Merging with covers

$C : \langle c_1, c_2, \ldots, c_n \rangle$ sorted

Using c_i, c_{i+1}

i^{th} interval : $[c_i, c_{i+1})$

0^{th} interval : $(-\infty, c_1)$

n^{th} interval $[c_n, \infty)$

- sorted array C is a d-cover of sorted array A if $\leq d$ elements of A fall in any interval of C

A & B sorted arrays

C is a d-cover of A & B

$C \to A$ and $C \to B$ ✓

for $x \in C$

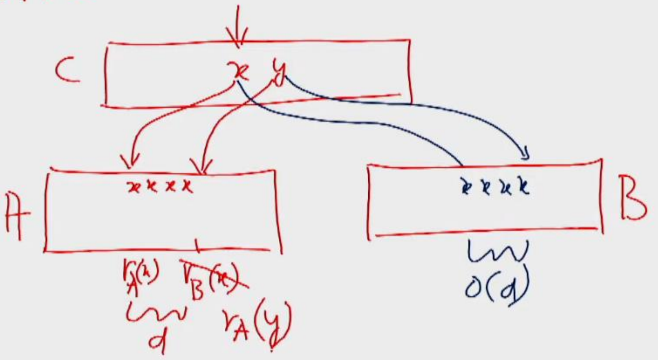
$r_A(x)$: the rank of x in A

$r_B(x)$: ———————— B

Merge A & B

- $|C|$ processors

Consider an interval $[x, y)$ of C



- $|A|$ processors

  $z \in A$. [z knows x is in charge of z

  z goes to C & obtains

  $r_B(x)$ and $r_B(y)$

  Consider positions

  $r_B(x)$ to $r_B(y)-1$ in B

z goes to B, $r_B(x)$

& searches for ~~x~~ z sequentially

$O(d)$. $z \to B$

$A \to B$ $O(d)$ time

- $|B|$ processors — $O(d)$ time $B \to A$

$|A| + |B| + |C|$ processors

$A \to B$, $B \to A$ can be computed in $O(d)$ time on a CREW PRAM

if $d = O(1)$ the time is $O(1)$

Claim 1 $\forall t > 0$, $S_{t-1}(u)$ is a 3-cover of $S_t(u)$

$h = 1$

Claim 2

$\forall t > 0$, h consecutive intervals of $S_{t-1}(u)$. contain at most $2h+1$ elements of $S_t(u)$

If ~~t~~ u becomes full at stage no. $(t-2)$ or earlier

/* $3k \leq t-2$ */

u emits every 4^{th}, 2^{nd}, every after this

[ )

u | y z a b c d e f g h i j

$\uparrow \; \uparrow \; \uparrow \; \uparrow \; \uparrow \; \uparrow \; \uparrow \; \uparrow \; \uparrow \; \uparrow \; \uparrow \; \uparrow$

between any two cons. elements of the old sample there are exactly one new element in the new sample

u becomes full at step $(t-1)$ or later

$\rightarrow$ sampling rate at u is 4

Take h cons. ints. of $S_{t-1}(u)$

$4h$ —— " —— $C_{t-2}(u)$

$= S_{t-2}(v) \cup S_{t-2}(w)$

Say i —— " —— $S_{t-2}(v)$ overlap these $4h$

j —— " —— $S_{t-2}(w)$ —— " ——

$\boxed{i+j = 4h+1}$

by ind., i cons. int. of $S_{t-2}(v)$ has $\leq 2i+1$ elements of $S_{t-1}(v)$

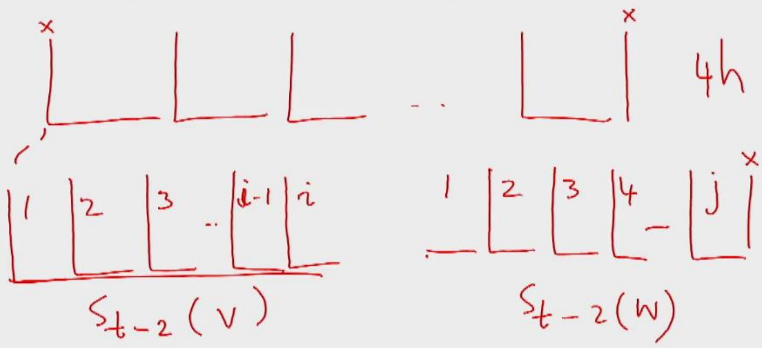
j —"— $S_{t-2}(w)$ has $\leq 2j+1$ —"— $S_{t-1}(w)$

the $4h$ —"— $C_{t-2}(u)$ has

$\leq 2(i+j)+2 = 8h+4$ items $C_{t-1}(u)$

h —"— $S_{t-1}(u)$ contains $\leq 2h+1$ items $S_t(u)$

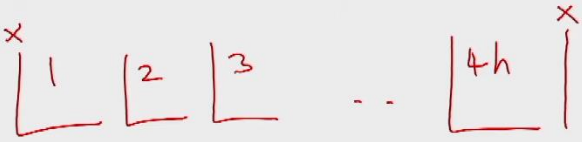
<u>Case 1</u>



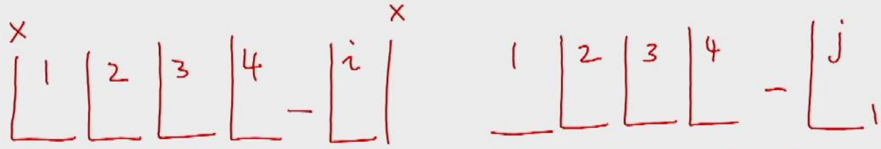
$4h$ cons. int of $C_{t-2}(u)$

$$\underline{i+j = 4h+1} \quad \checkmark$$

Case 2



$C_{t-2}(u)$



$S_{t-2}(v)$ $\qquad$ $S_{t-2}(w)$

$i+1$ $\qquad$ $j-1$

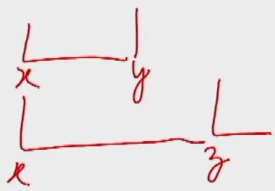
$\underline{i+j = 4h+1}$ ✓

Claim 3

$\forall t>0$, $C_{t-1}(u)$ is a 3-cover of $S_t(v)$ and $S_t(w)$ both.

$$C_{t-1}(u) = S_{t-1}(v) \cup S_{t-1}(w)$$

$$S_{t-1}(v) \xrightarrow{3} S_t(v) \qquad S_{t-1}(w) \xrightarrow{3} S_t(w)$$



$C_{t-1}(u)$ is a 3-cover of $S_t(v)$ & $S_t(w)$

the old cache of the parent is a 3-cover of the new samples of the children

Stage t for u at level k

Step 1.1 /* $C_{t-2}(u) \to C_{t-1}(u)$ */

Draw samples from $C_{t-1}(u)$ to $S_t(u)$

step 1.2 Rank $S_{t-1}(u) \to S_t(u)$

$x \in S_{t-1}(u)$. knows $x \to C_{t-2}(u)$

$x \to C_{t-1}(u)$

$x \to S_t(u)$

Step 1.3

$C_{t-1}(u) \rightarrow S_t(v)$ and $S_t(w)$

$C_{t-1}(u) = S_{t-1}(v) \cup S_{t-1}(w)$

$x \in C_{t-1}(u) \qquad x \in S_{t-1}(v)$

$x \rightarrow S_{t-1}(v)$

$x \rightarrow S_{t-1}(w)$

$S_{t-1}(v) \rightarrow S_t(v)$

$S_{t-1}(w) \rightarrow S_t(w)$

$x \rightarrow S_t(v)$ and $S_t(w)$

$O(1)$ with one processor for x

one prcr / cache of every node

<u>$O(1)$ time</u>

Step 2.1

$C_t(u) \leftarrow \underline{S_t(v) \cup S_t(w)}$

merging with covers

$C_{t-1}(u) \to S_t(v)$

and $S_t(w)$

1 pr / cache item

$O(1)$ time

Step 2.2 Rank $C_{t-1}(u) \to C_t(u)$

$x \in C_{t-1}(u)$

$x \to S_t(v)$

$x \to S_t(w)$

$x \to C_t(u)$

$O(1)$ time if we have as many processors as there are elements in caches & samples

Live node

a node is live if its cache nonempty & non full or it continues to emit samples

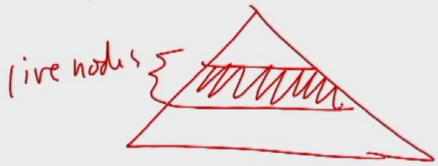
level k node $\Big]$ $\underbrace{2k+1 \quad 3k \quad 3k+3}$

samples

in any stage t

the live nodes are those at levels k where

$$2k+1 \leq t \leq 3k+3$$



elements in the live band is $O(n)$

$O(n)$ processors each stage runs in $O(1)$ time

n pr. $O(1)$ time/stage

n pr. $O(\log n)$ time