

Figure 2 GPP base band resource pool structure

Analysis of real-time signal processing based on GPP is shown in figure 3.

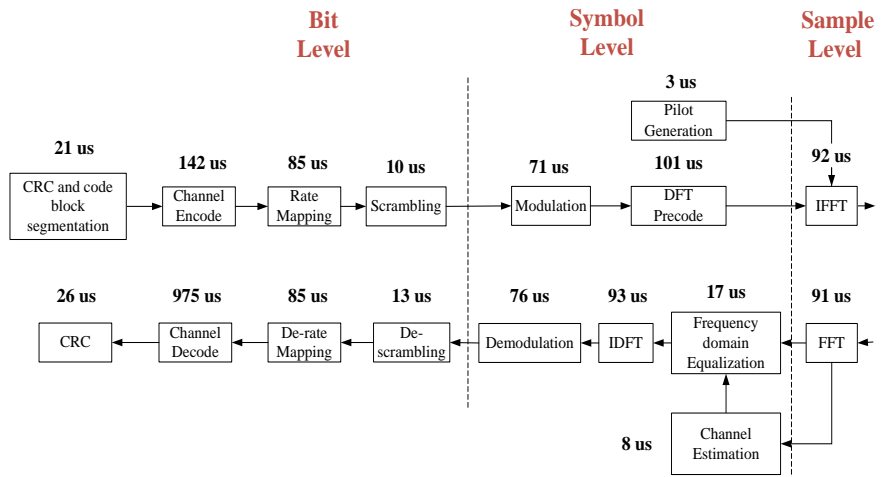


Figure 3 instances of LTE uplink

D. Connection technology of BBU and RRU

The network structure of C-RAN makes a new challenge to connection technology between BBU and RRU, the interface between them has the following several ways.

(1) Fiber optic drives directly. BBU and RRU in every site use connection of fiber point to point, using common light module, optical fiber demand is affected by RRU cascade ability and interface data, because of linear increment of sites, this connection is suitable for optical fiber resource-rich regions or areas like small-scale C-RAN gathering.

(2) The colourful drives directly. The colourful directly driving way belongs to the compromise of wavelength division technology and optical fiber, RRU and BBU use the colourful light module, and access to DADM, the maximum coarse direct-drive can support 18 waves, the maximum tight direct-drive can support 80 waves, numbers of optical fiber are not change along with increment numbers of sites, only two optical fibers can make a network. Without new transmission equipment, so the cost is low.

(3) Wavelength division. Using coarse wavelength division or dense wavelength division to carry CPRI data, it is needed to add wavelength division equipment at each site, the optical fiber resource utilization is high, but this way is limited by time delay of equipment and partial wave cascade loss, so it is mainly suitable for sharing business with others.

III. BASE BAND RESOURCE ALLOCATION METHOD BASED ON DYNAMIC PROJECTION OF BBU-RRU

A. Model of dynamic projection

According to its using resource, one BBU can support one or more RRU (the same color in figure 4), logical connections between the BBU and RRU can change, RRU does not belong to certain BBU, but it can be dynamically projected to the suitable BBU (RRU belongs to the same BBU is marked by dashed lines after changing mapping in figure 4).

BBU focuses on base band pool, it can be connected by fiber or strengthened X_2 (interface X_2 is used to realize interconnection between the eNB , it is equal to interface among base stations, consists of users plane X_2 and control plane X_2). RRU is the remote radio unit, which is located in the traditional station, it only can launch signals without signal processing. BBU and RRU are

connected by high-speed fiber-optic interfaces.

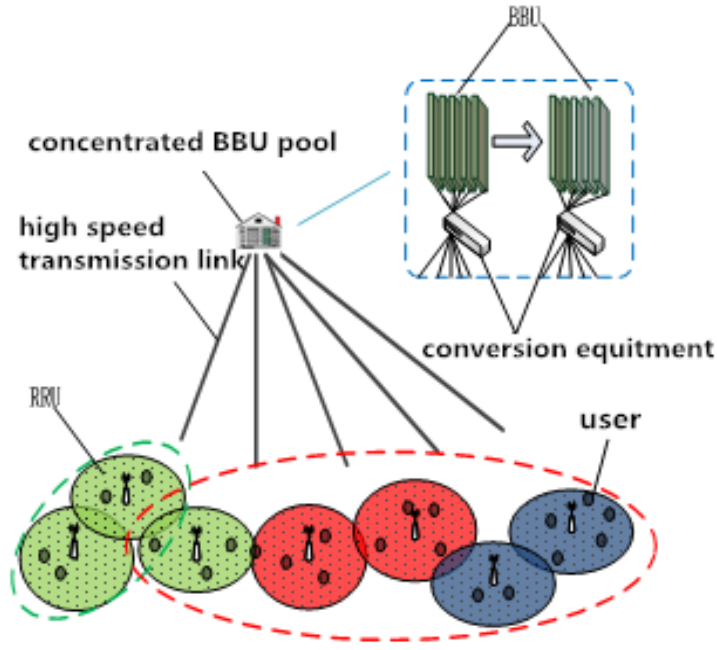


Figure 4 Dynamic projection model

B. System model

C-RAN base band resource pool contains a BBU collection with a number of m , each BBU consists of a "cluster" ("clusters" is included by several virtual machines and DSP (FPGA) chips).

$$\mathbf{B} = \{b_1, b_2, \dots, b_i, \dots, b_n\} \quad (1)$$

Assume calculation resource capacity of the BBU in the base band pool is collection C .

$$\mathbf{C} = \{c_1, c_2, \dots, c_j, \dots, c_m\} \quad (2)$$

BBU has two types of states, which are open and closed, and expressed by decision variable ω_j^t .

When at time of T , $\omega_j^t = 1$ represents BBU is open, $\omega_j^t = 0$ represents the BBU is closed,

$j \in \{1, 2, \dots, m\}, t \in \{1, 2, \dots, T\}$. Time-overhead exists at the time of BBU opening and closing, assume one time slot can only open (or close) one BBU.

In this model, cellular networks composed of n communities are considered, RRU is located in the center of each cellular community, regardless of the interference between RRU, RRU can be expressed by collection R . It is shown below.

$$\mathbf{R} = \{r_1, r_2, \dots, r_i, \dots, r_n\} \quad (3)$$

Matrix $\mathbf{P}_{n \times m}^t$ can represent dynamic changes of projection between BBU and RRU, θ_{ij}^t is decision variable in the matrix, at the time of T , r_i is projected to b_j when $\theta_{ij}^t = 1$, while, $\theta_{ij}^t = 0$ is on the contrary, in this matrix, $i \in \{1, \dots, n\}, j \in \{1, \dots, m\}, t \in \{1, \dots, T\}$.

$$\mathbf{P}_{n \times m}^t = \begin{bmatrix} \theta_{11}^t & \theta_{12}^t & \dots & \theta_{1m}^t \\ \theta_{21}^t & \theta_{22}^t & \dots & \theta_{2m}^t \\ \vdots & \vdots & \ddots & \vdots \\ \theta_{n1}^t & \theta_{n2}^t & \dots & \theta_{nm}^t \end{bmatrix} \quad (4)$$

In the model, users are randomly distributed in the community, at all time slots, number of users in every community can be represented by collection U^t at time of t . User traffic approaches according to the current time slot, total traffics have a certain regularity (day-night variability). Computing resources of every user can be expressed by matrix $\mathbf{K}_{n \times s^t}^t$. The needed computing resources of user l of r_i is obtained by k_{il}^t at time of t , in this matrix, $k_{i,l}^t \geq 0$, $s^t = \max(u_1^t, u_2^t, \dots, u_i^t, \dots, u_n^t)$, $i \in \{1, \dots, n\}, l \in \{1, \dots, s^t\}$.

$$\mathbf{U}^t = \{u_1^t, u_2^t, \dots, u_i^t, \dots, u_n^t\} \quad (5)$$

$$\mathbf{K}_{n \times s^t}^t = \begin{bmatrix} k_{11}^t & k_{12}^t & \dots & k_{1s^t}^t \\ k_{21}^t & k_{22}^t & \dots & k_{2s^t}^t \\ \vdots & \vdots & \ddots & \vdots \\ k_{n1}^t & k_{n2}^t & \dots & k_{ns^t}^t \end{bmatrix} \quad (6)$$

Taking operation time of traffic into account, traffic of the same user may appear at continuous time slot. Assume the maximum delay that users can tolerate is D time-slots, if the traffic arrives at the current time slot and can not be conducted, then it will be delayed to the next one, if the extension one is D time-slots, it will be discarded. Therefore, the user's traffic has three states, first of all, the new traffic at the time slot x_t ; second, the extension traffic x'_{t-d} ($0 \leq d \leq D-1$) that can not be carried out before; the third, expired traffic x''_{t-d} ($d \geq D$), then traffic in the current time slot is $y = x_t + x'_{t-d}$.

Executing time of users can be represented by a matrix $\mathbf{TIME}_{n \times s^t}^t$, The needed processing time of

user l of r_i is obtained by e_{il}^t at time of t , $e_{il}^t \geq 0$, $s^t = \max(u_1^t, u_2^t, \dots, u_i^t, \dots, u_n^t)$, $i \in \{1, \dots, n\}$, $l \in \{1, \dots, s^t\}$.

$$\mathbf{TIME}_{n \times s^t}^t = \begin{bmatrix} e_{11}^t & e_{12}^t & \cdots & e_{1s^t}^t \\ e_{21}^t & e_{22}^t & \cdots & e_{2s^t}^t \\ \vdots & \vdots & \ddots & \vdots \\ e_{n1}^t & e_{n2}^t & \cdots & e_{ns^t}^t \end{bmatrix} \quad (7)$$

Using the above model, at a time of t , the average resource utilization target function of BBU is

$$\bar{\eta}_t = \sum_{j=1}^m \sum_{i=1}^n \sum_{l=1}^{s^t} k_{il}^t \cdot \theta_{ij}^t / \sum_{j=1}^m w_j^t \cdot c_j, \text{ at each moment, average resource utilization of the BBU reaches}$$

the maximum when meeting restrictions.

$$\max \frac{1}{T} \lim_{T \rightarrow \infty} \sum_{t=1}^T \bar{\eta}_t \quad (8)$$

$$s.t. \quad \sum_{i=1}^n \sum_{l=1}^{s^t} k_{i,l}^t \cdot \theta_{ij}^t \leq c_j \quad \forall t, \quad \forall j \quad (9)$$

$$\sum_{j=1}^m \theta_{ij}^t \leq 1 \quad \forall i, \quad \forall t \quad (10)$$

$$\sum_{j=1}^m w_j^t \geq 1 \quad (11)$$

Constraint (9) shows that computing resources of users should be smaller than its system capacity at any time slots on any BBU; the constraint (10) shows one RRU mostly can be projected to one BBU at any time slots; constraint (11) can ensure services of users (Service Level Agreement, SLA for short).

This problem is equal to the number of the minimum BBU.

$$\min \quad z = \sum_{j=1}^m w_j^t \quad (12)$$

$$s.t. \quad \sum_{i=1}^n \sum_{l=1}^{s^t} k_{i,l}^t \cdot \theta_{ij}^t \leq c_j \quad \forall t, \quad \forall j \quad (13)$$

$$\sum_{j=1}^m \theta_{ij}^t \leq 1 \quad \forall i, \quad \forall t \quad (14)$$

$$w_j^t = \{0,1\} \quad \theta_{ij}^t = \{0,1\} \quad (15)$$

The meaning of constraints (13) and (14) are similar to the above.

C. Design of the base band resource allocation algorithm

FFD (First Fit Descending) pre-distribution algorithm designing is proposed in this paper. First, $queue_{1 \times n}^t$ is task queue that needs to be conducted in one time slot, then it will be entered, according to the traffic of each RRU in $queue_{1 \times n}^t$, RRU should be conducted with descending order; second, all RRU that cannot establish projection with BBU are composed to a collection, $capacity_{1 \times m}^t$ is the remaining resources matrix of BBU; the third, we check RRU to see if it has already established projection relationship successfully with BBU, then we should end the algorithm if it is, rather not, we should continue the algorithm. If computing resource demand of RRU at current time slot is less than remaining resources of BBU, which shows BBU and RRU have determined projection relationship, then we calculate remaining resources again in order to distribute it successfully, then it's time for the next RRU to assign according to the third step; fourth, if computing resource demand of RRU at current time slot is greater than remaining resources of BBU, then the next BBU should be opened, assignment is the same with the third step until the end.

Traffic distribution processing is designed, for traffic in waiting queue, a distribution variable $flag$ is set up for RRU with the original value of zero, $flag=1$ shows that RRU was successfully allocated on the BBU, otherwise is not successful; the second step, check if BBU and RRU exist projection relationship according to the BBU-RRU projection matrix, if it exists, then $flag=1$, traffic in waiting queue will be transferred, the corresponding location in waiting queue will be wiped clean, if it does not exist, skip to next BBU and repeat the second step. Finally, the value of $flag$ should be checked, if $flag=0$, then it will be delayed and transferred to a behind position in waiting queue, the corresponding position in waiting queue will be wiped, if there is also RRU without assignment, then we return to the first step or task delayed for D time slots in waiting queue will be transferred to dropping queue, the corresponding position in waiting queue will be wiped.

IV.CALCULATION AND ANALYSIS

In this section, by Matlab simulation, specific performance of the method proposed in this paper can be verified. Matlab is launched by MathWorks software company, it is a set of software for numerical analysis and calculation, it contains matrix operations, numerical analysis, graphics, programming calculation and others, for its powerful function, the majority of users can use it as a practical tool in scientific and engineering problems, analysis calculation and programming utility, it is also a new generation of software development platform with all language features and characteristics.

This paper uses its powerful data processing function and rich graphics to realize the research on the properties of resource allocation method. First, allocation algorithm is described with the Matlab language, and then program is written to simulate business of users, this program is as input allocation algorithm program, finally, the results of the distribution (like resource utilization, using numbers of BBU, executing business and discarded business, etc.) need to be compared and analyzed.

A. The simulation parameters

Wireless communication has obvious periodicity, for example, traffic during daytime is higher than the late-night. Assume the arrival traffic obeys a rule, we use a equation to explain the rule, it is described as follows.

$$\lambda(t) = A \sin \left[\frac{2\pi}{T}(t-t') \right] + B \sin \left[\frac{4\pi}{T}(t-t'') \right] + C \quad (16)$$

In the equation, T is the traffic period, t is the time slot, t', t'' are fixed values. According to the algorithm, in order to simulate easily, some parameters in the simulation are simplified while assuring the correctness of simulation. Execution time of traffic is randomly generated, which is not qualified in the simulation, parameters are shown in table 1.

Table 1 simulation parameter

name of parameter	parameter value
Simulation period T1(short)	288
simulation period T2(long)	1440

capacity of BBU c	1000
number of BBU m	28
number of RRU n	28
the maximum time delay business users can tolerate D	3
required computing resources of per user traffic k	1
The number of cycle	5
t'	144
t''	84

B. Analysis of results

a. Numbers of opening BBU changes following traffic analysis

In figure 5, it is a normalized curve that shows the relationship between traffic arriving in every time slot and numbers of BBU opened in the corresponding time slot. From the overall trend, number of BBU increases with the increment of business traffic, and decreases with the decline of business traffic. Numbers of BBU in every time slot are highly relevant with business traffic, which shows that the proposed method can adjust opening numbers of BBU in real-time according to business traffic changing, so the goal of designing algorithm is achieved. Meanwhile, curve of numbers of BBU appears a ladder-shape, which shows that numbers of BBU changing are later than business traffic changing, this is because it takes time to open or close BBU(open or close one BBU needs a time slot in simulation), opening numbers of BBU may remain constantly at some time slots. When business traffic arrives, using the method put forward in this paper to calculate the best numbers of BBU at current time slot, then we take numbers of BBU at last time slot into account to decide the numbers at current time slot, then the traffic at current time slot is distributed according to projection matrix.

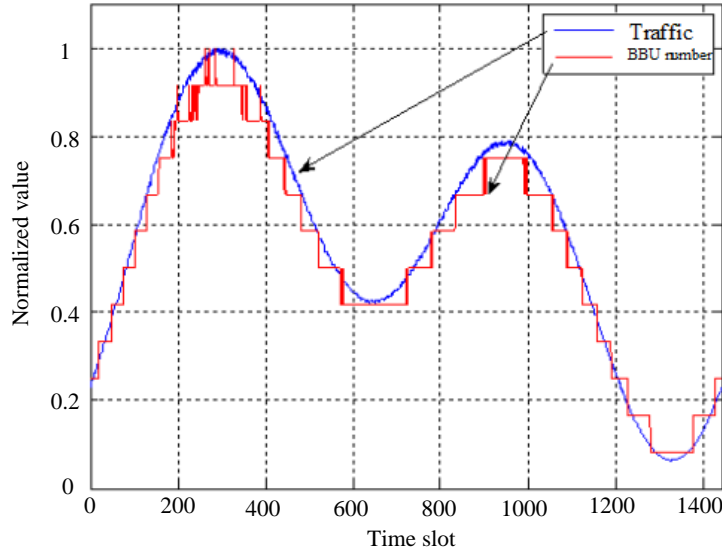


Figure 5 normalized curve of traffic and numbers of BBU

b. Average resource utilization of opening BBU analysis

In figure 6, we make a resource utilization comparison between the base band resource allocation method put forward in this paper and traditional distribution mechanism, the traditional method is one by one projection in community. From the figure, the blue line represents base band resource distribution under C-RAN, the red line represents the traditional distribution mechanism. Average resource utilization of BBU marked in red at every time slot maintains at a high level (more than 80%), average resource utilization of BBU is low only when traffic is low, its value is close to 50%; computing resource utilization in one by one projection distribution mechanism in traditional community (distributed community, D-RAN for short) changes along with traffic changing, computing resource utilization is high when traffic is high, this figure demonstrates that method put forward above can improve the base band resource utilization effectively. In the figure, there are about 1300 time slots when base band computing resource utilization is significantly lower than the average level, by checking the simulation data, only one BBU is opened. This is consistent with the theoretical analysis, in order to ensure service levels of customer (SLA), at least one BBU should be opened.

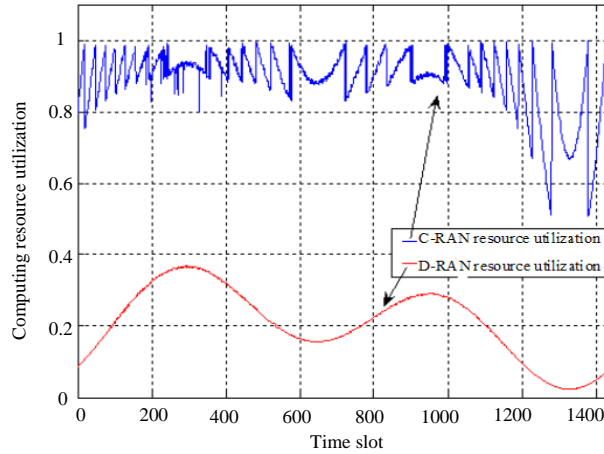


Figure6 resource utilization comparison curve

C. Operating and discard traffic analysis

Figure 7 represents operating and discarded business traffic analysis. Due to time-overhead during BBU opening when traffic is in rise stage, current time slot has no capacity to open enough numbers of BBU to process current traffic, and it cannot operate all traffics, which means a part of traffics will be delayed to the next time slot, if traffic of users can not finish operation in time of maximum delay when taking maximum delay that users can tolerance into account (set for 3 time slots in simulation), then it will be discarded. As can be seen from the diagram, only one or two time slots are discarded, this influence is very small in huge traffic system. In conclusion, the proposed distribution method can guarantee the quality of communication.

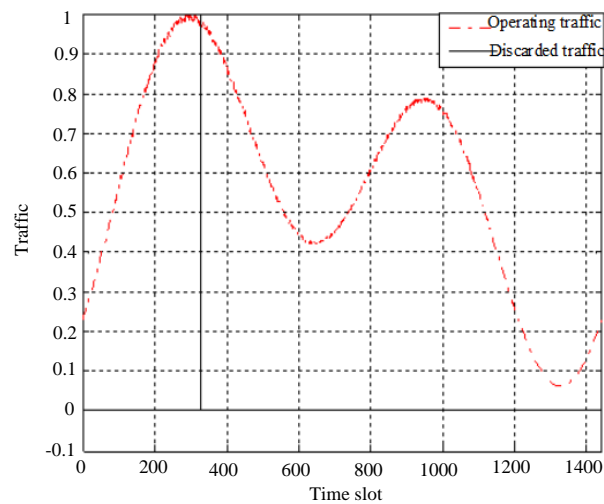


Figure 7 operating and discarding traffic normalized curves

D. Analysis of energy consumption

Figure 8 shows a histogram that represents the energy conservation, the BBU is all open in traditional community architecture. In simulation, assume consumption of electric is a constant during BBU opening in the paper, and it is not related to the processing traffic, unopened BBU does not consume electricity. Using the ratio of opening numbers of BBU at every time slot and the total numbers of BBU, energy consumption as percentage of energy consumption before resource allocation can be expressed by that ratio approximately, to some extent, it can illustrate the effectiveness of this method because this percentage is related to the capacity of the system, in this article, the system capacity is greater than the maximum traffic. From figure 1, the maximum resource utilization is 40% when the system is busy, therefore, the percentages calculated by this method may vary with the capacity of the system, this method can effectively save energy, with its low-difficulty calculation and low complexity, which indicates its feasibility.

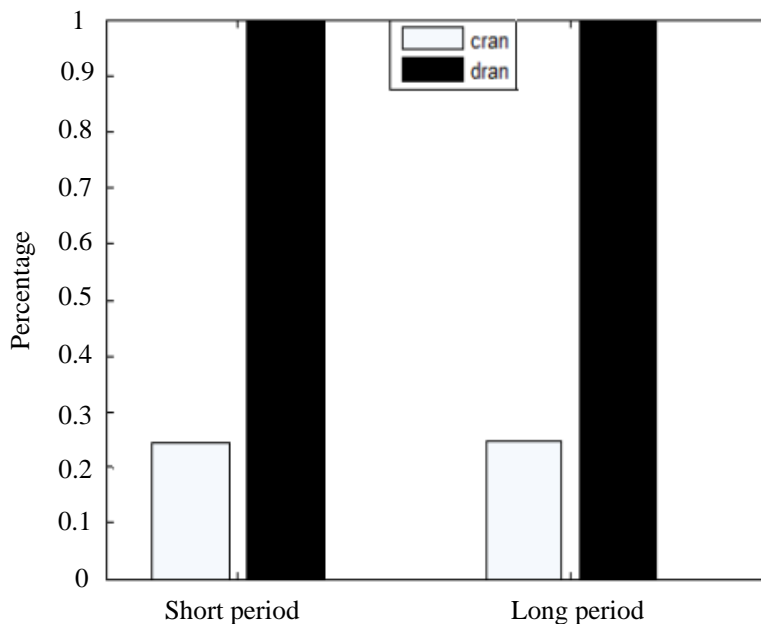


Figure 8 energy consumption comparison of C-RAN and D-RAN

In order to further illustrate this method on reducing energy consumption, three typical time slots are chosen in the following figure, which are at peak(time slot is 286), at the second peak(time slot is 950) and the average level(time slot is 750), figure 9 shows each of their energy consumption.

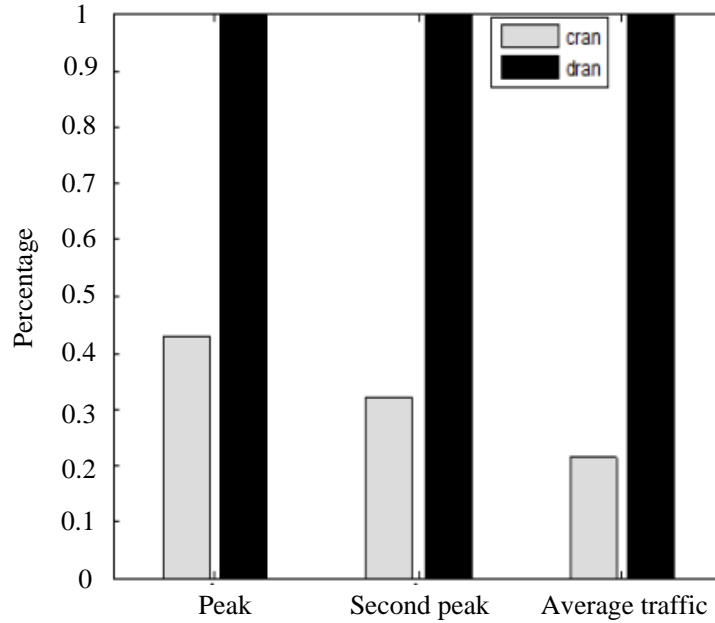


Figure 9 energy consumption contrasts of the three typical time slots

V. CONCLUSIONS

How to use the base band resource effectively and improve resource utilization become a hot research topic nowadays. C-RAN is a new access network architecture, this paper focuses on the base band resource allocation problem, introduces C-RAN technology. Mathematical model is set up. Combined with the C-RAN method, related algorithm steps are designed. The base band resource allocation method is described systematically in this paper, according to the scene proposed, base band resource is distributed properly taking heuristic order-decreasing first-fit algorithm as the pre-allocation algorithm into account. In order to understand simply, the algorithm steps are explained. Through simulation, four parts are analyzed, they are opening numbers of BBU changing rule, the average resource utilization, operating and discard traffic and energy consumption are assessed. In conclusion, the method proposed has characteristics of low complexity, performance stability, low packet lost and power consumption, while ensuring the quality of communication, it provides theoretical basis and practical design factors for the utilization of resources.

VI. REFERENCES

- [1] B. Sotomayor, K. Keahey, and I. Foster, "Combining batch execution and leasing using virtual machines", HPDC'08 Proceedings of the 17th international symposium on High performance distributed computing, pp. 87-96, 2008.
- [2] N. Q. Hung, N. Thoai and N. T. Son, "Performance constraint and power-aware allocation for user requests in virtual computing lab", Journal of Science and Technology, Vol.49, No.4, pp.383-392, 2011.
- [3] J. R. Correa, M. R. Wagner, "LP-based online scheduling: from single to parallel machines", Math Program, Vol.119, pp.109-136, 2009.
- [4] A. Schulz, "Scheduling unrelated machines by randomized rounding", SIAM Journal on Discrete Mathematics, Vol. 15, No. 4, pp.450, 2002.
- [5] I. Al Azzoni, "Power-Aware Linear Programming Based Scheduling for Heterogeneous Computer Clusters", Future Generation Computer Systems, Vol. 28, No. 5, pp.745-754, 2011.
- [6] B. Speitkamp, M. Bichler, "A Mathematical Programming Approach for Server Consolidation Problems in Virtual Data Centers", IEEE Transactions on Services Computing, Vol.3, No.4, pp.266-278, 2010.
- [7] A. Beloglazov, J. Abawajy and R. Buyya, "Energy-aware resource allocation heuristics for efficient management of data centers for Cloud computing", Future Generation Computer Systems, Vol. 28, No. 5, pp.755-768, 2012.
- [8] Xu P Y, Qian Z Z, "A relaxed co-scheduling method of virtual CPU on xen virtual machine", Computer Science, No.7, pp.297-301, 2012.
- [9] Shi X L, Xu K, "Utility maximization model of virtual machine scheduling in cloud environment", Computer Journal, No.2, pp.252-262, 2013.
- [10] Li J H, Li S F, "Research on the resource scheduling method of virtual machines under the cloud computing environment", Telecommunication Science, No.4, pp.78-82, 2013.
- [11] Xu B, Zhao C, "Virtual Machine Resource Scheduling Multi-objective Optimization in Cloud Computing", System Simulation Journal, No.3, pp.592-595, 2014.
- [12] Y. Lin, L. Shao, Z. Zhu, Q. Wang, *et al*, "Wireless network cloud: Architecture and system requirements", IBM Journal of Research and Development, Vol.54, No.1, pp.1-12, January-

February 2010.

- [13]Wang D W,Liu X M, “A resource scheduling strategy for cloud computing platform of power system simulation based on dynamic migration of virtual machine”,Automation of Electric Power Systems, No.12, pp.97-105, 2015.
- [14]Lou S T,Yao R Y.“MATLAB 7.x programming language”.Xidian University Journal, Vol 2006, pp.25-78.
- [15]Xiao P,Liu D B,“An virtual machine scheduling algorithm based on energy consumption ratio model in cloud computing”,Journal of Electronics, No.2, pp.305-311, 2015.
- [16]Jin H,Zhong H L,“Virtual Machine VCPU Scheduling in the Multi-core Environment: Issues and Challenges”,Computer Research and Development, No.7, pp.1216-1224, 2011.
- [17]Lei Q Y, Zhang Z Z,“5G radio access network architecture based on C-RAN”, Telecommunication Science, No.1, pp.112-121, 2015.
- [18]Wang M, “Transmission technology under C-RAN”, China Technology Investment, No.33, pp.18-19, 2012.
- [19]Fang Y Q,Tang D H,Ge J W,“Research on schedule strategy based on dynamic migration of virtual machines in cloud environment”, Microelectronics and Computers, No.4, pp.45-48, 2012.
- [20]Li M F,Bi J P,Li Z C,“Resource-Scheduling-Waiting-Aware Virtual Machine Consolidation”,journal of software, No.7, pp.1388-1402, 2014.