

# Internet das Coisas Médicas: Uma Avaliação de Desempenho Focando em Prioridades de Requisições

Lucas Santos, Brena Santos e Francisco Airton Silva

Laboratório de Pesquisas Aplicadas a Sistemas Distribuídos (PASID)  
Universidade Federal do Piauí (UFPI)  
Picos, PI, Brasil

{vinicius.lucas,brenamaia,faps}@ufpi.edu.br

**Abstract.** *The Internet of Medical Things (IoHT) is a subarea of the Internet of Things focused on the health context. IoT monitoring devices are equipped with sensors and actuators that can take a certain action and even save lives. However, there are medical situations that require more sophisticated computer systems that generate a large load of data and require real-time response. In this case, it is often necessary to evaluate possible distributed system architectures to support the sensor network. This work presents a queuing network model representing an IoHT architecture with priority requests. The model is highly configurable and can assist technology professionals in the context of IoHT.*

**Resumo.** *A Internet das Coisas Médicas (IoHT) é uma subárea da Internet das Coisas voltada para o contexto de saúde. Dispositivos de monitoramento IoT são dotados de sensores e atuadores que podem tomar uma determinada ação e até salvar vidas. No entanto, existem situações médicas que precisam de sistemas computacionais mais sofisticados que geram uma grande carga de dados e requerem resposta em tempo real. Neste caso, muitas vezes se faz necessário avaliar possíveis arquiteturas de sistemas distribuídos para dar suporte à rede de sensores. Este trabalho apresenta um modelo de rede de filas representando uma arquitetura IoHT com prioridade de requisições. O modelo é altamente parametrizável e pode auxiliar profissionais de tecnologia no contexto de IoHT.*

## 1. Introdução

A Internet das Coisas (IoT) é um dos grandes avanços atuais da tecnologia. A IoT é um paradigma adotado em diversas áreas práticas do cotidiano, tais como: segurança pública [Shafiq et al. 2020], linhas industriais [Khoa et al. 2020], educação [Mohammadian et al. 2020] e saúde [Uslu et al. 2020]. Particularmente, na saúde a IoT vem avançando de forma significativa provendo alternativas de monitoramento de sinais vitais das pessoas. A incorporação da IoT na área de saúde é comumente denominada de IoHT (Internet of Healthcare Things) [Yassein et al. 2019]. Sistemas que implementam a IoHT funcionam com a mesma ideia base de sistemas IoT tradicionais, porém, com um âmbito totalmente voltado para a medicina. Dispositivos inteligentes são utilizados de maneira que auxiliem no tratamento de pacientes em observação (por exemplo, marca-passo inteligente, medidor de glicose no sangue inteligente, etc.). Os dispositivos de um ambiente IoHT são responsáveis por monitorar e transmitir os dados relacionados aos pacientes diretamente para servidores remotos. As informações coletadas relacionadas aos

pacientes são então redirecionadas para médicos, enfermeiros ou cuidadores, otimizando o tratamento de pacientes em diferentes estados de saúde.

A IoHT obviamente apresenta grandes desafios. Na IoHT existe uma heterogeneidade significativa dos dispositivos conectados, o que coloca restrições na capacidade de armazenamento e processamento e resulta em desafios principalmente de desempenho [Yassein et al. 2019]. Estes desafios são mais visíveis em um contexto de saúde onde o monitoramento em tempo real e a tomada de decisões têm alta criticidade e podem resultar em mortes. Outros desafios comumente presentes neste contexto incluem interoperabilidade, protocolos de comunicação confiáveis e com eficiência energética, armazenamento de dados e gerenciamento de dados [de Morais Barroca Filho and de Aquino Junior 2017]. Uma solução comum para lidar com as limitações de armazenamento e processamento de dados é oferecer suporte a casos de uso de IoT com processamento em borda e principalmente na névoa. Recursos de névoa atuam como complementos em capacidade de processamento para os dispositivos locais, provendo maior confiabilidade e escalabilidade. Ademais, a névoa minimiza o tempo de solicitação-resposta de/para os aplicativos suportados [Iorga et al. 2018].

Portanto, quanto maior a criticidade de sistemas IoHT maior é a necessidade de realizar avaliações de vários requisitos não funcionais, como por exemplo relacionados a desempenho. No contexto de monitoramento de grande quantidade de pacientes em um hospital pode ser adicionado uma complexidade em relação à prioridade no tratamento de determinados dados. Em outras palavras, os dados podem ser processados remotamente (névoa por exemplo) de acordo com a prioridade da requisição. Em um contexto tão sofisticado e crítico se faz necessário adotar técnicas de avaliação avançadas antes mesmo da implantação dos sistemas. Modelos analíticos (ex: redes de filas, cadeias de Markov, redes de Petri) são alternativas viáveis para tal objetivo.

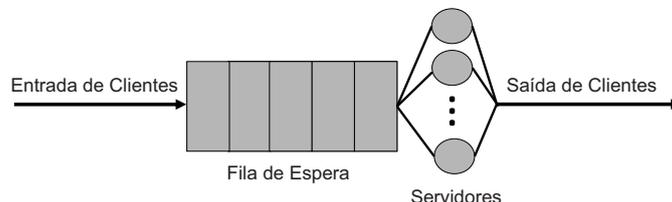
Este artigo usa redes de filas para avaliar um sistema IoHT usando critérios de prioridade entre as requisições que chegam ao sistema. O contexto de IoHT apresenta um deficit de trabalhos com temática de prioridades. O sistema engloba o monitoramento de um grande número de pacientes em diferentes níveis críticos de saúde. Os dados são processados em recursos de computação na névoa. As requisições são classificadas em duas prioridades (alta ou baixa). As requisições de alta prioridade devem ser as primeiras a serem concluídas, sendo postas em primeiro lugar nas filas de espera para atendimento. Quatro métricas de desempenho foram observadas: percentual de utilização, tempo médio de resposta, vazão de dados do sistema e taxa de descarte das requisições.

As seções seguintes estão organizadas da seguinte forma: A seção 2 apresenta um breve resumo sobre teoria de filas e métricas utilizadas neste artigo. A seção 3 apresenta alguns trabalhos relacionados. A seção 4 apresenta a arquitetura e modelagem de filas utilizadas neste artigo. A seção 5 apresenta os resultados das análises numéricas realizadas sobre o modelo. Por fim, a seção 6 traça algumas conclusões e trabalhos futuros.

## **2. Fundamentos da Teoria de Filas**

Uma fila é a implementação de uma lista de espera de trabalhos para realizar um serviço. Um ou mais servidores podem realizar esses serviços. Esse conjunto (fila + um ou mais servidores capazes de fornecer serviços) é denominado centro de atendimento. Se for feita uma analogia com o mundo real, pode-se imaginar um ou mais balcões de recepção que

são acessados por meio de uma fila. Todos os balcões funcionam em paralelo e prestam o mesmo serviço, independentemente da sua posição. Na teoria das filas, um centro de serviço é representado como na Figura 1 [Haragos and Cernazanu-Glavan 2012].



**Figura 1. Estação de atendimento (fila de espera + um ou vários servidores).**

Um modelo que usa um ou mais centros de serviço é chamado de modelo de filas. Várias métricas de desempenho são usadas para medir o desempenho de tal modelo. Abaixo, as métricas de desempenho mais críticas são apresentadas e eles podem caracterizar esse modelo de filas. O número de clientes em um sistema de tempo  $t$  é denotado por  $n_s$ . O valor esperado para o número de clientes que estão em um determinado momento no sistema é chamado de momento  $n^{th}$  sobre a origem de  $n$  e é calculado como segue:

$$E[n^k] = \sum_{i=0}^{\infty} i^k Prob\{n = i\} \quad (1)$$

Se for considerado que  $Prob\{n = i\}$  e  $E[n^k]$  são iguais às médias ao longo de um intervalo de tempo infinitamente longo, têm-se:

$$\lim_{s \rightarrow \infty} (Primeira\ vez\ quando\ os\ clientes\ estão\ no\ sistema) \quad (2)$$

$$E[n^k] = \lim_{s \rightarrow \infty} \frac{1}{S} \int_0^S [n_u]^k du \quad (3)$$

Cada cliente  $j$  que chega ao sistema deve gastar algum tempo antes de ser atendido. Este tempo é definido como tempo de resposta para o cliente  $j$  e é anotado como  $r_j$ . Se for considerado que  $Prob\{r \leq t\}$  e  $E[r^k]$  são iguais às médias ao longo de um infinito número de clientes (médias de clientes de longo prazo), a seguinte equação para a função de distribuição de probabilidade e o valor esperado é obtida:

$$Prob\{r \leq t\} = \lim_{j \rightarrow \infty} (Z) \quad (4)$$

A variável  $Z$  representa a fração dos primeiros  $J$  clientes a chegar cujo tempo de resposta é menor ou igual a  $t$ .

$$E[r^k] = \lim_{j \rightarrow \infty} \frac{1}{J} \sum_{j=1}^J r_j^k \quad (5)$$

Se o sistema for estável ( $E[n]$  e  $E[r]$  são números finitos), a vazão  $T$  deve ser igual à taxa de geração (TG) de clientes, onde  $T$  é definido como:

$$T = \lim_{s \rightarrow \infty} \frac{1}{S} (\text{Número de clientes que partiram nas primeiras unidades de tempo}) \quad (6)$$

O tempo que um servidor está ocupado é denominado de utilização da estação de atendimento. Se for denotado por  $b_s$ , o número de servidores ocupados no tempo  $s$ ,  $U$  (utilização) é definida como segue:

$$T = \lim_{s \rightarrow \infty} \frac{1}{S} \int_0^s b_u du \quad (7)$$

Neste trabalho, são usadas as quatro seguintes métricas de desempenho: tempo médio de resposta (MRT), taxa de descarte, utilização e vazão. A equação 8 define o tempo médio de resposta:

$$MRT = \frac{1}{\mu(1 - \rho)} \quad (8)$$

, onde  $\mu$  significa taxa de serviço e  $\rho$  representa taxa de uso do servidor. A equação 9 define a utilização dos componentes do sistema, onde  $\lambda_i = \frac{A_i}{T}$  representa a taxa de geração de dados (TG).

$$\rho = \frac{\lambda}{\mu} \quad (9)$$

### 3. Trabalhos Relacionados

Esta seção apresenta alguns trabalhos relacionados. A Tabela 1 resume um comparativo dos trabalhos relacionados sob quatro aspectos: métricas, contexto de IoHT, modelagem de filas, e critério de prioridade.

O primeiro critério de comparação se trata das **métricas** adotadas. Apesar de abrangentes, na maioria dos casos, os trabalhos não adotaram métricas fundamentais usadas no presente artigo. Apenas um trabalho ([El Kafhali and Salah 2018]) adotou o tempo médio de resposta, porém nenhum dos trabalhos coletados explorou as demais métricas aqui adotadas, que são: utilização de recursos, vazão de dados e taxa de descarte. A exploração conjunta de tais métricas é fundamental, pois há uma forte ligação entre elas. Por exemplo, o MRT pode ser baixo não pelo alto poder de processamento de um sistema, mas porque houve uma saturação de recursos, passando a existir descarte de dados.

O segundo critério de comparação refere-se ao **contexto de IoHT**. Todos os trabalhos focaram no contexto de IoHT. Apenas um artigo ([Adhikari et al. 2019]) não teve tal foco. O estudo de Adhikari et al. (2019) foi incluído nos trabalhos relacionados pois foi o único trabalho levantado que adotou redes de filas com critérios de prioridade.

Sobre a **modelagem com uso de redes de filas**, apenas os trabalhos de [El Kafhali and Salah 2018] e [Adhikari et al. 2019] adotaram modelos de filas. As redes de filas possuem certas limitações como por exemplo a representação de certas situações

com compartilhamento de recursos porém são uma alternativa muito poderosa e prática para sistemas de pequeno e médio porte [Sztrik et al. 2012]. Por fim, o último critério refere-se à adoção de **critério de prioridades** entre as requisições que chegam ao sistema. Apenas este trabalho apresenta uma arquitetura IoHT com modelagem de filas e com critérios de prioridade. Tal funcionalidade é configurada por meio do comportamento das classes que ditam como as requisições chegarão ao sistema. Neste caso o que fazemos é setar pesos para cada classe associada a um ponto de entrada do sistema.

**Tabela 1. Trabalhos relacionados com múltiplos critérios de comparação.**

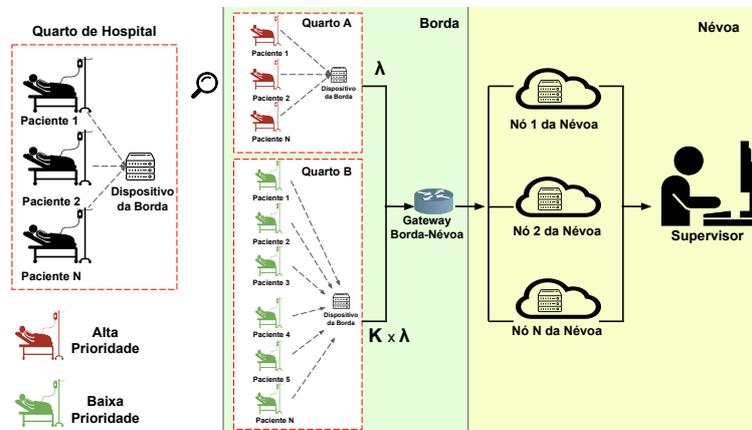
Trabalho	Métricas	Implementação de IoHT	Modelagem de Filas	Critério de Prioridade
[Baker et al. 2017]	Banda de operação, topologia, alcance, taxa de dados, recursos de segurança, adequação para saúde, direções de comunicação, capacidade de rede	Sim	Não	Não
[El Kafhali and Salah 2018]	Tempo médio de resposta, Variação de Capacidade	Sim	Sim	Não
[He et al. 2018]	Cálculo da força da senha	Sim	Não	Não
[Maktoubian and Ansari 2019]	Armazenamento de dados	Sim	Não	Não
[Adhikari et al. 2019]	Tempo médio de espera na fila, tempo médio de descarregamento, número de tarefas que atendem ao prazo de atraso, desempenho da taxa de transferência	Não	Sim	Sim
[Islam et al. 2020]	Dados de frequência cardíaca, dados de temperatura corporal, dados de umidade ambiente, taxa de erro	Sim	Não	Não
[Mukherjee et al. 2020]	Atraso total, consumo de energia, tempo médio de execução	Sim	Não	Não
<b>Este trabalho</b>	Percentual de utilização, tempo médio de resposta, vazão do sistema, taxa de descarte	Sim	Sim	Sim

#### 4. Arquitetura e Modelagem

Esta seção apresenta a arquitetura e um modelo de filas para processamento de dados de sensores em hospitais para o monitoramento de pacientes com diferentes graus de prioridade. A Figura 2 apresenta uma visão geral da arquitetura. A arquitetura é dividida em duas camadas: (i) uma camada de **computação em borda** contendo dispositivos responsáveis pelo monitoramento em tempo real de pacientes; (ii) uma camada de **computação em névoa** com equipamentos aptos a processar os dados advindos da borda.

Este trabalho considera que pacientes são separados em quartos de acordo com o grau de criticidade médica, algo que ocorre de fato em qualquer centro de internação hospitalar. Dados são gerados por sensores acoplados aos pacientes e são enviados a um *dispositivo de borda* localizado no mesmo local físico. A comunicação dos sensores com *dispositivo de borda* presume-se que seja por conexão sem fio. Os pacientes são divididos em duas categorias, podendo ser de alta prioridade (vermelho) ou baixa prioridade (verde), influenciando diretamente no tempo de atendimento que cada paciente recebe no sistema. Vale ressaltar que o projetista pode considerar um número arbitrário  $K$  de quartos e respectivas prioridades. Como simplificação, usamos apenas dois. A arquitetura do sistema foi projetada considerando que o número de pacientes com prioridade baixa seja maior que o número de pacientes com prioridade alta. Apesar da preferência, o número de requisições geradas pelos pacientes de baixa prioridade tende a ser maior, se comparado aos pacientes de prioridade mais alta. Por isso, a ilustração hipotética de um multiplicador  $K$  para a taxa de geração ( $k \times \lambda$ ) no quarto B, em relação ao quarto A.

Os dados são retransmitidos dos dispositivos de borda para um *gateway borda-névoa*. Considera-se que o *gateway* está localizado fisicamente dentro do hospital conectado aos dispositivos de borda por uma conexão cabeada. O *gateway* recebe os dados e retransmite-os para os nós da névoa aplicando um algoritmo de balanceamento de carga específico. Ao chegar nos nós da névoa os dados são processados. Tal processamento pode ser por exemplo a aplicação de algoritmos de *deep learning* para reconhecimento de padrões de anomalias dos sinais vitais dos pacientes [Rubin et al. 2017] [Sarmah 2020]. O projetista pode definir que os nós representam máquinas físicas que possuem vários núcleos de processamento, ou pode assumir que ao invés de núcleos, se tem vários contêineres que garantem um comportamento paralelo da operação. Por fim, os dados se tornam informações acessíveis por profissionais de saúde denominados de *supervisores*, que monitoram todos os pacientes através de uma plataforma *dashboard*.



**Figura 2. Visão geral da arquitetura loHT multicamadas modelada neste trabalho.**

A Figura 3 ilustra um modelo de rede de filas para a arquitetura proposta. Os componentes do modelo seguem os mesmos padrões de nomenclatura usados na descrição da arquitetura. O modelo possui dois pontos de entrada, correspondentes aos quartos do hospital, podendo ter  $N$  pacientes em cada quarto. Cada entrada gera dados constantes referentes aos pacientes, as entradas possuem taxas de chegada distintas. Cada quarto possui uma taxa de geração de dados  $\lambda$ . Pela notação de *Kendall*, a rede segue o padrão  $D/M/c/K/FCFS$ . A taxa de geração segue um padrão determinístico ( $D$ ), pois os sensores são calibrados para um intervalo de geração fixo. Como explicado na seção 2 de Fundamentos, uma fila em conjunto com um ou mais servidores é chamada de estação de atendimento. Os tempos de serviços dos servidores seguem uma distribuição exponencial. As estações de atendimento possuem um número  $c$  de servidores. As respectivas filas de espera possuem um tamanho  $k$  fixo e uma política *First-Come-First-Served*. As estações da névoa possuem uma política não preemptiva com prioridade [Bertoli et al. 2009], ou seja, as requisições são atendidas de acordo com seu tempo de chegada, mas as requisições com maior prioridade passa à frente das requisições com prioridade mais baixa (um número baixo para o peso da prioridade significa baixa prioridade). A requisição com prioridade mais baixa passa ao estado de espera até voltar a sua vez. Em caso de fila sobrecarregada e sem servidores disponíveis a requisição é descartada. No caso específico das estações da borda não há esta questão de tratamento de prioridades. O elemento *Delay Borda-Névoa* trata-se de um tempo de transmissão dos dados dos dispositivos de borda até o *gateway*.

Não foram considerados os tempos de *Delay Sensor-Borda* devido os tempos de transição serem ínfimos e irrelevantes para a análise proposta.

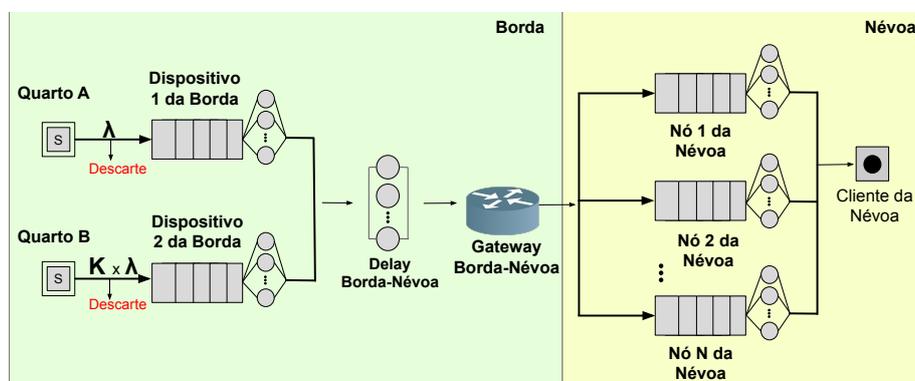


Figura 3. Rede de filas que representa uma arquitetura IoT.

## 5. Resultados

Essa seção apresenta um estudo de caso quanto à utilização do modelo de filas proposto. Foram configurados parâmetros específicos no modelo e realizada uma análise numérica sobre o mesmo. A ferramenta Java Modeling Tools (JMT) <sup>1</sup> foi utilizada para modelar e avaliar o modelo proposto. O JMT oferece um framework abrangente para avaliação de desempenho, modelagem de sistemas com técnicas analíticas e de simulação, planejamento de capacidade e estudos de caracterização de carga de trabalho [Bertoli et al. 2009]. A Tabela 2 mostra os parâmetros de configuração usados para cada elemento do modelo. A coluna de tempo representa o tempo de trabalho (tempo de serviço, geração ou transmissão) para cada componente. Vale destacar que o intervalo entre gerações do quarto A e o quarto B são distintas. O quarto A por ter mais pacientes, possui um intervalo menor do que o quarto B (com pacientes mais críticos). Em relação ao número de servidores (considerados aqui núcleos de CPU), foi configurado dois núcleos para dispositivos da borda e quatro para névoa. O caractere X indica que o componente não tem o parâmetro em questão. A mesma ideia aplica-se à capacidade das filas. O tamanho da fila de espera na névoa é três vezes maior do que a capacidade da borda por se tratarem de máquinas mais sofisticadas com maior poder de processamento.

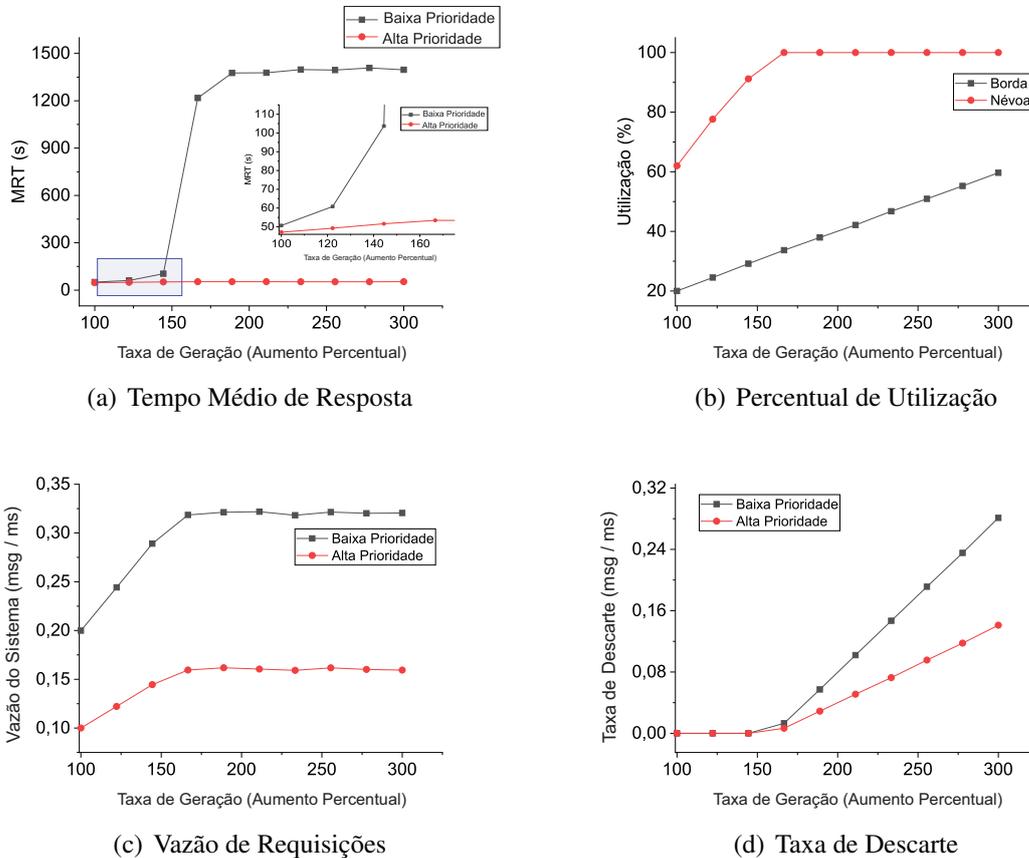
Tabela 2. Parâmetros de configuração para cada elemento do modelo.

Tipo de Componente	Componente	Tempo (ms)	Número de Servidores	Capacidade da Fila
Estação de Atendimento	Dispositivo da Borda	5 (tempo de serviço)	2	50
	Nó da Névoa	25 (tempo de serviço)	4	150
Fonte	Quarto A / Quarto B	5 / 10 (tempo entre gerações)	X	X
Delay	Delay Borda-Névoa	15 (tempo de transmissão)	X	X

A Figura 4 apresenta os resultados da análise do modelo. Quatro métricas foram calculadas: MRT, percentual de utilização (U) da borda e névoa, vazão de requisições e taxa de descarte. Um fator considerado refere-se à prioridade das requisições. Uma linha representa alta prioridade (quarto B) e baixa prioridade (quarto A). O segundo fator

<sup>1</sup> <http://jmt.sourceforge.net/>

é observado nos eixos x, correspondendo a um aumento percentual da taxa de geração (TG) das requisições no sistema. A taxa de geração do quarto A é de 1/5 e do quarto B é de 1/10. Ao longo dos eixos x de cada gráfico, estas taxas são aumentadas de 100 a 300%.



**Figura 4. Métricas de desempenho a partir do modelo de filas para IoHT.**

A Figura 4(a) apresenta o MRT para dois cenários: requisições com baixa e alta prioridade na fila de atendimento. O gráfico mostra que as requisições de baixa prioridade possuem um MRT muito superior às requisições com alta prioridade. Quando  $TG > \approx 150\%$ , o MRT alcança um pico de  $\approx 1500\text{ms}$  e permanece neste valor com taxas de gerações ainda maiores. Em contrapartida, as requisições de alta prioridade possuem valores de MRT inferiores e estáveis em todos os pontos com médias de  $\approx 50\text{ms}$ . O aumento significativo nos MRT de baixa prioridade ocorre devido tanto ao aumento das requisições de alta prioridade quanto à saturação dos recursos da névoa (vide Figura 4(b)).

A Figura 4(b) apresenta o percentual de utilização para as duas camadas (borda e névoa). Como mencionado anteriormente, dispositivos da névoa atingem 100% de utilização quando  $TG > 150\%$ . A utilização da névoa já inicia de certa forma alta com  $U = \approx 60\%$ , inicialmente até atingir 100%. Por outro lado, o percentual de utilização dos dispositivos da borda se manteve relativamente baixo, apresentando um uso máximo de  $\approx 60\%$ . Este baixo nível de utilização da borda pode ser explicado pelo baixo tempo de serviço em relação aos nós da névoa, pois a borda trata apenas de retransmitir os dados.

A Figura 4(c) apresenta os resultados para a vazão de requisições considerando alta e baixa prioridades. Neste trabalho, a vazão calcula a quantidade de mensagens

por milissegundo que o sistema pode processar, ou seja, relaciona-se ao número de requisições que saem do sistema. O gráfico mostra que as requisições com baixa prioridade possuem uma maior vazão, atingindo  $\approx 0,30$ msg/ms quando a taxa de chegada alcança  $\approx 150\%$ . As requisições com maior prioridade possuem uma vazão de dados bem menor, atingindo um pico máximo de  $\approx 0,15$ msg/ms quando a taxa de geração alcança  $\approx 150\%$ . A princípio, espera-se que as requisições com alta prioridade tenham uma vazão maior, porém é importante notar que existem mais pacientes com menor prioridade gerando dados no sistema. Portanto, observa-se que a vazão sofre um impacto muito maior da taxa de geração de requisições do que as prioridades de cada grupo de pacientes.

A Figura 4(d) apresenta a taxa de descarte.. As requisições de baixa prioridade possuem uma elevada taxa de descartes. Provavelmente estes descartes da baixa prioridade ocorre não por conta da prioridade em si mas pela taxa maior de geração de dados. Com  $TG > 150\%$ , o descarte de requisições se torna crescente para ambas prioridades até atingir um pico máximo de  $\approx 0,30$ msg/ms e  $\approx 0,12$ msg/ms. Vale ressaltar que mesmo havendo descarte em ambos os casos, a diferença entre tais resultados não é tão alta.

## 6. Conclusão

Este trabalho propôs um modelo de filas para uma arquitetura IoHT. As requisições do sistema foram classificadas em dois cenários: alta e baixa prioridade. O modelo permite configurar vários parâmetros relacionados a taxa de geração de dados, tempo de serviço, tamanho de filas e número de servidores. Neste trabalho foram calculadas as métricas de tempo médio de resposta, utilização de recursos, vazão e taxa de descarte. No entanto o modelo pode também calcular tempo médio de espera em fila e número de mensagens no sistema, porém tais métricas não foram observadas neste estudo. O modelo representa elementos relacionados, por exemplo, ao tempo de transmissão e componente de balanceamento de carga. Foi apresentado um estudo de caso que verificou o impacto de pesos atribuídos a prioridades de requisições juntamente com a variação da taxa de geração de dados. Os resultados indicam que no cenário estudado a prioridade de requisições teve grande influência sobre a métrica de tempo médio de resposta, porém não se sobressaiu em relação à grandeza da taxa de geração de dados nas demais métricas. Como trabalho futuro pretende-se avaliar as demais métricas não abordadas no presente estudo, como também estender o modelo incluindo camadas adicionais de nuvem pública e privada.

## Referências

- Adhikari, M., Mukherjee, M., and Srirama, S. N. (2019). Dpto: A deadline and priority-aware task offloading in fog computing framework leveraging multilevel feedback queueing. *IEEE Internet of Things Journal*, 7(7):5773–5782.
- Baker, S. B., Xiang, W., and Atkinson, I. (2017). Internet of things for smart healthcare: Technologies, challenges, and opportunities. *IEEE Access*, 5:26521–26544.
- Bertoli, M., Casale, G., and Serazzi, G. (2009). Jmt: performance engineering tools for system modeling. *SIGMETRICS Perform. Eval. Rev.*, 36(4):10–15.
- de Moraes Barroca Filho, I. and de Aquino Junior, G. S. (2017). Iot-based healthcare applications: a review. In *International conference on computational science and its applications*, pages 47–62. Springer.

- El Kafhali, S. and Salah, K. (2018). Performance modelling and analysis of internet of things enabled healthcare monitoring systems. *IET Networks*, 8(1):48–58.
- Haragos, I.-M. and Cernazanu-Glavan, C. (2012). Modelling road traffic using service center. *AECE 2012*, 2.
- He, D., Ye, R., Chan, S., Guizani, M., and Xu, Y. (2018). Privacy in the internet of things for smart healthcare. *IEEE Communications Magazine*, 56(4):38–44.
- Iorga, M., Feldman, L., Barton, R., Martin, M. J., Goren, N. S., and Mahmoudi, C. (2018). Fog computing conceptual model.
- Islam, M. M., Rahaman, A., and Islam, M. R. (2020). Development of smart healthcare monitoring system in iot environment. *SN computer science*, 1:1–11.
- Khoa, T. V., Saputra, Y. M., Hoang, D. T., Trung, N. L., Nguyen, D., Ha, N. V., and Dutkiewicz, E. (2020). Collaborative learning model for cyberattack detection systems in iot industry 4.0. In *2020 IEEE Wireless Communications and Networking Conference (WCNC)*, pages 1–6. IEEE.
- Maktoubian, J. and Ansari, K. (2019). An iot architecture for preventive maintenance of medical devices in healthcare organizations. *Health and Technology*, 9(3):233–243.
- Mohammadian, H. D., Mohammadian, F. D., and Assante, D. (2020). Iot-education policies on national and international level regarding best practices in german smes. In *2020 IEEE Global Engineering Education Conf. (EDUCON)*, pages 1848–1857. IEEE.
- Mukherjee, A., Ghosh, S., Behere, A., Ghosh, S. K., and Buyya, R. (2020). Internet of health things (ioht) for personalized health care using integrated edge-fog-cloud network. *Journal of Ambient Intelligence and Humanized Computing*, pages 1–17.
- Rubin, J., Abreu, R., Ganguli, A., Nelaturi, S., Matei, I., and Sricharan, K. (2017). Recognizing abnormal heart sounds using deep learning. *arXiv preprint arXiv:1707.04642*.
- Sarmah, S. S. (2020). An efficient iot-based patient monitoring and heart disease prediction system using deep learning modified neural network. *IEEE Access*, 8:135784–135797.
- Shafiq, M., Tian, Z., Sun, Y., Du, X., and Guizani, M. (2020). Selection of effective machine learning algorithm and bot-iot attacks traffic identification for internet of things in smart city. *Future Generation Computer Systems*, 107:433–442.
- Sztrik, J. et al. (2012). Basic queueing theory. *University of Debrecen, Faculty of Informatics*, 193:60–67.
- Uslu, B. Ç., Okay, E., and Dursun, E. (2020). Analysis of factors affecting iot-based smart hospital design. *Journal of Cloud Computing*, 9(1):1–23.
- Yassein, M. B., Hmeidi, I., Al-Harbi, M., Mrayan, L., Mardini, W., and Khamayseh, Y. (2019). Iot-based healthcare systems: a survey. In *Proceedings of the Second International Conf. on Data Science, E-Learning and Information Systems*, pages 1–9.