



Interactive Realtime Multimedia Applications  
on Service Oriented Infrastructures

# Interactive Realtime Multimedia Applications on SOIs

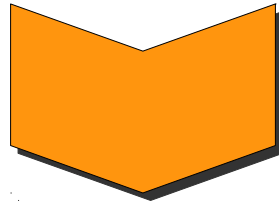
## Advance Reservations for Distributed Real-Time Workflows with Probabilistic Service Guarantees

**Tommaso Cucinotta**  
Real-Time Systems Laboratory  
Scuola Superiore Sant'Anna  
Pisa, Italy

**Kleopatra Kostanteli, Dora Varvarigou**  
National Technical University of Athens  
Athens, Greece

# Introduction

- High availability of broadband connections at affordable rates



- New paradigms of computing
  - **Distributed** computing
  - Not only **best-effort** remote access
  - But also remote **real-time interaction**

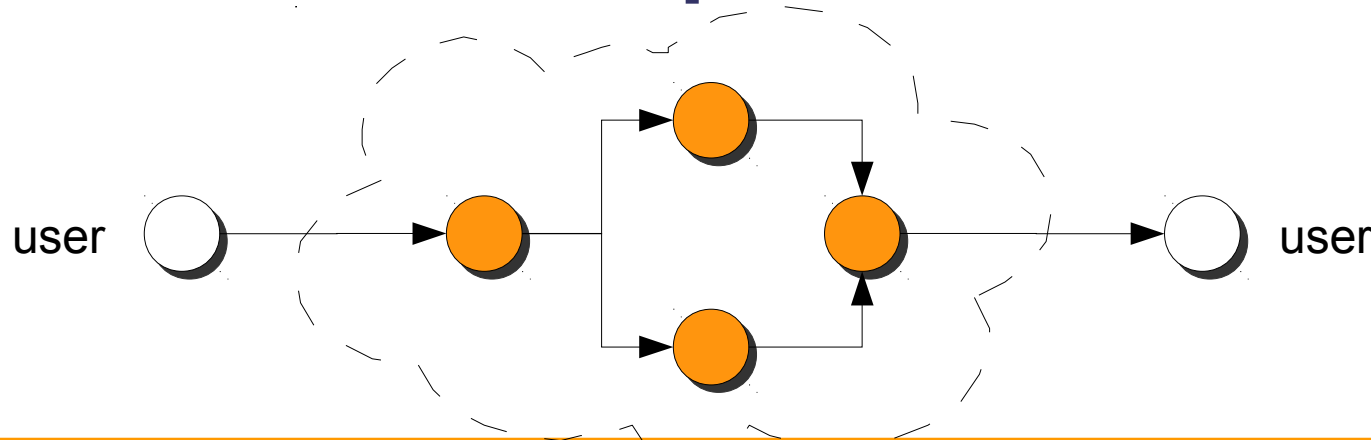
- New business models
  - From provisioning of network bandwidth
  - To provisioning of **distributed services** and **applications**
    - With **real-time/QoS constraints**
- User perspective/expectations
  - From buying costly equipments
  - To **renting** computing power, storage and services at **affordable rates**

- Provider perspective/expectations
  - High equipment (and infrastructure development) costs covered by renting them to **thousands** of users
- Resource management policies
  - High resource saturation levels
  - Overbooking strategies
    - Exploiting **statistical knowledge** about **actual usage** of services by users

# Problem presentation

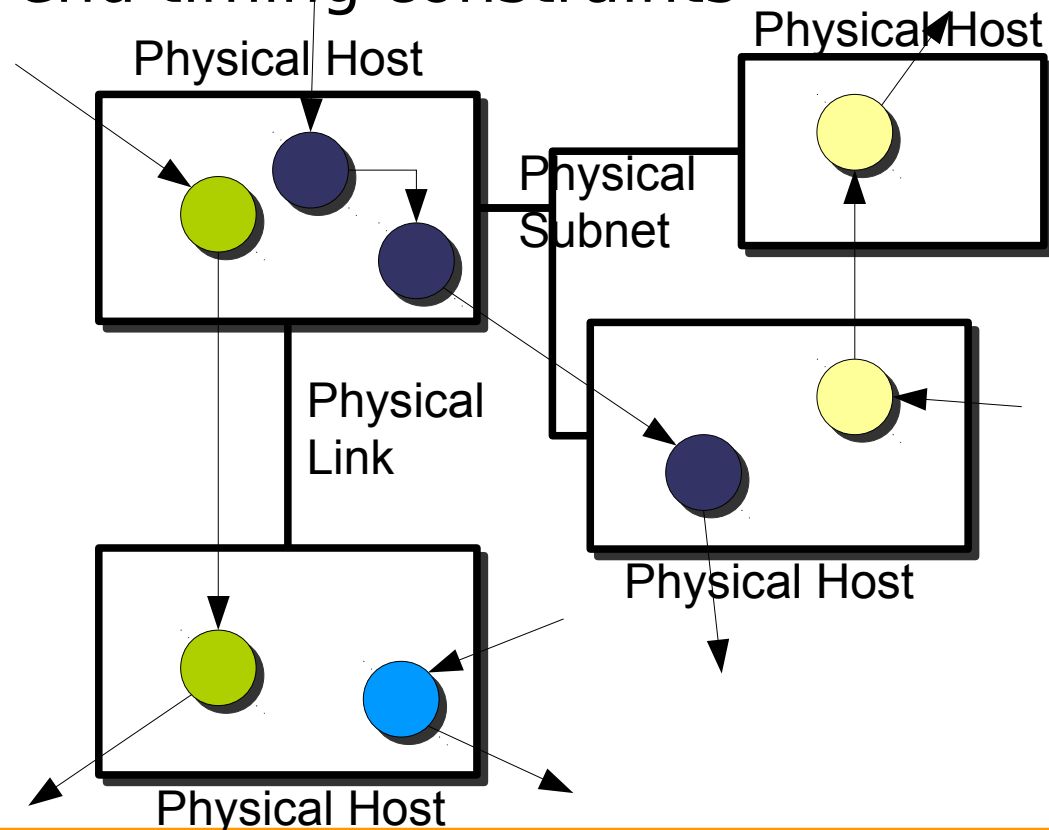
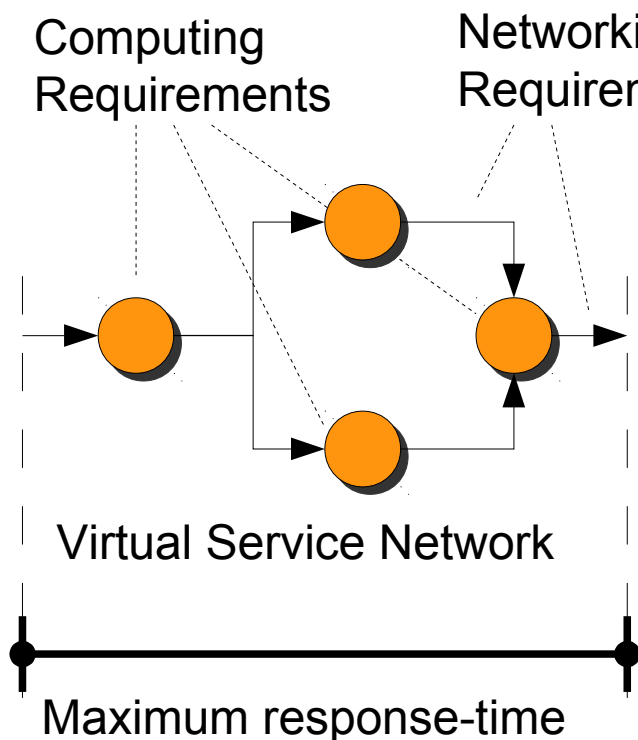
# Problem presentation

- Distributed real-time interactive applications characterised by:
  - **Periodic activation** of a distributed workflow
  - Low resource saturation levels
  - **End-to-end response-time constraints**



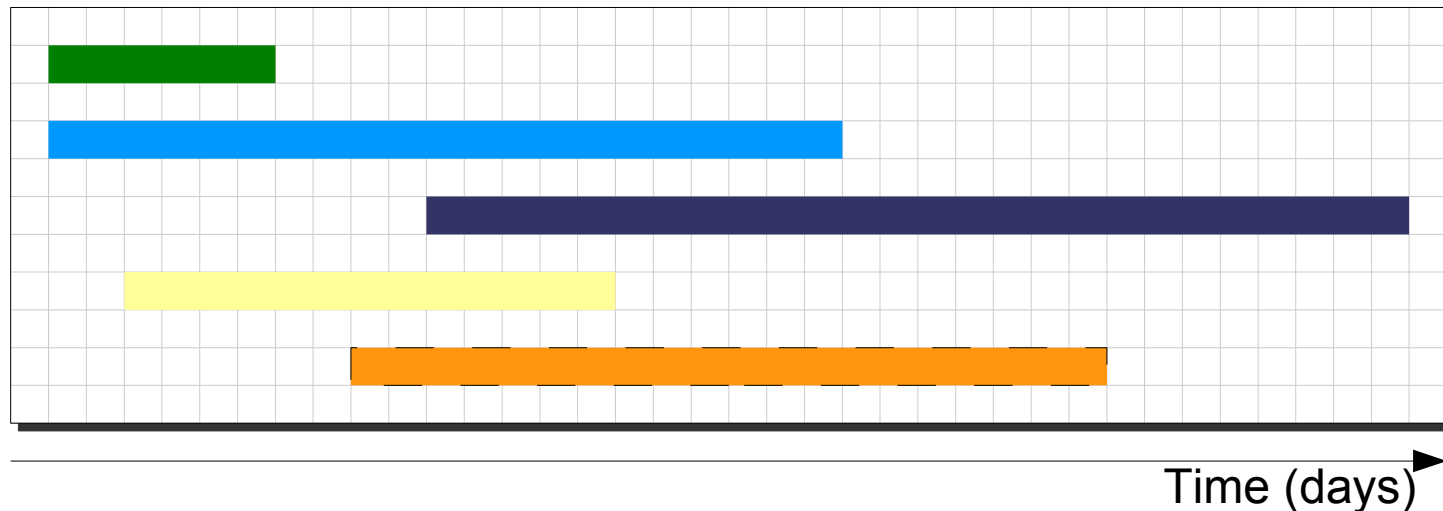
# Problem presentation

- Optimum deployment of VSNs on PNs
  - Given computing/network requirements
  - Respecting end-to-end timing constraints



# Problem presentation

- Optimum deployment of VSNs on PNs
  - Considering expected usage time-horizon (**advance reservations**)
  - Periods of overlapping reservations



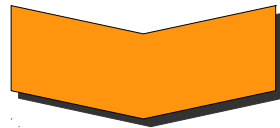
# Envisioned approach

- **Temporal isolation** among independent application workflows
  - Time-sharing of computing nodes
    - Through **real-time scheduling** at the OS/kernel level
  - Time-sharing of network links
    - Through **QoS-aware scheduling** of the medium (e.g., Wf<sup>2</sup>Q+)

- Widely available **POSIX schedulers**
  - Priority-based
  - No temporal isolation  
(high-priority tasks may arbitrarily delay low-priority ones)
  - Theoretical 69% utilisation bound  
(for real-time tasks)
- **IRMOS real-time scheduler**
  - Hierarchical deadline/priority-based
  - Provides temporal isolation/enforcement
  - Theoretical 100% utilisation bound

# Proposed approach

- Focus on **soft real-time** applications

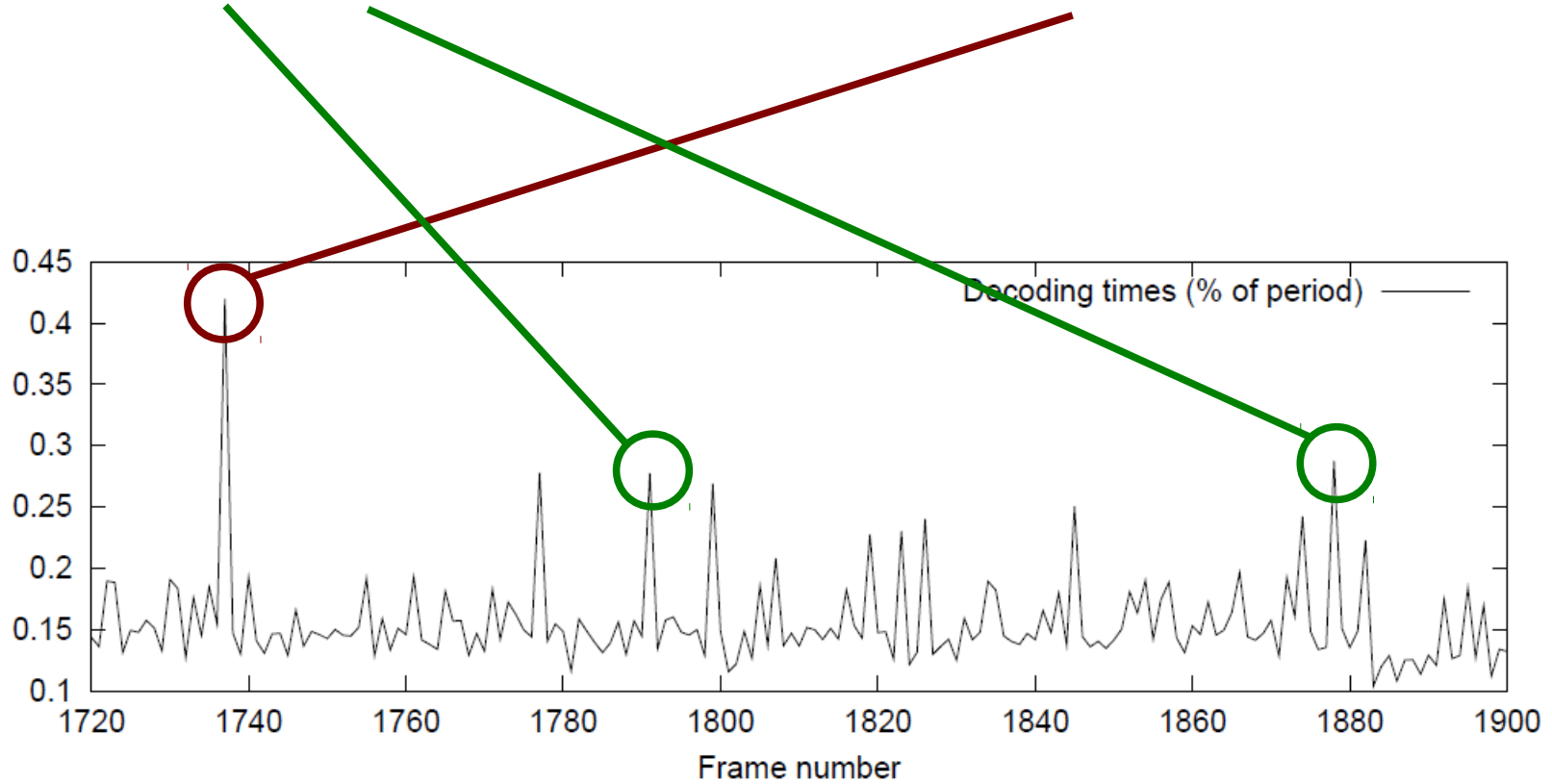


- Probabilistic guarantees
  - **Response-time** guarantees
    - Minimum probability of respecting the end-to-end deadline constraint (vs deterministic, WCET-based)
  - **Availability** guarantees
    - Minimum probability of finding the resources available when actually activating the service

# Probabilistic response-time guarantees

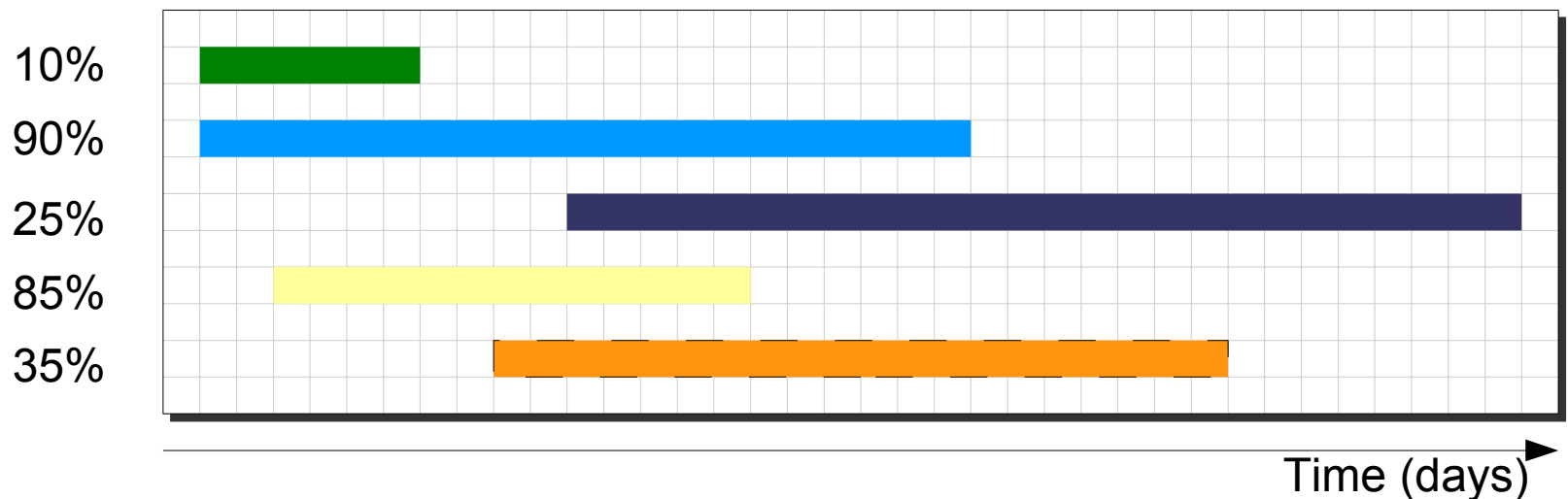


- Tune allocation on computation-time **percentile** (instead of WCET)



# Probabilistic availability guarantees

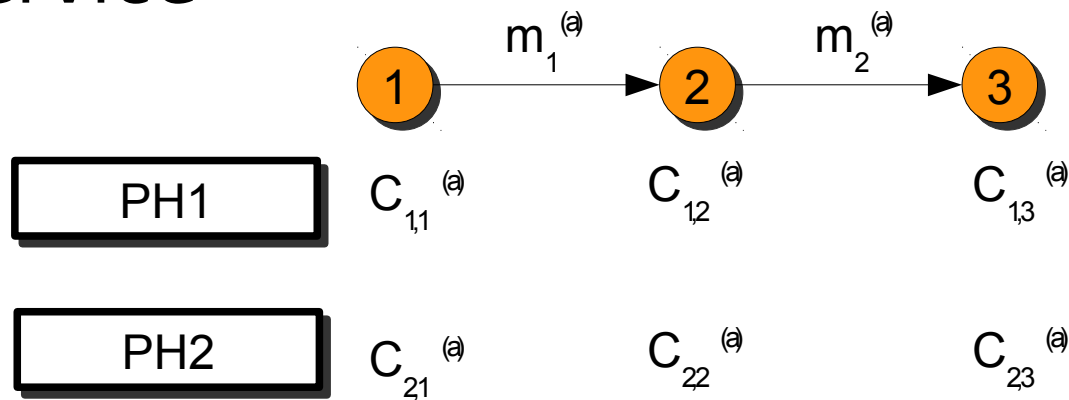
- Applications sharing the same PH may be independently activated
- Provider relies on actual probabilities of activation for admitted & new services



# Modelling

# Modelling real-time application workflows

- Application  $A^{(a)}$  is a pipeline of services
  - $C_{ij}^{(a)}$ : computation-time of  $i$ -th service when deployed on  $j$ -th PH
  - $m_i^{(a)}$ : size of data from  $i$ -th to  $(i+1)$ -th service



# Modelling computing response-time



## □ Reservation-based real-time scheduling

- A service is assigned  $(Q_i, d_i)$  parameters

- $Q$  time units (**budget**) reserved every time window of  $d$  time units (**period**)

- Service response-time due to computing:

$$\text{ceil}(C_{ij}/Q_i) * d_i$$

- If  $Q_i \geq C_{ij}$ , then response-time is  $d_i$

- Schedulability constraint:  
( $U_j$  is max CPU capacity)

$$\forall j \in \mathcal{H}, \sum_{a \in \mathcal{A}} \frac{\sum_{i \in \mathcal{A}(a)} x_{i,j}^{(a)} c_{i,j}^{(a)}}{d_i^{(a)}} \leq U_j$$

# Modelling network response-time

- Each transmission from  $i$ -th to  $(i+1)$ -th service is reserved a bandwidth of  $b_i$

- Data transmission time  $\frac{m_i^{(a)}}{b_i^{(a)}} + L_s$

( $L_s$  is a maximum fixed latency depending on the subnet)

- Schedulability constraint:  
( $B_s$  max link capacity)

$$\forall s \in \mathcal{S}, \quad \sum_{\substack{a \in \mathcal{A} \\ i \in \tilde{\mathcal{A}}(a)}} y_{i,s}^{(a)} b_i^{(a)} \leq B_s$$

# Modelling end-to-end response-time

## □ End-to-end response-time

$$\rho^{(a)} = \sum_{i \in \mathcal{A}^{(a)}} \left( d_i^{(a)} + \frac{m_i^{(a)}}{b_i^{(a)}} + \sum_{s \in \mathcal{S}} y_{i,s}^{(a)} L_s \right)$$

## □ Variables

- $d_i^{(a)}$  (real): relative computing deadline
- $b_i^{(a)}$  (real): network bandwidth
- $x_{ij}^{(a)}$  (boolean): i-th node on j-th host
- $y_{i,s}^{(a)}$  (boolean): i-th node on s-th subnet  
(derivate)

# Objective of optimization

- Cost due to turn-on of  $j$ -th host in each time-slot  $I_h$ :  $\zeta_j$
- Gain from accepting new service  $G^{(a)}$
- Minimize cost due to new hosts to turn on for admitting new services

$$\min_{x_{i,j}^{(a)}, y_{i,s}^{(a)}, d_i^{(a)}, b_i^{(a)}} \sum_{I_h \in \mathcal{G}} \sum_{j \in \mathcal{H}_{off}(I_h)} \zeta_j m_{j,h} - \sum_{a \in \mathcal{A}} x^{(a)} G^{(a)}$$

# Probabilistic response-time guarantee

## □ Deterministic setting

- Assumptions (Worst-Case figures):

- $C_{ij}^{(a)} \leq Q_i^{(a)} ; m_i^{(a)} / b_i^{(a)} + L_s^{(a)} \leq T^{(a)}$

- Goal:  $\rho^{(a)} \leq R^{(a)}$

## □ Probabilistic setting

- Assumptions (probabilistic figures):

- $\Pr\{C_{i,j}^{(a)} \leq Q_i^{(a)}\} \geq \alpha_i^{(a)}$

- $\Pr\{m_i^{(a)} \leq M_i^{(a)}\} \geq \beta_i^{(a)}$

- Goal  $\Pr\{\rho^{(a)} \leq R^{(a)}\} \geq \phi^{(a)}$

- Constraint:  $\prod_{i \in \mathcal{A}^{(a)}} \alpha_i^{(a)} \beta_i^{(a)} \geq \phi^{(a)}$

# Probabilistic availability guarantee

- Leverage of **actual average activation rates** of services  $r^{(a)} \ll 1/T^{(a)}$
- Probability that service is active in  $I_h$ :

$$\pi_i^{(a)} = r^{(a)} d_i^{(a)} \quad \pi_{i,j}^{(a)} \triangleq r^{(a)} d_i^{(a)} x_{i,j}$$

- For each time-slot  $I_h$ , prob. of enough bandwidth for all services in  $B(I_h)$ :

$$P_{j, \mathcal{B}(I_h)} = \prod_{a \in \mathcal{B}} \pi_{i,j}^{(a)} \prod_{a \in \mathcal{A}(I_h) \setminus \mathcal{B}} \overline{\pi_{i,j}^{(a)}}$$

# Probabilistic availability guarantee

- Probability  $\xi^{(a)}$  of having enough computing bandwidth for all services

$$\begin{aligned}
 & \sum_{I_h \in \mathcal{G}(a)} \frac{tn_h}{f^{(a)} - s^{(a)}} \prod_{j \in \mathcal{H}} \sum_{\mathcal{B} \subset \mathcal{A}(I_h) \setminus \{a\}} v_{\mathcal{B} \cup \{a\}}^j \cdot \\
 & \cdot \prod_{b \in \mathcal{B}} \prod_{i \in \mathcal{A}(b)} \pi_{i,j}^{(b)} \prod_{b \in \mathcal{A}(I_h) \setminus \{a\} \setminus \mathcal{B}} \prod_{i \in \mathcal{A}(b)} \overline{\pi_{i,j}^{(b)}} \cdot \\
 & \cdot \prod_{s \in \mathcal{S}} \sum_{\mathcal{B} \subset \mathcal{A}(I_h) \setminus \{a\}} w_{\mathcal{B} \cup \{a\}}^s \cdot \\
 & \cdot \prod_{b \in \mathcal{B}} \prod_{i \in \mathcal{A}(b)} \pi_{i, \{s\}}^{(b)} \prod_{b \in \mathcal{A}(I_h) \setminus \{a\} \setminus \mathcal{B}} \prod_{i \in \mathcal{A}(b)} \overline{\pi_{i, \{s\}}^{(b)}} \geq \xi^{(a)}, \forall a
 \end{aligned}$$

# Probabilistic optimization objective

- We introduce
  - Penalty  $P^{(a)}$  due to SLA violation
- Minimize expected gain minus costs:

$$\min_{x_{i,j}^{(a)}, y_{i,s}^{(a)}, d_i^{(a)}, b_i^{(a)}} \sum_{I_h \in \mathcal{G}} \sum_{j \in \mathcal{H}_{off}(I_h)} \zeta_j m_{j,h} - \sum_{a \in \mathcal{A}} x^{(a)} \left( G^{(a)} - \overline{\xi^{(a)}} P^{(a)} \right) \quad (20)$$

- Finally, we obtained a Mixed-Integer Geometric Programming (MIGP) optimization problem

# Conclusions and future work

- We tackled the problem of
  - allocation of distributed applications
  - with real-time timing constraints
  - over a physical network
  - under both deterministic and probabilistic guarantees in terms of
    - End-to-end response-time
    - Application availability at run-time
  - optimizing various system-wide metrics
- We modelled it as a MIGP problem

# Future work

- Validate the technique through simulation or real implementation
- Address scalability issues when deploying over large physical networks
  - via hierarchical approaches
  - via heuristics-based inexact solvers
- Refined optimization objectives
- Consider migration of already allocated virtualized services
- Extensions to non-linear workflows

# References



- T. Cucinotta, K. Konstanteli, T. Varvarigou, "*Advance Reservations for Distributed Real-Time Workflows with Probabilistic Service Guarantees*", to appear in IEEE International Conference on Service-Oriented Computing and Applications (SOCA 2009), December 2009, Taipei, Taiwan
- K. Kostanteli, D. Kyriazis, T. Varvarigou, T. Cucinotta, G. Anastasi, "*Real-time guarantees in flexible advance reservations*", 2<sup>nd</sup> IEEE International Workshop on Real-Time Service-Oriented Architecture and Applications (RTSOAA 2009), Seattle, Washington, July 2009
- F. Checconi, T. Cucinotta, D. Faggioli, G. Lipari, "*Hierarchical Multiprocessor CPU Reservations for the Linux Kernel*", in 5<sup>th</sup> International Workshop on Operating Systems Platforms for Embedded Real-Time Applications (OSPERT 2009), Dublin, Ireland, June 2009
- T. Cucinotta, G. Anastasi, L. Abeni, "*Real-Time Virtual Machines*", in 29<sup>th</sup> Real-Time System Symposium (RTSS 2008) -- Work in Progress Session, Barcelona, December 2008

**Thanks for your attention**

Questions ?