

ANEXO XIX **VOLUMETRIA ESTIMADA DE DADOS PARA CONVERSÃO**

Considerando a natureza da solução proposta, baseada em computação em nuvem, a precificação e dimensionamento dos serviços estão diretamente relacionados a métricas de consumo, tais como volume de dados, quantidade de usuários, capacidade de armazenamento e tráfego estimado.

Nesse contexto, a infraestrutura tecnológica atualmente em operação possui caráter meramente subsidiário, não sendo determinante para a formulação das propostas, uma vez que a solução pretendida independe da infraestrutura física existente.

Contudo, são disponibilizadas informações consolidadas do ambiente tecnológico atualmente utilizado pela Prefeitura Municipal e pela Câmara Municipal, abrangendo indicadores de volumetria de dados, crescimento histórico, quantitativo de usuários e demais métricas de infraestrutura disponíveis, com a finalidade de subsidiar o adequado dimensionamento técnico e econômico das propostas.

Ressalta-se que os sistemas atualmente utilizados são soluções desenvolvidas por fornecedores distintos, que adotam arquiteturas e modelos próprios de armazenamento e organização das informações. Assim, a distribuição dos dados entre bancos, tabelas e demais estruturas físicas não guarda, necessariamente, correspondência direta com os módulos funcionais das aplicações, razão pela qual a volumetria foi apurada de forma consolidada por ambiente de banco de dados.

VOLUME DE DADOS

Servidor 01:

Quantidade de Bancos de Dados e Volumetria Atual do Ambiente:

Quantidade de bancos usuário	Total de Dados Prefeitura(GB)	Total de Dados Câmara(GB)	Total Alocado (GB)
11	591,56		591,56

Servidor 02:

Quantidade de Bancos de Dados e Volumetria Atual do Ambiente:

Quantidade de bancos de usuário	Total de Dados Prefeitura(GB)	Total de Dados Câmara(GB)	Total Alocado (GB)
76	501,68	27,73	529,41

CONSIDERAÇÕES PARA O DIMENSIONAMENTO

Servidor 01:

O ambiente possui atualmente 11 bancos de dados, totalizando aproximadamente 591,56 GB de espaço alocado. O histórico disponível no banco de dados de sistema (msdb) contempla apenas os meses de abril a junho de 2026, não sendo suficiente para a determinação de uma taxa histórica de crescimento da base de dados.

Dessa forma, não é possível estabelecer uma projeção quantitativa do crescimento do ambiente com base em evidências históricas. Para fins de dimensionamento da solução em nuvem, recomenda-se que as licitantes considerem margem técnica compatível com a expansão natural das bases de dados durante a vigência contratual.

Servidor 02:

O ambiente possui atualmente 76 bancos de dados, totalizando aproximadamente 529,41 GB de espaço alocado. O histórico disponível contempla 12 meses de registros de backups, permitindo observar a estabilidade operacional do ambiente durante o período analisado.

Entretanto, considerando que o volume de backups não representa diretamente o crescimento efetivo das bases de dados, não é possível inferir uma taxa de crescimento anual apenas com essas informações. Assim, recomenda-se

que o dimensionamento da solução considere margem técnica para suportar a evolução natural do volume de dados ao longo da execução contratual.

DISTRIBUIÇÃO ESTIMADA DA VOLUMETRIA POR ÁREA FUNCIONAL

Com a finalidade de fornecer subsídios adicionais ao dimensionamento da solução, foi realizada uma estimativa da distribuição da volumetria dos dados por área funcional dos sistemas atualmente utilizados.

A estimativa foi elaborada a partir da análise técnica da estrutura dos bancos de dados atualmente existentes, considerando a organização lógica das informações, as funcionalidades suportadas e o conhecimento operacional do ambiente. Em razão da arquitetura adotada pelos sistemas atuais, determinadas estruturas de banco de dados são compartilhadas entre diferentes módulos de negócio, não sendo possível estabelecer uma segregação exata da volumetria por funcionalidade.

Dessa forma, os valores apresentados possuem caráter estimativo e destinam-se exclusivamente ao apoio ao dimensionamento da solução a ser contratada, não representando uma distribuição física ou lógica precisa dos dados.

Servidor 01:

No Servidor 01:

Verificou-se que o ambiente é dedicado às soluções da área tributária, contemplando os bancos de dados utilizados pelas funcionalidades de cadastro, arrecadação, dívida ativa, integrações, cálculos, emissão de relatórios e demais rotinas correlatas. Dessa forma, toda a volumetria apresentada refere-se ao ecossistema do Sistema Tributário.

Os valores apresentados possuem caráter estimativo e destinam-se exclusivamente ao apoio ao dimensionamento da solução, não representando, necessariamente, uma segregação física dos dados, uma vez que as estruturas de armazenamento podem ser compartilhadas entre diferentes funcionalidades do sistema.

Área Funcional	Prefeitura (GB)	Câmara (GB)
Tributário e Componentes Relacionados	591,56	-
Total	591,56	

Servidor 02:

Área Funcional	Prefeitura (GB)	Câmara (GB)
Materiais (Compras, Patrimônio e Almoxarifado)	169,74	2,94
Formalização de Demanda	99,16	-
Adiantamento de Despesas	19,00	-
Controle de Custos	25,13	-
Controle de Frotas	0,90	-
Pregão Eletrônico	5,25	-
Controle de Registro de Ponto	14,66	-
Contabilidade, Tesouraria, Recursos Humanos, Patrimônio Imobiliário e Protocolo	82,50	13,56
Demais Funcionalidades e Integrações	85,34	11,23
Total	501,68	27,73

Observação:

A distribuição apresentada possui caráter estimativo e foi elaborada exclusivamente para subsidiar o dimensionamento da solução. Os valores não correspondem, necessariamente, à estrutura física dos bancos de dados atualmente utilizados, uma vez que determinadas bases armazenam informações compartilhadas entre múltiplas áreas funcionais.

A volumetria total incluindo o (Servidor 1 + Servidor 2 + Dados da Câmara Municipal), equivale a 1.120,97 GB, considerando a volumetria total de Prefeitura e Câmara.

A volumetria externa total incluindo dados federais e estaduais é estimada por “IA – Gemini”, da ordem máxima de:

 Resumo da comparação (Porte de Olímpia):

CadÚnico: ~300 MB (Dados socioeconômicos puramente textuais e estáticos)

SED (Educação): ~2 a 5 GB (Histórico denso de notas, chamadas diárias e dados escolares)

e-SUS (Saúde): ~10 a 30 GB (Prontuários médicos altamente detalhados e receitas acumuladas)

JUSTIFICATIVA

ESTIMATIVAS DOS BANCOS DE DADOS DO GOVERNO FEDERAL

É possível ter uma estimativa do tamanho do Banco de Dados do SUS de um município do porte de Olímpia – SP, em GB?

Sim, é possível fazer uma estimativa técnica aproximada com base na estrutura de dados de prontuários eletrônicos e no volume populacional.

Para um município do porte de Olímpia - SP (cerca de 57 mil habitantes), estima-se que o banco de dados local centralizado (onde roda o e-SUS APS / PEC - Prontuário Eletrônico do Cidadão) tenha hoje entre 10 GB e 30 GB de tamanho em disco para dados puramente estruturados (textos, números e logs).

O cálculo estimado considera os seguintes fatores:

1. Os dados estruturados (O "grosso" do prontuário)

Bancos de dados relacionais de saúde (geralmente mantidos em PostgreSQL no e-SUS) armazenam texto de forma muito compacta.

Um cadastro de cidadão completo com histórico de consultas, receitas passadas, vacinas aplicadas e evolução médica consome, em média, entre 200 KB e 500 KB por paciente ao longo de alguns anos.

Multiplicando isso por cerca de 60.000 cadastros ativos e inativos em Olímpia, chegamos a algo em torno de 12 GB a 30 GB apenas de tabelas de texto e histórico clínico acumulado.

Nota de infraestrutura: O próprio Ministério da Saúde pontua nos manuais do e-SUS que o sistema pode rodar em servidores locais compactos e, historicamente, uma de suas versões suportava o banco Oracle XE, que possuía um limite estrito de 10 GB para armazenamento de tabelas em municípios de pequeno a médio porte.

2. O que pode fazer esse número explodir? (Dados Não Estruturados)

Se sairmos do banco de dados puramente clínico (texto/fichas) e olharmos para a infraestrutura digital completa da saúde do município, o armazenamento em Gigabytes muda drasticamente caso inclua:

Imagens Médicas (PACS/DICOM): Se o município armazenar localmente os arquivos digitais de exames de imagem (como Raio-X digital, Tomografias e Ultrassons realizados na rede municipal), o volume salta facilmente para a casa dos Terabytes (TB), pois uma única tomografia pode pesar mais de 500 MB.

Backups e Logs: Os backups diários e históricos de auditoria guardados pela equipe de TI da prefeitura multiplicam o tamanho do banco original muitas vezes.

Portanto, em termos de banco de dados de sistemas (fichas, consultas e cadastros), o tamanho de Olímpia é considerado leve para os padrões atuais de TI, girando na casa dos poucos gigabytes (10 a 30 GB), permitindo que seja gerenciado por servidores municipais comuns.

É possível ter uma estimativa do tamanho do Banco de Dados do CADÚNICO de um município do porte de Olímpia – SP, em GB?

Se para o SUS nós falávamos em Gigabytes, para o Cadastro Único (CadÚnico) de um município do porte de Olímpia - SP (cerca de 57 mil habitantes), a escala muda radicalmente: o tamanho desse banco de dados estruturado é medido em Megabytes (MB), e não em Gigabytes.

Uma estimativa realista para o tamanho do banco de dados do CadÚnico de Olímpia hoje gira em torno de 250 MB a 500 MB (menos de meio Gigabyte).

Por que o CadÚnico é tão "leve"?

O CadÚnico é um sistema puramente de caracterização socioeconômica. Ao contrário do SUS, ele não guarda históricos cumulativos pesados (como receitas médicas, prontuários de anos de evolução, vacinas ou imagens).

O sistema armazena essencialmente registros do tipo "texto" e "números":

Dados civis (Nome, CPF, RG, NIS).

Estrutura familiar (quem mora com quem).

Características do domicílio (se tem água encanada, número de cômodos, material das paredes).

Renda e escolaridade.

O Cálculo da Estimativa

População Alvo em Olímpia: O foco do CadÚnico são famílias com renda de até meio salário mínimo per capita. Em um município como Olímpia, estima-se que entre 25% e 35% da população total esteja inscrita no CadÚnico (entre beneficiários do Bolsa Família, BPC, Tarifa Social de Energia, etc.). Isso representa aproximadamente 15.000 a 20.000 pessoas cadastradas (divididas em cerca de 6.000 a 8.000 famílias).

Peso por Registro: Em bancos de dados relacionais bem estruturados (como os utilizados pela Caixa Econômica Federal e pelo Ministério do Desenvolvimento Social), uma ficha cadastral completa de uma pessoa com seu histórico básico consome cerca de 10 KB a 15 KB.

O Total: 20.000 pessoas × 15 KB por registro resulta em aproximadamente 300.000 KB, o que equivale a pouco menos de 300 Megabytes. Mesmo adicionando índices do banco, tabelas auxiliares do sistema e registros históricos de atualizações anteriores, o arquivo final mal encosta na marca de 0,5 GB.

💡 Curiosidade tecnológica: O banco de dados inteiro do CadÚnico de Olímpia caberia facilmente em um pen drive antigo de 512 MB ou em um único CD-ROM. Toda essa base municipal é periodicamente transmitida via internet para a base nacional da Caixa Econômica Federal de forma rápida exatamente por causa desse tamanho reduzido.

ESTIMATIVAS DOS BANCOS DE DADOS DO GOVERNO ESTADUAL (SED)

É possível ter uma estimativa do tamanho do Banco de Dados da SED - Secretaria Estadual de Educação do Estado de São Paulo, de um município do porte de Olímpia – SP, em GB?

Para estimar o tamanho da base de dados da SED (Secretaria Escolar Digital) do Estado de São Paulo correspondente ao município de Olímpia, precisamos entender como as informações dos alunos, professores e escolas estaduais são estruturadas.

Diferente do CadÚnico (que é levíssimo) e do SUS (que armazena históricos clínicos volumosos), o sistema de educação lida com notas, faltas, boletins, diários de classe digitais e materiais de apoio.

Uma estimativa realista para os dados estritos do município de Olímpia dentro da plataforma SED gira em torno de 2 GB a 5 GB de dados puramente estruturados (texto/tabelas).

O Cenário de Olímpia na Rede Estadual

Volume de Alunos: Olímpia possui cerca de 57 mil habitantes. O município conta com aproximadamente 6 a 7 escolas estaduais (como as E.E. Capitão Narciso Bertolino, E.E. Dr. Wilquem Manoel Neves, E.E. Dona Anita Costa, entre outras). Estima-se que a rede estadual na cidade atenda em torno de 4.000 a 5.500 estudantes (divididos entre Ensino Fundamental II, Ensino Médio e EJA).

Dados por Aluno: O prontuário digital de um aluno na SED inclui:

Cadastro pessoal e matrícula (RA, documentos, endereço).

Histórico escolar completo (notas de todos os bimestres, frequência diária detalhada por matéria).

Matrículas em itinerários formativos do Ensino Médio, atribuições de turmas e dados de responsáveis.

Por que o volume fica entre 2 GB e 5 GB?

Diferente de uma ficha estática, a SED registra a rotina escolar diária. Um banco relacional guarda milhões de linhas de "lançamentos de presença/ausência" e "notas" ao longo dos anos para esses milhares de alunos. Multiplicando o histórico acumulado de alunos ativos e dos que já se formaram na rede pública da cidade, o volume de texto e índices do banco chega facilmente à casa dos poucos gigabytes.

O Fator que muda tudo: E se incluirmos arquivos anexos?

Essa estimativa de até 5 GB considera apenas o banco de dados principal (as tabelas do SQL/PostgreSQL). O tamanho pode explodir para dezenas ou centenas de Gigabytes se forem contabilizados:

Upload de Documentos: Cópias escaneadas de RG, CPF, Certidões de Nascimento e Históricos Escolares antigos em PDF anexados no prontuário de cada aluno.

Plataformas de Aprendizagem: Se integrados os arquivos enviados pelos alunos em tarefas digitais (redações, imagens e trabalhos do CMSP - Centro de Mídias de São Paulo), o volume migra do banco de dados relacional para servidores de arquivos (storage), atingindo facilmente a escala de Terabytes (TB) no nível estadual, proporcionalmente pesando muito mais para o município.

USUÁRIOS

O quantitativo de usuários foi apurado com base em duas fontes: os cadastros dos sistemas corporativos e as informações de identidade do Active Directory (AD).

Nos sistemas corporativos foram identificados 878 usuários cadastrados, dos quais 516 encontram-se ativos. No Active Directory foram identificadas 1.894 contas de usuários, sendo 956 habilitadas. Destas, 940 registraram autenticação nas semanas recentes.

Os dados apresentados representam perspectivas distintas do ambiente, uma vez que os sistemas corporativos e o Active Directory possuem finalidades diferentes. Dessa forma, as informações foram utilizadas de forma complementar para o dimensionamento da solução, considerando tanto os usuários ativos nos sistemas quanto o universo de usuários habilitados na infraestrutura corporativa.

5 SOBRE A ENGENHARIA REVERSA PARA A MIGRAÇÃO DE DADOS

O Tribunal de Contas do Estado de São Paulo (TCESP) reconheceu a complexidade da migração de software na Administração Pública, especialmente quando a empresa fornecedora do sistema anterior não disponibiliza o "dicionário de dados" (a estrutura do banco de dados) para a nova contratada.

Sobre esse tema, a jurisprudência recente do TCESP estabeleceu as seguintes diretrizes:

Legitimidade da Engenharia Reversa: O TCESP validou a engenharia reversa como um procedimento técnico legítimo e necessário para garantir a continuidade do serviço público e a salvaguarda dos dados, permitindo que uma nova contratada mapeie e extraia as informações do banco de dados legado.

Exigência de Prazos Adequados: O Tribunal determinou que o prazo concedido para a migração de dados deve ser proporcional ao nível de informação fornecido

no edital. Como a Administração não possui a arquitetura do banco de dados atual para fornecer aos licitantes, exigir uma migração em um prazo curto (como 30 dias) é considerado irregular e insuficiente.

Alternativas para a Administração: Em decisões paradigmáticas recentes (como o Processo TC-011691.989.25-5 de Embu-Guaçu), o TCESP determinou que a Administração deve escolher entre duas opções: ou disponibiliza os parâmetros necessários (composição da estrutura do banco de dados e o respectivo dicionário) para manter um prazo curto de migração, ou amplia o prazo de execução para permitir que a empresa vencedora realize os procedimentos de engenharia reversa de forma segura.

Caso concreto do Município de Olímpia: Foi justamente para observar rigorosamente essas orientações do TCESP que o Município de Olímpia dilatou o prazo da etapa de conversão e migração de dados neste certame para aquisição de um novo sistema. Ao invés dos costumeiros 30 dias praticados no mercado com bases conhecidas, o município fixou um prazo de 120 dias para a migração. O Edital de Olímpia deixa claro que, como o município não dispõe do esquema dos dados existentes, esse interregno ampliado foi dimensionado especificamente para permitir que a adoção da engenharia reversa seja feita de forma viável e em conformidade com as exigências do Tribunal.