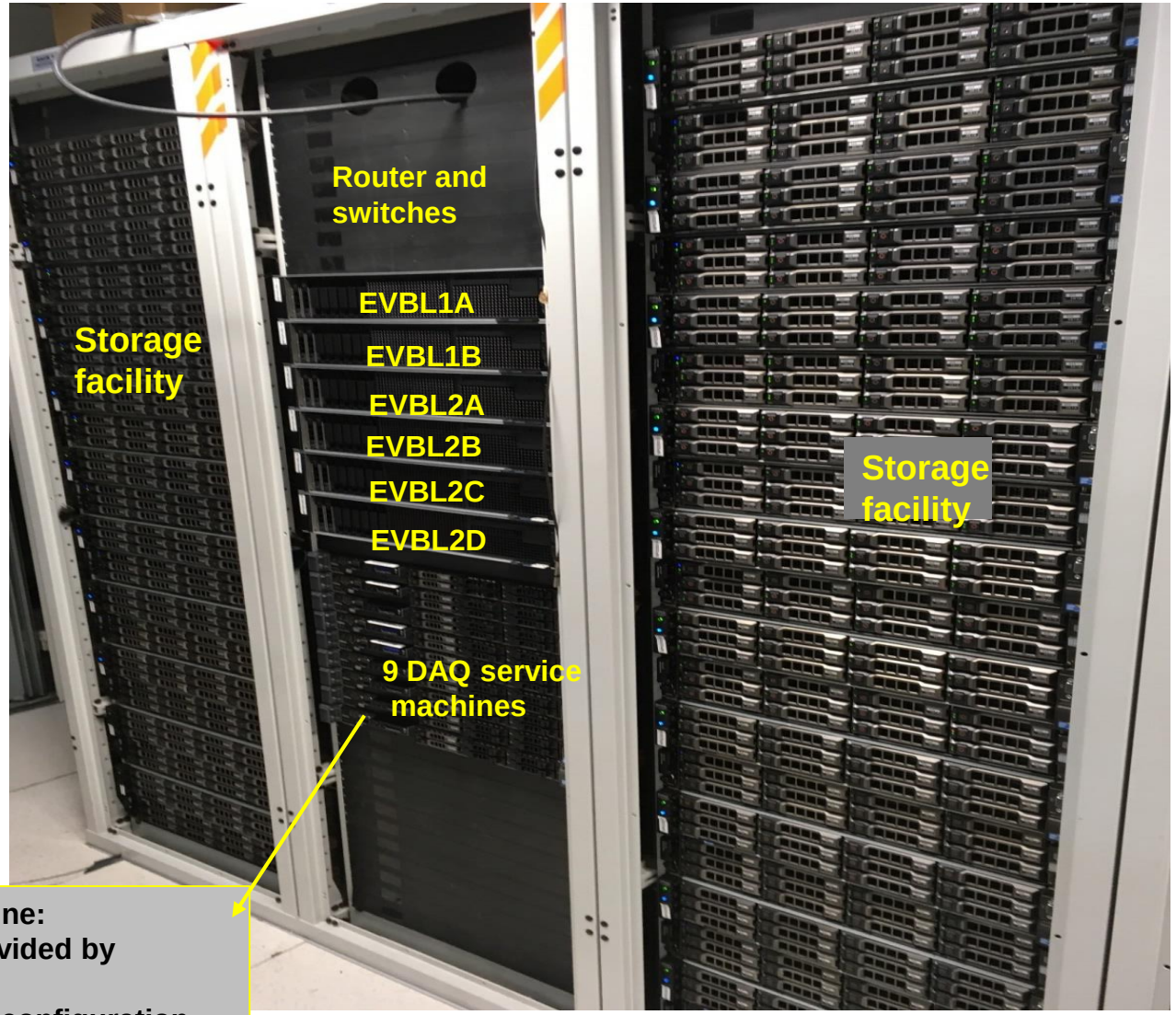
Joint data challenge 3: a first discussion on scope definition for NP02

Denis Pugnere, Elisabetta Pennacchio

IT coordination meeting, 25/04/2019

DAQ room @ EHN1



DAQ service machine:
9 servers R610 provided by CCIN2P3
(metadata servers, configuration server, online processing server...)

- The first (extremely useful) joint data challenge was performed in April 2018. The DAQ back-end system was already operational at that time, since April 2018 we have been continuing to perform “data challenge” exercises for any modification or addition of a new element → we do not expect to extract new information by performing a new data challenge apart a general verification that things are still fine, as in the past, on the side of the IT division (link to IT and central CERN EOS) and for the data transfer to FNAL

- We plan then to perform a new JDC (end of May- beginning of June). The start of NP02 data taking is foreseen for August 2019

- Before start discussing the scope of this new JDC it is important to remember what has been done since last year. A detailed description is given in next 3 slides (was shown during the DUNE collaboration meeting of September 2018)

- We will also briefly remind in the next slides the main parts of the NP02 system:
  front-end system: 12 mTCA crates, connected to back end system by 12 Ethernet optical links operating at 10 Gbit/s
  back-end system: →2 L1 event builders, each one assembling one half on the events
  →4 L2 event builders: they merge event halves, packing them in datafile of 3GB, that are then moved to the local eos instance
  →local eos instance (np02eos) : 1PB of disk space including RAID redundancy, 20 storage servers

- <u>The front end and the back-end systems have been installed and are operational.</u>

Description of ProtoDUNE-DP back-end system

- The ProtoDUNE-DP DAQ back-end system consists of two levels of event building machines (EVB L1 and EVB L2) plus the network infrastructure, and the online storage/processing facility. The event builders task is to receive in input the data flow from the front-end system, build the events and cluster them in data files, and write these data files into the local storage servers. This presentations aims to discuss how this is implemented in ProtoDUNE-DP .

- The DAQ in protoDUNE-DP is trigger based and the DAQ FE transmits the drift window corresponding to each event. Event size=15MB compressed (assuming a compression factor~10), rate 100 Hz

- INPUT to the back-end:

The Front-End system of ProtoDUNE-DP consists of 12 uTCA crates for the charge readout and 1 crate for the light readout. All these crates are connected to the back-end system with Ethernet optical links operating at 10 Gbit/s.

→ x12 10 Gbit links for charge readout + x1 10Gbit link for light readout

Event Building

Level 1: data are transferred from the network to the RAM of two level 1 machines. The task of each machine is to put together data from the uTCA crates for the same drift widow corresponding to half of the detector, and the light readout data.

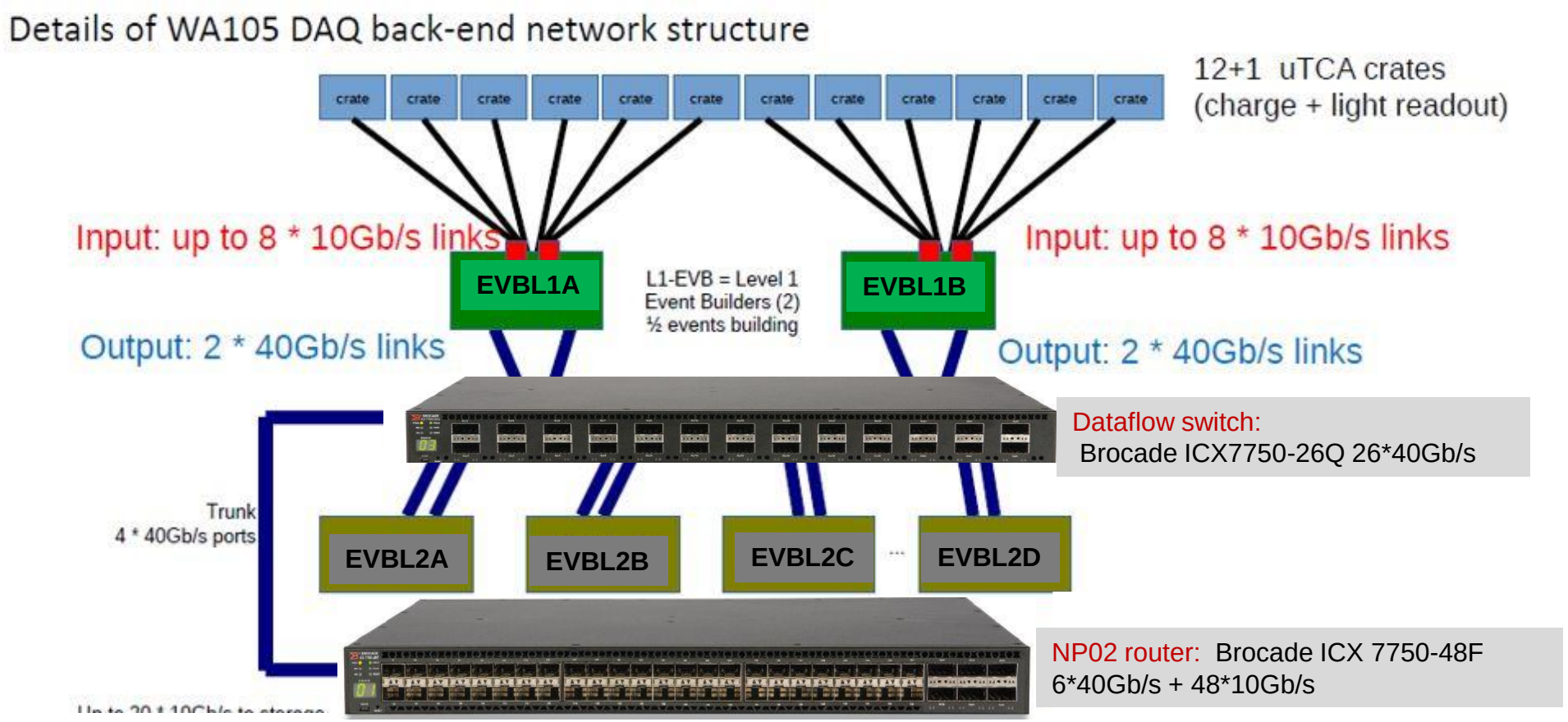
→ Two machines DELL R730 are used (384 GB RAM, 2 Intel cards R710, 2 Mellanox Connect X3 2 ports, 40Gb/s Ethernet QSFP+, CPU type Intel XEON Gold 5122 3.6 GHz, 4 cores, 8 threads)

Level 2: the data from the two L1 event builders are sent via the network to four level 2 machines working in parallel. The task of each machine is to put together the two events halves in a complete event and assemble multi-events files to be written on disk

→ Four machines DELL R730 are used: they have similar specifications as the LV1 but need less connectivity (since there is no need for the x8 10 Gbit/s links in input) and need less RAM memory for the event building (192 GB).
(192 GB RAM, 2 Mellanox Connect X3 , CPU type Intel XEON Gold 5122 3.6 GHz, 4 cores)

DELL R730

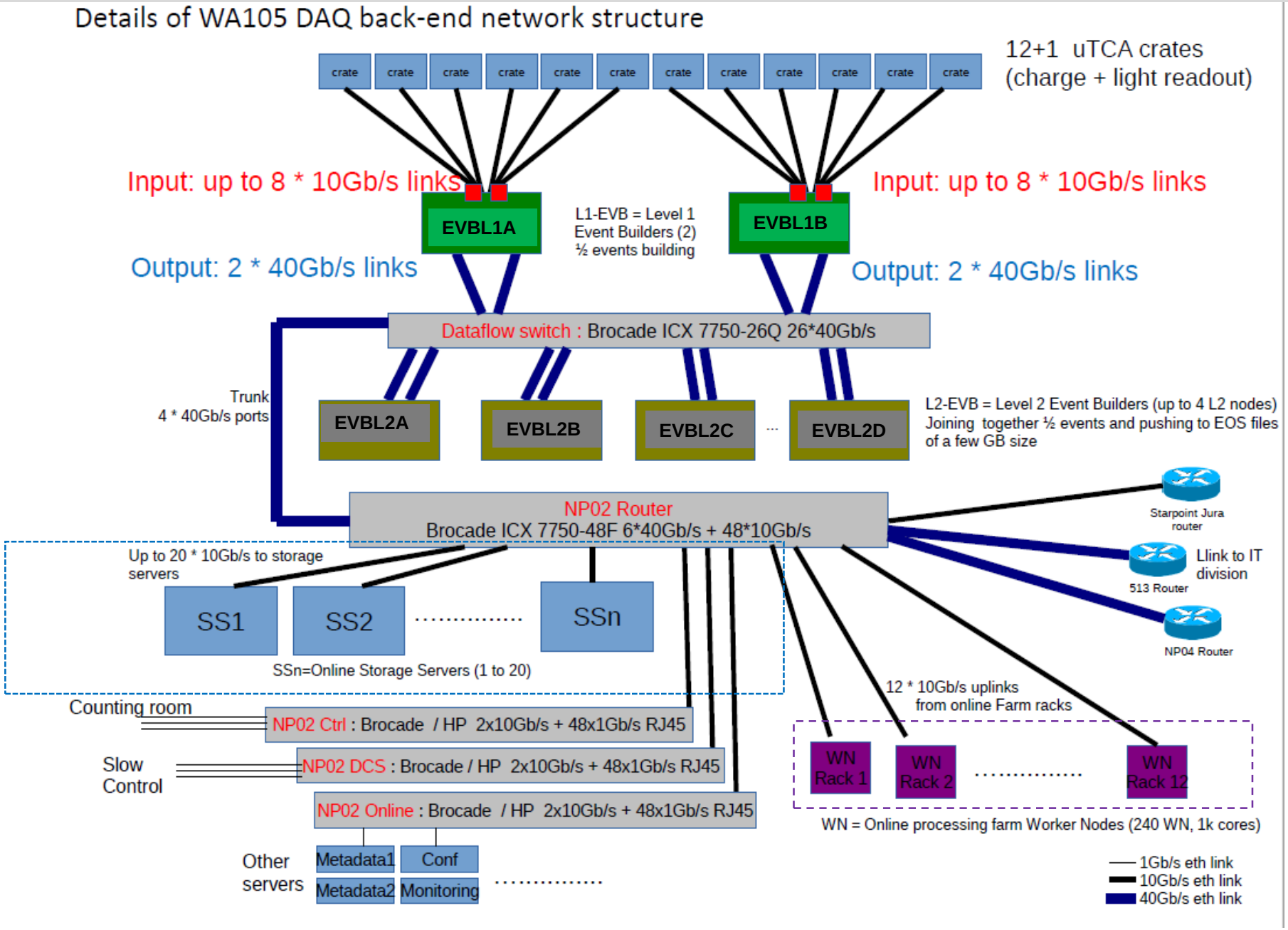
The events builders are interconnected among them by a dedicated switch:



The LV2 event builders loop in parallel on the list of events in the L1 EVBs memories following a precise ordering and assemble the event halves in event files of 3 GB size transmitted to the storage facility

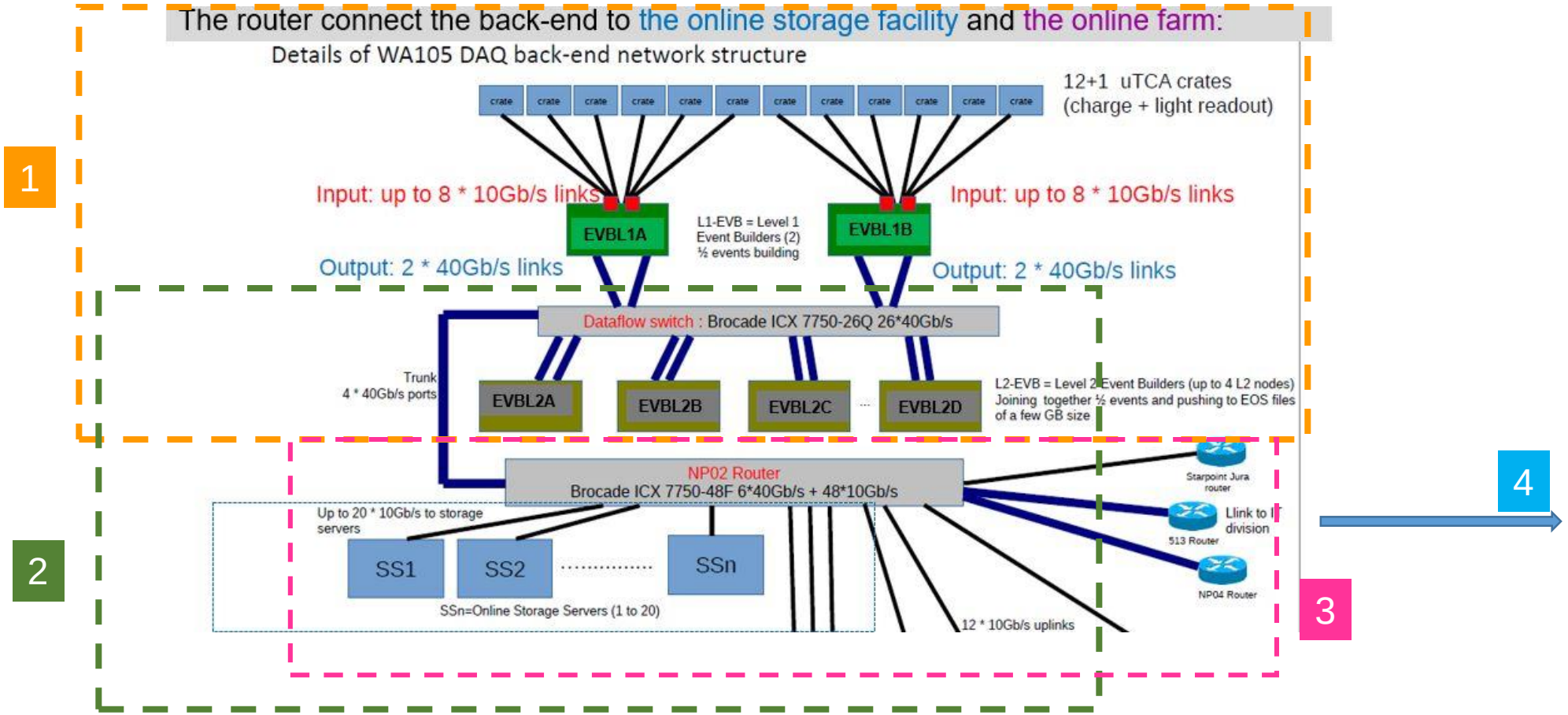
The configuration of the L2 machines has been optimized in terms on RAM and CPU so that four event builders can handle the data flow. Anyway, the system can be expanded: more L2 EVB machines can be added

The router connect the back-end to the online storage facility and the online farm:



4 main data-flow blocks in the NP02 DAQ system:

1. From the mTCA to the L1 Event builders and L2 Event builders (from bits to datafiles….)
2. From the L2 event builders to NP02 EOS instance
3. From NP02 EOS instance to central CERN EOS
4. From the central CERN EOS to CASTOR and Fermilab



1. **From the mTCA to the L1 Event builders and L2 Event builders (from bits to datafiles….)**

We are testing this part since a few weeks: we are now carefully checking the results to understand the system and find out its time stability

2. **From L2 event builders to NP02 EOS instance**
   This had been tested during several campaigns from December 2018 to February 2019, under different conditions (EOS configuration, optimization in the number of parallel transfers…). Tests have been resumed in these days and are going on right now

1+2 We have also been running the two parts together during the last weeks (typical single uninterrupted run duration ~ a few hours)

3. **From the NP02 EOS instance to central CERN EOS**
   We plan to use CERN FTS. We tested it last year in July, and this allowed to develop and set up all scripts. Some work has still to be done in order to automatize them (planned for May)

All these parts are under systematic tests now and these tests will continue during all the next month. We do not think that we need to put them in the scope of the joint data challenge, since they are mainly related to NP02 online, and we need to test and exercise them in a constant way for other reasons (stability and stress tests of the front-end)

→ We would like to check that on the side of the central CERN EOS there are no problems in transferring the data to CASTOR + FNAL and that we reproduce the expected performance

What we would then like to test in the JDC is the file transfer from CERN EOS to Castor and FNAL

Last year we tested the transfer to FNAL, so this will be mainly a repetition, basically to check again the metadata file generation, the directories in use and the permissions (last year NP04 was not able to remove the metadata files due to a permission issue)

The transfer to CASTOR is in our opinion the main point, since it has been never tested before.

These tests can last 3 days; JDC2 lasted one week, but everything was new (*This year is no longer needed to run long transfer test (24h) from NP02 EOS to CERN EOS to test the dedicated link as we did last year)*. It is mainly a repetition of what we already did, to check again everything before data taking