What it’s about…

Developing new technologies and optimization methodologies to develop next-generation energy-efficient datacenters.

Context and project goals

Energy efficiency in datacenters is of strategic importance to Switzerland, as over 75% of the Swiss economy is service-based and depends on information technology (IT), which makes Switzerland one of the top spenders of IT per capita. IT is also witnessing a major paradigm shift towards Cloud Computing with datacenters emerging as a key backbone of services at scale, making energy dissipated in datacenters a key concern even with moderate increases in IT electricity demands. In particular, because of Switzerland’s decision to abandon nuclear energy by 2034, improvements in energy efficiency are imperative to make up for half of the electricity otherwise furnished by nuclear power. Meanwhile, in Switzerland many IT departments for enterprises, research, and governmental organizations alike are at capacity with regards to their electricity budget. Thus, the only feasible solution to achieve energy sustainability is to drastically increase the power efficiency of data centers.

The design of datacenters today is a really complex process, where important opportunities exist both within and across various server and infrastructure components. Server software, system and silicon technologies as well as infrastructure for cooling and power delivery have historically been designed in isolation with an over provisioning of resources to guarantee a desired quality of service. Unfortunately, due to the diverse nature of workloads and demands on resources, such over provisioning results in prohibitive levels of waste in energy and efficiency. Modern volume server software and hardware is broadly based on designs primarily derived from the desktop market and are ill-suited for serviced-oriented server workloads. Similarly, while most datacenters make use of air-cooling technologies to ensure the correct running of the servers, air-cooling is reaching fundamental physical limits in efficiency with a continued increase in server density requiring innovation in cooling technologies.

In YINS, we propose to develop a radically new thermal-aware design approach for next generation energy-efficient datacenters. This new design approach tightly integrates the cooling infrastructure definition with holistic system-level power, performance and thermal management. This vertically-integrated system-level management paradigm goes beyond hardware and software boundaries by redesigning the entire datacenter to maximize performance given a target power, area and cost budget. Therefore, we propose to develop new server technologies based on Fully Depleted Silicon On Insulator (FDSOI) and specialized server architectures. Finally, these novel architectures interact with new on-chip microfluidic cooling delivery at server-level and passive thermosyphon cooling systems for the rack and room-level, as well as energy recovery strategies for the complete datacenter.

To realize its vision, YINS requires inter-disciplinary research at the boundaries of multiple scientific domains, as well as developing and integrating innovations in critical research areas, namely, computer systems, circuits and semiconductor technologies, and cooling technologies, large-scale simulation, software synthesis and optimization, statistical network modeling and model predictive control theory. Therefore, the YINS consortium includes six world-renown academic partners from EPFL and ETIHZ covering key research areas of computer, electrical and mechanical engineering, and three key industrial partners in Switzerland for datacenter design and large-scale IT banking services provisioning (Credit Suisse, Eaton and BrainServe), and three third-party industrial partners (Constellium, Froiotherm and Osram/de).
How it differentiates from similar projects in the field

The vertically-integrated system-level management paradigm that YINS targets for datacenters goes beyond competitors in the field by performing inter-disciplinary research at the boundaries of multiple scientific domains.

It integrates innovations in several research areas, namely, computer engineering and cooling design, large-scale computing system simulation, software generation and optimization, statistical network modeling and model predictive control theory.

Therefore, YINS includes six world-renowned academic partners from EPFL and ETHZ covering these key research areas and three key industrial partners in Switzerland for datacenter design and large-scale IT banking services provisioning (Credit Suisse, Eaton and BrainServe).

Quick summary of the project status

- The team has started developing an architectural simulating infrastructure for the modelling of heterogeneous large scale computing systems running computational kernels.
- The project validated the application of on-chip microfluidic fuel cell networks for joint cooling and power supply (i.e., localized power generation and delivery) of memories in multi-core computing servers.
- Joker, a framework for answering different "what-if" workload deployment configuration questions in data centers has been developed.
- Bulk Memory Page Access Prediction and Streaming (or BuMP) have been introduced to improve energy efficiency for in-memory workloads.
- A thermosyphon loop test bench has recently been developed and installed in the LTCM laboratory facilities.
- Analytical models and metrics relevant to the design and optimization of datacenters have been identified.

Success stories

A new wireless energy monitoring device and management system for datacenters has been developed to help companies. This new system, called Power System Monitoring and Management (PMSM), has been developed by the ESL-EPFL partner of YINS, and the new system is able to monitor and track the energy consumption by racks of servers and, even better, can help redistributing workload among servers to optimize energy use. PMSM has been already installed in two Credit Suisse datacenters and has been reported to enable cut power demand by 30 to 50 percent.

Within YINS, LTCM-EPFL members presented the first two-phase cooling prototype at server level. This prototype demonstrated that two-phase cooling can improve heat removal efficiency while requiring a lower flow rate (for lower operation cost) and enabling better temperature uniformity across the server chips. Their technology is showcased on the cover of the Electronics Cooling, a high-profile magazine dedicated to thermal management in electronics industry. In addition, Nicolas Lamaison, Jackson Marcinichen and John Thome have been awarded the Best Paper Award at InterPACK. The authors showcase in this work for the first time the need for transient modeling and control of on-chip liquid cooling in servers and propose, the use of online models for effective and accurate cooling control, which we will explore further in YINS in the coming reporting periods.

Presence in the media:

The “Power System Monitoring and Management (PMSM) device and system, developed at ESL-EPFL in cooperation with Credit Suisse, has been cited at Clean Technica, SNS Analytics International.

Main publications


Andrea Bartolini, Matteo Cacciari, Carlo Cavazzoni, Giampietro Tecchioli, Luca Benini, Unveiling eurora - thermal and power characterization of the most energy-efficient supercomputer in the world, DATE 2014.

Francesco Beneventi, Andrea Bartolini, Pascal Vivet, Denis Doutot, Luca Benini, Thermal analysis and model identification techniques for a logic+ WIDEIO stacked DRAM test chip, DATE 2014.


“Green servers and datacenters targeting the energy efficiency of portable cell phones”