Calcolo: Il caso RADIO

(Elaborazione di segnali ed immagini)

Mauro Nanni INAF-Ist. Radioastronomia

La riduzione e l'analisi dei dati negli Anni 80 - 90

40 - 60 Ricercatori si disputavano le risorse di calcolo di macchine molto costose





Applicazioni autoprodotte e Packages: AIPS (1977) / DIFMAP (1992)

La riduzione e l'analisi dei dati negli Anni 2000 -> 2010

L'elaborazione si sposta su strumenti individuali



Work-Station
10 X Convex C1

- >Costi Ridotti
- >Disponibilita'
- >Uso esclusivo delle risorse
- >Ottimizzazione
- <Installazioni personali
- <Risorse Ridotte

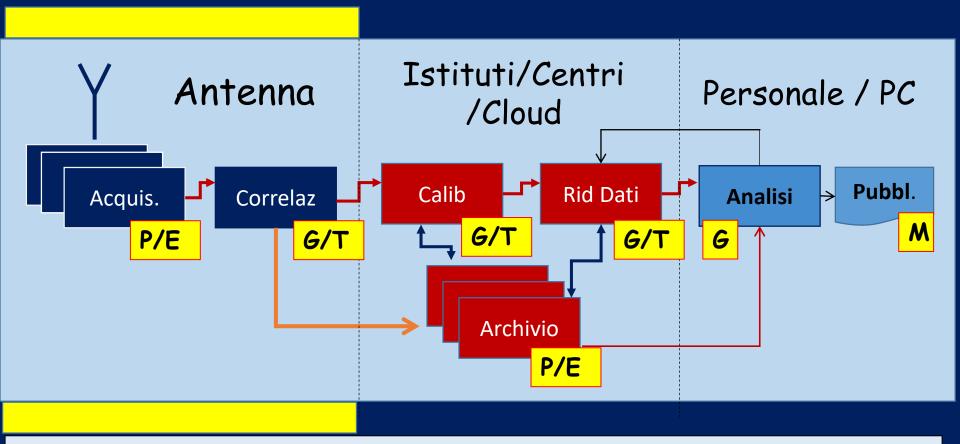


PC 100 X Convex C1

soluzione/mediazione nel "cluster di workstation" in rete

Packages: AIPS (1977) / DIFMAP (1992) / Miriad (1995)

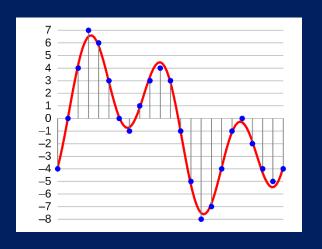
Flussi e dimensione dei dati



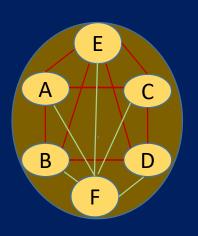
```
G Giga - Chiavetta
T Tera - Disco
P Peta - 1.000 Tera (con dischi da 10TBy si impegna 40% di un Rack )
E Exa - 1.000.000 Tera
```

La correlazione:

Mettere assieme i segnali (campionati) che giungono dalle single antenne per realizzare una antenna unica.



Campionamento del voltaggio del segnale di ogni antenna ogni N.microsecondi.



Calcolare il prodotto per ogni coppia di antenna

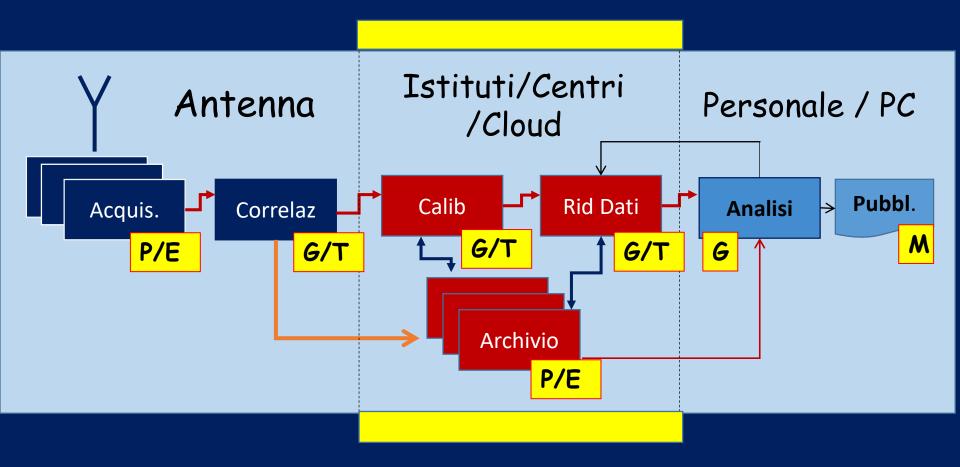
Il numero di coppie (baseline) e' dato da:

__(N+1) ! __2*(N-1) !



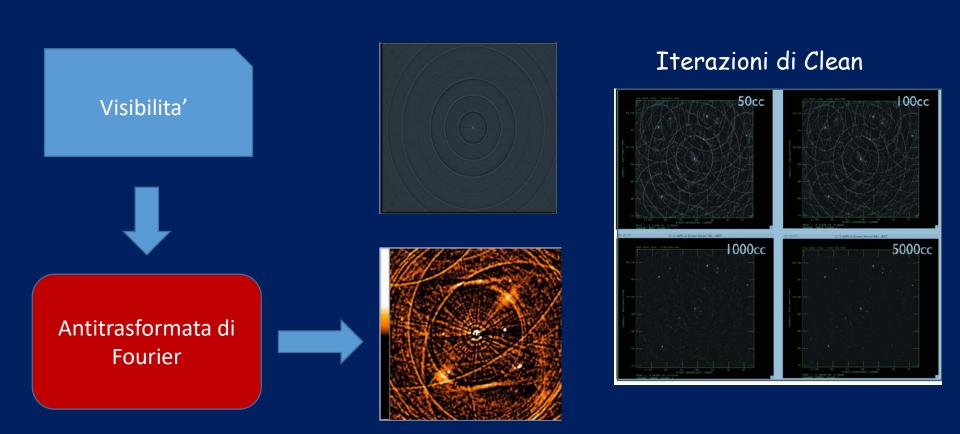
Numero Antenne	Numero Baseline
Antenne	Daseillie
2	1
3	3
4	6
5	10
6	15
12	78
32	528
64	2080
128	8256
254	32385
512	131228

Flussi e dimensione dei dati



Calibrazione e Riduzione Dati

Dalle "visibilita' " prodotte dal correlatore all'immagine



Pipeline iterattiva

Calibrazione / Riduzione dati : Il caso ALMA

Input Dataset: 10-100 GBy Work Space: 100 GBy - 1 TBy



Applicativo CASA: il meglio di AIPS++

- Facilmente installabile anche su PC
- Non separa Input / Work / Output
- Non sfrutta parallelismo
- Implementato skeduler

Implementazione LUSTRE NFS su Server con 24 dischi in Raid-5

Criticita': Conflitti di accesso ai dati in rete, rallentamenti non accettabili!



Scelta dei server indipendenti

- Work Space indipendente! (Dischi locali in raid-0)
- Pochi core x macchina (6 8)
- Prenotazione risorse
- Spazio Lustre per Input e Output (Liberare i dischi di lavoro !!)

Calibrazione / Riduzione dati : Il caso Lofar

Input Dataset 50 GBy - 400 GBy Work Space 0.5TBy - 2 TBy



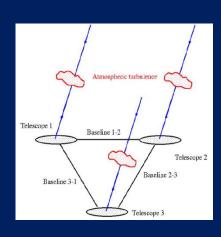
Applicativo Lofar (in sviluppo!)

- In Container (Simplicity)
- Uso della RAM come RamDisk
- Sfrutta il parallelismo

Scelta di server indipendenti

Criticita':

Dimensioni della RAM usata come RamDisk per ottimizzare I/O. Costo ~ 5-6E /Gby Limite attuali schede madri



- Memoria RAM 512 Gby
- Cores numerosi 23/64 Cores
- Work Space indipendente! (Dischi locali in raid-0)
- Prenotazione risorse
- Tempi di elaborazione : Giorni
- Spazio NAS per Input e Output (Liberare i dischi di lavoro !!)

By Elisabetta Liuzzo

EHT pipeline development

EHT is a new instrument operating at unprecedented sensitivity and angular resolution, required development of new software tools for calibration and imaging

Calibration: three dedicated calibration pipelines HOPS, CASA and AIPS



Haystack Observatory post-processing system

Common Astronomy Software Applications

Astronomical Image Processing system

Typical used working stations at Italian ARC for EHT calibration:

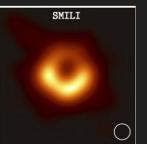
RAM 64Gb, CUP Intel Xeon E3-1275 v6, cores 4/8, clocks 3800, Data net 10GbE, work disk 22Tb, scratch disk 57Gb

CASA (McMullin+ 2007) pipeline rPicard (Janssen+ 2019), modular python-based pipeline using the *fringefit* algorithm (van Bemmel+ 2019)

Imaging: three dedicated image pipelines DIFMAP, eht-imaging, SMILI













Prodotti alle antenne 5 Pby, trasportati su disco e correlate ad ALMA

Trasferire e flussi di dati su lunghe distanze sulle reti

(progetti Express-Nexpres 2000-2010)

Avere collegamenti a 10 Gbit/s non e' sufficiente! Sulle tratte continentali non si e' soli e non si controlla la rete.

- Migliori prestazioni di trasporto con grandi pacchetti (Jumbo Frame 8-9000 Byte)
- Alcuni router deframmentano a 1500 Byte introducendo ritardi e sovraccarico alle CPU
- Le congestioni sui router portano ritardi a trasmissione TCP (salvaguardia dei dati) o perdita di pacchetti in UDP.

Migliori risultati (reti a 10Gbit/s):

- > In produzione 2 Gbit/sec verso Olanda e in Italia (UDP)
- > Esperimenti a 6 Gbit/sec con il Giappone (UDP)
- > Trasferimenti a 0,8 Gbit/s dal Sud Africa (UDP)

Salvare/archiviare i dati

E' necessario disporre di sistemi di grande capacita' e velocita'. Le due cose non vanno sempre d'accordo.

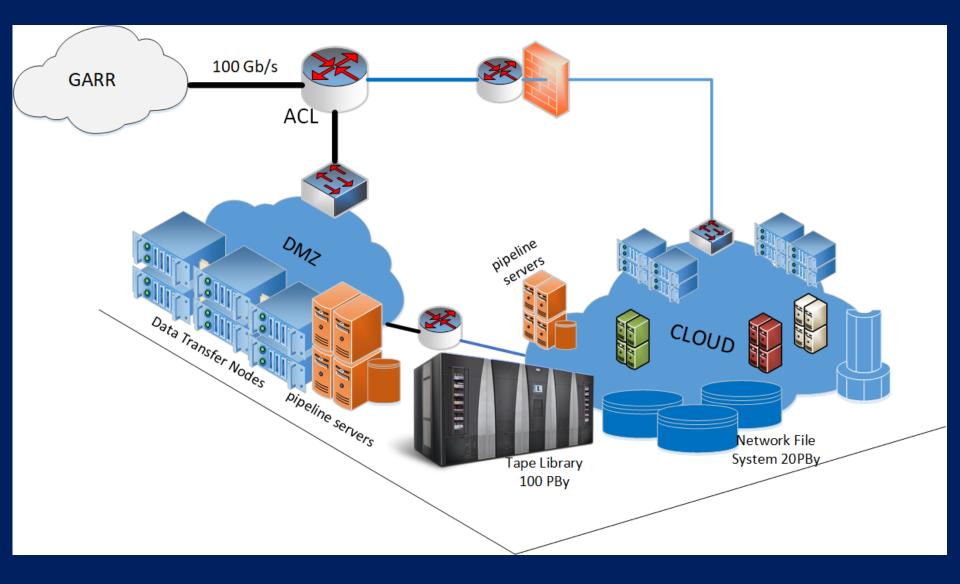
Soluzioni	Velocita' relativa	Costo al TBy	Note
HD	1	30	Grande capacita' 16TBY
SSD	5	120	Capacicta' max 4TBy, in crescita
NVME M2	25	200	Capacita' 2TB. Slot su scheda madre. Sovraccarico di CPU

Soluzioni	GBy/s	Minuti x Scrittura di 1 TBy	Note
Raid 24 Dischi SAS	2.2	8	Flusso costante, bassa latenza
Sistemi ibridi	18	1	Criticita' dimensione dataset
Network FS 40G	4.8	4	Latenza / Concorrenza tra applicazioni

Tape: Costo al Tby : 12 Euro

<u>Scri</u>ttura/Lettura di 1 Tby 15-45min

Un Data Center per SKA e per i maggiori progetti dell'INAF



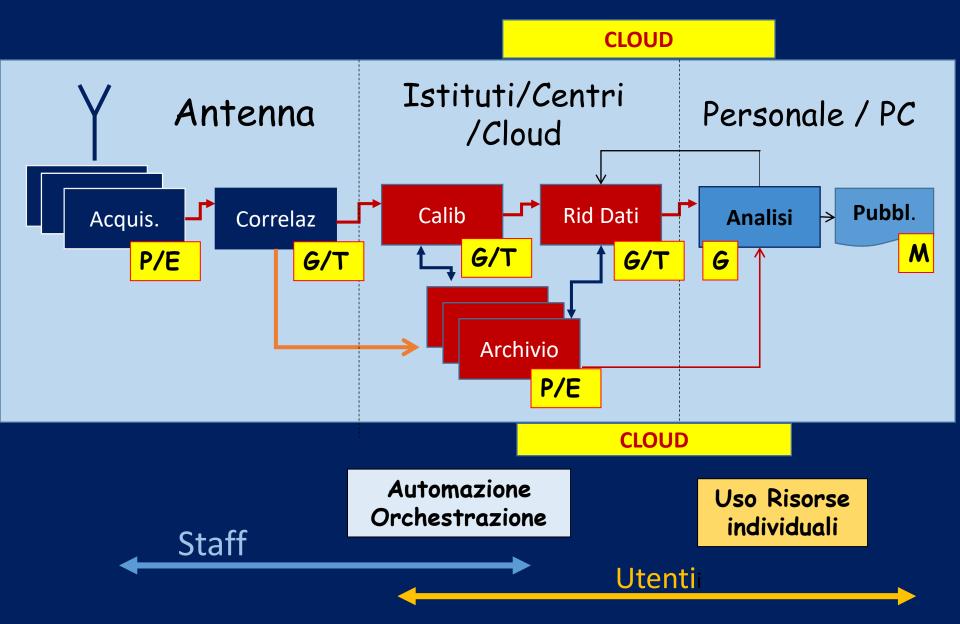
Il Cloud

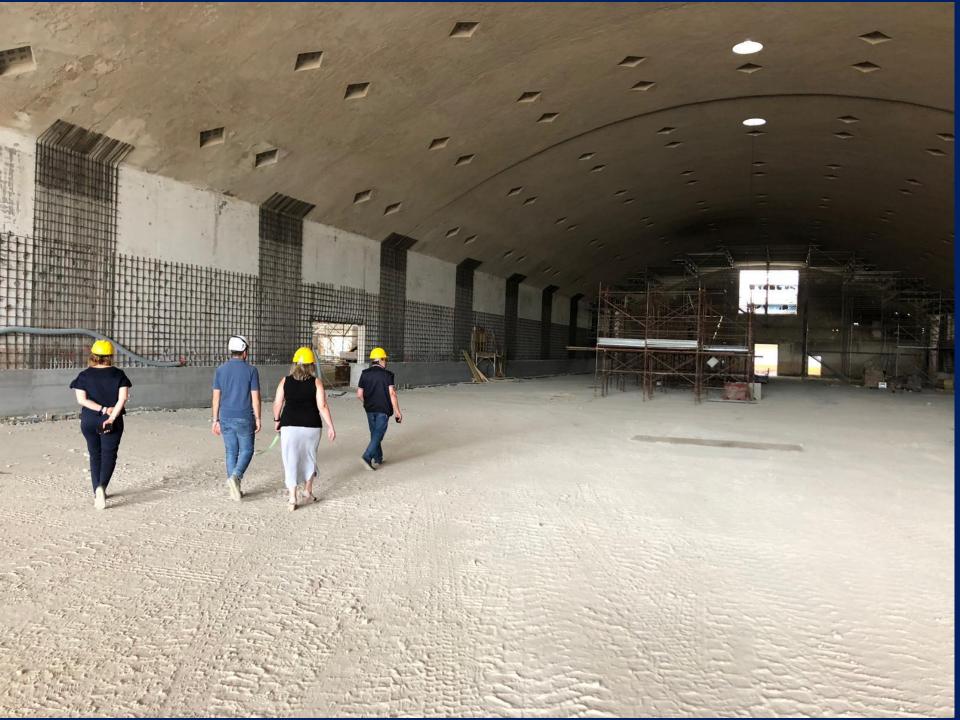
Con il termine **cloud computing** si indica un paradigma di erogazione di servizi offerti on demand da un fornitore ad un cliente finale attraverso la rete Internet (come l'archiviazione, l'elaborazione o la trasmissione dati), a partire da un insieme di risorse presistenti, configurabili e disponibili in remoto sotto forma di architettura distribuita.

Il «provider» deve Pianificare / prevedere:

- > Numero di Utenti
 - > La contemporaneita' di utilizzo risorse
 - > Calcolo: durata singolo task
 - Storage !!
- > Risorse da mettere a disposizione
 - > Eventuale degrado delle prestazioni dai layers software

Flussi e dimensione dei dati





I numeri di SKA sono fuori scala al 2025

Investire per creare e testare un «laboratorio» in cui iniziare a provare:

- Trasferimenti dei dati su reti geografiche a 10/100 Gby/sec
- Sistemi di storage basati su raid di SSD/NVME
- Gestione trasparente delle tape library con dati veri/simulati
- FAT Station (2 TB Mem ... 128 Cores)

Una scala 1/100 di come dovrebbe essere lo ska data center:

Nella sperimentazione usiamo i programmi che abbiamo oggi e che avremo coinvolgiamo «ricercatori veri»,

Contando di NON essere ancora qua

