

Analysis of possible join orders for Borrower |X| BookAuthor |X| Checkout

Assumptions:

$n(\text{Borrower}) = 2000$	$V(\text{borrowerID}, \text{Borrower}) = 2000$ (since a superkey)
$n(\text{BookAuthor}) = 10,000$	$V(\text{callNumber}, \text{BookAuthor}) = 5000$ (assumed)
$n(\text{Checkout}) = 1000$	$V(\text{borrowerID}, \text{Checkout}) = 100$ (assumed)
	$V(\text{callNumber}, \text{Checkout}) = 500$ (assumed)
	- $V(\text{borrowerID} + \text{callNumber}, \text{Checkout}) =$ $500 * 100 = 50,000$

Analysis of (Borrower |X| BookAuthor) |X| Checkout

- A temporary table (t) is created for Borrower |X| BookAuthor:
 - Cartesian product $n_t = 2000 * 10,000 = 20,000,000$
 $V(\text{borrowerID}, t) = 2000$ (all come from Borrower)
 $V(\text{callNumber}, t) = 5000$ (all come from BookAuthor)
 - $V(\text{borrowerID} + \text{callNumber}, t) =$
 $2000 * 5000 = 10,000,000$
- Result table from second join (r) = t |X| Checkout
 - join on BorrowerID + callNumber
 $n_r = 20,000,000 * 1000 / \max(50,000, 10,000,000) = 2000$

Analysis of Borrower |X| (BookAuthor |X| Checkout)

- A temporary table (t) is created for BookAuthor |X| Checkout
 - Join on callNumber: $n_t = 10,000 * 1000 / \max(5000, 500) = 2000$
 $V(\text{borrowerID}, t) = 100$ (all come from Checkout)
- Result table from second join (r) = Borrower |X| t
 - join on BorrowerID
 $n_r = 2000 * 2000 / \max(2000, 100) = 2000$

Analysis of (Borrower |X| Checkout) |X| BookAuthor

- A temporary table (t) is created for Borrower |X| Checkout
 - Join on borrowerID: $n_t = 2000 * 1000 / \max(2000, 100) = 1000$
 $V(\text{callNumber}, t) = 500$ (all come from Checkout)
- Result table from second join (r) = t |X| BookAuthor
 - join on callNumber
 $n_r = 1000 * 10,000 / \max(500, 5000) = 2000$

All three orders estimate the same size result table (which should be the case), but the third order produces a smaller temporary table so we prefer that!