

Конвейерное выполнение команд

1 Конвейер команд.

1.1 Этапы выполнения команды.

1.2 Диаграмма конвейера команд.

1.3 Эффект от конвейеризации вычислений.

2 Конфликты в конвейере команд.

2.1 Виды конфликтных ситуаций.

2.2 Структурный риск.

2.3 Риск по данным.

2.4 Риск по управлению.

3 Суперконвейерные процессоры.

- **Знать:** классическую диаграмму работы и основные показатели конвейера команд (ускорение, эффективность и пропускная способность); риски в конвейере команд: структурный, по данным и по управлению; понятие суперковейеризации, достоинства и проблемы суперковейеризации.
- **Уметь:.**
- **Помнить:** о причинах сбоя работы конвейера команд.
- **Литература:** [1, с. 413-446].

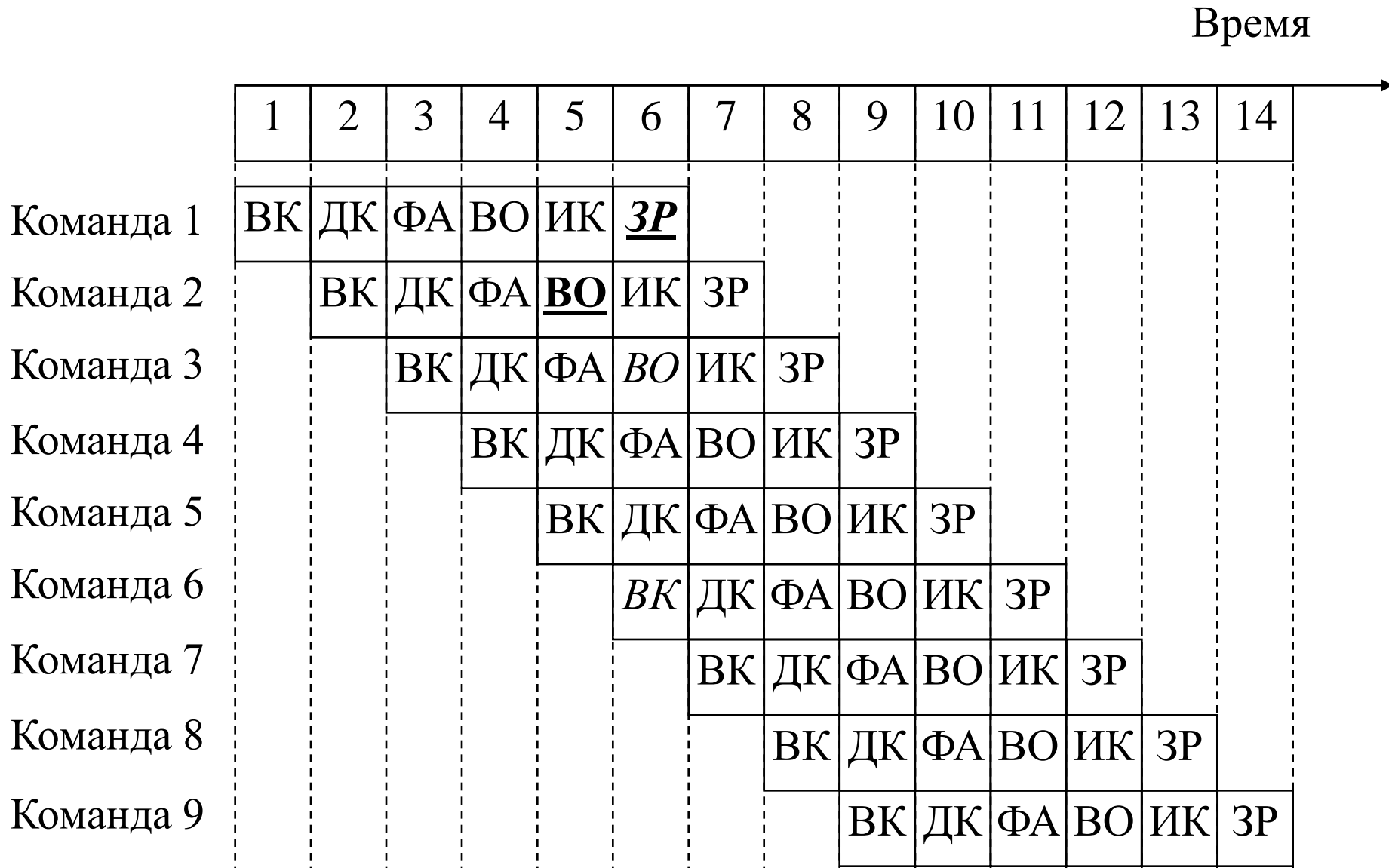
1 Конвейер команд

1.1 Этапы выполнения команды

- Идея конвейера команд была предложена в 1956 году академиком С.А. Лебедевым.
- Поскольку цикл команды представляет собой последовательность этапов, то, возложив реализацию каждого этапа на отдельное устройство, можно получить классическую схему конвейера команд.
- Предположим, что в цикле выполнения команды выделено шесть следующих этапов.

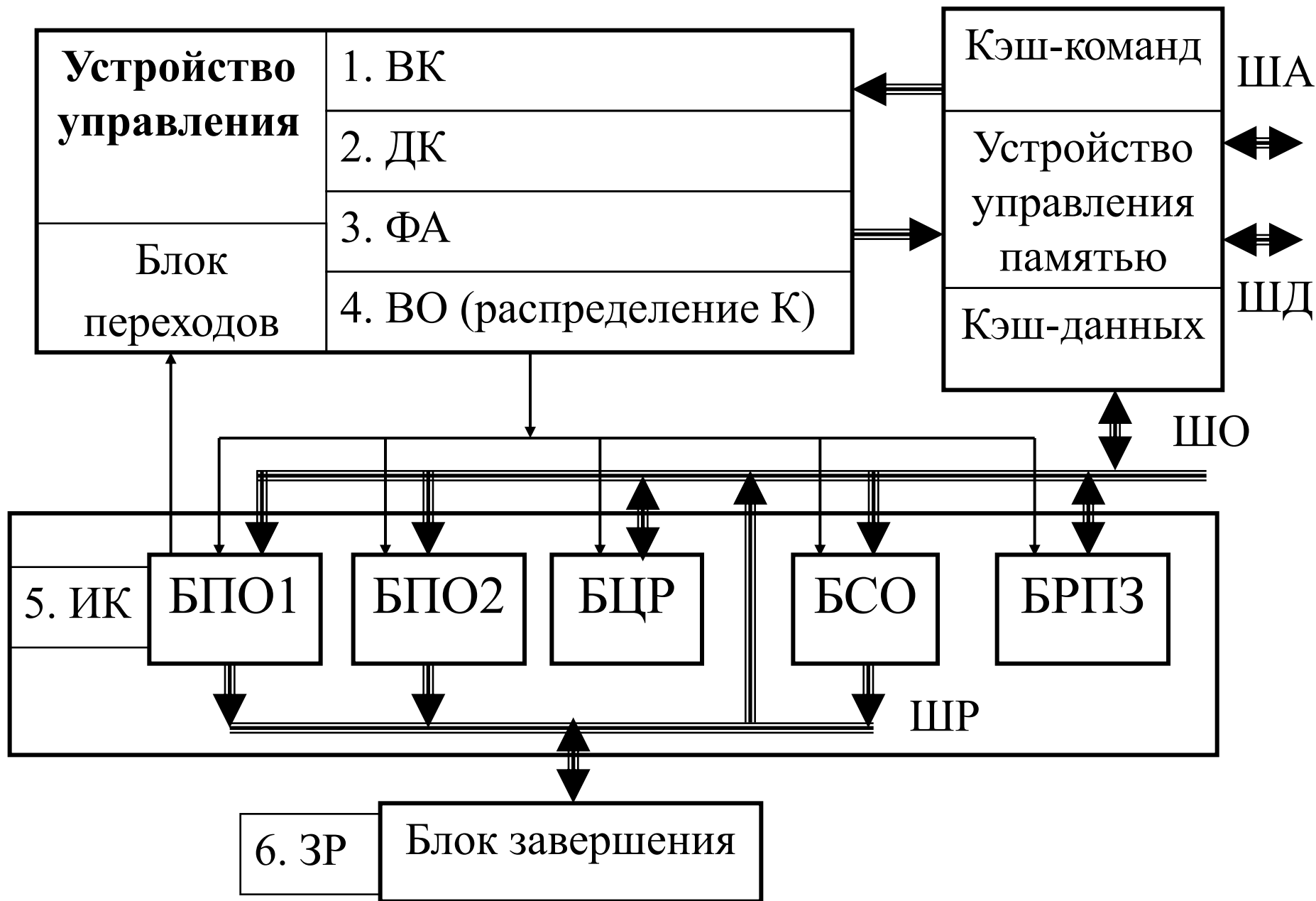
- ***Выборка команды (ВК).*** Чтение очередной команды из памяти и занесение ее в регистр команд.
- ***Декодирование команды (ДК).*** Определение кода операции и способов адресации.
- ***Формирование адресов операндов (ФА).*** Вычисление исполнительных адресов каждого из операндов в соответствии с указанным в команде способом их адресации.
- ***Выборка операндов (ВО).*** Извлечение операндов из памяти. Эта операция не нужна для операндов, находящихся в регистрах.
- ***Исполнение команды (ИК).*** Исполнение указанной операции.
- ***Запись результата (ЗР).*** Занесение результата в память.

1.2 Диаграмма конвейера команд



- На рисунке показана диаграмма работы конвейера команд с шестью ступенями.
- В диаграмме предполагается, что каждая команда обязательно проходит все шесть ступеней, хотя это случай не совсем типичен.
- Так, команда загрузки регистра не требует этапа записи результата (ЗР).
- Кроме того, на диаграмме принято, что все этапы могут выполняться одновременно.
- Без конвейеризации выполнение девяти команд заняло бы $9 \times 6 = 54$ единицы времени. Использование конвейера позволяет сократить время обработки до 14 единиц.

Пример структуры процессора



- На рисунке показана упрощенная схема конвейерного процессора с шестью ступенями, на которой приняты следующие обозначения.
- БПО1, БПО2 – блоки простых операций (логических, целочисленных арифметических и др.).
- БЦР – блок целочисленных регистров.
- БСО – блок сложных операций (операций с плавающей запятой).
- БРПЗ – блок регистров с плавающей запятой.
- ШО – шина операндов.
- ШР – шина результатов.
- ША – шина адреса, ШД – шина данных.

1.3 Эффект от конвейеризации вычислений

- Эффект от конвейеризации вычислений обычно характеризуют тремя показателями:
ускорением, эффективностью и производительностью.
- *Ускорение S*. Под ускорением понимают отношение времени обработки без конвейера и при его наличии.

Теоретически наилучшее время обработки входного потока из N значений на конвейере с K ступенями и тактовым периодом t можно определить выражением $T=(K+(N-1))t$.

Формула отражает тот факт, что до появления на выходе конвейера результата обработки первого элемента должно пройти K тактов, а последующие результаты будут следовать в каждом такте.

В процессоре без конвейера общее время выполнения составляет NKt . Таким образом, ускорение вычислений S за счет конвейеризации вычислений можно описать формулой

$$S = \frac{NKt}{(K + (N - 1))t} = \frac{NK}{(K + (N - 1))}.$$

При $N \rightarrow \infty$ ускорение стремится к величине, равной количеству ступеней в конвейере.

- **Эффективность E .** Эффективность – это доля ускорения, приходящаяся на одну ступень конвейера:

$$E = \frac{S}{K} = \frac{N}{(K + (N - 1))}.$$

- **Пропускная способность P .** Иначе ее называют производительностью, определяется как эффективность, деленная на длительность тактового периода:

$$P = \frac{E}{t} = \frac{N}{t(K + (N - 1))}.$$

При $N \rightarrow \infty$ эффективность стремится к единице, а производительность к частоте тактирования конвейера.

2 Конфликты в конвейере команд

2.1 Виды конфликтных ситуаций

- Полученное (см. диаграмму) число 14 характеризует потенциальную производительность конвейера команд, которую на практике в силу возникающих конфликтных ситуаций достичь не удастся.
- Конфликтные ситуации в конвейере часто называют рисками, и обусловлены они могут быть тремя причинами:
 - попыткой нескольких команд одновременно обратиться к одному и тому же ресурсу ЭВМ (структурный риск);
 - взаимосвязью команд по данным (риск по данным);
 - неоднозначностью при выборке следующей команды в случае команд перехода (риск по управлению).

2.2 Структурный риск

- Конфликт по ресурсам имеет место, когда несколько команд, находящихся на разных ступенях конвейера, пытаются одновременно использовать один и тот же ресурс, чаще всего – память.
- Так, в типовом цикле команды (см. диаграмму) сразу три этапа (ВК, ВО, ЗР) связаны с обращением к памяти.
- Диаграмма показывает, что все три обращения могут производиться одновременно, однако на практике это не всегда возможно.

Уменьшение структурного риска

- Подобных конфликтов частично удастся избежать за счет модульного построения основной памяти (ОП) и использования кэш-памяти. При этом имеется вероятность того, что команды будут обращаться либо к разным модулям ОП, либо одна из них станет обращаться к ОП, а другая – к кэш-памяти.
- С этих позиций выгоднее разделять кэш-память на кэш-память команд и кэш-память данных.
- Конфликты из-за одновременного обращения к памяти могут и не возникать, поскольку для многих команд ступени выборки операнда и записи результата часто не требуются.
- В целом, влияние структурного риска невелико.

2.3 Риск по данным

Типы конфликтов по данным

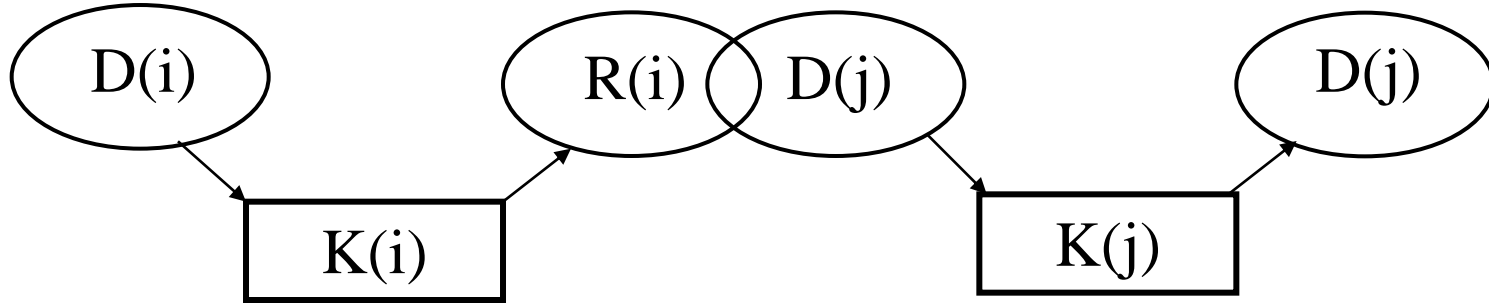
- Риск по данным – это типичная и регулярно возникающая в конвейере команд ситуация. Предположим, что две команды в конвейере (i и j) предусматривают обращение к одной и той же переменной X , причем команда i предшествует команде j .
- В общем случае между этими командами ожидаемы три типа конфликтов по данным:
- *"Чтение после записи" (ЧПЗ)*: команда j читает X до того, как команда i успела записать новое значение X , то есть команда j ошибочно получит старое значение X вместо нового;

- *"Запись после чтения" (ЗПЧ):* команда j записывает новое значение X до того, как команда i успела прочитать X , то есть команда i ошибочно получит новое значение X вместо старого;
- *"Запись после записи" (ЗПЗ):* команда j записывает новое значение X прежде, чем команда i успела записать в качестве X свое значение, то есть X ошибочно содержит i -е значение вместо j -го.
- Возможен и четвертый случай, когда команда j читает X прежде команды i . Этот случай не вызывает никаких конфликтов, поскольку обе команды получают верное значение X .

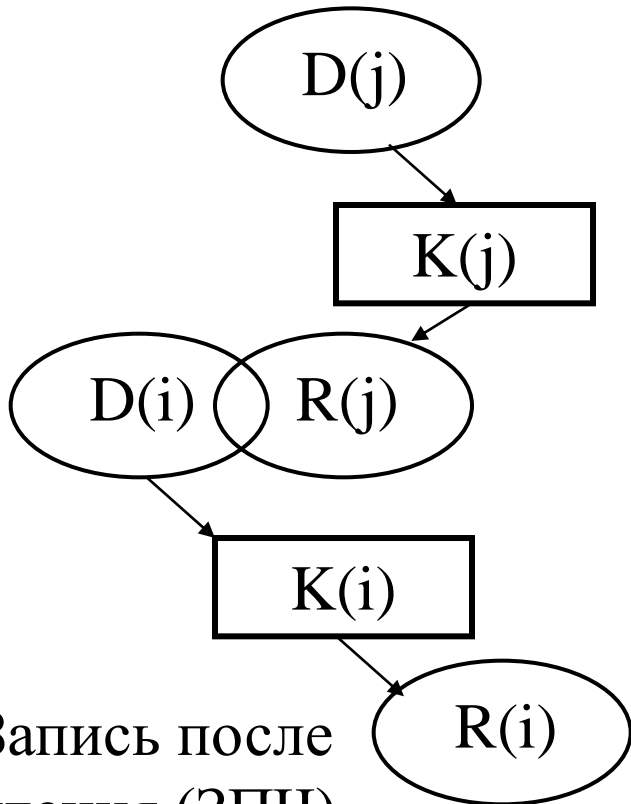
Области определения и значения команд

- *Область определения $D(i)$ команды i – это множество всех объектов (регистров, ячеек памяти и флагов), содержимое которых влияет на исполнение команды. Команда i читает эти объекты.*
- *Область значения $R(j)$ команды j – это множество всех объектов (регистров, ячеек памяти и флагов), содержимое которых может быть изменено в результате исполнения команды. Команда j модифицирует эти объекты.*
- *Условия возникновения помех между командой i и логически более поздней командой j :*
 - $R(i) \cap D(j) \neq \emptyset$ (чтение после записи);
 - $D(i) \cap R(j) \neq \emptyset$ (запись после чтения);
 - $R(i) \cap R(j) \neq \emptyset$ (запись после записи).

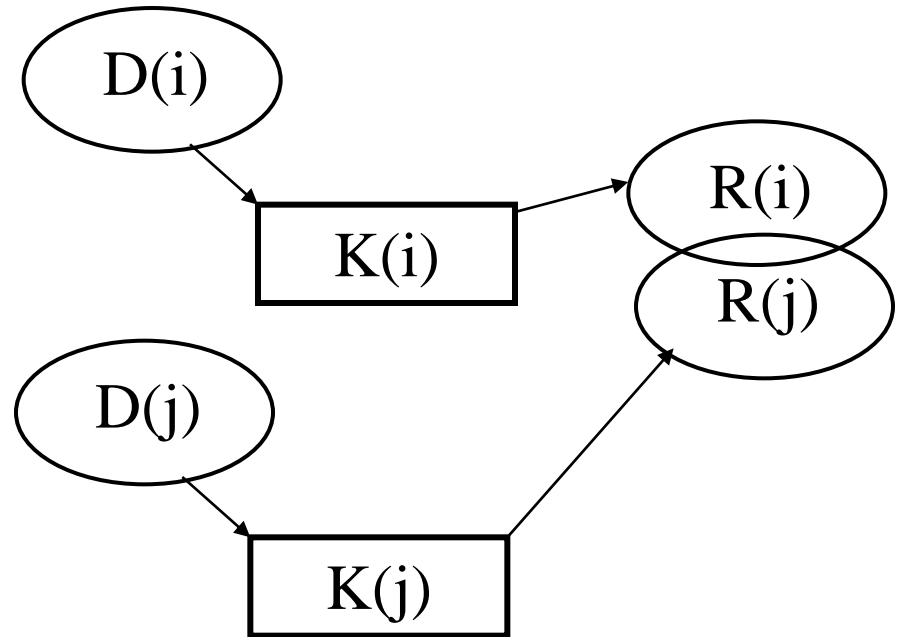
Графическое представление помех



Чтение после записи (ЧПЗ)



Запись после
чтения (ЗПЧ)



Запись после записи (ЗПЗ)

Примеры областей определения и значения команд

Команда	Область определения D	Множество значений R
Типа сложения (СЛ)	Ячейки памяти (операнды). Регистры (операнды). Регистры (флаги)	Регистры (результаты). Регистры (флаги)
Запись (ЗП)	Регистры (операнды). Регистры (флаги)	Ячейки памяти
БП	Регистры	Счетчик команд
УП	Регистры (флаги)	Счетчик команд

Примеры возможных помех

К(i)	К(j)			
	Типа СЛ	ЗП	БП	УП
Типа СЛ	ЧПЗ, ЗПЧ, ЗПЗ	ЧПЗ, ЗПЧ	ЧПЗ	ЧПЗ
ЗП	ЧПЗ, ЗПЧ	ЗПЗ	—	—
БП	ЗПЧ	—	ЗПЗ	ЗПЗ
УП	ЗПЧ	—	ЗПЗ	ЗПЗ

- Наиболее частый вид конфликтов по данным – ЧПЗ, поскольку операция чтения в цикле команды (этап ВО) предшествует операции записи (ЗР).
- По той же причине конфликты типа ЗПЧ большой проблемы не представляют. Сложности появляются, только если структура конвейера допускает запись прежде чтения или если команды в конвейере обрабатываются в последовательности, отличной от предписанной программой.
- Конфликт типа ЗПЗ также не вызывает особых проблем в конвейерах, где команды следуют в порядке, определенном программой, и могут производить запись только на этапе ЗП. В худшем случае, когда одна команда догоняет другую из-за приостановки последней, имеет место конфликт по ресурсу – попытка одновременного доступа к одной и той же ячейке.

Методы борьбы с конфликтами по данным

- Для борьбы с конфликтами по данным применяются как *программные*, так и *аппаратные методы*.
- *Программные методы.* Программные методы ориентированы на устранение самой возможности конфликтов еще на стадии компиляции программы.
- Оптимизирующий компилятор пытается создать такой объектный код, чтобы между командами, склонными к конфликтам, находилось достаточное количество нейтральных в этом плане команд. Если такое не удастся, то между конфликтующими командами компилятор вставляет необходимое количество команд типа "Нет операции".

- **Аппаратные методы.** Фактическое разрешение конфликтов возлагается на аппаратные методы.
- Наиболее очевидным решением является остановка команды j на несколько тактов с тем, чтобы команда i успела завершиться или, по крайней мере, миновать ступень конвейера, вызвавшую конфликт. Соответственно задерживаются и команды, следующие в конвейере за j -й командой. Данную ситуацию называют "пузырьком" в конвейере.
- Иногда приостанавливают только команду j , не задерживая следующие за ней команды. Это более эффективный прием, но его реализация усложняет конвейер. Основные усилия тратятся на устранение конфликтов типа ЧПЗ, поскольку такие конфликты возникают наиболее часто.

2.4 Риск по управлению

- Наибольшие проблемы обеспечения эффективной работы конвейера обусловлены командами, изменяющими естественный порядок вычислений.
- К таким командам (перехода) в частности относятся команды безусловного и условного переходов, обращения к подпрограмме и возврата из подпрограммы и другие. Доля подобных команд в программе оценивается как 10-20% и более.
- Команды вызывают приостановку работы конвейера и очистку его от команд, начало выполнения которых было сделано в предположении, что перехода не будет. Например, на диаграмме команда 3 — это команда условного перехода к команде 15.

Диаграмма конвейера с командой перехода

Время



Издержки перехода

Влияние команды перехода на работу конвейера

Время

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Команда 1	ВК	ДК	ФА	ВО	ИК	<u>ЗР</u>								
Команда 2		ВК	ДК	ФА	<u>ВО</u>	ИК	ЗР							
Команда 3			ВК	ДК	ФА	ВО	ИК	ЗР						
Команда 4				ВК	ДК	ФА	ВО	ИК	ЗР					
Команда 5					ВК	ДК	ФА	ВО	ИК	ЗР				
Команда 6						ВК	ДК	ФА	ВО	ИК	ЗР			
Команда 7							ВК	ДК	ФА	ВО	ИК	ЗР		
Команда 8								ВК	ДК	ФА	ВО	ИК	ЗР	
Команда 9									ВК	ДК	ФА	ВО	ИК	ЗР

Нет перехода

9 команд за 14
тактов

Время

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Команда 1	ВК	ДК	ФА	ВО	ИК	ЗР								
Команда 2		ВК	ДК	ФА	ВО	ИК	ЗР							
Команда 3 (УП15)			ВК	ДК	ФА	ВО	ИК	ЗР						
Команда 4				ВК	ДК	ФА	ВО							
Команда 5					ВК	ДК	ФА							
Команда 6						ВК	ДК							
Команда 7							ВК							
Команда 15								ВК	ДК	ФА	ВО	ИК	ЗР	
Команда 16									ВК	ДК	ФА	ВО	ИК	ЗР

Есть переход

5 команд за 14
тактов

Издержки перехода

- До завершения команды 3 невозможно определить, какая команда (4 или 15) должна выполняться следующей, поэтому в конвейер просто загружается следующая команда последовательности (4) и работа конвейера продолжается.
- На первой диаграмме (см. ранее) показан случай, когда переход не произошёл, и получена максимальная производительность.
- В данном варианте переход имеет место, но об этом не известно до 7 шага. Конвейер должен быть очищен от ненужных команд, выполнявшихся до 8 шага. Лишь на шаге 8 в конвейер поступает нужная команда 15, из-за чего в течение тактов 9-12 не будет завершена ни одна другая команда.
- Это и есть издержки из-за невозможности предвидения исхода команды условного перехода.

Ускорение выборки команды из точки перехода

- Для сокращения задержек, обусловленных выборкой команды из точки перехода, применяются несколько подходов:
- *вычисление исполнительного адреса перехода на ступени декодирования команды;*
- *использование буфера адресов перехода (кэш-память небольшого объема, в которую заносятся исполнительные адреса точек перехода нескольких последних команд, для которых переход имел место. Перед выборкой очередной команды содержимое счетчика команд сравнивается с содержимым тегов буфера и в случае совпадения ранее вычисленный адрес перехода берется из буфера);*

- *использование кэш-памяти для хранения команд, расположенных в точке перехода (это усовершенствованный вариант буфера адресов перехода, позволяющий ускорить выборку команд из точек перехода, что особенно важно для циклических участков);*
- *использование буфера цикла (представляет собой маленькую быстродействующую память, входящую в состав первой ступени конвейера, в которую копируются последовательности выбираемых команд для последующего повторного выполнения).*

3 Суперконвейерные процессоры

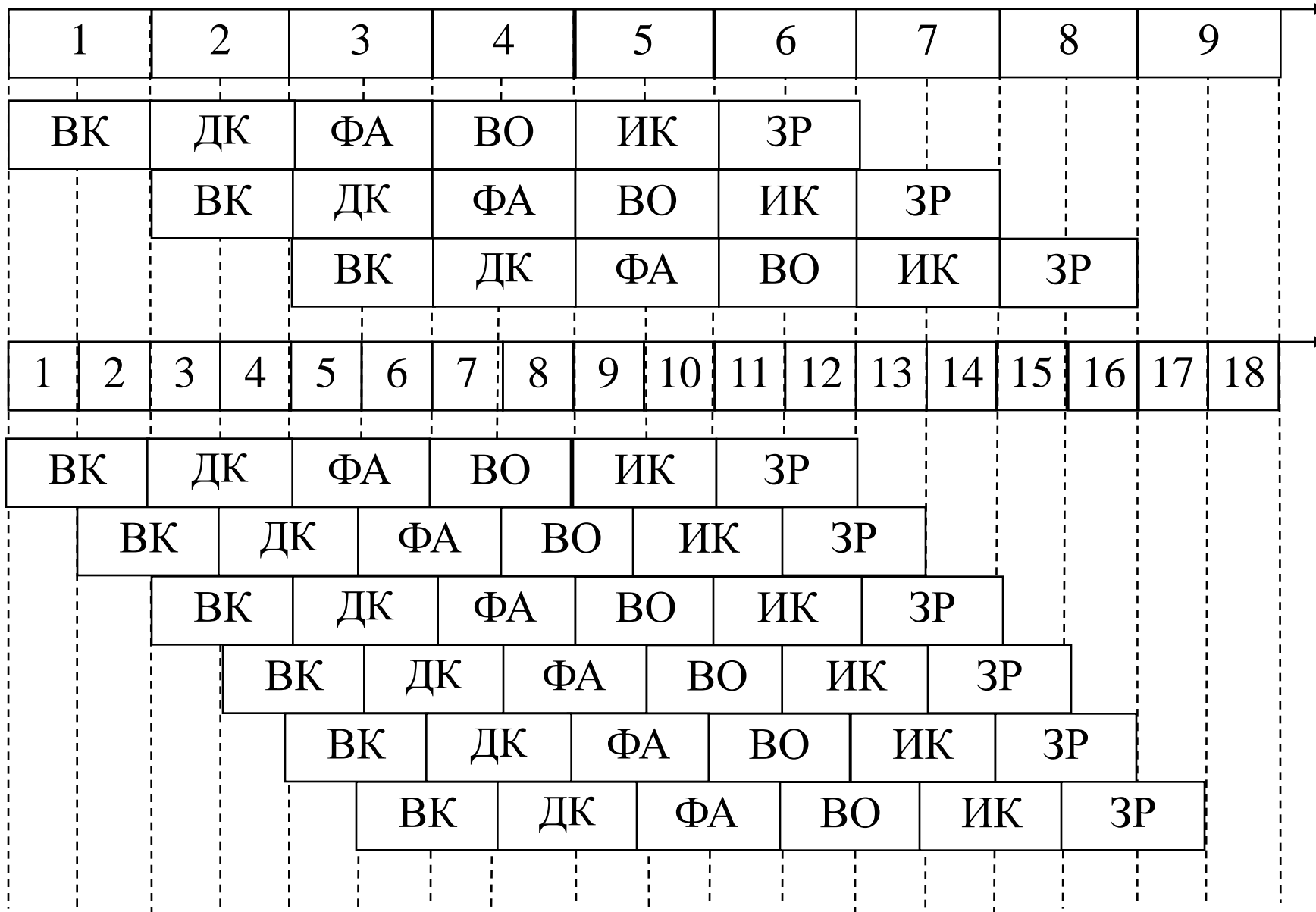
- Эффективность конвейера находится в прямой зависимости от того, с какой частотой на его вход подаются объекты обработки. Добиться n -кратного увеличения темпа работы конвейера можно двумя путями:
 - разбиением каждой ступени конвейера на n "подступеней" при одновременном повышении тактовой частоты внутри конвейера также в n раз;
 - включением в состав процессора n конвейеров, работающих с перекрытием.
- Первый из этих подходов получил название «суперконвейеризация».

Условия и критерий суперконвейерности

- Суперконвейеризация сводится к увеличению количества ступеней конвейера как за счет добавления новых ступеней, так и путем дробления имеющихся ступеней на несколько простых подступеней.
- Главное условие – возможность реализации операции в каждой подступени наиболее простыми техническими средствами, а значит с минимальными затратами времени.
- Вторым, не менее важным, условием является одинаковость задержки на всех ступенях.
- Критерием для причисления процессора к суперконвейерным служит число ступеней в конвейере команд. К суперконвейерным относят процессоры, где таких ступеней больше шести.

Диаграммы конвейеров

Время



Проблемы суперконвейерных процессоров

- Удлинение конвейера команд ведет не только к углублению проблем, характерных для любого конвейера, но и возникновению дополнительных сложностей.
- В длинном конвейере команд возрастает вероятность конфликтов. Дороже встает ошибка предсказания перехода (приходится очищать большее число ступеней, на что требуется больше времени). Усложняется логика взаимодействия ступеней конвейера.
- Тем не менее удастся справиться с большинством этих проблем. Так число ступеней в конвейерах команд процессоров Pentium III – 10, Pentium 4 – 20, P6 – 12.
- Фирма Intel в своих разработках сначала увеличила длину конвейера до 31 ступени (ядро Prescott), а в последующих моделях (процессор Core 2) сократила число ступеней до 14.