Reimplementing the bulk data system with DDS in ALMA ACS

Bogdan Jeram(bjeram@eso.org)
European Southern Observatory

G. Chiozzi, R. Javier Tobar, ESO, Garching bei Muenchen, Germany
R. Amestica, NRAO, Charlottesville, VA, U.S.A
M. Watanabe, NAOJ, Tokyo, Japan

Summary

- ALMA & ACS
- Bulk data – overview
- The “old” bulk data based on CORBA A/V
  - limitations
- The “new” bulk data based on DDS
  - Development process
  - Challenges
  - Improvements
- Conclusions

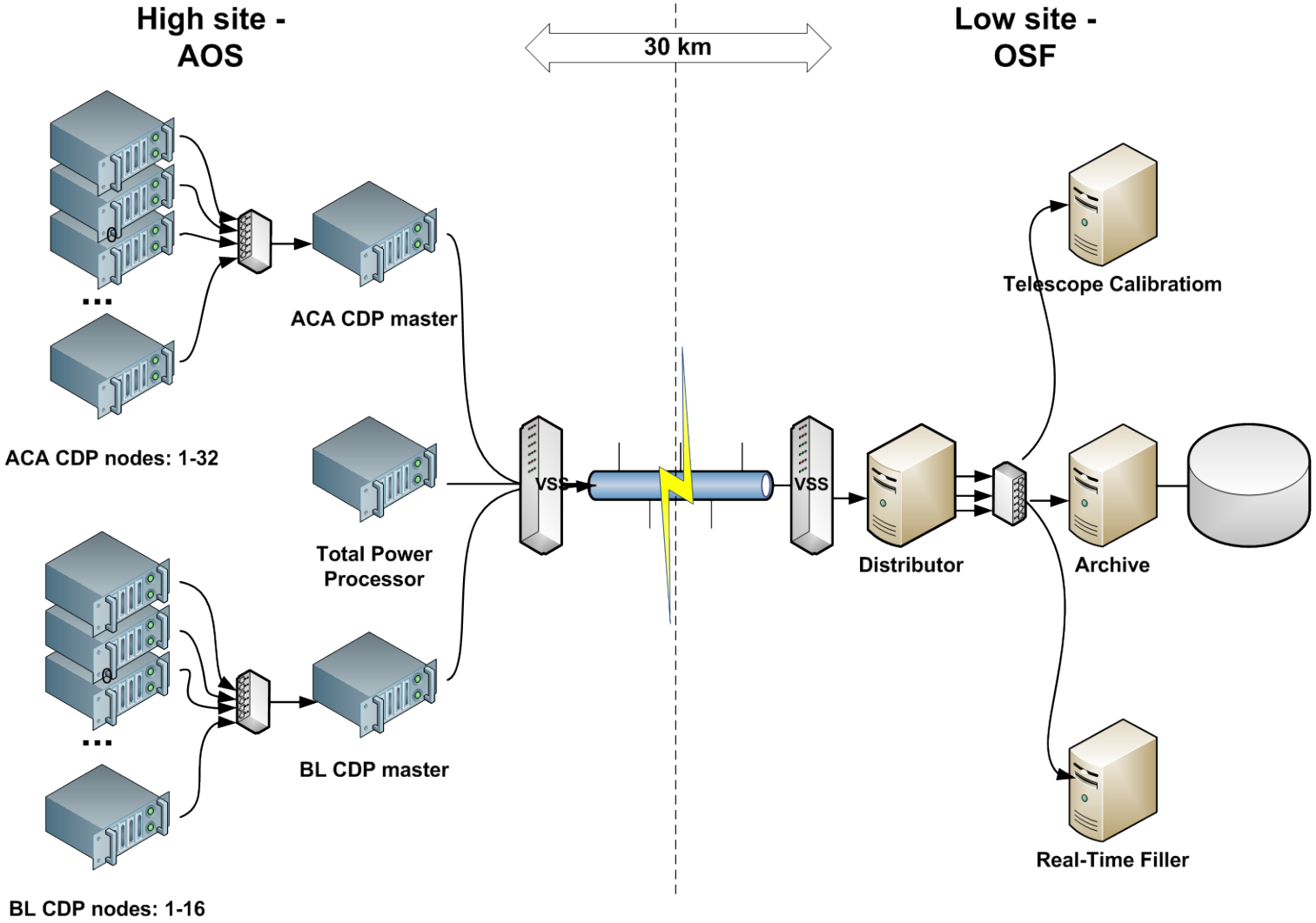
ALMA&ACS

- **A**tacama **L**arge **M**illimeter/submillimeter **A**rray
  - Collaboration between Europe, North America and East Asia organizations
  - 66 antennas: 54x12m + 12x7m
  - At 5000m on Chajnantor plateau in Atacama desert, Northern Chile
- **A**LMA **C**ommon **S**oftware
  - Container-component, CORBA based, middleware
  - Develop distributed (control) systems

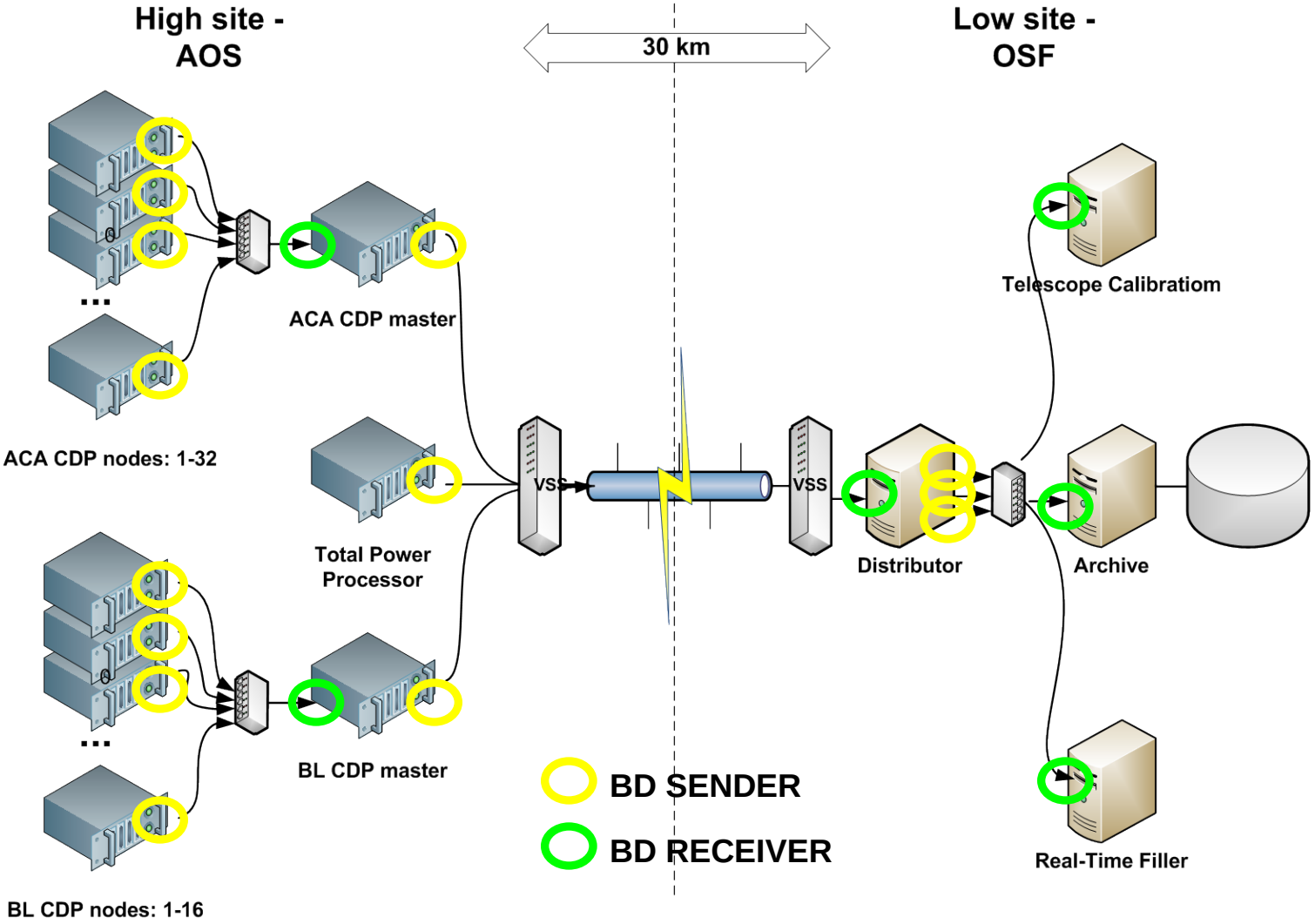
Bulk Data - BD

- ACS service for **reliable** and **concurrent** streaming of high volumes of (astronomical) data
- Used in two configurations:
  - Many senders to one receiver
  - One sender to many receivers - multicast
- Used by 6 ALMA SW sub-sytems
- Total peak data rate: 64MBytes/sec

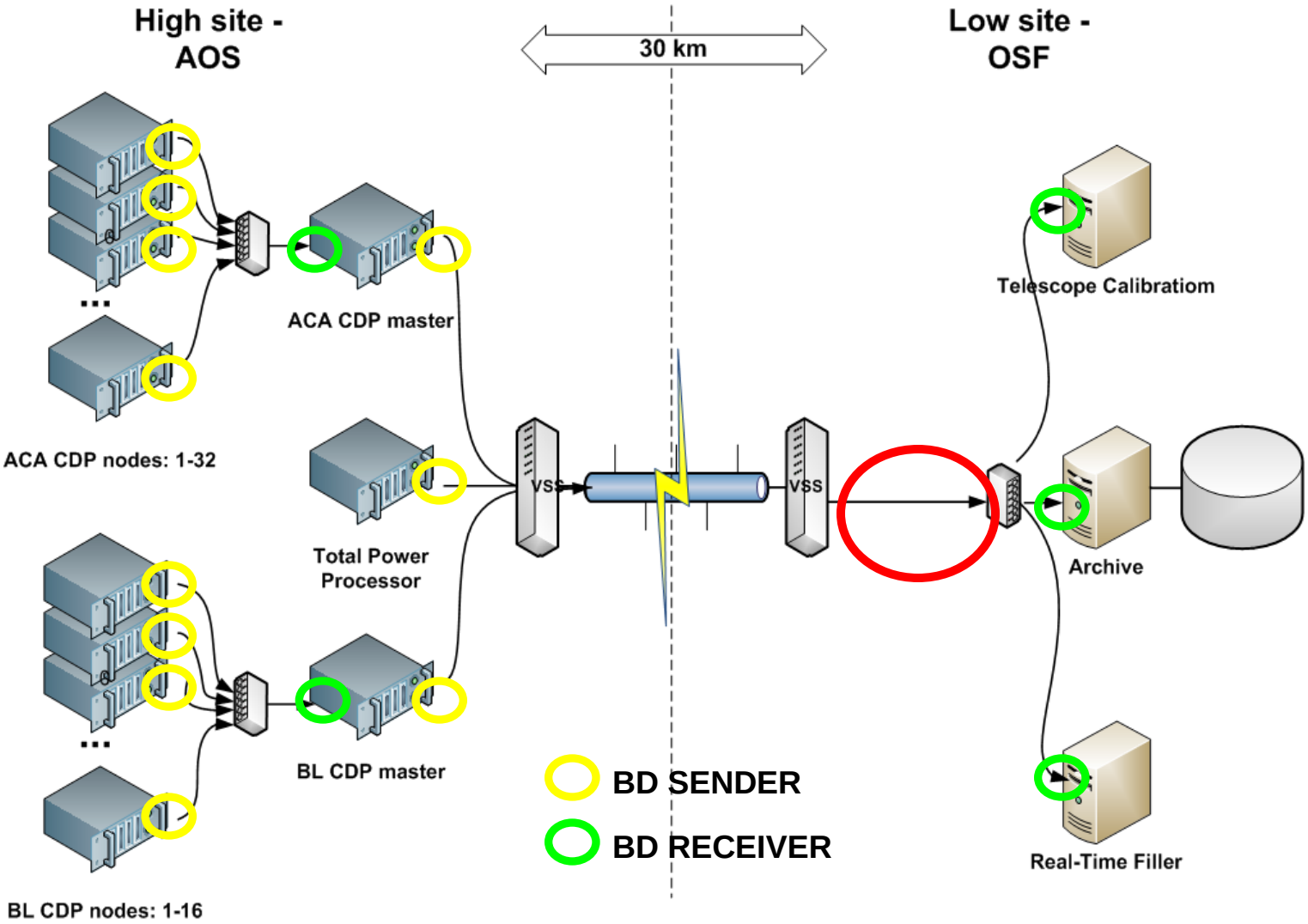
Deployment - old BD



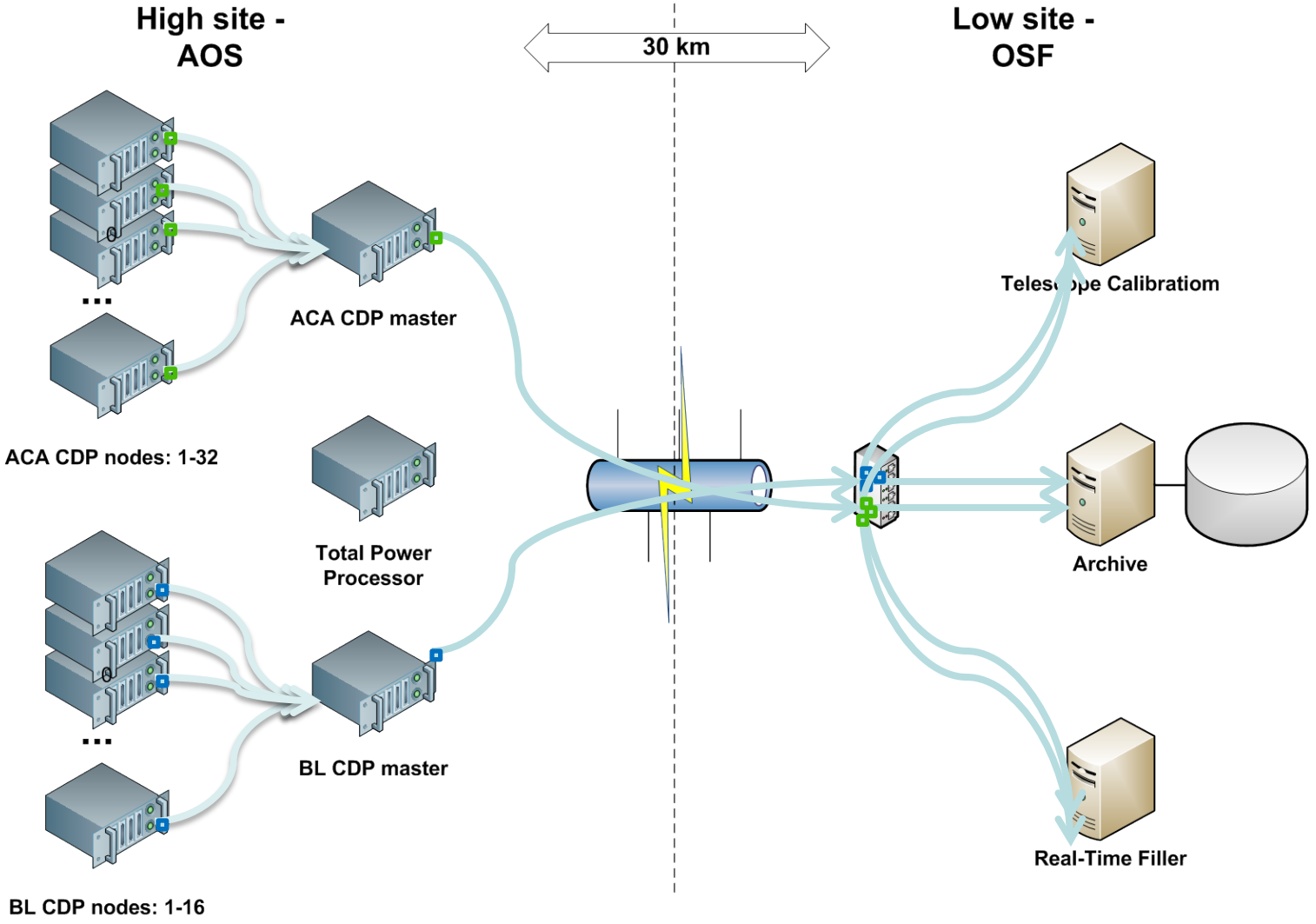
Deployment - old BD



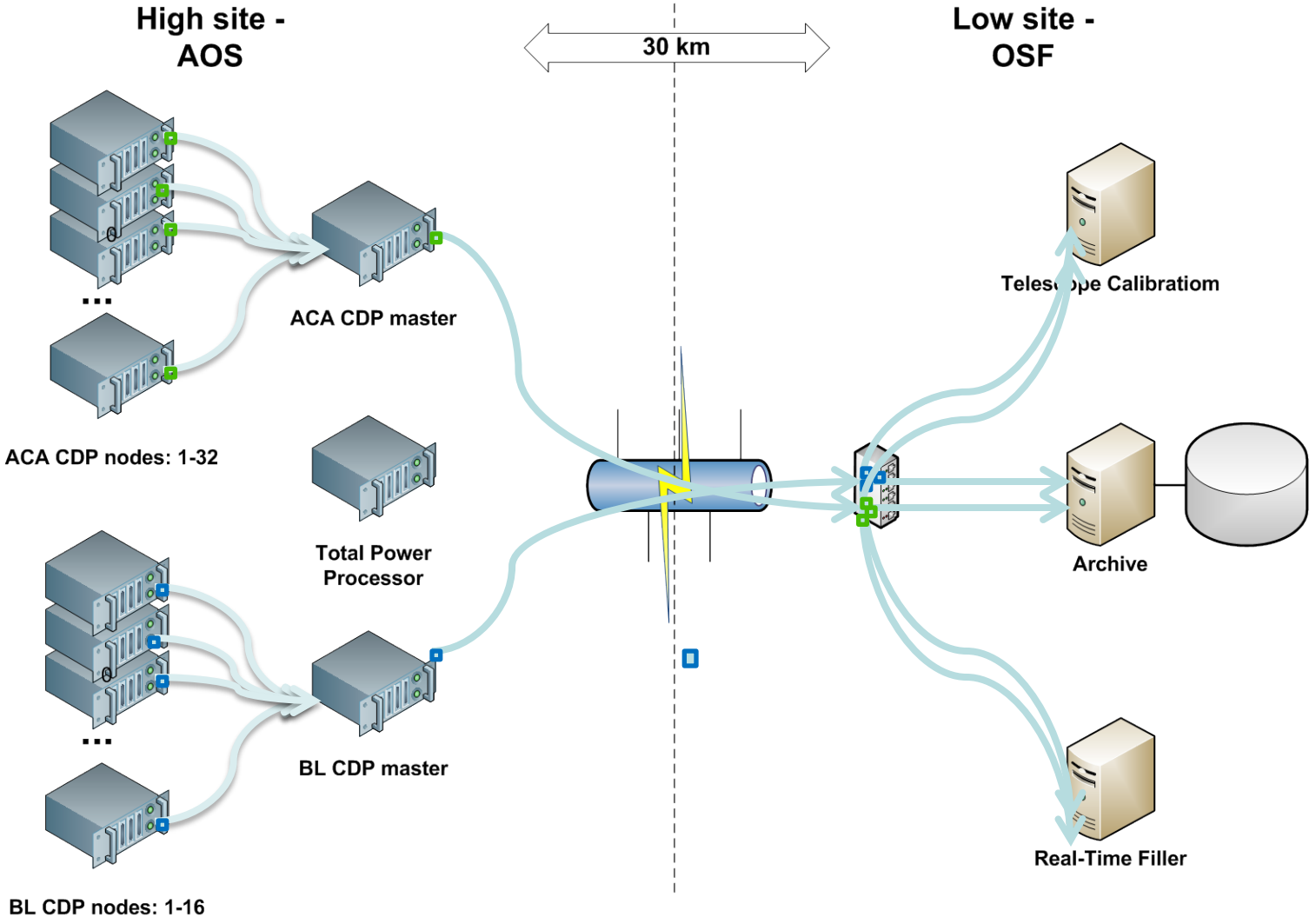
Deployment – new BD



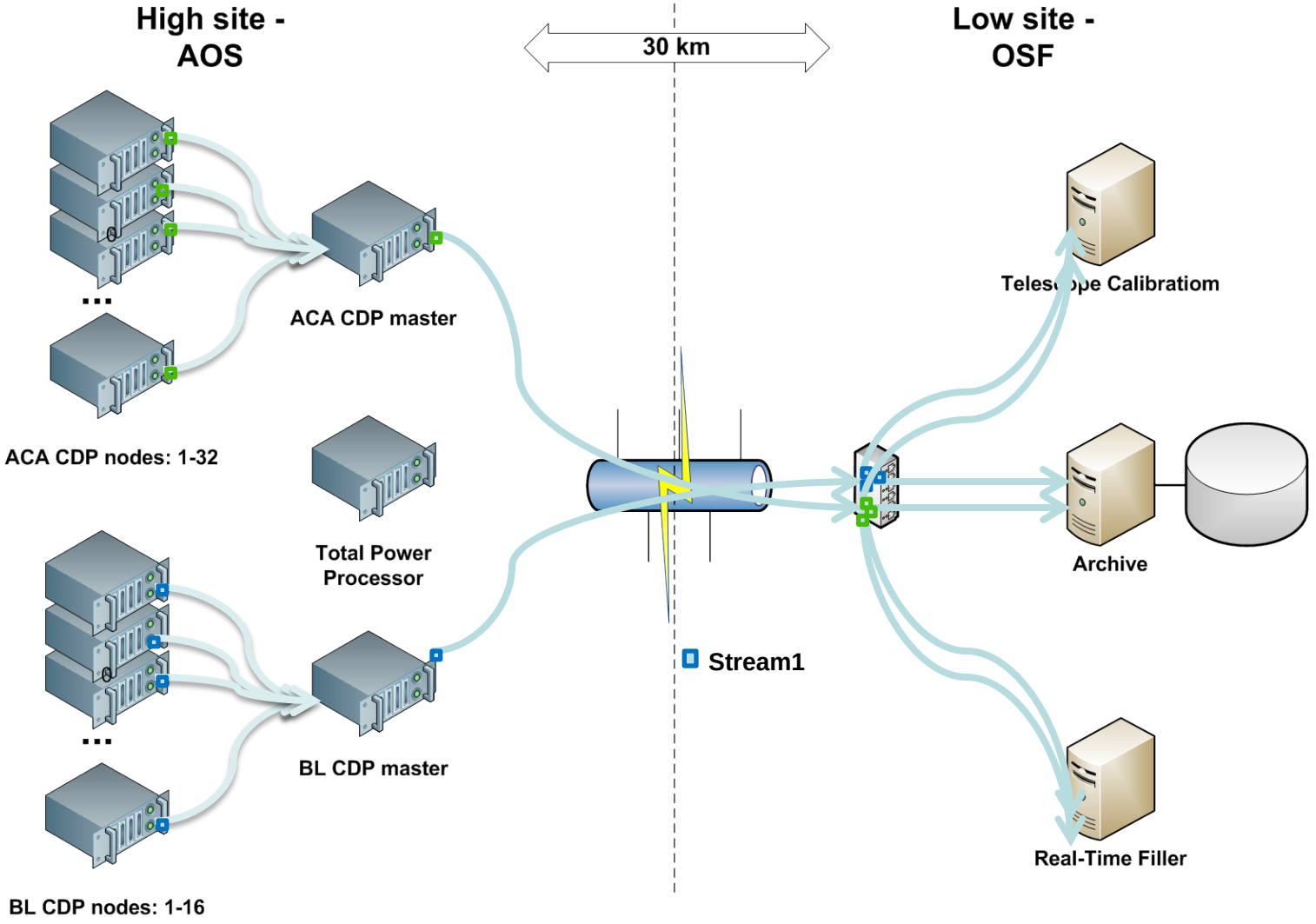
Data transfer



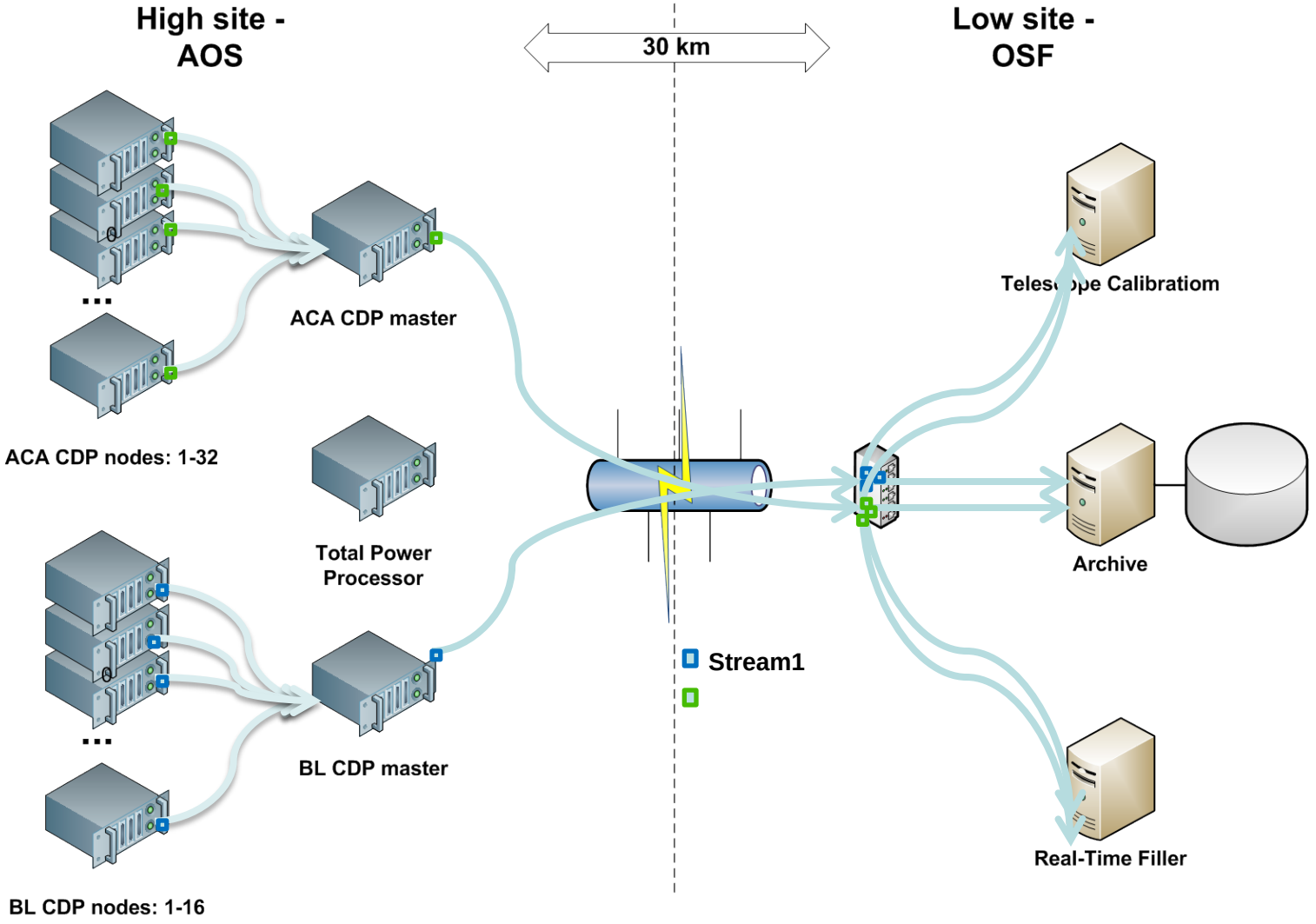
Data transfer



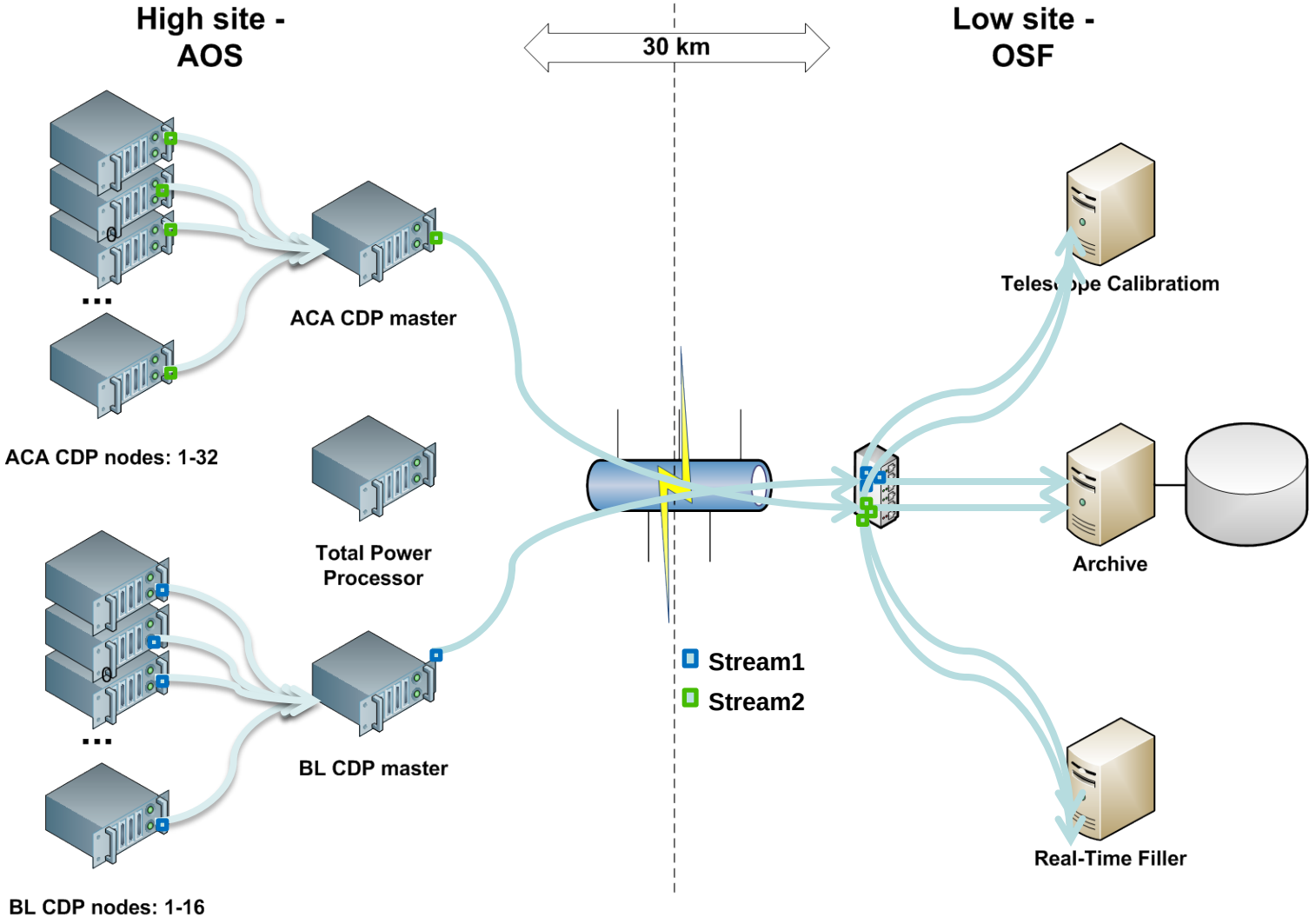
Data transfer



Data transfer

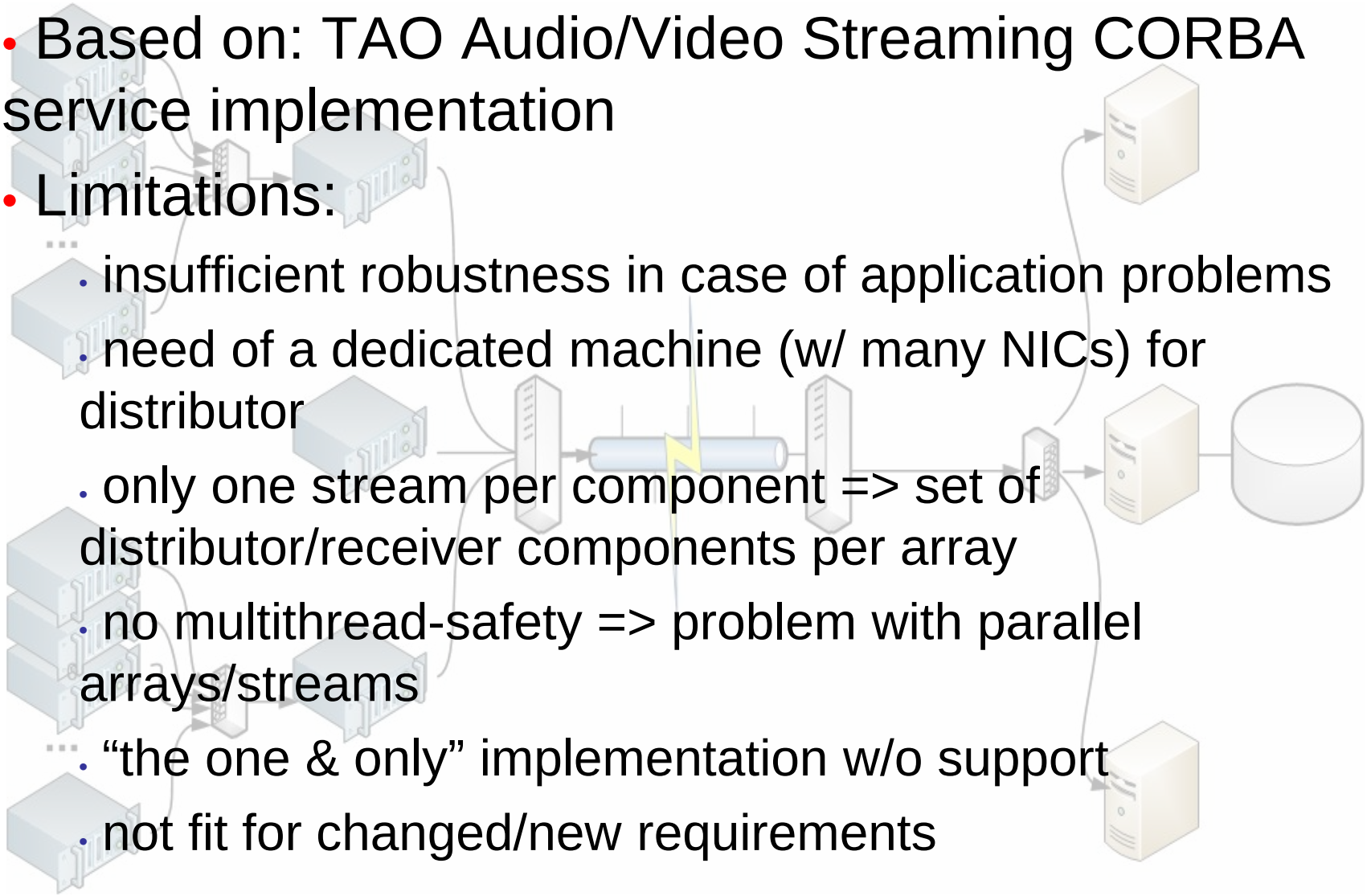


Data transfer



A/V based Implementation

- Based on: TAO Audio/Video Streaming CORBA service implementation
- Limitations:
  - insufficient robustness in case of application problems
  - need of a dedicated machine (w/ many NICs) for distributor
  - only one stream per component => set of distributor/receiver components per array
  - no multithread-safety => problem with parallel arrays/streams
  - “the one & only” implementation w/o support
  - not fit for changed/new requirements



New implementation
BulkData New Technology - BDNT

Critical phase – early science operation

- Carful planning
- Switch between old and new BD

New implementation
BulkData New Technology - BDNT

Critical phase – early science operation

✔ Carful planning

Switch between old and new BD

New implementation
BulkData New Technology - BDNT

Critical phase – early science operation

✔ Carful planning

✔ Switch between old and new BD

New implementation
BulkData New Technology - BDNT

Critical phase – early science operation

- ✔ Carful planning
- ✔ Switch between old and new BD

New implementation
BulkData New Technology - BDNT

2011

Critical phase – early science operation

- ✔ Carful planning
- ✔ Switch between old and new BD

New implementation
BulkData New Technology - BDNT

2011

Critical phase – early science operation

- ✔ Carful planning
- ✔ Switch between old and new BD

New implementation
BulkData New Technology - BDNT

2011

Critical phase – early science operation

- ✔ Carful planning
- ✔ Switch between old and new BD

New implementation
BulkData New Technology - BDNT

2011

Critical phase – early science operation

- ✔ Carful planning
- ✔ Switch between old and new BD

March 2013

BDNT- development process

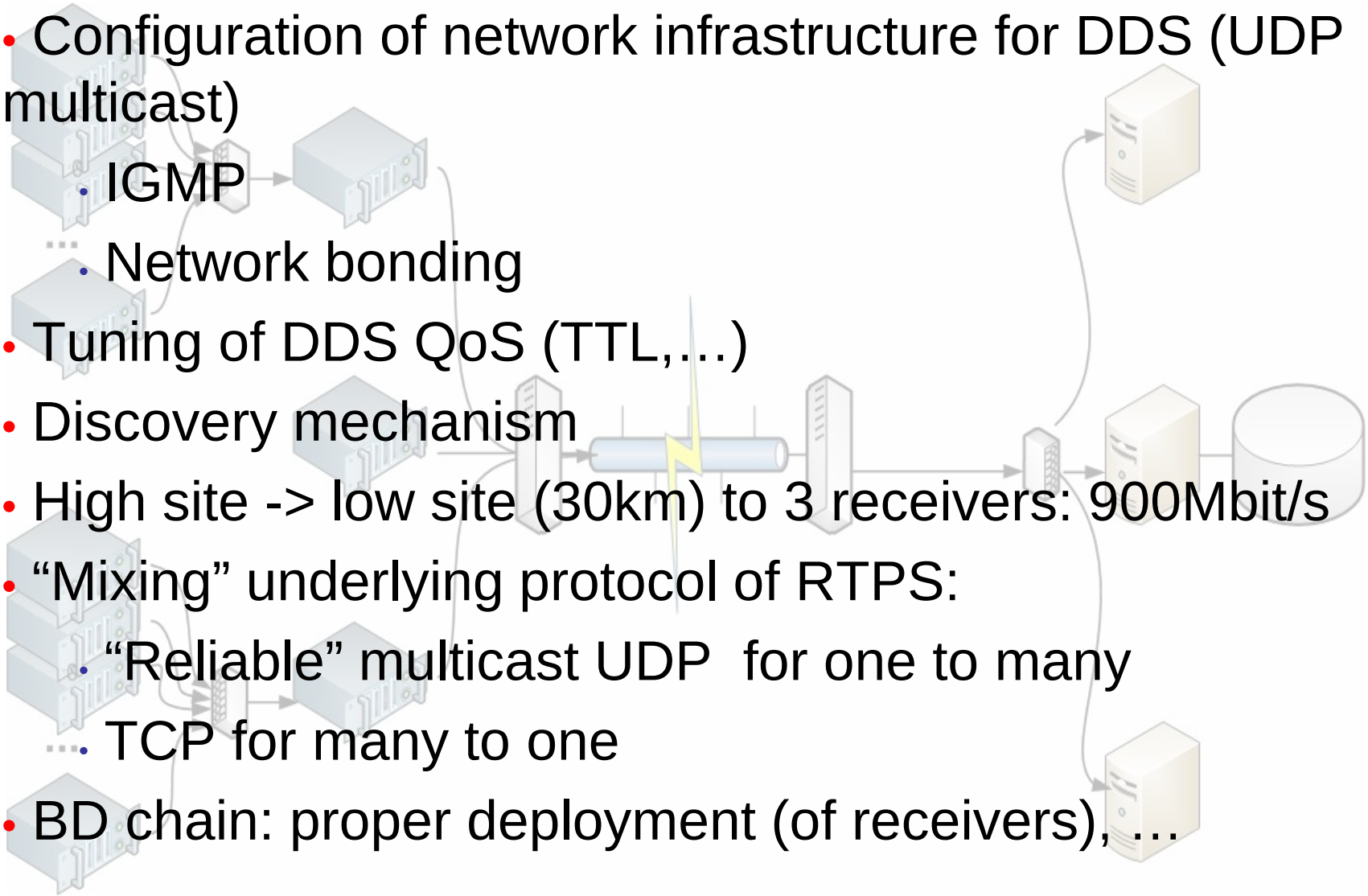
- Reviewed requirements
- Investigated possible technologies => **DDS**:
  - Well established standard
  - Several implementations w/ (commercial) support
  - Used in institutes/industry/military
  - Satisfy our requirements: reliable multicast, …
- Prototypes: **OpenDDS**, **RTI**, **OpenSplice** and **CoreDX**
- Selected specific implementation => **RTI DDS**:
  - Very stable/documented/supported
- Developed BDNT
- Ported applications: 6 sub-systems
- Tested in STEs and on site => **installed**

BDNT- benefits

- Improved error handling => improve stability & robustness
- Simpler and more intuitive API
- Reliable multicast instead of distributor & more streams per component => simplified deployment
- Built-in “support” for trouble shooting:
  - Early detection of problems: receiver processing throughput,…
  - Logs details control: configurable levels for detailed protocol diagnostic and statistics
- Generic sender & receiver => synthetic test

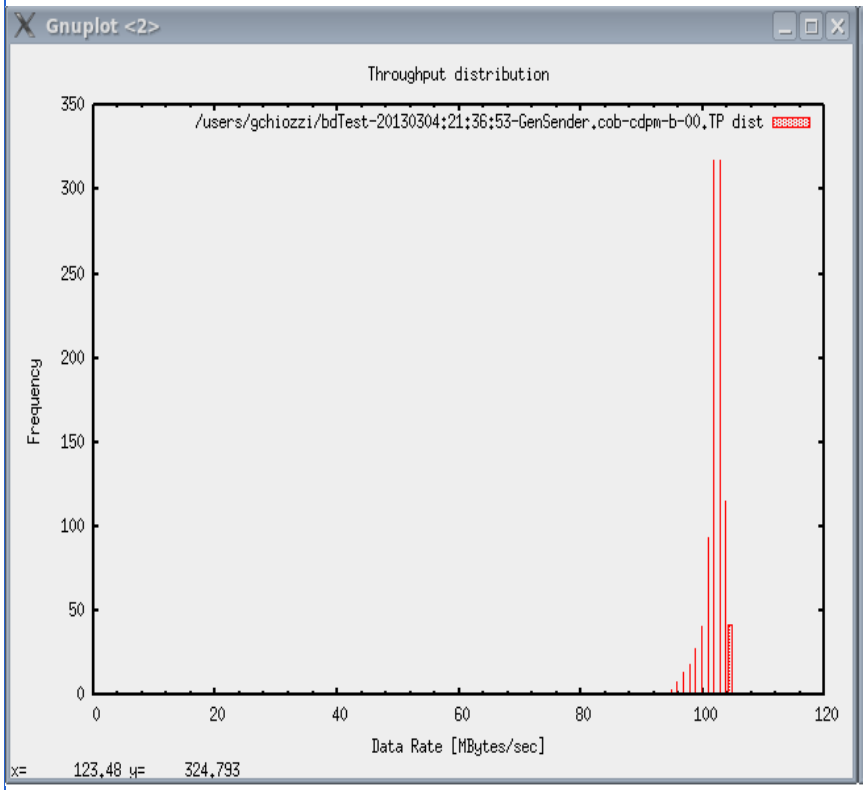
Challenges

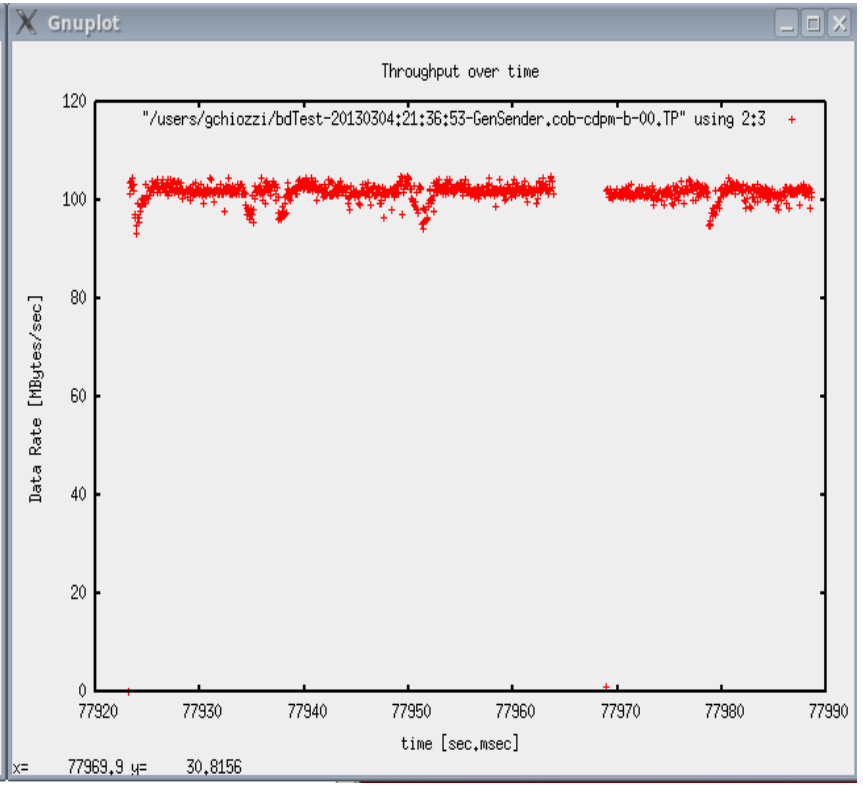
- Configuration of network infrastructure for DDS (UDP multicast)
  - IGMP
  - Network bonding
- Tuning of DDS QoS (TTL,…)
- Discovery mechanism
- High site -> low site (30km) to 3 receivers: 900Mbit/s
- “Mixing” underlying protocol of RTPS:
  - “Reliable” multicast UDP for one to many
  - TCP for many to one
- BD chain: proper deployment (of receivers), …



Synthetic test example

Problem multicasting from a machine with network bonding (2 NICs) to 3 machines with a single NIC





Problem solved by DDS QoS tuning

Conclusions

- In operation since March 2013
- Not yet tested at “full speed” in operation.
  Only with: 54 out of 66 antennas
  50% of the expected CORR data rate
- Big improvement in stability and availability of the whole SW: no need to restart the whole SW,…
- DDS is mature enough and the RTI implementation is very stable, well documented with good support
- Good understanding of what is going one in the whole BD chain, including applications, is necessary
- It is essential to have the right tools for troubleshooting

Questions