Laboratório de Pesquisa em Redes e Multimídia

O Protocolo ICMP

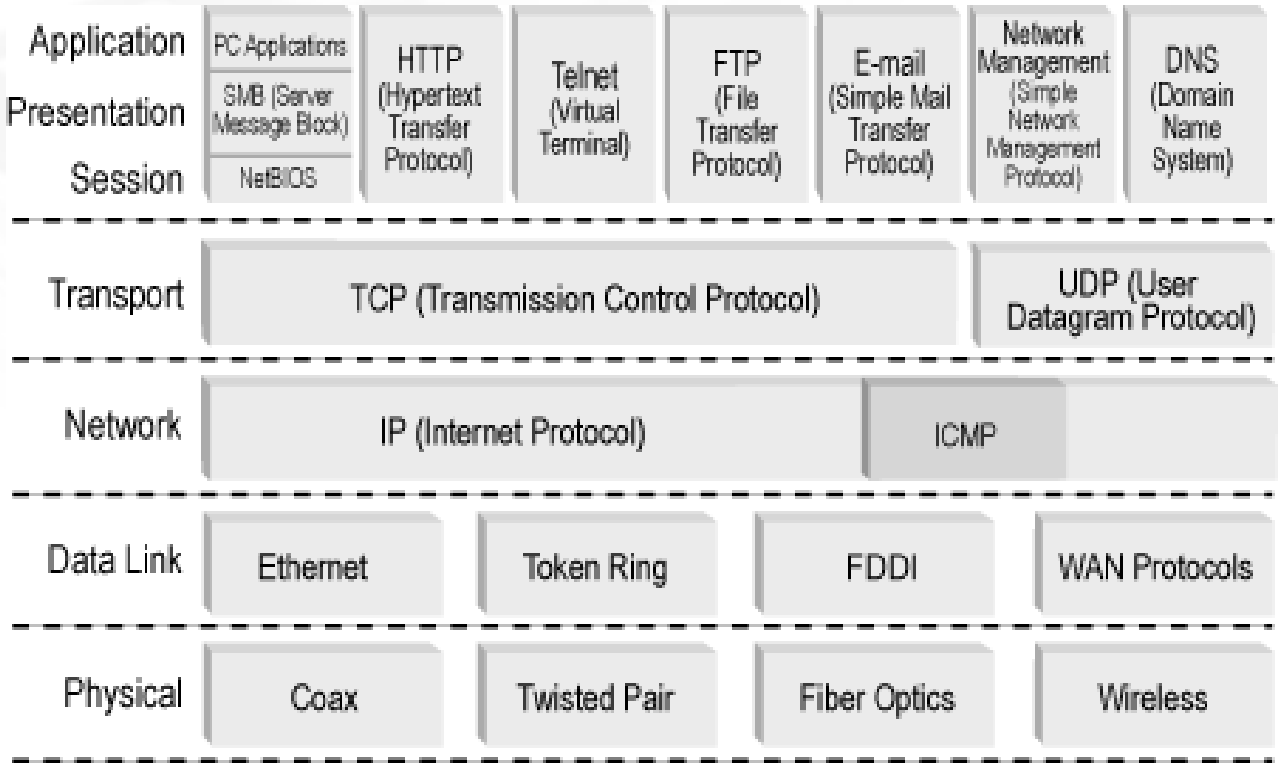
Universidade Federal do Espírito Santo
Departamento de Informática

O ICMP

- ICMP = "Internet Control Message Protocol"
- Através do ICMP, um roteador ou *host* destino pode reportar à estação origem uma condição de erro no processamento de um datagrama.
- O ICMP apenas *informa* erros ao nível IP de origem, não tendo qualquer responsabilidade sobre a correção dos mesmos.

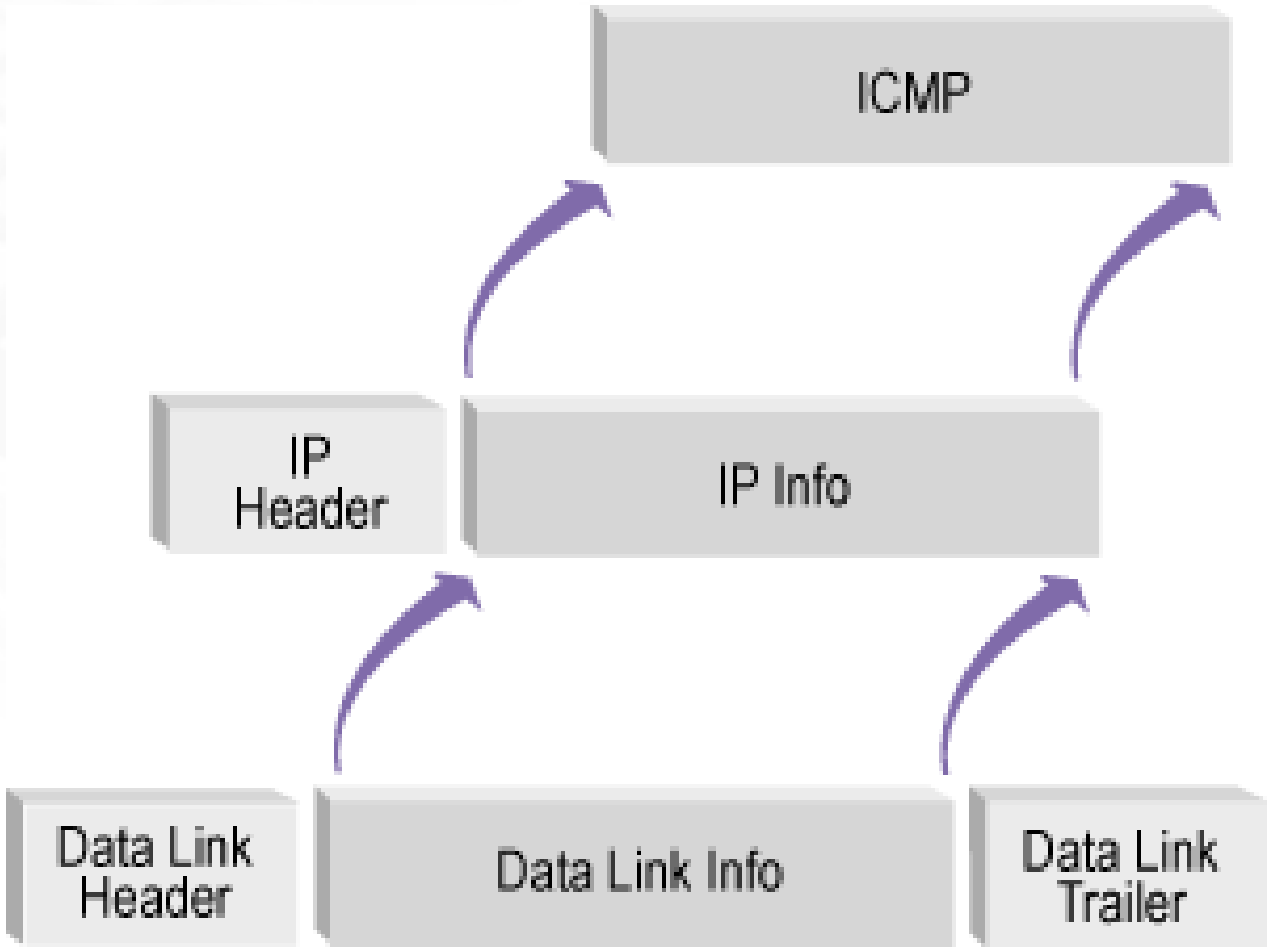
Nível do ICMP

- É um protocolo da camada de inter-redes da arquitetura TCP/IP.

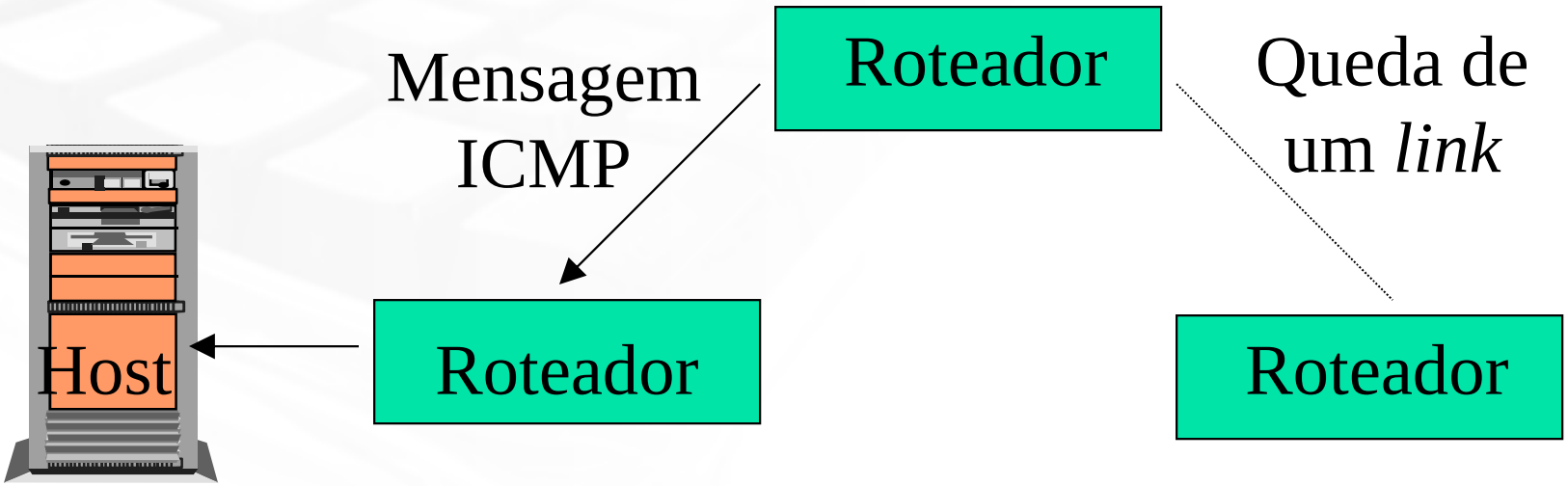


Encapsulamento do ICMP

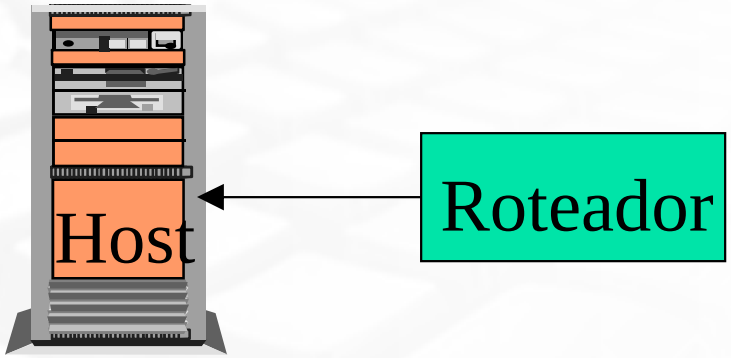
- Mensagens ICMP são encapsuladas na porção de dados do IP (*Protocol* = 1).



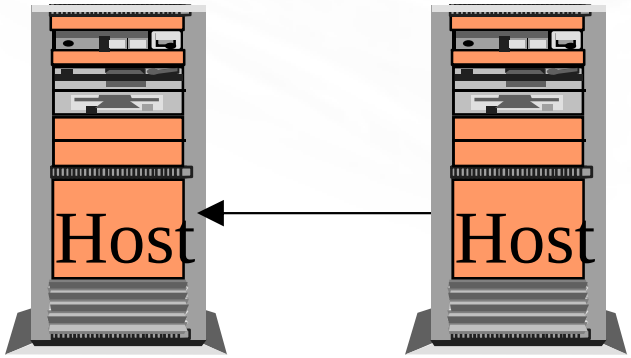
Cenários de Uso



Cenáros de Uso (cont.)



- Não consigo atingir o destino;
- TTL expirou;
- Parâmetro estranho;
- Existe rota melhor.



- Timer de remontagem expirou;
- Parâmetro estranho;
- Não consigo atingir o serviço.

Cenários de Uso (cont.)

- Quando um roteador descarta um pacote devido ao fato do TTL ter expirado.
- Quando o roteador não possui capacidade de *bufferização* para encaminhar o datagrama.
- Quando o roteador tem que fragmentar um datagrama com o bit "*don't fragment*" ligado.

Cenários de Uso (cont.)

- Quando o *host* ou o roteador descobrem um erro de sintaxe no cabeçalho do IP.
- Quando o roteador não tem uma rota para a rede destino na sua tabela de rotas.
- Quando o roteador solicita ao *host* fonte para usar uma outra rota de menor caminho.

Observações

- Mensagens ICMP são roteadas como outro datagrama qualquer, não sendo garantida a sua entrega ao seu destino final.
- Não existe mensagem ICMP para reportar erros ou descarte de pacotes ICMP.
- Erros somente são reportados num datagrama não fragmentado ou no primeiro fragmento de um datagrama.

Obrigatoriedade de Uso

- O padrão especifica que uma mensagem *pode* ("*should*") ser enviada na ocorrência de uma situação inesperada. Ele não obriga que todo erro *tenha* ("*must*") que resultar em uma mensagem ICMP.
- Essa escolha parece ser de bom senso pois a primeira prioridade de um roteador numa rede é encaminhar datagramas e não reportar erros.
- Um *host* congestionado, por sua vez, deve dar mais importância à entrega dos datagramas às suas aplicações do que às notificações de erros remotos.

Exemplo

```
> telnet 10.1.1.1
Trying 10.1.1.1
telnet: connect: Host is unreachable

> traceroute 10.1.1.1
traceroute to 10.1.1.1, 30 hops max, 40 bytes packets
1 nomad-gateway (128.121.50.50)  2 ms 2 ms 2 ms
2 liberty-gateway (130.94.40.250)  91 ms 11 ms 78 ms
3 border2-h0.NewYork.mci.net (204.70.45.9) !H !H !H
```

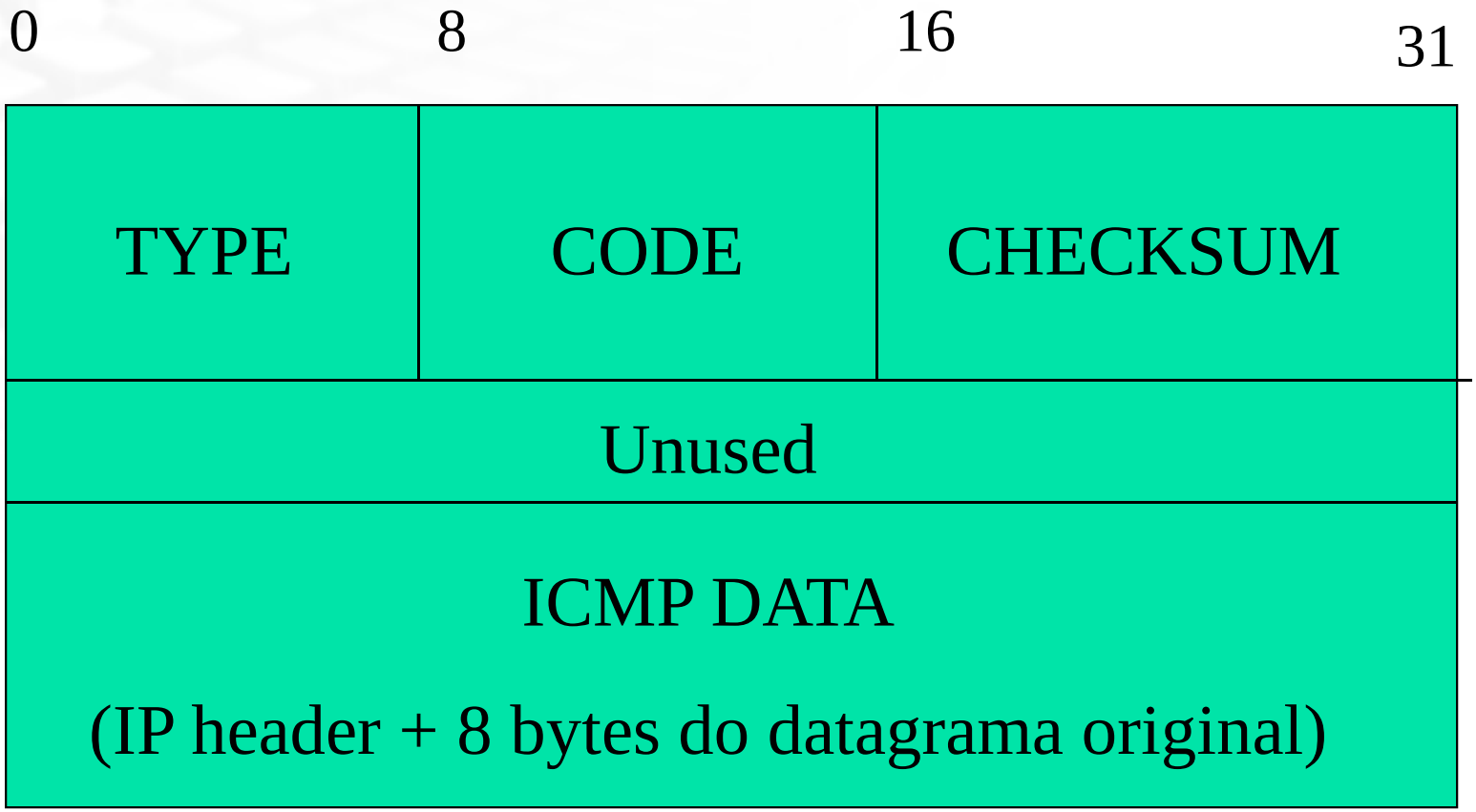
Quando não enviar mensagens ICMP

- É importante garantir que o tráfego ICMP não sobrecarregue a rede, tornando a situação ainda pior. Alguns limites óbvios devem ser estabelecidos no protocolo.
- Assim, o ICMP não deve reportar problemas causados por:
  - Roteamento ou entrega de mensagens ICMP;
  - Datagramas *broadcast* ou *multicast*;
  - Fragmentos de datagramas que não sejam os primeiros;
  - Mensagens cujo endereço fonte não identifica um único *host*. Por exemplo, 127.0.0.1 ou 0.0.0.0.

A Mensagem ICMP

- Cada mensagem ICMP tem seu próprio formato, mas todas começam com os seguintes campos:
  - TYPE (8 bits): identifica a mensagem;
  - CODE (8 bits): fornece mais informações sobre a mensagem;
  - CHECKSUM (16 bits).
- O formato do restante da mensagem é determinado pelo seu tipo.
- Uma mensagem ICMP reportando erro sempre inclui o cabeçalho e os primeiros 64 bits de dados do pacote causador do problema.

Formato da Mensagem



Formato da Mensagem (cont.)

- TYPE especifica o significado da mensagem e o formato do restante do pacote. Treze tipos foram definidos.
- CODE contém o código de erro para o datagrama, reportado pela mensagem ICMP. A interpretação desse campo é dependente do tipo da mensagem.
- CHECKSUM é aplicado à toda mensagem, iniciando a partir do campo TYPE. O algoritmo é o mesmo usado pelo IP para cálculo do *checksum* do cabeçalho IP.

Formato da Mensagem (cont.)

- ICMP DATA contém informações específicas desta mensagem ICMP.
- Tipicamente, contém parte da mensagem IP original para a qual a mensagem ICMP foi gerada (todo o cabeçalho IP + os primeiros 64 bits do campo de dados).
- O cabeçalho é incluído para que o host origem possa fazer um "*match*" da mensagem ICMP com o seu próprio *stream* de dados.
- Os primeiros 64 bits são incluídos porque eles contém o cabeçalho do TCP ou do UDP.

Campo Type

ICMP Header - Type Field	
Type Field	**Message Type**
0	Echo Reply
3	Destination Unreachable
4	Source Quench
5	Redirect
8	Echo Request
11	Time Exceeded
12	Parameter Problem
13	Timestamp Request
14	Timestamp Reply
15	Information Request
16	Information Reply
17	Address Mask Request
18	Address Mask Reply

Destination Unreachable (3)

- Usado em situações nas quais a entrega de um datagrama pode falhar:
  - Queda de um enlace, roteador desligado ou desconectado para manutenção, razões administrativas (segurança), etc.
- Se a mensagem é enviada por um roteador intermediário, isso significa que ele considerou o endereço IP destino inatingível.
- Se a mensagem é enviada pelo *host* destino, isso significa que o protocolo especificado no campo *Protocol* do datagrama original é inatingível.

Code	Descrição
0	Network unreachable
1	Host unreachable
2	Protocol unreachable (not supported at the destination)
3	Port unreachable (remote application may be unavailable)
4	Fragmentation needed but the *Don't Fragment* bit was set
5	Source route failed
6	Destination network unknown
7	Destination host unknown
8	Source host isolated (obsolete)
9	Destination network administratively prohibited
10	Destination host administratively prohibited
11	Network unreachable for this Type of Service
12	Host unreachable for this Type of Service
13	Communication administratively prohibited by filtering
14	Host precedence violation
15	Precedence cutoff in effect

Time Exceeded (11)

- Mensagem usada para reportar situações de *timeout* na entrega do datagrama.
- Um datagrama recebe *timeout* se o seu TTL atinge zero quando ele está em trânsito ou quando o *timer* de remontagem do *host* destino expira antes que todos os seus fragmentos tenham chegado.
- Code = 0: TTL zerou (enviado pelo roteador).
- Code = 1: Remontagem de fragmentos excedeu o tempo máximo (enviado pelo *host*).

Path MTU Discovery

- A fim de minimizar o *overhead* quanto maior tamanho dos datagramas melhor. Entretanto, há um limite imposto pelo MTU.
  - MTU – Maximum Transmission Unit
- Se o datagrama for muito grande ele poderá ser fragmentado, o que diminui o desempenho do sistema.

Path MTU Discovery (cont.)

- *Hosts* evitam a fragmentação usando um tamanho padrão de datagrama de 576 bytes para qualquer destino não local, o que diminui desnecessariamente o desempenho.
- *Path MTU Discovery* é o mecanismo usado para descobrir o maior tamanho de datagrama que pode ser enviado sobre um caminho.

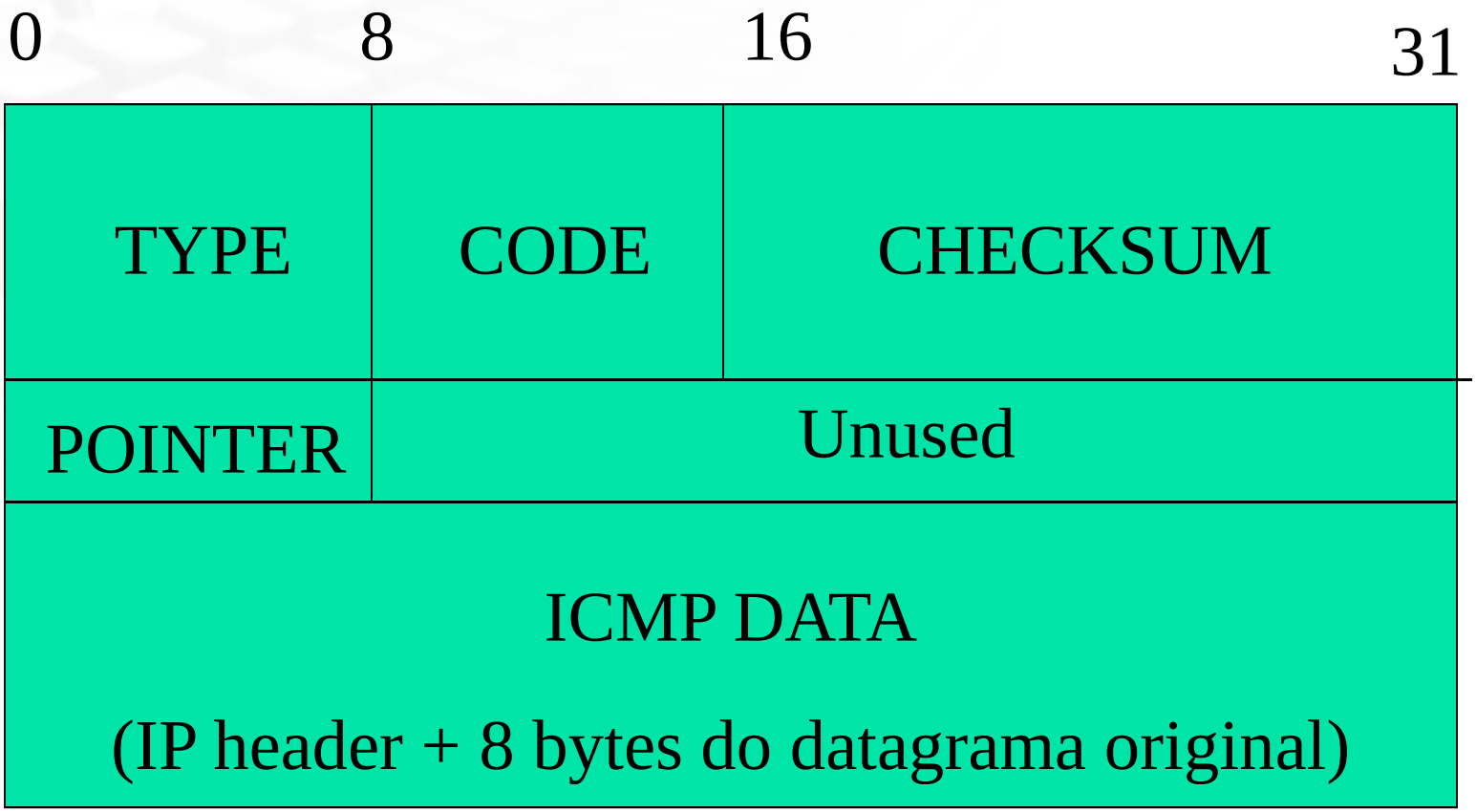
Algoritmo Path MTU Discovery

- O bit *don't fragment* é feito igual a 1.
- O *Path MTU* se inicia com o valor do MTU da interface local.
- Se o datagrama for muito grande para algum roteador, ele envia uma mensagem de *ICMP Destination Unreachable*, com Code = 4.
- O *host* origem recebe a mensagem, reduz o tamanho do datagrama e tenta novamente.

Path MTU Discovery

TYPE	CODE = 4	CHECKSUM
Unused		Next-Hop MTU
ICMP DATA (IP header + 8 bytes do datagrama original)		

Parameter Problem (12)



Parameter Problem (12) (cont.)

- Mensagem usada para reportar problemas não cobertos por nenhuma das outras mensagens de erro.
- Informação inconsistente em um campo de opções, pode tornar impossível o processamento correto do datagrama, forçando um descarte do mesmo.
- Em geral, problemas de parâmetros aparecem devido a erros de implementação no sistema que armazenou os parâmetros no cabeçalho IP.

Parameter Problem (12) (cont.)

- POINTER identifica o byte onde o erro foi detectado.
- CODE pode assumir um dos seguintes valores:

CODE	Descrição
0	O valor do campo Pointer identifica o byte onde ocorreu o erro.
1	Uma opção requerida está faltando (usado na comunidade militar para indicar a perda de uma opção de segurança)
2	Tamanho errado

Source Quench (4)

- A entrega de um pacote pelo IP nem sempre é feita de uma forma "suave", já que inúmeros problemas de congestionamento podem ocorrer:
  - uma conexão WAN de baixa velocidade entre duas LAN's pode criar um gargalo;
  - um ou mais *hosts* enviando tráfego UDP para um servidor lento pode sobrecarregar o servidor, causando o descarte de datagramas devido ao *overflow;*
  - um roteador pode ficar sem espaço de armazenamento e ser forçado a descartar alguns datagramas;

Source Quench (4) (cont.)

- Em suma, problemas de congestionamento podem causar o descarte de datagramas, resultando em retransmissões que poderão produzir maior tráfego e, conseqüentemente, um aumento do problema.
- Assim, em situações em que datagramas chegam a uma taxa muito alta a um roteador ou a um *host*, ocorre o descarte dos mesmos.

Source Quench (4) (cont.)

- A opção *source quench* implementa uma forma básica de controle de fluxo:
  - Mensagens são enviadas pelo receptor para requerer ao transmissor que reduza a sua taxa de transmissão.
- Usualmente, mensagens de erro ICMP avisam ao *host* origem o porquê de um dos seus datagramas ter sido descartado.

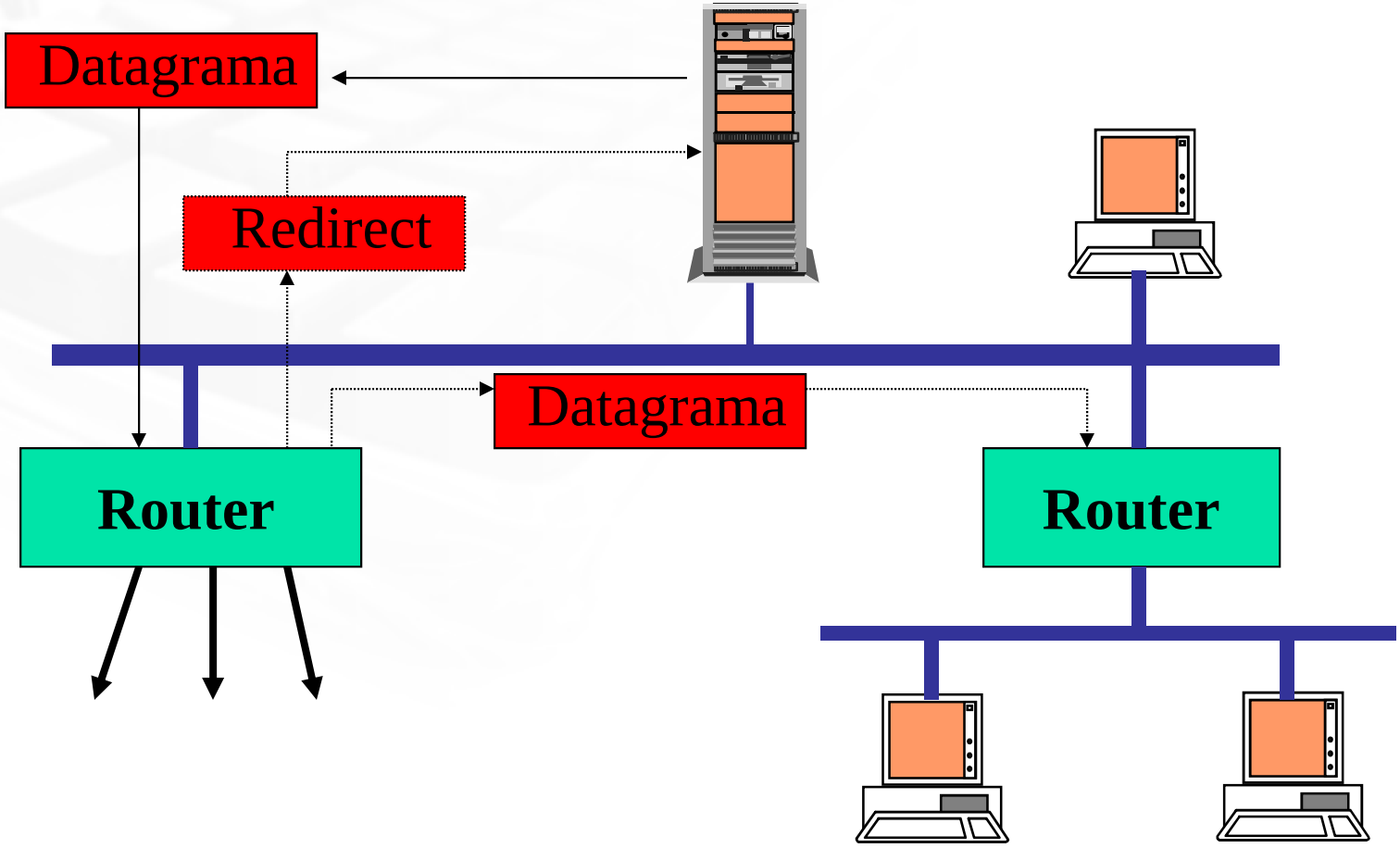
Source Quench (4) (cont.)

- Entretanto, em situações de congestionamento, é possível que os datagramas descartados não venham das máquinas que estão gerando o alto tráfego.
- Os detalhes de exatamente como o sistema congestionado deve executar um *source quench* é deixado para o implementador.
- Também é deixado em aberto a questão "*quando – e para quem – um roteador deve enviar a mensagem de source quench?*"

Redirect (5)

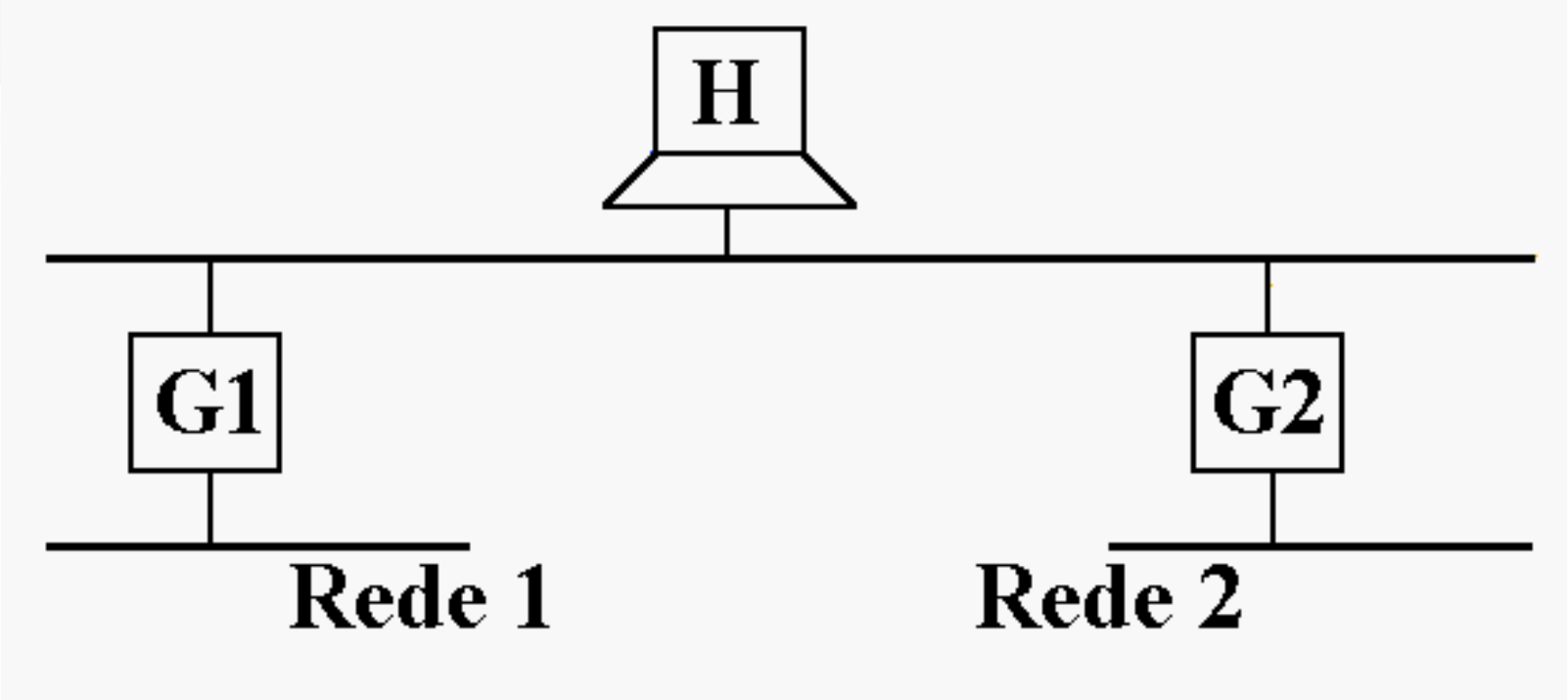
- Mensagem enviada por um roteador a um *host* da rede local quando ele detecta a existência de uma rota mais otimizada para a rede destino.
- O roteador envia um ICMP *redirect* requisitando que o *host* mude sua tabela de rotas. Ao recebê-la, o *host* muda a tabela caso esteja configurado para trabalhar com roteamento dinâmico.
- Esta mensagem está restrita a um roteador e um *host* na mesma rede. Mensagens *redirect* não são enviadas para outros roteadores (não são usadas para propagar informações de roteamento).

Redirect (5)



Redirect (5)

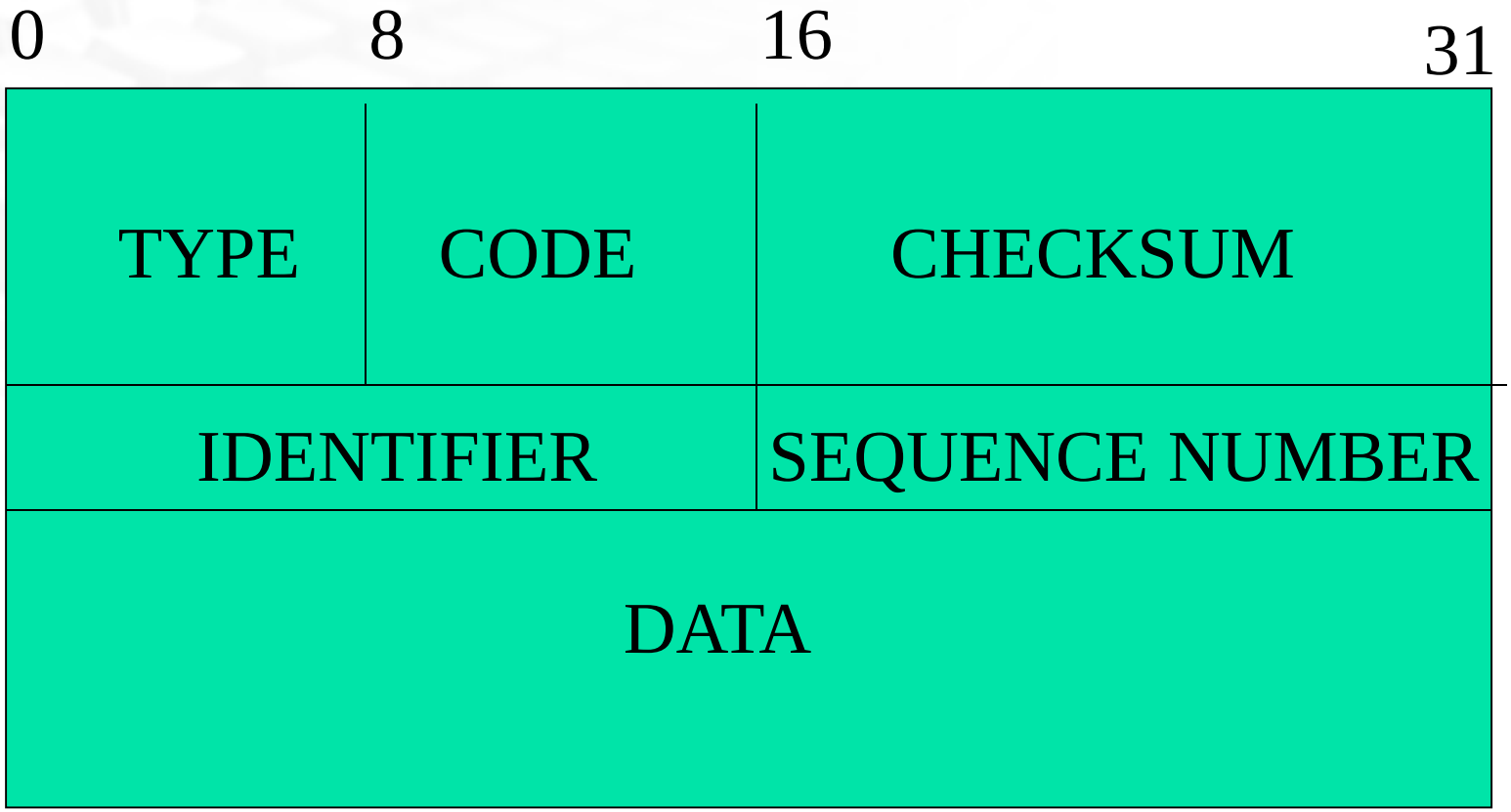
- Se o *host* H mandar um pacote para a Rede 1 via G2, G2 manda um ICMP REDIRECT para H e roteará o pacote via G1.



Echo Request (8) e Reply (9)

- Provê um mecanismo para determinar se é possível a comunicação entre dois *hosts*.
- Mensagem usada pelo comando *Ping* para verificar a alcançabilidade entre dois *hosts*.
- O campo *data* contém os dados a serem retornados ao remetente.
- Os campos *identifier* e *sequence number* são usados para "casar" *request* com *reply*.

Echo Request (8) e Reply (9) (cont.)



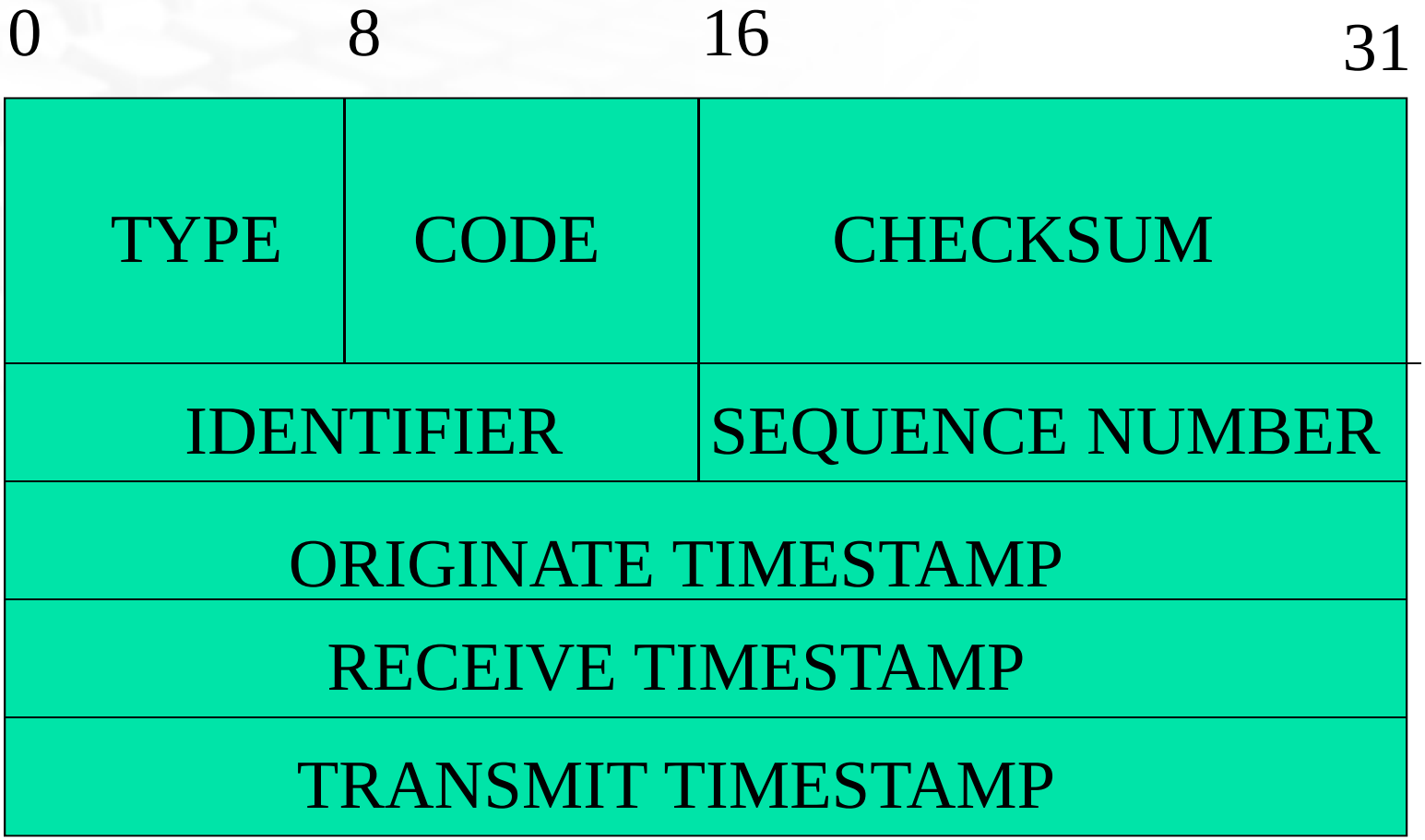
Exemplo – O Comando *Ping*

- Testar se *ring.bell.com* está ativo (*alive*).
- É enviada uma seqüência de 14 mensagens, cada uma contendo 64 bytes.
- No exemplo, as mensagens 0, 1 e 2 são perdidas.

O Comando ping

```
> ping ring.bell.com
ring.bell.com is alive
> ping –s ring.bell.com 64 14
64 bytes from ring.bell.com: icmp_seq = 3. time=21. ms
64 bytes from ring.bell.com: icmp_seq = 4. time=18. ms
64 bytes from ring.bell.com: icmp_seq = 5. time=17. ms
64 bytes from ring.bell.com: icmp_seq = 6. time=19. ms
64 bytes from ring.bell.com: icmp_seq = 7. time=17. ms
64 bytes from ring.bell.com: icmp_seq = 8. time=17. ms
64 bytes from ring.bell.com: icmp_seq = 9. time=17. ms
64 bytes from ring.bell.com: icmp_seq = 10. time=18. ms
64 bytes from ring.bell.com: icmp_seq = 11. time=17. ms
64 bytes from ring.bell.com: icmp_seq = 12. time=17. ms
64 bytes from ring.bell.com: icmp_seq = 13. time=17. Ms
---ring.bell.com PING statistics ---
14 packets transmmited, 11 packets received, 21% packet loss 4
round-trip (ms) min/avg/max = 17/17/21
```

Timestamp Request (13) e Reply (14)



Timestamp Request (13) e Reply (14)

- Mensagem usada para dar uma idéia do tempo que o sistema remoto gasta armazenando e processando o datagrama.
- Provê um mecanismo para verificar as características de *delay* da rede.
  - *Originate timestamp* é preenchido pelo remetente;
  - *Receive timestamp* e *Transmit timestamp* são preenchidos pelo destinatário no momento da chegada do *request* e no momento da saída do *reply*, respectivamente.
- Os tempos envolvidos são em milisegundos, a partir da meia noite, GMT.