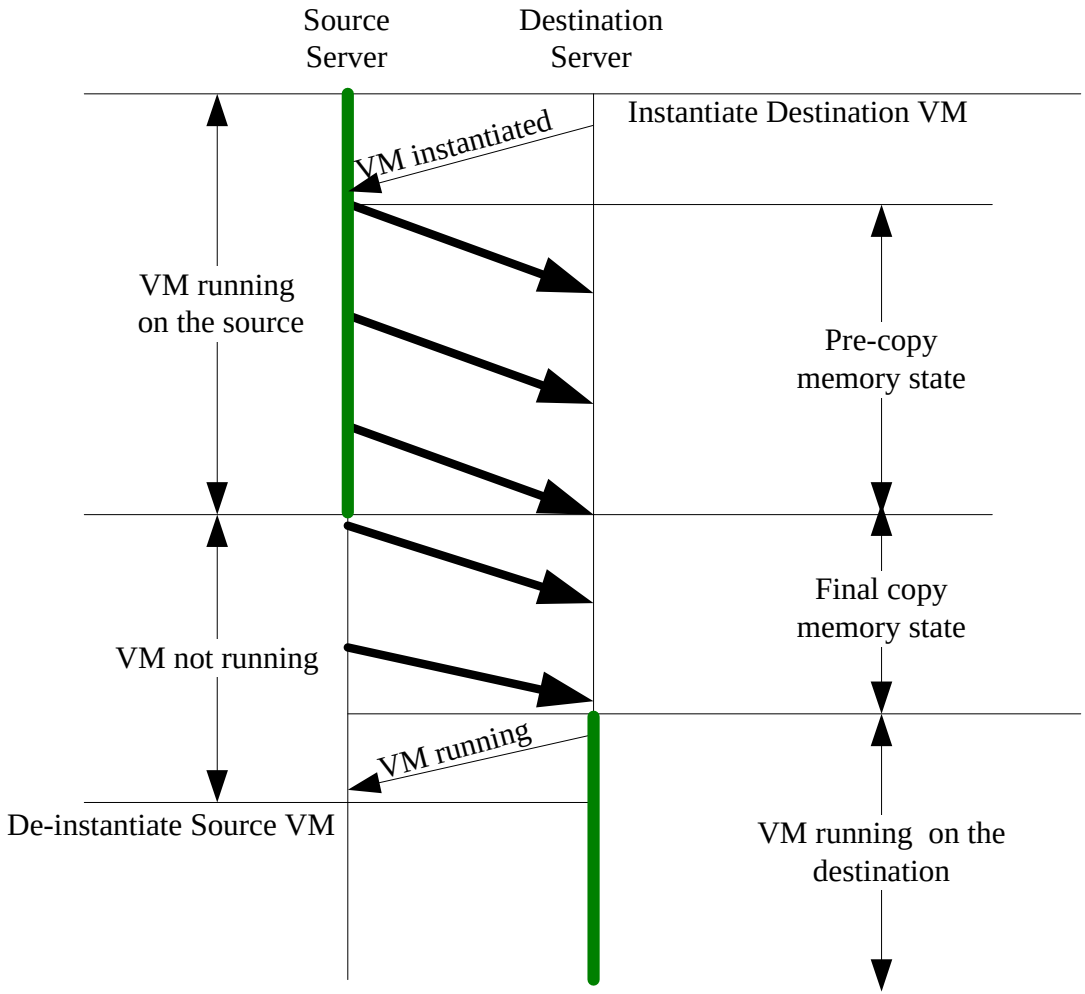
Dynamic Information Migration
June 2011

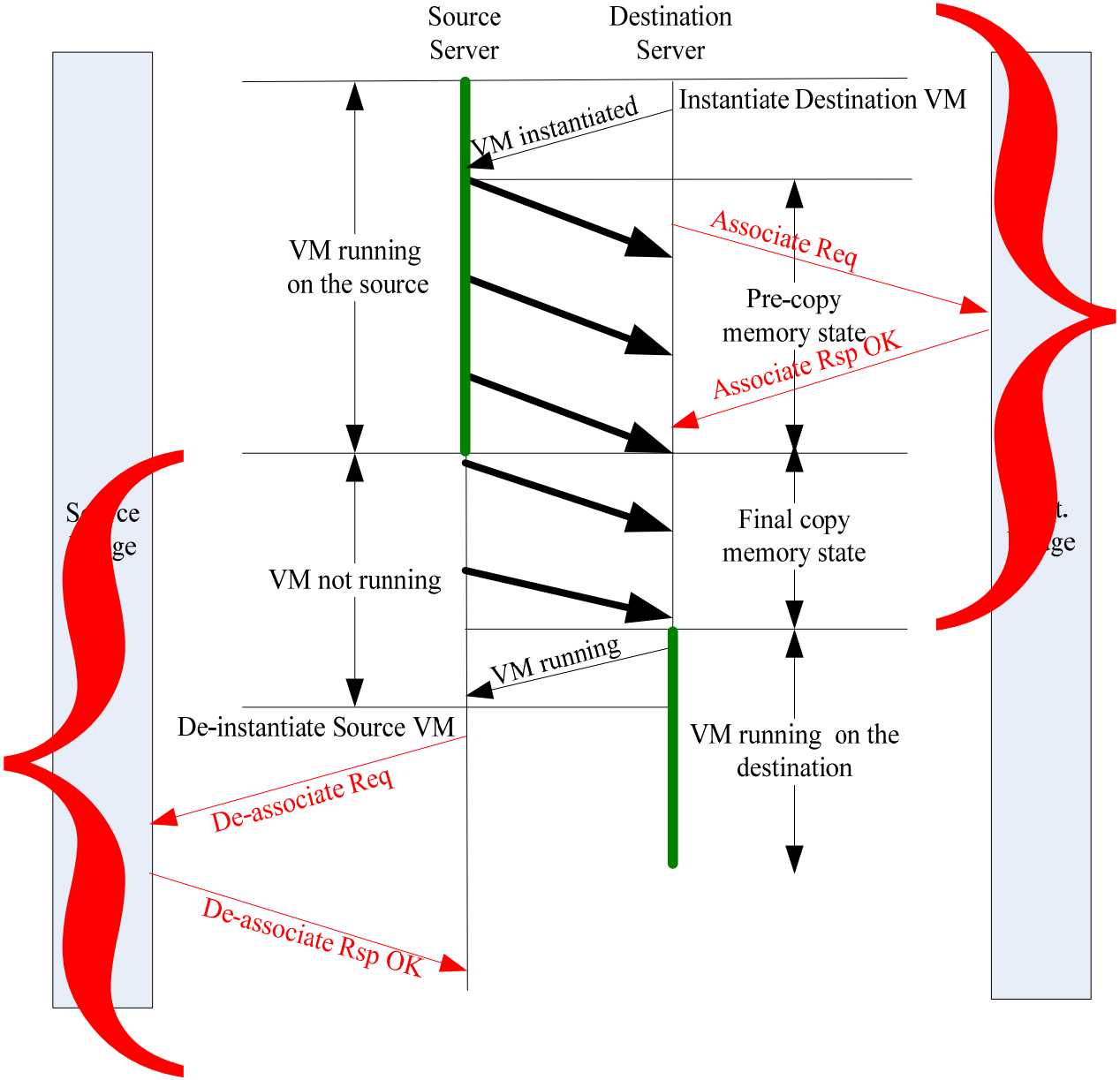
Gu Yingjie
Ben Mack-Crane
Bob Sultan (bsultan@huawei.com)

VM migration reference diagram

- Copy of VM state to destination VM cannot occur until destination VM has been instantiated;
- Destination hypervisor must notify source hypervisor that destination VM is instantiated before copy can start;
- Copy of VM state information is shown in bold arrows;
- Transfer of final (dynamic) portion of VM state cannot begin until source VM has stopped executing (but is still instantiated);
- Assume that VM state copy includes an indication signifying that the copy is complete;
- Destination VM can start executing after the copy is complete;
- Notification is sent from destination hypervisor to source hypervisor to indicate that destination VM is running;
- When source learns that destination VM is running, there is no longer the possibility of backout so source VM can be de-instantiated.



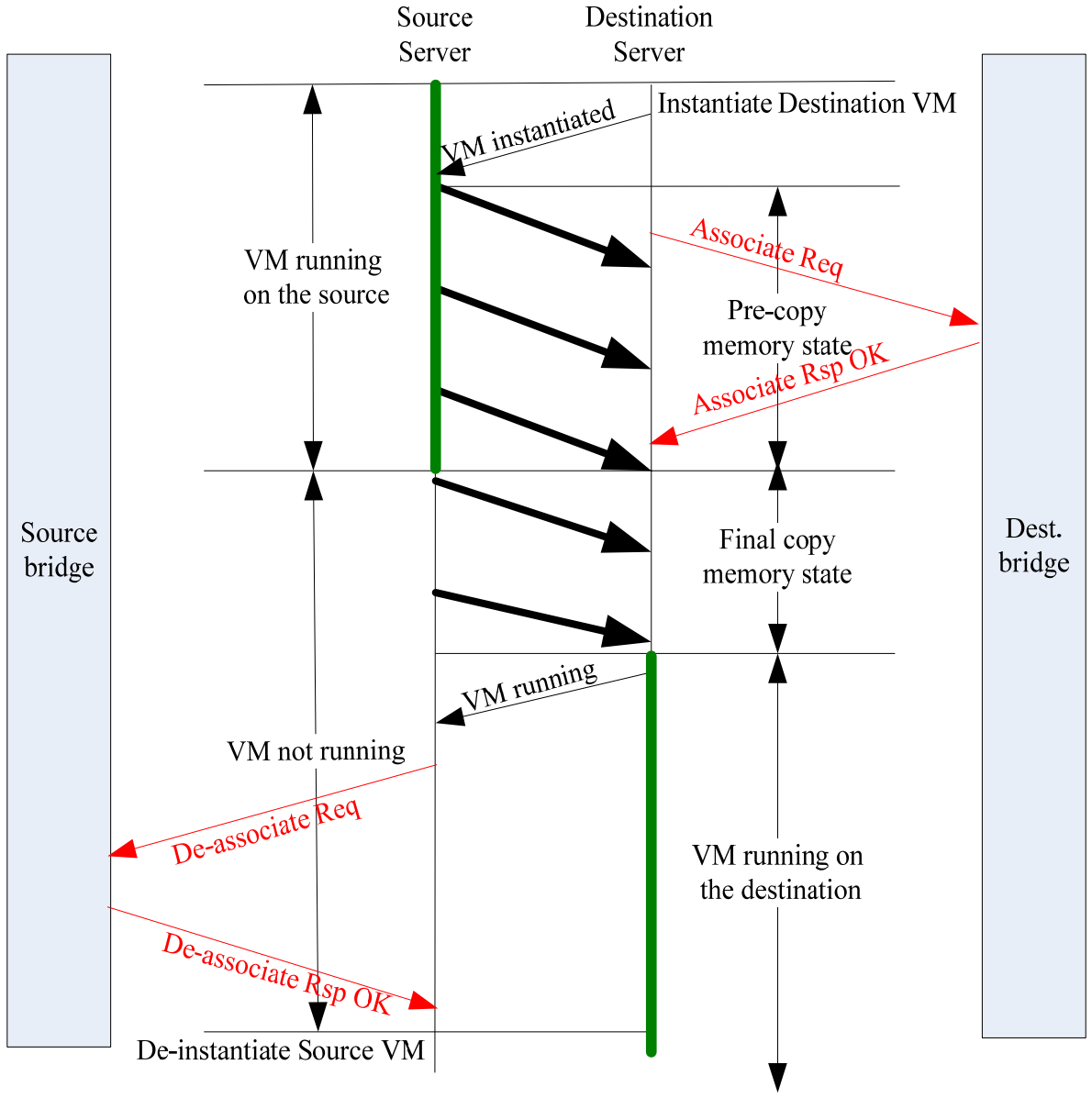
Associate/Deassociate in diagram



Associate must complete before the VM starts running on destination.

Deassociate request can be made anytime after VM stops. Here we show De-associate after destination VM is running so there is no possibility that backout will be needed.

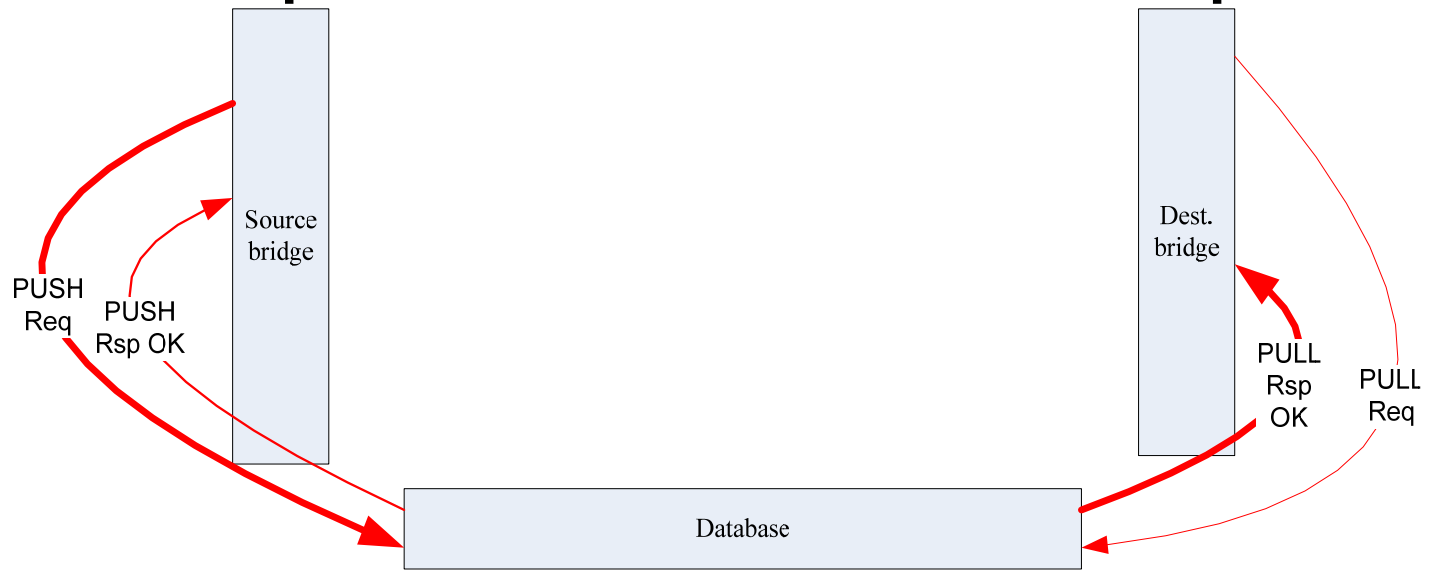
Or? (minor point)



De-instantiate of source VM should not occur until after de-associate completes??? Are Association and VM Instantiation are necessarily related?

Associate should not be requested until VM has been instantiated?? Again, I don't think this is necessary, but maybe someone can clarify.

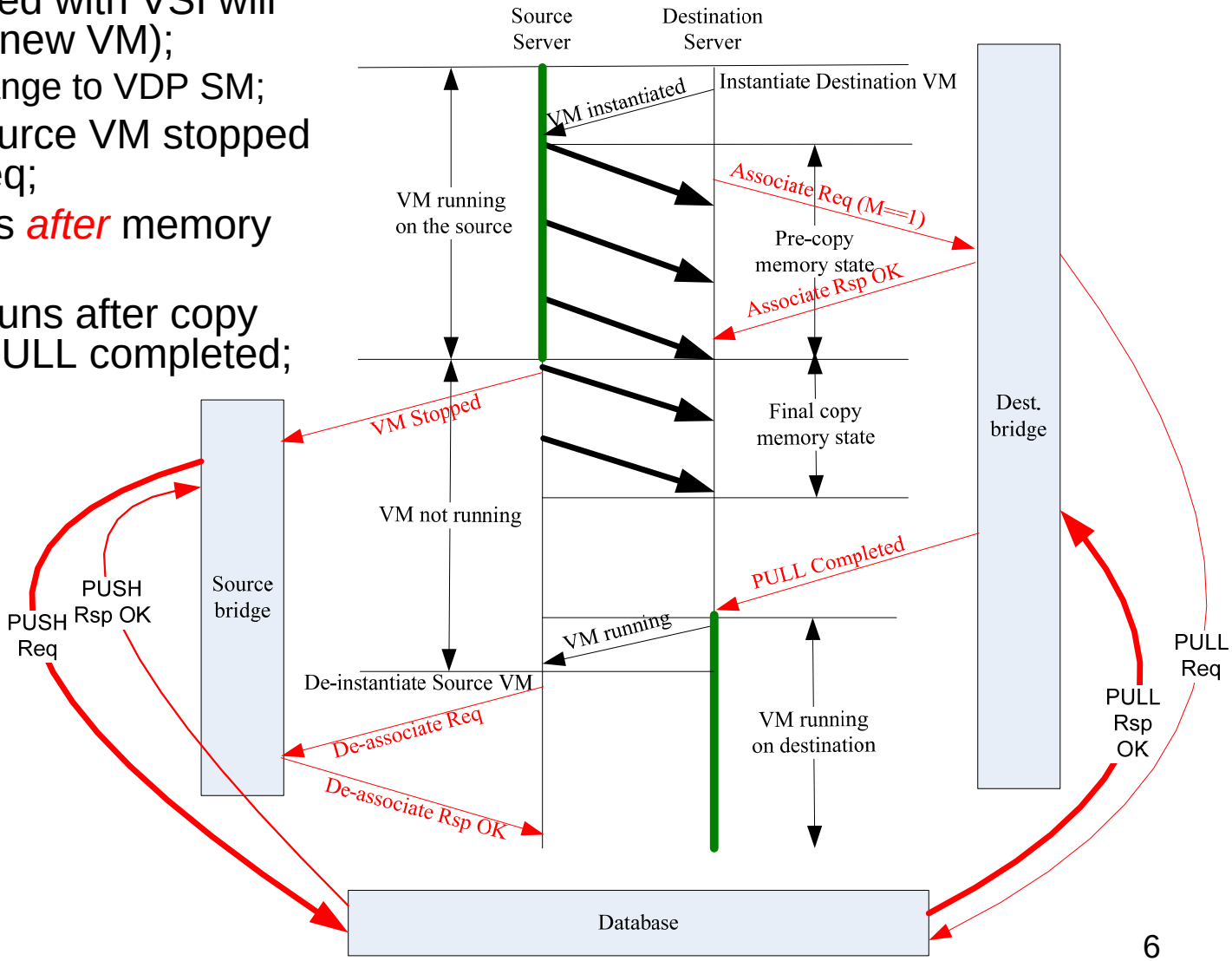
Assumptions for Database Update



- Database PUSH/PULL are asynchronous activities;
  - i.e., PULL request can precede PUSH;
- PUSH can occur *anytime* after notification is received that the source VM has stopped;
- PULL can occur *anytime* after notification is received that the destination VM is migrated (vs. new);
- PUSH Rsp can occur *anytime* after PUSH Req;
- PULL Rsp occurs after PUSH and PULL Requests;
- Destination VM cannot be started until Association and PULL Rsp completes;

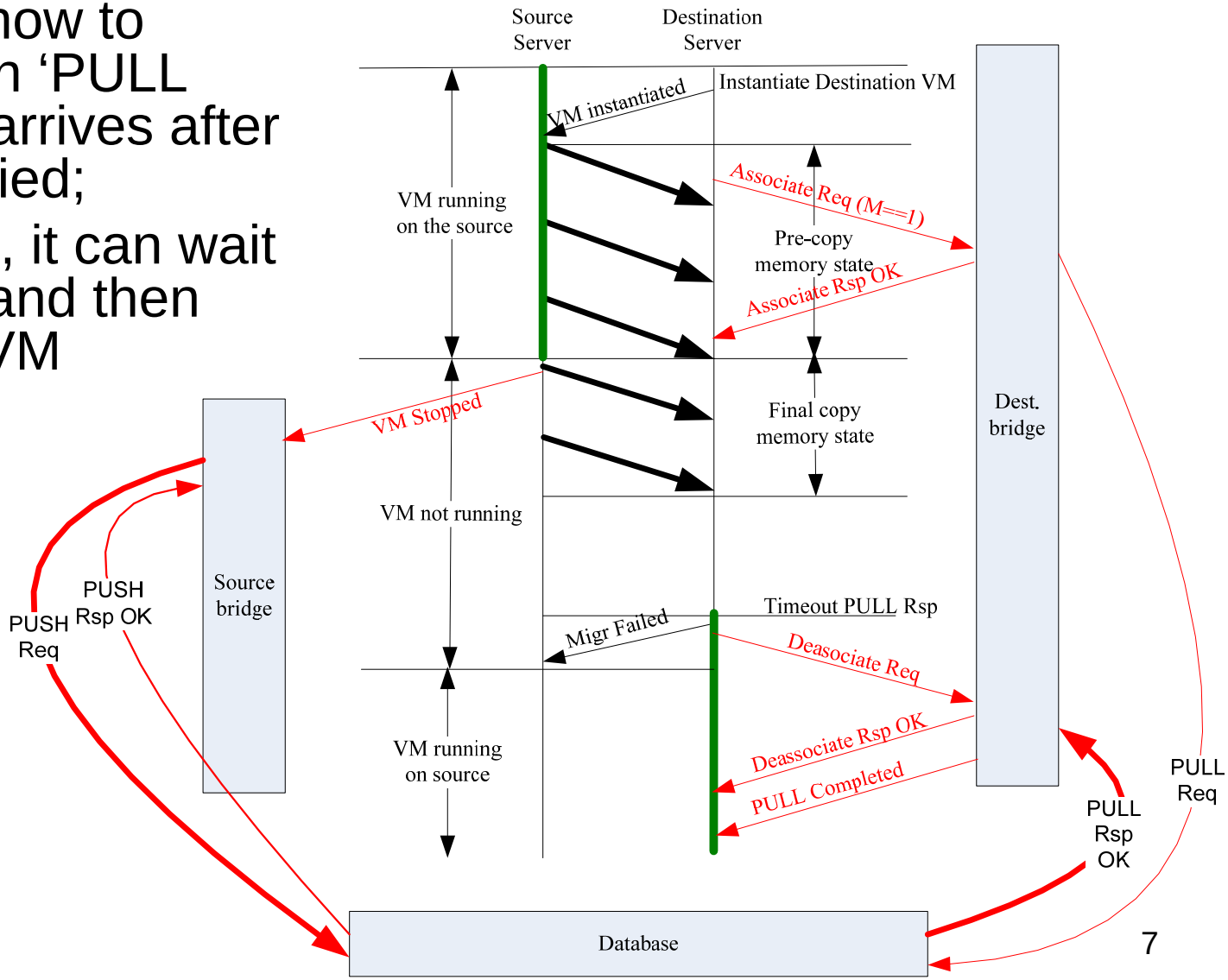
If PULL completes *after* memory copy

- Associate carries M-bit indicating that VM associated with VSI will be migrated (vs. new VM);
  - Implies no change to VDP SM;
- Notification of source VM stopped triggers PUSH req;
- PULL Rsp arrives *after* memory copy completes;
- Destination VM runs after copy completed *and* PULL completed;



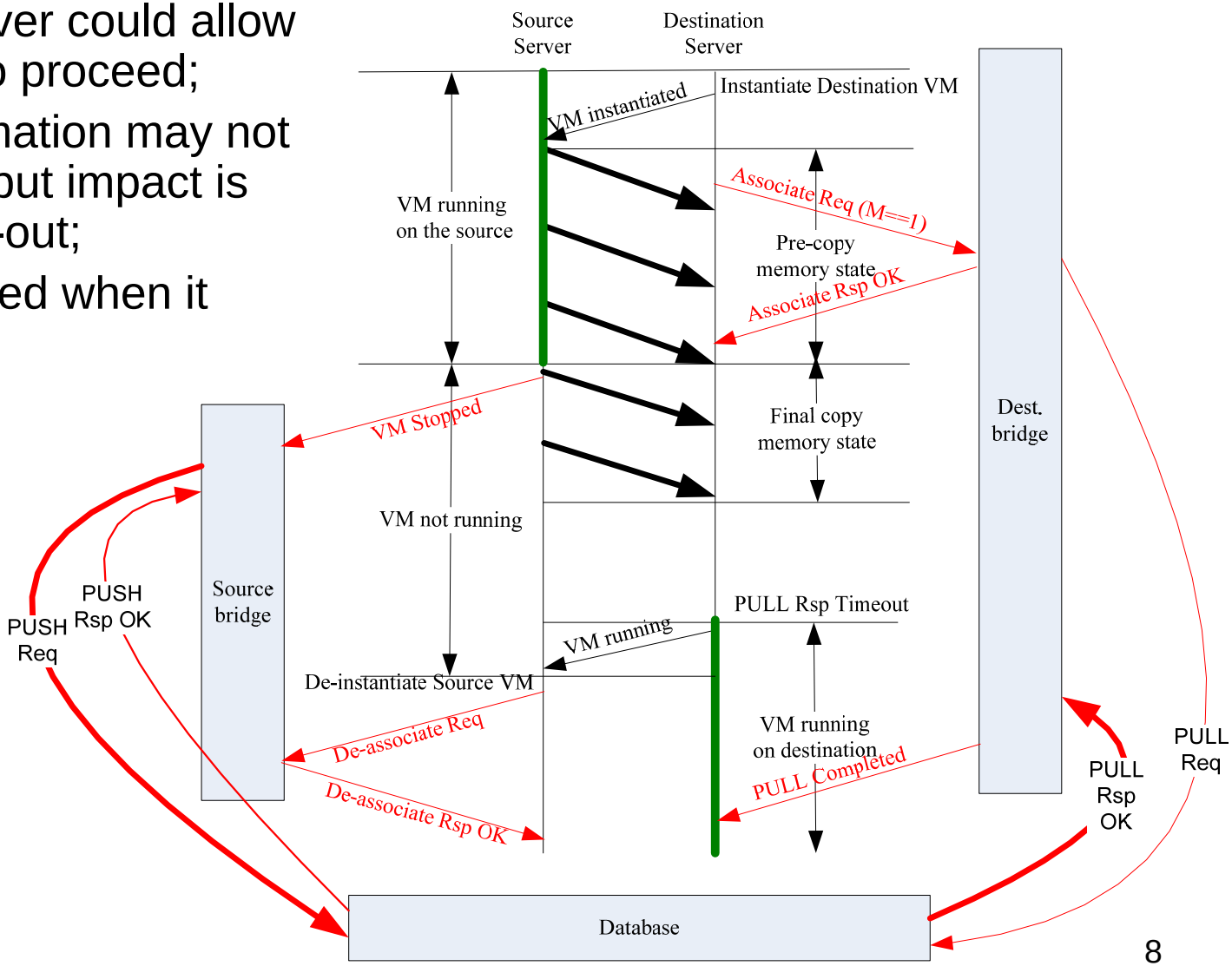
PULL completes *after* memory copy (alt 1)

- Destination server determines how to behave when ‘PULL Completed’ arrives after memory copied;
- For example, it can wait ‘n seconds’ and then back out of VM migration;



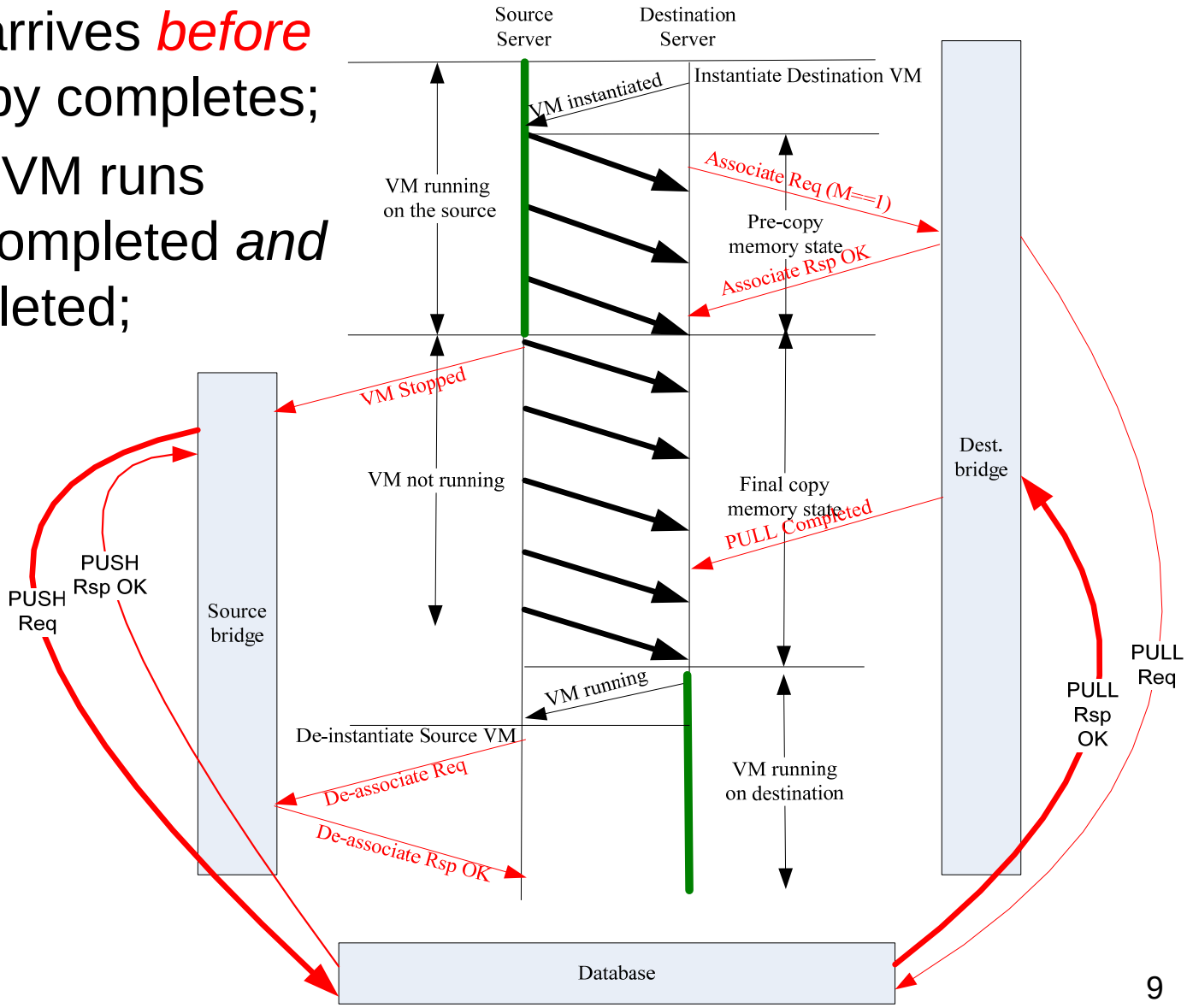
PULL completes *after* memory copy (alt 2)

- …or, after waiting ‘n seconds’ destination server could allow the migration to proceed;
- Dynamic Information may not be transferred but impact is not worth back-out;
- PULL rsp ignored when it arrives;



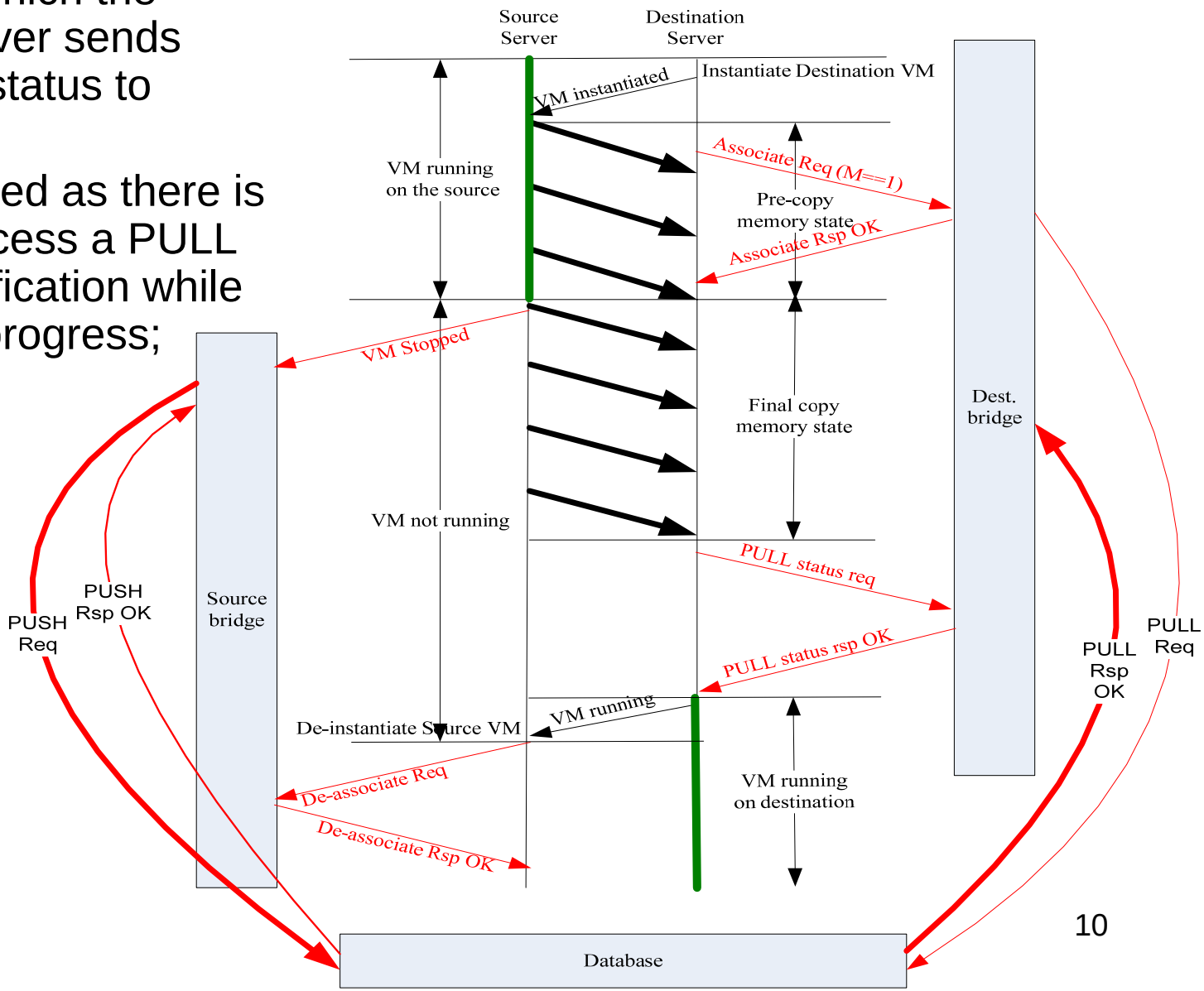
PULL completes *before* memory copy

- PULL Rsp arrives *before* memory copy completes;
- Destination VM runs after copy completed *and* PULL completed;



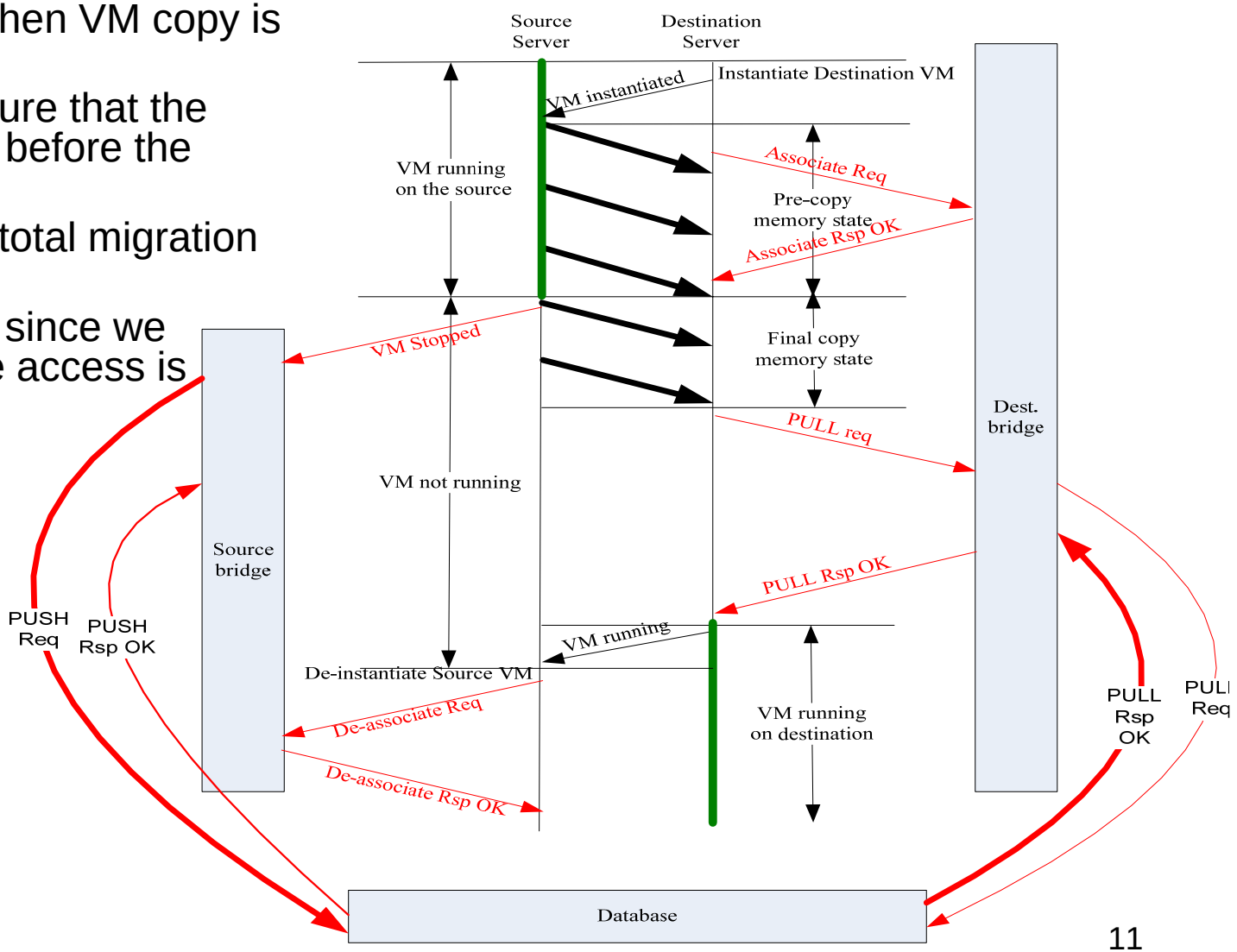
PULL completes *before* memory copy (alt)

- Alternative in which the destination server sends request PULL status to Bridge;
- Logic is simplified as there is no need to process a PULL completed notification while VM copy is in progress;



Mike Krause suggestion to serialize

- PUSH triggered when Source VM stops;
- PULL triggered when VM copy is complete???
- Purpose is to ensure that the PUSH completes before the PULL???
- Results in longer total migration time;
- Benefit isn't clear since we assume database access is asynchronous;



Conclusions

- Communicate M-bit on Associate indicating whether the VM associated with the VSI is migrating from another location (M = 1);
  - A piece of information requiring *no change* to VDP state machine;
- One-way notification (new TLV) sent from source server to source Bridge indicating that the source VM (1) is migrating and (2) has stopped;
- Either
  - One-way notification (same new TLV) sent from destination Bridge to source server indicating that the PULL has completed (or failed); OR
  - Request from destination server to destination Bridge asking whether the PULL has succeeded, failed, or is still pending, and the corresponding response from Bridge to server;
- Was not able to identify a reason to serialize the VM copy and the PULL as suggested in last meeting;