

DAVIDE MICHELINO



THE ITALIAN
EDUCATION
& RESEARCH
NETWORK

SCOReS: Studio di sistemi di Caching per l'utilizzo Ottimizzato di Risorse opportunistiche e siti senza pledged storage per applicazioni di e-Science



GIORNATA DI INCONTRO
BORSE DI STUDIO GARR
"ORIO CARLINI"
6 DICEMBRE 2018 ROMA

Roma, 06/12/2018

Borsisti Day 2018





Il progetto SCORes: obiettivi

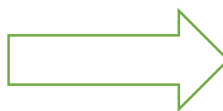
- **Studiare, progettare e mettere in esercizio** sistemi di Cache per lo stoccaggio temporaneo di data-set scientifici.
 - Sorgenti Multiple
 - Distribuiti geograficamente
 - Disponibili con interfacce cloud o Grid
- A supporto di siti e **infrastrutture sulla rete GARR**
- Utilizzo tecnologie standard
 - **HTTP**
 - Introduzione invisibile



Motivazioni ed opportunità

Indagine argomenti di tendenza in

- Comunità scientifiche
- Workshop/conferenze
- Working Group



Riduzione costi di:

- Manutenzione
- Forza Lavoro
- Requisiti di affidabilità

Dove poter intervenire?

- Storage
- Network



Tecnologie di cache da inserire nei workflow applicativi e nei computing model degli esperimenti



Limiti delle tecnologie di cache HTTP esistenti

Difficoltà di inserirle nel contesto dei computing model degli esperimenti:

- Mancanza supporto certificati X509/VOMS
- Gap utilizzo HTTP nel Data Analysis

SQUID[®]

NGINX

Nuovi tool per soluzioni alternative:

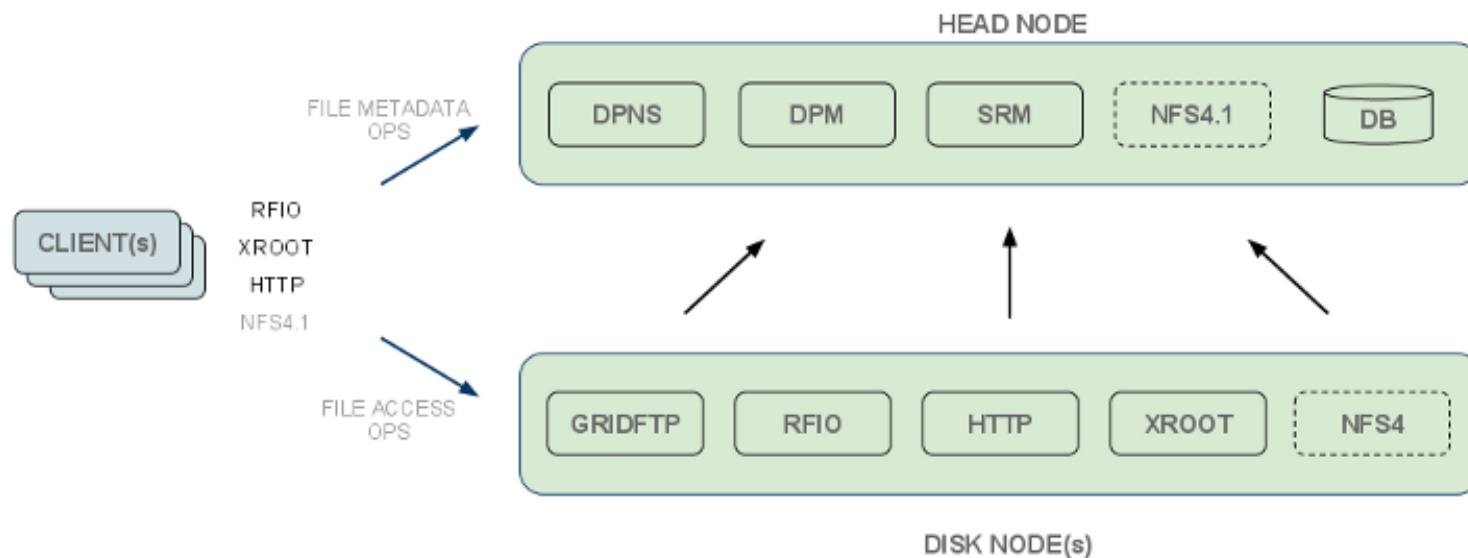
- Xcache
- DPM Volatile Pool
- Dynafed

DPM
Disk Pool Manager



Le tecnologie: DPM

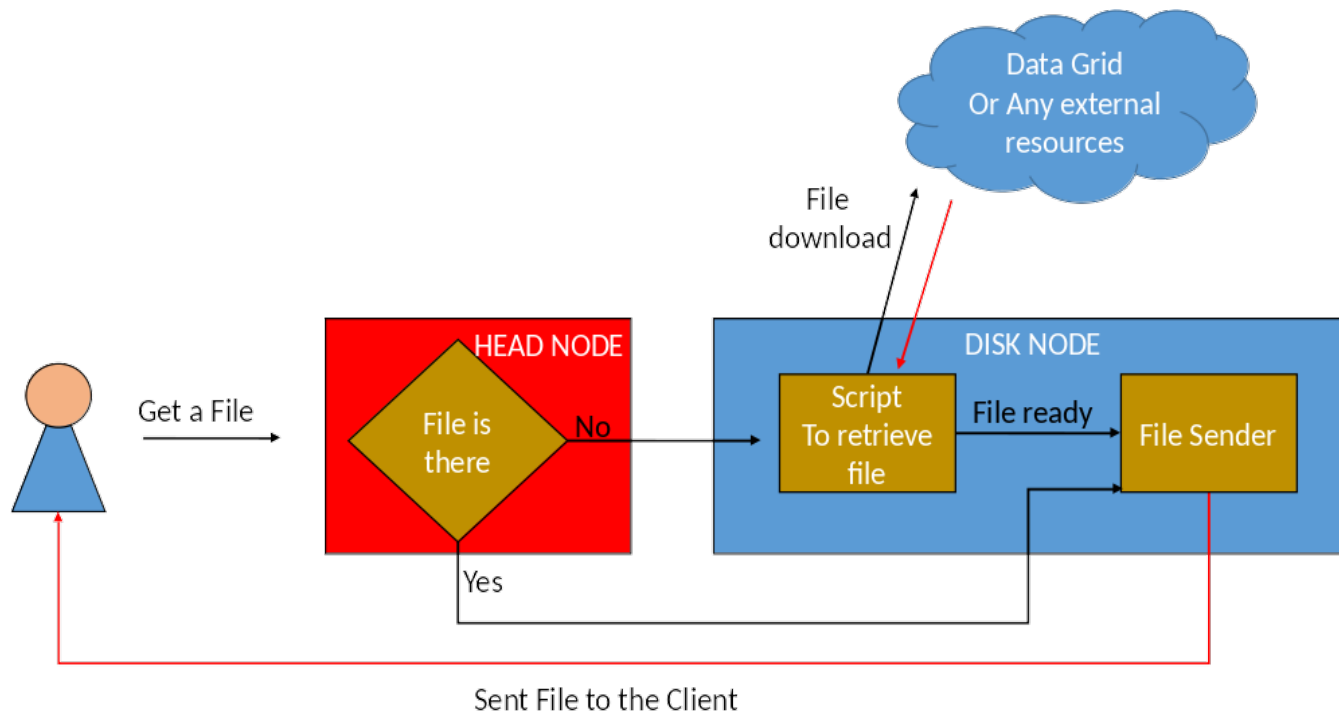
- Tecnologia di storage **largamente diffusa**
- In produzione con installazioni **nell'ordine dei PB**
- Aggrega dischi in **pool**
- Supporta **HTTP/X509/VOMS**





Le tecnologie: Pool Volatili di DPM

- Speciale pool che può recuperare file da **fonti esterne**
- Script personalizzati per STAT e GET

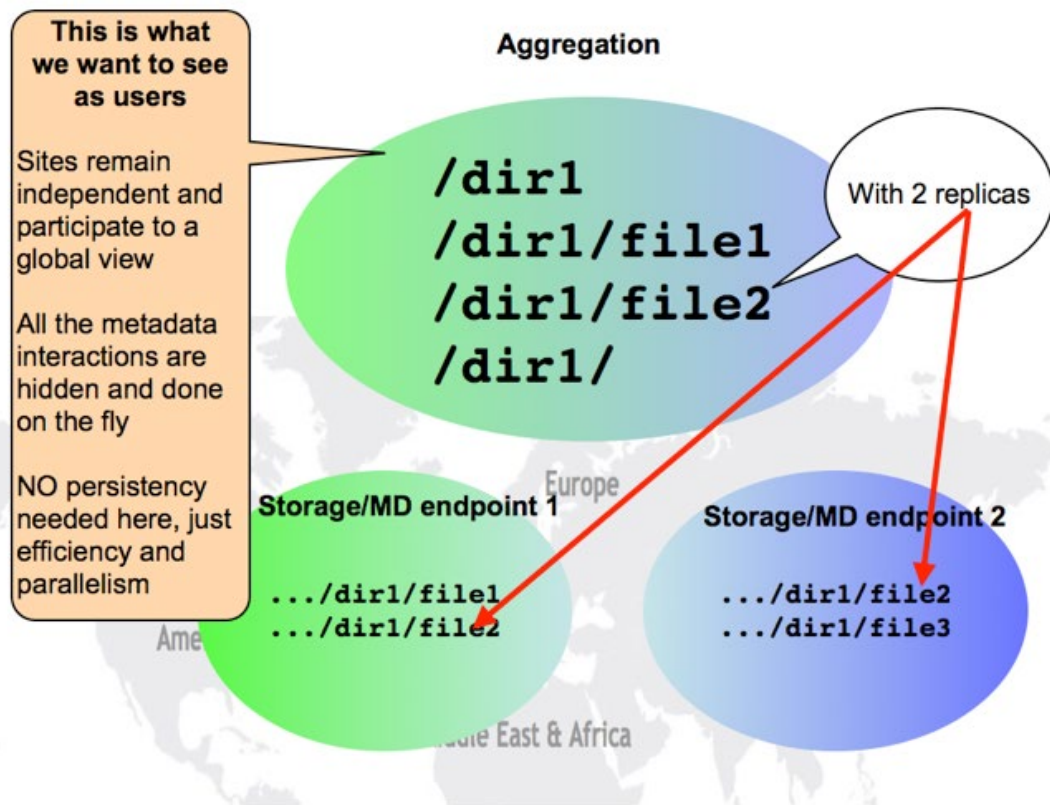




Le tecnologie: il federatore DynaFed

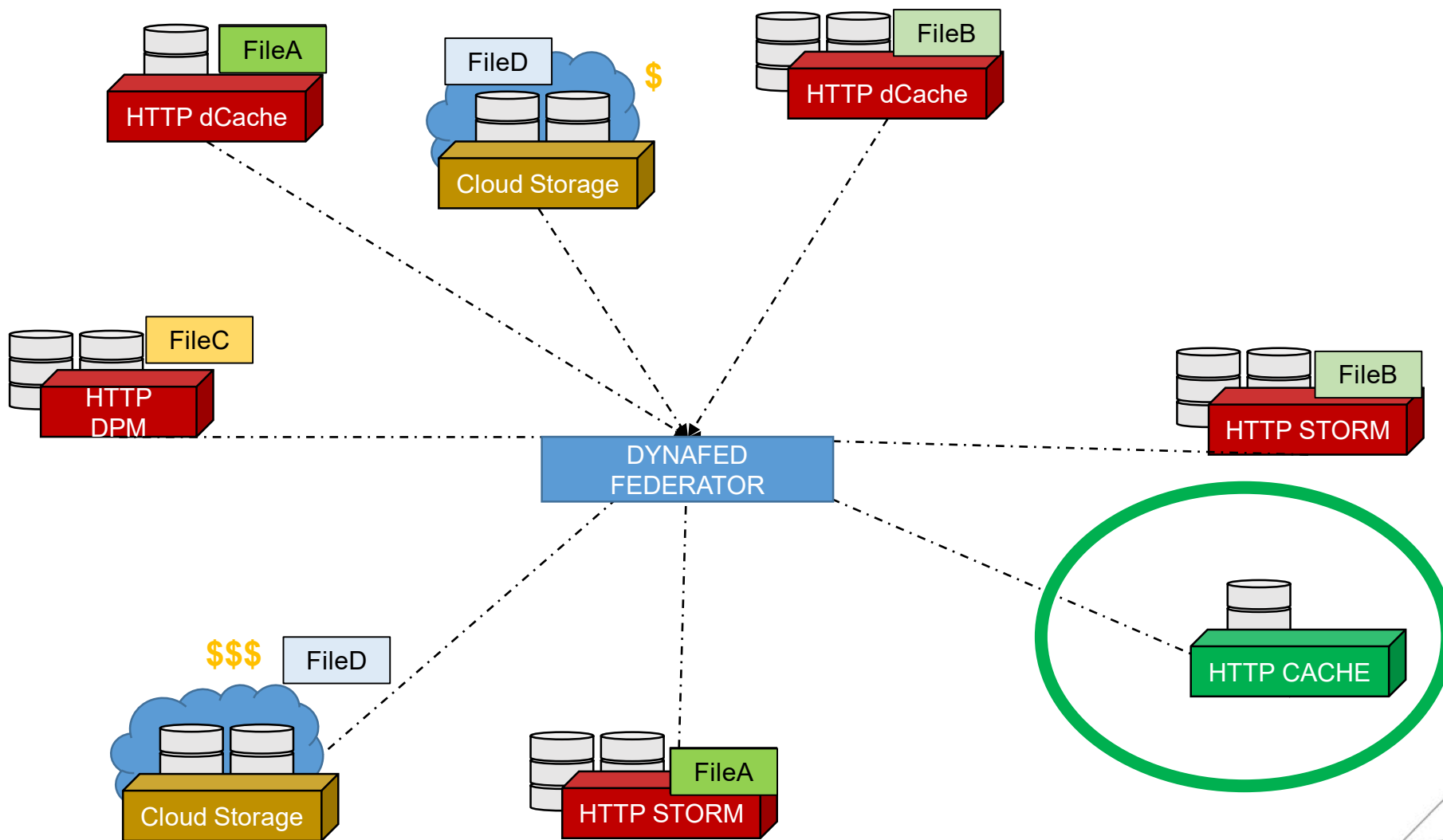
Prodotto opensource sviluppato dal CERN che aggrega endpoint di HTTP, WEBDAV e S3:

- Albero unificato
- Ogni file è un **metalink** con più repliche
- Caching dei metadati
- **Geolocalizzazione** client
- Supporta X.509 e VOMS



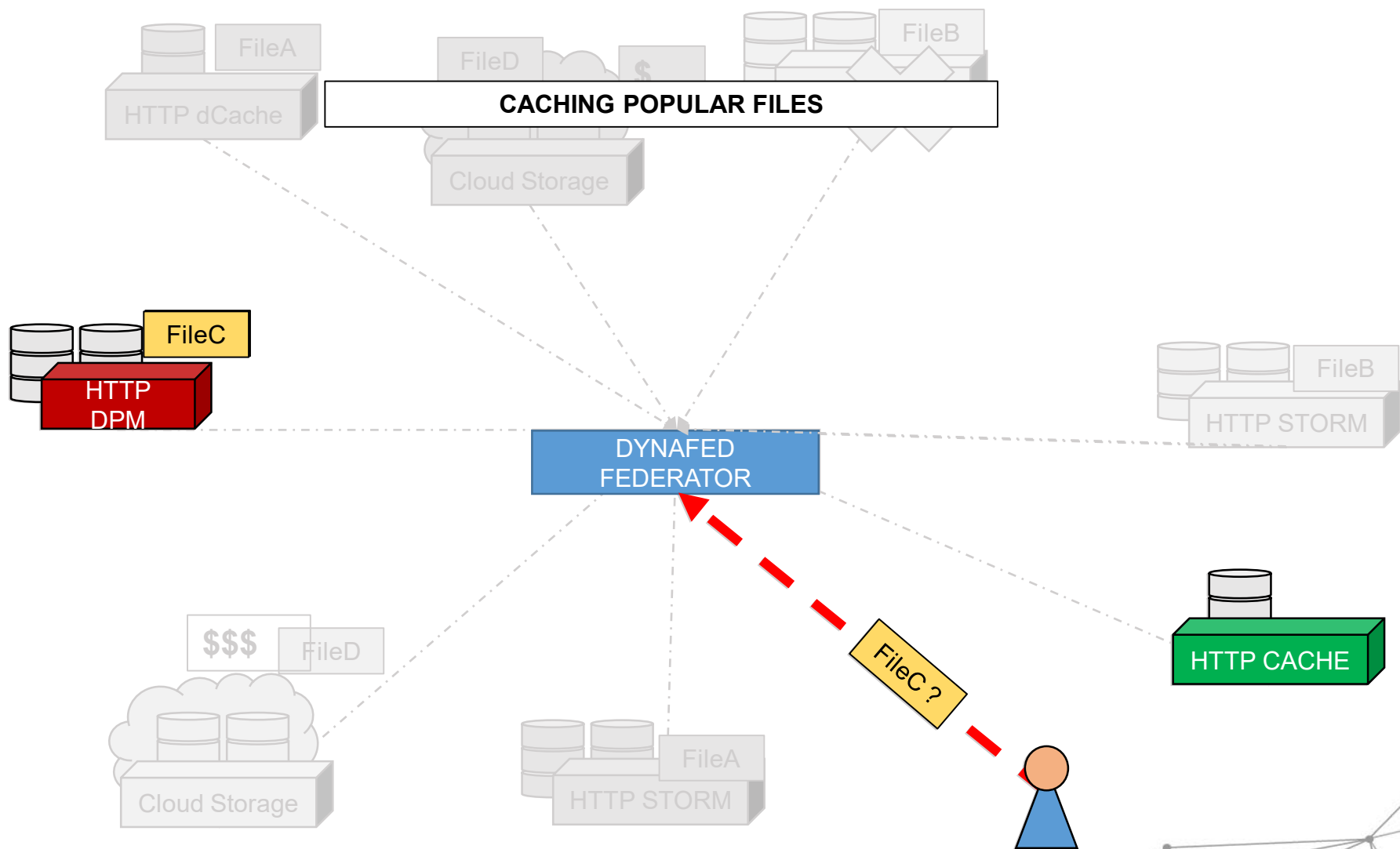


GIORNATA DI INCONTRO BORSE DI STUDIO GARR "ORIO CARLINI"
MARTEDÌ 6 DICEMBRE 2018 - ROMA



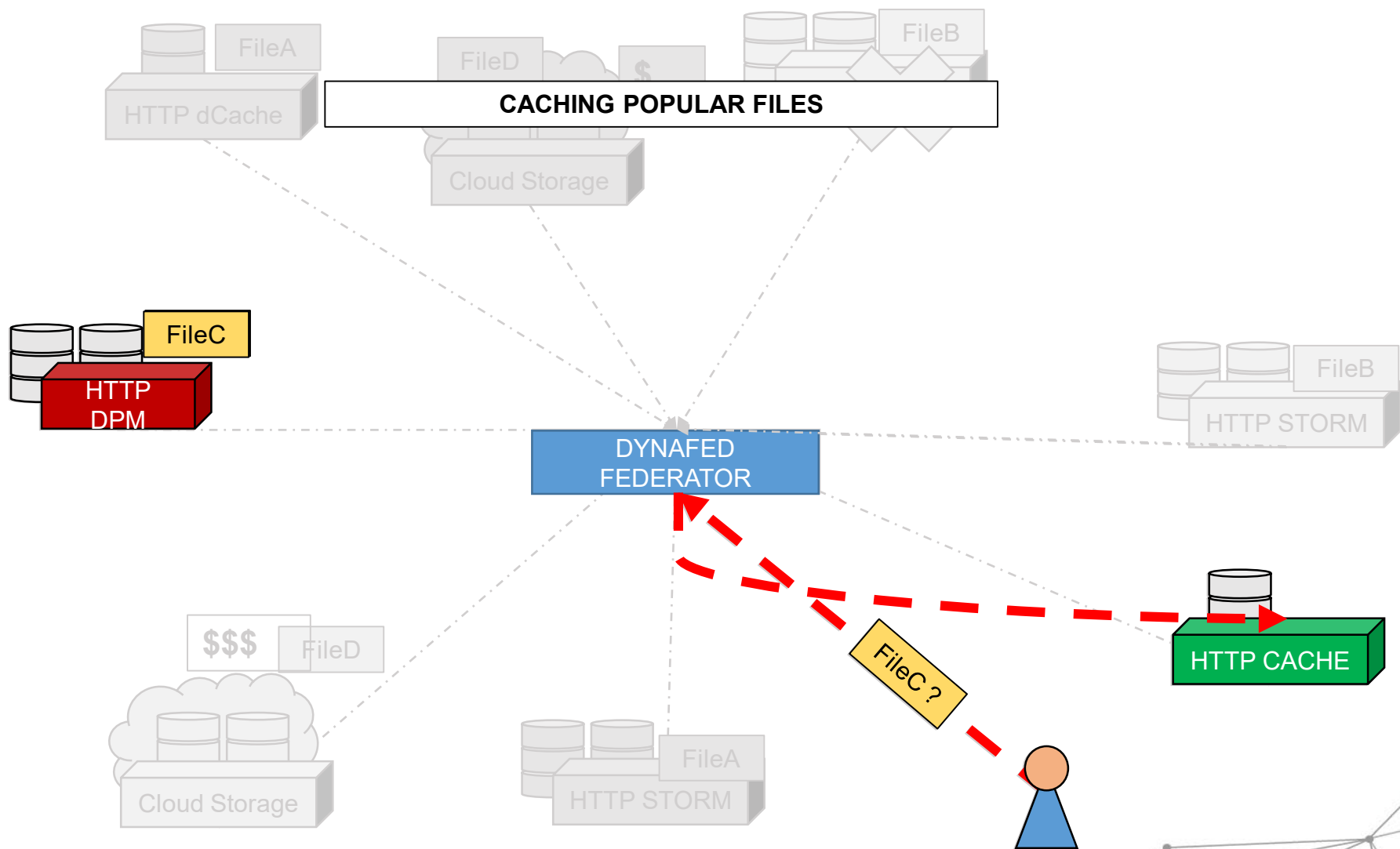


GIORNATA DI INCONTRO BORSE DI STUDIO GARR "ORIO CARLINI"
MARTEDÌ 6 DICEMBRE 2018 - ROMA



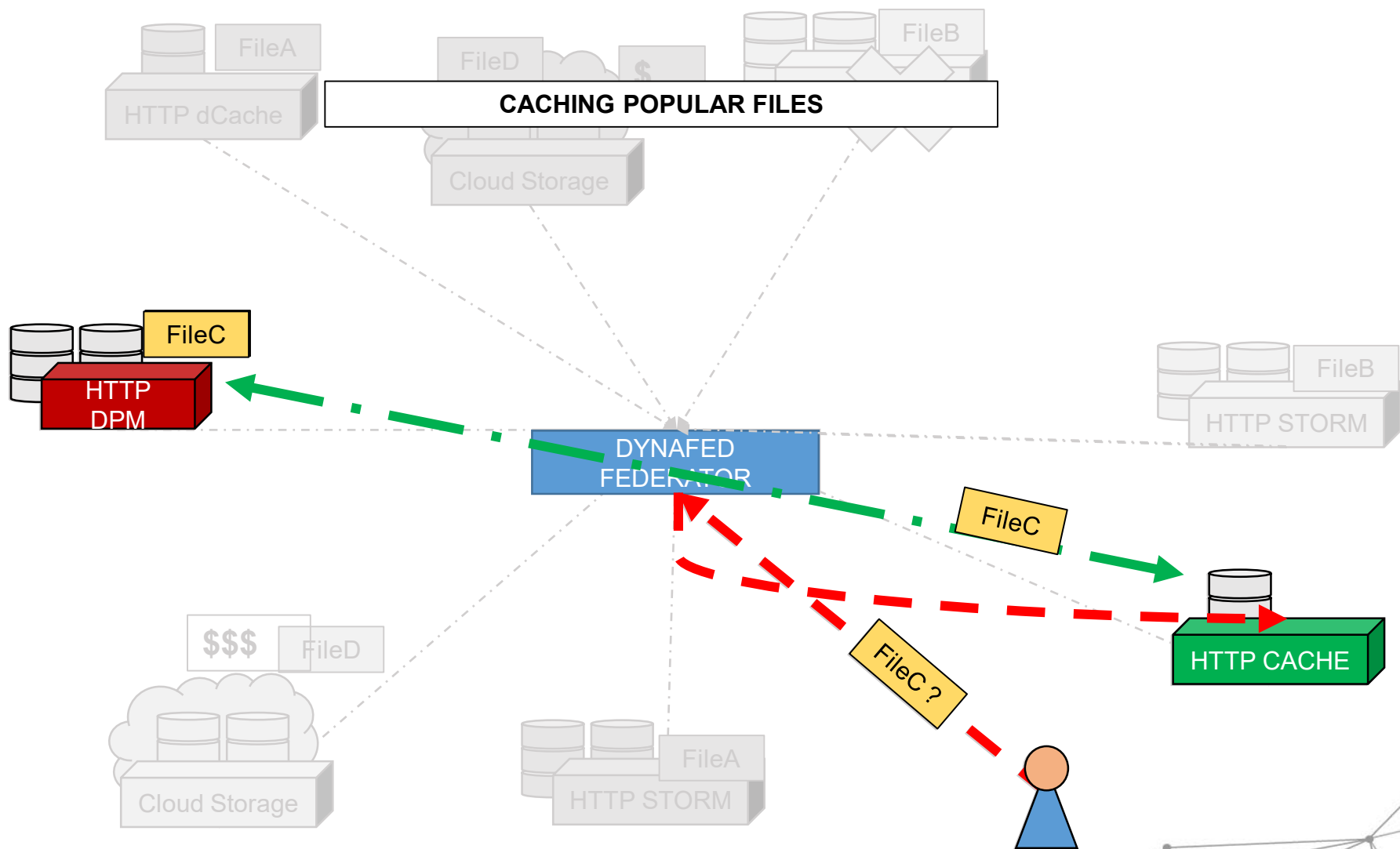


GIORNATA DI INCONTRO BORSE DI STUDIO GARR "ORIO CARLINI"
MARTEDÌ 6 DICEMBRE 2018 - ROMA



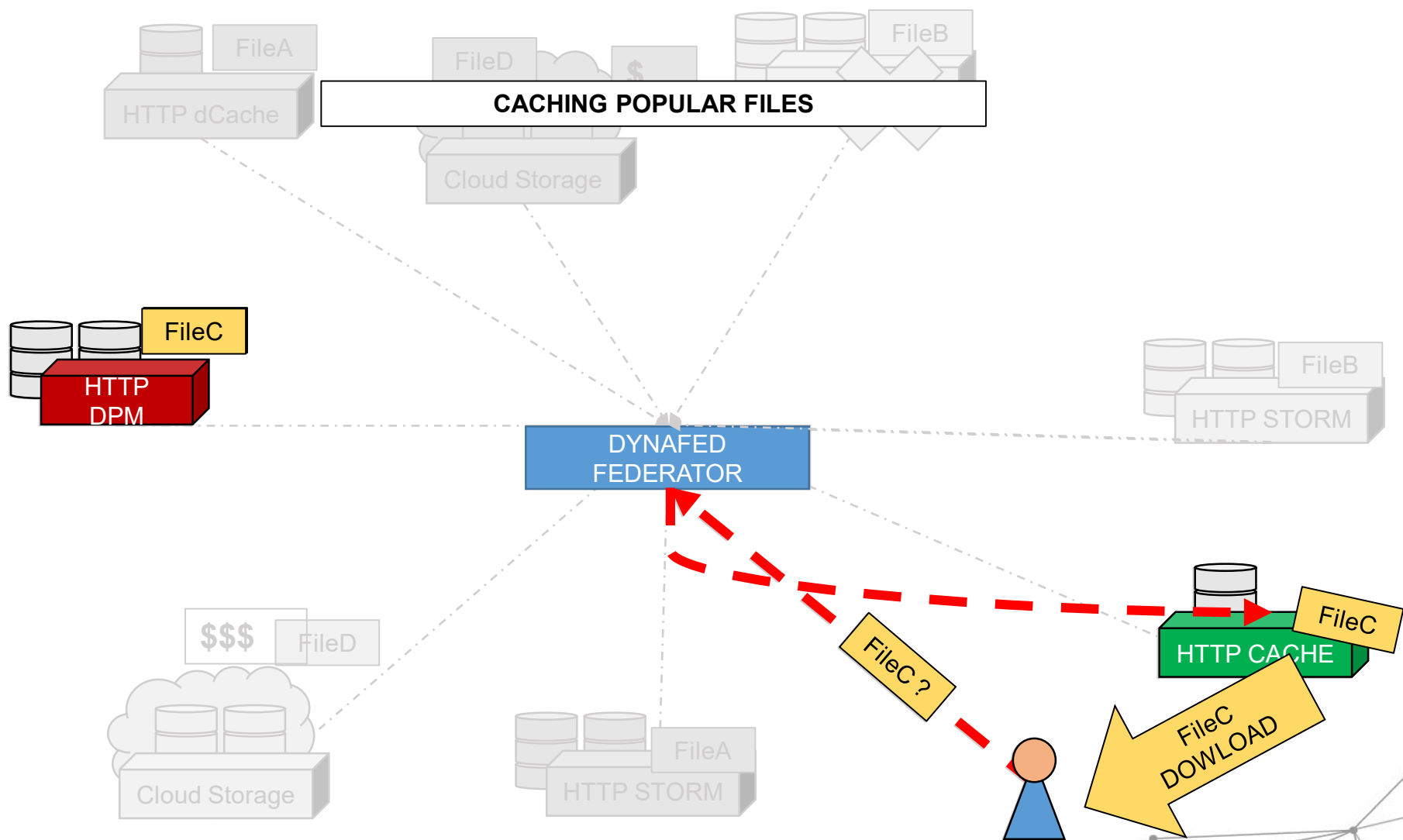


GIORNATA DI INCONTRO BORSE DI STUDIO GARR "ORIO CARLINI"
MARTEDÌ 6 DICEMBRE 2018 - ROMA



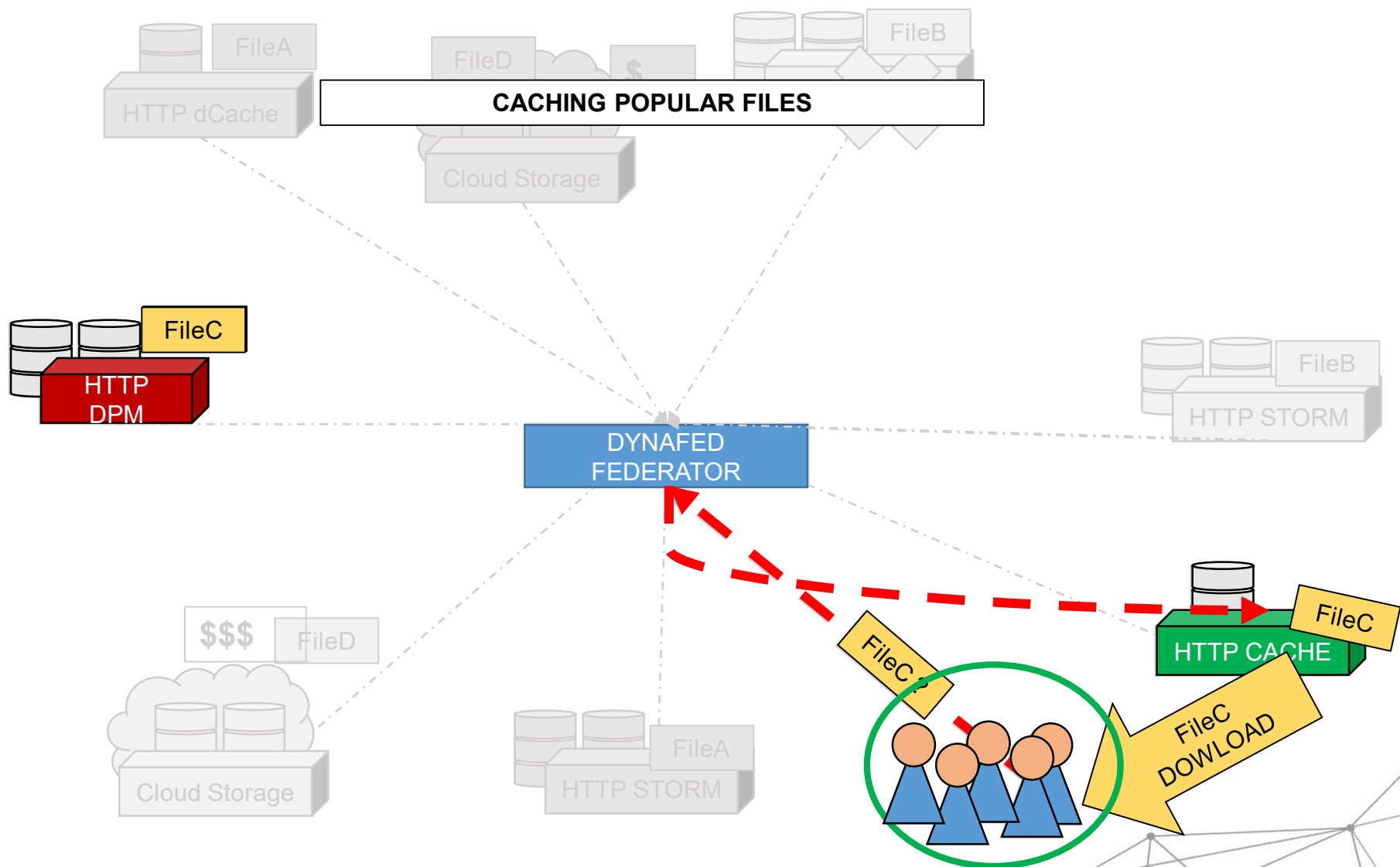


GIORNATA DI INCONTRO BORSE DI STUDIO GARR "ORIO CARLINI"
MARTEDÌ 6 DICEMBRE 2018 - ROMA



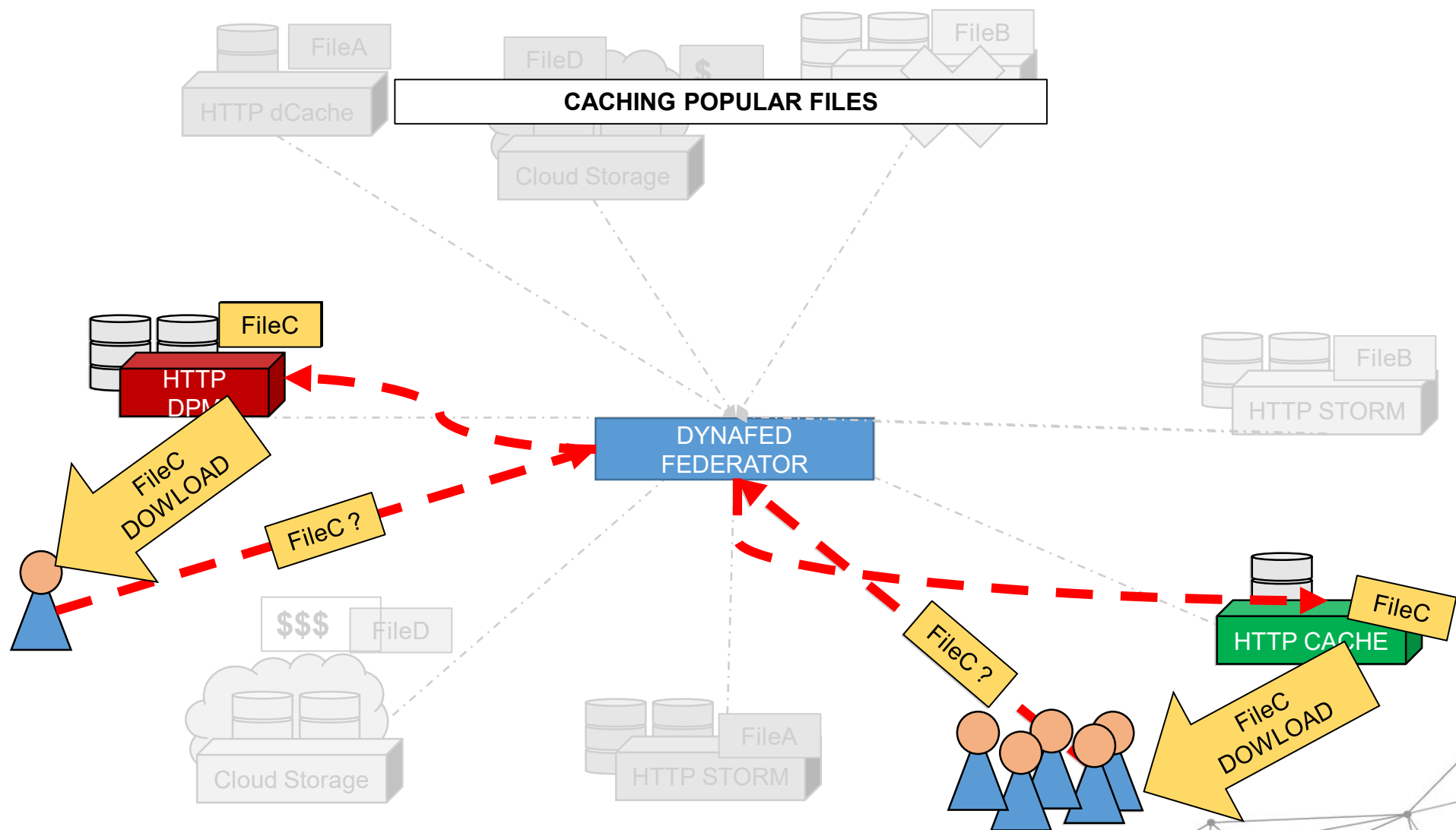


GIORNATA DI INCONTRO BORSE DI STUDIO GARR "ORIO CARLINI"
MARTEDÌ 6 DICEMBRE 2018 - ROMA



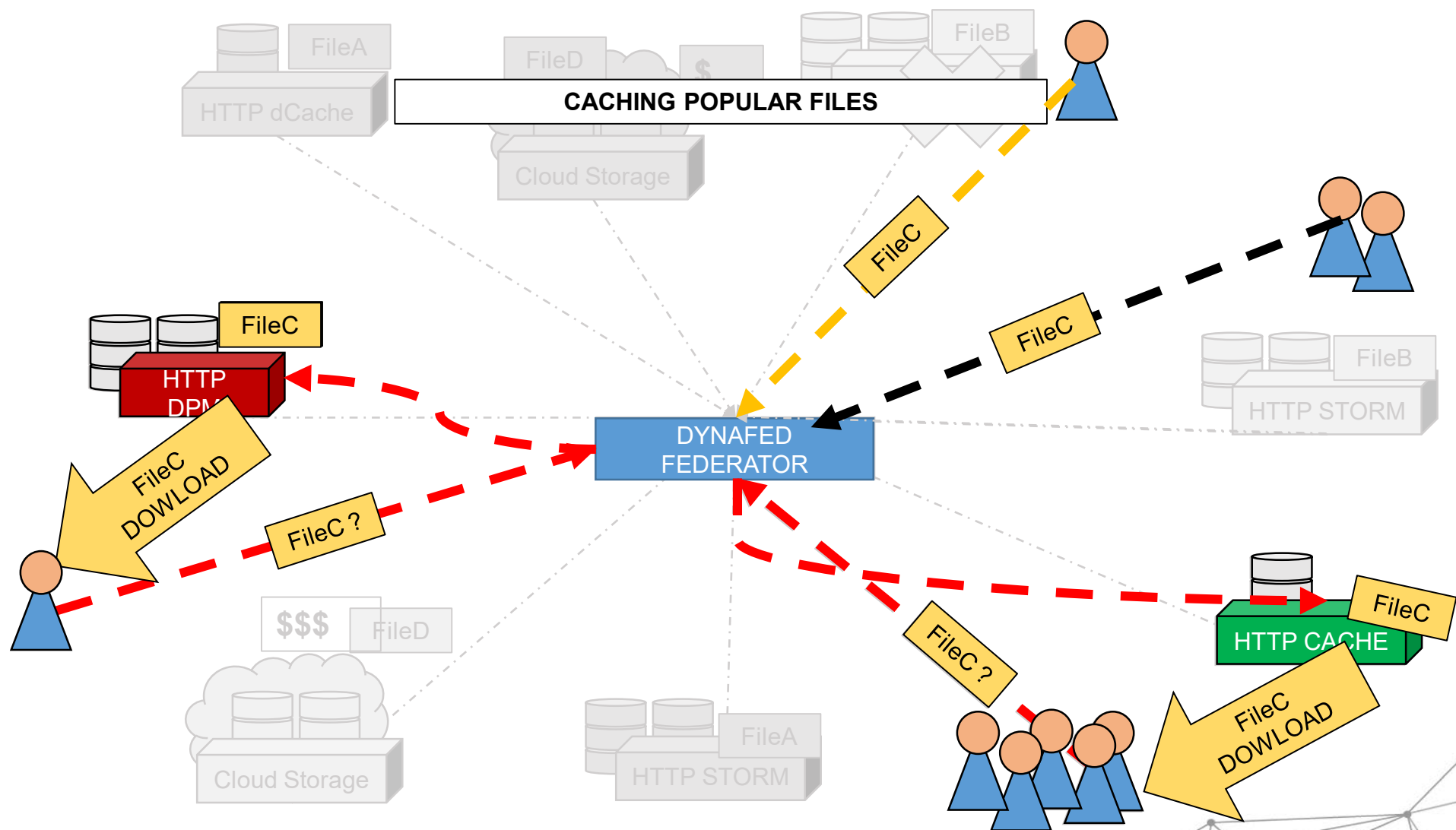


GIORNATA DI INCONTRO BORSE DI STUDIO GARR "ORIO CARLINI"
MARTEDÌ 6 DICEMBRE 2018 - ROMA



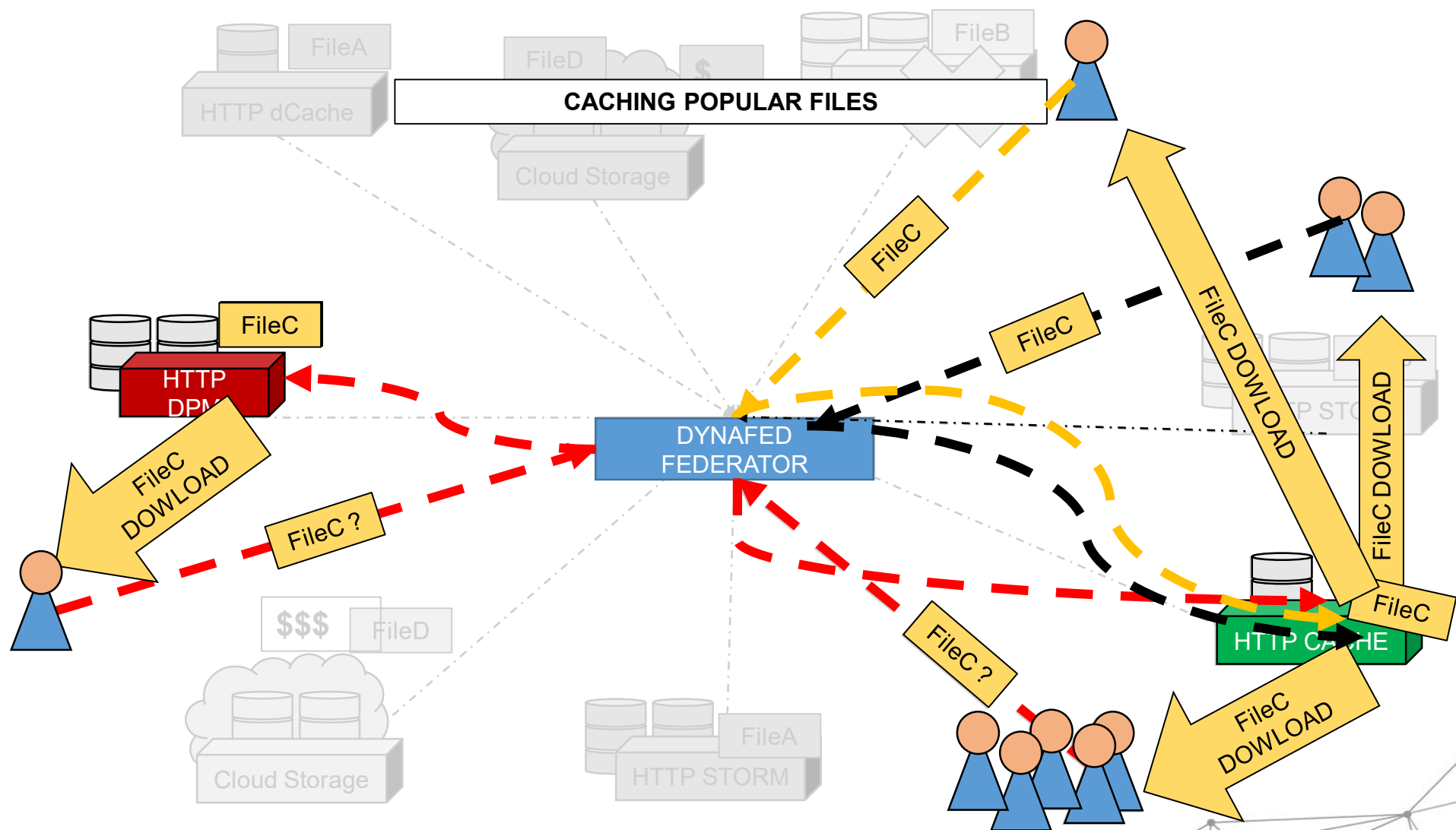


GIORNATA DI INCONTRO BORSE DI STUDIO GARR "ORIO CARLINI"
MARTEDÌ 6 DICEMBRE 2018 - ROMA



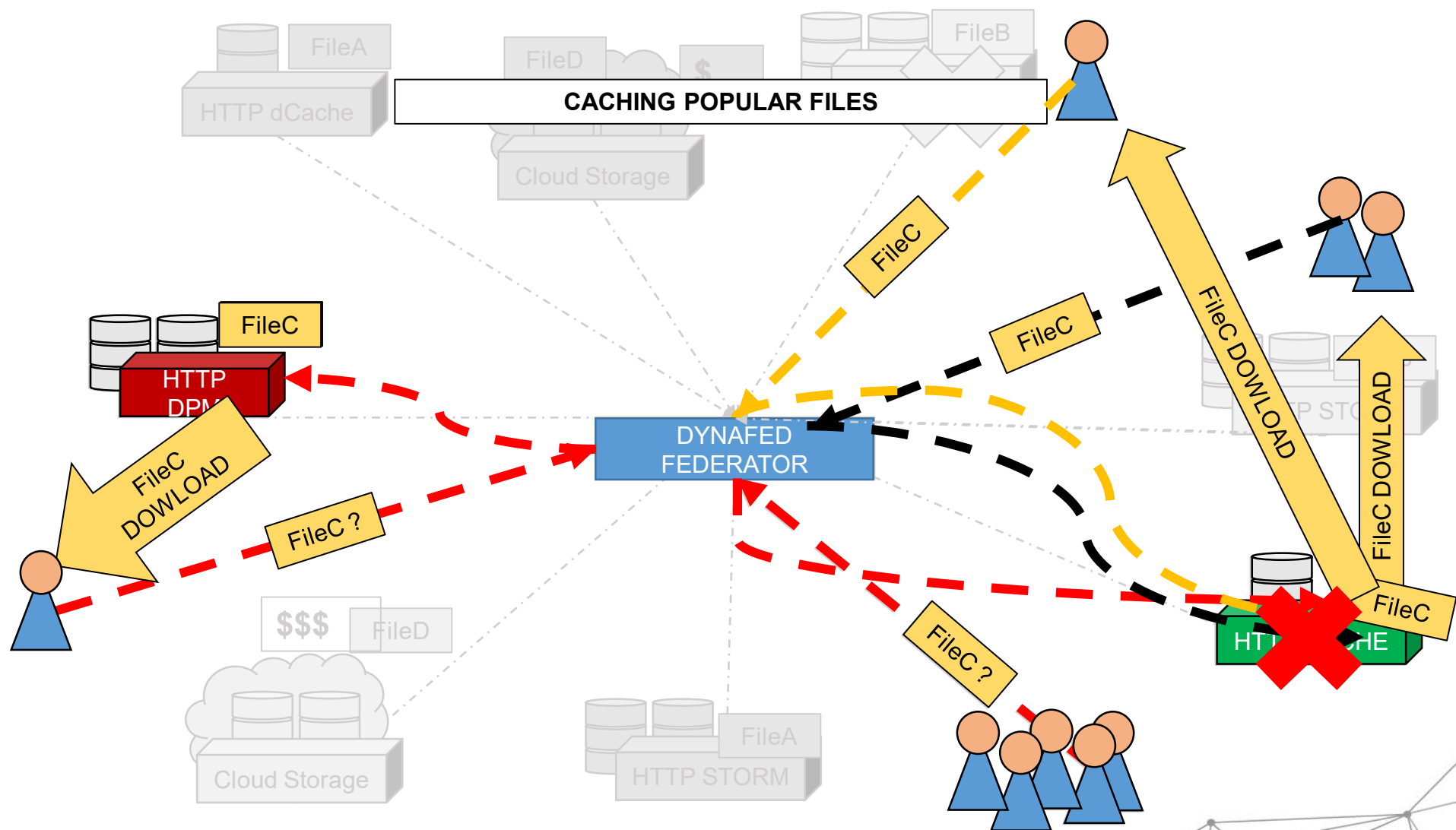


GIORNATA DI INCONTRO BORSE DI STUDIO GARR "ORIO CARLINI"
MARTEDÌ 6 DICEMBRE 2018 - ROMA



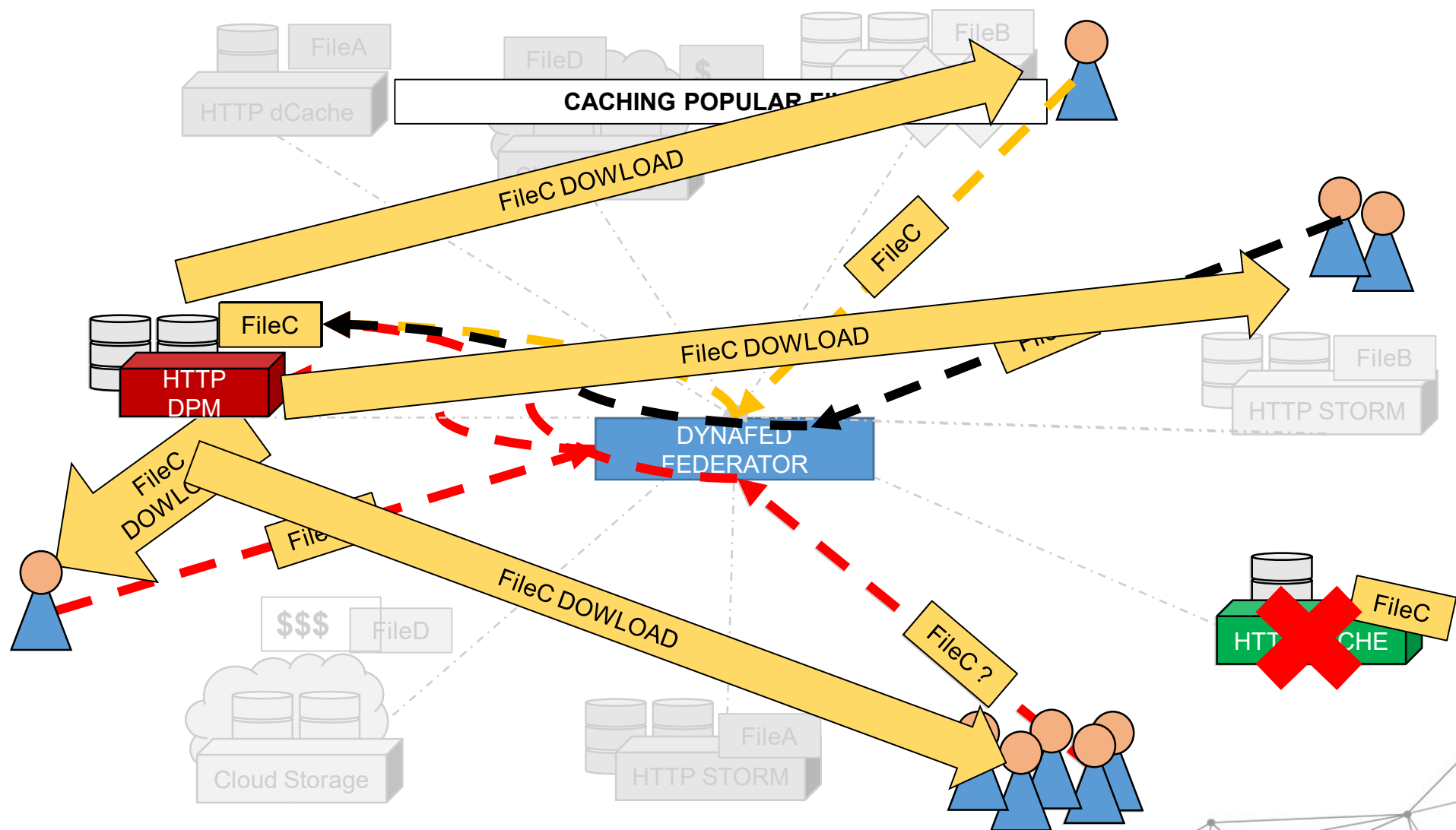


GIORNATA DI INCONTRO BORSE DI STUDIO GARR "ORIO CARLINI"
MARTEDÌ 6 DICEMBRE 2018 - ROMA



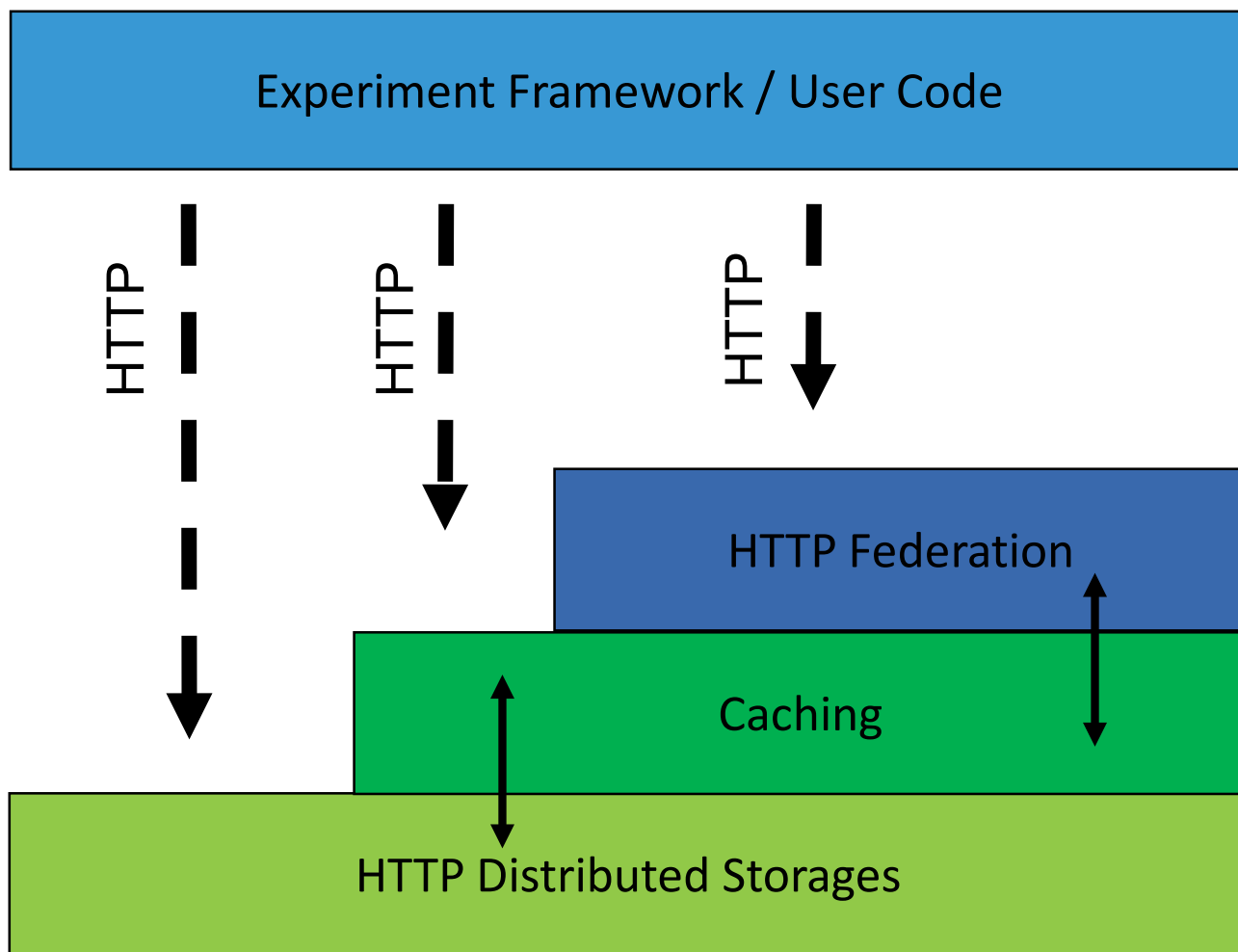


GIORNATA DI INCONTRO BORSE DI STUDIO GARR "ORIO CARLINI"
MARTEDÌ 6 DICEMBRE 2018 - ROMA



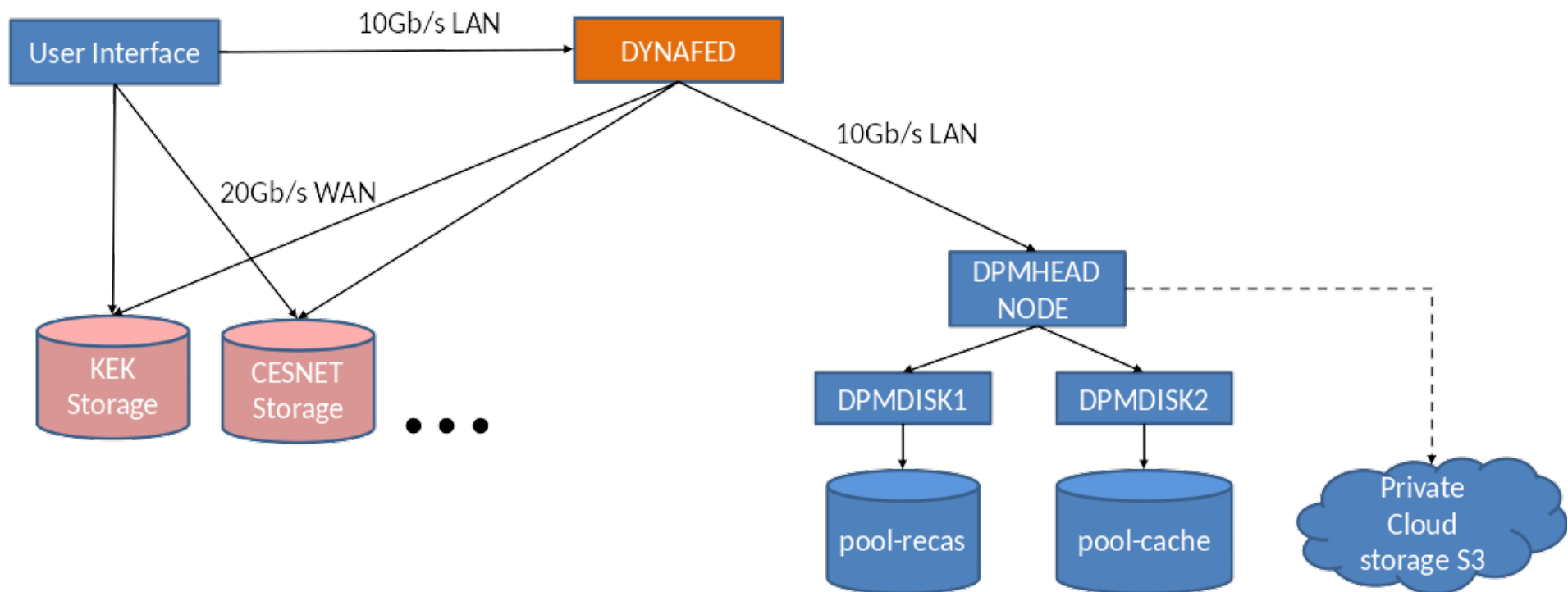


Il modello architetturale



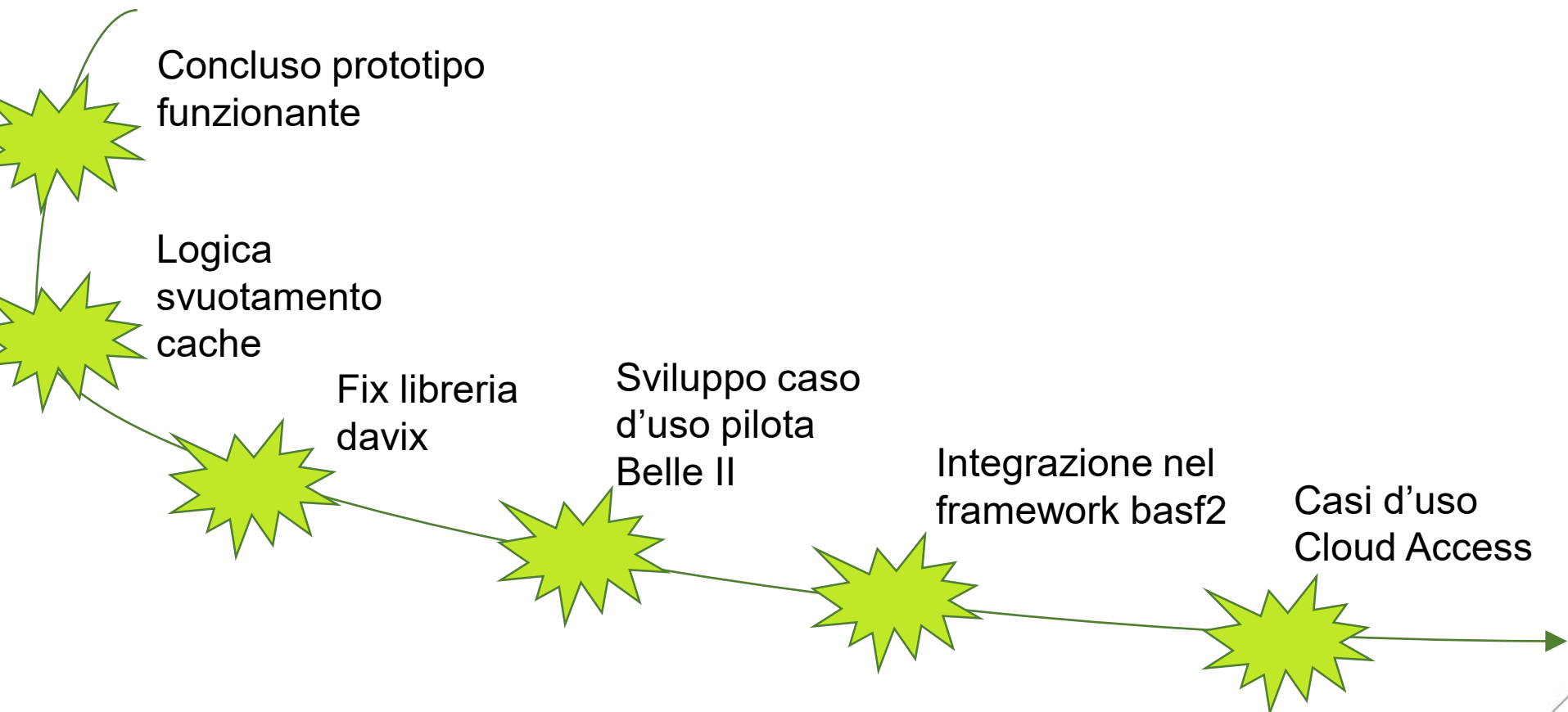


Il testbed implementato con DPM+DynaFED





Principali obiettivi raggiunti nel corso dell'attività





USE CASE Belle II: il testbed nell'ambito dell'esperimento

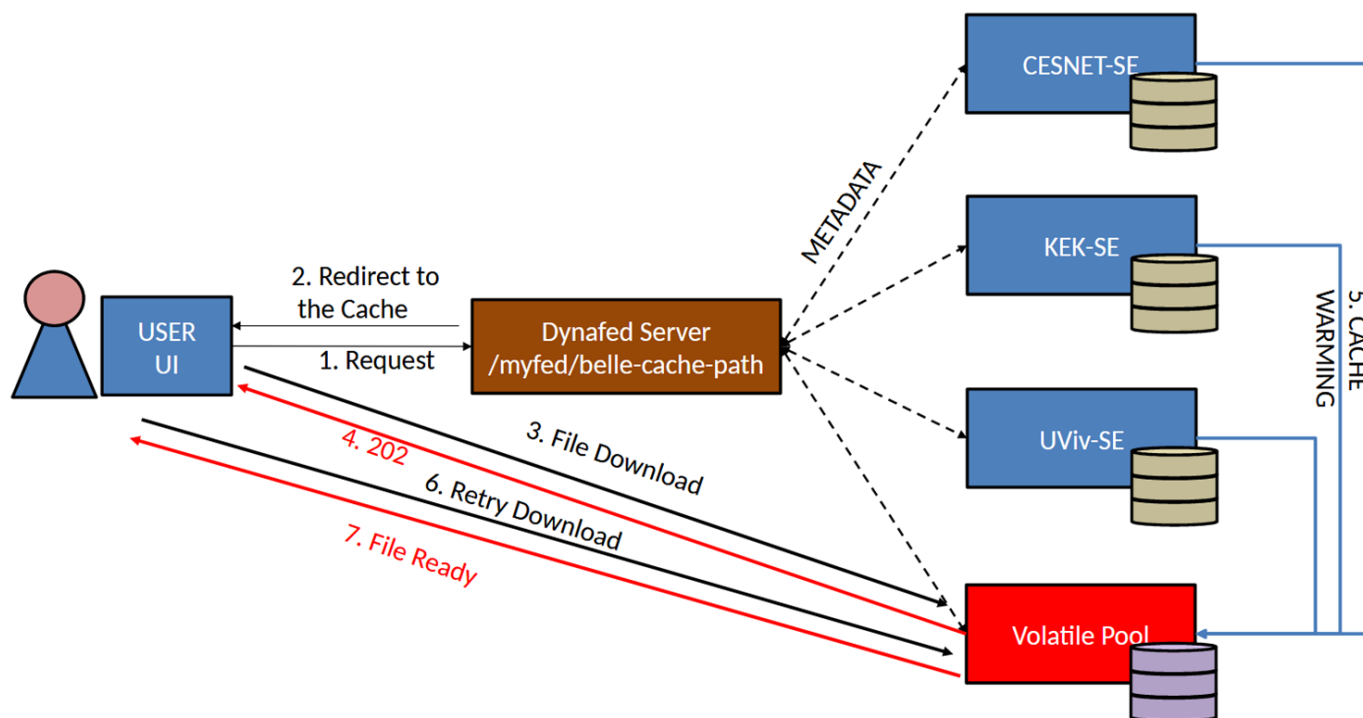
#	STORGE NAME	HOSTNAME	TYPE
1	DESY-DE	dcache-belle-webdav.desy.de	DCACHE
2	GRIDKA-SE	f01-075-140-e.gridka.de	DCACHE
3	NTU-SE	bgrid3.phys.ntu.edu.tw	DCACHE
4	SIGNET-SE	dcache.ijs.si	DCACHE
5	UVic-SE	charon01.westgrid.ca	DCACHE
6	BNL-SE	dcblldoor01.sdcc.bnl.gov	DCACHE
7	Adelaide-SE	coepp-dpm-01.ersa.edu.au	DPM
8	CESNET-SE	dpm1.egee.cesnet.cz	DPM
9	CYFRONNET-SE	dpm.cyf-kr.edu.pl	DPM
10	Frascati-SE	atlasse.lnf.infn.it	DPM
11	HEPHY-SE	hephyse.oew.ac.at	DPM
12	Melbourne-SE	b2se.mel.coepp.org.au	DPM
13	Napoli-SE	belle-dpm-01.na.infn.it	DPM
14	ULAKBIM-SE	torik1.ulakbim.gov.tr	DPM
15	IPHC-SE	sbgse1.in2p3.fr	DPM
16	CNAF-SE	ds-202-11-01.cr.cnaf.infn.it	STORM
17	ROMA3-SE	storm-01.roma3.infn.it	STORM
18	KEK-SE	Kek-se03.cc.kek.jp	STORM
19	McGill-SE	gridftp02.clumeq.mcgill.ca	STORM

- DynaFED aggrega quasi tutti gli **storage di produzione**
- Tecnologie **eterogenee**
- Storage **distribuiti** in tutto il globo



USE CASE Belle II: download e lettura locale

- Utilizzo client **davix**
- Download diretto da una **UI** utilizzando DynaFED
- Utilizzo del framework **basf2**





USE CASE Belle II: integrazione in DIRAC 1/3

Configurazioni provate:

- Registrare il pool volatile come un SE standard (perdita vantaggi DynaFED)
- Registrare DynaFED come SE (perdita controllo delle scritte)
- **Configurazione lettura/scrittura "asincrona" degli endpoint HTTP**

Resource Summary

Selectors

Name:

- Napoli-DAVS-ext-SE
- Napoli-DAVS-SE
- Dynafed-Napoli-SE

ResourceType:

- StorageElement

Status:

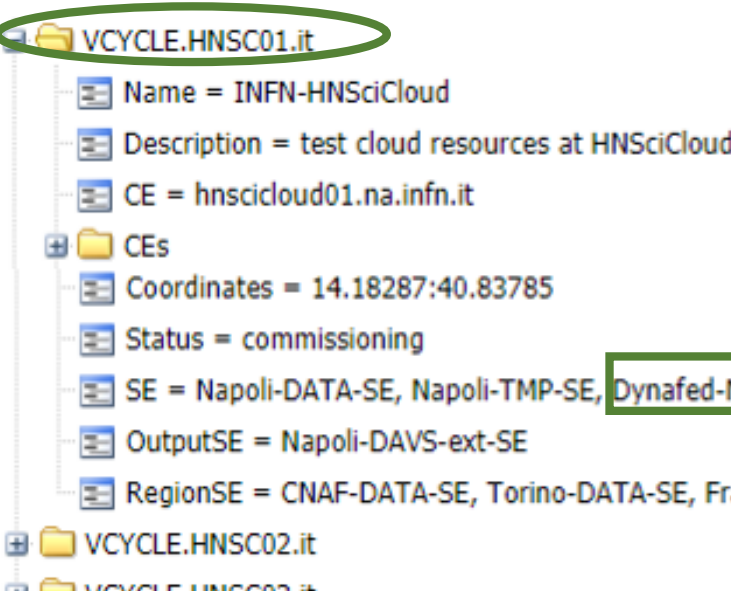
StatusType:

Items per page: 100 Page 1 of 1 Updated: 2018-07-11 07:3

Name	ResourceType	DateEffective	StatusT...	Status	Reason
Dynafed-Napoli...	StorageElement		4 elem...	Degraded	Not completely a
Dynafed-Napoli...	StorageElement		CheckAcc...	Active	No DownTime ann...
Dynafed-Napoli...	StorageElement		ReadAcc...	Active	No DownTime ann...
Dynafed-Napoli...	StorageElement		RemoveA...	Active	AlwaysBanned ###
Dynafed-Napoli...	StorageElement		WriteAcc...	Active	AlwaysBanned ###
Napoli-DAVS-SE	StorageElement		4 elem...	Active	All Active
Napoli-DAVS-ex...	StorageElement		4 elem...	Active	All Active



USE CASE Belle II: integrazione in DIRAC 2/3



The screenshot shows the DIRAC configuration for the site VCYCLE.HNSC01.it. The configuration details are as follows:

- Name = INFN-HNSciCloud
- Description = test cloud resources at HNSciCloud
- CE = hnscloud01.na.infn.it
- CEs
 - Coordinates = 14.18287;40.83785
 - Status = commissioning
 - SE = Napoli-DATA-SE, Napoli-TMP-SE, Dynafed-Napoli-SE, Napoli-DAVS-SE, Napoli-DAVS-ext-SE, KEK-DAVS-ext-SE
 - OutputSE = Napoli-DAVS-ext-SE
 - RegionSE = CNAF-DATA-SE, Torino-DATA-SE, Frascati-DATA-SE, Pisa-DATA-SE, Napoli-DAVS-SE, Napoli-DAVS-ext-SE, KEK-DAVS-ext-SE

Below the configuration, three other sites are listed: VCYCLE.HNSC02.it and VCYCLE.HNSC03.it.

Dirac è un **framework** per la gestione 'all-in-one' per il **calcolo distribuito**.

- Il sito VCYCLE.HNSC01.it nel DIRAC di PNNL è stato configurato per usare **endpoint HTTP**, incluso **DynaFed**.
- Sono stati copiati e registrati sullo storage **KEK-DAVS-SE** dei dataset con il comando `gb2_ds_put`.



USE CASE Belle II: integrazione in DIRAC 3/3

- Sottomissione di job **gbasf2** via **DIRAC** utilizzando la cache.
- Dataset replicati a **KEK**
- Job di analisi accedendo a **Dynafed via HTTP**
- Test **intero workflow di analisi**

Items per page: 25 Page 1 of 1 Updated: 2018-07-11 06:49 [UTC](0 0								
JobId ▾	Status	Min...	ApplicationSta...	Site	Job...	LastUpdate[UTC]	LastSignOfLife[UTC]	
70941	 Done	Exe...	Done	VCYCLE.HNSCO...	pro...	2018-07-10 14:38:54	2018-07-10 14:38:54	
70940	 Done	Exe...	Done	VCYCLE.HNSCO...	pro...	2018-07-10 14:30:11	2018-07-10 14:30:11	
70939	 Done	Exe...	Done	VCYCLE.HNSCO...	pro...	2018-07-10 13:48:33	2018-07-10 13:48:33	



USE CASE Cloud Access: riduzione costi di accesso

- Utilizzo del sistema di cache per l'accesso a **risorse in cloud**
- Utilizzate **risorse di T-System** nell'ambito del progetto HNSC
- Stima trasferimento **100GB/mese: 6.7€**

The screenshot shows the T-System Object Storage Service pricing calculator. It features a pink header with the T-System logo. The main content area is titled 'Object Storage Service' and includes a dropdown menu set to 'Standard'. Below this, there are three sections: 'Capacity', 'Lifecycle Requests', and 'Outbound Traffic'. Each section has a numeric input field, a unit, and a calculated cost. The 'Capacity' section shows 100 GB/Month with a cost of -2.30 €/month. The 'Lifecycle Requests' section shows 0 1000 Requests with a cost of -0.00 €/month. The 'Outbound Traffic' section shows 100 GB with a cost of -6.70 €/month. A final total of 9.00 €/month is displayed at the bottom right.

Resource	Unit	Cost
Capacity	100 GB/Month	- 2.30 €/month
Lifecycle Requests	0 1000 Requests	- 0.00 €/month
Outbound Traffic	100 GB	- 6.70 €/month
Total		= 9.00 €/month



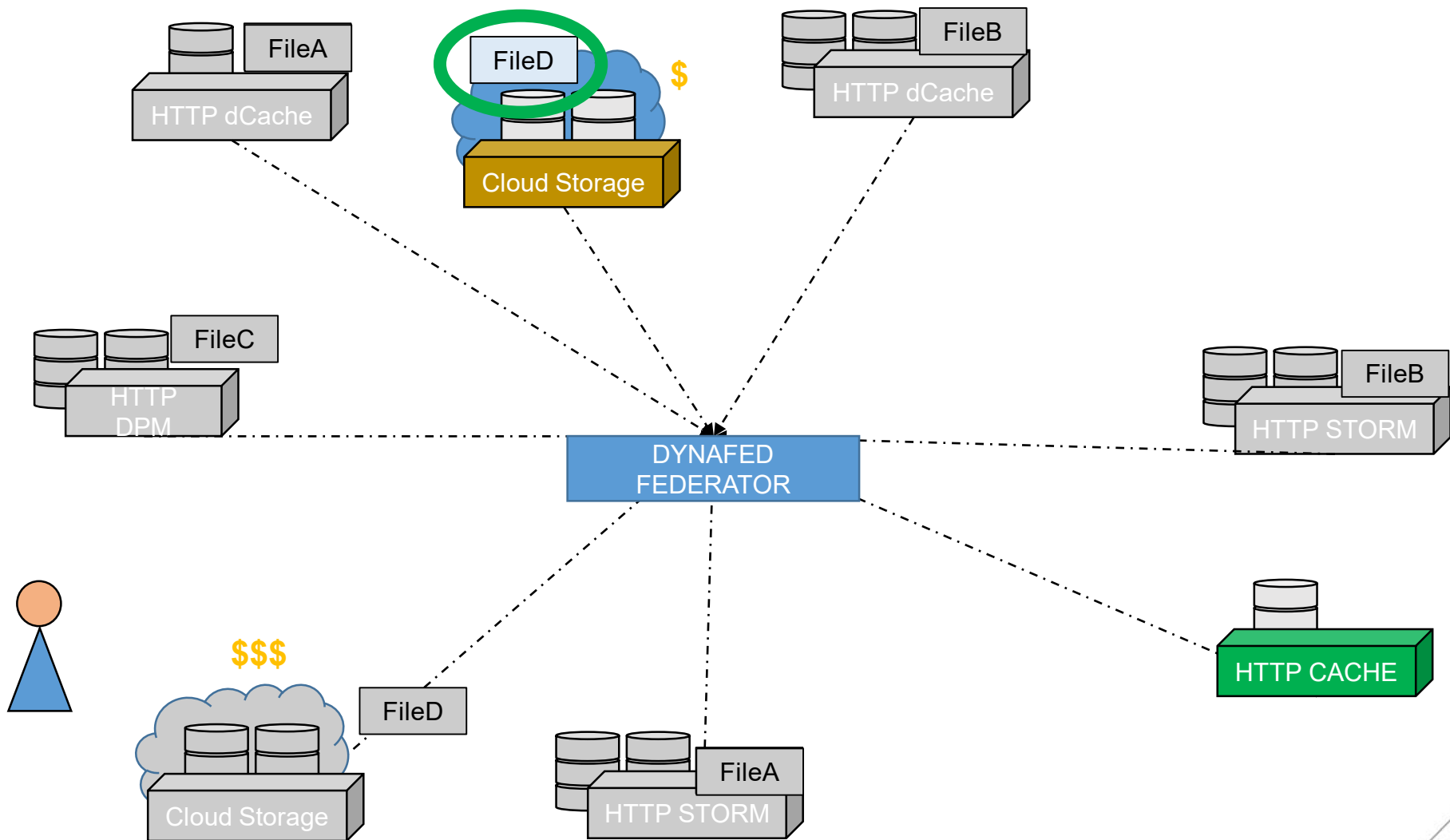
USE CASE Cloud Access: nuovi plugin di DynaFED

- Necessità di gestire **diversamente l'ordine delle repliche** dei metalink
- Sviluppo di **due nuovi plugin** per DynaFED:
 - **Price plugin**
 - **Default Plugin**

L'uso combinato di questi due plugin permette di creare numerosi nuovi scenari!

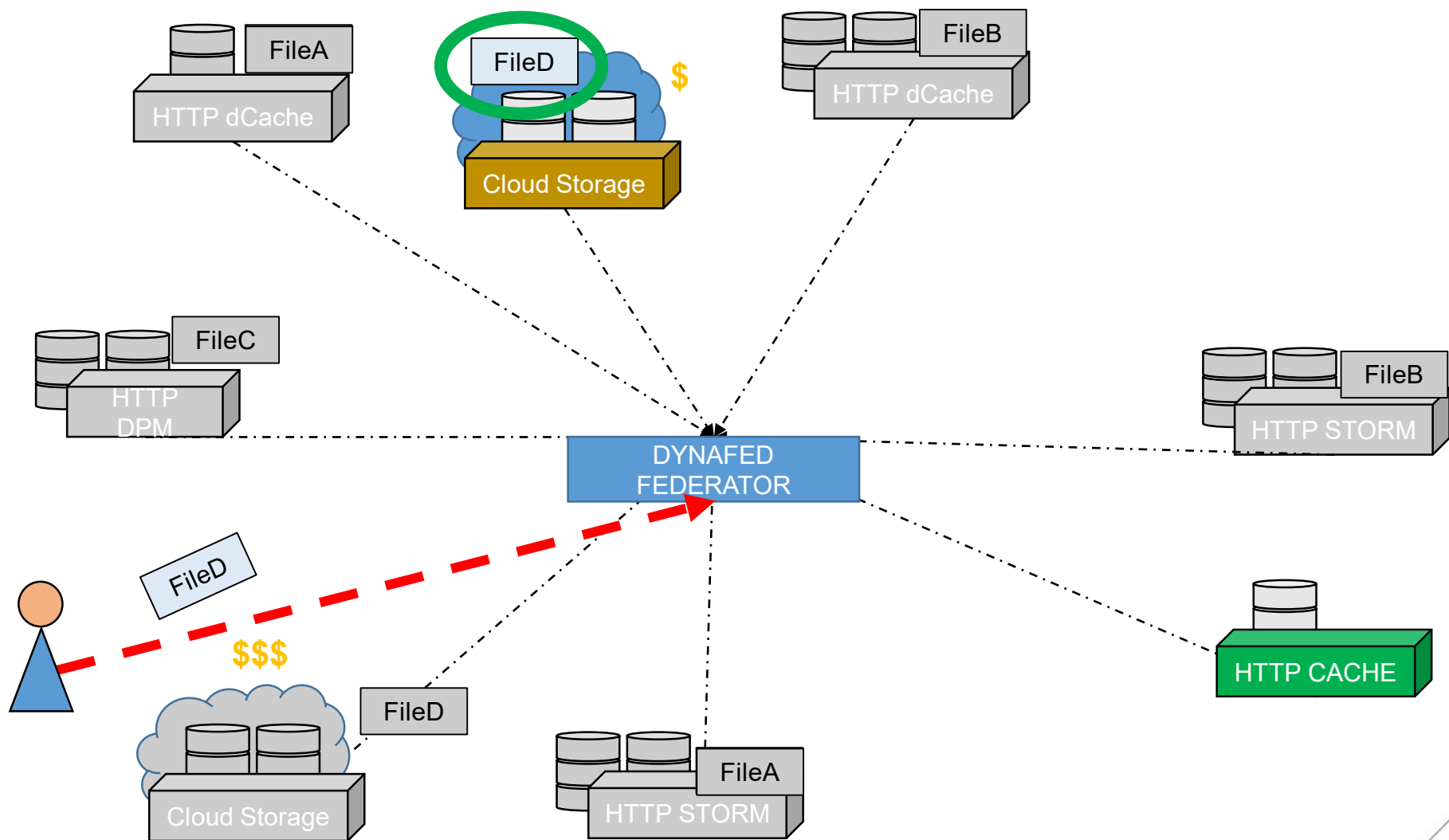


GIORNATA DI INCONTRO BORSE DI STUDIO GARR "ORIO CARLINI"
MARTEDÌ 6 DICEMBRE 2018 - ROMA



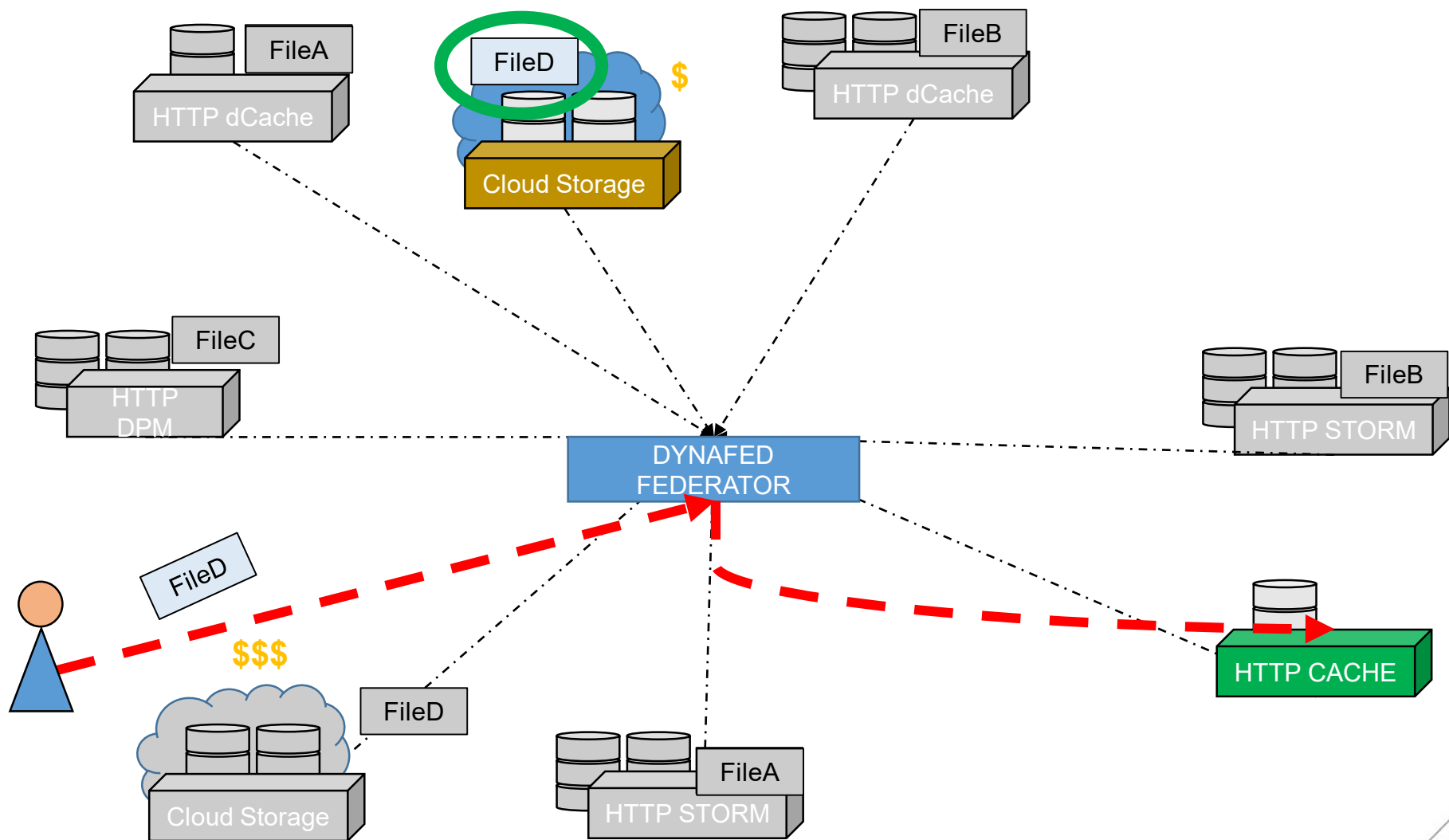


GIORNATA DI INCONTRO BORSE DI STUDIO GARR "ORIO CARLINI"
MARTEDÌ 6 DICEMBRE 2018 - ROMA



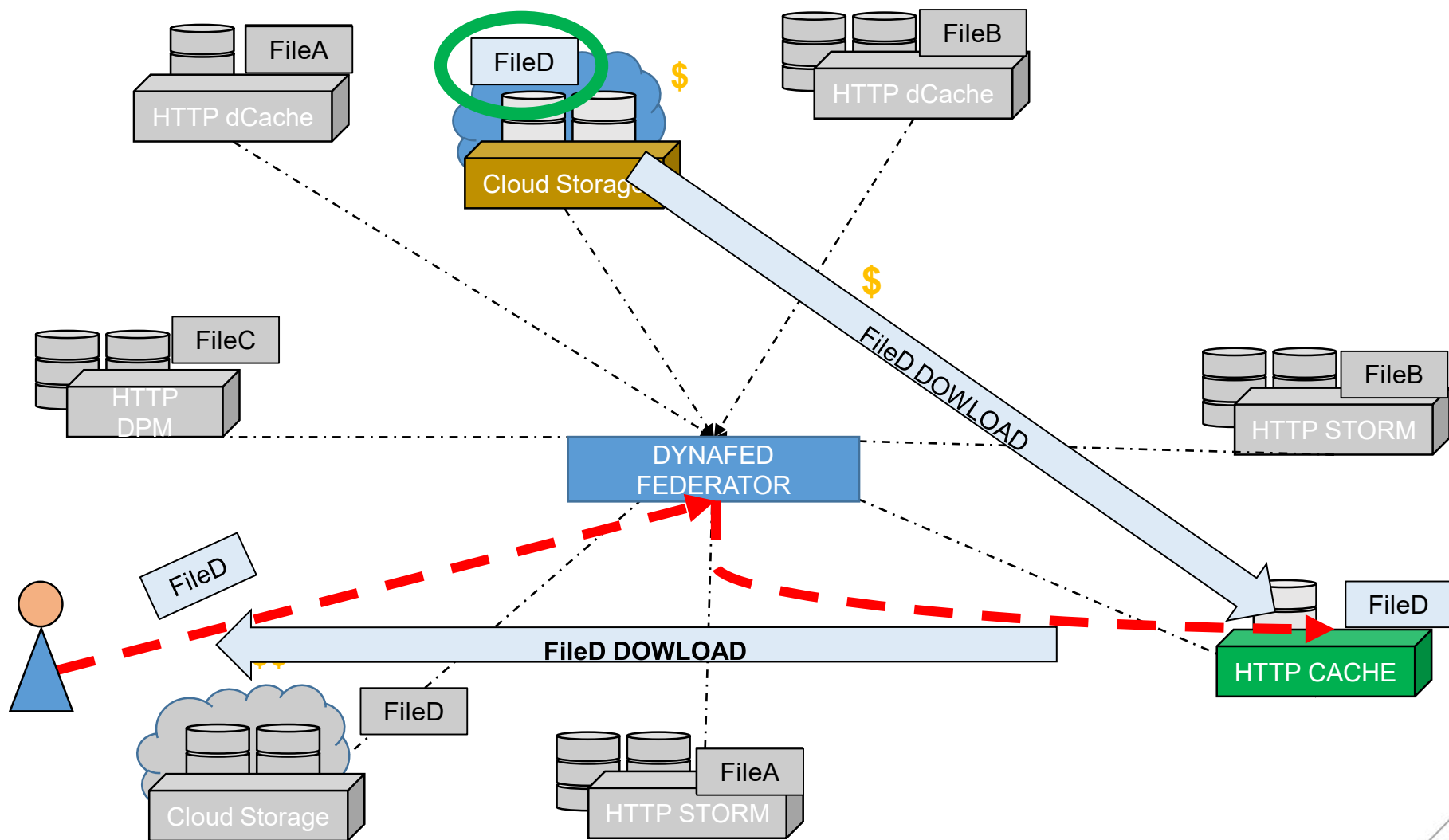


GIORNATA DI INCONTRO BORSE DI STUDIO GARR "ORIO CARLINI"
MARTEDÌ 6 DICEMBRE 2018 - ROMA



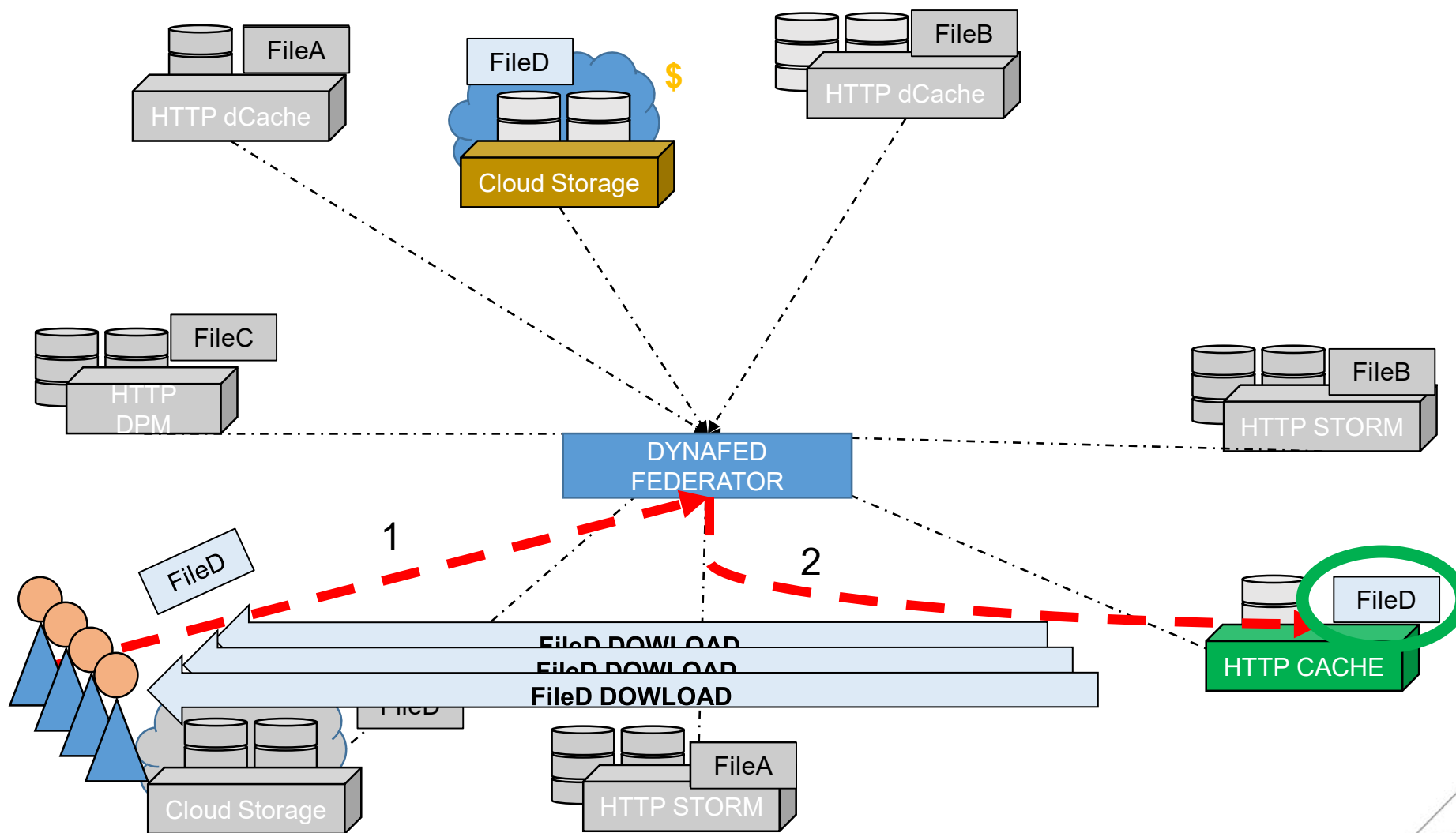


GIORNATA DI INCONTRO BORSE DI STUDIO GARR "ORIO CARLINI"
MARTEDÌ 6 DICEMBRE 2018 - ROMA





GIORNATA DI INCONTRO BORSE DI STUDIO GARR "ORIO CARLINI"
MARTEDÌ 6 DICEMBRE 2018 - ROMA





USE CASE Cloud Access: sperimentazione plugin

Stima costi da un client **sulla rete DESY** (Germania) usando i nuovi plugin:

	Total Size (GB)	Plugin	Costo I accesso	Costo II accesso	Costo III acceso
CLOUD	100	GeoIP	6,7 €	6,7 €	6,7 €
SCORES	100	GeoIP+Price/Default	6,7 €	0	0

Configurazione PRICE Plugin

recas-dpm-01.na.infn.it	0.20
dcache-belle-webdav.desy.de	0.40
kek2-se03.cc.kek.jp	0.50
dcache.ijs.si	0.50
charon01.westgrid.ca	0.50
dpm1.egee.cesnet.cz	0.50
davide.obs.otc.t-systems.com	0.80

Configurazione Default Plugin

131.169.168	recas-dpm-01.na.infn.it
79.23	kek2-se03.cc.kek.jp



Principali collaborazioni

Nel corso delle attività della borsa sono nate diverse **collaborazioni e confronti** sia con i **team di sviluppo** di dei principali software utilizzati che con vari **working group** internazionali che stanno trattando tematiche similari.

- Confronto con i team di sviluppo di **DPM**, **DynaFED** e librerie **Davix**.
- Attività di R&D nella **collaborazione Belle II** per lo sviluppo del caso d'uso principale.
- Working group UIVc, CERN, Belle II sull'utilizzo di **DynaFED**
- Working group DOMA (sottogruppo **DOMA-Access**) sull'utilizzo delle tecnologie di cache e data access.



Presentazioni, workshop e conferenze

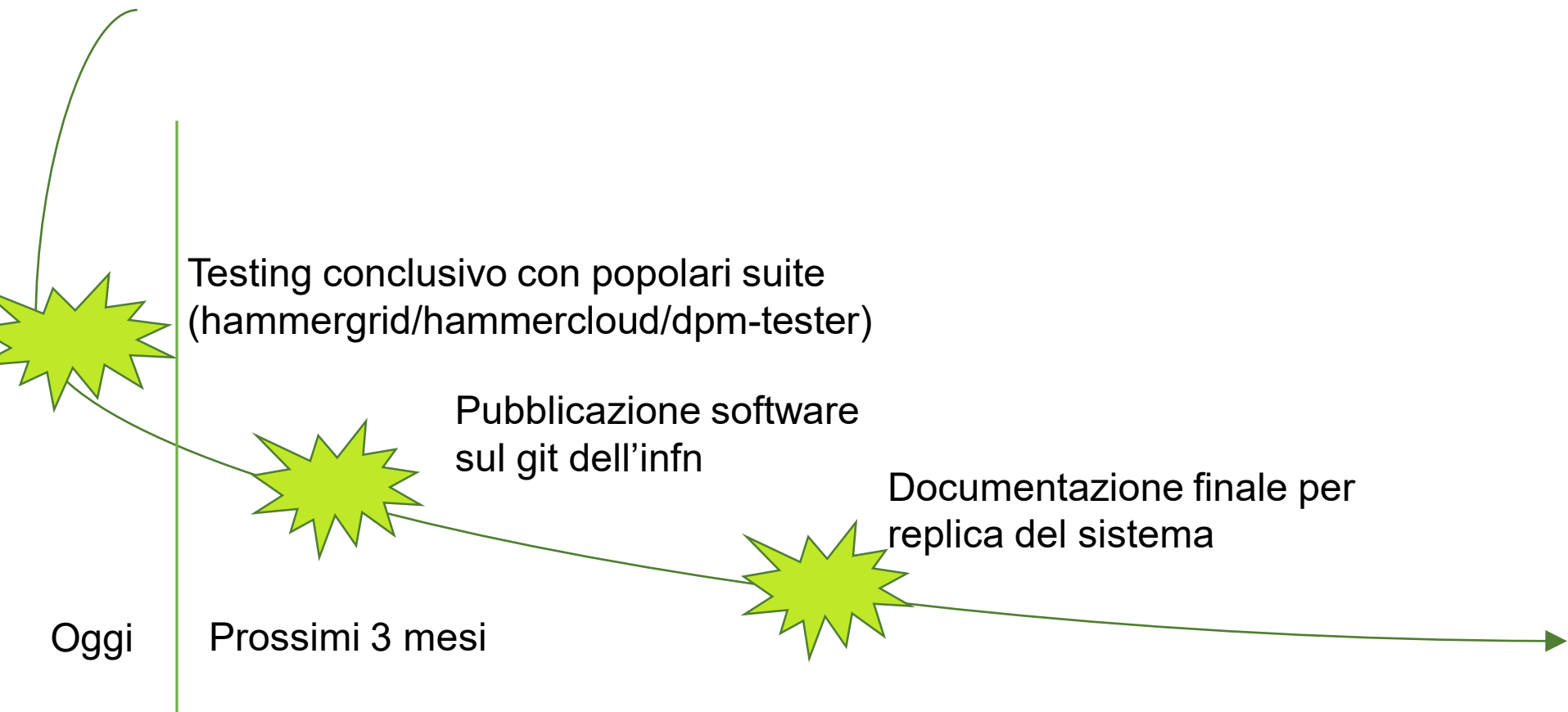
Nel corso della borsa le attività sono state condivise in diverse ambiti:

- WLCG Workshop - Napoli - Marzo 2018
- CCR INFN - Roma - Marzo 2018
- CHEP18 - Sofia, Bulgaria - Luglio 2018
- **TNC18 - Trondheim, Norvegia - Giugno 2018**
- BELLE II General Meeting 2018 - Japan, KEK





Attività in corso e conclusive della borsa





Potenziali attività per il 2019

Le tecnologie studiate trovano potenziale applicazione nell'ambito delle attività previste per il 2019 dall'INFN:

- Progetto Gruppo 5 "IDDLS: Italian Distributed Data Lake for Science", in collaborazione con il GARR.
- Progetto PRIN "HTTPhy" (in fase di valutazione), per un sistema di accesso ai dati basato sull'ecosistema HTTP e tecnologie di cache.
- Infrastruttura IBISCO - Progetto PON di Potenziamento dell'infrastruttura di Calcolo del Sud Italia per i BigData ed Data Science



Grazie per l'attenzione.
Domande?

Tutti i contenuti originali della presentazione sono coperti dalla licenza:
<https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode>

