C-RAN
Moving Baseband to the Cloud

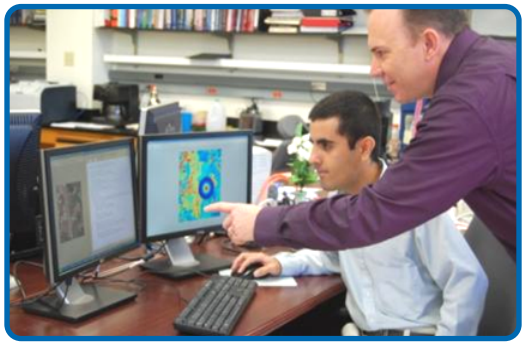
Matthew C. Valenti,
West Virginia University

Joint work with
Peter Rost, Nokia Networks
Aleksandra Checko, MTI Radiocomp

WVU Wireless Communications Research Lab

Faculty

- Matthew C. Valenti, Founder and Director
- Brian Woerner, Professor and Dept. Chair.
- Natalia Schmid, Professor
- Daryl Reynolds, Associate Professor
- Vinod Kulathumani, Associate Professor



Graduate Students

- Salvatore Talarico (interference modeling).
- Terry Ferrett (physical-layer network coding).
- Mohammad Fanaei (distributed detection).
- Veeru Talreja (biometrics in the cloud).

Research Topics

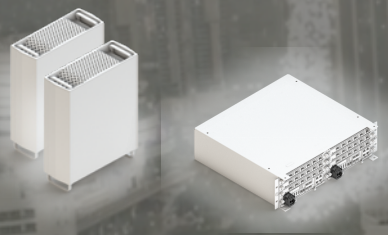
- Wireless and cellular networks.
- Cooperative communications.
- Coding and modulation.
- Spread spectrum systems.
  - Optimization of frequency hop systems.
- Wireless sensor networks
  - Distributed detection
- Cloud computing.
  - Cloud radio access networks
- Biometric and human activity detection.

Tools and Facilities

- 400-core cluster computer.
- Open source simulation software: CML.
- Web portal for researchers: WebCML.
- USRP software radio testbed.

Nokia AirScale Radio Access

AirScale Base Station



Cloud RAN

AirScale Wi-Fi

Common software supports evolution towards unified architecture and 5G

Business agility

Enables new services and business opportunities & agility to run shorter software and innovation cycles

High scalability

Enabler for efficient resource scalability. Meets the requirements of next generation heterogeneous access

Evolutionary

Combining licensed & un-licensed spectrum for maximum utilization of radio assets

5G readiness

A foundation for the next leap of spectral efficiency, peak rates & latency optimizations

MTI Radiocomp

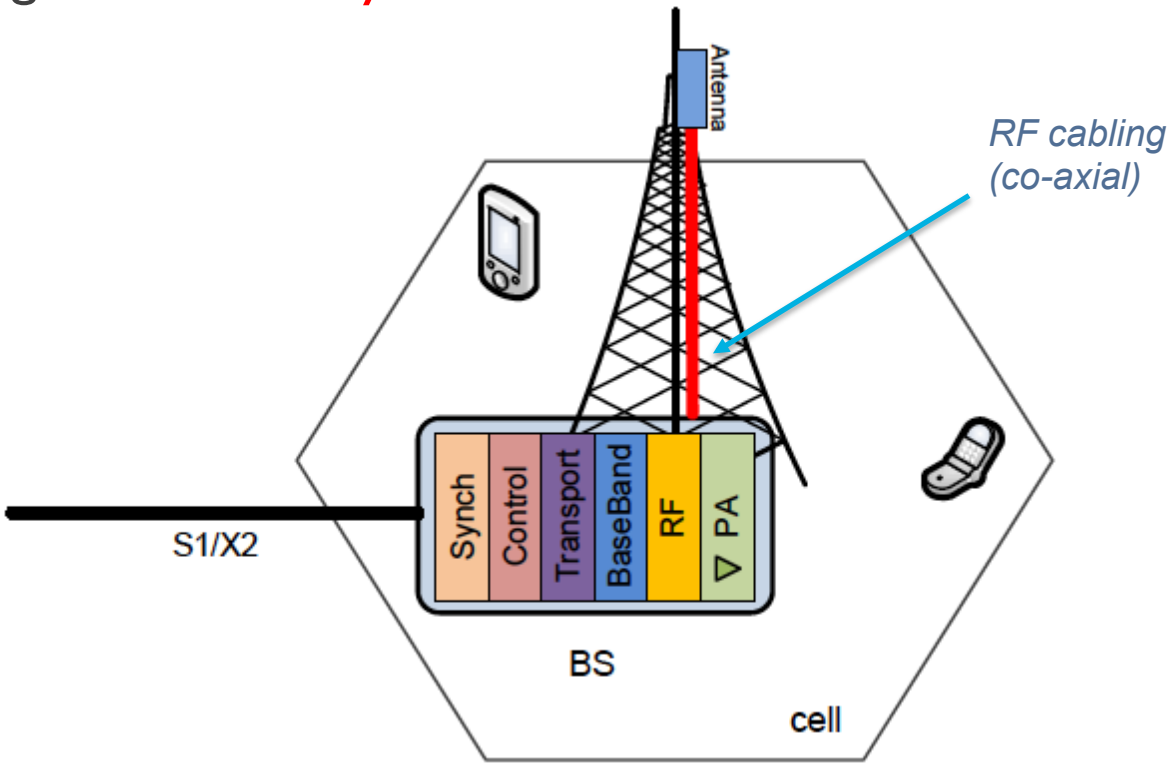
- Primary focus: research and development of Remote Radio Heads (RRHs) for LTE, WCDMA, WiMAX and towards 5G
- Global pioneer in the development and manufacturing of software-configurable and remotely tunable radio head systems
- One of the world's leading competence centers for CPRI, OBSAI, and JESD204A/B IPC controllers, core development, testing, and systems integration.
- Research on 5G networks, including C-RAN and fronthaul
- Danish site of MTI Mobile
- http://www.mti-mobile.com/

radiocomp
an MTI company

PART I: INTRO TO C-RAN

Traditional Base Stations

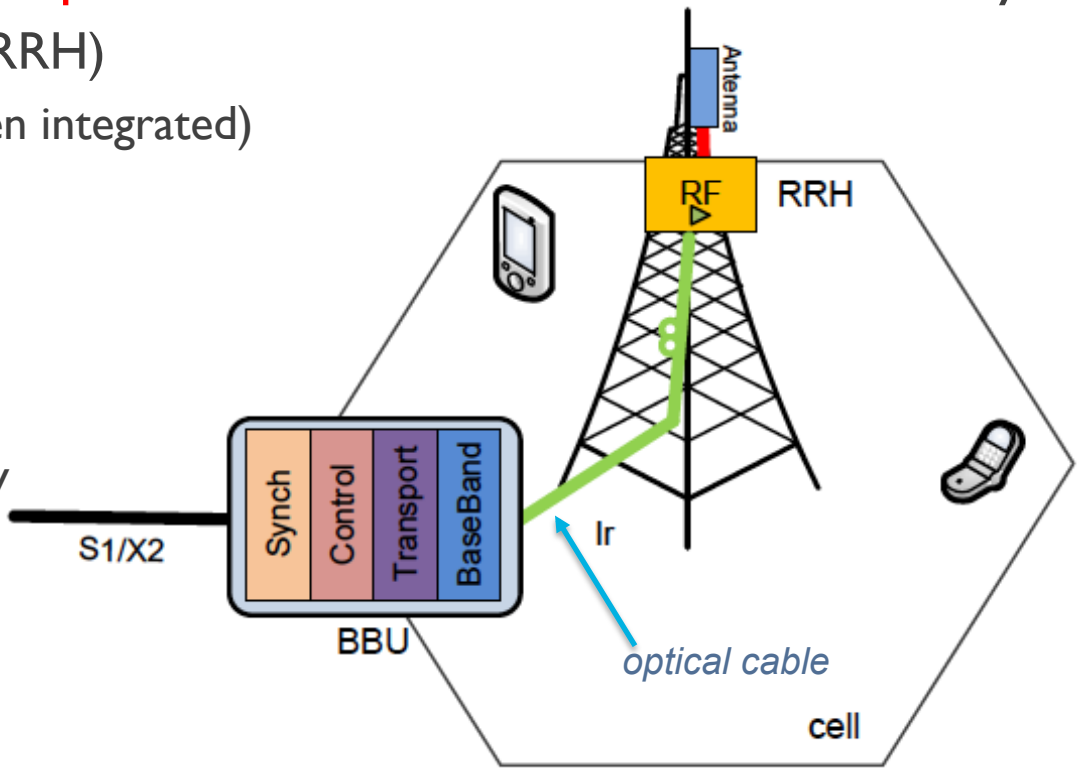
- Traditionally, all processing is done locally at the base station
  - Power Amplifier (PA)
  - RF
  - Baseband
  - Control Plane
  - Transport



[1]. A. Checko et al, "Cloud RAN for Mobile Networks – A Technology Overview," IEEE Comm. Surveys & Tutorials, First Quarter 2015.

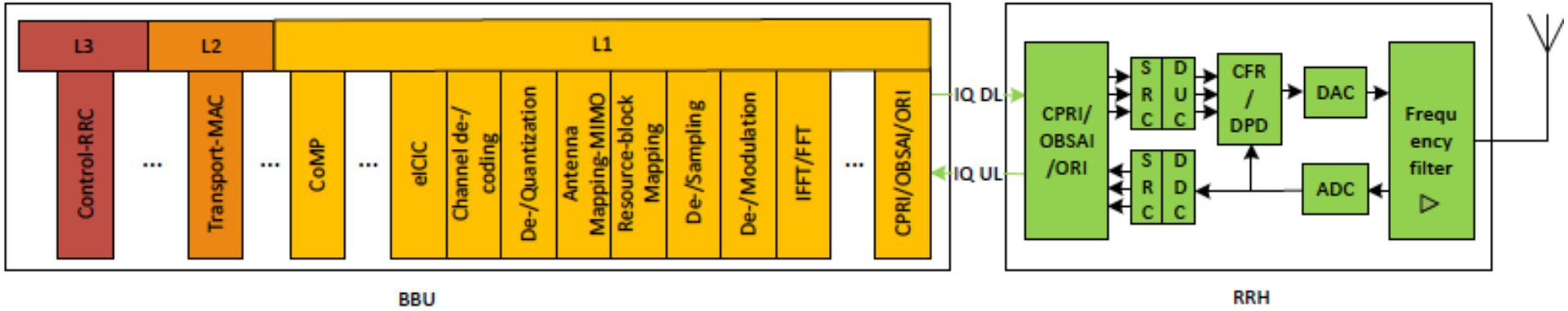
Separating the RRH from the BBU

- The modern trend is to separate the RF and baseband functionality
  - Remote Radio Head (RRH)
    - RF and antenna (often integrated)
    - Analog front-end
    - Power amplifier
  - Baseband Unit (BBU)
    - Baseband processing
    - Implemented digitally



[1]. A. Checko et al, "Cloud RAN for Mobile Networks – A Technology Overview," IEEE Comm. Surveys & Tutorials, 1Q 2015.

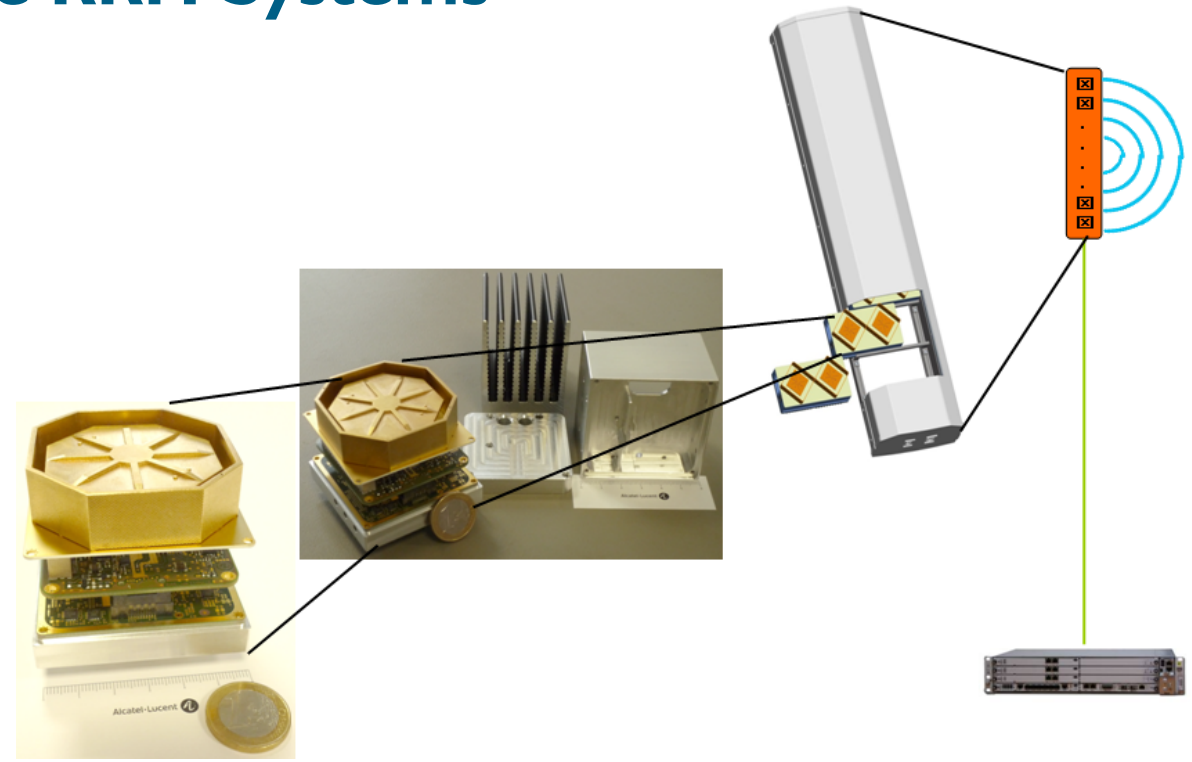
The RRH / BBU Functional Split



RRC	Radio Resource Control	SRC	Sampling Rate Conversion	DAC	Digital-to-Analog Converter
MAC	Media Access Control	DUC/DDC	Digital Up/Downconversion	ADC	Analog-to-Digital Converter
FFT	Fast Fourier Transform	CFR	Crest Factor Reduction	▷	Power Amplifier
		DPD	Digital Predistortion		

[1]. A. Checko et al, “Cloud RAN for Mobile Networks – A Technology Overview,” IEEE Comm. Surveys & Tutorials, 1Q 2015.

Example RRH Systems

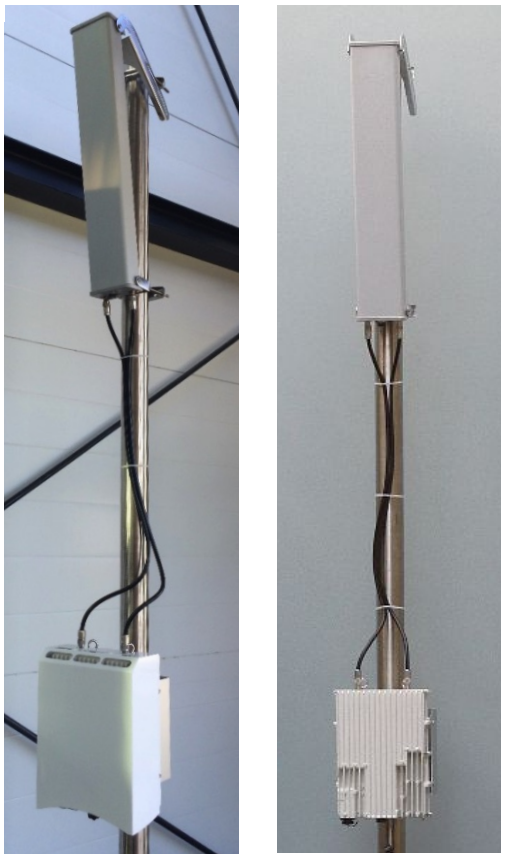


- Alcatel-Lucent (now Nokia) Light Radio

http://www.lgsinnovations.com/solutions/broadband-wireless-network-light-radio

Example RRH Systems

Telecom Italia Demo



MTI Radiocomp Demo

CPRI

- Common Public Radio Interface
- Features:
  - Constant bit-rate up to 12 Gbps
  - Transmission of I/Q samples
- Maximum distance between RRH and BBU
  - 40 km
  - 0.1 ms latency (speed of light in a fiber) and no repeater needed
- Other protocols for connecting RRH and BBU
  - Open Base Station Architecture Initiative (OBSAI)
  - Open Radio equipment Interface (ORI)

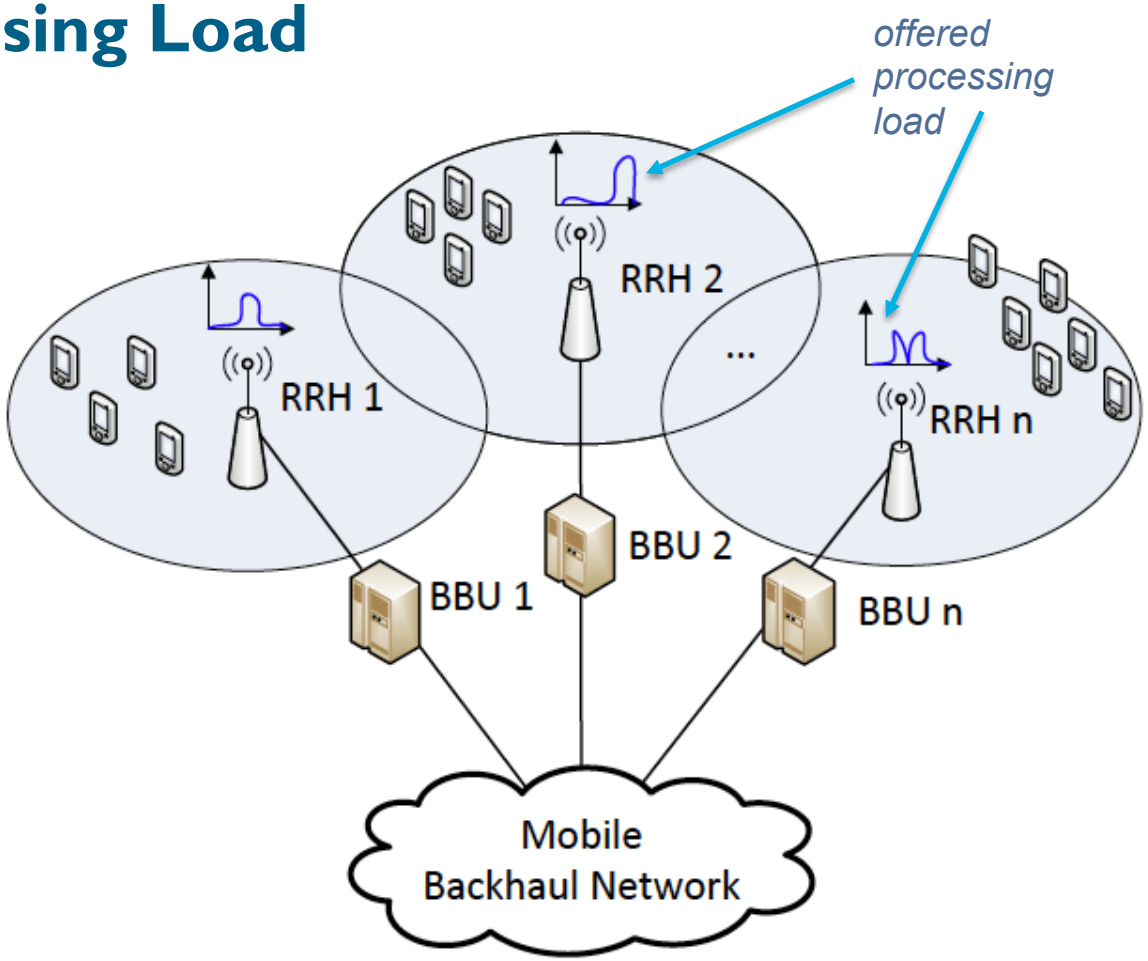
Required Rates for a CPRI-Based Fronthaul

Cell configuration	Bit rate
20 MHz LTE, 15+1 CPRI IQ Sample width, 10/8 line coding, 2x2 MIMO	2.5 Gbps
5x20 MHz LTE-A, 15 CPRI IQ Sample width, 2x2 MIMO, 3 sectors	13.8 Gbps
20 MHz LTE, 4x2 MIMO, 3 sectors	16.6 Gbps
TD-LTE, 3 sectors	30 Gbps
1.6 MHz TD-SCDMA, 8Tx/8Rx antennas, 4 times sampling rate	330 Mbps
TD-SCDMA S444, 3 sectors	6 Gbps
200 kHz GSM, 2Tx/2Rx antennas, 4x sampling rate	25.6 Mbps

[1]. A. Checko et al, “Cloud RAN for Mobile Networks – A Technology Overview,” IEEE Comm. Surveys & Tutorials, 1Q 2015.

Variability of Processing Load

- Each BBU has its own load
  - Depends on traffic in cell
  - As a result, the load on the different BBU's is highly variable

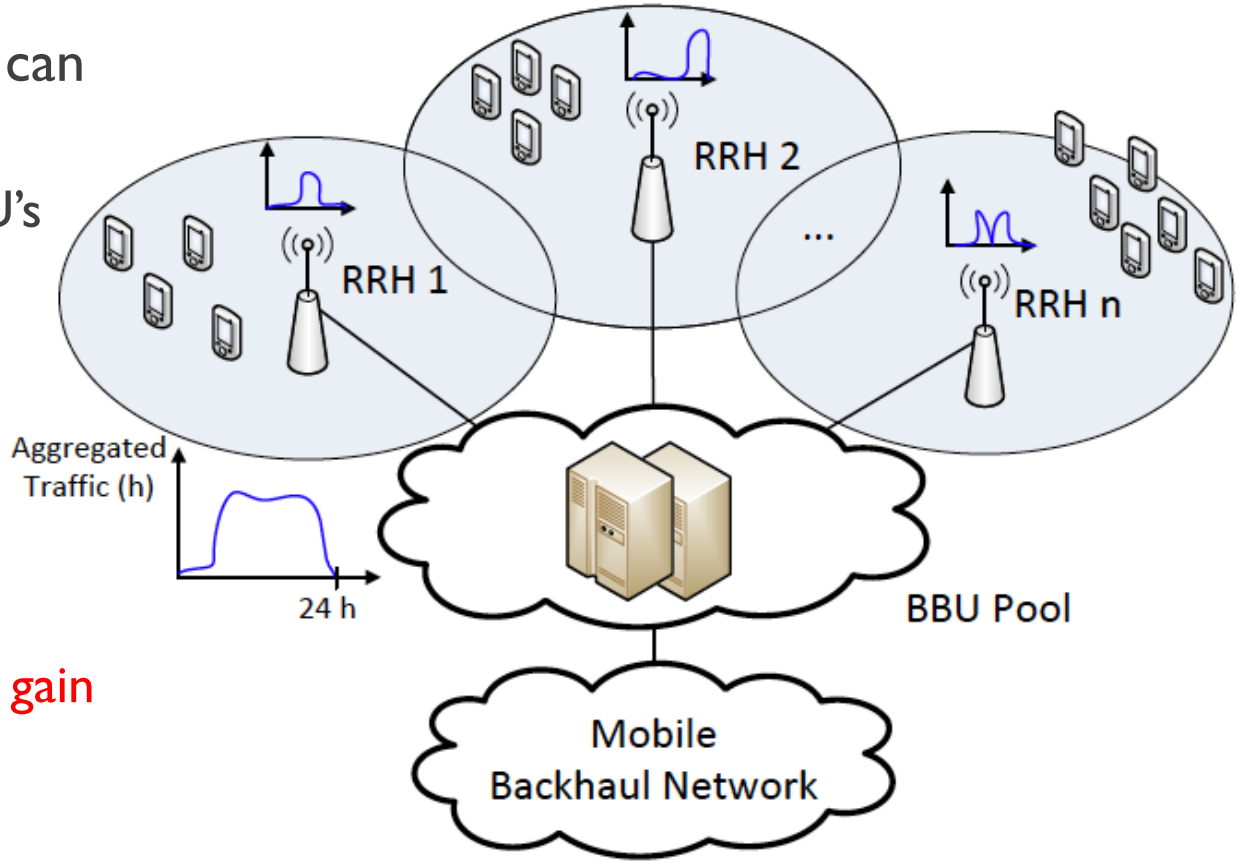


original

[1]. A. Checko et al, "Cloud RAN for Mobile Networks – A Technology Overview," IEEE Comm. Surveys & Tutorials, 1Q 2015.

The (Virtual) Baseband Hotel

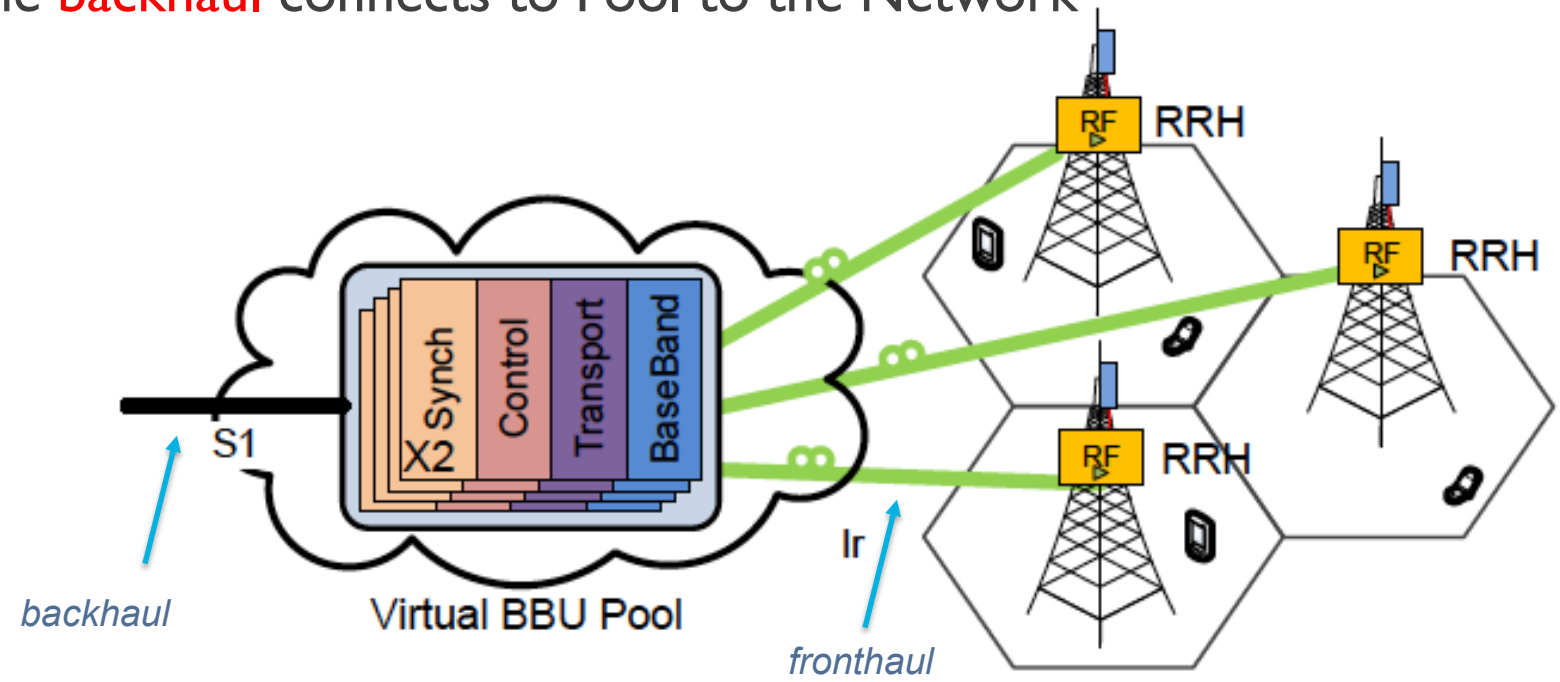
- Several Baseband Units can be co-located
  - Creates a pool of BBU's or BBU Hotel
  - BBU's could still be physically separated
  - However, combining BBU's into a shared resource provides a statistical multiplexing gain



[1]. A. Checko et al, "Cloud RAN for Mobile Networks – A Technology Overview," IEEE Comm. Surveys & Tutorials, 1Q 2015.

Fronthauls and Backhauls

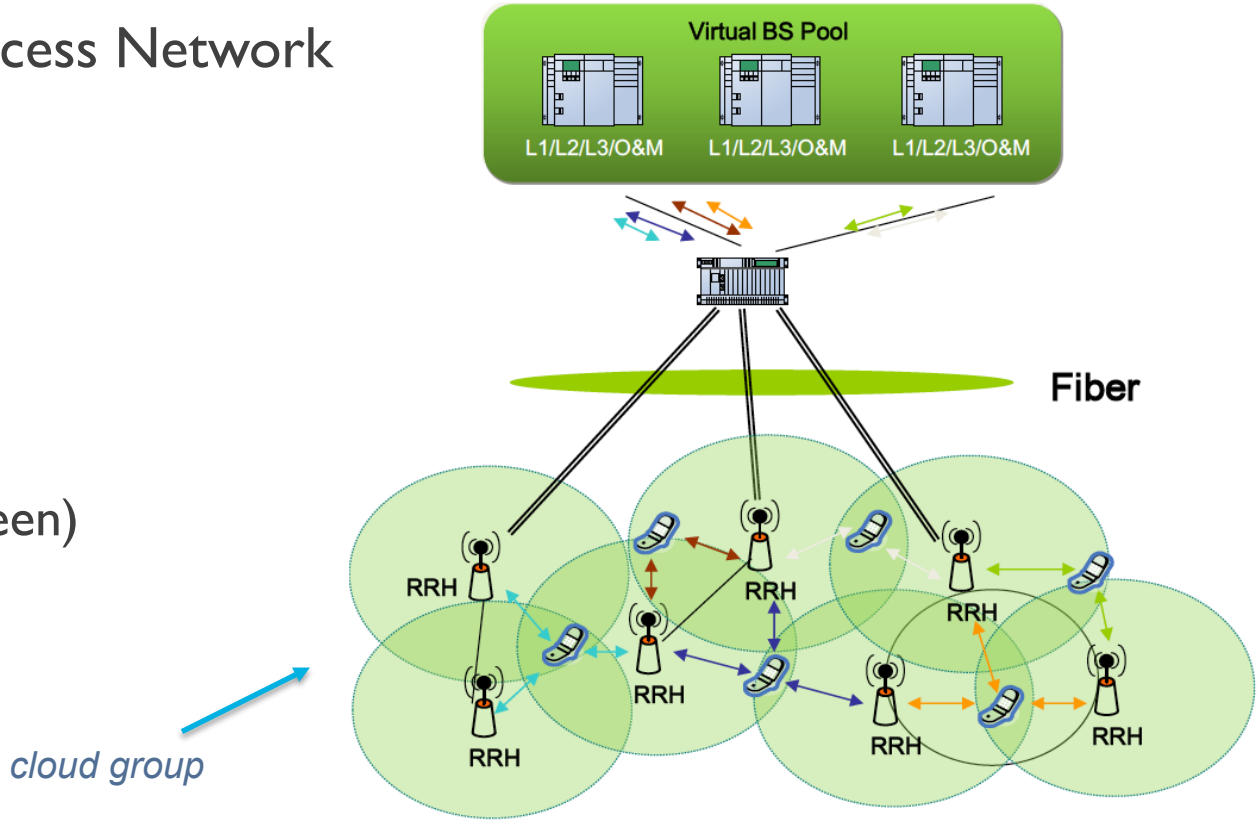
- The fronthaul connects the RRH's to the BBU Pool
- The backhaul connects to Pool to the Network



[1]. A. Checko et al, "Cloud RAN for Mobile Networks – A Technology Overview," IEEE Comm. Surveys & Tutorials, 1Q 2015.

C-RAN

- Aggressive use of virtual BBU pools is known as C-RAN
- RAN = Radio Access Network
- C = ?
  - Cloud
  - Central
  - Collaborative
  - Cooperative
  - Clean (i.e., Green)



[2]. China Mobile Research Institute, "C-RAN: The Road Towards Green RAN," White Paper, 2010.

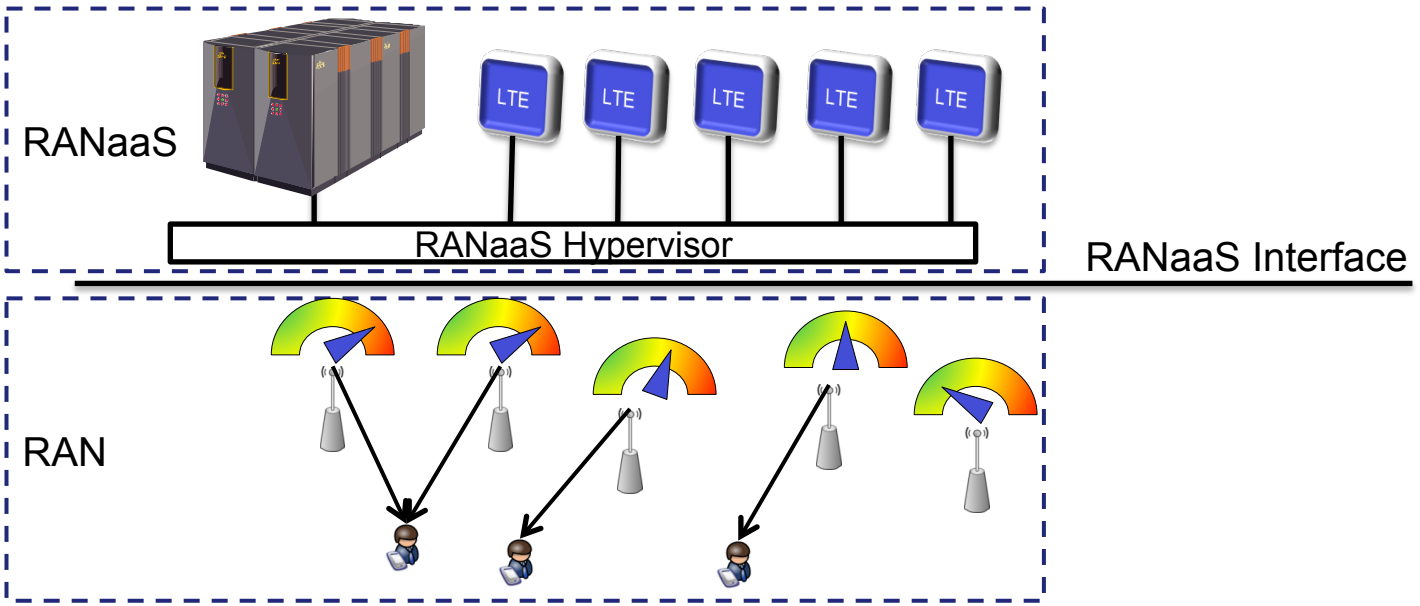
Benefits of C-RAN

- Centralized processing exploits statistical multiplexing gain
  - Processing resources provisioned for the average across the cloud group, not the peak requirement of any one cell
  - Adaptability to non-uniform and time-varying traffic
- Centralized management
  - Easier HW/SW upgrades (and they are decoupled)
  - Support for multiple standards
  - Facilitates self-organizing networks (SONs)
- Facilitates cooperative processing
  - COMP and eICIC
  - Easier handoffs
- Other applications and services can be integrated into the network edge

Perceived Benefits

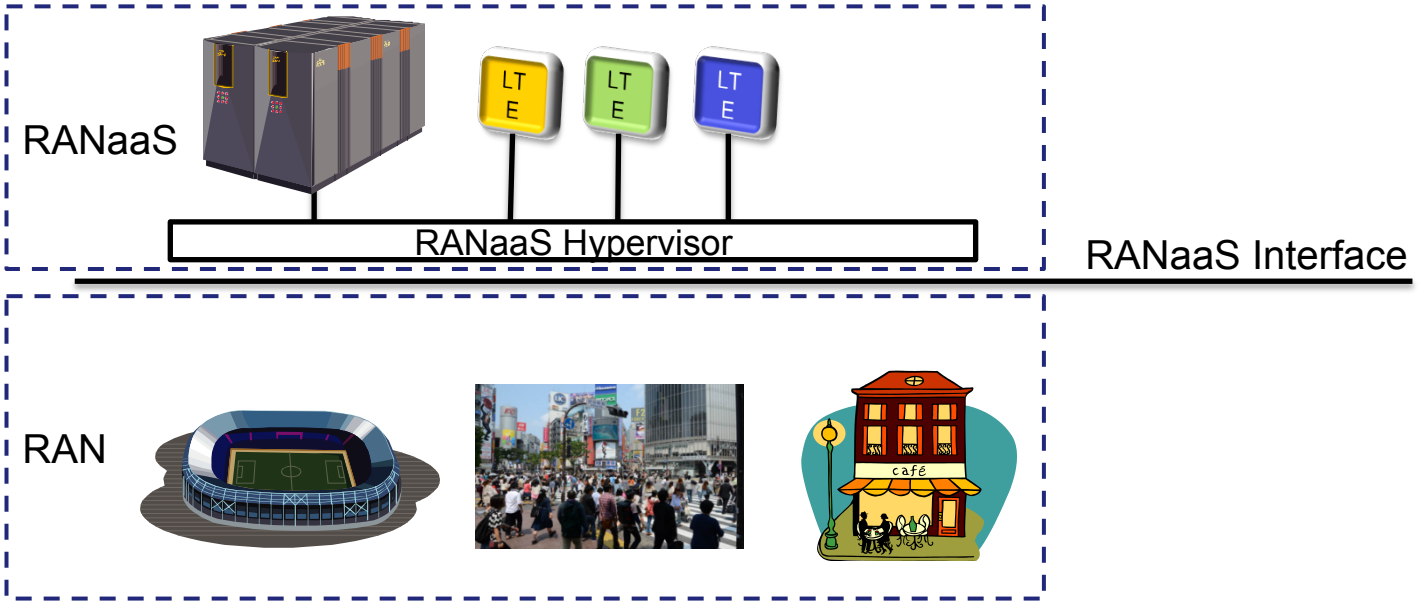
- Survey of Operator Motivations for C-RAN
  - 86% : Cost savings
  - 64% : Better interference management
  - 57% : More effective traffic management
  - 57% : Agility
  - 41% : Low power consumption

Statistical multiplexing gain (processing)



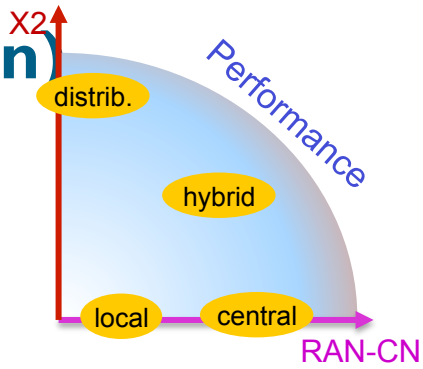
- Exploitation of temporal and spatial traffic fluctuations
- Efficiently use available resources, scale resource according to needs (resource pooling, elasticity)

Customization benefits (flexibility)

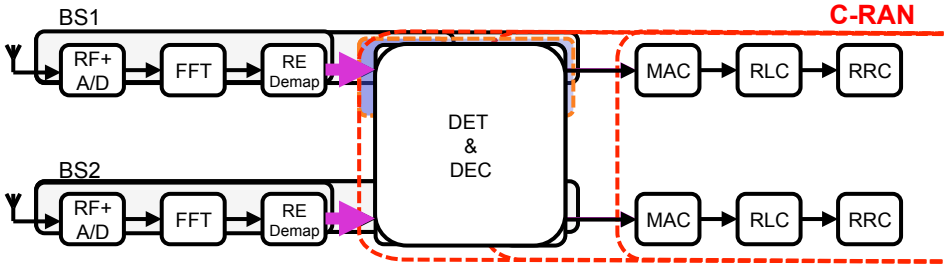


- Optimisation based on purpose, deployment, …
- Using software implementation rather than configuration (SON)
- Flexible software assignment over time and space

Performance improvement (cooperation)



- PHY processing approaches: Uplink
  - Local processing in BSs
  - Distributed processing among BSs
  - Central processing
  - Hybrid local / distributed / central processing

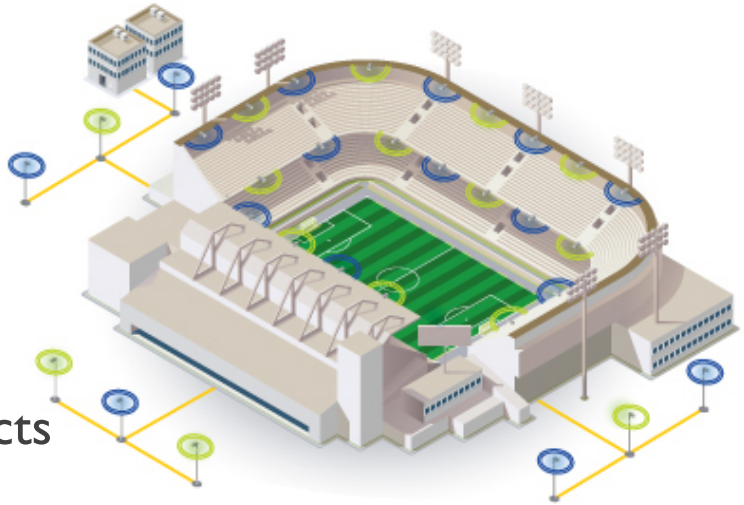


Challenges of C-RAN

- Requires significant infrastructure for fronthaul
  - High throughput requirements
  - Can be alleviated by data compression techniques
- Timing pressures
  - Data must be quickly transmitted over the fronthaul
  - Processing must be done in real time and within a hard deadline
  - Characterizing the performance of a system where missing the deadline is oftentimes more limiting than the degradation of the channel itself ("computational" outage)
- Virtualization technologies
  - Need to efficiently share the computing resources in real time

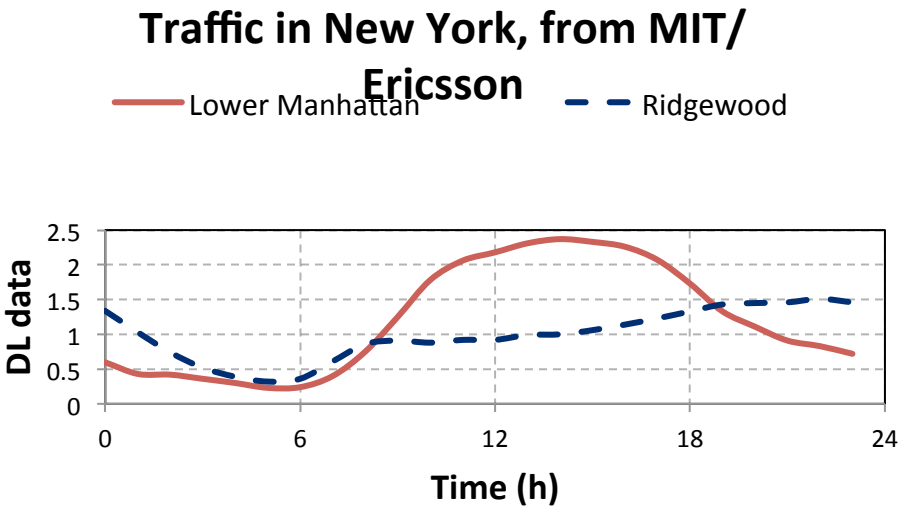
C-RAN vs. DAS

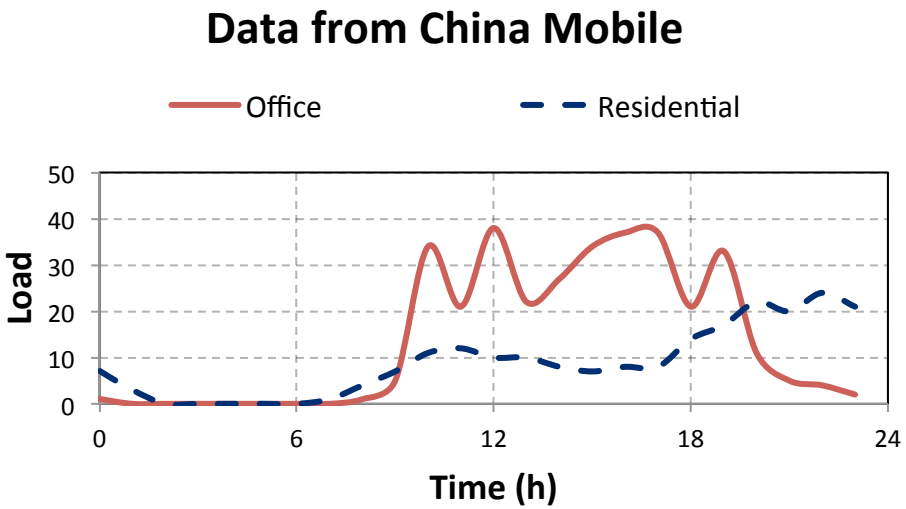
- C-RAN technology is reminiscent of Distributed Antenna Systems
  - DAS systems also connect multiple antennas to a centralized BBU.
  - Earlier DAS systems used the antennas to service a single cell, while in a C-RAN system, each antenna placement corresponds to a different logical cell.
  - More advanced DAS systems used antennas to service different sectors.
  - C-RAN can be considered to be an evolution of DAS, whereby the antennas serve different cells rather than merely different sectors, and optical interconnects are used between RRH and BBUs.



[3]. A. A. M. Saleh, A. J. Rustako and R. S. Roman, "Distributed Antennas for Indoor Radio Communications," *IEEE Transactions on Commun.*, vol. 35, pp. 1245-1251, Dec. 1987.

Tidal Usage Patterns

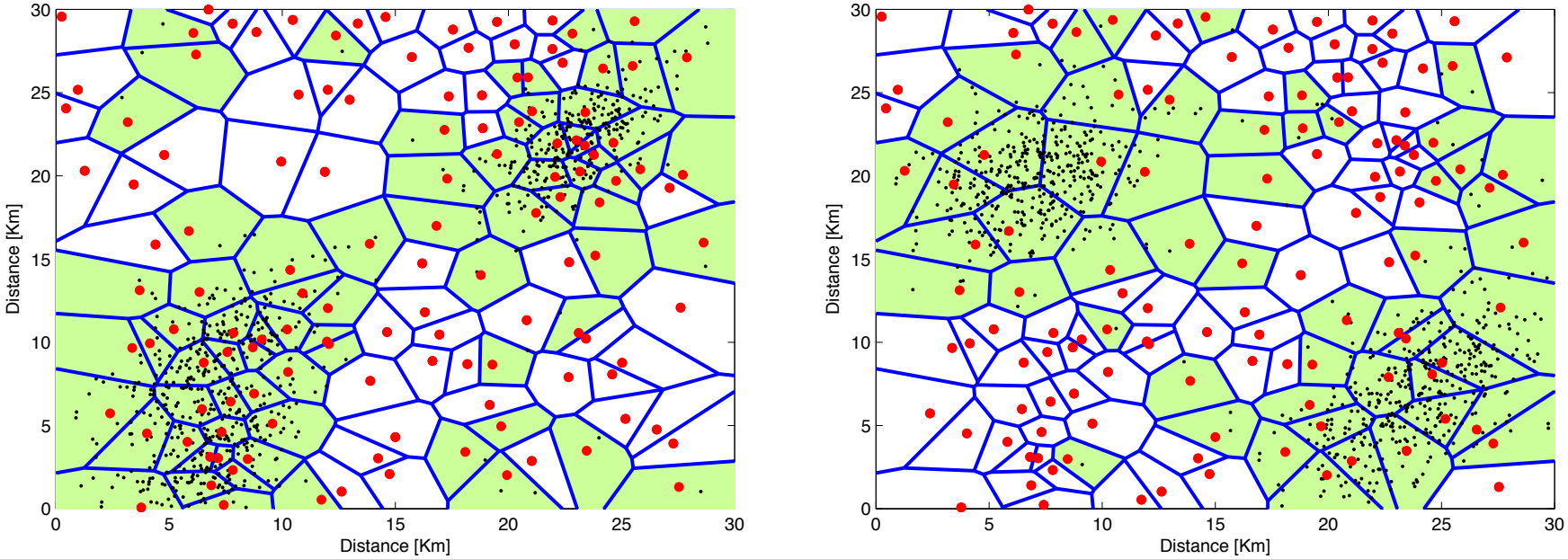




- Daily load depends on the base station location
- Peak traffic can be more than 10X that of off-peak traffic

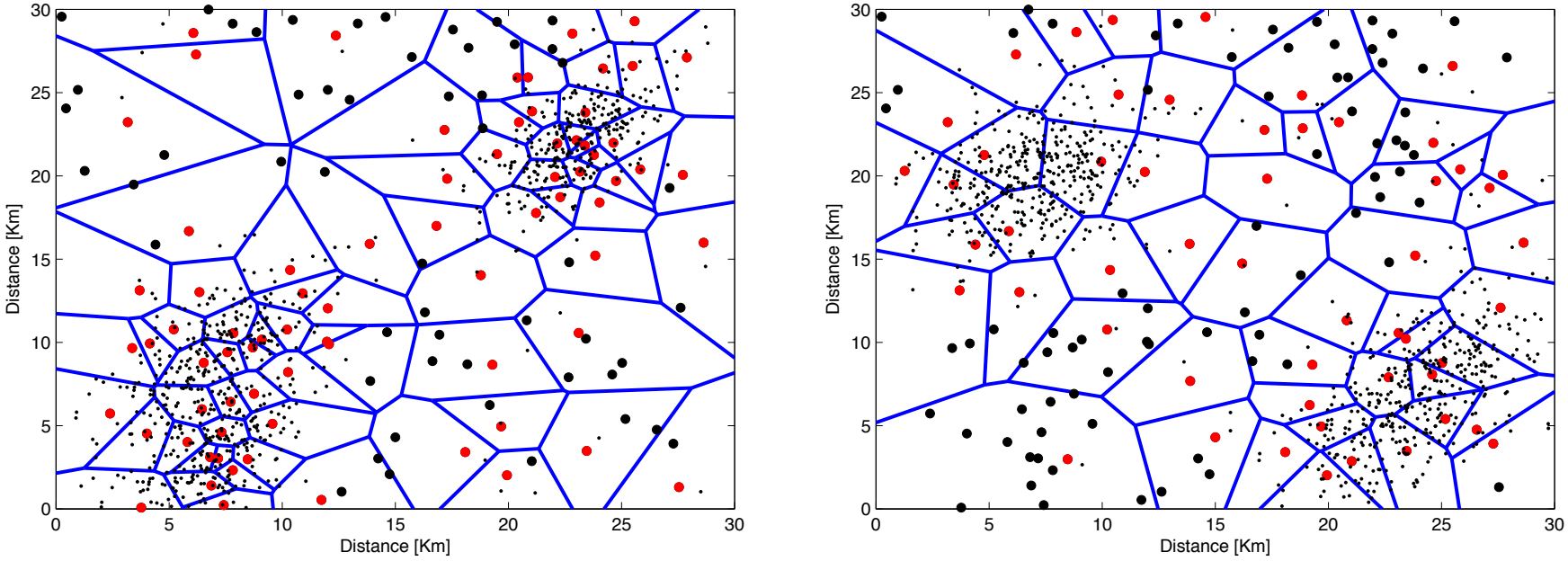
Sources: www.manycities.org, and [2] China Mobile Research Institute, "C-RAN: The Road Towards Green RAN," White Paper, 2010.

Tidal Usage Patterns



- Example showing the migration of users at two different times.
- Statistical multiplexing gain is enhanced by joining cells with disparate traffic patterns into the same cloud group

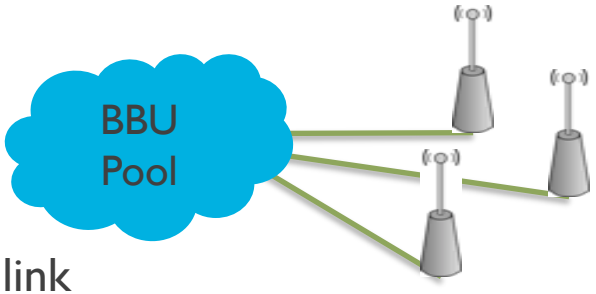
Switching off RRHs: Greener Networks



- Under-utilized RRH's can be switched off, thereby saving power
- This creates larger cells
- Particularly effective for dense deployments

Network Upgrades

- Upgrades can be done for two purposes
  - Increasing coverage
    - Add more RRH to the existing network
    - Need to also add the corresponding fronthaul link
    - RRH's are small and compact, since they have no baseband processing



  - Increasing capacity
    - Add more BBU to the pool
    - If virtualized, translates to adding more computational resources

Upgrading and Maintaining BBUs

- A failed compute server is easily swapped out
- Software easily upgraded when BBU hardware is centralized
  - Very little maintenance needed at the remote sites
- Compute hardware and software upgrades are decoupled
  - Different vendors for each?
- Additional comm standards supported through software update
  - New RRH typically only needed for new RF bands

The Cloud in C-RAN is Different than Cloud in General

	IT - Cloud computing	Telecom - Cloud RAN
Client/base station data rate	Mbps range, bursty, low activity	Gbps range, constant stream
Latency and jitter	Tens of ms	$<$ 0.5 ms, jitter in ns range
Life time of information	Long (content data)	Extremely short (data symbols and received samples)
Allowed recovery time	s range (sometimes hours)	ms range to avoid network outage
Number of clients per centralized location	Thousands, even millions	Tens, maybe hundreds

[1]. A. Checko et al, "Cloud RAN for Mobile Networks – A Technology Overview," IEEE Comm. Surveys & Tutorials, 1Q 2015.

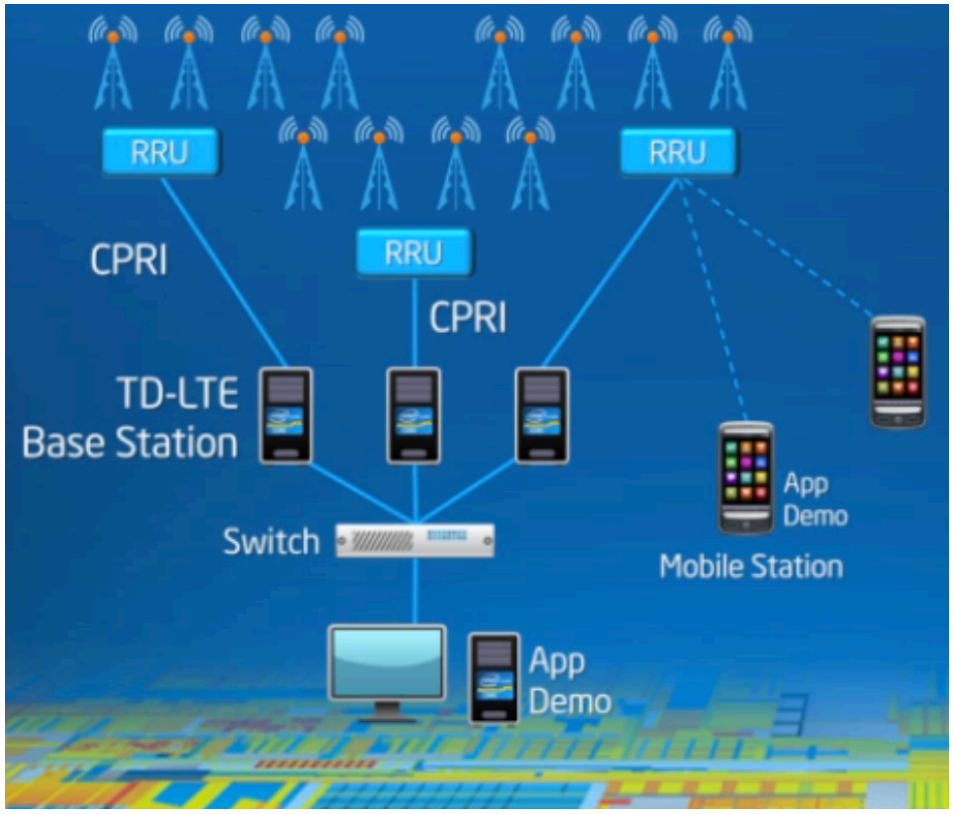
BBU Hardware Architectures

	DSP	GPP
Flexibility	dedicated solution	general purpose
Vendor compatibility	vendor specific, proprietary	higher compatibility between vendors
Backward compatibility	limited	assured
Power consumption	lower	higher
Real-time processing	optimized, achieved	only possible with high power hardware
Virtualization of BBU	possible	possible

- Digital Signal Processor (DSP) vs. General-Purpose Processor (GPP)
- Tradeoff between using commodity GPP hardware, vs. special-purpose DSP hardware.

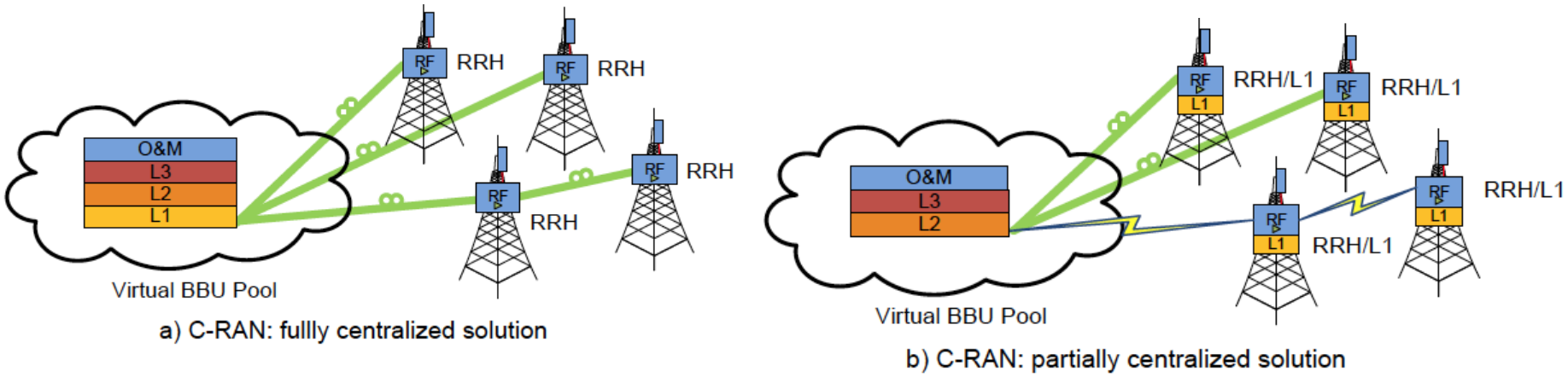
LTE on a GPP?

- Intel has developed the C-RAN LTE Base Station Architecture
  - A complete LTE base station with Intel Core i7 processor and software
  - Parallelizes baseband processing across multiple cores
  - No need to special-purpose hardware
  - Leverages unused CPU cycles to run other network applications



http://www.intel.com/content/www/us/en/communications/communications-c-ran-solution-video.html

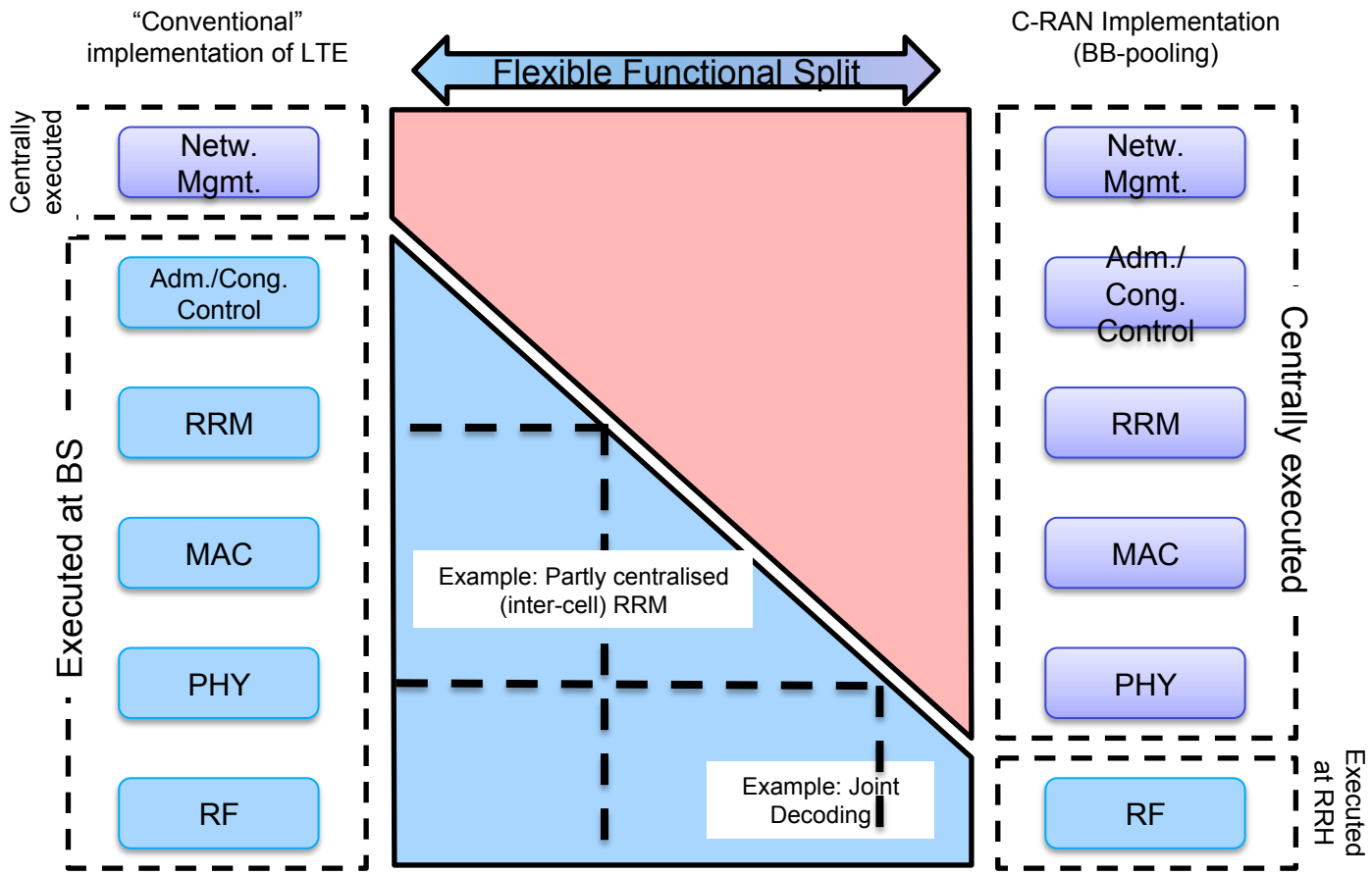
RRH / BBU Functional Split: Partial Centralization



- Rather than doing all L1 processing in the BBU Pool, some of the L1 processing can be done at the RRH.
- Reduces the load on the fronthaul.
- Increases computational load at the RRH.

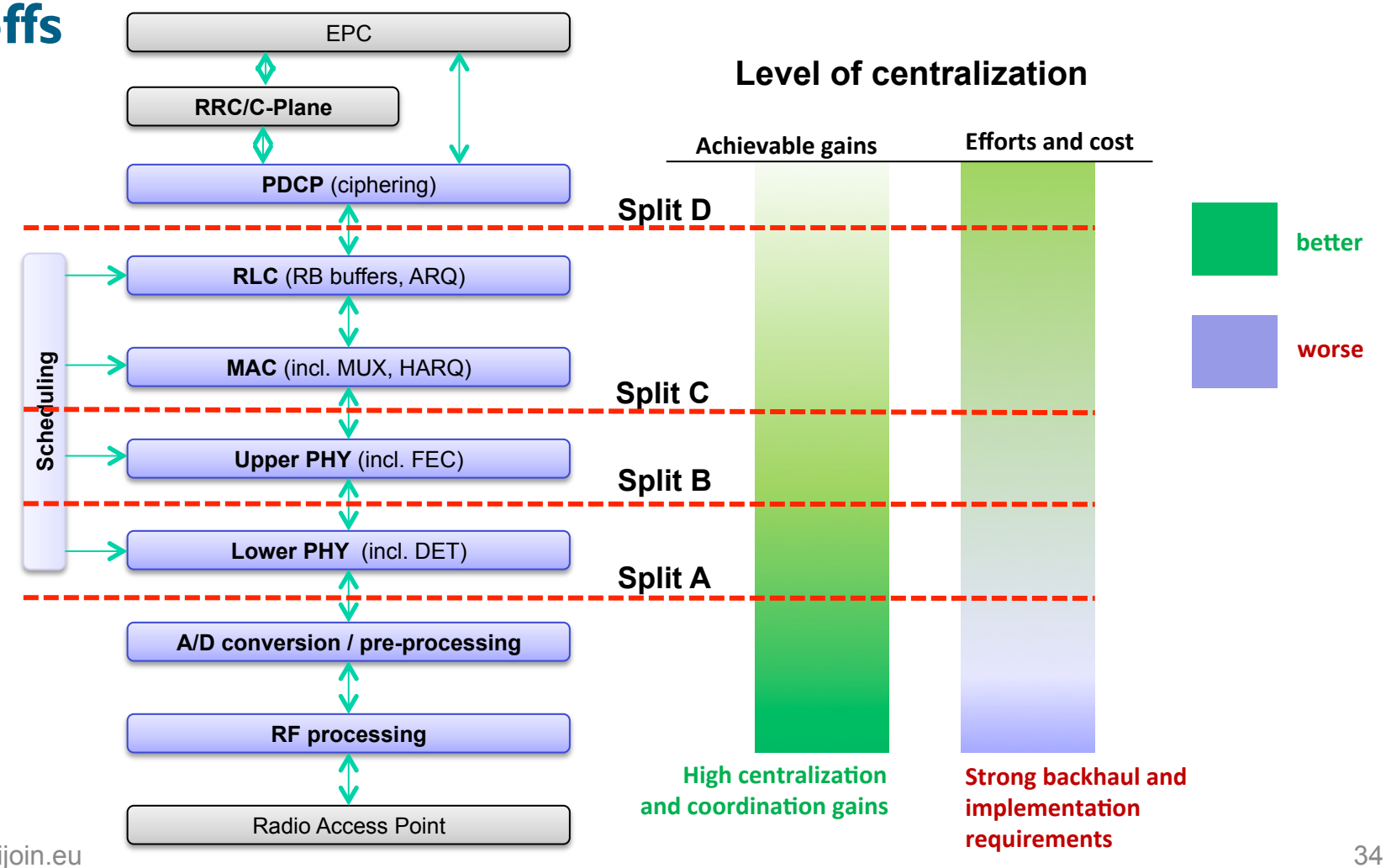
[1]. A. Checko et al, "Cloud RAN for Mobile Networks – A Technology Overview," IEEE Comm. Surveys & Tutorials, 1Q 2015.

RRH / BBU Functional Split: Partial Centralization



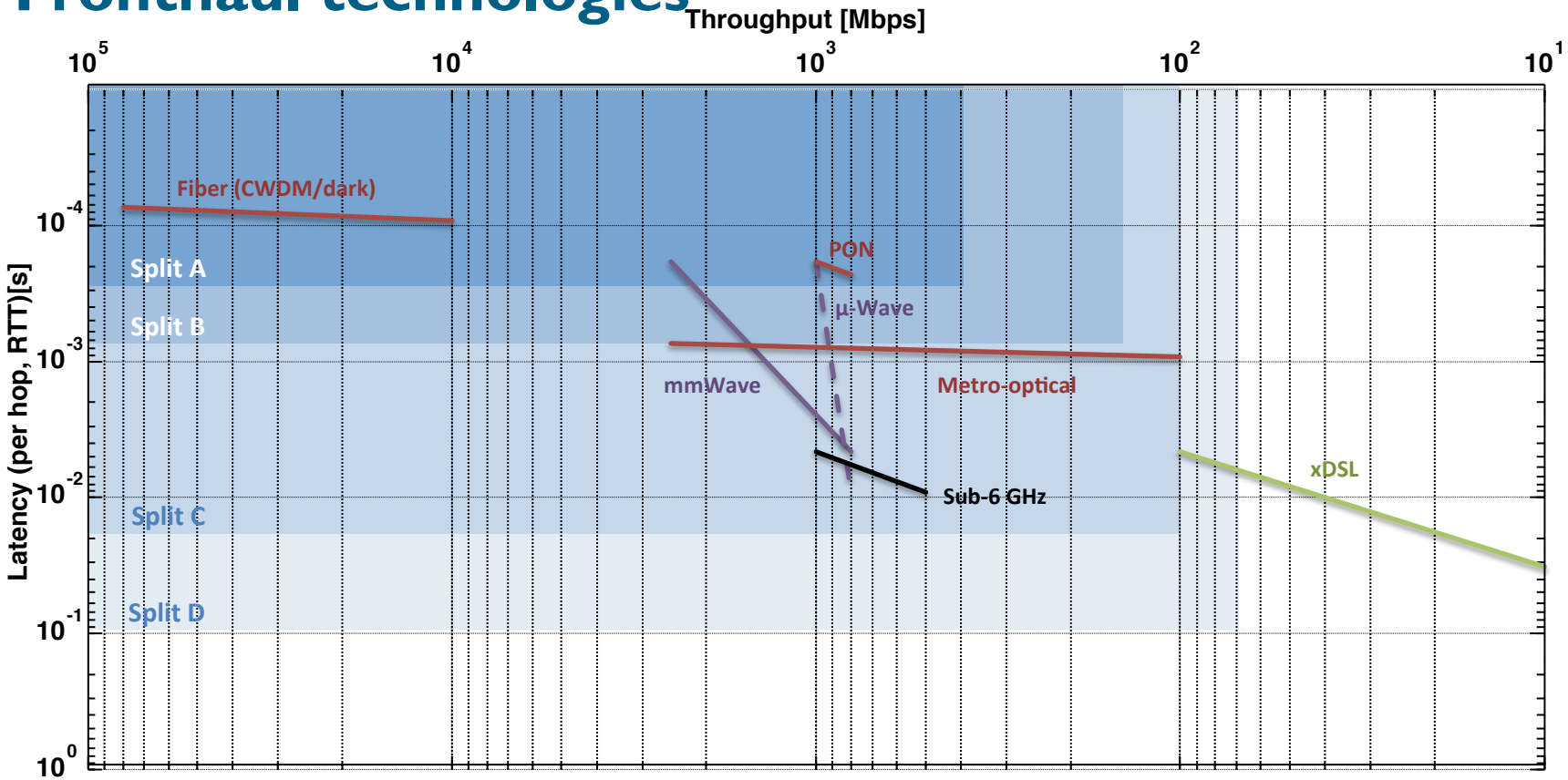
Source: www.ict-ijoin.eu

Tradeoffs



Source: www.ict-ijoin.eu

Fronthaul technologies



Source: www.ict-ijoin.eu

Fronthaul Compression

- To alleviate the demands on the fronthaul, compression can be used
- Standard compression techniques have limited success
  - Reduced sample rate and compressive sampling
  - Non-uniform and vector quantization
- Subcarrier compression
  - Requires FFT/IFFT at the RRH
  - Digital automatic gain control
- Special attention should be paid to the random access channel (RACH)
- Partial centralization can be considered a type of compression

Survey of Fronthaul Compression Techniques

Method	Techniques applied	Compression ratio	EVM
[10]	Not available	44%	Not available
[76]	removing redundancies in spectral domain	28%	3%
	preforming block scaling	23%	4%
	usage of non-uniform quantizer	17%	8%
[77]	removing redundancies in spectral domain	52%	$> 1.4\%$
	preforming block scaling	39%	$> 1.5\%$
	usage of non-uniform quantizer	30%	$> 2.5\%$
[78]	adaption of dynamic range of the signal	50%	0.5%
	removal of frequency redundancy	33%	3%
	IQ compression	25%	8%
[79]	removal of frequency redundancy optimized control information transmission IQ compression user detection	33% (100% cell load) 7% (20% cell load)	Not available
[80]	self-defined robust method performed jointly with base station selection algorithm	Not available	Not available

[1]. A. Checko et al, “Cloud RAN for Mobile Networks – A Technology Overview,” IEEE Comm. Surveys & Tutorials, 1Q 2015.

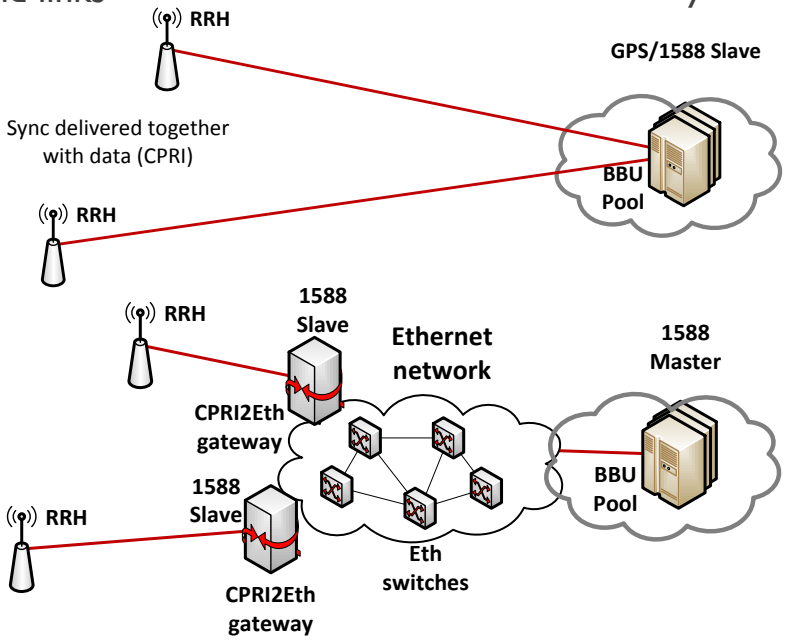
Fronthaul optimization

- Compression – up to 4:1, 3:1
- With load-dependent functional splits (Split B and above) higher bandwidth savings can be achieved on fronthaul
  - Up to reaching user data level – compression: 2.5Gbps/150 Mbps = 16:1
- Variable bit rate traffic fits packet-based transport solution

Shared Ethernet for cost-saving and flexibility

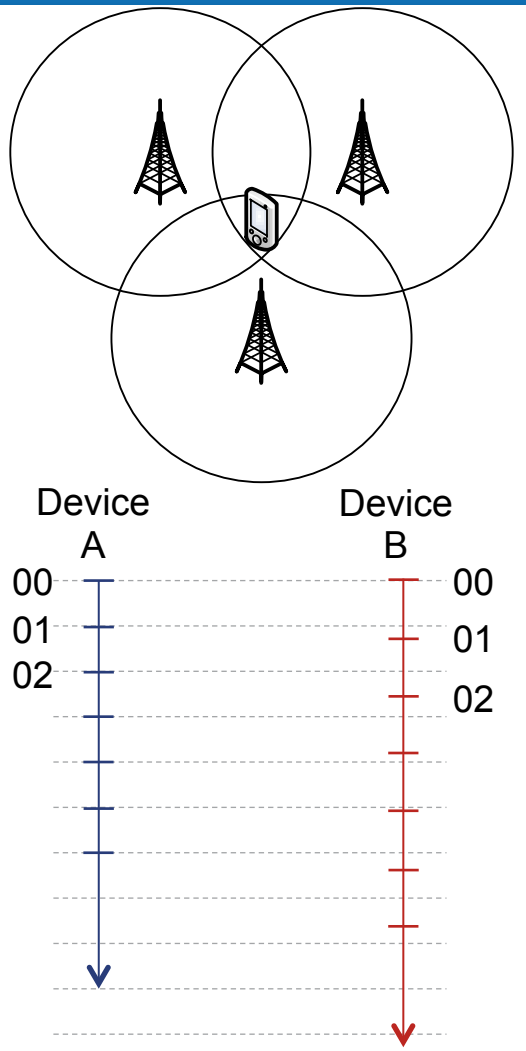
- ✓ Widely deployed (reuse!)
  - ✓ Dedicated links
  - ✓ Shared links
- ✓ Aggregation
  - ✓ Multiplexing gains on BBU and links
- ✓ Switching

- ! Fronthaul cost savings vs problems with delays and synchronization
  - ! Synchronous CPRI vs asynchronous Ethernet
  - ! Data delay: 100-400 us, ≈constant



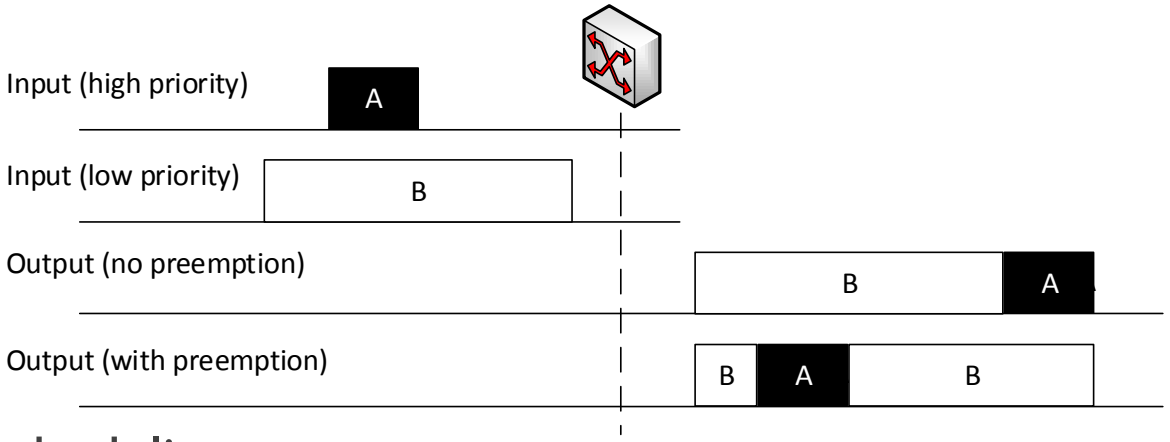
Timing in fronthaul

- Timing is really important
  - Frequency of transmission
  - Handover, coding, cooperative techniques, positioning
- Requirements (4G)
  - Frequency error LTE – A TDD/FDD: ±50 ppb
  - Phase error LTE-A with eICIC/CoMP: ± 1.5 - 5 μ s, MIMO: 65 ns, positioning: ± 30 ns
- Current solutions for timing distribution
  - GPS
  - PHY layer clock – SyncEth
  - Packet-based timing
    - IEEE 1588v2 (PTP)
  - Multiple

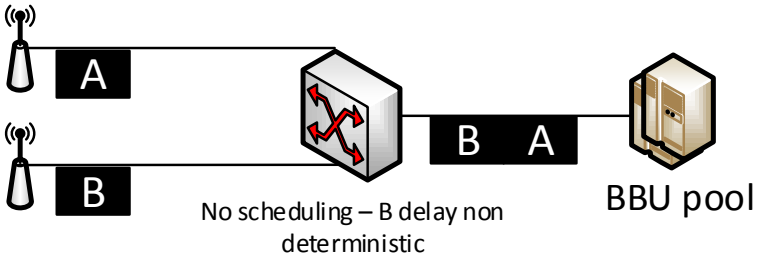


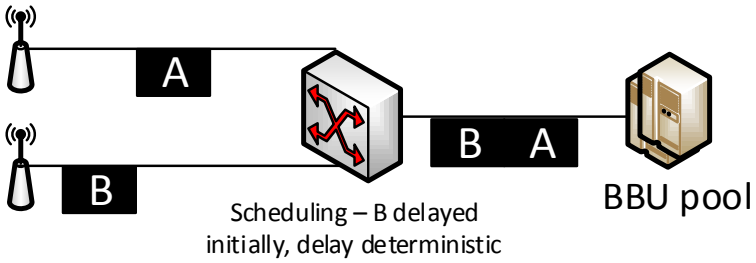
How to reduce queueing delays on a shared fronthaul?

- Preemption (switch upgrade required)



- Source scheduling





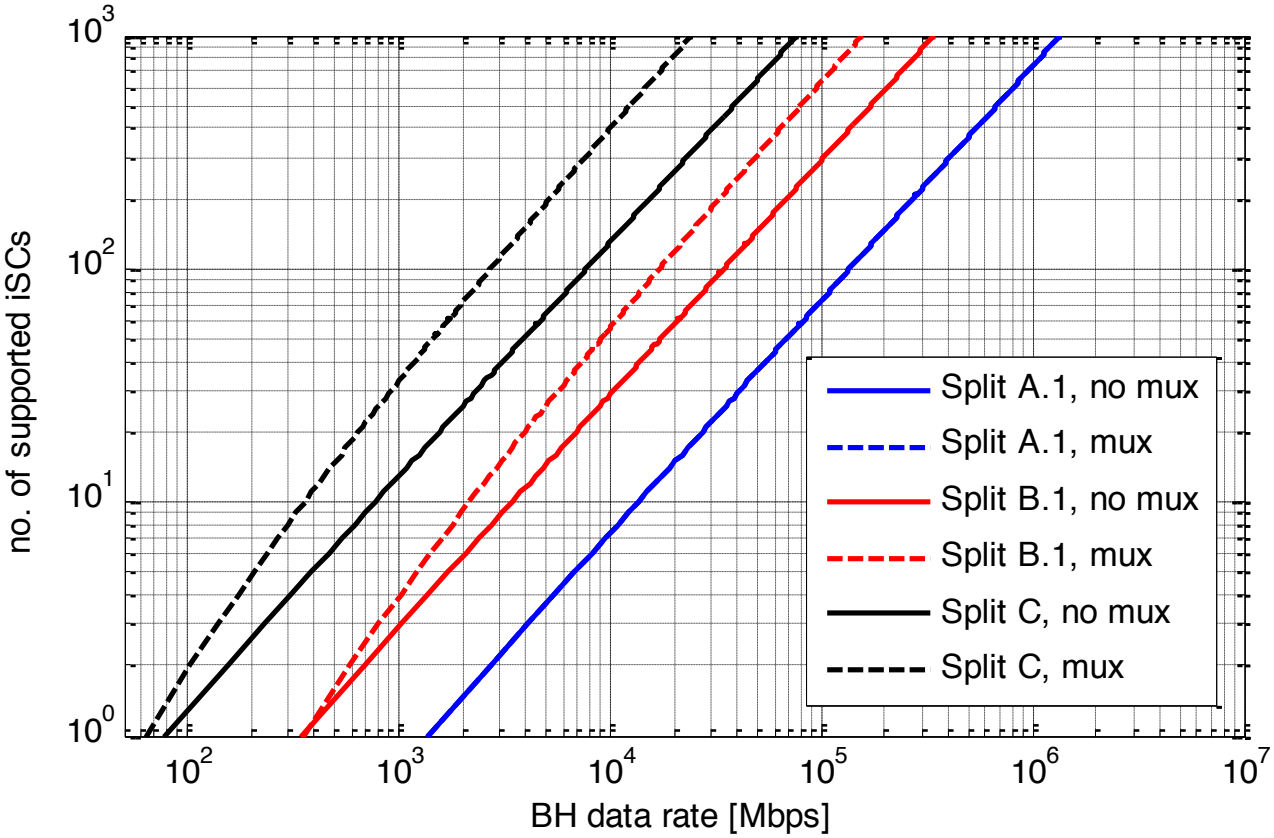
SDO activities on fronthaul

- IEEE 1904.3
  - Encapsulation and mapping of IQ data over Ethernet,
  - Fields need to be compatible with sync solution
- IEEE 1914 Next Generation Fronthaul Interface
  - Architectures for fronthaul transport
  - Requirements for the fronthaul network
  - Functional split
- IEEE 802.1, Time Sensitive Networking task force
  - Frame preemption (802.1Qbu)
  - Scheduled traffic (802.1Qbv)
  - Time-Sensitive Networking for Fronthaul (profile definition, 802.1CM)
- IEEE 1588v2 (-2008)
  - Packet based solution for time reference delivery
  - A working group preparing a new version of the standard with an expected completion date of 31 December 2017

→cooperation between activities

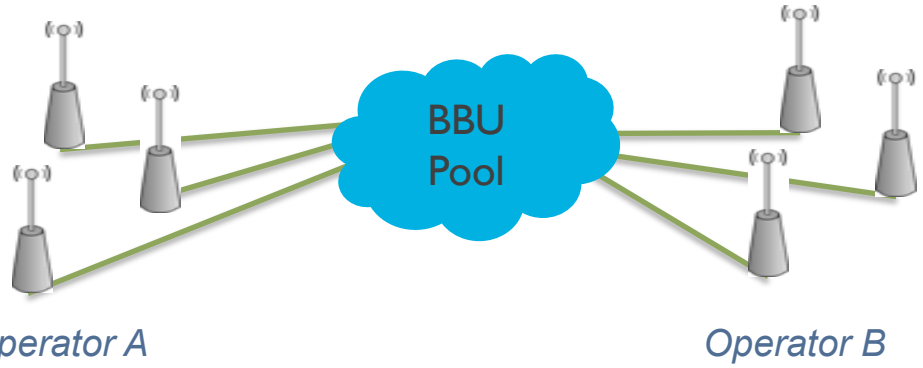
- Study item in China Communications Standards Association (CCSA)

Multiplexing gains (transport network)



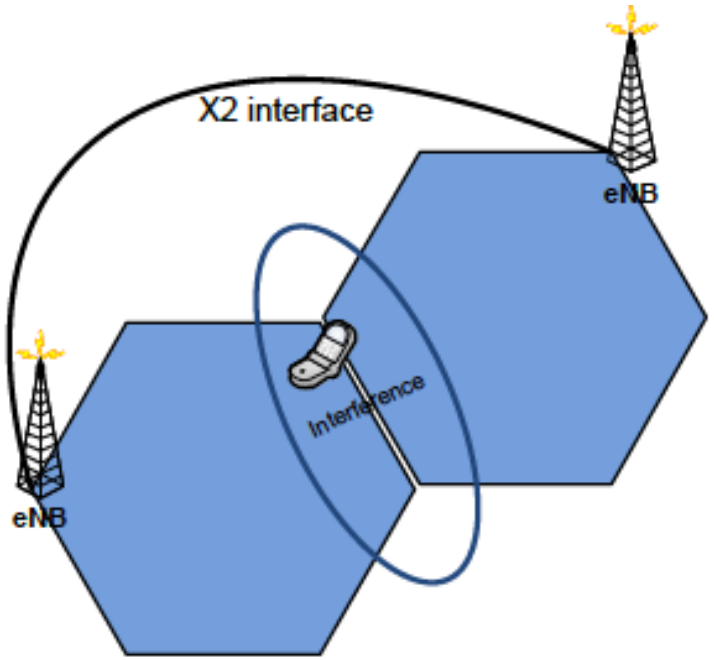
Sharing BBU Pools By Multiple Operators

- The BBU pool is not necessarily owned and maintained by the network operator.
  - A large metropolitan pool could handle baseband processing for multiple network operators.
  - Each operator would be a different tenant in the BB hotel.
  - BBU hosting is a new business model (RAN as a Service).



Enabling Tighter Inter-Cell Cooperation

- Inter-cell cooperation is standardized in LTE-A
  - eICIC: (enhanced) Inter-Cell Interference Coordination (R10)
  - COMP: CoOrdinated Multi-Point (R11)
- These features, as well as handover, require inter-cell communications
  - In a traditional architecture, done using X2 interface
- In a C-RAN environment, inter-cell coordination can be done without any additional latency
  - Opportunities for tighter coordination



[1]. A. Checko et al, "Cloud RAN for Mobile Networks – A Technology Overview," IEEE Comm. Surveys & Tutorials, 1Q 2015.

PART II: FROM C-RAN TO VIRTUALIZED MOBILE NETWORKS

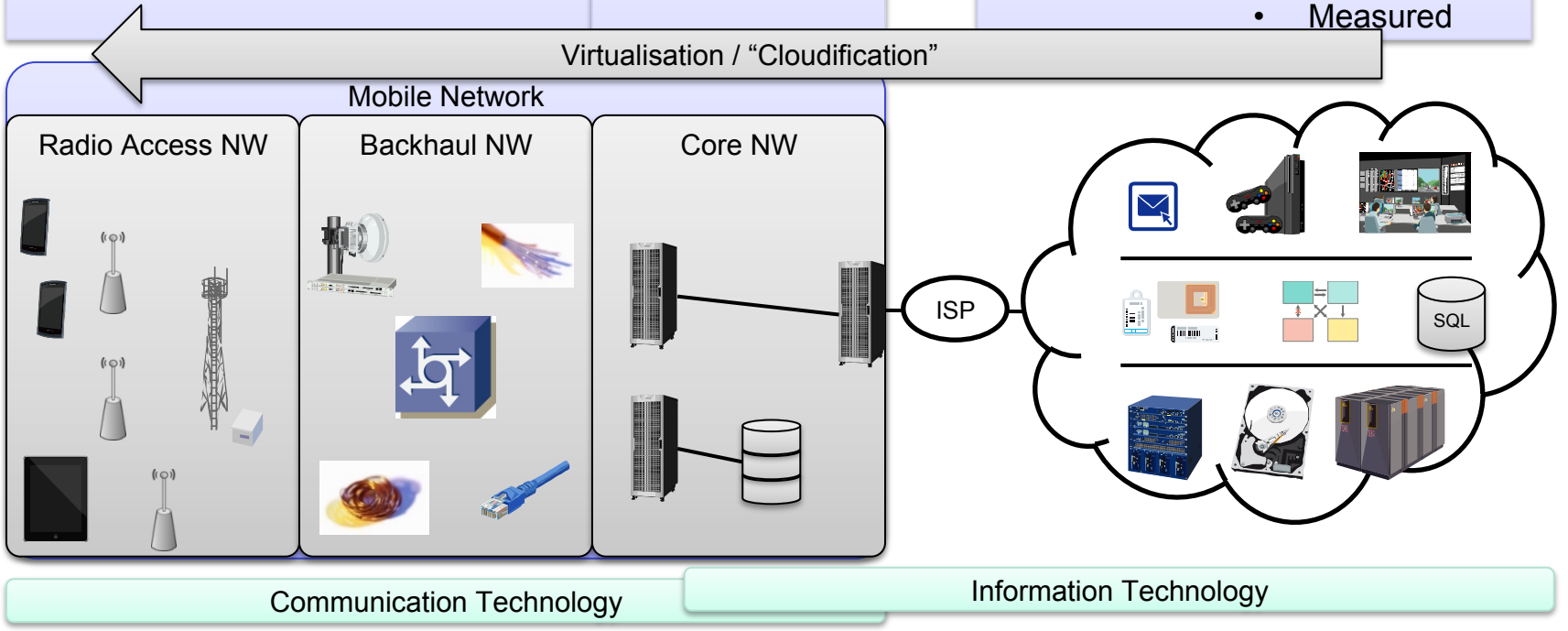
The cloudification of the network

- C-RAN
- RAN-Sharing
- SDN
- SDR

- NFV
- Soft-EPC
- SDN

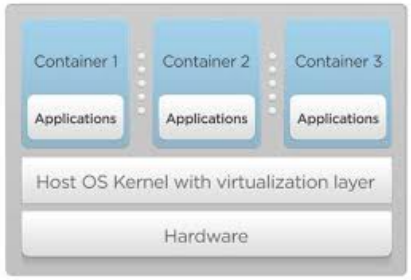
- SaaS
- PaaS
- IaaS

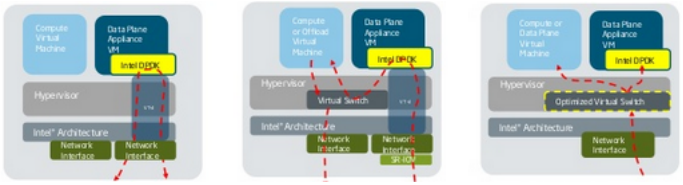
- On-demand
- Broad access
- Pooling
- Elasticity
- Measured



Cloud-Technologies

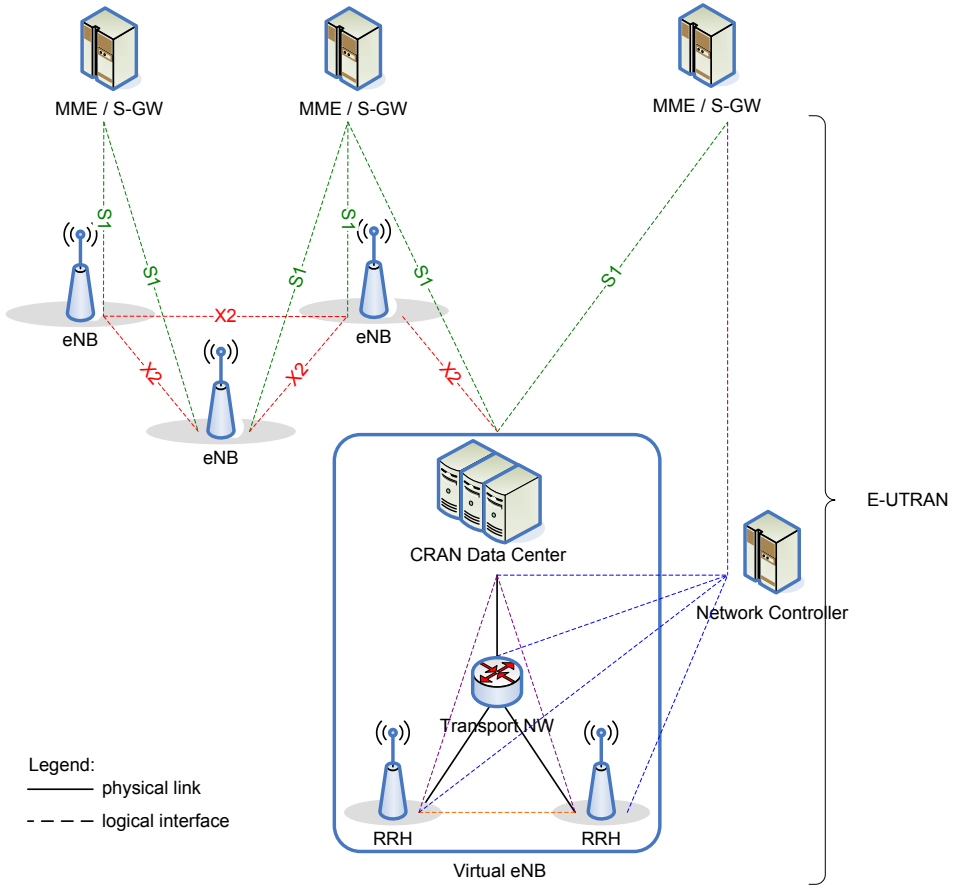
- Cloud computing will have a telecommunication specialized branch
  - Telco-grade clouds, telco over cloud,…
  - Self-healing systems
  - Carrier grade operating systems and hypervisors
  - High availability clouds
  - Quick response control planes
  - ……
- Container virtualization technology
  (virtualization of application layer)





- HW assisted virtualization
- Low-level, OS/hypervisor internal enhancements to overcome hypervisor limitations (e.g. interrupt latency optimization)

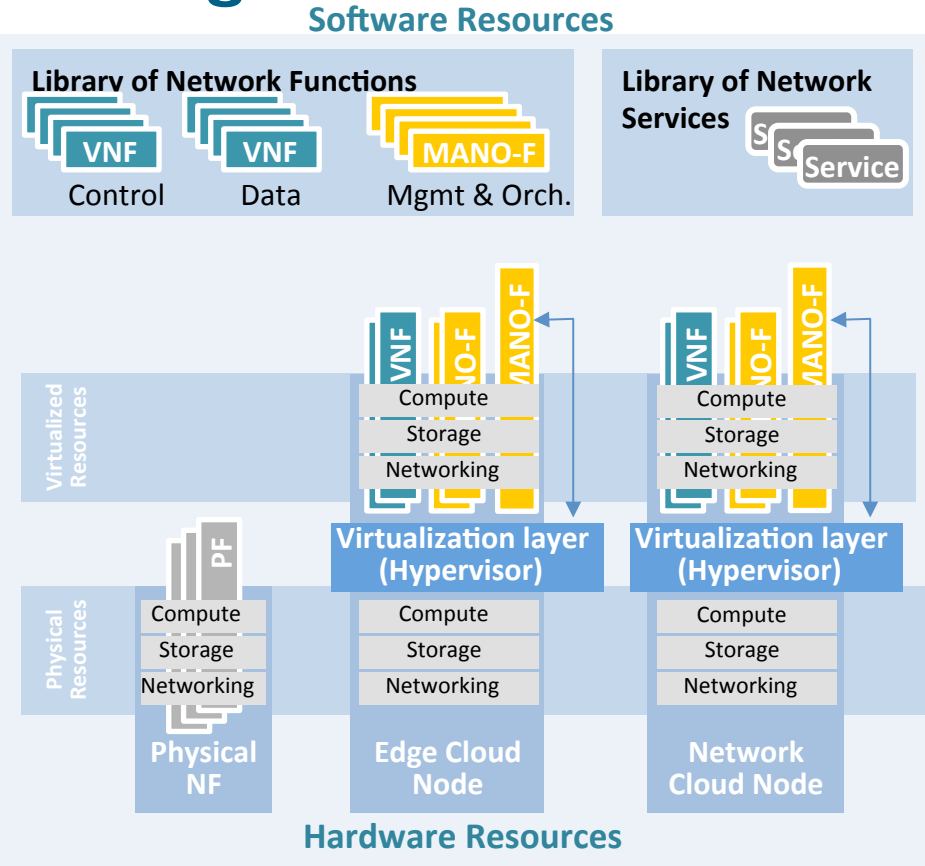
Integration in 3GPP LTE Architecture



Source: www.ict-ijoin.eu

Added

Integration in 5G Architecture

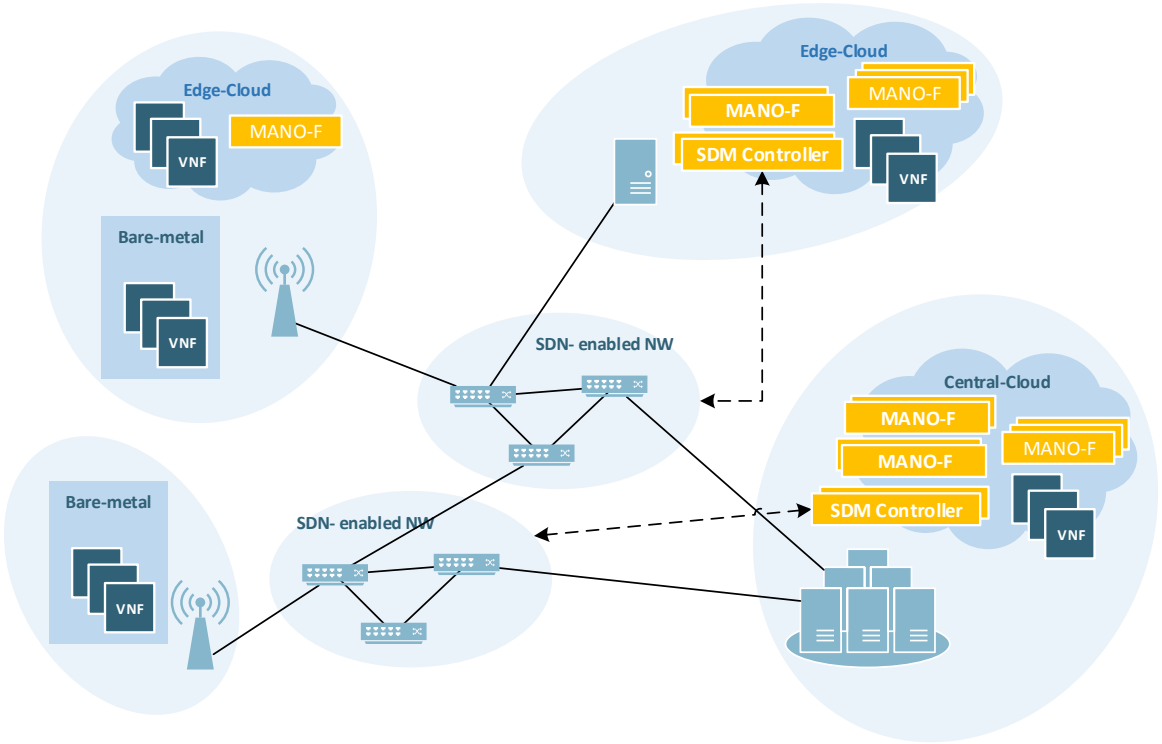


- Software libraries
  - functions
  - services
- Hardware resources
  - computing, storage, networking
  - virtualized and physical NF
- Edge and network clouds operate according to ETSI NFV principles
- Physical network functions
  - SW and HW cannot be decoupled
  - SW highly embedded in the HW

Source: 5gnorma.5g-ppp.eu

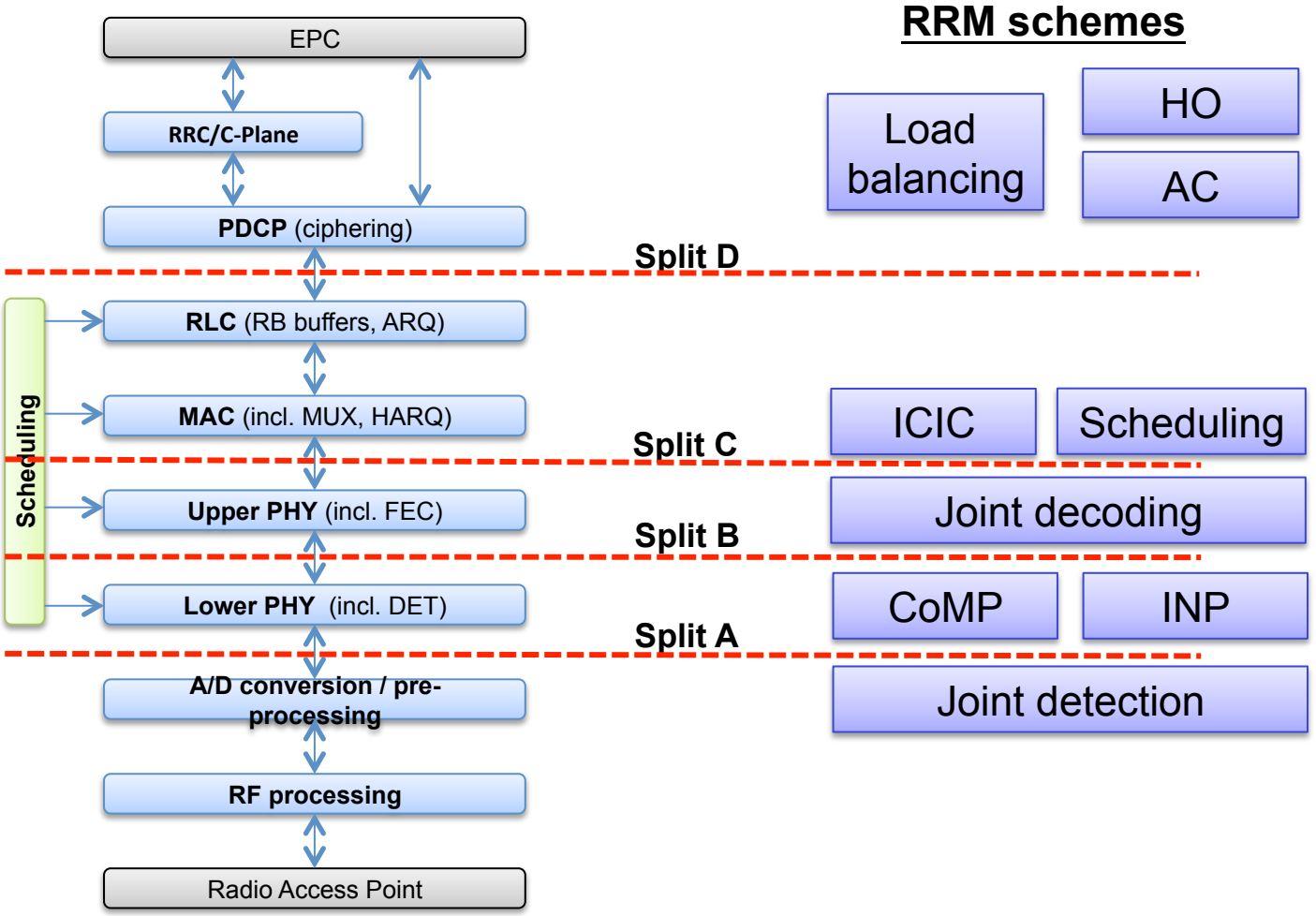
Integration in 5G Architecture

- Concrete illustration of network deployment and function placement
- Shows mapping of functions to geographical locations
- May be used to display per instance/per location properties
  - functions allocated
  - space demand
  - link capacity and latency



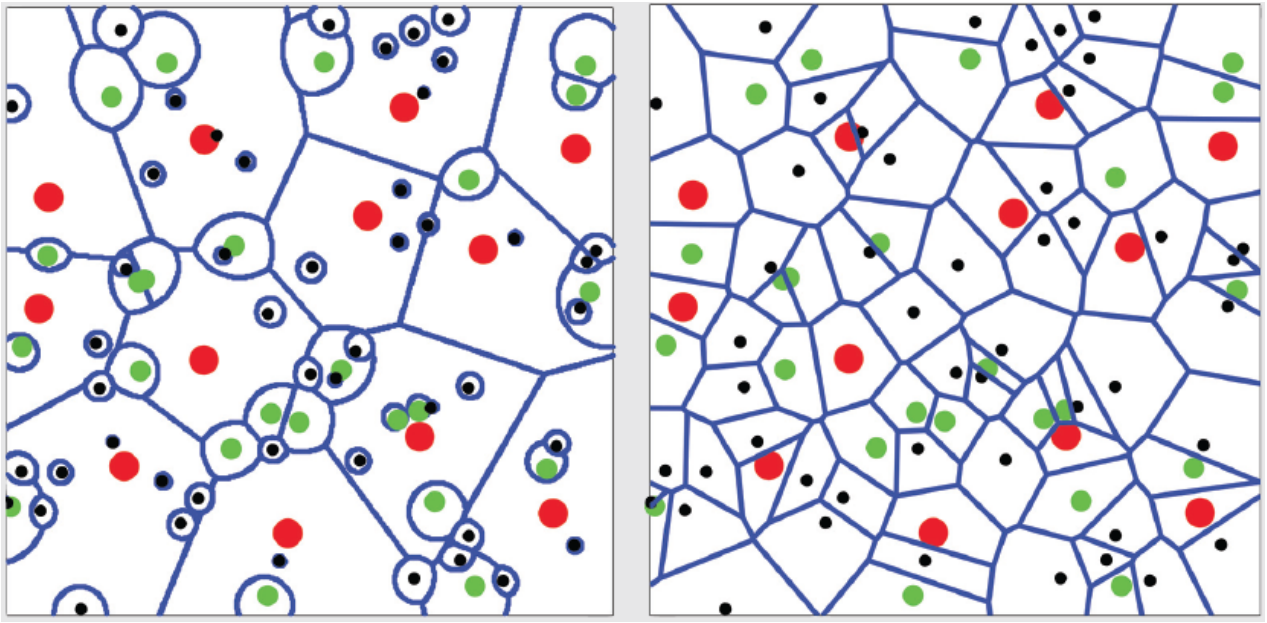
Source: 5gnorma.5g-ppp.eu

Functional Split Options



Decoupled Uplink and Downlink

Downlink



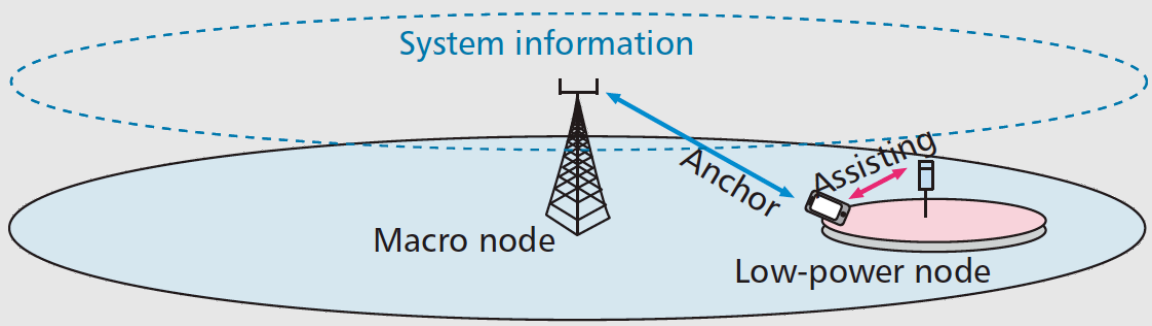
Uplink

- Uplink and downlink don't have to be serviced by the same RRH.
  - Uplink serviced by closest RRH
  - Dowlink serviced by the link with best SINR
  - Especially important in HetNet deployments

[4] J. Andrews, "Seven Ways that HetNets Are a Cellular Paradigm Shift," *IEEE Comm. Mag.*, 2013

Decoupling the Control and Data Planes

- Use high-powered microwave RRH for control information
- Lower-powered mmwave RRH for data transfer



- Device-Centric Networks ("Cell-Less")

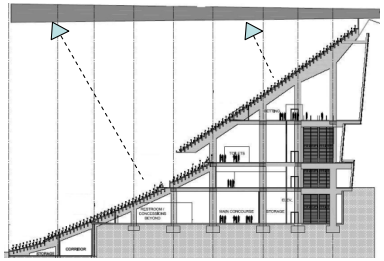
Exemplary deployments: Stadium

Use Case Description:

- Typical stadium: wide area (order of 50.000 m2)
- Average number of spectators: 40000.

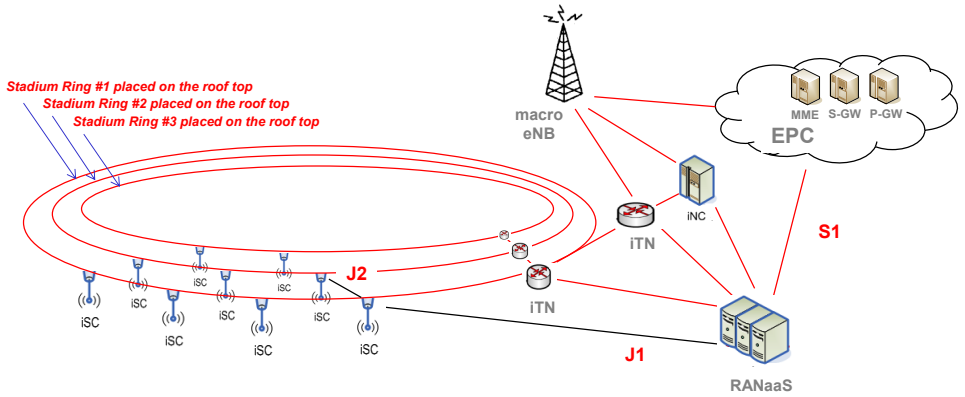
Physical Architecture Description:

- Architecture based on a number of Rings equipped with several RAP. Each Ring has a BH interface toward the EPC
- All the Rings and the Macro eNB coordinated by EMS and CRAN.



- very high density of RAPs that are connected through optical fibre in multiple rings
- local processing capabilities are provided in order to allow for breaking out local traffic and thereby reducing the traffic towards the core network

General Parameters	Value
UE density	~1 UE/m^2
Small Cells Number	50 - 300
Experienced UE DL+UL Throughput	1.5 – 10 Mbps
Traffic density	0.5 – 10 Mbps/m^2



Source: www.ict-ijoin.eu

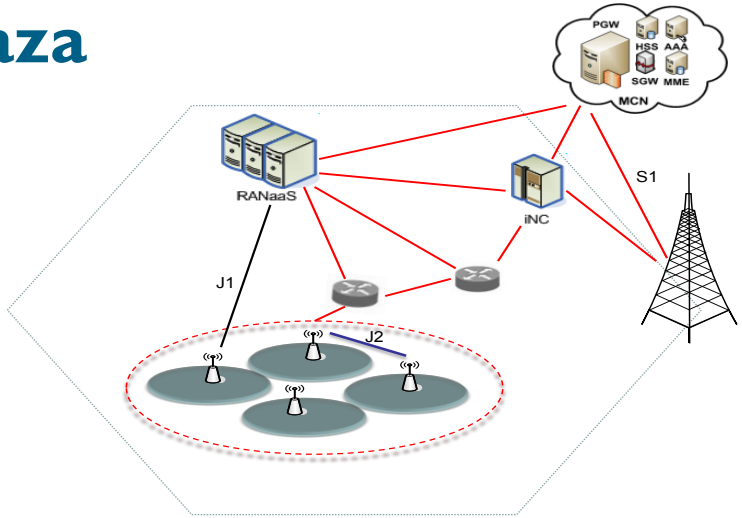
Exemplary deployments: Plaza

Use Case Description:

- Exemplary typical town square of 50.000 m2.
- Number of users: range from 1000 (off-peak hours) to 5000 (busy hours).
- Overall throughput in a square: from 1.5Gbps to 50Gbps (assuming 5% active UEs with 1.5-10 Mbps total throughput demand)

Physical Architecture Description:

- Random small cell deployment
- Random user deployment
- Static/Nomadic user
- Heterogeneous backhaul



General Parameters	Value
Number of small-cells	4-10 (sparse to dense deployment)
Number of UEs	15-30 (lightly to highly loaded scenarios)
Radius for small cell dropping in a cluster	50 m (3GPP TR 36.872 [1]) Random Dropping
Radius for UE dropping in a cluster	70 m (3GPP TR 36.872 [1]) Random Dropping
Minimum distance 3GPP TR 36.872 [1]	RAP-RAP 20 m UE-RAP 5 m Macro eNB-iSC cluster center 105 m
Backhaul Capacity / Latency	~50 - 100 Mbps, 1-10ms 10 Gbps, 5µs

Source: www.ict-ijoin.eu

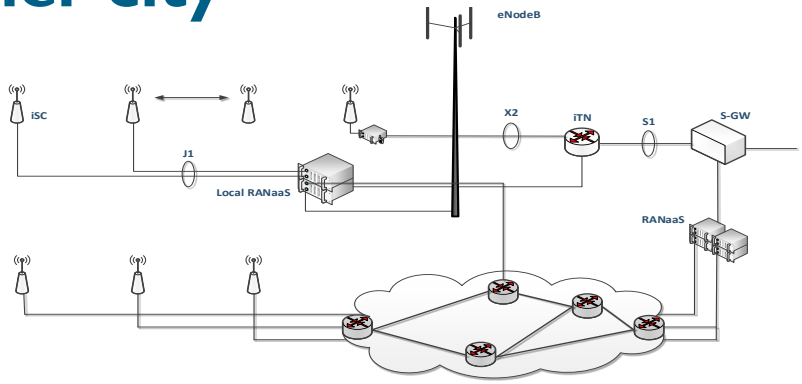
Exemplary deployments: Inner-city

Use Case Description:

- existence of hotspots and high demand areas, or coverage holes of the macro-cell layer
- Throughput: 10-15 Mbps (busy hour), assuming 10 MHz bandwidth.
- Layout: regular small cell deployment in a hexagonal grid, which covers 1 Km2.

Physical Architecture Description:

- Regular small cell deployment
- Random user deployment with Slow/High mobility
- Heterogeneous backhaul



General Parameters	Value
Number of small-cells	19
Number of UEs	1-3 UE per small cell (avg.)
Small cell dropping	Regular on Hexagonal Grid
ISD	50 m
UE dropping in a cluster	Random dropping
Minimum distance	UE-RAP 5 m
Backhaul Capacity / Latency	~~50 - 100 Mbps / 1-10ms ; 10 Gbps / 5µs

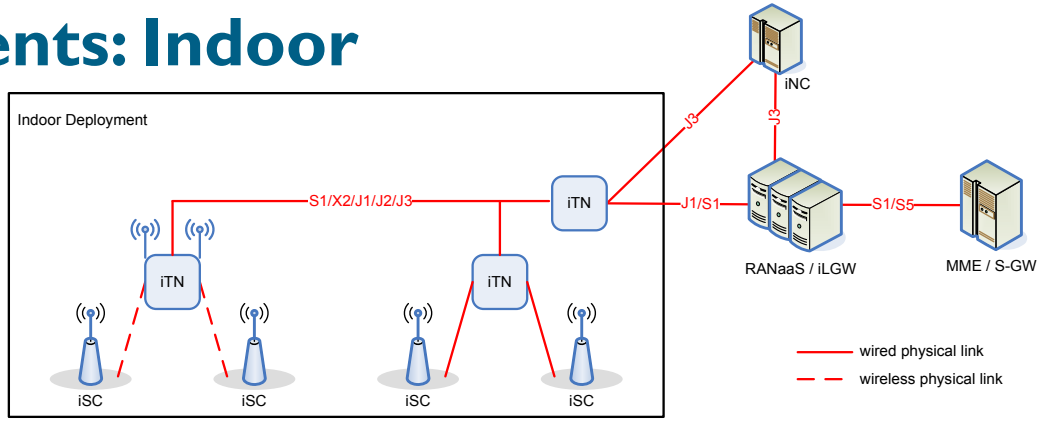
- Urban scenario where high capacity is continuously provided through RAPs
- Challenging scenario as each RAP needs to be connected to the backhaul network
- Number of aggregation points and CRAN data centers instances must be optimised

Source: www.ict-ijoin.eu

Exemplary deployments: Indoor

Use Case Description:

- Large amount of active users (range of 200-500) in one hall, with high communication activity.
- System throughput: from 200 Mbps to 5 Gbps (assuming throughput demand of 1 – 10 Mbps)



Physical Architecture Description:

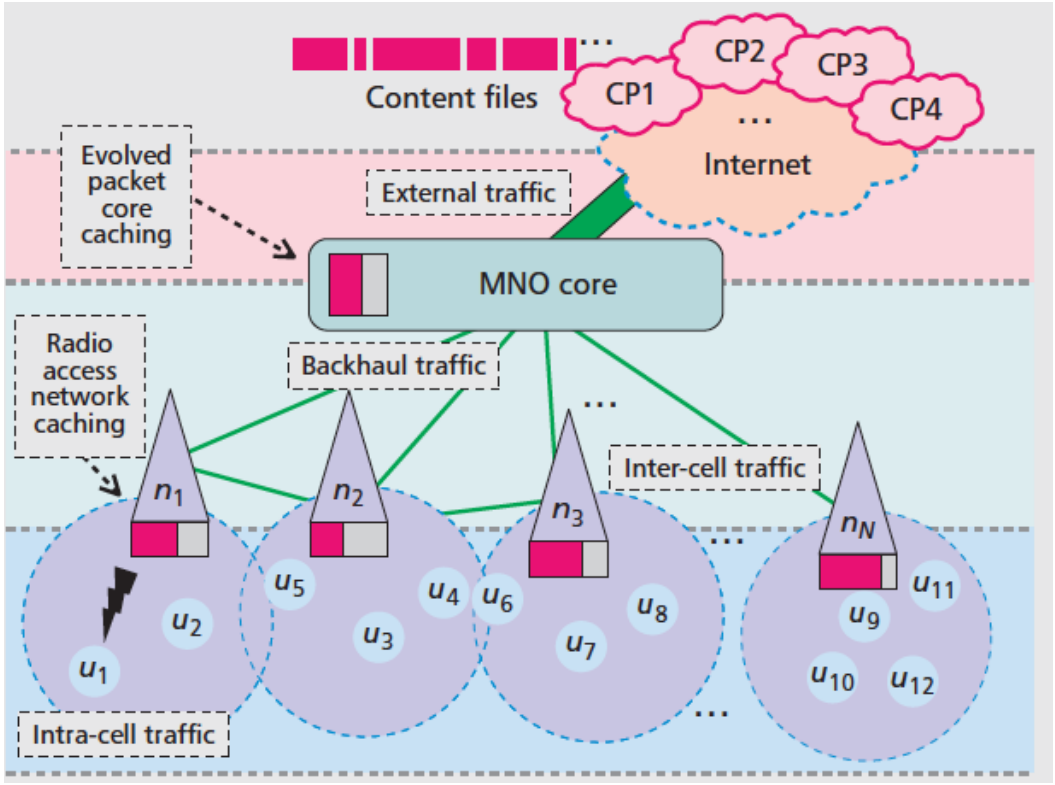
- Regular small cell deployment
- Random user deployment
- Nomadic user
- Wireline backhaul (optical fibre and ADSL)

General Parameters	Value
Number of areas per floor	16 Areas;1 or 2 Floors
Floor height	6 m
Area size	15 m X 15 m
Hall size	120 m X 20 m
Number of small cells	2 (sparse) / 4 (dense) per floor
Small cell dropping	Regular
Number of UEs	10 per small cell (sparse) 5/10 per small cell (dense)
UE dropping	Random
ISD	60 m (sparse) / 30 m (dense)
Minimum distance	UE-RAP 3 m
Backhaul Capacity / Latency	>100Mbps per RAP-EPC link <1, 10, and 50 ms

Source: www.ict-ijoin.eu

Enabling Higher-Layer Applications

- With C-RAN, can move applications and services from the core network to the network edge
  - RAN caching
- Examples
  - Voice-over-LTE
  - LTE-Broadcast
  - Content caching



[5] X. Wang et al, "Cache in the Air: Exploiting Content Caching and Delivery Techniques for 5G" *IEEE Comm. Mag.*, 2014

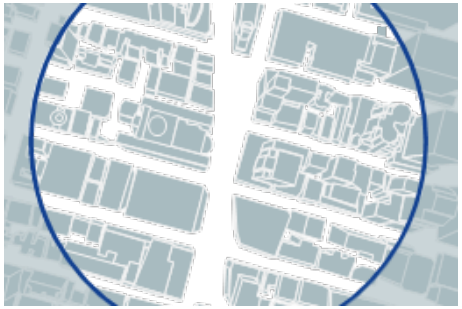
C-RAN Plays Well with Mobile Edge Computing

Hotspots

- Zonal applications
- E.g. special services in stadiums, exhibitions, malls, enterprise campuses
- Deployed in combination with Small Cell and Macro BTS (RRH, DAS)

Cities

- City applications
- E.g. IoT applications deployed as part of Smart City initiatives, or services for city residents and visitors
- Deployed at metro aggregation sites and baseband hotels

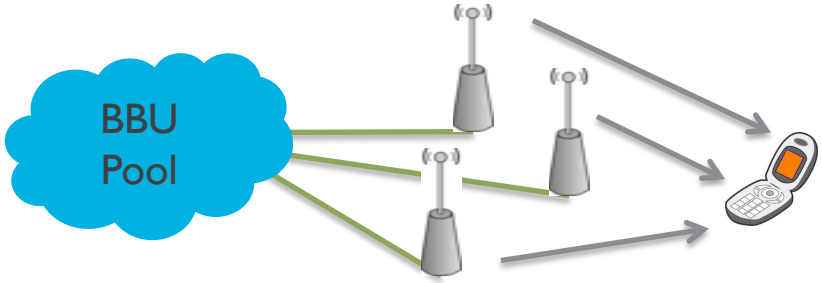


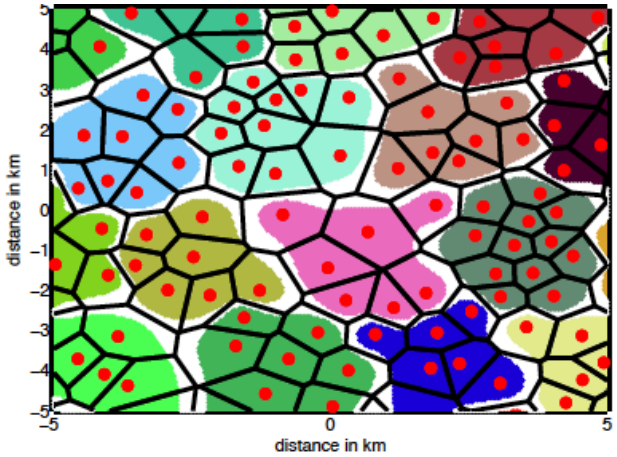
Network-wide

- Network-wide applications
- E.g. essential network functions, and ubiquitous services that require a consistent experience / performance
- Deployment in combination with Radio Cloud, or specific deployment patterns (e.g. Car2X along roads)

LTE Broadcast and Content Caching

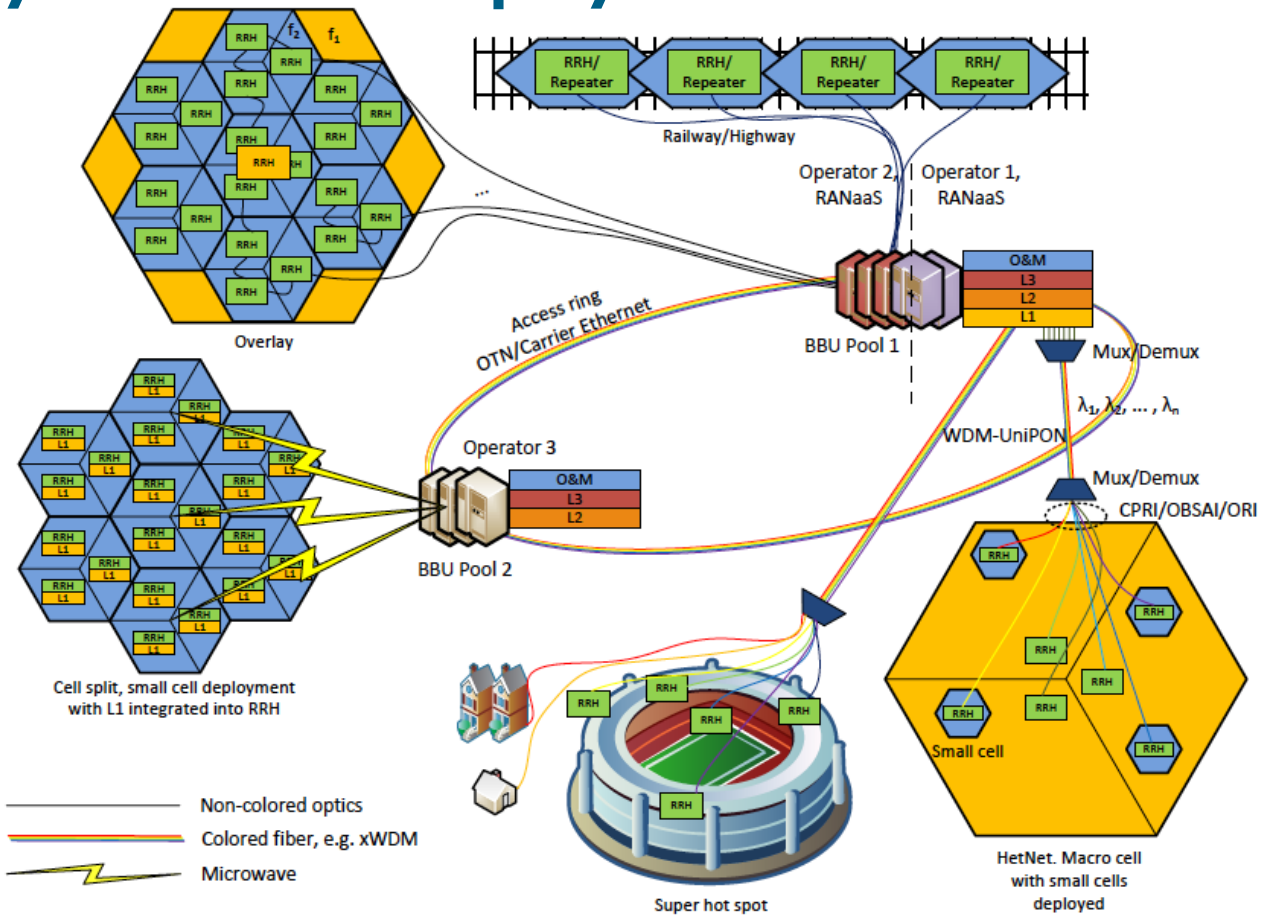
- LTE-Broadcast
  - aka enhanced Multimedia Broadcast Multicast Services (eMBMS) and MBSFN Single Frequency Network (MBSFN)
  - Special LTE subframes reserved for delivery of broadcast data
  - Same signal is transmitter from multiple base stations in an eMBMS area
  - C-RAN facilitates the transmission of identical signals





[6] S. Talarico and M.C. Valenti, "An accurate and efficient analysis of an MBSFN network, in *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, (Florence, Italy), May 2014.

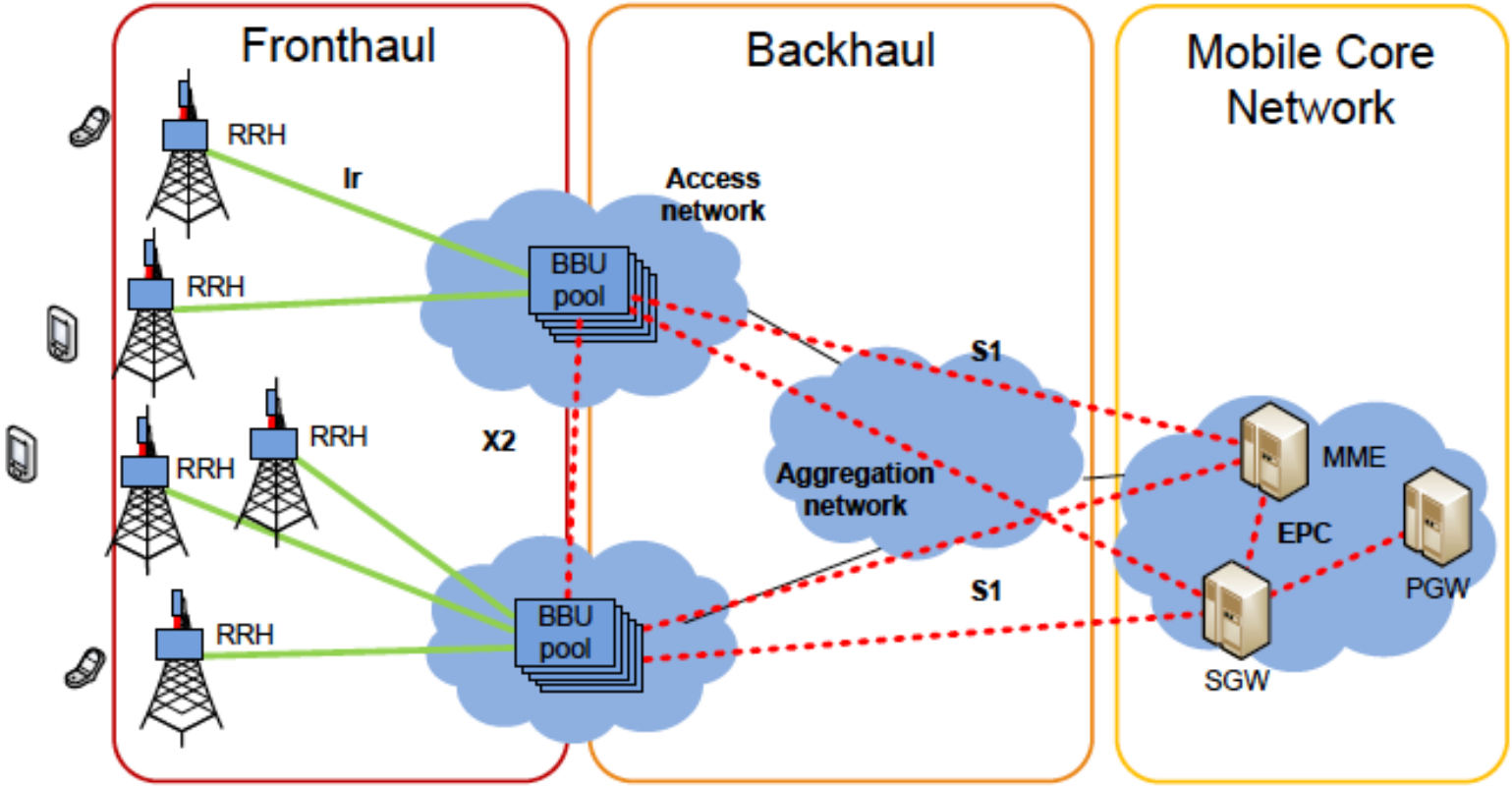
Summary of C-RAN Deployment Scenarios



[1]. A. Checko et al, "Cloud RAN for Mobile Networks – A Technology Overview," IEEE Comm. Surveys & Tutorials, 1Q 2015.

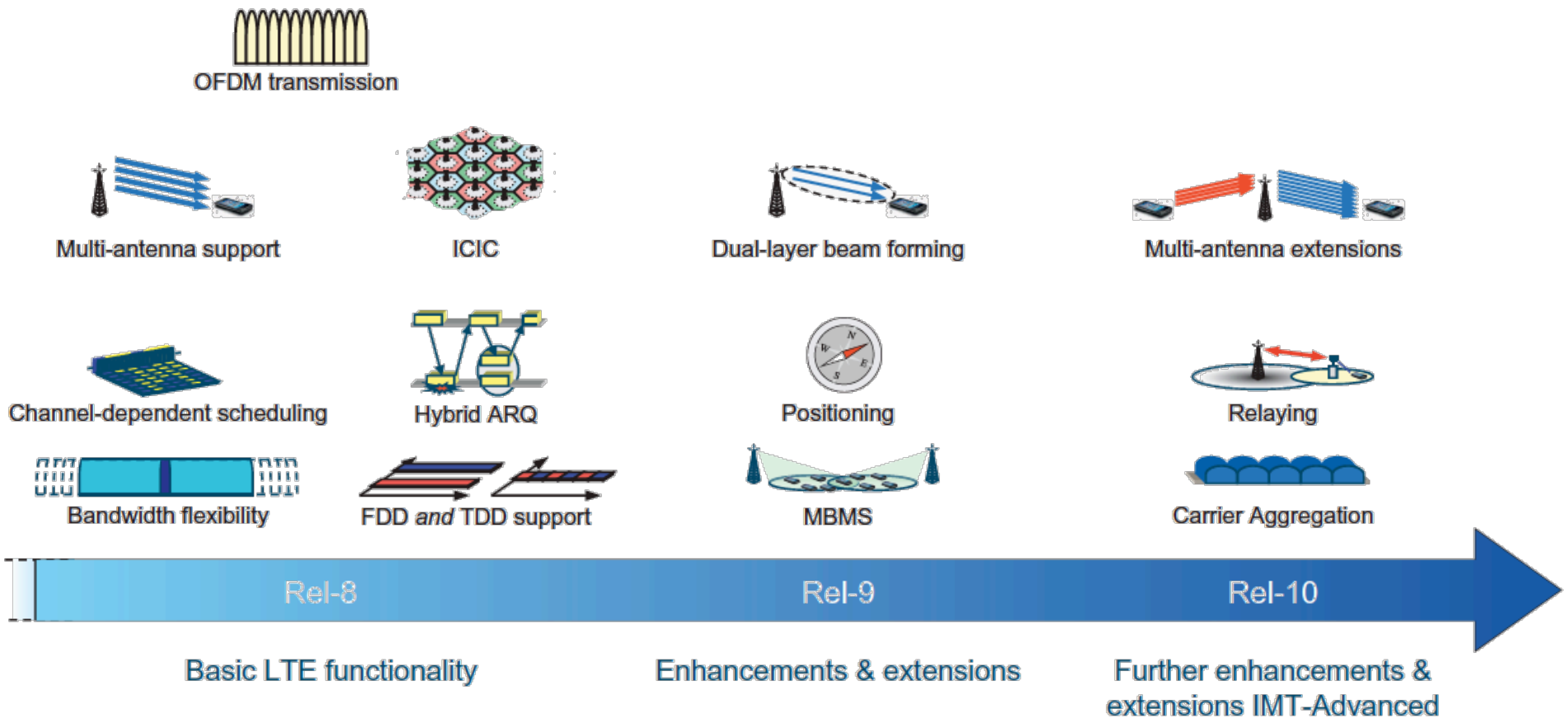
PART III: LTE CASE STUDY

LTE Case Study



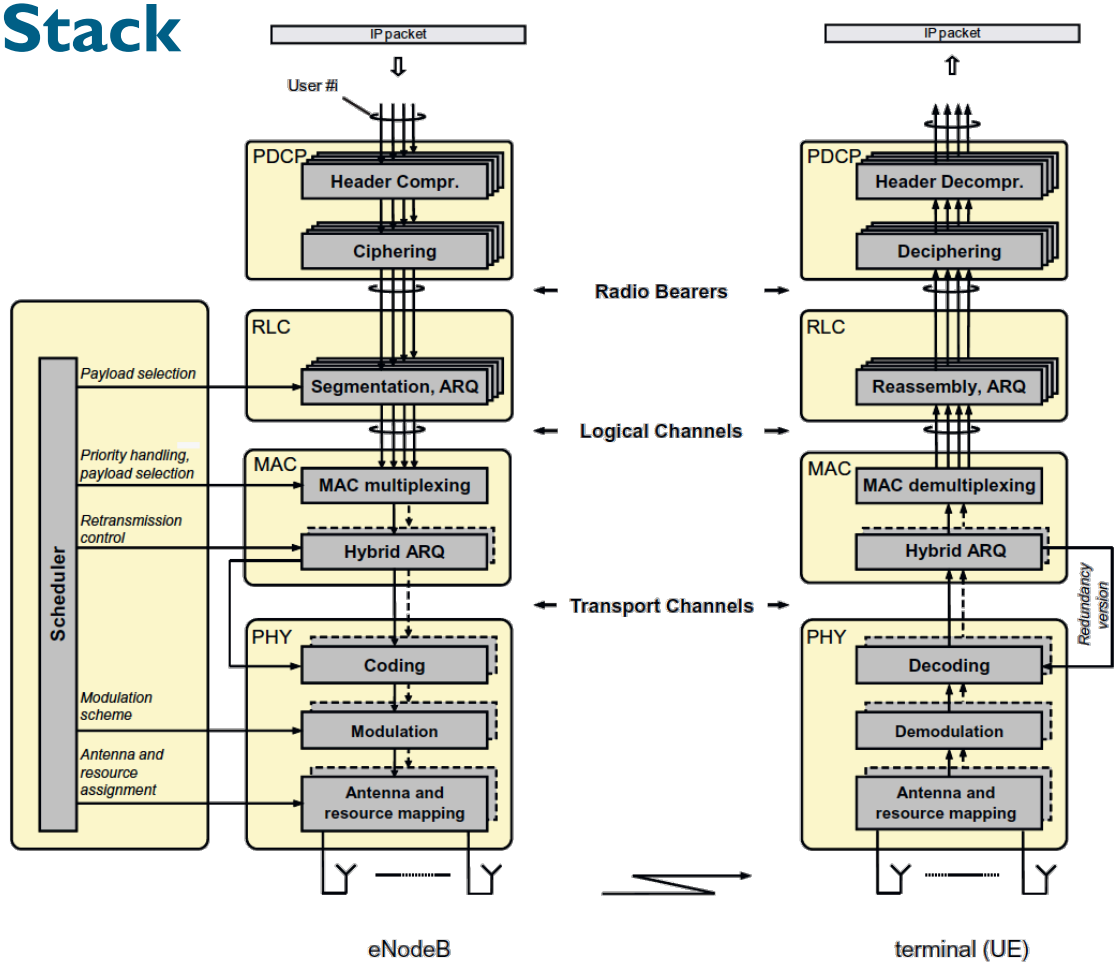
[1]. A. Checko et al, "Cloud RAN for Mobile Networks – A Technology Overview," IEEE Comm. Surveys & Tutorials, 1Q 2015.

LTE Releases



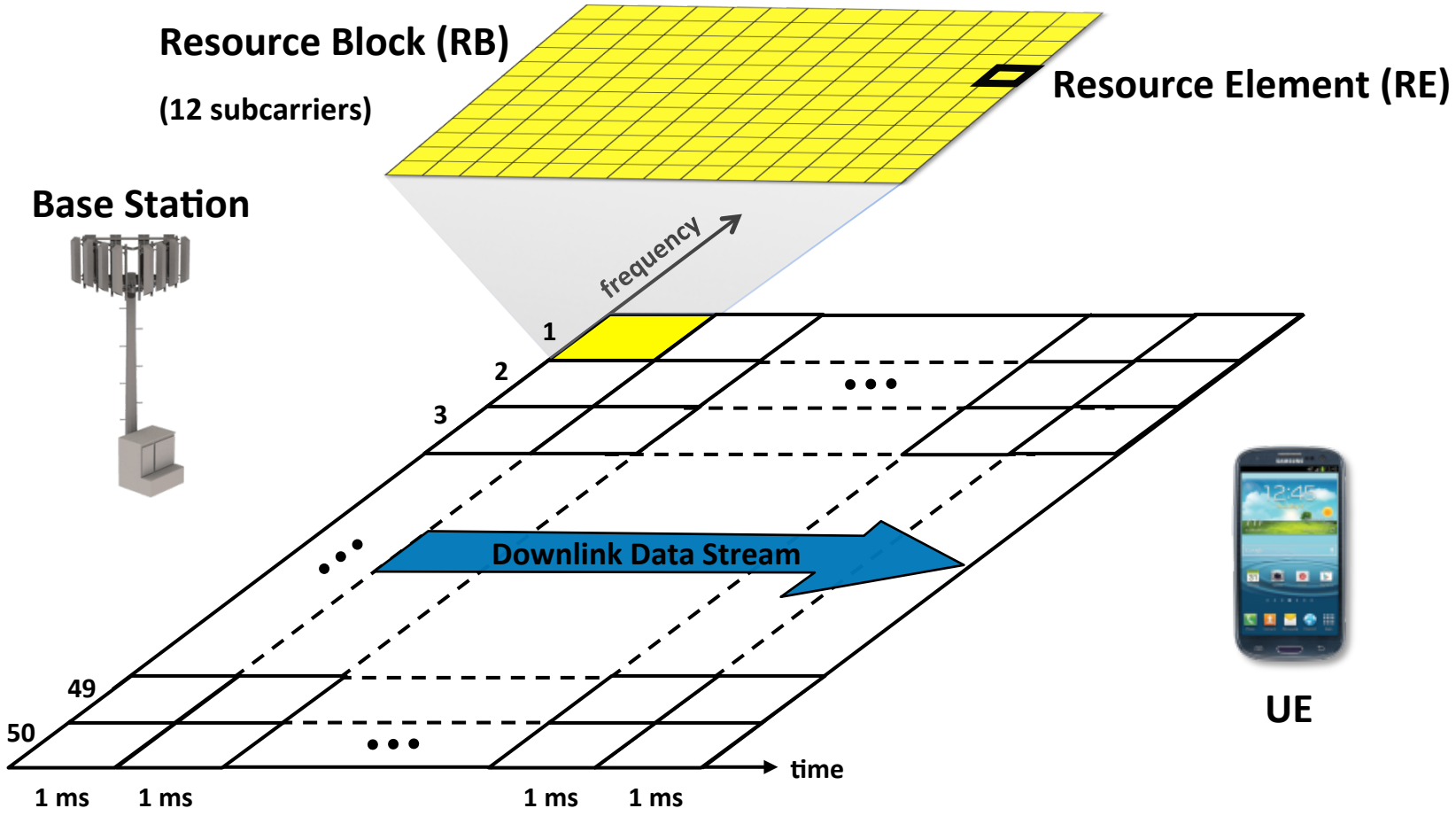
[7]. E. Dahlman, S. Parkvall, and J. Skold, *4G LTE / LTE-Advanced for Mobile Broadband*, second edition 2014.

LTE Protocol Stack

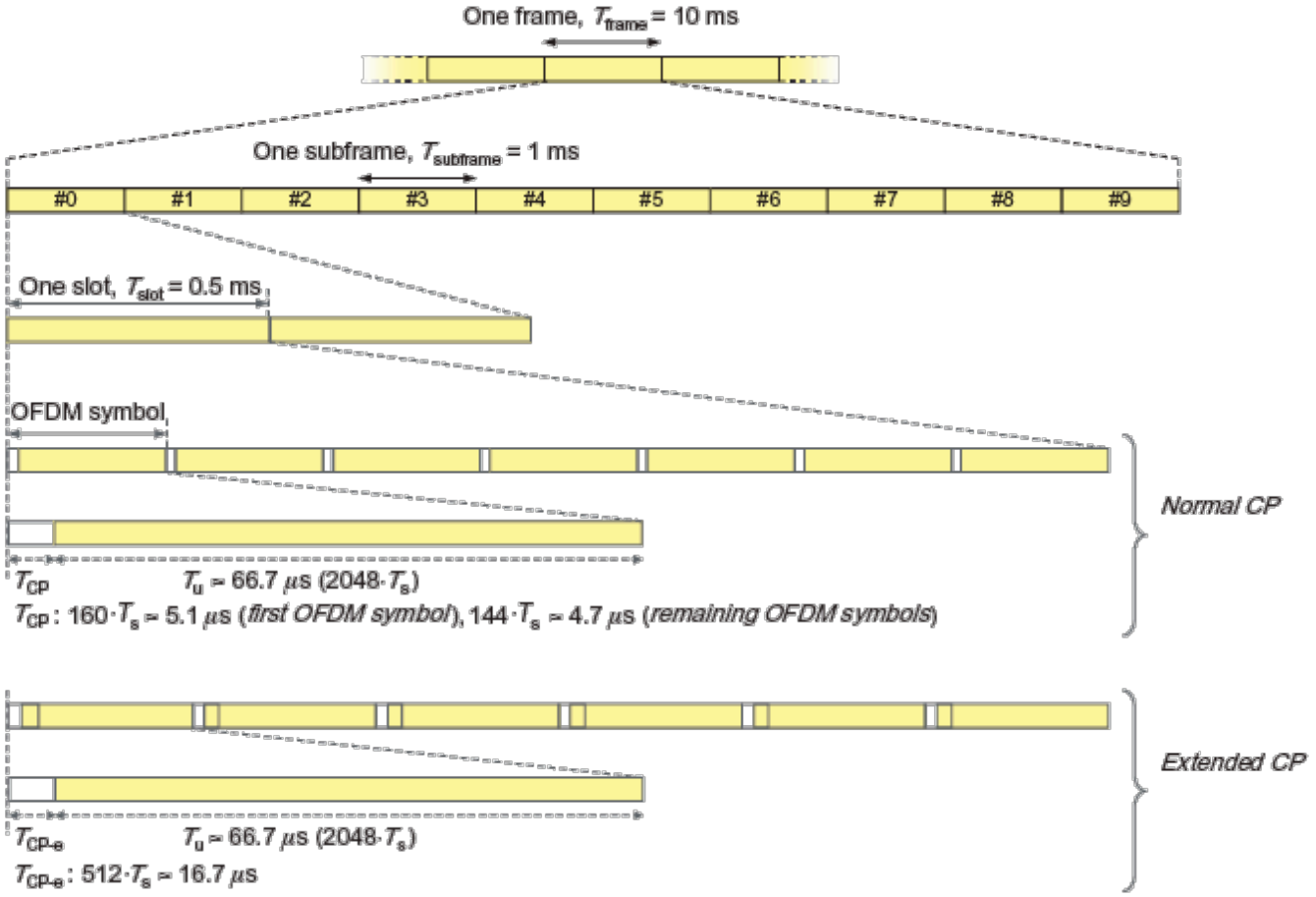


[7]. E. Dahlman, S. Parkvall, and J. Skold, *4G LTE / LTE-Advanced for Mobile Broadband*, second edition 2014.

LTE Time-Frequency Grid



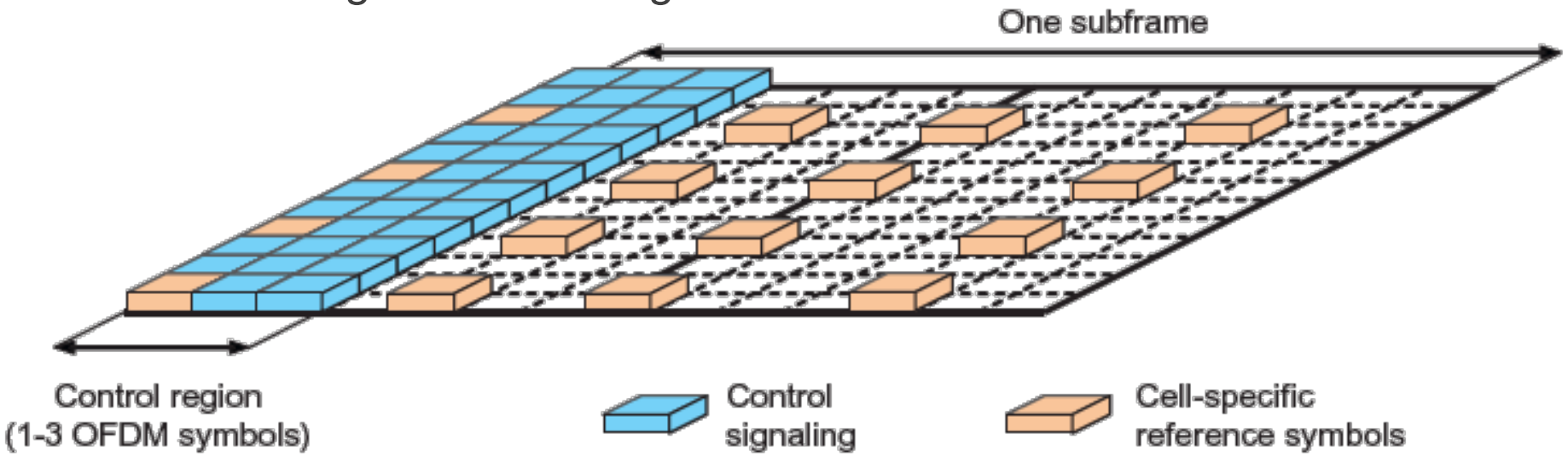
LTE Frame Structure



[7]. E. Dahlman, S. Parkvall, and J. Skold, *4G LTE / LTE-Advanced for Mobile Broadband*, second edition 2014.

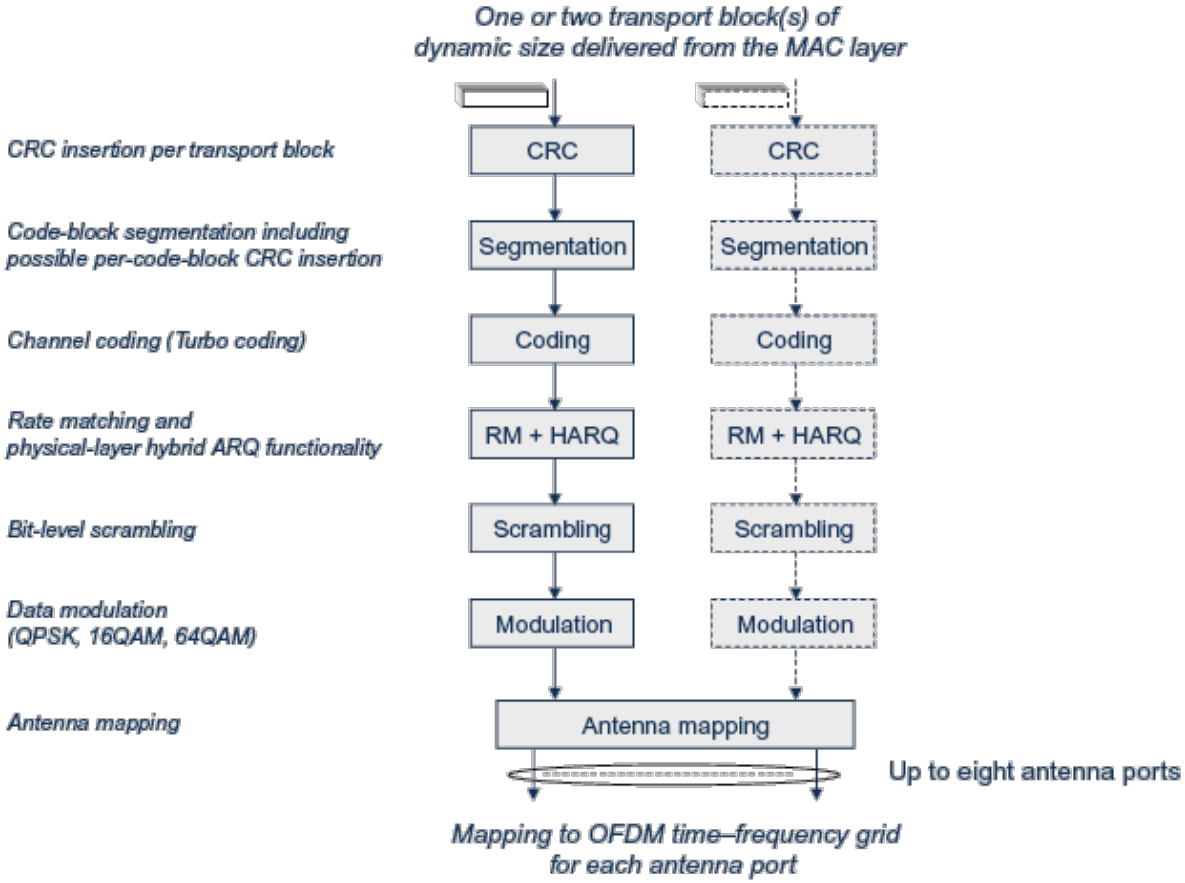
Downlink Arrangement

- Downlink subframe divided into control and data regions.
  - Control region is first 1 and 3 OFDM symbols of the subframe.
  - Physical Downlink Shared Channel (PDSCH) is in the region following the control region.



[7]. E. Dahlman, S. Parkvall, and J. Skold, *4G LTE / LTE-Advanced for Mobile Broadband*, second edition 2014.

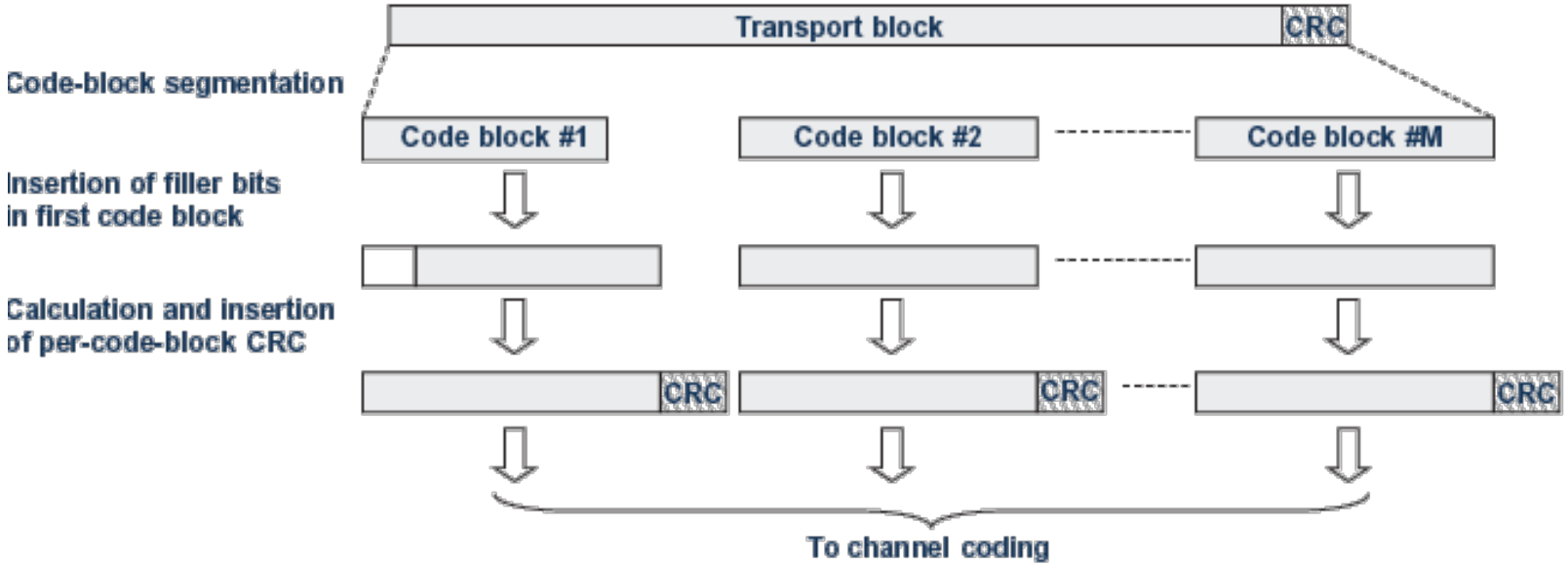
Baseband Processing for the Downlink



[7]. E. Dahlman, S. Parkvall, and J. Skold, *4G LTE / LTE-Advanced for Mobile Broadband*, second edition 2014.

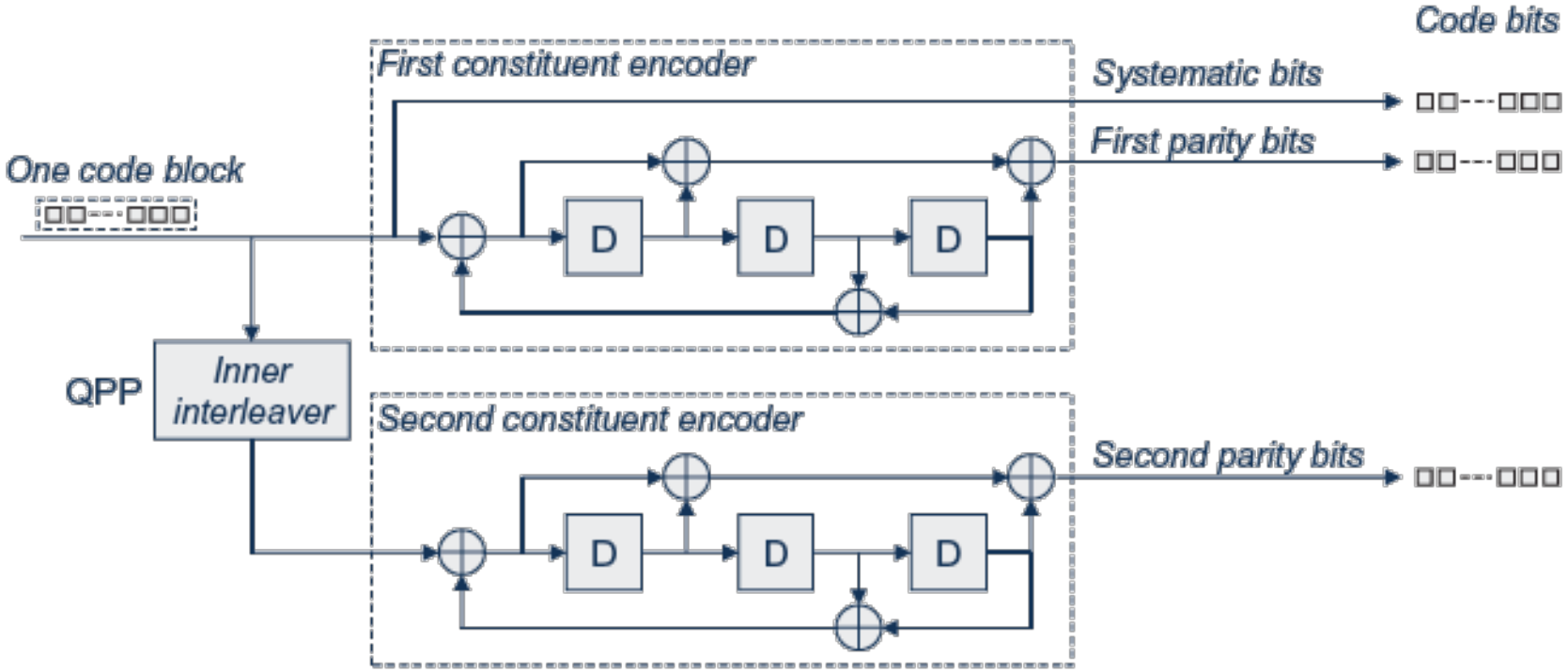
Code Block Segmentation

- Maximum code block size is 6144 bits
  - Maximum TB size is 75,376 bits.
  - May need to segment into as many as 13 code blocks.



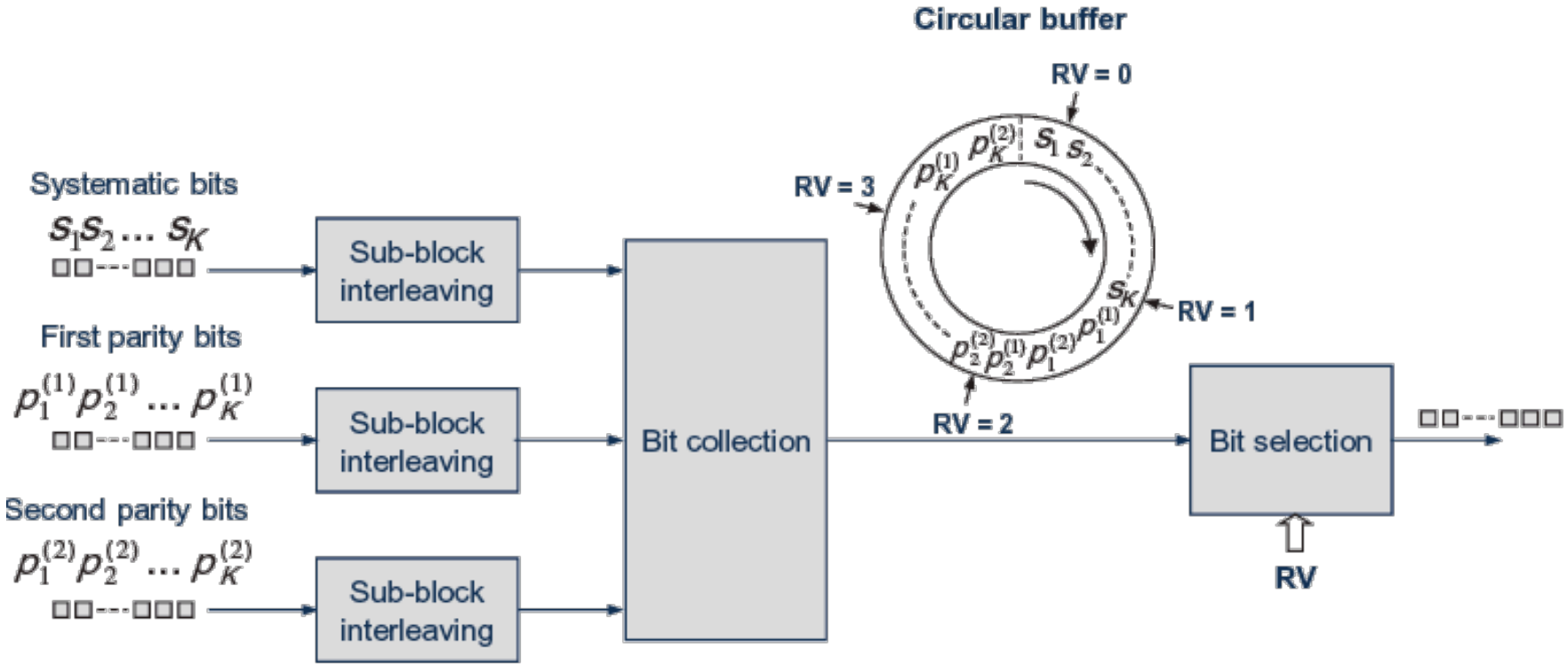
[7]. E. Dahlman, S. Parkvall, and J. Skold, *4G LTE / LTE-Advanced for Mobile Broadband*, second edition 2014.

Turbo Coding



[7]. E. Dahlman, S. Parkvall, and J. Skold, *4G LTE / LTE-Advanced for Mobile Broadband*, second edition 2014.

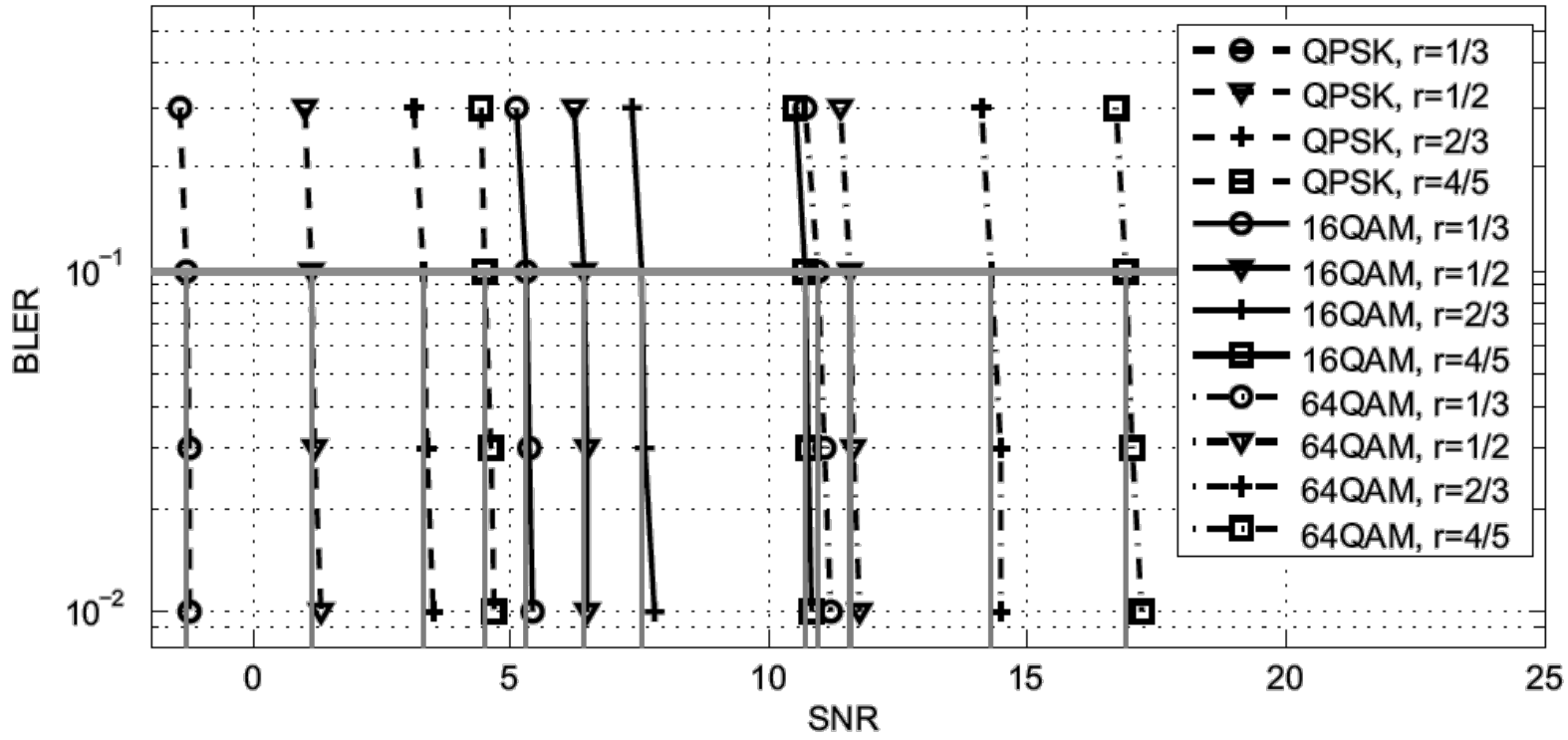
Rate Matching



[7]. E. Dahlman, S. Parkvall, and J. Skold, *4G LTE / LTE-Advanced for Mobile Broadband*, second edition 2014.

Adaptive Modulation and Coding

- The base station selects the highest Modulation and Coding Scheme (MCS) that can be supported at a 10% BLER

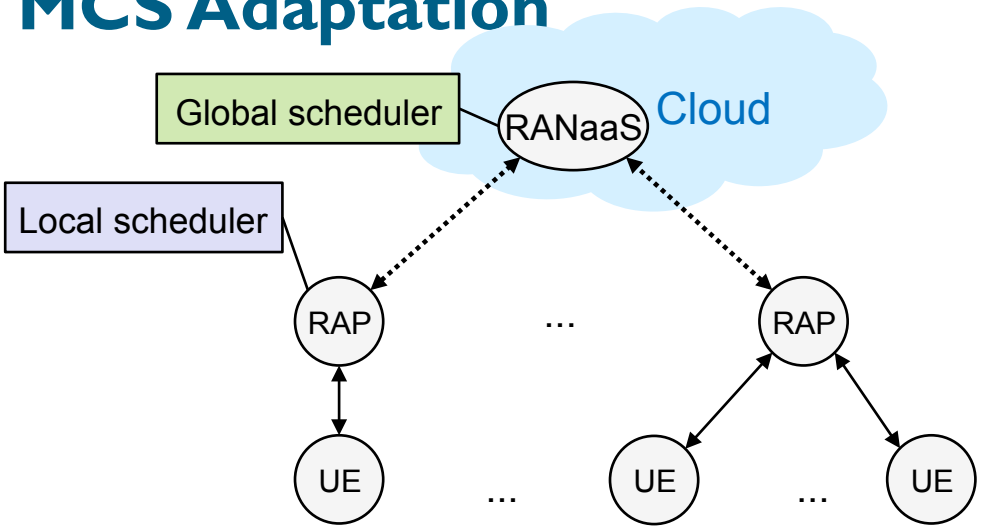


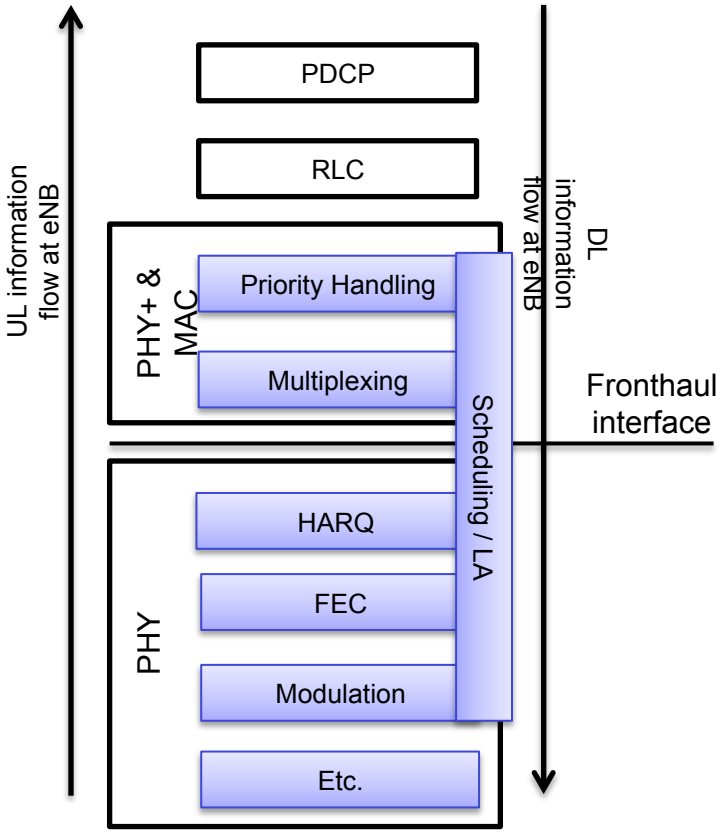
MCS Table

- The selected MCS is one of 27 options
- Indicated by a 5-bit quantity
- Here,
  - I_{mcs} = MCS index
  - Q_m = modulation order
  - C = number of code blocks
  - K_r = number data bits per block
  - η = spectral efficiency
  - An allocation of 45 RB is assumed

I_{mcs}	Q_m	C	K_r	η
0	2	1	1280	0.1938
1	2	1	1632	0.2481
2	2	1	2048	0.3123
3	2	1	2624	0.4012
4	2	1	3264	0.5000
5	2	1	4032	0.6185
6	2	1	4800	0.7370
7	2	1	5568	0.8556
8	2	2	3136	0.9568
9	2	2	3520	1.0753
10	2	2	4032	1.2333
11	4	2	4608	1.4111
12	4	2	5184	1.5889
13	4	2	5760	1.7667
14	4	3	4352	2.0000
15	4	3	4544	2.0889
16	4	3	4928	2.2667
17	4	3	5504	2.5333
18	4	3	5888	2.7111
19	4	4	4800	2.9444
20	4	4	5184	3.1815
21	6	4	5760	3.5370
22	6	5	4928	3.7802
23	6	5	5120	3.9284
24	6	5	5504	4.2247
25	6	5	5696	4.3728
26	6	6	5504	5.0704

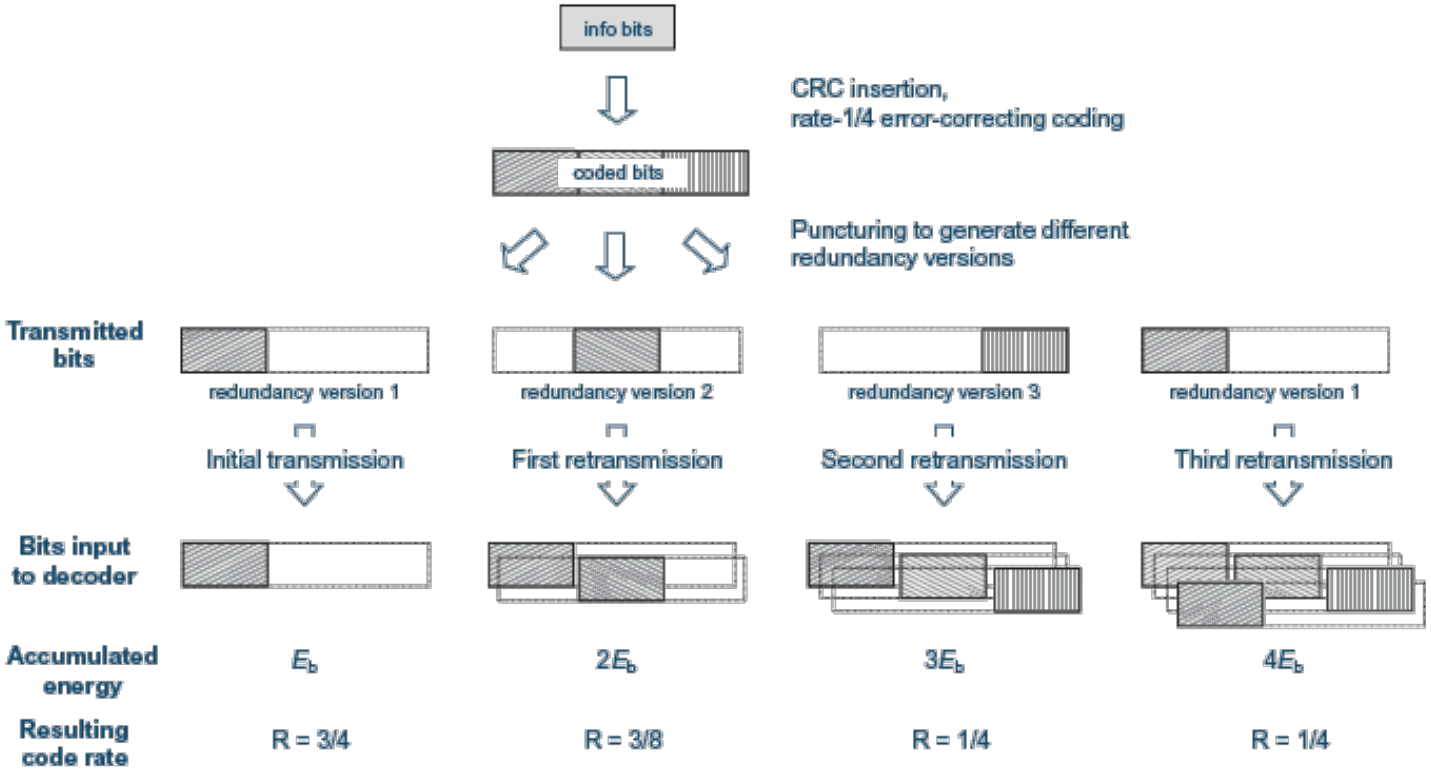
MCS Adaptation





- Limited capacity and latency affect link adaptation conditions
  - At the CRAN: outdated but global channel state information (CSI)
  - At the RAP: recent but local CSI

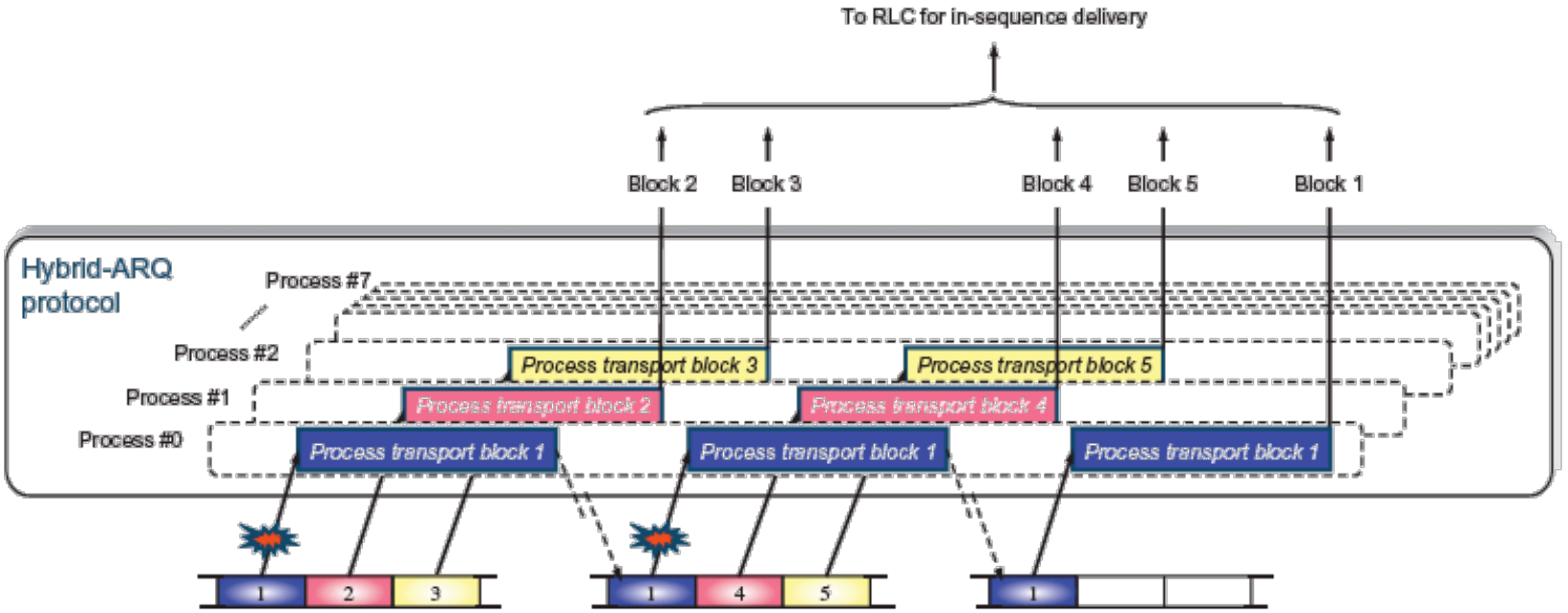
Hybrid-ARQ



[7]. E. Dahlman, S. Parkvall, and J. Skold, *4G LTE / LTE-Advanced for Mobile Broadband*, second edition 2014.

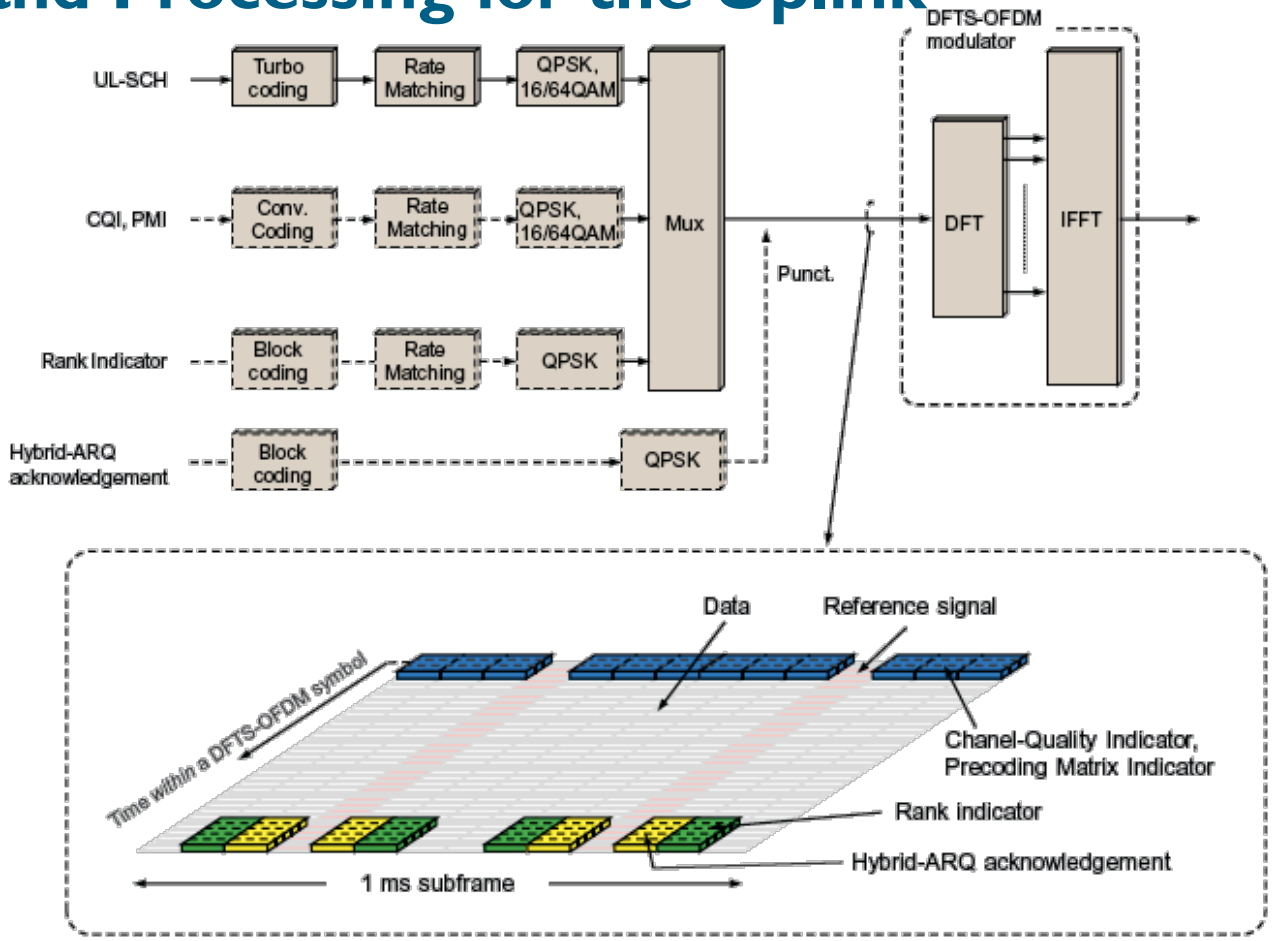
Multiple Parallel HARQ Processes

- 8 stop-and-wait processes run in parallel



[7]. E. Dahlman, S. Parkvall, and J. Skold, *4G LTE / LTE-Advanced for Mobile Broadband*, second edition 2014.

Baseband Processing for the Uplink



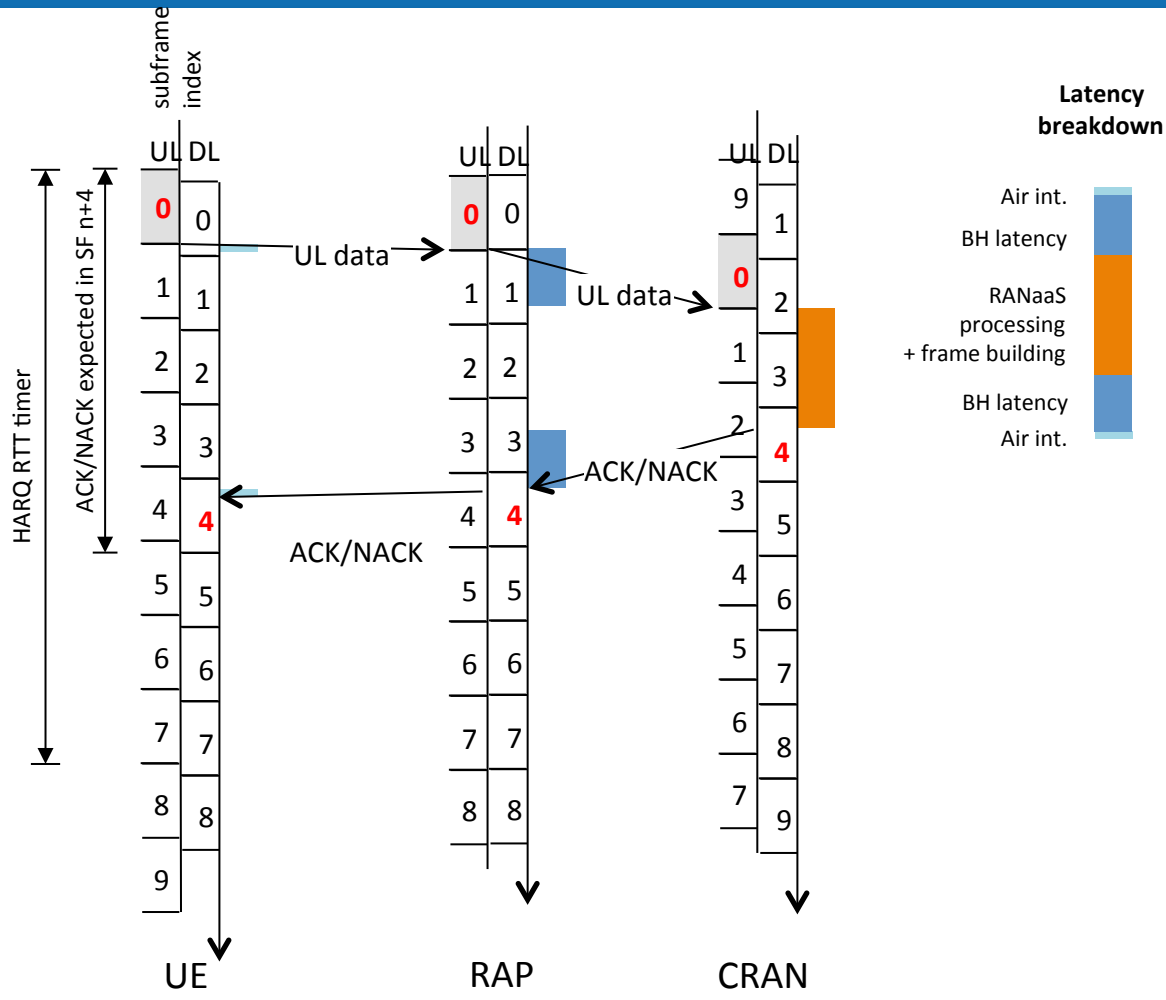
[7]. E. Dahlman, S. Parkvall, and J. Skold, *4G LTE / LTE-Advanced for Mobile Broadband*, second edition 2014.

HARQ Timing

- Downlink
  - HARQ are asynchronous
  - Retransmission can occur at any time
  - Must be flagged with a 3-bit HARQ process ID
- Uplink
  - Retransmissions are synchronous
  - Processes serviced in round-robin fashion: Sent once every 8 ms
  - Transmissions flagged with a 1-bit new data indicator
  - ACK/NACK indicator received within 4 ms
- Key uplink timing constraint:
  - *Uplink must be processed and downlink ACK sent within 4 ms.*

HARQ Timing

Added



RAN Timing Constraints

	Timer	Purpose	Min	Max	Default
PHY	Subframe	Physical subframe length	1 ms (fix)		
	Frame	Physical frame length	10 ms (fix)		
MAC	HARQ RTT Timer	When an HARQ process is available	In FDD: 8ms In TDD: FFS		
RLC	t-PollRetransmit	For AM RLC, poll for retransmission @tx side (if no status report received)	5ms	500ms	45ms
	t-Reordering	For UM/AM RLC, RLC PDU loss detection @rx side	0ms	200ms	35ms
	t-StatusProhibit	Prohibit generation of a status report @rx side	0ms	500ms	0ms
PDCP	discardTimer	At UE side in UL. Start at reception of PDCP SDU from upper layer. Discard PDCP SDU / PDU if expiration or successful transmission	50ms	Infinity	
RRC	T300	RRCConnectionRequest. If expire, reset MAC & signal RRC connection failure	100ms	2s	
	T301	RRCConnectionReestablishmentRequest If expire, go to RRC_IDLE	100ms	2s	
	T304	RRCConnectionReconfiguration	50ms/100ms	2s/8s	
	T310	Detection of physical problem (successive out-of-sync from lower layers)	0ms	2s	1s
	T311	RRC connection reestablishment (E-UTRA or another RAT).	1s	3s	1s

Turbo Decoding is Iterative

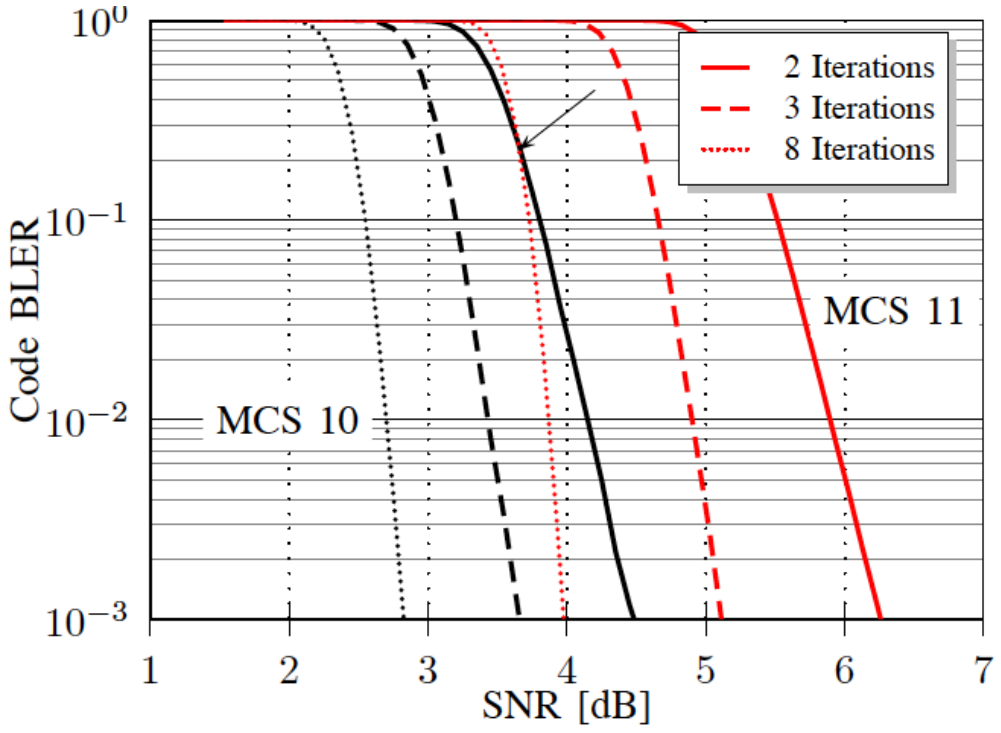


Fig. 2. Code block error rate as a function of SNR for MCS 10 and MCS 11 after 2, 3, and 8 decoder iterations. The arrow shows where the CBLER of MCS 10 after 2 iterations is the same as that of MCS 11 after 8 iterations.

- On the uplink, around 50% of compute load is due to turbo decoding.
- Because a CRC is used to halt decoding, the load is directly proportional to the number of iterations
- Operating with a higher SNR margin reduces the number of iterations

[8]. M.C. Valenti, S. Talarico, and P. Rost, "The role of computational outage in dense cloud-based centralized radio access networks," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, (Austin, TX), Dec. 2014.

Computational Load for Turbo Decoding

- The load to decode a given transport block is:

$$\mathcal{C}(\gamma) \quad = \quad \sum_{r=1}^{C} K_r I_r$$

- Where:
  - Load depends on SINR γ and the selected MCS
  - C is the number of code blocks after segmentation
  - K_r is the number of information bits in the rth code block
  - I_r is the number of decoding iterations for the rth code block
- Load is in units of bit-iterations
  - Relation between bit-iterations and CPU cycles is implementation dependent, but fixed for a given architecture

Computational Outage

- If a transport block is not decoded before the deadline, then a computational outage occurs
- From a systems perspective, a computational outage is no different than any other kind of outage (e.g., due to fading or interference)
- For a conventional (locally processed / non-pooled) system, a computational outage occurs when the following condition occurs

$$C(\gamma) > C_{max}$$

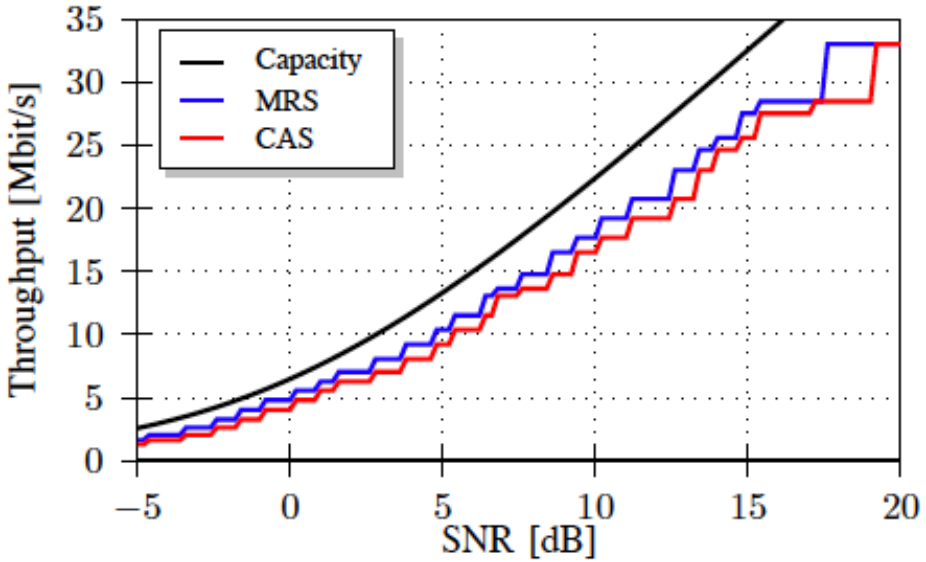
  where C_{max} is the maximum number of bit–iterations that can be supported within the deadline
- The computational outage probability is the probability of this event

A Few Details About Computational Outage

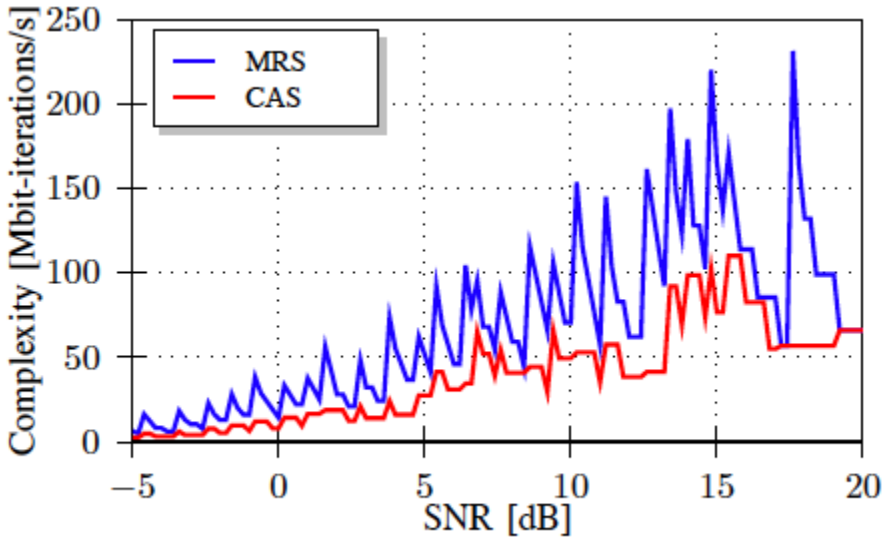
- The deadline must account for
  - The time from receiving the uplink transmission until when the HARQ ACK/NACK signal must be sent. 3 msec

  - Delays in the fronthaul: 0.1 msec each way
  - Other processing delays
- C_{max} should characterize
  - The amount of computing resources available for turbo decoding, which has the most variability.
  - Other operations (typically of fixed complexity) should not be incorporated into C_{max}

Scheduling Policy Influences the Load



(a) Raw throughput



(b) Computational effort

- MRS = max-rate scheduling
  - Target 10^{-1} BLER after 8 iterations
- CAS = computationally aware scheduling
  - Target 10^{-1} BLER after just 2 iterations

[8]. M.C. Valenti, S. Talarico, and P. Rost, "The role of computational outage in dense cloud-based centralized radio access networks," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, (Austin, TX), Dec. 2014.

Conservative Scheduling Helps if Compute Limited

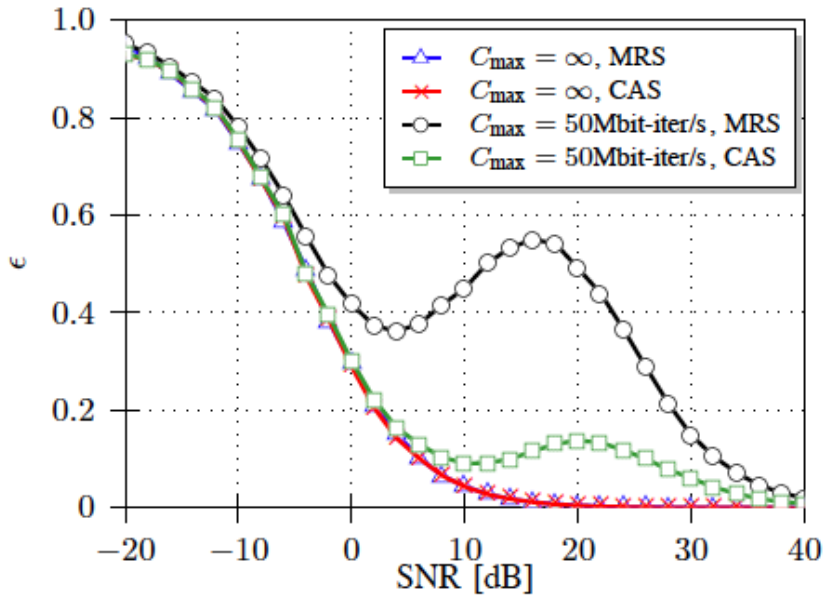


Fig. 4. Outage probability as a function of average SNR in the presence of Rayleigh fading, both with and without a complexity constraint.

- Comparison
  - Unlimited compute power
  - Compute limited
- Channel model
  - Block Rayleigh fading
  - Perfect T-CSI
  - No interference
- Outages can be due to channel or compute effects
- Outage probability with CAS is much lower when compute limited

[8]. M.C. Valenti, S. Talarico, and P. Rost, "The role of computational outage in dense cloud-based centralized radio access networks," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, (Austin, TX), Dec. 2014.

Influence on Throughput

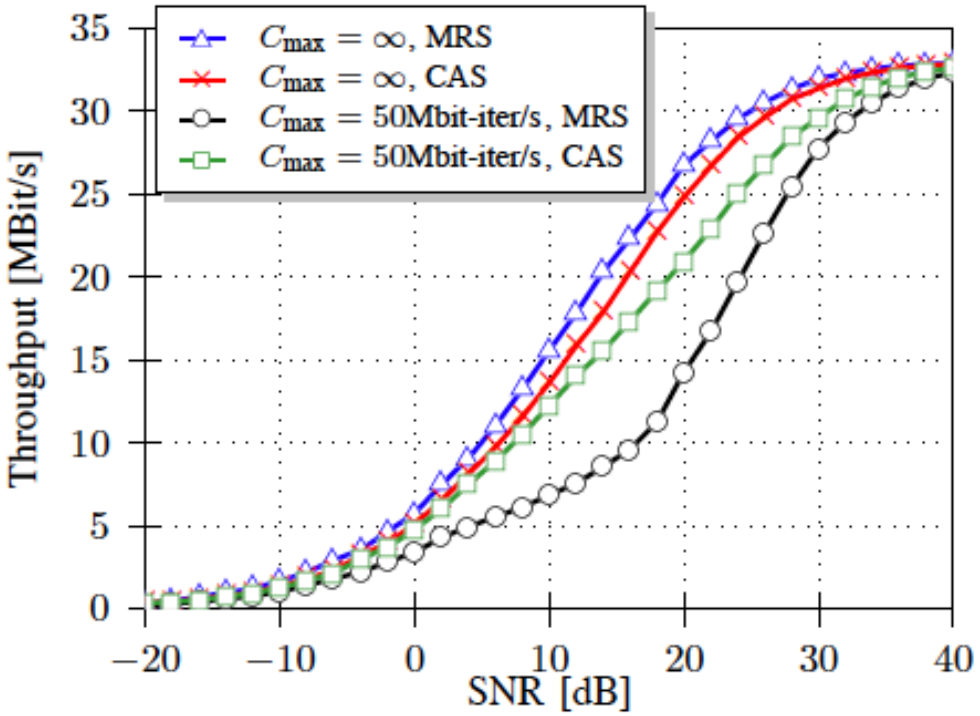


Fig. 5. Effective throughput as a function of average SNR in the presence of Rayleigh fading, both with and without a complexity constraint.

- Throughput is the rate of *correct* data transfer
- Even though CAS has a lower peak rate, its throughput is better due to reduced occurence of compute outage

[8]. M.C. Valenti, S. Talarico, and P. Rost, "The role of computational outage in dense cloud-based centralized radio access networks," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, (Austin, TX), Dec. 2014.

Influence on the Required Complexity

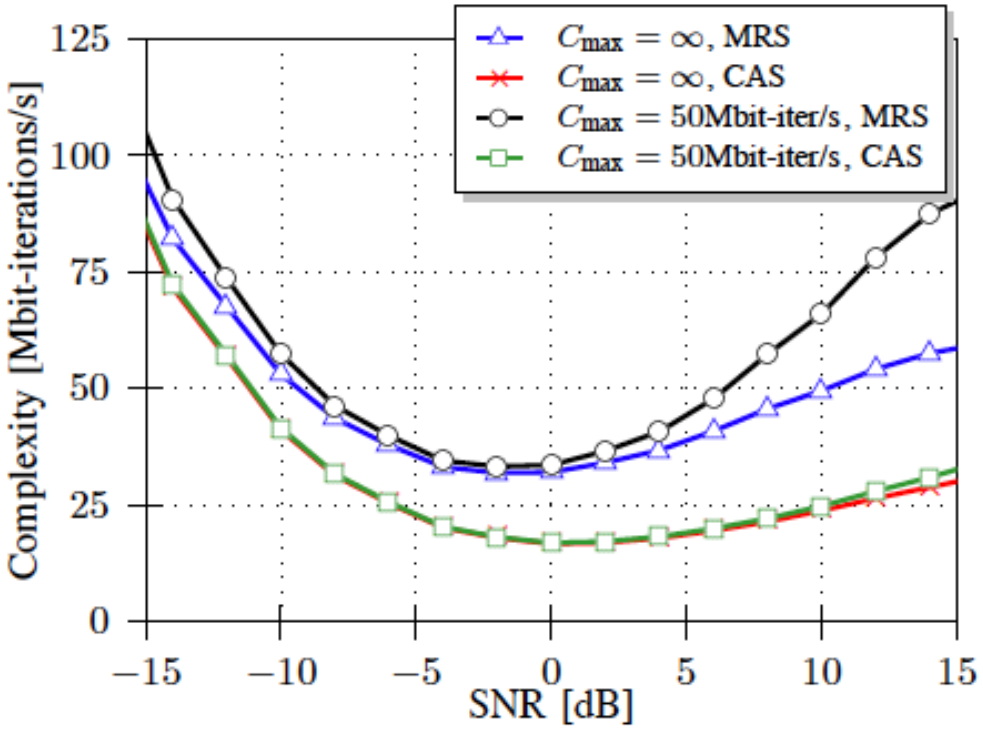


Fig. 6. Average complexity required in Rayleigh fading for a successfully decoded codeword as function of average SNR.

- Complexity required to correctly decode a transmission
- MRS can be computationally costly because each retransmission must be processed

[8]. M.C. Valenti, S. Talarico, and P. Rost, "The role of computational outage in dense cloud-based centralized radio access networks," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, (Austin, TX), Dec. 2014.

Computational Outage in a C-RAN Environment

- Computational resources are shared by the pool
- Let N_{cloud} be the number of RRH serviced by the pool
- A computational outage occurs when

$$\sum_{i=1}^{N_{\text{cloud}}} \mathcal{C}(\gamma_i) > N_{\text{cloud}} \cdot \mathcal{C}_{\text{max}}$$

- where γ_i is the SINR at the ith RRH and C_{max} is the available computing per RRH
- By exploiting the statistical multiplexing gain, it may be possible to reduce C_{max} --- but by how much?

Role of Interference

- The uplink SINR in a multi-cell network is

$$\gamma_j = \frac{P_j g_{j,j} |Y_j - X_j|^{-\alpha}}{W + \sum_{i \neq j} P_i g_{i,j} |Y_j - X_i|^{-\alpha}}$$

$$= \frac{g_{j,j} |Y_j - X_j|^{\alpha(s-1)}}{\Gamma^{-1} + \sum_{i \neq j} g_{i,j} |Y_j - X_i|^{-\alpha} |Y_i - X_i|^{s\alpha}}$$

- Where
  - Y_j is the jth RRH and X_j is the mobile served by it
  - $g_{i,j}$ is the fading gain between X_i and Y_j
  - α is the path-loss exponent
  - s is partial-power control compensation factor (s=1 for full PC)
  - Γ is the SNR at the RRH

Local Processing vs. Centralized Processing

- Example scenario
  - N = 129 base stations (actual locations from UK)
  - Ncloud = 8 in the center are considered
  - Can be processed centrally (CP) or locally (LP)
- Simulation parameters
  - Mobile devices placed according to a Poisson Point Process (PPP)
  - Density λ devices per km^2
  - Just one device serviced per cell (TDMA scheduling)
  - $\alpha = 3.7$ and $s=0.1$
  - $\Gamma = 20$ dB

Sum Throughput as a Function of Compute Power

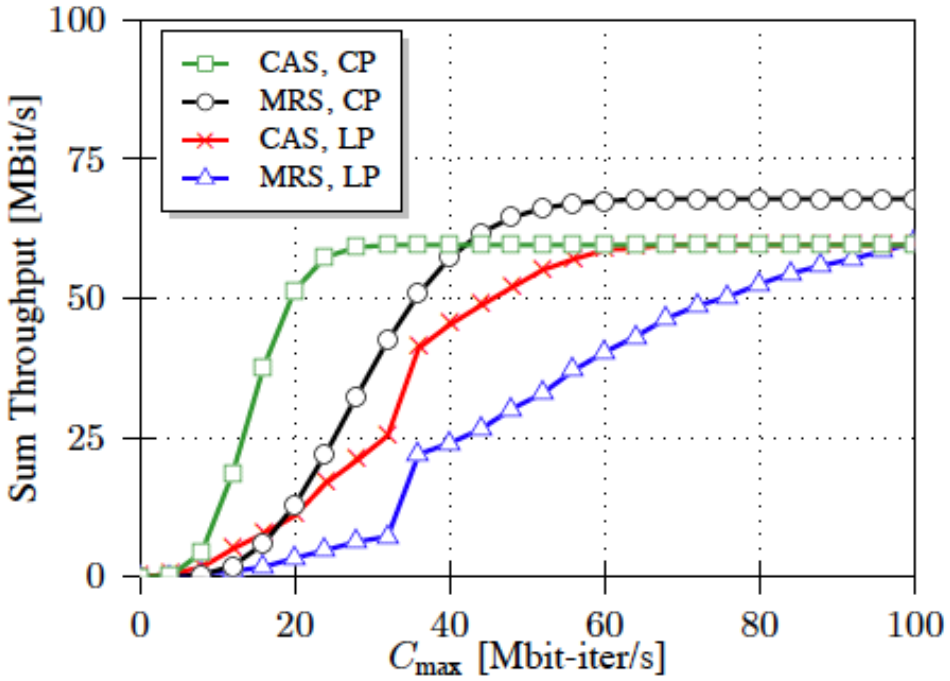


Fig. 8. Sum throughput as function of the per-RAP complexity constraint when $N_{\text{cloud}} = 8$, with local processing (LP) and with cloud processing (CP). Two MCS-selection schemes are considered: computationally aware selection (CAS) and max-rate selection (MRS).

- Fixed density of mobiles

  $\lambda = 0.1$ per km^2
- Central Processing always outperforms Local Processing
- CAS scheduling better than MRS when compute resources are constrained

[8]. M.C. Valenti, S. Talarico, and P. Rost, "The role of computational outage in dense cloud-based centralized radio access networks," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, (Austin, TX), Dec. 2014.

Effect of Mobile Density

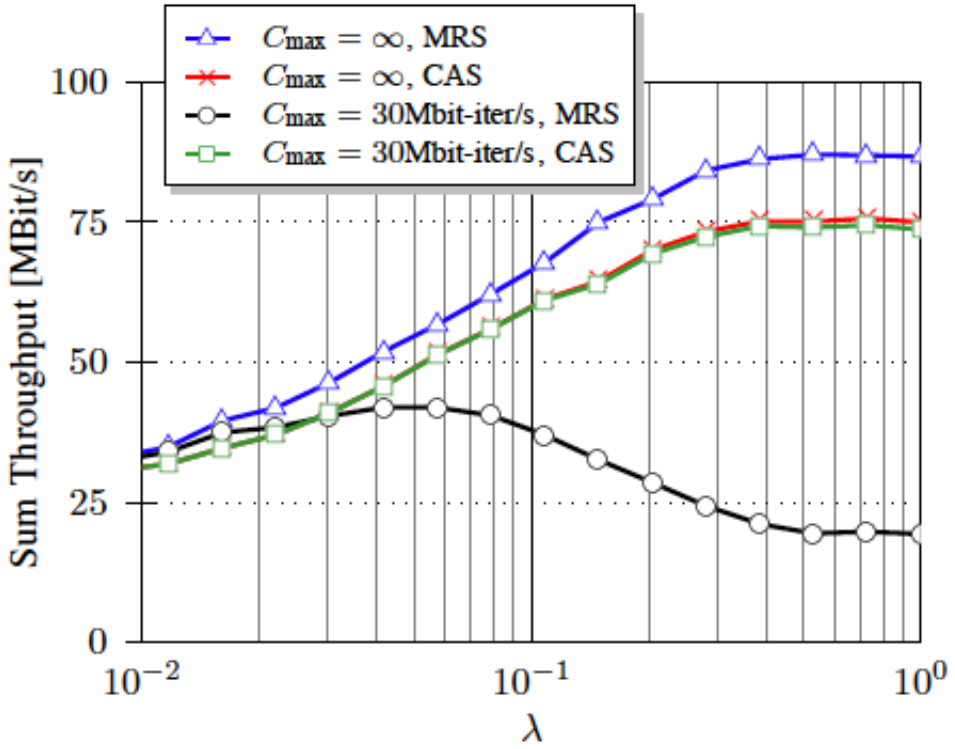


Fig. 9. Sum throughput with cloud processing as function of the density of UEs when $N_{\text{cloud}} = 8$. Two MCS-selection schemes are considered: computationally aware selection (CAS) and max-rate selection (MRS).

- Centrally processed
- Variable density of mobile devices
- When constrained, MRS degrades with increasing user density

[8]. M.C. Valenti, S. Talarico, and P. Rost, "The role of computational outage in dense cloud-based centralized radio access networks," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, (Austin, TX), Dec. 2014.

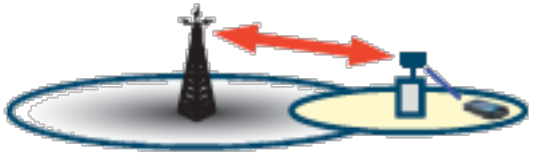
Considerations for LTE-Advanced

Carrier Aggregation

- Can combine up to 5 carriers to achieve 100 MHz of useful bandwidth.

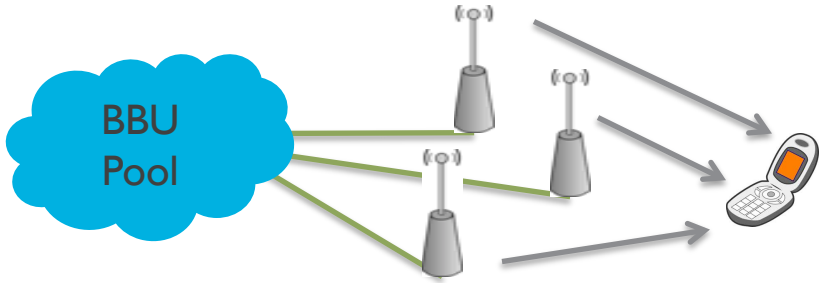
HetNets & Relaying

- Better support for small cells.
- Support for relays.



Coordinated Multipoint

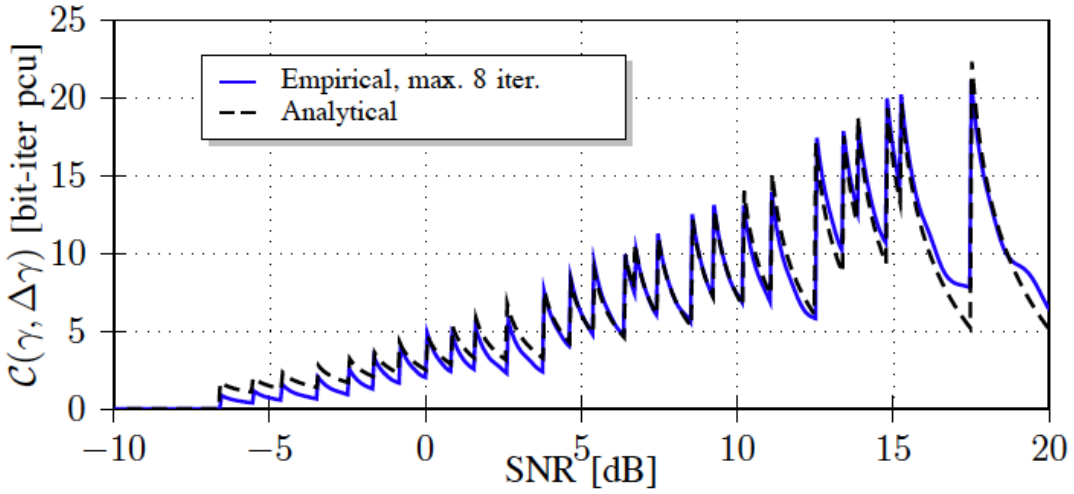
- Neighboring base stations simultaneously transmit to the same UE.



Enhanced Spatial Multiplexing

- Up to 8 layer MIMO on the downlink.
- Up to 4 layer MIMO on the uplink.

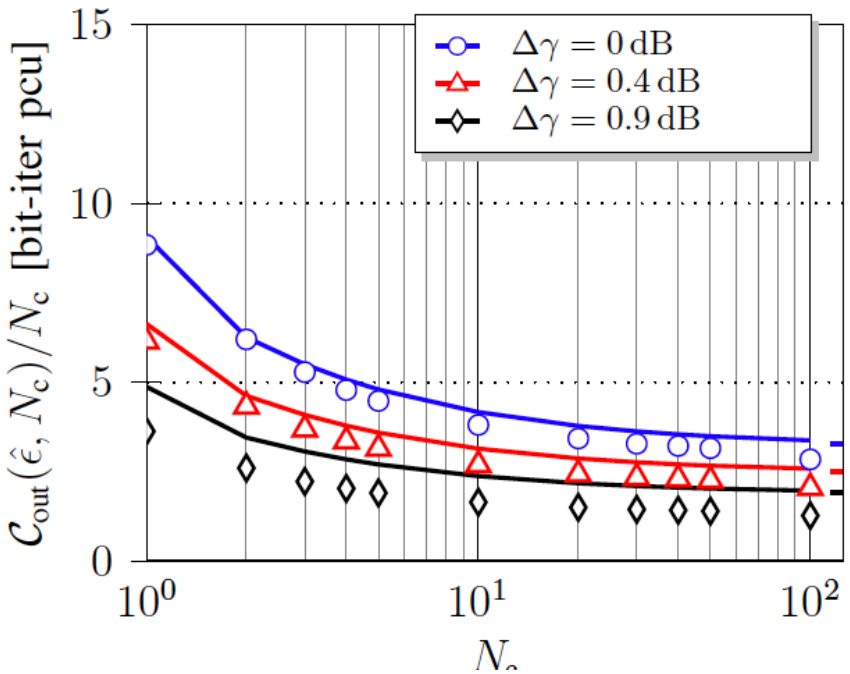
Towards a Theory for Computational Outage



- The complexity of decoding can be modeled statistically
- Similar to modeling the channel statistically
- By using the statistical model, analytical insight can be obtained without resorting to simulation

[9]. P. Rost, S. Talarico, and M.C. Valenti , "The complexity-rate tradeoff of centralized radio access networks," *IEEE Transactions on Wireless Communications,* vol. 14, no. 11, pp. 6164-6176, Nov. 2015.

Outage Complexity



- Outage complexity is the amount of computing power required to achieve a desired computational outage probability
- Analogous to outage capacity
- Useful to plot as a function of the cloud group size N_{cloud}
- Can be used to rapidly determine compute power needed

Fig. 5. Outage complexity to ensure per-cell outage constraint $\hat{\epsilon} = 0.1$ as function of the number of RAPs, whose signals are centrally processed. Solid lines are evaluated analytically, while dots are obtained through simulations using one million trials. The notches on the right side of each sub-figure show the behavior as $N_c \to \infty$.

[9]. P. Rost, S. Talarico, and M.C. Valenti , "The complexity-rate tradeoff of centralized radio access networks," *IEEE Transactions on Wireless Communications,* vol. 14, no. 11, pp. 6164-6176, Nov. 2015.

Optimal Scheduling

- The scheduling problem can be formulated as

$$\mathcal{R}_{\text{opt}} = \arg\max_{\mathcal{R}\in\mathcal{R}'} \sum_{r_k\in\mathcal{R}} r_k,$$
$$\text{s.t.} \sum_{r_k\in\mathcal{R}} \mathcal{C}_k \leq \mathcal{C}_{\text{server}}.$$

  where r_k is the rate of the kth user, C_k is its offered computational load, and C_{server} is the total available computing power

- Optimal solution results in a water-filling algorithm
- A heuristic alternative solution is to simply pick the user with highest complexity and back off its rate until the complexity constraint is satisfied

Benefits of Optimal Scheduling

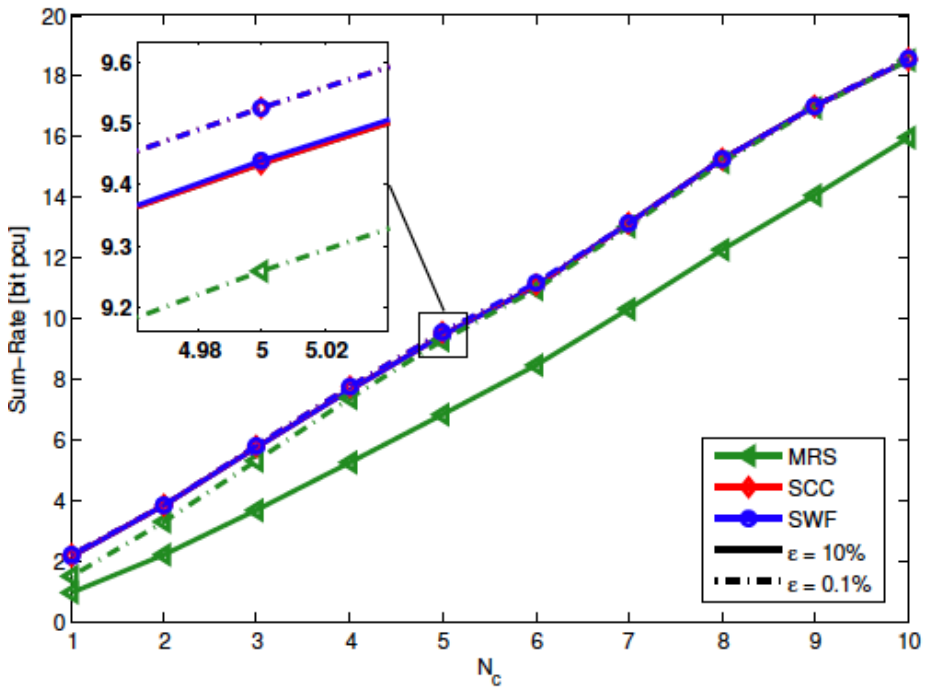
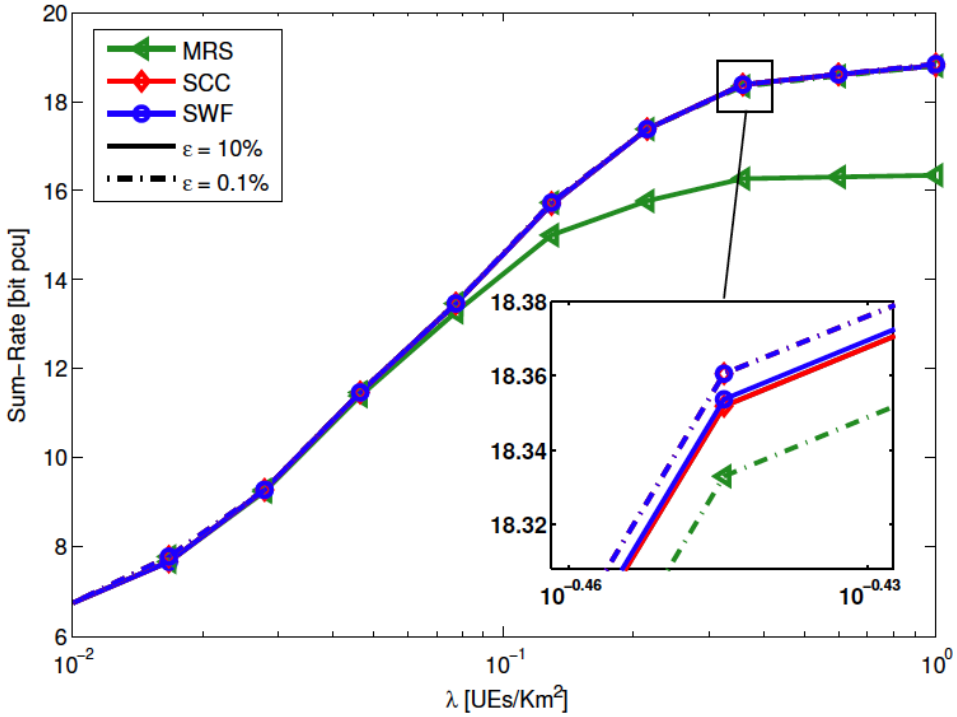


Fig. 4. Sum-rate as a function of N_c.

- Three approaches
  - max-rate scheduling (MRS)
  - scheduling with complexity cutoff (SCC)
  - scheduling with water filling (SWF)
- Cellular network
  - 129 actual base stations
  - λ = 1 device per km^2
- Variable N_{cloud}

[10] P. Rost, A. Maeder, M.C. Valenti, and S. Talarico, "Computationally aware sum-rate optimal scheduling for centralized radio access networks," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, (San Diego, CA), Dec. 2015.

Scheduling More Important As Network Densifies



- Fix N_{cloud} = 10
- Vary the user density
- Optimal scheduling provides 20% higher throughput for highly dense networks

[10] P. Rost, A. Maeder, M.C. Valenti, and S. Talarico, "Computationally aware sum-rate optimal scheduling for centralized radio access networks," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, (San Diego, CA), Dec. 2015.

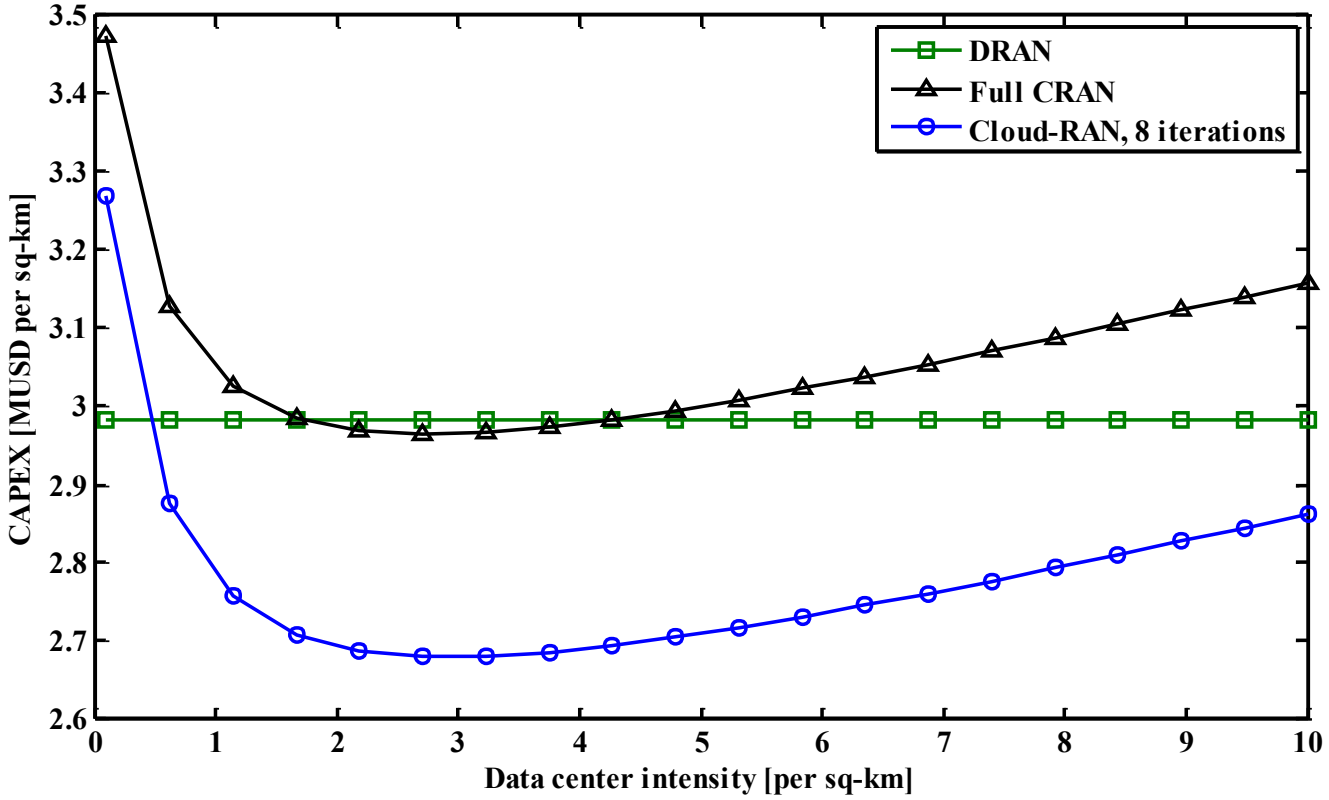
The Economics of C-RAN

- The costs of C-RAN are determined by
  - Cost per RRH
  - Cost per BBU server
  - Cost for fronthaul / backhaul (per km)

Type of cost	DRAN	Cloud-RAN
Macro base station	$50k	$25k
Micro base station	$20k	$10k
Microwave BH	$50k per link plus $5k per kilometer	
Optical fiber BH	$5k per link plus $100k per kilometer	
Data center		$40k
Server blades		$20k each

[11] P. Rost, I. Berberana, A. Maeder, H. Paul, V. Suryaprakash, M.C. Valenti, D. Wubben, A. Dekorsy, and G. Fettweis, "Benefits and challenges of virtualization in 5G radio access networks," *IEEE Communications Magazine*, vol. 53, no. 12, Communications Standards Supplement, pp. 75-82, Dec. 2015.

Economic Analysis: An Example



[11] P. Rost, I. Berberana, A. Maeder, H. Paul, V. Suryaprakash, M.C. Valenti, D. Wubben, A. Dekorsy, and G. Fettweis, "Benefits and challenges of virtualization in 5G radio access networks," *IEEE Communications Magazine*, vol. 53, no. 12, Communications Standards Supplement, pp. 75-82, Dec. 2015.

Industry/research interest

- NGMN
- Small Cell Forum
- Several EU projects
  - Mobile Cloud Networking (MCN)
  - IJOIN
  - HARP
  - Any many others
- Major equipment vendors
- Ongoing prototyping and field trials by operators

Concluding remarks

- C-RAN is coming
  - It can be implemented with current technology
  - Trials are underway, products to follow
- C-RAN is beneficial
  - Creates opportunities for enhanced collaboration
- C-RAN is green
  - Reduces the computational load
  - Opportunities to selectively turn off sites
- C-RAN is challenging
  - Trades benefit on the wireless side for increased demands on the wired fronthaul

QUESTIONS?

Cited References

[1] A. Checko, H.L. Christiansen, Y. Yan, L. Scolari, G. Kardaras, M.S. Berger, and L. Dittmann,, "Cloud RAN for Mobile Networks – A Technology Overview," IEEE Comm. Surveys & Tutorials, vol. 17, no. 1, pp. 405-426, First Quarter 2015.

[2] China Mobile Research Institute, "C-RAN: The Road Towards Green RAN," White Paper, 2010.

[3] A. A. M. Saleh, A. J. Rustako and R. S. Roman, "Distributed Antennas for Indoor Radio Communications," IEEE Transactions on Commun., vol. 35, pp. 1245-1251, Dec. 1987.

[4] J. Andrews, "Seven ways HeTNets are a cellular paradigm shift," IEEE Communications Magazine, 2013.

[5] X. Wang et al, "Cache in the Air: Exploiting Content Caching and Delivery Techniques for 5G" IEEE Comm. Mag., 2014.

[6] S. Talarico and M.C. Valenti, "An accurate and efficient analysis of an MBSFN network, in Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP), (Florence, Italy), May 2014.

[7] H. Paul, B.-S. Shin, D. Wübben, A. Dekorsy: In-Network-Processing for Small Cell Cooperation in Dense Networks, CLEEN 2013, Las Vegas, USA, September 2013.

[8] M.C. Valenti, S. Talarico, and P. Rost, "The role of computational outage in dense cloud-based centralized radio access networks," in Proc. IEEE Global Commun. Conf. (GLOBECOM), (Austin, TX), Dec. 2014.

[9] P. Rost, S. Talarico, and M.C. Valenti , "The complexity-rate tradeoff of centralized radio access networks," IEEE Transactions on Wireless Communications, vol. 14, no. 11, pp. 6164-6176, Nov. 2015.

[10] P. Rost, A. Maeder, M.C. Valenti, and S. Talarico, "Computationally aware sum-rate optimal scheduling for centralized radio access networks," in Proc. IEEE Global Commun. Conf. (GLOBECOM), (San Diego, CA), Dec. 2015.

[11] P. Rost, I. Berberana, A. Maeder, H. Paul, V. Suryaprakash, M.C. Valenti, D. Wubben, A. Dekorsy, and G. Fettweis, "Benefits and challenges of virtualization in 5G radio access networks," IEEE Communications Magazine, vol. 53, no. 12, Communications Standards Supplement, pp. 75-82, Dec. 2015.

Additional References

[12] P. Rost, C.J. Bernardos, A. De Domenico, M. Di Girolamo, M. Lalam, A. Maeder, D. Sabella, D. Wübben: *Cloud Technologies for Flexible 5G Radio Access Networks*, 5G Special Issue of IEEE Communications Magazine, Vol. 52, No. 5, pp. 68-76, May 2014.

[13] D. Wübben, P. Rost, J. Bartelt, M. Lalam, V. Savin, M. Gorgoglione, A. Dekorsy, G. Fettweis: *Benefits and Impact of Cloud Computing on 5G Signal Processing*, Special Issue "The 5G Revolution" of the IEEE Signal Processing Magazine, vol. 31, no. 6, pp. 35-44, November 2014.

[15] M. Lalam, T. Lestable: *On-Demand Radio Resource Management for Multi-Point Turbo-Detection in Dense Small Cell Deployment*, submitted to IWCPM 2015, London, June 2015.

[16] J. Bartelt, P. Rost, D. Wübben, J. Lessmann, B. Melis, G. Fettweis: *Fronthaul and Backhaul Requirements of Flexibly Centralized Radio Access Networks*, IEEE Wireless Communications Magazine, vol. 22, no. 5, pp. 105-111, Oct. 2015.

[17] H. Paul, J. Fliege, A. Dekorsy, *In-Network-Processing: Distributed Consensus-Based Linear Estimation*, IEEE Communications Letters, vol.17, no.1, pp.59-62, January 2013.

[8] B.-S. Shin, H. Paul, D. Wübben, A. Dekorsy: *Reduced Overhead Distributed Consensus-Based Estimation Algorithm*, IWCPM 2013, Atlanta, USA, December 2013.

[9] G. Xu, H. Paul, D. Wübben, A. Dekorsy: *Fast Distributed Consensus-based Estimation (Fast-DiCE) for Cooperative Networks*, WSA2014, Erlangen, Germany, March 2014.

[10] D. Wübben, H. Paul, B.-S. Shin, G. Xu, A. Dekorsy: *Distributed Consensus-based Estimation for Small Cell Cooperative Networks*, BWA 2014, Austin, TX, USA, Dec. 2014.