



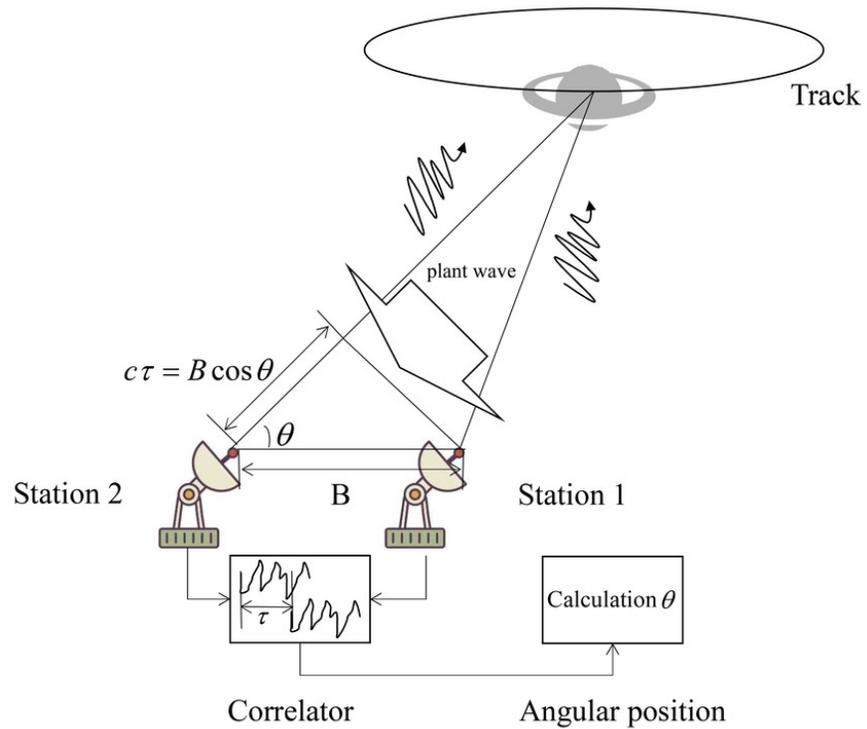
# Gestión de altos volúmenes de datos en observaciones VLBI

Javier González-García -  
Ingeniero soporte VLBI  
[j.gonzalez@oan.es](mailto:j.gonzalez@oan.es)

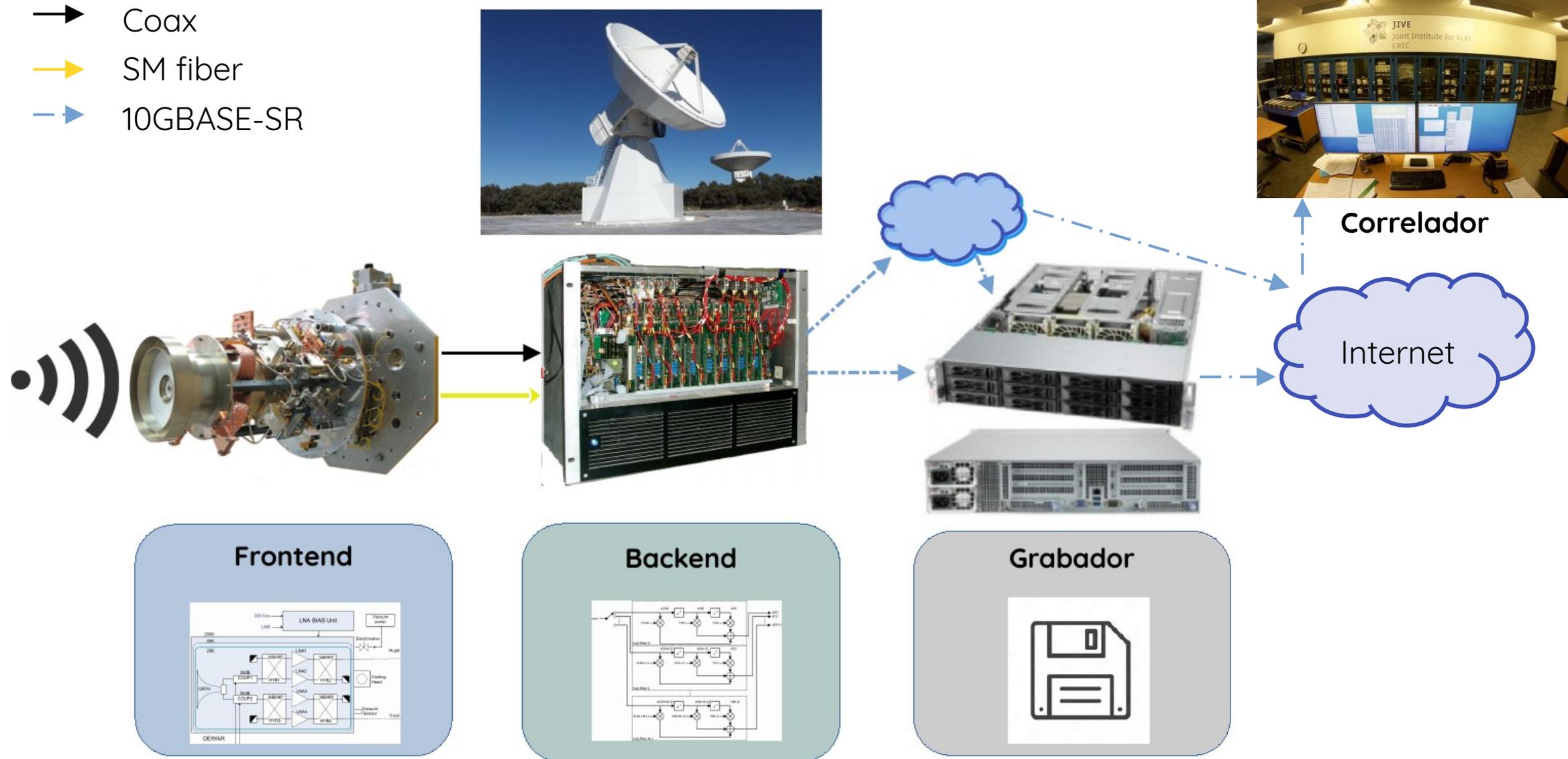


## Very Long **B**aseline Interferometry

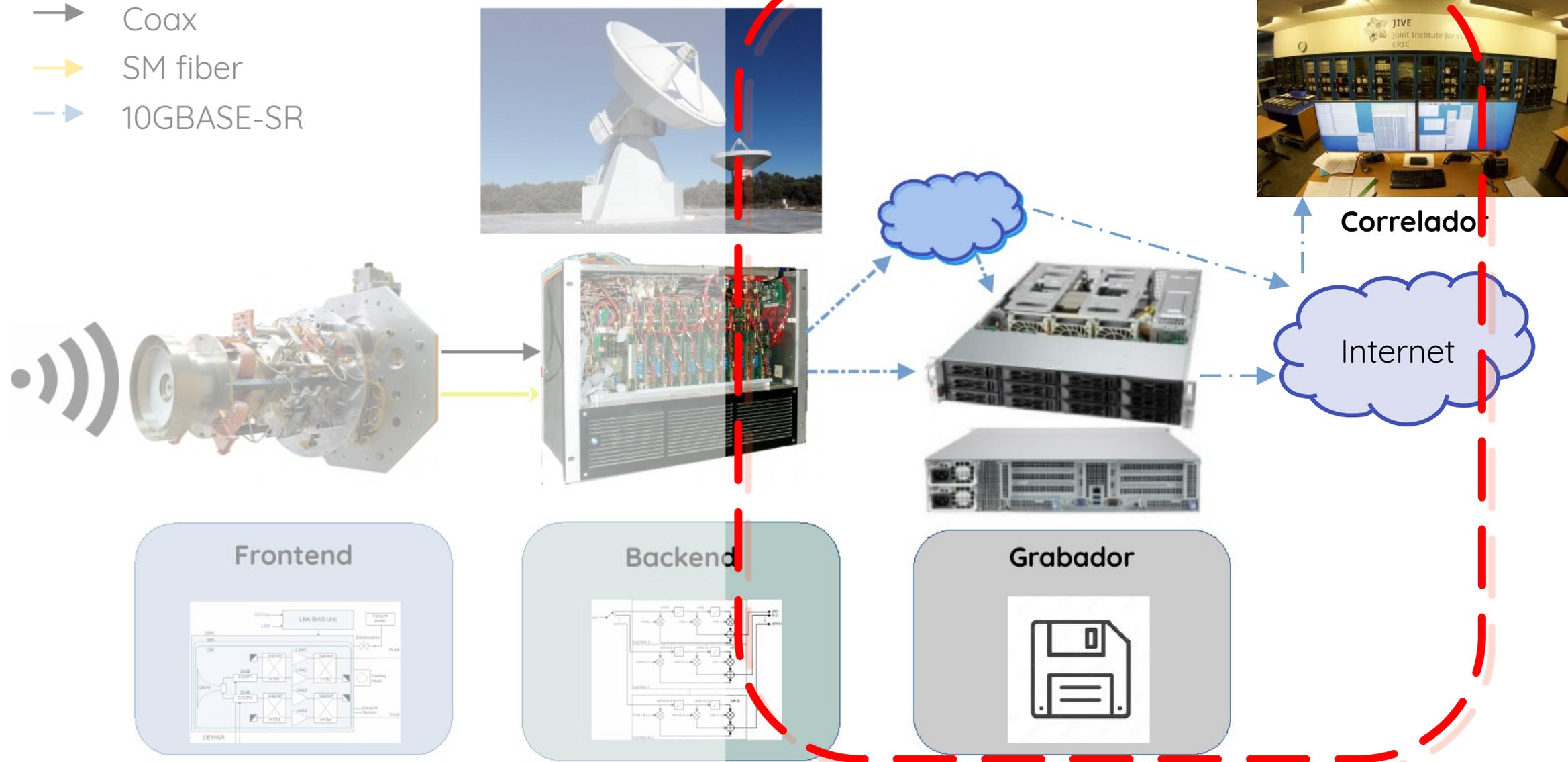
Técnica de interferometría radio NO conectada.



# Cadena de señal simplificada - VLBI



# Cadena de señal simplificada - VLBI

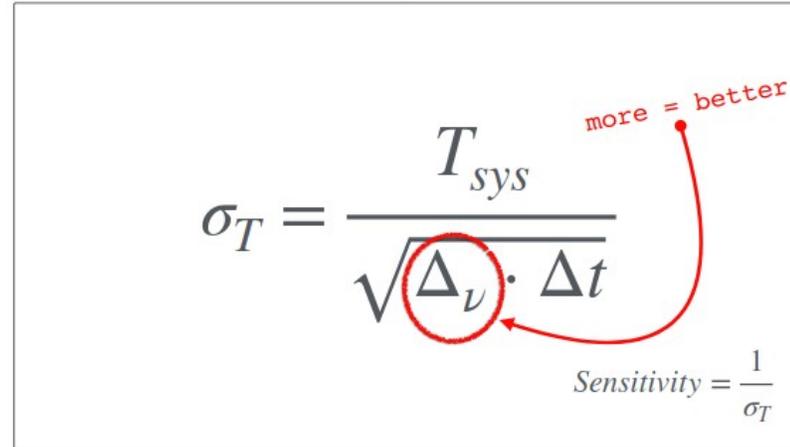


## Ecuación del radiómetro

$$\sigma_T = \frac{T_{sys}}{\sqrt{\Delta\nu \cdot \Delta t}}$$

*more = better*

*Sensitivity =  $\frac{1}{\sigma_T}$*



- 1) Aumentar el ancho de banda
- 2) Aumentar el tiempo de integración.

Ambas opciones suponen aumentar el número de bits a almacenar.

El modo de observación “más sensible” utilizado es: 1024 MHz x 2 pol x 2 Nyq x 2 bits = **8192 Mbps**

El tiempo de integración habitual es de 30 sec → 31 GB / scan. Tras 24 h de observación → **32 TB**

Ecuación del radiómetro



**Netflix UHD → 20 Mbps (equivalente a 4000 clientes!!)**

**Vídeo 4K durante 1 hora sin compresión ~ 22 GB**

- 1) Aumentar el ancho de banda
- 2) Aumentar el tiempo de integración.

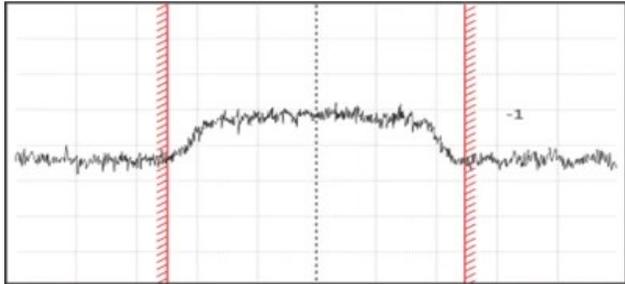
Ambas opciones suponen aumentar el número de bits a almacenar.

El modo de observación “más sensible” utilizado es:  $1024 \text{ MHz} \times 2 \text{ pol} \times 2 \text{ Nyq} \times 2 \text{ bits} = \mathbf{8192 \text{ Mbps}}$

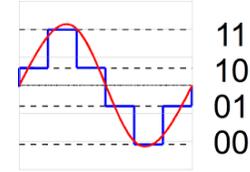
El tiempo de integración habitual es de 30 sec → 31 GB / scan. Tras 24 h de observación → **32 TB**

# Requerimiento alto ancho de banda

4 x 256 MHz x 2 pols = 2 GHz BW

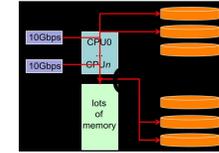


x2 muestreo = 2048 Msps  
 x2 bits/muestra = 8192 Gbps



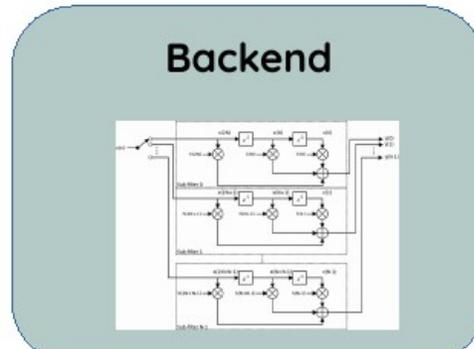
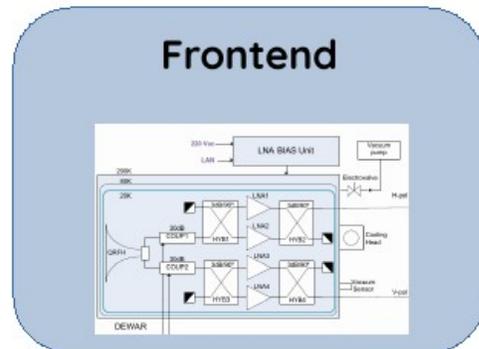
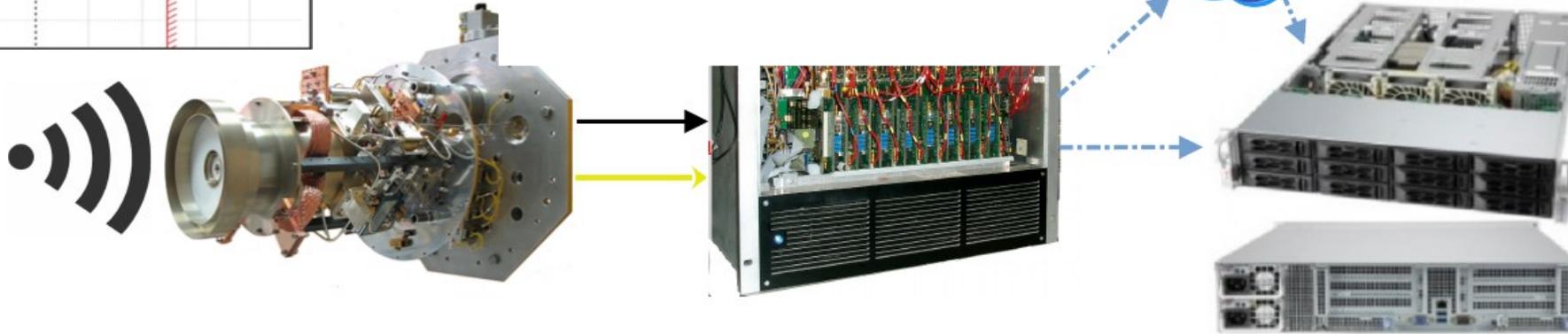
VDIF packet = 8224 B  
 32 B (header)  
 8192 B (payload)

Preamble	Destination MAC Address	Source MAC Address	Type	Data	CRC
8 bytes	6 bytes	6 bytes	2 bytes	46-1500 bytes	4 bytes



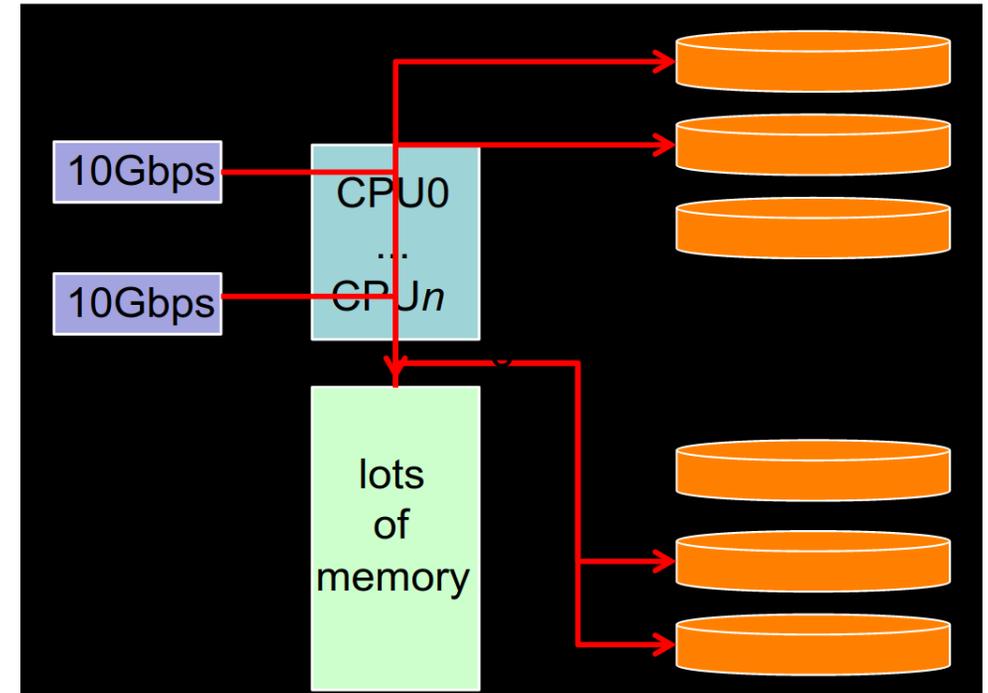
8 192 Gbps

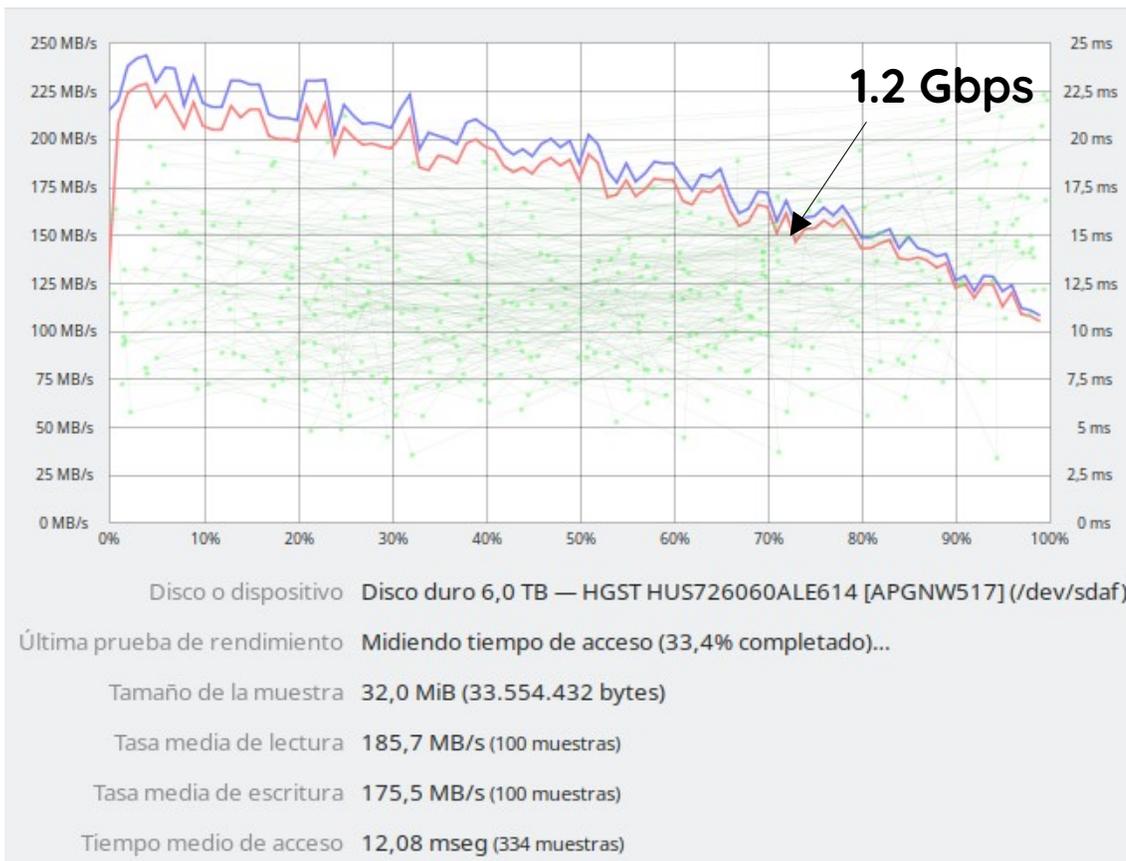
128 500 pkt/sec



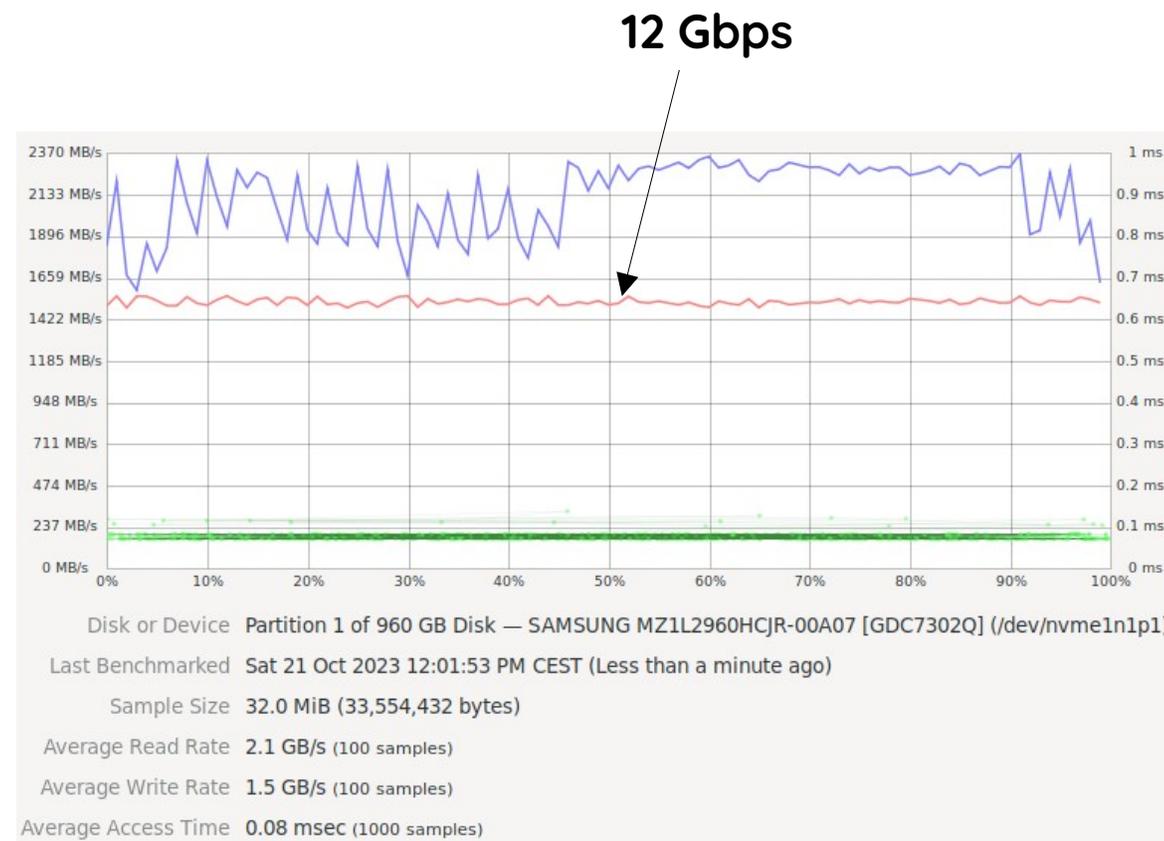
- Una observación a 8 Gbps de 24 h de duración requiere típicamente de ~32 TB (duty cycle de 36%, ~ 1000 scans de 30 sec)
- HDD actuales hasta 20 TB, Seagate anunció HDD de 32 TB para 3T2023 y 40 TB para 2024 en su roadmap. No es factible grabar a 8 Gbps en un HDD. La previsión en VLBI es llegar a tasas de **16** y **32 Gbps** en un futuro cercano.

- Una observación a 8 Gbps de 24 h de duración requiere típicamente de ~32 TB (duty cycle de 36%, ~ 1000 scans de 30 sec)
- HDD actuales hasta 20 TB, Seagate anunció HDD de 32 TB para 3T2023 y 40 TB para 2024 en su roadmap. No es factible grabar a 8 Gbps en un HDD. La previsión en VLBI es llegar a tasas de **16 y 32 Gbps** en un futuro cercano.
- La solución reside en utilizar **múltiples discos de forma simultánea y sacrificar la redundancia.**





## HDD SATA3 6TB



## SSD NVME 1TB

# Selección soporte almacenamiento



	HDD - SATA 3	SDD - SATA 3	SSD-NVME
Tasa escritura*	180 MB/s	550 MB/s	<b>5000 a 7300 MB/s</b>
\$/TB media [min - max]	<b>28 [14-66]</b>	85 [75-99]	82 [72-96] **
Tamaño máximo disponible (coste)	22 TB (550\$) WD Gold Enterprise Class	8TB (1k\$) Kingston DC600M	<b>30TB (5k\$)</b> <b>Micron 9400 SSD</b>
Capacidad Mejor \$/TB	8 TB (112\$)	4 TB (300\$)	4 TB (288\$)
Unidades necesarias (128 TB - 4 sesiones)	<b>16</b>	32	32
Coste total	<b>1792 \$</b>	9600 \$	9216 \$

\* Valor medio aproximado. La tasa real depende de múltiples factores.

\*\* Evaluadas capacidades solo de 1 a 8 TB

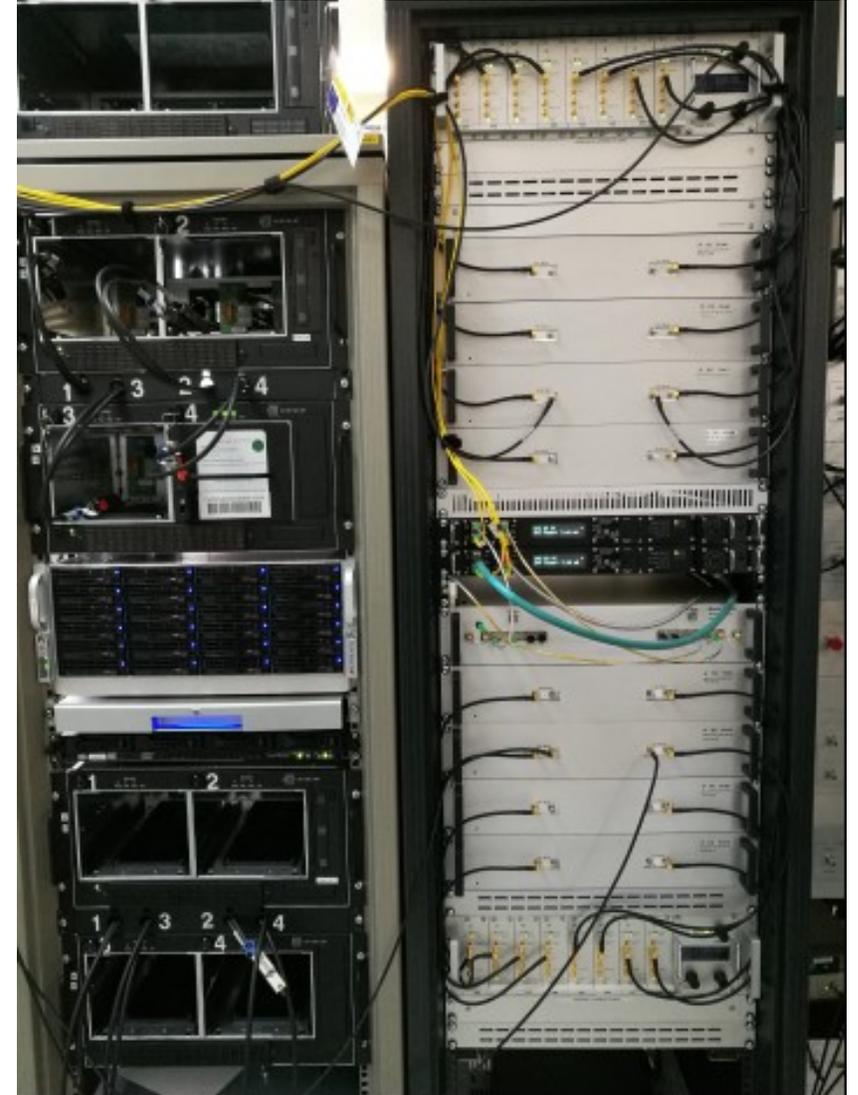
Fuente: Amazon.com (a través de diskprices.com), en 2023

## Mark6 → Conduant/MIT Haystack

- Solución propietaria
- ~ 30 k€ con 4x80 TB
- Configuración punto-punto fija (No implementa capas superiores IP) 😞
- Discos extraíbles

## Flexbuff → EVN/Joint Institute for VLBI in Europe (JIVE)

- Basado en hardware genérico COTS
- Software de código abierto multipropósito:  
<https://github.com/jive-vlbi/jive5ab>
- Totalmente configurable
- ~ 15 k€ con 36x10 TB
- Discos fijos

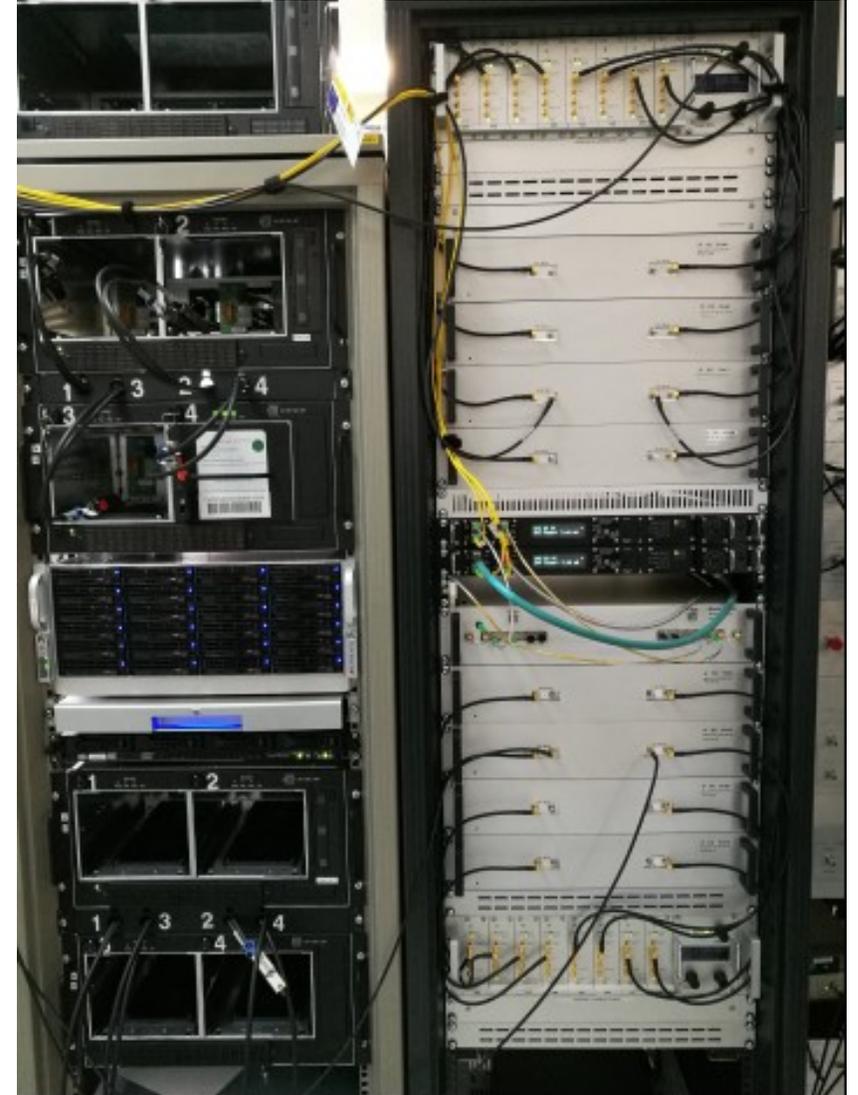


## Mark6 → Conduant/MIT Haystack

- Solución propietaria
- ~ 30 k€ con 4x80 TB
- Configuración punto-punto fija (No implementa capas superiores IP) 😞
- Discos extraíbles

## Flexbuff → EVN/Joint Institute for VLBI in Europe (JIVE)

- Basado en hardware genérico COTS
- Software de código abierto multipropósito:  
<https://github.com/jive-vlbi/jive5ab>
- Totalmente configurable
- ~ 15 k€ con 36x10 TB
- Discos fijos



El tráfico generado por el backend es UDP:

- No orientado a conexión
- Entrega no garantizada.
- Alta velocidad.
- Disponible broadcast y multicast.

Características del tráfico:

- Unidireccional.
- Tasa constante y mantenida.
- Pérdida de paquetes tolerable <0.1%
- Cuasi tiempo real.

Importante fijar el MTU en 9000 para evitar fragmentación (recordar paquete VDIF 8kB)

Pila TCP/IP montada en FPGA sobre silicio.

Tecnología 10GBASE-SR.



## Key Features

1. 4U chassis
2. 36 x 3.5" hot-swap HDD bays
3. 12-port 2U SAS 6Gbps direct-attached backplane, support up to 12x 3.5-inch SAS2/SATA3 HDD/SSD
4. 24-port 4U SAS 6Gbps direct-attached backplane, support up to 24x 3.5-inch SAS2/SATA3 HDD/SSD
5. 1U 1280W Redundant Platinum 95%+ Super Quiet Power Supply with PMbus
6. 7 low-profile expansion slot(s)
7. 7 x 8cm hot-swap redundant PWM cooling fan(s)



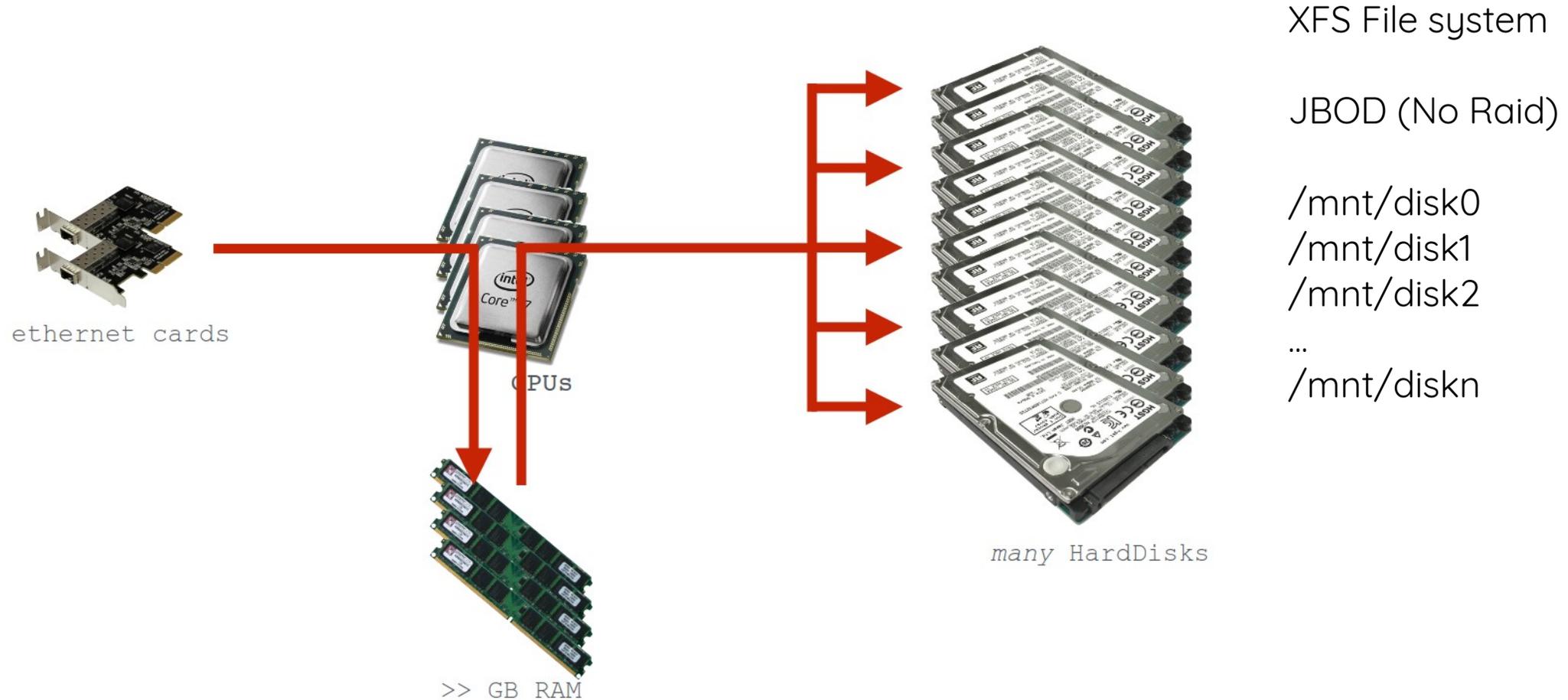
7 PCI-E 3.0 Slots

## Key Features

1. Intel® Xeon® Scalable Processors., Dual Socket P (LGA 3647)
2. Intel® C622
3. Up to 2TB 3DS ECC RDIMM, DDR4-2666MHz; Up to 2TB 3DS ECC LRDIMM, in 16 DIMM slots
4. 3 PCI-E 3.0 x16, 4 PCI-E 3.0 x8  
M.2 Interface: 2 PCI-E 3.0 x4  
M.2 Form Factor: 22110  
M.2 Key: M-Key (RAID 0,1 support)
5. 1 VGA port
6. Intel® C622 controller for 10 SATA3 (6 Gbps) ports; RAID 0,1,5,10
7. Dual LAN with 10GBase-T with Intel® X55

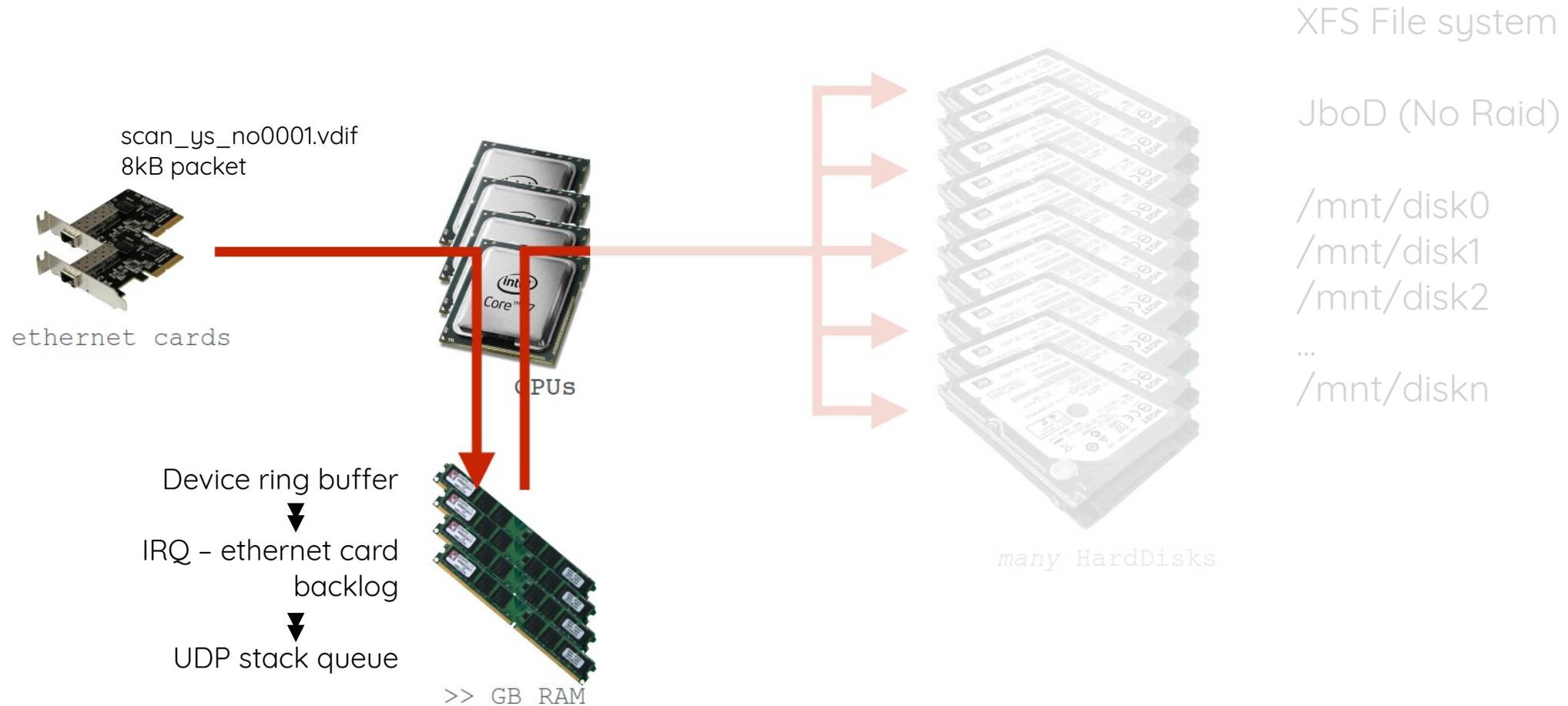
¿Cómo alcanzar tasas de registro de 8, 16 hasta 32 Gbps?

Solución: HW de computación genérico con **varias CPU**, varias interfaces de red y **múltiples discos** → High-speed packet recorder (Chia blockchain)



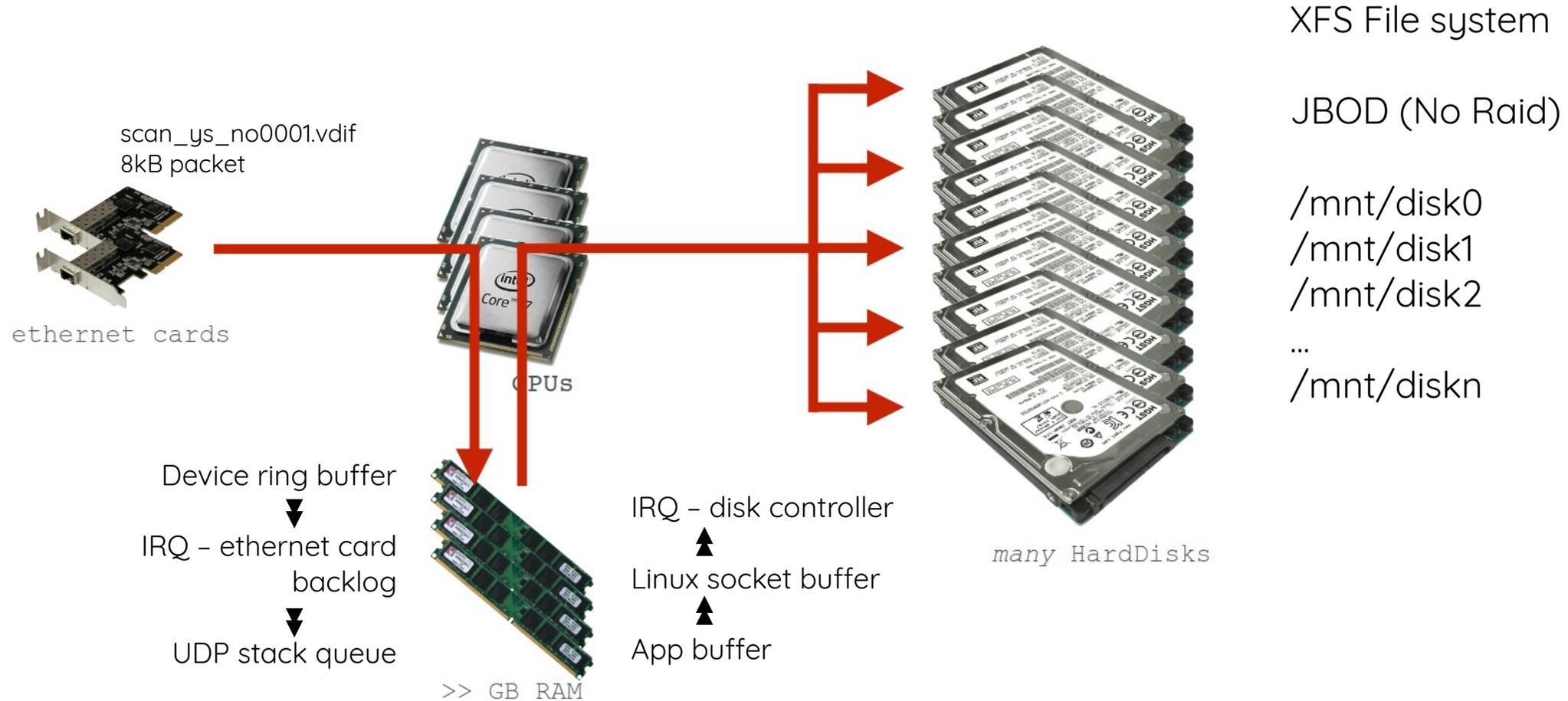
¿Cómo alcanzar tasas de registro de 8, 16 hasta 32 Gbps?

Solución: HW de computación genérico con **varias CPU**, varias interfaces de red y **múltiples discos** → High-speed packet recorder (Chia blockchain)



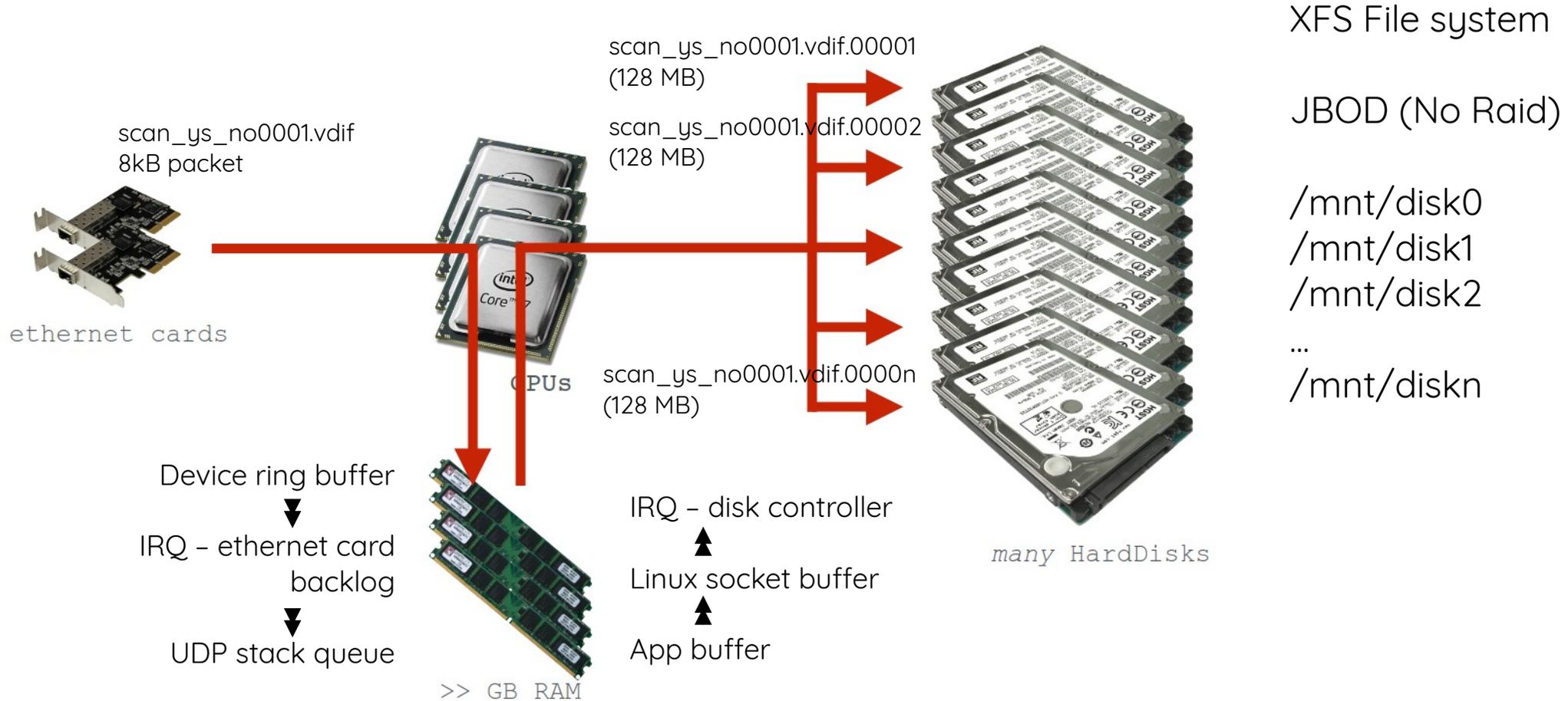
¿Cómo alcanzar tasas de registro de 8, 16 hasta 32 Gbps?

Solución: HW de computación genérico con **varias CPU**, varias interfaces de red y **múltiples discos** → High-speed packet recorder (Chia blockchain)



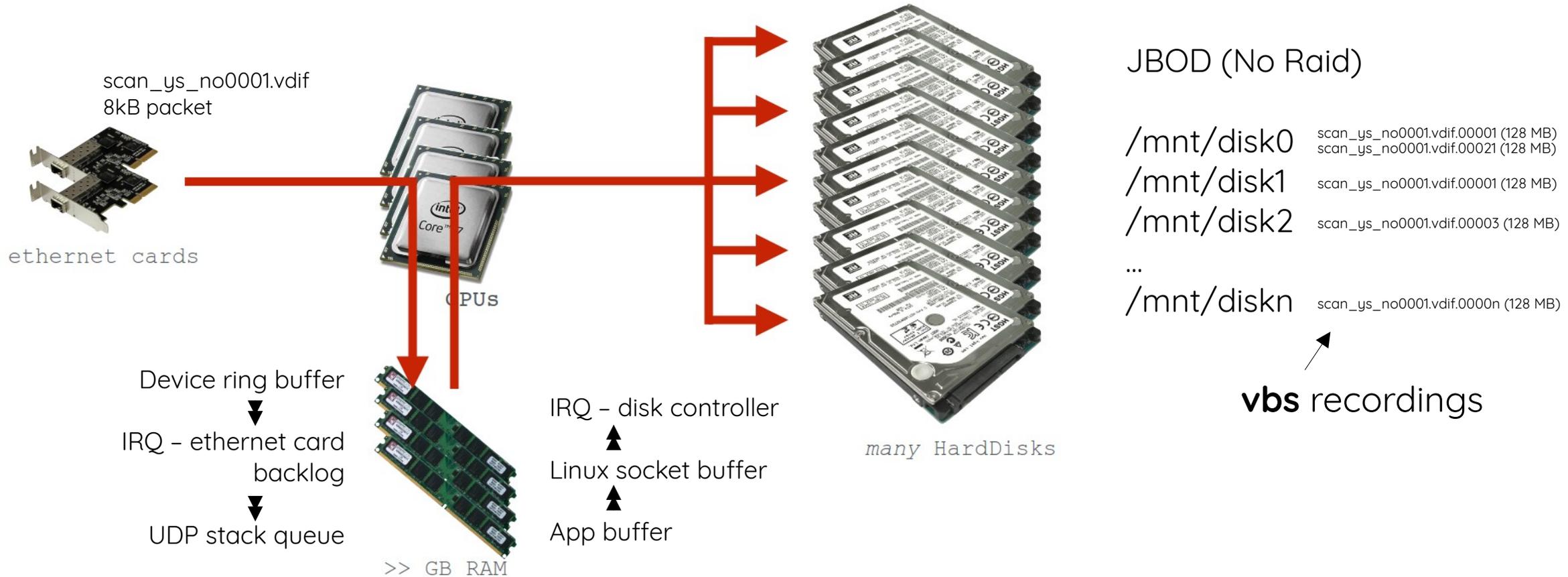
¿Cómo alcanzar tasas de registro de 8, 16 hasta 32 Gbps?

Solución: HW de computación genérico con **varias CPU**, varias interfaces de red y **múltiples discos** → High-speed packet recorder (Chia blockchain)



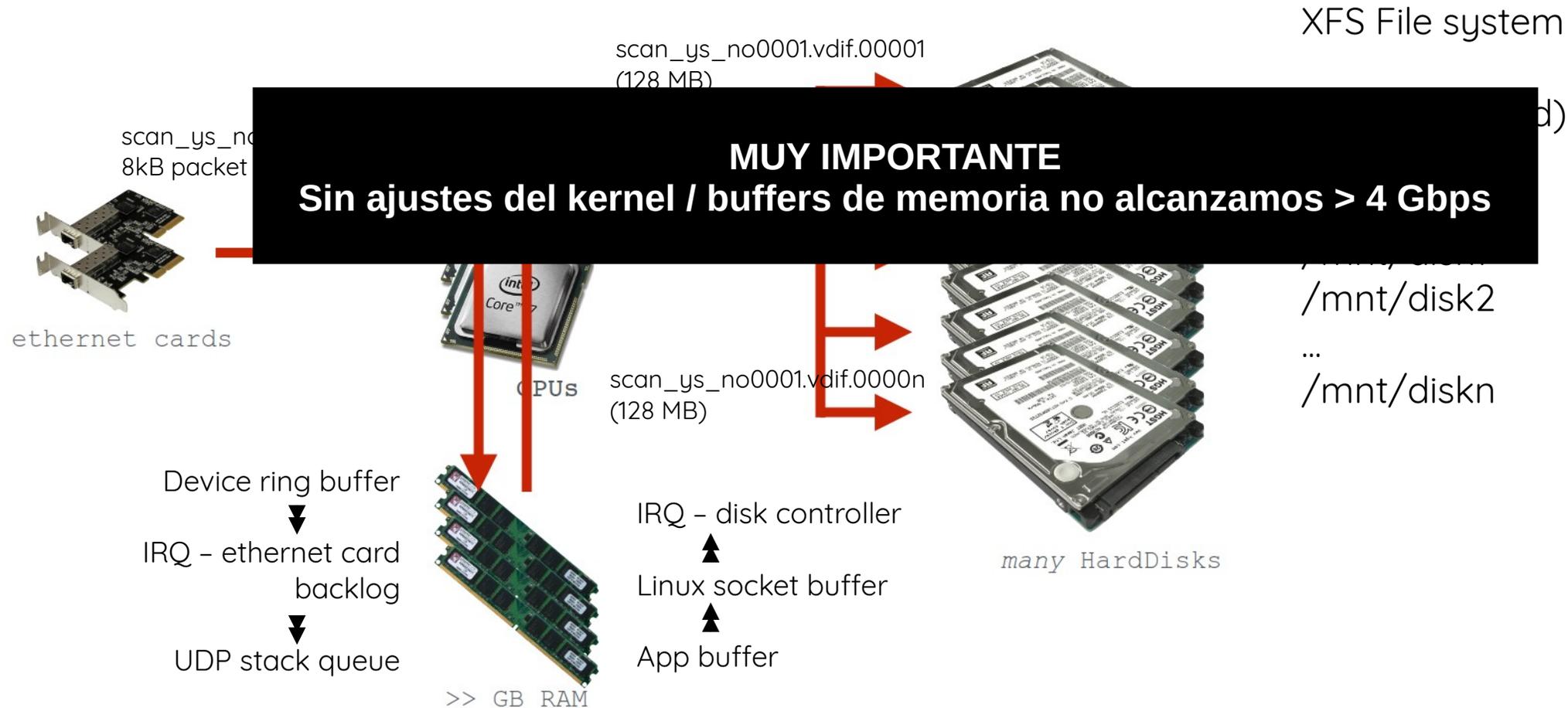
¿Cómo alcanzar tasas de registro de 8, 16 hasta 32 Gbps?

Solución: HW de computación genérico con **varias CPU**, varias interfaces de red y **múltiples discos** → High-speed packet recorder (Chia blockchain)



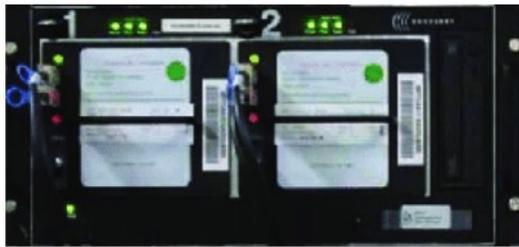
¿Cómo alcanzar tasas de registro de 8, 16 hasta 32 Gbps?

Solución: HW de computación genérico con **varias CPU**, varias interfaces de red y **múltiples discos** → High-speed packet recorder (Chia blockchain)



Siguiente paso: Entregar datos al correlador

- ✓ Fácil
- ✓ No requiere doblar la capacidad de almacenamiento

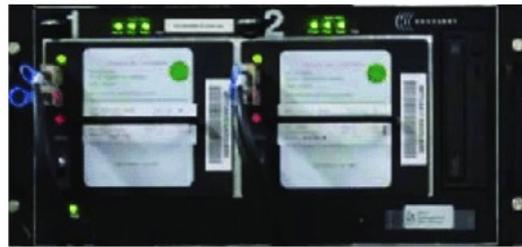




Fácil

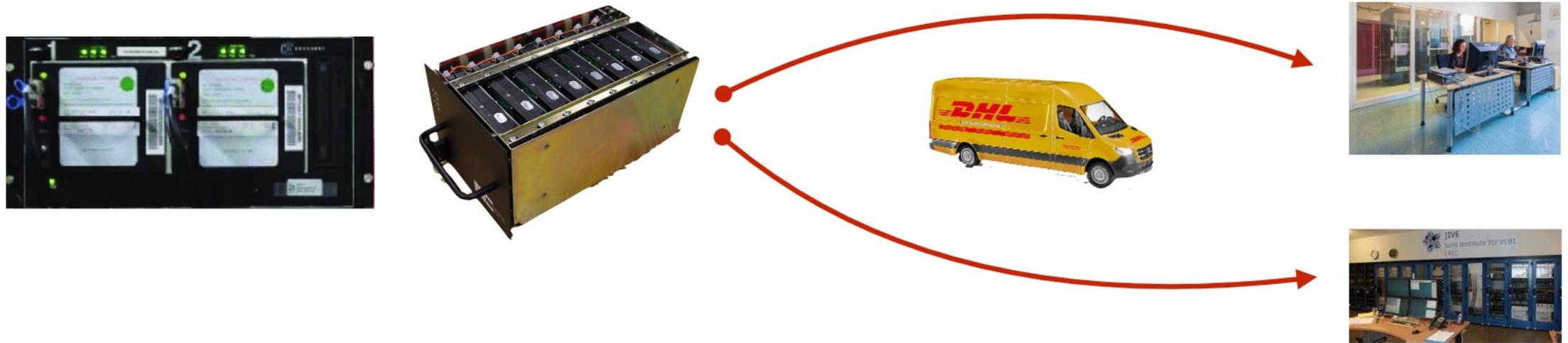


No requiere doblar la capacidad de almacenamiento



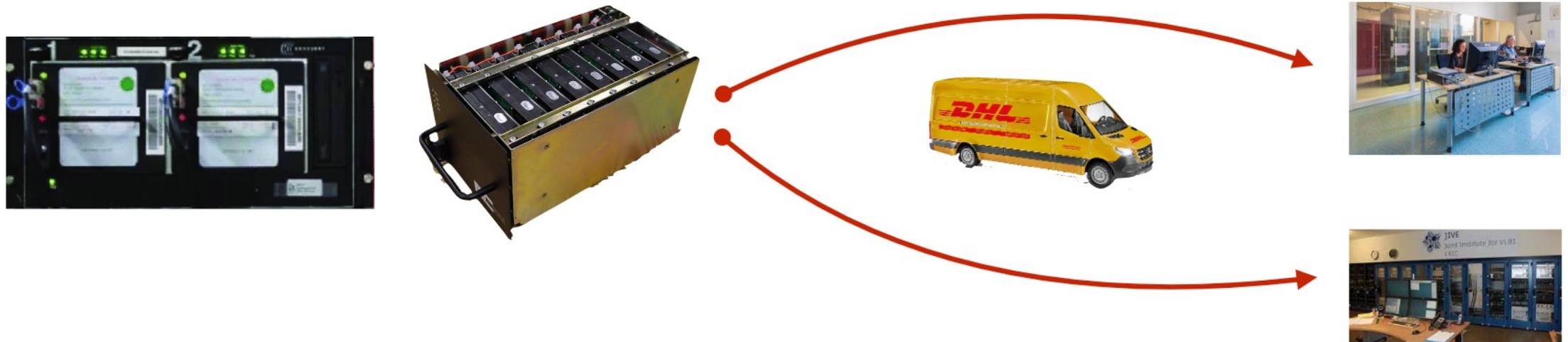
La estación pierde espacio de almacenamiento

- ✓ Fácil
- ✓ No requiere doblar la capacidad de almacenamiento



- ✗ La estación pierde espacio de almacenamiento
- ✗ Pérdida de control durante el tránsito

- ✓ Fácil
- ✓ No requiere doblar la capacidad de almacenamiento



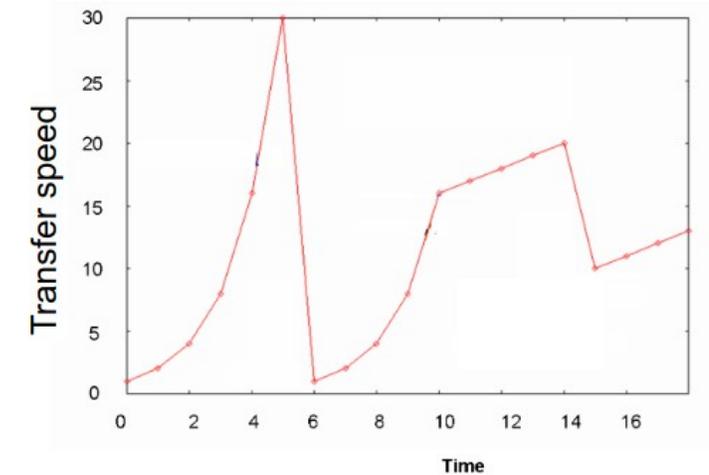
- ✗ La estación pierde espacio de almacenamiento
- ✗ Pérdida de control durante el tránsito
- ✗ Costoso



Las mismas reglas básicas que la transferencia entre backend y grabador aplican, salvo el requerimiento de cuasi tiempo real. La correlación puede retrasarse hasta recibir todos los datos.

Selección del protocolo de transporte:

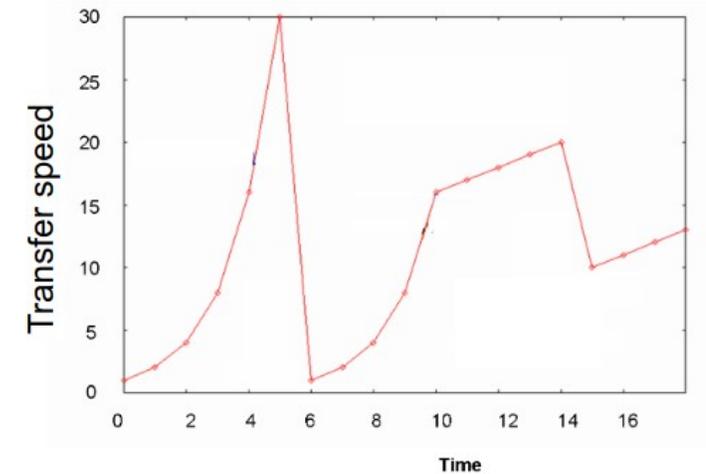
- TCP → Errático. Muere en distancias largas: scp, ftp, rsync



Las mismas reglas básicas que la transferencia entre backend y grabador aplican, salvo el requerimiento de cuasi tiempo real. La correlación puede retrasarse hasta recibir todos los datos.

Selección del protocolo de transporte:

- ~~TCP → Errático. Muere en distancias largas: scp, ftp, rsync~~



- UDP → Entrega no asegurada, sin control de flujo ni errores.

UDP es **acceptable** para datos VLBI pero es conveniente tener un cierto control de pérdidas → Solución:

Implementar una ligera capa de control sobre UDP: UDT

UDT es un protocolo de **capa de aplicación**.

Premio Supercomputing Bandwidth Challenge 2006, 2008, 2009

Implementa sobre UDP:

- Control de sesión.
- Control de congestión.
- Fiabilidad.

Take-away message:

*“For distributed data intensive applications in high speed networks. A small number of flows share the abundant bandwidth”*

[udt.sourceforge.net](http://udt.sourceforge.net)

¿Qué herramientas tenemos para transferir grandes volúmenes de datos usando UDT?

El propio *jive5ab* permite realizar transferencias usando UDT como cliente y servidor. Incluye script *m5copy* como CLI para manejar las transferencias. - Puede trabajar con ficheros VBS tanto en origen como en destino.

<https://github.com/jive-vlbi/jive5ab>

*Etransfer (etc/etd)* es otra herramienta que implementa UDT. Requiere un sistema FUSE si se trabaja con ficheros fragmentados VBS

<https://github.com/jive-vlbi/ettransfer>

Otra herramientas fuera de la comunidad:

Tsunami UDP, IBM© Aspera©, Data Expedition Inc, <https://github.com/oghazal/udtcats>



