Energy-Delay Tradeoffs of Base Stations in Cloud-Based Cellular Networks

Tao Zhao, Jian Wu, Sheng Zhou, and Zhisheng Niu

Network Integration for Ubiquitous Linkage and Broadband(NiuLab)
Tsinghua University, Beijing, China

Outline

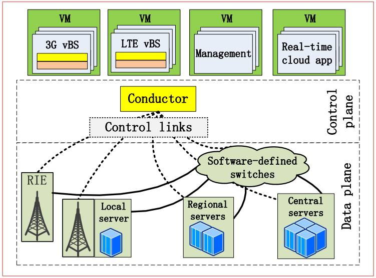
1. **Background**
2. **Energy-Delay Tradeoff**
   - System Model
   - Optimization Problem and Solution
   - Numerical Results
3. **Summary**

5G Challenges

- **Requirements & Challenges**
  - Massive data: 1000× capacity → Efficiency
  - Massive connections: U-shape traffic → Flexibility
  - Innovative applications: IT → operators → Sustainability
- **Conventional cellular architectures**
  - Peak load based provision → resource waste
  - Distributed BS decision → hard to coordinate or sleep
  - Tightly coupled software & hardware → hard to upgrade and deploy new services

Cloud-Based Cellular Architectures [1]–[3]

- **CONvergence of Cloud and cEllulaR sysTems**
  - Centralized cloud computing platform
  - Software defined services
  - Open IT technologies



---

[1] Y. Lin, L. Shao, Z. Zhu, *et al.*, "Wireless network cloud: architecture and system requirements," *IBM Journal of Research and Development*, vol. 54, no. 1, 4:1–4:12, 2010.

[2] China Mobile Research Institute, "C-RAN: the road towards green RAN," , White Paper, Dec. 2013, Version 3.0.

[3] J. Liu, T. Zhao, S. Zhou, *et al.*, "CONCERT: a cloud-based architecture for next-generation cellular systems," *IEEE Wireless Commun. Mag.*, 2014, To be published. eprint: arXiv:1410.0113.

Virtual Base Stations

- **Virtual BS (VBS)**
  - BS function: software in virtual machine
  - BBU (base band unit): commodity servers
  - RRH (remote radio head): universal hardware
- **Focus: energy-delay tradeoffs of BSs**
  - Total power consumption: BBU and RRH
  - Queueing delay
  - Load aware resource adaptation
- **Motivation**
  - What does the energy-delay tradeoff relationship look like for VBS?
  - How is the energy performance compared with conventional BS under delay constraints?

Outline

Power Consumption Model

- **Computational resource aware model**
  - RRH:

$$P_{\mathrm{R}} = \frac{P_{\mathrm{out}}}{\eta} + P_{\mathrm{RF}} \tag{1}$$

  - BBU: N_{c} CPU cores, CPU speed s, CPU load ρ_{c}

$$P_{\mathrm{B}} = N_{\mathrm{c}}(P_{\mathrm{Bm}} + \Delta_{P_{\mathrm{B}}}\rho_{\mathrm{c}} s^{\beta}) \tag{2}$$

$$\Delta_{P_{\mathrm{B}}} = (P_{\mathrm{BM}} - P_{\mathrm{Bm}})/s_0^{\beta} \tag{3}$$

$$\rho_{\mathrm{c}} = \frac{f(r)}{N_{\mathrm{c}} s} = \frac{c_0 + \kappa r}{N_{\mathrm{c}} s} \tag{4}$$

  - BS:

$$P = \begin{cases} P_{\mathrm{B}} + P_{\mathrm{R}}, & 0 < P_{\mathrm{out}} \leq P_{\max} \\ P_{\mathrm{sleep}}, & P_{\mathrm{out}} = 0 \end{cases} \tag{5}$$

Queueing Model

- **BBU pool**
  - One VBS on a server with N_c active CPU cores with speed s
- **Queueing model: M/G/1 Processor Sharing (PS)**
  - Flow arrival at BS: rate λ, average file size L
  - Data transmission rate r bps $(r_0 = 0; r_n = r, n > 0)$
  - Traffic load: $\rho = \lambda L / r$
  - Average queue length: $\mathbb{E}\{n\} = \frac{\rho}{1-\rho} = \frac{\lambda L}{r - \lambda L}$
  - Average delay: $\mathbb{E}\{D\} = \mathbb{E}\{n\}/\lambda$
- **Base station sleeping**
  - Cycle: $T_\mathrm{c} = T_\mathrm{a} + T_\mathrm{s}$
  - Switching cost: E_sw
  - Average power consumption:

$$\mathbb{E}\{P\} = p_\mathrm{active}(P_\mathrm{B} + P_\mathrm{R}) + p_\mathrm{sleep} P_\mathrm{sleep} + \frac{2E_\mathrm{sw}}{\mathbb{E}\{T_\mathrm{c}\}} \quad (6)$$

Cell Coverage Model

- **3GPP propagation model**
  - Path loss $L(d)$
  - Noise factor F
  - System bandwidth W
  - Downlink SINR

$$\mathrm{SINR}(d) = gP_{\mathrm{out}} = \frac{P_{\mathrm{out}}}{L(d)FN_0W} \tag{7}$$

- **Data rate**
  - User: assumed at the cell edge
  - Sum data rate at BS

$$r = W\log_2(1 + gP_{\mathrm{out}}) \tag{8}$$

Outline

1. Background
2. **Energy-Delay Tradeoff**
   - System Model
   - Optimization Problem and Solution
   - Numerical Results
3. Summary

Problem Formulation

- **Objective: system costs**
  - Power consumption
  - Average queue length
  - Weighting factor
- **Decision variables:**
  - Data transmission rate
  - Number of CPU cores

Optimization Problem

$$\min_{r, N_{\mathrm{c}}} \quad z = \mathbb{E}\{P\} + \alpha \mathbb{E}\{n\} \qquad (9)$$

Solving the Problem

- **Optimal rate**
  - Optimal rate satisfies:

$$\Omega\left(\frac{\alpha g\eta}{\mathrm{e}}\left(\frac{r^*}{r^*-\lambda L}\right)^2+\frac{g\eta P_{\mathrm{s}}-1}{\mathrm{e}}\right)=\frac{r^*\ln 2}{W}-1 \qquad (10)$$

  - where $\Omega(\cdot)$ is the principal branch of Lambert W function,

$$P_{\mathrm{s}}=P_{\mathrm{o}}-P_{\mathrm{sleep}}-2\lambda E_{\mathrm{sw}}, \qquad (11)$$

$$P_{\mathrm{o}}=N_{\mathrm{c}}P_{\mathrm{Bm}}+\Delta_{P_{\mathrm{B}}}c_0 s^{\beta-1}+P_{\mathrm{RF}}, \qquad (12)$$

Energy-Delay Tradeoff

- **Proposition**
  - There exists the unique energy optimal rate $r_{\rm e}^*$ when:

$$\lambda < \frac{P_{\rm o} - P_{\rm sleep}}{2E_{\rm sw}}, \tag{13}$$

$$L < \frac{W}{\lambda \ln 2}\left[\Omega\left(\frac{g\eta P_{\rm s} - 1}{\rm e}\right) + 1\right] \tag{14}$$

  The corresponding energy optimal rate is:

$$r_{\rm e}^* = \frac{W}{\ln 2}\left[\Omega\left(\frac{g\eta P_{\rm s} - 1}{\rm e}\right) + 1\right]. \tag{15}$$

  - Otherwise, the average power consumption is monotonically decreasing with the average delay.
  - In both cases, when the average delay goes to infinity, the average power consumption approaches
    $P_{\rm o} + \kappa \Delta_{P_{\rm B}} s^{\beta-1} \lambda L + \frac{2^{\frac{\lambda L}{W}} - 1}{g\eta}$.

Impact of Computational Resources

- **Impact of N_c**
  - Max rate $r_{\mathrm{M}}(N_{\mathrm{c}}) = \frac{N_{\mathrm{c}} s - c_0}{\kappa}$
  - Monotonicity: $\frac{\partial z}{\partial N_{\mathrm{c}}} = P_{\mathrm{Bm}} > 0$, $\frac{\partial r^*}{\partial N_{\mathrm{c}}} > 0$,

Joint Optimization Algorithm

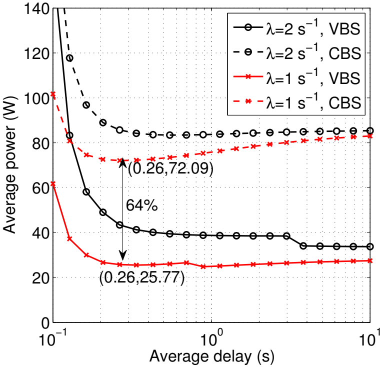
Set N_{cM}, $N_{\mathrm{c}} \leftarrow 1$, $S \leftarrow \Phi$
while $N_{\mathrm{c}} \leq N_{\mathrm{cM}}$ **do**
    $\hat{r}(N_{\mathrm{c}}) \leftarrow \mathrm{argmin}_r \, z(r, N_{\mathrm{c}})$
    if $\hat{r}(N_{\mathrm{c}}) \leq r_{\mathrm{M}}(N_{\mathrm{c}})$ **then**
        $S \leftarrow S \cup \{(\hat{r}(N_{\mathrm{c}}), N_{\mathrm{c}})\}$
        Break out of the loop
    else
        $S \leftarrow S \cup \{(r_{\mathrm{M}}(N_{\mathrm{c}}), N_{\mathrm{c}})\}$
        $N_{\mathrm{c}} \leftarrow N_{\mathrm{c}} + 1$
    end if
end while
return $(r^*, N_{\mathrm{c}}^*) = \mathrm{argmin}_{(r, N_{\mathrm{c}}) \in S} \, z(r, N_{\mathrm{c}})$

Outline

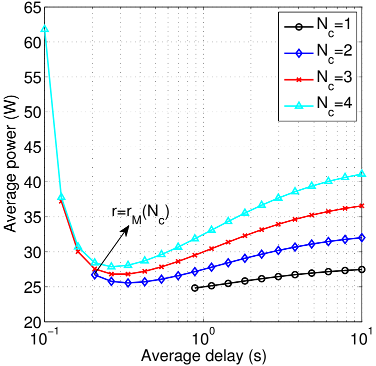
1. Background
2. **Energy-Delay Tradeoff**
   - System Model
   - Optimization Problem and Solution
   - Numerical Results
3. Summary

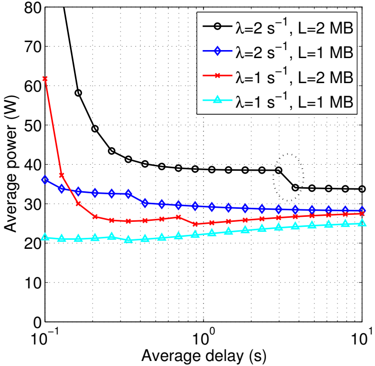
Numerical Results

- **Parameters**
  - BBU: commodity servers
  - Radio: LTE R11
- **Comparison**
  - Conventional BS (CBS): EARTH model



- **Energy-delay tradeoffs**





Summary

- Proposed a computational-resource-aware energy consumption model for VBSs in cloud-based cellular networks.
- Derived the explicit form of the optimal data rate, and observed the opportunity to achieve energy savings and reduce the average delay simultaneously.
- Proposed an efficient algorithm to jointly optimize the data rate and the number of CPU cores.

Network Integration for Ubiquitous Linkage and Broadband