The use of Ethernet in the DataFlow of the ATLAS Trigger & DAQ

Stefan Stancu	CERN, UPB Bucureşti
Bob Dobinson	CERN
Matei Ciobotaru	CERN, UPB Bucureşti
Krzysztof Korcyl	INP Cracow
Emil Knezo	CERN, SAI Eindhoven

Acknowledgments: ATLAS Trigger/DAQ Collaboration
Razvan Beuran – CERN, UPB Bucureşti
Mihail Ivanovici – CERN, UPB Bucureşti

March 21, 2003

Contents

The purpose of this talk is to describe the Ethernet features required for a proper operation of the ATLAS DataCollection network (described in the the ATLAS Baseline Architecture [2]). It will cover the following :

Contents

The purpose of this talk is to describe the Ethernet features required for a proper operation of the ATLAS DataCollection network (described in the the ATLAS Baseline Architecture [2]). It will cover the following :

Contents

The purpose of this talk is to describe the Ethernet features required for a proper operation of the ATLAS DataCollection network (described in the the ATLAS Baseline Architecture [2]). It will cover the following :

Contents

The purpose of this talk is to describe the Ethernet features required for a proper operation of the ATLAS DataCollection network (described in the the ATLAS Baseline Architecture [2]). It will cover the following :

Contents

The purpose of this talk is to describe the Ethernet features required for a proper operation of the ATLAS DataCollection network (described in the the ATLAS Baseline Architecture [2]). It will cover the following :

§1 Atlas Baseline Architecture & the Message Flow

§2 IEEE 802.3 Ethernet evolution

§3 Network testing equipment

§4 Ethernet features with great impact on the ATLAS Baseline Architecture's performance:

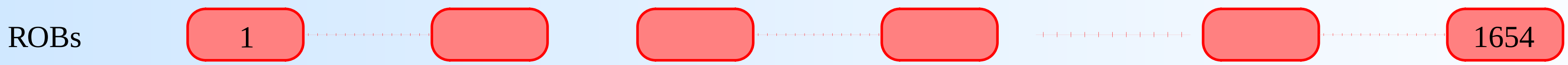
- Packet loss and latency in switches
- Virtual LANs (VLANs)
- Flow Control

Contents

The purpose of this talk is to describe the Ethernet features required for a proper operation of the ATLAS DataCollection network (described in the the ATLAS Baseline Architecture [2]). It will cover the following :

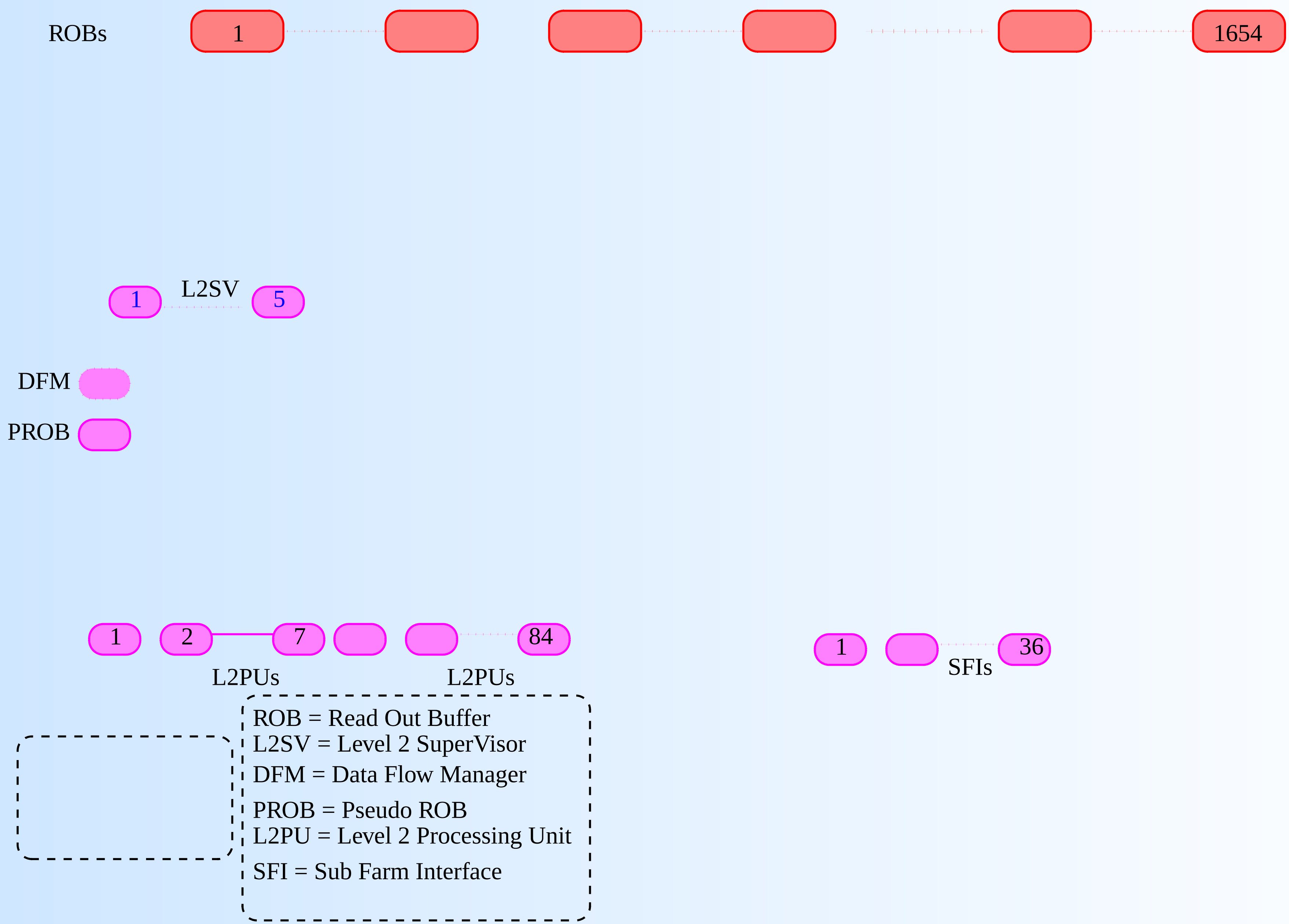
Atlas Baseline Architecture

Atlas Baseline Architecture

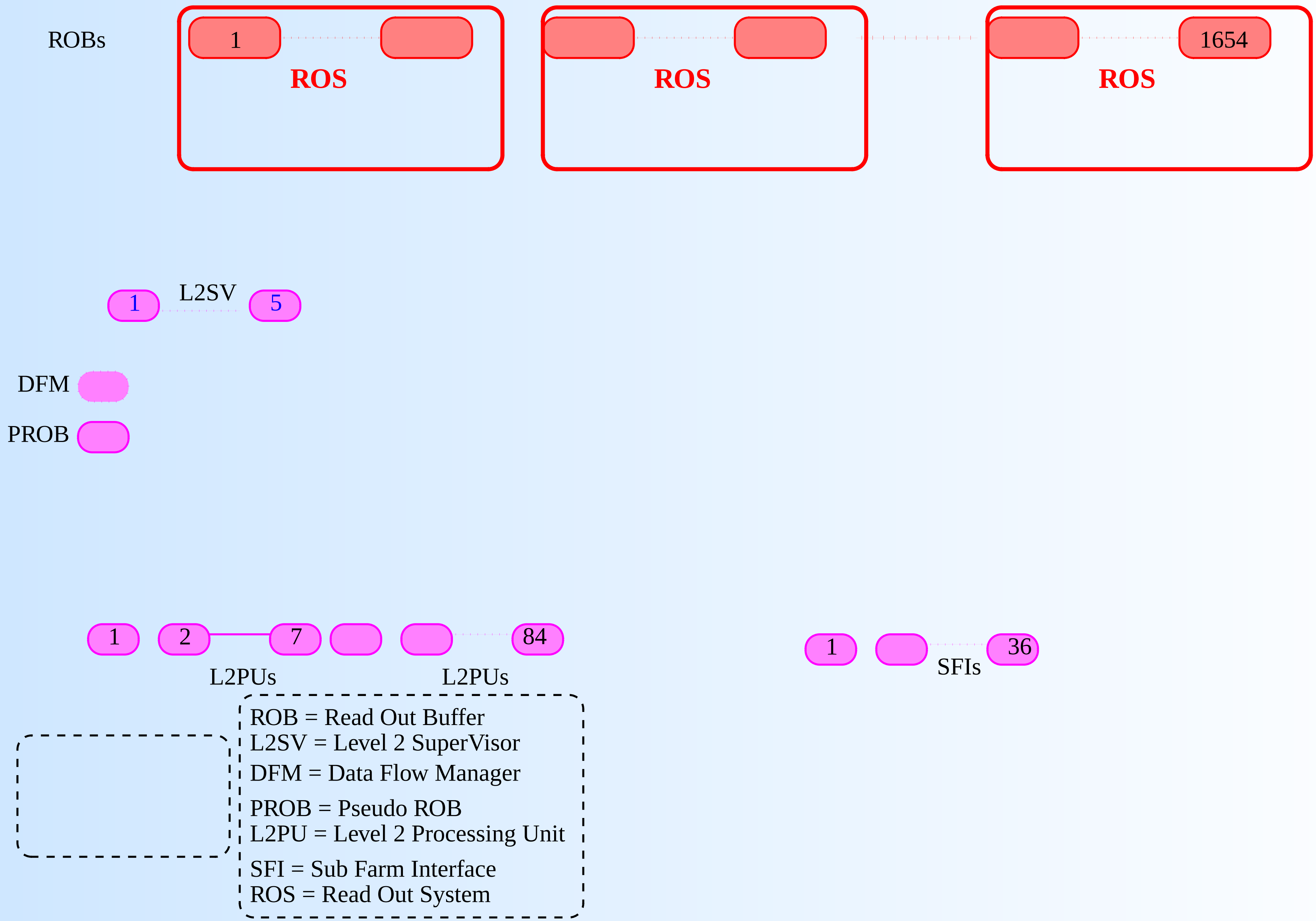


ROB = Read Out Buffer

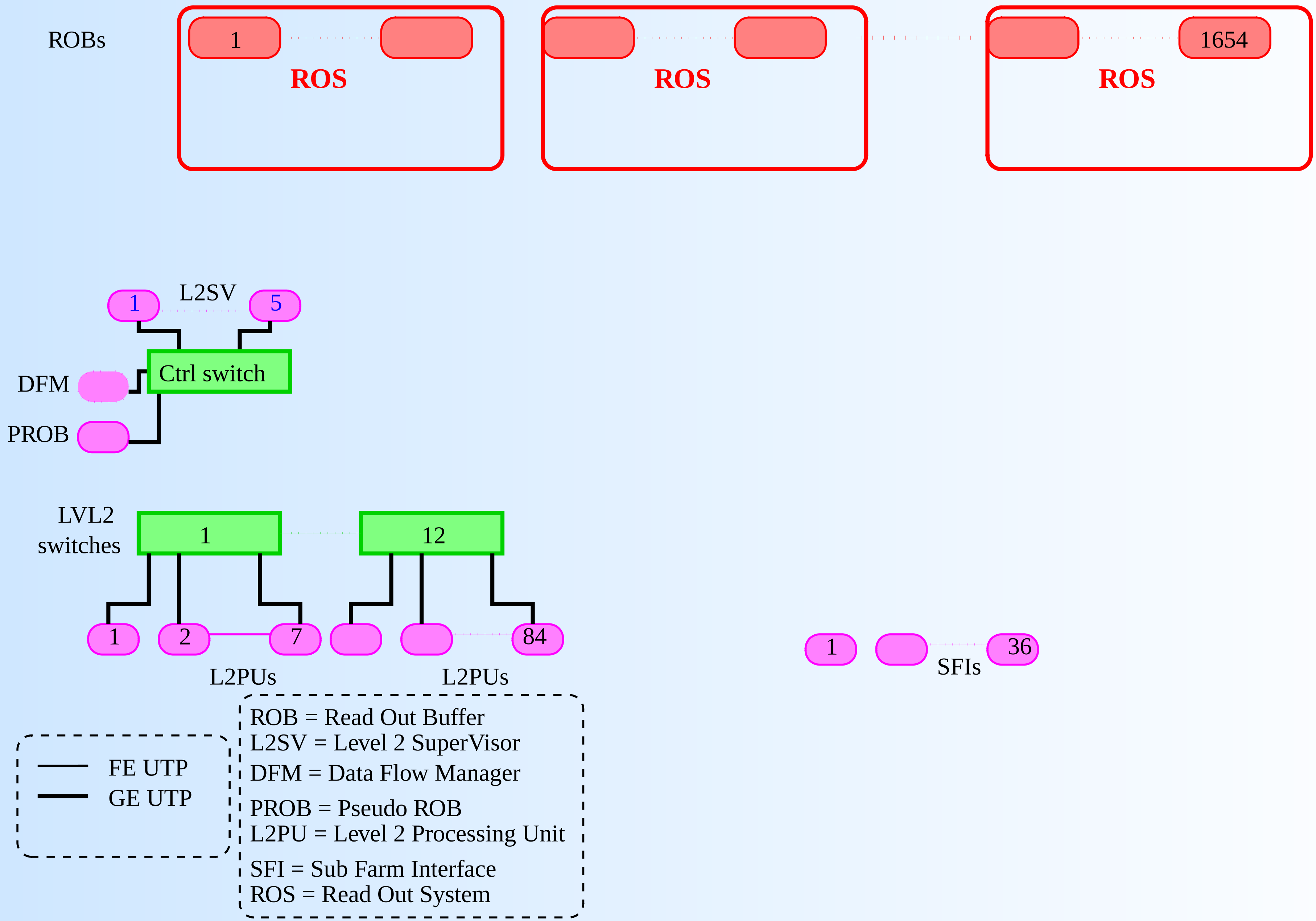
Atlas Baseline Architecture



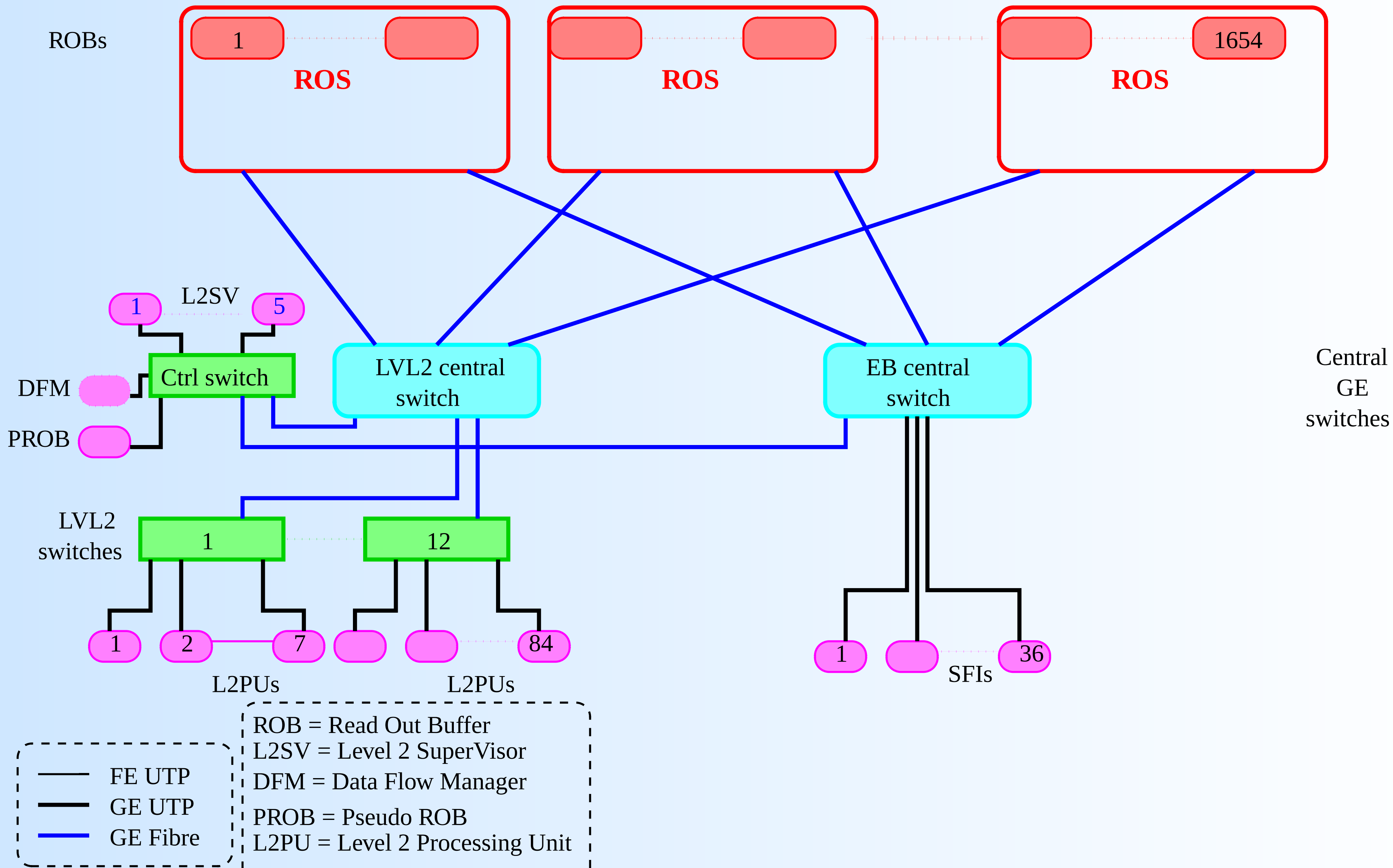
Atlas Baseline Architecture



Atlas Baseline Architecture

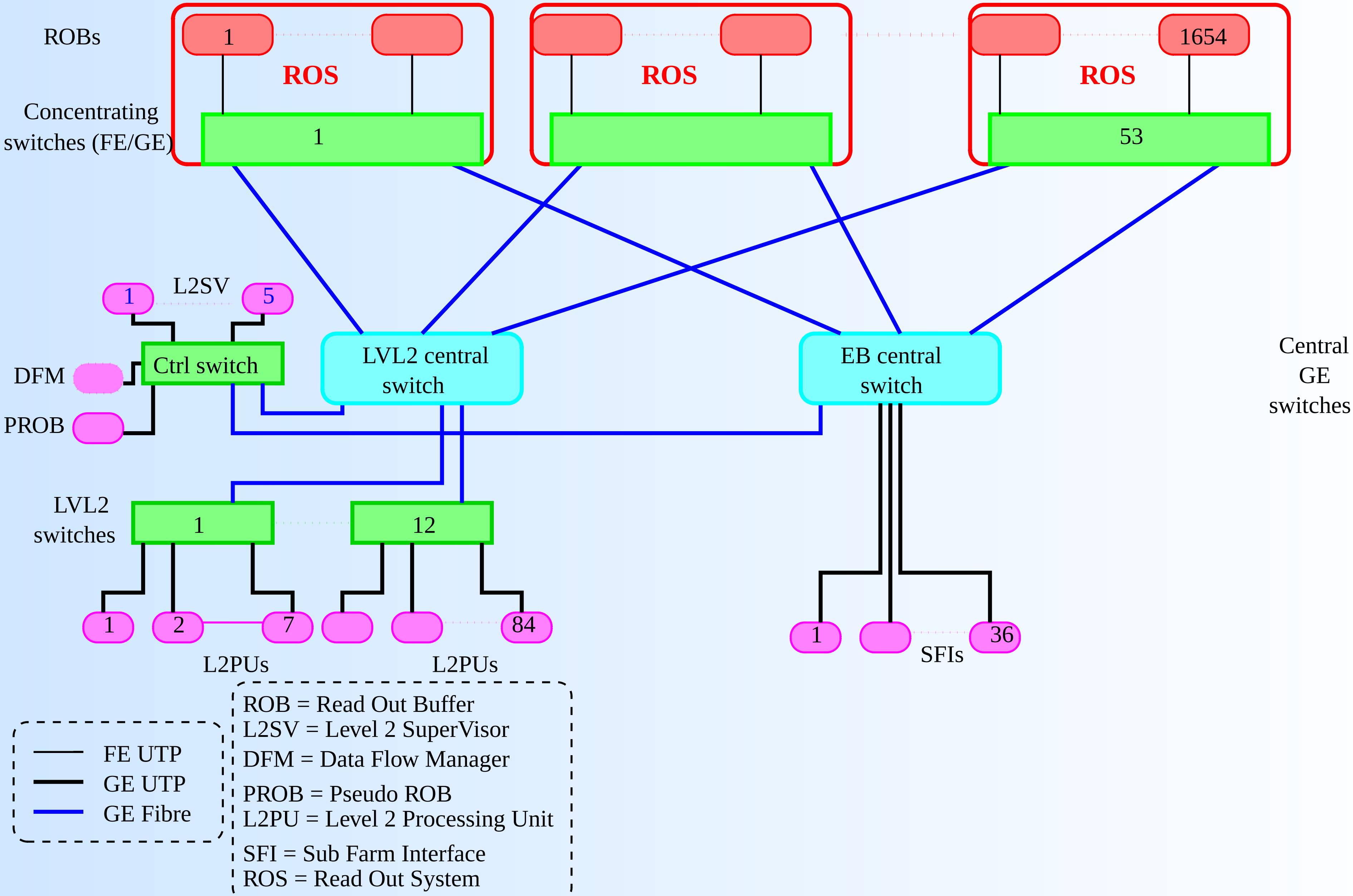


Atlas Baseline Architecture

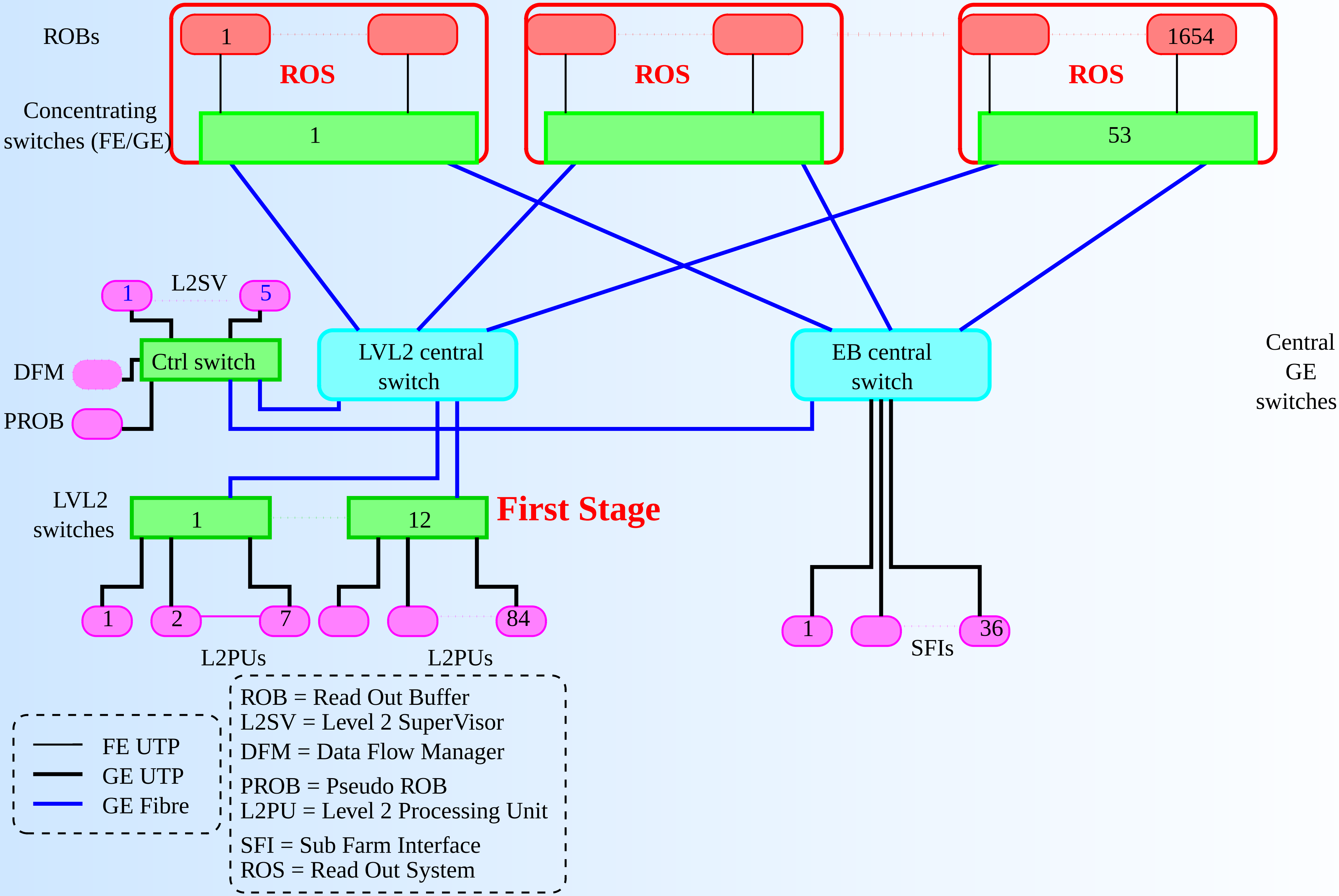


ROB = Read Out Buffer
L2SV = Level 2 SuperVisor
DFM = Data Flow Manager
PROB = Pseudo ROB
L2PU = Level 2 Processing Unit
SFI = Sub Farm Interface
ROS = Read Out System

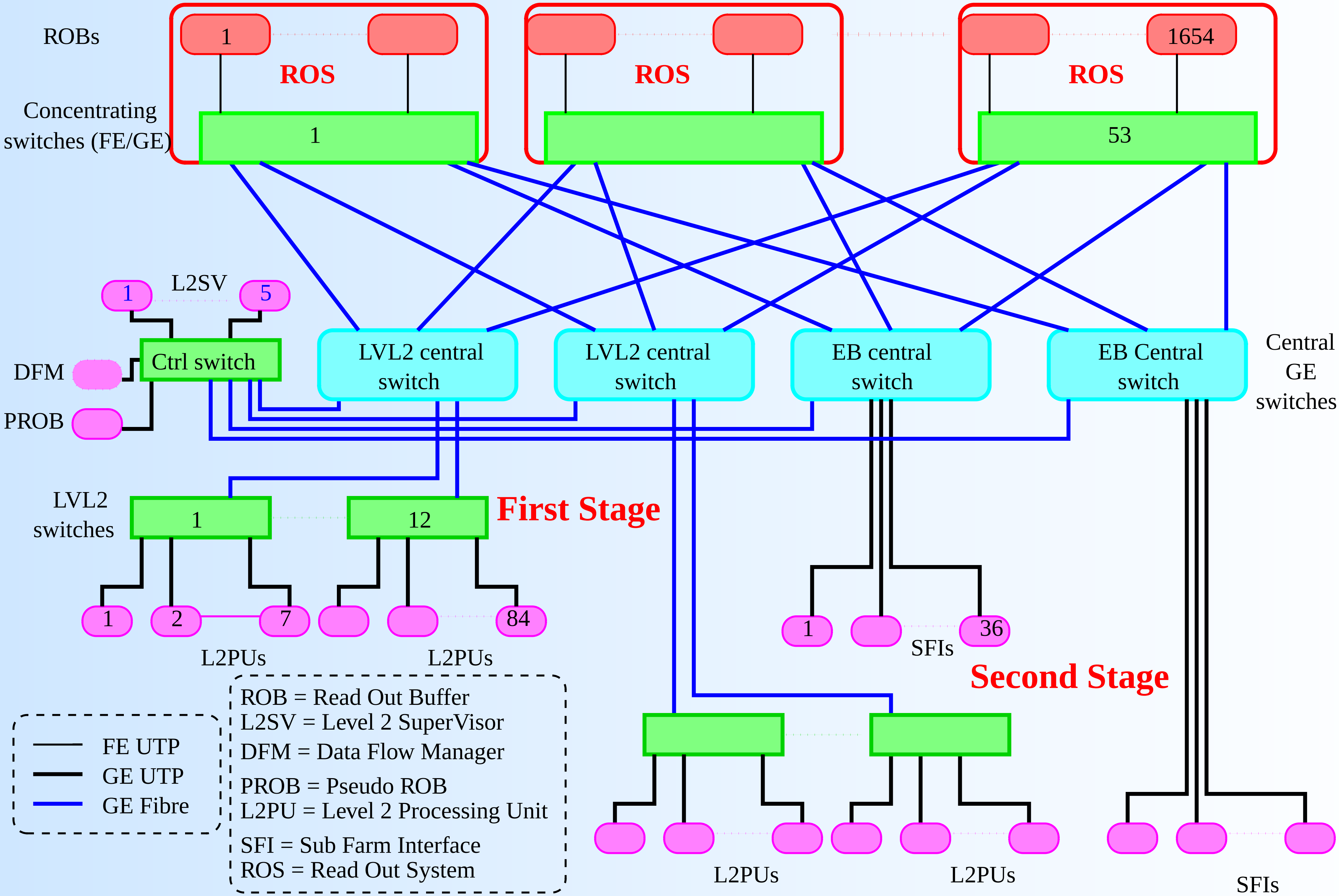
Atlas Baseline Architecture



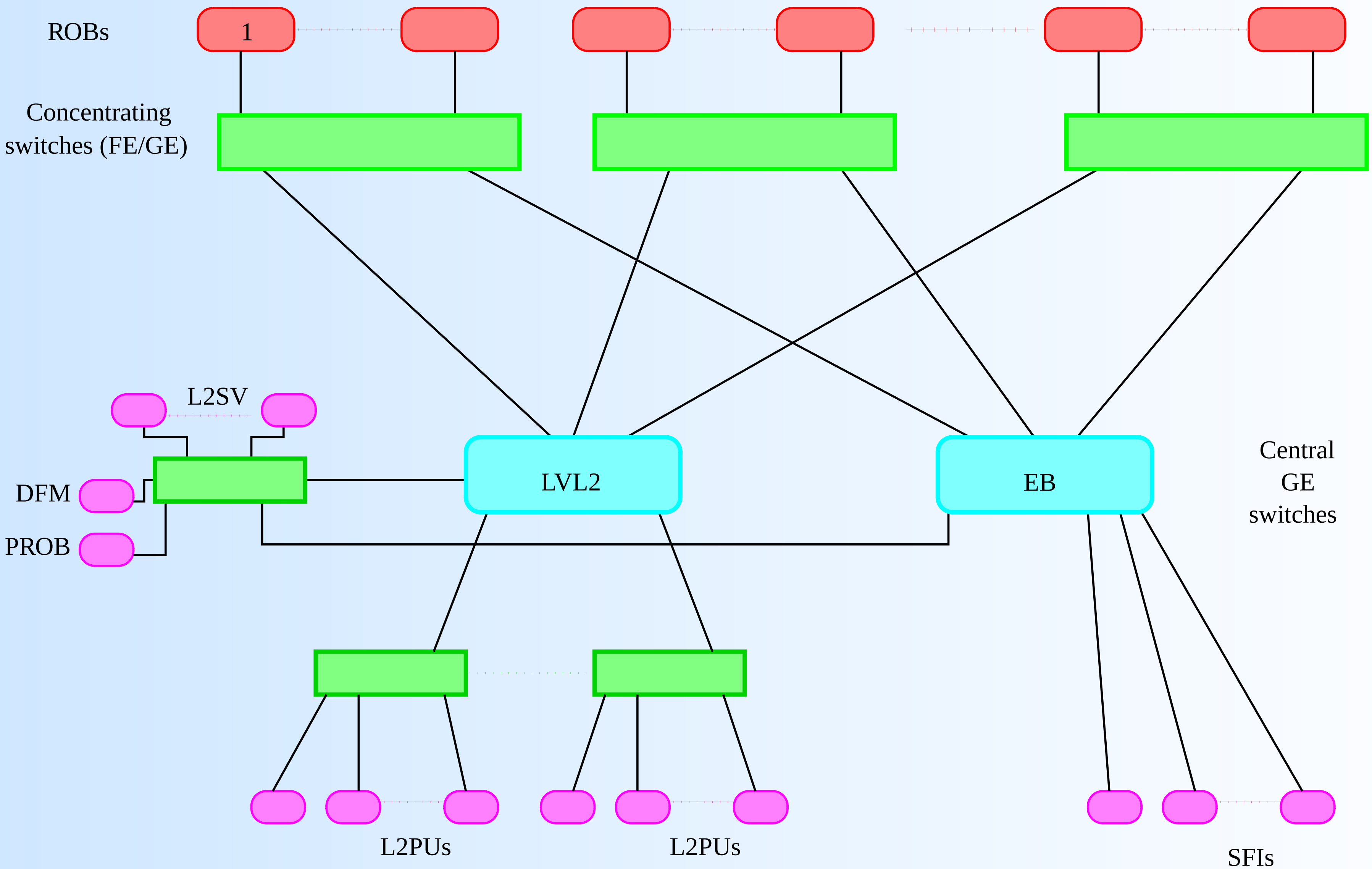
Atlas Baseline Architecture



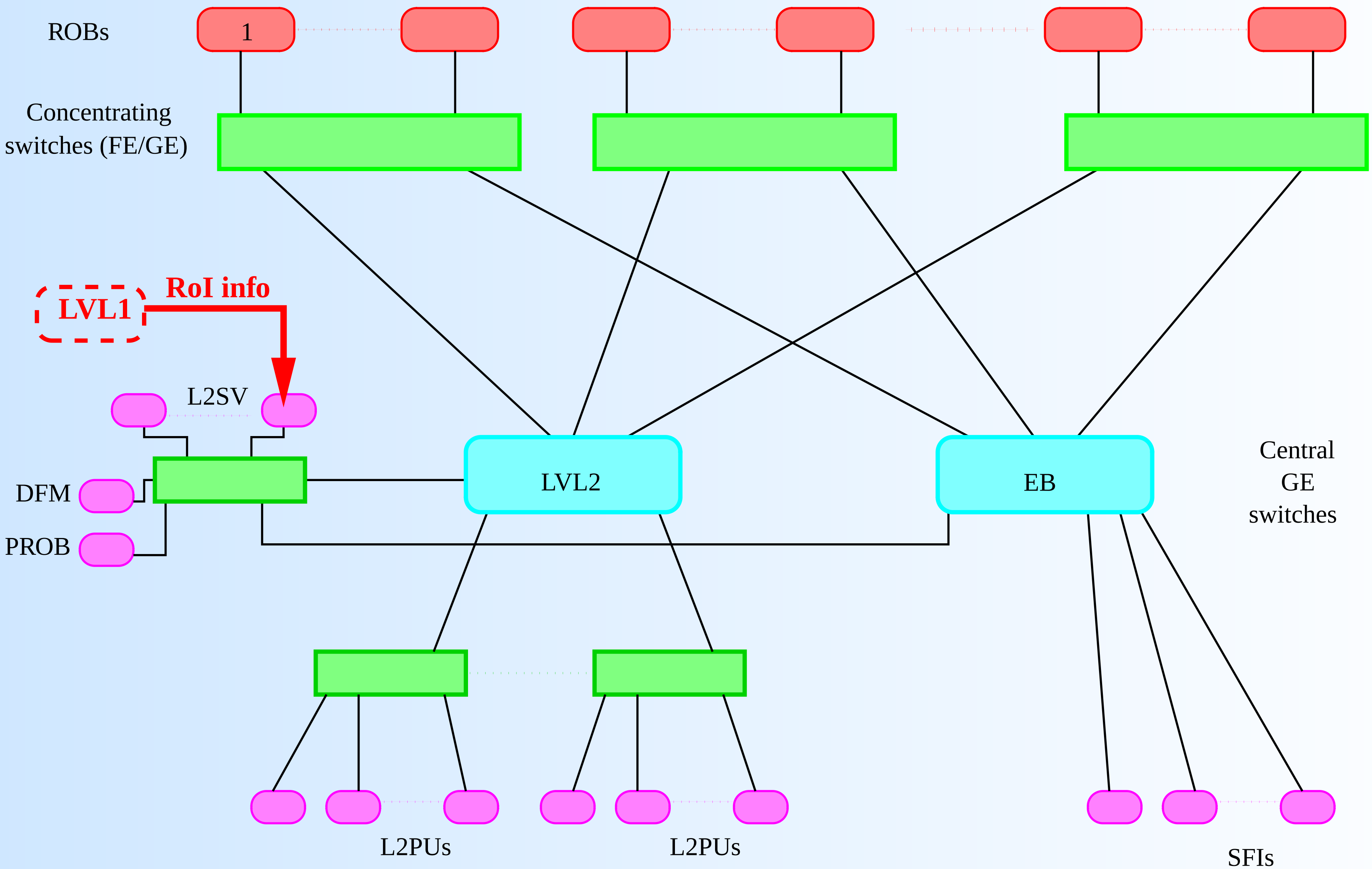
Atlas Baseline Architecture



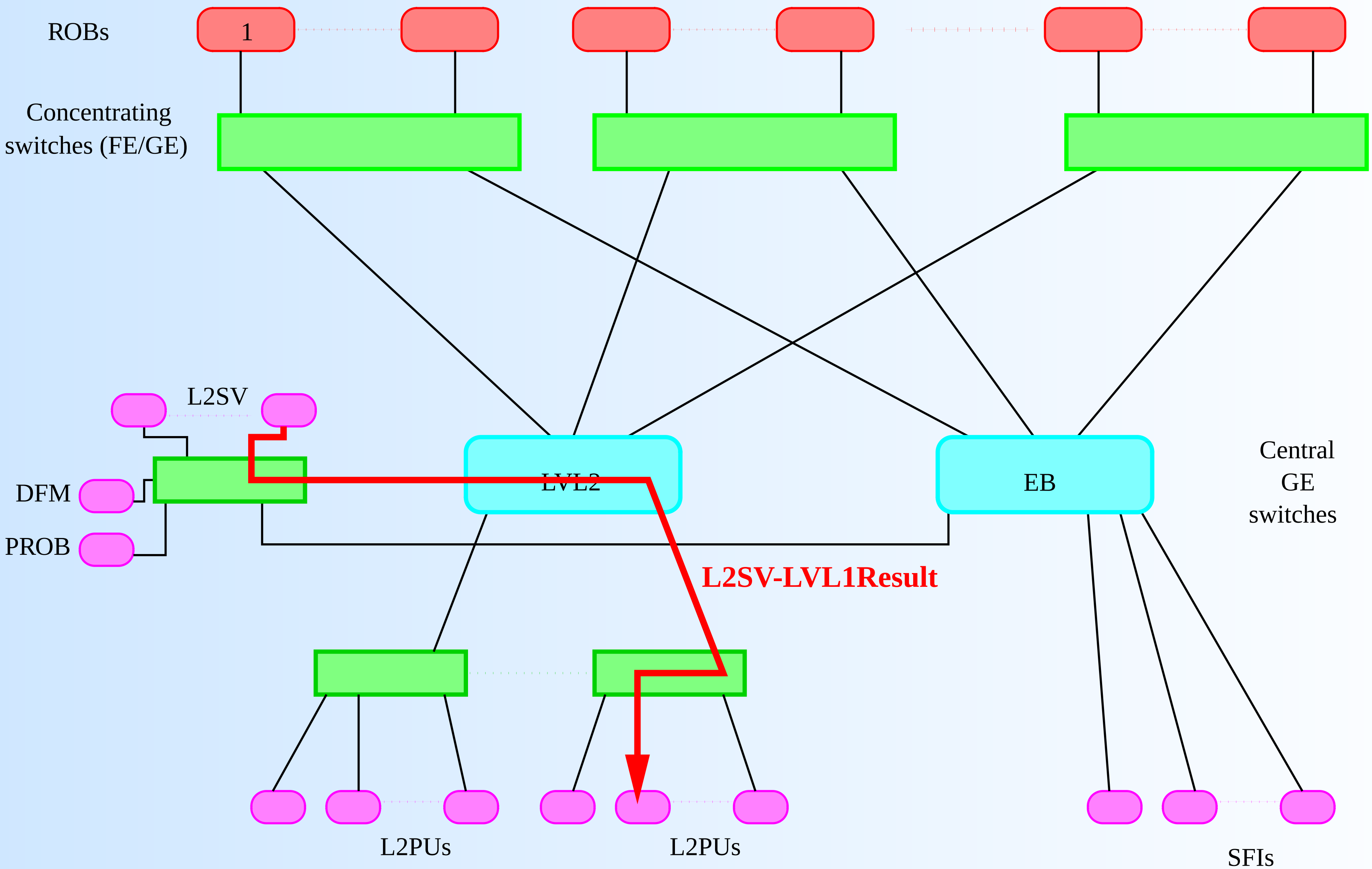
Message Flow [3]



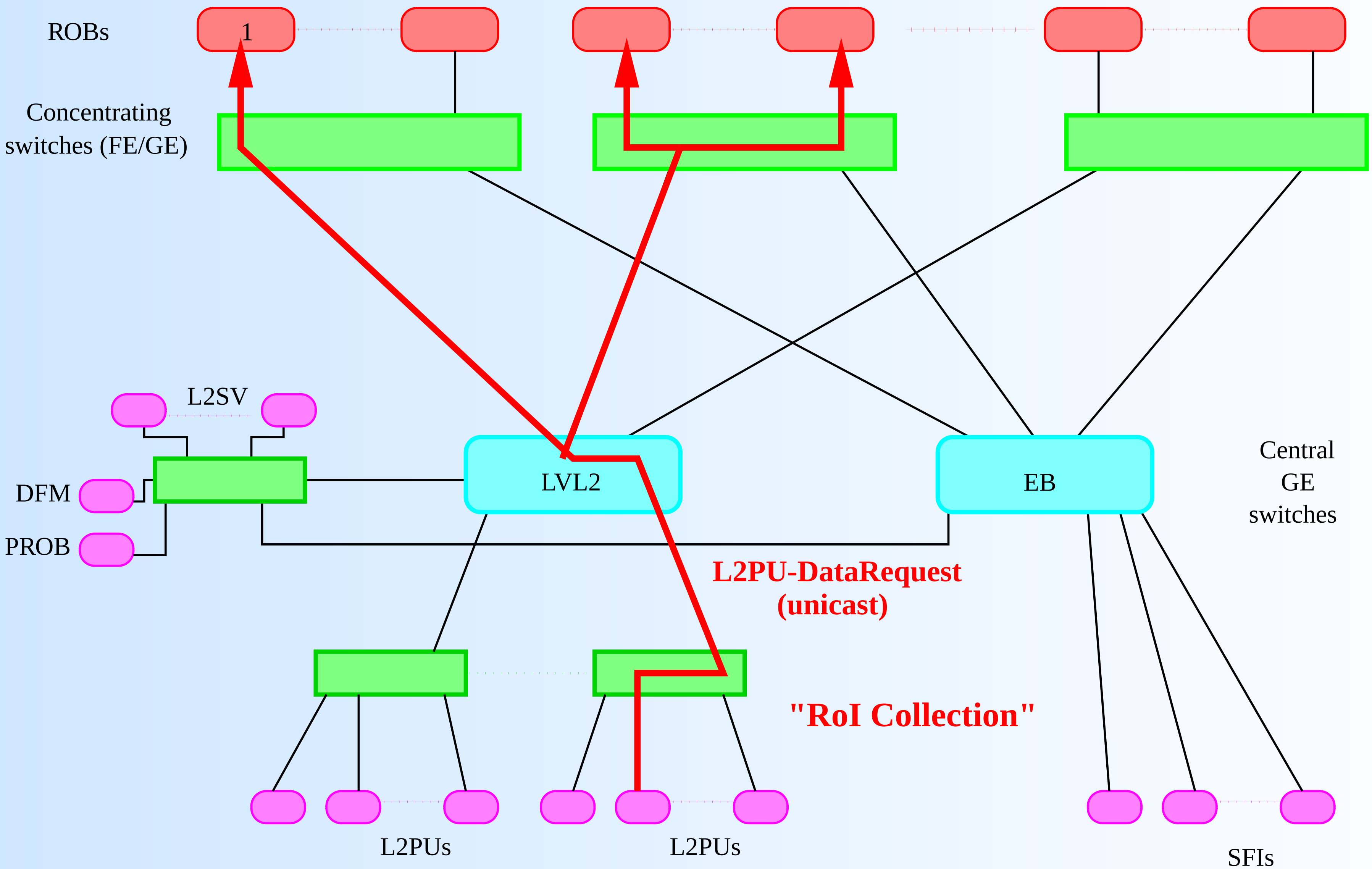
Message Flow [3]



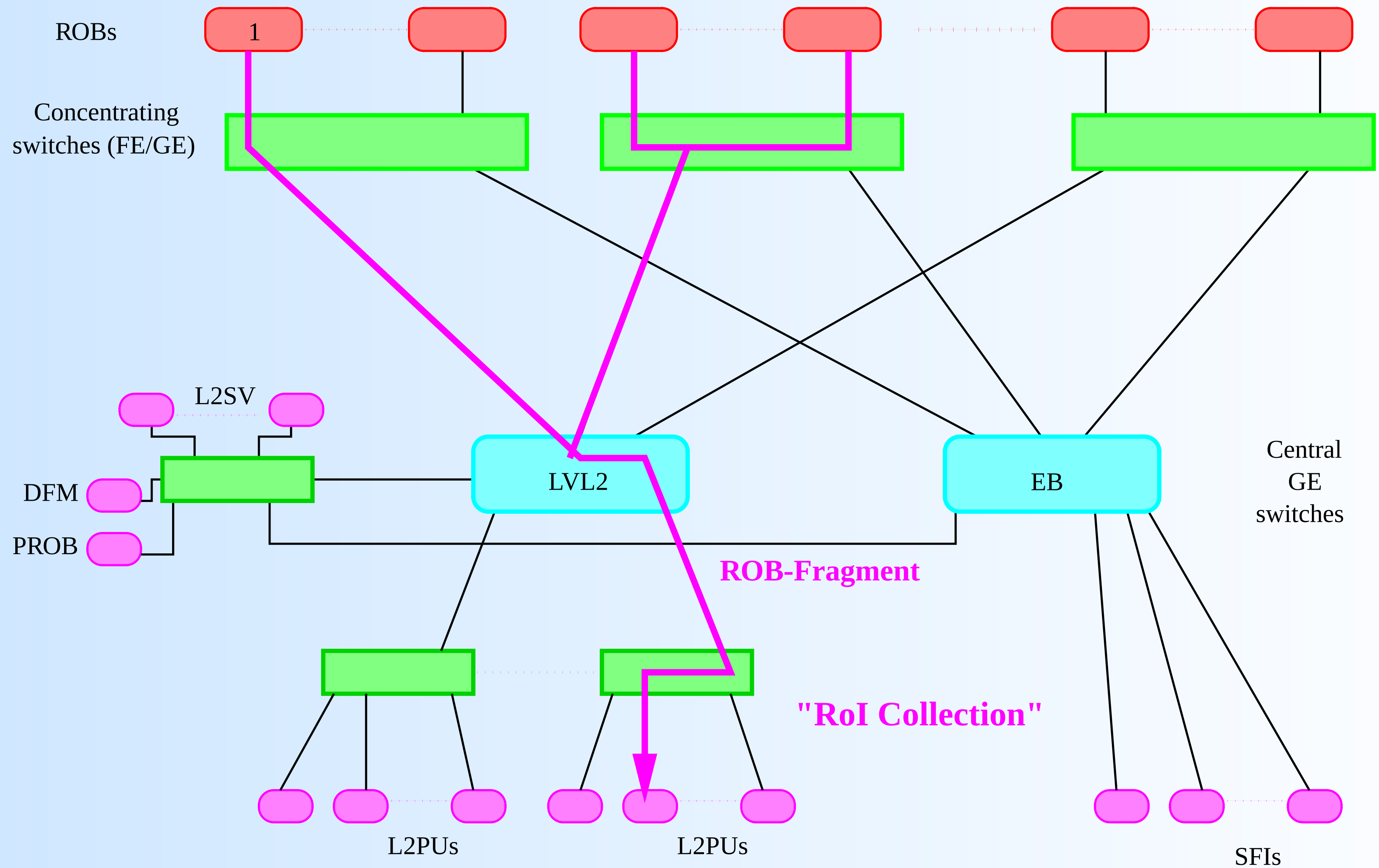
Message Flow [3]



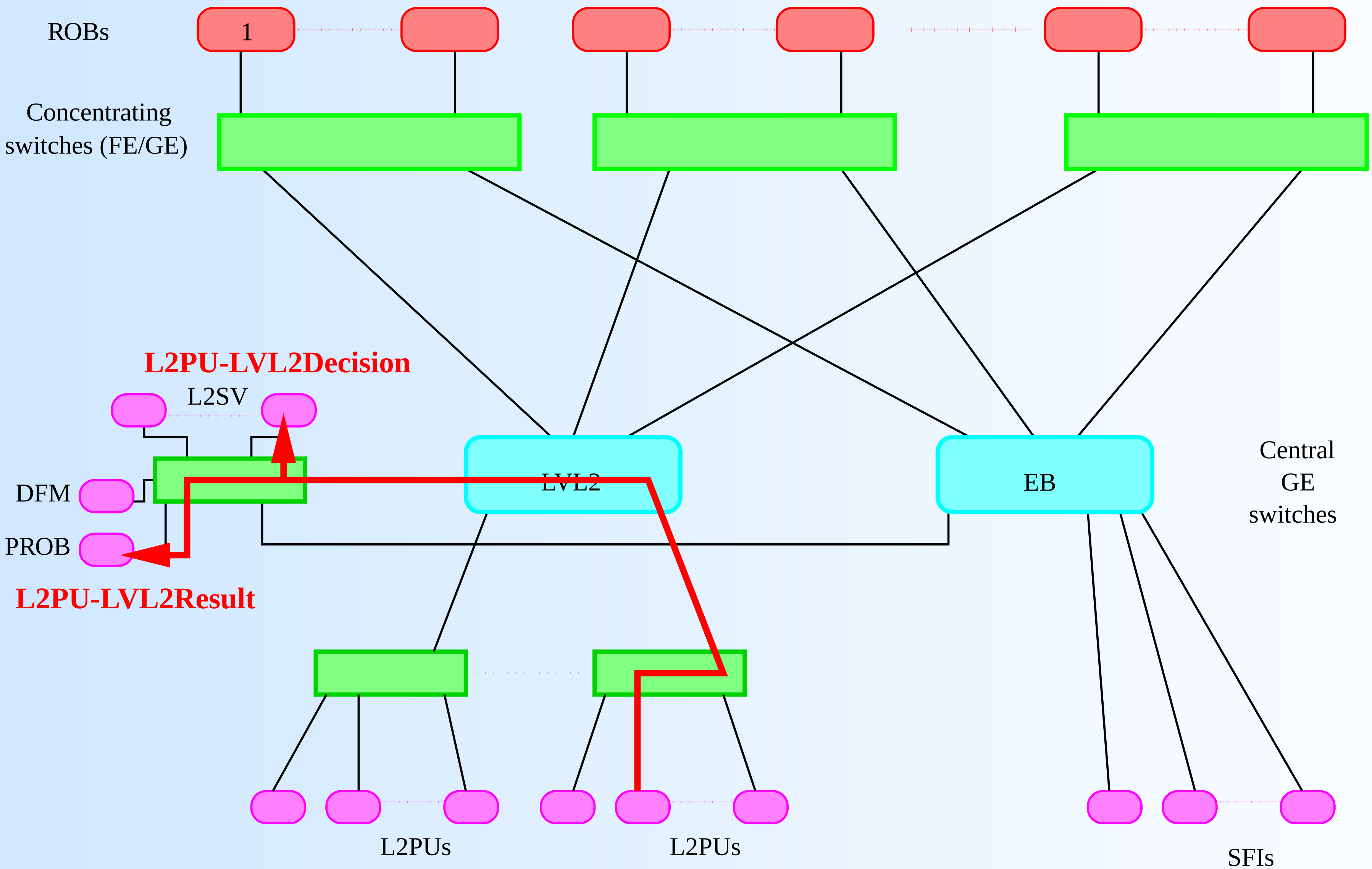
Message Flow [3]



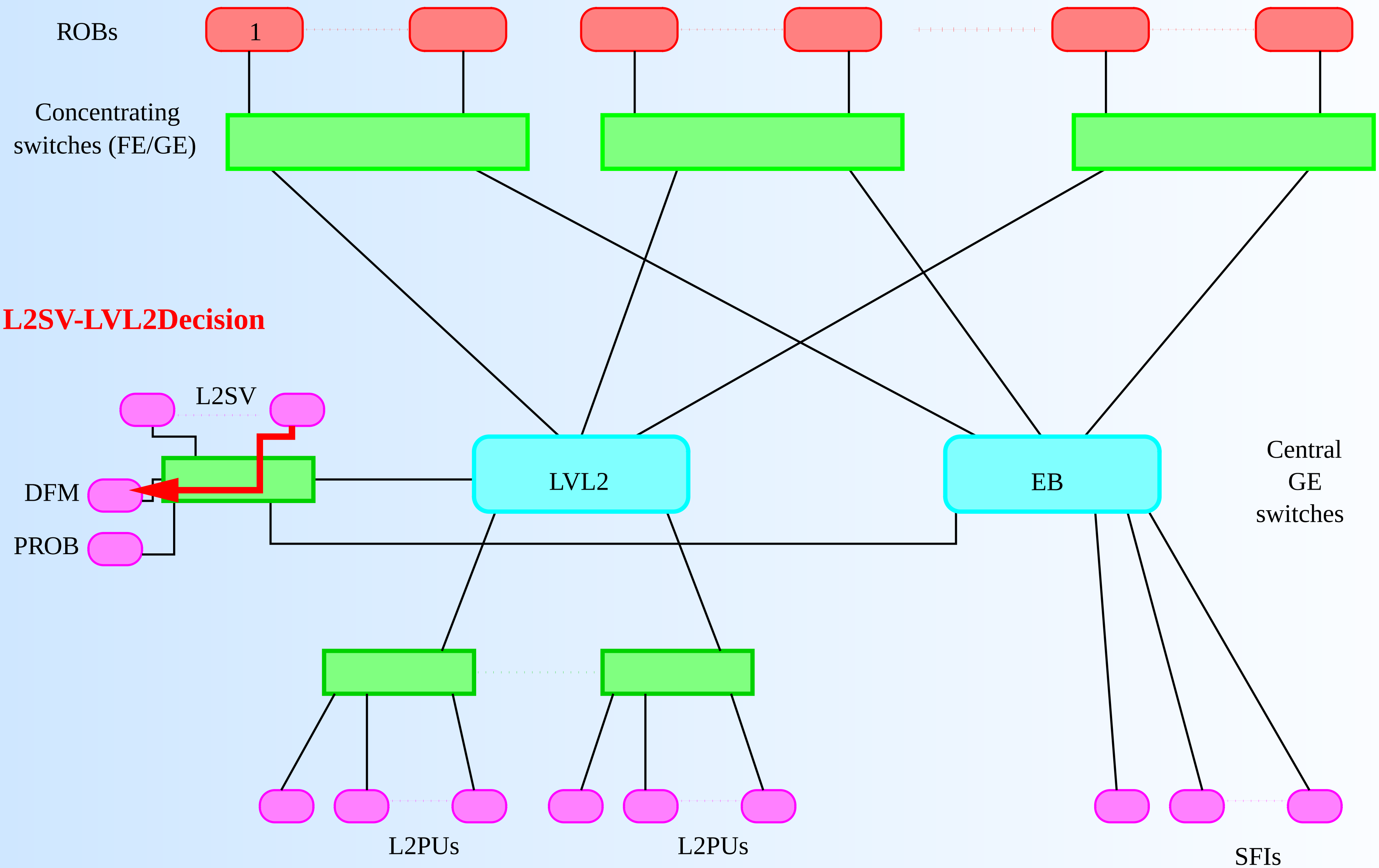
Message Flow [3]



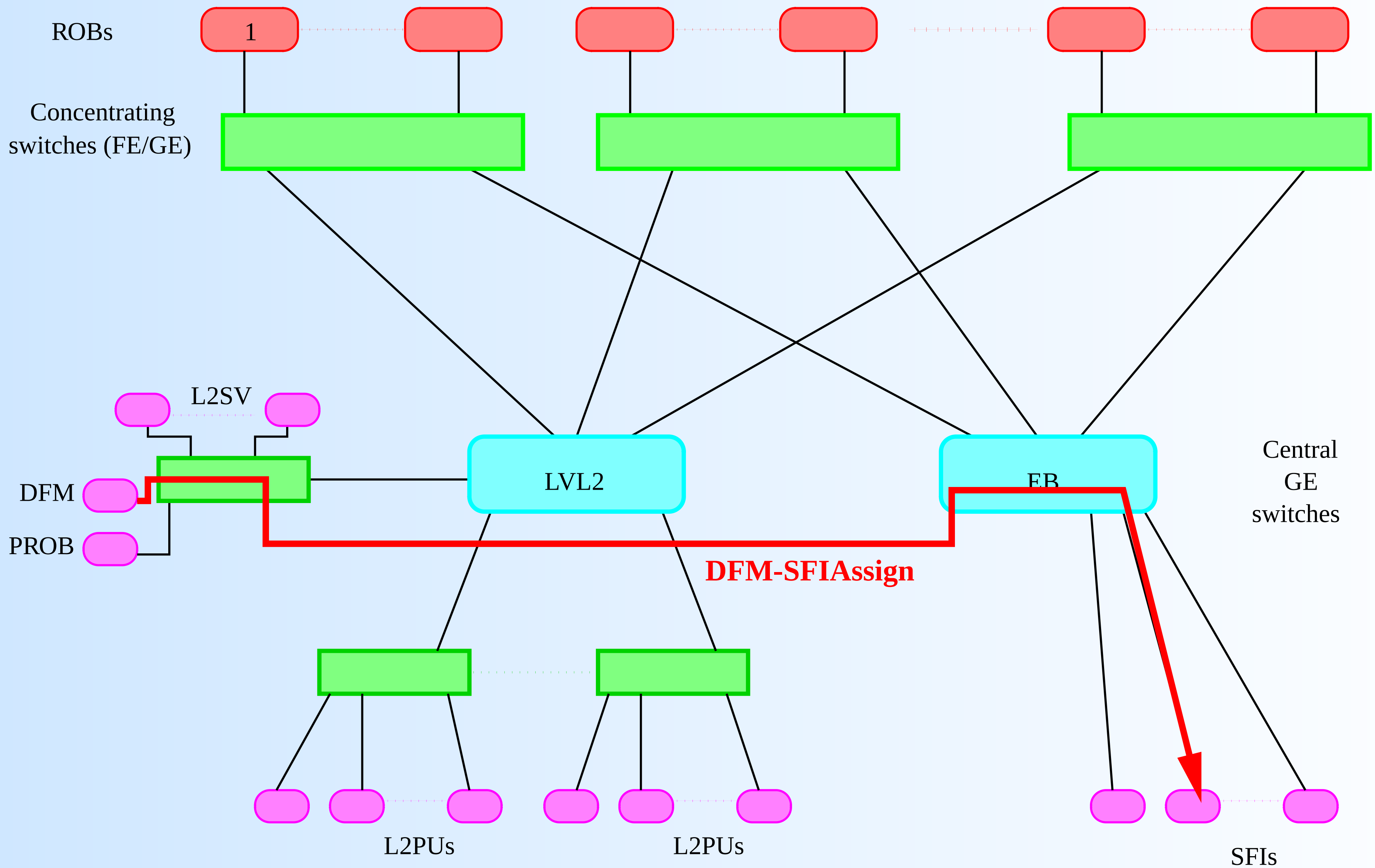
Message Flow [3]



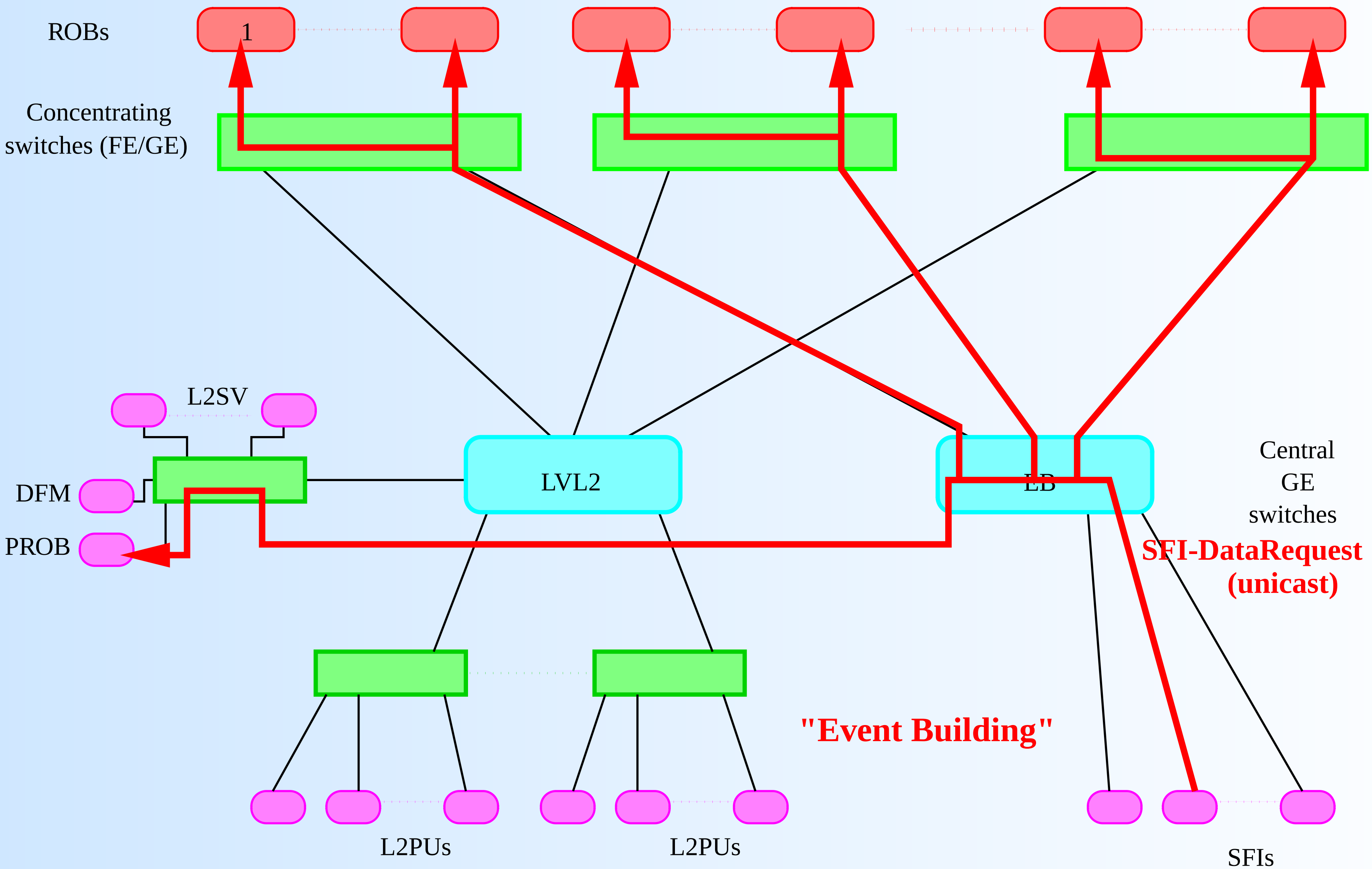
Message Flow [3]



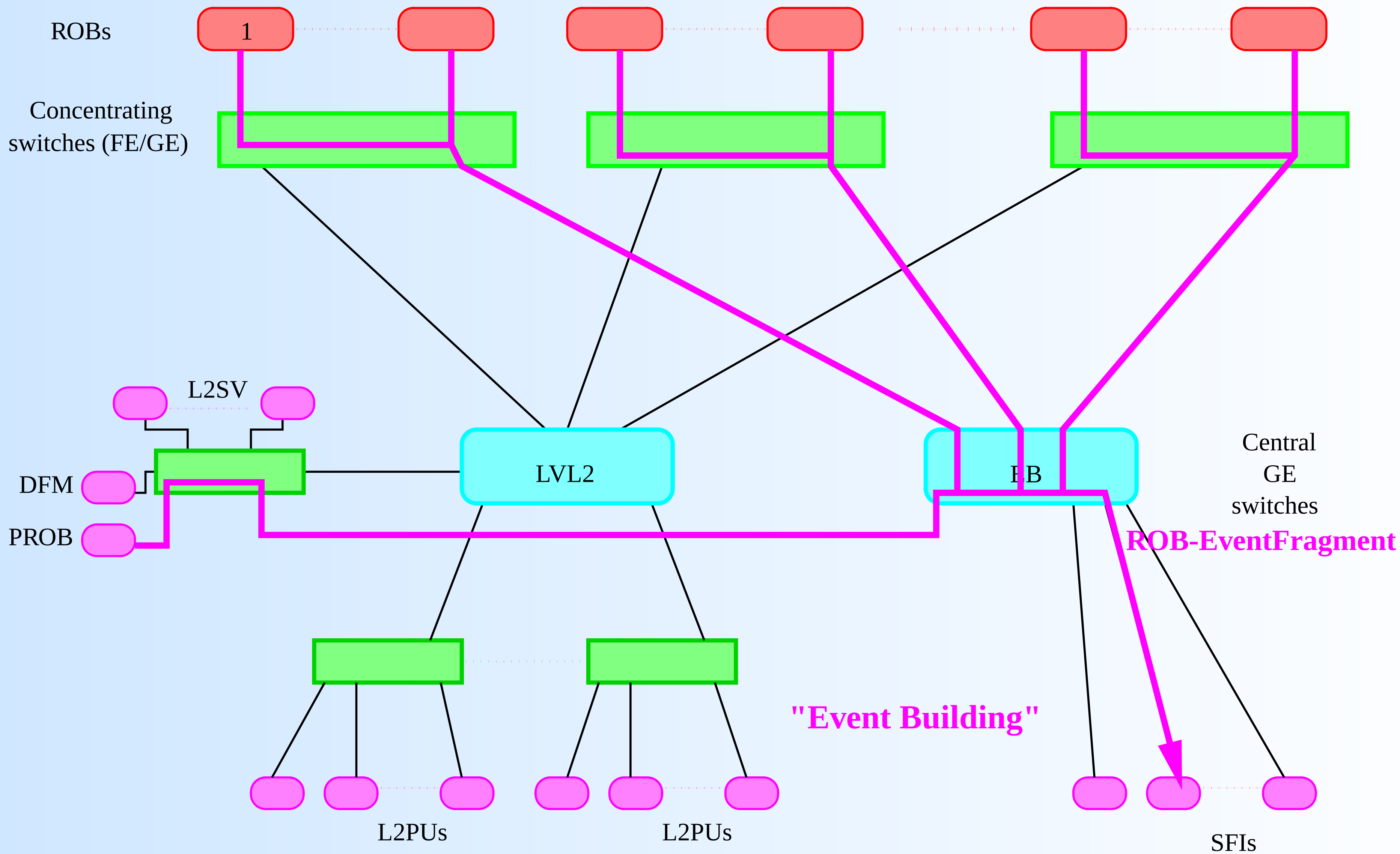
Message Flow [3]



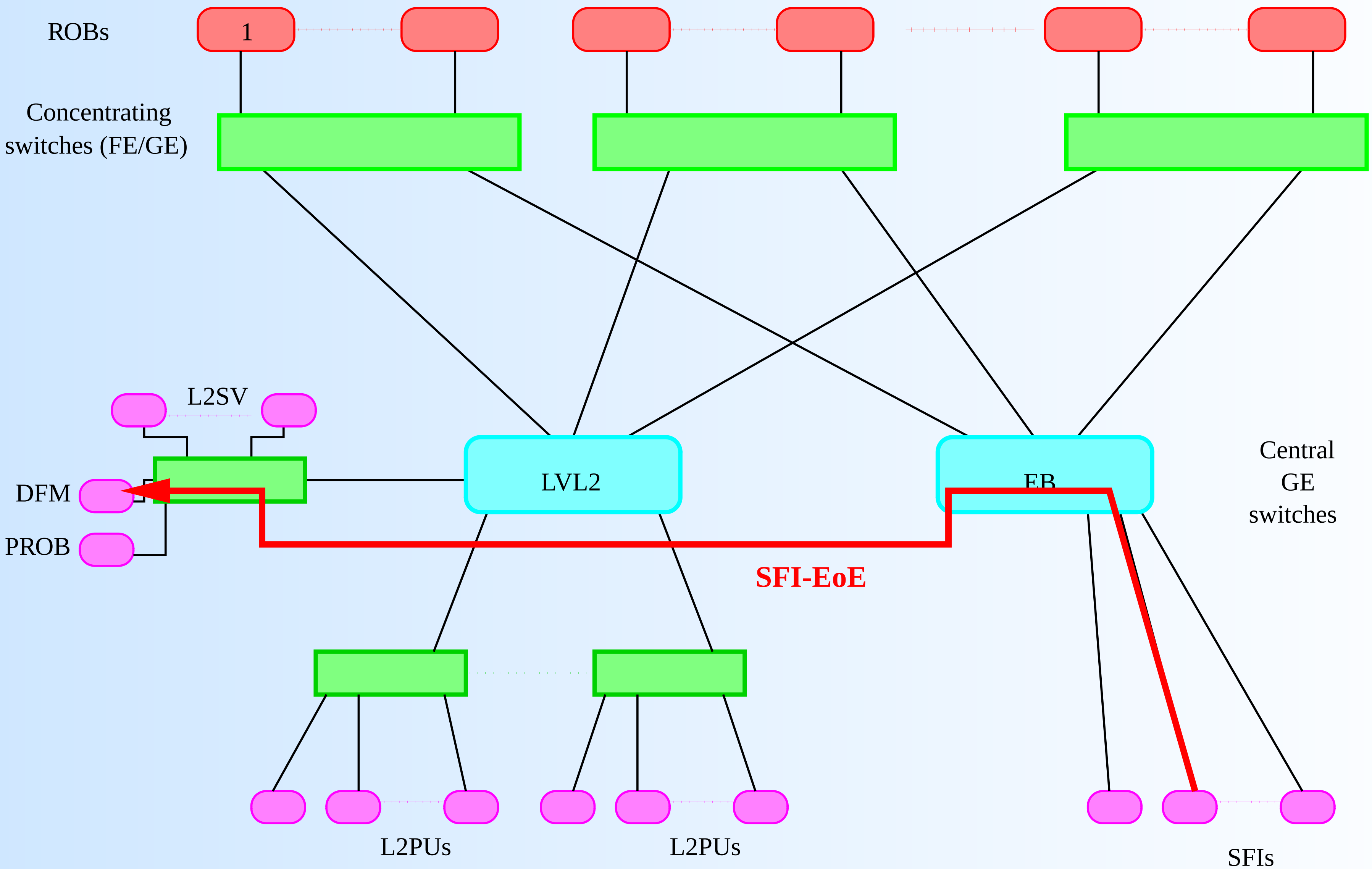
Message Flow [3]



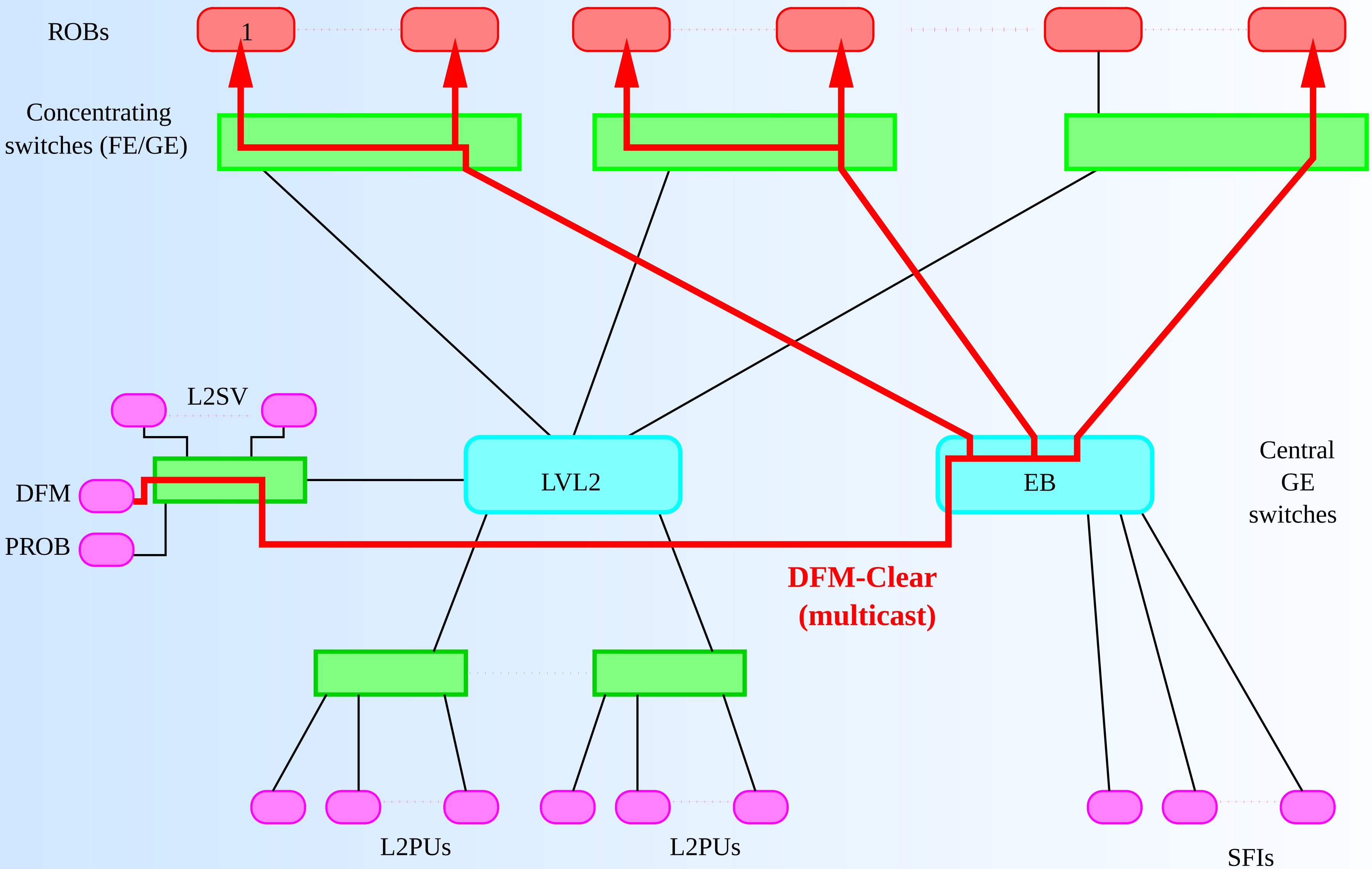
Message Flow [3]



Message Flow [3]



Message Flow [3]



Ethernet (IEEE 802.3) evolution

- **1973** – first lab implementation.

Ethernet (IEEE 802.3) evolution

- **1973** – first lab implementation.
- **1983** – IEEE 802.3: 10 Mbps, shared cable segment, half duplex, CSMA/CD, length limitation.

Ethernet (IEEE 802.3) evolution

- **1973** – first lab implementation.
- **1983** – IEEE 802.3: 10 Mbps, shared cable segment, half duplex, CSMA/CD, length limitation.
- **1990** – 10 Base T: 10 Mbps, UTP (unshielded twisted pair), half duplex, hub based star topology, bridges.

Ethernet (IEEE 802.3) evolution

- **1973** – first lab implementation.
- **1983** – IEEE 802.3: 10 Mbps, shared cable segment, half duplex, CSMA/CD, length limitation.
- **1990** – 10 Base T: 10 Mbps, UTP (unshielded twisted pair), half duplex, hub based star topology, bridges.
- **1995** – 100 Mbps: UTP and fibre, half/full duplex, point to point switched.

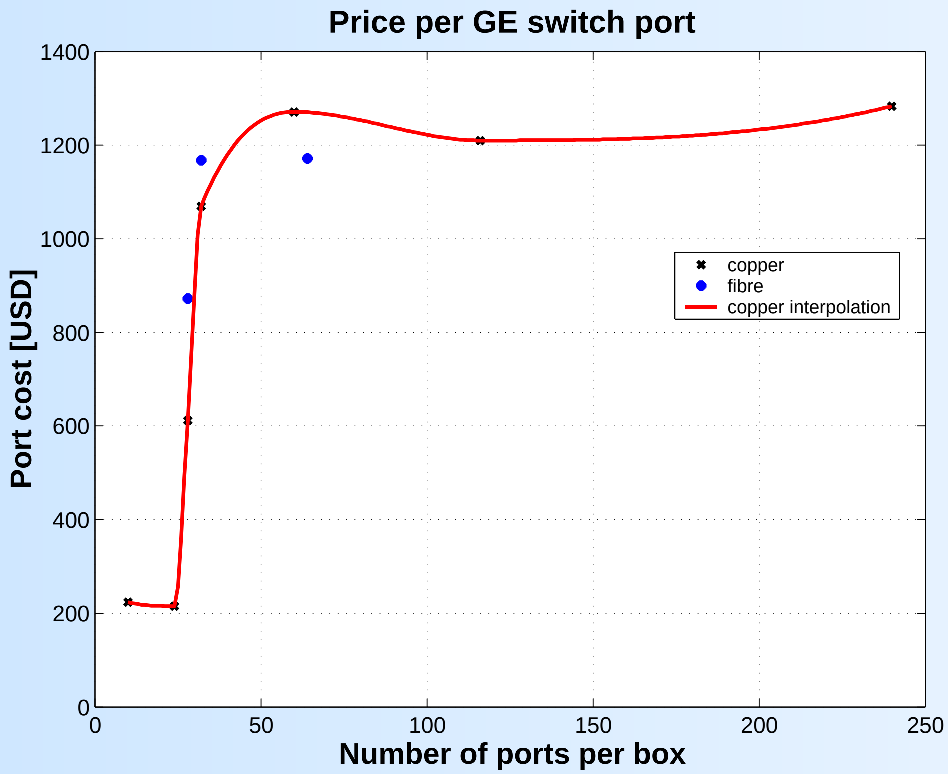
Ethernet (IEEE 802.3) evolution

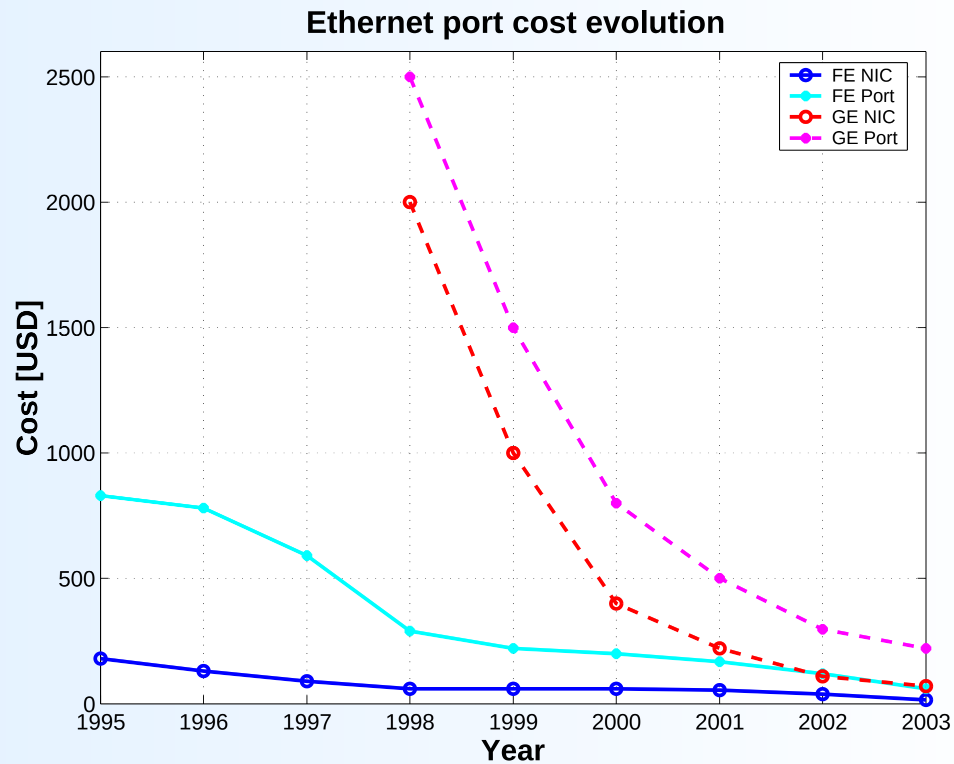
- **1973** – first lab implementation.
- **1983** – IEEE 802.3: 10 Mbps, shared cable segment, half duplex, CSMA/CD, length limitation.
- **1990** – 10 Base T: 10 Mbps, UTP (unshielded twisted pair), half duplex, hub based star topology, bridges.
- **1995** – 100 Mbps: UTP and fibre, half/full duplex, point to point switched.
- **1998** – GB Ethernet: fibre (also UTP in **1999**) half/full duplex (half duplex is not used in practice).

Ethernet (IEEE 802.3) evolution

- **1973** – first lab implementation.
- **1983** – IEEE 802.3: 10 Mbps, shared cable segment, half duplex, CSMA/CD, length limitation.
- **1990** – 10 Base T: 10 Mbps, UTP (unshielded twisted pair), half duplex, hub based star topology, bridges.
- **1995** – 100 Mbps: UTP and fibre, half/full duplex, point to point switched.
- **1998** – GB Ethernet: fibre (also UTP in **1999**) half/full duplex (half duplex is not used in practice).
- **2002** – 10 GB Ethernet: fibre, *only* full duplex.

Ethernet Port Cost





- Ethernet switches can be divided in:
  - ⋆ "Pizza" boxes (*concentrating switch*): $< 250 USD$ per port
  - ⋆ Big chassis (*central switch*): $> 1000 USD$ per port

- Most high end PCs have both FE and a GE NICs (Network Interface Cards) integrated on the motherboard.

Atlas Architecture – issues

- The *switches* must be able to handle the required throughput, with minimum packet loss and latency.

Atlas Architecture – issues

- The *switches* must be able to handle the required throughput, with minimum packet loss and latency.

- Special care should be taken in order to avoid packet loss, as it implies timeouts and retries from applications. Packet loss is almost entirely due to buffer overflow.

Atlas Architecture – issues

- The *switches* must be able to handle the required throughput, with minimum packet loss and latency.

- Special care should be taken in order to avoid packet loss, as it implies timeouts and retries from applications. Packet loss is almost entirely due to buffer overflow.
  - ⋆ Ethernet *Flow Control* offers protection only for short term congestions.
  - ⋆ Long term ones must be avoided by doing *traffic shaping* at the application level.

Atlas Architecture – issues

- The *switches* must be able to handle the required throughput, with minimum packet loss and latency.

- Special care should be taken in order to avoid packet loss, as it implies timeouts and retries from applications. Packet loss is almost entirely due to buffer overflow.
  - Ethernet *Flow Control* offers protection only for short term congestions.
  - Long term ones must be avoided by doing *traffic shaping* at the application level.

- We must avoid *multicast/broadcast* traffic, as it's rate is not guaranteed on all the switches. The main data flow is based on a *request–response* mechanism.

Atlas Architecture – issues

- The *switches* must be able to handle the required throughput, with minimum packet loss and latency.

- Special care should be taken in order to avoid packet loss, as it implies timeouts and retries from applications. Packet loss is almost entirely due to buffer overflow.
  - ⋆ Ethernet *Flow Control* offers protection only for short term congestions.
  - ⋆ Long term ones must be avoided by doing *traffic shaping* at the application level.

- We must avoid *multicast/broadcast* traffic, as it's rate is not guaranteed on all the switches. The main data flow is based on a *request–response* mechanism.

- *VLAN* support must be present in order to avoid Ethernet loops.

Atlas Architecture – issues

- The *switches* must be able to handle the required throughput, with minimum packet loss and latency.

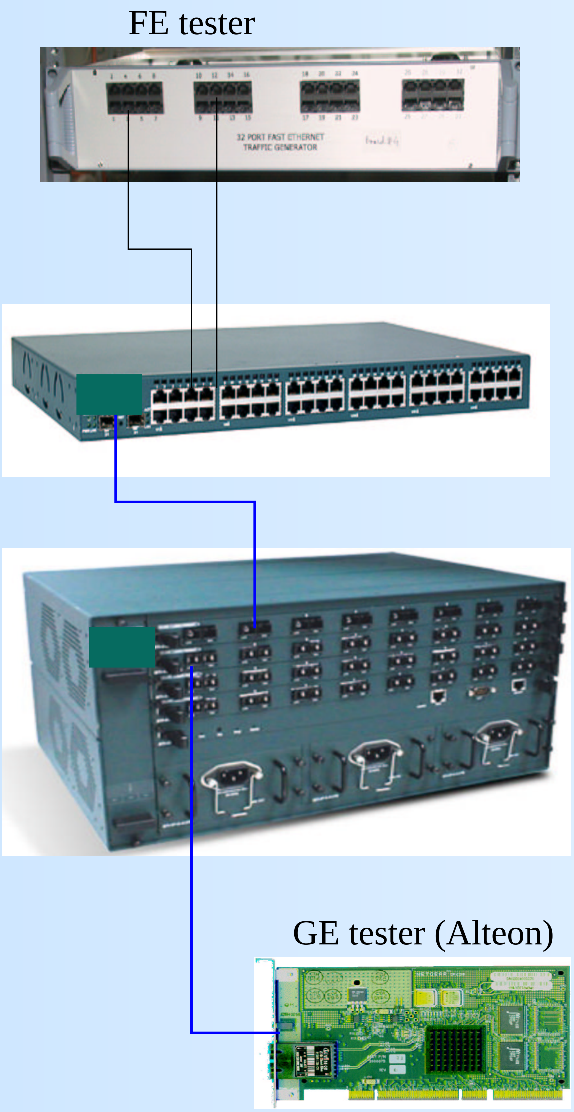
- Special care should be taken in order to avoid packet loss, as it implies timeouts and retries from applications. Packet loss is almost entirely due to buffer overflow.
  - ⋆ Ethernet *Flow Control* offers protection only for short term congestions.
  - ⋆ Long term ones must be avoided by doing *traffic shaping* at the application level.

- We must avoid *multicast/broadcast* traffic, as it's rate is not guaranteed on all the switches. The main data flow is based on a *request–response* mechanism.

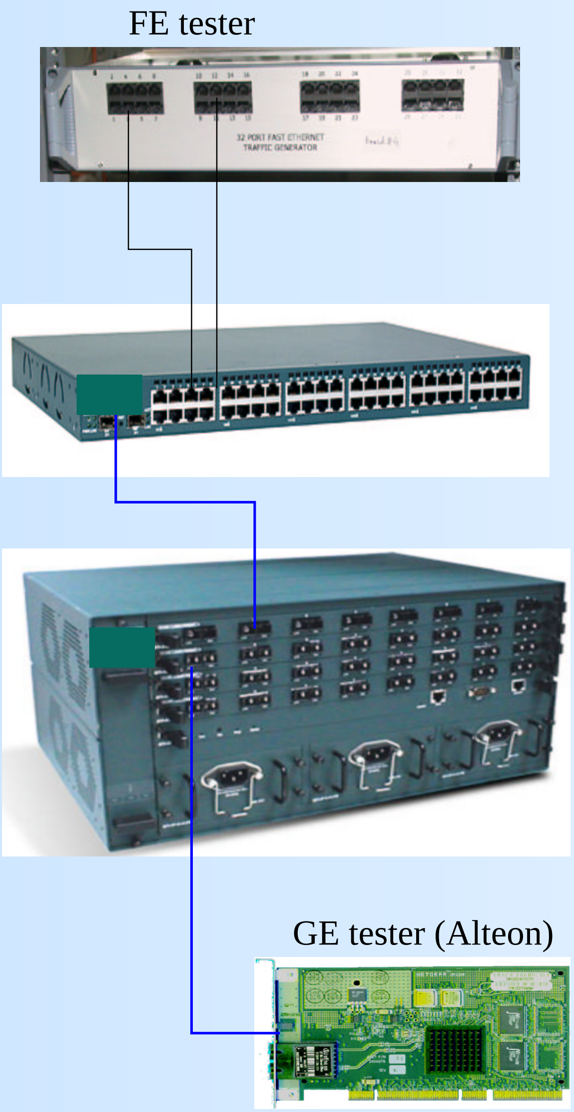
- *VLAN* support must be present in order to avoid Ethernet loops.

- The DataCollection network has a large number of nodes (≈ 2000), which must be memorized in the *MAC address table* of the switches.

Network Testing Equipment

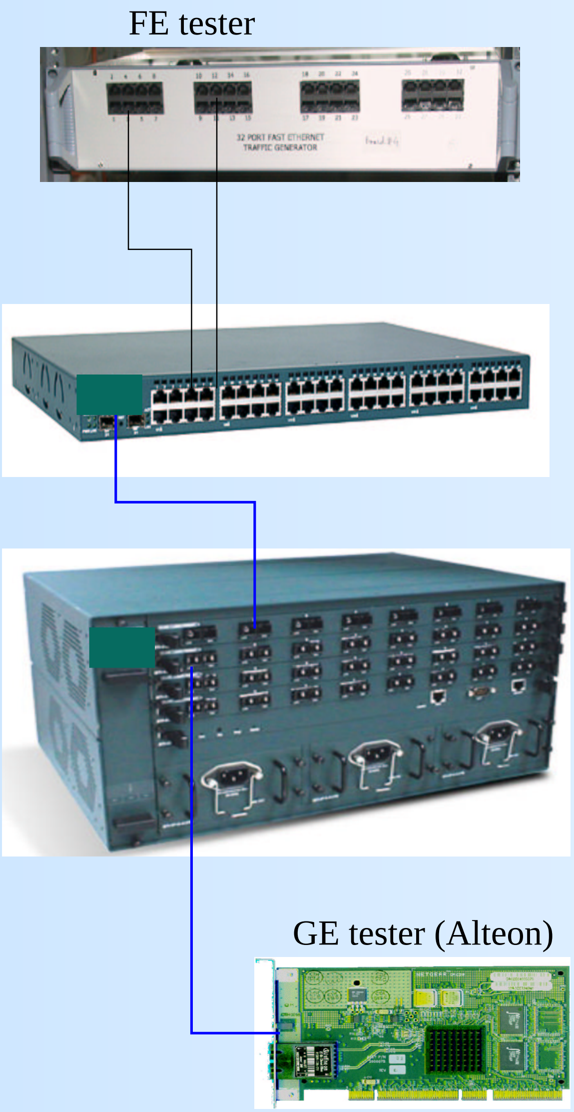


Network Testing Equipment



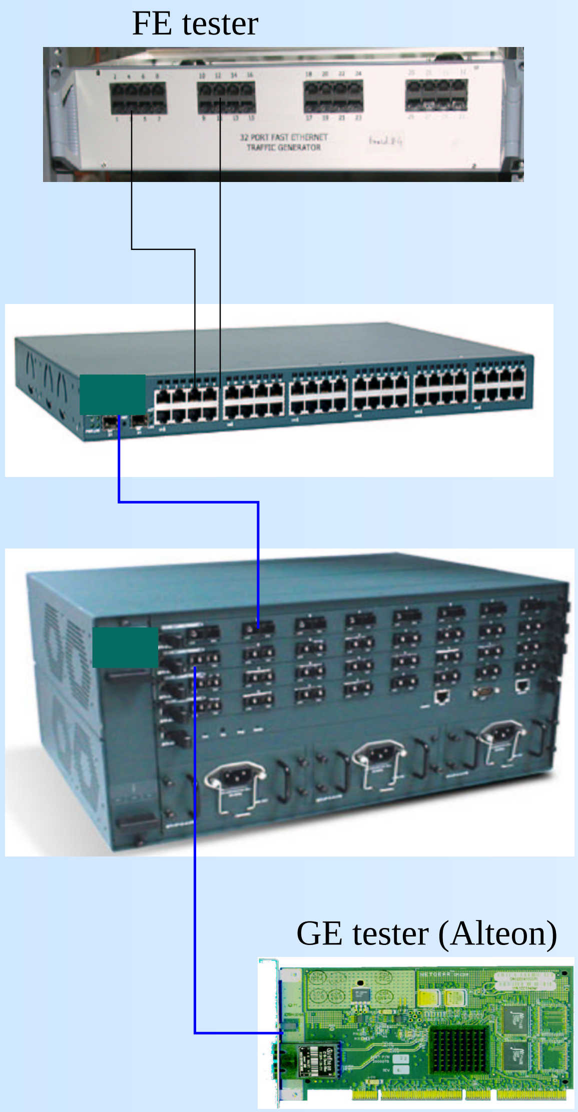
- FE tester (FPGA based) – 32 ports per board.

Network Testing Equipment



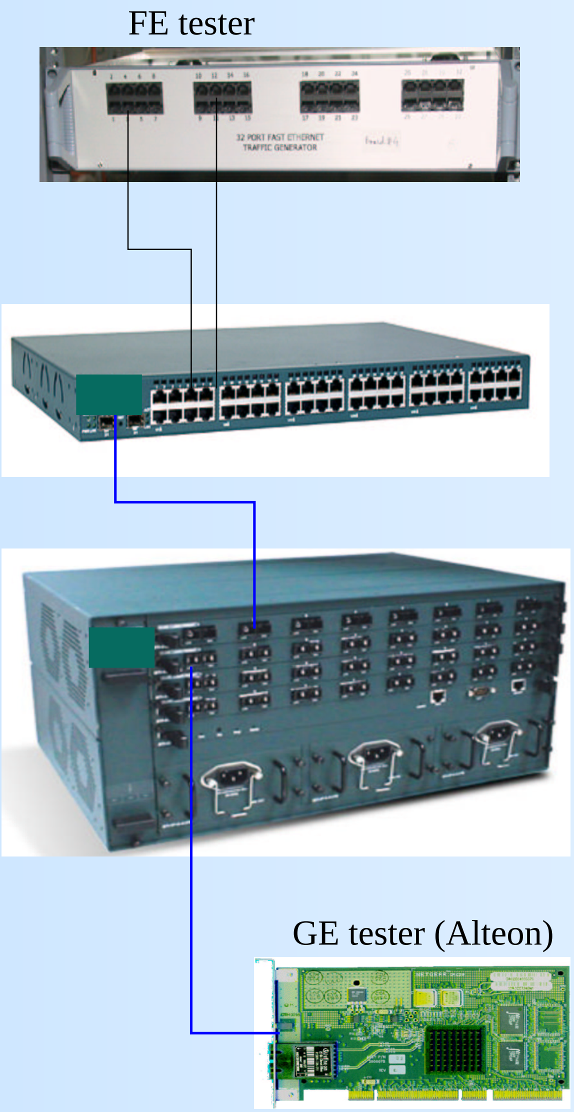
- FE tester (FPGA based) – 32 ports per board.
- GE tester (Alteon programmable NIC)– approx 30 NICs.

Network Testing Equipment



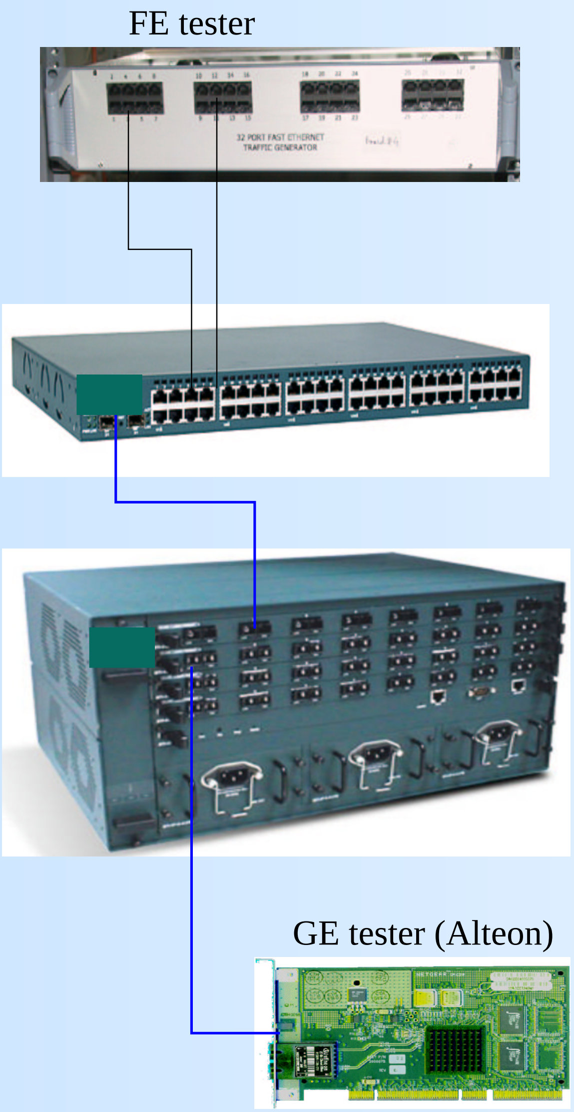
- FE tester (FPGA based) – 32 ports per board.
- GE tester (Alteon programmable NIC)– approx 30 NICs.
- A description of these testers can be found in [1].

Network Testing Equipment



- FE tester (FPGA based) – 32 ports per board.
- GE tester (Alteon programmable NIC)– approx 30 NICs.
- A description of these testers can be found in [1].
- The testers can:
  - ⋆ measure *packet loss*
  - ⋆ measure *average latency*
  - ⋆ histogram the latency (*jitter*).

Network Testing Equipment



- FE tester (FPGA based) – 32 ports per board.
- GE tester (Alteon programmable NIC)– approx 30 NICs.
- A description of these testers can be found in [1].
- The testers can:
  - ⋆ measure *packet loss*
  - ⋆ measure *average latency*
  - ⋆ histogram the latency (*jitter*).
- We have used this HW to investigate several central and concentrating switches from different manufacturers.

Switches

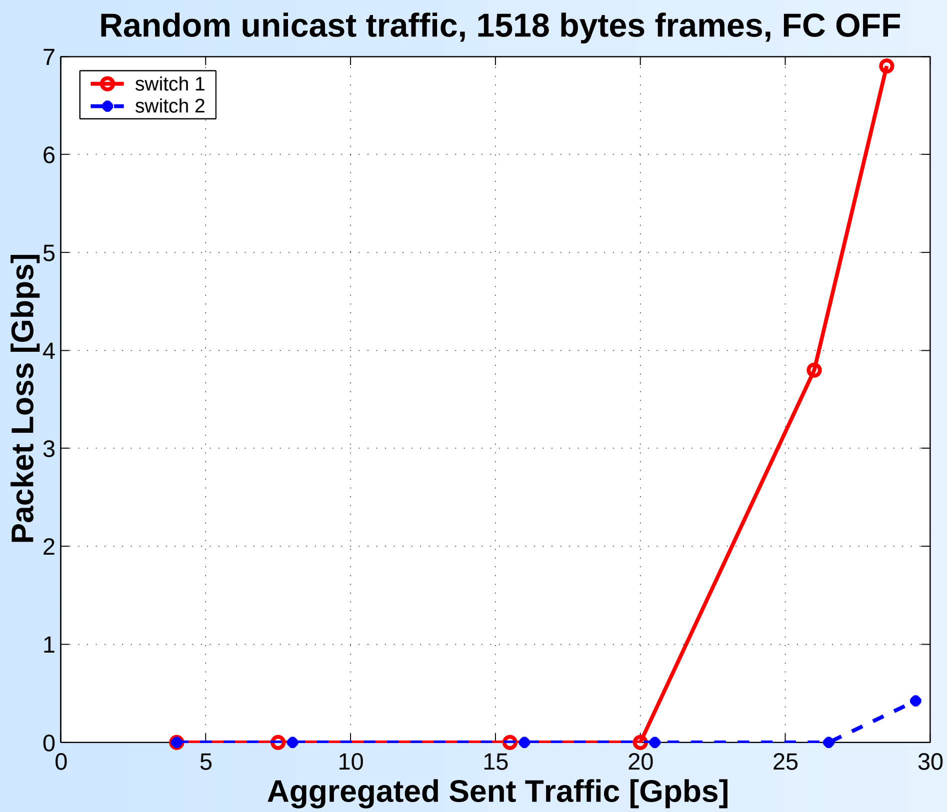
- They are the key element for the Baseline Architecture architecture:
  - ⋆ They must handle the throughput of the DataCollection network.
  - ⋆ The buffering capabilities and Flow Control implementation determine the network's resistance to packet loss.

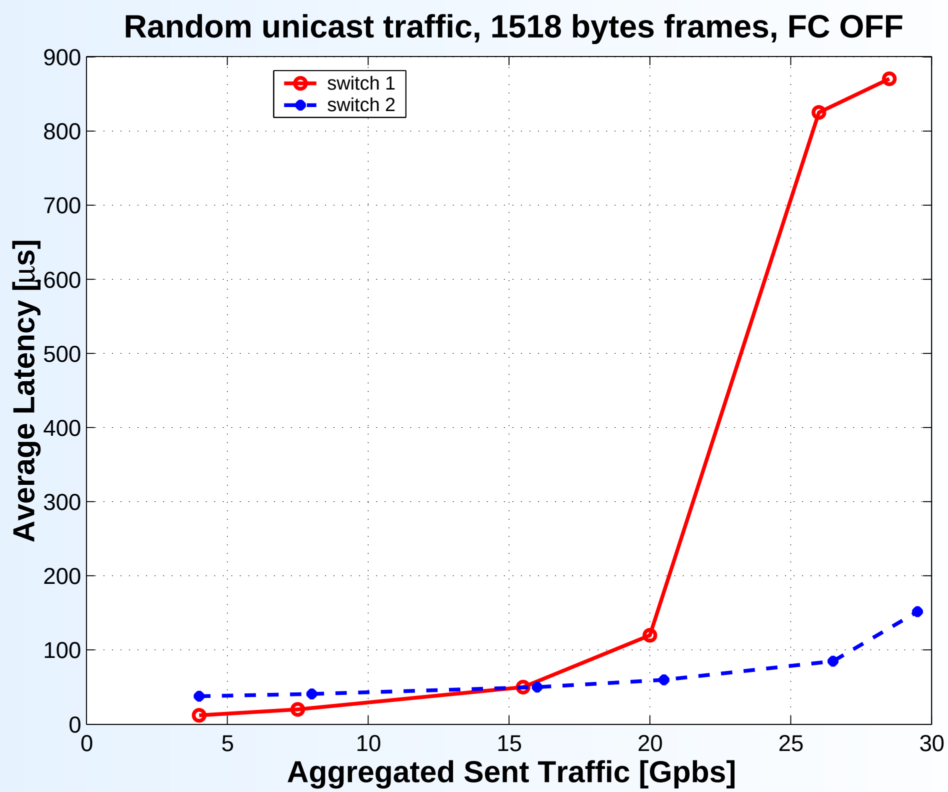
Switches

- They are the key element for the Baseline Architecture architecture:
  - They must handle the throughput of the DataCollection network.
  - The buffering capabilities and Flow Control implementation determine the network's resistance to packet loss.

- What we measure:
  - `throughput` vs `load`, for different frame sizes
  - `latency` vs `load` for different frames sizes
  - unicast, multicast and broadcast traffic
  - DataCollection traffic type:
    - request – response: small requests, large responses
    - low rate multicast (300 Hz) – distribution of clear messages to the ROBs.

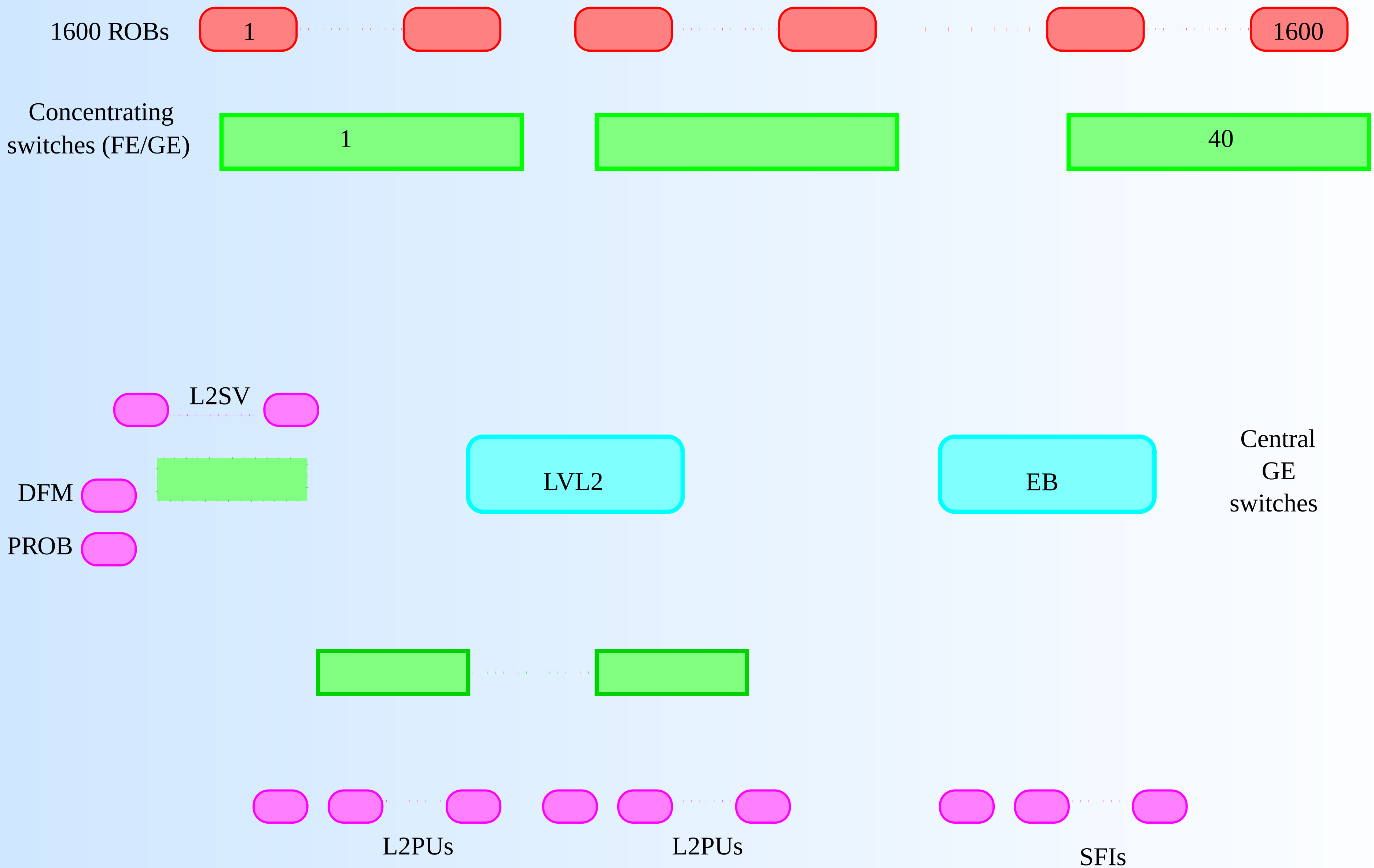
Switches – results



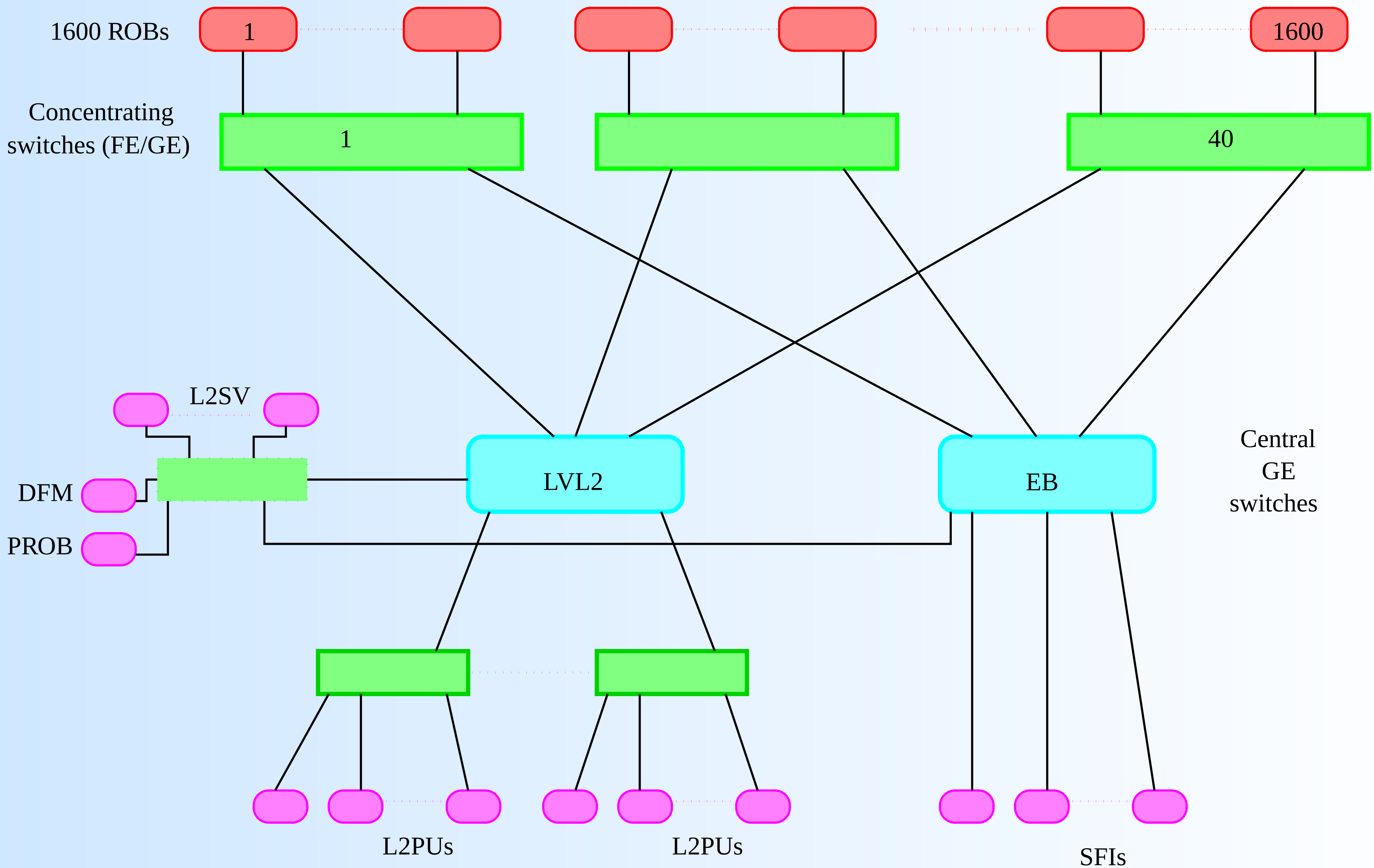


- 30 GE ports. Each one sends random traffic (Poisson) to all the other ports.
- At saturation latency increases and packets are dropped.
- We need to operate the switches *below saturation*.

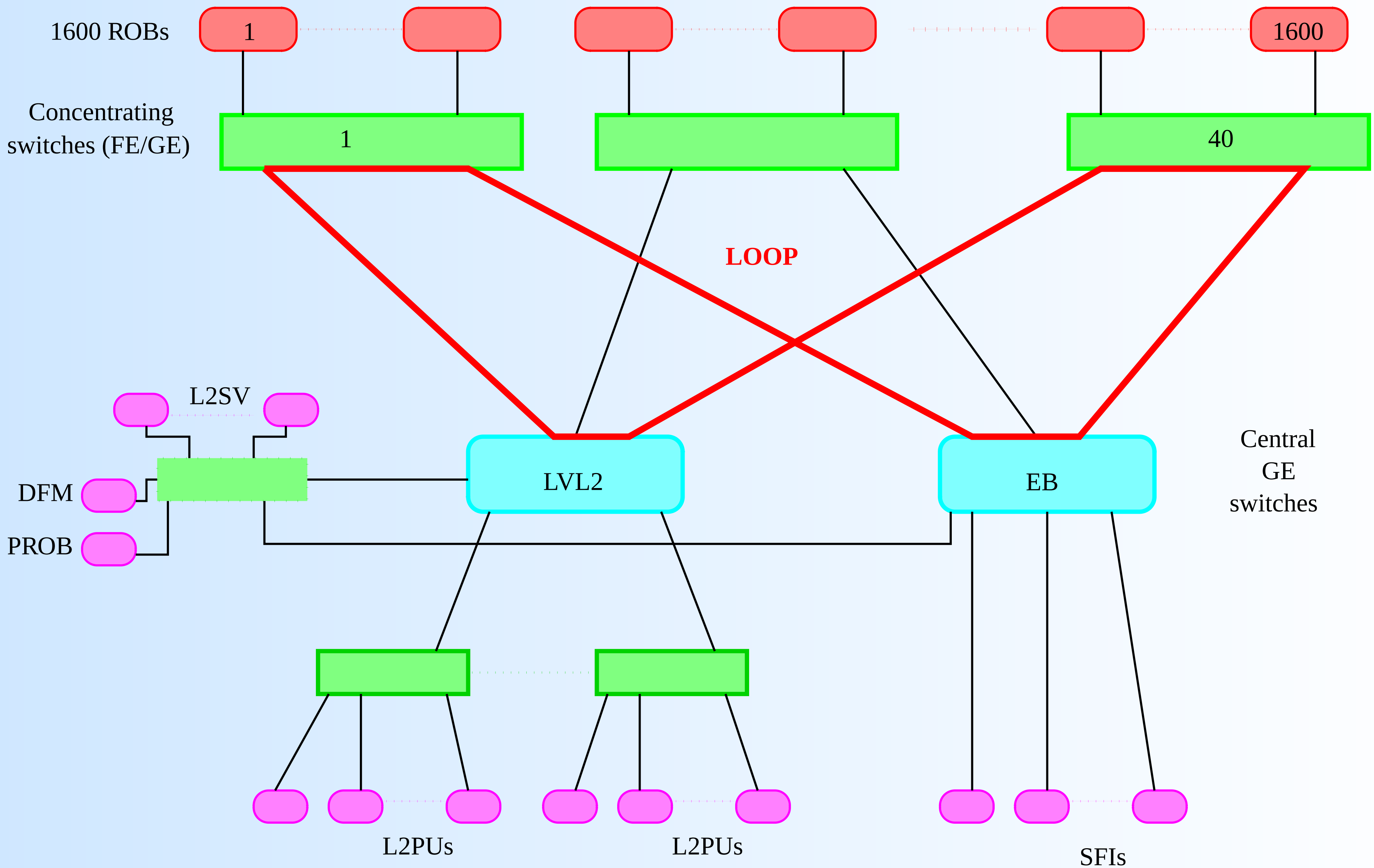
VLANs – loops


1600 ROBs
1
1600
Concentrating switches (FE/GE)
1
40
L2SV
DFM
PROB
LVL2
EB
Central GE switches
L2PUs
L2PUs
SFIs

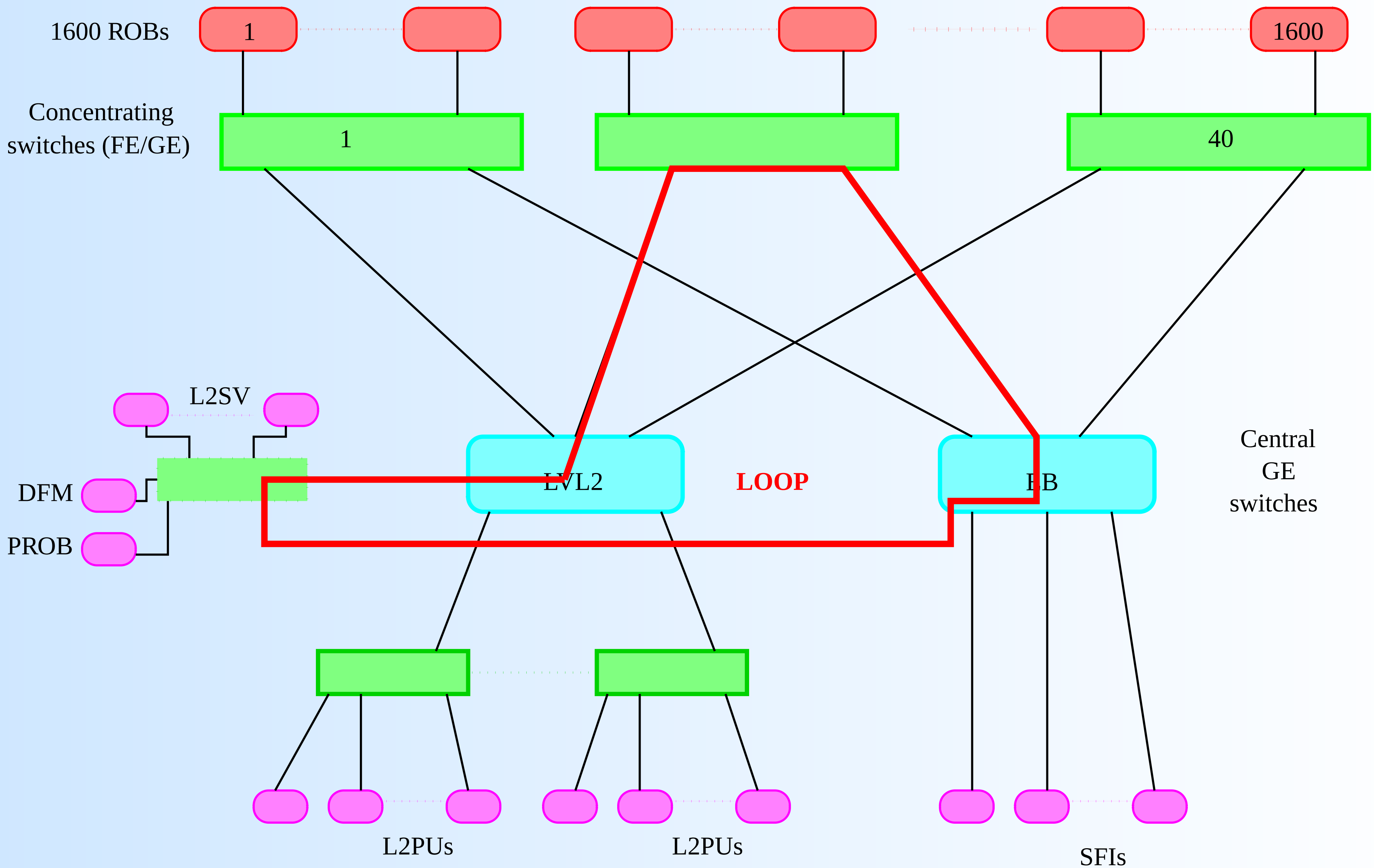
VLANs – loops


1600 ROBs
1
1600
Concentrating
switches (FE/GE)
1
40
L2SV
DFM
PROB
LVL2
EB
Central
GE
switches
L2PUs
L2PUs
SFIs

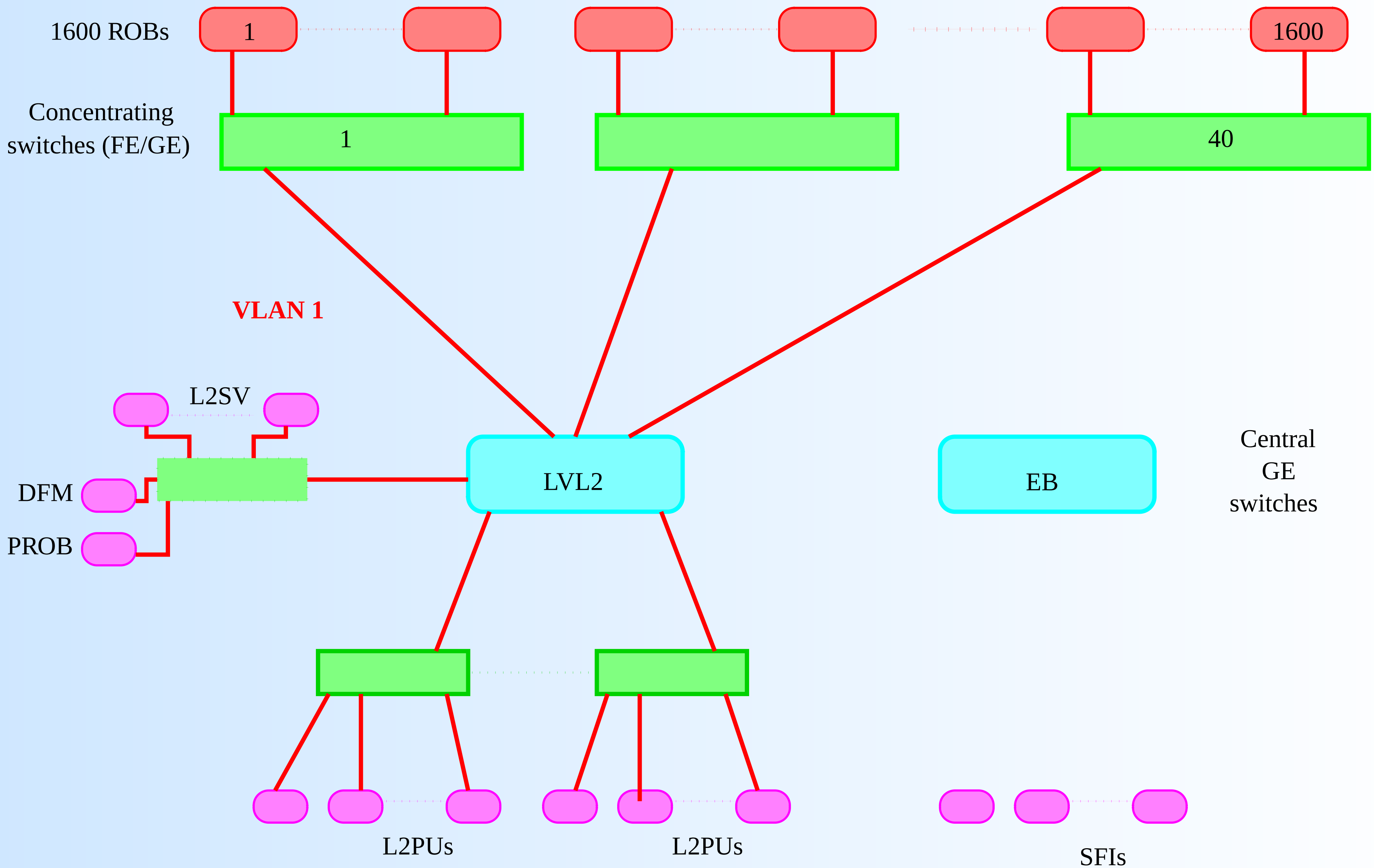
VLANs – loops


1600 ROBs
1
1600
Concentrating switches (FE/GE)
1
40
LOOP
L2SV
DFM
PROB
LVL2
EB
Central GE switches
L2PUs
L2PUs
SFIs

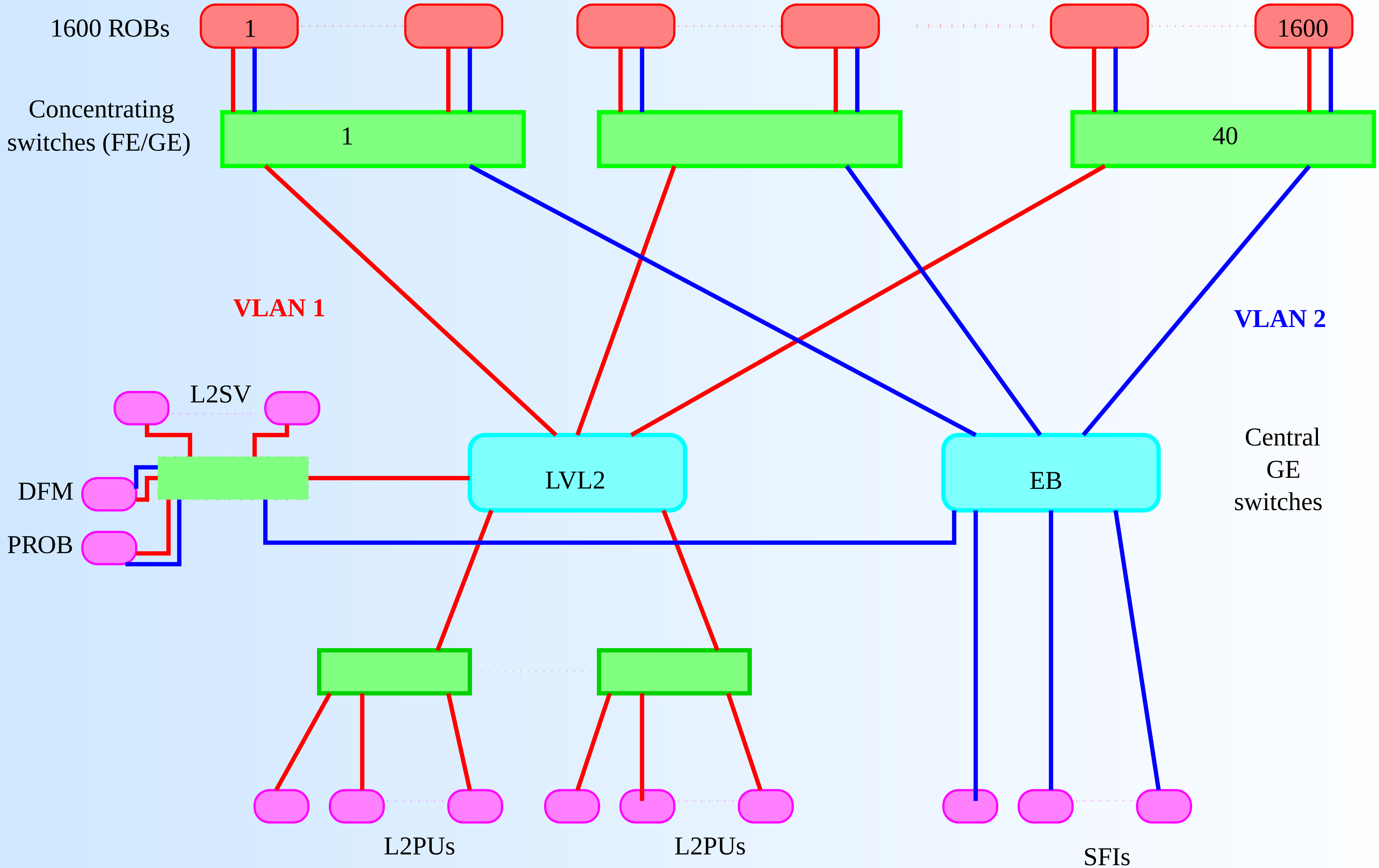
VLANs – loops


1600 ROBs
1
1600
Concentrating switches (FE/GE)
1
40
L2SV
DFM
PROB
LVL2
LOOP
EB
Central GE switches
L2PUs
L2PUs
SFIs

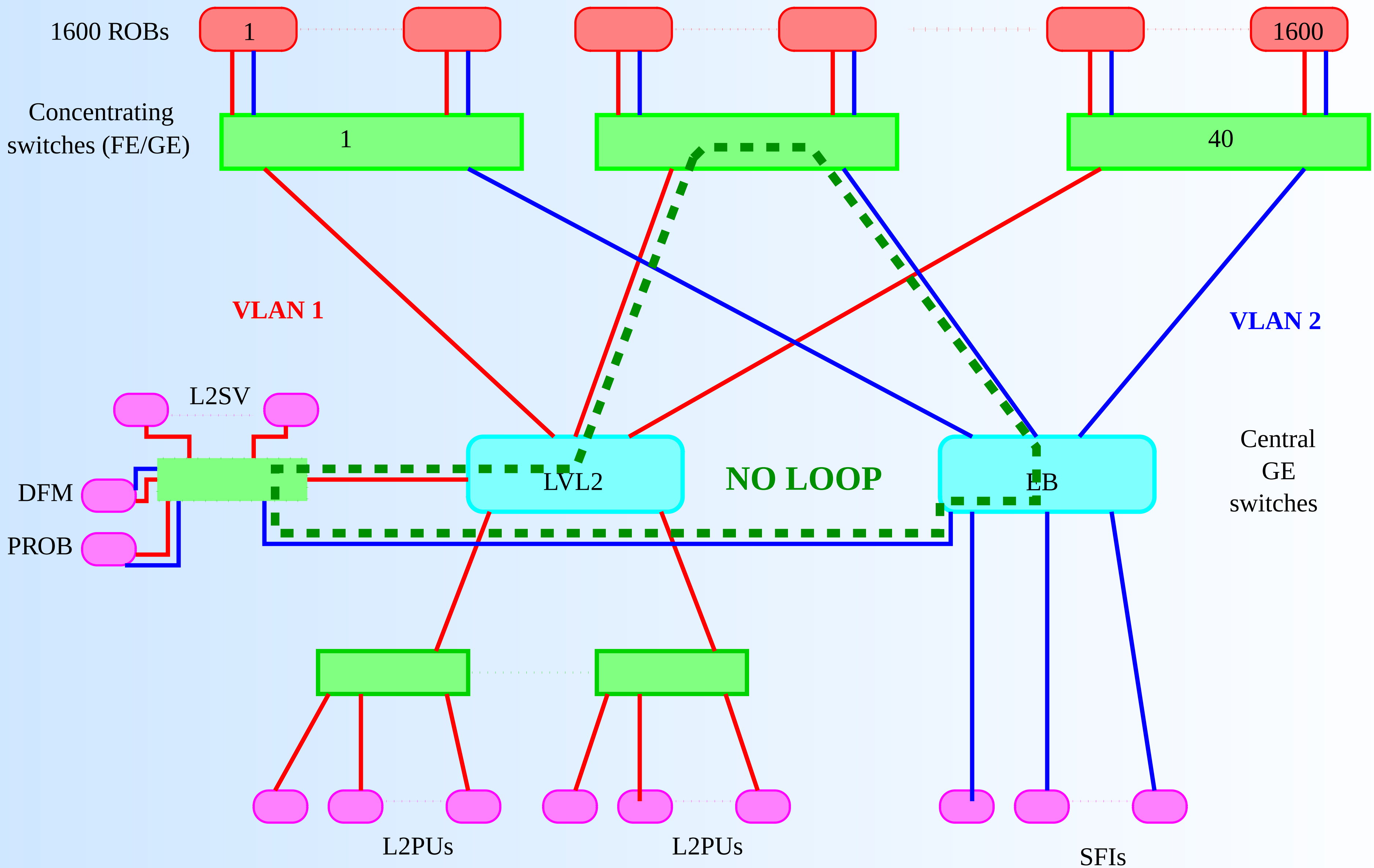
VLANs – loops


1600 ROBs
1
1600
Concentrating switches (FE/GE)
1
40
VLAN 1
L2SV
DFM
PROB
LVL2
EB
Central GE switches
L2PUs
L2PUs
SFIs

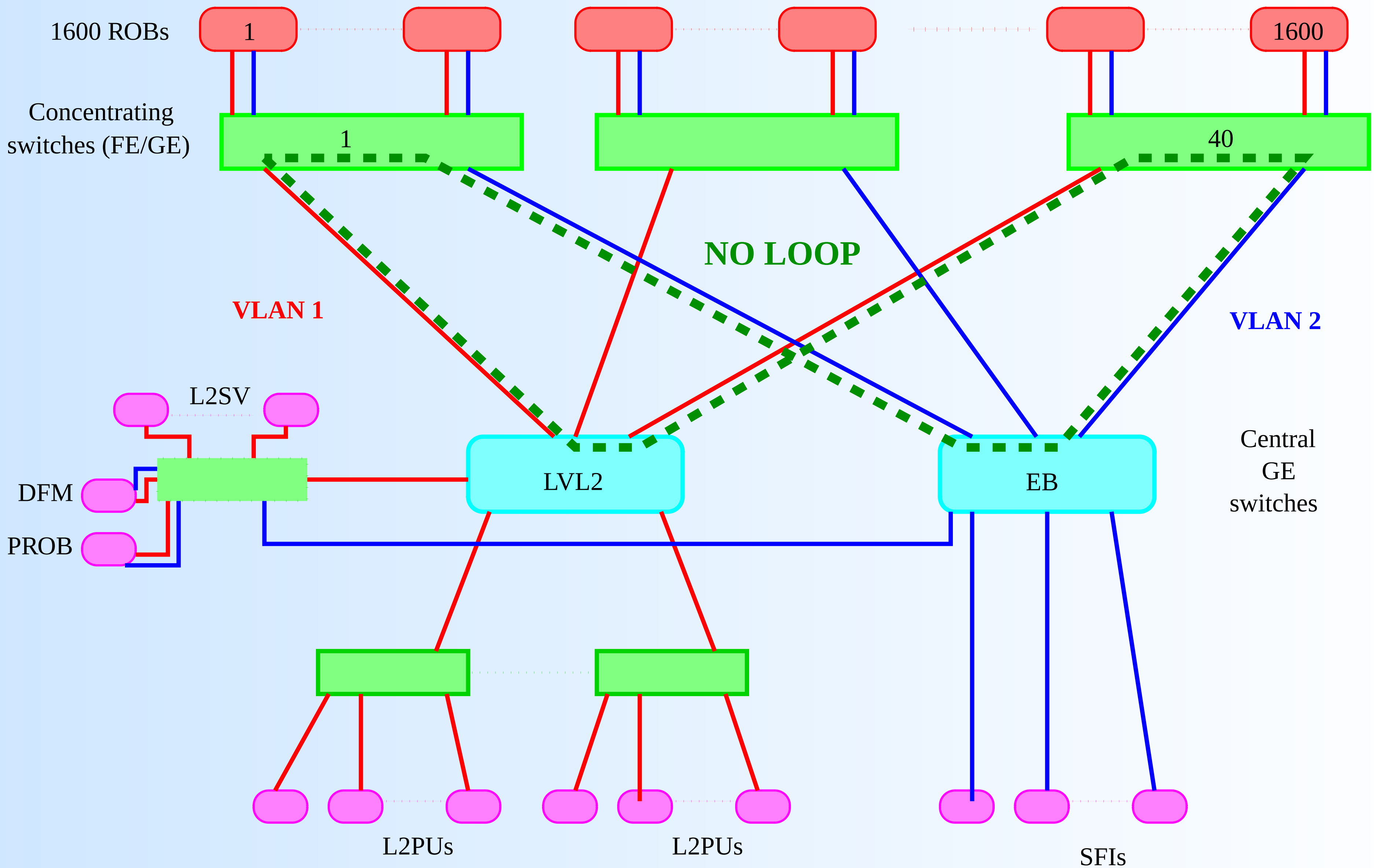
VLANs – loops


1600 ROBs
1
1600
Concentrating switches (FE/GE)
1
40
VLAN 1
VLAN 2
L2SV
DFM
PROB
LVL2
EB
Central GE switches
L2PUs
L2PUs
SFIs

VLANs – loops


1600 ROBs
1
1600
Concentrating
switches (FE/GE)
1
40
VLAN 1
VLAN 2
L2SV
DFM
PROB
LVL2
NO LOOP
EB
Central
GE
switches
L2PUs
L2PUs
SFIs

VLANs – loops


1600 ROBs
1
1600
Concentrating switches (FE/GE)
1
40
NO LOOP
VLAN 1
VLAN 2
L2SV
DFM
PROB
LVL2
EB
Central GE switches
L2PUs
L2PUs
SFIs

VLANs – loops & STP

- We need VLANs in order to avoid illegal Ethernet loops introduced by the use of more central switches. Ethernet loops are forbidden as they cause broadcast storms and corrupt the MAC address tables in the switches.

VLANs – loops & STP

- We need VLANs in order to avoid illegal Ethernet loops introduced by the use of more central switches. Ethernet loops are forbidden as they cause broadcast storms and corrupt the MAC address tables in the switches.

- The layer 2 spanning tree protocol (STP) disables the redundant links from a LAN maintaining a loop free topology. The ATLAS Architecture works if:
  - ⋆ STP is implemented per VLAN.
  - ⋆ STP is not implemented per VLAN, but we disable it and make sure there is no loop.

VLANs – loops & STP

- We need VLANs in order to avoid illegal Ethernet loops introduced by the use of more central switches. Ethernet loops are forbidden as they cause broadcast storms and corrupt the MAC address tables in the switches.

- The layer 2 spanning tree protocol (STP) disables the redundant links from a LAN maintaining a loop free topology. The ATLAS Architecture works if:
  - ⋆ STP is implemented per VLAN.
  - ⋆ STP is not implemented per VLAN, but we disable it and make sure there is no loop.

- VLANs are helpful to restrict flooding, broadcast and multicast.

VLANs – loops & STP

- We need VLANs in order to avoid illegal Ethernet loops introduced by the use of more central switches. Ethernet loops are forbidden as they cause broadcast storms and corrupt the MAC address tables in the switches.

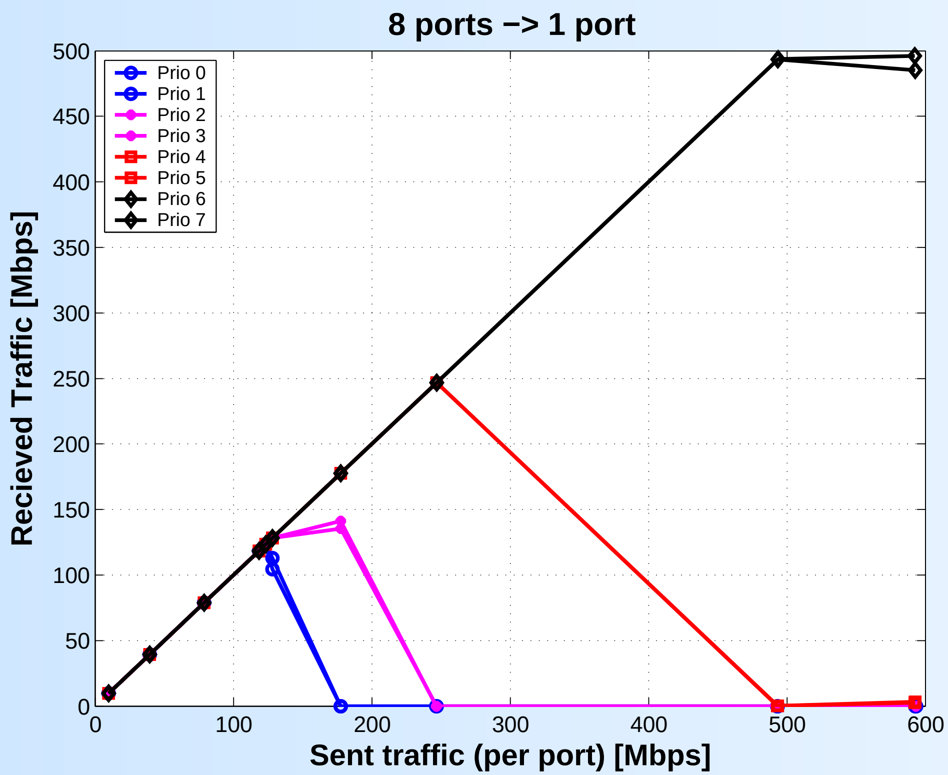
- The layer 2 spanning tree protocol (STP) disables the redundant links from a LAN maintaining a loop free topology. The ATLAS Architecture works if:
  - ⋆ STP is implemented per VLAN.
  - ⋆ STP is not implemented per VLAN, but we disable it and make sure there is no loop.

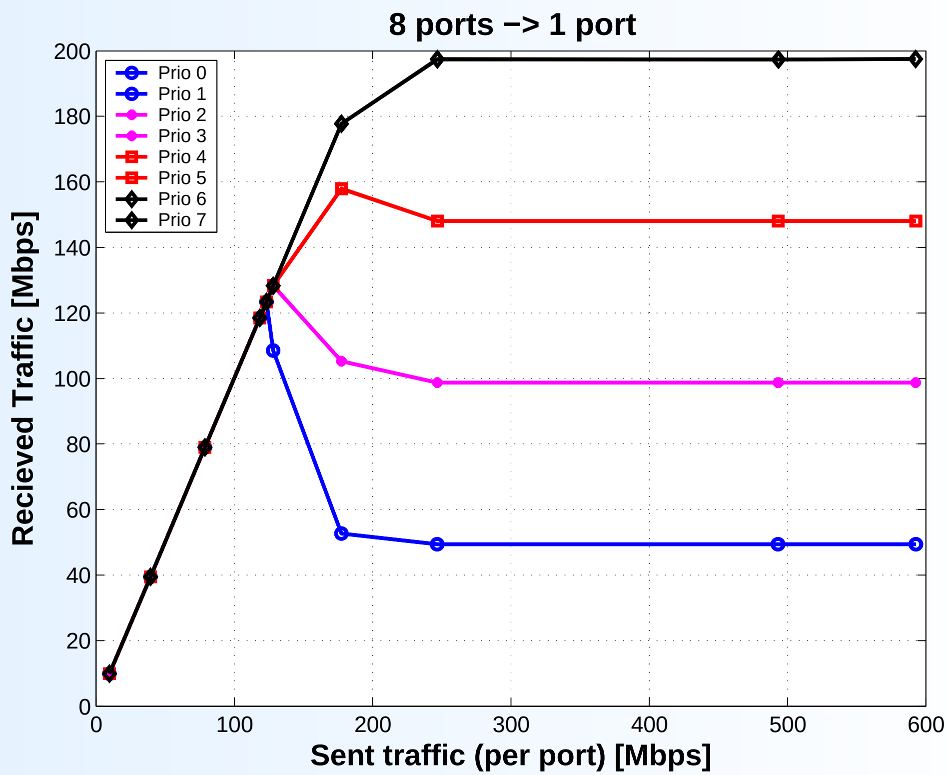
- VLANs are helpful to restrict flooding, broadcast and multicast.

- The *priority* field from the VLAN tag allows Ethernet traffic to be prioritized (QoS). We can assign a higher priority to the control messages with respect to the main data flow.

VLANs – QoS

- There are 8 transmitting ports and one receiving port. Each transmitter sends traffic (constant bit rate, 1518 bytes frames) with a different priority to the receiving port.

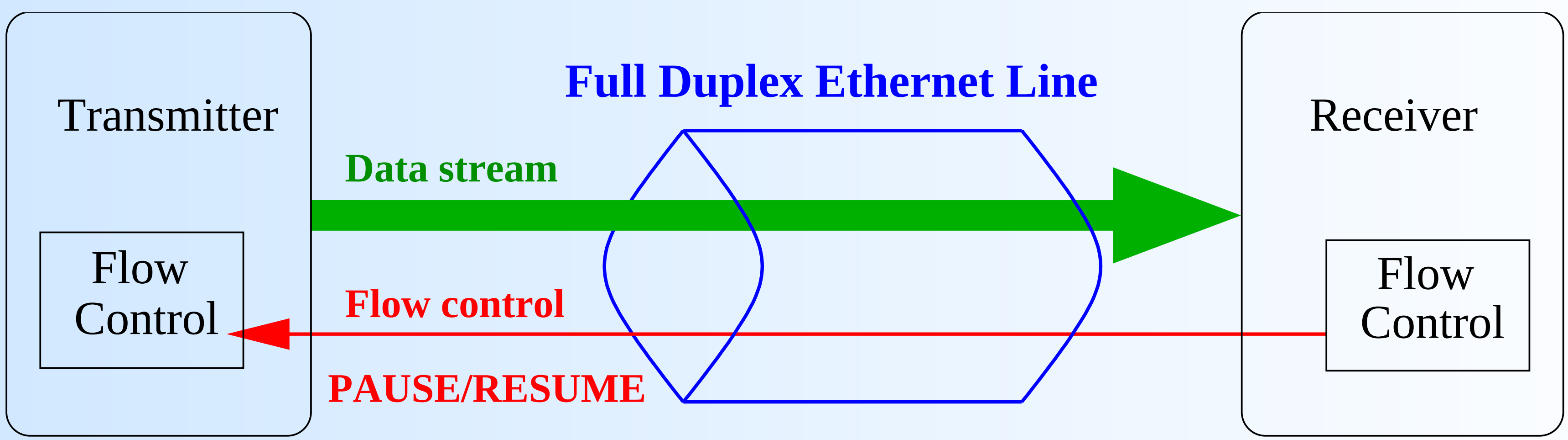




- Strict priority algorithm (priorities: *best-effort, normal, high, premium*)
- Weighted round robin algorithm (weights: *10, 20, 30, 40*)

Flow Control

- A slow receiver can limit the rate of a fast transmitter on full duplex operation using Ethernet Flow Control frames:
  - ⋆ **PAUSE** – the transmitter should stop sending
  - ⋆ **PAUSE CANCEL** – the transmitter can resume sending.



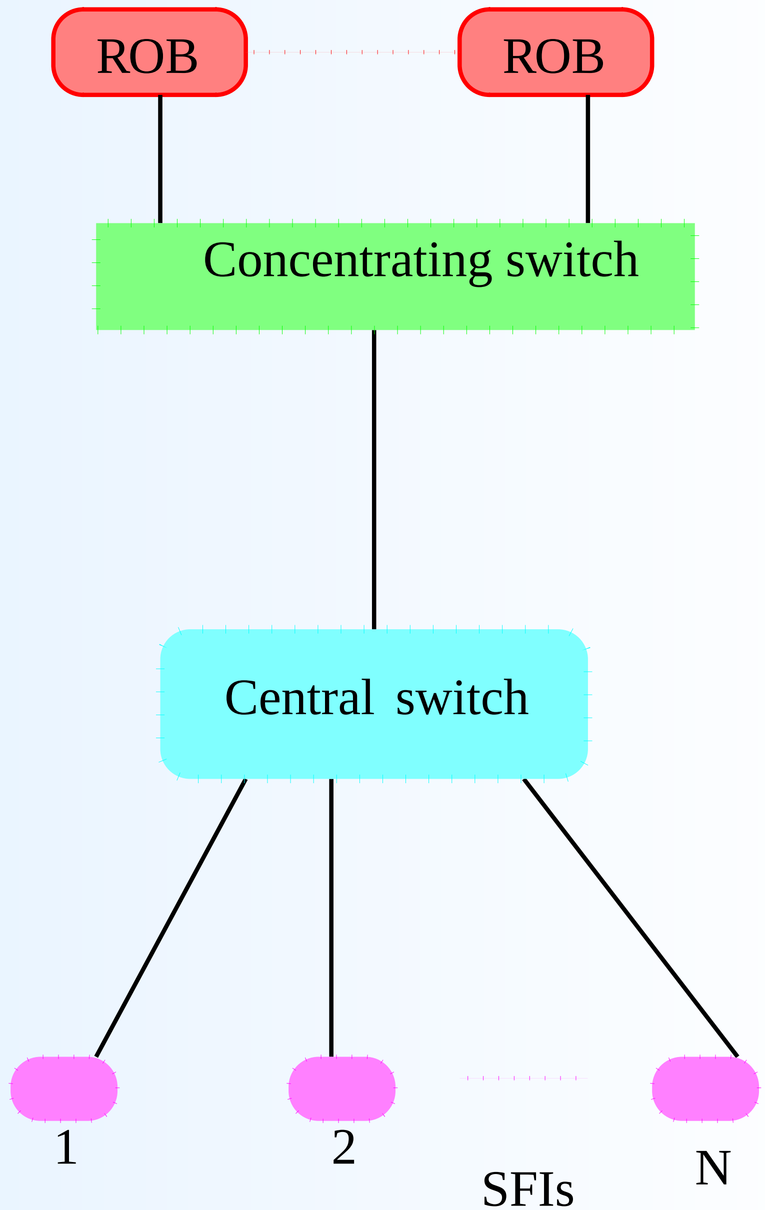
Flow Control – temporary congestion

- Packet loss implies time-outs and retries at the application level, having a great penalty on its performance. Therefore we want to minimize packet loss.

Flow Control – temporary congestion

- Packet loss implies time-outs and retries at the application level, having a great penalty on its performance. Therefore we want to minimize packet loss.

- Most of the loss is due to buffer overflow:

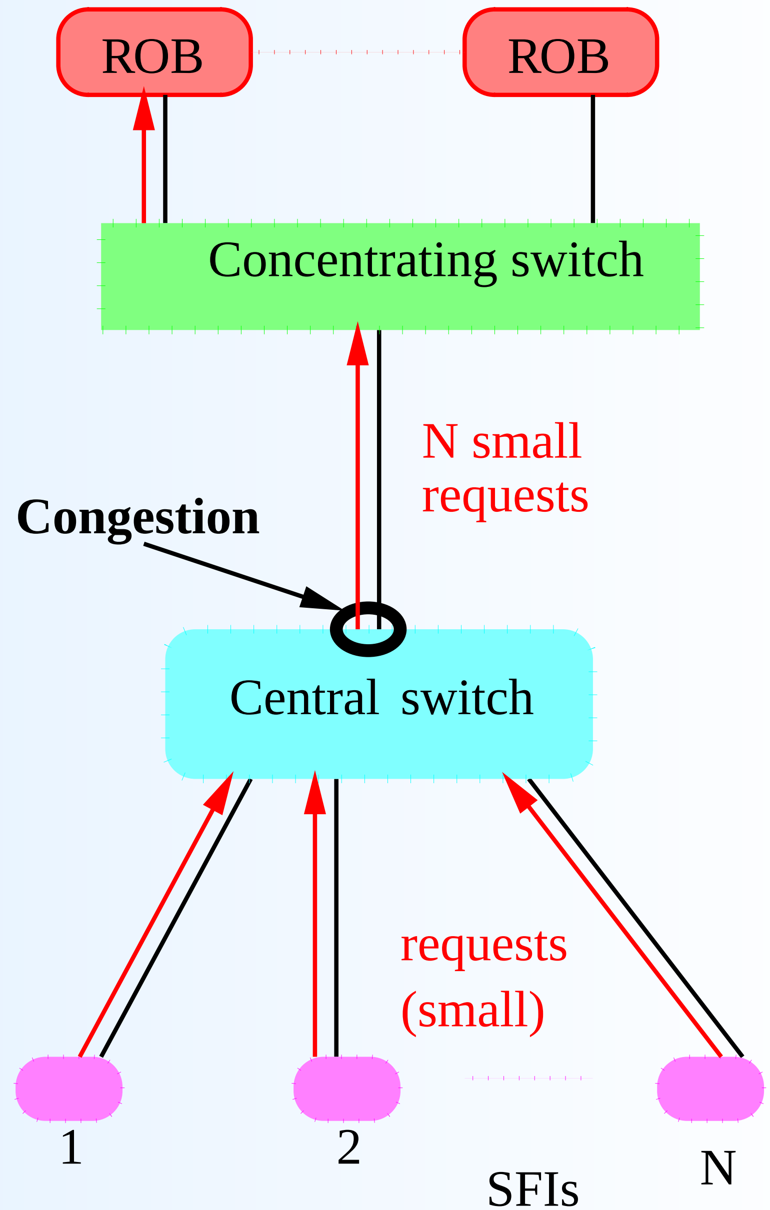


Flow Control – temporary congestion

- Packet loss implies time-outs and retries at the application level, having a great penalty on its performance. Therefore we want to minimize packet loss.

- Most of the loss is due to buffer overflow:

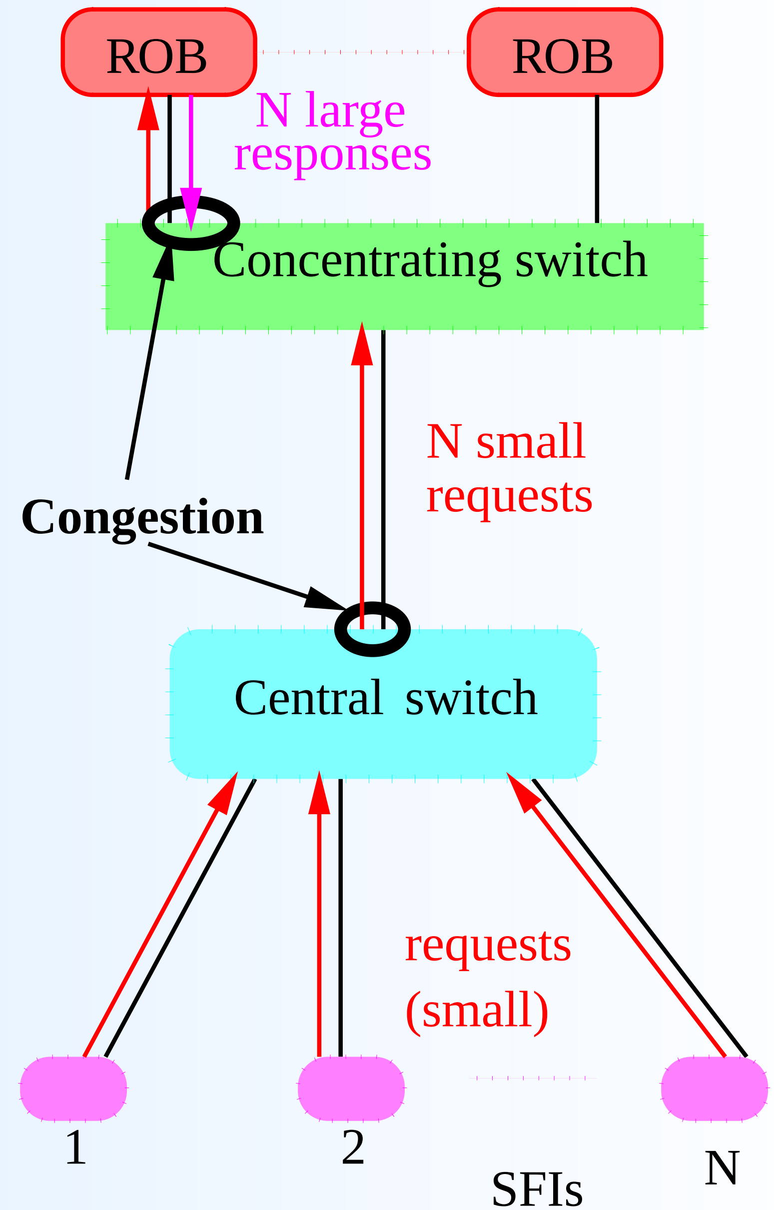
  ⋆ multiple requests funneling on one line



Flow Control – temporary congestion

- Packet loss implies time-outs and retries at the application level, having a great penalty on its performance. Therefore we want to minimize packet loss.

- Most of the loss is due to buffer overflow:
  - ⋆ multiple requests funneling on one line
  - ⋆ a ROB is asked faster than it can respond.

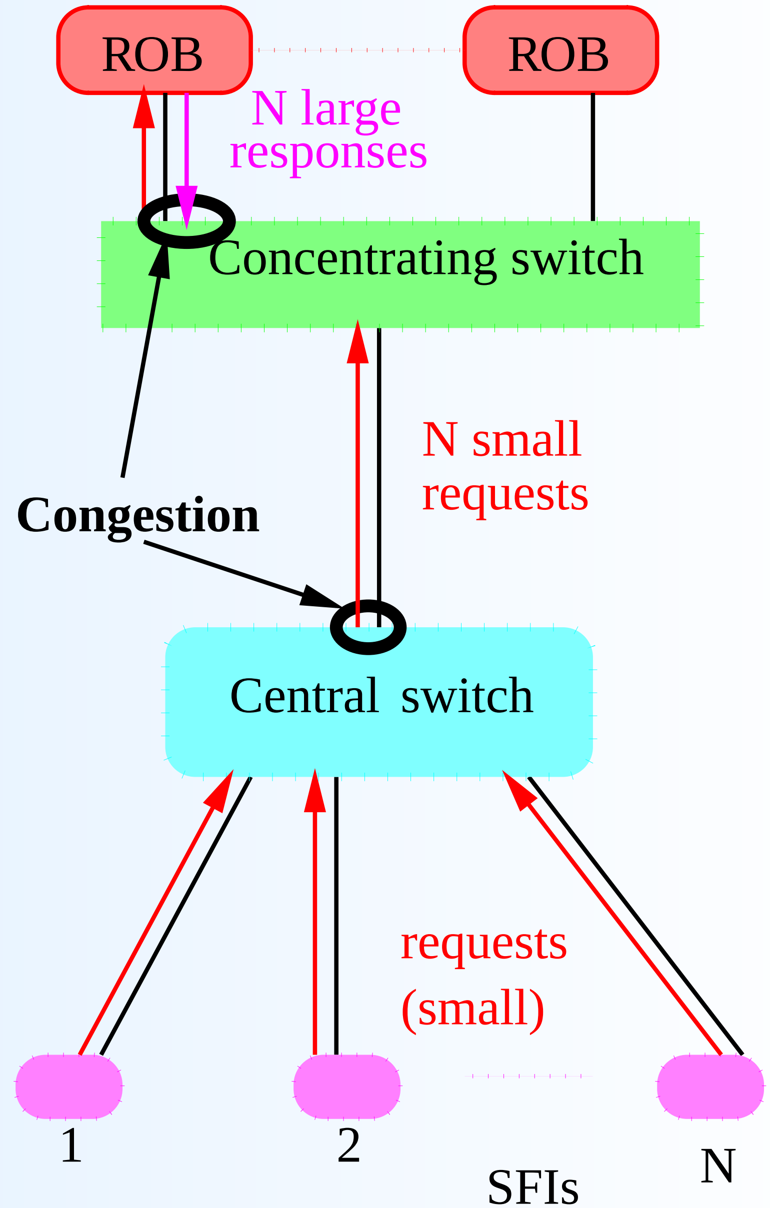


Flow Control – temporary congestion

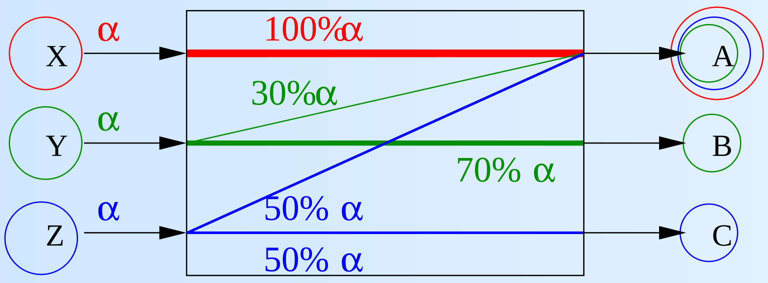
- Packet loss implies time-outs and retries at the application level, having a great penalty on its performance. Therefore we want to minimize packet loss.

- Most of the loss is due to buffer overflow:
  - ⋆ multiple requests funneling on one line
  - ⋆ a ROB is asked faster than it can respond.

- Flow Control prevents buffer overflow for short term congestions.

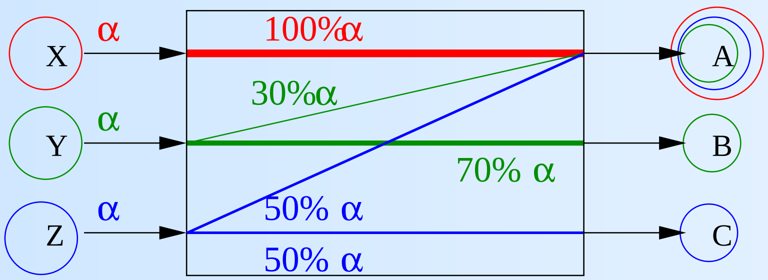


Flow Control – propagation through the switch

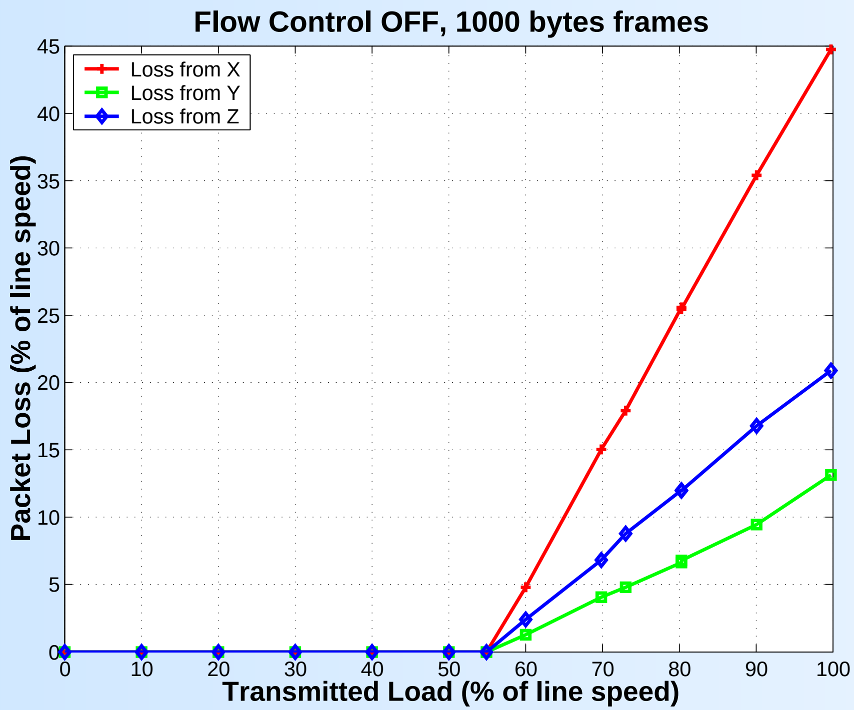


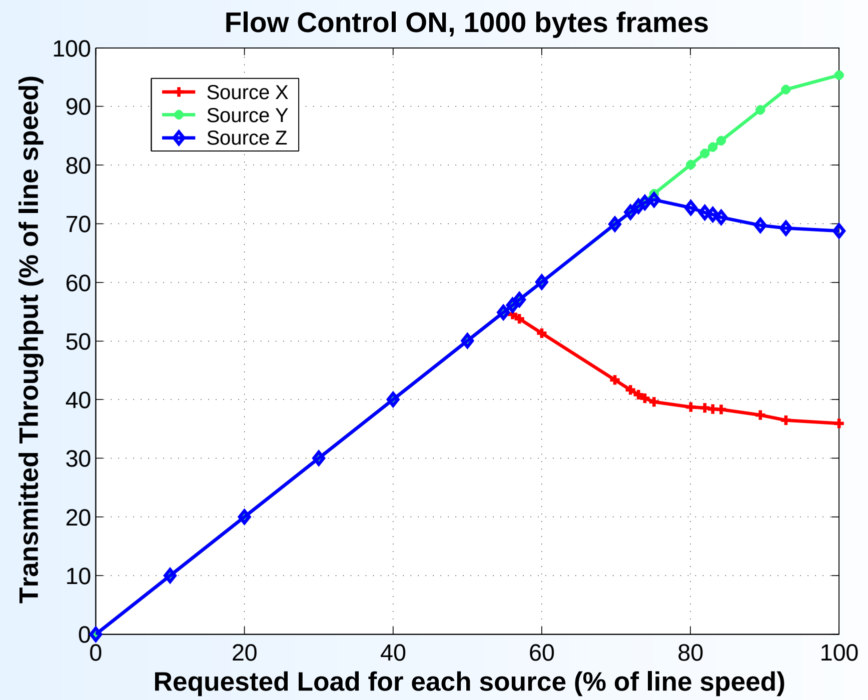
We gradually increase α until the line speed.

Flow Control – propagation through the switch



We gradually increase α until the line speed.

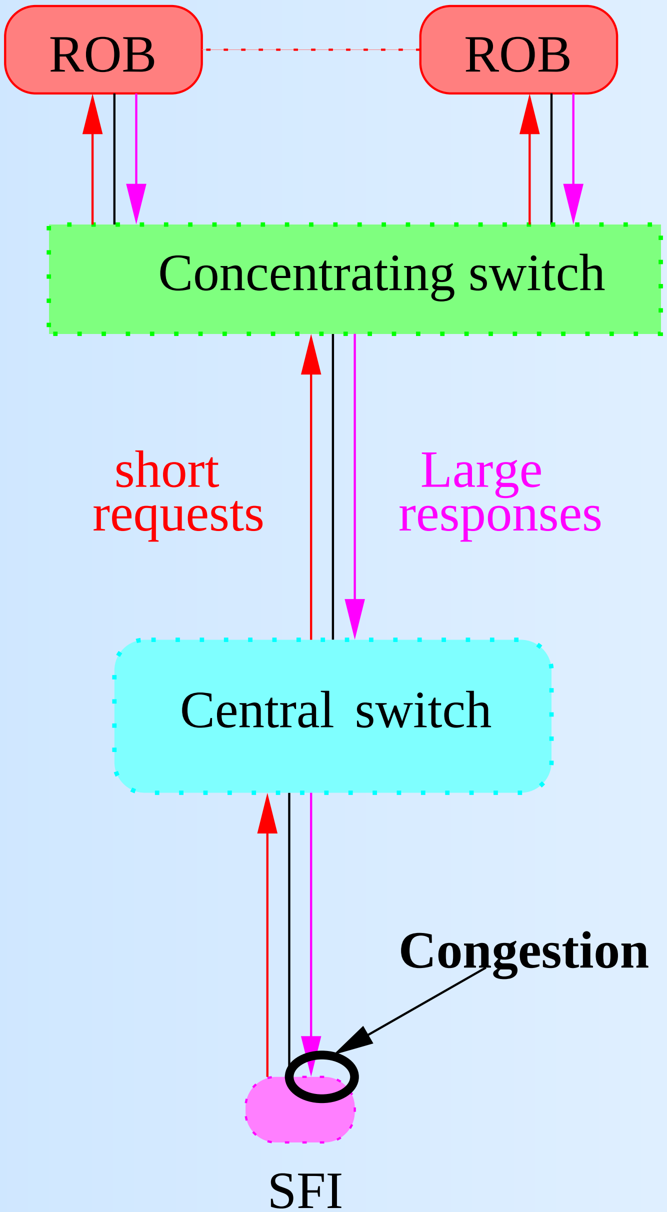




Packet loss is observed on the congested port.

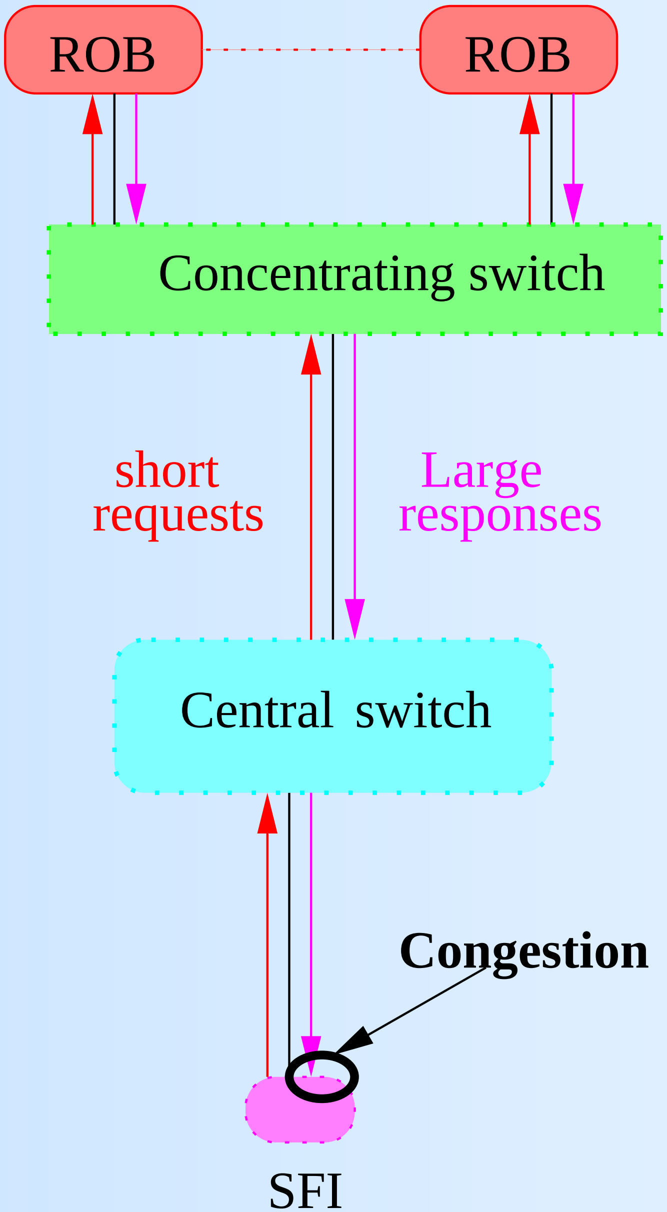
No packet loss, but the congestion spreads.

Traffic Shaping



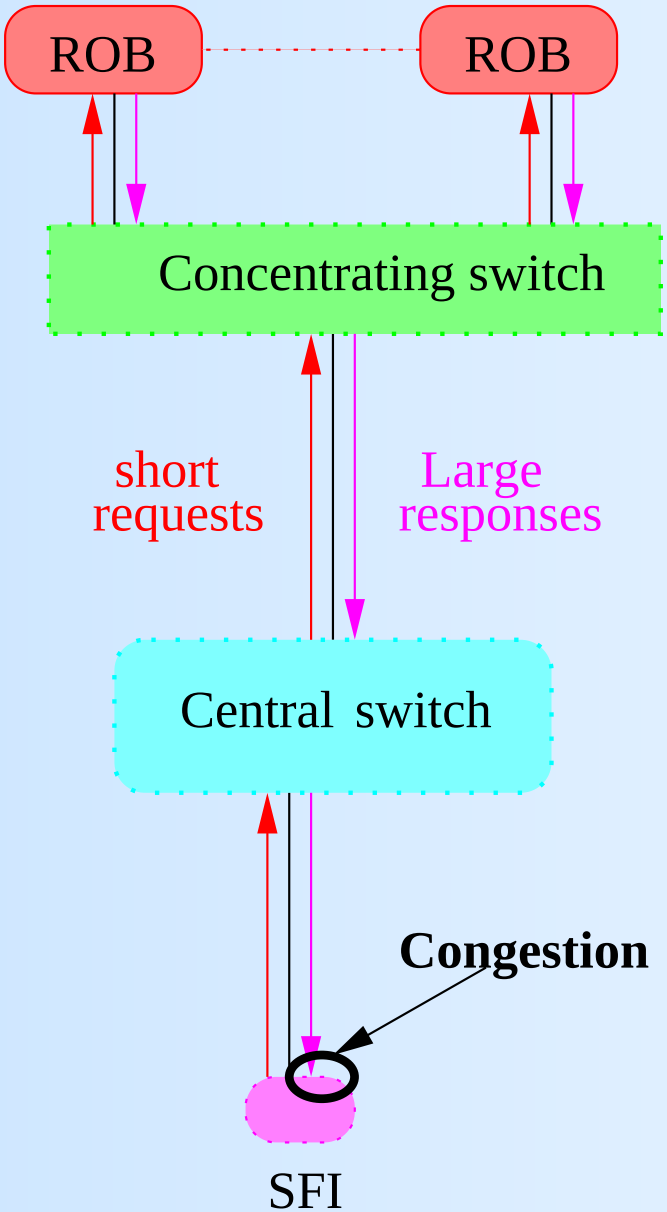
- If the application requests more than it can process congestion is created on the receiving PC.

Traffic Shaping



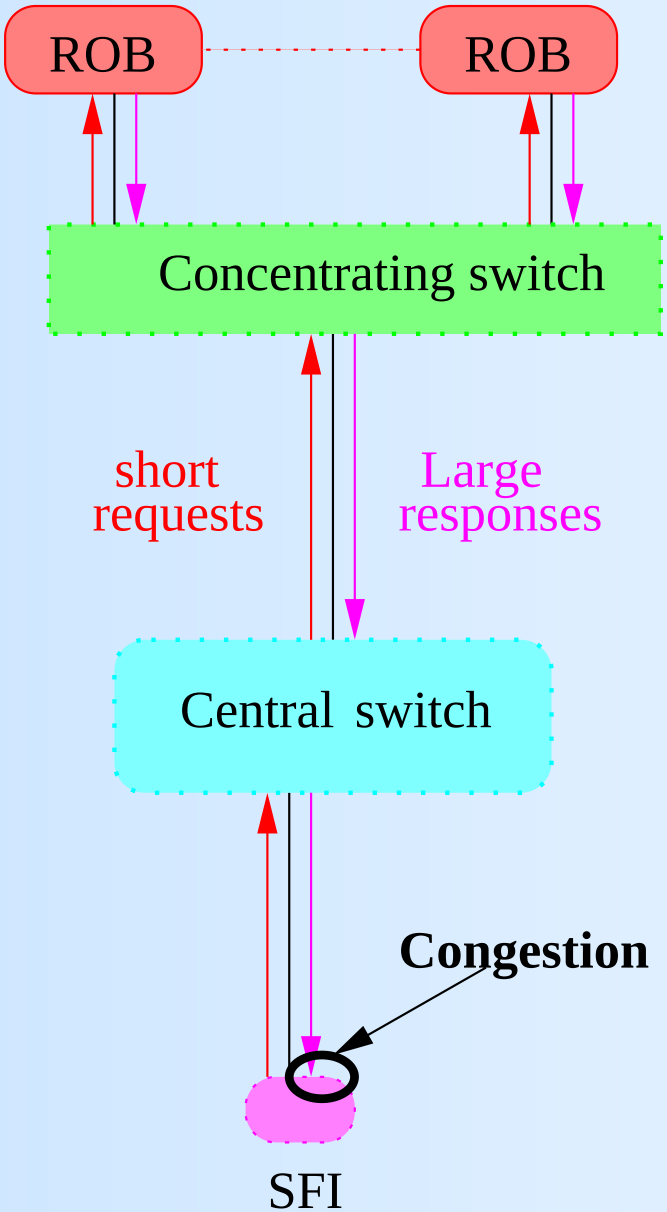
- If the application requests more than it can process congestion is created on the receiving PC.
- Flow control will be asserted if the Linux kernel or NIC cannot cope with the rate.

Traffic Shaping



- If the application requests more than it can process congestion is created on the receiving PC.
- Flow control will be asserted if the Linux kernel or NIC cannot cope with the rate.
- If the user level application cannot empty the kernel buffers, the kernel will silently drop the extra frames.

Traffic Shaping



- If the application requests more than it can process congestion is created on the receiving PC.
- Flow control will be asserted if the Linux kernel or NIC cannot cope with the rate.
- If the user level application cannot empty the kernel buffers, the kernel will silently drop the extra frames.
- Flow Control does not help for long term congestion. It is up to the application to make sure it can cope with the input rate. The application needs to implement traffic shaping (ex. limit the number of outstanding requests).

Conclusions

- The DataCollection network must obey strict constraints, as its nodes run real-time applications $\longrightarrow$ *the features presented in this talk must be verified on any switch before its integration to the network*.

Conclusions

- The DataCollection network must obey strict constraints, as its nodes run real-time applications $\longrightarrow$ *the features presented in this talk must be verified on any switch before its integration to the network*.

- Ethernet is the most suitable technology for this network:

Conclusions

- The DataCollection network must obey strict constraints, as its nodes run real-time applications $\longrightarrow$ *the features presented in this talk must be verified on any switch before its integration to the network*.

- Ethernet is the most suitable technology for this network:
  - ⋆ It satisfies the bandwidth requirements for the ATLAS DataFlow.

Conclusions

- The DataCollection network must obey strict constraints, as its nodes run real-time applications $\longrightarrow$ *the features presented in this talk must be verified on any switch before its integration to the network*.

- Ethernet is the most suitable technology for this network:
  - ⋆ It satisfies the bandwidth requirements for the ATLAS DataFlow.
  - ⋆ Segments with different speeds can be transparently interconnected via switches: 100 Mbps, 1 Gbps and 10 Gbps.

Conclusions

- The DataCollection network must obey strict constraints, as its nodes run real-time applications $\longrightarrow$ *the features presented in this talk must be verified on any switch before its integration to the network*.

- Ethernet is the most suitable technology for this network:
  - ⋆ It satisfies the bandwidth requirements for the ATLAS DataFlow.
  - ⋆ Segments with different speeds can be transparently interconnected via switches: 100 Mbps, 1 Gbps and 10 Gbps.
  - ⋆ It is multi-vendor technology with long term support.

Conclusions

- The DataCollection network must obey strict constraints, as its nodes run real-time applications $\longrightarrow$ *the features presented in this talk must be verified on any switch before its integration to the network*.

- Ethernet is the most suitable technology for this network:
  - ⋆ It satisfies the bandwidth requirements for the ATLAS DataFlow.
  - ⋆ Segments with different speeds can be transparently interconnected via switches: 100 Mbps, 1 Gbps and 10 Gbps.
  - ⋆ It is multi-vendor technology with long term support.
  - ⋆ Ethernet is commodity:
    - ∗ GE UTP NIC (network interface card) price $\approx$ 60 USD.
    - ∗ The GE switch port cost is affordable , and it's still getting cheaper.

Conclusions

- The DataCollection network must obey strict constraints, as its nodes run real-time applications $\longrightarrow$ *the features presented in this talk must be verified on any switch before its integration to the network*.

- Ethernet is the most suitable technology for this network:
  - ⋆ It satisfies the bandwidth requirements for the ATLAS DataFlow.
  - ⋆ Segments with different speeds can be transparently interconnected via switches: 100 Mbps, 1 Gbps and 10 Gbps.
  - ⋆ It is multi-vendor technology with long term support.
  - ⋆ Ethernet is commodity:
    - ∗ GE UTP NIC (network interface card) price $\approx$ 60 USD.
    - ∗ The GE switch port cost is affordable , and it's still getting cheaper.
  - ⋆ PCs have become fast enough to cope with the GE line speed (request-response traffic on a Dual P4, 2.4 GHz $\longrightarrow$ 70 Mbps incoming traffic).

Conclusions

- The DataCollection network must obey strict constraints, as its nodes run real-time applications $\longrightarrow$ *the features presented in this talk must be verified on any switch before its integration to the network*.

- Ethernet is the most suitable technology for this network:
  - ⋆ It satisfies the bandwidth requirements for the ATLAS DataFlow.
  - ⋆ Segments with different speeds can be transparently interconnected via switches: 100 Mbps, 1 Gbps and 10 Gbps.
  - ⋆ It is multi-vendor technology with long term support.
  - ⋆ Ethernet is commodity:
    - ∗ GE UTP NIC (network interface card) price $\approx$ 60 USD.
    - ∗ The GE switch port cost is affordable , and it's still getting cheaper.
  - ⋆ PCs have become fast enough to cope with the GE line speed (request-response traffic on a Dual P4, 2.4 GHz $\longrightarrow$ 70 Mbps incoming traffic).
  - ⋆ It has an evolutionary upgrade path to high speed.

Conclusions

- The DataCollection network must obey strict constraints, as its nodes run real-time applications → *the features presented in this talk must be verified on any switch before its integration to the network*.

- Ethernet is the most suitable technology for this network:
  - ⋆ It satisfies the bandwidth requirements for the ATLAS DataFlow.
  - ⋆ Segments with different speeds can be transparently interconnected via switches: 100 Mbps, 1 Gbps and 10 Gbps.
  - ⋆ It is multi-vendor technology with long term support.
  - ⋆ Ethernet is commodity:
    - ∗ GE UTP NIC (network interface card) price ≈ 60 USD.
    - ∗ The GE switch port cost is affordable , and it's still getting cheaper.
  - ⋆ PCs have become fast enough to cope with the GE line speed (request-response traffic on a Dual P4, 2.4 GHz → 70 Mbps incoming traffic).
  - ⋆ It has an evolutionary upgrade path to high speed.
  - ⋆ 10 GE is a candidate for the central switches if the price drops.

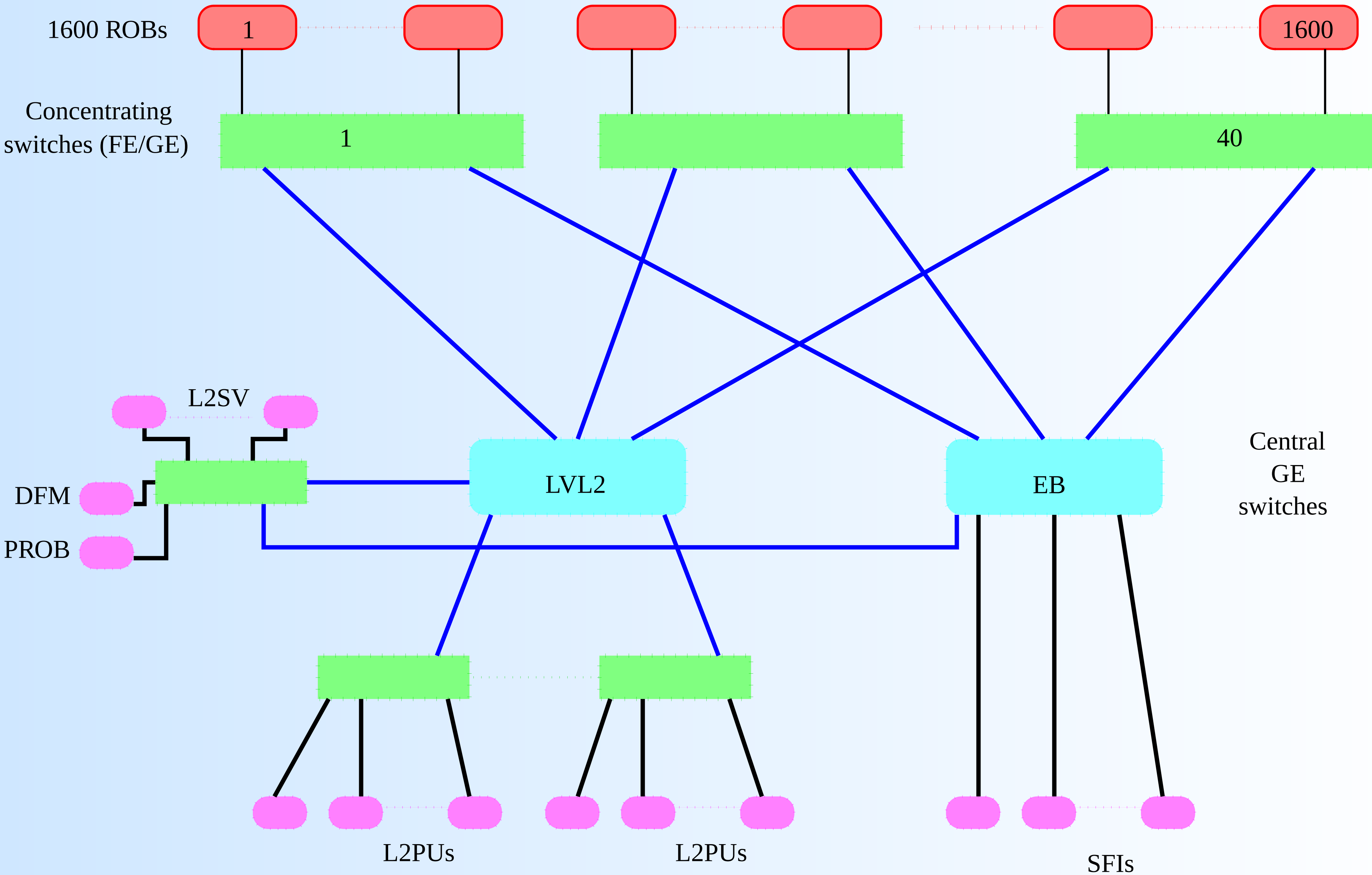
References

[1] F. Barnes, R. Beuran, R. Dobinson, M. J. LeVine, B. Martin, J. Lokier, and C. Meirosu. Testing ethernet networks for the atlas data collection system. In *IEEE Trans. on Nuclear Science*, volume 49, No. 2, page 516. IEEE Nuclear and Plasma Sciences Society, April 2002.

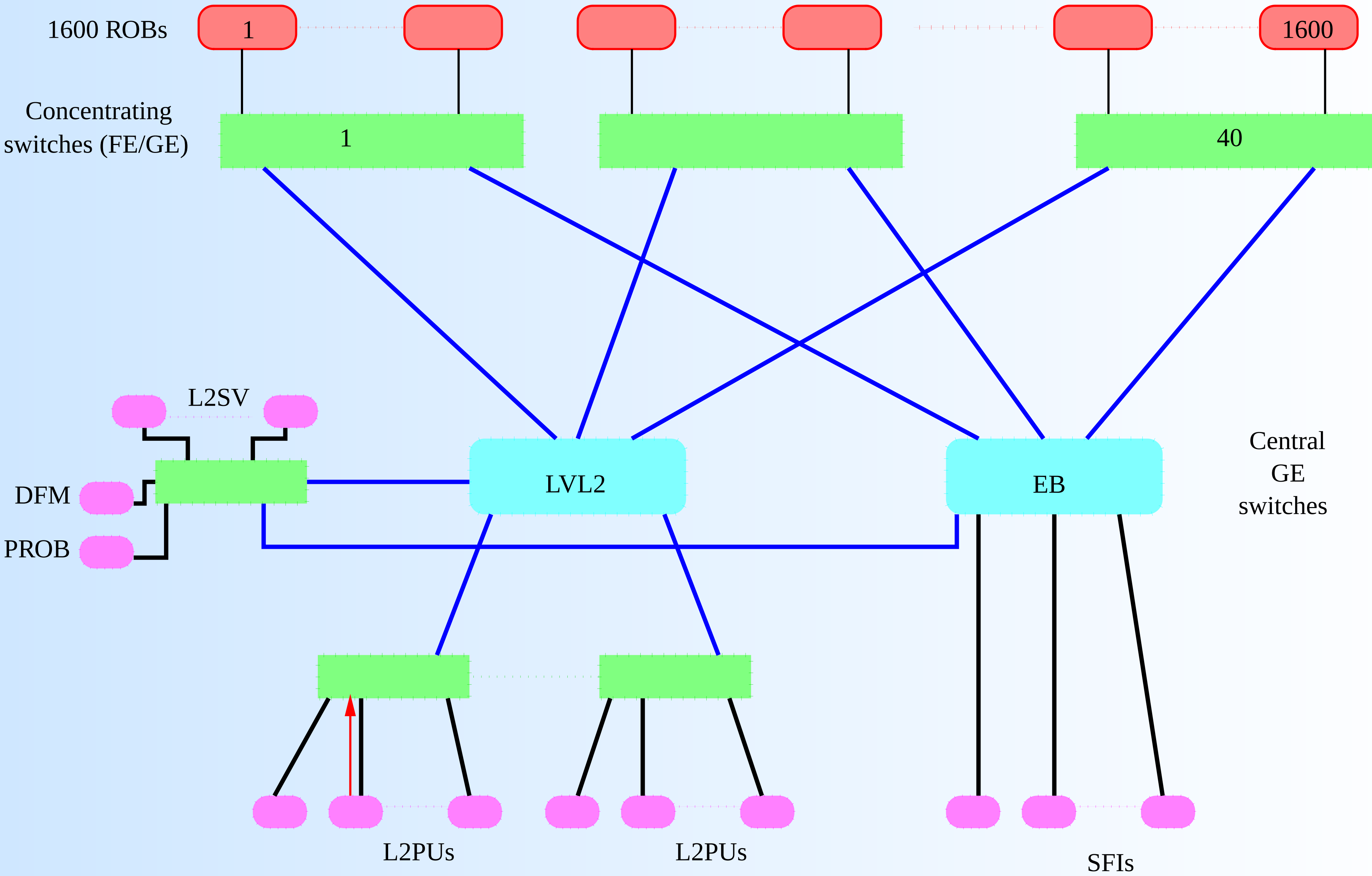
[2] H. Beck, B. Dobinson, K. Korcyl, and M. Levine. Atlas tdaq: A network-based architecture. *DC note 059*, February 27 2003. http://atlas.web.cern.ch/Atlas/GROUPS/DAQTRIG/DataFlow/DataCollection/docs/DC-059/DC-059.pdf.

[3] H. Beck and F. Wickens. High-level description of the flow of control and data messages for the atlas tdaq integrated prototypes. *Data Collection note 012*, June 19 2002. http://atlas.web.cern.ch/Atlas/GROUPS/DAQTRIG/DataFlow/DataCollection/docs/DC-012/DC-012.pdf.

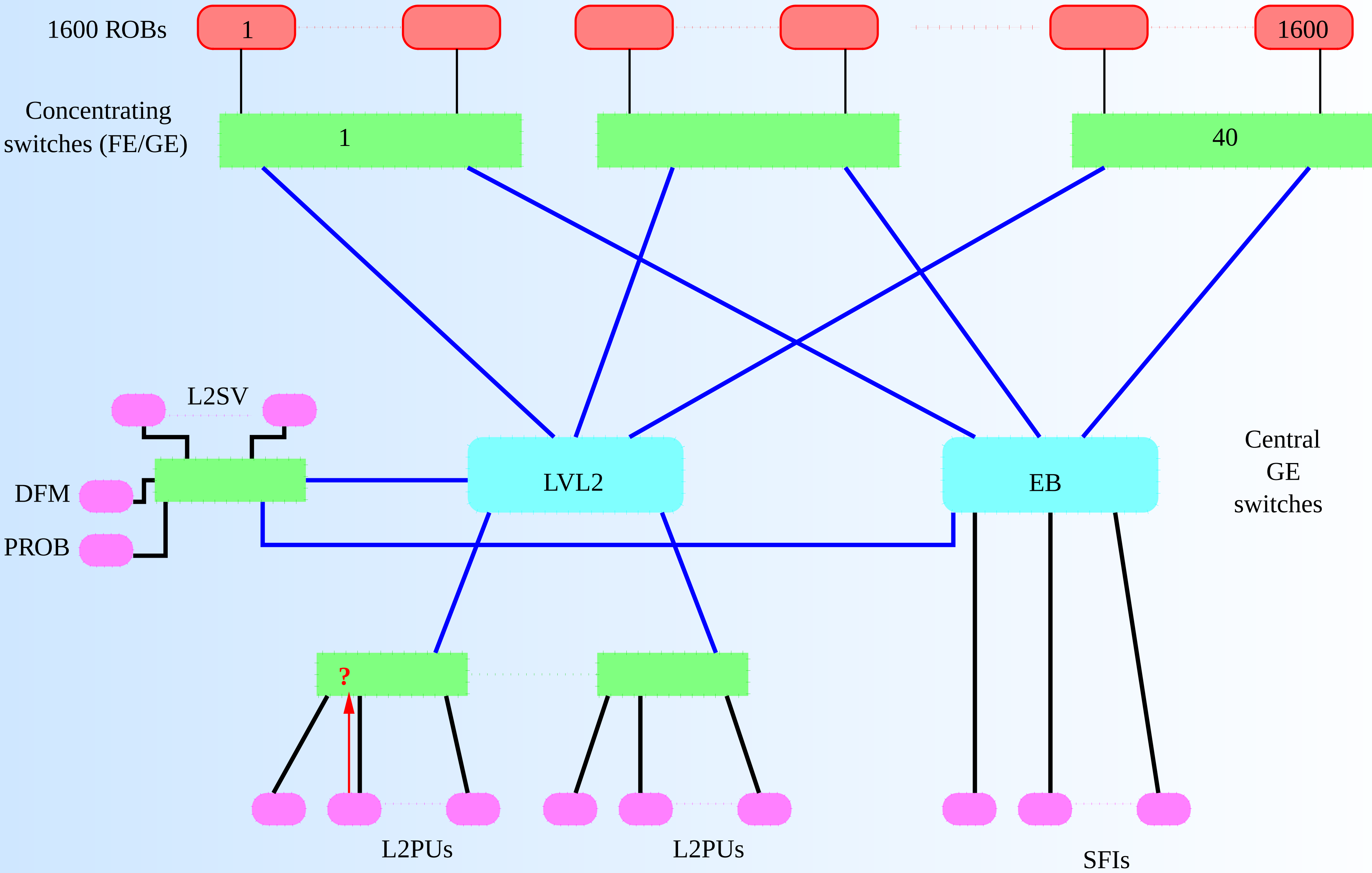
MAC address table size.



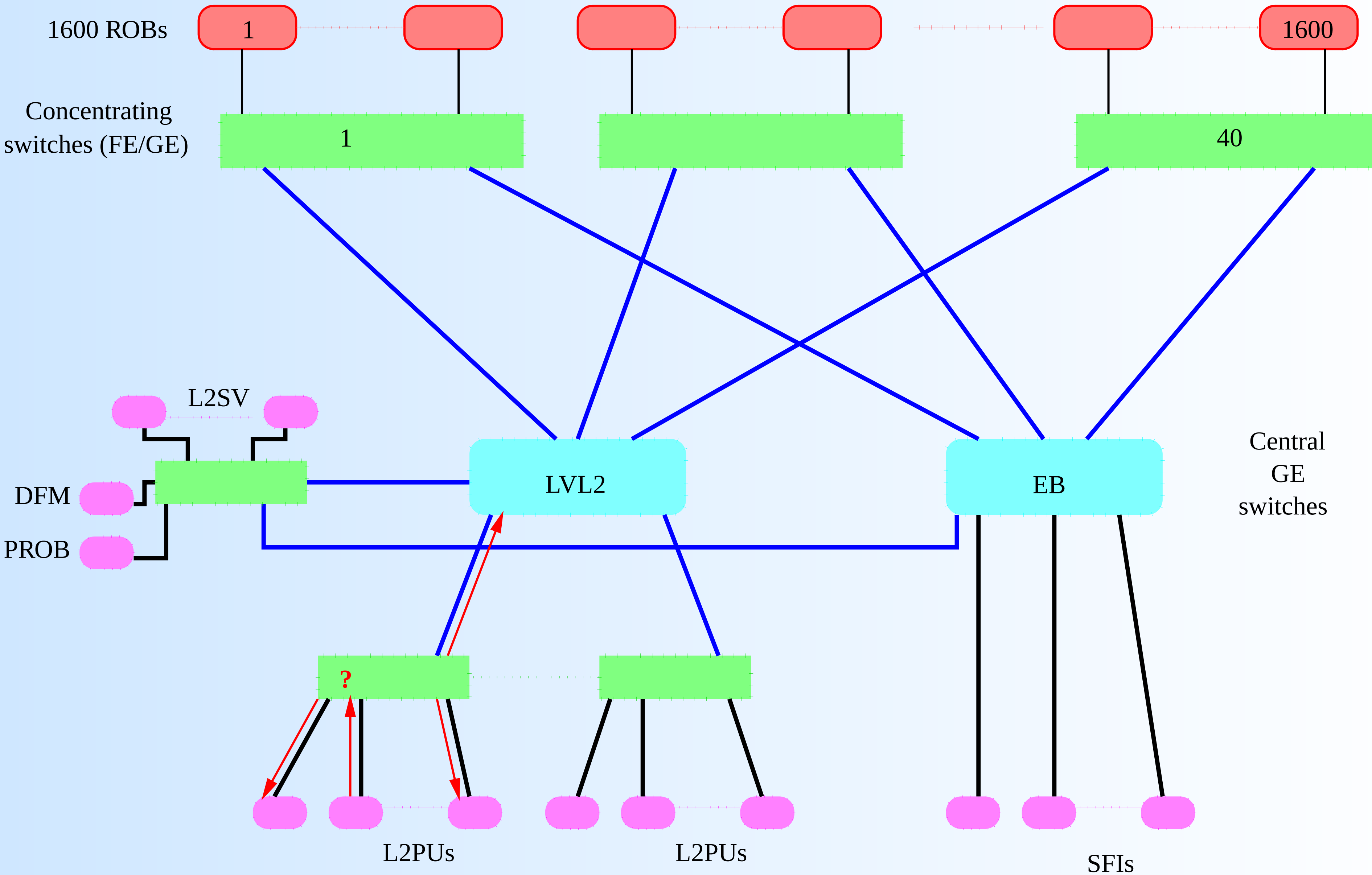
MAC address table size.



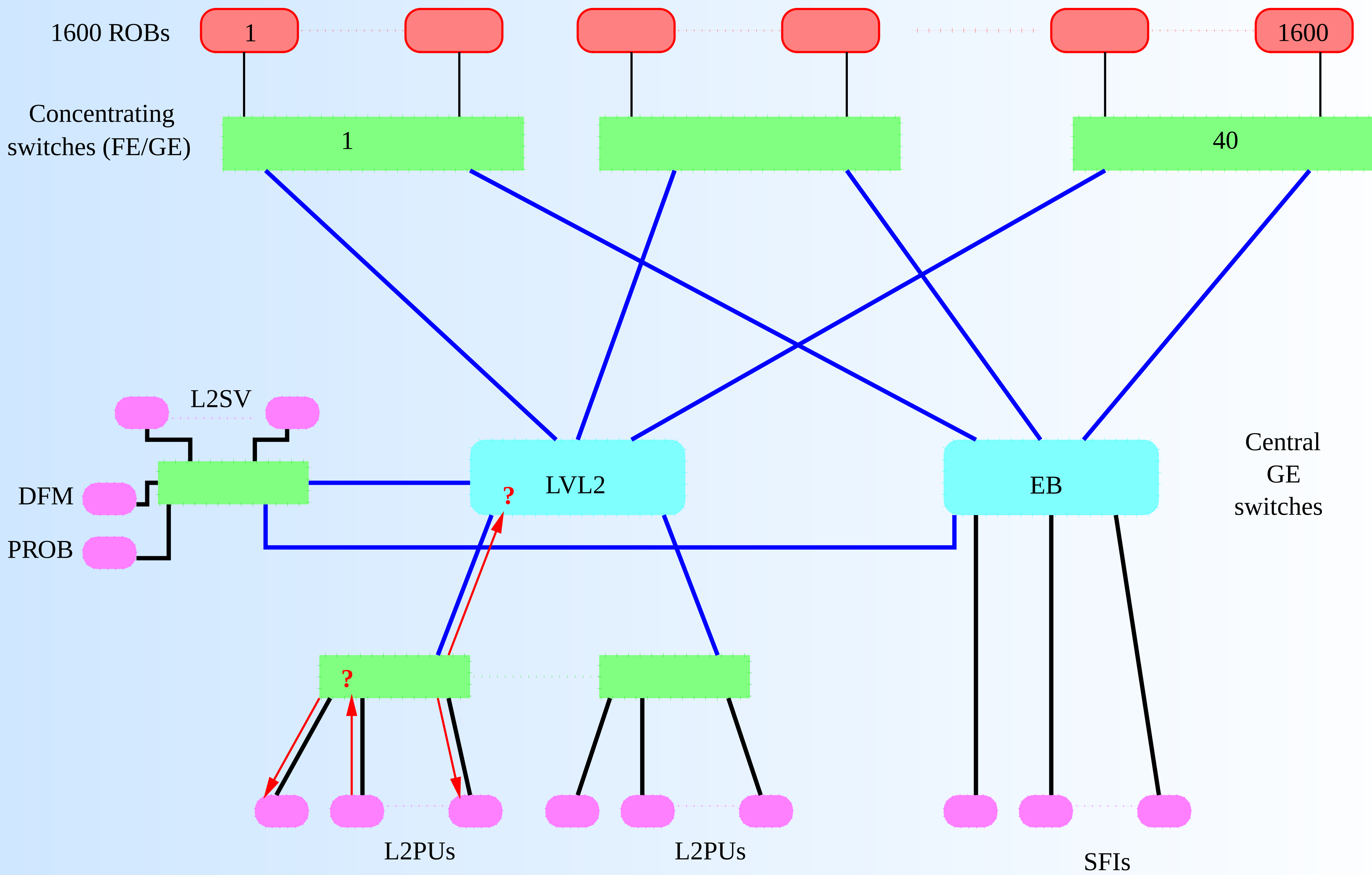
MAC address table size.



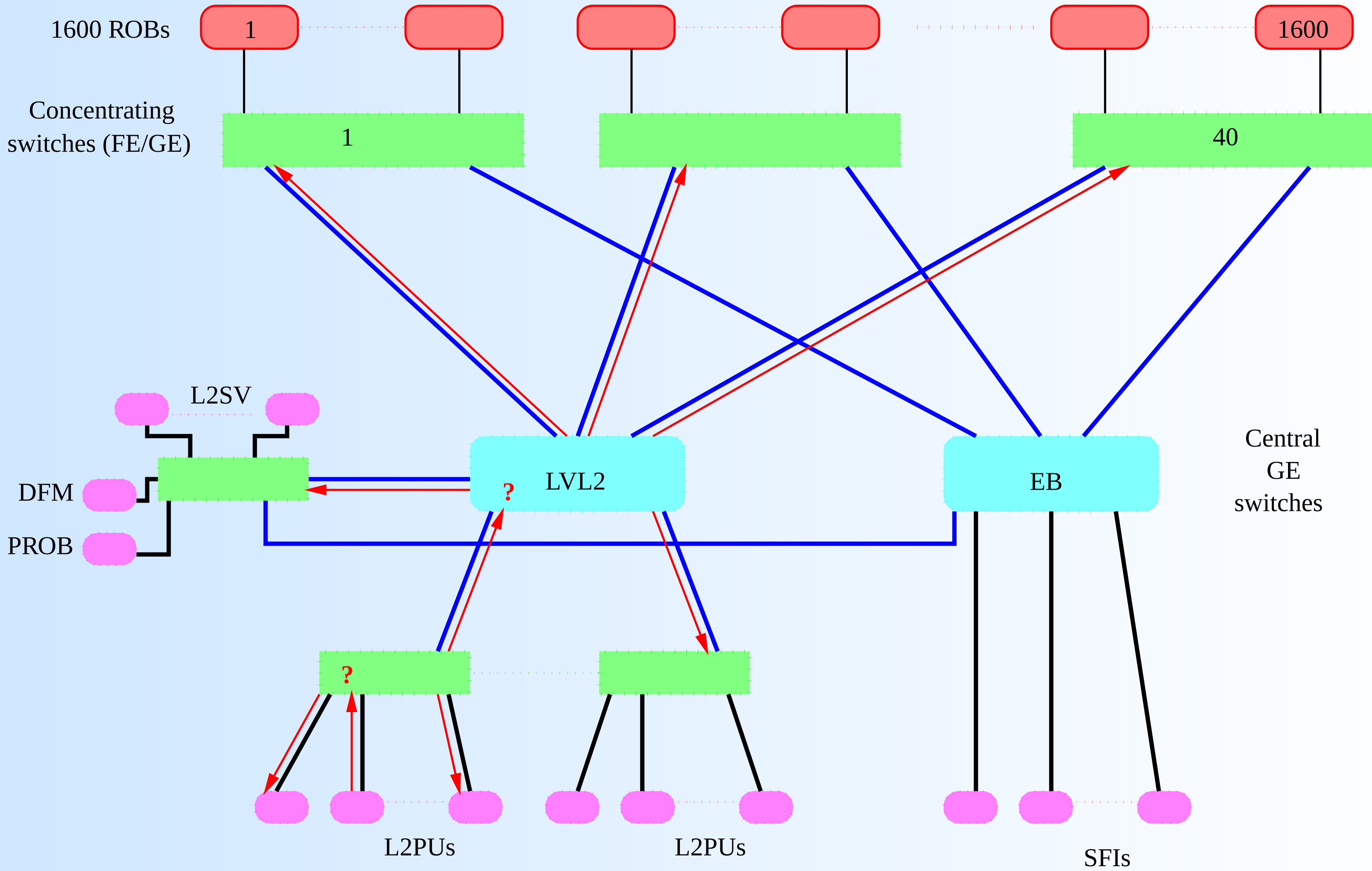
MAC address table size.



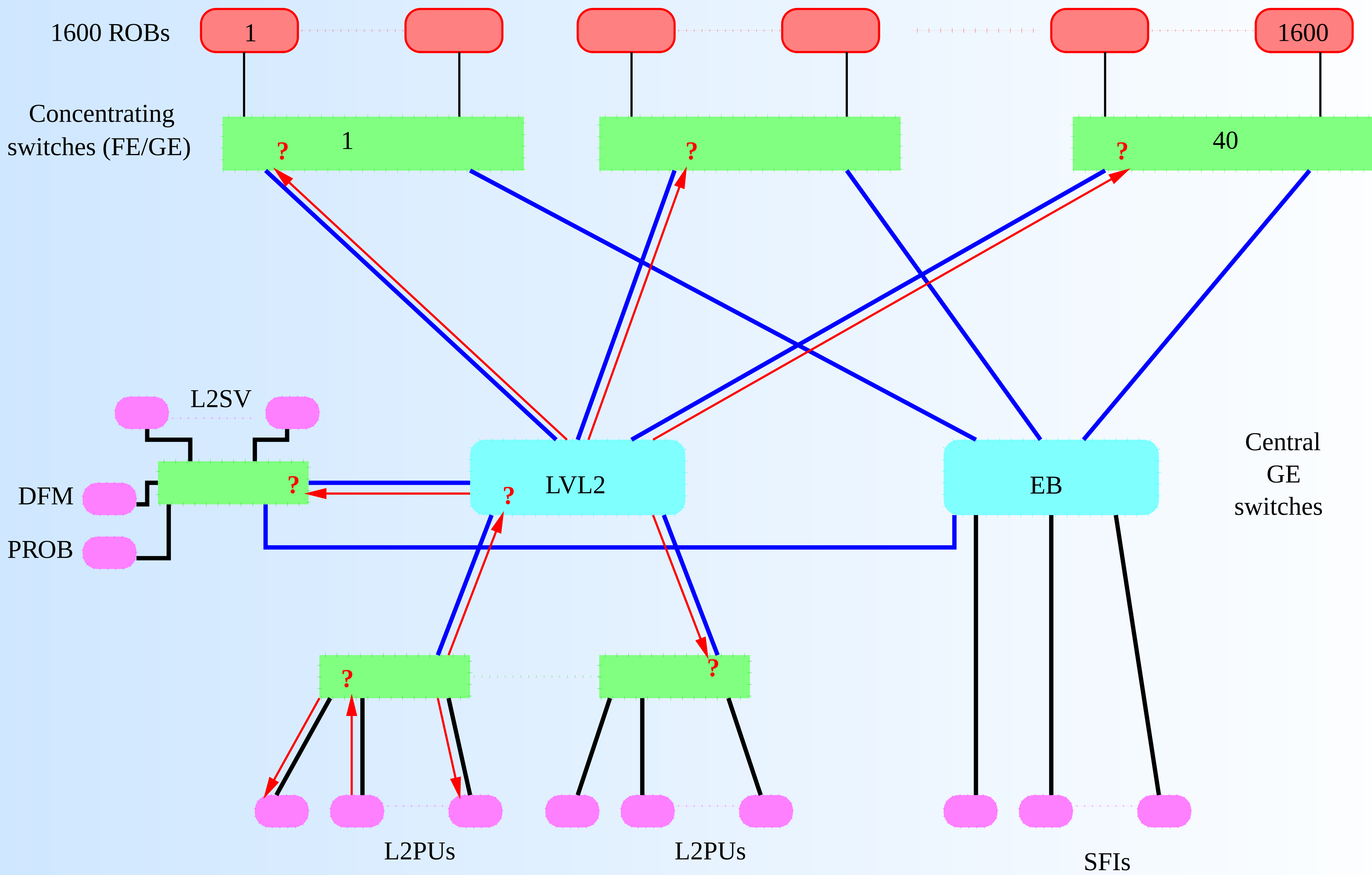
MAC address table size.



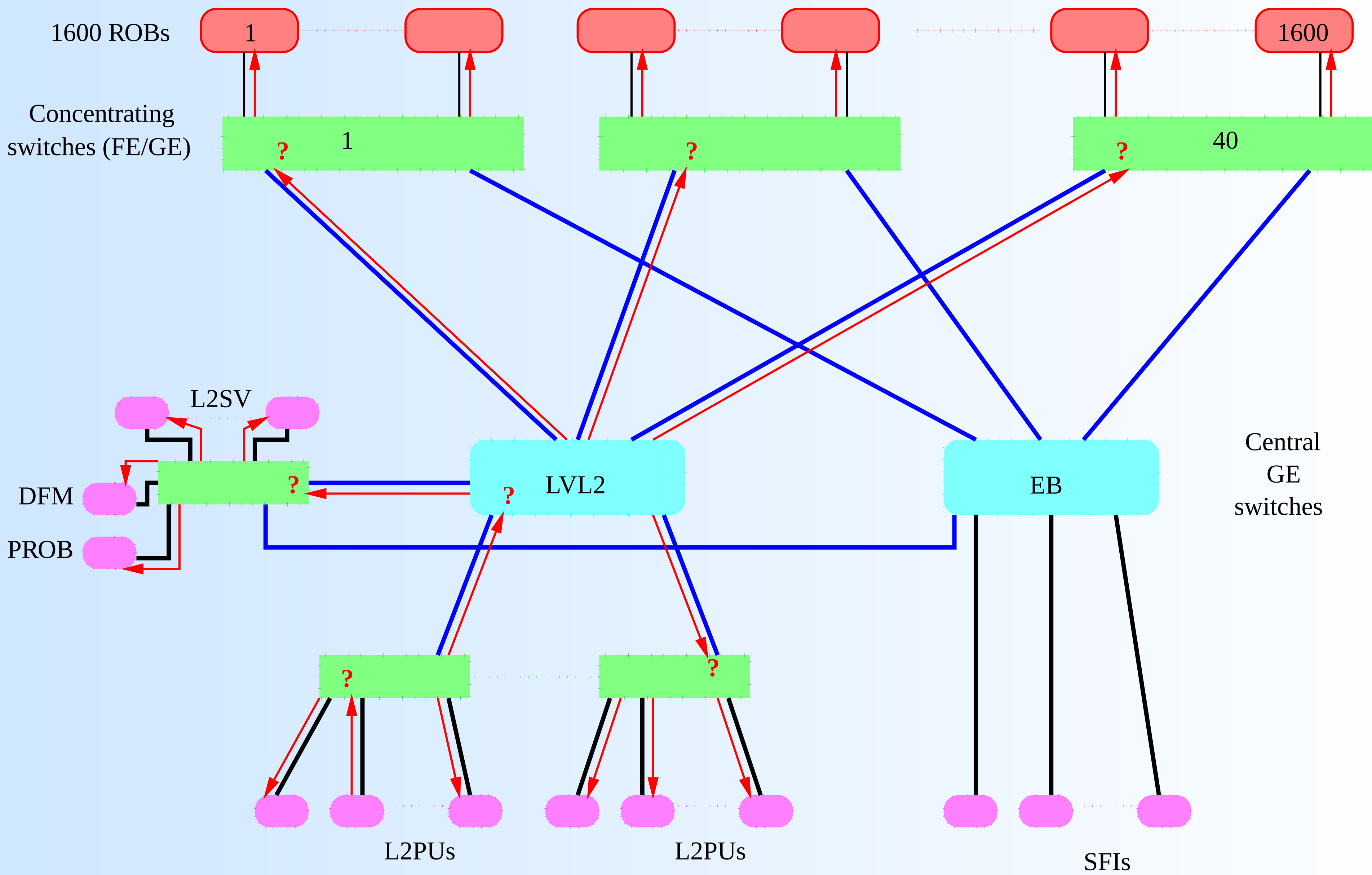
MAC address table size.



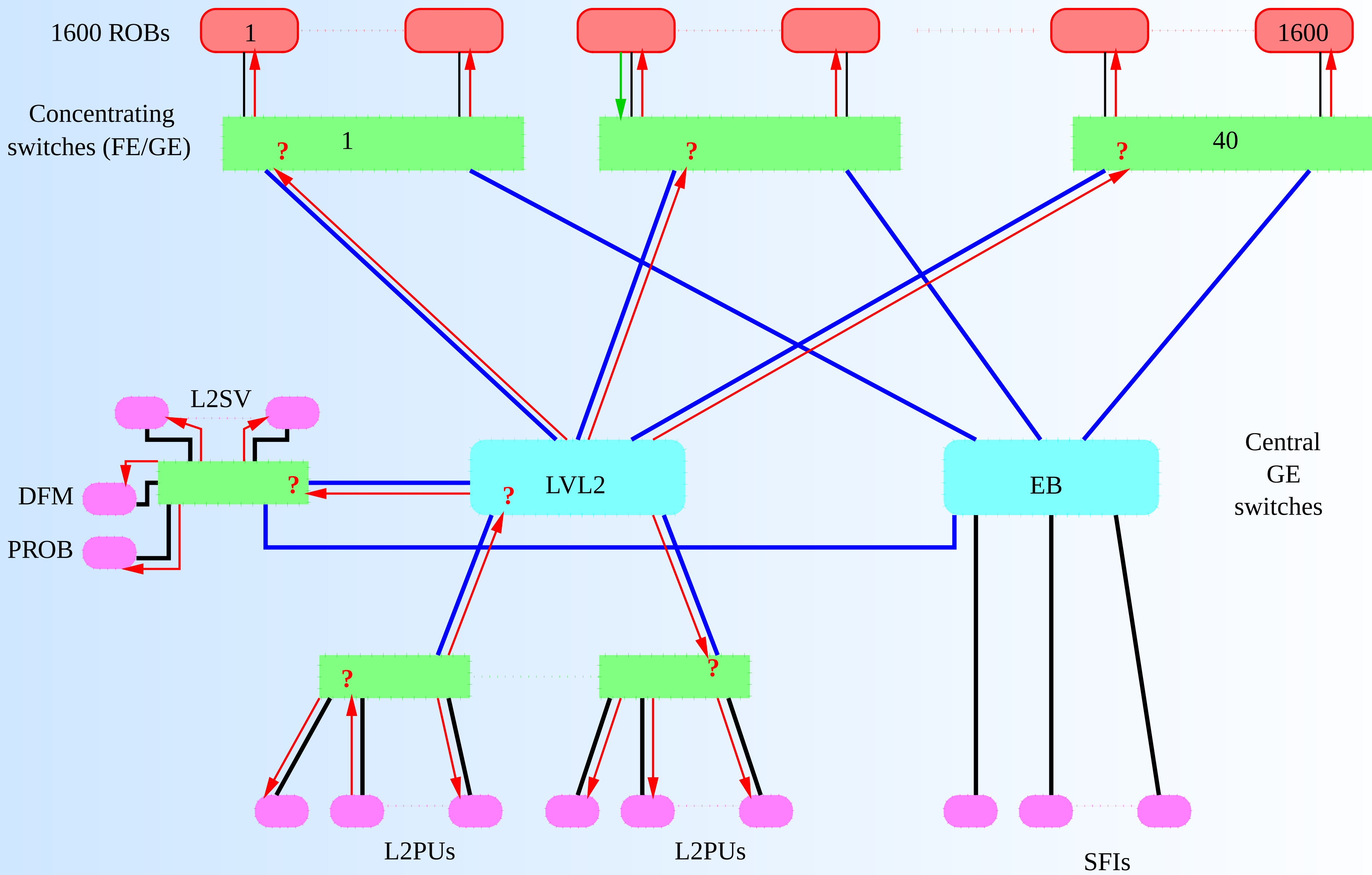
MAC address table size.



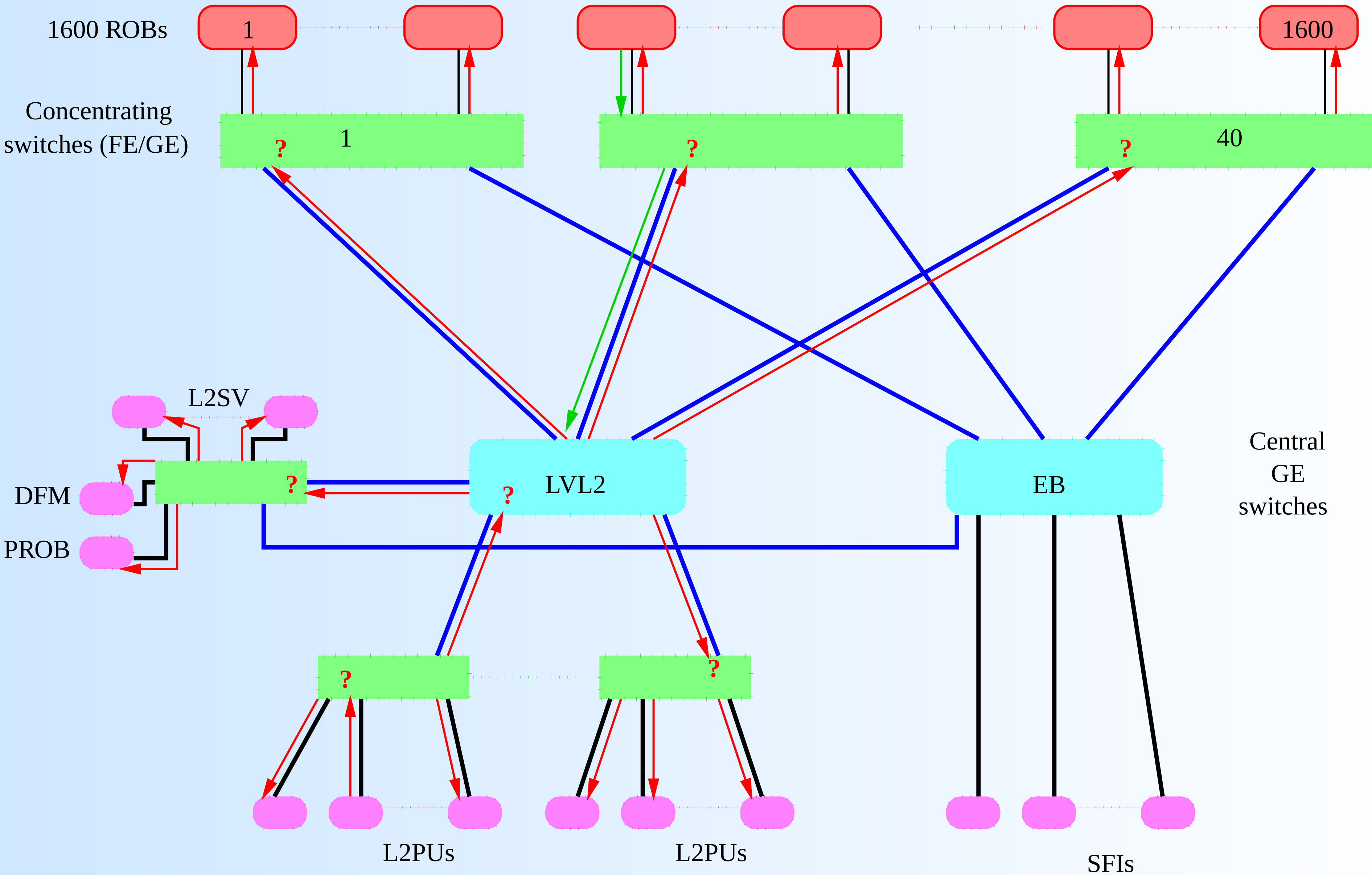
MAC address table size.



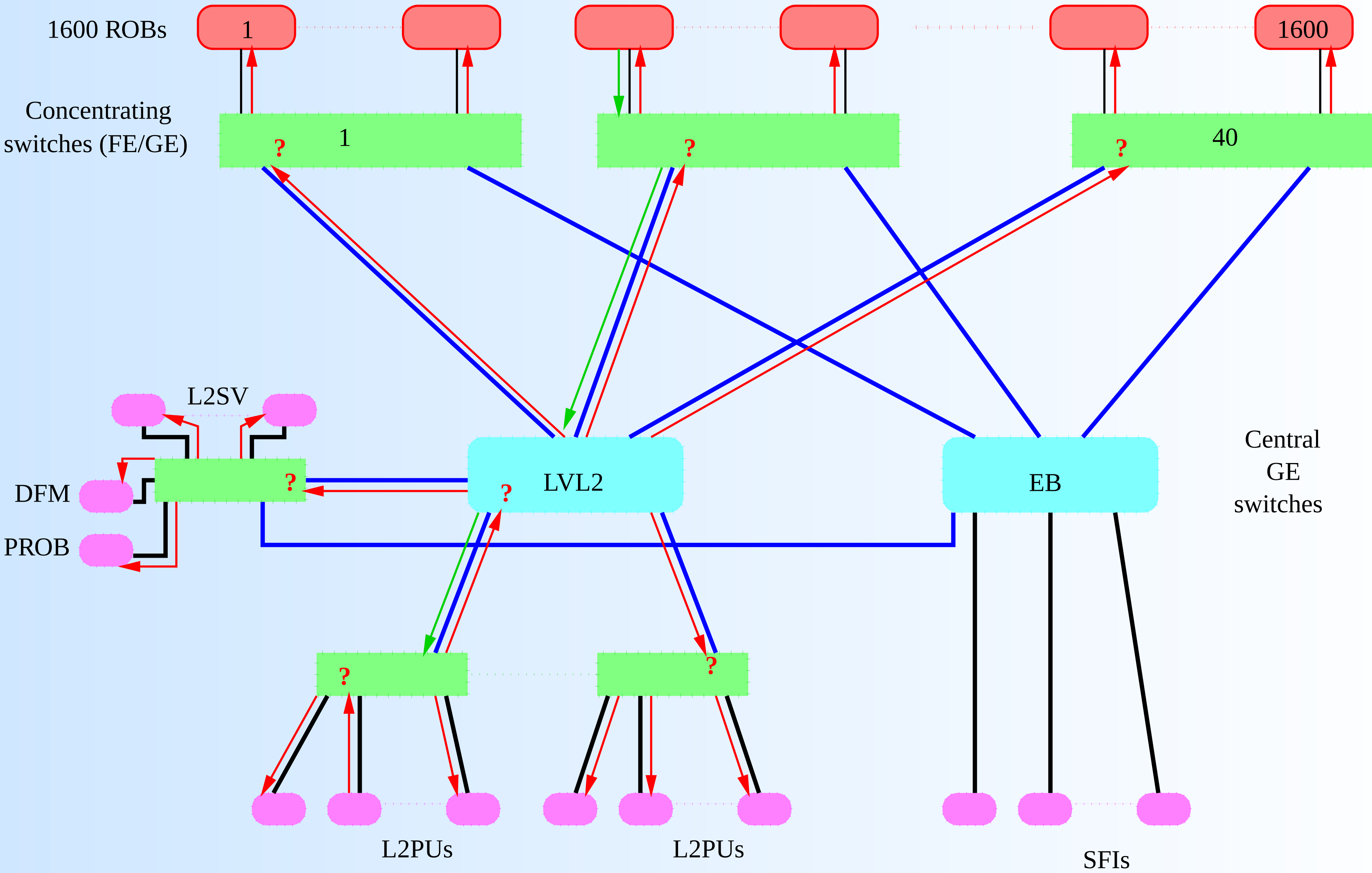
MAC address table size.



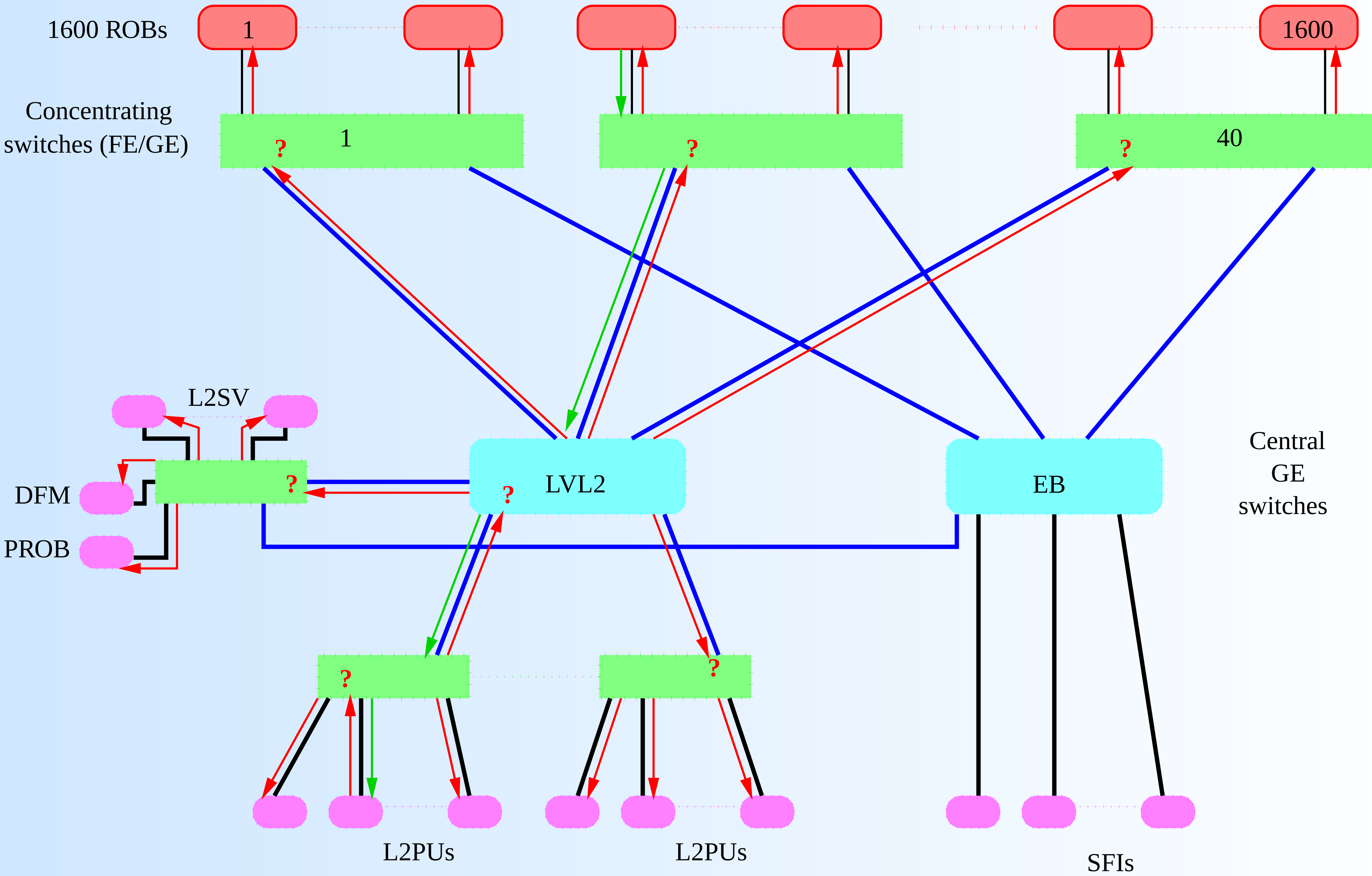
MAC address table size.



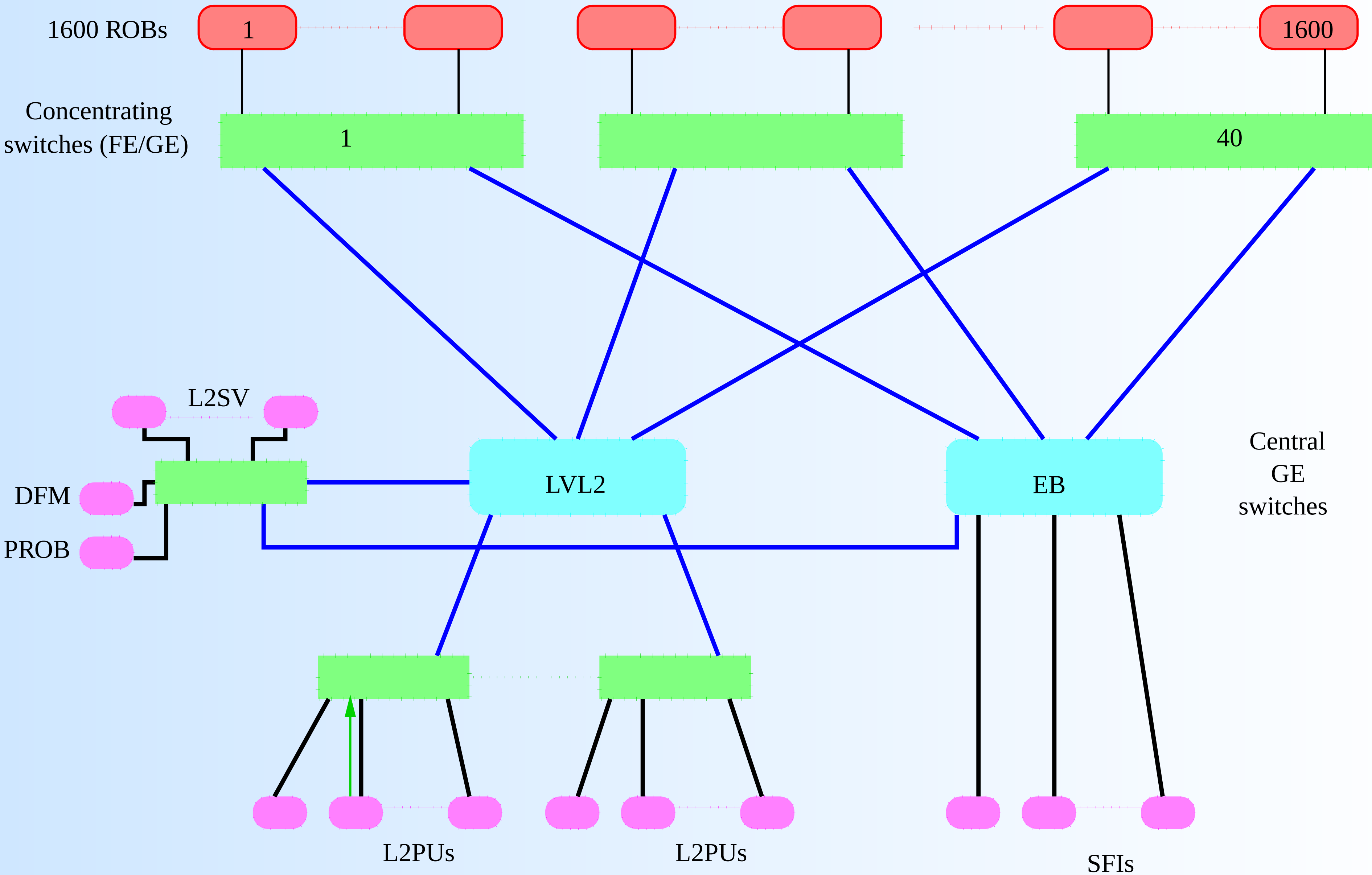
MAC address table size.



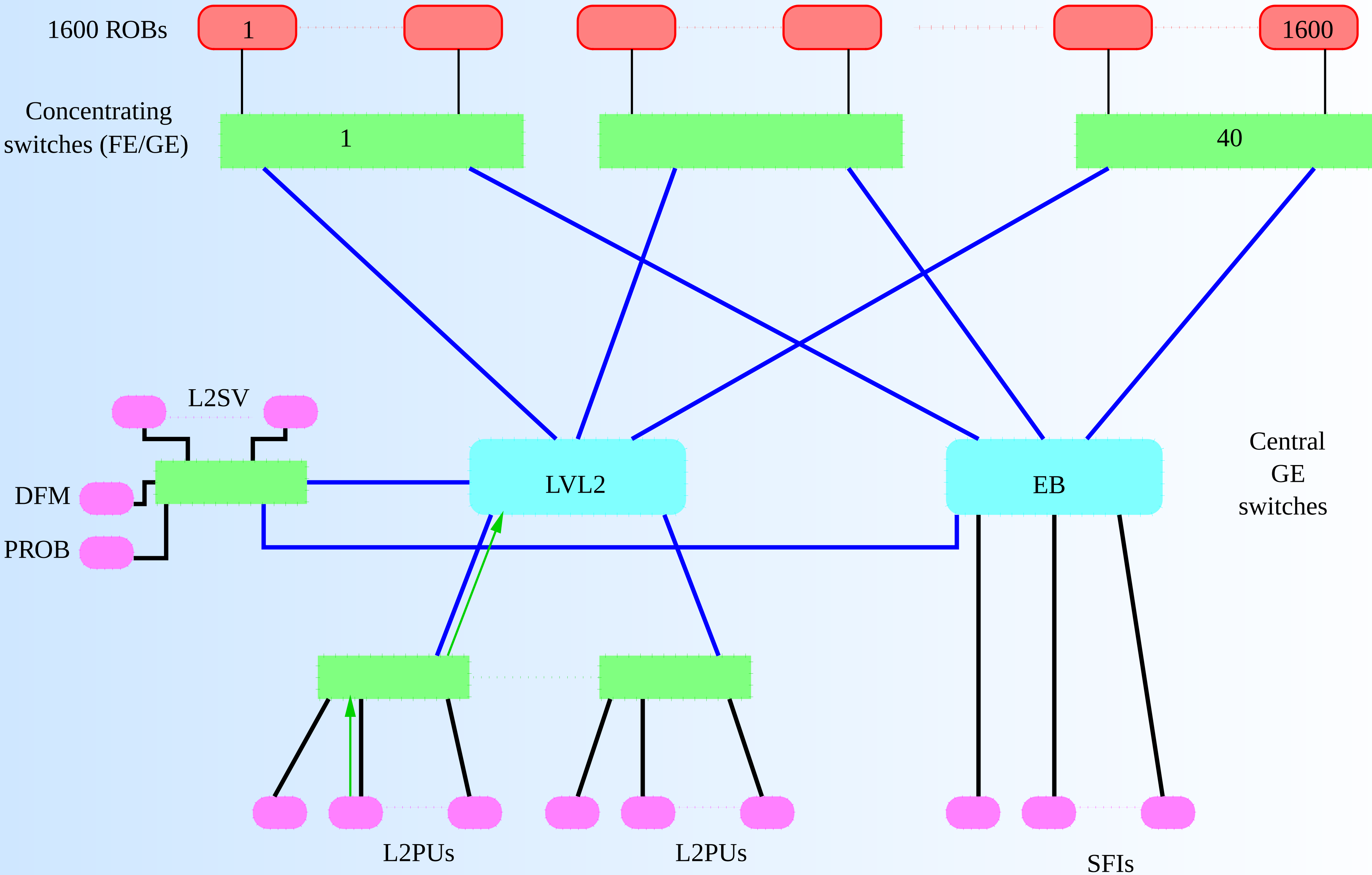
MAC address table size.



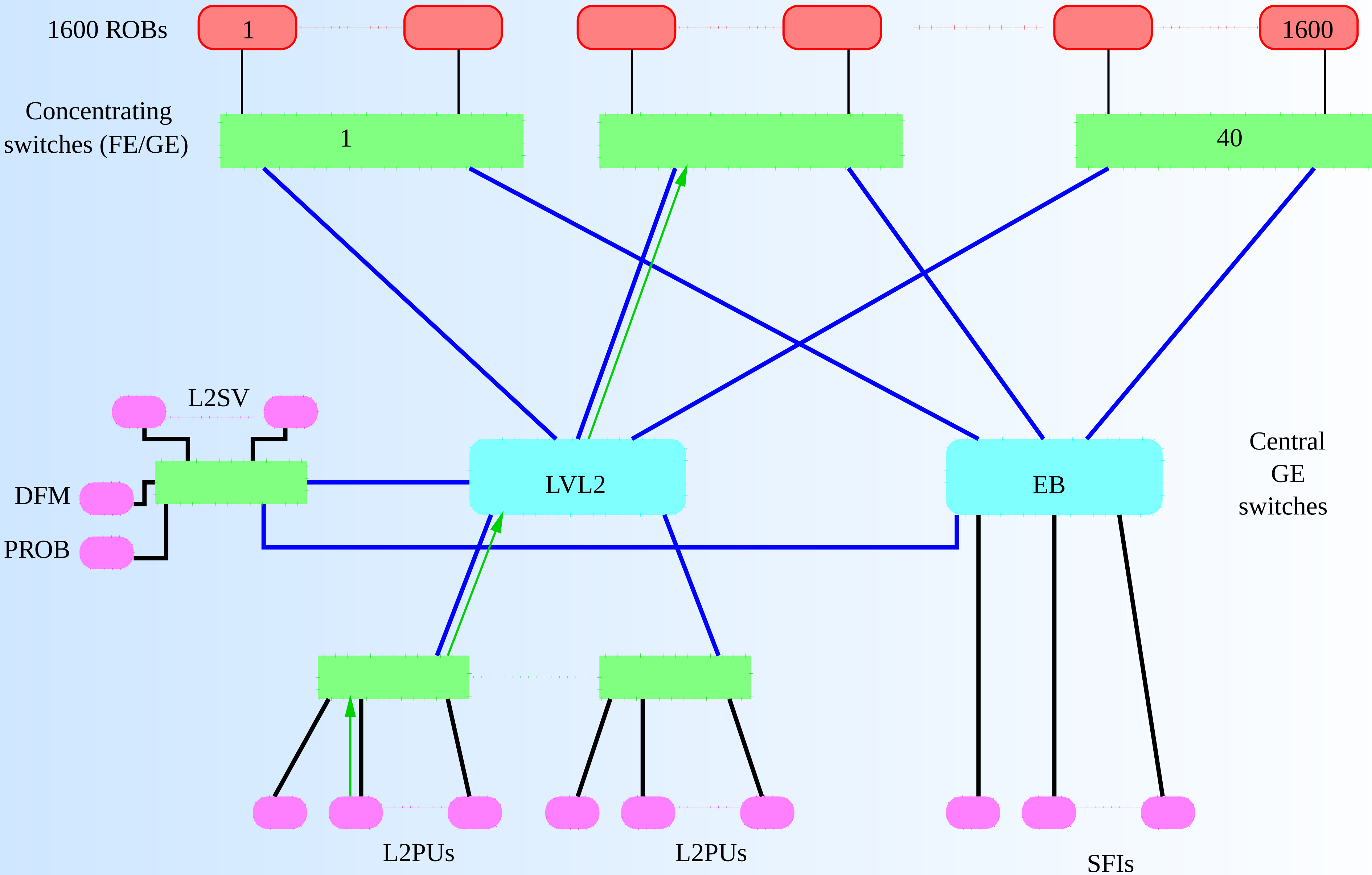
MAC address table size.



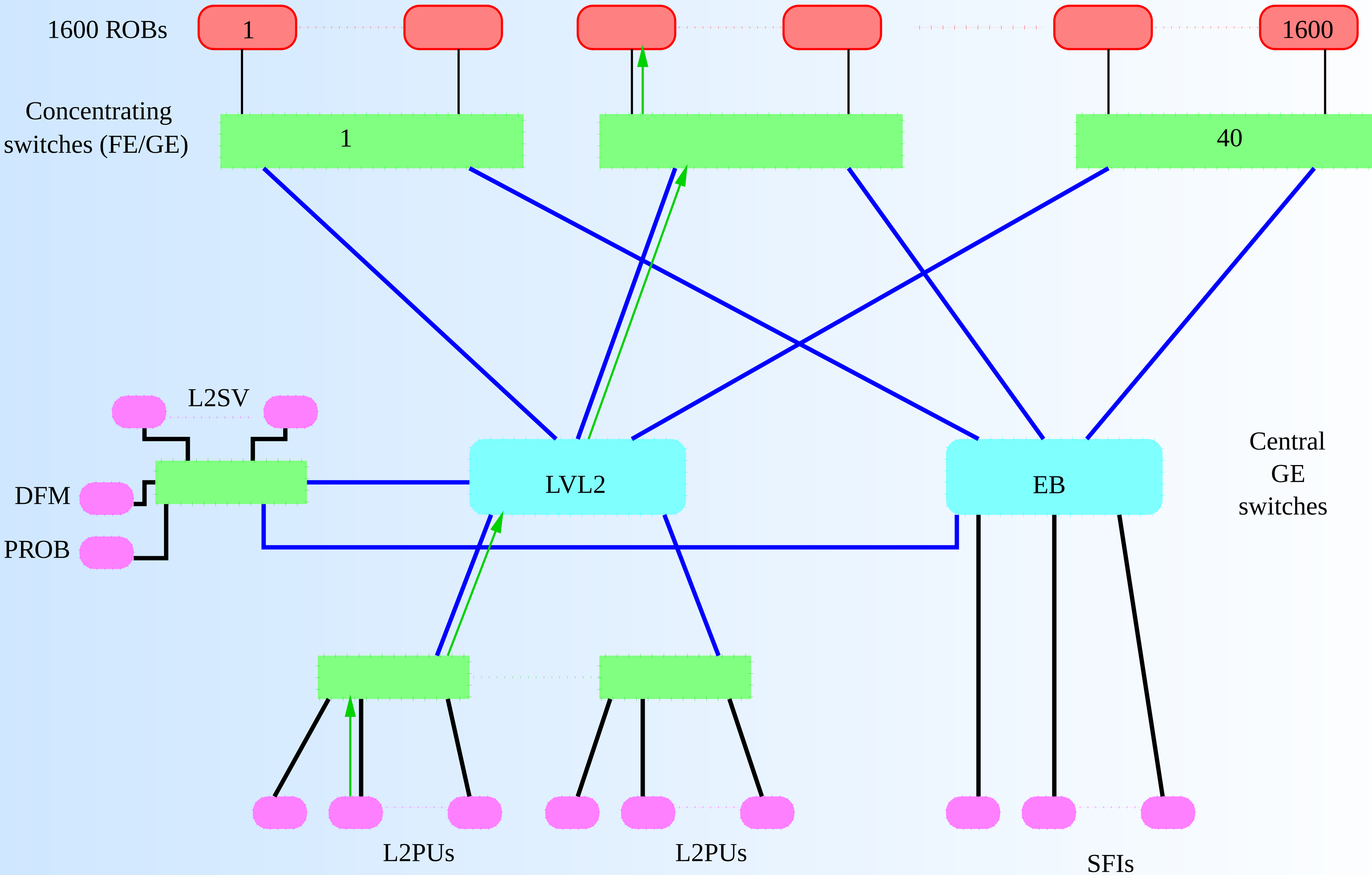
MAC address table size.



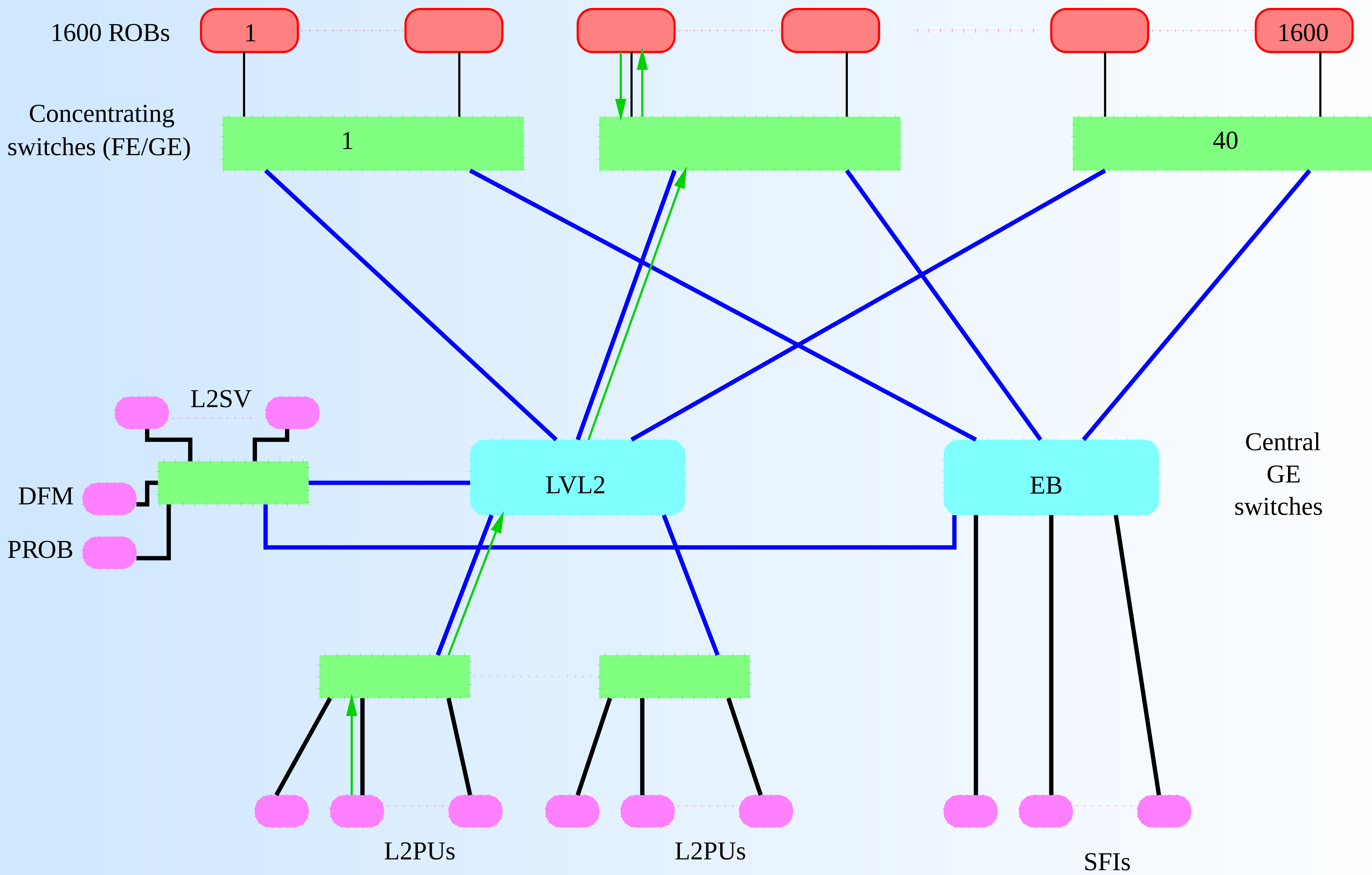
MAC address table size.



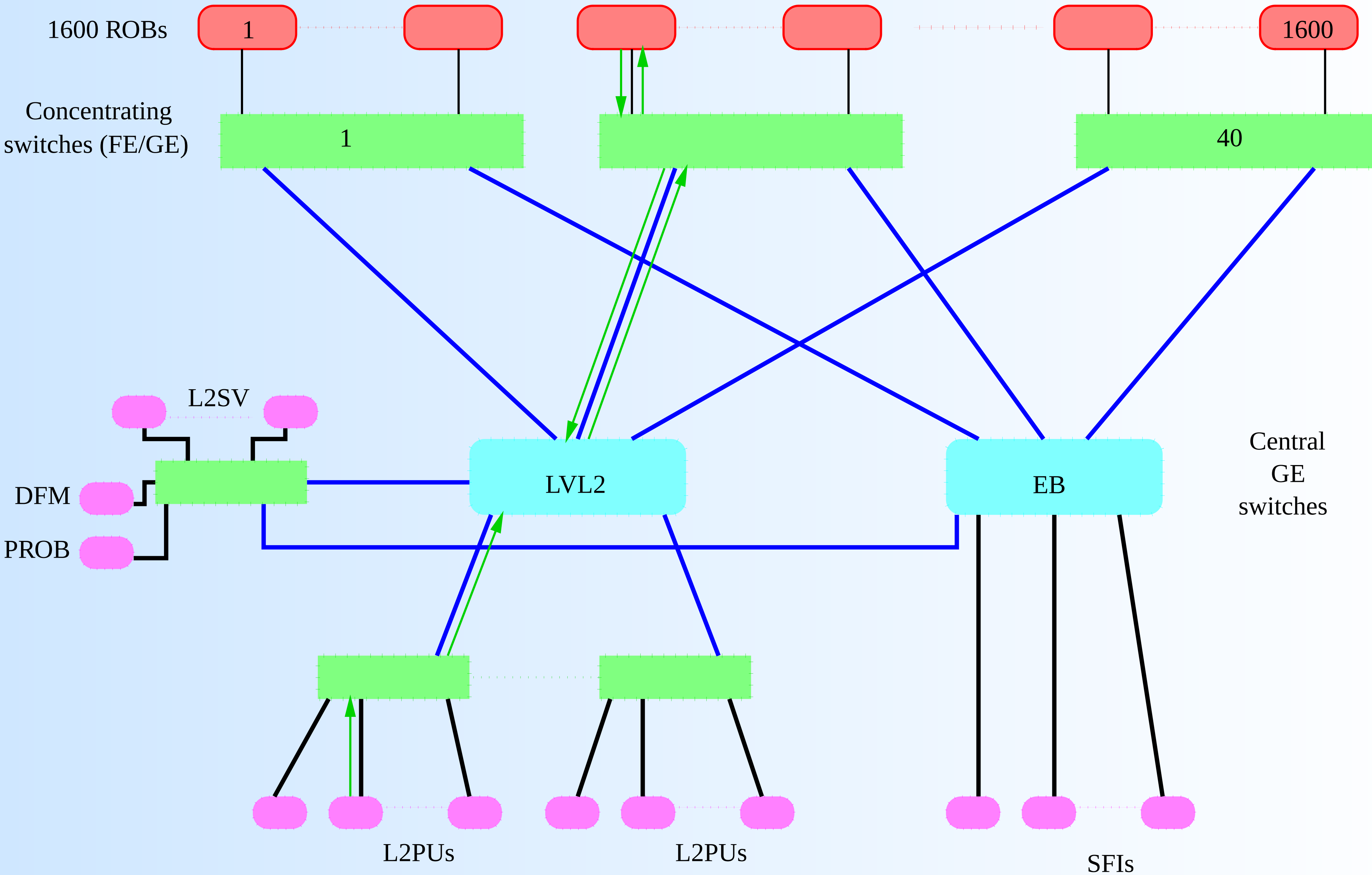
MAC address table size.



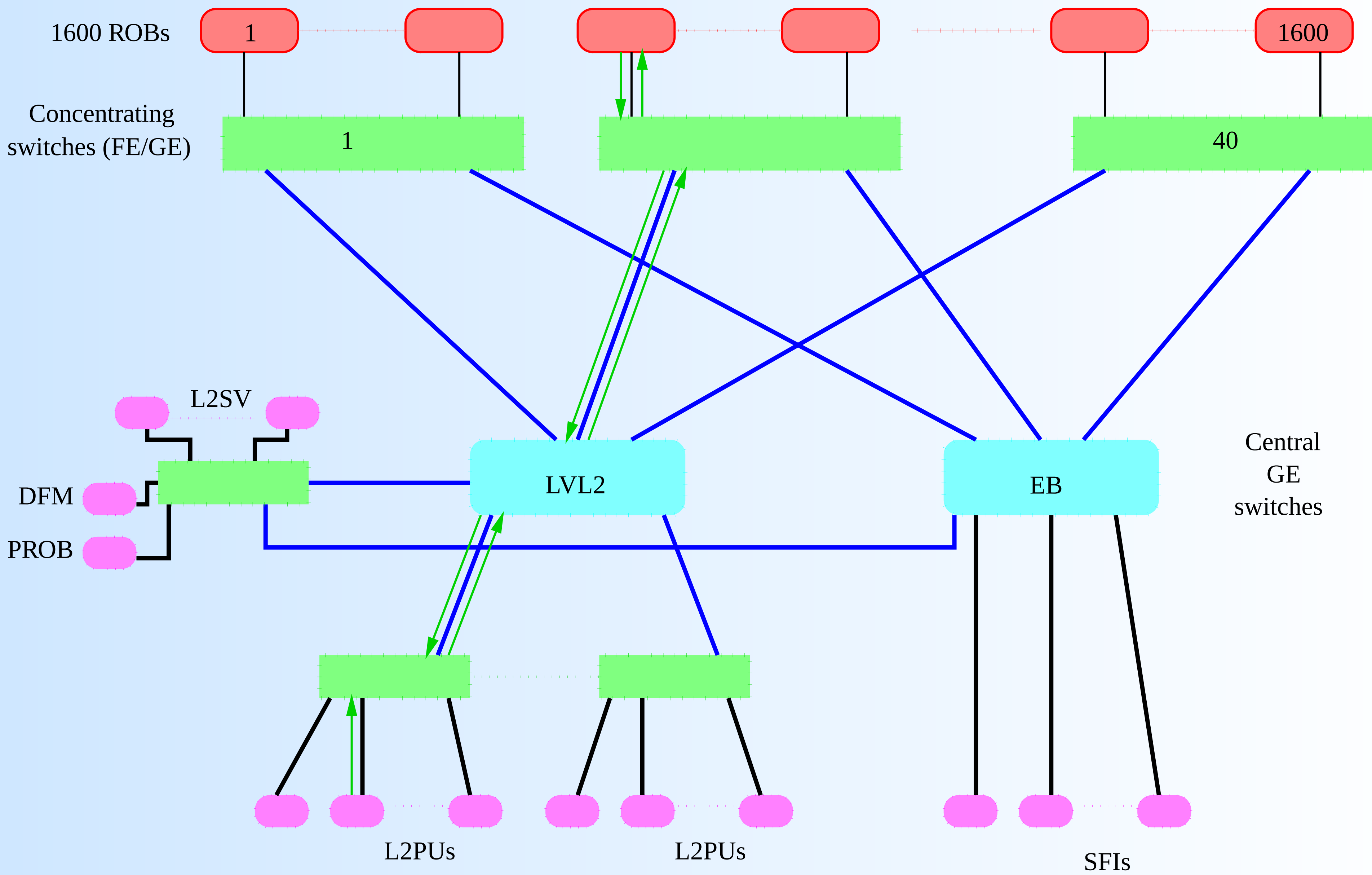
MAC address table size.



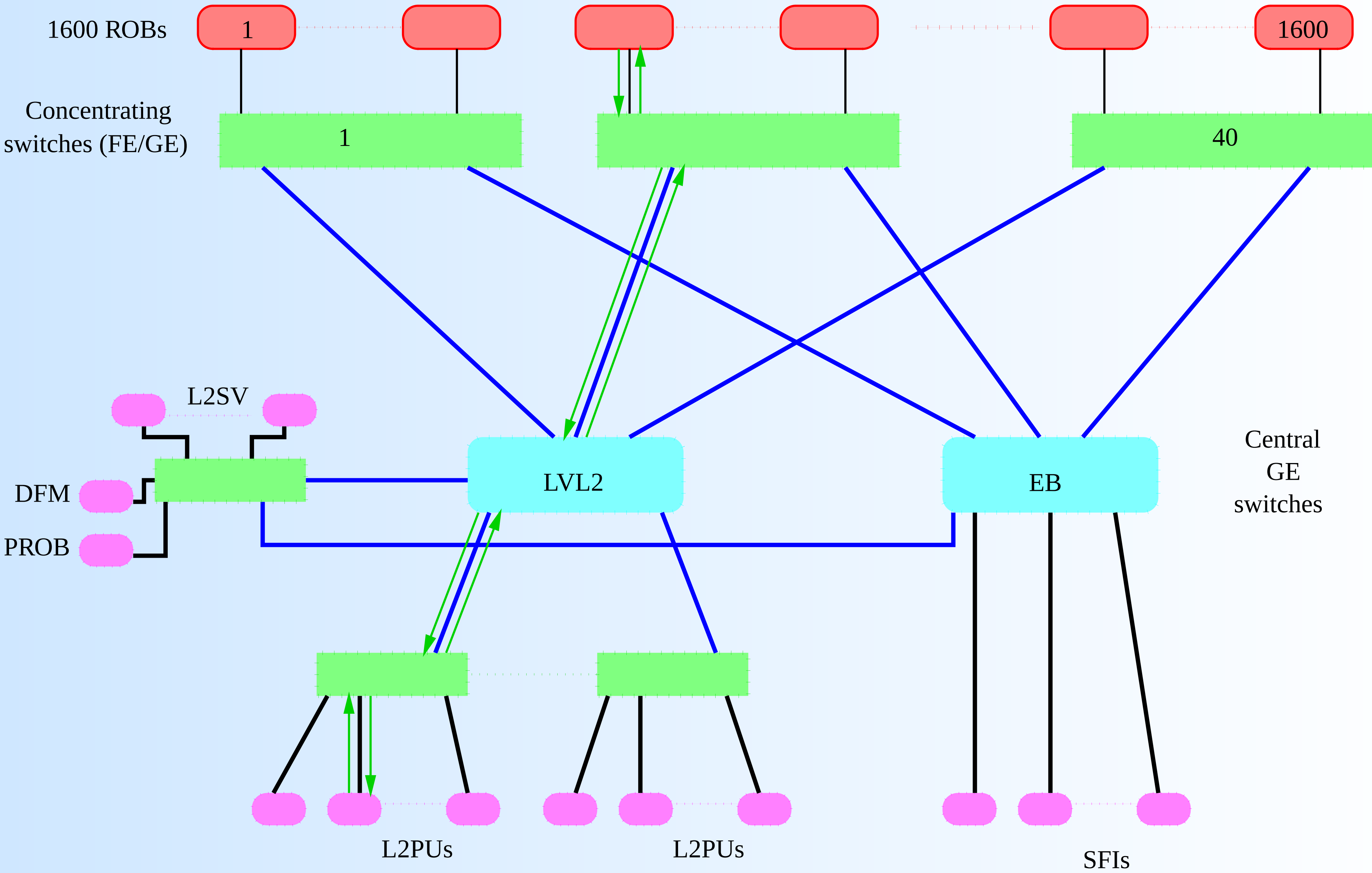
MAC address table size.



MAC address table size.



MAC address table size.



MAC address table size.

- The MAC address table contains the correspondence between the MAC addresses and the switch's ports associated to them.

MAC address table size.

- The MAC address table contains the correspondence between the MAC addresses and the switch's ports associated to them.
- If the switch has no knowledge about the Ethernet Destination address from a received frame, it will forward to all its active ports (*flooding*).

MAC address table size.

- The MAC address table contains the correspondence between the MAC addresses and the switch’s ports associated to them.
- If the switch has no knowledge about the Ethernet Destination address from a received frame, it will forward to all its active ports (*flooding*).
- In order to avoid flooding the size of the MAC address table should be large enough to accommodate all the MAC addresses from the DataCollection network (approx 2000 nodes).

MAC address table size.

- The MAC address table contains the correspondence between the MAC addresses and the switch's ports associated to them.

- If the switch has no knowledge about the Ethernet Destination address from a received frame, it will forward to all its active ports (*flooding*).

- In order to avoid flooding the size of the MAC address table should be large enough to accommodate all the MAC addresses from the DataCollection network (approx 2000 nodes).

- The MAC address table *aging time* is the lifetime of a MAC address in the table, after the reception of the last frame from that address. A common value is ≈ 300 seconds.

MAC address table size.

- The MAC address table contains the correspondence between the MAC addresses and the switch's ports associated to them.
- If the switch has no knowledge about the Ethernet Destination address from a received frame, it will forward to all its active ports (*flooding*).
- In order to avoid flooding the size of the MAC address table should be large enough to accommodate all the MAC addresses from the DataCollection network (approx 2000 nodes).
- The MAC address table *aging time* is the lifetime of a MAC address in the table, after the reception of the last frame from that address. A common value is ≈ 300 seconds.
- The request-response message flow scenario minimizes the probability of erasing the MAC address of a network node from the table.

Trunking (link aggregation)

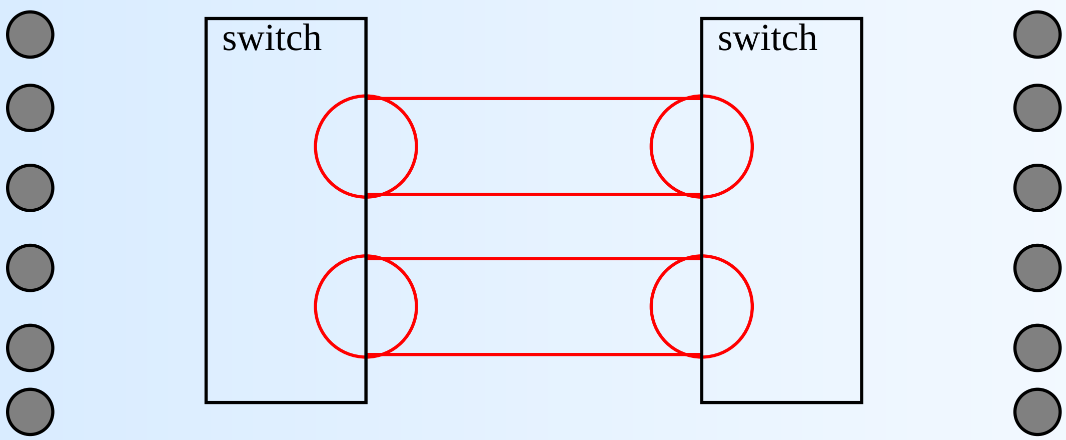
- More physical links are used to form one logical link with a higher bandwidth.

Trunking (link aggregation)

- More physical links are used to form one logical link with a higher bandwidth.
- Once a link is associated to a pair of MAC addresses , all the frames addressed to any of them will go on that link, until the switch forgets about one of the two addresses (aging).

Trunking (link aggregation)

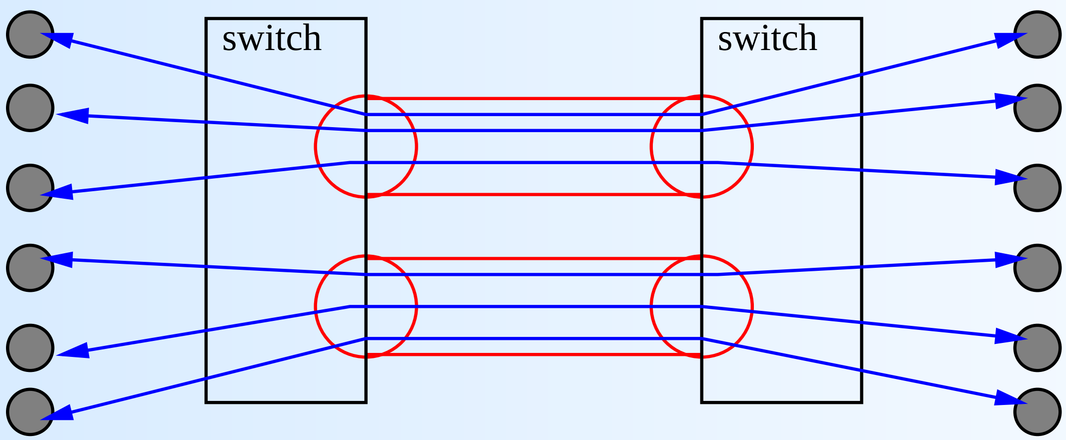
- More physical links are used to form one logical link with a higher bandwidth.
- Once a link is associated to a pair of MAC addresses , all the frames addressed to any of them will go on that link, until the switch forgets about one of the two addresses (aging).



Trunking (link aggregation)

- More physical links are used to form one logical link with a higher bandwidth.

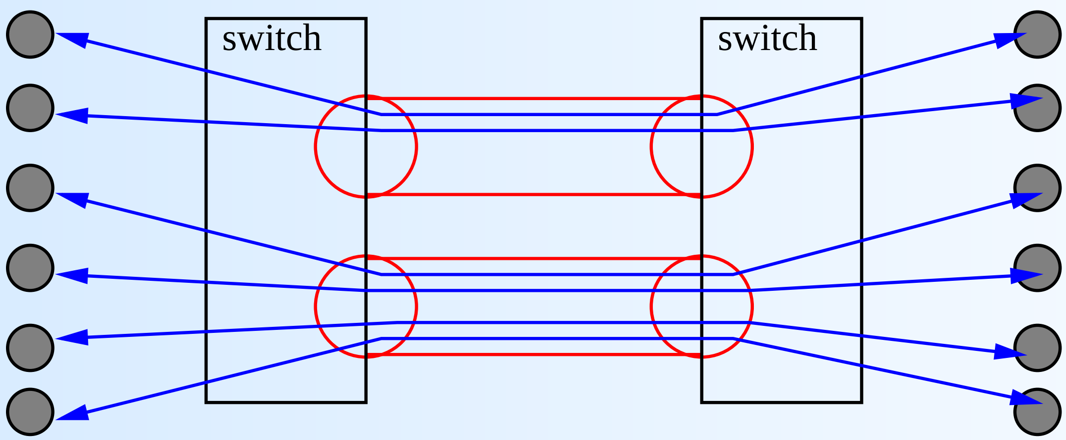
- Once a link is associated to a pair of MAC addresses , all the frames addressed to any of them will go on that link, until the switch forgets about one of the two addresses (aging).



Trunking (link aggregation)

- More physical links are used to form one logical link with a higher bandwidth.

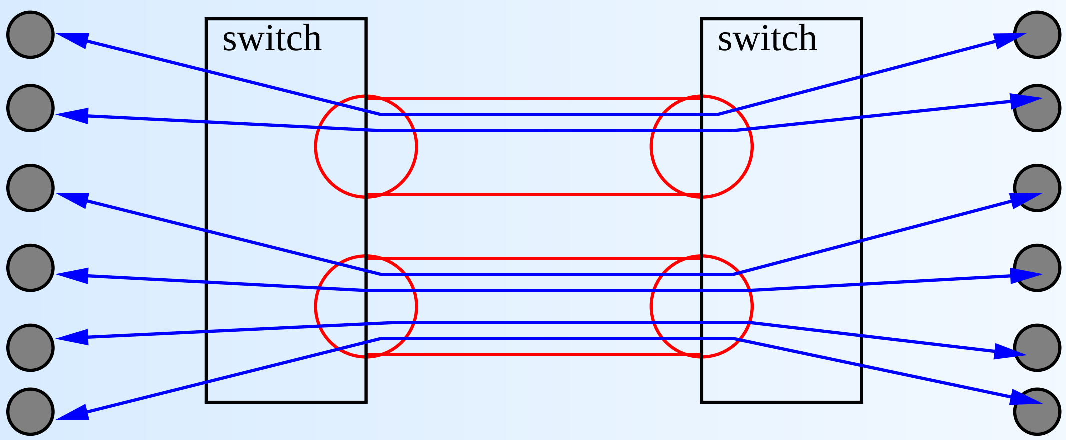
- Once a link is associated to a pair of MAC addresses , all the frames addressed to any of them will go on that link, until the switch forgets about one of the two addresses (aging).



Trunking (link aggregation)

- More physical links are used to form one logical link with a higher bandwidth.

- Once a link is associated to a pair of MAC addresses , all the frames addressed to any of them will go on that link, until the switch forgets about one of the two addresses (aging).



- Load balancing problem between the physical lines assigned to the trunk.