



Gestione dei dati e Cloud - ReCaS

Roberto Bellotti

Nicola Amoroso

Loredana Bellantuono

Giacinto Donvito

Giorgio Maggi

Dipartimento Interateneo di Fisica “M. Merlin”
Università degli Studi di Bari “Aldo Moro” & Politecnico di Bari
Istituto Nazionale di Fisica Nucleare
(Bari Medical Physics Group) --- <http://medphysics.ba.infn.it>

15 marzo 2019

Indice

- ✓ Il Data Center ReCaS
- ✓ I progetti in essere
- ✓ ReCaS per MUSICA
- ✓ MUSICA per gli obiettivi dello sviluppo sostenibile UN 2030
- ✓ Prospettive
- ✓ Conclusioni

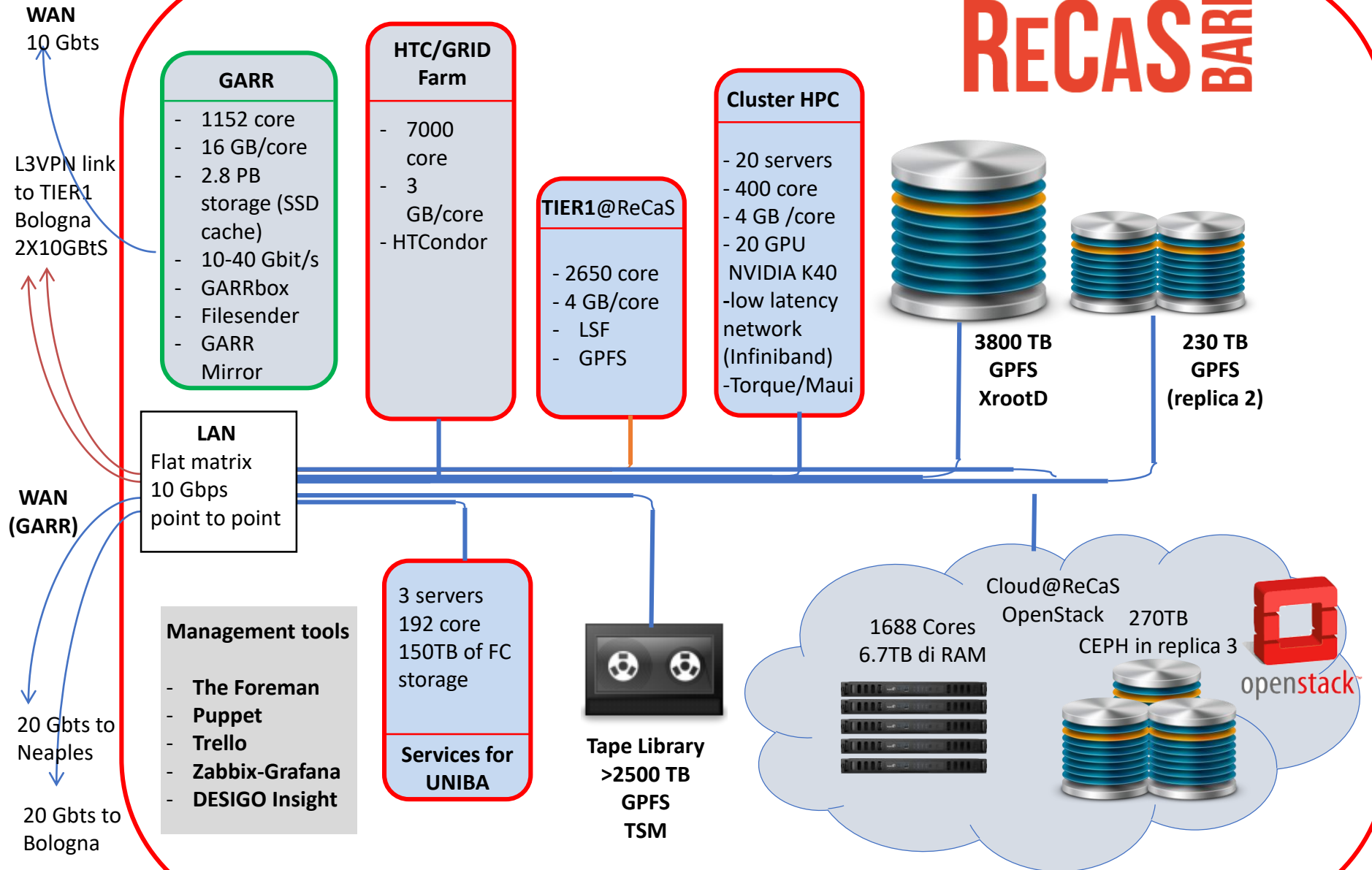
Il Data Center ReCaS @ Bari



9 luglio 2015

ReCaS – Bari

- ~ 300 servers
- ~ 11000 CPU cores (~ 4 GB RAM/core)
- ~ 5 PB storage
- Shared file system
- High availability for data and services
- Tape library
- 10 Gbit/s bandwidth between each two nodes
- 20 GPU NVIDIA K40
- Full inspection enabled firewall



ReCaS – Bari

Cloud@ReCaS Servizi IaaS

- 800 macchine virtuali
- 580 volumi virtuali (~ 180 TB)
- 244 progetti attivi
- 620 utenti

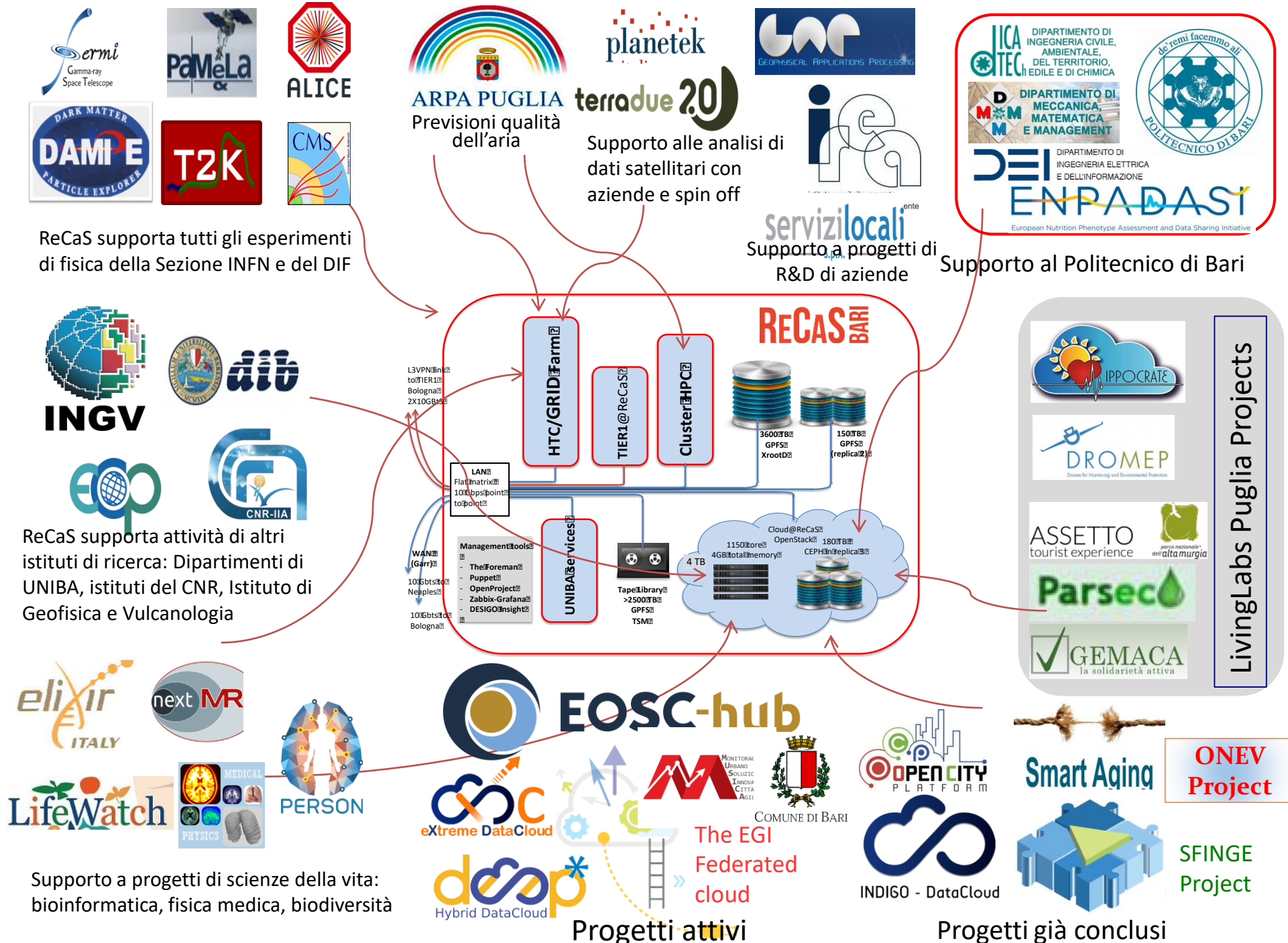


Servizi PaaS e SaaS

- Mesos cluster on-demand
 - Heat/Tosca template
 - Ansible roles and playbook
- Galaxy cluster on-demand
- RStudio on-demand
- Jupyterhub on-demand
- ShareLaTeX on-demand
- GitLab on-demand
- Dropbox-like service based on ownCloud
- Desktop as a Service (web based)
- Wordpress
- Moodle



I progetti in essere



ReCaS per MUSICA – risorse informatiche

11 macchine virtuali, circa 800 giorni di attività computazionale

Macchina virtuale	Data di avvio	Attività computazionale (giorni)*	RAM (GB)	CPU (core)	Storage (GB)
musica-bl	2017-02-28	743	16	8	50
musica-sgap	2017-02-16	755	16	8	500
musica-test	2017-02-01	770	32	16	1024
musica-pg	2017-01-23	779	16	8	500
musica-gs	2017-01-17	785	16	8	500
musica-es	2017-01-17	785	16	8	500
musica-vis	2017-01-17	785	16	8	50
musica-hb	2017-01-16	786	16	8	500
musica-dr	2017-01-13	789	16	8	50
musica-dss	2017-01-12	790	16	8	50
musica-das	2017-01-11	791	16	8	50

*dati aggiornati a dicembre 2018

ReCaS per MUSICA – dataset disponibili



SICUREZZA

Statistiche sentiment,
Anagrafe, BaRisolve,
Percezione Sicurezza,
Interventi Polizia
Municipale, Decoro
Urbano...



AMBIENTE

Cruscotto orario
PM2.5 e PM10, Indice
di Qualità dell'Aria,
Rumore, Aree Verdi,
Inquinamento,
Meteo...



ENERGIA

Dataset AMGAS –
Consumi, Dataset
ENEL – Consumi,
Emissioni CO₂,
Manutenzione
impianti illuminazione
pubblica...



MOBILITÀ

Dataset AMTAB –
Dettaglio percorrenze
e corse, Consumi
carburante, Traffico e
emissioni CO₂,
Biglietti...

Città e comunità sostenibili

11° obiettivo dell'Agenda 2030 delle Nazioni Unite per lo sviluppo sostenibile

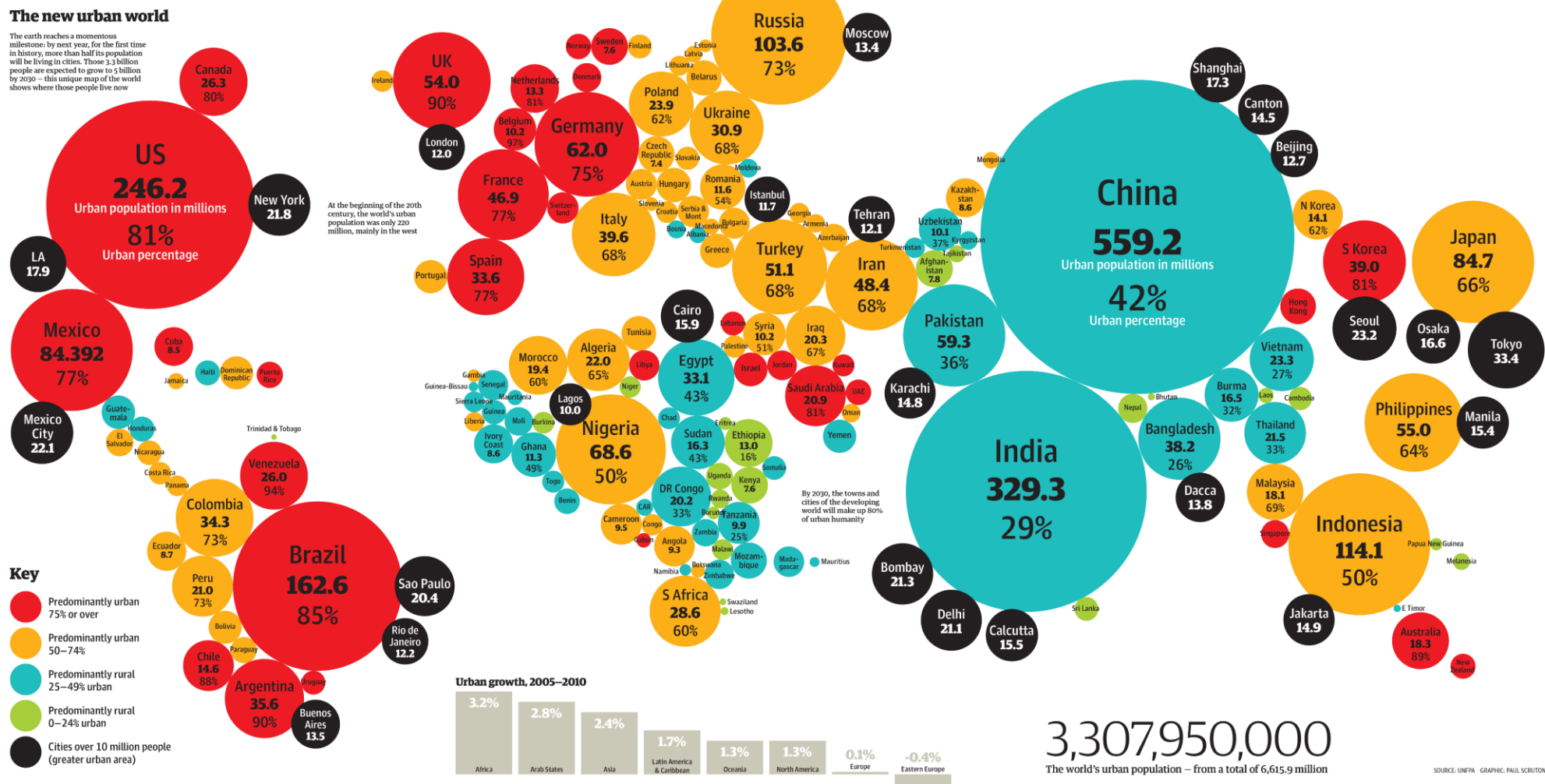


rendere le città e tutti gli insediamenti urbani

- ✓ inclusivi
- ✓ sicuri
- ✓ resilienti
- ✓ sostenibili

MUSICA per gli obiettivi dello sviluppo sostenibile UN 2030

Popolazione urbana nel mondo



MUSICA per gli obiettivi dello sviluppo sostenibile UN 2030

Emergenze della vita cittadina

Dall'Agenda 2030 delle Nazioni Unite per lo Sviluppo Sostenibile:

Disuguaglianze sociali

833 milioni di individui vivono in **periferie degradate**

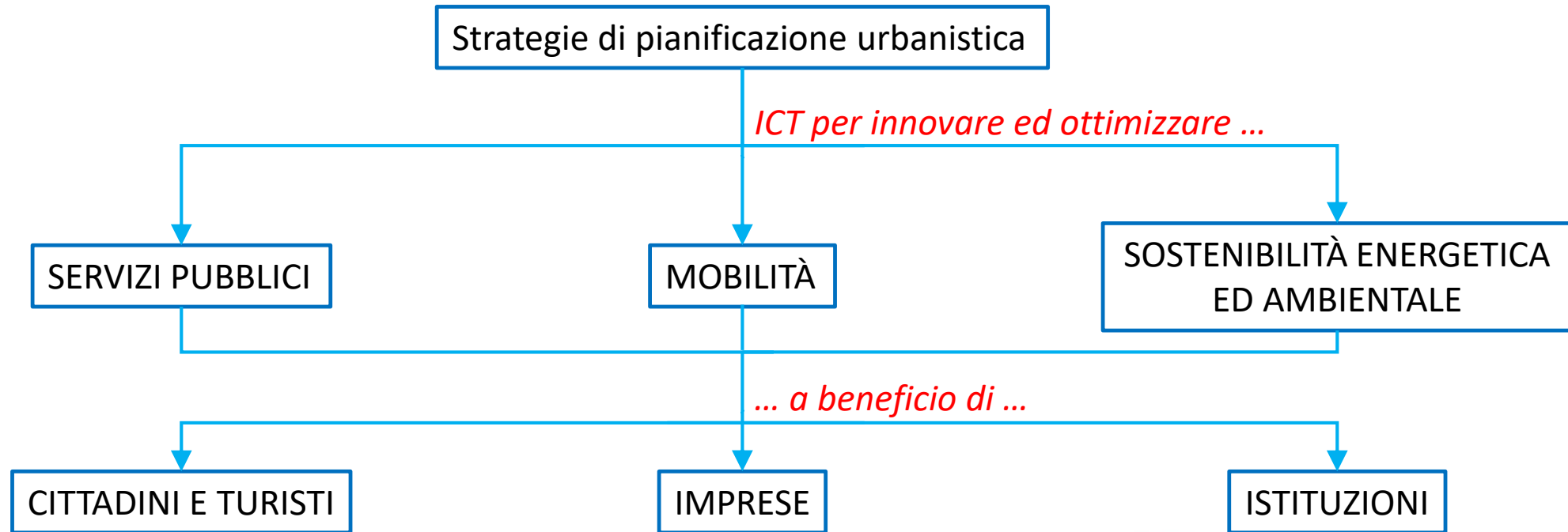


Sostenibilità energetica e inquinamento

le città occupano appena il **3%** della **superficie terrestre** ma coprono il **60-80%** del consumo di energia e il **75%** delle emissioni di carbonio.



Smart cities



Mobilità sostenibile

Efficientamento delle reti di trasporto pubblico urbano

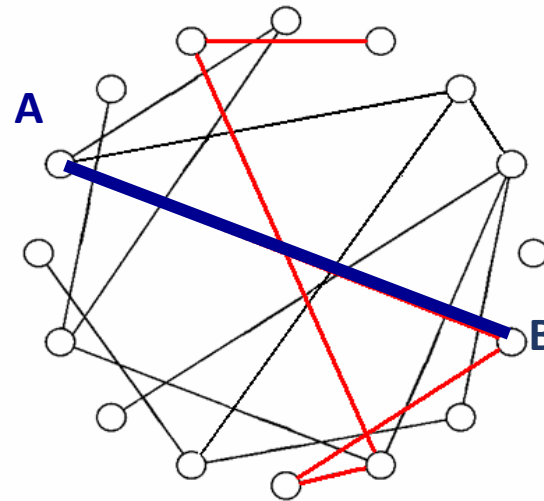
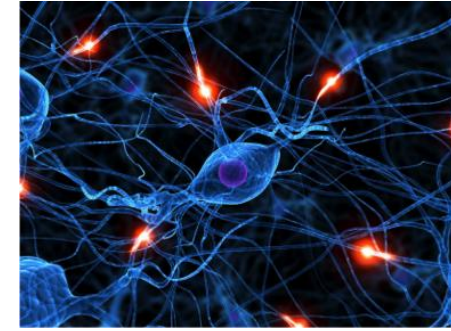
- ➡ contrasto all'**isolamento** delle **periferie**
- ➡ **riduzione** dei flussi di traffico e dell'**inquinamento** che ne deriva

Ruolo del Comune di Bari ➡️ azioni nell'ambito di **MUSICA**: analisi delle linee degli **autobus urbani** riguardo a luoghi/tempi di transito e all'utenza, monitoraggio dei consumi di carburante e delle emissioni di CO₂



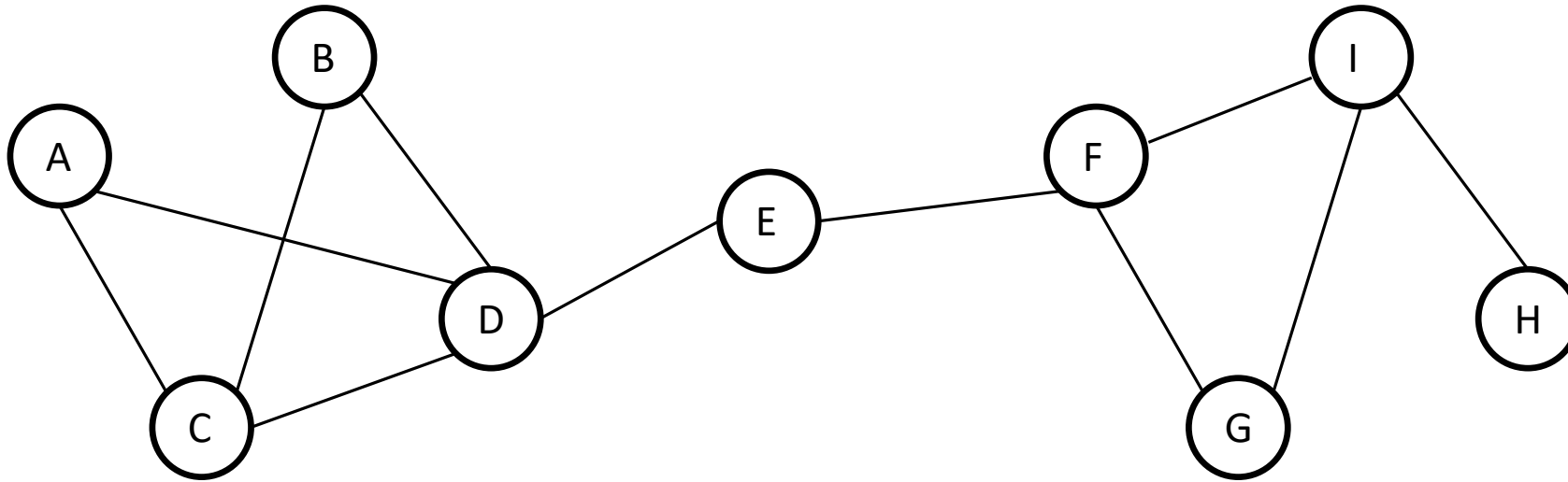
Big data tra loro connessi → reti complesse

- Trasporti
- Reti elettriche
- Mercati finanziari
- Sistemi biologici



e.g. La nazione **A** è collegata alla nazione **B** se **A** compra/vende un prodotto da **B**.

Centralità di un nodo



Node	Degree
A	1
B	2
C	3
D	4
E	2
F	3
G	3
H	2
I	3

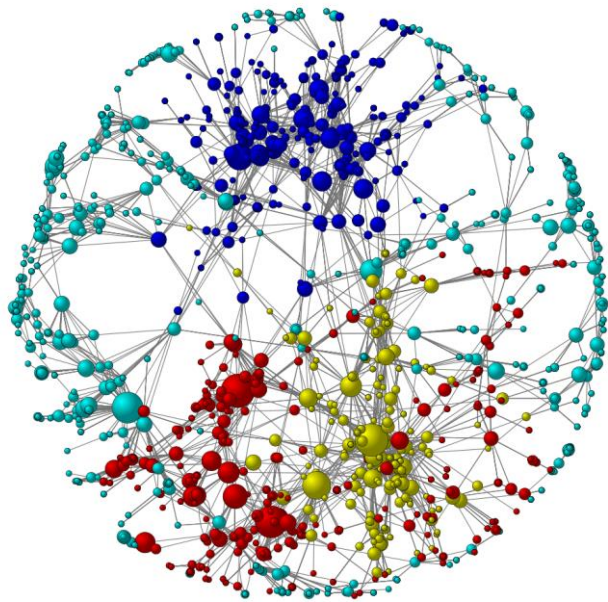
Node D has higher degree (4), yet is it the most important node of the network? A centrality measurement is also **betweenness**.

Betweenness measures the number of shortest paths within the network passing through each node.

Node	Betweenness
A	0
B	0
C	0.5
D	15.5
E	16
F	15.5
G	0
H	0
I	7

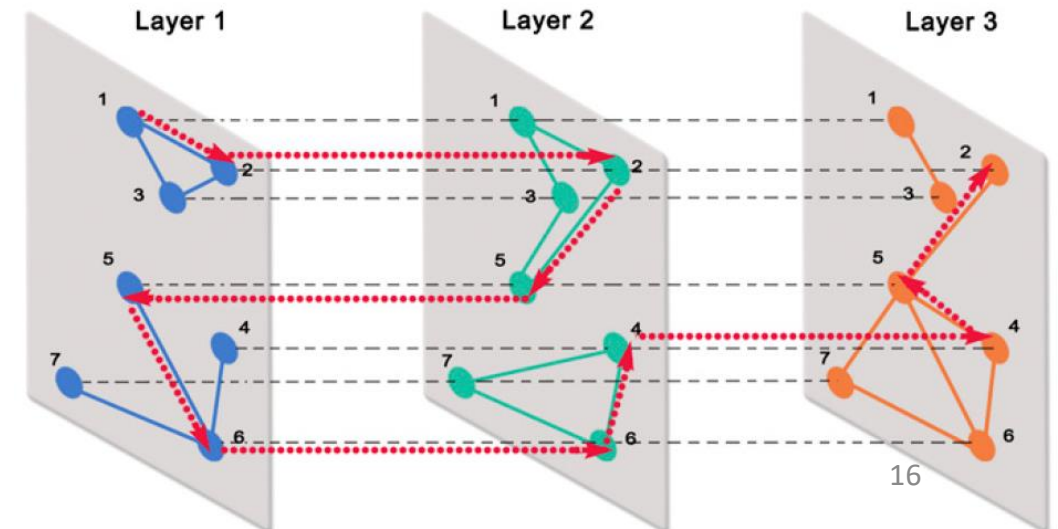
Complex Networks

Molti sistemi fisici, biologici e sociali sono costituiti da più unità elementari interagenti tra loro. Questa struttura può essere modellizzata mediante **reti complesse**: le singole componenti del sistema sono rappresentate come **nodi** e le relazioni tra di esse come **link**



Due nodi possono essere collegati tra loro con relazioni di diversa natura

➡ **Multiplex**, formate da più **reti complesse** disposte su strati diversi



Multiplex per il trasporto multimodale

Modello **multiplex** per il **trasporto multimodale** ➡ descrizione di centri urbani caratterizzati dalla **coesistenza di più reti di trasporto pubblico**

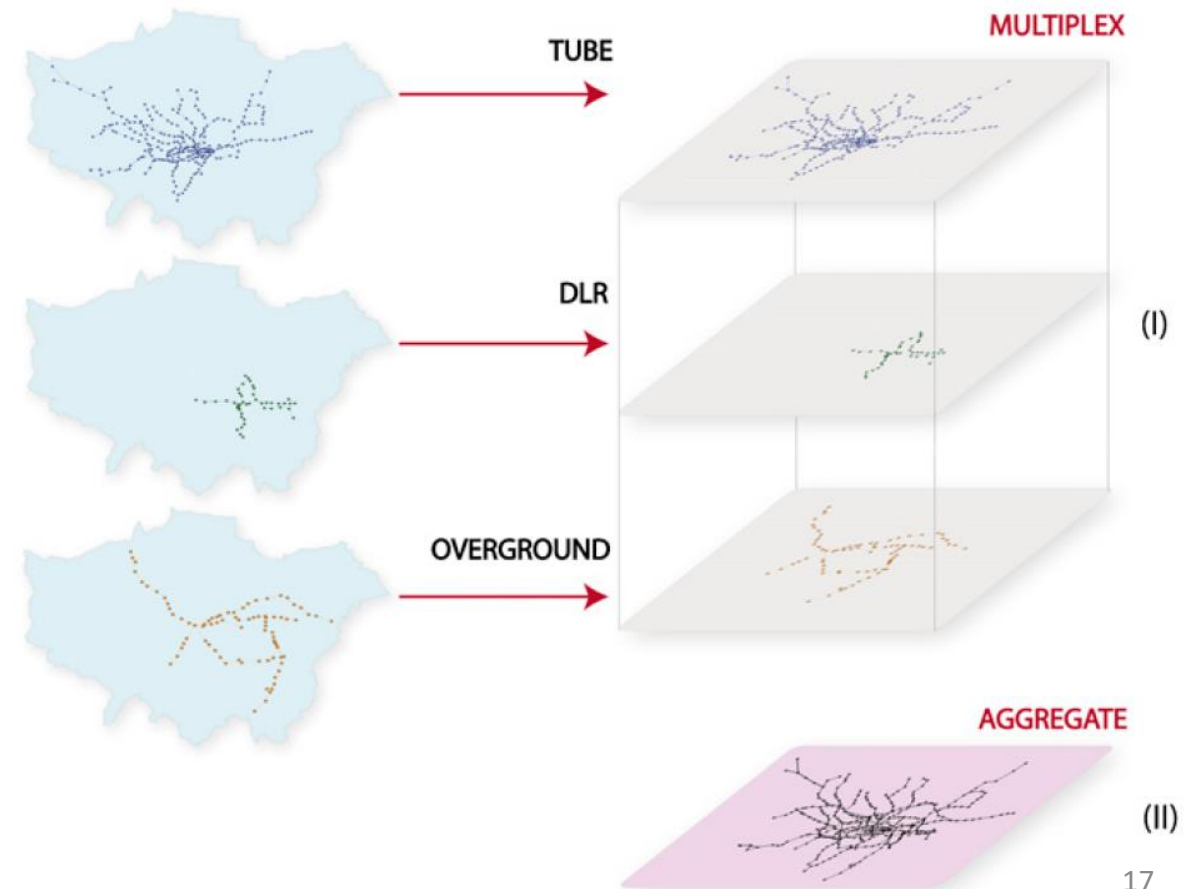
In ciascuno strato

Nodi ➡ fermate/stazioni delle reti di trasporto pubblico

Link ➡ percorsi del trasporto pubblico che connettono i nodi

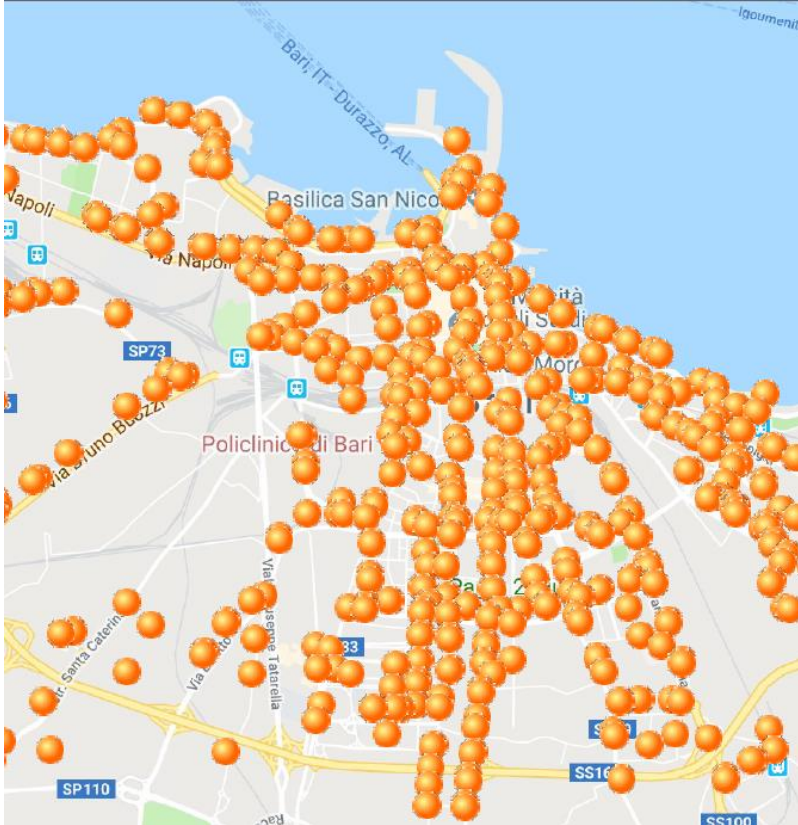
Link tra strati: collegamenti fra fermate/stazioni di reti diverse, separate da una distanza minore di 100 m.

➡ **Costruzione della rete aggregata**

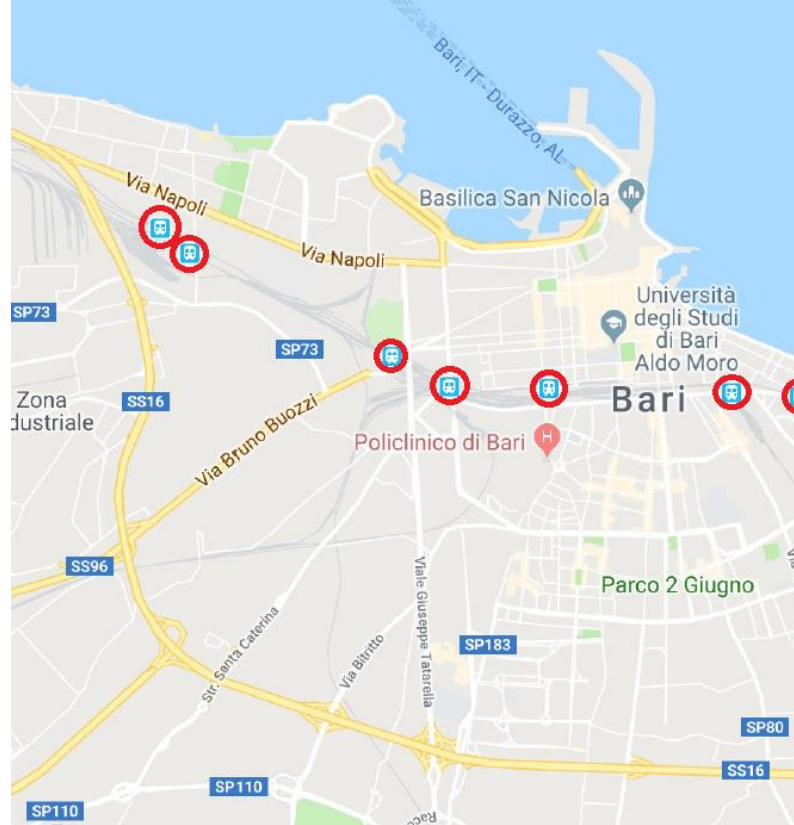


Reti di trasporto pubblico a Bari

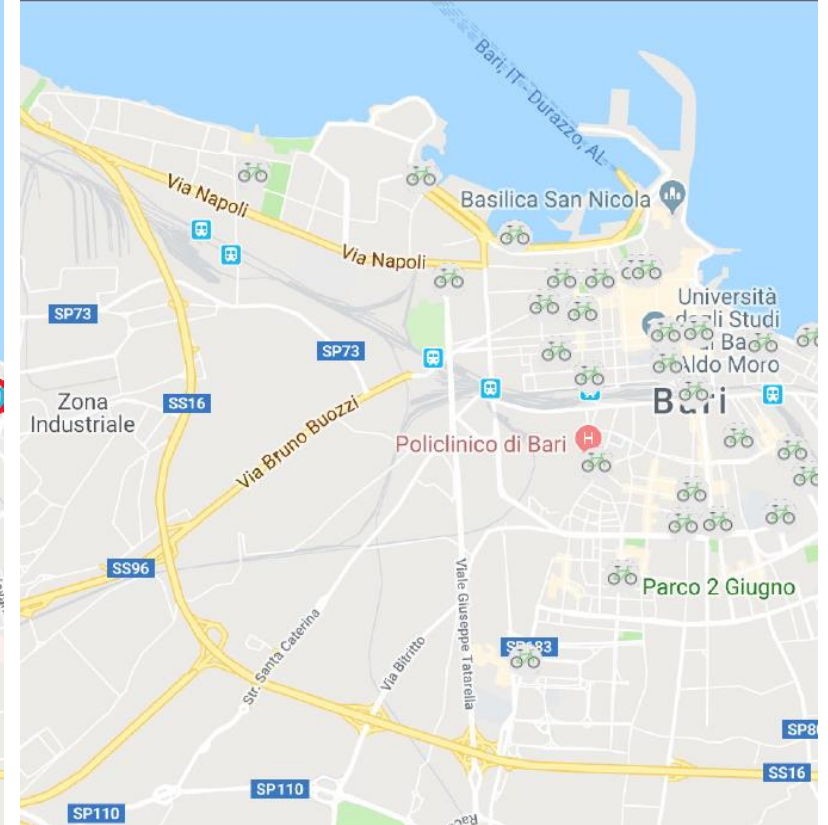
AUTOBUS



TRENI



BIKE-SHARING



Reti di trasporto pubblico a Bari: dataset usati

AUTOBUS

Fonte: opendata.comune.bari.it

ID Fermata	Descrizione Fermata	Latitudine	Longitudine
03393101	Viale O. Flacco, di fronte Civ. 4/A	41,1107	16,8618
06006001	Via Alberotanza, 5	41,0958	16,8718
05399002	Via Oberdan INPDAP	41,1177	16,8817
10208101	Via Delle Orchidee-ex Stanadyne	41,1013	16,7849
05510001	via Toscanini, 20	41,1029	16,9115
05500101	via Suglia, 4	41,1037	16,9128
03462001	S.P. Carbonara/Modugno BARIDOMANI	41,0883	16,8276
01143103	Via D'Annunzio svincolo Aeroporto Civile	41,1377	16,7785
01387101	Via Nicholas Green,54	41,1569	16,7416
10207102	Via Delle Ortensie, di fronte Pimespo	41,109	16,7737
04483001	Via S. Angelo, 39	41,0656	16,8660

BIKE-SHARING

Fonte: opendata.comune.bari.it

Denominazione	Latitudine	Longitudine	Numero Bici
Agraria	41,1114	16,8829	10
Area Sosta Mazzini	41,1249	16,8559	0
Area sosta Rossani	41,1162	16,8719	10
Ateneo	41,1212	16,8695	6
Camera di Commercio	41,1242	16,8722	10
Chiesa Russa	41,1092	16,8719	10
Cimitero	41,1233	16,8476	10
Città Vecchia	41,1273	16,8660	8
Comando Polizia Municipale	41,1058	16,9126	10
Economia	41,0941	16,8546	9
Garibaldi	41,1258	16,8620	9

TRENI

Fonti: trenitalia.com , ferrovienordbarese.it ,
ferrovieappulolucane.it

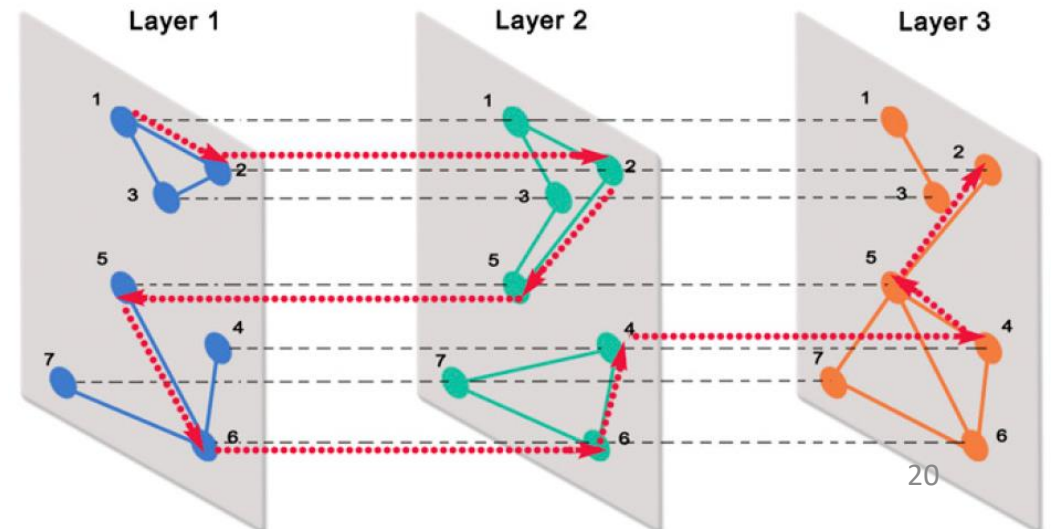
Resilienza della rete di trasporto pubblico

Analizzare la **robustezza della rete** rispetto a diversi processi:

- a) rimozione di nodi in maniera **casuale**
- b) rimozione di link in maniera **casuale**
- c) rimozione di **più nodi** secondo un **ordine sistematico**
- d) rimozione di **più link** secondo un **ordine sistematico**

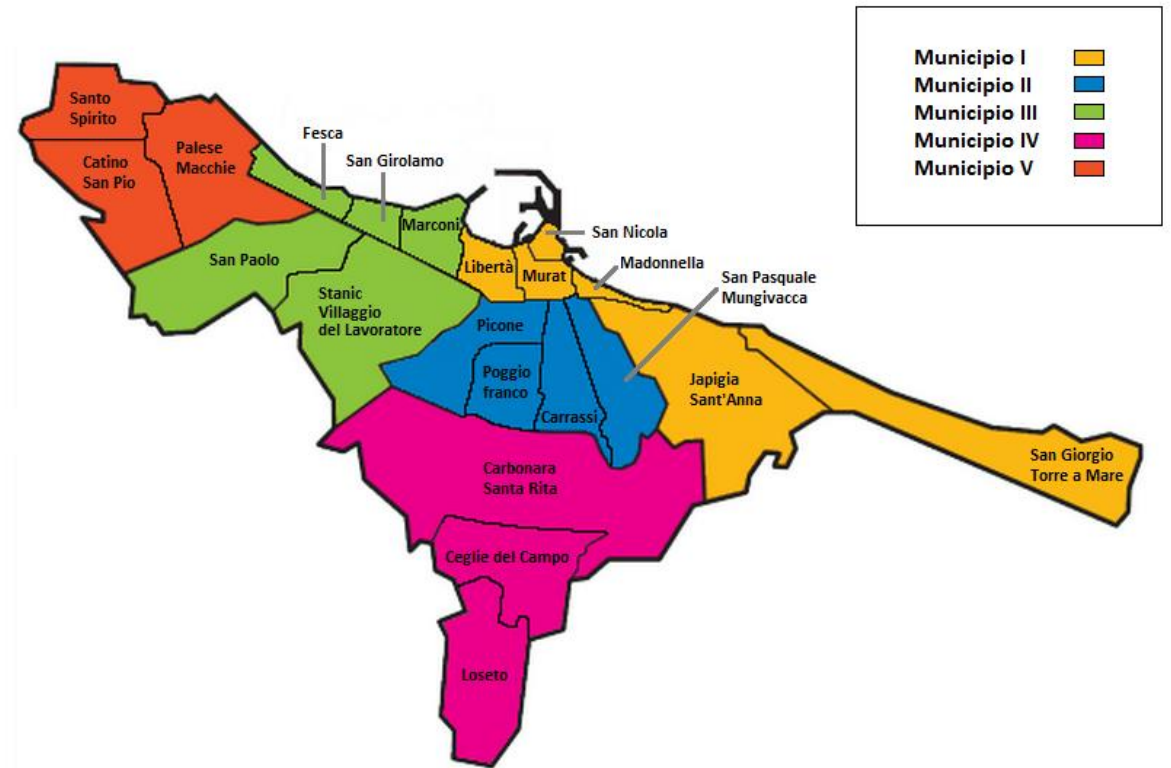
In un multiplex, i link relativi agli strati non coinvolti nella rimozione sopravvivono.

➡ Le **reti di trasporto ibrido** hanno una maggiore resilienza.



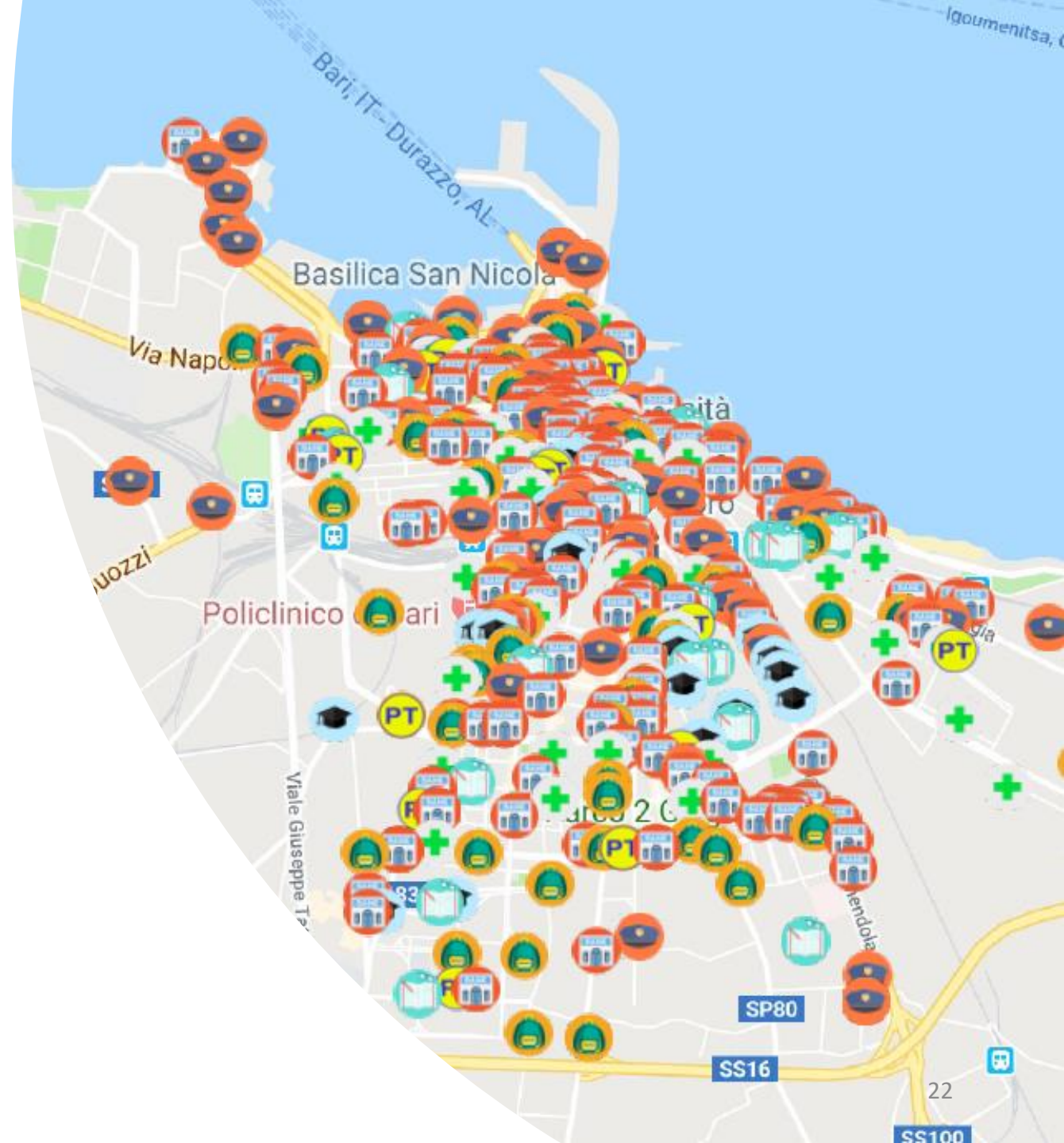
Resilienza della rete di trasporto pubblico

- La città di Bari presenta **quartieri periferici**, collegati al centro da **stazioni/fermate** e **collegamenti strategici**, la cui importanza è misurata dalla **betweenness**.
- Rimozione di nodi e link con **betweenness** più alta: impatto del malfunzionamento degli elementi strategici sulla mobilità cittadina



Prospettive

Aggiungere alla rete dei nuovi **nodi**, corrispondenti ai **luoghi di interesse pubblico** (scuole, ospedali, asl, uffici postali), in modo da fotografare l'**efficacia** del **sistema di trasporto** e individuare nuove **connessioni strategiche** da introdurre o rafforzare.



Conclusioni

- ✓ Il Data Center ReCaS è in funzione h24 dal 2015 (1 sola interruzione!)
- ✓ Fornisce risorse di calcolo e memorizzazione dati con alta affidabilità a molte comunità scientifiche e imprese
- ✓ Il gruppo di lavoro ReCaS è in grado di supportare altri gruppi di ricerca anche per attività di analisi di dati in svariati settori, in particolare su tematiche legate al machine learning e big data
- ✓ Entro pochi mesi aumenterà in modo sostanziale le risorse disponibili tramite i tre progetti PON appena sottomessi

Grazie per l'attenzione

roberto.bellotti@uniba.it

3386564596

Network Science: Introduction 2012