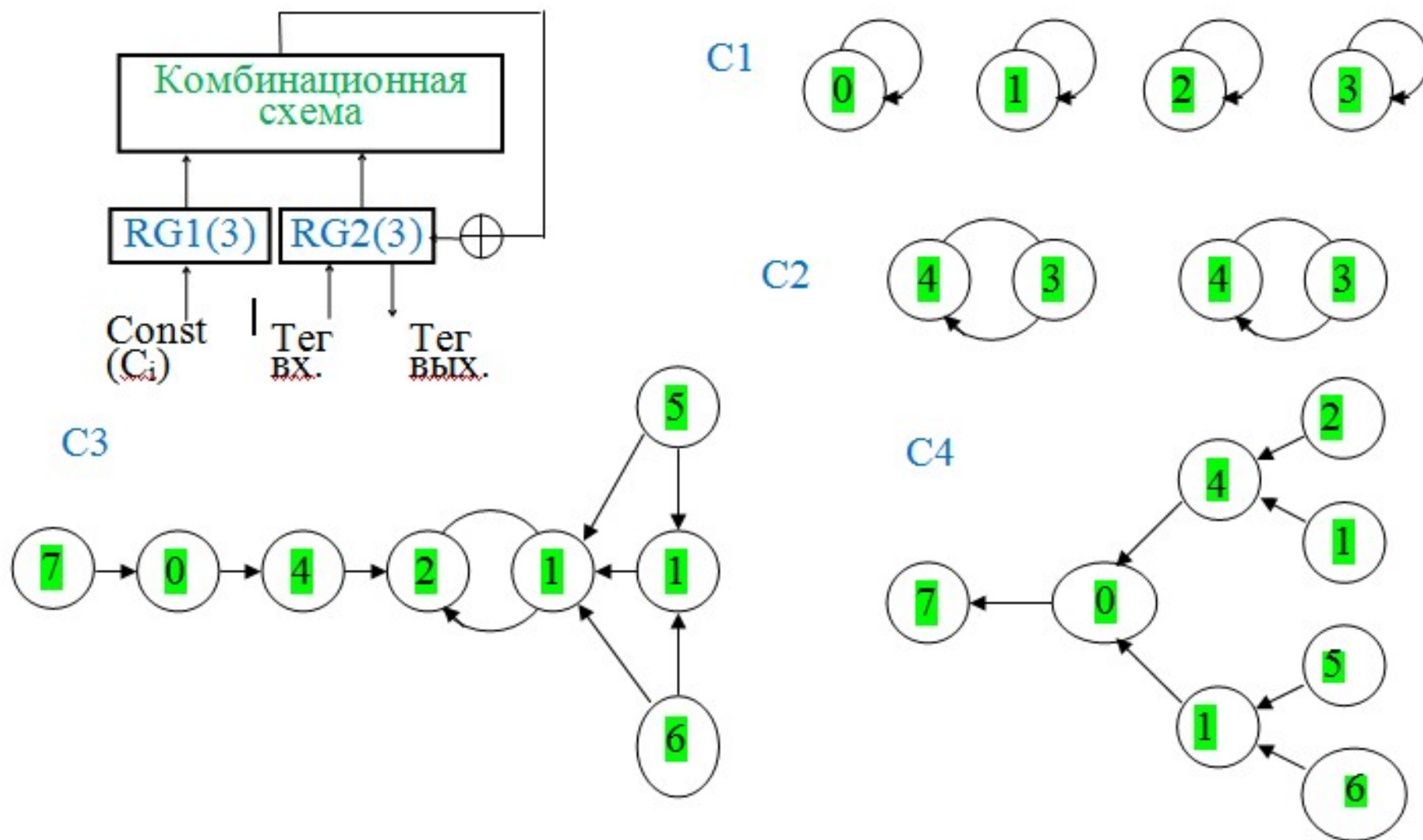
КЛАССИЧЕСКАЯ СТРУКТУРЫ ТОКЕНА



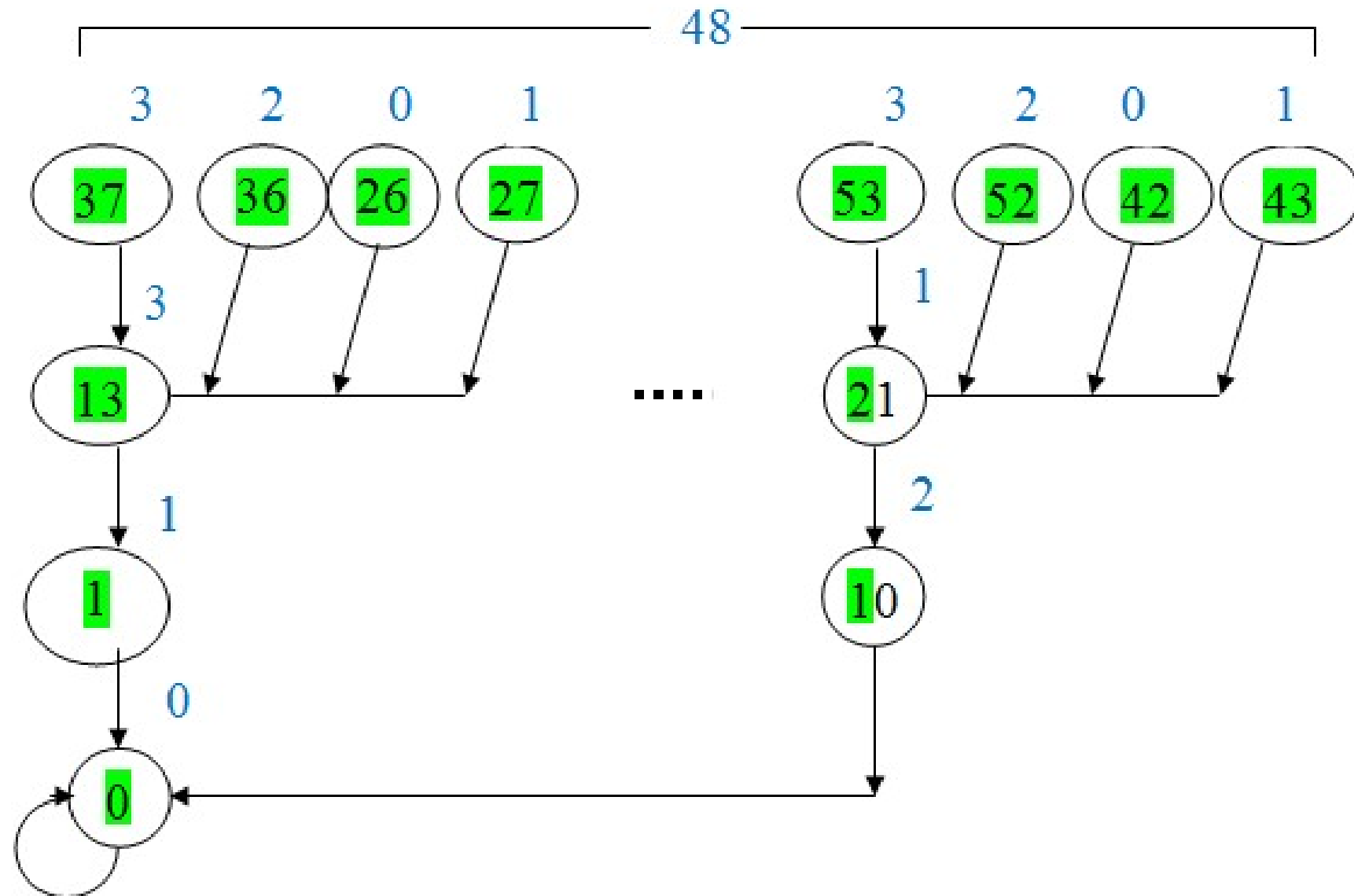
СТРУКТУРЫ ТОКЕНА В РПА



ПРЕОБРАЗОВАТЕЛЬ ТЕГОВ НАСТРАИВАЕМЫЙ



ПРЕОБРАЗОВАТЕЛЬ ТЕГОВ УНИВЕРСАЛЬНЫЙ



РЕЗУЛЬТАТЫ РЕАЛИЗАЦИИ АЛГОРИТМОВ РАСПОЗНАВАНИЯ СЛОВ-КОМАНД

Название алгоритма	Кол-во шагов для dsPic30F	Кол-во шагов для POУ		Коеф. ускорения	
		Вар. 1	Вар. 2	Вар. 1	Вар. 2
Баттеруорт (одна секция)	679	288	-	~2,36	-
Полосовой фильтр (одна полоса)	1428	420	-	3,4	-
Натуральный логарифм	36	12	26	3	~1,38* 4
RASTA фильтр	153	28	-	~5,46	-
Экспоненцирование	36	12	26	3	~1,38* 4
Косинусное ИДПФ	36	12	26	3	~1,38* 4
Рекурсия Дурбина-Скурра	~640	~110	-	~5,8	-
PLP параметры	144	32	-	4,5	-
Витерби (расчет решетки для текущего N *)	91*N-143	99*N	$(1-(8*N+143)/(99*N))*4$		

ЗАКЛЮЧЕНИЕ

- Реализация потокового процессора в виде двух-уровневой архитектуры с ведущим фон-неймановским процессором напозволило подтвердить потенциальную эффективность применения потоковой парадигмы в области ЦОС и обеспечить совместимость с существующими вычислительными средами
- Ориентация ЦСП на самосинхронную схемотехнику позволит реализовать концептуальное единство: самосинхронизация на логическом уровне (по готовности данных) хорошо сочетается самосинхронизацией на аппаратном уровне (по готовности результатов).

КОНТАКТЫ

- ▣ Адрес: Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук (ИПИ РАН), Россия, 119333, Москва, ул. Вавилова, д. 44, корпус 2
- ▣ Директор: академик Соколов И.А.
- ▣ Телефон: +7 (495) 137 34 94
- ▣ Fax: +7 (495) 930 45 05
- ▣ E-mail: ISokolov@ipiran.ru
- ▣ Докладчик: Дьяченко Ю. Г., +7(495) 671-15-20, YStepchenkov@ipiran.ru

Поддержка

Работа выполнена в рамках госзадания по НИР № 0063-2019-0010