6 Energy-Saving Solutions and Practices

Currently, 3GPP is actively progressing in the standardization of 5G NR. The concepts and many solutions proposed in previous chapters have been adopted in the NR technical report, including, for example, service-based network architecture, two-layer (CU and DU) RAN architecture and signaling interfaces, software-defined air interface with flexible frame structure, unified multiple access schemes with orthogonal and non-orthogonal solutions, unified signaling procedure for MIMO with hybrid beamforming structures, multiple waveform support, etc. Toward green and soft networks of 5G and beyond, more efforts are needed in both standardization and practical deployment.

In this chapter, some practices in CMCC's wireless communication networks are discussed, from 2G, 3G, 4G, even to future 5G. The existing energy-saving schemes and practices in cellular networks and WLAN are first surveyed. Then the field trial of C-RAN is presented, with test results discussed in various scenarios. This chapter continues to explore the feasibility of energy-saving in mobile applications running on mobile phones. Finally, a massive MIMO platform named "Invisible BS" is introduced, which is featured by low-power, low-cost UE-grade RF IC for power-saving, and flexibly configurable SmarTiles with integrated RF and antennas.

6.1 Green Wireless Technologies in Cellular Networks

6.1.1 Energy-Saving in GSM

Energy-saving techniques in GSM have been investigated extensively [1–6]. A GSM multi-carrier BS can utilize digital intermediate frequency (IF) and multi-carrier power amplification (MCPA) technology to achieve multi-carrier on a single radio frequency (RF) channel [6]. Attributed to high integration, it is more energy-efficient than a traditional GSM BS using a single-carrier power amplifier (SCPA). According to field test results, the energy efficiency of a multi-carrier BS is as high as 25%, while that of the traditional single carrier BS is only 15%. Additionally, the BS is able to adjust the operating voltage of the MCPA according to the resource utilization of multiple channels. Field test results show that applying this technique alone can help each BS save 438 kWh per year. At present, it has been widely implemented in many GSM networks.

https://doi.org/10.1017/9781108277389.007 Published online by Cambridge University Press

6.1.2 Energy-Saving in TD-SCDMA

There are mainly two schemes for energy-saving in TD-SCDMA systems, i.e., the BBU baseband deactivation technique and the RRU time slot deactivation technique based on power amplification (PA). The former can reallocate or migrate a small amount of traffic to one or a few baseband boards during low traffic load, and deactivate the idle baseband boards to save energy. When the traffic load rises, it will activate the sleeping boards in advance to avoid traffic congestion caused by insufficient resources. The latter scheme takes advantage of the key characteristics of TDD systems, i.e., the UL and DL time slots are utilized alternately. When there is no traffic on the DL, the boards can be turned off in corresponding time slots in real time, thereby reducing power consumption.

6.1.3 Energy-Saving in LTE

Compared with GSM and TD-SCDMA, the PA technology of LTE equipment is more mature and widely used. At present, the energy conversion efficiency for LTE RRU is up to 40%. From a hardware perspective, it is difficult to further improve the conversion efficiency. There are many energy-saving techniques in LTE [7–16]. The most straightforward solution is to turn off the BSs at various granularity, e.g., symbol-level turnoff, channel-level turnoff, and small BS shutdown. Symbol-level and channel-level turnoffs are supported and implemented within eNBs, while small BS turnoff may impact the interface among eNBs. This technique has been standardized by 3GPP. In the following subsections, we will discuss these energy-saving methods and their test results in CMCC's networks.

Symbol-Level Turnoff: PA Shut down on Symbol-Level

When there is no actual data transmission in certain symbols, PA can be powered off so as to reduce power consumption and possibly reduce interference to the adjacent cells. For LTE normal frames, each subframe contains 14 symbols, some of which include reference symbols (RS). When in high traffic load situations, all symbols are carrying data and PA cannot be shut down, whereas in low traffic load situations, PA can be powered off rapidly, such that power consumption can be reduced. In the next symbol or frame duration, PA can be powered on again if there is data to send. The trial results show that approximately 5% of energy consumption of RRU can be saved when the traffic load is 10%. The approximate annual electricity savings in CMCC's networks is a very substantial 30 million kWh.

Channel Shutdown

When under low traffic load, a cell can choose to use only parts of its total transmit channels to transmit data, and the rest can be shut down to save energy. As a result, the transmit/receive chain can be turned off, including PAs and other analog devices, since the LTE BS implements multiple RF chains to support multiple frequency channels. If we shut down some transmit channels without affecting the capacity and coverage of cells, we can achieve the goal of energy-saving effectively. Our test results show that

https://doi.org/10.1017/9781108277389.007 Published online by Cambridge University Press

approximately 15% of energy consumption of RRU can be saved when the traffic load is 10%.

Small Base Station Turnoff

As the BS take up a large amount of energy consumption in cellular networks, turning off the BSs according to the traffic patterns is an effective way for energy-saving. We specify three energy-saving scenarios: inter-RAT, inter-eNB, and intra-eNB.

Inter-RAT energy-saving

In the inter-RAT energy-saving scenario, the LTE BS is used for capacity enhancement and overlays with the existing 2G and 3G networks and there are different methods to allow the BS to enter or leave sleep mode:

- The OAM method: e-UTRAN cells get into sleep mode according to the decision by centralized OAM.
- The signaling method: in order to achieve energy-saving gain, the energy-saving target cell and adjacent cells need to exchange specific network condition parameters, such as the traffic load threshold, the length of energy-saving time, and current energy consumption.

Inter-eNB energy-saving

As in Fig. 6.1, in the inter-eNB energy-saving, the E-UTRAN cells A and B provide wide area coverage, and small cells C–G are deployed within the initial coverage area to enhance capacity in some locations. It is also possible that the coverage area of the energy-saving cell and the coverage compensation cell do not overlap completely. When powering off the energy-saving cells, the coverage compensation cell needs to take measures to provide consistent availability to users in the area of interest, as shown in Fig. 6.2.

For the OAM-based method, the cells are configured by OAM as either an energy-saving cell or a coverage compensation cell. Besides the OAM-based method, the cells'

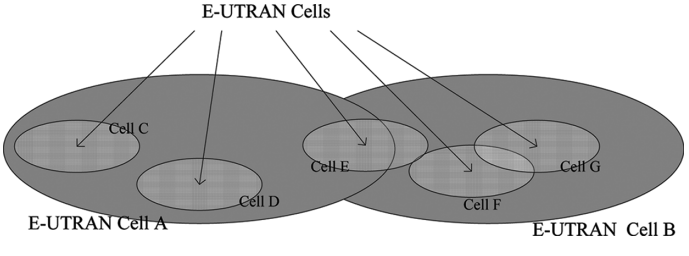

Figure 6.1 Inter-eNB energy-saving, where small cells C–G can be turned off.

https://doi.org/10.1017/9781108277389.007 Published online by Cambridge University Press

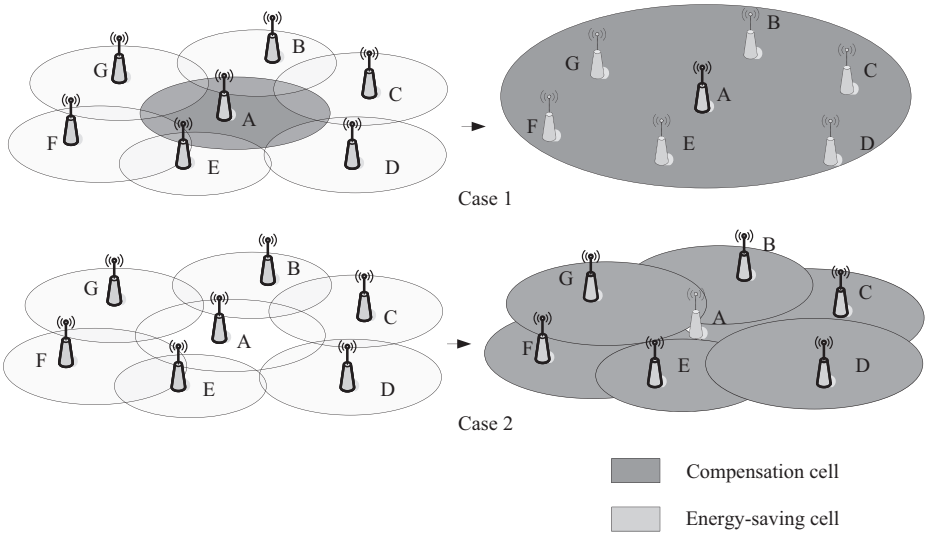

Figure 6.2 Illustration of coverage compensation in inter-eNB energy-saving.

role can also be dictated through the exchange of information signaling and/or OAM signaling.

Intra-eNB energy-saving

Intra-eNB energy-saving focuses on analyzing different methods that can be used by a single eNB. Possible solutions are:

- Multicast broadcast single frequency network (MBSFN) subframe configuration: When the normal frames change to MBSFN, power consumption can be reduced, since the overhead of the cell reference signal (CRS) in MBSFN is much reduced.
- The configuration of DwPTS length: for a TDD system, a special subframe includes DwPTS, UpPTS, and the guard period. A shorter DwPTS means less radiated power. Therefore, we can accomplish energy-saving by configuring fewer DwPTS when the traffic load is low.

6.2 Multi-RAT Cooperation Energy-Saving System (MCES)

In the previous section, we briefly introduced the concept and technique of energy-saving in a multi-RAT scenario. This section will put forth a novel system called MCES for effective improvement on the overall energy efficiency of mobile networks. MCES interacts with the RAN in real time, manner and can support multi-vendor equipments in 2G/3G/4G RAN, as shown in Fig. 6.3.

https://doi.org/10.1017/9781108277389.007 Published online by Cambridge University Press

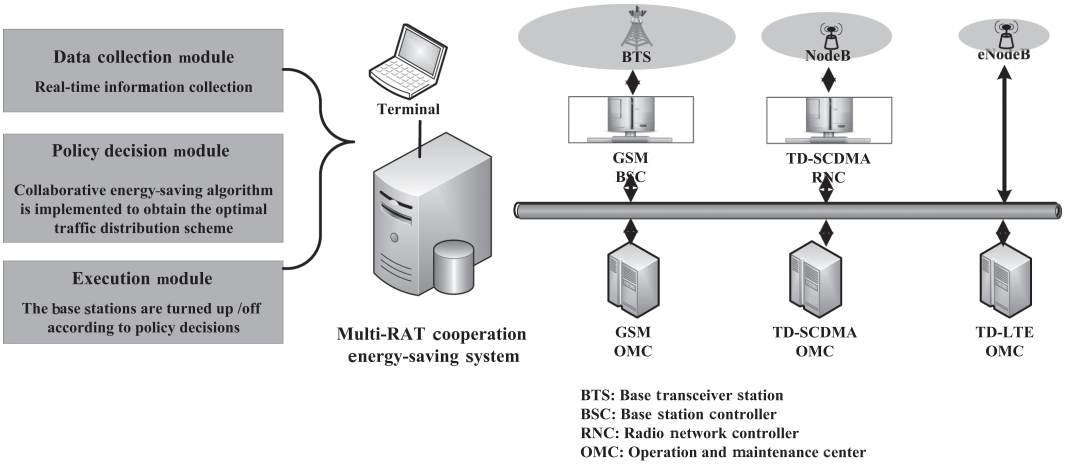

Figure 6.3 Topology of MCES.

6.2.1 Basic Principle

In recent years, with the rapid expansion of the customer base and soaring data traffic, the CMCC has deployed more than 1,500,000 TD-LTE BSs nationwide above the original TD-SCDMA and GSM BSs [17]. Current cellular wireless networks include GSM in 1,800/900 MHz, TD-SCDMA, and TD-LTE in D/F/E frequency bands. Multiple networks with overlapping coverage provide users with better Qos but also bring much-increased energy consumption. It is estimated that the CMCC consumes more than 20 billion kWh every year.

In this section, we first introduce the concept of network-level energy-saving. Network-level energy-saving technology refers to tight cooperation among BSs in single or multiple network topologies in order to improve the energy efficiency of the network [18]. To be specific, the saved power comes from putting some cells (identified as energy-saving cells) into idle mode and diverting its traffic to other cells (identified as compensation cells) [19]. An energy-saving cell usually has a traffic pattern with obvious tidal effect, i.e., the volume of traffic is very large during some time of the day and very small during the rest of day. The compensation cell provides basic coverage over the energy-saving cell, thus its coverage is generally slightly larger than the energy-saving cell. The relationship between the energy-saving cell and compensation cell is obtained by calculating the intercell coverage correlation, which is demonstrated in Fig. 6.4.

MCES is an application of the network-level energy-saving technology in a heterogeneous network consisting of GSM, TD-SCDMA, and TD-LTE. By collecting and analyzing the measurement report (MR) and the basic information on the 2G/3G/4G BSs, we have found that there are four types of overlapping scenarios in the network, which are listed in Table 6.1.

The technical precedure of MCES can be summarized as three sequential steps: energy-saving cell detection, energy-saving cell deactivation, and energy-saving cell activation:

https://doi.org/10.1017/9781108277389.007 Published online by Cambridge University Press

Table 6.1 Four types of overlapped scenarios.

	Energy-saving cell	Compensation cell
Type 1	GSM 1800 MHz band	GSM 900 MHz band
Type 2	TD-SCDMA	GSM 1800 MHz/900 MHz band
Type 3	TD-SCDMA	TD-LTE cell D/F band
Type 4	TD-LTE cell D band	TD-LTE cell F band

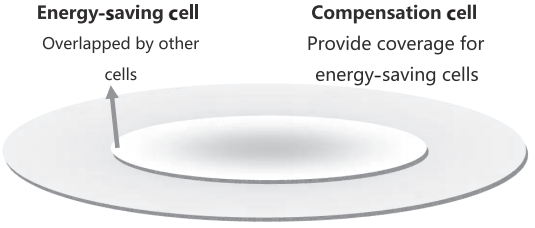

Figure 6.4 The coverage relationship between energy-saving cell and compensation cell.

1. Energy-saving cell detection: MCES collects the latitude, longitude, and MRs of all the BSs within the network of interest, and assigns the energy-saving cell-compensation cell pair. At the same time, MCES will continuously monitor the real-time traffic load of the energy-saving cell.
2. Energy-saving cell deactivation: When the energy-saving cell goes into a low-traffic load state, MCES can transfer its data traffic to the the compensation cells and deactivate the energy-saving cell.
3. Energy-saving cell activation: Thanks to the real-time monitoring function, MCES can wake up the sleeping energy-saving cell before high-data traffic arrives to ensure the QoS in the network.

6.2.2 Functional Architecture

MCES is divided into four modules: collection module, strategy module, execution module, and monitoring module, as illustrated in Fig. 6.5.

Collection module: It automatically collects all 2G/3G/4G cell performance and configuration information, MRs, system alarm information, etc.

Strategy module: A massive amount of data from the collection module is used as the input to the core algorithm for the purpose of guaranteeing the QoS in the network. First of all, according to different dimensions of data obtained, one of the two algorithms is selected to judge the relevance of cell coverage. Then, the combination of historical data and real-time data will be used to predict the traffic load of the energy-saving cells. In addition, the 2G/3G/4G network energy-saving parameters can be configured flexibly to meet the specific needs of different regions.

Execution module: According to the scheduling scheme generated by the strategy module, MCES can automatically complete the task of scheduling. Based on the specification of BS equipment, MCES can automatically send the instructions in parallel

https://doi.org/10.1017/9781108277389.007 Published online by Cambridge University Press

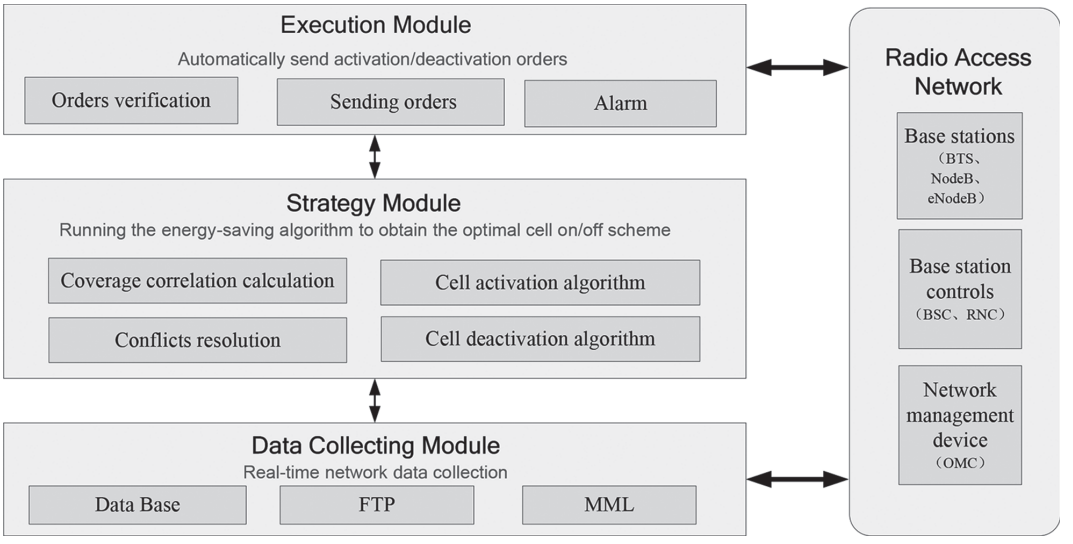

Figure 6.5 Functional architecture of MCES.

to the BS to shorten the execution time, and the BSs can automatically execute the deactivation and activation orders. Afterwards, MCES will keep tracks of the execution efficiency of such orders.

Monitoring module: MCES not only monitors processes, such as data collection, traffic prediction, and order execution, but also supervises performance statistics and alarm information in real time. If an abnormal event occurs, MCES will quickly provide the appropriate solution/action and notify the responsible parties.

6.2.3 Technical Characteristics

MCES is an energy-saving system based on real-time interaction with the RAN and is designed to support multi-vendor RAN equipments. MCES can turn off some of the overlapping cells in a timely manner to achieve network energy-saving in the scenario of significant coverage overlapping. Through analyzing massive MR data and the traffic profile, MCES can find the energy-saving cell and its corresponding compensating cell, and predict their traffic load trends. As for the energy-saving cell with low traffic load, MCES will migrate its traffic to the compensating cell and schedule the energy-saving cell to sleep mode. To ensure the reliability of the network, MCES also has a real-time monitoring function to turn on the sleeping cell when sudden high traffic occurs. Specifically, the MCES system has three major technical characteristics, which will be discussed next.

Network-Level Energy-Saving

Since traditional energy-saving technologies (e.g., micro DTX) are limited to single-network and single-cell scenarios, they cannot achieve optimal energy efficiency. MCES can oversee the wireless communication equipment that belongs to different networks

https://doi.org/10.1017/9781108277389.007 Published online by Cambridge University Press

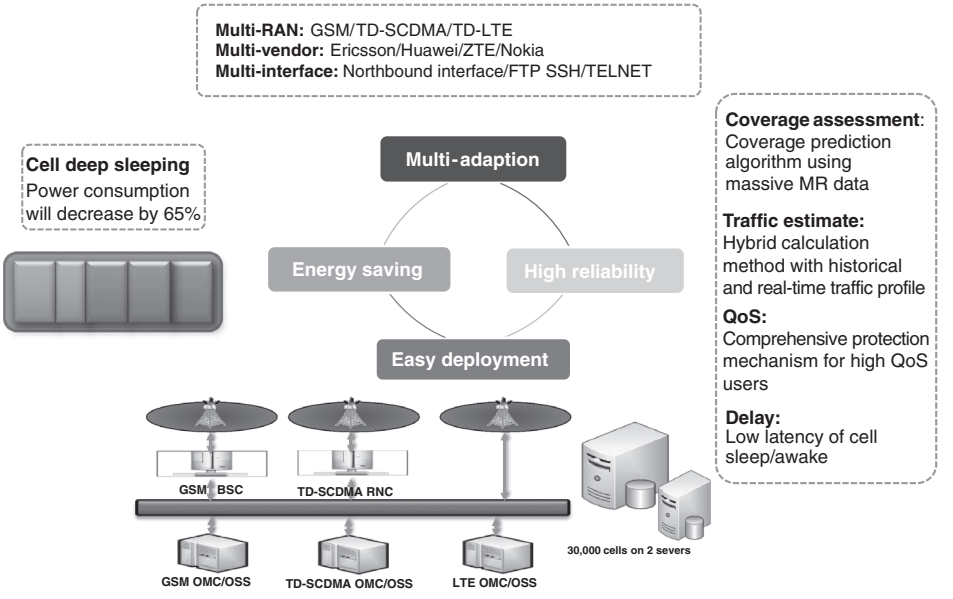

Figure 6.6 Technical characteristics of MCES.

and/or different vendors. Taking into consideration the overlay networks' respective status and the dynamic variations of traffic load, MCES can find the best actions for multiple networks and cells by analyzing the MR data.

Energy-Saving Cell Prediction Technologies Based on Big Data

MCES has access to cell configuration information, network performance statistics, and cell MRs, and could calculate intercell coverage correlation while locating the energy-saving cells and their corresponding cells in a multi-RAT network. The ML algorithm is used to make the accurate network traffic forecast for both types of cells through analyzing the historical traffic data and real-time traffic data.

Real-Time Cell On/Off Control

MCES can perform minimal periodic interaction with wireless network devices, obtain network performance data from the BS controller and OMC-R, and send instructions to the BS every 15 minutes. When the energy-saving cell goes into deep sleeping mode, its power consumption will immediately decrease by 65%.

6.2.4 Deployment Progress in China

Until now, MCES (see Fig. 6.6) has been deployed across about 7,000 cells in six provinces in China. The average annual electricity savings is 400,000 kWh per 10,000 cells.

In Nanning, for example, all 40,746 cells of CMCC's communication network can be accessed by MCES, where 6,704 cells could be turned off as energy-saving cells, and

https://doi.org/10.1017/9781108277389.007 Published online by Cambridge University Press

4,880.2 kWh can be saved daily on average. It is estimated that MCES will help to save 1,781,300 kWh every year, which is equivalent to 1,780,000 yuan. In the deployment process, both drive test results and the network performance KPIs were monitored. The results indicated that the cell throughput has not been affected by using MCES.

In Shanghai, the trial exceeds 20,000 cells, where 6,108 cells were identified as energy-saving cells, and 1,243 kWh could be saved daily on average. If MCES can reach the whole network of Shanghai, it is expected to achieve a daily energy-saving of 6,173 kWh and an annual cost savings of 2,250,000 yuan.

6.3 WLAN Energy-Saving Technology

With the development of mobile internet and the popularization of smart terminals, the scale and number of active WLAN networks have been rapidly increasing. Therefore, the soaring energy consumption of the WLAN access system has become increasingly prominent and power-saving solutions are in urgent need [20–25].

In the WLAN access system, the energy consumption of the access point (AP) takes a large proportion of the total energy consumption, and the radio frequency (RF) module accounts for the largest proportion of the total device energy consumption (up to 80%). Therefore, WLAN access system's energy-saving strategy mainly focused on how to reduce the idle energy consumption of the AP RF module so as to reduce energy waste in idle status. Several simple energy-saving mechanisms are briefly discussed in the following subsections.

6.3.1 AP Device Shutdown

An AP device could be shut down by configuring the on/off switch during a specific time window in a specific area based on the tidal effects of the traffic load. The function can be realized by upgrading the WLAN network management system, improving the power supply switch structure, or introducing a third-party control module. Nowadays, the last method has been deployed over a small scale of current network, aiming to maintain the original ONU/PoE architecture unchanged. It can implement power supply control and remote power failure monitoring, with the advantages of easy realization, low cost, and easy deployment. Energy-saving expectation depends on the ratio of off time and on time, which is estimated theoretically as about 30%.

6.3.2 AP RF Channel Shutdown

As the RF module accounts for the largest proportion of AP energy consumption, AP RF channels could be switched on/off to provide flexible and dynamic control strategy. Its implementation is relatively simple, such as by adjusting the PA bias voltage. The energy-saving expectation depends on the ratio of the AP RF module power consumption and shutdown time, which generally is about 50%.

https://doi.org/10.1017/9781108277389.007 Published online by Cambridge University Press

6.3.3 AP Single/Dual Band Selection

An AP single/dual band selection also helps to save energy in scenarios with tidal effects or intermittent service needs. Existing AP devices generally support 2.4 GHz and 5.8 GHz dual bands. According to a user's connection and signal strength, the 5.8 GHz band can be turned off or on automatically. The power consumption of the single-band APs could be reduced by at least 10% (depending on the different RF channel architectures) compared to that of the dual-band APs.

6.4 C-RAN Field Trials

6.4.1 Introduction

It has been pointed out that the key features of C-RAN include centralization, cooperation, and cloudification, as well as a clean system. With these features, C-RAN claims such advantages as cost reduction in terms of total cost of ownership (TCO), fast network deployment, improved system performance with support for cooperative technologies, on-demand and highly efficient resource utilization, and so on.

In order to demonstrate the features and verify the advantages of C-RAN, CMCC has been devoted to various field trials and prototype development since 2012.

The first step toward C-RAN is BBU centralization, which is relatively easy to implement and can be tested with existing 2G, 3G, and 4G systems. In the past few years, extensive field trials have been carried out in more than 10 cities in China using commercial 2G, 3G, and pre-commercial TD-LTE networks with different centralization scales. The main objective of C-RAN deployment in 2G and 3G is to demonstrate deployment benefits by centralization, including speedup of site construction and power consumption reduction. One example of the trial is in the city of Changchun, where 506 2G BSs in five counties were upgraded to a C-RAN-type architecture, i.e., centralized to several physical sites. In the largest central site, 21 BSs were aggregated to support 101 RRUs with a total of 312 carriers. It was observed that power consumption was reduced by 41% due to shared air conditioning. In addition, system performance in terms of call drop rate as well as DL data rate was enhanced using multiple-RRU-co-cell technologies. For detailed results and benefits demonstrated by centralization in 2G and 3G trials, readers can refer to [26, 27].

Similar results have also been verified in 4G LTE systems. However, a critical challenge for LTE C-RAN centralization in the past has been fronthaul (FH) transportation. FH is the link between BBU and RRUs. A typical FH interface includes CPRI and OBSAI, with CPRI being the most frequently adopted interface. FH transportation for 2G and 3G is not an issue, but becomes very challenging for LTE given the increasing number of antennas and bandwidth. Take CMCC's networks as an example. The RRUs are usually equipped with eight antennas with 20MHz of LTE bandwidth. As a result, the CPRI data rate is around 9.8 Gbps per carrier. It is clear that the larger scale of centralization, the higher the FH bandwidth is required. Therefore, fiber cores are required to enable C-RAN large-scale centralization, which unfortunately is rather expensive to

https://doi.org/10.1017/9781108277389.007 Published online by Cambridge University Press

most of the operators in the world due to the scarcity of fiber resources. FH transport had been an issue preventing C-RAN centralization.

Fortunately, with the advance in the transport area, low-cost FH transport solutions have gradually become feasible. In particular, WDM solutions have become more and more mature as mainstream FH solutions for 4G C-RAN centralization.

The basic idea of WDM is to use different wavelengths to transmit different signals within a single fiber core. There are two kinds of WDM solutions: passive WDM (p-WDM) and active WDM (a-WDM). Compared to the a-WDM solution, the passive solution does not need any power supply for the equipment. Therefore, it is more convenient for deployment; yet it usually offers less capacity than the active solution. For example, 12 wavelengths are typically supported by a set of p-WDM equipment while the capacity of active solutions is usually 48 wavelengths. In addition, compared to a-WDM, p-WDM is more stable and cost-effective.

6.4.2 Demonstration of WDM-Based FH Solutions

The test results for a-WDM-based C-RAN systems could be found in [28]. In this section, we will only provide the test results to demonstrate the applicability and the performance of the p-WDM solution for FH transportation. One of the reasons is that the p-WDM solution is the one widely adopted in CMCC's current 4G C-RAN networks. In addition, the tests are performed based on CMCC's commercial network, which could give the readers a more direct sense of how the real networks operate.

In the trial network, for each BS site, a pair of p-WDM boxes, called multiplexer/de-multiplexer, is needed. One box is put in the central office, together with the BBU pool, while the other is deployed on the remote site close to the RRUs. Instead of traditional gray optics, with the p-WDM solution, colored optical modules are needed on the BBU, RRU, and the multiplexer/de-multiplexer. Taking the UL as an example, the colored optics transform the radio signals to different colored wavelengths and multiplex them via the multiplexer. Then the multiplexed signal is transmitted via the single fiber core from the remote to the central site. In the central office, the wavelengths are de-multiplexed and distributed to different colored optics on the BBU where the signal will be processed. In CMCC's network, the data rate for each wavelength is 10 Gbps and each BS site consists of 3 RRUs with each covering one sector.

6.4.3 Test Methodology

The network in the test covers around 10 districts. In order to have a comprehensive test and evaluation of the p-WDM solution, in this test we consider and compare four scenarios: highway, suburbs, residential area, and enterprise buildings. For each scenario, the network consists of three types of deployment method: C-RAN, equipment room-based deployment, and outdoor cabinet-based deployment. The number of BSs for each deployment method and scenario is shown in Table 6.2. A set of performance indexes were tested, including RRC connection success rate, E-RAB connection success

https://doi.org/10.1017/9781108277389.007 Published online by Cambridge University Press

Table 6.2 Number of BSs with different deployment methods for different scenarios.

	DI	D2	D3	D4	D5	D6	D7	D8	D9	D10	Total
Highway	**1**	**3**	**6**	**0**	**2**	**2**	**0**	**1**	**0**	**2**	**17**
C-RAN	0	2	2	0	1	0	0	1	0	0	6
Outdoor cabinet	0	0	2	0	0	1	0	0	0	0	3
Traditional equipment room-based deployment	1	1	2	0	1	1	0	0	0	2	8
Suburb area	**1**	**1**	**0**	**2**	**0**	**1**	**0**	**1**	**0**	**0**	**6**
C-RAN	1	1	0	0	0	0	0	0	0	0	2
Outdoor cabinet	0	0	0	1	0	1	0	1	0	0	3
Traditional equipment room-based deployment	0	0	0	1	0	0	0	0	0	0	1
Residential area	**8**	**10**	**4**	**5**	**13**	**10**	**9**	**16**	**13**	**6**	**94**
C-RAN	3	3	1	4	5	5	1	7	3	2	34
Outdoor cabinet	3	4	2	1	5	3	4	4	5	4	35
Traditional equipment room-based deployment	2	3	1	0	3	2	4	5	5	0	25
Enterprise	**5**	**2**	**4**	**7**	**3**	**3**	**2**	**0**	**1**	**6**	**33**
C-RAN	1	0	1	0	2	1	0	0	1	2	8
Outdoor cabinet	2	1	1	3	0	0	1	0	0	1	9
Traditional equipment room-based deployment	2	1	2	4	1	2	1	0	0	3	16
Total	**15**	**16**	**14**	**14**	**18**	**16**	**11**	**18**	**14**	**14**	**150**

rate, average number of RRC connection, the maximum number of RRC connections, throughput, packet loss rate, delay, RRC connection establishment delay, block error rate, and so on.

Test Results

Network Fault Analysis

We tested the network for five months and continuously collected the statistics, as shown in Table 6.3.

From that table, the C-RAN network acted similar to traditional deployment methods (i.e., equipment room-based deployment) and was better than the outdoor cabinet solution in terms of network breakdown. Note that network breakdown mainly happens due to power failure or BBU equipment malfunction. Since the BBUs in a C-RAN network were centralized in the same office, it was thus easier for the engineers to pin down the faulty parts. As a result, the average recovery time for C-RAN was much shorter. Furthermore, there is no malfunction caused by high temperature of C-RAN. This is mainly because in the current C-RAN deployment the number of centralized BBUs is fewer than five (i.e., five sites centralized), and the air conditioning in the facilities in the central office in particular were much better than traditional BS sites.

https://doi.org/10.1017/9781108277389.007 Published online by Cambridge University Press

Table 6.3 Test results of the p-WDM-based C-RAN.

	Test item	Traditional equipment room-based deployment	Outdoor cabinet	C-RAN
Wireless	# of breakdown	79	16	73
	# of malfunction due to high temperature	4	1	0
	# of power failure	55	8	51
	# of BBU breakdown	20	7	22
	Average recovery time (hour)	19.9	4.2	12.4
Transport	# of breakdown	22	30	45
	# of optical link failure	19	29	42
	# of equipment breakdown	3	1	3
	Average recovery time (hour)	16.7	19.9	13.6

From the transport perspective, link errors happened much more frequent in C-RAN than on traditional sites. This is expected since in a C-RAN network the distance between the central office and the remote RRUs was much longer, leading to the increased possibility of link error. However, it turns out that the fault points were limited to some certain locations, which decreased the difficulty for fault detection and thus reduced the recovery time.

Performance Analysis

Through the five-month operation, we also collected extensive data to verify the performance of the p-WDM-based C-RAN networks. The data is shown in Table 6.4. Various indexes were collected from different perspectives. Note that there are two vendors involved in the test network by providing their C-RAN equipment. Both vendors' network was tested from a performance perspective.

From the test data in Table 6.4, the performance of p-WDM-based C-RAN is almost the same as the traditional deployment solutions.

6.4.4 C-RAN-Based UL CoMP Test

This trial was to demonstrate the performance gain of UL CoMP facilitated by the C-RAN architecture. The test was performed in a commercial network. The difference between inter-site and intra-site CoMP was also tested.

Test Environment and Network Configuration

The network in the test was a typical campus scenario that consisted of six 8T8R macro sites and one 2T2R small site. Here, "nTmR" means "n" transmission antennas and "m" receiving antennas. Each macro site contained three sectors while the small site was a single cell. In the central office, there were three BBUs, each controlling two macro

https://doi.org/10.1017/9781108277389.007 Published online by Cambridge University Press

Table 6.4 System performance for the C-RAN network with p-WDM as fronthaul solution.

Vendor 1		C-RAN	Outdoor cabinet	Traditional deployment	Total
Network completion rate-related	# of RRC connection request	4,989	14,878	14,571	11,765
	# of successful RRC connections	4,984	14,865	14,556	11,754
	RRC connection success ratio	99.91	99.91	99.91	99.91
	# of E-RAB request	4,555	13,926	14,347	11,227
	# of successful E-RAB connection	4,551	13,914	14,336	11,217
	E-RAB connection success rate	99.88	99.91	99.91	99.90
	Network completion ratio	99.78	99.82	99.83	99.81
Service-related	Average # of RRC connection	4.76	15.73	14.79	12.06
	Maximum number of RRC connections	16.44	40.65	38.26	32.45
	Total U-plane cell throughput (DL+UL) (bytes)	4,98,869	13,35,139	13,23,356	10,76,453
User experience-related	Cell-level DL packet loss rate	0.02	0.03	0.05	0.02
	Cell-level average DL delay (ms)	22.47	27.64	26.72	25.75
	MAC-layer UL block error rate	0.40	0.33	0.33	0.35
	MAC-layer DL block error rate	0.08	0.09	0.07	0.08
	Average time for RRC connection establishment (ms)	25.01	26.34	25.52	25.64
Vendor 2		**C-RAN**	**Outdoor cabinet**	**Traditional deployment**	**Total**
Network completion rate-related	# of RRC connection request	5,072	10,107	22,102	13,286
	# of successful RRC connections	5,056	10,091	22,025	13,246
	RRC connection success ratio	99.73	99.79	99.80	99.78
	# of E-RAB request	4,591	8,634	20,979	12,213
	# of successful E-RAB connection	4,587	8,631	20,893	12,178
	E-RAB connection success rate	99.93	99.94	99.87	99.91
	Network completion ratio	99.66	99.74	99.66	99.69
Service-related	Average # of RRC connection	5.35	10.07	26.09	14.86
	Maximum number of RRC connections	21.95	27.23	53.24	35.65
	Total U-plane cell throughput (DL+UL) (bytes)	4,48,195	7,82,517	21,79,566	12,20,804
User experience-related	Cell-level DL packet loss rate	0.05	0.01	0.02	0.02
	Cell-level average DL delay (ms)	31.78	33.19	35.36	33.63
	MAC-layer UL block error rate	0.23	0.34	0.22	0.26
	MAC-layer DL block error rate	0.29	0.26	0.21	0.25
	Average time for RRC connection establishment	24.79	26.66	29.27	27.14

https://doi.org/10.1017/9781108277389.007 Published online by Cambridge University Press

Table 6.5 System configuration.

Parameters	Value
Frequency	F and D band
Bandwidth	20MHz
Frame structure	• Configuration 2, Normal CP • Subframe configuration: DSUDDDSUDD
CFI	3
Antenna mode	DL：TM3/8 self-adaptive UL：SIMO
UL power control	On
HARQ	On
AMC	On
BS transmit power	8 × 5W
Handover	Contention-based

sites. In addition, there was a high-speed switch in the central office to connect the BBUs together for the sake of timely information exchange. Each macro site consisted of three sectors, with each sector having three cells on different frequency bands. As a result, the total number of macro cells was $6 * 3 * 3 = 54$. Together with a small cell, there were 55 cells in total in the trial network.

To have a better observation of the CoMP performance gain, the data was collected and analyzed during busy hours of the day. The system configuration is as shown in Table 6.5.

The test methodology is shown as follows.

- First, turn off the switch and any CoMP algorithms, then collect the statistics for one week;
- Second, turn on the intra-site CoMP while keeping the switch off and collect the statistics for one week;
- Third, turn on the inter-site CoMP and the switch to interconnect the BBUs, then collect the data for one week.

Note that as mentioned earlier, the data was collected during busy hours, which typically lasts for two hours.

Test Results

Without enabling any CoMP algorithms, the baseline performance is shown in Table 6.6. The performance gain with intra-site CoMP and inter-site CoMP open is shown in Table 6.7. It could be seen that CoMP has brought slight gain in terms of average UL cell spectral efficiency and average UL user experience rate. However, compared to intra-site CoMP, the performance by inter-site CoMP, which is enabled by the C-RAN architecture, is much higher.

When it comes to the performance gain for edge users, as shown in Table 6.8, the intra-site CoMP still brings a small gain. In comparison, C-RAN-enabled inter-site CoMP has much higher performance gain.

https://doi.org/10.1017/9781108277389.007 Published online by Cambridge University Press

Table 6.6 System performance with CoMP function off.

Average UL interference (dBm)	Average UL MCS	Average cell user number	Edge user (%)	Average utility rate for UL PRB (%)	Average throughput per E-RAB (KB)	Average UL throughput per PRB (bits)	Average UL throughput per PRB for edge users (bits)
-114.28	20.78	34.15	19.85	5.69	325	306	201

Table 6.7 Cell average UL CoMP gain.

UL CoMP	Average gain on cell UL spectral efficiency	Average gain on cell UL user experience rate
Intra-site UL COMP	2.42%	0.89%
Inter-site UL COMP	4.78%	10.91%

Table 6.8 UL CoMP gain for edge users.

ULCOMP	Average gain on UL spectral efficiency for edge users	Average gain on UL user experience rate for edge users
Intra-site UL COMP	2.90%	3.47%
Inter-site UL COMP	8.84%	17.25%

6.5 Green Application

As consumers rely more and more on mobile devices and applications (apps), their expectation of network connectivity is very high. Background services, software, and accessing data over the network are all part of the user experience. Both for smartphones and tablets, there is one common user focus: better battery life on mobile devices. In order to ensure that consumers get the longest battery life from their devices, feedback must be provided from all involving parties (network integrators, consumers, developers) (Fig. 6.7).

Consumers expect a device to feature outstanding battery life, as promised by the vendors. A typical claim is at least eight hours' usage on Wi-Fi, or four hours on LTE. Developers expect their application not to drain the battery too fast. Meanwhile, streaming media applications must meet users' expectations for performance. Network integrators focus on how much network activity is used by an application and how much signaling is generated. Research indicates that smartphones consume the same amount of data as feature terminals, while network signaling consumption is more than 10 times that.

To achieve better user experience, we believe a green application should not only be functionally complete, stable, and attractive but also power efficient. A green application will prolong battery life and generate less heat.

https://doi.org/10.1017/9781108277389.007 Published online by Cambridge University Press

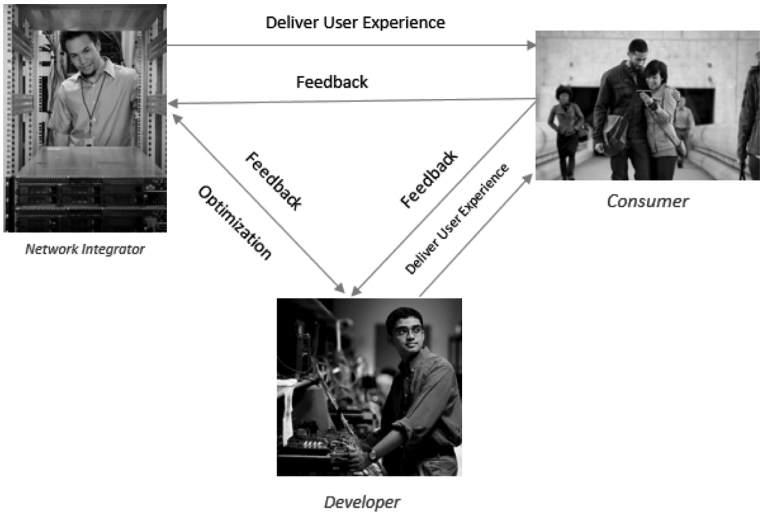

Figure 6.7 Consumer, developer, and network integrator dependencies.

This section describes an E2E system for green applications and intelligent network design, in order to optimize performance and minimize power consumption. The vision of this chapter is to construct an app-friendly ecosystem and a green community to help the whole industry understand how terminals, batteries, networks, and apps work together so as to guide developers to use a more efficient and friendly development mode. Consequently, both user experience and network efficiency can be improved.

6.5.1 Key Factors for App Power Consumption

The hardware architecture of a smartphone can be generally divided into three areas: application processor, communication processor, and peripherals. The application processor generally includes a central processing unit (CPU)/graphics processing unit (GPU), memory, etc. The communication processor mainly refers to network-communication-related modules. Peripherals are the screen, speaker, camera, sensors, and other parts.

Different types of smartphones in different operation modes have different power consumption ratios for the application processor, communication processor, and peripherals. For casual game playing, due to the short interaction time and few data exchanges with the server during the run time, the application processor plays an important role in power consumption. Therefore, for a smartphone running a casual game, the main performance optimization object is the application processor part. However, for apps such as online video, which requires constant high-rate network connection, the optimization focus should be shifted to the communication processor.

In the following subsections, we will separately discuss the impacts of the application processor and communication processor in terms of power consumption.

https://doi.org/10.1017/9781108277389.007 Published online by Cambridge University Press

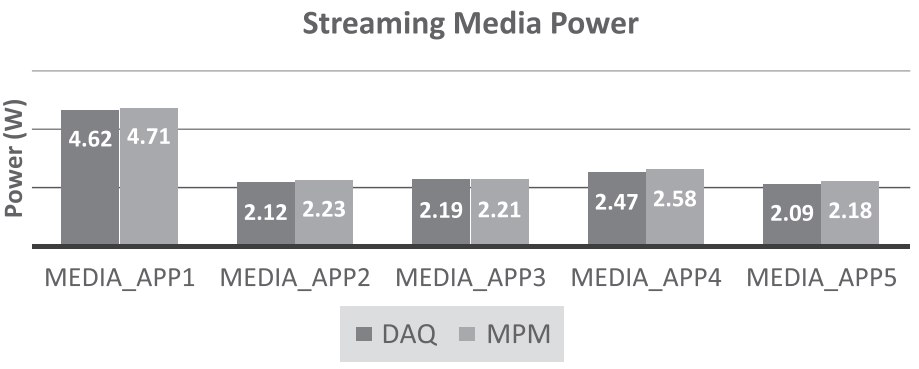

Figure 6.8 Streaming media power consumption in watts.

Application Processor

Consumers want smartphones to provide a great user experience with excellent battery life. Research indicates that battery life has become the single biggest reason people choose smartphones [29]. It is common that the performance of a best-in-class device is degraded by the application running on it. A poorly optimized application could destroy a good user experience. To understand how an application can impact the battery life, we have profiled various categories of applications for power consumption, including media streaming, casual games, and camera-based applications. For measurement, we use National Instrument's data acquisition (DAQ) and Intel's MPM. The goal of combining both hardware and software tools is to allow us to correlate power between the two different collection modes: application and hardware.

Category 1: Streaming Media

Media playback on smartphones or tablets is one of the top usages by consumers. Nearly 90% of usage on tablets is from media content [30]. To understand the impact of power consumption for media we tested similar video clips running at 1080p quality. There was a slight difference in quality, depending on the service provider who was streaming the video. Figure 6.8 shows the power consumption of several anonymous applications during media playback. Lower power indicates a longer expected battery life. MEDIA_APP2 can perform streaming and playback for 10 hours on a 20 Whr-battery while MEDIA_APP1 can run for less than 5 hours.

Category 2: Casual Game Play

Compared to streaming media, games are more computationally intensive. To understand the impact, we tested games with similar action-oriented game play. Figure 6.9 shows the power consumption of arcade games during normal usage.

Category 3: Camera Applications

More and more customers use smartphones and tablets for taking photos or videos. Even though the click takes hardly any time, the user might spend more time in the preview.

https://doi.org/10.1017/9781108277389.007 Published online by Cambridge University Press

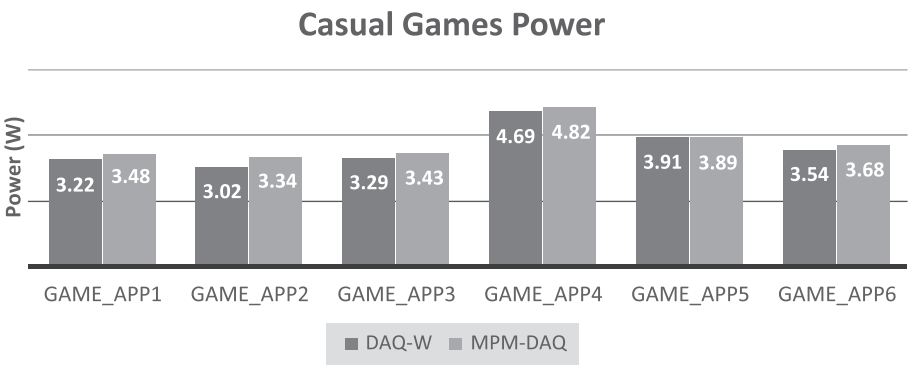

Figure 6.9 Casual game power consumption in watts.

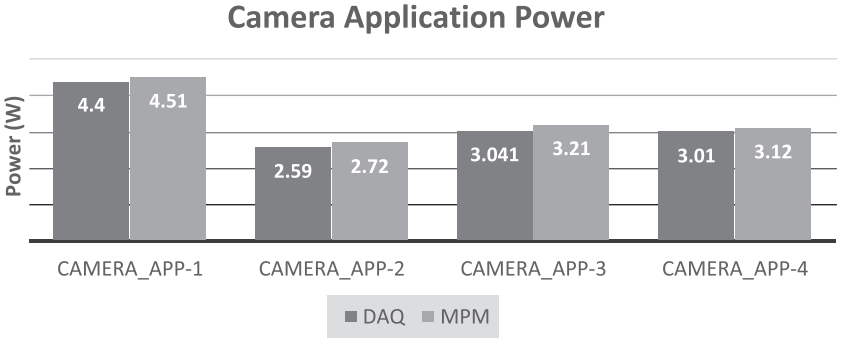

Figure 6.10 Camera application power consumption in watts.

Some applications add more editing features while others have predefined settings. We used a black-and-white effect to keep our data consistent between different applications. Figure 6.10 shows the power consumption of camera apps.

Communication Processor

Many mobile app developers design mobile communication similarly to how they would design for a wireless network, resulting in unfriendly network behaviors, such as higher power consumption, sustained mobile network signaling interaction, frequent small data transaction, and so on. They lack an appropriate understanding of the mobile network mechanism and the specific behavior of apps in a mobile network [31].

The key factor affecting application performance and network energy efficiency is the radio resource control (RRC) state machine whose purpose is to efficiently utilize limited radio resources and to improve smartphone's battery life. Application traffic patterns trigger RRC state transitions, which in turn affect radio resource utilization, smartphone power consumption, and user experience. Developers need to be aware of the reason and effect behind RRC state transition behavior [32–36].

https://doi.org/10.1017/9781108277389.007 Published online by Cambridge University Press

6.5.2 Optimization

There is no specific manual for optimizing applications, but there are guidelines that a developer should follow and then implement necessary changes in their applications. Note that different categories of applications will require different optimization. Different optimization methods can be adapted according to different application categories, as the following list shows.

Media: (Including streaming media, local media, audio streaming, photography, and video editing)

- Using the graphics and media pipeline; offloading to the GPU provides significant benefits for power and performance.
- Buffering during playback, record, or audio streaming; combining frames during processing saves significant power on the CPU and other components, such as storage and network access.
- Running simultaneously between CPU and graphics; overlapping calls helps to keep the CPU and GPU active at the same time.
- Using proper codecs and reducing total bandwidth; this results in running at a lower frequency and thus will make for lower power consumption.
- No overdrawing the user interface (UI) for the application; re-rendering the UI can cause unnecessary GPU usage.

Browsing: (Including watching media and casual browsing)

- No rendering the text or photo on the browser; this impacts graphics performance.
- No updating JavaScript animation during browsing.
- No buffering data over the network for advertisements and background processing.

Casual Gaming

- Using graphics for rendering: including UI rendering.
- No using the texture of very large sizes or unsuitable sizes (such as 13x19).
- Arranging draw call sequence with depth from near to far.
- No drawing regions invisible to end users.
- No doing alpha blending if unnecessary.
- Avoiding using branches in shading language if unnecessary.
- No using native screen resolution if up-sampling is acceptable since higher resolution can impact GPU usage.

Social, sharing, and emails

- Push notifications and background services causes high network usage and power.

Generic optimizations

- Optimizing timers, sleep, and spin to save power and increase performance.
- Minimizing background services updates to use less cache memory.
- No re-rendering UI that will impact frame-per-second and graphics utilization.

https://doi.org/10.1017/9781108277389.007 Published online by Cambridge University Press

6.6 "Invisible BS"

6.6.1 Motivation

In current cellular networks, there are a large quantities of BSs, including legacy 2G, 3G, and current 4G BSs. From the perspective of a network operator, a large portion of its capital expenditure (CAPEX) and operating expense (OPEX) lies in the purchase, deployment, and maintenance of such BS equipment. Moving forward with new BS technology, such as large scale antenna systems (LSASs), we can see capacity and power consumption being improved. However, the physical form factor and weight of a LSAS BS are of particular concern. The commercial deployment of new BSs faces resistance from the public, especially commercial property owners, regarding the aesthetics and potential/perceived health issues due to constant exposure to electromagnetic waves. The increasing footprint of the BSs will not only bring significant tower construction challenges but also greater confrontation. Moreover, it is more difficult to change the deployment site once chosen, which does not fit well with the flexibility requirement of next-generation communication networks.

In order to tackle these challenges, the concept of "invisible BS" has been proposed. By disguising the antenna system, the BSs can be made invisible to the public. A previous approach was to construct towers as fake trees, which are actually quite obvious. We modularize the BS, and design multiple antenna elements in the form of "tiles." By separating the LSAS panel into multiple tiles, one can flexibly deploy antennas in an irregular shape and as a part of the building facade or signage, thus blending into the surrounding environment. Another distinguishing feature is that low-power, low-cost RFIC is adopted in the "invisible BS" for further RF power-saving.

Figure 6.11 shows the system architecture of CMCC's hardware platform for an invisible BS. This platform most notably consists of a high-capacity software-configurable baseband unit and modular active antenna arrays, which is called SmartTile. The software-configurable baseband platform is responsible for baseband and physical-layer processing, particularly multiuser beamforming. Each SmarTile module contains eight RF chains and eight dual-polarized antenna elements. The baseband unit is connected to the SmartTile via the common public radio interface (CPRI) interface.

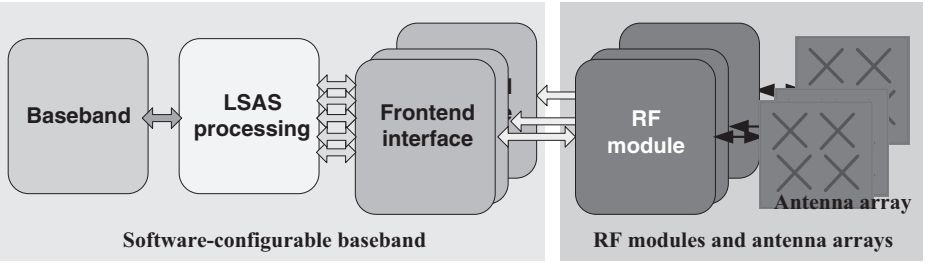

Figure 6.11 System architecture of CMCC's hardware platform for "invisible BS."

https://doi.org/10.1017/9781108277389.007 Published online by Cambridge University Press

The purpose of this system is to investigate the feasibility of invisible BSs and to verify different key 5G technologies, including massive MIMO, ultra wideband operation, RF requirements on large antenna arrays, OTA requirement, as well as highly effective heat dissipation techniques. There are several benefits to designing the LSAS system in this integrated manner:

- Dramatically increase system capacity by deploying multiple SmarTiles, equipped with many antennas and RF chains.
- Compact modular design, small form factor, easy deployment (e.g., on external building walls or logos), blends with surroundings, flexible site selection.
- Reduce power consumption due to the use of highly integrated, low-cost and low-power RFIC.
- Flexible arrangement of multiple modules.

6.6.2 Powerful Baseband Platform with a Unified Design

It is a great challenge for the 5G baseband platform to handle the amount of data generated from a large number of antenna elemnts (128/256/512) and a wide carrier bandwidth (e.g., above 100MHz). It is roughly estimated that tens of Gbits of data needs to be processed per second and several hundred Gbits of data needs to be transmitted per second. Our powerful baseband processing unit is designed with those high requirements in mind, using a universally acceptable design framework. A detailed system diagram is shown in Fig. 6.12.

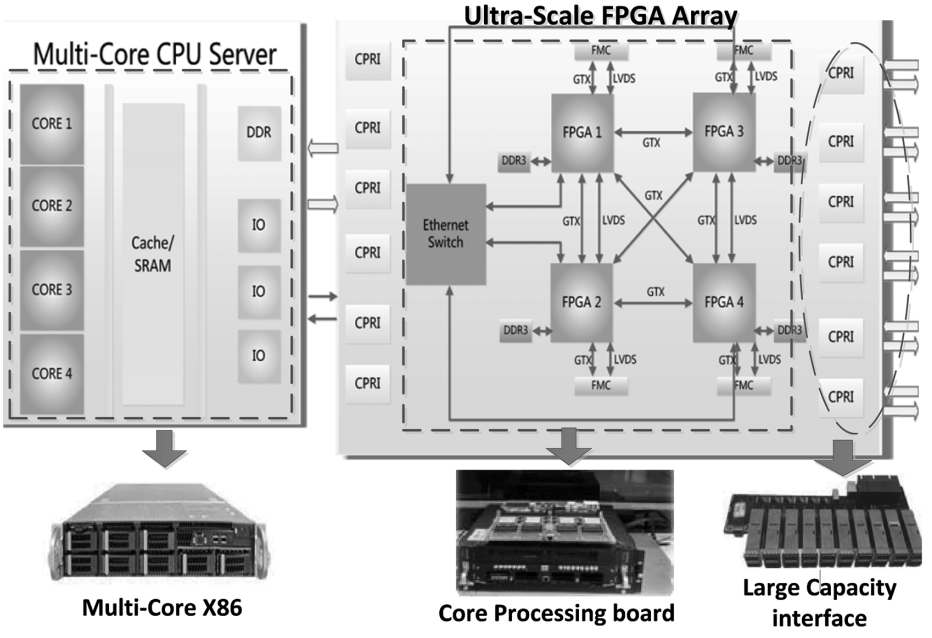

Figure 6.12 Baseband architecture.

https://doi.org/10.1017/9781108277389.007 Published online by Cambridge University Press

For the baseband platform, some other testbeds have also been proposed for massive MIMO verification. Wireless open-access research platform (WARP) [37] is a unified wireless network testbed for education, whose programmability and flexibility excels in implementing wireless communication systems quickly and cost-effectively. It has been used by Rice University in its ArgosV1 and ArgosV2 platforms [40]. However, such massive MIMO implementation becomes diffcult when the number of antenna elements increased significantly since the fronthaul interface and real-time baseband processing are the bottleneck. Lund University, in collaboration with National Instruments, has also developed a testbed named LuMaMi [37] for massive MIMO verification. It has the processing ability to support more than 100 antenna elements/RF chains for real-time wireless transmission [38]. The architecture design of the system is mainly based on FPGA, therefore, its compatibility with other platforms and softwares is rather weak.

Compared to other platforms, the interoperability and processing ability have both been taken into consideration for the design of the baseband platform described in Fig. 6.12. The baseband platform has a two-part architecture. The first level is the common processing architecture and the second is the boosted processing architecture. Based on this, the platform can support over 10 Gbps data processing and 800 Gbps fronthaul transmission. With the powerful hardware processing ability, various 5G key technologies can be verified, such as massive MIMO, new waveform, NoMA, etc.

Firstly, the platform has a multicore x86 server for tasks that are less computationally demanding, such as protocol stack processing and certain simple physical layer functions, since x86 CPU is good for task scheduling and multitask procedures. The obvious benefit here is the ability to place network functions and different implementations of protocol stacks in a flexible way and on the same server to ensure multi-vendor compatibility. Secondly, the boosted processing architecture focuses on highly complex signal processing, fast data exchange, and transmission. Therefore, it is mainly made of high-speed processing chips, such as FPGA, and super fast transmission interfaces, such as quad small form factor pluggable (QSFP+). It is based on the standard design architecture ATCA (advanced telecom computing architecture), and comprises an exchange processing board, a high-performance core processing board, and a high-capacity interface board. The exchange processing board mainly focuses on the scheduling and interactive processing of data; the high-capacity interface board is used at the connection between the RRU with interfaces, such as CPRI and serializer-deserializer (SERDES). The high-performance core processing board is the main computation hub within the BBU and is responsible for digital signal processing in the UL and DL transmission.

The high-performance core processing board, shown in Fig. 6.11, integrates with four high-performance FPGA chips of the Xilinx V7 series. The FPGA has abundant hardware resources and can handle complex baseband signal processing, such as high-order MU-MIMO, matrix decomposition, and channel estimation. For example, the calculation of inversion and singular-value decomposition (SVD) in matrix decomposition involves iterative complex calculation, such as multiplication, division, and extraction of the square root. The amount of computation required will likely increase with the antenna number. Large-scale real-time processing requires 128-IFFT/FFT processors

https://doi.org/10.1017/9781108277389.007 Published online by Cambridge University Press

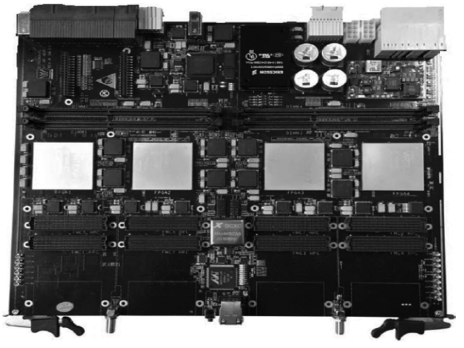

Figure 6.13 Baseband core processing board.

working at the same time. Each FPGA can provide millions of logical operations and thousands of DSP operations. As such, the core processing board with four such FPGA chips can provide baseband processing for massive MIMO systems. Besides, there is a total of 96 pairs of GTX (gigabit transceivers), providing 800 Gbps fronthaul transmission capability (theoretically equal to the data amount from 512 antennas), and LVDS (low-voltage differential signaling) for connections between the FPGAs and the ATCA back board.

To summarize, on one hand, it can support the complex calculations of massive MIMO. On the other hand, the x86 common server makes it much more flexible and compatible with different software providers and enables convenient software upgrade in the future.

6.6.3 SmarTile

The traditional BS baseband processing unit and antenna systems are highly integrated. However, we are seeing more and more benefits of heavy data processing distribution and BS deployment flexibility. Our SmarTile is an active antenna system that is composed of RF chains, antennas, self-calibration networks, and a baseband processor. It is among the first to realize such a compact modular system in the industry. Each unit can be used as an eight-antenna TD-LTE RRU, or an arbitrary number of units can be combined as a LSAS to suit diverse environments. Such design has been widely adopted in commercial equipment by vendors. The feature of being super compact (only 1.28L) and lightweight makes it rather portable and easy to mount on walls, roofs, and other structures such as lampposts, and thus well blended into the surrounding environment. Additionally, the superior energy efficiency demonstrates our ideal of being “green,” which is worth doing not only from an economic sense but also from the environmental sense. Table 6.9 lists some of the most important parameters of the SmarTile.

The SmarTile is so highly integrated that all on board components are assembled by hardware connections. The hardware structure of the SmarTile is shown in Fig. 6.14.

https://doi.org/10.1017/9781108277389.007 Published online by Cambridge University Press

Table 6.9 Detailed specifications of the SmarTile.

Frequency band	1880–1920MHz
Channel band	20MHz
RF chain number	8
Antenna number	8
Output power per chain	21dBm
Antenna gain	6dBi
Size	16×16×5cm
Power consumption	25 W

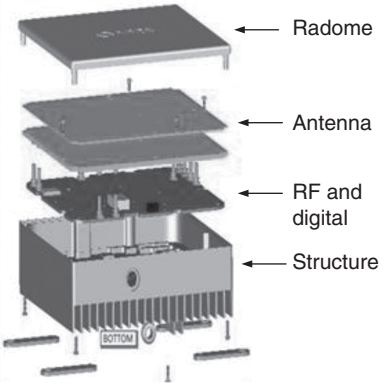

Figure 6.14 Structure of the SmarTile.

Hardware Design

The SmarTile's hardware is designed to reduce power consumption as much as possible. In this section, we look at a couple of design features of the SmarTile and explain in detail why it is unique.

RF Chain

To reduce the complexity of RF design, we choose a zero intermediate frequency (ZIF) structure to build the RF chain. A highly integrated transceiver is adopted, which consists of ADC, DAC, LO, IF amplifier, IF filter, and modulator. The ZIF structure can sample the signal with a lower rate, which can dramatically reduce the power consumption of ADC and DAC. Due to the phase and amplitude imbalance of the signal, the side image and DC offset are the most disadvantageous aspect of ZIF structure RF design, as they cannot be removed by filtering and must be calibrated out by a closed-loop algorithm from the baseband. This transceiver has the ability to remove the side image and DC offset, which makes the baseband design much easier. Each transceiver supports two RF chains with lower power consumption performance. The power amplifier and low-noise amplifier used in this platform are commercial components that are originally designed for user equipment, hence they have better power efficiency compared to BS amplifiers due to the power consumption limitation on user terminals.

https://doi.org/10.1017/9781108277389.007 Published online by Cambridge University Press

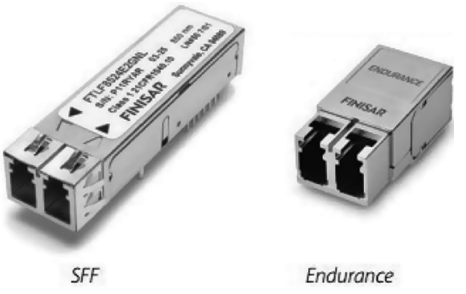

Figure 6.15 Compact SFP optical module.

Digital Hardware

The FPGA is used to process base-band data and is integrated with an ARM core; therefore no extra MCU is needed. The FPGA also contains two 10 Gpbs SERDES, which are responsible of transforming serial data to parallel data and vice visa. The interface between BBU and AAS is optical fiber via a 10 Gbps SERDES.

The majority of the functions at the intermediate frequency, such as DDC and DUC, are performed by RF transceivers, which also helps reduce power consumption.

Again for the purpose of size reduction, we have chosen a compact SFP optical module (shown in Fig. 6.13). Compared to the typical SFP module, the size is reduced from 56.5 mm to 26.41 mm.

Antenna and Calibration Network

The SmarTile has eight RF chains and eight antenna elements, as in shown Fig. 6.16. One patch antenna has two logical elements and is +/− 45-degree dual-polarized. Each element has a 6-dBi antenna gain with an 85-degree HPBW in both horizontal and vertical direction. The distance between patches is half the operating wavelength. We have measured the overall RF performance of this platform, and such a compact array has good array gain and side lobe suppression.

A calibration network is used to make sure all the RF chains have the same phase and amplitude output. The calibration network couples DL/UL signals from/to each antenna element. For better beamforming performance, the phase error should be less than five degrees, and the amplitude error should be less than 0.5dB.

Power Consumption Comparison

SmarTile is designed to be low-power and of small form factor. To demonstrate our superior power-saving performance, Tables 6.10 and 6.11 compare to a commercial eight-antenna BS.

Due to the lower power consumption, the size of the heat sink can be reduced to make the BS even smaller and lighter. It also reduces the cost of electricity and ultimately improve the carbon emissions.

https://doi.org/10.1017/9781108277389.007 Published online by Cambridge University Press

Table 6.10 The RF power consumption for commercial RRU.

Component	Power consumption	Number
Amplifier	100 mW	8
Mixer	100 mW	4
IQ modulator	200 mW	8
ADC	1.5 W	2
DAC	1 W	4
Total	9.8 W	

Table 6.11 The RF power consumption for SmarTile.

Component	Power consumption	Number
Transceiver	1.1 W	4
Total	4.4 W	

Table 6.12 RF performance for a SmarTile.

Test item	Specification
Output power	21.005dBm
ACLR upper	–47.25dB
ACLR lower	–48.56dB
EVM 64QAM	2.6%
EVM QPSK	1.13%

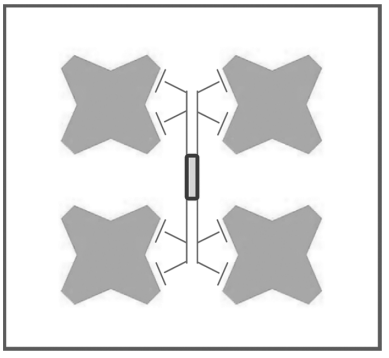

Figure 6.16 Antenna element and calibration network.

SmarTile RF Performance

The DL RF parameter is measured through the antenna ports by a spectrum analyzer. The test mode was ETM 3.1. The test results are shown in Table 6.12 and listed as follows.

Figure 6.17 shows the output power and the adjacent channel leakage ratio (ACLR) of the SmarTile, which is a key performance indicator (KPI) of the BS's transmitter quality. The output power of each RF chain is 21dBm with a bandwidth of 20 MHz,

https://doi.org/10.1017/9781108277389.007 Published online by Cambridge University Press

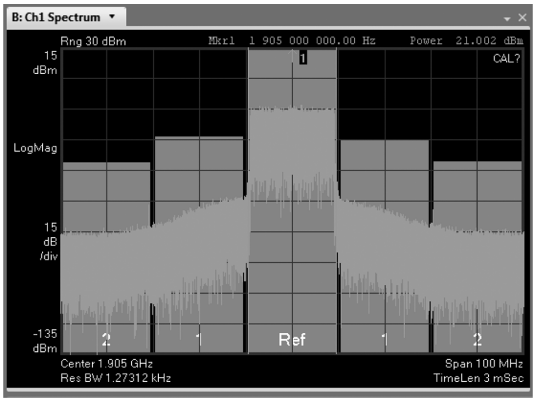

Figure 6.17 The output power and ACLR for a SmarTile.

and the theoretical output power of the whole BS is 30dBm. The measured ACLR was 47.25dBc, better than the 3GPP standardized requirement, which is −45dBc.

6.6.4 Flexible Over-the-Air Calibration Scheme

In the LSAS, the DL beamforming weights are derived from the UL channel estimation. For TDD systems, the channel information of UL can be used for DL precoding calculation because of UL and DL channel reciprocity. This is considered as the most significant advantage over FDD systems, where training pilots waste radio resource on the DL. Furthermore, the computation and overhead become prohibitive as the BS antenna elements grow large. However, for TDD systems, the RF paths of transmitter and receiver are asymmetric, and such discrepancies in the frequency and magnitude response leads to unacceptable channel mismatch between UL and DL. Moreover, the beamforming weights calculated based on the UL channel will no longer match the DL actual channel anymore, resulting in performance degradation. Therefore, TDD large-scale antenna system requires UL and DL calibration. The most common method is to estimate such difference beforehand and, periodically, to compensate the deviation in the DL beamforming parameters.

The mainstream antenna calibration scheme currently uses external measurement devices. The hardware calibration network is implemented physically for all RF paths [39, 40]. By performing predefined measurement procedures, the related parameters of each RF path, such as amplitude, phase, and frequency, can be obtained, which can be used to calculate the calibration coefficients for the BS. The architecture of the calibration network is shown in Fig. 6.18.

However, there are several drawbacks for this method. Firstly, extra hardware devices are needed, which means extra engineering actions, such as installation and debugging. Secondly, it will be a great challenge to calibrate antenna elements that are distributed far from each other physically. Moreover, the cost and complexity of such a hardware

https://doi.org/10.1017/9781108277389.007 Published online by Cambridge University Press

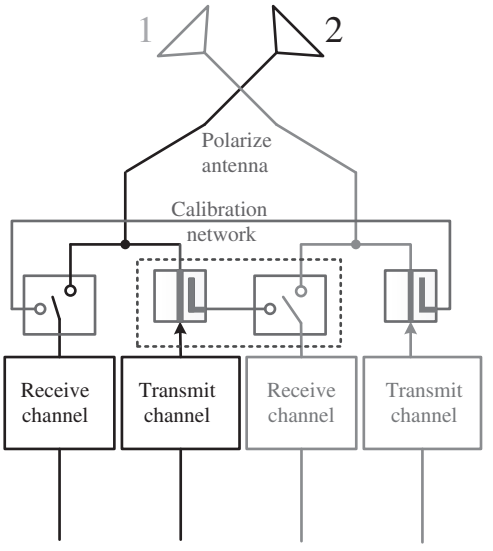

Figure 6.18 Antenna calibration network.

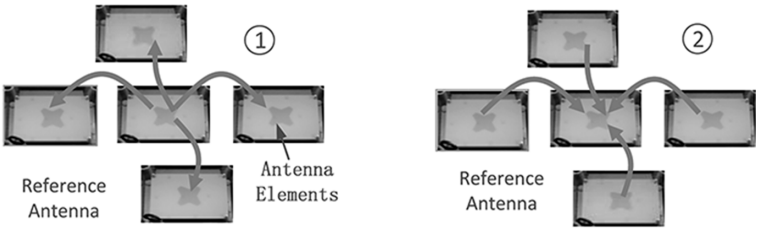

Figure 6.19 Current over-the-air calibration scheme.

network will increase with the number of antennas. Therefore, the old calibration network approach is no longer feasible for the distributed massive MIMO systems, which are envisioned as the most likely deployment scenarios for invisible BS.

Calibration over the air is the most promising scheme and was originally proposed in [41]. In this method, the reference signal for calibration interacts with wireless media other than the complex cabled network. The main step is to choose a reference antenna in advance, and the calibration signal will be transmitted and received between the reference antenna and all other antennas sequentially. Then, the calibration factor can be computed and compensated in the DL beamforming procedure. As in the Fig. 6.19, the antenna in the center is the reference element and the other four antennas around it are the to-be-calibrated antennas.

However, this OTA method still has some problems. The reference antenna is usually chosen as the center element of the array. The performance of the calibration is closely

https://doi.org/10.1017/9781108277389.007 Published online by Cambridge University Press

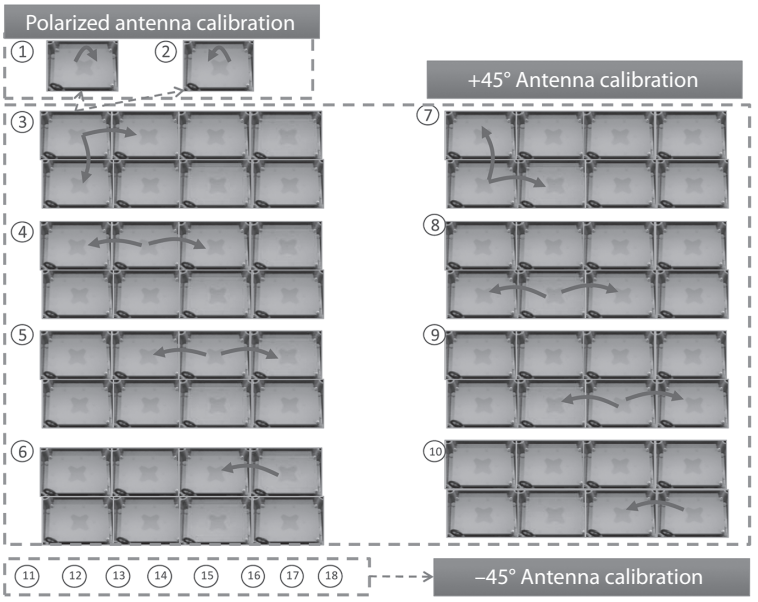

Figure 6.20 The procedure for flexible calibration scheme.

correlated with the location of the reference antenna. However, for a large antenna array, the distance to the reference antenna can be quite different for different elements. Besides, for the invisible BS, the distribution of antennas is irregular and the central element cannot not always be found.

To combat this issue, a flexible OTA calibration scheme is proposed for the massive MIMO system with an irregular antenna array in the invisible BS use case. The method is first proposed and published in [42]. The basic idea of this scheme is that the reference antenna is not fixed. Each RF path will send the calibration signal consecutively and the nearby antenna will receive it. Then the reciprocity factor can be calculated in each transmit-and-receive procedure. By a series of calculations, the calibration factor can be deduced. An example of eight SmarTiles (each SmarTile has a set of dual-polarized antennas) is given in Fig. 6.20.

In this example, several procedures are implemented and the whole process is introduced in detail:

1. The two polarized antenna calibration inside one SmarTile described as ① − ②;
2. The antenna calibration with 45 degrees polar, described as ③ − ⑩;
3. The antenna calibration with −45 degree polar, described as ⑪ − ⑱;
4. The calibration factor can be calculated and deducted;
5. The compensating factor is added in the DL beamforming.

The flexible calibration scheme has been verified to be very effective. For the different shapes of the antenna arrays, the calibration scheme performs well. The results

https://doi.org/10.1017/9781108277389.007 Published online by Cambridge University Press

Table 6.13 The performance of flexible air calibration.

Amplitude deviation	Phase deviation
0.0272dB	3.8842°

in Table 6.13 come from the real test in the system. It indicates that the performance of the proposed OTA calibration method can meet the requirement of LTE well. It can be used in a regular and irregular distributed antenna array without an external calibration network and reference antenna. It is first proposed and verified in the real testbed system and can be a potential scheme for commercialized devices. Moreover, for the standardization of 5G in 3GPP, lots of discussions[43–45] have been made for calibration in the Release 15 agenda item 7.1.2.5, and this scheme provides some reference and support for the discussions, which is very meaningful for communication systems in the future.

6.6.5 High-Efficiency Heat Dissipation Testbed

The designed power consumption of a SmarTile is 25 watts. Due to its small form factor, it is hard to dissipate the heat generated by operating with such high ouput power. Therefore, we have used a SmarTile and built a prototype to verify some new emerging heat-dissipation techniques.

Heat Dissipation Testbed

The prototype is composed of the SmarTile and temperature monitor system described in Fig. 6.21. Different thermal dissipation technologies can be verified on this prototype by looking at the temperature changes of the key components on the SmarTile, such as FPGA and clock modules.

The data processing board is a baseband prototype, which controls transmission and reception of data to/from the SmarTile. The reference signal source on the processing board supplies the whole system clock. Thermal couple wires were used to measure the temperature of each IC component.

For an outdoor BS, the main method for thermal dissipation is through natural heat radiation. There are three main types of material used to aid thermal radiation design: interface material, including thermal conductive adhesive, thermal pad, and phase change material; thermal-averaging material including graphite sheets and flat plate heat pipe; and a thermal radiation shell. Our prototype is used to perform thermal design research and test these three materials, as described in Fig. 6.22.

Interface Material

Interface material transfer/conducts thermal heat from the contact interface to the outside. In an ideal situation, the thermal source and the interface of the radiator will have the same temperature; that way, heat radiated by the source can be dissipated

https://doi.org/10.1017/9781108277389.007 Published online by Cambridge University Press

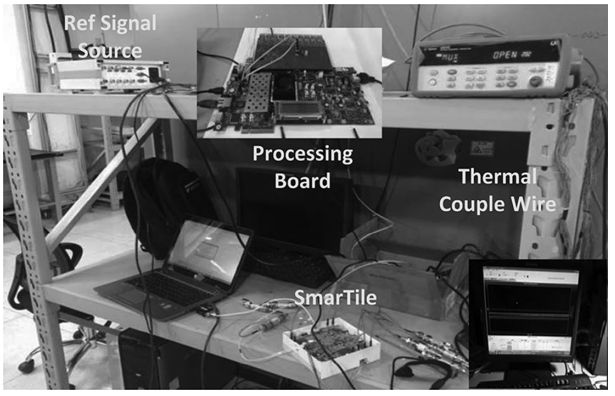

Figure 6.21 Heat dissipation testbed.

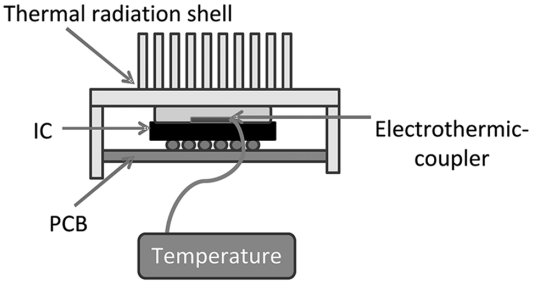

Figure 6.22 The thermal design of the prototype.

quickly. For better thermal dissipation, the interface material should have higher thermal conductivity.

An ordinary silicon-based thermal pad has a thermal conductivity of 2.0 W/m-k. In order to enhance thermal conductivity, we adopt a new interface material based on graphene with a thermal conductivity of 15.0 W/m-k as of now, which could significantly improve thermal performance. The lab test in Table 6.14 also verified such improvement.

Thermal Radiation Shell

The heat produced in the BS should be conducted to the outside shell effectively, and then heat can dissipate in the air through thermal exchange. Hence, the thermal resistance should be as low as possible. One way to lower the thermal resistance is to enlarge the size of the shell, and the other method is to increase the thermal conductivity of the shell. Because the size of the SmarTile is fixed, we strive to find a metal material that has higher thermal conductivity. Table 6.15 shows two types of aluminum alloy that have different thermal conductivities.

https://doi.org/10.1017/9781108277389.007 Published online by Cambridge University Press

Table 6.14 Thermal performance comparisons of different materials.

Thermal pad	FPGA (°C)	CLOCK (°C)	Transceiver (°C)
Silicon	61.95	58.24	61.38
Grapheme	53.41	54.90	59.57
Delta	**8.54**	**3.34**	**1.81**

Table 6.15 Thermal conductivity of two types of aluminum alloy.

	thermal conductivity（W/m-K）
Ordinary aluminum alloy	90–110
High-conductivity aluminum alloy	150–180

Through simulation by adopting high-conductivity aluminum alloy, the temperature of FPGA had a reduction of 1.5°C.

Thermal-Averaging Material

The thermal-averaging material is used to eliminate very high temperature points. For example, the graphite sheet is used widely in terminal design. In the future, the thermal-averaging material based on graphene will have even better thermal conductivity. And we hope that this kind of material will be used in the BS to further facilitate better heat dissipation performance.

References

[1] M. S. Ilyas et al., "Low-carb: Reducing energy consumption in operational cellular networks," *2013 IEEE Global Commun. Conf. (GLOBECOM)*, Atlanta, GA, pp. 2568–2573.

[2] T. Bitzer, J. Achard, and A. Pascht, "New energy-saving multicarrier transceivers and their standardization," *Bell Labs Tech. J.*, vol. 15, no. 2, pp. 53–58, Sept. 2010.

[3] C. Lubritto et al., "Telecommunication power systems: Energy saving, renewable sources and environmental monitoring," *IEEE 30th Int. Telecommunications Energy Conf.*, San Diego, CA, pp. 1–4, 2008.

[4] L. Anaya, D. Valerdi, L. Lezhan, X. Wenbo, and M. Torres, "Field validation of smart energy saving features in a GSM network," *2nd IEEE PES Int. Conf. and Exhibition on Innovative Smart Grid Technologies*, Manchester, 2011, pp. 1–5, 2011.

[5] A. Vidal, M. C. Purisima, M. T. Perez et al., "GSM timeslot detection and switching for power amplifier duty cycling in community cellular networks," *2017 Int. Conf. on Computer, Inf. and Telecommunication Systems (CITS)*, Dalian, China, pp. 200–204.

[6] Q. Lv, "A milestone of GSM BTS: Application of multi-carrier technology in GSM," *ZTE Technologies*, no. 2, pp. 1–5, 2010.

[7] F. Li, Y. Zhang, and L. Li, "Enhanced discontinuous reception mechanism for power saving in TD-LTE," *3rd Int. Conf. on Computer Science and Inf. Technol.*, Chengdu, China, pp. 682–686, 2010.

https://doi.org/10.1017/9781108277389.007 Published online by Cambridge University Press

[8] A. Virdis, G. Stea, D. Sabella, and M. Caretti, "A distributed power-Saving framework for LTE HetNets exploiting almost blank subframes," *IEEE Tran. on Green Commun. and Netw.*, vol. 1, no. 3, pp. 235–252, Sept. 2017.

[9] Y. B. Lin, L. C. Wang, and W. C. Chen, "eSES: Enhanced simple energy saving for LTE HeNBs," *IEEE Commun. Lett.*, vol. 21, no. 11, pp.2520–2523, Nov. 2017.

[10] P. K. Wali and D. Das, "Enhanced-power saving semi-persistent scheduler for VoLTE in LTE-Advanced," *IEEE Tran. on Wireless Commun.*, vol. 15, no. 11, pp. 7858–7871, Nov. 2016.

[11] V. J. Kotagi, R. Thakur, S. Mishra, and C. S. R. Murthy, "Breathe to save energy: Assigning downlink transmit power and resource blocks to LTE enabled IoT networks," *IEEE Commun. Lett.*, vol. 20, no. 8, pp. 1607–1610, Aug. 2016.

[12] L. You, L. Lei, and D. Yuan, "Optimizing power and user association for energy saving in load-coupled cooperative LTE," *2016 IEEE Int. Conf. on Commun. (ICC)*, Kuala Lumpur, pp. 1–6.

[13] K. Kanwal and G. A. Safdar, "Reduced early handover for energy saving in LTE networks," *IEEE Commun. Lett.*, vol. 20, no. 1, pp. 153–156, Jan. 2016.

[14] R. Vassoudevan and P. Samundiswary, "Performance analysis of DRX power saving technique for LTE based UE under bursty Web traffic," *2015 Int. Conf. on Commun. and Signal Process. (ICCSP)*, Melmaruvathur, India, pp. 924–928.

[15] A. Prasad and A. Maeder, "Energy saving enhancement for LTE-Advanced heterogeneous networks with dual connectivity," *2014 IEEE 80th Veh. Technol. Conf.*, Vancouver, BC, pp. 1–6.

[16] R. Imran, M. Shukair, N. Zorba, O. Kubbar, and C. Verikoukis, "A novel energy saving MIMO mechanism in LTE systems," *2013 IEEE Int. Conf. on Commun. (ICC)*, Budapest, pp. 2449–2453.

[17] CMCC, "China Mobile sustainable development report 2016." http://10086.cn/download/csrreport/cmcc_2016_csr_report_full_cn.pdf.

[18] S. Klein, E. Kuehn, and W. M. Wajda, "Energy savings in mobile networks based on adaptation to traffic statistics," *Bell Labs Tech. J.*, vol. 15, no. 2, pp. 77–94, Sept. 2010.

[19] Y. H. Hsu and K. Wang, "An adaptive energy saving mechanism for LTE-A self-organizing HetNets," *7th Int. Conf. on Ubiquitous and Future Networks*, Sapporo, Japan, pp. 289–294, 2015.

[20] R. P. Liu, G. J. Sutton, and I. B. Collings, "WLAN power save with offset listen interval for machine-to-machine communications," *IEEE Trans. on Wireless Commun.*, vol. 13, no. 5, pp. 2552–2562, May 2014.

[21] H. Tabrizi, G. Farhadi, and J. Cioffi, "An intelligent power save mode mechanism for IEEE 802.11 WLAN," *2012 IEEE Global Commun. Conf. (GLOBECOM)*, Anaheim, CA, pp. 3460–3464.

[22] F. Zhu and Z. Niu, "Priority based power saving mode in WLAN," in *2008 IEEE Global Commun. Conf. (GLOBECOM 2008)*, New Orleans, LA pp. 1–6.

[23] F. Zhu and Z. Niu, "Load-aware power saving mechanism in WLAN," *2007 IEEE Wireless Commun. and Netw. Conf.*, Kowloon, Hong Kong, pp. 2086–2091.

[24] J. H. Jun, Y. J. Choi, and S. Bahk, "Power-Saving schedulers for a WLAN with task-linking topology awareness," *IEEE Trans. on Veh. Technol.*, vol. 56, no. 3, pp. 1345–1356, May 2007.

[25] H. Kim and D. Cho, "Enhanced power-saving mechanism for broadcast and multicast service in WLAN," *IEEE Commun. Lett.*, vol. 9, no. 6, pp. 520–522, Jun. 2005.

[26] C.-L. I et al, "Recent progress on C-RAN centralization and cloudification," *IEEE Access*, vol. 2, pp.1030–1039, Aug. 2014."
[27] J. Wu et al., *Green Communication*, CRC Press, 2013.
[28] C.-L. I et al., "Recent progress on C-RAN centralization and cloudification," *IEEE Access*, vol. 2, pp. 1030–1039, Aug. 2014.
[29] G. Auer et al., "Energy efficiency analysis of the reference systems, areas of improvements and target breakdown," EARTH_WP2_D2.3_v2, Dec. 2010. www.ict-earth.eu/publications/deliverables/deliverables.html.
[30] G. Auer et al., "How much energy is needed to run a wireless network," *Wireless Commun.*, vol. 18, no. 5, pp. 40–49., Oct. 2011.
[31] 3GPP TS 23.401, "General packet radio service (GPRS) enhancements for evolved universal terrestrial radio access network (E-UTRAN) access (Release 8)."
[32] 3GPP TS 36.331, "Evolved universal terrestrial radio access (E-UTRA); Radio resource control (RRC); Protocol specification (Release 8)."
[33] 3GPP TS 36.104, "Evolved universal terrestrial radio access (E-UTRA); Base station (BS) radio transmission and reception (Release 8)."
[34] S. Sesia, I. Toufik, and M. Baker, *LTE–The UMTS Long Term Evolution: From Theory to Practice*, Wiley, 2nd Edition, 2011.
[35] E. C. Ifeachor and B. W. Jervis, *Digital Signal Processing: A Practical Approach, 2nd Edition*, Prentice Hall, 2002.
[36] ETSI, "Radio broadcasting systems; Digital audio broadcasting (DAB) to mobile, portable and fixed receivers," EN 300 401, 2004.
[37] K. Amiri, Y. Sun, P. Murphy et al., "WARP, A unified wireless network testbed for education and research," 2007 IEEE International Conference on Microelectronic Systems Education, Jun. 2007.
[38] S. Malkowsky, K. Nieman, Z. Miers et al., "A flexible 100-antenna testbed for Massive MIMO," 2014 *IEEE GLOBECOM Workshops*, Dec. 2014.
[39] J. Vieira, C. Shepard et al., "Argos: Practical many-antenna base stations," *Proc. 18th Annu. Int. Conf. Mobile Comput. Netw.*, pp. 53–64, 2012.
[40] M. Petermann et al., "Multi-user pre-processing in multi-antenna OFDM TDD systems with non-reciprocal transceivers," *IEEE Trans. Commun.*, vol. 61, no. 9, pp. 3781–3793, Sept. 2013.
[41] R. Rogalin et al., "Scalable synchronization and reciprocity calibration for distributed multiuser MIMO," *IEEE Trans. Wireless Commun.*, vol. 13, no. 4, pp. 1815–1831, Apr. 2014.
[42] D. Liu, W. Ma, S. Shao, Y. Shen, and Y. Tang, "Performance analysis of TDD reciprocity calibration for massive MU-MIMO systems with ZF beamforming," *IEEE Commun. Lett.*, vol. 20, no. 1, Jan. 2016.
[43] 3GPP RAN1 R1-1608809, "Over the air calibration for channel reciprocity in NR MIMO," Fujitsu.
[44] 3GPP RAN1 R1-1609903, "Partial reciprocity based UL MIMO schemes," InterDigital Communications.
[45] 3GPP RAN1 R1-1702603, "OTA calibration for multi-TRP transmission," Qualcomm.