



Asynchronous Design for IoT, or 25+ happy years in Newcastle and 100 Technical Reports on async.org.uk

Alex Yakovlev

uSystems Research Group

Newcastle University ALIOT Summer School July 2018

Outline

• There is no Outline today!!!

People involved

- Patrick Degenaar (CANDO) Optoneuro, Building implantable uSystems
- Andrey Mokhov (A4A, POETS) asynchronous, concurrency, graph algebras, high performance computing
- Rishad Shafik (PRiME) energy efficient many core, approximate computing, Machine learning
- Senior RAs and Cols: Fei Xia (PRiME), Delong Shang (SAVVIE), Danil Sokolov (A4A)

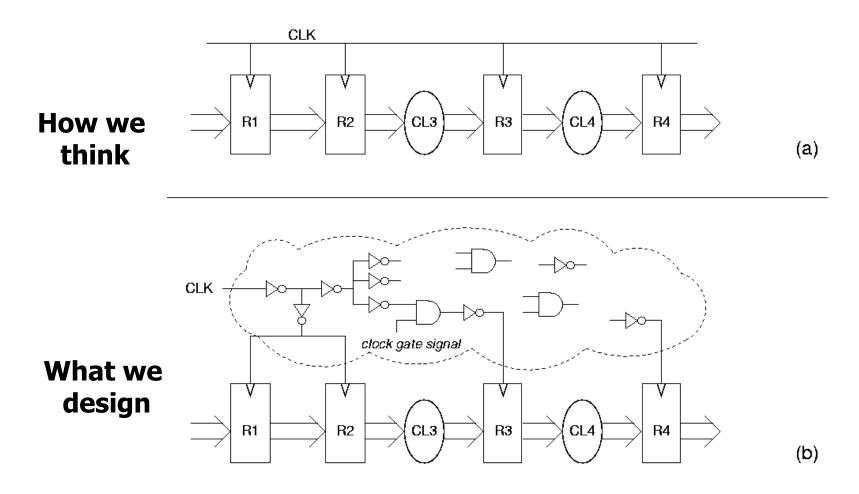
OUR RESEARCH VISION:

- "REAL POWER" (ENERGY-MODULATED) COMPUTING
- LOW ENERGY: MIN ENERGY POINT
- RESILIENCE TO VARIABILITY
- POWER-COMPUTE CO-DESIGN
- "APPROXIMATE" COMPUTING
- MIXED SIGNAL DESIGN AUTOMATION

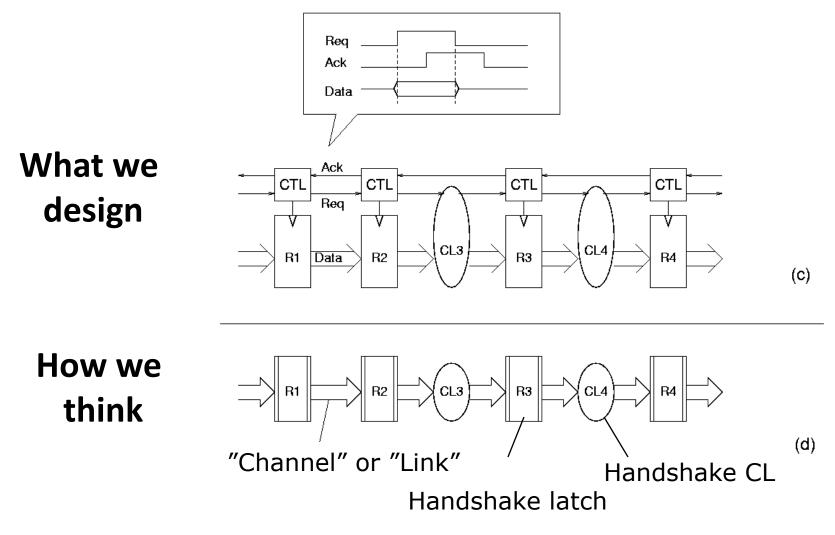
How it all began ...

- Prior to that:
 - David Kinniment's research on async processor designs and metastability and synchronsers in the 60s & 70s (Manchester)
 - My research on async design, models, interfaces in the 80s (St. Petersburg)
 - 1984-85 joint work (with Harry Whitfield, Albert Koelmans and Martin McLaughlin) on CAD for async (Newcastle) plus interactions with Brian Randell and Maciej Koutny on concurrency and dependability

Synchronous clocking



Asynchronous handshaking



Brief History of Asynchronous clocking

- In the 1950s there were no clocks in computers they were 'asynchronous'
- Theory of asynchronous switching circuits was created – David Muller et al from University of Illinois
- In the 1960s-70s circuits become more complex and designing them was easier with the clock
- In the 1980s VLSI circuits appeared and there began a battle between synchronous (circuits were complex to design and test; CAD tools started to appear based on clocked FSM paradigm) and asynchronous (interacting with the real world, metastability problems, working on average rather than worst case delays)

Brief History of Asynchronous clocking

- In the 1990s drive for low power began and async gave some advantages; later the problem of design reuse came about and GALS approach was propose
- In the 2000s drive for robustness against PVT variations, security, noise and EMI reductions gave more drive to async
- In the 2010s multi Billion transistor chips design many clock/Vdd domains, NOCS and GALS is a reality
- Push for ULTRA low power causes to work in nearthreshold and subthreshold
- Analog-Mixed Signal comes on chip needs little digital control – again asynchronous is important

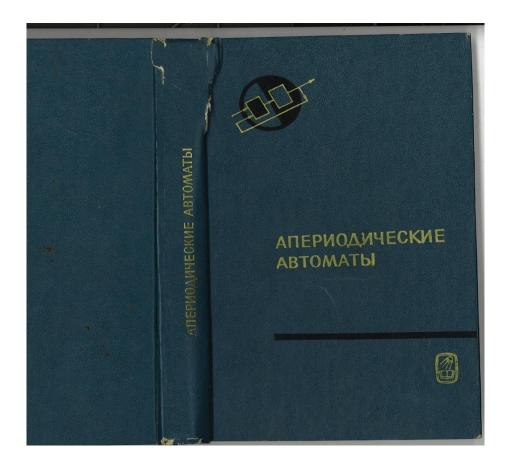
LETI people in async world: back then

- Victor Varshavsky and his group
 - Early work of Varshavsky on digital systems and automata
 - 1970s: work in the USSR Academy of Sciences Institute of Economics and Mathematics (LOCEMI) and Institute of Social-Economic Problems (ISEP)
 - 1980: the group moves to LETI (MO EVM)
 - 1980s: the group develops methods for synthesis, analysis, new tools, first VLSI chips (FIFO in PLMs), multiprocessor systems (fault tolerant token ring interface for airborne onboard computers) -> first GALS system – 1986
 - Early 1990s: Trassa co-operative: libraries of modules, Trassa synthesis-analysis tools under MS DOS

Victor Varshavsky: pioneer of asynchronous design



The first book on asynchronous design



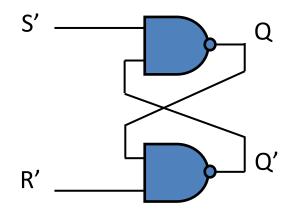
Aperiodic Automata, edited by V. Varshavsky, Nauka, Moscow, 1976

It contained:

Basic blocks with completion (flip-flops, registers, combinational)

Implementation of control (direct translation)

Asynchronous latches

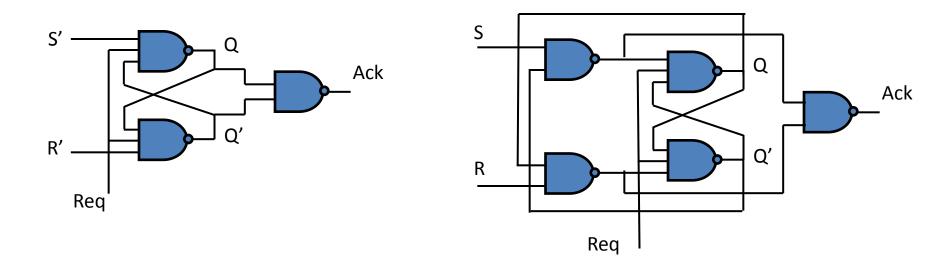


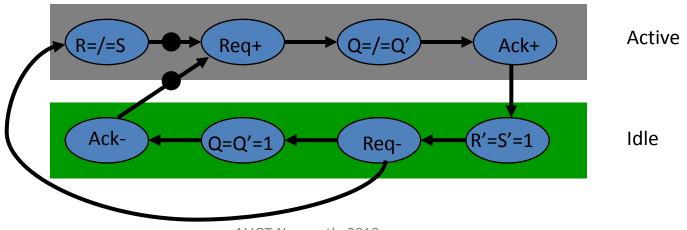
Classic RS-latch: observing outputs cannot tell if operation is completed

Completion detection

Partition computation into Idle and Active phases

Asynchronous latches

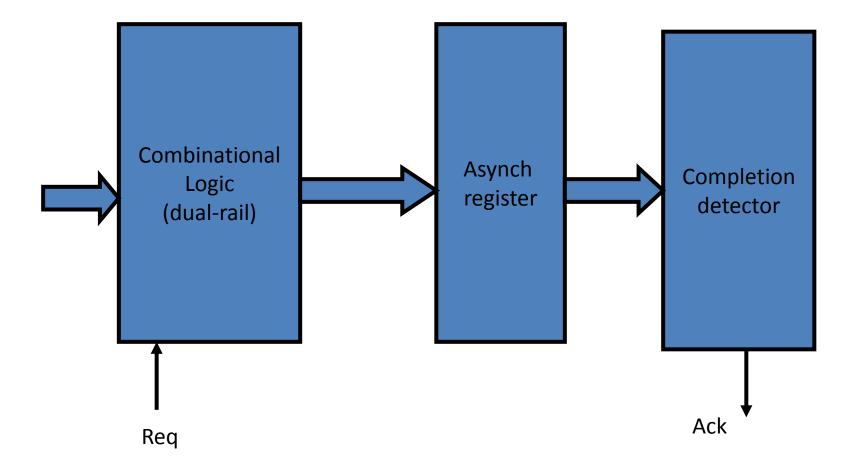




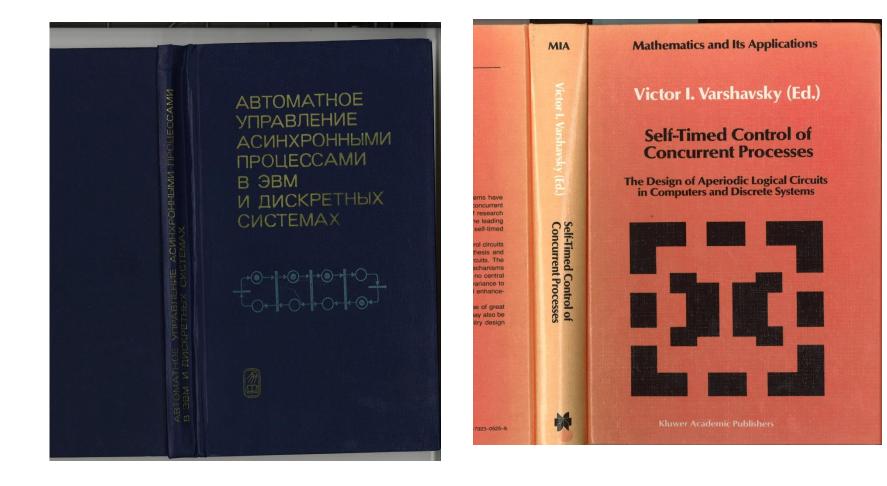
12/07/2018

ALIOT Newcastle 2018

Two-phase asynchronous FSMs

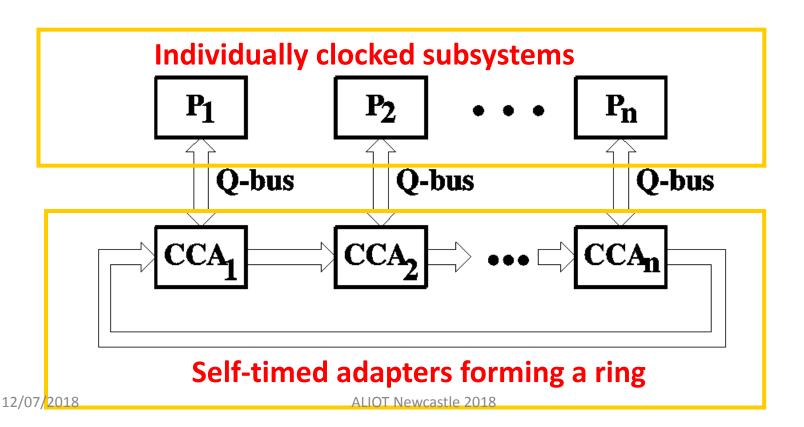


Second book (1986, 1990)

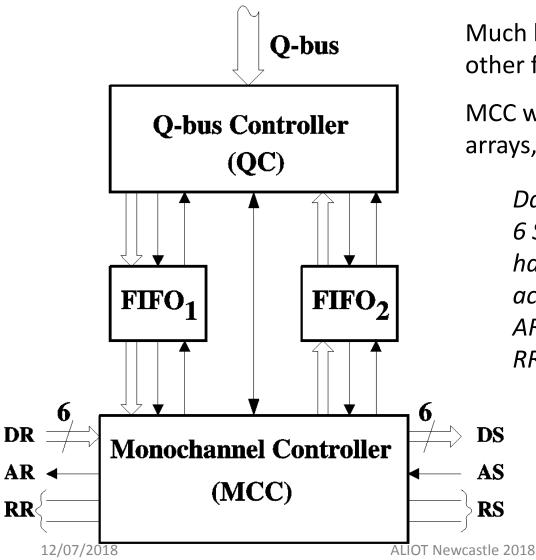


Fault-Tolerant Self-Timed Ring(1981-1986)

For an on-board airborne computer-control system which tolerates up to two faults. Self-timed ring was a **GALS system** with **self-checking and self-repair at the hardware level**



Communication Channel Adapter



Much higher reliability than a bus and other forms of redundancy

MCC was developed TTL-Schottky gate arrays, approx 2K gates.

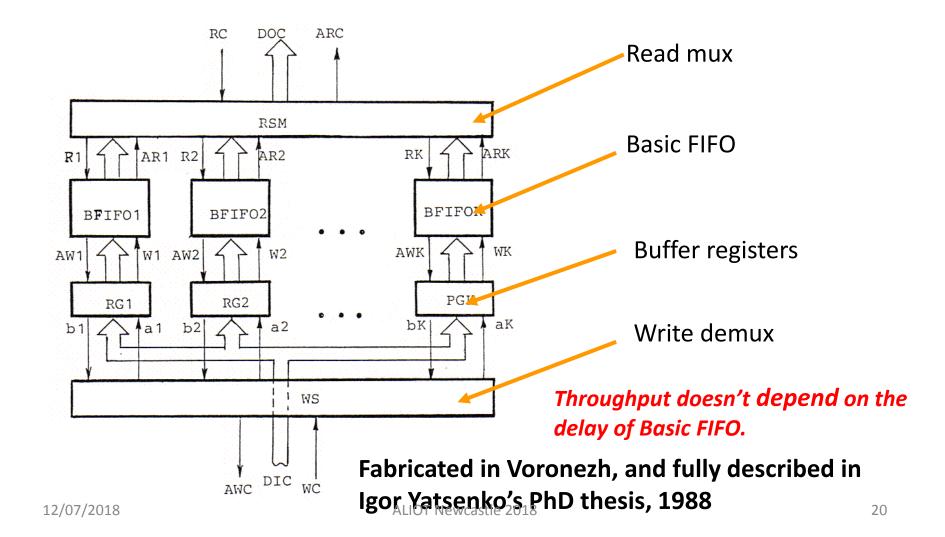
Data (DR,DS) is encoded using 3-of-6 Sperner code (16 data values for half-byte, plus 4 tokens for ring acquisition protocol) AR, AS – acknowledgements RR, RS – spare (for self-repair) lines

Self-Timed Ring published

- V.I. Varshavsky, V.B. Marakhovsky, L.Ya. Rosenblum, Yu.S. Tatarinov, V.Ya. Volodarskii, A.V. Yakovlev, THREE papers in Avtomatika & Vychislitelnaya Tekhnika, Control and Computer Science (Translated from Russian, Alerton Press), (1988-1989)
- A. Yakovlev, V. Varshavsky, V. Marakhovsky and A. Semenov, "Designing and asynchronous pipeline token ring interface", Proc. 2nd Working Conf. on Asynch. Design Methodologies, London, May 1995, IEEE CS Press, N.Y., pp. 32-41, 1995.

Self-Timed FIFO (1985-1988)

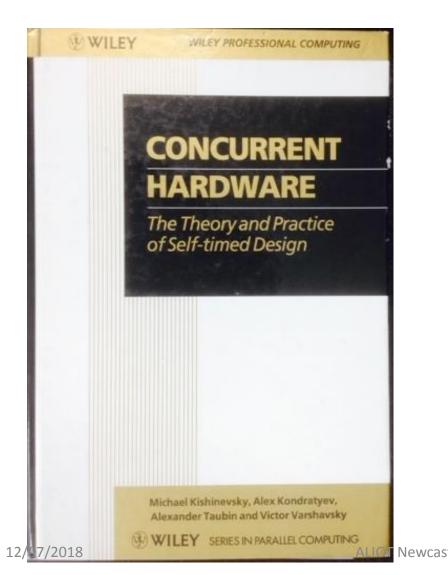
Basic FIFO are connected using a wagging-tail buffer method



LETI people in async world: recent past

- From 1993: Varshavsky is in Univ of Aizu in Japan with part of his group (Marakhovsky, Kishinevsky, Kondratyev, Taubin)
- From 1990: Rosenblum is in the USA
- From 1991: Yakovlev is in the UK
- Starodoubtsev is in Russia (NTO of Academy of Sciences)
- Active work on synthesis and verification methods for STGs
- Collaboration between Yakovlev, Kishinevsky, Kondratyev, Taubin and Cortadella (Barcelona) and Lavagno (Turin)
 - Many papers in IEEE Transactions and leading confs: ASYNC, DAC
 - Software tool Petrify
 - Springer book published in 2002

Group's Books on Async Design

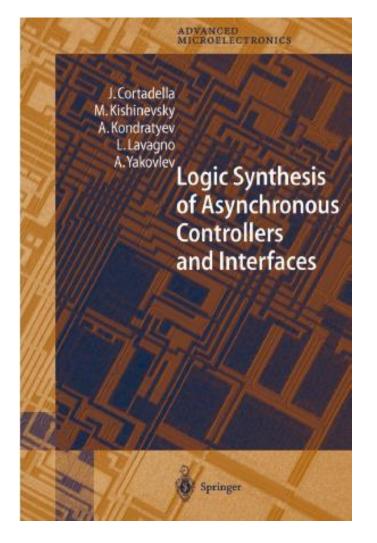


Hardware Design and Petri Nets

Edited by Alex Yakovlev Luis Gomes Luciano Lavagno

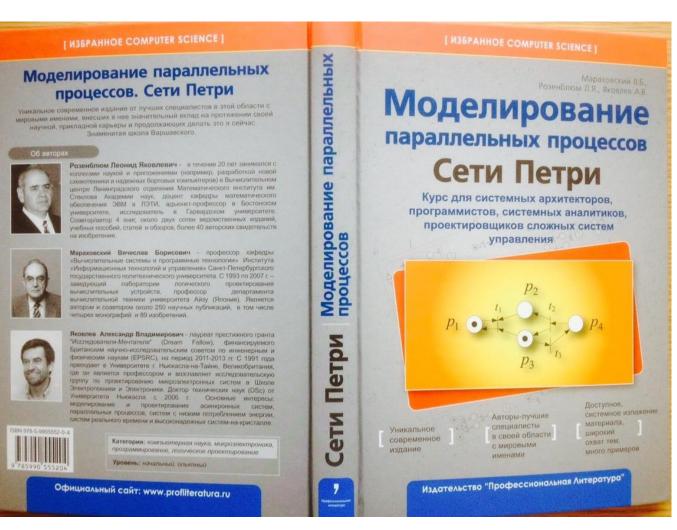
Kluwer Academic Publishers

Book on Petrify Method and Tools



- The book on Signal Graphs Methods and tools.
- Collaboration with Cortadella (UPC Barcelona) and Lavagno (Turin, Cadence)
- Kishinevsky (Intel), Kondratyev (Theseus Logic, Cadence)

The most recent book from the group

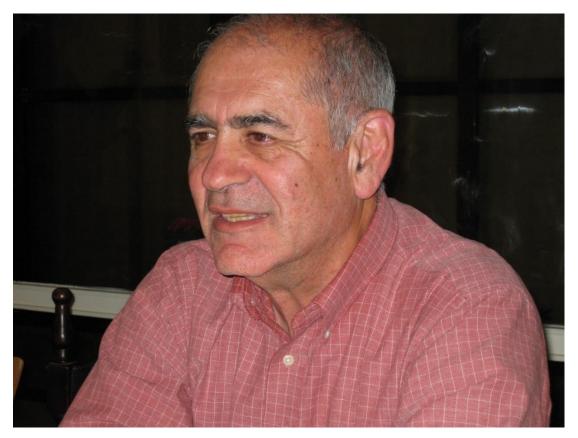


Some 'remainders' of Varshavsky's group



Victor Varshavsky with members of his group at our reunion in Eilat, Israel, in 2000; left to right: Alex Kondratyev, Alexander Taubin, Michael Kishinevsky, Victor Varshavsky, Alex Yakovlev (me) and Maria Yakovlev(my wife).

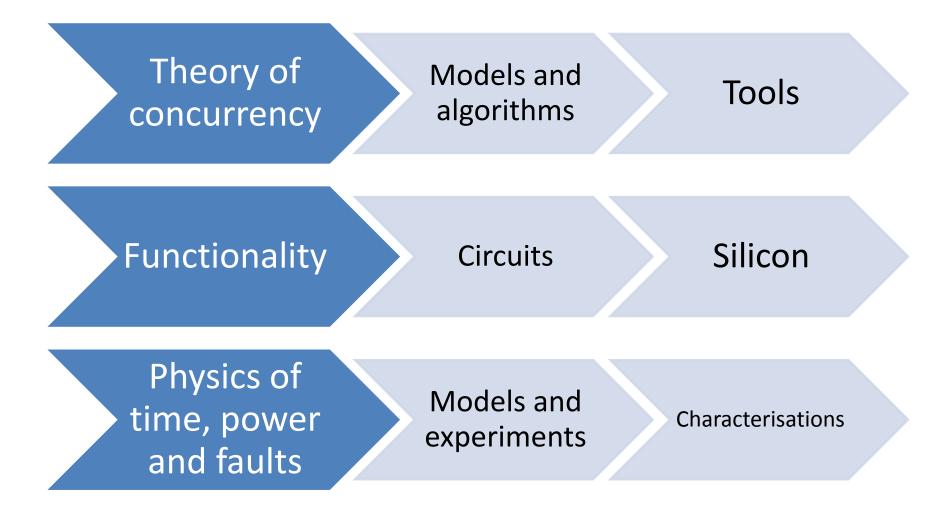
Leonid Rosenblum



Leonid Rosenblum, who introduced Petri nets into asynchronous circuit modelling and popularized Petri nets in USSR in the 70s and 80s.

ALIOT Newcastle 2018

Three main research threads



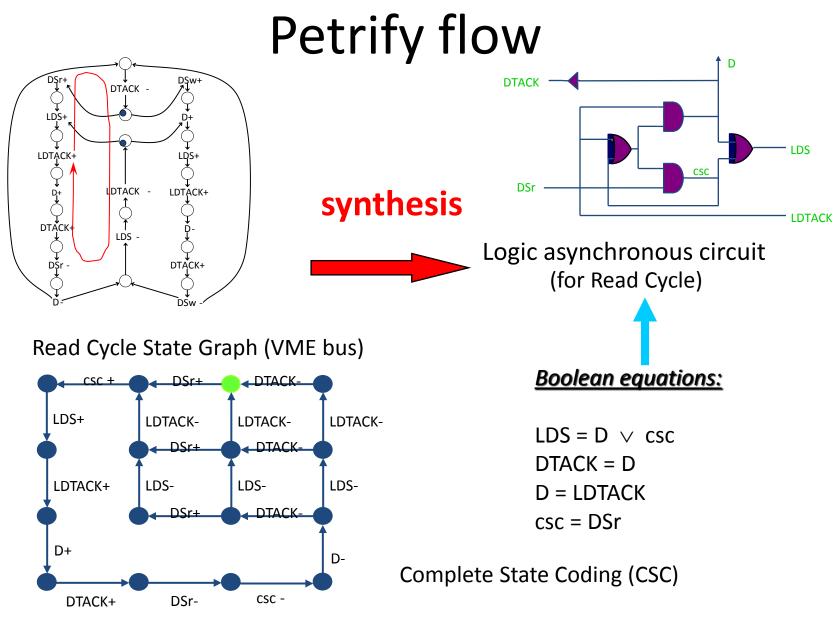
But how has it all unrolled in Newcastle ... from 1990s

The key element for success

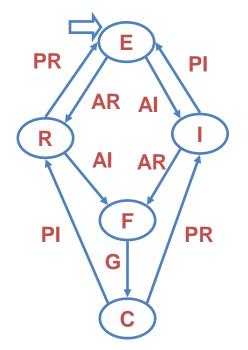
- Work of a team where all aspects of talent have a chance to develop and flourish
- The team both locally (Newcastle) and globally (worldwide)
- Async research is a wonderful area which combines the three main threads and different people can exploit the three main categories of strengths:
 - Computer programming
 - Electronic (Circuits and Systems) design
 - Analytical

1990s: Logic Synthesis of Asynchronous Control

- Formalisation of STGs and relation to state-based models; theory, models and tools for async controller synthesis, async processor prototyping
- Projects: ASAP, ASTI, ACiD-WG
- People: Jordi Cortadella, Luciano Lavagno, Alex Kondratyev, Mike Kishinevsky, Alexander Taubin, Albert Koelmans, Alexei Petrov, Alex Semenov, Enric Pastor, Lee Lloyd, Marta Pietkiewicz-Koutny
- Some highlights: synthesis flow, monotonic cover, CSC solving, regions for async, Or-causality, relative-timing ('lazy' approach), unfolding-based verification (incl. nets with read-arcs), tool Petrify; application to Molnar's counterflow pipeline design

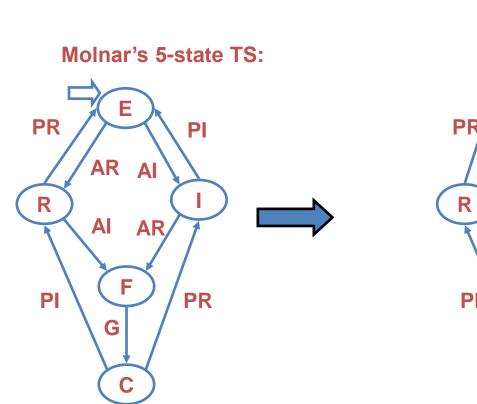


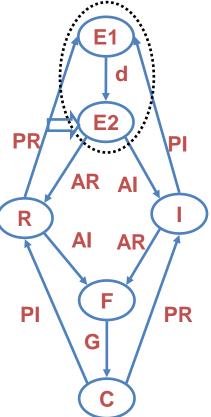
Molnar's 5-state TS:

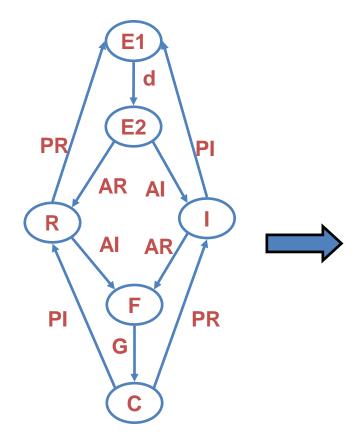


Solution:

Split a state (E) and insert a silent action (d), preserving behavioural (observational) equivalence

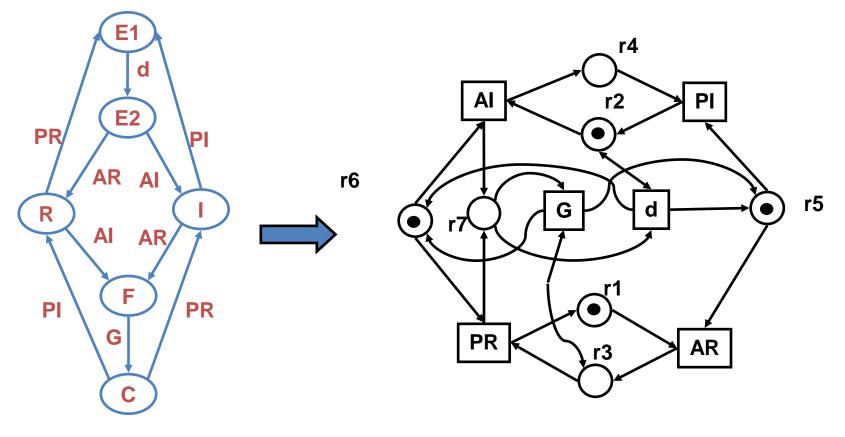






Semi-elementary TS

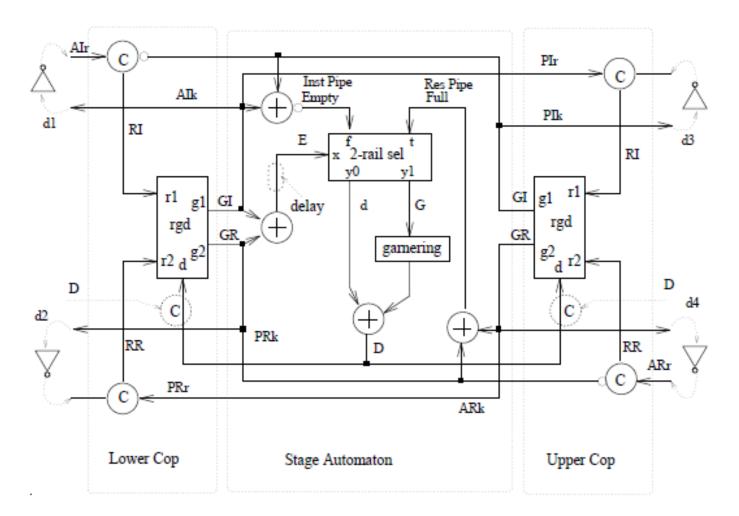
Minimal set of regions: r1 = {E1,E2,I} <- pre(AR), post(PR) r2 = {E1,E2,R} <- pre(AI), post (PI), co(d) r3 = {R,F,C} <- pre(PR), post(AR), co(G) r4 = {I,F,C} <- pre(PI), post(AI) r5 = {E2,I,C} <- pre(PI,AR), post(G,d) r6 = {E2,R,C} <- pre(PR,AI), post(G,d) r7 = {E1,I,F} <- pre(G,d), post(PR,AI)



Semi-elementary TS

Semi-elementary Petri net

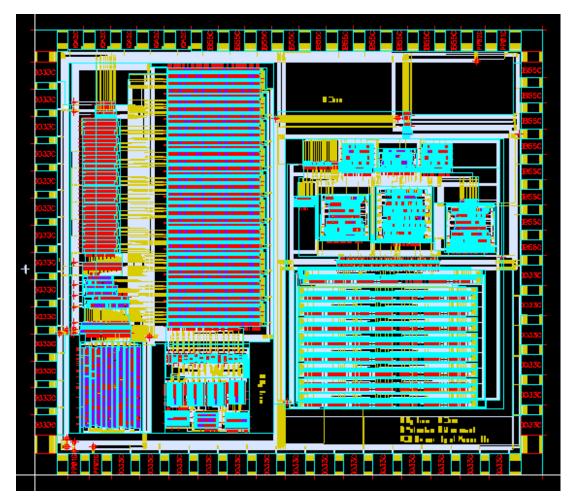
CFPP Implementation



1990s: Arbiter design

- Design of circuits with internal conflicts, all kinds of arbiters, modelling and analysis of arbiters asynchronous communication mechanisms
- Projects: HADES
- Highlights: method for factorising mutexes, new priority arbiters, 3 and 4 slot ACMs with mutexes
- People: David Kinniment, Fei Xia, Alex Bystrov, Delong Shang, Ian Clark, Bo Gao, Isi Mitrani, Tony Davies, Hugo Simpson and Eric Campbell (MBDA)
- First VLSI chip HADIC (0.6um CMOS, Europractice)

HADIC: Our first Newcastle's asynchronous chip



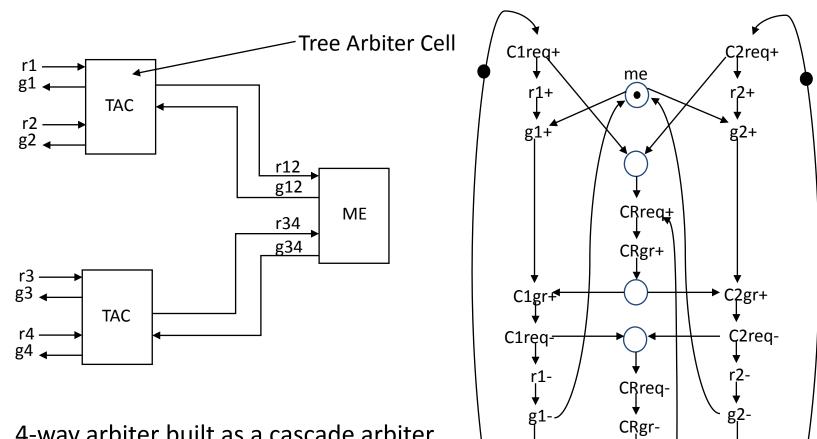
Designed using Europractice tools On a SUN w/station called Sadko

What HADIC had

- a **three-way ordered arbiter** comprising a three-way mutex element, based on a three-stable FF
- a **one-hot pipeline FIFO buffer**, and a request mask circuit, to demonstrate the effect of request-ordering in avoiding secondary metastability effects discussed,
- an **eight-way ordered arbiter with a low-latency tree arbiter** and a low-latency cyclic FIFO buffer, to demonstrate the role of low latency techniques in ordered arbitration with large numbers of channels,
- an **eight-way static priority arbiter** with a non-conventional (non-topological) priority allocation mapping to demonstrate the effect of static prioritisation defined as an arbitrary function,
- an **eight-way dynamic priority arbiter** with four priority levels (the priority data encoded in dual rail code) to study the variations of arbiter latency against different conditions of issuing prioritised requests and effects of early control signal propagation in the priority logic with a tree structure,
- an **eight-way token ring arbiter** with a one-level cascaded buffer, to study the effects of improved done-to-grant and practical fairness via buffering,
- a **comparator with a latch for an A/D converter**, to study the metastability effects in value/time domains,
- an on-chip tester, for arbiters and asynchronous communciation mechanisms,

^{12/07/2018} containing a number of original circuits to experiment with system-on-chip testing.³⁹

Eaxmple: Multiway Arbiters (tree)

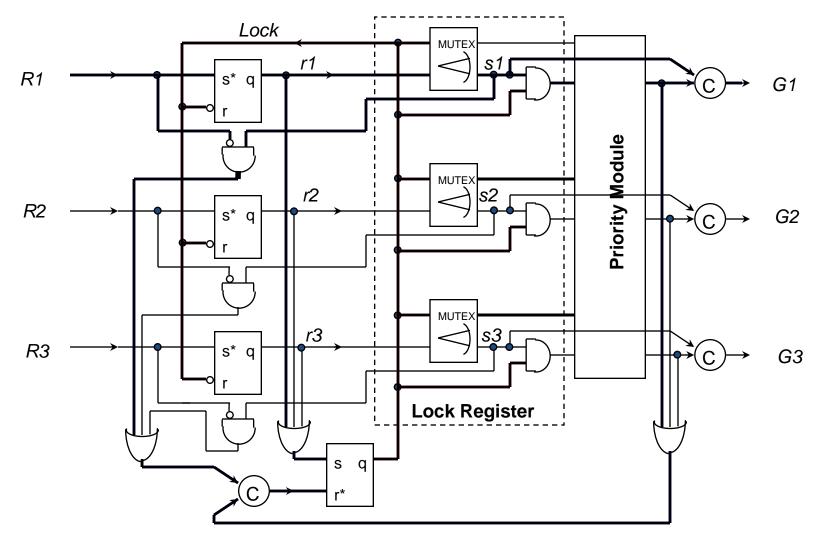


4-way arbiter built as a cascade arbiter with OR-causal request propagation

C1gr>

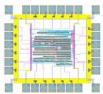
C2gr-

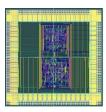
Example: static priority arbiter



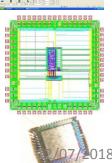




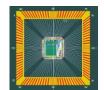


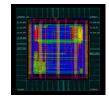


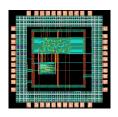


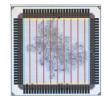


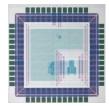
- Arbitration: HADIC1 (2000, CMOS 0.6um) asynchronous arbiters, ADC, Async communication mechanism
- Time measurement: TMC (2003, CMOS 0.2um, collab. with Sun) time amplifier and time-to-digital converter
- Synchronization: SYRINGE1 (2006, UMC 0.18um), SYRINGE2 (2008, UMC, 0.13um) metastability measurement, robust synchronizer
- Security: SCREEN1 (2005, AMS 0.35um), SCREEN2 (2006, AMS 0.18um) AES cores with power balancing; SURE (2008, UMC 0.13um), Galois Encoded Logic, Script1 &2 (2010, TSMC 90nm)
- Networks on Chip: NEGUS1 (2007, UMC 0.13um), phaseencoding signalling, NeuroNOC (2009, TSMC 90nm) reconfigurable neural network, 3D-chip (2011, MIT Lincoln Labs, 150nm with TSVs)
- Energy harvesting: Holistic1 (2011, UMC 90nm), self-timed SRAM, Holistic 2 (2012, UMC 0.18um), reference free voltage sensor
- Power proportionality: Async 8051 (2013, STM CMP 0.13um)
- SAVVIE and A4A chips: to be reported







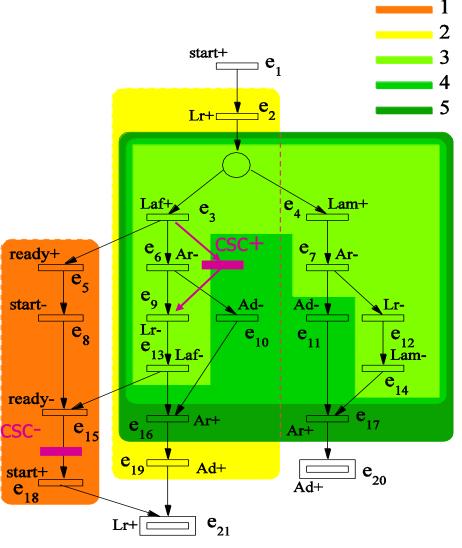




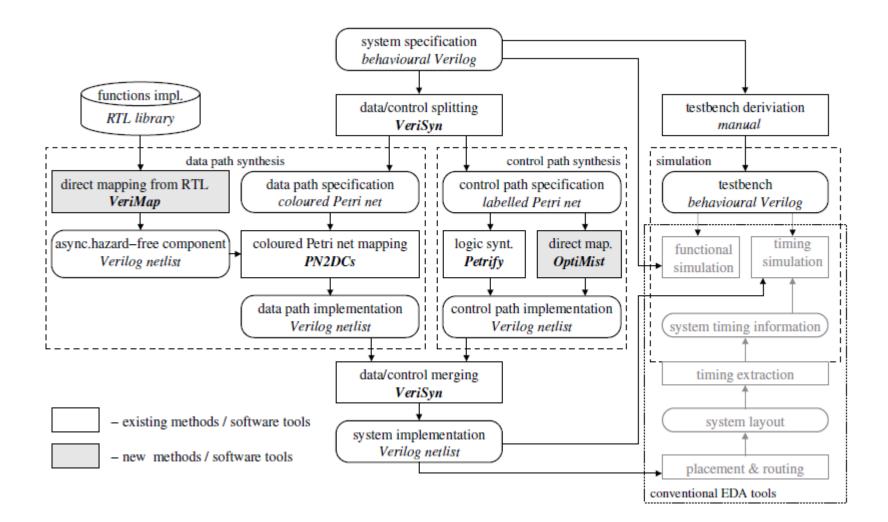
2000-2005: Synthesis and Testing

- Synthesis line: visualisation, unfolding-based synthesis, direct synthesis, negative gate synthesis, generalised ORcausality; regions, ACM synthesis
- Projects: BREACH, MOVIE, BESST, STELLA
- People: Maciej Koutny, Agnes Madalinski, Victor Khomenko, Danil Sokolov, Frank Burns, Deepali Koppad, Nikolai Starodoubtsev, Josep Carmona, Fei Hao, Kyller Gorgonyo
- Highlights: Core maps for visualization, Tracker-Bouncer for direct mapping, Lots of Tools: ConfRes, PUNF/MPSAT, Verisyn, Optimist, PN2DCs; application of direct mapping to S. Furber's duplex interface design; online testing for async protocols; generalised regions synthesis

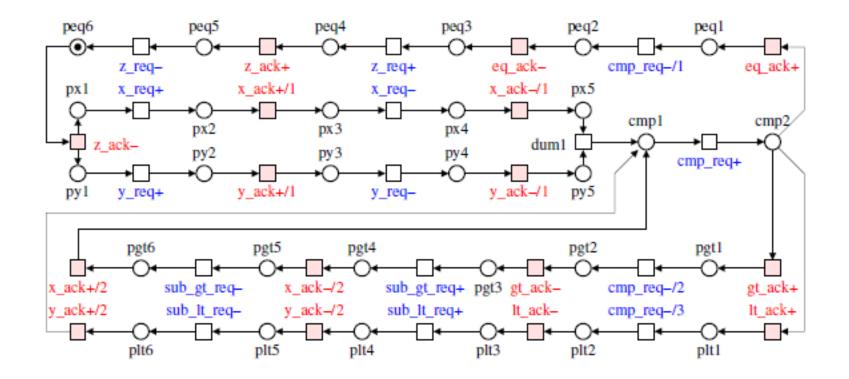
CSC Core visualisation using unfoldings



BESST Design Flow

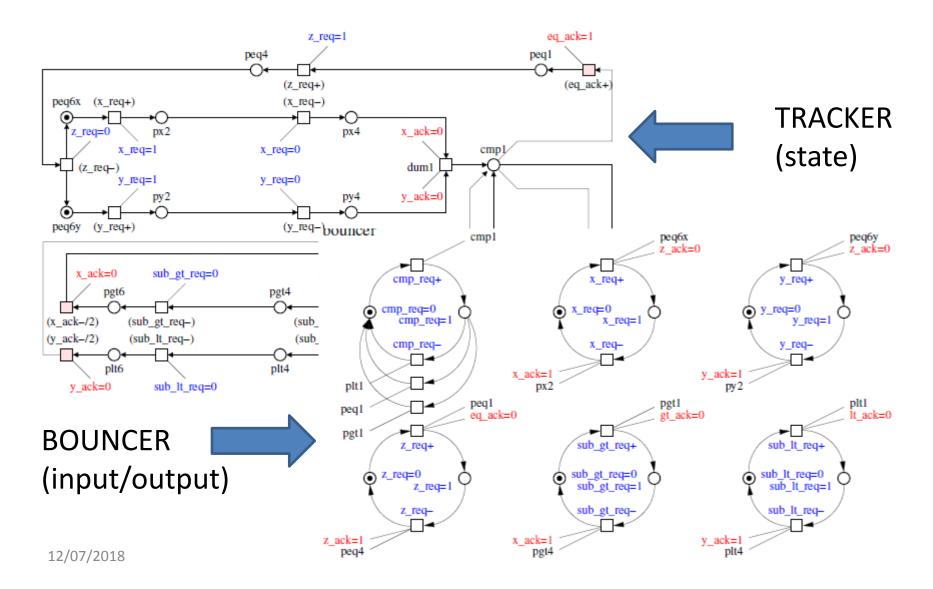


Example: GCD using Optimist flow

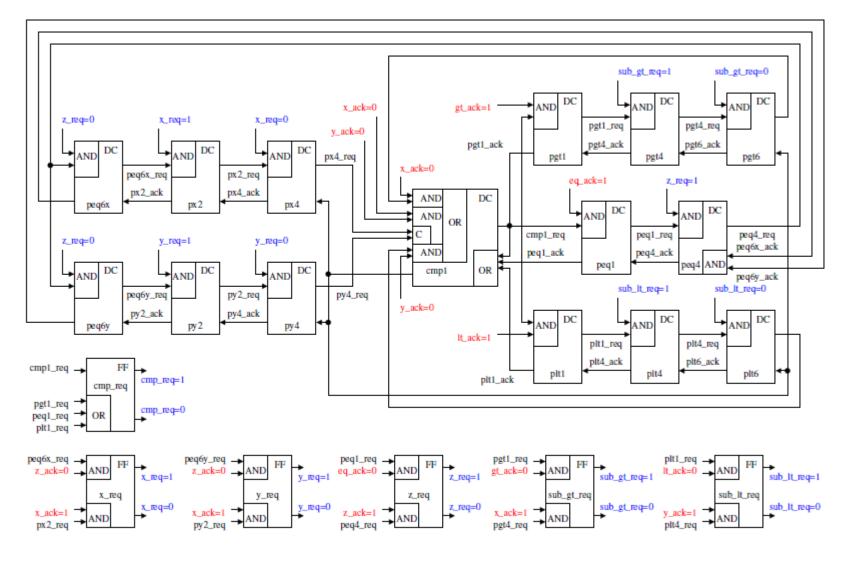


Range of Optimisations for latency and size could be applied

Example: GCD using Optimist



Direct mapping: David cell circuit

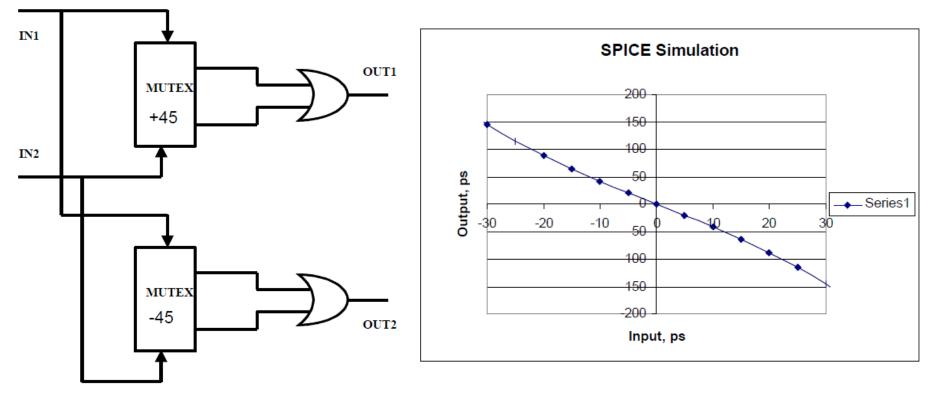


2000-2010: Metastability and Event Processing

- Synchronisation, time to digital conversion, time amplification, robust synchronization, speculative synchronization, dynamic systems models of analogue circuits
- Projects: COHERENT, STEP, SYRINGE, RelCel
- People: Gordon Russell, Graeme Chester, Oleg Maevsky, Amir Abbas, Nikos Minas, Keith Heron, Jun Zhou, Fei Hao, Yuan Chen, Stan Golubcovs, Ghaith Tarawneh, Ioannis Syranidis, Ahmed Alahmadi, Mohammed Alshaikh, James Guido; collaboration with Charles Dike (Intel), Ian Jones (Sun/Oracle) Marc Renaudin (Grenoble)
- Highlights: Metastability physical measurements, RNG, TD converters, time amplifier, FPGA synchronsers, speculative synchronizers; reconfigurable synchronizers, soft arbiters; Several China (win Europeration and Sup)

Chips (via Europractice and Sun) 12/07/2018 ALIOT Newcastle 2018

Time difference amplifier



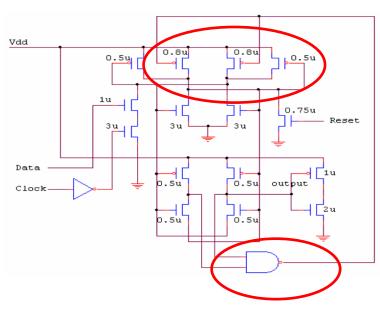
 $\Delta_{out} = \tau . \ln(45 + \Delta_{in}) - \tau . \ln(45 - \Delta_{in})$

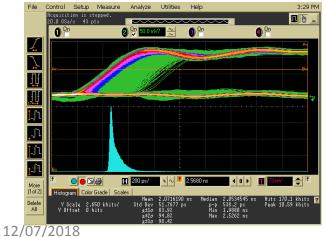
Was used as the first cascade of time2digital converter

ALIOT Newcastle 2018

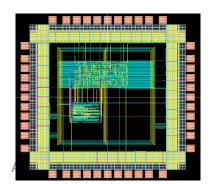
Robust synchronizer

Dynamic feedback control



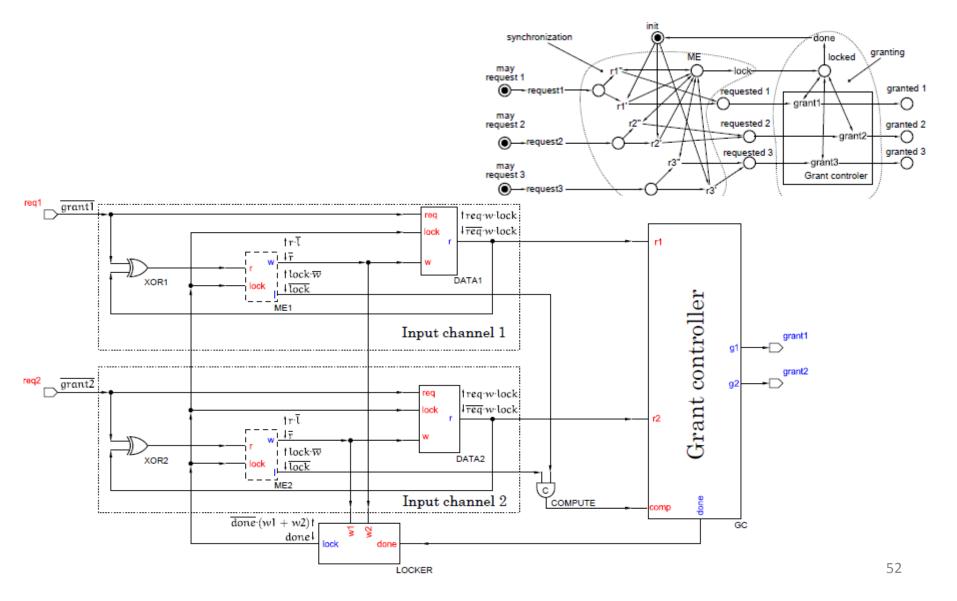


| | Measurement Results (ps) | | | |
|--------|--------------------------|----------------------|------------------------|----------------------|
| Vdd(v) | Jamb Latch B | | Robust Synchronizer | |
| | >10 ⁻¹⁴ s | <10 ⁻¹⁴ s | >10 ⁻¹⁴ s | <10 ⁻¹⁴ s |
| 1.8 | 19.44 | 35.55 | 15.27 | 34.92 |
| 1.7 | 21.75 | 37.29 | 16.53 | 35.76 |
| 1.6 | 25.64 | 40.93 | 19.38 | 38.25 |
| 1.5 | 28.77 | 52.36 | 20.29 | 43.07 |
| 1.4 | 36.22 | 66.17 | 23.75 | 50.36 |
| 1.35 | 45.43 | 75.35 | 28.51 | 58.19 |



τ (metastability time constant) vs Vdd

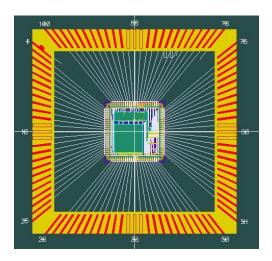
Multiway arbiters: general approach



2004-2009: Security

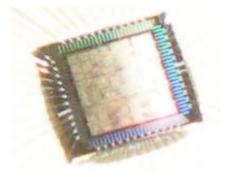
- Asynchronous circuits for power-attack resistance, mixed radix synthesis
- Projects: SCREEN, SURE, STRIP
- People: Julian Murphy, Ashur Rafiev, Atmel people (Ewart Grey, Russell Hobson, Steve Pickles)
- Highlights: Automatic translation from single to dual rail; Tools: Verimap; RMMixed; Many Chips: SCREEN1/2, SURE, STRIP1/2

Some Security Chips



SCREEN1: AES encryption core design using alternating spacer dual-rail signalling for power balancing (generated from synchronous RTL using in-house tool VeriMap)

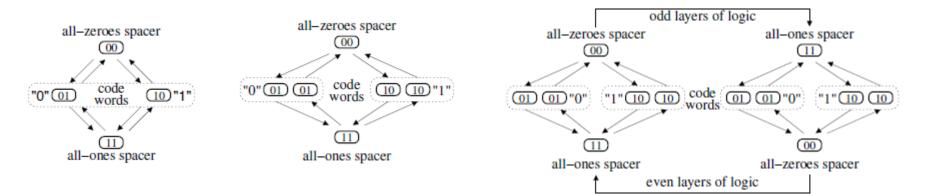
SCREEN 2: different (synchronous single-rail and power-balanced) versions of AES Sbox.



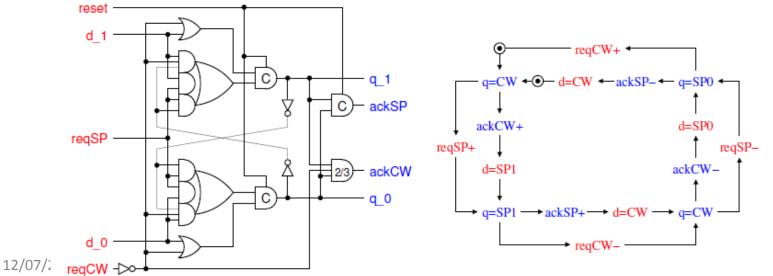
SURE:

Galois Encoded Logic AES-128 IP core. Implements side-channel secure "high radix" AES-128 encryption in a fully parallel architecture to investigate the security and study the market feasibility of the technology.

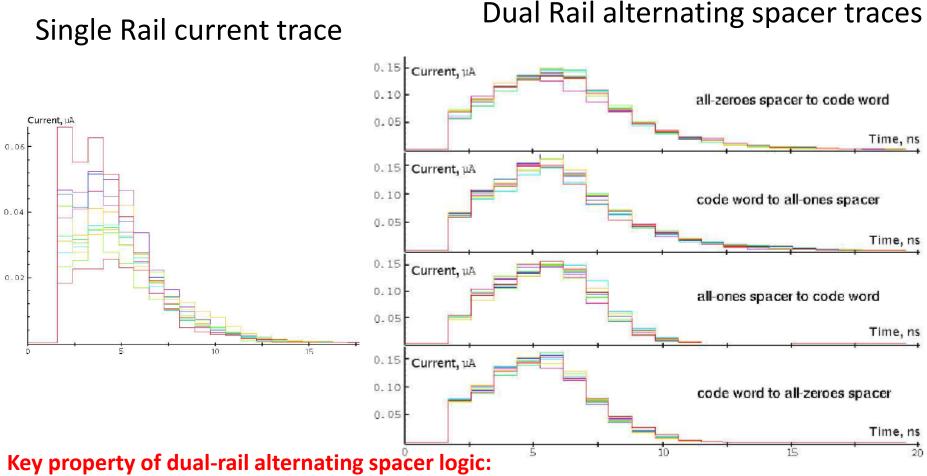
Verimap: Dual spacer and negative gate optimization



Alternating spacer dual-rail flip-flop



Example: Verimap AES design



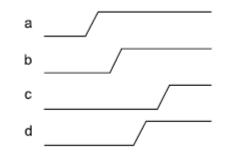
For every operational cycle of the circuit all logic nodes fire => maximum power balancing

12/07/2018

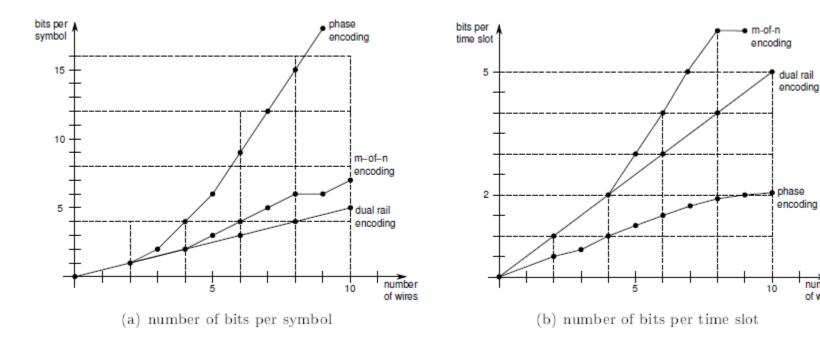
2005-2015: NoCs, FPGAs and New models

- New types of interconnects, communications, crosstalk, routing, neural network NoCs, surface wave comms, async FPGAs
- Projects: NEGUS
- People: Enzo D'Alessandro, Andrey Mokhov, Basel Halak, Sohini Dasgupta, Robin Emery, Terrence Mak, Ra'ed Aldujaili, Nizar Dahir, Ammar Karkar, Hock-Soon Low; Southampton (Bashir Al-Hashimi and Simon Ogg)
- Highlights: Phase-encoding, CPOGs, cross-talk resilient interconnects, models of GALS and wrappers, on-line deadlock detection; power/thermal model; Async FPGA flow; Chips: Negus1 and NeuralNoC

Phase encoding



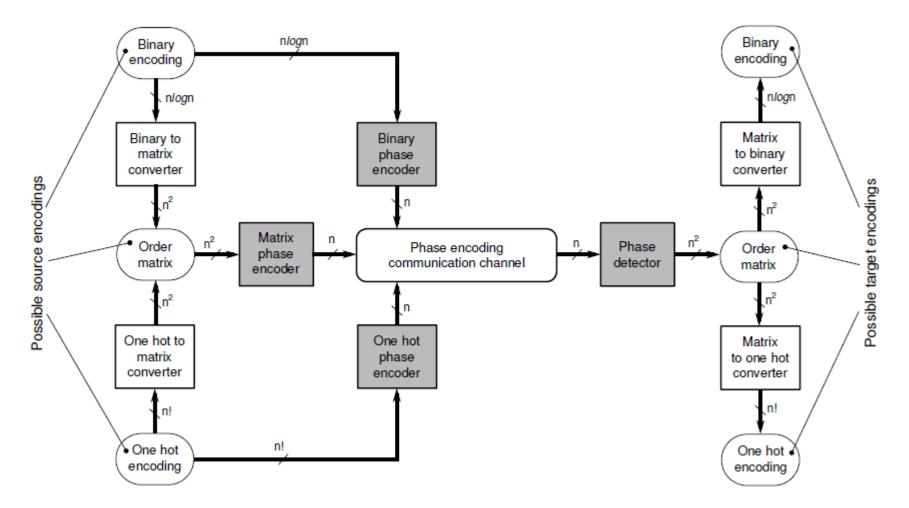
An example of phase encoded symbol "abdc" (total no of symbols is 4!=24)



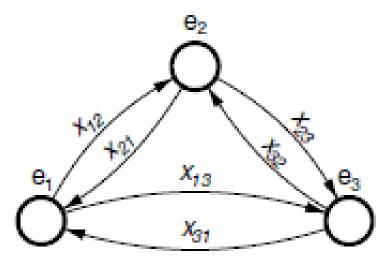
number

of wires

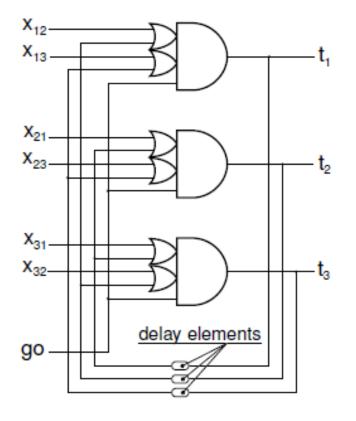
Phase encoding comms



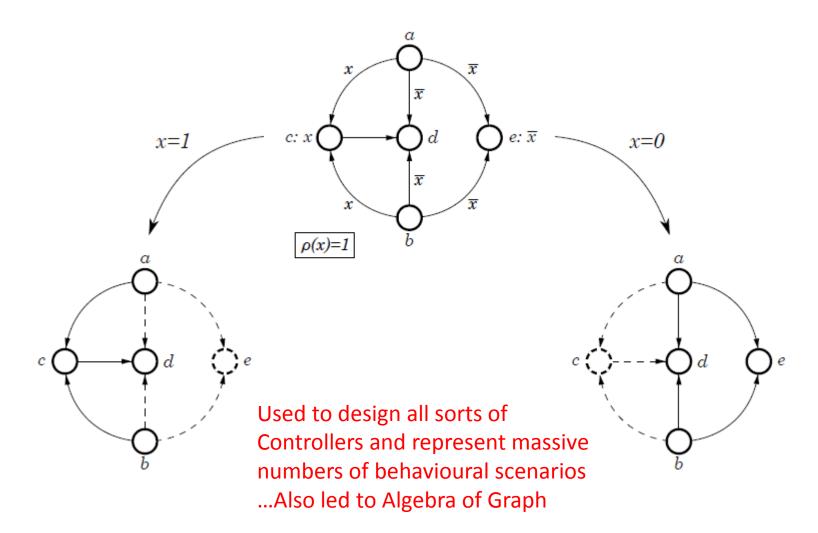
Synthesis of logic for phase encoding



(a) CPOG specification



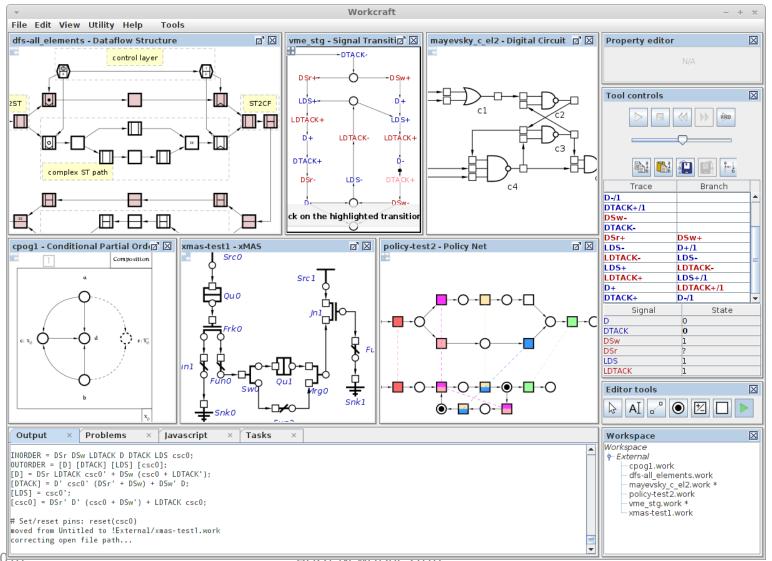
Conditional Partial Order Graphs



2005-2010: Synthesis

- Interpreted graph models, Compositionality, Data-path synthesis, SDFS and anti-tokens, Strong and Weak Indicatability, Synthesis of circuits with isochronic forks; Online data path testing
- Projects: SEDATE, VERDAD, GAELS
- People: Ivan Poliakov, Arseniy Alekseyev, Yu Zhou (Joe), Kelvin Gardiner, Yu Li; Manchester (Doug Edwards, Will Toms, Charlie Brej), Augusburg (Walter Vogler, Dominik Wist)
- Highlights: Theory for compositions, IF relaxation; Workcraft toolkit and many plugins

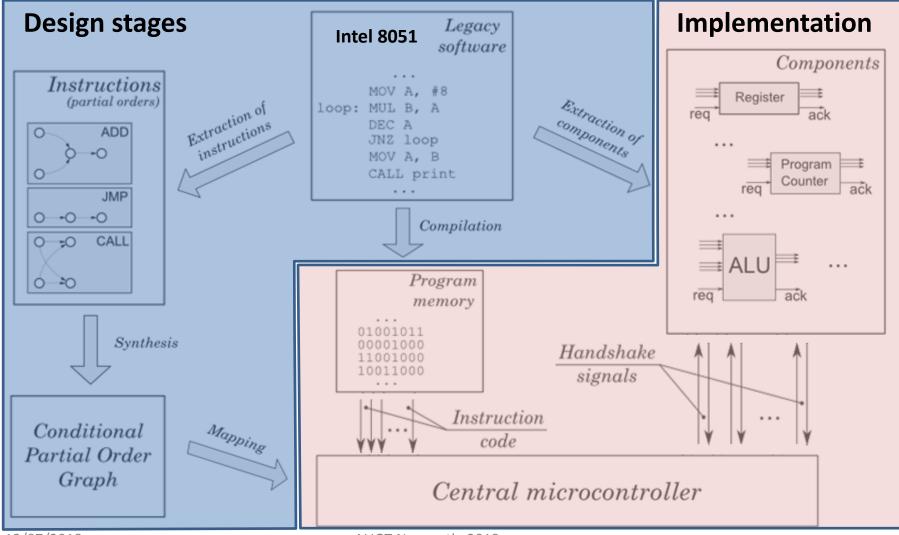
Workcraft toolkit



2007-2015: Low Power and Energymodulated computing

- Low power processors, energy harvesting and async, power-proportional design, power variation resilient circuits, new power supply methods, power gating and new
- People: Chihoon Shen (Korea), Reza Ramezani, Max Rykunov, Abdullah Baz, Xuefu Zhang, Haider Alrudainy, Alessandro De Gennaro
- Highlights: low power M430, Charge-to-code converter, Speed-independent SRAM, Power-proportional 8051, CBBs versus SCCs, NEMS/MEMS for power gating; Chips: Reference-free sensor, ASRAM, 8051, Async Pipeline

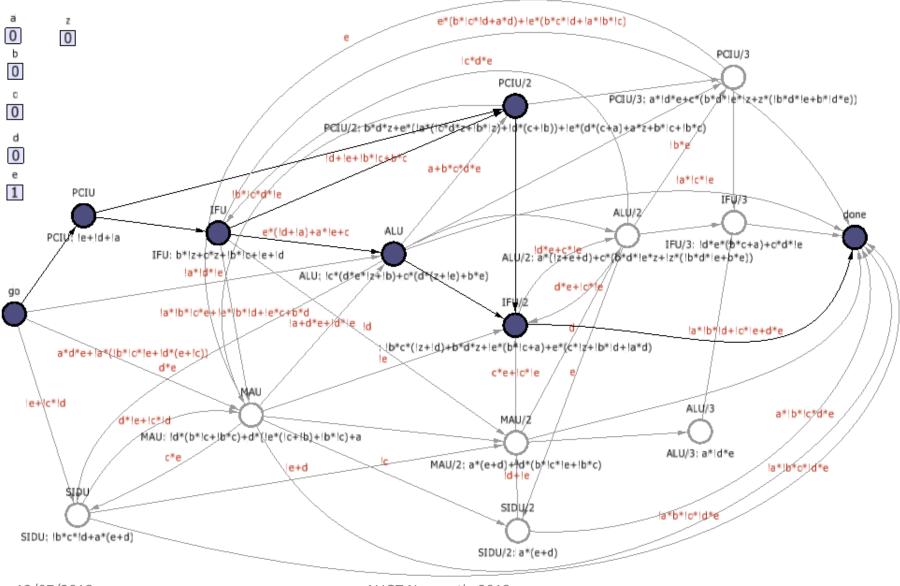
Conceptual view of the CPU design process



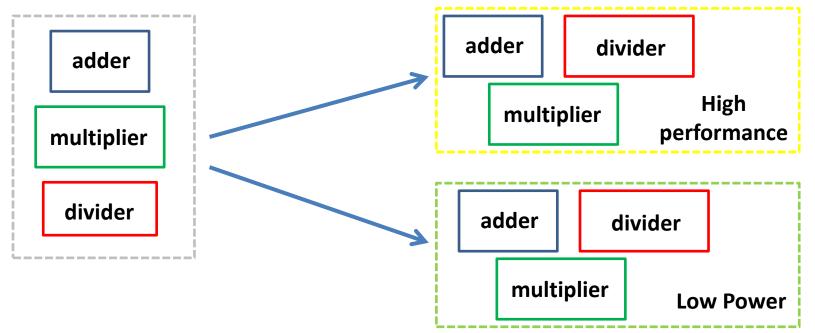
12/07/2018

ALIOT Newcastle 2018

Intel 8051 ISA...



Intel 8051 Datapath...



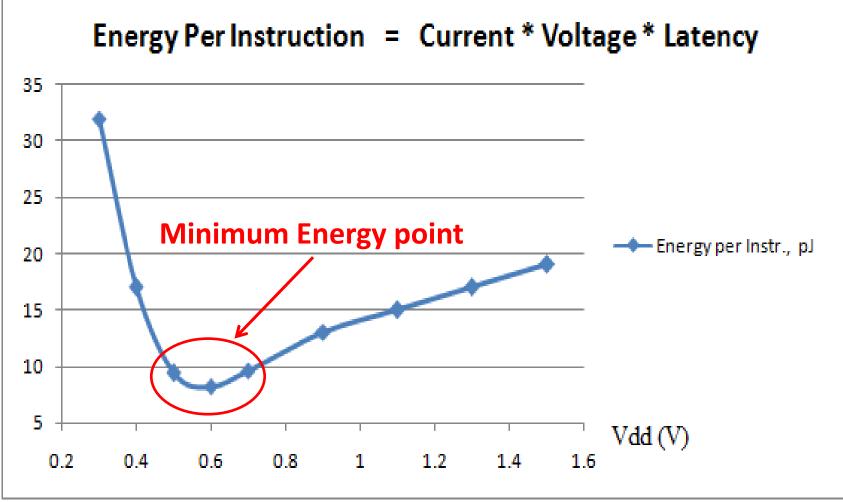
- ✓ Fully asynchronous implementation (bundled data protocol)
 - adjustable delay lines
- ✓ Fault tolerance operation

✓ DFT integration

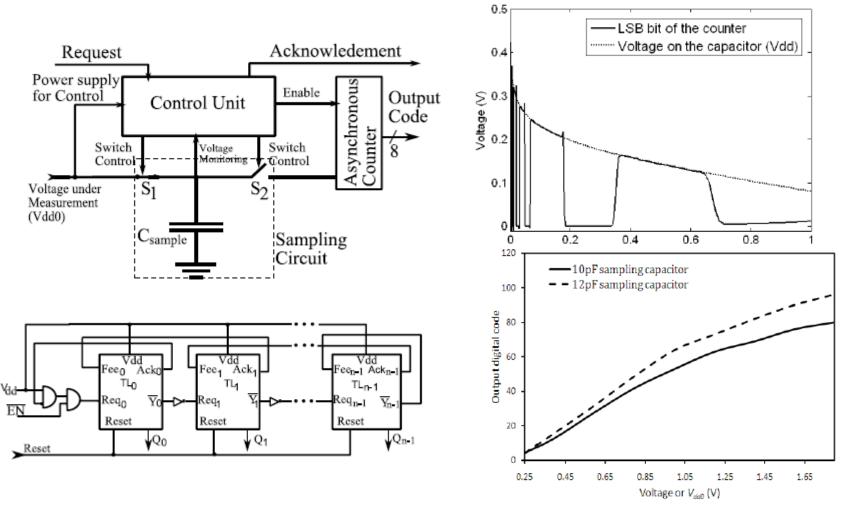
Some measurements...

- 0.89V to 1.5V: full capability mode.
- 0.74V to 0.89V: at 0.89V the RAM starts to fail, so the chip operates using
- 0.22V to 0.74V: at 0.74V the program counter starts to fail, however the control logic synthesised using the CPOG model continues to operate correctly down to 0.22V
- 67 MIPS at 1.2 V.
- ~2700 instructions per second at 0.25V.

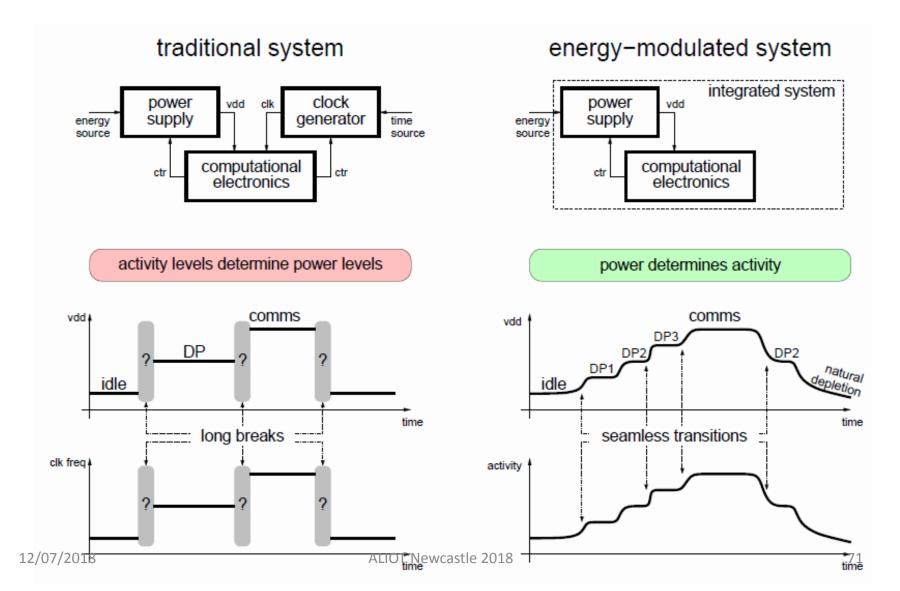
Energy Efficiency (measurements on real silicon - asynchronous 8051)



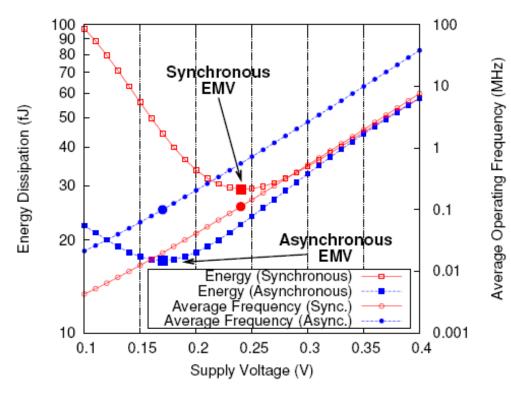
Reference-free Voltage Sensor



Traditional vs energy-modulated view



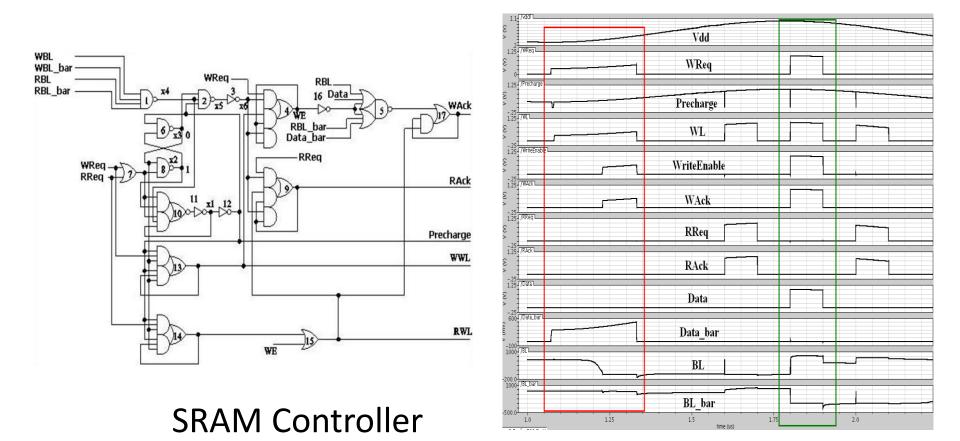
Synchronous vs Self-Timed Design (in terms of energy efficiency)



Asynchronous (selftimed) logic can provide completion detection and thus reduce the interval of leakage to minimum, thereby doing nothing well!

Source: Akgun et al, ASYNC'10

Self-timed SRAM

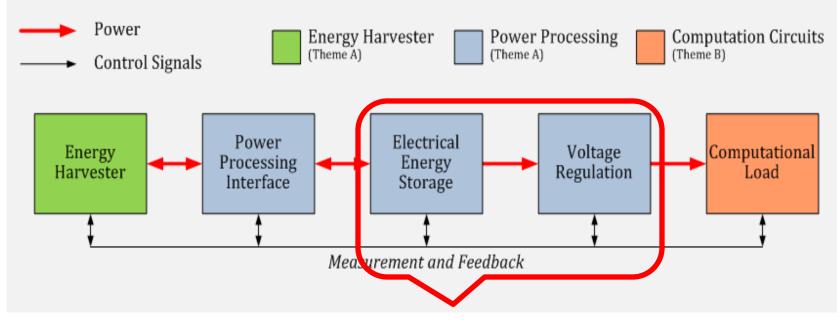


Today: living with Analog, new model mining, ...

- Asynchronous for "little digital", self-powered sensors, ADCs; synthesis of STG specifications
- Projects: SAVVIE, A4A
- People: Kaiyuan Gao, Yuqing Xu, Vladimir Dubikhin, Jonathan Beaumont, Austin Ogweno, Benafa Oyinkuro, Sergey Mileiko; Bernard Stark (Bristol)
- Industrial collaboration: Dialog Semiconductor

Energy harvesting systems

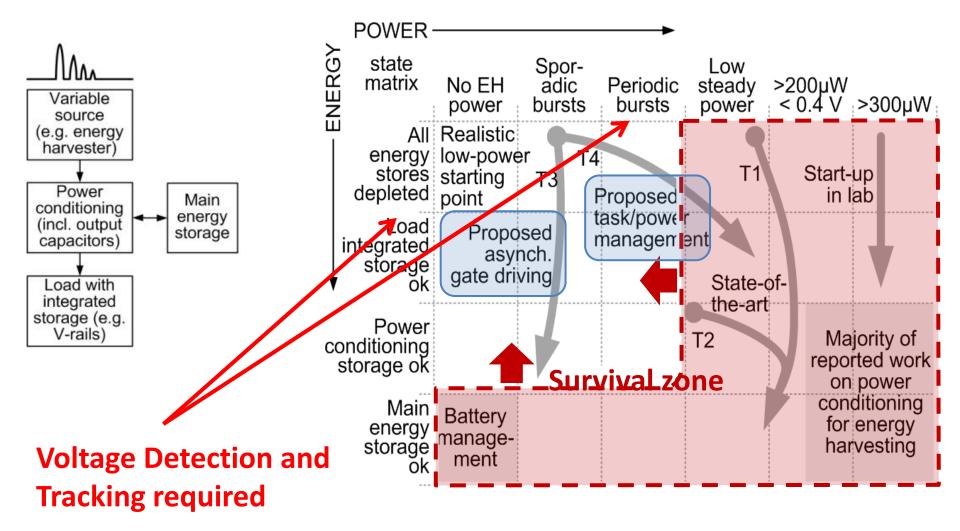
Using Holistic project vision:



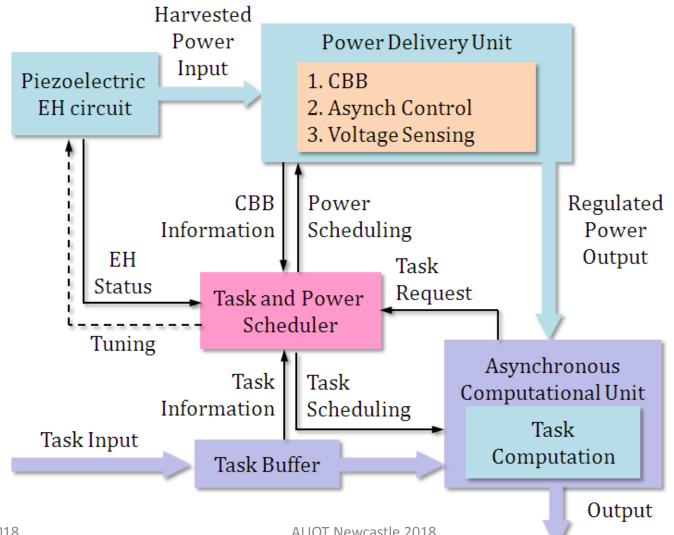
Sporadic source of energy does not allow for fancy power processing and therefore large storage

ALIOT Newcastle 2018

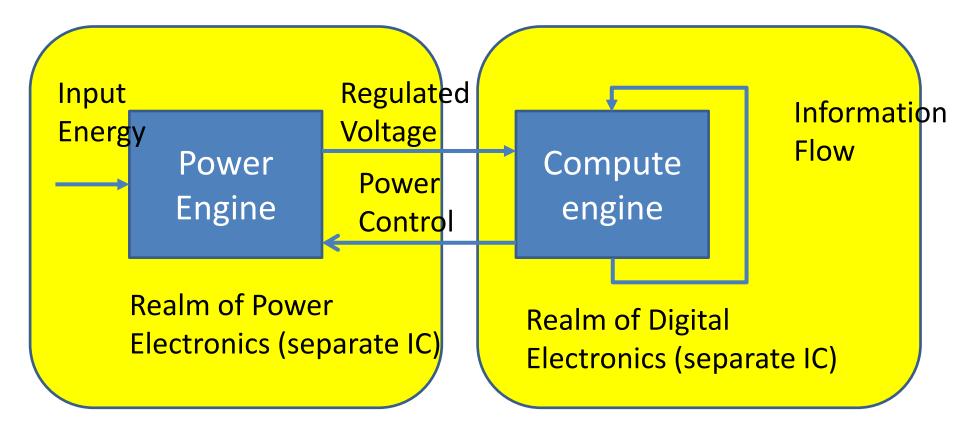
Staying alive in variable, intermittent, lowpower environments (Savvie Project)



From Holistic systems to Power-**Compute Co-design**

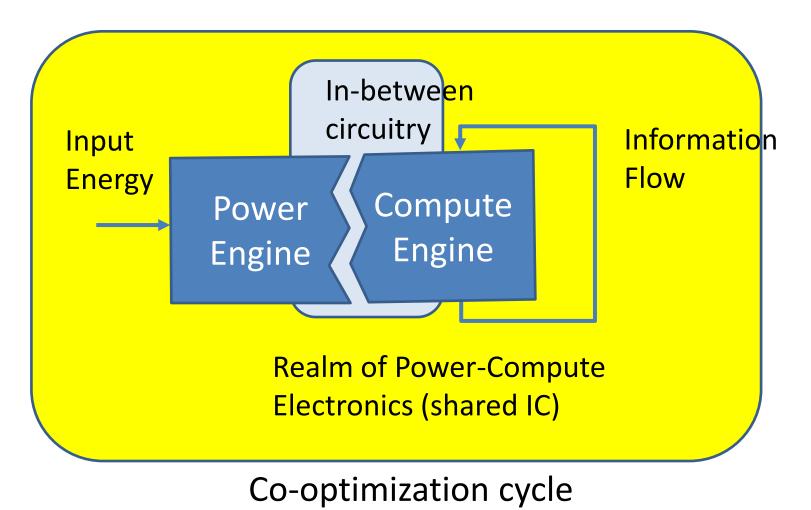


Traditional Power-Compute Divide



Separate optimization cycles

Power-Compute Co-design



Thank you!