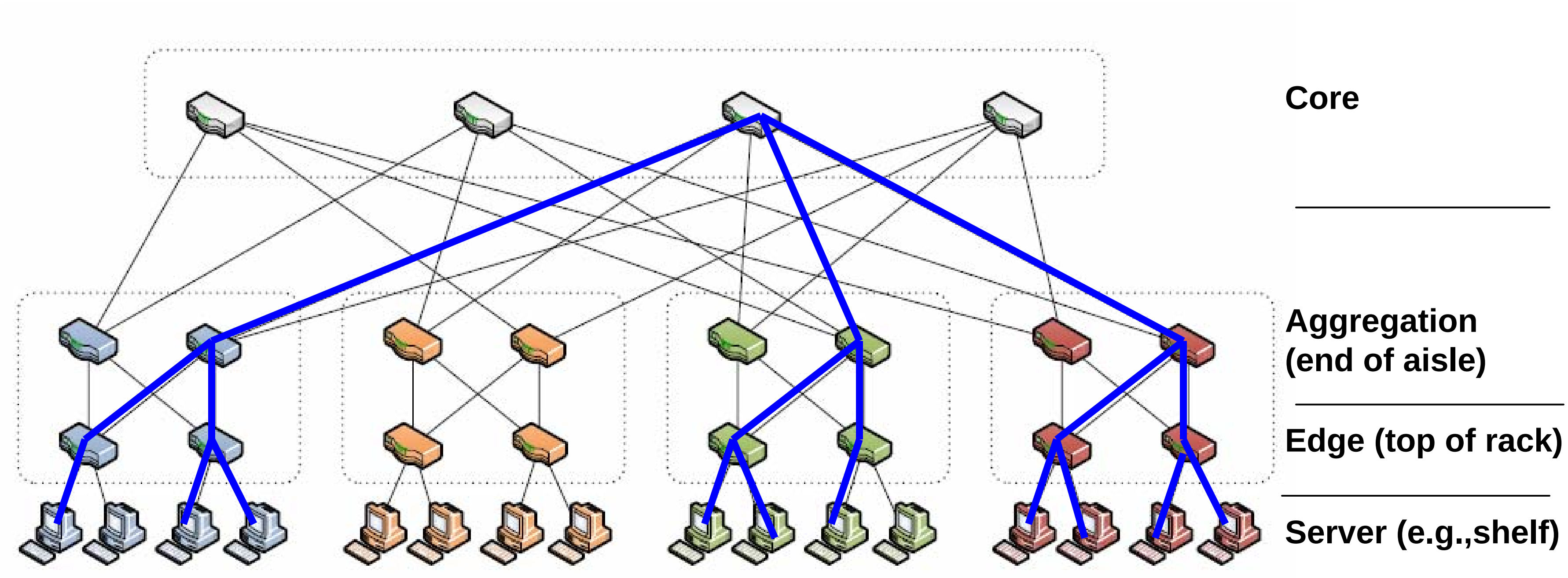
Flood Reduction with Districts

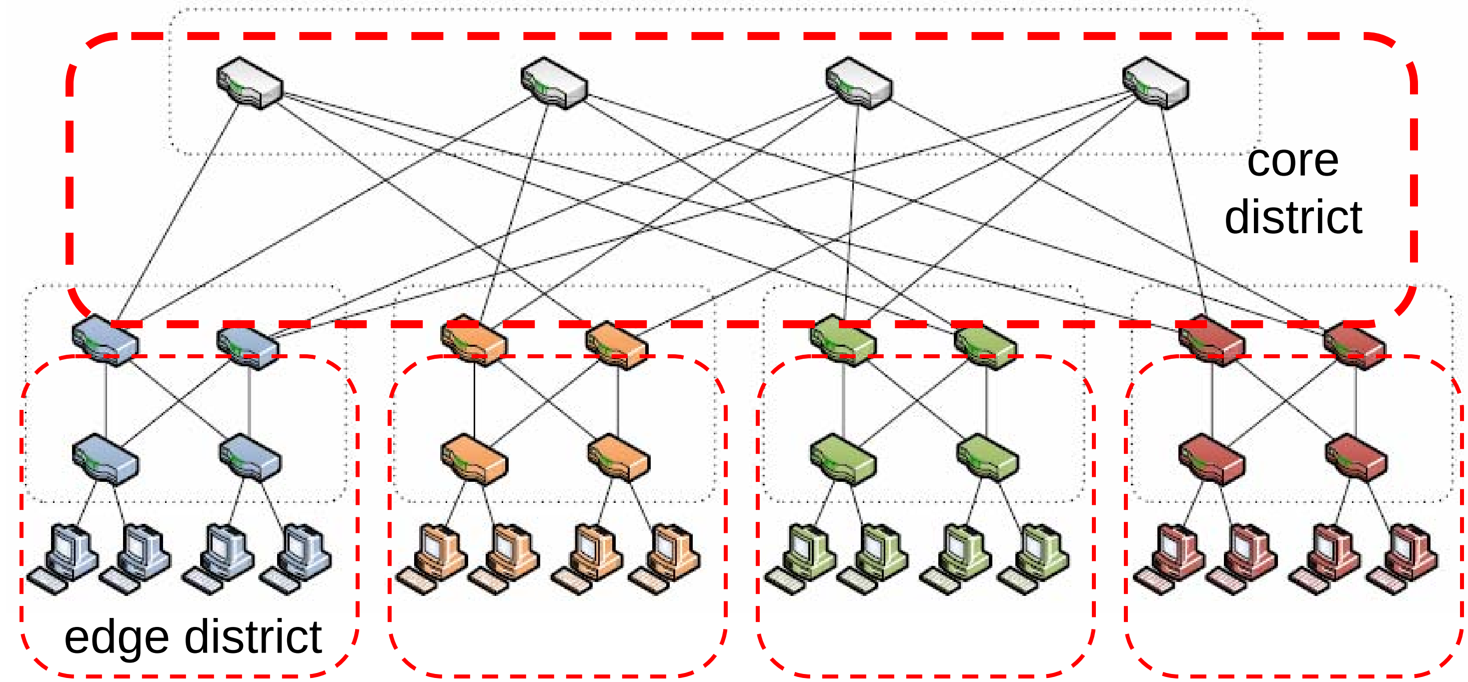
Bob Sultan (bsultan@huawei.com)

Problem



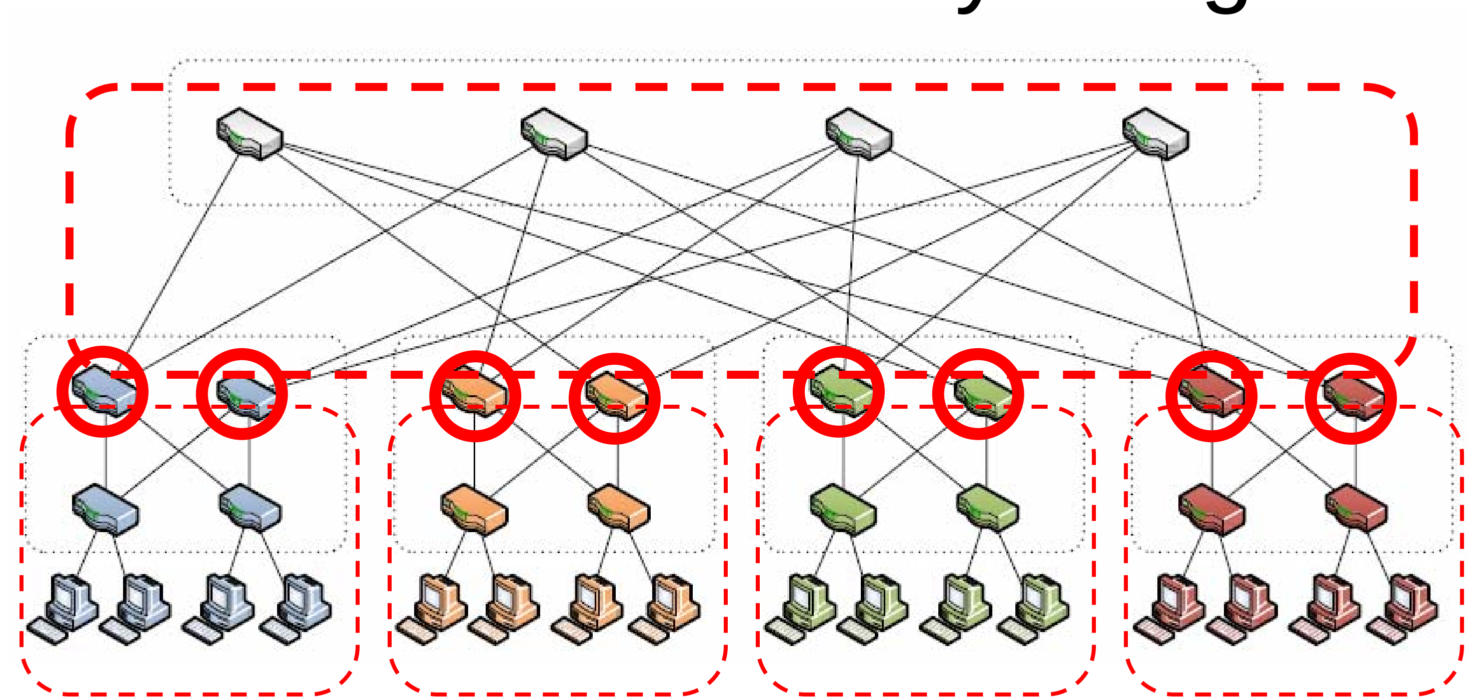
- The number of end-stations associated with the data center can grow large (>100K);
- Excessive flooding may result if Bridge FDB sizes are insufficient to cache large numbers of MAC addresses;
- Each flood has, as its scope, the entire VLAN of which the end-station is a member. Thus, the fraction of link capacity consumed by flooding increases with the number of end-stations;

Districts



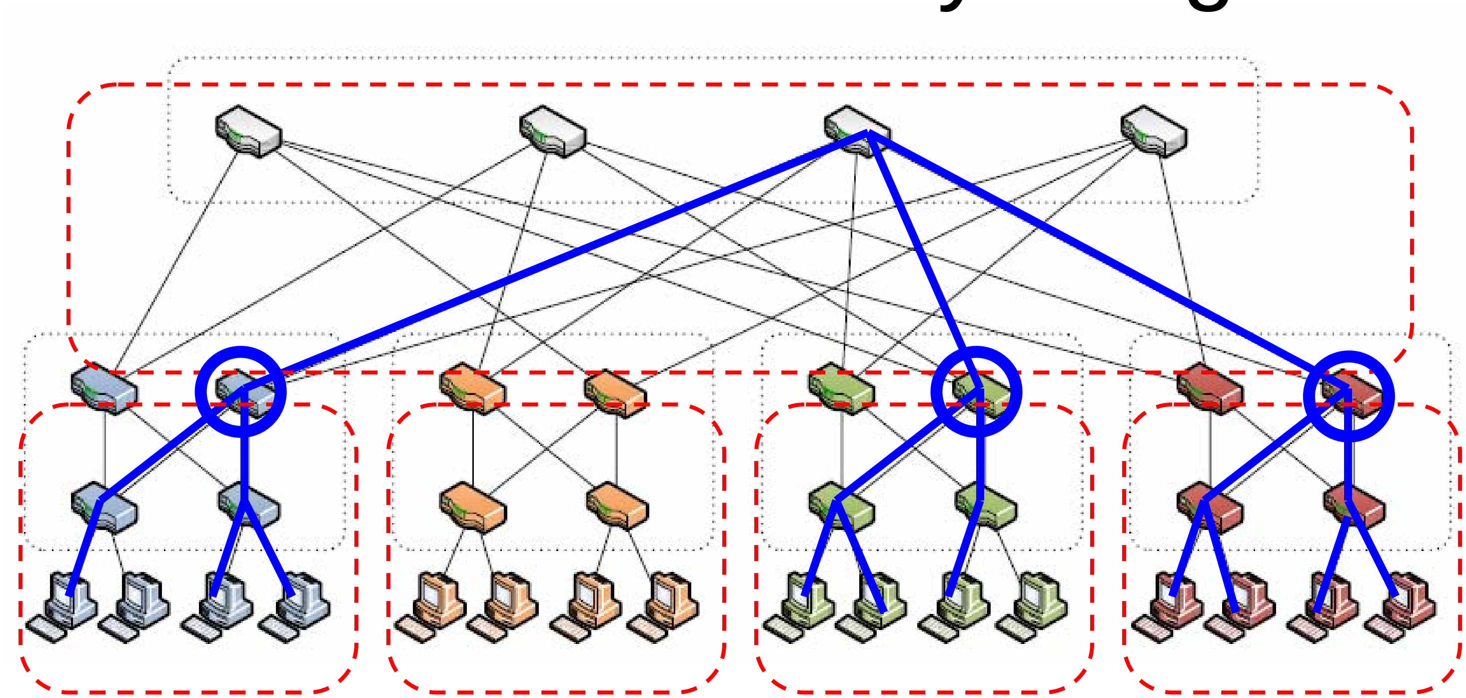
- The bridged network can be partitioned into *districts* (red dashed);
- An *edge district* (light dashed) supports end-station attachment;
- A *core district* (heavy dashed) interconnects edge districts;

District Boundary Bridge



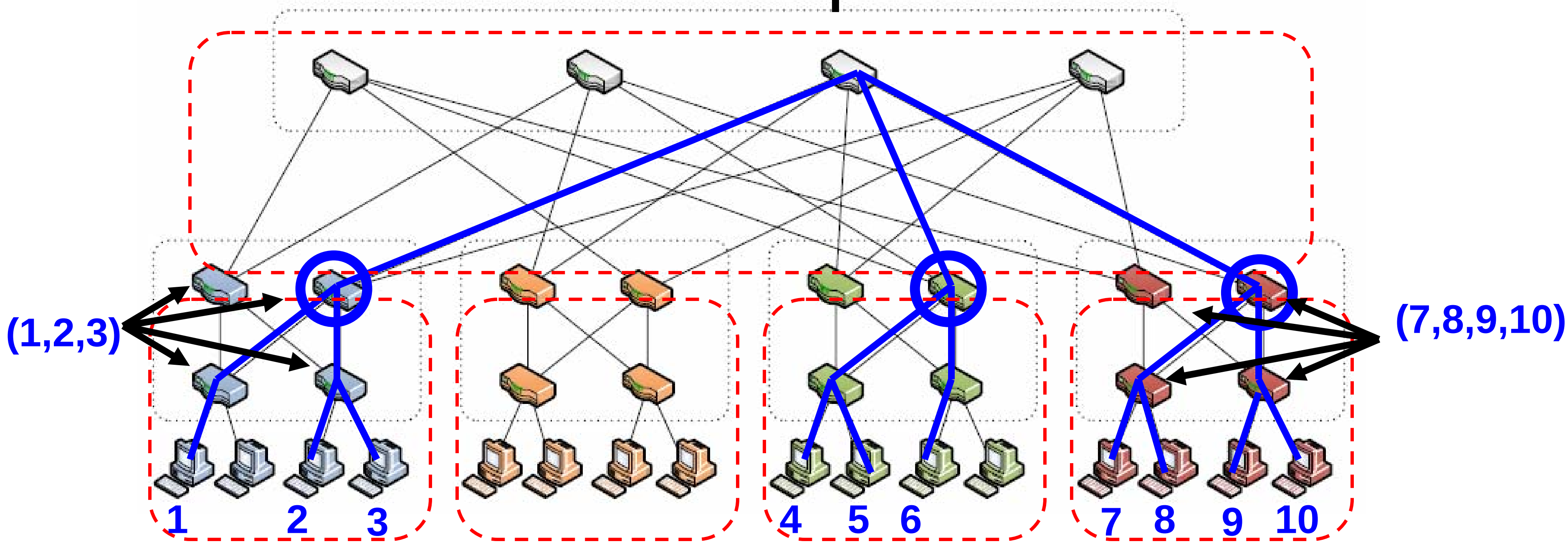
- A bridge participating in multiple districts is a *district boundary bridge* (DBB; red circle);

Active Boundary Bridge



- A DBB lying on the spanning tree associated with a particular VLAN (e.g., blue VLAN) is known as an Active Boundary Bridge (ABB; blue circle) with respect to that VLAN;

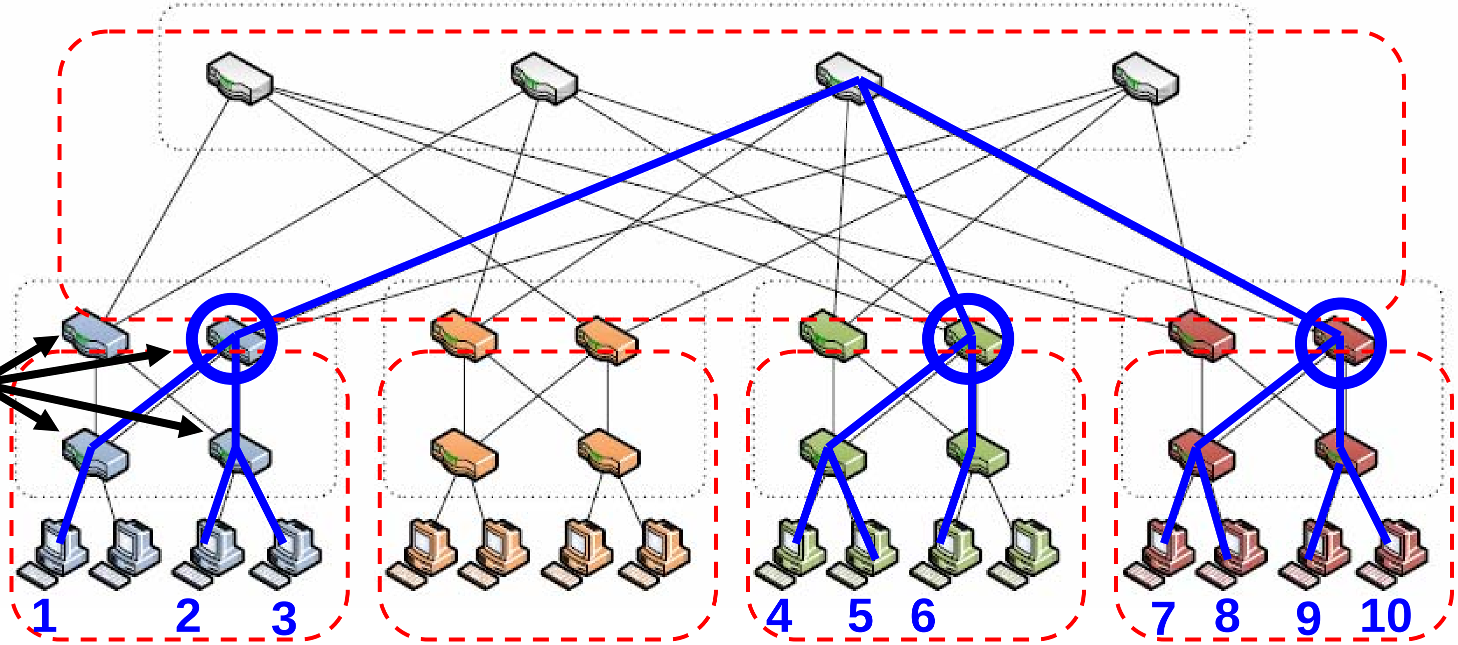
Assumption



- The <MAC, VID> identifying an end-station is known to all bridges in the district containing that end-station;
- This information can be (for example)
  - configured on the bridges of the edge district by the hypervisor;
  - registered by the end-station via MMRP to bridges in the edge district;
  - communicated by the end-station to the edge-bridge via EVB VDP and registered to other bridges in the edge district via MMRP;

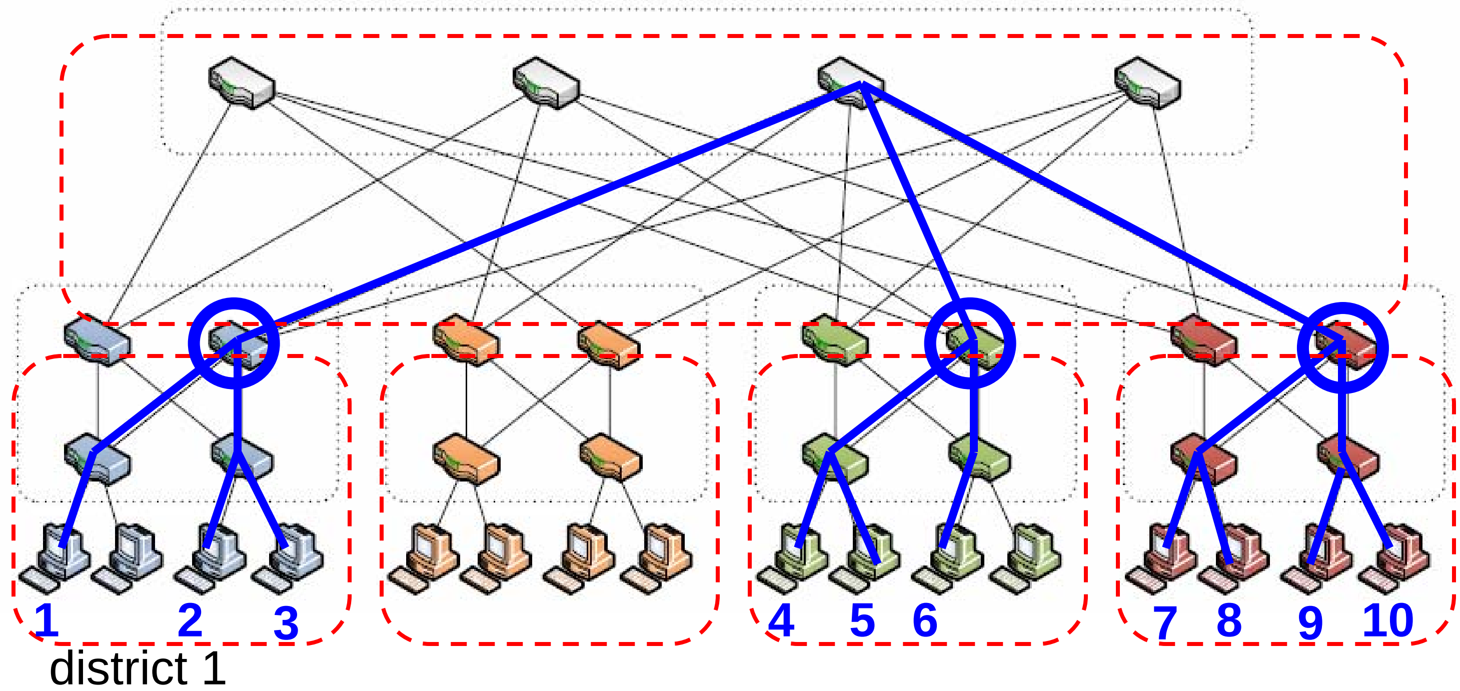
Maintaining end-station identities

DA	VID	Port Map
1	blue	NULL
2	blue	NULL
3	blue	NULL



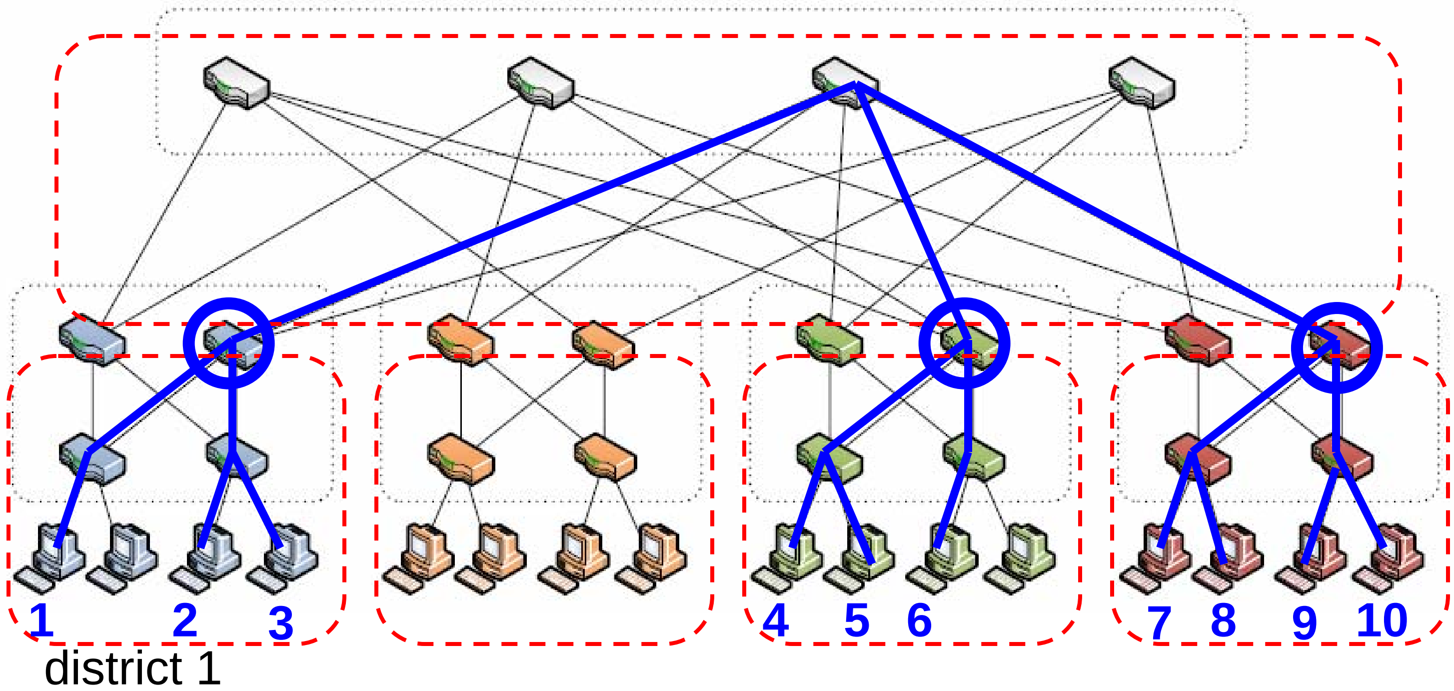
- A bridge can represent each end-station in the edge district using an FDB entry;
- If the FDB entry represents a <MAC, VID> that has not been learned, the outbound port value associated with that FDB entry is NULL;

SA and DA are in same edge district (1)



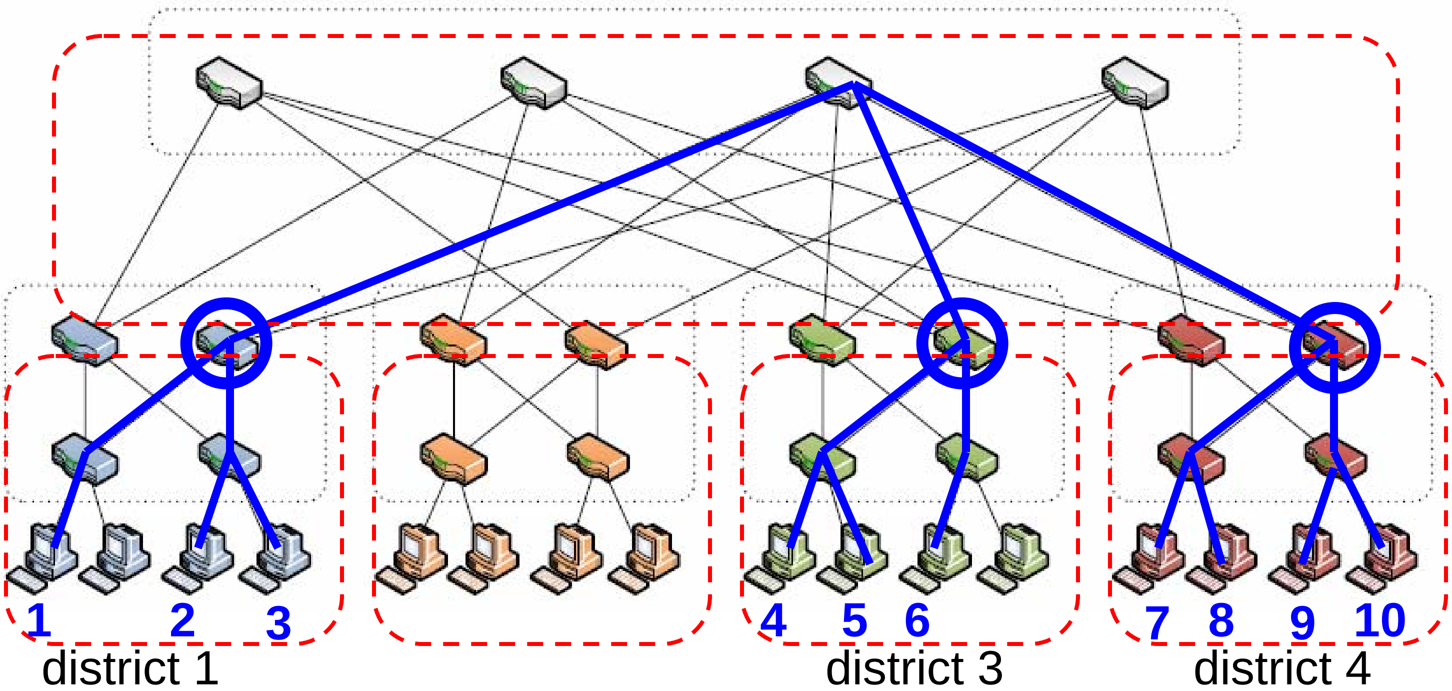
- End-station 1 sends frame to end-station 3;
- Bridge receiving frame in district 1 determines that DA *is* in district 1 by observing that an FDB entry exists;
- Bridge floods the frame if outbound port is NULL;
  - – or forwards the frame on specified outbound port;

SA and DA are in same edge district (2)



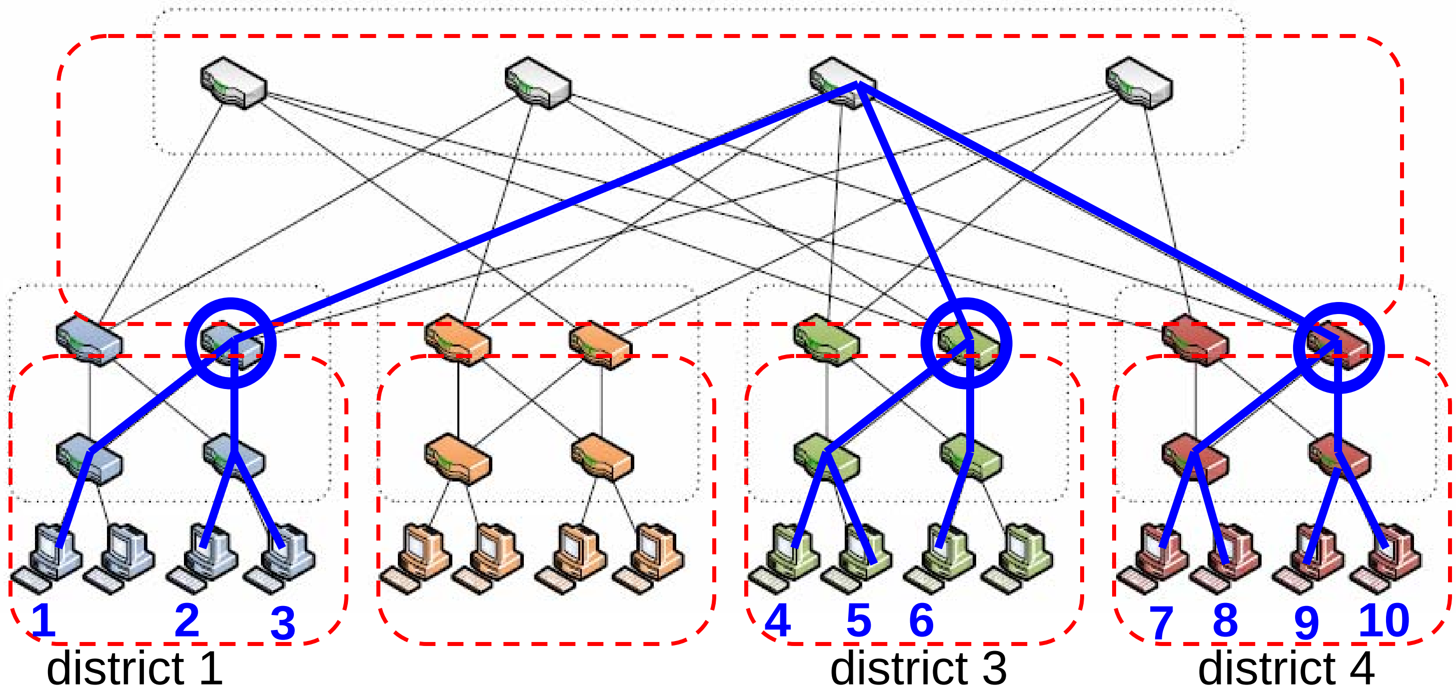
- Bridge learns the SA (overwriting the NULL value if the SA was not previously learned);
- District 1 ABB recognizes that DA is in district 1 and does not allow the frame to flood into the core district;

SA and DA in different edge-domains (1)



- End-station 1 sends frame to end-station 9;
- Bridges receiving the frame in district 1 forward frame towards ABB;
- The frame will be forwarded as usual across the core district (flooding and/or learning as appropriate);

SA and DA in different edge-domains (2)



- The ABB of district 3 recognizes that the DA is *not* in the district and drops the frame;
- The ABB of district 4 recognizes that the DA *is* in the district (no learning wrt district 4);
  - the frame is flooded or forwarded depending on whether or not the outbound port associated with the DA is NULL;
  - the SA is not learned (no FDB entry corresponding to the SA)

Rules

- If no FDB entry then forward towards ABB else
  - If outbound port is null then flood else
    - Forward on outbound port;
- ABB propagates frames outside the edge district iff the destination end-station is not in the edge district;
- Learning occurs in an edge district only if the source end-station is in the edge district;
- Learning/flooding takes place as usual in the core district;
- The ABB is assigned a well-known MAC address and floods a frame having that MAC address as SA into the edge district in order to allow bridges in the edge district to maintain an FDB entry for the ABB.

Key Points

- Maximum number of FDB entries for bridges in an edge district is capped at the number of end-stations in the edge district;
- Flooding of a frame is limited to the single destination edge district and the core district;
- The method assumes that each bridge in an edge district be aware of the <MAC, VID> associated with each end-station in the edge district; this requirement can be met by several different methods;