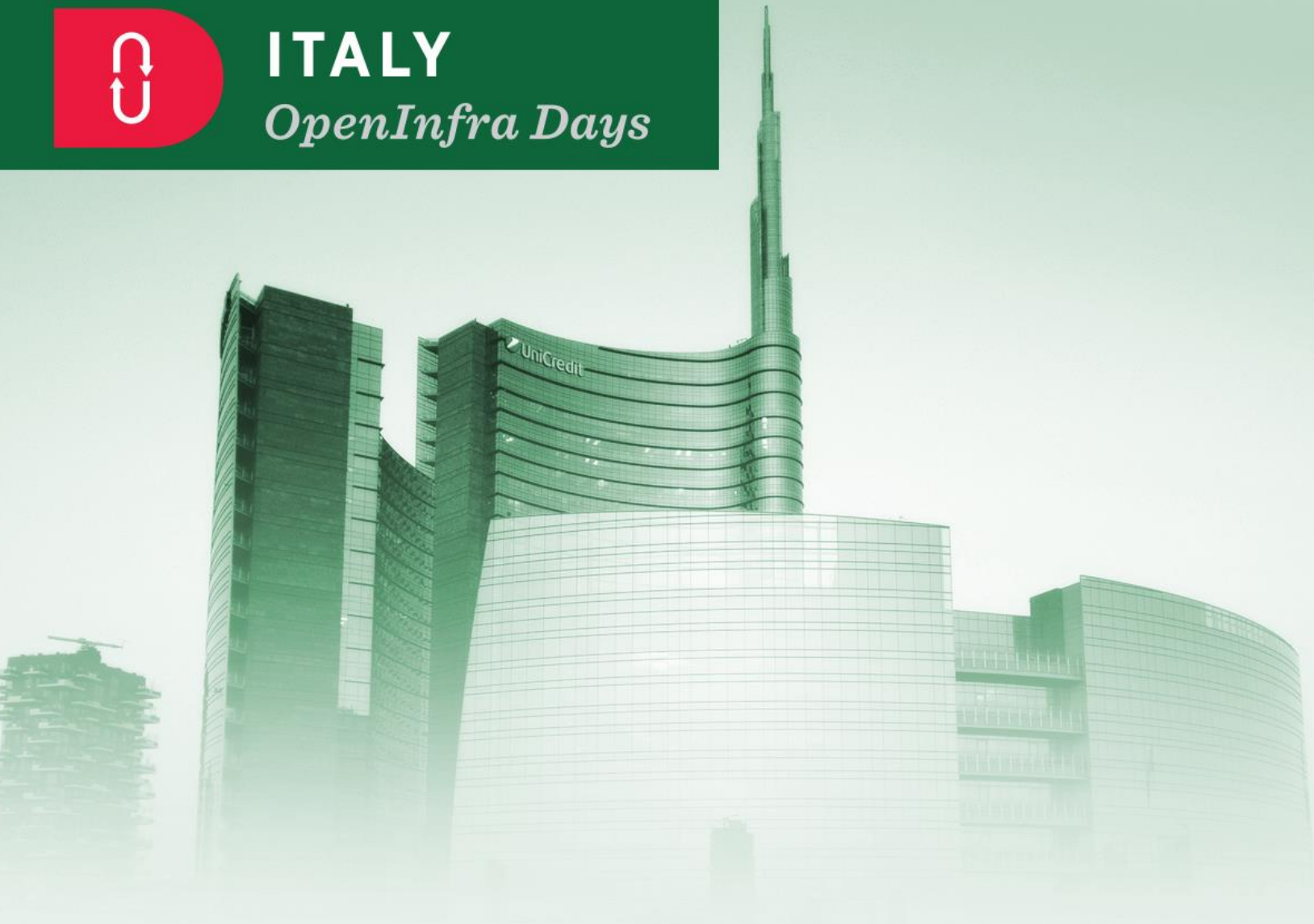
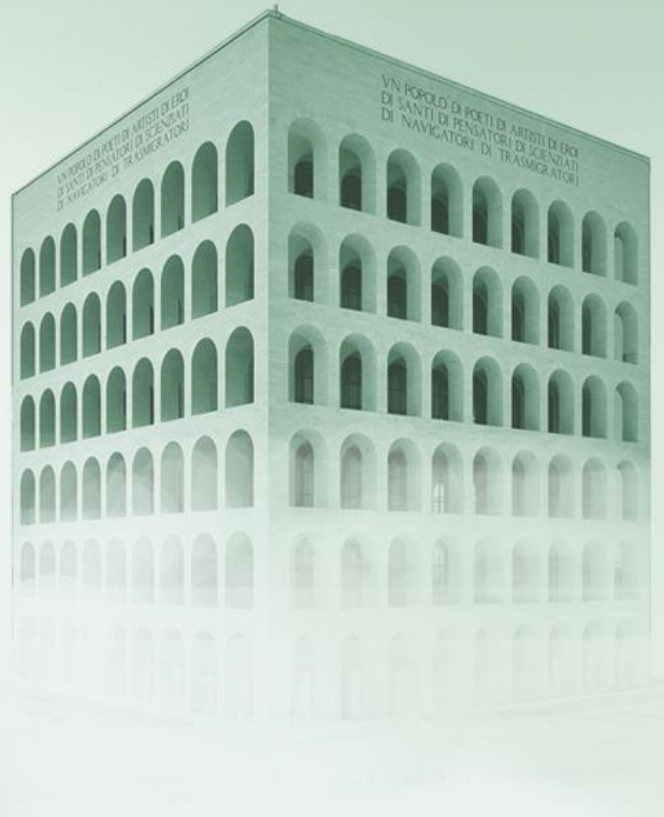
# ITALY
## *OpenInfra Days*

OPENINFRA DAYS
ITALY




**Organized by**

IRIDEOS

Binario Etico

**Under the patronage of**

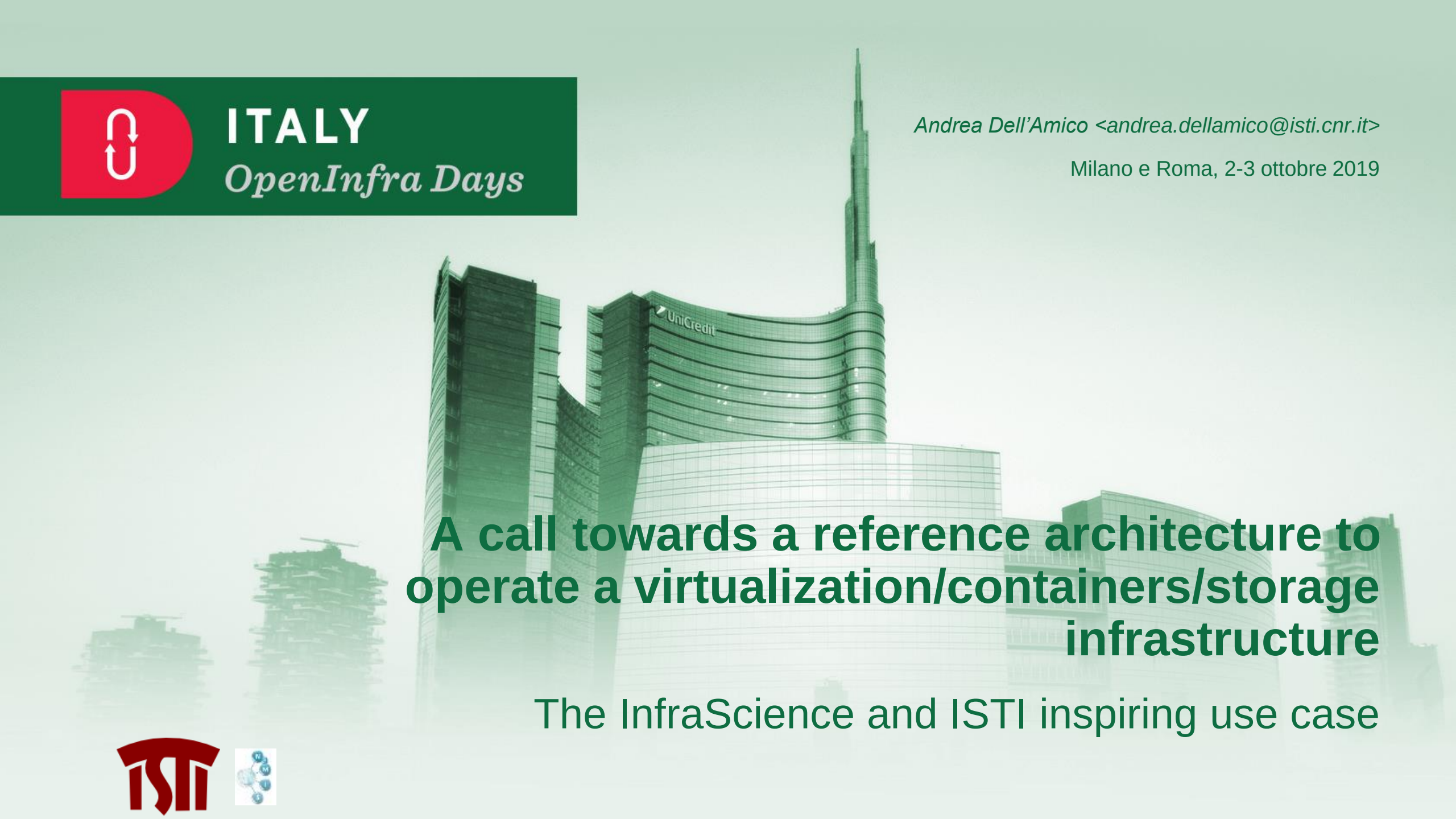
AGID | Agenzia per l'Italia Digitale

**Sponsored by**

Mellanox TECHNOLOGIES

MESOSPHERE

gci SERVICE FACTORY




ITALY
OpenInfra Days

*Andrea Dell'Amico <andrea.dellamico@isti.cnr.it>*

Milano e Roma, 2-3 ottobre 2019

# A call towards a reference architecture to operate a virtualization/containers/storage infrastructure

## The InfraScience and ISTI inspiring use case

# ISTI: Istituto di Scienza e Tecnologie per l'Informazione

- The biggest institute of the National Research Council (CNR)
- Thematic areas: *Networking, Software, Knownledge, High Performance Computing, Visual, Flight and Structural Mechanics*
- 12 research groups, many of them host more than one research group
- *Each research group manages its own hardware and services: plenty of security problems, abandoned hardware, under-utilised resources*

# D4Science.org

Integrated technologies that provide elastic access and usage of data and data-management capabilities

- Virtual Research Environments (VRE) that give access to multiple services
- Data discovery, accessing, analysis, and transformation in standard format
- Powered by gCube: https://www.gcube-system.org/

# OpenAIRE, https://www.openaire.eu

Operate a pan-European (and global) network for Open Science to articles and research data across countries and across research communities

- Definition and dissemination of guidelines for sharing scholarly products and links between them
- Provide services for populating and provide to the public an information graph of interlinked scholarly entities
- Provide services for assessing research impact of funders (the Commission in primis) and monitoring of open access trends
- Powered by D-Net: http://www.d-net.research-infrastructures.eu/





# D4Science + OpenAIRE infrastructure

- 500+ VMs, ~100 of them hosted by the GARR cloud
- 75% of them are production services (they must work)
- Migration, a year ago, from Xen/Aoe to OpenStack/Ceph





# ISTI infrastructure (1)

- It does not exist (yet):
  - Some services run on stand alone hypervisors
  - Mail + DNS run on legacy hardware
  - A lot of *unknown* hardware from different research groups
  - The commercial providers are very expensive when their usage increases





# ISTI infrastructure (2)

- We are building one. The goals:
  - More reliability
  - More efficient use of the existing hardware
  - More control on what's running
  - *The EU is asking for experiment replicability, data privacy, availability of code and data:*

  *this means more computing power, more storage capacity*

# A new migration, old problems

- Modern computing and storage infrastructures are still very complex and hard to maintain
- Italian research facilities are always understaffed (and underfunded)
- Administrative bureaucracy: inability to buy all the needed hardware to start a new infrastructure at the same time

# Reference architecture?

- Different hardware setups (10/50 hypervisors, 100s hypervisors?)
- Different usage scenarios: compute intensive, I/O intensive, general purpose? (most requests are just for standard VMs, with java/python/PHP web services)
- OpenStack and Ceph. Different software setups, eg:
  - Dedicated computing nodes
  - Dedicated storage nodes
  - Hyperconverged nodes





# Reference installation best practices?

- Documentation
- Working examples of provisioning of a complete infrastructure
- Backup practices
- Disaster recovery
- Upgrade (especially OpenStack, Ceph is usually less problematic)

# How we are going to contribute

- Document our installation/upgrade processes
- Free our provisioning tools (mostly ansible roles that complement the official tools)
- Document our configuration choices
- Document our hardware architecture and our architectural choices





# And then?

- Federation?
- Sharing of troubleshooting experiences
- Hybrid cloud setups
- Integration of the newest technologies:
  - Container orchestrators (yes, they are still a *new thing* in our world)
  - PAAS
  - SAAS
  - Lambda
  - …





# Questions?

My name: Andrea Dell'Amico <andrea.dellamico@isti.cnr.it>

The group I'm part of is reachable at s2i2s@isti.cnr.it