

# From the origins of performance evaluation to new Green ICT performance engineering

Carlos Juiz and Ramon Puigjaner

cjuiz@uib.es putxi@uib.cat

Universitat de les Illes Balears

Palma, Spain

**PERFORM 2010, October 14, 2010** 



#### From retired to retired or from emeritus to emeritus:

# **CONGRATULATIONS GÜNTER**

**PERFORM 2010, October 14, 2010** 





- > Introduction
- Erlang Works
- > Other Advances before the Computers
- > Performance Evaluation of Computers
- > Performance Evaluation of Networks
- > New Performance Challenges of ICT Systems
- Conclusions



#### Introduction



- > You cannot manage what you cannot measure
- Before installing a complex system it is good to have an estimation of what will be its behaviour.
- Since the first case of application based on the Erlang works on queuing theory to size the telephone system, a lot of techniques and tools have been developed for predicting the behaviour of ICT systems including the new challenges appeared during the last hundred of years.

# OUTLINE



- Introduction
- Erlang Works
- Other Advances before the Computers
- > Performance Evaluation of Computers
- > Performance Evaluation of Networks
- New Performance Challenges of ICT Systems
- Conclusions





- > We start our history with Agner Karup Erlang
- Erlang, Danish mathematician working for long time at the Copenhagen Telephone Company, had the opportunity to analyze the phenomena related to the new invention that was the telephone.



- > Among his works there are
  - the analysis allowing to represent the arrival of telephone communications like a Poisson distribution
  - the determination of how many circuits were needed to provide an acceptable telephone service (B and C Erlang formulae)
  - how many telephone operators were needed to handle a given volume of calls.



- The main limitation of the Erlang works is the assumption that both the service time and the interarrival time should be exponentially distributed.
- These results were intensively used for more than sixty years.



- Erlang was not a pure researcher but a hands-on one:
  - To verify his assumptions he did not refuse to conduct measurements that obliged him to climb into street manholes.
- He can be considered as the father of the queuing theory and the performance evaluation of ICT systems.

# OUTLINE



- Introduction
- Erlang Works
- Other Advances before the Computers
- > Performance Evaluation of Computers
- > Performance Evaluation of Networks
- > New Performance Challenges of ICT Systems
- Conclusions



Other Advances before the Computers



- Khinchine-Pollaczek Formula
  - Service times with any distribution
  - The result was not induced by a technical need but by the academic interest.
- Cobham Formula
  - Behaviour of a queuing system when there are customers with different priorities.

# OUTLINE



- > Introduction
- Erlang Works
- > Other Advances before the Computers
- > Performance Evaluation of Computers
- > Performance Evaluation of Networks
- New Performance Challenges of ICT Systems
- Conclusions



Performance Evaluation of Computers



- Predicting the performance of computers obliged to develop new techniques and tools to tackle the new challenges
  - Modelling
  - Benchmarking



- Queuing network modelling
  - Pioneering works of Jackson, and Gordon and Newell with the algorithmic complement of Buzen
  - Theorem of Baskett, Chandy, Muntz and Palacios (BCMP) on product form networks.
    - Complete basis for modelling the performance of a computer system allowing an analytical process
    - Restricted by an important set of conditions.
    - Convenient algorithms like those of Reiser and Kobayashi, and Reiser and Lavenberg.

UIB

- Operational analysis
  - Developed in parallel by Buzen and Denning
  - Experimental and intuitive approach without the strong mathematical and statistical conditions of the queuing network approach
  - Similar numerical results to those of BCMP



- To facilitate the use of these results, several tools were developed. Among them:
  - BEST-1 developed by BGS, the company created by J.P. Buzen,
  - RESQ developed by IBM
  - QNAP2 developed by INRIA.



- To relax the constraints of the product form queuing networks, different approximated methods were developed:
  - Decomposition-aggregation by Courtois,
  - Iterative by Marie,
  - Diffusion by Gelenbe and Mitrani
  - Etc.
  - Some of these algorithms were included in some of the previous tools.



Some of the mentioned tools included simulation capabilities to analyze queuing network models not able to be processed by any other technique. Performance Benchmarking



- Comparison of the computing capacity of different systems:
  - Initially evaluated by the execution time of a typical instruction (addition)
  - With the increasing complexity (clever architectures, sophisticated operating systems, etc.) this simple comparison was no more valid.
  - Computer customers proposed the execution of a significant sample of their real workload on the different systems to compare.
- Sophisticated techniques were developed to characterise the system workload.
  PERFORM 2010, October 14, 2010

Performance Benchmarking



- Next step: standardisation by consensus of programme sets depending on the type of workload.
  - LINPACK, for the scientific computation
  - Established by organizations participated by computer manufacturers and customers, like SPEC.

Performance Benchmarking



- The benchmarks proposed by SPEC were devoted to:
  - Initially to batch environments
  - Then to conversational systems
  - Currently cover most of the typical working environments of computer systems.
- Possibility of scaling the SPEC benchmarks to adapt the size to the particular needs of the customer.
- SPEC publishes results of its benchmarks with different sizes running on a variety of systems.
  PERFORM 2010, October 14, 2010

# OUTLINE



- Introduction
- Erlang Works
- > Other Advances before the Computers
- Performance Evaluation of Computers
- Performance Evaluation of Networks
- > New Performance Challenges of ICT Systems
- Conclusions



Performance Evaluation of Networks



- When computers were no more able to work isolated the need of connection initially with terminals and then with other computers appears.
- The complexity of the systems increased and consequently the complexity of the tools and techniques used to evaluate their performance:
  - Models
  - Benchmarks

PERFORM 2010, October 14, 2010

Performance Modelling of Networks



- With computers running in an isolated way it was quite easy to build analytical models able to be processed with the help of appropriated tools.
- However, with the increasing complexity of the networked systems (wired or wireless LANs, or ad-hoc networks, or sensor networks, etc.), it is much harder to find an enough accurate model for representing such systems and to solve them analytically.

Performance Modelling of Networks



- The increasing computer speed allowed building accurate simulation models and computing results with small enough confidence intervals in short enough computing times.
- Among the most popular simulation engines there are:
  - OPNET, proprietary software (there is a reduced free version for universities and research centres) that offers "black boxes" of models of frequently used networks and protocols.
  - ns-2, free software network simulator, with its typical advantages and inconveniences. Similar structure to OPNET.

Performance Benchmarking of Networks



- In terms of standard benchmarks, SPEC is proposing a variety of benchmarks to cover the most typical systems including a network in their structure.
- > Currently there are benchmarks for analyzing:
  - conversational systems
  - transactional systems
  - web systems (e-commerce sales systems)
  - etc.

# OUTLINE



- Introduction
- Erlang Works
- > Other Advances before the Computers
- Performance Evaluation of Computers
- Performance Evaluation of Networks
- New Performance Challenges of ICT Systems
- Conclusions





- Performance challenges are evolving with the changes of ICT systems
- > Current challenges are mainly in:
  - The domain of new network architectures.
  - The energy consumption in both systems and networks.
  - Etc.
- One of the problems is concerning the energy consumption in order to reduce the climate change
- > Why is this problem appearing?



- The growing of web systems degrades quality of service (e.g. response time) and the performance of any web application.
- The new challenge directly related to this branch of research emerges:
  - the relationship between performance and energy cost



- For mobile devices, battery capacity and energy use directly affect usability.
- Battery capacity determines how long devices last.
- As battery capacity is limited and improving slowly, device architects have concentrated on extracting greater energy efficiency from the components (processor, display, wireless subsystems in isolation, etc.).



- Some efforts are under way to establish benchmarks for energy efficiency in data centres.
- Other works have emphasized metrics such as the energy-delay product or performance per watt to capture energy efficiency for processors and servers without fixing a workload.
- There has been less focus on the I/O subsystem, which plays a significant role in total system power.



- Actions have to be taken to improve the energyefficiency of the distributed web servers:
  - At the networking level, the idea would be to decrease the overall usage of the network by:
    - Moving large amount of data on routes consuming less energy
    - Optimizing the placement replication of the data to minimize the equipment usage;
    - Minimizing the time to transfer the data,
    - Reducing the data to be transmitted.



- Actions have to be taken to improve the energyefficiency of the distributed web servers (cont.):
  - At the server level, to maximize energy efficiency by understanding the relationship between resource usage and system-level power consumption.
    - By consolidating workloads.
    - By building power-performance models for servers executing web different workloads.
    - By server virtualization.



- Actions have to be taken to improve the energyefficiency of the distributed web servers (cont.):
  - At the application level, context awareness of the applications should be favored. Developing energy-aware data mining techniques will support energy efficiency at the applications level.



- ➢ We propose an incomplete list of Green ICT performance topics putting together computing technologies, networking, optimization, and their performance evaluation and power issues.
  - Energy consumption/performance models.
  - New simulation, benchmarking and monitoring engines considering devices, networks and their corresponding energy consumption
  - Energy Efficient Computing Centres: this will be addressed at different levels, i.e. at a data centre level, workload and server management, respectively.



- Incomplete list of Green ICT performance topics (cont.):
  - Energy Efficient Networks at multiple levels (networking devices, communication and control protocols, network architectures design, and network performance).
  - Green Software Design: at compiler level on algorithmic and data structures optimizations for grid systems. Also, at application level to investigate the key aspects in the applications specifying the performance requirements/power consumption that user expects from those services



- Incomplete list of Green ICT performance topics (cont.):
  - Design for Energy Awareness: empowering social view in the use of products and services by increasing visibility of energy choices during operation.
  - ICT Management Green Metrics: indicators of performance/energy from data centre monitoring to ICT governance/management
  - Power-aware middleware for data centres: intelligent power management for virtualised machines, power-aware virtualisation and benchmarking.



- Incomplete list of Green ICT performance topics (cont.):
  - Power-efficient routing: design of packet classifiers for high speed and low power consumption.
  - Energy-efficient bandwidth allocation: making minimum consumption of energy resources in network equipment.
  - Energy and distributed computing: placement of mobile agents, workload placement, virtual machine migration, P2P resource allocation, etc, providing means to develop energy-aware allocation strategies.



- Incomplete list of Green ICT performance topics (cont.):
  - Eco Data Base Management Systems: managing the energy consumed during query processing and considering query aggregation in workload.
  - Energy-Efficient algorithms: Dynamic Speed scaling and power-down client integration but holding resilience of systems.
  - Dynamic performance management: programming models for the cloud rely on the use of tasks scheduled across large data centres, taking advantage of the parallelism/performance and power.
  - Etc.

# OUTLINE



- Introduction
- Erlang Works
- > Other Advances before the Computers
- Performance Evaluation of Computers
- Performance Evaluation of Networks
- New Performance Challenges of ICT Systems
- Conclusions



#### Conclusions



Computer performance evaluation concept has evolved along with the changes experienced by the changes in both the technology hardware and software and in the goals of the end users with respect to the systems they were using.

#### Conclusions



We have seen how the tools and techniques used today to evaluate or predict the computer performance have evolved and the new challenges derived from Green ICT will oblige to find new tools and techniques addressed to overcome the current challenges.



# THANK YOU FOR YOUR KIND ATTENTION QUESTIONS?

PERFORM 2010, October 14, 2010



# From the origins of performance evaluation to new Green ICT performance engineering

Carlos Juiz and Ramon Puigjaner

cjuiz@uib.es putxi@uib.cat

Universitat de les Illes Balears

Palma, Spain

**PERFORM 2010, October 14, 2010**