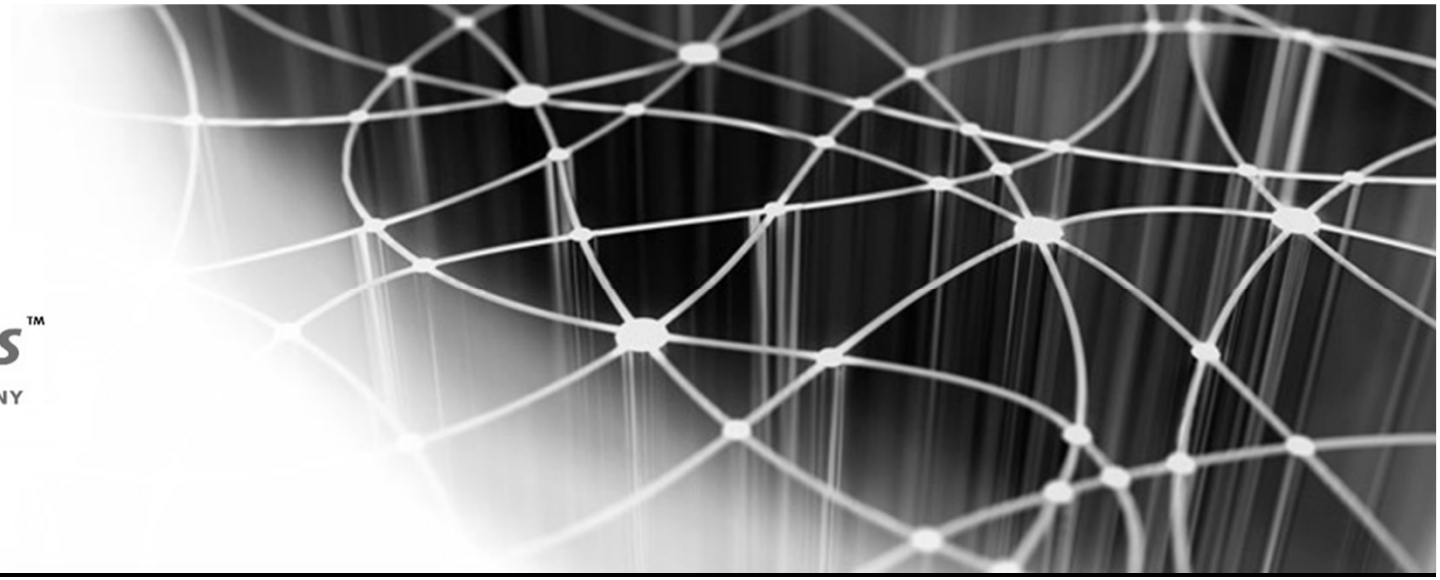


Arteris™
THE NETWORK-ON-CHIP COMPANY

Heterogeneous, Distributed and Scalable Cache-Coherent Interconnect

Scale system performance faster than Moore's Law will currently allow

K. Charles Janac
President and CEO, Arteris

MPSoC'16

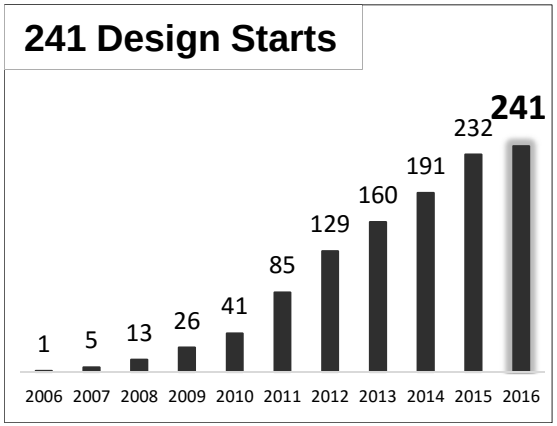
MPSoC Conference 2016
Nara, Japan, July 13, 2016

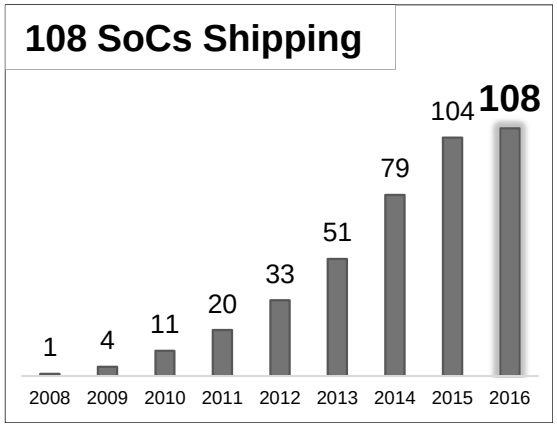
Copyright © 2016 Arteris

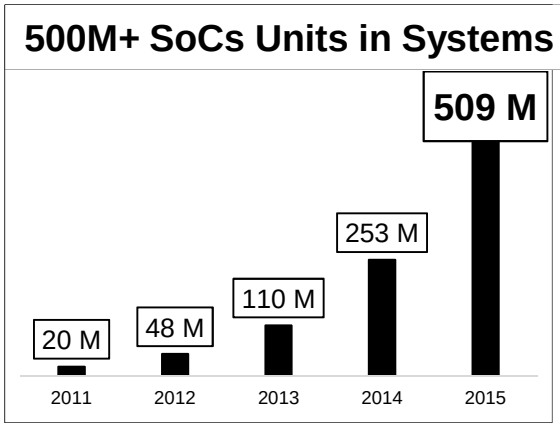
Arteris customers and their SoCs

Current as of 8 June 2016

SAMSUNG	QUALCOMM	HISILICON	SPREADTRUM	小米 xiaomi.com	Very Large SoC Maker	ZTE	HISILICON
MARVELL	Rockchip	MOBILEYE	RDA	LG	INFOTM	Rambus	Hisense
Allwinner Technology	Major Automotive OEM	Major Auto & CE SoC Maker	RENESAS	NXP	TEXAS INSTRUMENTS	SK hynix	新岸线 NUFRONT
ALTERA	LG	Major System OEM	Toshiba	Japan System OEM	Automotive SoC Maker	Schneider Electric	applied micro
Japan Tier 1 SoC Maker	SOCLE	Open-Silicon	MegaChips	GUC	Large Drone Maker	SEQUANS COMMUNICATIONS	ISTUARY INNOVATION GROUP
Major SSD Vendor	MONTAGE Technology	pixelworks	NTTElectronics	BLUWIRELESS TECHNOLOGY	Movidius	Major IP Provider	GCT
Defense Contractor	Defense Contractor	Major SSD Vendor	Defense Contractor	Silicon Foundry	SCKIPIO	炬力 Actions	LEADCORE 联芯科技

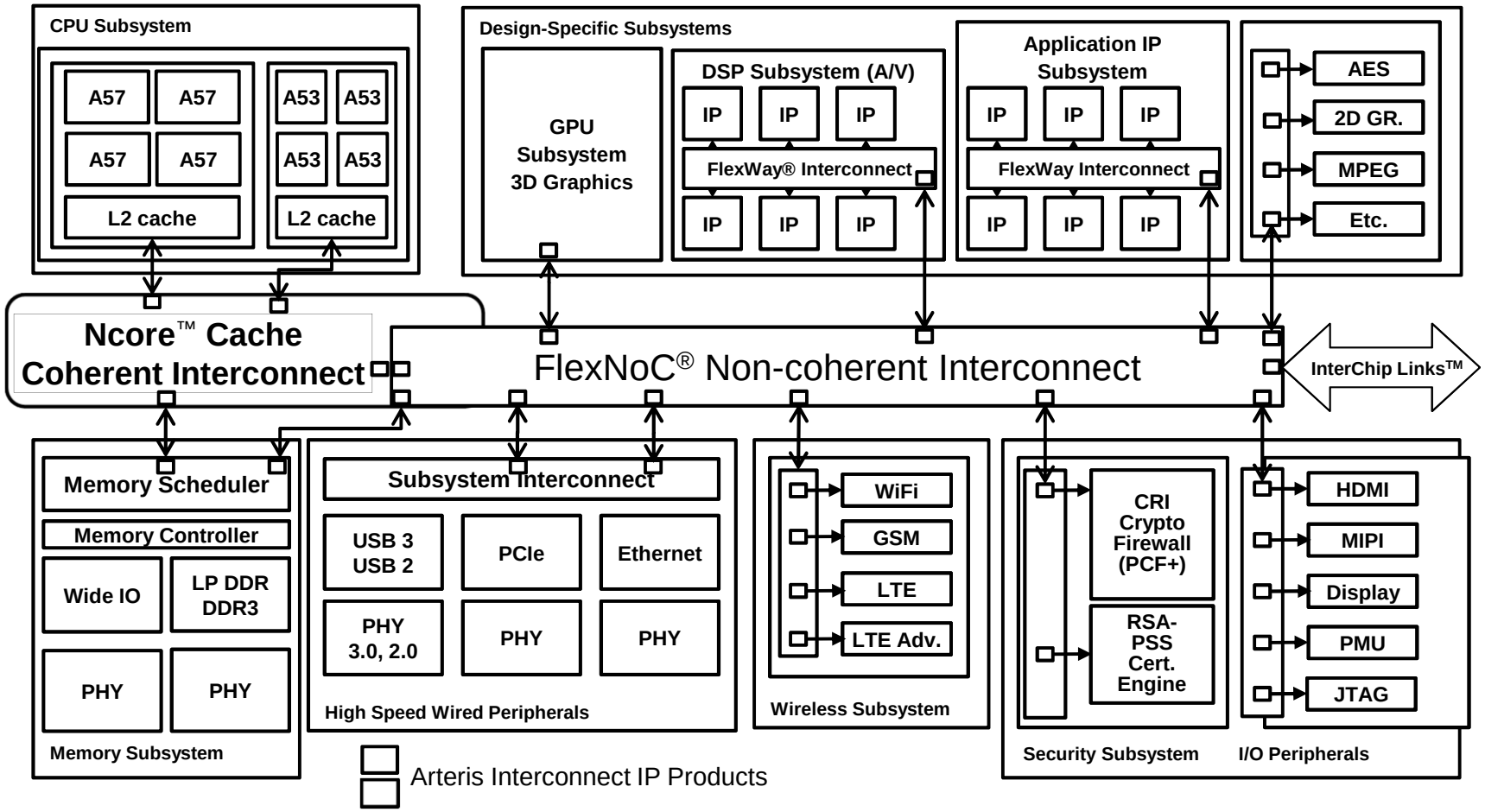






Copyright © 2016 Arteris

Arteris Interconnect IP: The easiest and most cost effective way to build differentiated SoCs

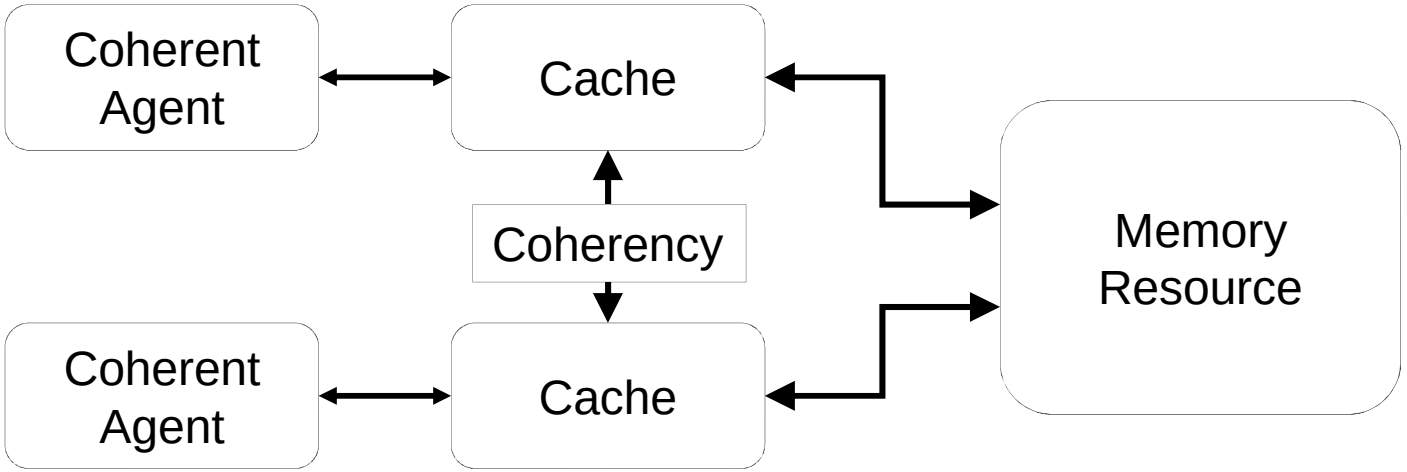


Copyright © 2016 Arteris, Strictly Confidential

Cache Coherency Primer

Copyright © 2016 Arteris

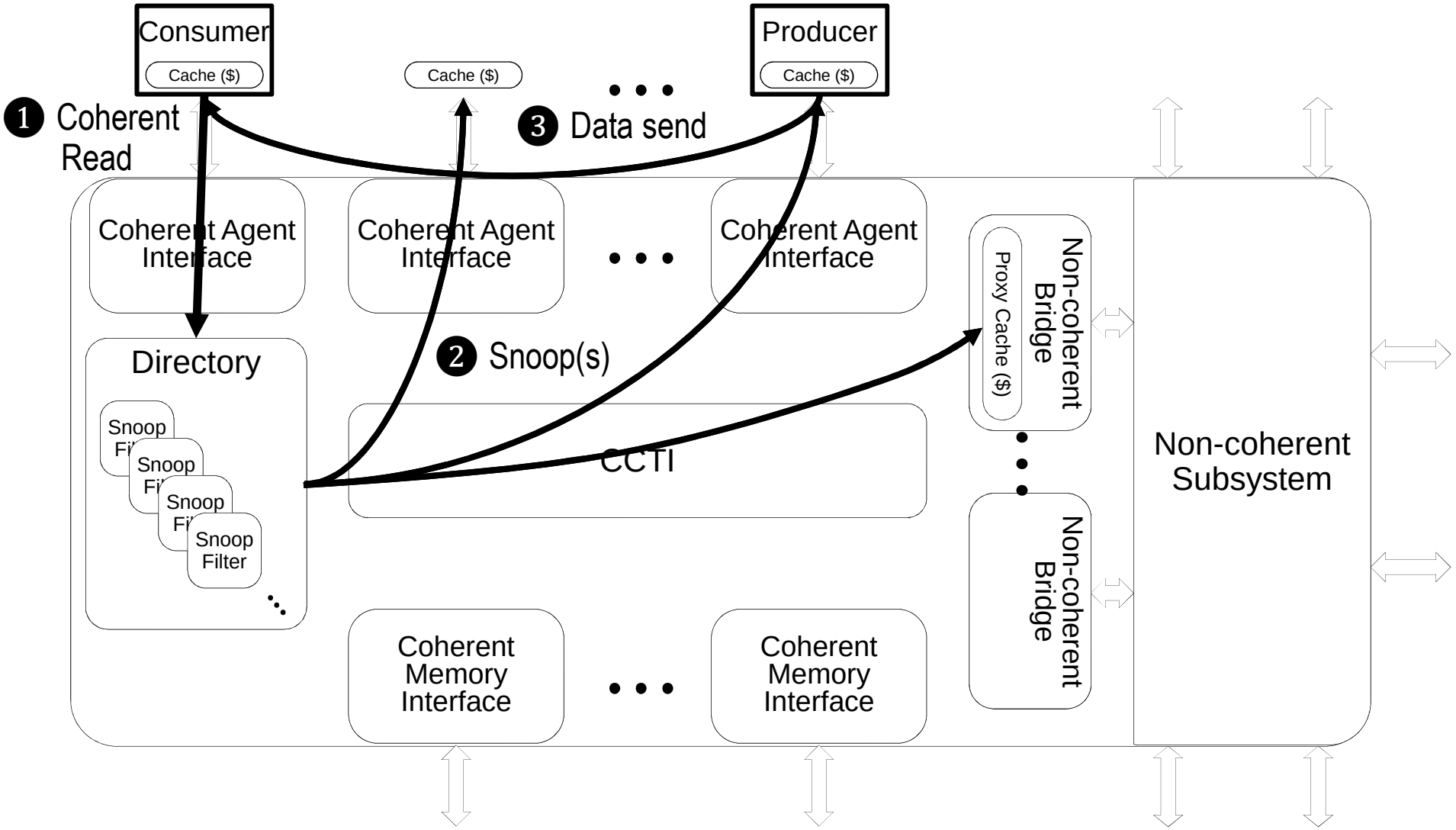
Cache coherency concept



- Source: Wikipedia

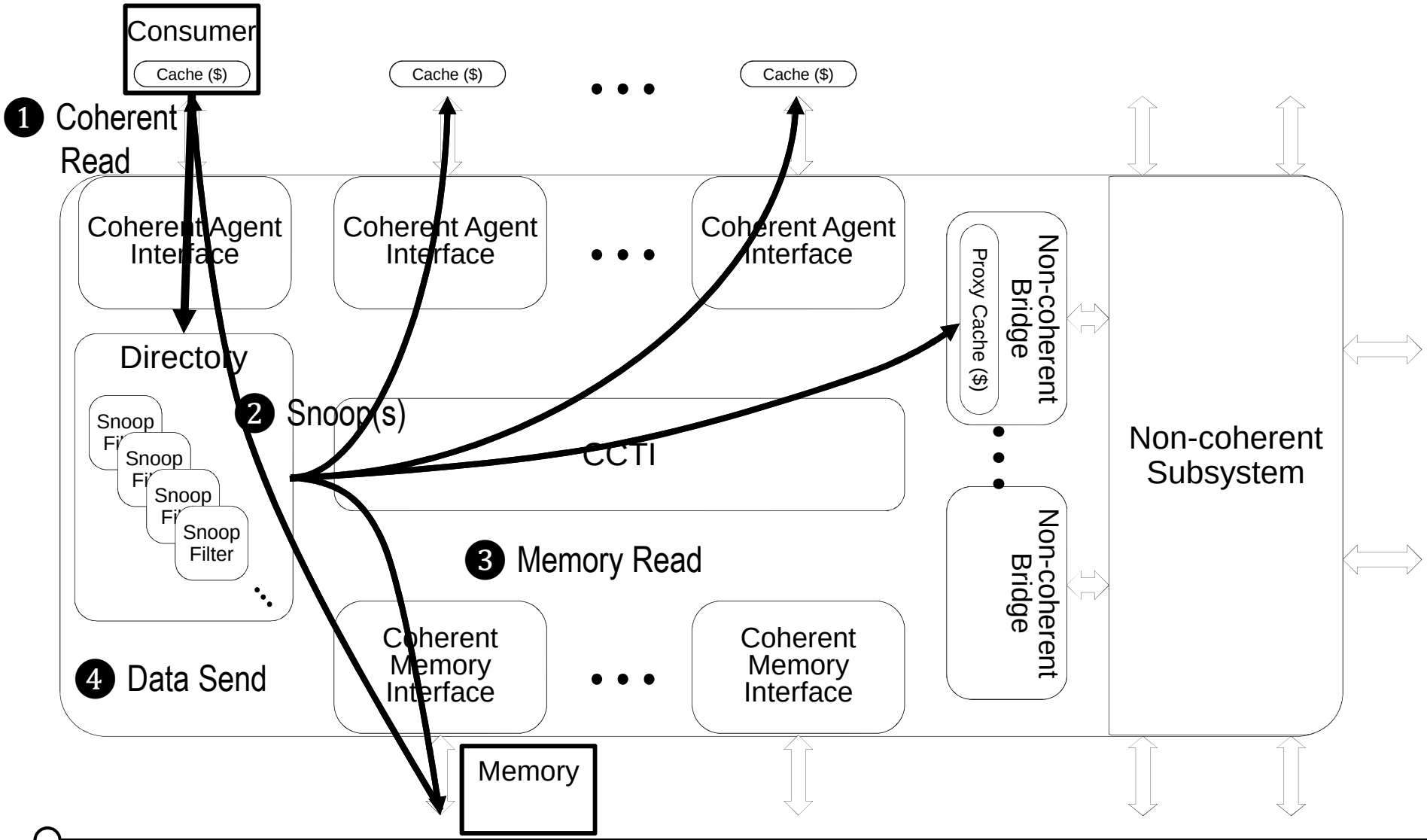
Copyright © 2016 Arteris

Coherent read example – cache hit



Copyright © 2016 Arteris

Coherent read example – cache miss

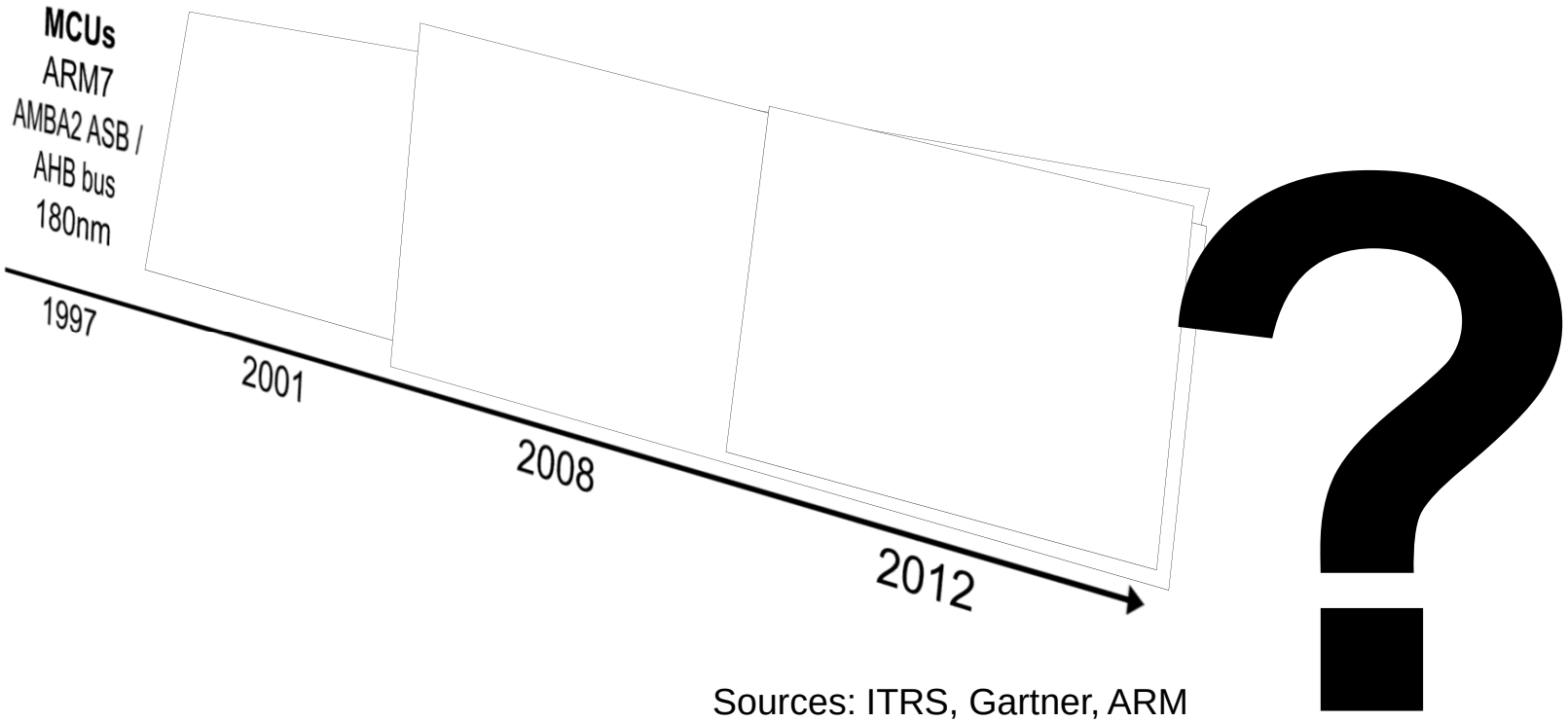


Copyright © 2016 Arteris

Who needs heterogeneous cache coherency?

Copyright © 2016 Arteris

How did we get here?

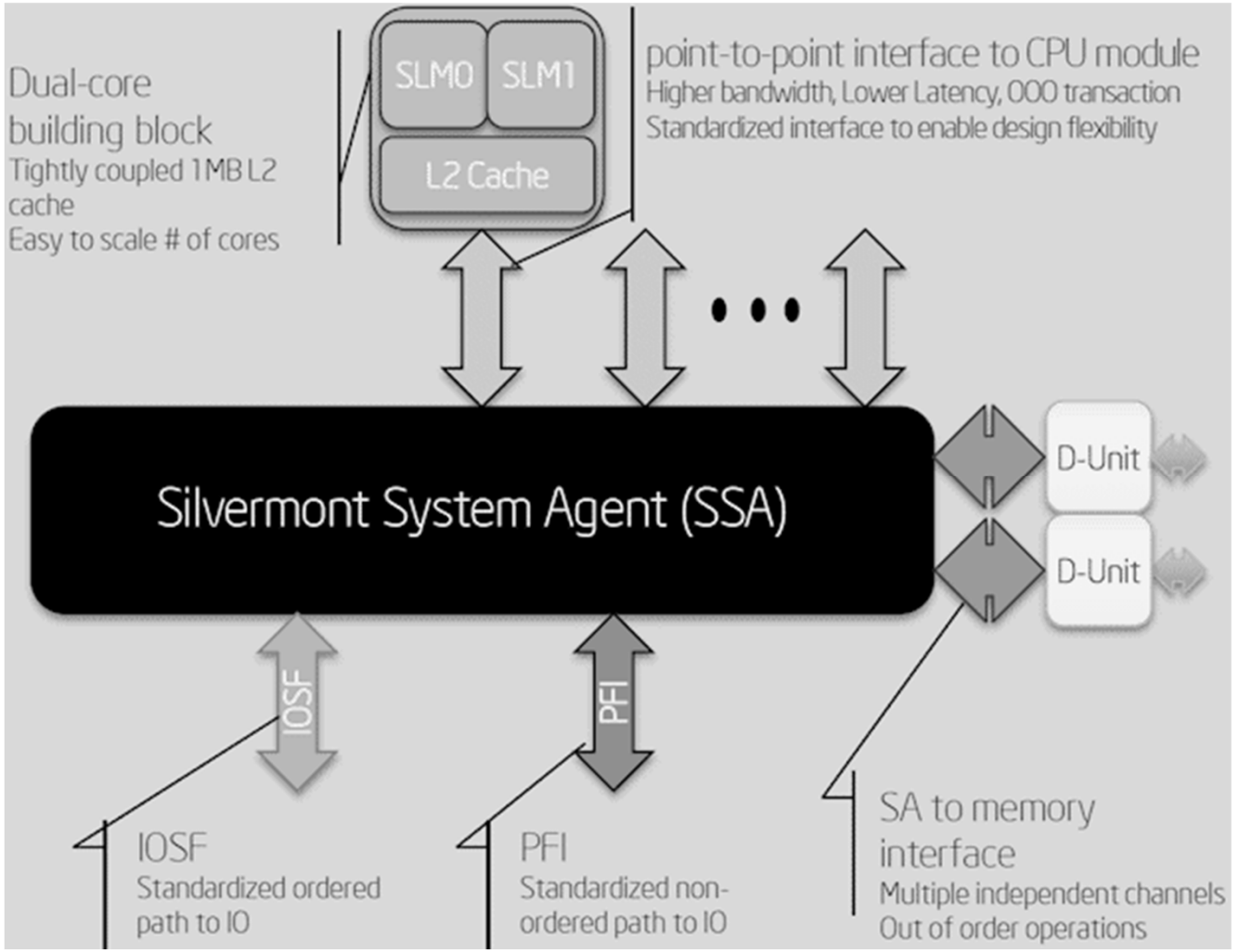


Sources: ITRS, Gartner, ARM

We can’t improve performance and economics by transistor scaling. No more free lunch!

Copyright © 2016 Arteris

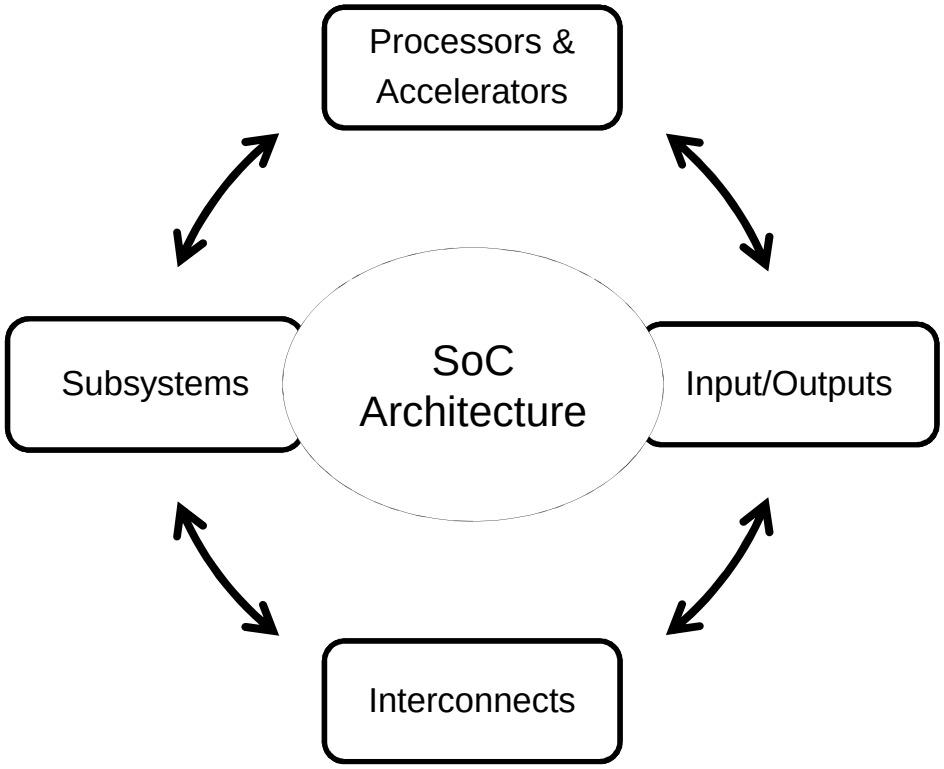
Cache coherency has been available



Copyright © 2015 Arteris

Keep this in mind…

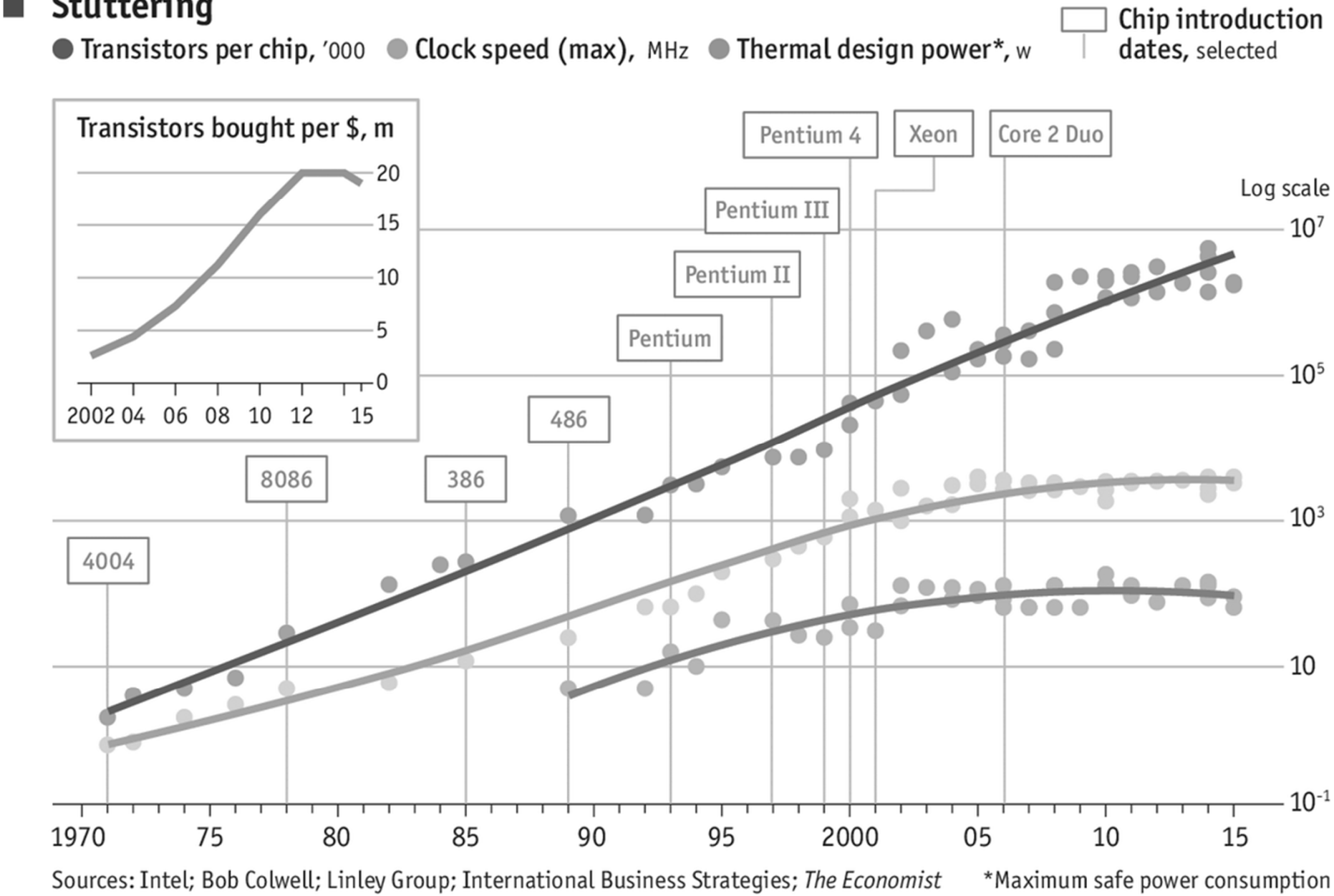
Software Simplicity



Moore’s Law

Copyright © 2016 Arteris

Stuttering

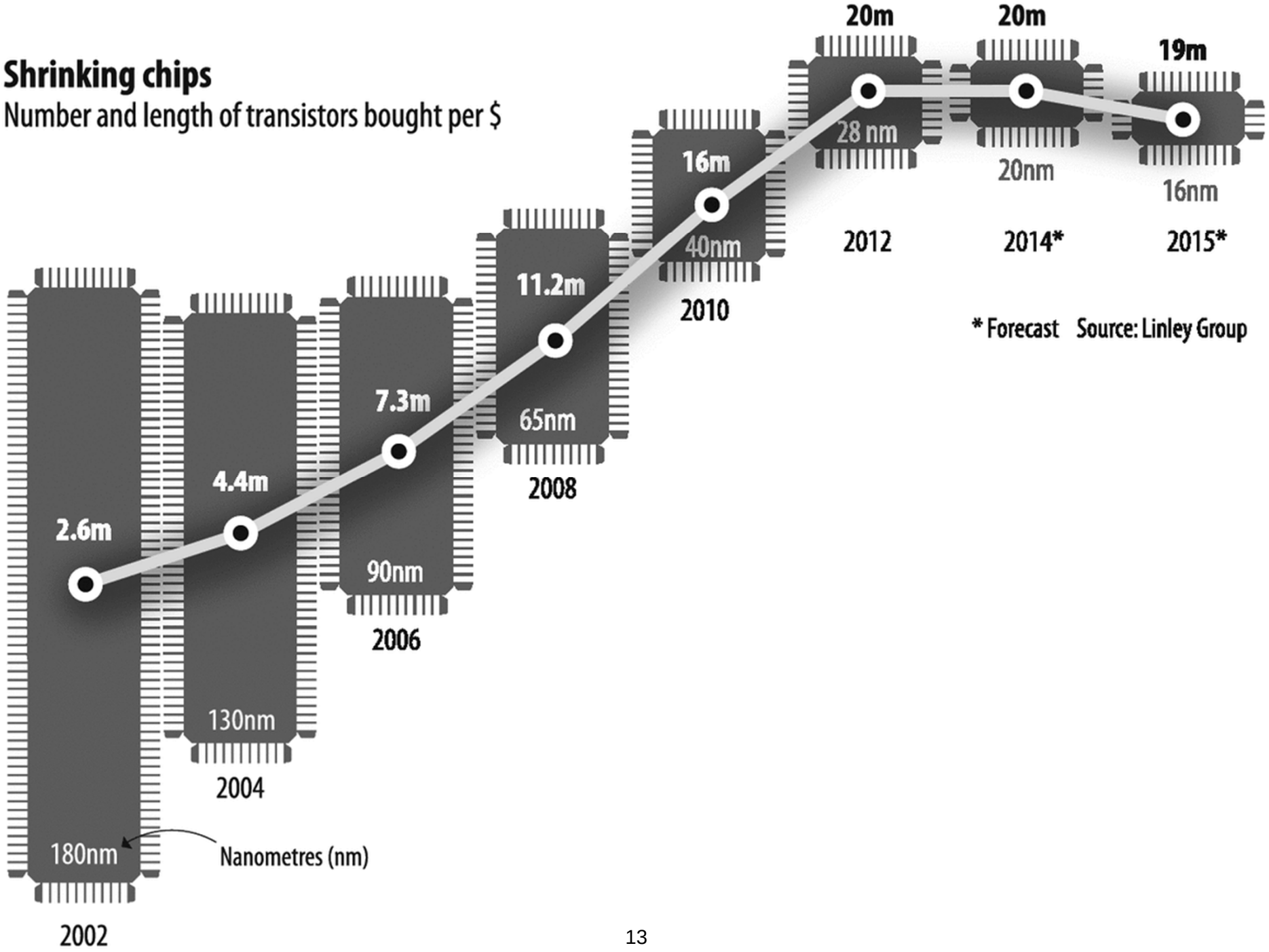


Sources: Intel; Bob Colwell; Linley Group; International Business Strategies; *The Economist*

*Maximum safe power consumption

Economist.com

Copyright © 2016 Arteris


Shrinking chips
Number and length of transistors bought per $
2.6m
180nm
2002
4.4m
130nm
2004
7.3m
90nm
2006
11.2m
65nm
2008
16m
40nm
2010
20m
28 nm
2012
20m
20nm
2014*
19m
16nm
2015*
* Forecast
Source: Linley Group
Nanometres (nm)

We need parallel computing to squeeze the most out of a process node

System	• HW + SW efficiency • Acceleration / CPU offload • Bandwidth & latency efficiency
Software	• Simplicity: Single view of memory • Reuse existing SW • “Software yield”: Useful work per LOC
Hardware	• Optimize for dissimilar HW – data sharing • Efficiency: Useful work per area (or mAh)

But we need to address **systemic** complexity in addition to scale complexity

Copyright © 2016 Arteris

Arteris™ The Network-on-Chip Company

Heterogeneous cache coherency enables simpler parallel computing

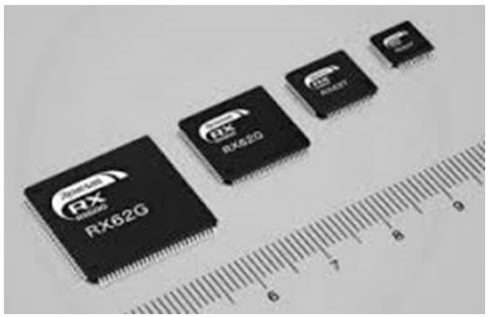
Multiple Coherence Models	Dissimilar Caching Agents	Non-coherent IP as coherent peers
• IP from different vendors / teams • Logical – coherence protocols / models • Generic system coherence model using lightweight messaging layer	• Physical – cache organization, transaction table sizes • Multiple configurable **snoop filters** save area	• Data sharing between non-coherent agents and coherent agents • Data sharing between non-coherent agents • Multiple configurable **proxy caches** minimize communication through DRAM

Copyright © 2016 Arteris

Arteris™
THE NETWORK-ON-CHIP COMPANY

Where is heterogeneous cache coherency relevant?

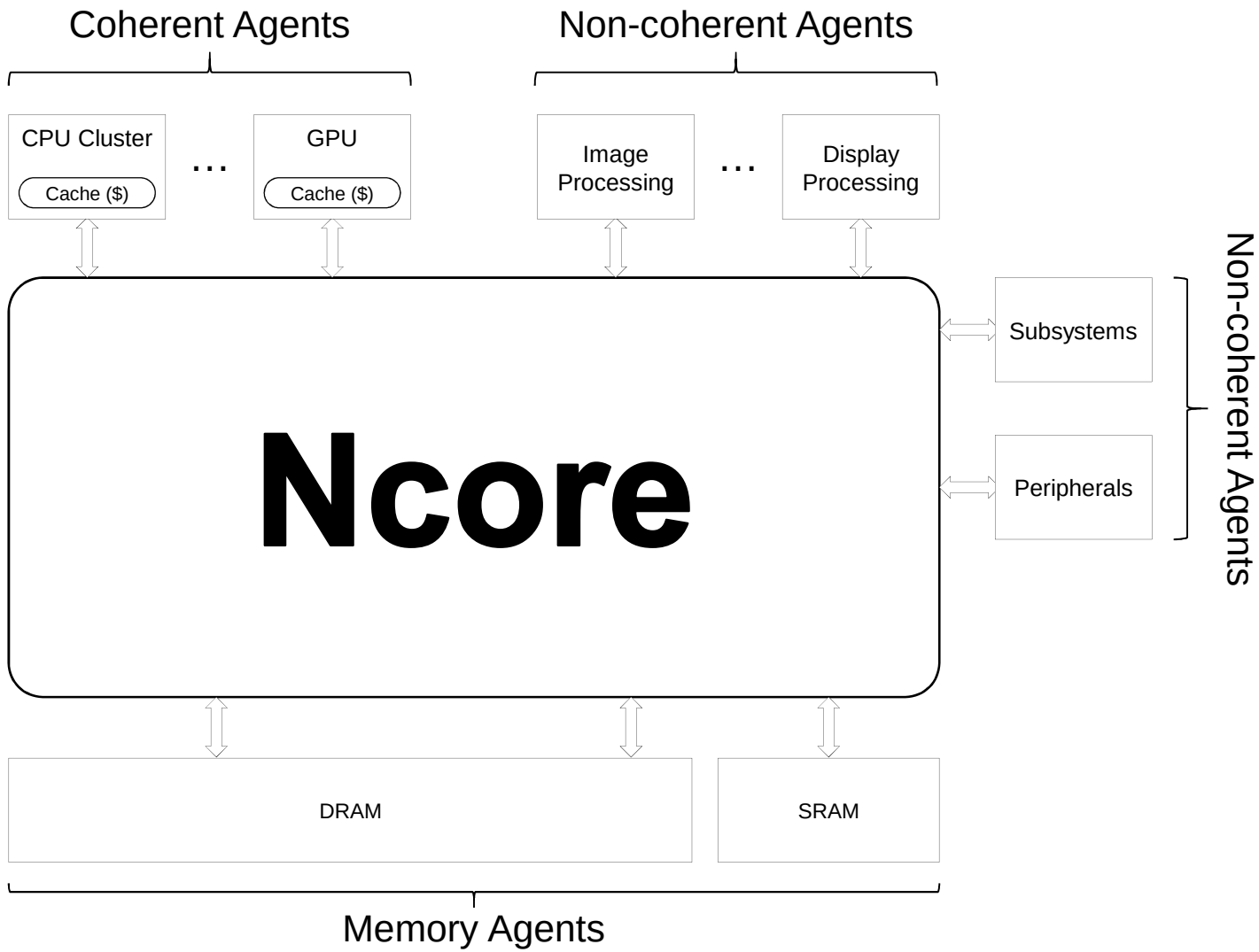
- Next generation automotive (ADAS)
- Mobility applications processors
- Virtual reality
- IoT
- Servers and others



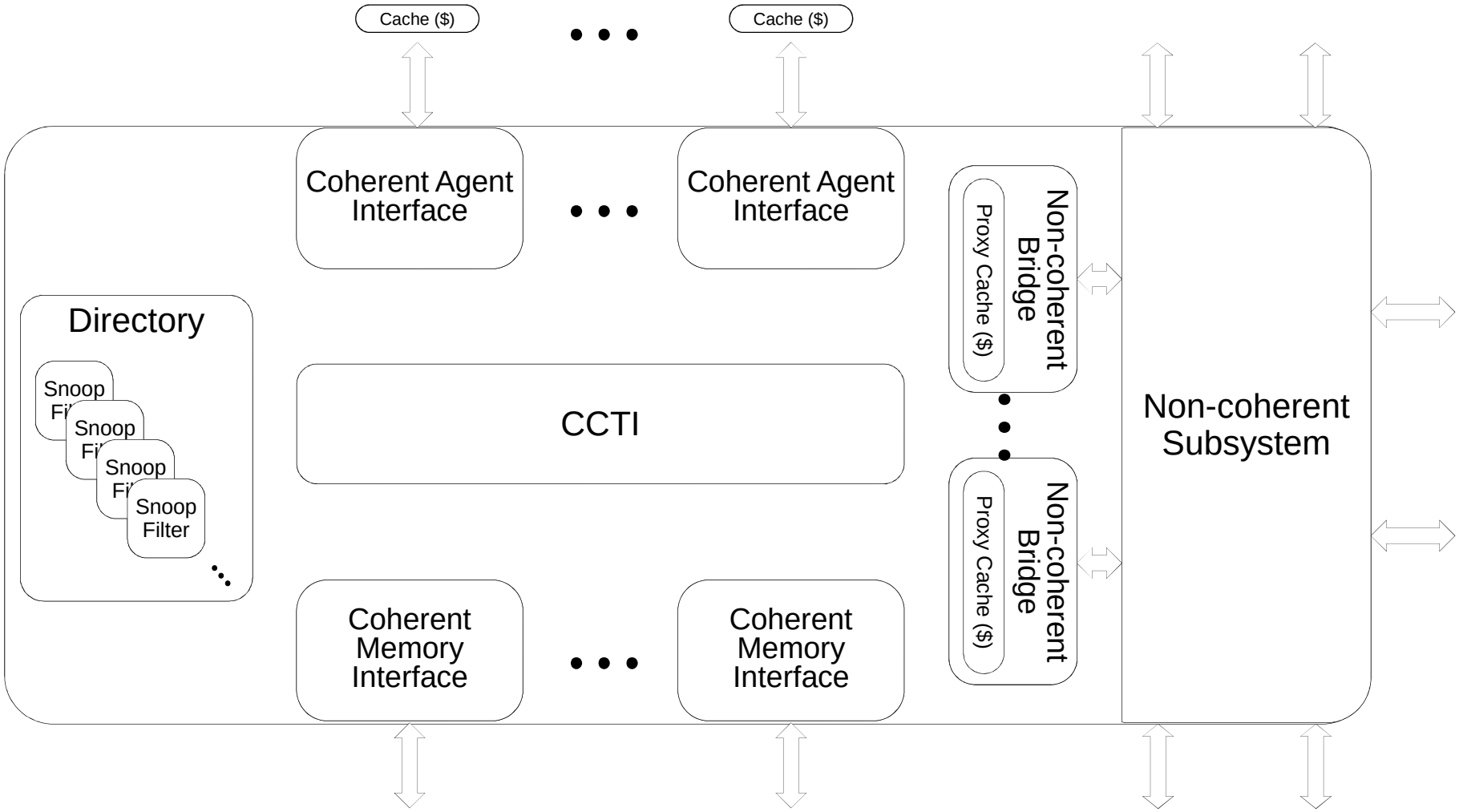
Heterogeneous Cache Coherency Implementation

Copyright © 2016 Arteris

Ncore Cache Coherent Interconnect IP



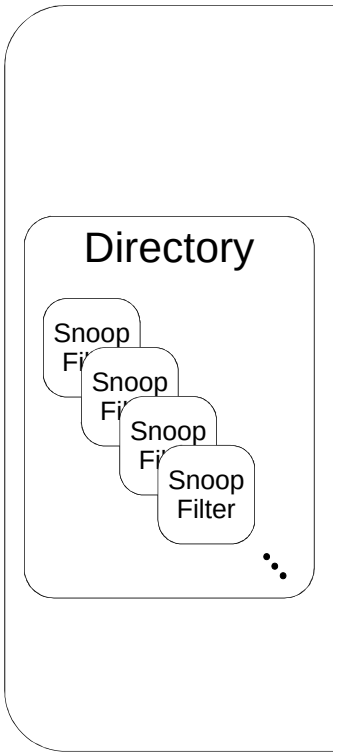
Ncore interconnect architecture



Copyright © 2016 Arteris

True heterogeneous coherency

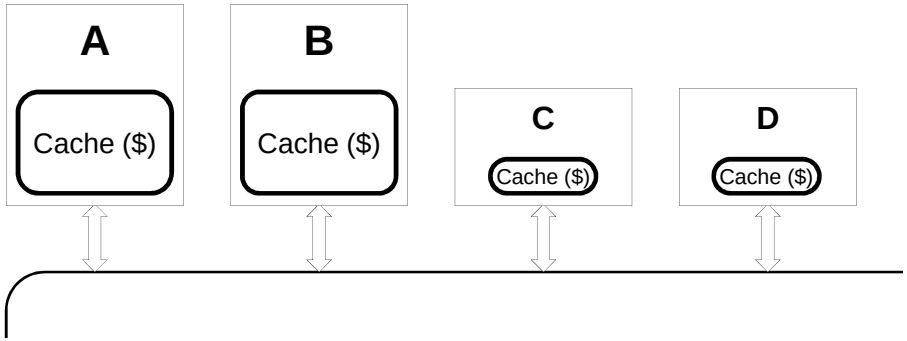
With multiple configurable snoop filters



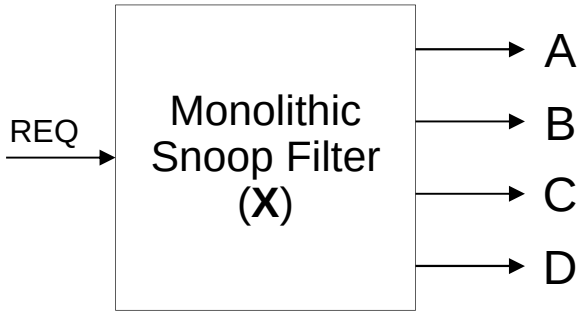
- Cache coherent agents can have very different behaviors
  - Cache organization
  - Coherency models
  - Workloads
- Associating caching agents that share common properties with individual snoop filters can consume **less die area** than a monolithic snoop filter

Copyright © 2016 Arteris

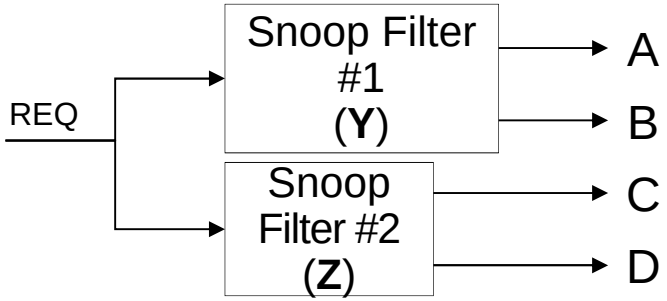
Multiple snoop filters are more area-efficient than one



Traditional Approach



Ncore Approach



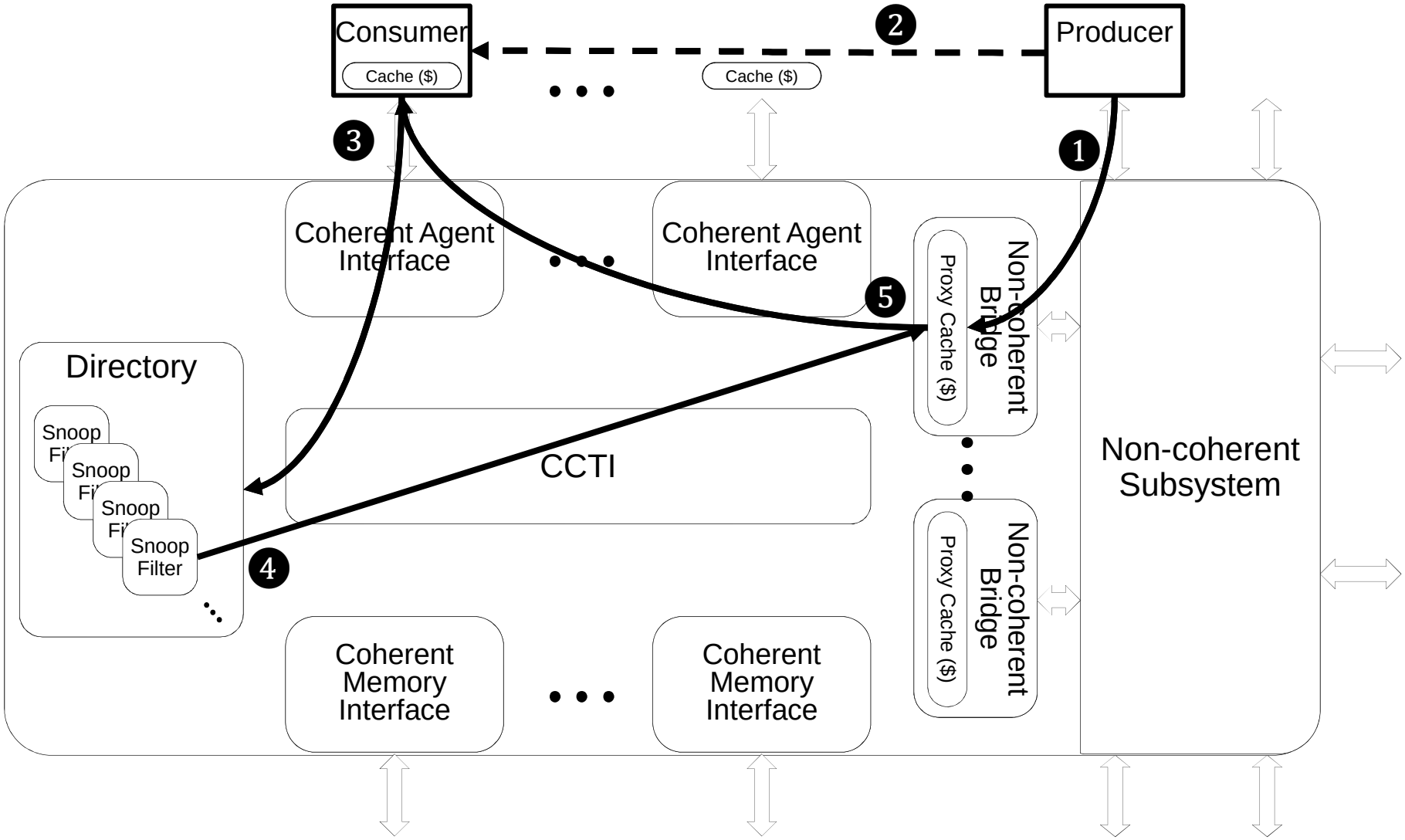
Multiple snoop filters are smaller: *area(Y+Z) < area (X)*

Copyright © 2016 Arteris

Higher performance with non-coherent IP

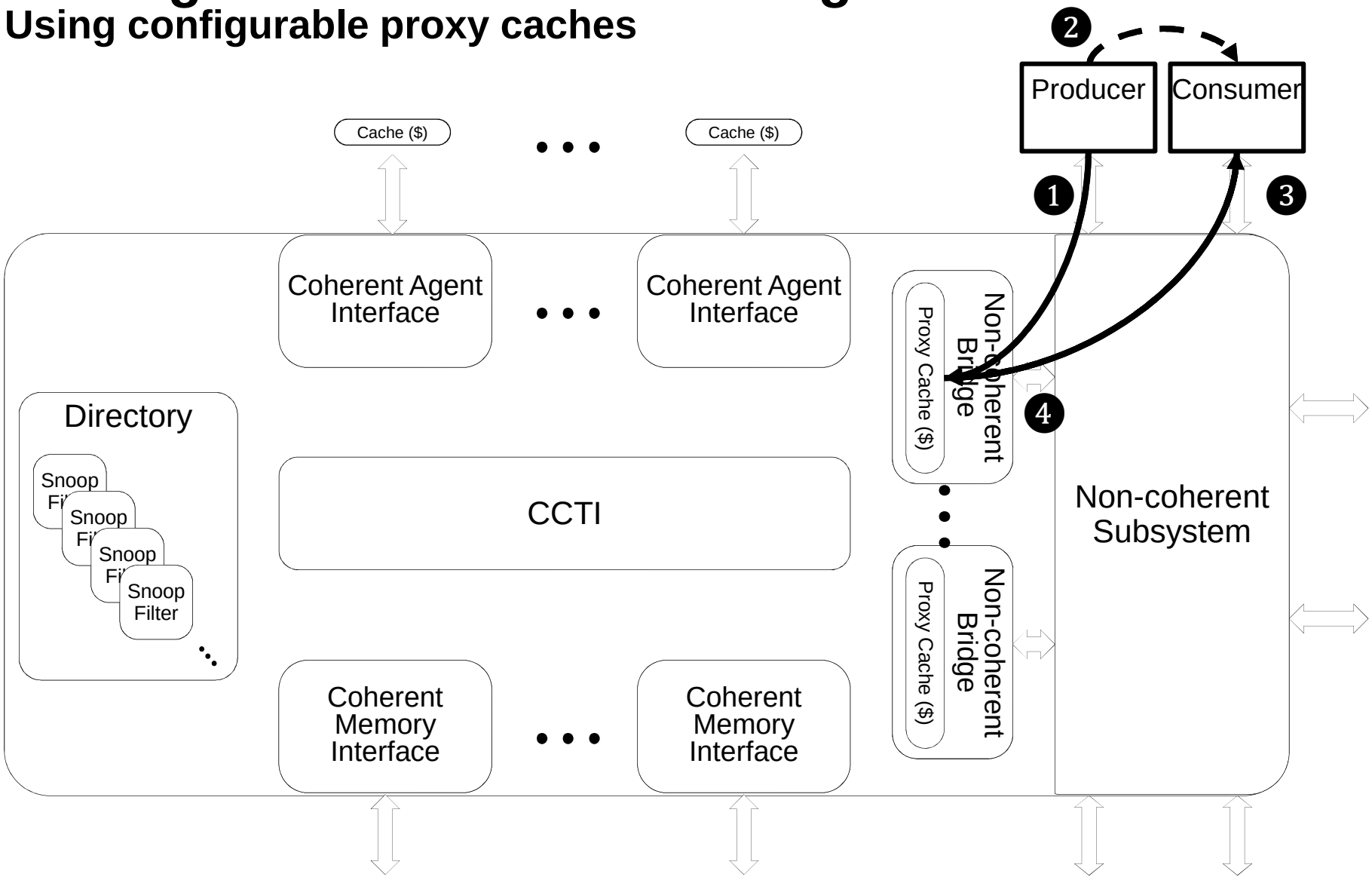
Sharing between non-coherent & coherent agents

Using configurable proxy caches

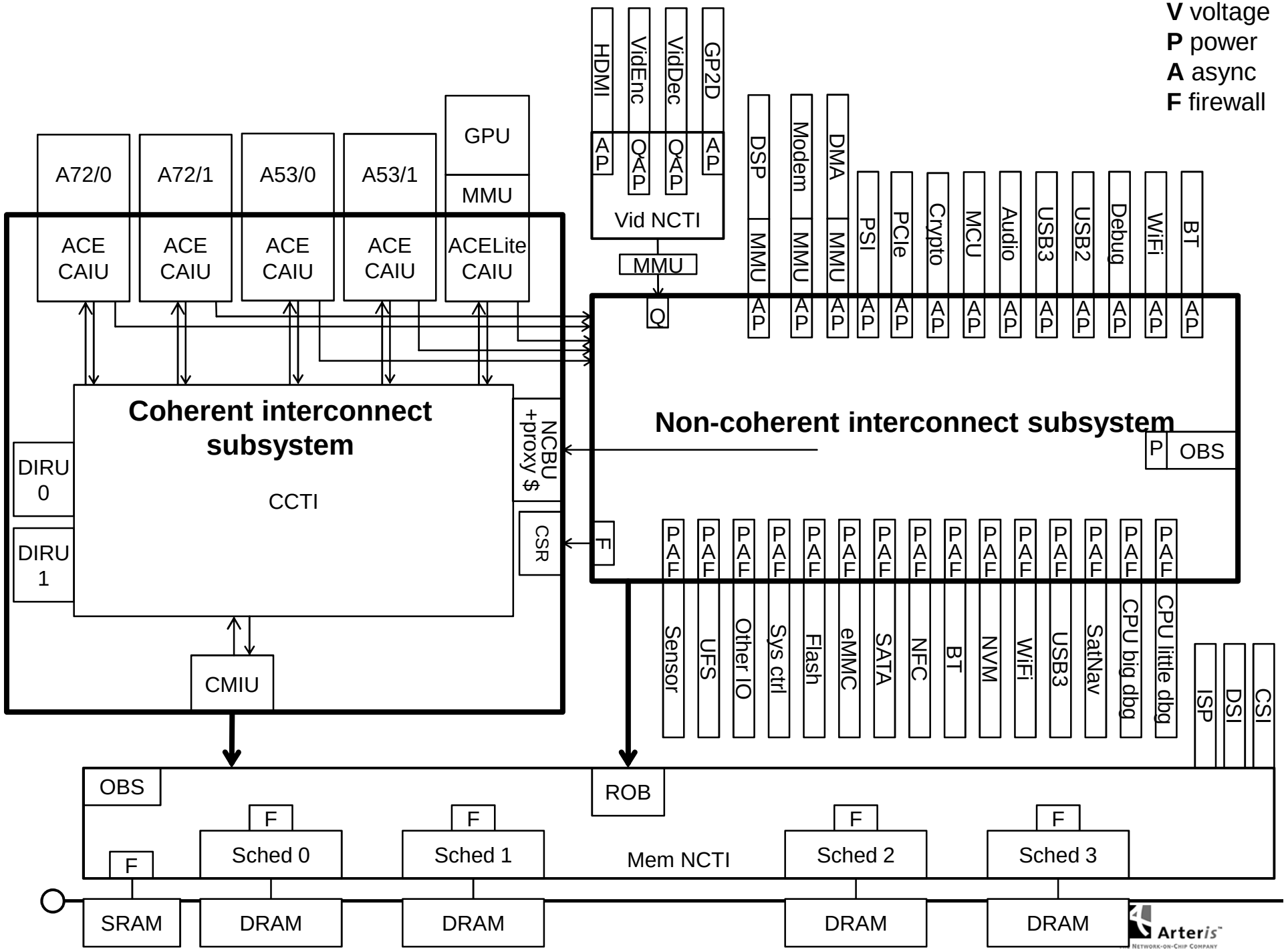


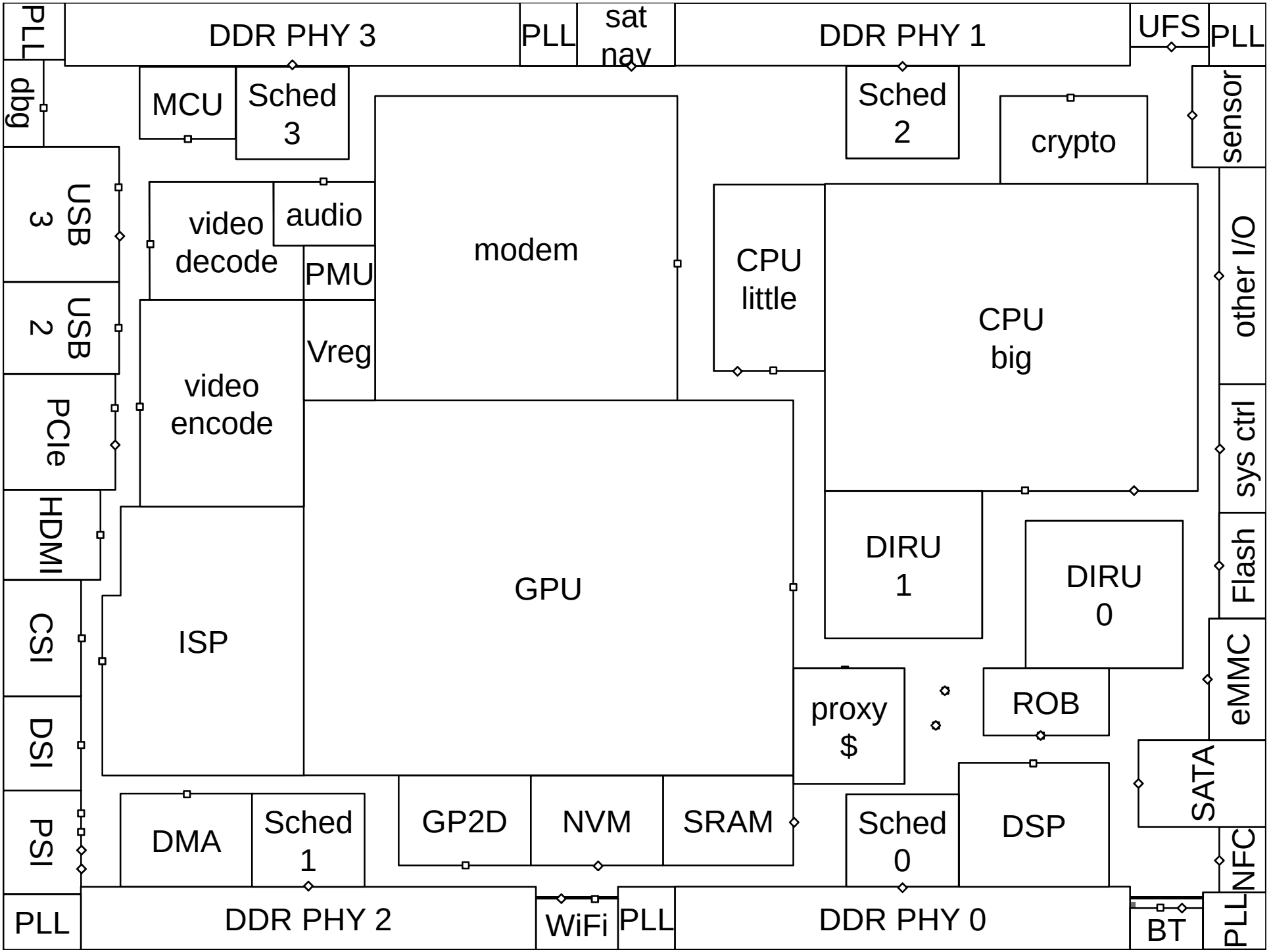
Copyright © 2016 Arteris

Sharing between non-coherent agents
Using configurable proxy caches



Copyright © 2016 Arteris


V voltage
P power
A async
F firewall
HDMI
VidEnc
VidDec
GP2D
AP
Vid NCTI
MMU
GPU
A72/0
A72/1
A53/0
A53/1
ACE CAIU
ACELite CAIU
DSP
Modem
DMA
PSI
PCIe
Crypto
MCU
Audio
USB3
USB2
Debug
WiFi
BT
Coherent interconnect subsystem
CCTI
NCBU +proxy $
CSR
DIRU 0
DIRU 1
CMIU
Non-coherent interconnect subsystem
P
OBS
Sensor
UFS
Other IO
Sys ctrl
Flash
eMMC
SATA
NFC
BT
NVM
WiFi
USB3
SatNav
CPU big dbg
CPU little dbg
ISP
DSI
CSI
ROB
Sched 0
Sched 1
Sched 2
Sched 3
Mem NCTI
SRAM
DRAM
Arteris
Network-on-Chip Company


PLL
DDR PHY 3
PLL
sat nav
DDR PHY 1
UFS
PLL
dbg
MCU
Sched 3
Sched 2
crypto
sensor
USB 3
video decode
audio
modem
CPU little
CPU big
other I/O
USB 2
PMU
Vreg
PCIe
video encode
sys ctrl
HDMI
DIRU 1
DIRU 0
Flash
CSI
ISP
GPU
eMMC
DSI
proxy $
ROB
SATA
PSI
DMA
Sched 1
GP2D
NVM
SRAM
Sched 0
DSP
NFC
PLL
DDR PHY 2
WiFi
PLL
DDR PHY 0
BT
PLL

Heterogeneous cache coherency required to maintain system power/performance trends

Ncore™ Cache Coherent Interconnect IP is targeted at heterogeneous SoCs.

Benefits

- Scalability
- Configurability
- Area efficiency
- High performance
- Optimal power consumption

Major Technologies

- Multiple configurable **snoop filters**
- Multiple configurable **proxy caches**
- Modular **distributed architecture**

RESULT:

Custom-configured interconnect IP that meets exact system requirements

Copyright © 2016 Arteris

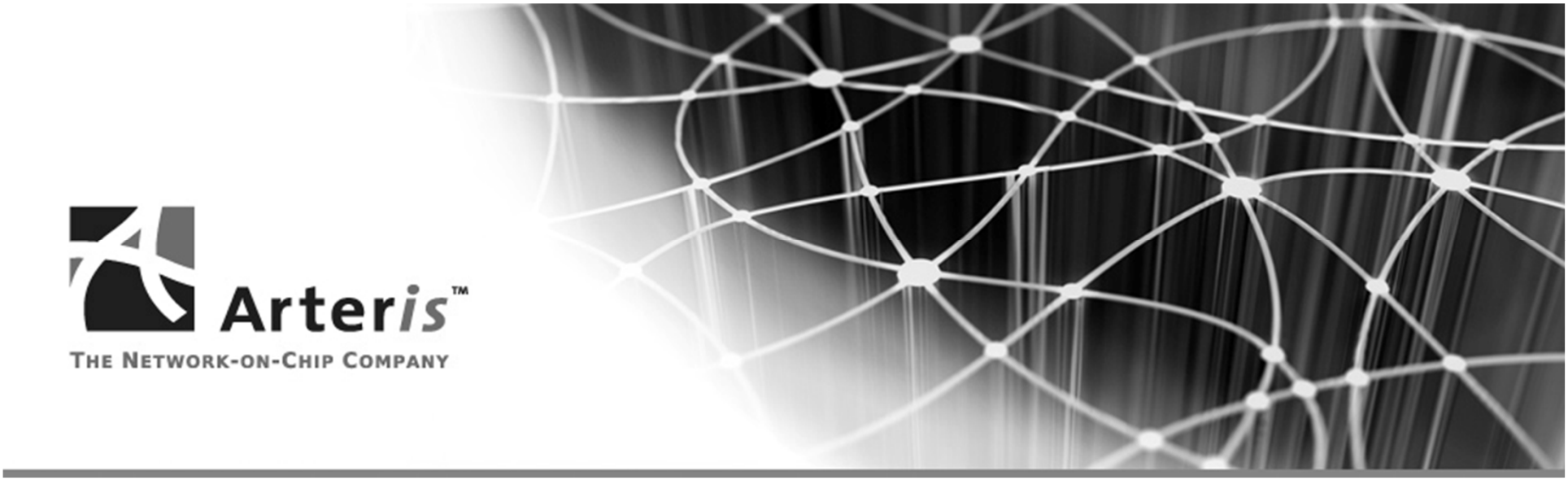
What’s next?

Copyright © 2016 Arteris

Upcoming challenges and trends

- Management of physical constraints
- Consolidation/de-consolidation of coherency protocols
- Hardware layer security
- Hardware layer resilience
- Lower power, lower power, lower power
- High performance
- Off-chip coherency
- 2.5/3D
- Constantly lowering the cost of coherency

Copyright © 2016 Arteris



Thank YOU

Copyright © 2016 Arteris