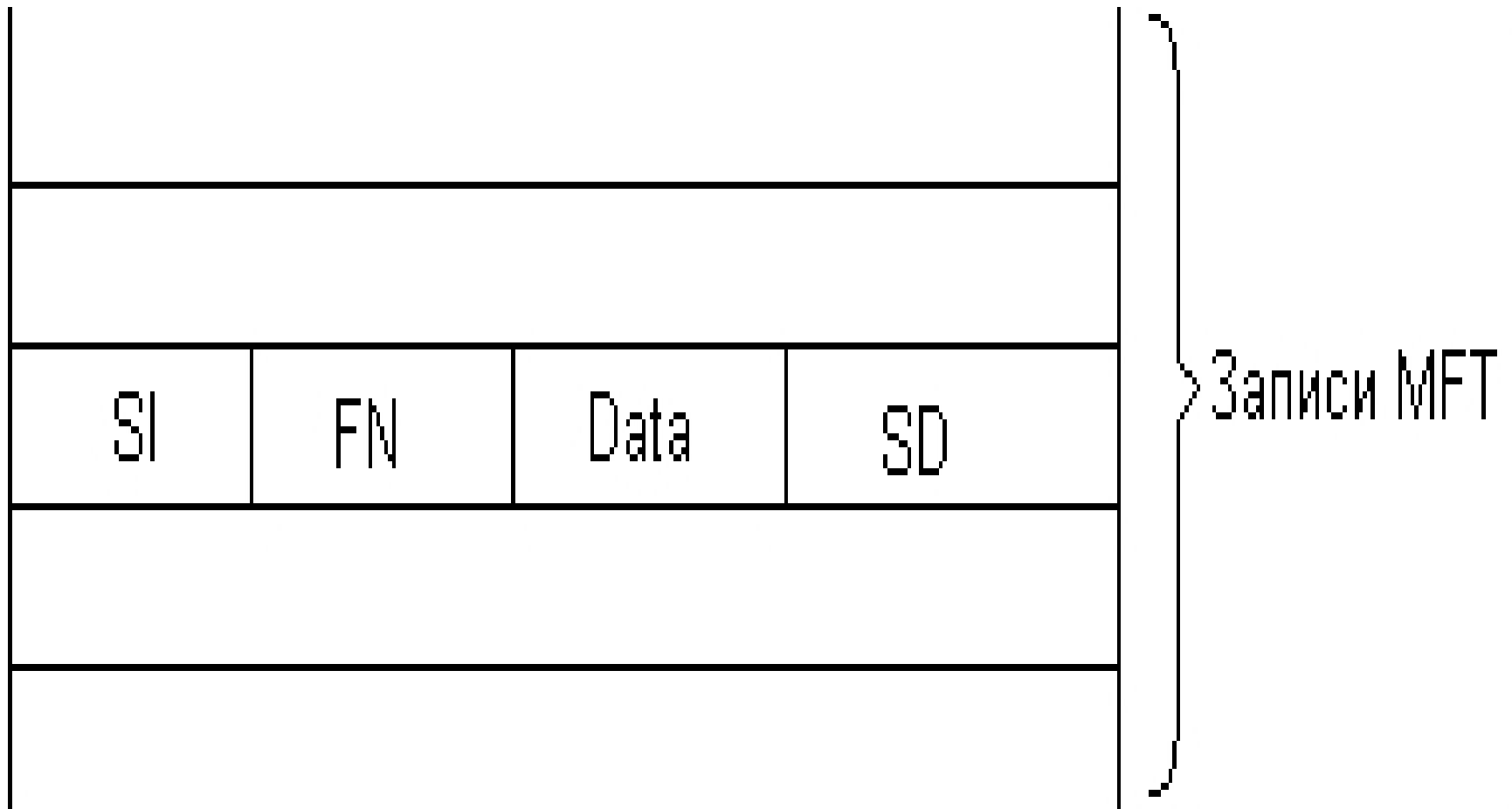




# Небольшой файл NTFS

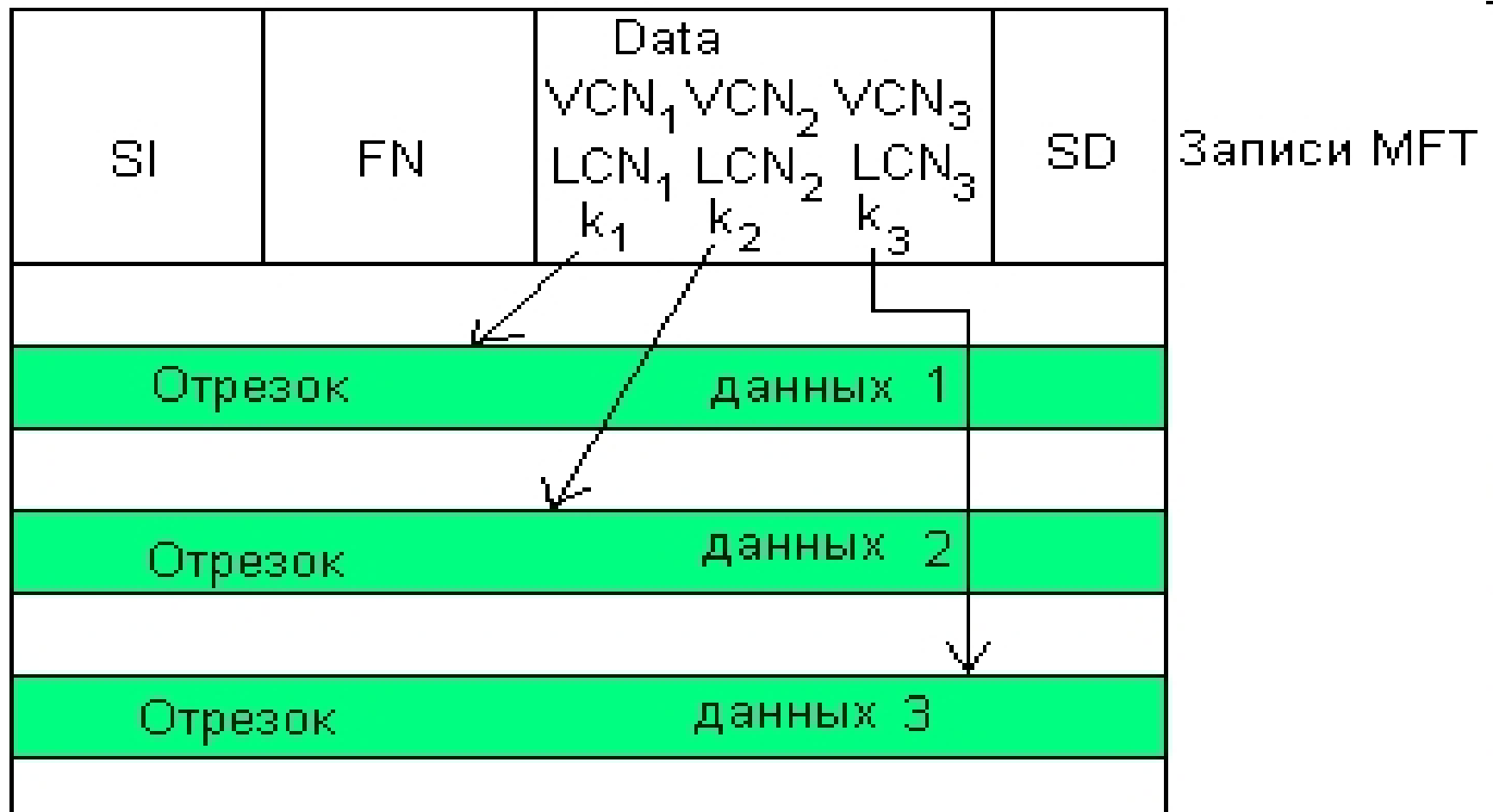


# Атрибуты небольшого файла

- стандартная информация (SI — standard information);
- имя файла (FN — file name);
- данные (Data);
- дескриптор безопасности (SD — security descriptor).

Большие файлы (large). Если данные файла не помещаются в одну запись MET, то этот факт отражается в заголовке атрибута Data, который содержит признак того, что этот атрибут является нерезидентным, то есть находится в отрезках вне таблицы MFT. В этом случае атрибут Data содержит адресную информацию (LCN, VCN, k) каждого отрезка данных

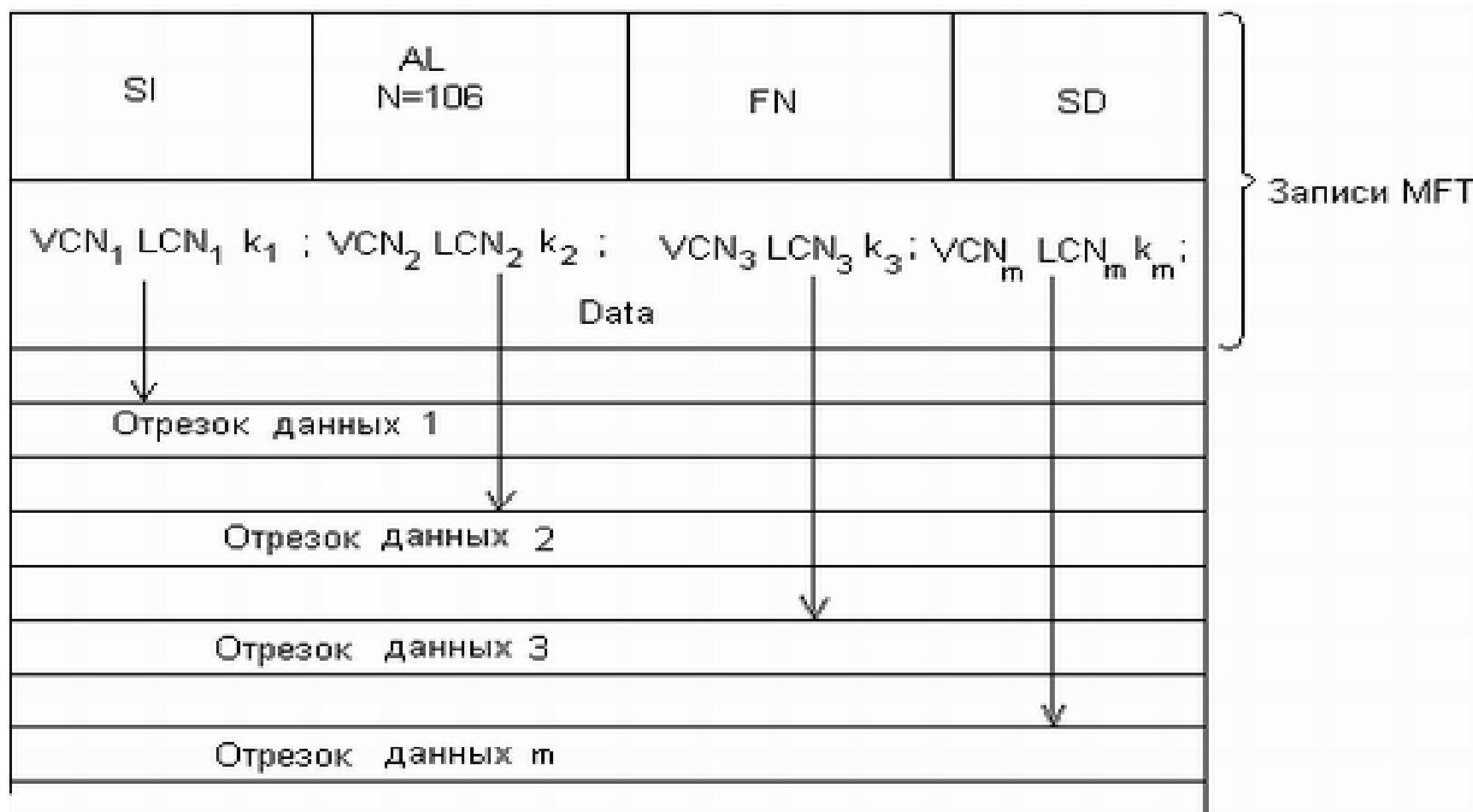
# Большой файл



Порядковый номер кластера в томе NTFS называется логическим номером кластера {Logical Cluster Number, LCN). Файл NTFS также состоит из последовательности кластеров, при этом порядковый номер кластера внутри файла называется виртуальным номером кластера (Virtual Cluster Number, VCN).

Базовая единица распределения дискового пространства для файловой системы NTFS — непрерывная область кластеров, называемая отрезком. В качестве адреса отрезка NTFS использует логический номер его первого кластера, а также количество кластеров в отрезке  $k$ , то есть пара  $(LCN, k)$ . Таким образом, часть файла, помещенная в отрезок и начинающаяся с виртуального кластера  $VCN$ , характеризуется адресом, состоящим из трех чисел:  $(VCN, LCN, k)$ .

# Очень большой файл





Очень большие файлы (huge). Если файл настолько велик, что его атрибут данных, хранящий адреса нерезидентных отрезков данных, не помещается в одной записи, то этот атрибут помещается в другую запись MFT, а ссылка на такой атрибут помещается в основную запись файла. Эта ссылка содержится в атрибуте Attribute List. Сам атрибут данных по-прежнему содержит адреса нерезидентных отрезков данных.

---



Сверхбольшие файлы (extremely huge). Для сверхбольших файлов в атрибуте Attribute List можно указать несколько атрибутов, расположенных в дополнительных записях MFT. Кроме того, можно использовать двойную косвенную адресацию, когда нерезидентный атрибут будет ссылаться на другие нерезидентные атрибуты, поэтому в NTFS не может быть атрибутов слишком большой для системы длины.

SI	FN	IR	IA	SD
		$\langle \text{fl.exe}, N_{\text{fl.exe}} \rangle$ $\langle \text{ltr.exe}, N_{\text{ltr.exe}} \rangle$ $\langle \text{####} \rangle$	$\text{VCN}_1, \text{LCN}_1, k_1$ $\text{VCN}_2, \text{LCN}_2, k_2$ $\text{VCN}_3, \text{LCN}_3, k_3$	
		IR $\langle \text{avia.doc}, N_{\text{avia.doc}} \rangle$  $\langle \text{az.exe}, N_{\text{az.exe}} \rangle$  $\langle \text{emax.exe}, N_{\text{emax.exe}} \rangle$ $\langle \text{####} \rangle$		
		IR $\langle \text{gl.htm}, N_{\text{gl.htm}} \rangle$  $\langle \text{green.com}, N_{\text{green.com}} \rangle$  $\langle \text{caw.doc}, N_{\text{caw.doc}} \rangle$ $\langle \text{####} \rangle$		
		IR $\langle \text{main1.c}, N_{\text{main1.c}} \rangle$  $\vdots$  $\langle \text{zero.txt}, N_{\text{zero.txt}} \rangle$ $\langle \text{####} \rangle$		

Большие каталоги (large indexes). По мере того как каталог растёт, список файлов может потребовать нерезидентной формы хранения. Однако начальная часть списка всегда остается резидентной в корневой записи каталога в таблице MFT

Имена файлов резидентной части списка файлов являются узлами так называемого В-дерева (двоичного дерева). Остальные части списка файлов размещаются вне MFT. Для их поиска используется специальный атрибут Index Allocation, представляющий собой адреса отрезков, хранящих остальные части списка файлов каталога.

Одни части списков являются листьями дерева, а другие являются промежуточными узлами, то есть содержат наряду с именами файлов атрибут Index Allocation, указывающий на списки файлов более низких уровней.

Узлы двоичного дерева делят весь список файлов на несколько групп. Имя каждого файла-узла является именем последнего файла в соответствующей группе. Считается, что имена файлов сравниваются лексикографически, то есть сначала принимаются во внимание коды первых символов двух сравниваемых имен, при этом имя считается меньшим, если код его первого символа имеет меньшее арифметическое значение, при равенстве кодов первых символов сравниваются коды вторых символов имен и т. д.



Например, файл f 1 .exe, являющийся первым узлом двоичного дерева, показанного на рис. 7.25, имеет имя, лексикографически большее имен avia.exe, az.exe,..., etax.exe, образующих первую группу списка имен каталога. Соответственно файл ltr.exe имеет наибольшее имя среди всех имен второй группы, а все файлы с именами, большими ltr.exe, образуют третью и последнюю группу.

Поиск в каталоге уникального имени файла, которым в NTFS является номер основной записи о файле в MFT, по его символьному имени происходит следующим образом. Сначала искомое символьное имя сравнивается с именем первого узла в резидентной части индекса. Если искомое имя меньше, то это означает, что его нужно искать в первой нерезидентной группе, для чего из атрибута Index Allocation извлекается адрес отрезка ( $VCN_j$ ,  $LCN^K$ ), хранящего имена файлов первой группы.

Среди имен этой группы поиск осуществляется прямым перебором имен и сравнением до полного совпадения всех символов искомого имени с хранящимся в каталоге именем. При совпадении из каталога извлекается номер основной записи о файле в MFT и остальные характеристики файла берутся уже оттуда.

- Если же искомое имя больше имени первого узла резидентной части индекса, то его сравнивают с именем второго узла, и если искомое имя меньше, то описанная процедура применяется ко второй нерезидентной группе имен, и т. д.
- В результате вместо перебора большого количества имен (в худшем случае — всех имен каталога) выполняется сравнение с гораздо меньшим количеством имен узлов и имен в одной из групп каталога.

Если одна из групп каталога становится слишком большой, то ее также делят на группы, последние имена каждой новой группы оставляют в исходном нерезидентном атрибуте Index Root, а все остальные имена новых групп переносят в новые нерезидентные атрибуты типа Index Root

К исходному нерезидентному атрибуту Index Root добавляется атрибут размещения индекса, указывающий на отрезки индекса новых групп. Если теперь при поиске искомого имени в нерезидентной части индекса первого уровня какое-либо сравнение показывает, что искомое имя оказывается меньше, чем одно из хранящихся там имен, то это говорит о том, что в данном атрибуте точного сравнения имени уже быть не может и нужно перейти к подгруппе имен следующего уровня дерева.