Integrated systems are critical to IT revolution

Embedded computing (cell phones, automotive, gaming)

High performance computing (scientific applications, forecasting, data centers)

Cyber-physical computing (healthcare, transportation, smart grid)

Three paradigm shifts, i.e. low-power (~90s), network-centric (~2K), and cyber-physical design (~2010). Exciting times ahead...
Low-power Design. Computation vs. communication

Multicore platforms are large scale distributed systems at nanoscale; they are dominated by communication costs

Last level cache (LLC)
Memory controller (MC) & channels
I/O controller(s)
QPI controller,
Power control unit (PCU), etc.

Need to understand the behavior of thousand core systems.

Network (routers+links) is the missing link in understanding.
CPS consist of networks of computational devices and networks that monitor and control physical processes

Multi-scale behavior has been observed in many technological, economical, and biological systems

MPSoCs as CPS: Current multicore platforms bring together computation, communication, and control

Processing cores interconnected via routers and links; they operate at various voltages/frequencies thus saving power
This presentation focuses on the new CPS paradigm and its impact on future integrated systems

Structure
Architecture and small world effects

Dynamics
Workloads and multiscale behavior

Control
Power and resource management

Our first insight into communication-based design came through architecture (topology, buffer, etc.) optimization

[Bolotin et al. DATE 2007]


[J. Kim et al. ISCA 2007]
Can induce small-world effects in regular NoCs. This brings huge performance improvements.

\[ b(n) = D(m,n) = m + n - 2 \]

\[ b(n) = D'(m,n) < m + n - 2 \]

\[ b(n) = \log_2 N \]

This way, the fundamental idea of Small World networks (aka “six degrees of separation”) enters the multicore world.

Small world effects can also be exploited to reduce hop count in 3D wireless NoCs and improve performance.

Wired and wireless NoCs can be used intra-chip, while inter-chip communication is based on wireless inductive-coupling.
Flow-control mechanisms, reconfigurability, adaptive routing have all been used to improve performance

[Q. Zhiliang et al. CODES-ISSS 2012]

One way or another, they all exploit the small world effects...

Optimum path-seeking behavior may be goal-oriented. Need to collect fitness information locally, at routing nodes

$\text{fitness}_i = \alpha \cdot m_{in} \cdot w_{in}^{-1}$

$m = \text{avg. number of free slots}$

$w = \text{avg. waiting time in channel}$

[T. Mak et al. IEEE Cir&Sys Mag.2011]
This presentation focuses on the new CPS paradigm and its impact on future integrated systems.

Packet inter-arrival times at interface queues play a fundamental part in network behavior.
Network behavior at high injection rates needs a non-equilibrium approach that accounts for fractal behavior.

High injection rates cause inter-arrival times deviate from exponential distribution and exhibit power law correlations.

Fractals are geometrical objects or stochastic processes displaying self-similar behavior over multiple scales.

Space

Time

Bellcore traffic
How long is the coast of Britain? Answering this question involves statistical self-similarity and fractal dimension.

Fractal dimension can be computed via box counting.

Choose an increasing set $R$ of edge lengths. For each size $r_i$ in $R$

- Super-impose a series of distinct squares (boxes) over the data.
- Count the minimum number of boxes needed to cover data and store it in vector $N$.

Compute fractal dimension by fitting the equation between $R$ and $N$.

$$ N(R) \approx R^{-D} $$

Box counting method can be applied to determine the fractal dimension of 1D, 2D, or 3D vectors.

Real world processes are not always smooth. Fractional dynamics needs a formalism stronger than integer calculus.

Fractions of behavior

Mandelbrot (1975)

$$ \langle \Delta x \rangle \propto \Delta t^H, \; 0 < H < 1 $$

$$ \frac{d^\alpha x(t)}{dt^\alpha} = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \sum_{j=0}^{[\Delta t]} (-1)^j \binom{\alpha}{j} x(t - j\Delta t) $$

For CPS, time is an intrinsic component of the programming model. Consequently, system dynamics becomes essential.
This presentation focuses on the new CPS paradigm and its impact on future integrated systems.

**Structure**
Architecture and small world effects

**Dynamics**
Workloads and multiscale behavior

**Control**
Power and resource management

---

8-Core Xeon® Processor has three clock domains and three voltage domains that help minimizing power.

8-Core Xeon® Processor advanced features:
- Three clock domains: Core Domain (MCLK), Un-Core Domain (UCLK), I/O Domain (QCLK)
- Different voltage domains:
  - Core Domain: 0.85-1.1V variable
  - Un-Core Domain: 0.9-1.1V fixed
  - I/O Domain: 1.1V fixed

Device count reported 2.3B transistors. The chip has 9 temperature sensors, one in each core hot spot and one in the die center.

[Rusu, ISSCC 2009]
Fine-grain power management becomes possible by exploiting workload variations

Minimizing energy consumption

Total energy consumption to be minimized

\[ E_{\text{Total}} = E_{\text{App}} + \sum_{i=1}^{\#VFIs} E_{\text{VFI}}(i) \]

- Application (useful) energy consumption (comp+comm)
- Overhead of \(i\)th VFI

\[ E_{\text{VFI}} = E_{\text{ClkGen}} + E_{\text{Vconv}} + E_{\text{MixClkFifo}} \]

Integrated power controller – independent of CPUs
- Single inductor multiple output (SIMO) converters
- Adaptive-output converters
- Voltage regulator controller (VRC) on Intel SCC

[D. Ma, Cir&Syst Mag, 2010]
Fine-grain power management can be implemented via control-theoretic approaches.

Voltage-Frequency Controller

Utilization (reference) values for interface FIFOs

\[ X_{\text{ref}} = [x_1^{\text{ref}}, x_2^{\text{ref}}, \ldots, x_N^{\text{ref}}]^T \]

Actual utilization of interface FIFOs

\[ X = [x_1, x_2, \ldots, x_N]^T \]

Inter-arrival times dictate the fractal exponent \((\alpha_k)\) of state equations; this is used to characterize the system dynamics.

V/F controller selects the minimum operating frequencies s.t. the queues reference values are satisfied

\[ \frac{d^\alpha x_k}{dt^\alpha} = G(x_k, f_j, f_i, t) \]

Applications (workload)

\[ V_1, f_1 \in [f_1^{\text{min}}, f_1^{\text{max}}] \]

\[ V_N, f_N \in [f_N^{\text{min}}, f_N^{\text{max}}] \]

Physical constraints

\[ 0 \leq y_i^{\text{min}} \leq y_i(t) \leq y_i^{\text{max}} \leq 1, \quad i = 1 + N_{PE} \]

\[ 0 \leq x_k^{\text{min}} \leq x_k(t) \leq x_k^{\text{max}} \leq 1, \quad k = 1 + 4, \quad j, l = 1 + N_r \]

\[ \min \left\{ \sum_{i=1}^{N_{PE}} \left[ \frac{1}{2} w_i (y_i(t) - y_i^{\text{ref}})^2 + \frac{1}{2} z_i f_i(t)^2 \right] \right\} dt \]

\[ \min \left\{ \sum_{i=1}^{N_{PE}} \left[ \frac{1}{2} r_i f_i(t)^2 + \sum_{k=1}^{4} q_k (x_k(t) - x_k^{\text{ref}})^2 \right] \right\} dt \]

Author: P. Bogdan et al. NOCS 2012
Accurate mathematical modeling and rigorous optimization can enable cross-layer power management

Queues utilization at tiles (0,0), (1,1) and (1,2) for a 4×4 mesh NoC running Apache HTTP webserver application.

FOC keeps the utilization of all queues below 0.1 by adjusting the operating frequencies of all PEs and routers. FOC consumes less power compared to a LQR controller.

How about core-to-core variability? Use frequency asymmetry to reduce performance loss.

Run leakier cores at lower voltages, but at higher than normal frequencies for those voltages.

[Herbert et al, IEEE TVLSI 2012]
For thousand core systems distributed approaches for power management are of crucial importance

15-router synch NoC that connects 22 processing units
[F. Clermidy et al., ISSCC’10]

Distributed objective function
Local maximization algorithm

Not only a power issue...

Distributed and hierarchical controllers:
- Energy Controller (EC)
  - Output Frequency $f_{EC}$
    - Minimize power – CPI based
    - Performance degradation < 5%
- Temperature Controller (TC)
  - Distributed MPC
  - Inputs:
    - $f_{EC}$, $T_{CORE}$, $T_{NEIGHBOURS}$
  - Output
    - Core frequency ($f_{TC}$)

[Slide courtesy of L. Benini, Bologna U.]
Peak temperature is important. DTM cannot be achieved by considering power alone. Physical context is crucial...

Agent-based systems can help system components (re)configure and optimize their resource usage independently.

Both centralized and distributed approaches have their own limitations

Scalability issues due to cost and latency of long wires. Long synchronization times

Highly scalable but potential problems in control performance

Need ‘best-of-both-worlds’ between fully-centralized and fully-distributed solutions. This is true for thermal management too.
An hierarchy of globally distributed locally centralized control may help the system self-organize

Local control w/ full state information, global control w/ partial information. Small world effects help convergence

<table>
<thead>
<tr>
<th>System Size</th>
<th>Flat Mesh (nJ)</th>
<th>WiNoC (nJ)</th>
<th>Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>128</td>
<td>1319</td>
<td>22.57</td>
<td>58x</td>
</tr>
<tr>
<td>256</td>
<td>2936</td>
<td>24.02</td>
<td>122x</td>
</tr>
<tr>
<td>512</td>
<td>4992</td>
<td>37.48</td>
<td>133x</td>
</tr>
</tbody>
</table>


This presentation focuses on the new CPS paradigm and its impact on future integrated systems

Structure
Architecture and small world effects

Dynamics
Workloads and multiscale behavior

Control
Power and resource management
A pacemaker analyzes the function of the heart; if necessary, sends signals to correct certain abnormalities.

A pacemaker (PM) is an electronic device implanted in the body to regulate the heart beat. A PM is not designed to defibrillate the heart by the delivery of shocks. It consists of a battery and electronic circuits enclosed in a hermetically sealed can. The PM delivers electrical stimuli over leads with electrodes in contact with the heart.

R-R interval: Time between an R wave and the next R wave.

Normal resting heart rate is between 60-100 bpm (0.6-1.0sec)
Heart rate is non-stationary. This means that its distribution and its higher order moments vary with shifts in time.

Statistics of R-R intervals vary with time as a function of various activities (sleep, running, etc.)

Heart rate datasets available at: http://www.physionet.org/

DFA can reveal the fractal dimension of R-R intervals

\[ y(k) = \sum_{i=1}^{n} [B(i) - B_{ave}] \]

\[ F(n) = \left( \frac{1}{N} \sum_{k=1}^{N} [y(k) - y_n(k)] \right)^2 \]

Interpretation
- slope \( \sim 0.5 \) white noise
- slope \( \sim 1 \) LRD
- slope \( >> 1 \) Brownian noise

\[ \log F(n) \]
\[ \log n \]
Breakdown of fractal physiological control mechanism can lead to highly periodic output (single scale) or randomness.

DFA can be applied to heart rate variability. Degradation of fractal correlations can be used as a quick diagnostic tool.

\[ F(n) \approx n^\alpha \]
Controller finds the pacing frequency which minimizes the deviation of ISE of heart rate from the reference value

\[ \min \int_{t_i}^{t_f} \left\{ \frac{1}{2} w (y(t) - y_{ref})^2 + \frac{1}{2} z f(t)^2 \right\} dt \]

\[ y(t_i) = y_0 \quad \text{and} \quad y(t_f) = y_1 \]

\[ f_{min} \leq f(t) \leq f_{max} \]

Fractal Optimal Controller

Model of Heart Rate Variability

Sensing, State Estimation, Model Identification

\[ \frac{d^\alpha y(t)}{dt^\alpha} = a(t)y(t) + b(t)f(t), \quad y_{min} \leq y(t) \leq y_{max} \]

Atrial fibrillation is characterized by short R-R intervals. If left untreated, it can lead to congestive heart failure

Healthy R-R intervals (0.66 to 1 sec)

Dangerous R-R intervals (as low as 0.20 sec)

\[ X = 0.096 \quad Y = 0.7166 \] (corresponds to a heart rate of 83 beats per min)

FOC brings the R-R interval from 0.40 to 0.80 secs (i.e., a healthy heart rate of about 75 beats per minute)
Workload analysis should not be an afterthought. In real CPS, network traffic is neither Poisson, nor stationary.

**Classical dynamics:** Linear Dependence & Exponential Inter-Event Distribution

\[
\frac{dP(a,t)}{dt} \propto P(a,t) \\
\frac{dM_1(t)}{dt} \propto M_1(t)
\]

**Fractal dynamics:** Linear Dependence & Power-Law Inter-Event Distribution

\[
\frac{d^\alpha P(a,t)}{dt^\alpha} \propto P(a,t) \\
\frac{d^\alpha M_1(t)}{dt^\alpha} \propto M_1(t)
\]

Statistical properties of the workload have deep implications in resource allocation, architectural design, RT scheduling, etc.

In summary, thousand core systems offer ample opportunities to bring science and engineering even closer.

**Hardware Capacity**

- microcontroller/microprocessor
- IP core
- System-on-Chip
- Networks-on-Chip
- memory

**Software Programming**

- single task uni-processor
- multiple tasks uni-processor (multi-threading)
- multiple tasks multi-processors (multi-processing)

User/application/OS optimization and resource management is the next frontier to conquer.
Finally...

Contributors
Paul Bogdan (Univ. of Southern Calif.), Umit Y. Ogras (Intel), Siddharth Garg (Univ. of Waterloo), Diana Marculescu (Carnegie Mellon Univ), Chen-Ling Chou (Qualcomm), Qian Zhiliang (Hong Kong Univ Sci&Tech), Chi-Ying Tsui (Hong Kong Univ Sci&Tech), Hiroki Matsutani (Keio Univ).

Relevant papers - www.ece.cmu.edu/~sld

Sponsors

Questions?

Thank you!